

Assortative Configuration Random Graphs

ASSORTATIVE CONFIGURATION RANDOM GRAPHS

BY

MOHAMMAD FARZAD POURBABAEE, B.Sc.

A THESIS

SUBMITTED TO THE DEPARTMENT OF MATHEMATICS AND STATISTICS

AND THE SCHOOL OF GRADUATE STUDIES

OF MCMASTER UNIVERSITY

IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

MASTER OF SCIENCE

© Copyright by Mohammad Farzad Pourbabae, April 2015

All Rights Reserved

Master of Science (2015)
(Mathematics & Statistics)

McMaster University
Hamilton, Ontario, Canada

TITLE: Assortative Configuration Random Graphs

AUTHOR: Mohammad Farzad Pourbabaee
B.Sc.,
Sharif University of Technology, Tehran, Iran

SUPERVISOR: Dr. Thomas R. Hurd

NUMBER OF PAGES: ix, 65

To my parents

Abstract

A random graph model is introduced, which is supposed to capture the assortativity effect in various real-world networks. An effect that appears when different types of vertices have non-uniform tendencies to connect to each other. Therefore, given the node- and edge-type distributions P, Q a discrete matching method is developed, so that the empirical distributions asymptotically passes to their given limits P, Q . Applications of martingale convergence methods are exploited to prove the large n limits of the discrete Markov setup. Finally, the relation between inhomogeneous random graph model as another recently studied topic and our model is investigated.

Acknowledgements

First and foremost, I would like to thank my esteemed supervisor Dr. Thomas Hurd, who has supported me since the moment I arrived in Canada. He has always kept track of my researches, and encouraged me on the problems even different from the mainstream of his own works. I should also recall his sharp intuition as well as persistence on mathematical rigor, that helped me all along this thesis. I would also like to express my deep gratitude to Dr. Traian Pirvu, who has never bothered to spend hours of discussing about technical subjects, student life and the proper way of doing research. His continued support led us to draft an article with Dr. Minsuk Kwak. My sincere appreciation is extended to the scholarly faculty members of McMaster Mathematics department, especially Dr. Fred Hoppe, Dr. Andrew Nicas and Dr. Lia Bronsard, whom I enjoyed a lot in their classes.

Also Many thanks to all of my friends for the great time I had with them in Hamilton Hall. Last but not least, my sincere gratitude goes to my sister, who made living in Canada a greater experience for me, and never refrained her supports.

Notations

$\mathbb{E}(X)$ Expected value of random variable X

$\mathbb{P}(A)$ Probability of event A

$\Omega(\cdot)$ $f(n) = \Omega(g(n))$ if for large enough n , $f(n) \geq Cg(n)$

$\omega(\cdot)$ $f(n) = \omega(g(n))$ if $\lim_{n \rightarrow \infty} \frac{g(n)}{f(n)} = 0$

$\Theta(\cdot)$ $f(n) = \Theta(g(n))$ if $\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = c$

$O(\cdot)$ $f(n) = O(g(n))$ if for large enough n , $f(n) \leq Cg(n)$

$o(\cdot)$ $f(n) = o(g(n))$ if $\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = 0$

Contents

| | |
|--|-----------|
| Abstract | iv |
| Acknowledgements | v |
| 1 Introduction | 1 |
| 1.1 Erdős-Renyi Graph | 4 |
| 1.2 Configuration Graph | 6 |
| 1.3 Preferential Attachment Model | 10 |
| 2 Assortative Configuration Graphs | 13 |
| 2.1 Degree Sequence | 14 |
| 2.2 Assortative Wiring | 20 |
| 2.3 Asymptotic Results | 23 |
| 2.3.1 Empirical edge type distribution | 23 |
| 2.3.2 Assortative wiring for undirected graphs | 30 |
| 2.3.3 Asymptotic distribution of self-loops | 32 |
| 3 Inhomogeneous Random Graphs | 37 |
| 3.1 Model Description | 37 |

| | |
|--|-----------|
| 3.2 Relation Between Assortative Configuration Model and IRG | 41 |
| Appendix A Proof of part i of Theorem 2.1.4 | 51 |
| Appendix B Proof of Proposition 2.3.1 | 53 |
| Appendix C Proof of Proposition 2.3.3 | 57 |

List of Figures

| | | |
|-----|--|----|
| 2.1 | directed graph | 15 |
| 2.2 | directed graph with self-loops | 33 |

Chapter 1

Introduction

The theory of random graph has been first investigated by the inspiring paper of Erdős and Renyi [1] in 1960. In their model it is assumed that the number of vertices is fixed and the edges are put between the vertices with given probabilities. As discussed in section 1.1, the significant assumption of homogeneous edge probabilities, that is the probability of existence of any edge is p all over the network, makes the analysis quite straightforward. The combinatorial works of Erdős opened a new perspective for studying large networks by bringing in methods which capture as many precise features of graph ensembles as possible. Usually the complete information about the number of individuals, their connectivity matrix and even more their types of connections is not available. Therefore, it is inevitable that a suitable q ensemble must be proposed of which the real graph would be any realization, i.e an element of this sample space. Depending on the proposed probability space, the accuracy of the reported average graph features might vary. The main topic of random graph theory, is to propose a graph ensemble that matches as closely as possible to characteristics of network types, such as biological networks, social graphs, citation networks and

financial networks. For instance, in the remarkable papers of Barabasi [2, 3, 4], the *scale free* graph arising from *preferential attachment* is shown to be a good fit for biological networks, the world wide web and friendship networks. This model is briefly explained in section 1.3.

It is not always the case that we are in a complete blindness about the topology of our networks. Sometimes we might be provided with partial information such as the degree sequence, i.e the sequence of node degrees, without knowing how these nodes are connected to each other. The *Configuration model* introduced in [5] addresses the problem of generating a random sample from an ensemble of graphs with the specified degree sequence. As briefly discussed in section 1.2, it is an inherent assumption of uniform attachments in it, which leads to equally probable graphs in the ensemble. Therefore the placements of edges in this model ignores the possible node and edge types, which are some extra information we have about the vertices and their connection tendencies. The main contribution of this thesis is to generalize the configuration model to allow edge placement to depend on node and edge type. Similar to the conventional model, the vertices have an associated number of *half-edges* (edges which are not yet paired), and then the pairings of the half-edges are performed according to the prescribed edge-type distribution. The aim is to achieve the given edge- and node-type distributions asymptotically as the graph size gets large. These distributions are usually determined from empirical measurements of a known real network, so as to include rather general graphs of similar types. The mathematical analysis of our matching algorithm in section 2.3 proves that for given distributions (having finite moments) the empirical distribution after pairings is concentrated with high probability around the nominal given distributions. Once we have an algorithm that

generates a random graph as close as desired to a real network. We can proceed with further investigations and study other problems on the random sample and expect to get similar results in real networks. For instance, we might want to find the size of the giant connected component in the graph, the probability of having cascading defaults in financial networks, the dynamics of disease contagion in biological networks, or several other interesting problems. Our model captures the *assortativity* of networks, mainly the tendency for individuals to connect to similar or opposite types. It is called *assortative configuration model* because it allows the half-edge pairings have the so called assortativity.

As another perspective, instead taking the edge-type distribution into account, one can work with node types, which results to the so called *inhomogeneous random graph* described in chapter 3. This model is easier to conceive and more natural to apply to real networks than the configuration model, but it can be regarded as a subset of the former model. The vertices are given with fixed or random weights, and the probability of having an edge between two nodes only depends through the node weights. This model seems more realistic because what one might know in the first place about any graph may be some information about vertices which can then be translated to node types. However in section 3.2 we show how this model can be related to assortative configuration model.

It is the hope of this thesis that the mathematical exploration of assortative graphs will establish the required foundation for applying this model to important real-world networks. Moreover, analytic formulations and their mathematical proofs provide strong basis for the Monte-Carlo simulations performed on random graphs.

1.1 Erdős-Renyi Graph

In this section, we are going to briefly address the so called Erdős-Renyi random graph $G_{n,p}$, which is considered to be the simplest model for random networks. The name originated from the inventors after their pioneering work [1], which established a new field of research. The underlying randomness in this model is homogeneous across the whole network: each pair of the n vertices is connected with occupation probability p ; also called *edge probability*. The described randomness is uniform and unrelated to the type of vertices, which makes the characteristics of this model such as degree distribution, phase transition and graph diameter easier to analyze. Given the independence of edge occupancies, the degree of any fixed node follows a Binomial distribution $\text{Bin}(n, p)$ with success probability p and $n - 1$ trials (total number of neighbours for each vertex). Denoting the degree of node v by D_v , then

$$\mathbb{P}(D_v = k) = \binom{n-1}{k} p^k (1-p)^{n-k}. \quad (1.1)$$

If the edge probability p scales down with n , i.e $p = \lambda/n$, then it is known as a basic probability theory property that the Binomial probabilities converges to a Poisson distribution with parameter λ :

$$\lim_{n \rightarrow \infty} \mathbb{P}(D_v = k) = e^{-\lambda} \frac{\lambda^k}{k!}. \quad (1.2)$$

One can also strengthen the results for empirical node distribution, and justify stricter methods of convergence such as convergence in probability for the number of vertices

with degree k , $u_k^{(n)}$. For any $\varepsilon > 0$,

$$\mathbb{P} \left(\left| \frac{u_k^{(n)}}{n} - e^{-\lambda} \frac{\lambda^k}{k!} \right| \geq \varepsilon \right) \rightarrow 0. \quad (1.3)$$

Since for large k the Poisson probability mass function is smaller than $k^{-\tau}$, for any positive τ , the Erdős-Renyi random graph is not in the category of *scale-free* graphs which have a fat tailed distribution, thus it can not be used as a suitable model for real-world networks which are showing heavy tail properties.

There is a less commonly used version also referred as Erdős-Renyi graph, which again has n vertices, but instead of taking each possible edge with probability p , one picks m of $n(n-1)/2$ possible edges between the vertices uniformly at random. The fixed number of edges in the second realization brings an annoying level of dependence, however when $p = 2m/n(n-1)$ the first model is closely related to the second one [6], and that is why much of the attention is given to the first model.

Since this section is aimed to express an overview of main features of the Erdős-Renyi model, we conclude by explaining an interesting result about a phase transition related to the emergence of a unique giant connected component in the network. The sequence of Erdos-Renyi graphs $G_{n,\lambda/n}$, where λ is constant, has a phase transition at $\lambda = 1$. We write $L_1(G)$ for the number of vertices in the largest component of a graph G . As shown first by Erdos-Renyi in [1], *with high probability* (whp), if $\lambda < 1$ then $L_1(G_{n,\lambda/n})$ is of logarithmic order, if $\lambda = 1$ it is of order $n^{2/3}$, while if $\lambda > 1$ then there is a unique giant component containing $\Theta(n)$ vertices.

The intuition behind this result is that, the average number of neighbours of a randomly chosen vertex is λ and allowing the randomness in edges occupation to be modelled by a Branching process with parameter λ , then the supercritical growth

of $\lambda > 1$ is associated to a giant connected component. A simple proof for the evolution of the random graph based on the Branching process approximation, and the associated phase transition results, is given in [7].

1.2 Configuration Graph

In this section our aim is to investigate a class of graphs which have a fixed degree sequence. In another words, given a sequence of node degrees, we want to construct such graphs with this degree sequence. Also it might be desirable in some situations that the ensemble of constructed graphs with the given degree sequence become equally probable, namely they are uniformly distributed. However not all the degree sequences are *graphical*, in the sense of being realizable as degrees of vertices in a graph. For instance $\mathbf{d} = (d_v)_{v \in [n]}$ is not graphical if the sum of degrees is an odd number.

Definition 1.2.1. A graph is said to be *simple*, if there are no any self-loops or multiple edges between any pairs of vertices. Otherwise it is referred to as a *multigraph*.

We would like to restrict the class of configuration graphs to simple graphs, and the question of obtaining a uniform measure over this class also needs to be answered. The configuration model as a possible realization of random graphs was first introduced in [5, 8] where both are inspired by [9]. Given the sequence of vertex degrees \mathbf{d} , let $\mathcal{H} = \cup_{j=1}^n \mathcal{H}_j$ be a fixed set of $2m = \sum_{j=1}^n d_j$ labelled half-edges, where $|\mathcal{H}_j| = d_j$. The so called *configuration* σ is a partition of \mathcal{H} into m pairs half-edges, which are the edges. A more sensible explanation for the formation of this model is that all vertices are equipped with certain number of stubs (half-edges) equal to their

degrees, then these stubs are attached to each other uniformly at random to build up the connections yielding the graph with prescribed degree sequence $\text{Conf}_n(\mathbf{d})$. So by labelling the set of all half-edges in $\{1, \dots, \ell_n\}$, where $\ell_n = \sum_{v \in [n]} d_v$, a specific configuration σ is denoted as

$$\sigma = \{(i, \sigma(i)) : i \in [\ell_n/2]\}. \quad (1.4)$$

where $\sigma(i)$ is the incident half-edge to i , which together form a full edge. The attachments are going to continue until all half-edges get connected to one another. Following the notations in [10], denote x_i as the half-edge which is picked at step i of pairings, and let y_i be the one to which x_i is paired. Then a pairing scheme $(x_i)_{i \in [\ell_n/2]}$ is said *adaptable* if the choice of x_m only depends on $(x_j, y_j)_{j=1}^{m-1}$. Moreover it is *uniform* when

$$\mathbb{P}(x_m \text{ is paired to } y_m \mid x_m, (x_j, y_j)_{j=1}^{m-1}) = \frac{1}{\ell_n - 2m + 1}. \quad (1.5)$$

At any step of the pairing, the drawn half-edge has no tendency to connect to specific types of half-edges, which means the remaining unpaired half-edges are equally probable to be connected the chosen half-edge. According to the given rule of attachment (1.5), the probability of any configuration σ , which occurs when an index j exists

such that $x_j = i$ and $y_j = \sigma(i)$, can be found:

$$\begin{aligned} \mathbb{P}(\text{Conf}_n(\mathbf{d}) = \sigma) &= \mathbb{P}(x_i \text{ is paired to } y_i \forall i \in [\ell_n/2]) \\ &= \prod_{i=1}^{\ell_n/2} \mathbb{P}(x_i \text{ is paired to } y_i \mid x_i, (x_j, y_j)_{j=1}^{i-1}) \\ &= \prod_{i=1}^{\ell_n/2} \frac{1}{\ell_n - 2i + 1} = \frac{1}{(\ell_n - 1)!!}. \end{aligned}$$

Therefore, all the configurations for a given degree sequence \mathbf{d} are equally probable. Assuming the degree sequence is graphical, following the uniform adaptable pairing scheme leads to a multigraph which is uniformly drawn from the sample space. However the expense of this uniformity is the emergence of possible self-loops and multiple edges, which can not be prevented during the matching algorithm. Moreover by ignoring the labels on stubs, different configurations might be corresponded to a same graph. This means that despite the uniform distribution of $\text{Conf}_n(\mathbf{d})$, their induced graphs are no longer equally probable. One can identify a (multi)graph by a collection of indicator events $(x_{ij})_{i,j \in [n]}$, such that $x_{ij} = 1$ when there exists an edge between vertices i, j . Therefore degree of vertex i is

$$d_i = x_{ii} + \sum_{j \in [n]} x_{ij}. \quad (1.6)$$

Proposition 1.2.2. Let $G = (x_{ij})_{i,j \in [n]}$ be a (multi)graph, which has the degree sequence $\mathbf{d} = (d_i)_{i \in [n]}$, then

$$\mathbb{P}(\text{Conf}_n(\mathbf{d}) = G) = \frac{\prod_{i \in [n]} d_i!}{(\ell_n - 1)!! \prod_{i \in [n]} 2^{x_{ii}} \prod_{1 \leq i < j \leq n} x_{ij}!}. \quad (1.7)$$

Proof. Following the proof of [11], each configuration has probability $1/(\ell_n - 1)!!$, so we only need to find the number of configurations which are equivalent to G upto relabeling the half-edges. Note that by permuting the half-edges incident to each vertex we get the same graph but a different configuration, where the numerator of (1.7) counts the number of permutations. Moreover the factor $x_{ij}!$ compensates for the fact that the multiple edges between i, j can be relabelled without affecting the graph structure but giving rise to different configuration. Likewise $2^{x_{ii}}$'s are due to the self-loops' half-edges which can be ordered reversely without changing the skeleton graph. ■

To get around the self-loops and multi-edges and get a simple graph with prescribed degree distribution, two methods have been presented. The first is the *erased configuration model* in which self-loops are removed and multi-edges are merged, so that the final generated graph is simple. The second is *repeated configuration model* in which the pairing algorithm is repeated until it produces a simple graph. Both methods are first introduced in [12], and the proof of convergence of empirical degree distribution to target distribution for the second case is done by Janson in [13]. In addition Janson derived the probability of a random configuration multigraph being simple and showed that it is bounded from below, which is a required condition for the feasibility of obtaining simple graphs by finite repetitions of the configuration pairings.

1.3 Preferential Attachment Model

In both previous sections, *static* models for random graphs are presented. That is the number of nodes is fixed and the degree sequence is given beforehand, so there is no place for dynamical growth of the model. However, these static models do not capture the evolution of networks. For instance, in social graphs the friendship relations are forming and deforming continuously, while in a citation network the number of authors and papers is growing by time. The *Preferential attachment* model is introduced by Barabasi and Albert in their highly cited paper [2], which gives rise to the random growth of a network and also explains why in their model the degree distribution decays proportionally to k^{-3} , known as power law. The growth mechanism is that at each time t a new vertex with certain number of edges is added to the network, and those edges are sequentially connected to already present vertices. The higher the degree of the receiving vertex is, the more chance of absorbing the edges of the new vertex it has. That is why this model sometimes referred as *rich-get-richer*. Since there are some imprecisions in how this model defined in [2], a rigorous treatment of the results which explores the complete formation from early stages and precise update formulations has been presented in [14]. Let us briefly explain the simplest expositions for the undirected preferential attachment model $PA_t(1)$, where at each step, one edge and its associated vertex are added to the network (the ‘1’ stands for the number of vertices added at each step). The set of vertices at time t is denoted by $\{v_1^{(1)}, \dots, v_t^{(1)}\}$, and degree of $v_i^{(1)}$ in $PA_t(1)$ is denoted by $D_i(t)$. At the beginning of the process, $PA_1(1)$ consists of a single vertex with a self-loop, and conditionally

on $\text{PA}_t(1)$, the growth rule for obtaining $\text{PA}_{t+1}(1)$ is as follows:

$$\mathbb{P}\left(v_{t+1}^{(1)} \sim v_i^{(1)} \mid \text{PA}_t(1)\right) = \begin{cases} \frac{1}{2t+1} & i = t+1 \\ \frac{D_i(t)}{2t+1} & i \in [t] \end{cases} \quad (1.8)$$

As it comes from (1.8) the vertices with higher degrees in $\text{PA}_t(1)$ have more chance of receiving the other side of the incident edge from $v_{t+1}^{(1)}$. There is also a positive probability of appearing a loop at $v_{t+1}^{(1)}$. This model has been generalized in [14] to $\text{PA}_t(m)$, by starting from $\text{PA}_{mt}(1)$ and collapsing the vertices $\{v_1^{(1)}, \dots, v_{mt}^{(1)}\}$ into groups of m . In other words, $v_1^{(1)}, \dots, v_m^{(1)}$ in $\text{PA}_{mt}(1)$ are merged into the vertex $v_1^{(m)}$ in $\text{PA}_t(m)$, similarly $v_{m+1}^{(1)}, \dots, v_{2m}^{(1)}$ are merged into $v_2^{(m)}$ and so forth until getting t vertices in $\text{PA}_t(m)$. Therefore, in $\text{PA}_t(m)$ there are mt edges, and t vertices and there might be some self-loops and multiple edges as well. The next theorem which can be found with its proof in [10] studies the asymptotic empirical degree distribution. Let us define $P_k(t)$ as the proportion of vertices with degree k at time t ,

$$P_k(t) = \frac{1}{t} \sum_{i=1}^t \mathbb{1}_{\{D_i(t)=k\}} \quad (1.9)$$

and denote $p_k = \frac{2m(m+1)}{k(k+1)(k+2)}$ as the limiting degree distribution.

Theorem 1.3.1. For fix $m \geq 1$, there exists a constant $C = C(m)$ such that as $t \rightarrow \infty$

$$\mathbb{P}\left(\max_k |P_k(t) - p_k| \geq C\sqrt{\log t/t}\right) = o(1). \quad (1.10)$$

This theorem confirms the decay of degree distribution according to k^{-3} , which has been stated earlier. The mainstream of all proposed models for preferential attachment is that the connection probabilities are proportional to degree of the receiving

vertex. Variations of this model have been presented to study the directed case [15] and nonlinear attachment probabilities [16].

Chapter 2

Assortative Configuration Graphs

The main contribution of thesis is described in this chapter, which is the extension of configuration model discussed in [1.2](#) to the directed and *assortative* case. The model introduced in this chapter is capable to capture the tendency of linkage between vertices according to their degrees. For instance, in social networks highly connected nodes tend to connect to other high degree nodes, or in biological networks high degree individuals usually connect to low degree nodes, thus vaccinations that target the high degree vertices in biological networks could quickly stop the disease contagion. Measures of assortativity and indications of intense percolation in such networks are studied in [\[17\]](#), but in this chapter we aim to introduce a quite general mathematical foundation, which allows almost all edge and node types behaviour. Therefore, we assume consistent node and edge type distributions P, Q are given, which can almost be any arbitrarily distributions satisfying some nonrestrictive conditions. Then a random discrete matching algorithm for half-edge pairings is suggested and shown to converge to target distributions. The chapter is organized as follows: in [section 2.1](#), general conditions on degree distribution P are imposed and a sampling method

to obtain a graphical bi-degree sequence is studied. In the second section 2.2, an assortative wiring of half-edges according to the edge-type distribution Q is presented, and in the next section 2.3 asymptotic results, which investigate the empirical edge type distribution are proved. Finally in 2.3.3 some applications such as limiting distribution of emerged self-loops is studied.

2.1 Degree Sequence

Chen et al. [18] studied the configuration model for directed graphs, where each vertex comprises two sets of half-edges: in-stubs and out-stubs, and a full edge is established by connecting an out-stub to an available (not yet paired) in-stub. In order to get a graphical realization for a directed graph, the total number of available in-stubs has to match the number of out-stubs, otherwise the matching process does not terminate. However, if we denote the bi-degree of vertex v by (j_v, k_v) , which is drawn at random from joint degree distribution P_{jk} , and let n be the total number of vertices, then

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\sum_{v \in [n]} (k_v - j_v) = 0 \right) = 0.$$

Unless $\mathbb{P}(k_v \neq j_v) = 0$, which typically doesn't happen. In this section we want to study an algorithm, starts off by sampling from degree distribution P , then develop a matching method which yields asymptotically to target edge distribution Q . So if we denote the set of vertices with in-degree j and out-degree k by \mathcal{V}_{jk} , and the set of edges made from pairing a k out-stub to a j in-stub by \mathcal{H}_{kj} ; from k out-stub we mean a half-edge which goes off from a vertex with out-degree k , similar notion also

exists for an in-stub; then in an infinite graph

$$\mathbb{P}(v \in \mathcal{V}_{jk}) = \mathbb{P}(d^-(v) = j, d^+(v) = k) = P_{jk}$$

$$\mathbb{P}(\ell \in \mathcal{H}_{kj}) = Q_{kj},$$

where v and ℓ are randomly chosen vertex and edge respectively. We say a vertex is of type (j, k) , if it belongs to \mathcal{V}_{jk} , and an edge is of type (k, j) , when it belongs to \mathcal{H}_{kj} . For example, as depicted in Fig. 2.1, vertex 1 is of type $(2, 3)$, and the dashed blue edge is of type $(3, 2)$.

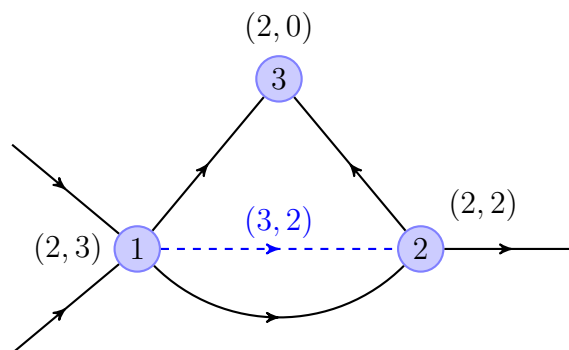


Figure 2.1: directed graph

As it was earlier, P and Q are bivariate node and edge type distributions, which have the following marginals:

$$\begin{aligned} P_k^+ &= \sum_j P_{jk} & P_j^- &= \sum_k P_{jk} \\ Q_k^+ &= \sum_j Q_{kj} & Q_j^- &= \sum_k Q_{kj}. \end{aligned} \tag{2.1}$$

Definition 2.1.1. The node and edge type probability laws P, Q are consistent if:

- The mean in- and out-degrees of nodes are equal: $z := \sum_j j P_j^- = \sum_k k P_k^+$.
- $Q_k^+ = k P_k^+ / z$, $Q_j^- = j P_j^- / z$ for all integer values of j, k .

Given the consistent laws P, Q we can draw n independent copies from P to form the bi-degree sequence $(j_v, k_v)_{v \in [n]}$. Some regularity conditions on the degree distribution, which relate the behaviour of the empirical degree distribution in a finite setup to the subjective probability measure, need to be imposed. Let u_{jk} be the number of type (j, k) nodes for the given bi-degree sequence $(j_v, k_v)_{v \in [n]}$, the shorthand quantities,

$$u_k^+ = \sum_j u_{jk} \quad u_j^- = \sum_k u_{jk} \quad (2.2)$$

are defined to better address the regularity conditions. Moreover, for the given sequence we denote the bi-degree of a uniformly chosen vertex U in $[n]$ by $D_n = (J_n, K_n) = (j_U, k_U)$. Thus one can associate an empirical measure to the finite graph, which is identified by the ratios u_{jk}/n .

Condition 2.1.2. For the given sequence of bi-degrees $(j_v, k_v)_{v \in [n]}$, the following conditions hold:

1. Convergence in distribution of the empirical degree to P -measure: $D_n \Rightarrow D$, where $D \sim P$. In other words:

$$\mathbb{P}(D_n = (j, k)) = \frac{u_{jk}}{n} \rightarrow P_{jk}. \quad (2.3)$$

2. Finite expectation: $\sum_k k P_k^+ = \sum_j j P_j^- = z < \infty$.
3. Convergence of the first moment: Both for the in- and out-degrees the following

convergence assertions hold:

$$\begin{aligned}\mathbb{E}(J_n) &= \frac{1}{n} \sum_j u_j^- \rightarrow \mathbb{E}(J) = \sum_j j P_j^- \\ \mathbb{E}(K_n) &= \frac{1}{n} \sum_k u_k^+ \rightarrow \mathbb{E}(K) = \sum_k k P_k^+.\end{aligned}\tag{2.4}$$

These set of constraints are the minimal assumptions we need in proceeding sections to prove our results. More conditions may be required depending on the type of the necessary boundedness. In the following two sections, our goal is to compose a finite graph with node and edge type empirical distribution as close as possible to P , Q . Therefore, two steps for this construction are studied. First, a realizable bi-degree sequence is obtained by slight modification of the original draw from P . Secondly the assortative wiring process for pairing the half-edges is introduced and shown to reach the target edge distribution Q .

As we have seen before if the bi-degree sequence is sampled arbitrarily from P , then with vanishing probability that sequence would lead to a graphical representation. So starting with $(j_v, k_v)_{v \in [n]}$, the sequence is modified slightly to form a feasible sequence $(\tilde{j}_v, \tilde{k}_v)_{v \in [n]}$. The following clipping algorithm is proposed in [18].

1. Draw a bi-degree sequence $(j_v, k_v)_{v \in [n]}$ from P distribution.
2. Let $D_n = \sum_{v \in [n]} (k_v - j_v)$ be the difference between out-degrees and in-degrees.
3. If $|D_n| > \Delta_n$ then go back to 1, otherwise continue to the next step.
4. Choose $|D_n|$ vertices, $\{i_1, \dots, i_{|D_n|}\}$ uniformly at random from the set of vertices

[n]. Define the modified degree sequence $(\tilde{j}_v, \tilde{k}_v)_{v \in [n]}$ like follows:

$$\begin{aligned} \tilde{j}_v &= \begin{cases} j_v + 1 & D_n \geq 0 \text{ and } v \in \{i_1, \dots, i_{|D_n|}\} \\ j_v & \text{otherwise} \end{cases} \\ \tilde{k}_v &= \begin{cases} k_v + 1 & D_n < 0 \text{ and } v \in \{i_1, \dots, i_{|D_n|}\} \\ k_v & \text{otherwise} \end{cases} \end{aligned} \quad (2.5)$$

The threshold level Δ_n has to be chosen, so that with high probability a randomly drawn bi-degree sequence is accepted in step 3 of the algorithm, namely probability of $\mathcal{D}_n = \{|D_n| \leq \Delta_n\}$ must converge to one. A suitable threshold is found in [18], but for the case when in- and out-degrees are independent from each other. Nevertheless they are not necessarily independent and might come from a joint measure P_{jk} . However, we can exploit their suggested threshold and generalize the proof of Lemma 2.1 in their paper, all that matters is to have $\Delta_n = o(n)$.

Remark 2.1.3. In order to prove $\mathbb{P}(\mathcal{D}_n) \rightarrow 1$, we need stricter tail boundedness conditions for the marginals of P , rather than merely having finite first moment. A sufficient condition is proposed in [18], as there exist slowly varying functions $L^+(\cdot)$ and $L^-(\cdot)$ such that

$$\sum_{k>x} P_k^+ \leq x^{-\alpha} L^+(x) \quad \sum_{j>x} P_j^- \leq x^{-\beta} L^-(x), \quad (2.6)$$

for all $x \geq 0$, where $\alpha, \beta > 1$.

Note that (2.6) ensures that marginals of P have finite moments of order greater than one. Although, it is not necessary for $k_v - j_v$ to have finite variance, but assuming it does we can show for $\Delta_n = n^{1/2+\delta}$, $\delta \in (0, 0.5)$ the acceptance probability in step 3

converges to 1. Equivalently we have:

$$\mathbb{P}(|D_n| \geq n^{1/2+\delta}) \leq \frac{\text{Var}(D_n)}{n^{1+2\delta}} = \frac{n\text{Var}(k_v - j_v)}{n^{1+2\delta}} \rightarrow 0. \quad (2.7)$$

Once we get a suitable sequence $(\tilde{j}_v, \tilde{k}_v)_{v \in [n]}$, we would like its joint distribution to converge asymptotically to i.i.d elements coming from P_{jk} . It is important to note that from the construction of the modified sequence, the vertices bi-degrees are no longer independent from each other. But, it is reasonable to expect some sort of similarity between the original degree sequence and the modified one, because the number of added half-edges in the algorithm is negligible in comparison with the number of vertices. Next proposition will let us to accept the modified sequence as a reasonable candidate for $(j_v, k_v)_{v \in [n]}$.

Theorem 2.1.4. Fix any chosen threshold level $\Delta_n = o(n)$, that satisfies the acceptance condition with probability $1 - o(1)$, then

- (i) For any fixed finite subset of $[n]$, say \mathcal{S} , $(\tilde{j}_v, \tilde{k}_v)_{v \in \mathcal{S}}$ converges in distribution to $(\hat{j}_v, \hat{k}_v)_{v \in \mathcal{S}}$, where (\hat{j}_v, \hat{k}_v) are i.i.d bivariate P -distributed random variables.
- (ii) If \tilde{u}_{jk} denotes the number of type (j, k) nodes after degree adjustment, then the following limit in probability holds:

$$\frac{\tilde{u}_{jk}}{n} \xrightarrow{p} P_{jk}. \quad (2.8)$$

Proof of Theorem 2.1.4. To show convergence in distribution in the first part, we have to prove that for any bounded function $f : \mathbb{N}^{2|\mathcal{S}|} \rightarrow [-M, M]$ the following equation

holds:

$$\left| \mathbb{E} \left(f((\tilde{j}_v, \tilde{k}_v)_{v \in \mathcal{S}}) - f((\hat{j}_v, \hat{k}_v)_{v \in \mathcal{S}}) \right) \right| \rightarrow 0. \quad (2.9)$$

The details of the proof is given in the appendix. For the second part it is important to note that, there are two sources of randomness. First is coming from the randomness lies on choosing the set $\{i_1, \dots, i_{|D_n|}\}$ in the clipping algorithm, and second one is due to the randomness of the original bi-degree sequence. Both parts of this theorem were studied in [18], though some minor modifications have to be made for extending the results to non-independent in- and out-degree case. ■

Remark 2.1.5. One could also demonstrate that condition on occurrence of \mathcal{D}_n the empirical mean of in- and out-degree after the clipping process converges in probability to z , because

$$\begin{aligned} \mathbb{P} \left(\left| \frac{1}{n} \sum_{v \in [n]} \tilde{k}_v - z \right| > \varepsilon \mid \mathcal{D}_n \right) &\leq \mathbb{P} \left(\left| \frac{1}{n} \sum_{v \in [n]} k_v - z \right| + \frac{|D_n|}{n} > \varepsilon \mid \mathcal{D}_n \right) \\ &\leq \frac{1}{\mathbb{P}(\mathcal{D}_n)} \mathbb{P} \left(\left| \frac{1}{n} \sum_{v \in [n]} k_v - z \right| + \frac{\Delta_n}{n} > \varepsilon \right) \\ &\leq \frac{1}{\mathbb{P}(\mathcal{D}_n)} \mathbb{P} \left(\left| \frac{1}{n} \sum_{v \in [n]} k_v - z \right| > \frac{\varepsilon}{2} \right) \rightarrow 0. \end{aligned} \quad (2.10)$$

Equation (2.10) holds because we can make Δ_n/n less than $\varepsilon/2$ for large n .

2.2 Assortative Wiring

The construction here is similar to the conventional configuration model, that is half-edges are initially detached from each other, and as time passes, they are going to

get connected. However, the main difference is the governing rule of wirings, which in the conventional case is uniform among the set of available in- and out-stubs, but in the assortative scheme the matchings are made according to general probability weights, which are not necessarily uniform.

To begin this process, all vertices are equipped with a bi-degree $(j_v, k_v)_{v \in [n]}$ which are i.i.d draws from P_{jk} . If this sequence passes the condition of clipping algorithm, then the modified degree sequence, $\tilde{\mathbf{d}} = (\tilde{j}_v, \tilde{k}_v)_{v \in [n]}$ is generated, otherwise the selection process is repeated. Although the results of next sections are derived for the modified sequence, but for notational simplicity all the variables are represented without tilde sign. Thus by $\mathbf{d} = (j_v, k_v)_{v \in [n]}$ we actually mean the modified feasible sequence $\tilde{\mathbf{d}} = (\tilde{j}_v, \tilde{k}_v)_{v \in [n]}$, which is resulted after the clipping algorithm while the tilde signs are dropped.

Definition 2.2.1. The following sets are defined, so that we can express the configuration's dynamics in terms of them:

- $\mathcal{H}_k^+(t) = \{\text{unpaired out-stubs with out-degree } k \text{ up to time } t\}$,
- $\mathcal{H}_j^-(t) = \{\text{unpaired in-stubs with in-degree } j \text{ up to time } t\}$,
- $\mathcal{E}_{kj}(t) = \{\text{type } (k, j) \text{ edges have been paired up to time } t\}$.

Moreover their cardinality are expressed by non-calligraphic letters, as $H_k^+(t)$, $H_j^-(t)$ and $E_{kj}(t)$.

The set of possible pairings at time t would be $\cup_k \mathcal{H}_k^+(t) \times \cup_j \mathcal{H}_j^-(t)$, with elements given weights proportional to $Q_{kj}/Q_k^+ Q_j^-$. In order to make these weights add up to

one at time t , all of them are normalized by

$$C(t) := \sum_{k,j} H_k^+(t) H_j^-(t) \frac{Q_{kj}}{Q_k^+ Q_j^-}. \quad (2.11)$$

Therefore, at time t the probability of choosing a type (k, j) edge among all possible pairings would be

$$\mathbb{P}(\mathcal{H}_k^+(t) \text{ connects to } \mathcal{H}_j^-(t)) = \frac{H_k^+(t) H_j^-(t) \frac{Q_{kj}}{Q_k^+ Q_j^-}}{C(t)}. \quad (2.12)$$

The wiring process stops at the terminal time $T(n) = \sum_{v \in [n]} k_v$, when there are no more available stubs, and all half-edges are wired by that time. One can model the dynamics of the wiring process acting on the above sets of available stubs by a Markov process with the given probability weights in (2.12), where the Markov state variables are

$$Y(t) = \left\{ (H_k^+(t))_{k=1}^{k_{(n)}}, (H_j^-(t))_{j=1}^{j_{(n)}}, (E_{kj}(t))_{k,j=1}^{k_{(n)}, j_{(n)}} \right\},$$

$$Y(0) = \left\{ (ku_k^+)_{k=1}^{k_{(n)}}, (ju_j^-)_{j=1}^{j_{(n)}}, (\mathbf{0})_{k,j=1}^{k_{(n)}, j_{(n)}} \right\}.$$

$k_{(n)}$ and $j_{(n)}$ are the n -th order statistics, i.e $k_{(n)} = \max_{v \in [n]} \{k_v\}$ and $j_{(n)} = \max_{v \in [n]} \{j_v\}$.

Transition probabilities can easily be found from (2.12). Conditioned on the graphical degree sequence \mathbf{d} and previous matchings $Y(t)$, for each time step $0 < t \leq T(n)$, we have

$$\mathbb{E}(H_k^+(t+1) - H_k^+(t) | \mathbf{d}, Y(t)) = -\frac{H_k^+(t) \sum_j H_j^-(t) \frac{Q_{kj}}{Q_k^+ Q_j^-}}{C(t)}$$

$$\mathbb{E}(H_j^-(t+1) - H_j^-(t) | \mathbf{d}, Y(t)) = -\frac{H_j^-(t) \sum_k H_k^+(t) \frac{Q_{kj}}{Q_k^+ Q_j^-}}{C(t)} \quad (2.13)$$

$$\mathbb{E}(E_{kj}(t+1) - E_{kj}(t) | \mathbf{d}, Y(t)) = \frac{H_j^-(t) H_k^+(t) \frac{Q_{kj}}{Q_k^+ Q_j^-}}{C(t)}.$$

These are the trend equations which represent the average rate of change of the Markov variables, conditioned on the history of previous pairings.

2.3 Asymptotic Results

2.3.1 Empirical edge type distribution

As it turns out the discrete matching algorithm we have defined can be analyzed for large n , using the techniques for “almost martingales” pioneered by Wormald [19]. The idea is to associate a differential equation system to the *trend* equations (2.13), and then demonstrate the claim that Markov variables lie with high probability around the solution of the differential system. In other words, we expect that the differential system could remarkably track the trajectories of the Markov processes. In our problem, it is intuitive to introduce the continuous time functions $(z_k^+(\tau), z_j^-(\tau), e_{kj}(\tau))$ with the intention that

$$H_k^+(t)/n = z_k^+(t/n) + o(1),$$

$$H_j^-(t)/n = z_j^-(t/n) + o(1),$$

$$E_{kj}(t)/n = e_{kj}(t/n) + o(1).$$

Therefore one can conjecture that the *ODE* system

$$\begin{aligned}
w(\tau) &:= \sum_{k \leq k(n), j \leq j(n)} z_k^+(\tau) z_j^-(\tau) \frac{Q_{kj}}{Q_k^+ Q_j^-} \\
\dot{z}_k^+(\tau) &= -\frac{z_k^+(\tau) \sum_{j \leq j(n)} z_j^-(\tau) \frac{Q_{kj}}{Q_k^+ Q_j^-}}{w(\tau)}, \quad z_k^+(0) = kP_k^+ \mathbf{1}_{\{k \leq k(n)\}} \\
\dot{z}_j^-(\tau) &= -\frac{z_j^-(\tau) \sum_{k \leq k(n)} z_k^+(\tau) \frac{Q_{kj}}{Q_k^+ Q_j^-}}{w(\tau)}, \quad z_j^-(0) = jP_j^- \mathbf{1}_{\{j \leq j(n)\}} \\
\dot{e}_{kj}(\tau) &= \frac{z_k^+(\tau) z_j^-(\tau) \frac{Q_{kj}}{Q_k^+ Q_j^-}}{w(\tau)}, \quad e_{kj}(0) = 0
\end{aligned} \tag{2.14}$$

could be the one whose solution is approximating the trajectories. The choice of the initial conditions actually seems natural, because from part [ii](#) of theorem [2.1.4](#) we expect $H_k^+(0) = nkP_k^+(1 + o(1))$, which is consistent with taking the deterministic initial condition $z_k^+(0) = kP_k^+$. Summarizing the representation of this system, one could consider

$$\begin{aligned}
\dot{x}(\tau) &= f(x(\tau)) \quad \text{where } f : \mathbb{R}^{k(n)+j(n)} \rightarrow \mathbb{R}^{k(n)+j(n)+k(n)j(n)} \\
x(0) &= \left(\{zQ_k^+\}_{k=1}^{k(n)}, \{zQ_j^-\}_{j=1}^{j(n)}, \{\mathbf{0}\}_{k,j=1}^{k(n),j(n)} \right),
\end{aligned} \tag{2.15}$$

where

$$x(\tau) := \left(\{z_k^+(\tau)\}_{k=1}^{k(n)}, \{z_j^-(\tau)\}_{j=1}^{j(n)}, \{e_{kj}(\tau)\}_{k,j=1}^{k(n),j(n)} \right),$$

as an alternative description of differential equations with the understood function f from R.H.S of (2.14). Moreover, the domain for this system is the bounded open set

$$\Omega_n^{(\varepsilon)} = \left\{ \left(\tau, \{z_k^+\}_{k=1}^{k(n)}, \{z_j^-\}_{j=1}^{j(n)}, \{e_{kj}\}_{k,j=1}^{k(n),j(n)} \right) : -\varepsilon < \tau < z(1 - \varepsilon), \quad -\varepsilon < z_k^+ < k \quad \forall k, \right. \\ \left. -\varepsilon < z_j^- < j \quad \forall j, -\varepsilon < e_{kj} < k \wedge j \quad \& \quad \sum_{j \leq j(n), k \leq k(n)} z_k^+ z_j^- \frac{Q_{kj}}{Q_k^+ Q_j^-} > \varepsilon \right\}. \quad (2.16)$$

Over this region, the denominators of autonomous functions of the differential system are bounded away from zero, so there is Lipschitz continuity over this set. Namely, there exists $c = c(\varepsilon)$ such that

$$\|f(x) - f(y)\| < c \|x - y\|$$

for every $x, y \in \Omega_n^{(\varepsilon)}$. Because of the equivalence of norms, let us agree on l_∞ norm from now on. Next proposition gives the solution for this system over the proposed region.

Proposition 2.3.1. Given the $\varepsilon > 0$, the system of ordinary differential equations in (2.14) has the unique solution

$$\begin{aligned} z_k^+(\tau) &= zQ_k^+ \left(1 - \frac{\tau}{z}\right) + o(1) \\ z_j^-(\tau) &= zQ_j^- \left(1 - \frac{\tau}{z}\right) + o(1) \\ e_{kj}(\tau) &= Q_{kj}\tau + o(1), \end{aligned} \quad (2.17)$$

for the time interval $\tau \in [0, z)$, which is extendible to the boundary of $\Omega_n^{(\varepsilon)}$. (see appendix for the proof)

Remark 2.3.2. The $o(1)$ factor in equation (B.9) of the proof of proposition 2.3.1 comes from the tail condition of Q , which is clearly going to zero as n grows. However, further on we may need to increase the Lipschitz factor with n , thus in order to get an $o(1)$ error in $x(\tau)$, we need to impose stricter tail condition on Q . For instance, if $\sum_{k>k_{(n)}} Q_k^+$ decays like $1/\log n$, then c can be chosen as $\log \log n/z$ and still we get an $o(1)$ error term in (B.9). Because the denominators are ε away from zero on $\Omega_n^{(\varepsilon)}$, then the Lipschitz constant would be of order $1/\varepsilon^2$. Thus, we can choose ε according to the desired Lipschitz constant, that gives us an $o(1)$ error in (B.9). From equation (B.9) and the fact $\tau \leq z$ on $\Omega_n^{(\varepsilon)}$, we get

$$\|x(\tau) - x^{(1)}(\tau)\| \leq \frac{z \log n}{\log \log n} \times \frac{1}{\log n} = o(1).$$

Thus the approximation error remains as $o(1)$.

Once the solution of (2.14) is known, we could go one step further to approximate the Markov processes with their continuous counterparts, by using Wormald's argument. However, there are two obstructions here. First, the proof for approximating the trajectories of random processes with the solution of corresponding differential system is brought for finite (n independent) number of variables in [19], but as mentioned in this paper the issue can be resolved without any trouble, and Amini has shown it in [20]. Second, as pointed out in [19] the growth of Lipschitz constant prevents us from choosing a domain Ω which extends to the natural end of the process. That is the reason $\Omega_n^{(\varepsilon)}$ is set to stay ε away from the actual terminal point of the process (when $\tau \rightarrow z$). But, eventually we would like to approach $\varepsilon \rightarrow 0$, and that forces the Lipschitz constant to blow up. So we have to prove a version of the theorem which allows the growth of the Lipschitz constant as n increases, while being able to

control the error probability as a function of Lipschitz factor. Therefore, an improved version of this theorem is expressed here, but proved in appendix.

Proposition 2.3.3. Let $b = b(n)$ be the number of random variables, and $\{Y_\ell^{(n)}(t)\}_{\ell=1}^b$ is a sequence of stochastic processes such that $0 \leq Y_\ell^{(n)}(t) \leq C_0 n$. Let \mathcal{F}_t be the σ -field induced by these processes upto time t . In addition, suppose there exists a bounded connected open set $D(n) \subseteq \mathbb{R}^{b+1}$ such that

$$\left\{ (0, z_1, \dots, z_b) : \mathbb{P} \left(Y_\ell^{(n)}(0) = nz_\ell, \ell = 1, \dots, b \right) \neq 0 \right\} \subseteq D(n), \quad (2.18)$$

and is a function of n . If $T_{D(n)} = \inf \left\{ t \geq 0 : \left(t/n, Y_1^{(n)}(t)/n, \dots, Y_b^{(n)}(t)/n \right) \notin D(n) \right\}$ would be the stopping time, which alarms the exit of variables from the controlled domain, and the following conditions hold for $t \leq T_{D(n)}$:

1. (Bounded Increments) for some constant $\beta(n)$:

$$\max_{1 \leq \ell \leq b} \left| Y_\ell^{(n)}(t+1) - Y_\ell^{(n)}(t) \right| \leq \beta(n).$$

2. (Trend) For some $\lambda_1(n) = o(1)$ and for all $1 \leq \ell \leq b$:

$$\left| \mathbb{E} \left(Y_\ell^{(n)}(t+1) - Y_\ell^{(n)}(t) \mid \mathcal{F}_t \right) - f_\ell \left(t/n, Y_1^{(n)}(t)/n, \dots, Y_b^{(n)}(t)/n \right) \right| \leq \lambda_1(n). \quad (2.19)$$

3. (Lipschitz Continuity) the family of functions $\{f_\ell\}$, $\ell = 1 \dots b(n)$ are Lipschitz continuous on $D(n)$ with all Lipschitz constants uniformly bounded by $\mu(n)$ (typically, this constant grows as n increases).

then it is possible to conclude:

- (i) For any initial condition $(0, \hat{z}_1, \dots, \hat{z}_b) \in D(n)$, the following system of differential equations has a unique solution on $D(n)$ which passes the initial points. i.e $z_k(0) = \hat{z}_k$, $k = 1, \dots, b$. Moreover, this solution is extendible to the boundary of $D(n)$.

$$\frac{dz_k}{ds} = f_k(s, z_1, \dots, z_b) \quad (2.20)$$

- (ii) For the given $\lambda > \lambda_1$ with $\lambda = o(1)$, and sufficient large C we can approximate the random processes with the solution to the above ODE such that:

$$\mathbb{P} \left(\left| Y_k^{(n)}(t) - nz_k(t/n) \right| \geq \lambda n \right) = O \left(\frac{b\rho\sigma}{\lambda} e^{-\frac{n\lambda^3}{\rho^3}} \right); \quad \rho(n) := \beta(n)\mu(n) \quad (2.21)$$

In other words with probability $1 - O \left(\frac{b\rho\sigma}{\lambda} e^{-\frac{n\lambda^3}{\rho^3}} \right)$ we have:

$$Y_k^{(n)}(t) = nz_k(t/n) + O(\lambda n), \quad (2.22)$$

for all $k \in \{1, \dots, b\}$ and $0 \leq t/n \leq \sigma(n)$. z_k is the solution of equation in (2.20), with initial condition $z_k(0) = Y_k^{(n)}(0)/n$. And $\sigma(n)$ determines the distance to the boundary of $D(n)$.

$$\sigma(n) = \sup\{s : \text{dist}_\infty((s, z_1(s), \dots, z_b(s)), \partial(D(n))) > C\lambda\} \quad (2.23)$$

Throughout the proof it is assumed that $\sigma(n) = O(1)$.

Now we are in a position to give a rough estimate about the eventual ratio of type (k, j) edges at the terminal time $T(n)$, which is basically the main result of this

thesis.

Theorem 2.3.4. The ratio of type (k, j) edges concentrates with high probability around the nominal edge type distribution Q_{kj} . To be more precise for any chosen $\lambda = o(1)$ with high probability of $1 - O\left(\frac{b\rho\sigma}{\lambda}e^{-n\lambda^3/\rho^3}\right)$ we have

$$\frac{E_{kj}(T(n))}{n} = zQ_{kj} + o(1), \quad (2.24)$$

where $\rho = \rho(n)$ is the uniform Lipschitz constant over $\Omega_n^{(\varepsilon)}$, and $b = b(n)$ is the dimension of system in (2.14).

Proof of Theorem 2.3.4. Since the denominators of autonomous functions in (2.14) are at least ε on $\Omega_n^{(\varepsilon)}$, the Lipschitz constant is bounded by $\rho = 1/\varepsilon^2$. Therefore, depending on the decay speed of Q 's tail, we could pick $\varepsilon(n)$ such that the error term in solution of (2.14) is kept at order $o(1)$, like the way we did in remark 2.3.2. Now because all the requirements of proposition 2.3.3 are satisfied by $\beta(n) = 1$ and $\lambda_1 = 0$, one can exploit it to establish the concentration region of E_{kj} for any $t/n \leq \sigma(n)$ as

$$\begin{aligned} \frac{E_{kj}(t)}{n} &= e_{kj}(t/n) + O(\lambda) \\ &= (t/n)Q_{kj} + o(1) + O(\lambda) = (t/n)Q_{kj} + o(1), \end{aligned} \quad (2.25)$$

where $\sigma(n)$ specifies the maximum time such that continuous solutions of (2.14) are at $C\lambda$ distance from the boundary of $\Omega_n^{(\varepsilon)}$. The only constraint in the set $\Omega_n^{(\varepsilon)}$, which is violated as the process keeps going on, is the approaching of τ to $z(1-\varepsilon)$. Therefore the boundary of $\Omega_n^{(\varepsilon)}$ is reached at $\hat{\tau} = z(1-\varepsilon)$, where $e_{kj}(\hat{\tau}) = zQ_{kj}(1-\varepsilon) + o(1)$.

Let us evaluate z_k^+ and z_j^- at the boundary:

$$\begin{aligned} z_k^+(\hat{\tau}) &= zQ_k^+\varepsilon + o(1) \\ z_j^-(\hat{\tau}) &= zQ_j^-\varepsilon + o(1) \end{aligned} \tag{2.26}$$

The complication arises right after $\hat{\tau}$, where we can not use (2.25), because τ is no longer in $\Omega_n^{(\varepsilon)}$. However, the point is that since $\varepsilon(n) = o(1)$ then for times around $n\hat{\tau}$, the remaining number of k out-stubs from (2.26) are going to be

$$\frac{H_k^+(t)}{n} = zQ_k^+\varepsilon + o(1) = o(1). \tag{2.27}$$

And we know that the difference between $E_{kj}(T(n))$ and $E_{kj}(t)$ is bounded by this number of k out-stubs, namely:

$$\frac{E_{kj}(T(n))}{n} - \frac{E_{kj}(t)}{n} \leq \frac{H_k^+(t)}{n} = o(1) \implies \frac{E_{kj}(T(n))}{n} = zQ_{kj} + o(1).$$

■

Remark 2.3.5. In order to obtain a negligible error probability in theorem 2.3.4, one has to choose λ , such that the exponent of e grows to infinity as n approaches infinity. And this needs to be done in accordance with the choice of $\rho(n)$, which has already been determined by picking suitable $\varepsilon(n)$ related to the tail of Q .

2.3.2 Assortative wiring for undirected graphs

As a more simplified case, one can observe that the results of the previous section hold for undirected graphs. Supposing the initial degree sequence $\mathbf{d} = (d_v)_{v \in [n]}$ satisfies

the only required condition for being graphical, namely $\sum_{v \in [n]} d_v$ be an even number, then the undirected graph can be constructed by assortative pairings according to the edge type distribution Q_{kj} , that in this case is symmetric, i.e $Q_{kj} = Q_{jk}$. In contrast to the directed case the degree distribution is univariate distribution, since there are no in- and out-degrees. The corresponding consistency between P, Q would be $Q_k = kP_k/z$ where z is the mean degree and Q_k is the marginal of Q , which is indeed independent of index taken to be fixed. The half-edges are connected to the vertices but not to each other at the beginning of the wiring process, however they are going to connect gradually during the wiring process. The set of un-paired half-edges upto time t , which are associated to a k -degree node is denoted by $\mathcal{H}_k(t)$, and its cardinality is shown by $H_k(t)$. Likewise before

$$\begin{aligned} \mathbb{P}(\mathcal{H}_k(t) \text{ connects to } \mathcal{H}_j(t)) &= \frac{2H_k(t)H_j(t)\frac{Q_{kj}}{Q_kQ_j}}{C(t)} \quad \text{for } k \neq j \\ \mathbb{P}(\mathcal{H}_k(t) \text{ connects to } \mathcal{H}_k(t)) &= \frac{H_k(t)(H_k(t) - 1)\frac{Q_{kk}}{Q_k^2}}{C(t)}, \end{aligned} \quad (2.28)$$

where the weighting factor is $C(t) = \sum_{k \neq j} H_k(t)H_j(t)\frac{Q_{kj}}{Q_kQ_j} + \sum_k H_k(t)(H_k(t) - 1)\frac{Q_{kk}}{Q_k^2}$. Similarly, the dynamics of wirings can be modelled by a Markov process with the given probabilities in (2.28), and the state variables are

$$\begin{aligned} Y(t) &= \left\{ (H_k(t))_{k=1}^{k(n)}, (E_{kj}(t))_{k,j=1}^{k(n),j(n)} \right\} \\ Y(0) &= \left\{ (H_k(0))_{k=1}^{k(n)}, (0)_{k,j=1}^{k(n),j(n)} \right\}. \end{aligned}$$

Such that $k(n) = j(n) = \max_{v \in [n]} d_v$ is the n -th order statistics, and $E_{kj}(t)$ is the number of type (k, j) edges paired up to time t ($E_{kj}(t) = E_{jk}(t)$). The terminal time of wiring is $T(n) = \sum_{v \in [n]} d_v$, and for all $t \leq T(n)$ the following trend equations hold:

$$\begin{aligned}
\mathbb{E}(H_k(t+1) - H_k(t) | \mathbf{d}, Y(t)) &= -\frac{H_k(t) \sum_{j \neq k} H_j(t) \frac{Q_{kj}}{Q_k Q_j}}{C(t)} - \frac{2H_k(t)(H_k(t) - 1) \frac{Q_{kk}}{Q_k^2}}{C(t)} \\
\mathbb{E}(E_{kj}(t+1) - E_{kj}(t) | \mathbf{d}, Y(t)) &= \frac{2H_j(t)H_k(t) \frac{Q_{kj}}{Q_k Q_j}}{C(t)} + \frac{H_k(t)(H_k(t) - 1) \frac{Q_{kk}}{Q_k^2}}{C(t)} \mathbb{1}_{\{k=j\}}.
\end{aligned} \tag{2.29}$$

Now one can express the associated continuous differential system to (2.29) together with a suitable region for the continuous variables like (2.16):

$$\begin{aligned}
w(\tau) &:= \sum_{k \neq j} z_k(\tau) z_j(\tau) \frac{Q_{kj}}{Q_k Q_j} + \sum_k z_k^2(\tau) \frac{Q_{kk}}{Q_k^2} \\
\dot{z}_k(\tau) &= -\frac{z_k(\tau) \sum_{j \neq k} z_j(\tau) Q_{kj} / Q_k Q_j}{w(\tau)} - \frac{2z_k^2(\tau) Q_{kk} / Q_k^2}{w(\tau)} \\
\dot{e}_{kj}(\tau) &= \frac{2z_j(\tau) z_k(\tau) Q_{kj} / Q_k Q_j}{w(\tau)} + \frac{z_k^2(\tau) Q_{kk} / Q_k^2}{w(\tau)} \mathbb{1}_{\{k=j\}}.
\end{aligned} \tag{2.30}$$

Following the similar analysis as proof of theorem 2.3.4, the terminal number of type (k, j) edges could be found with Wormald's method.

2.3.3 Asymptotic distribution of self-loops

After developing the mathematical foundation for assortative configuration, many other subgraphs can now be explored. Emergence of self-loops as depicted in Fig. 2.2 is one of the nuisance subgraphs that may happen over the process of random wirings. So it is plausible to count the number of them for each configuration. In [10] the argument for finding the asymptotic joint distribution of self-loops and multiple edges is presented for *conventional* configuration graphs. It is proved to be a joint *Poisson* distribution with two distinct parameters, related independently to self-loops and

multiple edges. Here we are going to find the asymptotic distribution of the number of self-loops for the assortative case, which is claimed to be a *Poisson* distribution with prescribed mean. The next theorem will depict our claim.

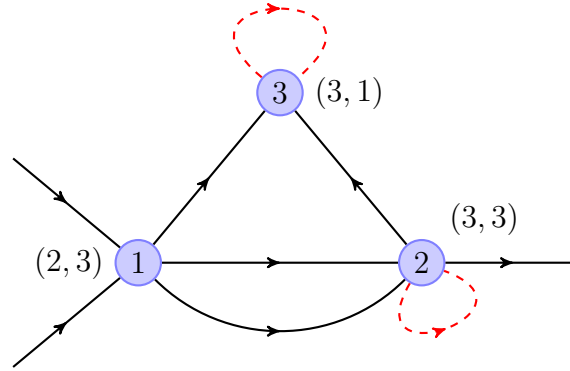


Figure 2.2: directed graph with self-loops

Theorem 2.3.6. If S_n is the number of self loops at the end of the process of assortative configuration, then S_n converges in distribution to a *Poisson* random variable, i.e:

$$S_n \implies S, \text{ such that } S \sim \text{Pois}(\mu) \ ; \ \mu = z \sum_{jk} \frac{P_{jk}}{P_j^- P_k^+} Q_{kj}$$

Proof of Theorem 2.3.6. To characterize the events yielding to a self loop, we define the index set \mathcal{S} :

$$\mathcal{S} = \{(\ell_v^+ \ell_v^-, v) : v \in [n], \ell_v^+ \in [k_v], \ell_v^- \in [j_v]\}.$$

Every self loop is an element of this set, such that ℓ_v^+ is wired to ℓ_v^- . Let us denote the set of the emerged self loops by $\mathcal{L}_n \subset \mathcal{S}$. So one can express S_n as the sum of

indicator functions over this index set (defining $I_s := \mathbb{1}_{\{s \in \mathcal{L}_n\}}$):

$$S_n = \sum_{s \in \mathcal{S}} \mathbb{1}_{\{s \in \mathcal{L}_n\}} = \sum_{s \in \mathcal{S}} I_s.$$

In order to show the convergence in distribution, it is helpful to find the r^{th} factorial $((S_n)_r := S_n! / (S_n - r)!)$ moment, i.e. $\mathbb{E}((S_n)_r)$:

$$\begin{aligned} \mathbb{E}((S_n)_r) &= r! \sum_{(s_1, \dots, s_r) \subset \mathcal{S}} \mathbb{P}(I_{s_1} = \dots = I_{s_r} = 1) \\ &= r! \sum_{(s_1, \dots, s_r) \subset \mathcal{S}} \prod_{t=1}^r \mathbb{P}\left(I_{s_t} = 1 \mid \prod_{i=1}^{t-1} I_{s_i} = 1\right) \\ &= \sum_{s_1} \mathbb{P}(I_{s_1} = 1) \sum_{s_2 \neq s_1} \mathbb{P}(I_{s_2} = 1 \mid I_{s_1} = 1) \dots \sum_{s_r \notin \{s_1, \dots, s_{r-1}\}} \mathbb{P}\left(I_{s_r} = 1 \mid \prod_{i=1}^{r-1} I_{s_i} = 1\right). \end{aligned} \quad (2.31)$$

To find the above probability, we would better to start by calculating the simplest case, namely $\sum_s \mathbb{P}(I_s = 1)$.

$$\begin{aligned} \sum_{s \in \mathcal{S}} \mathbb{P}(I_s = 1) &= \sum_{s \in \mathcal{S}} \frac{Q_{k_v j_v} / Q_{k_v}^+ Q_{j_v}^-}{\sum_{w \in [n]} j_w Q_{k_v j_w} / Q_{k_v}^+ Q_{j_w}^-} \\ &= \sum_{v \in [n]} k_v j_v \frac{Q_{k_v j_v} / Q_{j_v}^-}{\sum_{w \in [n]} j_w Q_{k_v j_w} / Q_{j_w}^-} \\ &= \sum_{v \in [n]} k_v j_v \frac{Q_{k_v j_v} / Q_{j_v}^-}{n \left[\mathbb{E}\left(\frac{J_{k_v J}}{Q_J}\right) + o(1) \right]} \\ &= \frac{1}{nz} \sum_{v \in [n]} k_v j_v \frac{Q_{k_v j_v}}{Q_{k_v}^+ Q_{j_v}^-} (1 + o(1)) \\ &= \frac{z}{n} \sum_{v \in [n]} \frac{Q_{k_v j_v}}{P_{k_v}^+ P_{j_v}^-} (1 + o(1)) = [z \sum_{jk} \frac{P_{jk}}{P_j^- P_k^+} Q_{kj}] (1 + o(1)) = \mu(1 + o(1)) \end{aligned}$$

We want to prove that the product of the terms in (2.31) is roughly equal to $(\sum_s \mathbb{P}(I_s = 1))^r$. In the remaining of the proof we are going to make this approximation more accurate, and finally we show that it holds within an $o(1)$ error term. The subtlety arises in computing conditional probabilities, when we know some half-edges have already been wired and are not counted as the available candidates of wiring, so we have to deduct their effects from the weighting factor in the denominators. Let us find one of the conditional probabilities for $q \leq r$:

$$\begin{aligned}
\sum_{s_q \notin \{s_1, \dots, s_{q-1}\}} \mathbb{P} \left(I_{s_q} = 1 \middle| \prod_{t=1}^{q-1} I_{s_t} = 1 \right) &= \sum_{s_q \notin \{s_1, \dots, s_{q-1}\}} \frac{Q_{k_q j_q} / Q_{k_q}^+ Q_{j_q}^-}{\sum_{w \in [n]} (j_w - O(q)) Q_{k_q j_w} / Q_{k_q}^+ Q_{j_w}^-} \\
&= \frac{1}{n} \sum_{s_q \notin \{s_1, \dots, s_{q-1}\}} \frac{Q_{k_q j_q} / Q_{k_q}^+ Q_{j_q}^-}{z[1 + o(1) + O(q/n)]} \\
&= \frac{1}{nz} \sum_{v \in [n]} \left((k_v j_v - O(q)) \frac{Q_{k_v j_v}}{Q_{k_v}^+ Q_{j_v}^-} [1 + o(1) + O(q/n)] \right) \\
&= \frac{1}{z} \mathbb{E} \left(K_v J_v \frac{Q_{K_v J_v}}{Q_{K_v}^+ Q_{J_v}^-} \right) (1 + o(1) + O(q/n)) \\
&= \left(z \sum_{jk} \frac{P_{jk}}{P_j^- P_k^+} Q_{kj} \right) (1 + o(1) + O(q/n)) \\
&= \mu(1 + o(1) + O(q/n)) = \mu(1 + o(1)).
\end{aligned} \tag{2.32}$$

Therefore one can conclude by putting together the last line of (2.31) and the manipulations in (2.32) that: $\mathbb{E}((S_n)_r) = \mu^r(1 + o(1))$, hence:

$$\lim_{n \rightarrow \infty} \mathbb{E}((S_n)_r) = \mu^r,$$

which is equivalent to showing that S_n converges in distribution to $\text{Pois}(\mu)$. ■

Corollary 2.3.7. As a conclusion of previous theorem, the probability of having no self-loops converges to $e^{-\mu}$; indicating the fact that for large n this probability stays bounded away from zero.

Chapter 3

Inhomogeneous Random Graphs

3.1 Model Description

Recalling the Erdős-Renyi graph with homogenous edge occupation probabilities, there is an extension, which assigns inhomogeneous edge probabilities adjusted by vertex weights. So the weights of two ends of an edge determine the edge probability between them, which seems quite natural for real applications. Perhaps we may expect that by assigning suitable equal weights the Erdős-Renyi model will be obtained. Given the degree weights $(w_i)_{i \in [n]}$, the probability of having an edge between nodes i, j would be

$$p_{ij} = \frac{w_i w_j}{n + w_i w_j}, \quad (3.1)$$

as suggested in [12]. In the sense of the odds ratio r_{ij} one could express the Bernoulli random variables X_{ij} showing the status of edge (i, j) by their occurrence ($X_{ij} = 1$)

probabilities, that is

$$\begin{aligned}\mathbb{P}(X_{ij} = 1) &= p_{ij} = \frac{r_{ij}}{1 + r_{ij}} \\ \mathbb{P}(X_{ij} = 0) &= q_{ij} = \frac{1}{1 + r_{ij}}.\end{aligned}\tag{3.2}$$

Then the probability of certain edge configurations $X = x = (x_{ij})$ would be

$$\mathbb{P}(X = x) = \prod_{i < j} p_{ij}^{x_{ij}} q_{ij}^{1-x_{ij}} = \prod_{i < j} (1 + r_{ij})^{-1} \prod_{i < j} r_{ij}^{x_{ij}}.\tag{3.3}$$

In the case of the given edge probabilities in (3.1), $r_{ij} = w_i w_j / n$ which is a multiplicative kernel for the odds ratio. By defining $u_i = w_i / \sqrt{n}$ we get the vector of node types $u = (u_i)_{i \in [n]}$, which is generated from the weights vector. Thus (3.3) would be

$$\begin{aligned}\mathbb{P}(X = x) &= \prod_{i < j} (1 + u_i u_j)^{-1} \prod_{i < j} (u_i u_j)^{x_{ij}} \\ &= G(u)^{-1} \prod_i u_i^{d_i(x)}.\end{aligned}\tag{3.4}$$

So from (3.4) $(d_i(X))_{i \in [n]}$ is the sufficient statistics for configuration probability, moreover from $\sum_x \mathbb{P}(X = x) = 1$ we get the important relation

$$G(u) = \sum_x \prod_i u_i^{d_i(x)}.\tag{3.5}$$

Lemma 3.1.1. Assuming the odds ratio is multiplicative, the set of graphs with degree sequence $(d_i)_{i \in [n]}$ are uniformly distributed.

Proof. We simply can find the conditional probability as:

$$\begin{aligned}
\mathbb{P}(X = x | d_i(X) = d_i \forall i \in [n]) &= \frac{\mathbb{P}(X = x)}{\mathbb{P}(d_i(x) = d_i \forall i \in [n])} \\
&= \frac{G(u)^{-1} \prod_i u_i^{d_i(x)}}{\sum_{x: d_i(x) = d_i \forall i \in [n]} G(u)^{-1} \prod_i u_i^{d_i(x)}} \\
&= \frac{1}{\#\{x : d_i(x) = d_i \forall i \in [n]\}}.
\end{aligned} \tag{3.6}$$

■

Given the degree weights or equivalently degree types, to investigate the joint degree distribution one could first find the formula for the generating function:

$$\begin{aligned}
\mathbb{E} \left(\prod_i t_i^{d_i(X)} \right) &= \sum_x \mathbb{P}(X = x) \prod_i t_i^{d_i(x)} \\
&= G(u)^{-1} \sum_x \prod_i (t_i u_i)^{d_i(x)} \\
&= \frac{G(tu)}{G(u)} = \prod_{i < j} \frac{1 + t_i u_i t_j u_j}{1 + u_i u_j}.
\end{aligned} \tag{3.7}$$

Keeping in mind that (3.7) is the conditional expectation provided the node types, we define *mixed Poisson* distribution as an appropriate candidate for vertex weight distribution.

Definition 3.1.2. Random variable X is said to have *mixed Poisson* distribution with mixing distribution F iff

$$\mathbb{P}(X = k) = \mathbb{E} \left(e^{-Z} \frac{Z^k}{k!} \right), \tag{3.8}$$

where Z is the mixing parameter and F -distributed.

The next theorem has been proved in [12], and is especially important because it studies the asymptotic degree distribution of a fixed vertex, given some regularity conditions on weight distribution and the multiplicative structure of the odds ratio. The weight or type sequence is no longer assumed to be deterministically given, but is random. Therefore, another expectation is needed to get the unconditional generation function in (3.7).

Theorem 3.1.3. Consider the generalized random graph on n vertices with edge probabilities defined by $r_{ij} = p_{ij}/q_{ij} = W_i W_j/n$, where $(W_i)_{i \in [n]}$ are i.i.d random variables with mean μ_w and finite moment of order $1 + \varepsilon$, for some $\varepsilon > 0$. Then

- (i) The limiting distribution of the k -th vertex degree D_k as $n \rightarrow \infty$ is *mixed Poisson* with mixing variable $W_k \mu_w$.
- (ii) For any finite m , the variables D_1, \dots, D_k are asymptotically independent, i.e their joint degree distribution can be factored into independent components.

Proof. As proved in [12] and [10], to capture only the effect of vertex k , we need to take other $t_i = 1$ for $i \neq k$ and take $t_k = t$, then from (3.7)

$$\mathbb{E}(t^{D_k}) = \mathbb{E} \left(\prod_{i \neq k} \frac{1 + W_i W_k t/n}{1 + W_i W_k/n} \right). \quad (3.9)$$

From Taylor expansion of $\log(1 + x) = x + O(x^2)$, we will get

$$\prod_{i \neq k} \frac{1 + W_i W_k t/n}{1 + W_i W_k/n} = \exp \left(\frac{W_k \sum_i W_i}{n} (t - 1) + R_n \right), \quad (3.10)$$

where $R_n = O(W_k^2 \sum_i W_i^2/n^2)$. Since,

$$\begin{aligned} \frac{\sum_{i \neq k} W_i^2}{n^2} &\leq \frac{\max\{W_i\} \sum_{i \neq k} W_i}{n^2} \\ &\approx \mu_w \frac{\max\{W_i\}}{n} \rightarrow 0, \end{aligned} \tag{3.11}$$

and $\max\{W_i\}/n \rightarrow 0$ because W has finite $1 + \varepsilon$ moment. Thus R_n converges almost surely to 0. Hence

$$\mathbb{E}(t^{D_k}) \rightarrow \mathbb{E}(e^{W_k \mu_w (t-1)}), \tag{3.12}$$

which is the transform of *mixed Poisson* distribution. Thus, the distribution of vertex k converges weakly to the *mixed Poisson* with parameter $W_k \mu_w$, which ends the proof of part (i). For the second part by taking $t_i = 1$ for $i > m$ and proceeding similar to the previous part we can get

$$\mathbb{E}\left(\prod_{i=1}^m t_i^{D_i}\right) \rightarrow \prod_{i=1}^m \mathbb{E}(e^{W_i \mu_w (t_i-1)}), \tag{3.13}$$

which verifies that the joint generating function asymptotically factorizes into the product of mixed Poisson generating functions. ■

3.2 Relation Between Assortative Configuration Model and IRG

The main study of this section is the relation between the IRG model and configuration graph. One can find the edge and more type probabilities in IRG model, and relate them to the similar distributions in assortative configuration model. However,

this investigation requires considering more general forms of odds ratios rather than the multiplicative one studied in previous section. So we assume that the $u = (u_v)_{v \in [n]}$ are iid non-negative samples from cumulative distribution function $F : \mathbb{R}_+ \rightarrow [0, 1]$, therefore the connection probabilities between node v and w can be identified with a *general* kernel function κ . That is,

$$p_{ij} = \frac{\kappa(u_i, u_j)}{1 + \kappa(u_i, u_j)}, \quad (3.14)$$

where $\kappa : \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is symmetric and non-decreasing in both variables. As a special case one can let $\kappa(u, u') = uu'$ to derive the Chung-Lu model [21]. Likewise, (3.3) the probability of a certain edge configuration $x = (x_{ij})_{i,j \in [n]}$ given the general kernel (3.14) would be

$$\begin{aligned} \mathbb{P}(X = x | u) &= \prod_{i < j} \left(\frac{\kappa(u_i, u_j)}{1 + \kappa(u_i, u_j)} \right)^{x_{ij}} \left(\frac{1}{1 + \kappa(u_i, u_j)} \right)^{1-x_{ij}} \\ &= \prod_{i < j} \frac{1}{1 + \kappa(u_i, u_j)} \prod_{i < j} \kappa(u_i, u_j)^{x_{ij}} \\ &= G(u)^{-1} \prod_{i < j} \kappa(u_i, u_j)^{x_{ij}}, \end{aligned} \quad (3.15)$$

and since the total conditional probability is 1, we get the identity

$$G(u) = \prod_{i < j} (1 + \kappa(u_i, u_j)) = \sum_x \prod_{i < j} \kappa(u_i, u_j)^{x_{ij}}. \quad (3.16)$$

In order to find some interesting results about the degree distribution we may want to find the generating function as we did in (3.7). Given the type sequence the

conditional generating function would be

$$\begin{aligned}
\mathbb{E} \left(\prod_{i \in [n]} t_i^{d_i(X)} \mid u \right) &= \mathbb{E} \left(\prod_{i < j} (t_i t_j)^{x_{ij}} \mid u \right) \\
&= \sum_x \mathbb{P}(X = x \mid u) \prod_{i < j} G(u)^{-1} (t_i t_j \kappa(u_i, u_j))^{x_{ij}} \\
&= G(u)^{-1} \sum_x \prod_{i < j} (t_i t_j \kappa(u_i, u_j))^{x_{ij}} \\
&= \prod_{i < j} \frac{1 + t_i t_j \kappa(u_i, u_j)}{1 + \kappa(u_i, u_j)}.
\end{aligned} \tag{3.17}$$

Node types may change as n changes. Generally they might scale with n like the usual case of previous section, $u = w/\sqrt{n}$. However, we allow any arbitrary scaling rule, which turns u into \tilde{u} as n independent random variables. Moreover, the kernel function might have some n dependencies, which together with the type scaling rule $\phi^{(n)}$ can be expressed as follows:

$$u = u^{(n)} = \phi^{(n)}(\tilde{u}); \quad \kappa^{(n)}(u, u') = (n-1)^{-1} \kappa(\tilde{u}, \tilde{u}'). \tag{3.18}$$

Therefore (3.17) is expressed as

$$\begin{aligned}
\mathbb{E} \left(\prod_{i \in [n]} t_i^{d_i(X)} \mid u \right) &= \mathbb{E} \left(\prod_{i \in [n]} t_i^{d_i(X)} \mid \tilde{u} \right) \\
&= \prod_{i < j} \frac{1 + t_i t_j (n-1)^{-1} \kappa(\tilde{u}_i, \tilde{u}_j)}{1 + (n-1)^{-1} \kappa(\tilde{u}_i, \tilde{u}_j)}.
\end{aligned} \tag{3.19}$$

The fact that the kernel function decays as $(n-1)^{-1}$ is important for the upcoming claims. For the next theorem, that has been proved in [22], we assume that the

underlying n independent kernel function has finite $1 + \alpha$ moment for some $\alpha > 0$,

$$\mathbb{E}(\kappa(\tilde{u}, \tilde{u}')^{1+\alpha}) = \int_{\mathbb{R}_+^2} \kappa(\tilde{u}, \tilde{u}')^{1+\alpha} dF(\tilde{u})dF(\tilde{u}') < \infty. \quad (3.20)$$

Theorem 3.2.1. In the IRG model with known kernel function $\kappa(\tilde{u}, \tilde{u}')$, which satisfies the scaling rule (3.18) as well as the $1 + \alpha$ moment condition,

- (i) the generating function for degree of vertex 1 would be

$$\psi(t_1) = \mathbb{E}\left(t_1^{d_1(X)}\right) = \mathbb{E}\left(e^{(t_1-1)h^{-1}(\tilde{U})}\right) (1 + o(1)), \quad (3.21)$$

where \tilde{U} is F distributed, and

$$h^{-1}(\tilde{u}) = \int_{\mathbb{R}_+} \kappa(\tilde{u}, \tilde{u}') dF(\tilde{u}'). \quad (3.22)$$

- (ii) For any fixed m , the joint degree distribution of d_v , $v \in [m]$ converges in distribution to the product of independent iid random variables.

Proof. To capture the generating function of $d_1(X)$, let $t_i = 1$ for $i \neq 1$ in (3.17).

Thus the unconditional generating function would be

$$\begin{aligned} \psi(t_1) &= \mathbb{E}\left(\mathbb{E}\left(t_1^{d_1(X)} \mid u\right)\right) = \mathbb{E}\left(\prod_{1 < j} \frac{1 + t_1 \kappa(u_1, u_j)}{1 + \kappa(u_1, u_j)}\right) \\ &= \mathbb{E}\left(\mathbb{E}\left(\prod_{1 < j} \frac{1 + t_1 \kappa(u_1, u_j)}{1 + \kappa(u_1, u_j)} \mid u_1\right)\right) \\ &= \mathbb{E}\left(\mathbb{E}\left(\frac{1 + t_1 \kappa(u_1, u')}{1 + \kappa(u_1, u')}\right)^{n-1} \mid u_1\right), \end{aligned} \quad (3.23)$$

where the last equation above comes from the independency of u elements. The

following integral addresses the inner conditional expectation in (3.23):

$$\begin{aligned} \mathbb{E} \left(\frac{1 + t_1 \kappa(u_1, u')}{1 + \kappa(u_1, u')} \middle| u_1 \right) &= \int_{\mathbb{R}_+} \frac{1 + t_1 \kappa(u_1, u')}{1 + \kappa(u_1, u')} dF(u') \\ &= \int_{\mathbb{R}_+} \frac{1 + t_1 (n-1)^{-1} \kappa(\tilde{u}_1, \tilde{u}')}{1 + (n-1)^{-1} \kappa(\tilde{u}_1, \tilde{u}')} dF(\tilde{u}') \end{aligned} \quad (3.24)$$

Now because $\frac{1+tx}{1+x} = 1 + (t-1)x + R(x)$ such that $R(x) = O(x^{1+\alpha})$, we can simplify the above integral as

$$\begin{aligned} \mathbb{E} \left(\frac{1 + t_1 \kappa(u_1, u')}{1 + \kappa(u_1, u')} \middle| u_1 \right) &= \int_{\mathbb{R}_+} (1 + (t_1 - 1)(n-1)^{-1} \kappa(\tilde{u}_1, \tilde{u}')) dF(\tilde{u}') \\ &\quad + \int_{\mathbb{R}_+} O((n-1)^{-(1+\alpha)} \kappa(\tilde{u}_1, \tilde{u}')^{1+\alpha}) dF(\tilde{u}') \\ &= 1 + (t_1 - 1)(n-1)^{-1} h^{-1}(\tilde{u}_1) + \tilde{R}. \end{aligned} \quad (3.25)$$

Because of the integrability condition on $1 + \alpha$ moment of κ , \tilde{R} would be of order $O((n-1)^{-(1+\alpha)})$, thus (3.23) is computed as

$$\begin{aligned} \psi(t_1) &= \mathbb{E} \left(\left(1 + (t_1 - 1)(n-1)^{-1} h^{-1}(\tilde{u}_1) + \tilde{R} \right)^{n-1} \right) \\ &= \mathbb{E} \left(e^{(t_1-1)h^{-1}(\tilde{U})} \right) (1 + O(n^{-\alpha})), \end{aligned} \quad (3.26)$$

which concludes the proof of first part. Deducing from this part, one can conclude that the asymptotic degree distribution of vertex 1 would be *mixed Poisson* with mixing parameter $h^{-1}(\tilde{U})$, that has the cumulative distribution $F(h(\cdot))$. By similar computations the second half of the theorem can be proved, which claims that the joint degree of a fixed number of vertices are asymptotically getting independent of each other. ■

In the sense of the parametrization of this theorem, one can define transformed $(\tilde{\kappa}, \tilde{F})$, such that under new parameters h becomes identity map. By the transformations

$$\tilde{\kappa}(v, v') = \kappa(h(v), h(v')); \quad \tilde{F} = F \circ h, \quad (3.27)$$

one can verify that the new pair leads to the same model:

$$\psi(t_1) = \mathbb{E} \left(e^{(t_1-1)V} \right) (1 + o(1)), \quad (3.28)$$

where V is \tilde{F} distributed. Moreover, under the new pair

$$v = \int_{\mathbb{R}_+} \tilde{\kappa}(v, v') d\tilde{F}(v'). \quad (3.29)$$

Since it is always possible to apply these transformations to a given pair, without any loss of generality we might assume that we can take (κ, F) such that h is the identity map.

We can go one step further to analyze the edge type distribution as desired at the beginning of this subsection. Hence, we should find the generating function for the *shifted* bivariate distribution of edge degrees by computing $\psi_{12}(t_1, t_2) =$

$\mathbb{E} (t_1^{d_1-1} t_2^{d_2-1} | v_1 \sim v_2)$ under the identity h map. Therefore, we first find the conditional generating function when $v_1 \sim v_2$ (v_1 is attached to v_2):

$$\begin{aligned}
\mathbb{E} \left(t_1^{d_1-1} t_2^{d_2-1} \prod_{i \geq 3} t_i^{d_i} \middle| v_1 \sim v_2 \right) &= (t_1 t_2)^{-1} \mathbb{E} \left(\prod_{i < j} (t_i t_j)^{x_{ij}} \middle| v_1 \sim v_2 \right) \\
&= (t_1 t_2)^{-1} \sum_x \mathbb{P}(X = x | x_{12} = 1) \prod_{i < j} (t_i t_j)^{x_{ij}} \\
&= (t_1 t_2)^{-1} \sum_{x: x_{12}=1} \frac{\mathbb{P}(X = x)}{\mathbb{P}(x_{12} = 1)} \prod_{i < j} (t_i t_j)^{x_{ij}} \\
&= \mathbb{P}(x_{12} = 1)^{-1} \sum_{x: x_{12}=1} \mathbb{E} \left(\frac{\kappa(u_1, u_2)}{1 + \kappa(u_1, u_2)} \prod_{\{i < j: (i,j) \neq (1,2)\}} \frac{(t_i t_j \kappa(u_i, u_j))^{x_{ij}}}{1 + \kappa(u_i, u_j)} \right) \quad (3.30) \\
&= \mathbb{P}(x_{12} = 1)^{-1} \mathbb{E} \left(\frac{\kappa(u_1, u_2)}{1 + \kappa(u_1, u_2)} \sum_{x: x_{12}=1} \prod_{\{i < j: (i,j) \neq (1,2)\}} \frac{(t_i t_j \kappa(u_i, u_j))^{x_{ij}}}{1 + \kappa(u_i, u_j)} \right) \\
&= \mathbb{P}(x_{12} = 1)^{-1} \mathbb{E} \left(\frac{\kappa(u_1, u_2)}{1 + \kappa(u_1, u_2)} \prod_{\{i < j: (i,j) \neq (1,2)\}} \frac{1 + t_i t_j \kappa(u_i, u_j)}{1 + \kappa(u_i, u_j)} \right).
\end{aligned}$$

The last inequality is an extended version of (3.16). Thus, we can find $\psi_{12}(t_1, t_2)$ by letting $t_i = 1$ for $i \geq 3$:

$$\psi_{12}(t_1, t_2) = \mathbb{P}(x_{12} = 1)^{-1} \mathbb{E} \left(\frac{\kappa(u_1, u_2)}{1 + \kappa(u_1, u_2)} \prod_{j \geq 3} \frac{1 + t_1 \kappa(u_1, u_j)}{1 + \kappa(u_1, u_j)} \frac{1 + t_2 \kappa(u_2, u_j)}{1 + \kappa(u_2, u_j)} \right). \quad (3.31)$$

And

$$\begin{aligned}
\mathbb{P}(x_{12} = 1) &= \mathbb{E} \left(\frac{\kappa(u_1, u_2)}{1 + \kappa(u_1, u_2)} \right) = \mathbb{E} \left(\frac{(n-1)^{-1} \kappa(\tilde{u}_1, \tilde{u}_2)}{1 + (n-1)^{-1} \kappa(\tilde{u}_1, \tilde{u}_2)} \right) \\
&= (n-1)^{-1} \mathbb{E}(\kappa(\tilde{u}_1, \tilde{u}_2)) (1 + o(1)).
\end{aligned} \quad (3.32)$$

The expectation term in $\psi_{12}(t_1, t_2)$ can be treated like in the proof of previous theorem

to obtain

$$\begin{aligned} \mathbb{E} \left(\frac{\kappa(u_1, u_2)}{1 + \kappa(u_1, u_2)} \prod_{j \geq 3} \frac{1 + t_1 \kappa(u_1, u_j)}{1 + \kappa(u_1, u_j)} \frac{1 + t_2 \kappa(u_2, u_j)}{1 + \kappa(u_2, u_j)} \right) = \\ \mathbb{E} \left((n-1)^{-1} \kappa(\tilde{u}_1, \tilde{u}_2) e^{(t_1-1)\tilde{u}_1} e^{(t_2-1)\tilde{u}_2} \right) (1 + o(1)). \end{aligned} \quad (3.33)$$

Therefore we find the expression

$$\begin{aligned} \psi_{12}(t_1, t_2) &= \mathbb{E} \left(t_1^{d_1-1} t_2^{d_2-1} | v_1 \sim v_2 \right) \\ &= [\mathbb{E} (\kappa(\tilde{u}_1, \tilde{u}_2))]^{-1} \mathbb{E} \left(\kappa(\tilde{u}_1, \tilde{u}_2) e^{(t_1-1)\tilde{u}_1} e^{(t_2-1)\tilde{u}_2} \right) (1 + o(1)), \end{aligned} \quad (3.34)$$

as the generating function for shifted edge degrees. Since we take (κ, F) such that h becomes identity, therefore the mixing parameters are \tilde{u}_1 and \tilde{u}_2 .

There is also another way around to find $\psi_{12}(t_1, t_2)$, which requires the edge type probabilities. Here is the point that we derive the relation between IRG and the assortative graph. The direct computation of the conditional expectation leads to

$$\begin{aligned} \psi_{12}(t_1, t_2) &= \mathbb{E} \left(t_1^{d_1-1} t_2^{d_2-1} | v_1 \sim v_2 \right) \\ &= \sum_{k, k' \geq 1} t_1^{k-1} t_2^{k'-1} \mathbb{P} (d_1 = k, d_2 = k' | v_1 \sim v_2) \\ &= \sum_{k, k' \geq 1} t_1^{k-1} t_2^{k'-1} Q_{kk'}^{(n)}. \end{aligned} \quad (3.35)$$

One can expand (3.34) to get

$$\begin{aligned}
\psi_{12}(t_1, t_2) &= [\mathbb{E}(\kappa(\tilde{u}_1, \tilde{u}_2))]^{-1} \mathbb{E}(\kappa(\tilde{u}_1, \tilde{u}_2) e^{(t_1-1)\tilde{u}_1} e^{(t_2-1)\tilde{u}_2}) (1 + o(1)) \\
&= [\mathbb{E}(\kappa(\tilde{u}_1, \tilde{u}_2))]^{-1} \mathbb{E} \left(\kappa(\tilde{u}_1, \tilde{u}_2) e^{-(\tilde{u}_1 + \tilde{u}_2)} \sum_{k, k'} \frac{(t_1 \tilde{u}_1)^k (t_2 \tilde{u}_2)^{k'}}{k! k'} \right) (1 + o(1)) \\
&= [\mathbb{E}(\kappa(\tilde{u}_1, \tilde{u}_2))]^{-1} \sum_{k, k'} t_1^k t_2^{k'} \mathbb{E} \left(\kappa(\tilde{u}_1, \tilde{u}_2) \frac{e^{-\tilde{u}_1} \tilde{u}_1^k}{k!} \frac{e^{-\tilde{u}_2} \tilde{u}_2^{k'}}{k'} \right) (1 + o(1)).
\end{aligned} \tag{3.36}$$

Letting $n \rightarrow \infty$, then by equating the coefficients of $t_1^{k-1} t_2^{k'-1}$ in (3.35) and (3.36) it follows

$$\begin{aligned}
Q_{kk'} &= [\mathbb{E}(\kappa(\tilde{u}_1, \tilde{u}_2))]^{-1} \mathbb{E} \left(\kappa(\tilde{u}_1, \tilde{u}_2) \frac{e^{-\tilde{u}_1} \tilde{u}_1^{k-1}}{(k-1)!} \frac{e^{-\tilde{u}_2} \tilde{u}_2^{k'-1}}{(k'-1)!} \right) \\
&= [\mathbb{E}(\kappa(\tilde{u}_1, \tilde{u}_2))]^{-1} \mathbb{E}(\kappa(\tilde{u}_1, \tilde{u}_2) \text{Pois}(\tilde{u}_1, k-1) \text{Pois}(\tilde{u}_2, k'-1)) \tag{3.37} \\
&= \frac{\int_{\mathbb{R}_+^2} \kappa(\tilde{u}_1, \tilde{u}_2) \text{Pois}(\tilde{u}_1, k-1) \text{Pois}(\tilde{u}_2, k'-1) dF(\tilde{u}_1) dF(\tilde{u}_2)}{\int_{\mathbb{R}_+^2} \kappa(\tilde{u}_1, \tilde{u}_2) dF(\tilde{u}_1) dF(\tilde{u}_2)}.
\end{aligned}$$

The above equation let us to conclude that the edge degree distribution can be found for a given pair (κ, F) in IRG model.

One can ignore the types of the nodes, then by picking these edge type factors in assortative discrete matching algorithm, we get a random graph being asymptotically equivalent to the degree sequence of the IRG model with the given pair. The only thing that can not be replicated through assortative matching would be the type labels of the nodes. To preserve the effect of node types, we have to apply a more generalized model, which contains both assortative and IRG graphs. Our suggestion is following the same assortative wiring, that was introduced in section 2.2, but with different edge weights $Q_{ku, k'u'}$. Therefore, aside from the in- and out-degree of the

edges, the types of incorporated nodes u, u' are also considered in edge weights. In this new model, we have finer resolution over the edge types, which includes the type of the incident nodes as well as their degrees. It can be readily shown that by ignoring the node types, assortative graphs would be attained. Moreover, one can speculate from (3.37) that the suitable edge factors with node types being considered would be:

$$Q_{k\tilde{u},k'\tilde{u}'} = \frac{\kappa(\tilde{u}_1, \tilde{u}_2) \text{Pois}(\tilde{u}_1, k-1) \text{Pois}(\tilde{u}_2, k'-1) dF(\tilde{u}_1) dF(\tilde{u}_2)}{\int_{\mathbb{R}_+^2} \kappa(\tilde{u}_1, \tilde{u}_2) dF(\tilde{u}_1) dF(\tilde{u}_2)}. \quad (3.38)$$

Therefore, one can pick the node joint degree-type distribution $P_{k\tilde{u}}$ to be consistent with the above edge factors. And expect to obtain a random graph with asymptotically equivalent degree distribution to the IRG model with the pair (κ, F) , that has already yielded to (3.38). It is important to note that the support of node types must be discrete, otherwise our large n analysis for the convergence of the discrete Markov variables does not hold anymore.

We are looking to study more the last proposed method, which results to both previously discussed models. Moreover, as the future works we speculate that this model would be a nice candidate for the skeleton graph of financial networks, since it captures both the assortativity effect as well as the agent types. Then, we would like to add on the random exposures between the agents, as independent random variables conditioned on the degree-type of the lender and borrower. Finally, this model would let us to explore the cascade dynamics in financial networks from a new perspective.

Appendix A

Proof of part **i** of Theorem 2.1.4

Proof.

$$\begin{aligned}
& \left| \mathbb{E} \left(f((\tilde{j}_i, \tilde{k}_i)_{i \in \mathcal{S}}) - f((\hat{j}_i, \hat{k}_i)_{i \in \mathcal{S}}) \right) \right| \\
&= \left| \mathbb{E} \left(f((\tilde{j}_i, \tilde{k}_i)_{i \in \mathcal{S}}) | \mathcal{D}_n \right) - \mathbb{E} \left(f((\hat{j}_i, \hat{k}_i)_{i \in \mathcal{S}}) \right) \right| \\
&\leq \left| \mathbb{E} \left(f((\tilde{j}_i, \tilde{k}_i)_{i \in \mathcal{S}}) | \mathcal{D}_n \right) - \mathbb{E} \left(f((j_i, k_i)_{i \in \mathcal{S}}) | \mathcal{D}_n \right) \right| \\
&+ \left| \mathbb{E} \left(f((j_i, k_i)_{i \in \mathcal{S}}) | \mathcal{D}_n \right) - \mathbb{E} \left(f((\hat{j}_i, \hat{k}_i)_{i \in \mathcal{S}}) \right) \right|
\end{aligned} \tag{A.1}$$

Let \mathcal{F}_n be the σ -field generated by $(j_i, k_i)_{i \in [n]}$, then it is possible to bound the first term in (A.1) as follows:

$$\begin{aligned}
& \mathbb{E} \left(\left| f((\tilde{j}_i, \tilde{k}_i)_{i \in \mathcal{S}}) - f((j_i, k_i)_{i \in \mathcal{S}}) \right| | \mathcal{D}_n \right) \\
&\leq 2M \mathbb{E} \left(\mathbf{1}_{\{\sum_{i \in \mathcal{S}} (\tilde{j}_i - j_i + \tilde{k}_i - k_i) \geq 0\}} | \mathcal{D}_n \right) \\
&= \frac{2M}{\mathbb{P}(\mathcal{D}_n)} \mathbb{E} \left(\mathbf{1}_{\{\sum_{i \in \mathcal{S}} (\tilde{j}_i - j_i + \tilde{k}_i - k_i) \geq 0\}} \mathbf{1}_{\mathcal{D}_n} \right) \\
&\leq \frac{2M}{\mathbb{P}(\mathcal{D}_n)} \mathbb{E} \left(\mathbf{1}_{\mathcal{D}_n} \mathbb{E} \left(\sum_{i \in \mathcal{S}} (\tilde{j}_i - j_i + \tilde{k}_i - k_i) | \mathcal{F}_n \right) \right) \\
&= \frac{2M}{\mathbb{P}(\mathcal{D}_n)} \mathbb{E} \left(\mathbf{1}_{\mathcal{D}_n} |\mathcal{S}| \frac{|D_n|}{n} \right) = 2M |\mathcal{S}| \frac{|D_n|}{n}.
\end{aligned} \tag{A.2}$$

Since $\mathbb{P}(\mathcal{D}_n) \rightarrow 1$, then the second term in (A.1) can also be bounded from above:

$$\begin{aligned} \mathbb{E}(f((j_i, k_i)_{i \in \mathcal{S}}) | \mathcal{D}_n) &= \mathbb{P}(\mathcal{D}_n)^{-1} \mathbb{E}(f((j_i, k_i)_{i \in \mathcal{S}}) \mathbf{1}_{\mathcal{D}_n}) \\ &= \mathbb{E}(f((j_i, k_i)_{i \in \mathcal{S}}) \mathbf{1}_{\mathcal{D}_n}) (1 + o(1)). \end{aligned} \quad (\text{A.3})$$

Moreover we have:

$$\mathbb{E}\left(f((\hat{j}_i, \hat{k}_i)_{i \in \mathcal{S}})\right) = \mathbb{E}\left(f((\hat{j}_i, \hat{k}_i)_{i \in \mathcal{S}}) \mathbf{1}_{\hat{\mathcal{D}}_n}\right) + \mathbb{E}\left(f((\hat{j}_i, \hat{k}_i)_{i \in \mathcal{S}}) \mathbf{1}_{\hat{\mathcal{D}}_n^c}\right). \quad (\text{A.4})$$

It should be clear what $\hat{\mathcal{D}}_n$ is indicating. It is exactly the same event like \mathcal{D}_n , but determined by $(\hat{j}_i, \hat{k}_i)_{i \in [n]}$, which is an independent copy generated from P distribution.

Hence we have:

$$\begin{aligned} &\left| \mathbb{E}(f((j_i, k_i)_{i \in \mathcal{S}}) | \mathcal{D}_n) - \mathbb{E}\left(f((\hat{j}_i, \hat{k}_i)_{i \in \mathcal{S}})\right) \right| \\ &\leq \mathbb{E}\left(\left| f((\hat{j}_i, \hat{k}_i)_{i \in \mathcal{S}}) \mathbf{1}_{\hat{\mathcal{D}}_n} \right|\right) o(1) + \mathbb{E}\left(f((\hat{j}_i, \hat{k}_i)_{i \in \mathcal{S}}) \mathbf{1}_{\hat{\mathcal{D}}_n^c}\right) \\ &\leq M \left(o(1) + \mathbb{P}\left(\hat{\mathcal{D}}_n^c\right) \right). \end{aligned} \quad (\text{A.5})$$

Then (A.2) together with (A.5) conclude the proof. ■

Appendix B

Proof of Proposition 2.3.1

Proof. The proof of the first claim in the proposition is trivial, since function f is Lipschitz on $\Omega_n^{(\varepsilon)}$ then the differential system in (2.15) has a unique solution on any compact subset of $\Omega_n^{(\varepsilon)}$, and it is uniquely extendible to the boundary of this set. For the proof of the second part let's denote the m -th approximate solution to the above system by $x^{(m)}(\tau) = \left(\left\{ z_k^{+, (m)}(\tau) \right\}_{k=1}^{k(n)}, \left\{ z_j^{-, (m)}(\tau) \right\}_{j=1}^{j(n)}, \left\{ e_{kj}^{(m)}(\tau) \right\}_{k,j=1}^{k(n), j(n)} \right)$, which can be computed throughout the following recursive equation

$$x^{(m+1)}(\tau) = x(0) + \int_0^\tau f(x^{(m)}(s)) ds, \quad (\text{B.6})$$

where $x^{(1)}(\tau)$ is the solution of the *unclipped* system, where the indices of the summations in numerator and denominators of (2.14) are not clipped by n -th order statistics, and the vector of initial values is not truncated as it was in (2.14). In other words, $x^{(1)}(\tau)$ is the solution to the infinite dimensional counterpart of (2.14). One could

readily check that $x^{(1)}(\tau)$ satisfies

$$\begin{aligned} z_k^{+, (1)}(\tau) &= zQ_k^+ \left(1 - \frac{\tau}{z}\right) \\ z_j^{-, (1)}(\tau) &= zQ_j^- \left(1 - \frac{\tau}{z}\right) \\ e_{kj}^{(1)}(\tau) &= Q_{kj}\tau. \end{aligned} \tag{B.7}$$

To calculate $x^{(2)}(\tau)$, we find an arbitrary element of $f(x^{(1)}(\tau))$, say $z_k^{+, (2)}(\tau)$. The R.H.S of the second equation in (2.14) would be equal to

$$-\frac{\sum_{j \leq j(n)} Q_{kj}}{\sum_{k \leq k(n), j \leq j(n)} Q_{kj}} = -\frac{Q_k^+ + o(1)}{1 + o(1)} = -Q_k^+ + o(1). \tag{B.8}$$

Which leads to

$$\begin{aligned} z_k^{+, (2)}(\tau) &= zQ_k^+ - \int_0^\tau (Q_k^+ + o(1)) ds \\ &= zQ_k^+ \left(1 - \frac{\tau}{z}\right) - \tau o(1), \end{aligned}$$

and we get

$$\left| z_k^{+, (2)}(\tau) - z_k^{+, (1)}(\tau) \right| = \tau o(1).$$

Therefore we have

$$\begin{aligned} \|x^{(2)}(s) - x^{(1)}(s)\| &= \max_{k \leq k(n), j \leq j(n)} \left\{ \left| z_k^{+, (2)}(s) - z_k^{+, (1)}(s) \right|, \left| z_j^{-, (2)}(s) - z_j^{-, (1)}(s) \right|, \left| e_{kj}^{(2)}(s) - e_{kj}^{(1)}(s) \right| \right\} \\ &= so(1). \end{aligned}$$

From equation (B.6)

$$\begin{aligned}
\|x^{(m+1)}(\tau) - x^{(m)}(\tau)\| &= \left\| \int_0^\tau (f(x^{(m)}(s)) - f(x^{(m-1)}(s))) ds \right\| \\
&\leq \int_0^\tau \|f(x^{(m)}(s)) - f(x^{(m-1)}(s))\| ds \\
&\leq c \int_0^\tau \|x^{(m)}(s) - x^{(m-1)}(s)\| ds \\
&\leq c^{m-1} \frac{\tau^m}{m!} o(1) \\
\implies \sum_{m=1}^{\infty} \|x^{(m+1)}(\tau) - x^{(m)}(\tau)\| &\leq e^{c\tau} o(1).
\end{aligned}$$

Therefore the sum $x^{(1)}(\tau) + \sum_{m=1}^{\infty} x^{(m+1)}(\tau) - x^{(m)}(\tau)$ is uniformly and absolutely convergent. It is not so difficult to check that, this infinite sum is the solution of the differential system, namely $x(\tau)$. Because

$$\left\| \int_0^\tau f(x(s)) ds - \int_0^\tau f(x^{(m)}(s)) ds \right\| \leq c \int_0^\tau \|x(s) - x^{(m)}(s)\| ds \rightarrow 0.$$

In addition we have

$$\begin{aligned}
x(\tau) &= \lim_{m \rightarrow \infty} x^{(m+1)}(\tau) \\
&= \lim_{m \rightarrow \infty} x(0) + \int_0^\tau f(x^{(m)}(s)) ds = x(0) + \int_0^\tau f(x(s)) ds.
\end{aligned}$$

Consequently, $x(\tau)$ which is found by the successive algorithm is the right solution of the system. Moreover, the error of approximating $x(\tau)$ with $x^{(m)}(\tau)$ can be bounded

as

$$\begin{aligned}
\|x(\tau) - x^{(m)}(\tau)\| &= \lim_{p \rightarrow \infty} \|x^{(p)}(\tau) - x^{(m)}(\tau)\| \\
&= \lim_{p \rightarrow \infty} \left\| \sum_{r=m}^p (x^{(r+1)}(\tau) - x^{(r)}(\tau)) \right\| \\
&\leq \lim_{p \rightarrow \infty} \sum_{r=m}^p \|x^{(r+1)}(\tau) - x^{(r)}(\tau)\| \\
&\leq \lim_{p \rightarrow \infty} \left[\sum_{r=m}^p \frac{c^{r-1} \tau^r}{r!} \right] o(1) \leq \frac{1}{c} e^{c\tau} o(1).
\end{aligned} \tag{B.9}$$

Since τ is bounded from above on $\Omega_n^{(\varepsilon)}$, then we can say from equation (B.9) that $\|x(\tau) - x^{(1)}(\tau)\| = o(1)$ which yields to $x(\tau) = x^{(1)}(\tau) + o(1)$. ■

Appendix C

Proof of Proposition 2.3.3

Proof. Since we have the Lipschitz property on $D(n)$, the proof of part [i](#) of the theorem is immediate in the theory of ordinary differential equations.

To bring the proof of the part [ii](#), we simply present the verifications for arbitrarily $\ell \in \{1, \dots, b(n)\}$. Taking $\lambda > \lambda_1$ as in part [ii](#), then define

$$\omega = \left\lceil \frac{n\lambda}{\rho} \right\rceil. \tag{C.10}$$

The assertion of the theorem would be trivial if $\rho/\lambda > n^{1/3}$, namely the probability in the conclusion is not restricted, therefore $w = \Omega(n^{2/3})$. Moreover, in the conclusion of part [ii](#), we would like to have $\lambda = o(1)$, otherwise the result would not be as interesting as expected. So we break up the interval $[0, n\sigma(n)]$ into pieces of length $w(n)$, and show the concentration of increments of $Y_\ell(t)$ over one of these pieces, say $[t, t + \omega]$:

$$Y_\ell(t + \omega) - Y_\ell(t).$$

We will assume that for sufficiently large C , $(t/n, Y_1(t)/n, \dots, Y_b(t)/n)$ is at $C\lambda$ distance of the boundary of $D(n)$, thus we can still use trend equation, bounded increment and Lipschitz property. Consequently, a supermartingale can be constructed from the trend equation in (2.19). For $0 \leq k < \omega$ we have:

$$\begin{aligned} \mathbb{E}(Y_\ell(t+k+1) - Y_\ell(t+k) | \mathcal{F}_t) &= f_\ell((t+k)/n, Y_1(t+k)/n, \dots, Y_b(t+k)/n) + O(\lambda_1) \\ &= f_\ell(t/n, Y_1(t)/n, \dots, Y_b(t)/n) + O\left(\frac{k\beta\mu}{n} + \lambda_1\right). \end{aligned} \tag{C.11}$$

The second equality above follows from Lipschitz property and bounded increments of all stochastic processes for $t \leq T_{D(n)}$, i.e. $|Y_\ell(t+k) - Y_\ell(t)| \leq k\beta$ for all ℓ . Thus, there exists a function

$$g(n) = O\left(\frac{w\beta\mu}{n} + \lambda_1\right) = O(\lambda), \tag{C.12}$$

such that conditioned on \mathcal{F}_t :

$$M_k = Y_\ell(t+k) - Y_\ell(t) - kf_\ell(t/n, Y_1(t)/n, \dots, Y_b(t)/n) - kg(n),$$

for $k = 1, \dots, \omega$ is supermartingale with respect to the sequence of σ -fields $\mathcal{F}_t, \dots, \mathcal{F}_{t+\omega}$.

The difference between successive elements of this supermartingale can be bounded

as:

$$\begin{aligned}
|M_{k+1} - M_k| &\leq \beta + |f_\ell(t/n, Y_1(t)/n, \dots, Y_b(t)/n) + g(n)| \\
&\leq \beta + |f_\ell(0, Y_1(0)/n, \dots, Y_b(0)/n)| + \frac{t\beta\mu}{n} + O(\lambda) \\
&\leq O(\beta) + \beta\sigma\mu + O(\lambda) = O(\beta\mu) = \kappa\rho(n)
\end{aligned} \tag{C.13}$$

for some κ . Last inequality above comes from $t/n \leq \sigma(n)$. Now from Azuma's inequality one could bound the upper tail of the supermartingale as

$$\mathbb{P}\left(M_w \geq \kappa\rho\sqrt{2\omega\alpha} \mid \mathcal{F}_t\right) \leq e^{-\alpha}. \tag{C.14}$$

The lower tail of M_w can also be bounded by considering $-M_k$ as a submartingale and applying the Azuma's inequality for submartingales with bounded increments. Then, we will get:

$$\mathbb{P}\left(|Y_\ell(t + \omega) - Y_\ell(t) - wf_\ell(t/n, Y_1(t)/n, \dots, Y_b(t)/n)| \geq wg(n) + \kappa\rho\sqrt{2\omega\alpha} \mid \mathcal{F}_t\right) \leq 2e^{-\alpha}. \tag{C.15}$$

We continue by setting

$$\alpha = \frac{n\lambda^3}{\rho^3}, \tag{C.16}$$

and representing the endpoints of the intervals by $k_i = i\omega$, $i = 0, 1, \dots, \lfloor n\sigma/w \rfloor$, then we next show by induction that

$$\mathbb{P}(\exists j \leq i \ \& \ 1 \leq \ell \leq b : |Y_\ell(k_j) - z_\ell(k_j/n)n| \geq B_j) = O(bie^{-\alpha}), \tag{C.17}$$

where

$$B_j = \frac{n}{\mu} \left(\lambda + \frac{\omega\mu}{n} \right) \left[\left(1 + \frac{B\omega\mu}{n} \right)^j - 1 \right]. \quad (\text{C.18})$$

The first step of the induction holds, because we have already set $z_\ell(0) = Y_\ell(0)/n$ for all ℓ . Note that

$$|Y_\ell(k_{i+1}) - z_\ell(k_{i+1}/n)| = |A_1 + A_2 + A_3 + A_4|, \quad (\text{C.19})$$

where

$$\begin{aligned} A_1 &= Y_\ell(k_i) - z_\ell(k_i/n)n \\ A_2 &= Y_\ell(k_{i+1}) - Y_\ell(k_i) - \omega f_\ell(k_i/n, Y_1(k_i)/n, \dots, Y_b(k_i)/n) \\ A_3 &= \omega z'_\ell(k_i/n) + z_\ell(k_i/n)n - z_\ell(k_{i+1}/n)n \\ A_4 &= \omega f_\ell(k_i/n, Y_1(k_i)/n, \dots, Y_b(k_i)/n) - \omega z'_\ell(k_i/n). \end{aligned} \quad (\text{C.20})$$

From the induction hypothesis $|A_1| < B_i$ for all ℓ , with probability $1 - O(bie^{-\alpha})$. For every single ℓ , A_2 can be bounded with high probability of $1 - O(e^{-\alpha})$ by concentration inequality in (C.15), and using (C.10), (C.12) and (C.16) like

$$\begin{aligned} |A_2| &\leq wg(n) + \kappa\rho\sqrt{2\omega\alpha} \\ &= wg(n) + \sqrt{2}\kappa\omega\lambda = O(\omega\lambda) \leq B'\omega\lambda, \end{aligned} \quad (\text{C.21})$$

for a universal constant B' (independent of n). One could claim that for all ℓ , A_2 can be bounded as above with probability $1 - O(be^{-\alpha})$, namely from (C.15) and (C.21)

$$\mathbb{P} \left(\begin{array}{c} \exists 1 \leq \ell \leq b : \\ |Y_\ell(k_{i+1}) - Y_\ell(k_i) - wf_\ell(k_i/n, Y_1(k_i)/n, \dots, Y_b(k_i)/n)| \geq B'\omega\lambda \end{array} \middle| \mathcal{F}_t \right) \leq O(be^{-\alpha}). \quad (\text{C.22})$$

Now since z_ℓ is the solution of differential equation system in part i, and on the other hand elements of the sequence $\{f_\ell\}$ satisfy Lipschitz condition with uniform constant $\mu(n)$ over all ℓ , then

$$\begin{aligned} |A_3| &= |\omega z'_\ell(k_i/n) + z_\ell(k_i/n)n - z_\ell(k_{i+1}/n)n| \\ &= \omega |z'_\ell(k_i/n) - z'_\ell(\zeta_i/n)| \\ &= \omega |f_\ell(k_i/n, z_1(k_i/n), \dots, z_b(k_i/n)) - f_\ell(\zeta_i/n, z_1(\zeta_i/n), \dots, z_b(\zeta_i/n))| \\ &= O\left(\frac{\omega^2 \mu}{n}\right) \leq \frac{B'' \omega^2 \mu}{n} \end{aligned} \quad (\text{C.23})$$

holds almost surely, with a suitable constant B'' for all ℓ . No need to mention that the second equality above follows from mean-value theorem for $k_i \leq \zeta_i \leq k_{i+1}$. The upper bound on $|A_4|$ is also readily drawn from the upper bound on $|A_1|$, like

$$\begin{aligned} |A_4| &= |\omega f_\ell(k_i/n, Y_1(k_i)/n, \dots, Y_b(k_i)/n) - \omega z'_\ell(k_i/n)| \\ &= \omega |f_\ell(k_i/n, Y_1(k_i)/n, \dots, Y_b(k_i)/n) - f_\ell(k_i/n, z_1(k_i/n), \dots, z_b(k_i/n))| \\ &\leq \omega \mu \max_\ell |Y_\ell(k_i)/n - z_\ell(k_i/n)| \leq \frac{B'' \omega \mu B_i}{n}, \end{aligned} \quad (\text{C.24})$$

(redefining B'' appropriately so as the last inequality in (C.24) holds). Note that because of exploiting the upper bound of $|A_1|$, subsequently equation (C.24) is also

true with the same probability as $|A_1| < B_i$ holds. By setting $B = \max\{B', B''\}$, and adding together these upper bounds, we will get:

$$\begin{aligned} |Y_\ell(k_{i+1}) - z_\ell(k_{i+1}/n)| &\leq B_i + B\omega\lambda + \frac{B\omega^2\mu}{n} + \frac{B\omega\mu B_i}{n} \\ &= B_i \left(1 + \frac{B\omega\mu}{n}\right) + B\omega \left(\lambda + \frac{\omega\mu}{n}\right) = B_{i+1}, \end{aligned} \quad (\text{C.25})$$

with probability $1 - O(b(i+1)e^{-\alpha})$, which ends the induction proof. Noticing that $B_i = O(\lambda n)$, the assertion of part [ii](#) of the theorem is verified for the endpoints of intervals. For any $t \leq n\sigma(n)$, put $i = \lfloor t/\omega \rfloor$, then by taking into account that the change in Y_ℓ and z_ℓ from time k_i to t is at most $\omega\beta = O(\lambda n)$, we have:

$$|Y_\ell(t) - z_\ell(t/n)n| = O(\lambda n) \quad \forall \ell \quad \text{with probability} \quad 1 - O\left(\frac{bn\sigma}{w}e^{-\alpha}\right). \quad (\text{C.26})$$

■

Bibliography

- [1] Paul Erdos and A Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hungar. Acad. Sci.*, 5:17–61, 1960.
- [2] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *science*, 286(5439):509–512, 1999.
- [3] Albert-Laszlo Barabasi and Zoltan N Oltvai. Network biology: understanding the cell’s functional organization. *Nature Reviews Genetics*, 5(2):101–113, 2004.
- [4] Réka Albert, Hawoong Jeong, and Albert-László Barabási. Internet: Diameter of the world-wide web. *Nature*, 401(6749):130–131, 1999.
- [5] Béla Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European Journal of Combinatorics*, 1(4):311–316, 1980.
- [6] Béla Bollobás. *Random graphs*. Springer, 1998.
- [7] Bela Bollobas and Oliver Riordan. A simple branching process approach to the phase transition in $g_{n,p}$. *Electronic Journal of Combinatorics*, 19:P21, 2012.
- [8] Nicholas C Wormald. Some problems in the enumeration of labelled graphs. *Bulletin of the Australian Mathematical Society*, 21(01):159–160, 1980.

-
- [9] Edward A Bender and E Rodney Canfield. The asymptotic number of labeled graphs with given degree sequences. *Journal of Combinatorial Theory, Series A*, 24(3):296–307, 1978.
- [10] Remco Van Der Hofstad. Random graphs and complex networks. Available on <http://www.win.tue.nl/rhofstad/NotesRGCN.pdf>, 2009.
- [11] Svante Janson, Donald E Knuth, Tomasz Łuczak, and Boris Pittel. The birth of the giant component. *Random Structures & Algorithms*, 4(3):233–358, 1993.
- [12] Tom Britton, Maria Deijfen, and Anders Martin-Löf. Generating simple random graphs with prescribed degree distribution. *Journal of Statistical Physics*, 124(6):1377–1397, 2006.
- [13] Svante Janson. The probability that a random multigraph is simple. *Combinatorics, probability & computing*, 18(1-2):205–225, 2009.
- [14] Béla Bollobás, Oliver Riordan, Joel Spencer, Gábor Tusnády, et al. The degree sequence of a scale-free random graph process. *Random Structures & Algorithms*, 18(3):279–290, 2001.
- [15] Béla Bollobás, Christian Borgs, Jennifer Chayes, and Oliver Riordan. Directed scale-free graphs. In *Proceedings of the fourteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 132–139. Society for Industrial and Applied Mathematics, 2003.
- [16] Paul L Krapivsky, Sidney Redner, and Francois Leyvraz. Connectivity of growing random networks. *Physical review letters*, 85(21):4629, 2000.

- [17] Mark EJ Newman. Assortative mixing in networks. *Physical review letters*, 89(20):208701, 2002.
- [18] Ningyuan Chen, Mariana Olvera-Cravioto, et al. Directed random graphs with given degree distributions. *Stochastic Systems*, 3(1):147–186, 2013.
- [19] Nicholas C Wormald. Differential equations for random processes and random graphs. *The annals of applied probability*, pages 1217–1235, 1995.
- [20] Hamed Amini. Bootstrap percolation and diffusion in random graphs with given vertex degrees. *the electronic journal of combinatorics*, 17(1):R25, 2010.
- [21] Fan Chung and Linyuan Lu. Connected components in random graphs with given expected degree sequences. *Annals of combinatorics*, 6(2):125–145, 2002.
- [22] Thomas R. Hurd. *Contagion! The Spread of Systemic Risk in Financial Networks*. Working project, 2015.