Weierstraß-Institut
für Angewandte Analysis und Stochastik

Leibniz-Institut im Forschungsverbund Berlin e. V.

Preprint ISSN 2198-5855

All spatial graphs with weak long-range effects have chemical distance comparable to Euclidean distance

Lukas Lüchtrath

submitted: December 18, 2024

Weierstrass Institute
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: lukas.luechtrath@wias-berlin.de

No. 3154
Berlin 2024

2020 *Mathematics Subject Classification.* 60K35, 90B15, 05C80.

Key words and phrases. Graph distances, spatial random graphs, Boolean model, weight-dependent random connection model, strong decay regime, polynomial correlations, long-range percolation.

Supported from the Leibniz Association within the Leibniz Junior Research Group on *Probabilistic Methods for Dynamic Communication Networks* as part of the Leibniz Competition (grant no. J105/2020).

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 20372-303
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

All spatial graphs with weak long-range effects have chemical distance comparable to Euclidean distance

Lukas Lüchtrath

Abstract

This note provides a sufficient condition for linear lower bounds on chemical distances (compared to the Euclidean distance) in general spatial random graphs. The condition is based on the scarceness of long edges in the graph and weak correlations at large distances and is valid for all translation invariant and locally finite graphs that fulfil these conditions. The proof is based on a renormalisation scheme introduced by Berger [arXiv: 0409021 (2004)].

1 Introduction and statement of results

A fundamental question in percolation theory concerns the scaling relationship between the graph distance, often referred to as the *chemical distance*, and the Euclidean distance between two vertices. Understanding this relationship provides insights into the geometric and structural properties of the infinite cluster in supercritical percolation models. Models for which the chemical distances were studied are for instances classical Bernoulli percolation [1, 10], its continuum analogue [30], random interlacements [6], and the Gaussian free field [9]. Although some of the models have long-range interactions, these models have in common that edges are of bounded length, and thus the chemical distance of two distant vertices of the infinite component is typically of the same order as their Euclidean distance. Conversely, introducing edges of unbounded lengths can lead to drastically different scaling behaviours, and chemical distances may depend logarithmically or even iterated logarithmically on the Euclidean distances [3, 8, 11, 24]. This property may be seen as a spatial version of the famous (ultra) *small world* property of complex networks [7]. However, unbounded edge lengths do not always lead to significantly shorter graph distances. Consider for instance long-range percolation [28], in which each pair of lattice sites is connected by an edge with a probability that decays polynomially with power $-d\delta$ in the vertices' distance. If $1 < \delta < 2$, then the graph distance between two distant vertices x, y is of order $\log^{\Delta} |x - y|$ for $\Delta = 1/\log_2(2/\delta)$ [3], while for $\delta > 2$ the graph distance is given by a linear function of the Euclidean distance [2, 4]. The reason for this dramatic change of behaviour is simply that for $\delta > 2$ long edges are too rare to give a significant advantage over the bounded edge lengths model.

This note builds on the work [2] and demonstrates that such linear scaling is a general feature of spatial graphs where long edges are sufficiently rare. Using the framework of [19], which relates long edges to subcritical annulus-crossing probabilities in graphs with weak correlations, we establish linear lower bounds on the chemical distance under broad assumptions. Furthermore, we demonstrate that the probabilistic rate of this is governed by the rate at which long edges are present, which has not been known even for long-range percolation before.

Framework. We use the framework of [19] and consider general translation-invariant models defined on some appropriate probability space whose probability measure we denote by $\mathbb{P}$. We aim to study

DOI 10.20347/WIAS.PREPRINT.3154 Berlin 2024

a countably infinite random graph $\mathscr{G} = (\mathscr{V}, \mathscr{E})$ whose vertex set $\mathscr{V}$ is built on the points of a suitable point process on $\mathbb{R}^d$. We make the following standing assumptions:

(G1) The locations of the vertices in $\mathscr{V}$ are given by a *locally finite, stationary, and simple point process* on $\mathbb{R}^d$ of intensity $\lambda > 0$. That is, the intensity measure is the Lebesgue measure multiplied by λ [25]. Note that this refers to the location of the vertices only and the vertices may carry additional marks or weights. In fact, in most examples, vertex marks are needed to model a vertex' attraction or influence in some way. Throughout the manuscript, we denote vertices by $\boldsymbol{x} \in \mathscr{V}$. For a vertex $\boldsymbol{x} \in \mathscr{V}$, we denote its location by $x \in \mathbb{R}^d$. Although $\mathscr{V}$ refers to the whole vertex set and may contain additional markings, we still write $\boldsymbol{x} \in \mathscr{V} \cap A$ for a vertex $\boldsymbol{x}$ with location $x \in A \subset \mathbb{R}^d$.

(G2) We assume that $\mathscr{G}$ is *translation invariant*. That is, $\mathscr{G}$ and $\mathscr{G} + x$ have the same distribution for each $x \in \mathbb{R}^d$, where $\mathscr{G} + x$ is the graph constructed on the points of $\mathscr{V} + x$, i.e., the connection mechanism does not change if the location of each vertex is shifted by x.

(G3) The graph $\mathscr{G}$ is *locally finite*. That is, all vertices have finite degree, almost surely. We also refer to this property as *sparseness*.

Remark 1.1.

(a) The parameter λ is the usual intensity parameter of continuum percolation. As studying the graph distance of far apart vertices is only reasonable if there is an infinite connected component containing both vertices, one may always assume that λ is chosen large enough to guarantee supercriticality. However, we only prove a lower bound on the graph distance that does not necessarily require the vertices being in the same component. Therefore, λ plays no particular role in the following and we thus drop it from the notation.

(b) Our setup also allows for models based on some site-percolation process on the lattice $\mathbb{Z}^d$, when translation invariance is with respect to shifts $x \in \mathbb{Z}^d$. In this case $\lambda = \mathbb{P}(o \in \mathscr{V})$, i.e. the probability that the origin survives percolation. If the remaining properties are fulfilled, all results hold verbatim for continuum or lattice models.

(c) The two standard examples for the underlying point process that are most frequently used in the literature are the homogeneous Poisson point process and the Bernoulli site-percolated lattice. Other examples are discussed in [19].

Quantifying long-rang effects. Next, we formulate the crucial properties for our main result, which are subject to the long-range effects of the graph. Typically, there are two ways in which these interactions can arise. One is through the presence of long edges in the graph that connect vertices far apart from each other. The other is via correlations of the local configurations of the graph across distant regions. While the amount of long edges is most relevant in order to deduce the typical graph distances' scale, we still require some control over the influence of the latter type on the graph topology. To employ our proof, we require both effects to vanish on a polynomial scale. More precisely, let $\Lambda_m(o)$ and $\Lambda_m(mx)$ be two boxes of side length m, centred in the origin o and in mx for some $|x| > 2$ respectively. Note that the two boxes are disjoint. Let $\mathcal{E}(\Lambda_m(o))$ be a local event, meaning that it can be decided by the realisation of the internal vertices and edges of $\Lambda_m(o)$ alone. Then, we say that $\mathscr{G}$ is *polynomially mixing* with mixing exponent $\xi < 0$ if there exists a suitable constant C_{mix} such that for all $|x| > 2$ and sufficiently large m,

$$\operatorname{Cov}\left(\mathbb{1}_{\mathcal{E}(\Lambda_m(o))}, \mathbb{1}_{\mathcal{E}(\Lambda_m(mx))}\right) \leq C_{\text{mix}}\, m^{-\xi}. \tag{$\mathscr{PM}$}$$

If $\mathscr{G}$ is polynomially mixing with exponent ξ, we also say that $\mathscr{G}$ has Property $\mathscr{PM}^{\xi}$. Let us further quantify the occurrence of long edges. To this end, define, for $n, m \in \mathbb{N}$, the event

$$\mathcal{L}(m,n) := \big\{\exists \boldsymbol{x} \sim \boldsymbol{y} : x, y \in \Lambda_m(o) \text{ and } |x-y| > n\big\},$$

where $\boldsymbol{x} \sim \boldsymbol{y}$ denotes the existence of an edge connecting the vertices $\boldsymbol{x}$ and $\boldsymbol{y}$. Therefore, $\mathcal{L}(m,n)$ describes the occurrence of long edges in a box. Henceforth, we say that $\mathscr{G}$ has the 'no long edges' Property $\mathscr{PL}^{\mu}$, if there exists an exponent $\mu < -d$ and a constant $C_{\mathcal{L}} > 0$ such that for all $m \geq 1$ and sufficiently large n, we have

$$\mathbb{P}(\mathcal{L}(m,n)) \leq C_{\mathcal{L}} m^d n^{\mu}. \tag{$\mathscr{PL}$}$$

Main result. We denote the graph distance in the graph $\mathscr{G}$ by $\mathrm{d} = \mathrm{d}_{\mathscr{G}}$, i.e.

$$\mathrm{d}(\boldsymbol{x},\boldsymbol{y}) := \mathrm{d}_{\mathscr{G}}(\boldsymbol{x},\boldsymbol{y}) = \min\big\{m : \exists \text{ a path } \boldsymbol{x} = \boldsymbol{x}_0 \sim \boldsymbol{x}_1 \sim \cdots \sim \boldsymbol{x}_m = \boldsymbol{y} \text{ in } \mathscr{G}\big\} \quad \text{for } \boldsymbol{x}, \boldsymbol{y} \in \mathscr{G},$$

with the usual convention $\min \emptyset = \infty$. In order to formulate our main theorem, let us define the event

$$\mathcal{D}^{\eta}_{L,M}(m) = \big\{\, \mathrm{d}(\boldsymbol{x},\boldsymbol{y}) \geq \eta|x-y|, \text{ for all } \boldsymbol{x} \in \Lambda_L \text{ and } \boldsymbol{y} \in \Lambda^{\mathsf{c}}_{Mm}\big\},$$

stating that all paths connecting a vertex located relatively close to the origin to a vertex faraway have length lower bounded by a constant multiple of the Euclidean distance of their end-vertices.

Theorem 1.2 (Linear graph distances, lower bound)**.** *If $\mathscr{G} = (\mathscr{V}, \mathscr{E})$ has the Properties $\mathscr{PM}^{\xi}$ and $\mathscr{PL}^{\mu}$, then there exist some constants $\eta > 0$ and $M \in \mathbb{N}$, depending only on model parameters, such that for all $L \in \mathbb{N}$,*

$$\limsup_{m\to\infty} \frac{\log \mathbb{P}\big(\neg \mathcal{D}^{\eta}_{L,M}(m)\big)}{\log m} \leq \xi \vee (d+\mu). \tag{1}$$

Let us, for the moment, assume that $d + \mu > \xi$. Then, Theorem 1.2 essentially states that the probability of finding a path that leaves a box of side length m with 'only a few steps' decays at least at the same rate as $\mathbb{P}(\mathcal{L}(m,m))$, which is the probability of having at least one edge of length m in said box. Put differently, either all paths contain a number of edges proportional to the box's side length, or there is one single edge spanning the whole Euclidean distance alone. A similar behaviour was observed in [19], where the existence of a subcritical annulus-crossing phase, i.e.

$$\widehat{\lambda}_c := \inf\big\{\lambda > 0 : \lim_{m\to\infty} \mathbb{P}(\text{Annulus with radii } m \text{ and } 2m \text{ is crossed by a path}) > 0\big\} > 0,$$

was investigated. In that article, the authors consider the 'long edges' event $\mathcal{L}(m) := \mathcal{L}(m,m)$ and show that, if $\mathbb{P}(\mathcal{L}(m)) \to 0$ as $m \to \infty$ uniformly in small intensities λ, then $\widehat{\lambda}_c > 0$, provided the model mixes in the sense that the defining covariance in ($\mathscr{PM}$) tends uniformly to zero. Conversely, if the convergence of $\mathbb{P}(\mathcal{L}(m))$ does not hold for any intensity λ, then there is no such phase, and annuli are typically crossed by a single long edge. The reason we require a slightly finer version of the 'long edges' event and the polynomial decay lies in the renormalisation scheme of Berger [2], our proof is built on. In order to employ it, we must exclude the existence of long edges that are on a smaller scale than the boxes, in which they are contained. However, both versions of the event are closely related. Clearly, Property $\mathscr{PL}^{\mu}$ implies polynomial decay of $\mathbb{P}(\mathcal{L}(m))$, and furthermore, polynomial decay of $\mathbb{P}(\mathcal{L}(m))$ also implies Property $\mathscr{PL}^{\mu}$ in all our examples below. Let us additionally note that, except for boundary cases, typical models of our interest either satisfy ($\mathscr{PL}$) with a polynomial decay or contain many long-edges.

Although the decay of the mixing term of Property $\mathscr{PM}^{\xi}$ could also be the dominant term in (1), we believe this to be a technical result and only the presence of long edges is truly significant. It is demonstrated in [9] that models with polynomial correlations, which are monotone in the intensity, can still exhibit *weak correlations* in the sense that monotone events in disjoint regions are independent after a sprinkling, depending on the Euclidean distance of the regions, with a stretched exponential error term. If the required sprinkling is summable over the scales, one should be able to adapt our proof in a way that the long edges always yield the dominant contribution. Note, however, that this property must be proved individually for each model, which is why we work with the weaker assumption of polynomially bounded correlations.

Finally, if one sets $a = -2/(\xi \vee (d+\mu))$ and $m_k = \lceil k^a \rceil$, then $\mathbb{P}(\mathcal{D}_{L,M}(m_k))$ is summable by Theorem 1.2. Thus, by the Borel-Cantelli lemma, almost surely, all paths connecting the vertices in Λ_L to a large distance become bounded from below by a constant multiple of their end-vertices' Euclidean distance eventually. This recovers the original result of Berger for long-range percolation [2] and extends it to the whole class of models that satisfy our assumptions.

Remark 1.3 (On the upper bound)**.** *Theorem 1.2 only establishes a linear lower bound on graph distances in terms of the Euclidean distance. The natural question is whether there exists a corresponding upper bound. This is believed to hold for all the models we are aware of. However, deriving this in full generality is challenging. Let us only consider models constructed on either a Poisson process or a site-percolated lattice, in which edges do only depend on their end vertices to ensure a sufficient amount of independence. For long-range percolation on the lattice, the upper bound and the corresponding shape-theorem were only recently proven in [4]. The proof relies on the continuity of the critical value, a property that is generally hard to prove. This continuity allows the removal of edges longer than an appropriate threshold, hence approximating the graph by one with bounded edge lengths. For this approximating graph, one can adapt the established results of [1, 5] to conclude the proof. Similarly, in [9] linear upper bounds for correlated percolation models on the lattice are proven in general, provided the model exhibits the aforementioned weak correlations of monotone events and local uniqueness of the infinite component. The latter property is again closely related to the continuity of the critical value. However, this continuity has only been established in regimes with many long-edges [26], and not in our regime, where long edges are rare. In a nutshell, long edges are too rare to use them constructively as in [26] but they still prevent us from applying localisation arguments from finite-range models.*

It is worth noting that typical examples often dominate either a long-range percolation model, a Gilbert graph or a model in the framework of [9]. If these underlying models are supercritical, one immediately observes linear upper bounds. Hence, in most examples chemical distances are linear in the Euclidean distance for sufficiently large intensities $\lambda \gg \lambda_c$, and we believe this to hold for all $\lambda > \lambda_c$.

2 Examples

Let us apply our result to a couple of examples. We shall only briefly sketch these examples and refer the reader to [19] for a more detailed discussion on the models.

The weight-dependent random connection model. This model was first introduced in [12, 15] as a general framework for inhomogeneous random connection models. It contains many models from the literature as special instances, cf. [12, Tab. 1]. In this model, $\mathscr{V}$ is either a standard Poisson point process or a (Bernoulli site-percolated) lattice where each vertex carries additionally an independent uniformly distributed vertex mark. We denote a vertex by $\boldsymbol{x} = (x, u_x) \in \mathbb{R}^d \times (0,1)$. Here, the mark

u_x models the *inverse weight* of a vertex, thus the smaller u_x the more attractive for connections the vertex $\boldsymbol{x}$ is. Currently, the model is most frequently studied using the following connection rule [14]: Given $\mathscr{V}$, each pair of vertices $\boldsymbol{x}$ and $\boldsymbol{y}$ is connected by an edge independently with probability

$$\rho\big((u_x \wedge u_y)^\gamma (u_x \vee u_y)^\alpha |x-y|^d\big), \qquad \text{where } \gamma \in [0,1) \text{ and } \alpha \in [0, 2-\gamma). \tag{2}$$

Here, ρ is a decreasing and integrable *profile function*, typically chosen to be either $\rho(x) = 1 \wedge |x|^{-\delta}$ for some $\delta > 1$ or $\rho = \mathbb{1}_{[0,1]}$. If one identifies the indicator function case with $\delta = \infty$, one can identify and compare many models from the literature by only three real parameters [14, Tab. 1].

It is clear that local events in disjoint regions are independent in this setting, and thus $\mathscr{PM}^{-\infty}$ is satisfied. It therefore remains to check for which choices of parameters Property $\mathscr{PL}^\mu$ holds in order to apply Theorem 1.2. To this end, the *downwards vertex boundary* exponent ζ, introduced in [21, 22], can be used. Loosely speaking, ζ quantifies the number of vertices within a box Λ_r that are connected to a vertex at distance r, which is weaker (i.e. has larger mark) than themselves (thus the term 'downward'). If $\zeta < 0$, then there are only a few such vertices, and long edges are rare. More precisely, it is shown in [19] that $\mathbb{P}(\mathcal{L}(m,m)) \asymp m^{d\zeta}$ in this case. Using the calculations and definitions of [19], it is not hard to deduce that $\zeta < 0$ implies Property $\mathscr{PL}^\mu$ for $\mu = d(\zeta - 1)$. More precisely, one can identify an optimal u_n, the mark a vertex should have in order to connect by a single edge to distance n. Following the reasoning of [21, 19], these optimal-mark vertices dominate the picture and, using their derived optimal mark, the number of such optimal-mark vertices in the box of volume m^d is of order $m^d n^{d(\zeta-1)}$. For the choice $m = n$, we arrive back at the original exponent $d\zeta = d + \mu$. Using the results of [19, Prop. 23], precisely identifying the $\zeta < 0$ regime of the model, we obtain:

Corollary 2.1. *Consider the weight-dependent random connection model* (2)*. If $\delta > 2$, $\gamma < 1 - 1/\delta$, and $\alpha < 1 - \gamma$, then*

$$\limsup_{m\to\infty} \log\big(\mathbb{P}\big(\neg \mathcal{D}^\eta_{L,M}(m)\big)\big) / \log(m) \le d\zeta.$$

Note that the model (2) always dominates a Gilbert graph and a linear upper bound on the chemical distances follows by domination for all large enough λ, cf. Remark 1.3 and [30]. Let us consider two models from the class of weight-dependent random connection models more closely.

The Boolean model corresponds to the choice of $\gamma > 0$, $\alpha = 0$, and $\delta = \infty$ [16]. That is, each vertex is assigned an independent Pareto distributed radius with tail exponent d/γ and each vertex is centre of a ball of its associated radius. Any two vertices are connected by an edge if the centre of the smaller ball is contained in the larger ball. It was shown in [18] that the chemical distances are lower bounded by $|x|/\log^p |x|$ for any power $p > 0$. However, it was conjectured that the logarithmic correction was a result of the method and that graph distances are indeed linearly lower bounded, cf. [18, Conj. 5]. Corollary 2.1 proves the conjecture.

The soft Boolean model corresponds to the choice of $\gamma > 0$, $\alpha = 0$, and $\delta \in (1, \infty)$ [11], which is a model that interpolates between the Boolean model and long-range percolation [20]. It was shown in [11] that chemical distances are doubly logarithmic in the Euclidean distance for all $\delta > 1$ and $\gamma > 1 - 1/(\delta+1)$. Further, [11] elaborates that distances are at least logarithmic in the Euclidean distance if $\gamma < 1 - 1/(\delta+1)$. We find with Corollary 2.1 that graph distances are at least linear if $\delta > 2$ and $\gamma < 1 - 1/\delta$, specifying the result of [11]. Recent results for scale-free percolation ($\alpha = \gamma > 0$) suggest that for $\delta < 2$ and $\gamma < 1/2$, graph distances are a power of the logarithm of the Euclidean distance, where the power is the one of long-range percolation [24]. It remains an interesting open problem to identify the power of the logarithm for the regime $\delta > 2$ and $1/\delta < \gamma < 1/(\delta+1)$. We leave this for future work.

Soft Boolean model with local interference. This model is an example for a *generalised weight-dependent random connection model*, in which edges not only depend on their end-vertices but also on surrounding vertex clouds [19]. The vertex set is the same as for the soft Boolean model and a vertex $\boldsymbol{x} = (x, u_x)$ still has assigned a sphere of influence of radius $u_x^{-d/\gamma}$ that is used together with additional long-range effects to form connections. Additionally, the vertex has assigned a sphere of interference of radius $u_x^{-d/\beta}$ for some $\beta < 1$. The vertices located in the latter sphere interfere and make it harder for the vertex to form connections. More precisely, two vertices $\boldsymbol{x} = (x, u_x)$ and $\boldsymbol{y} = (y, u_y)$ with $u_x < u_y$ are connected by an edge with probability

$$\mathbf{p}(\boldsymbol{x}, \boldsymbol{y}, \mathscr{V}) = \frac{1 \wedge (u_x^{\gamma}|x-y|^d)^{-\delta}}{\sharp\{\mathbf{z} \in \mathscr{V} : |x-z|^d < u_x^{-\beta}\}}.$$

It is elaborated in [19] that the exponent ζ can straight-forwardly be generalised to such a setting. It is further shown that the model mixes polynomially with exponent $\xi = 1 - 1/\beta$ and that it has $\zeta < 0$ if $\gamma < (\delta + \beta - 1)/\delta$. Hence, Theorem 1.2 applies in this case and proves linear lower bounds on the chemical distance, with probabilistic rate no slower than polynomially with exponent $\xi \vee \zeta$.

Ellipses percolation. Introduced in [29], this model can be seen as a generalisation of the planar Boolean model. Instead of a ball, each vertex is the centre of an ellipsis with $\mathrm{Pareto}(2/\gamma)$ distributed major axis, minor axis equal to one, and uniformly distributed orientation. Intriguingly, replacing balls by ellipses introduces additional correlations with interesting effects. The paper [17] elaborates graph distances in the regime $\gamma \in (1, 2)$ and shows that these scale doubly logarithmically in the Euclidean distance. In the $\gamma < 1$ regime however Property $\mathscr{PL}^{-2/\gamma}$ is satisfied, cf. [19], and Theorem 1.2 shows linear lower bounds on the graph distances. For a more detailed discussion on the correlations involved, their interesting effects, and generalisations of the model, we refer to the recent preprint [13].

More general underlying vertex locations. Theorem 1.2 also applies to models that are constructed on more correlated point clouds than a Poisson point process or a Bernoulli site-percolated lattice. For instance one could consider models constructed on a Cox process, a Gibbs process, or a lattice based on percolation of worms [27], which is done in greater detail in [19].

3 Proof of Theorem 1.2

In this section, we employ the renormalisation scheme that goes back to [2]. We start by defining the scales. Fix a large $K \in \mathbb{N}$ to be specified later and set $K_0 := K$ as well as $K_n := K(n!)^2$. A stage-n box is a box $B_n(x) := \Lambda_{K_n}(x) = x + [-K_n/2, K_n/2)^d$ of side length K_n, centred at $x \in \mathbb{R}^d$. We abbreviate $B_n = B_n(o)$. Let us define the notion of a *good* box; we refer to a box that is not good as being *bad*.

Definition 3.1 (Good boxes)**.** *Let $x \in \mathbb{R}^d$ and consider the boxes $B_n(x)$. We say that*

(i) the stage-0 box $B_0(x)$ is good, if it contains no internal edge longer than $K_0/100$.

(ii) We say that the stage-n box $B_n(x)$ is good if all 3^d boxes $B_n^{\mathbf{j}}(x) := B_n(x + \frac{K_{n-1}}{2}\mathbf{j})$, for $\mathbf{j} \in \{-1, 0, 1\}^d$, satisfy the following two conditions:

(a) No edge internal to $B_n^{\mathbf{j}}(x)$ is longer than $K_{n-1}/100$.

(b) Of all the n^{2d} stage-$(n-1)$ boxes contained in $B_n^{\mathbf{j}}(x)$, no more than 3^d are bad themselves.

Our first result is to determine the probability that a box centred at the origin is bad. Define

$$\psi(n) := \mathbb{P}(B_n \text{ is bad}).$$

Lemma 3.2. *If $\mathscr{G}$ has the Properties $\mathscr{PM}^\xi$ and $\mathscr{PL}^\mu$, then there exists $K \in \mathbb{N}$ such that*

$$\psi(n) \le \big((n+1)!\big)^{\xi \vee (d+\mu)}.$$

Proof. Let us define, for each $n \in \mathbb{N}$, the event $\mathcal{L}_n := \mathcal{L}(K_n, K_{n-1}/100)$, where $\mathcal{L}(m,n)$ is the defining event of Property $\mathscr{PL}^\mu$. We immediately infer

$$\mathbb{P}(\mathcal{L}_n) \le 100^{-\mu} C_{\mathcal{L}} K_n^d K_{n-1}^\mu = 100^{-\mu} C_{\mathcal{L}} n^{2d} K_{n-1}^{d+\mu}. \tag{3}$$

Observe that this implies, for each fixed n, that $\mathbb{P}(\mathcal{L}_n) \downarrow 0$, as $K \to \infty$. Let us further derive, for $n \ge 2$, a recursive formula for the probability of the stage-n box being bad. Such a box is bad, if one of the 3^d translated boxes B_n^{j} fails either Property (a) or (b). A union bound over the 3^d many boxes B_n^{j} combined with translation invariance (G2), thus yields

$$\psi(n) \le 3^d \big(\mathbb{P}(\mathcal{L}_n) + \mathbb{P}(\text{at least } 3^d+1 \text{ stage-}(n-1) \text{ boxes contained in } B_n \text{ are bad}) \big).$$

By construction, if there are at least 3^d+1 bad stage-$(n-1)$ boxes contained in B_n, then at least two of these are at distance K_{n-1}. Since there are no more than $\binom{n^{2d}}{2} \le n^{4d}$ many possibilities to choose two such boxes, we infer with a union bound, the mixing property $\mathscr{PM}^\xi$, and translation invariance (G2),

$$\psi(n) \le 3^d \mathbb{P}(\mathcal{L}_n) + 3^d n^{4d} \big(\psi(n-1)^2 + C_{\mathsf{mix}} K_{n-1}^\xi \big). \tag{4}$$

We now prove the claim by induction. Set $C = 3^d(100^{-\mu} C_{\mathcal{L}} + 1 + C_{\mathsf{mix}})$, write $\alpha = \xi \vee (d+\mu) < 0$, and let N be the smallest integer such that

$$C(n+1)^{-2\alpha+4d} ((n-1)!)^\alpha \le 1, \qquad \text{for all } n \ge N,$$

noting that this N exists. For $n = 1$, we now have

$$\psi(1) \le 3^d \mathbb{P}(\mathcal{L}_1) \le (2!)^\alpha,$$

for large enough K. Similarly, for $n = 2$, we find using (4) and $\psi(1) \le 3^d \mathbb{P}(\mathcal{L}_1')$,

$$\psi(2) \le 3^d \mathbb{P}(\mathcal{L}_2) + 3^d \cdot 2^{4d} \big(\psi(1)^2 + C_{\mathsf{mix}} K_1^\xi \big) \le 3^d \mathbb{P}(\mathcal{L}_2) + 3^{2d} \cdot 2^{4d}\, \mathbb{P}(\mathcal{L}_1)^2 + C_{\mathsf{mix}} K_1^\xi \le (3!)^\alpha,$$

perhaps increasing K if necessary. We proceed similarly N times, adapting K at each step if needed and infer $\psi(n) \le (n+1)!^\alpha$ for all $n = 1, \dots, N$. Note that K is finite still and remains unchanged from this point onwards. Now assume that the claim has been proven until some $n-1 \ge N$. Then, applying (3) and the induction hypothesis to (4), we infer

$$\begin{aligned} \psi(n) &\le 3^d\, 100^{-\mu} C_{\mathcal{L}} n^{2d} K^{d+\mu} (n-1)!^{2(d+\mu)} + 3^d \big(n!^{2\alpha} + C_{\mathsf{mix}} K^\xi (n-1)!^{2\xi} \big) \\ &\le C n^{-\alpha+4d} n!^{2\alpha} \le (n+1)!^\alpha \big(C(n+1)^{-2\alpha+4d} (n-1)!^\alpha \big) \le (n+1)!^\alpha, \end{aligned}$$

since $n > N$. This concludes the proof. □

In the remainder of this section, we shall always assume that K is chosen sufficiently large for Lemma 3.2 to hold. To make use of good boxes, we prove the following proposition, which is a stronger version of [2, Lemma 2] in the spirit of [23, Proposition 2.6] for cost distances in one-dependent first-passage percolation. Note that the following result holds deterministically on each realisation that fulfils the assumptions on the good boxes, the proposition requires.

Proposition 3.3. *Fix $N > (2d+1)9^d$. Consider $\boldsymbol{x}, \boldsymbol{y} \in \mathscr{V}$ such that $y \in B_n(x)$ and $|x-y| > K_{n-1}/8$ for some $n \geq N$. Assume that both boxes $B_n(x)$ and $B_{n-1}(x)$ are good. Then there exists a constant C_1, which does not depend on $\boldsymbol{x}$ and $\boldsymbol{y}$, such that each path connecting $\boldsymbol{x}$ and $\boldsymbol{y}$ that is completely contained in $B_n(x)$ consists of at least $C_1|x-y|$ many edges.*

Proof. We start by showing, inductively for all $n \geq N$, that if B_n' is some good box with $x, y \in B_n'$ and $|x-y| > K_{n-1}/16$, then every path π connecting the two vertices inside B_n' has length at least

$$\ell(\pi) := \sharp\{\text{edges of } \pi\} \geq C'\Lambda(n)|x-y|, \text{ for some } C' > 0, \text{ and } \Lambda(n) := \prod_{h=N}^{n} \left(1 - \tfrac{N}{h^2}\right). \quad (5)$$

We start with the *base case* $N = n$. Since B_N' is good, there is no edge contained in it longer than $K_{N-1}/100$. Furthermore, as $|x-y| > K_N/16$, we immediately infer $\ell(\pi) \geq (25/4)N^2$. Additionally, $|x-y| \leq \sqrt{d}K_N$. Thus, the claim follows for $C' = 1/(\sqrt{d}K_{N-1})$ as this implies $(25/4)N^2 \geq C'(1-1/N)|x-y|$.

For the *induction step*, assume that (5) holds for all stages up to stage $n-1$, for some $n > N$. Let $\boldsymbol{x}, \boldsymbol{y} \in B_n'$ with $|x-y| > K_n/16$ and B_n' be good. Let $\pi = (\boldsymbol{v}_0, \dots, \boldsymbol{v}_\ell)$ be a path connecting $\boldsymbol{x} = \boldsymbol{v}_0$ and $\boldsymbol{y} = \boldsymbol{v}_\ell$ that is completely contained in B_n'. Since B_n' is good, $|\boldsymbol{v}_i - \boldsymbol{v}_{i-1}| \leq K_{n-1}/100$. Moreover, at most 3^d of the stage-$(n-1)$ sub boxes in each translation $(B_n')^{\mathsf{j}}$ are bad. Hence, there are no more than 9^d bad sub boxes in the union of all translations. We denote the bad sub boxes by $Q_1, \dots, Q_j$ (for $j \leq 9^d$) and set $Q := Q_1 \cup \dots \cup Q_j$. We decompose π into alternating *good segments* π_s and *bad segments* σ_t such that $\pi = (\pi_1, \sigma_1, \pi_2, \dots, \sigma_T, \pi_S)$ for some S, T, where the last vertex in each segment is also the first vertex in the subsequent one. As, by construction, $S - 1 \leq T \leq 9^d$, we may simplify $S = T = 9^d$ in the following as this may only introduce additional empty segments at the end of the decomposition. With the same reasoning, we may always start the decomposition with a good segment because, if the first segment is bad, then simply $\pi_1 = \emptyset$. We use the following procedure to obtain the desired decomposition. Firstly, if $\pi \cap Q = \emptyset$, we simply choose $\pi_1 = \pi$ and all the other segments to be empty. Otherwise, define $i_1 = \min\{i \leq \ell : v_i \in Q\}$ the index of the first vertex contained in the bad region and j_1 the index of the bad box it is contained in, i.e. $v_{i_1} \in Q_{j_1}$, where we choose the smallest such index if two or more bad boxes overlap. However, our result does not depend on the precise ordering of the bad regions and the result holds verbatim for any other decision rule. We set $\pi_1 = (\boldsymbol{v}_0, \dots, \boldsymbol{v}_{i_1-1})$ (or $\pi_1 = \emptyset$ if $i_1 = 0$). Define further $k_1 = \max\{i : v_i \in Q_{j_1}\}$ and set $\sigma_1 = (\boldsymbol{v}_{i_1-1}, \dots, \boldsymbol{v}_{k_1+1})$. Let us remark that this segment may also leave and reenter the bad box Q_{j_1} multiple times but the path never returns to Q_{j_1} after the vertex with index k_1. Inductively, define $i_s = \min\{i > k_{s-1} : v_i \in Q\}$, $j_s = \min\{i : v_{i_s} \in Q_i\}$, and $k_s = \max\{i : v_i \in Q_{j_s}\}$, as well as the good segment $\pi_s = (\boldsymbol{v}_{k_{s-1}+1}, \dots, \boldsymbol{v}_{i_s-1})$, and the bad segment $\sigma_s = (\boldsymbol{v}_{i_s-1}, \dots, \boldsymbol{v}_{k_s+1})$ all the way up to S and T.

Denote for a path ρ by $\mathsf{dist}(\rho)$ the Euclidean distance between its two endpoints. The triangle inequality gives $|x-y| \leq \sum_1^{9^d} \mathsf{dist}(\pi_s) + \sum_1^{9^d} \mathsf{dist}(\sigma_t)$. Since each Q_{j_s} (a bad stage-$(n-1)$ box) has diameter $\sqrt{d}K_{n-1}$ and all edge lengths are bounded by $K_{n-1}/100$, the contribution of any bad segment is bounded by

$$\begin{aligned}\mathsf{dist}(\sigma_t) &\leq |v_{i_t-1} - v_{i_t}| + |v_{i_t} - v_{k_t}| + |v_{k_t} - v_{k_t+1}| \leq K_{n-1}/100 + \sqrt{d}K_{n-1} + K_{n-1}/100 \\ &\leq 2dK_{n-1}.\end{aligned}$$

Let $I = \{s : \mathsf{dist}(\pi_s) > K_{n-1}/2\}$. Then trivially, $\sum_{s \notin I} \mathsf{dist}(\pi_s) \leq 9^d K_{n-1}/2$. Combined with the

assumption $|x-y| > K_n/16$ and the fact $K_n = n^2 K_{n-1}$, we infer by use of the triangle inequality

$$\begin{aligned}\sum_{s\in I} \mathsf{dist}(\pi_s) &\geq |x-y| - \sum_{t=1}^{9^d} \mathsf{dist}(\sigma_t) - \sum_{s\notin I} \mathsf{dist}(\pi_s) \geq |x-y| - 2d\cdot 9^d K_{n-1} - \tfrac{9^d}{2} K_{n-1} \\ &\geq |x-y|\big(1-\tfrac{N}{n^2}\big)\end{aligned} \tag{6}$$

by our choice of N. Let us mention here, that for $|x-y| \geq 2NK_{n-1}$, we further get with the same calculation

$$\sum_{s\in I} \mathsf{dist}(\pi_s) \geq |x-y| - NK_{n-1} \geq |x-y|/2, \tag{7}$$

which will be useful below. To finish the induction proof of (5), we show that for each $s \in I$

$$\ell(\pi_s) \geq \mathsf{dist}(\pi_s)\cdot C'\Lambda(n-1) \tag{8}$$

since the claim then follows by summing (8) over the segments of I and applying (6). For $\boldsymbol{v}, \boldsymbol{w} \in \pi_s$, we write $\pi_s[\boldsymbol{v}, \boldsymbol{w}]$ for the path segment from $\boldsymbol{v}$ to $\boldsymbol{w}$. Observe that there exists a collection $\boldsymbol{v}_{k_{s-1}+1} = \boldsymbol{w}_1, \dots, \boldsymbol{w}_m = \boldsymbol{v}_{i_s-1}$ of vertices of π_s, such that for every i

(1) $|w_{i+1} - w_i| > K_{n-1}/16$, and

(2) $|v - w_i| < K_{n-1}/2$ for all $\boldsymbol{v} \in \pi_s[\boldsymbol{w}_i, \boldsymbol{w}_{i+1}]$.

That is, the $\boldsymbol{w}_i$ divide the good segment π_s into sub-segments $\pi_s[\boldsymbol{w}_i, \boldsymbol{w}_{i+1}]$ such that the whole sub-segment is contained in the ball of radius $K_{n-1}/2$, centred at w_i, but still a large Euclidean distance is bridged. This sequence can for instance be constructed using a greedy algorithm [23, Prop. 2.6]. To this end, one assumes inductively that $\boldsymbol{w}_1, \dots, \boldsymbol{w}_i$ have already been found and the segment π_s is not fully covered yet. Then, the remaining path is either contained in the ball of radius $K_{n-1}/2$, centred in $\boldsymbol{v}_{i_s-1}$ (the last vertex on π_s), in which case we choose $\boldsymbol{w}_{i+1} = \boldsymbol{v}_{i_s-1}$, or this is not the case and we follow the path and pick $\boldsymbol{w}_{i+1}$ to be the first vertex that fulfils both properties. Note that this is always possible as π_s is a good path segment and the assumption implies that either $\boldsymbol{v}_{i_s-1}$ is at a far distance from $\boldsymbol{w}_i$ or the paths wanders far off before reaching $\boldsymbol{v}_{i_s-1}$.

Now, by Property **(2)** and the choice of π_s, there exists a good stage-$(n-1)$ box containing $\pi_s[\boldsymbol{w}_i, \boldsymbol{w}_{i+1}]$. By **(1)** and the induction hypothesis (5) for $n-1$, we have

$$\ell(\pi_s[\boldsymbol{w}_i, \boldsymbol{w}_{i+1}]) \geq \mathsf{dist}(\pi_s[\boldsymbol{w}_i, \boldsymbol{w}_{i+1}])\cdot C'\Lambda(n-1),$$

using $|w_i - w_{i+1}| = \mathsf{dist}(\pi_s[\boldsymbol{w}_i, \boldsymbol{w}_{i+1}])$. Using $\ell(\pi_s) = \ell(\pi_s[\boldsymbol{w}_1, \boldsymbol{w}_2]) + \cdots + \ell(\pi_s[\boldsymbol{w}_{m-1}, \boldsymbol{w}_m])$ together with the triangle inequality, thus proves the claim (8).

To finish the proof of the proposition, set $C_2 := C'\Lambda(\infty)$, let $\boldsymbol{x}, \boldsymbol{y} \in \mathscr{V}$ such that $|x-y| > K_{n-1}/8$ and $B_{n-1}(x)$, $B_n(x)$ are good, and let π be a path connecting $\boldsymbol{x}$ and $\boldsymbol{y}$ within $B_n(x)$. Assume first that $|x-y| \geq 2NK_{n-1}$. Then, by combining (7) and (8), and the fact that $\Lambda(n)$ is decreasing in n,

$$\ell(\pi) \geq C_2 \sum_{s\in I} \mathsf{dist}(\pi_s) \geq \frac{C_2}{2}|x-y|.$$

If instead $|x-y| < 2NK_{n-1}$, then by the fact that $B_n(x)$ is good and no edge is longer than $K_{n-1}/100$, there must exists a vertex $\boldsymbol{v} \in \pi$, such that $K_{n-1}/16 < |x-v| \leq K_{n-1}/8$. If $\boldsymbol{v}$ is the first such vertex on the path, then additionally $\pi[\boldsymbol{x}, \boldsymbol{v}] \subset B_{n-1}(x)$. By assumption, the box $B_{n-1}(x)$ is good and we can apply (5) and the bounds on the Euclidean distance to obtain

$$\ell(\pi) \geq \ell(\pi[\boldsymbol{x}, \boldsymbol{v}]) \geq C_2|x-v| \geq C_2\frac{K_{n-1}}{16} \geq C_2\frac{|x-y|}{32N}.$$

Hence, the claim follows for $C_1 = C_2/32N$. □

Proof of Theorem 1.2. Fix $L \in \mathbb{N}$ and choose some $M \geq 2K$. For $m \in \mathbb{N}$, let $n = n_m$ be the largest integer satisfying $n! \leq m$. Assume n (and thus m) to be large enough such that $\Lambda_L \subset B_{n-3}$. Consider now a realisation ω, in which all the boxes $B_{n-1}, B_n, \dots$ are good. Let $\boldsymbol{x} \in \Lambda_L$ and $\boldsymbol{y} \in B_n^{\mathsf{c}}$ be two vertices in this realisation. Let n_y be the first index such that $\boldsymbol{y} \in B_{n_y} \setminus B_{n_y-1}$. Particularly, this implies $n_y \geq n$ as well as $\sqrt{d}(K_{n-1} - L)/2 < |x - y| < \sqrt{d}(K_{n_y} + L)/2$. Let further π be a path from $\boldsymbol{x}$ to $\boldsymbol{y}$ and let $n_\pi \geq n_y$ be the first index such that $\pi \subset B_{n_\pi}$. By the definition of n_π, there exists a vertex $\boldsymbol{v} \in \pi \cap B_{n_\pi} \setminus B_{n_\pi - 1}$. If $n_\pi = n_y$, we pick $\boldsymbol{v} = \boldsymbol{y}$. Otherwise, $n_\pi > n_y$ and we pick the first $\boldsymbol{v}$ in π with this property, and consequently,

$$|v - x| \geq (K_{n_\pi - 1} - L)/2 \geq (K_{n_y} - L)/2 \geq |y - x|/\sqrt{d}.$$

In both cases Proposition 3.3 applies and we find $\ell(\pi) \geq \eta |x - y|$, where $\eta = C_1/\sqrt{d}$. Note that, although we may not have centred the boxes in question at x, the box Λ_L is only a small box of constant size close to the origin. As $\boldsymbol{x} \in \Lambda_L$ and n is chosen large enough, the proof of Proposition 3.3 still applies. As a result, $\mathcal{D}^\eta_{L,M}(m)$ always occurs, if the boxes $B_{n-1}, B_n, \dots$ are all good. Consequently, $\mathcal{D}^\eta_L(m)$ can only fail to occur if at least one of these boxes is bad. Abbreviating $\alpha = \xi \vee (d + \mu)$, we derive for the latter event by Lemma 3.2,

$$\begin{aligned} \mathbb{P}\Big(\bigcup_{k \geq n-1} \{B_k \text{ is bad}\} \Big) &\leq \sum_{k=n}^{\infty} (n!)^\alpha \leq C \sum_{k=n}^{\infty} (\sqrt{k})^\alpha e^{\alpha k (\log k - 1)} \leq C (\sqrt{n})^\alpha \sum_{k=n}^{\infty} e^{\alpha k (\log n - 1)} \\ &\leq C \frac{(\sqrt{n})^\alpha}{\log n} e^{\alpha n (\log n - 1)} \leq C (n!)^\alpha, \end{aligned} \tag{9}$$

for large enough n, where we used Stirling's formula and denoted by C a constant that may have changed from line to line. By definition, we have $m < (n+1)!$ and thus $m/(n+1) \leq n!$. As the factorial grows faster than exponentially, we clearly have $n + 1 \leq \log m$, therefore $m/\log(m) \leq n!$. Plugging this back into (9), we obtain

$$\mathbb{P}(\neg \mathcal{D}^\eta_{L,M}(m)) \leq \mathbb{P}\Big(\bigcup_{k \geq n-1} \{B_k \text{ is bad}\} \Big) \leq C \log(m)^{-\alpha} m^\alpha.$$

Taking the logarithms and dividing both sides by $\log(m)$ concludes the proof of Theorem 1.2. □

Acknowledgement. I would like to thank Peter Gracar and Jonas Köppl for their valuable support and helpful discussions.

References

[1] Peter Antal and Agoston Pisztora. On the chemical distance for supercritical Bernoulli percolation. *Ann. Probab.*, 24(2):1036 – 1048, 1996.

[2] Noam Berger. A lower bound for the chemical distance in sparse long-range percolation models, 2004.

[3] Marek Biskup. On the scaling of the chemical distance in long-range percolation models. *Ann. Probab.*, 32(4):2938 – 2977, 2004.

[4] Johannes Bäumler. Continuity of the critical value for long-range percolation, 2023.

[5] Raphaël Cerf and Marie Théret. Weak shape theorem in first passage percolation with infinite passage times. *Ann. Inst. Henri Poincaré Probab. Stat.*, 52(3):1351–1381, 2016.

[6] Jiŗí Černý and Serguei Popov. On the internal distance in the interlacement set. *Electron. J. Probab.*, 17:1 – 25, 2012.

[7] Fan Chung and Linyuan Lu. *Complex Graphs and Networks*, volume 107 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 2006.

[8] Maria Deijfen, Remco van der Hofstad, and Gerard Hooghiemstra. Scale-free percolation. *Ann. Inst. Henri Poincaré Probab. Stat.*, 49(3):817–838, 2013.

[9] Alexander Drewitz, Balázs Ráth, and Artëm Sapozhnikov. On chemical distances and shape theorems in percolation models with long-range correlations. *J. Math. Phys.*, 55(8):083307, 2014.

[10] Olivier Garet and Régine Marchand. Large deviations for the chemical distance in supercritical Bernoulli percolation. *Ann. Probab.*, 35(3):833 – 866, 2007.

[11] Peter Gracar, Arne Grauer, and Peter Mörters. Chemical distance in geometric random graphs with long edges and scale-free degree distribution. *Comm. Math. Phys.*, 395(2):859–906, 2022.

[12] Peter Gracar, Markus Heydenreich, Christian Mönch, and Peter Mörters. Recurrence versus transience for weight-dependent random connection models. *Electron. J. Probab.*, 27:1 – 31, 2022.

[13] Peter Gracar, Marilyn Korfhage, and Peter Mörters. Robustness in the poisson boolean model with convex grains, 2024.

[14] Peter Gracar, Lukas Lüchtrath, and Christian Mönch. Finiteness of the percolation threshold for inhomogeneous long-range models in one dimension, 2022.

[15] Peter Gracar, Lukas Lüchtrath, and Peter Mörters. Percolation phase transition in weight-dependent random connection models. *Adv. in Appl. Probab.*, 53(4):1090–1114, 2021.

[16] Peter Hall. On continuum percolation. *Ann. Probab.*, 13(4):1250–1266, 1985.

[17] Marcelo Hilário and Daniel Ungaretti. Euclidean and chemical distances in ellipses percolation, 2021.

[18] Christian Hirsch. From heavy-tailed Boolean models to scale-free Gilbert graphs. *Braz. J. Probab. Stat.*, 31(1):111 – 143, 2017.

[19] Emmanuel Jacob, Benedikt Jahnel, and Lukas Lüchtrath. Subcritical annulus crossing in spatial random graphs, 2024.

[20] Benedikt Jahnel, Lukas Lüchtrath, and Marcel Ortgiese. Cluster sizes in subcritical soft boolean models, 2024.

[21] Joost Jorritsma, Júlia Komjáthy, and Dieter Mitsche. Cluster-size decay in supercritical kernel-based spatial random graphs, 2023.

[22] Joost Jorritsma, Júlia Komjáthy, and Dieter Mitsche. Large deviations of the giant in supercritical kernel-based spatial random graphs, 2024.

[23] Júlia Komjáthy, John Lapinskas, Johannes Lengler, and Ulysse Schaller. Polynomial growth in degree-dependent first passage percolation on spatial random graphs, 2024.

[24] Kostas Lakis, Johannes Lengler, Kalina Petrova, and Leon Schiller. Improved bounds for poly-logarithmic graph distances in scale-free percolation and related models, 2024.

[25] Günter Last and Mathew Penrose. *Lectures on the Poisson Process*. Cambridge University Press, 10 2017.

[26] Christian Mönch. Inhomogeneous long-range percolation in the weak decay regime. *Probab. Theory Relat. Fields*, 189(3):1129–1160, 2024.

[27] Balázs Ráth and Sándor Rokob. Percolation of worms. *Stochastic Process. Appl.*, 152:233–288, 2022.

[28] L. S. Schulman. Long range percolation in one dimension. *J. Phys. A*, 16(17):L639–L641, 1983.

[29] Augusto Teixeira and Daniel Ungaretti. Ellipses percolation. *J. Stat. Phys.*, 168(2):369–393, 2017.

[30] Chang-Long Yao, Ge Chen, and Tian-De Guo. Large deviations for the graph distance in supercritical continuum percolation. *J. Appl. Probab.*, 48(1):154 – 172, 2011.