Eötvös Loránd University

Institute of Mathematics

Bence Péter Rozner

Random graphs with multiple type edges: asymptotic properties and epidemic spread

Ph.D. Thesis

DOI: 10.15476/ELTE.2021.196

Supervisor:

Ágnes Backhausz, Assistant Professor, Ph.D.

Doctoral School of Mathematics

Director: Tibor Jordán, Professor, D.Sc.

Doctoral Program of Applied Mathematics

Director: János Karátson, Professor, D.Sc.

Department of Probability Theory and Statistics

Budapest, 2021

Acknowledgement

I would like to express my sincere gratitude to my supervisor, Ágnes Backhausz for her guidance and continuous support throughout my PhD studies.

This research was partially supported by Pallas Athene Domus Educationis Foundation. The views expressed are those of the author's and do not necessarily reflect the official opinion of Pallas Athene Domus Educationis Foundation.

Supported by the project "Integrált kutatói utánpótlás-képzési program az informatika és számítástudomány diszciplináris területein (Integrated program for training new generation of researchers in the disciplinary fields of computer science)", No. EFOP-3.6.3-VEKOP-16-2017-00002. The project has been supported by the European Union and co-funded by the European Social Fund.

The research was partially supported by the NKFIH "Élvonal" KKP 133921 grant.

Contents

CHAPTER 1

Random graphs and networks

1.1. Introduction

The topic of random graphs has become an intensively researched field of mathematics in recent decades, mainly due to its wide applicability in solving practical problems, like understanding the spread of infectious diseases or information, related to large-scale networks such as the internet or different kind of biological, financial and social networks. The structure of different random graph models can be very diverse and various problems can be raised that may be of interest both theoretically and practically. The various aspects and applications of random graph models are discussed for example in the following books: Béla Bollobás [**16**], Rick Durrett [**26**] and Remco van der Hofstad [**37**].

One such possible question is associated to the degrees of the vertices of the graph model. In some of the applications, we may be interested in the proportion of the vertices of a specific degree, which is provided by the so-called degree distribution, i.e. the number of vertices with degree d divided by the total number of vertices for every non-negative d. In theoretical problems this distribution can be generalized for infinitely many vertices, that leads us to the notion of the asymptotic degree distribution. Also, the average and the maximum degree of the vertices are well-studied quantities that can be important in some of the applications.

Although the degrees of the vertices provide some information about the structure of the graph, they do not reveal much about the wider neighbourhood of the vertices. The so-called clustering coefficient is a metric that indicates the extent to which the vertices of the graph tend to cluster together. There are several different ways to define the clustering coefficient, i.e. there are both local and global versions as well.

The local clustering coefficient of a specific vertex is defined as the proportion of edges between its neighbours divided by the number of all edges that could possibly exist between all its neighbours. The global clustering coefficient can be calculated by using the number of open and closed triplets in the graph. A triplet is a set of three vertices that can be either open, i.e. the vertices are connected by exactly two edges, or closed, i.e. the vertices are connected by three edges. The global clustering coefficient is defined as the proportion of closed triplets divided by the number of all triplets (open or closed) in the graph. There are many large real-world networks, particularly social networks, that tend to be highly clustered.

Another way to get to know more about the structure of a random graph is to examine the lengths of the paths between every pair of vertices. Since either the set of vertices or the edges (or both of them) are random, the distances in question are (possibly) random variables, so we may be interested in the distribution of the length of the path connecting two specific vertices, or two vertices chosen uniformly at random (or according to any suitable distribution). The diameter of the graph is defined as the maximum of all distances between any pair of vertices, which is also an interesting measure that can be useful in solving specific problems. The notion of small world graph is motivated by existence of certain networks in which the length of the path between any two vertices is relatively short. There are many interesting examples for these small world networks such as the collaboration graph or the network of actors. In these models two vertices are adjacent if the corresponding authors share a common publication or the represented actors play in the same movie, see e.g. in [**37**].

Some other problems relevant to the connection between the vertices of the graph belong to the question of robustness of the structure. In many applications, we may be interested in how sensitive the network is to potential deletion of some of the edges or some of the vertices and the edges that are incident to them. These problems are often related to the spread of infectious diseases on social networks or the spread of information on communications networks. Sometimes, when the structure of the graph is modified, we can inspect a phase transition, i.e. a significant change occurs in the topological properties of the graph when the level of modifications reach a so-called critical point.

Due to the recent Covid-19 pandemic the analysis spread of infectious diseases has gained a lot of importance and it became an intensively studied research area. There are several different ways of modelling such a disease. One of the most popular approaches is using the so-called compartmental models, see e.g. [**40**]. Another way is to use stochastic processes on large (random) graphs, [**15, 26, 40**]. A social network can be modelled by a graph, where vertices represent the individuals in the population, and two vertices are connected if there is a relationship between the two corresponding entities. In order to understand the spread of an infectious disease on the graph, we can label the vertices with different states (e.g. susceptible, infectious, recovered, carrier, exposed and so on). In the labelled graph, a discrete or a continuous-time stochastic process is defined on the phase space of the states of vertices, that we can use to model the spread of the epidemics on the structure of the underlying graph.

In the next section, we explain why it can be useful to enhance the structure of a given model by assigning types to the edges of the graph in order to have a more adequate model to have a better understanding of a real large network, such as the internet and various biological and social networks. We also will describe different families of well-known random graphs and summarize some of the results related to these models, then we show the generalizations of these graphs by labelling the edges of these graphs with different types.

1.2. Random graphs with multiple types of edges

In this thesis we will examine the asymptotic properties of random graph models with multiple type edges and the spread of epidemics processes on these structures by using the tools of probability theory, especially martingale theory and urn processes.

In many applications, it is natural to assign some kind of characteristics to the vertices or to the edges of the graph. For example, the strength of a connection may be represented by weights on the edges, or vertices can have different fitness, which has an impact on their degrees, see e.g. [**25, 34**]. Another class of random graphs are those in which the evolution of the graph depends on groups of existing vertices, e.g. triangles, or even more general configurations formed by the vertices, see [**6, 31, 32**]. A different way to enhance the structure of the graph is embedding the set of vertices and edges into a properly chosen space. There are many large networks where space contains more information than the topology of the graph, see e.g. [**14**].

It may also happen that the type of a vertex or an edge is chosen from a finite set of possibilities. This leads to different phenomena as if we assign weights to the vertices or to the edges. For example, in a social network, the vertices can be considered as males or females, and the edges can be considered as family or work relationships. Another example is the network of financial systems, where the systemic risk is examined, see e.g. [**3**]. To understand these kind of financial systems it is common to use graphs where the vertices represent financial institutions (e.g. banks), and the edges correspond to different types of financial instruments traded by the institutions. The risk arising from these assets (bonds, stocks or options etc.) can be different, which must be taken into account in the calculation of the systemic risk.

1.3. Erdős–Rényi graph

First, let us recall the definition of the classical Erdős–Rényi graph in the single type case, then we define the generalized version in the multi-type case. Some of the first applications of the probabilistic method in graph theory are related to Pál Erdős and Alfréd Rényi, see e.g. [**27, 28**], and also to Edgar Gilbert [**35**]. In the Erdős–Rényi graph, denoted by $G_n(m)$, the number of vertices n is fixed and exactly m pairs of vertices is connected by an edge, where every possible configuration is equally likely. In the Gilbert graph, denoted by $H_n(p)$, the number of vertices n is still fixed, and every pair of vertices is connected with probability p, independently of each other.

In order to obtain the multi-type generalization of the Erdős–Rényi graph (with N different types of edges), let us have $m_1, \ldots, m_N$ positive integers, such that $\sum_{j=1}^{N} m_j = m$. The N-type Erdős–Rényi graph, denoted by $\boldsymbol{G}_n^{(N)}(m_1, \ldots, m_N)$, is a graph on n vertices with exactly m_j edges of type j, chosen uniformly at random from all possible configurations. Similarly, the N-type Gilbert graph, denoted by $\boldsymbol{H}_n^{(N)}(p_1, \ldots, p_N)$ (where $\sum_{j=1}^{N} p_j = p$), is a graph on n vertices, where every pair of vertices is connected with probability p, independently of each other, and conditionally on being connected, the edge is of type j with probability p_j.

1.4. Scale-free graphs

Scale-free graphs form a wide range of random networks, in which the proportion of vertices of degree k approximately equals to $C \cdot k^{-\gamma}$ for sufficiently large k, where γ is known as the characteristic exponent. In other words, it means that the empirical distribution is almost independent of the number of the size of the graph, if the number of vertices is sufficiently large. Notice that the Erdős–Rényi graph is not scale-free, because the proportion of vertices of a given degree follows an exponential law, instead of a power law. There are many scale-free graphs that are widely used in modelling real networks, e.g. the Barabási–Albert graph [**13**].

1.4.1. Preferential attachment graph. The analysis of the preferential attachment graphs is motivated by large real networks, in which vertices of larger degree have more chance to be connected to new vertices. Various types of random graphs with preferential attachment dynamics have been examined in the last decade, see e.g. [**13, 22, 26, 29, 37**].

There are some multi-type preferential attachment graph models that have been investigated in which only the vertices have types. Antunović, Mossel and Rácz introduced a model of competition on growing networks in [**5**]. In their model, when a new vertex is born, it attaches to the old vertices by preferential attachment, and selects its type based on the number of its initial neighbours of each type. Their main interest is the question of coexistence, i.e. the probability that one of the types dies out asymptotically. Abdullah, Bode and Fountoulakis present a model in [**1**], but they use a different rule for choosing the types. At each step, a new vertex is born, it polls some of the old vertices and takes the majority type. A multi-type preferential attachment model was introduced by Rosengren in [**46**] which has similar dynamics to the model presented in [**5**]. The asymptotic degree distribution is examined by using methods from the theory of multi-type branching processes.

1.4.2. Model of independent edges. The model of independent edges has been introduced by Zsolt Katona and Tamás Móri in [**38**]. They consider a random graph evolving in discrete time steps in which a new vertex is born in every steps, and it is connected to all existing vertices with probabilities proportional to the degrees of the other vertices, independently of each other. One possible way to define a multi-type version of the model of independent edges is to enhance the dynamics of the graph by assigning a type to the new edges with probabilities proportional to the current number of edges of different types connected to the existing vertices. In Section 2.2.2, we define another version of the model of independent edges in which the new vertex is connected to all existing vertices with Poisson number of edges of different types, where the distributions of the number of new edges depend on the actual configuration of the graph.

1.4.3. Random graph with duplications and deletions. There are other random graphs evolving in discrete time steps with different kind of dynamics compared to the models described in the previous sections, e.g. the random graph with duplications and deletions examined in [**9, 10**]. In this model, at every step, we choose a vertex v uniformly at random. With probability ϑ we duplicate vertex v; i.e. we add a new vertex and connect it to the neighbours of v and to v itself with single edges. Otherwise, with probability $1-\vartheta$, all the edges of v are deleted. As for the multi-type generalization of the model, one possible approach is to simply copy the types of the original edges whenever we duplicate an existing vertex. These kind of random graphs can be used to model the structure of proteomes.

1.5. Spread of epidemics on random graphs

Random graphs enhanced with multiple type of edges can be used to model the spread of epidemics. In a social network, infectious diseases spread through human contact. Since the relationships between the individuals can be different in nature, the probability of propagation is also different among different people. In this thesis, we examine various types of epidemics on random graphs with multiple type of edges. Then, the infectious disease spreads among the vertices of the graph, so that the probability of infection is different on different types of edges. By using stochastic simulations, we examine the behaviour of the spread of epidemics, when there is also a connection between the types of the edges and the parameters of the process.

In some applications, we can control the spread of the disease, up to a certain level, by separating infected individuals in order to slow down the contagion. We can also assign a state to the edges of the graph, i.e. active or inactive. We assume that the virus cannot spread on inactive edges. At this point, it is clear that if all the edges of the graph are inactive, then the epidemic cannot spread further and all the infected individuals will recover in time, but in practice our goal is to slow down the spread of the infection by eliminating as few connections as possible. Again, by stochastic simulations, we examine the effect of separation (or quarantine), which can be considered as a graph with two types of edges, dynamically changing over time.

We will show that the spread of the epidemic depends on the structure of the underlying graph model, and the introduction of the types of the edges (with the different propagation probabilities) or the quarantine can lead to different results.

There are several articles on models describing the spread of epidemics that include quarantine. One possible direction of the modelling of the spread of infectious diseases is the use of so-called compartmental models. In this approach, we use differential equations to define the dynamics of the change in the number (or in the proportion) of individuals of a given state. In [**36**], [**41**], [**39**] the state of quarantine is also introduced in order to enhance the SIS-, SIR- and $SEIR$-processes.

In [**4**], they investigate the effects of individual decisions on social distancing and isolation in graphs with multi-type vertices. In [**2**], they include groups of age and risk in the SIR-model and find the optimal strategies for quarantine.

1.6. Main results

In this thesis we examine some properties of random graphs that have edges labelled with N different types. We assume that there is a connection between the evolution of the structure of the graph and the types of the edges. In the N-type case, we define the (generalized) degree of a given vertex as $\boldsymbol{d} = (d_1, d_2, \ldots, d_N)$, where d_k is the number of edges of type k connected to it. By using martingale techniques, we prove the existence of an almost sure asymptotic degree distribution. More precisely, we show that for every $\boldsymbol{d}$, the proportion of vertices with generalized degree $\boldsymbol{d}$ tends to some random variable in certain random graph models with multiple type edges as the number of steps (or equivalently the number of vertices) goes to infinity. We also provide recurrence equations for the asymptotic degree distribution. The results are verified not just for particular graph models; instead, we follow a model-free approach and formulate sufficient conditions for the existence of asymptotic degree distribution. Then we give two applications: for a multi-type version of the Barabási–Albert random graph, and for a preferential attachment model with Poisson number of edges. These examples show a new phenomenon: in the multi-type case it can happen that the asymptotic degree distribution is not deterministic, which is the case in many well-known models in the single-type case. We show that the asymptotic degree distribution in the generalized Barabási–Albert random graph and in the model of independent edges also depends on the asymptotic proportion of edges of type k which makes it a stochastic distribution.

In Section 2.4 we are interested in a version of robustness in preferential attachment graph models with multi-type edges. Our aim is to compare a model in which the probability of choosing a type is exactly the proportion of the current type among the edges going out from the endpoint of the new edge; and its modified version, when, after this step, types can change with certain probability. In particular, we introduce perturbation in the multi-type Barabási–Albert random graph, and prove that this shows different phenomena than the original version. That is, errors in the dynamics of multi-type random graphs can lead to essential changes in the asymptotic behaviour of the model. We prove the existence of the asymptotic degree distribution in the perturbed Barabási–Albert random graph, and we also provide recurrence equations for the asymptotic degree distribution. The main difference between the perturbed and the non-perturbed Barabási–Albert random graph is the deterministic or stochastic nature of the asymptotic degree distribution. The reason for that is the asymptotic behaviour of the proportion of edges of different types which can be described by using an urn model. If there is no perturbation, then the proportion of edges of a given type converges to a non-degenerate random variable. On the other hand, if there is perturbation, then it converges to a deterministic constant almost surely. This is based on the properties of the underlying urn models which is explained in more details in [**42**]. In the current thesis, we generalize the results of [**42**] about the almost sure limit of the proportion of edges of different types (or colours) for the case when we also allow multiple drawings with replacement.

Ostroumova, Ryabchenko and Samosvat [**44**] propose a general class of preferential attachment models with single-type edges. They also introduce perturbation in the dynamics, which is different from the one that we have in our model. They assume that the error terms converge to zero with rate $O(1/n)$, where n is the size of the graph. In the perturbed (multi-type) Barabási–Albert random graph, we assume that the probability of errors converges to a positive number.

In Chapter 3 we examine the spread of an infectious disease on several random graph models with multiple type edges. As mentioned before, the introduction of the types of the edges allows us to use more adequate models, because the probabilities of the propagations may depend on the variety of the connections in the graph. First, we generalize the SIR-process for graphs with multi-type edges. Then, we further generalize the process by introducing latency (i.e. infected individuals do not show symptoms for a random period of time) and quarantine (i.e. infected individuals who show symptoms are temporarily separated from the population). Finally, empirical results of some stochastic simulations related to the different processes and underlying structures are presented.

CHAPTER 2

Asymptotic degree distribution in preferential attachment graphs

In this chapter we define a general family of preferential attachment models with multi-type edges, and examine the existence and some properties of the (generalized) asymptotic degree distribution. Two specific graph models, the multi-type versions of the Barabási–Albert graph and the model of independent edges are examined in more details, which are special cases of the general graph model. Then the scale-free property of these models is considered. Finally, the asymptotic properties of a perturbed version of the multi-type Barabási–Albert graph is compared to the non-perturbed case.

Throughout the thesis, $\mathbb{N}$ will denote the set of non-negative integers, furthermore $\boldsymbol{e}_k$ will denote the k^{th} unit vector in $\mathbb{R}^N$ and $\mathbf{1}$ will be the vector with entries all equal to 1.

2.1. Notation and assumptions

Let $(G_n)_{n=0}^{\infty}$ be a sequence of finite random graphs. The sets of the vertices and the edges of G_n are denoted by V_n and E_n, respectively. Throughout the sequel, the number of possible types of edges, denoted by N, will be fixed. For every $k \in [N] = \{1, \ldots, N\}$ let $E_n^{(k)}$ be the set of edges of type k in G_n. For every n we have $E_n = \bigcup_{k\in[N]} E_n^{(k)}$ and for every k, l we have $E_n^{(k)} \cap E_n^{(l)} = \emptyset$ whenever $k \neq l$. That is the different types form a partition of the edges (where we allow empty sets in the partition).

In the following definition we generalize the notion of degree for graphs with multiple types of edges.

DEFINITION 1. For every graph $G(V, E)$ with N different types of edges, the generalized degree of a vertex $v \in V$ in G is defined as

$$\mathbf{deg}(v) = \left(\deg^{(1)}(v), \deg^{(2)}(v), \ldots, \deg^{(N)}(v)\right)^T \in \mathbb{N}^N,$$

where $\deg^{(k)}(v)$ is the number of edges of type k connected to v in G.

REMARK. For a sequence of graphs $(G_n(V_n, E_n))_{n=0}^{\infty}$ with N different types of edges, for every n the generalized degree of a vertex $v \in V_n$ will be denoted by $\mathbf{deg}_n(v)$.

In order to define the (generalized) asymptotic degree distribution of a sequence of random graphs with N different types of edges, we need to introduce the following quantity: for every $\boldsymbol{d} \in \mathbb{Z}^N$ let us have

$$X_n(\boldsymbol{d}) = \left|\{v \in V_n : \mathbf{deg}_n(v) = \boldsymbol{d}\}\right|.$$

This is the number of vertices in G_n with generalized degree $\boldsymbol{d}$. Notice that $X_n(\boldsymbol{d}) = 0$ if there is at least one $k \in [N]$ such that $d_k < 0$.

DEFINITION 2. Let us have a sequence of graphs $(G_n)_{n=0}^{\infty}$ with N different types of edges. If for every $\boldsymbol{d} \in \mathbb{N}^N$ we have

$$\lim_{n\to\infty} \frac{X_n(\boldsymbol{d})}{|V_n|} = x(\boldsymbol{d}) \text{ a.s.},$$

then the family of (possibly) random variables $\left\{x(\boldsymbol{d}), \boldsymbol{d} \in \mathbb{N}^N\right\}$ is called the (generalized) asymptotic degree distribution of $(G_n)_{n=0}^{\infty}$.

Example I. We may consider the following simple 2-type graph model: the initial configuration consists of one single edge of the 1st type. In every step we add a new isolated edge to the graph that is chosen to be of the 1st or 2nd type based on the parity of the index of the step, i.e. for every $n \geq 1$, if n is odd, then the new edge is of the 1st type, and if n is even, then the new edge is of the 2nd type. In this graph sequence, in every step the degrees of all the vertices are equal to one, and approximately half of the vertices have generalized degree $(1, 0)$, and for the other half it is $(0,1)$. More precisely, we have

$$x(1,0) = \lim_{n\to\infty} \frac{X_n(1,0)}{2n} = \frac{1}{2} \quad \text{and} \quad x(0,1) = \lim_{n\to\infty} \frac{X_n(0,1)}{2n} = \frac{1}{2},$$

thus the generalized asymptotic degree distribution is $x(1,0) = \frac{1}{2}$, $x(0,1) = \frac{1}{2}$ and for every $\boldsymbol{d} \in \mathbb{N}^2 \setminus \{(1,0),(0,1)\}$ we have $x(\boldsymbol{d}) = 0$.

Example II. The generalized asymptotic degree distribution does not always exist. As for a possible counterexample we may examine the following simple 2-type graph sequence with alternating types. In the initial step the graph contains one single edge of the 1^{st} type. In every step we add one isolated edge to the graph, however the types of all the edges are chosen to be of the 1^{st} type in every odd step, and the are chosen to be of the 2^{nd} type in every even step, i.e. the types of the already existing edges are also changed in every step in an alternating way. One can see that for every n we have

$$\frac{X_n(1,0)}{2n} = \begin{cases} 1 & \text{if } n \text{ is even} \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad \frac{X_n(0,1)}{2n} = \begin{cases} 1 & \text{if } n \text{ is odd} \\ 0 & \text{otherwise} \end{cases}$$

which means that

$$x(1,0) = \lim_{n\to\infty} \frac{X_n(1,0)}{2n} \quad \text{and} \quad x(0,1) = \lim_{n\to\infty} \frac{X_n(0,1)}{2n}$$

do not exist, thus the generalized asymptotic degree distribution does not exist either.

2.2. Asymptotic degree distribution in the general model

In this section we introduce a general family of random graph models with multiple types of edges evolving in discrete time steps according to a stochastic dynamics. First we determine the initial configuration, then we list the assumptions on the dynamics of the models that determines the evolution of the structure of the graph sequences.

Initial configuration. The initial configuration is denoted by $G_0(V_0, E_0)$, where $V_0 = \{u_1, u_2, \ldots, u_s\}$ and $s \geq 1$ are fixed. As for the structure of the graph, we allow multiple edges, but loops are forbidden. We assume that for every $k \in [N]$ we have $\left|E_0^{(k)}\right| > 0$, which means that there is at least one edge of each type. If the last condition does not hold, then we may omit every type that is not present in the initial configuration, since new types will not be added during the evolution of the graph model.

Dynamics. The dynamics of the general family of random graphs, which is the subject of the thesis, is the following: we add only one vertex in every step which is connected to some of the old vertices with possibly multiple edges, then we assign a type to every new edge. We also assume that new edges are not added or deleted among the pairs of already existing vertices and the types of the edges do not change during the evolution of the structure of the graph, i.e. we have $E_n^{(k)} \subseteq E_{n+1}^{(k)}$ for every $n \geq 0$ and for every $k \in [N]$.

The dynamics of the evolution of the general model can be described in the following way: for every n, in the n^{th} step,

(1) a new vertex, denoted by v_n, is added to the set of vertices, thus we have $V_n = V_0 \cup \{v_1, v_2, \ldots, v_n\}$.
(2) The new vertex v_n is attached to some of the existing vertices with at least one edge, so every element of the edge set $E_n \setminus E_{n-1}$ is connected to v_n.
(3) Every new edge gets a type according to a stochastic rule. For example, we may consider the following case: for every n, in the n^{th} step, any edge between the new vertex v_n and an existing vertex $v \in V_{n-1}$ will be assigned to type k with probabilities proportional to $\deg_{n-1}^{(k)}(v)$ for every $k \in [N]$.

Assumptions. In order to formulate the conditions on the evolution of the structure of the graph sequence we are going to introduce some notations. For every $n \geq 1$ let $\mathcal{F}_n$ denote the σ-algebra generated by the first n graphs with labelled edges, and let $\mathcal{F}_0$ be the trivial σ-algebra, thus $\mathcal{F} = (\mathcal{F}_n)_{n=0}^{\infty}$ is a filtration. The sum of the coordinates of a given vector $\boldsymbol{x} = (x_1, \ldots, x_N)^T \in \mathbb{Z}^N$ will be denoted by

$s(\boldsymbol{x}) = \boldsymbol{x}^T\mathbf{1} = \sum_{k=1}^N x_k$. Now we list the assumptions related to the dynamics of the general model that we are going to use throughout the thesis.

(GM1): For every n, we assume that the conditional distribution of the number of new edges of type l connected to $v \in V_{n-1}$, conditionally with respect to $\mathcal{F}_{n-1}$, depends only on $\mathbf{deg}_{n-1}(v)$ for every $l \in [N]$. By using this assumption, we denote by $p_{\boldsymbol{d}}^{(n)}(\boldsymbol{i})$, where $\boldsymbol{i} = (i_1, \dots, i_N)^T$, the conditional probability that a vertex with generalized degree $\boldsymbol{d}$ gets exactly i_l new edges of type l for every $l \in [N]$, conditionally with respect to $\mathcal{F}_{n-1}$. In other words, for every existing vertex the probability of having a new edge only depends on the actual generalized degree of the vertex.

(GM2): For every $\boldsymbol{d} \in \mathbb{N}^N$, there exists $\delta > 0$ and $C > 0$, such that

$$\mathbb{E}\left(|X_n(\boldsymbol{d}) - X_{n-1}(\boldsymbol{d})|^2 \middle| \mathcal{F}_{n-1}\right) \leq Cn^{1-\delta}$$

holds almost surely for every n. This means that the difference of the number of vertices with generalized degree $\boldsymbol{d}$ in the actual step and in the previous step is bounded in some sense. If the degree of the new vertex is uniformly bounded, then this condition is automatically fulfilled.

(GM3): For every $\boldsymbol{d} \in \mathbb{N}^N$, we define the sequence $(u_n(\boldsymbol{d}))_{n=1}^{\infty}$ by the following equality:

$$p_{\boldsymbol{d}}^{(n)}(\mathbf{0}) = 1 - \frac{u_n(\boldsymbol{d})}{n}.$$

The sequence $(u_n(\boldsymbol{d}))_{n=1}^{\infty}$ is non-negative and predictable with respect to the filtration $\mathcal{F}$. We assume that there is a positive random variable denoted by $u(\boldsymbol{d})$, such that $u_n(\boldsymbol{d}) \to u(\boldsymbol{d})$ almost surely as $n \to \infty$. This means that for a vertex with generalized degree $\boldsymbol{d}$ the probability of not receiving any new edge stabilizes in some sense.

(GM4): For every $\boldsymbol{d} = (d_1, \ldots, d_N)^T \in \mathbb{N}^N$ such that $s(\boldsymbol{d}) \geq 1$, let us introduce the following set of indices:

$$H(\boldsymbol{d}) = \left\{ \boldsymbol{i} = (i_1, \ldots, i_N)^T \in \mathbb{N}^N : \forall k \in [N] : i_k \leq d_k \text{ and } s(\boldsymbol{i}) \geq 1 \right\}.$$

We assume that for every $\boldsymbol{d} \in \mathbb{N}^N$ such that $s(\boldsymbol{d}) \geq 1$, and for every $\boldsymbol{i} \in H(\boldsymbol{d})$ there are families of non-negative random variables denoted by $r^{(k)}(\cdot)$ where $k \in [N]$ with the following property:

$$\lim_{n\to\infty} np^{(n)}_{\boldsymbol{d}-\boldsymbol{i}}(\boldsymbol{i}) = \begin{cases} r^{(k)}(\boldsymbol{d}-\boldsymbol{e}_k) & \text{if } \boldsymbol{i} = \boldsymbol{e}_k \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

holds almost surely. Assumption **(GM4)** states that for an existing vertex the probability of receiving edges of different types in order to obtain generalized degree $\boldsymbol{d}$ has a non-trivial limit if and only if one of the edges of a given type is missing. That is, although it may happen that a vertex gets more than one edges in a step, this has probability of $o\left(\frac{1}{n}\right)$ and disappears from the asymptotic equations.

(GM5): For every n and for every $\boldsymbol{d} \in \mathbb{N}^N$, we denote by $q^{(n)}(\boldsymbol{d})$ the conditional probability that the new vertex v_n is connected to the existing vertices with exactly d_l edges of type l, conditionally with respect to $\mathcal{F}_{n-1}$. We assume that there exists a non-negative random variable denoted by $q(\boldsymbol{d})$ such that $q^{(n)}(\boldsymbol{d}) \to q(\boldsymbol{d})$ almost surely as $n \to \infty$. This means that for the new vertex in the n^{th} step the probability of having generalized degree $\boldsymbol{d}$ converges as $n \to \infty$.

Remark. For every $\boldsymbol{d} = (d_1, \ldots, d_N)^T \in \mathbb{N}^N$ if there is at least one $k \in [N]$, such that $d_k < 0$, then $u(\boldsymbol{d})$ and $q(\boldsymbol{d})$ are chosen to be zero. Similarly, for every $\boldsymbol{d} \in \mathbb{N}^N$ and for every $k \in [N]$ if there is at least one $l \in [N]$, such that $(\boldsymbol{d} - \boldsymbol{e}_k)_l < 0$, then $r^{(k)}(\boldsymbol{d} - \boldsymbol{e}_k)$ is chosen to be zero.

Remark. There are many other interesting features that could be included in the dynamics of the graph models, e.g. the deletion of the edges can be also introduced, or we may allow the edges to be added between already existing vertices, and it may also happen that existing edges change their types during the evolution of the structure of the graph.

Now we can formulate our main theorem on the asymptotic degree distribution.

THEOREM 1. *If a sequence of N-type random graphs, denoted by $(G_n(V_n, E_n))_{n=0}^{\infty}$, satisfies the assumptions **(GM1)-(GM5)**, then for every $\boldsymbol{d} \in \mathbb{N}^N$ we have*

$$\lim_{n\to\infty} \frac{X_n(\boldsymbol{d})}{|V_n|} = x(\boldsymbol{d}) \ \ a.s.$$

The random variables $\left\{x(\boldsymbol{d}), \boldsymbol{d} \in \mathbb{N}^N\right\}$ satisfy the following recurrence equations:

$$x(\boldsymbol{d}) = \frac{1}{u(\boldsymbol{d})+1}\left[\sum_{k=1}^{N} r^{(k)}(\boldsymbol{d}-\boldsymbol{e}_k)x(\boldsymbol{d}-\boldsymbol{e}_k) + q(\boldsymbol{d})\right].$$

REMARK. Notice that the initial condition of the recurrence equations has not been defined explicitly. If there is at least one $l \in [N]$ such that $(\boldsymbol{d}-\boldsymbol{e}_k)_l < 0$, then the first term in the sharp bracket equals to zero, and if there is at least one $l \in [N]$ such that $d_l < 0$, then we also have $q(\boldsymbol{d}) = 0$.

Preliminaries for the Proof of Theorem 1

In this section we present the tools that play an essential role in the proof of out main theorem.

DEFINITION 3. Two sequences $(a_n)_{n=1}^{\infty}$ and $(b_n)_{n=1}^{\infty}$ are asymptotically equal (denoted by $a_n \sim b_n$) if they are positive except finitely many terms, and we have $\lim_{n\to\infty} \frac{a_n}{b_n} = 1$.

DEFINITION 4. A sequence $(\beta_n)_{n=1}^{\infty}$ is called regularly varying with exponent κ if $\beta_n \sim \gamma_n n^{\kappa}$, where $(\gamma_n)_{n=1}^{\infty}$ is a slowly varying sequence. A sequence $(\gamma_n)_{n=1}^{\infty}$ is slowly varying if for every positive s we have $\lim_{n\to\infty} \frac{\gamma_{[sn]}}{\gamma_n} = 1$.

We will use the following theorem, see also [**25**] for a similar statement.

LEMMA 1 (Lemma 1 in [**7**]). *Let $\mathcal{F} = (\mathcal{F}_n)_{n=1}^\infty$ be a filtration, $(\xi_n)_{n=1}^\infty$ a nonnegative adapted process with respect to $\mathcal{F}$. Let $(w_n)_{n=1}^\infty$ be a regularly varying sequence of positive numbers with exponent $\kappa > -1$. Suppose that for every $n \geq 1$,*

$$\mathbb{E}\left((\xi_n - \xi_{n-1})^2 \middle| \mathcal{F}_{n-1}\right) = O\left(n^{1-\delta+2\kappa}\right) \tag{2}$$

holds with some $\delta > 0$. Let $(u_n)_{n=1}^\infty$, $(v_n)_{n=1}^\infty$ be nonnegative predictable processes with respect to $\mathcal{F}$ such that $u_n < n$ for all $n \geq 1$.

(a) Suppose that

$$\mathbb{E}\left(\xi_n \middle| \mathcal{F}_{n-1}\right) \leq \left(1 - \frac{u_n}{n}\right)\xi_{n-1} + v_n,$$

and $\lim_{n\to\infty} u_n = u$, $\limsup_{n\to\infty} \frac{v_n}{w_n} \leq v$ with some random variables $u > 0$, $v \geq 0$. Then we have

$$\limsup_{n\to\infty} \frac{\xi_n}{n w_n} \leq \frac{v}{u + \kappa + 1} \quad a.s.$$

(b) Suppose that

$$\mathbb{E}\left(\xi_n \middle| \mathcal{F}_{n-1}\right) \geq \left(1 - \frac{u_n}{n}\right)\xi_{n-1} + v_n,$$

and $\lim_{n\to\infty} u_n = u$, $\liminf_{n\to\infty} \frac{v_n}{w_n} \geq v$ with some random variables $u > 0$, $v \geq 0$. Then we have

$$\liminf_{n\to\infty} \frac{\xi_n}{n w_n} \geq \frac{v}{u + \kappa + 1} \quad a.s.$$

We will use this lemma for the sequence $w_n \equiv 1$ and $\kappa = 0$.

Proof of Theorem 1. We prove the theorem by induction on the value of $s(\boldsymbol{d}) = \boldsymbol{d}^T\mathbf{1} = \sum_{k=1}^N d_k$. The initial step of the induction, when we set $s(\boldsymbol{d}) = 0$, is trivial. Let $\boldsymbol{d} = (d_1, \ldots, d_N)^T \in \mathbb{N}^N$ be a fixed vector such that $s(\boldsymbol{d}) > 0$. Notice that, for every $n \geq 1$, in the n^{th} step, the value of $X_n(\boldsymbol{d})$ may change due to one of the following events:

- an existing vertex with generalized degree $\boldsymbol{d}$ is connected to the new vertex;

- an existing vertex with generalized degree $\boldsymbol{d}-\boldsymbol{i}=(d_k-i_k)_{k=1}^N$ is chosen, and it gets i_k new edges of type k;
- the new vertex is attached to the old vertices with d_k edges of type k for every $k\in[N]$.

For every $n\geq 1$, in the n^{th} step, we have

$$\mathbb{E}\left[X_n(\boldsymbol{d})\middle|\mathcal{F}_{n-1}\right]=X_{n-1}(\boldsymbol{d})p_{\boldsymbol{d}}^{(n)}(\boldsymbol{0})+\left[\sum_{\boldsymbol{i}\in H(\boldsymbol{d})}X_{n-1}(\boldsymbol{d}-\boldsymbol{i})p_{\boldsymbol{d}-\boldsymbol{i}}^{(n)}(\boldsymbol{i})\right]+q^{(n)}(\boldsymbol{d}), \tag{3}$$

where

$$H(\boldsymbol{d})=\left\{\boldsymbol{i}=(i_1,\dots,i_N)\in\mathbb{N}^N:\forall k\in[N]:i_k\leq d_k\text{ and }s(\boldsymbol{i})\geq 1\right\}.$$

Assumption **(GM2)** implies that there exists a positive δ and a positive C such that for every $n\geq 1$ we have

$$\mathbb{E}\left(\left|X_n(\boldsymbol{d})-X_{n-1}(\boldsymbol{d})\right|^2\middle|\mathcal{F}_{n-1}\right)\leq Cn^{1-\delta}.$$

With this δ, equation (2) in Lemma 1 is satisfied with $\xi_n=X_n(\boldsymbol{d})$. We want to rewrite equation (3) in the following form:

$$\mathbb{E}\left[X_n(\boldsymbol{d})\middle|\mathcal{F}_{n-1}\right]=X_{n-1}(\boldsymbol{d})\left[1-\frac{u_n(\boldsymbol{d})}{n}\right]+v_n(\boldsymbol{d}),$$

where the processes $(u_n(\boldsymbol{d}))_{n=1}^\infty$ and $(v_n(\boldsymbol{d}))_{n=1}^\infty$ satisfy the assumptions of Lemma 1. Recall the definition of $u_n(\boldsymbol{d})$ from Assumption **(GM3)**. It is easy to see that this process is predictable with respect to $\mathcal{F}$. Assumption **(GM3)** implies that there exists a positive random variable $u(\boldsymbol{d})$ such that $u_n(\boldsymbol{d})\to u(\boldsymbol{d})$ almost surely as $n\to\infty$. We define $H'(\boldsymbol{d})=H(\boldsymbol{d})\setminus\{\boldsymbol{e}_k,k\in[N]\}$.

We define

$$v_n(\boldsymbol{d})=\sum_{k=1}^N X_{n-1}(\boldsymbol{d}-\boldsymbol{e}_k)p_{\boldsymbol{d}-\boldsymbol{e}_k}^{(n)}(\boldsymbol{e}_k)+\left[\sum_{\boldsymbol{i}\in H'(\boldsymbol{d})}X_{n-1}(\boldsymbol{d}-\boldsymbol{i})p_{\boldsymbol{d}-\boldsymbol{i}}^{(n)}(\boldsymbol{i})\right]+q^{(n)}(\boldsymbol{d}).$$

It is easy to see that this process is predictable with respect to $\mathcal{F}$. Using Assumptions **(GM4)** and **(GM5)** and the induction hypothesis, we conclude that there exists a non-negative random variable $v(\boldsymbol{d})$, such that

$$v_n(\boldsymbol{d}) \to v(\boldsymbol{d}) = \sum_{k=1}^{N} r^{(k)}(\boldsymbol{d} - \boldsymbol{e}_k)x(\boldsymbol{d} - \boldsymbol{e}_k) + q(\boldsymbol{d}) \text{ a.s.}$$

as $n \to \infty$. Lemma 1 implies that

$$\lim_{n\to\infty} \frac{X_n(\boldsymbol{d})}{n} = \frac{v(\boldsymbol{d})}{u(\boldsymbol{d}) + 1} \text{ a.s.}$$

Since $|V_n| \sim n$, the proof of Theorem 1 is complete. □

2.2.1. Generalized Barabási–Albert random graph. This is a multi-type version and a generalization (or modification) of the graph model in [**13**], specified in [**18**] (see also [**37, 30, 44**] for general setups).

The dynamics of this model is the following:

- for every $n \geq 1$, in the n^{th} step, the new vertex v_n is connected with M_n (not necessarily different) edges to some of the already existing vertices, where M_n is a positive integer-valued random variable, which is independent of $\mathcal{F}_{n-1}$.
- The endpoints of the M_n new edges are chosen independently. The endpoint of each edge is chosen among the existing vertices with probabilities proportional to the degrees. (Notice that we do not update degrees until the end of step.)
- The types of the new edges are chosen independently, and the probability of each type is its proportion among the edges of the already existing endpoint of the new edge (not counting the edges added in the actual step).

Now, we list the assumptions on the sequence of random variables $(M_n)_{n=1}^{\infty}$.

Assumption (BA1) We assume that M_n is a positive integer-valued random variable, which is independent of $\mathcal{F}_{n-1}$ for every $n \geq 1$.

Assumption (BA2) We assume that there exists a positive random variable, denoted by M, such that $M_n \to M$ in distribution, and for every $p \geq 1$ we have $\mathbb{E}(M_n^p) \to \mathbb{E}(M^p) < \infty$ as $n \to \infty$. The expected value of M will be denoted by $m = \mathbb{E}(M)$.

We will use the following lemma in order to understand the asymptotic behaviour of the proportion of edges of type k as the number of steps (or equivalently the number of vertices) goes to infinity.

LEMMA 2. *For every $k \in [N]$ let us define $\zeta_n^{(k)} = \frac{|E_n^{(k)}|}{|E_n|}$, i.e. the proportion of the number of edges of type k in the generalized Barabási–Albert random graph. For every $k \in [N]$ there exists a random variable $\zeta^{(k)}$ such that $\zeta_n^{(k)} \to \zeta^{(k)}$ almost surely as $n \to \infty$.*

REMARK. If we have $M_n \equiv 1$ for all $n \geq 1$, and the initial configuration is a tree, i.e. the model is an N-type Barabási–Albert random tree, then $\left(\zeta^{(k)}, k \in [N]\right)$ has a Dirichlet distribution with parameters $\left(|E_0^{(k)}|, k \in [N]\right)$. In this case the number of edges with different types follows a Pólya urn process.

Proof of Lemma 2. First, let us fix $k \in [N]$. For every $n \geq 1$ the distribution of the number of new edges of type k in the n^{th} step conditionally with respect to $\mathcal{F}_{n-1}^+$ is $\text{Bin}\left(M_n, \frac{\left|E_{n-1}^{(k)}\right|}{|E_{n-1}|}\right)$. For every $n \geq 1$ we have

$$\mathbb{E}\left(\frac{\left|E_n^{(k)}\right|}{|E_n|} \middle| \mathcal{F}_{n-1}^+\right) = \frac{\left|E_{n-1}^{(k)}\right|}{|E_{n-1}| + M_n} + \frac{M_n \frac{\left|E_{n-1}^{(k)}\right|}{|E_{n-1}|}}{|E_{n-1}| + M_n} = \frac{\left|E_{n-1}^{(k)}\right| \left(1 + \frac{M_n}{|E_{n-1}|}\right)}{|E_{n-1}| + M_n} = \frac{\left|E_{n-1}^{(k)}\right|}{|E_{n-1}|}.$$

This is $\mathcal{F}_{n-1}$-measurable, hence this yields

$$\mathbb{E}\left(\frac{\left|E_n^{(k)}\right|}{|E_n|} \middle| \mathcal{F}_{n-1}\right) = \frac{\left|E_{n-1}^{(k)}\right|}{|E_{n-1}|}.$$

We conclude that $\left(\zeta_n^{(k)}, \mathcal{F}_n\right)_{n=1}^{\infty}$ is a nonnegative martingale, thus it is convergent almost surely. Let us denote its limit by $\zeta^{(k)} \geq 0$. The proof of Lemma 2 is complete. □

Asymptotic degree distribution in the generalized Barabási–Albert random graph.

THEOREM 2. *If the assumptions **(BA1)** and **(BA2)** on the sequence* $(M_n)_{n=1}^{\infty}$ *are satisfied, then in the generalized Barabási–Albert model for every* $\boldsymbol{d} = (d_1, \ldots, d_N)^T \in \mathbb{N}^N$ *we have*

$$\lim_{n\to\infty} \frac{X_n(\boldsymbol{d})}{|V_n|} = x(\boldsymbol{d}) \ \text{a.s.}$$

The random variables $\left\{x(\boldsymbol{d}), \boldsymbol{d} \in \mathbb{N}^N\right\}$ *satisfy the following recurrence equations:*

$$x(\boldsymbol{d}) = \sum_{k=1}^{N} \frac{d_k - 1}{D+2} x(\boldsymbol{d} - \boldsymbol{e}_k) + \frac{2}{D+2} \mathbb{P}\left(M = D\right) \frac{D!}{\prod_{k=1}^{N} d_k!} \prod_{k=1}^{N} \left(\zeta^{(k)}\right)^{d_k},$$

where $\zeta^{(k)}$ *is defined in Lemma 2 and* $D = s(\boldsymbol{d}) = \sum_{k=1}^{N} d_k$.

Preliminaries for the proof of Theorem 2. First, for every $n \geq 0$ we define the following σ-algebra: $\mathcal{F}_n^+ = \sigma(\mathcal{F}_n, M_{n+1})$. Then we will show that we have $|E_n| \sim mn$, where $m = \mathbb{E}(M)$. For every $n \geq 1$ the number of edges equals to $|E_n| = \sum_{k=1}^{N} \left|E_0^{(k)}\right| + \sum_{i=1}^{n} M_i$. By the assumptions of the model, the sequence $(M_n)_{n=1}^{\infty}$ satisfies the following conditions:

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}(M_i) = \mathbb{E}(M) = m > 0 \quad \text{and} \quad \sum_{n=1}^{\infty} \frac{\operatorname{Var}(M_n)}{n^2} < \infty.$$

Therefore Kolmogorov's theorem can be applied (Theorem 6.7. in [**45**]) for the sequence $(M_n)_{n=1}^{\infty}$, thus we have $|E_n| \sim mn$.

We will use the following lemma, which can be proved by Bonferroni's inequality.

LEMMA 3. *For every* $n \geq 1$ *and* $x \in [0,1]$ *we have*

$$|(1-x)^n - (1-nx)| \leq \binom{n}{2} x^2.$$

Proof of Theorem 2. We will directly use Theorem 1, so it is required to check if the assumptions of the general model are fulfilled.

Assumption (GM1). By the definition of the dynamics of the multi-type Barabási–Albert model, it is easy to see that Assumption **(GM1)** trivially holds.

Assumption (GM2). Assumption **(BA2)** implies that, for every $n \geq 1$ and $\boldsymbol{d} \in \mathbb{N}^N$ we have

$$\mathbb{E}\left(\left|X_n(\boldsymbol{d}) - X_{n-1}(\boldsymbol{d})\right|^2 \middle| \mathcal{F}_{n-1}\right) \leq \mathbb{E}(M_n^2) \to \mathbb{E}(M^2) < \infty$$

as $n \to \infty$. If we choose $\delta = 1$, then Assumption **(GM2)** is satisfied.

Assumption (GM3). For every $n \geq 1$ and $\boldsymbol{d} \in \mathbb{N}^N$ the distribution of the number of new edges added to an existing vertex with generalized degree $\boldsymbol{d}$ follows a binomial distribution, thus the probability that a vertex with such a degree does not receive any new edges equals to

$$p_{\boldsymbol{d}}^{(n)}(\boldsymbol{0}) = \mathbb{E}\left[\left(1 - \frac{s(\boldsymbol{d})}{2|E_{n-1}|}\right)^{M_n} \middle| \mathcal{F}_{n-1}\right],$$

where $s(\boldsymbol{d}) = \sum_{i=1}^N d_i$. To calculate the expected value above, we will use the following formula:

$$\mathbb{E}\left[\left(1 - \frac{s(\boldsymbol{d})}{2|E_{n-1}|}\right)^{M_n} \middle| \mathcal{F}_{n-1}\right] = \mathbb{E}\left(1 - M_n \frac{s(\boldsymbol{d})}{2|E_{n-1}|} \middle| \mathcal{F}_{n-1}\right) + \eta_n(\boldsymbol{d}),$$

where

$$\eta_n(\boldsymbol{d}) = \mathbb{E}\left[\left(1 - \frac{s(\boldsymbol{d})}{2|E_{n-1}|}\right)^{M_n} - \left(1 - M_n \frac{s(\boldsymbol{d})}{2|E_{n-1}|}\right) \middle| \mathcal{F}_{n-1}\right].$$

Lemma 3 implies that for every $n \geq s(\boldsymbol{d})$ we have

$$\left|\left(1 - \frac{s(\boldsymbol{d})}{2|E_{n-1}|}\right)^{M_n} - \left(1 - M_n \frac{s(\boldsymbol{d})}{2|E_{n-1}|}\right)\right| \leq \binom{M_n}{2} \left(\frac{s(\boldsymbol{d})}{2|E_{n-1}|}\right)^2.$$

By using the above bound, we obtain that

$$\begin{aligned}|\eta_n(\boldsymbol{d})| &\leq \mathbb{E}\left[\left|\left(1-\frac{s(\boldsymbol{d})}{2|E_{n-1}|}\right)^{M_n}-\left(1-M_n\frac{s(\boldsymbol{d})}{2|E_{n-1}|}\right)\right|\middle|\mathcal{F}_{n-1}\right]\\ &\leq \mathbb{E}\left[\binom{M_n}{2}\left(\frac{s(\boldsymbol{d})}{2|E_{n-1}|}\right)^2\middle|\mathcal{F}_{n-1}\right]=\left(\frac{s(\boldsymbol{d})}{2|E_{n-1}|}\right)^2\mathbb{E}\left[\binom{M_n}{2}\right]\\ &\leq \left(\frac{s(\boldsymbol{d})}{2|E_{n-1}|}\right)^2\mathbb{E}(M_n^2)\end{aligned}$$

almost surely, by using the fact that $|E_{n-1}| \sim mn$, and also assumption **(BA2)**. The definition of $u_n(\boldsymbol{d})$ and $\eta_n(\boldsymbol{d})$ implies that

$$\begin{aligned}u_n(\boldsymbol{d}) &= n\left(1-\left[\mathbb{E}\left(1-M_n\frac{s(\boldsymbol{d})}{2|E_{n-1}|}\middle|\mathcal{F}_{n-1}\right)+\eta_n(\boldsymbol{d})\right]\right)\\ &= n\frac{s(\boldsymbol{d})}{2}\cdot\frac{\mathbb{E}(M_n)}{|E_{n-1}|}-n\cdot\eta_n(\boldsymbol{d}).\end{aligned}$$

This is $\mathcal{F}_{n-1}$-measurable, hence $(u_n(\boldsymbol{d}))_{n=1}^{\infty}$ is a predictable process with respect to the filtration $\mathcal{F}$. Recall that $|E_{n-1}| \sim mn$ and $n \cdot |\eta_n(\boldsymbol{d})| = o(1)$ almost surely. Assumption **(BA2)** implies that

$$u(\boldsymbol{d}) = \lim_{n\to\infty} u_n(\boldsymbol{d}) = \frac{s(\boldsymbol{d})}{2} \text{ a.s.}$$

Assumption (GM4). First, we fix $k \in [N]$. For every $n \geq 1$ and $\boldsymbol{d} \in \mathbb{N}^N$, where $s(\boldsymbol{d}) \geq 1$, by using the fact that the distribution of the number of new edges added to an existing vertex with generalized degree $\boldsymbol{d}$ follows a binomial distribution, the probability that a vertex with such a degree does not receive exactly one new edge equals to

$$\begin{aligned}p^{(n)}_{\boldsymbol{d}-\boldsymbol{e}_k}(\boldsymbol{e}_k) &= \mathbb{E}\left[M_n\left(\frac{d_k-1}{2|E_{n-1}|}\right)\left(1-\frac{s(\boldsymbol{d})-1}{2|E_{n-1}|}\right)^{M_n-1}\middle|\mathcal{F}_{n-1}\right]\\ &= \left(\frac{d_k-1}{2|E_{n-1}|}\right)\mathbb{E}\left[M_n\left(1-\frac{s(\boldsymbol{d})-1}{2|E_{n-1}|}\right)^{M_n-1}\middle|\mathcal{F}_{n-1}\right].\end{aligned}\tag{4}$$

Similarly to the previous case, we obtain that

$$\begin{aligned}&\mathbb{E}\left[M_n\left(1-\frac{s(\boldsymbol{d})-1}{2|E_{n-1}|}\right)^{M_n-1}\middle|\mathcal{F}_{n-1}\right]\\&=\mathbb{E}\left[M_n\left(1-(M_n-1)\frac{s(\boldsymbol{d})-1}{2|E_{n-1}|}\right)\middle|\mathcal{F}_{n-1}\right]+\eta_n(\boldsymbol{d}),\end{aligned}$$

where

$$\eta_n(\boldsymbol{d})=\mathbb{E}\left[M_n\left(1-\frac{s(\boldsymbol{d})-1}{2|E_{n-1}|}\right)^{M_n-1}-M_n\left(1-(M_n-1)\frac{s(\boldsymbol{d})-1}{2|E_{n-1}|}\right)\middle|\mathcal{F}_{n-1}\right],$$

which is not the same sequence as the η's from the previous section. Lemma 3 implies that for every $n\geq s(\boldsymbol{d})$ we have

$$\begin{aligned}&\left|M_n\left(1-\frac{s(\boldsymbol{d})-1}{2|E_{n-1}|}\right)^{M_n-1}-M_n\left(1-(M_n-1)\frac{s(\boldsymbol{d})-1}{2|E_{n-1}|}\right)\right|\\&\leq M_n\binom{M_n-1}{2}\left(\frac{s(\boldsymbol{d})-1}{2|E_{n-1}|}\right)^2.\end{aligned}$$

Combining this with assumption **(BA2)**, we obtain that

$$\begin{aligned}|\eta_n(\boldsymbol{d})|&\leq\mathbb{E}\left[\left|M_n\left(1-\frac{s(\boldsymbol{d})-1}{2|E_{n-1}|}\right)^{M_n-1}-M_n\left(1-(M_n-1)\frac{s(\boldsymbol{d})-1}{2|E_{n-1}|}\right)\right|\middle|\mathcal{F}_{n-1}\right]\\&\leq\mathbb{E}\left[M_n\binom{M_n-1}{2}\left(\frac{s(\boldsymbol{d})-1}{2|E_{n-1}|}\right)^2\middle|\mathcal{F}_{n-1}\right]=\left(\frac{s(\boldsymbol{d})-1}{2|E_{n-1}|}\right)^2\mathbb{E}\left[M_n\binom{M_n-1}{2}\right]\\&\leq\left(\frac{s(\boldsymbol{d})-1}{2|E_{n-1}|}\right)^2\mathbb{E}(M_n^3)=o\left(\frac{1}{n}\right)\ \text{a.s.}\end{aligned}$$

Getting back to equation (4), we conclude that

$$\begin{aligned}
p^{(n)}_{\boldsymbol{d}-\boldsymbol{e}_k}(\boldsymbol{e}_k) &= \left(\frac{d_k-1}{2|E_{n-1}|}\right)\mathbb{E}\left[M_n\left(1-\frac{s(\boldsymbol{d})-1}{2|E_{n-1}|}\right)^{M_n-1}\middle|\mathcal{F}_{n-1}\right] \\
&= \left(\frac{d_k-1}{2|E_{n-1}|}\right)\left(\mathbb{E}\left[M_n\left(1-(M_n-1)\frac{s(\boldsymbol{d})-1}{2|E_{n-1}|}\right)\middle|\mathcal{F}_{n-1}\right]+\eta_n(\boldsymbol{d})\right) \\
&= \left(\frac{d_k-1}{2|E_{n-1}|}\right)\left(\mathbb{E}(M_n)-\mathbb{E}\left[M_n(M_n-1)\frac{s(\boldsymbol{d})-1}{2|E_{n-1}|}\middle|\mathcal{F}_{n-1}\right]+\eta_n(\boldsymbol{d})\right) \\
&\sim \frac{d_k-1}{2}\cdot\frac{1}{n}+o\left(\frac{1}{n}\right) \text{ a.s.}
\end{aligned}$$

Therefore, for every $k\in[N]$ we have

$$\lim_{n\to\infty} np^{(n)}_{\boldsymbol{d}-\boldsymbol{e}_k}(\boldsymbol{e}_k) = r^{(k)}(\boldsymbol{d}-\boldsymbol{e}_k) = \frac{d_k-1}{2} \text{ a.s.}$$

Let $\boldsymbol{i}\in H'(\boldsymbol{d})$, i.e. $\forall k\in[N]: 0\le i_k\le d_k$ and $s(\boldsymbol{i})\ge 2$. In this case, we can bound the conditional expectation as follows:

$$\begin{aligned}
&p^{(n)}_{\boldsymbol{d}-\boldsymbol{i}}(\boldsymbol{i}) \\
&= \mathbb{E}\left[\frac{M_n!}{\left(\prod_{k=1}^N i_k!\right)(M_n-s(\boldsymbol{i}))!}\left[\prod_{k=1}^N\left(\frac{d_k-i_k}{2|E_{n-1}|}\right)^{i_k}\right]\left(1-\frac{s(\boldsymbol{d}-\boldsymbol{i})}{2|E_{n-1}|}\right)^{M_n-s(\boldsymbol{i})}\middle|\mathcal{F}_{n-1}\right] \\
&\le \prod_{k=1}^N\left(\frac{d_k-i_k}{2|E_{n-1}|}\right)^{i_k}\mathbb{E}\left[\frac{M_n!}{\left(\prod_{k=1}^N i_k!\right)(M_n-s(\boldsymbol{i}))!}\right] \le \frac{\prod_{k=1}^N(d_k-i_k)^{i_k}}{(2|E_{n-1}|)^{s(\boldsymbol{i})}}\mathbb{E}\left(M_n^{s(\boldsymbol{i})}\right).
\end{aligned}$$

This yields

$$\lim_{n\to\infty} np^{(n)}_{\boldsymbol{d}-\boldsymbol{i}}(\boldsymbol{i}) = 0 \text{ a.s.},$$

due to assumption **(BA2)** and the fact that $|E_{n-1}|\sim mn$.

Assumption (GM5). By the dynamics of the multi-type Barabási–Albert model, we conclude that for every $n \geq 1$ and $\boldsymbol{d} \in \mathbb{N}^N$ the following holds:

$$q^{(n)}(\boldsymbol{d}) = \mathbb{E}\left[I\left(M_n = s(\boldsymbol{d})\right) \frac{s(\boldsymbol{d})!}{\prod_{k=1}^N d_k!} \prod_{k=1}^N \left(\frac{\left|E_{n-1}^{(k)}\right|}{\left|E_{n-1}\right|} \right)^{d_k} \middle| \mathcal{F}_{n-1} \right]$$
$$= \mathbb{P}\left(M_n = s(\boldsymbol{d})\right) \frac{s(\boldsymbol{d})!}{\prod_{k=1}^N d_k!} \prod_{k=1}^N \left(\frac{\left|E_{n-1}^{(k)}\right|}{\left|E_{n-1}\right|} \right)^{d_k}.$$

Assumption **(BA1)** implies that $\mathbb{P}\left(M_n = s(\boldsymbol{d})\right) \to \mathbb{P}\left(M = s(\boldsymbol{d})\right)$ as $n \to \infty$. It follows from Lemma 2 that

$$q(\boldsymbol{d}) = \lim_{n\to\infty} q^{(n)}(\boldsymbol{d}) = \mathbb{P}\left(M = s(\boldsymbol{d})\right) \frac{s(\boldsymbol{d})!}{\prod_{k=1}^N d_k!} \prod_{k=1}^N \left(\zeta^{(k)}\right)^{d_k} \text{ a.s.}$$

This yields

$$u(\boldsymbol{d}) = \frac{s(\boldsymbol{d})}{2},$$
$$r^{(k)}(\boldsymbol{d} - \boldsymbol{e}_k) = \frac{d_k - 1}{2} \quad (\forall k \in [N])$$
$$q(\boldsymbol{d}) = \mathbb{P}\left(M = s(\boldsymbol{d})\right) \frac{s(\boldsymbol{d})!}{\prod_{k=1}^N d_k!} \prod_{k=1}^N \left(\zeta^{(k)}\right)^{d_k}.$$

Applying Theorem 1 we get Theorem 2. □

2.2.2. Model of independent edges. This model is an enhanced and a multi-type version of the models in [**24**] and [**38**], where the new vertex is connected to the existing ones independently, with probability depending on the edges of the actual vertex. In this model, instead of connecting with a single edge with a given probability, we add a Poisson number of new edges, with the multiplicative parameter chosen randomly.

In the model of independent edges, we have the following dynamics:

- for every $n \geq 1$, in the n^{th} step, the new vertex v_n attaches to all of the already existing vertices with some new edges of type k independently.

- For any existing vertex $w \in V_{n-1}$ let $\Delta_n^{(k)}(w)$ be the number edges of type k between the vertices v_n and w. We assume that, conditionally with respect to $\mathcal{F}_{n-1}$, for every $k \in [N]$ we have

$$\Delta_n^{(k)}(w) \sim \text{Poi}\left(\lambda_n \frac{\deg_{n-1}^{(k)}(w)}{2|E_{n-1}|}\right),$$

where λ_n is a positive random variable. We also assume that for every w, the random variables $\left(\Delta_n^{(k)}(w)\right)_{k=1}^{N}$ are conditionally independent of each other with respect to $\mathcal{F}_{n-1}$.

Remark. Since the Poisson distribution can be approximated by the binomial distribution we expect a similar behaviour from the two graphs models, although they are clearly different in some sense.

Let $\lambda_1, \lambda_2, \lambda_3, \ldots$ be a sequence of independent random variables. Similarly to the previously discussed model, we need a few assumptions on their distribution.

Assumption (IE1) For every $n \geq 1$ the random variable λ_n is positive and independent of $\mathcal{F}_{n-1}$.

Assumption (IE2) We assume that there exists a positive random variable, denoted by λ, such that $\lambda_n \to \lambda$ in distribution, and for every $p \geq 1$ we have $\mathbb{E}(\lambda_n^p) \to \mathbb{E}(\lambda^p) < \infty$ as $n \to \infty$. The expected value and the variance of λ will be denoted by $\mu = \mathbb{E}(\lambda)$ and $\sigma^2 = \text{Var}(\lambda)$, respectively.

For every $n \geq 1$ we define $\mathcal{F}_{n-1}^+ = \sigma(\mathcal{F}_{n-1}, \lambda_n)$. Let Δ_n be the number of new edges in the n^{th} step, and let $\Delta_n^{(k)}$ denote the number of new edges of type k in the n^{th} step. For every $n \geq 1$ we have $\Delta_n | \mathcal{F}_{n-1}^+ \sim \text{Poi}(\lambda_n)$, furthermore for every $k \in [N]$ we have

$$\Delta_n^{(k)} | \mathcal{F}_{n-1}^+ \sim \text{Poi}\left(\lambda_n \frac{|E_{n-1}^{(k)}|}{|E_{n-1}|}\right).$$

Note that the random variables $\left(\Delta_n^{(k)}\right)_{k=1}^{N}$ are conditionally independent given $\mathcal{F}_{n-1}^+$.

Again, we need the following lemma to understand the asymptotics of the proportion of edges of type k as the number of steps goes to infinity.

LEMMA 4. *For every $k \in [N]$ let us have $\hat{\zeta}_n^{(k)} = \frac{|E_n^{(k)}|}{|E_n|}$, i.e. the proportion of the number of edges of type k in the model of independent edges. For every $k \in [N]$ there exists a random variable $\hat{\zeta}^{(k)}$ such that $\hat{\zeta}_n^{(k)} \to \hat{\zeta}^{(k)}$ almost surely as $n \to \infty$.*

Proof of Lemma 4. Recall that for a fixed $k \in [N]$ we have

$$\Delta_n^{(k)} \sim \text{Poi}\left(\lambda_n \frac{|E_n^{(k)}|}{|E_n|}\right) \qquad \text{and} \qquad \Delta_n - \Delta_n^{(k)} \sim \text{Poi}\left(\lambda_n \left[1 - \frac{|E_n^{(k)}|}{|E_n|}\right]\right),$$

furthermore $\Delta_n^{(k)}$ and $\Delta_n - \Delta_n^{(k)}$ are conditionally independent given $\mathcal{F}_{n-1}$. Because of this, it is enough to prove this lemma for $N = 2$, which means there are only two types.

We are going to show that we have $\Delta_n^{(1)} \Big| \mathcal{F}_{n-1}^{+} \sim \text{Bin}\left(\Delta_n, \frac{|E_{n-1}^{(1)}|}{|E_{n-1}|}\right)$. For every $n \geq 1$ we define $\mathcal{F}_{n-1}^{++} = \sigma(\mathcal{F}_{n-1}^{+}, \Delta_n)$. For all $i \leq j$ the conditional distribution can be calculated as follows:

$$\mathbb{P}\left(\Delta_n^{(1)} = i \Big| \Delta_n = j, \mathcal{F}_{n-1}^{+}\right) = \frac{\mathbb{P}\left(\Delta_n^{(1)} = i, \Delta_n = j \Big| \mathcal{F}_{n-1}^{+}\right)}{\mathbb{P}\left(\Delta_n^{(1)} + \Delta_n^{(2)} = j \Big| \mathcal{F}_{n-1}^{+}\right)}$$

$$= \frac{\mathbb{P}\left(\Delta_n^{(1)} = i, \Delta_n^{(2)} = j - i \Big| \mathcal{F}_{n-1}^{+}\right)}{\mathbb{P}\left(\Delta_n^{(1)} + \Delta_n^{(2)} = j \Big| \mathcal{F}_{n-1}^{+}\right)} = \frac{\mathbb{P}\left(\Delta_n^{(1)} = i \Big| \mathcal{F}_{n-1}^{+}\right) \cdot \mathbb{P}\left(\Delta_n^{(2)} = j - i \Big| \mathcal{F}_{n-1}^{+}\right)}{\mathbb{P}\left(\Delta_n^{(1)} + \Delta_n^{(2)} = j \Big| \mathcal{F}_{n-1}^{+}\right)}$$

$$= \frac{\frac{\left(\lambda_n \frac{|E_{n-1}^{(1)}|}{|E_{n-1}|}\right)^i}{i!} \cdot \exp\left(-\lambda_n \frac{|E_{n-1}^{(1)}|}{|E_{n-1}|}\right) \cdot \frac{\left(\lambda_n \frac{|E_{n-1}^{(2)}|}{|E_{n-1}|}\right)^{j-i}}{(j-i)!} \cdot \exp\left(-\lambda_n \frac{|E_{n-1}^{(2)}|}{|E_{n-1}|}\right)}{\frac{\lambda_n^j}{j!} \cdot \exp\left(-\lambda_n\right)}$$

$$= \binom{j}{i} \left(\frac{|E_{n-1}^{(1)}|}{|E_{n-1}|}\right)^i \left(1 - \frac{|E_{n-1}^{(1)}|}{|E_{n-1}|}\right)^{j-i}.$$

For all $n \geq 1$, similarly to the proof of Lemma 2, we have

$$\mathbb{E}\left(\frac{|E_n^{(1)}|}{|E_n|} \Bigg| \mathcal{F}_{n-1}^{++}\right) = \frac{|E_{n-1}^{(1)}|}{|E_{n-1}| + \Delta_n} + \frac{\Delta_n \frac{|E_{n-1}^{(1)}|}{|E_{n-1}|}}{|E_{n-1}| + \Delta_n} = \frac{|E_{n-1}^{(1)}|}{|E_{n-1}|}.$$

Notice that E_{n-1} and $E^{(1)}_{n-1}$ are $\mathcal{F}_{n-1}$-measurable, which implies that

$$\mathbb{E}\left(\frac{|E^{(1)}_n|}{|E_n|}\middle|\mathcal{F}_{n-1}\right) = \frac{|E^{(1)}_{n-1}|}{|E_{n-1}|}.$$

We conclude that $\left(\hat{\zeta}^{(1)}_n, \mathcal{F}_n\right)_{n=1}^{\infty}$ is a non-negative martingale, thus it is convergent almost surely. Let $\hat{\zeta}^{(1)} \geq 0$ be its limit. The proof of Lemma 4 is complete. $\square$

Asymptotic degree distribution in the model of independent edges.

THEOREM 3. *If the assumptions **(IE1)** and **(IE2)** on the sequence $(\lambda_n)_{n=1}^{\infty}$ are satisfied, then in the model of independent edges for every $\boldsymbol{d} = (d_1, \ldots, d_N)^T \in \mathbb{N}^N$ we have*

$$\lim_{n\to\infty} \frac{X_n(\boldsymbol{d})}{|V_n|} = x(\boldsymbol{d}) \text{ a.s.}$$

The random variables $\left\{x(\boldsymbol{d}), \boldsymbol{d} \in \mathbb{N}^N\right\}$ satisfy the following recurrence equations:

$$x(\boldsymbol{d}) = \sum_{k=1}^{N} \frac{d_k - 1}{D+2} x(\boldsymbol{d} - \boldsymbol{e}_k) + \frac{2}{D+2} \frac{\prod_{k=1}^{N} \left(\hat{\zeta}^{(k)}\right)^{d_k}}{\prod_{k=1}^{N} d_k!} \mathbb{E}\left(\lambda^D e^{-\lambda}\right),$$

where $\hat{\zeta}^{(k)}$ is defined in Lemma 4 and $D = s(\boldsymbol{d}) = \sum_{k=1}^{N} d_k$.

REMARK. For the calculation of the last term we can use the following observation. Let us denote by g_λ the moment generating function of the random variable λ, i.e. $g_\lambda(t) = \mathbb{E}(e^{t\lambda})$, where $t \in \mathbb{R}$. Let us have $B = \{t \in \mathbb{R} : g_\lambda(t) < \infty\}$, i.e. the set of finiteness of g_λ, and let B_0 be the interior of B. Suppose that $-1 \in B_0$. It is well known that in this case $g_\lambda(t)$ is infinitely differentiable at $t = -1$, furthermore, we have

$$g^{(D)}_\lambda(-1) = \mathbb{E}\left(\lambda^D e^{-\lambda}\right),$$

where $D = s(\boldsymbol{d})$ and $g^{(D)}_\lambda$ is the derivative of order D of the moment generating function g_λ.

Preliminaries for the proof of Theorem 3. We will use the following lemma.

LEMMA 5. *For the number of edges of the model of independent edges, we have the following asymptotics:* $|E_n| \sim \mu n$.

PROOF. Let us have $\Delta_0 = |E_0|$ and $\lambda_0 = 0$. We define the following process:

$$Z_n = \sum_{i=0}^{n} \Delta_i - \lambda_i = |E_n| - \sum_{i=1}^{n} \lambda_i.$$

We show that $(Z_n, \mathcal{F}_n)_{n=1}^{\infty}$ is a square integrable martingale, i.e. $(Z_n, \mathcal{F}_n)_{n=1}^{\infty}$ is a martingale, and we have $\mathbb{E}(Z_n^2) < \infty$ for every $n \geq 1$.

For every $n \geq 1$ we have

$$\begin{aligned}\mathbb{E}(Z_n|\mathcal{F}_{n-1}) &= \mathbb{E}(Z_{n-1} + \Delta_n - \lambda_n|\mathcal{F}_{n-1}) = \\ &= Z_{n-1} + \mathbb{E}\left[\mathbb{E}(\Delta_n|\mathcal{F}_{n-1}^{+}) - \lambda_n|\mathcal{F}_{n-1}\right] = Z_{n-1},\end{aligned}$$

by using the fact that $\Delta_n|\mathcal{F}_{n-1}^{+} \sim \mathrm{Poi}(\lambda_n)$.

Furthermore, we can bound the expectation of the squares in the following way:

$$\begin{aligned}\mathbb{E}(Z_n^2) &= \mathbb{E}\left[\left(\sum_{i=1}^{n} \Delta_i - \lambda_i\right)^2\right] = \mathbb{E}\left[\sum_{i=1}^{n}(\Delta_i - \lambda_i)^2 + 2\sum_{i<j}(\Delta_i - \lambda_i)(\Delta_j - \lambda_j)\right] \\ &= \sum_{i=1}^{n} \mathbb{E}\left[(\Delta_i - \lambda_i)^2\right] + 2\sum_{i<j} \mathbb{E}\left[(\Delta_i - \lambda_i)(\Delta_j - \lambda_j)\right] \\ &= \sum_{i=1}^{n} \mathbb{E}\left(\mathbb{E}\left[(\Delta_i - \lambda_i)^2 \middle| \mathcal{F}_{i-1}^{+}\right]\right) + 2\sum_{i<j} \mathbb{E}\left(\mathbb{E}\left[(\Delta_i - \lambda_i)(\Delta_j - \lambda_j) \middle| \mathcal{F}_{j-1}^{+}\right]\right) \\ &= \sum_{i=1}^{n} \mathbb{E}(\lambda_i) < \infty,\end{aligned}$$

hence $(Z_n, \mathcal{F}_n)_{n=1}^{\infty}$ is a square integrable martingale. The increasing process associated with Z_n^2 by the Doob decomposition is the following:

$$\begin{aligned}
A_n &= \sum_{i=1}^{n} \mathrm{Var}\left(\Delta_i | \mathcal{F}_{i-1}\right) = \sum_{i=1}^{n} \mathbb{E}(\Delta_i^2 | \mathcal{F}_{i-1}) - \mathbb{E}^2(\Delta_i | \mathcal{F}_{i-1}) \\
&= \sum_{i=1}^{n} \mathbb{E}\left[\mathbb{E}(\Delta_i^2 | \mathcal{F}_{i-1}^{+}) \Big| \mathcal{F}_{i-1}\right] - \mathbb{E}^2\left[\mathbb{E}(\Delta_i | \mathcal{F}_{i-1}^{+}) \Big| \mathcal{F}_{i-1}\right] \\
&= \sum_{i=1}^{n} \mathbb{E}(\lambda_i^2 + \lambda_i) - \mathbb{E}^2(\lambda_i) \\
&= \sum_{i=1}^{n} \mathrm{Var}(\lambda_i) + \mathbb{E}(\lambda_i) \le n(\mu + \sigma^2).
\end{aligned}$$

By using [**43**], Proposition VII-2-4, we conclude that $|E_n| = \left(\sum_{i=1}^{n} \lambda_i\right) n + o\left(n^{1/2+\varepsilon}\right)$ almost surely as $n \to \infty$ on the event $\{A_n \to \infty\}$ for all $\varepsilon > 0$.

For every $n \ge 1$ we have $|E_n| = \sum_{k=1}^{N} \left|E_0^{(k)}\right| + \sum_{i=1}^{n} \Delta_i$. By the assumptions of the model, the sequence $(\lambda_i)_{i=1}^{n}$ satisfies the following conditions:

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}(\lambda_i) = \mathbb{E}(\lambda) = \mu \quad \text{and} \quad \sum_{n=1}^{\infty} \frac{\mathrm{Var}(\lambda_n)}{n^2} < \infty.$$

Therefore, Kolmogorov's theorem can be applied (Theorem 6.7. in [**45**]) for the sequence $(\lambda_n)_{n=1}^{\infty}$, similarly to the previous sections. We conclude that we have $|E_n| \sim \mu n$. □

Proof of Theorem 3. We will use Theorem 1, so we have to check the assumptions of the general model.

Assumption (GM1). By the definition of the dynamics of the model of independent edges, it is easy to see that Assumption **(GM1)** trivially holds.

Assumption (GM2). By using the fact that $\Delta_n | \mathcal{F}_{n-1}^{+} \sim \mathrm{Poi}(\lambda_n)$, we obtain that for every $\boldsymbol{d} \in \mathbb{N}^N$ we have

$$\begin{aligned}
&\mathbb{E}\left[\left|X_n(\boldsymbol{d}) - X_{n-1}(\boldsymbol{d})\right|^2 \Big| \mathcal{F}_{n-1}\right] \le \mathbb{E}(\Delta_n^2 | \mathcal{F}_{n-1}) = \mathbb{E}\left[\mathbb{E}(\Delta_n^2 | \mathcal{F}_{n-1}^{+}) \Big| \mathcal{F}_{n-1}\right] \\
&= \mathbb{E}(\lambda_n^2 + \lambda_n | \mathcal{F}_{n-1}) = \mathbb{E}(\lambda_n^2 + \lambda_n) \to \sigma^2 + \mu^2 + \mu < \infty
\end{aligned}$$

as $n \to \infty$. If we choose $\delta = 1$, then assumption **(GM2)** is satisfied.

Assumption (GM3). For every $n \geq 1$ and $\boldsymbol{d} \in \mathbb{N}^N$ we have

$$p_{\boldsymbol{d}}^{(n)}(\mathbf{0}) = \mathbb{E}\left[\exp\left(-\lambda_n \frac{s(\boldsymbol{d})}{2|E_{n-1}|}\right)\middle|\mathcal{F}_{n-1}\right].$$

We will use Taylor expansion. In order to do this, we write the expectation in the following form:

$$\mathbb{E}\left[\exp\left(-\lambda_n \frac{s(\boldsymbol{d})}{2|E_{n-1}|}\right)\middle|\mathcal{F}_{n-1}\right] = \mathbb{E}\left(1 - \lambda_n \frac{s(\boldsymbol{d})}{2|E_{n-1}|}\middle|\mathcal{F}_{n-1}\right) + \eta_n(\boldsymbol{d}),$$

where

$$\eta_n(\boldsymbol{d}) = \mathbb{E}\left[\exp\left(-\lambda_n \frac{s(\boldsymbol{d})}{2|E_{n-1}|}\right) - \left(1 - \lambda_n \frac{s(\boldsymbol{d})}{2|E_{n-1}|}\right)\middle|\mathcal{F}_{n-1}\right].$$

It is well known that for all $x \geq 0$ we have $|e^{-x} - (1-x)| \leq \frac{x^2}{2}$, which implies that

$$\left|\exp\left(-\lambda_n \frac{s(\boldsymbol{d})}{2|E_{n-1}|}\right) - \left(1 - \lambda_n \frac{s(\boldsymbol{d})}{2|E_{n-1}|}\right)\right| \leq \frac{1}{2}\left(\lambda_n \frac{s(\boldsymbol{d})}{2|E_{n-1}|}\right)^2.$$

By using the above inequality, we obtain that

$$\begin{aligned}
|\eta_n(\boldsymbol{d})| &\leq \mathbb{E}\left[\left|\exp\left(-\lambda_n \frac{s(\boldsymbol{d})}{2|E_{n-1}|}\right) - \left(1 - \lambda_n \frac{s(\boldsymbol{d})}{2|E_{n-1}|}\right)\right|\middle|\mathcal{F}_{n-1}\right] \\
&\leq \mathbb{E}\left[\frac{1}{2}\left(\lambda_n \frac{s(\boldsymbol{d})}{2|E_{n-1}|}\right)^2\middle|\mathcal{F}_{n-1}\right] = \mathbb{E}\left[\lambda_n^2 \frac{(s(\boldsymbol{d}))^2}{8|E_{n-1}|^2}\middle|\mathcal{F}_{n-1}\right] \\
&= \mathbb{E}(\lambda_n^2)\frac{(s(\boldsymbol{d}))^2}{8|E_{n-1}|^2} = o\left(\frac{1}{n}\right) \quad \text{a.s.}
\end{aligned}$$

by the assumption **(IE2)** and $|E_{n-1}| \sim \mu n$. The definition of $u_n(\boldsymbol{d})$ and $\eta_n(\boldsymbol{d})$ implies that

$$u_n(\boldsymbol{d}) = n\left(\frac{s(\boldsymbol{d})}{2} \cdot \frac{\mathbb{E}(\lambda_n)}{|E_{n-1}|} - \eta_n(\boldsymbol{d})\right).$$

We can see that $u_n(\boldsymbol{d})$ is $\mathcal{F}_{n-1}$-measurable, hence $(u_n(\boldsymbol{d}))_{n=1}^{\infty}$ is a predictable process with respect to $\mathcal{F}$. Recall that $|E_{n-1}| \sim \mu n$, and $n \cdot \eta_n(\boldsymbol{d}) = o(1)$ almost surely.

Assumption **(IE2)** implies that

$$u(\boldsymbol{d}) = \lim_{n\to\infty} u_n(\boldsymbol{d}) = \frac{s(\boldsymbol{d})}{2} \text{ a.s.}$$

Assumption (GM4). First, let us fix $k \in [N]$. For every $n \geq 1$ and $\boldsymbol{d} \in \mathbb{N}^N$, where $s(\boldsymbol{d}) \geq 1$, we have

$$\begin{aligned} (5) \qquad p^{(n)}_{\boldsymbol{d}-\boldsymbol{e}_k}(\boldsymbol{e}_k) &= \mathbb{E}\left[\lambda_n \frac{d_k - 1}{2|E_{n-1}|} \cdot \exp\left(-\lambda_n \frac{d_k - 1}{2|E_{n-1}|}\right) \middle| \mathcal{F}_{n-1}\right] \\ &= \frac{d_k - 1}{2|E_{n-1}|} \mathbb{E}\left[\lambda_n \cdot \exp\left(-\lambda_n \frac{d_k - 1}{2|E_{n-1}|}\right) \middle| \mathcal{F}_{n-1}\right]. \end{aligned}$$

Similarly to the previous case, we obtain that

$$\mathbb{E}\left[\lambda_n \cdot \exp\left(-\lambda_n \frac{d_k - 1}{2|E_{n-1}|}\right) \middle| \mathcal{F}_{n-1}\right] = \mathbb{E}\left[\lambda_n \left(1 - \lambda_n \frac{d_k - 1}{2|E_{n-1}|}\right) \middle| \mathcal{F}_{n-1}\right] + \eta_n(\boldsymbol{d}),$$

where

$$\eta_n(\boldsymbol{d}) = \mathbb{E}\left[\lambda_n \cdot \exp\left(-\lambda_n \frac{d_k - 1}{2|E_{n-1}|}\right) - \lambda_n \left(1 - \lambda_n \frac{d_k - 1}{2|E_{n-1}|}\right) \middle| \mathcal{F}_{n-1}\right].$$

Again, by using $|e^{-x} - (1 - x)| \leq \frac{x^2}{2}$ for all $x \geq 0$, we conclude that

$$\left|\lambda_n \cdot \exp\left(-\lambda_n \frac{d_k - 1}{2|E_{n-1}|}\right) - \lambda_n \left(1 - \lambda_n \frac{d_k - 1}{2|E_{n-1}|}\right)\right| \leq \frac{\lambda_n}{2} \left(\lambda_n \frac{d_k - 1}{2|E_{n-1}|}\right)^2.$$

Combining this with assumption **(IE2)**, we obtain that

$$\begin{aligned} |\eta_n(\boldsymbol{d})| &\leq \mathbb{E}\left[\left|\lambda_n \cdot \exp\left(-\lambda_n \frac{d_k - 1}{2|E_{n-1}|}\right) - \lambda_n \left(1 - \lambda_n \frac{d_k - 1}{2|E_{n-1}|}\right)\right| \middle| \mathcal{F}_{n-1}\right] \\ &\leq \mathbb{E}\left[\frac{\lambda_n}{2} \left(\lambda_n \frac{d_k - 1}{2|E_{n-1}|}\right)^2 \middle| \mathcal{F}_{n-1}\right] = \mathbb{E}\left[\lambda_n^3 \frac{(d_k - 1)^2}{8|E_{n-1}|^2} \middle| \mathcal{F}_{n-1}\right] \\ &= \mathbb{E}(\lambda_n^3) \frac{(d_k - 1)^2}{8|E_{n-1}|^2} = o\left(\frac{1}{n}\right) \text{ a.s.} \end{aligned}$$

By using this we conclude that

$$\begin{aligned}
p^{(n)}_{\boldsymbol{d}-\boldsymbol{e}_k}(\boldsymbol{e}_k) &= \frac{d_k-1}{2|E_{n-1}|}\mathbb{E}\left[\lambda_n \cdot \exp\left(-\lambda_n \frac{d_k-1}{2|E_{n-1}|}\right)\middle|\mathcal{F}_{n-1}\right] \\
&= \frac{d_k-1}{2|E_{n-1}|}\left(\mathbb{E}\left[\lambda_n\left(1-\lambda_n\frac{d_k-1}{2|E_{n-1}|}\right)\middle|\mathcal{F}_{n-1}\right]+\eta_n(\boldsymbol{d})\right) \\
&\sim \frac{d_k-1}{2}\cdot\frac{1}{n}+o\left(\frac{1}{n}\right) \text{ a.s.}
\end{aligned}$$

Putting this together, we obtain that for every $k \in [N]$ we have

$$\lim_{n\to\infty} np^{(n)}_{\boldsymbol{d}-\boldsymbol{e}_k}(\boldsymbol{e}_k) = r^{(k)}(\boldsymbol{d}-\boldsymbol{e}_k) = \frac{d_k-1}{2} \text{ a.s.}$$

Now let $\boldsymbol{i} \in H'(\boldsymbol{d})$, i.e. $\forall k \in [N] : 0 \le i_k \le d_k$ and $s(\boldsymbol{i}) \ge 2$. For every $n \ge 1$ we have

$$\begin{aligned}
p^{(n)}_{\boldsymbol{d}-\boldsymbol{i}}(\boldsymbol{i}) &= \mathbb{E}\left[\prod_{k=1}^{N}\frac{1}{i_k!}\left(\lambda_n\frac{d_k-i_k}{2|E_{n-1}|}\right)^{i_k}\exp\left(-\lambda_n\frac{d_k-i_k}{2|E_{n-1}|}\right)\middle|\mathcal{F}_{n-1}\right] \\
&= \frac{\prod_{k=1}^{N}(d_k-i_k)^{i_k}}{(2|E_{n-1}|)^{s(\boldsymbol{i})}\prod_{k=1}^{N} i_k!}\mathbb{E}\left[\lambda_n^{s(\boldsymbol{i})}\cdot\exp\left(-\lambda_n\frac{s(\boldsymbol{d}-\boldsymbol{i})}{2|E_{n-1}|}\right)\middle|\mathcal{F}_{n-1}\right] \\
&\le \frac{\prod_{k=1}^{N}(d_k-i_k)^{i_k}}{(2|E_{n-1}|)^{s(\boldsymbol{i})}\prod_{k=1}^{N} i_k!}\mathbb{E}\left(\lambda_n^{s(\boldsymbol{i})}\right),
\end{aligned}$$

which implies that

$$\lim_{n\to\infty} np^{(n)}_{\boldsymbol{d}-\boldsymbol{i}}(\boldsymbol{i}) = 0 \text{ a.s.}$$

Assumption (GM5). By the dynamics of the model, for every $n \geq 1$ and $\boldsymbol{d} \in \mathbb{N}^N$, the following holds:

$$\begin{aligned}
q^{(n)}(\boldsymbol{d}) &= \mathbb{E}\left[\mathbb{P}\left(\bigcap_{k=1}^{N}\left\{\Delta_n^{(k)} = d_k\right\}\middle|\mathcal{F}_{n-1}^{+}\right)\middle|\mathcal{F}_{n-1}\right] \\
&= \mathbb{E}\left[\prod_{k=1}^{N}\mathbb{P}\left(\Delta_n^{(k)} = d_k\middle|\mathcal{F}_{n-1}^{+}\right)\middle|\mathcal{F}_{n-1}\right] \\
&= \mathbb{E}\left[\prod_{k=1}^{N}\frac{1}{d_k!}\left(\lambda_n\frac{\left|E_{n-1}^{(k)}\right|}{|E_{n-1}|}\right)^{d_k}\exp\left(-\lambda_n\frac{\left|E_{n-1}^{(k)}\right|}{|E_{n-1}|}\right)\middle|\mathcal{F}_{n-1}\right] \\
&= \frac{1}{\prod_{k=1}^{N} d_k!}\prod_{k=1}^{N}\left(\frac{\left|E_{n-1}^{(k)}\right|}{|E_{n-1}|}\right)^{d_k}\mathbb{E}\left(\lambda_n^{s(\boldsymbol{d})}\cdot\exp\left(-\lambda_n\right)\middle|\mathcal{F}_{n-1}\right).
\end{aligned}$$

By Lemma 4 and the independence of λ_n and $\mathcal{F}_{n-1}$, we have

$$\begin{aligned}
q(\boldsymbol{d}) = \lim_{n\to\infty} q^{(n)}(\boldsymbol{d}) &= \lim_{n\to\infty}\frac{\prod_{k=1}^{N}\left(\hat{\zeta}_{n-1}^{(k)}\right)^{d_k}}{\prod_{k=1}^{N} d_k!}\mathbb{E}\left(\lambda_n^{s(\boldsymbol{d})}e^{-\lambda_n}\right) \\
&= \frac{\prod_{k=1}^{N}\left(\hat{\zeta}^{(k)}\right)^{d_k}}{\prod_{k=1}^{N} d_k!}\mathbb{E}\left(\lambda^{s(\boldsymbol{d})}e^{-\lambda}\right) \text{ a.s.,}
\end{aligned}$$

since the function $t^{s(\boldsymbol{d})}e^{-t}$ is bounded and continuous and $\lambda_n \to \lambda$ in distribution.

We obtain that

$$\begin{aligned}
u(\boldsymbol{d}) &= \frac{s(\boldsymbol{d})}{2}, \\
r^{(k)}(\boldsymbol{d} - \boldsymbol{e}_k) &= \frac{d_k - 1}{2} \quad (\forall k \in [N]) \\
q(\boldsymbol{d}) &= \frac{\prod_{k=1}^{N}\left(\hat{\zeta}^{(k)}\right)^{d_k}}{\prod_{k=1}^{N} d_k!}\mathbb{E}\left(\lambda^{s(\boldsymbol{d})}e^{-\lambda}\right).
\end{aligned}$$

Applying Theorem 1 we get Theorem 3. □

2.3. Scale-free property of random graphs in the multi-type case

A scale-free graph model is a random graph whose degree distribution follows a power law, i.e. the proportion of vertices with degree d asymptotically equals to $Cd^{-\gamma}$, where $\gamma > 0$ is a deterministic constant. It is well known that many large real networks have this property, see e.g. [**37**], although there are discussions about how common they are [**20**].

The formal definition of scale-free property of random graphs with no types is the following.

Definition 5. We assume that the proportion of vertices with degree d converges to a deterministic constant c_d a.s. for all $d \geq 0$, and the sum of the sequence $(c_d)_{d=0}^{\infty}$ equals to 1. In this case the sequence $(c_d)_{d=0}^{\infty}$ is an asymptotic degree distribution. Furthermore, if $c_d d^{\gamma} \to C$ as $d \to \infty$ holds with some positive C, then the model has the scale-free property, and γ is the so-called characteristic exponent.

In the following section we generalize the scale-free property for graphs with multi-type edges. First, let us define the (asymptotic) marginal degree distribution for a given type in a multi-type graph model.

Definition 6. Let us have a sequence of N-type random graphs, denoted by $(G_n(V_n, E_n))_{n=0}^{\infty}$. The marginal degree distribution for the edges of type $k \in [N]$ is defined as the following sequence of possibly random variables:

$$\left(x_l^{(k)}(n)\right)_{l=0}^{\infty} = \left(\frac{X_n^{(k)}(l)}{|V_n|}\right)_{l=0}^{\infty},$$

where we have

$$X_n^{(k)}(l) = \left| \left\{ v \in V_n : \deg_n^{(k)}(v) = l \right\} \right|,$$

i.e. the number of vertices in G_n with l edges of type k connected to them.

The asymptotic marginal degree distribution for the edges of type $k \in [N]$ is defined as

$$x_l^{(k)} = \lim_{n\to\infty} \frac{X_n^{(k)}(l)}{|V_n|} \text{ a.s.},$$

where $l \geq 0$.

Definition 7. Let us have a sequence of N-type random graphs, denoted by $(G_n(V_n, E_n))_{n=0}^{\infty}$. If for every $k \in [N]$ the marginal distribution of the edges of type k has scale-free distribution, moreover, the degree distribution of the graph without taking into consideration the types of the edges also has the scale-free property (in the single-type sense), then we say that $(G_n(V_n, E_n))_{n=0}^{\infty}$ has the scale-free property in the multi-type sense.

We are going to use the following theorem in order to prove that the generalized Barabási–Albert random graph and the model of independent edges have the scale-free property.

Theorem A (Theorem 1 in [**8**]). *Consider the following recurrence equation:*

$$x_n = \sum_{j=1}^{n-1} w_{n,j} x_{n-j} + r_n, \qquad w_{n,j} = a_j + \frac{b_j}{n} + c_{n,j}, \qquad (n = 1, 2, 3, \dots),$$

where $w_{n,j} \geq 0$, and a_n, b_n, $c_{n,j}$, r_n satisfy the following conditions.

- **(r1):** *$a_n \geq 0$ for $n \geq 1$, and the greatest common divisor of the set $\{n : a_n > 0\}$ is 1;*
- **(r2):** *$r_n \geq 0$, and there exists such an n that $r_n > 0$;*
- **(r3):** *there exists $z > 0$ such that*

$$1 < \sum_{n=1}^{\infty} a_n z^n < \infty, \qquad \sum_{n=1}^{\infty} |b_n| z^n < \infty,$$

$$\sum_{n=1}^{\infty} \sum_{j=1}^{\infty} |c_{n,j}| z^j < \infty, \qquad \sum_{n=1}^{\infty} r_n z^n < \infty.$$

Suppose that the sequence $(x_n)_{n=1}^{\infty}$ *satisfies the recurrence equation, conditions* $(r1)$-$(r3)$ *hold, and* $(x_n)_{n=1}^{\infty}$ *has infinitely many positive terms. Then* $x_n n^{-\gamma} q^n \to C$ *as* $n \to \infty$*, where* C *is a positive constant,* q *is the positive solution of equation* $\sum_{n=1}^{\infty} a_n q^n = 1$*, and*

$$\gamma = \frac{\sum_{n=1}^{\infty} b_n q^n}{\sum_{n=1}^{\infty} n a_n q^n}.$$

2.3.1. Scale-free property of the generalized Barabási–Albert random graph. In addition to the assumptions in Section 2.2.1, we require the following assumption:

Assumption (BA+): we assume that $(M_n)_{n=1}^{\infty}$ is a sequence of identically distributed random variables and there exists $z > 1$ such that we have

$$\sum_{l=1}^{\infty} \frac{\mathbb{E}(M_1^l)}{l \cdot l!} z^l < \infty.$$

The above assumption is trivially fulfilled if we have $\sup_l \mathbb{E}(M_1^l) < \infty$.

First, let us fix $k \in [N]$. Recall that in every step the endpoints of the new edges are chosen independently of each other and the degrees of the existing vertices are not updated until the end of the step. By using this, we conclude that for every $l \geq 0$ the change in the value of $X_n^{(k)}(l)$ only depends on the edges of type k, thus we have

$$\begin{aligned}(6)\qquad \mathbb{E}\left[X_n^{(k)}(l)|\mathcal{F}_{n-1}\right] &= X_{n-1}^{(k)}(l)\mathbb{E}\left[\left(1-\frac{l}{2|E_{n-1}|}\right)^{M_n}\middle|\mathcal{F}_{n-1}\right] \\ &+\sum_{i=1}^{l-1} X_{n-1}^{(k)}(l-i)\cdot\mathbb{E}\left[\binom{M_n}{i}\left(\frac{l-i}{2|E_{n-1}|}\right)^i\left(1-\frac{l-i}{2|E_{n-1}|}\right)^{M_n-i}\middle|\mathcal{F}_{n-1}\right] \\ &+\mathbb{E}\left[\binom{M_n}{k}\left(\zeta_{n-1}^{(k)}\right)^l\left(1-\zeta_{n-1}^{(k)}\right)^{M_n-l}\middle|\mathcal{F}_{n-1}\right],\end{aligned}$$

where $\zeta_{n-1}^{(k)}$ is the proportion of edges of type k in G_{n-1}. By using Lemma 1 and the same arguments as in the proof of Theorem 2, we can show that $x_l^{(k)}$ exists for all l and we can find the recurrence equations for the asymptotic degree distribution.

The only part which is different compared to the previous sections is finding the almost sure limit of the last term in equation (6) as the number of steps goes to infinity. Since M_n is independent of $\mathcal{F}_{n-1}$ and $\zeta_{n-1}^{(k)}$ is measurable with respect to $\mathcal{F}_{n-1}$, we have

$$(7)\quad \mathbb{E}\left[\binom{M_n}{l}\left(\zeta_{n-1}^{(k)}\right)^l\left(1-\zeta_{n-1}^{(k)}\right)^{M_n-l}\Big|\mathcal{F}_{n-1}\right]=\mathbb{E}\left[\binom{M_n}{l}t^l(1-t)^{M_n-l}\right]\Bigg|_{t=\zeta_{n-1}^{(k)}}.$$

Recall that $(M_n)_{n=1}^{\infty}$ is a sequence of identically distributed random variables, thus we define $f(t)=\mathbb{E}\left[\binom{M_1}{l}t^l(1-t)^{M_1-l}\right]$, where $t\in[0,1]$.

In order to show that $f(t)$ is a continuous function on $[0,1]$, we will use the Weierstrass M-test.

THEOREM 4 (Weierstrass M-test). *Suppose that $(\varphi_n)_{n=1}^{\infty}$ is a sequence of real-valued functions defined on a set denoted by T, furthermore there is a sequence of non-negative numbers denoted by $(\alpha_n)_{n=1}^{\infty}$ satisfying the following conditions:*

- *for every $n\geq 1$ and $t\in T$ we have $|\varphi_n(t)|\leq\alpha_n$ and*
- *$\sum_{n=1}^{\infty}M_n$ converges.*

Then the series defined as $\sum_{n=1}^{\infty}\varphi_n(t)$ converges absolutely and uniformly on T.

In many cases, the Weierstrass M-test is used in combination with the well-known uniform limit theorem. If we also assume that T is a topological space and the functions φ_n are continuous on T, then $\sum_{n=1}^{\infty}\varphi_n(t)$ is an absolutely continuous function on T.

For all $t\in[0,1]$, we have

$$\begin{aligned}f(t)&=\mathbb{E}\left[\binom{M_1}{l}t^l(1-t)^{M_1-l}\right]=\sum_{i=l}^{\infty}\binom{i}{l}t^l(1-t)^{i-l}\mathbb{P}(M_1=i)\\&\leq\sum_{i=l}^{\infty}\binom{i}{l}\mathbb{P}(M_1=i)=\mathbb{E}\left[\binom{M_1}{l}\right]\leq\mathbb{E}\left(M_1^l\right)<\infty,\end{aligned}$$

by the assumption **(BA2)**, thus f is continuous. Since $\zeta_n^{(k)} \to \zeta^{(k)}$ almost surely as $n \to \infty$ and f is continuous, equation (7) implies that

$$\lim_{n\to\infty} \mathbb{E}\left[\binom{M_n}{l}\left(\zeta_{n-1}^{(k)}\right)^l\left(1-\zeta_{n-1}^{(k)}\right)^{M_n-l}\Big|\mathcal{F}_{n-1}\right] = \mathbb{E}\left[\binom{M_1}{l}\left(\zeta^{(k)}\right)^l\left(1-\zeta^{(k)}\right)^{M_1-l}\right] \text{ a.s.}$$

By using Lemma 1, for every $l \geq 0$ we have

$$x_l^{(k)} = \frac{l-1}{l+2}x_{l-1}^{(k)} + \frac{2}{l+2}\mathbb{E}\left[\binom{M_1}{l}\left(\zeta^{(k)}\right)^l\left(1-\zeta^{(k)}\right)^{M_1-l}\right]. \tag{8}$$

We are going to apply Theorem A. Notice that the last equation can be written as

$$x_l^{(k)} = \left[1 - \frac{3}{l} + \left(\frac{3}{l} - \frac{3}{l+2}\right)\right]x_{l-1}^{(k)} + \frac{2}{l+2}\mathbb{E}\left[\binom{M_1}{l}\left(\zeta^{(k)}\right)^l\left(1-\zeta^{(k)}\right)^{M_1-l}\right].$$

If we choose the following parameters as described below:

- $a_1 = 1$ and $a_j = 0$ $(j \geq 2)$,
- $b_1 = -3$ and $b_j = 0$ $(j \geq 2)$,
- $c_{l,1} = \frac{3}{l} - \frac{3}{l+2}$ and $c_{l,j} = 0$ $(j \geq 2)$,
- $r_l^{(k)} = \frac{2}{l+2}\mathbb{E}\left[\binom{M_1}{l}\left(\zeta^{(k)}\right)^l\left(1-\zeta^{(k)}\right)^{M_1-l}\right]$,

then the assumptions (**r1**) and (**r3**) of Theorem A are fulfilled by also using Assumption (**BA+**). We know that there exists $l > 0$ such that $\mathbb{P}(M_1 = l) > 0$. By using Lemma 2, we conclude that $\zeta^{(k)}|M_1 = l$ is positive with positive probability thus $r_l^{(k)} > 0$ and the assumption (**r2**) is satisfied.

Remark. If in addition to the assumptions in Section 2.2.1 we also assume that $M_i \equiv M$ for all $i \geq 1$ where M is a positive integer, then in this case the proportion of edges of type k has an absolutely continuous almost sure limit (see e.g. Theorem 3 in [**21**]), thus none of the types die out asymptotically with probability one.

By using Theorem A, we conclude that for every $k \in [N]$ we have

$$x_l^{(k)} l^3 \to C_k$$

as $l \to \infty$ for some positive C_k, thus the characteristic exponent equals to 3.

Finally, for all $d \geq 0$ we define $Z_n(d) = \left| \left\{ v \in V_n : \sum_{k=1}^{N} \deg_n^{(k)}(v) = d \right\} \right|$, i.e. the number of vertices in G_n with d edges connected to them. This way we get back to the single-type graph models. The asymptotic degree distribution is denoted by $(z_d)_{d=0}^{\infty}$, where z_d is defined as the almost sure limit of the sequence $\left(\frac{Z_n(d)}{|V_n|}\right)_{n=0}^{\infty}$ as $n \to \infty$. For every $d \geq 0$ we have

$$\mathbb{E}\left[Z_n(d)|\mathcal{F}_{n-1}\right] = Z_{n-1}(d)\mathbb{E}\left[\left(1 - \frac{d}{2|E_{n-1}|}\right)^{M_n} \middle| \mathcal{F}_{n-1}\right]$$
$$+ \sum_{i=1}^{d-1} Z_{n-1}(k-i)\mathbb{E}\left[\binom{M_n}{i}\left(\frac{d-i}{2|E_{n-1}|}\right)^{i}\left(1 - \frac{d-i}{2|E_{n-1}|}\right)^{M_n - i} \middle| \mathcal{F}_{n-1}\right]$$
$$+ \mathbb{P}\left(M_n = d|\mathcal{F}_{n-1}\right).$$

By using the same argument as in the previous section, we have

$$z_d d^3 \to C$$

as $d \to \infty$ for some positive C, thus the characteristic exponent equals to 3. This provides a generalization on some of the results of preferential attachment models (see e.g. [**37**]). As the calculation above shows, this model fits into the general framework of [**30**] or [**44**] for single-type preferential attachment random graphs.

2.3.2. Scale-free property of the model of independent edges. In the model of independent edges we can use the same arguments as in the previous section. In addition to the assumptions in Section 2.2.2, we have the following additional assumption:

Assumption (IE+): we also assume that $(\lambda_n)_{n=1}^{\infty}$ is a sequence of identically distributed random variables and there exists $z > 1$ such that we have

$$\sum_{l=1}^{\infty} \frac{\mathbb{E}(\lambda_1^l)}{l \cdot l!} z^l < \infty.$$

In the model of independent edges for every type k we have

$$r_l^{(k)} = \frac{2}{l+2}\mathbb{E}\left[\frac{\left(\lambda_1\hat{\zeta}^{(k)}\right)^l}{l!}e^{-\lambda_1\hat{\zeta}^{(k)}}\right],$$

where $\hat{\zeta}^{(k)}$ is the asymptotic proportion of edges of type k. Similarly to the previous subsection, by using Lemma 4, we know that $\left(\hat{\zeta}^{(k)}, \mathcal{F}_n\right)_{n=1}^{\infty}$ is a martingale and for every $k \in [N]$ we have $\left|E_0^{(k)}\right| > 0$, thus $\hat{\zeta}^{(k)}$ is positive with positive probability and the last assumption of Theorem A is fulfilled.

In this special case we can prove the same results as in the previous subsection. For every $k \in [N]$ we have

$$\hat{x}_l^{(k)} \to \hat{C}_k$$

as $l \to \infty$ for some positive $\hat{C}_k$, and the characteristic exponent equals to 3. Again, for every $d \geq 0$ we define $\hat{Z}_n(d) = \left|\left\{v \in V_n : \sum_{k=1}^N \deg_n^{(k)}(v) = d\right\}\right|$. The asymptotic degree distribution is $(\hat{z}_d)_{d=0}^{\infty}$, where $\hat{z}_d$ is defined as the almost sure limit of the sequence $\left(\frac{\hat{Z}_n(d)}{|V_n|}\right)_{n=0}^{\infty}$ as $n \to \infty$. By using the same argument as in the previous subsection, we have

$$\hat{z}_d d^3 \to \hat{C}$$

as $d \to \infty$ for some positive $\hat{C}$, thus the characteristic exponent equals to 3.

2.4. Generalized Barabási–Albert random graph with perturbation

In this section we compare the asymptotic degree distributions of the multi-type Barabási–Albert graph and a perturbed version of this model. First, let us describe the dynamics and list the assumptions of this model.

2.4.1. Assumptions. In order to introduce perturbation, we have to define the matrices of error probabilities denoted by $\boldsymbol{F}_n$. For every $n \geq 1$ let

$$\boldsymbol{F}_n = (\varepsilon_{k,l}^{(n)} : k, l \in [N]) \in [0,1]^{N \times N}$$

be a matrix such that for every fixed $k \in [N]$ we have $\sum_{l=1}^{N} \varepsilon_{k,l}^{(n)} = 1$. That is, $\varepsilon_{k,l}^{(n)}$ is the probability that a type k edge becomes type l in the n^{th} step. We assume that there exists a matrix denoted by

$$\boldsymbol{F} = (\varepsilon_{k,l} : k, l \in [N]) \in [0,1]^{N \times N}$$

such that for every fixed $k \in [N]$ we have $\sum_{l=1}^{N} \varepsilon_{k,l} = 1$ and for every $k, l \in [N]$ we have $\varepsilon_{k,l}^{(n)} \to \varepsilon_{k,l}$ as $n \to \infty$. That is, for every k, l, the probability that a type k edge becomes type l, converges to $\varepsilon_{k,l}$.

The dynamics of the perturbed Barabási–Albert random graph is the following. Let us fix a positive integer denoted by M. In the n^{th} step

(1) a new vertex v_n is born.
(2) The vertex v_n attaches to some of the already existing vertices with M (not necessarily different) edges with probabilities proportional to the actual degrees of the existing vertices. The endpoints of the M new edges are chosen independently. We do not update the degrees of the vertices until the end of the n^{th} step.
(3) Every new edge gets a type randomly. The types of the new edges are chosen independently, and the probability of each type is its proportion among the types of the edges of the already existing endpoint of the new edge (not counting the edges added in the actual step).
(4) The types of the new edges change independently of each other with probabilities given by $\boldsymbol{F}_n$, i.e. if there is a new edge of type k, then its type after perturbation is l with probability $\varepsilon_{k,l}^{(n)}$.

2.4.2. Asymptotic degree distribution in the perturbed Barabási–Albert model.

First, recall the definition of irreducible matrices.

Definition 8. A matrix with non-negative entries denoted by $\boldsymbol{F} \in \mathbb{R}_{+}^{N \times N}$ is called irreducible if for every i, j there exists m such that $(\boldsymbol{F}^m)_{i,j} > 0$.

We are now ready to state our main theorem on the asymptotic degree distribution of the perturbed Barabási–Albert random graph.

THEOREM 5. *In the perturbed Barabási–Albert random graph, if we assume that* $\boldsymbol{F}$ *is irreducible, then for every* $\boldsymbol{d} = (d_1, \ldots, d_N)^T \in \mathbb{N}^N$

$$\lim_{n\to\infty} \frac{X_n(\boldsymbol{d})}{|V_n|} = x(\boldsymbol{d}) \ \textit{a.s.}$$

holds for a deterministic $x(\boldsymbol{d}) \in [0,1]$. *Furthermore, for every* $\boldsymbol{d} \in \mathbb{N}^N$ *we have the following recurrence equation:*

$$\textit{if } s(\boldsymbol{d}) = M, \textit{ then} \quad x(\boldsymbol{d}) = \frac{2\cdot M!}{M+2}\left[\prod_{l=1}^{N} \frac{1}{d_l!}\left(\sum_{k=1}^{N} \psi^{(k)} \cdot \varepsilon_{k,l}\right)^{d_l}\right]$$

$$\textit{if } s(\boldsymbol{d}) > M, \textit{ then} \quad x(\boldsymbol{d}) = \sum_{l=1}^{N} \frac{(\boldsymbol{d}-\boldsymbol{e}_l)^T \boldsymbol{F}_{\bullet,l}}{s(\boldsymbol{d})+2} x(\boldsymbol{d}-\boldsymbol{e}_l),$$

where

- $s(\boldsymbol{d}) = \boldsymbol{d}^T \boldsymbol{1} = \sum_{l=1}^{N} d_l$,
- $\psi^{(k)}$ *is the almost sure limit of the proportion of edges of type* k, *which is a deterministic constant and*
- $\boldsymbol{F}_{\bullet,l}$ *denotes the* l^{th} *column of the matrix* $\boldsymbol{F}$.

If we also assume that $\boldsymbol{F}$ *is symmetric, then for every* $\boldsymbol{d} = (d_1, \ldots, d_N)^T \in \mathbb{N}^N$ *we have the following recurrence equation:*

$$\textit{if } s(\boldsymbol{d}) = M, \textit{ then} \quad x(\boldsymbol{d}) = \frac{2\cdot M!}{M+2} \cdot \frac{1}{\prod_{l=1}^{N} d_l!}\left(\frac{1}{N}\right)^{M}$$

$$\textit{if } s(\boldsymbol{d}) > M, \textit{ then} \quad x(\boldsymbol{d}) = \sum_{l=1}^{N} \frac{(\boldsymbol{d}-\boldsymbol{e}_l)^T \boldsymbol{F}_{\bullet,l}}{s(\boldsymbol{d})+2} x(\boldsymbol{d}-\boldsymbol{e}_l).$$

REMARK. Notice that $x(\boldsymbol{d}) = 0$ if $s(\boldsymbol{d}) < M$, because every new vertex that is added to the initial configuration is attached to the existing vertices with exactly M new edges and all the vertices only receive new edges throughout the evolution of the graph, i.e. the proportion of vertices with less than M edges tend to disappear.

For comparison, let us assume that there is no perturbation, i.e. $\boldsymbol{F}_n$ is the identity matrix for every n. It also means that the condition on the irreducibility of $\boldsymbol{F}$ fails in Theorem 5. However, Theorem 1 in Section 2.2 describes the asymptotic degree distribution in the non-perturbed version of the model. Recall that, in the multi-type Barabási–Albert random graph for every $\boldsymbol{d} \in \mathbb{N}^N$, we have

$$\lim_{n\to\infty} \frac{X_n(\boldsymbol{d})}{|V_n|} = x(\boldsymbol{d}) \text{ a.s., where now } x(\boldsymbol{d}) \text{ is a non-deterministic random variable.}$$

The random variables $x(\boldsymbol{d})$ satisfy the following recurrence equation for every possible $\boldsymbol{d} \in \mathbb{N}^N$:

$$x(\boldsymbol{d}) = \sum_{l=1}^{N} \frac{d_l - 1}{s(\boldsymbol{d}) + 2} x(\boldsymbol{d} - \boldsymbol{e}_l) + \frac{2}{s(\boldsymbol{d}) + 2} \mathbb{P}(M = s(\boldsymbol{d})) \frac{s(\boldsymbol{d})!}{\prod_{l=1}^{N} d_l!} \prod_{l=1}^{N} \left(\psi^{(l)}\right)^{d_l},$$

where $\psi^{(l)}$ is the almost sure limit of the proportion of the edges of type l. In this case the asymptotic degree distribution is random, which means that it also depends on the asymptotic proportion of edges of different types. If $M = 1$, that is, the graph is a tree, then $(\psi^{(l)}, l \in [N])$ has Dirichlet distribution with parameters $(E_0^{(l)}, l \in [N])$. However, in the perturbed Barabási–Albert random graph the asymptotic degree distribution is deterministic.

A general urn model. In order to prove Theorem 5, we need to understand the asymptotic behaviour of the composition of edges of different types. We introduce a general urn model to describe the proportion of the edges in the multi-type perturbed Barabási–Albert random graph. This model is a generalization of a special case of the urn model introduced by Laruelle and Pagès in [**42**]. We remark that the generalization of the results of [**33**] could also be used for our purposes.

We assume that there are N colours represented by the elements of the set $\{1, 2, \dots, N\}$. The composition vector of the urn in the n^{th} step is denoted by $\boldsymbol{C}_n \in \mathbb{N}^N$, i.e. $C_{n,i}$ is the number of balls of colour i. The total number of balls in the urn in the n^{th} step is denoted by $s(\boldsymbol{C}_n) = \sum_{i=1}^{N} C_{n,i}$. We assume that in every step we draw M balls, with replacement, independently of each other and at the end of the step we add some additional balls to the urn. In the n^{th} step for trial i (where $i \in [M]$), let $\boldsymbol{\chi}_n^{(i)}$ be the N dimensional indicator vector of the colour drawn, and let $\boldsymbol{R}_n^{(i)}$ be the $N \times N$ dimensional replacement matrix with possibly random but non-negative entries. This means that $\left(\boldsymbol{R}_n^{(i)}\right)_{k,l}$ is the number of balls of colour l added to the urn if a ball of colour k was chosen in the n^{th} step for trial i.

For every n, we have

$$\boldsymbol{C}_{n+1} = \boldsymbol{C}_n + \sum_{i=1}^{M} \boldsymbol{R}_{n+1}^{(i)} \boldsymbol{\chi}_{n+1}^{(i)}. \tag{9}$$

For every n, we denote by $\mathcal{G}_n$ the σ-algebra of the draws and replacements in the first n steps, that is, the σ-algebra generated by

$$\boldsymbol{C}_0, \left(\boldsymbol{\chi}_j^{(i)}, i = 1, 2, \dots, M\right)_{j=1}^{n} \text{ and } \left(\boldsymbol{R}_j^{(i)}, i = 1, 2, \dots, M\right)_{j=1}^{n}.$$

We have the following assumptions on the urn model:

- **(U1):** the initial configuration $\boldsymbol{C}_0$ is non-negative and at least one of the coordinates is positive;
- **(U2):** for every n, i and j we have $\mathbb{P}\left(\boldsymbol{\chi}_n^{(i)} = \boldsymbol{e}_j \middle| \mathcal{G}_{n-1}\right) = \frac{C_{n-1,j}}{s(\boldsymbol{C}_{n-1})}$, where $\boldsymbol{e}_j$ is the j^{th} unit vector in $\mathbb{R}^N$; that is, the probability of choosing a colour is its proportion in the urn;
- **(U3):** for every n, the random variables $\boldsymbol{\chi}_n^{(1)}, \boldsymbol{\chi}_n^{(2)}, \dots, \boldsymbol{\chi}_n^{(M)}$ are identically distributed given $\mathcal{G}_{n-1}$, and similarly the random matrices denoted by $\boldsymbol{R}_n^{(1)}, \boldsymbol{R}_n^{(2)}, \dots, \boldsymbol{R}_n^{(M)}$ are identically distributed given $\mathcal{G}_{n-1}$, furthermore we assume that

$$\boldsymbol{\chi}_n^{(1)}, \boldsymbol{\chi}_n^{(2)}, \dots, \boldsymbol{\chi}_n^{(M)}, \boldsymbol{R}_n^{(1)}, \boldsymbol{R}_n^{(2)}, \dots, \boldsymbol{R}_n^{(M)}$$

  are conditionally independent given $\mathcal{G}_{n-1}$. This implies that, even if the replacement matrix is random, it is independent of the actual draw.

For every n, we define the generating matrices as the conditional expectation of the replacement matrix, i.e. $\boldsymbol{H}_n = \mathbb{E}\left(\boldsymbol{R}_n^{(1)} \middle| \mathcal{G}_{n-1}\right)$. We assume that

(U4): for every n and i the replacement matrix $\boldsymbol{R}_n^{(i)}$ has non-negative values almost surely;

(U5): for every n and i every column of the replacement matrix has the same weight almost surely, i.e. for every n, i and j we have $s\left((\boldsymbol{R}_n^{(i)})_{\bullet,j}\right) = \gamma_1$, that is, the number of balls added to the urn is constant;

(U6): every column of the generating matrices has the same weight almost surely, i.e. for every n and j we have $s\left((\boldsymbol{H}_n)_{\bullet,j}\right) = \gamma_2$. This constant is also known as the balance of the urn.

Finally, we assume that

(U7): there exists an irreducible $N \times N$ matrix denoted by $\boldsymbol{H}$ such that

$$\boldsymbol{H}_n \xrightarrow[n\to\infty]{a.s.} \boldsymbol{H}.$$

We denote by $\boldsymbol{v}^*_{\boldsymbol{H}}$ the normalized eigenvector of $\boldsymbol{H}$ corresponding to the largest eigenvalue of $\boldsymbol{H}$ such that $\|\boldsymbol{v}^*_{\boldsymbol{H}}\|_2 = 1$.

REMARK. If $M = 1$, then we get back the urn model in [**42**].

The next theorem states that the asymptotic composition of colours can be described with the normalized eigenvector corresponding to the largest eigenvalue.

THEOREM 6. *For all integers $M > 0$, assumptions* (**U1**)-(**U7**) *imply that*

$$\frac{\boldsymbol{C}_n}{s(\boldsymbol{C}_n)} \xrightarrow[n\to\infty]{a.s.} \boldsymbol{v}^*_{\boldsymbol{H}}.$$

To prove this theorem, we will use the same method as Laruelle and Pagès in [**42**], also known as the ordinary differential equation (ODE) method, which is a powerful tool of stochastic approximation.

Let us have a filtered probability space denoted by $(\Omega, (\mathcal{G}_n)_{n\geq 0}, \mathbb{P})$ and consider the following recurrence equation:

$$\boldsymbol{\vartheta}_{n+1} = \boldsymbol{\vartheta}_n - \gamma_{n+1} h(\boldsymbol{\vartheta}_n) + \gamma_{n+1}\left(\Delta \boldsymbol{M}_{n+1} + \boldsymbol{r}_{n+1}\right)$$

for $n \geq n_0$, where $h : \mathbb{R}^N \to \mathbb{R}^N$ is a locally Lipschitz continuous function, $\boldsymbol{\vartheta}_{n_0}$ is an $\mathbb{R}^N$-valued $\mathcal{G}_{n_0}$ measurable random variable, $(\gamma_n)_{n\geq n_0+1}$ is a (deterministic) sequence of positive numbers, $(\Delta \boldsymbol{M}_n)_{n\geq n_0+1}$ is a martingale difference in $(\mathcal{G}_n)_{n\geq n_0}$ and finally $(\boldsymbol{r}_n)_{n\geq n_0+1}$ is a sequence of $(\mathcal{G}_n)_{n\geq n_0+1}$-adapted random variables.

Theorem B (Almost sure convergence with ODE method, Theorem A.1 in [**42**]). *Assume that we have*

$$\boldsymbol{r}_n \xrightarrow[n\to\infty]{a.s.} 0, \qquad \sup_{n\geq n_0} \mathbb{E}\left(\|\Delta \boldsymbol{M}_{n+1}\|_2^2 \middle| \mathcal{G}_n \right) < \infty \ a.s.$$

and the sequence $(\gamma_n)_{n\geq n_0}$ satisfies the following assumptions

$$\sum_{n=n_0}^{\infty} \gamma_n = \infty \qquad \text{and} \qquad \sum_{n=n_0}^{\infty} \gamma_n^2 < \infty.$$

We denote by Θ_∞ the almost sure limiting values of the sequence $(\vartheta_n)_{n\geq n_0}$ as $n \to \infty$. Then Θ_∞ is almost surely a compact connected set.

Let us have a look at the following ordinary differential equation:

$$\dot{\boldsymbol{\vartheta}}(t) = -h(\boldsymbol{\vartheta}(t)), \text{ where } t \geq t_0. \tag{10}$$

The flow of the above differential equation on Θ_∞ is $\varphi(t, t_0, \boldsymbol{\vartheta}_0) = \boldsymbol{\vartheta}(t)$, if $\boldsymbol{\vartheta}(t)$ is the solution of the this differential equation with initial value $\boldsymbol{\vartheta}(t_0) = \boldsymbol{\vartheta}_0$.

We assume that for every $\boldsymbol{\vartheta}_0 \in \Theta_\infty$ the flow $\varphi(t, t_0, \boldsymbol{\vartheta}_0)$ is stable, i.e. for every $\varepsilon > 0$ and $t_1 > t_0$ there exists $\delta > 0$ such that

$$\text{if } |\boldsymbol{\tau} - \varphi(t_1, t_0, \boldsymbol{\vartheta}_0)| < \delta, \text{ then } |\varphi(t, t_1, \boldsymbol{\tau}) - \varphi(t, t_0, \boldsymbol{\vartheta}_0)| < \varepsilon \text{ for every } t \geq t_1.$$

If $\boldsymbol{\vartheta}^* \in \Theta_\infty$ is a uniformly stable equilibrium on Θ_∞ of the ordinary differential equation defined in (10)*, i.e.*

$$\sup_{\boldsymbol{\vartheta}_0 \in \Theta_\infty} |\varphi(t, t_0, \boldsymbol{\vartheta}_0) - \boldsymbol{\vartheta}^*| \xrightarrow[t\to\infty]{} 0,$$

then we have

$$\boldsymbol{\vartheta}_n \xrightarrow[n\to\infty]{a.s.} \boldsymbol{\vartheta}^*.$$

PROOF OF THEOREM 6. For every $n \geq 1$ we have

$$\boldsymbol{C}_{n+1} = \boldsymbol{C}_n + \sum_{i=1}^{M} \boldsymbol{R}_{n+1}^{(i)} \boldsymbol{\chi}_{n+1}^{(i)} = \boldsymbol{C}_n + \mathbb{E}\left(\sum_{i=1}^{M} \boldsymbol{R}_{n+1}^{(i)} \boldsymbol{\chi}_{n+1}^{(i)} \middle| \mathcal{G}_n\right) + \Delta \boldsymbol{M}_{n+1}, \tag{11}$$

where

$$\Delta \boldsymbol{M}_{n+1} = \sum_{i=1}^{M} \boldsymbol{R}_{n+1}^{(i)} \boldsymbol{\chi}_{n+1}^{(i)} - \mathbb{E}\left(\sum_{i=1}^{M} \boldsymbol{R}_{n+1}^{(i)} \boldsymbol{\chi}_{n+1}^{(i)} \middle| \mathcal{G}_n\right).$$

Recall that the generating matrices are defined as $\boldsymbol{H}_n^{(i)} = \mathbb{E}\left(\boldsymbol{R}_n^{(i)} \middle| \mathcal{G}_{n-1}\right)$. By using this and assumption **(U3)** on the conditional independence of $\boldsymbol{R}_{n+1}^{(i)}$ and $\boldsymbol{\chi}_{n+1}^{(i)}$, we have

$$\begin{aligned}
\mathbb{E}\left(\sum_{i=1}^{M} \boldsymbol{R}_{n+1}^{(i)} \boldsymbol{\chi}_{n+1}^{(i)} \middle| \mathcal{G}_n\right) &= \sum_{i=1}^{M} \mathbb{E}\left(\boldsymbol{R}_{n+1}^{(i)} \boldsymbol{\chi}_{n+1}^{(i)} \middle| \mathcal{G}_n\right) \\
&= \sum_{i=1}^{M} \sum_{j=1}^{N} \mathbb{E}\left(\boldsymbol{R}_{n+1}^{(i)} \mathrm{Ind}(\boldsymbol{\chi}_{n+1}^{(i)} = \boldsymbol{e}_j) \boldsymbol{e}_j \middle| \mathcal{G}_n\right) \\
&= \sum_{i=1}^{M} \left[\sum_{j=1}^{N} \mathbb{E}\left(\boldsymbol{R}_{n+1}^{(i)} \middle| \mathcal{G}_n\right) \mathbb{P}\left(\boldsymbol{\chi}_{n+1}^{(i)} = \boldsymbol{e}_j \middle| \mathcal{G}_n\right) \boldsymbol{e}_j\right] \\
&= \sum_{i=1}^{M} \left[\boldsymbol{H}_{n+1}^{(i)} \sum_{j=1}^{N} \frac{C_{n,j}}{s(\boldsymbol{C}_n)} \boldsymbol{e}_j\right] = \left(\sum_{i=1}^{M} \boldsymbol{H}_{n+1}^{(i)}\right) \frac{\boldsymbol{C}_n}{s(\boldsymbol{C}_n)}.
\end{aligned}$$

By normalizing equation (11), we have

$$\frac{\boldsymbol{C}_{n+1}}{s(\boldsymbol{C}_{n+1})} = \frac{\boldsymbol{C}_n}{s(\boldsymbol{C}_n)} + \frac{1}{s(\boldsymbol{C}_{n+1})} \left[\left(\sum_{i=1}^{M} \boldsymbol{H}_{n+1}^{(i)}\right) - M \boldsymbol{I}_N\right] \frac{\boldsymbol{C}_n}{s(\boldsymbol{C}_n)} + \frac{\Delta \boldsymbol{M}_{n+1}}{s(\boldsymbol{C}_{n+1})}, \tag{12}$$

where $\boldsymbol{I}_N$ is the N-dimensional identity matrix. To verify the above reformulation we can check that

$$\frac{\boldsymbol{C}_n}{s(\boldsymbol{C}_n)} - \frac{1}{s(\boldsymbol{C}_{n+1})} \cdot M \boldsymbol{I}_N \frac{\boldsymbol{C}_n}{s(\boldsymbol{C}_n)} = \frac{\boldsymbol{C}_n}{s(\boldsymbol{C}_{n+1})},$$

by using the fact that $s(\boldsymbol{C}_{n+1}) = s(\boldsymbol{C}_n) + M$.

Let us define $\widetilde{\boldsymbol{C}}_n = \boldsymbol{C}_n / s(\boldsymbol{C}_n)$. Equation (12) can be rewritten as the canonical stochastic approximation process in the following way:

$$\begin{aligned}\widetilde{\boldsymbol{C}}_{n+1} &= \widetilde{\boldsymbol{C}}_n + \frac{1}{s(\boldsymbol{C}_{n+1})}\left[\left(\sum_{i=1}^{M} \boldsymbol{H}_{n+1}^{(i)}\right) - M\boldsymbol{I}_N\right]\widetilde{\boldsymbol{C}}_n + \frac{\Delta \boldsymbol{M}_{n+1}}{s(\boldsymbol{C}_{n+1})} \\ &= \widetilde{\boldsymbol{C}}_n - \frac{1}{s(\boldsymbol{C}_{n+1})} M\left(\boldsymbol{I}_N - \boldsymbol{H}\right)\widetilde{C}_n + \frac{1}{s(\boldsymbol{C}_{n+1})}\left(\Delta \boldsymbol{M}_{n+1} + \boldsymbol{r}_{n+1}\right)\end{aligned}$$

with step size $\gamma_n = 1/s(\boldsymbol{C}_n)$ and the error term is defined as

$$\boldsymbol{r}_{n+1} = \left[\left(\sum_{i=1}^{M} \boldsymbol{H}_{n+1}^{(i)}\right) - M\boldsymbol{H}\right]\widetilde{\boldsymbol{C}}_n.$$

To apply the ODE method we need to check the assumptions of Theorem B.

Since $\widetilde{\boldsymbol{C}}_n$ is bounded a.s., by using assumption **(U7)** we have $\boldsymbol{r}_n \xrightarrow[n\to\infty]{a.s.} 0$. Notice that, we have

$$\left\| \sum_{i=1}^{M} \boldsymbol{R}_{n+1}^{(i)} \boldsymbol{\chi}_{n+1}^{(i)} \right\|_2^2 \leq M^2,$$

since the number of balls added in a step equals to M, which is fixed. Consequently we have

$$\sup_{n\geq 1} \mathbb{E}\left(\left\| \sum_{i=1}^{M} \boldsymbol{R}_{n+1}^{(i)} \boldsymbol{\chi}_{n+1}^{(i)} \right\|_2^2 \middle| \mathcal{G}_n \right) < \infty,$$

thus we obtain that $\sup_{n\geq 1} \mathbb{E}\left(\|\Delta \boldsymbol{M}_{n+1}\|_2^2 \middle| \mathcal{G}_n \right) < \infty$ almost surely.

It is obvious that the (almost sure) limiting values of $\widetilde{\boldsymbol{C}}_n$ as $n \to \infty$ are in the N-dimensional simplex denoted by $\mathcal{S} = \left\{ u \in \mathbb{R}_+^N | s(u) = 1 \right\}$. Let us have a look at the following ordinary differential equation:

$$\dot{y} = -M(\boldsymbol{I}_N - \boldsymbol{H})y,$$

where $y : \mathbb{R}^N \to \mathbb{R}^N$ is a differentiable function. By using assumption **(U7)** we obtain that $\boldsymbol{v}_{\boldsymbol{H}}^*$ is the unique zero of the following function: $y \mapsto -M(\boldsymbol{I}_N - \boldsymbol{H})y$

on $y \in \mathcal{S}$. Let us take the restriction of the above differential equation to the set $\mathcal{V}_0 = \left\{u \in \mathbb{R}_+^N | s(u) = 0\right\}$. By using assumption **(U7)** we conclude that the left eigenvalues of $M(\boldsymbol{I}_N - \boldsymbol{H})$ have positive real part. As a consequence we get that $\boldsymbol{v}_{\boldsymbol{H}}^*$ is a uniformly stable equilibrium of the equation on $\mathcal{S}$. By using the ODE method we conclude that

$$\frac{\boldsymbol{C}_n}{s(\boldsymbol{C}_n)} \xrightarrow[n\to\infty]{a.s.} \boldsymbol{v}_{\boldsymbol{H}}^*.$$

$\square$

Proof of the main theorem. First, we need to prove the following lemma on the asymptotic proportion of edges of different types. This is where we use Theorem 6 for the urn models.

LEMMA 6. *In the perturbed Barabási–Albert random graph, if we assume that* $\boldsymbol{F} \in (0,1)^{N\times N}$, *then for every* $l \in [N]$ *we have* $\psi_n^{(l)} = \frac{|E_n^{(l)}|}{|E_n|} \to \psi^{(l)}$ *almost surely as* $n \to \infty$, *where* $\psi^{(l)} \in (0,1)$ *is a deterministic constant. If we also assume that* $\boldsymbol{F} = (\varepsilon_{k,l})_{k,l=1}^N$ *is symmetric, then for every* $l \in [N]$ *we have* $\psi^{(l)} = \frac{1}{N}$.

PROOF. In the perturbed Barabási–Albert model, we can use the following urn model to understand the asymptotic composition of the number edges of type l for every $l \in [N]$. Let us have $\boldsymbol{C}_0 = (|E_0^{(l)}|, l \in [N])$ and for every $n \geq 1$ and $i \in [M]$ we define $\boldsymbol{R}_n^{(i)} = (\tau_{n;k,l}^{(i)})_{k,l=1}^N$, where $\tau_{n;k,l}^{(i)}$ is a Bernoulli distributed random variable with expectation equal to $\varepsilon_{l,k}^{(n)}$, furthermore we assume that for every $l \in [N]$ we have $\sum_{k=1}^N \tau_{n;k,l}^{(i)} = 1$ and the columns of the matrix $\boldsymbol{R}_n^{(i)}$ are independent of each other. Clearly, we have

$$\boldsymbol{H}_n = \mathbb{E}\left(\boldsymbol{R}_n^{(1)} | \mathcal{G}_{n-1}\right) = \boldsymbol{F}_n^T.$$

To apply Theorem 6 we have to check the assumptions of the general urn model and find $\boldsymbol{v}_{\boldsymbol{H}}^*$ to complete the proof of Lemma 6.

Assumption **(U1)** holds due to the fact that there is at least one edge of each type in the initial configuration of the perturbed Barabási–Albert random graph. By the dynamics of the model assumptions **(U2)-(U3)** hold (recall that we do not update the degrees of the vertices until the end of the steps). Assumptions **(U4)-(U6)** hold because of the choice of $\boldsymbol{R}_n^{(i)}$. Notice that in this case $\gamma_1 = \gamma_2 = 1$. For assumption **(U7)** we need to show that there exists an irreducible $N \times N$ matrix denoted by $\boldsymbol{H}$ such that $\boldsymbol{H}_n \xrightarrow[n\to\infty]{a.s.} \boldsymbol{H}$. In the perturbed Barabási–Albert random graph we assumed that $\boldsymbol{H}_n^T = \boldsymbol{F}_n \to \boldsymbol{F}$ in every entry. Since $\boldsymbol{F}$ is irreducible, we can choose $\boldsymbol{H} = \boldsymbol{F}^T$. Hence all assumptions of Theorem 6 hold.

The normalized (right) eigenvector of $\boldsymbol{H}$ corresponding to the eigenvalue with the largest real part is $\boldsymbol{v}_{\boldsymbol{H}}^* = \left(\psi^{(1)}, \dots, \psi^{(N)}\right)$. Notice that this is also the normalized left eigenvector of $\boldsymbol{F}$ corresponding to the same eigenvalue. By using Theorem 6 we get the first part of the lemma.

If we also assume that $\boldsymbol{F}$ is symmetric, then $\boldsymbol{F}$ is a double-stochastic matrix. It follows that for every $l \in [N]$ we have $\psi^{(l)} = \frac{1}{N}$. □

Now, we can prove our main result on the asymptotic degree distribution of the perturbed Barabási–Albert random graph.

To prove the existence of the asymptotic degree distribution, we can use Theorem 1 in Section 2.2.

In the proof of the main theorem, we will use Lemma 3. Recall the statement of this lemma which claims that for every $n \geq 1$ and $x \in [0, 1]$, we have

$$|(1-x)^n - (1-nx)| \leq \binom{n}{2} x^2.$$

Proof of Theorem 5. We need to check assumptions (**GM1**)-(**GM5**) of the general model. We will use the following set of indices: for any $\boldsymbol{d} \in \mathbb{N}^N$ (where $s(\boldsymbol{d}) \geq 1$), we define

$$\alpha(\boldsymbol{d}) = \left\{ \boldsymbol{i} = (i_1, \dots, i_N)^T \in \mathbb{N}^N : i_l \leq d_l \quad \forall l \in [N] \quad \text{and} \quad 1 \leq s(\boldsymbol{i}) \leq M \right\}.$$

For every $\boldsymbol{i} \in \mathbb{N}^N$ we define

$$\beta(\boldsymbol{i}) = \left\{\boldsymbol{I} = (i_{k,l})_{k,l=1}^N \in \mathbb{N}^{N\times N} : s\left(\boldsymbol{I}_{\bullet,l}\right) = i_l \quad \forall l \in [N]\right\}.$$

By the dynamics of the perturbed Barabási–Albert random graph, assumption (**GM1**) trivially holds.

To see that assumption (**GM2**) is satisfied, notice that for every n we have

$$\mathbb{E}\left[(X_n(\boldsymbol{d}) - X_{n-1}(\boldsymbol{d}))^2 \middle| \mathcal{F}_{n-1}\right] \leq M^2 < \infty,$$

because there are M new edges.

For assumption (**GM3**), we need to show that $u_n(\boldsymbol{d}) \to u(\boldsymbol{d}) > 0$ almost surely as $n \to \infty$, where

$$1 - \frac{u_n(\boldsymbol{d})}{n} = p_{\boldsymbol{d}}^{(n)}(\boldsymbol{0}) = \left(1 - \frac{s(\boldsymbol{d})}{2|E_{n-1}|}\right)^M.$$

To find the almost sure limit of $u_n(\boldsymbol{d})$ as $n \to \infty$, we can use the following formula:

$$\left(1 - \frac{s(\boldsymbol{d})}{2|E_{n-1}|}\right)^M = 1 - \frac{Ms(\boldsymbol{d})}{2|E_{n-1}|} + \eta_n(\boldsymbol{d}),$$

where

$$\eta_n(\boldsymbol{d}) = \left(1 - \frac{s(\boldsymbol{d})}{2|E_{n-1}|}\right)^M - \left[1 - \frac{Ms(\boldsymbol{d})}{2|E_{n-1}|}\right].$$

By using Lemma 3 and the fact that $|E_n| \sim Mn$, we obtain that

$$|\eta_n(\boldsymbol{d})| \leq \binom{M}{2}\left(\frac{s(\boldsymbol{d})}{2|E_{n-1}|}\right)^2 \leq M^2\left(\frac{s(\boldsymbol{d})}{2|E_{n-1}|}\right)^2 = o\left(\frac{1}{n}\right),$$

which yields

$$\begin{aligned}
u_n(\boldsymbol{d}) = n[1 - p_{\boldsymbol{d}}^{(n)}(\boldsymbol{0})] &= n\left[1 - \left(1 - \frac{s(\boldsymbol{d})}{2|E_{n-1}|}\right)^M\right] \\
&= n\left[1 - \left(1 - \frac{Ms(\boldsymbol{d})}{2|E_{n-1}|} + \eta_n(\boldsymbol{d})\right)\right] = n\left[\frac{Ms(\boldsymbol{d})}{2|E_{n-1}|} - \eta_n(\boldsymbol{d})\right] \\
&= n\frac{Ms(\boldsymbol{d})}{2|E_{n-1}|} - n\eta_n(\boldsymbol{d}) \to u(\boldsymbol{d}) = \frac{s(\boldsymbol{d})}{2} > 0
\end{aligned}$$

almost surely as $n \to \infty$.

For assumption (**GM4**), we will show that for every $\boldsymbol{i} \in \mathbb{N}^N$, such that $s(\boldsymbol{i}) \geq 1$

$$\lim_{n\to\infty} n p^{(n)}_{\boldsymbol{d}-\boldsymbol{i}}(\boldsymbol{i}) = \begin{cases} \frac{1}{2}(\boldsymbol{d}-\boldsymbol{e}_l)^T \boldsymbol{F}_{\bullet,l} & \text{if } \boldsymbol{i} = \boldsymbol{e}_l, \\ 0 & \text{otherwise} \end{cases}$$

holds almost surely. Notice that if $\boldsymbol{d}_k - \boldsymbol{i}_k < 0$ for any $k \in [N]$ or $\boldsymbol{d} - \boldsymbol{i} = \boldsymbol{0}$, then $p^{(n)}_{\boldsymbol{d}-\boldsymbol{i}}(\boldsymbol{i}) = 0$ for every n.

Let us fix $\boldsymbol{I} = (i_{k,l})_{k,l=1}^N \in \beta(\boldsymbol{i})$ where $\boldsymbol{i} \in \alpha(\boldsymbol{d})$. In the perturbed Barabási–Albert random graph for a fixed vertex and for every $k, l \in [N]$ we denote by $i_{k,l}$ the number of edges connected to the given vertex which were originally of type k and changed their types to l. In this case the value of $p^{(n)}_{\boldsymbol{d}-\boldsymbol{i}}(\boldsymbol{i})$ is given by

$$p^{(n)}_{\boldsymbol{d}-\boldsymbol{i}}(\boldsymbol{i}) = \sum_{\boldsymbol{I}\in\beta(\boldsymbol{i})} \hat{p}^{(n)}_{\boldsymbol{d}-\boldsymbol{i}}(\boldsymbol{I}),$$

where

$$\hat{p}^{(n)}_{\boldsymbol{d}-\boldsymbol{i}}(\boldsymbol{I}) = \frac{M!}{\left(\prod_{k,l=1}^N i_{k,l}!\right)(M-s(\boldsymbol{i}))!} \left[\prod_{k,l=1}^N \left[\left(\frac{\boldsymbol{d}_k - \boldsymbol{i}_k}{2|E_{n-1}|}\right)\varepsilon^{(n)}_{k,l}\right]^{i_{k,l}}\right] \left[1 - \frac{s(\boldsymbol{d}-\boldsymbol{i})}{2|E_{n-1}|}\right]^{M-s(\boldsymbol{i})}.$$

First, let us fix $l \in [N]$. Similarly to the previous calculations, we can use the following formula

$$\begin{aligned} p^{(n)}_{\boldsymbol{d}-\boldsymbol{e}_l}(\boldsymbol{e}_l) &= M\left[\sum_{k=1}^N \frac{\boldsymbol{d}_k - (\boldsymbol{e}_l)_k}{2|E_{n-1}|}\varepsilon^{(n)}_{k,l}\right]\left(1 - \frac{s(\boldsymbol{d})-1}{2|E_{n-1}|}\right)^{M-1} \\ &= M\left[\sum_{k=1}^N \frac{\boldsymbol{d}_k - (\boldsymbol{e}_l)_k}{2|E_{n-1}|}\varepsilon^{(n)}_{k,l}\right]\left(1 - \frac{(M-1)(s(\boldsymbol{d})-1)}{2|E_{n-1}|} + \eta'_n(\boldsymbol{d})\right) \end{aligned}$$

where

$$\eta'_n(\boldsymbol{d}) = \left(1 - \frac{s(\boldsymbol{d})-1}{2|E_{n-1}|}\right)^{M-1} - \left[1 - \frac{(M-1)(s(\boldsymbol{d})-1)}{2|E_{n-1}|}\right]$$

and $(\boldsymbol{e}_l)_k$ denotes the k^{th} element of $\boldsymbol{e}_l$, i.e. $(\boldsymbol{e}_l)_k = \begin{cases} 1 & \text{if } k = l \\ 0 & \text{if } k \neq l \end{cases}$.

Again, by using Lemma 3 and the fact that $|E_n| \sim Mn$, we get that

$$|\eta'_n(\boldsymbol{d})| \leq \binom{M-1}{2} \left(\frac{s(\boldsymbol{d})-1}{2|E_{n-1}|}\right)^2 \leq (M-1)^2 \left(\frac{s(\boldsymbol{d})-1}{2|E_{n-1}|}\right)^2 = o\left(\frac{1}{n}\right).$$

Recall that $\varepsilon_{k,l}^{(n)} \to \varepsilon_{k,l} \in (0,1)$ as $n \to \infty$ for every $k, l \in [N]$. We conclude that

$$np^{(n)}_{\boldsymbol{d}-\boldsymbol{e}_l}(\boldsymbol{e}_l) = nM \left[\sum_{k=1}^{N} \frac{\boldsymbol{d}_k - (\boldsymbol{e}_l)_k}{2|E_{n-1}|} \varepsilon_{k,l}^{(n)}\right] \left(1 - \frac{(M-1)(s(\boldsymbol{d})-1)}{2|E_{n-1}|} + \eta'_n(\boldsymbol{d})\right)$$
$$\to \sum_{k=1}^{N} \frac{\boldsymbol{d}_k - (\boldsymbol{e}_l)_k}{2} \varepsilon_{k,l} = \frac{(\boldsymbol{d}-\boldsymbol{e}_l)^T \boldsymbol{F}_{\bullet,l}}{2}$$

as $n \to \infty$.

Fix $\boldsymbol{i} \in \alpha(\boldsymbol{d}) \setminus \{\boldsymbol{e}_l, l \in [N]\}$. We need to prove that in this case $\lim_{n\to\infty} np^{(n)}_{\boldsymbol{d}-\boldsymbol{i}}(\boldsymbol{i}) = 0$. Recall that

$$p^{(n)}_{\boldsymbol{d}-\boldsymbol{i}}(\boldsymbol{i}) = \sum_{\boldsymbol{I}\in\beta(\boldsymbol{i})} \hat{p}^{(n)}_{\boldsymbol{d}}(\boldsymbol{I}).$$

Because of the choice of $\boldsymbol{i}$, we have $s(\boldsymbol{i}) = s(\boldsymbol{I}) \geq 2$. By using this and the fact that $|E_n| \sim Mn$, we conclude that

$$\hat{p}^{(n)}_{\boldsymbol{d}-\boldsymbol{i}}(\boldsymbol{I}) = \frac{M!}{\left(\prod_{k,l=1}^{N} i_{k,l}!\right)(M-s(\boldsymbol{i}))!} \left[\prod_{k,l=1}^{N} \left[\left(\frac{\boldsymbol{d}_k - \boldsymbol{i}_k}{2|E_{n-1}|}\right) \varepsilon_{k,l}^{(n)}\right]^{i_{k,l}}\right] \left[1 - \frac{s(\boldsymbol{d}-\boldsymbol{i})}{2|E_{n-1}|}\right]^{M-s(\boldsymbol{i})}$$
$$\leq \frac{M!}{\left(\prod_{k,l=1}^{N} i_{k,l}!\right)(M-s(\boldsymbol{i}))!} \left[\prod_{k,l=1}^{N} \left[\left(\frac{\boldsymbol{d}_k - \boldsymbol{i}_k}{2|E_{n-1}|}\right) \varepsilon_{k,l}^{(n)}\right]^{i_{k,l}}\right] = o\left(\frac{1}{n}\right).$$

This shows that (**GM4**) holds with $r^{(l)}(\boldsymbol{d}-\boldsymbol{e}_l) = \frac{1}{2}(\boldsymbol{d}-\boldsymbol{e}_l)^T \boldsymbol{F}_{\bullet,l}$.

Finally, for assumption (**GM5**), we have to find the almost sure limit of $q^{(n)}(\boldsymbol{d})$ as $n \to \infty$. Recall that, in the n^{th} step, every new edge will be of type l with probability $\psi_n^{(l)} = \frac{|E_n^{(l)}|}{|E_n|}$.

In the perturbed Barabási–Albert random graph we have

$$q(\boldsymbol{d}) = \lim_{n\to\infty} q^{(n)}(\boldsymbol{d}) = \lim_{n\to\infty} \left[\operatorname{Ind}(M = s(\boldsymbol{d})) \sum_{\boldsymbol{D}\in\beta(\boldsymbol{d})} \hat{q}^{(n)}(\boldsymbol{D})\right],$$

where

$$\hat{q}^{(n)}(\boldsymbol{D}) = \frac{s(\boldsymbol{D})!}{\prod_{k,l=1}^{N} d_{k,l}!} \left[\prod_{k,l=1}^{N} \left(\psi_n^{(k)} \cdot \varepsilon_{k,l}^{(n)} \right)^{d_{k,l}} \right].$$

Notice that $s(\boldsymbol{D}) = s(\boldsymbol{d})$. By using Lemma 6, the multinomial theorem and the fact that $\varepsilon_{k,l}^{(n)} \to \varepsilon_{k,l}$ almost surely as $n \to \infty$, we get that

$$\begin{aligned}
q(\boldsymbol{d}) &= \operatorname{Ind}(M = s(\boldsymbol{d})) \sum_{\boldsymbol{D} \in \beta(\boldsymbol{d})} \hat{q}(\boldsymbol{D}) \\
&= \operatorname{Ind}(M = s(\boldsymbol{d})) \sum_{\boldsymbol{D} \in \beta(\boldsymbol{d})} \frac{s(\boldsymbol{D})!}{\prod_{k,l=1}^{N} d_{k,l}!} \left[\prod_{k,l=1}^{N} \left(\psi^{(k)} \cdot \varepsilon_{k,l} \right)^{d_{k,l}} \right] \\
&= \operatorname{Ind}(M = s(\boldsymbol{d})) M! \left[\prod_{l=1}^{N} \frac{1}{d_l!} \left(\sum_{k=1}^{N} \psi^{(k)} \cdot \varepsilon_{k,l} \right)^{d_l} \right].
\end{aligned}$$

We conclude that

$$\begin{aligned}
u_n(\boldsymbol{d}) &\to u(\boldsymbol{d}) = \frac{s(\boldsymbol{d})}{2} \\
q^{(n)}(\boldsymbol{d}) &\to q(\boldsymbol{d}) = \operatorname{Ind}(M = s(\boldsymbol{d})) M! \left[\prod_{l=1}^{N} \frac{1}{d_l!} \left(\sum_{k=1}^{N} \psi^{(k)} \cdot \varepsilon_{k,l} \right)^{d_l} \right]
\end{aligned}$$

as $n \to \infty$. For the quantity defined in equation (1), we have

$$\lim_{n \to \infty} n p_{\boldsymbol{d}-\boldsymbol{i}}^{(n)}(\boldsymbol{i}) = \begin{cases} \frac{(\boldsymbol{d}-\boldsymbol{e}_l)^T \boldsymbol{F}_{\bullet,l}}{2} & \text{if } \boldsymbol{i} = \boldsymbol{e}_l, \\ 0 & \text{otherwise} \end{cases}$$

almost surely, that is, $r^{(l)}(\boldsymbol{d} - \boldsymbol{e}_l) = \frac{1}{2}(\boldsymbol{d} - \boldsymbol{e}_l)^T \boldsymbol{F}_{\bullet,l}$.

Applying Theorem 1, we get Theorem 5. □

CHAPTER 3

Epidemic spread on random graphs with multiple type edges

In this chapter, we examine the spread of epidemics (or information) on random graphs with multiple type edges. Due to the recent Covid-19 pandemic the spread of infectious diseases has became an intensively studied research area. The introduction of different types of the edges may result in more adequate models that can be used in the analysis. Since the probability (or the intensity) of the spread of the virus is different between individuals who live in the same household, work together or only meet rarely, labelling the edges with different propagation probabilities requires having various type of edges.

After summarizing the most important properties of the structures of the underlying graph models, we have a look at the different versions of the spread of epidemic processes on these graphs. The high-level overview of the approach is the following: we assign a state to the vertices of the graph, e.g. susceptible, infectious and recovered, then the spread of epidemic is modelled as a process on the phase space of the vertices. The different dynamics of these processes result in individual versions of these processes that can be used to model different phenomena. Then, we have a look at the empirical results of stochastic simulations related to these models and spread of epidemic processes with different parameters and examine the sensitivity on these specific parameters we will discuss in details. Results can also be found in [**47**].

3.1. Models

In this section, we define the graphs that are used as the underlying structures in the modelling of the spread of epidemics. These are the (multi-type) preferential attachment graph (which is different from those that we have seen in the previous sections), the model of independent edges (see Section 2.2.2) and a generalized version of the duplication and deletion model (which is a modification of the model defined in [**7**]). In terms of their definitions, these graph models can be either static or dynamic. This means that the structure of the graph of a given size is either defined by a specified rule, or as an element of a sequence of graphs where the set of vertices or edges are evolving according to a given dynamics.

Let us recall some notations from the previous chapter that we will use for the dynamic graph models. Let $(G_n)_{n=0}^{\infty}$ be a sequence of finite random graphs. The set of vertices and edges of G_n are denoted by V_n and E_n, respectively. The number of different types of edges, denoted by N, will be fixed. For every $k \in [N] = \{1, 2, \dots, N\}$ let $E_n^{(k)}$ denote the set of edges of type k in G_n. We assume that the different types form a partition of the edges, i.e. for every n we have $E_n = \bigcup_{k=1}^{N} E_n^{(k)}$ and for every k, l we have $E_n^{(k)} \cap E_n^{(l)} = \emptyset$ whenever $k \neq l$. We assume that the initial configuration, denoted by G_0, is a finite deterministic graph, moreover for every $k \in [N]$ we have $\left|E_0^{(k)}\right| > 0$. Finally, for every n let $\mathcal{F}_n$ denote the σ-algebra generated by the first n labelled graphs. We may choose $\mathcal{F}_0$ to be the trivial σ-algebra (since G_0 is deterministic), thus $\mathcal{F} = (\mathcal{F}_n)_{n=0}^{\infty}$ is a filtration.

For the static graph models, we simply omit the notation indicating the size of the graph from the indices of the vertex and edge sets.

3.1.1. Preferential attachment graph. In this section, we are going to define the multi-type preferential attachment graph model. The single-type version of this model is defined in [**19**]. First, let us have a look at the definition of the single-type version, then we define the multi-type preferential attachment graph model. Let $\beta > 0$ be a fixed parameter and let $V = \{v_1, v_2, \dots, v_n\}$ be the set of vertices.

In order to construct the set of edges we are going to create a sequence $(v_i^*)_{i=1}^{2m}$ from the elements of V. We start with the empty sequence. If the current length of the sequence equals to k, then the next element v_{k+1}^* is chosen to be equal to $v \in V$ with probability $\frac{d(n)+\beta}{k+n\beta}$, where $d(v)$ is the multiplicity of v in the sequence $v_1^*, \dots, v_k^*$. The edge set is defined as $E = \big\{\{v_{2i-1}^*, v_{2i}^*\}, i = 1, 2, \dots, m\big\}$. The single-type preferential attachment graph with n vertices, m edges and parameter β is denoted by $\mathrm{PAG}_\beta(n, m)$.

Remark. We are going to choose $m = n$, that means we use the sparse preferential attachment graph, since the simulation processes are much faster and more reliable.

Remark. Notice that the single-type preferential attachment graph is not the same, however it is motivated by the (sparse) Barabási–Albert graph model in [**13**], specified in [**18**].

In order to obtain a multi-type preferential attachment graph denoted by

$$N\text{-}\mathrm{PAG}_\beta(n, m_1, m_2, \dots, m_N),$$

we construct N independent single-type preferential attachment graphs on the vertex set V. Let us denote these independent graphs by

$$\mathrm{PAG}_\beta^{(1)}(n, m_1),\ \mathrm{PAG}_\beta^{(2)}(n, m_2), \dots,\ \mathrm{PAG}_\beta^{(N)}(n, m_N).$$

Then, the different edges which belong to $\mathrm{PAG}_\beta^{(k)}(n, m_k)$ form the set of edges of type k, where $k \in [N]$. Notice that the number of edges of N-$\mathrm{PAG}_\beta(n, m_1, m_2, \dots, m_k)$ equals to $N \cdot \sum_{k=1}^{N} m_k$.

3.1.2. Model of independent edges. The model of independent edges is another dynamic graph model. There are two different versions. The 2nd one was defined in Section 2.2.2.

Version I. This graph model is a modification and a multi-type version of the models defined in [**24**] and [**38**]. Let $\lambda > 0$ be a fixed parameter. In the n^{th} step, we have the following dynamics:

(i): a new vertex v_n is born, thus $V_n = V_0 \cup \{v_1, \ldots, v_n\}$;

(ii): every existing vertex $v \in V_{n-1}$, independently of each other, is connected to v_n with an edge of type k with probability equal to $\frac{\deg^{(k)}_{n-1}(v)}{2|E_{n-1}|}$. The choices for the edges of different types are also independent of each other.

Version II. Another version of the model was defined in Section 2.2.2. Let us recall the basic aspects of this model. We have a sequence $(\lambda_n)_{n=1}^{\infty}$ which meets certain conditions. In this chapter, we assume that $\lambda_n = \lambda$ for every n, where $\lambda > 0$ is fixed. In the n^{th} step we have the following dynamics:

(i): a new vertex v_n is born, thus $V_n = V_0 \cup \{v_1, \ldots, v_n\}$;

(ii): every existing vertex $v \in V_{n-1}$, independently of each other, is connected to v_n with $\Delta^{(k)}_n(v)$ edges of type k, where $\Delta^{(k)}_n(v) \sim \text{Poi}\left(\lambda \frac{\deg^{(k)}_{n-1}(v)}{2|E_{n-1}|}\right)$. The number of edges of different types are also independent of each other.

3.1.3. Duplication model. Let us describe the dynamics of the duplication model. For every vertex $v \in V$ we denote by $\mathcal{N}_n(v)$ the set of neighbours of v in G_{n-1}. Let us assume that there is an initial configuration with N different types of edges.

In the n^{th} step we have the following dynamics:

(i): a new vertex v_n is born, thus $V_n = V_0 \cup \{v_1, \ldots, v_n\}$.

(ii): We choose a vertex v from V_{n-1} uniformly at random and we connect v_n to every vertex in $\mathcal{N}_{n-1}(v)$. The type of the new edges will be exactly the same as the type of the edges connected to v. (*Duplication.*)

(iii): We choose a vertex w from V_{n-1} uniformly at random and we delete all the edges which are incident to w in G_{n-1}. Notice that the vertices v and w are not necessarily different. (*Deletion.*)

3.2. Epidemic spread

In this section we introduce the processes that may be suitable for modelling the spread of infectious diseases. These processes can be categorized according to the possible states of vertices of the underlying graph model.

For every epidemic spread process, we have an underlying graph to model the structure of the individuals of the population. Let us have a finite random graph on n vertices, denoted by G_n, with multi-type edges. Recall that the set of vertices and the set of edges of type k are denoted by V and $E^{(k)}$, where $k \in [N]$, respectively.

We will use the following notations:

$$\Pi^k = \left\{ \pi = (\pi_1, \pi_2, \ldots, \pi_k) \in \mathcal{P}^n(V) : \bigcup_{i=1}^{n} \pi_i = V \text{ and } \pi_i \cap \pi_j = \emptyset \text{ for every } i \neq j \right\}$$

$$\Sigma^k = \left\{ \sigma = (\sigma_1, \sigma_2, \ldots, \sigma_k) \in \left(\mathbb{N}_0^+\right)^k : \sum_{i=1}^{k} \sigma_i = n \right\}.$$

3.2.1. ***SIR*-process.** In this spread of epidemic process there are three different states for the vertices, these are susceptible, infectious and recovered. Susceptible vertices represent individuals who are healthy, but can be infected. Infectious vertices play the role of entities who are infected and infectious, i.e. they can spread the infection to susceptible vertices. Finally, recovered vertices represent the individuals who are not infectious any longer, and immune, i.e. cannot be infected again. The transitions between the different states can be represented by the following flow diagram.

Let us fix $J \in \mathbb{N}^+$, i.e. the total number of steps. For every $j \in [J]$, in the j^{th} step the set of susceptible, infectious and recovered vertices are denoted by $\mathcal{S}_j$, $\mathcal{I}_j$ and $\mathcal{R}_j$, respectively. We will also use the notations $S_j = |\mathcal{S}_j|$, $I_j = |\mathcal{I}_j|$ and $R_j = |\mathcal{R}_j|$. Since the structure of the underlying graph does not change during the spread of the epidemics, for every $j \in [J] \cup \{0\}$, we have $\mathcal{S}_j \cup \mathcal{I}_j \cup \mathcal{R}_j = V$, thus $S_j + I_j + R_j = n$.

Let us define the following (discrete-time) stochastic processes:

$$\mathbf{X} : \{0\} \cup [J] \to \Pi^3, \qquad \mathbf{X} = (X_j)_{j=0}^J = (\mathcal{S}_j, \mathcal{I}_j, \mathcal{R}_j)_{j=0}^J$$

$$\mathbf{Y} : \{0\} \cup [J] \to \Sigma^3, \qquad \mathbf{Y} = (Y_j)_{j=0}^J = (S_j, I_j, R_j)_{j=0}^J .$$

Remark. Notice that we have $Y_j \sim \sigma(X_j)$ for every $j \in [J]$, i.e. Y_j is measurable to X_j, which means that the value of Y_j can be calculated from X_j.

Recall that the σ-algebra generated by the underlying random multi-type graph G is denoted by $\mathcal{F}_n$. Let us define $\mathcal{G}_j = \sigma\left(\mathcal{F}_n, (X_i)_{i=0}^j\right)$, i.e. the σ-algebra generated by G and the first j steps of the spread of epidemic process $\mathbf{X}$.

We assume that the initial sets of the vertices of different states $\mathcal{S}_0$, $\mathcal{I}_0$ and $\mathcal{R}_0$ are given at the beginning. For every $j \in [J]$, in the j^{th} step we have the following:

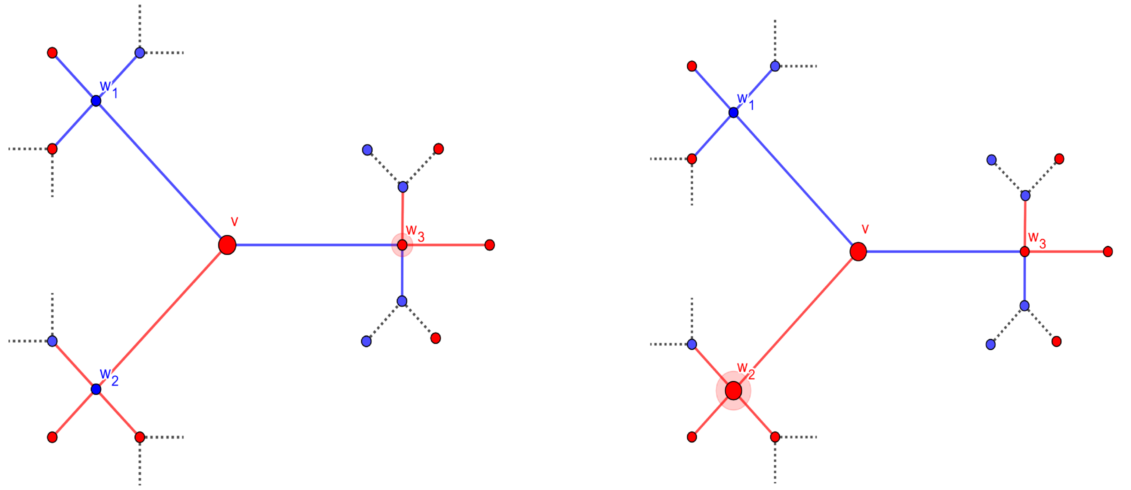
(i): every susceptible vertex $v \in \mathcal{S}_{j-1}$ becomes infectious with probability

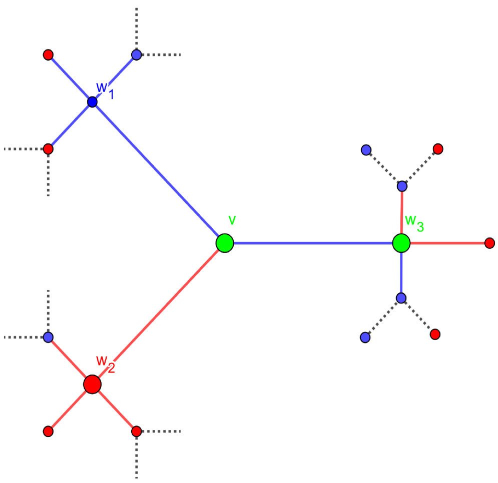
$$\mathbf{P}\left(v \in \mathcal{I}_j \middle| \mathcal{G}_{j-1}\right) = 1 - \prod_{k=1}^{N} (1-p_k)^{i_{j-1}^{(k)}(v)},$$

where $i_{j-1}^{(k)}(v)$ is the number of edges of type k which connect v to an infectious vertex in step $j-1$. Notice that the types of edges do not change, but the states of the vertices may be different over the steps.

(ii): Every infectious vertex $v \in \mathcal{I}_{j-1}$ becomes recovered with probability $\mathbf{P}\left(v \in \mathcal{R}_j \middle| \mathcal{G}_{j-1}\right) = q$.

An illustration of the dynamics of the SIR-process can be seen in the figures below.





Top left: Vertex v may infect w_1 and w_2 with different probabilities, because they are connected to v with different types of edges. Here, vertex w_3 cannot be infected, because it has already been infected.

Top right: Let us assume that vertex w_2 has been infected by v.

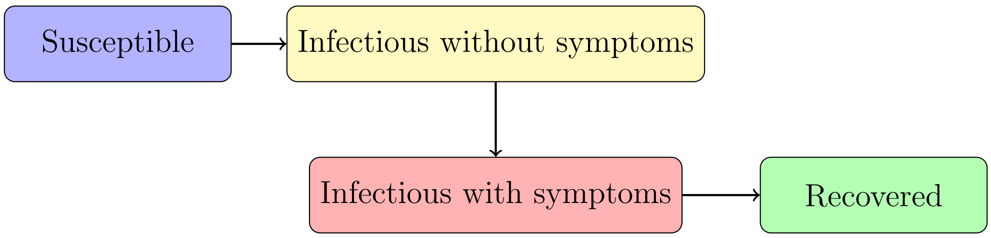
Bottom: After a while infectious vertices become recovered and they can no longer become infectious again. Here, vertices v an w_3 became recovered.

In the SIR-process, the parameters are the propagation probabilities of the spread of infection on the different types of edges $(p_1, p_2, \dots, p_N)$, the probability of recovery (q) (which is the same for all vertices), the finite time horizon (T) and the underlying graph.

3.2.2. Dynamical SI_1I_2R-process. Many infectious diseases are known in which the infected patient does not initially produce symptoms but the patient is contagious. Epidemics caused by such diseases are particularly difficult to control. It may be useful to separate any infectious patient from other people as soon as possible, but in this case the fact of the infection is initially unknown. In this section, we present a modification of the previously discussed model that can be used to model the spread of infectious diseases that have a latency period.

This is a modified version of the SIR-process. There are four different states for the vertices: susceptible, infectious without symptoms (I_1), infectious with symptoms (I_2) and recovered.

The following flow diagram shows the transitions between the different states.



In this model we assume that the edges of the underlying graph model are either open or closed. Closed edges represent the separation of the corresponding points, i.e. the disease cannot spread through these edges.

Let us fix $J \in \mathbb{N}^+$, i.e. the total number of steps. For every $j \in [J]$, in the j^{th} step the set of vertices of state susceptible, infectious without symptoms, infectious with symptoms and recovered are denoted by $\mathcal{S}_j$, $\mathcal{I}_j^{(1)}$, $\mathcal{I}_j^{(2)}$ and $\mathcal{R}_j$, respectively. We will also use the notations $S_j = |\mathcal{S}_j|$, $I_j^{(1)} = |\mathcal{I}_j^{(1)}|$, $I_j^{(2)} = |\mathcal{I}_j^{(2)}|$ and $R_j = |\mathcal{R}_j|$. As discussed in the previous section, for every $j \in [J] \cup \{0\}$, again we have $\mathcal{S}_j \cup \mathcal{I}_j^{(1)} \cup \mathcal{I}_j^{(2)} \cup \mathcal{R}_j = V$, thus $S_j + I_j^{(1)} + I_j^{(2)} + R_j = n$.

We define the following (discrete-time) stochastic processes:

$$\mathbf{X} : \{0\} \cup [J] \to \Pi^4, \qquad \mathbf{X} = (X_j)_{j=0}^J = \left(\mathcal{S}_j, \mathcal{I}_j^{(1)}, \mathcal{I}_j^{(2)}, \mathcal{R}_j\right)_{j=0}^J$$

$$\mathbf{Y} : \{0\} \cup [J] \to \Sigma^4, \qquad \mathbf{Y} = (Y_j)_{j=0}^J = \left(S_j, I_j^{(1)}, I_j^{(2)} R_j\right)_{j=0}^J.$$

As in the previous section, we have $\mathcal{G}_j = \sigma\left(\mathcal{F}_n, (X_i)_{i=0}^j\right)$, i.e. the σ-algebra generated by G and the first j steps of the process $\mathbf{X}$.

We assume that the initial sets of the vertices of different states $\mathcal{S}_0$, $\mathcal{I}_0^{(1)}$, $\mathcal{I}_0^{(2)}$ and $\mathcal{R}_0$ are given. For every $j \in [J]$, in the j^{th} step we have the following:

(i): every susceptible vertex $v \in \mathcal{S}_{j-1}$ becomes infectious without symptoms with probability

$$\mathbf{P}\left(v \in \mathcal{I}_j^{(1)} \middle| \mathcal{G}_{j-1}\right) = 1 - \prod_{k=1}^{N} (1 - p_k)^{i_{j-1}^{(k)}(v)},$$

where $i_{j-1}^{(k)}(v)$ is the number of open edges of type k which connect v with an I_2-vertex in the step $j-1$. Notice that the types of edges remain unchanged, but the states of the vertices may be different over the steps.

(ii): Every I_1-vertex $v \in \mathcal{I}_{j-1}^{(1)}$ becomes an I_2-vertex with probability

$$\mathbf{P}\left(v \in \mathcal{I}_j^{(2)} \middle| \mathcal{G}_{j-1}\right) = r.$$

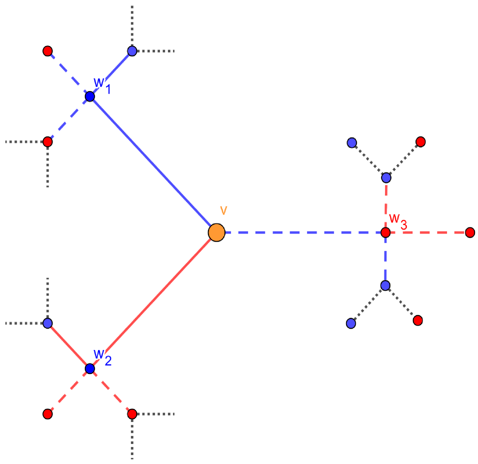
Then, every edge becomes closed which is incident to the vertex v.

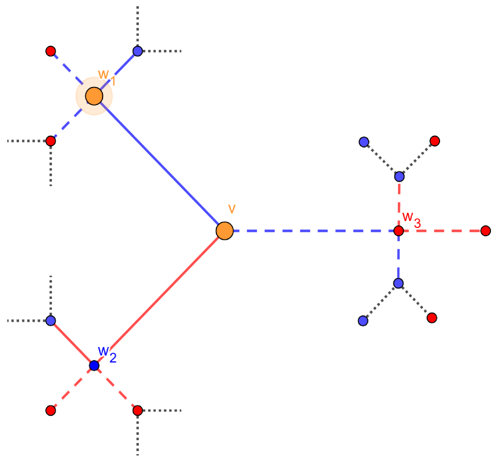
(iii): Every I_2-vertex $v \in \mathcal{I}_{j-1}^{(2)}$ becomes recovered with probability

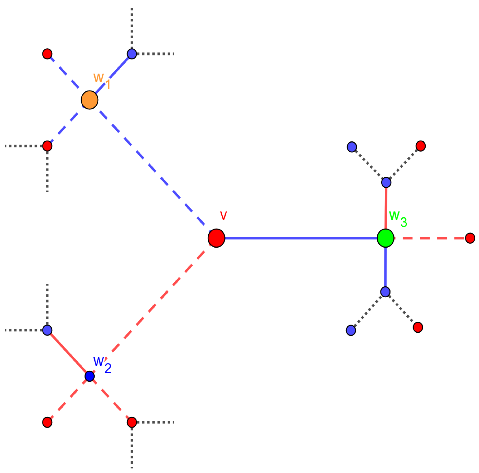
$$\mathbf{P}\left(v \in \mathcal{R}_j \middle| \mathcal{G}_{j-1}\right) = q.$$

Then, every edge becomes open which is incident to v, except those which is incident to an I_2-vertices.

As in the previous section, an illustration of the dynamics of the SI_1I_2R-process can be seen in the figures below.







Top left: Vertex v may infect w_1 and w_2 with different probabilities, because they are connected to v with different types of edges. Here, vertex w_3 has already been infected, thus it is temporarily removed from the rest of the graph like all other I_2-vertices.

Top right: Let us assume that vertex w_1 has been infected by v.

Bottom: After a while I_1-vertices become I_2-vertices and they show symptoms of the disease. Then, they are separated from the rest of the graph, like vertex v in this graph. The I_2-vertices become recovered after a random period of time. Then, they are reconnected to the rest of the graph, except to their neighbours which are I_2-vertices. Here, vertex w_3 became recovered.

In the SI_1I_2R-process, the parameters are the probabilities of the spread of infection on the different types of edges $(p_1, p_2, \ldots, p_N)$, the probability of appearance of symptoms (r), the probability of recovery (q) (which are the same for all vertices), the finite time horizon (T) and the underlying graph.

3.3. Sensitivity analysis of the parameters

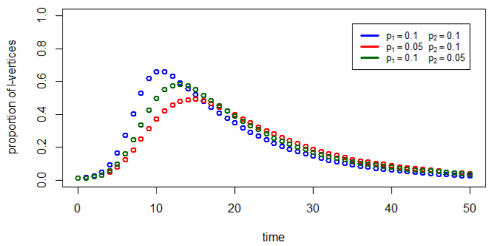
In this section, we examine the results of some stochastic simulations related to the spread of epidemics. We generated random graphs with two types according to the models in scope, then we simulated the spread of epidemics on these graphs.

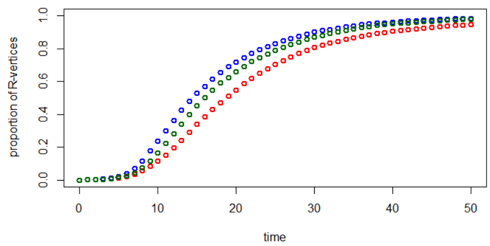
First, we examine the SIR-process for each graph model. Recall that in the SIR-process the symptoms of the infectious disease become visible immediately after the infection. After recovery, the individuals become immune to the disease which means that they can no longer become infected. In the following sections, we have a look at the evolution of the proportion of vertices with different states over time, and we compare the results of different parametrizations. Then, we also have a look at the SI_1I_2R-process. Recall that in the SI_1I_2R-process it takes some random number of steps until the symptoms of the infectious disease become visible after the infection. Similarly to the SIR-process, the individuals become immune to the disease after recovery so that they can no longer become infected.

3.3.1. Preferential attachment graph. In this section, we use the multi-type preferential attachment model as the underlying graph of the process. For a given parametrization, we generated 10 random graphs on 1000 vertices with two types of edges. We had 2000 edges of the first type and 1000 edges of the second one. We examined three different parametrizations of the SIR-process:

	p_1	p_2	q	T
1^{st} parametrization	0.1	0.1	0.1	50
2^{nd} parametrization	0.05	0.1	0.1	50
3^{rd} parametrization	0.1	0.05	0.1	50

For the SIR-processes, 10% of the vertices are infectious and all the other vertices are susceptible at the beginning, and for the SI_1I_2R-processes, 10% of the vertices are infectious without symptoms and the rest of the vertices are susceptible. The trajectories of the average of the scenarios for the three parametrizations can be seen on the following images. The blue, the red and the green trajectories represent the 1^{st}, the 2^{nd} and the 3^{rd} parametrization, respectively.

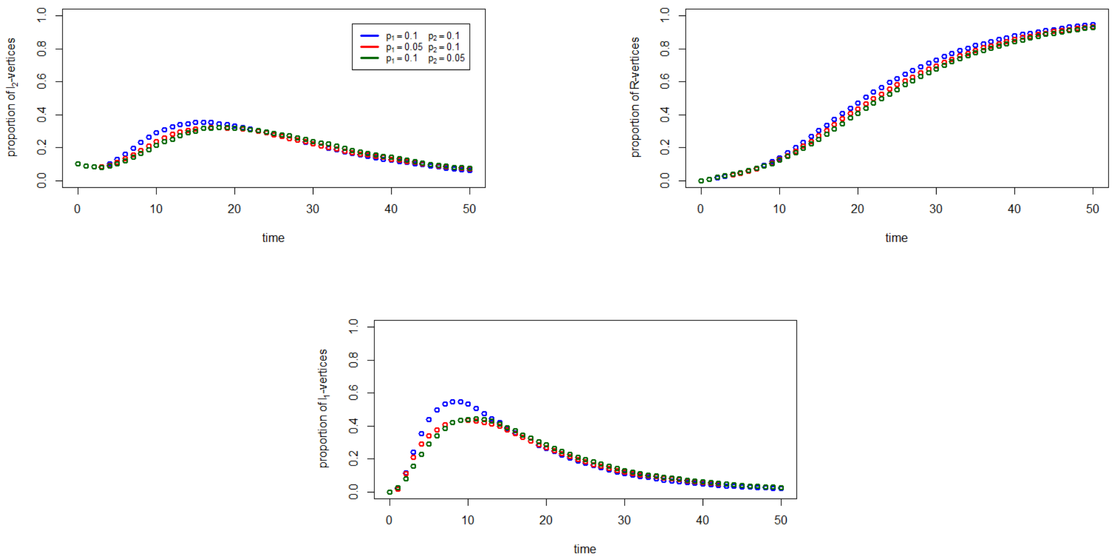




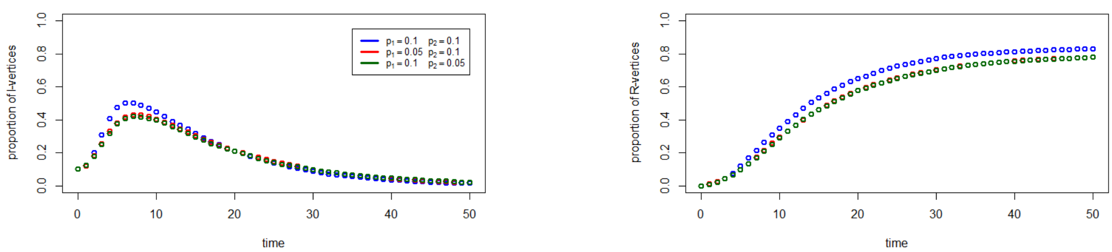
The blue trajectory represents the "single-type case", i.e. we have $p_1 = p_2$. We can see the constant decrease of the proportion of susceptible vertices and the constant increase of the recovered vertices. However, as for the infectious vertices, we observe the rapid increase and then the slow decrease.

Since the number of the edges of the 1^{st} is twice as mush as for the 2^{nd} types, decreasing the probability of propagation on the edges of the 1^{st} type has a more severe impact than the same decrease on the edges of the 2^{nd} type.

We use the same parametrization for the SI_1I_2R-process with an additional parameter, which is $r = 0.1$ (the probability that symptoms of an infected individual become detectable). As a result of the isolation of patients, who are infectious with symptoms, the epidemic curve flattens. Even though there is a latency period, if we isolate people when the symptoms appear, it already improves a lot. This holds in general, regardless of the type of edges. We can see that the introduction of quarantine makes the model less sensitive to changes in the parameters. However, the slowing effect of the quarantine is much severe for the 2^{nd} and 3^{rd} than it is in the single-type case.

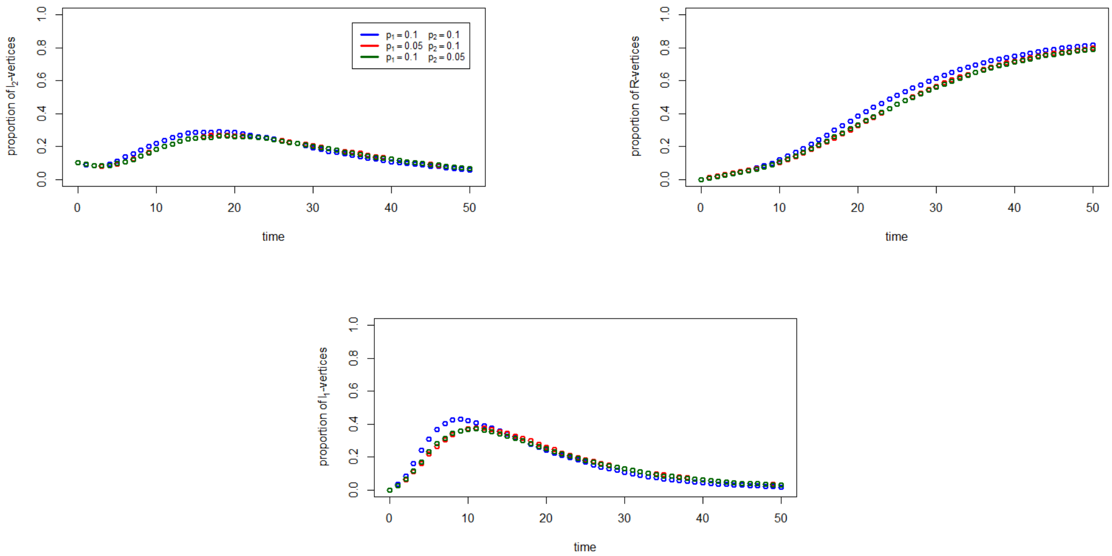


3.3.2. Model of independent edges - version I. In this section, we use the first version of the model of independent edges as the underlying graph of the process. Again, for a given parametrization, we generated 10 random graphs on 1000 vertices with two types of edges. The parametrization of the SIR-process is the same as in the previous section.



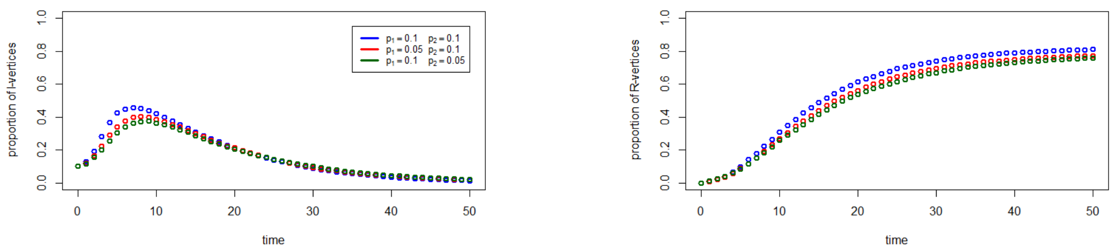
The structure of both versions of the model of independent edges depend on the finite initial configuration. In the simulations, this initial configuration is a graph with two vertices which are connected with 2 edges of the 1^{st} and 1 edge of the 2^{nd} type. Similarly to the preferential attachment model in the previous section, the underlying graph contains more edges of the 1^{st} type than the 2^{nd} type, but the results are less sensitive to changes in the propagation probabilities. We can also see that the spread of the infection is slower even in the single-type case compared to the preferential attachment model. This is due to the fact that edges of the same type are more likely grouped in the model of independent edges.

For the SI_1I_2R-process, we use the same parametrization as in the previous section. In this case, the impact of having two different types for the edges is less severe. The observable data is almost the same as it is in the single-type case.

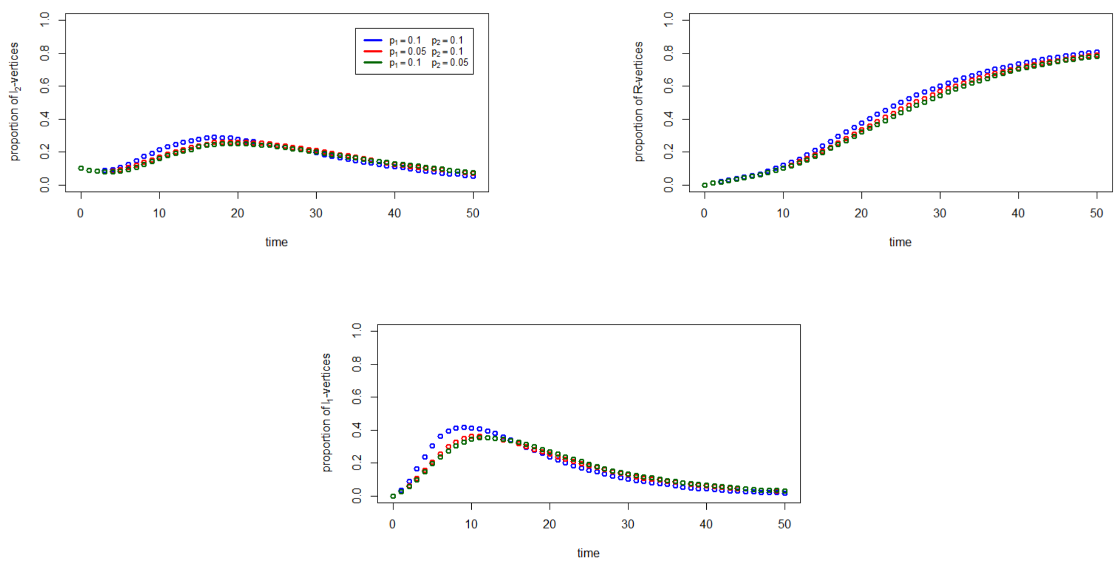


3.3.3. Model of independent edges - version II.

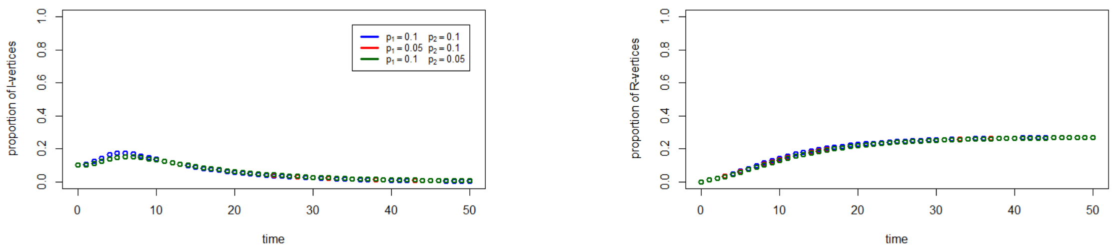
In this section, we use the second version of the model of independent edges as the underlying graph of the process. Again, for a given parametrization, we generated 10 random graphs on 1000 vertices with two types of edges. We have chosen $\lambda = 1$, i.e. the fixed parameter of the graph model.



Although the construction of the underlying graph is different, the results are very similar to the other version of the model of independent edges. In the first version of the model, the (random) degrees of the vertices are binomially distributed, while in the second version they follow Poisson distribution. In this case when the number of vertices is sufficiently large, these distributions have almost the same behaviour.

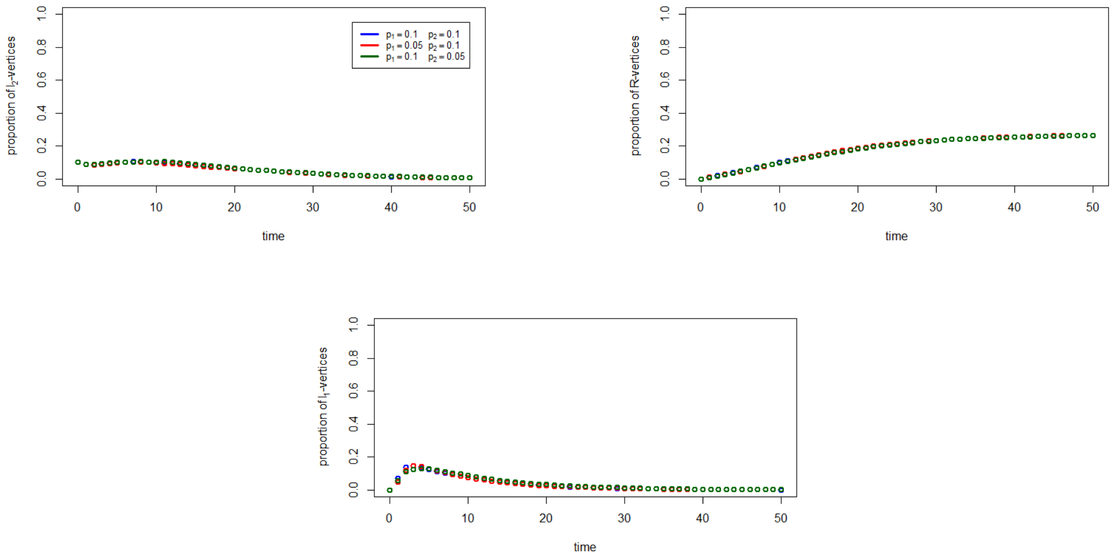


3.3.4. Duplication model. In this section, we use the duplication model as the underlying graph of the process. Again, for a given parametrization, we generated 10 random graphs on 1000 vertices with two types of edges. Similarly to the model of independent edges, the structure of the duplication model depends on the initial finite configuration. In our simulations, the initial configuration contains two independent Erdős–Rényi graphs on 900 vertices, and the probability that a pair of vertices is connected is set to 0.1. Because of the choice of parameters, both graphs contain a single giant connected component with high probability. Then, we apply 100 duplication and deletion steps. The edges of these graphs define the set of edges of different types in the multi-type configuration. The resulting graph is highly clustered, i.e. it mainly consists of independent cliques. This corresponds to a model where we segregate well-isolated groups through restrictive measures, such as the individuals only meet with those who are living in the same household, or the schools are partially open but the classes are isolated. If there are many independent components and only a negligible part of the vertices are infectious, then the results depend on the structure of the typical components and not on the global structure of the graph.



Because of the clustering properties of the underlying graph model, the results are not sensitive to the changes in probabilities of propagation. Initially, 10% of the vertices is infectious (for the SIR-process) or infectious without symptoms (for the SI_1I_2R-process). Within the dense connected components which contain some infectious vertices at the beginning, the epidemic will spread, no matter how small the infection probabilities are.

Because of the clustering properties of the underlying model, the effect of quarantine is less significant. In the connected components, the epidemic will spread among the individuals before we can detect the symptoms and remove the edges between some of the vertices. It also means that if we create isolated bubbles with the help of restrictive measures (e.g. school classes are well-separated), then no more quarantine is necessary, and it is not a problem if there are edges with higher propagation probabilities, the epidemic will spread too a much smaller extent.



3.4. Further research possibilities for epidemic spread

We have seen that the introduction of different types of edges and the presence of segregated groups can have a severe impact on the spread of the epidemic.

Due to the diversity of processes describing the spread of the epidemic, there are many opportunities for further research. One possibility is to introduce a model in which a group of vertices of the graph represent the medical employees. In some applications, the infected individuals require some kind of medical treatment. In this case, the medical employees (doctors, nurses, etc.) are assigned to the infected individuals. We may assume that the medical employees can also be infected, and then, they also require medical treatment. One can examine how much capacity is required in the healthcare for the infected patients to receive appropriate treatment under different parameters of the infection.

Another possibility is to study continuous-time models, i.e. models in which the events describing the infections and recoveries occur on a continuous time scale. For example, we can use exponentially distributed random times that determine when these events occur. Continuous-time models are typically more complex than the discrete-time versions, but they often describe epidemics in reality in a more natural way.

Summary

In this thesis, we investigated the asymptotic properties of a general family of random graphs with multiple type edges.

In Chapter 2 we defined the generalized asymptotic degree distribution for graphs with multiple type edges. Then we defined a general random graph model evolving in discrete time steps by listing some assumptions related to the dynamics of the evolution of the structure. We proved the existence of the generalized asymptotic degree distribution in the general model and provided recurrence equations that are satisfied by this distribution. Then we examined two random graphs in more details that are special cases of the general model: a multi-type version of the well-known Barabási–Albert graph and the model of independent edges. These examples show a new phenomenon, which is the stochastic nature of the asymptotic degree distribution. More precisely the asymptotic degree distribution depends on the asymptotic proportion of edges of different types, which is random due to the dynamics of these two graph models. Then we generalized the scale-free property of random graphs in the presence of different types of edges and showed that the multi-type generalization of the Barabási–Albert graph and the model of independent edges have this property. Afterwards, we defined a perturbed version of the Barabási–Albert model with multiple type edges and compared our results to the previously examined model. Due to the difference between the asymptotic behaviour of the proportion of edges of different types in the two models we obtain a deterministic asymptotic degree distribution in contrast to the non-perturbed model.

In Chapter 3 we investigated the spread of infectious disease on several random graph models with multiple type edges. By using stochastic simulations we examined the epidemic spread process in that case when the probabilities of the propagation depend on the types of the edges of the underlying structure. Empirical results of these simulations are presented by using various type of epidemic spread processes with different parametrizations.

Összefoglalás (in Hungarian)

Ebben az értekezésben többtípusú élekkel rendelkező véletlen gráfok egy általános családjának aszimptotikus tulajdonságait vizsgáltuk meg.

A 2. fejezetben definiáltuk a általánosított aszimptotikus fokszámeloszlást olyan gráfokra, amelyek élei többféle típusba sorolhatók. Ezután definiáltunk egy diszkrét lépésekben fejlődő általános gráfmodellt a fejlődési dinamikára vonatkozó néhány feltétel felsorolásával. Bizonyítottuk az általánosított aszimptotikus fokszámeloszlás létezését és megadtuk rekurziós egyenletek egy rendszerét, amit kielégít a kérdéses eloszlás. Ezután részletesebben megvizsgáltunk két gráfmodellt, amelyek az általános modell speciális esetei: egy többtípusú változata a jól ismert Barabási–Albert-modellnek és a független élek modellje. Ezek a példák rámutatnak egy új jelenségre, ami az aszimptotikus fokszámeloszlás sztochasztikus jellege. Pontosabban, az aszimptotikus fokszámeloszlás függvénye a különböző típusú élek aszimptotikus arányától, ami az említett gráfok fejlődési dinamikája miatt függ a véletlentől. Ezt követően általánosítottuk a skálafüggetlenség fogalmát olyan gráfokra, amelyek többféle típusú éllel rendelkeznek, majd megmutattuk, hogy a többtípusú Barabási–Albert-gráf és a független élek modellje rendelkeznek ezzel a tulajdonsággal. Ezután definiáltuk a többtípusú Barabási–Albert-gráf egy perturbált változatát, és összehasonlítottuk az erre vonatkozó eredményeket a korábban megvizsgált modellével. A különböző típusú élek aszimptotikus arányának eltérő viselkedése miatt azt kaptuk, hogy az aszimptotikus fokszámeloszlás már nem függ a véletlentől a nem-perturbált változattal ellentétben.

A 3. fejezetben megvizsgáltuk fertőző betegségek terjedését különféle gráfmodelleken, amelyek többtípusú élekkel rendelkeznek. Sztochasztikus szimulációk segítségével megvizsgáltuk a járványterjedési folyamatot olyan esetben, amikor a fertőzési valószínűségek függnek a kérdéses gráf éleinek típusaitól. Ezen szimulációk empirikus eredményeit mutattuk be különféle járványterjedési folyamatok és paraméterezések esetén.

Bibliography

[1] **M. A. Abdullah, M. Bode and N. Fountoulakis**, Local majority dynamics on preferential attachment graphs. International Workshop on Algorithms and Models for the Web-Graph. Springer, Cham, (2015).

[2] **D. Acemoglu, V. Chernozhukov, I. Werning and M. D. Whinston**, A multi-risk SIR model with optimally targeted lockdown. Technical report, National Bureau of Economic Research (2020).

[3] **D. Acemoglu, A. Ozdaglar and A. Tahbaz-Salehi**, Systemic risk and stability in financial networks, American Economic Review **105.2** (2015), 564-608.

[4] **H. Amini and A. Minca**, Epidemic Spreading and Equilibrium Social Distancing in Heterogeneous Networks, arXiv preprint arXiv:2007.04210v2 (2020).

[5] **T. Antunović, E. Mossel and M. Rácz**, Coexistence in preferential attachment networks. Combinatorics, Probability and Computing **25.6** (2016), 797–822.

[6] **Á. Backhausz and T. F. Móri**, A random graph model based on 3-interactions. Annales Univ. Sci. Budapest., Sect. Comp., **36**, pp. 41-52, (2012).

[7] **Á. Backhausz and T. F. Móri**, A random model of publication activity, Discrete Appl. Math. **162** (2014), 78–89.

[8] **Á. Backhausz and T. F. Móri**, Asymptotics of a renewal-like recursion and an integral equation, Applicable Analysis and Discrete Mathematics, **8** (2014), 200-223.

[9] **Á. Backhausz and T. F. Móri**, Asymptotic properties of a random graph with duplications. Journal of Applied Probability. **52(2)**, 375-390, (2015).

[10] **Á. Backhausz and T. F. Móri**, Further properties of a random graph with duplications and deletions. Stochastic Models. **32(1)**, 99-120, (2016).

[11] **Á. Backhausz and B. Rozner**, Asymptotic degree distribution in preferential attachment graph models with multiple type edges. *Stochastic Models* (2019), 1-27.

[12] **Á. Backhausz and B. Rozner**, Barabási–Albert random graph with multiple type edges with perturbation. Acta Mathematica Hungarica. **161**(1), 212-229 (2020).

[13] **A-L. Barabási and R. Albert**, Emergence of scaling in random networks, Science, **286** (1999) 509–512.

[14] **M. Barthélemy**, Spatial networks. Physics Reports, **499**(1-3), 1-101, (2011).

[15] **N. Berger, C. Borgs, J. Chayes, A. Saberi**, On the spread of viruses on the internet, *Proceedings of the 16th ACM-SIAM Symposium on Discrete Algorithm (SODA)* (2005).

[16] **B. Bollobás**, *Random graphs.* Cambridge University Press, Cambridge, (2001).

[17] **B. Bollobás, C. Borgs, J. Chayes and O. Riordan**, Directed scale-free graphs, Proceedings of the 14th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), (2003), 132-139.

[18] **B. Bollobás, O. Riordan, J. Spencer and G. Tusnády**, The degree sequence of a scale-free random graph process, Random Structures Algorithms **18** (2001), no. 3, 279–290.

[19] **C. Borgs, J. Chayes, L. Lovász, V. T. Sós, K. Vesztergombi**, Limits of randomly grown graph sequences. Eur. J. Combin. **32(7)**, 985–999. (2011).

[20] **A. D. Broido and A. Clauset**, Scale-free networks are rare, arXiv preprint arXiv:1801.03400, (2018).

[21] **M.-R. Chen and M. Kuba**, On generalized Pólya urn models, J. Appl. Probab. **50** (2013), no. 4, 1169–1186.

[22] **C. Cooper and A. Frieze**, A general model of web graphs, Random Structures Algorithms **22** (2003), no. 3, 311–335.

[23] **F. David, C. Hagendorf and K. J. Wiese**, A growth model for rna secondary structures. Journal of Statistical Mechanics Theory and Experiment, **128**, (2008)

[24] **S. Dereich and P. Mörters**, Random networks with sublinear preferential attachment: degree evolutions, Electron. J. Probab. **14** (2009), no. 43, 1222–1267.

[25] **S. Dereich and M. Ortgiese**, Robust analysis of preferential attachment models with fitness. Combinatorics, Probability and Computing **23.3** (2014), 386–411.

[26] **R. Durrett**, *Random graph dynamics*, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge Univ. Press, Cambridge, (2007).

[27] **P. Erdős and A. Rényi**, On random graphs. I. Publ. Math. Debrecen, 6:290–297, (1959)

[28] **P. Erdős and A. Rényi**, On the evolution of random graphs. Magyar Tud. Akad. Mat. Kutató Int. Közl., 5:17–61, (1960)

[29] **A. Frieze and M. Karoński**, *Introduction to random graphs,* Cambridge Univ. Press, Cambridge, (2015).

[30] **I. Fazekas, Cs. Noszály and A. Perecsényi**, A population evolution model and its applications to random networks, arXiv preprint arXiv:1604.01579, (2016).

[31] **I. Fazekas, Cs. Noszály and A. Perecsényi**, Weights of cliques in a random graph model based on three-interactions. Lithuanian Mathematical Journal **55.2**, (2015), 207–221.

[32] **I. Fazekas, B. Porvázsnyik**, Scale-free property for degrees and weights in an N-interactions random graph model. Journal of Mathematical Sciences **214** (2016) no. 1, 69–82.

[33] **U. Gangopadhyay and K. Maulik**, Stochastic approximation with random step sizes and urn models with random replacement matrices, arxiv preprint, arXiv:1709.00467

[34] **A. Garavaglia, R. van der Hofstad and G. Woeginger**, The dynamics of power laws: fitness and aging in preferential attachment trees, J. Stat. Phys. **168** (2017), no. 6, 1137–1179.

[35] **E. N. Gilbert**, Random graphs. Ann. Math. Statist., 30:1141–1144, (1959).

[36] **F. Hamelin, A. Iggidr, A. Rapaport, G. Sallet and M. O. Souza**, Identifiability and observability of the SIR model with quarantine, arXiv preprint arXiv:2103.04614v1 (2021).

[37] **R. van der Hofstad**, *Random graphs and complex networks*, Cambridge series in statistical and probabilistic mathematics **43** (2016).

[38] **Z. Katona and T. F. Móri**, A new class of scale free random graphs, Statist. Probab. Lett. **76** (2006), no. 15, 1587–1593.

[39] **D. Kiouach, Y. Sabbar and S. El Azami El-idrissi**, New results on the asymptotic behavior of an SIS epidemiological model with quarantine strategy, stochastic transmission, and Lévy disturbance, arXiv preprint arXiv:2012.00875v1 (2020).

[40] **I. Z. Kiss, J. Miller, P. Simon**, *Mathematics of Epidemics on Networks*, Springer (2017).

[41] **M. Kumar and S. Abbas**, Analysis of steady state solutions to an age structured SEQIR model with optimal vaccination, arXiv preprint arXiv:2012.10068v1 (2020).

[42] **S. Laruelle and G. Pagès**, Randomized urn models revisited using stochastic approximation. *The Annals of Applied Probability* **23.4** (2013), 1409-1436.

[43] **J. Neveu**, *Discrete-parameter martingales*, translated from the French by T. P. Speed., revised edition, North-Holland Publishing Co., Amsterdam, (1975).

[44] **L. Ostroumova, A. Ryabchenko and E. Samosvat**, Generalized preferential attachment: tunable powerlaw degree distribution and clustering coefficient, in Algorithms and models for the web graph, 185–202, Lecture Notes in Comput. Sci., 8305, Springer, Cham.

[45] **V. V. Petrov**, *Limit theorems of probability theory*, Oxford Studies in Probability, 4, The Clarendon Press, Oxford University Press, New York, (1995).

[46] **S. Rosengren**, A multi-type preferential attachment model. Internet Mathematics (2018).

[47] **B. Rozner**, Stochastic processes on random graphs with multiple type edges, Annales Universitatis Scientarium Budapestinesis de Rolando Eotvos Nominatae Section Computatorica **52** (2021), 293-312.

[48] **T. Wang and S. Resnick**, Degree growth rates and index estimation in a directed preferential attachment model, arXiv preprint arXiv:1808.01637, (2018).