Extremal linkage networks
Heydenreich, Markus; Hirsch, Christian

Published in:
Extremes

DOI:
10.1007/s10687-021-00433-3

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2022

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
Heydenreich, M., & Hirsch, C. (2022). Extremal linkage networks. Extremes, 25, 229–255. Advance online publication. https://doi.org/10.1007/s10687-021-00433-3

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment.

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 16-09-2023
Extremal linkage networks

Markus Heydenreich1 · Christian Hirsch2,3

Received: 4 June 2020 / Revised: 15 August 2021 / Accepted: 28 September 2021 / Published online: 3 December 2021
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract
We demonstrate how sophisticated graph properties, such as small distances and scale-free degree distributions, arise naturally from a reinforcement mechanism on layered graphs. Every node is assigned an a-priori i.i.d. fitness with max-stable distribution. The fitness determines the node attractiveness w.r.t. incoming edges as well as the spatial range for outgoing edges. For max-stable fitness distributions, we thus obtain a complex spatial network, which we coin extremal linkage network.

Keywords Spatial network · Random tree · Max-stable fitnesses · Coalescence · Small-world graph

AMS 2020 Subject Classifications 60G70 · 05C80 · 60K35

1 Motivation

While extreme value theory proved to be highly fruitful in the analysis of time series or unstructured data, it has more recently expanded far beyond this specific domain. Now, we see fascinating connections between extreme value theory and stochastic geometry (Calka and Chenavier 2014; Chenavier and Nagel 2019; Mayer and Molchanov 2007; Thomas and Owada 2021), as well as close interaction with the statistical analysis of networks (Wang and Resnick 2019). On the other hand, we currently witness vigorous research activity related to complex networks that come with a spatial structure. In this article, we illustrate that

1 Mathematisches Institut, Universität München, Theresienstr. 39, München 80333, Germany
2 Bernoulli Institute and CogniGron, University of Groningen, Nijenborgh 9, 9747 AG Groningen, The Netherlands
3 Department of Mathematics, Aarhus University, Ny Munkegade 118, 8000 Aarhus C, Denmark
classical methods from extreme value theory such as max-stable distributions also have a major impact in this novel research field.

Complex networks are mathematical models that are aimed at describing real-world networks such as social networks, electrical networks, the world-wide web, and others. We think of complex networks as graphs, with vertices (or nodes) representing individuals, relay stations, or websites, and edges (or links) representing material or immaterial connections between these vertices. It is also possible to rely on this framework for describing brain networks: There is a set of neurons, each of them equipped with one axon and a number of dendrites connected to axons of other neurons (see Fig. 1). Pairs of axons and dendrites may form synapses, i.e., functional connections between neurons. However, not all geometric connections necessarily also form functional connections. The resulting network can be interpreted as a directed graph with neurons as nodes and synapses as edges (directed from axon to dendrite).

1.1 Layered networks

Many complex networks are naturally organized into several layers, thereby leading to the concept of layered networks (Kivelä et al. 2014). Depending on the specific application context, the individual layers can carry very different meanings. For instance, in air transportation networks, the layers could represent airline companies offering specific flight connections (Cardillo et al. 2013). In different context, many network structures in the brain, such as the visual cortex, are organized in a layered architecture (Felleman and Van Essen 1991). Here, the lower layers perform the task of elementary transformation of sensory input. The output of the lower layers serves as input for the higher layers, which are devoted to the task of processing increasingly abstract entities. In general, it is also possible to think of layers defined in terms of properties of the nodes such as the number of incoming or outgoing connections. In this context, it has been found that the human cortical network of healthy subjects exhibits a clear layered structure, and that a loss of this organization can be an indicator for diseases such as schizophrenia (Bassett et al. 2008).

Fig. 1 An idealized, simplified representation of neurons and synaptic connections
1.2 Properties of complex networks

Most remarkably, although from very different backgrounds, many real networks share a number of ubiquitous features that are fairly universal. Therefore, any meaningful mathematical network model must reflect these features to some extent, and we shall describe them now. For this work, we specifically think of complex networks in neuroscience, see e.g. Loeffler et al. (2020), but we work out universal patterns that apply to complex networks in great generality.

To make the exposition more precise, we now introduce formally the framework for layered networks to be considered in this work. We define a random network on an infinite set of layers, each consisting of \( N \geq 1 \) nodes labeled \( \{0, \ldots, N-1\} \). Hence, the overall set of nodes is given by \( \{0, \ldots, N-1\} \times \mathbb{Z} \).

In brain networks, the layer \( h \in \mathbb{Z} \) and the label \( i \in \{0, \ldots, N-1\} \) could represent the degree of abstractness, and the location of an individual neuron in this layer, respectively. In our study, we focus on directed connections of the form \( (i, h) \rightarrow (i', h+1) \) leading from node \( i \) in layer \( h \) to node \( i' \) in layer \( h+1 \). In order to provide a more precise description of network properties, we now introduce additional terminology.

**Scale-free networks.** First, we consider the random variable

\[ D_N := \#\{i \in \{0, \ldots, N-1\} : (i, -1) \rightarrow (0, 0)\}, \]

which is the typical indegree for networks that are homogeneous in the location and layer. This homogeneity implies that replacing \( (i, -1) \rightarrow (0, 0) \) by \( (i, 0) \rightarrow (0, 1) \) would lead to the same distribution. In most of the models considered in literature (including the present work), the typical indegrees converges in distribution as \( N \rightarrow \infty \) to a limiting degree distribution \( D_\infty \). The scale-free property encodes that the limit \( D_\infty \) is Pareto-distributed, that is,

\[ \mathbb{P}(D_\infty \geq k) = k^{-\tau + o(1)} \]

for a certain parameter \( \tau > 0 \).

**Small-world networks.** Second, we let \( H_N \) denote the typical coalescence time, i.e., the first layer where two directed paths emanating from two random nodes from the initial layer \( \{0, \ldots, N-1\} \times \{0\} \) coalesce. For clarity, we note that the coalescence time is closely linked to the shortest distance between two nodes at layer 0 in a general graph. Indeed, the latter distance equals twice the coalescence time. The small-world property means that the graph distance between different network nodes are typically short (say, subpolynomial). That is, \( H_N \in o(N^\delta) \) with high probability for every \( \delta > 0 \). For instance, a number of experiments revealed that the resulting neural network is rather sparse and very well connected (Bullmore and Sporns 2009; van der Hofstad 2017). These features allow for very fast and efficient signal processing.

**Geometric clustering.** Assuming that vertices are embedded in an underlying geometric space (which could represent their physical location), geometric clustering attributes the property that geometric vicinity of vertices results in a higher probability for establishing an edge between them. For a general discussion of mathematical models of complex networks and these properties we refer to Cohen and Havlin (2010); van der Hofstad (2017).
At this point, we stress that the above definitions of $D_N$ and $H_N$ are tied to a specific data analysis procedure, taking into account that in practice, we only observe a finite number of layers. More precisely, we assume that we can observe all nodes within a range of layers so that it makes sense to select two nodes at random from layer 0. Then, the small-world property means that if the number of observed layers is at least $N^{1/3}$, then we avoid that small disconnected trees lead to undesirable noise in the estimation. The restriction to this specific data analysis procedure is one of the main limitation of our investigation, since it is not designed for situations where we only partially observe the data. Indeed, if the network dataset we observe is a connected hierarchical tree, we might understand this as a selection of nodes in the model that span from a common ancestry.

1.3 The formation of complex networks

A variety of (layered and non-layered) network models have been proposed addressing all or at least some of the features above, and these appear in contexts from a wide range of domains. It is nevertheless not clear why real-world networks share these ubiquitous features. In order to resolve this question, a key challenge is to explain the mechanism behind the formation of complex networks.

The preferential attachment model, introduced by Barabási and Albert (1999), establishes reinforcement mechanisms as an attempt to explain the universality of these features from an algorithmic point of view. Indeed, preferential attachment graphs exhibit the scale-free and small-world properties (Bollobás and Riordan 2004). Spatial versions exhibit even geometric clustering (Aiello et al. 2008; Jacob and Mörters 2013). Although the preferential attachment mechanism presents a compelling explanation for network formation, the ramifications of selecting nodes proportionally to their fitness can sometimes be prohibitively challenging to handle from a mathematical point of view. Here, the fitness can be thought of as some general mark associated with a vertex that can influence the process of network formation.

We focus on an activity-based reinforcement model inspired from synaptic plasticity in neuroscience: In an early stage, there is a (theoretical) all-to-all geometrical connectivity. Stimulation and transmission of signals enhance certain touches to ultimately form functional connections, which results in a sparse network of actual synapses. This describes brain plasticity at an early development stage. It is clear that the actual formation of the brain involves highly complex processes that are beyond the scope of a rigorous treatment. Yet, we aim at clarifying which network characteristics can be explained by a simple reinforcement scheme, and which cannot. The mathematical question that we put forward is:

Can a simple reinforcement mechanism give rise to complex network properties such as the scale-free and small-world properties?

A naive modelling using so-called $(W, A)$-reinforcement models (or ‘WARM’) suggests a negative answer. In this model involving Pólya urns with graph-based competition, there are only two regimes: the strong reinforcement regime, where the
selection of the edge to be reinforced is proportional to a super-linear function of the edge weights, and the weak reinforcement regime, where the selection is proportional to a sub-linear function. In the strong reinforcement regime, the process is supported on small isolated islands (Hirsch et al. 2021; van der Hofstad et al. 2016; Holmes and Kleptsyn 2017), whereas in the weak reinforcement regime the support is on the entire graph (Couzinié and Hirsch 2021; Holmes and Kleptsyn 2017). There is thus no regime in which a subgraph with suitable properties emerges.

The situation changes dramatically when looking at a slightly different setup. In an earlier work (Heydenreich and Hirsch 2019), we investigated a WARM-type model on a layered network. On this layered network, we proved rigorously that sufficiently strong reinforcement is responsible for logarithmic distances, and thus the small-world property applies for the resulting random graph. However, one may criticize that many of the final findings in Heydenreich and Hirsch (2019) were already hard-wired exogenously into the model from the beginning. In particular, each layer is assigned a (fixed) scope, which grows exponentially in the number of layers, and this scope determines how far a vertex can connect.

This raises the question, whether it is possible to recover many of the desirable features through a much simpler game. For instance, can we replace the proportional selection by simply picking the node with the maximum fitness in a spatial neighborhood? Although at first sight, this may seem like an entirely different story, it actually approximates preferential attachment in a regime, where the selection occurs according to power-weighted fitnesses with a large exponent. Naturally, this maximality-based selection scheme opens the door towards connections with extreme value theory, and we explore this path in detail.

1.4 A network evolving from max-stable distributions

We now outline the model definition in the present work and then showcase how extreme value theory enters the stage in the form of the max-stability of the Fréchet distribution. Instead of activity-based reinforcement, like in Heydenreich and Hirsch (2019), we simplify the interaction by assigning each vertex an a-priori fitness, which has two functions: first it measures the node attractiveness in comparison with the fitnesses of neighboring nodes; secondly, it encodes how far a connection from this vertex may reach. Then, we draw one directed edge from vertex \((i, h) \in \mathbb{Z} \times \mathbb{Z}\) with fitness \(F_{i, h} > 0\) (where \(i\) is the spatial location and \(h\) is the layer) to the vertex with highest fitness in the set

\[
\{(i - \lceil F_{i, h} \rceil, h + 1), \ldots, (i + \lceil F_{i, h} \rceil, h + 1)\},
\]

see Fig. 2.

This resembles a strong reinforcement regime if the fitnesses have a max-stable distribution, which we henceforth assume. We prove the scale-free and small-world property for such networks under suitable parameters. Since the max-stable fitness distribution is instrumental both for modelling the reinforcement and also for obtaining the desired behavior, we coin this network model extremal linkage networks.
Finally, in practice, it is unrealistic to assume that the fitness distribution follows exactly a Fréchet distribution. In fact, it may even happen that different layers are equipped with different distributions. We hypothesize that the implications of such model extensions depend sensitively on how rapidly the distributions change when moving from one layer to another. Indeed if such changes are not visible on the typical scale of \( H_N \), then the small-world property should still be valid. On the other hand, a rapidly evolving fitness distribution could substantially influence the graph distances. In Heydenreich and Hirsch (2019), we could illustrate that working with inhomogeneous layers is indeed feasible.

### 2 Model and results

We define our extremal linkage network as a random layered network, where the node \( i \in \{0, \ldots, N - 1\} \) in layer \( h \in \mathbb{Z} \) has a fitness \( F_{i,h} \), where we assume the family \( \{F_{i,h}\}_{i \in \{0, \ldots, N-1\}, h \in \mathbb{Z}} \) to be independent and identically distributed (i.i.d.). We say that \((j, h + 1)\) is visible from \((i, h)\) if \( d_N(i,j) \leq \lceil F_{i,h} \rceil \), where \( d_N \) is the distance on the discrete torus \( \mathbb{Z}/N\mathbb{Z} \). Then, the number of nodes on layer \( h + 1 \) that are visible for the \( i \)th node in layer \( h \), which we call the scope of \((i, h)\), is given by \( \varphi(F_{i,h}) \wedge N \), where

\[
\varphi(f) = 1 + 2\lceil f \rceil.
\]

Now, \((i, h)\) connects to precisely one visible node \((j, h + 1)\) in layer \( h + 1 \), namely the one of maximum fitness. In other words,

\[
F_{j,h+1} = \max_{j' : d_N(i,j') \leq \lceil F_{i,h} \rceil} F_{j',h+1}.
\]

We show a realization of such an extremal linkage network in Fig. 3, which illustrates that nodes with high fitness have a larger scope. To ease notation, we write \((i, h) \rightarrow (j, h + 1)\) and think of the directed edges as arrows. In this work, we identify the asymptotic degree and distance distribution as \( N \to \infty \).

#### 2.1 Degree distribution

To study the typical indegree \( D_N \), we assume that the fitnesses are i.i.d. copies of a random variable \( F \) with tails

\[\varphi(f) = 1 + 2\lceil f \rceil.\]
for some $\delta > 0$. Here and throughout we write $f(s) \asymp g(s)$ whenever there exist constants $c_1, c_2 > 0$ such that $c_1 \leq f(s)/g(s) \leq c_2$ uniform in $s$. As $N \to \infty$, the typical indegree converges in distribution to a non-degenerate random variable $D_\infty$ with the following tail behavior.

**Theorem 1 (Degree distribution)** The typical indegree $D_N$ converges in distribution to a non-degenerate random variable $D_\infty$. Moreover,

$$\mathbb{P}(D_\infty > s) \asymp s^{-\delta}.$$  \hspace{1cm} (2)

One possible application of Theorem 1 is that it paves the way to estimate the model parameter $\delta$. We elaborate on this point in the heavy-tailed regime $\delta \leq 1$, which we deem to be the most relevant one from the perspective of complex networks. More precisely, we simulated a network of $N = 10,000,000$ nodes in the base layer and determined the indegrees $D_{0,N}, \ldots, D_{N-1,N}$ in the base layer. Then, $\hat{p}_s = N^{-1}\#\{i \leq N : D_{i,N} = s\}$ is the empirical probability mass of the nodes of indegree $s$. Then, the right panel of Fig. 4 shows that for $\delta = 1$, we see a clear linear relation between $\log \hat{p}_s$ and $s$, as predicted by Theorem 1.
On the other hand, for $\delta < 1$, Theorem 1 suggests that the dependence between $\log \hat{p}_s$ and $s$ should be considered on a log-log scale. For $\delta = 0.5$, this is clearly visible from Fig. 4, and also for $\delta = 0.75$ it is plausible after disregarding the small degrees. In fact, the proof of Theorem 1 shows that asymptotically the slope of the log-log plot equals $-(2 - \delta)/(1 - \delta)$, so that also in the very heavy-tailed regime our results can be leveraged to extract the model parameter $\delta$ from data.

2.2 Distances

Since from each node, we draw precisely one arrow to the next layer, extremal linkage networks do not contain cycles. Hence, such a network is a collection of random trees. More precisely, since eventually the paths started from any two nodes become connected at a sufficiently high layer, this collection is in fact a single random tree.

Henceforth, we assume the fitnesses to follow a Fréchet distribution with tail index $\delta > 0$. That is,

$$\mathbb{P}(F \leq s) = \exp(-s^{-\delta}). \quad (3)$$

In particular, $G = \log(2F^\delta)$ follows a Gumbel distribution with mean $\log(\mu) := \mathbb{E}[G] > 0$. We recall that a family of random variables $\{X_i\}_I$ is tight if for every $\epsilon > 0$ there exists a compact set $K$ such that $\sup_{i \in I} \mathbb{P}(X_i \notin K) < \epsilon$.

**Theorem 2 (Distances)** If the fitness distribution is given by (3), then the typical distance $H_N$ is almost surely finite for every $N \geq 1$. Moreover,

1. if $\delta = 1$, then asymptotically almost surely,

   $$\frac{H_N}{\log_{\mu}(N)} \xrightarrow{N \to \infty} 1;$$

2. if $\delta < 1$, then $\{H_N - \log_{1/\delta} \log(N)\}_{N \geq 1}$ is tight in $\mathbb{R}$;

3. if $\delta \in (1, 2)$, then $\{H_N/N^\delta\}_{N \geq 1}$ is tight in $(0, \infty)$;

4. if $\delta > 2$, then $\{H_N/N^2\}_{N \geq 1}$ is tight in $(0, \infty)$.

Our results show that scale-free and small-world behavior is present if $\delta < 1$. A softer version of small-world behavior (with logarithmic distances) is present in the border case $\delta = 1$. Geometric clustering is incorporated through the network mechanism.

The proof for $\delta > 2$ is based on the central limit theorem, while the case $\delta \in (1, 2)$ is based on a corresponding stable limit theorem. It is plausible that an analog result is true for the case $\delta = 2$, however, logarithmic corrections might appear. We did not pursue this further, because from a modeling perspective our results are most interesting for $\delta \leq 1$. 

Springer
2.3 Organization

We prove the asymptotic degree distribution of Theorem 1 in Sect. 3. For the distance result in Theorem 2, we give separate proofs for the lower and upper bound in Sects. 4 and 5, respectively. In order to improve the readability, we collect some of the technical proofs in an appendix.

3 Degrees

As a first step, we describe the limiting distribution $D_\infty$. This description rests on the observation that, as $N \to \infty$, the $N$-torus $\mathbb{Z}/N\mathbb{Z}$ converges locally to $\mathbb{Z}$. More precisely, consider an i.i.d. family $\{F_{i,h}\}_{(i,h)\in \mathbb{Z}\times \mathbb{Z}}$ of Fréchet random variables with distribution as in (2). Then, we let

$$L := \sup\{i \leq 0 : F_{i,0} > F_{0,0}\} \quad \text{and} \quad R := \inf\{i \geq 0 : F_{i,0} > F_{0,0}\}$$

denote the locations of the first points to the left and to the right of $(0, 0)$ with a higher fitness. The vertex $(i, -1)$ with $i \in [L, R]$ is connected to $(0, 0)$ whenever its fitness $F_{i,-1}$ satisfies $|F_{i,-1}| \geq i$ (i.e., $(0, 0)$ is in the scope) but $|F_{i,-1}| < |i - L|$ and $|F_{i,-1}| < |i - R|$ (i.e., neither $(L, 0)$ nor $(R, 0)$ are in the scope). Hence the set of in-neighbors of $(0, 0)$ is given by

$$V := \{i \in [L, R] : F_{i,-1} \in I_i(L, R)\}$$

with $I_i(L, R) := (|i| - 1, (-L + i - 1) \land (R - i - 1)]$. In particular, the indegree at the origin in the unbounded model equals $D_\infty := |V|$.

Proof (Theorem 1) We couple the infinite model with a model on the $N$-torus based on fitnesses $\{F_{i,h}^N\}_{i \in \{0, \ldots, N-1\}, h \in \mathbb{Z}}$ defined through $F_{i,h}^N := F_{i,h}$ if $i \leq N/2$ and $F_{i,h}^N := F_{i-N,h}$ otherwise. Then, under the event that $\{R \leq N/4\} \cap \{L \geq -N/4\}$, we see that the indegree in the torus model coincides with the indegree in the infinite model. That is,

$$\mathbb{P}(D_\infty \neq D_N) \leq \mathbb{P}(L \leq -N/4) + \mathbb{P}(R \geq N/4),$$

and the right-hand side tends to 0 as $N \to \infty$. Then, a coupling argument shows that $D_N$ converges in distribution to $D_\infty$. Equipped with this knowledge, we now establish the tail behavior of $D_\infty$. We only establish the upper bound because the arguments for the lower bound are similar.

First, we note that $D_\infty \leq D_\infty^+$, where $D_\infty^+ := \sum_{i \in [L, R]} Y_i$ is a sum of independent Poisson random variables with $\mathbb{P}(Y_i \neq 0) = p_i(L, R) := \mathbb{P}(F_{0,0} \in I_i(L, R))$. Then, conditioned on $(L, R)$, also $D_\infty^+$ is a Poisson random variable and has parameter $\lambda(L, R) := -\sum_{i \in [L, R]} \log(1 - p_i(L, R))$.

If $\delta > 1$, then $\sup_{r \geq 1} \lambda(l, r) < \infty$ so that the asserted tail behavior follows from the Poisson concentration inequality [Penrose 2003, Lemma 1.2].
If \( \delta \leq 1 \), then we leverage the bound

\[
\mathbb{P}(D^+_\infty > s) \leq \mathbb{P}(\lambda(L, R) > 3s/4) + \mathbb{P}(\lambda(L, R) \leq 3s/4, D^+_\infty > s),
\]

where by the Poisson concentration inequality [Penrose 2003, Lemma 1.2], the second summand decays exponentially fast in \( s \). We note that \( p_i(L, R) = 0 \) for \( i < L/2 \) or \( i > R/2 \). In particular, since \( -\log(1 - x) \leq 1.01x \) for sufficiently small \( x > 0 \), there exists some large \( s_0 > 0 \) such that for all \( s \geq s_0 \) we have that \( p(L, R) \geq s/2 \) whenever \( \lambda(L, R) \geq 3s/4 \), where we set \( p(L, R) := \sum_{i \in \{L/2, R/2\}} p_i(L, R) \). Setting \( a(x) := \sum_{1 \leq i \leq \lfloor x/2 \rfloor} \mathbb{P}(F_{0,0} > j) \), we also obtain that \( p(L, R) \leq 1 + a(-L) + a(R) \). In particular, for all large \( s > 0 \),

\[
\mathbb{P}(\lambda(L, R) \geq 3s/4) \leq \mathbb{P}(a(R) \geq s/5) + \mathbb{P}(a(-L) \geq s/5) = 2\mathbb{P}(a(R) \geq s/5).
\]

We now claim that \( \mathbb{P}(R \geq m) = 1/m \). Indeed, the event \( \{R \geq m\} \) encodes that the fitness at the origin is largest among the fitnesses of the first \( m \) nodes. That is,

\[
\{R \geq m\} = \{F_{0,0} = \max_{i \leq m-1} F_{i,0}\}.
\]

Then, since fitnesses are identically distributed,

\[
\mathbb{P}(F_{0,0} = \max_{i \leq m-1} F_{i,0}) = \mathbb{P}(F_{1,0} = \max_{i \leq m-1} F_{i,0}) = \cdots = \mathbb{P}(F_{m-1,0} = \max_{i \leq m-1} F_{i,0}) = \frac{1}{m},
\]

so that \( \mathbb{P}(R \geq m) = 1/m \) as claimed. Now, for \( \delta = 1 \), we have \( \log(a^{-1}(r)) \asymp r \), where \( a^{-1} \) denotes the generalized inverse of \( a \), so that for all large \( s > 0 \),

\[
\log \mathbb{P}(\lambda(L, R) > 3s/4) \leq \log \left( 2\mathbb{P}(\log(R) > \log(a^{-1}(s/5))) \right) \asymp -s,
\]

as asserted. Leveraging \( \log(a^{-1}(r)) \asymp \log(r) \), a similar argument concludes the proof for \( \delta < 1 \).

### 4 Distances – Lower bounds

To prove the lower bounds in Theorem 2, we relate the graph distance \( H_N \) to the coalescence of two walkers. By symmetry, we may assume one of the randomly chosen nodes in the initial layer to be at position 0. Next, let \( X_{h}^{\leq N} \in \{0, \ldots, N - 1\} \) denote the position after \( h \geq 0 \) steps of a walker starting from 0 and following the arrows. More precisely, we put recursively \( X_{0}^{\leq N} = 0 \), and then \( X_{h+1}^{\leq N} \in \{0, \ldots, N - 1\} \) such that

\[
F_{h+1}^{\leq N} := F_{X_{h+1}^{\leq N}} = \max_{i : d_h(i, X_{h+1}^{\leq N}) \leq \lfloor F_{i,h+1}^{\leq N} \rfloor} F_{i,h+1}.
\]

Additionally,

\[
G_{h}^{\leq N} := \log_{p}(2F_{h}^{\leq N})
\]
denotes the log-fitness. Similarly to \( \{X^L_h\}_{h \geq 0} \), we let \( \{X^R_h\}_{h \geq 0} \) denote the walker started from a uniformly chosen random position \( X^R_0 \) in layer 0.

To prove the lower bound of Theorem 2, note that for every \( h \geq 0 \) and \( \eta > 0 \),
\[
\mathbb{P}(H_N \leq h) \leq \mathbb{P}(d_N(0, X_h^L) \geq \eta N) + \mathbb{P}(d_N(0, X_h^R) \geq \eta N) + \mathbb{P}(d_N(0, X^R_0) \leq 2\eta N)
\]
Since \( X^R_0 \) is uniformly distributed in an \( N \)-torus, we see that \( \mathbb{P}(d_N(0, X^R_0) \leq 2\eta N) \leq 4\eta \), so that together with \( \mathbb{P}(d_N(0, X_h^L) \geq \eta N) = \mathbb{P}(d_N(0, X^R_0, X_h^R) \geq \eta N) \), we arrive at
\[
\mathbb{P}(H_N \leq h) \leq 2\mathbb{P}(d_N(0, X_h^L) \geq \eta N) + 4\eta. \tag{4}
\]

For \( \delta \geq 1 \), we establish highly accurate upper bounds on the growth of the fitnesses \( F_h^{L,N} \) as a function of the layer \( h \). Consequently, we also obtain bounds on the location
\[
d_N(0, X_h^L) \leq F_0^{L,N} + \cdots + F_{h-1}^{L,N} + h \tag{5}
\]
after \( h \) steps, where the addition of \( h \) on the right hand side arises from rounding.

Using the same fitnesses attached to the nodes yields a natural coupling between the model on a finite torus and the limit model on the integers \( \mathbb{Z} \). We write \( \{F_h^\ell\}_{h \geq 0} \) for the fitnesses in this limit model. As long as \( \varphi(F_i^{L,N}) \leq N \), the wrapping around the torus is not observable so that, for \( \eta \in (0, 1) \),
\[
F_0^{L,N} + \cdots + F_{h-1}^{L,N} + h \geq \eta N
\]
if and only if in the coupled limit model \( F_0^\ell + \cdots + F_{h-1}^\ell + h \geq \eta N \). Hence, it suffices to study the limit model.

### 4.1 Proof for \( \delta = 1 \).

To bound the right-hand side in (5), we show that the log-fitnesses concentrate sharply around the current layer \( h \).

**Lemma 1 (Fluctuations of log-fitnesses)** For \( \delta = 1 \) and \( G_h^\ell := \log_\mu(2F_h^\ell) \),
\[
\mathbb{P}\left( \limsup_{h \to \infty} h^{-2/3}\left| G_h^\ell - h \right| = 0 \right) = 1.
\]

We prove Lemma 1 in Sect. 1 of the appendix. We now explain how to derive from it the lower bound on the distances.

**Proof** (Theorem 2; lower bound; \( \delta = 1 \)) Let \( I \geq 0 \) be an almost surely finite random variable such that, by Lemma 1, \( \left| G_h^\ell - h \right| \leq h^{2/3} \) holds for all \( h \geq I \). Then, inserting into (5),
\[ \mathbb{P}\left(h + \sum_{i < h} F^c_i \geq \eta N \right) \leq \mathbb{P}\left( \sum_{i \leq h} F^c_i > \eta N/2 \right) + \left\{ h + \sum_{i \leq h} \mu^{i + 2/3} > \eta N/2 \right\} \]
\[ \leq \mathbb{P}\left( \sum_{i \leq h} F^c_i > \eta N/2 \right) + \left\{ h + h\mu^{h + 2/3} > \eta N/2 \right\}. \]

The probability on the right hand side vanishes as \(N \to \infty\). Further, we fix \(\epsilon' > 0\) and let \(h = \log_\mu(N)(1 - \epsilon')\), then the indicator vanishes as well in the limit \(N \to \infty\) for arbitrary \(\eta > 0\). Inserting this into (4) proves the claim.

4.2 Proof for \(\delta < 1\).

In the heavy-tailed setting, we need a strong tightness property for the log-fitnesses.

Lemma 2 (Tightness for heavy tails) For \(\delta < 1\),
\[ \mathbb{P}\left( \lim_{h \to \infty} \delta^h G^c_h \in (0, \infty) \right) = 1. \]

We prove Lemma 1 in Sect. 1 of the appendix. We now explain how Lemma 2 enters the proof of the lower bound.

Proof (Theorem 2; lower bound; \(\delta < 1\)) Recall that \(\log(\mu) := \mathbb{E}[G]\). Let \(I\) be a strictly positive random variable such that \(\delta^h G^c_h \leq I\) holds almost surely for all \(h \geq 0\). Fix \(K \geq 1\) and write \(h = \log(N) - K\). If \(h \geq 0\), then inserting the bound from Lemma 2 into the representation from (5) gives that
\[ \mathbb{P}(h + \sum_{j < h} F^c_j > \eta N) \leq \mathbb{P}(h + \sum_{j \leq h} \mu^{j + 1} > \eta N) \leq \mathbb{P}(h + h\mu^{h + 1} > \eta N) = \mathbb{P}(h(1 + N^{\delta^h I}) > \eta N). \]

Now, \(\limsup_{N \to \infty} \mathbb{P}(h(1 + N^{\delta^h I}) > \eta N) \leq \mathbb{P}(\delta^h I \geq 1/2)\), so that we conclude the proof by noting that \(\mathbb{P}(\delta^h I \geq 1/2)\) tends to 0 as \(K \to \infty\).

4.3 Proof for \(\delta > 2\).

In the light-tailed setting, we show that the suitably rescaled walker \(\{X^c_h\}_{h \geq 0}\) satisfies the invariance principle.

Lemma 3 (Invariance principle) Let \(\delta > 2\). Then, \(\{h^{-1/2}X^c_h\}_{r \leq 1}\) converges in distribution as \(h \to \infty\) to some Brownian motion \(\{B_t\}_{t \leq 1}\).

Throughout we write \(ht\) for \([ht]\). Before establishing Lemma 3, we explain how to conclude the proof of the lower bound. Mind that, for the lower bound, a central limit theorem suffices. However, for the proof of the upper bound in the next section, we need a full functional CLT.

\[\includegraphics[width=\textwidth]{M. Heydenreich, C. Hirsch.png} \]
Proof of Theorem 2; lower bound; \(\delta > 2\) The invariance principle in the form of Lemma 3 for \(t = 1\) gives
\[
\lim_{\epsilon \to 0} \lim_{N \to \infty} \mathbb{P}(X_{\epsilon N^2}^{c} \geq \epsilon^{1/4} N) = \lim_{\epsilon \to 0} \lim_{N \to \infty} \mathbb{P}(X_{\epsilon N^2}^{c} (\epsilon N^2)^{-1/2} \geq \epsilon^{-1/4}) = 0,
\]
as asserted. \(\square\)

In order to prove Lemma 3, we rely on the general martingale functional CLT from [Meyn and Tweedie 2009, Theorem D.6.4]. To cast this problem in the setting of the present context, we let
\[
\mathcal{F}_h = \sigma(\{F_j^c, X_j^c\}_{j \leq h}).
\]
denote the information provided by the positions of the walker and the corresponding fitnesses up to layer \(h\). Then, \(M_h : = X_h^c\) form a square-integrable martingale with respect to the filtration \(\{\mathcal{F}_h\}_{h \geq 0}\). In order to apply [Meyn and Tweedie 2009, Theorem D.6.4], we need to verify two conditions.

(M1) Almost surely,
\[
\lim_{h \to \infty} \frac{1}{h} \sum_{j \leq h} \mathbb{E}[(M_j - M_{j-1})^2 \mid \mathcal{F}_{j-1}] = \gamma^2,
\]
for some constant \(0 < \gamma^2 < \infty\).

(M2) Almost surely, for every \(\epsilon > 0\),
\[
\lim_{h \to \infty} \frac{1}{h} \sum_{j \leq h} \mathbb{E}[(M_j - M_{j-1})^2 \mathbb{1}\{ (M_j - M_{j-1})^2 \geq \epsilon h \} \mid \mathcal{F}_{j-1}] = 0.
\]

Note that, when fixing any \(\epsilon_0 > 0\), condition (M2) follows from the following Lyapunov-type condition.

(M2’) Almost surely,
\[
\lim_{h \to \infty} \frac{1}{h^{1+\epsilon_0}} \sum_{j \leq h} \mathbb{E}[(M_j - M_{j-1})^{2+\epsilon_0} \mid \mathcal{F}_{j-1}] = 0.
\]

We verify conditions conditions (M1) and (M2’) in Sect. 1 in Appendix A.

4.4 Proof for \(\delta \in (1, 2)\)

Finally, we deal with the stable case, i.e., \(1 < \delta < 2\). In the light-tailed setting, a key ingredient was the invariance principle in the form of Lemma 3, which stated that the rescaled walker \(\{h^{-1/2} X_h^c\}_{h \in [0,1]}\) converges to Brownian motion as \(h \to \infty\). Now, we need a stable analog of this result. More precisely, we establish convergence to a symmetric \(\delta\)-stable process with Lévy measure
\[
\frac{\nu(dx)}{dx} = c(\delta) |x|^{\delta+1} \mathbb{1}\{x \neq 0\},
\]
for some \(c(\delta) > 0\).
Lemma 4 (Stable limit) Let $\delta \in (1, 2)$. Then, $\left\{ h^{-1/\delta} X_{ht}^\varepsilon \right\}_{t \leq 1}$ converges in distribution to symmetric $\delta$-stable processes.

Before establishing Lemma 4, we elucidate how it gives the tightness of $(H_N/N^\delta)$ away from 0. Essentially, this relies on the same line of arguments that we have seen in Sect. 4.3.

Proof (Theorem 2; lower bound; $\delta \in (1, 2)$) Invoking Lemma 4 gives that

$$\lim_{\varepsilon \to 0} \lim_{N \to \infty} \Pr(X_{tN^\delta}^\varepsilon \geq \varepsilon^{1/(2\delta)} N) = \lim_{\varepsilon \to 0} \lim_{N \to \infty} \Pr(X_{tN^\delta}^\varepsilon (\varepsilon N^\delta)^{-1/\delta} \geq \varepsilon^{-1/(2\delta)}) = 0,$$

as asserted.

In order to prove Lemma 4, we proceed as in Gayrard and Hartung (2019) and apply the versatile functional limit theorem [Durrett and Resnick 1978, Theorem 4.1]. We state the corresponding conditions here, and verify them in Sect. 1 in Appendix A.

We now formulate a version of [Durrett and Resnick 1978, Theorem 4.1], where we adapted (actually simplified) the conditions to our needs: Let $\{Z_{j,h}\}_{1 \leq j \leq h}$ be a triangular array of centered random variables and let $\{F_{j,h}\}_{1 \leq j \leq h}$ be a triangular array of $\sigma$-algebras such that $Z_{j,h}$ is $F_{j,h}$ measurable. Now, assume the following conditions:

- (D1) There exists a symmetric measure $\nu$ such that for all $x > 0$ and $t \leq 1$,

$$\sum_{j \leq h} \Pr(\{Z_{j,h} > x \mid F_{j-1,h}\} \to t\nu([x, \infty))) \text{ in probability as } h \to \infty.$$

- (D2) For $\varepsilon > 0$,

$$\sum_{j \leq h} \Pr(|Z_{j,h}| > \varepsilon \mid F_{j-1,h})^2 \to 0 \text{ in probability as } h \to \infty.$$

- (D3) For $\eta, \varepsilon > 0$,

$$\lim_{\eta \to 0} \lim_{h \to \infty} \Pr\left( \sum_{j \leq h} \mathbb{E}\left[Z_{j,h}^2 \mathbb{I}\{|Z_{j,h}| \leq \eta\} \mid F_{j-1,h}\right] > \varepsilon \right) = 0.$$

Then, $\left\{ \sum_{j \leq h} Z_{j,h} \right\}_{1 \leq 1}$ converges in distribution as $h \to \infty$ to a symmetric stable process with Lévy measure $\nu$.

5 Distances – Upper bound

For the lower bound on the distances in Sect. 4, it was sufficient to control the deviation of a single walker. Establishing the upper bound is substantially more involved, as we need to understand the joint movements of the left and the right walker. To lighten notation, we omit the torus size $N$ in the quantities $X^L_N$, $X^R_N$, $F^L_N$ and $F^R_N$. 

Springer
5.1 Proof for $\delta < 1$

We begin by discussing the heavy-tailed setting, as the argument is particularly short. The reason herefor lies in the rapid growth of the fitnesses following from recursion (10). In particular, the upper bound $N$ for the fitnesses becomes absorbing: after reaching it, it remains there for a long period of time.

**Lemma 5 (Absorbing upper fitness bound)** For $\delta < 1$,

$$\lim_{N \to \infty} \Pr(\varphi(F^L_h) \geq N > \varphi(F^L_{h+1}) \text{ for some } h \leq N) = 0.$$  

**Proof** Although the recursion leading to the bound (10) refers to the limiting model where the torus is replaced by the integers, we obtain a finite-volume version by the same arguments together with an additional truncation:

$$F^L_{h+1} = (\langle \varphi(F^L_h) \wedge N \rangle F^L_{h+1})^{1/\delta}.$$  

(7)

Thus,

$$\Pr(\varphi(F^L_h) \geq N > \varphi(F^L_{h+1})) \leq \Pr(F_{h+1} < N^{\delta-1}) = \exp\left(-N^{1-\delta}\right),$$

so that invoking the union bound over $h \leq N$ concludes the proof. \hfill \qed

Equipped with this auxiliary result, we now establish the tightness asserted in Theorem 2.

**Proof** (Theorem 2; upper bound; $\delta < 1$) First, the walkers coalesce with certainty once the scopes reach $N$. Hence, by Lemma 5, it suffices to show that

$$\lim_{K \to \infty} \lim_{N \to \infty} \sup \Pr(\sup_{h \log_{1/\delta}\log\mu(N)+K} \varphi(F^L_h) < N) = 0.$$  

Now, as long as $\varphi(F^L_h) \leq N$, the fitness in the torus model coincides with the one in the infinite limit. Next, by Lemma 2, the fitnesses in the limit model grow as $\mu^{Z\delta^{-h}}$ for a positive random variable $Z$. We conclude by noting that

$$\lim_{K \to \infty} \sup_{N \geq 1} \Pr(\mu^{Z\delta^{-\log_{1/\delta}\log\mu(N)+K}} < N) = \lim_{K \to \infty} \Pr(Z\delta^{-K} < 1) = 0.$$  

\hfill \qed

5.2 Proof for $\delta = 1$

To prove the upper bound for $\delta = 1$, we first show that, with high probability, the walker’s fitness is close to $N$ after at most
\[ \tau_0 := \tau_0(N) := \log_\mu(N) + \log_\mu(N)^{7/8} \]

steps.

**Lemma 6 (Lower bound on fitnesses)** For \( \delta = 1 \),

\[
\lim_{N \to \infty} \mathbb{P}\left( \inf \{ F^L_h : \tau_0(N) \leq h \leq 2 \log_\mu(N) \} \geq N \exp(- \log_\mu(N)^{3/4}) \right) = 1.
\]

The second ingredient is a specifically constructed coupling between \( \{((X^L_{h,i}, F^L_{h,i}),(X^R_{h,i}, F^R_{h,i}))\}_{h \geq 0} \) and independent walkers \( \{((X^L_{h,i}, F^L_{h,i}),(X^R_{h,i}, F^R_{h,i}))\}_{h \geq 0} \), the latter moving w.r.t. two independent copies of the fitnesses \( (F_i)_{i,h} \). We write

\[
E^\text{suc} := \left\{ \left( (X^L_{h,i}, F^L_{h,i}),(X^R_{h,i}, F^R_{h,i}) \right) = \left( (X^L_{h,i}, F^L_{h,i}),(X^R_{h,i}, F^R_{h,i}) \right) \right\},
\]

and

\[
E^\text{fail} := \left\{ \left( (X^L_{h,i}, F^L_{h,i}),(X^R_{h,i}, F^R_{h,i}) \right) \neq \left( (X^L_{h,i}, F^L_{h,i}),(X^R_{h,i}, F^R_{h,i}) \right) \right\}
\]

for the events that the coupling succeeds, respectively fails at level \( h \). Moreover, let

\[
\mathcal{F}^\text{coup}_h := \sigma \left( (X^L_{i}, F^L_{i}, X^R_{i}, F^R_{i}) \right)_{i \leq h}
\]

denote the \( \sigma \)-algebra of information on the coupled walkers up to level \( h \). Note that \( F^L_h = F^R_h \) if coalescence occurs at level \( h \), whereas \( F^L_h \neq F^R_h \) almost surely, by absolute continuity of the fitnesses. Therefore, \( \{X^L_h = X^R_h\} \subset E^\text{fail}_h \). The crux of the coupling is that, whenever it fails, the walkers coalesce with probability at least 1/4.

**Lemma 7 (Coupling with independent walkers)** There is a coupling between the true walkers \( \{((X^L_{h,i}, F^L_{h,i}),(X^R_{h,i}, F^R_{h,i}))\}_{h \geq 0} \) and independent walkers \( \{((X^L_{h,i}, F^L_{h,i}),(X^R_{h,i}, F^R_{h,i}))\}_{h \geq 0} \) such that almost surely on the event \( E^\text{suc}_h \),

\[
\mathbb{P}(X^L_{h+1} = X^R_{h+1} \mid \mathcal{F}^\text{coup}_h) \geq \frac{1}{4} \mathbb{P}(E^\text{fail}_h \mid \mathcal{F}^\text{coup}_h).
\]

Finally, we show that for the independent walkers starting from fitnesses at least \( N \exp(- \log_\mu(N)^{3/4}) \), with high probability \( F^L_h = F^R_h = N \) for some \( h \leq \log_\mu(N)^{7/8} \).

**Lemma 8 (Absence of long excursions)** For \( \delta = 1 \) and every \( \varepsilon > 0 \) there exists \( N^* = N^*(\varepsilon) \) such that if \( N \geq N^* \), then

\[
\mathbb{P}(F^L_{h,i} = F^R_{h,i} = N \text{ for some } h \leq \log_\mu(N)^{7/8} \mid F^R_0, F^L_0) \geq 1 - \varepsilon
\]

holds almost surely on the event

\[
E = E(N) = \{ \min \{ F^R, F^L \} \geq N \exp(- \log_\mu(N)^{3/4}) \}.
\]

To understand Lemma 8, we recall that, similarly as in the case \( \delta < 1 \), the fitnesses are absorbed by \( N \). We prove Lemmas 6, 7 and 8 in Sect. 1 in Appendix B. Now, we complete the proof of Theorem 2.
Extremal linkage networks

\textbf{Proof} (Theorem 2; upper bound; \(\delta = 1\)) By Lemma 6, we may assume that \(\min\{T_h^L, T_h^R\} \geq N \exp(-\log_\mu(N)^{3/4})\) for all \(\tau_0 \leq h \leq 2 \log_\mu(N)\). Now, if the coupling fails without coalescence at time \(h \geq \tau_0\), then we restart it by initializing \((X_h^L,i), X_h^R\) at time \(h + 1\) with the values of the true system \((X^L_h, X^R_h)\). Proceeding recursively, this produces stopping times \(\{\tau^*_f\}_{i \geq 1}\) encoding the sequence of coupling failures, where we impose that \(\tau_0^* = \tau_0\). To ease notation, we let the sequence \(\tau_i^*\) be constant after the index \(i_0\) where \(\tau_i^* = H_N\), i.e., where coalescence occurs.

In particular, applying Lemma 8 with \(\varepsilon/K^*\) in place of \(\varepsilon\), we obtain that
\[
\mathbb{P}(H_N \geq \log_\mu(N) + K^* \log_\mu(N)^{7/8}) \leq \mathbb{P}(\tau_{K^*}^* \neq H_N) + \sum_{i \leq K^*} \mathbb{P}(\tau_i^* - \tau_{i-1}^* \geq \log_\mu(N)^{7/8})
\]
\[
\leq \mathbb{P}(\tau_{K^*}^* \neq H_N) + \varepsilon
\]
holds for all sufficiently large \(N\). We now claim that \(\mathbb{P}(\tau_i^* \neq H_N) \leq 3/4\). Once this claim is shown, we deduce that \(\mathbb{P}(\tau_{K^*}^* \neq H_N) \leq (3/4)^{-K^*}\) since we restart the coupling after each failure. Finally, by Lemma 7,
\[
\mathbb{P}(\tau_1^* = H_N) = \sum_{h \geq \tau_0} \mathbb{P}(X_{h+1}^L = X_{h+1}^R, E_{h}^{\text{succ}}) \geq 1/4 \sum_{h \geq \tau_0} \mathbb{P}(E_{h+1}^{\text{all}}, C_h^{\text{succ}}) = 1/4 \mathbb{P}(\tau_1^* < \infty) = 1/4,
\]
as asserted. \(\square\)

5.3 \textbf{Proof for} \(\delta \in (1, 2) \cup (2, \infty)\)

For \(\delta > 2\) we think of \(\{X_h^L - X_h^R\}_{h \geq 0}\) as a centered random walk whose step size admits a finite second moment, so that we obtain convergence to Brownian motion just as in Lemma 3. For \(\delta \in (1, 2)\) we proceed similarly, but now with the stable limit law (Lemma 4) rather then the invariance principle of Lemma 3. Since \(\{X_h^L\}_{h \geq 0}\) and \(\{X_h^R\}_{h \geq 0}\) are independent, we conclude from Lemma 4 that also \(\{N^{-1/\delta} (X_h^L - X_h^R)\}_{h \leq 1}\) converges in distribution to a \(\delta\)-stable process.

\textbf{Proof} (Theorem 2; upper bound; \(\delta \in (1, 2) \cup (2, \infty)\)) We write
\[
T_N := \{h \geq 0 : X_h^R \leq X_h^L\}
\]
for the first time, where the left walker moves past the right one. Note that when the left walker moves past the right one, then necessarily the scopes must intersect and there is a positive probability of coalescence. Thus, it suffices to derive bounds on \(T_N\). To that end, note that the distance \(|X_0^L - X_0^R|\) of the initial locations is at most \(N\).

Now, we write \(T_N\) for the first time that Brownian motion, respectively a symmetric \(\delta\)-stable process exceeds \(1\), so that
\[
\limsup_{N \to \infty} \mathbb{P}(T_N \geq KN^{2\lambda\delta}) \leq \mathbb{P}(T \geq K).
\]
Finally, we leverage that hitting times are almost surely finite, which is true not only for Brownian motion but also for the symmetric $\delta$-stable process because it is recurrent when $\delta \in (1, 2)$, see [Bertoin 1996, Theorem I.17]. Therefore, the right-hand side tends to 0 as $K \to \infty$.

\[
\square
\]

Appendix A: Proofs of auxiliary results from Sect. 4

The key towards obtaining the bounds on the scopes in Lemma 1 is the max-stability of the Fréchet distribution: if $F_1, \ldots, F_m$ are i.i.d. Fréchet random variables with tail index $\delta = 1$, then $\max\{F_1, \ldots, F_m\}$ has the same distribution as $mF$, where $F$ is again a Fréchet random variable with tail index 1. Moreover, $G_i = \log_\mu(2F_i)$ follows a Gumbel distribution.

In particular, writing $m = \varphi(F^c_h)$ (recall (1) for the definition of $\varphi$) and $F = F_{h+1}$, we represent $F^c_{h+1}$ recursively as

\[
F^c_{h+1} = \varphi(F^c_h)F_{h+1},
\]

so that

\[
0 \leq G^c_{h+1} - G^c_h - G_{h+1} \leq \log_\mu \left( \frac{1 + 2[F^c_h]}{2F^c_h} \right) \leq \log_\mu(\rho_h),
\]

where $\rho_h := 1 + \frac{3}{2F^c_h}$. Starting from this observation, we now prove Lemma 1.

**Proof** (Lemma 1) In order to develop an intuition for the proof, we first establish the asserted lower bound on the growth. That is,

\[
P\left( \liminf_{h \to \infty} h^{-2/3}(G^c_h - h) \geq 0 \right) = 1.
\]

Indeed, applying the bound (9),

\[
G^c_h - h \geq \sum_{i \leq h} (G_i - 1),
\]

where $\{G_h - 1\}_{h \geq 0}$ is an i.i.d. sequence of centered random variables with finite exponential moments. Hence, by moderate deviations [Rassoul-Agha and Seppäläinen 2015, Theorem 11.2], almost surely,

\[
\lim_{h \to \infty} h^{-2/3} \sum_{j \leq h} (G_j - 1) = 0.
\]

Moreover, now the lower bound on the growth of $G^c_h$ implies that the error terms of the form $h^{-2/3} \sum_{j \leq h} \log_\mu(\rho_j)$ in (9) tend to 0 as $h \to \infty$, thereby concluding the proof. \[
\square
\]
Next, we show Lemma 2. The proof mimics the arguments presented in Lemma 1. Therefore, we present in detail only those arguments that are substantially different. The key identity now reads

\[ F_{h+1}^{c} = \left( \varphi(F_{h}^{c})F_{h+1} \right)^{1/\delta}, \]  

so that, as in (9),

\[ 0 \leq \delta G_{h+1}^{c} - G_{h}^{c} - G_{h+1}^{c} \leq \log_{\mu}(\rho_{h}), \]

where \( \rho_{h} := 1 + \frac{3}{2F_{h}^{c}} \).

**Proof (Lemma 2)** First, we iterate (11) to get that for every \( h_{2} \geq h_{1} \geq 1, \)

\[ 0 \leq \delta^{h_{2}} G_{h_{2}}^{c} - \delta^{h_{1}} G_{h_{1}}^{c} - \sum_{h_{1} \leq j < h_{2}} \delta^{j} G_{j+1}^{c} \leq \sum_{h_{1} \leq j < h_{2}} \delta^{j} \log_{\mu}(\rho_{j}). \]  

(12)

Now, the key step is to show that

\[ \mathbb{P}\left( \liminf_{h \to \infty} \delta^{h} G_{h}^{c} > 0 \right) = 1. \]  

(13)

Then, almost surely,

\[ \lim_{h \to \infty} \sup_{h_{2} \geq h_{1} \geq h} \sum_{h_{1} \leq j < h_{2}} \delta^{j} \log_{\mu}(\rho_{j}) = 0. \]

Moreover, by the Borel-Cantelli lemma also

\[ \sup_{h_{2} \geq h_{1} \geq h} \sum_{h_{1} \leq j < h_{2}} \delta^{j} G_{j+1}^{c} \]

tends to 0 almost surely as \( h \to \infty \). Hence, \( \delta^{h} G_{h}^{c} \) converges to an almost surely finite limit.

It remains to show (13). To achieve this goal, we assert that there exists an almost surely finite random variable \( I \) such that

\[ \min \left\{ G_{I}^{c}, \sum_{j \geq I} \delta^{j} G_{j+1}^{c} \right\} > 0. \]

(14)

Once (14) is established, we obtain that

\[ \delta^{h} G_{h+1}^{c} \geq \delta^{I} G_{I}^{c} + \sum_{I \leq j < h+1} \delta^{j} G_{j+1}^{c} \geq \sum_{I \leq j < h+1} \delta^{j} G_{j+1}^{c}, \]

so that taking the limit as \( h \to \infty \) concludes the proof.

To prove (14), we may first apply the Borel-Cantelli lemma to see that the sum \( \sum_{j \geq 0} \delta^{j} G_{j+1}^{c} \) converges almost surely. Hence,
In particular, the Kolmogorov 0-1-law yields an almost surely finite random variable \( I \) such that
\[
\min \left\{ G_I, \sum_{j \geq 1} \delta^i G_{j+1} \right\} > 0.
\]

Since \( G_i^e \geq G_i \) for every \( i \geq 1 \), this observation concludes the proof of (14). \( \square \)

Next, we prove Lemma 3. Before establishing the invariance principle for the random walk \( \{ X^e_h \}_{h \geq 0} \), we first show that the underlying Markov chain of fitnesses \( \{ F^e_h \}_{h \geq 0} \) satisfies a Foster-Lyapunov drift condition, thereby forming the basis for a Markov-chain LLN.

**Lemma 9 (Drift condition)** Let \( \delta > 1, \beta < \delta \) and set \( V(f) = 1 + f^\beta \). Then, there exists \( K = K(\beta) > 0 \) such that for all \( f > 0 \),
\[
\mathbb{E} \left[ V(F^e_1) \mid F^e_0 = f \right] \leq \frac{1}{2} V(f) + K \mathbb{1} \{ f \leq K \}. \tag{15}
\]

**Proof** To bound \( \mathbb{E} \left[ V(F^e_1) \mid F^e_0 = f \right] \), we leverage recursion (10) to deduce that for every \( f > 3 \),
\[
\mathbb{E} \left[ (F^e_1)^{\beta} \mid F^e_0 = f \right] = \varphi(f)^{\beta/\delta} \mathbb{E} \left[ F^{\beta/\delta}_1 \right] \leq \mathbb{E} \left[ F^{\beta/\delta}_1 \right] 3^{\beta/\delta} f^{\beta/\delta}.
\]

Hence, since \( \delta > 1 \), there exists \( K > 3 \) such that
\[
\mathbb{E} \left[ (F^e_1)^{\beta} \mid F^e_0 = f \right] \leq \frac{1}{2} f^\beta
\]
holds for all \( f > K \), thereby verifying the drift inequality (15). \( \square \)

Now, we have collected all ingredients to prove the invariance principle, i.e., Lemma 3.

**Proof** (Lemma 3) **Condition (M1).** First, we note that conditioned on \( \mathcal{F}_{j-1} \), the random variable \( M_j = X^e_j \) is uniformly distributed on \( \varphi(F^e_{j-1}) \) many integers. Hence,
\[
\mathbb{E} \left[ (M_j - M_{j-1})^2 \mid \mathcal{F}_{j-1} \right] = w(F^e_{j-1}),
\]
where
\[
w(r) := \frac{\varphi(r)^2 - 1}{12}
\]
denotes the variance of the uniform distribution on \( \varphi(r) \) consecutive integers. Then, condition (M1) becomes
\[
\lim_{h \to \infty} \frac{1}{h} \sum_{j \leq h} w(F^\epsilon_j) = \gamma^2.
\]  

(16)

This is a prototypical Markov-chain LLN that follows from the drift condition (15) through [Meyn and Tweedie 2009, Theorem 17.0.1]. More precisely, we deduce from (15) and [Meyn and Tweedie 2009, Theorem 9.1.8] that the chain \( \{F^\epsilon_h\}_{h \geq 0} \) is Harris recurrent. Next, by [Meyn and Tweedie 2009, Theorem 14.0.1], it is also positive recurrent with an invariant measure \( \pi \) satisfying \( \int_0^\infty x^2 \pi(dx) < \infty \). Hence, the asserted LLN in (16) follows from [Meyn and Tweedie 2009, Theorem 17.1.7].

**Condition (M2').** Similarly, let now \( w_{\epsilon_0}(r) \) denote the centered \((2 + \epsilon_0)\)-th moment of a uniform random variable on \( \varphi(r) \) consecutive integers. Then, (M2') becomes

\[
\lim_{h \to \infty} \frac{1}{h^{1+\epsilon_0}} \sum_{j \leq h} w_{\epsilon_0}(F^\epsilon_j) = 0,
\]

which again follows from the Markov LLN [Meyn and Tweedie 2009, Theorem 17.1.7].

Finally, we prove Lemma 4 by verifying conditions (D1)–(D3). We now use this criterion to prove convergence with the \( \delta \)-stable Lévy measure \( \nu \) as in (6). To this end, we set

\[
Z_{j,h} := h^{-1/\delta}(X^\epsilon_j - X^\epsilon_{j-1})
\]

and let

\[
\mathcal{F}_{j,h} := \sigma(Z_{1,h}, \ldots, Z_{j,h}, F^\epsilon_{1,h}, \ldots, F^\epsilon_{j-1,h})
\]

be the \( \sigma \)-algebra generated by the increments up to layer \( j \) and the fitnesses up to layer \( j - 1 \). To verify conditions (D1)–(D3), we rely on explicit computations with Fréchet random variables that we present as a separate auxiliary result.

**Lemma 10 (Frechet computations)** Let \( F \) be a standard Fréchet random variable with tail index 1. Then,

1. for every \( \delta > 1 \),

\[
\lim_{a \to \infty} a \mathbb{E}[(1 - (a/F)^{1/\delta})_+] = \frac{1}{\delta + 1}.
\]

2. for every \( \delta \in (1, 2) \) and \( \eta > 0 \),

\[
\lim_{a \to \infty} a \mathbb{E}[\eta^2 \wedge (F/a)^{2/\delta}] = \frac{2\eta^{2-\delta}}{2 - \delta}.
\]
We postpone the proof of the lemma and first show how it implies the proof of Lemma 4.

**Verification of condition (D1)**

To verify condition (D1), we need to compute the conditional expectation $\mathbb{P}(Z_{j,h} > x \mid \mathcal{F}_{j-1,h})$. Now, similar to the proof for $\delta > 2$, the key insight is that $X_j^e$ is distributed uniformly in the scope of size $\varphi(F_{j-1}^e)$. As an initial observation, we note that $\max_{j \leq h} F_j^e / h \in o_h(1)$ with high probability. Indeed, by the Markov inequality, for any $\beta \in (1, \delta)$,

$$\mathbb{P}(\max_{j \leq h} F_j^e > h) \leq \sum_{j \leq h} \mathbb{P}((F_j^e)^\beta > h^\beta) \leq \frac{1}{h^\beta} \sum_{j \leq h} \mathbb{E}[(F_j^e)^\beta],$$

so that similarly to the arguments in Sect. 4.3, we may invoke the Markov ergodic theorem, [Meyn and Tweedie 2009, Theorem 14.0.1].

We also recall from (10) that $F_{j-1}^e = \varphi(F_{j-2})^{1/\delta} F_{j-1}^{1/\delta}$. Hence,

$$\mathbb{P}(Z_{j,h} > x \mid \mathcal{F}_{j-1,h}) = \mathbb{E}\left[(1 - \frac{h^{1/\delta}X}{2\varphi(F_{j-2})^{1/\delta} F_{j-1}^{1/\delta}}) + \mathcal{F}_{j-1,h}\right](1 + o_h(1)).$$

Now, applying part 1. of Lemma 10 with $a = hx^\delta / (2^\delta \varphi(F_{j-2}))$ shows that

$$\mathbb{E}\left[(1 - \frac{h^{1/\delta}X}{2\varphi(F_{j-2})^{1/\delta} F_{j-1}^{1/\delta}}) + \mathcal{F}_{j-1,h}\right] = \frac{2^\delta \varphi(F_{j-2})}{hx^\delta(\delta + 1)}(1 + o_h(1)). \tag{17}$$

Finally, as in the computations in Sect. 4.3, we deduce that there is an LLN, so that $\frac{1}{h} \sum_{j \leq h} \varphi(F_{j-2}^e)$ converges weakly.

**Verification of condition (D2)**

We use (17) to get

$$\mathbb{P}(Z_{j,h} > x \mid \mathcal{F}_{j-1,h}) = \frac{F_{j-2}^e}{2^\delta hx^\delta(\delta + 1)}(1 + o_h(1)).$$

We observed already before that $\max_{j \leq h} F_j^e / h \in o_h(1)$ with high probability. Hence,

$$\max_{j \leq h} \mathbb{P}(Z_{j,h} > x \mid \mathcal{F}_{j-1,h}) \to 0,$$

which, together with (D1), implies the desired claim.
Verification of condition (D3)

Finally, we show that

\[
\lim_{\eta \to 0} \limsup_{h \to \infty} \mathbb{P}\left( \sum_{j \leq h} \mathbb{E}\left[ Z_{j,h}^2 \mathbb{1}\{|Z_{j,h}| \leq \eta\} | \mathcal{F}_{j-1,h} \right] > \varepsilon \right) = 0.
\]

First, conditioned on the event \(\{|Z_{j,h}| \leq \eta\}\), the increment \(Z_{j,h}\) is uniformly distributed in an interval of length \(\eta \wedge F_{j-1,h}^{-1/\delta}\). Hence, it suffices to show that

\[
\limsup_{h \to \infty} \mathbb{P}\left( \sum_{j \leq h} \mathbb{E}\left[ \eta^2 \wedge (F_{j-1,h}^{-1/\delta})^2 | \mathcal{F}_{j-1,h} \right] > \varepsilon \right)
\]

tends to 0 as \(\eta \to 0\). By part (2) of Lemma 10, the conditional expectation inside the probability becomes

\[
\frac{2\eta^{2-\delta} F_{j-2}^\delta}{(2-\delta)hx^\delta}.
\]

We may once more cite the Markov LLN for the weak convergence of \(\frac{1}{h} \sum_{j \leq h} F_{j-2}^\delta\) to deduce that (18) tends to 0 as \(\eta \to 0\).

It remains to establish the limits in Lemma 10.

**Proof** (Lemma 10) **Part (1).** Integration with respect to the Fréchet density yields that

\[
\mathbb{E}[1 - (a/F)^{1/\delta}] = \int_{a}^{\infty} (1 - (a/x)^{1/\delta})x^{-2} \exp(-x^{-1})dx.
\]

For large \(x\), the exponential factor approaches 1 and

\[
\int_{a}^{\infty} (1 - (a/x)^{1/\delta})x^{-2}dx = a^{-1} - \frac{\delta}{\delta + 1}a^{-1} = \frac{1}{\delta + 1}a^{-1}.
\]

**Part (2).** We split the expectation depending on which of the two contributions in the minimum becomes relevant. First, as in part (1),

\[
\lim_{a \to \infty} a\eta^2 \mathbb{P}(F \geq a\eta^\delta) = \eta^{2-\delta}.
\]

Hence, it remains to determine

\[
\lim_{a \to \infty} a^{1-2/\delta} \mathbb{E}[F^{2/\delta} \mathbb{1}\{F \leq a\eta^\delta\}].
\]

By l’Hôpital’s rule, we see that

\[
\lim_{a \to \infty} (a\eta^\delta)^{1-2/\delta} \int_{0}^{a\eta^\delta} x^{2/\delta}x^{-2} \exp(-x^{-1})dx = \frac{\delta}{2 - \delta}.
\]
Combining the latter with the result in (19) concludes the proof. □

Appendix B: Proofs of auxiliary results from Sect. 5

We start by proving Lemma 6.

Proof (Lemma 6) First,

\[ \mathbb{P}\left( F_h \leq 1/(2 \log \log(N)) \right) = \exp\left( -2 \log \log(N) \right) = \log(N)^{-2}, \]

so that we may assume \( F_h \geq 1/(2 \log \log(N)) \) for all \( h \leq 2 \log \mu(N) \).

By Lemma 1, the event \( \{ \sup_{h \leq r_0} \phi(F_h) \geq N \} \) occurs with high probability. Hence, it remains to show that

\[ \lim_{N \to \infty} \mathbb{P}\left( \inf_{h, h' \leq 2 \log \mu(N)} (F^{L}_{h+h'} / F^{L}_h) \leq \exp(-\log \mu(N)^{3/4}/2) \right) = 0. \]

To that end, we fix \( h, h' \leq 2 \log \mu(N) \) and apply the union bound afterwards. First, by (8) in companion with (7),

\[ \mathbb{P}\left( F^{L}_{h+h'} / F^{L}_h \leq \exp(-\log \mu(N)^{3/4}/2) \right) \leq \mathbb{P}\left( \prod_{h < i \leq h+h'} (2F_i) \leq \exp(-\log \mu(N)^{3/4}/2) \right). \]

If \( h' \leq \log(N)^{2/3} \), then

\[ \prod_{h < i \leq h+h'} (2F_i) \geq (\log \log(N))^{-h'} > \exp(-\log \mu(N)^{3/4}/2). \]

To conclude the proof, we show that if \( h' \geq \log(N)^{2/3} \) and \( N \) is sufficiently large, then

\[ \mathbb{P}\left( \prod_{h < i \leq h+h'} (2F_i) \leq 1 \right) \leq \exp(-\log(N)^{1/2}). \]

First, setting \( G_i := \log \mu(2F_i) \) and noting that \( \{ G_i - 1 \}_{i \geq 1} \) forms an iid sequence of centered random variables with finite second moment, we reduce the task to showing that

\[ \mathbb{P}\left( \sum_{i \leq h'} (G_i - 1) \leq -h' \right) \leq \exp\left( -\log(N)^{1/2} \right). \]

(20)

Now, \( G_i - 1 \) is Gumbel distributed and therefore its moment generating function is finite around 0. Hence, we may apply moderate deviations as in [Rassoul-Agha and Seppäläinen 2015, Theorem 11.2] at scale \( (h')^{7/8} \), to deduce that

\[ \mathbb{P}\left( \sum_{i \leq h'} (G_i - 1) \leq -h' \right) \leq \exp(-(h')^{3/4}) \]
for all sufficiently large $h'$. Since $h' \geq \log(N)^{2/3}$, we conclude the proof of (20). \hfill $\Box$

To construct the coupling, we write

$$\mathcal{N}^L_h := \{ i \in \{0, \ldots, N-1 \} : d_N(i, X^L_h) \leq \lfloor F^L_h \rfloor \}$$

for the scope of $X^L_h$ and define $\mathcal{N}^R_h$ accordingly. We also put

$$\mathcal{N}^\cap_h := \mathcal{N}^L_h \cap \mathcal{N}^R_h,$$

and

$$\mathcal{N}^\cup_h := \mathcal{N}^L_h \cup \mathcal{N}^R_h,$$

for the intersection and union of the scopes of the two walkers at time $h$.

**Proof** (Lemma 7) Write $n^L_h, n^R_h, n^\cap_h, n^\cup_h \geq 0$ for the cardinalities of the scopes $\mathcal{N}^L_h, \mathcal{N}^R_h, \mathcal{N}^\cap_h, \mathcal{N}^\cup_h$, respectively. Then, the coalescence event $X^L_{h+1} = X^R_{h+1}$ means that the maximum of all fitnesses in $\mathcal{N}^\cup_h$ is contained in $\mathcal{N}^\cap_h$. This event is of probability $n^\cap_h / n^\cup_h$. To prove the assertion, we therefore need to construct a coupling such that $\mathbb{P}(\mathcal{E}_h^{\text{coup}} \mid \mathcal{F}^c) \leq 4n^\cap_h / n^\cup_h$ holds under $\mathcal{E}_h^{\text{suc}}$.

To that end, we proceed inductively and assume the coupling to be constructed until step $h$ and work under the event of coupling success $\mathcal{E}_h^{\text{suc}}$. In order to define the true process $(X^L_{h+1}, F^L_{h+1}), (X^R_{h+1}, F^R_{h+1})$, we let $F_1, \ldots, F_{n^h_h}$ be a sequence of i.i.d. fitnesses. Here, we think of $F_1, \ldots, F_{n^h_h}$ to be in the scope of the left walker and $F_{n^h_h}, F_{n^h_h}+1, \ldots, F_{n^h_h}$ to be in the scope of the right walker. See Fig. 5. To recover the true process, each of the walkers selects the maximum fitness within its scope. Note that $\mathcal{N}^\cap_h$ is empty if $d_N(X^L_h, X^R_h) > \lfloor F^L_h \rfloor + \lfloor F^R_h \rfloor$.

To construct the coupling, we may w.l.o.g. assume that $n^L_h \leq n^R_h$. Otherwise, reverse the roles of $L$ and $R$ in the following argument. In order to define the independent process $((X^L_{h+1}, F^L_{h+1}), (X^R_{h+1}, F^R_{h+1}))$, we need to remove the dependence coming from the observation that in the true process, both the left and the right walker query the fitnesses $F_{n^L_h}, F_{n^L_h}+1, \ldots, F_{n^L_h}$. To this end, we introduce a further copy of $n^L_h$ independent uniform fitnesses $F_{n^L_h}, \ldots, F_{n^L_h}$. Then, the left walker selects the maximum fitness among $F_1, \ldots, F_{n^L_h}$ and the right walker selects the maximum fitness among $F_{n^L_h}, \ldots, F_{n^L_h}$. Hence, under $\mathcal{E}_h^{\text{suc}}$, we may express the probability of a coupling failure succinctly as

\[ \mathcal{N}^\cap_h \]

\[ \mathcal{N}^\cup_h \]

\[ \mathcal{N}^L_h \]

\[ n^L_h - n^R_h + 1 \]

\[ n^L_h \]

\[ n^R_h \]

\[ \mathcal{N}^\cup_h \]

\[ \mathcal{N}^L_h \]

\[ \mathcal{N}^\cap_h \]

\[ \mathcal{N}^R_h \]

Fig. 5 An illustration of the coupling in the proof of Lemma 6
\[ \mathbb{P}(E_{h+1}^{\text{fail}} \mid F_{h}^{\text{coup}}) = \mathbb{P}\left( \{ M_{R} \in N_{h}^{\gamma} \} \cup \{ M_{R,j} \geq n_{h}^{U} + 1 \} \right), \]

where \( M_{R} \in N_{h}^{\gamma} = \{ n_{h}^{U} - n_{h}^{R} + 1, \ldots, n_{h}^{U} \} \) and \( M_{R,j} \in \{ n_{h}^{L} + 1, \ldots, n_{h}^{L} + n_{h}^{R} \} \) denote the locations of the maxima of \( F_{n_{h}^{U} - n_{h}^{R} + 1}, \ldots, F_{n_{h}^{U}} \) and \( F_{n_{h}^{L} + 1}, \ldots, F_{n_{h}^{L} + n_{h}^{R}} \), respectively. Since both \( M_{R} \) and \( M_{R,j} \) are uniformly distributed in their domains of size \( n_{h}^{R} \), we deduce that

\[ \mathbb{P}\left( \{ M_{R} \in N_{h}^{\gamma} \} \cup \{ M_{R,j} \geq n_{h}^{U} + 1 \} \right) \leq \mathbb{P}(M_{R} \in N_{h}^{\gamma}) + \mathbb{P}(M_{R,j} \geq n_{h}^{U} + 1) = \frac{2n_{h}^{U}}{n_{h}^{R}} \leq \frac{4n_{h}^{U}}{n_{h}^{U}}, \]

the last step relying on the assumption that \( n_{h}^{L} \leq n_{h}^{R} \).

Finally, we establish Lemma 8.

**Proof** (Lemma 8) Recalling the recursion formula (7) for \( \delta = 1 \), we introduce a modified fitness by setting \( F_{0}^{L,m} = (2F_{0}^{L,1} \land N)/N \) and then \( F_{h+1}^{L,m} = ((2F_{h}^{L,m}F_{h+1}^{L}) \land 1) \). Then, by induction, \( F_{h}^{L,m} \leq \varphi(F_{h+1}^{L})/N \) for all \( h \geq 0 \).

Defining \( F_{h}^{R,m} \) similarly, it therefore suffices to prove the assertion with \( \varphi(F_{h}) \) and \( \varphi(F_{h}) \) replaced by \( N F_{h}^{L,m} \) and \( N F_{h}^{R,m} \).

Now, \( G_{h}^{L,m} := \log_{\varphi}(F_{h}^{L,m}) \) is a random walk in \((-\infty,0] \) truncated at 0 with drift to the right. Hence, there exists \( K > 0 \) such that for \( h_{0} = \log_{\varphi}(N)^{7/8} / 2 \) we have \( \mathbb{P}(G_{h_{0}}^{L,m} \leq -K) \leq \epsilon \). In particular, \( \mathbb{P}\left( (C_{h_{0}}^{L,m}, G_{h_{0}}^{R,m}) \in [-K,0)^{2} \right) \geq 1 - 2\epsilon \). Finally, note that the set \([-K,0]^{2} \) has finite expected return time and that there is a positive probability \( p(K) > 0 \) such that almost surely on the event \((C_{h_{0}}^{L,m}, G_{h_{0}}^{R,m}) \in [-K,0)^{2} \) we have

\[ \mathbb{P}\left( G_{h+1}^{L,m} = G_{h+1}^{R,m} = 0 \mid G_{h}^{L,m}, G_{h}^{R,m} \right) \geq p(K). \]

This gives domination by a geometric random variable and thereby concludes the proof.

**Acknowledgements** The authors are indebted to Lisa Hartung for pointing us to the relevant stable limit laws for martingales. We also thank both anonymous referees for their constructive feedback. CH acknowledges the financial support of the CogniGron research center and the Ubbo Emmius Funds (University of Groningen).

**Declaration**

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

**References**

Aiello, W., Bonato, A., Cooper, C., Janssen, J., Prałat, P.: A spatial web graph model with local influence regions. Internet Math. 5(1–2), 175–196 (2008)
Barabási, A.L., Albert, R.: Emergence of scaling in random networks. Science 286, 509–512 (1999)
Bassett, D.S., Bullmore, E., Verchinski, B.A., Mattay, V.S., Weinberger, D.R., Meyer-Lindenberg, A.: Hierarchical organization of human cortical networks in health and schizophrenia. Journal of Neuroscience 28(37), 9239–9248 (2008)
Bertoin, J.: Lévy Processes. Cambridge University Press, Cambridge (1996)

Springer
Bollobás, B., Riordan, O.: The diameter of a scale-free random graph. Combinatorica 24(1), 5–34 (2004)
Bullmore, E., Sporns, O.: Complex brain networks: graph theoretical analysis of structural and functional systems. Nat. Rev. Neurosci. 10(3), 186–198 (2009)
Calka, P., Chenavier, N.: Extreme values for characteristic radii of a Poisson-Voronoi tessellation. Extremes 17(3), 359–385 (2014)
Cardillo, A., Zanin, M., Gómez-Gardenes, J., Romance, M., del Amo, A.J.G., Boccaletti, S.: Modeling the multi-layer nature of the European air transport network: Resilience and passengers re-scheduling under random failures. Eur. Phys. J. Special Topics 215(1), 23–33 (2013)
Chenavier, N., Nagel, W.: The largest order statistics for the inradius in an isotropic STIT tessellation. Extremes 22(4), 571–598 (2019)
Cohen, R., Havlin, S.: Complex Networks: Structure, Robustness and Function. Cambridge University Press (2010)
Couzinié, Y., Hirsch, C.: Weakly reinforced Pólya urns on countable networks. Electron. Commun. Probab. 26, 1–10 (2021)
Durrett, R., Resnick, S.I.: Functional limit theorems for dependent variables. Ann. Probab. 6(5), 829–846 (1978)
Felleman, D.J., Van Essen, D.C.: Distributed hierarchical processing in the primate cerebral cortex. Cereb. Cortex 1(1), 1–47 (1991)
Gayrard, V., Hartung, L.: Dynamic phase diagram of the REM. In: V. Gayrard, L.P. Arguin, N. Kistler, I. Kourkova (eds.) Statistical Mechanics of Classical and Disordered Systems, pp. 111–170. Springer Publishing, Cham (2019)
Heydenreich, M., Hirsch, C.: A spatial small-world graph arising from activity-based reinforcement. In: K. Avrachenkov, P. Prałat, N. Ye (eds.) Proceedings of the 16th International Workshop, WAW 2019, Brisbane, QLD, Australia, July 6–7, 2019, Lecture Notes in Comput. Sci., vol. 11631, pp. 102–114. Springer, Cham (2019)
Hirsch, C., Holmes, M., Kleptsyn, V.: Absence of WARM percolation in the very strong reinforcement regime. Ann. Appl. Probab. 31(1), 199–217 (2021)
vander Hofstad, R.: Random graphs and complex networks. Vol. I. Cambridge University Press, Cambridge (2017)
vander Hofstad, R., Holmes, M., Kuznetsov, A., Ruszel, W.: Strongly reinforced Pólya urns with graph-based competition. Ann. Appl. Probab. 26(4), 2494–2539 (2016)
Holmes, M., Kleptsyn, V.: Infinite WARM graphs II. Critical regime. In preparation
Holmes, M., Kleptsyn, V.: Proof of the WARM whisker conjecture for neuronal connections. Chaos 27(4), 43104 (2017)
Jacob, E., Mörters, P.: A spatial preferential attachment model with local clustering. In: A. Bonato, M. Mitzenmacher, P. Prałat (eds.) Proceedings of the 10th International Workshop (WAW 2013) held at Harvard University, Cambridge, MA, December 14–15, 2013, pp. 14–25. Springer, Berlin (2013)
Kivelä, M., Arenas, A., Barthelemy, M., Gleeson, J.P., Moreno, Y., Porter, M.A.: Multilayer networks. J. Complex Netw. 2(3), 203–271 (2014)
Loeffler, A., Zhu, R., Hochstetter, J., Li, M., Fu, K., Diaz-Alvarez, A., Nakayama, T., Shine, J.M., Kuncic, Z.: Topological properties of neuromorphic nanowire networks. Frontiers in Neuroscience 14, 184 (2020)
Mayer, M., Molchanov, I.: Limit theorems for the diameter of a random sample in the unit ball. Extremes 10(3), 129–150 (2007)
Meyn, S., Tweedie, R.L.: Markov Chains and Stochastic Stability, second edn. Cambridge University Press, Cambridge (2009)
Penrose, M.: Random Geometric Graphs. Oxford University Press, Oxford (2003)
Rassoul-Agha, F., Seppäläinen, T.: A Course on Large Deviations with an Introduction to Gibbs Measures. American Mathematical Society, Providence, RI (2015)
Thomas, A.M., Owada, T.: Functional strong laws of large numbers for Euler characteristic processes of extreme sample clouds. Extremes (2021, forthcoming)
Wang, T., Resnick, S.I.: Consistency of Hill estimators in a linear preferential attachment model. Extremes 22(1), 1–28 (2019)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.