CURVATURE BOUNDS FOR CONFIGURATION SPACES

MATTHIAS ERBAR AND MARTIN HUESMANN

ABSTRACT. We show that the configuration space Υ over a manifold M inherits many curvature properties of the manifold. For instance, we show that a lower Ricci curvature bound on M implies a lower Ricci curvature bound on Υ in the sense of Lott–Sturm–Villani, the Bochner inequality, gradient estimates and Wasserstein contraction. Moreover, we show that the heat flow on Υ can be identified as the gradient flow of the entropy.

1. INTRODUCTION

The configuration space Υ over a manifold M is the space of all locally finite point measures, i.e.

$$\Upsilon = \{\gamma \in \mathcal{M}(M) : \gamma(K) \in \mathbb{N}_0 \text{ for all compact } K \subset M\}.$$

In the seminal paper [1] ALBEVERIO–KONDRATIEV–RÖCKNER identified a natural geometry on Υ by "lifting" the geometry of M to Υ. In particular, there is a natural gradient ∇^Υ, divergence $\operatorname{div}^\Upsilon$ and Laplace operator Δ^Υ on the configuration space. It turns out that the Poisson measure π is the unique (up to the intensity) measure on Υ under which the gradient and divergence become dual operators in $L^2(\Upsilon, \pi)$. Hence, the Poisson measure is the natural volume measure on Υ and Υ can be seen as an infinite dimensional Riemannian manifold. The canonical Dirichlet form

$$\mathcal{E}(F) = \int_\Upsilon |\nabla^\Upsilon F|_\gamma^2 \, \pi(d\gamma)$$

induces the heat semigroup T_t^Υ and a Brownian motion on Υ which can be identified with the independent infinite particle process. The intrinsic metric $d_\Upsilon(\gamma, \omega)$ between two configurations γ and ω with respect to $\mathcal{E}$ is the non-normalized L^2 Wasserstein distance between the two measures γ and ω. Typically, d_Υ will attain the value ∞.

In this article, we are interested in the curvature of Υ. We will not try to define a curvature tensor. Instead, we will show that many analytic and geometric estimates that characterize lower curvature bounds on Riemannian manifolds *lift* to natural analogues on the configuration space.

We will first consider *sectional curvature*. There are many equivalent ways of characterizing a global lower bound $K \in \mathbb{R}$ on the sectional curvature using only the Riemannian distance d, e.g. Toponogov's Theorem on triangle comparison. This

This material is based upon work supported by the National Science Foundation under Grant No. 0932078 000, while both authors were in residence at the Mathematical Sciences Research Institute in Berkeley, California, in the fall of 2013. M.E. gratefully acknowledges funding from the European Research Council under the European Communitys Seventh Framework Programme (FP7/2007-2013) / ERC Grant agreement GeMeThnES No. 246923; M.H. from the CRC 1060.

allows to define a generalized sectional curvature bound also for metric spaces leading to the notion of Alexandrov geometry, we point the reader to [8] for a detailed account. Our first result is that sectional curvature bounds lift from M to Υ.

Theorem 1.1. *If M has sectional curvature bounded below by $K \in \mathbb{R}$ then the configuration space Υ has Alexandrov curvature bounded below by* $\min\{K, 0\}$.

From now on we will be concerned with lower bounds on the *Ricci curvature*. They allow to control various analytic, stochastic and geometric quantities, like the volume growth and the heat kernel. A uniform lower bound $\mathrm{Ric} \geq K$ can be encoded in many different ways. Let us recall some of them.

(BI) *Bochner's inequality:* for every smooth function $u : M \to \mathbb{R}$

$$\frac{1}{2}\Delta|\nabla u|^2 - \langle \nabla u, \nabla \Delta u\rangle \geq K|\nabla u|^2 \ .$$

(GE) *Gradient estimate:* for every smooth function u

$$|\nabla T_t^M u|^2 \leq \mathrm{e}^{-2Kt} T_t^M |\nabla u|^2 \ .$$

Here $T_t^M = \mathrm{e}^{t\Delta}$ denotes the heat semigroup on M. (BI) is easily seen to be equivalent to $\mathrm{Ric} \geq K$ by noting that the left hand side equals $\mathrm{Ric}[\nabla u] + \|\mathrm{Hess}\, u\|_{HS}^2$. The equivalence of (BI) and (GE) is due to a classic interpolation argument of Bakry–Émery [7].

Other ways of encoding a lower Ricci bound involve the action of the (dual) heat semigroup on probability measures and the L^2-transportation distance between probability measures. For $\mu \in \mathscr{P}(M)$ the probability measure $H_t^M \mu$ is defined via $\int f \, \mathrm{d}H_t^M \mu = \int T_t^M f \, \mathrm{d}\mu$. Given $\mu_0, \mu_1 \in \mathscr{P}(M)$ their L^2-transportation distance associated to the Riemannian distance d is defined by

$$W_{2,d}^2(\mu_0, \mu_1) \ = \ \inf \left\{ \int d^2(x,y) \, \mathrm{d}q(x,y) \right\} \ ,$$

where the infimum is taken over all couplings of μ_0, μ_1. Recall also the relative entropy of a measure $\mu = \rho m$ w.r.t. the volume measure m given by $\mathrm{Ent}(\mu|m) = \int \rho \log \rho \, \mathrm{d}m$. Then, a lower bound $\mathrm{Ric} \geq K$ is equivalent to

(WC) *$W_{2,d}$-contractivity:* for all $\mu_0, \mu_1 \in \mathscr{P}(M)$ and $t > 0$:

$$W_{2,d}(H_t^M \mu, H_t^M \nu) \ \leq \ \mathrm{e}^{-Kt} W_{2,d}(\mu, \nu) \ .$$

(GC) *Geodesic convexity of* Ent*:* for every (constant-speed) geodesic $(\mu_s)_{s\in[0,1]}$ in $(\mathscr{P}(M), W_{2,d})$ and all $s \in [0,1]$:

$$\mathrm{Ent}(\mu_s|m) \leq (1-s)\,\mathrm{Ent}(\mu_0|m) + s\,\mathrm{Ent}(\mu_1|m) - \frac{K}{2}s(1-s)W_{2,d}^2(\mu_0,\mu_1) \ .$$

These equivalences have been established in [34, 10]. Finally, (WC) and (GC) can be captured in a single inequality

(EVI) *Evolution Variational Inequality:* for all $\mu, \sigma \in \mathscr{P}(M)$ with finite second moment and a.e. $t > 0$:

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{2}W_{2,d}^2(H_t^M\mu, \sigma) + \frac{K}{2}W_{2,d}^2(H_t^M\mu,\sigma) \ \leq \ \mathrm{Ent}(\sigma|m) - \mathrm{Ent}(H_t^M\mu|m) \ .$$

The last property (EVI) was first established in the Riemannian setting in [27, 11]. It is also a way of stating that the heat flow is the gradient flow of the entropy in the metric space $(\mathscr{P}(M), W_{2,d})$ and thus a reformulation of the celebrated result by Jordan–Kinderlehrer–Otto [15].

Notably, the property (GC) does not use the differential structure of M and can be formulated in the framework of metric measure spaces. Sturm [33] and Lott–Villani [23] used this observation to define a notion of lower Ricci curvature bound for metric measure spaces. The stronger property (EVI) was studied on metric measure spaces in a series of papers by Ambrosio–Gigli–Savaré [6, 4, 5]. There the authors show the equivalence of (EVI) with suitable weak forms of (BI) and (GE) for the canonical linear heat flow on such spaces.

Unfortunately, most of this theory does not apply to the configuration space since $(\Upsilon, d_\Upsilon, \pi)$ is only an extended metric measure space, the distance d_Υ can attain the value ∞. However, due to the rich structure of Υ we can establish suitable analogues of the various manifestations of Ricci bounds.
Denote by $T_t^\Upsilon = \mathrm{e}^{t\Delta^\Upsilon}$ the heat semigroup on the configuration space. For an absolutely continuous probability measure $\mu \in \mathscr{P}(\Upsilon)$ with $\mu = f\pi$ define the dual semigroup $H_t^\Upsilon \mu = (T_t^\Upsilon f)\pi$. Moreover, let now denote W_{2,d_Υ} the L^2-transportation distance on $\mathscr{P}(\Upsilon)$ built from d_Υ. The domain of the Dirichlet form $\mathcal{E}$ will be denoted by $\mathcal{F}$.

Theorem 1.2. *Assume that M has Ricci curvature bounded below by $K \in \mathbb{R}$. Then the following hold:*

(i) Bochner inequality: For all cylinder functions $F : \Upsilon \to \mathbb{R}$ we have

$$\frac{1}{2}\Delta^\Upsilon |\nabla^\Upsilon F| - \langle \nabla^\Upsilon F, \nabla^\Upsilon \Delta^\Upsilon F\rangle \ \geq \ K\ |\nabla^\Upsilon F|^2 \ .$$

(ii) Gradient estimate on Υ: For any function $F \in \mathcal{F}$ we have

$$|\nabla^\Upsilon T_t^\Upsilon F|^2 \ \leq \ \mathrm{e}^{-2Kt} T_t^\Upsilon |\nabla^\Upsilon (F)|^2 \quad \pi - a.e.$$

(iii) Wasserstein contraction: For all $\mu, \nu \ll \pi$ we have:

$$W_{2,d_\Upsilon}(H_t^\Upsilon \mu, H_t^\Upsilon \nu) \ \leq \ \mathrm{e}^{-Kt} W_{2,d_\Upsilon}(\mu,\nu) \ .$$

The Bochner inequality, the gradient estimate and the Wasserstein contraction will be derived by a suitable "lifting" of the corresponding statements on M. For the latter two this relies on a representation of the heat semigroup T_t^Υ as an infinite product of independent copies of the heat semigroup on M that will be established in Theorem 2.4. To our knowledge this identification is new in the present generality assuming a (possibly negative) Ricci bound.
Somehow surprisingly, there does not seem to be a straightforward way to "lift" the EVI or the convexity of the relative entropy from M to Υ. Nevertheless, using a careful approximation procedure it is possible to adapt the techniques of [5] to the setting of the configuration space to derive it from the gradient estimate established in Theorem 1.2.

Theorem 1.3. *Assume that M has Ricci curvature bounded below by $K \in \mathbb{R}$. Then the heat flow is the gradient flow of the entropy in the sense of the EVI_K: For all $\mu, \sigma \in \mathscr{P}(\Upsilon)$ with $\mathrm{Ent}(\sigma) < \infty$ and $W_{2,d_\Upsilon}(\mu,\sigma) < \infty$ and a.e. $t > 0$:*

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{2}W^2_{2,d_\Upsilon}(H_t^\Upsilon \mu, \sigma) + \frac{K}{2}W^2_{2,d_\Upsilon}(H_t^\Upsilon \mu, \sigma) \ \leq \ \mathrm{Ent}(\sigma|\pi) - \mathrm{Ent}(H_t^\Upsilon \mu|\pi) \ .$$

Note that a priori the dual semigroup is only defined on measures with density. Using a careful approximation technique given in Lemma 5.1 and Wasserstein contractivity we can extend it to all measures at finite distance to the domain of the entropy. This is also the maximal set of measures for which EVI can be stated. As a direct consequence we obtain

Corollary 1.4. *The entropy is (strongly) $K-$convex on $\big(\mathscr{P}(\Upsilon), W_{2,d_\Upsilon}\big)$. More precisely, for all $\mu_0, \mu_1 \in D(\mathrm{Ent})$ with $W_{2,d_\Upsilon}(\mu_0,\mu_1) < \infty$ and any geodesic $(\mu_s)_{s\in[0,1]}$ connecting them we have for all $s \in [0,1]$:*

$$\mathrm{Ent}(\mu_s|\pi) \;\le\; (1-s)\,\mathrm{Ent}(\mu_0|\pi) + s\,\mathrm{Ent}(\mu_1|\pi) - \frac{K}{2}s(1-s)W^2_{2,d_\Upsilon}(\mu_0,\mu_1)\;.$$

In particular, we see that $(\Upsilon, d_\Upsilon, \pi)$ is an extended metric measure space satisfying the synthetic Ricci bound $\mathrm{CD}(K,\infty)$ in the sense of Sturm and Lott–Villani. Since the configuration space naturally appears (see e.g. [2, 25, 26]) as the state space for infinite systems of interacting Brownian motions, our results can be interpreted as a first step in order to make tools from optimal transportation available for infinite particle systems. In fact, for the case of no interaction Theorem 1.3 is the realization of the famous heat flow interpretation of Jordan–Kinderlehrer–Otto for an infinite system of Brownian motions. It is a challenge for future work to incorporate interactions in this picture.

Remark 1.5. It would be natural to consider more generally as base space a *weighted* Riemannian manifold $(M, d, \mathrm{e}^{-V}m)$, with $V : M \to \mathbb{R}$ say of class C^2, and equip the configuration space (Υ, d_Υ) with the Poisson measure π_V built from the reference measure $\mathrm{e}^{-V}m$. This corresponds to a system of independent Brownian motions with drift. We expect that all the results presented here continue to hold under the assumption of a lower bound of the weighted Ricci curvature

$$\mathrm{Ric} + \mathrm{Hess}\,V \;\ge\; K\;.$$

The only thing that does not adapt immediately is the control on the tail of the heat kernel in Lemma 2.5 needed for the explicit representation of the heat semigroup. In fact, the validity of such a heat kernel bound under weighted Ricci bounds is interesting in itself and seems to be open in this generality. Since settling this question is not in the scope of this paper we chose to work with unweighted manifolds.

Connection to the literature. Even though the article [1] triggered off an enormous amount of research, the curvature of the "lifted" geometry on the configuration space has - to our knowledge - not yet been analyzed. Privault [28] derived a Weitzenböck type formula on the configuration space; however, his analysis is based on a different geometry which does not directly relate to the geometry introduced in [1].

Spaces satisfying (synthetic) lower Ricci curvature bounds are currently a hot topic of research and many impressive results have been obtained, e.g. see [6, 4, 5, 14]. However, most of the applications and examples are finite dimensional. So far the Wiener space was the only known example of a truly infinite dimensional CD space. Recently, also path spaces over a Riemannian manifold have been investigated by Naber [24] where he characterizes simultaneous lower and upper Ricci curvature bounds via gradient estimates and spectral gap estimates on the path space.

The geometry on the configuration space is very similar to the geometry of the Wasserstein space. However, due to the fact that every point in a configuration gets mass at least one the lower sectional curvature bound is stable even for negative lower bounds in contrast to the Wasserstein space, see Proposition 2.10 in [33]. Moreover, the Wasserstein space together with the entropic measure is known to not admit any Ricci lower bounds [9] which is again in sharp contrast to Theorem 1.2 and Theorem 1.3.

Outline. In Section 2 we start by explaining the "lifted" geometry on Υ. Using a version of Rademachers Theorem on the configuration space we show that differential structure and the metric structure fit together by proving that the Cheeger energy and the Dirichlet form coincide. Subsequently, we discuss the heat semigroup in some detail and give a useful point wise representation in terms of the semigroup on the base space M. We close this section by collecting some tools we need in the proof of the main theorems.
In Section 3, we collect and adapt results on optimal transport to the configuration space setting.
In Section 4, we prove Theorem 1.2, the different manifestations of curvature bounds which can be deduced by "lifting" of the corresponding results on M.
Finally in Section 5, we show that the EVI_K holds on the configuration space, i.e. we prove Theorem 1.3.
The Appendix contains the proof of the approximation result needed to extend the dual semigroup beyond measures with density.

Acknowledgements. The authors would like to thank Theo Sturm and Fabio Cavalletti for several fruitful discussions on the subject of this paper.

2. Preliminaries

2.1. **Differentiable structure of configuration space.** Let M be a smooth complete and connected Riemannian manifold. We denote by $\langle \cdot, \cdot \rangle_x$ the metric tensor at x, d is the Riemannian distance and m the volume measure. We assume that M is non-compact and $m(M) = \infty$.[1] The configuration space Υ over the base space (M, d, m) is the set of all locally finite counting measures, i.e.

$$\Upsilon := \{\gamma \in \mathcal{M}(M) : \gamma(K) \in \mathbb{N}_0 \;\text{ for all } K \subset M \text{ compact } \} .$$

Each $\gamma \in \Upsilon$ can be represented as $\gamma = \sum_{i=1}^n \delta_{x_i}$ for some $n \in \mathbb{N}_0 \cup \{\infty\}$, and suitable points x_i in M. Here $n = 0$ corresponds to the empty configuration. To be more precise, let $\mathcal{A}$ be the set of finite and infinite sequences in M without accumulation points and let

$$l : \mathcal{A} \to \Upsilon,\ (x_1, x_2, , \ldots) = \mathbf{x} \mapsto \gamma = \sum_i \delta_{x_i} .$$

Then any $\mathbf{x} \in l^{-1}(\gamma)$ is called a labeling of γ. We can decompose the configuration space as $\Upsilon = \cup_{n\in\mathbb{N}_0\cup\{\infty\}} \Upsilon^{(n)}$ where $\Upsilon^{(n)} = \{\gamma \in \Upsilon : \gamma(M) = n\}$.
We endow the configuration space with the vague topology which makes it a Polish space as a closed subset of a Polish space (e.g. see [16, Theorem A2.3]). This means that $\gamma_n \to \gamma$ if and only if $\int f \; d\gamma_n \to \int f \; d\gamma =: \gamma(f)$ for all $f \in C_c(M)$.
There is a natural probability measure on Υ, the Poisson measure π. It can be defined via its Laplace transform

$$\int \exp(\gamma(f)) \;\, \mathrm{d}\pi(\gamma) = \exp\left(\int \exp(f(x)) - 1 \;\, \mathrm{d}m(x) \right) .$$

[1]The results also hold in the case that M is compact. However, they can be derived much easier.

Equivalently, we can characterize π as follows: for any choice of disjoint Borel sets $A_1, \ldots, A_k \subset M$ with $m(A_i) < \infty$ the family of random variables $\gamma(A_1), \ldots, \gamma(A_k)$ is independent and $\gamma(A_i)$ is Poisson distributed with parameter $m(A_i)$. In particular, given a Borel set A of finite volume and condition on the event that $\gamma(A) = n < \infty$ then the n points are iid uniformly distributed in A.
Note that the analysis of Υ is most interesting when M in non-compact and $m(M) = \infty$ since in this case configurations consist typically of infinitely many points, i.e. we have $\pi(\Upsilon^{(n)}) = 0$ for all $n \in \mathbb{N}$ and $\pi(\Upsilon^{(\infty)}) = 1$.
The *tangent space* $T_\gamma\Upsilon$ of Υ at a configuration γ is defined to be the space of all γ-square integrable sections of the tangent bundle TM of M, i.e.

$$T_\gamma\Upsilon = \{V : M \to TM, \int_M \langle V, V\rangle_x \ \mathrm{d}\gamma(x) < \infty\}.$$

Equivalently, we can write $T_\gamma\Upsilon = L^2(\bigoplus_{x\in\gamma} T_xM, \gamma)$. We will denote the scalar product on $T_\gamma\Upsilon$ by

$$\langle V_1, V_2\rangle_\gamma := \int_M \langle V_1(x), V_2(x)\rangle_x \ \mathrm{d}\gamma(x).$$

We also sometimes write $\|V\|_\gamma^2 := \langle V, V\rangle_\gamma$. Note that this is a non-trivial structure. The tangent spaces vary with γ even if M is Euclidean.
Next we introduce an important class of "test functions". A smooth *cylinder function* is a function $F : \Upsilon \to \mathbb{R}$ that can be written as

$$F(\gamma) = g_F(\gamma(\varphi_1), \ldots, \gamma(\varphi_n)) ,$$

for some $n \in \mathbb{N}, g_F \in C_b^\infty(\mathbb{R}^n)$ and $\varphi_1, \ldots, \varphi_n \in C_c^\infty(M)$. The set of all smooth cylinder functions will be denoted by $\mathsf{Cyl}^\infty(\Upsilon)$. For $F \in \mathsf{Cyl}^\infty(\Upsilon)$ we define the *gradient* of F by

$$\nabla^\Upsilon F(\gamma; x) := \sum_{i=1}^n \partial_i g_F(\gamma(\varphi_1), \ldots, \gamma(\varphi_n))\nabla\varphi_i(x) \quad \gamma \in \Upsilon,\ x \in M .$$

Here ∂_i denotes the partial derivative in the i-th direction and ∇ denotes the gradient on M. Alternatively, we can define the gradient using directional derivatives. To this end denote the set of all smooth and compactly supported vector fields on M by $\mathcal{V}_0(M)$. For $V \in \mathcal{V}_0(M)$ let ψ_t be the flow of diffeomorphisms generated by V. For fixed $\gamma \in \Upsilon$, this generates a curve $\psi_t^*\gamma = \gamma \circ \psi_t^{-1}, t \in \mathbb{R}$ on Υ. Then we have for $F \in \mathsf{Cyl}^\infty(\Upsilon)$ and $\gamma \in \Upsilon$

$$\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} F(\psi_t^*\gamma) = \langle \nabla^\Upsilon F(\gamma), V\rangle_\gamma =: \nabla_V^\Upsilon F(\gamma) .$$

Similarly, we can introduce the *divergence* div^Υ on Υ. For $F_i \in \mathsf{Cyl}^\infty(\Upsilon)$ and $V_i \in \mathcal{V}_0(M)$ we define for $\gamma \in \Upsilon$

$$\mathrm{div}^\Upsilon\left(\sum_{i=1}^n F_i \cdot V_i\right)(\gamma) := \sum_{i=1}^n \nabla_{V_i}^\Upsilon F_i(\gamma) + F_i(\gamma)\cdot\gamma(\mathrm{div}^M(V_i)) .$$

It is proven in [1] that the Poisson measure π is (up to the intesity) the unique measure such that div^Υ and ∇^Υ are adjoint in $L^2(\pi)$. We also define the *Laplace operator* $\Delta^\Upsilon := \mathrm{div}^\Upsilon\nabla^\Upsilon$.

With this differential structure at hand we can talk about *Dirichlet forms*. For a cylinder function $F \in \mathsf{Cyl}^\infty(\Upsilon)$ we define the pre-Dirichlet form

$$\mathcal{E}(F,F) \ := \ \int \langle \nabla^\Upsilon F, \nabla^\Upsilon F\rangle_\gamma \ \pi(d\gamma) \ .$$

It is shown in [1, Cor. 4.1] that $(\mathcal{E}, \mathsf{Cyl}^\infty(\Upsilon))$ is closable and its closure $(\mathcal{E},\mathcal{F})$ is a Dirichlet form. By [29, Proposition 1.4 (iv)], for every $F \in \mathcal{F}$ there exists a measurable section $\nabla^\Upsilon F : \Upsilon \to T\Upsilon$ such that $\mathcal{E}(F) = \int |\nabla^\Upsilon F|_\gamma^2 \ \mathrm{d}\pi(\gamma)$. Thus $\mathcal{E}$ admits a carré du champs operator $\Gamma^\Upsilon : \mathcal{F} \to L^1(\Upsilon,\pi)$ given by $\Gamma^\Upsilon(F)(\gamma) = |\nabla^\Upsilon F|_\gamma^2$.
We will denote the semigroup in $L^2(\Upsilon,\pi)$ associated to the Dirichlet form $(\mathcal{E},\mathcal{F})$ by $T_t^\Upsilon = \exp(t\Delta^\Upsilon)$ and call it the *heat semigroup* on Υ. Its generator is the Friedrichs extension of Δ^Υ which we again denote by Δ^Υ. It is a selfadjoint and closed operator.

2.2. **Metric structure of Υ and compatibility.** A natural distance on the configuration space is given by the non-normalized L^2-transportation distance, defined for two measures $\gamma,\eta \in \Upsilon$ by

$$d_\Upsilon^2(\gamma,\eta) = \inf_{q\in \mathsf{Cpl}(\gamma,\eta)} \int d^2(x,y) \ q(dx,dy) \ ,$$

where $\mathsf{Cpl}(\gamma,\eta)$ denotes the set of all couplings between γ and η. Note that $d_\Upsilon : \Upsilon \to [0,+\infty]$ is an *extended* distance, i.e. it is symmetric, vanishes precisely on the diagonal and satisfies the triangle inequality. It can take the value $+\infty$, e.g. we have $d_\Upsilon(\gamma,\eta) = \infty$ if $\gamma \in \Upsilon^{(n)}$ and $\eta \in \Upsilon^{(m)}$ with $m \neq n$. We denote the set of all optimal couplings between γ and η by $\mathsf{Opt}(\gamma,\eta)$.
We denote by $C(\Upsilon)$ the set of all continuous functions on Υ w.r.t. the vague topology. We say that a function $F : \Upsilon \to \mathbb{R}$ is d_Υ-Lipschitz iff

$$|F(\gamma) - F(\eta)| \le C d_\Upsilon(\gamma,\eta) \quad \forall \gamma,\eta \in \Upsilon \ , \tag{2.1}$$

for some constant $C \ge 0$. The set of all d_Υ-Lipschitz functions will be denoted by $\mathsf{Lip}(\Upsilon)$ and the set of bounded d_Υ-Lipschitz functions by $\mathsf{Lip}_b(\Upsilon)$. For $F \in \mathsf{Lip}(\Upsilon)$ the global Lipschitz constant $\mathsf{Lip}(F)$ is the smallest C such that (2.1) holds and we define the local Lipschitz constant by

$$|DF|(\gamma) := \limsup_{d_\Upsilon(\eta,\gamma)\to 0} \frac{|F(\gamma) - F(\eta)|}{d_\Upsilon(\gamma,\eta)} \ . \tag{2.2}$$

The compatibility of the differential and metric structure of the configuration space is given by the following Rademacher theorem which we quote from [29, Thm. 1.3, Thm. 1.5].

Theorem 2.1. *(i) Suppose $F \in L^2(\pi) \cap \mathsf{Lip}(\Upsilon)$. Then $F \in \mathcal{F}$. Moreover, there exists a measurable section $\nabla^\Upsilon F$ of $T\Upsilon$ such that*
a) $\sqrt{\Gamma^\Upsilon(F)}(\gamma) = |\nabla^\Upsilon F(\gamma)|_\gamma \le \mathsf{Lip}(F)$ for π-a.e. γ.
b) If $V \in \mathcal{V}_0(M)$ generates the flow $(\psi_t)_{t\in\mathbb{R}}$, then for π-a.e. γ and all $s \in \mathbb{R}$:

$$\frac{F(\psi_t^*\gamma) - F(\gamma)}{t} \to \langle \nabla^\Upsilon F(\gamma), V\rangle_\gamma, \qquad \textit{as } t \to 0 \textit{ in } L^2(\pi \circ (\psi_s^*)^{-1}) \ .$$

(ii) *If $F \in \mathcal{F}$ satisfies $\Gamma^\Upsilon(F) \le C^2, \pi$-a.e. and if F has a d_Υ-continuous π-version, then there exists a π-measurable π-version $\tilde{F}$ which is d_Υ-Lipschitz with $Lip(\tilde{F}) \le C$.*

(iii) *d_Υ coincides with the intrinsic metric of the Dirichlet form $(\mathcal{E}, \mathcal{F})$, i.e. for all $\gamma, \eta \in \Upsilon$:*

$$d_\Upsilon(\gamma, \eta) = \sup\left\{F(\gamma) - F(\eta) : F \in \mathcal{F} \cap C(\Upsilon)\ ,\ \Gamma^\Upsilon(F) \le 1\ \pi\text{-a.e.}\right\}\ .$$

As a consequence we obtain the following pointwise comparison of the Lipschitz constant and the Gamma operator.

Lemma 2.2. *For all $F \in \mathsf{Lip}_b(\Upsilon)$ and π-a.e. γ we have*

$$|DF|(\gamma) \ge \|\nabla^\Upsilon F\|_\gamma = \sqrt{\Gamma^\Upsilon(F)}(\gamma)\ . \tag{2.3}$$

Proof. By [29, Prop. 5.4] for every $\gamma, \eta \in \Upsilon$ with $d_\Upsilon(\gamma, \eta) < \infty$ and every $\epsilon > 0$ there is a $V \in \mathcal{V}_0(M)$ generating the flow $(\psi_t)_{t\in\mathbb{R}}$ such that $d_\Upsilon(\psi_1^*\gamma, \eta) < \epsilon$ and $\|V\|_{\psi_t^*\gamma} = d_\Upsilon(\psi_1^*\gamma, \gamma)$ for all $t \in [0,1]$. Hence, by d_Υ-continuity of F we have

$$|DF|(\gamma) = \limsup_{d_\Upsilon(\eta,\gamma)\to 0} \frac{|F(\eta) - F(\gamma)|}{d_\Upsilon(\eta, \gamma)} = \limsup_{V\in\mathcal{V}_0(M), \|V\|_\gamma \to 0} \frac{|F(\psi_1^*\gamma) - F(\gamma)|}{\|V\|_\gamma}\ .$$

By part (i) b) of Theorem 2.1, we have for π-a.e. γ and all $V \in \mathcal{V}_0(M)$

$$|DF(\gamma)| \ge \lim_{t\to 0} \frac{F(\psi_t^*\gamma) - F(\gamma)}{t\|V\|_\gamma} = \frac{1}{\|V\|_\gamma}\langle \nabla^\Upsilon F(\gamma), V\rangle_\gamma\ .$$

Hence, taking the supremum over V we get $|DF|(\gamma) \ge \|\nabla^\Upsilon F\|_\gamma$ for π-a.e. γ. □

In [6] Ambrosio, Gigli and Savaré develop a calculus on (extended) metric measure spaces and study the "heat flow" in this setting. A crucial result is the construction of a natural candidate for a Dirichlet form starting only from a metric and a measure. Their work is the foundation for studying Riemannian Ricci curvature bounds via optimal transport on (non-extended) metric measure spaces in [4, 5]. Here we make the connection to this approach, showing that the triple $(\Upsilon, d_\Upsilon, \pi)$ fits into the framework of [6] and that the Dirichlet form $\mathcal{E}$ coincides with its metric counterpart constructed from d_Υ.

First note that (Υ, d_Υ) equipped with the vague topology is a Polish extended space in the sense of [6, Def. 2.3]: it is complete, i.e. every d_Υ-convergent sequence has a limit in Υ, $d_\Upsilon(\gamma_n, \gamma) \to 0$ implies that $\gamma_n \to \gamma$ vaguely for all sequences $(\gamma_n) \subset \Upsilon$ and $\gamma \in \Upsilon$, and d_Υ is lower semi continuous w.r.t. the vague topology.

The Cheeger energy Ch introduced in [6] is given on the configuration space as a functional $\mathsf{Ch} : L^2(\Upsilon, \pi) \to [0, +\infty]$ defined via

$$\mathsf{Ch}(F) := \inf\left\{\liminf_{n\to\infty} \frac{1}{2}\int_\Upsilon |DF_n|^2\, d\pi\ :\ F_n \in \mathsf{Lip}_b(\Upsilon),\ F_n \to F \text{ in } L^2(\Upsilon, \pi)\right\}\ . \tag{2.4}$$

Proposition 2.3. *The Cheeger energy associated to d_Υ coincides with the Dirichlet form $\mathcal{E}$, i.e. $\mathcal{E}(F) = 2\mathsf{Ch}(F)$ for all $F \in L^2(\Upsilon, \pi)$.*

Proof. Let us first show that $\mathcal{E} \le 2\mathsf{Ch}$. By definition for $F \in L^2(\Upsilon, \pi)$ with $\mathsf{Ch}(F) < \infty$ there is a sequence of bounded Lipschitz functions $(F_n)_{n\in\mathbb{N}}$ such that

$F_n \to F$ in $L^2(\Upsilon,\pi)$ and $\lim_n \mathsf{Ch}(F_n) = \mathsf{Ch}(F)$. By (2.3) of Lemma 2.2 and lower semicontinuity of $\mathcal{E}$ in $L^2(\Upsilon,\pi)$ we obtain

$$2\mathsf{Ch}(F) = \lim_n \int |DF_n|^2 \, d\pi \geq \liminf_n \int |\nabla^\Upsilon F_n|^2 \, d\pi \geq \mathcal{E}(F) \, .$$

To prove the converse inequality $\mathcal{E} \geq 2\mathsf{Ch}$, note that by definition for $F \in L^2(\Upsilon,\pi)$ with $\mathcal{E}(F) < \infty$ there exists a sequence of cylinder functions $(F_n)_{n\in\mathbb{N}}$ such that $F_n \to F$ in $L^2(\Upsilon,\pi)$ and $\lim_n \mathcal{E}(F_n) = \mathcal{E}(F)$. Note that any cylinder function F_n is d_Υ-Lipschitz with $|DF_n|(\gamma) = \|\nabla^\Upsilon F_n\|_\gamma$. Thus we obtain from the definition of Ch and its lower semicontinuity in $L^2(\Upsilon,\pi)$ (see [6, Thm. 4.5]):

$$2\mathsf{Ch}(F) \leq \liminf_n 2\mathsf{Ch}(F_n) \leq \liminf_n \int |DF_n|^2 \, d\pi = \liminf_n \mathcal{E}(F_n) = \mathcal{E}(F) \, .$$

□

Having identified the Dirichlet form $\mathcal{E}$ with the Cheeger Ch energy build from the distance d_Υ in particular yields that the semigroup T_t^Υ coincides with the gradient flow of Ch in $L^2(\Upsilon,\pi)$. This will be used in Section 5.

2.3. **The heat semigroup.** In this section we establish an explicit representation of the Markov semigroup T_t^Υ associated to the Dirichlet form $\mathcal{E}$. We identify it with the semigroup of the independent particle process obtained by starting in each point of a configuration independent Brownian motions. This identification is non-trivial when $m(M) = \infty$. While the first lives by definition on the configuration space, the latter a priory lives in the larger space of counting measures that are not necessarily locally finite. We will show that whenever $\mathrm{Ric}_M \geq K$ for some $K \in \mathbb{R}$ the independent particle process can be started in a subset of Υ of full π measure and stays there for all time.

Consider the infinite product $M^{\mathbb{N}}$ equipped with the cylinder σ-algebra $\mathcal{C}(M^{\mathbb{N}})$. We put $\mathcal{A} \in \mathcal{C}(M^{\mathbb{N}})$ to be the set of all sequences $(x_n)_{n=1}^\infty \in M^{\mathbb{N}}$ which have no accumulation points. Recall the labeling map $l : \mathcal{A} \to \Upsilon$ given by

$$l : \ (x_n)_{n=1}^\infty \ \mapsto \sum_{n=1}^\infty \delta_{x_n} \, .$$

Note that $\pi(l(\mathcal{A})) = 1$. Let $p_t^M(x,y)$ denote the heat kernel on the manifold M. Moreover, we denote by

$$p_t^M(x,A) \ = \ \int_A p_t^M(x,y) \, \mathrm{d}m(y)$$

the semigroup of transition kernels. This gives rise to a family of probability measure on $\big(M^{\mathbb{N}}, \mathcal{C}(M^{\mathbb{N}})\big)$ by considering the product measures

$$p_t^{\mathbb{N}}\big((x_n)_n, \cdot\big) := \bigotimes_{n=1}^\infty p_t^M(x_n, \cdot) \, .$$

Given $\gamma \in \Upsilon$ we can define a probability measure on Υ via

$$p_t^\Upsilon(\gamma, G) \ := \ p_t^{\mathbb{N}}\big((x_n)_n, l^{-1}(G)\big) \qquad G \in \mathcal{B}(\Upsilon) \, , \tag{2.5}$$

where $\gamma = l\big((x_n)\big)$, provided that $p_t^{\mathbb{N}}\big((x_n)_n, \mathcal{A}\big) = 1$ for all $t \geq 0$. Our goal will be to show that for a large class of γ the latter indeed holds.

We fix a point $x_0 \in M$ and denote by $B_r = B(x_0, r)$ the closed ball around x_0 with radius r. Define for each $\alpha \geq 1$:

$$\Theta_\alpha := \left\{\gamma \in \Upsilon \ : \ \exists C > 0 : \ \forall r \in \mathbb{N} : \gamma\big(B_r\big) \leq C\mathrm{e}^{\alpha r}\right\} \ .$$

Since $\Theta_\alpha \subset \Theta_\beta$ for $\alpha \leq \beta$ it makes sense to define also

$$\Theta := \bigcup_{\alpha \geq 1} \Theta_\alpha \ . \tag{2.6}$$

We call Θ the set of *good configurations*. Note that the Poisson measure is concentrated on configurations satisfying $\gamma(B_r) \sim \mathrm{vol}(B_r)$ as $r \to \infty$. Since we assume $\mathrm{Ric} \geq K$, the Bishop–Gromov volume comparison theorem (see Lemma 2.5 below) implies that $\mathrm{vol}(B_r) \leq C\mathrm{e}^{\alpha r}$ for suitable constants C, α. Thus, we conclude that $\pi(\Theta_\alpha) = 1$ for α sufficiently large and in particular $\pi(\Theta) = 1$. The following is a slight generalization of [19, Thm. 2.2, 4.1].

Theorem 2.4. *Assume that $\mathrm{Ric}_M \geq K$ for some $K \in \mathbb{R}$. Then for each $\gamma \in \Theta$ and all $t > 0$ the measure $p_t^\Upsilon(\gamma, \cdot)$ defined in (2.5) is a probability measure on Υ. Moreover, $(p_t^\Upsilon)_{t\geq 0}$ is a Markov semigroup of kernels on $\big(\Theta, \mathcal{B}(\Theta)\big)$. For each $F \in L^2(\Upsilon, \pi)$ the function*

$$\Theta \ni \gamma \mapsto \tilde{T}_t^\Upsilon F(\gamma) = \int_\Theta F(\xi) p_t^\Upsilon(\gamma, \ \mathrm{d}\xi)$$

is a π-version of the function $T_t^\Upsilon F \in L^2(\Upsilon, \pi)$.

Proof. Fix $\epsilon > 0$. Let us write $|x| := d(x, x_0)$, where x_0 is the point chosen in the definition of Θ. We will first prove that for any $\gamma \in \Theta$ and $t \in (0, \varepsilon)$:

$$\sum_{x\in\gamma} p_t^M\big(x, \complement B(x, |x|/2)\big) \ < \ \infty \ . \tag{2.7}$$

To this end let $\gamma \in \Theta$ and let $(x_n)_n$ be a labeling of γ. We can assume that $|x_n| \leq |x_{n+1}|$ for all n. There exists C, α such that $\gamma(B_r) \leq C\mathrm{e}^{\alpha r}$ for all $r \in \mathbb{N}$. For $n \in \mathbb{N}$ let us set:

$$r_n \ := \ \left\lfloor \frac{1}{\alpha} \log\Big(\frac{n}{C}\Big) \right\rfloor \ .$$

This implies that $\gamma(B_{r_n}) < n$ and hence we have $x_n \notin B_{r_n}$ and $|x_n| \geq r_n$. Using Lemma 2.5 below we obtain that for constants C_1, C_2 (possibly changing from line to line):

$$\begin{aligned}\sum_{x\in\gamma} p_t^M\big(x, \complement B(x, |x|/2)\big) \ &= \ \sum_{n=1}^\infty p_t^M\big(x_n, \complement B(x_n, |x_n|/2)\big) \ \leq \ \sum_{n=1}^\infty p_t^M\big(x_n, \complement B(x_n, r_n/2)\big) \\ \leq \ \sum_{n=1}^\infty C_2 \exp(-C_1 r_n^2) \ &\leq \ \sum_{n=1}^\infty C_2 \exp\Big(-C_1 \log(n)^2\Big) \ < \ \infty \ ,\end{aligned}$$

which proves (2.7).

Now, we want to prove that for any $(x_n)_n \in l^{-1}(\Theta)$ we have

$$p_t^{\mathbb{N}}\Big((x_n)_n, l^{-1}(\Theta)\Big) \ = \ 1 \ . \tag{2.8}$$

So fix $(x_n)_n \in l^{-1}(\Theta)$ and set

$$\begin{aligned} \mathcal{A}_n &:= \big\{(y_k)_k \in M^{\mathbb{N}} \ : \ y_n \in B(x_n, |x_n|/2)\big\} \ , \\ \mathcal{A}' &:= \liminf_n \mathcal{A}_n \ . \end{aligned}$$

From (2.7) and the Borel–Cantelli lemma we infer that for any $t \in (0, \varepsilon)$:

$$p_t^{\mathbb{N}}\Big((x_n)_n, \mathcal{A}'\Big) \ = \ 1 \ .$$

By definition of Θ we have $|x_n| \to \infty$ as $n \to \infty$ and so no sequence in $\mathcal{A}'$ has accumulation points which means $\mathcal{A}' \subset \mathcal{A}$. To prove (2.8) it is sufficient to show that $\mathcal{A}' \subset l^{-1}(\Theta)$. So fix $(y_n)_n \in \mathcal{A}'$ and let k be the number of those n such that $y_n \notin B(x_n, |x_n|/2)$. Putting $\gamma = l\big((x_n)_n\big)$ and $\gamma' = l\big((y_n)_n\big)$ and using (2.9) we can estimate:

$$\begin{aligned} \gamma'(B_r) \ &\leq \ \gamma(B_{2r}) + k \\ &\leq \ C\mathrm{e}^{2\alpha r} + k \\ &\leq \ C'\mathrm{e}^{2\alpha r} \ , \end{aligned}$$

for a suitable $C' > 0$ and all $r \in \mathbb{N}$. Hence we have $\gamma' \in \Theta_{2\alpha} \subset \Theta$ and this proves (2.8). Thus (2.5) defines a probability measure on Υ concentrated on Θ. It then follows easily from the semigroup property of $p_t^{\mathbb{N}}$ that p_t^{Υ} can be defined for all $t > 0$ and is a Markov semigroup of kernels on Θ. The last statement of the theorem is proven as in [19, Thm. 2.1]. □

Lemma 2.5. *Assume that* $\mathrm{Ric}_M \geq -K$ *for some* $K \in [0, \infty)$. *Then there is a constant* c *such that*

$$\mathrm{vol}\big(B(x,r)\big) \ \leq \ \mathrm{vol}\big(B(x,1)\big) \cdot \mathrm{e}^{cr} \qquad \forall x \in M, \ r \geq 1 \ . \tag{2.9}$$

Moreover, for any $T > 0$ *there are constants* c_1, c_2 *such that:*

$$\sup_{t \in (0,T]} \sup_{x \in M} p_t^M\big(x, \complement B(x,r)\big) \ \leq \ c_1 \mathrm{e}^{-c_2 r^2} \qquad \forall r > 0 \ . \tag{2.10}$$

Proof. The estimate (2.9) follows from the Bishop–Gromov volume comparison theorem [20, Lem. 5.3.bis].
The second estimate (2.10) is a consequence of the following result (see relation (8.65) in [31]): Fix $x \in M$ and let $(B_t^x)_{t \geq 0}$ be a Brownian motion started from x. Then for any $\lambda \in (0,1)$ and $r > 0$ we have:

$$\mathbb{P}\Big[\sup_{0 \leq s \leq t} d(B_s^x, x) \geq r\Big] \ \leq \ \frac{2}{\sqrt{1-\lambda}} \exp\left(-\frac{\lambda r^2}{2t} + \frac{\lambda\big(2d + Kd^2 t\big)}{1-\lambda}\right) \ , \tag{2.11}$$

where $d = \dim M$. This implies (2.10) immediately, since

$$p_t^M\Big(x, \complement B(x,r)\Big) \ = \ \mathbb{P}\Big[d(B_t^x, x) \geq r\Big] \ \leq \ \mathbb{P}\Big[\sup_{0 \leq s \leq t} d(B_s^x, x) \geq r\Big] \ .$$

□

2.4. **Additional tools.**

Lemma 2.6. *For every $\gamma,\omega \in \Upsilon$ with $d_\Upsilon(\gamma,\omega) < \infty$ there exists an optimal coupling q which is a matching, i.e. $d^2_\Upsilon(\gamma,\omega) = \int d^2(x,y)\ \mathrm{d}q(x,y)$ and for all $\{x,y\} \in M\times M$ we have $q(\{x,y\}) \in \{0,1\}$.*

As an immediate consequence we obtain that

$$d^2_\Upsilon(\gamma,\omega) \ = \ \min\left\{\sum_{i=1}^n d^2(x_i,y_i)\ :\ \gamma=\sum_i \delta_{x_i}, \omega = \sum_i \delta_{y_i}\right\}\ ,$$

provided $d_\Upsilon(\gamma,\omega) < \infty$ and $\gamma(M) = \omega(M) = n$.

Proof. By [18], the set of doubly stochastic matrices is the closure of convex combinations of permutations matrices, i.e. doubly stochastic matrices whose entries are precisely 0 or 1, with respect to the locally convex topology which makes all elements, row sums and column sums of the matrix continuous. Call this the τ topology. Now take $q \in \mathsf{Cpl}(\gamma,\omega)$ and $f \in C_c(M\times M)$. Then $\int f\ \mathrm{d}q = \sum f(x_i,y_j)q(x_i,y_j)$, for some labeling $(x_i)_i, (y_j)_j$ of γ and ω respectively. Then $a_{ij} = q(x_i,y_j)$ defines a doubly stochastic matrix. Fixing the labeling, a doubly stochastic matrix defines a coupling between γ and ω. Moreover, as $\sum f(x_i,y_j)a_{ij}$ is a finite sum, convergence in the τ topology implies convergence in the vague topology.
Take $q' \in \mathsf{Opt}(\gamma,\omega)$. By the results of [18], there exists a sequence of couplings $(q'_n)_n$ converging vaguely to q' such that each q'_n can be written as a (finite) convex combination of matchings (which correspond to permutation matrices). By the linearity of $q\mapsto \int d^2\ \mathrm{d}q$, this implies the existence of a sequence of matchings $(q_n)_n$ of γ and ω such that $\int d^2\ \mathrm{d}q_n \le \int d^2\ \mathrm{d}q'_n \searrow \int d^2\ \mathrm{d}q'$. Hence, we have a uniform bound on the transportation cost and there is a converging subsequence which we denote again by $(q_n)_n$. Denote by q its limit. By lower semicontinuity, we have

$$\int d^2\ \ \mathrm{d}q \le \liminf \int d^2\ \ \mathrm{d}q_n = \int d^2\ \ \mathrm{d}q'\ ,$$

so that $q \in \mathsf{Opt}(\gamma,\omega)$. As all the q_n are matchings also q has to be a matching which can be seen by testing against functions $f_{i,j} \in C_c(M\times M)$ which satisfy $f_{i,j}(x_l,y_k) = \delta_{x_i,y_j}(x_k,y_k)$ for the fixed labeling $(x_i)_i$ and $(y_j)_j$ of γ and ω. $\square$

Corollary 2.7. (Υ, d_Υ) *is a geodesic space, i.e any pair γ_0,γ_1 with $d_\Upsilon(\gamma_0,\gamma_1) < \infty$ can be connected by a curve $(\gamma_t)_{t\in[0,1]}$ such that for all $s,t\in[0,1]$ we have $d_\Upsilon(\gamma_s,\gamma_t) = |t-s|d_\Upsilon(\gamma_0,\gamma_1)$.*

Proof. Choose labelings $(x^j_i)_i$ of γ_j such that $d^2_\Upsilon(\gamma_0,\gamma_1) = \sum_i d^2(x^0_i,x^1_i)$. For each i choose a geodesic $(x^t_i)_{t\in[0,1]}$ and put $\gamma_t = \sum_i \delta_{x^t_i}$. Then $(\gamma_t)_t$ is a geodesic in Υ. Indeed,

$$d^2_\Upsilon(\gamma_s,\gamma_t)\ \le\ \sum_i d^2(x^s_i,x^t_i)\ =\ |t-s|^2\sum_i d^2(x^0_i,x^1_i)\ =\ |t-s|^2 d^2_\Upsilon(\gamma_0,\gamma_1)\ .$$

The reverse inequality follows from the triangle inequality. $\square$

3. Optimal transport on configuration space

We denote the set of probability measures on Υ by $\mathscr{P}(\Upsilon)$. For $\mu, \nu \in \mathscr{P}(\Upsilon)$ the L^2-Wasserstein distance is defined via

$$W_2^2(\mu,\nu) := \inf_{q\in\mathsf{Cpl}(\mu,\nu)} \int d_\Upsilon^2(\gamma,\eta)\ q(d\gamma,d\eta)\ ,$$

where $\mathsf{Cpl}(\mu,\nu)$ denotes the set of all couplings between μ and ν. A minimizer is called optimal coupling and the set of all optimal couplings between μ and ν will be denoted by $\mathsf{Opt}(\mu,\nu)$. This transportation problem has been studied in the case of $M = \mathbb{R}^k$ in [12]; the generalization to Riemannian manifolds is straightforward. The main result states

Theorem 3.1 ([12])**.** *Let $\mu, \nu \in \mathscr{P}(\Upsilon)$ with $W_2(\mu,\nu) < \infty$. Assume that $\mu \ll \pi$. Then, there is a unique optimal coupling q which is induced by a transportation map, i.e. q is given as the pushforward $q = (id, T)_*\mu$ of μ under the map (id, T). More precisely, for all measurable $f : \Upsilon \times \Upsilon \to \mathbb{R}$.*

$$\int f(\gamma,\eta)\,\mathrm{d}q(\gamma,\eta) \ = \ \int f\big(\gamma,T(\gamma)\big)\,\mathrm{d}\mu(\gamma)\ .$$

3.1. **Duality and Hopf–Lax semigroup.** By general theory, see [17, Thm. 2.2], we have the following Kantorovich duality

Theorem 3.2. *Let $\mu, \nu \in \mathscr{P}(\Upsilon)$ such that $W_2(\mu,\nu) < \infty$. Then we have*

$$\frac{1}{2}W_2^2(\mu,\nu) \ = \ \sup\left\{\int \varphi^c\,\mathrm{d}\nu + \int \varphi\,\mathrm{d}\mu \ : \ \varphi \in C_b(\Upsilon)\right\}\ , \tag{3.1}$$

where the c-transform *of φ is defined by*

$$\varphi^c(\gamma) \ = \ \inf_{\eta\in\Upsilon}\left\{\frac{1}{2}d_\Upsilon^2(\gamma,\eta) - \varphi(\eta)\right\}\ .$$

It is not known if the supremum is attained or not. For a function $f : \Upsilon \to \mathbb{R}\cup\{\infty\}$ we define the Hopf–Lax semigroup

$$Q_t f(\gamma) \ = \ \inf_{\eta\in\Upsilon}\left\{f(\eta) + \frac{d_\Upsilon^2(\eta,\gamma)}{2t}\right\}\ .$$

The function $Q_t f$ is non trivial on the set

$$\mathcal{D}(f) := \{\gamma \in \Upsilon : d_\Upsilon(\gamma,\omega) < \infty \text{ for some } \omega \text{ with } f(\omega) < \infty\}.$$

For $\gamma \in \mathcal{D}(f)$ we set

$$t_*(\gamma) := \sup\{t > 0 : Q_t f(\gamma) > -\infty\}$$

with the convention that $t_*(\gamma) = 0$ if $Q_t f(\gamma) = -\infty$ for all $t > 0$. If f is bounded also $Q_t f$ is bounded, even d_Υ-Lipschitz (with global Lipschitz bound $\mathrm{Lip}(Q_t f) \leq 2\sqrt{\mathrm{osc}(f)/t}$ where $\mathrm{osc}(f) = \sup f - \inf f$), and $t_* = \infty$ for all γ. Note that if f is d_Υ-Lipschitz, so is $Q_t f$ with a priori bound ([6, Prop. 3.4])

$$|DQ_s\varphi| \ \leq \ 2\,\mathsf{Lip}(\varphi)\ . \tag{3.2}$$

Since (Υ, d_Υ) is a length space, this implies $\mathsf{Lip}(Q_s\varphi) \leq 2\,\mathsf{Lip}(\varphi)$. For more details we refer to Section 3 of [6]. In particular, if $f \in C_b(\Upsilon)$ then $Q_1(-f) = f^c$ is $d_\Upsilon-$Lipschitz. Hence, we have

Corollary 3.3. *Let $\mu, \nu \in P(\Upsilon)$ such that $W_2(\mu,\nu) < \infty$. Then we have*

$$\frac{1}{2}W_2^2(\mu,\nu) \;=\; \sup\left\{\int \varphi^c \,\mathrm{d}\nu + \int \varphi \,\mathrm{d}\mu \;:\; \varphi \in \mathsf{Lip}_b(\Upsilon) \cap C(\Upsilon)\right\}\,, \tag{3.3}$$

Recall the local Lipschitz constant from (2.2). The next proposition states that the Hopf–Lax semigroup yields a solution of the Hamilton–Jacobi equation.

Proposition 3.4. [6, Thm. 3.6] *For $\gamma \in \mathcal{D}(f)$ and $t \in (0, t_*(\gamma))$ it holds that*

$$\frac{\mathrm{d}}{\mathrm{d}t} Q_t f(\gamma) + \frac{|DQ_t f(\gamma)|^2(\gamma)}{2} \;=\; 0$$

with at most countably many exceptions in $(0, t_*(\gamma))$.

4. Manifestations of curvature on Υ

In this section we derive several curvature properties of the configuration space by "lifting" the corresponding statement from the base manifold M to Υ.

4.1. **Sectional curvature bounds.** We start by showing that the configuration space inherits Alexandrov curvature bounds from the base space.
By Toponogov's triangle comparison theorem a lower bound on the sectional curvature of a Riemannian manifold can be characterized by a condition involving only the distance function. This allows to generalize the notion of sectional curvature bounds to metric spaces and gives rise to Alexandrov spaces. Loosely put, an Alexandrov space with curvature bounded below by $K \in \mathbb{R}$ is a complete length space (X,d) in which triangles are "thicker" than in the space form of constant curvature K. We refer to [8] for a nice and comprehensive treatment of Alexandrov geometry. There are various equivalent ways of characterizing Alexandrov curvature. We will use the following taken from [21]:

Definition 4.1. *A complete length space (X,d) is an Alexandrov space with curvature bounded below by $K \in \mathbb{R}$ iff the following holds: For each quadruple of points $x_0, x_1, x_2, x_3 \in X$ we have:*

$$\begin{aligned} \sum_{i=1}^{3} d^2(x_0,x_i) \;&\geq\; \frac{1}{6}\sum_{i,j=1}^{3} d^2(x_i,x_j)\,, && \textit{if } K = 0\,,\\ \left(\sum_{i=1}^{3}\cosh\left(\sqrt{|K|}d(x_0,x_i)\right)\right)^2 \;&\geq\; \sum_{i,j=1}^{3}\cosh\left(\sqrt{|K|}d(x_i,x_j)\right)\,, && \textit{if } K < 0\,.\\ \left(\sum_{i=1}^{3}\cos\left(\sqrt{K}d(x_0,x_i)\right)\right)^2 \;&\leq\; \sum_{i,j=1}^{3}\cos\left(\sqrt{K}d(x_i,x_j)\right)\,, && \textit{if } K > 0\,. \end{aligned} \tag{4.1}$$

Remark 4.2. There is a variant of this characterization by Sturm, [32]. The proof of Theorem 4.3 adapts with only minor changes.

Note in particular that the Riemannian manifold M has sectional curvature bounded below by K if and only if its Riemannian distance d satisfies (4.1). Definition 4.1 does not apply immediately to extended metric spaces such as the configuration space (Υ, d_Υ). However, considering the fibers $\Upsilon_\sigma := \{\gamma \in \Upsilon : d_\Upsilon(\gamma,\sigma) < \infty\}$, we note that $(\Upsilon_\sigma, d_\Upsilon)$ is a complete length metric space for each $\sigma \in \Upsilon$.

Theorem 4.3. *Assume that the base manifold M has sectional curvature bounded below by $K \in \mathbb{R}$. Then (any fiber of) (Υ, d_Υ) is an Alexandrov space with curvature bounded below by* $\min\{K, 0\}$ *in the sense of Definition 4.1.*

Proof. We will only consider the case $K < 0$, the case $K = 0$ follows by similar arguments or alternatively can be obtained from this by letting $K \nearrow 0$. Obviously the case $K > 0$ is reduced immediately to $K = 0$. We will verify the quadruple comparison inequality. So let $\gamma_0, \gamma_1, \gamma_2, \gamma_3 \in \Upsilon$ such that $d_\Upsilon(\gamma_0, \gamma_i) < \infty$ for $i = 1, 2, 3$ (and hence also $d_\Upsilon(\gamma_i, \gamma_j) < \infty$). In particular, we have $\gamma_i(M) = C$ for all $i = 0, 1, 2, 3$ and some $C \in \mathbb{N} \cup \{+\infty\}$. We will assume $C = +\infty$, the case $C < \infty$ follows from the same arguments and is simpler. Using Lemma 2.6 we can choose labelings $\gamma_i = \sum_n \delta_{x_n^i}$ for $i = 0, 1, 2, 3$ such that

$$d_\Upsilon^2(\gamma_0, \gamma_i) \;=\; \sum_{n=1}^{\infty} d^2\big(x_n^0, x_n^i\big) \;<\; \infty\,. \tag{4.2}$$

Further we can estimate for $i, j = 1, 2, 3$:

$$d_\Upsilon^2(\gamma_i, \gamma_j) \;\le\; \sum_{n=1}^{\infty} d^2\big(x_n^i, x_n^j\big) \;<\infty \quad , \tag{4.3}$$

where finiteness follows from the triangle inequality in (M, d) and (4.2). Using the fact that for any $N \in \mathbb{N}$ the product manifold M^N with Riemannian distance $d_N^2\big((x_1, \cdots, x_N), (y_1, \cdots, y_N)\big) = \sum_{n=1}^N d^2(x_n, y_n)$ has sectional curvature bounded below by K and thus satisfies quadruple comparison, we get setting $\lambda = \sqrt{|K|}$:

$$\begin{aligned}
\left(\sum_{i=1}^{3} \cosh\big(\lambda d_\Upsilon(\gamma_0, \gamma_i)\big)\right)^2 \;&=\; \lim_{N\to\infty} \sum_{i=1}^{3} \cosh\Big(\lambda \sqrt{\sum_{n=1}^{N} d^2(x_n^0, x_n^i)}\Big) \\
&\ge\; \lim_{N\to\infty} \sum_{i,j=1}^{3} \cosh\Big(\lambda \sqrt{\sum_{n=1}^{N} d^2(x_n^i, x_n^j)}\Big) \\
&=\; \sum_{i,j=1}^{3} \cosh\Big(\lambda \sqrt{\sum_{n=1}^{\infty} d^2(x_n^i, x_n^j)}\Big) \\
&\ge\; \sum_{i,j=1}^{3} \cosh\big(\lambda d_\Upsilon(\gamma_i, \gamma_j)\big)\;,
\end{aligned}$$

where the last inequality follows from (4.3) and the fact that cosh is increasing. This finishes the proof. □

4.2. **Bochner inequality on configuration space.** Starting from this section we will be concerned with lower bounds on the Ricci curvature. Let us recall the Bochner–Weitzenböck identity which asserts that for every smooth function $u : M \to \mathbb{R}$ on the Riemannian manifold M we have:

$$\frac{1}{2}\Delta|\nabla u|^2 - \langle \nabla u, \nabla \Delta u\rangle = \|\mathrm{Hess}\, u\|_{HS}^2 + \mathrm{Ric}[\nabla u, \nabla u]\;,$$

where $\|\cdot\|_{HS}$ denotes the Hilbert–Schmidt norm and Ric denotes the Ricci tensor. Thus a lower bound on the Ricci curvature in the form $\mathrm{Ric}[\nabla u, \nabla u] \ge K|\nabla u|^2$ is seen to be equivalent to the *Bochner inequality*

$$\frac{1}{2}\Delta|\nabla u|^2 - \langle \nabla u, \nabla \Delta u\rangle \geq K|\nabla u|^2 \ .$$

It will be convenient to introduce the *carré du champ operators*, defined for smooth functions $\varphi, \psi : M \to \mathbb{R}$ via

$$\begin{aligned} \Gamma(\varphi,\psi) \ &:= \ \frac{1}{2}\left[\Delta\big(\varphi\psi\big) - \varphi\Delta\psi - \psi\Delta\varphi\right] \ = \ \langle\nabla\varphi,\nabla\psi\rangle \ , \\ \Gamma_2(\varphi,\psi) \ &:= \ \frac{1}{2}\left[\Delta\Gamma(\varphi,\psi) - \Gamma(\varphi,\Delta\psi)\Gamma(\psi,\Delta\varphi)\right] \ . \end{aligned}$$

In particular, writing $\Gamma(\varphi) = \Gamma(\varphi,\varphi)$ and $\Gamma_2(\varphi) = \Gamma_2(\varphi,\varphi)$ we see $\Gamma_2(\varphi) = \frac{1}{2}\Delta|\nabla\varphi|^2 - \langle\nabla\varphi,\nabla\Delta\varphi\rangle$. Thus the Bochner inequality takes the form

$$\Gamma_2(\varphi) \ \geq \ K \ \Gamma(\varphi) \ .$$

The latter inequality has been used extensively in the study of general Markov semigroups and diffusions, originating in the work of Bakry–Émery [7], where Δ is replaced by the generator of the semigroup.
The aim of this section is to prove the natural analogue of Bochner's inequality on the configuration space. For smooth cylinder functions $F, G \in \mathsf{Cyl}^\infty(\Upsilon)$ we define

$$\begin{aligned} \Gamma^\Upsilon(F,G) \ &:= \ \frac{1}{2}\left[\Delta^\Upsilon(FG) - F\Delta^\Upsilon G - G\Delta^\Upsilon F\right] \ = \ \langle\nabla^\Upsilon F,\nabla^\Upsilon G\rangle \ , \\ \Gamma_2^\Upsilon(F,G) \ &:= \ \frac{1}{2}\left[\Delta^\Upsilon\Gamma^\Upsilon(F,G) - \Gamma^\Upsilon(F,\Delta^\Upsilon G) - \Gamma^\Upsilon(G,\Delta^\Upsilon F)\right] \ . \end{aligned}$$

Note that Γ^Υ coincides with the carré du champ operator of the Dirichlet form $\mathcal{E}$ introduced in Section 2.1.

Proposition 4.4. *Assume that M has Ricci curvature bounded below by K. Then any cylinder function $F \in \mathsf{Cyl}^\infty(\Upsilon)$ satisfies the following Bochner inequality:*

$$\Gamma_2^\Upsilon(F)(\gamma) \ \geq K \ \Gamma^\Upsilon(F)(\gamma) \qquad \forall\gamma\in\Upsilon \ . \tag{4.4}$$

Proof. The cylinder function F takes the form $F(\gamma) = g\big(\langle\varphi_1,\gamma\rangle,\ldots,\langle\varphi_n,\gamma\rangle\big)$, where $g \in C^\infty(\mathbb{R}^n)$ and $\varphi_i \in C_c^\infty(M)$ for $i = 1,\ldots,n$. From the definition of gradient and divergence on Υ a direct calculation yields:

$$\Gamma^\Upsilon(F)(\gamma) \ = \ \sum_{i,j} g_i(\varphi)g_j(\varphi)\langle\nabla\varphi_i,\nabla\varphi_j\rangle_\gamma \ = \ \sum_{i,j} g_i(\varphi)g_j(\varphi)\langle\Gamma(\varphi_i,\varphi_j),\gamma\rangle \ ,$$

where we write $g_i = \partial_i g$. Moreover, we obtain

$$
\begin{aligned}
&\Gamma_2^\Upsilon(F)(\gamma)\\
&= \sum_{i,j} g_i(\varphi)g_j(\varphi)\langle \frac{1}{2}\Delta\langle\nabla\varphi_i,\nabla\varphi_j\rangle - \langle\nabla\varphi_i,\nabla\Delta\varphi_j\rangle,\gamma\rangle\\
&\quad + \sum_{i,j,k,l} g_{ik}(\varphi)g_{jl}(\varphi)\langle\nabla\varphi_i,\nabla\varphi_j\rangle_\gamma\langle\nabla\varphi_k,\nabla\varphi_l\rangle_\gamma\\
&\quad + \sum_{i,j,k} g_i(\varphi)g_{jk}(\varphi)\Big[2\langle\nabla\langle\nabla\varphi_i,\nabla\varphi_k\rangle,\nabla\varphi_j\rangle_\gamma - \langle\nabla\langle\nabla\varphi_j,\nabla\varphi_k\rangle,\nabla\varphi_i\rangle_\gamma\Big]\\
&= \sum_{i,j} g_i(\varphi)g_j(\varphi)\langle\Gamma_2(\varphi_i,\varphi_j),\gamma\rangle + \sum_{i,j,k,l} g_{ik}(\varphi)g_{jl}(\varphi)\langle\Gamma(\varphi_i,\varphi_j),\gamma\rangle\langle\Gamma(\varphi_k,\varphi_l),\gamma\rangle\\
&\quad + \sum_{i,j,k} g_i(\varphi)g_{jk}(\varphi)\Big[\langle 2\Gamma\big(\varphi_j,\Gamma(\varphi_i,\varphi_k)\big),\gamma\rangle - \langle\Gamma\big(\varphi_i,\Gamma(\varphi_j,\varphi_k)\big),\gamma\rangle\Big]\ .
\end{aligned}
$$

Choose a compact set K containing all the supports of φ_i for $i = 1,\ldots,n$. Fix a configuration γ, let $N = \gamma(K)$ and write $\gamma|_K = \sum_{\alpha=1}^N \delta_{x_\alpha}$. Define functions $\psi_i : M^N \to \mathbb{R}$ via $\psi_i(y_1,\cdots,y_N) = \sum_{\alpha=1}^N \varphi_i(y_\alpha) = \langle\varphi_i,\gamma\rangle$. By the tensorization property (4.7) and the chain rule (4.6) of the carré du champ operators given by Lemma 4.5 below we obtain for $\mathbf{x} = (x_1,\cdots,x_N)$:

$$
\begin{aligned}
\Gamma_2^\Upsilon(F)(\gamma) &= \sum_{i,j} g_i(\psi)g_j(\psi)\Gamma_2^{(N)}(\psi_i,\psi_j)(\mathbf{x})\\
&\quad + \sum_{i,j,k,l} g_{ik}(\psi)g_{jl}(\psi)\Gamma^{(N)}(\psi_i,\psi_j)(\mathbf{x})\Gamma^{(N)}(\psi_k,\psi_l)(\mathbf{x})\\
&\quad + \sum_{i,j,k} g_i(\psi)g_{jk}(\psi)\Big[2\Gamma^{(N)}\big(\psi_j,\Gamma^{(N)}(\psi_i,\psi_k)\big)(\mathbf{x})\\
&\qquad\qquad - \Gamma^{(N)}\big(\psi_i,\Gamma^{(N)}(\psi_j,\psi_k)\big)(\mathbf{x})\Big]\\
&= \Gamma_2^{(N)}(g(\psi))(\mathbf{x})\ .
\end{aligned}
$$

Applying Bochner's inequality on M^N, which has Ricci curvature bounded below by K as well, and using (4.7), (4.5) we get:

$$
\begin{aligned}
\Gamma_2^\Upsilon(F)(\gamma) &= \Gamma_2^{(N)}(g(\psi))(\mathbf{x}) \ \geq\ K\ \Gamma^{(N)}(g(\psi))(\mathbf{x})\\
&= \sum_{i,j} g_i(\varphi)g_j(\varphi)\langle\Gamma(\varphi_i,\varphi_j),\gamma\rangle \ =\ K\ \Gamma^\Upsilon(F)(\gamma)\ ,
\end{aligned}
$$

which finishes the proof. □

The following lemma summarizes tensorization properties and a chain rule for the carré du champ operators which are readily verified by direct computations.

Lemma 4.5. *Let M be a smooth Riemannian manifold. Let $g \in C^\infty(\mathbb{R}^n)$ and $\psi_i \in C_c^\infty(M)$ for $i = 1,\ldots,n$ and write $\psi = \big(\psi_1,\cdots,\psi_n\big) \in C_c^\infty(M,\mathbb{R}^n)$. Then we* *have:*

(4.5)

$$\Gamma\big(g(\psi)\big) \;=\; \sum_{i,j=1}^{n} g_i(\psi)g_j(\psi)\Gamma(\psi_i,\psi_j)\;,$$

$$\Gamma_2\big(g(\psi)\big) \;=\; \sum_{i,j=1}^{n} g_i(\psi)g_j(\psi)\Gamma_2(\psi_i,\psi_j) + \sum_{i,j,k,l=1}^{n} g_{ik}(\psi)g_{jl}(\psi)\Gamma(\psi_i,\psi_j)\Gamma(\psi_k,\psi_l)$$

$$+\sum_{i,j,k=1}^{n} g_i(\psi)g_{jk}(\psi)\left[2\Gamma\big(\psi_j,\Gamma(\psi_i,\psi_k)\big)-\Gamma\big(\psi_i,\Gamma(\psi_j,\psi_k)\big)\right]\;. \tag{4.6}$$

Moreover, for $N\in\mathbb{N}$ let M^N be the N-fold tensor product of the Riemannian manifold M and denote by $\Gamma^{(N)},\Gamma_2^{(N)}$ the carré du champ operators associated to the Laplace–Beltrami operator on M^N. Let $\psi: M^N\to\mathbb{R}$ be given for $\mathbf{x}=(x_1,\cdots,x_N)$ by $\psi(\mathbf{x})=\sum_{\alpha=1}^N\varphi(x_\alpha)$ for a function $\varphi\in C_c^\infty(M)$. Then we have:

$$\Gamma^{(N)}(\psi)(\mathbf{x}) \;=\; \sum_{\alpha=1}^{N}\Gamma(\varphi)(x_\alpha)\;,\qquad \Gamma_2^{(N)}(\psi)(\mathbf{x}) \;=\; \sum_{\alpha=1}^{N}\Gamma_2(\varphi)(x_\alpha)\;. \tag{4.7}$$

More generally we have the following weak form of Bochner's inequality.

Proposition 4.6. *Assume that $\mathrm{Ric}_M\geq K$. Then for all non-negative $G\in D(\Delta^\Upsilon)$ with $G,|\nabla^\Upsilon G|,\Delta^\Upsilon G\in L^\infty(\Upsilon,\pi)$ and all $F\in D(\Delta^\Upsilon)$ we have:*

(4.8)

$$\int \frac{1}{2}\Delta^\Upsilon G|\nabla^\Upsilon F|^2 + G(\Delta^\Upsilon F)^2 + \Delta^\Upsilon F\langle\nabla^\Upsilon G,\nabla^\Upsilon F\rangle\,\mathrm{d}\pi \;\geq\; K\int G|\nabla^\Upsilon F|^2\,\mathrm{d}\pi\;.$$

Proof. First let F be a cylinder function. Multiplying (4.4) by G and integrating we obtain (4.8) immediately by applying the Leibniz rule

$$G\langle\nabla^\Upsilon F,\Delta^\Upsilon\nabla^\Upsilon F\rangle = \langle\nabla^\Upsilon F,\nabla^\Upsilon(G\Delta^\Upsilon F)\rangle - \Delta^\Upsilon F\langle\nabla^\Upsilon G,\nabla^\Upsilon F\rangle$$

and an integration by parts. For general $F\in D(\Delta^\Upsilon)\subset\mathcal{F}$ we argue by approximation. We can take a sequence $(F_n)\subset\mathsf{Cyl}^\infty(\Upsilon)$ such that $F_n\to F$, $|\nabla^\Upsilon F_n|\to|\nabla^\Upsilon F|$ and $\Delta^\Upsilon F_n\to\Delta^\Upsilon F$ in $L^2(\Upsilon,\pi)$ (recall that Δ^Υ is a closed operator). By the boundedness of $G,|\nabla^\Upsilon G|$ and $\Delta^\Upsilon G$ we can pass to the limit in the integrals and obtain (4.8). □

4.3. **Gradient estimates on Υ.** It is well known that the lower curvature bound $\mathrm{Ric}_M\geq K$ is equivalent to the following gradient estimate for the heat semigroup $T_t^M=\mathrm{e}^{t\Delta}$ on M, see e.g. [34, Thm. 1.3] and the discussion thereafter. For all smooth $f:M\to\mathbb{R}$, all $x\in M$ and $t>0$:

$$\Gamma\big(T_t^M f\big)(x) \;\leq\; \mathrm{e}^{-2Kt}T_t^M\Gamma\big(f\big)(x)\;. \tag{4.9}$$

The aim of this section is to show the gradient estimate for the heat semigroup T_t^Υ on the configuration space. Recall that the Dirichlet form admits a carré du champs operator Γ^Υ such that for all $u\in\mathcal{F}$ we have $\Gamma^\Upsilon(u)(\gamma)=|\nabla^\Upsilon u|_\gamma^2$. We have the following

Theorem 4.7. *Assume that $\mathrm{Ric}_M\geq K$. Then for any function $F\in\mathcal{F}$ and all $t>0$ we have:*

$$\Gamma^\Upsilon\big(T_t^\Upsilon F\big) \;\leq\; \mathrm{e}^{-2Kt}T_t^\Upsilon\Gamma^\Upsilon\big(F\big)\quad \pi\text{-a.e.} \tag{4.10}$$

The strategy we follow will be to use the explicit representation of the semigroup T_t^Υ as an infinite product of one-particle semigroups and the tensorization property of the gradient estimate. Before we give the proof we need to introduce some notation. Recall that $\pi(\Upsilon^{(\infty)}) = 1$. To a measurable function F on $\Upsilon^{(\infty)}$ we associate $\hat{F} : M^{\mathbb{N}} \to \mathbb{R}$ via

$$\hat{F}(\mathbf{x}) := F\left(\sum_{i\geq 1} \delta_{x_i}\right), \quad \mathbf{x} = (x_i)_{i\geq 1} \in M^{\mathbb{N}} ,$$

which is measurable with respect to the product $\sigma-$algebra on $M^{\mathbb{N}}$. Then, also the function $\hat{F}^i_{\mathbf{x}} : M \to \mathbb{R}$ defined by

$$\hat{F}^i_{\mathbf{x}}(y) := F\left(\sum_{j\geq 1, j\neq i} \delta_{x_j} + \delta_y\right)$$

is measurable. We say that $\hat{F}$ is differentiable in $\mathbf{x}$ if for each $i \geq 1$ the gradient in the i-th direction

$$\nabla^i \hat{F}(\mathbf{x}) := \nabla \hat{F}^i_{\mathbf{x}}(x_i) ,$$

exists. We say that $\hat{F}$ is differentiable with finite gradient if additionally

$$|\nabla^{\mathbb{N}} \hat{F}|^2(\mathbf{x}) := \sum_{i\geq 1} |\nabla^i \hat{F}|^2_{x_i}(\mathbf{x}) < \infty .$$

Then, for every $F \in \mathsf{Cyl}^\infty(\Upsilon), \gamma \in \Upsilon^{(\infty)}$ and $\mathbf{x} \in l^{-1}(\gamma)$ we have

$$\Gamma^\Upsilon(F) = |\nabla^\Upsilon F|^2_\gamma = |\nabla^{\mathbb{N}} \hat{F}|^2(\mathbf{x}) .$$

We will put

$$T^i_t \hat{F}(\mathbf{x}) = T^M_t \hat{F}^i_{\mathbf{x}}(x_i) ,$$

i.e. the action of the one-particle semigroup in the i-th coordinate. With this notation we can express the semigroup $T^{\mathbb{N}}_t$ introduced in Section 2.3 as $T^{\mathbb{N}}_t = \Pi_{j\in\mathbb{N}} T^j_t$, the iterated application of the one-particle semigroup in all directions. For $i \in \mathbb{N}$ we will also put

$$T^{\check{i}}_t = \prod_{j\in\mathbb{N}, j\neq i} T^j_t .$$

Proof of Theorem 4.7. Let us first assume that $F \in \mathsf{Cyl}^\infty(\Upsilon)$ and start by establishing a gradient estimate for $\hat{F}$. First note that by (4.9) for any $i \in \mathbb{N}$ and $\mathbf{x} \in M^{\mathbb{N}}$:

$$|\nabla^i T^i_t \hat{F}|^2(\mathbf{x}) \;=\; |\nabla T^M_t \hat{F}^i_{\mathbf{x}}|^2(x_i) \;\leq\; \mathrm{e}^{-2Kt} T^M_t |\nabla \hat{F}^i_{\mathbf{x}}|^2(x_i) \;=\; \mathrm{e}^{-2Kt} T^i_t |\nabla^i \hat{F}|^2(\mathbf{x}) .$$

By Jensen's inequality this yields

$$|\nabla^i T^{\mathbb{N}}_t \hat{F}|^2(\mathbf{x}) \;\leq\; T^{\check{i}}_t |\nabla^i T^i_t \hat{F}|^2(\mathbf{x}) \;\leq\; \mathrm{e}^{-2Kt} T^{\mathbb{N}}_t |\nabla^i \hat{F}|^2(\mathbf{x}) ,$$

and summing over i we obtain

$$|\nabla^{\mathbb{N}} T^{\mathbb{N}}_t \hat{F}|^2(\mathbf{x}) \;\leq\; \mathrm{e}^{-2Kt} T^{\mathbb{N}}_t |\nabla^{\mathbb{N}} \hat{F}|^2(\mathbf{x}) \;<\; \infty . \tag{4.11}$$

In particular $T^{\mathbb{N}}_t \hat{F}$ is differentiable with finite gradient. Note that the right hand side is also bounded above by a constant. We now want to pass from the estimate on $M^{\mathbb{N}}$ to an estimate on Υ. Note that for any good configuration $\gamma \in \Theta$ and $\mathbf{x} \in l^{-1}(\gamma)$:

$$T_t^{\mathbb{N}}|\nabla^{\mathbb{N}}\hat{F}|^2(\mathbf{x}) \;=\; \tilde{T}_t^{\Upsilon}|\nabla^{\Upsilon}F|^2(\gamma) \;=:\; G(\gamma)\;.$$

We claim that $\tilde{T}_t^{\Upsilon}F$ is d_{Υ}-Lipschitz on Θ and that $|D\tilde{T}_t^{\Upsilon}F| \leq \mathrm{e}^{-2Kt}G$. By Lemma 2.2 this will suffice to show (4.10). Indeed, consider $V \in \mathcal{V}_0(M)$ and its flow $(\psi_t)_t$. Then we have

$$\begin{aligned}
|\tilde{T}_t^{\Upsilon}F(\psi_1^*\gamma) - \tilde{T}_t^{\Upsilon}F(\gamma)| \;&\leq\; \left|\int_0^1 \frac{\mathrm{d}}{\mathrm{d}s}\tilde{T}_t^{\Upsilon}F(\psi_s^*\gamma)\,\mathrm{d}s\right|\\
&=\; \left|\int_0^1 \sum_i \langle \nabla^i T_t^{\mathbb{N}}\hat{F}, V\rangle(\psi_s^*\mathbf{x})\,\mathrm{d}s\right|\\
&\leq\; \mathrm{e}^{-Kt}\int_0^1 \sqrt{G(\psi_s^*\gamma)}|V|_{\psi_s^*\gamma}\,\mathrm{d}s\\
&=\; d_{\Upsilon}(\psi_1^*\gamma,\gamma)\mathrm{e}^{-Kt}\int_0^1 \sqrt{G(\psi_s^*\gamma)}\,\mathrm{d}s\;.
\end{aligned}$$

Thus $\tilde{T}_t^{\Upsilon}F$ is Lipschitz by the boundedness of G. Arguing as in the proof of Lemma 2.2 by letting $|V|_\gamma \to 0$ yields the claim by continuity of G.
Now take $F \in \mathcal{F}$. Then there is a sequence $(F_n)_{n\in\mathbb{N}} \subset \mathsf{Cyl}^\infty(\Upsilon)$ such that $F_n \to F$ in $L^2(\Upsilon,\pi)$ and $\mathcal{E}(F - F_n) \to 0$. Therefore, denoting by Λ the measure $\Lambda(dx, d\gamma) := \gamma(dx)\pi(d\gamma)$, $\nabla^{\Upsilon}F_n$ is a Cauchy sequence in $L^2(M\times\Upsilon \to TM, \Lambda)$. Therefore, there is a limit, denoted by $\nabla^{\Upsilon}F$, such that $\mathcal{E}(F) = \int |\nabla^{\Upsilon}F|^2\ \mathrm{d}\pi$. As T_t^{Υ} is a contraction also $T_t^{\Upsilon}F_n \to T_t^{\Upsilon}F$, $T_t^{\Upsilon}|\nabla^{\Upsilon}F_n|^2 \to T_t^{\Upsilon}|\nabla^{\Upsilon}F_n|^2$ and $\nabla^{\Upsilon}T_t^{\Upsilon}F_n$ is a Cauchy sequence with some limit G, by (4.10). By lower semicontinuity of the carré du champ operator (see e.g. [5, (2.17)]) we have $\Gamma^{\Upsilon}(T_t^{\Upsilon}F)(\gamma) \leq |G|^2(\gamma)$ π-a.e.. In the first part of the proof, we saw that (4.10) holds for all F_n, i.e.

$$|\nabla^{\Upsilon}T_t^{\Upsilon}F_n|^2 \leq \mathrm{e}^{-2Kt}T_t^{\Upsilon}(|\nabla^{\Upsilon}F_n|^2) \quad \pi\text{-a.e.}$$

Extracting a subsequence, this yields

$$|\nabla^{\Upsilon}T_t^{\Upsilon}F|^2(\gamma) \leq |G|^2(\gamma) \leq \mathrm{e}^{-2Kt}T_t^{\Upsilon}(|\nabla^{\Upsilon}F|^2)(\gamma) \quad \pi\text{-a.e.}$$

□

Remark 4.8. Alternatively, the gradient estimate could have been derived from the Bochner inequality from the previous section. In fact, a classical interpolation argument due to Bakry–Émery yields the equivalence of the Γ_2-inequality (4.4) and the gradient estimate (4.10). The idea is to consider

$$\varphi(s) = \mathrm{e}^{-2Ks}T_s^{\Upsilon}\Gamma^{\Upsilon}(T_{t-s}^{\Upsilon}F)$$

and note that $\varphi'(s) = \mathrm{e}^{-2Ks}T_s^{\Upsilon}\left[\Gamma_2^{\Upsilon}(T_{t-s}^{\Upsilon}F) - K\Gamma^{\Upsilon}(T_{t-s}^{\Upsilon})\right]$. For a detailed proof in a general setting see e.g. [5, Cor. 2.3]. However, in order to apply this in the present setting one would need to extend (4.4) (in a weak form) to a larger class of functions.

4.4. **Wasserstein contraction.** In [34] it has been shown that a lower bound on the Ricci curvature is also equivalent to expansion bounds in Wasserstein distance for the heat kernel. More precisely, [34, Cor. 1.4] states that $\mathrm{Ric}_M \geq K$ if and only if

$$W_{p,d}(H_t^M\mu, H_t^M\nu) \;\leq\; \mathrm{e}^{-Kt} W_{p,d}(\mu,\nu) \quad \forall t>0,\ \mu,\nu \in \mathscr{P}_p(M)\ . \tag{4.12}$$

Here for $p \geq 1$ we denote by $W_{p,d}$ the L^p-Wasserstein distance built from the Riemannian distance d and by $\mathscr{P}_p(M)$ the space of probability measures μ on M satisfying $\int d^p(x_0,x)\,\mathrm{d}\mu(x) < \infty$ for some (hence any) $x_0 \in M$. $H_t^M\mu \in \mathscr{P}(M)$ denotes the probability measure defined by

$$(H_t^M\mu)(A) = \int_A \int_M p_t^M(x,y)\,\mathrm{d}\mu(x)\,\mathrm{d}\,\mathrm{vol}(y) \qquad \forall A \in \mathcal{B}(M)\ ,$$

where p_t^M is the heat kernel on M.
Here we will show that the heat semigroup on the configuration space has the corresponding expansion bound in Wasserstein distance provided $\mathrm{Ric}_M \geq K$. Recall the set of good configurations Θ from (2.6). A probability measure μ on Υ is called *good* if it is concentrated on the good configurations, i.e. $\mu(\Theta)=1$. We denote the set of all good probability measures by $\mathscr{P}_g(\Upsilon)$. Note that in particular any measure absolutely continuous w.r.t. π and all Dirac measures $\delta_\gamma \in \mathscr{P}(\Upsilon)$ with $\gamma \in \Theta$ are good. Let p_t^Υ be the semigroup of Markov kernels on Θ given by Theorem 2.4 (i.e. the transition probabilities of the independent particle process). Given $\mu \in \mathscr{P}_g(\Upsilon)$ we define $H_t^\Upsilon\mu$ via

$$H_t^\Upsilon\mu(A) = \int_\Theta p_t^\Upsilon(\gamma, A)\,\mathrm{d}\mu(\gamma)\ . \tag{4.13}$$

Theorem 4.9. *Assume that* $\mathrm{Ric}_M \geq K$. *Then for all* $\mu,\nu \in \mathscr{P}_g(\Upsilon)$ *we have:*

$$W_{2,d_\Upsilon}(H_t^\Upsilon\mu, H_t^\Upsilon\nu) \;\leq\; \mathrm{e}^{-Kt} W_{2,d_\Upsilon}(\mu,\nu) \quad \forall t>0\ . \tag{4.14}$$

Proof. First we show that for all $\gamma,\sigma \in \Theta$ we have:

$$W_{2,d_\Upsilon}\big(p_t^\Upsilon(\gamma,\cdot), p_t^\Upsilon(\sigma,\cdot)\big) \;\leq\; \mathrm{e}^{-Kt} d_\Upsilon(\gamma,\sigma) \quad \forall t>0\ . \tag{4.15}$$

We can assume that $d_\Upsilon(\gamma,\sigma)<\infty$ and consider only the case $\gamma(M)=\sigma(M)=\infty$. Then by Lemma 2.6 there exist labelings $\gamma = \sum_{i=1}^\infty \delta_{x_i}$ and $\sigma = \sum_{i=1}^\infty \delta_{y_i}$ such that $d_\Upsilon^2(\gamma,\sigma) = \sum_{i=1}^\infty d^2(x_i,y_i)$. Now, for any i choose an optimal coupling $q_i \in \mathscr{P}(M\times M)$ of $p_t^M(x_i,\cdot)$ and $p_t^M(y_i,\cdot)$ such that

$$W_{2,d}^2\big(p_t^M(x_i,\cdot), p_t^M(y_i,\cdot)\big) = \int d^2(u,v)\,\mathrm{d}q_i(u,v)\ .$$

Let $q^{\mathbb{N}} = \bigotimes_{i=1}^\infty q_i \in \mathscr{P}(M^{\mathbb{N}}\times M^{\mathbb{N}})$ and set $q = (l\times l)_\# q^{\mathbb{N}}$, where l is the labelling map. Then $q \in \mathscr{P}(\Theta\times\Theta)$ defines a coupling of $p_t^\Upsilon(\gamma,\cdot)$ and $p_t^\Upsilon(\sigma,\cdot)$. Now we can estimate:

$$\begin{aligned}
W_{2,d_\Upsilon}^2\big(p_t^\Upsilon(\gamma,\cdot), p_t^\Upsilon(\sigma,\cdot)\big) \;&\leq\; \int d_\Upsilon^2\,\mathrm{d}q \;=\; \int d_\Upsilon^2\big(l(\boldsymbol{u}), l(\boldsymbol{v})\big)\,\mathrm{d}q^{\mathbb{N}}(\boldsymbol{u},\boldsymbol{v})\\
&\leq \sum_{i=1}^\infty \int d^2(u_i,v_i)\,\mathrm{d}q_i(u_i,v_i) \;=\; \sum_{i=1}^\infty W_{2,d}^2\big(p_t^M(x_i,\cdot), p_t^M(y_i,\cdot)\big)\\
&\leq\; \mathrm{e}^{-2Kt}\sum_{i=1}^\infty d^2(x_i,y_i) \;=\; \mathrm{e}^{-2Kt} d_\Upsilon^2(\gamma,\sigma)\ .
\end{aligned}$$

Here we have estimated d_Υ by the choice of a special labeling in the second inequality and used (4.12) in the third inequality. Finally, to prove (4.14) we can again assume that $W_{2,d_\Upsilon}(\mu,\nu)<\infty$ and choose an optimal coupling q of μ and ν. Then by convexity of the squared Wasserstein distance we get

$$\begin{aligned} W^2_{2,d_\Upsilon}(H^\Upsilon_t\mu,H^\Upsilon_t\nu) \;&\leq\; \int W^2_{2,d_\Upsilon}\big(p^\Upsilon_t(\gamma,\cdot),p^\Upsilon_t(\sigma,\cdot)\big)\,\mathrm{d}q(\gamma,\sigma)\\ &\leq\; \mathrm{e}^{-2Kt}\int d^2_\Upsilon(\gamma,\sigma)\,\mathrm{d}q(\gamma,\sigma) \;=\; \mathrm{e}^{-2Kt}W^2_{2,d_\Upsilon}(\mu,\nu)\;. \end{aligned}$$

□

Remark 4.10. The same argument as in the previous proof yields that for any $p\in[1,\infty]$ and any $\mu,\nu\in\mathscr{P}_g(\Upsilon)$:

$$W_{p,d_{\Upsilon,p}}(H^\Upsilon_t\mu,H^\Upsilon_t\nu)\;\leq\;\mathrm{e}^{-Kt}W_{p,d_{\Upsilon,p}}(\mu,\nu)\quad\forall t>0\;,$$

where $d_{\Upsilon,p}$ is the L^p-transport distance between non-normalized measures.
Moreover, combining the construction of the semigroup in (2.5) with [34, Cor. 1(x)] one can show along the lines of the previous proof that for any two good configurations γ and σ there exist a coupling $(\boldsymbol{B}^\gamma_t,\boldsymbol{B}^\sigma_t)$ of the two copies of the independent particle process in Θ on some probability space $(\Omega,\mathcal{F},\mathbb{P})$ starting in γ respectively σ such that

$$d_\Upsilon(\boldsymbol{B}^\gamma_t,\boldsymbol{B}^\sigma_t)\leq\mathrm{e}^{-Kt}\;d_\Upsilon(\gamma,\sigma)\quad\mathbb{P}\text{ a.s.}$$

5. Synthetic Riemannian Ricci curvature

It has been proven in [34] that M has Ricci curvature bounded below by K, if and only if the entropy is K-convex along geodesics in $\big(\mathscr{P}_2(M),W_2\big)$. This result has been the starting point for Sturm [33] and Lott–Villani [23] to define a notion of Ricci curvature for metric measure spaces.
The goal of this section is to show that the configuration space satisfies (a version for extended metric measure spaces of) this so-called $\mathrm{CD}(K,\infty)$ condition, provided $\mathrm{Ric}_M\geq K$. Unlike the previous results we will not obtain this by "lifting" the corresponding statement from the base space. Instead we will follow the approach in [5] and derive the so-called Evolution Variational Inequality starting from the gradient estimates established in Theorem 4.7. This will yield geodesic convexity as an immediate consequence and as a side product give the characterization of the heat semigroup on Υ as the gradient flow of the entropy.
The argument will follow closely the lines of [5, Sec. 4]. A careful inspection of the proofs given there in the case of a Dirichlet form with finite intrinsic distance, reveals that most of them carry over to the present setting of an extended metric measure space. We give a sketch of the arguments in Section 5.2 to make this transparent. However, in the configuration space setting we need to work significantly more to establish the required regularization properties of the heat semigroup. This is the purpose of Section 5.1. We assume from now on that $\mathrm{Ric}_M\geq K$.

5.1. **Regularizing properties of the dual semigroup.** We denote the set of probability measures absolutely continuous w.r.t. π by $\mathscr{P}_{ac}(\Upsilon)$. Given $\mu\in\mathscr{P}_{ac}(\Upsilon)$ with $\mu=f\pi$ we define the action of the dual semigroup H^Υ_t via

$$H^\Upsilon_t\mu\;=\;(T^\Upsilon_t f)\pi\;.$$

Note that this coincides with $H_t^\Upsilon \mu$ defined for good probability measures μ in (4.13). Thanks to the Wasserstein contractivity (4.14) we can extend H_t^Υ to a contractive semigroup on the closure of $\mathscr{P}_{ac}(\Upsilon)$ w.r.t. W_2.
Given $\mu \in \mathscr{P}_{ac}(\Upsilon)$ with $\mu = f\pi$ the *relative entropy* w.r.t. π is defined by

$$\mathrm{Ent}(\mu) \;=\; \int f \log f \,\mathrm{d}\pi \;.$$

If μ is not absolutely continuous we set $\mathrm{Ent}(\mu) = \infty$. Note that $\mathrm{Ent}(\mu) \geq 0$ for all $\mu \in \mathscr{P}(\Upsilon)$ since π is a probability measure. We write $D(\mathrm{Ent}) = \{\mu : \mathrm{Ent}(\mu) < \infty\}$. We will denote by $\mathscr{P}_e$ the set of all probability measures whose fiber contains a measure of finite entropy,

$$\mathscr{P}_e = \{\mu \in \mathscr{P}(\Upsilon) : \exists \nu \in D(\mathrm{Ent}),\; W_2(\mu,\nu) < \infty\}\;.$$

Lemma 5.1. *For any $\mu \in \mathscr{P}_e$ there exists a sequence of probability measures $(\mu_n)_n \in D(\mathrm{Ent})$ with $W_2(\mu_n,\mu) \to 0$ as $n \to \infty$. In particular, $H_t^\Upsilon \mu$ is defined for any such μ.*

The proof relies on an explicit construction but is rather lengthy and we postpone it to the appendix. Note that $\mathscr{P}_e$ is obviously the maximal set of measures that can be approximated in this way. To prove regularization properties of H_t^Υ we need to collect some estimates. The *Fisher information* of $\mu = f\pi$ with $\sqrt{f} \in \mathcal{F}$ is defined via

$$I(\mu) := 4\mathcal{E}(\sqrt{f})\;,$$

Otherwise we set $I(\mu) = \infty$. Note that we can also write

$$I(\mu) \;=\; \int_{\{f>0\}} \frac{\Gamma^\Upsilon(f)}{f} \,\mathrm{d}\pi\;,$$

and that I is convex on $\mathscr{P}_{ac}(\Upsilon)$, see [5, Prop. 4.1].
A curve $\mu : J \to \mathscr{P}(\Upsilon)$ is called *p-absolutely continuous* w.r.t. W_2 on an interval J, written $\mu \in AC^p\big(J, (\mathscr{P}(\Upsilon), W_2)\big)$, if there exist $a \in L^p(J, \mathrm{Leb})$ such that for all $s,t \in J$:

$$W_2(\mu_s,\mu_t) \;\leq\; \int_s^t a(r)\,\mathrm{Leb}(\,\mathrm{d}r)\;.$$

For any absolutely continuous curve $\mu : J \to \mathscr{P}(\Upsilon)$ the *metric derivative* defined by

$$|\dot\mu_t| \;=\; \lim_{h\to 0} \frac{W_2(\mu_{t+h},\mu_t)}{h}$$

exists for a.e. $t \in J$, see [3, Thm. 1.1.2].
Having identified the Dirichlet form $\mathcal{E}$ with the Cheeger energy Ch constructed from d_Υ in Proposition 2.3 yields in particular that T_t^Υ coincides with the gradient flow of Ch in $L^2(\Upsilon,\pi)$. This allows us to apply useful estimates for this gradient flow established in [6].

Lemma 5.2. *Let $\mu = f\pi$ with $\mathrm{Ent}(\mu) < \infty$ and set $\mu_t = (T_t^\Upsilon f)\pi$. Then the map $t \mapsto \mathrm{Ent}(\mu_t)$ is non-increasing, locally absolutely continuous. Moreover, we have for all $T > 0$:*

$$\int_0^T I(\mu_t)\,\mathrm{d}t \;\leq\; 2\,\mathrm{Ent}(\mu_0)\;. \tag{5.1}$$

The curve $t \mapsto \mu_t$ is absolutely continuous w.r.t. W_2 and for a.e. t:

$$|\dot{\mu}_t|^2 \ \leq \ I(\mu_t)\ . \tag{5.2}$$

Proof. That the entropy is non-increasing and (5.1) holds for $f \in L^1(\pi) \cap L^2(\pi)$ are proven in [6, Lem. 4.19, Prop. 4.22]. The general statement follows by a truncation argument using the lower semicontinuity of I in $L^1(\Upsilon,\pi)$ and of Ent w.r.t. weak convergence (and thus also in $L^1(\Upsilon,\pi)$). Finally [6, Lem. 6.1] gives (5.2). □

The following log-Harnack and entropy–cost inequalities will be crucial for the regularizing properties of the dual semigroup.

Lemma 5.3. *For any bounded Borel-measurable function $f : \Upsilon \to \mathbb{R}$ all $t > 0$ and $\gamma, \sigma \in \Theta$ we have:*

$$(\tilde{T}_t^\Upsilon \log f)(\gamma) \ \leq \ \log\big(\tilde{T}_t^\Upsilon f(\sigma)\big) + \frac{K}{2(1-\mathrm{e}^{-2Kt})}\ d_\Upsilon^2(\gamma,\sigma)\ . \tag{5.3}$$

In particular, for any $\mu \in \mathscr{P}_{ac}(\Upsilon)$ and $\nu \in D(\mathrm{Ent})$ we have:

$$\mathrm{Ent}(H_t^\Upsilon \mu) \ \leq \ \mathrm{Ent}(\nu) + \frac{K}{2(1-\mathrm{e}^{-2Kt})}\ W_2^2(\mu,\nu)\ . \tag{5.4}$$

Proof. (5.3) is proven in [13, Theorem 2.2] for $f \geq 1$. For general measurable $f \geq 0$ we apply this result to $f_\varepsilon = \varepsilon^{-1}(f+\varepsilon) \geq 1$ and obtain

$$\big(\tilde{T}_t^\Upsilon \log(f+\epsilon)\big)(\gamma) \ \leq \ \log\big(\tilde{T}_t^\Upsilon(f(\sigma)+\epsilon)\big) + \frac{K}{2(1-\mathrm{e}^{-2Kt})}\ d_\Upsilon^2(\gamma,\sigma)\ .$$

Letting $\varepsilon \to 0$ yields (5.3). To prove (5.4) consider $\mu = f\pi \in \mathscr{P}_{ac}(\Upsilon)$ and $\nu = g\pi \in D(\mathrm{Ent})$ with $W_2(\mu,\nu) < \infty$. Applying (5.3) with $\tilde{T}_t^\Upsilon f$ and integrating against an optimal coupling q of $H_t^\Upsilon \mu = (\tilde{T}_t^\Upsilon f)\pi$ and ν we obtain:

$$\begin{aligned}\mathrm{Ent}(H_t^\Upsilon \mu) \ &= \ \int \tilde{T}_t^\Upsilon f \log \tilde{T}_t^\Upsilon f \ \mathrm{d}\pi \\ &\leq \ \int \big(\log \tilde{T}_{2t}^\Upsilon f\big)\ \mathrm{d}\nu + \frac{K}{2(1-\mathrm{e}^{-2Kt})}\ W_2^2(\mu,\nu)\ .\end{aligned}$$

Using Jensen's inequality and the fact that $\int \tilde{T}_{2t}^\Upsilon f \ \mathrm{d}\pi = 1$ we estimate:

$$\begin{aligned}\int \big(\log \tilde{T}_{2t}^\Upsilon f\big)\ \mathrm{d}\nu \ &= \ \int \log\left(\frac{\tilde{T}_{2t}^\Upsilon f}{g}\right)\ \mathrm{d}\nu + \int \log g \ \mathrm{d}\nu \\ &\leq \ \log\left(\int \frac{\tilde{T}_{2t}^\Upsilon f}{g}\ \mathrm{d}\nu\right) + \mathrm{Ent}(\nu) \ = \ \mathrm{Ent}(\nu)\ ,\end{aligned}$$

which proves the claim. □

Consider the following mollification of the semigroup, defined for $\varepsilon > 0$ and $f \in L^p(\Upsilon,\pi)$, $1 \leq p \leq \infty$ via:

$$h^\varepsilon f \ = \ \int_0^\infty \frac{1}{\varepsilon}\eta\left(\frac{t}{\varepsilon}\right)\tilde{T}_t^\Upsilon f \ \mathrm{d}t\ , \tag{5.5}$$

with a non-negative kernel $\eta \in C_c^\infty(0,\infty)$ satisfying $\int_0^\infty \eta(t)\ \mathrm{d}t = 1$. Combining the convexity of I with (5.4) and (5.1) we obtain that for all $t > 0$ and all non-negative

$f \in L^1(\Upsilon, \pi)$ the measure $\mu = (h^\varepsilon f)\pi$ satisfies (see also [5, Lem. 4.9]):

$$I(H_t^\Upsilon \mu) \;\le\; C(\varepsilon)\Big(W_2^2(f\pi, \nu) + \mathrm{Ent}(\nu)\Big)\;, \tag{5.6}$$

where the constant $C(\varepsilon)$ on the right hand side depends only on ε.

Lemma 5.4. *For any $\mu \in \mathscr{P}_e$ and $t > 0$ we have $H_t^\Upsilon \mu \in D(\mathrm{Ent})$, $W_2(H_t^\Upsilon \mu, \mu) < \infty$ and moreover, $W_2(H_t^\Upsilon \mu, \mu) \to 0$ as $t \to 0$.*

Proof. First assume that $\mu = f\pi$ and $\mathrm{Ent}(\mu) < \infty$ and set $\mu_t = H_t^\Upsilon \mu$. Since H_t^Υ decreases the entropy we have also $\mu_t \in D(\mathrm{Ent})$. Further, we obtain by (5.2) and Hölder's inequality:

$$W_2^2(H_t^\Upsilon \mu, \mu) \;\le\; \int_0^t |\dot{\mu}_s|\,\mathrm{d}s \;\le\; \sqrt{t}\left(\int_0^t I(\mu_s)\,\mathrm{d}s\right)^{\frac{1}{2}}\;,$$

which goes to zero as $t \to 0$ by (5.1) and thus the lemma is established for $\mu \in D(\mathrm{Ent})$.

Now consider the general case where μ does not belong to $D(\mathrm{Ent})$. By Lemma 5.1, we can approximate it in W_2 by measures $\mu_n \in D(\mathrm{Ent})$. By Lemma 5.3 we have for some $\nu \in D(\mathrm{Ent})$ and all n

$$\mathrm{Ent}(H_t^\Upsilon \mu_n) \le \mathrm{Ent}(\nu) + \frac{K}{2(1 - \mathrm{e}^{-2Kt})} W_2(\mu_n, \nu)\;.$$

The left hand side is uniformly bounded in n because $W_2(\mu_n, \mu) \to 0$. Hence, the entropies stay bounded as well. Moreover, by the Wasserstein contractivity of H_t^Υ we have

$$W_2(H_t^\Upsilon \mu_n, H_t^\Upsilon \mu) \le \mathrm{e}^{-2Kt}\; W_2(\mu_n, \mu) \to 0\;,$$

implying the weak convergence of $H_t^\Upsilon \mu_n$ to $H_t^\Upsilon \mu$. By lower semicontinuity of the entropy we can derive in the limit $n \to \infty$:

$$\mathrm{Ent}(H_t^\Upsilon \mu) \le \liminf \mathrm{Ent}(H_t^\Upsilon \mu_n) \le \mathrm{Ent}(\nu) + \frac{K}{2(1 - \mathrm{e}^{-2Kt})} W_2(\mu, \nu) < \infty\;.$$

Finally, from the triangle inequality together with Wasserstein contraction we infer:

$$\begin{aligned} W_2(H_t^\Upsilon \mu, \mu) &\le W_2(H_t^\Upsilon \mu, H_t^\Upsilon \mu_n) + W_2(H_t^\Upsilon \mu_n, \mu_n) + W_2(\mu_n, \mu) \\ &\le (1 + \mathrm{e}^{-2Kt}) W_2(\mu, \mu_n) + W_2(H_t^\Upsilon \mu_n, \mu_n)\;, \end{aligned}$$

The right hand side can be made arbitrarily small by first choosing n so big such that the first term is small uniformly for $t \in [0, 1]$ and then taking t small to make the second term small by the first part of the proof. This proves the last claim of the lemma. □

We will now describe the regularization procedure needed in the sequel. We will use the notion of *regular curve* as introduced in [5, Def. 4.10]. Briefly, we call a curve $(\mu_s)_{s\in[0,1]}$ with $\mu_s = f_s \pi$ regular if the following are satisfied:

- (μ_s) is 2-absolutely continuous in $(\mathscr{P}(\Upsilon), W_2)$,
- $\mathrm{Ent}(\mu_s)$ and $I(H_t^\Upsilon \mu_s)$ are bounded for $s \in [0, 1], t \in [0, T]$ for any $T > 0$,
- $f \in C^1\big([0, 1], L^1(\Upsilon, \pi)\big)$ and $\Delta^{(1)} f \in C\big([0, 1], L^1(\Upsilon, \pi)\big)$,
- $f_s = h^\varepsilon \tilde{f}_s$ for some $\tilde{f}_s \in L^1(\Upsilon, \pi)$ and $\varepsilon > 0$.

Here $\Delta^{(1)}$ denotes the generator of the semigroup T_t^Υ in $L^1(\Upsilon,\pi)$ and h^ε is the mollification of the semigroup introduced in (5.5).

In the sequel we will denote by $\dot{f}_s$ the derivative of $[0,1]\ni s\mapsto f_s\in L^1(\Upsilon,\pi)$. We will need the following result which is an adaption and slight improvement of [5, Prop. 4.11].

Lemma 5.5 (Approximation by regular curves)**.** *Let $(\mu_s)_{s\in[0,1]}$ be a 2-absolutely continuous curve in $\big(\mathscr{P}(\Upsilon),W_2\big)$ such that $\mu_s\in\mathscr{P}_e$ for some (hence any) $s\in[0,1]$. Then there exists a sequence of regular curves (μ_s^n) with the following properties. As $n\to\infty$ we have for any $s\in[0,1]$:*

$$W_2(\mu_s^n,\mu_s)\ \to\ 0\ , \tag{5.7}$$

$$\limsup|\dot{\mu}_s^n|\ \le\ |\dot{\mu}_s|\quad a.e.\ in\ [0,1]\ . \tag{5.8}$$

Moreover, if $\mathrm{Ent}(\mu_0),\mathrm{Ent}(\mu_1)<\infty$ we have:

$$\mathrm{Ent}(\mu_0^n)\to\mathrm{Ent}(\mu_0)\ ,\quad \mathrm{Ent}(\mu_1^n)\to\mathrm{Ent}(\mu_1)\ . \tag{5.9}$$

Proof. Following [5, Prop. 4.11] we employ a threefold regularization procedure. Given n, we construct a curve $(\mu_s^{n,0})_s$ with $s\in[-\frac1n,1+\frac1n]$ by setting

$$\mu_s^{n,0}\ =\ \begin{cases}\mu_0\ , & -\frac1n\le s\le\frac1n\ ,\\ \mu_{(s-\frac1n)/(1-\frac2n)}\ , & \frac1n\le s\le 1-\frac1n\ ,\\ \mu_1\ , & 1-\frac1n\le s\le 1+\frac1n\ .\end{cases}$$

Then, for $s\in[0,1]$ we first define $\mu_s^{n,1}=H_{1/n}\mu_s^{n,0}=f_s^{n,1}\pi$, which is absolutely continuous w.r.t. π by Lemma 5.4. The second step consists in a convolution in the time parameter. We set

$$\mu_s^{n,2}=f_s^{n,2}\pi\ ,\qquad f_s^{n,2}\ =\ \int f_{s-s'}^{n,1}\psi_n(s')\,\mathrm{d}s'\ ,$$

where $\psi_n(s)=n\cdot\psi(ns)$ for some smooth kernel $\psi:\mathbb{R}\to\mathbb{R}_+$ supported in $[-1,1]$ with $\int\psi(s)\,\mathrm{d}s=1$. Finally, we set

$$\mu_s^n\ =\ f_s^n\pi\ ,\qquad f_s^n\ =\ h^{1/n}f_s^{n,2}\ ,$$

where h^ε denotes a mollification of the semigroup given by (5.5). Following the argument in [5, Prop. 4.11] one sees that $(\mu_s^n)_{s\in[0,1]}$ constructed in this way is a regular curve and that (5.7) holds. Note that in our setting the convergence (5.7) relies on Lemma 5.1, the uniform bounds on entropy and Fisher information are ensured by the $L\log L$-regularization (5.4) and the estimate (5.6). (5.8) follows from the convexity properties of W_2^2 and the K-contractivity of the heat flow. To prove (5.9), simply note that for $i=0,1$

$$\mathrm{Ent}(\mu_i^n)=\mathrm{Ent}(h^{1/n}H_{1/n}\mu_i)\le\mathrm{Ent}(\mu_i)$$

since H_t^Υ and hence also h_t decreases the entropy by Lemma 5.2. This together with (5.7) and lower semicontinuity of Ent implies (5.9). ☐

5.2. **Action estimate.** Here we establish the key action estimate, Proposition 5.9, which allows us to derive the Evolution Variational Inequality in the next section. We proceed very closely along the lines of [5, Sec. 4.3] where the corresponding result has been proven in the setting of a Dirichlet form with a finite intrinsic metric. However, a careful inspection of the proofs reveals that the same arguments work almost verbatim in the present context of an extended intrinsic distance. We give a

sketch of the main steps in the argument to make the line of reasoning transparent. We refer to [5, Sec. 4.3] for detailed proofs.
For the following lemmas let $(\mu_s)_{s\in[0,1]}$ be a regular curve and write $\mu_s = f_s\pi$. We set $\mu_{s,t} = H^\Upsilon_{st}\mu_s = f_{s,t}\pi$. Moreover, let $\varphi : \Upsilon \to \mathbb{R}$ be bounded and d_Υ-Lipschitz. We set $\varphi_s = Q_s\varphi$ for $s \in [0, 1]$, where

$$Q_s\varphi(\gamma) \;:=\; \inf_{\sigma\in\Upsilon}\left[\varphi(\sigma) + \frac{d^2_\Upsilon(\gamma,\sigma)}{2s}\right]$$

denotes the Hopf-Lax semigroup as recalled in Section 3.1.
Following [5, Lem. 4.13] we first obtain the following estimate.

Lemma 5.6. *For any $t > 0$ the map $s \mapsto \int \varphi_s \,\mathrm{d}\mu_{s,t}$ is absolutely continuous and we have for a.e $s \in (0, 1)$:*

(5.10)
$$\frac{\mathrm{d}}{\mathrm{d}s}\int \varphi_s \,\mathrm{d}\mu_{s,t} \;=\; \int \dot f_s T^\Upsilon_{st}\varphi_s \,\mathrm{d}\pi - \frac{1}{2}\int \Gamma^\Upsilon(\varphi_s)\,\mathrm{d}\mu_{s,t} - t\int 2\sqrt{f_{s,t}}\,\Gamma^\Upsilon\big(\sqrt{f_{s,t}},\varphi_s\big)\,\mathrm{d}\pi \;.$$

We need to use a regularization E_ε of the entropy functional where the singularities of the logarithm are truncated. Let us define $e_\varepsilon : [0,\infty) \to \mathbb{R}$ by setting $e'_\varepsilon(r) = \log(\varepsilon + r \wedge \varepsilon^{-1}) + 1$ and $e_\varepsilon(0) = 0$. Then for any $\mu = f\pi \in \mathscr{P}(\Upsilon)$ we define

$$E_\varepsilon(\mu) := \int e_\varepsilon(f)\,\mathrm{d}\pi \;.$$

Moreover we set $p_\varepsilon(r) = e'_\varepsilon(r^2) - \log\varepsilon - 1$. Note that for any $\mu \in D(\mathrm{Ent})$ we have $E_\varepsilon(\mu) \to \mathrm{Ent}(\mu)$ as $\varepsilon \to 0$.
Following [5, Lem. 4.15], we obtain an estimate for the derivative of the regularized entropy E_ε along the curve $s \mapsto \mu_{s,t}$.

Lemma 5.7. *For any $t > 0$ we have*

$$E_\varepsilon(\mu_{1,t}) - E_\varepsilon(\mu_{0,t}) \;\le\; \int_0^1 \left[\int T^\Upsilon_{st}(g^\varepsilon_{s,t})\dot f_s \,\mathrm{d}\pi - t\int \Gamma^\Upsilon(g^\varepsilon_{s,t})\,\mathrm{d}\mu_{s,t}\right]\mathrm{d}s \;, \tag{5.11}$$

where we put $g^\varepsilon_{s,r} = p_\varepsilon(\sqrt{f_{s,r}})$.

The following estimate follows parallel to [5, Lem. 4.12] building on Lisini's theorem for extended metric spaces from [22].

Lemma 5.8. *For any curve $(\mu_s)_{s\in[0,1]}$ in $AC^2\big([0,1],(\mathscr{P}(\Upsilon),W_2)\big)$ with $\mu_s = f_s\pi$ and $f \in C^1\big((0,1),L^1(\Upsilon,\pi)\big)$ and any d_Υ-Lipschitz function φ we have*

$$\left|\int \dot f_s\varphi\,\mathrm{d}\pi\right| \;\le\; |\dot\mu_s|\cdot\sqrt{\int \Gamma^\Upsilon(\varphi)f_s\,\mathrm{d}\pi} \;. \tag{5.12}$$

Now we can establish the action estimate by following [5, Thm. 4.16].

Proposition 5.9. *For any regular curve $(\mu_s)_{s\in[0,1]}$ in $\mathscr{P}(\Upsilon)$ and all $t > 0$ we have:*

$$\frac{1}{2}W_2^2(\mu_{1,t},\mu_{0,t}) - \frac{1}{2}\int_0^1 \mathrm{e}^{-2Kst}|\dot\mu_s|^2\,\mathrm{d}s \;\le\; t\left[\mathrm{Ent}(\mu_{0,t}) - \mathrm{Ent}(\mu_{1,t})\right] \;. \tag{5.13}$$

Proof. Fix a function $\varphi : \Upsilon \to \mathbb{R}$ which is bounded and d_Υ-Lipschitz. Applying Lemma 5.6 and Lemma 5.7 we first obtain

$$\begin{aligned}
&\int \varphi_1 \,\mathrm{d}\mu_{1,t} - \int \varphi_0 \,\mathrm{d}\mu_{0,t} + t\left[E_\varepsilon(\mu_{1,t}) - E_\varepsilon(\mu_{0,t})\right] - \int_0^1 \frac{1}{2}\mathrm{e}^{-2Kst}|\dot\mu_s|^2 \,\mathrm{d}s \\
&= \int_0^1 \frac{\mathrm{d}}{\mathrm{d}s}\left[\int \varphi_s \,\mathrm{d}\mu_{s,t} + tE_\varepsilon(\mu_{s,t})\right] - \frac{1}{2}\mathrm{e}^{-2Kst}|\dot\mu_s|^2 \,\mathrm{d}s \\
&\le \int_0^1 \Bigg[\int \dot f_s T^\Upsilon_{st}(\varphi_s + t g^\varepsilon_{s,t}) \,\mathrm{d}\pi - \frac{1}{2}\mathrm{e}^{-2Kst}|\dot\mu_s|^2 \\
&\qquad -t^2\int \Gamma^\Upsilon(g^\varepsilon_{s,t}) \,\mathrm{d}\mu_{s,t} - \frac{1}{2}\int \Gamma^\Upsilon(\varphi_s) \,\mathrm{d}\mu_{s,t} - t\int 2\sqrt{f_{s,t}}\Gamma^\Upsilon\big(\sqrt{f_{s,t}}, \varphi_s\big) \,\mathrm{d}\pi\Bigg] \,\mathrm{d}s \\
&=: A + B\,,
\end{aligned}$$

where A and B denote the sums of the terms in the first and second line respectively. Let us put $q_\varepsilon(r) = \sqrt{r}\big(2 - \sqrt{r}p'_\varepsilon(\sqrt{r})\big)$. Then we have by the chain rule

$$2\sqrt{f_{s,t}}\Gamma^\Upsilon(\sqrt{f_{s,t}}, \varphi_s) \;=\; f_{s,t}\Gamma^\Upsilon(g^\varepsilon_{s,t}, \varphi_s) + q_\varepsilon(f_{s,t})\Gamma^\Upsilon(\sqrt{f_{s,t}}, \varphi_s)\,.$$

Using this and completing the square we obtain

$$B \;\le\; \int_0^1 \left[-\frac{1}{2}\int \Gamma^\Upsilon(\varphi_s + t g^\varepsilon_{s,t}) \,\mathrm{d}\mu_{s,t} - t\int q_\varepsilon(f_{s,t})\Gamma^\Upsilon\big(\sqrt{f_{s,t}}, \varphi_s\big) \,\mathrm{d}\pi\right] \,\mathrm{d}s\,. \tag{5.14}$$

Using (5.12), Young's inequality as well as the gradient estimate (4.10) from Theorem 4.7 we infer that

$$\begin{aligned}
A \;&\le\; \int_0^1 \left[\frac{1}{2}\mathrm{e}^{2Kst}\int \Gamma^\Upsilon\big(T^\Upsilon_{st}(\varphi_s + t g^\varepsilon_{s,t})\big) f_s \,\mathrm{d}\pi\right] \,\mathrm{d}s \\
&\le\; \int_0^1 \left[\frac{1}{2}\int \Gamma^\Upsilon\big(\varphi_s + t g^\varepsilon_{s,t}\big) \,\mathrm{d}\mu_{s,t}\right] \,\mathrm{d}s\,.
\end{aligned}\tag{5.15}$$

Combining (5.15) and (5.14) we obtain that for any $\delta > 0$:

$$\begin{aligned}
A + B \;&\le\; \int_0^1 \left[-t\int q_\varepsilon(f_{s,t})\Gamma^\Upsilon\big(\sqrt{f_{s,t}}, \varphi_s\big) \,\mathrm{d}\pi\right] \,\mathrm{d}s \\
&\le\; t\int_0^1 \int |q_\varepsilon(f_{s,t})|\sqrt{\Gamma^\Upsilon\big(\sqrt{f_{s,t}}\big)\Gamma^\Upsilon\big(\varphi_s\big)} \,\mathrm{d}\pi \,\mathrm{d}s \\
&\le\; \int_0^1 \left[\frac{t\delta}{8}I(\mu_{s,t}) + \frac{t}{2\delta}\int q^2_\varepsilon(f_{s,t})\Gamma^\Upsilon\big(\varphi_s\big) \,\mathrm{d}\pi\right] \,\mathrm{d}s\,,
\end{aligned}$$

where we have used Young's inequality again. Now, using that $q^2_\varepsilon(r) \le r$ and $q_\varepsilon(r) \to 0$ as $\varepsilon \to 0$ we can pass to the limit first as $\varepsilon \to 0$ and then as $\delta \to 0$ to arrive at

$$\int \varphi_1 \,\mathrm{d}\mu_{1,t} - \int \varphi_0 \,\mathrm{d}\mu_{0,t} - \frac{1}{2}\int_0^1 \mathrm{e}^{-2Kst}|\dot\mu_s|^2 \,\mathrm{d}s \;\le\; t\left[\mathrm{Ent}(\mu_{0,t}) - \mathrm{Ent}(\mu_{1,t})\right]\,.$$

Finally, taking the supremum with respect to φ and invoking the Kantorovich duality Cor. 3.3 we get (5.13). □

5.3. **EVI, geodesic convexity and gradient flows.** We can now prove the main result of this section.

Theorem 5.10. *Assume that* $\mathrm{Ric}_M \geq K$. *Then the dual heat semigroup* $(H_t^\Upsilon)_t$ *satisfies the following Evolution Variational Inequality. For all* $\sigma \in D(\mathrm{Ent})$ *and* $\mu \in \mathscr{P}(\Upsilon)$ *with* $W_2(\mu,\sigma) < \infty$*:*

$$\frac{\mathrm{d}^+}{\mathrm{d}t}\frac{1}{2}W_2^2(H_t^\Upsilon\mu,\sigma) + \frac{K}{2}W_2^2(H_t^\Upsilon\mu,\sigma) \;\leq\; \mathrm{Ent}(\sigma) - \mathrm{Ent}(H_t^\Upsilon\mu) \quad \forall t > 0\,. \tag{5.16}$$

Here we denote by

$$\frac{\mathrm{d}^+}{\mathrm{d}t}f(t) \;=\; \limsup_{h\searrow 0}\frac{f(t+h)-f(t)}{h}$$

the upper right derivative.

Proof. By Lemma 5.1 we have that $H_t^\Upsilon\mu$ is well defined, belongs to $D(\mathrm{Ent})$ and $W_2(H_t^\Upsilon\mu,\sigma) < \infty$ for all $t \geq 0$. By the semigroup property it is sufficient to assume $\mu \in D(\mathrm{Ent})$ and prove (5.16) at $t = 0$. Let $(\mu_s)_s$ be a curve in $AC^2\big([0,1],\big(\mathscr{P}(\Upsilon),W_2\big)\big)$ connecting $\mu_0 = \sigma$ to $\mu_1 = \mu$. By Lemma 5.5 we can find approximating regular curves $(\mu_s^n)_s$ and applying Proposition 5.9 to the curves $\mu_{s,t}^n = H_{st}^\Upsilon\mu_s^n$ we find:

$$\frac{1}{2}W_2^2(\mu_{1,t}^n,\mu_{0,t}^n) - \frac{1}{2}\int_0^1 \mathrm{e}^{-2Kst}|\dot\mu_s^n|^2\,\mathrm{d}s \;\leq\; t\left[\mathrm{Ent}(\mu_{0,t}^n) - \mathrm{Ent}(\mu_{1,t}^n)\right]\,.$$

Passing to the limit $n \to \infty$ and using the convergences (5.7), (5.8) and (5.9) as well as lower semicontinuity of Ent we get:

$$\frac{1}{2}W_2^2(H_t^\Upsilon\mu,\sigma) - \frac{1}{2}\int_0^1 \mathrm{e}^{-2Kst}|\dot\mu_s|^2\,\mathrm{d}s \;\leq\; t\left[\mathrm{Ent}(\sigma) - \mathrm{Ent}(H_t^\Upsilon\mu)\right]\,.$$

Minimizing over the curve $(\mu_s)_s$ and using the fact that $(\mathscr{P}(\Upsilon),W_2)$ is a length space we obtain

$$\frac{1}{2}W_2^2(H_t^\Upsilon\mu,\sigma) - \frac{1}{2}\frac{\mathrm{e}^{2Kt}-1}{2Kt}W_2^2(\mu,\sigma) \;\leq\; t\left[\mathrm{Ent}(\sigma) - \mathrm{Ent}(H_t^\Upsilon\mu)\right]\,.$$

Dividing by t and letting $t \searrow 0$ finally yields (5.16). ☐

As a direct consequence we obtain convexity of the entropy along geodesics.

Corollary 5.11. *For all* $\mu_0,\mu_1 \in D(\mathrm{Ent})$ *with* $W_2(\mu_0,\mu_1) < \infty$ *and any geodesic* $(\mu_s)_{s\in[0,1]}$ *connecting them we have for all* $s \in [0,1]$*:*

$$\mathrm{Ent}(\mu_s) \;\leq\; (1-s)\,\mathrm{Ent}(\mu_0) + s\,\mathrm{Ent}(\mu_1) - \frac{K}{2}s(1-s)W_2^2(\mu_0,\mu_1)\,. \tag{5.17}$$

Proof. This follows from the very same argument as in [11, Thm. 3.2]. Since all the distances appearing are finite, the fact that we deal with extended metric spaces does not play a role. To make this clear we give a sketch of the proof.
Multiplying (5.16) with e^{Kt} and integrating from 0 to t yields that for every $\sigma \in D(\mathrm{Ent})$ and $\mu \in \mathscr{P}(\Upsilon)$ with $W_2(\mu,\sigma) < \infty$:

$$\frac{\mathrm{e}^{Kt}}{2}W_2^2(H_t^\Upsilon\mu,\sigma) - \frac{1}{2}W_2^2(\mu,\sigma) \;\leq\; \frac{\mathrm{e}^{Kt}-1}{K}\Big(\mathrm{Ent}(\sigma) - \mathrm{Ent}(H_t^\Upsilon\mu)\Big)\,.$$

Applying this with $\mu = \mu_s$ and $\sigma = \mu_0$ or $\sigma = \mu_1$ respectively and taking a convex combination of the resulting inequalities we get

$$\begin{aligned}
&\frac{e^{Kt}-1}{K}\Big((1-s)\operatorname{Ent}(\mu_0) + s\operatorname{Ent}(\mu_1) - \operatorname{Ent}(H_t^\Upsilon \mu_s)\Big)\\
&\geq \frac{e^{Kt}}{2}\Big((1-s)W_2^2(H_t^\Upsilon\mu_s,\mu_0) + sW_2^2(H_t^\Upsilon\mu_s,\mu_1)\Big)\\
&-\frac{1}{2}\Big((1-s)W_2^2(\mu_s,\mu_0) + sW_2^2(\mu_s,\mu_1)\Big)\\
&\geq \frac{e^{Kt}-1}{2}s(1-s)W_2^2(\mu_0,\mu_1)\ .
\end{aligned}$$

In the last step we have used the elementary inequality

$$(1-s)a^2 + sb^2 \ \geq\ s(1-s)(a+b)^2 \quad \forall a,b>0,\ s\in[0,1]\ ,$$

the triangle inequality and the fact that $(\mu_s)_s$ is a constant speed geodesic. Dividing by $e^{Kt}-1$ and letting $t \searrow 0$ then yields (5.17). □

Remark 5.12. We have obtained that $(\Upsilon, d_\Upsilon, \pi)$ is an extended metric measure space satisfying the $\mathrm{CD}(K,\infty)$ curvature bound in the sense of Lott–Villani and Sturm, see also [6, Def. 9.1] for an extension of the definition to extended metric measure spaces. Moreover, it is a *strong* $\mathrm{CD}(K,\infty)$ space in the sense that convexity holds along all geodesics.

6. Appendix

Proof of Lemma 5.1 Given $\mu \in \mathscr{P}_e$ we will construct a sequence of measures $\mu_n \in D(\operatorname{Ent})$ such that $W_2(\mu_n,\mu) \to 0$ as $n\to\infty$. So let us fix $\nu\in D(\operatorname{Ent})$ with $W_2(\mu,\nu)<\infty$. The strategy of the proof is to find a big bounded set in the base space M in which most of the transport happens. In this set we can approximate the measure μ nicely. Outside this set we will keep the ν-points to end up with a measure in the support of the entropy.

Construction of μ_n:
Fix $x_0 \in M$ and $n\in\mathbb{N}$ and set $B = B(x_0,n)$. Choose an optimal coupling $q\in\mathsf{Opt}(\mu,\nu)$. By Lemma 2.6 we can choose for each $(\gamma,\omega)\in\operatorname{supp}(q)$ an optimal matching $\eta\in\mathsf{Opt}(\gamma,\omega)$. By Lemma 6.1, the map $(\gamma,\omega)\mapsto\eta$ can be chosen measurable. Denoting by proj_i the projection onto the i-th component, define a map $(\gamma,\omega)\mapsto\xi\in\Upsilon$ via

$$\xi \ :=\ \mathrm{proj}_1(\mathbf{1}_{B\times B}\eta)\cup\mathrm{proj}_2(\mathbf{1}_{\complement B\times\complement B}\eta)\cup\mathrm{proj}_2(\mathbf{1}_{\complement B\times B}\eta)\cup\mathrm{proj}_2(\mathbf{1}_{B\times\complement B}\eta) \tag{6.1}$$

For a Borel set $V\subset M$ we define the restriction map $r_V:\Upsilon\to\Upsilon$ by $r_V(\gamma)=\gamma|_V$. We will often use the short hand notation $r_V(\gamma)=\gamma_V$. By construction we have

$$\xi_{\complement B} \ =\ \omega_{\complement B}\ , \tag{6.2}$$

$$\xi(B) \ =\ \omega(B)\ . \tag{6.3}$$

Let us set $\alpha := 1/(2\sqrt{n\xi(B)})$. For $x\in\xi\cap B$ put

$$\chi(x) = \begin{cases} x\ , & \text{if } d(x,\complement B)>\alpha\ ,\\ x_\alpha\ , & \text{otherwise}\ \ ,\end{cases}$$

where x_α is the point on the geodesic between x and x_0 satisfying $d(x,x_\alpha)=\alpha$. Clearly $B(\chi(x),\alpha)\subset B$. Denote the uniform distribution on $B(x,\alpha)$ by $U_{x,\alpha}$.

Given $(\gamma,\omega) \in \mathrm{supp}(q)$ we define a probability measure $\mathcal{U}^n_{\gamma,\omega} \in \mathscr{P}(\Upsilon)$ as follows. Let $\xi(\gamma,\omega)$ be defined as in (6.1) and write $\xi_B = \sum_{i=1}^k \delta_{x_i}$ and $\xi_{\complement B} = \sum_{i=k+1}^\infty \delta_{x_i}$. Given $(y_1,\dots,y_k) \in M$ we put

$$\tilde{\xi}(y_1,\dots,y_k) \ = \ \sum_{i=1}^k \delta_{y_i} + \sum_{i=k+1}^\infty \delta_{x_i} \in \Upsilon \ .$$

Then we define

$$\mathcal{U}^n_{\gamma,\omega} \ := \ \int \Pi_{i=1}^k \delta_{\tilde{\xi}(y_1,\dots,y_k)} U_{\chi(x_i),\alpha}(\,\mathrm{d}y_i) \ .$$

Note that the map $T : (\gamma,\omega) \mapsto \mathcal{U}^n_{\gamma,\omega}$ is measurable. We finally define

$$\mu_n := \int T(\gamma,\omega)\ q(d\gamma,d\omega) \ \in \mathscr{P}(\Upsilon) \ .$$

The proof of Lemma 5.1 will be finished once we have established the following claims.

Claim 1. $W_2(\mu,\mu_n) \to 0$ as $n \to \infty$.

Claim 2. For all n we have $\mathrm{Ent}(\mu_n) < \infty$.

Proof of Claim 1. Define for $\gamma,\omega \in \Upsilon$

$$c_n(\gamma,\omega) := \inf \left\{ \int_{B_n \times B_n} d^2(x,y)\ \eta(dx,dy),\ \eta \in \mathsf{Opt}(\gamma,\omega) \right\} \ .$$

By the same reasoning as for Lemma 2.6 there is a matching realizing the infimum. By the compactness of $\mathsf{Opt}(\gamma,\omega)$ we have the pointwise convergence $c_n(\gamma,\omega) \nearrow c(\gamma,\omega) = d^2_\Upsilon(\gamma,\omega)$. For $q \in \mathsf{Opt}(\mu,\nu)$ we have

$$\int c_n\ dq \nearrow \int c\ dq \ .$$

For $\epsilon > 0$ choose n large enough such that

$$\int c\ dq - \epsilon \le \int c_n\ dq \ .$$

By construction we have for any $(\gamma,\omega) \in \mathrm{supp}(q)$ and $\xi = \xi(\gamma,\omega)$ as defined in (6.1)

$$W_2^2(\delta_\xi,\mathcal{U}^n_{\gamma,\omega}) \ \le \ 4\xi(B)\alpha^2 \ = \ \frac{1}{n} \ , \tag{6.4}$$

$$W_2^2(\delta_\gamma,\delta_\xi) \ = \ d^2_\Upsilon(\gamma,\xi) \ \le \ d^2_\Upsilon(\gamma,\omega) - c_n(\gamma,\omega) \ . \tag{6.5}$$

Consider the coupling $Q := (\mathrm{proj}_1, T)_* q$ between μ and μ_n. Using (6.4) and (6.5) and the convexity of W_2^2 we can deduce

$$\begin{aligned} W_2^2(\mu,\mu_n) \ &\le \ \int W_2^2(\delta_\gamma,\mathcal{U}^n_{\gamma,\omega})\,\mathrm{d}q(\gamma,\omega) \\ &\le \ \frac{2}{n} + 2\int c(\gamma,\omega) - c_n(\gamma,\omega)\,\mathrm{d}q(\gamma,\omega) \ \le \ \frac{2}{n} + 2\epsilon \ , \end{aligned}$$

which finishes the proof. ∎

Proof of Claim 2. Note that (6.2) implies that $(r_{\complement B})_*\mu_n = (r_{\complement B})_*\nu =: \nu_{\complement B}$. Therefore, we can disintegrate μ_n with respect to $\nu_{\complement B}$ and get

$$\mu_n(\,\mathrm{d}\omega) \;=\; (\mu_n)_{\omega_{\complement B}}(\,\mathrm{d}\omega_B)\nu_{\complement B}(\,\mathrm{d}\omega_{\complement B})\;.$$

Denote by $(q_\omega)_\omega$ the disintegration of q with respect to ν and by $\nu_{B,\omega_{\complement B}}$ the disintegration of ν with respect to $\nu_{\complement B}$. Then, we have

$$(\mu_n)_{\omega_{\complement B}}(\,\mathrm{d}\omega_B) \;=\; \int T\big(\gamma,(\omega_B,\omega_{\complement B})\big)\; q_{\omega_B,\omega_{\complement B}}(\,\mathrm{d}\gamma)\;\nu_{B,\omega_{\complement B}}(\,\mathrm{d}\omega_B)\;. \tag{6.6}$$

By disintegration we have

$$\mathrm{Ent}(\mu_n|\pi) = \int \mathrm{Ent}((\mu_n)_{\gamma_{\complement B}}|\pi_B)\;\nu_{\complement B}(\,\mathrm{d}\gamma_{\complement B}) + \mathrm{Ent}(\nu_{\complement B}|\pi_{\complement B})\;, \tag{6.7}$$

where $\pi_B = (r_B)_*\pi$. By monotonicity of the entropy under push forward, it holds that $\mathrm{Ent}(\nu_{\complement B}|\pi_{\complement B}) \le \mathrm{Ent}(\nu|\pi) < \infty$. Thus it remains to show that the first term is finite. We will derive an estimate on $\mathrm{Ent}((\mu_n)_{\gamma_{\complement B}}|\pi_B)$ which is integrable w.r.t. $\nu_{\complement B}$ yielding the result.

We fix $\gamma_{\complement B} = \omega_{\complement B}$ and write – for notational convenience – $(\mu_n)_{\omega_{\complement B}} = \theta$. The configuration space over the set B will be denoted by Υ_B. It can be decomposed into $\bigcup_{k\ge0}\Upsilon_B^{(k)}$, where $\Upsilon_B^{(k)} = \{\gamma\in\Upsilon_B : \gamma(B) = k\}$. Note that for all $\rho = f\pi_B \in \mathscr{P}(\Upsilon_B)$ we have

$$\mathrm{Ent}(\rho|\pi_B) \;=\; \sum_{k\ge0}\rho_k\Big[\int_{\Upsilon^{(k)}} f_k\log f_k\;\mathrm{d}\pi_{B,k} + \log\frac{\rho_k}{\pi_k}\Big]\;, \tag{6.8}$$

where for each k we have set $\pi_k = \pi_B(\Upsilon_B^{(k)})$, $\pi_{B,k} = \pi_k^{-1}(\pi_B)_{\llcorner\Upsilon_B^{(k)}}$, i.e. the normalized restriction of π_B to the set $\Upsilon_B^{(k)}$. Moreover, we set $\rho_k = \rho(\Upsilon_B^{(k)})$ and $\rho_k^{-1}\rho = \pi_k^{-1}f_k\;\pi_B$ on $\Upsilon_B^{(k)}$. By (6.3), we have that $\theta(\gamma:\gamma(B)=k) = \nu_{B,\omega_{\complement B}}(\gamma:\gamma(B)=k)$ for all k, i.e. $\theta_k = (\nu_{B,\omega_{\complement B}})_k$. Since $\nu\in D(\mathrm{Ent})$, the formulas (6.8) and (6.7) imply that

$$\sum_k\theta_k\log\frac{\theta_k}{\pi_k} \;\le\; \mathrm{Ent}(\nu_{B,\omega_{\complement B}}|\pi_B) \;\in\; L^1(\nu_{\complement B})\;.$$

By (6.8), we therefore need to find a good estimate on $\mathrm{Ent}(\theta_k^{-1}\theta|\pi_{B,k})$ for all k. Put $A_k := T^{-1}(\Upsilon^{(k)}\cup\omega_{\complement B})$. By Jensen's inequality and (6.6) we have

$$\begin{aligned}&\mathrm{Ent}(\theta_k^{-1}\theta|\pi_{B,k})\\ &\quad\le \int_{A_k} 1/\theta_k\;\;\mathrm{Ent}((r_B)_*T(\gamma,(\omega_B,\omega_{\complement B}))|\pi_{B,k})\;q_{\omega_B,\omega_{\complement B}}(d\gamma)\nu_{B,\omega_{\complement B}}(d\omega_B)\;.\end{aligned}$$

Hence, we need to estimate the entropy of $(r_B)_*T(\gamma,(\omega_B,\omega_{\complement B}))$ which is a random k-point configuration, where each point of the configuration is uniformly distributed on a ball of radius $\alpha = 1/(2\sqrt{n\xi(B)})$ independently of the others. Putting $\tilde m = m_{\llcorner B}/m(B)$ and $U_i = U_{\chi(x_i),\alpha}$ for $\xi(\gamma,\omega)\cap B = \sum_{i=1}^k\delta_{x_i}$ we get using $m(B(x,r)) \ge \kappa r^N$ uniformly in $x\in B$ and $r\in[0,1/2]$ for some constants κ and N

$$\begin{aligned}\mathrm{Ent}((r_B)_*T(\gamma,(\omega_B,\omega_{\complement B}))|\pi_{B,k}) \;&=\; \mathrm{Ent}(\Pi_{i=1}^k U_i|\tilde m^{\otimes k})\\ &=\; \sum\mathrm{Ent}(U_i|\tilde m) \le Ck(\log k + \log n),\end{aligned}$$

for some constant C depending only on B. Putting everything together we get

$$\begin{aligned}\mathrm{Ent}(\theta|\pi_B) &\leq C\sum_{k\geq 0}\pi_k k(\log k+\log n)+\sum_{k\geq 0}\theta_k\log\frac{\theta_k}{\pi_k}\\ &\leq C'+\mathrm{Ent}(\nu_{B,\omega_{\complement B}}|\pi_B)\ ,\end{aligned}$$

which is in $L^1(\nu_{\complement B})$ by (6.7) and the assumption that $\nu\in D(\mathrm{Ent})$. This finishes the proof. □

Lemma 6.1. *Let $\mu,\nu\in\mathscr{P}(\Upsilon)$ with $W_2(\mu,\nu)<\infty$ and $q\in\mathsf{Opt}(\mu,\nu)$. Then there is a measurable selection $S:\mathrm{supp}(q)\to\Upsilon_{M^2}$ of optimal matchings.*

Proof. Take $(\gamma,\omega)\in\mathrm{supp}(q)$. Any matching of γ and ω can be identified with an element of the configuration space over M^2, denoted by Υ_{M^2}. Note that the map assigning to η its marginals $p_1(\eta)$ and $p_2(\eta)$ is measurable w.r.t. the vague topologies on Υ_{M^2} and Υ_M. Moreover, by Lemma 4.1 (i) and (vi) of [29] the mappings

$$G:\Upsilon_{M^2}\to[0,\infty]\quad \eta\mapsto\int d^2(x,y)\,\mathrm{d}\eta(x,y)$$

and

$$\tilde F:\Upsilon_M\times\Upsilon_M\to[0,\infty]\quad(\gamma,\omega)\mapsto d_\Upsilon(\gamma,\omega)$$

are lower semicontinuous. Hence, the function $F=\tilde F\circ(p_1(\cdot),p_2(\cdot))$ is measurable w.r.t. the vague topology on Υ_{M^2}. (Note that we always have $F(\eta)\leq G(\eta)$.) Then the set

$$L=\{(\gamma,\omega,\eta):(p_1(\eta),p_2(\eta))=(\gamma,\omega),\ F(\eta)=G(\eta)\}$$

is Borel measurable. Moreover, as the set of optimal matchings of (γ,ω) is closed (even compact) we can use the selection theorem by Kuratowski and Ryll-Nardzewski (e.g. [30, Thm. 5.2.1]) to get the desired map. □

References

[1] S. Albeverio, Y. G. Kondratiev, and M. Röckner. Analysis and geometry on configuration spaces. *J. Funct. Anal.*, 154(2):444–500, 1998.

[2] S. Albeverio, Y. G. Kondratiev, and M. Röckner. Analysis and geometry on configuration spaces: the Gibbsian case. *J. Funct. Anal.*, 157(1):242–291, 1998.

[3] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, second edition, 2008.

[4] L. Ambrosio, N. Gigli, and G. Savaré. Metric measure spaces with Riemannian Ricci curvature bounded from below. *Preprint at arXiv:1109.0222*, 2011.

[5] L. Ambrosio, N. Gigli, and G. Savaré. Bakry-Émery curvature-dimension condition and Riemannian Ricci curvature bounds. *Preprint at arXiv:1209.5786*, 2012.

[6] L. Ambrosio, N. Gigli, and G. 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.

[7] D. Bakry and M. Émery. Diffusions hypercontractives. In *Séminaire de probabilités, XIX, 1983/84*, volume 1123 of *Lecture Notes in Math.*, pages 177–206. Springer, Berlin, 1985.

[8] D. Burago, Y. Burago, and S. Ivanov. *A course in metric geometry*, volume 33. American Mathematical Society Providence, 2001.

[9] O. Chodosh. A lack of Ricci bounds for the entropic measure on Wasserstein space over the interval. *Journal of Functional Analysis*, 262(10):4570 – 4581, 2012.

[10] D. Cordero-Erausquin, R. J. McCann, and M. Schmuckenschläger. A Riemannian interpolation inequality à la Borell, Brascamp and Lieb. *Invent. Math.*, 146(2):219–257, 2001.

[11] S. Daneri and G. Savaré. Eulerian calculus for the displacement convexity in the Wasserstein distance. *SIAM Journal on Mathematical Analysis*, 40(3):1104–1122, 2008.

[12] L Decreusefond. Wasserstein distance on configuration space. *Potential Analysis*, 28(3):283–300, 2008.
[13] C.-S. Deng. Harnack inequality on configuration spaces: The coupling approach and a unified treatment. *Stochastic Processes and their Applications*, 124(1):220 – 234, 2014.
[14] M. Erbar, K. Kuwada, and K.-Th. Sturm. On the equivalence of the entropic curvature-dimension condition and Bochner's inequality on metric measure spaces. *arXiv preprint arXiv:1303.4382*, 2013.
[15] R. Jordan, D. Kinderlehrer, and F. Otto. The variational formulation of the Fokker–Planck equation. *SIAM journal on mathematical analysis*, 29(1):1–17, 1998.
[16] O. Kallenberg. *Foundations of modern probability.* springer, 2002.
[17] H. G. Kellerer. Duality theorems for marginal problems. *Z. Wahrsch. Verw. Gebiete*, 67(4):399–432, 1984.
[18] D. G. Kendall. On infinite doubly-stochastic matrices and Birkhoff's problem 111. *J. London Math. Soc.*, 35:81–84, 1960.
[19] Yu. G. Kondratiev, E. Lytvynov, and M. Röckner. Non-equilibrium stochastic dynamics in continuum: The free case. *Condensed Matter Physics*, 11(4):701–721, 2008.
[20] J. LaFontaine, M. Katz, M. Gromov, S.M. Bates, P. Pansu, and S. Semmes. *Metric structures for Riemannian and non-Riemannian spaces.* Springer, 2007.
[21] N. Lebedeva and A. Petrunin. Curvature bounded below: a definition a la Berg-Nikolaev. *Electron. Res. Announc. Math. Sci.*, 17:122–124, 2010.
[22] S. Lisini. Absolutely continuous curves in extended Wasserstein–Orlicz spaces. *arXiv preprint arXiv:1402.7328*, 2014.
[23] J. Lott and C. Villani. Ricci curvature for metric-measure spaces via optimal transport. *Ann. Math. (2)*, 169(3):903–991, 2009.
[24] A. Naber. Characterizations of bounded Ricci curvature on smooth and nonsmooth spaces. *arXiv preprint arXiv:1306.6512*, 2013.
[25] H. Osada. Infinite-dimensional stochastic differential equations related to random matrices. *Probab. Theory Related Fields*, 153(3-4):471–509, 2012.
[26] H. Osada. Interacting Brownian motions in infinite dimensions with logarithmic interaction potentials. *Ann. Probab.*, 41(1):1–49, 2013.
[27] F. Otto and M. Westdickenberg. Eulerian calculus for the contraction in the Wasserstein distance. *SIAM J. Math. Anal.*, 37(4):1227–1255, 2005.
[28] N. Privault. Connections and curvature in the Riemannian geometry of configuration spaces. *Journal of Functional Analysis*, 185(2):367 – 403, 2001.
[29] M. Röckner and A. Schied. Rademacher's theorem on configuration spaces and applications. *Journal of Functional Analysis*, 169(2):325–356, 1999.
[30] Sashi Mohan Srivastava. *A course on Borel sets*, volume 180. Springer, 1998.
[31] D. W. Stroock. *An introduction to the analysis of paths on a Riemannian manifold*, volume 74 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2000.
[32] K.-Th. Sturm. Metric spaces of lower bounded curvature. *Exposition. Math.*, 17(1):35–47, 1999.
[33] K.T. Sturm. On the geometry of metric measure spaces.I. *Acta Math.*, 196(1):65–131, 2006.
[34] M.-K. von Renesse and K.-Th. Sturm. Transport inequalities, gradient estimates, entropy, and Ricci curvature. *Comm. Pure Appl. Math.*, 58(7):923–940, 2005.

University of Bonn, Institute for Applied Mathematics, Endenicher Allee 60, 53115 Bonn, Germany
E-mail address: erbar@iam.uni-bonn.de
E-mail address: huesmann@iam.uni-bonn.de