Bernoulli **26**(2), 2020, 927–947
https://doi.org/10.3150/19-BEJ1121

Distances and large deviations in the spatial preferential attachment model

CHRISTIAN HIRSCH[1] and CHRISTIAN MÖNCH[2]

[1] *Institut für Mathematik, Universität Mannheim, B6, 26, 68161 Mannheim, Germany.*
E-mail: hirsch@uni-mannheim.de
[2] *Institut für Mathematik, Johannes Gutenberg-Universität, Staudingerweg 9, 55128 Mainz, Germany.*
E-mail: cmoench@uni-mainz.de

This paper considers two asymptotic properties of a spatial preferential-attachment model introduced by E. Jacob and P. Mörters (In *Algorithms and Models for the Web Graph* (2013) 14–25 Springer). First, in a regime of strong linear reinforcement, we show that typical distances are at most of doubly-logarithmic order. Second, we derive a large deviation principle for the empirical neighbourhood structure and express the rate function as solution to an entropy minimisation problem in the space of stationary marked point processes.

Keywords: distances; large deviation principle; Poisson point process; preferential attachment

1. Introduction

In the present paper, we investigate typical distances and large deviation principles (LDPs) in the *spatial preferential attachment models (S-PAMs)* model introduced in [11]. Firstly, we verify a conjecture in [13], Remark 3, about typical distances in the largest connected component of the S-PAM. As made precise in Theorem 2.1 below, the asymptotic behaviour depends both on the strength of the preferential attachment and on the influence of vertex distances on the connection probability. This is in contrast to the situation for the asymptotic degree distribution, where only the strength of the preferential attachment but not the geometry affects the power-law exponent [12], Remark 1.

Secondly, we establish an LDP for the in-degree evolution and the evolution of the neighbourhood structure of a typical vertex. The most fundamental building blocks for this result are the process-level LDP of a marked Poisson point process [9] and the contraction principle. The rate function is expressed via a constraint minimisation problem of the specific relative entropy. Loosely speaking, large deviations from the neighbourhood structure are induced by stationary modifications of the original Poisson point process of vertices. Asymptotically, the configurations in a rare event minimise the specific-entropy costs of these modifications subject to the constraint of leading to the considered rare event. Hence, our approach offers a novel complementing perspective to previously used martingale techniques [2,4].

The rest of the paper is organised as follows. In Section 2, we provide precise definitions of the S-PAM and state our two main results. Sections 3 and 4 contain the proofs for the distance asymptotics and the large deviation principle. Finally, highly technical proofs and auxiliary results are available in a short supplementary document that is available on the journal webpage.

1350-7265

2. Model definition and main results

2.1. Definition

We consider the S-PAM from [11–13]. More precisely, the torus $\mathbb{T}_n = [-n^{1/d}/2, n^{1/d}/2]^d/\sim$ of side length $n > 0$ in dimension $d \geq 1$ is the ambient space of the model. For the construction of the S-PAM, we view the process of network nodes as a space-time process of points arriving sequentially in $\mathbb{T}_n$. More precisely, they form a homogeneous Poisson point process $X = X_n$ on $\mathbb{T}_n \times [0, 1]$ with intensity 1. Formally, a point $(x, s) \in X$ is a *vertex* at *position* $x \in \mathbb{T}_n$ and *birth time* $s \in [0, 1]$. For brevity, we often write just $x \in X$ to denote the a.s. unique vertex (x, s) in position $x \in \mathbb{T}_n$.

We identify X with the vertex set of a random geometric graph $G_n = (X, E)$ obtained by the following distance-dependent preferential attachment mechanism.

The model is parametrised by an affine function $f : \mathbb{Z}_{\geq 0} \to (0, \infty)$, $z \mapsto \gamma z + \gamma'$ inducing the PA mechanism, where $\gamma \in (0, 1)$, $\gamma' > 0$, and a decreasing profile function $\varphi : [0, \infty) \to [0, 1]$ incorporating the spatial effects. We assume power decay of the profile function in the sense that $\varphi(x) = \min\{\kappa x^{-\delta}, 1\}$ for some $\delta > 1$ with normalising constant κ chosen such that $\int_0^\infty \varphi(x)\,\mathrm{d}x = 1/2$. With these settings, the edges in the S-PAM are obtained as follows, where by a slight abuse of notation we write $|\cdot - \cdot|$ for the Euclidean distance on the torus. Initially, the edge set is empty. Whenever a new vertex $(y, t) \in X$ is born, it connects to each vertex $(x, s) \in X$ with $s < t$ independently with probability

$$\varphi\left(\frac{t|x - y|^d}{f(Z_x(t-))}\right), \tag{1}$$

where $Z_x(t-)$ denotes the *in-degree* of (x, s) at time $t-$, that is, the number of connections it has already received from vertices born during (s, t). These dynamics give rise to an increasing process $(G_n(t))_{t\in[0,1]}$ of geometric graphs. We usually write G_n for $G_n(1)$.

2.2. Typical distances

One of the central findings in [13] is that the S-PAM undergoes a phase transition: if the attachment function f increases quickly and the profile function φ decreases slowly, then the connected component C_n of the oldest vertex in $G_n = G_n(1)$ grows linearly in n. Moreover, this component is *robust* under site percolation in the sense that it remains of linear size even after any nontrivial i.i.d. Bernoulli thinning of the vertices.

By [13], Theorem 1, and [12], Theorem 7, for $\gamma > \delta/(1 + \delta)$ we have

$$\mathbb{P} - \lim_{n\to\infty} \frac{\#C_n}{n} = \theta \in (0, 1), \tag{2}$$

where $\mathbb{P} - \lim$ is shorthand for limit in probability. In this regime, it is conjectured [13], Remark 3, that typical distances are of doubly-logarithmic order. We verify this conjecture here. The graph distance in G_n is denoted by $\mathsf{dist}_n(\cdot, \cdot)$.

Theorem 2.1 (Distances for $\gamma > \delta/(1+\delta)$). *Let Y, Y' be uniformly chosen vertices of C_n. Then, with high probability as $n \to \infty$,*

$$\mathsf{dist}_n\big(Y, Y'\big) \le \big(4 + o(1)\big)\frac{\log\log n}{\log \frac{\gamma}{\delta(1-\gamma)}}.$$

By checking the error terms in the proofs of Propositions 3.2 and 3.3 below, one can see that Theorem 2.1 extends to the slightly more general choices of attachment and profile function studied in [13]. Thus, the following extension of Theorem 2.1 holds true.

Corollary 2.2. *Let the preferential attachment function f satisfy $\lim_{k\to\infty} \frac{f(k)}{k} = \gamma \in (0, 1)$ and the profile function φ $\varphi(x) = L(x)x^{-\delta}$, for some $\delta \in (1, \infty)$ and some slowly varying function $L : (0, \infty) \to (0, \infty)$, then the conclusion of Theorem 2.1 remains valid.*

It is an intriguing open problem to complement Theorem 2.1 with a matching lower bound. An important step in this direction would be to derive a corresponding result in the more tractable age-dependent random connection model as suggested in [10].

2.3. Large deviations principle

As our second main result, we derive an LDP for the evolution of the neighbourhood structure in the S-PAM as defined in Section 2.1, in the vein of [1]. This complements results for combinatorial sparse graphs discussed in [1] by a class of random graphs with an underlying geometry. In particular, as a corollary we obtain an LDP for the evolution of the empirical in-degree distribution over time, similar to the setting in [2].

Next, we introduce notation related to local convergence of graphs. We let $\mathcal{G}^*$ denote the family of *rooted graphs*, that is, of locally finite and connected graphs with a distinguished vertex. We write g_h for the subgraph of a rooted graph $g \in \mathcal{G}^*$ obtained as the union of all paths in g connecting to the root in at most $h \ge 0$ hops. We equip $\mathcal{G}^*$ with the *local topology*, the topology generated by the functions $\mathsf{ev}_{h,g'} : g \mapsto 1\{g_h \simeq g'_h\}$, where $h \in \mathbb{Z}_{\ge 0}$ and $g' \in \mathcal{G}^*$. Then, writing $[G_n(t), x]$ for the spatial PAM at time t with distinguished vertex $x \in X$, the *evolution of the empirical neighbourhood structure*

$$L_n^{\mathsf{neighb}}(\cdot) = \frac{1}{n}\sum_{x\in X} \delta_{[G_n(\cdot),x]} \tag{3}$$

defines a random variable in the product space $\mathcal{M}(\mathcal{G}^*)^{[0,1]}$, where $\mathcal{M}(\mathcal{G}^*)$ is the family of finite measures on $\mathcal{G}^*$ endowed with the vague topology. In other words, $\mathcal{M}(\mathcal{G}^*)^{[0,1]}$ carries the smallest topology such that for each $t \in [0, 1]$, $h \in \mathbb{Z}_{\ge 0}$ and rooted graph $g \in \mathcal{G}^*$ the evaluation maps $\mathsf{ev}_{t,h,g} : \mathcal{M}(\mathcal{G}^*)^{[0,1]} \to [0, \infty)$, $\nu \mapsto \nu_t(g_h)$ are continuous. Also note that since we work in a Poisson setting, in (3) we normalise by the window size n rather than the random number of vertices.

By the contraction principle [3], Theorem 4.2.10, the LDP for L_n^{neighb} becomes a consequence of the LDP for marked Poisson point processes [9], Theorem 3.1. Hence, we introduce common notation in this setting. First, to realise the independent connections with the probability described in (1), we proceed as in the random connection model [8] and introduce a family of auxiliary random variables. More precisely, we augment each vertex $(x, s) \in X$ independently with a collection of iid random variables $\{V_{x,y}\}_{y\in X}$ such that each $V_{x,y} \sim \mathbf{U}([0, 1])$ is uniformly distributed on $[0, 1]$. In other words, X becomes an $[0, 1]^{\mathbb{Z}_{\geq 0}}$-marked Poisson point process. As a new vertex $(y, t) \in X$ arrives, it connects to $(x, s) \in X$ if and only if $s < t$ and $V_{x,y}$ is smaller than the threshold given in (1).

In the limit $n \to \infty$ the torus $\mathbb{T}_n$ approaches $\mathbb{R}^d$. Therefore, the limiting objects appearing in the LDP live in $\mathcal{P}_\theta$, the space of all distributions of stationary $[0, 1]^{\mathbb{Z}_{\geq 0}}$-marked point processes on $\mathbb{R}^d$ endowed with the $\tau_{\mathcal{L}}$-topology of local convergence. This topology is generated by the evaluations $\mathsf{ev}_f : \mathcal{P}_\theta \to [0, \infty)$, $\mathbb{Q} \mapsto \int_{\mathcal{C}} f(\psi)\mathbb{Q}(\mathrm{d}\psi)$, where $\mathcal{C}$ is the space of configurations in the space $\mathbb{R}^d \times [0, 1]^{\mathbb{Z}_{\geq 0}}$ that are locally finite in the first component and f is any nonnegative measurable function depending only on the configuration in a bounded domain [9]. Additionally, $\mathbb{Q}^*$ denotes the unnormalised *Palm version* of a stationary point process $\mathbb{Q} \in \mathcal{P}_\theta$ [14], Section 9. That is, $\mathbb{Q}^*$ is determined by the disintegration identity

$$\int_{\mathcal{C}} f(\psi)\mathbb{Q}^*(\mathrm{d}\psi) = \int_{\mathcal{C}} \int_{[0,1]^d} f(\theta_x \psi)\psi(\mathrm{d}x)\mathbb{Q}(\mathrm{d}\psi),$$

where $\theta_x : \mathbb{R}^d \to \mathbb{R}^d$, $y \mapsto y - x$ denotes the shift by $x \in \mathbb{R}^d$. Then, for each time $t \in [0, 1]$, considering the neighbourhood structure at the origin $o \in \mathbb{R}^d$ under the Palm measure $\mathbb{Q}^*$ yields an element in $\mathcal{M}(\mathcal{G}^*)$. Hence, letting t vary, we associate to $\mathbb{Q} \in \mathcal{P}_\theta$ the evolution of the neighbourhood structure $\mathbb{Q}^{*,\mathsf{neighb}} \in \mathcal{M}(\mathcal{G}^*)^{[0,1]}$.

Finally, the rate function in the LDP is expressed in terms of the *specific relative entropy* of stationary marked point processes. More precisely, writing $\restriction_{[-n/2,n/2]^d}$ for the restriction to the box $[-n/2, n/2]^d$, for $\mathbb{Q} \in \mathcal{P}_\theta$ we put

$$H(\mathbb{Q}) = \lim_{n\to\infty} n^{-d} \int \log\left(\frac{\mathrm{d}\mathbb{Q}\restriction_{[-n/2,n/2]^d}}{\mathrm{d}\mathsf{Pois}\restriction_{[-n/2,n/2]^d}}(\psi)\right)\mathbb{Q}\restriction_{[-n/2,n/2]^d}(\mathrm{d}\psi),$$

tacitly applying the convention that $H(\mathbb{Q}) = \infty$ if the Radon–Nikodym derivative of the restricted point processes does not exist. By means of the contraction principle, the LDP for marked Poisson point processes [9], Theorem 3.1, now gives rise to the LDP for the evolution of the empirical neighbourhood structure.

Theorem 2.3. *The empirical neighbourhood structure* $\{L_n^{\mathsf{neighb}}\}_{n\geq 1}$ *satisfies the LDP in the product space* $\mathcal{M}(\mathcal{G}^*)^{[0,1]}$ *with good rate function* $\nu \mapsto \inf_{\mathbb{Q}\in\mathcal{P}_\theta, \mathbb{Q}^{*,\mathsf{neighb}}=\nu} H(\mathbb{Q})$.

Since the degree of the root in a rooted graph is nothing more than the size of the 1-neighbourhood, after another application of the contraction principle, Theorem 2.3 yields an

LDP for the *evolution of empirical in-degrees*

$$L_n^{\mathsf{deg}}(\cdot) = \frac{1}{n}\sum_{k\geq 0} \#\big\{x \in X : Z_x(\cdot) = k\big\}\delta_k. \tag{4}$$

Corollary 2.4. *The empirical in-degree evolution* $\{L_n^{\mathsf{deg}}\}_{n\geq 1}$ *satisfies the LDP in the product space* $\mathcal{M}(\mathbb{Z}_{\geq 0})^{[0,1]}$ *with good rate function* $\nu \mapsto \inf_{\mathbb{Q}\in\mathcal{P}_\theta, \mathbb{Q}^{*,\mathsf{deg}}=\nu} H(\mathbb{Q})$.

Since $\mathcal{M}(\mathbb{Z}_{\geq 0})^{[0,1]}$ carries the product topology, Corollary 2.4 only provides access to crude information on the time evolution of the empirical degree distributions. For this reason, we next deduce a more refined LDP based on the Skorohod topology [7]. Since this topology requires an underlying metric space, we consider only the setting of a priori bounded in-degrees. More precisely, for $k \geq 0$ let

$$L_n^{\mathsf{deg};\leq k}(\cdot) = \Big(\frac{1}{n}\#\big\{x \in X : Z_x(\cdot) = 0\big\}, \dots, \frac{1}{n}\#\big\{x \in X : Z_x(\cdot) = k\big\}\Big)$$

denote the evolution of the vector containing the normalised in-degree evolutions truncated at the kth in-degree. Then, $L_n^{\mathsf{deg};\leq k}$ is a random element of the Skorohod space $\mathbb{D}_{k+1}$ of functions $f : [0, 1] \to [0, \infty)^{k+1}$ that are càdlàg in each coordinate.

Corollary 2.5. *For every* $k \geq 0$ *the truncated empirical in-degree evolution* $\{L_n^{\mathsf{deg}}\}_{n\geq 1}$ *satisfies the LDP in the Skorohod topology with good rate function*

$$\mathbf{f} = \big(f_0(\cdot), \dots, f_k(\cdot)\big) \mapsto \inf_{\substack{\mathbb{Q}\in\mathcal{P}_\theta \\ \mathbb{Q}^{*,\deg;\leq k}=\mathbf{f}}} H(\mathbb{Q}).$$

3. Proof of Theorem 2.1

We prove Theorem 2.1 in several steps. First, in Section 3.1 we explain the overall idea and give the proof subject to intermediate results. Then, Section 3.2 introduces sprinkling and monotonicity as central tools for the arguments in the subsequent sections. Finally, Sections 3.3 and 3.4 contain the proofs of the intermediate results.

3.1. The main argument

To establish the upper bound on typical distances, we first show that almost all vertices in the giant component are within bounded distance of a fairly old vertex of high degree. Then, we proceed to show that each such high-degree node is at distance at most $2(\rho + o(1)) \log\log n$ of the oldest vertex in G_n with high probability, where

$$\rho = \frac{1}{\log(\gamma/(\delta(1-\gamma))}. \tag{5}$$

In essence, this argument is already outlined in [13], cf. Remark 3, and can be traced in the proofs of Propositions 13 and 15 therein. However, the arguments given in [13] to establish the existence of a giant component only require the oldest vertex to connect to sufficiently many lower degree vertices. To show this, only a bounded number of search steps are necessary. To prove Theorem 2.1 along similar lines, we analyse the probabilities of adverse events occurring during the search for connecting vertices more thoroughly than is required for the robustness results in [13]. To keep the different stages of our search algorithm sufficiently independent, we rely on a sprinkling construction in the vein of [13]. More precisely, for some small $r > 0$ we colour each vertex in X independently *red* with probability r and *black* with probability $b = 1 - r$. Then, G_n^r and G_n^b denote the S-PAMs constructed on the red and black vertices, respectively. The reasoning behind this will be explained in Section 3.2.

Let us make the overall argument precise. To start the construction, we need to find an old black vertex near a uniformly chosen vertex $Y \in C_n$. By stationarity, we may assume that $Y = (o, U)$ is located at the origin $o \in \mathbb{R}^d$ with U uniform in [0, 1] and consider the Poisson process X under the corresponding Palm distribution $\mathbb{P}_{(o,U)}$. A vertex $(x, s) \in G_n^b$ is *D-reachable* if it connects to (o, U) by a path in G_n^b in at most D hops. For ease of reference, we introduce the events

$$E_n^b(D, s) = \big\{\text{some vertex } Y_0 \in G_n^b \text{ born before time } s \text{ is } D\text{-reachable}\big\}.$$

If there are several reachable vertices, Y_0 denotes the one with minimal birth time.

Proposition 3.1 (Connection to good vertices). *Let $b, s > 0$. Then, there exists an almost surely finite random variable $D = D^b(s)$*

$$\lim_{n\to\infty} \mathbb{P}_{(o,U)}\big(\{(o, U) \in C_n^b\} \setminus E_n^b\big(D^b(s), s\big)\big) = 0,$$

where C_n^b denotes the connected component of the oldest vertex in G_n^b.

Having reached a sufficiently old black vertex, we proceed as in the above heuristic argument. For the rest of this section, $g : (0, \infty) \to (0, \infty)$ denotes the sub-polynomially growing function introduced in Lemma A.2, which is parametrised by γ, δ and r only. Extending a notion from [13], we say a vertex $(x, s) \in G_n$ is *r-good* if $s < 1/2$ and it has at least $s^{-\gamma}/g(s^{-1})$ red neighbours with birth times in $(s, 1/2)$. It is *locally r-good* if it remains r-good after removing all edges of the form $y \to x$ with $y \notin [x - s^{-1/d}, x + s^{-1/d}]^d$. Loosely speaking, exploring possible paths along good vertices ensures a sufficient number of outgoing connections to choose from. Additionally, local goodness allows us to scan X for good vertices while keeping the explored areas sufficiently localised to leverage on the spatial independence of Poisson points.

We build up a hierarchical connection path along r-good vertices of increasing age joined by young red vertices born after time 1/2. Writing $\mathsf{rgood}_n \subset X$ for the subset of all red r-good vertices, we introduce a hierarchy of layers

$$L_1^r \subset L_2^r \subset \cdots \subset \mathsf{rgood}_n$$

of red r-good vertices, parametrised by their age. The first layer L_1^r contains the vertices of highest degree, i.e. near n^γ. With increasing index, the layers $\{L_i^r\}_{i\geq 1}$ contain more and more vertices of lower and lower degrees. More precisely, in the robust regime $\gamma > \delta/(1+\delta)$, we can fix global parameters

$$\alpha \in \left(1, \frac{\gamma}{\delta(1-\gamma)}\right), \qquad \beta \in \left(\alpha, \frac{\gamma}{\delta} + \alpha\gamma\right), \tag{6}$$

and then set

$$L_k^r = \left\{(x, s) \in \mathsf{rgood}_n : s \le n^{-\alpha^{-k}}\right\}$$

and

$$K = \min\left\{k \ge 1 : n^{-\alpha^{-k}} \ge (\log n)^{-\nu^{-1}}\right\} - 2,$$

where

$$\nu = \min\left\{-\beta\delta + \gamma - \alpha\gamma\delta, \frac{\beta - \alpha}{d}\right\} > 0.$$

Starting from an old r-good vertex, we typically reach L_K^r in at most $C(\alpha, \beta, r) \log\log\log n$ steps, where $C(\alpha, \beta, r)$ is a sufficiently large constant. In particular, in Section 3.4, we explicitly specify the scheme for establishing these connections. For the moment, assume that $C(\alpha, \beta, r)$ is given and call an r-good vertex at distance at most $C(\alpha, \beta, r) \log\log\log n$ from L_K^r *well-connected*.

Proposition 3.2 (Well-connectedness). *Let $b > 0$. Then,*

$$\lim_{s\to 0} \liminf_{n\to\infty} \mathbb{E}\left[\mathbb{P}_{Y_0}\left(Y_0 \text{ is well-connected} \mid G_n^b\right) 1_{E_n^b(D^b(s),s)}\right] = 1.$$

Having established a path from (o, U) to L_K^r, we now bound the diameter of L_K^r.

Proposition 3.3 (Final layer diameter). *With high probability, we have that*

$$\mathsf{diam}_n\left(L_K^r\right) \le 4K.$$

Combining Propositions 3.1–3.3, we now complete the proof of Theorem 2.1.

Proof of Theorem 2.1. Since K is of order $(1 + o(1)) \log\log n / \log\alpha$, by Proposition 3.3 it suffices to show that a uniformly chosen $Y \in C_n$ connects to L_K^r in at most $o(\log\log n)$ hops with high probability. In particular, this occurs under the event $E \cap F$ where $E = E_n^b(D^b(s), s)$ and $F = \{Y_0 \text{ is well-connected}\}$. In other words, it suffices to show that

$$\lim_{b\to 1} \liminf_{s\to 0} \liminf_{n\to\infty} \mathbb{P}_{(o,U)}\left(\left\{(o, U) \in C_n^b\right\} \cap E \cap F\right) = \theta.$$

To achieve this goal, decompose the left-hand side as

$$\begin{aligned} &\mathbb{P}_{(o,U)}\left(\left\{(o, U) \in C_n^b\right\} \cap E \cap F\right) \\ &\quad \ge \mathbb{P}_{(o,U)}\left((o, U) \in C_n^b\right) - \mathbb{P}_{(o,U)}\left(\left\{(o, U) \in C_n^b\right\} \setminus E\right) - \mathbb{P}_{(o,U)}(E \setminus F). \end{aligned} \tag{7}$$

By Proposition 3.1, the second summand tends to 0 as $n \to \infty$. Since E is measurable with respect to G_n^b, the third contribution equals

$$\mathbb{P}_{(o,U)}(E \setminus F) = \mathbb{E}_{(o,U)}\big[\big(1 - \mathbb{P}_{Y_0}\big(F \mid G_n^b\big)\big)1\{E\}\big],$$

which tends to 0 as $s \to 0$ and $n \to \infty$ by Proposition 3.2. Hence, (7) gives that

$$\liminf_{s\to 0}\liminf_{n\to\infty}\mathbb{P}_{(o,U)}\big(\big\{(o,U) \in C_n^b\big\} \cap E \cap F\big) \geq \theta^b.$$

Here θ^b denotes the asymptotic proportion of vertices in the giant component of G_n^b, which by continuity of the percolation probability [13], Proposition 7, tends to θ as $b \to 1$. $\square$

3.2. Sprinkling and monotonicity

In this subsection, we highlight sprinkling and monotonicity as central tools entering the proofs of Propositions 3.2 and 3.3 and in particular we explain which role the colouring plays in the proofs. To explain the idea behind sprinkling, assume we explore two disjoint subgraphs of G_n. If we want to show that these subgraphs are connected to each other, we would like to use that they are independent. However, it is sometimes challenging to exclude hidden dependencies in the construction or definition of the subgraphs in question. It would be much easier to sample edges independently between them. This is where the sprinkling technique enters the stage.

Since G_n is built from a Poisson point process, it is easiest to add a few additional points to X, which potentially results in additional edges in the S-PAM. It is convenient here to consider again the alternative formulation of the model obtained by first considering the set $X \times X$ of 'potential edges' and then assigning the collection of i.i.d. weights $V_{X\times X} = \{V_{x,y}\}_{x,y\in X}$ to the potential edges, with $V_{x,y}$ uniform on [0, 1] and an edge is added between (x, s) and (y, t) with $s < t$ if and only if $V_{x,y} \leq \varphi(t|x - y|^d/f(Z_x(t-)))$.

Formally, we consider, for b close to 1, the independent colouring of the Poisson process X described in the previous section, i.e. each node is either *black* with probability b or *red* with probability $r = 1 - b$. Hence, the thinning theorem for Poisson processes [14], Corollary 5.9, decomposes $X = X^b \cup X^r$ into a black and an independent red Poisson process with parameters b and r, respectively. Recall that G_n^b denotes the S-PA graph built from X^b only.

By continuity, the S-PAM G_n^b obtained from X^b resembles G_n, for b close to 1 and we view the edges sent from a red to a black vertex as a version of sprinkling. Now, the following monotonicity principle holds, where for two geometric graphs G, H we write $G \subset H$ if every vertex and every edge of G is also contained in H.

Lemma 3.4 (Monotonicity). *The graphs G_n, G_n^b and G_n^r can be defined on the same probability space in such a way that almost surely $G_n^b \cup G_n^r \subset G_n$.*

Proof. First, represent the Poisson process as $X = X^b \cup X^r$, where the latter is an independent superposition of the black Poisson process and the red Poisson process. We sample the edge variables in a consistent manner with the above decomposition. That is, $V_{X^b\times X^b} \subset V_{X\times X}$. Now, we couple G_n^b, G_n via the sequential PA construction. By monotonicity of the attachment and a suitable coupling of the corresponding in-degree evolutions, $G_n^b(t) \subset G_n(t)$ holds for all $t \le 1$. It only remains to note that we can also construct G_n^r in a consistent manner such that $G_n^r \subset G_n$. To achieve this, we use an identical copy of X^r and the corresponding restricted weights and just run the construction of G_n^r alongside the construction of G_n^b, G_n above and observe that by monotonicity of the attachment rule, any edge drawn in G_n^r is also drawn in G_n. □

3.3. Proof of Proposition 3.1

We prove a slightly stronger assertion based on the modified reachability events

$$E_n^*(D, N, s) = \{\text{at least } N \text{ black vertices born before time } s \text{ are } D\text{-reachable}\}.$$

Proof of Proposition 3.1. In [12,13], it is shown that as $n \to \infty$, the finite graphs $\{G_n^b\}_{n\ge 1}$ converge weakly to a local limit graph H_∞ with vertices in $\mathbb{R}^d \times [0,1]$ and $\theta^b \in (0,1)$ in (2) equals the proportion of vertices contained in the unique infinite component K_∞ of H_∞. Let us consider this limit graph and introduce the events $E_\infty^*(D, N, s)$ corresponding to $E_n^*(D, N, s)$ in H_∞. On the event $(o, U) \in K_\infty$, we already know that there exist (shortest) paths connecting (o, U) to at least N black vertices born before time s, for N, s fixed. This is a consequence of [13], Prop. 13, in which an infinite path containing arbitrarily old vertices is shown to be contained in H_∞ and of [13], Prop. 4, which asserts the uniqueness of the infinite cluster in H_∞. In particular,

$$D_\infty = \min\{D : \text{the event } E_\infty^*(D, N, s) \text{ occurs}\}$$

is finite. To use the random variables D_∞ in the finite graphs $\{G_n^b\}_{n\ge 1}$ we use a coupling from [12], Section 4.1: given a Poisson point process X^b of intensity b on $\mathbb{R}^d \times [0,1]$ with edge weights $V_{X^b\times X^b}$ we may construct both H_∞ and $\{G_n^b\}_{n\ge 1}$ from $(X^b, V_{X^b\times X^b})$ in a consistent manner by restricting X^b to the torus $\mathbb{T}_n$ for each n. Under this coupling, local weak convergence of $\{G_n^b\}_{n\ge 1}$ to H_∞ is turned into almost sure convergence of any finite vector of edge indicators [12], Proposition 5. Since the event $E_\infty^*(D, N, s)$ depends only on finitely many edges, we conclude that $E_n^*(D_\infty, N, s)$ occurs eventually if $(o, U) \in K_\infty$. Hence, in the upper bound

$$\begin{aligned}
&\mathbb{P}_{(o,U)}\big(\{(o,U) \in C_n^b\} \setminus E_n^*(D_\infty, N, s)\big) \\
&\quad \le \mathbb{P}_{(o,U)}\big(\{(o,U) \in K_\infty\} \setminus E_n^*(D_\infty, N, s)\big) + \mathbb{P}_{(o,U)}\big((o,U) \in C_n^b \setminus K_\infty\big),
\end{aligned}$$

the first term tends to 0 as $n \to \infty$. Moreover, the second term converges to 0 by (2) and uniqueness of the infinite component in H_∞. □

3.4. Proofs of Propositions 3.2 and 3.3

In what follows, we again apply Lemma 3.4 to embed the red graph G_n^r into G_n. We show that high-degree nodes are connected in two steps. First, Lemma 3.5 states that two moderately old red vertices of high degree are likely to both connect to a young red vertex. Hence, they are at graph distance at most 2 from each other in G_n^r, as long as they are sufficiently close in $\mathbb{T}_n$. Second, by Lemma 3.6, red high-degree vertices are well spread out such that a red high-degree vertex has a red vertex with a much higher degree not too far away in $\mathbb{T}_n$. This is reminiscent of the robustness proof in [13]. Nevertheless, due to the different nature of our goal, we conduct a more refined analysis.

For vertices $(x,s),(y,t) \in X^r$ set $\Psi^r(x,s) = Z_x^r(1/2)s^{(\beta-\alpha\gamma)\delta}$, where $Z_\cdot^r(\cdot)$ denotes in-degree evolutions in G_n^r. The following lemma is a variant of [13], Lemma 11, and follows from the more general Lemma A.1 in the appendix. Here, we say x and y are 2-*connected* if $x \leftarrow z \rightarrow y$ in G_n^r for some $(z,r) \in X^r\restriction_{\mathbb{T}_n\times[1/2,1]}$.

Lemma 3.5 (2**-connections).** *There exists a constant $c>0$ such that for every sufficiently large n and every locally r-good $(x,s),(y,t) \in X^r$ with $s,t \le 1/4$, $Z_x^r(1/2)^\alpha \le Z_y^r(1/2)$ and $|x-y|^d \le s^{-\beta}$ we have*

$$\mathbb{P}\big(x \textit{ and } y \textit{ are } 2\text{-connected} \mid X^r \cap \big(\mathbb{T}_n \times [0,1/2]\big)\big) \ge 1 - \mathrm{e}^{-cr\Psi^r(x,s)}.$$

Proceeding as in [13], Proposition 13, we now discover locally good vertices in X^r. Since in this section, we always explore X by moving 'towards the right', that is, by increasing the first space-coordinate and since we determine local goodness of a vertex (x,s) by peeking into a cube of volume $1/s$ around x, the following σ-algebra naturally captures the information collected during the exploration process in G_n^r:

$$\mathcal{F}\big(x^-\big) = \sigma\big(X'(x), V\restriction_{X'(x)}\big)$$

with

$$X'(x) = X^r\restriction_{[0,x]\times\mathbb{R}^{d-1}\times[0,1]} \cup X^r\restriction_{[x-s^{-1/d},x+s^{-1/d}]^d\times[0,1/2]}).$$

We write lgood_n^r for the family of locally good vertices in G_n^r and recall from (6) that the connection scheme between the high-degree vertices relies on the parameters α, β.

Lemma 3.6. *Let $r>0$ be arbitrary. Then, there exists a constant $q \in (0,1)$ with the following property. If $(x,s) \in G_n^r$ is any vertex with $n^{-1/\beta} < s \le 1/4$, then*

$$\mathbb{P}\big(\mathsf{lgood}_n^r \cap \big(B_{s^{-\beta/d}}(x) \times \big[0,s^\alpha\big]\big) \ne \varnothing \mid \mathcal{F}\big(x^-\big)\big) \ge 1 - q^{\frac{1}{7}s^{-(\beta-\alpha)/d}}.$$

Proof. For $\eta = \beta - \alpha$, we select $M = \lfloor s^{-\eta/d}/6 - 1\rfloor$ disjoint sub-intervals $I_1,\dots,I_M$ with midpoints $a_1,\dots,a_M$ within the interval $[s^{-\alpha/d}, s^{-\beta/d}]$. Now, define blocks

$$A_k = (a_k,0,\dots,0) + \big[-3s^{-\alpha/d}, 3s^{-\alpha/d}\big]^d, \quad \text{and}$$
$$B_k = (a_k,0,\dots,0) + \big[-s^{-\alpha/d}, s^{-\alpha/d}\big]^d.$$

The blocks $x + A_k$ are disjoint and their point configuration is independent of $\mathcal{F}(x^-)$. Next, note that the total number of points in $(x+B_k) \times (s^\alpha/2, s^\alpha)$ is Poisson distributed with parameter of constant order proportional to r. Moreover, by Corollary A.3, the expected number among them that are locally r-good is bounded away from 0. Hence, each of the blocks $x + B_k$ contains a locally good vertex born in $(s^\alpha/2, s^\alpha)$ with probability at least $1 - q \in (0, 1)$.

Now, to check local goodness we need to check in the worst case a set of diameter $2^{1/d} s^{-\alpha/d} \le 2s^{-\alpha/d}$, that is, these vertices occur independently for different blocks. Hence, the number of locally good vertices at distance at most $s^{-\beta/d}$ from x dominates a multinomial random variable with M trials and success probability $1 - q$, which exceeds 0 with probability $1 - q^M$. □

We are now in the position to prove the main result of this section, Proposition 3.2. Since, $Z^r_y(1/2) \ge Z^r_x(1/2)^\alpha$ holds for any $x \in L^r_{k+1} \setminus L^r_k$ and $y \in L^r_k$, by Lemma 3.5, with high probability, a vertex in L^r_{k+1} is 2-connected to a vertex in L^r_k.

Proof of Proposition 3.2. Let $Y_0 = (x_0, t_0)$ denote the vertex guaranteed by the event $E^b_n(s)$. We wish to apply first Lemma 3.6 and then Lemma 3.5 to find a locally good red vertex (x_1, t_1) with $|x_0 - x_1|^d \le t_0^{-\beta}$ and $t_1 \le t_0^\alpha$ that 2-connects to (x_0, t_0), thus establishing that (x_0, t_0) and (x_1, t_1) are at distance at most 2 in G^r_n. The probability that this fails is bounded by

$$e_1 = q^{\frac{1}{7} t_0^{-(\beta-\alpha)/d}} + \exp\bigl(-cr\Psi^r(x_0, t_0)\bigr). \tag{8}$$

Iteration yields

$$e_j = q^{\frac{1}{7} t_{j-1}^{-(\beta-\alpha)/d}} + \exp\bigl(-cr\Psi^r(x_{j-1}, t_{j-1})\bigr).$$

Note that $\Psi(x_{j-1}, t_{j-1}) \ge t_{j-1}^{\beta\delta-\gamma-\alpha\gamma\delta}/g(t_{j-1}^{-1})$ and we recall that

$$\nu = \min\bigl\{-\beta\delta + \gamma + \alpha\gamma\delta, (\beta - \alpha)/d\bigr\} > 0.$$

Hence, we can find a small number $q > 0$ with

$$e_j \le 2\exp\bigl(-qt_{j-1}^{-\nu}\bigr) \le 2\exp\bigl(-qt_0^{-\nu\alpha^{j-1}}\bigr).$$

The probability of failing to reach L^r_K from (x_0, t_0) in G^r_n is thus bounded by

$$2\sum_{j\ge1} \exp\bigl(-qt_0^{-\nu\alpha^{j-1}}\bigr) \le 2\sum_{j\ge1} \exp\bigl(-qs^{-\nu\alpha^{j-1}}\bigr),$$

which can be made arbitrarily small by lowering s, see Lemma A.4. Note that it takes at most $O(\log\log\log n)$ iterations to arrive at a vertex with birth time $1/(\log n)^C$ for any $C > 0$, since the birth time of the freshly discovered vertex is lower by at least a fixed power than the birth time of the last vertex in each iteration. □

For Proposition 3.3, we proceed similarly as in Proposition 3.2.

Proof of Proposition 3.3. Let a vertex in $(x, s) \in L^r_K$ be given. We need to consider two cases: $s < n^{-1/\beta}$ and $s \in (n^{-1/\beta}, n^{-1/\alpha^K})$. In the first case, we argue directly as in the proof of [13], Proposition 15, to obtain that (x, s) is either the oldest vertex in X^r or, by Lemma A.1, it connects to it with probability exceeding $1 - \varepsilon^{\log s^2}$. Note that there are at most $O(n^{1-1/\beta})$ such vertices, that is, this argument holds with high probability simultaneously for all of them.

Let us now consider $s \in (n^{-1/\beta}, n^{-1/\alpha^K})$. An iteration as in the proof of Proposition 3.2 yields a chain of 2-connections connecting (x, s) to the oldest vertex in at most K steps. Since the error bound is weakest in the first step, we may bound the total probability that the desired path does not exist by

$$\bar{q}(n) = K \exp\bigl(-q\bigl((\log n)^{\alpha/\nu}\bigr)^{\nu}\bigr).$$

Since $\alpha > 1$ we have $n\bar{q}(n) = o(1)$. Thus, the probability of failing to connect *any* node in L_K^r to the oldest vertex in K steps vanishes, as with high probability there are at most of order n vertices in the system. Thus, the diameter of L_K^r is at most $4K$ in G_n^r. □

4. Proof of Theorem 2.3

In this section, we prove the LDP asserted in Theorem 2.3 and deduce Corollary 2.5 by applying the LDP for the empirical field of a marked Poisson point process [9], Theorem 3.1. We first introduce an approximated network dynamic, where connections appear only up to a finite distance. In a second step, we show that this modified dynamic forms an exponentially good approximation in the sense of [3], Definition 4.2.14.

In order to prove Theorem 2.3, we rely on the LDP for the empirical field of a marked Poisson point process in the $\tau_{\mathcal{L}}$-topology of local convergence [9], Theorem 3.1. However, this result is not directly applicable in the present setting. Indeed, the $\tau_{\mathcal{L}}$-topology captures only interactions of bounded range, whereas the polynomial decay of the profile function φ allows for arbitrarily long edges.

The proof of Theorem 2.3, proceeds in two steps. First, in Proposition 4.1, we see that after truncating edges longer than a fixed distance, the resulting neighbourhood evolution is continuous in the input data. In particular, the contraction principle yields an LDP in the truncated setting. Second, Proposition 4.2 shows that changes induced by the truncations are asymptotically negligible in the sense of exponentially good approximations [3], Definition 4.2.14. Since the sprinkling construction does not appear in this section, we overwrite the previous notation G^r. This approximation has the advantage of exhibiting only local dependencies.

The truncated S-PAM G^r suppresses potential connections longer than a fixed threshold $r > 0$. That is, we consider the dynamics (1), except that $|x - y|$ is replaced by ∞ if $|x - y| > r$. For $t \le 1$ and $\mathbb{Q} \in \mathcal{P}_\theta$, we write $\mathbb{Q}^{*,r-\mathsf{neighb}}(t)$ for the measure on $\mathcal{G}^*$ determined by the rooted graph $[G^r(t), o]$ under the Palm measure $\mathbb{Q}^*$. Moreover, $\mathbb{Q}^{*,r-\mathsf{neighb}}(t, h)$ denotes the projection of this measure under the map of taking the h-neighbourhood.

Proposition 4.1 (LDP for finite-range model). *The approximated neighbourhood evolution $\mathbb{Q}^{*,r-\mathsf{neighb}} \in \mathcal{M}(\mathcal{G}^*)^{[0,1]}$ is continuous in $\mathbb{Q} \in \mathcal{P}_\theta$ under the $\tau_{\mathcal{L}}$-topology.*

Proof. Fix $t \in [0, 1]$, $h \ge 0$ and $g \in \mathcal{G}^*$. Then, the indicator of the event that $\{[G^r(t), o]_h \simeq g_h\}$ is a local observable. Hence, $\mathbb{Q}^{*,r-\mathsf{neighb}}(t, h)(g_h)$ is continuous in $\mathbb{Q}$ under the $\tau_{\mathcal{L}}$-topology, as asserted. □

Since the empirical field induced by the marked Poisson point process satisfies the LDP in the $\tau_{\mathcal{L}}$-topology with specific entropy as good rate function [9], Theorem 3.1, combining Proposition 4.1 with the contraction principle implies that $L_n^{r-\mathsf{neighb}}$ satisfies the LDP with good rate function

$$\nu \mapsto \inf_{\substack{\mathbb{Q}\in\mathcal{P}_\theta \\ \mathbb{Q}^{*,r-\mathsf{neighb}}=\nu}} H(\mathbb{Q}).$$

In order to bridge the gap between $L_n^{r-\mathsf{neighb}}$ and L_n^{neighb}, we rely on the machinery of exponentially good approximation [3,6].

Proposition 4.2. *Let $t \in [0,1]$, $h \geq 0$ and $g \in \mathcal{G}^*$. Then, the random variables $L_n^{r-\mathsf{neighb}}(t, h)(g_h)$ are an exponentially good approximation of $L_n(t,h)(g_h)$.*

Before establishing Proposition 4.2, we prove Theorem 2.3.

Proof of Theorem 2.3. Let $t \in [0,1]$, $h \geq 0$ and $g \in \mathcal{G}^*$. First, by [6], Corollary 1.11, it suffices to prove for every $\alpha > 0$ that

$$\lim_{r\to\infty} \sup_{\substack{\mathbb{Q}\in\mathcal{P}_\theta \\ H(\mathbb{Q})\leq\alpha}} \left|\mathbb{Q}^{*,r-\mathsf{neighb}}(t,h)(g_h) - \mathbb{Q}^{*,\mathsf{neighb}}(t,h)(g_h)\right| = 0.$$

Now, for different values of r the approximations G^r are coupled in the sense that for $r' \geq r$ both $\mathbb{Q}^{*,r-\mathsf{neighb}}(t,h)(g_h)$ and $\mathbb{Q}^{*,r'-\mathsf{neighb}}(t,h)(g_h)$ integrate suitable indicators with respect to $\mathbb{Q}^*$. In particular, by the dominated convergence theorem,

$$\lim_{r\to\infty} \mathbb{Q}^{*,r-\mathsf{neighb}}(t,h)(g_h) = \mathbb{Q}^{*,\mathsf{neighb}}(t,h)(g_h).$$

Hence, it suffices to show that for $\varepsilon > 0$ there exists $r_0 = r_0(\varepsilon)$ with the following property. If $\mathbb{Q} \in \mathcal{P}_\theta$ satisfies $H(\mathbb{Q}) \leq \alpha$, then

$$\sup_{r'\geq r\geq r_0} \left|\mathbb{Q}^{*,r-\mathsf{neighb}}(t,h)(g_h) - \mathbb{Q}^{*,r'-\mathsf{neighb}}(t,h)(g_h)\right| \leq \varepsilon. \tag{9}$$

By Proposition 4.2, there exists $r_0 > 0$ such that

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(\left|L_n^{r-\mathsf{neighb}}(t,h)(g_h) - L_n^{r'-\mathsf{neighb}}(t,h)(g_h)\right| > \varepsilon\right) < -\alpha \tag{10}$$

holds for every $r' \geq r \geq r_0$. As can be deduced from Proposition 4.1, not only the random variable $L_n^{r-\mathsf{neighb}}(t,h)(g_h)$ but also the difference $L_n^{r-\mathsf{neighb}}(t,h)(g_h) - L_n^{r'-\mathsf{neighb}}(t,h)(g_h)$ satisfies the LDP and the rate function equals

$$a \mapsto \inf_{\substack{\mathbb{Q}\in\mathcal{P}_\theta \\ \mathbb{Q}^{*,r-\mathsf{neighb}}(t,h)(g_h)-\mathbb{Q}^{*,r'-\mathsf{neighb}}(t,h)(g_h)=a}} H(\mathbb{Q}).$$

In particular, (10) gives that

$$-\inf_{\substack{\mathbb{Q}\in\mathcal{P}_\theta \\ |\mathbb{Q}^{*,r-\mathsf{neighb}}(t,h)(g_h)-\mathbb{Q}^{*,r'-\mathsf{neighb}}(t,h)(g_h)|>\varepsilon}} H(\mathbb{Q}) < -\alpha,$$

so that the asserted upper bound (9) holds for every $\mathbb{Q}\in\mathcal{P}_\theta$ with $H(\mathbb{Q})\le\alpha$. □

4.1. Proof of Proposition 4.2

To prove exponentially good approximation, we compare the S-PAM with the Poisson random connection model (RCM) [8]. In general, the S-PAM differs from the RCM substantially because preferential attachment leads to high-degree nodes. However, checking whether the h-neighbourhood of a given vertex is of a certain form entails a uniform bound on the maximum size of the relevant in-degrees, so that any discrepancy between G_n and G_n^r must come from an edge of length at least r in the RCM. The integrability of the profile function implies that this is a rare event.

Carrying out this program rigorously involves several intermediate steps that we state now and prove later in this section. First, the proportion of nodes arriving at early times is asymptotically negligible.

Lemma 4.3 (Early vertices). *Let* $\varepsilon > 0$. *Then,*

$$\limsup_{\sigma\to 0}\limsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}\big(\#\big(X\cap\big(\mathbb{T}_n\times[0,\sigma]\big)\big)>\varepsilon n\big)=-\infty.$$

Second, contributions from neighbourhoods around vertices located in highly dense regions of $\mathbb{T}_n$ can also be ignored. To simplify the presentation, we assume in the following that $n'=n^{1/d}$ is an integer. Now, we partition $\mathbb{T}_n$ into cubes $Q_z = z+[-1/2,1/2]^d$ centred at sites of discrete torus $\mathbb{Z}^d/n'$. Moreover, we let $N_z=\#(X\cap(Q_z\times[0,1]))$ denote the number of vertices in Q_z. For a threshold $m\ge 1$, we define $z\in\mathbb{Z}^d/n'$ to be *m-dense*, in symbols $z\in D_m$, if $N_z\ge m$ and *m-sparse* otherwise.

Lemma 4.4 (High-density regions). *Let* $\lambda > 0$. *Then,*

$$\lim_{m\to\infty}\sup_{n\ge 1}\frac{1}{n}\log\mathbb{E}\Big[\exp\Big(\lambda\sum_{z\in D_m}N_z\Big)\Big]=0.$$

Third, we deduce the finiteness of an exponential moment, thereby helping to bound the number of edges emanating from a given vertex. Fix $b, m>0$ and let $\{N_z'\}_{z\in\mathbb{Z}^d}$ be a family of independent random variables, where N_z' follows the distribution of a binomial random variable with m trials and success probability $\varphi(b(|z|-\sqrt{d})_+^d)$.

Lemma 4.5 (Exponential moment). *Let $b, m, \lambda > 0$. Then,*

$$\mathbb{E}\Big[\exp\Big(\lambda \sum_{z\in\mathbb{Z}^d} N'_z\Big)\Big] < \infty.$$

Finally, we bound the number of long edges ending in an m-sparse cube. For $x \in X$ we define $z(x) \in \mathbb{Z}^d/n'$ to be such that $x \in Q_{z(x)}$. In words, $z(x)$ is the centre of the cube containing x.

Lemma 4.6 (Long edges). *Let $b, m, \lambda > 0$. Then,*

$$\lim_{r\to\infty}\sup_{n\geq 1}\frac{1}{n}\log\mathbb{E}\Big[\exp\Big(\lambda \sum_{\substack{x,y\in X\\|x-y|>r}} 1\big\{z(y)\notin D_m \text{ and } V_{x,y}\leq \varphi\big(b|x-y|^d\big)\big\}\Big)\Big] = 0.$$

Before establishing Lemmas 4.3–4.6, we prove the main result.

Proof of Proposition 4.2. Without loss of generality, set $t = 1$. We derive a bound for the number of *bad vertices* $x \in X$, that is, vertices whose h-neighbourhood is isomorphic to g_h in G_n but not in G_n^r. The corresponding bound with interchanged roles of G_n and G_n^r follows from similar arguments. Hence, for any such x there exists a vertex x' in the $(h-1)$-neighbourhood of x in G_n^r and a vertex y' with $|y'-x'| > r$ such that there is an edge between y' and x' in G_n but not in G_n^r. By Lemma 4.3, we may restrict our attention to vertices born after time σ. In particular, writing ℓ for the number of vertices in g_h, we see that y' is adjacent to x' in the RCM G^{rc} with vertex set X and in which there is an edge from (x_1, s_1) to (x_2, s_2) if and only if $V_{x_1,x_2} \leq \varphi_*(|x_1-x_2|^d)$, where $\varphi_*(\rho) = \varphi(\sigma\rho/f(\ell))$.

To obtain bounds on exponential moments, we aim to restrict our attention to neighbourhoods intersecting only m-sparse cubes. More precisely, a self-avoiding path $\pi = ((x_0, s_0), \dots, (x_j, s_j))$ in G^{rc} is *m-sparse* if

1. $|x_i - x_{i+1}| \leq r$ and $s_{i+1} \leq s_i$ for every $i \geq 0$,
2. $z(x_i) \notin D_m$ for every $i \geq 1$.

It is on purpose, that we do not impose $z(x_0) \notin D_m$. With this definition, every bad vertex is contained in an m-sparse connected path that either starts at a vertex contained in an m-dense cube or features an in-going edge of length at least r in the RCM G^{rc}. Writing $\mathcal{C}_1$ and $\mathcal{C}_2$ for the families of vertices contained in these types of paths, it suffices to provide upper bounds for $\#\mathcal{C}_1$ and $\#\mathcal{C}_2$. More precisely, by the exponential Markov inequality it suffices to show that for every $\lambda > 0$,

$$\limsup_{m\to\infty}\limsup_{n\to\infty}\frac{1}{n}\log\mathbb{E}\big[\exp(\lambda\#\mathcal{C}_1)\big] = 0, \tag{11}$$

and

$$\limsup_{r\to\infty}\limsup_{n\to\infty}\frac{1}{n}\log\mathbb{E}\big[\exp(\lambda\#\mathcal{C}_2)\big] = 0. \tag{12}$$

To begin with, we show (11) and introduce the *m-sparse connected component C_x at $x \in X$* as the union of all vertices of m-sparse paths in G^{rc} starting at $x \in X$ and consisting of at most h

hops. First, recalling the definition of N'_z from the paragraph preceding Lemma 4.5, the number of edges in G^{rc} from any $x' \in X$ to a vertex in $z \notin D_m$ is stochastically dominated by $N'_{z(x')-z}$. In particular, the independence of the family $\{V_{x,y}\}_{x,y\in X}$ implies that conditioned on X, the size of C_x is stochastically dominated by the offspring until generation h of a Galton–Watson process with offspring distribution $N' = \sum_{z\in\mathbb{Z}^d} N'_z$. Despite the independence assumption on the collection $\{V_{x,y}\}_{x,y\in X}$, the m-sparse connected components C_x at different points $x \in X$ are not independent because they can share common vertices. Nevertheless, the size of their union is stochastically dominated by the sum of the component sizes, when each component is explored independently, see [5], Lemma 2.3. Hence, noting that Lemma 4.5 yields the finiteness of the cumulant generating function $c(\lambda) = \log \mathbb{E}[\exp(\lambda N')]$, we arrive at

$$\frac{1}{n}\log\mathbb{E}\big[\exp(\lambda\#\mathcal{C}_1)\big] \le \frac{1}{n}\log\mathbb{E}\big[\exp\big(c^{(h)}(\lambda)\#\big\{x\in X : z(x)\in D_m\big\}\big)\big], \tag{13}$$

where $c^{(h)}$ denotes the h-fold iteration of the function c. By Lemma 4.4, the right-hand side tends to 0 as $m \to \infty$.

In order to show (12), we proceed in precisely the same way until arriving at the analog of (13), where the number of vertices contained in an m-dense cube is replaced by the number of vertices that are contained in an m-sparse cube and are the endpoint of an edge in G^{rc} of length at least r. Instead of Lemma 4.4, now Lemma 4.6 implies that the resulting expression tends to 0 as $r \to \infty$. □

Finally, we prove Lemmas 4.3–4.6.

Proof of Lemma 4.3. For fixed $\sigma > 0$, the number $\#(X \cap (\mathbb{T}_n \times [0,\sigma]))$ of Poisson points born before time σ is a Poisson random variable with parameter $n\sigma$. In particular, by the Poisson concentration inequality

$$\frac{1}{n}\log\mathbb{P}\big(\#\big(X\cap\big(\mathbb{T}_n\times[0,\sigma]\big)\big) > \varepsilon n\big) \le -\frac{\varepsilon}{2}\log\big(\varepsilon\sigma^{-1}\big),$$

where the right-hand side tends to $-\infty$ as $\sigma \to 0$. □

Proof of Lemma 4.4. First, by the independence property of the Poisson process,

$$\frac{1}{n}\log\mathbb{E}\bigg[\exp\bigg(\lambda\sum_{z\in D_m} N_z\bigg)\bigg] = \log\mathbb{E}\big[\exp\big(\lambda 1\{o\in D_m\}N_o\big)\big].$$

As N_o has exponential moments, we conclude by the monotone convergence theorem. □

Proof of Lemma 4.5. By independence of the $\{N'_z\}_{z\in\mathbb{Z}^d}$, we can expand the exponential moment as

$$\begin{aligned}\log\mathbb{E}\bigg[\exp\bigg(\sum_{z\in\mathbb{Z}^d} N'_z\bigg)\bigg] &= m\sum_{z\in\mathbb{Z}^d}\log\big(1+\varphi\big(b\big(|z|-\sqrt{d}\big)_+^d\big)\big)\\ &\le m\sum_{z\in\mathbb{Z}^d}\varphi\big(b\big(|z|-\sqrt{d}\big)_+^d\big).\end{aligned}$$

Now, $\int_{\mathbb{R}^d} \varphi(|y|^d)\,\mathrm{d}y = \int_0^\infty \kappa_d \varphi(x)\,\mathrm{d}x < \infty$. Since φ is also decreasing, we conclude that the infinite series in the last line converges. □

Proof of Lemma 4.6. Since the $\{V_{x,y}\}_{x,y\in X}$ are i.i.d., we see that conditioned on X the events

$$\{V_{x,y} \le \varphi\big(b|x-y|^d\big)\}$$

are independent for different values of x and y. Hence, the sum in the exponential is stochastically dominated by

$$\sum_{x\in X} \sum_{z':|z(x)-z'|>r/2} N_{x,z'},$$

where the $N_{x,z'}$ are independent binomial random variables with m trials and success probability $\varphi(b(|z(x)-z'|-\sqrt{d})^d)$. In particular, by the formula for the characteristic function of compound Poisson sums,

$$\begin{aligned}
&\frac{1}{n}\log \mathbb{E}\Big[\exp\Big(\lambda \sum_{x\in X}\sum_{z':|x-z'|>r/2} N_{x,z}\Big)\Big] \\
&= \mathbb{E}\Big[\exp\Big(\lambda \sum_{z:|z|>r/2} N_{o,z}\Big)\Big] - 1 \\
&= \exp\Big(m \sum_{z:|z|>r/2} \log\big(1+\varphi\big(b\big(|z|-\sqrt{d}\big)^d\big)\big(e^\lambda - 1\big)\big)\Big) - 1 \\
&\le \exp\Big(m\big(e^\lambda - 1\big) \sum_{z:|z|>r/2} \varphi\big(b\big(|z|-\sqrt{d}\big)^d\big)\Big) - 1,
\end{aligned}$$

where the last sum converges by the integrability assumption on the profile function. Sending $r \to \infty$ concludes the proof. □

Appendix A: Auxiliary results

Here, we provide statements from the literature and auxiliary calculations used in the main text. We start by a refined version of Lemma 3.5, where we write $\kappa_d = |B_1(o)|$ for the volume of the unit ball in $\mathbb{R}^d$.

Lemma A.1. *Denote by G_n^λ the S-PAM built from a homogeneous Poisson point process X^λ with intensity $\lambda > 0$. Let $(x,s),(y,t) \in X^\lambda$ be two vertices satisfying $s,t \le 1/2$. Moreover, let $Z_x(1/2)$ and $Z_y(1/2)$ denote their respective in-degrees in G_n^λ at time $1/2$ and define*

$$k(x,y) = f\big(Z_x(1/2)\big)\varphi\Big(\frac{(f(Z_x(1/2)))^{1/d} + |x-y|)^d}{f(Z_y(1/2))}\Big),$$

and

$$Q(x, y) = \frac{\varphi(1)\kappa_d}{2}\big(k(x, y) \vee k(y, x)\big).$$

Then, conditional on $X^\lambda \cap (\mathbb{T}_n \times [0, 1/2])$, x *and* y *are* 2*-connected in* G_n^λ *by using only vertices from* $X^\lambda \cap (\mathbb{T}_n \times [1/2, 1])$ *with probability exceeding* $1 - \mathrm{e}^{-\lambda Q(x,y)}$.

Proof. Set $z_x = Z_x(1/2)$, $z_y = Z_y(1/2)$ and let X° denote those vertices (w, r) of X^λ which lie in $B_{f(z_x)^{1/d}}(y) \times [1/2, 1]$ and satisfy $V_{x,w} \le \varphi(1)$ and $V_{y,w} \le \varphi(r|y - w|^d/f(z_y))$. In particular, all $(w, r) \in X^\circ$ are 2-connectors. By the restriction theorem [14], Theorem 5.2, X° forms a Poisson point process with intensity

$$\int_{B_{f(z_x)^{1/d}}(y)} \frac{\lambda\varphi(1)}{2}\varphi\big(r|y - w|^d/f(z_y)\big)\,\mathrm{d}w \ge \frac{\lambda\varphi(1)\kappa_d f(z_x)}{2}\varphi\bigg(\frac{(f(z_x)^{1/d} + |x - y|)^d}{f(z_y)}\bigg).$$

In particular,

$$\mathbb{P}\big(X^\circ = \varnothing\big) \le \exp\bigg(-\frac{1}{2}\lambda\varphi(1)\kappa_d k(x, y)\bigg),$$

so that reversing the roles of x and y yields the assertion of the lemma. □

The next statement ensures that old vertices tend to be good. Let $Z^n(s, \cdot)$ denote the generic in-degree evolution of a vertex born at time s in G_n, noting that its spatial position has no influence on $Z^n(s, \cdot)$. Assume that G_n is built from a Poisson process of intensity $\lambda > 0$.

Lemma A.2 ([13], Lemma 24). *Let* $(x, s) \in G_n$ *be born at time* $s \le 1/2$. *There exists a function* $g = g_\lambda$ *decaying faster than any power at* ∞ *such that*

$$\sup_{\substack{n\ge1\\ n\log n\ge s^{-1}}} \mathbb{P}_{(x,s)}\big(Z^n(s, 1/2) \le s^{-\gamma}/g\big(s^{-1}\big)\big) \xrightarrow{s\to0} 0.$$

Consequently,

$$\sup_{\substack{n\ge1\\ n\log n\ge s^{-1}}} \mathbb{P}_{(x,s)}\big((x, s) \text{ is not good}\big) \xrightarrow{s\to0} 0.$$

Proof. See [13], page 1720, for the proof with $\lambda = 1$. A higher intensity increases the degree of (x, s). That lowering the intensity makes no difference to the sub-polynomial decay can be seen easily, since g may be replaced by any other increasing function of sufficiently slow decay, cf. the proofs of Lemma 23 and 24, [13], page 1720. In particular, reducing the intensity of the Poisson process can be compensated for by increasing g by a constant factor. This has no influence on its sub-polynomial decay. □

The following corollary is obtained directly from the proof of [13], Lemma 24: for a given birth time s, using a scaling property of the degree evolutions, it is actually sufficient to consider connections to (x, s) in an $s^{-1/d}$ environment of x. This fact is also used, without explicit mentioning, in the proof of [13], Proposition 13.

Corollary A.3. *For any $x \in \mathbb{T}_n$ we have*

$$\inf_{\substack{s<1/2 \\ n\geq 1}} \mathbb{P}_{(x,s)}\big((x, s) \textit{ is locally good}\big) > 0.$$

The following short calculation shows that the error in the proof of Proposition 3.2 can be made arbitrarily small.

Lemma A.4. *For any $q, \varepsilon > 0$ and $\alpha > 1$ we have*

$$\lim_{s\to 0} \sum_{k\geq 1} \exp\big(-q s^{-\varepsilon\alpha^k}\big) = 0.$$

Proof. Clearly $r(x) = \exp(-qx^{-\varepsilon}) \to 0$ as $x \to 0$, and

$$\sum_{k\geq 1} r\big(s^{\alpha^k}\big) \leq \sum_{k:\alpha^k \leq k} r\big(s^{\alpha^k}\big) + \frac{1}{1 - r(s)} - 1,$$

which vanishes as $r(s) \to 0$. □

Appendix B: Proof of Corollary 2.5

Since Theorem 2.4 already provides an LDP for fixed times, the proof of Corollary 2.5 reduces to verifying exponential tightness in the Skorohod topology.

Proof of Corollary 2.5. By Corollary 2.4, the rescaled number of nodes $L_n^{\mathrm{deg};\leq k}$ of degree at most k satisfies an LDP in the product topology. In particular, by [3], Corollary 4.2.6, it suffices to establish exponential tightness in the Skorohod topology. To this end, we use a criterion established in [7], Theorem 4.1:

1. $L_n^{\mathrm{deg};\leq k}(t)$ is exponentially tight for every $t \leq 1$, and
2. $\limsup_{\eta\to 0} \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(w'_\eta(L_n^{\mathrm{deg};\leq k}) > \varepsilon) = -\infty$,

where

$$w'_\eta\big(L_n^{\mathrm{deg};\leq k}\big) = \inf_{0=t_0<\cdots<t_j=1:\min_{i\leq k}|t_i - t_{i-1}|>\eta} \max_{i\leq j} \sup_{s,t\in[t_{i-1},t_i)} \big|L_n^{\mathrm{deg};\leq k}(s) - L_n^{\mathrm{deg};\leq k}(t)\big|$$

denotes the Skorohod-adapted modulus of continuity.

Exponential tightness of $|L_n^{\mathrm{deg};\le k}(t)| \le n^{-1}\#X$ is a consequence of the exponential tightness of the rescaled Poisson random variable $n^{-1}\#X$. Therefore, we concentrate on item (2). Here, we proceed along the lines of the proof of Proposition 4.2. Fixing an interval $I \subset [0,1]$ of length η, we distinguish two cases. First, assume that $I \subset [0, 2\sigma]$ for some small $\sigma > 0$. Then, during the time interval I, the truncated in-degree of each node can change by at most $k \ge 1$. Therefore,

$$\frac{1}{n}\log \mathbb{P}\big(w'_\eta\big(L_n^{\mathrm{deg};\le k}\big) > \varepsilon\big) \le n^{-1}\log \mathbb{P}\big(k\#\big(X \cap \big(\mathbb{T}_n \times [0,\sigma]\big)\big) > n\varepsilon\big),$$

which by Lemma 4.3 tends to $-\infty$ as $\sigma \to 0$.

Hence, from now on we may assume that $I \subset [\sigma, 1]$. Then, setting again $\varphi_*(\rho) = \varphi(\sigma\rho/f(\ell))$, we proceed similarly to Proposition 4.2 and introduce the quantity

$$N = \#\big\{x \in X : V_{x,y} \le \varphi_*\big(|x-y|^d\big) \text{ and } Z_x(t-) \le k \text{ for some } (y,t) \in X \cap (\mathbb{T}_n \times I)\big\}.$$

Since the total number of in-degree changes during the time interval I is at most kN, it suffices to derive a suitable upper bound for the latter. In particular, $N \le N^{(\mathsf{s})} + \sum_{z \in D_m} N_z$, where

$$N^{(\mathsf{s})} = \sum_{\substack{x,y \in X\\ z(x) \notin D_m}} 1\big\{(y,t) \in X \cap (\mathbb{T}_n \times I) \text{ and } V_{x,y} \le \varphi_*\big(|x-y|^d\big)\big\}.$$

Using the exponential Markov inequality and Lemma 4.4, it suffices to show that for any fixed $\lambda > 0$ we have that

$$\log \mathbb{E}\big[\exp\big(\lambda N^{(\mathsf{s})}\big)\big] \le 2n$$

if η is chosen sufficiently small. To achieve this goal, we note that as in the proof of Lemma 4.6 the random variable $N^{(\mathsf{s})}$ is stochastically dominated by

$$\sum_{(y,t) \in X \cap (\mathbb{T}_n \times I)} \sum_{z \in \mathbb{Z}^d/n'} N_{y,z},$$

where again the $N_{y,z}$ are independent binomial random variables with m trials and success probability $\varphi(b(|z| - \sqrt{d})_+^d)$. Since $X \cap (\mathbb{T}_n \times I)$ is a Poisson point process with intensity ηn, this time we arrive at

$$\frac{1}{n}\log \mathbb{E}\big[\exp\big(\lambda N^{(\mathsf{s})}\big)\big] \le \eta\left(\mathbb{E}\left[\exp\left(\lambda \sum_{z \in \mathbb{Z}^d} N'_z\right)\right] - 1\right),$$

which by Lemma 4.5 tends to 0 as $\eta \to 0$. □

Acknowledgements

We thank the anonymous referee for carefully reading the manuscript and making highly valuable suggestions that helped to substantially improve the quality of the manuscript. CH is supported by The Danish Council for Independent Research – Natural Sciences, grant DFF – 7014-00074 *Statistics for point processes in space and beyond*, and by the *Centre for Stochastic Geometry and Advanced Bioimaging*, funded by grant 8721 from the Villum Foundation.

References

[1] Bordenave, C. and Caputo, P. (2015). Large deviations of empirical neighborhood distribution in sparse random graphs. *Probab. Theory Related Fields* **163** 149–222. MR3405616 https://doi.org/10.1007/s00440-014-0590-8

[2] Choi, J. and Sethuraman, S. (2013). Large deviations for the degree structure in preferential attachment schemes. *Ann. Appl. Probab.* **23** 722–763. MR3059274 https://doi.org/10.1214/12-AAP854

[3] Dembo, A. and Zeitouni, O. (1998). *Large Deviations Techniques and Applications*, 2nd ed. *Applications of Mathematics* (*New York*) **38**. New York: Springer. MR1619036 https://doi.org/10.1007/978-1-4612-5320-4

[4] Dereich, S. and Mörters, P. (2009). Random networks with sublinear preferential attachment: Degree evolutions. *Electron. J. Probab.* **14** 1222–1267. MR2511283 https://doi.org/10.1214/EJP.v14-647

[5] Deuschel, J.-D. and Pisztora, A. (1996). Surface order large deviations for high-density percolation. *Probab. Theory Related Fields* **104** 467–482. MR1384041 https://doi.org/10.1007/BF01198162

[6] Eichelsbacher, P. and Schmock, U. (1998). Exponential approximations in completely regular topological spaces and extensions of Sanov's theorem. *Stochastic Process. Appl.* **77** 233–251. MR1649006 https://doi.org/10.1016/S0304-4149(98)00047-7

[7] Feng, J. and Kurtz, T.G. (2006). *Large Deviations for Stochastic Processes*. *Mathematical Surveys and Monographs* **131**. Providence, RI: Amer. Math. Soc. MR2260560 https://doi.org/10.1090/surv/131

[8] Franceschetti, M. and Meester, R. (2007). *Random Networks for Communication*: *From Statistical Physics to Information Systems*. *Cambridge Series in Statistical and Probabilistic Mathematics* **24**. Cambridge: Cambridge Univ. Press. MR2398551

[9] Georgii, H.-O. and Zessin, H. (1993). Large deviations and the maximum entropy principle for marked point random fields. *Probab. Theory Related Fields* **96** 177–204. MR1227031 https://doi.org/10.1007/BF01192132

[10] Gracar, P., Grauer, A., Lüchtrath, L. and Mörters, P. (2018). The age-dependent random connection model. Preprint. Available at arXiv:1810.03429.

[11] Jacob, E. and Mörters, P. (2013). A spatial preferential attachment model with local clustering. In *Algorithms and Models for the Web Graph*. *Lecture Notes in Computer Science* **8305** 14–25. Cham: Springer. MR3163707 https://doi.org/10.1007/978-3-319-03536-9_2

[12] Jacob, E. and Mörters, P. (2015). Spatial preferential attachment networks: Power laws and clustering coefficients. *Ann. Appl. Probab.* **25** 632–662. MR3313751 https://doi.org/10.1214/14-AAP1006

[13] Jacob, E. and Mörters, P. (2017). Robustness of scale-free spatial networks. *Ann. Probab.* **45** 1680–1722. MR3650412 https://doi.org/10.1214/16-AOP1098

[14] Last, G. and Penrose, M. (2017). *Lectures on the Poisson Process*. *Institute of Mathematical Statistics Textbooks* **7**. Cambridge: Cambridge Univ. Press. MR3791470

Received September 2018 and revised January 2019