Non-metricity in the continuum limit of randomly-distributed point defects

Raz Kupferman · **Cy Maor** · **Ron Rosenthal**

Abstract We present a homogenization theorem for isotropically-distributed point defects, by considering a sequence of manifolds with increasingly dense point defects. The loci of the defects are chosen randomly according to a weighted Poisson point process, making it a continuous version of the first passage percolation model. We show that the sequence of manifolds converges to a smooth Riemannian manifold, while the Levi-Civita connections converge to a non-metric connection on the limit manifold. Thus, we obtain rigorously the emergence of a non-metricity tensor, which was postulated in the literature to represent continuous distribution of point defects.

1 Introduction

The study of defects in solids with imperfections is a longstanding theme in material science. One of the prototypical crystalline defects are point defects (see e.g. [14,15]). In crystalline materials, point defects may be caused by vacancies, interstitials, or impurities. In amorphous materials, point defects can be viewed as either a contraction or a dilatation of local equilibrium

R. Kupferman
Institute of Mathematics, The Hebrew University.

C. Maor
Institute of Mathematics, The Hebrew University.
E-mail: cy.maor@mail.huji.ac.il

R. Rosenthal
Department Mathematik, ETH Zürich.

distances between adjacent material elements. Assuming that a defect-free body is modeled by a smooth Euclidean manifold (see e.g. [13]), a body containing isotropic (i.e. ball-shaped) point defects is modeled by a subset $D \subset \mathbb{R}^d$ endowed with a Riemannian metric of the form,

$$\mathfrak{g}_R = \phi \cdot \mathfrak{e}, \qquad \phi(x) = \begin{cases} \xi^2 & |x - x_i| \leq \frac{1}{R} \text{ for some } i = 1, \ldots, m \\ 1 & |x - x_i| > \frac{1}{R} \text{ for every } i = 1, \ldots, m, \end{cases} \tag{1}$$

where $\mathfrak{e}$ is the Euclidean metric on $\mathbb{R}^d$, the points $x_1, \ldots, x_m$ are the centers of the defects, $|\cdot|$ is the Euclidean norm, $1/R$ is the radius of a defect, and ξ is the dilatation factor. In this work we focus on defects of vacancy type, hence $\xi < 1$.

Material scientists are usually interested in materials that contain distributed defects (see e.g. [19,24]). In continuum models, bodies with distributed defects are modeled as smooth manifolds, in which the singularities are smoothed out (or homogenized) and their density is represented by an additional geometric field. For example, bodies with distributed dislocation-type defects have been modeled since the 1950s as Riemannian manifolds endowed with a metrically-consistent, *non-symmetric* flat connections (e.g. [20,1,4]). In this model, the density of the dislocations is represented by the torsion-field of the connection.

A model for distributed point defects has been much more elusive. It has been suggested that bodies with distributed point defects could be modeled as Riemannian manifolds with a flat, symmetric, *non-metric* connection (e.g. [14] p. 300–304 and [18]). There is, however, a big difference between the continuum models of dislocations and point defects. Since the 1950s, there has been a clear rationale—even if not a rigorous derivation—relating dislocations to torsion. We are not aware of a similar rationale relating point defects to non-metricity. In the words of Kröner in his seminal review ([14] p. 304):

> *We are, however, completely aware of the fact that these ... identifications [of point defects and non-metricity] have not the same degree of certainty as the ... identification ... of dislocations and torsion.*

In this paper we present a rigorous analysis of the homogenization of point defects. Similarly to the homogenization of edge-dislocations [11,12], we obtain a manifold endowed with a non-metric connection as a limit of manifolds with distributed defects. Specifically, we consider a sequence of manifolds with increasingly-dense point defects. As the density of the defects tends to infinity, the sequence of locally-smooth manifolds with singularities converges to a smooth Riemannian manifold. The Levi-Civita connection on each manifold in the sequence is, wherever it is defined, the Levi-Civita connection of $\mathbb{R}^d$. When the distribution of the point defects is not uniform, the Levi-Civita connection of $\mathbb{R}^d$ is *inconsistent* with the limit metric, i.e. has a non-zero non-metricity tensor. This is the source of non-metricity in the limit. A surprising feature of our result is that the limit metric (and hence the non-metricity tensor) is not that expected from volume vs. length considerations, see Section 2.1.3.

In this paper, we investigate isotropic distributions of isotropic point defects. That is, the distribution is locally invariant to rotations (isotropy of the distribution), and the defects are ball shaped (isotropy of the defects). A natural way to achieve such a distribution is to randomly select the loci of the defects using a weighted Poisson point process. The precise model, which is detailed in Section 2.2, turns out to be a continuous version of the first passage percolation model on the Euclidean lattice, thus making it an interesting probabilistic model on its own; see Subsection 2.2 for more details.

The structure of this paper is as follows: In Section 2.1 we give a rather informal presentation of the main results (without getting into the probabilistic details), and discuss their geometric and material-science consequences. Section 2.1 is the most relevant for the geometric and material-science-oriented reader. In Section 2.2 we describe the probabilistic model for the distribution of the point defects and discuss its connections to the probabilistic literature. After a list of definitions and notations in Section 2.3, we state the main results in Section 2.4. Since the proofs are rather technical, we provide in Section 3 a sketch of the proof of our main results. The detailed proofs are presented in Sections 4–8.

2 Setting and main results

2.1 Overview of the results and discussion

Let $d \geq 2$ be the dimension and let $\xi \in (0, 1)$ be the dilatation factor of the point defects. Ignoring momentarily the probabilistic details, our main result (Theorem 1) is roughly as follows:

There exists a $u_* > 0$ and a continuous monotonically-decreasing function $\eta : [0, u_*) \to (0, 1]$ such that the following holds: Let $D \subset \mathbb{R}^d$ be a compact d-dimensional manifold with corners. Let $u : D \to [0, u_*)$ be a continuous function. Let $(D, \mathfrak{g}_R)_{R>0}$ be a family of manifolds containing $R^d \cdot \int_D u$ point defects of intensity ξ and radius $1/R$, randomly distributed in D with distribution u. Then $(D, \mathfrak{g}_R)$ converges (in the Gromov-Hausdorff sense) as $R \to \infty$ to the Riemannian manifold $(D, (\eta \circ u) \cdot \mathfrak{e})$.

This result holds also if $\xi = 0$ (i.e., if point defects correspond to the removal of a ball and the identification of its boundary as a single point). The case of $\xi = 0$ involves however semi-metrics, hence to simplify the presentation we will only consider in this subsection the case $\xi > 0$.

The next subsections discuss geometric and material-science consequences of our main theorem.

2.1.1 Non-metricity

The Riemannian (Levi-Civita) connection for each of the manifolds $(D, \mathfrak{g}_R)$ coincides with the Euclidean Levi-Civita connection ∇ of $\mathbb{R}^d$, whenever it is

defined, i.e., everywhere except for the boundaries of the defects. Thus, as $R \to \infty$, the connection converges trivially to the Euclidean connection. If u is not constant (i.e. the point defects are distributed non-uniformly), then the limiting Riemannian metric is not Euclidean, hence the limit connection does not coincide with the Levi-Civita connection of the limit metric. In other words, parallel transport with respect to the limit connection ∇ is not an isometry in the limit manifold.

In fact, one can consider the convergence of Riemannian manifolds with connections,

$$(D, \mathfrak{g}_R, \nabla) \to (D, (\eta \circ u) \cdot \mathfrak{e}, \nabla),$$

in which case Riemannian manifolds with metric connections converge to a Riemannian manifold with a non-metric connection.

If $\eta \circ u$ is differentiable, then the non-metricity tensor Q_{kij} of ∇ is given in coordinates by

$$Q_{kij} = \nabla_k \mathfrak{g}_{ij} = \partial_k \mathfrak{g}_{ij} - \Gamma^l_{ik} \mathfrak{g}_{lj} - \Gamma^l_{jk} \mathfrak{g}_{il} = \partial_k (\eta \circ u) \, \delta_{ij} = \frac{\partial_k (\eta \circ u)}{\eta \circ u} \, \mathfrak{g}_{ij}. \qquad (2)$$

where $\mathfrak{g}_{ij}$ are the coordinate components of the metric $(\eta \circ u) \cdot \mathfrak{e}$, and Γ^l_{jk} are the Christoffel symbols of ∇, which are identically zero.

In particular, the non-metricity tensor is diagonal with respect to the metric. This is consistent with the model presented in [25] for bodies with distributed point defects: Riemannian manifolds with flat, symmetric, non-metric connections with a non-metricity tensor diagonal with respect to the metric (in [25] such manifolds are called Weyl manifolds).

We believe that the fact that the off-diagonal components of the non-metricity tensor are zero is only a result of our choice of *isotropic* point defects (i.e. ball-shaped). For different choices of point defects we expect the emergence of non-diagonal non-metricity tensors; see open questions below.

2.1.2 Curvature

An immediate corollary of our main result is that any subset of $\mathbb{R}^d$ endowed with a conformally-flat Riemannian metric can be obtained as a limit of Euclidean manifolds with point defects.

In particular, the limit manifold can have non-zero curvature, even though the point defects do not carry any curvature charge. This is similar to the case of dislocations, where it is only the limit connection that is flat—not the limit metric (see [12]).

2.1.3 Length-volume inconsistency

The Gromov-Hausdorff convergence of $(D, \mathfrak{g}_R)$ to $(D, (\eta \circ u) \cdot \mathfrak{e})$ as $R \to \infty$ is a convergence of distance functions in metric spaces. Another property that converges as $R \to \infty$ is the measure ν_R induced by the Riemannian metric $\mathfrak{g}_R$.

It weakly converges to the measure $\mu_{\sigma \circ u}$ induced by the Riemannian metric $(\sigma \circ u) \cdot \mathfrak{e}$ on D, where

$$\sigma(u) = \left(e^{-u\kappa_d} + \xi^d(1 - e^{-u\kappa_d})\right)^{1/d},$$

and $\kappa_d = \frac{\pi^{d/2}}{\Gamma(d/2+1)}$ is the volume of the d-dimensional unit ball. Indeed, for constant distribution u, the defects cover as $R \to \infty$ a fraction of $(1 - e^{-u\kappa_d})$ of the manifold; this remains true locally for non-constant u. In other words, the sequence of measured metric spaces $(D, \text{dist}_R^D, \mu_R)$ converges in the measured-Gromov-Hausdorff topology (see Definition 1 below) to the measured metric space $(D, d_{\eta \circ u}^D, \mu_{\sigma \circ u})$, where dist_R^D and $d_{\eta \circ u}^D$ are the intrinsic distance functions induced on D by $\mathfrak{g}_R$ and $(\eta \circ u) \cdot \mathfrak{e}$, respectively. See part 2 of Theorem 1 and Corollary 1 for details.

A naïve guess would be that η is equal to σ. Even though our analysis does not yield an explicit formula for η, we show that, in fact, $\eta < \sigma$ (part 3 of Theorem 1). Hence, the limit metric and the limit measure are inconsistent with each other. This can be viewed as another type of non-metricity, not related to the connection, which, to our knowledge, has not been mentioned in the material-science literature so far.

2.1.4 Open questions

We conclude this section by raising several natural questions awaiting further analysis:

1. In the present work we assume that point defects are spherically-symmetric ("isotropic" point defects). A natural question is what is the limit if one takes non-isotropic point defects, say ellipsoids. Our analysis predicts in the non-isotropic case convergence to some metric space, which is not conformally-Euclidean, as in the isotropic case. It is not clear, however, whether the limit distance function is induced by a Riemannian metric (a plausible alternative would be a Finsler metric).
   If the limit distance function is induced by a Riemannian metric, then we expect the resulting non-metricity tensor not to be diagonal with respect to the metric as in (2). Either way, whether the limit metric is Riemannian or Finsler, this will show that Weyl manifolds (in the sense of [25]) are not the most general model for distributed point defects, as suggested in [25].
2. A similar question arises if the defects are placed on a grid deterministically (or if the distribution that is not isotropic). As the grid spacing tends to zero, we conjecture the appearance of a non-Riemannian limit metric, even if the entire structure is symmetric (say, cubic defects on a cubic grid).
3. This work focuses on the phenomenological question of *describing* bodies with distributed point defects. Another natural question is how do the non-metricity and the length-volume inconsistency manifest in the *mechanical*, or elastic properties of the body. That is, if each manifold $(D, \mathfrak{g}_R)$ represents an elastic body with some elastic energy density related to its metric, what is the elastic energy functional in the $R \to \infty$ limit?

2.2 Probabilistic settings

Let $d \geq 2$. We consider the space of locally finite point measures,

$$\Omega = \left\{ \omega = \sum_{i \geq 0} \delta_{x_i} \; : \; \begin{array}{l} x_i \in \mathbb{R}^d \text{ for all } i \geq 0 \text{ and } \omega(A) < \infty \\ \text{for all compact } A \subset \mathbb{R}^d \end{array} \right\} \tag{3}$$

with its natural σ-algebra $\mathcal{F}$ generated by the evaluation maps $\omega \mapsto \omega(A)$, with A running over all Borel-measurable sets in $\mathbb{R}^d$.

For $x \in \mathbb{R}^d$ and $r \geq 0$, let

$$\overline{B(x,r)} = \{y \in \mathbb{R}^d \; : \; |x - y| \leq r\}$$

be the closed Euclidean ball of radius r around x.

Let $\xi \in [0,1)$. Given $\omega \in \Omega$ and $R > 0$ we denote

$$\mathcal{S}_R(\omega) = \bigcup_{x \in \text{supp}(\omega)} \overline{B(x, 1/R)},$$

and define a Riemannian (semi-)metric on $\mathbb{R}^d$

$$\mathfrak{g}_R(x; \omega) = \begin{cases} \mathfrak{e} & x \notin \mathcal{S}_R(\omega), \\ \xi^2 \cdot \mathfrak{e} & x \in \mathcal{S}_R(\omega), \end{cases} \tag{4}$$

where $\mathfrak{e}$ is the Euclidean metric on $\mathbb{R}^d$. We will often remove ω from the notation when no confusion occurs. For $\xi > 0$, $\mathfrak{g}_R$ is a Riemannian metric, and for $\xi = 0$ it is a semi-metric. Let dist_R be the distance function induced by $\mathfrak{g}_R$, that is

$$\text{dist}_R(x, y; \omega) = \inf\{\text{len}_R(\gamma) \; : \; \gamma \in \Gamma(x, y)\}, \tag{5}$$

where $\Gamma(x, y)$ denotes all the paths from x to y, and $\text{len}_R(\gamma)$ is the length of γ induced by $\mathfrak{g}_R$. If $\xi = 0$, dist_R is a semi-distance.

A note on nomenclature: the term "metric" is commonly used in two different contexts—for a Riemannian metric on a smooth manifold and for a distance function in a metric space. Since the distinction between the two is at the heart of the present work, we will consistently call the first a metric and the second a distance.

Finally, denote by ν_R the measure on $\mathbb{R}^d$ induced by $\mathfrak{g}_R$,

$$\nu_R(A; \omega) = \text{Leb}_d(A \setminus \mathcal{S}_R(\omega)) + \xi^d \, \text{Leb}_d(A \cap \mathcal{S}_R(\omega)), \tag{6}$$

where Leb_d is the d-dimensional Lebesgue measure and $A \subseteq \mathbb{R}^d$ is a Lebesgue measurable set.

The triple $(\mathbb{R}^d, \text{dist}_R, \nu_R)$ is a measured metric space if $\xi > 0$. In order to obtain a measured metric space for $\xi = 0$, we define the equivalence relation,

$$x \overset{\omega, R}{\sim} y \quad \Leftrightarrow \quad x = y \text{ or } x, y \text{ are in the same connected component of } \mathcal{S}_R(\omega). \tag{7}$$

In other words, for every $x \in \text{supp}(\omega)$, we identify all the points in $\overline{B(x, 1/R)}$.

For given $\omega \in \Omega$ and $R > 0$ the equivalence relation yields a measured metric space $(\mathcal{M}_R, \text{dist}_R, \nu_R)$, where

$$\mathcal{M}_R = \mathbb{R}^d / \overset{R}{\sim} . \tag{8}$$

Denoting by $\pi_R : \mathbb{R}^d \to \mathcal{M}_R$ the equivalence class map associated with $\overset{R}{\sim}$ (and with a slight abuse of notation),

$$\text{dist}_R(x, y) := \text{dist}_R(\pi_R^{-1}(x), \pi_R^{-1}(y)), \quad \forall x, y \in \mathcal{M}_R. \tag{9}$$

and

$$\nu_R(A) := \nu_R(\pi_R^{-1}(A)), \quad \forall \text{ Borel set } A \subset \mathcal{M}_R. \tag{10}$$

Note that in (9) both $\pi_R^{-1}(x)$ and $\pi_R^{-1}(y)$ may contain more than one element. However, the distance doesn't depend on the choice of the representatives by Definition (5) of dist_R. We denote by $\mathcal{M}_R'$ the set $\{x \in \mathcal{M}_R : |\pi_R^{-1}(x)| = 1\}$.

Remark 1 In simple words, given $\omega \in \Omega$ and $R > 0$ the measured metric space $\mathcal{M}_R = \mathcal{M}_R(\omega)$ is the measured metric space induced from $\mathbb{R}^d$ by identifying all the points in the closed balls $\overline{B(x_i, 1/R)}$, where x_i are the points in the support of ω.

Similarly to the full space, for a path-connected Lebesgue-measurable subset $D \subset \mathbb{R}^d$, we denote by dist_R^D the intrinsic distance/semi-distance induced by $\mathfrak{g}_R$ on D, that is

$$\text{dist}_R^D = \inf\{\text{len}_R(\gamma) : \gamma \in \Gamma(x, y), \gamma \subset D\}, \tag{11}$$

and denote by ν_R^D the restriction of ν_R to D. For $\xi > 0$, this yields a measured metric space $(D, \text{dist}_R^D, \nu_R^D)$. For $\xi = 0$, we denote

$$D_R = D / \sim_R, \tag{12}$$

obtaining a measured metric space $(D_R, \text{dist}_R^D, \nu_R^D)$, where dist_R^D and ν_R^D are defined (with a slight abuse of notation) as the pullback of dist_R^D and ν_R^D by π_R, similarly to the definitions in (9) and (10).

Note: in order to address at the same time the cases $\xi = 0$ and $\xi > 0$, we will sometimes write D_R and π_R even if $\xi > 0$. In this case $D_R = D$ and π_R is the identity map (when $\xi > 0$, all points in D are distinct).

Given a function $u : \mathbb{R}^d \to [0, \infty)$ and $R > 0$, we denote by $\mathbf{P}_{u,R}$ the probability measure on $(\Omega, \mathcal{F})$ under which ω is a Poisson point process with intensity $R^d u(x) \cdot \text{Leb}_d(dx)$, see [22, Chapter 3] for details on Poisson point processes. We denote by $\mathbf{E}_{u,R}$ the corresponding expectation. Note that as $R \to \infty$, the number of point defects grows like R^d whereas the volume in $\mathbb{R}^d$ of each point defect scales like $1/R^d$, which is why we expect the measure to converge.

We will show below (Lemma 1) that for $\xi = 0$ and for every density function $u : \mathbb{R}^d \to [0, \infty)$ taking values in a compact set of an interval $[0, u_*)$,

$u_* = u_*(d) > 0$, the metric space $\mathcal{M}_R$ is $\mathbf{P}_{u,R}$-a.s. simply-connected and locally isometric to the Euclidean space, everywhere except for a nowhere dense set. This is also trivially true for the case $\xi > 0$, in which $(\mathbb{R}^d, \mathfrak{g}_R)$ is locally isometric to the Euclidean plane at every continuity point of $\mathfrak{g}_R$.

2.2.1 Discussion on the probabilistic model

If the function $u : \mathbb{R}^d \to [0, \infty)$ is spatially invariant (i.e. constant), one can view our model as a continuous version of first passage percolation on $\mathbb{Z}^d$, in which one associate a random weight with each edge independently according to some weight distribution F, see for example [8]. Our specific choice of points defects is related to the function $F(\xi) = p$ and $F(1) = 1-p$ for a certain choice of p that depends on u. Note however that one can easily generalize the model to general weights by sampling the dilation factor of each ball randomly according to some distribution.

The discrete model of first passage percolation was introduced by Hammersley and Welsh [7], and was studied extensively since then; see e.g. [8,5,9, 2] and the references therein. The continuous version of this model is based on its continuous counterpart for percolation (known as the boolean model); see [6,17]. To the best of our knowledge, there is no existing work on a continuous version of first passage percolation.

For a general continuous function u, the discrete counterpart of the model is first passage percolation with independent but not identically distributed edge weights.

2.3 Notations and definitions

In this subsection we list notations and definitions that will be used throughout the paper.

2.3.1 Norms

- $|\cdot|$ – The inner-product (Euclidean) norm on $\mathbb{R}^d$.
- $|\cdot|_\infty$ – The supremum norm on $\mathbb{R}^d$.
- $\|\cdot\|_\infty$ – The supremum norm on $C(K)$ for some compact set K.

2.3.2 Riemannian metrics

- $\mathfrak{e}$ – The Euclidean metric on $\mathbb{R}^d$.
- $\mathfrak{g}_R = \mathfrak{g}_R(\cdot; \omega)$ – The (random) metric on $\mathbb{R}^d$ induced by point defects according to (4) (or equivalently (1)).

2.3.3 Distance functions

- Given a continuous function $\rho : \mathbb{R}^d \to (0, \infty)$,

$$d_\rho = \begin{array}{l} \text{the distance function induced} \\ \text{by the Riemannian metric } \rho(x) \cdot \mathfrak{e}. \end{array} \qquad (13)$$

- For a path-connected compact subset $D \subset \mathbb{R}^d$,

$$d_\rho^D = \begin{array}{l} \text{the intrinsic distance function induced} \\ \text{by the Riemannian metric } \rho(x) \cdot \mathfrak{e} \text{ on } D. \end{array} \qquad (14)$$

- $\text{dist}_R = \text{dist}_R(\cdot; \omega)$ – The distance function induced by $\mathfrak{g}_R$ on $\mathbb{R}^d$ or $\mathcal{M}_R$ (see (5) and (9)).
- dist – a shortened notation for dist_1 used in Sections 4–6.
- $\text{dist}_R^D = \text{dist}_R^D(\cdot; \omega)$ – The intrinsic distance function induced by $\mathfrak{g}_R$ on a path connected Lebesgue-mesurable set $D \subset \mathbb{R}^d$, or on D_R (see (11) and (12)).
- For two compact sets $A, B \subset \mathbb{R}^d$, we denote by $d_H(A, B)$ their Hausdorff distance with respect to the Euclidean metric on $\mathbb{R}^d$,

$$d_H(A, B) = \inf \left\{ \varepsilon > 0 \ : \ \begin{array}{l} \forall x \in A \, \exists y \in B, \text{ s.t. } |x - y| < \varepsilon \text{ and} \\ \forall x \in B \, \exists y \in A, \text{ s.t. } |x - y| < \varepsilon \end{array} \right\}.$$

- For compact metric spaces X and Y, we denote by $d_{GH}(X, Y)$ their Gromov-Hausdorff distance (see Definition 1 below).

2.3.4 Measures

- Leb_d – The Lebesgue measure on R^d.
- μ_ρ – The measure induced on $\mathbb{R}^d$ by the volume form of the Riemannian metric $\rho \cdot \mathfrak{e}$.
- μ_ρ^D – The restriction of μ_ρ to D.
- $\nu_R = \nu_R(\cdot; \omega)$ – The measure induced by $\mathfrak{g}_R$ on $\mathbb{R}^d$ or $\mathcal{M}_R$ (see (6) and (10)).
- $\nu_R^D = \nu_R^D(\cdot; \omega)$ – The restriction of ν_R to a subset $D \subset \mathbb{R}^d$ (or to D_R).

2.3.5 Other notations

- $B(x, r)$ – The (open) Euclidean ball or radius $r > 0$ around $x \in \mathbb{R}^d$.
- $\overline{A}$ – The closure of a set $A \subset \mathbb{R}^d$.
- $\text{diam}_\mathfrak{e}(A) = \sup\{|x - y| \ : \ x, y \in A\}$ – The Euclidean diameter of a set $A \subset \mathbb{R}^d$.
- S^{d-1} – The Euclidean sphere $\{x \in \mathbb{R}^d \ : \ |x| = 1\}$.
- $SO(d)$ – The special orthogonal group in $\mathbb{R}^d$ with respect to the Euclidean inner-product.
- $\kappa_d = \frac{\pi^{d/2}}{\Gamma(d/2+1)}$ – The volume of the d-dimensional unit ball.
- By a path in $\mathbb{R}^d$ we will always mean a continuous parametrized path $[0, 1] \to \mathbb{R}^d$.

- $\mathrm{len}_e(\gamma)$ – The Euclidean length of a path γ.
- $\mathrm{len}_R(\gamma) = \mathrm{len}_R(\gamma; \omega)$ – The length of a path γ induced by $\mathfrak{g}_R$.
- $\mathrm{len}(\gamma)$ – a shortened notation for $\mathrm{len}_1(\gamma)$ used in Sections 4–6. .
- $[x, y]$ – The linear segment connecting x and y in $\mathbb{R}^d$.
- $\mathrm{Lip}(f)$ – The Lipschitz constant of a continuous function $f : D \to \mathbb{R}$ for some $D \subset \mathbb{R}^d$.
- $C(D)$ - The space of continuous functions on D.

2.3.6 Measured Gromov-Hausdorff convergence

Definition 1 In the following (Z, d_Z) is a compact metric space and (Z, d_Z, μ_Z) is a compact measured metric space.

1. For a function $f : (X, d_X) \to (Y, d_Y)$ between two metric spaces, we define the **distortion** of f by

$$\mathrm{dis}\, f = \sup_{x,y \in X} |d_X(x, y) - d_Y(f(x), f(y))|. \tag{15}$$

   A function $f : D(f) \subset X \to Y$ is called an ε-**approximation** if $\mathrm{dis}\, f < \varepsilon$, $B_X(D(f), \varepsilon) = X$ and $B_Y(\mathrm{Im}(f), \varepsilon) = Y$, where $B_X(A, \varepsilon)$ is the ε-neiborhood in X around a subset A, with a similar definition for neighborhoods of sets in Y.
2. The **Gromov-Hausdorff distance** is a distance function between isometry classes of compact metric spaces. For the purpose of this paper, it is enough state that $d_{GH}((X, d_X), (Y, d_Y)) < 4\varepsilon$ if there exists an ε-approximation $f : D(f) \subset X \to Y$. For further details see e.g. [21], Chapter 10.
3. A sequence (X_n, d_n, μ_n) of compact measured metric spaces converges to a compact measured metric space (X, d, μ) in the **measured Gromov-Hausdorff topology** if there exists a sequence $f_n : D(f_n) \subset X_n \to X$ of ε_n-approximations, with $\varepsilon_n \to 0$, such that the pushforward measures $(f_n)_\# \mu_n$ weakly converge to μ. That is, for every continuous function $\Phi : X \to \mathbb{R}$,

$$\lim_{n \to \infty} \int_{D(f_n)} \Phi \circ f_n \, d\mu_n = \int_X \Phi \, d\mu.$$

2.3.7 Remark about constants

Throughout the paper, constants are denoted by C and c. The dependence of constants on parameters will be denoted by brackets. For example, $C = C(d)$ implies that C only depends on the dimension d. Note that the value of such constants may change from one line to the next. Numbered constants $C_1, C_2, \ldots$ have a fixed value which is determined in their first appearance.

2.4 Main results

Our main result is the following:

Theorem 1 *Let $d \geq 2$. There exists a real number $u_* > 0$ depending only on d, and there exists a continuous, non-constant, monotonically-decreasing function $\eta : [0, u_*) \to (0, 1]$ with $\eta_0 = 1$ that depends only on d and $\xi \in [0, 1)$, such that for every compact d-dimensional submanifold $D \subset \mathbb{R}^d$ with corners and every continuous function $u : D \to [0, u_*)$:*

1. *Metric convergence: For every $\varepsilon > 0$,*

$$\lim_{R \to \infty} \mathbf{P}_{u,R} \left(\sup_{x,y \in D} \left| d_{\eta \circ u}^D(x, y) - \operatorname{dist}_R^D(\pi_R(x), \pi_R(y)) \right| < \varepsilon \right) = 1. \qquad (16)$$

2. *Measure convergence: For every $\varepsilon > 0$,*

$$\lim_{R \to \infty} \mathbf{P}_{u,R} \left(\forall f \in W(D), \left| \int_D f \, d\mu_{\sigma \circ u}^D - \int_{D_R} f \circ \pi_R^{-1} \, dv_R^D \right| < \varepsilon \right) = 1, \qquad (17)$$

where

$$W(D) = \left\{ f : D \to \mathbb{R}^d : f \text{ is continuous, } \|f\|_\infty \leq 1 \text{ and } Lip(f) \leq 1 \right\},$$

and

$$\sigma(u) = \left(e^{-u\kappa_d} + \xi^d (1 - e^{-u\kappa_d}) \right)^{1/d} .$$

3. *Length-volume incompatibility:*

$$\eta(u) \leq e^{-u\kappa_d} + \xi(1 - e^{-u\kappa_d}),$$

and in particular $\eta(u) < \sigma(u)$.

4. *The following is of interest when $\xi = 0$: for every $\varepsilon > 0$,*

$$\lim_{R \to \infty} \mathbf{P}_{u,R} \left(d_H(D, \pi_R^{-1}(D_R')) < \varepsilon \right) = 1, \qquad (18)$$

where $D_R' = \{x \in D_R : |\pi_R^{-1}(x)| = 1\}$. In other words, $\pi_R^{-1} : D_R' \to D$ is asymptotically surjective.

Theorem 1 implies the following corollary:

Corollary 1

1. *Part 1 of Theorem 1 implies that for every $\varepsilon > 0$,*

$$\lim_{R \to \infty} \mathbf{P}_{u,R} \left(d_{GH} \left((D_R, \operatorname{dist}_R^D), (D, d_{\eta \circ u}^D) \right) < \varepsilon \right) = 1. \qquad (19)$$

Moreover, for every sequence $(\omega_n, R_n, \varepsilon_n)$, where $\omega_n \in \Omega$, $R_n > 0$ and $\varepsilon_n \to 0$ such that the events in (16)–(18) hold with ε_n and R_n, the sequence of measured metric spaces $(D_{R_n}, \operatorname{dist}_{R_n}^D, \mu_{R_n}^D)(\omega_n)$ converges in the measured-Gromov-Hausdorff topology to the measured metric space $(D, d_{\eta \circ u}^D, \mu_{\sigma \circ u}^D)$.

2. *Part 3 in Theorem 1 implies that while the measures $\mu^D_{R_n}$ and the distance functions $\mathrm{dist}^D_{R_n}$ on D_{R_n} are consistent, the limit measure $\mu_{\sigma \circ u}$ is strictly larger than the measure $\mu_{\eta \circ u}$ induced by the metric $(\eta \circ u) \cdot \mathrm{\mathcal{e}}$ associated with the limit distance $\mathbf{d}^D_{\eta \circ u}$.*

3 Sketch of the proof

In the remaining sections of this paper, we prove Theorem 1 and Corollary 1. For conciseness, we assume that $\xi = 0$, which is the most difficult case; the proofs for $\xi > 0$ are similar, and in certain parts, much simpler. Since the proof is long and technical, this section describe its main stages.

In Sections 4–7, we consider a uniform distribution of point defects, that is, the case where u is a constant function. In Section 8, we generalize the results to arbitrary continuous function u.

Section 4: For a constant u, a uniform rescaling of $\mathbb{R}^d$ enables the introduction of a natural coupling between $(\mathbf{P}_{u,R})_{R>0}$ using a single measure $\mathbf{P}_u$ that is used in the rest of the proof for the uniform case.

We show the existence of a sub-critical regime, i.e., a constant $u_* = u_*(d) > 0$, such that for $u \in [0, u_*)$, the process $\mathbf{P}_u$-a.s. does not percolate. In this regime, $\mathcal{M}_R$ is "nice"; it is $\mathbf{P}_u$-a.s. a simply-connected metric space, which is locally isometric to the Euclidean space but for a nowhere-dense set. From this point onward we work only in the subcritical regime $u < u_*$.

We then use the subadditive ergodic theorem to obtain the existence of a limit distance function in $\mathcal{M}_R$. We show that for every $u \in [0, u_*)$ there exists an $\eta(u) \in [0, 1)$ such that

$$\mathbf{P}_u \left(\lim_{R \to \infty} \mathrm{dist}_R(\pi_R(x), \pi_R(y)) = \eta(u) \cdot |x - y|, \quad \forall x, y \in \mathbb{R}^d \right) = 1.$$

Moreover, this limit is uniform in every compact $K \subset \mathbb{R}^d$. This establishes the $\mathbf{P}_u$-a.s. Gromov-Hausdorff convergence of (K_R, dist_R) to $(K, \mathbf{d}_{\eta(u)})$.

Note, however, that this is not quite the metric convergence we want in (16) for a uniform u, since we want to prove that (K_R, dist^D_R) converges to $(K, \mathbf{d}^D_{\eta(u)})$ (i.e. the intrinsic distances converge and not only the induced distances). Moreover, we need to prove that $\eta(u) > 0$ for $u \in (0, u_*)$. These are done in the next sections.

Section 5: In this section we prove large deviation results for the distance function in $\mathcal{M}_R$. An immediate corollary is that $\eta(u) > 0$ for $u \in (0, u_*)$. The key idea here is to exploit the independence structure of the Poisson point process, manifested in the BK inequality, in order to show that distances do not deviate significantly from their expected value.

Section 6: In this section we prove several results regarding the geometry of geodesics in $(\mathcal{M}_R, \mathrm{dist}_R)$, and properties of the function η that controls the limit metric.

By using the concentration results of Section 5, we prove that geodesics in $\mathcal{M}_R$ are, with high probability, very close to straight lines between their endpoints.

Using a coupling between the probability measures $(\mathbf{P}_u)_{u \geq 0}$, we prove that $\eta(u)$ is a continuous, monotonically-nondecreasing function, and give an upper bound on the value of $\eta(u)$ that implies that $\eta < \sigma$ (which proves part 3 of Theorem 1).

Finally, we use the ergodicity of the model to prove that $\nu_R([0,1]^d)$ converges $\mathbf{P}_u$-a.s. to $\mu_{\sigma(u)}([0,1]^d)$ (which is a step towards proving the measure convergence in part 2 of Theorem 1).

Section 7: In this section we prove Parts 1 and 4 of Theorem 1 for uniform distributions on a convex, compact d-dimensional manifold with corners $D \subset \mathbb{R}^d$. The idea behind the proof of Part 1 is as follows: from Section 4 we know that $\mathbf{P}_u$-a.s. (D_R, dist_R) Gromov-Hausdorff converges to $(D, \mathbf{d}_{\eta(u)})$. Since D is convex and u is constant, $\mathbf{d}_{\eta(u)} = \mathbf{d}^D_{\eta(u)}$ on D, so we only need to replace (D_R, dist_R) with (D_R, dist_R^D). We do so by showing that the identity mapping between (D_R, dist_R) and (D_R, dist_R^D) has $\mathbf{P}_u$-a.s. vanishing distortion. Here we use the result from Section 6 that geodesics in $\mathcal{M}_R$ are very close to straight lines, which implies that dist_R-geodesics between points tend to remain within D, i.e. they are also dist_R^D-geodesics with high probability.

Section 8: In this section, we conclude the proof of Theorem 1 and Corollary 1 for a general continuous distribution u of defects over a d-dimensional manifold with corners $D \subset \mathbb{R}^d$.

The idea is the following: we partition D into small cubes, such that u is approximately constant in each cube. We show that the results of Sections 6 and 7 apply approximately to each of the cubes, in the sense that for every cube $\square$, when R is large, the distance between the measured metric spaces $(\square_R, \mathrm{dist}_R^\square, \nu_R^\square)$ and $(\square, \mathbf{d}^\square_{\eta \circ u}, \mu^\square_{\sigma \circ u})$ is bounded with high probability by the variation of u in $\square$.

Finally, we glue the cubes together and obtain results for the whole manifold D. For measure convergence, the gluing is straightforward. For metric convergence, we use the control on the distortion in each cube and a bound on the number of cubes each geodesic crosses to control the Gromov-Hausdorff distance between (D_R, dist_R^D) and $(D, \mathbf{d}^D_{\eta(u)})$ (in a similar way as in [11, 12]).

4 Uniform distribution of point defects

In this section, as well as in the three to follow, we study the simplest version of the model: we assume that the function $u : \mathbb{R}^d \to (0, \infty)$ is constant. In

this section, we prove the existence of a subcritical regime (Lemma 1), and obtain our first main result regarding distances in the manifolds $\mathcal{M}_R$ (Theorem 2). Many of the results obtained in these sections are adaptations to the continuous setting of the results obtained in [7,8] for the discrete case.

4.1 Rescaling for uniform distributions

The parameter $R > 0$ is a scaling factor that affects both the density of the point defects and their magnitude, or volume. In particular, to every R corresponds a different probability measure $\mathbf{P}_{u,R}$. For spatially invariant u, the intensity measures are translationally-invariant. By a uniform rescaling of space, we may obtain a probabilistic model that doesn't depend on R, and in particular allows us to construct a natural coupling of the measures $\mathbf{P}_{u,R}$.

This is done as follows: construct a Poisson point process on $\mathbb{R}^d$ with intensity $u \cdot \mathrm{Leb}_d(dx)$, i.e., set $R = 1$. We denote the probability measure by $\mathbf{P}_u = \mathbf{P}_{u,1}$, and the corresponding expectation by $\mathbf{E}_u = \mathbf{E}_{u,1}$. Two points in $\mathcal{M}_1$ are identified, $x \overset{\omega,1}{\sim} y$, if they are in the same ball of radius $R = 1$ centered at a point in the support of ω. We denote by $\mathrm{dist} = \mathrm{dist}_1$ the corresponding distance function on $\mathbb{R}^d$, which can also be written as

$$\mathrm{dist}(x, y; \omega) = \inf\left\{\sum_{i=1}^{N} |x_i - y_i| \ : \ N \in \mathbb{N}, \ x_1 \sim x, \ x_{i+1} \sim y_i, \ y_N \sim y\right\} \quad (20)$$

Given $R > 0$, let $T_R : \mathbb{R}^d \to \mathbb{R}^d$ be defined by $T_R(x) = Rx$. Then, $T_{1/R}^* \omega$ is a Poisson point process with intensity $R^d u \cdot \mathrm{Leb}_d(dx)$, namely $\mathbf{P}_{u,R} = (T_{1/R})_\# \mathbf{P}_u$. Similarly, we identify $\mathcal{M}_R$ with $\mathcal{M}_1$ via the scaling $T_{1/R}$. The distance function dist_R on $\mathbb{R}^d$ is derived from the distance function $\mathrm{dist} = \mathrm{dist}_1$ by the following relation,

$$\mathrm{dist}_R(x, y; \omega) = \frac{1}{R} \mathrm{dist}\left(R\,\pi_R^{-1}(x), R\,\pi_R^{-1}(y); \omega\right), \qquad \forall x, y \in \mathcal{M}_R. \quad (21)$$

4.2 The sub-critical regime

We start our analysis by proving the existence of a subcritical regime:

Lemma 1 *For $\omega \in \Omega$ let $\mathcal{S}(\omega) = \mathcal{S}_1(\omega) = \bigcup_{x \in \mathrm{supp}(\omega)} \overline{B(x,1)}$. Then*

(1) Each of the connected component of $\mathcal{S}(\omega)$ is $\mathbf{P}_u$-almost surely closed.
(2) There exists a constant $u_* > 0$, depending only on d, such that for all $u \in (0, u_*)$

$$\mathbf{P}_u \left(\begin{array}{c} \text{The set } \mathcal{S} \text{ doesn't contain an} \\ \text{infinite connected component} \end{array}\right) = 1 \quad (22)$$

and for every $u > u_*$

$$\mathbf{P}_u \left(\begin{array}{c} \textit{The set } \mathcal{S} \textit{ contains a unique} \\ \textit{infinite connected component} \end{array} \right) = 1. \tag{23}$$

In particular, whenever $u \in (0, u_*)$ the simply connected metric space $\mathcal{M}_R$ is $\mathbf{P}_u$-a.s. locally isometric to the Euclidean space, up to a nowhere dense set.

Remark 2 The subset $\mathcal{M}'_R$ of $\mathcal{M}_R$ that is locally isometric to the Euclidean space can be identified with $\mathbb{R}^d \setminus \mathcal{S}(\omega)$, in the sense they are diffeomorphic. However, they are not globally isometric. Note that $\mathcal{M}'_R$ is $\mathbf{P}_u$-a.s. not connected.

Proof The fact that each connected component of $\mathcal{S}(\omega)$ is $\mathbf{P}_u$-almost surely closed follows from the fact that with $\mathbf{P}_u$-probability one, every bounded set contains only finitely many points of supp(ω). This implies that the complement of $\mathcal{S}(\omega)$ is $\mathbf{P}_u$-almost surely open.

The existence of $u_* > 0$ such that $\mathcal{S}(\omega)$ doesn't contain $\mathbf{P}_u$-a.s. infinite clusters for every $u \in (0, u_*)$ and contains $\mathbf{P}_u$-a.s. an infinite cluster for $u > u_*$ is the content of [6]; see also Theorem 3.3 in [17]. In fact, for $u \in (0, u_*)$, the expected number of balls in the cluster containing the origin is finite, see [17] Theorem 3.2. The uniqueness of the infinite cluster for $u > u_*$ can be found in [17] Theorem 3.6.

Since $\mathcal{M}_R$ is obtained from $\mathbb{R}^d$ by a similarity transformation followed by an identification of points in simply-connected subsets, $\mathcal{M}_R$ is simply connected. As proved above, $\mathbb{R}^d \setminus \mathcal{S}(\omega)$, which is identical to $\mathcal{M}'_R$ up to a similarity transformation, is open. It follows that $\mathcal{M}'_R$ is locally isometric to Euclidean space.

Finally, we need to show that for $u \in (0, u_*)$ the set $\mathcal{M}_R \setminus \mathcal{M}'_R$ is nowhere dense. This follows from the fact that any compact subset of $\mathbb{R}^d$ contains $\mathbf{P}_u$-a.s. only finitely many points in supp(ω), hence every compact subset of $\mathcal{M}_R$ contains only finitely many points in $\mathcal{M}_R \setminus \mathcal{M}'_R$.

4.3 Distances in $\mathcal{M}_R$ in the uniform case

The main result of this section proves the existence of a limit distance function in $\mathcal{M}_R$. The precise statement is as follows:

Theorem 2 *For every $u \in [0, u_*)$ there exists an $\eta(u) \in [0, 1]$ such that*

$$\mathbf{P}_u \left(\lim_{R \to \infty} \text{dist}_R(\pi_R(x), \pi_R(y)) = \eta(u) \cdot |x - y|, \quad \forall x, y \in \mathbb{R}^d \right) = 1. \tag{24}$$

The limits also exist in the $L^1(\Omega, \mathbf{P}_u)$ sense: for every $x, y \in \mathbb{R}^d$,

$$\lim_{R \to \infty} \mathbf{E}_u \left[\left| \text{dist}_R(\pi_R(x), \pi_R(y)) - \eta(u) \cdot |x - y| \right| \right] = 0. \tag{25}$$

Furthermore, the convergence of the distance function is uniform in every compact
$K \subset \mathbb{R}^d$ *(and in particular in S^{d-1}),*

$$\mathbf{P}_u \left(\begin{array}{c} \forall \epsilon > 0 \; \exists R_0 \; \text{such that} \; \forall R > R_0, \\ \sup_{x,y \in K} \left| \text{dist}_R(\pi_R(x), \pi_R(y)) - \eta(u) \cdot |x-y| \right| < \epsilon \end{array} \right) = 1 \qquad (26)$$

and

$$\lim_{R \to \infty} \sup_{x,y \in K} \mathbf{E}_u \left[\left| \text{dist}_R(\pi_R(x), \pi_R(y)) - \eta(u) \cdot |x-y| \right| \right] = 0. \qquad (27)$$

In Sections 5-6 below we prove certain properties of $\eta(u)$, and in particular
that $\eta(u) > 0$ for every $u \in [0, u_*)$.

An immediate corollary of Theorem 2 is:

Corollary 2 *For every $u \in [0, u_*)$ and compact $K \subset \mathbb{R}^d$, the sequence of metric
spaces (K_R, dist_R) defined by (12) $\mathbf{P}_u$-a.s. Gromov-Hausdorff converges to $(K, \mathbf{d}_{\eta(u)})$,
where $\eta(u)$ should be considered as a constant function on $\mathbb{R}^d$.*

Note that this convergence is with respect to the induced distances dist_R
on K_R and not with respect to the intrinsic metric dist_R^K on K_R, defined in (11).
Proving the convergence of the intrinsic metric is more involved and requires
more assumptions on K, and is treated in Section 7.

4.4 Proof of Theorem 2

We start by reformulating Theorem 2 using the relation between dist and dist_R
and the definition of π_R:

Theorem 3 (Rephrasing of Theorem 2) *For every $u \in [0, u_*)$ there exists $\eta(u) \in
[0, 1]$ such that*

$$\mathbf{P}_u \left(\lim_{R \to \infty} \frac{\text{dist}(Rx, Ry)}{R} = \eta(u) \cdot |x-y|, \quad \forall x, y \in \mathbb{R}^d \right) = 1. \qquad (28)$$

The limits also exist in $L^1(\Omega, \mathbf{P}_u)$: for every $x, y \in \mathbb{R}^d$,

$$\lim_{R \to \infty} \mathbf{E}_u \left[\left| \frac{\text{dist}(Rx, Ry)}{R} - \eta(u) \cdot |x-y| \right| \right] = 0.$$

*Furthermore, the convergence of the distance function is uniform in every compact
$K \subset \mathbb{R}^d$ (and in particular in S^{d-1}).*

The proof of Theorem 3 is separated into several parts and starts with the
observation that the system $(\Omega, \mathcal{F}, \mathbf{P}_u)$ is ergodic with respect to translations.

Lemma 2 *For $x \in \mathbb{R}^d$ define $\tau_x : \Omega \to \Omega$ by $\tau_x \left(\sum_{i \geq 0} \delta_{x_i} \right) = \sum_{i \geq 0} \delta_{x_i - x}$. Then, for
every $x \in \mathbb{R}^d \setminus \{0\}$ the quartet $(\Omega, \mathcal{F}, \mathbf{P}_u, \tau_x)$, defines a translation invariant ergodic
system.*

Proof See for example [17] Proposition 2.6.

Next, we prove the existence of the limit in (28) for $x = 0$.

Lemma 3 *Let $u \in [0, u_*)$. For every $y \in \mathbb{R}^d$ the limit*

$$\rho_u(y) = \lim_{R \to \infty} \frac{\text{dist}(0, Ry)}{R} \tag{29}$$

exists $\mathbf{P}_u$-a.s. and in $L^1(\Omega, \mathbf{P}_u)$.

Proof If $u = 0$, then $\omega(\mathbb{R}^d) = 0$ $\mathbf{P}_u$-a.s., i.e., $\text{dist}(x, y) = |x - y|$ with $\mathbf{P}_u$-probability one, which implies that

$$\lim_{R \to \infty} \frac{\text{dist}(0, Ry)}{R} = |y|, \qquad \mathbf{P}_u\text{-a.s.}$$

Thus the result holds with $\rho_0(y) = |y|$.

We turn to the case $u > 0$. For $y = 0$ the statement is trivial, so fix $y \in \mathbb{R}^d \setminus \{0\}$ and define for $0 \leq m < n$

$$Y_{m,n} = \text{dist}(my, ny). \tag{30}$$

The triangle inequality for dist implies that

$$Y_{0,n} \leq Y_{0,m} + Y_{m,n}, \quad \forall 0 \leq m < n. \tag{31}$$

Thus we are in a good position to use Kingman's subadditive ergodic theorem [10]. More specifically we will exploit Liggett's version [16], which states that if $(X_{m,n})_{0 \leq m < n}$ are nonnegative random variables such that

(1) $X_{0,n} \leq X_{0,m} + X_{m,n}$ for all $0 < m < n$,
(2) $\{X_{nk,(n+1)k} \ : \ n \geq 1\}$ is stationary and ergodic for each $k \geq 1$,
(3) the law of $\{X_{m,m+k} \ : \ k \geq 1\}$ is independent of $m \geq 1$, and
(4) $E[X_{0,1}] < \infty$,

then the limit $\lim_{n \to \infty} X_{0,n}/n$ exists almost surely and in L^1 and equals

$$\inf_{n > 0} \frac{E[X_{0,n}]}{n} = \lim_{n \to \infty} \frac{E[X_{0,n}]}{n} < \infty$$

almost surely.

Taking $X_{m,n} = Y_{m,n}$, (1) is given by (31). For (2) and (3) note that

$$\text{dist}(u + z, v + z; \omega) = \text{dist}(u, v; \tau_z \omega), \quad \forall u, v, z \in \mathbb{R}^d, \tag{32}$$

and therefore $Y_{m,n} = Y_{0,n-m} \circ \tau_y^m$. Condition (2) and (3) then hold by the translation invariance and the ergodicity of the law $\mathbf{P}_u$ under the shift τ_y, see Lemma 2. As for (4), it follows immediately from the fact that $Y_{0,1} = \text{dist}(0, y) \leq |y|$, $\mathbf{P}_u$-a.s.

Thus, the family $\{Y_{m,n}\}_{0 \leq m < n}$ satisfies conditions (1)–(4) and the limit

$$\lim_{n \to \infty} n^{-1} \text{dist}(0, ny),$$

which we denote by $\rho_u(y)$, exists $\mathbf{P}_u$-a.s. and in L^1. We then prove that

$$\rho_u(y) = \lim_{R \to \infty} \frac{\text{dist}(0, Ry)}{R}, \quad \mathbf{P}_u\text{-a.s.} \tag{33}$$

Indeed, $\rho_u(y) = \lim_{R \to \infty} \frac{\text{dist}(0, \lfloor R \rfloor y)}{\lfloor R \rfloor}$ and since $\text{dist}(\lfloor R \rfloor y, Ry) \le |\lfloor R \rfloor y - Ry| \le |y|$, it follows that

$$\left| \frac{\text{dist}(0, Ry)}{R} - \frac{\text{dist}(0, \lfloor R \rfloor y)}{\lfloor R \rfloor} \right| \le \frac{2|y|}{R}.$$

This completes the proof.

Lemma 4 *The function ρ_u in Lemma 3 satisfies*

$$\rho_u(\alpha y) = \alpha \rho_u(y), \quad \forall \alpha \in (0, \infty), y \in \mathbb{R}^d, \tag{34}$$

and

$$\rho_u(y + z) \le \rho_u(y) + \rho_u(z), \quad \forall y, z \in \mathbb{R}^d. \tag{35}$$

Proof The positive homogeneity follows from the existence of the limit since

$$\rho_u(\alpha y) = \lim_{R \to \infty} \frac{\text{dist}(0, R\alpha y)}{R} = \lim_{R \to \infty} \alpha \frac{\text{dist}(0, R\alpha y)}{R\alpha} = \alpha \rho_u(y).$$

For the triangle inequality note that by the translation invariance of $\mathbf{P}_u$

$$\begin{aligned}
\mathbf{E}_u[\text{dist}(0, R(y + z))] &\le \mathbf{E}_u[\text{dist}(0, Ry)] + \mathbf{E}_u[\text{dist}(Ry, R(y + z))] \\
&= \mathbf{E}_u[\text{dist}(0, Ry)] + \mathbf{E}_u[\text{dist}(0, Rz)].
\end{aligned} \tag{36}$$

Diving both sides by R, taking the limit $R \to \infty$ and using the L^1 convergence of $\text{dist}(0, Ry)/R$ gives the required inequality.

Lemma 5 *The function ρ_u in Lemma 3 is invariant under the action of $SO(d)$.*

Proof Since $\mathbf{P}_u$ is invariant under the action of $SO(d)$ it follows that for every two points $y_1, y_2 \in \mathbb{R}^d$ with the property that there exists $\mathcal{R} \in SO(d)$ such that $y_2 = \mathcal{R}y_1$,

$$\text{dist}(0, y_2) \stackrel{d}{=} \text{dist}(0, y_1).$$

Thus, ρ_u inherits the symmetries of $SO(d)$ as the $L^1(\Omega, \mathbf{P}_u)$ limit in (3).

Corollary 3 *For $u \in [0, u_*)$, the function ρ_u in Lemma 3 is of the form*

$$\rho_u(y) = \eta(u) \cdot |y| \tag{37}$$

for some $\eta(u) \in [0, 1]$.

Proof This follows from the positive homogeneity, the sub-additivity and the isotropy of ρ_u proved in Lemmas 4 and 5.

Next, we consider general $x, y \in \mathbb{R}^d$:

Lemma 6 *Fix $u \in [0, u_*)$. Then, for every $x, y \in \mathbb{R}^d$ the limit*

$$\lim_{R \to \infty} \frac{\text{dist}(Rx, Ry)}{R} = \eta(u) \cdot |x - y|$$

exists $\mathbf{P}_u$-a.s. and in L^1.

Proof For every $x, y \in \mathbb{R}^d$ and $\omega \in \Omega$ the relation $\text{dist}(x, y; \omega) = \text{dist}(0, y - x, \tau_x \omega)$ holds and therefore

$$\frac{\text{dist}(Rx, Ry; \omega)}{R} = \frac{\text{dist}(0, R(y - x); \tau_x \omega)}{R},$$

which implies that the limit $R \to \infty$ exists $\mathbf{P}_u$-a.s., is in $L^1(\Omega, \mathbf{P}_u)$ and equals $\eta(u) \cdot |x - y|$.

Proof of Theorem 3: By Lemmas 3–6 the $R \to \infty$ limit of $\text{dist}(Rx, Ry)/R$ exists both $\mathbf{P}_u$-a.s. and in L^1 for every fixed pair of points $x, y \in \mathbb{R}^d$. It remains to verify that the limit exists $\mathbf{P}_u$-a.s. simultaneously for all pairs of points $x, y \in \mathbb{R}^d$.

To this end, let $\varepsilon \in (0, 2\pi)$ and let $(v_i)_{i=1}^N$ with $N = \lceil c(d)/\varepsilon^d \rceil$ be a set of points on the unit sphere S^{d-1} that form an $\varepsilon/2$-net for S^{d-1} ($c(d)$ is a constant that depends only on d). Since the set $(v_i)_{i=1}^N$ is finite, it follows from Lemma 6 that

$$\mathbf{P}_u \left(\lim_{R \to \infty} \frac{\text{dist}(Rv_i, Rv_j)}{R} = \eta(u) \cdot |v_i - v_j|, \quad \forall 1 \le i, j \le N \right) = 1. \qquad (38)$$

Given $x, y \in \mathbb{R}^d \backslash \{0\}$, there exist $1 \le i, j \le N$ such that

$$|\hat{x} - v_i| \le \varepsilon \text{ and } \left| \hat{y} - v_j \right| \le \varepsilon, \qquad (39)$$

where $\hat{x} = x/|x|$ and $\hat{y} = y/|y|$. By the triangle inequality,

$$\left| \frac{\text{dist}(Rx, Ry)}{R} - \frac{\text{dist}(R|x|v_i, R|y|v_j)}{R} \right| \le \frac{\text{dist}(Rx, R|x|v_i)}{R} + \frac{\text{dist}(Ry, R|y|v_i)}{R}$$

$$\le \frac{|Rx - R|x|v_i)|}{R} + \frac{|Ry - R|y|v_i|}{R} \qquad (40)$$

$$\le \varepsilon(|x| + |y|),$$

and

$$\left| \eta(u) \cdot |x - y| - \eta(u) \cdot \||x|v_i - |y|v_j\| \right| \le \varepsilon(|x| + |y|). \qquad (41)$$

Combining both estimates, it follows from (38) that

$$\mathbf{P}_u \left(\limsup_{R \to \infty} \left| \frac{\text{dist}(Rx, Ry)}{R} - \eta(u) \cdot |x - y| \right| \le 2\varepsilon(|x| + |y|), \quad \forall x, y \in \mathbb{R}^d \backslash \{0\} \right) = 1.$$

The cases $x = 0$, $y = 0$ can be included by a similar argument, except that no approximation for 0 is needed. Since this holds for every $\varepsilon > 0$, the limit as $R \to \infty$ exists $\mathbf{P}_u$-a.s.

To justify the uniformity over compact sets, note that there are only finitely many v_i's for a fixed $\varepsilon > 0$ and that for compact sets we have a uniform bound on the Euclidean norm of both x and y. The same applies for the $L^1(\Omega, \mathbf{P}_u)$ convergence. ∎
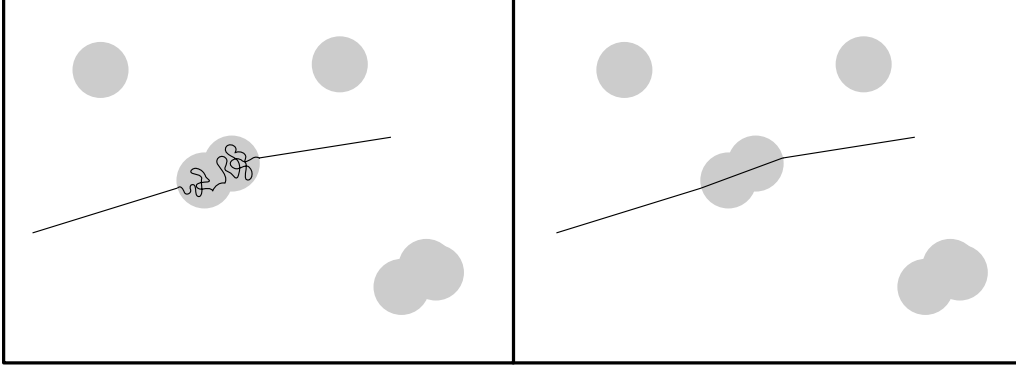
Fig. 1 The path on the left is a geodesic in $\mathcal{M}_R$, where the gray regions correspond to $\mathcal{S}_R(\omega)$. Note that the part in $\mathcal{S}_R$ has zero length. The path on the right is the corresponding true geodesic.

5 Large deviation results

In this section we prove large deviation results for the distance in $\mathcal{M}_R$. As an immediate corollary we obtain that $\eta(u) > 0$ for every $u \in [0, u_*)$.

The main result of this section is the following.

Theorem 4 *For every $u \in [0, u_*)$ and every $\varepsilon > 0$ there exists a positive constant c_1, depending only on d, u and ε, such that for large enough R,*

$$\mathbf{P}_u\left(\exists x \in \mathbb{R}^d \ : \ |x| = R, \ |\operatorname{dist}(0, x) - \eta(u)R| > \varepsilon R\right) < e^{-c_1 R}. \tag{42}$$

The proof of Theorem 4 follows the ideas developed for the discrete case by Kesten[8].

In addition, we will need a large deviation result for the existence of very long geodesics in the Euclidean sense. In order to state it we need another definition: Geodesics in $\mathcal{M}_R$ can be identified with geodesics in $\mathbb{R}^d$ with respect to the semi-distance function dist. Such geodesics are highly degenerate, as there is nothing that limits their behavior in $\mathcal{S}(\omega)$. For $x, y \in \mathbb{R}^d$ we denote by $\Gamma_0(x, y)$ the set of geodesics between x and y with respect to dist that minimize the Euclidean distance inside $\mathcal{S}(\omega)$. We will call such paths *true geodesics* (see Figure 1).

Theorem 5 *For every $u \in [0, u_*)$ there exists positive constants α, C_2, c_3, depending only on d and u, such that*

(1) For every $R > 0$,

$$\mathbf{P}_u\left(\begin{array}{c}\exists \text{ a true geodesic path starting at } 0 \\ \text{such that } \operatorname{len}_e(\gamma) \geq R \text{ and } \operatorname{len}(\gamma) < \frac{1}{\alpha}\operatorname{len}_e(\gamma)\end{array}\right) \leq C_2 e^{-c_3 R}. \tag{43}$$

(2) For every $x, y \in \mathbb{R}^d$ and for every $R > \alpha|x - y|$

$$\mathbf{P}_u\left(\exists \gamma \in \Gamma_0(x, y) \ : \ \operatorname{len}_e(\gamma) > R\right) \leq C_2 e^{-c_3 R}. \tag{44}$$

As a corollary we obtain:

Corollary 4 $\eta(u)$ *is strictly positive for every* $u \in [0, u_*)$.

Proof Let A_R be the event in (43). By Part (1) of Theorem 5,

$$\mathbf{P}_u(A_R) \leq C_2 e^{-c_3 R}, \quad \forall R > 0.$$

Since the Euclidean length of a true geodesic connecting 0 and Re_1 is at least R, for $\omega \in A_R^c$,

$$\mathrm{dist}(0, Re_1) \geq \frac{1}{\alpha} R,$$

Thus, for large enough R

$$\mathbf{E}_u[\mathrm{dist}(0, Re_1)] \geq \mathbf{E}_u[\mathrm{dist}(0, Re_1) \cdot \mathbb{1}_{A_R^c}] \geq \frac{1}{\alpha} R \cdot \mathbf{P}_u(A_R^c) \geq \frac{1}{2\alpha} R,$$

hence

$$\eta(u) = \lim_{R \to \infty} \frac{\mathbf{E}_u[\mathrm{dist}(0, Re_1)]}{R} \geq \frac{1}{2\alpha} > 0.$$

The proofs of Theorem 4 and Theorem 5 are quite technical. The main idea behind the proof is to exploit the independence structure of the Poisson point process, manifested in the BK inequality, in order to show that geodesics whose lengths deviate from the expected distance in the sense of (42), contain sufficiently many disjoint sub-paths (which are roughly independent) whose total length deviates significantly from its expected value. Such an event is highly unlikely due to large deviations results for independent random variables.

Since the proofs in this section are technical and since one can use the theorems as "black boxes" in the rest of the paper, the reader might wish to skip the rest of this section in a first reading.

5.1 The BK inequality

In this subsection we state the continuous version of the well-known BK inequality for product measures. For this, we need some additional definitions.

There is a natural partial ordering on Ω, which we denote by $\preceq$, under which $\omega \preceq \omega'$ if and only if $\mathrm{supp}(\omega) \subseteq \mathrm{supp}(\omega')$. Using it one can define increasing and decreasing events in $\mathcal{F}$. An event $A \in \mathcal{F}$ is said to be increasing (respectively decreasing) if for every $\omega \preceq \omega'$, $\omega \in A$ implies $\omega' \in A$ (i.e., A is closed under increasing support).

For any bounded Borel set $Y \subset \mathbb{R}^d$, define the set

$$\omega_Y = \mathrm{supp}(\omega) \cap Y,$$

and for $\omega \in \Omega$ and Y as above let

$$L(\omega, Y) = \{\omega' \in \Omega : \mathrm{supp}(\omega) \cap Y \subseteq \mathrm{supp}(\omega') \cap Y\}.$$

In words, the event $L(\omega, Y)$ contains all configurations that inside U are larger than ω. We say that an event A is an increasing event on Y if $\omega \in A$ implies that $L(\omega, Y) \subseteq A$.

Definition 2 Let A and B be two increasing events on a bounded Borel set Y. Then

$$A \circ B = \left\{ \omega \in \Omega \; : \; \begin{array}{c} \text{there are disjoint sets } V, W \subset \mathbb{R}^d \text{ such that} \\ V \text{ and } W \text{ are finite unions of rational} \\ \text{cubes and } L(\omega, V) \subset A, L(\omega, W) \subset B \end{array} \right\},$$

where by a rational cube we mean an open d-dimensional cube with rational coordinates. When $A \circ B$ occurs, we say that A and B occur disjointly.

To exemplify this definition, let $d = 2$, let $Y = [0, 1]^2$, and let A be the event that there exists a path in Y connecting the left boundary of Y to its right boundary, whose length is less than x. Clearly, A is an increasing event, as increasing the support of ω can only shorten paths. For $\omega \in A$ and $V \subset Y$, it is generally not true that $L(\omega, V) \subset A$; $L(\omega, V) \subset A$ only if there exists a path in V connecting the left boundary of U to its right boundary, whose length is less than x. The event $A \circ A$ occurs if there exist two disjoint paths connecting the left boundary of Y to its right boundary, whose length is at most x.

Theorem 6 (BK inequality) *Suppose Y is a bounded Borel set in $\mathbb{R}^d$ and A, B are two increasing events on Y. Then for every $u > 0$*

$$\mathbf{P}_u(A \circ B) \leq \mathbf{P}_u(A)\mathbf{P}_u(B).$$

A proof of this inequality in a more general setting can be found in [17], Theorem 2.3.

5.2 Key proposition

The proofs of Theorem 4 and Theorem 5 exploit the invariance of the law of $\mathbf{P}_u$ under translations and rotations. This implies that we only need to take care of large deviation results for the distance between the origin and points of the form Re_1, where $e_1 = (1, 0, \ldots, 0)$. We therefore restrict ourselves to the above case and define the following:

For $r \in \mathbb{R}$ let

$$H_r = \{x \in \mathbb{R}^d \; : \; \langle x, e_1 \rangle = r\}.$$

Given $r < s$ and a path γ we write

$$H_r < \gamma < H_s$$

if all points of γ, except possibly its endpoints, lie strictly between the hyperplanes H_r and H_s.

The following variants on the distance will stand in the core of the proofs. For $N, M > 0$ define the random variables,

$$s_{M,N} = \inf \left\{ \operatorname{len}(\gamma) \; : \; \begin{array}{l} \gamma \text{ is a path from } \{0\} \times [0, N]^{d-1} \\ \text{to } H_M \text{ such that } H_0 < \gamma < H_M \end{array} \right\} \tag{45}$$

and

$$\widehat{s}_{M,N} = \inf \left\{ \operatorname{len}(\gamma) \; : \; \begin{array}{l} \gamma \text{ is a path from } \{0\} \times [0, N]^{d-1} \text{ to } H_M \text{ such} \\ \text{that with the exception of its endpoints,} \\ \gamma \subset (0, M) \times [-4M, 4M]^{d-1} \end{array} \right\}. \tag{46}$$

It follows from the definitions of $s_{M,N}$ and $\widehat{s}_{M,N}$ that

$$s_{M,N} \leq \widehat{s}_{M,N}. \tag{47}$$

The main estimate used in the proofs of both theorems is stated in the next proposition:

Proposition 1 *Let* $\left(X_q(M, N)\right)_{q \geq 0}$ *and* $\left(\widehat{X}_q(M, N)\right)_{q \geq 0}$ *be sequences of independent random variables having the same distribution as* $s_{M,N}$ *and* $\widehat{s}_{M,N}$*. For every two integers* $2 \leq N \leq M \leq R/2$ *and every real number* $x \geq 0$,

$$\mathbf{P}_u \left(\operatorname{dist}(0, Re_1) < x \right) \leq \sum_{Q \geq \frac{R}{M+N} - 1} \left(2d \left(16 \frac{M}{N} \right)^d \right)^Q \mathbf{P}_u \left(\sum_{q=0}^{Q-1} \widehat{X}_q(M, N) < x \right)$$

$$\leq \sum_{Q \geq \frac{R}{M+N} - 1} \left(2d \left(16 \frac{M}{N} \right)^d \right)^Q \mathbf{P}_u \left(\sum_{q=0}^{Q-1} X_q(M, N) < x \right). \tag{48}$$

Proof Since $s_{M,N} \leq \widehat{s}_{M,N}$, the second inequality is immediate. Assume $\gamma : [0, 1] \to \mathbb{R}^d$ is a simple path from the origin to H_R. We choose a sequence of points $(x_0, x_1, \dots, x_Q)$ along γ as follows: Define $x_0 = \gamma(0) = 0$. Assume that $(x_0, x_1 \dots, x_q)$ have already been chosen such that $x_i = \gamma(t_i)$ with $0 = t_0 < t_1 < \dots < t_q$. Then we define

$$t_{q+1} = \min\{t \in (t_q, 1] \; : \; |\gamma(t) - x_q|_\infty = M + N\}, \quad x_{q+1} = \gamma(t_{q+1}), \tag{49}$$

provided such time t exists. If no such t exists, i.e., $|\gamma(t) - x_q|_\infty < M + N$ for all $t \in (t_q, 1]$, then we set $Q = q$ and stop the process.

It follows from the definition of the points x_q that

$$Q \geq \frac{R}{M + N} - 1. \tag{50}$$

Indeed, the distance between $\langle x_q, e_1 \rangle$ and $\langle x_{q+1}, e_1 \rangle$ is at most $M + N$ and $\langle x_Q, e_1 \rangle \geq R - (M + N)$.

Next, we analyze the path within the time interval $[t_q, t_{q+1}]$ for $0 \le q \le Q - 1$. Since, by definition, $|x_q - x_{q+1}|_\infty = M + N$, it follows that there exist $j = j(q) \in \{1, 2, \ldots, d\}$ and $\sigma = \sigma(q) \in \{-1, 1\}$ such that

$$\langle x_{q+1} - x_q, e_{j(q)} \rangle = \sigma(q)(M + N).$$

We introduce the hyperplanes,

$$H_r^j = \{x \in \mathbb{R}^d \mid \langle x, e_j \rangle = r\}, \qquad j = 1, \ldots, d, \qquad r \in \mathbb{R},$$

so that, in particular, $H_r^1 = H_r$. The path $\gamma|_{[t_q, t_{q+1}]}$ is strictly restricted between the hyperplanes

$$H(q) = H_{\langle x_q, e_{j(q)} \rangle}^{j(q)} \qquad \text{and} \qquad H'(q) = H_{\langle x_q, e_{j(q)} \rangle + \sigma(q)(M+N)}^{j(q)} .$$

Since the distance between these hyperplanes is $M + N$, we can find two hyperplanes $H''(q)$ and $H'''(q)$ at a distance M one of the other, that are strictly between $H(q)$ and $H'(q)$ (see Figure 2). Further, we denote by $[t''(q), t'''(q)]$ the sub-interval of $[t_q, t_{q+1}]$ in which the path γ contains a unique crossing between the two hyperplanes $H''(q)$ and $H'''(q)$.

Specifically, let $m_{j(q)}(q)$ be the integer m such that $H_{mN}^{j(q)}$ is the hyperplane bounded between $H(q)$ and $H'(q)$ that is the closest to $H(q)$. Formally,

$$m_{j(q)}(q) = \left\lfloor \frac{\langle x_q, e_{j(q)} \rangle}{N} \right\rfloor + \frac{1}{2}(\sigma(q) + 1).$$

We then define

$$H''(q) = H_{mN}^{j(q)} \qquad \text{and} \qquad H'''(q) = H_{mN + \sigma(q)M}^{j(q)},$$

and let $[t''(q), t'''(q)] \subset [t_q, t_{q+1}]$ be the subinterval defined by

$$t''(q) = \max \left\{ t \in [t_q, t_{q+1}] \; : \; \gamma(t) \in H''(q) \right\},$$

$$t'''(q) = \min \left\{ t \in (t''(q), t_{q+1}) \; : \; \gamma(t) \in H'''(q) \right\}.$$

Finally, denote by γ_q the path γ restricted to the time interval $[t''(q), t'''(q)]$.

We are interested in a bound on paths starting within a $(d-1)$-dimensional box of side length N within the hyperplane H_0, see (46). We denote by $m(q) = (m_1(q), m_2(q), \ldots, m_q(d))$ the unique point in $\mathbb{Z}^d$ satisfying

$$\gamma''(t) \in \Lambda(q) = H''(q) \cap \left(m(q)N + [0, N)^d \right).$$

where $\Lambda(q)$ is a $(d-1)$-dimensional box of side length N within the hyperplane $H''(q)$.

Exploiting all the above definitions, the path segments γ_q satisfy the following properties:

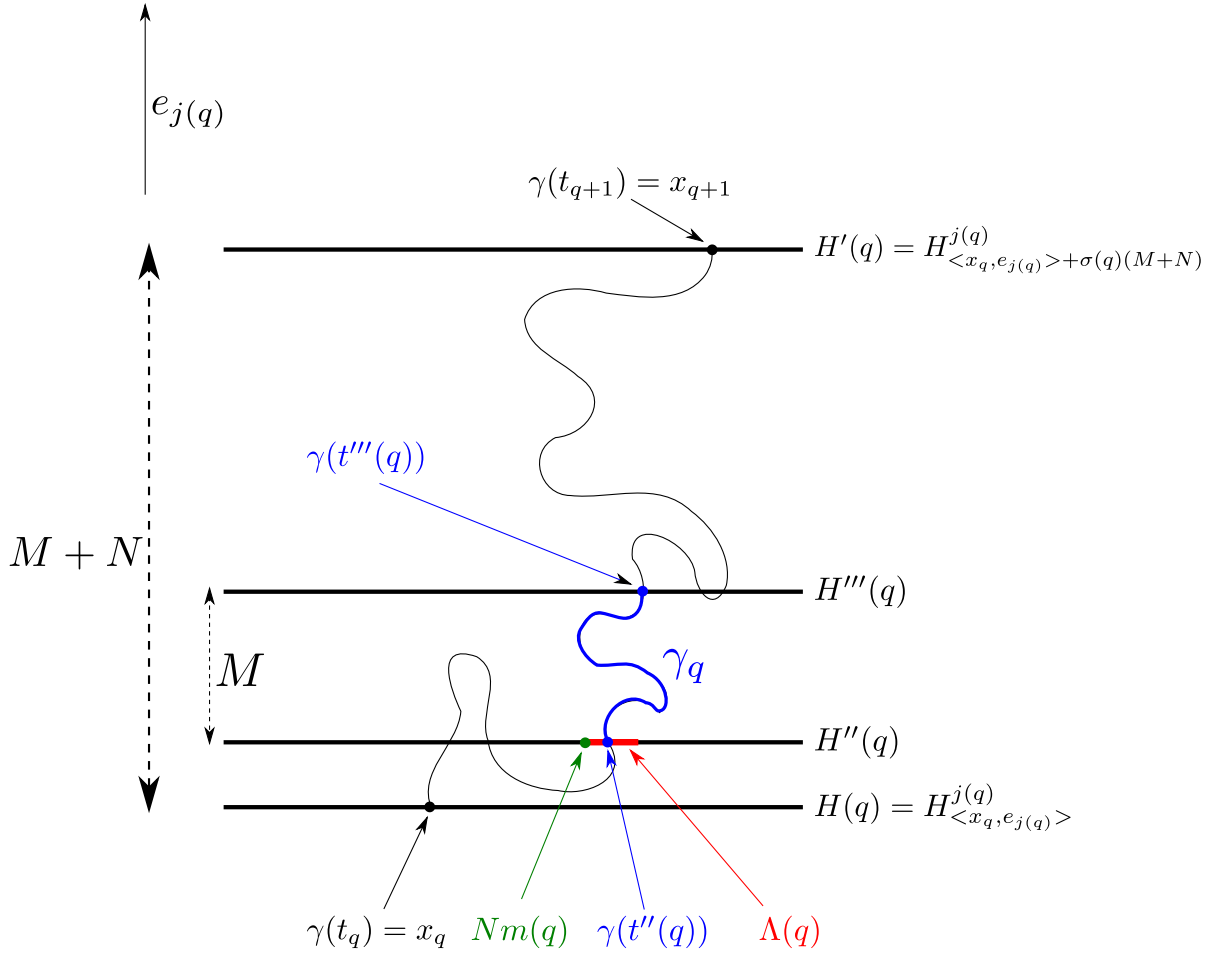– γ_q are pairwise disjoint.

Fig. 2 Illustration of the times $t_q, t_{q+1}, t''(q), t'''(q)$, the points $x_q, x_{q+1}, \gamma(t''(q)), \gamma(t'''(q))$, the hyperplanes $H(q), H'(q), H''(q), H'''(q)$, the box $\Lambda(q)$ with its corner $Nm(q)$ and the path γ_q crossing from $\Lambda(q)$ to the $H'''(q)$.

- γ_q connects $\Lambda(q) \subset H''(q)$ to $H'''(q)$.
- The path γ_q lies strictly between $H''(q)$ and $H'''(q)$, except for its endpoints.
- The path γ_q is contained in the box

$$B(q) = Nm(q) + [-4M, 4M]^{j(q)-1} \times [0, \sigma(j)M] \times [-4M, 4M]^{d-j(q)},$$

where we used the fact that $N + M \leq 2M$.
- The total length of the paths γ_q satisfies

$$\sum_{q=0}^{Q-1} \text{len}(\gamma_q) \leq \text{len}(\gamma).$$

Fix $x > 0$, fix $Q \geq \frac{R}{M+N} - 1$ and fix $(j(q), \sigma(q), m(q))$, $q = 0, \dots, Q - 1$. We denote by $A(j, \sigma, m, x)$ the event that there exists a piecewise-linear simple path γ containing disjoint segments $\gamma_q \subset B(q)$ crossing the box $B(q)$ in the $j(q)$ direction from $\Lambda(q) \subset H''(q)$ to $H'''(q)$, with $\sum_{q=0}^{Q-1} \mathrm{len}(\gamma_q) \leq x$.

Since every true geodesic is piecewise-linear, we conclude that the event

$$\{\mathrm{dist}(0, Re_1) < x\}$$

occurs only if there exists a $Q \geq \frac{R}{M+N} - 1$ and there exist $(j(q), \sigma(q), m(q))$, $q = 0, \dots, Q - 1$ such that $A(j(q), \sigma(q), m(q), x)$ occurs, namely,

$$\mathbf{P}_u(\mathrm{dist}(0, Re_1) < x) \leq \sum_{Q \geq \frac{R}{M+N} - 1} \sum_{(j(q), \sigma(q), m(q))} \mathbf{P}_u(A(j, \sigma, m, x)). \qquad (51)$$

The event $A(j, \sigma, m, x)$ is characterized by the existence of a path whose properties are specified over disjoint segments. Denote by $A_q(j, \sigma, m, x)$ the event that there exists a piecewise-linear simple path γ containing a segment $\gamma_q \subset B(q)$ crossing the box $B(q)$ in the $j(q)$ direction from $\Lambda(q) \subset H''(q)$ to $H'''(q)$ and satisfies $\mathrm{len}(\gamma_q) < x$.

Then,

$$\mathbf{P}_u(A(j, \sigma, m, x)) = \bigcup_{\substack{r_0, r_1, \dots, r_{Q-1} \in \mathbb{Q}_+ \\ \sum_{q=0}^{Q-1} r_q < x}} \mathbf{P}_u(A_0(j, \sigma, m, r_0) \circ A_1(j, \sigma, m, r_1) \circ \cdots \circ A_{Q-1}(j, \sigma, m, r_{Q-1})).$$

$$(52)$$

Since the paths γ_q are disjoint and piecewise linear, the conditions of the BK inequality are satisfied and we conclude that

$$\mathbf{P}_u(A(j, \sigma, m, x)) \leq \sum_{\substack{r_0, r_1, \dots, r_{Q-1} \in \mathbb{Q}_+ \\ \sum_{q=0}^{Q-1} r_q < x}} \prod_{q=0}^{Q-1} \mathbf{P}_u(A_q(j, \sigma, m, r_q)).$$

Noting that for a given configuration (j, σ, m) the minimal length of a path connecting $\Lambda(q)$ to $H''(q)$ inside $B(q)$ has the same distribution as $\widehat{s}_{M,N}$, we conclude that

$$\mathbf{P}_u(A(j, \sigma, m, x)) \leq \sum_{\substack{r_0, r_1, \dots, r_{Q-1} \in \mathbb{Q}_+ \\ \sum_{q=0}^{Q-1} r_q < x}} \prod_{q=0}^{Q-1} \mathbf{P}_u(\widehat{X}_q(M, N) \leq r_q) = \mathbf{P}_u\left(\sum_{q=0}^{Q-1} \widehat{X}_q(M, N) < x\right),$$

which combined with (51), yields

$$\mathbf{P}_u(\mathrm{dist}(0, Re_1) < x) \leq \sum_{Q = \frac{R}{M+N} - 1}^{\infty} \left(\begin{array}{c} \text{number of choices} \\ \text{for } (j, \sigma, m) \end{array}\right) \cdot \mathbf{P}_u\left(\sum_{q=0}^{Q-1} \widehat{X}_q(M, N) < x\right).$$

$$(53)$$

To complete the proof, we need to show that the number of ways to choose the triplets (j, σ, m) is bounded by $(2d(16M/N)^d)^Q$. To this end, assume that (j, σ, m) has already been chosen. By the definition of the times $t''(q)$, $t''(q+1)$, t_q and t_{q+1} and the points $m(q)$ and $m(q+1)$,

$$
\begin{aligned}
N|m(q+1) - m(q)|_\infty &\leq |Nm(q+1) - \gamma(t''(q+1))|_\infty + |\gamma(t''(q+1)) - \gamma(t_{q+1})|_\infty \\
&\quad + |\gamma(t_{q+1}) - \gamma(t_q)|_\infty + |\gamma(t_q) - \gamma(t''(q))|_\infty \\
&\quad + |Nm(q) - \gamma(t''(q))|_\infty \\
&\leq 3M + 5N.
\end{aligned}
$$

Since $m(q+1) \in \mathbb{Z}^d$, it follows that there are at most $\left(6\frac{M}{N} + 10\right)^d$ ways to choose $m(q+1)$ given $m(q)$. Moreover, there are at most d choices for $j(q+1)$ and 2 choice for $\sigma(q+1)$, hence given $(j(q), \sigma(q), m(q))$ there are at most

$$
2d\left(6\frac{M}{N} + 10\right)^d \leq 2d\left(16\frac{M}{N}\right)^d
$$

choices for $(j(q+1), \sigma(q+1), m(q+1))$ and in total at most

$$
\left(2d\left(16\frac{M}{N}\right)^d\right)^Q
$$

choices for the whole sequence.

5.3 Proof of Theorem 4

We separate the proof of Theorem 4 into two parts: a lower bound estimation and an upper bound estimation. We start with the first.

Formally, the claim for the lower bound is that for every $u \in [0, u^*)$ and every $\varepsilon > 0$ there exists a constant c, depending only on d, u and ε, such that for R large enough

$$
\mathbf{P}_u\left(\exists x \in \mathbb{R}^d \ : \ |x| = R, \ \mathrm{dist}(0, x) < (\eta(u) - \varepsilon)R\right) < e^{-cR}. \tag{54}
$$

We start by showing that $\eta(u)$ can also be obtained as the limiting distance between a point and a hyperplane:

Lemma 7 *For every $u \in [0, u_*)$,*

$$
\lim_{R \to \infty} \frac{\mathrm{dist}(0, H_R)}{R} = \eta(u), \quad \mathbf{P}_u\text{-a.s.} \tag{55}
$$

Proof From the definition of $\mathrm{dist}(0, H_R)$,

$$
\mathrm{dist}(0, H_R) \leq \mathrm{dist}(0, Re_1),
$$

and therefore

$$\limsup_{R\to\infty} \frac{\mathrm{dist}(0,H_R)}{R} \le \lim_{R\to\infty} \frac{\mathrm{dist}(0,Re_1)}{R} = \eta(u), \quad \mathbf{P}_u\text{-a.s.}$$

If $\eta(u) = 0$, then there is nothing left to prove since $\mathrm{dist}(0,H_R) \ge 0$, hence assume that $\eta(u) > 0$. Let $\omega \in \Omega$ be a realization such that

$$\liminf_{R\to\infty} \frac{\mathrm{dist}(0,H_R)}{R} = \eta(u) - 2\delta, \quad \text{for some } \delta > 0. \tag{56}$$

Then, there exists an increasing sequence R_k that tends to infinity, a sequence of points $z_k \in H_{R_k}$, and a sequence of paths γ_k from 0 to z_k such that

$$\mathrm{dist}(0,z_k) = \mathrm{len}(\gamma_k) \le \mathrm{dist}(0,H_{R_k}) + \delta R_K \le R_k(\eta(u) - \delta).$$

This however implies that

$$\limsup_{k\to\infty} \frac{\mathrm{dist}(0,z_k)}{|z_k|} \le \limsup_{k\to\infty} \frac{R_k(\eta(u) - \delta)}{|z_k|} \le \eta(u) - \delta,$$

contradicting the uniform convergence of the distance function on S^{d-1} from Theorem 3. Thus the event in (56) has probability zero and the claim follows.

Next, we show that the value of $s_{M,N}$ defined in (45) cannot be much smaller than $M\eta(u)$.

Lemma 8 *For every $\varepsilon > 0$*

$$\lim_{M\to\infty} \max_{N \le M} \mathbf{P}_u\left(s_{M,N} \le M(\eta(u) - \varepsilon)\right) = 0. \tag{57}$$

Proof If $\eta(u) = 0$, then there is nothing to prove since $s_{M,N} \ge 0$. We therefore assume that $\eta(u) > 0$. Let

$$r_{M,N} = \inf\left\{\mathrm{len}(\gamma) \ : \ \gamma \text{ is a path from } \{0\} \times [0,N]^{d-1} \text{ to } H_M\right\}. \tag{58}$$

Since $r_{M,N}$ is an infimum of path lengths over a set larger than that defining $s_{M,N}$, it follows that $r_{M,N} \le s_{M,N}$, therefore

$$\mathbf{P}_u(s_{M,N} \le M(\eta(u) - \varepsilon)) \le \mathbf{P}_u(r_{M,N} \le M(\eta(u) - \varepsilon)) \le \mathbf{P}_u(r_{M,M} \le M(\eta(u) - \varepsilon)),$$

where in the last passage we used the fact that $r_{M,N}$ is decreasing in N. Hence it is enough to prove that

$$\lim_{M\to\infty} \mathbf{P}_u(r_{M,M} \le M(\eta(u) - \varepsilon)) = 0.$$

We now show that, in fact, it is sufficient to prove that for some fixed choice of $\delta = \delta(\varepsilon) \in (0, 1/2)$,

$$\lim_{M\to\infty} \mathbf{P}_u(r_{M,2\delta M} \le M(\eta(u) - \varepsilon)) = 0. \tag{59}$$

Indeed, let $\delta \in (0, 1/2)$ and for $v \in H_0$ define $r_M(v) = \text{dist}(v, H_M)$. Since $\{0\} \times [0, M]^{d-1}$ is contained in the union of boxes $\{\{0\} \times (2\delta M k + [0, 2\delta M]^{d-1})\}_{k \in A(\delta)}$ with

$$A(\delta) = \left\{ k = (k_2, \ldots, k_d) \in \mathbb{Z}^{d-1} : 0 \le k_i \le \lceil 1/(2\delta) \rceil, \quad \forall 2 \le i \le d \right\}$$

it follows that

$$\begin{aligned} r_{M,M} &= \inf\{r_M(v) : v \in \{0\} \times [0, M]^{d-1}\} \\ &\ge \min_{k \in A(\delta)} \inf\{r_M(v) : v \in \{0\} \times (2\delta M k + [0, 2\delta M]^{d-1})\}. \end{aligned}$$

Since the set $A(\delta)$ is finite, and since each infimum for a fixed k has the same distribution as $r_{M,\delta M}$ it is indeed sufficient to prove (59).

Finally, let us prove (59). For every $v \in H_0$ such that $|v|_\infty \le 2\delta M$ we have $\text{dist}(0, v) \le 2\delta M$, and therefore

$$r_M(0) \le r_M(v) + 2\delta M.$$

Taking the infimum over all such v's yields

$$r_M(0) \le r_{M,2\delta M} + 2\delta M,$$

and therefore for $\delta < \varepsilon/4$,

$$\begin{aligned} \mathbf{P}_u(r_{M,2\delta M} \le M(\eta(u) - \varepsilon)) &\le \mathbf{P}_u(r_M(0) - 2\delta M \le M(\eta(u) - \varepsilon)) \\ &= \mathbf{P}_u(\text{dist}(0, H_M) - 2\delta M \le M(\eta(u) - \varepsilon)) \\ &\le \mathbf{P}_u(\text{dist}(0, H_M) \le M(\eta(u) - \varepsilon/2)) \end{aligned}$$

which, by Lemma 7, tends to zero as M tends to infinity.

Proof of Theorem 4 (lower bound): For $\eta(u) = 0$ the claim is trivial since $\mathbf{P}_u(\text{dist}(0, x) < 0) = 0$. For $\eta(u) > 0$, let $N = \min\left\{M, \left\lfloor \frac{M\varepsilon}{4\eta(u)} \right\rfloor\right\}$. For $Q \ge \frac{R}{M+N} - 1$ and $\beta > 0$,

$$\begin{aligned} \mathbf{P}_u\left(\sum_{q=0}^{Q-1} X_q(M, N) < R(\eta(u) - \varepsilon)\right) &\le e^{\beta R(\eta(u) - \varepsilon)} \cdot \mathbf{E}_u\left[\exp\left(-\beta \sum_{q=0}^{Q-1} X_q(M, N)\right)\right] \\ &= e^{\beta R(\eta(u) - \varepsilon)} \cdot \mathbf{E}_u\left[e^{-\beta X_1(M,N)}\right]^Q \\ &\le e^{\beta(Q+1)(M+N)(\eta(u) - \varepsilon)} \cdot \mathbf{E}_u\left[e^{-\beta X_1(M,N)}\right]^Q \\ &= e^{\beta(M+N)(\eta(u) - \varepsilon)} \cdot \left(e^{\beta(M+N)(\eta(u) - \varepsilon)} \cdot \mathbf{E}_u\left[e^{-\beta X_1(M,N)}\right]\right)^Q \\ &\le e^{\beta(M+N)(\eta(u) - \varepsilon)} \cdot \left(e^{\beta(M+N)(\eta(u) - \varepsilon)} \cdot \left(e^{-\beta M(\eta(u) - \varepsilon/2)} + \mathbf{P}_u(X_1(M, N) < M(\eta(u) - \varepsilon/2))\right)\right)^Q \\ &\le e^{\beta(M+N)(\eta(u) - \varepsilon)} \cdot \left(e^{-\beta M\varepsilon/4} + e^{2\beta M(\eta(u) - \varepsilon)} \mathbf{P}_u(X_1(M, N) < M(\eta(u) - \varepsilon/2))\right)^Q. \end{aligned}$$

$$(60)$$

On the first line we used Markov's inequality; in the passage to the second line we used the fact that the $X_q(M, N)$ are i.i.d.; in the passage to the third

line we used the fact that $R \le (Q+1)(M+N)$; the passage to the fourth line is an immediate algebraic identity; the passage to the fifth line follows from the identity $\mathbf{E}_u[e^{-\beta X}] \le e^{-\beta a} + \mathbf{P}_u(X < a)$, valid for every positive random variable X; finally, the passage to the sixth line follows from the choice of N.

Recalling that $X_1(M, N) \sim s_{M,N}$, we obtain from Lemma 8 that $0 < \mathbf{P}_u(X_1(M, N) < M(\eta(u) - \varepsilon/2)) < 1$ for large enough values of M. For every such value of M one can find β_M, depending only on M, d, ε, u, such that

$$(\mathbf{P}_u(X_1(M, N) < M(\eta(u) - \varepsilon/2))^{-1/3} \le e^{2\beta_M M(\eta(u) - \varepsilon)} \le (\mathbf{P}_u(X_1(M, N) < M(\eta(u) - \varepsilon/2))^{-1/2}.$$

When combined with (60) this implies

$$\mathbf{P}_u\left(\sum_{q=0}^{Q-1} X_q(M, N) < R(\eta(u) - \varepsilon)\right) \le e^{\beta_M (M+N)(\eta(u) - \varepsilon)}$$

$$\cdot \left((\mathbf{P}_u(X_1(M, N) < M(\eta(u) - \varepsilon/2)))^{\frac{3\varepsilon}{24(\eta(u) - \varepsilon)}} + (\mathbf{P}_u(X_1(M, N) < M(\eta(u) - \varepsilon/2)))^{1/2}\right)^Q.$$

$$(61)$$

Using Lemma 8 one more time, we can choose $M = M(\varepsilon, d, u)$ large enough so that

$$(\mathbf{P}_u(X_1(M, N) < M(\eta(u) - \varepsilon/2)))^{\frac{3\varepsilon}{24\eta(u) - \varepsilon}} + (\mathbf{P}_u(X_1(M, N) < M(\eta(u) - \varepsilon/2)))^{1/2}$$

$$\le \left(32d \cdot \max\left\{2, \frac{8\eta(u)}{\varepsilon}\right\}\right)^{-d} \le \left(32d\frac{M}{N}\right)^{-d}.$$

$$(62)$$

For such choices of M and β, we get from (61), (62) and Proposition 1

$$\mathbf{P}_u\left(\text{dist}(0, Re_1) < R(\eta(u) - \varepsilon)\right) \le e^{\beta(M+N)(\eta(u) - \varepsilon)} \cdot \sum_{Q \ge \frac{R}{M+N} - 1} \left(2d\left(16\frac{M}{N}\right)^d\right)^Q \left(32d\frac{M}{N}\right)^{-dQ}$$

$$\le e^{\beta(M+N)(\eta(u) - \varepsilon)} \cdot \sum_{Q \ge \frac{R}{M+N} - 1} 2^{-dQ}$$

$$\le e^{\beta(M+N)(\eta(u) - \varepsilon)} \cdot 2^{-\frac{dR}{2M} + d + 1}.$$

Recalling that M, N and β are fixed, this gives the desired exponential decay in R.

Given the result for $x = Re_1$ we turn to deal with general points, $x \in \mathbb{R}^d$, $|x| = R$. Due to the invariance of $\mathbf{P}_u$ under rotations, $\text{dist}(0, x)$ for $|x| = R$ has the same distribution as $\text{dist}(0, Re_1)$, and therefore for large enough R

$$\mathbf{P}_u(\text{dist}(0, x) < (\eta(u) - \varepsilon/2)R) < e^{-c_1 R}, \quad \forall x \in \mathbb{R}^d, \text{ such that } |x| = R.$$

Taking an $\varepsilon/2$-net $\mathcal{N}$ on S^{d-1} such that $|\mathcal{N}| \le \frac{C(d)}{\varepsilon^d}$ we get that

$$\mathbf{P}_u(\exists x \in \mathcal{N} \text{ such that } \text{dist}(0, Rx) < (\eta(u) - \varepsilon/2)R) \le \frac{C(d)}{\varepsilon^d} e^{-c_1 R}.$$

For every $x \in \mathbb{R}^d$ such that $|x| = R$ there exists a $y \in \mathcal{N}$, such that $|x/R - y| < \varepsilon/2$, and therefore

$$\text{dist}(0, Ry) \le \text{dist}(0, x) + |Ry - x| \le \text{dist}(0, x) + \frac{\varepsilon}{2}R.$$

Hence,

$$
\begin{aligned}
\mathbf{P}_u &\left(\exists x \in \mathbb{R}^d \; : \; |x| = R, \;\; \text{dist}(0, x) < (\eta(u) - \varepsilon)R \right) \\
&\le \mathbf{P}_u \left(\exists y \in \mathcal{N} \text{ such that } \text{dist}(0, Ry) < (\eta(u) - \varepsilon/2)R \right) \\
&\le \frac{C}{\varepsilon^d} e^{-c_1 R},
\end{aligned}
$$

which concludes the proof. ∎

Next, we turn to prove the upper bound in Theorem 4, which states that for every $u \in [0, u^*)$ and every $\varepsilon > 0$ there exists a constant c, depending only on d, u and ε, such that for R large enough

$$\mathbf{P}_u \left(\exists x \in \mathbb{R}^d \; : \; |x| = R, \;\; \text{dist}(0, x) > (\eta(u) + \varepsilon)R \right) < e^{-c_1 R}. \tag{63}$$

We start with the following lemma.

Lemma 9 *For $0 < S < R$ let*

$$p_{S,R} = \inf \left\{ \text{len}(\gamma) \; : \; \begin{matrix} \gamma \text{ is a path from } Se_1 \text{ to } Re_1 \\ \text{such that } H_S < \gamma < H_R \end{matrix} \right\}.$$

Then

$$\lim_{R \to \infty} \frac{\mathbf{E}_u[p_{0,R}]}{R} = \eta(u).$$

Proof We follow [7, Theorem 4.3.7]. Since $\text{dist}(0, Re_1) \le p_{0,R}$,

$$\liminf_{R \to \infty} \frac{\mathbf{E}_u[p_{0,R}]}{R} \ge \lim_{R \to \infty} \frac{\mathbf{E}_u[\text{dist}(0, Re_1)]}{R} = \eta(u). \tag{64}$$

For $k \ge 0$ define

$$p_{S,R}^k = \inf \left\{ \text{len}(\gamma) \; : \; \begin{matrix} \gamma \text{ is a path from } Se_1 \text{ to } Re_1 \\ \text{such that } H_{S-k} < \gamma < H_{R+k} \end{matrix} \right\},$$

so in particular, $p_{S,R} = p_{S,R}^0$. Since for every $k \ge 0$ and $0 < R_1 < R_2$,

$$p_{0,R_1+R_2}^k \le p_{0,R_1}^k + p_{R_1,R_1+R_2}^k \tag{65}$$

one can apply the sub-additive ergodic theorem for $p_{0,R}^k$ to obtain

$$\lim_{R \to \infty} \frac{\mathbf{E}_u[p_{0,R}^k]}{R} = \eta^k(u)$$

which, by the definition of $p_{0,R}^k$, satisfies

$$\eta(u) \le \eta^k(u) \le \eta^{k-1}(u) \le \eta^0(u), \quad \forall k \ge 1. \tag{66}$$

Noting that for every $R, k > 0$

$$p_{-k,R+k} \le p_{-k,0} + p_{0,R}^k + p_{R,R+k},$$

it follows by fixing k, taking expectation, dividing by R and taking the limit $R \to \infty$ that

$$\eta^0(u) \le \eta^k(u),$$

which together with (66) implies that $\eta^0(u) = \eta^k(u)$ for every fixed $k > 0$. Since for every fixed $\omega \in \Omega$ and $R > 0$, $p_{0,R}^k$ is a monotonically decreasing function in k converging to $\text{dist}(0, Re_1)$, it follows from the monotone convergence theorem that

$$\lim_{k \to \infty} \mathbf{E}_u[p_{0,R}^k] = \mathbf{E}_u[\text{dist}(0, Re_1)]. \tag{67}$$

Using once again the sub-additivity (65) we conclude that

$$\frac{\mathbf{E}_u[p_{0,R}^k]}{R} \ge \eta^k(u) = \eta^0(u)$$

for every $R > 0$ by Fekete's subadditive lemma. Combined with (67) this yields

$$\frac{\mathbf{E}_u[\text{dist}(0, Re_1)]}{R} \ge \eta^0(u).$$

Letting $R \to \infty$ implies $\eta^0(u) \le \eta(u)$. Together with (64), this completes the proof.

Proof of Theorem 4 upper bound: Proving the upper bound is in fact much simpler than proving the lower bound. Fix $\varepsilon > 0$. By Lemma 9 there exists $R_0 > 2$ such that

$$\frac{\mathbf{E}[p_{0,R_0-2}]}{R_0 - 2} < \eta(u) + \frac{\varepsilon}{5}. \tag{68}$$

For $i \ge 0$ let $X_i = p_{iR_0+2,(i+1)R_0}$. The random variables $(X_i)_{i \ge 0}$ are i.i.d. (with the same distribution as p_{0,R_0-2}), and

$$\text{dist}(0, Re_1) \le \sum_{i=0}^{\lfloor R/R_0 \rfloor - 1} X_i + R_0. \tag{69}$$

Using (68) and (69) we deduce that for every $R > 5R_0/\varepsilon$,

$$\mathbf{P}_u\left(\text{dist}(0, Re_1) > (\eta(u) + \varepsilon)R\right) \le \mathbf{P}_u\left(\sum_{i=0}^{\lfloor R/R_0 \rfloor - 1} X_i > (\eta(u) + \varepsilon - R_0/R)R\right)$$

$$\le \mathbf{P}_u\left(\sum_{i=0}^{\lfloor R/R_0 \rfloor - 1} (X_i - \mathbf{E}_u[X_i]) > \left(\eta(u) + \varepsilon - \frac{R_0}{R} - \frac{\mathbf{E}_u[X_1]}{R_0}\right)R\right)$$

$$\le \mathbf{P}_u\left(\sum_{i=0}^{\lfloor R/R_0 \rfloor - 1} (X_i - \mathbf{E}_u[X_i]) > \frac{3}{5}\varepsilon R\right).$$

Applying the function $x \mapsto e^{\beta x}$ (for some $\beta > 0$) to both sides in the last term, using the Markov inequality and then the independence of the X_i's, we can bound the last term on the right-hand side by

$$
e^{-\frac{3\beta\varepsilon}{5}R} \cdot \mathbf{E}_u\left[e^{\beta\left(\sum_{i=0}^{\lfloor R/R_0\rfloor-1}(X_i-\mathbf{E}_u[X_i])\right)}\right] = e^{-\frac{3\beta\varepsilon}{5}R} \cdot \mathbf{E}_u\left[e^{\beta(p_{0,R_0}-\mathbf{E}_u[p_{0,R_0}])}\right]^{\lfloor\frac{R}{R_0}\rfloor}
$$
$$
\leq e^{-\frac{3\beta\varepsilon}{5}R} \cdot \mathbf{E}_u\left[e^{\beta(p_{0,R_0}-\mathbf{E}_u[p_{0,R_0}])}\right]^{\frac{R}{R_0}}.
$$
(70)

Since $\mathbf{E}_u\left[e^{\beta(p_{0,R_0}-\mathbf{E}_u[p_{0,R_0}])}\right] \leq \mathbf{E}_u\left[e^{\beta p_{0,R_0}}\right] \leq e^{\beta R_0} < \infty$ we can expand the exponential inside the expectation into a power series in β and obtain that

$$
\mathbf{E}_u\left[e^{\beta(p_{0,R_0}-\mathbf{E}_u[p_{0,R_0}])}\right] = 1 + O(\beta^2).
$$

Hence (70) is bounded by $\exp\left(\left(\frac{O(\beta^2)}{R_0} - \frac{3\beta\varepsilon}{5}\right)R\right)$. By taking $\beta = \beta(\varepsilon) > 0$ small enough we can make the last term exponentially decaying in R, thus completing the proof. ∎

5.4 Proof of Theorem 5

The proof of Theorem 5.2 also follows ideas of Kesten [8]. Unlike the proof of Theorem 5.1, some additional elements are needed in order to apply to the continuous case. We start with some preliminary results.

Lemma 10 *There exists a constant $C_4 = C_4(d)$ such that for every $M > 0$, the Euclidean length of every true geodesic contained in a box of side length M is at most $C_4 M^{d+1}$.*

Proof Denote the box by B_M. Given a configuration $\omega \in \Omega$, and partition $\mathcal{S}(\omega) \cap B_M$ to convex connected subsets $A = \{A_i\}_i$. A true geodesic accumulates inside each A_i a Euclidean length of at most the diameter of the box, $\sqrt{d}M$. In addition, the length of the geodesic (with respect to dist) is bounded by the diameter of the box, hence the total Euclidean length accumulated outside $\mathcal{S}(\omega)$ is at most $\sqrt{d}M$. Thus, we can find a new geodesic contained in B_M whose Euclidean length is bounded by

$$
\sqrt{d}M\left(|A| + 1\right).
$$

One can choose A such that $|A|$ is bounded by the number of unit balls that can be fitted inside B_M without overlapp. By a volume consideration, this number is at most $\kappa_d \cdot M^d$, where κ_d is the volume of the Euclidean unit ball in $\mathbb{R}^d$.

We also recall the following variant on a result by Roy [23]; see also [17, Lemma 3.3]:

Theorem 7 ([23]) *For $x \in \mathbb{R}^d$ denote*

$$\mathcal{S}(x; \omega) = \begin{cases} \text{the connected component of } x \text{ in } \mathcal{S}(\omega), & x \in \mathcal{S}(\omega) \\ \emptyset, & x \notin \mathcal{S}(\omega) \end{cases}$$

Then, there exist for every $u < u_*$ positive constants C, c depending only on u and d such that

$$\mathbf{P}_u(\text{diam}_e(\mathcal{S}(0; \omega)) > t) \leq Ce^{-ct}.$$

In order to prove Theorem 5, as well as for future use, we will need a stronger version of Theorem 7:

Lemma 11 *Let*

$$W = \bigcup_{x \in [0,1]^d} \mathcal{S}(x; \omega).$$

Then, for every $u < u_*$ there exist positive constants C_5, c_6 depending only on u and d, such that

$$\mathbf{P}_u(\text{diam}_e(W) > t) \leq C_5 e^{-c_6 t}.$$

Proof Observe that every unit ball $\overline{B(x, 1)}$ for $x \in \mathbb{R}^d$ contains a point in the grid $\frac{1}{2}\mathbb{Z}^d$. Therefore, each connected component in W must contain a point from the set

$$L = \left\{ q \in \frac{1}{2}\mathbb{Z}^d \ : \ |q|_2 < 2 \right\}$$

whose size is finite (and depends only on d).

Assume that $\text{diam}_e(W) > t$. Since the number of connected components intersecting $[0, 1]^d$ is bounded by $|L|$, there exists a connected component whose Euclidean diameter is at least $(t - \sqrt{D})/|L|$. By Theorem 7,

$$\mathbf{P}_u(\text{diam}_e(W) > t) \leq \sum_{x \in L} \mathbf{P}_u(\text{diam}_e(\mathcal{S}(x) > (t - \sqrt{D})/|L|) \leq Ce^{-(c/|L|)t}.$$

Proof of Theorem 5: The proof follows by an argument very similar to the one used in the proof of Theorem 4. We start with Part (1), showing that there exist constants α, C_2 and c_3, such that for every $R > 0$,

$$\mathbf{P}_u \left(\begin{array}{c} \exists \text{ a true geodesic path starting at } 0 \\ \text{such that } \text{len}_e(\gamma) \geq R \text{ and } \text{len}(\gamma) < \frac{1}{\alpha} \text{len}_e(\gamma) \end{array} \right) \leq C_2 e^{-c_3 R}.$$

First observe that

$$\mathbf{P}_u \left(\begin{array}{c} \exists \text{ a true geodesic path starting at } 0 \\ \text{such that } \text{len}_e(\gamma) \geq R \text{ and } \text{len}(\gamma) < \frac{1}{\alpha} \text{len}_e(\gamma) \end{array} \right)$$

$$\leq \sum_{S=\lfloor R \rfloor}^{\infty} \mathbf{P}_u \left(\begin{array}{c} \exists \text{ a true geodesic path starting at } 0 \\ \text{such that } \text{len}_e(\gamma) \in [S, S+1] \text{ and } \text{len}(\gamma) < \frac{1}{\alpha}(S+1) \end{array} \right)$$

so it is enough to show that there exist constants α, C, c_3 such that for every natural number S

$$\mathbf{P}_u \left(\begin{array}{c} \exists \text{ a true geodesic path starting at } 0 \\ \text{such that } \text{len}_e(\gamma) \in [S, S+1] \text{ and } \text{len}(\gamma) < \frac{1}{\alpha}(S+1) \end{array} \right) \leq Ce^{-c_3 S}.$$

Fix $S \in \mathbb{N}$ and let γ be a true geodesic path starting at the origin such that $\text{len}_e(\gamma) \in [S, S+1]$. Fix $M \in \mathbb{N}$ and $N = 1$ and define the sequences t_q and x_q, $q = 1, \ldots, Q$ in the same way as in the proof of Proposition 1. By construction, each of the segments $\gamma|_{[t_q, t_{q+1}]}$ is contained in a box of side length $M + 1$. By Lemma 10 the Euclidean length of a true geodesic in each of the boxes is at most $C_4(M+1)^d$, hence

$$Q \geq \frac{S}{C_4(M+1)^{d+1}} - 1.$$

By repeating the argument of Proposition 1 we get that for every $\beta \geq 0$

$$\mathbf{P}_u \left(\begin{array}{c} \exists \text{ a true geodesic path starting from } 0 \\ \text{such that } \text{len}_e(\gamma) \in [S, S+1] \text{ and } \text{len}(\gamma) < \frac{1}{\alpha}(S+1) \end{array} \right)$$

$$\leq \sum_{Q \geq \frac{S}{C_4(M+1)^{d+1}} - 1} \left(2d\,(16M)^d \right)^Q \mathbf{P}_u \left(\sum_{q=0}^{Q-1} X_q(M, 1) < \frac{1}{\alpha}(S+1) \right)$$

$$\leq \sum_{Q \geq \frac{S}{C_4(M+1)^{d+1}} - 1} \left(2d\,(16M)^d \right)^Q e^{\beta(S+1)/\alpha} \mathbf{E}_u \left[\prod_{q=0}^{Q-1} e^{-\beta X_q(M,1)} \right]$$

$$= e^{\beta(S+1)/\alpha} \cdot \sum_{Q \geq \frac{S}{C_4(M+1)^{d+1}} - 1} \left(2d\,(16M)^d \right)^Q \mathbf{E}_u \left[e^{-\beta X_1(M,1)} \right]^Q.$$

With W as in Lemma 11,

$$\mathbf{P}_u \left(\text{diam}_e(W) > \frac{M}{2} \right) \leq C_5 e^{-c_6 M/2}.$$

Each connected component of $\mathcal{S}$ is composed $\mathbf{P}_u$-a.s. of finitely many balls; see Lemma 1. Hence connected components of $\mathcal{S}$, as well as W, are compact sets. Therefore, one can find $\delta = \delta(M) > 0$ such that

$$\mathbf{P}_u(\text{dist}(W, \mathcal{S} \setminus W) \geq \delta(M)) \geq 1 - C_5 e^{-c_6 M/2}.$$

It follows that with $\mathbf{P}_u$-probability at least $1 - 2C_5 e^{-c_6 M/2}$ the set W is contained within the box $[-M/2, M/2]^d$ and is at a distance at least $\delta(M)$ from any other cluster of $\mathcal{S}$, which in particular implies that $e^{-\beta X_1(M,1)} \leq e^{-\beta\delta}$. Consequently,

$$\mathbf{E}_u[e^{-\beta X_1(M,1)}] \leq e^{-\beta\delta(M)} + 2C_5 e^{-c_6 M/2}$$

and thus

$$\mathbf{P}_u \left(\begin{array}{c} \exists \text{ a true geodesic path starting at } 0 \\ \text{such that } \mathrm{len}_e(\gamma) \in [S, S+1] \text{ and } \mathrm{len}(\gamma) < \tfrac{1}{\alpha}(S+1) \end{array} \right)$$

$$\leq e^{\beta(S+1)/\alpha} \sum_{Q \geq \frac{S}{C_4(M+1)^d} - 1} \left(2d(16M)^d \right)^Q \left(e^{-\beta\delta(M)} + 2C_5 e^{-c_6 M/2} \right)^Q.$$

Taking M large enough so that $2d(16M)^d \cdot 4C_5 e^{-c_6 M/2} < \frac{1}{4}$, and then β large enough so that $e^{-\beta\delta(M)} \leq 2C_5 e^{-c_6 M/2}$, and finally α large enough so that $e^{\beta/\alpha} < 2^{\frac{1}{C_4(M+1)^{d+1}}}$ gives

$$\mathbf{P}_u \left(\begin{array}{c} \exists \text{ a true geodesic path starting from } 0 \\ \text{such that } \mathrm{len}_e(\gamma) \in [S, S+1] \text{ and } \mathrm{len}(\gamma) < \tfrac{1}{\alpha}(S+1) \end{array} \right) \leq e^{\beta/\alpha} \left(\frac{1}{2} \right)^{\frac{R}{C_4(M+1)^{d+1}}},$$

proving Part (1) of Theorem 5.

We next prove Part (2), namely that for every $x, y \in \mathbb{R}^d$ and for every $R > \alpha|x - y|$

$$\mathbf{P}_u \left(\exists \gamma \in \Gamma_0(x, y) : \mathrm{len}_e(\gamma) > R \right) \leq C_2 e^{-c_3 R}.$$

By the invariance of the measure $\mathbf{P}_u$ under translations and rotations, it is sufficient to prove the result for the case $x = 0$, $y = Se_1$ for some $S > 0$. If $\gamma \in \Gamma_0(0, Se_1)$ is a path such that $\mathrm{len}_e(\gamma) > R > \alpha S$, then we can find times t_q and points $x_q, q = 0, \ldots, Q$, as in Proposition 1 (with $N = 1$), with $Q > \alpha S(M+1) - 1$. Since $\mathrm{len}(\gamma) = \mathrm{dist}(0, Se_1) \leq S$ the result follows by the same argument used to prove Part (1). ∎

6 Further result for uniform point distributions

In this section we exploit the results obtained in the previous two sections to prove more results for uniform point distributions. These include the geometric concentration of geodesics and further properties of the function $\eta(u)$.

6.1 Geometry of geodesics

In order to obtain results on the limiting distance in the setting of general intensity measures, we will need control over the geometry of geodesics, or approximate geodesics, connecting pairs of points. For $x, y \in \mathbb{R}^d$, we call a path γ between x and y an ε-geodesic if

$$\mathrm{len}(\gamma) - \mathrm{dist}(x, y) \leq \varepsilon|x - y|. \tag{71}$$

We denote by $\Gamma_\varepsilon(x, y)$ the set of ε-geodesics which are also Euclidean geodesics inside $\mathcal{S}$. In particular, $\Gamma_0(x, y)$ is the set of true geodesics between x and y defined in Section 5.

The goal of this subsection is to show that there exist geodesics that do not deviate significantly from Euclidean segments:

Proposition 2 *Let $u \in [0, u_*)$. There exist $C_7, c_8 > 0$ such that for every $\varepsilon > 0$ and sufficiently large (depending on ε) $|x - y|$,*

$$\mathbf{P}_u \left(\forall \gamma \in \Gamma_\varepsilon(x, y) \ : \ d_H(\gamma, [x, y]) > \varepsilon |x - y| \right) \leq C_7 e^{-c_8 |x-y|}, \tag{72}$$

where $[x, y]$ is the linear segment connecting x and y.

Proof We follow the proof of [3, Proposition 3.2]. Fix $\varepsilon > 0$ and $x, y \in \mathbb{R}^d$. We will show that if $|x - y|$ is large enough, then there exists with $\mathbf{P}_u$-probability $\geq 1 - C_7 e^{-c_8 |x-y|}$ a curve $\gamma \in \Gamma_\varepsilon(x, y)$ satisfying

$$d_H(\gamma, [x, y]) \leq \varepsilon |x - y|.$$

Let $N = \left\lceil \frac{20}{\varepsilon \alpha} \right\rceil$, with α as in Theorem 5. Define the sequence of vertices z_k, $k = 0, \ldots, N$, along $[x, y]$ by

$$z_k = \left(1 - \frac{k}{N} \right) x + \frac{k}{N} y, \quad \forall 0 \leq k \leq N.$$

For $0 \leq k \leq N - 1$, let $\gamma_k = \gamma_k(\omega) \in \Gamma_0(z_k, z_{k+1})$ be a true geodesic connecting z_k and z_{k+1} and define $\gamma = (\ldots (\gamma_0 * \gamma_1) * \ldots *) * \gamma_{N-1})$ to be their concatenation, connecting x and y.

Since $|z_k - z_{k+1}| = |x - y|/N$ is of order $|x - y|$ it follows from Theorem 4 that once $|x - y|$ is large enough

$$|\operatorname{dist}(z_k, z_{k+1}) - \eta(u) \cdot |z_k - z_{k+1}|| \leq \frac{\varepsilon}{2} |z_k - z_{k+1}|, \quad \forall 0 \leq k \leq N - 1, \tag{73}$$

with $\mathbf{P}_u$-probability at least $1 - e^{-c_1 |x-y|}$. Similarly, whenever $|x - y|$ is large enough we have with $\mathbf{P}_u$-probability at least $1 - e^{-c_1 |x-y|}$ that

$$|\operatorname{dist}(x, y) - \eta(u) \cdot |x - y|| \leq \frac{\varepsilon}{2} |x - y|. \tag{74}$$

Consequently, under the events in (73) and (74),

$$\begin{aligned}
\operatorname{len}(\gamma) &\leq \sum_{k=0}^{N-1} \operatorname{len}(\gamma_k) = \sum_{k=0}^{N-1} \operatorname{dist}(z_k, z_{k+1}) \\
&\leq \sum_{k=0}^{N-1} \left(\eta(u) \cdot |z_k - z_{k+1}| + \frac{\varepsilon}{2} |z_k - z_{k+1}| \right) \\
&= \eta(u) \cdot |x - y| + \frac{\varepsilon}{2} |x - y| \\
&\leq \operatorname{dist}(x, y) + \varepsilon |x - y|,
\end{aligned} \tag{75}$$

which implies that $\gamma \in \Gamma_\varepsilon(x, y)$. Finally, by Theorem 5(2) we have with $\mathbf{P}_u$-probability at least $1 - C_2 e^{-c_3 |x-y|}$ that

$$\operatorname{len}_e(\gamma_k) \leq \frac{\varepsilon}{10} |x - y|, \quad \forall 0 \leq k \leq N - 1 \tag{76}$$

which implies that for every $0 \le k \le N - 1$,

$$\gamma_k \subset [z_k, z_{k+1}] + B\left(0, \frac{\varepsilon}{5}|x - y|\right), \tag{77}$$

and thus

$$\gamma \subset [x, y] + B\left(0, \frac{\varepsilon}{5}|x - y|\right), \tag{78}$$

thus proving the existence of the required ε-geodesic.

6.2 An upper bound on $\eta(u)$

The goal of this subsection is to prove an upper bound on the function $\eta(u)$:

Proposition 3 *For every $u \in [0, u_*)$*

$$\eta(u) \le e^{-u\kappa_d}.$$

Proof First observe that $x \in \mathcal{S}(\omega)$ if and only if there is a point $y \in \mathrm{supp}(\omega) \cap \overline{B(x, 1)}$. Since the the number of points in the support of ω inside $\overline{B(x, 1)}$ is distributed like a Poisson random variable with parameter $u\kappa_d$ it follows that

$$\mathbf{P}_u(x \notin \mathcal{S}(\omega)) = e^{-u\kappa_d}.$$

For $R > 0$ let $\gamma_R : [0, 1] \to \mathbb{R}^d$ denote the path $\gamma_R(t) = tRe_1$. Then

$$\frac{\mathrm{dist}(0, Re_1)}{R} \le \frac{\mathrm{len}(\gamma_R)}{R} = \frac{1}{R} \int_0^R \mathbb{1}_{xe_1 \notin \mathcal{S}(\omega)} \, \mathrm{Leb}_1(dx).$$

Taking expectation on both sides and using Fubini's theorem,

$$\frac{\mathbf{E}_u[\mathrm{dist}(0, Re_1)]}{R} \le \frac{1}{R} \int_0^R e^{-u\kappa_d} \, \mathrm{Leb}_1(dx) = e^{-u\kappa_d}.$$

Taking the limit $R \to \infty$ and using Theorem 3 gives $\eta(u) \le e^{-u\kappa_d}$.

6.3 Continuity of $\eta(u)$

Proposition 4 *The function $\eta : [0, u_*) \to (0, 1]$ is continuous. In addition, η is monotonically decreasing in u and $\eta_0 = 1$.*

We start the proof by introducing a natural coupling of the probability measures $\mathbf{P}_u$ for $u \ge 0$. Let

$$\widehat{\Omega} = \left\{ \widehat{\omega} = \sum_{i \ge 0} \delta_{(x_i, u_i)} \ : \ \begin{array}{l} x_i \in \mathbb{R}^d, \ u_i \in [0, \infty) \text{ for all } i \ge 0 \text{ and } \widehat{\omega}(A \times [0, u]) < \infty \\ \text{for all bounded, Borel-measurable } A \subset \mathbb{R}^d \text{ and } u \ge 0 \end{array} \right\}$$

and for $u \geq 0$ and $\widehat{\omega} \in \widehat{\Omega}$ define

$$\widehat{\omega}_u = \sum_{\substack{(x_i, u_i) \in \mathrm{supp}\,\widehat{\omega}, \\ u_i \leq u}} \delta_{x_i}.$$

Note that $\widehat{\omega}_u \in \Omega$, so $\mathrm{dist}(\cdot\,; \widehat{\omega}_u)$ is well-defined. Finally, let $\mathbf{P}$ be the probability measure on $\widehat{\Omega}$ under which $\widehat{\omega}$ is distributed like a Poisson point process on $\mathbb{R}^d \times [0, \infty)$ with intensity $\mathrm{Leb}_d(dx) \times \mathrm{Leb}_1(dx)|_{[0,\infty)}$.

One can verify that for every $u \geq 0$, the distribution of ω under $\mathbf{P}_u$ is the same as the distribution of $\widehat{\omega}_u$ under $\mathbf{P}$. That is, we constructed a coupling of the probability measures $\mathbf{P}_u$ for $u \geq 0$ under which

$$\widehat{\omega}_u \leq \widehat{\omega}_{u'}, \quad \forall 0 \leq u \leq u'. \tag{79}$$

We will denote by $\mathbf{E}$ expectation with respect to the probability measure $\mathbf{P}$.

Proof of Proposition 4: Fix $u \in [0, u_*)$ and $\varepsilon > 0$. Choose $\delta_0 > 0$ such that $[u - \delta_0, u + \delta_0] \subset [0, u_*)$ if $u > 0$ and $[0, \delta_0] \subset [0, u_*)$ if $u = 0$. We denote this compact interval by I

Denote

$$A_R = \{\exists \gamma \in \Gamma_0(0, Re_1) \,:\, \mathrm{len}_\varepsilon(\gamma) > \alpha R\} \cup \{|\,\mathrm{dist}(0, Re_1) - \eta(u)R| > \varepsilon R/6\}, \tag{80}$$

where α is as in Theorem 5(2).

For every $u' \in I$ we have

$$\left| \eta_{u'} - \eta_u \right| = \left| \mathbf{E}\left[(\eta_{u'} - \eta_u) \mathbb{1}_{A_R^c}(\widehat{\omega}_{u'}) \mathbb{1}_{A_R^c}(\widehat{\omega}_u) \right] \right| + \mathbf{P}(A_R(\widehat{\omega}_{u'})) + \mathbf{P}(A_R(\widehat{\omega}_u))$$

$$\leq \frac{1}{R} \left| \mathbf{E}\left[[\mathrm{dist}(0, Re_1; \widehat{\omega}_{u'}) - \mathrm{dist}(0, Re_1; \widehat{\omega}_u)] \cdot \mathbb{1}_{A_R^c}(\widehat{\omega}_{u'}) \mathbb{1}_{A_R^c}(\widehat{\omega}_u) \right] \right| \tag{81}$$

$$+ \frac{\varepsilon}{3} + \mathbf{P}(A_R(\widehat{\omega}_{u'})) + \mathbf{P}(A_R(\widehat{\omega}_u)).$$

By going back to the proofs of Theorem 4 and Theorem 5(2) one can verify that both α and the constants c_1, C_2, c_3 can be chosen uniformly on the compact interval I. That is, there exist positive constants α, C_9 and c_{10} depending on u and δ_0 such that for every $u' \in I$, $\mathbf{P}_{u'}(A_R) \leq C_9 e^{-c_{10}R}$.

Due to the uniform bound on the probability of A_R on the interval I one can choose R large enough (depending only on $\varepsilon > 0$ and I) so that

$$\mathbf{P}(A_R(\widehat{\omega}_{u'})) + \mathbf{P}(A_R(\widehat{\omega}_u)) \leq \frac{\varepsilon}{3}. \tag{82}$$

Combining (81) and (82), it is enough to show the existence of $\delta > 0$ such that for $|u - u'| < \delta$,

$$\frac{1}{R} \left| \mathbf{E}\left[[\mathrm{dist}(0, Re_1; \widehat{\omega}_{u'}) - \mathrm{dist}(0, Re_1; \widehat{\omega}_u)] \cdot \mathbb{1}_{A_R^c}(\widehat{\omega}_{u'}) \mathbb{1}_{A_R^c}(\widehat{\omega}_u) \right] \right| < \frac{\varepsilon}{3}.$$

This holds since on the event $\widehat{\omega}_{u'}, \widehat{\omega}_u \in A_R^c$ the values of $\mathrm{dist}(0, Re_1; \widehat{\omega}_{u'})$ and $\mathrm{dist}(0, Re_1; \widehat{\omega}_u)$ are the same whenever

$$\mathrm{supp}(\widehat{\omega}_{u'}) \cap B(0, R\alpha + 1) = \mathrm{supp}(\widehat{\omega}_u) \cap B(0, R\alpha + 1),$$

and therefore

$$\frac{1}{R} |\mathbf{E}[\mathrm{dist}(0, Re_1; \widehat{\omega}_{u'}) \cdot \mathbb{1}_{A_R^c}(\widehat{\omega}_{u'}) - \mathrm{dist}(0, Re_1; \widehat{\omega}_u) \cdot \mathbb{1}_{A_R^c}(\widehat{\omega}_u)]|$$
$$\leq 2 \cdot \mathbf{P}\left(\mathrm{supp}(\widehat{\omega}_{u'}) \cap B(0, R\alpha + 1) \neq \mathrm{supp}(\widehat{\omega}_u) \cap B(0, R\alpha + 1)\right).$$

Assume without loss of generality that $u' \leq u$. Due to the coupling, the point measure $\widehat{\omega}_u$ is obtained from $\widehat{\omega}_{u'}$ by adding to it an additional independent point measure in Ω which is distributed as a Poisson point process with intensity measure $(u - u') \cdot \mathrm{Leb}_d(dx)$. In particular, the probability that there is an additional point inside the ball $B(0, R\alpha + 1)$ is

$$1 - e^{-(u - u')\kappa_d (R\alpha + 1)^d} \leq 1 - e^{-\delta \kappa_d (R\alpha + 1)^d}.$$

Thus, with probability at most $1 - e^{-\delta \kappa_d (R\alpha + 1)^d}$ the point measures $\widehat{\omega}_{u'}$ and $\widehat{\omega}_u$ do not coincide inside the ball $B(0, R\alpha + 1)$. Recalling that R depends only on ε and I we can choose δ small enough so that $2(1 - e^{-\delta \kappa_d (R\alpha + 1)^d}) < \varepsilon/3$, thus completing the proof. ∎

6.4 Volume convergence

Proposition 5 *For every $u \geq 0$, $\mathbf{P}_u$-almost surely*

$$\lim_{R \to \infty} \nu_R([0, 1]^d) = \lim_{M \to \infty} \frac{\mathrm{Leb}_d([0, M]^d \setminus \mathcal{S}(\omega))}{\mathrm{Leb}_d([0, M]^d)} = e^{-u\kappa_d} = \mu_{\sigma(u)}([0, 1]^d). \qquad (83)$$

Proof As observed in Proposition 3

$$\mathbf{P}_u(0 \notin \mathcal{S}(\omega)) = e^{-u\kappa_d}.$$

Using the ergodicity of the model, see Lemma 2, and the ergodic theorem we can conclude that

$$\lim_{M \to \infty} \frac{\mathrm{Leb}_d([0, M]^d \setminus \mathcal{S}(\omega))}{\mathrm{Leb}_d([0, M]^d)} = \lim_{M \to \infty} \frac{1}{\mathrm{Leb}_d(B(0, M))} \int_{B(0,M)} \mathbb{1}_{x \notin \mathcal{S}(\omega)} \mathrm{Leb}_d(dx)$$
$$= \mathbf{E}_u[\mathbb{1}_{0 \notin \mathcal{S}(\omega)}] = e^{-u\kappa_d}.$$

7 Convergence for uniform point distributions

In this section, we exploit the results proved in Sections 4–6 to prove Parts 1 and 4 of Theorem 1 for the case when u is uniform and D is convex. These assumptions are relaxed in the next section. Since for uniform u, we have a natural coupling of the measures $\mathbf{P}_{u,R}$ using a single measure $\mathbf{P}_u$, we will have in fact a slightly stronger result than stated in Theorem 1.

The main result of this section is the following:

Theorem 8 *Let $u \in [0, u_*)$ and let $D \subset \mathbb{R}^d$ be a convex, compact d-dimensional manifold with corners. Then*

$$\lim_{R \to \infty} \sup_{x,y \in D} \left| d^D_{\eta(u)}(x, y) - \mathrm{dist}^D_R(\pi_R(x), \pi_R(y)) \right| = 0, \qquad \mathbf{P}_u\text{-a.s.} \qquad (84)$$

In particular, since π_R is onto, the sequence (D_R, dist^D_R) converges $\mathbf{P}_u$-a.s. to $(D, d^D_{\eta(u)})$ with respect to the Gromov-Hausdorff metric, where $d^D_{\eta(u)}$ should be interpreted as in (14).

Proof Denote $\mathcal{A}^D_R = (D_R, \mathrm{dist}^D_R)$, $\mathcal{B}^D_R = (D_R, \mathrm{dist}_R)$ and $C^D = (D, d^D_{\eta(u)})$. We will use the space $\mathcal{B}^D_R$ as an intermediate metric space in order to bound the distortion of $\pi^D_R : C^D \to \mathcal{A}^D_R$. We denote by $\tilde{\pi}^D_R$ the projection $D \to D_R$ when considered as a mapping between C^D and $\mathcal{B}^D_R$ and by Id the identity from D_R to itself when considered as a map between $\mathcal{B}^D_R$ and $\mathcal{A}^D_R$. See Figure 3 for an illustration.
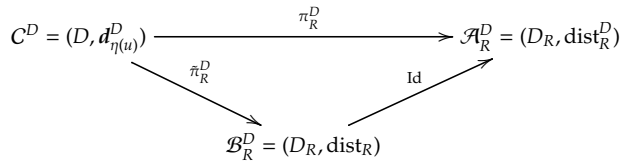
$$C^D = (D, d^D_{\eta(u)}) \xrightarrow{\quad \pi^D_R \quad} \mathcal{A}^D_R = (D_R, \mathrm{dist}^D_R)$$

Fig. 3 The spaces $\mathcal{A}^D_R, \mathcal{B}^D_R$ and C^D.

By the triangle inequality,

$$\mathrm{dis}\, \pi^D_R \le \mathrm{dis}\, \tilde{\pi}^D_R + \mathrm{dis}\, \mathrm{Id},$$

and thus it is enough to prove that with $\mathbf{P}_u$-probability one, both $\mathrm{dis}\, \tilde{\pi}^D_R$ and $\mathrm{dis}\, \mathrm{Id}$ go to zero as $R \to \infty$.

For $\mathrm{dis}\, \tilde{\pi}^D_R$, observe that D is compact and convex, and since u is constant, it follows that D is convex with respect to $d_{\eta(u)}$, hence $d^D_{\eta(u)} = d_{\eta(u)}$. The mapping $\tilde{\pi}^D_R : C^R_D \to \mathcal{B}^R_D$ is onto and by Theorem 2 has $\mathbf{P}_u$-a.s. vanishing distortion.

The rest of the proof shows that the distortion of $\mathrm{Id} : \mathcal{A}^D_R \to \mathcal{B}^D_R$ also vanishes asymptotically. We use the concentration result Proposition 2 as

follows: Fix $\delta > 0$, and denote by $D(\delta)$ the set of points in D whose distance from the boundary ∂D is greater than δ. For $R > 0$, let $N_{R,\delta}$ denote a finite $1/\sqrt{R}$-net of $D(\delta)$ such that $|N_{R,\delta}| < CR^{d/2}$ and for every $x, y \in N_{R,\delta}$, $|x - y| > c/\sqrt{R}$, with C and c depending only on d. Denote $0 < \varepsilon = \delta/2\,\mathrm{diam}(D)$, and let $x, y \in N_{R,\delta}$. By Proposition 2 we have that for $R > R_0(\delta)$,

$$\mathbf{P}_u\left(\forall \gamma \in \Gamma_\varepsilon(Rx, Ry)\ :\ d_H(\gamma, [Rx, Ry]) > \varepsilon R|x - y|\right) \leq C_7 e^{-c_8 R|x-y|}$$
$$< C_9 e^{-c_{10}\sqrt{R}}, \tag{85}$$

and therefore by a union bound argument we have

$$\mathbf{P}_u\left(\exists x, y \in N_{R,\delta}\ \forall \gamma \in \Gamma_\varepsilon(Rx, Ry)\ :\ d_H(\gamma, [Rx, Ry]) > \varepsilon R|x - y|\right) \leq C_{11} R^d e^{-c_{10}\sqrt{R}}. \tag{86}$$

Considering the sequence of events in (86) with R replaced by $m \in \mathbb{N}$, we get that the sum of the probabilities is finite, hence by the Borel-Cantelli lemma, we have that $\mathbf{P}_u$-a.s. there exists $M_0(\delta, \omega)$ such that for every $m \geq M_0(\delta, \omega)$,

$$\forall x, y \in N_{m,\delta}\ \exists \gamma \in \Gamma_\varepsilon(mx, my)\ :\ d_H(\gamma, [mx, my]) \leq \varepsilon m|x - y| < \frac{m\delta}{2}.$$

Since D is convex, $[x, y] \subset D$. Therefore by the definition of δ, if follows that for every $m \geq M_0$ and every $x, y \in N_{m,\delta}$ there exists an ε-geodesic with respect to dist_R that remains in D_R for R large enough, hence

$$\lim_{m\to\infty} \sup_{x,y\in N_{m,\delta}} |\,\mathrm{dist}_m(x, y) - \mathrm{dist}_m^D(x, y)| \leq \varepsilon, \quad \mathbf{P}_u\text{-a.s.}$$

Since $N_{m,\delta}$ is an $1/\sqrt{m}$-net of $D(\delta)$ is follows that

$$\lim_{m\to\infty} \sup_{x,y\in D(\delta)} |\,\mathrm{dist}_m(x, y) - \mathrm{dist}_m^D(x, y)| \leq \varepsilon, \quad \mathbf{P}_u\text{-a.s.}$$

and therefore

$$\lim_{m\to\infty} \sup_{x,y\in D} |\,\mathrm{dist}_m(x, y) - \mathrm{dist}_m^D(x, y)| \leq \varepsilon + 4\delta = \delta\left(4 + \frac{1}{2\,\mathrm{diam}(D)}\right), \quad \mathbf{P}_u\text{-a.s.}$$

Since for every $R > 0$ there exists an m such that $|m - R| < 1$, it follows that for such a choice of m, for every $x, y \in D$,

$$|\,\mathrm{dist}_m(x, y) - \mathrm{dist}_R(x, y)| = \left|\frac{\mathrm{dist}(mx, my)}{m} - \frac{\mathrm{dist}(Rx, Ry)}{R}\right|$$
$$\leq \frac{\mathrm{dist}(mx, Rx)}{m} + \frac{\mathrm{dist}(my, Ry)}{m} + \mathrm{dist}(Rx, Ry)\left|\frac{1}{m} - \frac{1}{R}\right|$$
$$\leq \frac{|x|_\infty}{R-1} + \frac{|y|_\infty}{R-1} + \frac{|x-y|_\infty}{R-1} \leq \frac{4\sup_{z\in D}|z|_\infty}{R-1}.$$

Similarly,

$$|\,\mathrm{dist}_m^D(x, y) - \mathrm{dist}_R^D(x, y)| \leq \frac{4\sup_{z\in D}|z|_\infty}{R-1}.$$

It follows that

$$\limsup_{R\to\infty} \sup_{x,y\in D} |\operatorname{dist}_R(x,y) - \operatorname{dist}_R^D(x,y)| \le \varepsilon + 4\delta = \delta\left(4 + \frac{1}{2\operatorname{diam}(D)}\right), \quad \mathbf{P}_u\text{-a.s.}$$

Since δ is arbitrary we finally have that

$$\lim_{R\to\infty} \sup_{x,y\in D} |\operatorname{dist}_R(x,y) - \operatorname{dist}_R^D(x,y)| = 0, \quad \mathbf{P}_u\text{-a.s.} \tag{87}$$

This shows that the identity mapping $\operatorname{Id} : \mathcal{B}_R^D \to \mathcal{A}_R^D$ has a vanishing distortion $\mathbf{P}_u$-a.s., which completes the proof.

The following proposition proves Part 4 of Theorem 1 for the case of constant u, that is, that $(\pi_R^D)^{-1} : D_R' \to D$ is asymptotically surjective.

Proposition 6
$$\lim_{R\to\infty} d_H(D, \pi_R^{-1}(D_R')) = 0, \quad \mathbf{P}_u\text{-a.s.}$$

Proof It follows from Lemma 11 that

$$\mathbf{P}_u\left(\begin{array}{l}\exists \text{ a connected component } C \text{ in } \bigcup_{x\in[-M,M]^d} \mathcal{S}(x;\omega) \\ \text{such that } \operatorname{diam}_e(C) \ge \log^2 M\end{array}\right)$$

$$\le \sum_{z\in\mathbb{Z}^d\cap[-M,M]^d} \mathbf{P}_u\left(\begin{array}{l}\exists \text{ a connected component } C \text{ in } \bigcup_{x\in z+[0,1]^d} \mathcal{S}(x;\omega) \\ \text{such that } \operatorname{diam}_e(C) \ge \log^2 M\end{array}\right)$$

$$\le C_5(2M+1)^d e^{-c_6 \log^2(M)}.$$

Taking $M = KR$, with $K = \operatorname{diam}_e(D)$ and $R > 0$, we get that with $\mathbf{P}_u$-probability at least $1 - C_5(2KR+1)^d e^{-c_6 \log^2(KR)}$ the distance between any point in $\mathcal{S} \cap R \cdot D$ and $R \cdot D \setminus \mathcal{S}$ is at most $\log^2(KR)$. Thus

$$\mathbf{P}_u\left(d_H(D, (\pi_R^D)^{-1}(D_R')) \ge \frac{\log^2(KR)}{R}\right) \le C_5(2KR+1)^d e^{-c_6 \log^2(KR)}. \tag{88}$$

Given $\varepsilon > 0$, consider the sequence $\left(d_H(D, (\pi_{m\varepsilon}^D)^{-1}(D_{m\varepsilon}'))\right)_{m\ge 1}$. Since the righthand side of (88) is summable in m, it follows from the Borel-Cantelli lemma, that $\mathbf{P}_u$-almost surely $d_H(D, (\pi_{m\varepsilon}^D)^{-1}(D_{m\varepsilon}')) \le \frac{\log^2(K\varepsilon m)}{\varepsilon m}$ for all but finitely many m's and in particular that

$$\lim_{m\to\infty} d_H(D, (\pi_{m\varepsilon}^D)^{-1}(D_{m\varepsilon}')) = 0.$$

Since for every $0 < R \le S$ we have $d_H((\pi_R^D)^{-1}(D_R'), (\pi_S^D)^{-1}(D_S')) \le |R - S|$ and since the sequence $\left(d_H(D, (\pi_{m\varepsilon}^D)^{-1}(D_{m\varepsilon}'))\right)_{m\ge 1}$ is ε dense in $[0,\infty)$ it follows that $\mathbf{P}_u$-a.s. $\limsup_{R\to\infty} d_H(D, (\pi_R^D)^{-1}(D_R')) \le \varepsilon$. Since $\varepsilon > 0$ is arbitrary we get that $\mathbf{P}_u$-a.s. $\lim_{R\to\infty} d_H(D, (\pi_R^D)^{-1}(D_R')) = 0$.

8 Proof of Theorem 1

In this section we prove Parts 1,2 and 4 of Theorem 1. Part 3 was proved in Proposition 3. Let $D \subset \mathbb{R}^d$ a compact d-dimensional manifold with corners, and let $u : D \to [0, u_*)$ be a continuous function. Since D is compact, we can always extend u continuously to $\mathbb{R}^d$ without enlarging its upper bound. Therefore, we can assume without loss of generality that $u : \mathbb{R}^d \to [0, u_*)$, with $\sup u < u_*$. The parameters u and $R > 0$ define a process with probability measure $\mathbf{P}_{u,R}$. Similarly to the proof of Proposition 4, we start by introducing a natural coupling of the probability measures $\mathbf{P}_{u,R}$ for a given $R > 0$.

Let

$$\widehat{\Omega} = \left\{ \widehat{\omega} = \sum_{i \geq 0} \delta_{(x_i, u_i)} \; : \; \begin{array}{l} x_i \in \mathbb{R}^d, \; u_i \in [0, \infty) \text{ for all } i \geq 0 \text{ and } \widehat{\omega}(A \times [0, u]) < \infty \\ \text{for all compact } A \subset \mathbb{R}^d \text{ and } u \geq 0 \end{array} \right\}$$

and for a continuous function $u : \mathbb{R}^d \to [0, \infty)$ and $\widehat{\omega} \in \widehat{\Omega}$ define

$$\widehat{\omega}_u = \sum_{\substack{(x_i, u_i) \in \operatorname{supp}\widehat{\omega}, \\ u_i \leq u(x_i)}} \delta_{x_i}.$$

Finally, let $\mathbf{P}_R$ be the probability measure on $\widehat{\Omega}$ under which $\widehat{\omega}$ is distributed like a Poisson point process on $\mathbb{R}^d \times [0, \infty)$ with intensity $R^d \operatorname{Leb}_d(dx) \times \operatorname{Leb}_1(dx)|_{[0,\infty)}$.

One can now verify, that for every continuous $u : \mathbb{R}^d \to [0, \infty)$, the distribution of ω under $\mathbf{P}_{u,R}$ is the same as the distribution of $\widehat{\omega}_u$ under $\mathbf{P}_R$. That is, we constructed a coupling of the probability measures $\mathbf{P}_{u,R}$, for $u : \mathbb{R}^d \to [0, \infty)$ under which

$$\widehat{\omega}_u \leq \widehat{\omega}_{u'}, \quad \forall u, u' : \mathbb{R}^d \to [0, \infty), \; u \leq u'.$$

8.1 Metric convergence

In this section we prove Part 1 of Theorem 1, which with the coupling constructed above states that:

For every $\varepsilon > 0$,

$$\lim_{R \to \infty} \mathbf{P}_R \left(\sup_{x,y \in D} \left| d^D_{\eta \circ u}(x, y) - \operatorname{dist}^D_R(\pi_R(x), \pi_R(y); \widehat{\omega}_u) \right| < \varepsilon \right) = 1. \qquad (89)$$

In particular, this implies that for every $\varepsilon > 0$,

$$\lim_{R \to \infty} \mathbf{P}_R (d_{GH}((D_R, \operatorname{dist}^D_R(\cdot; \widehat{\omega}_u)), (D, d^D_{\eta \circ u}) > \varepsilon) = 0.$$

Theorem 8 states that this holds when D is convex and $u = u_0$ is uniform.

We start with a proposition showing that if D is convex and $u : D \to [0, u_*)$ only takes values within a small interval $[u_{\min}, u_{\max}]$, then we can bound

(with high probability) the distortion of the projection π_R between $(D, \boldsymbol{d}^D_{\eta \circ u})$ and $(D_R(\widehat{\omega}_u), \text{dist}^D_R(\cdot; \widehat{\omega}_u))$.

Proposition 7 *Assume that D is a compact, path-connected and convex set, and let $u_{\min} = \min_D u \geq 0$ and $u_{\max} = \max_D u < u_*$. Consider π_R as a function $(D, \boldsymbol{d}^D_{\eta \circ u}) \to (D_R(\widehat{\omega}_u), \text{dist}^D_R(\cdot; \widehat{\omega}_u))$. Then*

$$\lim_{R \to \infty} \mathbf{P}_R(\text{dis } \pi_R > 3 \Delta \eta \cdot \text{diam}_\mathfrak{e} D) = 0,$$

where $\Delta \eta = \eta_{u_{\min}} - \eta_{u_{\max}}$.

Proof Consider the coupling as above between $\widehat{\omega}_u$, $\widehat{\omega}_{u_{\min}}$ and $\widehat{\omega}_{u_{\max}}$, with $u_{\min}$ and $u_{\max}$ viewed as constant functions. Then,

$$\text{dist}^D_R(\cdot; \widehat{\omega}_{u_{\min}}) \geq \text{dist}^D_R(\cdot; \widehat{\omega}_u) \geq \text{dist}^D_R(\cdot; \widehat{\omega}_{u_{\max}}) \quad \text{and} \quad \boldsymbol{d}^D_{\eta_{u_{\min}}} \geq \boldsymbol{d}^D_{\eta \circ u} \geq \boldsymbol{d}^D_{\eta_{u_{\max}}},$$

and in particular,

$$|\text{dist}^D_R(\cdot; \widehat{\omega}_{u_{\min}}) - \text{dist}^D_R(\cdot; \widehat{\omega}_u)| \leq |\text{dist}^D_R(\cdot; \widehat{\omega}_{u_{\min}}) - \text{dist}^D_R(\cdot; \widehat{\omega}_{u_{\max}})|, \quad (90)$$

and

$$|\boldsymbol{d}^D_{\eta_{u_{\min}}} - \boldsymbol{d}^D_{\eta \circ u}| \leq |\boldsymbol{d}^D_{\eta_{u_{\min}}} - \boldsymbol{d}^D_{\eta_{u_{\max}}}| \leq \Delta \eta \cdot \text{diam}_\varepsilon D. \quad (91)$$

Using (90) and (91), and by a repeated application of the triangle inequality,

$$\begin{aligned}
\text{dis } \pi_R &= \sup_{x,y \in D} |\boldsymbol{d}^D_{\eta \circ u}(x, y) - \text{dist}^D_R(x, y; \widehat{\omega}_u)| \\
&\leq \sup_{x,y \in D} |\boldsymbol{d}^D_{\eta \circ u}(x, y) - \boldsymbol{d}^D_{\eta_{u_{\min}}}(x, y)| + |\boldsymbol{d}^D_{\eta_{u_{\min}}}(x, y) - \text{dist}^D_R(x, y; \widehat{\omega}_{u_{\min}})| \\
&\quad + |\text{dist}^D_R(x, y; \widehat{\omega}_{u_{\min}}) - \text{dist}^D_R(x, y; \widehat{\omega}_u)| \\
&\leq \sup_{x,y \in D} |\boldsymbol{d}^D_{\eta_{u_{\max}}}(x, y) - \boldsymbol{d}^D_{\eta_{u_{\min}}}(x, y)| + |\boldsymbol{d}^D_{\eta_{u_{\min}}}(x, y) - \text{dist}^D_R(x, y; \widehat{\omega}_{u_{\min}})| \\
&\quad + |\text{dist}^D_R(x, y; \widehat{\omega}_{u_{\min}}) - \text{dist}^D_R(x, y; \widehat{\omega}_{u_{\max}})| \\
&\leq \sup_{x,y \in D} 2|\boldsymbol{d}^D_{\eta_{u_{\max}}}(x, y) - \boldsymbol{d}^D_{\eta_{u_{\min}}}(x, y)| + 2|\boldsymbol{d}^D_{\eta_{u_{\min}}}(x, y) - \text{dist}^D_R(x, y; \widehat{\omega}_{u_{\min}})| \\
&\quad + |\boldsymbol{d}_{\eta_{u_{\max}}}(x, y) - \text{dist}^D_R(x, y; \widehat{\omega}_{u_{\max}})| \\
&\leq 2 \Delta \eta \cdot \text{diam}_\varepsilon D + 2|\boldsymbol{d}^D_{\eta_{u_{\min}}}(x, y) - \text{dist}^D_R(x, y; \widehat{\omega}_{u_{\min}})| \\
&\quad + |\boldsymbol{d}_{\eta_{u_{\max}}}(x, y) - \text{dist}^D_R(x, y; \widehat{\omega}_{u_{\max}})|.
\end{aligned}$$
$$(92)$$

The result now follows by applying Theorem 8 to the last two addends.

We now prove Part 1 of Theorem 1. The idea behind the proof is the following: We partition D into small convex sets, in each u varies only a little. Proposition 7 states that with high probability, the distortion of the projection in each set is small. We then glue the sets together and show that the accumulated distortion remains small. A technical complication arises when D cannot be partitioned into finitely many convex sets. We overcome this problem by considering sets slightly larger and slightly smaller than D, denoted by $\overline{D}$ and $\underline{D}$, that can be partitioned in such a way.

Step I: Partitioning D. Let n be a large natural number to be chosen later (independent of R). Cover D with cubes of edge length $1/n$, whose corners are on the lattice $\frac{1}{n}\mathbb{Z}^d$; henceforth, "vertices" refers to the corners of the cubes. Denote the cubes that intersect D but not ∂D by $\square_{n,1}, \dots, \square_{n,k_n}$, and those that intersect ∂D by $\square_{n,k_n+1}, \dots, \square_{n,k_n+m_n}$. Since D is compact, there exists a slightly larger compact set $\square$ containing $\square_n = \bigcup_{i=1}^{k_n+m_n} \square_{n,i}$ for all $n \geq 1$. The function $\eta \circ u$ varies on each cube $\square_{n,i}$, by some $\Delta_{n,i}$. Denote $\Delta_n = \max_i \Delta_{n,i}$. Since $\eta \circ u$ is continuous (Proposition 4) and $\square$ is compact, $\Delta_n \to 0$ as $n \to \infty$. Denote by T_n the union of the facets of the cubes and let

$$\underline{D} = \bigcup_{i=1}^{k_n} \square_{n,i}, \qquad \overline{D} = \bigcup_{i=1}^{k_n+m_n} \square_{n,i}.$$

It follows from the definition of $\underline{D}$ and $\overline{D}$ that $\underline{D} \subset D \subset \overline{D}$ and both $|\operatorname{diam}\overline{D} - \operatorname{diam} D|$ and $|\operatorname{diam}\underline{D} - \operatorname{diam} D|$ are of order $1/n$. Moreover, we claim that

$$\lim_{n\to\infty} \sup_{x,y\in\underline{D}} |d_{\eta\circ u}^{\overline{D}}(x,y) - d_{\eta\circ u}^{\underline{D}}(x,y)| = 0. \tag{93}$$

Indeed, note that $d_{\eta\circ u}^{\overline{D}} \leq d_{\eta\circ u}^{\underline{D}}$ for pair of points in $\underline{D}$, and therefore we only need to prove that for every $x, y \in \underline{D}$ and every simple curve $\gamma \subset \overline{D}$ between x and y, there exists a simple curve $\gamma' \subset \underline{D}$ between x and y such that

$$\operatorname{len}_{d_{\eta\circ u}}(\gamma') < \operatorname{len}_{d_{\eta\circ u}}(\gamma) + o(1), \tag{94}$$

where $o(1)$ is with respect to n, independent of R. To prove (94), note that since $D \subset \mathbb{R}^d$ is compact, $\overline{D}$ and $\underline{D}$ are homotopic whenever n is large enough (∂D is a compact submanifold with corners and therefore its $1/n$-neighborhood is homotopic to itself for n large enough) and $d_H(\overline{D}, \underline{D}) < \sqrt{d}/n$. Then, for every simple curve $\gamma \subset \overline{D}$ between x and y in $\underline{D}$, there exists a simple curve $\gamma' \subset \underline{D}$ between x and y such that the $d_H(\gamma, \gamma') < 2\sqrt{d}/n$ and $\operatorname{len}_e(\gamma') < \operatorname{len}_e(\gamma) + c/n$ for some $c = c(d, D) > 0$ (see Figure 4). Since $\eta \circ u$ is continuous, γ and γ' are simple, and their Hausdorff distance is $o(1)$, (94) follows.

Step II: Each geodesic intersects only $O(n)$ cubes (in each of which the distortion is small). Consider a vertex of one of the cubes, and a facet of the same cube that does not intersect it. Since distances with respect to $d_{\eta\circ u}$ are bounded from below by $\underline{\eta} = \min_{\overline{D}} \eta(u)$ times the Euclidean distances, the distance between the vertex and the facet is at least $\underline{\eta}/n$. Therefore, a ball in $(D, d_{\eta\circ u}^D)$ of radius $\underline{\eta}/2n$ intersects at most 2^d cubes. It follows that all geodesics in $(D, d_{\eta\circ u}^D)$ (or more generally, every curve of length $\operatorname{diam}_\varepsilon D$ or less) intersect at most

$$\frac{\operatorname{diam}_e D}{\underline{\eta}/2n} \cdot 2^d = \frac{2^{d+1}\operatorname{diam}_\varepsilon D}{\underline{\eta}} \cdot n$$
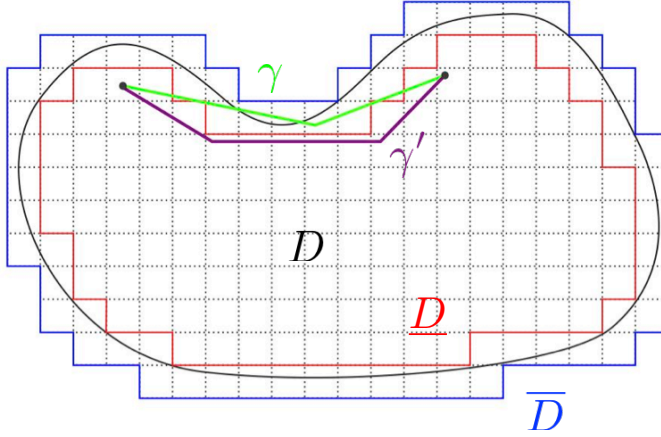
Fig. 4 Illustration of the construction of step I of the proof. The boundary of the domain D is marked in black, and the dotted lines represent the grid that forms $\square_{n,i}$. The boundaries of $\overline{D}$ and $\underline{D}$ are marked with blue and red lines, respectively. For every simple curve $\gamma \subset \overline{D}$ between two points in $\underline{D}$ there is a curve $\gamma' \subset \underline{D}$ of similar length which is close to γ in the Hausdorff distance (see (94)). Therefore, the distortion of the inclusion of $\underline{D}$ in $\overline{D}$ has a vanishing distortion as $n \to \infty$, as stated in (93).

of the cubes $\square_{n,1}, \ldots, \square_{n,k_n+m_n}$ (see Lemma 5.3 in [12] for a similar argument). By the same reasoning, each geodesic in either $(\overline{D}, d^{\overline{D}}_{\eta \circ u})$ or $(\underline{D}, d^{\underline{D}}_{\eta \circ u})$ intersects at most $2^{d+1} n \operatorname{diam} D/\eta$ cubes.

Next, cover similarly D_R by "cubes" $\square^R_{n,i} = \pi_R(\square_{n,i})$; also, define $\underline{D}_R$ and $\overline{D}_R$ in an analogous way. Applying Proposition 7 to each cube $\square_{n,i}$, we obtain that

$$\lim_{R \to \infty} \mathbf{P}_{u,R}(\forall i = 1, \ldots, k_n + m_n \ \operatorname{dis} \pi_R|_{\square_{n,i}} < 3\sqrt{d}\Delta_n/n) = 1. \tag{95}$$

For a given R, denote

$$A_R = \left\{ \forall i = 1, \ldots, k_n + m_n \ \operatorname{dis} \pi_R|_{\square_{n,i}} < 3\sqrt{d}\Delta_n/n \right\}.$$

Assume that A_R holds, and consider a vertex in one of the cubes in $\overline{D}_R$, and a facet of the same cube that does not intersect the vertex. The distance between the vertex and the facet is at least $(\eta - 3\sqrt{d}\Delta_n)/n$, which is positive for large enough n. By the same reasoning as above, each geodesic in either $\underline{D}_R$ or $\overline{D}_R$ intersects at most $2^{d+1} n \operatorname{diam} D/(\eta - 3\sqrt{d}\Delta_n)$ cubes.

Step III: Bounding the distortion. We now want to show that when A_R holds, i.e., the distortion within each cube is small, then the distortion of $\pi_R : (\overline{D}, d^{\overline{D}}_{\eta \circ u}) \to (D_R, \operatorname{dist}^{\overline{D}}_R)$ is small as well (and similarly with $\underline{D}$). The idea, which is similar

to the proof of Theorem 3.1 in [11], is that a geodesic γ in $(\overline{D}, d_{\eta \circ u}^{\overline{D}})$ crosses $O(n)$ cubes at points $x_1, \ldots x_m$, and at each crossing it accumulates a distance $d_{\eta \circ u}(x_i, x_{i+1})$. When A_R holds,

$$|d_{\eta \circ u}(x_i, x_{i+1}) - \operatorname{dist}_R^D(\pi_R(x_i), \pi_R(x_{i+1}))| \leq \frac{3\sqrt{d}\Delta_n}{n},$$

i.e. the distortion accumulated in each cube is $O(\Delta_n/n)$, so the total distortion is $O(n) \cdot O(\Delta_n/n) = O(\Delta) \to 0$.

Formally, let $x, y \in \overline{D}$, and let γ be a geodesic in $(\overline{D}, d_{\eta \circ u}^{\overline{D}})$ between x and y. If γ does not intersects the facets of the boxes transversely, we can take γ to be infinitesimally longer such that it does. Therefore, without loss of generality γ intersects the facets of the boxes at a finite number of points $x_1, \ldots, x_m$, with $m < 2^{d+1} n \operatorname{diam} D/\underline{\eta} = O(n)$. Denote $x = x_0$ and $y = x_{m+1}$. When the event A_R holds,

$$
\begin{aligned}
\operatorname{dist}_R^{\overline{D}}(\pi_R(x), \pi_R(y)) &\leq \sum_{i=1}^m \operatorname{dist}_R^D(\pi_R(x_i), \pi_R(x_{i+1})) \leq \sum_{i=1}^m \left(d_{\eta \circ u}(x_i, x_{i+1}) + \frac{3\sqrt{d}\Delta_n}{n} \right) \\
&\leq \sum_{i=1}^m d_{\eta \circ u}^{\overline{D}}(x_i, x_{i+1}) + m \cdot O(\Delta_n/n) = \operatorname{len}_{d_{\eta \circ u}^{\overline{D}}}(\gamma) + O(\Delta_n) \\
&= d_{\eta \circ u}^{\overline{D}}(x, y) + O(\Delta_n).
\end{aligned}
$$

A similar argument, on a geodesic σ between $\pi_R(x)$ and $\pi_R(y)$, shows that

$$d_{\eta \circ u}^{\overline{D}}(x, y) \leq \operatorname{dist}_R^{\overline{D}}(\pi_R(x), \pi_R(y)) + O(\Delta_n).$$

Therefore, we obtain that when A_R holds,

$$\sup_{x, y \in \overline{D}} |d_{\eta \circ u}^{\overline{D}}(x, y) - \operatorname{dist}_R^{\overline{D}}(\pi_R(x), \pi_R(y))| \leq C\Delta_n, \tag{96}$$

for some $C > 0$ independent of n. Similarly, when A_R holds,

$$\sup_{x, y \in \underline{D}} |d_{\eta \circ u}^{\underline{D}}(x, y) - \operatorname{dist}_R^{\underline{D}}(\pi_R(x), \pi_R(y))| \leq C\Delta_n. \tag{97}$$

If $D = \overline{D}$, then the proof is complete. The rest of the proof deals with the case where $\underline{D} \subsetneq D \subsetneq \overline{D}$. In $\underline{D}$ we have that $d_{\eta \circ u}^{\overline{D}} \leq d_{\eta \circ u}^D \leq d_{\eta \circ u}^{\underline{D}}$, hence

$$|d_{\eta \circ u}^D - d_{\eta \circ u}^{\underline{D}}| \leq |d_{\eta \circ u}^{\underline{D}} - d_{\eta \circ u}^{\overline{D}}|. \tag{98}$$

Similarly, $\operatorname{dist}_R^{\overline{D}} \leq \operatorname{dist}_R^D \leq \operatorname{dist}_R^{\underline{D}}$ in $\underline{D}_R$, hence

$$|\operatorname{dist}_R^D - \operatorname{dist}_R^{\underline{D}}| \leq |\operatorname{dist}_R^{\underline{D}} - \operatorname{dist}_R^{\overline{D}}|. \tag{99}$$

Using (93), (96)–(99), we obtain (similarly as in (92)) that when A_R holds,

$$\sup_{x,y\in \underline{D}} |d^D_{\eta\circ u}(x, y) - \text{dist}^D_R(\pi_R(x), \pi_R(y))|$$

$$\leq \sup_{x,y\in \underline{D}} |d^D_{\eta\circ u}(x, y) - d^D_{\overline{\eta}\circ u}(x, y)| + |d^D_{\overline{\eta}\circ u}(x, y) - \text{dist}^D_R(\pi_R(x), \pi_R(y))|$$

$$+ |\text{dist}^D_R(\pi_R(x), \pi_R(y)) - \text{dist}^D_R(\pi_R(x), \pi_R(y))|$$

$$\leq \sup_{x,y\in \underline{D}} |d^{\overline{D}}_{\overline{\eta}\circ u}(x, y) - d^D_{\overline{\eta}\circ u}(x, y)| + |d^D_{\overline{\eta}\circ u}(x, y) - \text{dist}^D_R(\pi_R(x), \pi_R(y))|$$

$$+ |\text{dist}^D_R(\pi_R(x), \pi_R(y)) - \text{dist}^{\overline{D}}_R(\pi_R(x), \pi_R(y))|$$

$$\leq \sup_{x,y\in \underline{D}} 2|d^{\overline{D}}_{\overline{\eta}\circ u}(x, y) - d^D_{\overline{\eta}\circ u}(x, y)| + 2|d^D_{\overline{\eta}\circ u}(x, y) - \text{dist}^D_R(\pi_R(x), \pi_R(y))|$$

$$+ |d^{\overline{D}}_{\overline{\eta}\circ u}(x, y) - \text{dist}^{\overline{D}}_R(\pi_R(x), \pi_R(y))|$$

$$\leq \sup_{x,y\in \underline{D}} 2|d^{\overline{D}}_{\overline{\eta}\circ u}(x, y) - d^D_{\overline{\eta}\circ u}(x, y)| + 3C\Delta_n = o(1),$$

where $o(1)$ here is with respect to n and independent of R. The first inequality is a triangle inequality; the second follows from (98) and (99); the third is again a triangle inequality; and the last one follows from (96) and (97).

Since $\underline{D}$ is a $1/n$-net in D and $\underline{D}_R$ is a $1/n$-net in D_R, we obtain that for $\pi_R : (D, d^D_{\eta\circ u}) \rightarrow (D_R, \text{dist}^D_R)$,

$$\text{dis} \, \pi_R = o(1). \tag{100}$$

Therefore, for every $\varepsilon > 0$,

$$\mathbf{P}_{u,R}(d_{GH}((D_R, \text{dist}^D_R), (D, d^D_{\eta\circ u})) > \varepsilon) \leq \mathbf{P}_{u,R}(\text{dis} \, \pi_R > \varepsilon) \leq \mathbf{P}_{u,R}(A^c_R) \rightarrow 0,$$

where A_R is with respect to n large enough, chosen according to ε, so that the $o(1)$ on the right-hand side of (100) is less than ε. Note that this choice is independent of R.

8.2 Measure convergence

In this section we prove part 2 of Theorem 1:

For every $\varepsilon > 0$,

$$\lim_{R\to\infty} \mathbf{P}_{u,R}\left(\exists f \in W(D), \ s.t. \ \left|\int_D f \, d\mu_{\sigma\circ u} - \int_{D_R} f \circ \pi_R^{-1} \, d\nu_R\right| > \varepsilon\right) = 0,$$

where

$$W(D) = \{f \in C(D) : \|f\|_\infty \leq 1 \ and \ Lip(f) \leq 1\},$$

$\sigma(u) = e^{-u\kappa_d/d}$, *and κ_d is the volume of the d-dimensional unit ball.*

We start with a proposition which is a measure analog of Proposition 7.

Proposition 8 *Let D be a cube in $\mathbb{R}^d$ and let $u_{\min} = \min_D u \geq 0$ and $u_{\max} = \max_D u < u_*$. Then*

$$\lim_{R \to \infty} \mathbf{P}_{u,R} \left(\left| \nu_R(D_R) - \mu_{\sigma \circ u}(D) \right| > 5 \Delta \sigma \, \mathrm{Leb}_d(D) \right) = 0,$$

where $\Delta\sigma = \sigma(u_{\min})^d - \sigma(u_{\max})^d$.

Proof As in the proof of Proposition 7 consider the coupling between $\widehat{\omega}_u$, $\widehat{\omega}_{u_{\min}}$ and $\widehat{\omega}_{u_{\max}}$. Since $u_{\min} \leq u \leq u_{\max}$, it follows that $\nu_R(\cdot; \widehat{\omega}_{u_{\min}}) \geq \nu_R(\cdot; \widehat{\omega}_u) \geq \nu_R(\cdot; \widehat{\omega}_{u_{\max}})$. Then,

$$\left| \nu_R(D_R; \widehat{\omega}_u) - \mu_{\sigma \circ u}(D) \right|$$
$$\leq \left| \nu_R(D_R; \widehat{\omega}_u) - \nu_R(D_R; \widehat{\omega}_{u_{\max}}) \right| + \left| \nu_R(D_R; \widehat{\omega}_{u_{\max}}) - \mu_{\sigma(u_{\max})}(D) \right| + \left| \mu_{\sigma(u_{\max})}(D) - \mu_{\sigma \circ u}(D) \right|$$
$$\leq \left| \nu_R(D_R; \widehat{\omega}_{u_{\min}}) - \nu_R(D_R; \widehat{\omega}_{u_{\max}}) \right| + \left| \nu_R(D_R; \widehat{\omega}_{u_{\max}}) - \mu_{\sigma(u_{\max})}(D) \right| + \left| \mu_{\sigma(u_{\max})}(D) - \mu_{\sigma(u_{\min})}(D) \right|$$
$$\leq \left| \nu_R(D_R; \widehat{\omega}_{u_{\min}}) - \mu_{\sigma(u_{\min})}(D) \right| + 2 \left| \nu_R(D_R; \widehat{\omega}_{u_{\max}}) - \mu_{\sigma(u_{\max})}(D) \right| + 2 \left| \mu_{\sigma(u_{\max})}(D) - \mu_{\sigma(u_{\min})}(D) \right|$$
$$\leq \left| \nu_R(D_R; \widehat{\omega}_{u_{\min}}) - \mu_{\sigma(u_{\min})}(D) \right| + 2 \left| \nu_R(D_R; \widehat{\omega}_{u_{\max}}) - \mu_{\sigma(u_{\max})}(D) \right| + 2 \Delta\sigma \, \mathrm{Leb}_d(D),$$

from which the result follows by Proposition 5.

We now prove part 3 of Theorem 1. In order to simplify the notation, we will consider ν_R as a measure on D (assigning zero measure to $\mathcal{S}_R$), rather than on D_R. In this way there is no composition with π_R^{-1}.

As in the proof of the first part, partition D into cubes $\square_{n,1}, \ldots, \square_{n,k_n+m_n}$, and let $\overline{D} = \bigcup_{i=1}^{k_n+m_n} \square_{n,i}$. We extend both $\mu_{\sigma \circ u}$ and ν_R to $\overline{D}$ by setting $\overline{D} \setminus D$ to be a null-set.

Let $\varepsilon > 0$ and choose $n > \sqrt{d}/\varepsilon$ large enough such that $\Delta\sigma < \varepsilon$ in each $\square_{n,i} \cap D$ and $\mathrm{Leb}_d(\overline{D}) < 2 \, \mathrm{Leb}_d(D)$. By applying Proposition 8 to each $\square_{n,i}$ we obtain that

$$\lim_{R \to \infty} \mathbf{P}_{u,R}(G_R) = 1,$$

where

$$G_R = \left\{ \left| \nu_R(\square_{n,i}) - \mu_{\sigma \circ u}(\square_{n,i}) \right| < 5\varepsilon \, \mathrm{Leb}_d(\square_{n,i}) \; \forall i = 1, \ldots k_n + m_n \right\}.$$

Let $f \in W(D)$. Due to the choice of n and since $\mathrm{Lip}(f) \leq 1$, it follows that $\sup_{\square_{n,i} \cap D} f - \inf_{\square_{n,i} \cap D} f < \varepsilon$ for every i. Denote $f_i = \inf_{\square_{n,i} \cap D} f$. Assume that G_R

holds. Then,

$$
\left| \int_D f \, d\mu_{\sigma \circ u} - \int_D f \, d\nu_R \right| \le \sum_{i=1}^{k_n + m_n} \left| \int_{\Box_{n,i}} f \, d\mu_{\sigma \circ u} - \int_{\Box_{n,i}} f \, d\nu_R \right|
$$

$$
\le \sum_{i=1}^{k_n + m_n} \left[\left| \int_{\Box_{n,i}} f \, d\mu_{\sigma \circ u} - f_i \, \mu_{\sigma \circ u}(\Box_{n,i}) \right| + |f_i| \left| \mu_{\sigma \circ u}(\Box_{n,i}) - \nu_R(\Box_{n,i}) \right| + \left| f_i \, \nu_R(\Box_{n,i}) - \int_{\Box_{n,i}} f \, d\nu_R \right| \right]
$$

$$
\le \sum_{i=1}^{k_n + m_n} \left[2\varepsilon \, \mathrm{Leb}_d(\Box_{n,i}) + \|f\|_\infty \left| \mu_{\sigma \circ u}(\Box_{n,i}) - \nu_R(\Box_{n,i}) \right| \right]
$$

$$
\le 7\varepsilon \sum_{i=1}^{k_n + m_n} \mathrm{Leb}_d(\Box_{n,i}) = 7\varepsilon \, \mathrm{Leb}_d(\overline{D}) < 14\varepsilon \, \mathrm{Leb}_d(D).
$$

Therefore,

$$
\mathbf{P}_{u,R} \left(\exists f \in W(D), \text{ s.t. } \left| \int_D f \, d\mu_{\sigma \circ u} - \int_D f \, d\nu_R \right| > 14\varepsilon \, \mathrm{Leb}_d(D) \right) \le \mathbf{P}_{u,R}(G_R^c) \to 0,
$$

which completes the proof.

8.3 Asymptotic surjectivity

In this section we prove Part 4 of Theorem 1:

For every $\varepsilon > 0$,

$$
\lim_{R \to \infty} \mathbf{P}_{u,R} \left(d_H(D, \pi_R^{-1}(D_R')) > \varepsilon \right) = 0, \tag{101}
$$

where $D_R' = \{ x \in D_R \ : \ |\pi_R^{-1}(x)| = 1 \}$ and d_H is the Hausdorff distance in $\mathbb{R}^d$.

This is an immediate consequence of Proposition 6, using the above coupling between u and $u_{\max} = \max_D u$. Denote $D_R' = \{ x \in D_R(\widehat{\omega}_u) \ : \ |\pi_R^{-1}(x; \widehat{\omega}_u)| = 1 \}$ and $\overline{D}_R' = \{ x \in D_R(\widehat{\omega}_{u_{\max}}) \ : \ |\pi_R^{-1}(x; \widehat{\omega}_{u_{\max}})| = 1 \}$. Then

$$
\pi_R^{-1}(\overline{D}_R') \subset \pi_R^{-1}(D_R') \subset D. \tag{102}
$$

By Proposition 6 for every $\varepsilon > 0$

$$
\lim_{R \to \infty} \mathbf{P}_R \left(d_H(D, \pi_R^{-1}(\overline{D}_R')) > \varepsilon \right) = 0, \tag{103}
$$

and combining (103) with (102) we obtain (101).

8.4 Proof of Corollary 1

We follow here definitions and notations of Definition 1.

Since $\pi_R : D \to D_R$ is always defined on the whole space D and always onto D_R, Part 1 of Theorem 1 as stated in (89) immediately implies (19), which proves the first part of Corollary 1.

For the second part, we need to prove that

(i) $\pi_R^{-1} : D'_{R_n} \to D$ is an ε_n-approximation
(ii) For every $f \in C(D)$,

$$\lim_{n \to \infty} \int_{D'_{R_n}} f \circ \pi_R^{-1} \, d\nu_R = \int_D f \, d\mu_{\sigma \circ u}. \tag{104}$$

Note that since $\nu_R(D_R \setminus D'_R) = 0$ by definition, we can replace the integrals on the left-hand side by integrals on D_R.

Part 1 (the event in (16)) implies that dis $\pi_{R_n}^{-1} < \varepsilon_n$. Part 4 (the event in (18)) implies that the ε_n-neighborhood of $\pi_{R_n}^{-1}(D'_{R_n})$ with respect to the Euclidean metric in $\mathbb{R}^d$ contains D, hence also with respect to $d^D_{\eta \circ u}$. By applying π_{R_n}, it follows that the ε_n-neighborhood of D'_{R_n} in $D_{R_n} = \pi_{R_n}(D)$ with respect to $d^D_{R_n}$ is indeed the whole D_{R_n}. This shows that $\pi_R^{-1} : D'_{R_n} \to D$ is indeed an ε_n-approximation.

To prove (104), note that it suffices to prove it for Lipschitz functions on D. Indeed, suppose we proved (104) for Lipschitz functions, and let $f \in C(D)$. For every $\varepsilon > 0$ there is a Lipschitz function g on D such that $\|f - g\|_\infty < \varepsilon$, hence

$$\limsup_n \left| \int_{D_{R_n}} f \circ \pi_R^{-1} \, d\nu_R - \int_D f \, d\mu_{\sigma \circ u} \right|$$

$$\leq \limsup_n \left| \int_{D_{R_n}} g \circ \pi_R^{-1} \, d\nu_R - \int_D g \, d\mu_{\sigma \circ u} \right| + \varepsilon \nu_R(D_{R_n}) + \varepsilon \mu_{\sigma \circ u}(D)$$

$$\leq 2\varepsilon \, \mathrm{Leb}_d(D).$$

To prove (104) for a Lipschitz function f, note that for $M > 0$ large enough, $f/M \in W(D)$. Therefore, when Part 2 (the event in (17)) holds,

$$\left| \int_{D_{R_n}} f \circ \pi_R^{-1} \, d\nu_R - \int_D f \, d\mu_{\sigma \circ u} \right| \leq M \varepsilon_n \to 0,$$

which completes the proof.

Disclosure of potential conflict of interest

None of the authors has any potential conflict of interest.

References

1. B. A. Bilby, R. Bullough, and E. Smith. Continuous distributions of dislocations: a new application of the methods of non-riemannian geometry. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 231, pages 263–273. The Royal Society, 1955.

2. I. Benjamini, G. Kalai, and O. Schramm. First passage percolation has sublinear distance variance. *Ann. Probab.*, 31(4):1970–1978, 2003.

3. M. Biskup, O. Louidor, E. B. Procaccia, and R. Rosenthal. Isoperimetry in two-dimensional percolation. *Communications on Pure and Applied Mathematics*, 68(9):1483–1531, 2015.

4. B. A. Bilby and E. Smith. Continuous distributions of dislocations. iii. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 236, pages 481–505. The Royal Society, 1956.

5. G. Grimmett and H. Kesten. First-passage percolation, network flows and electrical resistances. *Z. Wahrsch. Verw. Gebiete*, 66(3):335–366, 1984.

6. P. Hall. On continuum percolation. *Ann. Probab.*, 13(4):1250–1266, 1985.

7. J. M. Hammersley and D. J. A. Welsh. First-passage percolation, subadditive processes, stochastic networks, and generalized renewal theory. In *Proc. Internat. Res. Semin., Statist. Lab., Univ. California, Berkeley, Calif*, pages 61–110. Springer-Verlag, New York, 1965.

8. H. Kesten. Aspects of first passage percolation. In *École d'été de probabilités de Saint-Flour, XIV—1984*, volume 1180 of *Lecture Notes in Math.*, pages 125–264. Springer, Berlin, 1986.

9. H. Kesten. On the speed of convergence in first-passage percolation. *Ann. Appl. Probab.*, 3(2):296–338, 1993.

10. J. F. C. Kingman. Subadditive ergodic theory. *Ann. Probability*, 1:883–909, 1973. With discussion by D. L. Burkholder, Daryl Daley, H. Kesten, P. Ney, Frank Spitzer and J. M. Hammersley, and a reply by the author.

11. R. Kupferman and C. Maor. The emergence of torsion in the continuum limit of distributed dislocations. *J. Geo. Mech.*, 7(3):361–387, 2015.

12. R. Kupferman and C. Maor. Riemannian surfaces with torsion as homogenization limits of locally-euclidean surfaces with dislocation-type singularities. to appear in *Proc. Roy. Soc. Edin. A*, 2015.

13. R. Kupferman, M. Moshe and J.P. Solomon, Metric description of defects in amorphous materials. *Arch. Rat. Mech. Anal.*, 216:1009–1047, 2015.

14. E. Kröner et al. Continuum theory of defects. *Physics of defects*, 35:217–315, 1981.

15. E. Kröner. The differential geometry of elementary point and line defects in bravais crystals. *International Journal of Theoretical Physics*, 29(11):1219–1237, 1990.

16. T. M. Liggett. An improved subadditive ergodic theorem. *Ann. Probab.*, 13(4):1279–1285, 1985.

17. R. Meester and R. Roy. *Continuum percolation*, volume 119 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1996.

18. M.F. Miri and N. Rivier. Continuum elasticity with topological defects, including dislocations and extra-matter. *Journal of Physics A: Mathematical and General*, 35(7):1727, 2002.

19. W. Noll. Materially uniform simple bodies with inhomogeneities. *Arch. Rat. Mech. Anal.*, 27(1):1–32, 1967.

20. J. F. Nye. Some geometrical relations in dislocated crystals. *Acta metallurgica*, 1(2):153–162, 1953.

21. P. Petersen. *Riemannian geometry*, volume 171. Springer Science & Business Media, 2006.

22. S. I. Resnick. *Extreme values, regular variation, and point processes*, volume 4 of *Applied Probability. A Series of the Applied Probability Trust*. Springer-Verlag, New York, 1987.

23. R. Roy. The Russo-Seymour-Welsh theorem and the equality of critical densities and the "dual" critical densities for continuum percolation on $\mathbf{R}^2$. *Ann. Probab.*, 18(4):1563–1575, 1990.

24. C. C. Wang. On the geometric structures of simple bodies, a mathematical foundation for the theory of continuous distributions of dislocations. In *Arch. Rat. Mech. Anal.*, 27(1):33–94, 1967.

25. A. Yavari and A. Goriely. Weyl geometry and the nonlinear mechanics of distributed point defects. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, page rspa20120342. The Royal Society, 2012.