Bipartite Lattice Model for Mycorrhizal Networks

David Högberg

Tohoku University

August 2020

1   Introduction

Many real-world structures can be represented as networks, consisting of a set of nodes connected by links. In social networks, for example, nodes may represent people having links to other nodes in the form of social relationships. In communication networks, on the other hand, nodes may represent computers or routers connected through communication links. As it turns out, studying the statistical properties of networks often reveals characteristic patterns common to a wide range of observed structures. Most importantly, networks capture how things are interconnected in a system, which may not be immediately clear from studying its individual parts.

In this sense, the theory of networks has potential applicability in ecology, since understanding ecological structures requires not only an understanding of the individual species and organisms involved, but also how they interact. For example, many species engage in mutualistic interactions that are crucial for maintaining ecosystem functionality. Specifically, one of the oldest recorded mutualistic interactions is that between trees and fungi, forming what is called mycorrhiza [1], which take the form of physical connections between the root systems of the trees and the fungal mycelium. The mutualism enables exchange of photosynthetically acquired carbohydrates needed for fungal growth in exchange for e.g. phosphorus and nitrogen based compounds in the soil that may be inaccessible for the trees. A general network representation of these webs of interactions comes in the form of organisms as nodes connected through mutualistic links, forming a mycorrhizal network that channel the flow of nutrients within the ecosystem and adds evolutionary benefit to the species involved.

An important statistical property of networks is the degree distribution [2], which determines the probability that a node in the network will have a certain number of links to other nodes in the network, i.e. a certain degree. If links are formed randomly in the network, most nodes will have similar degrees close to some average value and the distribution of degrees within the network will have a sharp peak around the average. That is, the network will be homogeneous. However, most real networks have degree distributions that deviate in characteristic ways from randomly formed networks, which points to the fact that there are mechanisms involved in shaping those non-random aspects of networks. For example, social networks often include nodes with significantly higher degrees compared to other nodes, forming hubs in the network. Creating a network model for the phenomenon of interest that accurately recreates the observed patterns may therefore provide important insight into why certain network patterns emerge.

In the following, a basic model for mycorrhizal network formation is proposed that specifically examines the effect of spatial heterogeneity and a variable fitness within the population on the resulting degree distribution of the generated network. Trees and fungi are represented as nodes of two different partitions in a bipartite network, where links between each type of node indicate a mutualistic interaction in the form of nutrient exchange. The aim is first and foremost to introduce a basic formalism for studying and reasoning about the structural properties of mycorrhizal networks using network theory. In combination with empirical observations on actual networks, such models may provide valuable insight into why certain patterns are observed under certain environmental conditions.

2   Network Theory

Before proceeding with the model, we introduce a few concepts important in network theory and important for our specific purposes.

2.1   Networks and Degree Distributions

Formally, a *network* is defined by a set of nodes and a set of pairs of nodes that constitute the links in the network. In graph theory, the convention is instead to refer to nodes as vertices and links as edges (see Figure 1). Here, we assume unordered pairs of nodes, which means that the network is *undirected*.

Each node has a number of links to other nodes that defines its *degree*. The probability distribution of node degrees over the whole network, in turn, is termed the *degree distribution* of the network. That is, choosing a node at random in the network, that node will have a certain number of connections to other nodes with a probability given by the degree distribution. As with any probability distribution, there will be an average degree of the network, but deviations from that average may vary depending on the specific form of the distribution.
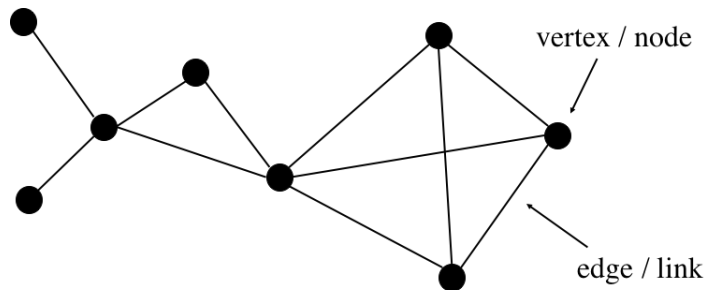

Figure 1: Networks consists of a set of nodes (vertices) connected by links (edges).

2.2   Bipartite Networks

In a *bipartite network*, the set of nodes that defines the network is divided into two partitions, with links only between nodes belonging to different partitions. As with unipartite networks, the nodes in

a partition will have a number of links to nodes in the other partition that defines its degree. Each partition will also have a corresponding degree distribution.

In a *unipartite projection* of a bipartite network, common neighbours between two nodes in the same partition are represented as direct links between those two nodes. That is, if two nodes in the same partition have links to the same node in the other partition, they have a direct link in the corresponding projected network. Since nodes may have several common neighbours, the resulting projected network is a *multigraph*, which allows more than one link between the same two nodes.

3   Network Models

Network models provide ways of generating network structures with certain degree distributions, in order to better understand the underlying principles of network formation. Earlier models have focused on two main classes of networks with qualitatively different degree distributions: random networks and scale-free networks.

3.1   Erdös-Rényi Model for Random Networks

The theory of random networks (random graphs) was initiated in 1959 by Erdös and Rényi [3] as well as independently by Gilbert [4], with important influences on the development of network theory. A random network can be obtained by starting with a fixed set of nodes and adding links between each pair of nodes with some fixed probability. The degree distribution of the resulting network will follow a binomial distribution, which in the limit of a large and sparse network is well approximated by a Poisson distribution. Therefore, the probability $p(k)$ of finding a node with degree k in the network is given by

$$p(k) = \frac{\bar{k}^k}{k!} \cdot e^{-\bar{k}}, \tag{1}$$

where $\bar{k}$ is the average degree of the network.

3.2   Barabási-Albert Model for Scale-Free Networks

As mentioned, however, real networks often display degree distributions that deviate in significant ways from randomly formed networks. For example, many networks display scale-free degree distributions, meaning they have degree distributions that follow a power law,

$$p(k) \sim k^{-\gamma}. \tag{2}$$

The constant γ is called the degree exponent, which typically has a value between 2 and 3 for real networks [2].

For this reason, recent work have focused on methods for generating networks that more closely resemble observed structures. Most notably, Barabási and Albert [5] developed a model based on the principle of preferential attachment in the formation process of a growing network, resulting in networks with nodes that acquire links at a higher rate as their degree increases.

An important aspect of scale-free networks is that a power law decays slower than an exponential function, which means that the probability of finding nodes with high degrees is significantly larger compared to other common distributions where the density function has an exponential form. In

essence, this means that scale-free networks will contain hubs, i.e. nodes with significantly higher degree than most other nodes in the network.

3.3 Hidden Parameter Model

Other models have made similar findings for generating heterogeneous degree distributions by assigning hidden parameters to each node that influence the probability of forming links. These parameters may figure as a measure of "fitness" or "importance" of a node [6].

In one formulation of this hidden parameter model, nodes are assigned fitness values η from some probability distribution. Links between node i and j are then formed in the network based on a connection probability

$$p(\eta_i, \eta_j) = \frac{\eta_i \eta_j}{\langle \eta \rangle N}, \tag{3}$$

where N is the size of the network. By choosing an appropriate form of the fitness distribution, the degree distribution of the resulting network can be tuned to acquire a specific form [2].

3.3.1 Latent Geometry of Networks

If the node-specific parameters in the hidden parameter model correspond to coordinates in a metric space, the network acquires a latent geometric structure. This enables the formulation of a distance metric between nodes that may also serve as a measure of node similarity or closeness. In fact, this latent geometry can under certain circumstances account for the heterogeneity observed in real networks [7].

4 Mycorrhizal Network Model

Mycorrhiza are mutualistic connections formed between the root system of trees and the fungal mycelium in order to aid nutrient uptake for trees and supply fungi with sufficient carbon for growth. When several trees are connected through the same fungal species, this forms a mycorrhizal network with important consequences for carbon, nutrient and water cycling within the ecosystem. Interestingly, there is growing evidence also for interplant transfer of carbon and signalling via the mycorrhizal connections, that may aid seedling establishment and increase disease resistance of plants [1].

The interconnected structure of the mycorrhizal mutualism can be suitably represented as a bipartite network with two types of nodes, trees and fungi, and with links between nodes indicating mutualistic connections. Also, the unipartite projection of the network would serve as a representation of the potential common mycorrhizal connections between trees.

In the following, a model for mycorrhizal network formation is introduced that may serve as a general framework for exploring hypotheses related to the mycorrhizal mutualism.

4.1 Ecological Assumptions

The aim of the model is to provide a basic formalism for understanding certain qualitative aspects of mycorrhizal networks using network theory. Therefore, as with any model, ecological assumptions and restrictions will be necessary. Specifically, we make the following assumptions:

- The mycorrhizal network consists of a fixed number of tree nodes in a top partition and a fixed number of fungal nodes in a bottom partition as part of a bipartite network. That is, there is no reproduction or removal of nodes.

- Each partition of the network has a spatial structure in the form of a circular lattice.

- Connections are formed in the network based on the activity of the fungi, which means there is no active partner discrimination from the perspective of the trees.

- The behaviour of a specific fungus is independent of any other fungus (i.e. there is no intraspecific competition among fungi).

- Without any specific restrictions, the fungi will form links to a certain average number of trees. That is, there is an upper limit to the number of connections a fungus would form even if there were no specific costs associated with forming links.

In reality, of course, these assumptions are too restrictive. The network would change size and connections with time and both fungi and trees would be involved in maintaining and forming links. There would also be competition among fungi for mutualistic connections. Therefore, as with any mathematical model of natural phenomena, any qualitative implications of the model have to take into account this divergence from the actual ecological context that is being modelled.

4.2   Mathematical Setup

Based on [8] and [9], and with the ecological assumptions in mind, we now formulate an extension of the hidden parameter model with a latent geometry to bipartite networks that may serve as a foundation for our modelling.

Consider a bipartite network with two types of nodes, forming a top partition ($\top$) and a bottom partition ($\perp$) of the network. In this case, top nodes correspond to trees and bottom nodes correspond to fungi. The size of each partition is fixed and given by $N_\top$ for top nodes and $N_\perp$ for bottom nodes, respectively.

Each node is assigned an angular coordinate between 0 and 2π corresponding to a position on a ring lattice (see Figure 2). The angular coordinates for each partition are drawn from two probability distributions with density functions $\rho_\top(\theta)$ and $\rho_\perp(\phi)$, respectively. For simplicity, we always denote top node positions with θ and bottom node positions with ϕ, also omitting the subscripts for each density function so that $\rho_\top(\theta) = \rho(\theta)$ and $\rho_\perp(\phi) = \rho(\phi)$.

Furthermore, we define a distance measure

$$d = d(\theta, \phi) \tag{4}$$

between nodes of different partitions based on their angular coordinates.

Similarly, each node is assigned a fitness value $\eta_\top$ and $\eta_\perp$ from distributions with density functions $\rho(\eta_\top)$ and $\rho(\eta_\perp)$, respectively for each partition.

Finally, we define a connection probability between nodes dependent on the distance between them as well as their respective fitness values,

$$r(d(\theta, \phi), \eta_\top, \eta_\perp) = r(\theta, \phi, \eta_\top, \eta_\perp). \tag{5}$$
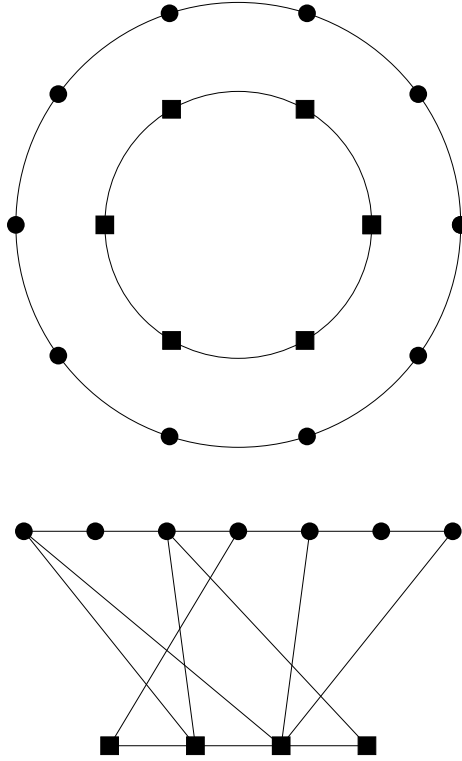
Figure 2: Illustration of the latent geometry of the model in the form of a ring lattice for each partition, seen from above as well as from the side. Round nodes correspond to nodes in a top partition of the network, whereas square nodes correspond to nodes in a bottom partition. Links are formed in the network based on a distance measure that influences the connection probability between nodes of different partitions.

For each pair of nodes from different partitions, links are then added to the network based on an evaluation of the connection probability, where each connection trial is independent of any other trial. Therefore, given specific forms of the distance function, the connection probability and each distribution function, a network will be formed.

4.3 Degree Distributions of Network Partitions

In order to derive a formula for the degree distribution of each network partition, we follow the approach given in [8]. The full derivation is given in Appendix A.

First, we consider the degree distribution for a node in the bottom partition, or equivalently in the top partition, with a fixed position ϕ and fitness value $\eta_\perp$. This node will have an average degree given by

$$\bar{k}_\perp(\phi, \eta_\perp) = N_\top \sum_{\theta, \eta_\top} \rho(\theta)\rho(\eta_\top)r(\theta, \phi, \eta_\top, \eta_\perp). \tag{6}$$

Then, assuming the connection probability is sufficiently small for large network size, the degree distribution will approximately follow a Poisson distribution with parameter $\bar{k}_\perp(\phi, \eta_\perp)$, i.e.

$$f(k|\phi, \eta_\perp) \approx \frac{1}{k!} \cdot [\bar{k}_\perp(\phi, \eta_\perp)]^k \cdot e^{-\bar{k}_\perp(\phi, \eta_\perp)}, \tag{7}$$

where $f(k|\phi, \eta_\perp)$ denotes the conditional degree distribution of a fungal node given position ϕ and fitness value $\eta_\perp$.

The average degree and the degree distribution for the whole partition is then given by the sum over all values of ϕ and $\eta_\perp$:

$$\bar{k}_\perp = \sum_{\phi, \eta_\perp} \rho(\phi)\rho(\eta_\perp)\bar{k}_\perp(\phi, \eta_\perp), \tag{8}$$

$$f(k) = \sum_{\phi, \eta_\perp} \rho(\phi)\rho(\eta_\perp)f(k|\phi, \eta_\perp). \tag{9}$$

4.4 Model Functions and Parameters

The model so far applies to general forms of the distance measure, the connection probability and each probability distribution. More specifically, therefore, we define the distance between two nodes to be the shortest angular difference between their respective positions, i.e.

$$d(\theta, \phi) = \pi - |\pi - |\theta - \phi||. \tag{10}$$

Moreover, we assume that the probability of forming links decays exponentially with larger distance and that the probability is proportional to the product of each node's fitness value. Also, we let the probability be inversely proportional to the size of the tree partition, so that for higher availability of trees, the probability of a fungi making a connection to any specific tree is lower. This leads to the following form of the connection probability:

$$r(\theta, \phi, \eta_\top, \eta_\perp) = \frac{\eta_\top \eta_\perp}{\langle \eta \rangle} \cdot \frac{r_\perp}{N_\top} \cdot e^{-\alpha \cdot d(\theta, \phi)}, \tag{11}$$

where $\langle \eta \rangle$ denotes the average of the product between the fitness values. The fraction $r_\perp/N_\top$ sets the probability of a fungus forming a link to a specific tree without any effect of distance or fitness, with $r_\perp$ average number of connections. The parameter α adjusts the cost of forming links (or the spatial scale of the circle).

The concept of fitness, in this case, may incorporate certain physiological or genetic advantages that increase the probability of successfully acquiring mutualistic links. In the current model, for fungi, this corresponds to a capacity of forming links whereas for trees, it corresponds to a capacity of attracting links.

A summary of model parameters and functions is given in Table 1.

5 Special Cases of Spatial- and Fitness Distributions

Specifically, we now examine the effects of different forms of ρ on the degree distribution of trees, which we donte by $g(k)$. Similar results would apply to the degree distribution of fungi. For complete derivations of all degree distributions, see Appendix B.

Table 1: Notation for parameters and functions of the model.

$N_\top$	Size of top partition (trees)
$N_\perp$	Size of bottom partition (fungi)
θ	Angular coordinates of trees
ϕ	Angular coordinates of fungi
$\eta_\top$	Fitness values of trees
$\eta_\perp$	Fitness values of fungi
ρ	Density function for coordinates or fitness values
d	Distance measure
r	Connection probability
α	Regulates the cost of distance
$r_\perp$	Regulates the average number of links for fungi without restrictions
$f(k)$	Degree distribution of fungi
$g(k)$	Degree distribution of trees

5.1 Uniform Spatial Distribution

First, we introduce a uniform spatial distribution for both trees and fungi as well as the same fitness values for all nodes of a partition. This may serve as a base case for comparisons in order to isolate the effects of spatial heterogeneity and a variable fitness among nodes. Therefore, let

$$
\begin{aligned}
\rho(\phi) = \rho(\theta) &= \frac{1}{2\pi}, \\
\rho(\eta_\top) &= \delta(\bar{\eta}_\top - \eta_\top), \\
\rho(\eta_\perp) &= \delta(\bar{\eta}_\perp - \eta_\perp),
\end{aligned}
\tag{12}
$$

where $\bar{\eta}_\top$ and $\bar{\eta}_\perp$ are constants and δ is the Dirac delta function.

The connection probability now simplifies to

$$
r(\theta, \phi, \eta_\top, \eta_\perp) = r(\theta, \phi) = \frac{r_\perp}{N_\top} \cdot e^{-\alpha \cdot d(\theta, \phi)},
\tag{13}
$$

and the average degree of a fungal node with position ϕ is given by

$$
\bar{k}_\perp(\phi) = \frac{N_\top}{2\pi} \int_0^{2\pi} r(\theta, \phi) d\theta = \frac{r_\perp}{\alpha \pi} \cdot [1 - e^{-\alpha \pi}],
\tag{14}
$$

using equation (6).

In this case, $\bar{k}_\perp(\phi) = \bar{k}_\perp$, since the conditional average is independent of position. Therefore, using equation (7), the degree distribution is a Poisson distribution with average $\bar{k}_\perp$.

Similarly, using the identity $N_\top \bar{k}_\top = N_\perp \bar{k}_\perp$, the degree distribution of trees is a Poisson distribution with average

$$
\bar{k}_\top = \frac{N_\perp}{N_\top} \cdot \frac{r_\perp}{\alpha \pi} \cdot [1 - e^{-\alpha \pi}] = \frac{r_\top}{\alpha \pi} \cdot [1 - e^{-\alpha \pi}],
\tag{15}
$$

where $r_\top = N_\perp r_\perp / N_\top$ sets the average number of connections for trees without any distance effect.

Figure 3 shows the dependence of the average on the parameter α, with a larger sensitivity for values between 0 and 1. Furthermore, Figure 4 shows the degree distribution of trees for different values of α. The second plot additionally shows the same degree distribution using a logarithmic scale. In this case, the degrees of the network are clearly restricted by the parameter $r_\top$, which in this case is set to 20.
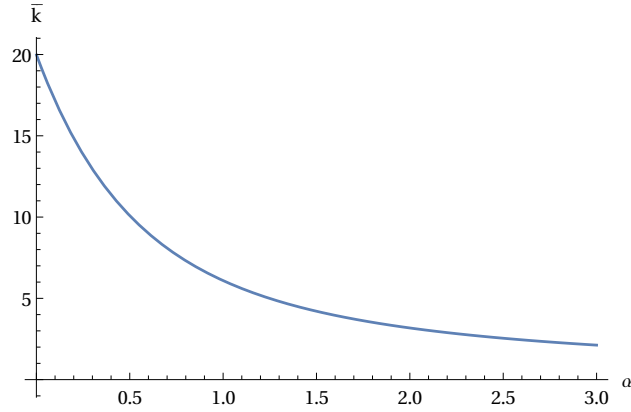
Figure 3: The average degree of network partitions as a function of α. In this case, $r_\perp$ (or $r_\top$) is set to 20, which is also the limit when $\alpha \to 0$.

5.2 Concentrated Spatial Distribution

Retaining the delta functions as fitness distributions, consider now a highly concentrated spatial distribution of fungi. For mathematical tractability, we let every fungi have the same position at $\phi = 0$, so that the density function is defined by a delta function on the circle. Even though this is a highly unrealistic proposition, it may serve as an extreme case of a concentrated point habitat. The density functions therefore take the following forms:

$$
\begin{aligned}
\rho(\theta) &= \frac{1}{2\pi}, \\
\rho(\phi) &= \text{III}_{2\pi}(\phi), \\
\rho(\eta_\top) &= \delta(\bar{\eta}_\top - \eta_\top), \\
\rho(\eta_\perp) &= \delta(\bar{\eta}_\perp - \eta_\perp),
\end{aligned}
\tag{16}
$$

where $\text{III}_{2\pi}(\phi)$ denotes the Dirac comb with period 2π.

Due to symmetry around 0, we can consider spatial positions of trees between 0 and π, since this simplifies the distance formula. The average degree of a tree with position $\theta \in [0, \pi]$ is then given by

$$
\bar{k}_\top(\theta) = \frac{N_\perp}{N_\top} \cdot r_\perp \cdot e^{-\alpha\theta} = r_\top \cdot e^{-\alpha\theta}.
\tag{17}
$$

The average degree for a tree nodes therefore decreases exponentially for larger distance from the point habitat, see Figure 5. As $\alpha \to 0$, the average approaches $r_\top$ for all positions.
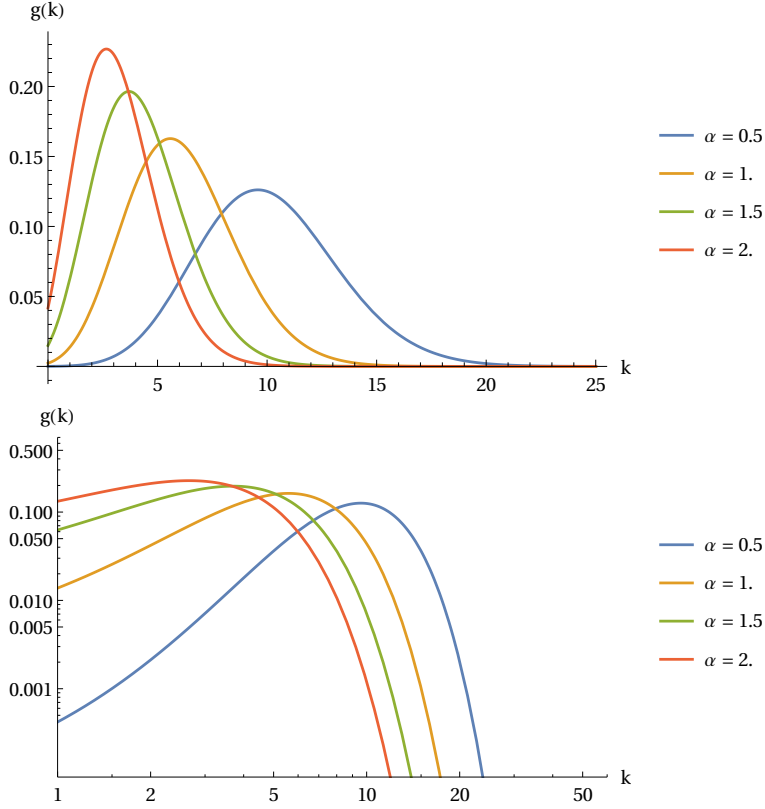
9

Figure 4: Degree distribution of trees for the uniform case and different values of α, displaying a Poisson distribution with average $\bar{k}_\top$. The parameter $r_\top$ is set to 20, which puts a restriction on the maximum degree of the network.

Using equation (9), the degree distribution of trees is then

$$g(k) = \frac{1}{\alpha \pi k!} \cdot \left[\Gamma(k, r_\top e^{-\alpha \pi}) - \Gamma(k, r_\top) \right] = \frac{1}{\alpha \pi k} \cdot \left[Q(k, r_\top e^{-\alpha \pi}) - Q(k, r_\top) \right], \tag{18}$$

where $\Gamma(k, x)$ denotes the incomplete gamma function and $Q(k, x)$ the regularized incomplete gamma function.

The regularized gamma function is related to the cumulative distribution function of a Poisson distribution in that if $X \sim \text{Poisson}(\lambda)$, then

$$Q(k, \lambda) = \Pr(X < k). \tag{19}$$

Therefore, for $k > 0$ and α sufficiently large, $Q(k, r_\top e^{-\alpha \pi}) \approx 1$, so that

$$g(k) \approx \frac{1}{\alpha \pi k} \cdot \left[1 - Q(k, r_\top) \right]. \tag{20}$$

Moreover, if $X \sim \text{Poisson}(r_\top)$, then
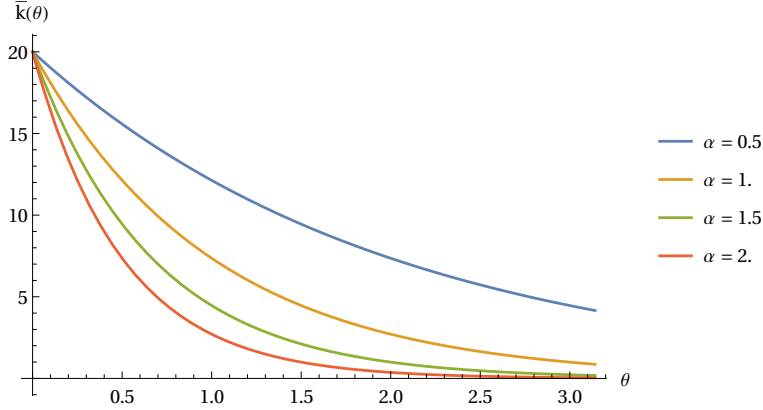
$$1 - Q(k, r_\top) = \Pr(X \geq k). \tag{21}$$

Figure 5: Average degree for tree nodes with a given distance θ from the point habitat, for different values of α and for $r_\top = 20$.

Assuming $r_\top$ is sufficiently large, this probability is close to 1 for $k \ll r_\top$ and close to 0 for $k \gg r_\top$, since the Poisson distribution has a sharp peak around $k = r_\top$. For $k \ll r_\top$, then, the degree distribution has a power law regime with degree exponent 1, i.e.

$$g(k) \sim k^{-1}, \tag{22}$$

followed by a high degree truncation for $k \gg r_\top$. In fact, from a mathematical point of view, a high degree truncation is the only way of having power law distributions with degree exponent 1, since the corresponding pure power law is not normalizable.

Figure 6 shows the resulting degree distribution of trees using both a linear and a logarithmic scale. The form of the degree distribution resembles the uniform case for smaller α, whereas for higher values of α, the power law nature becomes more apparent. For $\alpha = 2$, the logarithmic plot shows a straight line corresponding to the power law regime, followed by a sharp truncation for degrees $k > r_\top$, where $r_\top$ is set to 20.

5.3   Exponential Fitness Distribution

Consider again the case of a uniform spatial distribution for both partitions but a general fitness distribution. In this case, the average degree of a node given a specific position and fitness value will be proportional to the node's fitness value, i.e.

$$\bar{k}_\top(\theta, \eta_\top) = \bar{k}_\top(\eta_\top) \propto \eta_\top. \tag{23}$$

This means that the final degree distribution will be highly influenced by the distribution of the fitness values.

As an example, consider the case with an exponential fitness distribution for trees, i.e. let

$$\rho(\theta) = \rho(\phi) = \frac{1}{2\pi},$$
$$\rho(\eta_\top) = \lambda_\top e^{-\lambda_\top \eta_\top}, \tag{24}$$
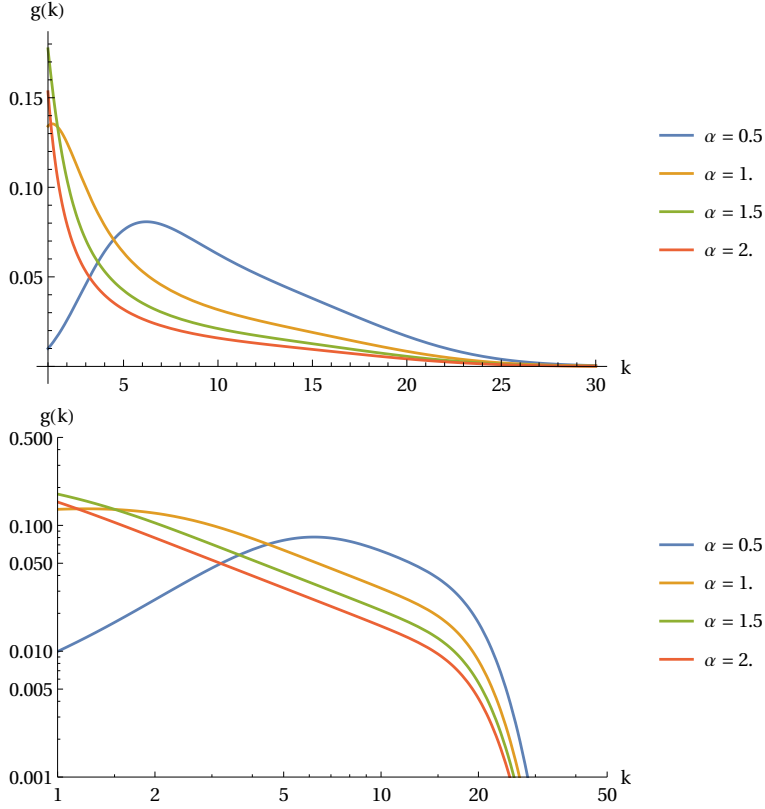$$\rho(\eta_\perp) = \delta(\bar{\eta}_\perp - \eta_\perp),$$

11

Figure 6: Degree distribution of trees for the case of a concentrated point habitat and for different values of α. For α sufficiently large, the degree distribution displays a power law regime followed by a high degree truncation for $k > r_\top$, where $r_\top$ is set to 20.

for some constant $\lambda_\top$.

The degree distribution for tree nodes is then given by

$$g(k) = \frac{1}{1 + K} \cdot (1 + 1/K)^{-k}, \tag{25}$$

where

$$K = \frac{r_\top}{\alpha \pi} \cdot [1 - e^{-\alpha \pi}] \tag{26}$$

is the average degree for the homogeneous case. Therefore, we see that

$$g(k) \sim e^{-\beta k}, \tag{27}$$

for a constant β, which takes the form of an exponential distribution, i.e. similar to the form of the fitness distribution.

Again, the distribution is show in Figure 7 using two different scales.
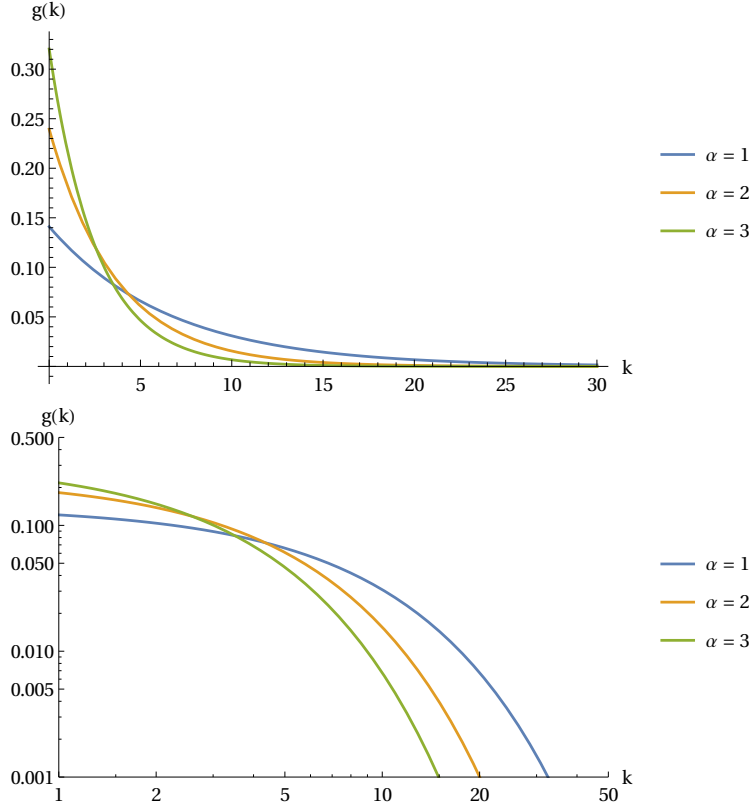
Figure 7: Degree distribution of trees for the case with a uniform spatial distribution and exponentially distributed fitness values for trees. The parameter $r_\top$ is set to 20.

5.4 Heterogeneous Spatial- and Fitness Distribution

Finally, we consider the combined effect of a concentrated spatial distribution in the form of a point habitat as well as exponentially distributed fitness values, i.e.

$$
\begin{aligned}
\rho(\theta) &= \frac{1}{2\pi}, \\
\rho(\phi) &= \text{III}_{2\pi}(\phi), \\
\rho(\eta_\top) &= \lambda_\top e^{-\lambda_\top \eta_\top}, \\
\rho(\eta_\perp) &= \delta(\bar{\eta}_\perp - \eta_\perp).
\end{aligned}
\tag{28}
$$

This introduces an exponential factor to the conditional average, so that

$$
\bar{k}_\top(\theta, \eta_\top) \propto \eta_\top \cdot e^{-\alpha\theta}.
\tag{29}
$$

The resulting degree distribution of tree nodes is then

$$
g(k) = \frac{1}{\alpha\pi k} \cdot [(1 + r_\top^{-1})^{-k} - (1 + r_\top^{-1} e^{\alpha\pi})^{-k}].
\tag{30}
$$

Moreover, assuming $k > 0$ and $e^{\alpha\pi} \gg r_\top$,

$$
g(k) \approx \frac{1}{\alpha\pi k} \cdot (1 + r_\top^{-1})^{-k} \sim C k^{-1} e^{-\beta k},
\tag{31}
$$

13

for some constant C. That is, the degree distribution is a power law distribution with an exponential cut-off [2]. Taking the logarithm,

$$\ln g(k) \sim \ln C - \ln k - \beta k, \tag{32}$$

we see that for $k \ll 1/\beta$, the power law behaviour dominates, whereas for $k \gg 1/\beta$, the βk term takes over, resulting in an exponential cut-off.

This is visible also in Figure 8, showing the degree distribution using a linear as well as logarithmic scale. A pure power law would follow a straight line using a logarithmic plot. In this case, however, the exponential factor introduces a high degree saturation. Interestingly, however, the saturation is slower in this case compared to the spatially heterogeneous case without fitness effects, which follows a high-degree truncation similar to a Poisson distribution.
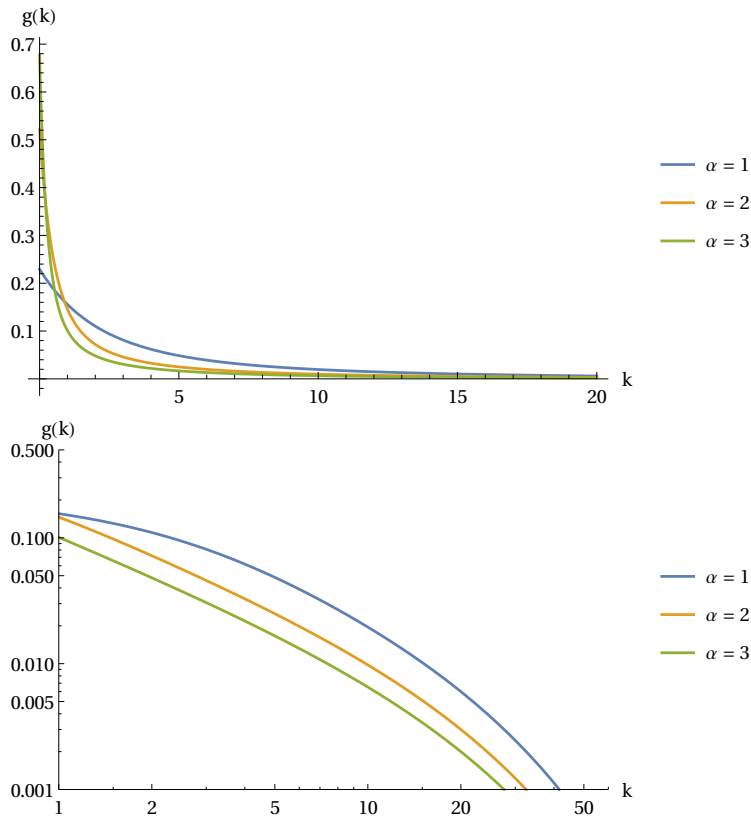

Figure 8: Degree distribution of trees for the case with a heterogeneous spatial distribution of fungi and exponentially distributed fitness values for trees, for different values of the parameter α. The logarithmic plot displays a power law regime with a high degree saturation in the form of an exponential cut-off. The parameter $r_\top$ is set to 20.

5.5   Numerical Simulations

Finally, we compare the analytical results for the degree distributions with numerical simulations. These can be performed in a straightforward way by counting the number of successful connection

trials, based on an evaluation of the connection probability, for each node in a network of fixed size.

Specifically, we assume $N_\top = N_\perp = 1000$ for network partitions and $r_\perp = 20$, as previously. Higher values of $r_\perp$ will mainly influence the average degree of the network. The parameter α, on the other hand, may significantly influence also the shape of the distribution. Therefore, we perform simulations for $\alpha = 1$ (blue colour), $\alpha = 2$ (orange colour) and $\alpha = 3$ (green colour).

Since $r_\top = r_\perp N_\perp / N_\top$, its value is determined by the other parameters. The ratio between the number of fungal nodes and the number of tree nodes will therefore affect the average degree for trees. Also, the introduced Poisson distribution assumes a large size for the network partition of trees. Therefore, the numerical results will coincide better with analytical formulas for larger values of $N_\top$.

The results of the numerical calculations are presented in Appendix C.

6 Conclusions

The presented model provides a simple formalism for studying the effects of spatial heterogeneity and a variable fitness on the structural properties of mycorrhizal networks. Specifically, given particular forms of spatial- and fitness distributions, the degree distributions of the network partitions range from homogeneous Poisson distributions to more heterogeneous power laws with a high degree truncation or cut-off.

The simplest case with no differences in fitness values and a uniform spatial distribution resembles the case of a random network in that the local situation for each fungi will be the same. Also, there is no reason for a fungus to choose any specific mutualistic partner other than those that are spatially closer. This excludes the possibility of high degree nodes in the network.

Introducing spatial heterogeneity in the form of a concentrated habitat for fungi means that trees located closer to the habitat will attract more links, forming nodes with a larger number of mutualistic connections than most other nodes. This effect gets more prevalent for a higher cost of forming long-distance connections for fungi (i.e. for higher values of α). The introduced spatial heterogeneity also means that trees located far away from the habitat will attract fewer links, with potential consequences for their survival.

The parameters $r_\perp$ and $r_\top$ were introduced in order to enforce certain limitations on the total number of links that can be formed in theory for a single node. The degree distributions of other mutualistic networks may therefore display similar high degree cut-offs due to similar restrictions in the underlying phenomena or in the surrounding environment. For this reason, fitting a pure power law to empirical data on mutualistic networks may in some cases be misguided.

Unfortunately, empirical data on mycorrhizal networks is rare due to the difficulties in accurately measuring underground mycorrhizal connections. The results in [10], however, suggests a scale-free degree distribution for a mycorrhizal network in a Douglas-fir forest, relating it also to the age distribution of trees within the forest. Also, in [11], data on mycorrhizal associations for an oak woodland was analysed, finding a homogeneous degree distribution for trees and a scale-free distribution for fungi. In this case, potential links between trees are assumed based on the existence of the same fungal morphotype on the root system of both trees. Given more data on actual mycorrhizal networks, therefore, the current model may provide some insight into how specific environmental circumstances may influence the observed degree distributions. As mentioned, however, the provided ecological assumptions are highly restrictive, and any comparisons with empirical data must therefore take into

account the ways in which the theoretical model diverges from the specific ecological context under consideration.

Finally, on a more general note, examining the degree distribution of mutualistic networks may be valuable from the perspective of conservation biology in that random networks and scale-free networks have qualitatively different error tolerances. Specifically, scale-free network structures are robust against random node removal in comparison to random networks, in terms of retaining network connectivity. However, scale-free structures may easily disintegrate as a consequence of a selective removal of high degree nodes in the network [12]. Uncautiously removing tree nodes that may serve as hubs in a mycorrhizal network may therefore have deleterious environmental implications.

7 Extensions and Future Research

The formulated model provides a general representation of mycorrhizal networks that can be used to further explore qualitative aspects of the mycorrhizal mutualism. For example, there is growing evidence that mycorrhiza may serve a function also in enabling nutrient transfer between trees, forming common mycorrhizal networks between trees [1]. Trees can therefore aid the survival of other trees that are low on certain resources if the two trees share a mutualistic partner. The resulting network structure formed between trees is exactly the unipartite projection of the bipartite network. Given specific forms of the degree distributions of both partitions, therefore, structural properties of the unipartite projection can be explored using tools of network theory. In particular, see [13] for general methods on deriving degree distributions for network projections.

Furthermore, given that the network display heterogeneous degree distributions under certain circumstances, an interesting question is whether high degree nodes tend to be connected to other high degree nodes or to lower degree nodes. This property is termed assortativity or disassortativity, respectively [2]. Social networks, for example, tend to be assortative in that people with a high number of connections often share links with other high degree nodes in the network. For common mycorrhizal networks, on the other hand, there is some empirical evidence that older trees with a more established root system form indirect connections to younger trees through mycorrhizal networks in order to aid seedling establishment [1]. The conditions under which the projected networks display disassortativity may therefore be explored using mycorrhizal network models.

The extension to 2-dimensional coordinates would also be straighforward. In particular, see [7] for an extension to a hyperbolic geometry.

Additionally, there is a need to ease some of the ecological restrictions made in the development of the model, in order to approach a more realistic scenario. In particular, there is going to be competition among fungi when forming links. Also, adding time dynamics to the model, nodes will reproduce and be removed from the network and the internal structure of links may change with time. In fact, the concept of partner choice plays an important role in discussions about mutualisms, both in terms of initiating interactions and re-evaluating already formed links [14].

Finally, the structure of the network in the end influences the dynamics that take place on the network, in this case with respect to the flow of nutrients between organisms. A fundamental question, therefore, is how the structure of the network influences the flow of nutrients taking place on the network, and vice versa. This would be a more realistic example of an adaptive network [15], where dynamics of the network and dynamics on the network influence eachother.

References

[1] Suzanne W. Simard, Kevin J. Beiler, Marcus A. Bingham, Julie R. Deslippe, Leanne J. Philip, and François P. Teste. Mycorrhizal networks: Mechanisms, ecology and modelling. *Fungal Biology Reviews*, 26(1):39 – 60, 2012. Hyphal networks: mechanisms, modelling and ecology.

[2] A-L. Barabási and M. Pósfai. *Network Science*. Cambridge University Press, 2016.

[3] P. Erdös and A. Rényi. On random graphs i. *Publicationes Mathematicae Debrecen*, 6:290–297, 1959.

[4] E. N. Gilbert. Random graphs. *Annals of Mathematical Statistics*, 30(4):1141–1144, 12 1959.

[5] A-L. Barabasi and Reka A. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.

[6] G. Caldarelli, A. Capocci, P. De Los Rios, and M. A. Muñoz. Scale-free networks from varying vertex intrinsic fitness. *Phys. Rev. Lett.*, 89:258702, Dec 2002.

[7] D. Krioukov, F. Papadopoulos, M. Kitsak, A. Vahdat, and M. Boguñá. Hyperbolic geometry of complex networks. *Physical Review E*, 82:036106, Sep 2010.

[8] M. Kitsak and D. Krioukov. Hidden variables in bipartite networks. *Physical Review E*, 84:026114, Aug 2011.

[9] M. Kitsak, F. Papadopoulos, and D. Krioukov. Latent geometry of bipartite networks. *Physical Review E*, 95:032309, Mar 2017.

[10] K. J. Beiler, D. M. Durall, S. W. Simard, S. A. Maxwell, and A. M. Kretzer. Architecture of the wood-wide web: Rhizopogon spp. genets link multiple douglas-fir cohorts. *New Phytologist*, 185(2):543–553, 2010.

[11] D. Southworth, X-H. He, W. Swenson, C. S. Bledsoe, and W. R. Horwath. Application of network theory to potential mycorrhizal networks. *Mycorrhiza*, 15(8):589 – 595, 2005.

[12] R. Albert, H. Jeong, and A. Barabási. Error and attack tolerance of complex networks. *Nature*, 406:378 – 382, 2000.

[13] D. Vasques Filho and D. R. J. O'Neale. Degree distributions of bipartite networks and their projections. *Physical Review E*, 98, Aug 2018.

[14] G. D. A. Werner and E. T. Kiers. Partner selection in the mycorrhizal mutualism. *New Phytologist*, 205(4):1437–1442, 2015.

[15] T. Gross and B. Blasius. Adaptive coevolutionary networks: a review. *Journal of The Royal Society Interface*, 5(20):259–271, 2008.

A   Degree Distributions of Network Partitions

Following the approach in [8], we can derive a formula for the degree distribution of each network partition. For simplicity, we consider the network model first without fitness parameters, since the extension to two parameters is straightforward. Also, deriving the degree distribution for one partition will follow the same procedure as for the other.

First, however, a general note on probability generating functions.

A.1   Generating Functions

Generating functions provide a way of encoding information about a probability distribution in the form of a power series. For a discrete random variable X with probability mass function $p(k) = Pr(X = k) = p_k$, the corresponding probability generating function (PGF) is defined as

$$F(z) = \mathrm{E}[z^X] = \sum_{k=0}^{\infty} p_k z^k. \tag{33}$$

Each value p_k can be retrieved by taking the kth derivative of $F(z)$ evaluated at $z = 0$:

$$p_k = \frac{1}{k!} \frac{d^k}{dz^k} F(z) \bigg|_{z=0}. \tag{34}$$

Also, the expected value of the distribution is given by

$$\mathrm{E}[X] = \sum_{k=0}^{\infty} k p_k = \frac{d}{dz} F(z) \bigg|_{z=1}. \tag{35}$$

An important property of probability generating functions is that given n independent random variables $X_1, X_2, \ldots, X_n$, with PGFs $F_1(z), F_2(z), \ldots, F_n(z)$, respectively, the PGF $F(z)$ for the sum of those variables will be the product of their individual PGFs, i.e.

$$F(z) = \prod_{i=1}^{n} F_i(z). \tag{36}$$

This is a simple consequence of the multiplicativity of the expected value for independent random variables.

A.2   Deriving the Degree Distribution

In order to derive the degree distribution for the bottom partition of the network (or equivalently for the top partition) we first define a function $f(k|\phi)$ for the probability of a node having degree k given that it has position ϕ. The complete degree distribution is then given by

$$f(k) = \sum_{\phi} \rho(\phi) f(k|\phi). \tag{37}$$

Given a node in the bottom partition with fixed position ϕ, the degree of that node will depend on the outcome of a series of trials with probability $r = r(\theta, \phi)$, for each value of θ in the network. With $\rho(\theta) N_\top$ expected number of nodes with position θ, the total degree of a node with position ϕ

is therefore given by a sum of binomially distributed variables with probability $r(\theta, \phi)$ and $\rho(\theta)N_\top$ number of trials, over all values of θ.

Since the generating function for a sum of independent random variables is the product of the generating functions of each term in the sum, the generating function for the conditional degree distribution is given by

$$F(z|\phi) = \prod_\theta [1 + r(\theta, \phi)(z - 1)]^{\rho(\theta)N_\top}, \tag{38}$$

using the PGF for a binomial distribution. From equation (35), the average degree of a bottom node given position ϕ is then

$$\bar{k}_\perp(\phi) = \left.\frac{d}{dz}F(z|\phi)\right|_{z=1} = N_\top \sum_\theta \rho(\theta)r(\theta, \phi), \tag{39}$$

and the average degree of the whole partition is calculated as

$$\bar{k}_\perp = \sum_\phi \rho(\phi)\bar{k}_\perp(\phi). \tag{40}$$

In its current form, the average degree increases with the size of the other partition. However, assuming the connection probability is inversely proportional to the size of the other partition, i.e. $r(\theta, \phi) \propto 1/N_\top$, the average number of links per node will be restricted even for large network size and the network will be sparse. With the assumption of a sparse network, and taking the logarithm of equation (38), we get

$$\ln F(z|\phi) = N_\top \sum_\theta \rho(\theta) \ln[1 + r(\theta, \phi)(z - 1)] \approx \bar{k}_\perp(\phi)(z - 1). \tag{41}$$

The generating function for the conditional degree distribution is therefore

$$F(z|\phi) \approx e^{\bar{k}_\perp(\phi)(z-1)}, \tag{42}$$

which is the generating function for a Poisson distribution with parameter $\bar{k}_\perp(\phi)$. This means that

$$f(k|\phi) \approx \frac{1}{k!}[\bar{k}_\perp(\phi)]^k e^{-\bar{k}_\perp(\phi)}. \tag{43}$$

The full degree distribution is now given by equation (37).

A.3 Adding Fitness Parameters

Including also a second parameter in the form of a fitness value, we first calculate the conditional degree distribution for a node with a specific position and fitness value, leading to an average degree

$$\bar{k}_\perp(\phi, \eta_\perp) = N_\top \sum_{\theta, \eta_\top} \rho(\theta)\rho(\eta_\top)r(\theta, \phi, \eta_\top, \eta_\perp), \tag{44}$$

Similarly, assuming $r(\theta, \phi, \eta_\top, \eta_\perp)$ is sufficiently small for large network size, the conditional degree distribution is a Poisson distribution with parameter $\bar{k}_\perp(\phi, \eta_\perp)$:

$$f(k|\phi, \eta_\perp) \approx \frac{1}{k!} \cdot [\bar{k}_\perp(\phi, \eta_\perp)]^k \cdot e^{-\bar{k}_\perp(\phi, \eta_\perp)}. \tag{45}$$

The average degree and the degree distribution for the whole partition is now given by the sum over all values of ϕ and $\eta_\perp$:

$$\bar{k}_\perp = \sum_{\phi,\eta_\perp} \rho(\phi)\rho(\eta_\perp)\bar{k}_\perp(\phi,\eta_\perp), \tag{46}$$

$$f(k) = \sum_{\phi,\eta_\perp} \rho(\phi)\rho(\eta_\perp)f(k|\phi,\eta_\perp). \tag{47}$$

B Calculations of Degree Distributions for Specific Cases

B.1 Uniform Spatial Distribution

With a uniform spatial distribution for both partitions and a delta function as fitness distribution, the connection probability is

$$r(\theta, \phi, \eta_\top, \eta_\perp) = r(\theta, \phi) = \frac{r_\perp}{N_\top} \cdot e^{-\alpha \cdot d(\theta, \phi)}. \tag{48}$$

Assuming, without loss of generality, a fixed $\phi \in [0, \pi]$, the average degree for a fungal node with that position is given by

$$
\begin{aligned}
\bar{k}_\perp(\phi) &= \frac{N_\top}{2\pi} \int_0^{2\pi} r(\theta, \phi) d\theta = \frac{r_\perp}{2\pi} \int_0^{2\pi} e^{-\alpha \cdot d(\theta, \phi)} d\theta \\
&= \frac{r_\perp}{2\pi} \cdot \left[\int_0^\phi e^{-\alpha \cdot (\phi - \theta)} d\theta + \int_\phi^{\phi+\pi} e^{-\alpha \cdot (\theta - \phi)} d\theta + \int_{\phi+\pi}^{2\pi} e^{-\alpha \cdot (2\pi - (\theta - \phi))} d\theta \right] = \dots \\
&= \frac{r_\perp}{\alpha\pi} \cdot [1 - e^{-\alpha\pi}],
\end{aligned} \tag{49}
$$

using equation (39).

Similarly, the average degree for a tree node with fixed position θ is

$$\bar{k}_\top(\theta) = \frac{N_\perp}{N_\top} \cdot \frac{r_\perp}{\alpha\pi} \cdot [1 - e^{-\alpha\pi}] = \frac{r_\top}{\alpha\pi} \cdot [1 - e^{-\alpha\pi}], \tag{50}$$

where $r_\top = r_\perp N_\perp / N_\top$. The difference comes from the fact that the average grows with the size of the bottom partition rather than the size of the top partition in equation (39).

Since the conditional average is independent of position, $\bar{k}_\perp = \bar{k}_\perp(\phi)$ and $\bar{k}_\top = \bar{k}_\top(\theta)$. Therefore, from equations (37) and (43), the two degree distributions are Poisson distributions with average $\bar{k}_\perp$ and $\bar{k}_\top$, respectively.

B.2 Spatial Heterogeneity

Given delta functions for each fitness distribution, uniform spatial distribution for trees and $\rho(\phi) = \text{III}_{2\pi}(\phi)$, now focusing on the degree distribution of trees, the average degree given position θ simplifies to

$$\bar{k}_\top(\theta, \eta_\top) = \bar{k}_\top(\theta) = r_\top e^{-\alpha\theta}, \tag{51}$$

where $r_\top = r_\perp N_\perp / N_\top$. Using equation (37), the full degree distribution is now given by

$$
\begin{aligned}
g(k) &= \frac{1}{\pi} \int_0^\pi \frac{1}{k!} [\bar{k}_\top(\theta)]^k e^{-\bar{k}_\top(\theta)} d\theta \\
&= \frac{1}{\pi} \int_0^\pi \frac{1}{k!} \left[r_\top e^{-\alpha\theta} \right]^k \exp\left[-r_\top e^{-\alpha\theta} \right] d\theta
\end{aligned} \tag{52}
$$

Making use of a substitution $x = r_\top e^{-\alpha\theta}$,

$$
\begin{aligned}
g(k) &= -\frac{1}{\alpha\pi k!} \int_{r_\top}^{r_\top e^{-\alpha\pi}} x^{k-1} e^{-x} dx \\
&= \frac{1}{\alpha\pi k!} \cdot \left[\Gamma(k, r_\top e^{-\alpha\pi}) - \Gamma(k, r_\top) \right],
\end{aligned} \tag{53}
$$

where Γ denotes the incomplete gamma function.

B.3 Exponential Fitness Distribution

With a general fitness distribution, the average degree of a tree node given position θ and fitness value $\eta_\top$ is given by

$$
\begin{aligned}
\bar{k}_\top(\theta, \eta_\top) &= \frac{\eta_\top}{\langle \eta \rangle} r_\top \int_{\eta_\perp} \eta_\perp \rho(\eta_\perp) d\eta_\perp \int_\phi e^{-\alpha \cdot d(\theta,\phi)} d\phi \\
&= \frac{\eta_\top}{\langle \eta_\top \rangle} r_\top \int_\phi e^{-\alpha \cdot d(\theta,\phi)} d\phi
\end{aligned}
\tag{54}
$$

With a uniform spatial distribution for fungi,

$$
\bar{k}_\top(\theta, \eta_\top) = \bar{k}_\top(\eta_\top) = \frac{\eta_\top}{\langle \eta_\top \rangle} \frac{r_\top}{\alpha \pi} \cdot [1 - e^{-\alpha \pi}] = K \cdot \frac{\eta_\top}{\langle \eta_\top \rangle},
\tag{55}
$$

where $K = r_\top [1 - e^{-\alpha \pi}]/(\alpha \pi)$.

Assuming also a uniform distribution of trees, the total degree distribution is given by

$$
\begin{aligned}
g(k) &= \int_{\eta_\top} \frac{1}{k!} \cdot [\bar{k}_\top(\eta_\top)]^k e^{-\bar{k}_\top(\eta_\top)} \rho(\eta_\top) d\eta_\top \\
&= \int_{\eta_\top} \frac{1}{k!} [K\eta_\top/\langle \eta_\top \rangle]^k \exp[-K\eta_\top/\langle \eta_\top \rangle] \rho(\eta_\top) d\eta_\top
\end{aligned}
\tag{56}
$$

Finally, assuming $\rho(\eta_\top) = \lambda_\top e^{-\lambda_\top \eta_\top}$, and making a substitution $x = K\eta_\top/\langle \eta_\top \rangle$, the degree distribution is given by

$$
\begin{aligned}
g(k) &= \frac{1}{K \cdot k!} \int_0^\infty x^k e^{-x(1+1/K)} dx = \{y = x(1 + 1/K)\} \\
&= \frac{1}{K \cdot k!} \cdot (1 + 1/K)^{-k-1} \int_0^\infty y^k e^{-y} dy \\
&= \frac{1}{1 + K} \cdot (1 + 1/K)^{-k},
\end{aligned}
\tag{57}
$$

since the second integral equals $\Gamma(k + 1)$.

B.4 Point Habitat with Exponential Fitness Distribution

Assuming $\rho(\phi) = \text{III}_{2\pi}(\phi)$ and $\rho(\eta_\top) = \lambda_\top e^{-\lambda_\top \eta_\top}$,

$$
\bar{k}_\top(\theta, \eta_\top) = \frac{\eta_\top}{\langle \eta \rangle} r_\top \int_{\eta_\perp} \eta_\perp \rho(\eta_\perp) d\eta_\perp \int_\phi e^{-\alpha \cdot d(\theta,\phi)} d\phi = \frac{\eta_\top}{\langle \eta_\top \rangle} r_\top e^{-\alpha\theta} = K(\theta) \cdot \frac{\eta_\top}{\langle \eta_\top \rangle},
\tag{58}
$$

where $K(\theta) = r_\top e^{-\alpha\theta}$. In this case, we have a dependence on both θ and $\eta_\top$. Similar to before,

$$
\begin{aligned}
g(k) &= \frac{1}{\pi} \int_0^\pi \frac{1}{K(\theta)} \cdot (1 + 1/K(\theta))^{-k-1} d\theta = \{x = 1/K(\theta)\} \\
&= \frac{1}{\alpha\pi} \int_{r_\top^{-1}}^{r_\top^{-1} e^{\alpha\pi}} (1 + x)^{-k-1} dx \\
&= \frac{1}{\alpha\pi k} \cdot \left[(1 + r_\top^{-1})^{-k} - (1 + r_\top^{-1} e^{\alpha\pi})^{-k} \right]
\end{aligned}
\tag{59}
$$

C   Numerical Simulations

Below, figures are presented displaying the results of numerical calculations for the different choices of spatial- and fitness distributions.

C.1   Homogeneous Spatial Distribution

The first case with a homogeneous spatial distribution is shown in Figure 9. A higher value of α shifts the distribution to the left.
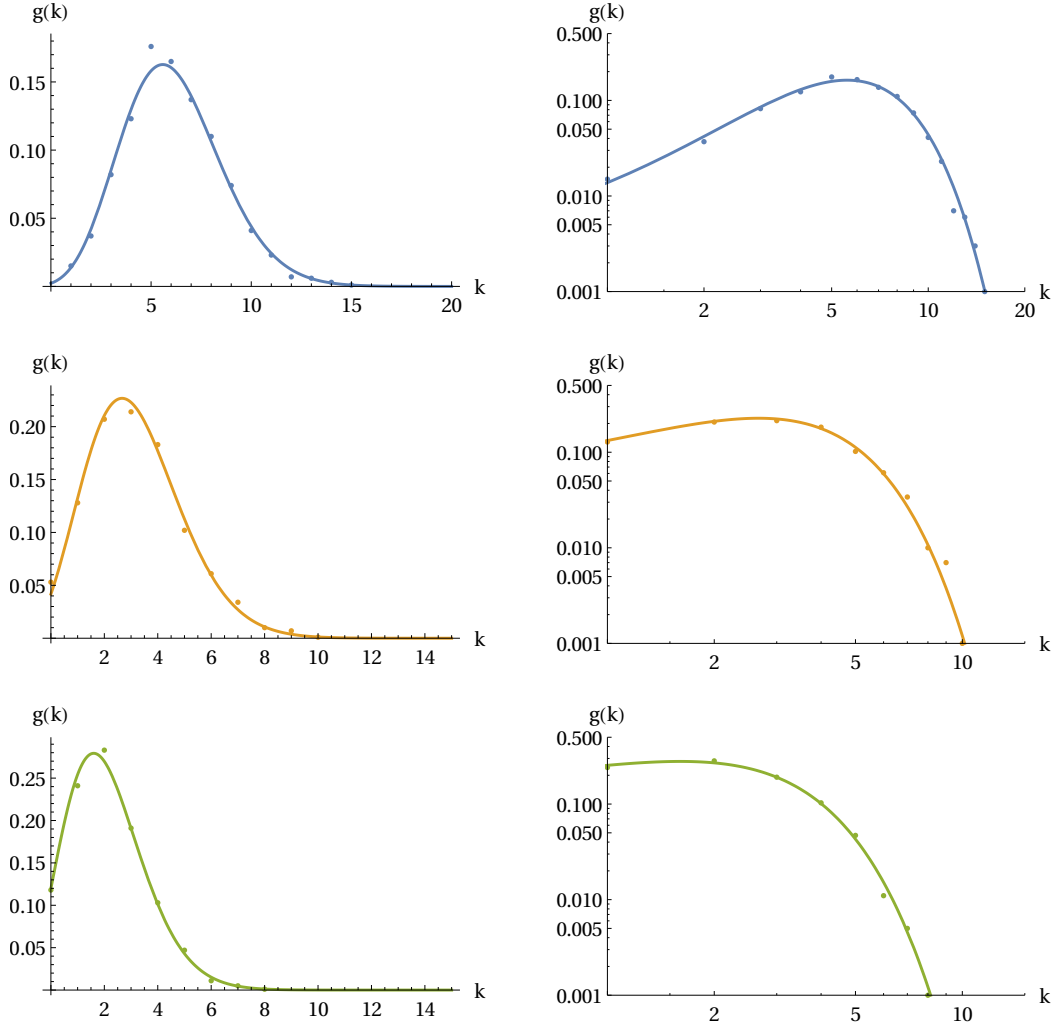

Figure 9: Numerical simulations for the degree distribution of trees in comparison with analytical results in the case of a homogeneous spatial distribution and no difference in fitness values. The simulations were performed with $N_\top = N_\perp = 1000$, $r_\perp = 20$ and three different values for α: $\alpha = 1$ (blue), $\alpha = 2$ (orange) and $\alpha = 3$ (green). Each case is displayed using both a linear (left) and a logarithmic scale (right).

C.2   Heterogeneous Spatial Distribution

Next, the case introducing a heterogeneous spatial distribution of fungi is shown in Figure 10. Here, the power law regime is more apparent for higher α.
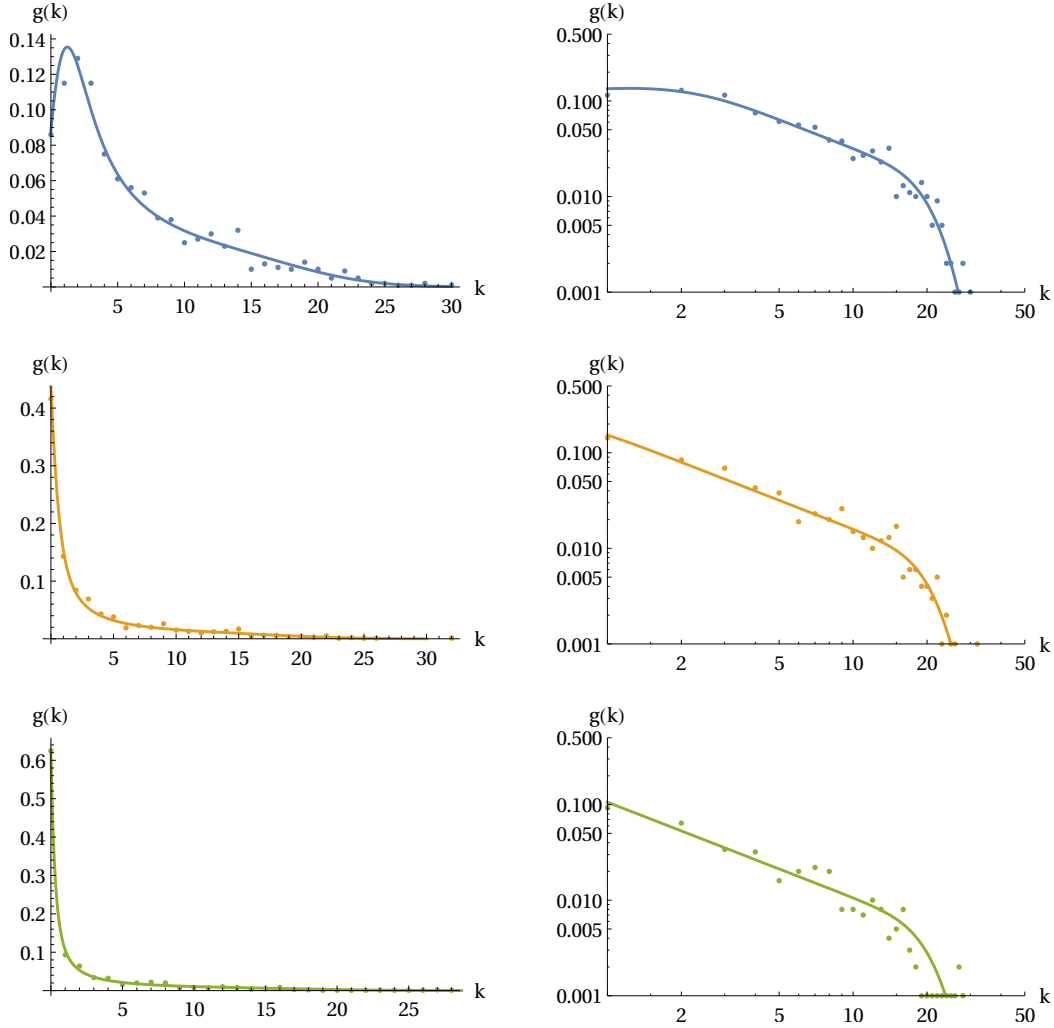

Figure 10: Numerical simulations for the degree distribution of trees in comparison with analytical results in the case of a point habitat for fungi, a uniform distribution for trees and no difference in fitness values. The simulations were performed with $N_\top = N_\perp = 1000$, $r_\perp = 20$ and three different values for α: $\alpha = 1$ (blue), $\alpha = 2$ (orange) and $\alpha = 3$ (green). Each case is displayed using both a linear (left) and a logarithmic scale (right).

C.3   Exponential Fitness Distribution

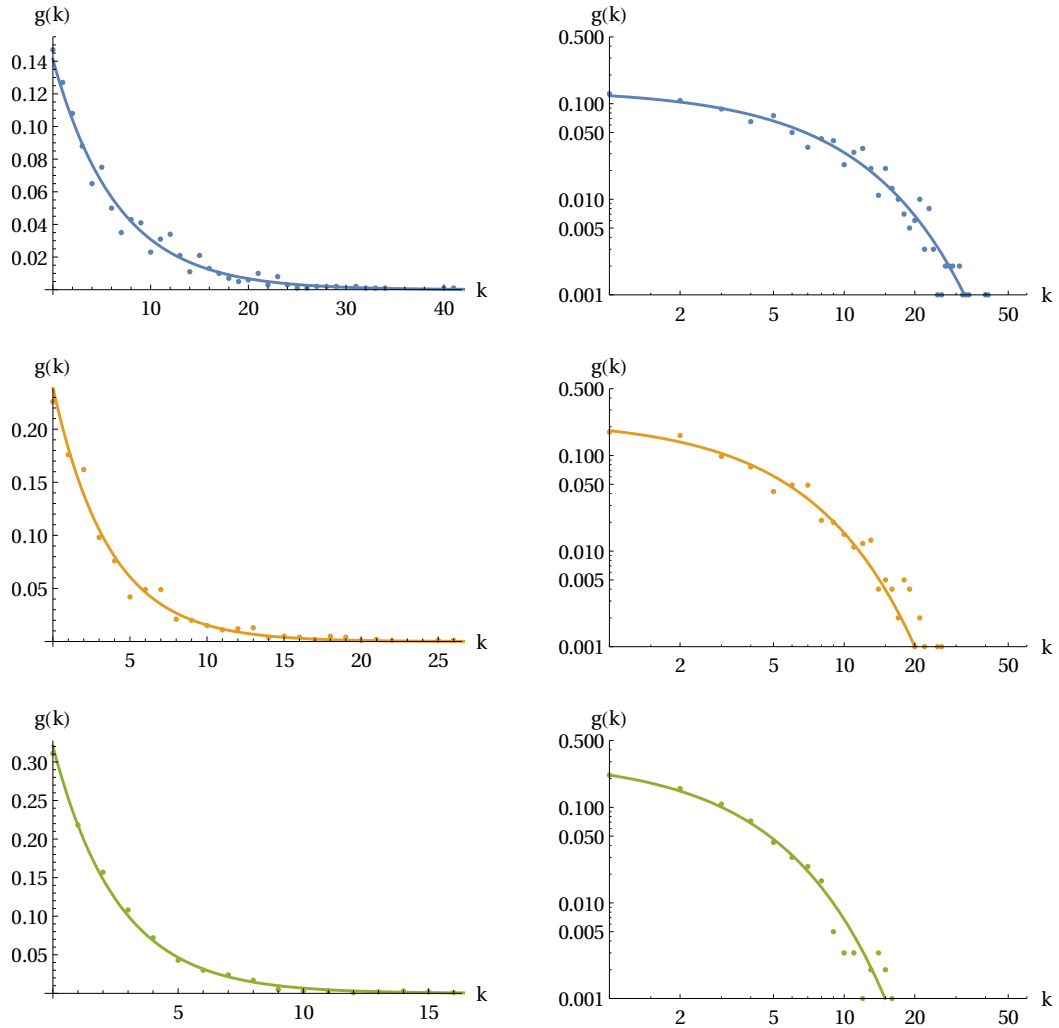The results for an exponential fitness distribution are presented in Figure 11.

Figure 11: Numerical simulations for the degree distribution of trees in comparison with analytical results in the case of an exponential fitness distribution for trees and uniform spatial distributions for both trees and fungi. The simulations were performed with $N_\top = N_\perp = 1000$, $r_\perp = 20$ and three different values for α: $\alpha = 1$ (blue), $\alpha = 2$ (orange) and $\alpha = 3$ (green). Each case is displayed using both a linear (left) and a logarithmic scale (right).

C.4 Point Habitat with Exponential Fitness Distribution

Finally, simulation results for the combined effects of spatial heterogeneity and exponentially distributed fitness values are displayed in Figure 12. Interestingly, the simulation results include nodes of fairly high degree compared to the other cases.
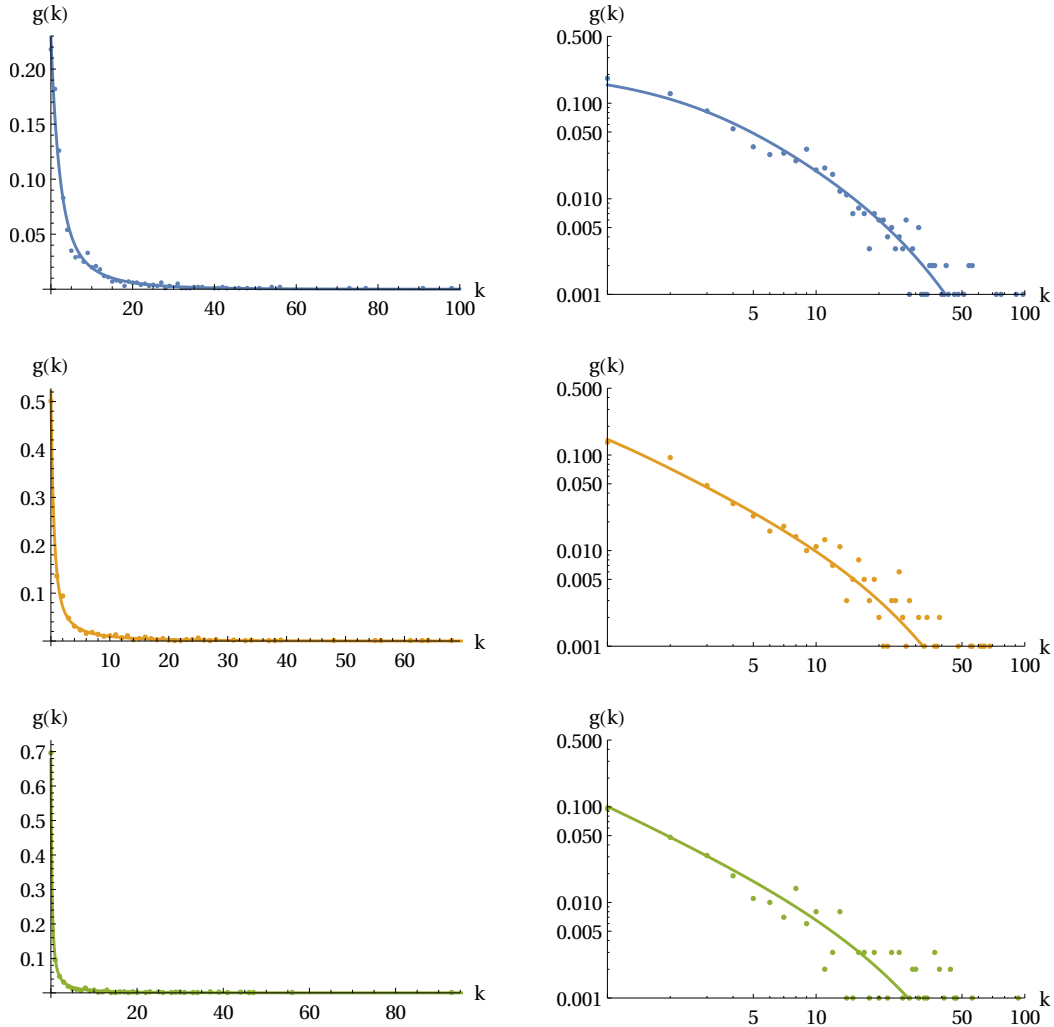

Figure 12: Numerical simulations for the degree distribution of trees in comparison with analytical results in the case of a point habitat for fungi, a uniform spatial distribution of trees and exponentially distributed fitness values for tree nodes. The simulations were performed with $N_\top = N_\perp = 1000$, $r_\perp = 20$ and three different values for α: $\alpha = 1$ (blue), $\alpha = 2$ (orange) and $\alpha = 3$ (green). Each case is displayed using both a linear (left) and a logarithmic scale (right).