Evolving Inhomogeneous Random Structures

By

Tejas Iyer

A thesis submitted to
the University of Birmingham
for the degree of
DOCTOR OF PHILOSOPHY

Combinatorics, Probability and Algorithms Group
School of Mathematics
College of Engineering and Physical Sciences
University of Birmingham
February 2021

ABSTRACT

We introduce general models of evolving, inhomogeneous random structures, where in each of the models either one or several nodes arrive at a time, and are equipped with random, independent weights. In the two evolving tree models we study, an existing vertex is chosen at each time-step with probability proportional to its fitness function, which is a function of its weight, and possibly the weights of its neighbours, and the newly arriving node(s) connect to it. The third models, with parameter d consist of evolving sequences of $(d-1)$-dimensional simplicial complexes. At each time-step a $(d-1)$-simplex is sampled with probability proportional to a function of the weights of the vertices the $(d-1)$-simplex contains. In both variants, Model **A** and Model **B**, for each subset S of size $(d-2)$, we add the simplex consisting of S and the single new-coming vertex. Additionally, in Model **B**, the selected simplex is removed from the simplicial complex.

In each of the models we study the limiting proportion of vertices in the structure with a given degree, showing that, in general, this limit exists in probability, and behaves like a type of *generalised geometric distribution.* In the evolving tree models, we actually study a more general quantity: the empirical measures associated with the number of vertices with a given degree and weight. With regards to this quantity, when normalised by the size of the network, we also show that the limit exists and belongs to a certain universal class. Depending on various assumptions, we prove that for any measurable set, the measure of that set converges either almost surely or in probability to its measure under this deterministic limit.

In the evolving tree models, we also study another quantity: the empirical measure corresponding to the proportion of edges in the structure with endpoint having a given weight. We show that, when normalised by the number of edges in the tree, under certain assumptions, this quantity also converges to a deterministic limiting measure, in the sense that for any measurable set, the measure of that set converges either almost surely. However, when the trees take certain forms, which we call the GPAF-tree, or the PANI-tree, we show that interesting, non-trivial behaviour can emerge when these assumptions fail. In particular, with regards to the GPAF-tree, we show that this model can exhibit *condensation* where a positive proportion of edges accumulate around vertices with weight that maximises the reinforcement of their fitness, or, more drastically, have a *degenerate* limiting degree distribution where the entire proportion of edges accumulate around these vertices. We also show that the *condensation* phenomenon extends to the more general PANI-tree model. As we will show, the latter two models have limiting distribution of degrees that behaves like an ‘averaged’ power law, which may be of interest when considering them as toy models for the evolution of complex networks.

DEDICATION

Dedicated to my thathas, and Tuffy.

ACKNOWLEDGMENTS

First, I'd like to thank Nikos for his guidance, support and encouragement, and calm, relaxed demeanor throughout the highs and lows of this PhD. I really appreciate it. I'd also like to thank Cécile and Henning for their guidance when working on the dynamical simplicial complexes project. Thanks also to Richard for acting as my co-supervisor.

I'd also like to thank all my friends, who have made my stay in Birmingham a memorable one. Thanks to Dryden, Bill and Spencer, for all the antics at the Spinney; Dave and Rahim for all the memories of watching/discussing/continuing to discuss the cricket and all the members of the Birmingham salsa community, including DCL. I'd also like to thank the occupants of 318, and my fellow library study buddies: Yahya, Gimmy and Soodeh for all the coffee breaks, and Yahya again, for doubling up as a gym/falafel munch/last-minute-lease takeover buddy, and always being around to talk to. Thanks also to Pearl for our talks over tea, and her valuable contributions to organising last minute games of Avalon; and to Padraig, for all the adventures we had over two years, the morning coffees and the lessons about Irish culture. Thanks to Zalesh, for the constant stream of messages over Whatsapp, and Ashwin and Rory for their continued friendship over the years, and finally, the numerous other people I haven't mentioned who have made the last three years a positive one.

Finally, thanks to my family for all the support and encouragement over the years, including the housing during the pandemic. Thanks to Shreyas, for being a source of sage advice, to Dad, who decided I shouldn't be alone and embarked on a PhD adventure with me, and Mum for all the phone calls and the visits, the travels around Europe, putting up with my complaints, and all the food.

Through the unknown, remembered gate
When the last of the earth left to discover
Is that which was the beginning

T.S. Elliot

STATEMENT OF ORIGINALITY

All of the original mathematical results in this thesis come from papers where I was a contributor, namely [43], where I was a sole author, [36], which was a collaborative project with my supervisor Nikolaos Fountoulakis, and [37], which was a collaborative project with Nikolaos Fountoulakis, Cécile Mailler and Henning Sulzbach.

The contents of Chapter 1 is wholly my own contribution, except for Section 1.2.4, which includes parts of the introduction of [37]. Chapter 2 includes the results from [43], except for the proof of Lemma 2.4.5, which comes from [37]. Chapter 3 includes results from [36] and is also mostly my own contribution, except for the calculations of the limiting vectors related to Urn E following the statement of Corollary 3.2.8. Finally, Chapter 4 includes results from [37], and thus, as with the other parts of this thesis sourced from this paper, may be regarded as the equal contribution of Nikolaos Fountoulakis, Cécile Mailler, Henning Sulzbach and myself.

Contents

Chapter One

Introduction

This chapter is an important foundational chapter in the reading of this thesis. In Section 1.1, we start with some motivation behind the areas of study this thesis concerns, namely, the probabilistic analysis of evolving inhomogeneous structures inspired by complex networks found in many applications. This section will be a rather gentle reading, and in Section 1.1.1 we include a number of pictures as illustrative examples. Section 1.2 may be regarded as a general review of the mathematical, and some of the physics literature related to this area. In Section 1.2, we start with some useful definitions in Section 1.2.1, review the well known *preferential attachment* and other recursive models in Section 1.2.2, review some evolving *inhomogeneous* models in Section 1.2.3 and, finally, some 'higher dimensional' models in Section 1.2.4. Then, in Section 1.3, we describe the models we introduce in this thesis, with helpful illustrations. In Section 1.3.1, we introduce some notation used throughout the thesis, the model of *generalised recursive trees with fitnesses* in Section 1.3.2, the model of *preferential attachment with neighbourhood influence* in Section 1.3.3 and finally, the *dynamical models of random simplicial complexes* in Section 1.3.4. Next, in Section 1.4 we describe the major quantities of interest in this thesis, namely, *degree distributions* in Section 1.4.1 and *edge distributions* in Section 1.4.2. Finally, in Section 2.1.2, we provide an general overview of the results of this thesis, stated and proved in the subsequent chapters.

In general, in this thesis, we will assume the reader has a good understanding of probability theory, including, for example, theory related to ‘couplings’, Markov chains and martingales, and a rudimentary, minimal understanding of graph theory. This chapter, and especially Section 1.1, however, are quite mild. The subsequent chapters in this thesis are ordered by increasing difficulty, and the interested reader may wish to skip some of the more technical arguments in Chapter 4 upon first reading.

1.1 Introduction to Complex Networks

Networks are ubiquitous structures, found almost everywhere in nature and society. When used to model complex systems, networks find applications in areas as diverse as computer science, biology and sociology. Advances in science over the last 30 years have led to an increased understanding of the properties of these networks, see, for example, [66, 77, 16, 67]. These advances have shown that while these networks may come from diverse settings, they possess typical, non-trivial features. In particular, they are generally *large*, of the order of billions of nodes; yet *sparse*, which means that the number of links in the network is at most the same order of magnitude as the size of the network. They are also *dynamic*, which refers to the fact that the nodes and links in a network are constantly evolving. In addition, networks are known to exhibit a *small world* phenomenon. This phenomenon, first popularised by Milgram in [60], refers to the fact that, despite the large size of the network and the fact that it is sparse, the typical distance between nodes is generally very ‘small’. Finally, these networks are known to display *scale-free degree distributions*. The *degree* of a node is the number of links incident to it, and this latter property refers to the fact that the proportion of nodes of degree k in the network tends to scale like $k^{-\alpha}$ for some $\alpha > 0$; often with α between 2 and 3. This latter property means that, if one plots the logarithm of number of nodes against the logarithm of the degree, one obtains a linear plot, as illustrated in Figure 1.1 below. Indeed, if N_k denotes the number of nodes with degree k, then if $N_k \approx k^{-\alpha}$,

$$\log N_k \approx -\alpha \log k,$$

which results in a linear relationship.

Scale-Free Degree Distributions

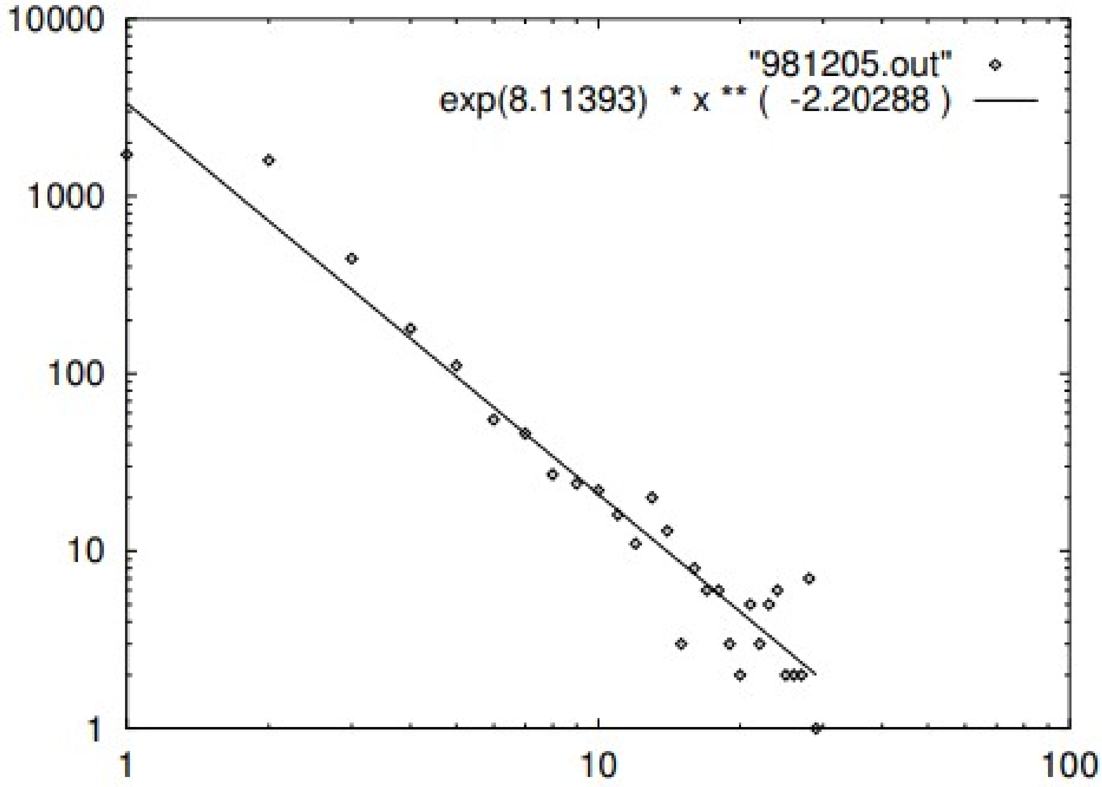


Figure 1.1: This plot, from a well known paper [35], is a log-log plot of number of nodes against their degree in a sub-network of the internet known as an ‘autonomous system’. The data seems to indicate a power law relationship.

1.1.1 Illustrative Examples of Complex Networks

Below are some illustrative examples of complex networks. The first example relates to the 'blogosphere', consisting of nodes from the internet corresponding to 'blogs'.

The Blogosphere

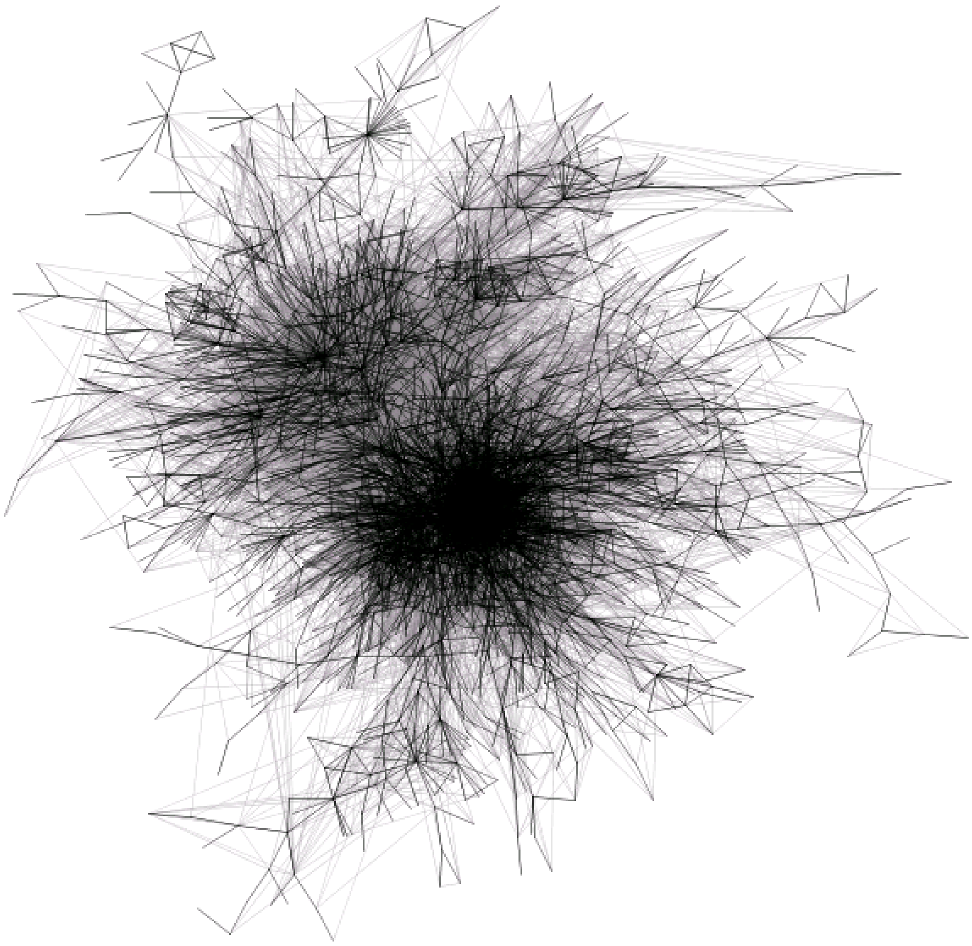


Figure 1.2: This illustration shows the links in the network associated with the blogosphere, where two nodes, associated with blogs, are linked one blog refers to the other. Taken from https://datamining.typepad.com/gallery/blog-map-gallery.htm - [42].

Our next examples are 'protein-protein interaction' network, which are common networks found in biological applications. In these networks, the nodes represent proteins and two nodes are connected by a link if their respective proteins take part in a common chemical reaction.

Protein-Protein Interaction Network: Yeast Cell

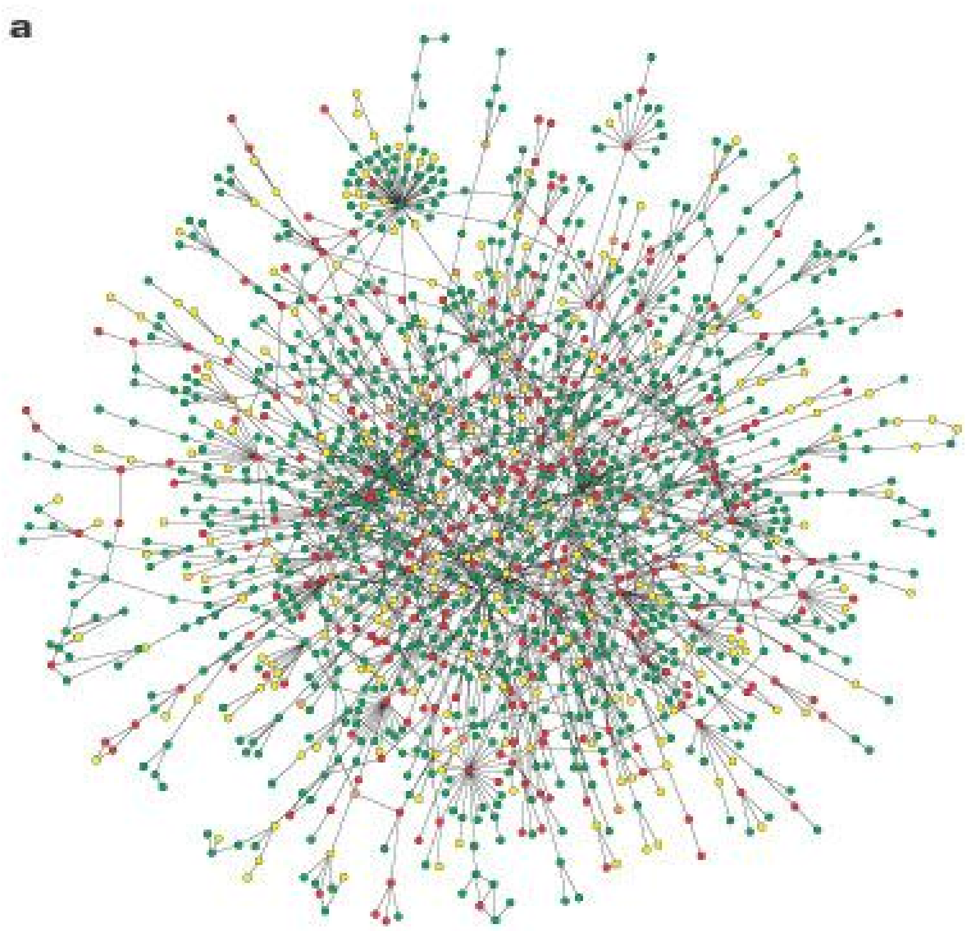


Figure 1.3: This illustration shows the nodes and links in the protein-protein interaction network associated with a yeast cell. Taken from [46].

Protein-Protein Interaction Network: Human Body

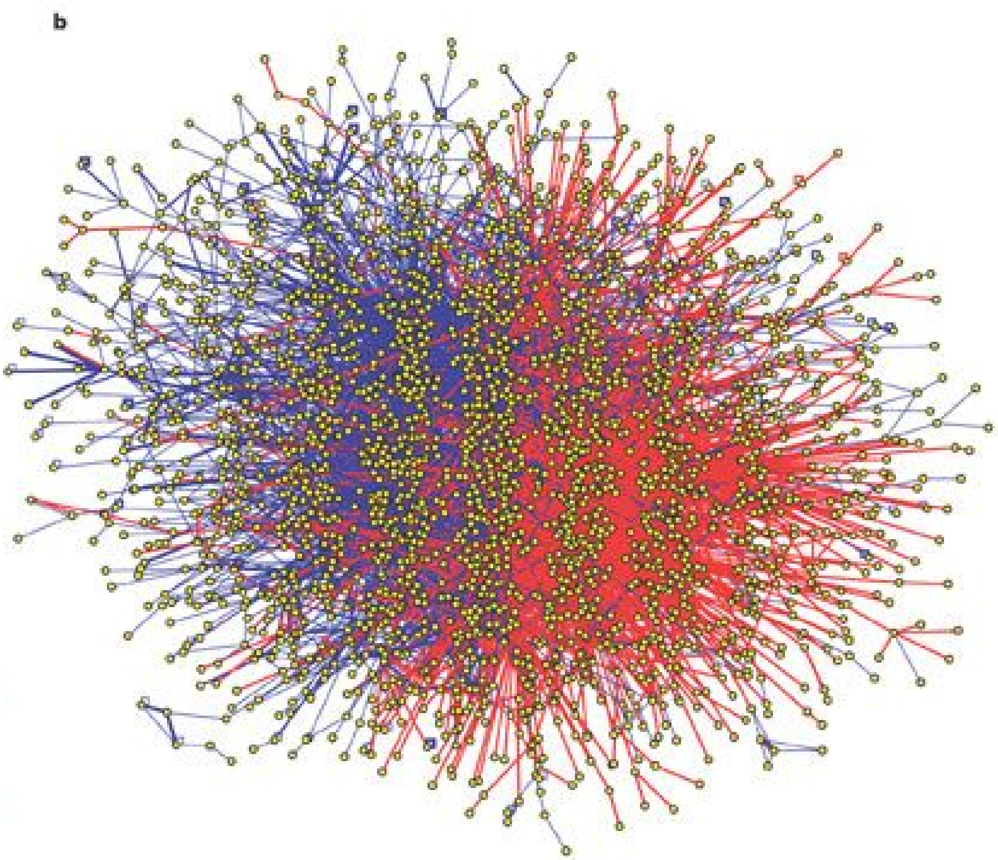


Figure 1.4: This illustration shows the nodes and links in the protein-protein interaction network associated with the human body. Taken from [71].

1.2 Generative Models of Evolving Complex Networks

There are a number of existing models in the literature that aim to generate networks with similar properties to the complex networks described in the previous section. The benefit of these models is that they offer insights into the possible mechanisms that lead to the emergence of some of the particular features associated with complex networks, which may in turn yield a deeper understanding of the way these networks behave. In this section we describe some of these models and some of the mathematical results associated with them. First, however, we provide a brief overview of definitions related to trees, graphs and simplicial complexes, as these structures will be the main object of study in this thesis.

1.2.1 Trees, Graphs and Simplicial Complexes

We first recall the definitions of *graphs* and *directed graphs*.

Definition 1.2.1. *A* graph $G = (V, E)$ *is an ordered pair, where* V *is a finite set of* vertices*, and* E *is a finite set of pairs* $\{v, v'\} \subseteq V$*. A* directed graph*, or* digraph D *is an ordered pair* (V, A)*, where* V *is a finite set of vertices and* A *is a set of* directed edges *or* arcs *consisting of ordered pairs of vertices in* V*.*

Simplicial complexes are defined somewhat similarly:

Definition 1.2.2. *An* abstract simplicial complex $\mathcal{K} = (V, F)$*, where* V *is a finite set of* vertices *and* F *is a family of subsets of* V*, called* faces*, that is* downwards closed*, which means that for any* $\sigma \in F$*, if* $\sigma' \subseteq \sigma$ *then* $\sigma' \in F$*. A vertex set* V *together with an arbitrary family* F *may be turned into a simplicial complex in the natural way by taking the* downwards closure*, that is, adding the minimal number of subsets to* F *to make* F *downwards closed.*

Often, to simplify notation with graphs (or digraphs), we simply write G for a graph (V, E), and to specify a particular edge, we write $e \in G$ rather than $e \in E$. We apply a similar convention with simplicial complexes, so that, to specify a face σ in a simplicial complex, we write $\sigma \in \mathcal{K}$. Note also that there is a natural simplicial complex obtained from a graph, by choosing the set of faces to be the downwards closure of the set of edges corresponding to the graph.

Definition 1.2.3. *Given a face σ in a simplicial complex $\mathcal{K}$, we say σ has* dimension *s if it has cardinality $s+1$. We also call it an s*-face *or an s*-simplex. *For $s \in \mathbb{N} \cup \{0, -1\}$, we denote by $\mathcal{K}^{(s)}$ the subset of $\mathcal{K}$ consisting of all its s-faces. The* dimension *of $\mathcal{K}$ is defined to be the maximum s such that $\mathcal{K}^{(s)}$ is non-empty. If $\mathcal{K} = \varnothing$ we say it has dimension -1.*

Just as one often interprets, or visualises, a graph geometrically as a collection of 'dots', representing vertices, connected by 'lines' representing edges, it is often useful to identify simplicial complexes with their *geometric realisation*, which means that we view a d-face as the convex hull of $d+1$ points in $\mathbb{R}^d$. Thus, a 0-face may be interpreted as a point, a 1-face as a line, a 2-face as a triangle and a 3-face as a tetrahedron. This is also the reason for the use of the term 'dimension'.

Simplices in Dimensions 0, 1, 2 and 3.

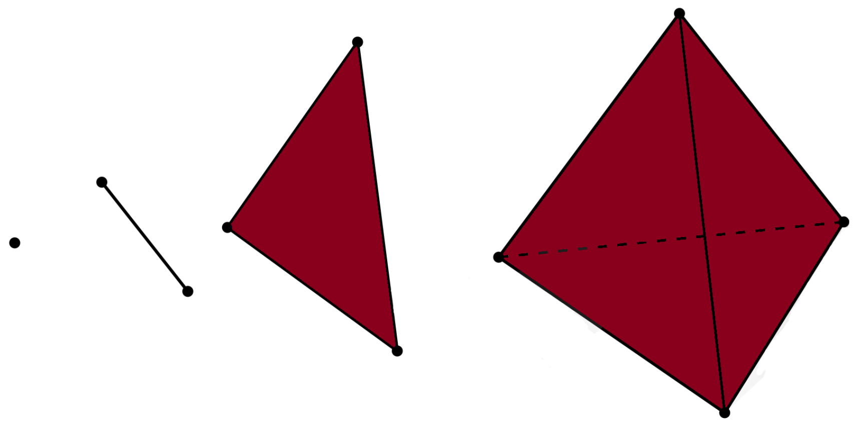


Figure 1.5: This illustration shows how one may interpret the faces of dimension 0, 1, 2 and 3 in a simplicial complex.

Finally, we recall the important concepts of *neighbourhood* and *degree.*

Definition 1.2.4. *Given a vertex v in a graph G, the* neighbourhood *of v in G is the set $\mathcal{N}(v, G) := \{v' \in G : \{v, v'\} \in G\}$. Likewise, if D is a directed graph, given a vertex $v \in D$,* *the* out-neighbourhood *of v in D is the set $\mathcal{N}^+(v, D) := \{v' \in D : (v, v') \in D\}$, and similarly the* in-neighbourhood *of v in D is the set $\mathcal{N}^-(v, D) := \{v' \in D : (v', v) \in D\}$. Finally, the* s-neighbourhood *of a vertex v in a simplicial complex $\mathcal{K}$ is the set $\mathcal{N}^{(s)}(v, \mathcal{K}) := \{\sigma \in \mathcal{K} : \sigma \cup \{v\} \in \mathcal{K}^{(s+1)}\}$.*

Thus, the 0-neighbourhood of a vertex v in a simplicial complex $\mathcal{K}$ coincides with the neighbourhood of the vertex v in the graph underlying the simplicial complex. We call this graph the *skeleton graph* associated with the complex. Finally, the *degree* corresponds to the size of the relevant neighbourhood:

Definition 1.2.5. *Given a vertex v in a graph G, the* degree *of v in G is* $\deg(v, G) := |\mathcal{N}(v, G)|$. *Likewise, for a vertex v in a directed graph D, the* out-degree *of v in D is* $\deg^+(v, D) := |\mathcal{N}^+(v, D)|$ *and similarly, the* in-degree *of v is* $\deg^-(v, D) := |\mathcal{N}^-(v, D)|$. *Finally, the s-*degree *of a vertex v in a simplicial complex $\mathcal{K}$ is* $\deg^{(s)}(v, \mathcal{K}) := \left|\mathcal{N}^{(s)}(v, \mathcal{K})\right|$. *For brevity, we also write* $\deg(v, \mathcal{K}) := \deg^{(0)}(v, \mathcal{K})$.

1.2.2 Preferential Attachment and other Recursive Models

A common framework for generating graphs that behave like complex networks is to consider evolving models where vertices arrive one at a time, and connect to existing vertices in the graph. These models are inherently dynamic, by construction, and if the number of edges added at each time-step is uniformly bounded from above, will also produce sparse graphs. In addition, in their seminal paper [8], Albert and Barabási, observed that the properties of being scale-free and having a small-world phenomenon emerged naturally in a model where vertices arrive one at a time, and display a "preference" to popular vertices - more precisely, connect to existing vertices with probability proportional to their degree. This model was later studied rigorously in [19, 62]. One of the main implications of this research is that it offers a possible explanation as to *why* complex networks display the features that they do: it is the result of the 'rich-gets-richer' postulate, that is, the simple hypothesis that more popular nodes are more likely to acquire more neighbours, and thus become even more popular over time. Indeed this so called "preferential attachment" model has been applied in other contexts, outside the generation of networks, to explain the emergence of power law distributions: first by Yule in the context of evolution in [79] and by Simon in [74], and Price in [27], who both observed the these distributions in a variety of contexts.

An example of the preferential attachment model, is that of an evolving tree, where one vertex arrives at a time and connects to a single existing vertex with probability proportional to its degree. This is a particular example of a *recursive tree model*, where an existing vertex is chosen according to an arbitrary probability distribution. Recursive trees generated in this manner have attracted widespread study, motivated by, for example, their applications to the evolution of languages [64], the analysis of algorithms [56] and the study of complex networks, see, for example, [78, Chapter 8.1]. Other applications include modelling the spread of epidemics, pyramid schemes and constructing family trees of ancient manuscripts (e.g. [33, page 14]). Whilst recursive tree models may display an inherent deficiency, as real world networks are hardly ever trees, they are often easier to analyse than more general evolving graph models. In addition, these models may be extended so that newly arriving vertices make $m \geqslant 1$ new connections. One way of doing this is to consider m copies of the new vertex each throwing one new connection to the existing network and then identifying them as one vertex, hence forming a *multigraph*. See Chapter 8 in [77] for a detailed description.

In the context of recursive trees, the preferential attachment model has been studied many times, under various guises: under the name *nonuniform recursive trees* by Szymański in [76], random *plane oriented recursive trees* in [55, 57], random *heap ordered recursive trees* [24] and *scale-free trees* [19, 75, 18]. Random ordered recursive trees, or plane-oriented recursive trees, are so named because the process stopped after n vertices arrive is distributed like a tree chosen at random from the set of rooted labelled trees on n vertices embedded in the plane where descendants of a node are ordered from left to right. This model has been extended to a number of interesting generalisations of the classical preferential attachment model, including the case that vertices are chosen according to a *super-linear* function of their degree in [68], or indeed any positive function of the degree [72], assuming a certain technical condition is satisfied. In [41], the latter model is generalised to arbitrary non-negative functions of the degree and is referred to as *generalised preferential attachment.*

1.2.3 Inhomogeneous Models

Models Exhibiting Condensation

Whilst the preferential attachment model is successful in reproducing the properties of complex networks, it is generally the earlier arriving vertices that are more likely to have higher degrees, since they have more time to acquire new neighbours, which in turn reinforces the growth of their degree. In other words, they have extra time to become 'rich' which allows them to acquire more 'wealth'. Indeed, a result of [30] shows that, from a certain time point onward, the vertex with maximal degree remains fixed in this model. Whilst this may be a realistic assumption in the context of the distribution of wealth in the world, in the context real world models it is often newly arriving nodes that quickly acquire a large number of links, for example, in the world wide web. Motivated by this, in [11], Bianconi and Barabási introduced their well-known *inhomogeneous* model, sometimes called *preferential attachment with multiplicative fitness.* There, vertices arrive one at a time, and, upon arrival, each vertex is equipped with a random weight sampled independently from a fixed distribution. At each time-step, the newly arriving vertex u connects to an existing vertex v with probability proportional to the product of the weight of v and its degree. Thus, the random weight may be interpreted as a measure of the intrinsic "attractiveness" of a vertex. Bianconi and

Barabási postulated the emergence of an interesting dichotomy in this model which they called *Bose-Einstein condensation*, motivated by similar phenomena in statistical physics.

This *condensation* phenomenon refers to the fact that under a certain critical condition on the weight distribution, a positive proportion of all the edges in tree accumulate around vertices of maximum weight. This dichotomy was first proved rigorously by Borgs et al. in [20] in the case that the weight distribution is supported on an interval, and absolutely continuous with respect to Lebesgue measure. However, they note that other classes of weight distribution are possible. They also showed that in this model, the degree distribution of vertices with a given weight follows an 'averaged' power law, with exponent depending on the weights of the vertex. A similar condensation phenomenon was observed in a variant of this model by Dereich in [28], and later, in a more general, robust setting, (in the sense that the results apply to wide variety of model specifications) in [31].

The *condensation* phenomenon observed by Bianconi and Barabási is closely related to the condensation phenomenon observed in other models. Indeed, it was first studied in a similar, yet simpler manner, in the context of evolution by Kingman in [51]. In [29], the authors studied condensation in models of *reinforced branching processes* that generalises a branching process associated with the Bianconi-Barabási model, showing that the condensation is *non-extensive*: whilst a positive proportion of edges in the family tree of the process accumulate around vertices of maximal weight, the maximal degree of the tree remains sublinear. Thus, this condensation phenomenon seems to be ubiquitous, and associated with other models outside the arena of complex networks.

Inhomogeneous models have also been studied in the context of models with *choice* in [38, 40], with the appearance of more fascinating condensation phenomena. In this model vertices are equipped with weights, at each time step r vertices are chosen with probability proportional to their degree, and out of these r vertices, a random vertex is chosen as the neighbour of the new-coming vertex. Here, the probability distribution by which the random vertex is chose, may depend on the weights of the vertices. In [38], the authors showed that, in the case that the maximal weight vertex is chosen, *extensive condensation* may occur, that is, under a critical condition on the weight distribution, a positive proportion of edges accumulate around the vertex of maximal degree. In addition, in [40], the authors showed that in certain cases, with random choice rules, the distribution of edges with endpoint having certain weight converges weakly to a *random measure* where *multiple condensation* can occur with positive probability, that is, positive proportions of edges accumulate around vertices of multiple weights. In addition, they showed that multiple condensation cannot occur when deterministic choice rules are used, and there exist phase transitions for condensation occurring with probability 0 or 1.

Other Inhomogeneous Recursive Models

There are a number of other interesting variations of inhomogeneous recursive tree models. In the *preferential attachment with additive fitness* introduced by Ergün and Rodgers in [34], newly arriving vertices now connect to existing vertices with probability proportional to the sum of their weight and degree, whilst in the *weighted recursive tree* introduced in [21], newly arriving vertices now connect to existing vertices with probability proportional to just their weight. In [73], Sénizergues showed that the preferential attachment with additive fitness with deterministic weights, is equal in distribution to a particular weighted random recursive tree with random weights, and used this to derive results related to a number of properties of both models, such as the *degree sequence* and the *height.* Moreover, recently in [69], Pain and Sénizergues derived sharper estimates for the heights of both models, in the case of random, identically distributed weights. Finally, in [54, 53], Lodewijks and Ortgiese uncovered an interesting dichotomy in the *maximal degrees* of these models, in a robust, evolving graph setting.

In [47], Jordan studies a model of preferential attachment where vertices belong to two types, and new vertices connect to one according to an additive fitness mechanism, and the other via a multiplicative fitness. Geometric models have also been considered in [48]: here, new vertices are equipped with a location in a metric space, and connect to existing vertices with probability proportional to the product of their degree, and a positive function of the distance between them. This positive function is known as an attractiveness function. In [48], the authors demonstrate a dichotomy, depending on the attractiveness function, between behaviour according to the model of Albert and Barabási, and a well known geometric model known as the *on line nearest neighbour model.*

1.2.4 Higher Dimensional Preferential Attachment Mechanisms

All the previously described models are 1-dimensional in the sense that newly arriving vertices are attached to single vertices. Our motivation is to consider attachment mechanisms in which newly arriving vertices join *groups* of vertices, where the attachment takes into account intrinsic features of a group of vertices, and thus encodes more complexity.

Simplicial complexes are a natural choice for incorporating this *higher dimensional* complexity at a local level. Furthermore, complex networks appearing in applications are typically *locally dense*: that is, although they form sparse graphs, the neighbourhood of a typical vertex is dense. This is usually measured by the *clustering coefficient.* The classic preferential attachment models do not satisfy this, as the graph that is formed is tree-like within a short distance from a randomly chosen vertex. However, this 'local density' arises naturally from the fact that simplicial complexes are *downwards closed.* Hence, a preferential attachment model which involves higher order interactions encapsulates these features naturally. Additionally, (random) simplicial complexes have already been used in applications such as topological data analysis (see, for example, [22]), and recent theories of quantum gravity (see, for example, [1]).

One model that realises higher order interactions is the Random Apollonian Network. It was first introduced in [4] and independently in [32] as a model for complex networks and was subsequently extended by Zhang et al. [80, 81]. Here, in dimension d, we begin with a d-simplex, all of whose $(d-1)$-dimensional faces are *active*. In each step, an active $(d-1)$-dimensional face is selected uniformly at random and d new $(d-1)$-faces are formed by the union of a new-coming vertex and each subset of the selected face of size $d-1$. Subsequently, the selected $(d-1)$-dimensional face is *deactivated*, so that the number of active $(d-1)$-faces in the complex increases by $d-1$ at each step. As each of the d new $(d-1)$-faces, together with the selected face σ form a d-face, we can interpret this step geometrically as a d-face being 'glued' onto the face σ, with the set of active faces being the boundary of the complex see Figure 4.1, in Section 1.3 below. Note that, when a node v enters the network, its degree is equal to d and the number of active faces containing it is equal to d. Moreover, every time an active face containing v is selected, the degree of v increases by one and the number of active faces containing v increases by $d-2$. Therefore, the number of active faces containing a given vertex v is $(d-2)\deg(v) - d(d-3)$. Thus, if $d > 2$ the number of active faces containing a vertex is proportional to its degree, and hence this model gives rise to a preferential attachment mechanism. In [52] and independently in [39], the authors determined that the degree distribution of this model for $d > 2$, gives rise to a power law with exponent $\tau = \frac{2d-3}{d-2} = 2 + \frac{1}{d-2}$.[1] For $d = 3$ the same model has been studied under the name *random stack-triangulations* by Albenque and Marckert in [2], where they proved that the sequence of complexes with *graph distance metric* rescaled by $\sqrt{n}$ considered as a *compact metric space* converges in the *Gromov-Hausdorff topology* to the *continuum random tree* of Aldous [3].

[1] Note that often in the literature surrounding Apollonian networks, rather than using the dimension of the initial simplex, authors use the number of vertices in an 'active' face as the parameter of the model. Thus the Apollonian network with parameter d is the same as the Apollonian network in dimension $d-1$.

Inhomogeneous Higher Dimensional Evolving Models

In the Apollonian network the choice among the active $(d-1)$-faces is uniform. In particular, there is no preferential attachment mechanism directly associated with the evolution of the vertices. This motivates the study of mechanisms in which these high-dimensional substructures are *inhomogeneous* and have some intrinsic fitness which is a function of the weights of their members.

Specific implementations of this idea were introduced by Bianconi, Rahmede, and other co-authors motivated by applications in physics ([12, 15, 25, 13, 14, 26]). For example, random triangulations have been considered in the context of quantum gravity [1]. The model of *Complex Quantum Network Manifolds (CQNMs)* described in [12] in dimension $d > 1$ can be viewed as a generalisation of the Random Apollonian Network, where vertices are equipped with independent, identically distributed (i.i.d.) weights, called *energies* in this context, and each $(d-1)$-face σ of the evolving d-dimensional simplicial complex has energy ϵ_σ given by the sum of the energies of its vertices. The simplicial complex evolves in the same way as the Random Apollonian network, with the only difference being that at each time-step, a new vertex selects an active $(d-1)$-face σ with probability proportional to $e^{-\beta\epsilon_\sigma}$ instead of uniformly at random; where $\beta \geqslant 0$ is a fixed constant, usually interpreted as the "inverse temperature". In [12], the authors argue that when $d = 2$ the underlying graph has degree distribution with *exponential tail* whilst, when $d \geqslant 3$ the degree distribution follows a power law with exponent that depends on d, β and the distribution of the weights. In this thesis, we verify a rigorous version of this result when the energies are bounded (see Section 4.2.3).

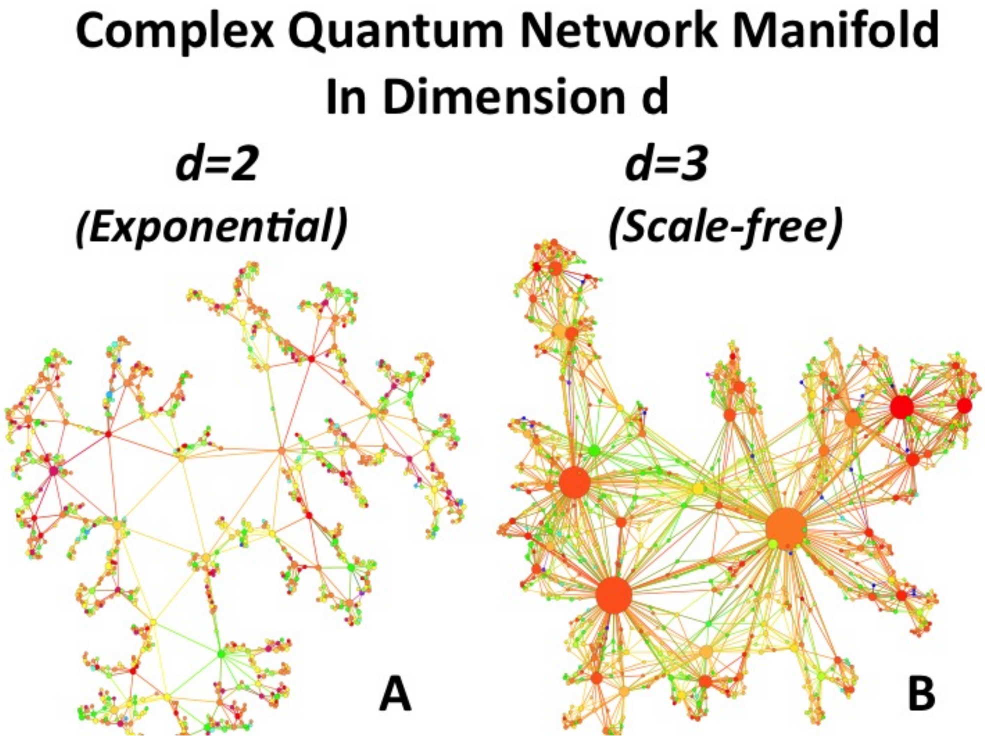


Figure 1.6: This illustration shows the different behaviour of Complex Quantum Network Manifolds in dimension 2 vs dimension 3, observed by the authors of [12]. In dimension 3, we obtain a model with scale-free degree distributions, reminiscent of complex networks in real world applications, whilst in dimension 2 we obtain a model with degree distributions having exponential tails. Image sourced from [12].

In [13], Bianconi and Rahmede introduce a more general model called the *network geometry with flavour (NGFs)*. The network geometry with flavour, in dimension d and *flavour* $s \in \{-1, 0, 1\}$ proceeds as follows. As before, vertices are equipped with i.i.d. *energies* and each $(d-1)$-face σ of the evolving d-dimensional simplicial complex has energy ϵ_σ which is equal to the sum of the energies of its vertices. At each time-step, a new vertex selects a $(d-1)$-face σ with probability proportional to $e^{-\beta\epsilon_\sigma}(1 + s\deg_d(\sigma) - s)$, where $\beta \geqslant 0$ is a fixed constant. In the case $s = -1$, Bianconi and Rahmede [12] argue that when $d = 2$

the underlying skeleton graph has degree distribution with exponential tail, whilst when $d \geqslant 3$ the degree distribution obeys a power law, with an exponent that depends on d as well as on β and the distribution of the weights. Moreover, in [15], Bianconi, Rahmede and Wu argue that for $d = 2$, if $s = -1$ the underlying skeleton graph has degree distribution with exponential tail, whilst if $s = 0$, the underlying skeleton graph has power law tails. We will prove weaker versions of both these results rigorously in this thesis, in the sense that the degree distribution has a tail bounded from above and below by a power law. See Section 4.2.3 for more details.

1.3 Our Models: Evolving Inhomogeneous Random Structures

In this thesis, we study evolving, inhomogeneous models that are closely related to many of the models studied in Section 1.2. In this section we provide a formal description of each of these models, and indicate the chapters associated with each model. We first provide a brief overview of the notation used in this thesis. Although the notation we introduce is closely related across each of the models, some notation varies depending on the context; however, this should be clear based on which model the notation relates. Subsequently we provide an overview of the main types of results we will prove in this thesis in Section 1.4.

1.3.1 Notation Applied Throughout the Thesis

In this thesis we generally set $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$ and $\mathbb{R}_+ := [0, \infty)$. In addition, for $s \in \mathbb{N}$, we denote by $[s]$ the set $\{1, \dots, s\}$. In addition, for $\ell \in \mathbb{N}$, we denote by $[s]^{\ell}$ the ℓ-fold Cartesian product $[s] \times \cdots \times [s]$. Given a set $S \subset \mathcal{S}$, we denote by S^c the complement of this set, and, if $\mathcal{S}$ has a topology made clear from context, we denote by $\overline{S}$ the topological closure of S. Finally, given a set S, we denote by $\mathbf{1}_S(x)$ the indicator function associated with this set, so that $\mathbf{1}_S(x) = 1$ if $x \in S$ and 0 otherwise. Moreover, if $\mathbf{1}_S(x)$ is a random variable on a probability space $(\Omega, \mathbb{F}, \mathbb{P})$, we omit the dependence on $x \in \Omega$, and simply write $\mathbf{1}_S$.

Weights, Weight Distribution, Support, Essential Supremum

In this thesis we will consider *inhomogeneous models* where vertices have weights assigned to them. In general, these weights take values in $\mathbb{R}_+$ and are sampled from a fixed probability measure μ. We generally denote by W a generic random variable sampled from μ.

In general, we assume that the space $\mathbb{R}_+$ is equipped with its Borel sigma algebra $\mathscr{B}$. Often it will be the case that we need to deal with weights that take bounded values. We denote by $\mathrm{Supp}(\mu)$ the *support* of the measure μ, that is the set of all points x in $\mathbb{R}_+$, for whom every open neighbourhood O_x has positive measure

$$\mathrm{Supp}(\mu) := \{x \in \mathbb{R}_+ : \mu(O_x) > 0, \text{ for all open sets } O_x \text{ such that } x \in O_x\}.$$

In certain cases, we will need to assume that the support is bounded, so that $\mathrm{Supp}(\mu) \subseteq [0, w^*]$, where $w^* := \sup(\mathrm{Supp}(\mu))$. Moreover, for a measurable function $g : \mathbb{R}_+ \to \mathbb{R}_+$ we define $\operatorname{ess\,sup}(g)$ such that

$$\operatorname{ess\,sup}(g) := \inf\{a \in \mathbb{R}_+ : \mu(\{x : g(x) > a\}) = 0\}.$$

1.3.2 Generalised Recursive Trees with Fitness

Our first model, which we study in Chapter 2 is a unified model that encompasses most of the models described in Section 1.2.2 and Section 1.2.3 above.

In order to define the model, we first require a probability measure μ supported on $\mathbb{R}_+$ and a *fitness function*, which is a measurable function $f : \mathbb{N}_0 \times \mathbb{R}_+ \to \mathbb{R}_+$. We consider evolving sequences of *weighted oriented trees* $\mathcal{T} := (\mathcal{T}_n)_{n\in\mathbb{N}_0}$; these are trees with *directed edges*, where vertices have real valued weights assigned to them. The model also has an additional parameter $\ell \in \mathbb{N}$. We start with an initial tree $\mathcal{T}_0$ consisting of a single vertex 0 with weight W_0 sampled from μ. To ensure that the evolution of the model is well-defined, we assume $f(0, W_0) > 0$ almost surely. Then, we define $\mathcal{T}_{n+1}$ recursively as follows:

(i) Sample a vertex j from $\mathcal{T}_n$ with probability

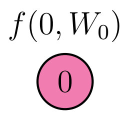
$$\frac{f(\deg^+(j, \mathcal{T}_n)/\ell, W_j)}{\mathcal{Z}_n},$$

where $\deg^+(j, \mathcal{T}_n)$ denotes the out-degree of the vertex j in the oriented tree $\mathcal{T}_n$ and $\mathcal{Z}_n := \sum_{j=0}^{\ell n} f(\deg^+(j, \mathcal{T}_n)/\ell, W_j)$ is the *partition function* associated with the process.

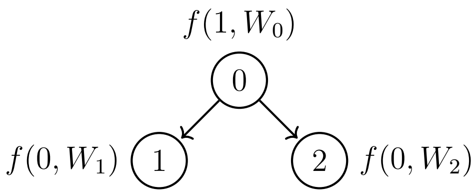
(ii) Introduce ℓ new vertices $n + 1, n + 2, \ldots, n + \ell$ with weights $W_{n+1}, W_{n+2}, \ldots, W_{n+\ell}$ sampled independently from μ and the directed edges $(j, n+1), (j, n+2), \ldots, (j, n+\ell)$ oriented towards the newly arriving vertices. We say that j is the *parent* of the new-coming vertices, and that the new-coming vertices are its *offspring*.

Note that, since ℓ new vertices are connected to a parent at each time-step, for any vertex i in the tree, ℓ divides the out-degree of i. Moreover, the evolution of the out-degree of vertex i with weight W_i is determined by the values $(f(j, W_i))_{j\in\mathbb{N}_0}$. In general, when the distribution μ, fitness function f and ℓ are specified, we refer to this model as a (μ, f, ℓ)-*recursive tree with independent fitnesses*, often abbreviated as a "(μ, f, ℓ) - RIF tree" for brevity. Here 'independent fitnesses' refers to the fact that the fitness associated with a given vertex does not depend on the weights of its neighbours, in contrast to, for example, the other models of *preferential attachment with neighbourhood influence* and *dynamical simplicial complexes* we will study. The following figure illustrates a possible evolution of this model over the first three steps.

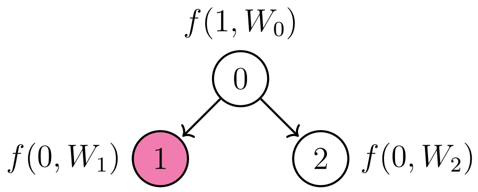
A Sample Evolution of the (μ, f, ℓ) - RIF tree with $\ell = 2$



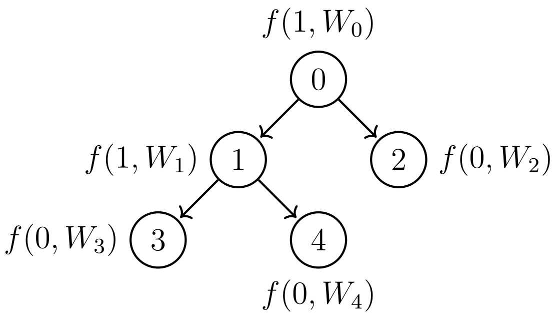
(a): At time 0, there is only one vertex with weight W_0 and fitness $f(0, W_0) > 0$, so this vertex is selected in the first step.



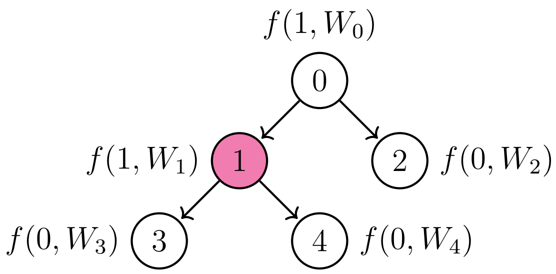
(b) This vertex connects to two new neighbours 1 and 2 with weights W_1, and W_2 and fitnesses $f(0, W_1)$ and $f(0, W_2)$. The fitness associated with 0 is now updated to $f(1, W_0)$.



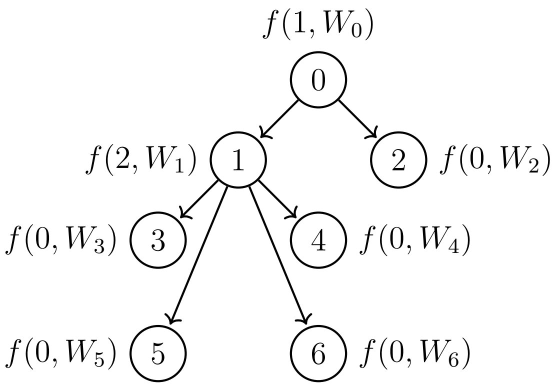
(c) A vertex is selected with probability proportional to its fitness function, and note that it now may be the case that $f(1, W_0) = 0$. In this case, vertex 1 is selected.



(d) Vertex 1 produces offspring 3 and 4, and its fitness is updated accordingly.



(e) Again, a vertex is sampled with probability proportional to its fitness. Here, vertex 1 is selected.



(f) Vertex 1 produces offspring 5 and 6, and its fitness is adjusted accordingly.

Figure 1.7: A sample evolution of the first three steps of the (μ, f, ℓ) - RIF tree when $\ell = 2$. Steps (b), (d) and (f) illustrate the trees $\mathcal{T}_1, \mathcal{T}_2$ and $\mathcal{T}_3$ respectively.

1.3.3 Preferential Attachment Trees with Neighbourhood Influence

A particular case of the (μ, f, ℓ) - RIF tree introduced in Section 1.3.2 is the case that f is affine, of the form $g(W)i + h(W)$, where g and h are measurable functions. As we will show in Section 2.3 in Chapter 2, this particular case of the model displays many interesting properties, including a *condensation phenomenon*. We call this *generalised preferential attachment with fitness*, or GPAF-tree.

This motivates us to consider a 'higher dimensional' form of this model, which we call preferential attachment tree with neighbourhood influence, or PANI-tree, where the attachment mechanism considers not only the weight of a given vertex, but also the weights of its *neighbours*. For brevity, in this model we only consider the case where only a single vertex arrives at each time-step ; in the context of the (μ, f, ℓ) - RIF tree this corresponds to the case that $\ell = 1$.

As in Section 1.3.2, we consider a model of *weighted directed trees* $(\mathcal{T}_n)_{n \in \mathbb{N}_0}$. Let $\mathbb{T}$ denote the set of all such weighted trees, and given a tree $\mathcal{T} \in \mathbb{T}$ and a vertex $j \in \mathcal{T}$, (abusing the notation for the out-neighbourhood slightly) let $\mathcal{N}^+(j, \mathcal{T})$ be the weighted tree consisting of j and all of its *out-neighbours*. In order to define the model, we will require a probability measure μ, which is supported on a subset of an interval $[0, w^*]$, for some $w^* > 0$ and a *fitness function* $f : \mathbb{T} \to \mathbb{R}_+$. One may interpret this as an analogue of the fitness function in Section 1.3.2 that may take into account the weights of neighbours of a given vertex. In the model we consider, we start with an initial tree $\mathcal{T}_0$ consisting of a single vertex with random weight W_0 sampled from μ. Then, given $\mathcal{T}_i$, the model proceeds recursively as follows:

(i) Sample a vertex j from $\mathcal{T}_i$ with probability $\frac{f(\mathcal{N}^+(j,\mathcal{T}_i))}{\mathcal{Z}_i}$, where $\mathcal{Z}_i := \sum_{k=0}^{i} f(\mathcal{N}^+(k, \mathcal{T}_i))$ is the *partition function* associated with the process.

(ii) Form $\mathcal{T}_{i+1}$ by adding the edge $(j, i+1)$, and assigning vertex $i+1$ weight W_{i+1} sampled independently from μ.

In this thesis, with regards to this model, we define f so that

$$f(\mathcal{N}^+(v,T)) = h(W_v) + \sum_{(v,u)\in T} g(W_v, W_u), \tag{1.1}$$

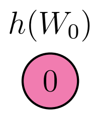
where $h : [0, w^*] \to [0, \infty)$ and $g : [0, w^*] \times [0, w^*] \to [0, \infty)$ are bounded and measurable. To ensure that the evolution of the model is well-defined, in all of our results we condition on W_0 satisfying $h(W_0) > 0$, which we assume is an event that has positive probability.

Remark 1.3.1. *The form of the fitness function in* (1.1) *is sufficiently general to encompass some existing models. In the case where g and h are a single constant, we obtain the classic preferential attachment tree of Albert and Barabási. The case $g(x, y) = h(x) = x$ is the Bianconi-Barabási model, whilst the case $g(x, y) \equiv 1, h(x) = x$ is the preferential attachment tree with additive fitness. Finally, the case $g(x, y) = g'(x)$, for some bounded measurable function of a single variable is a particular case of the* (μ, f, ℓ) - RIF *tree we call the* GPAF-*tree, which is studied in Section 2.3 of Chapter 2.*

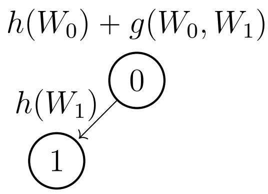
Remark 1.3.2. *As in the* (μ, f, ℓ) - RIF *tree, we may also analyse this model when ℓ vertices connect to the selected vertex during each time-step. However, for brevity, we restrict our analysis to the case that $\ell = 1$.*

We illustrate a possible evolution of this model below.

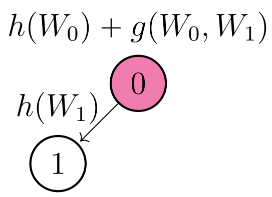
A Sample Evolution of the PANI-Tree



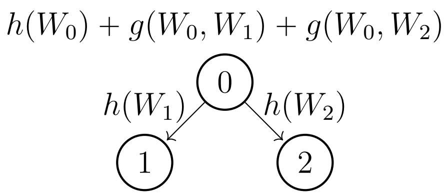
(a): At time 0, there is only one vertex with weight W_0 and fitness $h(W_0) > 0$, so this vertex is selected in the first step.



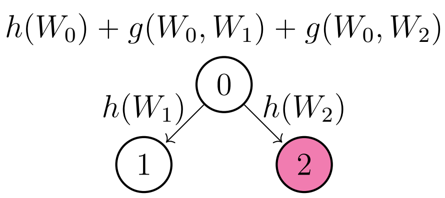
(b) This vertex connects to a new neighbours 1 with weight W_1 and fitness $h(W_1)$. The fitness associated with 0 is now increased by $g(W_0, W_1)$; note that, unlike the (μ, f, ℓ)-RIF tree illustrated in Figure 1.7, this change also depends on W_1.



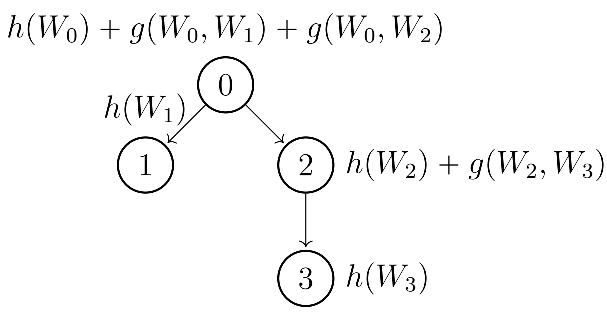
(c) A vertex is selected with probability proportional to its fitness function; note that either vertex may be selected with positive probability. In this case, vertex 0 is selected.



(d) Vertex 0 connects to the new vertex 2, and its fitness is updated accordingly.



(e) Again, a vertex is sampled with probability proportional to its fitness. Here, vertex 2 is selected.



(f) Vertex 2 connects to 3, and its fitness is adjusted accordingly.

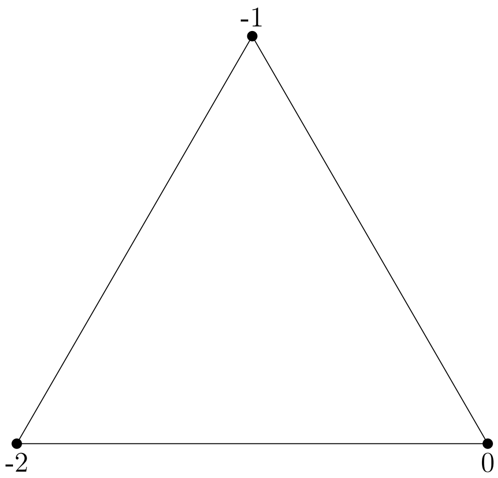
Figure 1.8: A sample evolution of the first three steps of the preferential attachment model with local dependencies. Steps (b), (d) and (f) illustrate the trees $\mathcal{T}_1, \mathcal{T}_2$ and $\mathcal{T}_3$ respectively.

1.3.4 Dynamical Models for Random Simplicial Complexes

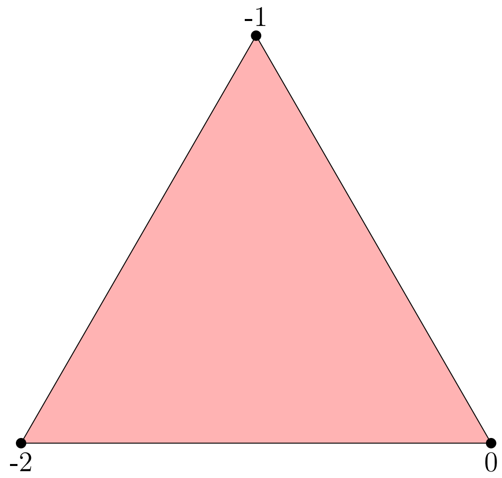
The final model we consider in this thesis involves even more dependence between the evolution of vertices and their neighbours: we consider a sequences of simplicial complexes $(\mathcal{K}_n)_{n\geqslant 0}$ of fixed parameter $d \geqslant 0$. In this case, again we assume that the weight distribution μ is supported on a subset of an interval $[0, w^*]$, and, as an additional parameter we have a *fitness function*, which in this context is a positive, symmetric function $f : [0, w^*]^d \to \mathbb{R}_+$. For all $n \geqslant 0$, $\mathcal{K}_{n+1}$ is obtained by adding one vertex labelled $n+1$ to $\mathcal{K}_n$ and assigning that vertex a random weight sampled independently according to μ.

At each time-step n, a $(d-1)$-face σ is sampled from the complex $\mathcal{K}_n$ with probability proportional to its fitness $f(\sigma)$, which is the image by f of the vector of the weights of the vertices that belong to σ (as the function f is symmetric, this image does not depend on the order of the weights in the vector). Then a new vertex $n+1$ arrives, with an associated independent weight W_{n+1}, and *subdivides* the selected face, as illustrated in Figure 1.9 below. In Model **A**, the selected face σ remains in the complex, whilst in Model **B** the selected face is removed from the complex.

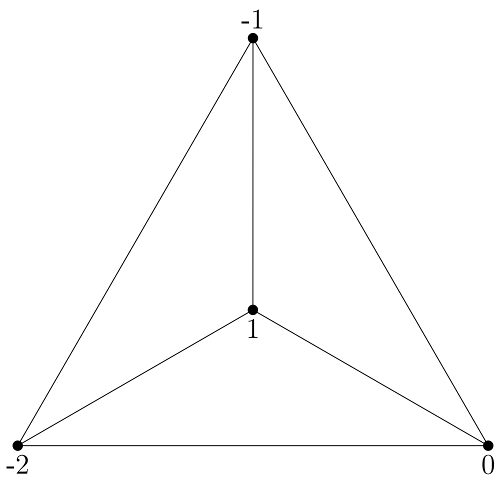
A Sample Evolution of the Dynamical Simplex Model in Dimension 3



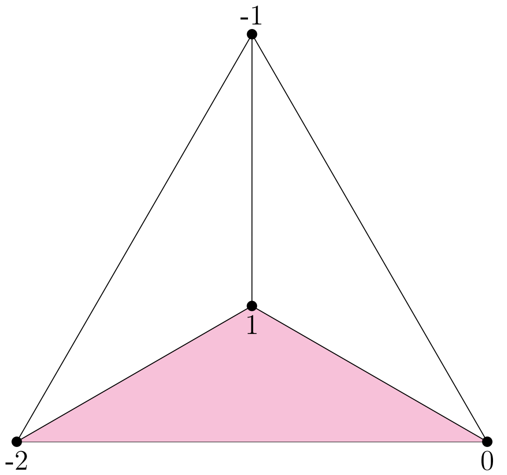
(a): At time 0 we begin with an arbitrary $(d-1)$-dimensional simplicial complex with vertices labelled by non-positive integers. In this case, we have a 2-simplex.



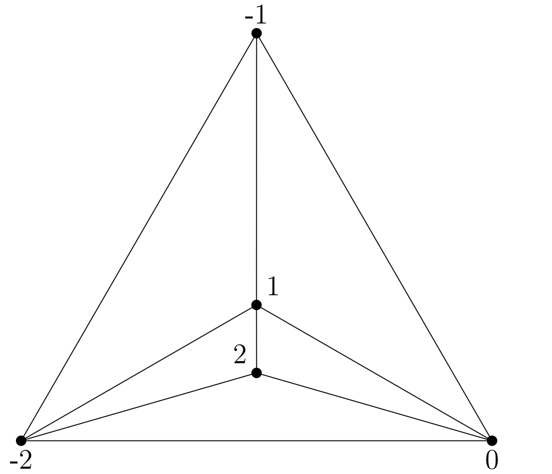
(b) A $(d-1)$-face σ is sampled with probability proportional to its fitness $f(\sigma)$, a positive function of the weights of the vertices in σ. In this case, there is only one 2-face, $\{-2, -1, 0\}$, which must be selected.



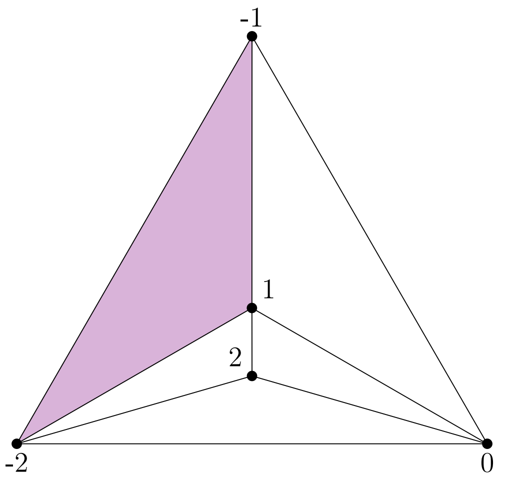
(c) A new coming vertex 1 arrives, and for each subset σ' of 2 of the selected face $\sigma = \{-2, -1, 0\}$, we add the face $\sigma' \cup \{1\}$. In Model **B**, the selected face is also removed from the complex. We may interpret this geometrically as a 3-dimensional tetrahedron being 'glued' onto the 2-face; thus in Model **B** we may associate the set of faces in the complex with the boundary of a 3-dimensional simplex.



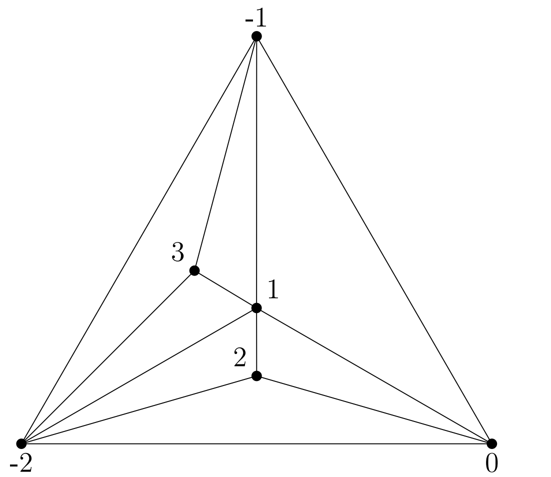
(d) Now, the face $\{-2, 0, 1\}$ is selected.



(e) A new-coming vertex 2 arrives, and again *subdivides* the selected face.



(f) Next, the fact $\{-2, -1, 1\}$ is selected.



(g) This face is subdivided by the vertex 3.

Figure 1.9: A sample evolution of the dynamical simplex model with parameter 3. This particular evolution may be an instance of either Model **A** or Model **B**.

1.4 Important Quantities of Interest in this Thesis

Despite the variations in each of the models we have described, we will see in this thesis that their recursive nature means that each of these models are amenable to similar techniques.

In general in this thesis we will be interested in two main quantities: the distribution of the proportion of nodes with a given *degree* and weight and the distribution of the proportion of *edges* with endpoint having a given weight. As we will see, the prior quantity seem to have a universal limiting behaviour, described by $p_k^\lambda(\cdot)$ defined in (1.4), below.

1.4.1 Degree Distributions

The first main quantity we will be concerned with in this thesis relates to *degree distributions.* In general in this thesis, we denote by $N_k(n)$ the number of vertices in the respective model at time n that have been selected k times in the evolution of this model, and $N_k(n, \cdot)$ the *empirical measure* corresponding to the number in the respective model at time n that have been selected k times with a given weight. We will also use the notation $N_{\geqslant k}(n)$ and $N_{\geqslant k}(n, \cdot)$ to denote the number of vertices selected at least k times, and the number of vertices with a given weight selected at least k times, respectively.

More precisely,

1. With regards to the (μ, f, ℓ) - RIF tree, given a Borel set $B \subseteq \mathbb{R}_+$, the quantity $N_k(n, B)$ denotes the number of vertices v in the tree $\mathcal{T}_n$ with out-degree $k\ell$ and weight $W_v \in B$, that is,
$$N_k(n, B) := \sum_{v \in \mathcal{T}_n : \deg^+(v, \mathcal{T}_n) = k\ell} \mathbf{1}_B(W_v). \tag{1.2}$$
Also, $N_k(n) := N_k(n, \mathbb{R}_+)$. With regards to the preferential attachment model with neighbourhood influence, or PANI-tree, $N_k(n, B)$ is defined identically, however, we have $\ell = 1$.

2. Similarly, the quantity $N_{\geqslant k}(n, B)$ is defined such that

$$N_{\geqslant k}(n, B) := \sum_{v \in \mathcal{T}_n : \deg^+(v, \mathcal{T}_n) \geqslant k\ell} \mathbf{1}_B(W_v),$$

and with $\ell = 1$ in the PANI-tree.

3. In the dynamical simplices model, up to a constant factor depending on the initial complex $\mathcal{K}_0$, the quantity $N_k(n)$ denotes the number of vertices with degree (or 0-degree) $k + d$. For brevity, with regards to this model we will generally state and prove results for $N_k(n)$, although similar analysis may be performed for quantities analogous to $N_k(n, \cdot)$.

Now, suppose V_n denotes the vertex set in each of the models, so that in the (μ, f, ℓ) - RIF tree, $|V_n|$ scales like ℓn, whilst in the other models, $|V_n|$ scales like n. We will then be interested in the limiting behaviour of the quantity $N_k(n, B)$ when re-scaled by the size of the network, $|V_n|$, in each of the models. It is reasonable to expect that the almost sure limit of $\frac{N_k(n,B)}{|V_n|}$ behaves like its expected value

$$\sum_{i=0}^{|V_n|} \mathbb{P}\left(W_i \in B, \{\text{vertex } i \text{ has been selected exactly } k \text{ times}\}\right) / |V_n|. \tag{1.3}$$

Suppose that the probability of selecting vertex i, with weight W_i, once this vertex has already been selected j times is approximately $(C_j(W_i))_{j \geqslant 0}$. Also, if we informally, suppose that the partition function $\mathcal{Z}_n$ behaves like λn, for some $\lambda > 0$, the probability of a vertex i, with weight W_i, arriving at i_0 and receiving out-neighbours at times $i_1, \dots, i_k$, is approximately

$$\prod_{j=1}^{i_1 - i_0 - 1} \left(1 - \frac{C_0(W_i)}{\lambda(i_0 + j)}\right) \frac{C_0(W_i)}{\lambda i_1} \cdot \prod_{j=1}^{i_2 - i_1 - 1} \left(1 - \frac{C_1(W_i)}{\lambda(i_1 + j)}\right) \frac{C_1(W_i)}{\lambda i_2} \cdots$$
$$\cdots \prod_{j=1}^{i_k - i_{k-1} - 1} \left(1 - \frac{C_{k-1}(W_i)}{\lambda(i_{k-1} + j)}\right) \frac{C_{k-1}(W_i)}{\lambda i_k} \cdot \prod_{j=1}^{n - i_k} \left(1 - \frac{C_k(W_i)}{\lambda(i_k + j)}\right).$$

Now, if we can approximate the expected value in (1.3) by considering summands $i > \eta n$, where η is a 'small' constant, we may write the products in the previous display as ratios of Gamma functions, which may then be approximated using Stirling's approximation. Then, for each i, taking the sum over possible choices $(i_1, \ldots, i_k)$, by applying suitable summation arguments, i.e., Corollary 2.4.6 in Section 2.4.2, Chapter 2, we obtain

$$\frac{\lambda}{C_k(W_i) + \lambda} \prod_{j=0}^{k-1} \frac{C_j(W_i)}{C_j(W_i) + \lambda}.$$

Taking expectations over $W_i \in B$, it is therefore reasonable to expect that the limit of $\frac{N_k(n,B)}{|V_n|}$ belongs to the family

$$p_k^{\lambda}(B) := \mathbb{E}\left[\frac{\lambda}{C_k(W) + \lambda} \prod_{j=0}^{k-1} \frac{C_j(W)}{C_j(W) + \lambda} \mathbf{1}_B(W)\right], \tag{1.4}$$

for $\lambda > 0$. The expectation on the right hand side of (1.4) is with regards to the path of a suitable random *companion process* $(C_j(W_i))_{j \geqslant 0}$, depending on the weight W_i. The precise form of the companion process depends on the model we consider. In particular, this companion process is such that

1. In the (μ, f, ℓ)- RIF tree the value $C_j(W_i)$ is W_i-measurable, and given by $f(j, W_i)$.

2. In the PANI-tree, $C_0(W_i) = h(W_i)$, and, given $C_j(W_i)$, $C_{j+1}(W_i) = g(W_i, W') + C_j(W_i)$, where W' is sampled independently from μ. Thus, $C_j(W_i) - h(W_i) = \sum_{\ell=1}^{\ell} g(W_i, W'_\ell)$, where each W'_ℓ is independently sampled from μ. In particular, $C_j(W_i) - h(W_i)$ is given by a sum of random variables, which are conditionally independent and identically distributed given W_i.

3. In the dynamical simplicial complex model, the values of $C_j(W_i)$ depend on the fitnesses in the $(d-1)$-neighbourhood of i. Thus, $C_j(W_i)$ is a process that depends on the 'typical' evolution of the $(d-1)$-neighbourhood of a vertex arriving sufficiently 'late'.

In this thesis, we will prove various forms of the limiting degree distribution, showing that the family $(p_k^\lambda(\cdot))_{k\in\mathbb{N}_0}$ is universal across all models. We also make the intuition outlined before (1.4) rigorous in Section 2.4 in Chapter 2 and Chapter 4. The assumption that the partition function $\mathcal{Z}_n$ behaves like λn, for some $\lambda > 0$, is made rigorous by requiring that

$$\frac{\mathcal{Z}_n}{n} \to \lambda \quad \text{almost surely,} \tag{1.5}$$

and applying Egorov's theorem. The convergence in (1.5) is assumed directly in Section 2.4 in Chapter 2, while proved in various forms in Section 4.3 in Chapter 4.

1.4.2 Edge Distributions and Condensation

With regards to the evolving tree models we study in this thesis, i.e, the (μ, f, ℓ)-RIF tree and the PANI-tree, we will also be interested in another quantity: the distribution of the proportion of edges with endpoint having a given weight.

1. In both the (μ, f, ℓ)-RIF tree and the PANI-tree, given a Borel set $B \subseteq \mathbb{R}_+$, the quantity $\Xi(n, B)$ will denote the number of directed edges (v, v') in the respective tree model $\mathcal{T}_n$ such that $W_v \in B$, that is,

$$\Xi(n, B) := \sum_{(v,v')\in\mathcal{T}_n} \mathbf{1}_B(W_v). \tag{1.6}$$

2. With regards to the PANI-tree, we will also study a higher dimensional analogue of this quantity: given a Borel set $A \subseteq \mathbb{R}_+^2$, the quantity $\Xi^{(2)}(n, A)$ will denote the number of edges (v, v') in the tree $\mathcal{T}_n$ such that $(W_v, W_{v'}) \in A$, that is,

$$\Xi^{(2)}(n, A) := \sum_{(v,v')\in\mathcal{T}_n} \mathbf{1}_A(W_v, W_{v'}).$$

Our emphasis will be on results related to the quantity $\Xi(n, B)$. Suppose ℓ corresponds to the parameter ℓ when referring to the (μ, f, ℓ)-RIF tree, and 1 when referring to the PANI-tree. Then, note that for every $n \in \mathbb{N}_0$, by computing the number of directed edges (v, v') in $\mathcal{T}_n$ with $W_v \in B$ in two different ways, we have

$$\Xi(n, B) = \sum_{k=0}^{n} \ell k N_k(n, B). \tag{1.7}$$

When we normalise by the number of vertices in the tree, $|V_n| = \ell n$, if, for $k \in \mathbb{N}_0$ the limit of $\frac{N_k(n,B)}{|V_n|}$ is $p_k^{\alpha}(B)$, as described in (1.4), by an application of Fatou's lemma we get

$$\liminf_{n\to\infty} \frac{\Xi(n, B)}{\ell n} \geqslant \sum_{k=0}^{\infty} \ell k p_k^{\alpha}(B), \tag{1.8}$$

which motivates the definition of the following family:

$$m(\lambda, B) := \sum_{k=0}^{\infty} \ell k p_k^{\lambda}(B) = \ell \cdot \mathbb{E}\left[\sum_{n=1}^{\infty} \prod_{j=0}^{n-1} \frac{C_j(W)}{C_j(W) + \lambda} \mathbf{1}_B(W)\right]. \tag{1.9}$$

Now, if the limit exists, since we add ℓ directed edges at each time-step, the measures $\Xi(n, \cdot)/\ell n$ are probability measures. However, if $m(\lambda, \cdot)$ is not a probability distribution (applying a similar argument to the proof of Theorem 2.2.2 in Section 2.2 of Chapter 2) we can show that there exists a measurable set B such that

$$\limsup_{n\to\infty} \frac{\Xi(n, B)}{\ell n} > m(\lambda, B).$$

In this case, the inequality in (1.8) is strict, so that, after normalising by ℓn, the operations of taking limits in k and in n in (1.7) do not commute. Thus, the set B has acquired additional "mass" in the limit, and this phenomenon is known as *condensation*. In Section 2.3.2 of Chapter 2 we derive an example of this in the GPAF-tree, i.e., the (μ, f, ℓ) - RIF tree in the case that $f(i, W) = g(W)i + h(W)$ for measurable functions g and h. In this case, we assume that g and h are bounded and non-decreasing. As the PANI-tree generalises this model further, we undertake a more refined analysis of the condensation phenomenon in Chapter 3 in Section 3.3.

Example: the (μ, f, ℓ) - RIF tree when $\ell = 2$

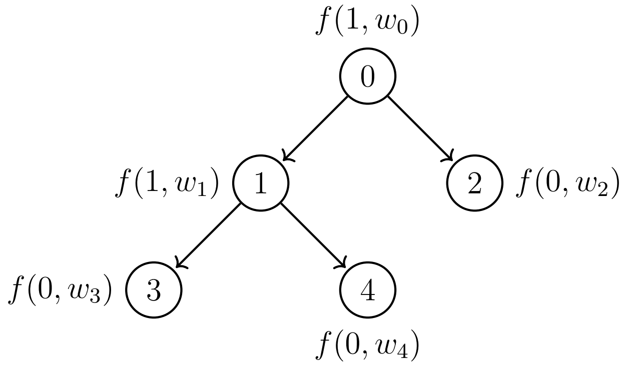


Figure 1.10: In the above instance of $\mathcal{T}_4$ in the (μ, f, ℓ) - RIF tree, $N_1(4, \cdot) = \delta_{w_0}(\cdot) + \delta_{w_1}(\cdot)$ and $\Xi(4, \cdot) = 2\,(\delta_{w_0}(\cdot) + \delta_{w_1}(\cdot))$.

1.5 Overview of Thesis

In this thesis we analyse the quantities outlined in Section 1.4, in each of the models described in Section 1.3. In particular,

- In Chapter 2 we analyse the (μ, f, ℓ) - RIF tree.
- In Chapter 3 we analyse the PANI-tree. The results of this chapter may be read independently of Chapter 2, however, are closely related to the results of Section 2.3.2 of Chapter 2, and as a result, we encourage the reader to at least review this section.
- In Chapter 4 we analyse the dynamical simplices model. However, the results of this chapter rely on certain results proved and stated in Chapter 2. In particular, the analysis in Section 4.4 is closely related to the analysis presented in Section 2.4 of Chapter 2, and applies the summation arguments proved in Section 2.4.2. In addition, the analysis in Section 4.3 of Chapter 4 applies results related to *Pólya urns*, and these

stochastic processes play a crucial role in the analysis of Chapter 3, in particular, in Section 3.2. We thus encourage the reader to read Chapter 4 after reading Chapter 2 and Chapter 3. Moreover, as previously mentioned, the interested reader may wish to skip some of the more technical proofs in this chapter upon first reading.

Note that each of the chapters rely closely on the specification of the model in Section 1.3 and the definitions of the quantities outlined in Section 1.4. The information in Section 1.2 may also be useful, especially the definitions in Section 1.2.1 - in particular with regards to the dynamical simplicial complexes model in Chapter 4.

Chapter Two

Generalised Recursive Trees with Fitness

2.1 Introduction

In this chapter, we consider the model of the generalised recursive tree with fitness described in Section 1.3.2 of Chapter 1, and prove limiting results regarding the degree distributions and edge distributions in relation to this model when re-scaled by the number of edges in the model, ℓn. Here we recall that these quantities, and their expected limiting behaviour was described in Section 1.4 of Chapter 1.

In relation to the (μ, f, ℓ) - RIF tree, the candidates $p_k^\lambda(\cdot)$ and $m(\lambda, \cdot)$, described in (1.4) and (1.9) of Chapter 1 have a specific form; in particular, we have

$$p_k^\lambda(B) = \mathbb{E}\left[\frac{\lambda}{f(k,W)+\lambda}\prod_{i=0}^{k-1}\frac{f(i,W)}{f(i,W)+\lambda}\mathbf{1}_B(W)\right], \tag{2.1}$$

and

$$m(\lambda, B) = \sum_{k=0}^{\infty} \ell k p_k^\lambda(B) = \ell \cdot \mathbb{E}\left[\sum_{n=1}^{\infty}\prod_{i=0}^{n-1}\frac{f(i,W)}{f(i,W)+\lambda}\mathbf{1}_B(W)\right]. \tag{2.2}$$

Since we only study the (μ, f, ℓ) - RIF tree in this chapter, in this chapter we may regard (2.1) and (2.2) as the *definitions* of the quantities $p_k^\lambda(\cdot)$ and $m(\lambda, \cdot)$ respectively. Moreover, using the heuristic outlined in Section 1.4.1 of Chapter 1, we expect the limiting behaviour of the re-scaled degree distribution $\frac{N_k(n,\cdot)}{\ell n}$ to belong to the family (2.1), for a suitable choice $\lambda = \alpha > 0$. In addition, if no *condensation* occurs, i.e., if $m(\alpha, \cdot)$ is a probability distribution, we expect the limit of $\frac{\Xi(n,\cdot)}{\ell n}$ to be $m(\alpha, \cdot)$.

2.1.1 Open Problems

We conjecture that, in general, the parameter α makes $m(\lambda, \cdot)$ 'as close as possible' to a probability distribution, so that

$$\alpha = \begin{cases} \inf\{\lambda > 0 : m(\lambda, \mathbb{R}_+) \leqslant 1\} & \text{if } m(\lambda, \mathbb{R}_+) < \infty \text{ for some } \lambda > 0 \\ \infty & \text{otherwise.} \end{cases} \tag{2.3}$$

Conjecture 2.1.1. *Let $\mathcal{T}$ be a (μ, f, ℓ) - RIF tree, with α as defined in* (2.3)*. Then, for each $k \in \mathbb{N}_0$ and measurable set B, almost surely, we have*

$$\frac{N_k(n,B)}{\ell n} \xrightarrow{n\to\infty} \begin{cases} p_k^{\alpha}(B), & \text{if } \alpha < \infty, \\ \mu(B)\mathbf{1}_{\{0\}}(k), & \text{otherwise.} \end{cases}$$

The conjectured limit in the case when $\alpha = \infty$ is obtained by taking the limit of $p_k^{\alpha}(B)$ as $\alpha \to \infty$. This limit is 0 unless $k = 0$, in which case it is $\mu(B)$.

The discussion in Section 1.4 of Chapter 1 described the parameter α as being closely related to the partition function $(\mathcal{Z}_n)_{n\in\mathbb{N}_0}$. As a result, we also conjecture:

Conjecture 2.1.2. *Let $\mathcal{T}$ be a (μ, f, ℓ) - RIF tree, with α as defined in* (2.3)*. Then we have*

$$\frac{\mathcal{Z}_n}{n} \xrightarrow{n\to\infty} \alpha, \qquad \textit{almost surely.}$$

2.1.2 Important Technical Conditions and Overview of Results

In this chapter, we make partial progress towards the proofs of Conjecture 2.1.1 and Conjecture 2.1.2. We will refer to the following technical conditions:

C1 With $m(\lambda, \cdot)$ as defined in (2.2), there exists some $\lambda > 0$ such that

$$1 < m(\lambda, \mathbb{R}_+) < \infty. \tag{2.4}$$

Under this condition, by monotonicity, there exists a unique $\alpha > 0$ such that $m(\alpha, \mathbb{R}_+) = 1$, we call this the *Malthusian parameter* associated with the process.

C2 There exists $\alpha > 0$ such that

$$\lim_{n\to\infty} \frac{\mathcal{Z}_n}{n} = \alpha.$$

Note that in (2.3), Conditions **C1** and **C2**, we use the same symbol α. This is because we conjecture that these coincide in general. In general, as we only assume either **C1** or **C2** at a time, the definition will be clear from context.

The chapter will be structured as follows:

Section 2.2: We analyse the model under Condition **C1**.

- In Theorem 2.2.1 we prove Conjecture 2.1.1 under Condition **C1**, and as a consequence, in Theorem 2.2.2 we show that for any measurable set B, $\Xi(n, B)/\ell n$ converges almost surely to $m(\alpha, B)$.
- In Theorem 2.2.5 we derive a condition under which **C1** implies **C2**. In particular, this proves Conjecture 2.1.2 under this condition and **C1**.

- The approaches used in this section are well-established, applying classical results in the theory of *Crump-Mode-Jagers branching processes*, in a similar manner to the approaches taken by the authors of [72, 41, 9, 29]. Nevertheless, these theorems have novel applications: we apply these theorems to the evolving Cayley tree considered by Bianconi in Example 2.2.4 and the *weighted random recursive tree.*

Section 2.3: We analyse a particular case of the model when the fitness function $f(i, W) = g(W)i + h(W)$, which we call the *generalised preferential attachment tree with fitness* (GPAF-tree). This extends the existing models of preferential attachment with additive fitness, i.e., $f(i, W) = i + 1 + W$, and multiplicative fitness, i.e., $f(i, W) = (i + 1)W$. When the functions g, h are non-decreasing, we also treat the cases where Condition **C1** can fail. Let α be as defined in (2.3), and also define $\Lambda := \{\lambda > 0 : m(\lambda, \mathbb{R}_+) < \infty\}$.

- We consider the situation in which Condition **C1** fails by having $m(\lambda, \mathbb{R}_+) \leqslant 1$ for all $\lambda \in \Lambda$. In this case, $m(\lambda, \mathbb{R}_+)$ converges for some $\lambda > 0$, but never exceeds 1, so that $m(\alpha, \mathbb{R}_+) \leqslant 1$. In Theorem 2.3.1 we prove Conjecture 2.1.1 and Conjecture 2.1.2 in this case, showing, in particular, that if $m(\alpha, \mathbb{R}_+) < 1$ the GPAF-tree exhibits a *condensation* phenomenon.

- Alternatively, Condition **C1** may fail by having $\alpha = \infty$. Theorem 2.3.3 also confirms Conjecture 2.1.1 in this case, showing that the limiting degree distribution is *degenerate*: almost surely the proportion of leaves in the tree tends to 1. Moreover, we show that the *fittest take all* of the mass of the distribution of edges according to weight, in the sense that *all* of the edges accumulate around vertices with maximum weight.

- The techniques in this section are inspired by the coupling techniques exploited in [20] and [29], and extend the well known phase transition associated with the model of preferential attachment with multiplicative fitnesses studied in [20, 31, 29]. This generalisation shows that the phase-transition depends on the parameter h too, so that, in some circumstances, condensation occurs, but vanishes if h is increased enough pointwise (see Section 2.3.2). This is interesting because $h(W)$ may be interpreted as the 'initial' popularity of a vertex when it arrives in the tree, showing that in order for the condensation to occur, there needs to be sufficiently many vertices of 'low enough' initial popularity. As far as the author is aware, these results are not only novel in the mathematical literature, but also in the general scientific literature concerning complex networks.

Section 2.4: We analyse the model under Condition **C2**, proving general results for the distribution of vertices with a given degree and weight.

- If the term α in Condition **C2** is finite, Theorem 2.4.1 and Theorem 2.4.4 confirm a weaker analogue of Conjecture 2.1.1 under this condition.

2.2 Analysis of (μ, f, ℓ) - RIF trees assuming C1

In order to apply Condition **C1** in this section, we study a branching processes with a *family tree* made up of individuals and their offspring whose distribution is identical to the discrete time model at the times of the branching events. In Section 2.2.1, we describe this continuous time model, state Theorem 2.2.1 and state and prove Theorem 2.2.2. In Section 2.2.2 we include the relevant theory of *Crump-Mode-Jagers* branching processes and use this to prove Theorem 2.2.1. In Section 2.2.3 we apply the same theory, along with some technical lemmas to state and prove a strong law of large numbers for the partition function in Theorem 2.2.5. We conclude the section with some interesting examples in Section 2.2.4.

2.2.1 Description of Continuous Time Embedding

In the continuous time approach, we begin with a population consisting of a single vertex 0 with weight W_0 sampled from μ and an associated exponential clock with parameter $f(0, W_0)$. Then recursively, when the ith birth event occurs in the population, with the ringing of an exponential clock associated to vertex j:

(i) Vertex j produces offspring $\ell(i-1)+1, \ldots, \ell i$ with independent weights $W_{\ell(i-1)+1}, \ldots, W_{\ell i}$ sampled from μ and exponential clocks with parameters $f(0, W_{\ell(i-1)+1}), \ldots, f(0, W_{\ell i})$.

(ii) Suppose the number of offspring of j before the birth event was m, so that its out-degree in the family tree is m. Then, the exponential random variable associated with j is updated to have rate $f(m/\ell+1, w_j)$. If $f(m/\ell+1, w_j) = 0$, then j ceases to produce offspring and we say j has *died*.

Now, if we let $\mathcal{Z}_{i-1}$ denote the sum of rates of the exponential clocks in the population when the population has size $i-1$, the probability that the clock associated with j is the first to ring is $f(m/\ell, W_j)/\mathcal{Z}_{i-1}$. Hence, the family tree of the continuous time model at the times of the birth events $(\sigma_i)_{i \geqslant 0}$ has the same distribution as the associated (μ, f, ℓ)-RIF tree. The continuous time branching process is actually a Crump-Mode-Jagers branching process, which we will describe in more depth in Section 2.2.2.

To describe the evolution of the degree of a vertex in the continuous time model, we define the pure birth process with underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and state space $\ell\mathbb{N}$ as follows: first sample a weight W and set $Y(0) = 0$. Let $\mathbb{P}_w$ denote the probability measure associated with the process when the weight sampled is w. Then, define the birth rates of Y such that

$$\mathbb{P}_w \left(Y(t+h) = (k+1)\ell \mid Y(t) = k\ell \right) = f(k, w)h + o(h). \tag{2.5}$$

In other words, the time taken to jump from $k\ell$ to $(k+1)\ell$ is exponentially distributed with parameter $f(k, w)$.

Let ρ denote the point process corresponding to the times of the jumps in Y and denote by $\mathbb{E}_w\left[\rho(\cdot)\right]$ the intensity measure when the weight $W = w$. Also, denote by $\hat{\rho}_w$ the Laplace-Stieltjes transform, i.e.,

$$\hat{\rho}_w(\lambda) := \int_0^\infty e^{-\lambda t}\mathbb{E}_w\left[\rho(\mathrm{d}t)\right].$$

Note that, by Fubini's theorem, we have

$$\begin{aligned}\hat{\rho}_w(\lambda) = \int_0^\infty \left(\int_t^\infty \lambda e^{-\lambda s}\mathrm{d}s\right)\mathbb{E}_w\left[\rho(\mathrm{d}t)\right] &= \int_0^\infty \lambda e^{-\lambda s}\left(\int_0^s \mathbb{E}_w\left[\rho(\mathrm{d}t)\right]\right)\mathrm{d}s \\ &= \int_0^\infty \lambda e^{-\lambda s}\mathbb{E}_w\left[Y(s)\right]\mathrm{d}s.\end{aligned} \tag{2.6}$$

Moreover, if we write τ_k for the time of the kth jump in Y, we have $\rho = \sum_{k=0}^\infty \ell\delta_{\tau_k}$. Note that, if the weight of Y is w, τ_k is distributed as a sum of independent exponentially distributed random variables with rates $f(0,w), f(1,w), \ldots, f(k-1,w)$ (we follow the convention that an exponential distributed random variable with rate 0 is ∞). Thus, we have that

$$\hat{\rho}_w(\lambda) = \ell\sum_{n=1}^\infty \mathbb{E}_w\left[e^{-\lambda\tau_n}\right] = \ell\sum_{n=1}^\infty\prod_{i=0}^{n-1}\frac{f(i,w)}{f(i,w)+\lambda}, \tag{2.7}$$

where in the last equality we have used the facts that a Laplace-Stieltjes transform of a convolution of measures is the product of Laplace-Stieltjes transforms and the Laplace-Stieltjes transform $\hat{X}(\lambda)$ of an exponential distributed random variable with parameter s is $\int_0^\infty e^{-\lambda t}se^{-st}\mathrm{d}t = \frac{s}{\lambda+s}$. Therefore, we see that $\mathbb{E}\left[\hat{\rho}_W(\lambda)\right] = m(\lambda, \mathbb{R}_+)$ as defined in (2.4), and Condition **C1** implies that there exists some $\lambda > 0$ such that $1 < \mathbb{E}\left[\hat{\rho}_W(\lambda)\right] < \infty$. In addition, the Malthusian parameter α appearing in Condition **C1** is the unique solution such that

$$\mathbb{E}\left[\hat{\rho}_W(\alpha)\right] = m(\alpha, \mathbb{R}_+) = \ell\cdot\mathbb{E}\left[\sum_{n=1}^\infty\prod_{i=0}^{n-1}\frac{f(i,W)}{f(i,W)+\alpha}\right] = 1. \tag{2.8}$$

Our first result is the following:

Theorem 2.2.1 (Convergence of the Degree Distribution under **C1**)**.** *Let $\mathcal{T}$ be a (μ, f, ℓ) -* RIF *tree satisfying Condition **C1** with Malthusian parameter α. Then, with $N_k(n, B)$ as defined in* (1.2) *and $p_k^\alpha(B)$ as defined in* (1.4)*, we have*

$$\frac{N_k(n, B)}{\ell n} \xrightarrow{n\to\infty} p_k^\alpha(B),$$

almost surely.

The limiting formula for Theorem 2.2.1 has appeared in a number of contexts, and generalises many known results. Under Condition **C1** this result was proved by Rudas, Tóth and Valkó [72] in the case that W is constant and $\ell = 1$. The cases $f(i, W) = W(i+1)$ and $f(i, W) = i+1+W$ with $\ell = 1$ correspond respectively to the preferential attachment models with multiplicative and additive fitness mentioned in the introduction. In the multiplicative model, the result was first proved in [20] and later in [9]. In [9], Bhamidi also first proved the result for the case $f(i, W) = i + 1 + W$. These models are examples of the generalised preferential attachment tree with fitness, which we study in more depth in Section 2.3. Finally, the case $f(i, W) = W$, $\ell = 1$ corresponds to a model of weighted random recursive trees (see Example 2.2.4). We postpone the proof of Theorem 2.2.1 to the end of Section 2.2.2.

Remark 2.2.1. *The limiting value has an interesting interpretation as a generalised geometric distribution. Consider an experiment where W is sampled from μ and, given W, coins are flipped, where the probability of heads in the ith coin flip is proportional to $f(i, W)$ and tails proportional to α. Then, the limiting distribution in Theorem 2.2.1 is the distribution of first occurrence of tails. Note that, by **C1**, the probability of infinite sequences of heads is* 0.

Remark 2.2.2. *Note that $Y(t) < \infty$ for all $t \geqslant 0$ almost surely if $\tau_\infty := \lim_{k\to\infty} \tau_k = \infty$ almost surely. The latter is satisfied if there exists $\lambda > 0$ such that for almost all w*

$$\mathbb{E}_w\left[e^{-\lambda\tau_\infty}\right] = \lim_{n\to\infty} \mathbb{E}_w\left[e^{-\lambda\tau_n}\right] = \lim_{n\to\infty} \prod_{i=0}^{n} \frac{f(i, w)}{f(i, w) + \lambda} = 0,$$

which is implied by **C1**. *In the literature concerning pure-birth Markov chains, this property is known as* non-explosivity.

Remark 2.2.3. *In this chapter, we have considered the case where the function f, and thus the birth process Y as defined in* (2.5)*, depends on a single random variable W taking values in $\mathbb{R}_+$. However, there is no loss of generality in assuming the random variable W takes values in an arbitrary measure space, so long as the function f is measurable. In particular, we may consider the case where the weight is given by a vector (W_1, W_2) where W_1 and W_2 are possibly correlated random variables.*

Now, recall the definitions of $\Xi(n, \cdot)$ from (1.6) and $m(\alpha, \cdot)$ from (1.9). In the case that $m(\alpha, \cdot)$ is a probability distribution, the almost sure convergence of $N_k(n, B)/\ell n$ to $p_k^\alpha(B)$ for any measurable set B is enough to imply that for any measurable set B the quantity $\Xi(n, B)$ converges almost surely to $m(\alpha, B)$. Note that this condition is weaker than directly assuming **C1**. In particular, we have the following.

Theorem 2.2.2. *Assume $\mathcal{T}$ is a* (μ, f, ℓ)-RIF *tree with limiting degree distribution of the form $(p_k^\alpha(\cdot))_{k\in\mathbb{N}_0}$ and such that the quantity $m(\alpha, \mathbb{R}_+) = 1$. Then, for any measurable set B, almost surely, we have*

$$\frac{\Xi(n, B)}{\ell n} \xrightarrow{n\to\infty} m(\alpha, B).$$

To prove this theorem, we will apply the following elementary lemma:

Lemma 2.2.3. *For any two sequences $(a_n)_{n\in\mathbb{N}}, (b_n)_{n\in\mathbb{N}}$, such that either $\liminf_{n\to\infty} a_n > -\infty$ or $\limsup_{n\to\infty} b_n < \infty$, we have*

$$\liminf_{n\to\infty}(a_n + b_n) \leqslant \liminf_{n\to\infty} a_n + \limsup_{n\to\infty} b_n \leqslant \limsup_{n\to\infty}(a_n + b_n).$$

Proof. We only prove the left inequality, as the right inequality is similar (or indeed is implied by the left combined with the fact that, for any sequence $(a_n)_{n\in\mathbb{N}}$, $\limsup_{n\to\infty}(-a_n) =$

$-\liminf_{n\to\infty} a_n)$. Let $\varepsilon > 0$ be given and suppose $\limsup_{n\to\infty} b_n = b$. Then, by definition, there exists $N > 0$ such that for all $n > N$ we have $b_n \leqslant b + \varepsilon$. But then,

$$\liminf_{n\to\infty} (a_n + b_n) \leqslant \liminf_{n\to\infty} (a_n + b + \varepsilon) = \liminf_{n\to\infty} a_n + b + \varepsilon.$$

Sending ε to 0 proves the result. $\square$

Proof of Theorem 2.2.2. Recall that, by (1.7), for each n, we have $\Xi(n, B) = \sum_{k=1}^{n} k\ell N_k(n, B)$. Also note that

$$\begin{aligned}
\sum_{k=0}^{\infty} k\ell p_k^{\alpha}(B) &= \ell \cdot \mathbb{E}\left[\left(\sum_{k=1}^{\infty} \frac{k\alpha}{f(k,W)+\alpha} \prod_{i=0}^{k-1} \frac{f(i,W)}{f(i,W)+\alpha}\right) \mathbf{1}_B(W)\right] \\
&= \ell \cdot \mathbb{E}\left[\left(\sum_{k=1}^{\infty} k \cdot \left(1 - \frac{f(k,W)}{f(k,W)+\alpha}\right) \prod_{i=0}^{k-1} \frac{f(i,W)}{f(i,W)+\alpha}\right) \mathbf{1}_B(W)\right] \\
&= \ell \cdot \mathbb{E}\left[\sum_{k=1}^{\infty} \left(k \prod_{i=0}^{k-1} \frac{f(i,W)}{f(i,W)+\alpha} - k \prod_{i=0}^{k} \frac{f(i,W)}{f(i,W)+\alpha}\right) \mathbf{1}_B(W)\right] \\
&= \ell \cdot \mathbb{E}\left[\left(\sum_{k=1}^{\infty} \prod_{i=0}^{k-1} \frac{f(i,W)}{f(i,W)+\alpha}\right) \mathbf{1}_B(W)\right] = m(\alpha, B),
\end{aligned}$$

where the second to last equality follows from the telescoping nature of the sum inside the expectation. Thus, by Fatou's lemma, almost surely we have

$$m(\alpha, B) = \sum_{k=0}^{\infty} k\ell p_k^{\alpha}(B) = \sum_{k=0}^{\infty} k\ell \liminf_{n\to\infty} \frac{N_k(n,B)}{\ell n} \leqslant \liminf_{n\to\infty} \frac{\Xi(n,B)}{\ell n}; \tag{2.9}$$

and likewise, almost surely, $\liminf_{n\to\infty} \frac{\Xi(n,B^c)}{\ell n} \geqslant m(\alpha, B^c)$. Now, since we add ℓ edges at every time-step, $\Xi(n, \mathbb{R}_+) = \ell n$. Thus, by Lemma 2.2.3

$$\begin{aligned}
1 = \liminf_{n\to\infty} \left(\frac{\Xi(n,B)}{\ell n} + \frac{\Xi(n,B^c)}{\ell n}\right) &\leqslant \liminf_{n\to\infty} \frac{\Xi(n,B^c)}{\ell n} + \limsup_{n\to\infty} \frac{\Xi(n,B)}{\ell n} \\
&\leqslant \limsup_{n\to\infty} \left(\frac{\Xi(n,B)}{\ell n} + \frac{\Xi(n,B^c)}{\ell n}\right) = 1.
\end{aligned}$$

But, $m(\alpha, \cdot)$ is a probability measure, this is only possible if

$$\liminf_{n\to\infty} \frac{\Xi(n,B^c)}{\ell n} = m(\alpha, B^c) \text{ and } \limsup_{n\to\infty} \frac{\Xi(n,B)}{\ell n} = m(\alpha, B) \text{ almost surely.} \tag{2.10}$$

Combining (2.9) and (2.10) completes the proof. $\square$

2.2.2 Crump-Mode-Jagers Branching Processes

In the continuous time setting, it is convenient to not only identify individuals of the branching process according to the order they were born, but also record their lineage, in such a way that the labelling encodes the structure of the tree. Therefore we also identify individuals of the branching process with elements of the infinite *Ulam-Harris* tree $\mathcal{U} := \bigcup_{n\geqslant 0}\mathbb{N}^n$, where $\mathbb{N}^0 = \varnothing$ is the *root.* In this case, an individual $u = u_1u_2\dots u_k$ is to be interpreted recursively as the u_kth child of the $u_1\dots u_{k-1}$. For example, $1, 2, \dots$ represent the offspring of $\varnothing$.

In *Crump-Mode-Jagers (CMJ)* branching processes, individuals $u \in \mathcal{U}$ are equipped with independent copies of a random point process ξ on $\mathbb{R}_+$. The point process ξ associates *birth times* to the offspring of a given individual, and we also may assume that ξ has some dependence on a random weight W associated with that individual. The process, together with birth times may be regarded as a random variable in the probability space $(\Omega, \Sigma, \mathbb{P}) = \prod_{x\in\mathcal{U}}(\Omega_x, \Sigma_x, \mathbb{P}_x)$ where each $(\Omega_x, \Sigma_x, \mathcal{P}_x)$ is a probability space with (ξ_x, W_x) having the same distribution as (ξ, W). We denote by $(\sigma_i^x)_{i\in\mathbb{N}}$ points ordered in the point process ξ_x and, for brevity, assume that $\xi(\{0\}) = 0$. We also drop the superscript when referring to the point process associated to $\varnothing$, so that $\sigma_i := \sigma_i^{\varnothing}$. Now, we set $\sigma_{\varnothing} := 0$ and recursively, for $x \in \mathcal{U}$, $\sigma_{xi} := \sigma_x + \sigma_i^x$. Finally, we set $\mathbb{T}_t = \{x \in \mathcal{U} : \sigma_x \leqslant t\}$ and note that for each $t \geqslant 0$, $\mathbb{T}_t$ may be identified with the *family tree* of the process in the natural way. Informally, $\mathbb{T}_t$ can be described as follows: at time zero, there is one vertex $\varnothing$, which reproduces according to $(\xi_{\varnothing}, W_{\varnothing})$. Thereafter, at times corresponding to points in $\xi_{\varnothing}$, descendants of $\varnothing$ are formed, which in turn produce offspring according to the same law. A crucial aspect of the study of CMJ processes are *characteristics* ϕ_x associated to each element $x \in \mathcal{U}$. For $x \in \mathcal{U}$, let $\mathcal{U}_x := \{xu : u \in \mathcal{U}\}$. Then, the processes ϕ_x are identically distributed, non-negative stochastic processes on the space $(\Omega, \Sigma, \mathbb{P})$ associated with individuals x, which may depend on $(\xi_z, W_z)_{z\in\mathcal{U}_x}$. Intuitively, these are processes that track 'characteristics' not only of the individual x, but on its potential offspring $\{xy : y \in \mathcal{U}\}$. We then define the *general branching process counted with characteristic* as

$$Z^{\phi}(t) := \sum_{x \in \mathcal{U}: \sigma_x \leqslant t} \phi_x(t - \sigma_x);$$

thus this function keeps a 'score' of characteristics of individuals in the family tree associated with the process up to time n. Let ν be the intensity measure of ξ, that is, $\nu(B) := \mathbb{E}[\xi(B)]$ for measurable sets $B \subseteq \mathbb{R}_+$. A crucial parameter in the study of CMJ processes is the *Malthusian parameter* α defined as the solution (if it exists) of

$$\mathbb{E}\left[\int_0^\infty e^{-\alpha u} \xi(\mathrm{d}u)\right] = 1.$$

Assume that ν is not supported on any lattice, i.e., for any $h > 0$ $\mathrm{Supp}(\nu) \subsetneq \{0, h, 2h, \ldots\}$, and that the first moment of $e^{-\alpha u}\nu(\mathrm{d}u)$ is finite, i.e., $\int_0^\infty u e^{-\alpha u} \nu(\mathrm{d}u) < \infty$. Nerman [65] proved the following theorem.

Theorem 2.2.4 ([65, Theorem 6.3])**.** *Suppose that there exists $\lambda < \alpha$ satisfying*

$$\mathbb{E}\left[\int_0^\infty e^{-\lambda s} \xi(\mathrm{d}s)\right] < \infty. \tag{2.11}$$

Then, for any two càdlàg characteristics $\phi^{(1)}, \phi^{(2)}$ such that $\mathbb{E}\left[\sup_{t \geqslant 0} e^{-\lambda t} \phi^{(i)}(t)\right] < \infty$, $i = 1, 2$, we have

$$\lim_{n \to \infty} \frac{Z^{\phi^{(1)}}(t)}{Z^{\phi^{(2)}}(t)} = \frac{\int_0^\infty \mathrm{e}^{-\alpha s} \mathbb{E}\left[\phi^{(1)}(s)\right] \mathrm{d}s}{\int_0^\infty \mathrm{e}^{-\alpha s} \mathbb{E}\left[\phi^{(2)}(s)\right] \mathrm{d}s},$$

almost surely on the event $\{|\mathbb{T}_t| \to \infty\}$.

Recall the definition of ρ as the point process associated with the jumps in the process Y defined in (2.5). Then, the continuous time model outlined in Section 2.2.1 is a CMJ process having ρ as its associated random point process and weight W. In this case, the Malthusian parameter is given by α in (2.8) and moreover, Condition **C1** implies that the first moment $\int_0^\infty t e^{-\alpha t} \hat{\rho}_\mu(\mathrm{d}t) < \infty$.

Theorem 2.2.1 is now an immediate application of Theorem 2.2.4.

Proof of Theorem 2.2.1. Consider the continuous time branching process outlined in Section 2.2.1 and denote by $\sigma'_1 < \sigma'_2 \cdots$ the times of births of individuals in the process. Then, $\mathcal{T}_n$ has the same distribution as the family tree $\mathbb{T}_{\sigma'_n}$. For any measurable set $B \subseteq \mathbb{R}$, define the characteristics $\phi^{(1)}(t) = \mathbf{1}_{\{Y(t)=k\ell, W\in B\}}$ and $\phi^{(2)}(t) = \mathbf{1}_{\{t\geqslant 0\}}$, where W denotes the weight of the process Y. Note that, $Z^{\phi^{(1)}}(t)$ is the number of individuals with $k\ell$ offspring and weight belonging to B up to time t, while $Z^{\phi^{(2)}}(t) = |\mathbb{T}_t|$. Thus,

$$\lim_{t\to\infty} \frac{Z^{\phi^{(1)}}(t)}{Z^{\phi^{(2)}}(t)} = \lim_{n\to\infty} \frac{N_k(n,B)}{\ell n}.$$

Note that both $\phi^{(1)}(t)$ and $\phi^{(2)}(t)$ are càdlàg and bounded and moreover, Condition **C1** implies that (2.11) is satisfied. Moreover, the assumption that $f(0,W) > 0$ almost surely implies that $|\mathbb{T}_t| \to \infty$ almost surely. Thus, by applying Theorem 2.2.4,

$$\lim_{t\to\infty} \frac{Z^{\phi^{(1)}}(t)}{Z^{\phi^{(2)}}(t)} = \alpha \int_0^\infty \mathrm{e}^{-\alpha s}\mathbb{E}\left[\mathbf{1}_{\{Y(s)=k\ell, W\in B\}}\right] \mathrm{d}s = \mathbb{E}\left[\mathbb{E}_W\left[\left(e^{-\alpha\tau_k} - e^{-\alpha\tau_{k+1}}\right)\right]\mathbf{1}_B(W)\right] \quad (2.12)$$

where the last equality follows from Fubini's theorem and we recall that τ_k is the time of the kth event in the process $Y_W(t)$. Now, since, when $W = w$, τ_k is distributed as a sum of independent exponentially distributed random variables with rates $f(0,w), f(1,w) \ldots$, we have

$$\mathbb{E}\left[\mathbb{E}_W\left[e^{-\alpha\tau_k}\right]\mathbf{1}_B(W)\right] = \mathbb{E}\left[\left(\prod_{i=0}^{k-1} \frac{f(i,W)}{f(i,W)+\alpha}\right)\mathbf{1}_B(W)\right]. \quad (2.13)$$

The result follows from combining (2.12) and (2.13). □

Remark 2.2.4. *As noted by the authors of [72], Theorem 2.2.4 can be applied to deduce a number of other properties of the tree, in particular the analogue of [72, Theorem 1] applies in this case as well.*

2.2.3 A Strong Law for the Partition Function

We can also apply Theorem 2.2.4 to show that the Malthusian parameter α emerges as the almost sure limit of the partition function, under certain conditions on the fitness function

f.

Theorem 2.2.5. *Let $(\mathcal{T}_n)_{n\geqslant 0}$ be a (μ, f, ℓ) - RIF tree satisfying **C1** with Malthusian parameter α. Moreover, assume that there exists a constant $C < \alpha$ and a non-negative function φ with $\mathbb{E}\left[\varphi(W)\right] < \infty$ such that, for all $k \in \mathbb{N}_0$, $f(k, W) \leqslant Ck + \varphi(W)$ almost surely. Then, almost surely*

$$\frac{\mathcal{Z}_n}{n} \xrightarrow{n\to\infty} \alpha.$$

In order to apply Theorem 2.2.4, we need to bound $\mathbb{E}\left[\sup_{t\geqslant 0} e^{-\lambda t}\phi^{(1)}(t)\right]$ for an appropriate choice of characteristic $\phi^{(1)}$ that tracks the evolution of the partition function associated with the process. In order to do so, using the assumptions on $f(i, W)$, we will couple the process Y defined in (2.5) with an appropriate pure birth process $(\mathcal{Y}(t))_{t\geqslant 0}$ (Lemma 2.2.9) and apply Doob's maximal inequality to a martingale associated with $(\mathcal{Y}(t))_{t\geqslant 0}$ (Lemma 2.2.8).

In order to define $\mathcal{Y}(t)$, first sample a weight W and set $\mathcal{Y}(0) = 0$. Then, if $\mathbb{P}_w$ denotes the probability measure associated with the process when the weight is w, define the rates such that

$$\mathbb{P}_w\left(\mathcal{Y}(t+h) = k+1 \mid \mathcal{Y}(t) = k\right) = (Ck + \varphi(w))h + o(h).$$ [1]

We also let $\mathcal{Y}_w$ denote the process with the same transition rates, but deterministic weight w.

It will be beneficial to state a more general result, about pure birth processes $(\mathcal{X}(t))_{t\geqslant 0}$ with linear rates, from the paper by Holmgren and Janson [41]. For brevity, we adapt the notation and only include some specific statements from both theorems.

Lemma 2.2.6 ([41, Theorem A.6 & Theorem A.7])**.** *Let $(\mathcal{X}(t))_{t\geqslant 0}$ be a pure birth process with $\mathcal{X}(0) = x_0$ and rates such that*

$$\mathbb{P}\left(\mathcal{X}(t+h) = k+1 \mid \mathcal{X}(t) = k\right) = (c_1 k + c_2)h + o(h),$$

[1] This process, when $C = 1$ and $\varphi(w) \equiv 0$, is often known as a Yule process.

for some constants $c_1, c_2 > 0$. Then, for each $t \geqslant 0$

$$\mathbb{E}\left[\mathcal{X}(t)\right] = \left(x_0 + \frac{c_2}{c_1}\right) e^{c_1 t} - \frac{c_2}{c_1}. \tag{2.14}$$

Moreover, if $x_0 = 0$ the probability generating function is given by

$$\mathbb{E}\left[z^{\mathcal{X}(t)}\right] = \left(\frac{e^{-c_1 t}}{1 - z\left(1 - e^{-c_1 t}\right)}\right)^{c_2/c_1}. \tag{2.15}$$

We also state a version of Doob's maximal inequality.

Lemma 2.2.7 (Doob's L^p Maximal Inequality, e.g. [Proposition 6.16, [49]])**.** *Let $(X_t)_{t\geqslant 0}$ be a sub-martingale and $S_t := \sup_{0\leqslant s\leqslant t} X_s$. Then, for any $T \geqslant 0$, $p > 1$*

$$\mathbb{E}\left[|S_T|^p\right] \leqslant \left(\frac{p}{p-1}\right)^p \mathbb{E}\left[|X_T|^p\right].$$

Finally, we will require Lemma 2.2.8 and Lemma 2.2.9.

Lemma 2.2.8. *For any $w > 0$, the process $(e^{-Ct}\left(\mathcal{Y}_w(t) + \varphi(w)/C\right))_{t\geqslant 0}$ is a martingale with respect to its natural filtration $(\mathcal{F}_t)_{t\geqslant 0}$. Moreover,*

$$\mathbb{E}\left[\sup_{t\geqslant 0}\left(e^{-Ct}\mathcal{Y}(t)\right)\right] < \infty.$$

Proof. The process $(\mathcal{Y}_w(t))_{t\geqslant 0}$ is a pure birth process satisfying the assumptions of Lemma 2.2.6, with $c_1 = C$ and $c_2 = \varphi(w)$. Therefore, by (2.14) and the Markov property, for any $t > s > 0$ we have

$$\mathbb{E}\left[\mathcal{Y}_w(t) \mid \mathcal{F}_s\right] = \mathbb{E}\left[\mathcal{Y}_w(t) \mid \mathcal{Y}_w(s)\right] = \left(\mathcal{Y}_w(s) + \frac{\varphi(w)}{C}\right) e^{C(t-s)} - \frac{\varphi(w)}{C},$$

which implies the martingale statement.

Moreover, applying (2.15) for the probability generating function, differentiating twice and evaluating at $z = 1$, we obtain

$$\mathbb{E}\left[\mathcal{Y}_w(t)\left(\mathcal{Y}_w(t) - 1\right)\right] = \frac{\varphi(w)\left(C + \varphi(w)\right)}{C^2}\left(e^{Ct} - 1\right)^2,$$

and thus

$$\mathbb{E}\left[(\mathcal{Y}(t)+\varphi(w)/C)^2\right] = \frac{\varphi(w)\left(C+\varphi(w)\right)}{C^2}\left(e^{Ct}-1\right)^2 + \left(2\varphi(w)/C+1\right)\frac{\varphi(w)}{C}\left(e^{Ct}-1\right) + \left(\varphi(w)/C\right)^2.$$

after some manipulations, we find that for all $t \geqslant 0$

$$\mathbb{E}\left[e^{-2Ct}\left(\mathcal{Y}_w(t)+\varphi(w)/C\right)^2\right] \leqslant \frac{\varphi(w)^2}{C^2} + \frac{\varphi(w)}{C}\left(1-e^{-Ct}\right).$$

Thus, we find that there exist constants A, B depending only on C such that for all $t \geqslant 0$

$$2\sqrt{\mathbb{E}\left[e^{-2Ct}\left(\mathcal{Y}_w(t)+\varphi(w)/C\right)^2\right]} \leqslant A + B\varphi(w).$$

Combining this L^2 quadratic bound with Doob's maximal inequality, we have

$$\begin{aligned}
\mathbb{E}\left[\sup_{t\geqslant 0}\left(e^{-Ct}\mathcal{Y}_w(t)\right)\right] &\leqslant \mathbb{E}\left[\sup_{t\geqslant 0}\left(e^{-Ct}\left(\mathcal{Y}_w(t)+\varphi(w)/C\right)\right)\right] \\
&\leqslant \sqrt{\mathbb{E}\left[\left(\sup_{t\geqslant 0}\left(e^{-Ct}\left(\mathcal{Y}_w(t)+\varphi(w)/C\right)\right)\right)^2\right]} \\
&\leqslant 2\sqrt{\mathbb{E}\left[e^{-2Ct}\left(\mathcal{Y}_w(t)+\varphi(w)/C\right)^2\right]} \\
&\leqslant A + B\varphi(w).
\end{aligned}$$

Thus,

$$\mathbb{E}\left[\sup_{t\geqslant 0}\left(e^{-Ct}\mathcal{Y}(t)\right)\right] = \mathbb{E}\left[\sup_{t\geqslant 0}\left(e^{-Ct}\mathcal{Y}_W(t)\right)\right] \leqslant A + B\mathbb{E}\left[\varphi(W)\right] < \infty.$$

□

Lemma 2.2.9. *Recall the definition of Y in (2.5) and assume that there exists a constant $C < \alpha$ and a non-negative function φ with $\mathbb{E}[\varphi(W)] < \infty$ such that, for all $k \in \mathbb{N}_0$, $f(k, W) \leqslant Ck + \varphi(W)$ almost surely. Then, there exists a coupling $(\hat{Y}(t), \hat{\mathcal{Y}}(t))_{t\geqslant 0}$ of $(Y(t))_{t\geqslant 0}$ and $(\mathcal{Y}(t))_{t\geqslant 0}$ such that, for all $t \geqslant 0$*

$$\hat{Y}(t) \leqslant \ell \cdot \hat{\mathcal{Y}}(t).$$

In the following proof, we denote by $\mathrm{Exp}\,(r)$ the exponential distribution with parameter r.

Proof. First, we sample $\hat{W}$ from μ and use this as a common weight for $\hat{Y}$ and $\hat{\mathcal{Y}}$. Now, let $(\varsigma_i)_{i\geqslant 0}$ be independent $\mathrm{Exp}\left(f(i,\hat{W})\right)$ distributed random variables. Then, for all $k>0$ set $\hat{\tau}_k=\sum_{i=0}^{k-1}\varsigma_i$ and

$$\hat{Y}(t)=\sum_{k=1}^{\infty}k\ell\mathbf{1}_{\hat{\tau}_k\leqslant t<\hat{\tau}_{k+1}}.$$

The ς_i can be interpreted as the intermittent time between jumps from state i to $i+\ell$. For all $t>0$ construct the jump times of $(\hat{\mathcal{Y}}(t))_{t\geqslant 0}$ iteratively as follows:

- Note that by assumption $f(0,\hat{W})\leqslant\varphi(\hat{W})$. Let $e_0\sim\mathrm{Exp}\left(\varphi(\hat{W})-f(0,\hat{W})\right)$ and set $\varsigma_0'=\min\{e_0,\varsigma_0\}$. We may interpret ς_0' as the time for $\hat{\mathcal{Y}}$ to jump from 0 to 1.
- Given $\varsigma_0',\ldots,\varsigma_j'$, let $q_j:=\sum_{i=0}^{j}\varsigma_i'$ and define $m_j:=\hat{Y}(q_j)/\ell$, i.e., the value of $\hat{Y}/\ell$ once $\hat{\mathcal{Y}}$ has reached $j+1$. Assume inductively that $m_j\leqslant j+1$ and set

$$e_{j+1}\sim\mathrm{Exp}\left(C(j+1)+\varphi(\hat{W})-f(m_j,\hat{W})\right)\quad\text{and}\quad\varsigma_{j+1}'=\min\left\{e_j,\varsigma_{m_j}\right\}.$$

Observe that, since $\varsigma_{j+1}'\leqslant\varsigma_{m_j+1}$, we have $m_{j+1}\leqslant j+2$, so we may iterate this procedure.

It is clear that $(\hat{Y}(t))_{t\geqslant 0}$ is distributed like $(Y(t))_{t\geqslant 0}$ and using the properties of the exponential distribution, one readily confirms that $(\hat{\mathcal{Y}}(t))_{t\geqslant 0}$ is distributed like $(\mathcal{Y}(t))_{t\geqslant 0}$. Finally, the desired inequality follows from the fact that $\hat{\mathcal{Y}}(t)$ always jumps before or at the same time as $\hat{Y}(t)$. □

Proof of Theorem 2.2.5. Consider the continuous time embedding of the (μ,f,ℓ)-RIF tree and define the characteristics $\phi^{(1)}(t):=\sum_{k=0}^{\infty}f(k,W)\mathbf{1}_{\{Y(t)=k\ell\}}$ and $\phi^{(2)}(t):=\mathbf{1}_{\{t\geqslant 0\}}$. Recall that we denote by $(\tau_i)_{i\geqslant 1}$ the times of the jumps in Y and that, for all $k\geqslant 0$, $f(k,W)\leqslant$

$Ck + \varphi(W)$. Then, by Lemma 2.2.9, Lemma 2.2.8 and the assumptions of the theorem,

$$\mathbb{E}\left[\sup_{t\geqslant 0}\left(e^{-Ct}\phi^{(1)}(t)\right)\right] \overset{\text{Lem. 2.2.9}}{\leqslant} \mathbb{E}\left[\sup_{t\geqslant 0}\left(e^{-Ct}\left(C\mathcal{Y}_W(t)+\varphi(W)\right)\right)\right] \overset{\text{Lem. 2.2.8}}{<} \infty.$$

Now, in this case $Z^{\phi^{(1)}}(t)$ is the total sum of fitnesses of individuals born up to time t, while $Z^{\phi^{(2)}}(t) = |\mathcal{T}_t|$. Thus, by Theorem 2.2.4 and Fubini's theorem in the second equality, almost surely we have

$$\begin{aligned}\lim_{n\to\infty}\frac{\mathcal{Z}_n}{\ell n} &= \alpha\int_0^{\infty} e^{-\alpha s}\mathbb{E}\left[\sum_{k=0}^{\infty} f(k,W)\mathbf{1}_{\{Y(s)=k\ell\}}\right]\mathrm{d}s \qquad (2.16)\\ &= \mathbb{E}\left[\sum_{k=0}^{\infty} f(k,W)\left(e^{-\alpha\tau_k}-e^{-\alpha\tau_{k+1}}\right)\right] = \mathbb{E}\left[\sum_{k=1}^{\infty}\frac{\alpha f(k,W)}{f(k,W)+\alpha}\prod_{i=0}^{k-1}\frac{f(i,W)}{f(i,W)+\alpha}\right].\end{aligned}$$

Now, recall that by (2.8) we have

$$\mathbb{E}\left[\sum_{k=1}^{\infty}\frac{f(k,W)}{f(k,W)+\alpha}\prod_{i=0}^{k-1}\frac{f(i,W)}{f(i,W)+\alpha}\right] = \frac{1}{\ell},$$

and combining this with (2.16) proves the result. □

2.2.4 Examples of Applications of Theorem 2.2.1

Weighted Cayley Trees

Consider the model where $f(k,W) = 0$ for $k \geqslant 1$ and $f(0,W) = g(W)$. Thus, at each step, a vertex with degree 0 is chosen and produces ℓ children and thus this model produces an $(\ell+1)$-*Cayley* tree, i.e., a tree in which each node that is not a leaf has degree $\ell+1$. Without loss of generality, by considering the pushforward of μ under g if necessary, we may assume that $g(W) = W$. In this case, $\hat{\rho}_\mu(\lambda) = \ell\cdot\mathbb{E}\left[\frac{W}{W+\lambda}\right]$ and thus **C1** is satisfied as long as $\ell \geqslant 2$. Thus, $p_k^{\alpha}(B) = 0$ for all $k \geqslant 2$ and

$$p_0(B) = \mathbb{E}\left[\frac{\alpha}{W+\alpha}\mathbf{1}_B(W)\right], \quad p_1(B) = \mathbb{E}\left[\frac{W}{W+\alpha}\mathbf{1}_B(W)\right].$$

This rigorously confirms a result of Bianconi [10]. Note however, that in [10], α is described as the almost sure limit of the partition function and we may only apply Theorem 2.2.5 under the assumption that $\mathbb{E}[W] < \infty$.

In the notation of [10], the weights W are called 'energies', using the symbol ϵ, the function $g(\epsilon) := e^{\beta\epsilon}$, where $\beta > 0$ is a parameter of the model, and $\alpha := e^{\beta\mu_F}$ is described as the limit of the partition function. Thus, the proportion of vertices with out-degree 0 with 'energy' belonging to some measurable set B is

$$\mathbb{E}\left[\frac{1}{e^{\beta(\epsilon-\mu_F)}+1}\mathbf{1}_B(W)\right],$$

which is known as a *Fermi-Dirac* distribution in physics.

Weighted Random Recursive Trees

In the case that $f(k, W) = W$, we obtain a model of *weighted random recursive trees* with independent weights and **C1** is satisfied with $\alpha = \mathbb{E}[W]$ provided $\mathbb{E}[W] < \infty$. Theorem 2.2.1 then implies that

$$\frac{N_k(n, B)}{\ell n} \xrightarrow{n\to\infty} \mathbb{E}\left[\frac{\ell\mathbb{E}[W]\, W^k}{(W + \ell\mathbb{E}[W])^{k+1}}\mathbf{1}_B(W)\right],$$

almost surely. This was observed in the case $\ell = 1$ by the authors of [37, Proposition 3]. Note also that in this case Theorem 2.2.5 coincides with the usual strong law of large numbers.

The weighted random recursive tree has a natural generalisation to affine fitness functions. This is the topic of the next section.

2.3 Generalised Preferential Attachment Trees with Fitness

In this section, we study (μ, f, ℓ) - RIF trees in the specific case when the function f takes an affine form, that is, $f(i, W) = ig(W) + h(W)$, for positive, measurable functions g, h. We call this particular case of the model a *generalised preferential attachment tree with fitness* (which we abbreviate as a GPAF-tree). The affine form of this model mean that it is tractable to apply the coupling methods outlined in Section 2.3.2, when Condition **C1** fails, and the functions g and h are non-decreasing. Moreover, this model is general enough to be an extension of not only the weighted random recursive tree, but also of the additive and multiplicative models studied in [20, 9].

The results, and techniques used in this section will inspire us to study a further generalisation of this model, the preferential attachment tree with neighbourhood influence (PANI-tree) in Chapter 3; in the latter the fitness function is affine, but also incorporates information about the weights of the neighbours of a given vertex. Below, in Section 2.3.1 we apply the theory of the previous section to this model when **C1** is satisfied. In the rest of Section 2.3, we assume that the associated functions g and h are non-decreasing. In Section 2.3.2, we analyse the model when Condition **C1** fails by having $m(\lambda, \mathbb{R}_+) \leqslant 1$ for all $\lambda > 0$ such that $m(\lambda, \mathbb{R}_+) < \infty$, stating and proving Theorem 2.3.1. Then, in Section 2.3.3 we analyse the model when Condition **C1** fails by having $m(\lambda, \mathbb{R}_+) = \infty$ for all $\lambda > 0$, stating and proving Theorem 2.3.3.

Note that in this section, we formulate our results in terms of functions g and h depending on a random variable W taking values in $\mathbb{R}_+$. However, in the vein of Remark 2.2.3, we expect these results to extend to cases where g and h may depend on more general random variables. For example, there is no loss of generality in assuming g and h depend on possibly correlated random variables W_1 and W_2 assigned to a given vertex. In this case, the coupling technique applied in Section 2.3.2 needs to be adjusted accordingly, with appropriate "truncations" of the vector (W_1, W_2).

2.3.1 When the GPAF-tree satisfies Condition C1

In the context of the GPAF-tree, Condition **C1** states that there exists $\lambda > 0$ such that

$$m(\lambda, \mathbb{R}_+) = \ell \cdot \mathbb{E}\left[\sum_{n=1}^{\infty} \prod_{i=0}^{n-1} \frac{g(W)i + h(W)}{g(W)i + h(W) + \lambda}\right] > 1.$$

First recall the definition of the birth process Y from (2.5) in Section 2.2, with $f(k, w) = g(W)k + h(W)$. By applying (2.14) from Lemma 2.2.6 and the initial condition $Y(0) = 0$, for any $w \in \mathbb{R}_+$ we have

$$\mathbb{E}_w\left[Y(t)\right] = \left(\frac{h(w)}{g(w)}\right) e^{\ell g(w) t} - \frac{h(w)}{g(w)}.$$

Now, (2.6) and (2.7) in Section 2.2 showed that

$$\ell \cdot \sum_{n=1}^{\infty} \prod_{i=0}^{n-1} \frac{g(W)i + h(W)}{g(W)i + h(W) + \lambda} = \int_0^{\infty} \lambda e^{-\lambda s} \mathbb{E}_w\left[Y(s)\right] \mathrm{d}s = \begin{cases} \frac{h(w)}{\lambda/\ell - g(w)} & \text{if } \lambda/\ell > g(w); \\ \infty & \text{otherwise.} \end{cases} \tag{2.17}$$

For a measurable function $g : \mathbb{R}_+ \to \mathbb{R}_+$ we define $\operatorname{ess\,sup}(g)$ such that

$$\operatorname{ess\,sup}(g) := \inf\left\{a \in \mathbb{R}_+ : \mu\left(\{x : g(x) > a\}\right) = 0\right\}.$$

Therefore by (2.17), for $\lambda \geqslant \ell \cdot \operatorname{ess\,sup}(g)$ we have $m(\lambda, \mathbb{R}_+) = \mathbb{E}\left[\frac{h(W)}{\lambda/\ell - g(W)}\right]$, while if $\lambda < \ell \cdot \operatorname{ess\,sup}(g)$ we have $m(\lambda, \mathbb{R}_+) = \infty$. Thus, Condition **C1** is satisfied if $\operatorname{ess\,sup}(g) < \infty$, $\mathbb{E}\left[h(W)\right] < \infty$ and, for some $\lambda \geqslant \ell \cdot \operatorname{ess\,sup}(g)$

$$1 < \mathbb{E}\left[\frac{h(W)}{\lambda/\ell - g(W)}\right] < \infty.$$

As a result, the Malthusian parameter α appearing in Condition **C1** is given by the unique $\alpha > 0$ such that

$$\mathbb{E}\left[\frac{h(W)}{\alpha/\ell - g(W)}\right] = 1. \tag{2.18}$$

Note that the parameter ℓ in the model has the effect of re-scaling the Malthusian parameter α. Also, since $\alpha \geqslant \ell \cdot \operatorname{ess\,sup}(g)$, if $\mathbb{E}\left[h(W)\right] < \infty$, Theorem 2.2.5 applies and α may also be interpreted as the almost sure limit of the partition function associated with the process. Now, in the context of this model, the limiting value $p_k^{\alpha}(\cdot)$ from Theorem 2.2.1 is such that

$$p_k^{\alpha}(B) = \mathbb{E}\left[\frac{\alpha}{g(W)k + h(W) + \alpha}\prod_{i=0}^{k-1}\frac{g(W)i + h(W)}{g(W)i + h(W) + \alpha}\mathbf{1}_B(W)\right]. \tag{2.19}$$

Now, recall Stirling's approximation, which states that

$$\Gamma(z) = \left(1 + O(1/z)\right) z^{z-\frac{1}{2}} e^{-z}. \tag{2.20}$$

If $g(W) > 0$ on B, by dividing the numerator and denominator of terms inside the product in (2.19), we obtain a ratio of Gamma functions. Thus, by applying Stirling's approximation, on any measurable set B on which g, h are bounded, we have

$$p_k^{\alpha}(B) = \left(1 + O(1/k)\right)\mathbb{E}\left[c_B k^{-\left(1+\frac{\alpha}{g(W)}\right)}\mathbf{1}_B(W)\right],$$

where c_B, which comes from the term outside the product in (2.19), depends on g and h but not k. Thus, the distribution of $(p_k^{\alpha}(B))_{k\in\mathbb{N}_0}$ follows what one might describe as an 'averaged' power law. Moreover, in the case $\ell = 1$, $\alpha \geqslant \operatorname{ess\,sup}(g)$, thus,

$$\mathbb{E}\left[c_B k^{-\left(1+\frac{\alpha}{g(W)}\right)}\mathbf{1}_B(W)\right] \geqslant c' k^{-2}$$

for some $c' > 0$. It has been observed that real world complex networks, have power law degree distributions where the observed power law exponent lies between 2 and 3 (see, for example, [77]). Note that by (2.18), α depends on both h and g, so that keeping g fixed and making h smaller has the effect of reducing the exponent of the power law.

In the remainder of this section we set $\ell = 1$, for brevity. The arguments may be adapted in a similar manner to the case $\ell > 1$.

2.3.2 A Condensation Phenomenon when Condition C1 Fails

Recall that, in the GPAF-tree, if $\lambda \geqslant \operatorname{ess\,sup}(g)$ we have

$$m(\lambda, \mathbb{R}_+) = \mathbb{E}\left[\frac{h(W)}{\lambda - g(W)}\right], \tag{2.21}$$

and if $\lambda < \operatorname{ess\,sup}(g)$, we have $m(\lambda, \mathbb{R}_+) = \infty$. If we define

$$\Lambda := \{\lambda > 0 : m(\lambda, \mathbb{R}_+) < \infty\},$$

in this subsection, we consider the case that the GPAF-tree fails to satisfy Condition **C1** by having $m(\lambda, \mathbb{R}_+) \leqslant 1$ for all $\lambda \in \Lambda$. We show that in this case the GPAF-tree satisfies a formula for the degree distribution of the same form as (1.4). Moreover, if $\lambda^* := \inf(\Lambda)$ and $m(\lambda^*, \mathbb{R}_+) < 1$, this model exhibits a condensation phenomenon, as described in Theorem 2.3.1. We remark that such results have been proved for the case of the preferential attachment tree with multiplicative fitness, i.e., the case $h \equiv g$, in [31], in a more general framework; that is to say encompassing other models apart from a tree.

In Section 2.3.2 we state our main result, Theorem 2.3.1 and discuss interesting implications in Section 2.3.2. In Section 2.3.2 we state and prove Lemma 2.3.2 which is the crucial tool used in proofs of the theorem. The proof of Theorem 2.3.1 is deferred to Section 2.3.2.

Note that in the case that g and h are bounded, we have $\tilde{\lambda} = \operatorname{ess\,sup}(g) < \infty$. Without loss of generality, we re-scale the measure μ and re-define g and h such that $\operatorname{Supp}(\mu) \subseteq [0, w^*]$, where $w^* := \sup(\operatorname{Supp}(\mu)) < \infty$. For example, we may replace W by $\arctan(W)$ and g and h by $g \circ \tan$ and $h \circ \tan$. Such a re-scaling does not affect the monotonicity of g, h and the boundedness assumption implies that $g(w^*), h(w^*) < \infty$. Moreover, if $\mathcal{T}$ does not satisfy **C1**, the monotonicity of g implies that μ does not have an atom at w^*, since in this case $\operatorname{ess\,sup}(g) = g(w^*)$. Thus, for each $\varepsilon > 0$, we have

$$\mu([w^* - \varepsilon, w^*]) > 0, \tag{2.22}$$

and, re-defining g such that $g(w^*) = \lim_{\varepsilon \to 0} g(w^* - \varepsilon)$ if necessary, we may assume without loss of generality that g is continuous at w^*. We adopt these assumptions for the rest of this subsection.

Theorem 2.3.1: Condensation in the GPAF-tree

Our main result in this subsection is the following theorem, which demonstrates the possibility of condensation in this model. Define the measure $\pi(\cdot)$ such that, for any measurable set B,

$$\pi(B) = \mathbb{E}\left[\frac{h(W)}{g(w^*) - g(W)}\mathbf{1}_B(W)\right] + \left(1 - \mathbb{E}\left[\frac{h(W)}{g(w^*) - g(W)}\right]\right)\delta_{w^*}(B).$$

Theorem 2.3.1. *Suppose $\mathcal{T} = (\mathcal{T}_n)_{n \geqslant 0}$ is a* GPAF*-tree, with associated functions g, h, where g, h are non-decreasing and bounded and Condition* ***C1*** *fails. Then we have the following assertions:*

- *With regards to the weak topology,*

  $$\frac{\Xi(n, \cdot)}{\ell n} \xrightarrow{n \to \infty} \pi(\cdot), \quad \textit{almost surely.}$$

  In particular, if $\mathbb{E}\left[\frac{h(W)}{g(w^*) - g(W)}\right] < 1$, this model exhibits a condensation phenomenon, as described before Conjecture 2.1.1 in Section 1.4.

- *For any measurable set B, almost surely we have*

  $$\frac{N_k(n, B)}{n} \xrightarrow{n \to \infty} \mathbb{E}\left[\frac{g(w^*)}{g(W)k + h(W) + g(w^*)}\prod_{i=0}^{k-1}\frac{g(W)i + h(W)}{g(W)i + h(W) + g(w^*)}\mathbf{1}_B(W)\right],$$

  i.e., $\frac{N_k(n,B)}{n} \xrightarrow{n \to \infty} p_k^{g(w^*)}(B)$ almost surely.

- *The partition function*

  $$\frac{\mathcal{Z}_n}{n} \xrightarrow{n \to \infty} g(w^*), \quad \textit{almost surely.}$$

Remark 2.3.1. *By applying a more refined coupling argument to the one presented in Lemma 2.3.2, we can actually improve this result to remove the assumption that h is non-decreasing. We omit the details, but instead refer the reader to Section 3.3 in Chapter 3, where we present a more refined coupling.*

Some Interesting Implications of the Condensation Phenomenon

The condensation result in Theorem 2.3.1 has interesting implications for the GPAF-tree. Informally, the parameter $g(w)$ measures the extend to which the 'popularity' of a vertex with weight w is reinforced by the number of its neighbours, while the parameter $h(w)$ represents its 'initial popularity'. The condensation phenomenon then depends on both μ and h, in the sense that condensation occurs if vertices of high weight are 'rare enough' and the initial popularity is 'low enough'. More precisely, if we assume g, h are non-decreasing and bounded, we can see two particular regimes of the tree:

1. If μ is such that $\mathbb{E}\left[\frac{1}{g(w^*)-g(W)}\right] = \infty$, then, for any non-decreasing bounded function h, Condition **C1** is satisfied in this model, and thus, the model does not demonstrate a condensation phenomenon.

2. Otherwise, if g is such that $\mathbb{E}\left[\frac{1}{g(w^*)-g(W)}\right] = C < \infty$, then either

$$\mathbb{E}\left[\frac{h(W)}{g(w^*) - g(W)}\right] > 1 \quad \text{or} \quad \mathbb{E}\left[\frac{h(W)}{g(w^*) - g(W)}\right] \leqslant 1.$$

   In the first case, Condition **C1** is satisfied, but fails in the second case. However, in the second case, if the inequality is strict, condensation arises. Therefore, for fixed g, condensation in this model arises by reducing h sufficiently point-wise, for example, by replacing h by $K \cdot h$ where $K < 1/C$ is a constant.

Remark 2.3.2. *Note that the first regime shows that whenever g attains its essential supremum on a set of positive measure, Condition **C1** is satisfied. This will be important in the* *couplings employed in in the rest of the section.*

A Coupling Lemma

In order to prove Theorem 2.3.1, we first prove an additional lemma. For each $\varepsilon > 0$ such that $\varepsilon < w^*$, let $\mathcal{T}^{+\varepsilon} = (\mathcal{T}_n^{+\varepsilon})_{n \geqslant 0}$ and $\mathcal{T}^{-\varepsilon} = (\mathcal{T}_n^{-\varepsilon})_{n \geqslant 0}$ denote GPAF-trees with the same functions g, h, but with weights $W^{(+\varepsilon)}, W^{(-\varepsilon)}$ distributed like

$$W \mathbf{1}_{[0, w^* - \varepsilon]}(W) + w^* \mathbf{1}_{(w^* - \varepsilon, w^*]}(W) \quad \text{and} \quad W \wedge (w^* - \varepsilon) \quad \text{respectively.}$$

The motivation behind these choices of $\mathcal{T}^{+\varepsilon}$ and $\mathcal{T}^{-\varepsilon}$ is that they have distributions with atoms at the value maximising g almost everywhere. Thus, by (2.22) and Remark 2.3.2, these trees satisfy Condition **C1**, and we may apply the theorems from Section 2.2 with regards to these trees. Then, provided these trees provide sufficiently good 'approximations' of the tree $\mathcal{T}$, we may deduce certain results by sending ε to 0.

In this vein, let $N^{+\varepsilon}_{\geqslant k}(n, B), N_{\geqslant k}(n, B)$ and $N^{-\varepsilon}_{\geqslant k}(n, B)$ denote the number of vertices with out-degree $\geqslant k$ and weight belonging to the set B in $\mathcal{T}_n^{+\varepsilon}, \mathcal{T}_n$ and $\mathcal{T}_n^{-\varepsilon}$ respectively. In their respective trees, we also denote by $W_i^{(+\varepsilon)}, W_i$ and $W_i^{(-\varepsilon)}$ the weight of a vertex i and $\mathcal{Z}_n^{+\varepsilon}, \mathcal{Z}_n$ and $\mathcal{Z}_n^{-\varepsilon}$ the partition functions at time n. Finally, for brevity, we write $f_n^{(+\varepsilon)}(v), f_n(v)$ and $f_n^{(-\varepsilon)}(v)$ for the fitness of a vertex v at time n in each of these models. In other words, $f_n(v) = g(W_v) \deg^+(v, \mathcal{T}_n) + h(W_v)$.

Lemma 2.3.2. *There exists a coupling $(\hat{\mathcal{T}}^{+\varepsilon}, \hat{\mathcal{T}}, \hat{\mathcal{T}}^{-\varepsilon})$ of these processes such that for all $n \in \mathbb{N}_0$,*

- *For any $x < w^* - \varepsilon$ we have $\Xi^{+\varepsilon}(n, [0, x]) \leqslant \Xi(n, [0, x]) \leqslant \Xi^{-\varepsilon}(n, [0, x])$,*
- *For all measurable sets $B \subseteq [0, w^* - \varepsilon)$ and $k \in \mathbb{N}_0$, we have*

$$N^{+\varepsilon}_{\geqslant k}(n, B) \leqslant N_{\geqslant k}(n, B) \leqslant N^{-\varepsilon}_{\geqslant k}(n, B),$$

- $\mathcal{Z}_n^{-\varepsilon} \leqslant \mathcal{Z}_n \leqslant \mathcal{Z}_n^{+\varepsilon}$.

Proof of Lemma 2.3.2. Initialise the trees with a vertex 0 having weight W_0 sampled independently from μ in $\hat{\mathcal{T}}_0$ and weights $W_0^{(+\varepsilon)} = W_0 \mathbf{1}_{[0,w*-\varepsilon]}(W_0) + w^* \mathbf{1}_{(w*-\varepsilon,w*]}(W_0)$ and $W_0^{(-\varepsilon)} = W_0 \wedge (w^* - \varepsilon)$ in $\hat{\mathcal{T}}_0^{+\varepsilon}$ and $\hat{\mathcal{T}}_0^{-\varepsilon}$. Assume, that at the nth time-step,

$$(\hat{\mathcal{T}}_t^{+\varepsilon})_{0 \leqslant t \leqslant n} \sim (\mathcal{T}_t^{+\varepsilon})_{0 \leqslant t \leqslant n}, \quad (\hat{\mathcal{T}}_t)_{0 \leqslant t \leqslant n} \sim (\mathcal{T}_t)_{0 \leqslant t \leqslant n} \quad \text{and} \quad (\hat{\mathcal{T}}_t^{-\varepsilon})_{0 \leqslant t \leqslant n} \sim (\mathcal{T}_t^{-\varepsilon})_{0 \leqslant t \leqslant n}.$$

In addition, assume, by induction, that we have $\mathcal{Z}_n^{-\varepsilon} \leqslant \mathcal{Z}_n \leqslant \mathcal{Z}_n^{+\varepsilon}$ and for each vertex v with $W_v^{(+\varepsilon)} = W_v = W_v^{(-\varepsilon)} < w^* - \varepsilon$ we have

$$\deg^+(v, \hat{\mathcal{T}}_n^{+\varepsilon}) \leqslant \deg^+(v, \hat{\mathcal{T}}_n) \leqslant \deg^+(v, \hat{\mathcal{T}}_n^{-\varepsilon}). \tag{2.23}$$

Note that (2.23) implies the first and the second assertions of the lemma up to time n. As a result, for each vertex v with $W_v < w^* - \varepsilon$ we have $f_n^{(+\varepsilon)}(v) \leqslant f_n(v) \leqslant f_n^{(-\varepsilon)}(v)$. Now, for the $(n+1)$st step

- Introduce a vertex $n+1$ with weight W_{n+1} sampled independently from μ and set $W_{n+1}^{(+\varepsilon)} = W_{n+1} \mathbf{1}_{[0,w*-\varepsilon]}(W_{n+1}) + w^* \mathbf{1}_{(w*-\varepsilon,w*]}(W_{n+1})$ and $W_{n+1}^{(-\varepsilon)} = W_{n+1} \wedge (w^* - \varepsilon)$.

- Form $\hat{\mathcal{T}}_{n+1}^{-\varepsilon}$ by sampling the parent v of $n+1$ independently according to the law of $\mathcal{T}^{-\varepsilon}$, i.e., with probability proportional to $f_n^{(-\varepsilon)}(v)$. Then, in order to form $\hat{\mathcal{T}}_{n+1}$ sample an independent uniformly distributed random variables U_1 on $[0, 1]$.

  - If $U_1 \leqslant \frac{\mathcal{Z}_n^{-\varepsilon} f_n(v)}{\mathcal{Z}_n f_n^{(-\varepsilon)}(v)}$ and $W_v^{(-\varepsilon)} < w^* - \varepsilon$, select v as the parent of $n+1$ in $\hat{\mathcal{T}}_{n+1}$ as well.

  - Otherwise, form $\hat{\mathcal{T}}_{n+1}$ by selecting the parent v' of $n+1$ with probability proportional to $f_n(v')$ out of all all the vertices with weight $W_{v'} \geqslant w^* - \varepsilon$.

- Then form $\hat{\mathcal{T}}_{n+1}^{+\varepsilon}$ in a similar manner. Sample an independent uniform random variable U_2 on $[0, 1]$.

– If a vertex v with weight $W_v < w^* - \varepsilon$ was chosen as the parent of $n+1$ in $\hat{\mathcal{T}}_{n+1}$ and also $U_2 \leqslant \frac{\mathcal{Z}_n f_n^{(+\varepsilon)}(v)}{\mathcal{Z}_n^{+\varepsilon} f_n(v)}$, also select v as the parent of $n+1$ in $\hat{\mathcal{T}}_{n+1}^{+\varepsilon}$.

– Otherwise, form $\hat{\mathcal{T}}_{n+1}^{+\varepsilon}$ by selecting the parent v'' of $n+1$ with probability proportional to $f_n^{(+\varepsilon)}(v'')$ out of all all the vertices with weight $W_{v''} = w^*$.

It is clear that $\hat{\mathcal{T}}_{n+1}^{-\varepsilon} \sim \mathcal{T}_{n+1}^{-\varepsilon}$. On the other hand, in $\hat{\mathcal{T}}_{n+1}$ the probability of choosing a parent v of $n+1$ with weight $W_v < w^* - \varepsilon$ is

$$\frac{\mathcal{Z}_n^{-\varepsilon} f_n(v)}{\mathcal{Z}_n f_n^{(-\varepsilon)}(v)} \times \frac{f_n^{(-\varepsilon)}(v)}{\mathcal{Z}_n^{-\varepsilon}} = \frac{f_n(v)}{\mathcal{Z}_n},$$

whilst the probability of choosing a parent v' with weight $W_{v'} \geqslant w^* - \varepsilon$ is

$$\begin{aligned}
\frac{f_n(v')}{\sum_{v:W_v \geqslant w^*-\varepsilon} f_n(v)} & \left(\sum_{v:W_v^{(-\varepsilon)} < w^*-\varepsilon} \left(1 - \frac{\mathcal{Z}_n^{-\varepsilon} f_n(v)}{\mathcal{Z}_n f_n^{(-\varepsilon)}(v)}\right) \frac{f_n^{(-\varepsilon)}(v)}{\mathcal{Z}_n^{-\varepsilon}} \right) \\
& \qquad + \frac{f_n(v')}{\sum_{v:W_v \geqslant w^*-\varepsilon} f_n(v)} \left(\sum_{v:W_v^{(-\varepsilon)} = w^*-\varepsilon} \frac{f_n^{(-\varepsilon)}(v)}{\mathcal{Z}_n^{-\varepsilon}} \right) \\
&= \frac{f_n(v')}{\sum_{v:W_v \geqslant w^*-\varepsilon} f_n(v)} \left(\sum_v \frac{f_n^{(-\varepsilon)}(v)}{\mathcal{Z}_n^{-\varepsilon}} - \sum_{v:W_v^{(-\varepsilon)} < w^*-\varepsilon} \frac{f_n(v)}{\mathcal{Z}_n} \right) \\
&= \frac{f_n(v')}{\sum_{v:W_v \geqslant w^*-\varepsilon} f_n(v)} \left(1 - \frac{\sum_{v:W_v^{(-\varepsilon)} < w^*-\varepsilon} f_n(v)}{\mathcal{Z}_n} \right) = \frac{f_n(v')}{\mathcal{Z}_n},
\end{aligned}$$

where we use the fact that $\sum_v f_n(v) = \mathcal{Z}_n$. Thus, we have $\hat{\mathcal{T}}_{n+1} \sim \mathcal{T}_{n+1}$. Moreover, either the same vertex is chosen as the parent of $n+1$ in both $\hat{\mathcal{T}}_{n+1}^{-\varepsilon}$ and $\hat{\mathcal{T}}_{n+1}$, or a vertex of higher weight, at least $w^* - \varepsilon$, is chosen as the parent of $n+1$ in $\hat{\mathcal{T}}_{n+1}$. This implies the left inequality in (2.23) and in addition, when combined with the fact that $W_{n+1}^{(-\varepsilon)} \leqslant W_{n+1}$ and g, h are non-decreasing, guarantees that $\mathcal{Z}_{n+1}^{-\varepsilon} \leqslant \mathcal{Z}_{n+1}$. The proof that $\hat{\mathcal{T}}_{n+1}^{+\varepsilon} \sim \mathcal{T}_{n+1}^{+\varepsilon}$, the right inequality in (2.23) and $\mathcal{Z}_{n+1} \leqslant \mathcal{Z}_{n+1}^{+\varepsilon}$ are similar, so we may thus iterate the coupling. $\square$

Proof of Theorem 2.3.1

In order to prove Theorem 2.3.1, we first define the auxiliary GPAF-trees $\mathcal{T}^{+\varepsilon}$ and $\mathcal{T}^{-\varepsilon}$ according to Lemma 2.3.2.

Proof of Theorem 2.3.1. For the first assertion, by the definition of weak convergence, we need only check that

$$\frac{\Xi(n,[0,x])}{\ell n} \xrightarrow{n\to\infty} \pi([0,x])$$

almost surely, at any point where $x \mapsto \pi([0,x])$ is continuous. Suppose $x < w^*$. For $\varepsilon > 0$ sufficiently small that $x < w^* - \varepsilon$, define the corresponding quantities $\Xi^{+\varepsilon}(n,\cdot)$, $\Xi^{-\varepsilon}(n,\cdot)$ associated with $\mathcal{T}^{+\varepsilon}$ and $\mathcal{T}^{-\varepsilon}$. Then, from the coupling in Lemma 2.3.2, we have

$$\frac{\Xi^{+\varepsilon}(n,[0,x])}{n} \leqslant \frac{\Xi(n,[0,x])}{n} \leqslant \frac{\Xi^{-\varepsilon}(n,[0,x])}{n}.$$

Note that the auxiliary trees $\mathcal{T}^{+\varepsilon}$ and $\mathcal{T}^{-\varepsilon}$ have associated weight distributions which contain an atom at their maximum value and thus, by Remark 2.3.2, satisfy Condition **C1**, with Malthusian parameters $\alpha^{(-\varepsilon)} > g(w^* - \varepsilon)$ and $\alpha^{(+\varepsilon)} > g(w^*)$. Moreover, note that, by the definition of $W^{(-\varepsilon)}$,

$$\mathbb{E}\left[\frac{h(W^{(-\varepsilon)})}{g(w^*) - g(W^{(-\varepsilon)})}\right] \leqslant \mathbb{E}\left[\frac{h(W)}{g(w^*) - g(W)}\right] \leqslant 1,$$

so that, recalling (2.18), $\alpha^{(-\varepsilon)} \leqslant g(w^*)$. Thus, since $x < w^* - \varepsilon$, by Lemma 2.3.2, dominated convergence and continuity of g at w^*, almost surely we have

$$\limsup_{n\to\infty} \frac{\Xi(n,[0,x])}{n} \leqslant \lim_{\varepsilon\to 0} \mathbb{E}\left[\frac{h(W)}{\alpha^{(-\varepsilon)} - g(W)}\mathbf{1}_{[0,x]}(W)\right] = \mathbb{E}\left[\frac{h(W)}{g(w^*) - g(W)}\mathbf{1}_{[0,x]}(W)\right].$$

Now, $\alpha^{(+\varepsilon)}$ is non-increasing in ε, and we have $\lim_{\varepsilon\to 0}\alpha^{(+\varepsilon)} = g(w^*)$. Indeed, suppose by way of a contradiction that $\lim_{\varepsilon\to 0}\alpha^{(+\varepsilon)} = \alpha' > g(w^*)$. Then,

$$\frac{h(w^*)}{\alpha' - g(w^*)} < \infty,$$

and thus by dominated convergence,

$$1 = \lim_{\varepsilon \to 0} \mathbb{E}\left[\frac{h(W^{(+\varepsilon)})}{\alpha^{(+\varepsilon)} - g(W^{(+\varepsilon)})}\right] = \mathbb{E}\left[\frac{h(W)}{\alpha' - g(W)}\right].$$

But then, (2.18) is satisfied for λ such that $g(w^*) < \lambda < \alpha'$, contradicting the assumption that Condition **C1** fails for $\mathcal{T}$.

It follows that $\lim_{\varepsilon \to 0} \alpha^{(+\varepsilon)} = g(w^*)$ and thus, by Lemma 2.3.2 and dominated convergence, almost surely we have

$$\liminf_{n \to \infty} \frac{\Xi(n, [0, x])}{n} \leqslant \lim_{\varepsilon \to 0} \mathbb{E}\left[\frac{h(W)}{\alpha^{(+\varepsilon)} - g(W)} \mathbf{1}_{[0,x]}(W)\right] = \mathbb{E}\left[\frac{h(W)}{g(w^*) - g(W)} \mathbf{1}_{[0,x]}(W)\right].$$

The first assertion follows.

For the second assertion, given a measurable set B, for each $\varepsilon > 0$, set $B^{\varepsilon} := B \cap [0, w^* - \varepsilon)$. In addition, note that, conditional on taking values in B^{ε} the random variables $W, W^{(-\varepsilon)}$ and $W^{(+\varepsilon)}$ are identically distributed. Combining these facts with Lemma 2.3.2, almost surely we have

$$\begin{aligned}
\limsup_{n \to \infty} \frac{N_{\geqslant k}(n, B)}{n} &\leqslant \liminf_{\varepsilon \to 0} \left(\mathbb{E}\left[\prod_{i=0}^{k-1} \frac{g(W^{(-\varepsilon)})i + h(W^{(-\varepsilon)})}{g(W^{(-\varepsilon)})i + h(W^{(-\varepsilon)}) + \alpha^{(-\varepsilon)}} \mathbf{1}_{B^{\varepsilon}}(W)\right] + \mu([w^* - \varepsilon, w^*]\right) \\
&= \liminf_{\varepsilon \to 0} \mathbb{E}\left[\prod_{i=0}^{k-1} \frac{g(W)i + h(W)}{g(W)i + h(W) + \alpha^{(-\varepsilon)}} \mathbf{1}_{B^{\varepsilon}}(W)\right] \\
&= \mathbb{E}\left[\prod_{i=0}^{k-1} \frac{g(W)i + h(W)}{g(W)i + h(W) + g(w^*)} \mathbf{1}_{B}(W)\right],
\end{aligned}$$

where we have applied dominated convergence in the final equality. Similarly, almost surely,

$$\begin{aligned}
\liminf_{n \to \infty} \frac{N_{\geqslant k}(n, B)}{n} &\geqslant \limsup_{\varepsilon \to 0} \mathbb{E}\left[\prod_{i=0}^{k-1} \frac{g(W^{(+\varepsilon)})i + h(W^{(+\varepsilon)})}{g(W^{(+\varepsilon)})i + h(W^{(+\varepsilon)}) + \alpha^{(+\varepsilon)}} \mathbf{1}_{B^{\varepsilon}}(W)\right] \\
&= \limsup_{\varepsilon \to 0} \mathbb{E}\left[\prod_{i=0}^{k-1} \frac{g(W)i + h(W))}{g(W)i + h(W) + \alpha^{(+\varepsilon)}} \mathbf{1}_{B^{\varepsilon}}(W)\right] \\
&= \mathbb{E}\left[\prod_{i=0}^{k-1} \frac{g(W)i + h(W))}{g(W)i + h(W) + g(w^*)} \mathbf{1}_{B}(W)\right].
\end{aligned}$$

Finally, for the last assertion, by Lemma 2.3.2, for each $n \in \mathbb{N}_0$ we have

$$\frac{\mathcal{Z}_n^{-\varepsilon}}{n} \leqslant \frac{\mathcal{Z}_n}{n} \leqslant \frac{\mathcal{Z}_n^{+\varepsilon}}{n}.$$

Taking limits as n goes to infinity and applying Theorem 2.2.5, the result follows in a similar manner to the previous assertions. □

2.3.3 Degenerate Degrees when Condition C1 Fails

In this subsection, we show that if the GPAF-tree fails to satisfy Condition **C1** by having $m(\lambda, \mathbb{R}_+) = \infty$ for all $\lambda > 0$, almost surely the proportion of vertices that are leaves tends to 1. Consequentially, the limiting mass of edges 'escapes to infinity', as described in Theorem 2.3.3 below. Note that Condition **C1** fails in this manner in the GPAF tree if $\operatorname{ess\,sup}(g) = \infty$ or $\mathbb{E}[h(W)] = \infty$. We remark that similar results to Theorem 2.3.3 have been shown in preferential attachment model with multiplicative fitness with μ having finite support [20, Theorem 6] and preferential attachment model with additive fitness (the *extreme disorder* regime in [54, Theorem 2.6]. These cases correspond to $h(x) \equiv 0$ and $g(x) \equiv 1$ respectively.

As in the previous subsection, we re-scale the measure μ and re-define g and h such that $\operatorname{Supp}(\mu) \subseteq [0, w^*]$, where $w^* := \sup(\operatorname{Supp}(\mu))$. In this case, however, we have either $g(w^*) = \infty$ or $h(w^*) = \infty$, and since $g(W), h(W) < \infty$ almost surely in order for the model to be well-defined, this implies that μ does not contain an atom at w^*.

Theorem 2.3.3. *Suppose $\mathcal{T} = (\mathcal{T}_n)_{n \geqslant 0}$ is a* GPAF*-tree, with associated functions g, h, with g, h non-decreasing such that* $\operatorname{ess\,sup}(g) = \infty$ *or* $\mathbb{E}[h(W)] = \infty$. *Then we have the following assertions:*

- *With regards to the weak topology*

$$\frac{\Xi(n, \cdot)}{\ell n} \xrightarrow{n \to \infty} \delta_{w^*}(\cdot), \quad \textit{almost surely.}$$

- *For any measurable set $B \subseteq [0, w^*]$, we have*

$$\frac{N_0(n,B)}{n} \xrightarrow{n\to\infty} \mu(B), \quad \textit{almost surely.} \tag{2.24}$$

Proof. This is similar to the proof of Theorem 2.3.1. For each $\varepsilon > 0$ set $B^{\varepsilon} := B \cap [0, w^* - \varepsilon]$, let $\mathcal{T}^{-\varepsilon} = (\mathcal{T}_n^{-\varepsilon})_{n \geqslant 0}$ denote the GPAF-tree, with weights $W^{(-\varepsilon)}$ distributed like $W \wedge (w^* - \varepsilon)$. Let $N_{\geqslant k}^{-\varepsilon}(n,B)$, $N_{\geqslant k}(n,B)$ denote the number of vertices with out-degree $\geqslant k$ and weight belonging to B in $\mathcal{T}_n^{-\varepsilon}$ and $\mathcal{T}_n$ respectively. The following claim follows in an analogous manner to Lemma 2.3.2:

Claim. *There exists a coupling $(\hat{\mathcal{T}}, \hat{\mathcal{T}}^{-\varepsilon})$ of $\mathcal{T}$ and $\mathcal{T}^{-\varepsilon}$ such that for all $n \in \mathbb{N}_0$ we have the following:*

- *For all $x < w^* - \varepsilon$ we have $\Xi(n,[0,x]) \leqslant \Xi^{-\varepsilon}(n,[0,x])$.*

- *For all measurable sets $B \subseteq [0, w^* - \varepsilon)$ we have $N_{\geqslant k}(n,B) \leqslant N_{\geqslant k}^{-\varepsilon}(n,B)$.*

Now note that $\mathcal{T}^{-\varepsilon}$ has a weight distribution with an atom at its maximum value, and thus, by Remark 2.3.2, satisfies **C1**, with Malthusian parameter $\alpha^{(-\varepsilon)}$. Moreover, note $\alpha^{(-\varepsilon)}$ is monotonically increasing as ε decreases. In addition, the assumptions on g and h imply that $m(\lambda, \mathbb{R}_+)$ as defined in (2.21) is infinite for all $\lambda > 0$. Therefore,

$$\lim_{\varepsilon \to 0} \alpha^{(-\varepsilon)} = \infty.$$

Now, for the first assertion, as in the proof of Theorem 2.3.1, we need only check that

$$\frac{\Xi(n,[0,x])}{\ell n} \xrightarrow{n\to\infty} 0,$$

almost surely, for all $x < w^*$. But now, for ε sufficiently small that $x < w^* - \varepsilon$, by the claim we have

$$\limsup_{n\to\infty} \frac{\Xi(n,[0,x])}{n} \leqslant \limsup_{n\to\infty} \frac{\Xi^{-\varepsilon}(n,[0,x])}{n} = \mathbb{E}\left[\frac{h(W)}{\alpha^{(-\varepsilon)} - g(W)} \mathbf{1}_{[0,x]}(W)\right].$$

Taking the limit as $\varepsilon \to 0$ proves the result.

For the second assertion, by the claim and applying, for example, dominated convergence in the right hand inequality, for all $k \geqslant 1$ we have

$$\begin{aligned}\limsup_{n\to\infty} \frac{N_{\geqslant k}(n,B)}{n} \\ \leqslant \liminf_{\varepsilon\to 0}\left(\mathbb{E}\left[\prod_{i=0}^{k-1}\frac{g(W^{(-\varepsilon)})i+h(W^{(-\varepsilon)})}{g(W^{(-\varepsilon)})i+h(W^{(-\varepsilon)})+\alpha^{(-\varepsilon)}}\mathbf{1}_{B^\varepsilon}(W)\right]+\mu(B\backslash B^\varepsilon)\right)=0.\end{aligned}$$

Then (2.24) follows from the strong law of large numbers, which implies that $\frac{N_{\geqslant 0}(n,B)}{n} \to \mu(B)$ almost surely. □

2.4 Analysis of (μ, f, ℓ) - RIF trees assuming C2

By Theorem 2.2.5, under certain conditions on the fitness function f and **C1**, Condition **C2** is satisfied, i.e.,

$$\frac{\mathcal{Z}_n}{n} \xrightarrow{n\to\infty} \alpha, \quad \text{almost surely.}$$

However, Theorem 2.3.1 shows that this condition may be satisfied despite Condition **C1** failing. Therefore, in this section, we analyse the model under Condition **C2**. In particular, we make the heuristic outlined in Section 1.4.1 of Chapter 1 precise, showing that the limit of $N_k(n,\cdot)/\ell n$ is closely linked to the almost sure limit of the partition function.

The methods applied in this section are closely related to those of Section 4.4 of Chapter 4, which also apply the summation arguments stated and proved in Section 2.4.2 below. However, the results in this section have significantly fewer technical difficulties, and, in addition, we present a much shorter proof of convergence of the mean of $N_k(n,B)/\ell n$. Therefore, we recommend the reader study this section closely before reading Chapter 4. We state and prove Theorem 2.4.1 below and state Theorem 2.4.4, leaving the details to the reader. These proofs rely on Proposition 2.4.2, proved in Section 2.4.3 and Section 2.4.4; and Proposition 2.4.3, proved in Section 2.4.5.

2.4.1 Convergence in probability of $N_k(n,B)/\ell n$ under C2

Theorem 2.4.1. *Assume **C2**. Then, for any measurable set B we have*

$$\frac{N_k(n,B)}{\ell n} \xrightarrow{n\to\infty} \mathbb{E}\left[\frac{\alpha}{f(k,W)+\alpha}\prod_{s=0}^{k-1}\frac{f(s,W)}{f(s,W)+\alpha}\mathbf{1}_B(W)\right] = p_k^{\alpha}(B), \quad \textit{in probability.}$$

In order to prove Theorem 2.4.1, we define the following family of sets:

$$\mathscr{F} := \{B : B \text{ is measurable and } \forall s \in \mathbb{N}_0,\ f(s,w) \text{ is bounded for } w \in B\}. \tag{2.25}$$

We also require Proposition 2.4.2 and Proposition 2.4.3, proved in Section 2.4.4 and Section 2.4.5. These proofs rely on the results stated in Section 2.4.2 and Section 2.4.3.

Proposition 2.4.2. *For any set $B \in \mathscr{F}$, for each $k \in \mathbb{N}_0$ we have*

$$\lim_{n\to\infty}\frac{\mathbb{E}\left[N_k(n,B)\right]}{\ell n} = p_k^{\alpha}(B).$$

Proposition 2.4.3. *For any $B \in \mathscr{F}$ and $k \in \mathbb{N}_0$ we have*

$$\lim_{n\to\infty}\mathbb{E}\left[\frac{(N_k(n,B))^2}{\ell^2 n^2}\right] = (p_k^{\alpha}(B))^2.$$

Proof of Theorem 2.4.1. The result follows for all $B \in \mathscr{F}$ by combining Proposition 2.4.2, Proposition 2.4.3 and applying Chebyshev's inequality.

Now, let B be an arbitrary measurable set and let $\varepsilon > 0$ be given. Then, since, by the definition of the model in Section 1.3.2 of Chapter 1, for each $s \in \{1,\ldots,k\}$ the map $w \mapsto f(s,w)$ is measurable, by Lusin's theorem we can find a compact set $E \subseteq B$ such that

$\mu(B \cap E^c) < \varepsilon/3$ and for each $s \in \{1, \ldots, k\}$ the map $w \mapsto f(s, w)$ is continuous on E. Moreover, note that $p_k^\alpha(B) - p_k^\alpha(B \cap E) \leqslant \mu(B \cap E^c) < \varepsilon/3$. Then,

$$\begin{aligned}
&\mathbb{P}\left(\left|\frac{N_k(n,B)}{\ell n} - p_k^\alpha(B)\right| > \varepsilon\right) \\
&\leqslant \mathbb{P}\Bigg(\Bigg(\left|\frac{N_k(n,B)}{\ell n} - \frac{N_k(n, B\cap E)}{\ell n}\right| + \left|\frac{N_k(n, B\cap E)}{\ell n} - p_k^\alpha(B \cap E)\right| \\
&\qquad\qquad + |p_k^\alpha(B\cap E) - p_k^\alpha(B)|\Bigg) > \varepsilon\Bigg) \\
&\leqslant \mathbb{P}\left(\left|\frac{N_k(n, B\cap E)}{\ell n} - p_k^\alpha(B \cap E)\right| > \varepsilon/3\right) \\
&\qquad\qquad + \mathbb{P}\left(\left|\frac{N_k(n,B)}{\ell n} - \frac{N_k(n, B\cap E)}{\ell n}\right| > \varepsilon/3\right). \qquad (2.26)
\end{aligned}$$

Now, note that by the strong law of large numbers, and Egorov's theorem, for any $\delta > 0$ there exists an event G with $\mathbb{P}(G) < \delta$ such that

$$\limsup_{n\to\infty}\left(\frac{N_k(n,B)}{\ell n} - \frac{N_k(n,B\cap E)}{\ell n}\right) = \limsup_{n\to\infty}\frac{N_k(n, B\cap E^c)}{\ell n} \leqslant \mu(B \cap E^c)$$

on the complement of G. Therefore, the result follows from (2.26), Proposition 2.4.2 and Proposition 2.4.3 by taking limits as n tends to infinity. □

Using the approach to the upper bound for the mean in the next subsection, and applying Corollary 2.4.6 stated below with $k = 1$ and $e_0, e_1 = 0$, if $N_{\geqslant 1}(n, B)$ denotes the number of vertices of out-degree at least 1 in the tree with weight belonging to B, we actually have

$$\limsup_{n\to\infty}\frac{\mathbb{E}\left[N_{\geqslant 1}(n,B)\right]}{\ell n} \leqslant \frac{1}{\alpha'}\mathbb{E}\left[f(0,W)\mathbf{1}_B(W)\right],$$

as long as $\liminf_{n\to\infty}\frac{\mathcal{Z}_n}{n} \geqslant \alpha'$. By sending α' to infinity, this yields the following analogue of Theorem 2.3.3:

Theorem 2.4.4. *Suppose $\mathcal{T}$ is a (μ, f, ℓ)-RIF tree such that $\lim_{n\to\infty}\frac{\mathcal{Z}_n}{n} = \infty$. Then for any measurable set $B \subseteq [0, \infty)$, we have*

$$\frac{N_0(n,B)}{n} \xrightarrow{t\to\infty} \mu(B), \quad \textit{in probability.}$$

2.4.2 Summation Arguments

Here we state and prove some summation arguments required for the subsequent proofs, in particular, the proofs in the rest of this section, as well as in the proofs of Section 4.4 of Chapter 4. For $e_0, \dots, e_k \geqslant 0, 0 \leqslant \eta < 1$, let

$$\mathcal{S}_n(e_0, \dots, e_k, \eta) := \frac{1}{n} \sum_{\eta n < i_0 < \dots < i_k \leqslant n} \prod_{j=0}^{k-1} \left(\left(\frac{i_j}{i_{j+1}} \right)^{e_j} \cdot \frac{1}{i_{j+1} - 1} \right) \left(\frac{i_k}{n} \right)^{e_k}.$$

Lemma 2.4.5. *Uniformly in $e_0, \dots, e_k \geqslant 0$, $0 \leqslant \eta \leqslant 1/2$, we have*

$$\mathcal{S}_n(e_0, \dots, e_k, \eta) = \prod_{j=0}^{k} \frac{1}{e_j + 1} + \theta(\eta) + O\left(\frac{1}{n^{1/(k+2)}} + \frac{\sum_{j=0}^{k} e_j \log^{k+1}(n)}{n} \right).$$

Here, $\theta(\eta)$ is a term satisfying $|\theta(\eta)| \leqslant M\eta^{1/(k+2)}$ for some universal constant M depending only on k.

Corollary 2.4.6. *For $e_0, \dots, e_k, f_0, \dots, f_{k-1} \geqslant 0$, $0 \leqslant \eta \leqslant 1/2$, we have*

$$\frac{1}{n} \sum_{\eta n < i_0 \leqslant n} \sum_{\mathcal{I}_k \in \binom{\{i_0+1, \dots, n\}}{k}} \prod_{j=0}^{k-1} \left(\left(\frac{i_j}{i_{j+1}} \right)^{e_j} \cdot \frac{f_j}{i_{j+1} - 1} \right) \left(\frac{i_k}{n} \right)^{e_k}$$
$$= \frac{1}{e_k + 1} \prod_{j=0}^{k-1} \frac{f_j}{e_j + 1} + \theta'(\eta) + O\left(\frac{1}{n^{1/(k+2)}} \right).$$

Here, $\theta'(\eta)$ is a term satisfying $|\theta'(\eta)| \leqslant M'\eta^{1/(k+2)}$ for some universal constant M' depending only on k and $f_0, \dots, f_{k-1}$, and the constant in the big O-term may depend on $e_0, \dots, e_k, f_0, \dots, f_k$.

To prepare the proof of the lemma, we rewrite the relevant sums using probabilistic language. Let $U_0, \dots, U_k$ be $k+1$ independent random variables uniformly distributed on $[0, 1]$. We write $U_{(0)} \leqslant \dots \leqslant U_{(k)}$ for their order statistics. Let $I_j = \lceil U_{(j)} n \rceil, j \in \{0, \dots, k\}$. Then, $I_n = (I_0, \dots, I_k)$ is the vector of order statistics of $k+1$ independent random variables with uniform distribution on $\{1, \dots, n\}$. Let A_n be the event that these random variables are distinct. Then, for $e_0, \dots, e_k \geqslant 0, 0 < \eta \leqslant 1/2$, we have

$$\begin{aligned} \mathcal{S}_n(e_0, \dots, e_k, \eta) &= \frac{1}{n} \sum_{\eta n < i_0 < \dots < i_k \leqslant n} \prod_{j=0}^{k-1} \left(\left(\frac{i_j}{i_{j+1}} \right)^{e_j} \cdot \frac{1}{i_{j+1} - 1} \right) \left(\frac{i_k}{n} \right)^{e_k} \\ &= \frac{1}{(k+1)!} \cdot \mathbb{E} \left[\prod_{j=0}^{k-1} \left(\left(\frac{I_j}{I_{j+1}} \right)^{e_j} \cdot \frac{n}{I_{j+1} - 1} \right) \left(\frac{I_k}{n} \right)^{e_k} \mathbf{1}_{A_n} \mathbf{1}_{I_0 > \eta n} \right]. \end{aligned}$$

Here, the $(k+1)!$ term corresponds to the $(k+1)!$ ways a vector of $k+1$ uniform random variables on $\{1, \dots, n\}$ can be $(e_0, \dots, e_k)$. Note that, given $U_{(i)}, U_{(i+1)}, \dots, U_{(k)}$, the random variables $U_{(0)}, \dots, U_{(i-1)}$ are distributed like the order statistics of i independent random variables with the uniform distribution on $[0, U_{(i)}]$. Now, $U_{(k)}$ is distributed like $U^{1/(k+1)}$, where U follows the uniform distribution on $[0,1]$; indeed, for any $x \in [0,1]$

$$\mathbb{P}\left(U_{(k)} \leqslant x\right) = x^{k+1} = \mathbb{P}\left(U^{1/(k+1)} \leqslant x\right).$$

Moreover, for any $i \in \{0, \dots, k-1\}$,

$$\mathbb{P}\left(U_{(i)} \leqslant x \,|\, U_{(i+1)}\right) = \left(\frac{x}{U_{(i+1)}} \right)^{i+1} \wedge 1 = \mathbb{P}\left(U_i^{1/i+1} \cdot U_{(i+1)} \leqslant x \,|\, U_{(i+1)}\right),$$

for an independent random variable U_i uniformly distributed on $[0,1]$. Thus, setting

$$V_i := U_i^{1/(i+1)} U_{i+1}^{1/(i+2)} \cdots U_k^{1/(k+1)}, \quad \text{for} \quad i \in \{0, \dots, k\},$$

the random vectors $(U_{(0)}, \dots, U_{(k)})$ and $(V_0, \dots, V_k)$ are equal in distribution. Therefore, by applying the dominated convergence theorem, for $\eta = 0$ we have

$$\lim_{n \to \infty} \mathcal{S}_n(e_0, \dots, e_k, 0) = \frac{1}{(k+1)!} \cdot \mathbb{E} \left[\prod_{j=0}^{k-1} \left(\left(\frac{U_{(j)}}{U_{(j+1)}} \right)^{e_j} \cdot \frac{1}{U_{(j+1)}} \right) U_{(k)}^{e_k} \right].$$

The last term is equal to

$$\begin{aligned} \frac{1}{(k+1)!} \cdot \mathbb{E} \left[\prod_{j=0}^{k-1} \left(\frac{V_j}{V_{j+1}} \right)^{e_j} \cdot V_k^{e_k} \prod_{j=0}^{k-1} \frac{1}{V_{j+1}} \right] &= \frac{1}{(k+1)!} \cdot \mathbb{E} \left[\prod_{j=0}^{k} U_j^{e_j/(j+1)} \prod_{j=0}^{k} U_j^{-j/(j+1)} \right] \\ &= \prod_{j=0}^{k} \frac{1}{e_j + 1}. \end{aligned}$$

Proof of Lemma 2.4.5. We start with the term involving η. Note that $\prod_{j=0}^{k-1} \frac{n}{I_{j+1}-1}\mathbf{1}_{A_n} \leqslant 2\prod_{j=0}^{k-1} U_{(j+1)}^{-1}$, since on the event A_n, we have $I_1 \geqslant 2$. Thus,

$$
\begin{aligned}
&\mathbb{E}\left[\prod_{j=0}^{k-1}\left(\left(\frac{I_j}{I_{j+1}}\right)^{e_j}\cdot\frac{n}{I_{j+1}-1}\right)\left(\frac{I_k}{n}\right)^{e_k}\mathbf{1}_{A_n}\mathbf{1}_{I_0\leqslant \eta n}\right] \\
&\leqslant 2\mathbb{E}\left[\prod_{j=0}^{k-1} U_{(j+1)}^{-1}\mathbf{1}_{I_0\leqslant \eta n}\right] \leqslant 2\mathbb{E}\left[\prod_{j=0}^{k-1} U_{(j+1)}^{-(k+2)/(k+1)}\right]^{(k+1)/(k+2)} \mathbb{P}\left(I_0 \leqslant \eta n\right)^{1/(k+2)} \\
&\leqslant 2\left(k+1\right)^{(1+k(k+1))/(k+2)}\eta^{1/(k+2)}.
\end{aligned}
$$

Here, in the last step, we have used $\mathbb{P}\left(I_0 \leqslant \eta n\right) \leqslant \mathbb{P}\left(U_{(0)} \leqslant \eta\right) = 1-(1-\eta)^{k+1} \leqslant (k+1)\eta$. Next, let $\Delta_{j+1} = \frac{n}{I_{j+1}-1} - \frac{1}{U_{(j+1)}}$. In the computation of

$$
\mathbb{E}\left[\prod_{j=0}^{k-1}\left(\left(\frac{I_j}{I_{j+1}}\right)^{e_j}\cdot\frac{n}{I_{j+1}-1}\right)\left(\frac{I_k}{n}\right)^{e_k}\mathbf{1}_{A_n}\right],
$$

we can now successively replace $\frac{n}{I_{j+1}-1}$ by $\frac{1}{U_{(j+1)}} + \Delta_{j+1}$ for $j \in \{0,\dots,k-1\}$. As $\Delta_{j+1} \to 0$ almost surely, it follows from the dominated convergence theorem, that

$$
\begin{aligned}
&\mathbb{E}\left[\prod_{j=0}^{k-1}\left(\left(\frac{I_j}{I_{j+1}}\right)^{e_j}\cdot\left(\frac{1}{U_{(j+1)}}+\Delta_{j+1}\right)\right)\left(\frac{I_k}{n}\right)^{e_k}\mathbf{1}_{A_n}\right] \\
&= \mathbb{E}\left[\prod_{j=0}^{k-1}\left(\left(\frac{I_j}{I_{j+1}}\right)^{e_j}\cdot\left(\frac{1}{U_{(j+1)}}\right)\right)\left(\frac{I_k}{n}\right)^{e_k}\mathbf{1}_{A_n}\right] + o(1).
\end{aligned}
$$

As $\mathbb{E}\left[|\Delta_{j+1}|\mathbf{1}_{\{U_{(0)}>1/n\}}\right] = O(\log n/n)$, it follows easily that the convergence rate in the last display is $O(\log n/n)$. Next, let $\Delta'_j = \frac{I_j}{I_{j+1}} - \frac{U_{(j)}}{U_{(j+1)}}$. Note that, for any positive real numbers x, y, we have

$$
\frac{-y}{(x+1)x} \leqslant \frac{\lceil y\rceil}{\lceil x\rceil} - \frac{y}{x} \leqslant \frac{1}{x},
$$

and thus, on A_n

$$
\Delta'_j \in [-(nU_{(j+1)})^{-1}, (nU_{(j+1)})^{-1}].
$$

Hence, by the mean value theorem, if $s \geqslant 1$, for $j \in \{0,\dots,k-1\}$, $\left|\left(\frac{I_j}{I_{j+1}}\right)^s - \left(\frac{U_{(j)}}{U_{(j+1)}}\right)^s\right| \leqslant s/(nU_{(j+1)})$. In the case that $s < 1$, observe that

$$
\min\left(\frac{I_j}{I_{j+1}}, \frac{U_{(j)}}{U_{(j+1)}}\right) \geqslant \frac{nU_{(j)}}{nU_{(j+1)}+1} \geqslant \frac{U_{(j)}}{2U_{(j+1)}},
$$

since $I_1 > 1$, and thus,

$$\max\left(\left(\frac{I_j}{I_{j+1}}\right)^{s-1}, \left(\frac{U_{(j)}}{U_{(j+1)}}\right)^{s-1}\right) \leqslant \left(\frac{U_{(j)}}{2U_{(j+1)}}\right)^{s-1} \leqslant \frac{2U_{(j+1)}}{U_{(j)}}.$$

Thus, by a similar application of the mean value theorem, if $0 \leqslant s \leqslant 1$, then,

$$\left|\left(\frac{I_j}{I_{j+1}}\right)^{s} - \left(\frac{U_{(j)}}{U_{(j+1)}}\right)^{s}\right| \leqslant 2s/(nU_{(j)}).$$

Now, for $j \in \{0, \ldots, k\}$, we have

$$\mathbb{E}\left[U_{(j)}^{-1}\prod_{i=0}^{k-1} U_{(i+1)}^{-1}\mathbf{1}_{A_n}\mathbf{1}_{\{I_0>1\}}\right] \leqslant \mathbb{E}\left[\prod_{i=0}^{k} U_i^{-1}\mathbf{1}_{\{U_i>n^{-i}\}}\right] = O(\log^{k+1}(n)).$$

Note that we only need $I_0 > 1$ when $s < 1$, in order to ensure that $U_{(0)} > 1/n$. Thus, successively replacing $\frac{I_j}{I_{j+1}}$ by $\frac{U_{(j)}}{U_{(j+1)}} + \Delta'_j$ shows

$$\begin{aligned}
&\mathbb{E}\left[\prod_{j=0}^{k-1}\left(\left(\frac{I_j}{I_{j+1}}\right)^{e_j}\cdot\left(\frac{1}{U_{(j+1)}}\right)\right)\left(\frac{I_k}{n}\right)^{e_k}\mathbf{1}_{A_n}\mathbf{1}_{\{I_0>1\}}\right] \\
&= \mathbb{E}\left[\prod_{j=0}^{k-1}\left(\frac{U_{(j)}}{U_{(j+1)}}\right)^{e_j}\cdot\prod_{j=0}^{k-1}\frac{1}{U_{(j+1)}}\left(\frac{I_k}{n}\right)^{e_k}\mathbf{1}_{A_n}\mathbf{1}_{\{I_0>1\}}\right] + O\left(\frac{\sum_{j=0}^{k-1} e_j \log^{k+1}(n)}{n}\right).
\end{aligned}$$

Replacing I_k/n by $U_{(k)}$ gives rise to an error term of order at most $e_k \log^{k+1}(n)/n$. As $\mathbb{P}(A_n^c) = O(1/n)$ and $\mathbb{P}(I_0 = 1) = O(1/n)$, an application of Hölder's inequality shows that we may drop the indicators $\mathbf{1}_{A_n}$ and $\mathbf{1}_{\{I_0>1\}}$ at the cost of an error term of order $n^{-1/(k+2)}$. □

2.4.3 Upper bound for the Mean of $N_k(n, B)/\ell n$

In the following subsections, unless otherwise specified, we let B denote an arbitrary element of the family $\mathscr{F}$ defined in (2.25). Let $N_{\eta,k}(n, B)$ be the number of vertices of degree $k\ell$ with weight in B that arrived after time ηn. Then, since $N_{\eta,k}(n, B) \leqslant N_k(n, B) \leqslant N_{\eta,k}(n, B) + \eta\ell n$, we have

$$\mathbb{E}\left[\left|\frac{N_{\eta,k}(n, B)}{\ell n} - \frac{N_k(n, B)}{\ell n}\right|\right] \leqslant \eta. \tag{2.27}$$

Thus, to obtain an upper bound for the convergence of the mean, it suffices to prove that

$$\limsup_{\eta\to 0}\limsup_{n\to\infty}\mathbb{E}\left[\frac{N_{\eta,k}(n,B)}{\ell n}\right] = p_k^{\alpha}(B).$$

In what follows, we use the notation $d_i(n)$ to denote the out-degree at time n of the vertex i born at time $i_0 := \lfloor i/\ell \rfloor$. We then have,

$$\mathbb{E}\left[N_{\eta,k}(n,B)\right] = \sum_{\eta n < i_0 \leqslant n-k} \ell \cdot \mathbb{P}\left(d_i(n) = k, W_i \in B\right),$$

since the probability is identical for each of the ℓ vertices born at each time i_0. In what follows, for a given i we denote by $\mathcal{I}_k := \{i_1, \ldots, i_k\}$ a collection of natural numbers $i_0 < i_1 < \ldots < i_k \leqslant n$. For ease of notation we exclude the dependence of $\mathcal{I}_k$ on i.

For a natural number $s > i_0$, we use the notation $i \sim s$ to denote that i is the vertex chosen at the sth time-step, hence i gains ℓ new neighbours at time s. Likewise, the notation $i \nsim s$ denotes that i is not chosen at the sth time-step. Then, let $\mathcal{E}_i(\mathcal{I}_k, B)$ denote the event that $W_i \in B$ and for all $s \in \{i_0+1, \ldots, n\}$, $i \sim s$ if and only if $s \in \mathcal{I}_k$. Clearly, we have

$$\mathbb{P}\left(d_i(n) = k, W_i \in B\right) = \sum_{\mathcal{I}_k \in \binom{\{i_0+1,\ldots,n\}}{k}} \mathbb{P}\left(\mathcal{E}_i(\mathcal{I}_k, B)\right).$$

where $\binom{\{i_0+1,\ldots,n\}}{k}$ denotes the set of all subsets of $\{i_0+1, \ldots, n\}$ of size k. For $\varepsilon > 0$ and $n \geqslant 0$ and natural numbers $N_1 \leqslant N_2$, we let

$$\mathcal{G}_\varepsilon(n) = \left\{|\mathcal{Z}_n - \alpha n| < \varepsilon\alpha n\right\}, \text{ and } \mathcal{G}_\varepsilon(N_1, N_2) = \bigcap_{t=N_1}^{N_2} \mathcal{G}_\varepsilon(n). \tag{2.28}$$

Moreover, for $n \geqslant 1$, we denote by $\mathscr{T}_n$ the σ-field generated by $(\mathcal{T}_s)_{1\leqslant s\leqslant n}$, containing all the information generated by the process up to time n. By the assumption of almost sure convergence and Egorov's theorem, for any $\delta, \varepsilon > 0$, there exists $N' = N'(\varepsilon, \delta)$ such that, for all $n \geqslant N'$, $\mathbb{P}\left(\mathcal{G}_\varepsilon(N', n)\right) \geqslant 1 - \delta$. Thus, for $n \geqslant N'/\eta$, we have

$$\begin{aligned}\mathbb{E}\left[N_{\eta,k}(n,B)\right] &\leqslant \mathbb{E}\left[N_{\eta,k}(n,B)\mathbf{1}_{\mathcal{G}_\varepsilon(N',n)}\right] + \ell n\left(1 - \mathbb{P}\left(\mathcal{G}_\varepsilon(N',n)\right)\right) \qquad (2.29)\\ &\leqslant \ell\left(\sum_{\eta n < i_0 \leqslant n}\ \sum_{\mathcal{I}_k \in \binom{\{i_0+1,\ldots,n\}}{k}} \mathbb{P}\left(\mathcal{E}_i(\mathcal{I}_k, B) \cap \mathcal{G}_\varepsilon(i_0, n)\right) + \delta n\right).\end{aligned}$$

We use the shorthand $\alpha_{\pm\varepsilon} := (1 \pm \varepsilon)\alpha$.

Proposition 2.4.7. *Let $B \in \mathscr{F}$ and $0 < \varepsilon, \eta \leqslant 1/2$. As $n \to \infty$, uniformly in $\eta n < i_0 \leqslant n-k, \mathcal{I}_k = \{i_1, \ldots, i_k\} \in \binom{\{i_0+1,\ldots,n\}}{k}$ and the choice of ε, we have*

$$\mathbb{P}\left(\mathcal{E}_i(\mathcal{I}_k, B) \cap \mathcal{G}_\varepsilon(i_0, n)\right) \leqslant (1 + O(1/n))\,\mathbb{E}\left[\left(\frac{i_k}{n}\right)^{f(k,W)/\alpha_{+\varepsilon}} \prod_{j=0}^{k-1} \left(\frac{i_j}{i_{j+1}}\right)^{f(j,W)/\alpha_{+\varepsilon}} \frac{f(j,W)}{\alpha_{-\varepsilon}(i_{j+1}-1)} \mathbf{1}_B(W)\right].$$

Corollary 2.4.8. *Let $B \in \mathscr{F}$ and $0 < \delta, \varepsilon, \eta \leqslant 1/2$. Then, there exists $N = N(\delta, \varepsilon, \eta)$ such that, for all $n \geqslant N$,*

$$\frac{\mathbb{E}\left[N_{\eta,k}(n,B)\right]}{\ell n} \leqslant (1+\delta)\left(\frac{1+\varepsilon}{1-\varepsilon}\right)^k \mathbb{E}\left[\frac{\alpha_{+\varepsilon}}{f(k,W)+\alpha_{+\varepsilon}} \prod_{j=0}^{k-1} \frac{f(j,W)}{f(j,W)+\alpha_{+\varepsilon}} \mathbf{1}_B(W)\right] + C\eta^{1/(k+2)} + \delta,$$

where the constant C may depend on k and B but not on n and not on the choices of $\delta, \varepsilon, \eta$. In particular, for each $B \in \mathscr{F}$ and $k \in \mathbb{N}_0$,

$$\limsup_{n\to\infty} \mathbb{E}\left[N_k(n,B)\right]/\ell n \leqslant p_k^\alpha(B).$$

Proof. This follows from applying (2.29) and Proposition 2.4.7 and then applying Corollary 2.4.6 with $e_j = f(j,W)/\alpha_{+\varepsilon}$ and $f_j = f(j,W)/\alpha_{-\varepsilon}$ to bound the sum over the collection of indices. Note that the term $\left(\frac{1+\varepsilon}{1-\varepsilon}\right)^k$ comes from replacing $\alpha_{-\varepsilon}$ by $\alpha_{+\varepsilon}$. □

We proceed towards the proof of Proposition 2.4.7. Let ε, η be given such that $0 < \varepsilon, \eta \leqslant 1/2$. For $\eta n < i_0 \leqslant n$ and $\mathcal{I}_k = \{i_1, \ldots, i_k\} \in \binom{\{i_0+1,\ldots,n\}}{k}$ for each $s \in \{i_0+1, \ldots, n\}$, we define

$$\mathcal{D}_s := \begin{cases} \{i \sim s\}, & \text{if } s \in \mathcal{I}_k, \\ \{i \nsim s\}, & \text{otherwise,} \end{cases}$$

and $\tilde{\mathcal{D}}_s = \mathcal{D}_s \cap \mathcal{G}_\varepsilon(s)$. We also define $\tilde{\mathcal{D}}_{i_0} = \mathcal{G}_\varepsilon(i_0) \cap \{W_i \in B\}$, and for simplicity of notation, write D_j and $\tilde{D}_j$ for the indicator random variables $\mathbf{1}_{\mathcal{D}_j}$ and $\mathbf{1}_{\tilde{\mathcal{D}}_j}$ respectively. Note that

$\mathcal{E}_i(\mathcal{I}_k, B) \cap \mathcal{G}_\varepsilon(i_0, n) = \bigcap_{j=i_0}^n \tilde{\mathcal{D}}_j$. To bound the probability of this event, we define

$$X_s = \mathbb{E}\left[\prod_{j=i_s+1}^{n} \tilde{D}_j \,\middle|\, \mathscr{T}_{i_s}\right] \tilde{D}_{i_s}, \quad s \in \{0, \ldots, k\}$$

and observe that $\mathbb{E}[X_0] = \mathbb{P}\left(\bigcap_{s=i_0}^n \tilde{\mathcal{D}}_s\right)$ is the sought after probability.

Lemma 2.4.9. *For $s \in \{0, \ldots, k\}$, we have*

$$X_s \leqslant \prod_{u=i_k+1}^{n}\left(1 - \frac{f(k,W)}{\alpha_{+\varepsilon}(u-1)}\right)\left(\prod_{j=s}^{k-1} \frac{f(j,W)}{\alpha_{-\varepsilon}(i_{j+1}-1)} \prod_{j'=i_j+1}^{i_{j+1}-1}\left(1 - \frac{f(j,W)}{\alpha_{+\varepsilon}(j'-1)}\right)\right)\tilde{D}_{i_s}, \quad (2.30)$$

where we interpret any empty products (for example when $i_k = n$) as equal to 1*. In particular,*

$$\mathbb{E}[X_0] \leqslant \mathbb{E}\left[\prod_{u=i_k+1}^{n}\left(1 - \frac{f(k,W)}{\alpha_{+\varepsilon}(u-1)}\right)\left(\prod_{j=0}^{k-1} \frac{f(j,W)}{\alpha_{-\varepsilon}(i_{j+1}-1)} \prod_{j'=i_j+1}^{i_{j+1}-1}\left(1 - \frac{f(j,W)}{\alpha_{+\varepsilon}(j'-1)}\right)\right)\mathbf{1}_B(W)\right]. \quad (2.31)$$

Proof. We prove (2.30) by backwards induction. For the base case, $s = k$, if $i_k = n$, the inequality is trivial, as $X_k = \tilde{D}_{i_k}$. Thus, assuming $i_k < n$, by the tower property,

$$\begin{aligned}
\mathbb{E}\left[\prod_{j=i_k+1}^{n} \tilde{D}_j \,\middle|\, \mathscr{T}_{i_k}\right] &= \mathbb{E}\left[\mathbb{E}\left[\tilde{D}_n \,\middle|\, \mathscr{T}_{n-1}\right] \prod_{j=i_k+1}^{n-1} \tilde{D}_j \,\middle|\, \mathscr{T}_{i_k}\right] \\
&\leqslant \mathbb{E}\left[\mathbb{E}\left[D_n \,\middle|\, \mathscr{T}_{n-1}\right] \prod_{j=i_k+1}^{n-1} \tilde{D}_j \,\middle|\, \mathscr{T}_{i_k}\right] \\
&= \mathbb{E}\left[\left(1 - \frac{f(k,W)}{\mathcal{Z}_{n-1}}\right) \prod_{j=i_k+1}^{n-1} \tilde{D}_j \,\middle|\, \mathscr{T}_{i_k}\right] \\
&\leqslant \left(1 - \frac{f(k,W)}{\alpha_{+\varepsilon}(n-1)}\right) \mathbb{E}\left[\prod_{j=i_k+1}^{n-1} \tilde{D}_j \,\middle|\, \mathscr{T}_{i_k}\right],
\end{aligned}$$

and iterating this argument with the conditional expectation on the right hand side proves the base case. Now, note that for $s \in \{0, \ldots, k-1\}$

$$X_s = \mathbb{E}\left[X_{s+1} \prod_{j=i_s+1}^{i_{s+1}-1} \tilde{D}_j \,\middle|\, \mathscr{T}_{i_s}\right] \tilde{D}_{i_s}.$$

Applying the induction hypothesis, it suffices to bound the term $\mathbb{E}\left[\prod_{j=i_s+1}^{i_{s+1}} \tilde{D}_j \;\middle|\; \mathscr{T}_{i_s}\right]$, and, similar to the base case, we may assume $i_s < i_{s+1} - 1$. But, then, we have

$$\begin{aligned}
\mathbb{E}\left[\prod_{j=i_s+1}^{i_{s+1}} \tilde{D}_j \;\middle|\; \mathscr{T}_{i_s}\right] &= \mathbb{E}\left[\mathbb{E}\left[\tilde{D}_{i_{s+1}} \;\middle|\; \mathscr{T}_{i_{s+1}-1}\right] \prod_{j=i_s+1}^{i_{s+1}-2} \tilde{D}_j \;\middle|\; \mathscr{T}_{i_s}\right] \\
&\leqslant \mathbb{E}\left[\mathbb{E}\left[D_{i_{s+1}} \;\middle|\; \mathscr{T}_{i_{s+1}-1}\right] \prod_{j=i_s+1}^{i_{s+1}-2} \tilde{D}_j \;\middle|\; \mathscr{T}_{i_s}\right] \\
&\leqslant \frac{f(s,W)}{\alpha_{-\varepsilon}(i_{s+1}-1)} \mathbb{E}\left[\prod_{j=i_s+1}^{i_{s+1}-2} \tilde{D}_j \;\middle|\; \mathscr{T}_{i_s}\right] \\
&\leqslant \frac{f(s,W)}{\alpha_{-\varepsilon}(i_{s+1}-1)} \mathbb{E}\left[\mathbb{E}\left[D_{i_{s+1}-1} \;\middle|\; \mathscr{T}_{i_{s+1}-1}\right] \prod_{j=i_s+1}^{i_{s+1}-2} \tilde{D}_j \;\middle|\; \mathscr{T}_{i_s}\right] \\
&\leqslant \frac{f(s,W)}{\alpha_{-\varepsilon}(i_{s+1}-1)} \left(1 - \frac{f(s,W)}{\alpha_{+\varepsilon}\left(i_{s+1}-2\right)}\right) \mathbb{E}\left[\prod_{j=i_s+1}^{i_{s+1}-2} \tilde{D}_j \;\middle|\; \mathscr{T}_{i_s}\right].
\end{aligned}$$

Iterating these bounds the inductive step follows in a similar manner to the base case. Finally, noting that $\mathbf{1}_{\tilde{\mathcal{D}}_i} \leqslant \mathbf{1}_B(W)$ proves (2.31). □

The next lemma follows from a simple application of Stirling's formula, i.e., (2.20):

Lemma 2.4.10. *Let $\eta, C > 0$. Then, uniformly over $\eta n \leqslant a \leqslant b$ and $0 \leqslant \beta \leqslant C$, we have*

$$\prod_{j=a+1}^{b-1} \left(1 - \frac{\beta}{j-1}\right) = \left(\frac{a}{b}\right)^{\beta} \left(1 + O\left(\frac{1}{n}\right)\right).$$

□

Proof of Proposition 2.4.7. We take the upper bound $\mathbb{E}[X_0]$ from Lemma 2.4.9 and bound each of the products by applying Lemma 2.4.10. □

2.4.4 Deducing Convergence of the Mean of $N_k(n, B)/\ell n$

In this subsection we deduce a lower bound on $\liminf_{n\to\infty} \mathbb{E}[N_k(n,B)]/\ell n$ on measurable sets $B \in \mathscr{F}$. In what follows, denote by $N_{\geqslant M}(n, B)$ the number of vertices of out-degree

$\geqslant \ell M$ with weight belonging to B. Moreover, let $N(n, B) = N_{\geqslant 0}(n, B)$ denote the total number of vertices at time n with weight belonging to B.

Lemma 2.4.11. *For any measurable set B, we have,* $\limsup_{n\to\infty} \frac{N_{\geqslant M}(n,B)}{\ell n} \leqslant \frac{1}{M}$ *almost surely.*

Proof. Since we add ℓ vertices at each time-step, we have $\limsup_{n\to\infty} \frac{|\mathcal{T}_n|}{\ell n} = 1$. However, $|\mathcal{T}_n| \geqslant M N_{\geqslant M}(n, \mathbb{R})$, since the right-side only provides a lower bound for the number of vertices in the tree incident to those with out-degree at least M. The result follows by dividing both sides by $M\ell n$ and sending n to infinity. □

Proof of Proposition 2.4.2

Proof of Proposition 2.4.2. Recall that Corollary 2.4.8 showed that for each $B \in \mathscr{F}$ and $k \in \mathbb{N}_0$,

$$\limsup_{n\to\infty} \mathbb{E}\left[N_k(n, B)\right]/\ell n \leqslant p_k^{\alpha}(B).$$

Now, suppose that Proposition 2.4.2 fails, so that, in particular there exists some set $B' \in \mathscr{F}$ and $k' \in \mathbb{N}_0$ such that

$$\liminf_{n\to\infty} \frac{\mathbb{E}\left[N_{k'}(n, B')\right]}{\ell n} < p_{k'}^{\alpha}(B').$$

Thus, for some $\epsilon' > 0$, we have $\liminf_{n\to\infty} \frac{\mathbb{E}[N_{k'}(n,B')]}{\ell n} \leqslant p_{k'}^{\alpha}(B) - \epsilon'$. Now, using Lemma 2.4.11, choose $M > \max\left\{k', \frac{2}{\epsilon'}\right\}$, so that $\limsup_{n\to\infty} \frac{N_{\geqslant M}(n,B')}{\ell n} < \epsilon'/2$. Then, recalling Lemma 2.2.3,

$$\begin{aligned} \liminf_{n\to\infty} \mathbb{E}\left[\sum_{k=0}^{M} \frac{N_k(n, B')}{\ell n}\right] &\leqslant \liminf_{n\to\infty} \mathbb{E}\left[\frac{N_{k'}(n, B')}{\ell n}\right] + \sum_{k\neq k'} \limsup_{n\to\infty} \mathbb{E}\left[\frac{N_k(n, B')}{\ell n}\right] \\ &\leqslant \left(\sum_{k=0}^{\infty} p_k^{\alpha}(B')\right) - \epsilon' \leqslant \mu(B') - \epsilon'. \end{aligned} \tag{2.32}$$

On the other hand, by Fatou's Lemma, we have

$$\liminf_{n\to\infty} \mathbb{E}\left[\sum_{k=0}^{M} \frac{N_k(n,B')}{\ell n}\right] \geqslant \mathbb{E}\left[\liminf_{n\to\infty} \sum_{k=0}^{M} \frac{N_k(n,B')}{\ell n}\right] \tag{2.33}$$
$$= \mathbb{E}\left[\liminf_{n\to\infty}\left(\frac{N(n,B')}{\ell n} - \frac{N_{\geqslant M}(n,B')}{\ell n}\right)\right] \geqslant \mu(B') - \epsilon'/2,$$

where the last inequality follows from the strong law of large numbers. But then, combining (2.32) and (2.33), we have $\mu(B') - \epsilon' \geqslant \mu(B') - \epsilon'/2$, a contradiction. □

2.4.5 Second Moment Calculations

In order to bound the second moment, we apply similar calculations to the start of the section to compute asymptotically the number of pairs of vertices of out-degree $k\ell$ born after time ηn. For vertices i and j, as in Section 2.4.3, we set $i_0 := \lfloor i/\ell \rfloor$ and $j_0 := \lfloor j/\ell \rfloor$, and note that

$$\mathbb{E}\left[(N_{\eta,k}(n,B))^2\right] = \sum_{\eta n < i_0, j_0 \leqslant n-k} \sum_{j : \lfloor j/\ell \rfloor = j_0} \sum_{i : \lfloor i/\ell \rfloor = i_0} \mathbb{P}\left(d_i(n) = k, W_i \in B, d_j(n) = k, W_j \in B\right). \tag{2.34}$$

Note that, in a similar manner to (2.27), we have

$$\mathbb{E}\left[\left|\frac{(N_{\eta,k}(n,B))^2}{\ell^2 n^2} - \frac{(N_k(n,B))^2}{\ell^2 n^2}\right|\right] \leqslant \eta$$

so that it suffices to prove that

$$\limsup_{\eta\to 0} \limsup_{n\to\infty} \mathbb{E}\left[\frac{(N_{\eta,k}(n,B))^2}{\ell^2 n^2}\right] \leqslant (p_k^{\alpha}(B))^2.$$

Recall that, for a given i we denote by $\mathcal{I}_k$ a collection of natural numbers $i_0 < i_1 < \cdots < i_k \leqslant n$. Moreover, for a given j, we denote by $\mathcal{J}_k$ a collection of natural numbers $j_0 < j_1 < \cdots < j_k \leqslant n$. Similar to Section 2.4.3, for $s > j$ we use the notation $j \sim s$ to denote that j is the vertex chosen at the sth time-step and likewise, we let $\mathcal{E}_j(\mathcal{J}_k, B)$ denote the event that $W_j \in B$ and for all $s \in \{j_0 + 1, \ldots, n\}$, $j \sim s$ if and only if $s \in \mathcal{J}_k$. Then we have

$$\begin{aligned}&\mathbb{P}\left(d_i(n) = k, W_i \in B, d_j(n) = k, W_j \in B\right)\\ &\qquad = \sum_{\mathcal{J}_k \in \binom{\{j_0+1,\dots,n\}}{k}} \sum_{\mathcal{I}_k \in \binom{\{i_0+1,\dots,n\}}{k}} \mathbb{P}\left(\mathcal{E}_i(\mathcal{I}_k, B) \cap \mathcal{E}_j(\mathcal{J}_k, B)\right).\end{aligned}$$

Note that the contribution to the above sum corresponding to terms with $\mathcal{I}_k \cap \mathcal{J}_k \neq \varnothing$, and $i \neq j$, is zero, since it is impossible for distinct vertices to be chosen in a single time-step. But then, the terms corresponding to $i = j$ contribute at most $\mathbb{E}\left[N_{\eta,k}(n, B)\right] \leqslant \ell n$ to the right side of (2.34). Next, for any choice of indices with $\mathcal{I}_k \cap \mathcal{J}_k = \varnothing$, there are at most ℓ^2 pairs of vertices (i, j) born at respective times (i_0, j_0) contributing to the sum in (2.34). Recalling the definitions of $\mathcal{G}_\varepsilon(n), \mathcal{G}_\varepsilon(N_1, N_2)$ and $N' = N'(\varepsilon, \delta)$ from (2.28) and below in the previous subsection, in a similar manner to (2.29) we have, for $n \geqslant N'/\eta$,

$$\begin{aligned}&\mathbb{E}\left[(N_{\eta,k}(n, B))^2\right]\\ &\quad \leqslant \ell^2 \left(\sum_{\eta n < i_0, j_0 \leqslant n-k} \sum_{\mathcal{I}_k \cap \mathcal{J}_k = \varnothing} \mathbb{P}\left(\mathcal{E}_i(\mathcal{I}_k, B) \cap \mathcal{E}_j(\mathcal{J}_k, B) \cap \mathcal{G}_\varepsilon(i_0, n)\right) + \delta n^2 \right) + \ell n. \quad (2.35)\end{aligned}$$

We then have the following:

Proposition 2.4.12. *Let $B \in \mathscr{F}$ and $0 < \varepsilon, \eta \leqslant 1/2$. As $n \to \infty$, uniformly in $\eta n < i_0 \leqslant j_0 \leqslant n - k$ and $\mathcal{I}_k \in \binom{\{i_0+1,\dots,n\}}{k}$, $\mathcal{J}_k \in \binom{\{j_0+1,\dots,n\}}{k}$ such that $\mathcal{I}_k \cap \mathcal{J}_k = \varnothing$, and the choice of ε, we have*

$$\begin{aligned}&\mathbb{P}\left(\mathcal{E}_i(\mathcal{I}_k, B) \cap \mathcal{E}_j(\mathcal{J}_k, B) \cap \mathcal{G}_\varepsilon(i_0, n)\right)\\ &\qquad \leqslant (1 + O(1/n))\mathbb{E}\left[\left(\frac{i_k}{n}\right)^{f(k,W)/\alpha_{+\varepsilon}} \cdot \prod_{s=0}^{k-1} \left(\left(\frac{i_s}{i_{s+1}}\right)^{f(s,W)/\alpha_{+\varepsilon}} \frac{f(s, W)}{\alpha_{-\varepsilon}(i_{s+1} - 1)} \right) \mathbf{1}_B(W)\right]\\ &\qquad\quad \times \mathbb{E}\left[\left(\frac{j_k}{n}\right)^{f(k,W)/\alpha_{+\varepsilon}} \cdot \prod_{s=0}^{k-1} \left(\left(\frac{j_s}{j_{s+1}}\right)^{f(s,W)/\alpha_{+\varepsilon}} \frac{f(s, W)}{\alpha_{-\varepsilon}(j_{s+1} - 1)} \right) \mathbf{1}_B(W)\right].\end{aligned} \quad (2.36)$$

We leave the details of the proof of this proposition to the reader, as it follows an analogous approach to the proof of Proposition 2.4.7, using a backwards induction argument.

Proof Sketch. Let $u_1, \ldots, u_{2k}$ denote the indices in $\mathcal{I}_k \cup \mathcal{J}_k$, and $f_x(i), f_x(j)$ denote the fitnesses associated with vertex i and vertex j at time x. Then, when we bound the probabilities $\{i \not\to x\} \cap \{j \not\to x\}$ for all $x \in \{u_s + 1, \ldots, u_{s+1} - 1\}$ from above we obtain terms of the form

$$\prod_{x=u_s+1}^{u_{s+1}-1} \left(1 - \frac{f_x(i) + f_x(j)}{\alpha_{+\varepsilon}(x-1)}\right) = \left(\frac{u_s}{u_{s+1}}\right)^{f_x(i)+f_x(j)} \left(1 + O\left(\frac{1}{n}\right)\right),$$

where the right side follows from Lemma 2.4.10. Then, when we evaluate the expectation analogous to the expectation appearing in (2.31), we obtain an expectation involving products of terms dependent on W_i and W_j, i.e., the weights associated with vertex i and vertex j. These terms separate into a product of expectations by the independence of the random variables W_i, W_j, and finally, many of the products telescope to yield the right side of (2.36). $\square$

Proof of Proposition 2.4.3

Proof. We apply Proposition 2.4.12 to bound the summands in (2.35). Then, as we are looking for an upper bound, we may drop the condition $\mathcal{I}_k \cap \mathcal{J}_k = \varnothing$ when evaluating the sum. But then, by Corollary 2.4.6, we have, uniformly in ε and η,

$$\begin{aligned}
&\sum_{\eta n < i_0, j_0 \leqslant n} \sum_{\mathcal{I}_k, \mathcal{J}_k} \mathbb{E}\left[\left(\frac{i_k}{n}\right)^{f(k,W)/\alpha_{+\varepsilon}} \cdot \prod_{s=0}^{k-1} \left(\frac{i_s}{i_{s+1}}\right)^{f(s,W)/\alpha_{+\varepsilon}} \frac{f(s,W)}{\alpha_{-\varepsilon}(i_{s+1}-1)} \mathbf{1}_B(W)\right] \\
&\qquad \times \mathbb{E}\left[\left(\frac{j_k}{n}\right)^{f(k,W)/\alpha_{+\varepsilon}} \cdot \prod_{s=0}^{k-1} \left(\frac{j_s}{j_{s+1}}\right)^{f(s,W)/\alpha_{+\varepsilon}} \frac{f(s,W)}{\alpha_{-\varepsilon}(j_{s+1}-1)} \mathbf{1}_B(W)\right] \\
&\leqslant \left(\frac{1+\varepsilon}{1-\varepsilon}\right)^{2k} \left(\mathbb{E}\left[\frac{\alpha_{+\varepsilon}}{f(k,W)+\alpha_{+\varepsilon}} \prod_{s=0}^{k-1} \frac{f(s,W)}{f(s,W)+\alpha_{+\varepsilon}} \mathbf{1}_B(W)\right]\right)^2 + O\left(n^{-1/(k+2)}\right) + C'\eta^{1/k+2},
\end{aligned}$$

for some universal constant $C' > 0$, depending only on B, f. The result follows. $\square$

Chapter Three

Preferential Attachment Trees with Neighbourhood Influence

3.1 Introduction

In Section 2.3 of Chapter 2, we saw that the particular case of the (μ, f, ℓ)-RIF tree when f is affine displays many interesting properties, including the condensation phenomenon, proved in Section 2.3.2. This motivates our study of the 'higher dimensional' analogue of this model, the PANI-tree, as described in Section 1.3.3 of Chapter 1. Note that in this chapter, for brevity, we only consider the case that 1 vertex arrives at each time-step, corresponding to the case that $\ell = 1$ in the GPAF-tree. However, the description of the model, and analogues of the statements we prove may readily be generalised to the case that $\ell > 1$ using the same techniques. We first briefly recall the dynamics of this model, but, for a more precise description, encourage the reader to refer back to Section 1.3.3 of Chapter 1.

Recall that in this model, at each time-step n a vertex v is selected with probability proportional to its fitness $f(\mathcal{N}^+(v, \mathcal{T}_n))$, which is a function of the weights of the vertices in the out-neighbourhood of v. In this model, we define f such that

$$f(\mathcal{N}^+(v, \mathcal{T}_n)) := h(W_v) + \sum_{(v,u)\in\mathcal{T}_n} g(W_v, W_u), \tag{3.1}$$

where $h : [0, w^*] \to [0, \infty)$ and $g : [0, w^*] \times [0, w^*] \to [0, \infty)$ are bounded and measurable. A newcomer, $n + 1$ then arrives, with its own independent weight $W_{n+1} \in [0, w^*]$ sampled independently from the weight distribution μ, and the directed edge $(v, n + 1)$ is added to $\mathcal{T}_n$ to form $\mathcal{T}_{n+1}$.

Dynamics of the PANI**-Tree**

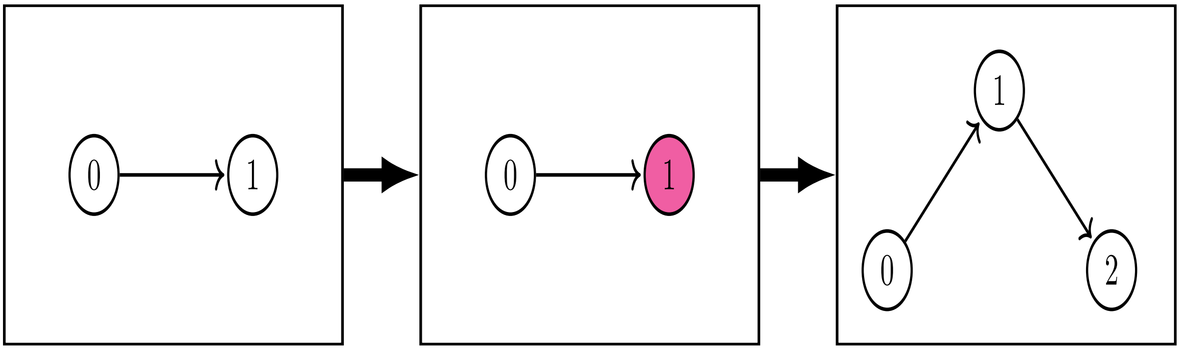


Figure 3.1: A sample transition from $\mathcal{T}_1$ to $\mathcal{T}_2$. In $\mathcal{T}_1$, 0 is chosen with probability proportional to $f(N^+(0, \mathcal{T}_1)) = h(W_0) + g(W_0, W_1)$, while 1 is chosen with probability proportional to $f(N^+(1, \mathcal{T}_1)) = h(W_1)$. In this evolution, 1 is chosen, so the newcomer 2 arrives as an out-neighbour of 1.

Remark 3.1.1. *One may interpret $(\mathcal{T}_n)_{n\in\mathbb{N}_0}$ in the context of reinforced branching processes as follows: we begin with an* individual 0 *belonging to its own* family *that reproduces after an exponentially distributed amount of time, with parameter $h(W_0)$. We say that the* ancestral weight *of the family is W_0. Then, recursively, when a birth event occurs in the ith family, with ancestral weight W_i, a new individual with random weight W joins the ith family, reproducing after an* $\mathrm{Exp}(g(W_i, W))$ *distributed amount of time, where* $\mathrm{Exp}(g(W_i, W))$ *denotes* *the exponential distribution with parameter* $g(W_i, W)$*; and simultaneously, an individual of weight* W *begins its own family, with ancestral weight* W*. The out-neighbourhood of a vertex* i *in the tree* $\mathcal{T}_n$*, including the vertex* i *itself, then represents individuals in the* i*th family in the branching process, at the time of the* n*th birth event.*

Remark 3.1.2. *One can extend the model from the previous remark further by supplanting it with constants* $0 \leqslant \beta, \gamma \leqslant 1$*, so that when a birth event occurs, independently with probability* β*, an individual with random weight* W *joins the* i*th family, and with probability* γ*, an individual with random weight* W' *(also sampled from* μ*) initiates its own family with ancestral weight* W'*. While not immediately clear from the way we have defined the model, our methods also extend to this case - this link becomes clearer when viewing individuals as "loops" and "edges" in a Pólya urn similar to Urn E see Figure 3.2 in Section 3.2.1 below). In this extended model, the case* $g(x, y) = h(x) = x$*, and this terminology, was introduced in [29], as a stochastic analogue of the model of Kingman [51].*

3.1.1 Statements of Main Results

The results in this chapter depend on two sets of conditions. One set of conditions describes the 'non-condensation' regime, which one might interpret as the analogue of Condition **C1** with regards to the GPAF-tree analysed in Section 2.3.1 of Chapter 2, whilst the other describes the 'condensation' regime which one might interpret as an analogue of the condensation phenomenon analysed in Section 2.3.2 of Chapter 2. Note that, with regards to the GPAF-tree we also studied a third phenomenon when Condition **C1** fails in Section 2.3.3 of Chapter 2: degenerate degrees. We expect a similar phenomenon to be generalised to the PANI-tree, but do not pursue this in this chapter.

In order to emphasise the connection between the PANI-tree and the (μ, f, ℓ) - RIF tree of the previous chapter, we incorporate some of the same notation: the Condition **C1** appearing below may be interpreted as an analogue of the Condition **C1** defined in Chapter 2. However, one should not similarly interpret Condition **C2** appearing below as an analogue of **C2** as these conditions are very different.

The Non-Condensation Regime of the PANI-tree

The first main conditions are the following: recalling g and h as defined in (3.1), assume

C1 There exists some $\lambda^* > \tilde{g}^*$ such that

$$\mathbb{E}\left[\frac{h(W)}{\lambda^* - \tilde{g}(W)}\right] = 1,$$

where $\tilde{g}(x) := \mathbb{E}\left[g(x, W)\right]$ and $\tilde{g}^* := \mathbb{E}\left[\sup_{x \in [0,w^*]} g(x, W)\right]$. We call λ^* the *Malthusian parameter* of the process.

C2 For some $J > 0, N \in \mathbb{N}$, there exist measurable functions $\phi_j^{(i)} : [0, w^*] \to [0, J]$, $j = 1, 2$, $i \in [N]$, and a bounded continuous function $\kappa : [0, J]^{2N} \to \mathbb{R}_+$ such that

$$g(x, y) = \kappa\left(\phi_1^{(1)}(x), \ldots, \phi_1^{(N)}(x), \phi_2^{(1)}(y), \ldots, \phi_2^{(N)}(y)\right).$$

Remark 3.1.3. *We expect similar results under the weaker hypothesis that g and h are measurable and bounded rather than Condition **C2**. However, this condition still allows many "reasonable" choices of bounded measurable functions g. This includes the* GPAF*-tree of Section 2.3, Chapter 2, the case where g is continuous, as well as functions of the form $g(x, y) = \alpha\phi_1(x) + \beta\phi_2(y)$ or $g(x, y) = \phi_1(x)\phi_2(y)$, where ϕ_1, ϕ_2 are bounded and measurable and $\alpha, \beta \geqslant 0$.*

Our first theorem concerns the partition function of the process,

Theorem 3.1.1. *Assume Conditions **C1** and **C2**. Then we have*

$$\lim_{n\to\infty} \frac{\mathcal{Z}_n}{n} \to \lambda^*$$

almost surely, where $\mathcal{Z}_n$ and λ^* respectively denote the partition function and Malthusian parameter of the process.

Recall from Section 1.4.2 in Chapter 1 that in the PANI-tree we also study a higher dimensional analogue of the edge distribution $\Xi(n, \cdot)$: given a product, Borel measurable set A, the quantity $\Xi^{(2)}(n, A)$ denotes the number of edges (v, v') in the tree $\mathcal{T}_n$ such that $(W_v, W_{v'}) \in A$, that is,

$$\Xi^{(2)}(n, A) := \sum_{(v,v') \in \mathcal{T}_n} \mathbf{1}_A(W_v, W_{v'}).$$

Under this notation, we have $\Xi(n, B) = \Xi^{(2)}(n, B \times [0, w^*])$ almost surely. Also, define $\psi(x) := h(x)/(\lambda^* - \tilde{g}(x))$, and denote by $\psi_* \mu$ the pushforward measure of μ under ψ - i.e. the measure such that for any measurable set A

$$(\psi_* \mu)(A) = \mathbb{E}\left[\frac{h(W)}{\lambda^* - \tilde{g}(W)} \mathbf{1}_A(W)\right].$$

Theorem 3.1.2. *Assume Conditions **C1** and **C2**. Then, with $\Xi^{(2)}(n, \cdot)$ as defined in* (1.6), *we have*

$$\frac{\Xi^{(2)}(n, \cdot)}{n} \to (\psi_* \mu \times \mu)(\cdot),$$

almost surely, in the sense of weak convergence. Here $\psi_* \mu \times \mu$ denotes the product measure of $\psi_* \mu$ and μ on $[0, w^*]^2$ equipped with the Borel sigma algebra.

We include the proofs of Theorem 3.1.1 and Theorem 3.1.2 in Section 3.2.2 and Section 3.2.2. We also prove theorems related to the degree distribution. In view of Section 1.4.1 of Chapter 1, in order to describe this result, we first describe a *companion process*

$(S_i(w))_{i\geqslant 0}$ that describes the evolution of the *fitness* of a vertex with weight w as its neighbourhood changes. First, let $W_1, W_2, \ldots$ be independent μ-distributed random variables and let $w \in [0, w^*]$. We then define the random process $(S_i(w))_{i\geqslant 0}$ inductively so that

$$S_0(w) := h(w); \quad S_{i+1}(w) := S_i(w) + g(w, W_{i+1}),\ i \geqslant 0. \tag{3.2}$$

In the following theorem, $\mathbb{E}[\cdot]$ denotes expectation with respect to the path of $S_i(W_0)$, i.e., expectations with respect to the product measure involving the terms $W_0, W_1, W_2, \ldots, W_{k-1}$. Also recall that $N_{\geqslant k}(n, B)$ denotes the number of vertices of out-degree at least k in the tree $\mathcal{T}_n$ with weight belonging to B.

We then have the following theorem:

Theorem 3.1.3. *Assume Conditions **C1** and **C2**. Then, for any measurable set $B \subseteq [0, w^*]$, we have*

$$\lim_{n\to\infty} \frac{N_{\geqslant k}(n, B)}{n} = \mathbb{E}\left[\prod_{i=0}^{k-1}\left(\frac{S_i(W_0)}{S_i(W_0) + \lambda^*}\right)\mathbf{1}_B(W_0)\right], \tag{3.3}$$

almost surely.

A particular consequence of Theorem 3.1.3 is that, for any measurable set B, almost surely, we have

$$\frac{N_k(n, B)}{n} \to p_k^{\lambda^*}(B).$$

where $p_k^{\lambda^*}(\cdot)$ is the quantity described in (1.4) of Section 1.4.1, Chapter 1. We prove Theorem 3.1.3 in Section 3.2.3.

Remark 3.1.4. *One may interpret the right hand side of* (3.3) *as the probability of a sequence of at least k consecutive heads before a first tail when, sampling W_0 at random, and flipping the ith coin heads with probability proportional to $S_{i-1}(W_0)$.*

In a manner analogous to the end of Section 2.2.1 in Chapter 2, Theorem 3.1.3 allows us to deduce, for any measurable set B, almost sure convergence of the quantity $\Xi(n, B)/n$. First we require the following lemma, which may be of independent interest:

Lemma 3.1.4. *Let $(S_i(w))_{i \geqslant 0}$ denote the process defined in (3.2) in terms of bounded, measurable functions g, h, suppose $\tilde{g}(x) := \mathbb{E}\left[g(x, W)\right]$ and $\tilde{g}_+ = \sup_{x \in [0,w^*]} \tilde{g}(x)$. Then, for any $w \in [0, w^*]$, and $\lambda \geqslant \tilde{g}_+$ we have*

$$\sum_{k=1}^{\infty} \mathbb{E}\left[\prod_{i=0}^{k-1}\left(\frac{S_i(w)}{S_i(w)+\lambda}\right)\right] = \frac{h(w)}{\lambda - \tilde{g}(w)}, \tag{3.4}$$

where the right hand side is infinite if $g(w) = \tilde{g}_+$ and $\lambda = \tilde{g}_+ = \tilde{g}(w)$. In particular,

$$\sum_{k=1}^{\infty} \mathbb{E}\left[\prod_{i=0}^{k-1}\left(\frac{S_i(W_0)}{S_i(W_0)+\lambda}\right)\mathbf{1}_B(W_0)\right] = \mathbb{E}\left[\frac{h(W_0)}{\lambda - \tilde{g}(W_0)}\mathbf{1}_B(W_0)\right].$$

As the proof of this lemma detracts from the main techniques used in this chapter, we delay its proof to the end of the chapter, in Section 3.4.1.

Remark 3.1.5. *One may interpret (3.4) as a generalisation of the classic geometric series formula: if we set $g(x, y) \equiv 0$, and $q := h(w)/(h(w) + \lambda)$, the left hand side of (3.4) is $\sum_{i=1}^{\infty} q^i = \frac{h(w)}{\lambda} = \frac{q}{1-q}$. Indeed, as Remark 3.1.4 shows, one may interpret the left hand side as the expected value of a generalised geometrically distributed random variable.*

Lemma 3.1.4 allows us to strengthen the weak convergence result of Theorem 3.1.2. One may interpret this result as an analogue of Theorem 2.2.2 from Chapter 2, indeed the proof of this theorem is almost identical to the proof of Theorem 2.2.2.

Theorem 3.1.5. *Assume Condition **C1**. Then, for any measurable set $A \subseteq [0, w^*]$ we have*

$$\frac{\Xi(n, A)}{n} \to (\psi_* \mu)(A),$$

almost surely.

Remark 3.1.6. *Lemma 3.1.4 shows that the limiting measure $(\psi_*\mu)(\cdot)$ is the same as the quantity $m(\lambda^*, \cdot)$, where $m(\lambda^*, \cdot)$ is the quantity described in (1.9) of Section 1.4.1, Chapter 1.*

Remark 3.1.7. *As the limiting measure appearing in Theorem 2.3.1 is absolutely continuous with respect to μ, and hence almost surely with respect to the measures $\Xi(n, \cdot)$, one might* *expect to improve this convergence to almost sure convergence in the total variation norm. Indeed, in the simplified model first analysed by Kingman in [51] the convergence takes place in the total variation norm (in this context, however, the sequence of measures he considered was deterministic). Note that Kingman described the non-condensation regime as the "democratic" regime.*

The Condensation Regime of the PANI-tree

In this chapter we undertake a more nuanced investigation into the *condensation* phenomenon in the GPAF-tree, from Section 2.3.2 of Chapter 2. We first make a more precise definition of what *condensation* means.

Definition 3.1.6. *Suppose we are given a μ-null set $S \subseteq [0, w^*]$. We say that* condensation *occurs around the set S, if for some nested collection of sets $(S_\varepsilon)_{\varepsilon \geqslant 0}$,* [1] *with $S_\varepsilon \downarrow S$ as $\varepsilon \to 0$ we have*

$$\lim_{\varepsilon \to 0} \lim_{n \to \infty} \frac{\Xi(n, S_\varepsilon)}{n} > 0,$$

with positive probability.

Remark 3.1.8. *Informally, condensation means that, in the limit of the random measure $\Xi(n, \cdot)/n$, the set S acquires more mass than one 'would expect'. Indeed, if we swap limits,*

$$\lim_{n \to \infty} \lim_{\varepsilon \to 0} \frac{\Xi(n, S_\varepsilon)}{n} = \lim_{n \to \infty} \frac{\Xi(n, S)}{n} = 0,$$

almost surely, since $\mu(S) = 0$.

Our main assumptions are now as follows:

D1 We have

$$\mathbb{E}\left[\frac{h(W)}{\tilde{g}^* - \tilde{g}(W)}\right] < 1. \tag{3.5}$$

[1] That is, a collection of sets such that if $\varepsilon_1 < \varepsilon_2$, $S_{\varepsilon_1} \subseteq S_{\varepsilon_2}$.

D2 The function g satisfies Condition **C2**.

D3 There exists a (maximal) set of points $\mathcal{M} \subseteq \mathrm{Supp}\,(\mu)$, such that, for any $x^* \in \mathcal{M}$,

$$\max_{p \in [0,w^*]} g(p, W) = g(x^*, W) \quad \mathbb{P} - \text{a.s.}$$

We denote by x^* a generic point in $\mathcal{M}$.

D4 For all $\varepsilon > 0$ sufficiently small, and a measurable function $u_\varepsilon : [0, w^*] \to \mathbb{R}_+$ with $\lim_{\varepsilon \to 0} u_\varepsilon = 0$ pointwise, for $x^* \in \mathcal{M}$, we have

$$\begin{aligned} \mathcal{M}_\varepsilon &:= \{x : \mathbb{P}\left(g(x^*, W) - g(x, W) < u_\varepsilon(W)\right) = 1\} \\ &= \{x : \mathbb{P}\left(g(x^*, W) - g(x, W) < u_\varepsilon(W)\right) > 0\}\,. \end{aligned} \tag{3.6}$$

Under this assumption, we have $\mu(\mathcal{M}_\varepsilon) > 0$.

Remark 3.1.9. *Note that, by the measurability of $g(\cdot, q)$ for any $q \in [0, w^*]$, the function*

$$p \mapsto \operatorname{ess\,sup}_{q \in [0,w^*]} \{g(x^*, q) - g(p, q) - u_\varepsilon(q)\}$$

is also measurable - see, e.g. [17, Theorem 4.7.1.]. This ensures that the set $\mathcal{M}_\varepsilon$ is measurable.

Example 3.1.10. *In the case that $g(x, y) = \phi_1(x)\phi_2(y)$ for bounded, measurable ϕ_1, ϕ_2, if $\phi_1(x)$ is maximised on a set $\mathcal{M}$ and $\phi_2(y) > 0$ μ-a.e., for $\varepsilon > 0$ and $x^* \in \mathcal{M}$ we may take $u_\varepsilon = \varepsilon \cdot \phi_2$ and*

$$\mathcal{M}_\varepsilon := \{x : \phi_1(x^*)\phi_2(W) - \phi_1(x)\phi_2(W) < \varepsilon\phi_2(W)\} = \{x : \phi_1(x^*) - \phi_1(x) < \varepsilon\}\,.$$

A condition that guarantees that this set has positive measure is assuming continuity of ϕ_1 at some point $x^* \in \mathcal{M}$, as this implies that $\mathcal{M}_\varepsilon$ is a neighbourhood of x^*.

Remark 3.1.11. *Conditions **D1** and **D2** may be interpreted as analogues of Conditions **C1** and **C2** in the condensation regime. One may regard $\mathcal{M}$ from **D3** as a "dominating set",* *in the sense that $\mathbb{P}$-a.s., upon arrival of a new vertex into its neighbourhood, the change of the fitness of any vertex is at most the change of the fitness of a vertex with weight with weight in $\mathcal{M}$. Condition **D4** ensures that this "dominating property" is captured by sets $\mathcal{M}_\varepsilon$ of positive measure.*

Indeed the right hand side of (3.6) *implies that the change of the fitness of any vertex with weight in $\mathcal{M}_\varepsilon^c$ is at most the change of the fitness of a vertex having weight in $\mathcal{M}_\varepsilon$. Note that $\mathcal{M}_\varepsilon \downarrow \mathcal{M}$ as $\varepsilon \to 0$. This accounts for the formation of the condensate in Theorem 3.1.7 below, since $\tilde{g}$ is maximised on $\mathcal{M}$, by **D1** it must be the case that $\mu(\mathcal{M}) = 0$.*

The following theorem may be viewed as an analogue of Theorem 2.3.1 from Chapter 2.

Theorem 3.1.7. *Assume Conditions **D1-D4**. Then,*

- *We have $\lim_{n\to\infty} \frac{\mathcal{Z}_n}{n} \to \tilde{g}^* = g(x^*)$, almost surely.*
- *For any measurable set $A \subseteq [0, w^*]$ such that, for $\varepsilon > 0$ sufficiently small $A \cap \mathcal{M}_\varepsilon = \varnothing$, we have*
$$\frac{\Xi(n, A)}{n} \to (\psi_* \mu)(A), \quad \textit{almost surely.} \tag{3.7}$$
In addition,
$$\lim_{\varepsilon\to 0} \lim_{n\to\infty} \frac{\Xi(n, \mathcal{M}_\varepsilon)}{n} = 1 - (\psi_*\mu)([0, w^*]) > 0, \tag{3.8}$$
so that condensation occurs around $\mathcal{M}$.
- *For any measurable set B, almost surely, we have*
$$\lim_{n\to\infty} \frac{N_{\geqslant k}(n, B)}{n} = \mathbb{E}\left[\prod_{i=0}^{k-1} \left(\frac{S_i(W)}{S_i(W) + \tilde{g}^*}\right) \mathbf{1}_B(W)\right].$$

Remark 3.1.12. *As the condensation occurs around the "dominating set" $\mathcal{M}$, in the context of reinforced branching processes as described in Remark 3.1.1 and Remark 3.1.2, one may interpret this is families with maximum reinforced 'fitness' acquiring a positive proportion of* *individuals in the population in the limit. In this context, 'fitness' refers to the ability of an individual to produce offspring quickly. This has an interesting interpretation in the context of evolution.*

We have the following corollary:

Corollary 3.1.8. *Assume Conditions **D1-D4**, and the sets $\mathcal{M}_\varepsilon$ in **D4** are such that $\overline{\mathcal{M}_\varepsilon} \downarrow \mathcal{M}$ as $\varepsilon \to 0$, recalling that $\overline{\mathcal{M}_\varepsilon}$ denotes the topological closure of $\mathcal{M}_\varepsilon$. Also, suppose that $\mathcal{M} = \{x^*\}$, and define the measure $\Pi(\cdot)$ such that, for any measurable set $B \subseteq [0, w^*]$*

$$\Pi(B) = (\psi_*\mu)(B) + (1 - (\psi_*\mu)([0, w^*]))\, \delta_{x^*}(B).$$

Then,

$$\frac{\Xi(n, \cdot)}{n} \to \Pi(\cdot) \quad \textit{almost surely},$$

in the sense of weak convergence.

Example 3.1.13. *In the case that $g(x, y) = \phi_1(x)\phi_2(y)$ for a bounded, continuous function ϕ_1 and bounded measurable function ϕ_2, if $\phi_1(x)$ is maximised at a unique point x^* and $\phi_2(y) > 0$ μ-a.e., we may take u_ε and $\mathcal{M}_\varepsilon$ as defined in Example 3.1.10. Indeed, in this case*

$$\overline{\mathcal{M}_\varepsilon} = \{x : \phi_1(x^*) - \phi_1(x) \leqslant \varepsilon\},$$

so that $\overline{\mathcal{M}_\varepsilon} \downarrow \{x^*\}$ as $\varepsilon \to 0$.

3.1.2 An Informal Discussion of the Main Results

In this subsection, we provide an informal discussion of some of the implications of our main results.

Averaged Power-Law Degrees in the PANI-tree

First note that by Theorem 3.1.3, almost surely

$$\lim_{n\to\infty}\frac{N_k(n,B)}{n} = p_k^{\lambda^*}(B) = \mathbb{E}\left[\frac{\lambda^*}{S_k(W)+\lambda^*}\prod_{i=0}^{k-1}\left(\frac{S_i(W)}{S_i(W)+\lambda^*}\right)\mathbf{1}_B(W)\right].$$

Now by the strong law of large numbers, one would expect, at least asymptotically, $S_i(W) \sim h(W) + i\tilde{g}(W)$, and thus it is natural to expect

$$\lim_{n\to\infty}\frac{N_k(n,B)}{n} \sim \mathbb{E}\left[\frac{\lambda^*}{k\tilde{g}(W)+\lambda^*}\prod_{i=0}^{k-1}\left(\frac{h(W)+i\tilde{g}(W)}{h(W)+i\tilde{g}(W)+\lambda^*}\right)\right].$$

We therefore expect the degrees in this model to behave asymptotically like the GPAF-tree analysed in Section 2.3 of Chapter 2, with $\ell = 1$ and associated functions h and $\tilde{g}$. Recall that in Section 2.3.1 of Chapter 2, we showed that on any measurable set B where $\tilde{g}$ and h are bounded

$$\mathbb{E}\left[\frac{\lambda^*}{k\tilde{g}(W)+\lambda^*}\prod_{i=0}^{k-1}\left(\frac{h(W)+i\tilde{g}(W)}{h(W)+i\tilde{g}(W)+\lambda^*}\right)\right] = \mathbb{E}\left[c_B k^{-(1+\lambda^*/\tilde{g}(W))}\mathbf{1}_B(W)\right],$$

where c_B depends on g and h but not k. Thus, informally, like the GPAF-tree, the PANI-tree displays a degree distribution that satisfies an 'averaged' power law that depends on the distribution μ. Noting also that $\lambda^*/\tilde{g}(W) > 1$, the exponent of this power law is larger than 2. A similar analysis can be applied to the condensation regime by applying Theorem 3.1.7.

The Growth of the Neighbourhood of Fixed Vertex in the PANI-tree

In the following proposition, we let $f_n(v) = f(N^+(v,\mathcal{T}_n))$ denote the fitness, as defined in (3.1), of a vertex labelled $v \in \mathbb{N}_0$, with weight w_v in the tree at time n. In addition, let $(R_i)_{i\geqslant v}$ denote the filtration generated by the tree process $(\mathcal{T}_i)_{i\geqslant v}$. Next, set

$$M_n(v) := \frac{f_n(v)}{\prod_{s=v}^{n-1}\left(\frac{\mathcal{Z}_s+\tilde{g}(w_v)}{\mathcal{Z}_s}\right)}.$$

Proposition 3.1.9. *For any vertex v, $(M_n(v))_{n \geqslant v}$ is a martingale with respect to the filtration $(R_i)_{i \geqslant v}$.*

Proof. Using the definition of the process, for $n \geqslant v$ we compute

$$\begin{aligned} \mathbb{E}\left[f_{n+1}(v) | R_n\right] &= \frac{f_n(v)}{\mathcal{Z}_n}\left(f_n(v) + \tilde{g}(w_v)\right) + \left(1 - \frac{f_n(v)}{\mathcal{Z}_n}\right) f_n(v) \\ &= f_n(v)\left(\frac{\mathcal{Z}_n + \tilde{g}(w_v)}{\mathcal{Z}_n}\right). \end{aligned}$$

The result follows from the definition of $(M_n(v))_{n \geqslant v}$. □

Now, here we note two things: first, if $\deg_t^+(v)$ denotes the out-degree of vertex v at time n, then we expect $f_n(v) \sim \deg_n^+(v)$. In fact, by applying Wald's lemma, one can show $\mathbb{E}\left[f_n(v)\right] = h(w_v) + \mathbb{E}\left[\deg_n^+(v)\right] \tilde{g}(w_v)$. Second, by Theorems 3.1.1 and 3.1.7, we expect $\mathcal{Z}_i \sim \lambda^* i$ and $\tilde{g}^* i$ in the non-condensation and condensation regimes respectively. Thus, we expect

$$\deg_n^+(v) \sim \prod_{s=v}^{n-1}\left(\frac{\mathcal{Z}_s + \tilde{g}(w_v)}{\mathcal{Z}_s}\right) \sim \begin{cases} n^{\tilde{g}(w_v)/\lambda^*}, & \text{under Conditions } \mathbf{C1} \text{ and } \mathbf{C2}; \\ n^{\tilde{g}(w_v)/\tilde{g}^*}, & \text{under Conditions } \mathbf{D1\text{-}D4}. \end{cases}$$

Therefore, in the non-condensation regime, we expect each individual vertex to grow like $n^{\tilde{g}(w_v)/\lambda^*} \leqslant n^{\tilde{g}^*/\lambda^*} < n$, whereas, in the condensation regime, vertices with weight w_v such that $g(w_v)$ is closer and closer to $\tilde{g}^*$ grow at a rate closer and closer to linearity with respect to the size of the network. Note that to turn this argument into a rigorous result in terms of $\mathbb{E}\left[\deg_n^+(v)\right]$, one requires L^1 convergence of the martingale in Proposition 3.1.9.

3.1.3 Overview and Techniques

Overview of this Chapter

In Section 3.2 we prove results about the model related to the non-condensation regime. We first review some background theory about *Pólya urns* in Section 3.2.1, and then, the results of Section 3.2.2 are used in order to prove Theorem 3.1.1 and Theorem 3.1.2 in Section 3.2.2 and Section 3.2.2 respectively. Next, the results of Section 3.2.3 are used to prove Theorem 3.1.3 and Theorem 3.1.5 in Section 3.2.3 and Section 3.2.3. In Section 3.3 we extend these results to the condensation regime, proving Theorem 3.1.7 and Corollary 3.1.8 Section 3.3.1 and Section 3.3.2 and respectively. Finally, we prove Lemma 3.1.4 in Section 3.4.1.

Techniques used in this Chapter

The results in this chapter generalise the techniques used in [20] for the study of the Bianconi-Barabási model, using a *Pólya urn approximation.* However, the generalisation of this model to bounded measurable functions h, functions g satisfying Condition **C2**, and the possibility of arbitrary weight distributions lead to technical challenges, somewhat analogous to those arising from using a measure-theoretic approach to integration as opposed to the Riemann integral. Applying this approach to studying the degree distribution in the case of uncountably supported weight distributions also appears to be novel. The couplings used in the Pólya urn approximation, Proposition 3.2.6 and Proposition 3.2.12 and the coupling used to extend the results to the condensation regime, Lemma 3.3.2, are closely related to that used in Lemma 2.3.2 in Chapter 2, and thus we encourage the reader to quickly review the latter coupling before reading the rest of this chapter.

One might imagine that many of the results here may follow easily from an application of the theory of Crump-Mode-Jagers branching processes, for example as in Section 2.2 of

Chapter 2. However, the dependence between the point processes associated with a parent and its offspring means that the classic theory is not immediately applicable. This in turn raises the question of whether one can develop a theory of C-M-J branching processes with *dependencies* between the point-processes associated with individuals.

3.2 The Non-Condensation Regime

3.2.1 A Brief Introduction to Generalised Pólya Urns

Generalised Pólya urns are a well studied family of stochastic processes representing the composition of an *urn* containing balls with certain *types*. If $\mathscr{T}$ denotes the set of possible types, associated to a ball of type $t \in \mathscr{T}$ is a non-negative *activity* $\mathbf{a}(t)$, which depends on the type. The process then evolves in discrete time so that, at each time-step, a ball of type t is sampled at random from the urn with probability proportional to its activity $\mathbf{a}(t)$, and replaced with balls of a number of different types according to a possibly random *replacement rule*.

In the case that $\mathscr{T}$ is finite, the configuration of the urn after n replacements may be represented as a *composition vector* $(X_n)_{n\in\mathbb{N}_0}$ with entries labelled by type, and the activities encoded in an *activity vector* $\mathbf{a}$. In this vector, the ith entry corresponds to the number of balls of type $i \in \mathscr{T}$. Let $(\xi_{ij})_{i,j\in\mathscr{T}}$ be the matrix whose ijth component denotes the random number of balls of type j added, if a ball of type i is drawn, and (following the notation of Janson in [45]) define the matrix A such that $A_{ij} := a_j\mathbb{E}\left[\xi_{ji}\right]$. The expected evolution of the urn in the $(n+1)$st step, may therefore be obtained by applying the matrix A to the composition vector X_n. A type $i \in \mathscr{T}$ is said to be *dominating* if, for any $j \in \mathscr{T}$, it is possible to obtain a ball of type j starting with a ball of type i. If we write $i \sim j$ for the equivalence relation where $i \sim j$ if it is possible to obtain j starting from a ball of type i, and vice versa. This partitions the types into equivalence *classes*. A class $\mathscr{C} \subseteq \mathscr{T}$ is *dominating* if, for every $i \in \mathscr{C}$, i is dominating. Moreover, the eigenvalues of A may be obtained by the restriction of A to its classes; we say an eigenvalue belongs to a *dominating class* if it is an eigenvalue of the restriction of A to this class. Finally, we say that the urn, or the matrix A, is *irreducible* if there is only one dominating class. Note the difference when compared to irreducible matrices in the context of Markov chains: here it is possible for diagonal entries to be negative. Now, assume the following conditions are satisfied:

(A1) For all $i, j \in \mathscr{T}$, $\xi_{ij} \geqslant 0$ if $i \neq j$ and $\xi_{ii} \geqslant -1$.

(A2) For all $i, j \in \mathscr{T}$, $\mathbb{E}\left[\xi_{ij}^2\right] < \infty$.

(A3) The largest real eigenvalue λ_1 of A is positive.

(A4) The largest real eigenvalue λ_1 is simple.

(A5) We start with at least one ball of a dominating type.

(A6) λ_1 belongs to the dominating class.

The following is a well known result of Janson from 2004 building on previous work by by Athreya and Karlin (for example, [6, Proposition 2] and [5, Theorem 5]):

Theorem 3.2.1 ([45, Theorem 3.16]). *Assume Conditions (A1)-(A6), and suppose that v_1 denotes the right eigenvector, corresponding to the leading eigenvalue λ_1 of A, normalised so that $\mathbf{a}^T v_1 = 1$. Then, we have*

$$\frac{X_n}{n} \xrightarrow{n\to\infty} \lambda_1 v_1,$$

almost surely, conditional on essential non-extinction, i.e., non-extinction of balls of dominating type.

In addition, the following lemma by Janson provides convenient criteria for satisfying (A1)-(A6):

Lemma 3.2.2 ([45, Lemma 2.1])**.** *If A is irreducible, (A1) and (A2) hold, $\sum_{j\in\mathscr{T}} \mathbb{E}[\xi_{ij}] \geqslant 0$ for all $i \in \mathscr{T}$, with the inequality being strict for some $i \in \mathscr{T}$, then (A1) - (A6) are satisfied and essential extinction does not occur.*

We will not only analyse the PANI-tree using generalised Pólya urns, but also the dynamical model of random simplicial complexes, in Section 4.3 of Chapter 4.

Analysing the PANI-tree using Pólya Urns

The idea behind analysing the distribution of edges with a given weight, and the degree distribution in this model, is to consider two different types of Pólya urns, which we call *Urn E* and *Urn D* respectively. We illustrate the evolution of both these urns below. Recall, Figure 3.1 illustrates a possible evolution of a step of the process $(\mathcal{T}_i)_{i\in\mathbb{N}_0}$; Figures 3.2 and 3.3 illustrate the corresponding steps in Urn E and Urn D.

In Urn E, we consider a generalised Pólya urn with balls of two types: singletons x, and tuples (x, y), corresponding to 'edges' and 'loops'. A ball of type (x, y) has activity $g(x, y)$ and a ball of type x has activity $h(x)$. At each step, if a ball of type given by either x or (x, y) is selected, we introduce two new balls, of which one has random type W, and the other has type (x, W). In relation to the evolving tree, this corresponds to the event that a vertex of weight x has been sampled in the subsequent step.

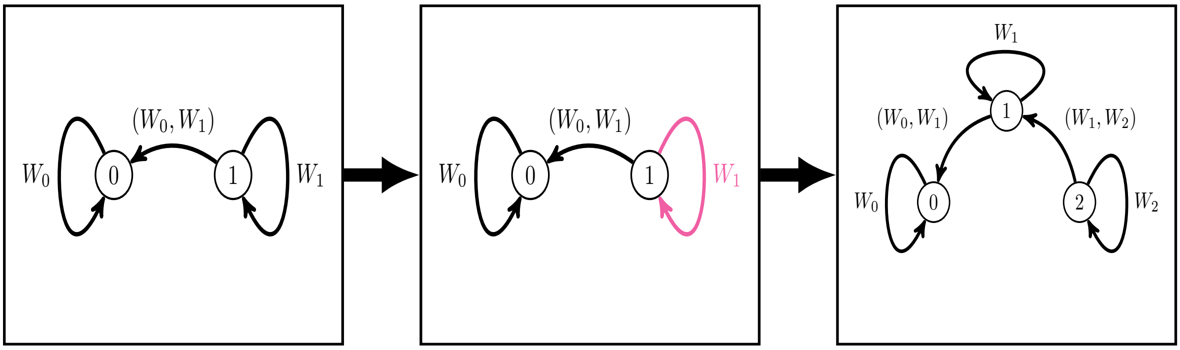


Figure 3.2: The evolution of the tree from $\mathcal{T}_1$ to $\mathcal{T}_2$ from Figure 3.1 viewed as a transition in Urn E. The event vertex 1 is selected may be interpreted as the event that the 'loop' W_1 is selected in the Pólya urn - and thus the arrival of the vertex 2 corresponds to the arrival of the 'loop' W_2 and the 'edge' (W_1, W_2) in the Pólya urn.

In Urn D, we consider a generalised Pólya urn with balls of types corresponding to tuples of varying lengths. A ball of type $(x_0, \ldots, x_k)$ has activity $h(x_0) + \sum_{i=1}^{k} g(x_0, x_i)$, and at each step, if a ball this type is selected, we remove it and introduce two new balls: one of random type W, and one of type $(x_0, \ldots, x_k, W)$. In relation to the evolving tree, this corresponds to the event that a vertex v of weight x_0 has been sampled when proceeding to the subsequent step, with neighbours of v listed in order of arrival having weights $x_1, \ldots, x_k$.

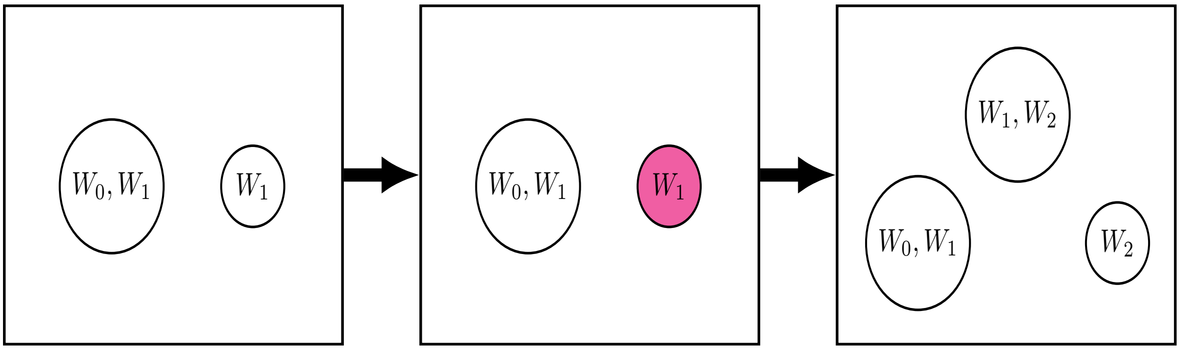


Figure 3.3: The evolution of the tree from $\mathcal{T}_1$ to $\mathcal{T}_2$ from Figure 3.1 viewed as a transition in Urn D. The event vertex 1 is selected may be interpreted as the event that the ball W_1 is selected in the Pólya urn - and thus the arrival of the vertex 2 corresponds to the addition of the balls W_2 and (W_1, W_2). The latter ball represents the addition of vertex 2 into the neighbourhood of vertex 1.

Note that, in the manner we have described Urns E and D, the set of possible types may be infinite: the measure μ may have infinite support so that W may take on infinite values, and the neighbourhoods of vertices (in Urn D) may be infinite. Whilst there is some theory related to infinite type Pólya urns within the framework of measure-valued Pólya processes (see, for example, [59]), these results are often non-trivial to apply in practice, as we will see in Section 4.3 of Chapter 4. We instead opt for a different approach by approximating these infinite urns with urns of finitely many types - enough to approximate the sigma algebras generated by $W, g(W, W')$ and $h(W)$, where W, W' are i.i.d random variables sampled according to μ. In Section 3.2.2 we apply this analysis to Urn E, and in Section 3.2.3 we apply it to Urn D. We first introduce some extra notation specific to this section.

Some More Notation and Terminology used in this Section

Recall from Section 1.3.1 of Chapter 2, that for a natural number $N \in \mathbb{N}$, we denote by $[N]$ the set $\{1, \ldots, N\}$. In order to apply the finite Pólya urn theory, given a set of types $\mathscr{T}$, we denote by $\mathbb{V}_{\mathscr{T}}$ the *free vector space* over the field $\mathbb{R}$ generated by $\mathscr{T}$, i.e., the vector space where vectors are indexed by the elements of $\mathscr{T}$. We will generally view an urn with types $\mathscr{T}$ as a stochastic process taking values in $\mathbb{V}_{\mathscr{T}}$. In addition we will generally identify vectors $\mathbf{v} \in \mathbb{V}_{\mathscr{T}}$ interchangeably with functions $\mathbf{v} : \mathscr{T} \to \mathbb{R}$. Thus, for $x \in \mathscr{T}$, $\mathbf{v}(x)$ denotes the entry of the vector corresponding to x, and for $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{V}_{\mathcal{B}}$, we have $(\mathbf{v}_1\mathbf{v}_2)(x) = \mathbf{v}_1(x)\mathbf{v}_2(x)$. For $x \in \mathscr{T}$, we define $\delta_x \in \mathbb{V}_{\mathscr{T}}$ such $\delta_x(y) = 1$ if $y = x$ and 0 otherwise.

For a Borel measurable set $S \subseteq \mathbb{R}$, and a finite set $\mathcal{A}$ of Borel measurable subsets of S, we say that $\mathcal{A} = \{A_1, \ldots, A_s\}$ forms a *good partition* of S if, given any two nonempty sets $A_i, A_j \in \mathcal{A}$, $A_i \cap A_j \neq \varnothing \implies A_i = A_j$, and $\bigcup_{i=1}^{s} A_i = S$. Note that, given two good partitions $\mathcal{A}_1, \mathcal{A}_2$ of S, the set

$$\{A_1 \cap A_2 : A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2\} \tag{3.9}$$

also forms a good partition of S. In addition, if $\mathcal{A}$ is a good partition of S, we say that $\mathcal{A}'$ forms a *refined good partition* of $\mathcal{A}$, if, for any $A' \in \mathcal{A}'$ there exists $A \in \mathcal{A}$ such that $A' \subseteq A$. Often, we will simply write *refined partition* for a *refined good partition.* The following lemma, which is well-known, justifies the use of the word 'refined'.

Lemma 3.2.3. *Suppose $\mathcal{A}$ is a good partition of a set S, and $\mathcal{A}'$ is a refined partition of $\mathcal{A}$. Then, for any set $A \in \mathcal{A}$, there exist sets $X_1, \ldots, X_s \in \mathcal{A}'$ such that $A = \bigcup_{i=1}^{s} X_i$. In particular, $\{X_i\}_{i\in[s]}$ forms a good partition of A.*

Proof. For $A \in \mathcal{A}$, define the sub-family $\mathcal{X} := \{A' \in \mathcal{A}' : A' \subseteq A\}$. Suppose $U := (\bigcup_{X\in\mathcal{X}} X) \neq A$. Then, there exists $x \in A \backslash U$, and since $\mathcal{A}'$ partitions S, $x \in V'$, for some set $V' \in \mathcal{A}'$ with $V' \nsubseteq A$. But then, since $\mathcal{A}'$ is a refined partition of $\mathcal{A}$, $V' \subseteq V$ for some $V \in \mathcal{A}$. But then, this implies that either $V \cap A \neq \varnothing$, contradicting the fact that $\mathcal{A}$ is a good partition of S, or $V = A$, contradicting the fact that $V' \nsubseteq A$. □

3.2.2 Analysing the PANI-tree by Coupling with Urn E

In this subsection we will refer to Conditions **C1** and **C2**. We will analyse the process under these conditions by coupling the tree process $(\mathcal{T}_n)_{n\in\mathbb{N}_0}$ with Pólya urn processes, parameterised by $m \in \mathbb{N}$. These may be interpreted as finite approximations of Urn E. Now, for each $x \in \mathbb{R}$ and $m \in \mathbb{N}$ we define a good partition of the interval $[0, x]$ into 2^m intervals, i.e., a *dyadic partition.* Set

$$\mathcal{D}_1^m(x) := [0, 2^{-m}x], \quad \text{and} \quad \mathcal{D}_i^m(x) := ((i-1)\cdot 2^{-m}x, i\cdot 2^{-m}x],\ i \in [2^m]\backslash\{1\}.$$

For $i \in [2^m]$, we also denote the closure of $\mathcal{D}_i^m(x)$ by $\overline{\mathcal{D}}_i^m(x)$, so that

$$\overline{\mathcal{D}}_i^m(x) = [(i-1)\cdot 2^{-m}x, i\cdot 2^{-m}x].$$

Supposing $h : [0, w^*] \to \mathbb{R}_+$ takes values in $[0, h_{\max}]$, and recalling the functions $\phi_1^{(j)}, \phi_2^{(j)}, j \in [N]$ from Condition **C2**, for each $i \in [2^m]$, $j \in [N]$ and $k \in [2]$, we set

$$\mathcal{H}_i^m := h^{-1}\left(\mathcal{D}_i^m(h_{\max})\right) \quad \text{and } \Phi_k^m(i,j) := \left(\phi_k^{(j)}\right)^{-1}\left(\mathcal{D}_i^m(J)\right).$$

By the measurability assumptions on the functions $\phi_k^{(j)}$ and h, for each $i \in [2^m]$, $j \in [N]$ and $k \in [2]$, the sets $\mathcal{H}_i^m$ and $\Phi_i^m(j,k)$ are measurable, and thus, the collections of sets $\{\mathcal{H}_i^m\}_{i\in[2^m]}$ and $\{\Phi_k^m(i,j)\}_{i\in[2^m]}$ form good partitions of $[0, w^*]$. We now split the latter family of sets to form a refined partition: for $\mathbf{i} = (i_1, \dots, i_N)$, $\mathbf{j} = (j_1, \dots, j_N) \in [2^m]^N$, if we set

$$\begin{aligned} \Phi_1^m(\mathbf{i}) &= \Phi_1^m(i_1,1) \cap \Phi_1^m(i_2,2) \cap \cdots \cap \Phi_1^m(i_N,N) \quad \text{and,} \\ \Phi_2^m(\mathbf{j}) &= \Phi_2^m(j_1,1) \cap \Phi_2^m(j_2,2) \cap \cdots \cap \Phi_2^m(j_N,N), \end{aligned} \tag{3.10}$$

by iteratively applying (3.9), the families of sets $\{\Phi_1^m(\mathbf{i})\}_{\mathbf{i}\in[2^m]^N}$ and $\{\Phi_2^m(\mathbf{j})\}_{\mathbf{j}\in[2^m]^N}$ also form good partitions of $[0, w^*]$. Now, given $\mathbf{v} = (v_1, \dots, v_N) \in [2^m]^N$, set

$$\overline{\mathcal{D}}_{\mathbf{v}}^m(J) := \overline{\mathcal{D}}_{v_1}^m(J) \times \overline{\mathcal{D}}_{v_2}^m(J) \times \cdots \times \overline{\mathcal{D}}_{v_N}^m(J),$$

and observe that, given $\mathbf{i}, \mathbf{j} \in [2^m]^N$, the construction of the sets in (3.10) are such that $(x, y) \in \Phi_1^m(\mathbf{i}) \times \Phi_2^m(\mathbf{j})$ implies that

$$\left(\phi_1^{(1)}(x), \dots, \phi_1^{(N)}(x), \phi_2^{(1)}(y), \dots, \phi_2^{(N)}(y)\right) \in \overline{\mathcal{D}}_{\mathbf{i}}^m(J) \times \overline{\mathcal{D}}_{\mathbf{j}}^m(J)$$

Now, recalling the function $\kappa : [0, J]^{2N} \to [0, g_{\max}]$ from Condition **C2**, for each $\mathbf{i}, \mathbf{j} \in [2^m]^N$, by continuity on the compact set $\overline{\mathcal{D}}_{\mathbf{i}}^m(J) \times \overline{\mathcal{D}}_{\mathbf{j}}^m(J)$, for $(x, y) \in \Phi_1^m(\mathbf{i}) \times \Phi_2^m(\mathbf{j})$ we have

$$\begin{aligned} \kappa\left(\phi_1^{(1)}(x), \dots, \phi_1^{(N)}(x), \phi_2^{(1)}(y), \dots, \phi_2^{(N)}(y)\right) &\geqslant \inf_{\mathbf{u},\mathbf{v}\in\overline{\mathcal{D}}_{\mathbf{i}}^m(J)\times\overline{\mathcal{D}}_{\mathbf{j}}^m(J)} \{\kappa(\mathbf{u}, \mathbf{v})\} \\ &= \min_{\mathbf{u},\mathbf{v}\in\overline{\mathcal{D}}_{\mathbf{i}}^m(J)\times\overline{\mathcal{D}}_{\mathbf{j}}^m(J)} \{\kappa(\mathbf{u}, \mathbf{v})\} =: \kappa^-(\mathbf{i}, \mathbf{j}), \quad (3.11) \end{aligned}$$

and likewise,

$$\begin{aligned} \kappa\left(\phi_1^{(1)}(x), \dots, \phi_1^{(N)}(x), \phi_2^{(1)}(y), \dots, \phi_2^{(N)}(y)\right) &\leqslant \sup_{\mathbf{u},\mathbf{v}\in\overline{\mathcal{D}}_{\mathbf{i}}^m(J)\times\overline{\mathcal{D}}_{\mathbf{j}}^m(J)} \{\kappa(\mathbf{u}, \mathbf{v})\} \\ &= \max_{\mathbf{u},\mathbf{v}\in\overline{\mathcal{D}}_{\mathbf{i}}^m(J)\times\overline{\mathcal{D}}_{\mathbf{j}}^m(J)} \{\kappa(\mathbf{u}, \mathbf{v})\} =: \kappa^+(\mathbf{i}, \mathbf{j}). \quad (3.12) \end{aligned}$$

Now, set

$$g^-(x, y) := \sum_{\mathbf{i},\mathbf{j}\in[2^m]^N} \kappa^-(\mathbf{i}, \mathbf{j}) \mathbf{1}_{\Phi_1^m(\mathbf{i})\times\Phi_2^m(\mathbf{j})}(x, y), \quad g^+(x, y) := \sum_{\mathbf{i},\mathbf{j}\in[2^m]^N} \kappa^+(\mathbf{i}, \mathbf{j}) \mathbf{1}_{\Phi_1^m(\mathbf{i})\times\Phi_2^m(\mathbf{j})}(x, y);$$

and

$$h^-(x) := \sum_{i=1}^{2^m} (i - 1) \cdot 2^{-m} h_{\max} \mathbf{1}_{\mathcal{H}_i}(x), \quad h^+(x) := \sum_{i=1}^{2^m} i \cdot 2^{-m} h_{\max} \mathbf{1}_{\mathcal{H}_i}(x).$$

One should interpret these functions as *lower* and *upper* approximations to g and h, indeed, by construction, we now have the following lemma:

Lemma 3.2.4. *We have $g^- \uparrow g$, $h^- \uparrow h$, $g^+ \downarrow g$ and $h^+ \downarrow h$ uniformly, as $m \to \infty$.*

Proof. We prove the statements regarding h^- and g^-; the others follow analogously (in the case of g^+ using (3.12) instead of (3.11)). Since the sets $(\mathcal{H}_i^m)_{i\in[2^m]}$ form a good partition of $[0, w^*]$, for each $m \in \mathbb{N}$, given $x \in [0, w^*]$, we have $x \in \mathcal{H}_j^m$ for some $j \in [2^m]$, and thus

$$h^-(x) = (j-1) \cdot 2^{-m} h_{\max} \leqslant h(x) \leqslant h^-(x) + 2^{-m} h_{\max}.$$

The convergence result for h^- follows. Now, note that by uniform continuity of κ on the compact set $[0, J]^{2N}$, for $\varepsilon > 0$, let M be sufficiently large so that for all $\mathbf{u}, \mathbf{v} \in [0, J]^{2N}$

$$\|\mathbf{u} - \mathbf{v}\| < \sqrt{2N} \cdot 2^{-M} J \qquad \Longrightarrow \qquad |\kappa(\mathbf{u}) - \kappa(\mathbf{v})| < \varepsilon.$$

Now, for any $m > M$, given $(x, y) \in [0, w^*] \times [0, w^*]$, there exists a unique set $\Phi_1^m(\mathbf{i}) \times \Phi_2^m(\mathbf{j})$ containing (x, y), which implies that

$$\left(\phi_1^{(1)}(x), \ldots, \phi_1^{(N)}(x), \phi_2^{(1)}(y), \ldots, \phi_2^{(N)}(y)\right) \in \overline{\mathcal{D}}_{\mathbf{i}}^m(J) \times \overline{\mathcal{D}}_{\mathbf{j}}^m(J).$$

Thus, for each $j \in [N]$, combining this equation with the definition of $\kappa^-(\mathbf{i}, \mathbf{j})$ from (3.11), we have

$$\kappa^-(\mathbf{i}, \mathbf{j}) \leqslant \kappa\left(\phi_1^{(1)}(x), \ldots, \phi_1^{(N)}(x), \phi_2^{(1)}(y), \ldots, \phi_2^{(N)}(y)\right) \leqslant \kappa^-(\mathbf{i}, \mathbf{j}) + \varepsilon,$$

and thus

$$g^-(x, y) \leqslant g(x, y) \leqslant g^-(x, y) + \varepsilon.$$

The result now follows. □

Now, using the good partitions $\{\mathcal{H}_i^m\}_{i\in[2^m]}$, $\{\Phi_1^m(\mathbf{i})\}_{\mathbf{i}\in[2^m]^N}$, $\{\Phi_2^m(\mathbf{j})\}_{\mathbf{j}\in[2^m]^N}$ and $\{\mathcal{D}_i^m(w^*)\}_{i\in[2^m]}$, we will form an even more refined partition, which we will use as the "building blocks" of the evolution of the Pólya urn approximations. For each m, define the good partition $\mathscr{I}^m$ such that

$$\mathscr{I}^m := \left\{I \subseteq [0, w^*] : I = \mathcal{H}_p^m \cap \mathcal{D}_q^m(w^*) \cap \Phi_1^m(\mathbf{i}) \cap \Phi_2^m(\mathbf{j}),\ p, q \in [2^m], \mathbf{i}, \mathbf{j} \in [2^m]^N\right\}. \quad (3.13)$$

Intuitively, this family of sets is such that the finite σ-algebra $\sigma(\mathscr{I}^m)$, is "fine enough" to approximate the Borel sigma algebra on $[0, w^*]$, and also capture the behaviour of g and h. Observe that, for $m_1 < m_2$, $\mathscr{I}^{m_2}$ is a refined partition of $\mathscr{I}^{m_1}$.

Suppose $|\mathscr{I}^m| = D_m$; then we label the sets in $\mathscr{I}^m$ arbitrarily as $(\mathcal{I}_i^m)_{i \in [D_m]}$. Now, for each $(x, y) \in \mathcal{I}_i^m \times \mathcal{I}_j^m$, $g^-(x, y)$ and $g^+(x, y)$ are constant, depending only on (i, j), and likewise, for each $x \in \mathcal{I}_\ell^m$, $h^-(x)$ and $h^+(x)$ are constant, depending on ℓ. Motivated by this, for each $(i, j) \in [D_m] \times [D_m]$, we define the following quantities:

$$g_{\min}(i, j) := g^-(x, y), \quad g_{\max}(i, j) := g^+(x, y), \quad (x, y) \in \mathcal{I}_i^m \times \mathcal{I}_j^m,$$

and likewise, for each $\ell \in [D_m]$, we define

$$h_{\min}(\ell) := h^-(x), \quad h_{\max}(\ell) := h^+(x), \quad x \in \mathcal{I}_\ell^m,$$

We also set

$$r(x) := \sum_{i=1}^{D_m} i \mathbf{1}_{\mathcal{I}_i^m}(x),$$

so that $r(x) = i$ if $x \in \mathcal{I}_i^m$. In addition, set

$$p_i^m := \mu(\mathcal{I}_i^m), i \in [D_m], \quad g^*(j) := \max_{i \in [D_m]} \{g_{\max}(i, j)\},$$

$$\tilde{g}_-(i) := \sum_{j=1}^{D_m} p_j^m g_{\min}(i, j), \quad \tilde{g}_+(i) := \sum_{j=1}^{D_m} p_j^m g_{\max}(i, j), \quad \text{and} \quad \tilde{g}_+^* := \sum_{j=1}^{D_m} p_j^m g^*(j). \tag{3.14}$$

Recall that $\tilde{g}(x) = \mathbb{E}[g(x, W)]$, and note that $\tilde{g}_-(r(x)) = \mathbb{E}[g^-(x, W)]$, $\tilde{g}_+(r(x)) = \mathbb{E}[g^+(x, W)]$ and $\tilde{g}_+^* = \mathbb{E}\left[\max_{x \in [0, w^*]} g^+(x, W)\right]$. Then, observe that by Lemma 3.2.4 and dominated convergence, $\tilde{g}_-(r(x)) \uparrow \tilde{g}(x)$, $\tilde{g}_+(r(x)) \downarrow \tilde{g}(x)$ and

$$\tilde{g}_+^* \downarrow \mathbb{E}\left[\sup_{x \in [0, w^*]} g(x, W)\right] = \tilde{g}^*, \quad \text{as } m \to \infty.$$

The Definition of Urn E

We are now ready to define the urn process $(\mathcal{U}_n)_{n \in \mathbb{N}_0}$. For $i \in \mathbb{N}$, set

$$[D_m]^i := [D_m] \times [D_m] \cdots \times [D_m] = \{(u_0, \ldots u_{i-1}) : u_0, \ldots, u_{i-1} \in [D_m]\},$$

and

$$\mathcal{B} := [D_m] \cup [D_m]^2 \cup (\{D_m + 1\} \times [D_m])\,;$$

this will represent the set of types in Urn E. We now define parameters $\boldsymbol{\gamma}$ such that, for $x \in [D_m] \cup [D_m] \times [D_m]$,

$$\boldsymbol{\gamma}(x) = \begin{cases} \frac{g_{\min}(i,j)}{g_{\max}(i,j)}, & x = (i,j) \in [D_m]^2, g_{\max}(i,j) > 0; \\ \frac{h_{\min}(i)}{h_{\max}(i)}, & x = i \in [D_m], h_{\max}(i) > 0; \\ 0, & \text{otherwise.} \end{cases} \tag{3.15}$$

Then, we define the urn process $(\mathcal{U}_n^m)_{n \in \mathbb{N}_0}$ as the urn process with *activities* $\mathbf{a}$ such that

$$\mathbf{a}(x) = \begin{cases} g_{\max}(i,j) & \text{if } x = (i,j),\ i,j \in [D_m] \\ g^*_{\max}(j) & \text{if } x = (i,j),\ i = D_m + 1, j \in [D_m] \\ h_{\max}(i) & \text{if } x = i \in [D_m]; \end{cases} \tag{3.16}$$

and a *replacement matrix* M such that, for $x, x' \in \mathbb{V}_{\mathcal{B}}$,

$$M_{x',x} = \begin{cases} (\boldsymbol{\gamma}\mathbf{a})(x) p_\ell^m, & \text{if } x' = (i,\ell), x \in (\{i\} \times [D_m]) \cup \{i\}, i, \ell \in [D_m]; \\ (\mathbf{a} - \boldsymbol{\gamma}\mathbf{a})(x) p_\ell^m, & \text{if } x' = (D_m + 1, \ell)\,, x \in \mathcal{B}; \\ \mathbf{a}(x) p_\ell^m, & \text{if } x' = \ell, x \in \mathcal{B}; \\ 0 & \text{otherwise.} \end{cases}$$

Note that it is not necessarily the case that M is irreducible: it may be the case that $\mathbf{a}(x) = 0$ for certain $x \in \mathcal{B}$ (this is possible if $h_{\max}(i) = 0$ or $g_{\max}(i,j) = 0$), or it may be the case that $p_\ell^m = 0$ for certain choices of ℓ. We therefore define the following subsets of $\mathcal{B}$:

$$\mathscr{U}_1 := \{x \in \mathcal{B} : M_{x'x} = 0\ \forall x' \in \mathcal{B}\} = \{x \in \mathcal{B} : \mathbf{a}(x) = 0\}\,,$$

and

$$\mathscr{U}_2 := \{x' \in \mathcal{B} : M_{x'x} = 0\ \forall x \in \mathcal{B}\}\,.$$

Also assume that $\mathscr{U}_1 \cap \mathscr{U}_2 = \varnothing$; if not, we replace $\mathscr{U}_1$ by $\mathscr{U}_1 \backslash \mathscr{U}_2$. We then set $R = \mathcal{B} \backslash (\mathscr{U}_1 \cup \mathscr{U}_2)$, and let M_R be the restriction of M to R. It is easy to check that M_R is irreducible, and thus, by Lemma 3.2.2, has a unique largest positive eigenvalue λ_m with corresponding eigenvector $\mathbf{u}_R$. But then, writing M in block form (with columns and rows labelled by $R, \mathscr{U}_1, \mathscr{U}_2$) for suitable matrices A, B, C, we have

$$M = \begin{array}{c} \begin{matrix} R & \mathscr{U}_1 & \mathscr{U}_2 \end{matrix} \\ \begin{pmatrix} M_R & 0 & B \\ A & 0 & C \\ 0 & 0 & 0 \end{pmatrix} \end{array} \begin{matrix} \\ R \\ \mathscr{U}_1 \\ \mathscr{U}_2 \end{matrix} .$$

Thus, M has the same largest positive eigenvalue, with corresponding right eigenvector given (in block form) by

$$\mathbf{u}_m = \begin{bmatrix} \mathbf{u}_R \\ \left(\lambda_R^{-1}\right) A \mathbf{u}_R \\ 0 \end{bmatrix} .$$

Here, we assume $\mathbf{u}_m$ is normalised so that $\mathbf{a} \cdot \mathbf{u}_m = 1$. In addition, assuming we begin with a single ball $x \in R$, one readily verifies that the restriction of M to R and $\mathscr{U}_1$ satisfies conditions (A1)-(A6) of Subsection 3.2.1. Note also, that at each time-step the probability of adding a ball of type $x \in \mathscr{U}_2$ is 0 and thus, for each $n \in \mathbb{N}_0$, $\mathcal{U}_n(x) = 0$ almost surely. Therefore, combining this fact with Theorem 3.2.1, we have the following corollary.

Corollary 3.2.5. *With $\mathbf{u}_m, \lambda_m$ and R as defined above, assuming we begin with a ball $x \in R$, we have*

$$\frac{\mathcal{U}_n^m}{n} \xrightarrow{n \to \infty} \lambda_m \mathbf{u}_m \tag{3.17}$$

almost surely. In particular, almost surely

$$\frac{\mathbf{a} \cdot \mathcal{U}_n^m}{n} \xrightarrow{n \to \infty} \lambda_m . \tag{3.18}$$

In the coupling below, the assumption of a ball $x \in R$ is met by the tree process being initiated by a vertex 0 with weight W_0 sampled at random from μ and satisfying $h(W_0) > 0$.

Coupling Urn E with the PANI-tree Process

For a product measurable set $A \subseteq [0, w^*] \times [0, w^*]$, recall the definition of $\Xi^{(2)}(A, n)$ from (1.6): this is the number of directed edges (v, v') of $\mathcal{T}_n$ where $(W_v, W_{v'}) \in A$.

Proposition 3.2.6. *There exists a coupling $((\hat{\mathcal{U}}_n^m)_{m\in\mathbb{N}}, \hat{\mathcal{T}}_n)_{n\in\mathbb{N}_0}$ of the Pólya urn processes $\{(\mathcal{U}_n^m)_{n\in\mathbb{N}_0}, m \in \mathbb{N}\}$ and the tree process $(\mathcal{T}_n)_{n\in\mathbb{N}_0}$ such that, for each $m \in \mathbb{N}$, almost surely (on the coupling space), $\hat{\mathcal{U}}_0^m = \delta_\ell$ for an initial ball of type $\ell \in R$ and, in addition, for $(i, j) \in [D_m]^2$, we have*

$$\hat{\mathcal{U}}_n^m((i,j)) \leqslant \Xi^{(2)}(n, \mathcal{I}_i^m \times \mathcal{I}_j^m), \tag{3.19}$$

$$\sum_{(i,j)\in[D_m]^2} \Big(\Xi^{(2)}(n, \mathcal{I}_i^m \times \mathcal{I}_j^m) - \hat{\mathcal{U}}_n^m((i,j))\Big) = \sum_{j=1}^{D_m} \hat{\mathcal{U}}_n^m((D_m+1, j)), \tag{3.20}$$

and

$$(\boldsymbol{\gamma}\mathbf{a}) \cdot \hat{\mathcal{U}}_n^m \leqslant \mathcal{Z}_n \leqslant \mathbf{a} \cdot \hat{\mathcal{U}}_n^m. \tag{3.21}$$

for all $n \in \mathbb{N}_0$.

Proof. First sample the entire tree process $(\hat{\mathcal{T}}_n)_{n\in\mathbb{N}_0}$; we will use this to define the evolution of the urn processes. Moreover, for $i \in [D_m]$ let

$$\eta_n(i) := \sum_{v\in\mathcal{T}_n : r(v)=i} f(N^+(v, \mathcal{T}_n));$$

i.e., the sum of fitnesses of vertices with weight belonging to $\mathcal{I}_i^m$. Also, for $i \in [D_m]$ define

$$\theta_n(i) := (\boldsymbol{\gamma}\,\mathbf{a}\hat{\mathcal{U}}_n^m)(i) + \sum_{j=1}^{D_m} (\boldsymbol{\gamma}\,\mathbf{a}\hat{\mathcal{U}}_n^m)((i,j)).$$

Finally, recall that $\mathcal{Z}_n$ denotes the partition function associated with the tree at time n. Assume that at time 0 the tree consists of a single vertex 0 such that $r(W_0) = \ell \in [D_m]$. Then, set $\hat{\mathcal{U}}_0^m = \delta_\ell$. Using the definition of r, since $W_0 \in \mathcal{I}_\ell^m$

$$0 < \mathcal{Z}_0 = h(W_0) \leqslant h_{\max}(\ell) = \mathbf{a} \cdot \hat{\mathcal{U}}_0^m,$$

and by the choice of $\boldsymbol{\gamma}$, we have

$$\eta_0(\ell) = h(W_0) \geqslant h_{\min}(\ell) = (\boldsymbol{\gamma}\,\mathbf{a}\hat{\mathcal{U}}_0^m)(\ell) = \theta_0(\ell).$$

In this case, (3.19) and (3.20) are trivially satisfied since both sides of both equations are 0. Now, assume inductively that after n steps in the urn process, (3.19) and (3.20) are satisfied, we have

$$\eta_n(k) \geqslant \theta_n(k) \quad \text{for each} \quad k \in [D_m], \tag{3.22}$$

and moreover, $\mathcal{Z}_n \leqslant \mathbf{a} \cdot \hat{\mathcal{U}}_n^m$. Note that (3.22) implies the left hand side of (3.21), since

$$(\boldsymbol{\gamma}\mathbf{a}) \cdot \hat{\mathcal{U}}_n^m = \sum_{k=1}^{D_m} \theta_n(k) \leqslant \sum_{k=1}^{D_m} \eta_n(k) = \mathcal{Z}_n.$$

Let s be the vertex sampled from $\mathcal{T}_n$ in the $(n+1)$st step, and assume that $r(W_s) = \ell'$, $r(W_{n+1}) = k$. Then, for the $(n+1)$th step in the urn: sample an independent random variable U_{n+1} uniformly distributed on $[0,1]$. Then:

- If $U_{n+1} \leqslant \frac{\theta_n(\ell')\mathcal{Z}_n}{\eta_n(\ell')\mathbf{a}\cdot\hat{\mathcal{U}}_n^m}$, add balls of type (ℓ', k) and k to the urn, i.e., set $\hat{\mathcal{U}}_{n+1}^m = \hat{\mathcal{U}}_n^m + \delta_{(\ell',k)} + \delta_k$.
- Otherwise, add balls of type $(D_m + 1, k), k$.

Note that, in the first case, we have

$$\Xi^{(2)}(n+1, \mathcal{I}_{\ell'}^m \times \mathcal{I}_k^m) = \Xi^{(2)}(n, \mathcal{I}_{\ell'}^m \times \mathcal{I}_k^m) + 1 \geqslant \hat{\mathcal{U}}_n^m((\ell', k)) + 1 = \hat{\mathcal{U}}_{n+1}^m((\ell', k))$$

and for $i \neq \ell'$ or $j \neq k$

$$\Xi^{(2)}(n+1, \mathcal{I}_i^m \times \mathcal{I}_j^m) = \Xi^{(2)}(n, \mathcal{I}_i^m \times \mathcal{I}_j^m) \geqslant \hat{\mathcal{U}}_n^m((i,j)) = \hat{\mathcal{U}}_{n+1}^m((i,j)).$$

Also, in this case

$$\eta_{n+1}(\ell') = \eta_n(\ell') + g(W_s, W_{n+1}) \geqslant \theta_n(\ell') + g_{\min}(\ell', k) = \theta_{n+1}(\ell'),$$

and similarly,

$$\eta_{n+1}(k) = \eta_n(k) + h(W_{n+1}) \geqslant \theta_n(k) + h_{\min}(k) = \theta_{n+1}(k),$$

so that (3.22) is satisfied. Finally, in this case,

$$\mathcal{Z}_{n+1} = \mathcal{Z}_n + g(W_s, W_{n+1}) + h(W_{n+1}) \leqslant \mathbf{a} \cdot \hat{\mathcal{U}}_n^m + g_{\max}(\ell', k) + h_{\max}(k) = \mathbf{a} \cdot \hat{\mathcal{U}}_{n+1}^m.$$

Meanwhile, in the second case $\Xi^{(2)}(n, \mathcal{I}_{\ell'}^m \times \mathcal{I}_k^m)$ and $\eta_n(\ell')$ increase, while $\sum_{j=1}^{D_m} \hat{\mathcal{U}}_n^m((\ell', j))$ and $\theta_n(\ell')$ remain the same, and thus (3.19) is satisfied and $\eta_{n+1}(\ell') \geqslant \theta_{n+1}(\ell')$. As this is the only case when $\Xi^{(2)}(n, \mathcal{I}_{\ell'}^m \times \mathcal{I}_k^m) - \hat{\mathcal{U}}_n^m((\ell', k))$ increases, and we add a ball of type $(D_m + 1, k)$, (3.20) also follows. Both $\eta_n(k)$ and $\theta_n(k)$ increase as in the first case. Next,

$$\mathcal{Z}_{n+1} = \mathcal{Z}_n + g(W_s, W_{n+1}) + h(W_{n+1}) \leqslant \mathbf{a} \cdot \hat{\mathcal{U}}_n^m + g_{\max}^*(k) + h_{\max}(k) = \mathbf{a} \cdot \hat{\mathcal{U}}_{n+1}^m.$$

As all other quantities remain the same, (3.22) is satisfied, and moreover, $\mathcal{Z}_{n+1} \leqslant \mathbf{a} \cdot \hat{\mathcal{U}}_{n+1}^m$. To complete the proof, it remains to prove the following claim.

Claim 3.2.7. *For each $m \in \mathbb{N}$, almost surely (on the coupling space), the urn process $\hat{\mathcal{U}}^m = (\hat{\mathcal{U}}_n^m)_{n \in \mathbb{N}_0}$ is distributed like the Pólya urn process $(\mathcal{U}_n^m)_{n \in \mathbb{N}_0}$ with $\mathcal{U}_0^m = \delta_\ell$ for an initial ball of type $\ell \in R$.*

Proof. First note that, since W_0 is sampled from μ, conditionally on the positive probability event $\{h(W_0) > 0\}$, we have

$$\mathbb{P}(W_0 \in \mathcal{I}_\ell^m, h(W_0) > 0) \leqslant \mathbb{P}(W_0 \in \mathcal{I}_\ell^m) = p_\ell^m,$$

and thus, $\mathbb{P}$-a.s., we have $W_0 \in \mathcal{I}_\ell^m$ with $p_\ell^m > 0$. This, combined with the fact that $0 < h(W_0) \leqslant h_{\max}(\ell)$, implies that $\mathbb{P}$-a.s., the initial ball $\ell \in R$.

Now, note that in every step in $(\hat{\mathcal{U}}_n^m)_{n\in\mathbb{N}_0}$, we add a ball of type k for $k \in [D_m]$ with probability p_k^m, which is the same as in $(\mathcal{U}_n^m)_{n\in\mathbb{N}_0}$. Moreover, given $\hat{\mathcal{U}}_n^m$, the probability of adding balls of type (k,ℓ) is

$$p_\ell^m \left(\frac{\eta_n(k)}{\mathcal{Z}_n} \times \frac{\theta_n(k)\mathcal{Z}_n}{\eta_n(k)\mathbf{a}\cdot\hat{\mathcal{U}}_n^m} \right) = p_\ell^m \frac{\theta_n(k)}{\mathbf{a}\cdot\hat{\mathcal{U}}_n^m},$$

which also agrees with the Pólya urn scheme. Finally, the probability of adding a ball of type (D_m+1,ℓ) is

$$p_\ell^m \sum_{j=1}^{D_m} \left[\left(1 - \frac{\theta_n(j)\mathcal{Z}_n}{\eta_n(j)\mathbf{a}\cdot\hat{\mathcal{U}}_n^m} \right) \frac{\eta_n(j)}{\mathcal{Z}_n} \right] = p_\ell^m \left(1 - \sum_{j=1}^{D_m} \frac{\theta_n(j)}{\mathbf{a}\cdot\hat{\mathcal{U}}_n^m} \right),$$

as required. □

□

Note also, that, since the functions h^+, g^+ are non-increasing pointwise in m, on the coupling we have that for any fixed n, $\mathbf{a}\cdot\mathcal{U}_n^m$ is non-increasing in m. Combining this result with Corollary 3.2.5, we have the following corollary.

Corollary 3.2.8. *The sequence $(\lambda_m)_{m\in\mathbb{N}}$ is non-increasing in m. In particular, there exists a limit $\lambda_\infty \geqslant 0$ such that*

$$\lambda_m \downarrow \lambda_\infty$$

as $m \to \infty$.

The Limiting Vectors of Urn Schemes Associated with Urn E

We now calculate the limiting vector $\mathbf{u}_m$ and the limiting eigenvalue λ_m. First note that by the definition of the urn process, for each $n \in \mathbb{N}_0$, $\ell \in [D_m]$ we have that $\mathcal{U}_{n+1}^m(\ell) - \mathcal{U}_n^m(\ell)$

is Bernoulli distributed with parameter p_ℓ^m. Thus, by the strong law of large numbers and Corollary 3.2.5, we have, for each $\ell \in [D_m]$,

$$\mathbf{u}_m(\ell) = \frac{p_\ell^m}{\lambda_m}. \tag{3.23}$$

Next, for any $i, j \in [D_m]$ using the definitions of $\boldsymbol{\gamma}$ and $\mathbf{a}$ ((3.15) and (3.16)) we have

$$\begin{aligned}
\lambda_m \mathbf{u}_m((i,j)) &= p_j^m \sum_{\ell=1}^{D_m} (\boldsymbol{\gamma}\, \mathbf{a}\mathbf{u}_m)((i,\ell)) + p_j^m (\boldsymbol{\gamma}\, \mathbf{a}\mathbf{u}_m)(i) \\
&= p_j^m \sum_{\ell=1}^{D_m} g_{\min}\,(i,\ell)\, \mathbf{u}_m((i,\ell)) + p_j^m h_{\min}\,(i)\, \mathbf{u}_m(i) \\
&\overset{(3.23)}{=} p_j^m \sum_{\ell=1}^{D_m} g_{\min}\,(i,\ell)\, \mathbf{u}_m((i,\ell)) + \frac{p_j^m p_i^m h_{\min}\,(i)}{\lambda_m}.
\end{aligned} \tag{3.24}$$

We now define

$$\mathcal{A}_i := \sum_{\ell=1}^{D_m} g_{\min}\,(i,\ell)\, \mathbf{u}_m((i,\ell)).$$

Multiplying both sides of (3.24) by $g_{\min}\,(i,j)$ and taking the sum over $j \in [D_m]$, recalling the definition of $\tilde{g}_-(i)$ in (3.14), we get

$$\begin{aligned}
\lambda_m \mathcal{A}_i &= \left(\mathcal{A}_i + \frac{p_i^m h_{\min}\,(i)}{\lambda_m}\right) \sum_{j=1}^{D_m} p_j^m g_{\min}\,(i,j) \\
&= \left(\mathcal{A}_i + \frac{p_i^m h_{\min}\,(i)}{\lambda_m}\right) \tilde{g}_-(i).
\end{aligned}$$

Thus, solving for $\mathcal{A}_i$

$$\mathcal{A}_i = \frac{p_i^m h_{\min}\,(i)\, \tilde{g}_-(i)}{\lambda_m(\lambda_m - \tilde{g}_-(i))}. \tag{3.25}$$

Substituting (3.25) into (3.24), we have

$$\begin{aligned}
\lambda_m \mathbf{u}_m((i,j)) &= p_j^m \left(\frac{p_i^m h_{\min}\,(i)\, \tilde{g}_-(i)}{\lambda_m(\lambda_m - \tilde{g}_-(i))} + \frac{p_i^m h_{\min}\,(i)}{\lambda_m} \right) \\
&= p_j^m \frac{p_i^m h_{\min}\,(i)}{\lambda_m - \tilde{g}_-(i)}.
\end{aligned} \tag{3.26}$$

Meanwhile, for each $j \in [D_m]$ we have

$$\begin{aligned}
\lambda_m \mathbf{u}_m((D_m+1,j)) &= p_j^m \left(\sum_{\ell=1}^{D_m} (\mathbf{a}\mathbf{u}_m)((D_m+1,\ell)) + \sum_{i=1}^{D_m}\sum_{\ell=1}^{D_m} (\mathbf{a} - \boldsymbol{\gamma}\,\mathbf{a})((i,\ell)) + \sum_{i=1}^{D_m} (\mathbf{a} - \boldsymbol{\gamma}\,\mathbf{a})(i) \right) \\
&= p_j^m \left(\sum_{\ell=1}^{D_m} g^*(\ell)\mathbf{u}_m((D_m+1,\ell)) + \sum_{i=1}^{D_m}\sum_{\ell=1}^{D_m} (g_{\max}\,(i,\ell) - g_{\min}\,(i,\ell))\mathbf{u}_m((i,\ell)) \right. \\
&\qquad \left. + \sum_{i=1}^{D_m} (h_{\max}\,(i) - h_{\min}\,(i))\mathbf{u}_m(i) \right) \\
&=: p_j^m \left(\mathcal{B}_m + \mathcal{E}_m \right); \qquad (3.27)
\end{aligned}$$

where, in the last equation we set

$$\mathcal{B}_m := \sum_{\ell=1}^{D_m} g^*(\ell)\mathbf{u}_m((D_m+1,\ell))$$

and

$$\mathcal{E}_m := \sum_{i=1}^{D_m}\sum_{\ell=1}^{D_m} (g_{\max}\,(i,\ell) - g_{\min}\,(i,\ell))\mathbf{u}_m((i,\ell)) + \sum_{i=1}^{D_m} (h_{\max}\,(i) - h_{\min}\,(i))\mathbf{u}_m(i).$$

Multiplying both sides of (3.27) by $g^*(j)$ and taking the sum over j, we have

$$\lambda_m \mathcal{B}_m = \left(\sum_{j=1}^{D_m} p_j^m g^*(j) \right) (\mathcal{B}_m + \mathcal{E}_m) = \tilde{g}_+^* (\mathcal{B}_m + \mathcal{E}_m)$$

and thus

$$\mathcal{B}_m = \frac{\tilde{g}_+^*}{\lambda_m - \tilde{g}_+^*} \mathcal{E}_m. \qquad (3.28)$$

Note that all of the previous analysis implicitly applied Condition **C2**. We now apply Condition **C1** in the following lemma:

Lemma 3.2.9. *Assume Conditions **C1** and **C2**. Then, we have* $\lambda_\infty := \lim_{m\to\infty} \lambda_m > \tilde{g}^*$.

Proof. Note that, since we add two balls to the urn at each time-step, we have

$$\|\mathcal{U}_{n+1}^m\|_1 - \|\mathcal{U}_n^m\|_1 = 2.$$

Thus, by (3.17), we have $\|\lambda_m \mathbf{u}_m\|_1 = 2$. Now, by (3.23), we have $\lambda_m \sum_{\ell=1}^{D_m} \mathbf{u}_m(\ell) = 1$, and thus, by (3.26), we have

$$\sum_{j=1}^{D_m}\sum_{i=1}^{D_m} \lambda_m \mathbf{u}_m((i,j)) = \mathbb{E}\left[\frac{h_{\min}\left(r\left(W\right)\right)}{\lambda_m - \tilde{g}_{-}\left(r\left(W\right)\right)}\right] \leqslant 1.$$

Note that as $m \to \infty$, $h_{\min}\left(r\left(W\right)\right) \uparrow h(W)$ and $\tilde{g}_{-}\left(r\left(W\right)\right) \uparrow \tilde{g}(W)$. Thus, by the monotone convergence theorem, we have

$$\mathbb{E}\left[\frac{h(W)}{\lambda_\infty - \tilde{g}(W)}\right] = \lim_{m\to\infty} \mathbb{E}\left[\frac{h_{\min}\left(r\left(W\right)\right)}{\lambda_m - \tilde{g}_{-}\left(r\left(W\right)\right)}\right] \leqslant 1.$$

Now, since the eigenvectors $\mathbf{u}_m$ are non-negative, by (3.28), we have

$$\lambda_m \geqslant \tilde{g}^*_+,$$

and thus, $\lambda_\infty = \lim_{m\to\infty} \lambda_m \geqslant \lim_{m\to\infty} \tilde{g}^*_+ = \tilde{g}^*$. But, if $\lambda_\infty = \tilde{g}^*$, since the expression in (2.4) is decreasing in λ^*, we would have a contradiction to Condition **C1**. The result follows. □

Lemma 3.2.10. *Assume Conditions **C1** and **C2**. Then, we have $\mathcal{B}_m \downarrow 0$ and $\mathcal{E}_m \downarrow 0$ as $m \to \infty$. In particular,*

$$\mathbb{E}\left[\frac{h(W)}{\lambda_\infty - \tilde{g}(W)}\right] = 1,$$

so that $\lambda_\infty = \lambda^*$.

Proof. First, note that by Corollary 3.2.8 and Lemma 3.2.9, for each $m \in \mathbb{N}$, we have $\lambda_m \geqslant \lambda_\infty > \tilde{g}^*$. Combining this fact with the boundedness of g and h we observe that

$$\sup_{x\in[0,w^*]}\left\{\frac{h(x)}{\lambda_m\left(\lambda_m - \tilde{g}(x)\right)}, \frac{1}{\lambda_m}\right\} < \sup_{x\in[0,w^*]}\left\{\frac{h(x)}{\tilde{g}^*\left(\lambda_\infty - \tilde{g}(x)\right)}, \frac{1}{\lambda_\infty}\right\} =: C < \infty,$$

where the bound on the right is independent of m. Now, given $\varepsilon > 0$, by applying Lemma 3.2.4, let m be sufficiently large that for all $x, y \in [0, w^*]$

$$\left(g^+(x,y) - g^-(x,y)\right) < \frac{\varepsilon}{2C} \quad \text{and} \quad \left(h^+(x) - h^-(x)\right) < \frac{\varepsilon}{2C}.$$

Then we have

$$\begin{aligned}\mathcal{E}_m &= \sum_{i=1}^{D_m}\sum_{j=1}^{D_m}\left(g_{\max}\,(i,j) - g_{\min}\,(i,j)\right)\mathbf{u}_m((i,j)) + \sum_{\ell=1}^{D_m}\left(h_{\max}\,(\ell) - h_{\min}\,(\ell)\right)\mathbf{u}_m(\ell) \\ &\overset{(3.23),(3.26)}{=} \sum_{i=1}^{D_m}\sum_{j=1}^{D_m}\left(g_{\max}\,(i,j) - g_{\min}\,(i,j)\right)\frac{h_{\min}(i)p_i^m p_j^m}{\lambda_m(\lambda_m - \tilde{g}_-(i))} + \sum_{\ell=1}^{D_m}\left(h_{\max}\,(\ell) - h_{\min}\,(\ell)\right)\frac{p_\ell^m}{\lambda_m} \\ &< \frac{\varepsilon}{2C}\cdot C\left(\sum_{i=1}^{D_m}\sum_{j=1}^{D_m} p_i^m p_j^m\right) + \frac{\varepsilon}{2C}\cdot C\left(\sum_{\ell=1}^{D_m} p_\ell^m\right) = \varepsilon.\end{aligned}$$

The result for $\mathcal{B}_m$ then follows from the fact that $\tilde{g}_+^* \downarrow \tilde{g}^*$, and Lemma 3.2.9. □

We are now ready to prove our main results of this subsection.

Proof of Theorem 3.1.1

Proof of Theorem 3.1.1. Note that, by (3.21) from Proposition 3.2.6, we have

$$0 \leqslant \mathbf{a}\cdot\mathcal{U}_n^m - \mathcal{Z}_n \leqslant (\mathbf{a} - \boldsymbol{\gamma}\mathbf{a})\cdot\mathcal{U}_n^m.$$

Dividing by n and taking limits as $n \to \infty$, by (3.18) we have

$$0 \leqslant \lambda_m - \limsup_{n\to\infty}\frac{\mathcal{Z}_n}{n} \leqslant \lambda_m - \liminf_{n\to\infty}\frac{\mathcal{Z}_n}{n} \leqslant \limsup_{n\to\infty}\left((\mathbf{a} - \boldsymbol{\gamma}\mathbf{a})\cdot\frac{\mathcal{U}_n^m}{n}\right) = \mathcal{B}_m + \mathcal{E}_m.$$

The result follows by applying Lemma 3.2.10. □

In addition, recalling the definition of $\mathscr{I}^m$ from (3.13), note that

$$\sigma(\mathscr{I}^m) = \left\{S \subseteq [0, w^*] : S = \bigcup_{i\in I}\mathcal{I}_i^m, I \subseteq [D_m]\right\}. \tag{3.29}$$

In other words, the σ-algebra generated by $\mathscr{I}^m$ is the set of finite unions of sets in $\mathscr{I}^m$. Recalling that $\mathscr{I}^{m_2}$ is a refined partition of $\mathscr{I}^{m_1}$ for $m_1 < m_2$, by Lemma 3.2.3 we have

$$\sigma(\mathscr{I}^{m_1}) \subseteq \sigma(\mathscr{I}^{m_2}). \tag{3.30}$$

We now prove Theorem 3.1.2.

Proof of Theorem 3.1.2

Proof of Theorem 3.1.2. We begin by proving the result for Cartesian products of the form $S \times S'$ with $S, S' \in \sigma(\mathscr{I}^{m'})$, for $m' \in \mathbb{N}$. Note that, by the definition of $\Xi^{(2)}(n, \cdot)$, we clearly have finite *additivity*, that is, for any measurable sets $S_1, S_2, S_3 \subseteq [0, w^*]$ if $S_1 \cap S_2 = \varnothing$, we have

$$\Xi^{(2)}(n, (S_1 \cup S_2) \times S_3) = \Xi^{(2)}(n, S_1 \times S_3) + \Xi^{(2)}(n, S_2 \times S_3), \quad \text{and similarly,}$$

$$\Xi^{(2)}(n, S_3 \times (S_1 \cup S_2)) = \Xi^{(2)}(n, S_3 \times S_1) + \Xi^{(2)}(n, S_3 \times S_2).$$

Combining these facts with Proposition 3.2.6, Corollary 3.2.5 and (3.26), for sets $S \times S'$ with $S, S' \in \sigma(\mathscr{I}^{m'})$ we have, for each $m > m'$,

$$\begin{aligned}\mathbb{E}\left[\frac{h^-(W)}{\lambda_m - \tilde{g}_-(r(W))}\mathbf{1}_S(W)\right]\mu(S') &\leqslant \liminf_{n\to\infty} \frac{\Xi^{(2)}(n, S \times S')}{n} \\ &\leqslant \limsup_{n\to\infty} \frac{\Xi^{(2)}(n, S \times S')}{n} \\ &\leqslant \mathbb{E}\left[\frac{h^-(W)}{\lambda_m - \tilde{g}_-(r(W))}\mathbf{1}_S(W)\right]\mu(S') + \mathcal{B}_m + \mathcal{E}_m.\end{aligned}$$

Taking limits as $m \to \infty$ and applying Lemma 3.2.10, this proves the result for this family of sets.

Now, by the Portmanteau Theorem, we need only prove that for all sets $U \in \mathcal{O}$, where $\mathcal{O}$ denotes the class of open subsets of $[0, w^*] \times [0, w^*]$, we have

$$\liminf_{n\to\infty} \frac{\Xi^{(2)}(n, U)}{n} \geqslant (\psi_* \mu \times \mu)(U). \tag{3.31}$$

Now, let

$$\mathcal{I}^m(U) := \bigcup_{i,j \in [D_m] : \mathcal{I}_i^m \times \mathcal{I}_j^m \subseteq U} \mathcal{I}_i^m \times \mathcal{I}_j^m.$$

Note that, since U is open, and $\mathscr{I}^m$ is fine enough that the set of dyadic intervals $\{\mathcal{D}_i^m(w^*)\}_{i \in [2^m]} \subseteq \sigma(\mathscr{I}^m)$, we have

$$\mathbf{1}_{\mathcal{I}^m(U)}(W) \uparrow \mathbf{1}_U(W) \quad \text{pointwise as } m \to \infty. \tag{3.32}$$

In addition, since $\mathcal{I}^m(U) \subseteq U$, for each $m \in \mathbb{N}$

$$(\psi_*\mu \times \mu)(\mathcal{I}^m(U)) = \liminf_{n\to\infty} \frac{\Xi^{(2)}(n, \mathcal{I}^m(U))}{n} \leqslant \liminf_{n\to\infty} \frac{\Xi^{(2)}(n, U)}{n}.$$

Then, (3.31) follows by taking limits as $m \to \infty$. □

3.2.3 Analysing the PANI-tree by Coupling with Urn D

In order to analyse the degree distribution in this model under Conditions **C1** and **C2**, we introduce another collection of Pólya urns $(\mathcal{V}_n^{K'})_{n\in\mathbb{N}_0}$, which not only depend on m, but also depends on a parameter $K' \in \mathbb{N}$. These may be regarded as finite approximations of Urn D. For brevity of notation, wherever possible in this subsection we will omit the dependence of these parameters on m. For $i \in \mathbb{N}$, define $[D_m]^i$ so that

$$[D_m]^i := \{(u_0, \dots u_{i-1}) : u_0, \dots, u_{i-1} \in [D_m]\}.$$

Now, we set

$$\mathcal{B}' := \left(\bigcup_{i=1}^{K'+1} [D_m]^i\right) \cup (\{D_m + 1\} \times [D_m]).$$

The urn process $(\mathcal{V}_n^{K'})_{n\geqslant 0}$ is then a vector-valued stochastic process taking values in $\mathbb{V}_{\mathcal{B}'}$. We now define the vectors $\mathbf{a}'$, $\boldsymbol{\gamma}'$ associated with the urn process such that

$$\mathbf{a}'(x) = \begin{cases} h_{\max}(u_0) + \sum_{j=1}^k g_{\max}(u_0, u_j) & \text{if } x = (u_0, \dots, u_k) \in [D_m]^{k+1} \\ g^*_{\max}(\ell) & \text{if } x = (D_m + 1, \ell); \end{cases}$$

and,

$$\boldsymbol{\gamma}'(x) = \begin{cases} \frac{h_{\min}(u_0) + \sum_{j=1}^k g_{\min}(u_0, u_j)}{h_{\max}(u_0) + \sum_{j=1}^k g_{\max}(u_0, u_j)}, & \text{if } x = (u_0, \dots, u_k) \in [D_m]^{k+1}, k < K', \mathbf{a}'(x) > 0; \\ 0, & \text{otherwise.} \end{cases}$$

Now, given $\mathbf{u} = (u_0, \dots, u_k) \in [D_m]^{k+1}$, $k < K'$, and $\ell \in [D_m]$, we define their *concatenation* $(\mathbf{u}, \ell) \in [D_m]^{k+2}$ such that

$$(\mathbf{u}, \ell) := (u_0, \dots, u_k, \ell).$$

Then, we define the replacement matrix M' of the urn $(\mathcal{V}_n^{K'})_{n\in\mathbb{N}_0}$ such that, given $x, x' \in \mathcal{B}'$,

$$M'_{x',x} = \begin{cases} -(\boldsymbol{\gamma}'\mathbf{a}')(x) & \text{if } x' = x, x \in [D_m]^k, k \leqslant K'; \\ (\boldsymbol{\gamma}'\mathbf{a}')(x)p_\ell^m, & \text{if } x' = (x, \ell), \ell \in [D_m], x \in \mathcal{B}'; \\ (\mathbf{a}' - \boldsymbol{\gamma}'\mathbf{a}')(x)p_\ell^m, & \text{if } x' = (D_m + 1, \ell), \ell \in [D_m], x \in \mathcal{B}'; \\ \mathbf{a}'(x)p_\ell^m, & \text{if } x' = \ell, x \in \mathcal{B}'; \\ 0 & \text{otherwise.} \end{cases}$$

Again, note that it may be the case that M' is not irreducible, if either $\mathbf{a}'(x) = 0$ for certain $x \in \mathcal{B}'$ or $p_\ell^m = 0$ for certain choices of ℓ. Nevertheless, we define the sets

$$\mathscr{U}_1' := \{x \in \mathcal{B}' : M'_{x'x} = 0 \ \forall x' \in \mathcal{B}'\} = \{x \in \mathcal{B}' : \mathbf{a}'(x) = 0\},$$

and

$$\mathscr{U}_2' := \{x' \in \mathcal{B}' : M'_{x'x} = 0 \ \forall x \in \mathcal{B}' \backslash \{x'\}\}.$$

Again, we assume that $\mathscr{U}_1' \cap \mathscr{U}_2' = \varnothing$; if not, we replace $\mathscr{U}_1'$ by $\mathscr{U}_1' \backslash \mathscr{U}_2'$. We then set $R' = \mathcal{B}' \backslash (\mathscr{U}_1' \cup \mathscr{U}_2')$, and let $M'_{R'}$ be the restriction of M' to R'. As in Section 3.2.2, $M'_{R'}$ satisfies the conditions of Lemma 3.2.2, and thus has a unique largest positive eigenvalue $\lambda'_{R'}$ with corresponding eigenvector $\mathbf{V}_{R'}$. But then, writing M' in block form in a manner analogous to Section 3.2.2, M has the same largest positive eigenvalue, with corresponding right eigenvector given, in block form, by

$$\mathbf{V}_{K'} = \begin{bmatrix} \mathbf{V}_{R'} \\ (\lambda'_{R'})^{-1} A' \mathbf{V}_{R'} \\ 0 \end{bmatrix}.$$

Here, we assume $\mathbf{V}_{K'}$ is normalised so that $\mathbf{a}' \cdot \mathbf{V}_{K'} = 1$. Also in a manner similar to the Section 3.2.2, assuming we begin with a ball of type $x \in R'$, one readily verifies that the restriction of M' to R' and $\mathscr{U}_1'$ satisfies conditions (A1)-(A6) of Section 3.2.1, and also, that for each $x \in \mathscr{U}_2'$ and $n \in \mathbb{N}_0$, $\mathcal{U}_n(x) = 0$ almost surely. Therefore, applying Theorem 3.2.1 again, we have the following corollary:

Corollary 3.2.11. *With $\mathbf{V}_{K'}, \lambda'_{K'}$ and R' as defined above, assuming we begin with a ball $x \in R'$, we have*

$$\frac{\mathcal{V}_n^{K'}}{n} \xrightarrow{n\to\infty} \lambda'_{K'}\mathbf{V}_{K'}$$

almost surely. In particular, we have

$$\frac{\mathbf{a} \cdot \mathcal{V}_n^{K'}}{n} \xrightarrow{n\to\infty} \lambda'_{K'}. \tag{3.33}$$

As in Section 3.2.2, in the coupling below, the assumption of a ball $x \in R'$ is met by the tree process being initiated by a vertex 0 with weight W_0 sampled at random from μ and satisfying $h(W_0) > 0$.

Coupling Urn D with the PANI-tree Process

Recall that we denote by $N_{\geqslant k}(n, B)$ the number of vertices of out-degree at least k having weight belonging to a measurable set $B \subseteq [0, w^*]$. We also define the analogue $\mathscr{D}_{\geqslant k}(n, j)$ for $n \in \mathbb{N}_0$ and $j \in [D_m]$ such that

$$\mathscr{D}_{\geqslant k}(n, j) := \sum_{j=k}^{K'+1} \sum_{\mathbf{u}_j \in [D_m]^j} \mathcal{V}_n^{K'}(\mathbf{u}_j)\mathbf{1}_{\{j\}}(u_0). \tag{3.34}$$

This represents the number of balls in the urn $\mathcal{V}_n^{K'}$ with type $\mathbf{u} = (u_0, \ldots)$ having dimension at least $k + 1$, with $u_0 = j$. We then have the following analogue of Proposition 3.2.6:

Proposition 3.2.12. *There exists a coupling $(\hat{\mathcal{V}}_n^{K'}, \hat{\mathcal{T}}_n)_{n\in\mathbb{N}_0}$ of the Pólya urn process $(\mathcal{V}_n^{K'})_{n\in\mathbb{N}_0}$ and the tree process $(\mathcal{T}_n)_{n\in\mathbb{N}_0}$ such that, almost surely (on the coupling space),*

$\mathcal{V}_0^{K'}$ *consists of a single ball* $\ell \in R'$ *and for all* $n \in \mathbb{N}_0$, $k \in \{0\} \cup [K']$, *we have*

$$\mathscr{D}_{\geqslant k}(n, j) \leqslant N_{\geqslant k}\left(n, \mathcal{I}_j^m\right) \quad \text{and} \tag{3.35}$$

$$\sum_{j=1}^{D_m} \left(N_{\geqslant k}\left(n, \mathcal{I}_j^m\right) - \mathscr{D}_{\geqslant k}(n, j)\right) \leqslant \sum_{j=1}^{D_m} \hat{\mathcal{V}}_n^{K'}((D_m + 1, j). \tag{3.36}$$

In addition, we have

$$(\boldsymbol{\gamma}'\mathbf{a}') \cdot \hat{\mathcal{V}}_n^{K'} \leqslant \mathcal{Z}_n \leqslant \mathbf{a}' \cdot \hat{\mathcal{V}}_n^{K'}. \tag{3.37}$$

Proof. We proceed in a somewhat similar manner to Proposition 3.2.6, however, in this case, we first introduce a "labelled" Pólya urn $(\mathcal{L}_n)_{n\geqslant 0}$ where balls carry *integer labels* from $\{-D_m, \ldots, 0, \ldots, n\}$. In addition, for $j \in \{0\} \cup [n]$, the label is independent of the *type* of the ball: we denote by $b_n(j)$ the *type* of a ball with label j at time n. One may interpret the ball with label j as representing the evolution of vertex j in the tree process - in this sense, the label may be interpreted as a "time-stamp". Balls of type $(D_m + 1, j)$, $j \in [D_m]$, however, are labelled $-j$ - we denote by $d_n(j)$ the number of balls with this label, since here, multiple balls may share the same label. We describe the labelled urn process $\mathcal{L}_n$ as an evolving vector in $\mathcal{B}' \times \mathbb{Z}$, so that $\mathcal{L}_n = \sum_{j=1}^{D_m} d_n(j) \cdot \delta_{(b_n(j),j)} + \sum_{i=0}^{n} \delta_{(b_n(i),i)}$. We set

$$\mathbf{a}'(\mathcal{L}_n) = \sum_{j=-D_m}^{-1} d_n(j) \cdot \mathbf{a}'(b_n(j)) + \sum_{i=0}^{n} \mathbf{a}'(b_n(i)), \text{ and } (\boldsymbol{\gamma}'\mathbf{a}')(\mathcal{L}_n) = \sum_{i=0}^{n} (\boldsymbol{\gamma}'\mathbf{a}')(b_n(i)).$$

Now, we use $\mathcal{L}_{n+1}$ to define $\hat{\mathcal{V}}_{n+1}^{K'}$ by "forgetting" labels, so that,

$$\text{if } \mathcal{L}_{n+1} = \sum_{j=-D_m}^{-1} d_n(j) \cdot \delta_{(b_{n+1}(j),j)} + \sum_{i=0}^{n+1} \delta_{(b_{n+1}(j),i)}, \text{ we set } \hat{\mathcal{V}}_{n+1}^{K'} = \sum_{j=-D_m}^{-1} d_n(j) \cdot \delta_{b_{n+1}(j)} + \sum_{i=0}^{n+1} \delta_{b_{n+1}(i)}.$$

Sample the entire tree process $(\hat{\mathcal{T}}_n)_{n\in\mathbb{N}_0}$. If, at time 0, the tree consists of a single vertex 0 with weight $W_0 \in I_\ell^m$ then, we set $\mathcal{L}_0 = \delta_{(\ell,0)}$, and note that we have

$$(\boldsymbol{\gamma}'\mathbf{a}')(\mathcal{L}_0) = h_{\min}(\ell) \leqslant h(W_0) = \mathcal{Z}_0 \leqslant \mathbf{a}'(\mathcal{L}_0) = h_{\max}(\ell),$$

and

$$f(N^+(0, \hat{\mathcal{T}}_0)) = h(W_0) \geqslant (\boldsymbol{\gamma}'\,\mathbf{a}')\,(b_0(0)) = h_{\min}(\ell).$$

Now, assume inductively that after n steps in the process, for each $i \in \{0\} \cup [n]$ we have

$$f(N^+(i, \hat{\mathcal{T}}_n)) \geqslant (\boldsymbol{\gamma}' \, \mathbf{a}') \, (b_n(i)), \quad \deg^+(i, \mathcal{T}_n) \geqslant \dim(b_n(i)) - 1, \tag{3.38}$$

$$\sum_{i=0}^{n} \Big(\deg^+(i, \mathcal{T}_n) - \dim(b_n(i)) + 1 \Big) = \sum_{j=1}^{D_m} \hat{\mathcal{V}}_n^{K'}((D_m + 1, j), \tag{3.39}$$

and (3.37) is satisfied.

Let s be the vertex sampled in the tree in the $(n+1)$st step, assume that $r(s) = \ell'$ and that $r(n+1) = k$. Then, for the $(n+1)$th step in the urn: sample an independent random variable U_{n+1} uniformly distributed on $[0,1]$. Then:

- If $\dim(b_n(s)) \leqslant K'$ and $U_{n+1} \leqslant \frac{(\boldsymbol{\gamma}'\mathbf{a}')(b_n(s))\mathcal{Z}_n}{f(N^+(s,\hat{\mathcal{T}}_n))\mathbf{a}'(\mathcal{L}_n)}$, remove the ball $(b_n(s), s)$ from the urn, and add balls $((b_n(s), k), s)$ and $(k, n+1)$ to the urn, i.e., set $\mathcal{L}_{n+1} = \mathcal{L}_n + \delta_{((b_n(s),\ell),s)} + \delta_{(k,n+1)} - \delta_{(b_n(s),s)}$. We call this step Case 1.
- Otherwise, add balls of type $((D_m + 1, k), -k), (k, n+1)$ - we call this Case 2.

First note that

$$(\boldsymbol{\gamma}'\mathbf{a}')(b_{n+1}(s)) - (\boldsymbol{\gamma}'\mathbf{a}')(b_n(s)) = \begin{cases} g_{\min}(\ell', k), & \text{in Case 1} \\ 0, & \text{in Case 2} \end{cases}$$

$$\leqslant g(W_s, W_{n+1}) = f(N^+(s, \hat{\mathcal{T}}_{n+1})) - f(N^+(s, \hat{\mathcal{T}}_n)),$$

and likewise

$$(\boldsymbol{\gamma}'\mathbf{a}')(b_{n+1}(n+1)) = h_{\min}(\ell) \leqslant h(W_{n+1}) = f(N^+(n+1, \hat{\mathcal{T}}_{n+1})).$$

Additionally, in Case 1 the dimension of $b_n(s)$ and the degree of s in $\hat{\mathcal{T}}_n$ both increase, whilst in Case 2 only the degree of s increases whilst the dimension of $b_n(s)$ remains the same. This proves (3.38) at time $n+1$. In addition, Case 2 coincides with the addition of a ball of type

$(D_m + 1, \ell)$, which yields (3.39). Finally,

$$
\begin{aligned}
(\boldsymbol{\gamma}'\mathbf{a}') \cdot \left(\hat{\mathcal{V}}_{n+1}^{K'} - \hat{\mathcal{V}}_n^{K'}\right) &= \begin{cases} h_{\min}(k) + g_{\min}(\ell', k), & \text{in Case 1} \\ h_{\min}(k), & \text{in Case 2} \end{cases} \\
&\leqslant h(W_{n+1}) + g(W_s, W_{n+1}) = \mathcal{Z}_{n+1} - \mathcal{Z}_n \\
&\leqslant \begin{cases} h_{\max}(k) + g_{\max}(\ell', k), & \text{in Case 1} \\ h_{\max}(k) + g^*_{\max}(k), & \text{in Case 2} \end{cases} \\
&\leqslant (\mathbf{a}') \cdot (\hat{\mathcal{V}}_{n+1}^{K'} - \hat{\mathcal{V}}_n^{K'});
\end{aligned}
$$

which shows that (3.37) is also satisfied at time $n + 1$.

Claim 3.2.13. *Almost surely (on the coupling space), the urn process $\hat{\mathcal{V}}^{K'} = (\hat{\mathcal{V}}_n^{K'})_{n\in\mathbb{N}_0}$ is distributed like the Pólya urn $(\mathcal{V}_n^{K'})_{n\in\mathbb{N}_0}$ with $\mathcal{V}_0^{K'}$ consisting of an initial ball $\ell \in R'$.*

Proof. The fact that, $\mathbb{P}$-a.s., the initial ball $\ell \in R'$ follows immediately from the fact that the initial weight W_0 is sampled from μ conditionally on the event $\{h(W_0) > 0\}$ (analogous to in Claim 3.2.7). Moreover, in every step in $\hat{\mathcal{V}}^{K'}$, we add a ball of type k for $k \in [D_m]$ with probability p_k^m, which is the same as in $\mathcal{V}^{K'}$. Furthermore, given $\hat{\mathcal{V}}_n^{K'}$ the probability of removing a ball of type $\mathbf{u}$ with $\dim \mathbf{u} \leqslant K'$ and adding a ball of type $(\mathbf{u}, \ell)$ is

$$
\begin{aligned}
p_\ell^m \sum_{s\in\mathcal{L}_n : b_n(s)=\mathbf{u}} \frac{(\boldsymbol{\gamma}'\mathbf{a}')(b_n(s))\mathcal{Z}_n}{f(N^+(s,\hat{\mathcal{T}}_n))\mathbf{a}'(\mathcal{L}_n)} \times \frac{f(N^+(s,\hat{\mathcal{T}}_n))}{\mathcal{Z}_n} &= p_\ell^m \sum_{s\in\mathcal{L}_n : b_n(s)=\mathbf{u}} \frac{(\boldsymbol{\gamma}'\mathbf{a}')(b_n(s))}{\mathbf{a}'(\mathcal{L}_n)} \\
&= p_\ell^m \frac{\hat{\mathcal{V}}_n^{K'}(\mathbf{u})}{\mathcal{Z}_n},
\end{aligned}
$$

which also agrees with the transition law of the Pólya urn scheme $\mathcal{V}$. Finally, the probability of adding a ball of type $(D_m + 1, \ell)$ is

$$\begin{aligned} p_\ell^m \sum_{s \in \mathcal{L}_n : \dim b_n(s) > K'} \frac{f(N^+(s, \hat{\mathcal{T}}_n))}{\mathcal{Z}_n} + p_\ell^m \sum_{s \in \mathcal{L}_n : \dim b_n(s) \leqslant K'} \left(1 - \frac{(\boldsymbol{\gamma}'\mathbf{a}')(b_n(s))\mathcal{Z}_n}{f(N^+(s, \hat{\mathcal{T}}_n))\mathbf{a}'(\mathcal{L}_n)}\right) \frac{f(N^+(s, \hat{\mathcal{T}}_n))}{\mathcal{Z}_n} \\ = p_\ell^m \sum_{s \in \mathcal{L}_n} \left(\frac{f(N^+(s, \hat{\mathcal{T}}_n))}{\mathcal{Z}_n}\right) - p_\ell^m \sum_{s \in \mathcal{L}_n : \dim b_n(s) \leqslant K'} \frac{(\boldsymbol{\gamma}'\mathbf{a}')(b_n(s))}{\mathbf{a}'(\mathcal{L}_n)} \\ = p_\ell^m \left(1 - \sum_{\mathbf{u} \in \hat{\mathcal{V}}_n^{K'} : \dim \mathbf{u} \leqslant K'} \frac{(\boldsymbol{\gamma}'\mathbf{a}')(\hat{\mathcal{V}}^K(\mathbf{u}))}{\mathbf{a}'(\hat{\mathcal{V}}_n^K)}\right), \end{aligned}$$

which agrees with transition rule of $\mathcal{V}^{K'}$. □

Finally, to complete the proof, we verify the following claim.

Claim 3.2.14. *For all $n \in \mathbb{N}_0$, (3.35) and (3.36) are satisfied for all $k \in \{0\} \cup [K']$.*

Proof. If we define $b_n(i)|_0$ such that $b_n(i)|_0 = x_0$ if $b_n(i) = (x_0, \ldots, x_k)$, then, by construction of the labelled urn process $(\mathcal{L}_n)_{n \in \mathbb{N}_0}$, $b_n(i)|_0 = x_0 \implies r(W_i) = x_0$, so that $W_i \in \mathcal{I}_{x_0}^m$. Therefore, for each $k \in \{0\} \cup [K'], j \in [D_m]$,

$$\mathscr{D}_{\geqslant k}(n, j) = \sum_{b_n(i) : \dim(b_n(i)) \geqslant k+1} \mathbf{1}_{\{j\}}(b_n(i)|_0) \overset{(3.38)}{\leqslant} \sum_{i : \deg^+(i, \hat{\mathcal{T}}_n) \geqslant k} \mathbf{1}_{\mathcal{I}_j^m}(W_i) = N_{\geqslant k}(n, \mathcal{I}_j^m).$$

Moreover, by (3.39),

$$\begin{aligned} \sum_{j=1}^{D_m} \hat{\mathcal{V}}_n^{K'}((D_m + 1, j) &= \sum_{i=0}^{n} \left(\deg^+(i, \hat{\mathcal{T}}_n) - \dim(b_n(i)) + 1\right) \\ &= \sum_{k=0}^{n} \sum_{j=1}^{D_m} \left(\left(N_{\geqslant k}\left(n, \mathcal{I}_j^m\right) - \mathscr{D}_{\geqslant k}(n, j)\right)\right), \end{aligned}$$

which implies (3.36). □

□

The Limiting Vectors of the Urn Schemes Associated with Urn D

We now calculate the limiting vector $\mathbf{V}_K$ and limiting eigenvalue λ'_K of the Pólya urn scheme $(\mathcal{V}_n^{K'})_{n\geqslant 0}$. We first introduce some more notation: for any vector $\mathbf{u} = (u_0, \ldots, u_{k-1}) \in [D_m]^k$, and $i \in \{0\} \cup [k-1]$, denote by $\mathbf{u}|_i := (u_0, \ldots, u_i) \in [D_m]^{i+1}$. We also define the following quantities:

$$\mathcal{R}_{K'} := \sum_{\ell=1}^{D_m} \mathbf{a}'((D_m+1,\ell))\mathbf{V}_{K'}((D_m+1,\ell)), \tag{3.40}$$

$$\mathcal{E}_{K'} := \sum_{\mathbf{u}:\dim \mathbf{u} \leqslant K'} (\mathbf{a}' - \boldsymbol{\gamma}'\mathbf{a}')(\mathbf{u})\mathbf{V}_{K'}(\mathbf{u}), \quad \text{and}$$

$$\mathcal{F}_{K'} := \sum_{\mathbf{v}:\dim \mathbf{v} = K'+1} \mathbf{a}'(\mathbf{v})\mathbf{V}_{K'}(\mathbf{v}). \tag{3.41}$$

Proposition 3.2.15. *Let $\lambda'_{K'}$ and $\mathbf{V}_{K'}$ denote the limiting leading eigenvalue and corresponding right eigenvector of M', respectively. Then, denoting the components of a vector $\mathbf{u}$ by $u_0, u_1, \ldots$, the eigenvector $\mathbf{V}_{K'}$ satisfies*

$$\lambda'_{K'}\mathbf{V}_{K'}(x) = \begin{cases} \frac{p_{u_k}\lambda'_{K'}}{(\boldsymbol{\gamma}'\mathbf{a}')(\mathbf{u})+\lambda'_{K'}} \prod_{i=0}^{k-1}\left[p^m_{u_i}\left(\frac{(\boldsymbol{\gamma}'\mathbf{a}')(\mathbf{u}|_i)}{(\boldsymbol{\gamma}'\mathbf{a}')(\mathbf{u}|_i)+\lambda'_{K'}}\right)\right], & x = \mathbf{u} \in [D_m]^{k+1}, 0 \leqslant k < K'; \\ p^m_{u_{K'}} \prod_{i=0}^{K'-1}\left[p^m_{u_i}\left(\frac{(\boldsymbol{\gamma}'\mathbf{a}')(\mathbf{u}|_i)}{(\boldsymbol{\gamma}'\mathbf{a}')(\mathbf{u}|_i)+\lambda'_{K'}}\right)\right], & x = \mathbf{u} \in [D_m]^{K'+1}, \end{cases} \tag{3.42}$$

where we set the empty product of terms, when $k = 0$ equal to 1. In addition, we have

$$\mathcal{R}_{K'} = \frac{\mathcal{E}_{K'} + \mathcal{F}_{K'}}{\lambda'_{K'} - g^*_+}. \tag{3.43}$$

Proof. First note that, for each $u_0 \in [D_m]$, since we add a ball of type u_0 with probability $p^m_{u_0}$ at each time-step, and remove such a ball with probability proportional to $(\boldsymbol{\gamma}'\mathbf{a}')(u_0)$, we have

$$\lambda'_{K'}\mathbf{V}_{K'}(u_0) = p^m_{u_0} - (\boldsymbol{\gamma}'\mathbf{a}')(u_0)\mathbf{V}_{K'}(u_0), \tag{3.44}$$

this implies the case $k = 0$ in (3.42). Next, for $k > 0$, we have

$$\lambda'_{K'}\mathbf{V}_{K'}(\mathbf{u}) = \begin{cases} p^m_{u_k}(\boldsymbol{\gamma}'\mathbf{a}')(\mathbf{u}|_{k-1})\mathbf{V}_{K'}(\mathbf{u}|_{k-1}) - (\boldsymbol{\gamma}'\mathbf{a}')(\mathbf{u})\mathbf{V}_{K'}(\mathbf{u}), & \mathbf{u} \in [D_m]^{k+1},\ k < K'; \\ p^m_{u_{K'}}(\boldsymbol{\gamma}'\mathbf{a}')(\mathbf{u}|_{K'-1})\mathbf{V}_{K'}(\mathbf{u}|_{K'-1}); & \mathbf{u} \in [D_m]^{K'+1}; \end{cases} \tag{3.45}$$

so that, if $\mathbf{u} \in [D_m]^{k+1}$, $1 \leqslant k \leqslant K' - 1$,

$$\mathbf{V}_{K'}(\mathbf{u}) = \frac{p^m_{u_k}(\boldsymbol{\gamma}'\mathbf{a}')(\mathbf{u}|_{k-1})\mathbf{V}_{K'}(\mathbf{u}|_{k-1})}{(\boldsymbol{\gamma}'\mathbf{a}')(\mathbf{u}) + \lambda'_{K'}}. \tag{3.46}$$

Applying (3.45) and (3.46), recursing backwards, and using the fact that $\mathbf{V}_{K'}(u_0) = p^m_{u_0}/((\boldsymbol{\gamma}'\mathbf{a}')(u_0) + \lambda'_{K'})$ from (3.44), completes the proof of (3.42). Finally, for each $j \in [D_m]$, we have

$$\begin{aligned}\lambda'_{K'}\mathbf{V}_{K'}((D_m+1, j)) =& p^m_j \left(\sum_{\ell=1}^{D_m} \mathbf{a}'((D_m+1,\ell))\mathbf{V}_{K'}((D_m+1,\ell)) \right. \\ & \left. + \sum_{\mathbf{u}:\dim \mathbf{u} \leqslant K'} (\mathbf{a}' - \boldsymbol{\gamma}'\mathbf{a}')(\mathbf{u})\mathbf{V}_{K'}(\mathbf{u}) + \sum_{\mathbf{v}:\dim \mathbf{v} = K'+1} \mathbf{a}'(\mathbf{v})\mathbf{V}_{K'}(\mathbf{v}) \right) \\ =& p^m_j \left(\mathcal{R}_{K'} + \mathcal{E}_{K'} + \mathcal{F}_{K'} \right); \end{aligned} \tag{3.47}$$

where, in the last equation we recall the definitions in (3.40) and (3.41). Now, multiplying both sides of (3.47) by $\mathbf{a}'((D_m+1,j)) = g^*(j)$ and taking the sum over j, we have

$$\lambda'_{K'}\mathcal{R}_{K'} = \left(\sum_{j=1}^{D_m} p^m_j g^*(j) \right) \left(\mathcal{R}_{K'} + \mathcal{E}_{K'} + \mathcal{F}_{K'} \right) = \tilde{g}^*_+ \left(\mathcal{R}_{K'} + \mathcal{E}_{K'} + \mathcal{F}_{K'} \right).$$

Rearranging this proves (3.43), thus completing the proof of the proposition. □

Now, we recall the definition of the *companion process* $(S_i(w))_{i \geqslant 0}$ from Section 3.1.1 in (3.2): Recall that $W_1, W_2, \ldots$ were defined to be independent μ-distributed random variables and let $w \in [0, w^*]$. We then defined the random process $(S_i(w))_{i \geqslant 0}$ inductively so that $S_0(w) = h(w)$ and for all $i \geqslant 0$, we have $S_{i+1}(w) = S_i(w) + g(w, W_{i+1})$. Now, we also define the *lower companion process* $(S^-_i(w))_{i \geqslant 0}$ in a similar way, but instead with functions h^-, g^- respectively, so that

$$S^-_0(w) := h^-(w); \quad S^-_{i+1}(w) := S^-_i(w) + g^-(w, W_{i+1}),\ i \geqslant 0. \tag{3.48}$$

Lemma 3.2.16. *Assume Conditions **C1** and **C2**. Then we have*

$$\lim_{K' \to \infty} \lim_{m \to \infty} \mathcal{F}_{K'} = 0.$$

Proof. Note that by (3.42), with J' being an upper bound on $\max\{h, g\}$, we have

$$\begin{aligned}
\mathcal{F}_{K'} &= \sum_{\mathbf{u}:\dim \mathbf{u}=K'+1} \mathbf{a}'(\mathbf{u})\mathbf{V}_{K'}(\mathbf{u}) \\
&= \sum_{\mathbf{u}:\dim \mathbf{u}=K'+1} \mathbf{a}'(\mathbf{u})\, p^m_{u_{K'}} \prod_{i=0}^{K'-1} \left[p^m_{u_i} \left(\frac{(\boldsymbol{\gamma}'\mathbf{a}')(\mathbf{u}|_i)}{(\boldsymbol{\gamma}'\mathbf{a}')(\mathbf{u}|_i) + \lambda'_{K'}} \right) \right] \\
&\leqslant J'(K'+1) \cdot \sum_{\mathbf{u}:\dim \mathbf{u}=K'+1} p^m_{u_{K'}} \prod_{i=0}^{K'-1} \left[p^m_{u_i} \left(\frac{(\boldsymbol{\gamma}'\mathbf{a}')(\mathbf{u}|_i)}{(\boldsymbol{\gamma}'\mathbf{a}')(\mathbf{u}|_i) + \lambda'_{K'}} \right) \right] \\
&= J'(K'+1) \cdot \sum_{\mathbf{u}:\dim \mathbf{u}=K'} \left(\sum_{u_{K'} \in [D_m]} p^m_{u_{K'}} \right) \prod_{i=0}^{K'-1} \left[p^m_{u_i} \left(\frac{(\boldsymbol{\gamma}'\mathbf{a}')(\mathbf{u}|_i)}{(\boldsymbol{\gamma}'\mathbf{a}')(\mathbf{u}|_i) + \lambda'_{K'}} \right) \right] \\
&= J'(K'+1) \cdot \sum_{\mathbf{u}:\dim \mathbf{u}=K'} \prod_{i=0}^{K'-1} \left[p^m_{u_i} \left(\frac{(\boldsymbol{\gamma}'\mathbf{a}')(\mathbf{u}|_i)}{(\boldsymbol{\gamma}'\mathbf{a}')(\mathbf{u}|_i) + \lambda'_{K'}} \right) \right] \\
&= J'(K'+1) \cdot \mathbb{E}\left[\prod_{i=0}^{K'-1} \left(\frac{S^-_i(W)}{S^-_i(W) + \lambda'_{K'}} \right) \right],
\end{aligned}$$

where we recall the definition of $(S^-_i(w))_{i \geqslant 0}$ from (3.48). Now, note that for all $m \in \mathbb{N}$, $S^-(W)$ is stochastically bounded above by $S(W)$, and by Theorem 3.1.1 and (3.33) and (3.37), $\lambda'_{K'}$ is bounded below by λ^* uniformly in m and K'. Therefore, since the function $x \mapsto \frac{x}{x+\lambda}$ is increasing in x and decreasing in λ, we may bound the previous display above so that

$$\begin{aligned}
J'(K'+1) \cdot \mathbb{E}\left[\prod_{i=0}^{K'-1} \left(\frac{S^-_i(W)}{S^-_i(W) + \lambda'_{K'}} \right) \right] &\leqslant J'(K'+1) \cdot \mathbb{E}\left[\prod_{i=0}^{K'-1} \left(\frac{S_i(W)}{S_i(W) + \lambda'_{K'}} \right) \right] \\
&\leqslant J'(K'+1) \cdot \mathbb{E}\left[\prod_{i=0}^{K'-1} \left(\frac{S_i(W)}{S_i(W) + \lambda^*} \right) \right].
\end{aligned}$$

We complete the proof by proving the following claim.

Claim 3.2.17. *We have*

$$\lim_{k \to \infty} k \cdot \mathbb{E}\left[\prod_{i=0}^{k-1} \left(\frac{S_i(W)}{S_i(W) + \lambda^*} \right) \right] = 0$$

Proof. First observe that

$$\mathbb{E}\left[\prod_{i=0}^{\infty} \left(\frac{S_i(W)}{S_i(W) + \lambda^*} \right) \right] \leqslant \prod_{i=1}^{\infty} \left(\frac{J'i}{J'i + \lambda^*} \right) = \prod_{i=0}^{\infty} \left(1 - \frac{\lambda^*}{J'i + \lambda^*} \right) \leqslant e^{-\sum_{i=1}^{\infty} \frac{\lambda^*}{J'i+\lambda^*}} = 0.$$

Therefore, we have

$$k \cdot \mathbb{E}\left[\prod_{i=0}^{k-1}\left(\frac{S_i(W)}{S_i(W)+\lambda^*}\right)\right] = k \cdot \sum_{j=k}^{\infty} \mathbb{E}\left[\left(1-\frac{S_j(W)}{S_j(W)+\lambda^*}\right)\prod_{i=0}^{j-1}\left(\frac{S_i(W)}{S_i(W)+\lambda^*}\right)\right]$$
$$= k \cdot \sum_{j=k}^{\infty} \mathbb{E}\left[\frac{\lambda^*}{S_j(W)+\lambda^*}\prod_{i=0}^{j-1}\left(\frac{S_i(W)}{S_i(W)+\lambda^*}\right)\right]$$
$$\leqslant \sum_{j=k}^{\infty} j \cdot \mathbb{E}\left[\frac{\lambda^*}{S_j(W)+\lambda^*}\prod_{i=0}^{j-1}\left(\frac{S_i(W)}{S_i(W)+\lambda^*}\right)\right].$$

The series on the right of the previous display consists of non-negative terms, and for each $N \in \mathbb{N}$, we have

$$\sum_{j=1}^{N} j \cdot \mathbb{E}\left[\frac{\lambda^*}{S_j(W)+\lambda^*}\prod_{i=0}^{j-1}\left(\frac{S_i(W)}{S_i(W)+\lambda^*}\right)\right] \tag{3.49}$$
$$= \sum_{j=1}^{N}\left(j \cdot \mathbb{E}\left[\prod_{i=0}^{j-1}\left(\frac{S_i(W)}{S_i(W)+\lambda^*}\right)\right] - j \cdot \mathbb{E}\left[\prod_{i=0}^{j}\left(\frac{S_i(W)}{S_i(W)+\lambda^*}\right)\right]\right)$$
$$= \sum_{j=1}^{N} \mathbb{E}\left[\prod_{i=0}^{j-1}\left(\frac{S_i(W)}{S_i(W)+\lambda^*}\right)\right] - N \cdot \mathbb{E}\left[\prod_{i=0}^{N}\left(\frac{S_i(W)}{S_i(W)+\lambda^*}\right)\right]$$
$$\leqslant \sum_{j=1}^{N} \mathbb{E}\left[\prod_{i=0}^{j-1}\left(\frac{S_i(W)}{S_i(W)+\lambda^*}\right)\right].$$

Now, note that by Lemma 3.1.4, we have

$$\sum_{j=1}^{\infty} \mathbb{E}\left[\prod_{i=0}^{j-1}\left(\frac{S_i(W)}{S_i(W)+\lambda^*}\right)\right] < \infty,$$

and thus by (3.49) and the monotone convergence theorem, we also have

$$\sum_{j=1}^{\infty} j \cdot \mathbb{E}\left[\frac{\lambda^*}{S_j(W)+\lambda^*}\prod_{i=0}^{j-1}\left(\frac{S_i(W)}{S_i(W)+\lambda^*}\right)\right] < \infty.$$

Therefore,

$$\lim_{k\to\infty} k \cdot \mathbb{E}\left[\prod_{i=0}^{k-1}\left(\frac{S_i(W)}{S_i(W)+\lambda^*}\right)\right] \leqslant \lim_{k\to\infty} \sum_{j=k}^{\infty} j \cdot \mathbb{E}\left[\frac{\lambda^*}{S_j(W)+\lambda^*}\prod_{i=0}^{j-1}\left(\frac{S_i(W)}{S_i(W)+\lambda^*}\right)\right] = 0.$$

□

□

Lemma 3.2.18. *Assume Conditions **C1** and **C2**. Then we have*

$$\lim_{K'\to\infty}\lim_{m\to\infty}\mathcal{E}_{K'}=0,\qquad \textit{and}\qquad \lim_{K'\to\infty}\lim_{m\to\infty}\mathcal{R}_{K'}=0. \tag{3.50}$$

In addition,

$$\lim_{K'\to\infty}\lim_{m\to\infty}\lambda'_{K'}=\lambda^{*}. \tag{3.51}$$

Proof. The proof is similar to that of Lemma 3.2.10. First, let $\varepsilon>0$ be given, and, by Lemma 3.2.4, let m be sufficiently large that for all $x,y\in[0,w^{*}]$

$$\big(g^{+}(x,y)-g^{-}(x,y)\big)<\frac{\varepsilon\lambda'_{K'}}{K'}\ \text{ and }\ \big(h^{+}(x)-h^{-}(x)\big)<\frac{\varepsilon\lambda'_{K'}}{K'}. \tag{3.52}$$

The inequalities in (3.52) now imply that for any $\mathbf{u}=(u_0,\ldots,u_{K'-1})\in[D_m]^{K'}$, and each $i\in\{0\}\cup[K'-1]$ we have (taking the empty sum to be zero when $i=0$)

$$\begin{aligned}(\mathbf{a}'-\boldsymbol{\gamma}'\mathbf{a}')(\mathbf{u}|_i)&=h_{\max}(u_0)-h_{\min}(u_0)+\sum_{j=1}^{i-1}\left(g_{\max}(u_0,u_j)-g_{\min}(u_0,u_j)\right)\\&<\frac{\varepsilon\lambda'_{K'}}{K'}\cdot K'=\varepsilon\lambda'_{K'}\end{aligned} \tag{3.53}$$

Now, using the $\mathbf{u}|_i$ notation as a shorthand, we can write

$$\begin{aligned}\mathcal{E}_{K'}&=\sum_{\mathbf{u}\in[D_m]^{K'}}\sum_{i=0}^{K'-1}\left((\mathbf{a}'-\boldsymbol{\gamma}'\mathbf{a}')(\mathbf{u}|_i)\right)\mathbf{V}_{K'}(\mathbf{u}|_i)\\&\overset{(3.42)}{=}\sum_{\mathbf{u}\in[D_m]^{K'}}\sum_{i=0}^{K'-1}\frac{\left((\mathbf{a}'-\boldsymbol{\gamma}'\mathbf{a}')(\mathbf{u}|_i)\right)p^m_{u_i}}{(\boldsymbol{\gamma}'\mathbf{a}')(\mathbf{u}|_i)+\lambda'_{K'}}\prod_{j=0}^{i-1}\left[p^m_{u_j}\left(\frac{(\boldsymbol{\gamma}'\mathbf{a}')(\mathbf{u}|_j)}{(\boldsymbol{\gamma}'\mathbf{a}')(\mathbf{u}|_j)+\lambda'_{K'}}\right)\right]\\&\overset{(3.53)}{\leqslant}\varepsilon\cdot\sum_{\mathbf{u}\in[D_m]^{K'}}\sum_{i=0}^{K'-1}\frac{\lambda'_{K'}p^m_{u_i}}{(\boldsymbol{\gamma}'\mathbf{a}')(\mathbf{u}|_i)+\lambda'_{K'}}\prod_{j=0}^{i-1}\left[p^m_{u_j}\left(\frac{(\boldsymbol{\gamma}'\mathbf{a}')(\mathbf{u}|_j)}{(\boldsymbol{\gamma}'\mathbf{a}')(\mathbf{u}|_j)+\lambda'_{K'}}\right)\right]\\&=\varepsilon\cdot\mathbb{E}\left[\sum_{i=0}^{K'-1}\frac{\lambda'_{K'}}{S^{-}_i(W)+\lambda'_{K'}}\prod_{j=0}^{i-1}\frac{S^{-}_j(W)}{S^{-}_j(W)+\lambda'_{K'}}\right]<\varepsilon,\end{aligned}$$

where we recall the definition of $(S^{-}_j(w))_{j\geqslant 0}$ from (3.48), and observe that the sum in the final line of the display telescopes. The first equation in (3.50) follows. Next, (3.43), Lemma 3.2.16, and the facts that $\lambda'_{K'} \geqslant \lambda^*$ and $\lim_{m\to\infty} \tilde{g}^*_+ = \tilde{g}^* < \lambda^*$ together imply the second limit in (3.50). Finally, by (3.37), Proposition 3.2.15 and Theorem 3.1.1 we have

$$\lambda'_{K'} - \lambda^* \leqslant \mathcal{E}_{K'} + \mathcal{F}_{K'} + \mathcal{R}_{K'},$$

so that (3.51) follows by taking limits as $m \to \infty$ and $K' \to \infty$. □

Proof of Theorem 3.1.3

Proof of Theorem 3.1.3. First, recalling the definition of $\mathscr{D}_{\geqslant k}(n, \cdot)$ from (3.34), by Proposition 3.2.15 for any $\ell \in [D_m]$ we have

$$\begin{aligned}\lim_{n\to\infty} \frac{\mathscr{D}_{\geqslant k}(n,\ell)}{n} &= \sum_{j=k}^{K'} \sum_{\mathbf{u}\in[D_m]^{K'+1}} \mathbf{V}_{K'}(\mathbf{u}|_j)\mathbf{1}_{\{\ell\}}(u_0) \\ &= \sum_{\mathbf{u}\in[D_m]^{K'+1}} \left(p^m_{u_{K'}} \prod_{i=0}^{K'-1} \left[p^m_{u_i} \left(\frac{(\boldsymbol{\gamma}'\mathbf{a}')(\mathbf{u}|_i)}{(\boldsymbol{\gamma}'\mathbf{a}')(\mathbf{u}|_i) + \lambda'_{K'}} \right) \right] \right. \\ &\qquad \left. + \sum_{j=k}^{K'-1} \frac{p^m_{u_j} \lambda'_{K'}}{(\boldsymbol{\gamma}'\mathbf{a}')(\mathbf{u}|_j) + \lambda'_{K'}} \prod_{i=0}^{j-1} \left[p^m_{u_i} \left(\frac{(\boldsymbol{\gamma}'\mathbf{a}')(\mathbf{u}|_i)}{(\boldsymbol{\gamma}'\mathbf{a}')(\mathbf{u}|_i) + \lambda'_{K'}} \right) \right] \right) \mathbf{1}_{\{\ell\}}(u_0).\end{aligned}$$

Now, as with the proofs of Lemma 3.2.16 and Lemma 3.2.18, recalling the definition of $(S^-_i(w))_{i\geqslant 0}$ from (3.48), we may write the last equation as

$$\begin{aligned} &= \mathbb{E}\left[\prod_{i=0}^{K'-1} \left(\frac{S^-_i(W)}{S^-_i(W) + \lambda'_{K'}} \right) \mathbf{1}_{\mathcal{I}^m_\ell}(W) \right] \\ &\qquad + \sum_{j=k}^{K'-1} \mathbb{E}\left[\frac{\lambda'_{K'}}{S^-_j(W) + \lambda'_{K'}} \prod_{i=0}^{j-1} \left(\frac{S^-_i(W)}{S^-_i(W) + \lambda'_{K'}} \right) \mathbf{1}_{\mathcal{I}^m_\ell}(W) \right] \\ &= \mathbb{E}\left[\prod_{i=0}^{k-1} \left(\frac{S^-_i(W)}{S^-_i(W) + \lambda'_{K'}} \right) \mathbf{1}_{\mathcal{I}^m_\ell}(W) \right]. \end{aligned} \tag{3.54}$$

For $m' \in \mathbb{N}$, (3.54) allows us to prove the result for sets $S \in \sigma(\mathscr{I}^{m'})$, where we recall the definition of $\mathscr{I}^{m'}$ in (3.13), and (3.29) and (3.30). Since $N(n, \cdot)$ is finitely additive, if

$S \in \sigma(\mathscr{I}^m)$, by (3.35) and (3.54) we have

$$\mathbb{E}\left[\prod_{i=0}^{k-1}\left(\frac{S_i^-(W)}{S_i^-(W) + \lambda'_{K'}}\right)\mathbf{1}_S(W)\right] \leqslant \liminf_{n\to\infty}\frac{N_{\geqslant k}(n,S)}{n} \leqslant \limsup_{n\to\infty}\frac{N_{\geqslant k}(n,S)}{n}$$
$$\leqslant \mathbb{E}\left[\prod_{i=0}^{k-1}\left(\frac{S_i^-(W)}{S_i^-(W) + \lambda'_{K'}}\right)\mathbf{1}_S(W)\right] + \mathcal{R}_{K'} + \mathcal{E}_{K'} + \mathcal{F}_{K'}.$$

Taking limits as $m \to \infty$ and then as $K' \to \infty$, and applying Lemma 3.2.16 and Lemma 3.2.18 now proves the result for sets in $\sigma(\mathscr{I}^{m'})$. Now, note that for each $k \in \mathbb{N}_0$, and measurable sets $S' \subseteq [0, w^*]$, we have

$$\limsup_{n\to\infty}\frac{N_{\geqslant k}(n,S')}{n} \leqslant \limsup_{n\to\infty}\frac{N_{\geqslant 0}(n,S')}{n} = \mu(S') \quad \text{almost surely,} \tag{3.55}$$

where the last equality applies the strong law of large numbers.

We now prove the result for sets $U \in \mathcal{O}$ where $\mathcal{O}$ denotes the class of all open subsets of $[0, w^*]$. For a fixed open set $U \in \mathcal{O}$, and $m \in \mathbb{N}$, recall that $\mathcal{I}^m(U) := \bigcup_{j\in[D_m]:\mathcal{I}_j^m\subseteq U}\mathcal{I}_j^m$. Also recall (3.32), which states that $\mathbf{1}_{\mathcal{I}^m(U)}(W) \uparrow \mathbf{1}_U(W)$ pointwise as $m \to \infty$. Now, since each $\mathcal{I}^m(U) \in \sigma(\mathscr{I}^m)$, by applying (3.55) for each $k \leqslant K'$ we have

$$\mathbb{E}\left[\prod_{i=0}^{k-1}\left(\frac{S_i(W)}{S_i(W) + \lambda'_{K'}}\right)\mathbf{1}_{\mathcal{I}^m(U)}(W)\right] \leqslant \liminf_{n\to\infty}\frac{N_{\geqslant k}(n,U)}{n} \leqslant \limsup_{n\to\infty}\frac{N_{\geqslant k}(n,U)}{n}$$
$$\leqslant \mathbb{E}\left[\prod_{i=0}^{k-1}\left(\frac{S_i(W)}{S_i(W) + \lambda'_{K'}}\right)\mathbf{1}_{\mathcal{I}^m(U)}(W)\right] + \mu(U\backslash\mathcal{I}^m(U)).$$

Taking limits as $m \to \infty$ and then $K' \to \infty$ now proves the result for sets belonging to $\mathcal{O}$.

Finally, note that since μ is a *regular* measure, for any measurable set $A \subseteq [0, w^*]$ we have

$$\mu(A) = \inf_{U\in\mathcal{O}:A\subseteq U}\{\mu(U)\}.$$

Thus, for a given measurable set A, and any $\varepsilon > 0$, there exists an open set U_ε such that

$$\mu(U_\varepsilon\backslash A) \leqslant \varepsilon.$$

Therefore by finite additivity and (3.55)

$$\lim_{n\to\infty}\frac{N_{\geqslant k}(n,U_\varepsilon)}{n} - \varepsilon \leqslant \liminf_{n\to\infty}\frac{N_{\geqslant k}(n,A)}{n} \leqslant \limsup_{n\to\infty}\frac{N_{\geqslant k}(n,A)}{n} \leqslant \lim_{n\to\infty}\frac{N_{\geqslant k}(n,U_\varepsilon)}{n}.$$

The proof for the general case now follows by applying the result for the class $\mathcal{O}$, and sending $\varepsilon \to 0$. □

Theorem 3.1.3 now allows us to prove Theorem 3.1.5.

Proof of Theorem 3.1.5

The proof of this theorem is almost identical to that of Theorem 2.2.2 in Chapter 2. Recall that, if $N_k(n, A)$ denotes the number of vertices of out-degree k in the tree at time n having weight in A, by counting the edges in the tree in two ways we have

$$\Xi(n, A) = \sum_{k=1}^{n} k N_k(n, A) = \sum_{k=1}^{n} N_{\geqslant k}(n, A).$$

Proof of Theorem 3.1.5. By Lemma 3.1.4, and using Fatou's Lemma in the last inequality, we have,

$$\begin{aligned}(\psi_*)\mu(A) = \mathbb{E}\left[\frac{h(W)}{\lambda^* - \tilde{g}(W)} \mathbf{1}_A(W)\right] &= \sum_{k=1}^{\infty} \mathbb{E}\left[\prod_{i=0}^{k-1}\left(\frac{S_i(W)}{S_i(W) + \lambda^*}\right) \mathbf{1}_A(W)\right] \\ &= \sum_{k=1}^{\infty} \liminf_{n\to\infty} \frac{N_{\geqslant k}(n, A)}{n} \leqslant \liminf_{n\to\infty} \frac{\Xi(n, A)}{n};\end{aligned}$$

and likewise, $\liminf_{n\to\infty} \frac{\Xi(n, A^c)}{n} \geqslant (\psi_*\mu)(A^c)$. Now, since we add one edge at each time-step, it follows that $\Xi(n, [0, w^*]) = n$. Thus, by finite additivity,

$$\begin{aligned}1 = \liminf_{n\to\infty}\left(\frac{\Xi(n, A)}{n} + \frac{\Xi(n, A^c)}{n}\right) &\leqslant \limsup_{n\to\infty} \frac{\Xi(n, A)}{n} + \liminf_{n\to\infty} \frac{\Xi(n, A^c)}{n} \\ &\leqslant \limsup_{n\to\infty}\left(\frac{\Xi(n, A)}{n} + \frac{\Xi(n, A^c)}{n}\right) = 1.\end{aligned}$$

But, since (2.4) implies that $(\psi_*\mu)(\cdot)$ is a probability measure, this is only possible if

$$\limsup_{n\to\infty} \frac{\Xi(n, A)}{n} = (\psi_*\mu)(A) \text{ and } \liminf_{n\to\infty} \frac{\Xi(n, A^c)}{n} = (\psi_*\mu)(A^c) \text{ almost surely.}$$

The result follows. □

3.3 The Condensation Regime

In this section, we extend the results of Section 3.2 to the condensation regime. This section is closely related to Section 2.3.2 of Chapter 2, and indeed, Lemma 3.3.2 should be viewed as the analogue of Lemma 2.3.2, as we also couple the PANI-tree process $\mathcal{T}$ with auxiliary processes $\mathcal{T}^{(\varepsilon)}, \mathcal{T}^{(-\varepsilon)}, \varepsilon > 0$. However, the coupling we present is a refinement: rather than constructing the trees with truncated weights as we did in Lemma 2.3.2, we instead use the *same* weights, but instead adjust the function g in the processes $\mathcal{T}^{(\varepsilon)}$ and $\mathcal{T}^{(-\varepsilon)}$.

In particular, given $\varepsilon > 0$, and $\mathcal{M}_\varepsilon$ as defined in (3.6), define the functions $g_\varepsilon, g_{-\varepsilon}$ such that

$$g_\varepsilon(p, q) := \mathbf{1}_{\mathcal{M}_\varepsilon^c}(p)g(p, q) + \mathbf{1}_{\mathcal{M}_\varepsilon}(p)g(x^*, q)$$

and

$$g_{-\varepsilon}(p, q) := \mathbf{1}_{\mathcal{M}_\varepsilon^c}(p)g(p, q) + \mathbf{1}_{\mathcal{M}_\varepsilon}(p)(g(x^*, q) - u_\varepsilon(q));$$

and let $\mathcal{T}^{(\varepsilon)}, \mathcal{T}^{(-\varepsilon)}$ be the evolving trees with measure μ, and associated functions g_ε, h and $g_{-\varepsilon}, h$ respectively. We also denote by $(\mathcal{Z}_n^{(\varepsilon)})_{n\geqslant 0}$ and $(\mathcal{Z}_n^{(-\varepsilon)})_{n\geqslant 0}$ the partition functions associated with $\mathcal{T}^{(\varepsilon)}, \mathcal{T}^{(-\varepsilon)}$, respectively.

Lemma 3.3.1. *Assume Conditions **D1-D4**. Then, for each $\varepsilon > 0$ sufficiently small, $\mathcal{T}^{(\varepsilon)}$ and $\mathcal{T}^{(-\varepsilon)}$ satisfy Conditions **C1** and **C2**. In addition, if $\lambda_\varepsilon, \lambda_{-\varepsilon}$ denote the Malthusian parameters associated with $\mathcal{T}^{(\varepsilon)}, \mathcal{T}^{(-\varepsilon)}$, then $\lambda_\varepsilon \downarrow \tilde{g}^*$ and $\lambda_{-\varepsilon} \uparrow \tilde{g}^*$ as $\varepsilon \downarrow 0$.*

Proof. First, since by **D2** g satisfies Condition **C2**, we have

$$g(x, y) = \kappa\left(\phi_1^{(1)}(x), \ldots, \phi_1^{(N)}(x), \phi_2^{(1)}(y), \ldots, \phi_2^{(N)}(y)\right),$$

for measurable functions $\phi_j^i : [0, w^*] \to [0, J]$, $j = 1, 2$, $i \in [N]$ and a bounded continuous function $\kappa : [0, J]^{2N} \to \mathbb{R}_+$. Now, if we set $\phi_1^{(N+1)}(x) := \mathbf{1}_{\mathcal{M}_\varepsilon}(x), \phi_1^{(N+2)}(x) := \mathbf{1}_{\mathcal{M}_\varepsilon^c}(x)$,

$\phi_2^{(N+1)}(y) := g(x^*, y) - u_\varepsilon(y)$ and define κ' such that

$$\kappa'(c_1, \ldots, c_{N+2}, d_1, \ldots d_{N+1}) := c_{N+2}\kappa(c_1, \ldots, c_N, d_1, \ldots, d_N) + c_{N+1}d_{N+1},$$

we clearly have that $\phi_1^{(N+1)}, \phi_1^{(N+2)}, \phi_2^{(N+1)}$ are bounded, non-negative measurable functions, and κ' is bounded and continuous, taking values in $\mathbb{R}_+$. Noting that

$$g_{-\varepsilon}(x, y) = \kappa'\left(\phi_1^{(1)}(x), \ldots, \phi_1^{(N+2)}(x), \phi_2^{(1)}(y), \ldots, \phi_2^{(N+1)}(y)\right),$$

it follows that $g_{-\varepsilon}$ satisfies Condition **C2**. The proof of **C2** for g_ε is similar.

For **C1**, since h is bounded, for sufficiently large $\lambda > \tilde{g}^*$, we have

$$\mathbb{E}\left[\frac{h(W)}{\lambda - \tilde{g}_\varepsilon(W)}\right] < 1.$$

Meanwhile, since, by Condition **D4**, $\mu(\mathcal{M}_\varepsilon) > 0$ and $\tilde{g}_\varepsilon(x) = \tilde{g}^*$ for any $x \in \mathcal{M}_\varepsilon$, by monotone convergence

$$\lim_{\lambda \downarrow \tilde{g}^*} \mathbb{E}\left[\frac{h(W)}{\lambda - \tilde{g}_\varepsilon(W)}\right] = \mathbb{E}\left[\frac{h(W)}{\tilde{g}^* - \tilde{g}_\varepsilon(W)}\right] = \infty.$$

Thus, by continuity in λ, Condition **C1** is satisfied for $\mathcal{T}^{(\varepsilon)}$. A similar argument also works for $\mathcal{T}^{(-\varepsilon)}$: if $\tilde{g}^*_{-\varepsilon}$ denotes the maximum value of $\tilde{g}_{-\varepsilon}(x)$, then this value is also attained on $\mathcal{M}_\varepsilon$ which has positive measure. If $\lambda_\varepsilon, \lambda_{-\varepsilon}$ denote the associated Malthusian parameters associated with the trees, then, for each $\varepsilon > 0$, $\lambda_\varepsilon > \tilde{g}^*$ and $\lambda_{-\varepsilon} > \tilde{g}^*_{-\varepsilon}$. Moreover, since g_ε is non-increasing pointwise as ε decreases, λ_ε is non-increasing in ε; likewise, $\lambda_{-\varepsilon}$ is non-decreasing in ε. Now, suppose $\lim_{\varepsilon \downarrow 0} \lambda_\varepsilon = \lambda_+ > \tilde{g}^*$. Then we may apply dominated convergence, and

$$1 = \lim_{\varepsilon \downarrow 0} \mathbb{E}\left[\frac{h(W)}{\lambda_\varepsilon - \tilde{g}_\varepsilon(W)}\right] = \mathbb{E}\left[\lim_{\varepsilon \downarrow 0} \frac{h(W)}{\lambda_\varepsilon - \tilde{g}_\varepsilon(W)}\right] = \mathbb{E}\left[\frac{h(W)}{\lambda_+ - \tilde{g}(W)}\right],$$

contradicting (3.5). The case for $\lambda_{-\varepsilon}$ follows identically. □

Lemma 3.3.2. *There exists a coupling $(\hat{\mathcal{T}}^{(-\varepsilon)}, \hat{\mathcal{T}}, \hat{\mathcal{T}}^{(\varepsilon)})$ of these processes such that, almost surely (on the coupling space), for all $n \in \mathbb{N}_0$,*

$$\mathcal{Z}_n^{(-\varepsilon)} \leqslant \mathcal{Z}_n \leqslant \mathcal{Z}_n^{(\varepsilon)}, \tag{3.56}$$

and, for each vertex v with $W_v \in \mathcal{M}_\varepsilon^c$, we have

$$f(N^+(v, \hat{\mathcal{T}}_n^{(\varepsilon)})) \leqslant f(N^+(v, \hat{\mathcal{T}}_n)) \leqslant f(N^+(v, \hat{\mathcal{T}}_n^{(-\varepsilon)})) \tag{3.57}$$

and

$$\deg(v, \hat{\mathcal{T}}_n^{(\varepsilon)}) \leqslant \deg(v, \hat{\mathcal{T}}_n) \leqslant \deg(v, \hat{\mathcal{T}}_n^{(-\varepsilon)}). \tag{3.58}$$

Proof. We initialise the trees with a single vertex 0 having weight W_0 sampled independently from μ, conditioned on $\{h(W_0) > 0\}$ and will construct copies of these three tree processes on the same vertex set, which is identified with $\mathbb{N}_0$. Now, assume that at the nth time-step,

$$(\hat{\mathcal{T}}_j^{(-\varepsilon)})_{0 \leqslant j \leqslant n} \sim (\hat{\mathcal{T}}_j^{(-\varepsilon)})_{0 \leqslant j \leqslant n}, \quad (\hat{\mathcal{T}}_j)_{0 \leqslant j \leqslant n} \sim (\mathcal{T}_j)_{0 \leqslant j \leqslant n} \quad \text{and} \quad (\hat{\mathcal{T}}_j^{(\varepsilon)})_{0 \leqslant j \leqslant n} \sim (\mathcal{T}_j^{(\varepsilon)})_{0 \leqslant j \leqslant n}.$$

In addition, assume that (3.56) and (3.57) are satisfied up to time n.

Now, for the $(n+1)$st step:

- Introduce vertex $n+1$ with weight W_{n+1} sampled independently from μ in $\hat{\mathcal{T}}_n^{(-\varepsilon)}, \hat{\mathcal{T}}_n$ and $\hat{\mathcal{T}}_n^{(\varepsilon)}$.
- Form $\hat{\mathcal{T}}_{n+1}^{(-\varepsilon)}$ by sampling the parent v of $n+1$ independently according to the law of $\mathcal{T}^{(-\varepsilon)}$, i.e., with probability proportional to $f(N^+(v, \hat{\mathcal{T}}_n^{(-\varepsilon)}))$. Then, in order to form $\hat{\mathcal{T}}_{n+1}$ sample an independent uniformly distributed random variables U_1 on $[0, 1]$.
  - If $U_1 \leqslant \frac{\mathcal{Z}_n^{(-\varepsilon)} f(N^+(v, \hat{\mathcal{T}}_n))}{\mathcal{Z}_n f(N^+(v, \hat{\mathcal{T}}_n^{(-\varepsilon)}))}$ and $W_v \in \mathcal{M}_\varepsilon^c$, select v as the parent of $n+1$ in $\hat{\mathcal{T}}_{n+1}$ as well.
  - Otherwise, form $\hat{\mathcal{T}}_{n+1}$ by selecting the parent v' of $n+1$ with probability proportional to $f(N^+(v', \hat{\mathcal{T}}_n))$ out of all all the vertices with weight $W_{v'} \in \mathcal{M}_\varepsilon$.
- Then form $\hat{\mathcal{T}}_{n+1}^{(\varepsilon)}$ in a similar manner. Sample an independent uniform random variable U_2 on $[0, 1]$.

- If vertex v (with weight $W_v \in \mathcal{M}_\varepsilon^c$) was chosen as the parent of $n+1$ in $\hat{\mathcal{T}}_{n+1}$ and $U_2 \leqslant \frac{\mathcal{Z}_n f(N^+(v,\hat{\mathcal{T}}_n^{(\varepsilon)}))}{\mathcal{Z}_n^{(\varepsilon)} f(N^+(v,\hat{\mathcal{T}}_n))}$, also select v as the parent of $n+1$ in $\hat{\mathcal{T}}_{n+1}^{\varepsilon}$.
- Otherwise, form $\hat{\mathcal{T}}_{n+1}^{(\varepsilon)}$ by selecting the parent v'' of $n+1$ with probability proportional to $f(N^+(v'', \mathcal{T}_n^{(\varepsilon)}))$ out of all the vertices with weight $W_{v''} \in \mathcal{M}_\varepsilon$.

Clearly $\hat{\mathcal{T}}_{n+1}^{(-\varepsilon)} \sim \mathcal{T}_{n+1}^{(-\varepsilon)}$. On the other hand, in $\hat{\mathcal{T}}_{n+1}$ the probability of choosing a certain parent v of $n+1$ with weight $W_v \in \mathcal{M}_\varepsilon^c$ is

$$\frac{\mathcal{Z}_n^{(-\varepsilon)} f(N^+(v,\hat{\mathcal{T}}_n))}{\mathcal{Z}_n f(N^+(v,\hat{\mathcal{T}}_n^{(-\varepsilon)}))} \times \frac{f(N^+(v,\hat{\mathcal{T}}_n^{(-\varepsilon)}))}{\mathcal{Z}_n^{(-\varepsilon)}} = \frac{f(N^+(v,\hat{\mathcal{T}}_n))}{\mathcal{Z}_n},$$

whilst the probability of choosing a parent v' with weight $W_{v'} \in \mathcal{M}_\varepsilon$ is

$$\begin{aligned}
&\frac{f(N^+(v',\hat{\mathcal{T}}_n))}{\sum_{v':W_{v'}\in\mathcal{M}_\varepsilon} f(N^+(v',\hat{\mathcal{T}}_n))} \left(\sum_{v:W_v\in\mathcal{M}_\varepsilon^c} \left(1 - \frac{\mathcal{Z}_n^{(-\varepsilon)} f(N^+(v,\hat{\mathcal{T}}_n))}{\mathcal{Z}_n f(N^+(v,\hat{\mathcal{T}}_n^{(-\varepsilon)}))}\right) \frac{f(N^+(v,\hat{\mathcal{T}}_n^{(-\varepsilon)}))}{\mathcal{Z}_n^{(-\varepsilon)}} \right) \\
&\qquad\qquad + \frac{f(N^+(v',\hat{\mathcal{T}}_n))}{\sum_{v':W_{v'}\in\mathcal{M}_\varepsilon} f(N^+(v',\hat{\mathcal{T}}_n))} \left(\sum_{v:W_v\in\mathcal{M}_\varepsilon} \frac{f(N^+(v,\hat{\mathcal{T}}_n^{(-\varepsilon)}))}{\mathcal{Z}_n^{(-\varepsilon)}} \right) \\
&\qquad = \frac{f(N^+(v',\hat{\mathcal{T}}_n))}{\sum_{v':W_{v'}\in\mathcal{M}_\varepsilon} f(N^+(v',\hat{\mathcal{T}}_n))} \left(\sum_{v} \frac{f(N^+(v,\hat{\mathcal{T}}_n^{(-\varepsilon)})))}{\mathcal{Z}_n^{(-\varepsilon)}} - \sum_{v:W_v\in\mathcal{M}_\varepsilon^c} \frac{f(N^+(v,\hat{\mathcal{T}}_n))}{\mathcal{Z}_n} \right) \\
&\qquad = \frac{f(N^+(v',\hat{\mathcal{T}}_n))}{\sum_{v':W_{v'}\in\mathcal{M}_\varepsilon} f(N^+(v',\hat{\mathcal{T}}_n))} \left(1 - \frac{\sum_{v:W_v\in\mathcal{M}_\varepsilon^c} f(N^+(v,\hat{\mathcal{T}}_n))}{\mathcal{Z}_n} \right) = \frac{f(N^+(v',\hat{\mathcal{T}}_n))}{\mathcal{Z}_n},
\end{aligned}$$

where we use the fact that $\sum_v f(N^+(v,\hat{\mathcal{T}}_n)) = \mathcal{Z}_n$. Thus, we have $\hat{\mathcal{T}}_{n+1} \sim \mathcal{T}_{n+1}$. Now, note that if the parent v of $n+1$ in $\hat{\mathcal{T}}_{n+1}^{(-\varepsilon)}$ is such that $W_v \in \mathcal{M}_\varepsilon^c$, the same parent is chosen in $\hat{\mathcal{T}}_{n+1}$. Since $W_v \in \mathcal{M}_\varepsilon^c$, we have

$$\begin{aligned}
f(N^+(v,\hat{\mathcal{T}}_{n+1}^{(-\varepsilon)})) - f(N^+(v,\hat{\mathcal{T}}_n^{(-\varepsilon)})) &= g_{-\varepsilon}(W_v, W_{n+1}) = g(W_v, W_{n+1}) \\
&= f(N^+(v,\hat{\mathcal{T}}_{n+1})) - f(N^+(v,\hat{\mathcal{T}}_n)).
\end{aligned}$$

Otherwise, the parent of $n+1$ in $\hat{\mathcal{T}}_{n+1}$ has weight which belongs to $\mathcal{M}_\varepsilon$, and thus $f(N^+(v,\hat{\mathcal{T}}_n^{(-\varepsilon)}))$ increases whilst $f(N^+(v,\hat{\mathcal{T}}_n))$ stays the same. An increase in $f(N^+(v,\hat{\mathcal{T}}_n^{(-\varepsilon)}))$ coincides with the increase of $\deg(v,\hat{\mathcal{T}}_n^{(-\varepsilon)})$, and thus the right hand sides of (3.57) and (3.58) are satisfied for time $n+1$.

Now, note that

$$\mathcal{Z}^{(-\varepsilon)}_{n+1} - \mathcal{Z}^{(-\varepsilon)}_{n} = h(W_{n+1}) + g_{-\varepsilon}(W_v, W_{n+1}), \text{ and } \mathcal{Z}_{n+1} - \mathcal{Z}_n = h(W_{n+1}) + g(W_{v'}, W_{n+1}),$$

where v, v' denote the parent of $n+1$ in $\hat{\mathcal{T}}_n$ and $\hat{\mathcal{T}}^{(\varepsilon)}_n$ respectively. Then we either have:

- $v = v'$, so that $g_{-\varepsilon}(W_v, W_{n+1}) = g(W_{v'}, W_{n+1})$.
- $v \in \mathcal{M}^c_\varepsilon$ and $v' \in \mathcal{M}_\varepsilon$, in which case, $\mathbb{P}$-a.s, using **D4**

$$g_{-\varepsilon}(W_v, W_{n+1}) = g(W_v, W_{n+1}) \leqslant g(x^*, W_{n+1}) - u_\varepsilon(W_{n+1}) < g(W_{v'}, W_{n+1}).$$

- Both $v, v' \in \mathcal{M}_\varepsilon$, in which case, $\mathbb{P}$-a.s.,

$$g_{-\varepsilon}(W_v, W_{n+1}) = g(x^*, W_{n+1}) - u_\varepsilon(W_{n+1}) < g(W_{v'}, W_{n+1}).$$

In every case we have $\mathcal{Z}^{(-\varepsilon)}_{n+1} - \mathcal{Z}^{(-\varepsilon)}_{n} \leqslant \mathcal{Z}_{n+1} - \mathcal{Z}_n$, and thus (3.56) is also satisfied at time $n+1$.

Each of the statements concerning $\hat{\mathcal{T}}^{(\varepsilon)}$ follow in an analogous manner, applying Condition **D3**. □

3.3.1 Proof of Theorem 3.1.7

The proof of Theorem 3.1.7 uses the auxiliary trees $\mathcal{T}^{(\varepsilon)}$ and $\mathcal{T}^{(-\varepsilon)}$, and Lemma 3.3.2.

Proof of Theorem 3.1.7. For the first statement, note that by (3.56) in Lemma 3.3.2 and Theorem 3.1.1, for each $\varepsilon > 0$ we have, $\mathbb{P}$-a.s.,

$$\lambda_{-\varepsilon} = \lim_{n\to\infty} \frac{\mathcal{Z}^{(-\varepsilon)}_n}{n} \leqslant \liminf_{n\to\infty} \frac{\mathcal{Z}_n}{n} \leqslant \limsup_{n\to\infty} \frac{\mathcal{Z}_n}{n} = \lim_{n\to\infty} \frac{\mathcal{Z}^{(\varepsilon)}_n}{n} = \lambda_\varepsilon.$$

The statement follows by sending $\varepsilon \to 0$, using Lemma 3.3.1.

Next, by assumption, for each $\varepsilon > 0$ sufficiently small, we have $A \subseteq \mathcal{M}_{\varepsilon}^c$. Next, applying (3.58), if $\Xi^{(\varepsilon)}$ and $\Xi^{(-\varepsilon)}$ denote the edge distributions in the coupled trees $\hat{\mathcal{T}}^{(\varepsilon)}, \hat{\mathcal{T}}^{(-\varepsilon)}$, respectively, then for each $n \in \mathbb{N}_0$

$$\Xi^{(\varepsilon)}(n, A) \leqslant \Xi(n, A) \leqslant \Xi^{(-\varepsilon)}(n, A),$$

and thus, by Theorem 3.1.5, we have

$$\begin{aligned} \mathbb{E}\left[\frac{h(W)}{\lambda_{\varepsilon} - \tilde{g}_{\varepsilon}(W)} \mathbf{1}_A(W)\right] &\leqslant \liminf_{n \to \infty} \frac{\Xi(n, A)}{n} \\ &\leqslant \limsup_{n \to \infty} \frac{\Xi(n, A)}{n} \leqslant \mathbb{E}\left[\frac{h(W)}{\lambda_{-\varepsilon} - \tilde{g}_{-\varepsilon}(W)} \mathbf{1}_A(W)\right]. \end{aligned} \tag{3.59}$$

Now, noting that $\tilde{g}_{-\varepsilon} = \tilde{g} = \tilde{g}_{\varepsilon}$ on A, and $\lambda_{-\varepsilon} > \tilde{g}^*_{-\varepsilon} \geqslant \sup_{x \in A} \tilde{g}(x)$ and is non-decreasing in ε, by applying Lemma 3.3.1 and dominated convergence we have

$$\begin{aligned} \lim_{\varepsilon \to 0} \mathbb{E}\left[\frac{h(W)}{\lambda_{\varepsilon} - \tilde{g}_{\varepsilon}(W)} \mathbf{1}_A(W)\right] &= \lim_{\varepsilon \to 0} \mathbb{E}\left[\frac{h(W)}{\lambda_{-\varepsilon} - \tilde{g}_{-\varepsilon}(W)} \mathbf{1}_A(W)\right] \\ &= \mathbb{E}\left[\frac{h(W)}{\tilde{g}^* - \tilde{g}(W)} \mathbf{1}_A(W)\right]. \end{aligned} \tag{3.60}$$

Then, (3.7) follows by combining (3.59) and (3.60). Moreover, for each $\varepsilon' > 0$, by setting $A = \mathcal{M}_{\varepsilon'}^c$,

$$\lim_{n \to \infty} \frac{\Xi(n, \mathcal{M}_{\varepsilon'})}{n} = \lim_{n \to \infty} \left(1 - \frac{\Xi(n, \mathcal{M}_{\varepsilon'}^c)}{n}\right) = 1 - \mathbb{E}\left[\frac{h(W)}{\tilde{g}^* - \tilde{g}(W)} \mathbf{1}_{\mathcal{M}_{\varepsilon'}^c}(W)\right].$$

But then, again by dominated convergence,

$$\lim_{\varepsilon' \to 0} \mathbb{E}\left[\frac{h(W)}{\tilde{g}^* - \tilde{g}(W)} \mathbf{1}_{\mathcal{M}_{\varepsilon'}^c}(W)\right] = \mathbb{E}\left[\frac{h(W)}{\tilde{g}^* - \tilde{g}(W)}\right],$$

and (3.8) follows.

Finally, for the last statement, recall the definition of the companion process $(S_i)_{i \geqslant 0}$ in (3.2), and that, for any measurable $B \subseteq [0, w^*]$, $N_{\geqslant k}(n, B)$ denotes the number of vertices of out-degree at least k with weight belonging to B at time n. Then, for $\varepsilon > 0$, note that

$$\frac{N_{\geqslant k}(n, B \cap \mathcal{M}_{\varepsilon}^c)}{n} \leqslant \frac{N_{\geqslant k}(n, B)}{n} \leqslant \frac{N_{\geqslant k}(n, B \cap \mathcal{M}_{\varepsilon}^c)}{n} + \frac{N_{\geqslant 0}(n, \mathcal{M}_{\varepsilon})}{n}.$$

Now, by the strong law of large numbers, in the limit as $n \to \infty$, as in (3.55), the second quantity tends to $\mu(\mathcal{M}_\varepsilon)$, and thus,

$$\liminf_{n\to\infty} \frac{N_{\geqslant k}(n, B \cap \mathcal{M}_\varepsilon^c)}{n} \leqslant \limsup_{n\to\infty} \frac{N_{\geqslant k}(n, B)}{n} \leqslant \limsup_{n\to\infty} \frac{N_{\geqslant k}(n, B \cap \mathcal{M}_\varepsilon^c)}{n} + \mu(\mathcal{M}_\varepsilon). \tag{3.61}$$

Now, let $N_{\geqslant k}^{(-\varepsilon)}(n,\cdot), N_{\geqslant k}^{(\varepsilon)}(n,\cdot)$ denote the associated quantities in the trees $\mathcal{T}^{(-\varepsilon)}, \mathcal{T}^{(\varepsilon)}$, and denote by $(S_i^{(-\varepsilon)})_{i\geqslant 0}$ and $(S_i^{(\varepsilon)})_{i\geqslant 0}$ the companion processes defined in terms of the functions $h, g_{-\varepsilon}$ and $h, g_{+\varepsilon}$ respectively. Then, by (3.58), on the coupling in Lemma 3.3.2, we have

$$N_{\geqslant k}^{(\varepsilon)}(n, B \cap \mathcal{M}_\varepsilon^c) \leqslant N_{\geqslant k}(n, B \cap \mathcal{M}_\varepsilon^c) \leqslant N_{\geqslant k}^{(-\varepsilon)}(n, B \cap \mathcal{M}_\varepsilon^c).$$

Therefore, by Theorem 3.1.3, recalling the definitions of $\lambda_\varepsilon, \lambda_{-\varepsilon}$ in Lemma 3.3.1,

$$\begin{aligned}\mathbb{E}\left[\prod_{i=0}^{k-1}\left(\frac{S_i^{(\varepsilon)}(W)}{S_i^{(\varepsilon)}(W)+\lambda_\varepsilon}\right)\mathbf{1}_{B\cap\mathcal{M}_\varepsilon^c}(W)\right] &\leqslant \liminf_{n\to\infty}\frac{N_{\geqslant k}(n, B \cap \mathcal{M}_\varepsilon^c)}{n}\\ &\leqslant \limsup_{n\to\infty}\frac{N_{\geqslant k}(n, B \cap \mathcal{M}_\varepsilon^c)}{n}\\ &\leqslant \mathbb{E}\left[\prod_{i=0}^{k-1}\left(\frac{S_i^{(-\varepsilon)}(W)}{S_i^{(-\varepsilon)}(W)+\lambda_{-\varepsilon}}\right)\mathbf{1}_{B\cap\mathcal{M}_\varepsilon^c}(W)\right],\end{aligned}$$

and thus, by (3.61), we have

$$\begin{aligned}\mathbb{E}\left[\prod_{i=0}^{k-1}\left(\frac{S_i^{(\varepsilon)}(W)}{S_i^{(\varepsilon)}(W)+\lambda_\varepsilon}\right)\mathbf{1}_{B\cap\mathcal{M}_\varepsilon^c}(W)\right] &\leqslant \liminf_{n\to\infty}\frac{N_{\geqslant k}(n, B)}{n}\\ &\leqslant \limsup_{n\to\infty}\frac{N_{\geqslant k}(n, B)}{n}\\ &\leqslant \mathbb{E}\left[\prod_{i=0}^{k-1}\left(\frac{S_i^{(-\varepsilon)}(W)}{S_i^{(-\varepsilon)}(W)+\lambda_{-\varepsilon}}\right)\mathbf{1}_{B\cap\mathcal{M}_\varepsilon^c}(W)\right] + \mu(\mathcal{M}_\varepsilon).\end{aligned} \tag{3.62}$$

Now, by dominated convergence, as $\varepsilon \to 0$

$$\mathbb{E}\left[\prod_{i=0}^{k-1}\left(\frac{S_i^{(\varepsilon)}(W)}{S_i^{(\varepsilon)}(W)+\lambda_\varepsilon}\right)\mathbf{1}_{B\cap\mathcal{M}_\varepsilon}(W)\right] \to \mathbb{E}\left[\prod_{i=0}^{k-1}\left(\frac{S_i(W)}{S_i(W)+\tilde{g}^*}\right)\mathbf{1}_{B}(W)\right], \text{ and}$$
$$\mathbb{E}\left[\prod_{i=0}^{k-1}\left(\frac{S_i^{(-\varepsilon)}(W)}{S_i^{(-\varepsilon)}(W)+\lambda_{-\varepsilon}}\right)\mathbf{1}_{B\cap\mathcal{M}_\varepsilon}(W)\right] \to \mathbb{E}\left[\prod_{i=0}^{k-1}\left(\frac{S_i(W)}{S_i(W)+\tilde{g}^*}\right)\mathbf{1}_{B}(W)\right],$$

and, since, by (3.5), $\mathcal{M}$ is a μ-null set, $\mu(\mathcal{M}_\varepsilon) \to 0$. Combining these statements with (3.62) completes the proof. □

3.3.2 Proof of Corollary 3.1.8

Proof of Corollary 3.1.8. By the Portmanteau theorem, it suffices to show that, $\mathbb{P}$-a.s.

$$\lim_{n\to\infty} \frac{\Xi(n, A)}{n} = \Pi(A),$$

for any measurable set $A \subseteq [0, w^*]$ with $\mu(\partial A) = 0$. Now, since $\mu(\mathcal{M}) = 0$, it suffices to prove this equation for measurable sets $A \subseteq [0, w^*]$ with $\overline{A} \cap \mathcal{M} = \varnothing$. In view of Theorem 3.1.7, we need only show that for all $\varepsilon > 0$ sufficiently small, we have $\overline{A} \cap \mathcal{M}_{\varepsilon} = \varnothing$. Indeed, if this were not the case, then, since $(\overline{A} \cap \overline{\mathcal{M}}_{1/n})_{n\in\mathbb{N}}$ is a nested sequence of closed sets, by Cantor's intersection theorem,

$$\varnothing \neq \bigcap_{n\in\mathbb{N}} \left(\overline{A} \cap \overline{\mathcal{M}}_{1/n}\right) = \overline{A} \cap \bigcap_{n\in\mathbb{N}} \overline{\mathcal{M}}_{1/n} = \overline{A} \cap \mathcal{M},$$

a contradiction. □

3.4 A Generalised Geometric Series

3.4.1 Proof of Lemma 3.1.4

Lemma 3.1.4 may be interpreted as an extension of (2.17) in Section 2.3.1 of Chapter 2, where we proved an analogous result in regards to the companion process associated with the GPAF-tree. In that section, the approach was to apply the analysis of Section 2.2 in Chapter 2, computing the Laplace transform of an appropriate pure-jump process in two different ways. Here we adopt a slightly different approach: we also introduce an auxiliary piece-wise constant, continuous time Markov process but instead compute its expected value at an independent, exponentially distributed stopping time in two different ways.

More precisely, we define a process $(\mathcal{Y}_w(t), r_w(t))_{t\geqslant 0}$ taking values in $\mathbb{N} \times [0, \infty)$. Let $(W_i)_{i\geqslant 0}$ be independent μ-distributed random variables, and define $(S_i(w))_{i\geqslant 0}$ according to

(3.2), that is,

$$S_0(w) := h(w); \quad S_{i+1}(w) := S_i(w) + g(w, W_{i+1}),\ i \geqslant 0.$$

In addition, set $\tau_0 = 0$, and define $(\tau_i)_{i\geqslant 1}$ recursively so that, given $S_i(w)$

$$\tau_{i+1} - \tau_i \sim \mathrm{Exp}(S_i(w)); \tag{3.63}$$

where $\mathrm{Exp}(S_i(w))$ denotes the exponential distribution with parameter $S_i(w)$. Then, we set

$$\mathcal{Y}_w(t) := \sum_{n=1}^{\infty} \mathbf{1}_{[\tau_n,\infty)}(t), \quad \text{and} \quad r_w(t) := \sum_{n=0}^{\infty} S_n(w)\mathbf{1}_{[\tau_n,\tau_{n+1})}(t).$$

Now, let $(\mathcal{F}_t)_{t\geqslant 0}$ denote the filtration generated by the process $(\mathcal{Y}_w(t), r_w(t))_{t\geqslant 0}$.

Claim 3.4.1. *The process $\mathcal{Y}_w(t) - \int_0^t r_w(s)\mathrm{d}s$ is a martingale with respect to the filtration $(\mathcal{F}_t)_{t\geqslant 0}$.*

Proof. This follows from the fact that the difference between jump times is exponentially distributed, and by applying, for example, [44, Theorem 1.33, page 149]. □

In addition,

Claim 3.4.2. *For all $t \in [0,\infty)$, we have $\mathbb{E}[\mathcal{Y}_w(t)] < \infty$ almost surely. In particular, for each $t \in [0,\infty)$,*

$$\mathbb{E}[\mathcal{Y}_w(t)] = \int_0^t \mathbb{E}[r_w(s)]\,\mathrm{d}s. \tag{3.64}$$

Proof. Let α be an independent exponentially distributed random variable with parameter $a > 0$, and set $\mathcal{Y}_w(\alpha) := \inf_{t\geqslant\alpha}(\mathcal{Y}_w(t))$. Then,

$$\begin{aligned}\mathbb{E}\left[\mathbf{1}_{\{\mathcal{Y}_w(\alpha)\geqslant k\}} | S_{k-1}(w), \mathbf{1}_{\{\mathcal{Y}_w(\alpha)\geqslant k-1\}}\right] &= \mathbb{E}\left[\mathbf{1}_{\{\alpha\geqslant\tau_k\}} | S_{k-1}(w), \mathbf{1}_{\{\mathcal{Y}_w(\alpha)\geqslant k-1\}}\right] \\ &= \mathbb{P}\left(\min\left(\alpha - \tau_{k-1}, \tau_k - \tau_{k-1}\right) = \tau_k - \tau_{k-1} | S_{k-1}(w)\right) \\ &\qquad \times \mathbf{1}_{\{\mathcal{Y}_w(\alpha)\geqslant k-1\}} \\ &= \frac{S_{k-1}(w)}{a + S_{k-1}(w)} \mathbf{1}_{\{\mathcal{Y}_w(\alpha)\geqslant k-1\}}, \end{aligned} \tag{3.65}$$

where in the last equality we have used (3.63) and the memory-less property of the exponential distribution. Note also, that for any $j \leqslant k-1$, the random variables $(S_j(w), \dots, S_{k-1}(w))$ and $\mathbf{1}_{\{\mathcal{Y}_w(\alpha) \geqslant j\}}$ are conditionally independent given the random variables $S_{j-1}(w), \mathbf{1}_{\{\mathcal{Y}_w(\alpha) \geqslant j-1\}}$. Indeed, for each $\ell \in \{j, \dots, k-1\}$,

$$S_\ell(w) = S_{j-1}(w) + \sum_{i=j}^{\ell} g(w, W_i),$$

where $W_j, \dots, W_{k-1}$ are independent random variables sampled from μ, while

$$\mathbf{1}_{\{\mathcal{Y}_w(\alpha) \geqslant j\}} = \mathbf{1}_{\{\mathcal{Y}_w(\alpha) \geqslant j-1\}} \times \mathbf{1}_{\{\min(\tau_j - \tau_{j-1}, \alpha - \tau_{j-1}) = \tau_j - \tau_{j-1}\}},$$

where, we recall $\tau_j - \tau_{j-1}$ is an independent exponentially distributed random variable with parameter $S_{j-1}(w)$ and thus conditionally independent of $(S_j(w), \dots, S_{k-1}(w))$. As a result, we have

$$\begin{aligned} &\mathbb{E}\left[\left(\prod_{i=j}^{k-1} \frac{S_i(w)}{S_i(w)+a}\right) \mathbf{1}_{\{\mathcal{Y}_w(\alpha) \geqslant j\}} \middle| S_{j-1}(w), \mathbf{1}_{\{\mathcal{Y}_w(\alpha) \geqslant j-1\}}\right] \\ &= \mathbb{E}\left[\left(\prod_{i=j}^{k-1} \frac{S_i(w)}{S_i(w)+a}\right) \middle| S_{j-1}(w), \mathbf{1}_{\{\mathcal{Y}_w(\alpha) \geqslant j-1\}}\right] \mathbb{E}\left[\mathbf{1}_{\{\mathcal{Y}_w(\alpha) \geqslant j\}} \middle| S_{j-1}(w), \mathbf{1}_{\{\mathcal{Y}_w(\alpha) \geqslant j-1\}}\right]. \end{aligned} \tag{3.66}$$

Therefore, we have

$$\begin{aligned} \mathbb{P}(\mathcal{Y}_w(\alpha) \geqslant k) &= \mathbb{E}\left[\mathbf{1}_{\{\mathcal{Y}_w(\alpha) \geqslant k\}}\right] = \mathbb{E}\left[\mathbb{E}\left[\mathbf{1}_{\{\mathcal{Y}_w(\alpha) \geqslant k\}} | S_{k-1}(w), \mathbf{1}_{\{\mathcal{Y}_w(\alpha) \geqslant k-1\}}\right]\right] \\ &\overset{(3.65)}{=} \mathbb{E}\left[\frac{S_{k-1}(w)}{a + S_{k-1}(w)} \mathbf{1}_{\{\mathcal{Y}_w(\alpha) \geqslant k-1\}}\right] \\ &= \mathbb{E}\left[\mathbb{E}\left[\frac{S_{k-1}(w)}{a + S_{k-1}(w)} \mathbf{1}_{\{\mathcal{Y}_w(\alpha) \geqslant k-1\}} \middle| S_{k-2}(w), \mathbf{1}_{\{\mathcal{Y}_w(\alpha) \geqslant k-2\}}\right]\right] \\ &\overset{(3.66)}{=} \mathbb{E}\Bigg[\mathbb{E}\left[\frac{S_{k-1}(w)}{a + S_{k-1}(w)} \middle| S_{k-2}(w), \mathbf{1}_{\{\mathcal{Y}_w(\alpha) \geqslant k-2\}}\right] \\ &\qquad \times \mathbb{E}\left[\mathbf{1}_{\{\mathcal{Y}_w(\alpha) \geqslant k-1\}} | S_{k-2}(w), \mathbf{1}_{\{\mathcal{Y}_w(\alpha) \geqslant k-2\}}\right]\Bigg] \\ &\overset{(3.65)}{=} \mathbb{E}\left[\mathbb{E}\left[\frac{S_{k-1}(w)}{a + S_{k-1}(w)} \times \frac{S_{k-2}(w)}{a + S_{k-2}(w)} \mathbf{1}_{\{\mathcal{Y}_w(\alpha) \geqslant k-2\}} \middle| S_{k-2}(w), \mathbf{1}_{\{\mathcal{Y}_w(\alpha) \geqslant k-2\}}\right]\right] \\ &= \mathbb{E}\left[\frac{S_{k-1}(w)}{a + S_{k-1}(w)} \times \frac{S_{k-2}(w)}{a + S_{k-2}(w)} \mathbf{1}_{\{\mathcal{Y}_w(\alpha) \geqslant k-2\}}\right]. \end{aligned}$$

Iterating in this manner and noting that $\mathcal{Y}_w(\alpha) \geqslant 0$ almost surely, we deduce that the previous expression is $\mathbb{E}\left[\prod_{i=0}^{k-1} \frac{S_i(w)}{a+S_i(w)}\right]$. This now implies that

$$\mathbb{E}\left[\mathcal{Y}_w(\alpha)\right] = \sum_{k=1}^{\infty} \mathbb{E}\left[\prod_{i=0}^{k-1} \frac{S_i(w)}{a + S_i(w)}\right]. \tag{3.67}$$

Now, the display on the right is increasing in $S_i(w)$, and using the fact that g and h are bounded by J', we may bound this above by

$$\sum_{k=1}^{\infty} \prod_{i=1}^{k} \frac{J'i}{J'i + a} < \infty \quad \text{for all } a > J', \text{ by applying, for example, Stirling's approximation.}$$

Thus, for a suitable choice of a, $\mathbb{E}\left[\mathcal{Y}_w(\alpha)\right]$ is finite, so that, in particular, for each $t \in [0, \infty)$, since the random variable $\mathcal{Y}_w(t)$ is independent of the event $\{\alpha \geqslant t\}$ which occurs with positive probability,

$$\mathbb{E}\left[\mathcal{Y}_w(t)\right] \leqslant \frac{\mathbb{E}\left[Y_w(\alpha)\mathbf{1}_{\{\alpha \geqslant t\}}\right]}{\mathbb{P}\left(\alpha \geqslant t\right)} < \infty.$$

Now (3.64) follows from Claim 3.4.1. □

We require an additional claim:

Claim 3.4.3. *We have*

$$\mathbb{E}\left[r_w(t)\right] = h(w) + \mathbb{E}\left[g(w, W)\right]\mathbb{E}\left[\mathcal{Y}_w(t)\right] = h(w) + \tilde{g}(w)\mathbb{E}\left[\mathcal{Y}_w(t)\right]. \tag{3.68}$$

Proof. First note that, since $r_w(t)$ jumps by $g(w, W)$ whenever $\mathcal{Y}_w(t)$ jumps, we have

$$\mathbb{E}\left[r_w(t)\right] - h(w) = \mathbb{E}\left[\sum_{i=1}^{\mathcal{Y}_w(t)} g(w, W_i)\right].$$

Assume that $g(w, W_i)$ are bounded by J'. In addition, for each $n \in \mathbb{N}$,

$$\begin{aligned} \mathbb{E}\left[g(w, W_n)\mathbf{1}_{\{\mathcal{Y}_w(t) \geqslant n\}}\right] &= \mathbb{E}\left[g(w, W_n)\right] - \mathbb{E}\left[g(w, W_n)\mathbf{1}_{\{\mathcal{Y}_w(t) < n\}}\right] \\ &= \mathbb{E}\left[g(w, W_n)\right]\left(1 - \mathbb{P}\left(\mathcal{Y}_w(t) < n\right)\right) = \mathbb{E}\left[g(w, W_n)\right]\mathbb{P}\left(\mathcal{Y}_w(t) \geqslant n\right), \end{aligned}$$

where the second to last equality follows from the fact that the event $\{\mathcal{Y}_w(t) < n\}$ depends only on $(S_i(w))_{i=0,\ldots,n-1}$, and is thus independent of W_n. Finally, by Claim 3.4.2, $\mathbb{E}\left[Y_w(t)\right] < \infty$, and thus the result follows by applying Wald's Lemma. □

Proof of Lemma 3.1.4. First note that by (3.64) and (3.68), we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\left[\mathcal{Y}_w(t)\right] = \tilde{g}(w)\mathbb{E}\left[\mathcal{Y}_w(t)\right] + h(w),$$

and solving this differential equation, with initial condition $\mathbb{E}\left[\mathcal{Y}_w(0)\right] = 0$, we have

$$\mathbb{E}\left[\mathcal{Y}_w(t)\right] = \frac{h(w)}{\tilde{g}(w)}(e^{\tilde{g}(w)t} - 1). \tag{3.69}$$

Now, let Λ be an exponentially distributed random variable with parameter λ. Then, on the one hand, by (3.67)

$$\mathbb{E}\left[\mathcal{Y}_w(\Lambda)\right] = \sum_{k=1}^{\infty}\mathbb{E}\left[\prod_{i=0}^{k-1}\frac{S_i(w)}{S_i(w)+\lambda}\right].$$

On the other hand,

$$\mathbb{E}\left[\mathcal{Y}_w(\Lambda)\right] = \int_0^\infty \lambda e^{-\lambda u}\mathbb{E}\left[\mathcal{Y}_w(\Lambda)|\Lambda = u\right]\mathrm{d}u = \int_0^\infty \lambda e^{-\lambda u}\mathbb{E}\left[\mathcal{Y}_w(u)\right]\mathrm{d}u \overset{(3.69)}{=} \frac{h(w)}{\lambda - \tilde{g}(w)}$$

where, in order to evaluate the integral to get the last equality, we have used the fact that $\lambda > \tilde{g}_+$. The result follows. □

Chapter Four

Dynamical Models for Random Simplicial Complexes

4.1 Introduction

So far in this thesis we have studied evolving trees of a recursive nature, where one vertex arrives at each time-step. In this chapter we study the higher dimensional recursive models of *simplicial complexes*, described in Section 1.3.4 of Chapter 1. While the PANI-tree model studied in Chapter 3 also incorporated some degree of "neighbourhood influence", the models we study in this chapter have a lot more dependencies, and thus will require the use of more technical tools. As a result, for brevity we only study the quantity $N_k(n)$, the number of vertices with degree $k + d$ rather than empirical measure associated with the number of vertices with degree $k + d$ and a certain weight, although we remark similar analysis may be performed for the latter quantity. We first present a more formal description of the dynamics of the models.

4.1.1 Description of the Models

Recall from Section 1.3.4 of Chapter 1 that in the models of simplicial complexes we study, vertices are equipped with weights sampled independently from μ, supported on a subset of an interval $[0, w^*]$. Given a parameter $d \geqslant 1$, the models we study are of fixed dimension $(d-1) \geqslant 0$. In addition, the models also have a fitness function associated to them, which is a positive, symmetric function $f : [0, w^*]^d \to \mathbb{R}_+$. Using the weights of the vertices, we define the *fitness* of a face σ as the value of f when applied to the vector $\omega(\sigma)$ of the weights of the vertices that belong to that face. Abusing notation slightly, we sometimes write $f(\sigma)$ instead of $f(\omega(\sigma))$. Since f is assumed to be symmetric, the order of the coordinates of $\omega(\sigma)$ is not relevant.

Motivated by this symmetry, for all $s \geqslant 0$, we view the *type* $\omega(\sigma)$ of an s-dimensional face σ as an element of $\mathcal{C}_s := [0, w^*]^{s+1}/\sim$, where $\sim$ denotes the equivalence relation where vectors are the same under permutation of their entries. Unless otherwise stated, we identify entries of $\mathcal{C}_s$ with the set $\{(x_0, \ldots, x_s) \in [0, w^*]^{s+1} : x_0 \leqslant \ldots \leqslant x_s\}$ and equip $\mathcal{C}_s$ with the max-norm inherited from $[0, w^*]^{s+1}$.

We consider two versions of the model: Model **A** and Model **B**. These models are defined as follows: first, let $\mathcal{K}_0$ be an arbitrary $(d-1)$-dimensional simplicial complex, with finite vertex set $V_0 \subseteq -\mathbb{N}_0$ and each vertex assigned a fixed weight chosen from $\mathrm{Supp}\,(\mu)$. In this thesis, we will show that our limiting results do not depend on this choice of weights. Then, recursively for all $n \geqslant 0$:

(i) Define the random empirical measure

$$\Pi_n = \sum_{\sigma \in \mathcal{K}_n^{(d-1)}} \delta_{\omega(\sigma)} \tag{4.1}$$

on $\mathcal{C}_{d-1}$ and the associated probability measure on the set $\mathcal{K}_n^{(d-1)}$ of $(d-1)$-dimensional faces:

$$\hat{\Pi}_n = \frac{1}{\mathcal{Z}_n} \sum_{\sigma \in \mathcal{K}_n^{(d-1)}} f(\sigma)\delta_\sigma, \quad \text{where } \mathcal{Z}_n := \int_{\mathcal{C}_{d-1}} f(x)\mathrm{d}\Pi_n(x). \tag{4.2}$$

We call $\mathcal{Z}_n$ the *partition function* associated with the process $(\mathcal{K}_n)_{n\geqslant 0}$ at time n.

(ii) Select a face $\sigma' = (\sigma'_0, \dots, \sigma'_{d-1}) \in \mathcal{K}_n^{(d-1)}$ according to the measure $\hat{\Pi}_n$.

(iii) In both Models **A** and **B**, for each $\sigma'' \in \mathcal{K}_n^{(d-2)}$ such that $\sigma'' \subset \sigma'$, add the face $\sigma'' \cup \{n+1\}$ to $\mathcal{K}_n$ (here it may be useful to recall that $\mathcal{K}_n^{(-1)} = \varnothing$). Moreover, in Model **B** remove the set σ' from $\mathcal{K}_n$. Then, take the downwards closure, recalling Definition 1.2.2, to form $\mathcal{K}_{n+1}$.

Note that, in Model **A** the existing faces always remain in the complex, whilst in Model **B** the selected face is removed at every step. We call step (iii) applied to a chosen face σ' a *subdivision* of σ' by vertex $n+1$. Equivalently we say σ' has been *subdivided* by vertex $n+1$. Recall Figure 1.9 from Section 1.3.4 of Chapter 1 which illustrated a possible sample evolution of either of the models with parameter 3. We present a smaller illustration of this evolution in Figure 4.1 below.

Remark 4.1.1. *For general d, Model* ***A*** *may be considered as a generalisation of the* Network Geometry with Flavour model *introduced in [13], and outlined in Section 1.2.4, with flavour* $s = 0$*, and bounded energies. We recall that when* $s = 0$*, each face* σ *is selected with probability proportional to* $\mathrm{e}^{-\beta\epsilon_\sigma}$*, where* ϵ_σ *is the (random) energy of face* σ*. Model* ***B*** *may be considered as a generalisation of CQNMs with bounded energies (this model was also outlined in Section 1.2.4). However, note that for brevity, rather than 'deactivating' selected faces, we simply remove them from the complex: this does not affect any of the results we will be interested in this thesis.*

Remark 4.1.2. *The models we introduced can be further generalised. For example, instead of selecting a* $(d-1)$*-face to subdivide, one may consider a setting where a face of dimension* *s may be selected and subsequently subdivided, with the addition of an $(s+1)$-dimensional face.*

Dynamics of Model A and Model B with Parameter 3.

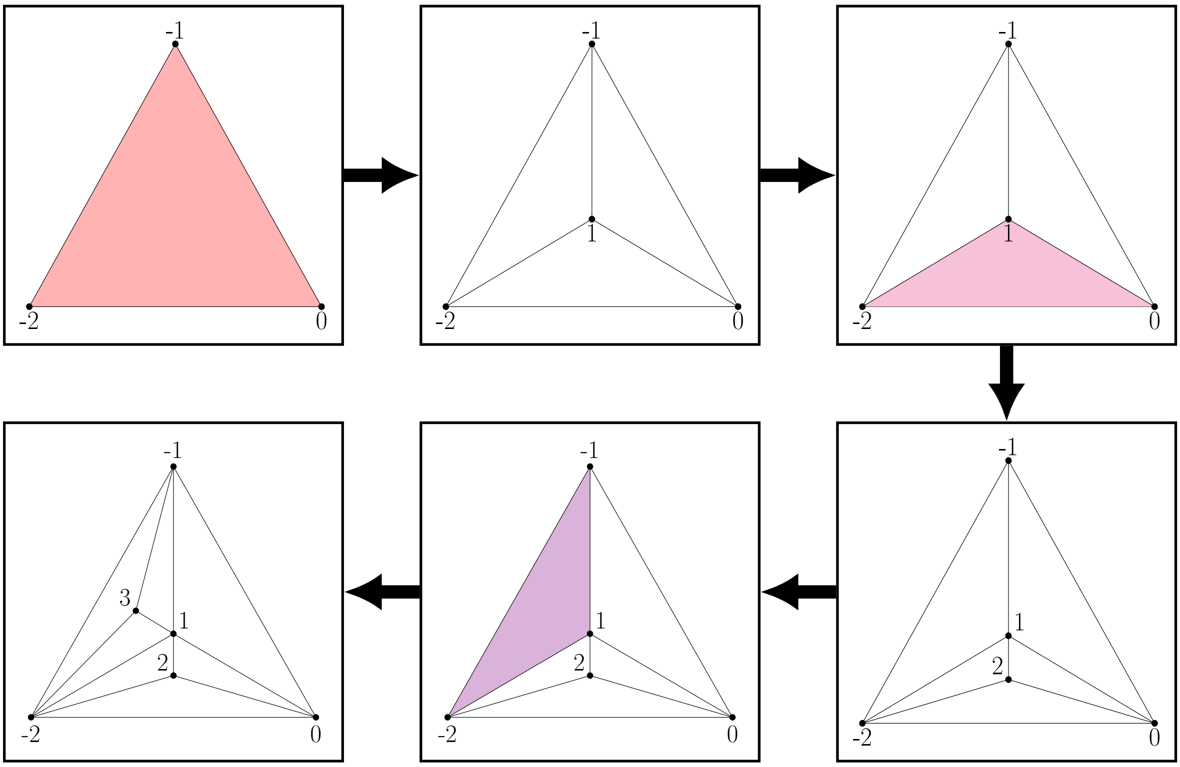


Figure 4.1: A possible evolution of steps $\mathcal{K}_0$ to $\mathcal{K}_3$ in either Model **A** or Model **B** with parameter 3. At each step, a 2-face (triangle) is chosen randomly according to step (i), and subdivided. In Model **B**, the chosen face is then removed from the complex.

Before we describe our main results we first introduce some notation specific to this chapter.

4.1.2 Some More Notation Specific to Chapter 4

Recall that for all $s \geqslant 0$, $\mathcal{C}_s = \{(x_0, \ldots, x_s) \in [0, w^*]^{s+1} : x_0 \leqslant \ldots \leqslant x_s\}$. For all $x = (x_0, \ldots, x_s) \in \mathcal{C}_s$ and $i \in \{0, \ldots, s\}$, we set $\tilde{x}_i := (x_0, \ldots, x_{i-1}, x_{i+1}, \ldots, x_s) \in \mathcal{C}_{s-1}$ and define the empirical measure $\nu_x = \sum_{i=0}^{s} \delta_{\tilde{x}_i}$ on $\mathcal{C}_{s-1}$. Next, for $w \geqslant 0$ and $y \in \mathcal{C}_s$, let $y \cup w \in \mathcal{C}_{s+1}$ denote the vector obtained by adding a coordinate equal to w to the vector y and reordering the coordinates of this $(s+1)$-dimensional vector in non-decreasing order. In addition, for $i \in \{0, \ldots, s\}$, we write $x_{i \leftarrow w} := \tilde{x}_i \cup w$. With this notation, when a face of type x is subdivided by a vertex of weight w, we add to the complex d new $(d-1)$-faces of respective types $x_{i \leftarrow w}$ for $i \in \{0, \ldots, d-1\}$. Moreover, for a vector $x = (x_0, \ldots, x_j, w, x_{j+1} \ldots, x_s) \in \mathcal{C}_s$, we denote by $x \backslash \{w\}$ the element $(x_0, \ldots, x_j, x_{j+1}, \ldots, x_s) \in \mathcal{C}_{s-1}$.

For a vertex v in a $(d-1)$-dimensional simplicial complex $\mathcal{K}$, we define the *star* of v in $\mathcal{K}$, which we denote by $\mathrm{st}_v(\mathcal{K})$, to be the subset of $\mathcal{K}^{(d-1)}$ consisting of those $(d-1)$-faces which contain v. Finally, we write $\mathbf{0}$ and $\mathbf{1}$ for the vectors $(0, \ldots, 0)$ and $(1, \ldots, 1)$ respectively, in any dimension.

4.1.3 Statements of Main Results of Chapter 4

This analysis, as we will see, applies the heuristic outlined in Section 1.4.1 of Chapter 1. Applying this approach requires two main steps, both of which are non-trivial: deriving a strong law of large numbers for the *partition function* associated with the model, and the empirical measure $(\Pi_n)_{n \geqslant 0}$, from (4.1), describing the *type* $\omega(\sigma)$ of a face σ to be chosen in the nth step; and an approach analogous to Section 2.4 of Chapter 2 to deduce convergence in probability of the degree distribution.

Part I: Convergence of the Partition Function

We will refer to the following hypotheses throughout the text:

H1. The measure μ is finitely supported, the fitness function f is positive and $|\mathcal{K}_n^{(d-1)}| \to \infty$ as $n \to \infty$, where we recall that $\mathcal{K}_n^{(d-1)}$ is the set of all $(d-1)$-faces in the random simplicial complex $\mathcal{K}_n$ at time n.

H2. The process $(\mathcal{K}_n)_{n \geqslant 0}$ evolves according to Model **A** and $\mu(\{1\}) = 0$. Moreover, the fitness function f is continuous, monotonically increasing in each argument, positive and such that, for a random variable W with distribution μ,

$$\mathbb{E}[f(\mathbf{1}_{0 \leftarrow W})] < \left(1 + \frac{1}{d}\right) \mathbb{E}[f(\mathbf{0}_{0 \leftarrow W})]. \tag{4.3}$$

Remark 4.1.3. *It is reasonable to believe that Assumption* **H2**, *and in particular* (4.3) *which ensures that the function f is not "too steep" on its domain of definition, is not necessary for our results to hold true. Our main result on the asymptotic degree distribution holds under Assumptions (a-d) of Remark 4.1.7 below. We use Assumption* **H2** *to show that Assumptions (c-d) hold: this is done in Proposition 4.1.1 and Proposition 4.1.2. Their proofs, in the case of μ having infinite support, rely on recent results of [59] on the convergence of infinitely many type Pólya urns; more precisely, Assumption* **H2** *ensures that the assumptions of [59, Theorem 1] hold. The case when μ has continuous support is more difficult to treat because the coupling arguments analogous to those applied in Section 3.2 of Chapter 3 allowing one to apply the theory of finite type Pólya urns, do not seem to work in this case.*

Note that $|\mathcal{K}_n^{(d-1)}| \to \infty$ as long as $d > 1$ in Model **B**, and for all $d \geqslant 1$ in Model **A**.

Proposition 4.1.1. *Assume* **H1** *or* **H2**, *and let $Y_n, n \geqslant 1$, be the $\mathcal{C}_{d-1}$-valued random variable that equals the type of the face chosen to be subdivided in the n-th step. Then, Y_n converges to a $\mathcal{C}_{d-1}$-valued random variable Y_∞ in distribution when n tends to infinity.*

Given any sub-complex $\tilde{\mathcal{K}} \subseteq \mathcal{K}_n$ define

$$F(\tilde{\mathcal{K}}) := \sum_{\sigma \in \tilde{\mathcal{K}}^{(d-1)}} f(\sigma). \tag{4.4}$$

and note that $F(\mathcal{K}_n) = \mathcal{Z}_n$, the partition function associated with the process defined in (4.2).

Proposition 4.1.2. *Assume* **H1** *or* **H2**. *Then, there exists* $\lambda > 0$ *such that, almost surely,*

$$\frac{\mathcal{Z}_n}{n} = \frac{F(\mathcal{K}_n)}{n} \longrightarrow \lambda, \quad \textit{as } n \to \infty.$$

Remark 4.1.4. *The distribution of the limiting random variable* Y_∞ *and the value of* λ *do not depend on the choice of the initial complex* $\mathcal{K}_0$.

Remark 4.1.5. *Because under either condition* **H1** *or* **H2** *the function* f *is bounded, we have trivial deterministic bounds on* $\mathcal{Z}_n = F(\mathcal{K}_n)$, *and therefore on* λ. *In particular, if we let*

$$f_{\min} = \min\{f(x) : x \in \mathcal{C}_{d-1}\} \quad \textit{and} \quad f_{\max} = \max\{f(x) : x \in \mathcal{C}_{d-1}\} \tag{4.5}$$

be the minimum and the maximum respectively of the fitness function on its domain of definition, then $\lambda \in [df_{\min}, df_{\max}]$ *in Model* ***A***, *whereas* $\lambda \in [(d-1)f_{\min}, (d-1)f_{\max}]$ *in Model* ***B***.

Remark 4.1.6. *The monotonicity requirement and* (4.3) *in* **H2** *may be used to cover a particular case of the Network Geometry with Flavour, the model from [13] outlined in Section 1.2.4 in Chapter 1. Namely, we may cover the case with 'flavour'* $s = 0$, *in which each face* σ *is selected with probability proportional to* $\mathrm{e}^{-\beta\epsilon_\sigma}$, *where* ϵ_σ *is the energy of face* σ, *and the selected faces remain in the complex. We may do this by setting the weights* $w_i := (1 - \epsilon_i)$ *where* ϵ_i *are the energies assigned to the vertices. We therefore assume that the distribution of* ϵ_i *does not have an atom at* 0, *the energies are bounded, and* (4.3) *is satisfied, that is, the "inverse temperature"* β *satisfies* $\beta < \frac{1}{d-1}\log\left(1 + \frac{1}{d}\right)$.

Both Proposition 4.1.1 and Proposition 4.1.2 are corollaries of a more general almost sure limit theorem for the empirical measure $\Pi_n, n \geqslant 0$ associated with the types of faces in the complex, namely Theorem 4.3.1 proved in Section 4.3. While this result, and therefore the two propositions, follows from the standard Pólya urn theory outlined in Section 3.2.1 of Chapter 3 under **H1**, for **H2** we need to make use of general results for measure-valued Pólya urn processes recently established in [59] to cover the general case. See, in particular, Section 4.3 in this work.

4.1.4 The companion star process

In this model the companion process that tracks the probability of selecting a vertex as its degree evolves (as outlined in Section 1.4.1 of Chapter 1) takes the form of a simplicial complex valued stochastic process $(S_n^*)_{n\geqslant 0}$. Informally, this process approximates the evolution of the star of a fixed vertex i in $(\mathcal{K}_n)_{n\geqslant 0}$, assuming that i is sufficiently large, namely, large enough for the distribution of Y_i, the type of the face selected by node i when it enters the network, to be close enough to the distribution of Y_∞ from Proposition 4.1.1). Let π_∞ denote the distribution of the random variable Y_∞. Then, sample a face type from π_∞, and form a $(d-1)$-simplex on vertex set $\{1-d,\ldots,0\}$ with weights corresponding to this type. Subdivide this face (using the mechanisms of Model **A** or **B**) by a new vertex labelled r with weight W sampled from μ, and form the simplicial complex S_0^* consisting of the $(d-1)$-faces containing r. We call r the *centre* of S_0^*. Then, recursively:

(i) Select a face σ from $(S_n^*)^{(d-1)}$ with probability proportional to its fitness, and subdivide it by a new vertex $n+1$ obeying the subdivision rules of Model **A** or Model **B** respectively.

(ii) Form the simplicial complex S_{n+1}^* consisting only of the $(d-1)$-faces containing r. Essentially this means removing all the $(d-1)$-faces formed during the subdivision step not containing r.

Dynamics of the Companion Process with Parameter 3.

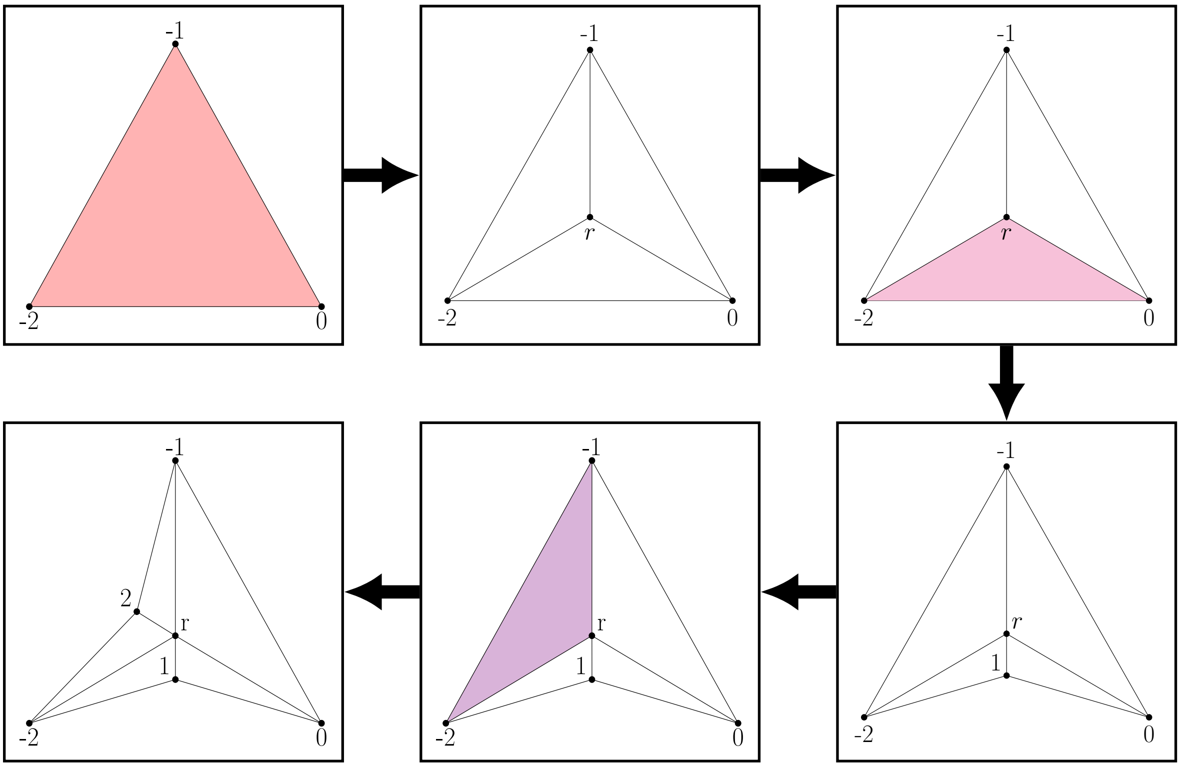


Figure 4.2: The evolution of the companion process, S_0^* to S_2^* in Model **B** with parameter 3. A face with type selected from π_∞ is formed on vertices $\{-2,-1,0\}$ and subdivided with a vertex labelled r to form S_0^* in the second square. Subsequently, a face is chosen randomly and subdivided according to step (i), and then faces not containing r are deleted. Since this is Model **B**, the chosen face is also removed from the complex.

A more formal construction of this process is provided in Section 4.3.3. We set

$$F(S_n^*) := \sum_{\sigma \in (S_n^*)^{(d-1)}} f(\sigma). \tag{4.6}$$

4.1.5 Main results, Part II: Convergence of the Degree Distribution

Theorem 4.1.3. *Assume* **H1** *or* **H2** *and for all $n \geqslant 1$, $k \geqslant 0$, let $N_k(n)$ denote the number of nodes of degree $k+d$ in the random simplicial complex $\mathcal{K}_n$ at time n. Then, for all $k \geqslant 0$, we have, with convergence in probability,*

$$\lim_{n\to\infty} \frac{1}{n} N_k(n) = \mathbb{E}\left[\frac{\lambda}{F(S_k^*)+\lambda} \prod_{j=0}^{k-1} \frac{F(S_j^*)}{F(S_j^*)+\lambda} \right] =: p_k,$$

where the star process S^* and its fitness function F are defined respectively in Section 4.1.4 and (4.6).

In fact, we have a more general result. Recall, from Definition 1.2.5 in Section 1.2.1 of Chapter 1, that the s-degree of a face is the number of distinct s-faces that contain it. Then, suppose that $N_k^{(s)}(n)$ denotes the number of vertices of s-degree $\binom{d}{s} + \binom{d-1}{s-1}k$, for $1 \leqslant s < d$.

Corollary 4.1.4. *Assume* **H1** *or* **H2**. *For all $k \geqslant 0$, we have, independent of the initial complex $\mathcal{K}_0$, with convergence in probability,*

$$\lim_{n\to\infty} \frac{1}{n} N_k^{(s)}(n) = p_k.$$

Remark 4.1.7. *In fact, in the proof of Theorem 4.1.3, we show that the conclusion of the theorem holds if one assumes the following weaker conditions instead of* **H1** *or* **H2***:*

(a) *The measure μ is an arbitrary probability distribution on $\mathbb{R}_+$.*

(b) *The fitness function f is non-negative, symmetric, bounded and continuous.*

(c) *If for all $n \geqslant 1$, Y_n is the type of face that is subdivided at time n, then $(Y_n)_{n\geqslant 1}$ converges in distribution when $n \to +\infty$.*

(d) *There exists $\lambda > 0$ such that, almost surely when $n \to +\infty$, $F(\mathcal{K}_n)/n \to \lambda$.*

One may interpret these assumptions as the analogue of Condition ***C2*** *used to analyse the* (μ, f, ℓ) - RIF *tree in Section 2.4 of Chapter 2.*

Remark 4.1.8. *Note that the boundedness of* f *implies that*

$$\begin{cases} (d+(d-1)n)f_{\min} \leqslant F(S_n^*) \leqslant (d+(d-1)n)f_{\max}, & \text{in Model } \boldsymbol{A}; \\ (d+(d-2)n)f_{\min} \leqslant F(S_n^*) \leqslant (d+(d-2)n)f_{\max}, & \text{in Model } \boldsymbol{B}, \end{cases} \tag{4.7}$$

where we recall that $f_{\min}$ *and* $f_{\max}$ *are the minimum and the maximum of the fitness function* f *(see* (4.5)*).*

Remark 4.1.9. *For an* r*-face* σ *with* $r < d-1$*, the* degree *of* σ *is the number of* $(d-1)$*-faces which contain* σ*. One can derive the analogue of Theorem 4.1.3 for the degree distribution of* r*-faces by considering a star companion process for an* r*-face. Here, the star of an* r*-face will simply consist of the* $(d-1)$*-faces that contain it. As long as the process is such that a.s. the total weight of the star tends to infinity, then one could derive a formula as in Theorem 4.1.3.*

Outline of the rest of Chapter 4

In Section 4.2 we discuss the connection of our main results to existing models. This will include classifying the values of d that ensure that the degree distributions follows a power law, which are consistent with analysis from [12] and [13].

Section 4.3 is dedicated to the study of the empirical measure Π_n, $n \geqslant 0$, and in particular, to the proofs of Proposition 4.1.1 and Proposition 4.1.2. As we remarked earlier (see Remark 4.1.3), these propositions make use of the recent theory of measure-valued Pólya processes. To our knowledge this is the first application of this theory, rather than finite type Pólya urns, in the context of evolving networks.

In Section 4.4 we apply the results of Section 4.3 to prove Theorem 4.1.3. This approach is similar to the approach used in Section 2.4 used in Chapter 2. However, due to the increased complexity in this model, there are additional technicalities used to find an upper bound for the limit of the mean of $N_k(n)/n$ in Section 4.4.2. Moreover, rather than applying the shorter, indirect approach used to deduce convergence of the mean applied in Section 2.4.4 of Chapter 2, we apply a more direct approach, finding a lower bound for the limit of the mean of $N_k(n)/n$ in Section 4.4.4. While details of the proof in Section 4.4.4 are much more technical, this approach is favourable as the methods used to derive a lower bound may be useful in other contexts, for example, in studying the evolution of the degree of a fixed vertex in related recursive network models.

We defer the proofs of some technical probabilistic lemmas to the end of the chapter, so as to not interrupt the general flow of the chapter.

4.2 Discussion and Examples

4.2.1 Constant fitness function

In the case that the fitness functions are constant, so that $f(x) = f_0$, we have deterministic formulas for $F(S_n^*)$ and λ. These cases correspond to models where the face chosen to be subdivided at time $n+1$ is chosen uniformly at random from the set $\mathcal{K}_n^{(d-1)}$. Here we use the asymptotic approximation of the ratio of two gamma functions: for fixed $a \in \mathbb{R}$ as $t \to \infty$

$$\frac{\Gamma(t+a)}{\Gamma(t)} = (1 + O(1/t))t^a. \tag{4.8}$$

This is a straightforward result of Stirling's approximation, i.e., (4.8) from Chapter 2, and will be used often throughout this paper.

1. In Model **A** we have $F(S_n^*) = ((d-1)n+d)f_0$, and $\lambda = df_0$. Theorem 4.1.3 implies that
$$p_k = \frac{d}{(d-1)k+2d}\prod_{j=0}^{k-1}\frac{(d-1)j+d}{(d-1)j+2d}.$$
If $d > 1$, using (4.8)
$$p_k = \left(1+\frac{1}{d-1}\right)\frac{\Gamma\left(k+\frac{d}{d-1}\right)\Gamma\left(\frac{2d}{d-1}\right)}{\Gamma\left(k+1+\frac{2d}{d-1}\right)\Gamma\left(\frac{d}{d-1}\right)} \sim k^{-\frac{2d-1}{d-1}}.$$
This is a new result. For $d = 1$ we obtain $p_k = 2^{-k}$, which is an old result of Na and Rapoport for the random recursive tree [63].

2. Model **B** with constant fitness function (with $\mathcal{K}_0$ given by a d-simplex) is the same as the Random Apollonian Network. In this case, if $d \geqslant 2$, $F(S_n^*) = ((d-2)n+d)f_0$ and $\lambda = (d-1)f_0$. Applying Theorem 4.1.3 we get,
$$p_k = \frac{d-1}{(d-2)k+2d-1}\prod_{j=0}^{k-1}\frac{(d-2)j+d}{(d-2)j+2d-1}.$$
Note that if $d = 1$, $\Pi_n(\mathcal{C}_{d-1}) = |V_0|$ (where V_0 is the set of vertices of the initial complex $\mathcal{K}_0$), so Theorem 4.1.3 does not apply. However, in this case it is easy to see that $p_1 = 1$. In the case $d = 2$, we have $p_k = \frac{2^{k-1}}{3^k}$. For $d \geqslant 3$, using (4.8), we get
$$p_k = \left(1+\frac{1}{d-2}\right)\frac{\Gamma\left(k+\frac{d}{d-2}\right)\Gamma\left(\frac{2d-1}{d-2}\right)}{\Gamma\left(k+1+\frac{2d-1}{d-2}\right)\Gamma\left(\frac{d}{d-2}\right)} \sim k^{-\frac{2d-3}{d-2}}.$$
This is the same exponent proved in [52] and [39].

4.2.2 Weighted Random Recursive Trees

The case $d = 1$ in Model **A** with initial simplicial complex given by a single vertex, is the *weighted random recursive tree*, the specific case of the (μ, f, ℓ) - RIF tree analysed in Section 2.2.4 of Chapter 2.[1] In this case, the fitness of the new vertex arriving at each time-step is independent of the rest of the complex, so the strong law of large numbers implies

[1] Note that Model **B** is trivial for $d = 1$ as the tree is a single path.

that λ in Proposition 4.1.2 is given by $\mathbb{E}\left[f(W)\right]$. Moreover, the simplicial complex $(S_j^*)_{j\geqslant 0}$ is a fixed vertex, so that $F(S_j^*) = f(W)$ for all $j \geqslant 0$, where W is the weight of the vertex. Thus, Theorem 4.1.3 implies the following:

Proposition 4.2.1. *As $n \to +\infty$, we have*

$$\frac{N_k(n)}{n} \to \mathbb{E}\left[\frac{\lambda f(W)^k}{(f(W)+\lambda)^{k+1}}\right], \quad \textit{in probability.}$$

This is a weaker version of the statements related to this model from Section 2.2.4 of Chapter 2.

4.2.3 Tails of the Distribution

In this subsection, we will require the additional assumption that

$$\left|\mathcal{K}_n^{(d-2)}\right| \overset{n\to\infty}{\longrightarrow} \infty. \tag{4.9}$$

Note that this assumption is satisfied as long as $d > 1$ in Model **A** and $d > 2$ in Model **B**. It is this assumption that leads to the emergence of scale-free behaviour for $d > 2$ in *Complex Quantum Network Manifolds* observed by Bianconi and Rahmede in [12] (recall Figure 1.6 from Chapter 1) and the scale-free behaviour for all $d > 1$ in the *Network Geometry with Flavour* from [13]. In the case μ is not finitely supported, we will require an analogue of (4.3). For brevity, we define the following additional hypotheses:

H1*. Assume **H1** and (4.9) holds.

H2*. Assume **H2** and (4.9) holds. Moreover, for all $w \in \mathrm{Supp}\,(\mu)$, the function $\tilde{f}_x : \mathcal{C}_{d-2} \to \mathbb{R}$, $\tilde{f}_x(v) = f(v \cup x)$ satisfies

$$\mathbb{E}[\tilde{f}_x(\mathbf{1}_{0\leftarrow W})] < \left(1 + \frac{1}{(d-1)}\right)\mathbb{E}[\tilde{f}_x(\mathbf{0}_{0\leftarrow W})].$$

Remark 4.2.1. *Similarly to* **H2***, we do not believe that Assumption* **H2*** *is necessary for our results to hold. We use it to apply [59, Theorem 1] in the proof of Proposition 4.2.2.*

In order to analyse the tails of the distribution from Theorem 4.1.3, we require the following proposition, similar to Proposition 4.1.2. In the statement of the following proposition, we allow S_0^* to have a centre with a fixed weight w instead of a random weight W with distribution μ. In the construction of S_0^*, however, we still choose the face according to π_∞. We use $\mathbb{P}_w$ and $\mathbb{E}_w$ for probabilities and expectations, respectively with regards to this initial state.

Proposition 4.2.2. *Assume* **H1*** *or* **H2****. Then, if the centre of* S_0^* *has weight* $w \in \operatorname{Supp}(\mu)$*, there exists* λ_w^* *such that,* $\mathbb{P}_w$*-almost surely*

$$\frac{F(S_n^*)}{n} \to \lambda_w^*.$$

We postpone the proof of Proposition 4.2.2 to Section 4.3.3. The following proposition holds under **H1***: Under Assumption **H1***, μ has finite support and thus $\max\{\lambda_w^* : w \in \operatorname{Supp}(\mu)\}$ exists and is attained at some value $w_+ \in \operatorname{Supp}(\mu)$; we set $\lambda_{w_+}^* = \max\{\lambda_w^* : w \in \operatorname{Supp}(\mu)\}$.

Proposition 4.2.3. *Assume* **H1****. With* p_k *as defined in Theorem 4.1.3, we have*

$$\liminf_{k\to\infty} \log_k p_k \geqslant -\left(1 + \frac{\lambda}{\lambda_{w_+}^*}\right). \tag{4.10}$$

Proof. Suppose $\mathbb{P}(W = w_+) = \kappa$ (recall that under **H1*** μ is finitely supported). Then, by the definition of p_k, we have

$$p_k = \mathbb{E}\left[\frac{\lambda}{F(S_k^*) + \lambda}\prod_{j=0}^{k-1}\frac{F(S_j^*)}{F(S_j^*) + \lambda}\right] \geqslant \mathbb{E}_{w_+}\left[\frac{\lambda}{F(S_k^*) + \lambda}\prod_{j=0}^{k-1}\frac{F(S_j^*)}{F(S_j^*) + \lambda}\right]\kappa.$$

Fix $\delta, \varepsilon' > 0$. By Proposition 4.2.2 (and Egorov's theorem), there exists $k_0 = k_0(\varepsilon, \delta)$ such that for all $k \geqslant k_0$

$$\mathbb{P}_{w_+}\left(\left|\frac{F(S_k^*)}{k} - \lambda_{w_+}^*\right| < \varepsilon\right) > 1 - \delta.$$

Let $\mathcal{G}^*_{\varepsilon,\delta}$ be the associated event in the previous display. We may bound the product $\prod_{j=0}^{k_0-1} \frac{F(S_j^*)}{F(S_j^*)+\lambda}$ below by a constant by applying (4.7). Moreover, for all $k > k_0$, on $\mathcal{G}^*_{\varepsilon,\delta}$, we have

$$\begin{aligned}\frac{\lambda}{F(S_k^*)+\lambda}\prod_{\ell=k_0}^{k-1}\frac{F(S_\ell^*)}{F(S_\ell^*)+\lambda} &> \frac{\lambda\left(k(\lambda^*_{w*}-\varepsilon)+\lambda\right)}{k(\lambda^*_{w*}+\varepsilon)+\lambda}\cdot\frac{1}{k(\lambda^*_{w*}-\varepsilon)+\lambda}\prod_{\ell=k_0}^{k-1}\frac{\ell(\lambda^*_{w_+}-\varepsilon)}{\ell(\lambda^*_{w_+}-\varepsilon)+\lambda}\\ &= \frac{k(\lambda^*_{w*}-\varepsilon)+\lambda}{k(\lambda^*_{w*}+\varepsilon)+\lambda}\cdot\frac{\lambda}{\lambda^*_{w*}-\varepsilon}\cdot\frac{\Gamma(k_0+\frac{\lambda}{\lambda^*_{w_+}-\varepsilon})}{\Gamma(k_0-1)}\frac{\Gamma(k)}{\Gamma(k+1+\frac{\lambda}{\lambda^*_{w_+}-\varepsilon})}.\end{aligned}$$

Therefore, by applying (4.8), we find that there exists a constant $c = c(k_0, \delta, \varepsilon, \kappa)$ such that

$$\log_k p_k \geqslant \log_k c - \left(1 + \frac{\lambda}{\lambda^*_{w_+} - \varepsilon}\right).$$

The equation (4.10) follows from taking limits as $k \to \infty$, and sending ε to 0. □

Further Discussion

Applying (4.7), it is easy to show that, whenever (4.9) holds,

$$\liminf_{k\to\infty} \log_k p_k \geqslant \begin{cases} -\left(1+\frac{\lambda}{(d-1)f_{\min}}\right), & \text{in Model } \mathbf{A};\\ -\left(1+\frac{\lambda}{(d-2)f_{\min}}\right), & \text{in Model } \mathbf{B},\end{cases}$$

and likewise,

$$\limsup_{k\to\infty} \log_k p_k \leqslant \begin{cases} -\left(1+\frac{\lambda}{(d-1)f_{\max}}\right), & \text{in Model } \mathbf{A};\\ -\left(1+\frac{\lambda}{(d-2)f_{\max}}\right), & \text{in Model } \mathbf{B}.\end{cases}$$

Thus, when $d > 1$ in Model **A** and $d > 2$ in Model **B**, the degree distribution is bounded above and below by a power law. This leads to the scale-free behaviour observed in [12] and [13].

In general, by counting the edges in the complex in two different ways, we find that $\sum_{k=0}^{\infty} kp_k \leqslant d$, so that p_k cannot obey a power law with a fixed exponent less than 2, otherwise the sum would diverge. However, we cannot deduce from these methods that the degree distribution in each case follows a power law with a fixed exponent. Instead, we believe that the degree distribution obeys an 'averaged' power law, as described in the GPAF-tree and the PANI-tree in Section 2.3.1 of Chapter 2 and Section 3.1.2 of Chapter 3 respectively.

4.3 Convergence of the empirical distribution

The aim of this section is to prove the following almost sure limit theorem for the empirical distribution Π_n.

Theorem 4.3.1. *Assume* **H1** *or* **H2**. *Then, there exists a deterministic, positive, finite measure π on $\mathcal{C}_{d-1}$, which does not depend on the choice of $\mathcal{K}_0$ such that, almost surely,*

$$\frac{\Pi_n}{n} \to \pi$$

with respect to the weak topology.

Proposition 4.1.2 and Proposition 4.1.1 both follow from Theorem 4.3.1 above, with $\lambda = \int_{\mathcal{C}_{d-1}} f(x)\,\mathrm{d}\pi(x)$ in Proposition 4.1.2 and Y_∞ from Proposition 4.1.1 having law π_∞ defined by

$$\pi_\infty(A) = \frac{\int_A f(x)\mathrm{d}\pi(x)}{\int_{\mathcal{C}_{d-1}} f(x)\mathrm{d}\pi(x)},$$

for any measurable set $A \subseteq \mathcal{C}_{d-1}$.

4.3.1 Proof of Theorem 4.3.1 Assuming Hypothesis H1

To prove Theorem 4.3.1 assuming **H1**, we view the collection of faces as balls in a *generalised Pólya urn process*, the family of stochastic processes previously applied in Section 3.2 (and briefly described in Section 3.2.1) of Chapter 3.

Recall from Section 3.2.1 of Chapter 3 that in this set-up, one considers an *urn* consisting of *balls* with a finite number of possible *types*. A ball of type j is sampled at random from the urn with probability proportional to its *activity* a_j, and replaced with balls of a number of different types according to a possibly random *replacement rule*. In the common set-up, the configuration of the urn after n replacements is represented as a *composition vector* X_n with entries labelled by type, and the activities associated with the types are encoded in an *activity vector* $\mathbf{a}$. In this vector, the ith entry corresponds to the number of balls of type i. Let (ξ_{ij}) be the matrix whose ijth component denotes the random number of balls of type j added, if a ball of type i is drawn. The following theorem is implied by Theorem 3.2.1 and Lemma 3.2.2 from Chapter 3, which we recall were due to Janson [45].

Theorem 4.3.2 ([45]). *Assume $\xi_{ii} \geqslant -1$, $\xi_{ij} \geqslant 0$ for $i \neq j$, and the matrix $A_{ij} := a_j \mathbb{E}\left[\xi_{ji}\right]$ is* irreducible. *Moreover, denote by λ_1 the principal eigenvalue of A, and v_1 the corresponding right-eigenvector normalised so that $\mathbf{a}^T v_1 = 1$. For any non-empty initial configuration of the urn, we have*

$$\frac{X_n}{n} \xrightarrow{n \to \infty} \lambda_1 v_1,$$

almost surely, and independently of the initial configuration of the urn.

Note that when μ is finitely supported, the number of possible face types $\omega(\sigma)$ in the complex is finite. We denote this finite set of possible types by $\mathcal{C}^f_{d-1} \subseteq \mathcal{C}_{d-1}$. The empirical distribution of face types then corresponds to the distribution of balls in a generalised Pólya urn; where the types of the balls in the urn correspond to the types of the $(d-1)$-faces, and the activities are the fitnesses. In each step, we draw a ball of type x in the urn with probability proportional to its activity $f(x)$, choose a weight W independently according to μ, and add d new balls of respective types $x_{i \leftarrow W}$, for $i \in \{0, \ldots, d-1\}$. In Model **B** we also remove the ball we drew from the urn.

Proof of Theorem 4.3.1, assuming **H1**. Recall that, under **H1**, the random weight W has finite support, and thus, for some $M > 0$, $W \in \{w_1, \ldots, w_M\}$ almost surely. Let $X_n = (X_x(n))_{x \in \mathcal{C}^f_{d-1}}$ denote the vector whose coordinate $X_x(n)$ counts the number of balls of type x in the urn after n steps. For $x \in \mathcal{C}^f_{d-1}$ and $k \in \{1, \ldots, M\}$, let $n_x(k)$ be the number of entries in x equal to w_k. We call $x \neq x'$ neighbours if x' can be obtained from x by changing exactly one entry $\ell_1 = \ell_1(x, x')$ into w_{ℓ_2}, where $\ell_2 = \ell_2(x, x')$ (and then re-ordering the entries in non-decreasing order).

In Model **A**, this urn has the following replacement rule:

$$\xi_{xx'} = \begin{cases} \sum_{k=1}^M n_x(k)\mathbf{1}_{\{w_k\}}(W) & x = x', \\ n_x(\ell_1)\mathbf{1}_{\left\{w_{\ell_2(x,x')}\right\}}(W) & \text{if } x, x' \text{ are neighbours}, \\ 0 & \text{otherwise;} \end{cases}$$

whilst in Model **B** the replacement rule is

$$\xi_{xx'} = \begin{cases} \sum_{k=1}^M n_x(k)\mathbf{1}_{\{w_k\}}(W) - 1 & x = x', \\ n_x(\ell_1)\mathbf{1}_{\left\{w_{\ell_2(x,x')}\right\}}(W) & \text{if } x, x' \text{ are neighbours}, \\ 0 & \text{otherwise.} \end{cases}$$

If we define the matrix $A_{xx'} = f(x')\mathbb{E}\left[\xi_{x'x}\right]$, since $f > 0$ it is easy to see that A is irreducible. Thus we may deduce Theorem 4.3.1 by applying Theorem 4.3.2. □

4.3.2 Proof of Theorem 4.3.1 Assuming Hypothesis H2

In order to prove Theorem 4.3.1 assuming **H2**, we show that $\Pi_n, n \geqslant 0$ is a *measure-valued Pólya process* (MVPP), a recent extension of the finite type generalised Pólya urn theory introduced in [7] and [58]. We then apply results from [59]. In the process, we will state a few lemmas, whose proofs we defer to the end of the section in Section 4.3.4. For brevity, for the rest of the section, we set

$$w^* = 1,$$

so that the maximum possible value a weight can take is 1. This is done purely for convenience of notation, and the results easily extend to other values of $w^* \in \mathbb{R}_+$.

Let $\mathcal{S}$ be a locally compact Polish space and $\mathcal{M}(\mathcal{S})$ be the set of finite, non-negative measures on $\mathcal{S}$. Recall that $\mathcal{M}(\mathcal{S})$ is also Polish when equipped with the Prokhorov metric, which metrises the weak topology when we view $\mathcal{M}(\mathcal{S})$ as the dual of the space of bounded continuous functions from $\mathcal{S}$ to $\mathbb{R}$. For a given kernel P on $\mathcal{S}$ and $\mu \in \mathcal{M}(\mathcal{S})$, we define the measure

$$(\mu \otimes P)(\cdot) := \int_{\mathcal{S}} P_x(\cdot)\,\mathrm{d}\mu(x).$$

Thanks to, e.g., [50, Section 4.1], and because of the local compactness, a random function R with values in $\mathcal{M}(\mathcal{S})$ is a random variable, i.e., measurable, if and only if, for all Borel sets $B \subseteq \mathcal{S}$, $R(B)$ is a real-valued random variable. We call a family $R_x, x \in \mathcal{S}$ of random variables with values in $\mathcal{M}(\mathcal{S})$ a *random kernel* if, almost surely, $x \mapsto R_x$ is continuous. Note that, for a random kernel R_x, $x \in \mathcal{S}$, the annealed quantity $\bar{R}_x(\cdot) = \mathbb{E}\left[R_x(\cdot)\right]$ is a kernel on $\mathcal{S}$ and the map $x \mapsto \bar{R}_x$ is continuous. We call two random kernels R_x, R'_x for $x \in \mathcal{S}$ independent if, for all $x \in \mathcal{S}$, the random measures R_x, R'_x are independent.

Definition 4.3.3. *Let $(R_x^{(n)}, x \in \mathcal{S})_{n \geqslant 1}$ be a sequence of i.i.d. random kernels. The measure-valued Pólya process with $m_0 \in \mathcal{M}(\mathcal{S})$ satisfying $m_0(\mathcal{S}) > 0$, replacement kernels $(R_x^{(n)}, x \in \mathcal{S})_{n \geqslant 1}$ and non-negative weight kernel P is the sequence of random non-negative measures $(m_n)_{n \geqslant 0}$ defined recursively as follows: given $m_{n-1}, n \geqslant 1$:*

(i) Sample a random variable ξ from $\mathcal{S}$ according to the probability measure

$$\frac{(m_{n-1} \otimes P)(\,\cdot\,)}{(m_{n-1} \otimes P)(\mathcal{S})}.$$

(ii) Set $m_n = m_{n-1} + R_\xi^{(n)}$.

The next lemma allows us to express the empirical distribution of the $(d-1)$-faces in Model **A** as an MVPP.

Lemma 4.3.4. *For all $n \geqslant 1$ and $x \in \mathcal{C}_{d-1}$ let*

$$R_x^{(n)} = \sum_{i=0}^{d-1} \delta_{x_{i\leftarrow W_n}}.$$

The sequence $\Pi_n, n \geqslant 0$, is the MVPP with initial composition Π_0, replacement kernel $(R_x^{(n)}, x \in \mathcal{C}_{d-1})_{n\geqslant 1}$ and weight kernel $P_x = f(x)\delta_x$, $x \in \mathcal{C}_{d-1}$.

Proof. Let σ be the face chosen and subdivided at step n and ξ be its type. By construction,

$$\Pi_n = \Pi_{n-1} + \sum_{i=0}^{d-1} \delta_{\xi_{i\leftarrow W_n}} = \Pi_{n-1} + R_\xi^{(n)},$$

and, for all Borel sets $B \subseteq \mathcal{C}_{d-1}$,

$$\mathbb{P}\left(\xi \in B | \Pi_{n-1}\right) = \frac{\sum_{\sigma\in\mathcal{K}_n^{(d-1)}} f(\sigma)\delta_{\omega(\sigma)}(B)}{\sum_{\sigma\in\mathcal{K}_n^{(d-1)}} f(\sigma)} = \frac{(\Pi_{n-1}\otimes P)(B)}{(\Pi_{n-1}\otimes P)(\mathcal{C}_{d-1})}.$$

This concludes the proof. □

We now state [59, Theorem 1]. We will apply this theorem to the MVPP $\Pi_n, n \geqslant 0$ to deduce Theorem 4.3.1. We require the following definitions. For an i.i.d. sequence of random kernels $(R_x^{(n)}, x \in \mathcal{S})_{n\geqslant 1}$ and a weight kernel P, let $\bar{R}_x(\cdot) = \mathbb{E}\left[R_x^{(1)}(\cdot)\right]$ and

$$Q_x^{(n)}(\cdot) := (R_x^{(n)} \otimes P)(\cdot) = \int_{\mathcal{S}} P_y(\cdot)\,\mathrm{d}R_x^{(n)}(y) \quad \text{and } \bar{Q}_x(\cdot) := (\bar{R}_x \otimes P)(\cdot) = \int_{\mathcal{S}} P_y(\cdot)\,\mathrm{d}\bar{R}_x(y).$$

Theorem 4.3.5 (Mailler & Villemonais [59])**.** *Let $(m_n)_{n\geqslant 0}$ be the MVPP on $\mathcal{S}$ with initial composition m_0, replacement kernel $(R_x^{(n)}, x \in \mathcal{S})_{n\geqslant 1}$ and weight kernel P. Assume that:*

A1 *For all $x \in \mathcal{S}$, $\bar{Q}_x(\mathcal{S}) \leqslant 1$, and there exists a probability distribution $\eta \neq \delta_0$ on $\mathbb{R}_+$ such that, for all $x \in \mathcal{S}$, the law of $Q_x^{(1)}(\mathcal{S})$ stochastically dominates η.*

A2 *The space $\mathcal{S}$ is compact.*

A3 *Denote by $(X_t)_{t\geqslant 0}$ the continuous-time Markov process defined on $\mathcal{S} \cup \{\varnothing\}$ absorbed at $\varnothing$ with infinitesimal generator given by $\bar{Q}_x - \delta_x + (1 - \bar{Q}_x(\mathcal{S}))\delta_\varnothing$. There exists a* *probability distribution ν such that*

$$\mathbb{P}_x(X_t \in \cdot \,|\, X_t \neq \varnothing) \to \nu(\cdot),$$

with respect to the total variation distance on $\mathcal{C}_{d-1}$ uniformly over $x \in \mathcal{C}_{d-1}$.

A4 *For all bounded and continuous functions $g : \mathcal{S} \to \mathbb{R}$, the functions $x \mapsto \int_{\mathcal{S}} g(y) d\bar{R}_x(y)$ and $x \mapsto \int_{\mathcal{S}} g(y) d\bar{Q}_x(y)$ are continuous.*

Then, almost surely as $n \to \infty$, m_n/n converges to $\nu \otimes \bar{R}$ with respect to the weak topology on $\mathcal{M}(\mathcal{S})$.

*Proof of Theorem 4.3.1, assuming **H2**.* The idea of the proof is to apply Theorem 4.3.5 to the MVPP $(\Pi_n)_{n\geqslant 0}$ (see Lemma 4.3.4). In this set-up, we have, for all $x \in \mathcal{C}_{d-1}$,

$$Q_x^{(n)}(\cdot) = (R_x^{(n)} \otimes P)(\cdot) = \sum_{i=0}^{d-1} f(x_{i\leftarrow W_n})\, \delta_{x_{i\leftarrow W_n}}(\cdot),$$

and

$$\bar{Q}_x(\cdot) = (\bar{R}_x \otimes P)(\cdot) = \mathbb{E}\left[\sum_{i=0}^{d-1} f(x_{i\leftarrow W})\, \delta_{x_{i\leftarrow W}}(\cdot)\right].$$

In order to satisfy the normalization requirements in Theorem 4.3.5, we consider a suitable re-scaling. We define

$$M = d \cdot \mathbb{E}[f(\mathbf{1}_{0\leftarrow W})], \tag{4.11}$$

and for all $n \geqslant 0$, set $\Pi'_n = \Pi_n/M$. It is immediate (using Lemma 4.3.4) that $(\Pi'_n)_{n\geqslant 0}$ is a MVPP with weight kernel P whose replacement kernel and associated Q-kernel are given by

$$\mathcal{R}_x^{(n)} = \frac{R_x^{(n)}}{M}, \quad \mathcal{Q}_x^{(n)} = \frac{Q_x^{(n)}}{M}.$$

The corresponding annealed kernels are defined analogously by $\bar{\mathcal{R}}_x(\cdot) = \mathbb{E}\left[\mathcal{R}_x^{(1)}(\cdot)\right]$ and $\bar{\mathcal{Q}}_x(\cdot) = \mathbb{E}\left[\mathcal{Q}_x^{(1)}(\cdot)\right]$. Note that, by monotonicity of f in all its coordinates, and symmetry,

$$\sup_{x\in\mathcal{C}_{d-1}} \mathbb{E}\left[\sum_{i=0}^{d-1} f(x_{i\leftarrow W})\right] \leqslant d \cdot \mathbb{E}\big[f(\mathbf{1}_{0\leftarrow W})\big],$$

implying that, for all $x \in \mathcal{C}_{d-1}$, $\bar{\mathcal{Q}}_x(\mathcal{C}_{d-1}) \leqslant 1$. We also have that, for all $x \in \mathcal{C}_{d-1}$, by monotonicity of f

$$\mathcal{Q}_x^{(1)}(\mathcal{C}_{d-1}) \geqslant \frac{d \cdot f(\mathbf{0})}{M} \overset{(4.11)}{=} \frac{d \cdot f(\mathbf{0})}{d \cdot \mathbb{E}[f(\mathbf{1}_{0 \leftarrow W})]} \geqslant \frac{f(\mathbf{0})}{f(\mathbf{1})} > 0,$$

implying that Assumption **A1** of Theorem 4.3.5 is satisfied with $\eta = \delta_{f(\mathbf{0})/f(\mathbf{1})}$. Assumption **A2** is immediately satisfied since $\mathcal{C}_{d-1}$ is compact. Next, as $\int_{\mathcal{C}_{d-1}} g(y) \mathrm{d}\bar{R}_x(y) = \sum_{i=0}^{d-1} \mathbb{E}\left[g(x_{i \leftarrow W})\right]$, continuity of $x \mapsto \int_{\mathcal{C}_{d-1}} g(y) \mathrm{d}\bar{R}_x(y)$ for a bounded and continuous function $g : \mathcal{C}_{d-1} \to \mathbb{R}$ is immediate. Analogously, one can prove the statement for the Q-kernel and establish Assumption **A4** as the rescaling leaves continuity properties unaltered.

It thus remains to check that the rescaled Pólya process $(\Pi'_n)_{n \geqslant 0}$ satisfies Assumption **A3**. Let $(X_t)_{t \geqslant 0}$ be the jump-process with infinitesimal generator $\bar{\mathcal{Q}}_x - \delta_x + (1 - \bar{\mathcal{Q}}_x(\mathcal{C}_{d-1}))\delta_\varnothing$, for all $x \in \mathcal{C}_{d-1}$. By definition, when X_t sits at x, it jumps to $\varnothing$ at rate

$$1 - \frac{1}{M} \sum_{i=0}^{d-1} \mathbb{E}[f(x_{i \leftarrow W})],$$

and, at rate $\frac{1}{M} \sum_{i=0}^{d-1} \mathbb{E}[f(x_{i \leftarrow W})]$, it jumps to a random position chosen according to the probability distribution

$$\frac{\sum_{i=0}^{d-1} \mathbb{E}[f(x_{i \leftarrow W}) \delta_{x_{i \leftarrow W}}(\cdot)]}{\sum_{i=0}^{d-1} \mathbb{E}[f(x_{i \leftarrow W})]}.$$

Thus, in total, X jumps at rate 1 at all times. In particular, discrete skeleton and jump times of the process are independent.

To prove **A3**, we apply [23, Theorem 3.5 and Lemma 3.6] to the jump process $(X_t)_{t \geqslant 0}$, where we take $t_1 = t_2 = 1$[2]. Since X is a pure jump process and satisfies the strong Markov property, condition (F0) in [23, Theorem 3.5] is satisfied. It is therefore enough to prove that there exist a set $L \subseteq \mathcal{C}_{d-1}$ and a probability measure ϱ on L such that:

- **B1** There exist $c_1 > 0$ such that, for all $x \in L$, $\mathbb{P}_x(X_1 \in \cdot) \geqslant c_1 \varrho(\cdot \cap L)$, where $\mathbb{P}_x(\cdot)$ denotes the probability measure associated with the Markov process X initiated by x.

[2] Note that, although this is not clear in the current version of [23], t_1 and t_2 need to be positive.

B2 There exist $0 < \gamma_1 < \gamma_2$ such that

$$\sup_{x \in \mathcal{C}_{d-1}} \mathbb{E}_x[\gamma_1^{-\tau_L \wedge \tau_\varnothing}] < +\infty, \text{ and } \gamma_2^{-t}\mathbb{P}_x(X_t \in L) \to +\infty \text{ when } t \to +\infty \, (\forall x \in L),$$

where $\tau_\varnothing$ and τ_L stand for the respective hitting times of $\varnothing$ and L.

B3 There exists $c_2 > 0$ such that

$$\sup_{t \geqslant 0} \frac{\sup_{y \in L} \mathbb{P}_y(t < \tau_\varnothing)}{\inf_{y \in L} \mathbb{P}_y(t < \tau_\varnothing)} \leqslant c_2.$$

In order to prove the above, we define the partial order '$\preccurlyeq$' on $\mathcal{C}_{d-1}$ such that for $x, y \in \mathcal{C}_{d-1}$, $x \preccurlyeq y$ if and only if, for all $i \in \{0, \ldots, d-1\}$, $x_i \leqslant y_i$ (recall that the coordinates of x and y are ordered in increasing order). We then define $L = L(\varepsilon) = \{x \in \mathcal{C}_{d-1} \colon x \preccurlyeq (1-\varepsilon)\mathbf{1}\}$.

Proof of B1: We denote by $(\sigma_i)_{i \geqslant 1}$ the random jump-times of X. In order for these times to be well-defined for all $n \geqslant 1$, we let the process jump from $\varnothing$ to $\varnothing$ at rate one. Fix a Borel set $B \subseteq \mathcal{C}_{d-1}$. Then, by monotonicity and symmetry, we have

$$\mathbb{P}_x(X_{\sigma_1} \in B) = \frac{1}{M} \sum_{i=0}^{d-1} \mathbb{E}[f(x_{i \leftarrow W})\mathbf{1}_B(x_{i \leftarrow W})] \geqslant \frac{f(\mathbf{0})}{M} \sum_{i=0}^{d-1} \mathbb{P}(x_{i \leftarrow W} \in B).$$

By the strong Markov property, we have

$$\mathbb{P}_x\left(X_{\sigma_2} \in B \,|\, X_{\sigma_1} = x'\right) = \frac{1}{M} \sum_{i=0}^{d-1} \mathbb{E}[f(x'_{i \leftarrow W})\mathbf{1}_B(x'_{i \leftarrow W})] \geqslant \frac{f(\mathbf{0})}{M} \sum_{i=0}^{d-1} \mathbb{P}\left(x'_{i \leftarrow W} \in B\right),$$

so that,

$$\begin{aligned} \int_{\mathcal{C}_{d-1}} \mathbb{P}_x\left(X_{\sigma_2} \in B \,|\, X_{\sigma_1} = x'\right) \mathbb{P}_x(X_{\sigma_1} \in \mathrm{d}x') &\geqslant \int_{\mathcal{C}_{d-1}} \frac{f(\mathbf{0})}{M} \sum_{i=0}^{d-1} \mathbb{P}\left(x'_{i \leftarrow W'} \in B\right) \mathbb{P}_x\left(X_{\sigma_1} \in \mathrm{d}x'\right) \\ &\geqslant \left(\frac{f(\mathbf{0})}{M}\right)^2 \sum_{0 \leqslant i,j \leqslant d-1} \mathbb{P}\left((x_{j \leftarrow W})_{i \leftarrow W'} \in B\right) \end{aligned}$$

for i.i.d copies W, W'. Iterating this argument, we obtain

$$\mathbb{P}_x(X_{\sigma_d} \in B) \geqslant \left(\frac{f(\mathbf{0})}{M}\right)^d \sum_{i_0, \ldots, i_{d-1} \in \{0, \ldots, d-1\}^d} \mathbb{P}\left(\left(\left((x_{i_0 \leftarrow W_0})_{i_1 \leftarrow W_1}\right) \ldots\right)_{i_{d-1} \leftarrow W_{d-1}} \in B\right),$$

where $W_0, \ldots, W_{d-1}$ are i.i.d. random variables with law μ. Let $W_{(0)} \leqslant W_{(1)} \leqslant \ldots \leqslant W_{(n)}$ denote the order statistics of $W_0, \ldots, W_{d-1}$. Then, for an appropriate (random) choice of $i_0, \ldots, i_{d-1}$ we have $\left(\left((x_{i_0 \leftarrow W_0})_{i_1 \leftarrow W_1}\right) \ldots\right)_{i_{d-1} \leftarrow W_{d-1}} = (W_{(0)}, \ldots, W_{(d-1)})$. Therefore

$$\mathbb{P}_x(X_{\sigma_d} \in B) \geqslant \left(\frac{f(\mathbf{0})}{M}\right)^d \mathbb{E}\left[\sum_{i_0,\ldots,i_{d-1} \in \{0,\ldots,d-1\}^d} \mathbf{1}_B\left(\left(\left((x_{i_0 \leftarrow W_0})_{i_1 \leftarrow W_1}\right) \ldots\right)_{i_{d-1} \leftarrow W_{d-1}}\right)\right]$$

$$\geqslant \left(\frac{f(\mathbf{0})}{M}\right)^d \mathbb{P}\left((W_{(0)}, \ldots, W_{(d-1)}) \in B\right).$$

As the probability that X jumps exactly d times before time 1 is positive and skeleton and jump times are independent, because X always jumps with rate 1, **B1** is satisfied with ϱ being the probability distribution induced by $\mu^{\otimes d}$ restricted to L in the natural way.

Proof of B2: For $x \in \mathcal{C}_{d-1}$, let $n_x(x_i)$ denotes the number of co-ordinates of x equal to x_i. X jumps from a position x such that $x_i > 1 - \varepsilon$ to a position $x_{i \leftarrow v}$ for some $v \leqslant 1 - \varepsilon$ at rate

$$\frac{n_x(x_i)\mathbb{E}\big[f(x_{i \leftarrow W})\mathbf{1}_{W \leqslant 1-\varepsilon}\big]}{M} \geqslant \frac{n_x(x_i)\mathbb{E}[f(\mathbf{0}_{0 \leftarrow W})\mathbf{1}_{W \leqslant 1-\varepsilon}]}{M} =: n_x(x_i)\varpi_\varepsilon,$$

for all $i \in \{0, \ldots, d-1\}$ (where we have applied the symmetry and monotonicity of f). Similarly, the walk jumps from a position x such that $x_i \leqslant 1 - \varepsilon$ to a position $x_{i \leftarrow v}$ for some $v > 1 - \varepsilon$ at rate

$$\frac{n_x(x_i)\mathbb{E}\big[f(x_{i \leftarrow W})\mathbf{1}_{W > 1-\varepsilon}\big]}{M} \leqslant \frac{n_x(x_i)\mathbb{E}[f(\mathbf{1}_{0 \leftarrow W})\mathbf{1}_{W > 1-\varepsilon}]}{M} =: n_x(x_i)\vartheta_\varepsilon,$$

for all $i \in \{0, \ldots, d-1\}$. Let $\mathscr{C}(X_t)$ denote the number of coordinates of X_t that are larger than $1 - \varepsilon$, where we set $\mathscr{C}(\varnothing) = 0$. Consider a pure jump Markov process with rates given in Figure 4.3.

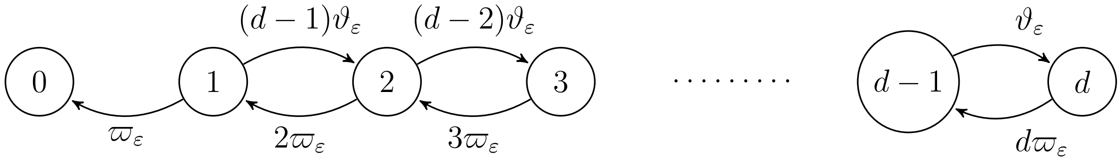


Figure 4.3: Jump rates of the associated Markov chain N^ε.

If for some $t \geqslant 0$ this Markov chain has the same non-zero value as $\mathscr{C}(X_t)$ then it jumps upwards (resp. downwards) at a faster (resp. lower) rate than $\mathscr{C}(X_t)$. This observation motivates the following lemma whose proof is given in Section 4.3.5. Note that $\tau_L \wedge \tau_\varnothing$ is the first time t when $\mathscr{C}(X_t) = 0$.

Lemma 4.3.6. *For all sufficiently small $\varepsilon > 0$, there exists a coupling of the process X with a realisation N^ε of the Markov process with jump rates given in Figure 4.3 and $N_0^\varepsilon = \mathscr{C}(X_0)$ such that, $\mathscr{C}(X_t) \leqslant N_t^\varepsilon$ for all $t \leqslant \tau_L \wedge \tau_\varnothing$.*

The proof of Lemma 4.3.6 is where we use the assumption $\mu(\{1\}) = 0$. By Lemma 4.3.6, we deduce that

$$\mathbb{P}_x\left(\tau_L \wedge \tau_\varnothing \geqslant t\right) \leqslant \mathbb{P}_{\mathscr{C}(x)}\left(N_t^\varepsilon \neq 0\right). \tag{4.12}$$

Here, we use the notation $\mathbb{P}_\ell$, $\ell \in \{0, \dots, d\}$ to indicate that the Markov process $N_t^\varepsilon, t \geqslant 0$ is initiated at position ℓ. Note that, since μ does not contain an atom at 1, we have $\vartheta_\varepsilon \to 0$ and $\varpi_\varepsilon \to \mathbb{E}[f(\mathbf{0}_{0\leftarrow W})]/M =: \varpi_0 \in (0, 1]$ as $\varepsilon \to 0$. Therefore, as $\varepsilon \to 0$ the generator $\mathcal{L}_\varepsilon$ of the Markov chain N^ε converges to the generator

$$\mathcal{L} = \begin{pmatrix} 0 & 0 & \dots & & & 0 \\ \varpi_0 & -\varpi_0 & 0 & \dots & & 0 \\ 0 & 2\varpi_0 & -2\varpi_0 & 0 & \dots & 0 \\ & & \ddots & \ddots & & \\ & & & \ddots & \ddots & \\ 0 & \dots & & 0 & d\varpi_0 & -d\varpi_0 \end{pmatrix}$$

whose eigenvalues are $0, -\varpi_0, \dots, -d\varpi_0$ (and thus whose spectral gap is ϖ_0), and whose stationary distribution on $\{0, \dots, d\}$ is given by δ_0 as 0 is an absorbing state.

Since $\mathcal{L}_\varepsilon$ converges entry-wise to $\mathcal{L}$ when $\varepsilon \to 0$, their respective characteristic polynomials converge, and thus the eigenvalues of $\mathcal{L}_\varepsilon$ converge to the eigenvalues of $\mathcal{L}$. Since the eigenvalues of $\mathcal{L}$ are all distinct it follows that for ε sufficiently small all eigenvalues of $\mathcal{L}_\varepsilon$ are simple. Thus, $\mathcal{L}_\varepsilon$ is diagonalisable, and may be written as $\mathcal{L}_\varepsilon = V_\varepsilon^{-1} D_\varepsilon V_\varepsilon$, where D_ε is a diagonal matrix consisting of the eigenvalues of $\mathcal{L}_\varepsilon$, and the rows of V_ε^{-1} are the corresponding unit-norm (left) eigenvectors. This condition allows us to apply [61, Theorem 3.1]. Since, for each $\varepsilon > 0$, the stationary distribution of N^ε is δ_0, for all $\ell \in \{0, \ldots, d\}$ and for all $t \geqslant 0$,

$$|\mathbb{P}_\ell(N_t^\varepsilon = 0) - 1| \leqslant C(\varepsilon)\mathrm{e}^{-\rho(\varepsilon)t}, \tag{4.13}$$

where $\rho(\varepsilon)$ is the spectral gap of the generator of N^ε, and $C(\varepsilon) = \|V_\varepsilon\|_\infty \|V_\varepsilon^{-1}\|_\infty$. Here $\|\cdot\|_\infty$ denotes the ∞-norm, i.e. maximum absolute row sum. Note that as $\varepsilon \to 0$, $\rho(\varepsilon) \to \varpi_0$. Moreover, using the basis of unit-norm (left) eigenvectors introduced above, we have $C(\varepsilon) = \|V_\varepsilon\|_\infty \|V_\varepsilon^{-1}\|_\infty \to C := \|V\|_\infty \|V^{-1}\|_\infty$, as $\varepsilon \to 0$, where the rows of V^{-1} are a basis of unit-norm (left) eigenvectors of $\mathcal{L}$. Now, by (4.12) and (4.13), we have

$$\mathbb{P}_x(\tau_L \wedge \tau_\varnothing \geqslant t) \leqslant \mathbb{P}_{\mathscr{C}(x)}(N_t^\varepsilon \neq 0) = 1 - \mathbb{P}_{\mathscr{C}(x)}(N_t^\varepsilon = 0) \leqslant C(\varepsilon)\exp(-\rho(\varepsilon)t). \tag{4.14}$$

Therefore, for all $\gamma_1 < 1$ and $x \in \mathcal{C}_{d-1}$, using the fact that $\log \gamma_1 < 0$ in the second equality,

$$\begin{aligned}\mathbb{E}_x[\gamma_1^{-\tau_L \wedge \tau_\varnothing}] &= 1 + \int_1^\infty \mathbb{P}_x(\gamma_1^{-\tau_L \wedge \tau_\varnothing} \geqslant u)\mathrm{d}u = 1 + \int_1^\infty \mathbb{P}_x\left(\tau_L \wedge \tau_\varnothing \geqslant \frac{\log u}{\log\left(\frac{1}{\gamma_1}\right)}\right)\mathrm{d}u \\ &\overset{(4.14)}{\leqslant} 1 + \int_1^\infty C(\varepsilon) u^{-\rho(\varepsilon)/\log\left(\frac{1}{\gamma_1}\right)}\mathrm{d}u < +\infty\end{aligned}$$

as long as $\log\left(\frac{1}{\gamma_1}\right) < \rho(\varepsilon)$. Also note that, for all $x \in L$,

$$\mathbb{P}_x(X_t \in L) \geqslant \mathbb{P}_x(X_{\sigma_i} \in L \text{ for all } 0 \leqslant i \leqslant N(t)),$$

where $N(t)$ is the number of jumps of X by time t, and

$$\begin{aligned}\mathbb{P}_x(X_{\sigma_1} \in L) &= \frac{1}{M}\sum_{i=0}^{d-1} \mathbb{E}[f(x_{i\leftarrow W})\mathbf{1}_{x_{i\leftarrow W} \in L}] = \frac{1}{M}\sum_{i=0}^{d-1} \mathbb{E}[f(x_{i\leftarrow W})\mathbf{1}_{W \leqslant 1-\varepsilon}] \\ &\overset{(4.11)}{\geqslant} \frac{\mathbb{E}[f(\mathbf{0}_{0\leftarrow W})\mathbf{1}_{W \leqslant 1-\varepsilon}]}{\mathbb{E}[f(\mathbf{1}_{0\leftarrow W})]} =: \chi_\varepsilon.\end{aligned}$$

Since the walk jumps at rate one, we have that the number of jumps before time t is Poisson distributed with parameter t. As skeleton and jump times are independent, it follows that, for all $x \in L$,

$$\mathbb{P}_x(X_t \in L) \geqslant \mathbb{P}_x(X_{\sigma_i} \in L \text{ for all } 0 \leqslant i \leqslant N(t)) \geqslant \mathbb{E}[\chi_\varepsilon^{N(t)}] = \mathrm{e}^{-(1-\chi_\varepsilon)t}.$$

If $1 - \chi_\varepsilon < \log\left(\frac{1}{\gamma_2}\right)$, then $\gamma_2^{-t}\mathbb{P}_x(X_t \in L) \to +\infty$ as required. In other words, **B2** is satisfied if we can choose $\gamma_1 < \gamma_2 < 1$ such that

$$1 - \chi_\varepsilon < \log\left(\frac{1}{\gamma_2}\right) < \log\left(\frac{1}{\gamma_1}\right) < \rho(\varepsilon).$$

As $\varepsilon \to 0$, we have $\chi_\varepsilon \to \mathbb{E}[f(\mathbf{0}_{0\leftarrow W})]/\mathbb{E}[f(\mathbf{1}_{0\leftarrow W})] = d\varpi_0$ while $\rho(\varepsilon) \to \varpi_0 > 1 - d\varpi_0$ by (4.3). It is thus possible to choose ε small enough such that $1 - \chi_\varepsilon < \rho(\varepsilon)$. For this value of ε, a choice of γ_1 and γ_2 is possible, which concludes the proof of **B2**.

Proof of B3: We require the following coupling lemma, where we adopt the convention that $\varnothing \preccurlyeq x$ for all $x \in \mathcal{C}_{d-1}$ and $\varnothing \preccurlyeq \varnothing$. We defer the proof of this lemma to Section 4.3.6

Lemma 4.3.7. *Let $x, y \in \mathcal{C}_{d-1}$ with $x \preccurlyeq y$. There exist processes $X^{(x)}, X^{(y)}$ such that $X^{(x)}$ is distributed as X with respect to $\mathbb{P}_x$ and $X^{(y)}$ is distributed as X with respect to $\mathbb{P}_y$ satisfying that, almost surely, $X_t^{(x)} \preccurlyeq X_t^{(y)}$ for all $t \geqslant 0$.*

Thanks to Lemma 4.3.7, we have that, if $x \preccurlyeq y \in \mathcal{C}_{d-1}$, then

$$\mathbb{P}_x(t < \tau_\varnothing) \leqslant \mathbb{P}_y(t < \tau_\varnothing). \tag{4.15}$$

In particular, this implies that

$$\inf_{y\in L} \mathbb{P}_y(t < \tau_\varnothing) = \mathbb{P}_{\mathbf{0}}(t < \tau_\varnothing), \text{ and } \sup_{y\in L} \mathbb{P}_y(t < \tau_\varnothing) = \mathbb{P}_{(1-\varepsilon)\mathbf{1}}(t < \tau_\varnothing).$$

Also, since $1 \in \operatorname{Supp}(\mu)$, with positive probability, every coordinate of $(X_t)_{t\geqslant 0}$ is at least $1-\varepsilon$ after d jumps. If we denote this probability by $\kappa_1 = \kappa_1(\varepsilon)$, we obtain

$$\mathbb{P}_{\mathbf{0}}(t < \tau_\varnothing) \geqslant \mathbb{P}_{\mathbf{0}}(\sigma_d < t < \tau_\varnothing) \geqslant \kappa_1 \mathbb{P}_{\mathbf{0}}\left(\sigma_d < t < \tau_\varnothing \,|\, (1-\varepsilon)\mathbf{1} \preccurlyeq X_{\sigma_d}\right),$$

where $(1-\varepsilon)\mathbf{1} \preccurlyeq X_{\tau_d}$ denotes the event that all coordinates of X_{τ_d} are at least $1-\varepsilon$. Next, observe that for all $t \leqslant 1$,

$$\frac{\mathbb{P}_{(1-\varepsilon)\mathbf{1}}\left(t < \tau_\varnothing\right)}{\mathbb{P}_{\mathbf{0}}\left(t < \tau_\varnothing\right)} \leqslant \frac{1}{\mathrm{e}^{-1}} = \mathrm{e},$$

since the probability the process has not jumped by time t is e^{-t}. Now, by (4.15) and the strong Markov property, for Lebesgue almost all $0 \leqslant u \leqslant 1 < t$,

$$\begin{aligned}\mathbb{P}_{\mathbf{0}}\left(t < \tau_\varnothing \,|\, (1-\varepsilon)\mathbf{1} \preccurlyeq X_{\sigma_d},\, \sigma_d = u\right) &= \mathbb{E}_{\mathbf{0}}\left[\mathbb{P}_{X_{\sigma_d}}\left(t-u < \tau_\varnothing\right) \,|\, (1-\varepsilon)\mathbf{1} \preccurlyeq X_{\sigma_d},\, \sigma_d = u\right] \\ &\geqslant \mathbb{P}_{(1-\varepsilon)\mathbf{1}}\left(t-u < \tau_\varnothing\right) \geqslant \mathbb{P}_{(1-\varepsilon)\mathbf{1}}\left(t < \tau_\varnothing\right).\end{aligned}$$

Thus, for $t > 1$, since jump times and skeleton are independent

$$\begin{aligned}\mathbb{P}_{\mathbf{0}}(t < \tau_\varnothing) &\geqslant \kappa_1 \mathbb{P}_{\mathbf{0}}(\sigma_d \leqslant 1 \leqslant t < \tau_\varnothing \,|\, (1-\varepsilon)\mathbf{1} \preccurlyeq X_{\sigma_d}) \\ &\geqslant \kappa_1 \int_0^1 \mathbb{P}_0\left(t < \tau_\varnothing \,|\, (1-\varepsilon)\mathbf{1} \preccurlyeq X_{\sigma_d},\, \sigma_d = u\right) \mathbb{P}_{\mathbf{0}}\left(\sigma_d \in \mathrm{d}u \,|\, (1-\varepsilon)\mathbf{1} \preccurlyeq X_{\sigma_d}\right) \\ &= \kappa_1 \int_0^1 \mathbb{P}_0\left(t < \tau_\varnothing \,|\, (1-\varepsilon)\mathbf{1} \preccurlyeq X_{\sigma_d},\, \sigma_d = u\right) \mathbb{P}_{\mathbf{0}}\left(\sigma_d \in \mathrm{d}u\right) \\ &= \kappa_1 \mathbb{P}_{\mathbf{0}}\left(\sigma_d < 1\right) \mathbb{P}_{(1-\varepsilon)\mathbf{1}}\left(t-u < \tau_\varnothing\right) \geqslant \kappa_1 \mathbb{P}_{\mathbf{0}}\left(\sigma_d < 1\right) \mathbb{P}_{(1-\varepsilon)\mathbf{1}}\left(t-u < \tau_\varnothing\right).\end{aligned}$$

Thus, if we set $\mathbb{P}_{\mathbf{0}}\left(\sigma_d < 1\right) := \kappa_2$, taking $c_2 = \max\left\{\frac{1}{\kappa_1\kappa_2}, \mathrm{e}\right\}$ completes the proof. ☐

4.3.3 The Star Process

We now revisit the companion Markov process $(S_n^*)_{n\geqslant 0}$ defined in Section 4.1.4. We wish to apply the same theory of Pólya processes to study the distribution of $(d-1)$-faces in $(S_n^*)_{n\geqslant 0}$. Note, however, that by definition, in this process every face contains the central vertex of S_0^*. Therefore, if the central vertex has weight x, we may view the empirical distribution of $(d-1)$-faces as a measure on $\mathcal{C}_{d-2}$, which represents the weights of the other vertices in the $(d-1)$-faces in S_n^*.

Thus, we can interpret the evolving empirical measure as a homogeneous Markov process $(S_n)_{n\geqslant 0}$ on $\mathcal{C}' := \mathbb{R}_+ \times \mathcal{M}(\mathcal{C}_{d-2})$, where we recall that $\mathcal{M}(\mathcal{C}_{d-2})$ is the space of non-negative, finite measures on $\mathcal{C}_{d-2}$.

Given $S_n = (x, \nu) \in \mathcal{C}'$ for some $n \geqslant 0$:

(i) Set $c^* = \int_{\mathcal{C}_{d-2}} f((x,y))\mathrm{d}\nu(y)$ and sample $z \in \mathcal{C}_{d-2}$ according to the distribution admitting density $f((x,y))/c^*$ with respect to ν.

(ii) Let W be a random variable with distribution μ which is independent of the past of the process. Then, set

$$S_{n+1} = \begin{cases} (x, \nu + \sum_{i=0}^{d-2} \delta_{z_i \leftarrow W}), & \text{in Model } \mathbf{A}, \\ (x, \nu + \sum_{i=0}^{d-2} \delta_{z_i \leftarrow W} - \delta_z), & \text{in Model } \mathbf{B}. \end{cases}$$

For a completely rigorous definition, we also set $S_{n+1} = S_n$ if the measure component of S_n is the zero measure and step (i) cannot be executed. We write $\mathbb{P}^*_{(x,\nu)}, \mathbb{E}^*_{(x,\nu)}$ for probabilities and expectations, respectively with respect to this process when the initial state S_0 satisfies $S_0 = (x, \nu)$. Note that this implies that the first component of S_n remains equal to x for all $n \geqslant 0$. Let us write $\mathbb{S}_n$ for the measure component of S_n. Then, provided that $\mathbb{S}_0$ is a non-trivial sum of Dirac measures, we have

$$\mathbb{S}_n(\mathcal{C}_{d-2}) = \begin{cases} (d-1)n + \mathbb{S}_0(\mathcal{C}_{d-2}), & \text{in Model } \mathbf{A}, \\ (d-2)n + \mathbb{S}_0(\mathcal{C}_{d-2}), & \text{in Model } \mathbf{B}. \end{cases}$$

Upon identifying faces with their types, we may consider $\mathrm{st}_i(\mathcal{K}_n)$ as a $\mathcal{C}'$-valued random variable by separating the weight of vertex i from the remaining vertices. Let $\tau_0 = i$ (which is the time of arrival of vertex i) and, for $n \geqslant 1$, let τ_n be the n-th time, the randomly chosen face in the construction of $(\mathcal{K}_m)_{m\geqslant 0}$ contains vertex i. Formally, letting σ_n denote the face chosen and subdivided in step n, we have

$$\tau_n := \inf\{m > \tau_{n-1} : i \in \sigma_m\}, \quad n \geqslant 1.$$

It is easy to see that $\tau_n < \infty$ almost surely for all $n \geqslant 1$. Indeed, under either Hypothesis **H1** or **H2**, we have $\mathcal{Z}_n = F(\mathcal{K}_n) \leqslant f_{\max}(n + |\mathcal{K}_0^{(d-1)}|)$, and if $\tau_{k-1} \leqslant n < \tau_k$, $F(\mathrm{st}_i(\mathcal{K}_n)) \geqslant f_{\min}(d-1)(k-1)$. Therefore, (analogous to proof of the Borel-Cantelli lemma) one can bound the probability

$$\mathbb{P}\left(\tau_k = \infty \,|\, \tau_{k-1} = N\right) \leqslant \prod_{j=N+1}^{\infty} \left(1 - \frac{f_{\min}(d-1)(k-1)}{f_{\max}(j + |\mathcal{K}_0^{(d-1)}|)}\right) \leqslant \mathrm{e}^{-\sum_{j=N+1}^{\infty} \frac{f_{\min}(d-1)(k-1)}{f_{\max}(j+|\mathcal{K}_0^{(d-1)}|)}} = 0;$$

and the result follows by induction on k.

Furthermore, the sequence of random variables

$$\left(W_i, \sum_{\sigma \in \mathrm{st}_i(\mathcal{K}_{\tau_n})} \delta_{\omega(\sigma)\setminus\{W_i\}}\right)_{n \geqslant 0}$$

is equal in distribution to $S_n, n \geqslant 0$ with respect to $\mathbb{P}^*_{(x,\nu)}$, when the configuration (x, ν) is chosen with respect to the law of $(W_i, \sum_{\sigma \in \mathrm{st}_i(\mathcal{K}_i)} \delta_{\omega(\sigma)\setminus\{W_i\}})$.

Let $\varphi : \mathbb{R}_+ \times \mathcal{C}_{d-1} \to \mathcal{C}' = \mathbb{R}_+ \times \mathcal{M}(\mathcal{C}_{d-2})$ be the map

$$\varphi(w, x) = \left(w, \sum_{i=0}^{d-1} \delta_{\tilde{x}_i}\right), \tag{4.16}$$

where we recall that for all $x \in \mathcal{C}_{d-1}$, $\tilde{x}_i \in \mathcal{C}_{d-2}$ is the vector x from which we have removed the i-th coordinate. We also let $\psi : \mathbb{R}_+ \times \mathcal{C}_{d-2} \to \mathcal{C}_{d-1}$ be such that

$$\psi(w, x) = w \cup x, \tag{4.17}$$

where we recall that $w \cup x$ is obtained by adding a coordinate equal to w to the vector x, and reordering the coordinates of the obtained vector in non-decreasing order. For $(w, \nu) \in \mathcal{C}'$, we define the fitness

$$F(w, \nu) = \int_{\mathcal{C}_{d-1}} f \,\mathrm{d}\psi_*(\delta_w \otimes \nu), \tag{4.18}$$

where $\psi_*(\delta_w \otimes \nu)$ is the pushforward of $\delta_w \otimes \nu$ under ψ. In other words, $\psi_*(\delta_w \otimes \nu)$ is the distribution of $\psi(w, X)$ where $X \in \mathcal{C}_{d-2}$ is a ν-distributed random variable). Note that, when S_0 is chosen according to the law of (W, Y_∞), we have $(F(S_n))_{n \geqslant 0} = (F(S_n^*))_{n \geqslant 0}$ in distribution. Moreover, for any $x \in \mathrm{Supp}((\mu))$, assuming **H1*** or **H2***, Theorem 4.3.1 implies almost sure convergence of the re-scaled measure valued process $(\frac{1}{n}\mathbb{S}_n)_{n>0}$ on $\mathcal{C}_{d-2}$ to a positive limiting measure depending on x. Thus, we get the following:

Theorem 4.3.8. *Assume* **H1*** *or* **H2*** *and recall the definition of* ψ *in* (4.17)*, and that* $\mathbb{S}_n$ *denotes the measure-valued component of the star process* $S_n \in \mathcal{C}'$*. Then, for any* $x \in \mathrm{Supp}((\mu))$*, there exists a positive measure* m^*_x *on* $\mathcal{C}_{d-1}$*, such that, for any positive non-zero measure* $\nu \in \mathcal{M}(\mathcal{C}_{d-2})$*, we have*

$$\frac{1}{n}\psi_*(\delta_x \otimes \mathbb{S}_n) \to m^*_x, \quad \mathbb{P}^*_{(x,\nu)}\text{-almost surely as } n \to \infty,$$

with respect to the weak topology.

By continuity and boundedness of f, this implies that

$$\frac{F(S_n)}{n} \to \lambda^*_x := \int_{\mathcal{C}_{d-1}} f(y)\,\mathrm{d}m^*_x(y) > 0, \quad \mathbb{P}^*_{(x,\nu)}\text{-almost surely when } n \to \infty.$$

This yields Proposition 4.2.2 by setting the initial state to be $S_0 = \varphi(w, Y_\infty)$, where Y_∞ is defined in Proposition 4.1.1 and φ in (4.16).

4.3.4 Proofs of Additional Lemmas used to prove Theorem 4.3.1

4.3.5 Proof of Lemma 4.3.6

For brevity, we omit the superscript ε when referring to the process N^ε, and in the notation of other parameters depending on ε.

Proof of Lemma 4.3.6. Let $\varepsilon > 0$ be small enough such that $\varpi > \vartheta$ (this is possible because μ does not contain an atom at 1). Then, $i\varpi + (d-i)\vartheta \leqslant 1$ for $i \in \{1, \ldots, d\}$. Let $\theta_i = 1 - i\varpi - (d-i)\vartheta, i \in \{0, \ldots, d\}$. The Markov chain N has the following dynamics: jump times are exponentially distributed with unit mean while the skeleton process performs a random walk on $\{0,\ldots,d\}$ according to the following rules: the process is absorbed at 0 and, given that its current state is $i \in \{1,\ldots,d\}$, it moves to $i+1$ with probability $(d-i)\vartheta$ and to $i-1$ with probability $i\varpi$, while it remains at i with probability θ_i.

We construct the process N from a realisation of X. First, we use the jump times $\sigma_n, n \geqslant 1$ of the X-process for the jump times of N. We define N_{σ_n} by induction, starting with $N_{\sigma_0} = \mathscr{C}(X_{\sigma_0})$, where $\sigma_0 := 0$. Let $n \geqslant 1$ and suppose $X_{\sigma_{n-1}} = \mathbf{x}$ and $\mathscr{C}(X_{\sigma_{n-1}}) = \mathbf{j}$ (recalling that $\mathscr{C}(\varnothing) = 0$). If $0 \leqslant \mathbf{j} < N_{\sigma_{n-1}}$, then choose N_{σ_n} arbitrarily obeying the dynamics of the random walk (for example by using additional external randomness). If $N_{\sigma_{n-1}} = 0$, set $N_{\sigma_n} = 0$. Finally, assume that $N_{\sigma_{n-1}} = \mathbf{j} > 0$. Let

$$s^{\uparrow} = \sum_{i=0}^{d-1-\mathbf{j}} \frac{\mathbb{E}\left[f(\mathbf{x}_{i\leftarrow W})1_{W>1-\varepsilon}\right]}{M} \leqslant (d-\mathbf{j})\vartheta, \quad s_{\downarrow} = \sum_{i=d-\mathbf{j}}^{d-1} \frac{\mathbb{E}\left[f(\mathbf{x}_{i\leftarrow W})1_{W\leqslant 1-\varepsilon}\right]}{M} \geqslant \mathbf{j}\varpi.$$

Let A be an event that has probability $\mathbf{j}\varpi/s_{\downarrow} \in [0,1]$ which is independent of the past of the process given $X_{\sigma_{n-1}}$.[3] Let

$$E = \{X_{\sigma_n} = \varnothing\} \cup (\{\mathscr{C}(X_{\sigma_n}) = \mathscr{C}(X_{\sigma_{n-1}}) - 1\} \cap A^c) \cup \{\mathscr{C}(X_{\sigma_n}) = \mathscr{C}(X_{\sigma_{n-1}})\}.$$

We first define $N(\sigma_n)$ on E^c as follows: we set

$$N_{\sigma_n} = \begin{cases} N_{\sigma_{n-1}} + 1 & \text{on } \{\mathscr{C}(X_{\sigma_n}) = \mathscr{C}(X_{\sigma_{n-1}}) + 1\}, \\ N_{\sigma_{n-1}} - 1 & \text{on } \{\mathscr{C}(X_{\sigma_n}) = \mathscr{C}(X_{\sigma_{n-1}}) - 1\} \cap \{X_{\sigma_n} \neq \varnothing\} \cap A. \end{cases}$$

Provided that $N_{\sigma_n} \in \{N_{\sigma_{n-1}}, N_{\sigma_{n-1}} + 1\}$ on E, this guarantees that $\mathscr{C}(X_{\sigma_n}) \leqslant N_{\sigma_n}$. Finally, we ensure that the coupling respects the dynamics of the process N by using additional randomness where required. For example, we can proceed as follows: let B be an event that has probability $((d-\mathbf{j})\vartheta - s^{\uparrow})/(1 - s^{\uparrow} - \mathbf{j}\varpi)$ which is independent of the past of the process given $X_{\sigma_{n-1}}$ (note that the denominator in the last expression is the probability of the event E given $X_{\sigma_{n-1}} = \mathbf{x}$). Then, set $N_{\sigma_n} = N_{\sigma_{n-1}} + 1$ on $B \cap E$ and $N_{\sigma_n} = N_{\sigma_{n-1}}$ on $B^c \cap E$. By construction, we have $\mathscr{C}(X_t) \leqslant N_t$ for all $t \leqslant \tau_L \wedge \tau_\varnothing$. □

[3] For example $A = \{U \in [0, \mathbf{j}\varpi/s_{\downarrow}]\}$ for an independent uniformly distributed random variable U.

4.3.6 Proof of Lemma 4.3.7

Proof of Lemma 4.3.7. First note that since both $X^{(x)}$ and $X^{(y)}$ jump at rate one, we can couple them so that they jump at the same times, which we denote by $(\sigma_i)_{i\in\mathbb{N}}$. At the first jump, for any measurable set $A \subseteq \mathcal{C}_{d-1}$ we should have

$$\mathbb{P}(X^{(x)}_{\sigma_1} \in A) = \frac{1}{M}\sum_{i=0}^{d-1}\mathbb{E}\big[f(x_{i\leftarrow W})\mathbf{1}_A(x_{i\leftarrow W})\big];\ \mathbb{P}(X^{(y)}_{\sigma_1} \in A) = \frac{1}{M}\sum_{i=0}^{d-1}\mathbb{E}\big[f(y_{i\leftarrow W})\mathbf{1}_A(y_{i\leftarrow W})\big],$$

and both processes jump to $\{\varnothing\}$ with probability equal to the remaining mass. We can interpret these measures as sums of $d+1$ measures given by $\big(\frac{1}{M}\mathbb{E}\big[f(x_{i\leftarrow W})\delta_{x_{i\leftarrow W}}(\cdot)\big]\big)_{0\leqslant i\leqslant d-1}$ and $c(x)\delta_{\varnothing}(\cdot)$, where $c(x) := 1-\sum_{i=0}^{d-1}\mathbb{E}\big[f(x_{i\leftarrow W})\big]/M$, for $X^{(x)}$; similarly for $X^{(y)}$. On Figure 4.4, we draw the unit interval vertically and divide it in sub-intervals of respective lengths $\mathbb{E}\big[f(y_{i\leftarrow W})\big]/M$. On each of these intervals, we draw, from bottom to top as i increases from 0 to $d-1$,

$$F^{(x)}_i\colon u \mapsto b_i + \int_{[0,u]} f(x_{i\leftarrow v})\mathrm{d}\mu(v)/M \quad \left(\text{resp. } F^{(y)}_i\colon u \mapsto b_i + \int_{[0,u]} f(y_{i\leftarrow v})\mathrm{d}\mu(v)/M\right)$$

in orange (resp. purple), where for $i \in \{0,\dots,d-1\}$, $b_i = \sum_{j=0}^{i-1}\mathbb{E}\big[f(y_{j\leftarrow W})\big]/M$. Note that, by monotonicity of f, both $F^{(x)}_i$ and $F^{(y)}_i$ are non-decreasing, and since $x \leqslant y$, $F^{(x)}_i \leqslant F^{(y)}_i$ pointwise.

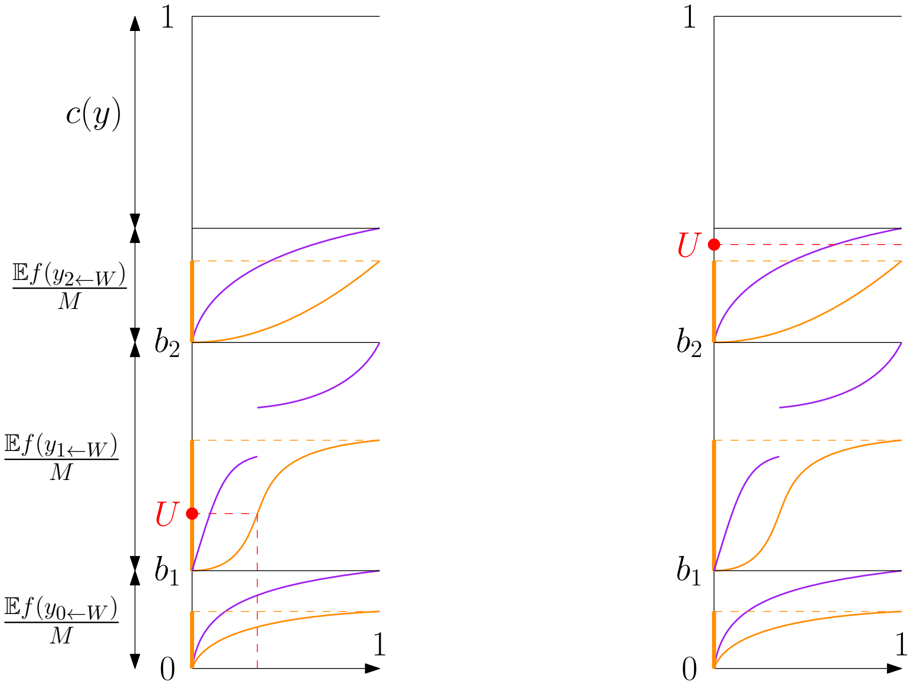


Figure 4.4: A visual aid for the proof of Lemma 4.3.7. For the sake of presentation, we have chosen $d = 3$.

Now, consider a uniformly distributed random variable U on $[0, 1]$. If U lands in the top-most interval (that is, if $U \geqslant \sum_{i=0}^{d-1} \mathbb{E}\left[f(y_{i\leftarrow W})\right]$), then we set $X_{\sigma_1}^{(x)} = X_{\sigma_1}^{(y)} = \varnothing$. If U lands in the i-th interval (numbered from the bottom of the picture), we consider two cases:

- If U lands into the orange part of the i-th interval (see left-hand-side of Figure 4.4), we set $X_{\sigma_1}^{(x)} = x_{i\leftarrow(F_i^{(x)})^{-1}(U)}$ and $X_{\sigma_1}^{(y)} = y_{i\leftarrow(F_i^{(x)})^{-1}(U)}$ (if $F_i^{(x)}$ is not strictly increasing, we choose the left-continuous version of the inverse $(F_i^{(x)})^{-1}(w) := \inf\{y \in [0, 1] \colon F_i^{(x)}(y) \geqslant w\}$).
- If U lands in the rest of the i-th interval (right-hand-side example on Figure 4.4), we set $X_{\sigma_1}^{(x)} = \varnothing$. Set $G_i = F_i^{(y)} - F_i^{(x)}$ and note that this function is non-negative on $[0, 1]$

and non-decreasing. Indeed, for all $u < v$, we have

$$G_i(v) - G_i(u) = \int_{(u,v]} \big(f(y_{i\leftarrow w}) - f(x_{i\leftarrow w})\big)\mathrm{d}\mu(w)/M \geqslant 0.$$

We can thus define the left-continuous inverse $G_i^{-1}(w) := \inf\{y \in [0,1] \colon G_i^{(x)}(y) \geqslant w\}$, and set $X_{\sigma_1}^{(y)} = y_{i\leftarrow G_i^{-1}(U - F_i^{(x)}(1))}$.

Let us prove that, with these definition, $X_{\sigma_1}^{(x)}$ and $X_{\sigma_1}^{(y)}$ have the correct distributions and that $X_{\sigma_1}^{(x)} \preccurlyeq X_{\sigma_1}^{(y)}$. First note that, if $X_{\sigma_1}^{(y)} = \varnothing$, then U fell into the topmost interval and thus $X_{\sigma_1}^{(x)} = \varnothing$, hence $X_{\sigma_1}^{(x)} \preccurlyeq X_{\sigma_1}^{(y)}$. If $X_{\sigma_1}^{(x)} \neq \varnothing$, then U fell in the orange part of an interval and thus $X_{\sigma_1}^{(x)} = x_{i\leftarrow V} \preccurlyeq y_{i\leftarrow V} = X_{\sigma_1}^{(y)}$ (where $V = (F_i^{(x)})^{-1}(U)$), since $x \preccurlyeq y$.

Let us now check that $X_{\sigma_1}^{(x)}$ defined in the coupling above has the right distribution. It is equal to $\varnothing$ if and only if U landed in the topmost interval, or it did not land in an orange sub-interval, and thus

$$\begin{aligned}
&\mathbb{P}(X_{\sigma_1}^{(x)} = \varnothing) = c(y) + \sum_{i=0}^{d-1} \big(F_i^{(y)}(1) - F_i^{(x)}(1)\big) \\
&= 1 - \frac{1}{M}\sum_{i=0}^{d-1} \mathbb{E}[f(y_{i\leftarrow W})] + \frac{1}{M}\sum_{i=0}^{d-1}\int_{[0,1]} f(y_{i\leftarrow v})\mathrm{d}\mu(v) - \frac{1}{M}\sum_{i=0}^{d-1}\int_{[0,1]} f(x_{i\leftarrow v})\mathrm{d}\mu(v) \\
&= 1 - \frac{1}{M}\sum_{i=0}^{d-1} \mathbb{E}[f(x_{i\leftarrow W})] = c(x).
\end{aligned}$$

For all Borel sets $A \subseteq \mathcal{C}_{d-1}$, we have

$$\begin{aligned}
\mathbb{P}(X_{\sigma_1}^{(x)} \in A) &= \sum_{i=0}^{d-1} \mathbb{P}(X_{\sigma_1}^{(x)} \in A \text{ and } F_i^{(x)}(0) \leqslant U \leqslant F_i^{(x)}(1)) \\
&= \sum_{i=0}^{d-1} \int_{F_i^{(x)}(0)}^{F_i^{(x)}(1)} \mathbf{1}_A\left(x_{i\leftarrow (F_i^{(x)})^{-1}(u)}\right)\mathrm{d}u \\
&= \sum_{i=0}^{d-1} \int_{[0,1]} \mathbf{1}_A(x_{i\leftarrow v}) f(x_{i\leftarrow v})\mathrm{d}\mu(v)/M,
\end{aligned}$$

by definition of $F_i^{(x)}$ and by the change of variable $u = F_i^{(x)}(v)$. This proves the claim.

Let us now check that $X_{\sigma_1}^{(y)}$ also has the right distribution under the coupling. First note that $\mathbb{P}(X^{(y)}_{\sigma_1} = \varnothing)$ is equal to the probability that U lands in the topmost interval, which is of length $c(y)$, and thus $\mathbb{P}(X^{(y)}_{\sigma_1} = \varnothing) = c(y)$.

For all Borel sets $A \subseteq \mathcal{C}_{d-1}$, we have

$$\begin{aligned}\mathbb{P}(X^{(y)}_{\sigma_1} \in A) = &\sum_{i=0}^{d-1} \mathbb{P}(X^{(y)}_{\sigma_1} \in A \text{ and } F_i^{(x)}(0) \leqslant U \leqslant F_i^{(x)}(1)) \\ &+ \sum_{i=0}^{d-1} \mathbb{P}(X^{(y)}_{\sigma_1} \in A \text{ and } F_i^{(x)}(1) < U \leqslant F_i^{(y)}(1)).\end{aligned}$$

The first sum is similar to the calculation above when checking the distribution of $X^{(x)}_{\sigma_1}$:

$$\sum_{i=0}^{d-1} \mathbb{P}(X^{(y)}_{\sigma_1} \in A \text{ and } F_i^{(x)}(0) \leqslant U \leqslant F_i^{(x)}(1)) = \frac{1}{M}\sum_{i=0}^{d-1} \mathbb{E}[f(x_{i\leftarrow W})\mathbf{1}_A(y_{i\leftarrow W})].$$

For the second sum, we have

$$\begin{aligned}&\sum_{i=0}^{d-1} \mathbb{P}(X^{(y)}_{\sigma_1} \in A \text{ and } F_i^{(x)}(1) < U \leqslant F_i^{(y)}(1)) \\ &= \sum_{i=0}^{d-1} \mathbb{P}(y_{i\leftarrow G_i^{-1}(U-F_i^{(x)}(1))} \in A \text{ and } F_i^{(x)}(1) < U \leqslant F_i^{(y)}(1)) \\ &= \sum_{i=0}^{d-1} \int_{F_i^{(x)}(1)}^{F_i^{(y)}(1)} \mathbf{1}_A\left(y_{i\leftarrow G_i^{-1}(u-F_i^{(x)}(1))}\right) \mathrm{d}u \\ &= \sum_{i=0}^{d-1} \int_{[0,1]} \mathbf{1}_A\left(y_{i\leftarrow v}\right)\,(f(y_{i\leftarrow v}) - f(x_{i\leftarrow v}))\mathrm{d}\mu(v)/M,\end{aligned}$$

by definition of G_i and by the change of variable $u = G_i(v) + F_i^{(x)}(1)$. We thus conclude that, in total,

$$\mathbb{P}(X^{(y)}_{\sigma_1} \in A) = \frac{1}{M}\sum_{i=0}^{d-1} \mathbb{E}[f(y_{i\leftarrow W})\mathbf{1}_A(y_{i\leftarrow W})],$$

as claimed. We can now iterate this coupling at each jump-time until $X^{(x)}$ becomes absorbed. After $X^{(x)}$ reaches $\varnothing$, we let $X^{(y)}$ evolve independently according to its dynamics. This concludes the proof. □

4.4 The degree profile

In this section we determine the degree profile associated with the sequence of simplicial complexes $(\mathcal{K}_n)_{n\geqslant 0}$. Throughout this section we assume that the conclusion of Theorem 4.3.1 holds, and that $f : [0, w^*]^d \to (0, \infty)$ is continuous and symmetric.

Let π^* be the distribution of the random variable $\varphi(W, Y_\infty)$, where W and Y_∞ are independent, W is μ-distributed and Y_∞ is as in Proposition 4.1.1. We now prove the following equivalent of Theorem 4.1.3; the only difference in the two statements being that we now use the notation of Section 4.3.3. In particular the process S with initial distribution π^* is equal in distribution to the process S^* from Theorem 4.1.3.

Theorem 4.4.1. *Denote by $N_k(n)$ the number of vertices of degree $d+k$ in $\mathcal{K}_n$. For all $k \geqslant 0$, we have, in probability,*

$$\lim_{n\to\infty} \frac{1}{n} N_k(n) = \mathbb{E}^*_{\pi^*}\left[\frac{\lambda}{F(S_k)+\lambda}\prod_{\ell=0}^{k-1}\frac{F(S_\ell)}{F(S_\ell)+\lambda}\right] = p_k$$

with λ as in Proposition 4.1.2.

Recall, from Remark 2.2.1 in Chapter 2 that $(p_k)_{k\geqslant 0}$ may thus be regarded as a *generalised geometric distribution*, where probability of success at the ith step is given by $\lambda/(F(S_{i-1})+\lambda)$.

The proof of Theorem 4.4.1 is analogous to the proof of Theorem 2.4.1 in Chapter 2. Recall that this approach was to first show convergence of the corresponding mean, and then study the variance of $N_k(n)$ to show convergence in probability by an application of Chebychev's inequality.

To prove convergence of the mean, as in Chapter 2, it is convenient to consider only vertices that arrive after a certain time ηn where $\eta > 0$ is a small constant; this allows us to work in the asymptotic regime of the sequence of simplicial complexes. Hence, let $N_{\eta,k}(n)$ be the number of vertices of degree $k+d$ in $\mathcal{K}_n$ which arrived after time ηn. Obviously,

$$N_{\eta,k}(n) \leqslant N_k(n) \leqslant \eta n + N_{\eta,k}(n),$$

and therefore,

$$\lim_{\eta\to 0} \limsup_{n\to\infty} \frac{1}{n} \left|\mathbb{E}\left[N_k(n)\right] - \mathbb{E}\left[N_{\eta,k}(n)\right]\right| = 0.$$

Most of this section is thus devoted to proving that, for all $k \geqslant 0$,

$$\lim_{\eta\to 0} \lim_{n\to\infty} \frac{1}{n} \mathbb{E}\left[N_{\eta,k}(n)\right] = p_k.$$

Let $\hat{d}_n(i)$ be the number of vertices which are neighbours of node i that arrived after node i. By construction, we have that

$$\mathbb{E}\left[N_{\eta,k}(n)\right] = \sum_{\eta n < i \leqslant n-k} \mathbb{P}\left(\hat{d}_n(i) = k\right). \tag{4.19}$$

Henceforth, we use the simplified notation $\mathcal{I}_k = \{i_1, \ldots, i_k\}$ for a collection of natural numbers $i < i_1 < \ldots < i_k \leqslant n$. Let $\mathcal{E}_i(\mathcal{I}_k)$ denote the event that $i \sim \ell$, that is ℓ connects to i, for all $\ell \in \mathcal{I}_k$ and $i \nsim \ell$ for all $\ell \notin \mathcal{I}_k$ with $\ell \in \{i+1, \ldots, n\}$. We have

$$\mathbb{P}\left(\hat{d}_n(i) = k\right) = \sum_{\mathcal{I}_k \in \binom{\{i+1,\ldots,n\}}{k}} \mathbb{P}\left(\mathcal{E}_i(\mathcal{I}_k)\right), \tag{4.20}$$

where $\binom{\{i+1,\ldots,n\}}{k}$ denotes the set of all subsets of $\{i+1, \ldots, n\}$ of size k. For $k = 0$, the sum consists only of the term $\mathcal{I}_0 = \varnothing$.

Overview of the proof of Theorem 4.4.1

The proof now consists of three steps. First, we provide sufficient upper and lower bounds for $\mathbb{P}(\hat{d}_n(i) = k)$ using the fact that, by Proposition 4.1.2, for $i > \eta n$, with high probability, for all $i \leqslant j \leqslant n$, the partition function $\mathcal{Z}_j$ is concentrated around λj. On the event of concentration, we can estimate the probability that insertions in the star of vertex i or its complement occur, similar to as in the proof of Theorem 2.4.1 in Chapter 2. Second, we use Proposition 4.1.1 to incorporate the stationary distribution of the Markov chain Y_n when passing to the limit as $n \to \infty$. Third, we apply a probabilistic argument to evaluate the sums in (4.19) and (4.20). In Section 4.4.1, we state the necessary tools to work out the second and third step. The proof of Proposition 4.4.2 may be omitted on first reading.

The main part of the work involves exploiting the concentration of the partition function to derive upper and lower bounds on (a variant of) $\mathbb{P}(\mathcal{E}_i(\mathcal{I}_k))$ and are proved in Section 4.4.2 and Section 4.4.4, respectively. Note that the proof of the upper bound in Section 4.4.2 is significantly less technical, as we can 'drop' the event of concentration from probability computations. We recommend the reader to study this case first. Second moment calculations which allow one to deduce stochastic convergence from convergence of the mean in Theorem 4.4.1 are presented in Section 4.4.3 and follow the arguments developed in Section 4.4.2 closely. The proof of the lower bound in Section 4.4.4 deviates from the indirect approach used in the proof of Theorem 2.4.1 in Section 2.4.4, and directly estimates the aforementioned variant of $\mathbb{P}(\mathcal{E}_i(\mathcal{I}_k))$. Thus, this proof requires additional work, due, in part, to the 'migration' of faces into the complement on the event of an insertion into the star of vertex i (see Figure 4.2). We deal with this technical challenge by bounding the total number of 'descendants' of a small number of faces by the sum of geometrically distributed random variables with sufficiently small success probability in Lemma 4.4.15 and Lemma 4.4.16). The rest of the proof then involves some lengthy computations to control error terms.

4.4.1 Technical Lemmas used in the proof of Theorem 4.4.1

This subsection is dedicated to the statements of some technical lemmas that will be important in the sequel. The proof of Lemma 4.4.2 may be omitted on first reading.

A Continuity Statement for the star Markov Chain

The following result concerns continuity of the k-step transition kernel of the star Markov chain with respect to its starting point. Recall that the function F is defined in (4.4), and the process $(S_n)_{n\geqslant 0}$ has been defined in Section 4.3.3.

Proposition 4.4.2. *Let $k \geqslant 0, w \in \mathbb{R}_+$ and $x, x_1, x_2, \ldots \in \mathcal{C}_{d-1}$ with $x_n \to x$. Then, in the sense of weak convergence on $\mathbb{R}_+^{k+1}$, we have, as $n \to \infty$,*

$$\mathbb{P}^*_{\varphi(w,x_n)}((F(S_0), F(S_1), \ldots, F(S_k)) \in \cdot) \to \mathbb{P}^*_{\varphi(w,x)}((F(S_0), F(S_1), \ldots, F(S_k)) \in \cdot).$$

Proof. Let $\mathcal{C}'_f \subseteq \mathcal{C}'$ be the set of elements of the form $(z, \sum_{i=1}^m \delta_{y_i})$ for $z \geqslant 0, m \geqslant 1$ and $y_1, y_2, \ldots, y_m \in \mathcal{C}_{d-2}$. Here, we view $\mathcal{M}(\mathcal{C}_{d-2})$ as a metric space under the Prokhorov metric, and view $\mathcal{C}' = \mathbb{R}_+ \times \mathcal{M}(\mathcal{C}_{d-2})$ as a product metric space with ∞ product metric (where the distance is the maximum co-ordinate wise distance). First of all, we prove that there exists a function $h : \mathcal{C}'_f \times [0,1] \times \mathbb{R}_+ \to \mathcal{C}'_f$ such that, for independent and identically distributed random variables $(U_1, W_1), (W_2, U_2) \ldots$, where U_i, W_i are independent, U_i has the uniform distribution on $[0, 1]$ and W_i follows the distribution μ (as before), we obtain a realisation of the Markov chain starting at $x' \in \mathcal{C}'_f$ by setting $S_0 = x'$ and, recursively, $S_{n+1} = h(S_n, U_{n+1}, W_{n+1})$ for $n \geqslant 0$. We then couple the two Markov chains started at $\varphi(w, x_n)$ and $\varphi(w, x)$ using the same sequence $(U_1, W_1), (U_2, W_2), \ldots$, and write $S_0^{(n)}, S_1^{(n)}, \ldots$ and $S_0, S_1, \ldots$ for these chains. The construction of h is straightforward. Let $x' = (z, \nu) \in \mathcal{C}'_f$

with $\nu = \sum_{i=1}^{m} \delta_{y_i} \in \mathcal{C}'_f$ and $u \in [0,1]$, $w' \geqslant 0$. Order $y_1, \ldots, y_m$ lexicographically and define

$$s_0 = 0 \text{ and } s_i = \sum_{j=1}^{i} f(y_j \cup z), 1 \leqslant i \leqslant m. \tag{4.21}$$

Then, let $1 \leqslant p \leqslant m$ be such that $s_{p-1} < us_m \leqslant s_p$. We now set

$$h((z,\nu),u,w') = \begin{cases} (z, \nu + \sum_{i=0}^{d-2} \delta_{(y_p)_{i\leftarrow w'}}), & \text{in Model } \mathbf{A}, \\ (z, \nu + \sum_{i=0}^{d-2} \delta_{(y_p)_{i\leftarrow w'}} - \delta_{y_p}), & \text{in Model } \mathbf{B}. \end{cases}$$

It follows immediately from the dynamics of the Markov chain, that the function h has the desired properties. Next, we show that, for the coupled Markov chains:

$$\text{for any } k \geqslant 0, \text{ we have } S_k^{(n)} \to S_k \text{ almost surely.} \tag{4.22}$$

By continuity of f, this implies that $F(S_k^{(n)}) \to F(S_k)$ almost surely, which concludes the proof. To prove (4.22), we proceed by induction. The case $k = 0$ is trivial as the function φ is continuous. Assume that we have already proved the statement for all $j \in \{0, \ldots, k-1\}$. Recall that $S_k = h(S_{k-1}, U_k, W_k)$ and $S_k^{(n)} = h(S_{k-1}^{(n)}, U_k, W_k)$. Conditioning on $S_{k-1}, S_{k-1}^{(0)}, S_{k-1}^{(1)}, \ldots$ shows that

$$\mathbb{P}\left(S_k^{(n)} \nrightarrow S_k\right) \leqslant \mathbb{E}[\mathrm{Leb}(\{u \in [0,1] : \text{ there exist } v_1, v_2, \ldots \in \mathcal{C}'_f \text{ and } w' \geqslant 0$$
$$\text{such that } \lim_{\ell\to\infty} v_\ell = S_{k-1} \text{ but } h(v_\ell, u, z) \nrightarrow h(S_{k-1}, u, z)\})]$$

We conclude the proof by showing that, almost surely, the set $u \in [0,1]$ for which $v_\ell, \ell \geqslant 1$ and $w' \geqslant 0$ exist satisfying $v_\ell \to S_{k-1}$ as $\ell \to \infty$ and $h(v_\ell, u, w') \nrightarrow h(S_{k-1}, u, w')$ is a Lebesgue null set. To this end, we prove the following stronger statement: for $x' = (z, \sum_{i=1}^{m} \delta_{y_i}) \in \mathcal{C}'_f$, we have that, for all $u \notin \{s_1/s_m, \ldots, 1\}$, where $s_1, \ldots, s_m$ are as in (4.21) for this particular x', it holds that, for any sequence $x'_\ell \to x'$ and $w' \geqslant 0$, we have $h(x'_\ell, u, w') \to h(x', u, w')$. To see this, let $x'_\ell = (z_\ell, \sum_{i=1}^{m_\ell} \delta_{y_i^{(\ell)}})$ be a sequence with $x'_\ell \to x'$. This implies that $m_n = m$ for all sufficiently large n and that $y_i^{(\ell)} \to y_i$ for all $1 \leqslant i \leqslant m$ as $\ell \to \infty$. By continuity of f, for the values $s_i^{(\ell)}$ defined in (4.21) for x'_ℓ, we have $s_i^{(\ell)} \to s_i$ for all $1 \leqslant i \leqslant m$. Hence, if $u \notin \{s_1/s_m, \ldots, 1\}$, again using continuity, we have that $p^{(\ell)} = p$ for all ℓ sufficiently large and the desired result follows. □

Summation Arguments

Here, we recall the statements of Lemma 2.4.5 and Corollary 2.4.6, which were proved in Section 2.4.2 of Chapter 2. Recall that for $e_0, \ldots, e_k \geqslant 0, 0 \leqslant \eta < 1$, let

$$\mathcal{S}_n(e_0, \ldots, e_k, \eta) := \frac{1}{n} \sum_{\eta n < i_0 < \cdots < i_k \leqslant n} \prod_{j=0}^{k-1} \left(\left(\frac{i_j}{i_{j+1}} \right)^{e_j} \cdot \frac{1}{i_{j+1} - 1} \right) \left(\frac{i_k}{n} \right)^{e_k}.$$

Lemma 4.4.3. *Uniformly in* $e_0, \ldots, e_k \geqslant 0$, $0 \leqslant \eta \leqslant 1/2$, *we have*

$$\mathcal{S}_n(e_0, \ldots, e_k, \eta) = \prod_{j=0}^{k} \frac{1}{e_j + 1} + \theta(\eta) + O\left(\frac{1}{n^{1/(k+2)}} + \frac{\sum_{j=0}^{k} e_j \log^{k+1}(n)}{n} \right).$$

Here, $\theta(\eta)$ *is a term satisfying* $|\theta(\eta)| \leqslant M\eta^{1/(k+2)}$ *for some universal constant* M *depending only on* k.

Corollary 4.4.4. *For* $e_0, \ldots, e_k, f_0, \ldots, f_{k-1} \geqslant 0$, $0 \leqslant \eta \leqslant 1/2$, *we have*

$$\begin{aligned} \frac{1}{n} \sum_{\eta n < i_0 \leqslant n} \sum_{\mathcal{I}_k \in \binom{\{i_0+1, \ldots, n\}}{k}} \prod_{j=0}^{k-1} \left(\left(\frac{i_j}{i_{j+1}} \right)^{e_j} \cdot \frac{f_j}{i_{j+1} - 1} \right) \left(\frac{i_k}{n} \right)^{e_k} \\ = \frac{1}{e_k + 1} \prod_{j=0}^{k-1} \frac{f_j}{e_j + 1} + \theta'(\eta) + O\left(\frac{1}{n^{1/(k+2)}} \right). \end{aligned}$$

Here, $\theta'(\eta)$ *is a term satisfying* $|\theta'(\eta)| \leqslant M'\eta^{1/(k+2)}$ *for some universal constant* M' *depending only on* k *and* $f_0, \ldots, f_{k-1}$, *and the constant in the big* O*-term may depend on* $e_0, \ldots, e_k, f_0, \ldots, f_k$.

4.4.2 Upper Bound for the Mean of $\mathbb{E}\left[N_{\eta,k}(n)\right]/n$

The aim of this section is to prove that

$$\lim_{\eta \to 0} \limsup_{n \to \infty} \mathbb{E}\left[N_{\eta,k}(n)\right]/n \leqslant p_k. \tag{4.23}$$

Recall that we write $\Pi_n = \sum_{\sigma \in \mathcal{K}_n^{(d-1)}} \delta_{w(\sigma)}$ for the empirical distribution of the weights of all $(d-1)$-faces in the complex after the nth step. We also define the partition function associated with $\mathcal{K}_n$ by $\mathcal{Z}_n = \int_{\mathcal{C}_{d-1}} f(x)\mathrm{d}\Pi_n(x)$. For $\varepsilon > 0$ and $n \geqslant 0$ and natural numbers $N_1 \leqslant N_2$, we let

$$\mathcal{G}_\varepsilon(n) = \{|\mathcal{Z}_n - \lambda n| < \varepsilon\lambda n\} \quad \text{ and } \quad \mathcal{G}_\varepsilon(N_1, N_2) = \bigcap_{n=N_1}^{N_2} \mathcal{G}_\varepsilon(n). \tag{4.24}$$

Moreover, for $n \geqslant 1$, we denote by $\mathscr{G}_n$ the σ-field generated by $(\mathcal{K}_\ell, W_\ell), 1 \leqslant \ell \leqslant n$ containing all information about the process up to time n.

By Proposition 4.1.2 and Egorov's theorem, for any $\delta, \varepsilon > 0$, there exists $N' = N'(\delta, \varepsilon)$ such that, for all $n \geqslant N'$, $\mathbb{P}(\mathcal{G}_\varepsilon(N', n)) \geqslant 1 - \delta$. Therefore, for all $n \geqslant N'/\eta$, we have

$$\begin{aligned}
\mathbb{E}\left[N_{\eta,k}(n)\right] &\leqslant \mathbb{E}\left[N_{\eta,k}(n)\mathbf{1}_{\mathcal{G}_\varepsilon(N',n)}\right] + n(1 - \mathbb{P}(\mathcal{G}_\varepsilon(N', n))) \\
&\leqslant \sum_{\eta n < i \leqslant n} \sum_{\mathcal{I}_k \in \binom{\{i+1,\dots,n\}}{k}} \mathbb{P}(\mathcal{E}_i(\mathcal{I}_k) \cap \mathcal{G}_\varepsilon(i, n)) + \delta n.
\end{aligned} \tag{4.25}$$

Finally, for $x > 0$ and $\alpha \in \mathbb{R}$, we set $\alpha_{\pm x} := \alpha(1 \pm x)$. The following proposition gives an upper bound on the summands in the right-hand side of (4.25). For simplicity, we subsequently write

$$\mathrm{st}_i(\mathcal{K}_n) = \Big(W_i, \sum_{\sigma \in \mathrm{st}_i(\mathcal{K}_n)} \delta_{\omega(\sigma)\setminus\{W_i\}}\Big) \in \mathcal{C}' = \mathbb{R}_+ \times \mathcal{M}(\mathcal{C}_{d-2}) \tag{4.26}$$

when considering the $\mathcal{C}'$-valued random variable associated with the star around vertex i at step n.

Proposition 4.4.5. *Let $0 < \varepsilon, \eta \leqslant 1/2$. As $n \to \infty$, uniformly in $\eta n < i \leqslant n - k$, $\mathcal{I}_k = \{i_0, \dots, i_{k-1}\} \in \binom{\{i+1,\dots,n\}}{k}$ and the choice of ε, we have*

$$\begin{aligned}
&\mathbb{P}\left(\mathcal{E}_i(\mathcal{I}_k) \cap \mathcal{G}_\varepsilon(i, n)\right) \\
&\leqslant \left(1 + O\left(\frac{1}{n}\right)\right) \mathbb{E}\left[\mathbb{E}^*_{\mathrm{st}_i(\mathcal{K}_i)}\left[\left(\frac{i_k}{i_{k+1}}\right)^{F(S_k)/\lambda_{+\varepsilon}} \cdot \prod_{\ell=0}^{k-1} \left(\frac{i_\ell}{i_{\ell+1}}\right)^{F(S_\ell)/\lambda_{+\varepsilon}} \frac{F(S_\ell)}{\lambda_{-\varepsilon}(i_{\ell+1} - 1)}\right]\right].
\end{aligned}$$

Applying Corollary 4.4.4 to this, we will deduce the following upper bound.

Corollary 4.4.6. *Let $0 < \delta, \varepsilon, \eta \leqslant 1/2$. Then, there exists $N = N(\delta, \varepsilon, \eta)$ such that, for all $n \geqslant N$,*

$$\frac{\mathbb{E}\left[N_{\eta,k}(n)\right]}{n} \leqslant (1+\delta)\left(\frac{1+\varepsilon}{1-\varepsilon}\right)^k \mathbb{E}^*_{\pi^*}\left[\frac{\lambda_{+\varepsilon}}{F(S_k)+\lambda_{+\varepsilon}}\prod_{\ell=0}^{k-1}\frac{F(S_\ell)}{F(S_\ell)+\lambda_{+\varepsilon}}\right] + C\eta^{1/(k+2)} + \delta,$$

where the constant C may depend on k, f and μ but not on n and not on the choices of $\delta, \varepsilon, \eta$. In particular, (4.23) *is satisfied.*

To prove Proposition 4.4.5, let $0 < \varepsilon, \eta \leqslant 1/2$. For $\eta n < i \leqslant n$ and $\mathcal{I}_k \in \binom{\{i+1,\dots,n\}}{k}$, set $i_0 := i, i_{k+1} := n+1$. Then, for $j \in \{i+1, \dots, n\}$, let

$$\mathcal{D}_j := \begin{cases} \{i \sim j\}, & \text{if } j \in \mathcal{I}_k, \\ \{i \nsim j\}, & \text{otherwise,} \end{cases} \qquad \text{and} \qquad \tilde{\mathcal{D}}_j = \mathcal{D}_j \cap \mathcal{G}_\varepsilon(j), \tag{4.27}$$

where $\mathcal{G}_\varepsilon(j)$ is defined as in (4.24). For simplicity, we write D_j and $\tilde{D}_j$ for the indicator random variables $\mathbf{1}_{\mathcal{D}_j}$ and $\mathbf{1}_{\tilde{\mathcal{D}}_j}$ respectively. Note that $\mathcal{E}_i(\mathcal{I}_k) \cap \mathcal{G}_\varepsilon(i,n) = \bigcap_{j=i}^n \tilde{\mathcal{D}}_j$. To estimate the probability of this event, we decompose the indices $j \in \{i, \dots, n\}$ into groups $\{i_\ell, \dots, i_{\ell+1}-1\}$ for $\ell \in \{0, \dots, k\}$. More precisely, we define

$$X_\ell = \mathbb{E}\left[\prod_{j=i_\ell+1}^{n} \tilde{D}_j \,\middle|\, \mathscr{G}_{i_\ell}\right] \tilde{D}_{i_\ell}, \quad \ell \in \{0, \dots, k\}.$$

To prove Proposition 4.4.5, we need to estimate $\mathbb{E}\left[X_0\right] = \mathbb{P}\left(\bigcap_{j=i}^n \tilde{\mathcal{D}}_j\right)$.

From the tower property of conditional expectation, it follows that

$$X_\ell = \mathbb{E}\left[\prod_{j=i_\ell+1}^{i_{\ell+1}-1} \tilde{D}_j \; X_{\ell+1} \,\middle|\, \mathscr{G}_{i_\ell}\right] \tilde{D}_{i_\ell}, \quad \ell \in \{0, \dots, k-1\}, \tag{4.28}$$

which suggests a backwards recursive approach. We need more notation: for $S \in \mathcal{C}' = \mathbb{R}_+ \times \mathcal{M}(\mathcal{C}_{d-2})$ and $\ell \in \{0, \dots, k\}$, we let

$$h_\ell(S) = \prod_{j=i_\ell+1}^{i_{\ell+1}-1}\left(1 - \frac{F(S)}{\lambda_{+\varepsilon}(j-1)}\right), \tag{4.29}$$

where F is as defined in (4.18), and set

$$f_k = h_k \quad \text{and} \quad f_\ell(S) = \frac{F(S)}{\lambda_{-\varepsilon}(i_{\ell+1}-1)} h_\ell(S), \quad 0 \leqslant \ell \leqslant k-1. \tag{4.30}$$

For the sake of presentation, we do not indicate that the definitions of the $\tilde{\mathcal{D}}_j, X_\ell, h_\ell, f_\ell$ depend on $\mathcal{I}_k$ and ε.

Lemma 4.4.7. *For $\ell \in \{0, \ldots, k\}$, and h_ℓ as defined in* (4.29)*, we have*

$$\mathbb{E}\left[\prod_{j=i_\ell+1}^{i_{\ell+1}-1} \tilde{D}_j \,\middle|\, \mathscr{G}_{i_\ell}\right] \leqslant h_\ell(\mathrm{st}_i(\mathcal{K}_{i_\ell})). \tag{4.31}$$

Recall that, by definition, $\mathrm{st}_i(\mathcal{K}_{i_\ell}) \in \mathcal{C}'$ (see (4.26)) and thus $h_\ell(\mathrm{st}_i(\mathcal{K}_{i_\ell}))$ is well-defined.

Proof. First note that for all $\ell \in \{1, \ldots, k\}$, by the tower property,

$$\begin{aligned}
\mathbb{E}\left[\prod_{j=i_\ell+1}^{i_{\ell+1}-1} \tilde{D}_j \,\middle|\, \mathscr{G}_{i_\ell}\right] &= \mathbb{E}\left[\mathbb{E}\left[\tilde{D}_{i_{\ell+1}-1} \,\middle|\, \mathscr{G}_{i_{\ell+1}-2}\right] \prod_{j=i_\ell+1}^{i_{\ell+1}-2} \tilde{D}_j \,\middle|\, \mathscr{G}_{i_\ell}\right] \\
&\leqslant \mathbb{E}\left[\mathbb{E}\left[D_{i_{\ell+1}-1} \,\middle|\, \mathscr{G}_{i_{\ell+1}-2}\right] \prod_{j=i_\ell+1}^{i_{\ell+1}-2} \tilde{D}_j \,\middle|\, \mathscr{G}_{i_\ell}\right],
\end{aligned}$$

where we have used the fact that, by definition, $\tilde{\mathcal{D}}_j = \mathcal{D}_j \cap \mathcal{G}_\varepsilon(j)$ and thus $\tilde{D}_j \leqslant D_j$ (recall that the latter denote the indicators of the events $\tilde{\mathcal{D}}_j$ and $\mathcal{D}_j$ respectively). If $i_{\ell+1} - 1 \notin \mathcal{I}_k$ we have that

$$\mathbb{E}\left[D_{i_{\ell+1}-1} \,\middle|\, \mathscr{G}_{i_{\ell+1}-2}\right] = \mathbb{P}(\mathcal{D}_{i_{\ell+1}-1} \,|\, \mathscr{G}_{i_{\ell+1}-2}) = 1 - \frac{F(\mathrm{st}_i(\mathcal{K}_{i_{\ell+1}-2}))}{\mathcal{Z}_{i_{\ell+1}-2}},$$

where we recall that $F(\mathrm{st}_i(\mathcal{K}_{i_{\ell+1}-2}))$ is the sum of the fitnesses of the faces in the complex that contains node i at time $i_{\ell+1} - 2$ (see (4.4)). Thus,

$$\begin{aligned}
\mathbb{E}\left[\prod_{j=i_\ell+1}^{i_{\ell+1}-1} \tilde{D}_j \,\middle|\, \mathscr{G}_{i_\ell}\right] &\leqslant \mathbb{E}\left[\left(1 - \frac{F(\mathrm{st}_i(\mathcal{K}_{i_{\ell+1}-2}))}{\mathcal{Z}_{i_{\ell+1}-2}}\right) \prod_{j=i_\ell+1}^{i_{\ell+1}-2} \tilde{D}_j \,\middle|\, \mathscr{G}_{i_\ell}\right] \\
&\leqslant \left(1 - \frac{F(\mathrm{st}_i(\mathcal{K}_{i_\ell}))}{\lambda_{+\varepsilon}(i_{\ell+1}-2)}\right) \mathbb{E}\left[\prod_{j=i_\ell+1}^{i_{\ell+1}-2} \tilde{D}_j \,\middle|\, \mathscr{G}_{i_\ell}\right],
\end{aligned}$$

where we recall that, by definition, $\lambda_{+\varepsilon} = \lambda(1+\varepsilon)$ and $F(\mathrm{st}_i(\mathcal{K}_{i_{\ell+1}-2})) = F(\mathrm{st}_i(\mathcal{K}_{i_\ell}))$. In the last inequality, we have used the fact that on the event $\tilde{\mathcal{D}}_{i_{\ell+1}-2}$, we have $\mathcal{Z}_{i_{\ell+1}-2} \leqslant \lambda_{+\varepsilon}(i_{\ell+1}-2)$. Iterating the argument shows the claim. □

We now use the Lemma 4.4.7 to derive an almost-sure upper bound for X_ℓ.

Proposition 4.4.8. *For $\ell \in \{0,\dots,k\}$, and f_ℓ as defined in* (4.30)*, we have*

$$X_\ell \leqslant \mathbb{E}^*_{\mathrm{st}_i(\mathcal{K}_{i_\ell})}\left[\prod_{j=\ell}^{k} f_j(S_{j-\ell})\right]\tilde{D}_{i_\ell}.$$

In particular,

$$\mathbb{E}\left[X_0\right] \leqslant \mathbb{E}\left[\mathbb{E}^*_{\mathrm{st}_i(\mathcal{K}_i)}\left[\prod_{j=0}^{k} f_j(S_j)\right]\right].$$

Proof. We proceed by backwards induction. For $\ell = k$, the statement is identical to the one in Lemma 4.4.7. Now, assume the claim holds for some $1 \leqslant \ell \leqslant k$. Using (4.28) and the induction hypothesis in the second inequality, we get

$$\begin{aligned} X_{\ell-1} &= \mathbb{E}\left[\prod_{j=i_{\ell-1}+1}^{i_\ell-1} \tilde{D}_j\, X_\ell \,\middle|\, \mathscr{G}_{i_{\ell-1}}\right]\tilde{D}_{i_{\ell-1}} \\ &\leqslant \mathbb{E}\left[\mathbb{E}\left[\mathbb{E}^*_{\mathrm{st}_i(\mathcal{K}_{i_\ell})}\left[\prod_{j=\ell}^{k} f_j(S_{j-\ell})\right] D_{i_\ell} \,\middle|\, \mathscr{G}_{i_\ell-1}\right]\prod_{j=i_{\ell-1}+1}^{i_\ell-1}\tilde{D}_j \,\middle|\, \mathscr{G}_{i_{\ell-1}}\right]\tilde{D}_{i_{\ell-1}}. \end{aligned} \tag{4.32}$$

The event $\mathcal{D}_{i_\ell} = \{i_\ell \sim i\}$ indicates that an insertion has been made into $\mathrm{st}_i(\mathcal{K}_{i_\ell-1})$. Therefore, conditionally on $\mathscr{G}_{i_\ell-1}$, on the event $\mathcal{D}_{i_\ell}$, the sequence $(S_0,\dots,S_{k-\ell})$ initiated by $\mathrm{st}_i(\mathcal{K}_{i_\ell})$ is equal in distribution to $(S_1,\dots,S_{k-\ell+1})$ initiated by $\mathrm{st}_i(\mathcal{K}_{i_\ell-1})$. Thus,

$$\begin{aligned} \mathbb{E}\left[\mathbb{E}^*_{\mathrm{st}_i(\mathcal{K}_{i_\ell})}\left[\prod_{j=\ell}^{k} f_j(S_{j-\ell})\right] D_{i_\ell} \,\middle|\, \mathscr{G}_{i_\ell-1}\right] &= \mathbb{P}\left(\mathcal{D}_{i_\ell} \,|\, \mathscr{G}_{i_\ell-1}\right)\mathbb{E}^*_{\mathrm{st}_i(\mathcal{K}_{i_\ell-1})}\left[\prod_{j=\ell}^{k} f_j(S_{j-\ell+1})\right] \\ &= \frac{F(\mathrm{st}_i(\mathcal{K}_{i_\ell-1}))}{\mathcal{Z}_{i_\ell-1}}\mathbb{E}^*_{\mathrm{st}_i(\mathcal{K}_{i_\ell-1})}\left[\prod_{j=\ell}^{k} f_j(S_{j-\ell+1})\right]. \end{aligned} \tag{4.33}$$

On the other hand, on the events $\bar{\mathcal{D}}_j$, $j \in \{i_{\ell-1}+1,\dots,i_\ell-1\}$, we have $\mathrm{st}_i(\mathcal{K}_{i_\ell-1}) = \mathrm{st}_i(\mathcal{K}_{i_{\ell-1}})$, and thus $F(\mathrm{st}_i(\mathcal{K}_{i_\ell-1})) = F(\mathrm{st}_i(\mathcal{K}_{i_{\ell-1}}))$. Combining (4.32) and (4.33) and the fact that on

$\tilde{\mathcal{D}}_{i_\ell - 1}$, $\mathcal{Z}_{i_\ell - 1} \geqslant \lambda_{-\varepsilon}(i_\ell - 1)$ in the first inequality, we obtain

$$\begin{aligned} X_{\ell-1} &\leqslant \mathbb{E}^*_{\mathrm{st}_i(\mathcal{K}_{i_{\ell-1}})}\left[\frac{F(S_0)}{\lambda_{-\varepsilon}(i_\ell - 1)}\prod_{j=\ell}^{k} f_j(S_{j-\ell+1})\right]\mathbb{E}\left[\prod_{j=i_{\ell-1}+1}^{i_\ell - 1}\tilde{D}_j \,\middle|\, \mathscr{G}_{i_{\ell-1}}\right]\tilde{D}_{i_{\ell-1}} \\ &\overset{(4.31)}{\leqslant} \mathbb{E}^*_{\mathrm{st}_i(\mathcal{K}_{i_{\ell-1}})}\left[\frac{F(S_0)}{\lambda_{-\varepsilon}(i_\ell - 1)}\prod_{j=\ell}^{k} f_j(S_{j-\ell+1})\right] h_{\ell-1}(\mathrm{st}_i(\mathcal{K}_{i_{\ell-1}}))\tilde{D}_{i_{\ell-1}} \\ &= \mathbb{E}^*_{\mathrm{st}_i(\mathcal{K}_{i_{\ell-1}})}\left[\prod_{j=\ell-1}^{k} f_j(S_{j-\ell+1})\right]\tilde{D}_{i_{\ell-1}}. \end{aligned}$$

This concludes the induction argument, and thus the proof. □

The following elementary lemma is an easy consequence of Stirling's approximation, using (4.8), so we state it without proof.

Lemma 4.4.9. *Let $\delta, C > 0$. Then, as $m \to \infty$, uniformly over $\delta m \leqslant a \leqslant b$ and $0 \leqslant \beta \leqslant C$, we have*

$$\prod_{j=a+1}^{b-1}\left(1 - \frac{\beta}{j-1}\right) = \left(\frac{a}{b}\right)^{\beta}\left(1 + O\left(\frac{1}{m}\right)\right).$$

The statement of Proposition 4.4.5 follows immediately from Proposition 4.4.8 and Lemma 4.4.9.

Proof of Corollary 4.4.6. In view of the statement of Proposition 4.4.5, it remains to replace $\mathrm{st}_i(\mathcal{K}_i)$ by its distributional limit $\varphi(W, Y_\infty)$ and to evaluate the sum over the possible values of $i, i_1, \dots, i_k$. We start with the first task and show that, for any $0 < \delta, \varepsilon, \eta \leqslant 1/2$, there exists $N = N(\delta, \eta)$ such that, for all $\eta n < i \leqslant n - k, \mathcal{I}_k \in \binom{\{i+1,\dots,n\}}{k}$ and $n \geqslant N$, we have

$$\begin{aligned} &\mathbb{P}\left(\mathcal{E}_i(\mathcal{I}_k) \cap \mathcal{G}_\varepsilon(i,n)\right) \\ &\qquad \leqslant (1+\delta)\,\mathbb{E}^*_{\pi^*}\left[\left(\frac{i_k}{i_{k+1}}\right)^{F(S_k)/\lambda_{+\varepsilon}} \cdot \prod_{\ell=0}^{k-1}\left(\frac{i_\ell}{i_{\ell+1}}\right)^{F(S_\ell)/\lambda_{+\varepsilon}} \frac{F(S_\ell)}{\lambda_{-\varepsilon}(i_{\ell+1}-1)}\right]. \end{aligned} \tag{4.34}$$

Note that the statement of Corollary 4.4.6 follows immediately from this identity and Corollary 4.4.4. To verify the last statement, let π_n^* be the law of $\mathrm{st}_n(\mathcal{K}_n)$ considered as

$\mathcal{C}'$-valued random variable, that is, $\varphi(W_n, Y_n)$ (see (4.16) for the definition of φ). Thanks to Proposition 4.4.5, it is sufficient to prove that, uniformly in $\eta n < i < i_1 < i_2 < \ldots < i_k \leqslant n$ and $\varepsilon \in (0, 1/2]$, as $n \to \infty$

$$\mathbb{E}^*_{\pi^*_i}\left[\left(\frac{i_k}{i_{k+1}}\right)^{F(S_k)/\lambda_{+\varepsilon}} \cdot \prod_{\ell=0}^{k-1}\left(\frac{i_\ell}{i_{\ell+1}}\right)^{F(S_\ell)/\lambda_{+\varepsilon}} F(S_\ell)\right] - \mathbb{E}^*_{\pi^*}\left[\left(\frac{i_k}{i_{k+1}}\right)^{F(S_k)/\lambda_{+\varepsilon}} \cdot \prod_{\ell=0}^{k-1}\left(\frac{i_\ell}{i_{\ell+1}}\right)^{F(S_\ell)/\lambda_{+\varepsilon}} F(S_\ell)\right] \to 0. \tag{4.35}$$

To this end, we prove the following stronger statement: uniformly in $\eta \leqslant x_0, \ldots, x_k \leqslant 1$ and the choice of ε, as $n \to \infty$,

$$\mathbb{E}^*_{\pi^*_n}\left[x_k^{F(S_k)/\lambda_{+\varepsilon}} \cdot \prod_{\ell=0}^{k-1} x_\ell^{F(S_\ell)/\lambda_{+\varepsilon}} F(S_\ell)\right] - \mathbb{E}^*_{\pi^*}\left[x_k^{F(S_k)/\lambda_{+\varepsilon}} \cdot \prod_{\ell=0}^{k-1} x_\ell^{F(S_\ell)/\lambda_{+\varepsilon}} F(S_\ell)\right] \to 0.$$

By continuity of φ, Proposition 4.1.1 and Proposition 4.4.2, we have $\mathbb{P}^*_{\pi^*_n}((F(S_0), \ldots, F(S_k)) \in \cdot) \to \mathbb{P}^*_{\pi^*}((F(S_0), \ldots, F(S_k)) \in \cdot)$ weakly. Note that, for all $0 \leqslant \ell \leqslant k$, $F(S_\ell) \leqslant C$, where $C = (d+1)(k+1)f_{\max}$ and we recall that $f_{\max}$ is the maximum of the fitness function f. For all $\eta \leqslant x_0, \ldots, x_k \leqslant 1$ and $0 \leqslant \varepsilon \leqslant 1/2$, the function $J(y_0, \ldots, y_k) = x_k^{y_k/\lambda_{+\varepsilon}} \cdot \prod_{\ell=0}^{k-1} x_\ell^{y_\ell/\lambda_{+\varepsilon}} y_\ell$ defined on $[0, C]^{k+1}$ satisfies

$$\|\nabla J\| \leqslant \alpha_\eta := C^k\,(1 - \log \eta/\lambda) \tag{4.36}$$

uniformly in $x_0, \ldots, x_k, \varepsilon$. For any two probability distributions ν and ν' on $[0, C]^{k+1}$, let

$$d(\nu, \nu') = \sup_{g \in \mathcal{F}}\left|\int g \mathrm{d}\nu - \int g \mathrm{d}\nu'\right| \tag{4.37}$$

$$\text{where } \mathcal{F} := \{g : [0, C]^{k+1} \to \mathbb{R} \mid \forall x, y \in [0, C]^{k+1} \quad |g(x) - g(y)| \leqslant \alpha_\eta \|x - y\|\}.$$

It is well-known that $d(\nu_n, \nu) \to 0$ if and only if $\nu_n \to \nu$ weakly (see for example, Example 19, page 74 [70]). This concludes the proof of (4.35) and of Corollary 4.4.6. □

4.4.3 Stochastic convergence: second moment calculations

By counting the number of unordered pairs of vertices with degree $d + k$, arguments similar to those applied in Section 4.4.2 allow us to compute asymptotically the second moment of

$N_{\eta,k}(n)$ (recall this is the number of vertices of degree $k+d$ in $\mathcal{K}_n$ that arrived after time ηn). Note that

$$\mathbb{E}\left[(N_{\eta,k}(n))^2\right] = \sum_{\eta n < i,j \leqslant n} \mathbb{P}\left(\hat{d}_n(i) = k, \hat{d}_n(j) = k\right).$$

We prove that

$$\lim_{\eta\to 0}\limsup_{n\to\infty} \frac{\mathbb{E}\left[(N_{\eta,k}(n))^2\right]}{n^2} \leqslant p_k^2. \tag{4.38}$$

This shows that $\lim_{n\to\infty} \mathbb{E}\left[(N_{\eta,k}(n))^2\right]/n^2 = p_k^2$ which is sufficient to deduce the convergence in probability stated in Theorem 4.4.1 from convergence of the mean by a standard application of Chebychev's inequality.

Recall that we use the notation $\mathcal{I}_k = \{i_1, \ldots, i_k\}$ for a collection of natural numbers $i < i_1 < \ldots < i_k \leqslant n$. Similarly, we write $\mathcal{J}_k = \{j_1, \ldots, j_k\}$ for a collection of natural numbers such that $j < j_1 < \ldots < j_k \leqslant n$. As before, we let $\mathcal{E}_i(\mathcal{I}_k)$ denote the event $i \sim \ell$ for $i < \ell \leqslant n$ if and only if $\ell \in \mathcal{I}_k$ and define the event $\mathcal{E}_j(\mathcal{J}_k)$ analogously for $j, j_1, \ldots, j_k$.

With these definitions, we have

$$\mathbb{E}\left[(N_{\eta,k}(n))^2\right] = \sum_{\eta n < i,j \leqslant n} \sum_{\mathcal{I}_k, \mathcal{J}_k} \mathbb{P}\left(\mathcal{E}_i\left(\mathcal{I}_k\right) \cap \mathcal{E}_j\left(\mathcal{J}_k\right)\right), \tag{4.39}$$

where the inner sum is over all $\mathcal{I}_k \in \binom{\{i+1,\ldots,n\}}{k}$ and $\mathcal{J}_k \in \binom{\{j+1,\ldots,n\}}{k}$. As in Section 4.4.2, we fix $0 \leqslant \delta, \varepsilon \leqslant 1/2$ and choose N' such that for all $n \geqslant N'$, $\mathbb{P}\left(\mathcal{G}_\varepsilon(N', n)\right) \geqslant 1 - \delta$.

Note that, on $\mathcal{E}_i(\mathcal{I}_k) \cap \mathcal{E}_j(\mathcal{J}_k)$, if $\mathcal{I}_k \cap \mathcal{J}_k \neq \varnothing$ we either have $i = j$ or $i \sim j$. If $i = j$ then $\mathcal{I}_k = \mathcal{J}_k$, and the contribution of these terms to the right hand side of (4.39) is at most $\mathbb{E}\left[N_{\eta,k}(n)\right] \leqslant n$. On the event $\{\hat{d}_n(i) = k\}$ we have $F(\mathrm{st}_i(\mathcal{K}_\ell)) \leqslant (k+1) d f_{\max}$ for all $i + 1 \leqslant \ell \leqslant n$. Therefore, for $\eta n < i < j \leqslant n$, we have

$$\begin{aligned} &\mathbb{P}\left(\left\{\hat{d}_n(i) = k\right\} \cap \left\{\hat{d}_n(j) = k\right\} \cap \{j \sim i\} \cap \mathcal{G}_\varepsilon(i,n)\right) \\ &\qquad\qquad \leqslant \mathbb{P}\left(\{j \sim i\} \mid \mathcal{G}_\varepsilon\left(i, j-1\right), \hat{d}_{j-1}(i) \leqslant k\right) \leqslant \frac{(k+1) d f_{\max}}{\lambda_{-\varepsilon} \eta n}. \end{aligned}$$

It follows that, for all n sufficiently large, depending on δ, ε and η,

$$\mathbb{E}\left[(N_{\eta,k}(n))^2\right] \leqslant 2 \sum_{\eta n < i < j \leqslant n} \sum_{\mathcal{I}_k \cap \mathcal{J}_k = \varnothing} \mathbb{P}\left(\mathcal{E}_i\left(\mathcal{I}_k\right) \cap \mathcal{E}_j\left(\mathcal{J}_k\right) \cap \mathcal{G}_\varepsilon(i,n)\right) + \delta n^2 + Cn/\eta,$$

for a constant $C \geqslant 0$ which is independent of n, δ, ε and η. The following proposition is the analogue of Proposition 4.4.5.

Proposition 4.4.10. *Let $0 < \varepsilon, \eta \leqslant 1/2$. As $n \to \infty$, uniformly in $\eta n < i < j \leqslant n-k$, $\mathcal{I}_k \in \binom{\{i+1,\ldots,n\}}{k}$ and $\mathcal{J}_k \in \binom{\{j+1,\ldots,n\}}{k}$ with $\mathcal{I}_k \cap \mathcal{J}_k = \varnothing$ and the choice of ε, we have*

$$\begin{aligned}
&\mathbb{P}\left(\mathcal{E}_i\left(\mathcal{I}_k\right) \cap \mathcal{E}_j\left(\mathcal{J}_k\right) \cap \mathcal{G}_\varepsilon(i,n)\right) \\
&\leqslant \left(1 + O\left(\frac{1}{n}\right)\right) \mathbb{E}\Bigg[\mathbb{E}^*_{\mathrm{st}_i(\mathcal{K}_i)}\left[\left(\frac{i_k}{n}\right)^{F(S_k)/\lambda_{+\varepsilon}} \cdot \prod_{\ell=0}^{k-1}\left(\frac{i_\ell}{i_{\ell+1}}\right)^{F(S_\ell)/\lambda_{+\varepsilon}} \frac{F(S_\ell)}{\lambda_{-\varepsilon}(i_{\ell+1}-1)}\right] \\
&\qquad \mathbb{E}^*_{\mathrm{st}_j(\mathcal{K}_j)}\left[\left(\frac{j_k}{n}\right)^{F(S_k)/\lambda_{+\varepsilon}} \cdot \prod_{\ell=0}^{k-1}\left(\frac{j_\ell}{j_{\ell+1}}\right)^{F(S_\ell)/\lambda_{+\varepsilon}} \frac{F(S_\ell)}{\lambda_{-\varepsilon}(j_{\ell+1}-1)}\right]\Bigg].
\end{aligned}$$

The proof of this proposition is completely analogous to the proof of Proposition 4.4.5 and relies on a backward induction argument and an application of Lemma 4.4.9. We omit the details as no new arguments are necessary at this point. We move on to show the following analogue of (4.34): for any $0 < \delta, \varepsilon, \eta \leqslant 1/2$, there exists $N = N(\delta, \eta)$ such that, for all $n \geqslant N$, $\eta n < i < j \leqslant n-k$ and disjoint sets $\mathcal{I}_k, \mathcal{J}_k$, we have

$$\begin{aligned}
&\mathbb{P}\left(\mathcal{E}_i\left(\mathcal{I}_k\right) \cap \mathcal{E}_j\left(\mathcal{J}_k\right) \cap \mathcal{G}_\varepsilon(i,n)\right) \\
&\leqslant (1+\delta)\Bigg(\mathbb{E}^*_{\pi^*}\left[\left(\frac{i_k}{n}\right)^{F(S_k)/\lambda_{+\varepsilon}} \cdot \prod_{\ell=0}^{k-1}\left(\frac{i_\ell}{i_{\ell+1}}\right)^{F(S_\ell)/\lambda_{+\varepsilon}} \frac{F(S_\ell)}{\lambda_{-\varepsilon}(i_{\ell+1}-1)}\right] \\
&\qquad \mathbb{E}^*_{\pi^*}\left[\left(\frac{j_k}{n}\right)^{F(S_k)/\lambda_{+\varepsilon}} \cdot \prod_{\ell=0}^{k-1}\left(\frac{j_\ell}{j_{\ell+1}}\right)^{F(S_\ell)/\lambda_{+\varepsilon}} \frac{F(S_\ell)}{\lambda_{-\varepsilon}(j_{\ell+1}-1)}\right]\Bigg).
\end{aligned} \tag{4.40}$$

The details are very similar to the approach in Section 4.4.2, and we only give the necessary additional results entering the proof.

Proposition 4.4.11. *As $n, m \to \infty$ with $n \neq m$, we have $(Y_n, Y_m) \to (Y_\infty, Y'_\infty)$, in distribution, for independent random variables Y_∞, Y'_∞ both distributed according to π^*.*

Proof. This follows easily from Theorem 4.3.1. Let $g_1, g_2 : \mathcal{C}_{d-1} \to \mathbb{R}$ be bounded and continuous and Y_∞, Y'_∞ be independent realisations of π^*. We have

$$\begin{aligned} &|\mathbb{E}\left[g_1(Y_n)g_2(Y_m)\right] - \mathbb{E}\left[g_1(Y_\infty)g_2(Y'_\infty)\right]| \qquad (4.41)\\ &\qquad\leqslant |\mathbb{E}\left[g_1(Y_n)g_2(Y_m)\right] - \mathbb{E}\left[g_1(Y_n)\right]\mathbb{E}\left[g_2(Y'_\infty)\right]| \\ &\qquad\quad + \left|\mathbb{E}\left[g_1(Y_n)\right]\mathbb{E}\left[g_2(Y'_\infty)\right] - \mathbb{E}\left[g_1(Y_\infty)g_2(Y'_\infty)\right]\right|. \end{aligned}$$

Since Y_∞, Y'_∞ are independent, the second term on the right hand side is equal to

$$|\mathbb{E}\left[g_2(Y_\infty)\right]| \cdot |\mathbb{E}\left[g_1(Y_n)\right] - \mathbb{E}\left[g_1(Y_\infty)\right]|. \qquad (4.42)$$

As $n \to \infty$, (4.42) converges to zero by Theorem 4.3.1. For $n < m$, we have $\mathbb{E}\left[g_1(Y_n)g_2(Y_m)\right] = \mathbb{E}\left[g_1(Y_n)\mathbb{E}\left[g_2(Y_m) \mid \mathscr{G}_{m-1}\right]\right]$. Hence, the first term on the right hand side of (4.41) is bounded from above by

$$\|g_1\| \cdot \mathbb{E}\left[|\mathbb{E}\left[g_2(Y_m) \mid \mathscr{G}_{m-1}\right] - \mathbb{E}\left[g_2(Y_\infty)\right]|\right]. \qquad (4.43)$$

Write ν_m for the law of Y_m given $\mathscr{G}_{m-1}$, that is, for all measurable $A \subseteq \mathcal{C}_{d-1}$,

$$\nu_m(A) = \frac{\int_A f(x)\mathrm{d}\Pi_{m-1}(x)}{\int_{\mathcal{C}_{d-1}} f(x)\mathrm{d}\Pi_{m-1}(x)}.$$

By Theorem 4.3.1, we have, almost surely, $\nu_m \to \pi^*$ weakly. Thus, $\mathbb{E}\left[g_2(Y_m) \mid \mathscr{G}_{m-1}\right] \to \mathbb{E}\left[g_2(Y_\infty)\right]$. Hence, by the dominated convergence theorem, (4.43) converges to zero as $m \to \infty$. This concludes the proof for $n, m \to \infty$ with $n < m$ and the case $n > m$ can be treated analogously. □

In the remainder, we write $\mathbb{P}^{**}_{x,x'}$ and $\mathbb{E}^{**}_{x,x'}$ with $x, x' \in \mathcal{C}'$ for probabilities and expectations, respectively, involving a pair of independent copies of the star Markov chain $(S_0, S'_0), (S_1, S'_1), \ldots$, where $S_0 = x$ and $S'_0 = x'$.

Proposition 4.4.12. *Let $k \geqslant 0$, $w, w' \geqslant 0$ and $x, x', x_1, x'_1, x_2, x'_2, \ldots \in \mathcal{C}_{d-1}$ with $x_n \to x$*

and $x'_n \to x'$*. Then, in the sense of weak convergence on* $\mathbb{R}_+^{2k+2}$*, we have, as* $n \to \infty$*,*

$$\mathbb{P}^{**}_{\varphi(w,x_n),\varphi(w',x'_n)}((F(S_0), F(S'_0), F(S_1), F(S'_1), \ldots, F(S_k), F(S'_k)) \in \cdot)$$
$$\to \mathbb{P}^{**}_{\varphi(w,x),\varphi(w',x')}((F(S_0), F(S'_0), F(S_1), F(S'_1), \ldots, F(S_k), F(S'_k)) \in \cdot).$$

Proof. This follows from the independence of the two star processes involved and Proposition 4.4.2. □

Using Proposition 4.4.11 and Proposition 4.4.12, the continuity of φ, and an argument analogous to the proof of Corollary 4.4.6 (using a probability metric similar to (4.37)), (4.40) follows upon verifying the following: For any $\eta \leqslant x_0, x'_0, \ldots, x_k, x'_k \leqslant 1$ and $0 \leqslant \varepsilon \leqslant 1/2$, with the function

$$J'(y_0, y'_0, \ldots, y_k, y'_k) = x_k^{y_k/\lambda_{+\varepsilon}} \cdot \prod_{\ell=0}^{k-1} x_\ell^{y_\ell/\lambda_{+\varepsilon}} y_\ell \cdot (x'_k)^{y'_k/\lambda_{+\varepsilon}} \cdot \prod_{\ell=0}^{k-1} (x'_\ell)^{y'_\ell/\lambda_{+\varepsilon}} y'_\ell$$

defined on $[0, C]^{2k+2}$, we have that $\|\nabla J'\|$ is bounded uniformly in $x_0, \ldots, x_k$, $x'_0, \ldots, x'_k$ and ε. This follows from that the fact that J' factorizes, $\|J'\| \leqslant C^{2k}$, and (4.36).

Now, when evaluating the sum over $\eta n < i \neq j \leqslant n$ and disjoint $\mathcal{I}_k \in \binom{\{i+1,\ldots,n\}}{k}$, $\mathcal{J}_k \in \binom{\{j+1,\ldots,n\}}{k}$ in (4.40), since the summands are non-negative, and we are looking for an upper bound, we may remove the conditions $i \neq j$ and $\mathcal{I}_k \cap \mathcal{J}_k = \varnothing$. But Corollary 4.4.4 shows that, uniformly in ε and η,

$$\sum_{\eta n < i, j \leqslant n} \sum_{\mathcal{I}_k, \mathcal{J}_k} \mathbb{E}^*_{\pi^*} \left[\left(\frac{i_k}{n}\right)^{F(S_k)/\lambda_{+\varepsilon}} \cdot \prod_{\ell=0}^{k-1} \left(\frac{i_\ell}{i_{\ell+1}}\right)^{F(S_\ell)/\lambda_{+\varepsilon}} \frac{F(S_\ell)}{\lambda_{-\varepsilon}(i_{\ell+1}-1)} \right]$$
$$\times \mathbb{E}^*_{\pi^*} \left[\left(\frac{j_k}{n}\right)^{F(S_k)/\lambda_{+\varepsilon}} \cdot \prod_{\ell=0}^{k-1} \left(\frac{j_\ell}{j_{\ell+1}}\right)^{F(S_\ell)/\lambda_{+\varepsilon}} \frac{F(S_\ell)}{\lambda_{-\varepsilon}(j_{\ell+1}-1)} \right]$$
$$\leqslant \left(\frac{1+\varepsilon}{1-\varepsilon}\right)^{2k} \left(\mathbb{E}^*_{\pi^*} \left[\frac{\lambda_{+\varepsilon}}{F(S_k) + \lambda_{+\varepsilon}} \prod_{\ell=0}^{k-1} \frac{F(S_\ell)}{F(S_\ell) + \lambda_{+\varepsilon}} \right] \right)^2 + O\left(n^{-1/(k+2)}\right) + C'\eta^{1/k+2},$$

for some universal constant $C' > 0$. From here, identity (4.38) follows easily as in Section 4.4.2.

4.4.4 Lower bound for the Mean of $N_k(n)/n$

In this section, we prove that, for all $k \geqslant 0$,

$$\lim_{\eta \to 0} \liminf_{n \to \infty} \frac{\mathbb{E}\left[N_{\eta,k}(n)\right]}{n} \geqslant p_k, \tag{4.44}$$

where we recall that $N_{\eta,k}(n)$ is the number of vertices of degree $k+d$ in $\mathcal{K}_n$ that arrived after time ηn, and p_k is defined in Theorem 4.4.1. Recall that in order to prove the analogue of (4.44) with regards to the (μ, f, ℓ) - RIF tree, we adopted an indirect approach, using a proof by contradiction in Section 2.4.4 of Chapter 2. This approach is also applicable here, and the interested reader may consider applying this approach as an exercise. However, in this subsection we adopt a more direct proof of (4.44). Whilst this proof is much more technical, this approach is favourable as the techniques may transfer to the analysis of other quantities related to recursive network models, for example, the study of the evolution of the degree of a fixed vertex.

To apply this approach, we need more notation. First, let $\mathbf{C}$ be the set of all finite $(d-1)$-dimensional simplicial complexes with integer vertices. To add weights, let $\mathbf{C}^w = \mathbf{C} \times \mathbb{R}_+^{\mathbb{Z}}$, where, for $t = (c, x) \in \mathbf{C}^w$, $x_i, i \in \mathbb{Z}$ keeps track of the weight assigned to the vertex i - if no such vertex exists, set $x_i = 0$. We then consider $\mathcal{K}_n$ as a $\mathbf{C}^w$-valued random variable incorporating vertex weights. For a simplicial complex $\mathcal{K} \in \mathbf{C}$, let $\mathcal{K}_{\setminus i} := \{\sigma \in \mathcal{K} : i \notin \sigma\}$ be the sub-complex obtained from $\mathcal{K}$, when we remove the faces which contain vertex i. We set $\mathcal{K}_{\setminus i} := \mathcal{K}$ if $i \notin \mathcal{K}$. When applied to the random dynamical process, we write $\mathcal{K}_{n\setminus i}$ for $(\mathcal{K}_n)_{\setminus i}$. Let

$$\Pi_{n\setminus i} = \sum_{\sigma \in \mathcal{K}_{n\setminus i}^{(d-1)}} \delta_{\omega(\sigma)}, \text{ and } \quad \mathcal{Z}_{n\setminus i} = \int_{\mathcal{C}_{d-1}} f(x) \mathrm{d}\Pi_{n\setminus i}(x)$$

be the empirical measure of the types of active faces in $\mathcal{K}_{n\setminus i}$ and the corresponding partition function, respectively. Note that $\mathcal{K}_n^{(d-1)} = \mathcal{K}_{n\setminus i}^{(d-1)} \cup \mathrm{st}_i(\mathcal{K}_n)$, where the union is disjoint and therefore $\mathcal{Z}_n = \mathcal{Z}_{n\setminus i} + F(\mathrm{st}_i(\mathcal{K}_n))$.

To prove a suitable lower bound on the probability that vertex i receives edges at certain times, we need to control $\mathcal{Z}_{n\setminus i}$ throughout the process. It is reasonable to expect $\mathcal{Z}_{n\setminus i}$ to behave similarly to $\mathcal{Z}_n$. To this end, for all $\varepsilon > 0$, $n \geqslant i \geqslant 1$ and $m \geqslant 1$, we let

$$\mathcal{G}_\varepsilon^{(i)}(n) = \left\{\left|\mathcal{Z}_{n\setminus i} - \lambda n\right| < \varepsilon\lambda n\right\} \quad \text{and} \quad \mathcal{G}_\varepsilon(n;m) = \left\{\left|\mathcal{Z}_n - \lambda m\right| < \varepsilon\lambda m\right\}. \tag{4.45}$$

Note the difference between the notation $\mathcal{G}_\varepsilon(n;m)$ and the notation for concentration along an interval $\mathcal{G}_\varepsilon(N_1, N_2)$ defined in Section 4.4.2.

For $1 \leqslant i \leqslant n$, $\mathcal{I}_k \in \binom{\{i+1,\ldots,n\}}{k}$ and $j = i, \ldots, n$, we let

$$p(j) \in \{0, \ldots, k\} \text{ be such that } i_{p(j)} \leqslant j \leqslant i_{p(j)+1} - 1, \tag{4.46}$$

where we recall that we use the conventions $i_0 = i$ and $i_{k+1} = n + 1$.

As opposed to the arguments in Section 4.4.2, the inductive proof in this section requires us to modify the value of ε in different intervals $\{i_\ell, \ldots, i_{\ell+1} - 1\}, \ell = 0, \ldots, k$. We thus need more notation. First, for a fixed $\varepsilon > 0$, and $\ell \in \{0, \ldots, k\}$ we set $\varepsilon_\ell := (1 + \ell)\varepsilon$. We only apply this notation to the symbol ε, to avoid confusion with subscripts. Next, for $j \in \{i+1, \ldots, n\}$, recalling the events $\mathcal{D}_j$ from (4.27), and $\mathcal{G}_\varepsilon^{(i)}(j)$, $\mathcal{G}_\varepsilon(i;i)$ from (4.45), we set

$$\bar{\mathcal{D}}_j(\varepsilon) = \mathcal{D}_j \cap \mathcal{G}^{(i)}_{\varepsilon_{p(j)}}(j) \quad \text{and} \quad \bar{\mathcal{D}}_i(\varepsilon) = \mathcal{G}_\varepsilon(i;i).$$

Similarly to before, we write $D_j(\varepsilon) := \mathbf{1}_{\mathcal{D}_j(\varepsilon)}$ and $\bar{D}_j(\varepsilon) := \mathbf{1}_{\bar{\mathcal{D}}_j(\varepsilon)}$. With this notation, we have

$$\mathbb{E}\left[N_{\eta,k}(n)\right] \geqslant \sum_{\eta n < i \leqslant n} \sum_{\mathcal{I}_k \in \binom{\{i+1,\ldots,n\}}{k}} \mathbb{P}\left(\bigcap_{j=i}^{n} \bar{\mathcal{D}}_j(\varepsilon)\right). \tag{4.47}$$

We then have the following analogue of Proposition 4.4.5.

Proposition 4.4.13. *Let* $0 < \delta, \varepsilon, \eta \leqslant 1/2$. *There exists a constant* $C' > 0$, $N = N(\delta, \varepsilon, \eta)$

and $0 \leqslant \varrho \leqslant 1$ *such that, for all* $n \geqslant N$,

$$\mathbb{E}\left[N_{\eta,k}(n)\right] \geqslant -C'\delta n + \varrho(1-\delta) \cdot \sum_{\eta n < i \leqslant n} \sum_{\mathcal{I}_k \in \binom{\{i+1,\dots,n\}}{k}} \mathbb{E}\left[\mathbb{E}^*_{\mathrm{st}_i(\mathcal{K}_i)}\left[\left(\frac{i_k}{i_{k+1}}\right)^{\frac{F(S_k)}{\lambda_{-\varepsilon_k}}} \cdot \prod_{\ell=0}^{k-1} \left(\frac{i_\ell}{i_{\ell+1}}\right)^{\frac{F(S_\ell)}{\lambda_{-\varepsilon_\ell}}} \frac{F(S_\ell)}{\lambda_{+\varepsilon_\ell}(i_{\ell+1}-1)}\right]\right], \tag{4.48}$$

where ϱ *depends only on* ε, η *and, for any fixed* $0 < \eta \leqslant 1/2$, *we have* $\varrho \to 1$ *as* $\varepsilon \to 0$.

Similar arguments leading from Proposition 4.4.5 to Corollary 4.4.6 then give the following result.

Corollary 4.4.14. *Let* $0 < \delta, \varepsilon, \eta \leqslant 1/2$. *Then, there exists* $N = N(\delta, \varepsilon, \eta)$ *and a universal constant* $C > 0$ *not depending on any of these parameters, such that, for all* $n \geqslant N$,

$$\frac{\mathbb{E}\left[N_{\eta,k}(n)\right]}{n} \geqslant \varrho(1-\delta)\left(\frac{1-\varepsilon_k}{1+\varepsilon_k}\right)^k \cdot \mathbb{E}^*_{\pi^*}\left[\frac{\lambda_{-\varepsilon_k}}{F(S_k)+\lambda_{-\varepsilon_k}} \prod_{\ell=0}^{k-1} \frac{F(S_\ell)}{F(S_\ell)+\lambda_{-\varepsilon_\ell}}\right] - C(\eta^{1/(k+2)} + 1/n^{1/(k+2)}) - \delta,$$

where ϱ *is as in the Proposition 4.4.13. In particular,* (4.44) *holds.*

We now define analogues of h_ℓ and f_ℓ from (4.29) and (4.30) in Section 4.4.2. Here, however, it is necessary to indicate the dependence of these functions on ε. For $S \in \mathcal{C}'$ and $\ell \in \{0, \dots, k\}$, let

$$\mathfrak{h}^\varepsilon_\ell(S) = \prod_{j=i_\ell+1}^{i_{\ell+1}-1} \left(1 - \frac{F(S)}{\lambda_{-\varepsilon_\ell}(j-1)}\right) \tag{4.49}$$

and, for $\ell \in \{0, \dots, k-1\}$,

$$\mathfrak{f}^\varepsilon_\ell(S) = \frac{F(S)}{F(S)+\lambda_{+\varepsilon_\ell}(i_{\ell+1}-1)} \mathfrak{h}^\varepsilon_\ell(S) \quad \text{while } \mathfrak{f}^\varepsilon_k = \mathfrak{h}^\varepsilon_k. \tag{4.50}$$

We follow the arguments from the proof of the upper bound in Section 4.4.2 and show analogues of Lemma 4.4.7 and Proposition 4.4.8. To this end, we need to make use of the more general framework introduced at the beginning of this subsection: we write $\mathbb{P}_x(\cdot), \mathbb{E}_x(\cdot)$ for probabilities and expectations respectively, when the initial weighted configuration is equal to $x = (c, z)$ with $c \in \mathbf{C}, z \in \mathbb{R}_+^{\mathbb{Z}}$. Here, if $m \in \mathbb{Z}$ is the maximum vertex label occurring in c, then the vertex inserted in step i of the process carries label $m + i$. Then, for a real-valued function g depending on the path of the process and $u(x) = \mathbb{E}_x[g((\mathcal{K}_n)_{n\geqslant 0})]$, we use the slightly inaccurate but standard notation $\mathbb{E}_X[g((\mathcal{K}_n)_{n\geqslant 0})]$ for $u(X)$ and a random variable X which is typically defined in terms of $\mathcal{K}_n, n \geqslant 0$. Probabilities $\mathbb{P}$ and expectations $\mathbb{E}$ appearing in the following without subscript are with respect to the initial process with given $\mathcal{K}_0$.

Proving analogues of Lemma 4.4.7 and Proposition 4.4.8 becomes more intricate since we can no longer drop the concentration conditions relying on the events $\mathcal{G}_\varepsilon(j)$ as we did in Section 4.4.2. Nevertheless, ignoring the dependency structure of the evolution of the process in the star of vertex i and outside, intuitively we still expect to bound $\mathbb{P}\left(\bigcap_{j=i}^n \bar{\mathcal{D}}_j\right)$ from below by a term similar to

$$\mathbb{E}\left[\mathbb{E}_{\mathcal{K}_{i\setminus i}}\left[\prod_{j=i+1}^{n-k} \mathbf{1}_{\mathcal{G}_{\varepsilon_{p(j)}}(j-i;j+p(j))}\right]\mathbb{E}^*_{\mathrm{st}_i(\mathcal{K}_i)}\left[\prod_{j=0}^{k} \mathfrak{f}_j^\varepsilon(S_j)\right]\right]. \tag{4.51}$$

The two main hurdles to prove such a lower bound are the following: first, while the process outside the star of vertex i follows the Markovian transition rule, there is a subtle dependence between the star and its complement as the addition of faces to the star adds faces to its complement. More formally, on $\mathcal{D}_{i_\ell}$, we have $\mathcal{K}_{i_\ell\setminus i} \neq \mathcal{K}_{(i_\ell-1)\setminus i}$. The reason is that when a face in $\mathrm{st}_i(\mathcal{K}_{i_\ell-1})$ is subdivided during step i_ℓ, one of the faces that are created does not contain vertex i and therefore migrates into $\mathcal{K}_{i_\ell\setminus i}$ (this is the face that is removed at each step in Figure 4.2). Second, in order to exploit the concentration of the partition function $\mathcal{Z}_j$ for $j \geqslant i > \eta n$, an argument is needed to replace $\mathbb{P}_{\mathcal{K}_{i\setminus i}}$ by $\mathbb{P}_{\mathcal{K}_i}$. In order to overcome these difficulties, we use the following two lemmas, whose proofs we delay to the end of the section.

Lemma 4.4.15. *For any $\delta, \varepsilon > 0, 0 < \eta < 1$, there exists $N = N(\delta, \varepsilon, \eta)$ such that, for all*

$n \geqslant N, \eta n < i < n-k$, we have

$$\mathbb{E}\left[\mathbb{P}_{\mathcal{K}_{i\backslash i}}\left(\bigcap_{j=i+1}^{n} \mathcal{G}_{\varepsilon}(j-i;j)\right)\right] \geqslant 1-\delta.$$

Lemma 4.4.16. *For any $\varepsilon_1, \varepsilon_2, \varepsilon_3 > 0, 0 < \eta_1 < 1$ and $C_1, C_2 > 0$, there exists N depending on these six quantities, such that the following is satisfied for all $n \geqslant N$: for any weighted simplicial complexes $\mathcal{X}, \mathcal{Y} \in \boldsymbol{C}^w$ such that*

(i) $|\mathcal{X}^{(d-1)} \triangle \mathcal{Y}^{(d-1)}| \leqslant C_1$, *where* $\mathcal{X}^{(d-1)} \triangle \mathcal{Y}^{(d-1)} = (\mathcal{X}^{(d-1)} \backslash \mathcal{Y}^{(d-1)}) \cup (\mathcal{Y}^{(d-1)} \backslash \mathcal{X}^{(d-1)})$;

(ii) *any vertex contained in a face in $\mathcal{X}^{(d-1)} \cap \mathcal{Y}^{(d-1)}$ has the same weight in both complexes;*

(iii) *each face in $\mathcal{X}^{(d-1)} \triangle \mathcal{Y}^{(d-1)}$ has at most fitness C_2 in the complex it belongs to;*

(iv) *$F(\mathcal{X}) \geqslant \varepsilon_1 u$ for some $\eta_1 n \leqslant u \leqslant n$ (where we recall that $F(\mathcal{X})$ is the sum of fitnesses of faces in $\mathcal{X}$),*

we have, for any $u < m \leqslant n$, that

$$\mathbb{P}_{\mathcal{X}}\left(\bigcap_{j=u+1}^{m} \mathcal{G}_{\varepsilon_2}(j-u;j)\right) \geqslant \mathbb{P}_{\mathcal{Y}}\left(\bigcap_{j=u+1}^{m} \mathcal{G}_{\varepsilon_2/2}(j-u;j)\right) - \varepsilon_3.$$

Intuitively, Lemma 4.4.15 states that, for the process initiated by $\mathcal{K}_{i\backslash i}$, the partition function remains concentrated with high probability at each of the $n-i$ steps after the arrival of vertex i. Lemma 4.4.16 states that any sufficiently large simplicial complexes $\mathcal{X}$ and $\mathcal{Y}$, in the sense of being linear in n, which differ by at most a constant number of faces, have partition functions that evolve in a similar manner. This is due to the fact that the contribution of the descendants of faces in $\mathcal{X} \triangle \mathcal{Y}$ may be bounded by the sum of geometrically distributed random variables with small success parameter, and is thus negligible.

For brevity, for all $\ell \in \{0, \ldots, k\}$ and $\varepsilon > 0$, recalling the definition of $p(j)$ in (4.46), we define

$$G_{\ell}(\varepsilon) = \bigcap_{j=i_{\ell}+1}^{n-(k-\ell)} \mathcal{G}_{\varepsilon_{p(j)}}(j-i_{\ell}; j+p(j)-\ell) \quad \text{and} \quad \alpha_{\ell}(\mathcal{K}, \varepsilon) = \mathbb{P}_{\mathcal{K}}(G_{\ell}(\varepsilon)), \quad \mathcal{K} \in \mathbf{C}^w. \quad (4.52)$$

Thus, in $\alpha_\ell(\mathcal{K}_{i_\ell\setminus i}, \varepsilon)$ the term $\mathcal{G}_{\varepsilon_{p(j)}}(j - i_\ell; j + p(j) - \ell)$ represents concentration of $\mathcal{Z}_{j-i_\ell}$ (initiated with $\mathcal{K}_{i_\ell\setminus i}$) around $\lambda(j + p(j) - \ell)$. When $p(j)$ increases, the values of $\varepsilon_{p(j)}$ and $j + p(j) - \ell$ change to account for the additional 'step' that has occurred in the underlying process without a step occurring in the process initiated with $\mathcal{K}_{i_\ell\setminus i}$. Lemma 4.4.16 has the following corollary which justifies this notation, showing that the migration of the additional face into $\mathcal{K}_{i_\ell\setminus i}$ at the step i_ℓ is insignificant.

Corollary 4.4.17. *For any $0 < \eta, \delta, \varepsilon' < 1$, there exists $N = N(\delta, \varepsilon', \eta)$ such that the following holds for all $n \geqslant N$: for all $0 < \varepsilon < 1/(2k+2)$, $\ell \in \{1, \dots, k\}$ and $\eta n < i < i_1 < \dots < i_k \leqslant n$, on the event $\mathcal{G}^{(i)}_{\varepsilon_\ell}(i_\ell)$, with α_ℓ as defined in* (4.52)*, we have*

$$\alpha_\ell(\mathcal{K}_{i_\ell\setminus i}, \varepsilon') \geqslant \alpha_\ell(\mathcal{K}_{(i_\ell-1)\setminus i}, \varepsilon'/4(k+1)) - \delta. \tag{4.53}$$

Proof. For sufficiently large n, depending on ε' and η, we clearly have that, for all $\mathcal{K} \in \mathbf{C}^w$

$$\alpha_\ell(\mathcal{K}, \varepsilon') \geqslant \mathbb{P}_{\mathcal{K}}\left(\bigcap_{j=i_\ell+1}^{n-(k-\ell)} \mathcal{G}_{3\varepsilon'_\ell/4}(j - i_\ell; j)\right)$$

and

$$\mathbb{P}_{\mathcal{K}}\left(\bigcap_{j=i_\ell+1}^{n-(k-\ell)} \mathcal{G}_{3\varepsilon'_\ell/8}(j - i_\ell; j)\right) \geqslant \alpha_\ell(\mathcal{K}, \varepsilon'/4(k+1)). \tag{4.54}$$

Note that, on $\mathcal{G}^{(i)}_{\varepsilon_\ell}(i_\ell)$, we have $\mathcal{Z}_{i_\ell\setminus i} \geqslant \lambda i_\ell/2$. Hence, Lemma 4.4.16 applied with $\varepsilon_1 = \lambda/2, \varepsilon_2 = 3\varepsilon'_\ell/4, \varepsilon_3 = \delta, u = i_\ell, \eta_1 = \eta, \mathcal{Y} = \mathcal{K}_{(i_\ell-1)\setminus i}, \mathcal{X} = \mathcal{K}_{i_\ell\setminus i}, C_1 = d+1, C_2 = f_{\max}$ shows that, on the event $\mathcal{G}^{(i)}_{\varepsilon_\ell}(i_\ell)$,

$$\mathbb{P}_{\mathcal{K}_{i_\ell\setminus i}}\left(\bigcap_{j=i_\ell+1}^{n-(k-\ell)} \mathcal{G}_{3\varepsilon'_\ell/4}(j - i_\ell; j)\right) \geqslant \mathbb{P}_{\mathcal{K}_{(i_\ell-1)\setminus i}}\left(\bigcap_{j=i_\ell+1}^{n-(k-\ell)} \mathcal{G}_{3\varepsilon'_\ell/8}(j - i_\ell; j)\right) - \delta \tag{4.55}$$

for n sufficiently large, depending on $\delta, \varepsilon', \eta$. Then the equations (4.54) and (4.55) together imply (4.53). □

Once we have Corollary 4.4.17, the arguments to prove the lower bound are similar to the upper bound, however, the details are more technical. The following lemma is the analogue of Lemma 4.4.7.

Lemma 4.4.18. *For any* $0, \delta, \eta < 1$ *and* $0 < \varepsilon < 1/(2k+2)$ *there exists* $N = N(\delta, \varepsilon, \eta)$*, such that, for all* $n \geqslant N$ *and* $\eta n < i < i_1 < \ldots < i_k \leqslant n$*, with* $\mathfrak{h}_j^{\varepsilon}$ *as defined in* (4.49)*, we have*

$$\mathbb{P}\left(\bigcap_{j=i_k+1}^{n} \bar{\mathcal{D}}_j(\varepsilon) \,\middle|\, \mathscr{G}_{i_k}\right) \bar{D}_{i_k}(\varepsilon) \geqslant (\alpha_k(\mathcal{K}_{(i_k-1)\setminus i}, \varepsilon/(4(k+1))) - \delta)\mathfrak{h}_k^{\varepsilon}(\mathrm{st}_i(\mathcal{K}_{i_k}))\bar{D}_{i_k}(\varepsilon) \quad (4.56)$$

and, for all $\ell \in \{1, \ldots, k-1\}$*,*

$$\mathbb{E}\left[\prod_{j=i_\ell+1}^{i_{\ell+1}-1} \bar{D}_j(\varepsilon)\ \alpha_{\ell+1}(\mathcal{K}_{(i_{\ell+1}-1)\setminus i}, \varepsilon) \,\middle|\, \mathscr{G}_{i_\ell}\right] \bar{D}_{i_\ell}(\varepsilon)$$

$$\geqslant (\alpha_\ell(\mathcal{K}_{(i_\ell-1)\setminus i}, (k+1)) - \delta)\mathfrak{h}_\ell^{\varepsilon}(\mathrm{st}_i(\mathcal{K}_{i_\ell}))\bar{D}_{i_\ell}(\varepsilon), \ \textit{while,}$$

$$\mathbb{E}\left[\prod_{j=i+1}^{i_1-1} \bar{D}_j(\varepsilon)\ \alpha_1(\mathcal{K}_{(i_1-1)\setminus i}, \varepsilon) \,\middle|\, \mathscr{G}_i\right] \bar{D}_i(\varepsilon) \geqslant \alpha_0(\mathcal{K}_{i\setminus i}, \varepsilon)\mathfrak{h}_0^{\varepsilon}(\mathrm{st}_i(\mathcal{K}_i))\bar{D}_i(\varepsilon).$$

Proof. We write $\bar{D}_j$ for $\bar{D}_j(\varepsilon)$ throughout the proof. If $i_k \neq n$, we have

$$\mathbb{E}\left[\prod_{j=i_k+1}^{n} \bar{D}_j \,\middle|\, \mathscr{G}_{i_k}\right] = \mathbb{E}\left[\mathbb{E}\left[\bar{D}_n \,\middle|\, \mathscr{G}_{n-1}\right] \prod_{j=i_k+1}^{n-1} \bar{D}_j \,\middle|\, \mathscr{G}_{i_k}\right]$$

$$= \mathbb{E}\left[\left(1 - \frac{F(\mathrm{st}_i(\mathcal{K}_{n-1}))}{\mathscr{Z}_{n-1}}\right) \mathbb{P}_{\mathcal{K}_{(n-1)\setminus i}}\left(\mathcal{G}_{\varepsilon_k}(1;n)\right) \prod_{j=i_k+1}^{n-1} \bar{D}_j \,\middle|\, \mathscr{G}_{i_k}\right], \quad (4.57)$$

because, by definition (see (4.45)), $\mathcal{G}_{\varepsilon_k}(1;n) = \{|\mathscr{Z}_1 - \lambda n| < \varepsilon_k \lambda n\}$. First note that, on the event $\bigcap_{j=i_k+1}^{n-1} \bar{\mathcal{D}}_j$, we have, for any $j = i_k+1, \ldots, n-1$, $F(\mathrm{st}_i(\mathcal{K}_j)) = F(\mathrm{st}_i(\mathcal{K}_{i_k}))$. On the event $\bar{\mathcal{D}}_j$ we have

$$1 - \frac{F(\mathrm{st}_i(\mathcal{K}_{n-1}))}{\mathscr{Z}_j} \geqslant 1 - \frac{F(\mathrm{st}_i(\mathcal{K}_{i_k}))}{\lambda_{-\varepsilon_k} j}. \quad (4.58)$$

Furthermore, by the tower property, we may substitute

$$\mathbb{E}\left[\mathbb{P}_{\mathcal{K}_{(n-1)\setminus i}}\left(\mathcal{G}_{\varepsilon_k}(1;n)\right) \bar{D}_{n-1} \,\middle|\, \mathscr{G}_{n-2}\right] \quad \text{for} \quad \mathbb{P}_{\mathcal{K}_{(n-1)\setminus i}}\left(\mathcal{G}_{\varepsilon_k}(1;n)\right) \bar{D}_{n-1}$$

inside the conditional expectation, and together with (4.57) and (4.58), this gives

$$\mathbb{E}\left[\prod_{j=i_k+1}^{n} \bar{D}_j \,\middle|\, \mathscr{G}_{i_k}\right] \geqslant \left(1 - \frac{F(\mathrm{st}_i(\mathcal{K}_{i_k}))}{\lambda_{-\varepsilon_k}(n-1)}\right) \mathbb{E}\left[\mathbb{E}\left[\mathbb{P}_{\mathcal{K}_{(n-1)\setminus i}}\left(\mathcal{G}_{\varepsilon_k}(1;n)\right) \bar{D}_{n-1} \,\middle|\, \mathscr{G}_{n-2}\right] \prod_{j=i_k+1}^{n-2} \bar{D}_j \,\middle|\, \mathscr{G}_{i_k}\right]. \quad (4.59)$$

Then, if $i_k \neq n-1$ we also have

$$\mathbb{E}\left[\mathbb{P}_{\mathcal{K}_{(n-1)\setminus i}}\left(\mathcal{G}_{\varepsilon_k}(1;n)\right)\bar{D}_{n-1}\,\middle|\,\mathscr{G}_{n-2}\right] = \left(1-\frac{F(\mathrm{st}_i(\mathcal{K}_{n-2}))}{\mathcal{Z}_{n-2}}\right)\mathbb{P}_{\mathcal{K}_{(n-2)\setminus i}}\left(\mathcal{G}_{\varepsilon_k}(1;n-1)\cap\mathcal{G}_{\varepsilon_k}(2;n)\right). \tag{4.60}$$

Thus, using (4.59) and (4.60) in the first inequality, and (4.58) in the second,

$$\begin{aligned}
&\mathbb{E}\left[\prod_{j=i_k+1}^{n}\bar{D}_j\,\middle|\,\mathscr{G}_{i_k}\right] \\
&\geqslant \left(1-\frac{F(\mathrm{st}_i(\mathcal{K}_{i_k}))}{\lambda_{-\varepsilon_k}(n-1)}\right)\mathbb{E}\left[\left(1-\frac{F(\mathrm{st}_i(\mathcal{K}_{n-2}))}{\mathcal{Z}_{n-2}}\right)\mathbb{P}_{\mathcal{K}_{(n-2)\setminus i}}\left(\mathcal{G}_{\varepsilon_k}(1;n-1)\cap\mathcal{G}_{\varepsilon_k}(2;n)\right)\prod_{j=i_k+1}^{n-2}\bar{D}_j\,\middle|\,\mathscr{G}_{i_k}\right] \\
&\geqslant \left(1-\frac{F(\mathrm{st}_i(\mathcal{K}_{i_k}))}{\lambda_{-\varepsilon_k}(n-1)}\right)\left(1-\frac{F(\mathrm{st}_i(\mathcal{K}_{i_k}))}{\lambda_{-\varepsilon_k}(n-2)}\right)\mathbb{E}\left[\mathbb{P}_{\mathcal{K}_{(n-2)\setminus i}}\left(\mathcal{G}_{\varepsilon_k}(1;n-1)\cap\mathcal{G}_{\varepsilon_k}(2;n)\right)\prod_{j=i_k+1}^{n-2}\bar{D}_j\,\middle|\,\mathscr{G}_{i_k}\right].
\end{aligned}$$

Iterating this process gives us

$$\mathbb{P}\left(\bigcap_{j=i_k+1}^{n}\bar{\mathcal{D}}_j(\varepsilon)\,\middle|\,\mathscr{G}_{i_k}\right)\bar{D}_{i_k}(\varepsilon)\geqslant\alpha_k(\mathcal{K}_{i_k\setminus i},\varepsilon)\mathfrak{h}_k^{\varepsilon}(\mathrm{st}_i(\mathcal{K}_{i_k}))\bar{D}_{i_k}.$$

Applying (4.53) from Corollary 4.4.17 concludes the proof of (4.56) as $\bar{\mathcal{D}}_{i_k}\subseteq\mathcal{G}^{(i)}_{\varepsilon_k}(i_k)$.

We use the same ideas to prove the general case, for $\ell\in\{0,\dots,k-1\}$. Here, we want to provide a lower bound to $\mathbb{E}\left[\alpha_{\ell+1}(\mathcal{K}_{(i_{\ell+1}-1)\setminus i},\varepsilon)\ \prod_{j=i_\ell+1}^{i_{\ell+1}-1}\bar{D}_j\,\middle|\,\mathscr{G}_{i_\ell}\right]$. First, for any $j=i_\ell+1,\dots,i_{\ell+1}-1$, we have $F(\mathrm{st}_i(\mathcal{K}_j))=F(\mathrm{st}_i(\mathcal{K}_{i_\ell})$. Thus, on the event $\bar{\mathcal{D}}_j$, we have

$$1-\frac{F(\mathrm{st}_i(\mathcal{K}_j))}{\mathcal{Z}_j}\geqslant 1-\frac{F(\mathrm{st}_i(\mathcal{K}_{i_\ell}))}{\lambda_{-\varepsilon_\ell}j}. \tag{4.61}$$

Second, using the tower property, we substitute

$$\mathbb{E}\left[\alpha_{\ell+1}(\mathcal{K}_{(i_{\ell+1}-1)\setminus i},\varepsilon)\ \bar{D}_{i_{\ell+1}-1}\,\middle|\,\mathscr{G}_{i_{\ell+1}-2}\right]\ \text{ for }\ \alpha_{\ell+1}(\mathcal{K}_{(i_{\ell+1}-1)\setminus i},\varepsilon)\bar{D}_{i_{\ell+1}-1} \tag{4.62}$$

inside the conditional expectation. Third, if $i_{\ell+1}-1\neq i_\ell$,

$$\begin{aligned}
&\mathbb{E}\left[\alpha_{\ell+1}(\mathcal{K}_{(i_{\ell+1}-1)\setminus i},\varepsilon)\ \bar{D}_{i_{\ell+1}-1}\,\middle|\,\mathscr{G}_{i_{\ell+1}-2}\right]=\left(1-\frac{F(\mathrm{st}_i(\mathcal{K}_{i_{\ell+1}-2}))}{\mathcal{Z}_{i_{\ell+1}-2}}\right)\times \\
&\mathbb{P}_{\mathcal{K}_{(i_{\ell+1}-2)\setminus i}}\left(\mathcal{G}_{\varepsilon_\ell}(1;i_{\ell+1}-1)\cap\bigcap_{j=i_{\ell+1}+1}^{n-(k-\ell-1)}\mathcal{G}_{\varepsilon_{p(j)}}(j-i_{\ell+1}+1;j+p(j)-\ell-1)\right).
\end{aligned}$$

So we write:

$$\begin{aligned}
&\mathbb{E}\left[\alpha_{\ell+1}(\mathcal{K}_{(i_{\ell+1}-1)\setminus i},\varepsilon)\prod_{j=i_\ell+1}^{i_{\ell+1}-1}\bar{D}_j\,\middle|\,\mathscr{G}_{i_\ell}\right]\\
&\overset{(4.62)}{=}\mathbb{E}\left[\mathbb{E}\left[\alpha_{\ell+1}(\mathcal{K}_{(i_{\ell+1}-1)\setminus i},\varepsilon)\;\bar{D}_{i_{\ell+1}-1}\,\middle|\,\mathscr{G}_{i_{\ell+1}-2}\right]\prod_{j=i_\ell+1}^{i_{\ell+1}-2}\bar{D}_j\,\middle|\,\mathscr{G}_{i_\ell}\right]\\
&\overset{(4.63)}{=}\mathbb{E}\Bigg[\left(1-\frac{F(\mathrm{st}_i(\mathcal{K}_{i_{\ell+1}-2}))}{\mathcal{Z}_{i_{\ell+1}-2}}\right)\prod_{j=i_\ell+1}^{i_{\ell+1}-2}\bar{D}_j\times\\
&\qquad\mathbb{P}_{\mathcal{K}_{(i_{\ell+1}-2)\setminus i}}\left(\mathcal{G}_{\varepsilon_\ell}(1;i_{\ell+1}-1)\cap\bigcap_{j=i_{\ell+1}+1}^{n-(k-\ell-1)}\mathcal{G}_{\varepsilon_{p(j)}}(j-i_{\ell+1}+1;j+p(j)-\ell-1)\right)\,\Bigg|\,\mathscr{G}_{i_\ell}\Bigg].
\end{aligned}$$

Now, the lower bound of (4.61) yields:

$$\begin{aligned}
&\mathbb{E}\left[\alpha_{\ell+1}(\mathcal{K}_{(i_{\ell+1}-1)\setminus i},\varepsilon)\prod_{j=i_\ell+1}^{i_{\ell+1}-1}\bar{D}_j\,\middle|\,\mathscr{G}_{i_\ell}\right]\\
&\geqslant\left(1-\frac{F(\mathrm{st}_i(\mathcal{K}_{i_\ell}))}{\lambda_{-\varepsilon_\ell}(i_{\ell+1}-2)}\right)\mathbb{E}\left[\mathbb{P}_{\mathcal{K}_{(i_{\ell+1}-2)\setminus i}}\left(\bigcap_{j=i_{\ell+1}-1}^{n-(k-\ell)}\mathcal{G}_{\varepsilon_{p(j)}}(j-i_{\ell+1}+2;j+p(j)-\ell)\right)\prod_{j=i_\ell+1}^{i_{\ell+1}-2}\bar{D}_j\,\middle|\,\mathscr{G}_{i_\ell}\right].
\end{aligned}$$

By the tower property again, we substitute

$$\mathbb{E}\left[\mathbb{P}_{\mathcal{K}_{(i_{\ell+1}-2)\setminus i}}\left(\bigcap_{j=i_{\ell+1}-1}^{n-(k-\ell)}\mathcal{G}_{\varepsilon_{p(j)}}(j-i_{\ell+1}+2;j+p(j)-\ell)\right)\bar{D}_{i_{\ell+1}-2}\,|\,\mathscr{G}_{i_{\ell+1}-3}\right]$$

$$\text{for}\quad\mathbb{P}_{\mathcal{K}_{(i_{\ell+1}-2)\setminus i}}\left(\bigcap_{j=i_{\ell+1}-1}^{n-(k-\ell)}\mathcal{G}_{\varepsilon_{p(j)}}(j-i_{\ell+1}+2;j+p(j)-\ell)\right)\bar{D}_{i_{\ell+1}-2}.$$

Also, if $i_{\ell+1}-2\neq i_\ell$,

$$\begin{aligned}
&\mathbb{E}\left[\mathbb{P}_{\mathcal{K}_{(i_{\ell+1}-2)\setminus i}}\left(\bigcap_{j=i_{\ell+1}-1}^{n-(k-\ell)}\mathcal{G}_{\varepsilon_{p(j)}}(j-i_{\ell+1}+2;j+p(j)-\ell)\right)\bar{D}_{i_{\ell+1}-2}\,|\,\mathscr{G}_{i_{\ell+1}-3}\right]=\\
&\left(1-\frac{F(\mathrm{st}_i(\mathcal{K}_{i_{\ell+1}-3}))}{\mathcal{Z}_{i_{\ell+1}-3}}\right)\mathbb{P}_{\mathcal{K}_{(i_{\ell+1}-3)\setminus i}}\left(\bigcap_{j=i_{\ell+1}-2}^{n-(k-\ell)}\mathcal{G}_{\varepsilon_{p(j)}}(j-i_{\ell+1}+3;j+p(j)-\ell)\right).
\end{aligned}$$

Bounding the first factor as in (4.61), and combining (4.64) and (4.65) give

$$\begin{aligned}
&\mathbb{E}\left[\alpha_{\ell+1}(\mathcal{K}_{(i_{\ell+1}-1)\setminus i},\varepsilon)\prod_{j=i_\ell+1}^{i_{\ell+1}-1}\bar{D}_j\,\middle|\,\mathscr{G}_{i_\ell}\right]\\
&\geqslant\left(1-\frac{F(\mathrm{st}_i(\mathcal{K}_{i_\ell}))}{\lambda_{-\varepsilon_\ell}(i_{\ell+1}-2)}\right)\left(1-\frac{F(\mathrm{st}_i(\mathcal{K}_{i_\ell}))}{\lambda_{-\varepsilon_\ell}(i_{\ell+1}-3)}\right)\times\\
&\qquad\mathbb{E}\left[\mathbb{P}_{\mathcal{K}_{(i_{\ell+1}-3)\setminus i}}\left(\bigcap_{j=i_{\ell+1}-2}^{n-(k-\ell)}\mathcal{G}_{\varepsilon_{p(j)}}(j-i_{\ell+1}+3;j+p(j)-\ell)\right)\prod_{j=i_\ell+1}^{i_{\ell+1}-3}\bar{D}_j\,\middle|\,\mathscr{G}_{i_\ell}\right].
\end{aligned}$$

Iterating the argument shows that the right hand side multiplied by $\bar{D}_{i_\ell}$ is bounded from below by $\alpha_\ell(\mathcal{K}_{i_\ell\setminus i},\varepsilon)\mathfrak{h}^\varepsilon_\ell(\mathrm{st}_i(\mathcal{K}_{i_\ell}))\bar{D}_{i_\ell}$. We conclude the proof by applying (4.53) from Corollary 4.4.17. □

Lemma 4.4.19. *For any $\delta > 0, 0 < \eta < 1$ and $0 < \varepsilon < 1/(2k+2)$, there exists $N = N(\delta,\varepsilon,\eta)$ such that, for all $n \geqslant N$, $\ell \in \{1,\ldots,k\}$ and $\eta n < i < i_1 < \ldots < i_k \leqslant n$, with $\mathfrak{f}^\varepsilon_j$ as defined in* (4.50) *we have*

$$\mathbb{E}\left[\alpha_{\ell+1}(\mathcal{K}_{(i_{\ell+1}-1)\setminus i},\varepsilon)\ \mathbb{E}^*_{\mathrm{st}_i(\mathcal{K}_{i_{\ell+1}})}\left[\prod_{j=\ell+1}^{k}\mathfrak{f}^\varepsilon_j(S_{j-\ell-1})\right]\prod_{j=i_\ell+1}^{\min(i_{\ell+1},n)}\bar{D}_j(\varepsilon)\,\middle|\,\mathscr{G}_{i_\ell}\right]\bar{D}_{i_\ell}(\varepsilon)$$
$$\geqslant (\alpha_\ell(\mathcal{K}_{(i_\ell-1)\setminus i},\varepsilon/(4(k+1)))-\delta)\mathbb{E}^*_{\mathrm{st}_i(\mathcal{K}_{i_\ell})}\left[\prod_{j=\ell}^{k}\mathfrak{f}^\varepsilon_j(S_{j-\ell})\right]\bar{D}_{i_\ell}(\varepsilon), \tag{4.66}$$

where we use the convention $\alpha_{k+1}(\cdot) = 1$, while

$$\mathbb{E}\left[\alpha_1(\mathcal{K}_{(i_1-1)\setminus i},\varepsilon)\ \mathbb{E}^*_{\mathrm{st}_i(\mathcal{K}_{i_1})}\left[\prod_{j=1}^{k}\mathfrak{f}^\varepsilon_j(S_{j-\ell-1})\right]\prod_{j=i+1}^{i_1}\bar{D}_j(\varepsilon)\,\middle|\,\mathscr{G}_i\right]\bar{D}_i(\varepsilon)$$
$$\geqslant \alpha_0(\mathcal{K}_{i\setminus i},\varepsilon)\mathbb{E}^*_{\mathrm{st}_i(\mathcal{K}_i)}\left[\prod_{j=0}^{k}\mathfrak{f}^\varepsilon_j(S_j)\right]\bar{D}_i(\varepsilon).$$

Proof. The inequality (4.66) coincides with (4.56) from Lemma 4.4.18 when $\ell = k$. Let $0 \leqslant \ell \leqslant k-1$. Note that, for all $1 \leqslant i \leqslant n$, we have $|\mathcal{Z}_{n\setminus i} - \mathcal{Z}_{(n-1)\setminus i}| \leqslant (d+1)f_{\max}$. Thus, for all n sufficiently large, depending on ε and η, we have

$$\mathcal{D}_{i_{\ell+1}} \cap \mathcal{G}^{(i)}_{\varepsilon_\ell}(i_{\ell+1}-1) \subseteq \mathcal{G}^{(i)}_{\varepsilon_{\ell+1}}(i_{\ell+1}). \tag{4.67}$$

Using this observation in the second step, we deduce

$$\mathbb{E}\left[\alpha_{\ell+1}(\mathcal{K}_{(i_{\ell+1}-1)\setminus i},\varepsilon)\ \mathbb{E}^*_{\mathrm{st}_i(\mathcal{K}_{i_{\ell+1}})}\left[\prod_{j=\ell+1}^{k}\mathfrak{f}^\varepsilon_j(S_{j-\ell-1})\right]\prod_{j=i_\ell+1}^{i_{\ell+1}}\bar{D}_j\,\middle|\,\mathscr{G}_{i_\ell}\right]\bar{D}_{i_\ell}$$
$$= \mathbb{E}\left[\mathbb{E}\left[\bar{D}_{i_{\ell+1}}\ \mathbb{E}^*_{\mathrm{st}_i(\mathcal{K}_{i_{\ell+1}})}\left[\prod_{j=\ell+1}^{k}\mathfrak{f}^\varepsilon_j(S_{j-\ell-1})\right]\,\middle|\,\mathscr{G}_{i_{\ell+1}-1}\right]\ \alpha_{\ell+1}(\mathcal{K}_{(i_{\ell+1}-1)\setminus i},\varepsilon)\prod_{j=i_\ell+1}^{i_{\ell+1}-1}\bar{D}_j\,\middle|\,\mathscr{G}_{i_\ell}\right]\bar{D}_{i_\ell}$$
$$\overset{(4.67)}{\geqslant} \mathbb{E}\left[\mathbb{E}\left[D_{i_{\ell+1}}\mathbb{E}^*_{\mathrm{st}_i(\mathcal{K}_{i_{\ell+1}})}\left[\prod_{j=\ell+1}^{k}\mathfrak{f}^\varepsilon_j(S_{j-\ell-1})\right]\,\middle|\,\mathscr{G}_{i_{\ell+1}-1}\right]\ \alpha_{\ell+1}(\mathcal{K}_{(i_{\ell+1}-1)\setminus i},\varepsilon)\prod_{j=i_\ell+1}^{i_{\ell+1}-1}\bar{D}_j\,\middle|\,\mathscr{G}_{i_\ell}\right]\bar{D}_{i_\ell}.$$

Recall that (analogous to in the Proof of Proposition 4.4.8), conditionally on $\mathscr{G}_{i_{\ell+1}-1}$, on the event $\mathcal{D}_{i_\ell+1}$, the random variable $\mathrm{st}_i(\mathcal{K}_{i_{\ell+1}})$ is distributed as S_1 for the star Markov process starting at $\mathrm{st}_i(\mathcal{K}_{i_{\ell+1}-1})$. This yields:

$$\mathbb{E}\left[D_{i_{\ell+1}}\mathbb{E}^*_{\mathrm{st}_i(\mathcal{K}_{i_{\ell+1}})}\left[\prod_{j=\ell+1}^{k}\mathfrak{f}_j^\varepsilon(S_{j-\ell-1})\right]\middle|\mathscr{G}_{i_{\ell+1}-1}\right] = \mathbb{P}\left(\mathcal{D}_{i_{\ell+1}}\,|\,\mathscr{G}_{i_{\ell+1}-1}\right)\cdot\mathbb{E}^*_{\mathrm{st}_i(\mathcal{K}_{i_{\ell+1}-1})}\left[\prod_{j=\ell+1}^{k}\mathfrak{f}_j^\varepsilon(S_{j-\ell})\right]$$
$$= \frac{F(\mathrm{st}_i(\mathcal{K}_{i_{\ell+1}-1}))}{\mathcal{Z}_{i_{\ell+1}-1}}\cdot\mathbb{E}^*_{\mathrm{st}_i(\mathcal{K}_{i_{\ell+1}-1})}\left[\prod_{j=\ell+1}^{k}\mathfrak{f}_j^\varepsilon(S_{j-\ell})\right].$$

We deduce that

$$\mathbb{E}\left[\alpha_{\ell+1}(\mathcal{K}_{(i_{\ell+1}-1)\setminus i},\varepsilon)\ \mathbb{E}^*_{\mathrm{st}_i(\mathcal{K}_{i_{\ell+1}})}\left[\prod_{j=\ell+1}^{k}\mathfrak{f}_j^\varepsilon(S_{j-\ell-1})\right]\prod_{j=i_\ell+1}^{i_{\ell+1}}\bar{D}_j\middle|\mathscr{G}_{i_\ell}\right]\bar{D}_{i_\ell}$$
$$\geqslant \mathbb{E}\left[\frac{F(\mathrm{st}_i(\mathcal{K}_{i_\ell}))}{\mathcal{Z}_{i_{\ell+1}-1}}\mathbb{E}^*_{\mathrm{st}_i(\mathcal{K}_{i_\ell})}\left[\prod_{j=\ell+1}^{k}\mathfrak{f}_j^\varepsilon(S_{j-\ell})\right]\alpha_{\ell+1}(\mathcal{K}_{(i_{\ell+1}-1)\setminus i},\varepsilon)\prod_{j=i_\ell+1}^{i_{\ell+1}-1}\bar{D}_j\middle|\mathscr{G}_{i_\ell}\right]\bar{D}_{i_\ell}.$$

But on the event associated with $\bar{D}_{i_{\ell+1}}$ we have

$$\frac{F(\mathrm{st}_i(\mathcal{K}_{i_\ell}))}{\mathcal{Z}_{i_{\ell+1}-1}} \geqslant \frac{F(\mathrm{st}_i(\mathcal{K}_{i_\ell}))}{F(\mathrm{st}_i(\mathcal{K}_{i_\ell}))+\lambda_{+\varepsilon_\ell}(i_{\ell+1}-1)}.$$

So the previous inequality continues as follows:

$$\frac{F(\mathrm{st}_i(\mathcal{K}_{i_\ell}))}{F(\mathrm{st}_i(\mathcal{K}_{i_\ell}))+\lambda_{+\varepsilon_\ell}(i_{\ell+1}-1)}\times$$
$$\mathbb{E}^*_{\mathrm{st}_i(\mathcal{K}_{i_\ell})}\left[\prod_{j=\ell+1}^{k}\mathfrak{f}_j^\varepsilon(S_{j-\ell})\right]\cdot\mathbb{E}\left[\alpha_{\ell+1}(\mathcal{K}_{(i_{\ell+1}-1)\setminus i},\varepsilon)\prod_{j=i_\ell+1}^{i_{\ell+1}-1}\bar{D}_j\middle|\mathscr{G}_{i_\ell}\right]\bar{D}_{i_\ell}.$$

We bound the last term from below using Lemma 4.4.18:

$$\mathbb{E}\left[\alpha_{\ell+1}(\mathcal{K}_{(i_{\ell+1}-1)\setminus i},\varepsilon)\prod_{j=i_\ell+1}^{i_{\ell+1}-1}\bar{D}_j\middle|\mathscr{G}_{i_\ell}\right]\bar{D}_{i_\ell}\geqslant(\alpha_\ell(\mathcal{K}_{(i_\ell-1)\setminus i},\varepsilon/(4(k+1)))-\delta)\mathfrak{h}_\ell^\varepsilon(\mathrm{st}_i(\mathcal{K}_{i_\ell}))\bar{D}_{i_\ell}.$$

By (4.50), we have

$$\frac{F(\mathrm{st}_i(\mathcal{K}_{i_\ell}))}{F(\mathrm{st}_i(\mathcal{K}_{i_\ell}))+\lambda_{+\varepsilon_\ell}(i_{\ell+1}+1)}\mathfrak{h}_\ell^\varepsilon(\mathrm{st}_i(\mathcal{K}_{i_\ell}))\mathbb{E}^*_{\mathrm{st}_i(\mathcal{K}_{i_\ell})}\left[\prod_{j=\ell+1}^{k}\mathfrak{f}_j^\varepsilon(S_{j-\ell})\right]=\mathbb{E}^*_{\mathrm{st}_i(\mathcal{K}_{i_\ell})}\left[\prod_{j=\ell}^{k}\mathfrak{f}_j^\varepsilon(S_{j-\ell})\right],$$

so the claim follows. □

The lemma allows us to bound $\mathbb{P}\big(\bigcap_{j=i+1}^{n} \bar{\mathcal{D}}_j\big)$ from below by a term similar to (4.51) using a backward induction argument which is of the same nature as the proof of Proposition 4.4.8. This result needs to be prepared with the following definition. For $0 < \varepsilon < 1/(2k+2), 0 < \eta < 1$ and $C > 0$, set

$$\gamma(\varepsilon, \eta, C) = \gamma_k(\varepsilon, \eta, C)^{k(k+1)/2}, \quad \gamma_\ell(\varepsilon, \eta, C) = (1-\varepsilon_\ell)\, \eta^{2C\varepsilon_\ell/\lambda}, \quad \ell = 1, \ldots, k. \tag{4.68}$$

Note that these terms decrease as ε or C increase.

Lemma 4.4.20. *For $0 < \varepsilon < 1/(2k+2), 0 < \eta < 1$ and $C > 0$ there exists $N = N(\varepsilon, \eta, C)$ such that, for all $n \geqslant N$, $\eta n < i < i_1 < \ldots < i_k \leqslant n$ and $0 < \varepsilon' \leqslant \varepsilon$*

$$\mathfrak{f}_\ell^{\varepsilon}(S) \geqslant \gamma_\ell(\varepsilon, \eta, C) \mathfrak{f}_\ell^{\varepsilon'}(S) \quad \textit{for all } S \in \mathcal{C}' \textit{ with } F(S) \leqslant C.$$

Proof. Recalling that $\lambda_{+\varepsilon_\ell} = \lambda(1+\varepsilon_\ell)$ we deduce that

$$\frac{F(S)}{F(S) + \lambda_{+\varepsilon_\ell}(i_{\ell+1}-1)} > \frac{F(S)}{(1+\varepsilon_\ell)(F(S) + \lambda(i_{\ell+1}-1))} > (1-\varepsilon_\ell)\, \frac{F(S)}{F(S) + \lambda(i_{\ell+1}-1)}.$$

This statement requires no bounds on $F(S)$ or i_ℓ. Hence, it is sufficient to prove that $\mathfrak{h}_\ell^{\varepsilon}(S) \geqslant \eta^{2C\varepsilon_\ell/\lambda} \mathfrak{h}_\ell^{\varepsilon'}(S)$ for sufficiently large n. By Lemma 4.4.9, we have

$$\mathfrak{h}_\ell^{\varepsilon}(S) = \left(\frac{i_\ell}{i_{\ell+1}}\right)^{F(S)/\lambda_{-\varepsilon_\ell}} \left(1 + O\left(\frac{1}{n}\right)\right),$$

where the O-term can be chosen uniformly in $\varepsilon, i_\ell, i_{\ell+1}$ and S for given η and C. Note that $\mathfrak{h}_\ell^{\varepsilon}(S)$ increases as ε decreases. Therefore, it is enough to prove that for each $\ell \in \{0, \ldots, k+1\}$

$$\left(\frac{i_\ell}{i_{\ell+1}}\right)^{F(S)/\lambda_{-\varepsilon_\ell}} > \eta^{2C\varepsilon_\ell/\lambda} \left(\frac{i_\ell}{i_{\ell+1}}\right)^{F(S)/\lambda}$$

for all S with $F(S) \leqslant C$. This follows easily from the bound on F, the fact that $\varepsilon < 1/(2k+2)$ (so that for each ℓ we have $1/(1-\varepsilon_\ell) \leqslant 2$) and each ratio satisfies $\eta \leqslant \frac{i_\ell}{i_{\ell+1}} < 1$. ☐

Proposition 4.4.21. *For $\delta > 0, 0 < \eta < 1$ and $0 < \varepsilon < 1/(2k+2)$, there exists $N = N(\delta, \varepsilon, \eta) > 0$ such that, for all $n \geqslant N$ and $\eta n < i \leqslant i_1 < \ldots < i_k \leqslant n$, with $\gamma_k =$*

$\gamma_k(\varepsilon, \eta, (d+1)(k+1)f_{\max})$ *and* $\gamma = \gamma(\varepsilon, \eta, (d+1)(k+1)f_{\max})$, *we have,*

$$\mathbb{P}\left(\bigcap_{j=i+1}^{n} \bar{\mathcal{D}}_j(\varepsilon)\right) \geqslant \gamma \mathbb{E}\left[\alpha_0\left(\mathcal{K}_{i\setminus i}, \varepsilon/(4(k+1))^{k+1}\right)\mathbb{E}^*_{\mathrm{st}_i(\mathcal{K}_i)}\left[\prod_{j=0}^{k} \mathfrak{f}_j^{\varepsilon}(S_j)\right]\bar{D}_i(\varepsilon/(4(k+1))^{k+1})\right]$$
$$- \delta \sum_{\ell=1}^{k} \mathbb{E}\left[\prod_{j=i+1}^{i_\ell} \bar{D}_j(\varepsilon)\mathbb{E}^*_{\mathrm{st}_i(\mathcal{K}_{i_\ell})}\left[\prod_{j=0}^{\ell} \mathfrak{f}_{k+j-\ell}^{\varepsilon/(4(k+1))^k}(S_j)\right]\bar{D}_i(\varepsilon)\right]. \tag{4.69}$$

Proof. By Lemma 4.4.18, we have

$$\mathbb{P}\left(\bigcap_{j=i+1}^{n} \bar{\mathcal{D}}_j(\varepsilon)\right) = \mathbb{E}\left[\mathbb{P}\left(\bigcap_{j=i_k+1}^{n} \bar{\mathcal{D}}_j(\varepsilon)\,\middle|\,\mathscr{G}_{i_k}\right)\prod_{j=i+1}^{i_k} \bar{D}_j(\varepsilon)\right]$$
$$\overset{(4.56)}{\geqslant} \mathbb{E}\left[\alpha_k(\mathcal{K}_{(i_k-1)\setminus i}, \varepsilon/(4(k+1)))\mathbb{E}^*_{\mathrm{st}_i(\mathcal{K}_{i_k})}[\mathfrak{f}_k^{\varepsilon}(S_0)]\prod_{j=i+1}^{i_k} \bar{D}_j(\varepsilon)\right]$$
$$- \delta \mathbb{E}\left[\mathbb{E}^*_{\mathrm{st}_i(\mathcal{K}_{i_k})}[\mathfrak{f}_k^{\varepsilon}(S_0)]\prod_{j=i+1}^{i_k} \bar{D}_j(\varepsilon)\right].$$

In order to apply Lemma 4.4.19 again in the first term, we may replace $\bar{D}_j(\varepsilon)$ by $\bar{D}_j(\varepsilon/(4(k+1)))$. Moreover, by Lemma 4.4.20 and as $F(S_\ell) \leqslant (d+1)(k+1)f_{\max}$ for $\ell \in \{0, \ldots, k\}$, we may replace $\mathfrak{f}_k^{\varepsilon}(S_0)$ by $\gamma_k \mathfrak{f}_k^{\varepsilon/(4(k+1))}(S_0)$ for sufficiently large n. Hence, applying Lemma 4.4.19 again after this step, we deduce that the first term in the last display is bounded from below by

$$\gamma_k \mathbb{E}\left[\alpha_{k-1}(\mathcal{K}_{(i_{k-1}-1)\setminus i}, \varepsilon/16)\mathbb{E}^*_{\mathrm{st}_i(\mathcal{K}_{i_{k-1}})}\left[\mathfrak{f}_{k-1}^{\varepsilon/(4(k+1))}(S_0)\mathfrak{f}_k^{\varepsilon/(4(k+1))}(S_1)\right]\prod_{j=i+1}^{i_{k-1}} \bar{D}_j(\varepsilon/(4(k+1)))\right]$$
$$- \delta\gamma_k \mathbb{E}\left[\mathbb{E}^*_{\mathrm{st}_i(\mathcal{K}_{i_{k-1}})}\left[\mathfrak{f}_{k-1}^{\varepsilon/(4(k+1))}(S_0)\mathfrak{f}_k^{\varepsilon/(4(k+1))}(S_1)\right]\prod_{j=i+1}^{i_{k-1}} \bar{D}_j(\varepsilon/(4(k+1)))\right].$$

We now iterate these steps until the main term contains α_0. In particular, with the leading term, at the $(\ell+1)$th step we get an expression of the form

$$\mathbb{E}\left[\alpha_{k-\ell}(\mathcal{K}_{(i_{k-\ell}-1)\setminus i}, \varepsilon/(4(k+1))^{\ell+1})\mathbb{E}^*_{\mathrm{st}_i(\mathcal{K}_{i_{k-\ell}})}\left[\prod_{j=0}^{\ell} \mathfrak{f}_{k+j-\ell}^{\varepsilon/(4(k+1))^{\ell}}(S_j)\right]\prod_{j=i+1}^{i_{k-\ell}} \bar{D}_j\left(\varepsilon/(4(k+1))^{\ell}\right)\right]$$

$$\geqslant \left(\prod_{j=0}^{\ell}\gamma_{k-j}\right)\mathbb{E}\left[\alpha_{k-(\ell+1)}(\mathcal{K}_{(i_{k-(\ell+1)}-1)\setminus i}, \varepsilon/(4(k+1))^{\ell+2})\right.$$

$$\left.\times\mathbb{E}^*_{\mathrm{st}_i(\mathcal{K}_{i_{k-(\ell+1)}})}\left[\prod_{j=0}^{\ell+1} \mathfrak{f}_{k+j-(\ell+1)}^{\varepsilon/(4(k+1))^{\ell+1}}(S_j)\right]\prod_{j=i+1}^{i_{k-(\ell+1)}} \bar{D}_j(\varepsilon/(4(k+1))^{\ell+1})\right]$$

$$-\,\delta\left(\prod_{j=0}^{\ell}\gamma_{k-j}\right)\mathbb{E}\left[\mathbb{E}^*_{\mathrm{st}_i(\mathcal{K}_{i_{k-(\ell+1)}})}\left[\prod_{j=0}^{\ell+1} \mathfrak{f}_{k+j-(\ell+1)}^{\varepsilon/(4(k+1))^{\ell+1}}(S_j)\right]\prod_{j=i+1}^{i_{k-(\ell+1)}} \bar{D}_j(\varepsilon/(4(k+1))^{\ell+1})\right].$$

Now, thanks to monotonicity, when we iterate this expression, we may do the following replacements in the procedure. First, for the term not involving δ, any factors of type $\gamma_\ell(\varepsilon', \eta, (d+1)(k+1)f_{\max})$ with $0 < \varepsilon' < \varepsilon$ may be bounded from below by γ_k. Thus, at the $(\ell+1)$th step, we multiply a product of $\gamma_k^{\ell+1}$ to the co-efficient of the main term, leading to the co-efficient γ as defined in (4.68). Moreover, in the final product $\prod_{j=0}^{k} \mathfrak{f}_j^{\varepsilon/(4(k+1))^k}(S_j)$, we may replace $\varepsilon/(4(k+1))^k$ by ε to get a lower bound. This leads to the first term in the statement of the proposition. Next, in the error term involving δ, we bound each γ_ℓ from above by 1, and bound each of the factors of the form $\mathfrak{f}_{k+j-\ell}^{\varepsilon/(4(k+1))^{\ell}}$ from above by $\mathfrak{f}_{k+j-\ell}^{\varepsilon/(4(k+1))^{k+1}}$. This gives us the error term as stated in (4.69). □

We are finally ready to prove Proposition 4.4.13. Recalling (4.47), we bound $\mathbb{E}[N_{\eta,k}(n)]$ from below by summing the lower bound stated in Proposition 4.4.21 over $\eta n < i < i_1 < \ldots < i_k \leqslant n$. We start with the error term. Upon dropping the indicator variables $\bar{D}_j(\varepsilon)$ and bounding $\mathfrak{f}_j^{\varepsilon}$ from above by f_j defined in (4.30) from Section 4.4.2, the absolute value of the error term is bounded from above by

$$\delta\sum_{\eta n<i<n}\ \sum_{\mathcal{I}_k\in\binom{\{i+1,\ldots,n\}}{k}}\mathbb{E}\left[\mathbb{E}^*_{\mathrm{st}_i(\mathcal{K}_i)}\left[\prod_{j=0}^{k} f_j(S_j)\right]\right]. \tag{4.70}$$

From the proof of Corollary 4.4.6 in Section 4.4.2, we know that the double sum converges after re-scaling by n. Hence, there exist $C_1 > 0$ and a natural number N both depending on

ε, η, such that, for all $n \geqslant N$, (4.70) is bounded from above by $C_1 \delta n$.

To treat the main term, assume for now that there exists a constant $C_2 = C_2(\varepsilon, \eta) > 0$ such that, for all $\eta n < i \leqslant n$, we have

$$\sum_{\mathcal{I}_k \in \binom{\{i+1,\ldots,n\}}{k}} \mathbb{E}^*_{\mathrm{st}_i(\mathcal{K}_i)} \left[\prod_{j=0}^{k} \mathfrak{f}^{\varepsilon}_j(S_j) \right] \leqslant C_2. \tag{4.71}$$

We shall use the following inequality: for a non-negative random variable X satisfying $X \leqslant C$, for some $C > 0$, and indicator random variables I_1, I_2 we have

$$\mathbb{E}\left[X\right] \leqslant \mathbb{E}\left[X I_1 I_2\right] + C(\mathbb{E}\left[1 - I_1\right] + \mathbb{E}\left[1 - I_2\right]).$$

Thanks to this inequality, the main term in the lower bound from Proposition 4.4.21 summed over $i < i_1 < \ldots < i_k \leqslant n$ (for fixed $\eta n < i \leqslant n$) can be bounded from below by

$$\gamma \sum_{\mathcal{I}_k \in \binom{\{i+1,\ldots,n\}}{k}} \mathbb{E}\left[\mathbb{E}^*_{\mathrm{st}_i(\mathcal{K}_i)} \left[\prod_{j=0}^{k} \mathfrak{f}^{\varepsilon}_j(S_j) \right] \right] - C_2 \gamma \left(1 - \mathbb{E}\left[\alpha_0 \left(\mathcal{K}_{i \setminus i}, \frac{\varepsilon}{4^{k+1}} \right) \right] + 1 - \mathbb{E}\left[\bar{D}_i \left(\frac{\varepsilon}{4^{k+1}} \right) \right] \right). \tag{4.72}$$

Let $\delta' > 0$. Thanks to Lemma 4.4.15 and the fact that $\mathbb{P}\left(\mathcal{G}^{(i)}_{\varepsilon/(4(k+1))^{k+1}}(i) \right) \to 1$ as $n \to \infty$ uniformly in $\eta n < i \leqslant n$, there exists a natural number $N = N(\delta', \varepsilon, \eta) > 0$ such that, for all $n \geqslant N$, the absolute value of the second term in (4.72) is bounded from above by $C_2 \gamma \delta' \leqslant C_2 \delta'$. Collecting all bounds and using Lemma 4.4.9 concludes the proof of (4.48) upon setting $\varrho = \gamma$. (Note that we may remove the additional $F(S_j)$ in the denominator of $\mathfrak{f}^{\varepsilon}_{\ell}(S_j)$ in the final statement as $F(S_j)$ is bounded by $(k+1)(d+1) f_{\max}$.) Therefore, it remains to establish the existence of C_2 satisfying (4.71). To this end, we shall bound $\mathfrak{f}^{\varepsilon}_j$ from above by f_j (as defined in (4.30)). Note that if $i \geqslant 2$, then $\frac{1}{i-1} \leqslant \frac{2}{\eta n}$. Thus, by applying

Stirling's formula and recalling that $F(S_\ell) \leqslant (d+1)(k+1)f_{\max}$ for all $\ell \in \{0,\dots,k\}$, we have

$$\begin{aligned}
&\sum_{\mathcal{I}_k \in \binom{\{i+1,\dots,n\}}{k}} \prod_{j=0}^{k} f_j(S_j) \\
&\qquad \leqslant \left(1 + O\left(\frac{1}{n}\right)\right) \sum_{i<i_1<\dots<i_k\leqslant n} \prod_{\ell=0}^{k-1} \left(\left(\frac{i_\ell}{i_{\ell+1}}\right)^{\frac{F(S_\ell)}{\lambda_{+\varepsilon}}} \cdot \frac{F(S_\ell)}{\lambda_{-\varepsilon}(i_{\ell+1}-1)} \right) \left(\frac{i_k}{n}\right)^{\frac{F(S_k)}{\lambda_{+\varepsilon}}} \\
&\qquad \leqslant \frac{2\prod_{\ell=0}^{k-1} F(S_\ell)}{\lambda_{-\varepsilon}\eta} \left(1 + O\left(\frac{1}{n}\right)\right) \times \\
&\qquad\qquad \frac{1}{n} \sum_{\eta n<i_0<\dots<i_{k-1}\leqslant n} \prod_{\ell=0}^{k-2} \left(\left(\frac{i_\ell}{i_{\ell+1}}\right)^{\frac{F(S_{\ell+1})}{\lambda_{+\varepsilon}}} \cdot \frac{1}{\lambda_{-\varepsilon}(i_{\ell+1}-1)} \right) \left(\frac{i_{k-1}}{n}\right)^{\frac{F(S_k)}{\lambda_{+\varepsilon}}},
\end{aligned}$$

where the O-term depends only on η. From Corollary 4.4.4 (applied with $k-1$ instead of k) it follows that the right hand side is uniformly bounded for any ε and η.

Proofs of Additional Lemmas used to prove Proposition 4.4.13

We conclude the section with the proofs of Lemmas 4.4.15 and 4.4.16.

Proof of Lemma 4.4.15. Let $i \in \mathbb{N}$ and $\mathcal{X} \in \mathbf{C}^w$ contain a vertex with label i and at most d active faces containing i, where each $(d-1)$-face containing i has fitness at most $f_{\max}$. In the random dynamical process $\mathcal{K}_j, j \geqslant 0$ initiated with complex $\mathcal{X}$, at time $j \geqslant 1$, to each face $\sigma \in \mathcal{K}_j^{(d-1)}$, we can associate a unique ancestral $(d-1)$-dimensional face in $\mathcal{X}$. (Formally, the ancestral face of a face in $\mathcal{X}$ is the face itself. The ancestral face of any other face σ is defined recursively as the ancestral face of the face which was subdivided when σ was formed.) Let $\mathcal{K}_{j\not\downarrow i} \subseteq \mathcal{K}_j$ be the sub-complex of faces of $\mathcal{K}_j$ whose ancestral face does not lie in $\mathrm{st}_i(\mathcal{X})$. Note that $\mathcal{K}_{j\not\downarrow i} \subseteq \mathcal{K}_{j\setminus i}$ and that this inclusion is typically strict due to migration of faces to the outside of the star at times of insertion in the star. For $j \geqslant 1$, let ς_j be j-th time the face chosen in the construction of the simplicial complex has its ancestral face in $\mathcal{X}_{\setminus i}$. Set $\varsigma_0 = 0$. Note that $\varsigma_j \geqslant j$ and that $\varsigma_j - j$ is non-decreasing in j. The crucial observation is that the sequence $\mathcal{K}_{\varsigma_j\not\downarrow i}, j \geqslant 0$ under $\mathbb{P}_{\mathcal{X}}$ is distributed as the sequence $\mathcal{K}_j, j \geqslant 0$ under

$\mathbb{P}_{\mathcal{X}_{\setminus i}}$ upon disregarding vertex labels which are irrelevant here. Formally, this follows from $\mathcal{K}_{\varsigma_0 \downarrow i} = \mathcal{X}_{\setminus i}$ under $\mathbb{P}_{\mathcal{X}}$ and the fact that $\mathcal{K}_{\varsigma_j \downarrow i}, j \geqslant 0$ is Markovian with the same transition rule as $\mathcal{K}_j, j \geqslant 0$. For an integer $K > 0$, on the event $\varsigma_\ell \leqslant \ell + K$ and for any initial configuration $\mathcal{X}$ as described at the beginning of the proof, we have $|F(\mathcal{K}_\ell) - F(\mathcal{K}_{\varsigma_\ell \downarrow i})| \leqslant (2d+1)K f_{\max}$. Hence, for all n sufficiently large, depending on ε, η and K,

$$\begin{aligned}
\mathbb{E}\left[\mathbb{P}_{\mathcal{K}_{i\setminus i}}\left(\bigcap_{j=i+1}^{n} \mathcal{G}_\varepsilon(j-i;j)\}\right)\right] &\geqslant \mathbb{E}\left[\mathbb{P}_{\mathcal{K}_i}\left(\bigcap_{j=i+1}^{n} \{|F(\mathcal{K}_{\varsigma_{j-i}\downarrow i}) - \lambda j| < \varepsilon\lambda j\}\right) \cdot \mathbf{1}_{|\varsigma_{n-i}-(n-i)|\leqslant K}\right] \\
&\geqslant \mathbb{E}\left[\mathbb{P}_{\mathcal{K}_i}\left(\bigcap_{j=i+1}^{n+K} \mathcal{G}_{\varepsilon/2}(j-i;j)\right)\right] \\
&\qquad - \mathbb{E}\left[\mathbb{P}_{\mathcal{K}_i}(|\varsigma_{n-i} - (n-i)| > K)\right] \\
&\geqslant \mathbb{E}\left[\mathbb{P}_{\mathcal{K}_i}\left(\bigcap_{j=i+1}^{\infty} \mathcal{G}_{\varepsilon/2}(j)\right)\right] - \mathbb{E}\left[\mathbb{P}_{\mathcal{K}_i}(|\varsigma_{n-i} - (n-i)| > K)\right].
\end{aligned}$$

By Proposition 4.1.2, for all n sufficiently large, the first term in the last display is at least $1 - \delta/2$ for all $\eta n < i \leqslant n$. Further, we can choose K large enough, such that the absolute value of the second term is bounded from above by $\delta/2$ for all $\eta n < i \leqslant n$ and all n sufficiently large. To see this, note that $\mathbb{P}_x(|\varsigma_n - n| \geqslant K)$ is the probability that the number of faces with ancestral face in $\mathrm{st}_i(x)$ chosen to be subdivided up to time n exceeds K. Let $1 \leqslant \tau_1 < \tau_2 < \cdots$ be the instances, when such faces are chosen. Then, the sought after quantity equals $\mathbb{P}_x(\tau_K \leqslant n)$. Note that τ_K can be bounded from below stochastically by $X_1 + \cdots + X_K$ for independent summands, where X_ℓ follows the geometric distribution with success parameter $\min((d+1)\ell f_{\max}/F(x), 1)$, which implies that $\mathbb{E}[X_1 + \cdots + X_K] \geqslant F(x)\frac{\log K}{(d+1)f_{\max}}$. Thus, if $F(x) \geqslant \lambda\eta n/2$, then, for a given $\varepsilon' > 0$, for any K large enough, depending on η, and all n sufficiently large, depending on ε', η and K, we have $\mathbb{P}_x(\tau_K \leqslant n) \leqslant \varepsilon'$ for all $n \geqslant 1$. This follows from a straightforward application of Chebychev's inequality, whose details we omit. The fact that $F(\mathcal{K}_i) \geqslant \lambda\eta n/2$ with high probability for sufficiently large n, depending on η, concludes the proof of the lemma. □

Proof of Lemma 4.4.16. The proof is very similar to the previous. Let $\mathcal{K}_{j\downarrow\mathcal{X}}$ be the sub-complex of $\mathcal{K}_j$ of faces whose ancestral face lies in $\mathcal{X}$. For $j \geqslant 1$, let $\varsigma_j^{\mathcal{X}}$ be the jth time a face with ancestral face in $\mathcal{X}$ is subdivided. Set $\varsigma_0^{\mathcal{X}} = 0$. As before, we have $\varsigma_j^{\mathcal{X}} \geqslant j$ and $\varsigma_j^{\mathcal{X}} - j$ is non-decreasing. Define $\mathcal{K}_{j\downarrow\mathcal{Y}}$ and $\varsigma_j^{\mathcal{Y}}$ analogously. Thanks to (ii), under $\mathbb{P}_{\mathcal{X}}$, the sequence $\mathcal{K}_{\varsigma_j^{\mathcal{Y}}\downarrow\mathcal{Y}}, j \geqslant 0$ is distributed as $\mathcal{K}_{\varsigma_j^{\mathcal{X}}\downarrow\mathcal{X}}, j \geqslant 0$ under $\mathbb{P}_{\mathcal{Y}}$. Thus, it is enough to show that, under the conditions (i) - (iv), for sufficiently large n, we have

$$\mathbb{P}_{\mathcal{Y}}\left(\bigcap_{j=u+1}^{m} \mathcal{G}_{\varepsilon_2}(j-u,j)\right) - \varepsilon_3/2 \leqslant \mathbb{P}_{\mathcal{Y}}\left(\bigcap_{j=u+1}^{m} \{|F(\mathcal{K}_{\varsigma_{j-u}^{\mathcal{X}}\downarrow\mathcal{X}}) - \lambda j| < 3\varepsilon_2 j/2\}\right)$$

and

$$\mathbb{P}_{\mathcal{X}}\left(\bigcap_{j=u+1}^{m} \{|F(\mathcal{K}_{\varsigma_{j-u}^{\mathcal{Y}}\downarrow\mathcal{Y}}) - \lambda j| < 3\varepsilon_2 j/2\}\right) \leqslant \mathbb{P}_{\mathcal{X}}\left(\bigcap_{j=u+1}^{m} \mathcal{G}_{2\varepsilon_2}(j-u,j)\right) + \varepsilon_3/2.$$

We only show the second statement, as the first can be proved by similar arguments. Note that, for any natural number K, we have

$$\begin{aligned}
&\mathbb{P}_{\mathcal{X}}\left(\bigcap_{j=u+1}^{m} \{|F(\mathcal{K}_{\varsigma_{j-u}^{\mathcal{Y}}\downarrow\mathcal{Y}}) - \lambda j| < 3\varepsilon_2\lambda j/2\}\right) \\
&\leqslant \sum_{p=0}^{K} \mathbb{P}_{\mathcal{X}}\left(\bigcap_{j=u+1}^{m} \{|F(\mathcal{K}_{\varsigma_{j-u}^{\mathcal{Y}}\downarrow\mathcal{Y}}) - \lambda j| < 3\varepsilon_2\lambda j/2, \varsigma_{n-u}^{\mathcal{Y}} = n-u+p\}\right) \\
&\qquad + \mathbb{P}_{\mathcal{X}}(|\varsigma_{n-u}^{\mathcal{Y}} - (n-u)| \geqslant K).
\end{aligned}$$

On $\varsigma_{n-u}^{\mathcal{Y}} = n-u+p$, $0 \leqslant p \leqslant K$, we have, using (i) and (iii),

$$|F(\mathcal{K}_{\varsigma_{j-u}^{\mathcal{Y}}\downarrow\mathcal{Y}}) - F(\mathcal{K}_{j-u})| \leqslant K(d+1)f_{\max} + F\left(\mathcal{X}^{(d-1)} \triangle \mathcal{Y}^{(d-1)}\right) \leqslant K(d+1)f_{\max} + C_1C_2.$$

Here, $F\left(\mathcal{X}^{(d-1)} \triangle \mathcal{Y}^{(d-1)}\right)$ denotes the sum of all finesses of faces in $\mathcal{X}^{(d-1)} \triangle \mathcal{Y}^{(d-1)}$. Thus, for all n sufficiently large, depending on η, ε_2 and K, we can bound the right hand side of the last display from above by

$$\begin{aligned}
&\sum_{p=0}^{K} \mathbb{P}_{\mathcal{X}}\left(\bigcap_{j=u+1}^{m+p} \mathcal{G}_{2\varepsilon_2}(j-u,j) \cap \{\varsigma_{n-u}^{\mathcal{Y}} = n-u+p\}\right) + \mathbb{P}_{\mathcal{X}}(|\varsigma_{n-u}^{\mathcal{Y}} - (n-i)| \geqslant K) \\
&\leqslant \mathbb{P}_{\mathcal{X}}\left(\bigcap_{j=u+1}^{m} \mathcal{G}_{2\varepsilon_2}(j-u,j)\right) + \mathbb{P}_{\mathcal{X}}(|\varsigma_{n-u}^{\mathcal{Y}} - (n-u)| \geqslant K).
\end{aligned}$$

Now, the same arguments relying on a stochastic bound involving sums of independent geometric random variables used in the previous proof show that the second summand can be made smaller than $\varepsilon_3/2$ for sufficiently large, but fixed, K and all n sufficiently large, depending on η, $\varepsilon_1, \varepsilon_3, C_1$ and C_2. Here, one uses (iv) and the fact that $F(\mathcal{X}^{(d-1)} \triangle \mathcal{Y}^{(d-1)}) \leqslant C_1C_2$ to bound the success probabilities of the geometric random variables suitably. $\square$

Bibliography

[1] M.E. Agistein and A.A. Migdal, *Simulations of four-dimensional simplicial quantum gravity as dynamical triangulation*, Modern Physics Letters A **07** (1992), no. 12, 1039–1061.

[2] Marie Albenque and Jean-François Marckert, *Some families of increasing planar maps*, Electron. J. Probab. **13** (2008), no. 56, 1624–1671. MR 2438817

[3] David Aldous, *The continuum random tree. III*, Ann. Probab. **21** (1993), no. 1, 248–289. MR 1207226

[4] José S. Andrade, Hans J. Herrmann, Roberto F. S. Andrade, and Luciano R. da Silva, *Apollonian networks: Simultaneously scale-free, small world, euclidean, space filling, and with matching graphs*, Phys. Rev. Lett. **94** (2005), 018702.

[5] Krishna B. Athreya and Samuel Karlin, *Limit theorems for the split times of branching processes*, J. Math. Mech. **17** (1967), 257–277. MR 0216592

[6] ———, *Embedding of urn schemes into continuous time Markov branching processes and related limit theorems*, Ann. Math. Statist. **39** (1968), 1801–1817. MR 232455

[7] Antar Bandyopadhyay and Debleena Thacker, *A new approach to Pólya urn schemes and its infinite color generalization*, arXiv preprint arXiv:1606.05317, 2016.

[8] Albert-László Barabási and Réka Albert, *Emergence of scaling in random networks*, Science **286** (1999), no. 5439, 509–512. MR 2091634

[9] Shankar Bhamidi, *Universal techniques to analyze preferential attachment trees: global and local analysis*, 2007, Preprint available at https://pdfs.semanticscholar.org/e7fb/8c999ff62a5f080e4c329a7a450f41fb1528.pdf.

[10] Ginestra Bianconi, *Growing Cayley trees described by a Fermi distribution*, Phys. Rev. E **66** (2002), 036116.

[11] Ginestra Bianconi and Albert-László Barabási, *Bose-Einstein condensation in complex networks.*, Phys. Rev. Lett. **86** (2001), 5632–5.

[12] Ginestra Bianconi and Christoph Rahmede, *Complex Quantum Network Manifolds in Dimension $d > 2$ are Scale-Free*, Sci. Rep. **5** (2015), 13979.

[13] Ginestra Bianconi and Christoph Rahmede, *Network geometry with flavor: from complexity to quantum geometry*, Phys. Rev. E **93** (2016), 032315. MR 3657791

[14] ———, *Emergent Hyperbolic Network Geometry*, Sci. Rep. **7** (2017), 41974.

[15] Ginestra Bianconi, Christoph Rahmede, and Zhihao Wu, *Complex quantum network geometries: Evolution and phase transitions*, Phys. Rev. E **92** (2015), 022815.

[16] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, *Complex networks: Structure and dynamics*, Physics Reports **424** (2006), no. 4, 175 – 308.

[17] V. I. Bogachev, *Measure theory. Vol. I, II*, Springer-Verlag, Berlin, 2007. MR 2267655

[18] Béla Bollobás and Oliver Riordan, *Shortest paths and load scaling in scale-free trees.*, Phys. Rev. E (3) **69** (2004), 036114.

[19] Béla Bollobás, Oliver Riordan, Joel Spencer, and Gábor Tusnády, *The degree sequence of a scale-free random graph process*, Random Structures Algorithms **18** (2001), no. 3, 279–290. MR 1824277

[20] Christian Borgs, Jennifer Chayes, Constantinos Daskalakis, and Sebastien Roch, *First to market is not everything: an analysis of preferential attachment with fitness*, STOC'07—Proceedings of the 39th Annual ACM Symposium on Theory of Computing, ACM, New York, 2007, pp. 135–144. MR 2402437

[21] K. A. Borovkov and V. A. Vatutin, *On the asymptotic behaviour of random recursive trees in random environments*, Adv. in Appl. Probab. **38** (2006), no. 4, 1047–1070. MR 2285693

[22] Gunnar Carlsson, *Topology and data*, Bull. Amer. Math. Soc. (N.S.) **46** (2009), no. 2, 255–308. MR 2476414

[23] Nicolas Champagnat and Denis Villemonais, *General criteria for the study of quasi-stationarity*, arXiv preprint arXiv:1712.08092, 2017.

[24] Wen-Chin Chen and Wen-Chun Ni, *Internal path length of the binary representation of heap-ordered trees*, Inform. Process. Lett. **51** (1994), no. 3, 129 – 132.

[25] Owen T. Courtney and Ginestra Bianconi, *Weighted growing simplicial complexes*, Phys. Rev. E **95** (2017), 062301.

[26] Diamantino C. da Silva, Ginestra Bianconi, Rui A. da Costa, Sergey N. Dorogovtsev, and José F. F. Mendes, *Complex network view of evolving manifolds*, Phys. Rev. E **97** (2018), 032316.

[27] Derek J. de Solla Price, *Networks of scientific papers*, Science **149** (1965), no. 3683, 510–515.

[28] Steffen Dereich, *Preferential attachment with fitness: unfolding the condensate*, Electron. J. Probab. **21** (2016), Paper No. 3, 38. MR 3485345

[29] Steffen Dereich, Cécile Mailler, and Peter Mörters, *Nonextensive condensation in reinforced branching processes*, Ann. Appl. Probab. **27** (2017), no. 4, 2539–2568. MR 3693533

[30] Steffen Dereich and Peter Mörters, *Random networks with sublinear preferential attachment: degree evolutions*, Electron. J. Probab. **14** (2009), no. 43, 1222–1267. MR 2511283

[31] Steffen Dereich and Marcel Ortgiese, *Robust analysis of preferential attachment models with fitness*, Combin. Probab. Comput. **23** (2014), no. 3, 386–411. MR 3189418

[32] Jonathan P. K. Doye and Claire P. Massen, *Self-similar disk packings as model spatial scale-free networks*, Phys. Rev. E (3) **71** (2005), no. 1, 016128, 12. MR 2139325

[33] Michael Drmota, *Random trees: An interplay between combinatorics and probability*, Mathematics and Statistics, Springer Vienna, 2009.

[34] G. Ergün and G.J. Rodgers, *Growing random networks with fitness*, Phys. A. **303** (2002), no. 1, 261 – 272.

[35] Michalis Faloutsos, Petros Faloutsos, and Christos Faloutsos, *On power-law relationships of the internet topology*, Proceedings of the Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication (New York, NY, USA), SIGCOMM '99, Association for Computing Machinery, 1999, p. 251–262.

[36] Nikolaos Fountoulakis and Tejas Iyer, *Condensation phenomena in preferential attachment trees with neighbourhood influence*, arxiv preprint arxiv:2101.02734, 2021.

[37] Nikolaos Fountoulakis, Tejas Iyer, Cécile Mailler, and Henning Sulzbach, *Dynamical models for random simplicial complexes*, arXiv preprint arXiv:1910.12715, 2019.

[38] Nic Freeman and Jonathan Jordan, *Extensive condensation in a model of preferential attachment with fitness*, Electron. J. Probab. **25** (2020), Paper No. 68, 42 pp. MR 4115737

[39] Alan Frieze and Charalampos E. Tsourakakis, *On certain properties of random Apollonian networks*, Algorithms and models for the web graph, Lecture Notes in Comput. Sci., vol. 7323, Springer, Heidelberg, 2012, pp. 93–112. MR 2983779

[40] John Haslegrave, Jonathan Jordan, and Mark Yarrow, *Condensation in preferential attachment models with location-based choice*, Random Structures Algorithms **56** (2020), no. 3, 775–795. MR 4084189

[41] Cecilia Holmgren and Svante Janson, *Fringe trees, Crump-Mode-Jagers branching processes and m-ary search trees*, Probab. Surv. **14** (2017), 53–154. MR 3626585

[42] Matthew Hurst, *Data mining: Mapping the blogosphere*, https://datamining.typepad.com/gallery/blog-map-gallery.htm, 2006, Accessed: 04-04-2019.

[43] Tejas Iyer, *Degree distributions in recursive trees with fitnesses*, arxiv preprint arxiv:2005.02197, 2020.

[44] Jean Jacod and Albert N. Shiryaev, *Limit theorems for stochastic processes*, second ed., Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 288, Springer-Verlag, Berlin, 2003. MR 1943877

[45] Svante Janson, *Functional limit theorems for multitype branching processes and generalized Pólya urns*, Stochastic Process. Appl. **110** (2004), no. 2, 177–245. MR 2040966

[46] H. Jeong, S. P. Mason, A. L. Barabási, and Z. N. Oltvai, *Lethality and centrality in protein networks*, Nature **411** (2001), no. 6833, 41–42.

[47] Jonathan Jordan, *Preferential attachment graphs with co-existing types of different fitnesses*, J. Appl. Probab. **55** (2018), no. 4, 1211–1227. MR 3899937

[48] Jonathan Jordan and Andrew R. Wade, *Phase transitions for random geometric preferential attachment graphs*, Adv. in Appl. Probab. **47** (2015), no. 2, 565–588. MR 3360390

[49] Olav Kallenberg, *Foundations of modern probability*, second ed., Springer, New York, 2002.

[50] ———, *Random measures, theory and applications*, Springer, 2017. MR 3642325

[51] J. F. C. Kingman, *A simple model for the balance between selection and mutation*, J. Appl. Probability **15** (1978), no. 1, 1–12. MR 465272

[52] István Kolossváry, Júlia Komjáthy, and Lajos Vágó, *Degrees and distances in random and evolving Apollonian networks*, Adv. in Appl. Probab. **48** (2016), no. 3, 865–902. MR 3568896

[53] Bas Lodewijks and Marcel Ortgiese, *The maximal degree in random recursive graphs with random weights*, arxiv preprint arxiv:2007.05438, 2020.

[54] ———, *A phase transition for preferential attachment models with additive fitness*, arXiv preprint arXiv:2002.12863, 2020.

[55] Hosam M. Mahmoud, *Distances in random plane-oriented recursive trees*, J. Comput. Appl. Math. **41** (1992), no. 1-2, 237–245, Asymptotic methods in analysis and combinatorics. MR 1181723

[56] ———, *Evolution of random search trees*, Wiley-Interscience series in Discrete Mathematics and Optimization, Wiley, 1992.

[57] Hosam M. Mahmoud, R. T. Smythe, and Jerzy Szymański, *On the structure of random plane-oriented recursive trees and their branches*, Random Structures Algorithms **4** (1993), no. 2, 151–176. MR 1206674

[58] Cécile Mailler and Jean-François Marckert, *Measure-valued Pólya urn processes*, Electron. J. Probab. **22** (2017), Paper No. 26, 33. MR 3629870

[59] Cécile Mailler and Denis Villemonais, *Stochastic approximation on non-compact measure spaces and application to measure-valued Pólya processes*, Ann. Appl. Probab. **20** (2020), no. 5, 2393–2438. MR 4149532

[60] Stanley Milgram, *The small-world problem*, Psychology Today **1** (1967), no. 1, 61–67.

[61] A. Yu. Mitrophanov, *Stability and exponential convergence of continuous-time Markov chains*, J. Appl. Probab. **40** (2003), no. 4, 970–979. MR 2012680

[62] T. F. Móri, *On random trees*, Studia Sci. Math. Hungar. **39** (2002), no. 1-2, 143–155. MR 1909153

[63] Hwa Sung Na and Anatol Rapoport, *Distribution of nodes of a tree by degree*, Math. Biosci. **6** (1970), 313–329. MR 278985

[64] D. Najock and C. C. Heyde, *On the number of terminal vertices in certain random trees with an application to stemma construction in philology*, J. Appl. Probab. **19** (1982), no. 3, 675–680. MR 664852

[65] Olle Nerman, *On the convergence of supercritical general (C-M-J) branching processes*, Z. Wahrsch. Verw. Gebiete **57** (1981), no. 3, 365–395. MR 629532

[66] M. E. J. Newman, *The structure and function of complex networks*, SIAM Review **45** (2003), no. 2, 167–256.

[67] Mark Newman, *Networks*, Oxford University Press, 2010.

[68] Roberto Oliveira and Joel Spencer, *Connectivity transitions in networks with superlinear preferential attachment*, Internet Math. **2** (2005), no. 2, 121–163. MR 2193157

[69] Michel Pain and Delphin Sénizergues, *Correction terms for the height of weighted recursive trees*, arXiv preprint arXiv:2101.01156, 2021.

[70] David Pollard, *Convergence of stochastic processes*, Springer Series in Statistics, Springer-Verlag, New York, 1984. MR 762984

[71] JF Rual, K Venkatesan, T Hao, T Hirozane-Kishikawa, A Dricot, N Li, GF Berriz, FD Gibbons, M Dreze, N Ayivi-Guedehoussou, N Klitgord, C Simon, M Boxem, S Milstein, J Rosenberg, DS Goldberg, LV Zhang, SL Wong, G Franklin, S Li, JS Albala, J Lim, C Fraughton, E Llamosas, S Cevik, C Bex, P Lamesch, RS Sikorski, J Vandenhaute, HY Zoghbi, A Smolyar, S Bosak, R Sequerra, L Doucette-Stamm, ME Cusick, DE Hill, FP Roth, and M. Vidal, *Towards a proteome-scale map of the human protein-protein interaction network*, Nature **437** (2005), no. 7062, 1173–1178.

[72] Anna Rudas, Bálint Tóth, and Benedek Valkó, *Random trees and general branching processes*, Random Structures Algorithms **31** (2007), no. 2, 186–202. MR 2343718

[73] Delphin Sénizergues, *Geometry of weighted recursive and affine preferential attachment trees*, arXiv preprint arXiv:1904.07115, 2019.

[74] Herbert A. Simon, *On a class of skew distribution functions*, Biometrika **42** (1955), 425–440. MR 73085

[75] Gábor Szabó, Mikko Alava, and János Kertész, *Shortest paths and load scaling in scale-free trees*, Phys. Rev. E (3) **66** (2002), 026101.

[76] Jerzy Szymański, *On a nonuniform random recursive tree*, Random graphs ’85 (Poznań, 1985), North-Holland Math. Stud., vol. 144, North-Holland, Amsterdam, 1987, pp. 297–306. MR 930497

[77] Remco van der Hofstad, *Random graphs and complex networks. Vol. 1*, Cambridge Series in Statistical and Probabilistic Mathematics, [43], Cambridge University Press, Cambridge, 2017. MR 3617364

[78] Remco van Der Hofstad, *Random Graphs and Complex Networks Volume 2*, 2020, Preprint available at https://www.win.tue.nl/~rhofstad/NotesRGCNII.pdf.

[79] G. Udny Yule, *A mathematical theory of evolution, based on the conclusions of Dr. J. C. Willis, F.R.S.*, Philosophical Transactions of the Royal Society of London. Series B, Containing Papers of a Biological Character **213** (1925), 21–87.

[80] Z. Zhang, L. Rong, and F. Comellias, *High-dimensional random apollonian networks*, Physica A **364** (2006), 610–618.

[81] Z. Zhang, L. Rong, and S. Zhou, *Evolving apollonian networks with small-world scale-free topologies*, Phys. Rev. E **74** (2006), 046105.