The Annals of Applied Probability
2024, Vol. 34, No. 4, 3870–3910
https://doi.org/10.1214/24-AAP2052
© Institute of Mathematical Statistics, 2024

RATIO CONVERGENCE RATES FOR EUCLIDEAN FIRST-PASSAGE PERCOLATION: APPLICATIONS TO THE GRAPH INFINITY LAPLACIAN

BY LEON BUNGERT[1,a], JEFF CALDER[2,b] AND TIM ROITH[3,c]

[1]*Institute of Mathematics, University of Würzburg,* [a]*leon.bungert@uni-wuerzburg.de*
[2]*School of Mathematics, University of Minnesota,* [b]*jwcalder@umn.edu*
[3]*Helmholtz Imaging, Deutsches Elektronen-Synchrotron DESY,* [c]*tim.roith@desy.de*

In this paper we prove the first quantitative convergence rates for the graph infinity Laplace equation for length scales at the connectivity threshold. In the graph-based semisupervised learning community this equation is also known as Lipschitz learning. The graph infinity Laplace equation is characterized by the metric on the underlying space, and convergence rates follow from convergence rates for graph distances. At the connectivity threshold, this problem is related to Euclidean first passage percolation, which is concerned with the Euclidean distance function $d_h(x, y)$ on a homogeneous Poisson point process on $\mathbb{R}^d$, where admissible paths have step size at most $h > 0$. Using a suitable regularization of the distance function and subadditivity we prove that $d_{h_s}(0, se_1)/s \to \sigma$ as $s \to \infty$ almost surely where $\sigma \geq 1$ is a dimensional constant and $h_s \gtrsim \log(s)^{1/d}$. A convergence rate is not available due to a lack of approximate superadditivity when $h_s \to \infty$. Instead, we prove convergence rates for the *ratio* $\frac{d_h(0,se_1)}{d_h(0,2se_1)} \to \frac{1}{2}$ when h is frozen and does not depend on s. Combining this with the techniques that we developed in (*IMA J. Numer. Anal.* **43** (2023) 2445–2495), we show that this notion of ratio convergence is sufficient to establish uniform convergence rates for solutions of the graph infinity Laplace equation at percolation length scales.

1. Introduction. In this paper we will use techniques from first-passage percolation to prove new and strong results in the field of partial differential equations on graphs. In more detail, we will exploit stochastic homogenization effects in Euclidean first-passage percolation on Poisson point clouds to derive uniform convergence rates for the infinity Laplacian equation on a random geometric graph with n vertices in $\mathbb{R}^d$ whose connectivity length scale ε_n is proportional to the connectivity threshold, that is,

$$\varepsilon_n \sim \left(\frac{\log n}{n}\right)^{\frac{1}{d}}.$$

Our approach is based on the insight from our previous work [15] that convergence rates for the graph distance function translate to convergence rates for solutions of the graph infinity Laplace equation which can be regarded as a generalized finite difference method and which, in the context of semisupervised learning, is also known as Lipschitz learning.

While the fields of percolation theory and partial differential equations (PDEs) on graphs (including finite difference methods and semisupervised learning) are very well developed, there are relatively few results that connect them, such as [13, 27] which deals with Gamma-convergence of discrete Dirichlet energies on Poisson clouds or [41] on distance learning from a Poisson cloud on an unknown manifold. In the following we give a brief overview of first-passage percolation and graph PDEs.

Received October 2022; revised January 2024.
MSC2020 subject classifications. Primary 35R02, 65N12, 60K35; secondary 60F10, 60G44, 68T05.
Key words and phrases. First-passage percolation, Poisson point process, concentration of measure, graph infinity Laplacian, Lipschitz learning, graph-based semisupervised learning.

First-passage percolation. First-passage percolation was introduced in [14, 43] as a model for the propagation of fluid through a random medium. In mathematical terms, the set-up is a graph $G = (V, E)$ whose edges are equipped with passage times $t(e) \in [0, \infty]$ and one would like to understand the graph distance function between vertices $x, y \in V$:

$$T(x, y) := \inf\left\{\sum_{i=1}^{m} t(e_i) : m \in \mathbb{N}, (e_1, \dots, e_m) \text{ connects } x \text{ and } y\right\}. \tag{1.1}$$

Typical questions address properties of geodesics, shape theorems, size of connected components of the graph, and large scale asymptotics of the graph distance.

Stochasticity can enter the model in different ways. In the simplest set-up the graph consists of the square lattice $\mathbb{Z}^d$ and the passage times $t(e)$ are *i.i.d.* random variables. This setting is well-understood (see the incomplete list of results [1, 2, 30, 49, 50] and the surveys [10, 48, 62]). Another way that randomness can enter the percolation model is through the vertices of the graph instead of its edge weights. This setting is known as Euclidean first-passage percolation and typically the vertices are assumed to constitute a Poisson point process X in $\mathbb{R}^d$, which possesses convenient isotropy properties [51]. The connectivity of the graph can be modelled in different ways but is typically assumed to follow deterministic rules (once the vertices are given).

In the works [45, 46] a fully connected graph together with power weighted passage times is considered, that is, $t(e) = |x - y|^\alpha$ where $e = (x, y)$ represents an edge in the graph and $\alpha \geq 1$ is a parameter. For $\alpha = 1$ long hops are possible and the corresponding graph distance $T(x, y)$ equal the Euclidean one $|x - y|$. To prevent this trivial behavior and enforce short hops, in almost all results it is assumed that $\alpha > 1$. More recent results and applications of this power weighted Euclidean first-passage percolation model can be found, for instance, in [47, 53]. It is also possible to replace the fully connected graph by a Delaunay triangulation subordinate to the Poisson point process, see, for example, [44, 57, 59].

Most relevant for us will be the setting of a random geometric graph. Here, the connectivity relies on some parameter $h > 0$ and admissible paths in the definition of the distance $T(x, y)$ cannot have hops of length larger than h. Such models were previously considered but much less is known, as compared to lattice percolation or Euclidean percolation with power weights. High probability bounds between the graph and Euclidean distance were proved in [35, 39] and large deviation results for the graph distance and a shape theorem were established in [63]. The central difficulty of this model is that the distance function is a random variable with infinite expectation with respect to the realizations of the Poisson point process. This makes standard techniques from subadditive ergodic theory inapplicable. Furthermore, establishing quantitative large deviation bounds for this graph distance is very challenging due to the fact that feasible paths on different scales h cannot be straightforwardly combined into a feasible path. In essence, this means that the stochastic processes, while still subadditive, do not readily admit any type of approximate superadditivity *across length scales*, which is needed to establish convergence rates.[1] This issue does not arise in lattice percolation [50], power weighted percolation [46], or related problems like the longest chain problem [12], since in these cases the connectivity structure does not involve a length scale h, and so approximate superadditivity is readily available. For additional convergence rate results in lattice percolation we also refer to [1, 2].

Let us mention that there is a history of ideas from percolation theory (e.g., subadditivity and concentration inequalities) finding important applications in the theory of PDEs. Recent

[1] As we show in this paper, approximate superadditivity does hold when the length scale h is fixed, which is sufficient for the ratio convergence results in this paper, but not for establishing convergence rates for the scaling limit.

results on stochastic homogenization theory for PDEs make use of subadditive quantities [3, 5, 6, 8], including homogenization of elliptic PDEs on percolation clusters [4, 32]. Subadditivity and concentration inequalities are also key tools in the convergence of data peeling processes to solutions of continuum PDEs [26, 28].

Graph PDEs, finite difference methods, and semisupervised learning. Recent years have seen a surge of interest and results in the field of PDEs and variational problems on graphs. This is based on the observations that, on one hand, PDEs on graphs generalize finite difference methods for the numerical solution of PDEs and, on the other hand, constitute efficient and mathematically well-understood tools for solving problems in machine learning, including data clustering, semisupervised learning, and regression problems, to name a few.

The first observation is easily understood, noting that any grid in $\mathbb{R}^d$ with neighbor relations—for instance, the rectangular regular grid $\varepsilon\mathbb{Z}^d$ where every point $x_0 \in \varepsilon\mathbb{Z}^d$ is connected to its $2d$ nearest neighbors $x_0 \pm \varepsilon e_i$ for $i = 1, \dots, d$ and their connection is weighted by their Euclidean distance ε—is a special case of a weighted graph. The Laplacian operator of a smooth function, for instance, can be approximated as

$$\Delta u(x_0) \approx \frac{1}{\varepsilon^2}\sum_{i=1}^{d}\big(u(x_0+\varepsilon e_i) - 2u(x_0) + u(x_0 - \varepsilon e_i)\big). \tag{1.2}$$

It is important to remark that graphs allow for richer models. For instance, if the points $\{x_1, \dots, x_n\}$ are *i.i.d.* samples from a probability density ρ, then the graph Laplacian offers an approximation of a density weighted Laplacian with high probability (see e.g, [17] and the references therein):

$$\frac{1}{\rho(x_0)}\operatorname{div}\big(\rho(x_0)^2\nabla u(x_0)\big) \approx \frac{1}{n\varepsilon^{d+2}}\sum_{\substack{1\le i\le n\\ |x_i - x_0|\le\varepsilon}}\big(u(x_i) - u(x_0)\big). \tag{1.3}$$

Furthermore, as opposed to standard finite difference methods, random graphs can possess an increased approximation and convergence behavior due to stochastic homogenization effects. In the context of the *infinity* Laplace operator, this is a key finding of the present paper.

The convergence analysis of finite difference methods for nonlinear PDEs like the p-Laplace and the infinity Laplace equations was revolutionized by the seminal work of Barles and Souganidis [11] on convergence of monotone schemes to viscosity solutions and sparked results like [54–56]. Furthermore, the dynamic programming principles and mean value formulas gave rise to new finite difference methods for p-Laplace equations [33, 34].

There are also close connections between graph PDEs and semisupervised learning (SSL). In SSL one is typically confronted with a relatively large collection of $n \in \mathbb{N}$ data points Ω_n, only few of which carry a label. The points with labels constitute the small subset $\mathcal{O}_n \subset \Omega_n$ (which can but does not have to depend on n). A prototypical example for this is the field of medical imaging where obtaining data is cheap but obtaining labels is expensive. Based on pairwise similarity or proximity of the data points, the whole data set is then turned into a weighted graph structure and one seeks to extend the label information by solving a "boundary" value problem on this graph, where the boundary data is given by the labels on the small labeled set. The abstract problem consists of finding a function $u_n : \Omega_n \to \mathbb{R}$ that solves the graph PDE

$$\begin{cases} \mathcal{L}_n u_n(x) = 0 & \text{for all } x \in X_n \setminus \mathcal{O}_n, \\ u_n(x) = g(x) & \text{for all } x \in \mathcal{O}_n, \end{cases}$$

where $\mathcal{L}_n$ is a suitable differential operator on a graph, for example, a version of the graph Laplacian [24, 25, 64], the graph p-Laplacian for $p \in (1, \infty)$ [17, 40, 42, 60], the graph infinity Laplacian [15, 18, 58], a Poisson operator [20, 23], or an eikonal-type operator [21, 36, 38].

Both in the context of finite difference methods and in graph-based semisupervised learning two main questions arise:

1. Under which conditions on the graph and the discrete operators do solutions converge to solutions of the respective continuum PDE?
2. What is the rate of convergence?

The answers to these questions, if they exist, typically involve two important parameters: The graph resolution δ_n, which describes how well the graph approximates the continuum domain in the Hausdorff distance, and the graph length scale ε_n, which encodes the maximum distance between neighbors in the graph. Note that for n *i.i.d.* samples from a positive distribution $\delta_n \sim (\log n/n)^{\frac{1}{d}}$ with high probability whereas for a regular grid $\delta_n \sim (1/n)^{\frac{1}{d}}$. The finite difference approximation of the Laplacian on a regular grid (1.2) where $\varepsilon_n \sim \delta_n$ is consistent with the Laplacian, where with consistency we mean that the application of the discrete operator to a smooth function converges to the application of the limiting operator to the same function. However, already for the graph Laplacian (1.3) on general point clouds or for nonlinear differential operators like the game theoretic p-Laplacian or the infinity Laplacian, one has to choose ε_n significantly larger than δ_n to ensure that the discrete operators are consistent with the continuum one, for example, $\varepsilon_n \gg \delta_n^{\frac{d}{d+2}}$ for the Laplacian [17], Theorem 5, $\varepsilon_n \gg \delta_n^{\frac{2}{3}}$ for the p-Laplacian [33], Theorem 1.1, and [18], Lemma 15, Theorem 17, for the infinity Laplacian.

Note that convergence rates can be proved for solutions of the graph Laplace equation by combing consistency with maximum principles, see, for example, [25], and spectral convergence rates for eigenfunctions are also available, see [22] and the references therein. Furthermore, in the consistent regime of the infinity Laplacian, rates of convergence were proved for $\varepsilon_n \gg \delta_n^{\frac{1}{2}}$ and a very restrictive setting in [61] and, recently, for general unstructured grids but very large length scales $\varepsilon_n \sim \delta_n^{\frac{1}{4}}$ in [52]. In [15] we established convergence rates in a general setting whenever $\varepsilon_n \gg \delta_n$.

As our result in [15] shows, overcoming the lower bounds imposed by consistency of the operator is clearly possible is some cases. For instance, when working with variational methods like Gamma-convergence, convergence can typically be established in the regime $\varepsilon_n \gg \delta_n$ [40, 58, 60], however, proving convergence rates is difficult due to the asymptotic nature of Gamma-convergence.

In our previous work [15] we proposed an entirely new approach based on ideas from homogenization theory. We defined a new homogenized length scale τ_n that is significantly larger then the graph length scale ε_n, that is, one has $\delta_n \ll \varepsilon_n \ll \tau_n$. We showed that solutions of the graph infinity Laplace equation

$$\begin{aligned} &\max_{1\leq j\leq n} \eta\big(|x_i - x_j|/\varepsilon_n\big)\big(u(x_j) - u(x_i)\big) \\ &\quad + \min_{1\leq j\leq n} \eta\big(|x_i - x_j|/\varepsilon_n\big)\big(u(x_j) - u(x_i)\big) = 0, \end{aligned} \tag{1.4}$$

where $\eta : (0, \infty) \to (0, \infty)$ is a decreasing function satisfying $\operatorname{supp} \eta \subset [0, 1]$ and some other mild conditions, give rise to approximate sub- and super-solutions of a nonlocal homogenized infinity Laplace equation for the operator

$$\Delta_\infty^{\tau_n} u(x) := \frac{1}{\tau_n^2}\Big(\sup_{y\in B(x;\tau_n)} \big(u(y) - u(x)\big) + \inf_{y\in B(x;\tau_n)} \big(u(y) - u(x)\big)\Big). \tag{1.5}$$

Loosely speaking the larger length scale τ_n can then be used to ensure consistency with the infinity Laplacian $\Delta_\infty u := \langle \nabla u, \nabla^2 \nabla u\rangle$ while at the same time allowing ε_n to arbitrarily close to δ_n as long as $\varepsilon_n \gg \delta_n$ is satisfied. The rate is then given by the optimal choice of τ_n in terms of ε_n and δ_n. The convergence rates obtained in our previous work [15] depend on quantities like the ratio $\frac{\delta_n}{\varepsilon_n}$, and are degenerate at the connectivity scaling $\varepsilon_n \sim \delta_n$. Establishing convergence rates at the length scale $\varepsilon_n \sim \delta_n$ is the main focus of this paper.

Structure of this paper. The rest of the paper is organized as follows: In Section 2 we explain our precise setup and our main results for Euclidean first-passage percolation Theorem 2.1 and the graph infinity Laplacian Theorem 2.3. We also discuss some open problems and extensions. Sections 3 to 5 are devoted to proving the percolation results, by first establishing asymptotics for the expected value of a regularized graph distance (which has finite expectation and coincides with the original distance with high probability), proving concentration of measure for this distance, and establishing quantitative convergence rates for the ratio of two distance functions. Finally, we apply our findings to get convergence rates for the graph infinity Laplace equation in Section 6.

2. Setup and main results.

In this section we introduce the different distance functions on Poisson point processes that we shall use in the course of the paper.

In large parts of this paper we let X be a Poisson point process on $\mathbb{R}^d$ with unit intensity. This means that X is a random at most countable collection of points such that the number of points in $X \cap A$, for a Borel set A, is a Poisson random variable with mean $|A|$, which denotes the Lebesgue measure of A. That is,

$$\mathbb{P}\big(\#(A \cap X) = k\big) = \frac{|A|^k}{k!}\exp\big(-|A|\big). \tag{2.1}$$

The Poisson process has the important property that for any $A \subset \mathbb{R}^d$, the intersection $X \cap A$ is also a Poisson point process with intensity function 1_A, or rather, a unit intensity Poisson point process on A [51]. This is not true for *i.i.d.* sequences, restrictions of which to subsets are, in fact, Binomial point processes.

2.1. *Paths and distances.*

Given a set of points $P \subset \mathbb{R}^d$, $x, y \in \mathbb{R}^d$, and a length scale $h > 0$, we denote the set of paths in P, connecting $x, y \in \mathbb{R}^d$ with steps of size less than or equal to h, by

$$\begin{aligned} \Pi_{h,P}(x,y) := \big\{ p \in P^m \ : m \in \mathbb{N}, p_1 \in \pi_P(x), p_m \in \pi_P(y), \\ |x - p_1| \le h/2, |y - p_m| \le h/2, \text{and} \\ |p_i - p_{i+1}| \le h \ \forall i = 1, \dots, m \big\}, \end{aligned} \tag{2.2}$$

where for $x \in \mathbb{R}^d$ the set $\pi_P(x) := \{\hat{x} \in P : |x - \hat{x}| = \min_{y \in P} |x - y|\}$ is the set of all closest points in P. For any path $p \in \Pi_{h,P}(x,y)$ with m elements we let

$$L(p) := \sum_{i=1}^{m-1} |p_{i+1} - p_i| \tag{2.3}$$

denote the length of the path. We now define

$$d_{h,P}(x,y) := \inf\big\{L(p) : p \in \Pi_{h,P}(x,y)\big\}, \quad x, y \in P, \tag{2.4}$$

to be the length of the shortest path in $\Pi_{h,P}(x,y)$ connecting $x \in \mathbb{R}^d$ and $y \in \mathbb{R}^d$. Whenever we are referring to the Poisson point process X, meaning $P = X$, we obfuscate the dependency on X by using the abbreviations $\Pi_h(x,y)$ and $d_h(x,y)$.

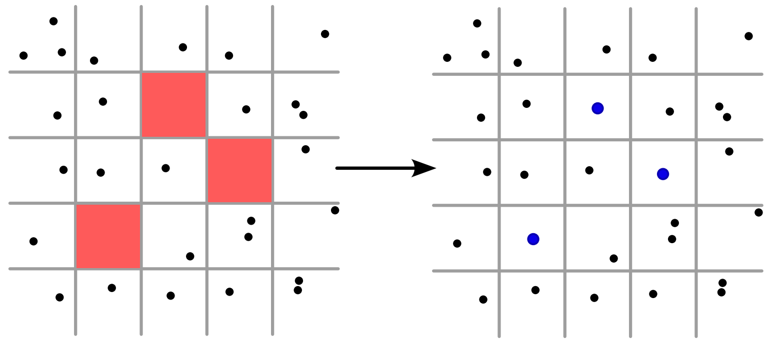

FIG. 1. *The covering of the Poisson process X on the left yields empty boxes in red. We define the enriched process $\mathcal{X}_s$ on the right by adding the points in blue.*

2.2. *Different distance-based random variables.* Thanks to the spatial homogeneity of the Poisson process it suffices to study the distance between the points $x = 0$ and $y = se_1$ where $e_1 = (1, 0, \dots, 0) \in \mathbb{R}^d$ denotes the first unit vector. This leads to the quantity $d_h(0, se_1)$ for $s \geq 0$. If the length scale $h > 0$ is fixed, most distances will be infinite with high probability when s is large. Therefore, we consider length scales $h \equiv h_s$ which depend on the distance.

Our main object of study is the random variable

$$T_s := d_{h_s}(0, se_1) = \inf\{L(p) : p \in \Pi_{h_s}(0, se_1)\}, \quad s \geq 0. \tag{2.5}$$

For properly chosen length scales h_s, roughly satisfying $\log(s)^{\frac{1}{d}} \lesssim h_s \ll s$, we show that $s - h_s \leq T_s \leq C_d s$ with high probability. Here $C_d > 0$ is a suitable dimensional constant, to be specified later. However, with small but positive probability there are no feasible paths and T_s is infinite which makes it meaningless to study its expectation $\mathbb{E}[T_s]$ and fluctuations around the expectation.

Therefore, we construct yet another distance function which always has feasible paths. For this, let us fix $s > 0$ and cover $\mathbb{R}^d$ with closed boxes $\{B_k\}_{k \in \mathbb{N}}$ of side length δ_s / C_d. Here $\delta_s \sim \log(s)^{\frac{1}{d}}$ will be specified later and $C_d > 0$ is a dimensional constant, sufficiently large such that the maximum distance of two points in two touching boxes is at most $\delta_s/2$. A possible choice is

$$C_d := 2\sqrt{d}. \tag{2.6}$$

The probability that all boxes B_k contain at least one point in X is zero and therefore we define an at most countable index set $\mathcal{I}_s$ such that $B_i \cap X = \varnothing$ for all $i \in \mathcal{I}_s$. For every $i \in \mathcal{I}_s$ we then add a point $x_i \in B_i$, for instance the center of the box, to the Poisson process X, see Figure 1, which leads to the enriched set of points

$$\mathcal{X}_s := X \cup \bigcup_{i \in \mathcal{I}_s} \{x_i\}. \tag{2.7}$$

With the notation $\mathcal{X}_s$ we emphasize that this enriched Poisson point process depends on s. We can hence consider the following distance function on the enriched Poisson process which we define for all scalings $h \geq \delta_s$ (which can but do not have to depend on s)

$$d_{h,\mathcal{X}_s}(x, y), \quad x, y \in \mathbb{R}^d, \tag{2.8}$$

and we define T'_s as

$$T'_s := d_{h_s,\mathcal{X}_s}(0, se_1). \tag{2.9}$$

TABLE 1
Different random variables used in this work. X denotes a unit intensity Poisson process, $\mathcal{X}_s$ an enrichment with additional points

Symbol	Meaning	Definition
T_s	graph distance on Poisson process	$d_{h_s,X}(0,se_1)$
T_s'	graph distance on enriched Poisson process	$d_{h_s,\mathcal{X}_s}(0,se_1)$
–	graph distance on enriched Poisson process with fixed step size	$d_{h,\mathcal{X}_s}(x,y)$

Later we shall express $T_s' - \mathbb{E}[T_s']$ as a sum over a *martingale difference sequence* with bounded increments. This will allow us to prove concentration of measure for T_s'.

Finally, we will synthesis these different results (see Table 1 for an overview of the different definitions) by utilizing that with high probability the random variables T_s and T_s' coincide.

2.3. *Constants and symbols*. We will encounter many constants in the paper, most of which are dimensional, that is, they depend on $d \in \mathbb{N}$. Out of all these constants, only C_d (which was already introduced above), C_d' (which shall be introduced in Section 3.2), and σ (which will arise as $\sigma = \lim_{s\to\infty} \frac{T_s}{s} \in [1, C_d]$) will keep their meaning throughout the whole paper. However, we do not claim that their values are optimal or analytically known. In many estimates and probabilities other (mostly dimensional) constants will appear and we number them as C_1, C_2, C_3, etc. Note, however, that their values change between the individual lemmas and theorems they appear in and also sometimes change in proofs, which we mention in the latter case. Since these constants are of no importance to us, we refrain from numbering them continuously, as its sometimes done. Finally, we sometimes write our inequalities in a more compact form by absorbing all constants into the symbols $\lesssim$, $\gtrsim$, or $\sim$, where, for instance, $f(s) \lesssim g(s)$ means $f(s) \le Cg(s)$ and $f(s) \sim g(s)$ means $f(s) = Cg(s)$ for a constant $C > 0$. Finally, we will use the symbol $\ll$ to denote

$$f(s) \ll g(s) \quad \iff \quad \lim_{s\to\infty} \frac{f(s)}{g(s)} = 0.$$

2.4. *Main results*. In this section we state our most important results for Euclidean first-passage percolation on a unit intensity Poisson process $X \subset \mathbb{R}^d$ and the application to the graph infinity Laplace equation.

Note that the first theorem is stated in a very compact form and some of the results will be proved in a slightly more general setting. Also the assumptions will be spelled out in a more quantitative form throughout the paper. The most important parts of this statement are the concentration of measure and the convergence rates for ratio convergence. Even though convergence rates for $d_{h_s}(0, se_1)$ are not available, the ratio convergence rates are sufficient for our application to Lipschitz learning.

THEOREM 2.1 (Euclidean first-passage percolation). *Let $s > 1$ and assume that $s \mapsto h_s$ is nondecreasing and satisfies*

$$\log(s)^{\frac{1}{d}} \lesssim h_s \ll s.$$

There exist dimensional constants $C_1, C_2 > 0$, not depending on s, such that:

1. (Convergence) *There exists a dimensional constant $\sigma \in [1, C_d]$ (depending on the choice of $s \mapsto h_s$) such that*

$$\lim_{s\to\infty} \frac{\mathbb{E}[d_{h_s,\mathcal{X}_s}(0, se_1)]}{s} = \sigma \quad \textit{and} \quad \lim_{s\to\infty} \frac{d_{h_s}(0, se_1)}{s} = \sigma \quad \textit{almost surely.}$$

2. (Concentration) *It holds for all* $t \geq h_s$

$$\mathbb{P}\left(\left|d_{h_s,\mathcal{X}_s}(0,te_1) - \mathbb{E}[d_{h_s,\mathcal{X}_s}(0,te_1)]\right| > \lambda\sqrt{\frac{\log(s)^{\frac{2}{d}}}{h_s}t}\right) \leq C_1 \exp(-C_2\lambda) \quad \forall \lambda \geq 0.$$

3. (Ratio convergence) *It holds for* $s > 1$ *sufficiently large*

$$\left|\frac{\mathbb{E}[d_{h_s,\mathcal{X}_s}(0,se_1)]}{\mathbb{E}[d_{h_s,\mathcal{X}_s}(0,2se_1)]} - \frac{1}{2}\right| \leq C_1\frac{h_s}{s} + C_2\sqrt{\frac{\log(s)^{\frac{2}{d}}}{h_s}}\frac{\log(s)}{\sqrt{s}}.$$

PROOF. The theorem collects results from Theorems 4.1 and 4.4 and Proposition 5.4. □

REMARK 2.2. Since $d_{h,\mathcal{X}_s}(x,y) = d_h(x,y)$ with high probability, the concentration of measure statement in Item 2 of Theorem 2.1 implies concentration of the standard distance function T_s around $\mathbb{E}[T_s']$. Furthermore, using concentration, Item 3 has a corresponding high probability versions for both distances.

Our second main result concerns convergence rates for solutions to the graph infinity Laplace equation. For this we let $X_n \subset \overline{\Omega}$ be a Poisson point process with density $n \in \mathbb{N}$ in an open and bounded domain $\Omega \subset \mathbb{R}^d$. For a bandwidth parameter $\varepsilon > 0$ and a function $u : X_n \to \mathbb{R}$ we define the graph infinity Laplacian of u as

$$\mathcal{L}_\infty^\varepsilon u(x) := \sup_{y\in B(x,\varepsilon)\cap X_n} \frac{u(y)-u(x)}{|y-x|} + \inf_{y\in B(x,\varepsilon)\cap X_n} \frac{u(y)-u(x)}{|y-x|}, \quad x \in X_n.$$

The infinity Laplacian operator of a smooth function $u : \Omega \to \mathbb{R}$ is defined as

$$\Delta_\infty u = \sum_{i,j=1}^d \partial_i u \partial_j u \partial_{ij}^2 u = \langle \nabla u, \nabla^2 u \nabla u\rangle.$$

The following theorem states quantitative high probability convergence rates of solutions to the equation $\mathcal{L}_\infty^\varepsilon u_n = 0$ to solutions of $\Delta_\infty u = 0$. Note that the theorem considers the boundary value problem associated with the infinity Laplace operator whereas in our previous work [15] we considered the setting where function values are prescribed in a very general closed set $\mathcal{O} \subset \overline{\Omega}$. While this is much more realistic in the context of semisupervised learning, the corresponding convergence proof requires precise control of graph distance functions close to the boundary of the domain. Achieving this control in the percolation setting is far beyond the scope of this paper since it would essentially require percolation results on Poisson point processes on half spaces together with suitable flattening techniques. Therefore, we focus on the setting of a boundary value problem, where boundary values for the discrete equation are prescribed in a tube around the boundary. This is in line with previous work for the linear Laplacian operator, for example, [13, 25].

THEOREM 2.3 (Convergence rates). *Let* $\Omega \subset \mathbb{R}^d$ *be an open and bounded domain. Let* $g : \overline{\Omega} \to \mathbb{R}$ *be a Lipschitz function and* $u : \Omega \to \mathbb{R}$ *be the unique viscosity solution of*

$$\begin{cases} \Delta_\infty u = 0 & \text{in } \Omega, \\ u = g & \text{on } \partial\Omega. \end{cases}$$

Let X_n *be a Poisson point process in* $\mathbb{R}^d$ *with density* $n \in \mathbb{N}$, *let* $\varepsilon > 0$ *and* $\tau > 0$ *satisfy*

$$K\left(\frac{\log n}{n}\right)^{\frac{1}{d}} \leq \varepsilon \leq \frac{1}{K}\tau, \quad 0 < \tau < 1,$$

and let

$$\mathcal{O}_n := \{x \in X_n \cap \overline{\Omega} : \operatorname{dist}(x, \partial\Omega) \le \varepsilon\}.$$

Let $u_n : X_n \to \mathbb{R}$ *be the unique solution of*

$$\begin{cases} \mathcal{L}_\infty^\varepsilon u_n = 0 & \text{in } \Omega \cap X_n \setminus \mathcal{O}_n, \\ u_n = g & \text{on } \mathcal{O}_n. \end{cases}$$

There exist dimensional constants $C_1, C_2, C_3, C_4, C_5 > 0$ *such that for* $n \in \mathbb{N}$, *for all* $\lambda \ge 0$, *and for* $K \ge 8$ *sufficiently large it holds*

$$\mathbb{P}\left(\sup_{x \in X_n} |u(x) - u_n(x)| \lesssim \tau + \sqrt[3]{(\log n + \lambda)\left(\frac{\log n}{n}\right)^{\frac{1}{d}} \frac{1}{\sqrt{\tau^3 \varepsilon}} + \frac{\varepsilon}{\tau^2}}\right)$$
$$\ge 1 - C_1 \exp(-C_2 K^d \log n) - C_3 \exp(-C_4 \lambda + C_5 \log n).$$

An important special case of Theorem 2.3 is the choice of $\varepsilon_n \sim (\frac{\log n}{n})^{\frac{1}{d}}$.

COROLLARY 2.4. *Under the conditions of Theorem 2.3 and for* $\varepsilon = \varepsilon_n = K(\frac{\log n}{n})^{\frac{1}{d}}$ *with* K *sufficiently large it holds for all* $\lambda \ge 0$ *that*

$$\mathbb{P}\left(\sup_{x \in X_n} |u(x) - u_n(x)| \lesssim (\log n + \lambda)^{\frac{2}{9}} \left(\frac{\log n}{n}\right)^{\frac{1}{9d}}\right)$$
$$\ge 1 - C_1 \exp(-C_2 K^d \log n) - C_3 \exp(-C_4 \lambda + C_5 \log n).$$

PROOF. For this choice of $\varepsilon = \varepsilon_n$ it holds that

$$\left(\frac{\log n}{n}\right)^{\frac{1}{d}} \frac{1}{\sqrt{\tau^3 \varepsilon}} \gtrsim \frac{\varepsilon}{\tau^2}$$

so we can ignore the second term under the root in Theorem 2.3. Optimizing the resulting error term over τ yields the optimal choice of $\tau_n := (\log n + \lambda)^{\frac{2}{9}} (\frac{\log n}{n})^{\frac{1}{9d}}$. For this choice both terms scale in the same way. □

REMARK 2.5. Corollary 2.4 shows that we get a convergence rate of $(\frac{\log n}{n})^{\frac{1}{9d}}$ (up to the log factor) at the connectivity scale $\varepsilon_n \sim (\frac{\log n}{n})^{\frac{1}{d}}$. Interestingly, this rate coincides with the best rate achievable using the techniques from our previous paper [15], though in that work we had to choose a much larger length scale $\varepsilon_n \sim (\frac{\log n}{n})^{\frac{5}{9d}}$ to obtain the rate. In any case, judging from our numerical experiments and simple examples we do not expect these rates to be optimal. In particular, it would be interesting to understand the degree of suboptimality which our techniques introduce when passing from rates of distance functions (or their ratio) to rates for the infinity Laplace equation.

We can obtain almost sure convergence rates by letting λ depend on n.

COROLLARY 2.6. *Under the conditions of Corollary 2.4 and for* $K > 0$ *sufficiently large it holds*

$$\limsup_{n \to \infty} \frac{\sup_{x \in X_n} |u(x) - u_n(x)|}{(\log n)^{\frac{2}{9}} (\frac{\log n}{n})^{\frac{1}{9d}}} < \infty \quad \textit{almost surely}.$$

PROOF. For $\lambda_n = C \log n$ with a large constant $C > 0$ and for $K > 0$ sufficiently large we can use the Borel–Cantelli lemma to conclude. □

While we have stated our results for Poisson point processes, it is straightforward to de-Poissonize and obtain the same results for *i.i.d.* sequences.

COROLLARY 2.7. *Assume the conditions of Theorem 2.3, except that X_n is defined instead as an* i.i.d. *sample of size n uniformly distributed on Ω. There exist dimensional constants $C_1, C_2, C_3, C_4, C_5 > 0$ such that for $n \in \mathbb{N}$ and $K > 0$ sufficiently large it holds*

$$\mathbb{P}\left(\sup_{x \in X_n} |u(x) - u_n(x)| \lesssim \tau + \sqrt[3]{(\log n + \lambda)\left(\frac{\log n}{n}\right)^{\frac{1}{d}} \frac{1}{\sqrt{\tau^3 \varepsilon}} + \frac{\varepsilon}{\tau^2}}\right)$$
$$\geq 1 - e^{\frac{1}{12}} C_1 \exp\left(-C_2 K^d \log n + \frac{1}{2}\log(n)\right) - C_3 \exp(-C_4 \lambda + C_5 \log n).$$

PROOF. Let $\widetilde{X}_n$ be a Poisson point process on $\mathbb{R}^d$ with intensity $\frac{n}{|\Omega|}$. Conditioned on $\#(\widetilde{X}_n \cap \Omega) = n$, both X_n and $\widetilde{X}_n \cap \Omega$ have the same distribution. By conditioning on $\#(\widetilde{X}_n \cap \Omega) = n$ and using Theorem 2.3 the probability of the event

$$\sup_{x \in X_n} |u(x) - u_n(x)| \gtrsim \tau + \sqrt[3]{(\log n + \lambda)\left(\frac{\log n}{n}\right)^{\frac{1}{d}} \frac{1}{\sqrt{\tau^3 \varepsilon}} + \frac{\varepsilon}{\tau^2}}$$

is bounded by

$$\mathbb{P}\big(\#(\widetilde{X}_n \cap \Omega) = n\big)^{-1}\big(C_1 \exp(-C_2 K^d \log n) + C_3 \exp(-C_4 \lambda + C_5 \log n)\big).$$

By Stirling's formula we have

$$\mathbb{P}\big(\#(\widetilde{X}_n \cap \Omega) = n\big)^{-1} = \frac{n! e^n}{n^n} \leq e^{\frac{1}{12}} \sqrt{n}.$$

Upon adjusting the values of C_3 and C_5, the proof is complete. □

2.5. *Outlook.* There are two central directions of future research that originate from this paper, namely further strengthening and generalizing our percolation results, and applying the techniques from this paper to prove convergence rates for other graph PDEs, like for instance the p-Laplace equation. With respect to the first direction, the ultimate goal would be to prove a strong approximate superadditivity result of the form (5.1) which in combination with the concentration of measure from Theorem 4.1 immediately yields convergence rates for the almost sure convergences $T_s'/s \to \sigma$ and $T_s/s \to \sigma$, as shown in [50]. Therefore, we formulate the following open problem:

OPEN PROBLEM. Does there exist a function $s \mapsto g(s)$, satisfying $\int_1^\infty g(s) s^{-2}\, \mathrm{d}s < \infty$, such that

$$\mathbb{E}[T_{2s}'] \geq 2\mathbb{E}[T_s'] - g(s), \quad s > 1?$$

This form of strong super-additivity is implied and roughly equivalent to establishing a modulus of continuity of the distance function with respect to the length scale, that is, for the function

$$h \mapsto \mathbb{E}[d_{h,\mathcal{X}_s}(0, s e_1)].$$

This problem is related to continuity of the time constant in first passage percolation, which was established for lattice percolation in [29, 31]. However, the notion of continuity in [29, 31] is nonquantitative, and taken with respect to the distribution of the *i.i.d.* edge weights, whereas in our setting we seek a *quantitative* continuity statement with respect to the *length scale h* that defines the connectivity structure. It seems that different techniques are required here.

Having this continuity at hand, it would be straightforward to extend the arguments of Section 6 to *inhomogeneous* Poisson point processes with intensity $n\rho$ where $n \in \mathbb{N}$ and ρ is a probability density with some regularity. Blowing up around a point shows that the graph distance can be bounded from above and from below with distances $d'_{h_i}(0, se_1)$ on a unit intensity process, albeit with two different but close length scales $h_1, h_2 > 0$.

It would be desirable to extend the percolation results to weighted distances of the form

$$d_h(x, y) := \inf\left\{\sum_{i=1}^{m} \frac{h}{\eta(|p_i - p_{i-1}|/h)} : p \in \Pi_h(x, y)\right\}. \tag{2.10}$$

For $\eta(t) := \frac{1}{t}$ this reduces to the distance that we considered here but it allows to generate a large class of commonly known graph distances where the weight of an edge (x, y) is given by $h^{-1}\eta(|x - y|/h)$. Most notably, if $\eta(t) = 1$ for $0 \le t \le 1$ and $\eta(t) = 0$ for $t > 1$ one obtains the hop counting distance, scaled with h. The analysis of (2.10) is complicated by the fact that they do not obey the triangle inequality and, furthermore, are inaccurate if $|x - y| \ll h$. Still, we expect that our results can be generalized to these distances relatively easily.

The question of whether and how percolation techniques can be applied to other graph PDEs (e.g., the Laplace or p-Laplace equations) seems much harder. Recent results in two dimensions show that at least Dirichlet energies Gamma-converge for percolation length scales [13, 27]. Combining quantitative versions of these arguments with the techniques from [19], Section 5.5, can potentially produce convergence rates.

3. Convergence in expectation.

In this section we prove that T'_s satisfies

$$\lim_{s\to\infty} \frac{\mathbb{E}[T'_s]}{s} = \sigma \in (0, \infty). \tag{3.1}$$

For this we use the subadditivity techniques from [16], Appendix A. It will turn out that $s \mapsto \mathbb{E}[T'_s]$ is only nearly subadditive which, however, is enough to establish (3.1). Note that we cannot hope for an analogous statement for T_s since $\mathbb{E}[T_s] = \infty$ for all $s > 0$.

3.1. *Bounds*.

First we prove coarse lower and upper bounds for T_s and T'_s which will be used to prove that, if the limit in (3.1) exists, then $0 < \sigma < \infty$ has to hold.

We start with a trivial lower bound which is true for any distance function, independent of the set of points which is used to construct it.

LEMMA 3.1 (Lower bound). *For any set of points $P \subset \mathbb{R}^d$, $x, y \in \mathbb{R}^d$, and $h > 0$ it holds*

$$d_{h,P}(x, y) \ge |x - y| - \operatorname{dist}(x, P) - \operatorname{dist}(y, P) \ge |x - y| - h,$$

and, in particular, for all $s \ge 0$

$$T_s \ge s - h_s, \qquad T'_s \ge s - h_s.$$

PROOF. We can assume that $d_{h,P}(x,y) < \infty$ since otherwise the inequality is trivially true. Let therefore $p \in \Pi_{h,P}(x,y)$ be a path with $m \in \mathbb{N}$ elements in X, the length of which realizes $d_{h,P}(x,y)$. Then it holds

$$\begin{aligned} d_{h,P}(x,y) &\geq \sum_{i=1}^{m-1} |p_{i+1} - p_i| \geq \left| \sum_{i=1}^{m-1} (p_{i+1} - p_i) \right| = |p_m - p_1| \\ &\geq |y - x + p_m - y - (p_1 - x)| \\ &\geq |x - y| - |p_m - y| - |p_1 - x| \\ &= |x - y| - \operatorname{dist}(x, P) - \operatorname{dist}(y, P) \\ &\geq |x - y| - h, \end{aligned}$$

using that the existence of a feasible path implies $\operatorname{dist}(x,P), \operatorname{dist}(y,P) \leq h/2$. The statements for T_s and T_s' follow from their definition as distances on $P := X$ and $P := \mathcal{X}_s$, respectively. □

Now we prove a high probability upper bound for the distance function on the Poisson point process which we will apply to T_s.

LEMMA 3.2 (Upper bound 1). *For all $x, y \in \mathbb{R}^d$ and $h > 0$ it holds*

$$\begin{aligned} \mathbb{P}\big(d_h(x,y) \leq C_d|x-y| + h\big) &\geq \mathbb{P}\big(d_h(x,y) \leq C_d|x-y| + \operatorname{dist}(x,X) + \operatorname{dist}(y,X)\big) \\ &\geq 1 - \exp\left(-\left(\frac{h}{C_d}\right)^d + \log\left(\frac{C_d|x-y|}{h}\right)\right), \end{aligned}$$

and, in particular, for all $s \geq 0$

$$\mathbb{P}(T_s \leq C_d s + h_s) \geq 1 - \exp\left(-\left(\frac{h_s}{C_d}\right)^d + \log\left(\frac{C_d s}{h_s}\right)\right) \quad \forall s > 0.$$

Here the constant C_d is defined in (2.6).

REMARK 3.3. The probability for this upper bound deteriorates for large distances if the step size $h > 0$ is fixed. Therefore, we have to use $h = h_s$ which shall be chosen as $h_s \sim \log(s)^{\frac{1}{d}}$ later.

PROOF OF LEMMA 3.2. Because of the spatial invariance of the Poisson process, it suffices to proof the statement for $d_h(0, se_1)$. We cover the line segment connecting 0 and se_1 by $M \in \mathbb{N}$ boxes $B_i := \{\frac{2i - Mr}{2M} e_1\} \oplus [-r, r]^{d-1}$, $i = 1, \ldots, M$, see Figure 2. The side length $r > 0$ is given by

$$r = \frac{h}{C_d}, \tag{3.2}$$

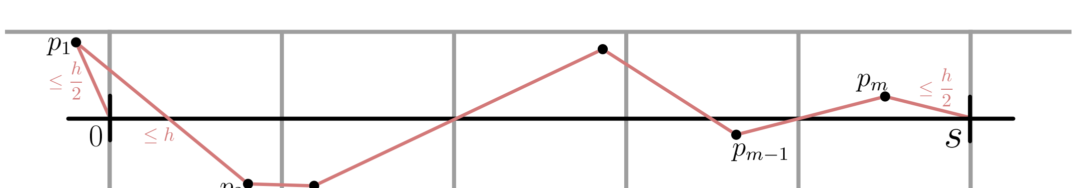

FIG. 2. *Boxes covering the line segment between 0 and se_1.*

where C_d given by (2.6) assures that the maximal distance of two points in two adjacent boxes B_i, B_{i+1} is bounded by h, and the maximal distance between 0 and the points in the first box B_1 and se_1 and the points in the last box B_M is bounded by $h/2$. Consequently, the number of boxes is

$$M = \frac{s}{r} = \frac{C_d s}{h}. \tag{3.3}$$

If each box contains a point from the Poisson cloud X, we can construct a valid path $p \in \Pi_h(0, se_1)$ which satisfies

$$\begin{aligned} d_h(0, se_1) &\le \operatorname{dist}(0, X) + \frac{h}{2} + (M-1)h + \frac{h}{2} + \operatorname{dist}(se_1, X) \\ &= C_d s + \operatorname{dist}(0, X) + \operatorname{dist}(se_1, X). \end{aligned}$$

Here, we used the triangle inequality to estimate $|p_1 - p_2| \le |p_1 - 0| + |0 - p_2| \le \operatorname{dist}(0, X) + h/2$ and similarly for the last term. Furthermore, the probabilility of this event is

$$\mathbb{P}(B_i \cap X \neq \varnothing \ \forall i \in \{1, \dots, M\}) = 1 - \mathbb{P}\left(\bigcup_{i=1}^{M} \{B_i \cap X = \varnothing\}\right).$$

Using a union bound, and (2.1) with $k = 0$ we obtain

$$\begin{aligned} \mathbb{P}\left(\bigcup_{i=1}^{M} \{B_i \cap X = \varnothing\}\right) &\le \sum_{i=1}^{M} \mathbb{P}(B_i \cap X = \varnothing) = \sum_{i=1}^{M} \exp(-\mu(B_i)) \\ &= M \exp(-r^d) = \frac{s}{r} \exp(-r^d) = \exp\left(-r^d + \log\left(\frac{s}{r}\right)\right). \end{aligned}$$

Furthermore, the definition of r in (3.2) implies

$$\mathbb{P}\left(\bigcup_{i=1}^{M} \{B_i \cap X = \varnothing\}\right) \le \exp\left(-\left(\frac{h}{C_d}\right)^d + \log\left(\frac{C_d s}{h}\right)\right).$$

□

LEMMA 3.4 (Upper bound 2). *For the constant C_d, defined in (2.6), for any $s > 1$, and for $h \ge \delta_s$ it holds almost surely*

$$d_{h,\mathcal{X}_s}(x, y) \le C_d |x - y| + h \quad \forall x, y \in \mathbb{R}^d,$$

and in particular for $h_s \ge \delta_s$

$$T'_s \le C_d s + h_s \quad \forall s > 0.$$

PROOF. The proof is the same as the one of Lemma 3.2 with the only difference being that the path which is constructed there uses the nonempty boxes from the definition of $\mathcal{X}_s$. □

REMARK 3.5 (Better upper bounds). It is important to remark that the upper bounds $d_h(x, y), d_{h,\mathcal{X}_s}(x, y) \le C_d |x - y| + h$ are quite coarse. Using the more careful strategy from [15], Lemma 5.5, one can obtain the (high probability) bounds

$$d_h(x, y), d_{h,\mathcal{X}_s}(x, y) \le \left(1 + C \frac{\delta_s}{h}\right) |x - y| + h,$$

where C is a dimensional constant. Since in our regime $h \sim \delta_s$ the constant in front of $|x - y|$ does not converge to 1 anyway, there is no need for us to use these improved bounds.

3.2. *The distances coincide with high probability.* It turns out that the two distance functions $d_h(\cdot,\cdot)$ and $d_{h,\mathcal{X}_s}(\cdot,\cdot)$ coincide with high probability. For this, we first show localization, that is, that optimal paths for the former distance lie in a sufficiently large ball with high probability.

LEMMA 3.6. *There exists a dimensional constant $C_d' \geq 1$ such that for $0 < h \leq |x-y|/2$ with probability at least $1 - \exp(-(\frac{h}{C_d})^d + \log(\frac{C_d|x-y|}{h}))$ any optimal path of $d_h(x,y)$ lies in $B(x, C_d'|x-y|)$.*

PROOF. Without loss of generality we assume $x = 0$ and $y = se_1$. By Lemma 3.2, with probability at least $1 - \exp(-(\frac{h}{C_d})^d + \log(\frac{C_d s}{h}))$ there exists an optimal path for $d_h(0, se_1)$ and it holds $d_h(0, se_1) \leq C_d s + h \leq (C_d + 1/2)s$. Let p be such an optimal path with $m := |p|$ elements. If p contained a point p_i outside $B(0, C_d' s)$ its length would satisfy

$$\begin{aligned} L(p) &\geq |p_1 - p_i| + |p_i - p_m| \geq 2|p_i| - |p_1| - |p_m| \geq 2|p_i| - \frac{h}{2} - \frac{h}{2} - |se_1| \\ &\geq 2C_d' s - h - s = (2C_d' - 1)s - h. \\ &= (2C_d' - 1)s\left(1 - \frac{h}{s}\right). \end{aligned}$$

By the assumption $h \leq s/2$ we get that the brackets are larger or equal than $\frac{1}{2}$. Hence, if we choose $C_d' \geq C_d + 3/2$ we get that

$$d_h(0, se_1) = L(p) \geq (C_d + 1)s$$

which is a contradiction. □

An analogous statement is satisfied by $d_{h,\mathcal{X}_s}(\cdot,\cdot)$, using the upper bound established in Lemma 3.4.

LEMMA 3.7. *Assume that $\delta_s \leq h \leq |x-y|/2$. Then any optimal path of $d_{h,\mathcal{X}_s}(x,y)$ lies in $B(x, C_d'|x-y|)$.*

PROOF. Using Lemma 3.4, the proof works exactly as the one of the previous lemma. □

Thanks to these two lemmata for any $x, y \in \mathbb{R}^d$ the distance $d_{h,\mathcal{X}_s}(x,y)$ in fact only depends on points in a compact set. Using properties of the Poisson process we can argue that the small boxes B_k from the definition of $d_{h,\mathcal{X}_s}(\cdot,\cdot)$ which fall into this compact set all contain a Poisson point with high probability. This then implies that $d_h(x,y) = d_{h,\mathcal{X}_s}(x,y)$ since no point has to be added to X.

LEMMA 3.8. *Let $x, y \in \mathbb{R}^d$ and $\delta_s \leq h \leq |x-y|/2$. Then it holds that*

$$\mathbb{P}\big(d_h(x,y) = d_{h,\mathcal{X}_s}(x,y)\big) \geq 1 - 2\exp\left(-\left(\frac{h}{C_d}\right)^d + d\log\left(\frac{2C_d C_d'|x-y|}{\delta_s}\right)\right).$$

PROOF. Again it suffices to prove the statement for $x = 0$ and $y = se_1$. Let E_s be the event any optimal path of $d_h(0, se_1)$ lies within $B(0, C_d' s)$. Then Lemma 3.6 shows

$$\mathbb{P}(E_s) \geq 1 - \exp\left(-\left(\frac{h_s}{C_d}\right)^d + \log\left(\frac{C_d s}{h}\right)\right). \tag{3.4}$$

After possibly enlarging C'_d a little we can assume that the box of side length $2C'_d s$ which contains $B(0, C'_d s)$ coincides with the union of $M \in \mathbb{N}$ boxes B_k which have a side length of δ_s/C_d. Here $M = (\frac{2C_d C'_d s}{\delta_s})^d$. As in the proof of Lemma 3.2, using (2.1) and a union bound shows that the probability that all of these boxes contain a point from X is at least

$$1 - M\exp\Big(-\Big(\frac{h}{C_d}\Big)^d\Big) = 1 - \exp\Big(-\Big(\frac{h}{C_d}\Big)^d + \log\Big(\Big(\frac{2C_d C'_d s}{\delta_s}\Big)^d\Big)\Big)$$
$$= 1 - \exp\Big(-\Big(\frac{h}{C_d}\Big)^d + d\log\Big(\frac{2C_d C'_d s}{\delta_s}\Big)\Big).$$

We call this event F_s and obtain

$$\mathbb{P}(F_s) \geq 1 - \exp\Big(-\Big(\frac{h}{C_d}\Big)^d + d\log\Big(\frac{2C_d C'_d s}{\delta_s}\Big)\Big). \tag{3.5}$$

Since according to Lemma 3.7 it holds $d_h(0, se_1) = d_{h,\mathcal{X}_s}(0, se_1)$ if all boxes contain a point from X, we obtain $E_s \cap F_s \subset \{d_h(0, se_1) = d_{h,\mathcal{X}_s}(0, se_1)\}$. Hence, using (3.4) and (3.5) and a union bound we get

$$\mathbb{P}\big(d_h(0, se_1) = d_{h,\mathcal{X}_s}(0, se_1)\big) \geq \mathbb{P}(E_s \cap F_s) = 1 - \mathbb{P}\big(E_s^c \cup F_s^c\big) \geq 1 - \mathbb{P}\big(E_s^c\big) - \mathbb{P}\big(F_s^c\big)$$
$$\geq 1 - \exp\Big(-\Big(\frac{h}{C_d}\Big)^d + \log\Big(\frac{C_d s}{h}\Big)\Big)$$
$$- \exp\Big(-\Big(\frac{h}{C_d}\Big)^d + d\log\Big(\frac{2C_d C'_d s}{\delta_s}\Big)\Big)$$
$$\geq 1 - 2\exp\Big(-\Big(\frac{h}{C_d}\Big)^d + d\log\Big(\frac{2C_d C'_d s}{\delta_s}\Big)\Big).$$

Here we also used that $d \geq 1$ and $2C'_d/\delta_s \geq 1/h$. □

3.3. *Approximate spatial invariance.* A main benefit of using distance functions over homogeneous Poisson point processes is their invariance with respect to isometric transformations like shifts, rotations, etc., which preserve the Lebesgue measure.

Using that the distance functions $d_{h,\mathcal{X}_s}(x, y)$ and $d_h(x, y)$ coincide with high probability, we can show that this invariance of $d_h(x, y)$ translates to $d_{h,\mathcal{X}_s}(x, y)$. In fact, we will need the slightly more general statement of the following lemma.

LEMMA 3.9. *Let $M \in \mathbb{N}$ and $x_i, y_i \in \mathbb{R}^d$ be points satisfying $|x_i - y_i| = \Delta$ for all $i = 1, \dots, M$ and $\delta_s \leq h \leq \Delta/2$. Let furthermore $\Phi : \mathbb{R}^d \to \mathbb{R}^d$ be an isometry. Then it holds*

$$\Big|\mathbb{E}\Big[\min_{1\leq i\leq M} d_{h,\mathcal{X}_s}(\Phi(x_i), \Phi(y_i))\Big] - \mathbb{E}\Big[\min_{1\leq i\leq M} d_{h,\mathcal{X}_s}(x_i, y_i)\Big]\Big|$$
$$\leq \exp\Big(-\Big(\frac{h}{C_d}\Big)^d + (d+1)\log\big(\max\{2C_d C'_d, 4C_d + 2\}\Delta\big) + \log M - d\log(\delta_s)\Big).$$

PROOF. Using that Φ is an isometry and applying Lemma 3.8 and a union bound, yields that the event

$$A := \{d_{h,\mathcal{X}_s}\big(\Phi(x_i), \Phi(y_i)\big) = d_h\big(\Phi(x_i), \Phi(y_i)\big) \text{ and } d_{h,\mathcal{X}_s}(x_i, y_i) = d_h(x_i, y_i)\ \forall i = 1, \dots, n\}$$

satisfies

$$\mathbb{P}(A) \geq 1 - 2M\exp\Big(-\Big(\frac{h}{C_d}\Big)^d + d\log\Big(\frac{2C_d C'_d \Delta}{\delta_s}\Big)\Big).$$

Hence, we can use the invariance of the distance function on the Poisson process X to get

$$\mathbb{E}\Big[\min_{1\le i\le M} d_h\big(\Phi(x_i),\Phi(y_i)\big)\mid A\Big]=\mathbb{E}\Big[\min_{1\le i\le M} d_h(x_i,y_i)\mid A\Big].$$

Therefore, we obtain

$$\begin{aligned}
&\mathbb{E}\Big[\min_{1\le i\le M} d_{h,\mathcal{X}_s}\big(\Phi(x_i),\Phi(y_i)\big)\Big]\\
&\quad=\mathbb{E}\Big[\min_{1\le i\le M} d_{h,\mathcal{X}_s}\big(\Phi(x_i),\Phi(y_i)\big)\mid A\Big]\mathbb{P}(A)+\mathbb{E}\Big[\min_{1\le i\le M} d_{h,\mathcal{X}_s}\big(\Phi(x_i),\Phi(y_i)\big)\mid A^c\Big]\mathbb{P}\big(A^c\big)\\
&\quad=\mathbb{E}\Big[\min_{1\le i\le M} d_{h}\big(\Phi(x_i),\Phi(y_i)\big)\mid A\Big]\mathbb{P}(A)+\mathbb{E}\Big[\min_{1\le i\le M} d_{h,\mathcal{X}_s}\big(\Phi(x_i),\Phi(y_i)\big)\mid A^c\Big]\mathbb{P}\big(A^c\big)\\
&\quad=\mathbb{E}\Big[\min_{1\le i\le M} d_{h}(x_i,y_i)\mid A\Big]\mathbb{P}(A)+\mathbb{E}\Big[\min_{1\le i\le M} d_{h,\mathcal{X}_s}\big(\Phi(x_i),\Phi(y_i)\big)\mid A^c\Big]\mathbb{P}\big(A^c\big)\\
&\quad=\mathbb{E}\Big[\min_{1\le i\le M} d_{h,\mathcal{X}_s}(x_i,y_i)\mid A\Big]\mathbb{P}(A)+\mathbb{E}\Big[\min_{1\le i\le M} d_{h,\mathcal{X}_s}\big(\Phi(x_i),\Phi(y_i)\big)\mid A^c\Big]\mathbb{P}\big(A^c\big)\\
&\quad=\mathbb{E}\Big[\min_{1\le i\le M} d_{h,\mathcal{X}_s}(x_i,y_i)\Big]\\
&\qquad+\Big(\mathbb{E}\Big[\min_{1\le i\le M} d_{h,\mathcal{X}_s}\big(\Phi(x_i),\Phi(y_i)\big)\mid A^c\Big]-\mathbb{E}\Big[\min_{1\le i\le M} d_{h,\mathcal{X}_s}(x_i,y_i)\mid A^c\Big]\Big)\mathbb{P}\big(A^c\big).
\end{aligned}$$

Reordering and trivially estimating $d_{h,\mathcal{X}_s}(\cdot,\cdot)$ using Lemma 3.4 we obtain

$$\begin{aligned}
&\Big|\mathbb{E}\Big[\min_{1\le i\le M} d_{h,\mathcal{X}_s}\big(\Phi(x_i),\Phi(y_i)\big)\Big]-\mathbb{E}\Big[\min_{1\le i\le M} d_{h,\mathcal{X}_s}(x_i,y_i)\Big]\Big|\\
&\quad\le 4M(C_d\Delta+h)\exp\Big(-\Big(\frac{h}{C_d}\Big)^d+d\log\Big(\frac{2C_dC_d'\Delta}{\delta_s}\Big)\Big)\\
&\quad=\exp\Big(-\Big(\frac{h}{C_d}\Big)^d+d\log\Big(\frac{2C_dC_d'\Delta}{\delta_s}\Big)+\log\big(4M(C_d\Delta+h)\big)\Big)\\
&\quad\le\exp\Big(-\Big(\frac{h}{C_d}\Big)^d+d\log\big(2C_dC_d'\Delta\big)+\log\big((4C_d+2)\Delta\big)+\log M-d\log(\delta_s)\Big)\\
&\quad\le\exp\Big(-\Big(\frac{h}{C_d}\Big)^d+(d+1)\log\big(\max\{2C_dC_d',4C_d+2\}\Delta\big)+\log M-d\log(\delta_s)\Big),
\end{aligned}$$

where we used the isometry of Φ and that $h\le\Delta/2$. $\square$

3.4. *Near subadditivity.* In this section we prove an approximate triangle inequality for the distance $d_h(\cdot,\cdot)$ which will then allow us to prove an *approximate subadditivity* property for $\mathbb{E}[T_s']$. Old results, which go back to Erdős and others, will then allow us to deduce (3.1).

First, we prove a general approximate triangle inequality for the distance function on an arbitrary set of points and different values of the length scale h.

LEMMA 3.10. *Let $P\subset\mathbb{R}^d$ be a set of points. Let $h_1,h_2>0$ and $h_3\ge\max(h_1,h_2)$. Then it holds*

$$d_{h_3,P}(x,y)\le d_{h_1,P}(x,z)+d_{h_2,P}(z,y)+h_3\quad\forall x,y,z\in\mathbb{R}^d. \tag{3.6}$$

PROOF. The statement follows from the simple observation that if $p\in\Pi_{h_1,P}(x,z)$ and $q\in\Pi_{h_2,P}(z,y)$ are optimal paths which realize $d_{h_1,P}(x,z)$ and $d_{h_2,P}(z,y)$ then $r:=(p,q)$ is a path in $\Pi_{h_3,P}(x,y)$. To see this, note that the last point in p has a distance of at most $h_1/2$

to z and the first point in q has a distance of at most $h_2/2$ to z. Using the triangle inequality the distance between the those two points is at most $h_1/2 + h_2/2 \leq h_3$ and consequently

$$d_{h_3,P}(x,y) \leq L(r) \leq L(p) + L(q) + h_3 = d_{h_1,P}(x,z) + d_{h_2,P}(z,y) + h_3. \qquad \square$$

A straightforward consequence of Lemma 3.10 would be that $s \mapsto \mathbb{E}[T_s]$ is near subadditive which by means of [16], Lemma A.2, implies that the limit $\lim_{s\to\infty} \frac{\mathbb{E}[T_s]}{s}$ exists. However, since there is a small but nonzero probability that $T_s = d_{h_s}(0, se_1) = \infty$, the expected value $\mathbb{E}[T_s]$ and this limit is infinite. Therefore, we investigate T_s' defined in (2.9).

From Lemma 3.8 we know that $T_s' = T_s$ with high probability and, furthermore, T_s' is always finite and satisfies $T_s' \leq T_s$. We introduce the error term $E_s := T_s - T_s' \geq 0$. For estimating it we now specify the choice of δ_s, the width of the boxes in the definition of T_s' in (2.9). We shall choose it in such a way that the error E_s is zero with high probability as $s \to \infty$.

ASSUMPTION 1. For a constant $k > 0$ and for $C_d'' := \max\{2C_d C_d', 4C_d + 2\}$ we choose

$$\delta_s = C_d \big(k \log(C_d'' s)\big)^{\frac{1}{d}}.$$

At this point we also fix the assumptions on the step size h_s:

ASSUMPTION 2. Let $s \mapsto h_s$ be nonincreasing and satisfy

$$\delta_s \leq h_s \ll s.$$

For these assumptions on δ_s and h_s (note that we are mainly interested in the case $h_s = \delta_s$) one can simplify the following term, which appears in a lot of probabilities:

$$\exp\left(-\left(\frac{h_s}{C_d}\right)^d + d\log\left(\frac{2C_d C_d' s}{\delta_s}\right)\right) \leq \frac{1}{\delta_s^d}\left(\frac{1}{2C_d C_d' s}\right)^{k-d} \tag{3.7}$$

and similarly for the error term in Lemma 3.9 with $M = 1$ and $s \geq \Delta/2$ we have

$$\exp\left(-\left(\frac{h}{C_d}\right)^d + (d+1)\log(C_d''\Delta) - d\log(\delta_s)\right) \leq \frac{2^k}{\delta_s^d}\left(\frac{1}{C_d''\Delta}\right)^{k-(d+1)}, \tag{3.8}$$

which is dominating (3.7). Using (3.7), the statement of Lemma 3.8 can be reformulated as follows

$$\mathbb{P}(E_s > 0) \leq \frac{2}{\delta_s^d}\left(\frac{1}{2C_d C_d' s}\right)^{k-d}. \tag{3.9}$$

Utilizing that the error E_s is zero with high probability and that we have the approximate triangle inequality from Lemma 3.10 we can show that $\mathbb{E}[T_s']$ is nearly subadditive.

PROPOSITION 3.11. *Under Assumptions 1 and 2 and for $k \geq d+1$ there exists a constant $C = C(d) > 0$ such that for all $s > 0$ sufficiently large and all $s \leq t \leq 2s$ it holds*

$$\mathbb{E}[T_{s+t}'] \leq \mathbb{E}[T_s'] + \mathbb{E}[T_t'] + Ch_{s+t}.$$

PROOF. We define the translation $\Phi_s : \mathbb{R}^d \to \mathbb{R}^d$, $x \mapsto x - se_1$ and note that it is a probability measure preserving transformation of $\mathbb{R}^d$. We define the event

$$A := \{T_{s+t}' = T_{s+t}\} \cap \{T_s' = T_s\} \cap \{T_t' \circ \Phi_s = T_t \circ \Phi_s\}, \tag{3.10}$$

abbreviating $T_t \circ \Phi_s := d_{h_t}(se_1, (s+t)e_1)$ and analogously $T_t' \circ \Phi_s := d_{h_t,\mathcal{X}_t}(se_1, (s+t)e_1)$.

Using the conditional expectation we obtain the following formula of total probability

$$\mathbb{E}[T'_{s+t}] = \mathbb{E}[T'_{s+t} \mid A]\mathbb{P}(A) + \mathbb{E}[T'_{s+t} \mid A^c]\mathbb{P}(A^c). \tag{3.11}$$

By definition of the T' random variable and the event A, and using the approximate triangle inequality from Lemma 3.10, we have

$$\begin{aligned} \mathbb{E}[T'_{s+t} \mid A] &= \mathbb{E}[T_{s+t} \mid A] = \mathbb{E}[d_{h_{s+t}}(0, (s+t)e_1) \mid A] \\ &\le \mathbb{E}[d_{h_s}(0, se_1) \mid A] + \mathbb{E}[d_{h_t}(se_1, (s+t)e_1) \mid A] + h_{s+t} \\ &= \mathbb{E}[T_s \mid A] + \mathbb{E}[T_t \circ \Phi_s \mid A] + h_{s+t} \\ &= \mathbb{E}[T'_s \mid A] + \mathbb{E}[T'_t \circ \Phi_s \mid A] + h_{s+t}. \end{aligned} \tag{3.12}$$

Combining (3.11) and (3.12), estimating $\mathbb{P}(A) \le 1$, and using also Lemmas 3.1 and 3.4 and the almost translation invariance of $\mathbb{E}[T'_s]$ from Lemma 3.9 in the case $M = 1$ together with (3.8) we get

$$\begin{aligned} \mathbb{E}[T'_{s+t}] &\le (\mathbb{E}[T'_s \mid A] + \mathbb{E}[T'_t \circ \Phi_s \mid A] + h_{s+t})\mathbb{P}(A) + \mathbb{E}[T'_{s+t} \mid A^c]\mathbb{P}(A^c) \\ &= \mathbb{E}[T'_s] + \mathbb{E}[T'_t \circ \Phi_s] + h_{s+t} \\ &\quad + (\mathbb{E}[T'_{s+t} \mid A^c] - \mathbb{E}[T'_s \mid A^c] - \mathbb{E}[T'_t \circ \Phi_s \mid A^c])\mathbb{P}(A^c) \\ &\le \mathbb{E}[T'_s] + \mathbb{E}[T'_t \circ \Phi_s] + h_{s+t} \\ &\quad + (C_d(s+t) + h_s - (s - h_s) - (t - h_t))\mathbb{P}(A^c) \\ &\le \mathbb{E}[T'_s] + \mathbb{E}[T'_t] + h_{s+t} + \frac{2^k}{\delta_t^d}\left(\frac{1}{C''_d t}\right)^{k-(d+1)} \\ &\quad + [(C_d - 1)(s+t) + 2h_s + 2h_t]\mathbb{P}(A^c). \end{aligned} \tag{3.13}$$

Using Lemma 3.8, Eq. (3.9), and Assumption 1 and the fact that $0 \le s \le t \le 2s$ we obtain that

$$\begin{aligned} \mathbb{P}(A^c) &\le \mathbb{P}(E_{s+t} > 0) + \mathbb{P}(E_s > 0) + \mathbb{P}(E_t > 0) \\ &\le C\left(\frac{(s+t)^{d-k}}{\log(C_d(s+t))} + \frac{s^{d-k}}{\log(C_d s)} + \frac{t^{d-k}}{\log(C_d t)}\right) \\ &\le C\frac{(s+t)^{d-k}}{\log(C_d(s+t)/3)}, \end{aligned}$$

where the constant C is dimensional and changes its value. Plugging this estimate into (3.13) and using Assumption 2, we obtain that

$$\mathbb{E}[T'_{s+t}] \le \mathbb{E}[T'_s] + \mathbb{E}[T'_t] + h_{s+t} + C\frac{(s+t)^{d+1-k}}{\log(C_d(s+t)/3)},$$

where C again changed its value. For $k \ge d+1$ and using Assumption 1 we can absorb the second error term into the first one. Changing C again concludes the proof. □

3.5. *Convergence.* Utilizing the bounds and the near subadditivity we obtain the following result:

PROPOSITION 3.12. *Assume that δ_s satisfies Assumption 1 with $k \ge d+1$ and h_s satisfies Assumption 2 with the additional requirement that for s sufficiently large it holds $h_s \le Cs^\alpha$ for some constant $C > 0$ and some $\alpha \in (0,1)$. Then the limit*

$$\sigma := \lim_{s\to\infty} \frac{\mathbb{E}[T'_s]}{s} \tag{3.14}$$

exists and satisfies $\sigma \in [1, C_d]$.

PROOF. For $\alpha \in (0,1)$ the function $g(z) := Cz^\alpha$ satisfies $\int_{z_0}^\infty g(z)z^{-2} < \infty$ for $z_0 > 0$. Hence, Proposition 3.11 and [16], Lemma A.2, imply that $\sigma := \lim_{s\to\infty} \frac{\mathbb{E}[T_s']}{s}$ exists. By Lemmas 3.1 and 3.4 the random variable T_s' satisfies the deterministic bounds

$$s - h_s \le T_s' \le C_d s + h_s.$$

Taking the expectation, dividing by s, using $h_s \ll s$ and $C_d \ge 1$ shows that $\sigma \in [1, C_d]$. □

REMARK 3.13 (The constant σ). As already pointed out in Remark 3.5, the constant σ can be brought arbitrarily close to 1 by multiplying h_s with a large constant. Note that it is not required to demand a quicker growth than $h_s \sim \log(s)^{\frac{1}{d}}$.

4. Concentration of measure and almost sure convergence. In this section we will prove concentration of measure for T_s' around its expectation. This will have two important consequences: First, it will allow us to show that

$$\lim_{s\to\infty} \frac{T_s}{s} = \sigma \quad \text{almost surely},$$

where σ is the constant from (3.1) and Proposition 3.12. Second, we will use concentration to prove an approximate superadditivity property which will be the key to convergence rates.

4.1. *Concentration of measure.* In this section we prove concentration of measure for T_s'. In fact, we will prove a slightly more general statement, namely concentration of measure for the distance function $d_{h_s,\mathcal{X}_s}(x,y)$. Since $T_s' = d_{h_s,\mathcal{X}_s}(0, se_1)$ concentration for T_s' will be a special case.

THEOREM 4.1 (Concentration of measure for $d_{h,\mathcal{X}_s}(\cdot,\cdot)$). *There exist dimensional constants $C_1, C_2 > 0$ such that for all $s > 0$ and all $x, y \in \mathbb{R}^d$ with $|x-y| \ge h \ge \delta_s$ it holds*

$$\mathbb{P}\left(\left|d_{h,\mathcal{X}_s}(x,y) - \mathbb{E}\left[d_{h,\mathcal{X}_s}(x,y)\right]\right| > \lambda\sqrt{\frac{\delta_s^2}{h}|x-y|}\right) \le C_1 \exp(-C_2\lambda) \quad \forall \lambda \ge 0.$$

REMARK 4.2. We will be mostly interested in the regime where $h = h_s = \delta_s = C_d(k\log(2C_dC_d's))^{\frac{1}{d}}$, in which case we get for $|x-y| \ge h_s$:

$$\mathbb{P}\left(\left|d_{h_s,\mathcal{X}_s}(x,y) - \mathbb{E}\left[d_{h_s,\mathcal{X}_s}(x,y)\right]\right| > \lambda\sqrt{h_s|x-y|}\right) \le C_1 \exp(-C_2\lambda) \quad \forall \lambda \ge 0.$$

This means that with high probability the fluctuations of $d_{h_s,\mathcal{X}_s}(x,y)$ around its expectation are of order $\sqrt{|x-y|}$ modulo a log factor. However, the general result from Theorem 4.1 can also be applied to larger length scales $h_s \gg \delta_s$ in which case the fluctuations are smaller.

PROOF. The proof relies on an application of the abstract martingale estimate from [16], Lemma B.1. We follow the proofs of [50], Theorem 1, or [46], Theorem 2.1. For a more compact notation we use the following abbreviation:

$$T := d_{h,\mathcal{X}_s}(x,y).$$

Step 1: We define a filtration $\mathbb{F} := \{\mathcal{F}_k\}_{k\in\mathbb{N}_0}$ of the probability space by setting $\mathcal{F}_0 := \{\emptyset, \Omega\}$, $\mathcal{F}_k := \mathcal{F}(B_1 \cup \cdots \cup B_k)$ for $k \ge 1$. By $\mathcal{F}(A)$ we refer to the σ-subfield of $\mathcal{F} := \mathcal{F}(\mathbb{R}^d)$ which is generated by events of the form $\{X \cap A \neq \emptyset\}$ for Borel sets $A \subset \mathbb{R}^d$. We also define the martingale $M_k := \mathbb{E}[T \mid \mathcal{F}_k] - \mathbb{E}[T]$ with $M_0 = 0$ and we define $\Delta_k := M_k - M_{k-1} = \mathbb{E}[T \mid \mathcal{F}_k] - \mathbb{E}[T \mid \mathcal{F}_{k-1}]$.

Step 2: We want to compute a constant $c > 0$ for which $|\Delta_k| \le c$.

Let us define the random variable

$$T^{(k)} := d_{h,\mathcal{X}_s \setminus B_k}(x, y)$$

as the graph distance on the enriched Poisson point process without the kth box. Using that $\mathbb{E}[T^{(k)} \mid \mathcal{F}_{k-1}] = \mathbb{E}[T^{(k)} \mid \mathcal{F}_k]$ and trivially $T \le T^{(k)}$ we have

$$\begin{aligned}\Delta_k &= \mathbb{E}[T \mid \mathcal{F}_k] - \mathbb{E}[T \mid \mathcal{F}_{k-1}] = \mathbb{E}\big[T - T^{(k)} \mid \mathcal{F}_k\big] + \mathbb{E}\big[T^{(k)} - T \mid \mathcal{F}_{k-1}\big] \\ &\le \mathbb{E}\big[T^{(k)} - T \mid \mathcal{F}_{k-1}\big]\end{aligned}$$

and analogously $-\Delta_k \le \mathbb{E}[T^{(k)} - T \mid \mathcal{F}_k]$. This implies

$$|\Delta_k| \le \mathbb{E}\big[T^{(k)} - T \mid \mathcal{F}_{k-1}\big] \vee \mathbb{E}\big[T^{(k)} - T \mid \mathcal{F}_k\big].$$

For bounding $T^{(k)} - T$ we argue as follows: By definition we know that there exist feasible paths of finite length for T and hence also optimal paths. Let F_k be the event that the optimal path $p = (p_1, \dots, p_m)$ for T contains at least one point $p_i \in B_k$ for $i \in \{1, \dots, m\}$. On F_k^c it obviously holds $T^{(k)} = T$. On F_k we can—using that all boxes contain a Poisson point—construct an alternative path around the no-go box B_k which is at most $C_d'''\delta_s$ longer than T and is feasible for $T^{(k)}$. Here C_d''' is a suitable dimensional constant, depending on C_d. In either case we have $T^{(k)} - T \le C_d'''\delta_s$ and therefore also

$$|\Delta_k| \le C_d'''\delta_s =: c. \tag{4.1}$$

Step 3: In this step we want to find a sequence of positive $\mathcal{F}$-measurable random variables $\{U_k\}_{k\in\mathbb{N}}$ such that $\mathbb{E}[\Delta_k^2 \mid \mathcal{F}_{k-1}] \le \mathbb{E}[U_k \mid \mathcal{F}_{k-1}]$.

For any two random variables X, Y with Y measurable with respect to a σ-field $\mathcal{G}$ the projection identity (see, e.g., [37], Prop. 1.26)

$$\mathbb{E}\big[\big(X - \mathbb{E}[X \mid \mathcal{G}]\big)^2 \mid \mathcal{G}\big] \le \mathbb{E}\big[(X - Y)^2 \mid \mathcal{G}\big]$$

holds. We shall use this with $X := \mathbb{E}[T \mid \mathcal{F}_k]$, $Y := \mathbb{E}[T^{(k)} \mid \mathcal{F}_{k-1}] = \mathbb{E}[T^{(k)} \mid \mathcal{F}_k]$, and $\mathcal{G} := \mathcal{F}_{k-1}$. Note that we have the tower property $\mathbb{E}[X \mid \mathcal{G}] = \mathbb{E}[\mathbb{E}[T \mid \mathcal{F}_k] \mid \mathcal{F}_{k-1}] = \mathbb{E}[T \mid \mathcal{F}_{k-1}]$. Hence, we can compute

$$\begin{aligned}\mathbb{E}\big[\Delta_k^2 \mid \mathcal{F}_{k-1}\big] &= \mathbb{E}\big[\big(\mathbb{E}[T \mid \mathcal{F}_k] - \mathbb{E}[T \mid \mathcal{F}_{k-1}]\big)^2 \mid \mathcal{F}_{k-1}\big] \\ &= \mathbb{E}\big[\big(X - \mathbb{E}[X \mid \mathcal{G}]\big)^2 \mid \mathcal{F}_{k-1}\big] \\ &\le \mathbb{E}\big[(X - Y)^2 \mid \mathcal{F}_{k-1}\big] \\ &= \mathbb{E}\big[\big(\mathbb{E}[T \mid \mathcal{F}_k] - \mathbb{E}\big[T^{(k)} \mid \mathcal{F}_k\big]\big)^2 \mid \mathcal{F}_{k-1}\big] \\ &= \mathbb{E}\big[\mathbb{E}\big[T - T^{(k)} \mid \mathcal{F}_k\big]^2 \mid \mathcal{F}_{k-1}\big] \\ &\le \mathbb{E}\big[\mathbb{E}\big[\big(T - T^{(k)}\big)^2 \mid \mathcal{F}_k\big] \mid \mathcal{F}_{k-1}\big] \\ &= \mathbb{E}\big[\big(T - T^{(k)}\big)^2 \mid \mathcal{F}_{k-1}\big] \\ &= \mathbb{E}[U_k \mid \mathcal{F}_{k-1}],\end{aligned}$$

using Jensen's inequality and again the tower property for the last two steps and defining the $\mathcal{F}$-measurable random variable $U_k := (T - T^{(k)})^2$.

Step 4: Here we want to find a constant $\lambda_0 \ge \frac{c^2}{4e}$ such that for all $K \in \mathbb{N}$

$$S_K := \sum_{k=1}^{K} U_k \le \lambda_0 \quad \text{almost surely.} \tag{4.2}$$

For this we first note that U_k equals zero whenever there exists an optimal path for T which does not use a point in the box B_k. We now fix an optimal path with the smallest number of elements, denote it by p, and abbreviate its number of points by $|p|$. We define the index set $\mathcal{K} := \{k \in \{1, \dots, K\} : \exists i \in \{1, \dots, |p|\}, p_i \in B_k\}$ and get

$$S_K = \sum_{k \in \mathcal{K}} \left(T^{(k)} - T\right)^2.$$

Using that $T^{(k)} - T \leq C_d''' \delta_s$ and that the cardinality of $\mathcal{K}$ is at most $|p|$, we obtain the bound

$$S_K \leq \left(C_d''' \delta_s\right)^2 |p|. \tag{4.3}$$

Our next goal is to upper-bound $|p|$ by a constant times T/h for which we basically want to argue that most of the hops in the path p have a length of order h.

Let us abbreviate $m := |p|$. Our first claim is that for every $\gamma > 0$ and $j \in \{1, \dots, m-2\}$ we have:

$$|p_{j+1} - p_j| \leq \gamma \quad \Longrightarrow \quad |p_{j+2} - p_{j+1}| > h - \gamma. \tag{4.4}$$

If this were not the case, then it would hold

$$|p_{j+2} - p_j| \leq |p_{j+1} - p_j| + |p_{j+2} - p_{j+1}| \leq \gamma + h - \gamma = h.$$

Hence, the path $q := (p_1, \dots, p_j, p_{j+2}, \dots, p_m)$ would be feasible for T, and would satisfy $L(q) \leq L(p)$ as well as $|q| = |p| - 1$. Since p is optimal and shortest, this is a contradiction.

With this at hand we define the index sets

$$\begin{aligned}
I_\gamma &:= \{j \in \{1, \dots, m-2\} : |p_{j+1} - p_j| \leq \gamma\}, \\
I_\gamma' &:= \{j + 1 : j \in I_\gamma\}, \\
I_\gamma'' &:= \{1, \dots, m-1\} \setminus \left(I_\gamma \cup I_\gamma'\right).
\end{aligned}$$

Note that for $0 < \gamma < h/2$, the implication (4.4) shows that I_γ and I_γ' are disjoint and their cardinalities coincide. We abbreviate the latter by $k \in \{0, \dots, m-2\}$ and note that the cardinality of I_γ'' equals $m - 1 - 2k$. Using this we can estimate T from below as follows:

$$\begin{aligned}
T &= \sum_{j=1}^{m-1} |p_{j+1} - p_j| \\
&= \sum_{j \in I_\gamma} \underbrace{|p_{j+1} - p_j|}_{\geq 0} + \sum_{j \in I_\gamma'} \underbrace{|p_{j+1} - p_j|}_{\geq h - \gamma} + \sum_{j \in I_\gamma''} \underbrace{|p_{j+1} - p_j|}_{> \gamma} \\
&\geq k(h - \gamma) + (m - 1 - 2k)\gamma \\
&= k(h - 3\gamma) + (m - 1)\gamma.
\end{aligned}$$

Choosing $\gamma = h/6$ the first term is nonnegative and we obtain the estimate

$$m \leq 6\frac{T}{h} + 1.$$

Plugging this into our previous bound (4.3) for S, we obtain

$$S_K \leq \left(C_d''' \delta_s\right)^2 \left(6\frac{T}{h} + 1\right) = 6\left(C_d''' \delta_s\right)^2 \frac{T}{h} + \left(C_d''' \delta_s\right)^2. \tag{4.5}$$

Utilizing that $\delta_s \le h \le |x-y|$ and, according to Lemma 3.4, also $T \le C_d|x-y|+h$, we get for a suitable constant $C_d'''' > 0$ that

$$S_K \le C_d'''' \frac{\delta_s^2}{h}|x-y|. \tag{4.6}$$

Using that $h \le |x-y|$ we can possibly enlarge C_d'''' a little such that

$$\lambda_0 := C_d'''' \frac{\delta_s^2}{h}|x-y| \tag{4.7}$$

satisfies $\lambda_0 \ge \frac{c^2}{4e}$ where $c = C_d'''\delta_s$ was defined in (4.1). Hence, we have established the almost sure bound (4.2).

Step 5: We have checked all assumptions for [16], Lemma B.1, which lets us conclude

$$\mathbb{P}\big(T - \mathbb{E}[T] > \varepsilon\big) \le C \exp\biggl(-\frac{1}{2\sqrt{e\lambda_0}}\varepsilon\biggr) \quad \forall \varepsilon \ge 0.$$

Plugging in $\lambda_0 = C_d'''' \frac{\delta_s^2}{h}|x-y|$ yields

$$\begin{aligned}\mathbb{P}\big(T - \mathbb{E}[T] > \varepsilon\big) &\le C \exp\biggl(-\frac{1}{2\sqrt{eC_d''''\frac{\delta_s^2}{h}|x-y|}}\varepsilon\biggr) \\ &= C \exp\biggl(-\frac{\sqrt{h}}{2\sqrt{eC_d''''|x-y|}\delta_s}\varepsilon\biggr) \quad \forall \varepsilon \ge 0.\end{aligned}$$

For the choice $\varepsilon = \lambda\sqrt{\frac{\delta_s^2}{h}|x-y|}$ we get

$$\mathbb{P}\biggl(T - \mathbb{E}[T] > \lambda\sqrt{\frac{\delta_s^2}{h}}\sqrt{|x-y|}\biggr) \le C\exp(-C_2\lambda) \quad \forall \lambda \ge 0.$$

Repeating the proof verbatim for $\mathbb{E}[T] - T$ we finally obtain

$$\mathbb{P}\biggl(\big|T - \mathbb{E}[T]\big| > \lambda\sqrt{\frac{\delta_s^2}{h}}\sqrt{|x-y|}\biggr) \le C_1\exp(-C_2\lambda) \quad \forall \lambda \ge 0,$$

where $C_1 := 2C$. □

The fact that with high probability T_s and T_s' coincide allows us to deduce concentration of T_s around $\mathbb{E}[T_s']$ (remember that the expectation of T_s is infinite which is why concentration around it is irrelevant).

COROLLARY 4.3 (Concentration of measure for T_s). *Under the assumptions of Theorem 4.1 and Lemma 3.8 it holds for $s > 0$ sufficiently large*

$$\begin{aligned}\mathbb{P}\biggl(\big|T_s - \mathbb{E}[T_s']\big| > \lambda\sqrt{\frac{\delta_s^2}{h_s}s}\biggr) \le 2\exp\biggl(-\Bigl(\frac{h_s}{C_d}\Bigr)^d + d\log\Bigl(\frac{2C_dC_d's}{\delta_s}\Bigr)\biggr) \\ + C_1\exp(-C_2\lambda) \quad \forall \lambda \ge 0.\end{aligned}$$

PROOF. Utilizing that $T_s' = d_{h_s,\mathcal{X}_s}(0, se_1)$ and using Theorem 4.1 and Lemma 3.8, for all $\lambda \geq 0$ it holds

$$\begin{aligned}
&\mathbb{P}\left(|T_s - \mathbb{E}[T_s']| > \lambda\sqrt{\frac{\delta_s^2}{h_s}}\sqrt{s}\right) \\
&\quad \leq \mathbb{P}(|T_s - T_s'| > \lambda\sqrt{h_s s}) + \mathbb{P}\left(|T_s' - \mathbb{E}[T_s']| > \lambda\frac{\delta_s}{\sqrt{h_s}}\sqrt{s}\right) \\
&\quad \leq \mathbb{P}(|T_s - T_s'| > 0) + \mathbb{P}\left(|T_s' - \mathbb{E}[T_s']| > \lambda\frac{\delta_s}{\sqrt{h_s}}\sqrt{s}\right) \\
&\quad \leq 2\exp\left(-\left(\frac{h_s}{C_d}\right)^d + d\log\left(\frac{2C_d C_d' s}{\delta_s}\right)\right) + C_1\exp(-C_2\lambda).
\end{aligned}$$

□

4.2. *Almost sure convergence.* Combining concentration of measure with the convergence in expectation from Proposition 3.12, we can now prove almost sure convergence of T_s'/s and even of T_s/s. Note that Kingman's subadditive ergodic theorem is not applicable in this case since the random variables T_s have infinite expectations, hence, are not in L^1. An additional difficulty arises from T_s and T_s' being stochastic processes with a continuous variable $s \in (0, \infty)$. We prove all statements for a subsequence of integers and to use Lipschitz regularity to extend to the real line.

THEOREM 4.4. *Assume that δ_s satisfies Assumption 1 with $k > d + 1$ and h_s satisfies Assumption 2 with the additional requirement that for s sufficiently large it holds $h_s \leq Cs^\alpha$ for some constant $C > 0$ and some $\alpha \in (0, 1)$. Then it holds*

$$\lim_{s\to\infty} \frac{T_s}{s} = \sigma \quad \textit{almost surely},$$

where σ denotes the constant from Proposition 3.12.

REMARK 4.5. As outlined in Remark 3.5 one can make sure that σ is arbitrarily close (but not equal) to 1 by multiplying h_s with a large constant.

PROOF. Let $\varepsilon > 0$ be arbitrary and choose $\lambda = \varepsilon\sqrt{s}\sqrt{\frac{h_s}{\delta_s^2}}$. Then Theorem 4.1 implies

$$\mathbb{P}(|T_s' - \mathbb{E}[T_s']| > s\varepsilon) \leq C_1\exp\left(-C_2\varepsilon\sqrt{s}\sqrt{\frac{h_s}{\delta_s^2}}\right).$$

Let now $s := n$ where $n \in \mathbb{N}$ is a natural number. By assumption we have $\delta_n \leq h_n \leq Cn^\alpha$ which implies that

$$\exp\left(-C_2\varepsilon\sqrt{n}\sqrt{\frac{h_n}{\delta_n^2}}\right) \leq \exp\left(-C_2\varepsilon\frac{\sqrt{n}}{\sqrt{\delta_n}}\right) = \exp(-C_2\varepsilon n^{\frac{1-\alpha}{2}}).$$

Now we use that for all $m \in \mathbb{N}$ and $x > 0$ it holds

$$\exp(-x) \leq \frac{m!}{x^m}$$

to obtain that

$$\exp\left(-C_2\varepsilon\sqrt{n}\sqrt{\frac{h_n}{\delta_n^2}}\right) \leq \frac{m!}{(C_2\varepsilon(n)^{\frac{1-\alpha}{2}})^m}.$$

If we choose $m > \frac{2}{1-\alpha}$ we obtain

$$\sum_{n=1}^{\infty} \exp\left(-C_2\varepsilon\sqrt{n}\sqrt{\frac{h_n}{\delta_n^2}}\right) \le \frac{m!}{(C_2\varepsilon)^m} \frac{1}{t^{\frac{m(1-\alpha)}{2}}} \sum_{n=1}^{\infty} \frac{1}{n^{\frac{m(1-\alpha)}{2}}} < \infty.$$

Hence, the Borel–Cantelli lemma allows us to conclude that

$$\mathbb{P}\Big(\limsup_{n\to\infty}\{|T_n' - \mathbb{E}[T_n']| > n\varepsilon\}\Big) = 0.$$

Since $\varepsilon > 0$ was arbitrary, we obtain that $|\frac{T_n'}{n} - \frac{\mathbb{E}[T_n']}{n}| \to 0$ almost surely as $n \to \infty$. Together with Proposition 3.12 this implies

$$\lim_{n\to\infty} \frac{T_n'}{n} = \sigma \quad \text{almost surely.} \tag{4.8}$$

We claim that we also have

$$\lim_{n\to\infty} \frac{T_n}{n} = \sigma \quad \text{almost surely.} \tag{4.9}$$

To see this, let $\varepsilon > 0$ be arbitrary and choose $\lambda = \varepsilon\sqrt{s}\sqrt{\frac{h_s}{\delta_s^2}}$. Then Corollary 4.3, Assumptions 1 and 2, and Eq. (3.7) imply

$$\mathbb{P}(|T_s - \mathbb{E}[T_s']| > s\varepsilon) \le \frac{2}{\delta_s^d}\left(\frac{1}{2C_dC_d's}\right)^{k-d} + C_1\exp\left(-C_2\varepsilon\sqrt{s}\sqrt{\frac{h_s}{\delta_s^2}}\right).$$

Let again $s := n$ where $n \in \mathbb{N}$ is a natural number. Using that $\delta_n \ge 1$ for n sufficiently large and that $k > d+1$ we get that

$$\sum_{n=1}^{\infty} \frac{2}{\delta_n^d}\left(\frac{1}{2C_dC_d'n}\right)^{k-d} < \infty.$$

Furthermore, by assumption we have $\delta_n \le h_n \le Cn^\alpha$ which implies that

$$\exp\left(-C_2\varepsilon\sqrt{n}\sqrt{\frac{h_n}{\delta_n^2}}\right) \le \exp\left(-C_2\varepsilon\frac{\sqrt{n}}{\sqrt{\delta_n}}\right) = \exp(-C_2\varepsilon n^{\frac{1-\alpha}{2}}).$$

Now we use that for all $m \in \mathbb{N}$ and $x > 0$ it holds

$$\exp(-x) \le \frac{m!}{x^m}$$

to obtain that

$$\exp\left(-C_2\varepsilon\sqrt{n}\sqrt{\frac{h_n}{\delta_n^2}}\right) \le \frac{m!}{(C_2\varepsilon(n)^{\frac{1-\alpha}{2}})^m}.$$

If we choose $m > \frac{2}{1-\alpha}$ we obtain

$$\sum_{n=1}^{\infty} \exp\left(-C_2\varepsilon\sqrt{n}\sqrt{\frac{h_n}{\delta_n^2}}\right) \le \frac{m!}{(C_2\varepsilon)^m} \frac{1}{t^{\frac{m(1-\alpha)}{2}}} \sum_{n=1}^{\infty} \frac{1}{n^{\frac{m(1-\alpha)}{2}}} < \infty.$$

Hence, the Borel–Cantelli lemma allows us to conclude that

$$\mathbb{P}\Big(\limsup_{n\to\infty}\{|T_n - \mathbb{E}[T_n']| > n\varepsilon\}\Big) = 0.$$

Since $\varepsilon > 0$ was arbitrary, we obtain that $|\frac{T_n}{n} - \frac{\mathbb{E}[T_n']}{n}| \to 0$ almost surely as $n \to \infty$. Together with Proposition 3.12 this establishes the claim, proving (4.9).

We now extend the limits to hold for real-valued $s \to \infty$. We first show that

$$\lim_{s\to\infty} \frac{T_s}{s} = \sigma \quad \text{almost surely.} \tag{4.10}$$

To see this, we let $B(x,t) := \{y \in \mathbb{R}^d : |x-y| \le t\}$ denote the closed ball around $x \in \mathbb{R}^d$ with radius $t > 0$ and let A_n denote the event that

$$X \cap B\left((n+1)e_1, \frac{h_n}{4}\right) \neq \varnothing.$$

By the law of the Poisson point process and the choice of scaling of h_n we have

$$\sum_{n=1}^{\infty} \mathbb{P}(A_n^c) = \sum_{n=1}^{\infty} \exp\left(-\omega_d \frac{h_n^d}{4^d}\right) < \infty, \tag{4.11}$$

where ω_d is the volume of the d-dimensional unit ball. When A_n occurs, let us denote by x_n any point in the intersection of X with $B((n+1)e_1, \frac{h_n}{4})$. We also assume n is large enough so that $\frac{h_n}{4} \ge 1$.

We now claim that whenever A_n occurs and T_n is finite, we have

$$T_{n+1} - h_n \le T_s \le T_n + h_n \quad \text{for all } n \le s \le n+1. \tag{4.12}$$

To see this, note that any optimal path for T_n must terminate at a point x within distance $\frac{h_n}{2}$ of ne_1. Since A_n occurs we can add the point x_n to this path to obtain a feasible path for T_s. Indeed, we simply note that $h_n \le h_s$ and compute

$$|x - x_n| \le |x - ne_1| + \left|ne_1 - (n+1)e_1\right| + \left|x_n - (n+1)e_1\right| \le \frac{h_n}{2} + 1 + \frac{h_n}{4} \le h_n \le h_s,$$

and

$$|x_n - se_1| \le \left|x_n - (n+1)e_1\right| + \left|(n+1)e_1 - se_1\right| \le \frac{h_n}{4} + 1 \le \frac{h_n}{2} \le \frac{h_s}{2}.$$

Note that we used that $\frac{h_n}{4} \ge 1$ in both inequalities. It follows that T_s is finite and

$$T_s \le T_n + h_n.$$

To prove the other inequality, we follow a similar argument, taking a path that is optimal for T_s, which must terminate at a point y that is within distance $\frac{h_n}{2}$ of se_1, and concatenating the point x_n to obtain a feasible path for T_{n+1}. This yields the inequality

$$T_{n+1} \le T_s + h_n,$$

which establishes the claim. The proof of (4.10) is completed by dividing by s in (4.12), recalling (4.11) and applying Borel–Cantelli. □

5. Near superadditivity and ratio convergence. In this section we prove a type of approximate superadditivity of the distance function with the aim of proving convergence rates. Ideally, we would like to show that for a slowly increasing function $s \mapsto g(s)$

$$\mathbb{E}[T_{2s}'] \ge 2\mathbb{E}[T_s'] - g(s) \tag{5.1}$$

holds true. Together with the near subadditivity from Proposition 3.11, the convergence from Proposition 3.12, and [16], Lemma A.2, this would directly imply quantitative convergence rates for $\frac{\mathbb{E}[T_s']}{s}$ to the constant σ. The concentration statement from Theorem 4.1 would then yield almost sure rates for $\frac{T_s'}{s}$ to σ.

Although we think that (5.1) might be true, a proof of this is very difficult since the distance functions in the definition of T_{2s}' and T_s' utilize the different length scales h_{2s} and h_s. Consequently, a path which realizes T_{2s}' is typically not feasible for the distance in T_s' which makes a construction of a suboptimal path for this distance such that (5.1) is satisfied hard. This is a specific problem of our sparse graph setting and can be avoided using a fully connected graph as, for example, in [46].

Therefore, we shall not work with the random variables T_s' or T_s in the following but rather work with a fixed length scale h and the distance function $d_{h,\mathcal{X}_s}(\cdot,\cdot)$ defined in (2.8). That is, we aim to prove near superadditivity of the form

$$\mathbb{E}\big[d_{h,\mathcal{X}_s}(0,2se_1)\big] \geq 2\mathbb{E}\big[d_{h,\mathcal{X}_s}(0,se_1)\big] - g(s), \tag{5.2}$$

where we emphasize that both distance functions on the left and on the right have the same length scale h.

5.1. *Near superadditivity.* We start by proving the following proposition which asserts near superadditivity of the form (5.2). The argument closely follows the proof given in [46], Lemma 4.1.

PROPOSITION 5.1 (Near superadditivity). *Let δ_s satisfy assumption Assumption 1 with $k > d+1$. There exist dimensional constants $C_1, C_2 > 0$ such that for all $s > 1$ sufficiently large with $\delta_s \leq h \leq s$ we have that*

$$\mathbb{E}\big[d_{h,\mathcal{X}_s}(0,2se_1)\big] \geq 2\mathbb{E}\big[d_{h,\mathcal{X}_s}(0,se_1)\big] - C_1 h - C_2\sqrt{\frac{\delta_s^2}{h}s}\,\log(s).$$

REMARK 5.2. In the case that $h = h_s = \delta_s$ we can subsume the error terms into one and for some dimensional constant $C > 0$ we have

$$\mathbb{E}\big[d_{h_s,\mathcal{X}_s}(0,2se_1)\big] \geq 2\mathbb{E}\big[d_{h_s,\mathcal{X}_s}(0,se_1)\big] - C\log(s)^{\frac{1+2d}{2d}}\sqrt{s}.$$

PROOF. Let $p_1,\dots,p_m$ be a path realizing the length $d_{h,\mathcal{X}_s}(0,2se_1)$. We now consider the balls $B(0,s)$, $B(2se_1,s)$ and denote by i_s, i_{2s} the indices such that

$$p_i \in B(0,s) \quad \forall i \leq i_s, \qquad p_{i_s+1} \notin B(0,s),$$
$$p_i \in B(2e_1,s) \quad \forall i \geq i_{2s}, \qquad p_{i_{2s}-1} \notin B(2se_1,s),$$

that is, the smallest index i_s after which the path leaves $B(0,s)$ and the largest index i_{2s} before which the path enters $B(2se_1,s)$. We note that by definition these indices exist and that $i_s < i_{2s}$ holds. The construction is illustrated in Figure 3. We now take the points where the path intersects the respective spheres,

$$x_s := \overline{p_{i_s}p_{i_s+1}} \cap \partial B(0,s),$$
$$x_{2s} := \overline{p_{i_{2s}-1}p_{i_{2s}}} \cap \partial B(2se_1,s),$$

for which we have

$$\begin{aligned} d_{h,\mathcal{X}_s}(0,2se_1) &= \sum_{i=1}^{m-1}|p_{i+1}-p_i| \\ &\geq \sum_{i=1}^{i_s-1}|p_{i+1}-p_i| + |p_{i_s}-x_s| + |x_{2s}-p_{i_{2s}}| + \sum_{i=i_{2s}}^{m-1}|p_{i+1}-p_i| \qquad (5.3)\\ &\geq d_{h,\mathcal{X}_s}(0,p_{i_s}) + d_{h,\mathcal{X}_s}(p_{i_{2s}},2se_1). \end{aligned}$$

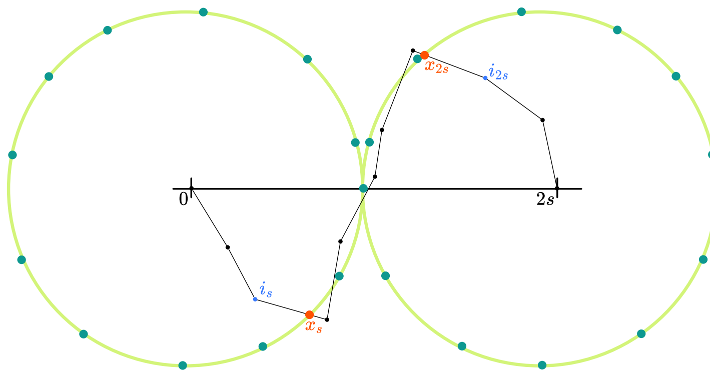

FIG. 3. *Construction in the proof of Proposition 5.1. The blue points on the spheres constitute the deterministic coverings x_i and x_i'.*

We choose a family of n_s points $\{x_i : 1 \le i \le n_s\} \subset \partial B(0,s)$ on the sphere $\partial B(0,s)$ such that $x_1 = se_1$ and the other points are distributed in such a way that for all $x \in \partial B(0,s)$ there exists a point x_i with $|x - x_i| \le h$. The construction of these points is straightforward: Given $\{x_1, \dots, x_k\}$ one chooses

$$x_{k+1} \in \partial B(0,s) \setminus \bigcup_{i=1}^{k} B(x_i, h/2).$$

Obviously, this process terminates after order $(s/h)^{d-1}$ iterations which means that

$$n_s \le C\left(\frac{s}{h}\right)^{d-1}. \tag{5.4}$$

Analogously one defines a covering of $\partial B(2se_1, s)$ into points $\{x_i' : 1 \le i \le n_s\}$ by reflecting the points x_i at the point se_1. Note that the value of n_s will turn out to be irrelevant, with the only important thing being that it is at most polynomially large in s.

For $i^* \in \{1, \dots, n_s\}$ chosen such that $|x_s - x_{i^*}| \le h$ it holds that

$$|p_{i_s} - x_{i^*}| \le |p_{i_s} - x_s| + |x_s - x_{i^*}| \le 2h.$$

It also holds that

$$d_{h,\mathcal{X}_s}(p_{i_s}, x_{i^*}) \le C_d |p_{i_s} - x_{i^*}|$$

which implies that

$$d_{h,\mathcal{X}_s}(0, x_{i^*}) \le d_{h,\mathcal{X}_s}(0, p_{i_s}) + d_{h,\mathcal{X}_s}(p_{i_s}, x_{i^*}) \le d_{h,\mathcal{X}_s}(0, p_{i_s}) + 2C_d h.$$

Analogously, for a suitable $i_* \in \{1, \dots, n_s\}$ one gets

$$d_{h,\mathcal{X}_s}(x_{i_*}', 2se_1) \le d_{h,\mathcal{X}_s}(p_{i_{2s}}, 2se_1) + 2C_d h.$$

Using these two inequalities together with (5.3) we obtain

$$\begin{aligned} d_{h,\mathcal{X}_s}(0, 2se_1) &\ge d_{h,\mathcal{X}_s}(0, p_{i_s}) + d_{h,\mathcal{X}_s}(p_{i_{2s}}, 2se_1) \ge d_{h,\mathcal{X}_s}(0, x_{i^*}) + d_{h,\mathcal{X}_s}(x_{i_*}', 2se_1) - 4C_d h \\ &\ge \min_{1 \le i \le n_s} d_{h,\mathcal{X}_s}(0, x_i) + \min_{1 \le i \le n_s} d_{h,\mathcal{X}_s}(x_i', 2se_1) - 4C_d h. \end{aligned}$$

Taking expectations and using Lemma 3.9 with $M = n_s$ as well as (5.4) we get

$$\begin{aligned}\mathbb{E}\big[d_{h,\mathcal{X}_s}(0, 2se_1)\big] &\geq 2\mathbb{E}\Big[\min_{1\leq i\leq n_s} d_{h,\mathcal{X}_s}(0, x_i)\Big] - 4C_d h \\ &\quad - C_1 \exp\biggl(-\biggl(\frac{h}{C_d}\biggr)^d + C_2 \log\biggl(\frac{s}{\delta_s}\biggr)\biggr) \\ &\geq 2\mathbb{E}\Big[\min_{1\leq i\leq n_s} d_{h,\mathcal{X}_s}(0, x_i)\Big] - C_1 h,\end{aligned}$$

where we used the assumption $\delta_s \leq h \leq s$ and $s > 1$ sufficiently large to simplify and absorb the rightmost term into the error term of order h. The constant C_1 changed its value several times. By adding two zeros and using that because of Lemma 3.9 with $M = 1$ it holds $|\mathbb{E}[d_{h,\mathcal{X}_s}(0, x_i)] - \mathbb{E}[d_{h,\mathcal{X}_s}(0, se_1)]| \leq C_1 h$ for all i, we can reorder this inequality in the following way:

$$\begin{aligned}\mathbb{E}\big[d_{h,\mathcal{X}_s}(0, 2se_1)\big] &\geq 2\mathbb{E}\big[d_{h,\mathcal{X}_s}(0, se_1)\big] \\ &\quad - 2\mathbb{E}\Big[\max_{1\leq i\leq n_s} \big(\mathbb{E}\big[d_{h,\mathcal{X}_s}(0, x_i)\big] - d_{h,\mathcal{X}_s}(0, x_i)\big)\Big] - C_1 h,\end{aligned}$$

where the constant C_1 again changed its value. We shall apply [16], Lemma B.2, to the random variables

$$Y_i^{(s)} := \frac{1}{\sqrt{\frac{\delta_s^2}{h} s}} d_{h,\mathcal{X}_s}(0, x_i), \quad 1 \leq i \leq n_s,$$

which satisfy

$$\mathbb{E}\big[Y_i^{(s)}\big] \leq \frac{C_d s}{\sqrt{\frac{\delta_s^2}{h} s}} \leq C_d \sqrt{\frac{h}{\delta_s^2}} \sqrt{s} \leq C s$$

for $s > 1$ with $s \geq h$ and some constant $C > 0$. Using also (5.4) and the concentration of measure from Theorem 4.1 we can apply [16], Lemma B.2, to get that

$$\mathbb{E}\Big[\max_{1\leq i\leq n_s} \big(\mathbb{E}\big[Y_i^{(s)}\big] - Y_i^{(s)}\big)\Big] \leq C_2 \log(s)$$

which translates to

$$\mathbb{E}\Big[\max_{1\leq i\leq n_s} \big(\mathbb{E}\big[d_{h,\mathcal{X}_s}(0, x_i)\big] - d_{h,\mathcal{X}_s}(0, x_i)\big)\Big] \leq C_2 \sqrt{\frac{\delta_s^2}{h} s}\, \log(s).$$

Hence, we obtain the desired inequality

$$\mathbb{E}\big[d_{h,\mathcal{X}_s}(0, 2se_1)\big] \geq 2\mathbb{E}\big[d_{h,\mathcal{X}_s}(0, se_1)\big] - C_1 h - C_2 \sqrt{\frac{\delta_s^2}{h} s}\, \log(s). \qquad \square$$

Similarly, one can prove near monotonicity of the function $s \mapsto \mathbb{E}[d_{h,\mathcal{X}_s}(0, se_1)]$. While we believe that this function should in fact be nondecreasing in s, the proof is not obvious. However, for our purposes the following approximate monotonicity statement is sufficient.

PROPOSITION 5.3 (Near monotonicity). *There exist dimensional constants $C_1, C_2 > 0$ such that for all $s > 1$ with $\delta_s \leq h \leq \frac{s}{C_d+2}$, and $0 \leq s' \leq s$ it holds*

$$\mathbb{E}\big[d_{h,\mathcal{X}_s}(0, s'e_1)\big] \leq \mathbb{E}\big[d_{h,\mathcal{X}_s}(0, se_1)\big] + C_1 h + C_2 \sqrt{\frac{\delta_s^2}{h} s}\, \log(s).$$

PROOF. We distinguish two cases, based on whether s' is smaller or larger than h.

Case 1, $s' \le h$: In this case we can perform trivial estimates:

$$\mathbb{E}\big[d_{h,\mathcal{X}_s}(0, s'e_1)\big] \le C_d h + h \le s - h + (C_d + 2)h - s \le s - h \le \mathbb{E}\big[d_{h,\mathcal{X}_s}(0, se_1)\big].$$

Case 2, $s' \ge h$: Using similar notation as in the proof of Proposition 5.1, we obtain

$$d_{h,\mathcal{X}_s}(0, se_1) \ge d_{h,\mathcal{X}_s}(0, p_{i_{s'}}) \ge d_{h,\mathcal{X}_s}(0, x_{i^*}) - 2C_d h \ge \min_{1\le i\le n_s} d_{h,\mathcal{X}_s}(0, x_i) - 2C_d h.$$

With the same arguments as in this previous proof and using that $|x_i| = s' \ge h$ we then obtain

$$\mathbb{E}\big[d_{h,\mathcal{X}_s}(0, se_1)\big] \ge \mathbb{E}\big[d_{h,\mathcal{X}_s}(0, s'e_1)\big] - C_1 h - C_2\sqrt{\frac{\delta_s^2}{h} s\log(s)},$$

where C_2 originates from an application of [16], Lemma B.2. Combining both cases completes the proof. □

5.2. *Ratio convergence rates.* We can use the previous superadditivity results to prove a convergence rate of the ratios of two distance functions:

$$\frac{\mathbb{E}[d_{h,\mathcal{X}_s}(0, se_1)]}{\mathbb{E}[d_{h,\mathcal{X}_s}(0, 2se_1)]} \to \frac{1}{2}, \quad s \to \infty.$$

Note that, in contrast to Proposition 3.12, the limiting constant σ does not appear in this ratio convergence.

PROPOSITION 5.4. *Under the conditions of Theorem 4.1 and Proposition 5.1 there exist dimensional constants $C_1, C_2 > 0$ such that it holds for all $s > 1$ sufficiently large with $\delta_s \le h \le s$ that*

$$\left|\frac{\mathbb{E}[d_{h,\mathcal{X}_s}(0, se_1)]}{\mathbb{E}[d_{h,\mathcal{X}_s}(0, 2se_1)]} - \frac{1}{2}\right| \le C_1\frac{h}{s} + C_2\sqrt{\frac{\delta_s^2}{h}}\frac{\log(s)}{\sqrt{s}}.$$

REMARK 5.5. In the case that $h = \delta_s$ we can again subsume the convergence rate into one term and for some dimensional constant $C_1 > 0$ we have for $s, t > 1$ with $s \ge h$ sufficiently large:

$$\left|\frac{\mathbb{E}[d_{h,\mathcal{X}_s}(0, se_1)]}{\mathbb{E}[d_{h,\mathcal{X}_s}(0, 2se_1)]} - \frac{1}{2}\right| \le C_1\frac{\log(s)^{\frac{1+2d}{2d}}}{\sqrt{s}}.$$

PROOF. The approximate triangle inequality from Lemma 3.10 implies that

$$d_{h,\mathcal{X}_s}(0, 2se_1) \le d_{h,\mathcal{X}_s}(0, se_1) + d_{h,\mathcal{X}_s}(se_1, 2se_1) + h,$$

where we remark that all three distance functions are defined on the same set of points $\mathcal{X}_s$. Taking expectations and using the approximate translation invariance from Lemma 3.9 with $M = 1$ yields

$$\mathbb{E}\big[d_{h,\mathcal{X}_s}(0, 2se_1)\big] \le 2\mathbb{E}\big[d_{h,\mathcal{X}_s}(0, se_1)\big] + C_1 h,$$

where we used the scaling assumption and (3.8) to estimate the error term by $C_1 h$. Using also Proposition 5.1 we get

$$\frac{1}{2} - \frac{C_1 h}{\mathbb{E}[d_{h,\mathcal{X}_s}(0, 2se_1)]} \le \frac{\mathbb{E}[d_{h,\mathcal{X}_s}(0, se_1)]}{\mathbb{E}[d_{h,\mathcal{X}_s}(0, 2se_1)]} \le \frac{\mathbb{E}[d_{h,\mathcal{X}_s}(0, se_1)]}{2\mathbb{E}[d_{h,\mathcal{X}_s}(0, se_1)] - C_1 h - C_2\sqrt{\frac{\delta_s^2}{h} s\log(s)}}$$

$$= \left(2 - C_1\frac{h}{\mathbb{E}[d_{h,\mathcal{X}_s}(0, se_1)]} - C_2\frac{\sqrt{\frac{\delta_s^2}{h} s\log(s)}}{\mathbb{E}[d_{h,\mathcal{X}_s}(0, se_1)]}\right)^{-1}$$

$$\leq \left(2 - C_1 \frac{h}{s-h} - C_2 \frac{\sqrt{\frac{\delta_s^2}{h} s} \log(s)}{s-h}\right)^{-1}$$

$$\leq \left(2 - C_1 \frac{h}{s(1-h/s)} - C_2 \frac{\sqrt{\frac{\delta_s^2}{h}} \log(s)}{\sqrt{s}(1-h/s)}\right)^{-1}.$$

For $s > 1$ sufficiently large we can assume that the two negative terms are smaller than $\frac{3}{2}$ and we can use the elementary inequality $\frac{1}{2-x} \leq \frac{1}{2} + x$ for $0 \leq x \leq \frac{3}{2}$ to obtain

$$\left|\frac{\mathbb{E}[d_{h,\mathcal{X}_s}(0, se_1)]}{\mathbb{E}[d_{h,\mathcal{X}_s}(0, 2se_1)]} - \frac{1}{2}\right| \leq C_1 \frac{h}{s} + C_2 \sqrt{\frac{\delta_s^2}{h}} \frac{\log(s)}{\sqrt{s}},$$

where we used that $2s - h \leq \mathbb{E}[d_{h,\mathcal{X}_s}(0, 2se_1)] \leq 2C_d s + h$ and increased the constants C_1, C_2 a little. □

6. Application to Lipschitz learning.

In this section we discuss an application of our results to the graph infinity Laplace equation which arises in the context of graph-based semisupervised learning. In particular, we will extend our previous results from [15] by proving uniform convergence rates for Lipschitz learning on graphs with bandwidths on the connectivity threshold. An alternative viewpoint of our results is that we prove that finite difference discretizations of the infinity Laplace equation on Poisson clouds converge at the percolation length scale. In particular, choosing large stencils—which is required for structured grids, see [54] but also our results in [15]—is *not* necessary for Poisson clouds.

For the readers' convenience we first translate the results of the present paper to Poisson processes with intensity $n \gg 1$ which is the natural setting when working on graphs in bounded domains.

6.1. *Rescaling to processes with higher intensity.* Let X_n be a Poisson point process with intensity n in $\mathbb{R}^d$. This means that

$$\mathbb{P}\big(\#(A \cap X_n) = k\big) = \frac{(n|A|)^k}{k!} e^{-n|A|} \quad \forall A \subset \mathbb{R}^d.$$

In expectation, the number of Poisson points in a set A equals $\mathbb{E}[\#(A \cap X_n)] = n|A|$. Given $x_0, x_1 \in \mathbb{R}^d$, we define the affine map

$$\Phi(x) := n^{\frac{1}{d}} R(x - x_0), \quad x \in \mathbb{R}^d,$$

where $R \in \mathbb{R}^{d\times d}$ is a suitable orthogonal matrix such that $\Phi(x_1) = n^{\frac{1}{d}}|x_1 - x_0|e_1$. Using the mapping theorem for Poisson point processes [51] we can connect the graph distance with step size $\varepsilon > 0$ on X_n with the graph distance on a unit intensity process, as studied in the previous sections. Defining the unit intensity Poisson point process $X := \Phi(X_n)$, the length and step size

$$s := n^{\frac{1}{d}}|x - x_0|, \tag{6.1}$$

$$h := n^{\frac{1}{d}}\varepsilon, \tag{6.2}$$

we have

$$d_{\varepsilon,X_n}(x_0, x) = n^{-\frac{1}{d}} d_{h,X}(0, se_1).$$

We also have a regularized version of the distance on X_n by defining

$$d'_\varepsilon(x_0, x) := n^{-\frac{1}{d}} d_{h,\mathcal{X}_s}(0, se_1),$$

where we suppress the dependency of the enriched Poisson process for a more compact notation. Note that for distances $|x-x_0|$ of order one the choice of $h=h_s \sim \log(s)^{\frac{1}{d}}$ translates to

$$\varepsilon = \frac{\log(n^{\frac{1}{d}}|x-x_0|)^{\frac{1}{d}}}{n^{\frac{1}{d}}} = \frac{(\frac{1}{d}\log(n)+\log|x-x_0|)^{\frac{1}{d}}}{n^{\frac{1}{d}}} \sim \left(\frac{\log(n)}{n}\right)^{\frac{1}{d}},$$

which is precisely the connectivity threshold for the graph X_n.

REMARK 6.1 (Change of notation). In what follows we will let ε denote the length scales used for the distance function on X_n. Furthermore, we will also suppress the dependency of the distance function on X_n and will simply write $d_\varepsilon(x_0, x)$.

Let us rephrase our previous results which are needed for the application to the graph infinity Laplacian in terms of the rescaled distance function. These are the localization results Lemmas 3.6 and 3.8, the concentration statement Theorem 4.1, the near monotonicity from Proposition 5.3, and the ratio convergence statement from Proposition 5.4.

THEOREM 6.2 (Properties of the distance function on X_n). *Let $x_0, x \in \mathbb{R}^d$ and assume*

$$K\left(\frac{\log n}{n}\right)^{\frac{1}{d}} \leq \varepsilon \leq |x-x_0|.$$

Then there exist dimensional constants $C_1, C_2 > 0$ which are independent of x_0 and x such that for $K>0$ sufficiently large:

1. (Concentration) *For all $\lambda > 0$ it holds*

$$\mathbb{P}\left(\left|d'_\varepsilon(x_0,x) - \mathbb{E}\left[d'_\varepsilon(x_0,x)\right]\right| > \lambda K\left(\frac{\log n}{n}\right)^{\frac{1}{d}}\sqrt{\frac{|x-x_0|}{\varepsilon}}\right) \leq C_1\exp(-C_2\lambda).$$

2. (Near monotonicity) *For n sufficiently large*, $x_0=0$, *and $x\in\mathbb{R}^d$ such that $(C_d+2)\varepsilon \leq |x| \leq 1$ it holds for all $x'\in\mathbb{R}^d$ with $|x'|\leq|x|$*:

$$\mathbb{E}\left[d'_\varepsilon(0,x')\right] \leq \mathbb{E}\left[d'_\varepsilon(0,x)\right] + C_1\varepsilon + C_2K\left(\frac{\log n}{n}\right)^{\frac{1}{d}}(\log n + \log|x|)\sqrt{\frac{|x|}{\varepsilon}}.$$

3. (Ratio convergence in expectation) *For n sufficiently large*, $x_0=0$, *and $x\in\mathbb{R}^d$ such that $\varepsilon\leq|x|$ it holds that*

$$\left|\frac{\mathbb{E}[d'_\varepsilon(0,x)]}{\mathbb{E}[d'_\varepsilon(0,2x)]} - \frac{1}{2}\right| \leq C_1\frac{\varepsilon}{|x|} + C_2K\left(\frac{\log n}{n}\right)^{\frac{1}{d}}\frac{\log n + \log|x|}{\sqrt{\varepsilon|x|}}.$$

4. (Localization) *For $|x-x_0|\geq 2\varepsilon$ it holds*

$$\begin{aligned}&\mathbb{P}\big(\text{any optimal path of } d_\varepsilon(x_0,x) \text{ lies in } B(x_0, C'_d|x_0-x|)\big)\\&\quad\geq 1-\exp\big(-C_1 n\varepsilon^d + C_2\log(n|x_0-x|)\big),\\&\mathbb{P}\big(d_\varepsilon(x_0,x) = d'_\varepsilon(x_0,x)\big) \geq 1 - 2\exp\big(-C_1 n\varepsilon^d + C_2\log(n|x_0-x|)\big).\end{aligned}$$

PROOF. One simply uses (6.1) and (6.2) and observes that $\delta_s = C_d(k\log(C''_d s))^{\frac{1}{d}} = K(\log n)^{\frac{1}{d}}$ for a suitable constant $K=K(d)$. □

6.2. *Convergence rates.* We still let X_n be a Poisson point process with intensity $n \in \mathbb{R}$ on $\mathbb{R}^d$ and let $\Omega \subset \mathbb{R}^d$ be an open and bounded domain. Remember that for a bandwidth parameter $\varepsilon > 0$ and a function $u_n : X_n \to \mathbb{R}$ we defined the graph infinity Laplacian of u_n as

$$\mathcal{L}_\infty^\varepsilon u_n(x) := \sup_{y \in B(x,\varepsilon)\cap X_n} \frac{u(y)-u(x)}{|y-x|} + \inf_{y \in B(x,\varepsilon)\cap X_n} \frac{u(y)-u(x)}{|y-x|}, \quad x \in X_n.$$

Solutions of the graph infinity Laplacian equation $\mathcal{L}_\infty^\varepsilon u_n = 0$ satisfy a special comparison principle with the graph distance function, called *comparison with cones*. To explain this, we introduce some terminology. For a subset $A \subset X_n$ we define its graph boundary and closure as

$$\begin{aligned} \mathrm{bd}_\varepsilon(A) &:= \{x \in X_n \setminus A : \exists y \in A, |x-y| \le \varepsilon\}, \\ \mathrm{cl}_\varepsilon(A) &:= A \cup \mathrm{bd}_\varepsilon(A). \end{aligned}$$

Furthermore, we refer to a subset $A \subset X_n$ as ε-connected if for all points $x, y \in A$ there exists a path in A which connects x and y and has hops of maximal size ε, in other words if $d_\varepsilon(x, y)[A] < \infty$.

We say that u_n satisfies comparison with cones on a subset $X_n' \subset X_n$ if for every subset $X_n'' \subset X_n'$, for all $a \ge 0$, and for all $z \in X_n' \setminus X_n''$ it holds

$$\max_{\mathrm{cl}_\varepsilon(X_n'')} \big(u_n - a d_\varepsilon(\cdot, z)\big) = \max_{\mathrm{bd}_\varepsilon(X_n'')} \big(u_n - a d_\varepsilon(\cdot, z)\big), \tag{6.3a}$$

$$\min_{\mathrm{cl}_\varepsilon(X_n'')} \big(u_n - a d_\varepsilon(\cdot, z)\big) = \min_{\mathrm{bd}_\varepsilon(X_n'')} \big(u_n - a d_\varepsilon(\cdot, z)\big). \tag{6.3b}$$

We have the following result:

THEOREM 6.3 ([15], Theorem 3.2). *Let $X_n' \subset X_n$ be an ε-connected subset of X_n and let $u_n : X_n' \to \mathbb{R}$ satisfy $\mathcal{L}_\infty^\varepsilon u_n(x) = 0$ for all $x \in X_n'$. Then u_n satisfies comparison with cones on X_n'.*

The goal of this section is to establish rates of convergence for solutions of $\mathcal{L}_\infty^\varepsilon u_n = 0$ to solutions of the infinity Laplacian equation $\Delta_\infty u = 0$, where $\Delta_\infty u := \sum_{i,j=1}^d \partial_i u \partial_j u \partial_{ij}^2 u$ for smooth functions u. Note that solutions to the infinity Laplacian equation are not C^2 in general which is why one typically uses the theory of viscosity solutions. However, solutions can be characterized through a comparison with cones property, as well. We refer to the seminal monograph [9] for this and other important properties of the infinity Laplacian equation.

For proving the rates we shall utilize the framework which we developed in [15] and which only relies on the comparison with cones property of the respective solutions. The novel idea there was the introduction of a homogenized length scale $\tau > \varepsilon$, a corresponding extensions u_n^τ of a graph solution u_n, and a homogenized infinity Laplacian operator Δ_∞^τ. The general recipe for getting rates as in [15] is the following:

1. Let $\mathcal{L}_\infty^\varepsilon u_n = 0$ and $\Delta_\infty u = 0$.
2. Use convergence of the distance function to prove that
$$-\Delta_\infty^\tau u_n^\tau \lesssim \mathrm{error}(n, \varepsilon, \tau) \quad \text{and} \quad \sup|u_n - u_n^\tau| \lesssim \tau.$$
3. Perturb the continuum solution u to a function $\tilde{u}$ which satisfies
$$-\Delta_\infty^\tau \tilde{u} \gtrsim \mathrm{error}(n, \varepsilon, \tau) \quad \text{and} \quad \sup|u - \tilde{u}| \lesssim \tau + \sqrt[3]{\mathrm{error}(n, \varepsilon, \tau)}.$$
4. Use a comparison principle for Δ_∞^τ and repeat the argument for $-u_n$ and $-u$ to get
$$\sup|u_n - u| \lesssim \tau + \sqrt[3]{\mathrm{error}(n, \varepsilon, \tau)}.$$

5. Optimize over n, ε, τ to get explicit rates.

Note that in [15] a careful analysis of boundary conditions and regularity is performed in order to be able to perform the arguments above all the way up to the boundary. Furthermore, the introduction of the homogenized operator allowed us to obtain convergence rates for arbitrary small graph bandwidths satisfying

$$\varepsilon \gg \left(\frac{\log n}{n}\right)^{\frac{1}{d}}.$$

The purpose of this section is to show how our results on Euclidean first-passage percolation allows to improve the error term $\operatorname{error}(n, \varepsilon, \tau)$ in order to allow for length scales of the form

$$\varepsilon \sim \left(\frac{\log n}{n}\right)^{\frac{1}{d}}.$$

Let us now introduce the homogenized quantities. For $\tau > 0$ we define extensions of the discrete function $u_n : X_n \to \mathbb{R}$ to functions $u_n^\tau, (u_n)_\tau : \mathbb{R}^d \to \mathbb{R}$ as follows:

$$u_n^\tau(x) := \sup_{B(x,\tau)\cap X_n} u_n, \quad x \in \mathbb{R}^d, \tag{6.4a}$$

$$(u_n)_\tau(x) := \inf_{B(x,\tau)\cap X_n} u_n, \quad x \in \mathbb{R}^d. \tag{6.4b}$$

Note that both extrema are attained if $B(x, \tau) \cap X_n \neq \varnothing$ since this set is of finite cardinality. We also define the nonlocal infinity Laplacian with respect to $\tau > 0$ of a function $u : \mathbb{R}^d \to \mathbb{R}$ as

$$\Delta_\infty^\tau u(x) := \frac{1}{\tau^2}\Big(\sup_{B(x,\tau)} u - 2u(x) + \inf_{B(x,\tau)} u\Big), \quad x \in \mathbb{R}^d. \tag{6.5}$$

Last, for a positive number $r > 0$ we define inner parallel sets of Ω as

$$\Omega^r := \{x \in \Omega : \operatorname{dist}(x, \partial\Omega) > r\}. \tag{6.6}$$

THEOREM 6.4. *Let $\Omega \subset \mathbb{R}^d$ be an open and bounded domain and X_n be a Poisson point process on $\mathbb{R}^d$ with density $n \in \mathbb{N}$. Assume that $\varepsilon > 0$ and $\tau > 0$ satisfy*

$$K\left(\frac{\log n}{n}\right)^{\frac{1}{d}} \le \varepsilon \le \frac{1}{K}\tau, \quad 0 < \tau < 1, \tag{6.7}$$

and define

$$\mathcal{O}_n := \{x \in X_n \cap \overline{\Omega} : \operatorname{dist}(x, \partial\Omega) \le \varepsilon\}.$$

Let $g : \overline{\Omega} \to \mathbb{R}$ be a Lipschitz function and $u_n : X_n \to \mathbb{R}$ solve

$$\begin{cases} \mathcal{L}_\infty^\varepsilon u_n = 0 & \text{on } X_n \setminus \mathcal{O}_n, \\ u_n = g & \text{on } \mathcal{O}_n. \end{cases}$$

Then there exist dimensional constants $C_1, C_2, C_3, C_4, C_5 > 0$ and $C_6 > 1$ such that for all $\lambda \ge 0$ and for $K \ge 8$ sufficiently large with probability at least

$$1 - C_1 \exp(-C_2 K^d \log n) - C_3 \exp(-C_4\lambda + C_5 \log n)$$

it holds for all $x_0 \in \Omega^{2C_6\tau}$ that

$$-\Delta_\infty^\tau u_n^\tau(x_0) \lesssim \operatorname{Lip}(g)\left((\log n + \lambda)\left(\frac{\log n}{n}\right)^{\frac{1}{d}} \frac{1}{\sqrt{\tau^3\varepsilon}} + \frac{\varepsilon}{\tau^2}\right), \tag{6.8a}$$

$$-\Delta_\infty^\tau (u_n)_\tau(x_0) \gtrsim -\operatorname{Lip}(g)\left((\log n + \lambda)\left(\frac{\log n}{n}\right)^{\frac{1}{d}} \frac{1}{\sqrt{\tau^3\varepsilon}} + \frac{\varepsilon}{\tau^2}\right). \tag{6.8b}$$

REMARK 6.5. Abbreviating $\delta_n := (\frac{\log n}{n})^{\frac{1}{d}}$ our result translates to

$$-\Delta_\infty^\tau u_n^\tau(x_0) \lesssim \mathrm{Lip}(g)\left((\log n + \lambda)\frac{\delta_n}{\sqrt{\tau^3 \varepsilon}} + \frac{\varepsilon}{\tau^2}\right).$$

In particular, we can choose $\varepsilon = \delta_n$ and the error term reduces to $\sqrt{\frac{\delta_n}{\tau^3}}$ which goes to zero if τ is sufficiently large compared to δ_n. In our previous work [15], Theorem 5.13, we proved an analogous result for arbitrary weighted graphs (whose vertices could also be deterministic) with connectivity radius δ_n, graph bandwidth ε, and a free parameter τ. There we proved that

$$-\Delta_\infty^\tau u_n^\tau(x_0) \lesssim \mathrm{Lip}(g)\left(\frac{\delta_n}{\tau\varepsilon} + \frac{\varepsilon}{\tau^2}\right)$$

and one observes that choosing $\varepsilon = \delta_n$ is not possible since then the right hand side would diverge as $\tau \to 0$.

PROOF. The proof follows very closely our earlier result [15], Theorem 5.13, but involves nontrivial adaptations.

It suffices to prove the first statement since the second one follows by changing the signs of u_n. Furthermore, it suffices to prove the statement for graph vertices $x_0 \in X_n$ and then use [15], Lemma 5.8, to extend it to continuum points, which does only incur error terms that are already present and increases the constant C_6.

Let us fix $x_0 \in \Omega^{2C_6\tau}$ where for now we assume that $C_6 > 1$. Utilizing that

$$\sup_{B(x_0,\tau)} u_n^\tau = \sup_{x\in B(x_0,\tau)} \sup_{B(x,\tau)\cap X_n} u_n = \sup_{B(x_0,2\tau)} u_n = u_n^{2\tau}(x_0),$$
$$\inf_{B(x_0,\tau)} u_n^\tau = \inf_{x\in B(x_0,\tau)} \sup_{B(x,\tau)\cap X_n} u_n \geq u_n(x_0),$$

we obtain

$$-\tau^2\Delta_\infty^\tau u_n^\tau(x_0) \leq 2u_n^\tau(x_0) - u_n^{2\tau}(x_0) - u_n(x_0). \tag{6.9}$$

To estimate this term, we turn our attention to the function u_n and the fact that it satisfies comparison with cones. For this we define the set $\mathsf{B}_n(x_0, 2\tau) \subset X_n$ as

$$\mathsf{B}_n(x_0, 2\tau) := \left\{x \in X_n \setminus \{x_0\} : d_\varepsilon(x_0, x) \leq \inf_{y\in B(x_0,2\tau-\varepsilon)^c} d_\varepsilon(x_0, y) - \varepsilon\right\}. \tag{6.10}$$

We start by recording a couple of properties of the set $\mathsf{B}_n(x_0, 2\tau)$:

First, we observe that

$$\mathsf{B}_n(x_0, 2\tau) \subset B(x_0, 2\tau - \varepsilon) \tag{6.11}$$

since otherwise there would be a point $x \in \mathsf{B}_n(x_0, 2\tau)$ such that $d_\varepsilon(x_0, x) \leq d_\varepsilon(x_0, x) - \varepsilon$ which is a contradiction.

Second, we claim that

$$\inf_{y\in B(x_0,2\tau-\varepsilon)^c} d_\varepsilon(x_0, y) = \inf_{y\in B(x_0,2\tau)\setminus B(x_0,2\tau-\varepsilon)} d_\varepsilon(x_0, y) \tag{6.12}$$

which is going to be relevant a little later. To see this, note that the left hand side is always smaller or equal than the right hand side. Furthermore, any feasible path from a point $y \in B(x_0, 2\tau - \varepsilon)^c$ to x_0 has to contain a point in $B(x_0, 2\tau)$ and can hence be truncated to obtain a feasible path for the right side.

Third, we claim that the (graph) boundary of $\mathsf{B}_n(x_0, 2\tau)$ satisfies

$$\operatorname{bd}_\varepsilon(\mathsf{B}_n(x_0, 2\tau)) \subset \Big\{x \in X_n \cap B(x_0, 2\tau) : \\ d_\varepsilon(x_0, x) > \inf_{y \in B(x_0, 2\tau-\varepsilon)^c} d_\varepsilon(x_0, y) - \varepsilon\Big\} \cup \{x_0\} =: \mathsf{B}' \tag{6.13}$$

and in particular $\mathsf{B}' \subset \Omega$. By definition, for $z \in \operatorname{bd}_\varepsilon(\mathsf{B}_n(x_0, 2\tau))$ there exists $x \in \mathsf{B}_n(x_0, 2\tau)$ with $|x - z| \leq \varepsilon$ and hence, using also (6.11), we get

$$|z - x_0| \leq |z - x| + |x_0 - x| \leq \varepsilon + 2\tau - \varepsilon = 2\tau,$$

which proves (6.13). In particular, we see by (6.7) and (6.13) that for $C_6 > 1$ sufficiently large it holds $\mathsf{B}' \cap \mathcal{O}_n = \varnothing$. We have the following trivial inequality:

$$u_n(x) \leq u_n(x_0) + \big(u_n^{2\tau}(x_0) - u_n(x_0)\big) \frac{d_\varepsilon(x_0, x)}{\inf_{y \in B(x_0, 2\tau-\varepsilon)^c} d_\varepsilon(x_0, y) - \varepsilon} \quad \forall x \in \mathsf{B}'.$$

Indeed, if $x = x_0$ the inequality is in fact an equality, and for all $x \in B(x_0, 2\tau)$ it is also true since $u_n^{2\tau}(x_0) \geq u_n(x)$. Consequently, since $\operatorname{bd}_\varepsilon(\mathsf{B}_n(x_0, 2\tau)) \subset \mathsf{B}' \subset (X_n \cap \overline{\Omega}) \setminus \mathcal{O}_n$ and u_n satisfies comparison with cones on this set, we infer that for all $x \in \operatorname{cl}_\varepsilon(\mathsf{B}_n(x_0, 2\tau))$ it holds

$$u_n(x) \leq u_n(x_0) + \big(u_n^{2\tau}(x_0) - u_n(x_0)\big) \frac{d_\varepsilon(x_0, x)}{\inf_{y \in B(x_0, 2\tau-\varepsilon)^c} d_\varepsilon(x_0, y) - \varepsilon}. \tag{6.14}$$

Without loss of generality we can assume that $C_d \leq 3/2$ (otherwise, one can increase K in the definition of ε, see Remark 3.5). Using Lemma 3.2 this ensures that for all $x \in B(x_0, \tau)$ we have with probability at least $1 - C_1 \exp(-C_2 K^d \log n)$ for some constants $C_1, C_2 > 0$ that

$$d_\varepsilon(x_0, x) \leq C_d \tau + \varepsilon \leq \frac{3}{2}\tau + \varepsilon = 2\tau - 3\varepsilon + 4\varepsilon - \frac{1}{2}\tau. \tag{6.15}$$

On the other hand, using Lemma 3.1 we also have

$$\inf_{y \in B(x_0, 2\tau-\varepsilon)^c} d_\varepsilon(x_0, y) - \varepsilon \geq \inf_{y \in B(x_0, 2\tau-\varepsilon)^c} |x_0 - y| - \operatorname{dist}(y, \mathcal{X}_s) - \varepsilon \\ \geq 2\tau - 3\varepsilon. \tag{6.16}$$

Since $\tau \geq K\varepsilon \geq 8\varepsilon$ we infer from (6.15) and (6.16) that

$$d_\varepsilon(x_0, x) \leq \inf_{y \in B(x_0, 2\tau-\varepsilon)^c} d_\varepsilon(x_0, y) - \varepsilon$$

and this implies $B(x_0, \tau) \subset \mathsf{B}_n(x_0, 2\tau)$. Consequently, we can maximize both sides in (6.14) over $x \in B(x_0, \tau) \cap X_n$ to get

$$\begin{aligned} u_n^\tau(x_0) &\leq u_n(x_0) + \big(u_n^{2\tau}(x_0) - u_n(x_0)\big) \frac{\sup_{x \in B(x_0, \tau) \cap X_n} d_\varepsilon(x_0, x)}{\inf_{y \in B(x_0, 2\tau-\varepsilon)^c} d_\varepsilon(x_0, y) - \varepsilon} \\ &\leq u_n(x_0) + \big(u_n^{2\tau}(x_0) - u_n(x_0)\big) \frac{\sup_{x \in B(x_0, \tau) \cap X_n} d_\varepsilon(x_0, x)}{\inf_{y \in B(x_0, 2\tau-\varepsilon)^c} d_\varepsilon(x_0, y)} \\ &\quad \times \left(1 + \frac{\varepsilon}{\inf_{y \in B(x_0, 2\tau-\varepsilon)^c} d_\varepsilon(x_0, y)}\right). \end{aligned}$$

In the last step we used the elementary inequality $\frac{1}{1-t} \leq 1 + 2t$ for $0 \leq t \leq 1/2$. Now we argue that we can replace d_ε by d_ε' in this expression with high probability: First, we finally use property (6.12) from above which tells us that the infimum $\inf_{y \in B(x_0, 2\tau-\varepsilon)^c} d_\varepsilon(x_0, y)$ can be restricted to the annulus $B(x_0, 2\tau) \setminus B(x_0, 2\tau - \varepsilon)$. Hence Item 4 in Theorem 6.2 implies that $d_\varepsilon(x_0, y) = d'_\varepsilon(x_0, y)$ for all $y \in B(x_0, 2\tau) \setminus B(x_0, 2\tau - \varepsilon)$ with probability at least $1 - C_1 \exp(-C_2 K^d \log n)$ where we possibly increase C_1 and C_2.

Second, we argue for the supremum. Possibly increasing C_1 and C_2 with probability at least $1 - C_1 \exp(-C_2 K^d \log n)$ it is finite and let us assume it is attained at a point $\hat{x} \in B(x_0, 2\varepsilon) \cap X_n$. Then using Lemma 3.2 with the same probability we have

$$2C_d\varepsilon \geq d_\cap(x_0, \hat{x}) = \sup_{x \in B(x_0,\tau)\cap X_n} d_\varepsilon(x_0, x) \geq d_\varepsilon(x_0, \tilde{x}) \geq (K-1)\varepsilon$$

for every point $\tilde{x} \in (B(x_0, K\varepsilon) \cap X_n) \setminus B(x_0, (K-1)\varepsilon)$. Note that if K is sufficiently large then such a point exists with the same probability.

This is a contradiction if $K > 2C_d + 1$ and so Item 4 in Theorem 6.2 again lets us replace $d_\varepsilon(x_0, x)$ by $d'_\varepsilon(x_0, x)$ for all $x \in B(x_0, \tau) \cap X_n$. Hence, we obtain

$$u_n^\tau(x_0) \leq u_n(x_0) + \big(u_n^{2\tau}(x_0) - u_n(x_0)\big) \frac{\sup_{x \in B(x_0,\tau)\cap X_n} d'_\varepsilon(x_0, x)}{\inf_{y \in B(x_0, 2\tau-\varepsilon)^c} d'_\varepsilon(x_0, y)} \times \left(1 + \frac{\varepsilon}{\inf_{y \in B(x_0,2\tau-\varepsilon)^c} d'_\varepsilon(x_0, y)}\right)$$

with probability at least $1 - C_1 \exp(-C_2 K^d \log n)$. Introducing the shortcut notation

$$\overline{d}_\tau(x_0) := \sup_{x \in B(x_0,\tau)\cap X_n} d'_\varepsilon(x_0, x), \tag{6.17a}$$

$$\underline{d}_{2\tau}(x_0) := \inf_{y \in B(x_0,2\tau-\varepsilon)^c} d'_\varepsilon(x_0, y), \tag{6.17b}$$

$$r_\tau(x_0) := \frac{\overline{d}_\tau(x_0)}{\underline{d}_{2\tau}(x_0)} - \frac{1}{2}, \tag{6.17c}$$

we can rewrite and continue the previous estimate as follows:

$$\begin{aligned}
u_n^\tau(x_0) &\leq u_n(x_0) + \big(u_n^{2\tau}(x_0) - u_n(x_0)\big)\left(r_\tau(x_0) + \frac{1}{2}\right)\left(1 + \frac{\varepsilon}{\inf_{y \in B(x_0,2\tau-\varepsilon)^c} d'_\varepsilon(x_0, y)}\right) \\
&\leq \big(u_n^{2\tau}(x_0) - u_n(x_0)\big) r_\tau(x_0) + \frac{1}{2}\big(u_n(x_0) + u_n^{2\tau}(x_0)\big) \\
&\quad + \big(u_n^{2\tau}(x_0) - u_n(x_0)\big)\left(r_\tau(x_0) + \frac{1}{2}\right)\frac{\varepsilon}{\inf_{y \in B(x_0,2\tau-\varepsilon)^c} d'_\varepsilon(x_0, y)}.
\end{aligned}$$

Returning to (6.9) we obtain

$$\begin{aligned}
-\tau^2 \Delta_\infty^\tau u_n^\tau(x_0) &\leq 2\big(u_n^{2\tau}(x_0) - u_n(x_0)\big) r_\tau(x_0) \\
&\quad + 2\big(u_n^{2\tau}(x_0) - u_n(x_0)\big)\left(r_\tau(x_0) + \frac{1}{2}\right)\frac{\varepsilon}{\inf_{y \in B(x_0,2\tau-\varepsilon)^c} d'_\varepsilon(x_0, y)} \\
&\leq 2 \operatorname{Lip}_n(u_n) d_\varepsilon\big(x_0, x_0^*\big) r_\tau(x_0) \\
&\quad + 2 \operatorname{Lip}_n(u_n) d_\varepsilon\big(x_0, x_0^*\big)\left(r_\tau(x_0) + \frac{1}{2}\right)\frac{\varepsilon}{\inf_{y \in B(x_0,2\tau-\varepsilon)^c} d'_\varepsilon(x_0, y)},
\end{aligned}$$

where we let $x_0^* \in B(x_0, 2\tau) \cap X_n$ be a point which realizes $u_n^{2\tau}(x_0)$ and define the graph Lipschitz constant

$$\operatorname{Lip}_n(u_n) := \max_{x, y \in X_n} \frac{|u_n(x) - u_n(y)|}{d_\varepsilon(x, y)}. \tag{6.18}$$

Since u_n solves the graph infinity Laplace equation it holds $\operatorname{Lip}_n(u_n) = \operatorname{Lip}_n(g)$ by [15], Proposition 3.8, and using Lemma 3.1 we get

$$\operatorname{Lip}_n(g) = \max_{x,y\in X_n} \frac{|g(x)-g(y)|}{d_\varepsilon(x,y)} \le \max_{x,y\in X_n} \frac{|g(x)-g(y)|}{|x-y|} \le \operatorname{Lip}(g).$$

We have the estimates $d_\varepsilon(x_0, x_0^*) \le 2C_d\tau$ with high probability and $\inf_{y\in B(x_0,2\tau-\varepsilon)^c} d'_\varepsilon(x_0, y) \ge 2\tau - 2\varepsilon$ which imply

$$\begin{aligned}
-\tau^2\Delta_\infty^\tau u_n^\tau(x_0) &\le 2C_d \operatorname{Lip}(g)\tau\, r_\tau(x_0) \\
&\quad + 4C_d^2 \operatorname{Lip}(g)\tau \frac{\varepsilon}{\inf_{y\in B(x_0,2\tau-\varepsilon)^c} d'_\varepsilon(x_0,y)}\left(r_\tau(x_0) + \frac{1}{2}\right) \\
&\lesssim \operatorname{Lip}(g)\left(\tau\, r_\tau(x_0) + \frac{\tau\varepsilon}{\tau-\varepsilon}\right) \\
&\lesssim \operatorname{Lip}(g)\big(\tau\, r_\tau(x_0) + \varepsilon\big).
\end{aligned}$$

In the second inequality we used trivial estimates on τ, ε, and $r_\tau(x_0)$ to absorb the second term into the first one, and we absorbed dimensional constants into the $\lesssim$ symbol. In the third inequality we used that $\tau \ge 8\varepsilon$ to simplify $\frac{\tau\varepsilon}{\tau-\varepsilon} = \varepsilon\frac{1}{1-\frac{\varepsilon}{\tau}} \le \frac{8}{7}\varepsilon \lesssim \varepsilon$. Dividing by τ^2 we obtain

$$-\Delta_\infty^\tau u_n^\tau(x_0) \lesssim \operatorname{Lip}(g)\left(\frac{r_\tau(x_0)}{\tau} + \frac{\varepsilon}{\tau^2}\right).$$

By [16], Lemma C.3, and a union bound there exist constants $C_3, C_4, C_5 > 0$ such that for all $\lambda \ge 0$ with probability at least $1 - C_3\exp(-C_4\lambda + \log(\tau/\varepsilon) + C_5\log n)$ it holds

$$r_\tau(x_0) \lesssim (\log n + \lambda)\left(\frac{\log n}{n}\right)^{\frac{1}{d}} \frac{\log n}{\sqrt{\tau\varepsilon}}.$$

Plugging this in we obtain

$$-\Delta_\infty^\tau u_n^\tau(x_0) \lesssim \operatorname{Lip}(g)\left((\log n + \lambda)\left(\frac{\log n}{n}\right)^{\frac{1}{d}} \frac{1}{\sqrt{\tau^3\varepsilon}} + \frac{\varepsilon}{\tau^2}\right).$$

We conclude the proof, noting that the last probability can be simplified using (6.7):

$$\begin{aligned}
\log(\tau/\varepsilon) + C_5\log n &\le \log\tau - \log\varepsilon + C_5\log n \\
&\le -\log K - (1/d)\log\log n + (1/d)\log n + C_5\log n \le C_5\log n
\end{aligned}$$

by changing the value of $C_5 > 0$ and choosing $K \ge 1$ and $n \ge 3$. Hence the last probability can be simplified to $1 - C_3\exp(-C_4\lambda + C_5\log n)$ and the final result is establish with another union bound. □

The proof of Theorem 2.3 is now identical to the one presented in our previous paper with the essential ingredient being Theorem 6.4.

PROOF SKETCH OF THEOREM 2.3. The proof works as in [15], Section 5.3.3, replacing ε there with $C_6\tau$. For completeness we sketch the proof below.

From Theorem 6.4 we obtain

$$-\Delta_\infty^\tau u_n^\tau \le C\operatorname{Lip}(g)\left((\log n + \lambda)\left(\frac{\log n}{n}\right)^{\frac{1}{d}} \frac{1}{\sqrt{\tau^3\varepsilon}} + \frac{\varepsilon}{\tau^2}\right) =: C_{n,\tau} \quad \text{in } \Omega^{2C_6\tau},$$

for some constant $C > 0$. The proof strategy is to perturb u to a strict supersolution associated to the operator $-\Delta_\infty^\tau$. For this we use [15], Lemma 4.8, Lemma 4.9, as in the proof of [15], Proposition 5.16, which allows us to choose $w : \Omega^{2C_6\tau} \to \mathbb{R}$ such that

$$-\Delta_\infty^\tau w \geq C_{n,\tau} \quad \text{in } \Omega^{2C_6\tau}, \qquad \|w - (u)_\tau\|_{L^\infty(\Omega^{2C_6\tau})} \lesssim \sqrt[3]{C_{n,\tau}}.$$

Since we now have $-\Delta_\infty^\tau u_n^\tau \leq C_{n,\tau} \leq -\Delta_\infty^\tau w$ we can invoke the comparison principle for the operator $-\Delta_\infty^\tau$, see [7], Corollary 3.3, to obtain that

$$\begin{aligned}\sup_{\Omega^{(2C_6-1)\tau}} \big(u_n^\tau - (u)_\tau\big) &\lesssim \sup_{\Omega^{(2C_6-1)\tau}} \big(u_n^\tau - w\big) + \sqrt[3]{C_{n,\tau}} = \sup_{\Omega^{(2C_6-1)\tau} \setminus \Omega^{2C_6\tau}} \big(u_n^\tau - w\big) + \sqrt[3]{C_{n,\tau}} \\ &\lesssim \sup_{\Omega^{(2C_6-1)\tau} \setminus \Omega^{2C_6\tau}} \big(u_n^\tau - (u)_\tau\big) + 2\sqrt[3]{C_{n,\tau}},\end{aligned}$$

where we also used the triangle inequality twice. Analogously, we obtain

$$\sup_{\Omega^{(2C_6-1)\tau}} \big(u^\tau - (u_n)_\tau\big) \lesssim \sup_{\Omega^{(2C_6-1)\tau} \setminus \Omega^{2C_6\tau}} \big(u^\tau - (u_n)_\tau\big) + 2\sqrt[3]{C_{n,\tau}}.$$

The next steps consists in getting rid of the extension operators at the scale of τ, for which we employ (approximate) Lipschitzness of u (and u_n). Utilizing [15], Lemma 5.9, Lemma 5.10, Lemma 5.11, this can be done at the cost of an additive error of order τ, for which we obtain

$$\sup_{X_n \cap \Omega^{(2C_6-1)\tau}} |u - u_n| \lesssim \tau + \sqrt[3]{C_{n,\tau}}.$$

Finally, we extend this result to $X_n \cap \overline{\Omega}$ using again Lipschitzness of u and the data g. Namely take $x \in X_n \cap \Omega$ and $\tilde{x} \in X_n \cap (\Omega \setminus \Omega^{(2C_6-1)\tau})$ such that $|x - \tilde{x}| \lesssim \tau$ which yields

$$\begin{aligned}\big|u(x) - u_n(x)\big| &\leq \big|u(x) - u(\tilde{x})\big| + \big|u(\tilde{x}) - u_n(\tilde{x})\big| + \big|u_n(\tilde{x}) - u_n(x)\big| \\ &\lesssim \operatorname{Lip}(g)\tau + \tau + \sqrt[3]{C_{n,\tau}},\end{aligned}$$

where we used that u_n satisfies an approximate Lipschitz estimate of the form

$$\big|u_n(x) - u_n(y)\big| \leq \operatorname{Lip}_n(u_n) d_\varepsilon(x, y) \lesssim \operatorname{Lip}_n(g)\big(|x - y| + \varepsilon\big) \lesssim \operatorname{Lip}(g)\tau.$$

Hence, we have showed

$$\sup_{X_n \cap \Omega} |u - u_n| \lesssim \operatorname{Lip}(g)\tau + \sqrt[3]{C_{n,\tau}}$$

which concludes the proof sketch. □

Funding. Part of this work was also done while LB and TR were in residence at Institut Mittag–Leffler in Djursholm, Sweden during the semester on *Geometric Aspects of Nonlinear Partial Differential Equations* in 2022, supported by the Swedish Research Council under grant no. 2016-06596. LB also acknowledges funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy—GZ 2047/1, Projekt-ID 390685813. Most of this study was carried out while LB was affiliated with the Hausdorff Center for Mathematics at the University of Bonn. JC acknowledges funding from NSF Grant DMS-1944925, the Alfred P. Sloan foundation, and a McKnight Presidential Fellowship. TR acknowledges support from DESY (Hamburg, Germany), a member of the Helmholtz Association HGF, by the German Ministry of Science and Technology (BMBF) under grant agreement No. 05M2020 (DELETO) and the European Unions Horizon 2020 research and innovation programme under the Marie Skłodowska–Curie grant agreement No. 777826 (NoMADS). Most of this study was carried out while TR was affiliated with the Friedrich–Alexander-Universität Erlangen–Nürnberg.

SUPPLEMENTARY MATERIAL

Appendix (DOI: 10.1214/24-AAP2052SUPPA; .pdf). The appendix collects important statements regarding (approximately) sub- and superadditive functions, an abstract concentration inequality for martingale difference sequences, some auxiliary estimates, and numerical illustrations.

Code for numerical examples (DOI: 10.1214/24-AAP2052SUPPB; .zip). The code for the numerical examples is provided in the supplementary material in the file `PercolationConvergenceRates.zip`. It can also be found at https://github.com/TimRoith/PercolationConvergenceRates.

REFERENCES

[1] ALEXANDER, K. S. (1993). A note on some rates of convergence in first-passage percolation. *Ann. Appl. Probab.* **3** 81–90. MR1202516

[2] ALEXANDER, K. S. (2011). Subgaussian rates of convergence of means in directed first passage percolation.

[3] ARMSTRONG, S. and CARDALIAGUET, P. (2018). Stochastic homogenization of quasilinear Hamilton–Jacobi equations and geometric motions. *J. Eur. Math. Soc.* (*JEMS*) **20** 797–864. MR3779686 https://doi.org/10.4171/JEMS/777

[4] ARMSTRONG, S. and DARIO, P. (2018). Elliptic regularity and quantitative homogenization on percolation clusters. *Comm. Pure Appl. Math.* **71** 1717–1849. MR3847767 https://doi.org/10.1002/cpa.21726

[5] ARMSTRONG, S., KUUSI, T. and MOURRAT, J.-C. (2017). The additive structure of elliptic homogenization. *Invent. Math.* **208** 999–1154. MR3648977 https://doi.org/10.1007/s00222-016-0702-4

[6] ARMSTRONG, S., KUUSI, T. and MOURRAT, J.-C. (2019). *Quantitative Stochastic Homogenization and Large-Scale Regularity*. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **352**. Springer, Cham. MR3932093 https://doi.org/10.1007/978-3-030-15545-2

[7] ARMSTRONG, S. N. and SMART, C. K. (2012). A finite difference approach to the infinity Laplace equation and tug-of-war games. *Trans. Amer. Math. Soc.* **364** 595–636. MR2846345 https://doi.org/10.1090/S0002-9947-2011-05289-X

[8] ARMSTRONG, S. N. and SMART, C. K. (2016). Quantitative stochastic homogenization of convex integral functionals. *Ann. Sci. Éc. Norm. Supér.* (4) **49** 423–481. MR3481355 https://doi.org/10.24033/asens.2287

[9] ARONSSON, G., CRANDALL, M. G. and JUUTINEN, P. (2004). A tour of the theory of absolutely minimizing functions. *Bull. Amer. Math. Soc.* (*N.S.*) **41** 439–505. MR2083637 https://doi.org/10.1090/S0273-0979-04-01035-3

[10] AUFFINGER, A., DAMRON, M. and HANSON, J. (2017). 50 *Years of First-Passage Percolation*. *University Lecture Series* **68**. Amer. Math. Soc., Providence, RI. MR3729447 https://doi.org/10.1090/ulect/068

[11] BARLES, G. and SOUGANIDIS, P. E. (1991). Convergence of approximation schemes for fully nonlinear second order equations. *Asymptot. Anal.* **4** 271–283. MR1115933

[12] BOLLOBÁS, B. and BRIGHTWELL, G. (1992). The height of a random partial order: Concentration of measure. *Ann. Appl. Probab.* **2** 1009–1018. MR1189428

[13] BRAIDES, A. and CAROCCIA, M. (2023). Asymptotic behavior of the Dirichlet energy on Poisson point clouds. *J. Nonlinear Sci.* **33** Paper No. 80, 57 pp. MR4617151 https://doi.org/10.1007/s00332-023-09937-7

[14] BROADBENT, S. R. and HAMMERSLEY, J. M. (1957). Percolation processes. I. Crystals and mazes. *Proc. Camb. Philos. Soc.* **53** 629–641. MR0091567 https://doi.org/10.1017/s0305004100032680

[15] BUNGERT, L., CALDER, J. and ROITH, T. (2023). Uniform convergence rates for Lipschitz learning on graphs. *IMA J. Numer. Anal.* **43** 2445–2495. MR4621850 https://doi.org/10.1093/imanum/drac048

[16] BUNGERT, L., CALDER, J. and ROITH, T. (2024). Supplement to “Ratio convergence rates for Euclidean first-passage percolation: Applications to the graph infinity Laplacian.” https://doi.org/10.1214/24-AAP2052SUPPA, https://doi.org/10.1214/24-AAP2052SUPPB

[17] CALDER, J. (2019). The game theoretic p-Laplacian and semi-supervised learning with few labels. *Nonlinearity* **32** 301–330. MR3893728 https://doi.org/10.1088/1361-6544/aae949

[18] CALDER, J. (2019). Consistency of Lipschitz learning with infinite unlabeled data and finite labeled data. *SIAM J. Math. Data Sci.* **1** 780–812. MR4039189 https://doi.org/10.1137/18M1199241

[19] CALDER, J. (2020). The calculus of variations.

[20] CALDER, J., COOK, B., THORPE, M. and SLEPČEV, D. (2020). Poisson learning: Graph based semi-supervised learning at very low label rates. In *Proceedings of the* 37*th International Conference on Machine Learning* **119** 1306–1316. PMLR .

[21] CALDER, J. and ETTEHAD, M. (2022). Hamilton–Jacobi equations on graphs with applications to semi-supervised learning and data depth. *J. Mach. Learn. Res.* **23** Paper No. [318], 62 pp. MR4577757

[22] CALDER, J. and GARCÍA TRILLOS, N. (2022). Improved spectral convergence rates for graph Laplacians on ε-graphs and k-NN graphs. *Appl. Comput. Harmon. Anal.* **60** 123–175. MR4393800 https://doi.org/10.1016/j.acha.2022.02.004

[23] CALDER, J., GARCÍA TRILLOS, N. and LEWICKA, M. (2022). Lipschitz regularity of graph Laplacians on random data clouds. *SIAM J. Math. Anal.* **54** 1169–1222. MR4384039 https://doi.org/10.1137/20M1356610

[24] CALDER, J. and SLEPČEV, D. (2020). Properly-weighted graph Laplacian for semi-supervised learning. *Appl. Math. Optim.* **82** 1111–1159. MR4167693 https://doi.org/10.1007/s00245-019-09637-3

[25] CALDER, J., SLEPČEV, D. and THORPE, M. (2023). Rates of convergence for Laplacian semi-supervised learning with low labeling rates. *Res. Math. Sci.* **10** Paper No. 10, 42 pp. MR4548608 https://doi.org/10.1007/s40687-022-00371-x

[26] CALDER, J. and SMART, C. K. (2020). The limit shape of convex hull peeling. *Duke Math. J.* **169** 2079–2124. MR4132581 https://doi.org/10.1215/00127094-2020-0013

[27] CAROCCIA, M. (2023). A compactness theorem for functions on Poisson point clouds. *Nonlinear Anal.* **231** Paper No. 113032, 18 pp. MR4575699 https://doi.org/10.1016/j.na.2022.113032

[28] COOK, B. and CALDER, J. (2022). Rates of convergence for the continuum limit of nondominated sorting. *SIAM J. Math. Anal.* **54** 872–911. MR4376298 https://doi.org/10.1137/20M1344901

[29] COX, J. T. (1980). The time constant of first-passage percolation on the square lattice. *Adv. in Appl. Probab.* **12** 864–879. MR0588407 https://doi.org/10.2307/1426745

[30] COX, J. T. and DURRETT, R. (1981). Some limit theorems for percolation processes with necessary and sufficient conditions. *Ann. Probab.* **9** 583–603. MR0624685

[31] COX, J. T. and KESTEN, H. (1981). On the continuity of the time constant of first-passage percolation. *J. Appl. Probab.* **18** 809–819. MR0633228 https://doi.org/10.1017/s0021900200034161

[32] DARIO, P. and GU, C. (2021). Quantitative homogenization of the parabolic and elliptic Green's functions on percolation clusters. *Ann. Probab.* **49** 556–636. MR4255127 https://doi.org/10.1214/20-aop1456

[33] DEL TESO, F. and LINDGREN, E. (2022). A finite difference method for the variational p-Laplacian. *J. Sci. Comput.* **90** Paper No. 67, 31 pp. MR4358715 https://doi.org/10.1007/s10915-021-01745-z

[34] DEL TESO, F., MANFREDI, J. J. and PARVIAINEN, M. (2022). Convergence of dynamic programming principles for the p-Laplacian. *Adv. Calc. Var.* **15** 191–212. MR4399821 https://doi.org/10.1515/acv-2019-0043

[35] DÍAZ, J., MITSCHE, D., PERARNAU, G. and PÉREZ-GIMÉNEZ, X. (2016). On the relation between graph distance and Euclidean distance in random geometric graphs. *Adv. in Appl. Probab.* **48** 848–864. MR3568895 https://doi.org/10.1017/apr.2016.31

[36] DUNBAR, O. R., ELLIOTT, C. M. and KREUSSER, L. M. (2022). Models for information propagation on graphs. Preprint. Available at arXiv:2201.07577.

[37] E, W., LI, T. and VANDEN-EIJNDEN, E. (2019). *Applied Stochastic Analysis*. *Graduate Studies in Mathematics* **199**. Amer. Math. Soc., Providence, RI. MR3932086 https://doi.org/10.1090/gsm/199

[38] FADILI, J., FORCADEL, N., NGUYEN, T. T. and ZANTOUT, R. (2023). Limits and consistency of nonlocal and graph approximations to the Eikonal equation. *IMA J. Numer. Anal.* **43** 3685–3728. MR4673672 https://doi.org/10.1093/imanum/drac082

[39] FRIEDRICH, T., SAUERWALD, T. and STAUFFER, A. (2013). Diameter and broadcast time of random geometric graphs in arbitrary dimensions. *Algorithmica* **67** 65–88. MR3072817 https://doi.org/10.1007/s00453-012-9710-y

[40] GARCÍA TRILLOS, N., SLEPČEV, D., VON BRECHT, J., LAURENT, T. and BRESSON, X. (2016). Consistency of Cheeger and ratio graph cuts. *J. Mach. Learn. Res.* **17** Paper No. 181, 46 pp. MR3567449

[41] GROISMAN, P., JONCKHEERE, M. and SAPIENZA, F. (2022). Nonhomogeneous Euclidean first-passage percolation and distance learning. *Bernoulli* **28** 255–276. MR4337705 https://doi.org/10.3150/21-bej1341

[42] HAFIENE, Y., FADILI, J. M. and ELMOATAZ, A. (2019). Continuum limits of nonlocal p-Laplacian variational problems on graphs. *SIAM J. Imaging Sci.* **12** 1772–1807. MR4025770 https://doi.org/10.1137/18M1223927

[43] HAMMERSLEY, J. M. and WELSH, D. J. A. (1965). First-passage percolation, subadditive processes, stochastic networks, and generalized renewal theory. In *Proc. Internat. Res. Semin., Statist. Lab., Univ. California, Berkeley, Calif.*, 1963 61–110. Springer, New York. MR0198576

[44] HIRSCH, C., NEUHÄUSER, D., GLOAGUEN, C. and SCHMIDT, V. (2015). First passage percolation on random geometric graphs and an application to shortest-path trees. *Adv. in Appl. Probab.* **47** 328–354. MR3360380 https://doi.org/10.1239/aap/1435236978

[45] HOWARD, C. D. and NEWMAN, C. M. (1997). Euclidean models of first-passage percolation. *Probab. Theory Related Fields* **108** 153–170. MR1452554 https://doi.org/10.1007/s004400050105

[46] HOWARD, C. D. and NEWMAN, C. M. (2001). Geodesics and spanning trees for Euclidean first-passage percolation. *Ann. Probab.* **29** 577–623. MR1849171 https://doi.org/10.1214/aop/1008956685

[47] HWANG, S. J., DAMELIN, S. B. and HERO, A. O. III (2016). Shortest path through random points. *Ann. Appl. Probab.* **26** 2791–2823. MR3563194 https://doi.org/10.1214/15-AAP1162

[48] KESTEN, H. (1981). Percolation theory for mathematicians. *Nieuw Arch. Wiskd.* (3) **29** 227–239. MR0643930

[49] KESTEN, H. (1986). Aspects of first passage percolation. In *École D'été de Probabilités de Saint-Flour, XIV—1984. Lecture Notes in Math.* **1180** 125–264. Springer, Berlin. MR0876084 https://doi.org/10.1007/BFb0074919

[50] KESTEN, H. (1993). On the speed of convergence in first-passage percolation. *Ann. Appl. Probab.* **3** 296–338. MR1221154

[51] KINGMAN, J. F. C. (1993). *Poisson Processes. Oxford Studies in Probability* **3**. The Clarendon Press, New York. MR1207584

[52] LI, W. and SALGADO, A. J. (2022). Convergent, with rates, methods for normalized infinity Laplace, and related, equations. Preprint. Available at arXiv:2209.06109.

[53] LITTLE, A., MCKENZIE, D. and MURPHY, J. M. (2022). Balancing geometry and density: Path distances on high-dimensional data. *SIAM J. Math. Data Sci.* **4** 72–99. MR4368991 https://doi.org/10.1137/20M1386657

[54] OBERMAN, A. M. (2005). A convergent difference scheme for the infinity Laplacian: Construction of absolutely minimizing Lipschitz extensions. *Math. Comp.* **74** 1217–1230. MR2137000 https://doi.org/10.1090/S0025-5718-04-01688-6

[55] OBERMAN, A. M. (2006). Convergent difference schemes for degenerate elliptic and parabolic equations: Hamilton–Jacobi equations and free boundary problems. *SIAM J. Numer. Anal.* **44** 879–895. MR2218974 https://doi.org/10.1137/S0036142903435235

[56] OBERMAN, A. M. (2013). Finite difference methods for the infinity Laplace and p-Laplace equations. *J. Comput. Appl. Math.* **254** 65–80. MR3061067 https://doi.org/10.1016/j.cam.2012.11.023

[57] PIMENTEL, L. P. R. (2011). Asymptotics for first-passage times on Delaunay triangulations. *Combin. Probab. Comput.* **20** 435–453. MR2784636 https://doi.org/10.1017/S0963548310000477

[58] ROITH, T. and BUNGERT, L. (2023). Continuum limit of Lipschitz learning on graphs. *Found. Comput. Math.* **23** 393–431. MR4568195 https://doi.org/10.1007/s10208-022-09557-9

[59] SERAFINI, H. C. (1997). First-passage percolation on the Delaunay graph of a d-dimensional Poisson process. New York University.

[60] SLEPČEV, D. and THORPE, M. (2019). Analysis of p-Laplacian regularization in semisupervised learning. *SIAM J. Math. Anal.* **51** 2085–2120. MR3953458 https://doi.org/10.1137/17M115222X

[61] SMART, C. K. (2010). On the infinity Laplacian and Hrushovski's fusion. PhD thesis, UC Berkeley. MR2941545

[62] SMYTHE, R. T. and WIERMAN, J. C. (2006). *First-Passage Percolation on the Square Lattice. Lecture Notes in Math.* **671**. Springer, Berlin. MR0513421

[63] YAO, C.-L., CHEN, G. and GUO, T.-D. (2011). Large deviations for the graph distance in supercritical continuum percolation. *J. Appl. Probab.* **48** 154–172. MR2809893 https://doi.org/10.1239/jap/1300198142

[64] ZHU, X., GHAHRAMANI, Z. and LAFFERTY, J. D. (2003). Semi-supervised learning using Gaussian fields and harmonic functions. In *Proceedings of the* 20*th International Conference on Machine Learning* (*ICML*-03) 912–919.