# On the Geometry of Metric Measure Spaces 

Karl-Theodor Sturm

Preprint No. 203, 2004


## Sonderforschungsbereich 611

Singuläre Phänomene und Skalierung in mathematischen Modellen

# On the Geometry of Metric Measure Spaces 

Karl-Theodor Sturm


#### Abstract

We introduce and analyze lower ('Ricci') curvature bounds $\mathbb{C u r v}(M, \mathrm{~d}, m) \geq K$ for metric measure spaces ( $M, \mathrm{~d}, m$ ). Our definition is based on convexity properties of the relative entropy $\operatorname{Ent}(. \mid m)$ regarded as a function on the $L_{2}$-Wasserstein space of probability measures on the metric space ( $M, \mathrm{~d}$ ). Among others, we show that $\mathbb{C u r v}(M, \mathrm{~d}, m) \geq K$ implies estimates for the volume growth of concentric balls. For Riemannian manifolds, $\mathbb{C u r v}(M, \mathrm{~d}, m) \geq K$ if and only if $\operatorname{Ric}_{M}(\xi, \xi) \geq K \cdot|\xi|^{2}$ for all $\xi \in T M$. The crucial point is that our lower curvature bounds are stable under an appropriate notion of $\mathbb{D}$-convergence of metric measure spaces. We define a complete and separable metric $\mathbb{D}$ on the family of all isomorphism classes of normalized metric measure spaces. The metric $\mathbb{D}$ has a natural interpretation, based on the concept of optimal mass transportation. We also prove that the family of normalized metric measure spaces with doubling constant $\leq C$ is closed under $\mathbb{D}$-convergence. Moreover, the family of normalized metric measure spaces with doubling constant $\leq C$ and radius $\leq R$ is compact under $\mathbb{D}$-convergence.


## 1 Introduction

The notion of a 'metric space' is one of the basic concepts of mathematics. Metric spaces play a prominent role in many fields of mathematics. In particular, they constitute natural generalizations of manifolds admitting all kinds of singularities and still providing rich geometric structures.
A. D. Alexandrov [A151] introduced the notion of lower curvature bounds for metric spaces in terms of comparison properties for geodesic triangles. These curvature bounds are equivalent to lower bounds for the sectional curvature in the case where the metric spaces are Riemannian manifolds, - and they may be regarded as generalized lower bounds for the 'sectional curvature' for general metric spaces. A fundamental observation is that these lower bounds are stable under an appropriate notion of convergence of metric spaces, the so-called Gromov-Hausdorff convergence, introduced by M. Gromov [Gr81a]. The family of manifolds with sectional curvature $\geq K$ is, of course, not closed under Gromov-Hausdorff convergence but the family of metric spaces with curvature $\geq K$ in the sense of Alexandrov is closed (for each $K \in \mathbb{R}$ ). Even more, the family of compact metric spaces with curvature $\geq K$, Hausdorff dimension $\leq N$ and diameter $\leq \Delta$ is compact (for any choice of $K, N, \Delta$ ).

For many fundamental results in geometric analysis, however, the crucial ingredients are not bounds for the sectional curvature but bounds for the Ricci curvature: estimates for heat kernels and Green functions, Harnack inequalities, eigenvalue estimates, isoperimetric inequalities, Sobolev inequalities, - they all depend on lower bounds for the Ricci curvature of the underlying manifolds as pointed out by S. T. Yau and others, see e.g. [LY86], [Ch93], [Da89], [SC02]. The family of Riemannian manifolds with given lower bound for the Ricci curvature is not closed under Gromov-Hausdorff convergence (nor it is closed under any other reasonable notion of convergence). One of the great challenges thus is to establish a generalized notion of lower Ricci curvature bounds for singular spaces. For detailed investigations and a survey of the state of
the art for this problem, we refer to the contributions by J. Cheeger and T. Colding [CC97/00]. Fascinating new developments have been outlined just recently by G. Perelman [Pe02] in the context of his work on the Poincaré conjecture.

Generalizations of lower Ricci curvature bounds should be formulated in the framework of metric measure spaces. These are triples $(M, \mathrm{~d}, m)$ where $(M, \mathrm{~d})$ is a metric space and $m$ is a measure on the Borel $\sigma$-algebra of $M$. We will always require that the metric space ( $M, \mathrm{~d}$ ) is complete and separable and that the measure $m$ is locally finite. Recall that for generalizations of sectional curvature bounds only the metric structure ( $M, \mathrm{~d}$ ) is required whereas for generalizations of Ricci curvature bounds in addition a reference measure $m$ has to be specified. In a certain sense, this phenomenon is well-known from the discussion of the curvature-dimension condition of D. Bakry and M. Emery [BE85] in the framework of Dirichlet forms and symmetric Markov semigroups. Of course, also the Bakry-Emery condition is a kind of generalized lower bound for the Ricci curvature (together with an upper bound for the dimension). However, it is not given in terms of the basic data ( $M, \mathrm{~d}, m$ ) but in terms of the Dirichlet form (or heat semigroup) derived from the original quantities in a highly non-trivial manner.

Metric measure spaces have been studied quite intensively in recent years. Of particular interest is the study of functional inequalities, like Sobolev and Poincaré inequalities, on metric measure spaces and the construction and investigation of function spaces of various types [HK95, HK00], [Ko00], [He01]. To some extent, doubling properties for the volume and scale invariant Poincaré inequalities on metric balls can be regarded as weak replacements of lower Ricci curvature bounds. Among others, they allow to construct Dirichlet forms, Laplacians and heat kernels on given metric measure spaces and to derive (elliptic and parabolic) Harnack inequalities as well as (upper and lower) Gaussian estimates for heat kernels [St98], [Ch99]. On the other hand, however, even in simplest cases doubling constant and Poincaré constant do not characterize spaces with lower bounded Ricci curvature: they always allow at least also metrics which are equivalent to the given ones.

The main results of this paper are:

- We define a complete and separable metric $\mathbb{D}$ on the family of all isomorphism classes of normalized metric measure spaces, Theorem 3.6. The metric $\mathbb{D}$ has a natural interpretation, based on the concept of optimal mass transportation.
- The family of normalized metric measure spaces with doubling constant $\leq C$ is closed under $\mathbb{D}$-convergence, Theorem 3.15. Moreover, the family of normalized metric measure spaces with doubling constant $\leq C$ and radius $\leq R$ is compact under $\mathbb{D}$-convergence (for any choice of real numbers $C, R$ ), Theorem 3.16.
- We introduce a notion of lower curvature bounds $\mathbb{C} \mathbb{C u r v}(M, \mathrm{~d}, m)$ for metric measure spaces ( $M, \mathrm{~d}, m$ ), based on convexity properties of the relative entropy $\operatorname{Ent}(. \mid m)$ w.r.t. the reference measure $m$. Here $\nu \mapsto \operatorname{Ent}(\nu \mid m)$ is regarded as a function on the $L_{2}$-Wasserstein space of probability measures on the metric space ( $M, \mathrm{~d}$ ). For Riemannian manifolds, $\underline{\operatorname{Curv}}(M, \mathrm{~d}, m) \geq K$ if and only if $\operatorname{Ric}_{M}(\xi, \xi) \geq K \cdot|\xi|^{2}$ for all $\xi \in T M$, Theorem 4.9.
- Local lower curvature bounds imply global lower curvature bounds, Theorem 4.17.
- Lower curvature bounds are stable under $\mathbb{D}$-convergence, Theorem 4.20.
- Lower curvature bounds of the form $\mathbb{C u r v}(M, \mathrm{~d}, m) \geq K$ imply estimates for the volume growth of concentric balls, for instance, if $K \leq 0$

$$
m\left(B_{r}(x)\right) \leq C(x) \cdot \exp \left(-K r^{2} / 2\right)
$$

for all $r \geq 1$, Theorem 4.24.
The concept of optimal mass transportation plays a crucial role in our approach. It originates in the classical transportation problems of G. Monge [M1781] and L. V. Kantorovich [Ka42]. The basic quantity for us is the so-called $L_{2}$-Wasserstein distance between two probability measures $\mu$ and $\nu$ on a given complete separable metric space ( $M, \mathrm{~d}$ ) defined as

$$
\mathrm{d}_{W}(\mu, \nu):=\inf _{q}\left(\int_{M \times M} \mathrm{~d}^{2}(x, y) d q(x, y)\right)^{1 / 2}
$$

where the infimum is taken over all couplings $q$ of $\mu$ and $\nu$. The latter are probability measures on the product space $M \times M$ whose marginals (i.e. image measures under the projections) are the given measures $\mu$ and $\nu$. One choice, of course, is $q=\mu \otimes \nu$ but in most cases this will be a very bad choice if one aims for minimal transportation costs. The $L_{2}$-Wasserstein distance can be interpreted as the minimal transportation costs (measured in $L_{2}$-sense) for transporting goods from producers at locations distributed according to $\mu$ to consumers at locations distributed according to $\nu$.

Two results may be regarded as milestones in the recent development of theory and application of mass transportation concepts; these results have raised an increasing interest in this topic of people from various fields of mathematics including pde's, geometry, fluid mechanics, and probability. See e.g. [Ta95], [Le01], [BG99], [OV00], [DD02], [CMS01], [AT04], [FÜ04a] and in particular the monograph by C. Villani [Vi03] which gives an excellent survey on the whole field. The first of these two results is the polar factorization of Y. Brenier [Br91] and its extension to the Riemannian setting by R. McCann [Mc95, Mc97]. The second one is F. Otto's [JKO98, Ot01] formal Riemannian calculus on the space $\mathcal{P}_{2}(M)$ of probability measures on $M$, equipped with the $L_{2}$-Wasserstein metric, and his interpretation of the heat equation (and of other nonlinear dissipative evolution equations) as gradient flow(s) of the relative entropy

$$
\operatorname{Ent}(\nu \mid m)=\int_{M} \frac{d \nu}{d m} \log \left(\frac{d \nu}{d m}\right) d m
$$

(or related functionals, resp.) on $\mathcal{P}_{2}(M)$.
It turned out that convexity properties of the function $\nu \mapsto \operatorname{Ent}(\nu \mid m)$ are intimately related to curvature properties of the underlying metric measure space ( $M, \mathrm{~d}, m$ ). If $M$ is a complete Riemannian manifold with Riemannian distance d and if $m=e^{-V} d x$ then M.-K. von Renesse and the author proved (see [RS04] for the case $V=0$ and $[\mathrm{St04}]$ for the general case) that the function $\nu \mapsto \operatorname{Ent}(\nu \mid m)$ is $K$-convex ${ }^{1}$ on $\mathcal{P}_{2}(M)$ if and only if

$$
\operatorname{Ric}_{M}(\xi, \xi)+\operatorname{Hess} V(\xi, \xi) \geq K \cdot|\xi|^{2}
$$

for all $\xi \in T M$. An heuristic argument for the 'if'-implication of this equivalence was presented in [OV00], based on the formal Riemannian calculus on $\mathcal{P}_{2}(M)$. In the particular case $K=0, V=0$ the 'if'-implication was proven in [CMS01].

[^0]Having in mind these results, it seems quite natural to say that an arbitrary metric measure space ( $M, \mathrm{~d}, m$ ) has curvature $\geq K$ if and only if for any pair $\nu_{0}, \nu_{1} \in \mathcal{P}_{2}(M)$ with $\operatorname{Ent}\left(\nu_{0} \mid m\right)<\infty$, $\operatorname{Ent}\left(\nu_{1} \mid m\right)<\infty$ there exists a geodesic $\Gamma:[0,1] \rightarrow \mathcal{P}_{2}(M)$ connecting $\nu_{0}$ and $\nu_{1}$ with

$$
\begin{equation*}
\operatorname{Ent}(\Gamma(t) \mid m) \leq(1-t) \operatorname{Ent}(\Gamma(0) \mid m)+t \operatorname{Ent}(\Gamma(1) \mid m)-\frac{K}{2} t(1-t) \mathrm{d}_{W}^{2}(\Gamma(0), \Gamma(1)) \tag{1.1}
\end{equation*}
$$

for all $t \in[0,1]$. In this case, we also briefly write $\mathbb{C u r v}(M, \mathrm{~d}, m) \geq K$.
A crucial property of this kind of curvature bound is its stability under convergence of metric measure spaces. Of course, this requires to have an appropriate notion of topology or distance on the family of all metric measure spaces. We define the distance between two normalized metric measure spaces by

$$
\mathbb{D}\left((M, \mathrm{~d}, m),\left(M^{\prime}, \mathrm{d}^{\prime}, m^{\prime}\right)\right):=\inf _{\mathrm{q}, \hat{\mathrm{~d}}}\left(\int_{M \times M^{\prime}} \hat{\mathrm{d}}^{2}(x, y) d q(x, y)\right)^{1 / 2}
$$

where the infimum is taken over all couplings $q$ of $m$ and $m^{\prime}$ and over all couplings $\hat{d}$ of $d$ and $\mathrm{d}^{\prime}$. The former are probability measures on $M \times M^{\prime}$ with marginals $m$ and $m^{\prime}$. The latter are pseudo metrics on the disjoint union $M \sqcup M^{\prime}$ which extend d and $\mathrm{d}^{\prime}$.
Also the distance $\mathbb{D}$ has an interpretation in terms of mass transportation: In order to realize the distance

$$
\mathbb{D}\left((M, \mathrm{~d}, m),\left(M^{\prime}, \mathrm{d}^{\prime}, m^{\prime}\right)\right)
$$

between two normalized metric measure spaces $(M, \mathrm{~d}, m)$ and ( $M^{\prime}, \mathrm{d}^{\prime}, m^{\prime}$ ) one first may use isometric transformations of ( $M, \mathrm{~d}$ ) and ( $M^{\prime}, \mathrm{d}^{\prime}$ ) to bring the images of $m$ and $m^{\prime}$ in optimal position to each other. (In the sense of transportation costs, these transformations are for free). Then one has to solve the usual mass transportation problem, trying to minimize the transportation costs in the $L_{2}$-sense.

It turns out that $\mathbb{D}$ is a complete separable metric on the family $\mathbb{X}_{1}$ of all isomorphism classes of normalized metric measure spaces. The family of normalized metric measure spaces with curvature $\geq K$ is closed under $\mathbb{D}$-convergence. Moreover, the family of normalized metric measure spaces with doubling constant $\leq C$ is closed under $\mathbb{D}$-convergence and the family of normalized metric measure spaces with doubling constant $\leq C$ and radius $\leq R$ is compact under $\mathbb{D}$-convergence (for any choice of real numbers $K, C, R$ ).
For various other distances on the family $\mathbb{X}_{1}$, see the additional chapter $3 \frac{1}{2}$ in [Gr99]. A completely different notion of distance between Riemannian manifolds was proposed by A. Kasue [KK94, Ka04], based on the short time asymptotics of the heat kernel. Yet another convergence concept was proposed by K. Kuwae and T. Shioya [KS03] extending the concept of $\Gamma$-convergence and Mosco convergence towards a notion of convergence of operators (or Dirichlet forms or heat semigroups) on varying spaces.

A major advantage of our distance $\mathbb{D}$ seems to be that it has a very natural geometric interpretation, namely, in terms of the above-mentioned mass transportation concept. We also expect that it is closely related to more analytic properties of metric measure spaces. Following [JKO98], the heat semigroup on a metric measure space $(M, \mathrm{~d}, m)$ should be obtained as the gradient flow on $\mathcal{P}_{2}(M)$ for the relative entropy $\operatorname{Ent}(. \mid m)$. Curvature bounds of the form $\mathbb{C u r v}(M, \mathrm{~d}, m) \geq K$ should e.g. imply $K$-contractivity of the heat flow

$$
\mathrm{d}_{W}\left(\mu p_{t}, \nu p_{t}\right) \leq e^{-K t} \mathbf{d}_{W}(\mu, \nu),
$$

gradient estimates for harmonic functions, isoperimetric inequalities, and volume growth estimates.
The basic concepts and main results of this paper have been presented in [St04a]. In a forthcoming paper [St05], we will study metric measure spaces satisfying a so-called curvature-dimension condition $(K, N)$ being more restrictive than the condition $\mathbb{C u r v}(M, \mathrm{~d}, m) \geq K$. The additional parameter $N$ plays the role of an upper bound for the dimension. Among others, this will lead to more precise volume growth estimates in the spirit of the Bishop-Gromov volume comparison theorem.

Here in this paper, we will proceed as follows:
In Chapter 2 we give a brief survey on the geometry of metric spaces, recalling the concepts of length and geodesic spaces, the Gromov-Hausdorff distance and the lower curvature bounds in the sense of Alexandrov. We introduce the $L_{2}$-Wasserstein space of probability measures on a given metric space and derive some of the basic properties.
Chapter 3 is devoted to the metric $\mathbb{D}$. The first main result states that it indeed defines a (complete and separable) metric on the family of isomorphism classes of normalized metric measure spaces. We collect several simple examples of $\mathbb{D}$-convergence with increasing and decreasing dimensions and we discuss closedness and compactness properties of the families of normalized metric measure spaces with the doubling property.
In Chapter 4 we study metric measure spaces with curvature bounds. First we introduce and discuss the relative entropy, then we present the definition of curvature bounds and analyze their behavior under various transformations (isomorphisms, scaling, weights, subsets, products). The main results are the Globalization Theorem and the Convergence Theorem. Finally, we deduce growth estimates for the volume of concentric balls.

## 2 On the Geometry of Metric Spaces

### 2.1 Length and Geodesic Spaces

Let us summarize some definitions and basic results on the geometry of metric spaces. For proofs and further details we refer to [BH99], [Gr99], and [BBI01].
Throughout this paper, a pseudo metric on a set $M$ will be a function $\mathrm{d}: M \times M \rightarrow[0, \infty]$ which is symmetric, vanishes on the diagonal and satisfies the triangle inequality. If it does not vanish outside the diagonal and does not take the value $+\infty$ then it is called metric. From now on, let ( $M, \mathrm{~d}$ ) be a metric space. Open balls in $M$ will be denoted by $B_{r}(x)=\{y \in M: \mathrm{d}(x, y)<r\}$, their closures by $\bar{B}_{r}(x) \subset\{y \in M: \mathrm{d}(x, y) \leq r\}$. A curve connecting two points $x, y \in M$ is a continuous map $\gamma:[a, b] \rightarrow M$ with $\gamma(a)=x$ and $\gamma(b)=y$. Obviously, then Length $(\gamma) \geq \mathrm{d}(x, y)$ with the length of $\gamma$ being defined as

$$
\operatorname{Length}(\gamma)=\sup \sum_{k=1}^{n} \mathrm{~d}\left(\gamma\left(t_{k-1}\right), \gamma\left(t_{k}\right)\right)
$$

where the supremum is taken over all partitions $a=t_{0}<t_{1}<\cdots<t_{n}=b$. If Length $(\gamma)<\infty$ then $\gamma$ is called rectifiable. In this case we can and will henceforth always assume that (after suitable reparametrization) $\gamma$ has constant speed, i.e. Length $\left(\left.\gamma\right|_{[s, t]}\right)=\frac{t-s}{b-a} \cdot \operatorname{Length}(\gamma)$ for all $a<s<t<b$. In general, we will not distinguish between curves and equivalence classes of curves which are reparametrizations of each other. The curve $\gamma:[a, b] \rightarrow M$ is called geodesic iff Length $(\gamma)=\mathrm{d}(\gamma(a), \gamma(b))$. A geodesic in this sense is always minimizing.

A metric space ( $M, \mathrm{~d}$ ) is called length space (or length metric space) iff for all $x, y \in M$

$$
\mathrm{d}(x, y)=\inf _{\gamma} \operatorname{Length}(\gamma)
$$

where the infimum is taken over all curves $\gamma$ in $M$ which connect $x$ and $y$. A metric space ( $M, \mathrm{~d}$ ) is called geodesic space (or geodesic metric space) iff each pair of points $x, y \in M$ is connected by a geodesic. (This geodesic is not required to be unique.)

Lemma 2.1. A complete metric space ( $M$, d) is a length space (or geodesic space) if and only if for each pair of points $x_{0}, x_{1} \in M$ and for each $\epsilon>0$ (or for $\epsilon=0$, resp.) there exists a point $y \in M$ satisfying for each $i=0,1$

$$
\begin{equation*}
\mathrm{d}\left(x_{i}, y\right) \leq \frac{1}{2} \mathrm{~d}\left(x_{0}, x_{1}\right)+\epsilon . \tag{2.1}
\end{equation*}
$$

Any such point $y$ will be called $\epsilon$-midpoint of $x_{0}$ and $x_{1}$. In the case $\epsilon=0$ it will be called midpoint of $x_{0}$ and $x_{1}$.

Remark 2.2. Given $x_{0}, x_{1} \in M$ then each $\epsilon$-midpoint $y \in M$ satisfies

$$
\begin{equation*}
\mathrm{d}^{2}\left(x_{0}, y\right)+\mathrm{d}^{2}\left(y, x_{1}\right) \leq \frac{1}{2} \mathrm{~d}^{2}\left(x_{0}, x_{1}\right)+\epsilon^{\prime} \tag{2.2}
\end{equation*}
$$

with $\epsilon^{\prime}=2 \epsilon \mathrm{~d}\left(x_{0}, x_{1}\right)+2 \epsilon^{2}$. Vice versa, each $y \in M$ which satisfies (2.2) is an $\epsilon$-midpoint with $\epsilon=\sqrt{\left(\mathrm{d}\left(x_{0}, x_{1}\right) / 2\right)^{2}-\epsilon^{\prime} / 2}-\mathrm{d}\left(x_{0}, x_{1}\right) / 2-\sqrt{\epsilon^{\prime} / 2}$.
Indeed, if $y$ is an $\epsilon$-midpoint then

$$
\mathrm{d}^{2}\left(x_{0}, y\right)+\mathrm{d}^{2}\left(y, x_{1}\right) \leq 2\left[\frac{1}{2} \mathrm{~d}\left(x_{0}, x_{1}\right)+\epsilon\right]^{2}=\frac{1}{2} \mathrm{~d}^{2}\left(x_{0}, x_{1}\right)+\epsilon^{\prime}
$$

with $\epsilon^{\prime}$ chosen as above. Conversely, if $y$ satisfies (2.2) then

$$
\begin{aligned}
\frac{1}{2} \mathrm{~d}^{2}\left(x_{0}, x_{1}\right)+\epsilon^{\prime} & \geq \mathrm{d}^{2}\left(x_{0}, y\right)+\mathrm{d}^{2}\left(y, x_{1}\right) \\
& =\frac{1}{2}\left[\mathrm{~d}\left(x_{0}, y\right)+\mathrm{d}\left(y, x_{1}\right)\right]^{2}+\frac{1}{2}\left[\mathrm{~d}\left(x_{0}, y\right)-\mathrm{d}\left(y, x_{1}\right)\right]^{2} \\
& \geq \frac{1}{2} \mathrm{~d}^{2}\left(x_{0}, x_{1}\right)+\frac{1}{2}\left[\mathrm{~d}\left(x_{0}, y\right)-\mathrm{d}\left(y, x_{1}\right)\right]^{2}
\end{aligned}
$$

Hence, $\left|\mathrm{d}\left(x_{0}, y\right)-\mathrm{d}\left(y, x_{1}\right)\right| \leq \sqrt{2 \epsilon^{\prime}}$ and

$$
2 \mathrm{~d}\left(x_{i}, y\right)-\sqrt{2 \epsilon^{\prime}} \leq \mathrm{d}\left(x_{0}, y\right)+\mathrm{d}\left(y, x_{1}\right) \leq \sqrt{\mathrm{d}^{2}\left(x_{0}, x_{1}\right)+2 \epsilon^{\prime}}
$$

for $i=0,1$.
Lemma 2.3. If ( $M, \mathrm{~d}$ ) is a complete length space then:
(i) The closure of $B_{r}(x)$ is $\{y \in M: \mathrm{d}(x, y) \leq r\}$.
(ii) $M$ is locally compact if and only if each closed ball in $M$ is compact.
(iii) If $M$ is locally compact then it is a geodesic space.

Recall that the Hausdorff distance between two subsets $A_{1}, A_{2}$ of a metric space ( $M, \mathrm{~d}$ ) is given by

$$
\mathrm{d}^{\mathrm{H}}\left(A_{1}, A_{2}\right)=\inf \left\{\epsilon>0: A_{1} \subset B_{\epsilon}\left(A_{2}\right), A_{2} \subset B_{\epsilon}\left(A_{1}\right)\right\}
$$

where $B_{\epsilon}(A):=\left\{x \in M: \inf _{y \in A} \mathrm{~d}(x, y)<\epsilon\right\}$ denotes the $\epsilon$-neighborhood of $A \subset M$. The Gromov-Hausdorff distance between two metric spaces $\left(M_{1}, \mathrm{~d}_{1}\right)$ and $\left(M_{2}, \mathrm{~d}_{2}\right)$ is defined by

$$
\mathrm{D}^{\mathrm{GH}}\left(\left(M_{1}, \mathrm{~d}_{1}\right),\left(M_{2}, \mathrm{~d}_{2}\right)\right)=\inf \quad \mathrm{d}^{\mathrm{H}}\left(j_{1}\left(M_{1}\right), j_{2}\left(M_{2}\right)\right)
$$

where the inf is taken over all metric spaces ( $M, \mathrm{~d}$ ) and over all isometric embeddings $j_{1}: M_{1} \hookrightarrow$ $M, j_{2}: M_{2} \hookrightarrow M$.

## Proposition 2.4.

(i) The Gromov-Hausdorff metric $\mathrm{D}^{G H}$ is a pseudo metric on the family X of isometry classes of metric spaces.
(ii) Let $\mathrm{X}_{l}$ (and $\mathrm{X}_{l c l}$ ) denote the family of isometry classes of (locally compact)complete length spaces. Then $\mathrm{X}_{l}$ and $\mathrm{X}_{\text {lcl }}$ are closed under $\mathrm{D}^{G \mathrm{H}}$-convergence.
(iii) Let $\mathrm{X}_{c}\left(\right.$ and $\left.\mathrm{X}_{f}\right)$ denote the family of isometry classes of compact (or finite, resp.) metric spaces. Then $\left(\mathrm{X}_{c}, \mathrm{D}^{6 H}\right)$ is a complete separable metric space. The family $\mathrm{X}_{f}$ is dense in $\mathrm{X}_{\mathrm{c}}$.

### 2.2 Alexandrov Spaces

Now let us briefly discuss metric spaces with lower curvature bounds in the sense of A.D. Alexandrov [A151]. The latter are generalizations of lower bounds for the sectional curvature for Riemannian manifolds. The results of this section will not be used in the sequel. The focus in this paper is on generalizations of lower bounds for the Ricci curvature. Partly, however, there will be some analogy to Alexandrov's generalizations of lower bounds for the sectional curvature. We summarize some of the basic properties of these metric spaces and refer to [BGP92], [GP97], [Gr99], [BBI01], [Pl02] for further details.

Given any $K \in \mathbb{R}$ we say that a metric space ( $M, \mathrm{~d}$ ) has curvature $\geq K$ iff for each quadruple of points $z, x_{1}, x_{2}, x_{3} \in M$

$$
\begin{equation*}
\varangle_{K}\left(z ; x_{1}, x_{2}\right)+\varangle_{K}\left(z ; x_{2}, x_{3}\right)+\varangle_{K}\left(z ; x_{3}, x_{1}\right) \leq 2 \pi . \tag{2.3}
\end{equation*}
$$

Here for any triple of points $z, x, y \in M$ with $K \cdot[\mathrm{~d}(z, x)+\mathrm{d}(x, y)+\mathrm{d}(y, z)]^{2}<(2 \pi)^{2}$ we denote by $\varangle_{K}(z ; x, y)$ the angle at $\bar{z}$ of a triangle $\Delta(\bar{z}, \bar{x}, \bar{y})$ with side lengths $\bar{z} \bar{x}=\mathrm{d}(z, x), \bar{z} \bar{y}=$ $\mathrm{d}(z, y), \bar{x} \bar{y}=\mathrm{d}(x, y)$ in the simply connected two-dimensional space of constant curvature $K$, i.e.

$$
\begin{equation*}
\varangle_{K}(z ; x, y)=\arccos \left(\frac{\cos (\sqrt{K} \mathrm{~d}(x, y))-\cos (\sqrt{K} \mathrm{~d}(z, x)) \cdot \cos (\sqrt{K} \mathrm{~d}(z, y))}{\sin (\sqrt{K} \mathrm{~d}(z, x)) \cdot \sin (\sqrt{K} \mathrm{~d}(z, y))}\right) \tag{2.4}
\end{equation*}
$$

(with appropriate interpretations/modifications if $K<0$ or $K=0$ ). If $K \cdot[\mathrm{~d}(z, x)+\mathrm{d}(x, y)+$ $\mathrm{d}(y, z)]^{2} \geq(2 \pi)^{2}$ we put $\varangle_{K}(z ; x, y):=-\infty$. We define

$$
\underline{\operatorname{curv}}(M, \mathrm{~d})=\sup \{K \in \mathbb{R}:(M, \mathrm{~d}) \text { has curvature } \geq K\} .
$$

Complete length spaces with curvature $\geq K$ and finite Hausdorff dimension are called Alexandrov spaces with curvature $\geq K$. For complete geodesic spaces there are several alternative (but
equivalent) ways to define this curvature bound: via triangle comparison, angle monotonicity, convexity properties of the distance. For instance, one can interpret it as a weak formulation of

$$
\begin{equation*}
\operatorname{Hess} \frac{1}{K} \cos (\sqrt{K} \mathrm{~d}(z, \cdot)) \geq-\cos (\sqrt{K} \mathrm{~d}(z, \cdot)) \tag{2.5}
\end{equation*}
$$

for all $z \in M$ (with appropriate modification in the case $K \leq 0$, e.g. $\operatorname{Hess}^{2}(z, \cdot) / 2 \leq 1$ if $K=0$ ) 。

Example 2.5. Let $M$ be a complete Riemannian manifold with Riemannian distance $d$ and dimension $n \geq 2$. Then curv $(M, \mathrm{~d})$ is the greatest lower bound for the sectional curvature of $M$.

Example 2.6. If $M$ is an interval $\subset \mathbb{R}$ or if $M$ is a circle of length $L$ then curv $(M, \mathrm{~d})=+\infty$. This example is regarded as pathological.

Similarly, one can define metric spaces of curvature $\leq K$ and a number $\overline{\operatorname{curv}}(M, \mathrm{~d}$ ) (which coincides with the least upper bound for the sectional curvature if $M$ is a Riemannian manifold). However, in this paper we concentrate on lower curvature bounds.

Proposition 2.7. For each complete length space ( $M, \mathrm{~d}$ ) the following properties hold:
(i) Scaling: $\underline{\text { curv }}(M, \alpha \mathrm{~d})=\alpha^{-2} \underline{\operatorname{curv}}(M, \mathrm{~d})$ for all $\alpha \in \mathbb{R}_{+}$.
(ii) Products: If $(M, \mathrm{~d})=\bigotimes_{i=1}^{n}\left(M_{i}, \mathrm{~d}_{i}\right)$ with complete length spaces $\left(M_{1}, \mathrm{~d}_{1}\right), \ldots,\left(M_{n}, \mathrm{~d}_{n}\right)$ consisting of more than one point then

$$
\underline{\operatorname{curv}}(M, \mathrm{~d})=\inf \left\{\underline{\operatorname{curv}}\left(M_{1}, \mathrm{~d}_{1}\right), \ldots, \underline{\operatorname{curv}}\left(M_{n}, \mathrm{~d}_{n}\right), 0\right\} .
$$

(iii) Local/global: If $M=\bigcup_{i \in I} M_{i}$ with open subsets $M_{i} \subset M$ then

$$
\underline{\operatorname{curv}}(M, \mathrm{~d})=\inf _{i \in I} \underline{\operatorname{curv}}\left(M_{i}, \mathrm{~d}_{i}\right)
$$

('Topogonov's globalization theorem').
(iv) Convergence: If $\left(\left(M_{n}, \mathrm{~d}_{n}\right)\right)_{n \in \mathbb{N}}$ is a sequence of complete length spaces with GromovHausdorff distance $\mathrm{D}^{\mathrm{GH}}\left(\left(M_{n}, \mathrm{~d}_{n}\right),(M, \mathrm{~d})\right) \rightarrow 0$ as $n \rightarrow \infty$ then

$$
\underline{\operatorname{curv}}(M, \mathrm{~d}) \geq \limsup _{n \rightarrow \infty} \underline{\operatorname{curv}}\left(M_{n}, \mathrm{~d}_{n}\right)
$$

In particular, for each $K \in \mathbb{R}$ the set

$$
\mathrm{X}_{c}(K)=\left\{(M, \mathrm{~d}) \in \mathrm{X}_{c}: \underline{\operatorname{curv}}(M, \mathrm{~d}) \geq K\right\}
$$

is a closed subset of $\left(\mathrm{X}_{c}, \mathrm{D}^{\mathrm{GH}}\right)$.
(v) Compactness: For each $K \in \mathbb{R}, N \in \mathbb{N}$ and $\Delta \in \mathbb{R}_{+}$the set $\mathrm{X}_{c}(K, N, \Delta)$ of all compact length spaces $(M, \mathrm{~d})$ with curvature $\geq K$, Hausdorff dimension $\leq N$ and diameter $\leq \Delta$ is compact w.r.t. $\mathrm{D}^{\mathrm{GH}}$ ('Gromov's compactness theorem').

Definition 2.8. A geodesic space ( $M, \mathrm{~d}$ ) is called nonbranching iff for each quadruple of points $z, x_{0}, x_{1}, x_{2}$ with $z$ being the midpoint of $x_{0}$ and $x_{1}$ as well as the midpoint of $x_{0}$ and $x_{2}$ it follows that $x_{1}=x_{2}$.

Remark 2.9. If a geodesic space has curvature $\geq K$ for some $K \in \mathbb{R}$ then it is nonbranching.

### 2.3 The $L_{2}$-Wasserstein Space

Probability measures on metric spaces will play an important role throughout this paper. We collect some definitions and the basic facts on the $L_{2}$-Wasserstein distance. For further reading we recommend [Du89], [KR57], [RR98], [Vi03] and [Wa69].
For the rest of this chapter, let ( $M, \mathrm{~d}$ ) be a complete separable metric space. A measure $\nu$ on $M$ will always mean a measure on $(M, \mathcal{B}(M))$ with $\mathcal{B}(M)$ being the Borel $\sigma$-algebra of $M$ (generated by the open balls in $M)$. Recall that $\operatorname{supp}[\nu]$, the support of $\nu$, is the smallest closed set $M_{0} \subset M$ such that $\nu\left(M \backslash M_{0}\right)=0$. The push forward of $\nu$ under a measurable map $f: M \rightarrow M^{\prime}$ into another metric space $M^{\prime}$ is the probability measure $f_{*} \nu$ on $M^{\prime}$ given by

$$
\left(f_{*} \nu\right)(A):=\nu\left(f^{-1}(A)\right)
$$

for all measurable $A \subset M^{\prime}$. Given two measures $\mu, \nu$ on $M$ we say that a measure $q$ on $M \times M$ is a coupling of $\mu$ and $\nu$ iff its marginals are $\mu$ and $\nu$, that is, iff

$$
q(A \times M)=\mu(A), \quad q(M \times A)=\nu(A)
$$

for all measurable sets $A \subset M$. (This in particular implies that the total masses coincide: $\nu(M)=q(M \times M)=\mu(M)$.) If $\mu$ and $\nu$ are probability measures then for instance one such coupling is the product measure $\mu \times \nu$.
The $L_{2}$-Wasserstein distance between $\mu$ and $\nu$ is defined as

$$
\begin{equation*}
\mathrm{d}_{W}(\mu, \nu)=\inf \left\{\left(\int_{M \times M} \mathrm{~d}^{2}(x, y) d q(x, y)\right)^{1 / 2}: q \text { is a coupling of } \mu \text { and } \nu\right\} . \tag{2.6}
\end{equation*}
$$

Note that $\mathrm{d}_{W}(\mu, \nu)=+\infty$ whenever $\mu(M) \neq \nu(M)$. We denote by $\mathcal{P}_{2}(M, \mathrm{~d})$ or briefly $\mathcal{P}_{2}(M)$ the space of all probability measures $\nu$ on $M$ with finite second moments:

$$
\int_{M} \mathrm{~d}^{2}(o, x) d \nu(x)<\infty
$$

for some (hence all) $o \in M$. The pair $\left(\mathcal{P}_{2}(M), \mathrm{d}_{W}\right)$ is called $L_{2}$-Wasserstein space over $(M, \mathrm{~d})$.
Proposition 2.10.
(i) $\left(\mathcal{P}_{2}(M), \mathrm{d}_{W}\right)$ is a complete separable metric space.

The map $x \mapsto \delta_{x}$ defines an isometric and totally geodesic embedding of ( $M, \mathrm{~d}$ ) into $\left(\mathcal{P}_{2}(M), \mathrm{d}_{W}\right)$.

The set of all normalized configurations $\mu=\frac{1}{n} \sum_{i=1}^{n} \delta_{x_{i}}$ with $n \in \mathbb{N}$ and $x_{1}, \ldots, x_{n} \in M$ is dense in $\mathcal{P}_{2}(M)$.
(ii) $\mathrm{d}_{W}$-convergence implies weak convergence (in the sense of measures). More precisely. if $\left(\mu_{n}\right)_{n \in \mathbb{N}}$ is a sequence in $\mathcal{P}_{2}(M)$ then $\mathrm{d}_{W}\left(\mu_{n}, \mu\right) \rightarrow 0$ if and only if $\mu_{n} \rightarrow \mu$ weakly and

$$
\begin{equation*}
\lim _{R \rightarrow \infty} \sup _{n} \int_{M \backslash B_{R}(o)} \mathrm{d}^{2}(o, x) d \mu_{n}(x)=0 \tag{2.7}
\end{equation*}
$$

for some (hence each) point $o \in M$. Note that obviously (2.7) is always satisfied if ( $M, \mathrm{~d}$ ) is bounded.
(iii) $\left(\mathcal{P}_{2}(M), \mathrm{d}_{W}\right)$ is a compact space or a length space if and only if $(M, \mathrm{~d})$ is so.
(iv) If $M$ is a length space with more than one point then

$$
\underline{\operatorname{curv}}\left(\mathcal{P}_{2}(M), \mathrm{d}_{W}\right)=0 \quad \Longleftrightarrow \quad \underline{\operatorname{curv}}(M, \mathrm{~d}) \geq 0
$$

and

$$
\underline{\operatorname{curv}}\left(\mathcal{P}_{2}(M), \mathrm{d}_{W}\right)=-\infty \quad \Longleftrightarrow \quad \underline{\operatorname{curv}}(M, \mathrm{~d})<0
$$

Proof. (i), (ii): [RR98], [Vi03].
$(\text { iii })_{a}$ The 'only if' statements follow from the fact that $M$ is isometrically embedded in $\mathcal{P}_{2}(M)$. $(i i i)_{b}$ Compactness of $M$ implies compactness of $\mathcal{P}_{2}(M)$ according to (ii) and Prohorov's theorem.
$(\text { iii })_{c}$ Assume that ( $M, \mathrm{~d}$ ) is a length space and let $\epsilon>0$ and $\mu, \nu \in \mathcal{P}_{2}(M)$ be given. We have to prove that there exists an $\epsilon$-midpoint $\eta$ of $\mu$ and $\nu$. Choose $n \in \mathbb{N}$ and $x_{1}, \ldots, x_{n}, y_{1}, \ldots, y_{n} \in M$ such that $\mathrm{d}_{W}(\mu, \bar{\mu}) \leq \epsilon / 3, \mathrm{~d}_{W}(\nu, \bar{\nu}) \leq \epsilon / 3$ and $\mathrm{d}_{W}^{2}(\bar{\mu}, \bar{\nu})=\frac{1}{n} \sum_{i=1}^{n} \mathrm{~d}^{2}\left(x_{i}, y_{i}\right)$ where $\bar{\mu}:=$ $\frac{1}{n} \sum_{i=1}^{n} \delta_{x_{i}}, \bar{\nu}:=\frac{1}{n} \sum_{i=1}^{n} \delta_{y_{i}}$ For each $i=1, \ldots, n$ let $z_{i}$ be an $\epsilon / 3$-midpoint of $x_{i}$ and $y_{i}$ and put $\eta:=\frac{1}{n} \sum_{i=1}^{n} \delta_{z_{i}}$. Then

$$
\begin{aligned}
\mathrm{d}_{W}(\bar{\mu}, \eta) & \leq\left(\frac{1}{n} \sum_{i=1}^{n} \mathrm{~d}^{2}\left(x_{i}, z_{i}\right)\right)^{1 / 2} \leq\left(\frac{1}{n} \sum_{i=1}^{n}\left[\frac{1}{2} \mathrm{~d}\left(x_{i}, y_{i}\right)+\frac{\epsilon}{3}\right]^{2}\right)^{1 / 2} \\
& \leq \frac{1}{2}\left(\frac{1}{n} \sum_{i=1}^{n} \mathrm{~d}^{2}\left(x_{i}, y_{i}\right)\right)^{1 / 2}+\frac{\epsilon}{3} \leq \frac{1}{2} \mathrm{~d}_{W}(\bar{\mu}, \bar{\nu})+\frac{\epsilon}{3}
\end{aligned}
$$

and thus $\mathrm{d}_{W}(\mu, \eta) \leq \frac{1}{2} \mathrm{~d}_{W}(\mu, \nu)+\epsilon$. Similarly, $\mathrm{d}_{W}(\nu, \eta) \leq \frac{1}{2} \mathrm{~d}_{W}(\mu, \nu)+\epsilon$. This proves the claim.
$(i v)_{a}$ Assume that $(M, \mathrm{~d})$ has curvature $\geq 0$. Then for each $n \in \mathbb{N}$ the space $M^{n}=M \times \cdots \times M$ has curvature $\geq 0$ (Proposition 2.7 (ii)). According to [St99], the latter is equivalent to

$$
\begin{equation*}
\sum_{i, j=1}^{l} \lambda_{i} \lambda_{j} \mathrm{~d}^{2}\left(y_{i}, y_{j}\right) \leq 2 \sum_{i=1}^{l} \lambda_{i} \mathrm{~d}^{2}\left(y_{i}, y_{0}\right) \tag{2.8}
\end{equation*}
$$

for all $l \in \mathbb{N}$, all $\lambda_{1}, \ldots, \lambda_{l} \in \mathbb{R}_{+}$with $\sum_{i=1}^{l} \lambda_{i}=1$ and all $y_{0}, y_{1}, \ldots, y_{l} \in M^{n}$. In order to prove that $\left(\mathcal{P}_{2}(M), \mathrm{d}_{W}\right)$ has curvature $\geq 0$, let $l \in \mathbb{N}, \lambda_{1}, \ldots, \lambda_{l} \in \mathbb{R}_{+}$and $\nu_{0}, \nu_{1}, \ldots, \nu_{l} \in \mathcal{P}_{2}(M)$ be given. For $\epsilon>0$ choose $n \in \mathbb{N}$ and $y_{0}=\left(y_{01}, \ldots, y_{0 n}\right), \ldots, y_{l}=\left(y_{l 1}, \ldots, y_{l n}\right) \in M^{n}$ such that $\mathrm{d}_{W}\left(\nu_{i}, \bar{\nu}_{i}\right) \leq \epsilon$ for all $i=0,1, \ldots, l$ and $\mathrm{d}_{W}^{2}\left(\bar{\nu}_{i}, \bar{\nu}_{0}\right)=\frac{1}{n} \sum_{k=1}^{n} \mathrm{~d}^{2}\left(y_{i k}, y_{0 k}\right)=\frac{1}{n} \mathrm{~d}^{2}\left(y_{i}, y_{0}\right)$ for all $i=1, \ldots, l$ where we put

$$
\bar{\nu}_{i}=\frac{1}{n} \sum_{k=1}^{n} \delta_{y_{i k}}
$$

Then $\mathrm{d}_{W}^{2}\left(\bar{\nu}_{i}, \bar{\nu}_{j}\right) \leq \frac{1}{n} \sum_{k=1}^{n} \mathrm{~d}^{2}\left(y_{i k}, y_{j k}\right)=\frac{1}{n} \mathrm{~d}^{2}\left(y_{i}, y_{j}\right)$ for all $i, j=1, \ldots, l$ and thus by $(2.8)$

$$
\begin{aligned}
& \sum_{i, j=1}^{l} \lambda_{i} \lambda_{j} \mathrm{~d}_{W}^{2}\left(\bar{\nu}_{i}, \bar{\nu}_{j}\right) \leq \frac{1}{n} \sum_{i, j=1}^{l} \lambda_{i} \lambda_{j} \mathrm{~d}^{2}\left(y_{i}, y_{j}\right) \\
& \leq \frac{2}{n} \sum_{i=1}^{l} \lambda_{i} \mathrm{~d}^{2}\left(y_{i}, y_{0}\right)=2 \sum_{i=1}^{l} \lambda_{i} \mathrm{~d}_{W}^{2}\left(\bar{\nu}_{i}, \bar{\nu}_{0}\right)
\end{aligned}
$$

In the limit $\epsilon \rightarrow 0$ this yields

$$
\sum_{i, j=1}^{l} \lambda_{i} \lambda_{j} \mathrm{~d}_{W}^{2}\left(\nu_{i}, \nu_{j}\right) \leq 2 \sum_{i=1}^{l} \lambda_{i} \mathrm{~d}_{W}^{2}\left(\nu_{i}, \nu_{0}\right)
$$

which (again by [St99]) proves the claim.
$(i v)_{b}$ Since ( $M, \mathrm{~d}$ ) is isometrically and totally geodesically embedded into $\left(\mathcal{P}_{2}(M), \mathrm{d}_{W}\right)$ it is obvious that $\underline{\operatorname{curv}}\left(\mathcal{P}_{2}(M), \mathrm{d}_{W}\right) \leq \underline{\operatorname{curv}}(M, \mathrm{~d})$.
$(i v)_{c}$ Assume that curv $(M, \mathrm{~d})<0$. Choose $K<0 \operatorname{such}$ that curv $(M, \mathrm{~d})<K$. Then there exist points $x_{0}, x_{1}, x_{2}, x_{3} \in M$ with

$$
\varangle_{K}\left(x_{0} ; x_{1}, x_{2}\right)+\varangle_{K}\left(x_{0} ; x_{2}, x_{3}\right)+\varangle_{K}\left(x_{0} ; x_{3}, x_{1}\right)>2 \pi .
$$

Choose a point $z \in M$ 'far away' from the $x_{i}$, say $\mathrm{d}\left(z, x_{i}\right) \geq 3 \mathrm{~d}\left(x_{i}, x_{j}\right)$ for all $i, j=0,1,2,3$. [This is always possible since $M$ is a length space and the $x_{i}$ for $i=1,2,3$ can be replaced by points $x_{i}^{\prime}$ lying arbitrarily close to $x_{0}$ on approximate geodesics connecting $x_{0}$ and $x_{i}$.]
For $t \in] 0,1]$ and $i=0,1,2,3$ define $\mu_{i}:=t \delta_{x_{i}}+(1-t) \delta_{z}$. Then

$$
\mathrm{d}_{W}^{2}\left(\mu_{i}, \mu_{j}\right)=t \mathrm{~d}^{2}\left(x_{i}, x_{j}\right)
$$

for all $i, j=0,1,2,3$ and thus according to formula (2.4)

$$
\varangle_{K}\left(x_{0} ; x_{i}, x_{j}\right)=\varangle_{K / t}\left(\mu_{0} ; \mu_{i}, \mu_{j}\right) .
$$

Therefore,

$$
\varangle_{K / t}\left(\mu_{0} ; \mu_{1}, \mu_{2}\right)+\varangle_{K / t}\left(\mu_{0} ; \mu_{2}, \mu_{3}\right)+\varangle_{K / t}\left(\mu_{0} ; \mu_{3}, \mu_{1}\right)>2 \pi
$$

which implies

$$
\underline{\operatorname{curv}}\left(\mathcal{P}(M), \mathrm{d}_{W}\right)<K / t .
$$

Since the latter holds for the chosen $K<0$ and all arbitrarily small $t>0$, it proves the claim. $(i v)_{d}$ Finally, it remains to prove that $\operatorname{curv}\left(\mathcal{P}_{2}(M), \mathrm{d}_{W}\right) \leq 0$ if $M$ has more than one point. Assume that $\operatorname{curv}\left(\mathcal{P}_{2}(M), \mathrm{d}_{W}\right) \geq K$ for some $K>0$ and that $x_{0}, x_{1} \in M$ with $x_{0} \neq x_{1}$. Given $\epsilon>0$ let $y$ be an $\epsilon$-midpoint of $x_{0}$ and $x_{1}$. Put $\nu_{0}=\delta_{x_{0}}, \nu_{1}=\delta_{x_{1}}, \mu=\delta_{y}$ and $\eta=\frac{1}{2} \delta_{x_{0}}+\frac{1}{2} \delta_{x_{1}}$. Then $\mathrm{d}_{W}\left(\nu_{0}, \nu_{1}\right)=\mathrm{d}\left(x_{0}, x_{1}\right), \mathrm{d}_{W}(\eta, \mu) \geq \frac{1}{2} \mathrm{~d}\left(x_{0}, x_{1}\right), \mathrm{d}_{W}\left(\nu_{i}, \eta\right)=\frac{1}{\sqrt{2}} \mathrm{~d}\left(x_{0}, x_{1}\right)$ and $\mathrm{d}_{W}\left(\nu_{i}, \mu\right) \leq$ $\frac{1}{2} \mathrm{~d}\left(x_{0}, x_{1}\right)+\epsilon$ for $i=0$, 1 . In particular, $\mu$ is an $\epsilon$-midpoint of $\nu_{0}$ and $\nu_{1}$.
Our curvature assumption on $\mathcal{P}_{2}(M)$ implies (via quadruple comparison for ( $\mu ; \nu_{0}, \nu_{1}, \eta$ ) or via triangle comparison for $\left.\left(\nu_{0}, \nu_{1}, \eta\right)\right)$ that

$$
\begin{aligned}
& 2 \cos \left(\frac{\sqrt{K}}{2} \mathrm{~d}_{W}\left(\nu_{0}, \nu_{1}\right)\right) \cdot \cos \left(\sqrt{K} \mathrm{~d}_{W}(\eta, \mu)\right) \\
& \quad \leq \cos \left(\sqrt{K} \mathrm{~d}_{W}\left(\eta, \nu_{0}\right)\right)+\cos \left(\sqrt{K} \mathrm{~d}_{W}\left(\eta, \nu_{1}\right)\right)+\epsilon^{\prime}
\end{aligned}
$$

with some $\epsilon^{\prime} \rightarrow 0$ as $\epsilon \rightarrow 0$. Therefore,

$$
\begin{aligned}
& 2 \cos \left(\frac{\sqrt{K}}{2} \mathrm{~d}\left(x_{0}, x_{1}\right)\right) \cdot \cos \left(\frac{\sqrt{K}}{2} \mathrm{~d}\left(x_{0}, x_{1}\right)\right) \\
& \quad \leq \cos \left(\sqrt{\frac{K}{2}} \mathrm{~d}\left(x_{0}, x_{1}\right)\right)+\cos \left(\sqrt{\frac{K}{2}} \mathrm{~d}\left(x_{0}, x_{1}\right)\right) .
\end{aligned}
$$

Now choosing $x_{0}, x_{1} \in M$ with sufficiently small $\mathrm{d}\left(x_{0}, x_{1}\right)$ leads to a contradiction.
Let us recall that a Markov kernel on $M$ is a map $Q: M \times \mathcal{B}(M) \rightarrow[0,1]$ (where $\mathcal{B}(M)$ denotes the Borel $\sigma$-algebra of $M$ ) with the following properties:

- for each $x \in M$ the map $Q(x, \cdot): \mathcal{B}(M) \rightarrow[0,1]$ is a probability measure on $M$, usually denoted by $Q(x, d y)$;
- for each $A \in \mathcal{B}(M)$ the function $Q(\cdot, A): M \rightarrow[0,1]$ is measurable.


## Lemma 2.11.

(i) For each pair $\mu, \nu \in \mathcal{P}_{2}(M)$ there exists a coupling $q$ (called 'optimal coupling') such that

$$
\mathrm{d}_{W}^{2}(\mu, \nu)=\int_{M \times M} \mathrm{~d}^{2}(x, y) d q(x, y)
$$

and there exist Markov kernels $Q, Q^{\prime}$ on $M$ ('optimal transport kernels') such that

$$
d q(x, y)=Q(x, d y) d \mu(x)=Q^{\prime}(y, d x) d \nu(x) .
$$

(In general, neither $q$ nor $Q, Q^{\prime}$ are unique.)
(ii) For each geodesic $\Gamma:[0,1] \rightarrow \mathcal{P}_{2}(M)$, each $l \in \mathbb{N}$ and each partition $0=t_{0}<t_{1}<\cdots<$ $t_{l}=1$ there exists a probability measure $\hat{q}$ on $M^{l+1}$ with the following properties:

- the projection on the $i$-th factor is $\Gamma\left(t_{i}\right)$ (for all $i=0,1, \ldots, l$ );
- for $\hat{q}$-almost every $x\left(x_{0}, \ldots, x_{l}\right) \in M^{l+1}$ and every $i, j=0,1, \ldots, l$

$$
\begin{equation*}
\mathrm{d}\left(x_{i}, x_{j}\right)=\left|t_{i}-t_{j}\right| \cdot \mathrm{d}\left(x_{0}, x_{l}\right) . \tag{2.9}
\end{equation*}
$$

In particular, for every pair $i, j \in\{0,1, \ldots, l\}$ the projection on the $i$-th and $j$-th factor is an optimal coupling of $\Gamma\left(t_{i}\right)$ and $\Gamma\left(t_{j}\right)$.
In the case $l=2$ and $t=\frac{1}{2}$, (2.9) states that for $\hat{q}$-a.e. $\left(x_{0}, x_{1}, x_{2}\right) \in M^{3}$ the point $x_{1}$ is a midpoint of $x_{0}$ and $x_{2}$.
(iii) If $M$ is a nonbranching geodesic space then in the previous situation for $\hat{q}$-almost every $\left(x_{0}, x_{1}, x_{2}\right)$ and $\left(y_{0}, y_{1}, y_{2}\right) \in M^{3}$

$$
x_{1}=y_{1} \quad \Rightarrow \quad\left(x_{0}, x_{2}\right)=\left(y_{0}, y_{2}\right) .
$$

Proof. (i) For the existence of optimal coupling, see [RR98] or [Du89], 11.8.2.
The existence of optimal transport kernels is a straightforward application of disintegration of measures on Polish spaces (or of the existence of regular conditional probabilities), namely, $Q$ is the disintegration of $q$ w.r.t $\mu$.
(ii) We assume $l=2$ and $t=\frac{1}{2}$. (The general case follows by iterated application and appropriate modifications.)
Let $q_{1}$ be an optimal coupling of $\Gamma(0)$ and $\Gamma\left(\frac{1}{2}\right)$ and let $q_{2}$ be an optimal coupling of $\Gamma\left(\frac{1}{2}\right)$ and $\Gamma(1)$. Then there exists a probability measure $\hat{q}$ on $M \times M \times M$ such that its projection on the first two factors is $q_{1}$ and the projection on last two factors is $q_{2}$ ([Du89], section 11.8). Hence, for $i=1,2,3$ the projection of $\hat{q}$ on the $i$-th factor is $\Gamma\left(\frac{i-1}{2}\right)$ and for $i=1,2$

$$
\mathrm{d}_{W}^{2}\left(\Gamma\left(\frac{i-1}{2}\right), \Gamma\left(\frac{i}{2}\right)\right)=\int_{M^{3}} \mathrm{~d}^{2}\left(x_{i-1}, x_{i}\right) d \hat{q}\left(x_{0}, x_{1}, x_{2}\right) .
$$

Then

$$
\begin{aligned}
\mathrm{d}_{W}(\Gamma(0), \Gamma(1)) \leq & {\left[\int \mathrm{d}^{2}\left(x_{0}, x_{2}\right) d \hat{q}\left(x_{0}, x_{1}, x_{2}\right)\right]^{\frac{1}{2}} } \\
\stackrel{(*)}{\leq} & {\left[\int\left[\mathrm{d}\left(x_{0}, x_{1}\right)+\mathrm{d}\left(x_{1}, x_{2}\right)\right]^{2} d \hat{q}\left(x_{0}, x_{1}, x_{2}\right)\right]^{\frac{1}{2}} } \\
\stackrel{(* *)}{\leq} & {\left[\int \mathrm{d}^{2}\left(x_{0}, x_{1}\right) d \hat{q}\left(x_{0}, x_{1}, x_{2}\right)\right]^{\frac{1}{2}} } \\
& +\left[\int \mathrm{d}^{2}\left(x_{1}, x_{2}\right) d \hat{q}\left(x_{0}, x_{1}, x_{2}\right)\right]^{\frac{1}{2}} \\
= & \mathrm{d}_{W}\left(\Gamma(0), \Gamma\left(\frac{1}{2}\right)\right)+\mathrm{d}_{W}\left(\Gamma\left(\frac{1}{2}\right), \Gamma(1)\right) .
\end{aligned}
$$

Since $\Gamma\left(\frac{1}{2}\right)$ is a midpoint of $\Gamma(0)$ and $\Gamma(1)$, the previous inequalities $\left({ }^{*}\right)$ and $\left({ }^{* *}\right)$ have to be equalities. From equality in $\left({ }^{*}\right)$ we conclude that $\hat{q}$-almost surely the point $x_{1}$ lies on some geodesic connecting $x_{0}$ and $x_{2}$. Equality in $\left({ }^{* *}\right)$ implies that $\hat{q}$-almost surely the point $x_{1}$ is a midpoint of $x_{0}$ and $x_{2}$.
(iii) Let $\eta=\Gamma\left(\frac{1}{2}\right)$ be the distribution of the midpoints and let $Q$ be a disintegration of $\hat{q}$ w.r.t. $\eta$, i.e.

$$
d \hat{q}(x, z, y)=Q(z, d(x, y)) d \eta(z) .
$$

We have to prove that for $\eta$-a.e. $z \in M$ the probability measure $Q(z, \cdot)$ is a Dirac measure (sitting on some $(x, y) \in M \times M)$. Denote the marginals of $Q(z, \cdot)$ by $p_{1}(z, \cdot)$ and $p_{2}(z, \cdot)$. Then

$$
\begin{aligned}
\int \mathrm{d}^{2}(x, y) Q(z, d(x, y)) & =\int\left[2 \mathrm{~d}^{2}(x, z)+2 \mathrm{~d}^{2}(z, y)\right] Q(z, d(x, y)) \\
& =\iint\left[2 \mathrm{~d}^{2}(x, z)+2 \mathrm{~d}^{2}(z, y)\right] p_{1}(z, d x) p_{2}(z, d y) \\
& \stackrel{(* * *)}{\geq} \iint \mathrm{d}^{2}(x, y) p_{1}(z, d x) p_{2}(z, d y) .
\end{aligned}
$$

The optimality of $\hat{q}$ implies that for $\eta$-a.e. $z \in M$ the measure $Q(z, \cdot)$ is an optimal coupling of $p_{1}(z, \cdot)$ and $p_{2}(z, \cdot)$. Hence, there has to be equality in $(* * *)$ which in turn implies that for $p_{1}(z,$.$) -a.e. x \in M$ and $p_{2}(z,$.$) -a.e. y \in M$ the point $z$ is a midpoint of $x$ and $y$. Since $M$ is nonbranching this implies that both $p_{1}(z,$.$) and p_{2}(z,$.$) are Dirac measures. Thus Q(z, \cdot)$ is also a Dirac measure. This proves the claim.

## Remark 2.12.

(i) Couplings $q$ of $\mu$ and $\nu$ are also called transportation plans from $\mu$ to $\nu$. If $\mu$ is the distribution of locations at which a good is produced and $\nu$ is the distribution of locations where it is consumed, then each coupling $q$ of $\mu$ and $\nu$ gives a plan how to transport the products to the consumer. More precisely, for each $x$ the kernel $Q(x, d y)$ determines how to distribute goods produced at the location $x$ to various consumers at location $y$.
(ii) The interpretation of Lemma 2.11 (ii) is that for each geodesic in $\mathcal{P}_{2}(M)$ the mass is transported along geodesics of the underlying space $M$. (iii) states that 'the paths of optimal mass transportation do not cross each other halfway'.
(iii) If $M$ is a complete Riemannian manifold with Riemannian volume $m$ then for each pair
$\mu, \nu \in \mathcal{P}_{2}(M)$ with $\mu \ll m$ there exists an optimal transport map $F_{1}: M \rightarrow M$ such that $d q(x, y)=Q(x, d y) d \mu(x)$ with

$$
Q(x, d y)=d \delta_{F_{1}(x)}(y)
$$

is the unique optimal coupling of $\mu$ and $\nu$.
More precisely, there exists a function $\varphi: M \rightarrow \mathbb{R}$ such that for $\mu$-a.e. $x \in M$ and $t \in[0,1]$

$$
F_{t}(x)=\exp _{x}(-t \nabla \varphi(x))
$$

exists and the unique geodesic $\Gamma$ in $\mathcal{P}_{2}(M)$ connecting $\mu=\Gamma(0)$ and $\nu=\Gamma(1)$ is given by

$$
\Gamma(t):=\left(F_{t}\right)_{*} \mu
$$

the push forward of $\mu$ under $F_{t}$, [CMS01].

## 3 Metric Measure Spaces

### 3.1 The Metric $\mathbb{D}$

Throughout this paper, a metric measure space will always be a triple ( $M, \mathrm{~d}, m$ ) where

- $(M, \mathrm{~d})$ is a complete separable metric space,
- $m$ is a measure on $(M, \mathcal{B}(M))$ which is locally finite in the sense that $m\left(B_{r}(x)\right)<\infty$ for all $x \in M$ and all sufficiently small $r>0$.

A metric measure space $(M, \mathrm{~d}, m)$ is called normalized iff $m(M)=1$. It is called compact or locally compact or geodesic iff the metric space $(M, \mathrm{~d})$ is compact or locally compact or geodesic, resp.
Two metric measure spaces $(M, \mathrm{~d}, m)$ and $\left(M^{\prime}, \mathrm{d}^{\prime}, m^{\prime}\right)$ are called isomorphic iff there exists an isometry $\psi: M_{0} \rightarrow M_{0}^{\prime}$ between the supports $M_{0}:=\operatorname{supp}[m] \subset M$ and $M_{0}^{\prime}:=\operatorname{supp}\left[m^{\prime}\right] \subset M^{\prime}$ such that

$$
\psi_{*} m=m^{\prime}
$$

The variance of a metric measure space $(M, \mathrm{~d}, m)$ is defined as

$$
\begin{equation*}
\mathbb{V a r}(M, \mathrm{~d}, m)=\inf \int_{M^{\prime}} \mathrm{d}^{\prime 2}(z, x) d m^{\prime}(x) \tag{3.1}
\end{equation*}
$$

where the inf is taken over all metric measure spaces $\left(M^{\prime}, \mathrm{d}^{\prime}, m^{\prime}\right)$ which are isomorphic to $(M, \mathrm{~d}, m)$ and over all $z \in M^{\prime}$. Note that a normalized metric measure space $(M, \mathrm{~d}, m)$ has finite variance if and only if

$$
\begin{equation*}
\int_{M} \mathrm{~d}^{2}(z, x) d m(x)<\infty \tag{3.2}
\end{equation*}
$$

for some (hence all) $z \in M$. Similarly, the radius of $(M, \mathrm{~d}, m)$ is defined as

$$
\begin{aligned}
\operatorname{rad}(M, \mathrm{~d}, m)=\inf \{R>0: \quad & \exists\left(M^{\prime}, \mathrm{d}^{\prime}, m^{\prime}\right) \text { isom. to }(M, \mathrm{~d}, m) \\
& \left.\exists z \in M^{\prime}: \operatorname{supp}\left[m^{\prime}\right] \subset \overline{B^{\prime}} R(z)\right\}
\end{aligned}
$$

The diameter of a metric measure space $(M, \mathrm{~d}, m)$ is defined as the diameter of the metric space $(\operatorname{supp}[m], d)$ :

$$
\operatorname{diam}(M, \mathrm{~d}, m)=\sup \{d(x, y): x, y \in \operatorname{supp}[m]\}
$$

Example 3.1. Let $M=\mathbb{R}^{2}$ with Euclidean distance d and $m=\frac{1}{3}\left(\delta_{x_{1}}+\delta_{x_{2}}+\delta_{x_{3}}\right)$ where $x_{1}, x_{2}, x_{3}$ are the vertices of an equilateral triangle of sidelength 1 . Then

$$
\operatorname{Var}(M, \mathrm{~d}, m)=\frac{1}{4} \quad \text { whereas } \quad \inf _{z \in M} \int \mathrm{~d}^{2}(z, x) d m(x)=\frac{1}{3} .
$$

(Hint: embed $\operatorname{supp}[m]$ isometrically into a graph or into a hyperbolic space with curvature close to $-\infty$.)

The family of all isomorphism classes of metric measure spaces will be denoted by $\mathbb{X}$. For each $\lambda \in \mathbb{R}_{+}$, let $\mathbb{X}_{\lambda}$ denote the family of isomorphism classes of metric measure spaces ( $M, \mathrm{~d}, m$ ) with finite variances and total mass $m(M)=\lambda$. Moreover, for $\Delta \in \mathbb{R}_{+}$, let $\mathbb{X}_{\lambda}(\Delta)$ denote the family of isomorphism classes of metric measure spaces $(M, \mathrm{~d}, m)$ with diameter $\leq \Delta$ and total mass $m(M)=\lambda$. If $\lambda \Delta \neq 0$, the map

$$
(M, \mathrm{~d}, m) \mapsto(M, \Delta \mathrm{~d}, \lambda m)
$$

defines a bijection between $\mathbb{X}_{1}$ and $\mathbb{X}_{\lambda}$ and also a bijection between $\mathbb{X}_{1}(1)$ and $\mathbb{X}_{\lambda}(\Delta)$.
For $\lambda>0$, the family $\mathbb{X}_{\lambda}$ contains a unique element with $\mathbb{V} \operatorname{ar}(M, \mathrm{~d}, m)=0$, namely, $m=\lambda \cdot \delta_{o}$ for some $o \in M$. Here a priori $M$ is an arbitrary nonempty set. But without restriction it contains just one point, say $M=\{o\}$. The family $\mathbb{X}_{0}$ is pathological: it contains only one element, the 'empty space'.

## Definition 3.2.

(i) Given two metric measure spaces $(M, \mathrm{~d}, m)$ and $\left(M^{\prime}, \mathrm{d}^{\prime}, m^{\prime}\right)$ we say that a measure $q$ on the product space $M \times M^{\prime}$ is a coupling of $m$ and $m^{\prime}$ iff

$$
\begin{equation*}
q\left(A \times M^{\prime}\right)=m(A), \quad q\left(M \times A^{\prime}\right)=m^{\prime}\left(A^{\prime}\right) \tag{3.3}
\end{equation*}
$$

for all measurable sets $A \subset M, A^{\prime} \subset M^{\prime}$. We say that a pseudo metric $\hat{\mathrm{d}}$ on the disjoint union $M \sqcup M^{\prime}$ is a coupling of d and $\mathrm{d}^{\prime}$ iff

$$
\begin{equation*}
\hat{\mathrm{d}}(x, y)=\mathrm{d}(x, y), \quad \hat{\mathrm{d}}\left(x^{\prime}, y^{\prime}\right)=\mathrm{d}^{\prime}\left(x^{\prime}, y^{\prime}\right) \tag{3.4}
\end{equation*}
$$

for all $x, y \in \operatorname{supp}[m] \subset M$ and all $x^{\prime}, y^{\prime} \in \operatorname{supp}\left[m^{\prime}\right] \subset M^{\prime}$.
(ii) We define the distance $\mathbb{D}$ between two metric measure spaces by

$$
\begin{aligned}
\mathbb{D}\left((M, \mathrm{~d}, m),\left(M^{\prime}, \mathrm{d}^{\prime}, m^{\prime}\right)\right)=\inf \left\{\left(\int_{M \times M^{\prime}}\right.\right. & \left.\hat{\mathrm{d}}^{2}(x, y) d q(x, y)\right)^{1 / 2}: \\
& \hat{\mathrm{d}} \text { is a coupling of } \mathrm{d} \text { and } \mathrm{d}^{\prime} \\
& \left.q \text { is a coupling of } m \text { and } m^{\prime}\right\} .
\end{aligned}
$$

Remark 3.3. (i) Note that the integrals involved in the definition of $\mathbb{D}$ are well-defined since each coupling $\hat{d}$ is a function on $\left(M \sqcup M^{\prime}\right) \times\left(M \sqcup M^{\prime}\right)=(M \times M) \sqcup\left(M \times M^{\prime}\right) \sqcup\left(M^{\prime} \times M\right) \sqcup\left(M^{\prime} \times M^{\prime}\right)$ and each coupling $q$ is a measure on $M \times M^{\prime}$.
(ii) In the definition of the distance $\mathbb{D}$ we may restrict ourselves to take the infimum over all complete separable metrics $\hat{\mathrm{d}}$ on $M \sqcup M^{\prime}$ which are couplings of d and $\mathrm{d}^{\prime}$. Indeed, given any (pseudo metric) coupling $\hat{d}$ of $d$ and $d^{\prime}$ and any $\epsilon>0$ we obtain a complete separable metric $d_{\epsilon}$ which is a coupling of $d$ and $d^{\prime}$ as follows:

$$
\hat{\mathrm{d}}_{\epsilon}= \begin{cases}\mathrm{d}, & \text { on }(M \times M) \sqcup\left(M^{\prime} \times M^{\prime}\right) \\ \mathrm{d}+\epsilon, & \text { on }\left(M \times M^{\prime}\right) \sqcup\left(M^{\prime} \times M\right) .\end{cases}
$$

(iii) One easily verifies that

$$
\begin{equation*}
\mathbb{D}\left((M, \mathrm{~d}, m),\left(M^{\prime}, \mathrm{d}^{\prime}, m^{\prime}\right)\right)=\inf \hat{\mathrm{d}}_{W}\left(\psi_{*} m, \psi_{*}^{\prime} m^{\prime}\right) \tag{3.5}
\end{equation*}
$$

where the inf is taken over all metric spaces ( $\hat{M}, \hat{\mathrm{~d}}$ ) with isometric embeddings $\psi: M_{0} \hookrightarrow \hat{M}$, $\psi^{\prime}: M_{0}^{\prime} \hookrightarrow \hat{M}$ of the supports $M_{0}$ and $M_{0}^{\prime}$ of $m$ and $m^{\prime}$, resp. Here $\hat{\mathrm{d}}_{W}$ denotes the $L_{2^{-}}$ Wasserstein distance for measures on $\hat{M}$ as introduced in Chapter 2. In other words,

$$
\begin{equation*}
\mathbb{D}\left((M, \mathrm{~d}, m),\left(M^{\prime}, \mathrm{d}^{\prime}, m^{\prime}\right)\right)=\inf \left(\int_{M \times M^{\prime}} \hat{\mathrm{d}}^{2}\left(\psi(x), \psi^{\prime}\left(x^{\prime}\right)\right) d q\left(x, x^{\prime}\right)\right)^{1 / 2} \tag{3.6}
\end{equation*}
$$

where the inf now is taken over all metric spaces ( $\hat{M}, \hat{\mathrm{~d}}$ ) with isometric embeddings $\psi: M_{0} \hookrightarrow \hat{M}$, $\psi^{\prime}: M_{0}^{\prime} \hookrightarrow \hat{M}$ and over all couplings $q$ of $m$ and $m^{\prime}$.
Indeed, the set $\hat{M}$ can always be chosen as the disjoint union of $M_{0}$ and $M_{0}^{\prime}$ (the supports of $m$ and $m^{\prime}$ ), i.e.

$$
\hat{M}=M_{0} \sqcup M_{0}^{\prime}
$$

and $\psi, \psi^{\prime}$ can be chosen as identities. Hence,

$$
\mathbb{D}\left((M, \mathrm{~d}, m),\left(M^{\prime}, \mathrm{d}^{\prime}, m^{\prime}\right)\right)=\inf _{\hat{\mathrm{d}}} \hat{\mathrm{~d}}_{W}\left(m, m^{\prime}\right)
$$

where the $\inf _{\hat{\mathrm{d}}}$ is taken over all metrics (or, equivalently, over all pseudo metrics) on $M_{0} \sqcup M_{0}^{\prime}$ the restriction of which coincide on $M_{0}$ with d and on $M_{0}^{\prime}$ with $\mathrm{d}^{\prime}$.

Let us summarize some elementary properties of $\mathbb{D}$.

## Lemma 3.4.

(i) If $(M, \mathrm{~d})=\left(M^{\prime}, \mathrm{d}^{\prime}\right)$ then $\mathbb{D}\left((M, \mathrm{~d}, m),\left(M, \mathrm{~d}, m^{\prime}\right)\right) \leq \mathrm{d}_{W}\left(m, m^{\prime}\right)$. In general, there will be no equality.
(ii) If $m(M) \neq m^{\prime}\left(M^{\prime}\right)$ then $\mathbb{D}\left((M, \mathrm{~d}, m),\left(M^{\prime}, \mathrm{d}^{\prime}, m^{\prime}\right)\right)=+\infty$.
(iii) For all $\alpha, \beta \in \mathbb{R}_{+}$

$$
\mathbb{D}\left((M, \alpha \mathrm{~d}, \beta m),\left(M^{\prime}, \alpha \mathrm{d}^{\prime}, \beta m^{\prime}\right)\right)=\alpha \sqrt{\beta} \cdot \mathbb{D}\left((M, \mathrm{~d}, m),\left(M^{\prime}, \mathrm{d}^{\prime}, m^{\prime}\right)\right)
$$

(iv) If $m=\sum_{n} m_{n}$ and $m^{\prime}=\sum_{n} m_{n}^{\prime}$ then

$$
\mathbb{D}^{2}\left((M, \mathrm{~d}, m),\left(M^{\prime}, \mathrm{d}^{\prime}, m^{\prime}\right)\right) \leq \sum_{n} \mathbb{D}^{2}\left(\left(M, \mathrm{~d}, m_{n}\right),\left(M^{\prime}, \mathrm{d}^{\prime}, m_{n}^{\prime}\right)\right) .
$$

In particular, with $M_{n}=\operatorname{supp}\left[m_{n}\right], M_{n}^{\prime}=\operatorname{supp}\left[m_{n}^{\prime}\right]$

$$
\mathbb{D}^{2}\left((M, \mathrm{~d}, m),\left(M^{\prime}, \mathrm{d}^{\prime}, m^{\prime}\right)\right) \leq \sum_{n} \mathbb{D}^{2}\left(\left(M_{n},\left.\mathrm{~d}\right|_{M_{n}}, m_{n}\right),\left(M_{n}^{\prime},\left.\mathrm{d}^{\prime}\right|_{M_{n}^{\prime}}, m_{n}^{\prime}\right)\right)
$$

Now let us concentrate on normalized metric measure spaces.

## Lemma 3.5.

(i) If $m(M)=1$ and $m^{\prime}=\delta_{o}$ for some $o \in M^{\prime}$ then

$$
\mathbb{D}^{2}\left((M, \mathrm{~d}, m),\left(M^{\prime}, \mathrm{d}^{\prime}, m^{\prime}\right)\right)=\mathbb{V} \operatorname{ar}(M, \mathrm{~d}, m) .
$$

(ii) The family $\mathbb{X}_{1, *}$ of isomorhpism classes of $(M, \mathrm{~d}, m)$ with finite supports $M_{0}$, say $\left\{x_{1}, \ldots, x_{n}\right\}$, and uniform distribution $m=\frac{1}{n} \sum \delta_{x_{i}}$ ('normalized configurations') is dense in $\mathbb{X}_{1}$.
(iii) For each $(M, \mathrm{~d}, m) \in \mathbb{X}_{1}$ let $X_{1}, X_{2}, \ldots$ be an independent sequence of random variables $X_{i}: \Omega \rightarrow M$ (defined on some probability space $\Omega$ with values in $M$ ) with distribution $m$ and let

$$
m_{n}(\omega, \cdot):=\frac{1}{n} \sum_{i=1}^{n} \delta_{X_{i}(\omega)}
$$

be their empirical distributions. Then for m-almost every $\omega \in \Omega$

$$
\left(M, \mathrm{~d}, m_{n}(\omega, \cdot)\right) \rightarrow(M, \mathrm{~d}, m)
$$

in $\left(\mathbb{X}_{1}, \mathbb{D}\right)$ as $n \rightarrow \infty$.
(iv) If $m=\frac{1}{n} \sum_{i=1}^{n} \delta_{x_{i}}$ and $m^{\prime}=\frac{1}{n} \sum_{i=1}^{n} \delta_{x_{i}^{\prime}}$ then

$$
\mathbb{D}\left((M, \mathrm{~d}, m),\left(M^{\prime}, \mathrm{d}^{\prime}, m^{\prime}\right)\right) \leq \sup _{i, j}\left|\mathrm{~d}_{i j}-\mathrm{d}_{i j}^{\prime}\right|
$$

where $\mathrm{d}_{i j}:=\mathrm{d}\left(x_{i}, x_{j}\right)$ and $\mathrm{d}_{i j}^{\prime}:=\mathrm{d}^{\prime}\left(x_{i}^{\prime}, x_{j}^{\prime}\right)$.
Proof. (i) is obvious.
(ii) Given $(M, \mathrm{~d}, m) \in \mathbb{X}_{1}$ we have $m \in \mathcal{P}_{2}(M, \mathrm{~d})$ by (3.2). Then by Proposition 2.10(i)

$$
\forall \epsilon>0: \exists n \in \mathbb{N}, x_{1}, \ldots, x_{n} \in M: \quad \mathrm{d}_{W}(m, \bar{m}) \leq \epsilon
$$

where $\bar{m}:=\frac{1}{n} \sum_{i=1}^{n} \delta_{x_{i}}$. Hence, $(M, \mathrm{~d}, \bar{m}) \in \mathbb{X}_{1, *}$ and $\mathbb{D}((M, \mathrm{~d}, m),(M, \mathrm{~d}, \bar{m})) \leq \mathrm{d}_{W}(m, \bar{m}) \leq \epsilon$. (iii) follows from the Empirical Law of Large Numbers or Varadarajan's Theorem, e.g. [Du89], Theorem 11.4.1.
(iv) Assume (without restriction) that $M=\left\{x_{1}, \ldots, x_{n}\right\}, M^{\prime}=\left\{x_{1}^{\prime}, \ldots, x_{n}^{\prime}\right\}$, and $\left|\mathrm{d}_{i j}-\mathrm{d}_{i j}^{\prime}\right| \leq \epsilon$ for all $i, j$ (with $\mathrm{d}_{i j}, \mathrm{~d}_{i j}^{\prime}$ as above). Define $\hat{\mathrm{d}}$ on $M \times M^{\prime}$ by

$$
\hat{\mathrm{d}}\left(x_{i}, x_{j}^{\prime}\right):=\inf _{k=1, \ldots, n}\left[\mathrm{~d}\left(x_{i}, x_{k}\right)+\mathrm{d}^{\prime}\left(x_{k}^{\prime}, x_{j}^{\prime}\right)\right]+\epsilon
$$

and analogously on $M^{\prime} \times M$. As usual, put $\hat{\mathrm{d}}:=\mathrm{d}$ on $M \times M$ and $\hat{\mathrm{d}}:=\mathrm{d}^{\prime}$ on $M^{\prime} \times M^{\prime}$. Moreover, put

$$
m=\frac{1}{n} \sum_{i=1}^{n} \delta_{\left(x_{i}, x_{i}^{\prime}\right)} .
$$

Then $\hat{d}$ is a coupling of $d$ and $d^{\prime}$ and $q$ is a coupling of $m$ and $m^{\prime}$. Thus

$$
\mathbb{D}^{2}\left((M, \mathrm{~d}, m),\left(M^{\prime}, \mathrm{d}^{\prime}, m^{\prime}\right)\right) \leq \int_{M \times M^{\prime}} \hat{\mathrm{d}}^{2}(x, y) d q(x, y)=\epsilon^{2}
$$

Theorem 3.6. $\left(\mathbb{X}_{1}, \mathbb{D}\right)$ is a complete separable metric space.
Proof.
(i) Obviously, $\mathbb{D}$ is well-defined and symmetric on $\mathbb{X}_{1} \times \mathbb{X}_{1}$ with values in $\mathbb{R}_{+}$.
(ii) The density of $\mathbb{X}_{1, *}$ in $\mathbb{X}_{1}$ follows from Lemma 3.5 (ii).
(iii) According to (ii), separability of $\mathbb{X}_{1}$ will follow from separability of $\mathbb{X}_{1, *}$. The latter is the disjoint union $\dot{U}_{n \in \mathbb{N}} \tilde{\mathcal{K}}(n)$ where $\tilde{\mathcal{K}}(n):=\left\{(M, \mathrm{~d}, m) \in \mathbb{X}_{1, *}: \operatorname{supp}[m]\right.$ has $n$ points $\}$. But $\tilde{\mathcal{K}}(n)$ can be identified with

$$
\begin{aligned}
& \mathcal{K}(n)=\left\{D=\left(D_{i j}\right)_{i, j} \in \mathbb{R}_{+}^{n \times n}: \quad \forall i, j, k \in\{1, \ldots, n\}:\right. \\
&\left.D_{i j}=D_{j i}, D_{i j}+D_{j k} \geq D_{i k} \text { and } D_{i j}=0 \Leftrightarrow i=j\right\} .
\end{aligned}
$$

Now each of the $\mathcal{K}(n)$ is separable (as a subset of $\mathbb{R}^{n \times n}$ ), hence, $\tilde{\mathcal{K}}(n)$ is separable (Lemma $3.5(\mathrm{iv})$ ) and thus finally $\mathbb{X}_{1, *}$ is separable.
(iv) In order to prove the triangle inequality let three metric measure spaces $\left(M_{i}, \mathrm{~d}_{i}, m_{i}\right) \in$ $\mathbb{X}_{1}, i=1,2,3$, be given. Without restriction, we may assume $M_{i}=\operatorname{supp}\left[m_{i}\right]$ for $i=1,2,3$. Then for each $\epsilon>0$ there exist a complete separable metric $\mathrm{d}_{12}$ on $M_{1} \sqcup M_{2}$ and a complete separable metric $\mathrm{d}_{23}$ on $M_{2} \sqcup M_{3}$ such that

$$
\begin{aligned}
& \mathbb{D}\left(\left(M_{1}, \mathrm{~d}_{1}, m_{1}\right),\left(M_{2}, \mathrm{~d}_{2}, m_{2}\right)\right) \geq \mathrm{d}_{12}^{W}\left(m_{1}, m_{2}\right)-\epsilon, \\
& \mathbb{D}\left(\left(M_{2}, \mathrm{~d}_{2}, m_{2}\right),\left(M_{3}, \mathrm{~d}_{3}, m_{3}\right)\right) \geq \mathrm{d}_{23}^{W}\left(m_{2}, m_{3}\right)-\epsilon,
\end{aligned}
$$

and $d_{i j}$ restricted to $M_{i}$ coincides with $d_{i}$, restricted to $M_{j}$ coincides with $d_{j}$ for $(i, j)=$ $(1,2)$ or $(2,3)$. (Here for typographical reasons, we use not a lower but an upper index to indicate the Wasserstein metric derived from a given metric.) Now define d on $M \times M$ with $M:=M_{1} \sqcup M_{2} \sqcup M_{3}$ by

$$
\mathrm{d}(x, y)= \begin{cases}\mathrm{d}_{12}(x, y), & \text { if } x, y \in M_{1} \sqcup M_{2} \\ \mathrm{~d}_{23}(x, y), & \text { if } x, y \in M_{2} \sqcup M_{3} \\ \inf _{z \in M_{2}}\left[\mathrm{~d}_{12}(x, z)+\mathrm{d}_{23}(z, y)\right], & \text { if } x \in M_{1}, y \in M_{3} \\ \inf _{z \in M_{2}}\left[\mathrm{~d}_{23}(x, z)+\mathrm{d}_{12}(z, y)\right], & \text { if } x \in M_{3}, y \in M_{1} .\end{cases}
$$

Obviously, d is a complete separable metric on $M$ and, restricted to $M_{i}$ it coincides with $\mathrm{d}_{i}$ (for each $i=1,2,3$ ).
Then by the triangle inequality for $\mathrm{d}^{W}$ (Proposition 2.10 (i))

$$
\begin{aligned}
& \mathbb{D}\left(\left(M_{1}, \mathrm{~d}_{1}, m_{1}\right),\left(M_{3}, \mathrm{~d}_{3}, m_{3}\right)\right) \\
& \leq \mathrm{d}^{W}\left(m_{1}, m_{3}\right) \\
& \leq \mathrm{d}^{W}\left(m_{1}, m_{2}\right)+\mathrm{d}^{W}\left(m_{2}, m_{3}\right) \\
& =\mathrm{d}_{12}^{W}\left(m_{1}, m_{2}\right)+\mathrm{d}_{23}^{W}\left(m_{2}, m_{3}\right) \\
& \leq \mathbb{D}\left(\left(M_{1}, \mathrm{~d}_{1}, m_{1}\right),\left(M_{2}, \mathrm{~d}_{2}, m_{2}\right)\right) \\
& \quad \quad+\mathbb{D}\left(\left(M_{2}, \mathrm{~d}_{2}, m_{2}\right),\left(M_{3}, \mathrm{~d}_{3}, m_{3}\right)\right)+2 \epsilon .
\end{aligned}
$$

This proves the claim.
(v) In order to prove completeness let $\left(\left(M_{n}, \mathrm{~d}_{n}, m_{n}\right)\right)_{n \in \mathbb{N}}$ be a Cauchy sequence in $\left(\mathbb{X}_{1}, \mathbb{D}\right)$. Let us choose a subsequence such that

$$
\mathbb{D}\left(\left(M_{n_{k}}, \mathrm{~d}_{n_{k}}, m_{n_{k}}\right),\left(M_{n_{k+1}}, \mathrm{~d}_{n_{k+1}}, m_{n_{k+1}}\right)\right) \leq 2^{-k-1}
$$

for all $k \in \mathbb{N}$. Then there exist a coupling $\hat{\mathrm{d}}_{k+1}$ of $\mathrm{d}_{n_{k}}, \mathrm{~d}_{n_{k+1}}$ and a coupling $\hat{q}_{k+1}$ of $m_{n_{k}}$, $m_{n_{k+1}}$ such that

$$
\left(\int \hat{\mathrm{d}}_{k+1}^{2}(x, y) d \hat{q}_{k+1}(x, y)\right)^{1 / 2} \leq 2^{-k}
$$

Without restriction $\hat{\mathrm{d}}_{k+1}$ is a complete separable metric. Let us define recursively a sequence of complete separable metric spaces $\left(M_{k}^{\prime}, \mathrm{d}_{k}^{\prime}\right)$ as follows: $\left(M_{1}^{\prime}, \mathrm{d}_{1}^{\prime}\right):=\left(M_{n_{1}}, \mathrm{~d}_{n_{1}}\right)$ and $M_{k+1}^{\prime}=M_{k}^{\prime} \sqcup M_{n_{k+1}} / \sim$ with $x \sim y$ iff $\mathrm{d}_{k+1}^{\prime}(x, y)=0$ where

$$
\mathrm{d}_{k+1}^{\prime}(x, y)= \begin{cases}\mathrm{d}_{k}^{\prime}(x, y), & \text { if } x, y \in M_{k}^{\prime} \\ \mathrm{d}_{k+1}(x, y) & \text { if } x, y \in M_{n_{k}} \sqcup M_{n_{k+1}} \\ \inf _{z \in M_{n_{k}}}\left[\mathrm{~d}_{k}^{\prime}(x, z)+\hat{\mathrm{d}}_{k+1}(z, y)\right], & \text { if } x \in M_{k}^{\prime}, y \in M_{n_{k}} \sqcup M_{n_{k+1}} .\end{cases}
$$

This way, $\left(M_{k}^{\prime}, \mathrm{d}_{k}^{\prime}\right)$ is a sequence of complete separable metric spaces with $M_{n_{k}} \subset M_{k}^{\prime}$ and $M_{k}^{\prime} \subset M_{k+l}^{\prime}$ for all $k, l$. Hence, $M^{\prime}=\cup_{k=1}^{\infty} M_{k}^{\prime}$ is naturally equipped with a metric $\mathrm{d}^{\prime}=\lim \mathrm{d}_{k}^{\prime}$.
Let ( $M, \mathrm{~d}$ ) be the completion of $\left(M^{\prime}, \mathrm{d}^{\prime}\right)$. Then $\left(M_{n_{k}}, \mathrm{~d}_{n_{k}}\right)$ is isometrically embedded in $(M, \mathrm{~d})$ for each $k \in \mathbb{N}$ and the measure $m_{n_{k}}$ on $M_{n_{k}}$ defines a push forward measure $\bar{m}_{n_{k}}$ on $M$. By construction

$$
\mathrm{d}_{W}\left(\bar{m}_{n_{k}}, \bar{m}_{n_{k+1}}\right) \leq\left(\int \hat{\mathrm{d}}_{k+1}^{2}(x, y) d \hat{q}_{k+1}(x, y)\right)^{1 / 2} \leq 2^{-k}
$$

for all $k \in \mathbb{N}$. Hence, $\left(\bar{m}_{n_{k}}\right)_{k \in \mathbb{N}}$ is a Cauchy sequence in $\left(\mathcal{P}_{2}(M), \mathrm{d}_{W}\right)$. According to Proposition 2.10 the latter is complete. That is, there exists a probability measure $m$ on ( $M, \mathrm{~d}$ ) such that

$$
\mathbb{D}\left(\left(M_{n_{k}}, \mathrm{~d}_{n_{k}}, m_{n_{k}}\right),(M, \mathrm{~d}, m)\right) \leq \mathrm{d}_{W}\left(\bar{m}_{n_{k}}, m\right) \rightarrow 0
$$

as $k \rightarrow \infty$. This in turn implies

$$
\mathbb{D}\left(\left(M_{n}, \mathrm{~d}_{n}, m_{n}\right),(M, \mathrm{~d}, m)\right) \rightarrow 0
$$

as $n \rightarrow \infty$ which proves the claim.
(vi) Nondegeneracy of $\mathbb{D}$ will follow from the corresponding property of Gromov's metric $\square_{1}$ together with the following Lemma.

Lemma 3.7. The metric $\mathbb{D}$ can be estimated from below in terms of Gromov's metric $\square_{1}$ ([Gr99], $3 \frac{1}{2} .12$ ) as follows:

$$
\frac{1}{2} \square_{1} \leq \mathbb{D}^{2 / 3} .
$$

Proof. In order to prove this estimate, let normalized metric measure spaces ( $M, \mathrm{~d}, m$ ) and $\left(M^{\prime}, \mathrm{d}^{\prime}, m^{\prime}\right)$ be given with $\mathbb{D}\left((M, \mathrm{~d}, m),\left(M^{\prime}, \mathrm{d}^{\prime}, m^{\prime}\right)\right)<\epsilon^{3 / 2}$ for some $\epsilon>0$. Then for some metric $\hat{\mathrm{d}}$ on $M \sqcup M^{\prime}$ extending d and $\mathrm{d}^{\prime}$ and for some coupling $q$ of $m$ and $m^{\prime}$

$$
\int \hat{\mathrm{d}}^{2}\left(x, x^{\prime}\right) d q\left(x, x^{\prime}\right)<\epsilon^{3} .
$$

Hence, $q\left(\left\{\left(x, x^{\prime}\right) \in M \times M^{\prime}: \hat{\mathrm{d}}\left(x, x^{\prime}\right) \geq \epsilon\right\}\right)<\epsilon$. Therefore there exists a measurable map $\Phi:\left[0,1\left[\rightarrow M \times M^{\prime}\right.\right.$ such that $\Phi_{*} \lambda=q$ where $\lambda$ denotes the Lebesgue measure on $[0,1[$ ('parametrization of $q^{\prime}$ ) and there exists a measurable set $X_{\epsilon} \subset\left[0,1\left[\right.\right.$ with $\lambda\left(X_{\epsilon}\right)<\epsilon$ such that for all $x \in\left[0,1\left[\backslash X_{\epsilon}\right.\right.$ :

$$
\begin{equation*}
\hat{\mathrm{d}}(\Phi(x))<\epsilon . \tag{3.7}
\end{equation*}
$$

If we write $\Phi(x)=\left(\varphi(x), \varphi^{\prime}(x)\right)$ with $\varphi:\left[0,1\left[\rightarrow M, \varphi^{\prime}:\left[0,1\left[\rightarrow M^{\prime}\right.\right.\right.\right.$ then $\varphi_{*} \lambda=m$ and $\varphi_{*}^{\prime} \lambda=m^{\prime}$. Moreover, for all $x, y \in\left[0,1\left[\backslash X_{\epsilon}\right.\right.$

$$
\begin{aligned}
& \left|\mathrm{d}(\varphi(x), \varphi(y))-\mathrm{d}^{\prime}\left(\varphi^{\prime}(x), \varphi^{\prime}(y)\right)\right| \\
\leq & \hat{\mathrm{d}}\left(\varphi(x), \varphi^{\prime}(x)\right)+\hat{\mathrm{d}}\left(\varphi(y), \varphi^{\prime}(y)\right) \\
= & \hat{\mathrm{d}}(\Phi(x))+\hat{\mathrm{d}}(\Phi(y))<2 \epsilon
\end{aligned}
$$

according to (3.7). This proves

$$
\square_{1}\left((M, \mathrm{~d}, m),\left(M^{\prime}, \mathrm{d}^{\prime}, m^{\prime}\right)\right)<2 \epsilon
$$

### 3.2 Examples for $\mathbb{D}$-Convergence

Let us demonstrate the notion of $\mathbb{D}$-convergence with various examples.
Example 3.8. ('Products')
Let $\left(M_{n}, \mathrm{~d}_{n}, m_{n}\right) \in \mathbb{X}_{1}$ for $n \in \mathbb{N}$. Then $\left(\bigotimes_{n=1}^{l}\left(M_{n}, \mathrm{~d}_{n}, m_{n}\right)\right)_{l \in \mathbb{N}}$ is a $\mathbb{D}$-Cauchy sequence in $\mathbb{X}_{1}$ if (and only if)

$$
\sum_{n=1}^{\infty} \mathbb{V} \operatorname{ar}\left(M_{n}, \mathrm{~d}_{n}, m_{n}\right)<\infty
$$

In this case, as $l \rightarrow \infty$

$$
\bigotimes_{n=1}^{l}\left(M_{n}, \mathrm{~d}_{n}, m_{n}\right) \xrightarrow{\mathbb{D}} \bigotimes_{n=1}^{\infty}\left(M_{n}, \mathrm{~d}_{n}, m_{n}\right) .
$$

Proof. Obviously, for all $k$ and $l$

$$
\begin{aligned}
& \mathbb{D}^{2}\left(\bigotimes_{n=1}^{l}\left(M_{n}, \mathrm{~d}_{n}, m_{n}\right), \bigotimes_{n=1}^{l+k}\left(M_{n}, \mathrm{~d}_{n}, m_{n}\right)\right) \\
& \quad \leq \mathbb{D}^{2}\left(\left(\{o\}, 0, \delta_{o}\right), \bigotimes_{n=l+1}^{l+k}\left(M_{n}, \mathrm{~d}_{n}, m_{n}\right)\right) \\
& \quad \leq \sum_{n=l+1}^{l+k} \mathbb{D}^{2}\left(\left(\{o\}, 0, \delta_{o}\right),\left(M_{n}, \mathrm{~d}_{n}, m_{n}\right)\right)=\sum_{n=l+1}^{l+k} \operatorname{Var}\left(M_{n}, \mathrm{~d}_{n}, m_{n}\right) .
\end{aligned}
$$

(Actually, all these inequalities are equalities.) This proves the claim(s).
Example 3.9. ('Dimension Increasing to Infinity')
Let $M_{n}=\mathbb{R}$ with Euclidean distance,

$$
d m_{n}(x)=\frac{1}{\sqrt{2 \pi} \sigma_{n}} \exp \left(-\frac{x^{2}}{2 \sigma_{n}^{2}}\right)
$$

and $M^{(l)}=\mathbb{R}^{l}$ with Euclidean distance,

$$
d m^{(l)}(x)=\frac{1}{(2 \pi)^{l / 2} \prod_{n=1}^{l} \sigma_{n}} \exp \left(-\frac{1}{2} \sum_{n=1}^{l}\left(\frac{x_{n}}{\sigma_{n}}\right)^{2}\right) d x
$$

Then

$$
\left(M^{(l)}, \mathrm{d}^{(l)}, m^{(l)}\right) \xrightarrow{\mathbb{D}}\left(M^{(\infty)}, \mathrm{d}^{(\infty)}, m^{(\infty)}\right)
$$

if and only if $\sum_{n=1}^{\infty} \sigma_{n}^{2}<\infty$.

Example 3.10. ('Increasing Finite Dimension')
(i) Let $M_{n}=\left(\frac{1}{n} \mathbb{Z} \cap[0,1]\right)^{k}$ be the rescaled $k$-dimensional lattice, d be the Euclidean distance in $\mathbb{R}^{k}$ and $m_{n}$ be the renormalized counting measure on $M_{n}$. Then as $n \rightarrow \infty$

$$
\left(M_{n}, \mathrm{~d}, m_{n}\right) \xrightarrow{\mathbb{D}}\left([0,1]^{k}, \mathrm{~d}, m\right)
$$

with $m$ being the $k$-dimensional Lebesgue measure in $[0,1]^{k}$.

(ii) Similarly, if $\tilde{M}_{n}$ denotes the graph obtained from $M_{n}$ with edges between next neighbors and $\tilde{m}_{n}$ being the 1-dimensional Lebesgue measure on the edges:

$$
\left(\tilde{M}_{n}, \mathrm{~d}, \tilde{m}_{n}\right) \xrightarrow{\mathbb{D}}\left([0,1]^{k}, \mathrm{~d}, m\right)
$$

as $n \rightarrow \infty$.


Example 3.11. ('Increasing to Fractal Dimension')
Let $\left(M_{n}\right)_{n \in \mathbb{N}}$ be the usual approximation of the Sierpinski gasket $M \subset \mathbb{R}^{2}$ by graphs $M_{n}$ with $3^{n}$ edges of sidelength $2^{1-n}, n \in \mathbb{N}$. To be more specific, $M_{1}$ is the equilateral triangle with

sidelength 1 and for each $n \in \mathbb{N}$, the graph $M_{n}$ is obtained from $M_{n-1}$ by gluing together 3 copies and rescaling the whole by the factor $\frac{1}{2}$. Let $\mathrm{d}_{n}$ be the distance from the ambient twodimensional Euclidean space (or alternatively the induced length distance on $M_{n}$ ) and let $m_{n}$ be normalized one-dimensional Lebesgue measure on $M_{n}$. Then

$$
\left(M_{n}, \mathrm{~d}_{n}, m_{n}\right) \xrightarrow{\mathbb{D}}(M, \mathrm{~d}, m)
$$

where $M$ is the Sierpinski gasket, d is the two-dimensional Euclidean distance restricted to $M$ (or the induced length distance on $M$, resp.) and $m$ is the normalized $\log 3 / \log 2$-dimensional Hausdorff measure on $M$.

Similarly, we can approximate the two-dimensional Sierpinski carpet $\tilde{M}$ (equipped with Euclidean distance $\tilde{\mathbf{d}}$ - or alternatively with the induced length distance - and with normalized $\log 8 / \log 3$-Hausdorff measure $\tilde{m})$ by graphs $\tilde{M}_{n}$ with sidelength $3^{-n}$. Here $\tilde{M}_{1}$ is the square with sidelength 1 and $\tilde{M}_{n}$ is obtained by gluing together 8 copies of $\tilde{M}_{n-1}$ and rescaling the whole by the factor $\frac{1}{3}$. Then

$$
\left(\tilde{M}_{n} \tilde{\mathrm{~d}}_{n}, \tilde{m}_{n}\right) \xrightarrow{\mathbb{D}}(\tilde{M}, \tilde{\mathrm{~d}}, \tilde{m})
$$

See for instance [Ki01].


Example 3.12. ('Decreasing Dimension, Collapse')
(i) For each metric measure space $(M, \mathrm{~d}, m)$ and each sequence $\left(M_{n}, \mathrm{~d}_{n}, m_{n}\right)$, $n \in \mathbb{N}$, with $\lim _{n \rightarrow \infty} \operatorname{Var}\left(M_{n}, \mathrm{~d}_{n}, m_{n}\right)=0$ one has

$$
\left(M \times M_{n}, \mathrm{~d} \otimes \mathrm{~d}_{n}, m \otimes m_{n}\right) \xrightarrow{\mathbb{D}}(M, \mathrm{~d}, m)
$$

as $n \rightarrow \infty$.
(ii) Let $M$ be a finite graph, embedded in $\mathbb{R}^{3}$, let d be the graph distance and $m$ be the 1-dimensional Lebesgue measure on $M$ normalized to 1 . Let

$$
M_{n}:=\left\{x \in \mathbb{R}^{3}: \mathrm{d}_{E u c l i d}(x, M) \leq \frac{1}{n}\right\}
$$

and

$$
\tilde{M}_{n}:=\left\{x \in \mathbb{R}^{3}: \mathrm{d}_{E u c l i d}(x, M)=\frac{1}{n}\right\}
$$

be the full (and surface, resp.) tubular neighborhood of $M$, let $\mathrm{d}_{n}$ (and $\tilde{\mathrm{d}}_{n}$ ) be the geodesic distance on $M_{n}$ (or $\tilde{M}_{n}$, resp.) induced by the Euclidean distance $\mathrm{d}_{\text {Euclid }}$ on the ambient space $\mathbb{R}_{\sim}^{3}$, and let $m_{n}\left(\right.$ and $\left.\tilde{m}_{n}\right)$ be the 3 - (or 2-, resp.) dimensional Lebesgue measure on $M$ (or $\tilde{M}$, resp.), normalized to 1 . Then

$$
\left(M_{n}, \mathrm{~d}_{n}, m_{n}\right) \xrightarrow{\mathbb{D}}(M, \mathrm{~d}, m)
$$

and

$$
\left(\tilde{M}_{n} \tilde{\mathrm{~d}}_{n}, \tilde{m}_{n}\right) \xrightarrow{\mathbb{D}}(M, \mathrm{~d}, m)
$$

as $n \rightarrow \infty$.


Being a length space is not preserved under isomorphisms of metric measure spaces. But the support $\operatorname{supp}[m]$ being a length space is preserved under isomorphisms. However, also this property is not preserved under $\mathbb{D}$-convergence.

Example 3.13. Let $M=\mathbb{R}$ with Euclidean distance $\mathrm{d}, d m_{n}(x)=\varphi_{n}(x) d x$ with

$$
\varphi_{n}(x)= \begin{cases}1 /(2 n), & x \in]-1,1[ \\ 1 / 2, & x \in\left[-2+\frac{1}{n},-1\right] \cup\left[1,2-\frac{1}{n}\right] \\ 0, & \text { else }\end{cases}
$$

and

$$
d m(x)=\frac{1}{2}\left(1_{[-2,-1]}(x)+1_{[1,2]}(x)\right) d x
$$

Then $\left(M, \mathrm{~d}, m_{n}\right) \xrightarrow{\mathbb{D}}(M, \mathrm{~d}, m)$ as $n \rightarrow \infty$ as well as $\left(\operatorname{supp}\left[m_{n}\right], \mathrm{d}, m_{n}\right) \xrightarrow{\mathbb{D}}(\operatorname{supp}[m], \mathrm{d}, m)$. However, $\left(\operatorname{supp}\left[m_{n}\right], d\right)$ as $n \rightarrow \infty$ does not converge w.r.t. $\mathrm{D}^{\mathrm{GH}}$ towards ( $\operatorname{supp}[m]$, d). It converges towards $([-2,2], \mathrm{d})$. On the other hand, Example 3.10 (i) demonstrates the opposite phenomenon: nonlength spaces converging to a length space.

### 3.3 Doubling Property under $\mathbb{D}$-Convergence

Definition 3.14. Given a number $C \in \mathbb{R}_{+}$, we say that a metric measure space ( $M, \mathrm{~d}, m$ ) has the restricted doubling property with doubling constant $C$ iff for all $x \in \operatorname{supp}[m]$ and all $r \in \mathbb{R}_{+}$

$$
m\left(B_{2 r}(x)\right) \leq C \cdot m\left(B_{r}(x)\right)
$$

A metric measure space $(M, \mathrm{~d}, m)$ has the restricted doubling property if and only if for all $x \in \operatorname{supp}[m]$ and all $r, R \in \mathbb{R}_{+}$

$$
m\left(B_{R}(x)\right) \leq C^{\left\lfloor\frac{\log R / r}{\log 2}+1\right\rfloor} \cdot m\left(B_{r}(x)\right)
$$

where $\lfloor a\rfloor$ denotes the greatest integer $\leq a$. It implies that for all $x$ and $r$ the sets $\bar{B}_{r}(x) \cap \operatorname{supp}[m\rfloor$ are compact.
Note that our definition differs from the usual definition of the doubling property: we only impose a condition on balls with center in the support of the measure. Some modification of this kind is necessary in order to obtain a property which is preserved under isomorphisms of metric measure spaces. The usual doubling property without this restriction implies that $\operatorname{supp}[m]=M$ whenever $m(M) \neq 0$.
For instance, let ( $M, \mathrm{~d}$ ) be the two-dimensional Euclidean plane and let $m$ be the one-dimensional Lebesgue measure on the $x_{1}$-axis. Then $(M, \mathrm{~d}, m)$ has the restricted doubling property but not the doubling property in the usual sense.
There is a huge literature on metric measure spaces which have the doubling property, see e.g. [He01] and references cited therein.

Theorem 3.15. The restricted doubling property is stable under $\mathbb{D}$-convergence.
That is, if for all $n \in \mathbb{N}$ the normalized metric measure spaces $\left(M_{n}, \mathrm{~d}_{n}, m_{n}\right)$ have the restricted doubling property with a common doubling constant $C$ and if $\left(M_{n}, \mathrm{~d}_{n}, m_{n}\right) \xrightarrow{\mathbb{D}}(M, \mathrm{~d}, m)$ as $n \rightarrow \infty$ then also $(M, \mathrm{~d}, m)$ has the restricted doubling property with the same constant $C$.

Proof. Assume that the normalized metric measure spaces $\left(M_{n}, \mathrm{~d}_{n}, m_{n}\right), n \in \mathbb{N}$, have the restricted doubling property with a common doubling constant $C$ and that

$$
\delta_{n}:=\mathbb{D}\left(\left(M_{n}, \mathrm{~d}_{n}, m_{n}\right),(M, \mathrm{~d}, m)\right) \rightarrow 0
$$

as $n \rightarrow \infty$. Then for each $n \in \mathbb{N}$ the spaces $(\operatorname{supp}[m], \mathrm{d})$ and $\left(\operatorname{supp}\left[m_{n}\right], \mathbf{d}_{n}\right)$ can isometrically be embedded into some space ( $\hat{M}, \hat{\mathrm{~d}}$ ) such that

$$
\hat{\mathrm{d}}_{W}\left(\hat{m}, \hat{m}_{n}\right) \leq 2 \delta_{n}
$$

where $\hat{m}$ and $\hat{m}_{n}$ denote the push forwards of the measures $m$ and $m_{n}$, resp., under the embedding maps $\psi$ and $\psi_{n}$, resp.
Let $x \in \operatorname{supp}[m], r>0, \epsilon>0$ and $\alpha<1$ be given (with $2 r \alpha^{2}-6 \epsilon>0$ for simplicity). Our first observation is that

$$
\begin{equation*}
\hat{m}_{n}\left(\hat{B}_{2 \epsilon}(\psi(x))\right) \geq \hat{m}\left(\hat{B}_{\epsilon}(\psi(x))\right)-\frac{1}{\epsilon^{2}} \hat{d}_{W}^{2}\left(\hat{m}, \hat{m}_{n}\right) \tag{3.8}
\end{equation*}
$$

since the mass which has to be transported from the interior of the small ball to the exterior of the large ball has to be moved by a distance of at least $\epsilon$. Moreover, we know that $\hat{m}\left(\hat{B}_{\epsilon}(\psi(x))\right)=$ $m\left(B_{\epsilon}(x)\right)>0$ since by assumption $x \in \operatorname{supp}[m]$. Hence, for $n$ large enough we conclude that $\hat{m}_{n}\left(\hat{B}_{2 \epsilon}(\psi(x))\right)>0$. Therefore, there exists a point $\hat{x}_{n} \in \operatorname{supp}\left[\hat{m}_{n}\right] \subset \hat{M}$ with $\hat{\mathrm{d}}\left(\hat{x}_{n}, \psi(x)\right) \leq 2 \epsilon$. In particular, we may apply the restricted doubling property for balls centered at $x_{n}$. This yields

$$
\begin{aligned}
m\left(B_{2 \alpha^{2} r-6 \epsilon}(x)\right) & =\hat{m}\left(\hat{B}_{2 \alpha^{2} r-6 \epsilon}(\psi(x))\right) \leq \hat{m}\left(\hat{B}_{2 \alpha^{2} r-4 \epsilon}\left(\hat{x}_{n}\right)\right) \\
& \leq \hat{m}_{n}\left(\hat{B}_{2 \alpha r-4 \epsilon}\left(\hat{x}_{n}\right)\right)+\frac{1}{\left(2 r \alpha-2 r \alpha^{2}\right)^{2}} \hat{\mathrm{~d}}_{W}^{2}\left(\hat{m}, \hat{m}_{n}\right) \\
& \leq C \cdot \hat{m}_{n}\left(\hat{B}_{\alpha r-2 \epsilon}\left(\hat{x}_{n}\right)\right)+\frac{1}{\left(2 r \alpha-2 r \alpha^{2}\right)^{2}} \hat{\mathrm{~d}}_{W}^{2}\left(\hat{m}, \hat{m}_{n}\right) \\
& \leq C \cdot \hat{m}_{n}\left(\hat{B}_{\alpha r}(\psi(x))\right)+\frac{1}{\left(2 r \alpha-2 r \alpha^{2}\right)^{2}} \hat{\mathrm{~d}}_{W}^{2}\left(\hat{m}, \hat{m}_{n}\right) \\
& \leq C \cdot \hat{m}\left(\hat{B}_{r}(\psi(x))\right)+\left[\frac{1}{\left(2 r \alpha-2 r \alpha^{2}\right)^{2}}+\frac{C}{(2 r-2 r \alpha)^{2}}\right] \hat{\mathrm{d}}_{W}^{2}\left(\hat{m}, \hat{m}_{n}\right) \\
& \leq C \cdot m\left(B_{r}(x)\right)+\frac{1+C \alpha^{2}}{[r \alpha(1-\alpha)]^{2}} \cdot \delta_{n}^{2} .
\end{aligned}
$$

In the limit $n \rightarrow \infty$ we obtain

$$
m\left(B_{2 \alpha^{2} r-6 \epsilon}(x)\right) \leq C \cdot m\left(B_{r}(x)\right)
$$

Since this holds for any $\alpha<1$ and any $\epsilon>0$ we conclude

$$
m\left(B_{2 r}(x)\right) \leq C \cdot m\left(B_{r}(x)\right) .
$$

We close this chapter with an important result on compactness under $\mathbb{D}$-convergence.

Theorem 3.16. ('Compactness')
For each pair $(C, R) \in \mathbb{R}_{+} \times \mathbb{R}_{+}$the family $\mathbb{X}_{1}(C, R)$ of all isomorphism classes of normalized metric measure spaces with ('restricted') doubling constant $\leq C$ and radius $\leq R$ is compact under $\mathbb{D}$-convergence.

Proof. Let $C$ and $R$ be given and consider a space $(M, \mathrm{~d}, m) \in \mathbb{X}_{1}(C, R)$. Without restriction, we may assume that $\operatorname{supp}[m]=M$. Then each ball $B_{\epsilon}(x) \subset M$ has volume

$$
m\left(B_{\epsilon}(x)\right) \geq\left(\frac{\epsilon}{2 R}\right)^{N}
$$

with 'doubling dimension' $N:=\log C / \log 2$. Hence, each family of disjoint balls of radius $\epsilon$ in $M$ contains at most $(2 R / \epsilon)^{N}$ elements which in turn implies that $M$ can be covered by $(2 R / \epsilon)^{N}$ balls of radius $2 \epsilon$. Firstly, this implies that $M$ is compact. Secondly, since this covering property holds true uniformly in $(M, \mathrm{~d}, m) \in \mathbb{X}_{1}(C, R)$, according to Gromov's compactness theorem ([Gr81a] or [BH99], Theorem 5.41) it implies compactness of the family $\mathbb{X}_{1}(C, R)$ under GromovHausdorff convergence. Due to the following Lemma 3.17 this in turn implies compactness under $\mathbb{D}$-convergence.

Lemma 3.17. Let $\left\{\left(M_{i}, \mathrm{~d}_{i}, m_{i}\right), i \in I\right\}$ be an arbitrary family of normalized compact metric measure spaces which is closed under $\mathbb{D}$-convergence. If the family $\mathrm{X}^{\prime}=\left\{\left(M_{i}, \mathrm{~d}_{i}\right): i \in I\right\}$ (of isometry classes) is compact w.r.t. Gromov-Hausdorff distance $\mathrm{D}^{\mathrm{GH}}$ then the family $\mathbb{X}^{\prime}=$ $\left\{\left(M_{i}, \mathrm{~d}_{i}, m_{i}\right): i \in I\right\}$ (of isomorphism classes) is compact w.r.t. $\mathbb{D}$.

Proof. Let a sequence $\left(\left(M_{n}, \mathrm{~d}_{n}, m_{n}\right)\right)_{n \in \mathbb{N}}$ in $\mathbb{X}^{\prime}$ be given. Then (by the assumption of the compactness of $\mathrm{X}^{\prime}$ ) there exists a subsequence $\left(\left(M_{n_{k}}, \mathrm{~d}_{n_{k}}, m_{n_{k}}\right)\right)_{k \in \mathbb{N}}$ and a compact metric space $(M, \mathrm{~d})$ such that $\left(M_{n_{k}}, \mathrm{~d}_{n_{k}}\right) \rightarrow(M, \mathrm{~d})$ in $\mathrm{D}^{\mathrm{GH}}$. Moreover, for an appropriately chosen subsequence one can achieve that $M$ as well as all $M_{n_{k}}, k \in \mathbb{N}$, are isometrically embedded in some other compact metric space ( $\hat{M}, \hat{\mathrm{~d}}$ ). Let us denote the embedded spaces and the associated push forward measures again by $\left(M_{n_{k}}, \mathrm{~d}_{n_{k}}\right)$ and $m_{n_{k}}$ (or $(M, \mathrm{~d})$ and $m$, resp.). Then by compactness of $\hat{M}$ there exists a subsequence, again denoted by $\left(m_{n_{k}}\right)_{k \in \mathbb{N}}$, converging to some $m \in \mathcal{P}_{2}(\hat{M})$. But then

$$
\mathbb{D}\left(\left(M_{n_{k}}, \mathrm{~d}_{n_{k}}, m_{n_{k}}\right),(M, \mathrm{~d}, m)\right) \leq \hat{\mathrm{d}}_{W}\left(m_{n_{k}}, m\right) \rightarrow 0
$$

as $k \rightarrow \infty$. This proves the compactness of $\mathbb{X}^{\prime}$.

## 4 Curvature Bounds for Metric Measure Spaces

### 4.1 The Relative Entropy

Recall that a metric measure space always means a triple $(M, \mathrm{~d}, m)$ where $(M, \mathrm{~d})$ is a complete separable metric space and $m$ is a locally finite measure on $M$ equipped with its Borel $\sigma$-algebra. To avoid pathologies, in the sequel we always exclude the case $m(M)=0$.
Given a metric measure space $(M, \mathrm{~d}, m)$ we denote by $\mathcal{P}_{2}(M, \mathrm{~d}, m)$ the subspace of all $\nu \in$ $\mathcal{P}_{2}(M, \mathrm{~d})$ which are absolutely continuous w.r.t. $m$, that is, which can be written as $\nu=\rho \cdot m$ with Radon-Nikodym density $\rho$. In other words, $\mathcal{P}_{2}(M, \mathrm{~d}, m)$ can be identified with the set of all $m$-equivalence classes of nonnegative Borel-measurable functions $\rho: M \rightarrow \mathbb{R}$ satisfying $\int \rho(x) d m(x)=1$ and $\int x^{2} \rho(x) d m(x)<\infty$.
For $\nu=\rho \cdot m \in \mathcal{P}_{2}(M, \mathrm{~d}, m)$ we define the relative entropy of $\nu$ w.r.t. $m$ by

$$
\begin{equation*}
\operatorname{Ent}(\nu \mid m):=\lim _{\epsilon \searrow 0} \int_{\{\rho>\epsilon\}} \rho \log \rho d m . \tag{4.1}
\end{equation*}
$$

This coincides with

$$
\int_{\{\rho>0\}} \rho \log \rho d m
$$

provided $\int_{\{\rho>1\}} \rho \log \rho d m<\infty$. Otherwise $\operatorname{Ent}(\nu \mid m):=+\infty$. For $\nu \in \mathcal{P}_{2}(M, \mathrm{~d}) \backslash \mathcal{P}_{2}(M, \mathrm{~d}, m)$ we also define $\operatorname{Ent}(\nu \mid m):=+\infty$. Finally, we put

$$
\mathcal{P}_{2}^{*}(M, \mathrm{~d}, m):=\left\{\nu \in \mathcal{P}_{2}(M, \mathrm{~d}): \operatorname{Ent}(\nu \mid m)<\infty\right\} .
$$

Lemma 4.1. If $m$ has finite mass, then the relative entropy $\operatorname{Ent}(\cdot \mid m)$ is lower semicontinuous and $\neq-\infty$ on $\mathcal{P}_{2}(M, \mathrm{~d})$. More precisely, for all $\nu \in \mathcal{P}_{2}(M, \mathrm{~d})$

$$
\begin{equation*}
\operatorname{Ent}(\nu \mid m) \geq-\log m(M) \tag{4.2}
\end{equation*}
$$

Proof. The lower estimate for the relative entropy is a simple application of Jensen's inequality. The lower semicontinuity is more subtle. For $N>1$ define $U_{N}(r)=-N r^{1-1 / N}$. Then $r \rightarrow U_{N}(r)$ is convex on $\mathbb{R}_{+}$, hence,

$$
\begin{equation*}
U_{N}(r) \geq U_{N}\left(r_{0}\right)-(N-1) r_{0}^{-1 / N}\left(r-r_{0}\right) \tag{4.3}
\end{equation*}
$$

for all $r, r_{0}$. Moreover,

$$
\lim _{N \rightarrow \infty}\left[N r+U_{N}(r)\right]=\sup _{N}\left[N r+U_{N}(r)\right]=r \log r .
$$

Now consider $S_{N}: \mathcal{P}_{2}(M, \mathrm{~d}) \rightarrow \mathbb{R}$ with

$$
S_{N}(\nu):=\int U_{N}(\rho) d m+N
$$

for $\nu=\nu_{0}+\nu_{*} \in \mathcal{P}_{2}(M, \mathrm{~d})$ with $\nu_{*} \perp m$ and $\nu_{0}=\rho m$. Note that

$$
N-N \nu_{0}(M)^{1-1 / N} m(M)^{1 / N} \leq S_{N}(\nu) \leq N
$$

for all $\nu$. Therefore, for all $\nu \in \mathcal{P}_{2}(M, \mathrm{~d})$

$$
\begin{equation*}
\operatorname{Ent}(\nu \mid m)=\lim _{N \rightarrow \infty} S_{N}(\nu)=\sup _{N} S_{N}(\nu) \tag{4.4}
\end{equation*}
$$

Hence, lower semicontinuity of $\operatorname{Ent}(\cdot \mid m)$ will follow from lower semicontinuity of $S_{N}$. In order to verify the latter, let $\nu_{n}=\rho_{n} m+\nu_{n}^{*}$ be any sequence in $\mathcal{P}_{2}(M, \mathrm{~d})$ which converges to some $\nu=\rho m+\nu^{*} \in \mathcal{P}_{2}(M, \mathrm{~d})$. It implies that $\nu_{n} \rightarrow \nu$ weakly in the sense of measures. Since $\rho$ is a subprobability density it lies in $L_{1-1 / N}(M, m)$ and it can be approximated in the metric $D_{1-1 / N}$ by nonnegative bounded, continuous $\rho^{(i)} \in L_{1-1 / N}(M, m)$. Here $D_{1-1 / N}(u, v):=\int|u-v|^{1-1 / N} d m$. Put $\rho_{n}^{(i)}:=\left|\rho_{n}-\rho+\rho^{(i)}\right|$. Then $\left|S_{N}\left(\rho^{(i)} m\right)-S_{N}(\rho m)\right| \leq N \cdot D_{1-1 / N}\left(\rho^{(i)}, \rho\right) \rightarrow 0$ as well as $\left|S_{N}\left(\rho_{n}^{(i)} m\right)-S_{N}\left(\rho_{n} m\right)\right| \leq N \cdot D_{1-1 / N}\left(\rho^{(i)}, \rho\right) \rightarrow 0$ for $i \rightarrow \infty$, uniformly in $n$. Moreover, since $\rho^{(i)}$ is continuous and bounded it follows from (4.3) that

$$
S_{N}\left(\rho_{n}^{(i)} m\right)-S_{N}\left(\rho^{(i)} m\right) \geq-(N-1) \int\left[\rho^{(i)}\right]^{1-1 / N}\left(\rho_{n}^{(i)}-\rho^{(i)}\right) d m \rightarrow 0
$$

for $n \rightarrow \infty$. Summing up, we obtain

$$
\liminf _{n \rightarrow \infty} S_{N}\left(\rho_{n} m\right) \geq S_{N}(\rho m)
$$

and thus finally as $N \rightarrow \infty$

$$
\liminf _{n \rightarrow \infty} \operatorname{Ent}\left(\rho_{n} m \mid m\right) \geq \operatorname{Ent}(\rho m \mid m)
$$

Remark 4.2. The relative entropy can in the same manner also be defined for finite, nonnormalized measures $\nu$ (and $m$ ) on $M$. Then for all $\alpha, \beta>0$

$$
\begin{equation*}
\operatorname{Ent}(\alpha \nu \mid \beta m)=\alpha \operatorname{Ent}(\nu \mid m)+(\log \alpha-\log \beta) \alpha \nu(M) \tag{4.5}
\end{equation*}
$$

Moreover, for all finite or countable sets $I$ and all finite measures $\nu_{i}, i \in I$,

$$
\begin{equation*}
\operatorname{Ent}\left(\sum_{i \in I} \nu_{i} \mid m\right) \geq \sum_{i \in I} \operatorname{Ent}\left(\nu_{i} \mid m\right) \tag{4.6}
\end{equation*}
$$

with equality if and only if the $\nu_{i}, i \in I$, are mutually singular, and

$$
\begin{equation*}
\operatorname{Ent}\left(\sum_{i \in I} \nu_{i} \mid m\right) \leq \sum_{i \in I} \operatorname{Ent}\left(\nu_{i} \mid m\right)-\sum_{i \in I} \nu_{i}(M) \log \nu_{i}(M) . \tag{4.7}
\end{equation*}
$$

For convex combinations of probability measures $\nu_{i}$ inequalities (4.6) and (4.7) read as follows

$$
\begin{equation*}
\sum \alpha_{i} \operatorname{Ent}\left(\nu_{i} \mid m\right)+\sum \alpha_{i} \log \alpha_{i} \leq \operatorname{Ent}\left(\sum \alpha_{i} \nu_{i} \mid m\right) \leq \sum \alpha_{i} \operatorname{Ent}\left(\nu_{i} \mid m\right) \tag{4.8}
\end{equation*}
$$

Proof. Indeed,

$$
\begin{aligned}
\operatorname{Ent}\left(\sum_{i \in I} \nu_{i} \mid m\right) & =\sum_{i \in I} \int \rho_{i} \log \left(\sum_{k \in I} \rho_{k}\right) d m \\
& \stackrel{(*)}{\geq} \sum_{i \in I} \int \rho_{i} \log \rho_{i} d m=\sum_{i} \operatorname{Ent}\left(\nu_{i} \mid m\right)
\end{aligned}
$$

with equality in $\left(^{*}\right)$ if and only if $\rho_{k} \rho_{i}=0 m$-a.e. on $M$ for all $k \neq i$. On the other hand, according to Jensen's inequality (applied with the convex function $\varphi(r)=r \log r$ ):

$$
\begin{aligned}
\operatorname{Ent}\left(\sum \nu_{i} \mid m\right) & =\int\left(\sum \alpha_{i} \bar{\rho}_{i}\right) \log \left(\sum \alpha_{i} \bar{\rho}_{i}\right) d m \\
& \leq \int \sum \alpha_{i} \bar{\rho}_{i} \log \bar{\rho}_{i} d m \\
& =\sum \int \rho_{i} \log \rho_{i} d m-\sum \alpha_{i} \log \alpha_{i}
\end{aligned}
$$

with $\alpha_{i}=\nu_{i}(M), \bar{\rho}_{i}=\frac{1}{\alpha_{i}} \rho_{i}$ and $\rho_{i}=\frac{d \nu_{i}}{d m}$.
Remark 4.3. (i) If $m$ has infinite mass then $\operatorname{Ent}(. \mid m)$ may exhibit strange behavior. In particular, it can attain the value $-\infty$ and also lower semicontinuity may fail. See the example below.
(ii) If $m$ is finite on all balls and if $\operatorname{Ent}(\nu \mid m)<\infty$ then

$$
\begin{equation*}
\operatorname{Ent}(\nu \mid m)=\lim _{R \rightarrow \infty} \int_{B_{R}(o)} \rho \log \rho d m \tag{4.9}
\end{equation*}
$$

for each $\nu=\rho m$ (with any $o \in M$ ). Indeed, due to the finiteness of $m$ on $B_{R}(o)$ the integral on the RHS exists for all $R$ and as $R \rightarrow \infty$ by monotone convergence

$$
\int_{B_{R}(o) \cap\{\rho>1\}} \rho \log \rho d m \rightarrow \int_{\{\rho>1\}} \rho \log \rho d m<\infty
$$

whereas

$$
\int_{B_{R}(o) \cap\{\rho<1\}} \rho \log \rho d m \rightarrow \int_{\{\rho<1\}} \rho \log \rho d m \leq \infty .
$$

Example 4.4. Let $M=\mathbb{R}$ with Euclidean distance d and $d m(x)=\exp \left(\exp \left(x^{2}\right)\right) d x, d \mu_{\alpha}(x)=$ $\frac{1}{2 \alpha} \exp \left(\frac{-|x|}{\alpha}\right) d x$.
(i) Then for all $\alpha>0, \mu_{\alpha} \in \mathcal{P}_{2}(M)$ with $\mathrm{d}_{W}\left(\mu_{\alpha}, \delta_{0}\right)=\sqrt{2} \alpha$ and

$$
\operatorname{Ent}\left(\mu_{\alpha} \mid m\right)=-\infty
$$

(ii) For each $\eta \in \mathcal{P}_{2}(M, d)$ with $\operatorname{Ent}(\eta \mid m)>-\infty$ the relative entropy $\operatorname{Ent}(\cdot \mid m)$ is not lower semicontinuous at $\eta$ since

$$
\eta_{\alpha}:=(1-\alpha) \eta+\alpha \mu_{\alpha} \rightarrow \eta \quad \text { in } \quad\left(\mathcal{P}_{2}(M), \mathrm{d}_{W}\right) .
$$

as $\alpha \rightarrow 0$ and

$$
-\infty=\lim _{\alpha \rightarrow 0} \operatorname{Ent}\left(\eta_{\alpha} \mid m\right)<\operatorname{Ent}(\eta \mid m)
$$

(iii) Moreover, given any $\nu_{0}, \nu_{1} \in \mathcal{P}_{2}(M, \mathrm{~d})$ there exists a midpoint $\eta$ of them. If $\operatorname{Ent}(\eta \mid m)<\infty$ then for each $\epsilon>0$ there exists an $\alpha>0$ such that $\eta_{\alpha}$ (defined as before) is an $\epsilon$-midpoint of $\nu_{0}$ and $\nu_{1}$ and

$$
-\infty=\operatorname{Ent}\left(\eta_{\alpha} \mid m\right) \leq \frac{1}{2} \operatorname{Ent}\left(\nu_{0} \mid m\right)+\frac{1}{2} \operatorname{Ent}\left(\nu_{1} \mid m\right)-\frac{K}{8} \mathrm{~d}_{W}^{2}\left(\nu_{0}, \nu_{1}\right)
$$

for each $K \in \mathbb{R}$.

### 4.2 Curvature Bounds

## Definition 4.5.

(i) We say that a metric measure space $(M, \mathrm{~d}, m)$ has curvature $\geq K$ for some number $K \in \mathbb{R}$ iff the relative entropy $\operatorname{Ent}(\cdot \mid m)$ is weakly $K$-convex on $\mathcal{P}_{2}^{*}(M, \mathrm{~d}, m)$ in the following sense: for each pair $\nu_{0}, \nu_{1} \in \mathcal{P}_{2}^{*}(M, \mathrm{~d}, m)$ there exists a geodesic $\Gamma:[0,1] \rightarrow \mathcal{P}_{2}^{*}(M, \mathrm{~d}, m)$ connecting $\nu_{0}$ and $\nu_{1}$ with

$$
\begin{equation*}
\operatorname{Ent}(\Gamma(t) \mid m) \leq(1-t) \operatorname{Ent}(\Gamma(0) \mid m)+t \operatorname{Ent}(\Gamma(1) \mid m)-\frac{K}{2} t(1-t) \mathrm{d}_{W}^{2}(\Gamma(0), \Gamma(1)) \tag{4.10}
\end{equation*}
$$

for all $t \in[0,1]$. To be more specific, we say that in the previous case the metric measure space ( $M, \mathrm{~d}, m$ ) has globally curvature $\geq K$. Moreover, we put

$$
\underline{\mathbb{C u r v}}(M, \mathrm{~d}, m):=\sup \{K \in \mathbb{R}:(M, \mathrm{~d}, m) \text { has curvature } \geq K\}
$$

( with $\sup \emptyset:=-\infty$ as usual). Note that then $(M, \mathrm{~d}, m)$ has curvature $\geq \mathbb{C u r v}(M, \mathrm{~d}, m)$. Occasionally, we use slightly modified concepts:
(ii) We say that a metric measure space $(M, \mathrm{~d}, m$ ) has (globally) curvature $\geq K$ in the lax sense iff for each $\epsilon>0$ and for each pair $\nu_{0}, \nu_{1} \in \mathcal{P}_{2}^{*}(M, \mathrm{~d}, m)$ there exists an $\epsilon$-midpoint $\eta \in \mathcal{P}_{2}^{*}(M, \mathrm{~d}, m)$ of $\nu_{0}$ and $\nu_{1}$ with

$$
\begin{equation*}
\operatorname{Ent}(\eta \mid m) \leq \frac{1}{2} \operatorname{Ent}\left(\nu_{0} \mid m\right)+\frac{1}{2} \operatorname{Ent}\left(\nu_{1} \mid m\right)-\frac{K}{8} \mathrm{~d}_{W}^{2}\left(\nu_{0}, \nu_{1}\right)+\epsilon . \tag{4.11}
\end{equation*}
$$

We denote the maximal $K$ with this property by $\underline{\mathbb{C u r v}}_{l a x}(M, \mathrm{~d}, m)$.
(iii) We say that a metric measure space $(M, \mathrm{~d}, m)$ has locally curvature $\geq K$ if each point of $M$ has a neighborhood $M^{\prime}$ such that ( $M^{\prime}, \mathrm{d}, m$ ) - with d and $m$ also denoting the restrictions of $d$ and $m$ onto $M^{\prime}$ - has globally curvature $\geq K$. The maximal $K$ with this property will be denoted by $\mathbb{C u r v}_{l o c}(M, \mathrm{~d}, m)$.
Remark 4.6. Let $(M, \mathrm{~d}, m)$ be a metric measure space of finite mass.
(i) Then $\underline{\operatorname{Curv}}(M, \mathrm{~d}, m) \geq K$ if and only if for each pair $\nu_{0}, \nu_{1} \in \mathcal{P}_{2}^{*}(M, \mathrm{~d}, m)$ there exists a midpoint $\eta \in \mathcal{P}_{2}^{*}(M, \mathrm{~d}, m)$ of $\nu_{0}$ and $\nu_{1}$ with

$$
\begin{equation*}
\operatorname{Ent}(\eta \mid m) \leq \frac{1}{2} \operatorname{Ent}\left(\nu_{0} \mid m\right)+\frac{1}{2} \operatorname{Ent}\left(\nu_{1} \mid m\right)-\frac{K}{8} \mathrm{~d}_{W}^{2}\left(\nu_{0}, \nu_{1}\right) . \tag{4.12}
\end{equation*}
$$

(ii) Similarly, $\mathbb{C u r v}_{\text {lax }}(M, \mathrm{~d}, m) \geq K$ if and only if for all $\epsilon>0$ and all $\nu_{0}, \nu_{1} \in \mathcal{P}_{2}^{*}(M, \mathrm{~d}, m)$ there exists a curve $\Gamma:[0,1] \rightarrow \mathcal{P}_{2}^{*}(M, \mathrm{~d}, m)$ connecting $\nu_{0}$ and $\nu_{1}$ with

$$
\begin{equation*}
\operatorname{Length}(\Gamma) \leq \mathrm{d}_{W}\left(\nu_{0}, \nu_{1}\right)+\epsilon \tag{4.13}
\end{equation*}
$$

and

$$
\begin{equation*}
\operatorname{Ent}(\Gamma(t) \mid m) \leq(1-t) \operatorname{Ent}\left(\nu_{0} \mid m\right)+t \operatorname{Ent}\left(\nu_{1} \mid m\right)-\frac{K}{2} t(1-t) \mathrm{d}_{W}^{2}\left(\nu_{0}, \nu_{1}\right)+\epsilon \tag{4.14}
\end{equation*}
$$

for all $t \in[0,1]$.
(iii) $\mathbb{C u r v}_{l a x}(M, \mathrm{~d}, m)>-\infty$ implies that $\mathcal{P}_{2}^{*}(M, \mathrm{~d}, m)$ is a length space (with metric $\left.\mathrm{d}_{W}\right)$ and that $M_{0}=\operatorname{supp}[m] \subset M$ is a length space (with metric d).

Proof. (i), (ii) We have to prove that the existence of (approximate) midpoints with property (4.12) (or (4.11)) implies the existence of (approximate) geodesics with property (4.10) (or (4.14), resp.).
Given $\epsilon=0$ (or $\epsilon>0$, resp.) define $\Gamma\left(\frac{1}{2}\right)$ as $\epsilon$-midpoint of $\Gamma(0):=\nu_{0}$ and $\Gamma(1):=\nu_{1}$ satisfying (4.11). Then define $\Gamma\left(\frac{1}{4}\right)$ as $\epsilon / 2$-midpoint of $\Gamma(0)$ and $\Gamma\left(\frac{1}{2}\right)$ satisfying (4.11) with $\epsilon / 2$ and define $\Gamma\left(\frac{3}{4}\right)$ as $\epsilon / 2$-midpoint of $\Gamma\left(\frac{1}{2}\right)$ and $\Gamma(1)$ satisfying (4.11) with $\epsilon / 2$. By iteration we obtain $\Gamma(t)$ for all dyadic $t \in[0,1]$. The continuous extension yields the required curve. [See for instance [St03], proof of Proposition 2.3, for a similar argument.] Lower semicontinuity of the relative entropy then proves the claim for all $t \in[0,1]$.
(iii) According to part (ii), it only remains to prove that $M_{0}$ is a length space. Given $x_{0}, x_{1} \in M_{0}$ let $\nu_{i}$ for $i=0,1$ be the normalized volume in $B_{\epsilon}\left(x_{i}\right)$, i.e. $\nu_{i}=\frac{1}{m\left(B_{\epsilon}\left(x_{i}\right)\right)} 1_{B_{\epsilon}\left(x_{i}\right)} m$, with $\epsilon>0$ to be chosen later. (Note that $m\left(B_{\epsilon}\left(x_{i}\right)\right)>0$ for all $\epsilon>0$ since $x_{i} \in M_{0}$ and $m\left(B_{\epsilon}\left(x_{i}\right)\right)<\infty$ for all sufficiently small $\epsilon>0$ since $m$ is locally finite.) Then $\nu_{i} \in \mathcal{P}_{2}^{*}(M, \mathrm{~d}, m)$. Hence, there exists $\eta \in \mathcal{P}_{2}^{*}(M, \mathrm{~d}, m)$ with

$$
\mathrm{d}_{W}\left(\nu_{i}, \eta\right) \leq \frac{1}{2} \mathrm{~d}_{W}\left(\nu_{0}, \nu_{1}\right)+\epsilon
$$

for $i=0,1$. Therefore

$$
\begin{aligned}
& \int\left[\mathrm{d}^{2}\left(x_{0}, y\right)+\mathrm{d}^{2}\left(x_{1}, y\right)\right] d \eta(y)=\mathrm{d}_{W}^{2}\left(\delta_{x_{0}}, \eta\right)+\mathrm{d}_{W}^{2}\left(\delta_{x_{1}}, \eta\right) \\
& \leq\left[\mathrm{d}_{W}\left(\nu_{0}, \eta\right)+\epsilon\right]^{2}+\left[\mathrm{d}_{W}\left(\nu_{1}, \eta\right)+\epsilon\right]^{2} \leq 2\left[\frac{1}{2} \mathrm{~d}_{W}\left(\nu_{0}, \nu_{1}\right)+2 \epsilon\right]^{2} \\
& \leq 2\left[\frac{1}{2} \mathrm{~d}\left(x_{0}, x_{1}\right)+3 \epsilon\right]^{2}=\frac{1}{2} \mathrm{~d}^{2}\left(x_{0}, x_{1}\right)+\epsilon^{\prime}
\end{aligned}
$$

for arbitrarily small $\epsilon^{\prime}>0$. It implies that there exists a point $y \in \operatorname{supp}[\eta]$ with $\mathrm{d}^{2}\left(x_{0}, y\right)+$ $\mathrm{d}^{2}\left(x_{1}, y\right) \leq \frac{1}{2} \mathrm{~d}^{2}\left(x_{0}, x_{1}\right)+\epsilon^{\prime}$. In other words, $y$ is an approximate midpoint and thus $M_{0}$ is a length space.

Lemma 4.7. If $M$ is compact then curvature bounds in the usual sense and in the lax sense coincide:

$$
\underline{\operatorname{Curv}}(M, \mathrm{~d}, m)=\underline{\operatorname{Curv}}_{l a x}(M, \mathrm{~d}, m) .
$$

Proof. Given $\nu_{0}, \nu_{1} \in \mathcal{P}_{2}^{*}(M, \mathrm{~d}, m)$ let $\eta^{(i)}$ be a family of $\epsilon$-midpoints of $\nu_{0}$ and $\nu_{1}$ satisfying (4.11) with $\epsilon=1 / i$. Let us consider the family of probability measures $Q:=\left\{\eta^{(i)}: i \in \mathbb{N}\right\}$. This family is tight. Indeed, we may assume without restriction that $M$ is a compact length space [otherwise, replace $M$ by $M_{0}=\operatorname{supp}[m$ ], see Remark 4.6(iii)]. Hence, there exists a suitable subsequence $\left(\eta^{\left(i_{j}\right)}\right)_{j \in \mathbb{N}}$ which converges to some $\eta \in \mathcal{P}_{2}(M, \mathrm{~d})$. Continuity of the distance $\mathrm{d}_{W}$ and lower semicontinuity of the relative entropy $\operatorname{Ent}(\cdot \mid m)$ imply that $\eta$ is a midpoint of $\nu_{0}$ and $\nu_{1}$ and (4.11) holds with $\epsilon=0$. Iterating this procedure yields a geodesic connecting $\nu_{0}$ and $\nu_{1}$ and satisfying (4.10).

The usual definition of $K$-convexity for the relative entropy would require that (4.10) holds for each geodesic connecting $\nu_{0}$ and $\nu_{1}$. This leads to the following definition which, however, will be not used in this paper.
We say that a metric measure space $(M, \mathrm{~d}, m)$ has (globally) curvature $\geq K$ in the restricted sense iff $\mathcal{P}_{2}^{*}(M, \mathrm{~d}, m)$ is a geodesic space and iff each geodesic $\Gamma$ in $\mathcal{P}_{2}^{*}(M, \mathrm{~d}, m)$ satisfies (4.10).

Remark 4.8. Assume that $M$ is a compact nonbranching geodesic space where each pair of points in $M$ is connected by a unique geodesic which depends continuously on the endpoints. Then curvature bounds in the restricted sense and curvature bounds in the usual sense coincide.

Proof. The (uniformly) continuous dependence of the geodesics on the endpoints implies (and actually is equivalent to the fact) that for each $\epsilon>0$ there exists $\delta>0$ such that the midpoint $z^{\prime}$ of $x^{\prime} \in B_{\delta}(x)$ and $y^{\prime} \in B_{\delta}(y)$ lies in $B_{\epsilon}(z)$ whenever $z$ is the midpoint of $x$ and $y$. Now let the probability measures $q$ on $M \times M$ and $\eta$ on $M$ be given which are an optimal coupling and a midpoint, resp., of some $\nu_{0}$ and $\nu_{1}$. Decompose $q$ into a sum $q=\sum_{i \in \mathbb{N}} q_{i}$ of mutually singular $q_{i}, i \in \mathbb{N}$, with $\operatorname{supp}\left[q_{i}\right] \subset B_{\delta}\left(x_{i}\right) \times B_{\delta}\left(y_{i}\right)$ for suitable $x_{i}, y_{i} \in M, i \in \mathbb{N}$. Let $\nu_{0, i}$ and $\nu_{1, i}$ denote the marginals of $q_{i}$. Assuming that $(M, \mathrm{~d}, m)$ has curvature $\geq K$ in the usual sense then implies that for each $i \in \mathbb{N}$ there exists a midpoint $\tilde{\eta}_{i}$ of $\nu_{0, i}$ and $\nu_{1, i}$ satisfying

$$
\operatorname{Ent}\left(\tilde{\eta}_{i} \mid m\right) \leq \frac{1}{2} \operatorname{Ent}\left(\nu_{0, i} \mid m\right)+\frac{1}{2} \operatorname{Ent}\left(\nu_{1, i} \mid m\right)-\frac{K}{8} \mathrm{~d}_{W}^{2}\left(\nu_{0, i}, \nu_{1, i}\right)
$$

The $\tilde{\eta}_{i}$ for $i \in \mathbb{N}$ are mutually singular since $M$ is nonbranching and since the $q_{i}$ are mutually singular (Lemma 2.11 (iii)). Hence, $\tilde{\eta}=\sum \tilde{\eta}_{i}$ satisfies

$$
\operatorname{Ent}(\tilde{\eta} \mid m) \leq \frac{1}{2} \operatorname{Ent}\left(\nu_{0} \mid m\right)+\frac{1}{2} \operatorname{Ent}\left(\nu_{1} \mid m\right)-\frac{K}{8} \mathrm{~d}_{W}^{2}\left(\nu_{0}, \nu_{1}\right)
$$

Moreover, $\mathrm{d}_{W}(\eta, \tilde{\eta}) \leq 2 \epsilon$ since for each $i \in \mathbb{N}, \operatorname{supp}\left[\eta_{i}\right] \subset B_{\epsilon}\left(z_{i}\right)$ as well as $\operatorname{supp}\left[\tilde{\eta}_{i}\right] \subset B_{\epsilon}\left(z_{i}\right)$ with $z_{i}$ being the midpoint of $x_{i}$ and $y_{i}$. By lower semicontinuity of $\operatorname{Ent}(\cdot \mid m)$ this implies

$$
\operatorname{Ent}(\eta \mid m) \leq \frac{1}{2} \operatorname{Ent}\left(\nu_{0} \mid m\right)+\frac{1}{2} \operatorname{Ent}\left(\nu_{1} \mid m\right)-\frac{K}{8} \mathrm{~d}_{W}^{2}\left(\nu_{0}, \nu_{1}\right)
$$

Theorem 4.9. ('Riemannian Spaces')
Let $M$ be a complete Riemannian manifold with Riemannian distance d and Riemannian volume $m$ and put $m^{\prime}=\exp (-V) \cdot m$ with a $\mathcal{C}^{2}$ function $V: M \rightarrow \mathbb{R}$. Then

$$
\begin{equation*}
\underline{\operatorname{Curv}}\left(M, \mathrm{~d}, m^{\prime}\right)=\inf \left\{\operatorname{Ric}_{M}(\xi, \xi)+\operatorname{Hess} V(\xi, \xi): \xi \in T M,|\xi|=1\right\} \tag{4.15}
\end{equation*}
$$

In particular, $(M, \mathrm{~d}, m)$ has curvature $\geq K$ if and only if the Ricci curvature of $M$ is $\geq K$.
Note that in the above Riemannian setting for each pair of points $\nu_{0}, \nu_{1}$ in $\mathcal{P}_{2}(M, \mathrm{~d}, m)$ there exists a unique geodesic connecting them. Hence, curvature bounds in the usual sense coincide with curvature bounds in the restricted sense. Moreover, note that in this setting, local curvature bounds always coincide with global curvature bounds.

Proof. Let us briefly sketch the main ideas of the proof, ignoring smoothness and regularity questions. For details, see $[\mathrm{RS} 04]$ for the case $V=0$ or [St04] for the general case.
Let $\nu_{0}=\rho_{0} m$ and $\nu_{1}=\rho_{1} m$ be given. According to Remark 2.12 (iii) there exists a function $\varphi: M \rightarrow \mathbb{R}$ such that

$$
\nu_{t}=\left(F_{t}\right)_{*} \nu_{0}
$$

with

$$
F_{t}(x)=\exp _{x}(-t \nabla \varphi(x))
$$

defines the unique geodesic $t \mapsto \nu_{t}$ in $\mathcal{P}_{2}(M, \mathrm{~d})$ connecting $\nu_{0}$ and $\nu_{1}$. Change of variable formula then gives

$$
\begin{equation*}
\operatorname{Ent}\left(\nu_{t} \mid e^{-V} m\right)=\int \rho_{0} \log \rho_{0} d m-\int y_{t} \rho_{0} d m+\int V\left(F_{t}\right) \rho_{0} d m \tag{4.16}
\end{equation*}
$$

with $y_{t}=\log \operatorname{det} d F_{t}$ being the logarithm of the determinant of the Jacobian of $F_{t}$ (in some weak sense). Now for $\nu_{0}$-a.e. $x \in M$ the function $t \mapsto y_{t}(x)$ satisfies the differential inequality

$$
\begin{equation*}
\ddot{y}_{t}(x) \leq-\frac{1}{n}\left(\dot{y}_{t}\right)^{2}(x)-\operatorname{Ric}\left(\dot{F}_{t}(x), \dot{F}_{t}(x)\right) \tag{4.17}
\end{equation*}
$$

Together with (4.16) this yields

$$
\begin{aligned}
\frac{\partial^{2}}{\partial t^{2}} \operatorname{Ent}\left(\nu_{t} \mid e^{-V} m\right) & \geq \int\left[\operatorname{Ric}\left(\dot{F}_{t}, \dot{F}_{t}\right)+\operatorname{Hess} V\left(\dot{F}_{t}, \dot{F}_{t}\right)\right] \rho_{0} d x \\
& \geq K \cdot \mathrm{~d}_{W}^{2}\left(\nu_{0}, \nu_{1}\right)
\end{aligned}
$$

provided $\operatorname{Ric}(\xi, \xi)+\operatorname{Hess} V(\xi, \xi) \geq K|\xi|^{2}$ for all $\xi \in T M$. This 'proves' the $K$-convexity of $\operatorname{Ent}\left(\cdot \mid e^{-V} m\right)$.

Some of the most simple examples are
Example 4.10. (i) If $M$ is a n-dimensional Riemannian manifold of constant sectional curvature $\kappa$ then

$$
\underline{\mathbb{C u r v}}(M, \mathrm{~d}, m)=(n-1) \kappa
$$

(ii) If $M$ is the Euclidean space $\mathbb{R}^{n}$ with the weighted measure $d m(x)=\exp \left(-K\|x\|^{2} / 2\right) d x$ then

$$
\underline{\mathbb{C u r v}}(M, \mathrm{~d}, m)=K .
$$

(iii) If $\operatorname{supp}[m]$ consists of one point then $\underline{\mathbb{C u r v}}(M, \mathrm{~d}, m)=+\infty$.

Remark 4.11. If $m$ is finite on all balls then it suffices to verify (4.11) for all $\nu_{i}=\rho_{i} m$ with bounded density $\rho_{i}$ and with bounded support $\operatorname{supp}\left[\nu_{i}\right], i=0,1$.

Proof. Let $\nu_{0}, \nu_{1} \in \mathcal{P}_{2}^{*}(M, \mathrm{~d}, m)$ be given, say $\nu_{0}=\rho_{0} m$ and $\nu_{1}=\rho_{1} m$. Fix $o \in M$ and define for $i=0,1$

$$
\rho_{i, R}=\frac{1}{\alpha_{i, R}} 1_{B_{R}(o)}\left[\rho_{i} \wedge R\right] \quad \text { with } \quad \alpha_{i, R}=\int_{B_{R}(o)}\left(\rho_{i} \wedge R\right) d m .
$$

Then according to Remark 4.3, $\alpha_{i, R} \rightarrow 1$ and $\operatorname{Ent}\left(\rho_{i, R} m \mid m\right) \rightarrow \operatorname{Ent}\left(\nu_{i} \mid m\right)$ as $R \rightarrow \infty$. Moreover, $\mathrm{d}_{W}\left(\rho_{i, R} m, \nu_{i}\right) \rightarrow 0$ and thus for sufficiently large $R$, each $\frac{\epsilon}{2}$-midpoint of $\rho_{0, R} m$ and $\rho_{1, R} m$ will be an $\epsilon$-midpoint of $\nu_{0}$ and $\nu_{1}$.

### 4.3 Basic Transformations

Proposition 4.12. ('Isomorphism')
If $(M, \mathrm{~d}, m)$ and $\left(M^{\prime}, \mathrm{d}^{\prime}, m^{\prime}\right)$ are isomorphic metric measure spaces then

$$
\underline{\operatorname{Curv}}(M, \mathrm{~d}, m)=\underline{\operatorname{Curv}}\left(M^{\prime}, \mathrm{d}^{\prime}, m^{\prime}\right) .
$$

Thus $\mathbb{C u r v}(\cdot)$ extends to a function on $\mathbb{X}$, the family of isomorphism classes of metric measure spaces.

Analogous statements hold true for $\mathbb{C u r v}_{l a x}(\cdot)$ and $\mathbb{C u r v}_{l o c}(\cdot)$.
Proof. Let $\Psi: M_{0} \rightarrow M_{0}^{\prime}$ be an isometry between $M_{0}:=\operatorname{supp}[m]$ and $M_{0}^{\prime}:=\operatorname{supp}\left[m^{\prime}\right]$ such that $\Psi_{*} m=m^{\prime}$. Then for all $\nu=\rho m \in \mathcal{P}_{2}(M, \mathrm{~d}, m)$ the push forward measure $\Psi_{*} \nu$ is absolutely continuous w.r.t. $m^{\prime}$ with density $\rho\left(\Psi^{-1}\right)$. Indeed, for all bounded measurable $f: M \rightarrow \mathbb{R}$ :

$$
\int f d\left(\Psi_{*} \nu\right)=\int f(\Psi) \rho d m=\int\left[f \cdot \rho\left(\Psi^{-1}\right)\right] \circ \Psi d m=\int f \cdot \rho\left(\Psi^{-1}\right) d m^{\prime}
$$

and for all $\nu_{0}, \nu_{1} \in \mathcal{P}_{2}(M, \mathrm{~d})$ :

$$
\mathrm{d}_{W}^{\prime}\left(\Psi_{*} \nu_{0}, \Psi_{*} \nu_{1}\right)=\mathrm{d}_{W}\left(\nu_{0}, \nu_{1}\right)
$$

This is, $\Psi$ induces an isometry $\nu \mapsto \Psi_{*} \nu$ between $\mathcal{P}_{2}(M, \mathrm{~d})$ and $\mathcal{P}_{2}\left(M^{\prime}, \mathrm{d}^{\prime}\right)$ which maps $\mathcal{P}_{2}(M, \mathrm{~d}, m)$ onto $\mathcal{P}_{2}\left(M^{\prime}, \mathrm{d}^{\prime}, m^{\prime}\right)$. Moreover,

$$
\operatorname{Ent}\left(\Psi_{*} \nu \mid m^{\prime}\right)=\operatorname{Ent}(\nu \mid m)
$$

since

$$
\operatorname{Ent}\left(\Psi_{*} \nu \mid m^{\prime}\right)=\int \rho\left(\Psi^{-1}\right) \log \left(\rho\left(\Psi^{-1}\right)\right) d m^{\prime} \int \rho \log \rho d m=\operatorname{Ent}(\nu \mid m)
$$

This proves the claim.
Proposition 4.13. ('Scaled spaces')
For each metric measure space ( $M, \mathrm{~d}, m$ ) and all $\alpha, \beta>0$ :

$$
\underline{\operatorname{Curv}}(M, \alpha \mathrm{~d}, \beta m)=\alpha^{-2} \underline{\operatorname{Curv}}(M, \mathrm{~d}, m) .
$$

Analogous statements hold true for $\mathbb{C u r v}_{l a x}(\cdot)$ and $\mathbb{C u r v}_{l o c}(\cdot)$.
Proof. Obviously, $\operatorname{Ent}(\nu \mid \beta m)=\operatorname{Ent}(\nu \mid m)-\log \beta$ and $(\alpha \cdot \mathrm{d})_{W}\left(\nu_{0}, \nu_{1}\right)=\alpha \cdot \mathrm{d}_{W}\left(\nu_{0}, \nu_{1}\right)$.

Proposition 4.14. ('Weighted Spaces')
For each metric measure space $(M, \mathrm{~d}, m)$ and each lower bounded, measurable function $V: M \rightarrow$ $\mathbb{R}$

$$
\underline{\operatorname{Curv}}\left(M, \mathrm{~d}, e^{-V} m\right) \geq \underline{\operatorname{Curv}}(M, \mathrm{~d}, m)+\underline{\operatorname{Hess} V}
$$

where $\underline{\operatorname{Hess}} V:=\sup \{K \in \mathbb{R}: V$ is $K$-convex on $\operatorname{supp}[m]\}$. If $V$ is locally bounded from below then an analogous statement holds for Curv $_{l o c}(\cdot)$.

Recall that a function $V: M \rightarrow]-\infty,+\infty]$, defined on a geodesic space $M$, is called $K$-convex for some $K \in \mathbb{R}$ iff for each geodesic $\gamma:[0,1] \rightarrow M$ and for each $t \in[0,1]$

$$
\begin{equation*}
V(\gamma(t)) \leq(1-t) V(\gamma(0))+t V(\gamma(1))-\frac{K}{2} t(1-t) \mathrm{d}^{2}(\gamma(0), \gamma(1)) \tag{4.18}
\end{equation*}
$$

Proof. A simple calculation yields

$$
\operatorname{Ent}\left(\nu \mid e^{-V} m\right)=\operatorname{Ent}(\nu \mid m)+\int V d \nu
$$

Moreover, $\mathcal{P}_{2}^{*}\left(M, \mathrm{~d}, e^{-V} m\right) \subset \mathcal{P}_{2}^{*}(M, \mathrm{~d}, m)$ according to lower boundedness of $V$. Now put $K_{0}:=\underline{\mathbb{C u r v}}(M, \mathrm{~d}, m)$ and $K_{1}:=\underline{\text { Hess } V}$. Given any geodesic $\Gamma$ in $\mathcal{P}_{2}^{*}(M, \mathrm{~d}, m)$ and any $t \in[0,1]$, choose an optimal coupling $\hat{q}$ on $M^{3}$ with marginals $\Gamma(0), \Gamma(t), \Gamma(1)$ in the sense of Lemma 2.11(ii). Then

$$
\begin{aligned}
& \operatorname{Ent}\left(\Gamma(t) \mid e^{-V} m\right)-(1-t) \operatorname{Ent}\left(\Gamma(0) \mid e^{-V} m\right)-t \operatorname{Ent}\left(\Gamma(1) \mid e^{-V} m\right) \\
&= \operatorname{Ent}(\Gamma(t) \mid m)-(1-t) \operatorname{Ent}(\Gamma(0) \mid m)-t \operatorname{Ent}(\Gamma(1) \mid m) \\
&+\int_{M} V d \Gamma(t)-(1-t) \int_{M} V d \Gamma(0)-t \int_{M} V d \Gamma(1) \\
&= \operatorname{Ent}(\Gamma(t) \mid m)-(1-t) \operatorname{Ent}(\Gamma(0) \mid m)-t \operatorname{Ent}(\Gamma(1) \mid m) \\
&+\int_{M^{3}}\left[V\left(x_{t}\right)-(1-t) V\left(x_{0}\right)-t V\left(x_{1}\right)\right] d \hat{q}\left(x_{0}, x_{t}, x_{1}\right) \\
& \stackrel{(*)}{\leq}-\frac{K_{0}}{2} t(1-t) \mathrm{d}_{W}^{2}(\Gamma(0), \Gamma(1))-\int_{M^{3}} \frac{K_{1}}{2} t(1-t) \mathrm{d}^{2}\left(x_{0}, x_{1}\right) d \hat{q}\left(x_{0}, x_{t}, x_{1}\right) \\
&=-\frac{K_{0}+K_{1}}{2} t(1-t) \mathrm{d}_{W}^{2}(\Gamma(0), \Gamma(1)) .
\end{aligned}
$$

The inequality $\left(^{*}\right.$ ) follows from $K_{1}$-convexity of $V$ since for $\hat{q}$-almost every $\left(x_{0}, x_{t}, x_{1}\right)$ the point $x_{t}$ lies on a geodesic connecting $x_{0}$ and $x_{1}$ (Lemma 2.11(ii)).

Proposition 4.15. ('Subsets')
Let $(M, \mathrm{~d}, m)$ be a metric measure space and let $M^{\prime}$ be a convex subset of $M$. Then

$$
\underline{\mathbb{C u r v}}\left(M^{\prime}, \mathrm{d}, m\right) \geq \underline{\operatorname{Curv}}(M, \mathrm{~d}, m) .
$$

Proof. Let $\nu_{0}, \nu_{1}$ be probability measures in $\mathcal{P}_{2}^{*}\left(M^{\prime}, \mathrm{d}, m\right)$. Regard them as probability measures on $M$. Let $\Gamma$ be a geodesic in $\mathcal{P}_{2}^{*}(M, \mathrm{~d}, m)$ connecting them and satisfying (4.10). It remains to prove that of each $t \in[0,1]$ the measures $\Gamma(t)$ is supported by $M^{\prime}$, i.e.

$$
\Gamma(t)\left(M \backslash M^{\prime}\right)=0
$$

According to Lemma 2.11(ii), there exists an optimal coupling $\hat{q}$ of $\Gamma(0)=\nu_{0}, \Gamma(t)$, and $\Gamma(1)=\nu_{1}$ such that for $\hat{q}$-a.e. $(x, z, y) \in M^{3}$ the point $z$ lies on some geodesic connecting the points $x$ and $y$. But then $\hat{q}$-almost surely $z$ has to lie in $M^{\prime}$ since $x$ and $y$ lie in $M^{\prime}$ and the latter is assumed to be convex. This proves that $\Gamma(t)\left(M \backslash M^{\prime}\right)=0$ and thus yields the claim for $\mathbb{C u r v}\left(M^{\prime}, \mathrm{d}, m\right)$.

Proposition 4.16. ('Products')
Let $\left(M_{i}, \mathrm{~d}_{i}, m_{i}\right)$ for $i=1, \ldots, l$ be metric measure spaces and

$$
(M, \mathrm{~d}, m)=\bigotimes_{i=1}^{l}\left(M_{i}, \mathrm{~d}_{i}, m_{i}\right) .
$$

Assume that $M$ is nonbranching and compact. Then

$$
\begin{equation*}
\underline{\operatorname{Curv}}(M, \mathrm{~d}, m)=\inf _{i \in\{1, \ldots, l\}} \underline{\operatorname{Curv}}\left(M_{i}, \mathrm{~d}_{i}, m_{i}\right) . \tag{4.19}
\end{equation*}
$$

Proof. (i) Let us first prove the inequality

$$
\underline{\operatorname{Curv}}(M, \mathrm{~d}, m) \leq \inf _{i \in\{1, \ldots, l\}} \underline{\operatorname{Curv}}\left(M_{i}, \mathrm{~d}_{i}, m_{i}\right) .
$$

Assume that this is not true. Then for some $K \in \mathbb{R}$ and $i \in\{1, \ldots, l\}$

$$
\begin{equation*}
\underline{\mathbb{C u r v}}(M, \mathrm{~d}, m) \geq K>\underline{\mathbb{C u r v}}\left(M_{i}, \mathrm{~d}_{i}, m_{i}\right) . \tag{4.20}
\end{equation*}
$$

Without restriction, we may assume $i=1$. Then the last inequality implies that there exist $\nu_{0}^{(1)}, \nu_{1}^{(1)} \in \mathcal{P}_{2}^{*}\left(M_{1}, \mathrm{~d}_{1}, m_{1}\right)$ such that for each midpoint $\eta^{(1)}$ in $\mathcal{P}_{2}^{*}\left(M_{1}, \mathrm{~d}_{1}, m_{1}\right)$ between $\nu_{0}^{(1)}, \nu_{1}^{(1)}$ the inequality (4.12) is violated, i.e.

$$
\begin{equation*}
\operatorname{Ent}\left(\eta^{(1)} \mid m_{1}\right)>\frac{1}{2} \operatorname{Ent}\left(\nu_{0}^{(1)} \mid m_{1}\right)+\frac{1}{2} \operatorname{Ent}\left(\nu_{1}^{(1)} \mid m_{1}\right)-\frac{K}{8} \mathrm{~d}_{W}^{2}\left(\nu_{0}^{(1)}, \nu_{1}^{(1)}\right) . \tag{4.21}
\end{equation*}
$$

Now for $j=0,1$ put $\nu_{j}:=\nu_{j}^{(1)} \otimes \bar{m}_{2} \otimes \ldots \otimes \bar{m}_{l}$ with normalized measures $\bar{m}_{i}=\frac{1}{m_{i}\left(M_{i}\right)} m_{i}$ for $i=2, \ldots, l$. Then obviously

$$
\begin{equation*}
\operatorname{Ent}\left(\nu_{j} \mid m\right)=\operatorname{Ent}\left(\nu_{j}^{(1)}\right)-\sum_{i=2}^{l} \log m_{i}\left(M_{i}\right) . \tag{4.22}
\end{equation*}
$$

and $\nu_{j} \in \mathcal{P}_{2}^{*}(M, \mathrm{~d}, m)$. Moreover

$$
\mathrm{d}^{W}\left(\nu_{0}, \nu_{1}\right)=\mathrm{d}_{1}^{W}\left(\nu_{0}^{(1)}, \nu_{1}^{(1)}\right)
$$

(where for typographical reasons, we replace the lower index 'W' by an upper index, again denoting the $L_{2}$-Wasserstein distances derived from d or $\mathrm{d}_{1}$, resp.).
Now the first inequality in (4.20) implies that there exists a midpoint $\eta$ of $\nu_{0}$ and $\nu_{1}$ satisfying (4.12). According to Remark 2.2 it implies

$$
\mathrm{d}^{W}\left(\nu_{0}, \eta\right)^{2}+\mathrm{d}^{W}\left(\eta, \nu_{1}\right)^{2} \leq \frac{1}{2} \mathrm{~d}^{W}\left(\nu_{0}, \nu_{1}\right)^{2}
$$

which in turn implies

$$
\mathrm{d}_{1}^{W}\left(\nu_{0}^{(1)}, \eta^{(1)}\right)^{2}+\mathrm{d}_{1}^{W}\left(\eta^{(1)}, \nu_{1}^{(1)}\right)^{2} \leq \frac{1}{2} \mathrm{~d}_{1}^{W}\left(\nu_{0}^{(1)}, \nu_{1}^{(1)}\right)^{2} .
$$

Again according to Remark 2.2 this yields that $\eta^{(1)}$ is a midpoint of $\nu_{0}^{(1)}$ and $\nu_{1}^{(1)}$. But (4.21) and (4.22) imply that (4.12) is violated, - contradicting our previous assertion. Thus $\underline{\operatorname{Curv}}(M, \mathrm{~d}, m)<K$ which proves our first claim.
(ii) To prove the reverse implication, we first treat the particular case $\nu_{0}=\nu_{0}^{(1)} \otimes \ldots \otimes \nu_{0}^{(l)}$ and $\nu_{1}=\nu_{1}^{(1)} \otimes \ldots \otimes \nu_{1}^{(l)}$. Assume that $\underline{\operatorname{Curv}}\left(M_{i}, d_{i}, m_{i}\right) \geq K$ for each $i=1, \ldots, l$. Then for each $i$ there exists a midpoint $\eta^{(i)}$ of $\nu_{0}^{(i)}$ and $\nu_{1}^{(i)}$ with

$$
\operatorname{Ent}\left(\eta^{(i)} \mid m_{i}\right) \leq \frac{1}{2} \operatorname{Ent}\left(\nu_{0}^{(i)} \mid m\right)+\frac{1}{2} \operatorname{Ent}\left(\nu_{1}^{(i)} \mid m\right)-\frac{K}{8} \cdot \mathrm{~d}_{i}^{W}\left(\nu_{0}^{(i)}, \nu_{1}^{(i)}\right)^{2} .
$$

Put $\eta:=\eta^{(1)} \otimes \ldots \otimes \eta^{(l)}$. Then $\eta$ is a midpoint of $\nu_{0}$ and $\nu_{1}$ since

$$
\mathrm{d}^{W}\left(\eta, \nu_{0}\right)^{2}=\sum_{i=1}^{l} \mathrm{~d}_{i}^{W}\left(\eta^{(i)}, \nu_{0}^{(i)}\right)^{2} \leq \sum_{i=1}^{l}\left[\frac{1}{2} \mathrm{~d}_{i}^{W}\left(\nu_{0}^{(i)}, \nu_{1}^{(i)}\right)\right]^{2}=\left[\frac{1}{2} \mathrm{~d}^{W}\left(\nu_{0}, \nu_{1}\right)\right]^{2} .
$$

Moreover,

$$
\operatorname{Ent}\left(\nu_{0} \mid m\right)=\sum_{i=1}^{l} \operatorname{Ent}\left(\nu_{0}^{(i)}, m_{i}\right), \quad \operatorname{Ent}\left(\nu_{1} \mid m\right)=\sum_{i=1}^{l} \operatorname{Ent}\left(\nu_{1}^{(i)}, m_{i}\right)
$$

and

$$
\operatorname{Ent}(\eta \mid m)=\sum_{i=1}^{l} \operatorname{Ent}\left(\eta^{(i)}, m_{i}\right) .
$$

Hence,

$$
\begin{aligned}
\operatorname{Ent}(\eta \mid m) & \leq \sum_{i=1}^{l}\left[\frac{1}{2} \operatorname{Ent}\left(\nu_{0}^{(i)} \mid m\right)+\frac{1}{2} \operatorname{Ent}\left(\nu_{1}^{(i)} \mid m\right)-\frac{K}{8} \cdot \mathrm{~d}_{i}^{W}\left(\nu_{0}^{(i)}, \nu_{1}^{(i)}\right)^{2}\right] \\
& \leq \frac{1}{2} \operatorname{Ent}\left(\nu_{0} \mid m\right)+\frac{1}{2} \operatorname{Ent}\left(\nu_{1} \mid m\right)-\frac{K}{8} \mathrm{~d}^{W}\left(\nu_{0}, \nu_{1}\right)^{2} .
\end{aligned}
$$

This proves the claim in the particular case.
(iii) Now let arbitrary $\nu_{0}, \nu_{1} \in \mathcal{P}_{2}^{*}(M, \mathrm{~d}, m)$ and $\epsilon>0$ be given. Then there exist

$$
\tilde{\nu}_{0}=\frac{1}{n} \sum_{j=1}^{n} \nu_{0, j}
$$

with mutually singular product measures $\nu_{0, j}, j=1, \ldots, n$ and

$$
\tilde{\nu}_{1}=\frac{1}{n} \sum_{j=1}^{n} \nu_{1, j}
$$

with mutually singular product measures $\nu_{1, j}, j=1, \ldots, n$ such that

$$
\begin{gathered}
\operatorname{Ent}\left(\tilde{\nu}_{0} \mid m\right) \leq \operatorname{Ent}\left(\nu_{0} \mid m\right)+\epsilon, \quad \operatorname{Ent}\left(\tilde{\nu}_{1} \mid m\right) \leq \operatorname{Ent}\left(\nu_{1} \mid m\right)+\epsilon, \\
\mathrm{d}^{W}\left(\nu_{0}, \tilde{\nu}_{0}\right) \leq \epsilon, \quad \mathrm{d}^{W}\left(\nu_{1}, \tilde{\nu}_{1}\right) \leq \epsilon
\end{gathered}
$$

and

$$
\mathrm{d}^{W}\left(\tilde{\nu}_{0}, \tilde{\nu}_{1}\right) \geq\left[\frac{1}{n} \sum_{j=1}^{n} \mathrm{~d}^{W}\left(\nu_{0, j}, \nu_{1, j}\right)^{2}\right]^{1 / 2}-\epsilon .
$$

Furthermore, since $\nu_{0}$ is the sum of mutually singular $\nu_{0, j}$

$$
\operatorname{Ent}\left(\tilde{\nu}_{0} \mid m\right)=\frac{1}{n} \sum_{j=1}^{n} \operatorname{Ent}\left(\nu_{0, j} \mid m\right)-\log n
$$

and similarly,

$$
\operatorname{Ent}\left(\tilde{\nu}_{1} \mid m\right)=\frac{1}{n} \sum_{j=1}^{n} \operatorname{Ent}\left(\nu_{1, j} \mid m\right)-\log n .
$$

According to part (ii) for each $j=1, \ldots, n$ there exists a midpoint $\eta_{j}$ of $\nu_{0, j}$ and $\nu_{1, j}$ satisfying

$$
\operatorname{Ent}\left(\eta_{j} \mid m\right) \leq \frac{1}{2} \operatorname{Ent}\left(\nu_{0, j} \mid m\right)+\frac{1}{2} \operatorname{Ent}\left(\nu_{1, j} \mid m\right)-\frac{K}{8} \cdot \mathrm{~d}^{W}\left(\nu_{0, j}, \nu_{1, j}\right)^{2} .
$$

According to Lemma 2.11(iii), since $M$ is nonbranching and since the $\nu_{0, j}$ for $j=1, \ldots, n$ are mutually singular, also the $\eta_{j}$ for $j=1, \ldots, n$ must be mutually singular. Hence,

$$
\eta:=\frac{1}{n} \sum_{j=1}^{n} \eta_{j}
$$

satisfies

$$
\operatorname{Ent}(\eta \mid m)=\frac{1}{n} \sum_{j=1}^{n} \operatorname{Ent}\left(\eta_{j} \mid m\right)-\log n
$$

and thus

$$
\begin{aligned}
\operatorname{Ent}(\eta \mid m) & \leq \frac{1}{2} \operatorname{Ent}\left(\tilde{\nu}_{0} \mid m\right)+\frac{1}{2} \operatorname{Ent}\left(\tilde{\nu}_{1} \mid m\right)-\frac{K}{8} \frac{1}{n} \sum_{j=1}^{n} \mathrm{~d}^{W}\left(\nu_{0, j}, \nu_{1, j}\right)^{2} \\
& \leq \frac{1}{2} \operatorname{Ent}\left(\nu_{0} \mid m\right)+\frac{1}{2} \operatorname{Ent}\left(\nu_{1} \mid m\right)-\frac{K}{8}\left[\mathrm{~d}^{W}\left(\nu_{0}, \nu_{1}\right) \mp 3 \epsilon\right]^{2}+\epsilon
\end{aligned}
$$

(with $\mp$ to be chosen according to the sign of $K$ ). Moreover, $\eta$ is an approximate midpoint of $\nu_{0}$ and $\nu_{1}$ :

$$
\begin{aligned}
2 \mathrm{~d}^{W}\left(\nu_{0}, \eta\right) & \leq 2\left[\frac{1}{n} \sum_{j} \mathrm{~d}^{W}\left(\nu_{0, j}, \eta_{j}\right)^{2}\right]^{1 / 2} \leq\left[\frac{1}{n} \sum_{j} \mathrm{~d}^{W}\left(\nu_{0, j}, \nu_{1, j}\right)^{2}\right]^{1 / 2} \\
& \leq \mathrm{d}^{W}\left(\tilde{\nu}_{0}, \tilde{\nu}_{1}\right)+\epsilon \leq \mathrm{d}^{W}\left(\nu_{0}, \nu_{1}\right)+3 \epsilon
\end{aligned}
$$

and similarly for $\mathrm{d}^{W}\left(\nu_{1}, \eta\right)$. This proves that $\mathbb{C u r v}_{l a x}(M, \mathrm{~d}, m) \geq K$. Together with compactness of $M$ this finally yields the claim.

### 4.4 From Local to Global

A crucial implication of our definition of lower curvature bounds for metric measure spaces is the following Globalization Theorem which states that local curvature bounds imply global curvature bounds. This is in the spirit of the Globalization Theorem of Topogonov for lower curvature bounds (in the sense of Alexandrov) for metric spaces.

Theorem 4.17. ('Globalization')
Let $(M, \mathrm{~d}, m)$ be a compact, nonbranching metric measure space and assume that $\mathcal{P}_{2}^{*}(M, \mathrm{~d}, m)$ is a geodesic space. Then

$$
\begin{equation*}
\underline{\mathbb{C u r v}}(M, \mathrm{~d}, m)=\underline{\operatorname{Curv}}_{l o c}(M, \mathrm{~d}, m) . \tag{4.23}
\end{equation*}
$$

Proof. Put $K=\mathbb{C u r v}_{l o c}(M, \mathrm{~d}, m)$ and consider for each number $k \in \mathbb{N} \cup\{0\}$ the following property:
$\mathrm{C}(\mathrm{k})$ : For each geodesic $\Gamma:[0,1] \rightarrow \mathcal{P}_{2}^{*}(M, \mathrm{~d}, m)$ and for each pair $s, t \in[0,1]$ with $0 \leq t-s \leq$ $2^{-k}$ there exists a midpoint $\eta(s, t)$ of $\Gamma(s)$ and $\Gamma(t)$ such that

$$
\begin{equation*}
\operatorname{Ent}(\eta(s, t) \mid m) \leq \frac{1}{2} \operatorname{Ent}(\Gamma(s) \mid m)+\frac{1}{2} \operatorname{Ent}(\Gamma(t) \mid m)-\frac{K}{8} \mathrm{~d}_{W}^{2}(\Gamma(s), \Gamma(t)) \tag{4.24}
\end{equation*}
$$

Our first claim is that

- For each $k \in \mathbb{N}: \mathrm{C}(\mathrm{k})$ implies $\mathrm{C}(\mathrm{k}-1)$.

In order to prove this claim, let $k \in \mathbb{N}$ be given with property $\mathrm{C}(\mathrm{k})$. Moreover, let a geodesic $\Gamma$ and numbers $s, t \in[0,1]$ be given with $0 \leq t-s \leq 2^{1-k}$. Define iteratively a sequence $\left(\Gamma^{(i)}\right)_{i \in \mathbb{N}}$ of geodesics in $\mathcal{P}_{2}^{*}(M, \mathrm{~d}, m)$ which coincide with $\Gamma$ on $[0, s] \cup[t, 1]$ as follows:
start with $\Gamma^{(0)}:=\Gamma$; assuming that $\Gamma^{(2 i)}$ is already given, let $\Gamma^{(2 i+1)}:[0,1] \rightarrow \mathcal{P}_{2}^{*}(M, \mathrm{~d}, m)$ be any geodesic which coincides with $\Gamma$ on $[0, s] \cup[t, 1]$ and for which $\Gamma^{(2 i+1)}\left(s+\frac{t-s}{4}\right)$ is a midpoint of $\Gamma(s)=\Gamma^{(2 i)}(s)$ and $\Gamma^{(2 i)}\left(s+\frac{t-s}{2}\right)$ and for which $\Gamma^{(2 i+1)}\left(s+3 \frac{t-s}{4}\right)$ is a midpoint of $\Gamma^{(2 i)}\left(s+\frac{t-s}{2}\right)$ and $\Gamma(t)=\Gamma^{(2 i)}(t)$ satisfying

$$
\begin{align*}
\operatorname{Ent}\left(\left.\Gamma^{(2 i+1)}\left(s+\frac{t-s}{4}\right) \right\rvert\, m\right) & \leq \frac{1}{2} \operatorname{Ent}(\Gamma(s) \mid m) \\
& +\frac{1}{2} \operatorname{Ent}\left(\left.\Gamma^{(2 i)}\left(s+\frac{t-s}{2}\right) \right\rvert\, m\right)-\frac{K}{32} \mathrm{~d}_{W}^{2}(\Gamma(s), \Gamma(t)) \tag{4.25}
\end{align*}
$$

and

$$
\begin{align*}
\operatorname{Ent}\left(\left.\Gamma^{(2 i+1)}\left(s+3 \frac{t-s}{4}\right) \right\rvert\, m\right) & \leq \frac{1}{2} \operatorname{Ent}(\Gamma(t) \mid m) \\
& +\frac{1}{2} \operatorname{Ent}\left(\left.\Gamma^{(2 i)}\left(s+\frac{t-s}{2}\right) \right\rvert\, m\right)-\frac{K}{32} \mathrm{~d}_{W}^{2}(\Gamma(s), \Gamma(t)) \tag{4.26}
\end{align*}
$$

Such midpoints exist according to assumption $\mathrm{C}(\mathrm{k})$.
Then let $\Gamma^{(2 i+2)}:[0,1] \rightarrow \mathcal{P}_{2}^{*}(M, \mathrm{~d}, m)$ be any geodesic which coincides with $\Gamma$ on $[0, s] \cup[t, 1]$ and for which $\Gamma^{(2 i+2))}\left(s+\frac{t-s}{2}\right)$ is a midpoint of $\Gamma^{(2 i+1)}\left(s+\frac{t-s}{4}\right)$ and $\Gamma^{(2 i+1)}\left(s+3 \frac{t-s}{4}\right)$ satisfying

$$
\begin{align*}
\operatorname{Ent}\left(\left.\Gamma^{(2 i+2)}\left(s+\frac{t-s}{2}\right) \right\rvert\, m\right) & \leq \frac{1}{2} \operatorname{Ent}\left(\left.\Gamma^{(2 i+1))}\left(s+\frac{t-s}{4}\right) \right\rvert\, m\right) \\
& +\frac{1}{2} \operatorname{Ent}\left(\left.\Gamma^{(2 i+1)}\left(s+3 \frac{t-s}{4}\right) \right\rvert\, m\right)-\frac{K}{32} d_{W}^{2}(\Gamma(s), \Gamma(t)) \tag{4.27}
\end{align*}
$$

Again, such a midpoint exists according to assumption C(k). This yields a sequence of geodesics $\Gamma^{(i)}, i \in \mathbb{N}$.
Combining (4.25) - (4.27) gives

$$
\begin{aligned}
\operatorname{Ent}\left(\left.\Gamma^{(2 i+2)}\left(s+\frac{t-s}{2}\right) \right\rvert\, m\right) & \leq \frac{1}{2} \operatorname{Ent}\left(\left.\Gamma^{(2 i))}\left(s+\frac{t-s}{2}\right) \right\rvert\, m\right) \\
& +\frac{1}{4} \operatorname{Ent}(\Gamma(s) \mid m)+\frac{1}{4} \operatorname{Ent}(\Gamma(t) \mid m)-\frac{K}{16} \mathrm{~d}_{W}^{2}(\Gamma(s), \Gamma(t)) .
\end{aligned}
$$

By iteration, it yields

$$
\begin{aligned}
\operatorname{Ent}\left(\left.\Gamma^{(2 i)}\left(s+\frac{t-s}{2}\right) \right\rvert\, m\right) & \leq 2^{-i} \operatorname{Ent}\left(\left.\Gamma\left(s+\frac{t-s}{2}\right) \right\rvert\, m\right) \\
& +\left(1-2^{-i}\right)\left[\frac{1}{2} \operatorname{Ent}(\Gamma(s) \mid m)+\frac{1}{2} \operatorname{Ent}(\Gamma(t) \mid m)-\frac{K}{8} d_{W}^{2}(\Gamma(s), \Gamma(t))\right] .
\end{aligned}
$$

Compactness of $\mathcal{P}_{2}(M, \mathrm{~d})$ now implies that there exists a subsequence of $\left(\Gamma^{(2 i)}\left(s+\frac{t-s}{2}\right)\right)_{i \in \mathbb{N}}$ converging to some $\eta \in \mathcal{P}_{2}(M, \mathrm{~d})$. Continuity of the distance implies that $\eta$ is a midpoint of
$\Gamma(s)$ and $\Gamma(t)$ (since each of the $\Gamma^{(2 i)}\left(s+\frac{t-s}{2}\right)$ is a midpoint) and lower semicontinuity of the relative entropy implies

$$
\operatorname{Ent}(\eta \mid m) \leq \frac{1}{2} \operatorname{Ent}\left(\nu_{0} \mid m\right)+\frac{1}{2} \operatorname{Ent}\left(\nu_{1} \mid m\right)-\frac{K}{8} \mathrm{~d}^{2}\left(\nu_{0}, \nu_{1}\right) .
$$

This proves property $\mathrm{C}(\mathrm{k}-1)$.
Now according to our curvature assumption, each point $x \in M$ has a neighborhood $M(x)$ such that $(M(x), \mathrm{d}, m)$ has globally curvature $\geq K$ in the usual sense. By compactness of $M$, there exist $\lambda>0$, finitely many disjoint sets $L_{1}, \ldots, L_{n}$ which cover $M$, and closed sets $M_{j} \supset B_{\lambda}\left(L_{j}\right)$ for $j=1, \ldots, n$ such that ( $M_{j}, \mathrm{~d}, m$ ) has globally curvature $\geq K$ in the usual sense. Choose $k^{\prime} \in \mathbb{N}$ such that

$$
\begin{equation*}
2^{-k^{\prime}} \cdot \operatorname{diam}(M, \mathrm{~d}, m) \leq \lambda \tag{4.28}
\end{equation*}
$$

Our next claim is that

- Property $\mathrm{C}\left(\mathrm{k}^{\prime}\right)$ is satisfied.

In order to prove this claim, fix $\Gamma$ as well as $s, t$ and let $\hat{q}$ be a coupling of $\Gamma(0), \Gamma(s), \Gamma(t), \Gamma(1)$ on $M^{4}$. Then according to Lemma 2.11, for $\hat{q}$-almost every $\left(x_{0}, x_{s}, x_{t}, x_{1}\right) \in M^{4}$ the points $x_{s}, x_{t}$ lie on some geodesic connecting $x_{0}$ and $x_{1}$ with

$$
\begin{equation*}
d\left(x_{s}, x_{t}\right)=|t-s| \cdot d\left(x_{0}, x_{1}\right) \leq 2^{-k^{\prime}} \cdot \operatorname{diam}(M, \mathrm{~d}, m) \leq \lambda \tag{4.29}
\end{equation*}
$$

according to (4.28). Define probability measures $\Gamma_{j}(s)$ and $\Gamma_{j}(t)$ for $j=1, \ldots, n$ by

$$
\Gamma_{j}(s)(A):=\frac{1}{\alpha_{j}} \Gamma(s)\left(A \cap L_{j}\right)=\frac{1}{\alpha_{j}} \hat{q}\left(M \times\left(A \cap L_{j}\right) \times M \times M\right)
$$

and

$$
\Gamma_{j}(t)(A):=\frac{1}{\alpha_{j}} \hat{q}\left(M \times L_{j} \times A \times M\right)
$$

provided $\alpha_{j}:=\Gamma_{s}\left(L_{j}\right) \neq 0$. [Otherwise, define $\Gamma_{j}(s)$ and $\Gamma_{j}(t)$ arbitrarily.] Then $\operatorname{supp}\left[\Gamma_{j}(s)\right] \subset$ $\bar{L}_{j}$ which together with (4.29) implies

$$
\begin{equation*}
\operatorname{supp}\left[\Gamma_{j}(s)\right] \cup \operatorname{supp}\left[\Gamma_{j}(t)\right] \subset \bar{B}_{\lambda}\left(L_{j}\right) \subset M_{j} . \tag{4.30}
\end{equation*}
$$

Therefore, for each $j \in\{1, \ldots, n\}$ the assumption $\underline{\operatorname{Curv}}\left(M_{j}, \mathrm{~d}, m\right) \geq K$ can be applied to the probability measures $\Gamma_{j}(s)$ and $\Gamma_{j}(t) \in \mathcal{P}_{2}^{*}\left(M_{j}, \mathrm{~d}, m\right)$. It yields the existence of a midpoint $\eta_{j}(s, t)$ of them with the property

$$
\begin{equation*}
\operatorname{Ent}\left(\eta_{j}(s, t) \mid m\right) \leq \frac{1}{2} \operatorname{Ent}\left(\Gamma_{j}(s) \mid m\right)+\frac{1}{2} \operatorname{Ent}\left(\Gamma_{j}(t) \mid m\right)-\frac{K}{8} \mathrm{~d}_{W}^{2}\left(\Gamma_{j}(s), \Gamma_{j}(t)\right) . \tag{4.31}
\end{equation*}
$$

Define

$$
\eta(s, t):=\sum_{j=1}^{n} \alpha_{j} \eta_{j}(s, t)
$$

Then $\eta(s, t)$ is a midpoint of $\Gamma(s)=\sum_{j=1}^{n} \alpha_{j} \Gamma_{j}(s)$ and $\Gamma(t)=\sum_{j=1}^{n} \alpha_{j} \Gamma_{j}(t)$. Moreover, since the $\Gamma_{j}(s)$ for $j=1, \ldots, n$ are mutually singular and since $M$ is nonbranching, also the $\eta_{j}(s, t)$ for $j=1, \ldots, n$ are mutually singular, Lemma 2.11(iii). Hence, by (4.6)

$$
\begin{equation*}
\operatorname{Ent}(\eta(s, t) \mid m)=\sum_{j} \alpha_{j} \operatorname{Ent}\left(\eta_{j}(s, t) \mid m\right)+\sum_{j} \alpha_{j} \log \alpha_{j} \tag{4.32}
\end{equation*}
$$

and

$$
\begin{equation*}
\operatorname{Ent}(\Gamma(s) \mid m)=\sum_{j} \alpha_{j} \operatorname{Ent}\left(\Gamma_{j}(s) \mid m\right)+\sum_{j} \alpha_{j} \log \alpha_{j} \tag{4.33}
\end{equation*}
$$

whereas

$$
\begin{equation*}
\operatorname{Ent}(\Gamma(t) \mid m) \geq \sum_{j} \alpha_{j} \operatorname{Ent}\left(\Gamma_{j}(t) \mid m\right)+\sum_{j} \alpha_{j} \log \alpha_{j} \tag{4.34}
\end{equation*}
$$

since the $\Gamma_{j}(t)$ for $j=1, \ldots, n$ are not necessarily mutually singular. Summing up (4.31) for $j=1, \ldots, n$ and using (4.32) - (4.34) yields (4.24). This proves property C(k').

In order to finish the proof of the Theorem, let two probability measures $\nu_{0}, \nu_{1} \in \mathcal{P}_{2}^{*}(M, \mathrm{~d}, m)$ be given. By assumption, there exists a geodesic $\Gamma$ in $\mathcal{P}_{2}^{*}(M, \mathrm{~d}, m)$ connecting them. According to our second claim, property $\mathrm{C}\left(\mathrm{k}^{\prime}\right)$ is satisfied and according to our first claim, this implies property $\mathrm{C}(\mathrm{k})$ for all $k=k^{\prime}-1, k^{\prime}-2, \ldots, 0$. Property $\mathrm{C}(0)$ finally states that there exists a midpoint $\eta$ of $\Gamma(0)$ and $\Gamma(1)$ with

$$
\begin{equation*}
\operatorname{Ent}(\eta \mid m) \leq \frac{1}{2} \operatorname{Ent}(\Gamma(0) \mid m)+\frac{1}{2} \operatorname{Ent}(\Gamma(1) \mid m)-\frac{K}{8} \mathrm{~d}_{W}^{2}(\Gamma(0), \Gamma(1)) \tag{4.35}
\end{equation*}
$$

This proves the Theorem.
Remark 4.18. Let $M$ be a compact space.
(i) The condition on $\mathcal{P}_{2}^{*}(M, \mathrm{~d}, m)$ to be a geodesic space is always satisfied if $\underline{\mathbb{C u r v}}(M, \mathrm{~d}, m)>$ $-\infty$ (Remark 4.6(iii)).
(ii) If $\mathcal{P}_{2}^{*}(M, \mathrm{~d}, m)$ is a geodesic space then $\operatorname{supp}[m]$ is a geodesic space. The converse is not true in general. Conjecture: it is true under the additional assumption $\mathbb{C u r v}_{l o c}(M, \mathrm{~d}, m)>-\infty$.
(iii) If $M_{0}:=\operatorname{supp}[m]$ is a geodesic space then $\mathcal{P}_{2}\left(M_{0}, \mathrm{~d}\right)$ is a geodesic space. Moreover, the space $\mathcal{P}_{2}^{*}\left(M_{0}, \mathrm{~d}, m\right)$ is dense in $\mathcal{P}_{2}\left(M_{0}, \mathrm{~d}\right)$. Indeed, given any $\mu \in \mathcal{P}_{2}\left(M_{0}, \mathrm{~d}\right)$ and any $\epsilon>0$ there exist $n \in \mathbb{N}$ and $x_{1}, \ldots, x_{n} \in M_{0}$ such that $\mathrm{d}_{W}\left(\mu, \mu^{\prime}\right) \leq \epsilon$ where $\mu^{\prime}:=\frac{1}{n} \sum_{i=1}^{n} \delta_{x_{i}}$. Moreover, $\mathrm{d}_{W}\left(\mu^{\prime}, \mu^{\prime \prime}\right) \leq \epsilon$ where $\mu^{\prime \prime}:=\frac{1}{n} \sum_{i=1}^{n} \frac{1}{m\left(B_{\epsilon}\left(x_{i}\right)\right)} \cdot 1_{B_{\epsilon}\left(x_{i}\right)} m$ and

$$
\operatorname{Ent}\left(\mu^{\prime \prime} \mid m\right) \leq \sup _{x \in M_{0}} \log \left[\frac{1}{n} \sum_{i=1}^{n} \frac{1}{m\left(B_{\epsilon}\left(x_{i}\right)\right)} \cdot 1_{B_{\epsilon}\left(x_{i}\right)}(x)\right] \leq-\inf _{i=1, \ldots, n} \log m\left(B_{\epsilon}\left(x_{i}\right)\right)<\infty
$$

That is, $\mu^{\prime \prime} \in \mathcal{P}_{2}^{*}\left(M_{0}, \mathrm{~d}, m\right)$ and $\mathrm{d}_{W}\left(\mu, \mu^{\prime \prime}\right) \leq 2 \epsilon$ which proves the density.

### 4.5 Stability under Convergence

One of the most important results in this paper is that our curvature bounds for metric measure spaces are stable under convergence. The key to this result is the fact that we are able to construct a transformation $Q^{\prime}$ from the $L_{2}$-Wasserstein space over one metric measure space ( $M, \mathrm{~d}, m$ ) to the $L_{2}$-Wasserstein space over any other metric measure space ( $M^{\prime}, \mathrm{d}^{\prime}, m^{\prime}$ ) which reduces the relative entropy and which is almost an isometry between these spaces, provided the underlying spaces are close in the metric $\mathbb{D}$.
Actually, this is quite easy in the particular case where $m$ is the push forward of $m^{\prime}$ under a map $\psi^{\prime}: M^{\prime} \rightarrow M$. In this case we can define (similarly to the construction in the proof of Proposition 4.12)

$$
Q^{\prime}: \mathcal{P}_{2}(M, \mathrm{~d}, m) \rightarrow \mathcal{P}_{2}\left(M^{\prime}, \mathrm{d}^{\prime}, m^{\prime}\right), \quad \rho \cdot m \mapsto\left(\rho \circ \psi^{\prime}\right) \cdot m^{\prime} .
$$

The general case is more subtle since we may not restrict ourselves to transformations derived from push forward maps. For instance, if $m^{\prime}$ is the Riemannian measure of a collapsed space then the Riemannian measure $m$ of the initial space cannot be represented as a push forward measure.

Given two normalized metric measure spaces $(M, \mathrm{~d}, m)$ and $\left(M^{\prime}, \mathrm{d}^{\prime}, m^{\prime}\right)$ we will define a canonical map

$$
Q^{\prime}: \mathcal{P}_{2}(M, \mathrm{~d}, m) \rightarrow \mathcal{P}_{2}\left(M^{\prime}, \mathrm{d}^{\prime}, m^{\prime}\right)
$$

as follows: Let $q$ be a coupling of $m$ and $m^{\prime}$ and $\hat{\mathrm{d}}$ be a coupling of d and $\mathrm{d}^{\prime}$ such that

$$
\int \hat{\mathrm{d}}^{2}\left(x, x^{\prime}\right) d q\left(x, x^{\prime}\right) \leq 2 \mathbb{D}^{2}\left((M, \mathrm{~d}, m),\left(M^{\prime}, \mathrm{d}^{\prime}, m^{\prime}\right)\right)
$$

Let $Q^{\prime}$ and $Q$ be the disintegrations of $q$ w.r.t. $m^{\prime}$ and $m$, resp., that is,

$$
d q\left(x, x^{\prime}\right)=Q^{\prime}\left(x^{\prime}, d x\right) d m^{\prime}\left(x^{\prime}\right)=Q\left(x, d x^{\prime}\right) d m(x)
$$

and let $\hat{\Delta}$ denote the $m$-essential supremum of the map

$$
x \mapsto\left[\int_{M^{\prime}} \hat{\mathrm{d}}^{2}\left(x, x^{\prime}\right) Q\left(x, d x^{\prime}\right)\right]^{1 / 2}
$$

In general, of course, $\hat{\Delta}$ may attain the value $\infty$. However, if both metric measure spaces have finite diameter we have

$$
\hat{\Delta} \leq \operatorname{diam}(M, \mathrm{~d}, m)+\operatorname{diam}\left(M^{\prime}, \mathrm{d}^{\prime}, m^{\prime}\right)<\infty .
$$

For $\nu=\rho m \in \mathcal{P}_{2}(M, \mathrm{~d}, m)$ define $Q^{\prime}(\nu) \in \mathcal{P}_{2}\left(M^{\prime}, \mathrm{d}^{\prime}, m^{\prime}\right)$ by $Q^{\prime}(\nu):=\rho^{\prime} m^{\prime}$ where

$$
\begin{equation*}
\rho^{\prime}\left(x^{\prime}\right):=\int_{M} \rho(x) Q^{\prime}\left(x^{\prime}, d x\right) . \tag{4.36}
\end{equation*}
$$

In other words, for all measurable $A \subset M^{\prime}$

$$
\begin{aligned}
Q^{\prime}(\nu)(A) & =\int_{M^{\prime}} 1_{A}\left(x^{\prime}\right) \rho^{\prime}\left(x^{\prime}\right) d m^{\prime}\left(x^{\prime}\right) \\
& =\int_{M^{\prime}} \int_{M} 1_{A}\left(x^{\prime}\right) \rho(x) Q^{\prime}\left(x^{\prime}, d x\right) d m^{\prime}\left(x^{\prime}\right) \\
& =\int_{M \times M^{\prime}} 1_{A}\left(x^{\prime}\right) \rho(x) d q\left(x, x^{\prime}\right) .
\end{aligned}
$$

Lemma 4.19. The map $Q^{\prime}$ defined as above satisfies $Q^{\prime}(m)=m^{\prime}$ and for all $\nu=\rho m$ :

$$
\begin{equation*}
\operatorname{Ent}\left(Q^{\prime}(\nu) \mid m^{\prime}\right) \leq \operatorname{Ent}(\nu \mid m) \tag{4.37}
\end{equation*}
$$

and

$$
\begin{equation*}
\mathrm{d}_{W}^{2}\left(\nu, Q^{\prime}(\nu)\right) \leq \frac{2+\hat{\Delta}^{2} \cdot \operatorname{Ent}(\nu \mid m)}{-\log \mathbb{D}\left((M, \mathrm{~d}, m),\left(M^{\prime}, \mathrm{d}^{\prime}, m^{\prime}\right)\right)} \tag{4.38}
\end{equation*}
$$

provided $\mathbb{D}\left((M, \mathrm{~d}, m),\left(M^{\prime}, \mathrm{d}^{\prime}, m^{\prime}\right)\right)<1$.

Proof. Inequality (4.37) is a consequence of Jensen's inequality, applied to the convex function $r \mapsto r \log r$, as follows

$$
\begin{aligned}
\operatorname{Ent}\left(Q^{\prime}(\nu) \mid m^{\prime}\right) & =\int \rho^{\prime} \log \rho^{\prime} d m^{\prime} \\
& =\int\left[\int \rho(x) Q^{\prime}\left(x^{\prime}, d x\right)\right] \cdot \log \left[\int \rho(x) Q^{\prime}\left(x^{\prime}, d x\right)\right] d m^{\prime}\left(x^{\prime}\right) \\
& \leq \iint \rho(x) \log \rho(x) Q^{\prime}\left(x^{\prime}, d x\right) d m^{\prime}\left(x^{\prime}\right) \\
& =\int \rho(x) \log \rho(x) d m(x) \\
& =\operatorname{Ent}(\nu \mid m) .
\end{aligned}
$$

Inequality (4.38) follows from the fact that the measure $\rho(x) d q\left(x, x^{\prime}\right)=\rho(x) Q\left(x, d x^{\prime}\right) d m(x)$ is a coupling of $\rho(x) d m(x)$ and $\rho^{\prime}\left(x^{\prime}\right) d m^{\prime}\left(x^{\prime}\right)=\int_{M} \rho(x) Q^{\prime}\left(x^{\prime}, d x\right) d m^{\prime}\left(x^{\prime}\right)=\int_{M} \rho(x) d q\left(x, x^{\prime}\right)$ and thus

$$
\mathrm{d}_{W}^{2}\left(\rho m, \rho^{\prime} m^{\prime}\right) \leq \iint \hat{\mathrm{d}}^{2}\left(x, x^{\prime}\right) Q\left(x, d x^{\prime}\right) \rho(x) d m(x)=: \Phi(\rho) .
$$

Now again with Jensen's inequality applied to the convex function $\psi(r):=r \log r$,

$$
\psi\left(\frac{\Phi(\rho)}{\Phi(1)}\right) \leq \frac{1}{\Phi(1)} \iint \hat{\mathrm{d}}^{2}\left(x, x^{\prime}\right) Q\left(x, d x^{\prime}\right) \psi(\rho(x)) d m(x) \leq \frac{\hat{\Delta}^{2}}{\Phi(1)} \operatorname{Ent}(\nu \mid m)
$$

Hence, since by assumption $\Phi(1) \leq 2 \mathbb{D}\left((M, \mathrm{~d}, m),\left(M^{\prime}, \mathrm{d}^{\prime}, m^{\prime}\right)\right)<2$

$$
\begin{aligned}
\mathrm{d}_{W}^{2}\left(\rho m, \rho^{\prime} m^{\prime}\right) & \leq \Phi(\rho) \leq \Phi(1) \cdot \psi^{-1}\left(\hat{\Delta}^{2} \operatorname{Ent}(\nu \mid m) / \Phi(1)\right) \\
& \leq \Phi(1) \cdot \psi^{-1}\left(\left[2+\hat{\Delta}^{2} \operatorname{Ent}(\nu \mid m)\right] / \Phi(1)\right) \\
& =\frac{2+\hat{\Delta}^{2} \operatorname{Ent}(\nu \mid m)}{\log \psi^{-1}\left(\left[2+\hat{\Delta}^{2} \operatorname{Ent}(\nu \mid m)\right] / \Phi(1)\right)} \\
& \leq 2 \frac{2+\hat{\Delta}^{2} \operatorname{Ent}(\nu \mid m)}{\log \left(\left[2+\hat{\Delta}^{2} \operatorname{Ent}(\nu \mid m)\right] / \Phi(1)\right)} \\
& \leq 2 \frac{2+\hat{\Delta}^{2} \operatorname{Ent}(\nu \mid m)}{\log \left(\left[2+\hat{\Delta}^{2} \operatorname{Ent}(\nu \mid m)\right] / 2 \mathbb{D}{ }^{2}\right)} \\
& \leq \frac{2+\hat{\Delta}^{2} \operatorname{Ent}(\nu \mid m)}{-\log \mathbb{D}}
\end{aligned}
$$

where we have used the abbreviation $\mathbb{D}:=\mathbb{D}\left((M, \mathrm{~d}, m),\left(M^{\prime}, \mathrm{d}^{\prime}, m^{\prime}\right)\right)$.
Theorem 4.20. ('Convergence')
Let $\left(\left(M_{n}, \mathrm{~d}_{n}, m_{n}\right)\right)_{n \in \mathbb{N}}$ be a sequence of normalized metric measure spaces with uniformly bounded diameter. If

$$
\left(M_{n}, \mathrm{~d}_{n}, m_{n}\right) \xrightarrow{\mathbb{D}}(M, \mathrm{~d}, m)
$$

as $n \rightarrow \infty$ then

$$
\limsup _{n \rightarrow \infty} \mathbb{C u r v}_{l a x}\left(M_{n}, \mathrm{~d}_{n}, m_{n}\right) \leq{\underset{\mathbb{C u r v}_{l a x}}{ }(M, \mathrm{~d}, m) . . . . ~}_{\text {. }}
$$

In particular, for each $K \in \mathbb{R}$ and $\Delta \in \mathbb{R}_{+}$the family $\mathbb{X}_{1}(K, \Delta)$ of isomorphism classes of normalized metric measure spaces with curvature $\geq K$ and diameter $\leq \Delta$ is closed w.r.t $\mathbb{D}$.

Proof. Let $\left(\left(M_{n}, \mathrm{~d}_{n}, m_{n}\right)\right)_{n \in \mathbb{N}}$ be a sequence in $\left(\mathbb{X}_{1}, \mathbb{D}\right)$ with $\left(M_{n}, \mathrm{~d}_{n}, m_{n}\right) \rightarrow(M, \mathrm{~d}, m)$ and assume that $\operatorname{diam}(M, \mathrm{~d}, m) \leq \Delta$ and $\mathbb{C u r v}_{\text {lax }}\left(M_{n}, \mathrm{~d}_{n}, m_{n}\right) \geq K$ for some $\Delta, K \in \mathbb{R}$ and all sufficiently large $n \in \mathbb{N}$. Now let $\epsilon>0$ and $\nu_{0}=\rho_{0} m, \nu_{1}=\rho_{1} m \in \mathcal{P}_{2}^{*}(M, \mathrm{~d}, m)$ be given. Choose $R$ with

$$
\begin{equation*}
\sup _{i=0,1} \operatorname{Ent}\left(\nu_{i} \mid m\right)+\frac{|K|}{8}\left[\mathrm{~d}_{W}\left(\nu_{0}, \nu_{1}\right)+2 \epsilon\right]^{2}+\epsilon \leq R . \tag{4.39}
\end{equation*}
$$

We have to deduce the existence of an $\epsilon$-midpoint $\eta$ which satisfies inequality (4.11). Choose $n \in \mathbb{N}$ with

$$
\begin{equation*}
\mathbb{D}\left(\left(M_{n}, \mathrm{~d}_{n}, m_{n}\right),(M, \mathrm{~d}, m)\right) \leq \exp \left(-\frac{2+\left(\Delta+\Delta^{\prime}\right)^{2} K}{\epsilon^{2}}\right) \tag{4.40}
\end{equation*}
$$

Define the map $Q_{n}^{\prime}: \mathcal{P}_{2}(M, \mathrm{~d}, m) \rightarrow \mathcal{P}_{2}\left(M_{n}, \mathrm{~d}_{n}, m_{n}\right)$ as in the previous lemma, now with $m_{n}$ in the place of $m^{\prime}$, and analogously the map $Q_{n}: \mathcal{P}_{2}\left(M_{n}, \mathrm{~d}_{n}, m_{n}\right) \rightarrow \mathcal{P}_{2}(M, \mathrm{~d}, m)$. Put

$$
\nu_{i, n}:=Q_{n}^{\prime}\left(\nu_{i}\right)=\rho_{i, n} m_{n}
$$

with $\rho_{i, n}(y)=\int \rho_{i}(x) Q_{n}^{\prime}(y, d x)$ for $i=0,1$ and let $\eta_{n}$ be an $\epsilon$-midpoint of $\nu_{0, n}$ and $\nu_{1, n}$ with

$$
\begin{equation*}
\operatorname{Ent}\left(\eta_{n} \mid m_{n}\right) \leq \frac{1}{2} \operatorname{Ent}\left(\nu_{0, n} \mid m_{n}\right)+\frac{1}{2} \operatorname{Ent}\left(\nu_{1, n} \mid m_{n}\right)-\frac{K}{8} \mathrm{~d}_{W}^{2}\left(\nu_{0, n}, \nu_{1, n}\right)+\epsilon \tag{4.41}
\end{equation*}
$$

From (4.38) - (4.40) we conclude

$$
\begin{aligned}
\mathrm{d}_{W}^{2}\left(\nu_{0}, \nu_{0, n}\right) & \leq \frac{2+\hat{\Delta}^{2} \cdot \operatorname{Ent}\left(\nu_{0} \mid m\right)}{-2 \log \mathbb{D}\left((M, \mathrm{~d}, m),\left(M_{n}, \mathrm{~d}_{n}, m_{n}\right)\right)} \\
& \leq \frac{2+\left(\Delta+\Delta^{\prime}\right)^{2} \cdot R}{-2 \log \mathbb{D}\left((M, \mathrm{~d}, m),\left(M_{n}, \mathrm{~d}_{n}, m_{n}\right)\right)} \leq \epsilon^{2}
\end{aligned}
$$

and analogously $\mathrm{d}_{W}^{2}\left(\nu_{1}, \nu_{1, n}\right) \leq \epsilon^{2}$. Moreover, (4.37) and (4.41) imply

$$
\begin{aligned}
\operatorname{Ent}\left(\eta_{n} \mid m_{n}\right) & \leq \frac{1}{2} \operatorname{Ent}\left(\nu_{0, n} \mid m_{n}\right)+\frac{1}{2} \operatorname{Ent}\left(\nu_{1, n} \mid m_{n}\right)-\frac{K}{8} \mathrm{~d}_{W}^{2}\left(\nu_{0, n}, \nu_{1, n}\right)+\epsilon \\
& \leq \frac{1}{2} \operatorname{Ent}\left(\nu_{0} \mid m\right)+\frac{1}{2} \operatorname{Ent}\left(\nu_{1} \mid m\right)-\frac{K}{8} \mathrm{~d}_{W}^{2}\left(\nu_{0}, \nu_{1}\right)+\epsilon^{\prime}
\end{aligned}
$$

with $\epsilon^{\prime}=\left[1+\frac{|K|}{2}\left(\mathrm{~d}_{W}\left(\nu_{0}, \nu_{1}\right)+\epsilon\right)\right] \cdot \epsilon$. Finally, put

$$
\eta=Q_{n}\left(\eta_{n}\right)
$$

Then again by (4.38) - (4.40) and by the previous estimate for $\operatorname{Ent}\left(\eta_{n} \mid m\right)$

$$
\begin{aligned}
\mathrm{d}_{W}^{2}\left(\eta_{n}, \eta\right) & \leq \frac{2+\hat{\Delta}^{2} \cdot \operatorname{Ent}\left(\eta_{n} \mid m\right)}{-2 \log \mathbb{D}\left((M, \mathrm{~d}, m),\left(M_{n}, \mathrm{~d}_{n}, m_{n}\right)\right)} \\
& \leq \frac{2+\left(\Delta+\Delta^{\prime}\right)^{2} \cdot R}{-2 \log \mathbb{D}\left((M, \mathrm{~d}, m),\left(M_{n}, \mathrm{~d}_{n}, m_{n}\right)\right)} \leq \epsilon^{2}
\end{aligned}
$$

Hence,

$$
\sup _{i=0,1} \mathrm{~d}_{W}\left(\eta, \nu_{i}\right) \leq \frac{1}{2} \mathrm{~d}_{W}\left(\nu_{0}, \nu_{1}\right)+4 \epsilon,
$$

i.e. $\eta$ is a (4 $\epsilon$-midpoint of $\nu_{0}$ and $\nu_{1}$. Furthermore, by (4.37)

$$
\begin{aligned}
\operatorname{Ent}(\eta \mid m) & \leq \operatorname{Ent}\left(\eta_{n} \mid m_{n}\right) \\
& \leq \frac{1}{2} \operatorname{Ent}\left(\nu_{0} \mid m\right)+\frac{1}{2} \operatorname{Ent}\left(\nu_{1} \mid m\right)-\frac{K}{8} \mathrm{~d}_{W}^{2}\left(\nu_{0}, \nu_{1}\right)+\epsilon^{\prime}
\end{aligned}
$$

with $\epsilon^{\prime}$ as above. This proves that $\mathbb{C u r v}_{l a x}(M, \mathrm{~d}, m) \geq K$.

As an immediate consequence of the previous theorem together with Proposition 4.16 we obtain:
Corollary 4.21. ('Infinite Products')
Let $(M, \mathrm{~d}, m)=\bigotimes_{n \in \mathbb{N}}\left(M_{n}, \mathrm{~d}_{n}, m_{n}\right)$ where $\left(M_{n}, \mathrm{~d}_{n}, m_{n}\right)$ for $n \in \mathbb{N}$ are normalized metric measure spaces with compact nonbranching $M_{n}$. Assume $\sum_{n \in \mathbb{N}} \mathbb{V} \operatorname{ar}\left(M_{n}, \mathrm{~d}_{n}, m_{n}\right)<\infty$. Then

$$
\underline{\operatorname{Curv}}(M, \mathrm{~d}, m)=\inf _{n \in \mathbb{N}} \underline{\operatorname{Curv}}\left(M_{n}, \mathrm{~d}_{n}, m_{n}\right) .
$$

Important infinite dimensional examples are given by abstract Wiener spaces. Let ( $M, H, m$ ) be an abstract Wiener space, that is, $M$ is a separable Banach space, $m$ is a Gaussian measure on $M$, and $H$ is a separable Hilbert space that is continuously and densely embedded in $M$ such that

$$
\int_{M} \exp (i\langle x, y\rangle) d m(x)=\exp \left(-\frac{1}{2}\|y\|_{H}^{2}\right)
$$

for any $y \in M^{*} \subset H$ (where we identify $H$ with its dual). For the classical Wiener space, $M=\mathcal{C}\left(\mathbb{R}_{+}, \mathbb{R}\right)$ is the path space of one-dimensional Brownian motion,

$$
H=\left\{u \in M: u \text { is abs. cont. with } \int_{\mathbb{R}_{+}}|\dot{u}(t)|^{2} d t<\infty\right\}
$$

is the Cameron-Martin space, and $m$ is the Wiener measure.
Given any abstract Wiener space $(M, H, m)$ define a pseudo metric on $M$ by

$$
\mathrm{d}(x, y):=\|x-y\|_{H}
$$

if $x-y \in H$ and $\mathrm{d}(x, y):=\infty$ else and consider the 'pseudo metric measure space' $(M, \mathrm{~d}, m)$. Of course, formally this does not fit in our framework. Nevertheless, the definition of the $L_{2^{-}}$ Wasserstein distance $\mathrm{d}_{W}$ derived from this pseudo metric d perfectly makes sense. It is a pseudo metric on the space of probability measures on $M$ and a metric on the subspace of all those probability measures which have finite $\mathrm{d}_{W}$-distance from $m$. Also the relative entropy Ent $(. \mid m)$ and the curvature bound $\mathbb{C u r v}(M, \mathrm{~d}, m)$ are well-defined. Particular attention, however, has to be paid to the fact that $\mathrm{d}_{W}$ is not continuous w.r.t. the weak topology of measures on $M$; it behaves very singular. For a detailed analysis, we refer to [FÜ04a, FÜ04b]. Here we restrict ourselves to the following result.

Proposition 4.22. ('Wiener Space')

$$
\mathbb{C u r v}_{l a x}(M, \mathrm{~d}, m) \geq 1 .
$$

Proof. Let $\nu_{0}, \nu_{1} \in \mathcal{P}_{2}^{*}(M, \mathrm{~d}, m)$ and $\epsilon>0$ be given. Choose a increasing ('total') sequence $\left(M_{n}\right)_{n}$ of regular finite dimensional subspaces with $\bigcup_{n} M_{n}$ being dense in $W$ and $H$. Let $m_{n}$, $\nu_{0, n}$ and $\nu_{1, n}$ be the images measures of $m, \nu_{0}$ and $\nu_{1}$, resp., under the projections $\pi_{n}: M \rightarrow M_{n}$. Then

$$
\mathrm{d}_{W}\left(\nu_{0, n}, \nu_{0}\right) \rightarrow 0, \quad \mathrm{~d}_{W}\left(\nu_{1, n}, \nu_{1}\right) \rightarrow 0
$$

and

$$
\operatorname{Ent}\left(\nu_{0, n} \mid m_{n}\right) \rightarrow \operatorname{Ent}\left(\nu_{0} \mid m\right), \quad \operatorname{Ent}\left(\nu_{1, n} \mid m_{n}\right) \rightarrow \operatorname{Ent}\left(\nu_{1} \mid m\right)
$$

as $n \rightarrow \infty$. The space $\left(M_{n}, M_{n}, m_{n}\right)$ is a finite dimensional abstract Wiener space. Therefore, it is isomorphic to

$$
\left(\mathbb{R}^{N}, \mathbb{R}^{N}, \exp \left(-\|x\|^{2} / 2\right) d x\right)
$$

for some $N=N(n) \in \mathbb{N}$ (where $d x$ denotes the Lebesgue measure in $\mathbb{R}^{N}$ ). Hence, according to Theorem 4.9

$$
\underline{\operatorname{Curv}}\left(M_{n}, \mathrm{~d}, m_{n}\right)=1 .
$$

Thus for each $n \in \mathbb{N}$ there exists a midpoint $\eta_{n}$ of $\nu_{0, n}$ and $\nu_{1, n}$ with

$$
\begin{aligned}
\operatorname{Ent}\left(\eta_{n} \mid m\right) & =\operatorname{Ent}\left(\eta_{n} \mid m_{n}\right) \\
& \leq \frac{1}{2} \operatorname{Ent}\left(\nu_{0, n} \mid m_{n}\right)+\frac{1}{2} \operatorname{Ent}\left(\nu_{1, n} \mid m_{n}\right)-\frac{K}{8} \mathrm{~d}_{W}^{2}\left(\nu_{0, n}, \nu_{1, n}\right) \\
& \leq \frac{1}{2} \operatorname{Ent}\left(\nu_{0} \mid m\right)+\frac{1}{2} \operatorname{Ent}\left(\nu_{1} \mid m\right)-\frac{K}{8} \mathrm{~d}_{W}^{2}\left(\nu_{0}, \nu_{1}\right)+\epsilon
\end{aligned}
$$

for $n$ large enough. This proves the claim since $\eta_{n}$ is an $\epsilon$-midpoint of $\nu_{0}$ and $\nu_{1}$ (again, for large enough $n$ ).

### 4.6 Volume Growth Estimates

In the Riemannian setting, it is well-known that lower bounds for the Ricci curvature of the underlying space imply upper bounds for the growth

$$
R \mapsto m\left(\bar{B}_{R}(x)\right)
$$

of the volume of concentric balls. In particular, this growth is at most exponentially in $R$. This is the content of the famous Bishop-Gromov volume comparison theorem.
Also for general metric measure spaces, lower bounds for the curvature will imply upper estimates for the volume growth of concentric balls. These estimates, however, have to take into account that in the general case (without any dimensional restriction) the volume can grow much faster than exponentially. For instance, already in the following standard example we observe squared exponential volume growth.

Example 4.23. Let ( $M, \mathrm{~d}$ ) be the one-dimensional Euclidean space equipped with the measure $d m(x)=\exp \left(-\frac{K}{2} x^{2}\right) d x$ for some $K \in \mathbb{R}$. Then $\underline{\mathbb{C u r v}}(M, \mathrm{~d}, m)=K$ and, if $K<0$

$$
m\left(\bar{B}_{R}(x)\right) \geq \exp \left(\frac{|K|}{2}\left(R-\frac{1}{2}\right)^{2}\right)
$$

for each $x \in M$ and $R \geq \frac{1}{2}$.
Theorem 4.24. Let $(M, \mathrm{~d}, m)$ be an arbitrary metric measure space with $\underline{\mathbb{C u r v}}(M, \mathrm{~d}, m) \geq K$ for some $K \leq 0$. For fixed $x \in \operatorname{supp}[m] \subset M$ consider the volume growth

$$
v_{R}:=m\left(\bar{B}_{R}(x)\right)
$$

of closed balls centered at $x$. Then for all $R \geq 2 \epsilon>0$

$$
\begin{equation*}
v_{R} \leq v_{2 \epsilon} \cdot\left(\frac{v_{2 \epsilon}}{v_{\epsilon}}\right)^{R / \epsilon} \cdot \exp \left(\frac{|K|}{2}\left(R+\frac{\epsilon}{2}\right)^{2}\right) \tag{4.42}
\end{equation*}
$$

In particular, each ball in $M$ has finite volume.
Proof. Apply the following Lemma with $r=\epsilon$.
Lemma 4.25. Let $(M, \mathrm{~d}, m), K$ and $x$ as in the above theorem. Then for all $\epsilon, R>0$ and all $t \in] 0,1]$

$$
\begin{equation*}
\log v_{R} \leq \frac{1}{t} \log v_{\epsilon+t(R+\epsilon)}+\left(1-\frac{1}{t}\right) \log v_{\epsilon}+\frac{|K|}{2}(1-t)(R+\epsilon)^{2} \tag{4.43}
\end{equation*}
$$

In other words, for all $\epsilon, r>0$ and all $R>\epsilon+r$

$$
\begin{equation*}
v_{R} \leq v_{\epsilon} \cdot\left(\frac{v_{\epsilon+r}}{v_{\epsilon}}\right)^{\frac{R+\epsilon}{r}} \cdot \exp \left(\frac{|K|}{2}(R+\epsilon-r)(R+\epsilon)\right) \tag{4.44}
\end{equation*}
$$

Proof. Fix $x \in \operatorname{supp}[m]$ and $\epsilon, R>0$. Let $\nu_{0}$ and $\nu_{1}$ denote the uniform distributions in $\bar{B}_{\epsilon}(x)$ and $\bar{B}_{R}(x)$, resp. That is,

$$
d \nu_{0}(x)=\frac{1}{v_{\epsilon}} \cdot 1_{\bar{B}_{\epsilon}(x)} d m(x), \quad d \nu_{1}(x)=\frac{1}{v_{R}} \cdot 1_{\bar{B}_{R}(x)} d m(x) .
$$

Then obviously $\nu_{0}, \nu_{1} \in \mathcal{P}_{2}^{*}(M, \mathrm{~d}, m)$ with

$$
\operatorname{Ent}\left(\nu_{0} \mid m\right)=-\log v_{\epsilon}, \quad \operatorname{Ent}\left(\nu_{1} \mid m\right)=-\log v_{R} .
$$

Let $\nu_{t}, t \in[0,1]$, be a geodesic in $\mathcal{P}_{2}^{*}(M, \mathrm{~d}, m)$ connecting $\nu_{0}, \nu_{1}$ such that

$$
\operatorname{Ent}\left(\nu_{t} \mid m\right) \leq(1-t) \operatorname{Ent}\left(\nu_{0} \mid m\right)+t \operatorname{Ent}\left(\nu_{1} \mid m\right)-\frac{K}{2} t(1-t) \mathrm{d}_{W}^{2}\left(\nu_{0}, \nu_{1}\right)
$$

Such a geodesic exists according to our curvature assumption. Since $\mathrm{d}_{W}\left(\nu_{0}, \delta_{x}\right) \leq \epsilon$ and $\mathrm{d}_{W}\left(\nu_{1}, \delta_{x}\right) \leq R$ it follows

$$
\begin{equation*}
\mathrm{d}_{W}\left(\nu_{0}, \nu_{1}\right) \leq R+\epsilon . \tag{4.45}
\end{equation*}
$$

Moreover, if $\hat{q}$ is an optimal coupling of $\nu_{0}, \nu_{t}, \nu_{1}$ then for $\hat{q}$-almost every $\left(y_{0}, y_{t}, y_{1}\right)$ the point $y_{t}$ lies on a geodesic connecting $y_{0}, y_{1}$ with $\mathrm{d}\left(y_{0}, y_{t}\right)=t \cdot \mathrm{~d}\left(y_{0}, y_{1}\right)$. Together with (4.45), the latter implies

$$
\begin{equation*}
\operatorname{supp}\left[\nu_{t}\right] \subset \bar{B}_{\epsilon+t(R+\epsilon)}(x) . \tag{4.46}
\end{equation*}
$$

Now according to Jensen's inequality, for all $\nu_{t}$ satisfying (4.46)

$$
\operatorname{Ent}\left(\nu_{t} \mid m\right) \geq \operatorname{Ent}\left(m_{t} \mid m\right)
$$

where $m_{t}:=\frac{1}{v_{\epsilon+t(R+\epsilon)}} 1_{\bar{B}_{\epsilon+t(R+\epsilon)}(x)} m$ denotes uniform distribution in the closed ball $\bar{B}_{\epsilon+t(R+\epsilon)}(x)$. Hence,

$$
\begin{aligned}
-\log v_{\epsilon+t(R+\epsilon)} & =\operatorname{Ent}\left(m_{t} \mid m\right) \leq \operatorname{Ent}\left(\nu_{t} \mid m\right) \\
& \leq(1-t) \operatorname{Ent}\left(\nu_{0} \mid m\right)+t \operatorname{Ent}\left(\nu_{1} \mid m\right)-\frac{K}{2} t(1-t) \mathrm{d}_{W}^{2}\left(\nu_{0}, \nu_{1}\right) \\
& \leq-(1-t) \log v_{\epsilon}-t \log v_{R}+\frac{|K|}{2} t(1-t)(R+\epsilon)^{2} .
\end{aligned}
$$

This proves the first claim. For the second claim, choose $t=\frac{r}{R+\epsilon}$ and apply the first claim.
Slightly modifying the previous arguments also yields estimates for the volume of spherical shells

$$
v_{R, \delta}:=m\left(\bar{B}_{R}(x) \backslash B_{R-\delta}(x)\right) .
$$

Let $\nu_{1}$ denote uniform distribution in the shell $\bar{B}_{R}(x) \backslash B_{R-\delta}(x)$ and let $\nu_{0}$ (as before) be uniform distribution in $\left.\bar{B}_{\epsilon}(x)\right)$. Then we now obtain

$$
R-\epsilon-\delta \leq \mathrm{d}_{W}\left(\nu_{0}, \nu_{1}\right) \leq R+\epsilon
$$

and

$$
\begin{equation*}
\operatorname{supp}\left[\nu_{t}\right] \subset \bar{B}_{\epsilon+t(R+\epsilon)}(x) \backslash B_{R-\delta-(1-t)(R+\epsilon)} \tag{4.47}
\end{equation*}
$$

for the probability measures on the geodesic connecting $\nu_{0}$ and $\nu_{1}$. Hence, arguing similarly as before, we deduce

Theorem 4.26. Let $(M, \mathrm{~d}, m)$ be an arbitrary metric measure space with $\underline{\mathbb{C u r v}}(M, \mathrm{~d}, m) \geq K$ for some $K \in \mathbb{R}$. For fixed $x \in \operatorname{supp}[m]$ consider $v_{R, \delta}:=m\left(\bar{B}_{R}(x) \backslash B_{R-\delta}(x)\right)$. Then for all $\epsilon, \delta, r>0$ and all $R>r>2 \epsilon+\delta$

$$
\begin{equation*}
v_{R, \delta} \leq v_{\epsilon} \cdot\left(\frac{v_{\epsilon+r, 2 \epsilon+\delta}}{v_{\epsilon}}\right)^{\frac{R+\epsilon}{r}} \cdot \exp \left(-\frac{K}{2}\left(1-\frac{r}{R+\epsilon}\right) \cdot\left(R-\frac{\delta}{2} \pm \frac{2 \epsilon+\delta}{2}\right)^{2}\right) \tag{4.48}
\end{equation*}
$$

where $\pm$ has to be chosen as + if $K \leq 0$ and as - if $K>0$.
Choosing $\epsilon=\delta=r / 2$ this yields in the case $K \geq 0$ for all $R \geq 3 \epsilon>0$

$$
\begin{equation*}
v_{R, \epsilon} \leq v_{3 \epsilon} \cdot\left(\frac{v_{3 \epsilon}}{v_{\epsilon}}\right)^{R / 2 \epsilon} \cdot \exp \left(-\frac{K}{2}\left[(R-3 \epsilon)^{2}-\epsilon^{2}\right]\right) \tag{4.49}
\end{equation*}
$$

In particular, $K>0$ implies that $m$ has finite mass and finite variance.
In general, estimating the volume of concentric balls in terms of squared exponential growing functions is best possible, as demonstrated in the previous Example. In a forthcoming paper [St05], we discuss metric measure spaces satisfying a so-called curvature-dimension condition $(K, N)$ (replacing the condition that the curvature is $\geq K$ ) with some additional number $N \in$ $\mathbb{R}_{+}$, playing the role of a dimension. We will prove that under this condition the volume of balls grows at most exponentially.

## References

[Al51] A. D. Alexandrov (1951): A theorem on triangles in a metric space and some applications. Trudy Math. Inst. Steklov 38, 5-23. (Russian; translated into German and combined with more material in [Al57])
[Al57] A. D. Alexandrov (1957): Über eine Verallgemeinerung der Riemannschen Geometrie. Schr. Forschungsinst. Math. Berlin 1, 33-84.
[AT04] L. Ambrosio, P. Tilli (2004): Topics on analysis in metric spaces. Oxford Lecture Series in Mathematics and its Applications 25, Oxford University Press, Oxford.
[BE85] D. Bakry, M. Émery (1985): Diffusions hypercontractives. Séminaire de probabilités XIX, Lecture Notes in Math. 1123, Springer, Berlin, 177-206.
[BG99] S. G. Bobkov, F. Götze (1999): Exponential integrability and transportation cost related to logarithmic Sobolev inequalities. J. Funct. Anal. 163, no. 1, 1-28.
[BH99] M. R. Bridson, A. Haefliger (1999): Metric spaces of non-positive curvature. Grundlehren der Mathematischen Wissenschaften, 319, Springer-Verlag, Berlin.
[Br91] Y. Brenier (2001): Polar factorization and monotone rearrangement of vector-valued functions. Comm. Pure Appl. Math. 44, no. 4, 375-417.
[BBI01] D. Burago, Y. Burago, S. Ivanov (2001): A course in metric geometry. Graduate Studies in Mathematics 33. American Mathematical Society, Providence, RI.
[BGP92] Y. Burago, M. Gromov, G. Perelman (1992): A. D. Aleksandrov spaces with curvatures bounded below. (Russian) Uspekhi Mat. Nauk 47, no. 2(284), 3-51, 222; translation in Russian Math. Surveys 47 (1992), no. 2, 1-58.
[Ch93] I. Chavel (1993): Riemannian geometry-a modern introduction. Cambridge Tracts in Mathematics 108, Cambridge University Press, Cambridge.
[Ch70] J. Cheeger (1970): Finiteness theorems for Riemannian manifolds. Amer. J. Math. 92, 61-74.
[Ch99] J. Cheeger (1999): Differentiability of Lipschitz functions on metric measure spaces. Geom. Funct. Anal. 9, no. 3, 428-517.
[CC97/00] J. Cheeger, T. H. Colding (1997,2000,2000): On the structure of spaces with Ricci curvature bounded below I, II, III. J. Differential Geom. 46, no. 3, 406-480; J. Differential Geom. 54, no. 1, 13-35; J. Differential Geom. 54, no. 1, 37-74.
[CMS01] D. Cordero-Erausquin, R. McCann, M. Schmuckenschläger (2001): A Riemannian interpolation inequality à la Borell, Brascamb and Lieb. Invent. Math. 146, 219-257.
[Da89] E. B. Davies (1989): Heat kernels and spectral theory. Cambridge Tracts in Mathematics 92, Cambridge University Press, Cambridge.
[DD02] M. Del Pino, J. Dolbeault (2002): Best constants for Gagliardo-Nirenberg inequalities and applications to nonlinear diffusions. J. Math. Pures Appl. 81, no. 9, 847-875.
[Du89] R. M. Dudley (1989): Real analysis and probability. The Wadsworth \& Brooks/Cole Mathematics Series. Wadsworth \& Brooks/Cole Advanced Books \& Software, Pacific Grove, CA.
[FÜ04a] D. Feyel, A. S. ÜstÜnel (2004): Monge-Kantorovitch measure transportation and Monge-Ampère equation on Wiener space. Probab. Theory Related Fields 128, no. 3, 347-385.
[FÜ04b] D. Feyel, A. S. Üstünel (2004): The strong solution of the Monge-Ampère equation on the Wiener space for log-concave densities. C. R. Math. Acad. Sci. Paris 339, no. 1, 49-53.
[Gr81a] M. Gromov (1981): Groups of polynomial growth and expanding maps. inst. Hautes Études Sci. Publ. Math. 53, 53-78
[Gr81] M. Gromov (1981): Structures Métriques des Variétes Riemanniennes. Edited by J. LaFontaine and P. Pansu. Textes Mathématiques, 1. CEDIC, Paris, 1981.
[Gr99] M. Gromov (1999): Metric structures for Riemannian and non-Riemannian spaces. Based on [Gr81] (English translation by Sean Michael Bates). With appendices by M. Katz, P. Pansu and S. Semmes. Birkhäuser Boston Inc., Boston, MA.
[GP97] K. Grove, P. Petersen (1997): Comparison geometry. Papers from the Special Year in Differential Geometry held in Berkeley, CA, 1993-94. Research Institute Publications, 30, Cambridge University Press, Cambridge.
[He01] J. Heinonen (2001): Lectures on analysis on metric spaces. Universitext, SpringerVerlag, New York.
[HK95] P. HajŁasz, P. Koskela (1995): Sobolev meets Poincaré. C. R. Acad. Sci. Paris Sr. I Math. 320, no. 10, 1211-1215.
[HK00] P. HajŁasz, P. Koskela (2000): Sobolev met Poincaré. Mem. Amer. Math. Soc. 145, no. 688.
[JKO98] R. Jordan, D. Kinderlehrer, F. Otto (1998): The variational formulation of the Fokker-Planck equation. SIAM J. Math. Anal. 29, no. 1, 1-17 (electronic).
[Ka42] L. V. Kantorovich (1942): On the translocation of masses. C. R. (Doklady) Acad. Sci. URSS (N.S.) 37, 199-201.
[KR57] L. V. Kantorovich, G. S. Rubinstein (1957): On a functional space and certain extremum problems (Russian). Dokl. Akad. Nauk SSSR (N.S.) 115, 1058-1061.
[KK94] A. Kasue, H. Kumura (1994): Spectral convergence of Riemannian manifolds. Tohoku Math. J. 46, no. 2, 147-179.
[Ka04] A. Kasue (2004): Convergence of Riemannian manifolds and Laplace operators II, Preprint 2004.
[Ki01] J. Kigami (2001): Analysis on fractals. Cambridge Tracts in Mathematics 143. Cambridge University Press, Cambridge.
[Ko00] P. Koskela (2000): Upper gradients and Poincaré inequalities. Lecture notes on analysis in metric spaces (Trento, 1999), 55-69, Appunti Corsi Tenuti Docenti Sc., Scuola Norm. Sup., Pisa.
[KS03] K. Kuwae, T. Shioya (2003): Convergence of spectral structures: a functional analytic theory and its applications to spectral geometry. Comm. Anal. Geom. 11, no. 4, 599-673.
[Le01] M. Ledoux (2001): The concentration of measure phenomenon. Mathematical Surveys and Monographs 89, American Mathematical Society, Providence, RI, 2001.
[LY86] P. Li, S.-T. Yau (1986): On the parabolic kernel of the Schrödinger operator. Acta Math. 156, no. 3-4, 153-201.
[Mc95] R. J. McCann (1995): Existence and uniqueness of monotone measure-preserving maps. Duke Math. J. 80, no. 2, 309-323.
[Mc97] R. J. McCann (1997): A convexity principle for interacting gases. Adv. Math. 128, no. 1, 153-179.
[M1781] G. Monge (1871): Mémoire sur la théorie des déblais et des remblais. Histoire de l'Académie Royale des Sciences, Paris.
[OV00] F. Оtto, C. Villani (2000): Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. J. Funct. Anal. 173, no. 2, 361-400.
[Ot01] F. Otто (2001): The geometry of dissipative evolution equations: the porous medium equation. Comm. Partial Differential Equations 26, no. 1-2, 101-174.
[Pe02] G. Perelman (2002): The entropy formula for the Ricci flow and its geometric applications, Preprint 2002.
[Pl02] C. Plaut (2002): Metric spaces of curvature $\geq k$, Chapter 19 in Handbook of Geometric Topology, Elsevier Science.
[RR98] S. T. Rachev, L. Rüschendorf (1998): Mass transportation problems. Vol. I. Theory. Probability and its Applications. Springer, New York.
[RS04] M.-K. von Renesse, K. T. Sturm (2003): Transport inequalities, gradient estimates, entropy and Ricci curvature. To appear in Comm. Pure Appl. Math.
[SC02] L. Saloff-Coste (2002): Aspects of Sobolev-type inequalities. London Mathematical Society Lecture Note Series 289, Cambridge University Press, Cambridge.
[St98] K. T. Sturm (1998): Diffusion processes and heat kernels on metric spaces. Ann. Probab. 26, no. 1, 1-55.
[St99] K. T. Sturm (1999): Metric spaces of lower bounded curvature. Exposition. Math. 17 no. 1, 35-47.
[St03] K. T. Sturm (2003): Probability Measures on Metric Spaces of Nonpositive Curvature. In "Heat Kernels and Analysis on Manifolds, Graphs, and Metric Spaces" (edited by P. Auscher, T. Coulhon, A. Grigor'yan) pp. 357-390. Contemporary Mathematics 338. AMS 2003.
[St04] K. T. Sturm (2004): Convex functionals of probability measures and nonlinear diffusions on manifolds, to appear in J. Math. Pures Appl.
[St04a] K. T. Sturm (2004): Generalized Ricci Bounds and Convergence of Metric Measure Spaces, to appear in C. R. Acad. Sci. Paris, Ser. I.
[St05] K. T. Sturm (2005): A curvature-dimension condition for metric measure spaces. In preparation.
[Ta95] M. Talagrand (1995): Concentration of measure and isoperimetric inequalities in product spaces. Inst. Hautes Études Sci. Publ. Math. 81, 73-205.
[Vi03] C. Villani (2003): Topics in Mass Transportation. Graduate Studies in Mathematics, American Mathematical Society.
[Wa69] L. N. Wasserstein [or Vasershtein] (1969): Markov processes over denumerable products of spaces describing large system of automata. Problems of Information Transmission 5 no. 3, 47-52 (translated from Problemy Peredači Informacii 5 (1969), no. 3, 64-72 (Russian))

Institut für Angewandte Mathematik
Universität Bonn
Wegelerstrasse 6
53115 Bonn
Germany
sturm@uni-bonn.de


[^0]:    ${ }^{1}$ By definition, this means that property (1.1) below holds for each geodesic $\Gamma$ in $\mathcal{P}_{2}(M)$.

