Available online at www.sciencedirect.com

ScienceDirect

Stochastic Processes and their Applications 126 (2016) 2527–2553

stochastic processes and their applications

www.elsevier.com/locate/spa

The gap between Gromov-vague and Gromov–Hausdorff-vague topology

Siva Athreya[a,*], Wolfgang Löhr[b], Anita Winter[b]

[a] *Indian Statistical Institute, 8th Mile Mysore Road, Bangalore 560059, India*
[b] *Fakultät für Mathematik, Universität Duisburg–Essen, Thea-Leymann-Str. 9, 45117 Essen, Germany*

Received 12 April 2015; received in revised form 22 December 2015; accepted 23 February 2016
Available online 8 April 2016

Abstract

In Athreya et al. (2015) an invariance principle is stated for a class of strong Markov processes on tree-like metric measure spaces. It is shown that if the underlying spaces converge Gromov vaguely, then the processes converge in the sense of finite dimensional distributions. Further, if the underlying spaces converge Gromov–Hausdorff vaguely, then the processes converge weakly in path space. In this paper we systematically introduce and study the Gromov-vague and the Gromov–Hausdorff-vague topology on the space of equivalence classes of metric boundedly finite measure spaces. The latter topology is closely related to the Gromov–Hausdorff–Prohorov metric which is defined on different equivalence classes of metric measure spaces.

We explain the necessity of these two topologies via several examples, and close the gap between them. That is, we show that convergence in Gromov-vague topology implies convergence in Gromov–Hausdorff-vague topology if and only if the so-called lower mass-bound property is satisfied. Furthermore, we prove and disprove Polishness of several spaces of metric measure spaces in the topologies mentioned above.

As an application, we consider the Galton–Watson tree with critical offspring distribution of finite variance conditioned to not get extinct, and construct the so-called Kallenberg–Kesten tree as the weak limit in Gromov–Hausdorff-vague topology when the edge length is scaled down to go to zero.

MSC: primary 60B05; 60B10; secondary 05C80; 60B99

Keywords: Metric measure spaces; Gromov-vague topology; Gromov–Hausdorff-vague; Gromov-weak; Gromov–Hausdorff-weak; Gromov–Prohorov metric; Lower mass-bound property; Full support assumption; Coding trees by excursions; Kallenberg tree

* Corresponding author.
E-mail addresses: athreya@isibang.ac.in (S. Athreya), wolfgang.loehr@uni-due.de (W. Löhr), anita.winter@uni-due.de (A. Winter).

http://dx.doi.org/10.1016/j.spa.2016.02.009
0304-4149/

1. Introduction

The paper introduces the *Gromov-vague* and the *Gromov–Hausdorff-vague* topology. These are two notions of convergence of (equivalence classes of) metric boundedly finite measure spaces. These are "localized" versions of the Gromov-weak topology and a topology closely related to the Gromov–Hausdorff–Prohorov topology on (equivalence classes of) metric finite measure spaces.

Gromov-weak convergence and sampling. The Gromov-weak topology originates from the weak topology in the space of probability measures on a fixed metric space. It is an example of a topology which comes with a canonical family of measures and convergence determining test functions. That is, given a complete, separable metric space, (X, r), we denote by $\mathcal{M}_1(X)$ the space of all Borel probability measures on X and by $\bar{\mathcal{C}}(X) := \bar{\mathcal{C}}_{\mathbb{R}}(X)$ the space of bounded, continuous $\mathbb{R}$-valued functions. A sequence of probability measures (μ_n) converges *weakly* to μ in $\mathcal{M}_1(X)$ (abbreviated $\mu_n \Longrightarrow \mu$), as $n \to \infty$, if and only if $\int \mathrm{d}\mu_n\, f \to \int \mathrm{d}\mu\, f$ in $\mathbb{R}$, as $n \to \infty$, for all $f \in \bar{\mathcal{C}}(X)$.

We wish to consider sequence of measures that live on different spaces. In such a case an immediate analogue of bounded continuous functions is not available. To still be in a position to imitate the notion of weak convergence, we rely on the following useful fact: for a sequence (μ_n) in $\mathcal{M}_1(X)$ and $\mu \in \mathcal{M}_1(X)$,

$$\mu_n \underset{n\to\infty}{\Longrightarrow} \mu \quad \text{if and only if} \quad \mu_n^{\otimes\mathbb{N}} \underset{n\to\infty}{\Longrightarrow} \mu^{\otimes\mathbb{N}}. \tag{1.1}$$

Indeed, the *"if"* direction follows by the fact that projections to a single coordinate are continuous. The *"only if"* direction follows as the set of bounded continuous functions $\varphi\colon X^{\mathbb{N}} \to \mathbb{R}$ of the form $\varphi((x_n)_{n\in\mathbb{N}}) = \prod_{i=1}^{N} \varphi_i(x_i)$ for some $N \in \mathbb{N}, \varphi_i : X \to \mathbb{R}, i = 1, \ldots, N$, separates points in $X^{\mathbb{N}}$ and is multiplicatively closed (see, for example, [31, Theorem 2.7] for an argument how to use [29] to conclude from here that integration over such test functions is even convergence determining for measures on $\mathcal{M}_1(X)$).

Consider now the set of bounded continuous functions $\varphi\colon X^{\mathbb{N}} \to \mathbb{R}$ of the following form

$$\varphi = \tilde{\varphi} \circ R^{(X,r)}, \tag{1.2}$$

where $R^{(X,r)}$ denotes the map that sends a vector $(x_n)_{n\in\mathbb{N}} \in X^{\mathbb{N}}$ to the matrix $(r(x_i, x_j))_{1\le i<j} \in \mathbb{R}_+^{\binom{\mathbb{N}}{2}}$ of mutual distances, and $\tilde{\varphi} \in \bar{\mathcal{C}}(\mathbb{R}_+^{\binom{\mathbb{N}}{2}})$ depends on finitely many coordinates only. A *(complete, separable) metric measure space* (X, r, μ) consists of a complete, separable metric space (X, r) and a Borel measure μ on X. Denote by $\mathbb{X}_1$ the space of measure preserving isometry classes of metric spaces equipped with a Borel probability measure. Then for each representative (X, r, μ) of an isometry class $\mathcal{X} \in \mathbb{X}_1$ the image measure $R_*^{(X,r)}\mu^{\otimes\mathbb{N}} = \mu^{\otimes\mathbb{N}} \circ (R^{(X,r)})^{-1} \in \mathcal{M}_1(\mathbb{R}_+^{\binom{\mathbb{N}}{2}})$ is the same and is referred to as the *distance matrix distribution* $\nu^{\mathcal{X}}$ of $\mathcal{X}$. It turns out that if the distance matrix distributions of two metric measure spaces coincide, then the metric measure spaces fall into the same isometry class. This is known as Gromov's reconstruction theorem (compare [23, Chapter $3\frac{1}{2}$]), and suggests to consider the *Gromov-weak topology*, which is the topology induced by the set of functions of the form

$$\Phi(X, r, \mu) = \int_{X^{\mathbb{N}}} \mathrm{d}\mu^{\otimes\mathbb{N}}\, \varphi = \int \mathrm{d}\nu^{\mathcal{X}}\, \tilde{\varphi}, \tag{1.3}$$

where φ is of the form (1.2). As this set is multiplicatively closed we can conclude once more that it is also convergence determining for metric measure spaces on $\mathbb{X}_1$.

The Gromov-weak topology on spaces of metric measure spaces, prescribed by test functions as in (1.3), originates from the work of Gromov in the context of metric geometry, where it is induced by so-called box metrics. In [20] the Gromov-weak topology on complete, separable metric measure spaces was reintroduced via convergence of the functions of the form (1.3), and metrized by the so-called *Gromov–Prohorov metric*. Recently, in [31], it was shown that Gromov's box metric and the Gromov–Prohorov metric are bi-Lipschitz equivalent.

Independently of Gromov's work, however, the idea of proving convergence of random 0-hyperbolic metric measure spaces (that means trees) via "finite dimensional distributions", i.e., with the help of test functions of the form (1.3), has been used before in probability theory. As the land mark we consider [4, Theorem 23], which states Gromov-weak convergence of suitably rescaled Galton–Watson trees towards the so-called Brownian continuum random tree (CRT), where the Galton–Watson trees are associated with an offspring distribution of finite variance, conditioned on a growing number of nodes and equipped with the uniform distribution on its nodes. Other results using test functions which imitate sampling include [20, Theorem 4], where the so-called Λ-coalescent tree is constructed as a Gromov-weak limit of finite trees. Furthermore, distributions of finite samples from metric measure spaces are used in hypothesis testing and for providing confidence intervals in the field of topological data analysis (see, for example, [7,9]).

From Gromov-weak to Gromov–Hausdorff weak convergence. The following embedding result is known from [20, Lemma 5.8]. A sequence $(\mathcal{X}_n)$ converges Gromov-weakly to $\mathcal{X}$ in $\mathbb{X}_1$ if and only if there is a complete, separable metric space (E, d) such that (representatives of) all $\mathcal{X}_n$ and $\mathcal{X}$ can be embedded measure-preserving isometrically into (E, d) in a way that the image measures under the isometries converge weakly to the image limit measure. Using this embedding procedure, we can also define a stronger topology: We say that a sequence $(\mathcal{X}_n)$ converges *Gromov–Hausdorff-weakly* to $\mathcal{X}$ in $\mathbb{X}_1$ if and only if there is a metric space (E, d) such that we can do the above embedding in a way that, additionally, the supports of the measures converge in Hausdorff distance.

This topology is closely related to the one introduced under the name *measured Hausdorff topology* in [17] in the context of studying the asymptotics of eigenvalues of the Laplacian on collapsing Riemannian manifolds, and extended from compact to Heine–Borel measure spaces in [28]. The difference to the Gromov–Hausdorff weak topology is that, instead of the supports, the whole spaces are required to converge in Hausdorff metric topology. This leads to different equivalence classes, and the connection is discussed extensively in Section 5. In probability theory, the measured (Gromov-)Hausdorff topology was reintroduced and further discussed in [16,34], and recently extended in [1] to complete, locally compact length spaces equipped with locally finite measures.

Verification of convergence. As for the Gromov–Hausdorff-weak topology no canonical family of convergence determining functions is available, a key question is how to actually verify convergence in Gromov–Hausdorff-weak topology? According to the definition, first an embedding of the whole sequence into the same metric space must be provided. For random forests there has been the tradition to encode them (if possible) as excursions on compact intervals, and showing then convergence of the associated excursions in the uniform topology. As the map that sends an excursion to a tree-like metric measure space is continuous with respect to the Gromov-(Hausdorff)-weak topology ([2, Proposition 2.9], [31, Theorem 4.8]), convergence statements obtained by re-scaling the associated excursions always imply convergence Gromov–Hausdorff-weakly. This approach has been successfully applied to branching forests with a particular off-spring distribution (see, for example, [14,13,22]). However, except for a few prototype models there is no obvious way to assign to a random graph model an excursion coming from a Markov process. In such a situation, Gromov–Hausdorff convergence and Gromov-weak convergence are shown separately (for example, [24,12,3]), or the scaling results are stated either without the measure, using Gromov–Hausdorff convergence (for example, [30,33,25]), or only in the Gromov-weak topology (for example, [21]).

Closing the gap. It is known that, if all considered metric measure spaces satisfy a (common) uniform volume doubling property, then Gromov-weak and Gromov–Hausdorff-weak topology are the same [36, Corollary 27.27]. "Volume doubling" is a standard property for Riemannian manifolds and regular, self-similar fractals. It is quite restrictive for random spaces, such as random recursive fractals or, important for us, random $\mathbb{R}$-trees. In particular, Aldous's Brownian CRT almost surely does not have the doubling property, as can be seen from the estimates in [11, Theorem 1.3] (see also [15] for stable Lévy trees).

If the uniform volume doubling property fails, Gromov–Hausdorff-weak convergence is in general not implied by Gromov-weak convergence. The gap between Gromov-weak and Gromov–Hausdorff-weak topology, however, sometimes matters a lot.

Important example. We have recently considered in [5] a class of strong Markov processes on natural scale with values in 0-hyperbolic compact metric spaces, which are uniquely determined by their speed measures. We obtained in [5, Theorem 1] an invariance principle which states convergence of such processes in path space provided the underlying metric (speed-)measure spaces converge Gromov–Hausdorff-weakly. If we only assume Gromov-weak convergence, the processes still converge in their finite dimensional distributions, but without the additional convergence of the supports, convergence in paths space fails.

The main goal of the present paper is to close this gap between Gromov–Hausdorff-weak and Gromov-weak topology. We show that provided metric measure spaces converge Gromov-weakly, they also converge Gromov–Hausdorff-weakly if and only if the so-called (global) lower mass-bound property (Definition 3.1) is satisfied. This allows to verify Gromov–Hausdorff weak convergence via the following two steps (Theorem 6.1):

1. Verify convergence of the test functions from (1.3) together with
2. an extra "tightness condition" given by this lower mass-bound property.

The same lower mass function also turns out to be useful for characterizing the metric measure spaces which are compact and Heine–Borel, respectively, and for proving that the subspaces consisting of these metric measure spaces are Lusin spaces but not Polish if equipped with the Gromov-weak topology. The lower mass-bound property also appears in a compactness condition for the Gromov–Hausdorff-weak topology (Corollary 5.7). Furthermore, we also extend the space of complete, separable metric probability measure spaces to complete, separable, metric boundedly finite measure spaces and equip the latter with the so-called Gromov–(Hausdorff)-vague topologies.

Outline. The paper is organized as follows: In Section 2 we recall the Gromov-weak topology on the space of metric finite measure spaces and then use it to define the Gromov-vague topology on the space of metric boundedly finite measure spaces. In Section 3 the global and local lower mass-bound properties are defined and used to characterize compact metric (finite) measure spaces and Heine–Borel metric boundedly finite measure spaces. In Section 4 we characterize

Gromov-vague convergence via isometric embeddings and deduce criteria for Gromov-vague compactness and Gromov-vague tightness, as well as Polishness of the space of metric boundedly finite measure spaces in Gromov-vague topology. Furthermore, we show that the subspaces of all compact and all Heine–Borel spaces, respectively, are Lusin but not Polish. In Section 5 we introduce the stronger Gromov–Hausdorff-vague topology, and clarify its relation to the measured Gromov–Hausdorff topology and the Gromov–Hausdorff–Prohorov metric. We also show that it is a Polish topology on the space of Heine–Borel boundedly finite measure spaces. For the measured Gromov–Hausdorff topology and the Gromov–Hausdorff–Prohorov metric, this means that restricting to spaces with measures of full support yields again a Polish space. In Section 6 we prove our main convergence criterion for Gromov–Hausdorff-weak and -vague convergence. Namely, given convergence in Gromov-weak or Gromov-vague topology, Gromov–Hausdorff-weak or Gromov–Hausdorff-vague convergence is equivalent to the global or local lower mass-bound property, respectively. In Section 7 we consider the construction of trees coded by continuous, transient excursions, and show that the map which sends an excursion to the corresponding metric boundedly finite measure space is continuous with respect to the Gromov–Hausdorff-vague topology. Finally, as an example, we present the Gromov–Hausdorff-vague convergence in distribution of suitably re-scaled finite-variance, critical Galton–Watson trees, which are conditioned on survival, to the so-called *continuum Kallenberg–Kesten tree*.

2. The Gromov-vague topology

In this section we define the (pointed) Gromov-vague topology. We first introduce pointed metric boundedly finite measure spaces, and the subspaces of interest. We recall the pointed Gromov-weak topology on pointed metric finite measure spaces (Definition 2.5). The pointed Gromov-vague topology is then defined based on the Gromov-weak topology via a "localization procedure" (Definition 2.7). We discuss the connection between both topologies (Remark 2.8), and present a perturbation result (Lemma 2.9).

A *(pointed, complete, separable) metric measure space* (X, r, ρ, μ) consists of a complete, separable metric space (X, r), a distinguished point $\rho \in X$ called the *root*, and a Borel measure μ on X. Since all our spaces are pointed, complete and separable, we usually drop these adjectives in the following when referring to metric measure spaces.

Definition 2.1 (*Equivalence of Metric Measure Spaces*)**.** Two metric measure spaces (X, r, ρ, μ) and (X', r', ρ', μ') are said to be *equivalent* if and only if there is an isometry $\phi\colon \operatorname{supp}(\mu) \cup \{\rho\} \to \operatorname{supp}(\mu') \cup \{\rho'\}$ such that $\phi(\rho) = \rho'$ and $\phi_*\mu = \mu'$, where as usual we denote by

$$\phi_*\mu := \mu \circ \phi^{-1} \tag{2.1}$$

the *push forward* of the measure μ under the measurable map ϕ. We denote the equivalence of metric measure spaces by $\cong$. Most of the time, however, we do not distinguish between a metric measure space and its equivalence class.

Recall that a *Heine–Borel space* is a metric space in which every bounded, closed set is compact. A Heine–Borel space is obviously complete, separable and locally compact. We consider the following subclasses of metric measure spaces.

Definition 2.2 ($\mathbb{X}, \mathbb{X}_{\mathrm{HB}}, \mathbb{X}_{\mathrm{c}}$)**.**

1. A metric measure space (X, r, ρ, μ) is called *boundedly finite* if the measure μ is finite on all bounded subsets of X. Let $\mathbb{X}$ be the set of (equivalence classes of) metric boundedly finite measure spaces.
2. $\mathcal{X} \in \mathbb{X}$ is called *Heine–Borel locally finite measure space* if the equivalence class contains a representative $\mathcal{X} = (X, r, \rho, \mu)$ such that (X, r) is a Heine–Borel space. Denote the subspace of Heine–Borel spaces in $\mathbb{X}$ by $\mathbb{X}_{\mathrm{HB}}$.
3. An equivalence class $\mathcal{X} \in \mathbb{X}_{\mathrm{HB}}$ is called *compact metric finite measure space* if it contains a representative $\mathcal{X} = (X, r, \rho, \mu)$ such that (X, r) is a compact space. Denote the subspace of compact spaces in $\mathbb{X}_{\mathrm{HB}}$ by $\mathbb{X}_{\mathrm{c}}$.

We illustrate this definition with an example which is useful for considering continuum limits of trees.

Example 2.3 (*Locally Compact Geodesic Spaces and $\mathbb{R}$-trees*)**.** Recall that a geodesic space is a metric space in which every two points are connected by an isometric path, i.e. a path with length equal to the distance between these points. A geodesic space is called $\mathbb{R}$-tree if there is, up to reparametrization, only one simple path between every pair of points. It is a classical fact that every complete, locally compact geodesic space is a Heine–Borel space. In particular, $\mathbb{X}_{\mathrm{HB}}$ contains the subclass of complete, locally compact $\mathbb{R}$-trees with Radon measures. $\square$

As every Heine–Borel space is locally compact, the local compactness assumption on the geodesic space is obviously essential. The following remark discusses why the completeness assumption is important as well.

Remark 2.4 (*Non-complete Spaces*)**.** We can allow also non-complete spaces as elements of $\mathbb{X}$ by identifying them with their respective completions. Note, however, that Radon measures on non-complete metric spaces are not boundedly finite in general. Consider for example the binary tree $T := \{\rho\} \cup \bigcup_{n\in\mathbb{N}}\{0, 1\}^n$ with edges connecting $w \in \{0, 1\}^n$ with $(w, 0) \in \{0, 1\}^{n+1}$ and $(w, 1) \in \{0, 1\}^{n+1}, n \in \mathbb{N}_0$, equipped with a metric determined by $r(w, (w, 0)) = r(w, (w, 1)) := c^{-n}$ if $w \in \{0, 1\}^n$, for some $c \in [\frac{1}{2}, 1)$, and equipped with the length measure (see, Example 5.15 for a detailed definition). The length measure is indeed a Radon measure as all compact subtrees are contained in a subtree spanned by finitely many vertices. On the other hand (T, r) is bounded, but the length measure is not finite. Thus non-complete, locally compact $\mathbb{R}$-trees with a Radon measure are not elements of $\mathbb{X}$ in general.

Moreover, non-complete, locally compact $\mathbb{R}$-trees with a boundedly finite measure are not elements of $\mathbb{X}_{\mathrm{HB}}$ in general, as their completions do not need to be locally compact. Take for example $T := (0, 1] \times \{0\} \cup \bigcup_{n\in\mathbb{N}}\{\frac{1}{n}\} \times [0, 1] \subseteq \mathbb{R}^2$, and let r be the intrinsic length metric on T (i.e., $r(x, y)$ is the Euclidean length of the shortest path within T connecting x and y). Then (T, r) is a non-complete $\mathbb{R}$-tree, and it is easy to see that it is locally compact. Its completion $\overline{T} = T \cup \{(0, 0)\}$, however, is not locally compact, because $(0, 0)$ does not possess any compact neighborhood. $\square$

We next recall the definition of the (pointed) Gromov-weak topology on metric finite measure spaces (see [20] and [32, Section 2.1] for more details). As with the metric measure spaces, we drop the adjective "pointed" in the following when referring to topologies on spaces of (pointed) metric measure spaces.

Definition 2.5 (*(Pointed) Gromov-weak Topology*). For $m \in \mathbb{N}$, the *m-point distance matrix distribution* of a metric finite measure space $\mathcal{X} = (X, r, \rho, \mu)$ is the finite measure on $\mathbb{R}_+^{\binom{m+1}{2}}$ defined by

$$\nu_m(\mathcal{X}) := \int_{X^m} \mu^{\otimes m}(\mathrm{d}(x_1, \ldots, x_m))\, \delta_{(r(x_i,x_j))_{0\le i<j\le m}}, \tag{2.2}$$

where $x_0 := \rho$ and δ is the Dirac measure. A sequence $(\mathcal{X}_n)_{n\in\mathbb{N}}$ of metric finite measure spaces *converges* to a metric finite measure space $\mathcal{X}$ *Gromov-weakly* if all m-point distance matrix distributions converge, i.e., if

$$\nu_m(\mathcal{X}_n) \underset{n\to\infty}{\Longrightarrow} \nu_m(\mathcal{X}), \tag{2.3}$$

for all $m \in \mathbb{N}$, where we write $\Rightarrow$ for weak convergence of finite measures.

Next we define the Gromov-vague topology on the space $\mathbb{X}$ of metric boundedly finite measure spaces. The construction is a straight-forward “localization” procedure, similar to the one used by Gromov for Gromov–Hausdorff convergence of pointed locally compact spaces (compare [23, Section 3B]).

Given a metric space (X, r), we use the notations $B_r(x, R)$ and $\overline{B}_r(x, R)$ for the open respectively closed ball around $x \in X$ of radius $R \ge 0$. If there is no risk of confusion, we sometimes drop the subscript r. The restriction of a metric measure space $\mathcal{X} = (X, r, \rho, \mu) \in \mathbb{X}$ to the closed ball $\overline{B}(\rho, R)$ of radius $R \ge 0$ around the root is denoted by

$$\mathcal{X}\restriction_R := (X, r, \rho, \mu\restriction_{\overline{B}(\rho,R)}) \cong \big(\overline{B}(\rho, R), r\restriction_{\overline{B}(\rho,R)^2}, \rho, \mu\restriction_{\overline{B}(\rho,R)}\big). \tag{2.4}$$

Generally (and informally), localization works as follows: given a topology on some class of spaces, the localized form of convergence is defined for those spaces $\mathcal{X}$, where for all $R > 0$, the restriction $\mathcal{X}\restriction_R$ falls into the original class. Such spaces converge in the localized topology if, for almost all $R > 0$, the restrictions converge. If d is a metric inducing the original topology, the localized convergence can therefore, for example, be induced by the metric

$$d^{\#}\big(\mathcal{X}, \mathcal{Y}\big) := \int_{\mathbb{R}_+} \mathrm{d}R\, e^{-R}\big(1 \wedge d(\mathcal{X}\restriction_R, \mathcal{Y}\restriction_R)\big). \tag{2.5}$$

We need the following lemma for our definition of Gromov-vague topology. Denote the Gromov–Prohorov metric, which we define in Section 4, by d_{GP}. For the moment, it is enough to know that it induces the Gromov-weak topology by [20, Theorem 5].

Lemma 2.6. *Let $(\mathcal{X}_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbb{X}$ and $\mathcal{X} = (X, r, \rho, \mu) \in \mathbb{X}$. The following are equivalent:*

1. $(\mathcal{X}_n)\restriction_R \xrightarrow[n\to\infty]{} \mathcal{X}\restriction_R$ *Gromov-weakly for all $R > 0$ with $\mu\big(S_r(\rho, R)\big) = 0$, where $S_r(\rho, R) = \overline{B}_r(\rho, R) \setminus B_r(\rho, R)$ is the sphere of radius R around ρ.*
2. $(\mathcal{X}_n)\restriction_R \xrightarrow[n\to\infty]{} \mathcal{X}\restriction_R$ *Gromov-weakly for all but countably many $R > 0$.*
3. $(\mathcal{X}_n)\restriction_R \xrightarrow[n\to\infty]{} \mathcal{X}\restriction_R$ *Gromov-weakly for Lebesgue-almost all $R > 0$.*
4. *There exists a sequence $R_k \to \infty$ such that $(\mathcal{X}_n)\restriction_{R_k} \xrightarrow[n\to\infty]{} \mathcal{X}\restriction_{R_k}$ Gromov-weakly for all $k \in \mathbb{N}$.*
5. $d^{\#}_{\mathrm{GP}}(\mathcal{X}_n, \mathcal{X}) \xrightarrow[n\to\infty]{} 0.$

Proof. The implications "1. ⇒ 2. ⇒ 3. ⇒ 4". are trivial.

"4. ⇒ 1". is a consequence of the Portmanteau theorem. Indeed, assume that $(\mathcal{X}_n)\lceil_{R_k} \xrightarrow[n\to\infty]{} \mathcal{X}\lceil_{R_k}$ Gromov-weakly along a sequence $R_k \to \infty$, and fix $R > 0$. Choose $k \in \mathbb{N}$ large enough such that $R_k \geq R$. Then, for every $m \in \mathbb{N}$, $\nu_m((\mathcal{X}_n)\lceil_{R_k}) \underset{n\to\infty}{\Longrightarrow} \nu_m(\mathcal{X}\lceil_{R_k})$. The first row of the m-point distance matrix ν_m contains, by definition, the distances to the root. Hence $\nu_m((\mathcal{X}_n)\lceil_R)$ is equal to the restriction of $\nu_m((\mathcal{X}_n)\lceil_{R_k})$ to the set of matrices with no entry in the first row exceeding R. The set of these matrices is closed, hence, by the Portmanteau theorem, the condition $\mu\big(S_r(\rho, R)\big) = 0$ implies the claimed convergence.

"3. ⇔ 5". follows directly from the fact that d_{GP} induces the Gromov-weak topology, the definition of $d^{\#}_{\mathrm{GP}}$ in (2.5), and the dominated convergence theorem. □

We are now in a position to define the Gromov-vague topology.

Definition 2.7 (*(Pointed) Gromov-vague Topology*)**.** We say that a sequence $(\mathcal{X}_n)_{n\in\mathbb{N}}$ in $\mathbb{X}$ converges to $\mathcal{X} \in \mathbb{X}$ *Gromov-vaguely* if the equivalent conditions of Lemma 2.6 hold.

Note that usually localized convergence is not strictly a generalization of the original one, because parts can "vanish at infinity" in the limit. For example, consider Gromov–Hausdorff convergence of (pointed) compact metric spaces, and a sequence of two-point spaces, where the distance between the two points tends to infinity. Such a sequence does not converge. In the localized Gromov–Hausdorff topology, however, it converges to the compact space consisting of only one point. A similar phenomenon arises for the Gromov-vague topology.

Remark 2.8 (*Gromov-vague Versus Gromov-weak*)**.** Consider the subspaces $\mathbb{X}_{\mathrm{fin}}$ and $\mathbb{X}_1$ of $\mathbb{X}$, consisting of spaces $\mathcal{X} = (X, r, \rho, \mu)$ where μ is a finite measure, respectively a probability. Then on $\mathbb{X}_1$, the induced Gromov-vague topology coincides with the Gromov-weak topology. On $\mathbb{X}_{\mathrm{fin}}$, and even on $\mathbb{X}_{\mathrm{c}}$, however, this is not the case, because the total mass is not preserved in the Gromov-vague convergence. In fact, for $\mathcal{X} = (X, r, \rho, \mu)$, $\mathcal{X}_n = (X_n, r_n, \rho_n, \mu_n) \in \mathbb{X}_{\mathrm{fin}}$, The Gromov-weak convergence $\mathcal{X}_n \to \mathcal{X}$ is equivalent to $\mathcal{X}_n \to \mathcal{X}$ Gromov-vaguely and $\mu_n(X_n) \to \mu(X)$. □

For a given metric space (X, r), denote by $d^{(X,r)}_{\mathrm{Pr}}$ the Prohorov-metric on the space of all finite measures on $(X, \mathcal{B}(X))$, i.e.,

$$d^{(X,r)}_{\mathrm{Pr}}\big(\mu, \mu'\big) := \inf\big\{\varepsilon > 0 : \ \mu(A) \leq \mu'(A^{\varepsilon}) + \varepsilon, \ \mu'(A) \leq \mu(A^{\varepsilon}) + \varepsilon \ \ \forall A \text{ closed}\big\}, \tag{2.6}$$

where $A^{\varepsilon} = \{x : d(x, A) \leq \varepsilon\}$ is the closed ε-neighborhood of A. Recall that the Prohorov metric induces weak convergence.

We conclude this section with a simple stability property of Gromov-vague convergence under perturbations of the measures in a localized Prohorov sense. We will illustrate this later in Section 7 with Example 5.15.

Lemma 2.9 (*Perturbation of Measures*)**.** *Consider* $\mathcal{X} = (X, r, \rho, \mu)$, $\mathcal{X}_n = (X_n, r_n, \rho_n, \mu_n) \in \mathbb{X}$, *and another sequence of boundedly finite measure* μ'_n *on* X_n, $n \in \mathbb{N}$. *Assume that* $\mathcal{X}_n \xrightarrow[n\to\infty]{} \mathcal{X}$ *Gromov-vaguely, and that there exists a sequence* $R_k \to \infty$ *such that for all* $k \in \mathbb{N}$,

$$\lim_{n\to\infty} d^{(X_n,r_n)}_{\mathrm{Pr}}\big(\mu_n\lceil_{R_k}, \mu'_n\lceil_{R_k}\big) = 0. \tag{2.7}$$

Then $\mathcal{X}'_n := (X_n, r_n, \rho_n, \mu'_n)$ *converges Gromov-vaguely to* $\mathcal{X}$.

Proof. Notice that for every fixed $k, n \in \mathbb{N}$,

$$\lim_{R \downarrow R_k} d_{\text{Pr}}^{(X_n, r_n)}\big(\mu_n\restriction_R, \mu_n\restriction_{R_k}\big) = 0. \tag{2.8}$$

We may therefore assume w.l.o.g. that (2.7) and $(\mathcal{X}_n)\restriction_{R_k} \xrightarrow[n\to\infty]{} \mathcal{X}\restriction_{R_k}$, Gromov-weakly, hold along the same sequence $(R_k)_{k\in\mathbb{N}}$. Thus for any fixed $k \in \mathbb{N}$,

$$(\mathcal{X}_n')\restriction_{R_k} \xrightarrow[n\to\infty]{} \mathcal{X}\restriction_{R_k}, \tag{2.9}$$

Gromov-weakly, by Theorem 5 of [20]. This, however, implies the claimed Gromov-vague convergence. $\square$

3. The lower mass-bound property

In this section we introduce the local and global lower mass-bound properties, and use them to characterize compact spaces and Heine–Borel spaces, respectively. These properties are formulated in terms of the following lower mass functions on the space of metric boundedly finite measure spaces. For $\delta, R > 0$, we define $\mathfrak{m}_\delta^R : \mathbb{X} \to \mathbb{R}_+ \cup \{\infty\}$ as

$$\mathfrak{m}_\delta^R\big((X, r, \rho, \mu)\big) := \inf\big\{\, \mu\big(\overline{B}_r(x, \delta)\big) : x \in B_r(\rho, R) \cap \operatorname{supp}(\mu) \,\big\}, \tag{3.1}$$

with the convention that the infimum of the empty set is ∞ (which may happen if $\rho \notin \operatorname{supp}(\mu)$). Furthermore, set

$$\mathfrak{m}_\delta := \lim_{R\to\infty} \mathfrak{m}_\delta^R = \inf_{R>0} \mathfrak{m}_\delta^R. \tag{3.2}$$

The following property plays an important rôle at several places in later arguments.

Definition 3.1 (*Lower Mass-bound Property*). A set $\mathbb{K} \subseteq \mathbb{X}$ of metric boundedly finite measure spaces satisfies the *local lower mass-bound property* if and only if

$$\inf_{\mathcal{X}\in\mathbb{K}} \mathfrak{m}_\delta^R(\mathcal{X}) > 0, \tag{3.3}$$

for all $R > \delta > 0$. It satisfies the *global lower mass-bound property* if and only if

$$\inf_{\mathcal{X}\in\mathbb{K}} \mathfrak{m}_\delta(\mathcal{X}) > 0, \tag{3.4}$$

for all $\delta > 0$. We say that a single metric measure space $\mathcal{X} \in \mathbb{X}$ satisfies the local/global mass-bound property if and only if $\mathbb{K} := \{\mathcal{X}\}$ does.

Notice that in the definition of $\mathfrak{m}_\delta^R$, we could have replaced the closed ball by an open ball and/or the open ball by a closed ball without changing the conditions (3.3) and (3.4). We made our choice such that $\mathfrak{m}_\delta^R$ is upper semi-continuous, which will be convenient in some proofs.

Lemma 3.2 (*Upper Semi-continuity*). *For every $R, \delta > 0$, the lower mass functions $\mathfrak{m}_\delta^R$ and $\mathfrak{m}_\delta$ are upper semi-continuous with respect to the Gromov-vague topology.*

Proof. Fix $R, \delta > 0$, and let $\mathcal{X}_n = (X_n, r_n, \rho_n, \mu_n) \to \mathcal{X} = (X, r, \rho, \mu)$ be a Gromov-vaguely converging sequence in $\mathbb{X}$. Then we can choose $R' > R + \delta$ such that $\mathcal{X}_n' := \mathcal{X}_n\restriction_{R'}$ converges Gromov-weakly to $\mathcal{X}' := \mathcal{X}\restriction_{R'}$. By Lemma 5.8 of [20], we can assume w.l.o.g. that $X, X_1, X_2, \ldots,$ are subspaces of some metric space (E, d), and $\mu_n' := \mu_n\restriction_{\overline{B}(\rho, R')}$ converges weakly to $\mu' := \mu\restriction_{\overline{B}(\rho,R')}$ on (E,d). We can then find for every $x \in \mathrm{supp}(\mu) \cap B_r(\rho,R)$ a sequence $x_n \to x$ with $x_n \in \mathrm{supp}(\mu_n)$ for all $n \in \mathbb{N}$. Thus

$$\begin{aligned}\mu\big(\overline{B}(x,\delta)\big) &= \inf_{\varepsilon>0} \mu'\big(\overline{B}(x,\delta+\varepsilon)\big)\\ &\geq \inf_{\varepsilon>0} \liminf_{n\to\infty} \mu'_n\big(\overline{B}(x,\delta+\varepsilon)\big)\\ &\geq \liminf_{n\to\infty} \mu_n\big(\overline{B}(x_n,\delta)\big)\\ &\geq \mathfrak{m}_\delta^R(\mathcal{X}_n),\end{aligned}\tag{3.5}$$

where we have applied the Portmanteau theorem in the second step, and used in the last step that $x_n \in B(\rho_n, R)$ for large enough n. Hence $\mathfrak{m}_\delta^R$ is upper semi-continuous. Therefore, $\mathfrak{m}_\delta = \inf_{R>0} \mathfrak{m}_\delta^R$ is also upper semi-continuous. □

Corollary 3.3 (*Lower Mass-bound Property is Preserved Under Closure*). *If $\mathbb{K} \subseteq \mathbb{X}$ satisfies the global or local lower mass-bound property, the same is true for its Gromov-vague closure $\overline{\mathbb{K}}$.*

Lemma 3.4 (*Characterization of Compact mm-spaces*). *Let $\mathcal{X} \in \mathbb{X}$. Then $\mathcal{X}$ is a compact metric finite measure space if and only if it has finite total mass, and satisfies the global lower mass-bound property.*

Proof. "⇒" Assume that $\mathcal{X} = (X, r, \rho, \mu)$ is compact. Then X is bounded, and hence μ is a finite measure. For every $\delta > 0$, the function $x \mapsto \mu\big(B(x,\delta)\big)$ is lower semi-continuous. Therefore, it attains its minimum on the compact set $\mathrm{supp}(\mu)$, and thus the global lower mass-bound property holds.

"⇐" Assume that μ is finite, and that the global lower mass-bound property holds. Then for all $\delta > 0$, we can cover $\mathrm{supp}(\mu)$ with finitely many balls of radius 2δ. To see this, notice that we can choose an at most countable covering $\{B(x,2\delta);\ x \in S \subseteq X\}$ of $\mathrm{supp}(\mu)$ with the property that the points in S have mutual distances at least 2δ. As $\{B(x,\delta);\ x \in S \subseteq X\}$ then consists of pairwise disjoint sets, each carrying μ-mass at least $\mathfrak{m}_\delta(\mathcal{X})$, the total mass of μ is at least $\mathfrak{m}_\delta(\mathcal{X}) \cdot \#S$. As μ is a finite measure, $\{B(x,2\delta);\ x \in S \subseteq X\}$ must be a finite set. Since $\mathrm{supp}(\mu)$ is complete, this means that $\mathrm{supp}(\mu)$ is actually compact. □

Lemma 3.5 (*Characterization of Heine–Borel mm-spaces*). *Let $\mathcal{X} \in \mathbb{X}$. Then $\mathcal{X}$ is a Heine–Borel locally finite measure space if and only if it satisfies the local lower mass-bound property.*

Proof. Given $R > 0$, $\mathcal{X}$ satisfies $\mathfrak{m}_\delta^R(\mathcal{X}) > 0$ for every $\delta > 0$ if and only if $\mathcal{X}\restriction_R$ satisfies the global lower mass-bound property. Hence by Lemma 3.4, $\mathcal{X}\restriction_R$ satisfies the global lower mass-bound property if and only if $\mathcal{X}\restriction_R$ is compact. Obviously, $\mathcal{X}\restriction_R$ is compact for all $R > 0$ if and only if $\mathcal{X}$ is Heine–Borel. □

Corollary 3.6 ($\mathbb{X}_{\mathrm{HB}}$ *and* $\mathbb{X}_{\mathrm{c}}$ *are Measurable*). *Both $\mathbb{X}_{\mathrm{HB}}$ and $\mathbb{X}_{\mathrm{c}}$ are measurable subsets of $\mathbb{X}$ with respect to Borel σ-field generated by the Gromov-vague topology.*

Proof. Notice that

$$\mathbb{X}_{\mathrm{HB}} = \bigcap_{R\in\mathbb{N}} \bigcap_{\delta>0} \bigcup_{a>0} \big\{\mathcal{X} \in \mathbb{X} :\ \mathfrak{m}_\delta^R(\mathcal{X}) \geq a\big\},\tag{3.6}$$

by Lemma 3.5. Since the lower mass functions are upper semi-continuous by Lemma 3.2, $A_{R,\delta,a} := \{\mathcal{X} \in \mathbb{X} : \mathfrak{m}_\delta^R(\mathcal{X}) \geq a\}$ is closed for all $\delta, R > 0$. Hence $\mathbb{X}_{\text{HB}}$ is measurable. The measurability of $\mathbb{X}_\text{c}$ follows analogously by noticing that

$$\mathbb{X}_\text{c} = \bigcap_{\delta>0} \bigcup_{a>0} \{\mathcal{X} \in \mathbb{X} : \mathfrak{m}_\delta(\mathcal{X}) \geq a,\ \mu(X) \leq a^{-1}\}, \tag{3.7}$$

by Lemma 3.4. □

4. Embeddings, compactness and polishness

Recall that weak convergence of finite measures on a complete, separable metric space is induced by the complete Prohorov metric (see, (2.6)). In the same spirit, the Gromov-weak topology is induced by the complete *Gromov–Prohorov metric*, which is defined for two metric finite measure spaces $\mathcal{X} = (X, r, \mu)$ and $\mathcal{X}' = (X', r', \mu')$ by

$$d_{\text{GP}}(\mathcal{X}, \mathcal{X}') := \inf_d d_{\text{Pr}}^{(X \sqcup X', d)}(\mu, \mu'), \tag{4.1}$$

where the infimum is taken over all metrics d on $X \sqcup X'$ that extend both r and r', and $\sqcup$ denotes the disjoint union (see [20, Theorem 5]).

The fact that d_{GP} induces the Gromov-weak topology immediately implies the following embedding result: for every Gromov-weakly convergent sequence, $((X_n, r_n, \mu_n))_{n\in\mathbb{N}}$, there exists a common complete, separable metric space (E, d) in which all (X_n, r_n) can be isometrically embedded such that (the push-forwards of) the measures μ_n converge weakly to a measure μ on (E, d) (compare [20, Lemma 5.8]).

In this section we show that an analogous statement (Proposition 4.1) is true for the Gromov-vague topology and, if the sequence satisfies the local lower mass-bound, (E, d) can be chosen as Heine–Borel space. We will apply this to characterize compact sets in $\mathbb{X}$ (Corollary 4.3), and to show that $\mathbb{X}$ is Polish (Proposition 4.8), while $\mathbb{X}_\text{c}$ and $\mathbb{X}_{\text{HB}}$ are Lusin spaces but not Polish (Corollary 4.9). We also prove a tightness criterion for probability measures on $\mathbb{X}$ (Corollary 4.6).

We start with the embedding result.

Proposition 4.1 (*Characterization via Isometric Embeddings*). *For each $n \in \mathbb{N} \cup \{\infty\}$, let $\mathcal{X}_n = (X_n, r_n, \rho_n, \mu_n) \in \mathbb{X}$. Then $(\mathcal{X}_n)_{n\in\mathbb{N}}$ converges to $\mathcal{X}_\infty$ Gromov-vaguely if and only if there exist a pointed complete, separable metric space (E, d, ρ) and isometries $\varphi_n\colon \text{supp}(\mu_n) \to E$ such that $\varphi_n(\rho_n) = \rho$ for $n \in \mathbb{N} \cup \{\infty\}$, and*

$$\big((\varphi_n)_*\mu_n\big)\restriction_{\overline{B}_d(R,\rho)} \underset{n\to\infty}{\Longrightarrow} \big((\varphi_\infty)_*\mu_\infty\big)\restriction_{\overline{B}_d(R,\rho)}, \tag{4.2}$$

for all but countably many $R \geq 0$. Furthermore, if $\{\mathcal{X}_n : n \in \mathbb{N}\}$ satisfies the local lower mass-bound property, then $\mathcal{X}_\infty \in \mathbb{X}_{\text{HB}}$ and E can be chosen as Heine–Borel space. In this case, (4.2) is equivalent to

$$(\varphi_n)_*\mu_n \xrightarrow{vag} (\varphi_\infty)_*\mu_\infty, \tag{4.3}$$

where $\xrightarrow{vag}$ denotes vague convergence of Radon measures on E.

Before we give the proof, we illustrate with an example that the local lower mass-bound property cannot be dropped without replacement in the second part of the proposition, even if the limit is assumed to be compact.

Example 4.2 (*E is not Heine–Borel Without Lower Mass-bound*)**.** Consider $\mathcal{X}_n = \big([0,1]^n, r_n, 0, \frac{n-1}{n}\delta_{0^n} + \frac{1}{n}\lambda_n\big)$, where r_n is the Euclidean metric, δ_{0^n} is the Dirac measure in $0^n = (0,\dots,0) \in [0,1]^n$, and λ_n is the n-dimensional Lebesgue measure. Then $\mathcal{X}_n$ is compact and obviously converges Gromov-vaguely (and Gromov-weakly) to the compact probability space consisting of only one point, but the embedding space (E,d) cannot be chosen as Heine–Borel space. □

Proof of Proposition 4.1. It is easy to see that (4.2) implies the Gromov-vague convergence, and that if E is a Heine–Borel space, (4.2) is equivalent to (4.3).

Conversely, assume that $\mathcal{X}_n \xrightarrow[n\to\infty]{} \mathcal{X}$ Gromov-vaguely, and abbreviate $X := X_\infty, r := r_\infty$ and $\rho := \rho_\infty$. Let $(R_k)_{k\in\mathbb{N}}$ be an increasing sequence of radii with $\lim_{k\to\infty} R_k = \infty$ and $\mu\big(\overline{B}_r(\rho,R_k)\setminus B_r(\rho,R_k)\big) = 0$. Using that the Gromov–Prohorov metric metrizes the Gromov-weak topology by [20, Theorem 5], we can construct for $n,k \in \mathbb{N}$ a metric $d_{n,k}$ on $X_n \sqcup X$ extending both r_n and r such that for all $l \in \{1,\dots,k\}$,

$$\lim_{n\to\infty} d_{\mathrm{Pr}}^{(X_n\sqcup X, d_{n,k})}\big(\mu_n\restriction_{R_l}, \mu\restriction_{R_l}\big) = 0, \tag{4.4}$$

where we use the abbreviation

$$\mu\restriction_R := \mu\restriction_{\overline{B}_d(\rho,R)}. \tag{4.5}$$

It is easy to check that we can do it such that ρ_n and ρ are identified. Using Cantor's diagonal argument, we can find a subsequence (k_n) such that $d_n := d_{n,k_n}$ satisfies $\lim_{n\to\infty} d_{\mathrm{Pr}}^{(X_n\sqcup X, d_n)}$ $(\mu_n\restriction_{R_k}, \mu\restriction_{R_k}) = 0$, for every $k \in \mathbb{N}$. Let $E := \bigsqcup_{n\in\mathbb{N}\cup\{\infty\}} X_n$, d the largest metric on E which extends all d_n, and $\varphi_n \colon X_n \to E$ the canonical injection. Then it is easy to check that (E,d) is a complete, separable metric space and (4.2) is satisfied.

Now assume that $\{\mathcal{X}_n : n \in \mathbb{N}\}$ satisfies the local lower mass-bound property. Then it is also satisfied for $\big\{\mathcal{X}_n : n \in \mathbb{N}\cup\{\infty\}\big\}$ by Corollary 3.3. Due to Lemma 3.5, we may assume that X_n and X are Heine–Borel spaces. We have to show that E is a Heine–Borel space as well. To this end, we show that every bounded sequence $(x_i)_{i\in\mathbb{N}}$ in E has an accumulation point. If infinitely many x_i are in a single $X_n, n \in \mathbb{N}\cup\{\infty\}$, this follows from the Heine–Borel property of X_n. Therefore, we can assume w.l.o.g. that $x_n \in X_n \cap B_d(\rho, R_k)$ for all n and some k. By (3.3) together with $(\mu_n)\restriction_{R_k} \Rightarrow \mu\restriction_{R_k}$ on E, we obtain $d(x_n, X) \to 0$. Hence there is $y_n \in X$ with $d(x_n,y_n) \to 0$ and, by the Heine–Borel property of X, $(y_n)_{n\in\mathbb{N}}$ has an accumulation point, which is also an accumulation point of $(x_n)_{n\in\mathbb{N}}$. □

From here we can easily characterize the relatively compact sets.

Corollary 4.3 (*Gromov-vague Compactness*)**.** *For a set $\mathbb{K} \subseteq \mathbb{X}$ the following are equivalent:*

1. *$\mathbb{K}$ is relatively compact in $\mathbb{X}$ equipped with the Gromov-vague topology.*
2. *For all $R > 0$, the set of restrictions $\mathbb{K}\restriction_R := \big\{\mathcal{X}\restriction_R : \mathcal{X} \in \mathbb{K}\big\}$ is relatively compact in the Gromov-weak topology.*
3. *$\mathbb{K}\restriction_{R_k}$ is relatively compact in the Gromov-weak topology for a sequence $R_k \to \infty$.*

Furthermore, a set $\mathbb{K} \subseteq \mathbb{X}_{\mathrm{HB}}$ which satisfies the local lower mass-bound property is relatively compact in $\mathbb{X}_{\mathrm{HB}}$ equipped with Gromov-vague topology if and only if the total masses of large balls are uniformly bounded, i.e., for all $R > 0$,

$$\sup_{(X,r,\rho,\mu)\in\mathbb{K}} \mu\big(B_r(\rho,R)\big) < \infty. \tag{4.6}$$

Remark 4.4 (*Gromov-weak Compactness*)**.** Criteria for relative compactness in the Gromov-weak topology are given in Theorem 2 and Proposition 7.1 of [20]. □

Remark 4.5 (*Convergence Without the Lower-mass Bound Property*)**.** As we have seen in Example 4.2, a Gromov-vaguely convergent sequence in $\mathbb{X}_{\mathrm{HB}}$ does not have to satisfy the local lower mass-bound property. Hence the local lower mass-bound property is not necessary for relative compactness in $\mathbb{X}_{\mathrm{HB}}$. □

Proof of Corollary 4.3. "1⇒2" Assume that $\mathbb{K}$ is relatively compact. Let $R > 0$, and consider a sequence $\big(\mathcal{X}_n = (X_n, r_n, \rho_n, \mu_n)\big)_{n\in\mathbb{N}}$ in $\mathbb{K}$. Then it possesses a Gromov-vague limit point $\mathcal{X} \in \mathbb{X}$, and, by passing to a subsequence, we may assume w.l.o.g. that $\mathcal{X}_n \xrightarrow[n\to\infty]{} \mathcal{X}$ Gromov-vaguely. By Proposition 4.1, we can assume that $X_n \subseteq E$ for some separable metric space E, and $\mu'_n := \mu_n{\restriction}_{B(\rho,R')} \underset{n\to\infty}{\Longrightarrow} \mu{\restriction}_{B(\rho,R')} =: \mu'$ for some $R' > R$. By the Prohorov theorem, the sequence $(\mu'_n)_{n\in\mathbb{N}}$ is tight. Thus $\big(\mu_n{\restriction}_{\overline{B}(\rho,R)}\big)_{n\in\mathbb{N}}$ is also tight. We can conclude once more with the Prohorov theorem that $\big(\mu_n{\restriction}_{\overline{B}(\rho,R)}\big)_{n\in\mathbb{N}}$ is relatively weakly compact. Consequently, $(\mathcal{X}_n{\restriction}_R)_{n\in\mathbb{N}}$ has a Gromov-weak limit point.

"2⇒3" is obvious.

"3⇒1" Let $(\mathcal{X}_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbb{K}$. By passing to a subsequence, we may assume that $\mathcal{X}_n{\restriction}_{R_k}$ converges Gromov-weakly to some metric finite measure space $\mathcal{X}^{(k)}$ for all k. Now it is easy to check that $\mathcal{X}^{(i)} = \mathcal{X}^{(k)}{\restriction}_{R_i}$ whenever $R_i \le R_k$ and that we can therefore construct $\mathcal{X} \in \mathbb{X}$ with $\mathcal{X}{\restriction}_{R_k} = \mathcal{X}^{(k)}$ for every $k \in \mathbb{N}$. By definition, $\mathcal{X}_n \to \mathcal{X}$ in Gromov-vague topology.

Now assume that $\mathbb{K} \subseteq \mathbb{X}_{\mathrm{HB}}$ satisfies the local lower mass-bound property and (4.6). Fix $R > 0$. Then for every $\varepsilon > 0$ we can find $N = N(\varepsilon, \mathbb{K}) \in \mathbb{N}$ such that for every $\mathcal{X} = (X, r, \rho, \mu) \in \mathbb{K}$, we can cover $\overline{B}_r(\rho, R)$ by N balls of radius ε. Hence $\mathbb{K}{\restriction}_R$ is relatively compact in Gromov-weak topology by Proposition 7.1 of [20]. Therefore, $\mathbb{K}$ is relatively compact in $\mathbb{X}$ with Gromov-vague topology. As also $\overline{\mathbb{K}}$ satisfies the local lower mass-bound property by Corollary 3.3, $\overline{\mathbb{K}} \subseteq \mathbb{X}_{\mathrm{HB}}$ by Lemma 3.5. □

Having a characterization of compactness at hand, we can also characterize tightness of probability measures on $\mathbb{X}$. Denote by $\mathbb{X}_{\mathrm{fin}}$ the subspace of metric finite measure spaces.

Corollary 4.6 (*Tightness of Measures on $\mathbb{X}$*)**.** *Let Γ be a family of probability measures on $\mathbb{X}$, and consider for each $R > 0$ the restriction map $\psi_R : \mathbb{X} \to \mathbb{X}_{\mathrm{fin}}$ given by $\mathcal{X} \mapsto \mathcal{X}{\restriction}_R$. Then the following are equivalent:*

1. *Γ is Gromov-vaguely tight.*
2. *The set $(\psi_R)_*(\Gamma)$ is Gromov-weakly tight for all $R > 0$.*
3. *For all $R, \varepsilon > 0$, there is a $\delta > 0$ such that*

$$\sup_{\mathbb{P}\in\Gamma} \mathbb{P}\Big\{(X, r, \rho, \mu) \in \mathbb{X} : \mu\big(B_r(\rho, R)\big) > \tfrac{1}{\delta}\Big\} \le \varepsilon, \tag{4.7}$$

$$\sup_{\mathbb{P}\in\Gamma} \mathbb{P}\Big\{(X, r, \rho, \mu) \in \mathbb{X} : \mu\Big\{x \in B_r(\rho, R) : \mu(B_r(x, \varepsilon)) \le \delta\Big\} \ge \varepsilon\Big\} \le \varepsilon.$$

Remark 4.7 (*Gromov-weak Tightness*)**.** A characterization of Gromov-weak tightness of probability measures of metric finite measures spaces is given in Theorem 3 in [20] (compare also [20, Remark 7.2(ii)]). □

Proof of Corollary 4.6. *"only if"* Assume that the family Γ is Gromov-vaguely tight. Then we can find for all $\varepsilon > 0$ a compact set $\mathbb{K}_\varepsilon$ with $\mathbb{P}(\mathbb{K}_\varepsilon) \ge 1-\varepsilon$ for all $\mathbb{P} \in \Gamma$. In particular, by Corollary 4.3, the sets $\mathbb{K}_\varepsilon\restriction_R$ are Gromov-weakly relatively compact for all $R > 0$. Because $(\psi_R)_*(\mathbb{P})(\mathbb{K}_\varepsilon\restriction_R) \ge \mathbb{P}(\mathbb{K}_\varepsilon) \ge 1-\varepsilon$ for all $\mathbb{P} \in \Gamma$, the set $(\psi_R)_*(\Gamma)$ is Gromov-weakly tight.

"if" Conversely if, for all $\varepsilon, R > 0, \mathbb{K}_{\varepsilon,R}$ are Gromov-weakly compact sets satisfying $(\psi_R)_*(\mathbb{P})(K_{\varepsilon,R}) \ge 1-\varepsilon$, then for all $\varepsilon > 0, \mathbb{K}_\varepsilon := \big\{\mathcal{X} \in \mathbb{X} : \mathcal{X}\restriction_n \in \mathbb{K}_{2^{-n}\varepsilon,n}\ \forall n \in \mathbb{N}\big\}$ is a Gromov-vaguely relatively compact set which satisfies $\mathbb{P}(K_\varepsilon) \ge 1-\varepsilon$ for all $\mathbb{P} \in \Gamma$.

The equivalence of (4.7) now follows from Theorem 3 in [20]. $\square$

Constructing a complete metric on $\mathbb{X}$ that metrizes the Gromov-vague topology is now standard.

Proposition 4.8 ($\mathbb{X}$ *is Polish*). *The space $\mathbb{X}$ of metric boundedly finite measure spaces equipped with the Gromov-vague topology is a Polish space.*

Proof. One possible choice of a complete metric is

$$d_{\mathrm{GP}}^{\#}(\mathcal{X},\mathcal{Y}) := \int_{\mathbb{R}_+} \mathrm{d}R\, e^{-R}\big(1 \wedge d_{\mathrm{GP}}(\mathcal{X}\restriction_R, \mathcal{Y}\restriction_R)\big). \tag{4.8}$$

Indeed, that $d_{\mathrm{GP}}^{\#}$ induces the Gromov-vague topology is shown in Lemma 2.6, and separability is obvious. To see completeness, consider a Cauchy sequence $(\mathcal{X}_n)_{n\in\mathbb{N}}$ in $\mathbb{X}$. Then $\mathcal{X}_n\restriction_R$ is a Cauchy sequence with respect to d_{GP} for Lebesgue-almost all $R > 0$. By completeness of $d_{\mathrm{GP}}, \{\mathcal{X}_n\restriction_R : n \in \mathbb{N}\}$ is relatively compact in the Gromov-weak topology for these $R > 0$. By Corollary 4.3, this implies relative compactness of $\{\mathcal{X}_n : n \in \mathbb{N}\}$ in the Gromov-vague topology. Hence the sequence converges Gromov-vaguely. $\square$

Unfortunately, the subspaces $\mathbb{X}_{\mathrm{HB}}$ and $\mathbb{X}_{\mathrm{c}}$ are not Polish, and hence it is impossible to find a complete metric inducing Gromov-vague topology on them. They are, however, Lusin spaces. Recall that a metrizable topological space is, by definition, a Lusin space if it is the image of a Polish space under a continuous, bijective map.

Corollary 4.9 ($\mathbb{X}_{\mathrm{HB}}$ *and* $\mathbb{X}_{\mathrm{c}}$ *are Lusin*). *The space $\mathbb{X}_{\mathrm{HB}}$ of Heine–Borel locally finite measure spaces, equipped with the Gromov-vague topology, is a Lusin space but not Polish. The same is true for the space $\mathbb{X}_{\mathrm{c}}$ of compact metric finite measure spaces.*

Proof. $\mathbb{X}_{\mathrm{HB}}$ and $\mathbb{X}_{\mathrm{c}}$ are measurable subsets (Corollary 3.6) of the Polish space $\mathbb{X}$. Hence they are Lusin by Theorem 8.2.10 of [10].

To see that $\mathbb{X}_{\mathrm{HB}}$ is not Polish, note that it is a dense subspace of $\mathbb{X}$, and using Lemma 3.5 we see that its complement, $\mathbb{X} \setminus \mathbb{X}_{\mathrm{HB}}$, contains a countable intersection of open dense sets, namely $G := \bigcap_{a>0,\, a\in\mathbb{Q}}\big\{\mathcal{X} \in \mathbb{X} : \mathfrak{m}_1^1(\mathcal{X}) < a\big\}$. Such a subspace cannot be Polish by standard arguments (see also [31, Remark 4.7]), which we recall for the reader's convenience. Assume for a contradiction that $\mathbb{X}_{\mathrm{HB}}$ is Polish. By the Mazurkiewicz theorem [10, Theorem 8.1.4], it is a countable intersection of open sets, $\mathbb{X}_{\mathrm{HB}} = \bigcap_{n\in\mathbb{N}} U_n$, say. Obviously, the U_n have to be dense, because $\mathbb{X}_{\mathrm{HB}}$ is. Now $\mathbb{X}_{\mathrm{HB}} \cap G$ is also a countable intersection of open dense sets, hence it is dense by the Baire category theorem [10, Theorem D.37]. This is a contradiction, because $G \subseteq \mathbb{X} \setminus \mathbb{X}_{\mathrm{HB}}$.

The same reasoning also applies to $\mathbb{X}_{\mathrm{c}}$, hence $\mathbb{X}_{\mathrm{c}}$ is also not Polish. $\square$

5. The Gromov–Hausdorff-vague topology

In this section we introduce with the *Gromov–Hausdorff-vague topology* a topology which is stronger than the Gromov-vague topology. The need for such a topology can be motivated by situations as in Example 4.2, and by the fact that there are sequences of finite measures $(\mu_n)_{n\in\mathbb{N}}$ on a common compact space (E,d), such that $\mu_n \Rightarrow \mu$, as $n \to \infty$, but their supports do not converge. The convergence of supports, however, plays a crucial rôle for the convergence of associated random walks to a Brownian motion on the limit space (see [5]). We define the stronger topology based on isometric embeddings, discuss its connection to the related measured (Gromov-)Hausdorff topology and to the Gromov–Hausdorff–Prohorov metric known from the literature, state a stability result, and characterize compact sets. A main result of this section is Polishness of the Gromov–Hausdorff-weak and -vague topologies (Propositions 5.5 and 5.12). Interpreted in terms of the Gromov–Hausdorff–Prohorov metric as used in [34], this means that the subspace of metric measure spaces with full support of the measure is Polish although it is not closed (Corollary 5.6).

We cannot, of course, build such a strong notion of convergence on the notion of sampling alone, and therefore rather use an isometric embedding approach (compare Proposition 4.1). Recall that the Hausdorff distance between two closed subsets A, B of a metric space (E,d) of bounded diameter is defined by

$$d_{\mathrm{H}}\big(A,B\big) := \inf\big\{\varepsilon > 0 : \ A^{\varepsilon} \supseteq B, \ \text{and } B^{\varepsilon} \supseteq A\big\}, \tag{5.1}$$

where once more $A^{\varepsilon} := \{x \in A : d(x,A) \le \varepsilon\}$ denotes the closed ε-neighborhood of A. Recall that $\mathbb{X}_{\mathrm{fin}}$ denotes the space of metric finite measure spaces.

Definition 5.1 (*(Pointed) Gromov–Hausdorff-weak Topology*)**.** Let for each $n \in \mathbb{N}\cup\{\infty\}$, $\mathcal{X}_n := (X_n, r_n, \rho_n, \mu_n) \in \mathbb{X}_{\mathrm{fin}}$. We say that $(\mathcal{X}_n)_{n\in\mathbb{N}}$ converges to $\mathcal{X}_\infty$ in *Gromov–Hausdorff weak topology* if and only if there exists a pointed metric space (E, d_E, ρ_E) and, for each $n \in \mathbb{N}\cup\{\infty\}$, an isometry $\varphi_n : \mathrm{supp}(\mu_n) \to E$ with $\varphi_n(\rho_n) = \rho_E$, and such that in addition to

$$(\varphi_n)_*\mu_n \underset{n\to\infty}{\Longrightarrow} (\varphi_\infty)_*\mu_\infty, \tag{5.2}$$

also

$$d_{\mathrm{H}}\Big(\varphi_n\big(\mathrm{supp}(\mu_n)\big), \ \varphi_\infty\big(\mathrm{supp}(\mu_\infty)\big)\Big) \underset{n\to\infty}{\longrightarrow} 0. \tag{5.3}$$

A very similar topology for compact metric measure spaces was first introduced in [17] under the name *measured Hausdorff topology* (often referred to as measured Gromov–Hausdorff topology) and further discussed in [16,34]. The definition of this topology is exactly the same as that of the Gromov–Hausdorff-weak topology, except that $\mathrm{supp}(\mu_n)$ is replaced by X_n, $n \in \mathbb{N}\cup\{\infty\}$. The consequence is that, when comparing compact metric measure spaces, the geometric structure outside the support is taken into account, while it is ignored by our definition. It is important to note that this leads to different equivalence classes, i.e., the measured Hausdorff topology is *not* defined on $\mathbb{X}_{\mathrm{c}}$, but rather on a space of equivalence classes with respect to the following equivalence relation. We say that two metric measure spaces (X, r, ρ, μ) and (X', r', ρ', μ') are *strongly equivalent* if and only if there is a surjective isometry $\phi : X \to X'$ such that $\phi(\rho) = \rho'$ and $\phi_*(\mu) = \mu'$. Define

$$\mathfrak{X}_c := \{\,\text{strong equivalence classes of compact metric measure spaces}\,\}. \tag{5.4}$$

it is well-known that the measured Hausdorff topology is induced by the so-called *Gromov–Hausdorff–Prohorov metric* defined on $\mathfrak{X}_c$ as follows. For $\mathcal{X} = (X, r, \rho, \mu)$, $\mathcal{X}' = (X', r', \rho', \mu') \in \mathfrak{X}_c$,

$$d_{\mathrm{GHP}}(\mathcal{X}, \mathcal{X}') := \inf_d d_{\mathrm{Pr}}^{(X \sqcup X', d)}(\mu, \mu') + d_{\mathrm{H}}^{(X \sqcup X', d)}(X, X') + d(\rho, \rho'), \tag{5.5}$$

where the infimum is taken over all metrics d on $X \sqcup X'$ that extend both r and r'. Note that $(\mathfrak{X}_c, d_{\mathrm{GHP}})$ is a complete, separable metric space (see [34, Proposition 8]).

Now we can easily identify $\mathbb{X}_c$ with the subspace of $\mathfrak{X}_c$ that consists of all (strong equivalence classes of) compact metric spaces with a measure of full support, i.e. with

$$\mathfrak{X}_c^{\mathrm{supp}} := \big\{ (X, r, \rho, \mu) \in \mathfrak{X}_c : X = \mathrm{supp}(\mu) \big\}, \tag{5.6}$$

by choosing representatives with full support from the larger equivalence classes of $\mathbb{X}_c$, i.e. via the injective map

$$\iota \colon \mathbb{X}_c \to \mathfrak{X}_c^{\mathrm{supp}}, \quad (X, r, \rho, \mu) \mapsto \big(\mathrm{supp}(\mu), r, \rho, \mu\big). \tag{5.7}$$

It is obvious that ι is a homeomorphism if we equip $\mathbb{X}_c$ with the Gromov–Hausdorff-weak and $\mathfrak{X}_c^{\mathrm{supp}}$ with the measured Hausdorff topology. Its inverse ι^{-1} can naturally be extended to all of $\mathfrak{X}_c$, but this extension looses continuity, as we show in the following remark.

Remark 5.2 (*Support Projection*)**.** Equip $\mathfrak{X}_c$ with the measured Hausdorff topology and $\mathbb{X}_c$ with the Gromov–Hausdorff-weak topology. The *support projection*

$$\pi^{\mathrm{supp}} \colon \mathfrak{X}_c \to \mathfrak{X}_c^{\mathrm{supp}}, \quad (X, r, \rho, \mu) \mapsto \big(\mathrm{supp}(\mu), r, \rho, \mu\big) \tag{5.8}$$

is an open map, but neither continuous nor closed. In particular, associating to a strong equivalence class of metric measure spaces in $\mathfrak{X}_c$ the corresponding equivalence class in $\mathbb{X}_c$ is not a continuous operation, although it induces a homeomorphism from $\mathfrak{X}_c^{\mathrm{supp}}$ onto $\mathbb{X}_c$.

Remark 5.3 (*Full Support Assumption*)**.** The requirement that the measure on a metric space has full support is not unnatural. It plays, for instance, a crucial rôle for defining Markov processes via Dirichlet forms (a particular example is [5]), and is even included in the definition of "Radon measure" in [18]. □

Note that $\mathfrak{X}_c^{\mathrm{supp}}$ is not closed in $\mathfrak{X}_c$, hence transporting the Gromov–Hausdorff–Prohorov metric d_{GHP} with ι back to $\mathbb{X}_c$ does not yield a complete metric. The following proposition shows, however, that we can find a different, complete metric for the Gromov–Hausdorff-weak topology on $\mathbb{X}_c$. This also implies that, although $(\mathfrak{X}_c^{\mathrm{supp}}, d_{\mathrm{GHP}})$ is not complete, it can still be used as a Polish state-space, because the induced topological space is Polish. To define the complete metric on $\mathbb{X}_c$, we use the global lower mass function $\mathfrak{m}_\delta$ from (3.2), the Gromov–Hausdorff–Prohorov metric d_{GHP} from (5.5), and the homeomorphism ι from (5.7). Recall that $\mathfrak{m}_\delta > 0$ on $\mathbb{X}_c$ for every $\delta > 0$ by Lemma 3.4.

Definition 5.4. For $\mathcal{X}, \mathcal{X}' \in \mathbb{X}_c$, let

$$d_{\mathrm{sGHP}}(\mathcal{X}, \mathcal{X}') := d_{\mathrm{GHP}}\big(\iota(\mathcal{X}), \iota(\mathcal{X}')\big) + \int_0^1 \mathrm{d}\delta\, 1 \wedge \left| \frac{1}{\mathfrak{m}_\delta(\mathcal{X})} - \frac{1}{\mathfrak{m}_\delta(\mathcal{X}')} \right|. \tag{5.9}$$

We call d_{sGHP} the *support Gromov–Hausdorff–Prohorov metric*.

Proposition 5.5 (($\mathbb{X}_c, d_{sGHP}$) *is a Complete Metric Space*). *The metric d_{sGHP} induces the Gromov–Hausdorff-weak topology on $\mathbb{X}_c$. Furthermore, $(\mathbb{X}_c, d_{sGHP})$ is a complete, separable metric space.*

Proof. Let $(\mathcal{X}_n)_{n\in\mathbb{N}}$ and $\mathcal{X}$ be in $\mathbb{X}_c$. Then $\mathfrak{m}_\delta(\mathcal{X}) > 0$, for all $\delta > 0$. Thus by definition, $d_{sGHP}(\mathcal{X}_n, \mathcal{X}) \xrightarrow[n\to\infty]{} 0$ if and only if

$$d_{GHP}\big(\iota(\mathcal{X}_n), \iota(\mathcal{X})\big) \xrightarrow[n\to\infty]{} 0, \tag{5.10}$$

and for almost all $\delta > 0$,

$$\mathfrak{m}_\delta(\mathcal{X}_n) \xrightarrow[n\to\infty]{} \mathfrak{m}_\delta(\mathcal{X}). \tag{5.11}$$

Because ι is a homeomorphism, (5.10) is equivalent to the Gromov–Hausdorff-weak convergence $\mathcal{X}_n \xrightarrow[n\to\infty]{} \mathcal{X}$. We have to show that this already implies (5.11), i.e. that $\mathcal{X}$ is continuity point of $\mathfrak{m}_\delta$ w.r.t. Gromov–Hausdorff-weak topology for almost all $\delta > 0$. To see this, recall that $\mathfrak{m}_\delta$ is upper semi-continuous w.r.t. Gromov-vague topology (Lemma 3.2), and a fortiori also w.r.t. Gromov–Hausdorff-weak topology. Assume that all $\mathcal{X}_n = (X_n, r_n, \rho_n, \mu_n)$ and $\mathcal{X} = (X, r, \rho, \mu)$ are embedded in some common space (E, d, ρ) such that μ_n converges weakly to μ and $\mathrm{supp}(\mu_n)$ in Hausdorff metric to $\mathrm{supp}(\mu)$. Then, for every $\hat{\delta} < \delta$ and n sufficiently large, every δ-ball around some $y \in \mathrm{supp}(\mu_n)$ contains a $\hat{\delta}$-ball around some $x \in \mathrm{supp}(\mu)$. Therefore, $\liminf_{n\to\infty} \mathfrak{m}_\delta(\mathcal{X}_n) \geq \mathfrak{m}_{\hat{\delta}}(\mathcal{X})$. This means that $\mathfrak{m}_\delta$ is Gromov–Hausdorff-weakly lower semi-continuous in $\mathcal{X}$ for every $\delta > 0$ with $\mathfrak{m}_\delta(\mathcal{X}) = \sup_{\hat{\delta}<\delta} \mathfrak{m}_{\hat{\delta}}(\mathcal{X})$. Because $\delta \mapsto \mathfrak{m}_\delta(\mathcal{X})$ is an increasing function, this is the case for almost all $\delta > 0$. This means that (5.11) is implied by Gromov–Hausdorff-weak convergence, and hence d_{sGHP} induces Gromov–Hausdorff-weak topology as claimed.

That $(\mathbb{X}_c, d_{sGHP})$ is a *separable* metric space is obvious, and it remains to show its *completeness*. Consider a d_{sGHP}-Cauchy sequence $(\mathcal{X}_n)_{n\in\mathbb{N}}$ in $\mathbb{X}_c$. Then, by completeness of d_{GHP} on $\mathfrak{X}_c$, the sequence $\big(\iota(\mathcal{X}_n)\big)_{n\in\mathbb{N}}$ converges in measured Hausdorff topology to some $\mathcal{Y} = (X, r, \rho, \mu) \in \mathfrak{X}_c$. We have to show $\mathcal{Y} \in \iota(\mathbb{X}_c) = \mathfrak{X}_c^{\mathrm{supp}}$. Assume for a contradiction that this is not the case, i.e. there exists $x \in X \setminus \mathrm{supp}(\mu)$. Then there is a $\delta > 0$ with $B(x, 2\delta) \cap \mathrm{supp}(\mu) = \emptyset$. By the measured Hausdorff convergence and the fact that $\iota(\mathcal{X}_n) \in \mathfrak{X}_c^{\mathrm{supp}}$ for all n, this clearly implies $\mathfrak{m}_\delta(\mathcal{X}_n) \xrightarrow[n\to\infty]{} 0$. This, however, cannot be the case because $(\mathcal{X}_n)_{n\in\mathbb{N}}$ is a Cauchy sequence w.r.t. d_{sGHP}. □

Corollary 5.6. *The set $\mathfrak{X}_c^{\mathrm{supp}}$ of (strong equivalence classes of) compact metric full-support measure spaces with the topology induced by the Gromov–Hausdorff–Prohorov metric d_{GHP} is a Polish space (although d_{GHP} restricted to $\mathfrak{X}_c^{\mathrm{supp}}$ is not complete).*

Corollary 5.7 (*Gromov–Hausdorff-weak Compactness*). *A set $\mathbb{K} \subseteq \mathbb{X}_c$ is relatively compact in the Gromov–Hausdorff-weak topology if and only if the following hold*

1. *The set of the total masses is uniformly bounded, i.e.,*

$$\sup_{(X,r,\rho,\mu)\in\mathbb{X}_c} \mu(X) < \infty. \tag{5.12}$$

2. *For all $\varepsilon > 0$ there exists an $N_{R,\varepsilon} \in \mathbb{N}$ such that for all $(X, r, \rho, \mu) \in \mathbb{K}$, $\mathrm{supp}(\mu)$ can be covered by $N_{R,\varepsilon}$ many balls of radius ε.*
3. *$\mathbb{K}$ satisfies the global lower mass-bound property.*

Proof. From Proposition 5.5, the definition of d_{sGHP}, and the fact that d_{GHP} induces the measured Hausdorff topology, we see that $\mathbb{K}$ is Gromov–Hausdorff-weakly relatively compact if and only if $\iota(\mathbb{K})$ is relatively compact in measured Hausdorff topology, and $1/\mathfrak{m}_\delta$ is bounded on $\mathbb{K}$. The latter is obviously equivalent to the global lower mass-bound 3. If $\mathbb{K} \subseteq \mathbb{X}_1$, the measured Hausdorff relative compactness of $\iota(\mathbb{K})$ is equivalent to 2 by [16, Proposition 2.4] together with [8, Theorem 7.4.15]. It is therefore easy to see that it is in general equivalent to 2 together with 1 (compare [20, Remark 7.2(ii)]). □

In the same way as we used the Gromov-weak topology to define the Gromov-vague topology, we also define the Gromov–Hausdorff-vague topology on $\mathbb{X}$ based on the Gromov–Hausdorff-weak topology on $\mathbb{X}_{\mathrm{fin}}$.

Definition 5.8 (*(Pointed) Gromov–Hausdorff-vague Topology*)**.** Let for each $n \in \mathbb{N}\cup\{\infty\}$, $\mathcal{X}_n := (X_n, r_n, \rho_n, \mu_n)$ be in $\mathbb{X}$. We say that $(\mathcal{X}_n)_{n\in\mathbb{N}}$ converges to $\mathcal{X}_\infty$ in *Gromov–Hausdorff vague topology* if and only if $(\mathcal{X}_n)\restriction_R \xrightarrow[n\to\infty]{} (\mathcal{X}_\infty)\restriction_R$ Gromov–Hausdorff-weakly for all but countably many $R > 0$.

The following embedding result and its corollary about Gromov–Hausdorff-vaguely compact sets are proved in the same way as Proposition 4.1 and Corollary 4.3.

Proposition 5.9 (*Isometric Embeddings; Gromov–Hausdorff–Prohorov Metric*)**.** *Let for each $n \in \mathbb{N}\cup\{\infty\}$, $\mathcal{X}_n := (X_n, r_n, \rho_n, \mu_n)$ be in $\mathbb{X}_{\mathrm{HB}}$. The following are equivalent:*

1. $\mathcal{X}_n \xrightarrow[n\to\infty]{} \mathcal{X}_\infty$*, Gromov–Hausdorff vaguely.*
2. *There exists a rooted Heine–Borel space (E, d_E, ρ_E) and for each $n \in \mathbb{N}\cup\{\infty\}$ isometries $\varphi_n : \mathrm{supp}(\mu_n) \to E$ with $\varphi_n(\rho_n) = \rho_E$, and such that in addition to (4.2), also*
$$d_{\mathrm{H}}\Big(\varphi_n(\mathrm{supp}\mu_n) \cap \overline{B}_{d_E}(\rho_E, R),\ \varphi_\infty(\mathrm{supp}\mu_\infty) \cap \overline{B}_{d_E}(\rho_E, R)\Big) \xrightarrow[n\to\infty]{} 0, \tag{5.13}$$
for all but countably many $R > 0$.
3. $d^{\#}_{\mathrm{sGHP}}(\mathcal{X}_n, \mathcal{X}_\infty) \xrightarrow[n\to\infty]{} 0$*, where for $\mathcal{X}, \mathcal{X}' \in \mathbb{X}_{\mathrm{HB}}$,*
$$d^{\#}_{\mathrm{sGHP}}(\mathcal{X}, \mathcal{X}') := \int \mathrm{d}R\, e^{-R}\big(1 \wedge d_{\mathrm{sGHP}}(\mathcal{X}\restriction_R, \mathcal{X}'\restriction_R)\big). \tag{5.14}$$

Corollary 5.10 (*Gromov–Hausdorff-vague Compactness*)**.** *For a set $\mathbb{K} \subseteq \mathbb{X}$ the following are equivalent:*

1. *$\mathbb{K}$ is relatively compact in $\mathbb{X}$ equipped with the Gromov–Hausdorff-vague topology.*
2. *For all $R > 0$, the set of restrictions $\mathbb{K}\restriction_R := \big\{\mathcal{X}\restriction_R : \mathcal{X} \in \mathbb{K}\big\}$ is relatively compact in the Gromov–Hausdorff-weak topology.*
3. *$\mathbb{K}\restriction_{R_k}$ is relatively compact in the Gromov–Hausdorff-weak topology for a sequence $R_k \to \infty$.*

Remark 5.11 (*Gromov–Hausdorff–Prohorov and Length Spaces*)**.** Under the name *Gromov–Hausdorff–Prohorov topology*, the measured Hausdorff topology was recently extended in [1] to the space of complete, locally compact length spaces equipped with locally finite measures. The extension was done with the same localization procedure that we use. Note the following:

1. Complete locally compact length spaces are Heine–Borel spaces and well suited for applications concerning $\mathbb{R}$-trees. The assumption of being a length space and thereby path-connected, however, is too restrictive in general. For example, in Theorem 1 of [5] we establish convergence in path space of continuous time random walks on discrete trees to time-changed Brownian motion on $\mathbb{R}$-trees (appearing as the Gromov–Hausdorff-vague limit of the discrete trees), where the underlying trees are encoded as metric spaces and jump rates and/or time-changes are encoded by the so-called speed measure. Since we need the speed measure to have full support, the situation is incompatible with a connectedness requirement.
2. In a general setting, the name Gromov–Hausdorff–Prohorov topology might be a bit misleading, as “Prohorov” suggests weak convergence, while the localized convergence is vague in the sense that mass can get lost. Also note that, if we drop the assumption of being length spaces, the localized convergence is not really an extension of measured Hausdorff convergence any more (compare Remark 2.8). □

Proposition 5.12 ($\mathbb{X}_{\mathrm{HB}}$ *with Gromov–Hausdorff-vague Topology is Polish*)**.** *The space* $\mathbb{X}_{\mathrm{HB}}$ *of Heine–Borel boundedly finite measure spaces equipped with the Gromov–Hausdorff-vague topology is a Polish space.*

Proof. We follow the proof of Proposition 4.8 and define

$$d^{\#}_{\mathrm{sGHP}}(\mathcal{X},\mathcal{Y}) := \int_{\mathbb{R}_+} \mathrm{d}R\, e^{-R}\big(1 \wedge d_{\mathrm{sGHP}}(\mathcal{X}\restriction_R, \mathcal{Y}\restriction_R)\big). \tag{5.15}$$

We know from Proposition 5.9 that $d^{\#}_{\mathrm{sGHP}}$ induces the Gromov–Hausdorff-vague topology. Separability and completeness follow from the corresponding properties of d_{sGHP} (Proposition 5.5) and the compactness criterion given in Corollary 5.10, in the same way as in the proof of Proposition 4.8. □

Even though the Gromov–Hausdorff-vague topology is nice (i.e. Polish) on $\mathbb{X}_{\mathrm{HB}}$ and defined on all of $\mathbb{X}$, it appears to be too strong to be useful on the larger space.

Remark 5.13 (*Gromov–Hausdorff-vague Topology is Non-separable on* $\mathbb{X}$)**.** The spaces $\mathbb{X}$ and $\mathbb{X}_1$, equipped with the Gromov–Hausdorff-vague topology, are not separable. In particular they are not Lusin spaces. Indeed, we can topologically embed the non-separable space l^{∞}_{+} into $\mathbb{X}_1$ as follows: for $n \in \mathbb{N}$ and $a \in \mathbb{R}_+$, let $A^n_a := \{n\} \times [0,a]^n$, and μ^n_a some measure on A^n_a with full support and total mass 2^{-n}. Define $\psi : l^{\infty}_{+} \to \mathbb{X}_1$ by $\psi(a) := \big(\bigcup_{n\in\mathbb{N}} A^n_{a_n}, r, \rho, \sum_{n\in\mathbb{N}} \mu^n_{a_n}\big)$, where $\rho = (1,0)$, and r is the supremum of the discrete metric on the first component and the Euclidean metric on the second component. It is straightforward to check that ψ is a homeomorphism onto its image. □

We know from Propositions 4.8, 5.5 and 5.12 that $\mathbb{X}$ with Gromov-vague topology, $\mathbb{X}_{\mathrm{c}}$ with Gromov–Hausdorff-weak topology, and $\mathbb{X}_{\mathrm{HB}}$ with Gromov–Hausdorff-vague topology are Polish spaces. Furthermore, it is known from [20, Theorem 1] that $\mathbb{X}_1$ and $\mathbb{X}_{\mathrm{fin}}$ with Gromov-weak topology are Polish. In the case of $\mathbb{X}_1$, this is also true for the Gromov-vague topology, although $\mathbb{X}_1$ is not Gromov-vaguely closed in $\mathbb{X}$ (see Remark 2.8). On the other hand, Corollary 4.9 proves that $\mathbb{X}_{\mathrm{c}}$ and $\mathbb{X}_{\mathrm{HB}}$ with Gromov-vague topology are Lusin but *not* Polish. Similar arguments also show that $\mathbb{X}_{\mathrm{fin}}$ with Gromov-vague topology and $\mathbb{X}_{\mathrm{c}}$ with Gromov-weak as well as with Gromov–Hausdorff-vague topology are Lusin and not Polish. Gromov–Hausdorff-vague topology is not even separable on $\mathbb{X}$ and $\mathbb{X}_1$ by Remark 5.13. We summarize the situation in Fig. 1.

	$\mathbb{X}$	$\mathbb{X}_{\mathrm{HB}}$	$\mathbb{X}_{\mathrm{c}}$	$\mathbb{X}_{\mathrm{fin}}$	$\mathbb{X}_1$	$\mathbb{X}_{\mathrm{c}} \cap \mathbb{X}_1$
Gv	Polish	Lusin	Lusin	Lusin	Polish	Lusin
Gw	–	–	Lusin	Polish	Polish	Lusin
GHv	non-sep.	Polish	Lusin	non-sep.	non-sep.	Lusin
GHw	–	–	Polish	non-sep.	non-sep.	Polish

Fig. 1. The table shows topological properties of different spaces of metric measure spaces in different topologies. Entries: "–" means not defined; "non-sep". means non-separable; "Lusin" means Lusin but *not* Polish. Spaces: the spaces are defined in Definition 2.2 and Remark 2.8. Topologies: Gv = Gromov-vague, Gw = Gromov-weak, GHv = Gromov–Hausdorff-vague, GHw = Gromov–Hausdorff-weak.

We conclude this section with the following stability property, which is the analogue of Lemma 2.9. It is an immediate consequence of the definition of the Gromov–Hausdorff-vague topology by means of isometric embeddings. The proof follows the same lines as the proof of Lemma 2.9 and is therefore omitted.

Lemma 5.14 (*Perturbation of Measures*). *Consider $\mathcal{X} = (X, r, \rho, \mu)$, $\mathcal{X}_n = (X_n, r_n, \rho_n, \mu_n) \in \mathbb{X}$, and another boundedly finite measure μ'_n on $X_n, n \in \mathbb{N}$. Assume that $\mathcal{X}_n \xrightarrow[n\to\infty]{} \mathcal{X}$ Gromov–Hausdorff-vaguely, and that there exists a sequence $R_k \to \infty$ such that for all $k \in \mathbb{N}$,*

$$d_{\mathrm{Pr}}^{(X_n,r_n)}\big(\mu_n\restriction_{R_k}, \mu'_n\restriction_{R_k}\big) \xrightarrow[n\to\infty]{} 0, \quad \textit{and} \quad d_{\mathrm{H}}^{(X_n,r_n)}\big(\mathrm{supp}(\mu_n\restriction_{R_k}), \mathrm{supp}(\mu'_n\restriction_{R_k})\big) \xrightarrow[n\to\infty]{} 0. \tag{5.16}$$

Then $(X_n, r_n, \rho_n, \mu'_n)$ converges Gromov–Hausdorff-vaguely to $\mathcal{X}$.

Example 5.15 (*Normalized Length Measure Versus Degree Measure*). Consider a graph theoretic tree T' which is locally finite, i.e. $\deg(v) < \infty$ for all $v \in T'$, where deg is the degree of a node. Equip T' with the graph distance r', i.e. the length of the shortest path, and fix a root $\rho' \in T'$. Recall the notion of $\mathbb{R}$-tree from Example 2.3. It is well known that (T', r') can be embedded isometrically into a complete, locally compact $\mathbb{R}$-tree (T, r) in an essentially unique way. Denote the image of ρ' by ρ and the image of T' by $\mathrm{nod}(T)$. On T', we consider two natural measures. The *node measure* $\mu_{T'}^{\mathrm{nod}}$, which is just the counting measure on the nodes (except the root), and the *degree measure* $\mu_{T'}^{\mathrm{deg}}$, which is proportional to the degree of the node. The push-forwards on T are given by

$$\mu_T^{\mathrm{nod}} := \sum_{x\in\mathrm{nod}(T)\setminus\{\rho\}} \delta_x \quad \text{and} \quad \mu_T^{\mathrm{deg}} := \tfrac{1}{2} \sum_{x\in\mathrm{nod}(T)} \deg(x)\cdot\delta_x. \tag{5.17}$$

Note that $(T', r', \rho', \mu_{T'}^{\mathrm{deg}}) \cong (T, r, \rho, \mu_T^{\mathrm{deg}})$, and similarly for the node measure. On T, there is also a third natural measure, namely the *length measure* $\lambda = \lambda_{(T,r)}$, which is the 1-dimensional Hausdorff measure on $T \setminus \mathrm{lf}(T)$, where $\mathrm{lf}(T) = \{x \in T : T \setminus \{x\} \text{ is connected}\}$ is the set of leaves of T. Note that $\lambda_{(T,r)}(T) = \mu_T^{\mathrm{nod}}(T) = \mu_T^{\mathrm{deg}}(T)$.

Now consider a sequence $(T'_n)_{n\in\mathbb{N}}$ of locally finite, graph theoretic trees, and the rooted $\mathbb{R}$-trees (T_n, r_n, ρ_n) constructed as above. We assume that there are two sequences $(\alpha_n)_{n\in\mathbb{N}}$ and $(\beta_n)_{n\in\mathbb{N}}$ of positive numbers, both of which converge to 0, such that

$$\big(T_n, \alpha_n r_n, \rho_n, \beta_n\lambda_{(T_n,r_n)}\big) \xrightarrow[n\to\infty]{} \mathcal{X}, \tag{5.18}$$

Gromov–Hausdorff-vaguely for some $\mathcal{X} = (T, r, \rho, \mu) \in \mathbb{X}_{\mathrm{HB}}$, which is necessarily an $\mathbb{R}$-tree. Such a convergence can often be deduced via convergence of excursions, see Proposition 7.5 and

Example 7.6. We claim that in this case, the length measure can be replaced by the degree measure or the node measure, i.e. that (5.18) implies the Gromov–Hausdorff-vague convergences

$$\big(T_n, \alpha_n r_n, \rho_n, \beta_n \mu^{\mathrm{deg}}_{T_n}\big) \xrightarrow[n\to\infty]{} \mathcal{X} \quad \text{and} \quad \big(T_n, \alpha_n r_n, \rho_n, \beta_n \mu^{\mathrm{nod}}_{T_n}\big) \xrightarrow[n\to\infty]{} \mathcal{X}. \tag{5.19}$$

Indeed, we have $\mathrm{supp}(\mu^{\mathrm{nod}}_{T_n}) = \mathrm{supp}(\mu^{\mathrm{deg}}_{T_n}) = \mathrm{nod}(T_n)$, $\mathrm{supp}(\lambda_{(T_n,r_n)}) = T_n$ and, for every $R > 0$,

$$d_{\mathrm{H}}\big(\mathrm{nod}(T_n) \cap B(\rho_n, R),\ B(\rho_n, R)\big) \le \alpha_n \xrightarrow[n\to\infty]{} 0. \tag{5.20}$$

For the Prohorov distance, assume first that the diameter of T_n is smaller than R. Then

$$d^{(T_n,r_n)}_{\mathrm{Pr}}(\mu^{\mathrm{deg}}_{T_n}, \lambda_{(T_n,r_n)}) \le \tfrac{1}{2}\alpha_n \quad \text{and} \quad d^{(T_n,r_n)}_{\mathrm{Pr}}(\mu^{\mathrm{nod}}_{T_n}, \lambda_{(T_n,r_n)}) \le \alpha_n. \tag{5.21}$$

In the general case, we have to take boundary effects into account. Using the annulus $S^{\varepsilon}(\rho_n, R) := \overline{B}(\rho_n, R + \frac{1}{2}\varepsilon) \setminus B(\rho_n, R - \frac{1}{2}\varepsilon)$, we obtain

$$d^{(T_n,r_n)}_{\mathrm{Pr}}(\mu^{\mathrm{deg}}_{T_n}\restriction_R,\ \lambda_{(T_n,r_n)}\restriction_R) \ \le\ \tfrac{1}{2}\alpha_n \vee \beta_n \cdot \lambda_{(T_n,r_n)}\big(S^{\alpha_n}(\rho_n, R)\big), \tag{5.22}$$

and a similar estimate for $\mu^{\mathrm{nod}}_{T_n}$ instead of $\mu^{\mathrm{deg}}_{T_n}$. Using (5.18) we see that $\beta_n \lambda_{(T_n,r_n)}\big(S^{\alpha_n}(\rho_n, R)\big)$ tends to zero for all R with $\mu\big(S(\rho, R)\big) = 0$. Therefore the claimed Gromov–Hausdorff-vague convergences (5.19) follow from (5.22), (5.20) and Lemma 5.14. □

6. Closing the gap

In this section we prove the main criterion for convergence in Gromov–Hausdorff-vague topology. We shall use notation used in (1.2), (1.3) and the definitions of the lower mass functions $\mathfrak{m}^R_\delta$ and $\mathfrak{m}_\delta$ from (3.1) and (3.2), respectively.

In order for a sequence $(\mathcal{X}_n)_{n\in\mathbb{N}} := (X_n, r_n, \rho_n, \mu_n)_{n\in\mathbb{N}}$ of compact metric finite measure spaces to converge in Gromov–Hausdorff-weak topology to a space $\mathcal{X} = (X, r, \rho, \mu) \in \mathbb{X}_{\mathrm{c}}$, it certainly has to converge in the weaker Gromov-weak topology. This is a kind of "finite-dimensional convergence", which is expressible in terms of sampling finite sub-spaces:

1. For all $k \in \mathbb{N}$, and $\varphi \in \bar{\mathcal{C}}(\mathbb{R}_+^{\binom{k+1}{2}})$

$$\int \mu_n^{\otimes k}(\mathrm{d}(x^n_1, \ldots, x^n_k))\, \varphi\big((r_n(x^n_i, x^n_j))_{0\le i<j\le k}\big)$$
$$\xrightarrow[n\to\infty]{} \int \mu^{\otimes k}(\mathrm{d}(x_1, \ldots, x_k))\, \varphi\big((r(x_i, x_j))_{0\le i<j\le k}\big) \tag{6.1}$$

   where we put $x^n_0 := \rho$ and $x_0 := \rho$.

We show in Theorem 6.1 that, given 1, Gromov–Hausdorff-weak convergence follows from a simple "tightness condition", which is given in terms of the lower mass function:

2. For all $\delta > 0$, $\liminf_{n\to\infty} \mathfrak{m}_\delta(\mathcal{X}_n) > 0$.

Note that for checking 1 and 2, we do not have to find any embedding into a common metric space. We actually show that 1 and 2 together are even equivalent to Gromov–Hausdorff-weak convergence, and this characterization even holds if the $\mathcal{X}_n$ are not compact (but $\mathcal{X}$ is).

Theorem 6.1 (*Gromov-weak Versus Gromov–Hausdorff-weak Convergence*)**.** *Let $\mathcal{X} = (X, r, \rho, \mu)$ and $\mathcal{X}_n = (X_n, r_n, \rho_n, \mu_n), n \in \mathbb{N}$, be metric finite measure spaces. Then the following are equivalent.*

1. *$(\mathcal{X}_n)_{n\in\mathbb{N}}$ converges in Gromov-weak topology to $\mathcal{X}$, and for all $\delta > 0$,*
$$\liminf_{n\to\infty} \mathfrak{m}_\delta(\mathcal{X}_n) > 0. \tag{6.2}$$
2. *$\mathcal{X}$ is compact, and $(\mathcal{X}_n)_{n\in\mathbb{N}}$ converges in Gromov–Hausdorff-weak topology to $\mathcal{X}$.*

If $\mathcal{X}_n$ is compact for all $n \in \mathbb{N}$, the following is also equivalent:

3. *$(\mathcal{X}_n)_{n\in\mathbb{N}}$ converges in Gromov-weak topology to $\mathcal{X}$, and $\{\mathcal{X}_n : n \in \mathbb{N}\}$ satisfies the global lower mass-bound property (Definition 3.1).*

Proof. "2⇒1" Assume $(\mathrm{supp}(\mu_n), r_n, \rho_n, \mu_n) \xrightarrow[n\to\infty]{} (\mathrm{supp}(\mu), r, \rho, \mu)$ Gromov–Hausdorff-weakly. W.l.o.g. we may assume that $X = \mathrm{supp}(\mu)$, $X_n = \mathrm{supp}(\mu_n)$, and that X_n and X are embedded into a complete, separable metric space (E, d) such that
$$d_{\mathrm{Pr}}^{(E,d)}\big(\mu_n, \mu\big) \xrightarrow[n\to\infty]{} 0, \quad \text{and} \quad d_{\mathrm{H}}^{(E,d)}\big(X_n, X\big) \xrightarrow[n\to\infty]{} 0. \tag{6.3}$$
Furthermore let (X, r) be compact. We need to show (6.2). Assume to the contrary that there exist $\delta > 0$ and $x_n \in X_n$ such that $\liminf_{n\to\infty} \mu_n\big(B(x_n, 2\delta)\big) = 0$. Due to (6.3) we can find $y_n \in X$ with $d(x_n, y_n) \xrightarrow[n\to\infty]{} 0$. Moreover by (6.3),
$$\liminf_{n\to\infty} \mu\big(B(y_n, \delta)\big) \le \liminf_{n\to\infty} \mu_n\big(B(x_n, 2\delta)\big) = 0. \tag{6.4}$$
As X is compact, we may assume w.l.o.g. that y_n converges to some $y \in X$. Then $\mu\big(B(y, \delta)\big) \le \liminf \mu\big(B(y_n, \delta)\big) = 0$, which contradicts $X = \mathrm{supp}(\mu)$.

"1⇒2" Assume that $\mathcal{X}_n \xrightarrow[n\to\infty]{} \mathcal{X}$ Gromov-weakly, and w.l.o.g. that $X = \mathrm{supp}(\mu)$, $X_n = \mathrm{supp}(\mu_n)$, and that X_n and X are embedded into a complete, separable metric space (E, d) such that $d_{\mathrm{Pr}}^{(E,d)}\big(\mu_n, \mu\big) \xrightarrow[n\to\infty]{} 0$, and that (6.2) holds. Then, for all $\varepsilon > 0$, we can find $n_0 = n_0(\varepsilon) \in \mathbb{N}$ such that for all $n \ge n_0$,
$$d_{\mathrm{Pr}}^{(E,d)}\big(\mu_n, \mu\big) < \varepsilon \wedge \inf_{y\in X_n} \mu_n\big(\overline{B}(y, \varepsilon)\big) \wedge \inf_{x\in X} \mu\big(\overline{B}(x, \varepsilon)\big), \tag{6.5}$$
where we used that
$$\inf_{x\in X} \mu\big(\overline{B}(x, \varepsilon)\big) \ge \liminf_{n\to\infty} \inf_{y\in X_n} \mu_n\big(\overline{B}(y, \tfrac{1}{2}\varepsilon)\big) > 0. \tag{6.6}$$
Then, for all $y \in X_n$, $\overline{B}(y, \varepsilon) \cap X^\varepsilon \ne \emptyset$, and thus also $\overline{B}(y, 2\varepsilon) \cap X \ne \emptyset$. Similarly, $\overline{B}(x, 2\varepsilon) \cap X_n \ne \emptyset$ for all $x \in X$, and hence $d_{\mathrm{H}}^{(E,d)}(X_n, X) \le 2\varepsilon$.

Compactness of $\mathcal{X}$ follows directly from (6.6) and Lemma 3.4.

"3⇔1" Obviously, the global lower mass-bound is equivalent to (6.2) together with $\mathfrak{m}_\delta(\mathcal{X}_n) > 0$ for all $n \in \mathbb{N}$ and $\delta > 0$. The last condition is satisfied for compact spaces by Lemma 3.4. □

The following corollaries are now obvious.

Corollary 6.2 (*Gromov-vague Versus Gromov–Hausdorff-vague Convergence*)**.** *Let $\mathcal{X} = (X, r, \rho, \mu)$ and $\mathcal{X}_n = (X_n, r_n, \rho_n, \mu_n), n \in \mathbb{N}$, be metric boundedly finite measure spaces. Then the following are equivalent.*

1. $(\mathcal{X}_n)_{n\in\mathbb{N}}$ *converges in Gromov-vague topology to* $\mathcal{X}$*, and for all* $\delta > 0$ *and* $R > 0$,
$$\liminf_{n\to\infty} \mathfrak{m}_\delta^R(\mathcal{X}_n) > 0. \tag{6.7}$$
2. $\mathrm{supp}(\mu)$ *is Heine–Borel, and* $\big(\mathrm{supp}(\mu_n), r_n, \mu_n\big)_{n\in\mathbb{N}}$ *converges in Gromov–Hausdorff-vague topology to* $\big(\mathrm{supp}(\mu), r, \mu\big)$.

If $\mathcal{X}_n$ *is Heine–Borel for all* $n \in \mathbb{N}$*, the following is also equivalent:*

3. $(\mathcal{X}_n)_{n\in\mathbb{N}}$ *converges in Gromov-vague topology to* $\mathcal{X}$*, and* $\{\mathcal{X}_n : n \in \mathbb{N}\}$ *satisfies the local lower mass-bound property* (3.3).

Corollary 6.3 (*Polish Subspaces*)**.** *Let* $\mathbb{K} \subseteq \mathbb{X}_{\mathrm{HB}}$ *be a space of Heine–Borel locally finite measure spaces satisfying the local lower mass-bound property* (3.3)*. Then its closure* $\overline{\mathbb{K}}$ *in* $\mathbb{X}$ *(w.r.t. the Gromov-vague topology) is a Polish subspace of* $\mathbb{X}_{\mathrm{HB}}$*. Furthermore, the Gromov-vague topology and the Gromov–Hausdorff-vague topology coincide on* $\overline{\mathbb{K}}$*.*

Corollary 6.4 (*Topologies Agree up to Exceptional Sets*)**.** *Let* $\mathcal{X} = (X, r, \rho, \mu)$ *and* $\mathcal{X}_n = (X_n, r_n, \rho_n, \mu_n)$*,* $n \in \mathbb{N}$*, be in* $\mathbb{X}$*. Then the following are equivalent:*

1. $\mathcal{X}_n \xrightarrow[n\to\infty]{} \mathcal{X}$*, Gromov-vaguely.*
2. *For each* $n \in \mathbb{N}$ *there is* $A_n \subseteq X_n$ *such that* $\mu_n(A_n \cap B_{r_n}(\rho_n, R)) \xrightarrow[n\to\infty]{} 0$ *for all* $R > 0$*, and*
$$\big(X_n \setminus A_n, r_n, \rho_n, \mu_n\restriction_{X_n\setminus A_n}\big) \xrightarrow[n\to\infty]{} \mathcal{X}, \tag{6.8}$$
Gromov–Hausdorff-vaguely.
3. *For each* $n \in \mathbb{N}$ *there is* $A_n \subseteq X_n$ *such that* $\mu_n(A_n \cap B_{r_n}(\rho_n, R)) \xrightarrow[n\to\infty]{} 0$ *for all* $R > 0$*, and* (6.8) *holds Gromov-vaguely.*

7. Application to trees coded by excursions

In this section we consider encodings of trees by means of excursions. To be in a position to consider locally compact rather than just compact $\mathbb{R}$-trees we consider possibly transient excursions, and conclude from uniform convergence on compacta of a sequence of excursions that the corresponding rooted boundedly finite $\mathbb{R}$-trees converge Gromov–Hausdorff-vaguely (Proposition 7.5). As an example we present a representation of the scaling limit of a size-biased Galton–Watson tree (Example 7.6).

Definition 7.1 (*(Transient) Excursions*)**.** A continuous function $e\colon \mathbb{R}_+ \to \mathbb{R}_+$ is called *(continuous) excursion* if $e(0) = 0$ and e is not identically 0. We refer to $\zeta(e) := \sup\big\{s > 0 : e(s) > 0\big\}$ as the *excursion length*, and to $I_e := [0, \zeta(e))$ as the *excursion interval*. If the excursion length is finite, we call the excursion *compactly supported*. If $\lim_{s\to\infty} e(s) = \infty$, the excursion is called *transient*.

Let
$$\mathcal{E} := \big\{e\colon \mathbb{R}_+ \to \mathbb{R}_+ \;\big|\; e \text{ is a continuous excursion}\big\}. \tag{7.1}$$
Given $e \in \mathcal{E}$, we define the pseudo-metric r'_e by letting for all $0 \le s \le t < \zeta(e)$,
$$r'_e(s,t) := e(s) + e(t) - 2 \inf_{u\in[s,t]} e(u). \tag{7.2}$$
We write $s \sim_e t$ if $r'_e(s,t) = 0$. Obviously, $s \sim_e t$ is an equivalence relation.

Definition 7.2 (*Glue Map*)**.** The *glue map* $g\colon \mathcal{E} \to \mathbb{X}$ sends an excursion to the complete, separable, rooted measure $\mathbb{R}$-tree

$$g(e) := (T_e, r_e, \rho_e, \mu_e), \tag{7.3}$$

where $T_e := I_e/\sim_e$, and r_e, μ_e, ρ_e are the push-forwards of r'_e, the Lebesgue measure λ_{I_e}, and 0, respectively, under the canonical projection $\pi_e\colon I_e \to T_e$.

Lemma 7.3 (*Excursions and Associated $\mathbb{R}$-trees*)**.** *Let $e \in \mathcal{E}$.*

1. *If e is compactly supported, then $g(e)$ is a pointed compact finite measure $\mathbb{R}$-tree, i.e. $g(e) \in \mathbb{X}_{\mathrm{c}}$.*
2. *If e is transient, then $g(e)$ is a pointed Heine–Borel boundedly finite measure $\mathbb{R}$-tree, i.e. $g(e) \in \mathbb{X}_{\mathrm{HB}}$.*
3. *If e is neither compactly supported nor transient, then $g(e) \notin \mathbb{X}_{\mathrm{HB}}$.*

Proof. 1. Follows from Lemma 3.1 in [16].

2. Assume that e is transient. Then for all $R > 0$, $\xi_e(R) := \sup\{s \geq 0 : e(s) < R\} < \infty$, and $A_R := \big\{ s \in [0, \infty) : e(s) \leq R \big\}$ is a closed subset of $[0, \xi_e(R)]$, and hence compact. Note that continuity of e implies continuity of the projection π_e. Therefore, $\overline{B}(\rho, R) = \pi_e(A_R)$ is also compact. Moreover, $\mu_e\big(\overline{B}(\rho, R)\big) \leq \xi(R) < \infty$. As any closed and bounded subset of (T_e, r_e) is a closed subset of a closed ball $\overline{B}(\rho, R)$ for some $R > 0$, it is compact as well. Thus $g(e) \in \mathbb{X}_{\mathrm{HB}}$.
3. Assume that e is such that $\zeta(e) = \infty$ but $a := \liminf_{t\to\infty} e(t) < \infty$, and define $b := \limsup_{t\to\infty} e(t)$. In the case $b > a$, (T_e, r_e) is not Heine–Borel (and therefore not locally compact). Indeed, there is an $\varepsilon > 0$ with $a + 3\varepsilon < b$, and an increasing sequence (t_n) in $\mathbb{R}_+$ with $e(t_n) \in [a + 2\varepsilon, a + 3\varepsilon]$ and $\inf_{u\in[t_n,t_{n+1}]} e(u) \leq a + \varepsilon$ for all $n \in \mathbb{N}$. This means that $x_n := \pi_e(t_n)$ defines a sequence of points in $\overline{B}(\rho, a + 3\varepsilon)$ with mutual distances at least ε. In the case $b = a$,
$$\mu_e\big(\overline{B}(\rho, b + 1)\big) = \lambda\big\{ s \in \mathbb{R}_+ : e(s) \leq b + 1 \big\} = \infty, \tag{7.4}$$
which means that μ_e is not boundedly finite. In both cases $g(e) \notin \mathbb{X}_{\mathrm{HB}}$. $\square$

Denote the space of continuous, transient excursions on $\mathbb{R}_+$ by

$$\mathcal{E}_{\mathrm{trans}} := \Big\{e\colon \mathbb{R}_+ \to \mathbb{R}_+ \;\Big|\; e \text{ is continuous, } e(0) = 0,\ \lim_{x\to\infty} e(x) = \infty\Big\}, \tag{7.5}$$

and let for $e \in \mathcal{E}_{\mathrm{trans}}$ and $R > 0$,

$$\xi_e(R) := \sup\{s \geq 0 : e(s) < R\} < \infty \tag{7.6}$$

denote the last visit to height $R > 0$.

Remark 7.4 (*$\mathbb{R}$-trees Under Transient Excursions*)**.** Let $e \in \mathcal{E}_{\mathrm{trans}}$. Then $g(e)$ is a Heine–Borel boundedly finite measure $\mathbb{R}$-tree with precisely one end at infinity, i.e., there is a unique isometry $\varphi\colon [0, \infty) \to T_e$ with $\varphi(0) = \rho_e$. Indeed, the map $\varphi_e := \pi_e \circ \xi_e$ is such an isometry. Assume that ψ is a further such isometry and fix $R > 0$. We show that $\psi(R) = \varphi_e(R)$. Choose $t \in \mathbb{R}_+$ with $\pi_e(t) = \psi(R)$. Because ψ is an isometry, we have $e(t) = R$, and consequently $t \leq \xi_e(R)$. Choose $S > \sup_{u\in[0,\xi_e(R)]} e(u)$ and $s \in \pi_e^{-1}\big(\psi(S)\big)$. Note that $s > \xi_e(R)$ and $e(s) = S$. Therefore $S - R = r_e\big(\psi(S), \psi(R)\big) = S + R - 2\inf_{u\in[t,s]} e(u)$, and hence $\inf_{u\in[t,\xi_e(R)]} e(u) \geq \inf_{u\in[t,s]} e(u) = R$. This implies $r_e\big(\psi(R), \varphi_e(R)\big) = 2R - \inf_{u\in[t,\xi_e(R)]} e(u) = 0$. $\square$

Proposition 7.5 (*Continuity of Glue Map*)**.** *The glue map $g\colon \mathcal{E}_{\text{trans}} \to \mathbb{X}_{\text{HB}}$ is continuous if $\mathcal{E}_{\text{trans}}$ is equipped with the topology of uniform convergence on compacta, and $\mathbb{X}_{\text{HB}}$ with the Gromov–Hausdorff-vague topology.*

Proof. Let $(e_n)_{n\in\mathbb{N}}$ and e in $\mathcal{E}_{\text{trans}}$ be such that $e_n \xrightarrow[n\to\infty]{} e$ uniformly on compacta. Put $\mathbb{W}_+ := \{R \geq 0 : \lambda\{s \geq 0 : e(s) = R\} > 0\}$. Standard arguments show that $\mathbb{W}_+$ is at most countable. Recall $\xi_e(R)$ from (7.6) and note that for all $R > 0$, the map $e \mapsto \xi_e(R)$ is continuous with respect to the uniform topology on compacta. Thus for all $R \in [0,\infty) \setminus \mathbb{W}_+$,

$$e_n\restriction_{[0,\xi_{e_n}(R)]} \xrightarrow[n\to\infty]{} e\restriction_{[0,\xi_e(R)]}, \tag{7.7}$$

which in turn implies that $g(e_n)\restriction_R \to g(e)\restriction_R$ Gromov–Hausdorff-weakly (see, for example, [2, Proposition 2.9]). Therefore $g(e_n) \xrightarrow[n\to\infty]{} g(e)$ Gromov–Hausdorff-vaguely by Definition 5.8. □

We illustrate the usefulness of Proposition 7.5 with an example about the scaling limit of a size-biased branching tree (compare [26,19] for a probabilistic representation of this tree).

Example 7.6 (*Kallenberg–Kesten Tree*)**.** Consider a (discrete time) Galton–Watson tree with a finite variance, mean 1 offspring distribution $p = (p_n)_{n\in\mathbb{N}_0}$. Let T' be the so-called *Kallenberg–Kesten tree*, which is a random graph theoretic tree that is distributed like this Galton–Watson tree conditioned on survival. The simple, nearest neighbor random walk on T', and scaling limits thereof, are of interest because of the "subdiffusive" behavior (see [27,6]). The random walk is associated to the degree measure, defined in Example 5.15, as "speed measure" (see [5, Section 7.4]). As in Example 5.15, we construct the (equivalent) rooted, measured $\mathbb{R}$-tree $(T, r, \rho, \mu_T^{\text{deg}})$, corresponding to T'. In the particular case of a geometric offspring distribution, i.e., $p_n := 2^{-(1+n)}$ for all $n \in \mathbb{N}_0$, we can code the tree with the length measure instead of the degree measure as follows:

$$(T, r, \rho, \lambda_{(T,r)}) \stackrel{\mathcal{L}}{=} g(\tilde{W}), \tag{7.8}$$

where $\stackrel{\mathcal{L}}{=}$ denotes equivalence in law and, for all $t \geq 0$, $\tilde{W}_t := W_t - 2\inf_{s\in[0,t]} W_s$, with a simple random walk path $(W_n)_{n\in\mathbb{N}}$ linearly interpolated. We refer to $\mathcal{T}_{\text{K}}^{\text{geom}(1/2)} := (T, r, \rho, \mu_T^{\text{deg}})$ as the *discrete Kallenberg tree* with geometric offspring distribution.

As W converges, after Brownian rescaling, weakly in path space towards standard Brownian motion B, we have

$$\big(n^{-1}\tilde{W}_{n^2 t}\big)_{t\geq 0} \underset{n\to\infty}{\Longrightarrow} \big(\tilde{B}_t\big)_{t\geq 0}, \tag{7.9}$$

where $\tilde{B}_t := B_t - 2\inf_{s\in[0,t]} B_s$. It is shown in [35] that $(\tilde{B}_t)_{t\geq 0}$ equals in law the unique strong solution of the stochastic differential equation

$$X_t := \tfrac{1}{X_t}\mathrm{d}t + \mathrm{d}B_t,\ \ t > 0, \qquad X_0 = 0. \tag{7.10}$$

Note that this solution is a three dimensional Bessel process (i.e. the radial path of a three dimensional Brownian motion). We refer to $g(X)$ as the *continuum Kallenberg–Kesten tree*, $\mathcal{T}_{\text{K}}$.

Because, almost surely, a realization $e := n^{-1}\tilde{W}_{n^2\cdot}$ has slope $\pm n$ almost everywhere, we have $\mu_e = n^{-1}\lambda_{(T_e,r_e)}$. Hence, by Proposition 7.5, (7.9) implies

$$(T, n^{-1}r, \rho, n^{-2}\lambda_{(T,r)}) \stackrel{\mathcal{L}}{=} g\big(n^{-1}\tilde{W}_{n^2\cdot}\big) \underset{n\to\infty}{\Longrightarrow} g(X) = \mathcal{T}_{\text{K}}, \tag{7.11}$$

Gromov–Hausdorff-vaguely. By Example 5.15, this also implies

$$(T, n^{-1}r, \rho, n^{-2}\mu_T^{\deg}) \underset{n\to\infty}{\Longrightarrow} \mathcal{T}_{\mathrm{K}}, \tag{7.12}$$

Gromov–Hausdorff-vaguely. In words, if we consider the discrete Kallenberg–Kesten tree and rescale the edge length to become n^{-1}, and then equip it with the measure which assigns mass $\frac{1}{2}n^{-2}\deg(x)$ to each branch point x, then this discrete measure tree converges weakly with respect to the Gromov–Hausdorff-vague topology to the continuum Kallenberg–Kesten tree. This implies that the simple, nearest neighbor random walk on the rescaled discrete Kallenberg–Kesten tree converges, if sped up by a factor of n^3, to the Brownian motion on $\mathcal{T}_{\mathrm{K}}$, according to Theorem 1 of [5]. See also Section 7.4 there. □

References

[1] Romain Abraham, Jean-François Delmas, Patrick Hoscheit, A note on the Gromov–Hausdorff–Prokhorov distance between (locally) compact metric measure spaces, Electron. J. Probab. 18 (14) (2013) 1–21.

[2] Romain Abraham, Jean-François Delmas, Patrick Hoscheit, Exit times for an increasing Lévy tree-valued process, Probab. Theory Related Fields 159 (1–2) (2014) 357–403.

[3] Lougi Addario-Berry, Nicolas Broutin, Christina Goldschmidt, Grégory Miermont, The scaling limit of the minimum spanning tree of the complete graph, 2013. arXiv:1301.1664.

[4] David Aldous, The continuum random tree III, Ann. Probab. 21 (1993) 248–289.

[5] Siva Athreya, Wolfgang Löhr, Anita Winter, Invariance principle for variable speed random walks on trees. Ann. Probab., 49 pages, 2015, (in press). arXiv:1404.6290.

[6] Martin T. Barlow, Takashi Kumagai, Random walk on the incipient infinite cluster on trees, Illinois J. Math. 50 (1) (2006) 33–65.

[7] Andrew J. Blumberg, Itamar Gal, Michael A. Mandell, Matthew Pancia, Robust statistics, hypothesis testing, and confidence intervals for persistent homology on metric measure spaces, Found. Comput. Math. 14 (4) (2014) 745–789.

[8] Dmitri Burago, Yuri Burago, Sergei Ivanov, A Course in Metric Geometry, in: Graduate Studies in Mathematics, vol. 33, AMS, Boston, MA, 2001.

[9] Gunnar Carlson, Topological pattern recognition for point cloud data, Acta Numer. 23 (2014) 289–368.

[10] Donald L. Cohn, Measure Theory, Birkhäuser, 1980.

[11] David A. Croydon, Volume growth and heat kernel estimates for the continuum random tree, Probab. Theory Related Fields 140 (2008) 207–238.

[12] Nicolas Curien, Bénédicte Haas, The stable trees are nested, Probab. Theory Related Fields 157 (2013) 847–883.

[13] Thomas Duquesne, A limit theorem for the contour process of conditioned Galton–Watson trees, Ann. Probab. 31 (2003) 996–1027.

[14] Thomas Duquesne, Jean-François Le Gall, Random trees, Lévy processes and spatial branching processes, Astérisque 281 (2002) vi+147.

[15] Thomas Duquesne, Guanying Wang, Exceptionally small balls in stable trees, Bull. Soc. Math. France 142 (2) (2014) 223–254.

[16] Steven N. Evans, Anita Winter, Subtree prune and re-graft: A reversible real-tree valued Markov chain, Ann. Probab. 34 (3) (2006) 918–961.

[17] Kenji Fukaya, Collapsing of Riemannian manifolds and eigenvalues of Laplace operators, Invent. Math. 87 (1987) 517–547.

[18] Masatoshi Fukushima, Yoichi Oshima, Masayoshi Takeda, Dirichlet Forms and Symmetric Markov Processes, second edition, in: de Gruyter Studies in Mathematics, vol. 19, Walter de Gruyter & Co., Berlin, 2011.

[19] Luis G. Gorostiza, Anton Wakolbinger, Persistence criteria for a class of critical branching particle systems in continuous time, Ann. Probab. 19 (1) (1991) 266–288.

[20] Andreas Greven, Peter Pfaffelhuber, Anita Winter, Convergence in distribution of random metric measure spaces (Λ-coalescent measure trees), Probab. Theory Related Fields 145 (1–2) (2009) 285–322.

[21] Andreas Greven, Peter Pfaffelhuber, Anita Winter, Tree-valued resampling dynamics: Martingale problems and applications, Probab. Theory Related Fields 155 (3–4) (2013) 789–838.

[22] Andreas Greven, Lea Popovic, Anita Winter, Genealogy of catalytic branching models, Ann. Appl. Probab. 19 (3) (2009) 1232–1272.

[23] Misha Gromov, Metric Structures for Riemannian and Non-Riemannian Spaces, in: Progress in Mathematics, vol. 152, Birkhäuser Boston Inc., Boston, MA, 1999, Based on the 1981 French original.
[24] Bénédicte Haas, Grégory Miermont, Scaling limits of Markov branching trees with applications to Galton–Watson and random unordered trees, Ann. Probab. 40 (6) (2012) 2589–2666.
[25] Hui He, Matthias Winkel, Invariance principles for pruning processes of Galton-Watson trees, 2014. arXiv:1409.1014.
[26] Olav Kallenberg, Stability of critical cluster fields, Math. Nachr. 77 (1977) 7–43.
[27] Harry Kesten, Subdiffusive behavior of random walk on a random cluster, Ann. Inst. Henri Poincaré Probab. Stat. 22 (4) (1986) 425–487.
[28] Kazhiro Kuwae, Takashi Shioya, Convergence of spectral structure: a functional analytic theory and its application to spectral geometry, Anal. Geom. 11 (4) (2003) 599–673.
[29] Lucien Le Cam, Convergence in distribution of stochastic processes, Univ. Calif. Publ. Stat. 2 (1957) 207–236.
[30] Jean-François Le Gall, The topological structure of scaling limits of large planar maps, Invent. Math. 169 (3) (2007) 621–670.
[31] Wolfgang Löhr, Equivalence of Gromov–Prohorov- and Gromov's $\underline{\square}_\lambda$-metric on the space of metric measure spaces, Electron. Commun. Probab. 18 (17) (2013) 1–10.
[32] Wolfgang Löhr, Guillaume Voisin, Anita Winter, Convergence of bi-measure $\mathbb{R}$-trees and the pruning process, Ann. Inst. Henri Poincaré Probab. Stat. 51 (4) (2015) 1342–1368.
[33] Jean-François Marckert, Grégory Miermont, The CRT is the scaling limit of unordered binary trees, Random Structures Algorithms 38 (4) (2011) 467–501.
[34] Grégory Miermont, Tessellations of random maps of arbitrary genus, Ann. Sci. Éc. Norm. Supér. (4) 42 (5) (2009) 725–781.
[35] James W. Pitman, One-dimensional Brownian motion and the three-dimensional Bessel process, Adv. Appl. Probab. 7 (1975) 511–526.
[36] Cédric Villani, Optimal transport, old and new, in: École d'Été de Probabilités de Saint Flour XXXV-2005, in: Lecture Notes in Mathematics, vol. 1920, Springer-Verlag, 2008.