



HAL
open science

Local matching algorithms on the configuration model

Mohamed Habib Aliou Diallo Aoudi

► **To cite this version:**

Mohamed Habib Aliou Diallo Aoudi. Local matching algorithms on the configuration model. Mathematics [math]. Université de technologie de Compiègne, 2023. English. NNT: . tel-04122066

HAL Id: tel-04122066

<https://hal.science/tel-04122066>

Submitted on 15 Jun 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

École doctorale 71 : Sciences pour l'Ingénieur

Doctorat UTC

THÈSE

pour l'obtention du grade de docteur, délivré par

Université de Technologie de Compiègne

Spécialité doctorale "Mathématiques et Applications"

présentée et soutenue publiquement par

Mohamed Habib Aliou Diallo Aoudi

le 13 Juin 2023

Local Matching Algorithms on the Configuration Model

Directeur de thèse : **Pascal Moyal**

Co-encadrant de thèse : **Vincent Robin**

Jury

M. Nicolas Broutin,	Professeur, Sorbonne Université	Rapporteur
Mme. Mariana Olvera-Cravioto,	Professeure, UNC Chapel Hill	Rapporteuse
Mme. Ghislaine Gayraud,	Professeure, UTC Compiègne	Examinatrice
Mme. Amandine Veber,	Senior researcher, CNRS	Examinatrice
Mme. Anna Ben-Hamou,	Maître de conférences, Sorbonne Université	Examinatrice
M. Vianney Perchet,	Professeur, ENSAE	Examinateur

Université de Technologie de Compiègne

Laboratoire de Mathématiques Appliquées de Compiègne

57 Av. de Landshut, 60200 Compiègne, France

Abstract

The present thesis constructs an alternative framework to online matching algorithms on large graphs. Using the configuration model to mimic the degree distributions of large networks, we are able to build algorithms based on local matching policies for nodes. Thus, we are allowed to predict and approximate the performances of a class of matching policies given the degree distributions of the initial network. Towards this goal, we use a generalization of the differential equation method to measure valued processes. Throughout the text, we provide simulations and a comparison to the seminal work of Karp, Vazirani and Vazirani based on the prevailing viewpoint in online bipartite matching.

Dedication

To all the dearly beloved, lost. You will always be missed.

Acknowledgements

I could never be grateful enough to my supervisors, Pascal Moyal and Vincent Robin, for their unwavering kindness, their guidance, and mentorship throughout my PhD. These last few years were not easy, especially on the personal level but their feedback, criticism, and more importantly their support were invaluable.

I would like to thank the members of my thesis committee. The remarks of Mr Broutin and Ms Olvera-Cravioto gave me invaluable insights on my work. I thank Ms Gayraud, Ms Ben Hamou, Ms Veber and Mr Perchet for their valuable comments and suggestions whether they were given directly through courses and comments or indirectly through works and conferences.

I am indebted to my colleagues in the Université de Technologie de Compiègne for their insight, their help and their opinions. Especially, I would like to thank Ms Schaefflen from LMAC, Ms Prella, Ms Kaczkowski and Ms Aaras-Andaloussi from the ED for they were always quick to give me their support in all matters.

I dare not say that I am grateful to my family. I wish to defend before picking a fight upon what is 'normal' to do for each other. What I will say is that I am blessed to have them.

I would also like to thank my friends for their emotional support and encouragement. They know I owe them this PHD and even more.

Thank you all for your contributions and support.

Contents

I	Introduction	13
I.1	The early days of matching theory	13
I.2	Birth of an algorithmic approach	16
I.3	Modern days	18
I.4	Motivation	21
I.5	Layout of the document	23
II	Local Matching Algorithms	25
II.1	Local Matching Algorithms on Given Graphs	25
II.1.1	Matching criteria	26
II.1.2	Local matching through exploration	27
II.2	Local matching on the configuration model	30
II.2.1	Matching criteria on the CM	31
II.2.2	The matching process on the CM	32
II.3	Simulations	36
II.3.1	Convergence of the Matching Size : a case study on 3-regular graphs	37
II.3.2	An Array of Degree Distributions	38
II.3.3	The Influence of Topology and Another Approach To Online Bipar- tite Matching: Optimal Algorithm by Karp, Vazirani and Vazirani .	41
II.3.4	Notes	43
III	Local Matching Algorithms as Markov processes	45
III.1	Notations and Hypotheses	46
III.1.1	Main Hypotheses	46
III.2	Discrete time Markov chain	48
III.2.1	Measure valued representation and Markov Chain dynamic	48
III.2.2	Moment estimates and properties	52
III.3	Normalization and scaling	56
III.3.1	Normalization	56

III.3.2	Continuous time Markov chain	58
III.3.3	Semi-martingale representation	60
IV	Performance Prediction and Concentration Results	65
IV.1	Main theorem	66
IV.1.1	Convergence of local matching algorithms	69
IV.1.2	Further simulations	70
IV.2	Proof of the main convergence theorem	72
IV.3	Illustrative example : GREEDY	83
IV.3.1	Proof of Theorem IV.3.1: Well behavedness and moment properties	84
IV.3.2	Proof of Theorem IV.3.1: Characterization of the solutions	85
V	Notes and conclusion	93
A	Appendix	95
A.1	The Configuration Model and Degree Sequences	95
A.1.1	Algorithmic Construction of a Configuration	96
A.1.2	The CM as an approximation for large simple graphs	96
A.1.3	The Independence Property	97
A.1.4	Graphicality of degree sequences	98
A.2	Tools for the convergence of processes	99
A.2.1	Topological Properties	99
A.2.2	Characterizing Convergence	100

Frequently used Notations and Conventions

We introduce several of the notations and conventions later used in the text. All of the notations introduced here will properly be reintroduced in context.

Graph related objects and sets are illustrated using the *mathcal* calligraphy. Thus, $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is a graph on the vertex set \mathcal{V} with edge set \mathcal{V} . For a node I , the notation $\mathcal{N}(I)$ is also often used. It denotes the *neighbors* of node I *apart from itself*. Oftentimes, objects will also be indexed by a time/iteration indicator. Thus, $\mathcal{N}_t(I)$ simply means the neighbors of I at time/iteration t .

\mathbb{N} and \mathbb{R} are respectively the set of integers and that of real numbers. Both are endowed with their natural topology.

Likewise, for any topological space E , $\mathcal{C}_b(E)$ (respectively $\mathcal{C}_c(E)$) is the space of real valued bounded (respectively compactly supported) functions on E endowed with the uniform topology.

The *Borel* sigma-algebra $\mathcal{B}(E)$ on E is the sigma-algebra generated by the topological open sets and we denote $\mathcal{M}_F(E)$ (resp. \mathcal{M}) the space of finite measures on E (resp. the finite measure with mass ≤ 1). Both the weak (induced by the duality with $\mathcal{C}_b(E)$) and vague (induced by $\mathcal{C}_c(E)$) topology will be relevant to our work.

$\mathbb{D}(\mathbb{R}^+, E)$ is the space of right continuous, left limited functions from \mathbb{R}^+ to E .

A generic probability space $(\Omega, \mathfrak{F}, \mathbb{P})$ is also used to handle all the random variables.

An E valued random variable is a measurable function from $(\Omega, \mathfrak{F}, \mathbb{P})$ to $(E, \mathcal{B}(E))$. And we define a random process as $\mathbb{D}(\mathbb{R}^+, E)$ valued random variable. Sometimes, processes are called '*dynamics*' or '*algorithms*'.

Chapter I

Introduction

I.1 The early days of matching theory

Considered as one of Leonhard Euler's most notorious papers, 1735's *Solutio problematis ad geometriam situs pertinentis* presents the solution to a puzzle: Euler established that there was no path which crossed all the seven bridges of Königsberg once and only once. Today, this paper is viewed as the earliest work both in graph theory and in topology. But, where topology broke free of its recreational roots and quickly expanded into its own fundamental branch of mathematics, the history of graph theory is more peculiar. Most of its early developments stayed rooted in recreational puzzles, consequently making it thorny to paint an accurate picture (see [Mitsuko, 2010]). In fact, it took more than a century and a half for *matchings* to become a definitive subject of study in graph theory.

The inception of graph theory as we know is widely attributed to Julius Petersen and Dénes König. *Matching theory*, an earlier focus of their work, thus became a foundational concept of graph theory. While Petersen worked mostly on regular graphs, he touched matching theory in an oblique way. By **Regular graphs**, it is meant graphs where all the vertices have the same degree. In 1891, [Petersen, 1891] Petersen reformulated an algebraic factorization problem due to Hilbert as a graph factorization problem, already hinting at deeper ties between algebra and the newborn graph theory. A **k-factor** of a graph \mathcal{G} is a subgraph \mathcal{F} such that every vertex has exactly degree k in \mathcal{F} . In particular, a **matching** is a 1-factor: each vertex in a matching has exactly one neighbor and thus is *matched* to it. In the same paper, Petersen remarkably proved that any regular graph of even degree is in fact the reunion of two edge disjoint 2-factors, tying his results to the work of Euler 150 years earlier. More importantly (for matching theory), he noticed that *odd degree regular graphs having no more than two cutlines* admitted a **perfect matching**, one that covered all the vertices.

Petersen then became interested in the Four Colors Theorem (1852), only a conjecture at the time. A detailed account of the early research on the conjecture is given in [R.J. et al., 1976]. In his pursuit of the conjecture, P. G Tait [Tait, 1880] in 1880 had claimed that every polyhedral cubic graph could be factorised in three disjoint perfect matchings (perfect means every vertex is covered in the matching). Petersen's answer came 10 years later, in *Sur le théorème de Tait* [Petersen, 1898], and it came in the form of a counterexample, perhaps the single most famous graph in graph theory rightly called *The Petersen graph*.

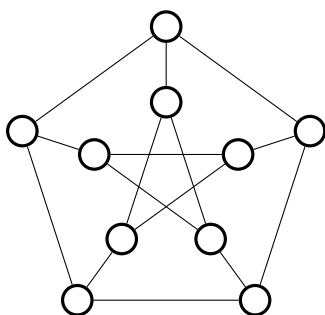


Figure I.1: The Petersen Graph

In parallel to the early works of Petersen, another wing of important results had sprouted under the *father of graph theory*, Dénes König. To this day, this branch stays the most studied part of matching theory. *Bipartite matching* is a fundamental affectation problem. As an illustration, we provide one of its formulations : the *marriage problem*. Say n men and n women are to get married, where n is an arbitrary number. Since they live in a conservative society, polyamorous and homosexual relationships are excluded and each person needs to get married to someone they are acquainted with. The marriage problem asks : *When is it possible for each individual to get married?* or reformulating in graph theoretical terms : *Under which conditions does a bipartite graph admit a perfect matching?*

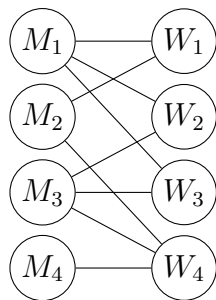


Figure I.2: The marriage problem: When do such bipartite graphs admit a perfect matching?

Again, the study of bipartite matching would find new roots in an algebraic problem. In 1912, Frobenius while working on the reducibility of determinants proved in [Frobenius, 1912] that if M is a $n \times n$ matrix with either zero or unique entries, then the $\det(M)$ is a factorisable polynomial of the non zero entries if and only if there exists a permutation of the rows and columns of M such that it exhibits a $p \times (n - p)$ sized zero block with $0 \leq p \leq n$. Frobenius' proof was hard to follow at the time and König, only 3 years later in [König, 1915], observing the combinatorial aspect of the reasoning, gave a much simpler proof using bipartite graphs by introducing a new representation today called *incidency matrices*. In retrospect, this approach is only a natural path to take and again a preface to the deep ties between linear algebra and graph theory. In the 1910s, König only faced the unwillingness of Frobenius to acknowledge the theory of *line systems* as a useful mathematical theory and his graph-theoretic proof was hence given little credence.

Nevertheless, only one year later the same method of using incidence matrices proved fruitful in establishing the first formal result directly addressing matching theory.

Theorem I.1.1 ([König, 1916]). *Every regular bipartite graph of degree $k \in \mathbb{N}$ is the (edge) disjoint union of k perfect matchings.*

In algebraic terms, this theorem meant every bi-stochastic square matrix (where each row and each column have the same sum) with non negative entries had to have a non zero term in its determinant. In particular, it implies that bi-stochastic square matrices are the convex sum of permutation matrices. This is the *Birkhoff-Von Neumann theorem* and it was rediscovered almost half a decade later by Von Neumann [von Neumann, 1953] and Birkhoff [Birkhoff, 1946].

A few months later in 1917, Frobenius published his own follow up result. In [Frobenius, 1917], he provided a simpler proof of his first lemma using a stronger theorem on determinants that were identically zero. Perhaps ironically, this theorem came to be known mostly in its combinatorial form as the *Marriage Theorem* and it is predominantly used in graph theory and in optimization. The marriage theorem stipulates that the marriage problem admits a solution if and only if each subgroup of k men collectively know at least k women.

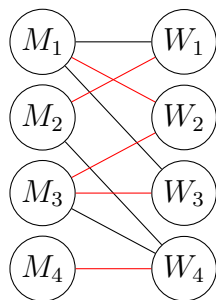


Figure I.3: The marriage problem: A solution

It was also the precursor (in fact, it was discovered to be equivalent to) the single most important characterization theorem for perfect matchings on bipartite graphs:

Theorem I.1.2 (Theorem on Distinct Representatives [Hall, 1935]). *Let S_1, \dots, S_n be a finite collection of sets, there exists a collection of distinct elements x_1, \dots, x_n such that $x_i \in S_i$ if and only if every union of k sets contains at least k elements.*

During the same time frame, through *König's Theorem* (1931), König [König, 1931] established the duality between the maximum matching problem and the minimal edge cover on bipartite graphs (generalized the same year to weighted graphs by Egerváry [Egerváry, 1931]). König's theorem inscribes itself in the line of *minmax* theorems which paved the way to the algorithmic paradigm not only in graph theory, but also in several other combinatorial domains such as linear programming or convex optimization. In fact, in the first graph theory textbook [König, 1936], König showed that both marriage theorems (by Hall and Frobenius) and Menger's graph connectivity theorem [Karl, 1927] were direct consequences of his minmax theorem.

I.2 Birth of an algorithmic approach

After König's book, matching theory stagnated until post world war II. Most results were still characterization theorems focused on regular and bipartite graphs. Then the first breakthrough for general graphs happened in 1947 with [Tutte, 1947], where Tutte generalized Hall's theorem to all graphs :

Theorem I.2.1 (Tutte). *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ a graph. \mathcal{G} has a perfect matching if and only if for every subset \mathcal{U} of \mathcal{V} , the subgraph $\mathcal{G} \setminus \mathcal{U}$ has at most $|\mathcal{U}|$ odd components (connected components having an odd number of vertices).*

It then took Ore eight years [Ore, 1955] to follow Tutte's result by a "defect" version of Hall's theorem. Ore's is called a defect version of Hall's theorem because it identifies

maximum (cardinality) matchings on bipartite graphs instead of *perfect* matchings. And then, where it took Hall's theorem decades to be generalized to all graphs, Berge [Berge, 1957] was able to produce a general defect version of Ore's theorem only one year later:

Theorem I.2.2 (Berge). *A matching \mathcal{M} in a graph \mathcal{G} is of maximum cardinality if and only if there is no augmenting path with \mathcal{M} (a path where endpoint vertices are unmatched and that alternates between edges in and out the matching).*

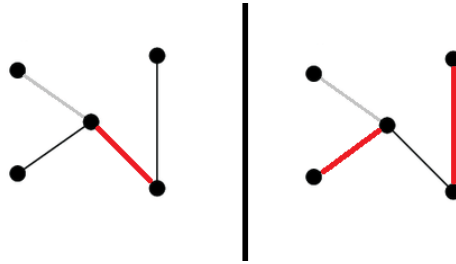


Figure I.4: Augmentation of a path

More than the identification of maximum matchings, Berge rightfully noticed that his theorem yielded a naive algorithm for finding maximum matchings. He pointed that to build a maximum matching, one simply had to start from any maximal matching \mathcal{M} then search for augmenting paths between any two pairs of edges. Once an augmenting path \mathcal{P} was found, $(\mathcal{M} \setminus \mathcal{P}) \cup (\mathcal{P} \setminus \mathcal{M})$ was a larger matching: the procedure is called an *augmentation* (see figure I.4) of the matching and it was to be repeated until a maximum matching was reached. To this day, most maximum matching algorithms are based either around augmentation or the use of equivalent objects to matchings (Maximum Flow for example).

Berge's observation appends the question that has been looming over matching theory for decades. The existence of maximum/perfect matchings could be fully characterized, yet a way to exhibit such objects was still to be explicated. Incidentally, these questions overlapped with the invention of computers immediately after World War II and were followed by a surge of interest in algorithms. It was simpler to build matching algorithms on bipartite graphs, since the basic steps had already been laid in the 1930s by König and Egervàry. Thus, when Kuhn [1955] and Hall [1956] built the first bipartite matching algorithm, Kuhn named it *The Hungarian Method* as an homage to both König and Egervary. The hungarian method also provided a way to cast bipartite matching as a linear programming question in the light of a primal-dual setting which could be solved by the simplex method (Dantzig [1951]), producing another equivalence for maximum bipartite matching. Almost simultaneously, Ford and Fulkerson published the first papers on The Theory of Network Flows [Ford and Fulkerson, 1956], [Ford and Fulkerson, 1957]

[Ford and Fulkerson, 1958] which again permitted another formulation of the already rich maximum matching problem with the use of the *Max Flow Min Cut theorem*.

As milestones were being set on bipartite graphs, matching on general graphs turned out to be a considerably more difficult problem. Almost a decade elapsed before Gallai published [Gallai, 1964] in 1964. Using another existence theorem, Gallai established that every graph admitted a canonical decomposition in terms of its maximum matchings. The *Gallai-Edmonds Structure Theorem* proved to be the much needed breakthrough because the next year, Edmonds [1965] was headlining BLOSSOM. BLOSSOM was the first non trivial algorithm for maximum matching on general graphs, and it settled as a staple not only in graph theory but also in the burgeoning *algorithmic complexity theory*. Its complexity set up *polynomial* as a criteria for performance and efficiency for algorithms.

After 1965, the classical theory of maximum matching mostly settled on using Berge’s Theorem in the most way efficient possible like BLOSSOM did. Research is still about improving the search for augmenting paths or attacking matchings using the ever expanding multitude of equivalent formulations (an example is given in Lovász Grötschel et al. [1981]). We defer to Dua and Pettie [2014] for an up to date survey of the modern techniques in the theory.

I.3 Modern days

Even as the classical matching theory was settling and computers were being commercialized to the public, the newborn computer science faced its first obstacles. Computers were severely limited by their hardware. In theory, a multitude of algorithms had been devised to solve problems. In practice, only a few were usable because the first computers simply could not handle them. This led to the emergence of newer techniques whose prime focus would be viability.

The first of such techniques that is relevant to our work is randomization (eg. Karp [Karp, 1991]). *Random algorithms* had existed alongside the deterministic ones from the early days of the theory. However, randomization as an important technique to drive down the cost of algorithms really started getting traction only after the fifties. In 1946, after a review of the performances of the ENIAC, the first computer, Ulam and later by Von Neumann saw the potential of the use of statistical sampling to simplify calculations. Later, as they needed a codename for their new method, they would rebrand it as the “Monte-Carlo” method ¹ [Metropolis and Ulam, 1949; Metropolis, 1987]. The Monte Carlo method is today a staple in almost all domains that require complicated calculations, but at its beginnings, the technique only gave a starting push to the use of randomization

¹Referring to the Monte Carlo Casino, where Ulam’s uncle would gamble.

in algorithms. Its importance as a central tool in algorithms would only get traction later, after both Solovay [Solovay and Strassen, 1977] in 1977 and Rabin [Rabin, 1980] in 1980 published randomized primality tests.

The second relevant technique is the use of online algorithms. The online setup is defined by the need for algorithms to process their input progressively without total knowledge of the data. It models situations where the resources are limited (for example the first computers were very limited in memory). The performance of online algorithms is then tested against that of offline optimal algorithms, algorithms to which the totality of the input is made available from the start. This is manifested by the *competitive ratio*, a ratio between the worst case result for the online algorithm and the optimal offline one. Competitive analysis for online algorithms had been first introduced in 1985 by Sleator and Tarjan [1985] for list updates and paging rules and after the paper of Manasse in 1986 Manasse et al. [1988], it is, to this day the main approach to the performance analysis of online algorithms. In the 1990s, the theory of online algorithms exploded with its use on various topics Borodin et al. [1996]; Karp [1992]. In particular, it resulted in the seminal paper for online bipartite matching: *An Optimal Algorithm for On-Line Bipartite Matching* [Karp et al., 1990] by Karp, Vazirani and Vazirani in 1990.

Karp et al. [1990] presented RANKING an optimal² algorithm for *Online bipartite matching*. In the paradigm they studied, the graph is split into *customers* and *items*. Customers arrive one by one, and it is the role of the matching algorithm to pick one of the items for a match. The use of RANKING led to the now classical $1 - 1/e$ bound for the *matching coverage*, namely, the ratio of matched items out of the total set.

Theorem I.3.1. *The performance of any on-line bipartite matching algorithm is $< n(1 - \frac{1}{e}) + o(n)$.*

Since then, applications in online advertising prompted the construction of algorithms that are able to beat this benchmark under various conditions. Such algorithms are the subject of an important line of research, see e.g. Goel and Mehta [2008]; Feldman et al. [2009] and references therein. More recently, various extensions of the online bipartite matching problem have been proposed among which, stochastic matching by Borodin et al. [2020] (meaning that each edge emanating from the online nodes exists with a given probability), random customer arrivals by Mahdian and Yan [2011], or models with patience times by Brubach et al. [2021]. Because of the importance of online advertisement, online bipartite matching is today, in terms of volume *the most used aspect of matching theory* (Mehta [2013]).

²In the competitive ratio sense

In computer science, online algorithms went through a rebranding. *Streaming algorithms* are algorithms that only process a stream of the data. This line of research mainly focuses on the type of input for the stream (either dynamic or enter only), its randomness or its relative size to the whole graph. Thus several authors have worked to apply them to the matching problem (see e.g. Chen et al. [2021], Farhadi et al. [2019] and Tirodkar [2018], and references therein).

Meanwhile, newer mathematical techniques introduced newer bounds for matching sizes on certain classes of graphs. These bounds furthered the computations of the competitive ratio of matching algorithms despite being overall graph dependent. The first model of random graph was popularized in the sixties by Erdős and Rényi (reprinted for example in Erdős and Rényi [2006]) and Erdős used it for the *Probabilistic Method*. The probabilistic method is a technique for proving the existence of a graph with certain properties (say a perfect matching) by constructing an appropriate probability space and showing that a randomly chosen element in this space had the desired properties with positive probability. This probability is viewed as a proportion of graphs with the given property on the space. Thus a non zero probability implied the existence of the desired object. In matching theory, the probabilistic method was used by Erdős and Rényi in [Erdős, 1965] to exhibit a size transition for the appearance of a perfect matching on the Erdős-Rényi random graph, almost surely characterizing the parameters at which a perfect matching appeared. The book of Lovász and Plummer [2009] contains a more precise account of the relevant techniques and how they relate to matching theory.

The usage of random graphs brought the necessary second shift in perspective leading to our work. Coming in the 20th century, Watts and Strogatz had introduced the small world model (Watts and Strogatz [1998]) and it was followed by the preferential attachment model of Albert and Barabási [2002]. These two models led to an explosion in the research because they exhibited properties observed in real world complex networks. Internet, social networks or protein chains all shared the property that they were difficult to describe because of their sheer size. As a consequence, their complete description was utterly impossible both in theory, because there was no unifying definition of limits for graphs, and in practice because it would cost an unfathomable amount of resources to have a complete and precise map of say the internet. This naturally led to a more statistical and local approach to networks : the theory would study the rules by which vertices were locally connected to one another and how those rules were to change when the graph size grew arbitrarily large. This is the so called *local limit* of graphs (in the Benjamini-Schramm sense). And, oftentimes the edge defining rules were defined through probabilistic means.

Thus, random graphs became the perfect tool for mimicking properties of real networks

since they naturally integrate a layer of randomness in their structure, akin to how the structure of complex networks is often studied only through sampling. The chosen random graph model (in our case the configuration model) is then to be viewed as the underlying model producing the complex network under study, similarly to statistics where samples are assumed to be generated from a given law³.

I.4 Motivation

In this thesis, we present an alternative to online algorithms using the local construction of maximal matchings provided by random graphs. Over the structural randomness of the configuration model, we construct local matching algorithms that are built around the choice of a matching criterion (a preferential choice of edges to add to the matching for every node).

Afterward, we will manage to predict the performance of chosen criteria given a graph (or at least its degree distribution). Moreover, our results apply directly to general graphs instead of being restricted to the usual bipartite matching problem. In the traditional online bipartite matching, there is a layer of randomness that comes from the order of arrivals, from the exploration of the graph. Our goal is to add a second layer, representing the exploration, to the first layer of randomness already acknowledging an uncertainty in the structure of large networks. To capture the randomness of the structure, we use the configuration model. Even for large networks, characteristics like degree distributions are more accessible than the whole topology and have huge effects on the networked systems. Because it emulates such an important characteristic, the configuration model is one of the most important theoretical models in the study of networks. In particular, since it can emulate a uniformly drawn graph from a degree sequence, it strikes an ideal balance between realism and simplicity. Hence, given a matching criterion, we *approximate* the matching size as a function of the degree sequence of the graph. This is done through a deterministic approximation of a stochastic process representing the matching procedure⁴, in the large graph asymptotic. The general idea is as follows: rather than precisely defining and keeping track of the whole geometry of the graph, we generate a graph from the configuration model (Bollobás and Béla [2001] and Hofstad [2016]). To this end, we use a classical uniform pairing procedure of the *half-edges* of the nodes (a more precise description is given in the next chapter). This construction leads a uniformly drawn realization of a graph having the prescribed degree distribution. In parallel, we simultaneously construct an online matching (meaning in this context, that each edge

³We are very grateful to Ms Olvera-Cravioto for her ideas on the subject

⁴given the chosen criterion

that is added to the matching cannot be erased afterward) on the resulting graph. This can be viewed as using simultaneously the two layers of randomness: the structure of the graph (1st layer) is discovered alongside the exploration process (2nd layer) associated to the matching criterion.

The whole procedure can be viewed as an exploration of the configuration model, where the edges to be added to the matching are marked following the designated criterion. Our idea of viewing the matching algorithms as exploration processes originate from the course *Limites d'échelles de graphes aléatoires* Broutin [2021] by N. Broutin at Sorbonne Universités. In an adjacent paper Broutin and Marckert [2014] to the course, an encoding for the exploration of trees with prescribed degree distributions is provided. In our case, the interest lied more in the number of marked edges than in the structure of the explored graph. Thus, the marking process is encoded in a different manner.

The simultaneous construction leads to a simple Markov representation, keeping track of the *remaining degrees* of the nodes that are not yet fully attached to the graph (a definition that will be made precise hereafter), provided that the matching algorithm depends only on these remaining degrees. This will be our preferred encoding of the exploration. We say in that case that the matching algorithm is *local*. The underlying Markov process is then measure-valued, where each measure is a sum of Dirac masses marking the remaining degrees of the nodes. By doing so, we do not need to keep track of the precise form of the constructed graph, and in fact, we do not need to have access to it. Then, using usual approximation tools for Markov processes, one can identify the approximation of the considered process as the solution of an ordinary differential equation. This results in a generalization, for measure-valued processes, of the celebrated *Differential Equation Method*, introduced by Kurtz [1970, 1969] and later popularized in the random graphs community by Wormald [1999]. By doing so, we retrieve an estimate of the resulting matching coverage as a simple function of the latter (deterministic) solution, without knowledge of the precise geometry of the graph at hand. Remarkably, we show hereafter that the resulting matching coverage has the same distribution as the one obtained when applying the corresponding online and local matching algorithm on a previously constructed graph \mathcal{G} , conditional on the fact that the resulting graph constructed by the CM is precisely \mathcal{G} . As a consequence, our estimate of the matching coverage by the differential equation method, provides a remarkably accurate estimation of the matching coverage of the considered local algorithm on a given graph, a result that we support with simulations.

In fact, at the same moment the present thesis was being written Noiry et al. [2021] wrote a paper constructing a greedy algorithm on the bipartite configuration model. Their work focused on the bipartite case using a representation of the matching process leaning

on generating functions. Despite using an overall different approach, the empirical results obtained are similar to those obtained in restricting our version to the greedy matching (Aoudi et al. [2022]). Hence, our aim in this thesis is a generalization of this approach to general graphs and to broader matching criteria. In any case, both the author and Noiry et al. [2021] go even deeper in the comparison of this alternative technique to existing methods in online bipartite matching.

The extension of the differential equation method to measure-valued processes, resulting from a simultaneous construction of the CM and an exploration algorithm on the latter, first appeared in a paper by Decreusefond et al. [2012] to describe the propagation of an SIR epidemics on an heterogeneous graph. A closely related idea was then applied in Bermolen et al. [2017] to approximate the size of maximal independent sets on graphs with given degree distributions (for a more direct use of the differential equation method on the same topic, see also Brightwell et al. [2015]). This led to various extensions, for example to address various coverage problems in CSMA-type algorithms for radio-mobile and ad-hoc communication networks, see Bermolen et al. [2016].

Before that, measure-valued processes Markov processes were first introduced in the queueing literature. Space of measures are amenable to showing weak convergence of sequences of processes under an appropriate scaling, and the exhaustive representation of queueing systems by point measures, in which each Dirac mass typically represents the characteristic of a customer in line, led to fruitful developments in fluid and diffusion approximations of the systems at hand, see e.g. Gromoll et al. [2002] for processor sharing queues, Doytchinov et al. [2001]; Decreusefond and Moyal [2008a] for queueing systems with impatient customers, Decreusefond and Moyal [2008b] for infinite-server queues, or Kaspi and Ramanan [2011] for many-server queues.

I.5 Layout of the document

This thesis is organized in three main chapters.

In Chapter II, we introduce the framework for the rest of the work. At first, the local marking process is provided for finite deterministic graphs. Here, we build an intuition for the behavior of matching criteria on finite graphs before transposing said criteria to the configuration model. The chapter ends with our first batch of simulations, illustrating the definitions given in the chapter for practical cases.

Chapter III is devoted to the representation of the matching process as a measure valued Markov process. It starts with a theorem linking the exploration of given graphs to that of the configuration model. Then, the properties of the associated process are studied in preparation for Chapter IV. The chapter is closed with a section dedicated to

a proper normalization of the dynamic.

Chapter IV presents the main result of our thesis. We are able to prove the convergence of the matching process to a deterministic function. By studying this function, we show that the matching coverage for given criteria converges to a value we are able to compute by solving an ordinary differential equation. This value serves as a predictor for the performance of the given matching criteria. The convergence is also illustrated empirically through a second batch of simulations.

Chapter II

Local Matching Algorithms

The naive algorithm for building a maximal matching consists of the repetition of the following iteration: **1.** an edge is added to the matching **2.** its neighboring edges are blocked. Maximum matching algorithms would then proceed to progressively *augment* the resulting maximal matching ¹. The present thesis studies local matching criteria and their influence on the size of the first maximal matching produced without augmentation. A *matching criterion* is a procedure selecting the edge added to the matching at each iteration. And it is *local* when it depends only on the neighborhood of an edge.

In this chapter, the general formula for the local matching algorithms studied hereafter is provided. Section II.1 defines the local approach on given graphs as a prelude (the graph is considered a realization of the CM that is simple). The context of given graphs can be viewed as conditioning the model by a given structure. Thus, the structure of the graph is known and available, so the edges locally explored and added to matching already exist and are only to be discovered. In II.2, the process is transposed to the configuration model where instead of discovering the graph, it is *built* alongside the matching, using the uniform pairing procedure (which symbolizes the randomness in structure). Thus, the configuration model serves as a placeholder for large complex networks where complete information is often inaccessible. Finally, we provide a batch of simulations illustrating the whole approach and comparing it to the optimal traditional online algorithm.

II.1 Local Matching Algorithms on Given Graphs

For the entirety of this section, $\mathring{\mathcal{G}} = (\mathring{\mathcal{V}}, \mathring{\mathcal{E}})$ represents a graph with vertex set $\mathring{\mathcal{V}}$ and edge set $\mathring{\mathcal{E}}$. We differentiate given graphs in this section from the configuration model generated ones later by using the $\mathring{\circ}$ symbol. Those graphs will be later addressed without

¹see Introduction/I.2.2

superscripts.

II.1.1 Matching criteria

Definition II.1.1. Let $\mathcal{U} \subset \mathcal{V}$ a collection of nodes, define the following core objects:

- A **choice function** Φ on \mathcal{U} is a probability law on \mathcal{U} . A choice following the function $\Phi_{\dot{i}}$ is a random variable \dot{I} whose law is $\Phi_{\dot{i}}$. For $u \in \mathcal{U}$, we write :

$$\mathbb{P}(\dot{I} = u) := \Phi_{\dot{i}}(\mathcal{U})(u) \quad (\text{II.1})$$

- A **matching criterion** $\Phi := (\Phi_{\dot{i}}(\mathcal{V}), \Phi_{\dot{i}'}(\mathcal{N}(\dot{I})))$ is a pair of choice functions such that $\Phi_{\dot{i}}(\mathcal{V})$ chooses a node \dot{I} among all vertex and $\Phi_{\dot{i}'}(\mathcal{N}(\dot{I}))$ chooses its **match** among its neighbors.

A matching criterion defines a method for picking an edge (\dot{I}, \dot{I}') which is later explored and added to the matching. The choice of \dot{I} fixes a priority order to the exploration of the graph since it dictates the first vertex seen at each iteration, it is then followed by the choice of \dot{I}' functionally completing the selection of the edge added to the matching.

Example II.1.1. • The **uniform choice** on \mathcal{U} is the choice function where every node is picked with the same probability:

$$\mathbb{P}_{\mathcal{U}}(\dot{I} = u) := \text{UNIF}_{\dot{i}}(\mathcal{U})(u) = \frac{\mathbb{1}_{\mathcal{U}}(u)}{|\mathcal{U}|} \quad (\text{II.2})$$

- The **maximal (degree) choice** selects uniformly amongst the maximal degree nodes:

$$\mathbb{P}_{\mathcal{U}}(\dot{I} = u) := \text{MAX}_{\dot{i}}(\mathcal{U})(u) = \text{UNIF}_{\dot{i}}(\mathcal{U}_{max})(u) = \frac{\mathbb{1}_{\mathcal{U}_{max}}(u)}{|\mathcal{U}_{max}|}. \quad (\text{II.3})$$

If we define the $d_{max}(\mathcal{U}) = \max_{u \in \mathcal{U}} d(u)$ as the maximal degree in \mathcal{U} , we write $\mathcal{U}_{max} := \{u \in \mathcal{U} \mid d(u) = d_{max}(\mathcal{U})\}$

- Likewise the **minimal (degree) choice** is a uniform choice amongst the minimal degree nodes:

$$\mathbb{P}_{\mathcal{U}}(\dot{I} = u) := \text{MIN}_{\dot{i}}(\mathcal{U})(u) = \text{UNIF}_{\dot{i}}(\mathcal{U}_{min})(u) = \frac{\mathbb{1}_{\mathcal{U}_{min}}(u)}{|\mathcal{U}_{min}|}. \quad (\text{II.4})$$

where $d_{min}(\mathcal{U}) = \min_{u \in \mathcal{U}} d(u)$ is the minimal degree in \mathcal{U} , we write $\mathcal{U}_{min} := \{u \in \mathcal{U} \mid d(u) = d_{min}(\mathcal{U})\}$

Remark II.1.1. *It is important to stress that we do not count a node among its own neighbors, even when there is a loop. This is to avoid adding loops to the matching. As we will later see, loops will occur in the configuration model and whenever they are marked, the algorithms would no longer produce a matching. Asymptotically, it will be irrelevant since the total number of loops will be negligible compared to the number of marked edges. Nevertheless we have to mention it, since loops appear with positive probability.*

Another way to circumvent the problem and produce a matching is to allow loops to be included but to later erase them whenever they are marked.

II.1.2 Local matching through exploration

We now construct the *local* matching algorithms later considered in this work. Fix a (non-oriented) graph $\mathring{\mathcal{G}} = (\mathring{\mathcal{V}}, \mathring{\mathcal{E}})$, where $\mathring{\mathcal{V}}$ is the set of nodes and $\mathring{\mathcal{E}} \subset \mathring{\mathcal{V}} \times \mathring{\mathcal{V}}$ is the set of edges and a matching criterion $\mathring{\Phi} := (\mathring{\Phi}_f(\mathring{\mathcal{V}}), \mathring{\Phi}_{\hat{I}}(\mathring{\mathcal{N}}(\hat{I}))$. Fix the size of $\mathring{\mathcal{V}}$ as $|\mathring{\mathcal{V}}| = n$ for $n \in \mathbb{N}$.

At any iteration t , are given two disjoint subgraphs of $\mathring{\mathcal{G}}$:

- $\mathring{\mathcal{G}}_t = (\mathring{\mathcal{U}}_t, \mathring{\mathcal{E}}_t)$ is the *undiscovered part of the graph* for our procedure at time t . The nodes of $\mathring{\mathcal{U}}_t$ are said to be *available* at t . Those are the nodes whose fate still is to be determined.
- $\mathring{\mathcal{G}}'_t = (\mathring{\mathcal{M}}_t, \mathring{\mathcal{G}}'_t)$ is the *matching* at time t . It is a $1 - factor$, a subgraph of $\mathring{\mathcal{G}}$ in which all nodes have only one neighbor with $\mathring{\mathcal{M}}_t$ containing all the matched nodes.

$\mathring{\mathcal{I}}_t$ is also defined as *isolated* nodes at t , that is nodes which will not be matched at all, because all of their neighbors have been matched. Thus,

$$\mathring{\mathcal{V}} = \mathring{\mathcal{U}}_t \cup \mathring{\mathcal{M}}_t \cup \mathring{\mathcal{I}}_t. \tag{II.5}$$

At first, the whole graph is available to be matched, meaning $\mathring{\mathcal{U}}_0 = \mathring{\mathcal{V}}$, $\mathring{\mathcal{M}}_0 = \emptyset$ and $\mathring{\mathcal{I}}_0 = \emptyset$. The matching process has yet to start, we also set $\mathring{\mathcal{E}}_0 = \mathring{\mathcal{E}}$ and $\mathring{\mathcal{E}}'_0 = \emptyset$, in a way that $\mathring{\mathcal{G}}_0 = \mathring{\mathcal{G}}$ and $\mathring{\mathcal{G}}'_0 = (\emptyset, \emptyset)$. The matching algorithm then proceeds by induction, repeating the following procedure:

At any time $t \in \llbracket 0, n - 1 \rrbracket$,

Step 1. A vertex \hat{I} is chosen following $\mathring{\Phi}_f(\mathring{\mathcal{U}}_t)$

Step 2. A match for \hat{I} is selected among the neighbors of \hat{I} , it is a choice \hat{I}' following $\mathring{\Phi}_{\hat{I}'}(\mathring{\mathcal{N}}(\hat{I}))$. Note that this procedure is *local*. It only depends on the current degree of the neighbors of \hat{I} . The next figure illustrates two criteria, the *match* is represented in red for GREEDY and MINRES, properly defined in Definition II.1.2.

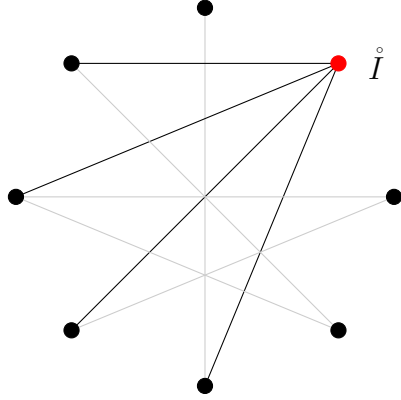
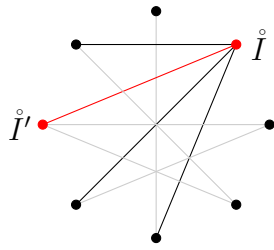
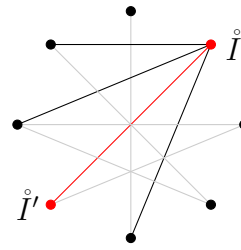


Figure II.1: Exploring a vertex



(a) GREEDY criterion

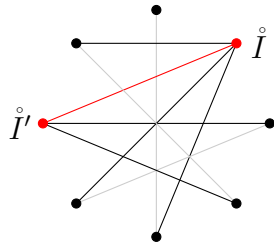


(b) MINRES criterion

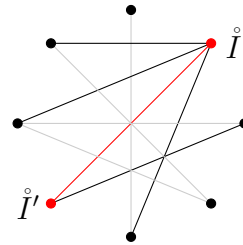
Figure II.2: Choice of an edge

Step 3. The matched nodes $\overset{\circ}{I}$ and $\overset{\circ}{I}'$ as well as their neighboring edges are explored. Those are removed from the unexplored graph (in grey). Specifically, we set

$$\begin{cases} \overset{\circ}{\mathcal{U}}_{t+1} = \overset{\circ}{\mathcal{U}}_t \setminus \{\overset{\circ}{I}, \overset{\circ}{I}'\}, & \overset{\circ}{\mathcal{G}}_{t+1} = \text{Induced subgraph of } \overset{\circ}{\mathcal{U}}_{t+1} \text{ in } \overset{\circ}{\mathcal{G}}_t, \\ \overset{\circ}{\mathcal{M}}_{t+1} = \overset{\circ}{\mathcal{M}}_t \cup \{\overset{\circ}{I}, \overset{\circ}{I}'\}, & \overset{\circ}{\mathcal{E}}'_{t+1} = \overset{\circ}{\mathcal{E}}'_t \cup \{(\overset{\circ}{I}, \overset{\circ}{I}')\}, \\ \overset{\circ}{\mathcal{I}}_{t+1} = \overset{\circ}{\mathcal{I}}_t \cup \{v \in \overset{\circ}{\mathcal{G}}_{t+1} \mid d_{t+1}(v) = 0\}. \end{cases}$$



(a) GREEDY: End of iteration



(b) MINRES: End of iteration

Figure II.3: Explored part of the graph after an iteration

Step 4. We set $t := t + 1$ and go back to step 1.

At the terminating point $t = n$, all the nodes are either matched or isolated, begetting

$$|\mathcal{I}_n| + |\mathcal{M}_n| = |\mathcal{V}|.$$

The *matching coverage* $\mathring{M}_{\Phi}^n(\mathring{\mathcal{G}})$, the proportion of initial nodes that ended up in the matching at the termination time n is expressed as a simple function of the sets we tracked during the matching dynamic.

$$\mathring{M}_{\Phi}^n(\mathring{\mathcal{G}}) = \frac{|\mathring{\mathcal{M}}_n|}{n} = 1 - \frac{|\mathring{\mathcal{I}}_n|}{n} \in [0, 1]. \quad (\text{II.6})$$

Such expression is crucial in the remainder of this work. It expresses the matching proportion as a function of the matching dynamic. Later on, concentration results will allow us to establish the convergence of the matching coverage given any criterion.

We can now introduce the main criteria that have been the focus of our simulations :

Definition II.1.2. • Φ is **greedy**, denoted $\Phi = \text{GREEDY}$, if both choices of \mathring{I} and \mathring{I}' are uniform. We can write

$$\text{GREEDY} = \left(\text{UNIF}_{\mathring{I}}(\mathring{\mathcal{V}}), \text{UNIF}_{\mathring{I}'}(\mathring{\mathcal{N}}(\mathring{I})) \right). \quad (\text{II.7})$$

- Φ is **minimal residual**, denoted $\Phi = \text{MINRES}$ if \mathring{I} is matched to its lower degree neighbor. The goal here is to obtain a better matching by decreasing the probability of augmenting paths. Since MINRES matches \mathring{I} to its lowest degree neighbor, it has a better chance of being matched to leaves², thus decreasing the chance of creating a situation where both leaves are left out of the matching, creating an augmenting path.

$$\text{MINRES} = \left(\text{UNIF}_{\mathring{I}}(\mathring{\mathcal{V}}), \text{MIN}_{\mathring{I}'}(\mathring{\mathcal{N}}(\mathring{I})) \right). \quad (\text{II.8})$$

²By leaf, it is meant a degree one node

Algorithm 1 General local matching algorithm for a choice criterion Φ

Require: Non-empty graph $\mathring{\mathcal{G}} = (\mathring{\mathcal{U}}, \mathring{\mathcal{E}})$
 Matching $\leftarrow \emptyset$;
for $t < n$ **do**
 Pick \mathring{I} following $\Phi_{\mathring{I}}(\mathring{\mathcal{V}})$;
 Discover its neighborhood $\mathring{\mathcal{N}}(\mathring{I})$ in $\mathring{\mathcal{U}}$;
 $\mathring{\mathcal{U}} \leftarrow \mathring{\mathcal{U}} \setminus \{\mathring{I}\}$;
 if $\mathring{\mathcal{N}}(\mathring{I})$ is empty **then**
 Do not do anything more
 else
 Pick \mathring{I}' following $\Phi_{\mathring{I}'}(\mathring{\mathcal{N}}(\mathring{I}))$;
 $\mathring{\mathcal{U}} \leftarrow \mathring{\mathcal{U}} \setminus \{\mathring{I}'\}$;
 Matching \leftarrow Matching $\cup (\mathring{I}, \mathring{I}')$;
 end if $t \leftarrow t + 1$;
end for

II.2 Local matching on the configuration model

The main feature of complex networks is their size. The larger they get, the more complicated it becomes to access their exact topology. This has prompted researchers to turn to other viewpoints. The local approach to large graphs is one such viewpoint: Even without full information, are there ways to describe the relationships between vertices?

One answer to this question is statistical. If we have access to a descriptor of the graph, for example its degree sequence, it can be used to pick uniformly a *similar* graph in some sense. If the degree distribution is known for example, one reasonable assumption to make is that typical vertex of the network and their neighborhoods should follow such distribution. From that assumption, concentration techniques can be used to obtain global properties on any procedure that is local. This led to random graph models being used as substitutes for complex networks.

The *configuration model* (Bollobás [1980]), also known as the *uniform pairing model*, is one of the most important theoretical models in the study of networks. Because, when conditioned on being simple, it represents a uniform drawing of a graph given its degree sequence, it is both realistic and simple enough to be used in a myriad of situations (Decreusefond et al. [2012] for the spread of an epidemic based on the SIR model or Bermolen et al. [2017] for a study of the jamming constant of general graphs). Moreover, the sequential nature of the construction of the configuration model makes the model at hand amenable to the fabrication of local algorithms (Appendix A.1), especially like we intended to do in the first section of the present chapter. This section uses this sequential

procedure to transpose the local matching algorithms to the configuration model. The algorithms are now built on a model with both layers of randomness.

Set ξ a probability measure on \mathbb{N} . Set n be a positive integer and $\mathfrak{d} := (d(1), \dots, d(n))$ a n -sample of the probability distribution ξ . We intend \mathfrak{d} to model the degree sequences of a given graph so it has to be *graphic*³.

Alternatively, we call $\mu := \sum_{i \leq n} \delta_{d(i)}$ the *measure valued representation* of \mathfrak{d} .

Thus, we let $\mathcal{V} = \{v_1, \dots, v_n\}$ be the set of nodes of the graph and for all $i \in \llbracket 1, n \rrbracket$, $d(i)$ being the *degree* of node v_i . A realization of the configuration model consists of giving every vertex v , $d(v)$ half-edges which are to be completed into edges through uniform pairing (see Appendix A.1).

II.2.1 Matching criteria on the CM

Definition II.2.1 (Matching criteria on degree sequences). *Let \mathfrak{d} a degree sequence, we can redefine matching criteria for degree sequences (or equivalently on their measure valued representation) as follows:*

- A **choice function** Φ on \mathfrak{d} is again a probability law on \mathfrak{d} . A random variable $K = d(I)$ is called *choice* following the function Φ_K when its law is Φ_K . Again, let $k \in \mathfrak{d}$ we write :

$$\mathbb{P}(K = k) := \Phi_K(\mathfrak{d})(k). \quad (\text{II.9})$$

abusing notations, for $\mu = \sum_{d \in \mathfrak{d}} \delta_d$, Φ can be written as a measure on $M_F(\mathbb{N})$:

$$\Phi_K(\mu) := \Phi_K(\mathfrak{d}) = \sum_k \Phi_K(\mathfrak{d})(k) \delta_k. \quad (\text{II.10})$$

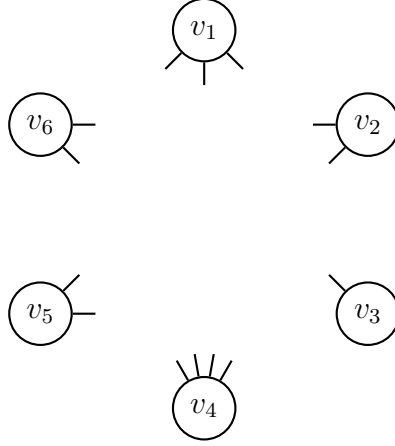
- Let $\mu' := \sum_{v \in \mathcal{N}(I)} \delta_{d(v)}$ the degree sequence of the neighbors of I . A **matching criterion** $\Phi := (\Phi_K(\mu), \Phi_{K'}(\mu'))$ is again a pair of choice functions such that a node $\Phi_K(\mu)$ chooses a vertices I of degree K among all vertices and $\Phi_{K'}(\mu')$ chooses the degree of its **match** among the degree of its neighbors.

This time, the matching criterion defines the degrees (K, K') of the edge (I, I') added to the matching.

Example II.2.1. • The **uniform choice** on \mathfrak{d} is the choice function where every degree is picked with the same probability. For $k \in \mathfrak{d}$, we use \mathfrak{d} and μ interchangeably:

³Graphicality of degree sequence has been extensively studied in the literature, we quickly present a characterization result and a drawing procedure in appendix A.1.4

Figure II.4: A running example: Initial state for $\mathfrak{d} = (3, 2, 1, 4, 2, 2)$



$$\text{UNIF}_K(\mu)(k) := \mathbb{P}_\mu(K_{\text{UNIF}} = k) = \frac{\langle \mu, \mathbb{1}_k \rangle}{\langle \mu, \mathbf{1} \rangle}. \quad (\text{II.11})$$

- The **maximal (degree) choice** selects uniformly amongst the maximal degrees:

$$\text{MAX}_K(\mu)(k) := \mathbb{P}_\mu(K_{\text{MAX}} = k) = \mathbb{1}_{d_{\text{max}}}(k). \quad (\text{II.12})$$

where $d_{\text{max}} := \max \mathfrak{d} = \max_i d_i$ with $d_i \in \mathfrak{d}$

- Again, the **minimal (degree) choice** selects uniformly amongst the minimal degrees:

$$\text{MIN}_K(\mu)(k) := \mathbb{P}_\mu(K_{\text{MIN}} = k) = \mathbb{1}_{d_{\text{min}}}(k). \quad (\text{II.13})$$

where $d_{\text{min}} := \min \mathfrak{d} = \min_i d_i$ with $d_i \in \mathfrak{d}$

II.2.2 The matching process on the CM

We can now set the trackers for the number of *available* stubs. Stubs (or half edges) are to be paired sequentially *uniformly* into edges, for all i we set $a_0(v_i) = d(i)$ and define the following set of initial *availabilities*:

$$\mathcal{A}_0 = \left\{ (v_1, a_0(v_1)), \dots, (v_n, a_0(v_n)) \right\}.$$

At first, all stubs are available and the \mathcal{G}_0 has no edges. We get $\mathcal{G}_0 := (\mathcal{V}, \emptyset)$. The matching is empty, and all the edges have to be discovered, hence we also set $\mathcal{M}_0 =$

\emptyset , $\mathcal{U}_0 = \mathcal{V}$ and \mathcal{G}'_0 is the empty graph, i.e. $\mathcal{G}'_0 = (\emptyset, \emptyset)$. To define the tempo of our construction, set t as an iteration count. In a similar fashion to the first section, we shall proceed by induction. At iteration t , we are given

- A multi-graph $\mathcal{G}_t = (\mathcal{V}, \mathcal{E}_t) = (\mathcal{M}_t \cup \mathcal{U}_t \cup \mathcal{I}_t, \mathcal{E}_t)$ representing the partially constructed connections between elements of \mathcal{V} , where we denote by:
 - \mathcal{M}_t the set of *matched* nodes at t which are nodes that are fully attached to the graph at t (no available stubs), and belong to the matching at t ;
 - \mathcal{U}_t the set of *unexplored* nodes at t , that is, nodes that do not belong to the matching at t , but can still be attached to it (it means they are unmatched but they have available stubs which can yet become edges added to the matching);
 - \mathcal{I}_t , the set of *isolated* nodes at t , that is, nodes that are already fully attached to the graph at t , but do not belong to the matching at t (unmatched and no more stubs to connect them to an eventual match).

By our very construction, all nodes of \mathcal{M}_t (if any) will have degree at least 1 in \mathcal{G}_t , and we let $\mathcal{U}_t = \{v_{i_1}, \dots, v_{i_p}\} \subset \mathcal{V}$. (We skip the dependance in t in the parameters v_{j_i} , for short.)

- A maximum *matching* $\mathcal{G}'_t = (\mathcal{M}_t, \mathcal{E}'_t)$ on the induced subgraph of \mathcal{M}_t in \mathcal{G}_t . In particular, \mathcal{E}'_t is a set of subsets of pairs of \mathcal{M}_t of the form $\{v_i, v_j\}$ for $v_i, v_j \in \mathcal{M}_t$ such that any element of \mathcal{M}_t appears in exactly one pair of \mathcal{E}'_t .
- Two sets of pairs, where all active nodes, nodes that have yet to be visited, are in

$$\mathcal{A}_t = \left\{ (v_{i_1}, a_t(v_{i_1})), \dots, (v_{i_p}, a_t(v_{i_p})) \right\}$$

and active nodes with stubs available are in

$$\mathcal{A}_t^* = \left\{ (v_{i_{1l}}, a_t(v_{i_{1l}})), \dots, (v_{i_{ml}}, a_t(v_{i_{ml}})) \right\}$$

with $\forall (v, a_t(v)) \in \mathcal{A}_t^*$, $a_t(v) > 0$. For any $l \in \llbracket 1, p \rrbracket$, $a_t(v_{i_l})$ is interpreted as the *availability* of node v_{i_l} at t , that is, the number of available stubs.

The next iteration is then triggered :

Step 1. An unexplored node I of positive availability K is selected following

$\Phi_K(\{a_t(v_{i_{11}}), \dots, a_t(v_{i_{1m}})\})$, this defines $(I, K) = (v_{i_q}, a_t(v_{i_q})) \in \mathcal{A}_t^*$ for some q . The uniform pairing procedure is used to complete the K stubs of I into edges

more specifically : We draw, **without repetition**, K elements uniformly at random among l_m bunches of elements of respective sizes $a_t(v_{i_l})$, $l \in \llbracket 1, l_m \rrbracket$ to construct the emanating edges of node I (recall that $l_m = |\mathcal{A}_t^*|$ is the number of nodes with available half edges). Then, we let $\mathcal{N}_t(I) := \{I'_1, \dots, I'_u\} \subset \{v_{i_{l_1}}, \dots, v_{i_{l_m}}\}$, where $u \leq K$ is the set of the u neighbors of I , i.e., the indexes of the bunches containing the chosen half-edges. Note that this operation may lead to parallel edges or loops whenever several elements of the same bunch of half-edges are chosen. In this case, $u < K$. However, is a well known fact that barring some restrictions on the initial degree sequence, the number of loops/multiple edges is $O(1)$ (see for example Angel et al. [2017]) so it does not affect the large graph matching coverage.

For all $l \in \llbracket 1, l_m \rrbracket$, we let $D_{I'_l}$ be the number of edges shared by I'_l with I , that is, the number of elements in the bunch I'_l chosen in the uniform pairing procedure.

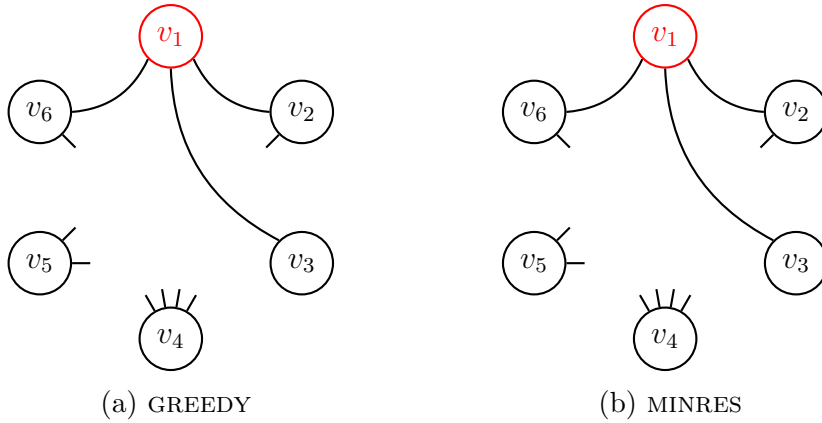


Figure II.5: Chosing $I = v_1$ and discovering its neighbors $\mathcal{N}_t(I) := \{v_2, v_3, v_6\}$. We also have that $K = 3$ and $a_t(v_2) = 2, a_t(v_3) = 1, a_t(v_6) = 2$.

Step 2. The match of I is chosen within the set $\{I'_1, \dots, I'_k\}$. Call it $I' = I'_m$, with availability K' , where K' is drawn following the law $\Phi_{K'}(\{a_t(I'_1), \dots, a_t(I'_u)\})$. Henceforth, both nodes I and I' together with the edge $\{I, I'\}$ can be added to the matching \mathcal{G}'_t , that is, we set

$$\mathcal{M}_{t+1} = \mathcal{M}_t \cup \{I, I'\} \quad \text{and} \quad \mathcal{E}'_{t+1} = \mathcal{E}_t \cup \{\{I, I'\}\}.$$

Step 3. Finally, we determine the neighboring edges on the I' side.

- If $K' > D_{I'}$, I' still has $K' - D_{I'}$ uncompleted half-edges. The procedure in **Step 1.** is reiterated. We draw at random the indexes I_1, \dots, I_ℓ of the neighbors of I' other than I using the same uniform pairing procedure. Namely, we draw $a_t(I') - D_{I'}$ elements uniformly at random among the $l_m - 1$ “bunches”

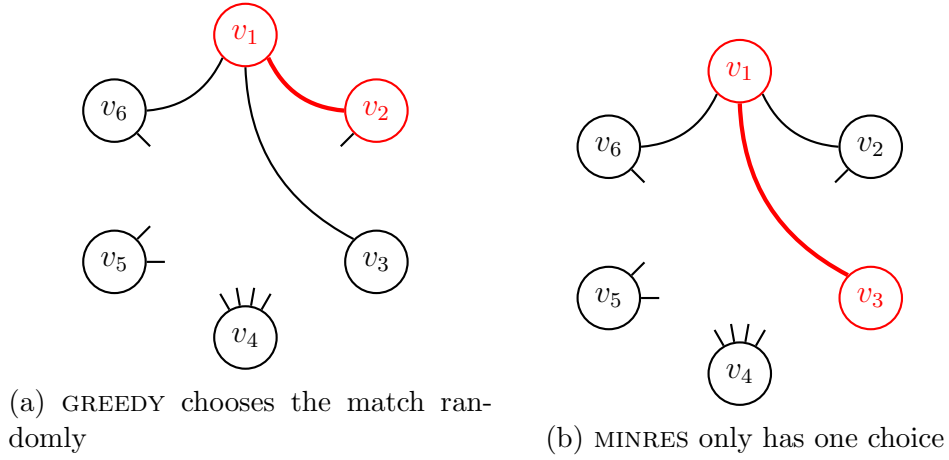


Figure II.6: Matching I with I'

(recall that l_m is the number of available bunches) of elements of respective sizes $a_t(v_{i_l}) - D_{i_l}$, $l \in \llbracket 1, l_m \rrbracket \setminus \{q\}$ (that is, all unexplored elements now that $I = v_{i_q}$ has been matched), to determine the other neighbors of I' .

We let $\{I_1, \dots, I_\ell\} \subset \{v_{i_1}, \dots, v_{i_m}\} \setminus \{i_q\}$ (for $\ell \leq a_t(I') - D_{I'}$) be the set of the ℓ neighbors of I , i.e., the indexes of the bunches containing the chosen half-edges. For all $l \in \llbracket 1, \ell \rrbracket$, we let F_{I_l} be the number of edges shared by I_l with I' , that is, the number of elements in the bunch I_l chosen in the uniform pairing procedure.

- If $K' = D_{I'}$ (which for example the case if $u = 1$), then I' has no more open half-edges to complete. In this case we do not do anything at this stage, and just set $\ell = 0$.

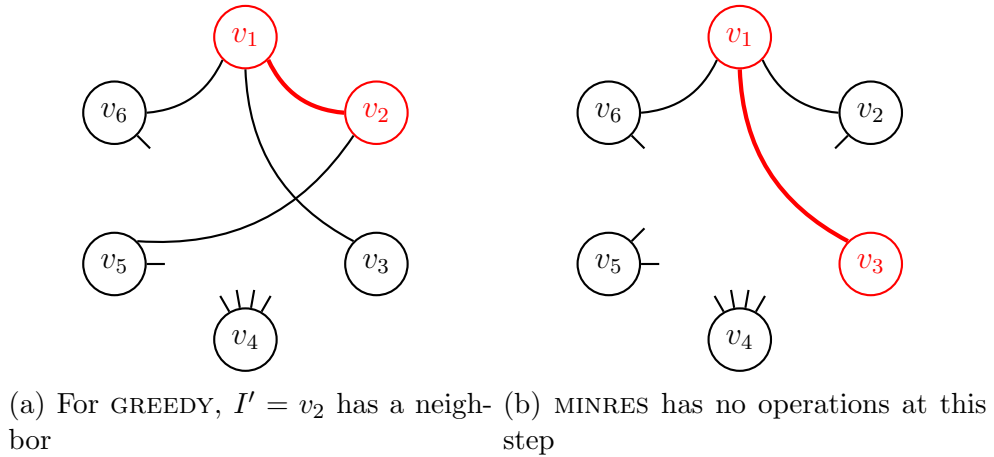


Figure II.7: Matching I with I'

After all those operations, in the illustrative example for GREEDY we have that $\mathcal{A}_{t+1} = \{(v_3, 0), (v_4, 3), (v_5, 1), (v_6, 1)\}$ and $\mathcal{A}_{t+1}^* = \{(v_4, 3), (v_5, 1), (v_6, 1)\}$. The other relevant sets are easily deduced as well.

In all cases, the neighboring edges of (I, I') are now complete, and we can summarize the operations

$$\mathcal{E}_{t+1} = \mathcal{E}_t \cup \{\{I, I'_1\}, \dots, \{I, I'_k\}\} \cup \{\{I_1, I'\}, \dots, \{I_\ell, I'\}\},$$

where the second set on the right-hand side above is understood as empty if $\ell = 0$. We also update the sets of availabilities \mathcal{A}_t by deleting the pairs corresponding to the newly matched nodes I and I' , and by updating the remaining number of stubs of the unmatched nodes connected to the two newly matched ones, if any. In other words, we set for the edges on both sides

$$\left\{ \begin{array}{l} \mathcal{A}_{t+1} = \mathcal{A}_t \cup \{(I_1, a_t(I_1) - F_{I_1}), \dots, (I_\ell, a_t(I_\ell) - F_{I_\ell})\} \\ \quad \setminus \{(I, a_t(I)), (I_1, a_t(I_1)), \dots, (I_\ell, a_t(I_\ell))\}; \\ \mathcal{A}_{t+1} = \mathcal{A}_{t+1} \cup \{(I'_1, a_t(I'_1) - D_1), \dots, (I'_u, a_t(I'_u) - D_u)\} \\ \quad \setminus \{(I', a_t(I')), (I'_1, a_t(I'_1)), \dots, (I'_u, a_t(I'_u))\}. \end{array} \right. \quad (\text{II.14})$$

The procedure terminates at most at time n , when \mathcal{U}_n becomes empty. One last operation is needed to complete the process. All the active nodes become isolated :

$$\begin{aligned} \mathcal{I}_n &= \mathcal{I}_n \cup \mathcal{A}_n, \\ \mathcal{A}_n &= \emptyset. \end{aligned} \quad (\text{II.15})$$

At that time, we end up with a multi-graph $\mathcal{G} := \mathcal{G}_n = (\mathcal{V}, \mathcal{E}_n)$, since all stubs have been completed. Moreover all the nodes are either matched or isolated.

The *matching coverage* is reexpressed the same way:

$$\mathbf{M}_{\Phi}^n(\mathring{\mathcal{G}}) = \frac{|\mathcal{M}_n|}{n} = 1 - \frac{|\mathcal{I}_n|}{n} \in [0, 1]. \quad (\text{II.16})$$

II.3 Simulations

To motivate the remainder of our work we provide some early simulation results. (Aoudi et al. [2022]) gives a more detailed account of the technical aspects. We chose to focus on bipartite graphs since it is the most relevant framework for practical uses. We test the

empirical convergence of the matching coverage under both GREEDY and MINRES. This section serves also as a prelude for testing the robustness of our method. For the entirety of this section, the work presented is restricted to the bipartite configuration model. Hence, the empirical degree sequence is divided accordingly in $\{+\}$ and $\{-\}$ sides, representing the bipartition. We write

$$\mathfrak{d} = (\mathfrak{d}^+, \mathfrak{d}^-), \quad (\text{II.17})$$

where \mathfrak{d} is the initial degree sequence. The criteria are also lined accordingly: Φ_I is restricted to \mathfrak{d}^+ .

II.3.1 Convergence of the Matching Size : a case study on 3-regular graphs

In this section, we use 3-regular graphs to test the convergence of $\mathbf{M}_{\Phi}^n(G)/n$, for G a 3-regular graph and the initial degree sequence $\mathfrak{d}^+ = \mathfrak{d}^- = 3^{\otimes n}$, for n the number of vertices on both sides.

To illustrate the convergence of $\mathbf{M}_{\Phi}^n(G)/n$, we proceed as follows: for each value of n from 10 to 10 000, we use the bipartite configuration model to draw a 3-regular graph as a realization of the CM. For various graph sizes, we ran $N = 50$ iterations of the previous procedure.

As the graph gets larger, the fluctuations get smaller and smaller, heuristically showing the **concentration** of the matching coverage $\mathbf{M}_{\Phi}^n(G)/n$ for both algorithms. Means and standard deviations of the corresponding statistical distributions for $\mathbf{M}_{\Phi}^n(G)/n$, are given in the table (II.1).

	Graph Size n	200	500	1000	3000	5000
$\mathbf{M}_{\text{GREEDY}}^n(G)/n$	Mean	0.8904	0.8916	0.8911	0.8897	0.8898
	Std Dev	0.0198	0.0109	0.009	0.0041	0.00311
$\mathbf{M}_{\text{MINRES}}^n(G)/n$	Mean	0.9356	0.9365	0.9396	0.9378	0.9385
	Std Dev	0.0148	0.0096	0.0052	0.0040	0.0025

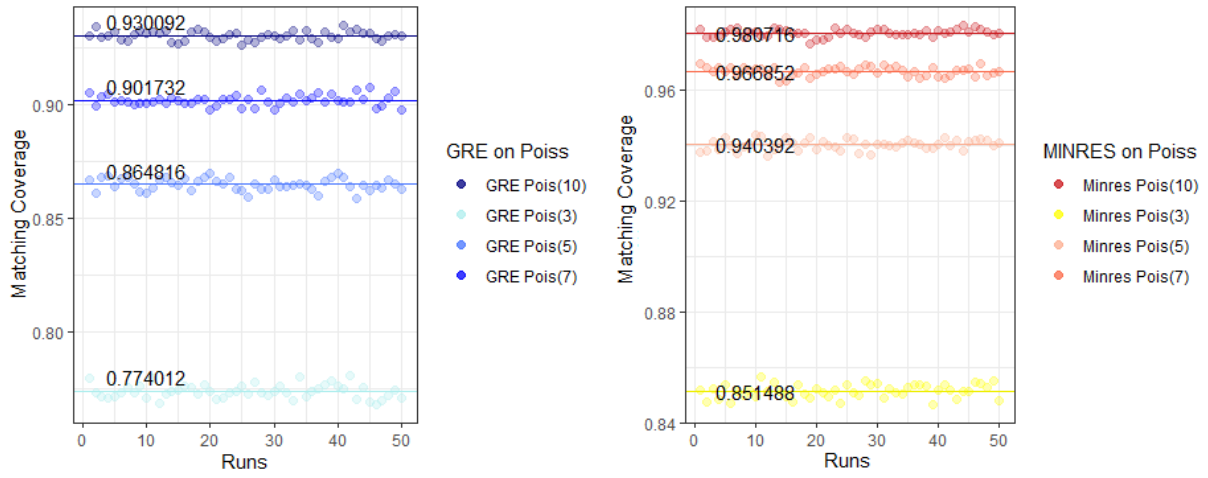
Table II.1: Recap. of Average Matching Size

Table II.1 hints at the convergence results we later establish in Chapter 3 and 4. The shrinking of the standard deviations confirm the heuristic convergence to a deterministic value. It also stresses on the better performances of MINRES with respect to GREEDY for this particular degree distribution. In the next part, we develop this comparison for a larger range of degree distributions.

II.3.2 An Array of Degree Distributions

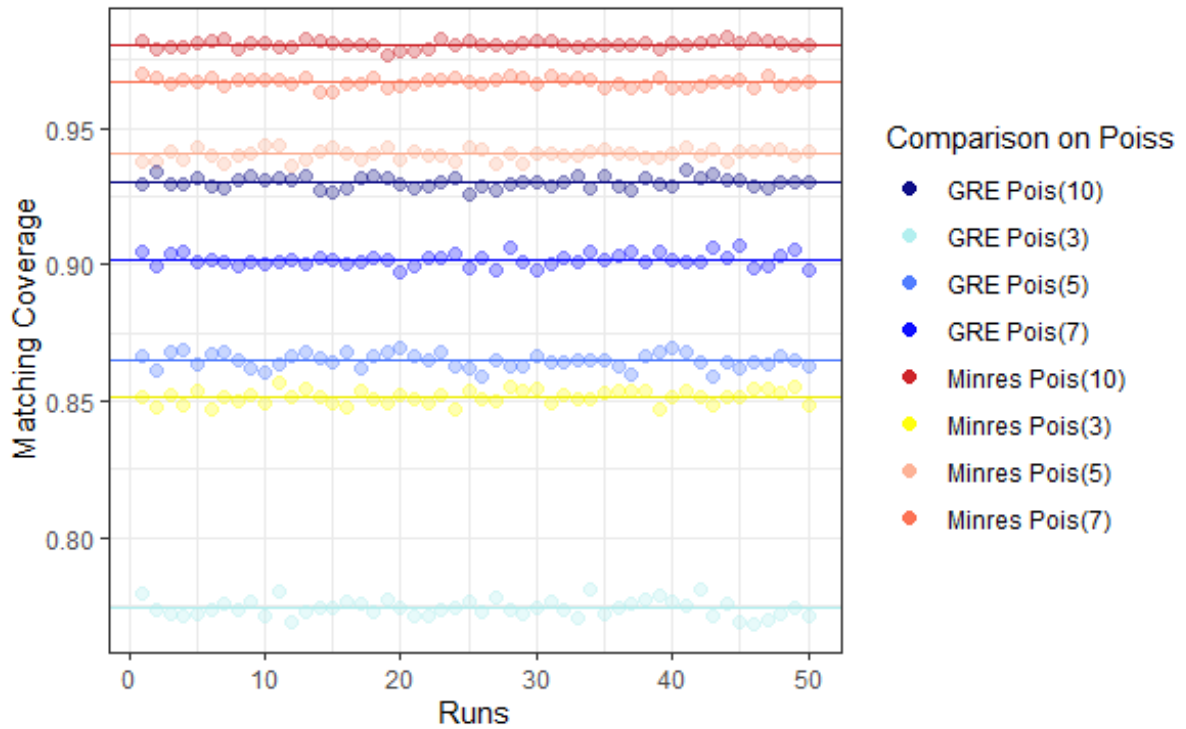
Here, we study the evolution of the performances of the two algorithms GREEDY and MINRES under varying parameters of the configuration model. We *empirically* identify the convergence along various degree distributions and various parameters. Our procedure is the following: for each degree distribution, we use still the bipartite CM to generate a large graph (of size $n = 10^4$ nodes), in which the degrees of the nodes form a n -sample of the prescribed degree distribution, after testing the graphicality of the latter degree distribution (i.e., the feasibility of the generation of the graph). For each distribution, again we ran 50 iterations of both algorithms, following the first construction of Section II.1. Each dot on the figure represents the matching coverage of a single run.

Poisson distributions. We first address the class of Poisson distributions, which are well known to be the asymptotic degree distributions of Erdős-Rényi graphs. Distributions of the matching coverage for the GREEDY (respectively, MINRES) algorithm are given in Figure II.8a (resp., Figure II.8b), for Poisson distributions of various parameters. To achieve the comparison of the two algorithms, these distributions are gathered in Figure II.8c.



(a) GREEDY on Poisson Laws

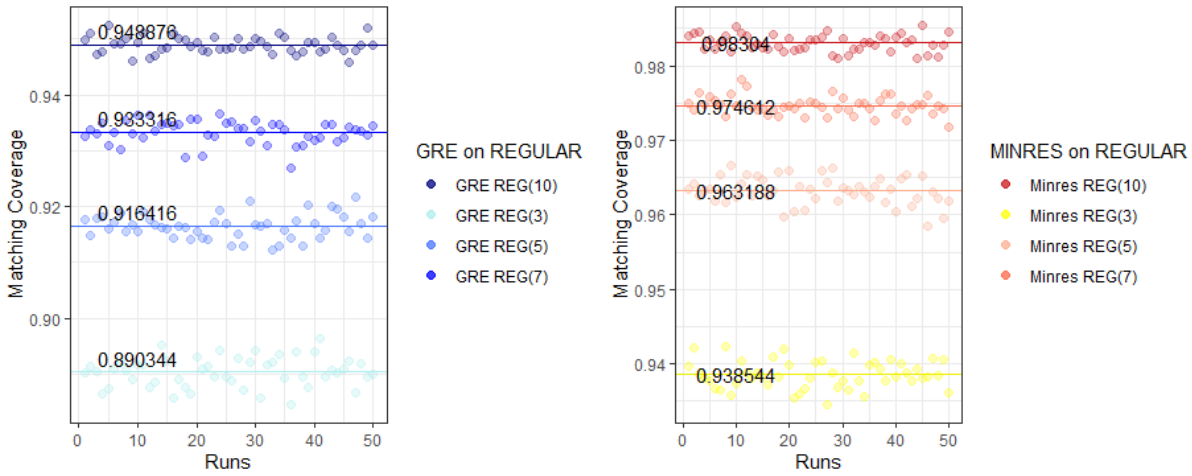
(b) MINRES on Poisson Laws



(c) Comparison on Poisson Law

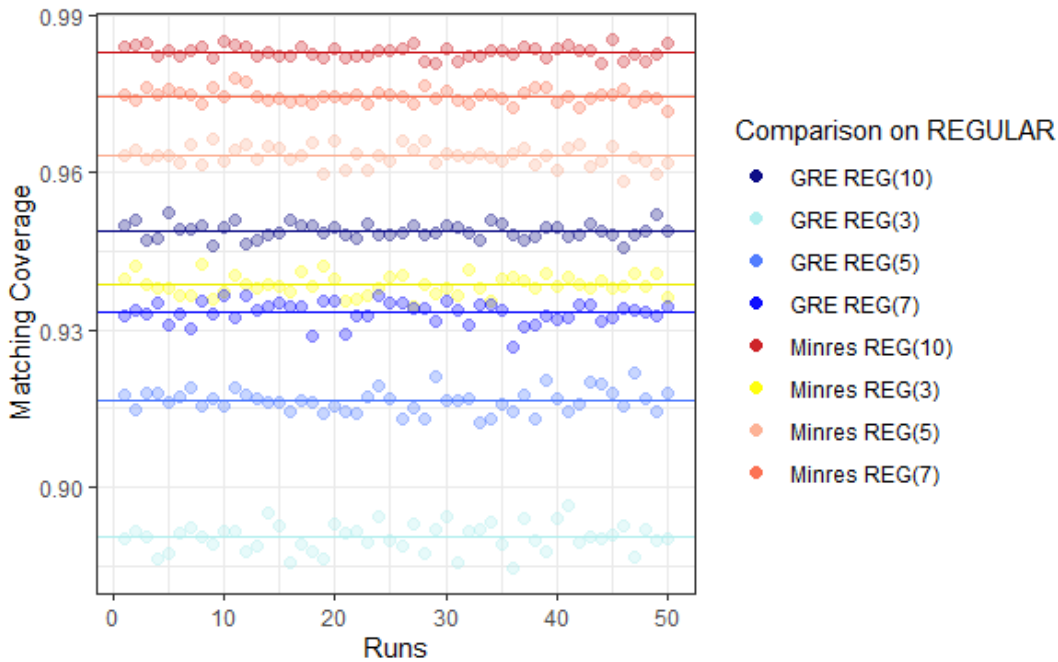
Figure II.8: Performance on Poisson Degree Distribution

Regular bipartite graphs. We now address various degree distributions that correspond to regular bipartite graphs: each node has the same degree, in other words we have $d^+ = d^- = p^{\otimes n}$, for some $p \in \mathbb{N}^*$, thereby generalizing the study of Sub-section II.3.1. Notice that such bipartite graphs always admit a perfect matching. The distributions of matching coverage are given in Figures II.9a, II.9b and II.9c.



(a) GREEDY on Regular Degree Distr.

(b) MINRES on Regular Degree Distr.



(c) Comparison on Regular Distribution

Figure II.9: Performance on Regular Degree Distr.

These results illustrate perfectly the growth of the performance of both algorithms as the average degree grows. For each designated distribution, we can also confirm that the matching coverage of MINRES is consistently larger than that of GREEDY.

Comparing the sub-figures of Figure II.8 with the corresponding sub-figures of Figure II.9 for the same average degree, we make the two following observations: First, both algorithms consistently perform better on regular graphs than on graphs with Poisson degree distributions. Regular graphs can be written as the disjoint union of perfect matchings so their structure might drastically increase the chance of building one. We

also conjecture that this phenomena is due to the variance of Poisson degrees in the first panel: by restricting choices, the optimal partners for certain nodes (which we are trying to reach) might get blocked, while regular distributions provide more latitude to chose a match without missing an optimal one ⁴. Second, for both algorithms, the distribution of matching coverages has higher variance on regular graphs than on graphs having Poisson degrees. It seems that a uniform initial degree distribution provides more opportunities to deflect from *typical* runs, while the variance of degrees often restricts the choices, creating a disparity between nodes.

II.3.3 The Influence of Topology and Another Approach To Online Bipartite Matching: Optimal Algorithm by Karp, Vazirani and Vazirani

In [Karp et al., 1990], Karp, Vazirani and Vazirani present a different approach for online bipartite matchings on graphs. The authors define online algorithms as a way of picking the matches of “girls” (i.e., nodes on the ‘+’ side) that arrive one by one, based only on the identity of their neighbors on the ‘boys’ side (i.e., the ‘-’ side). On the adjacency matrix of the graph, it translates to the fact that the columns are revealed one by one and the match is processed upon knowing the information of the current column. Our approach based on local algorithms has slight differences and actually uses more information. Indeed, we also consider the degree of each neighbor, of the incoming ‘girl’, and the neighbors of its match, of her ‘boy’ neighbor.

Second, the performance metric considered in [Karp et al., 1990] is the so-called *adversary approach*. It allows to get a lower bound on the performance of online algorithms, as defined above. Namely, Theorem 2 in [Karp et al., 1990] states that in that context, the matching coverage is at least $1 - 1/e - o(n)/n$ on average. This average happens to be the expected matching coverage attained by the RANDOM algorithm defined therein (which is roughly equivalent to our GREEDY). In our paper, a different metric is used: we establish convergence to a deterministic value of the average matching ratio, rather than a lower bound.

Therefore, it is clear that our framework differs from that of [Karp et al., 1990]. However, to gain some insights on how much our algorithms based on the degree distribution stand up against their counterparts on real graphs, we conducted the following study: we let G be a randomly generated graph with 5000 nodes, from an upper triangular adjacency matrix that is specified as follows: all diagonal elements are 1, thereby insuring the

⁴Konig’s theorem states that regular graphs can in fact be viewed as a union of p perfect matchings, thus the margin of error can be approximately known at each step

existence of a perfect matching, and all upper elements are Bernoulli(p) random variables, where p is so that the graph has average degree, say, 5. Such graphs provide the worst case scenarios for the framework in Karp et al. [1990]. Then, the exploration algorithms are run on G (Section II.1 of both GREEDY and MINRES). In parallel, we extract $(\mathbf{d}^+, \mathbf{d}^-)$ the degree distributions of the graph under consideration, on which we run both GREEDY and MINRES (the CM version of Section II.2), thereby constructing another (multi-)graph having the same degree distribution. Our results can be summarized in Figure II.10.

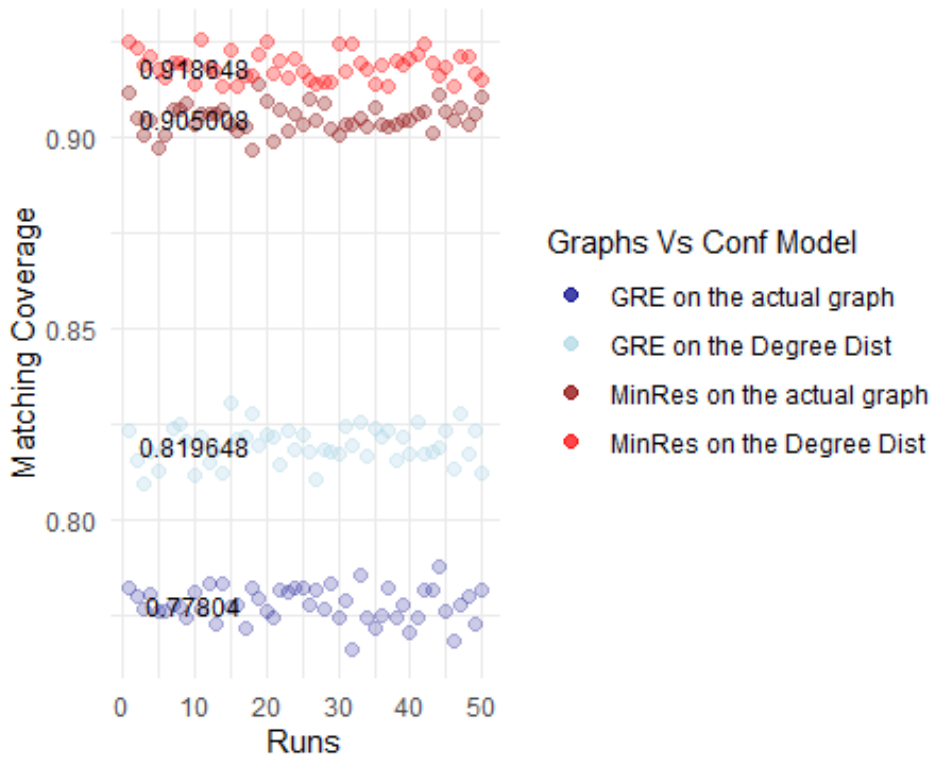


Figure II.10: Exploration Vs Conf Model

Figure II.10 shows that matching algorithms that are jointly constructed with the CM achieve a better matching coverage than on this given realization. In other words, matching algorithms typically perform badly on this particular graph, with respect to a graph that is obtained as a uniform draw amongst graphs having the same degree distribution, unsurprisingly hinting at the influence of the graph topology on the considered matching algorithms. Simulations seem to indicate that this influence is more so enhanced in GREEDY. Finally, we also observe that MINRES produces a better matching coverage than GREEDY anyway.

Notes II.3.1 (Choice of the Configuration Model). *In the introduction, we mentioned that random graphs were analogous to statistical models for complex networks. And statistical*

models are picked accordingly to the populations under study. Hence, the choice of the underlying random graph model should also depend on the network at hand. This means our approach has more legitimacy on networks that are well described by the CM.

We chose the CM mostly because of its local description of the neighborhood of nodes which ends up being convenient for building local algorithms. Our algorithms could be transposed to other asymptotic graph models with local descriptors, although it would perhaps require a different set of mathematical tools to study their performances.

Another observation to be made is that the matching problem we are studying is defined on simple graphs. This restricts the use cases for our model. For the CM to produce and asymptotically be 'close' to a simple graph, it is a well known fact that some moments restrictions have to be imposed on the degree distribution van der Hofstad [2016] which we will explore later. This is the reason why most of the distributions we picked for simulations admit at least moments of order 4.

II.3.4 Notes

In the next chapters, we will develop the procedure used to approximate and predict the matching coverage of local matching algorithms on graphs. Representing the matching algorithms as a dynamic on the degree distribution will allow us to predict the results of the considered algorithms with remarkable accuracy. In particular, the next chapter will delve into such representation. Notice that this results in a dramatic reduction of the problem complexity: as long as one is interested in the matching coverage of the algorithm under consideration, one only needs to keep track of the unexplored part of the graph.

As our simulations indicate, two natural and interesting problems already arise from this chapter: The first (and probably the harder) one is to quantify the influence of the graph topology on the average matching coverage for a given algorithm. How much does the structure of the network influence the performance of algorithms?

The second concerns the potential optimality of MINRES as a criterion. Among all the criteria we addressed, MINRES always seemed to give the better matching coverages *in average*. An interested follow up to this work could be the establishment of a (stochastic?) order on all the matching criteria.

Chapter III

Local Matching Algorithms as Markov processes

Chapter II described the algorithms we consider as local. Chapter III extends the representation of the latter algorithms as operations on empirical measures. Each iteration is written as a sequence of calculations on the current degree distribution of the graph. Random dynamics on spaces of measures are an immensely prominent wing in the study of stochastic processes. Being the prime representation for point processes, their uses range from the study of population processes in biology (Fournier and Méléard [2004]; Champagnat et al. [2008]) to statistical physics and interacting particle systems (Liggett [2012]). Likewise, a series of important results comes from the queuing literature, noticeably with the concept of fluid limits and its applications to several related problems (Gromoll et al. [2002]; Decreusefond and Moyal [2008b,a]). We intend to follow the works of Bermolen et al. [2017] and use the vast variety of tools for measure spaces on the particular study of local algorithms as we previously defined them.

Chapter III is comprised of two sections. The first section is an examination of matching algorithms as measure valued processes. The intent is to find a representation of the local algorithms as defined in the last chapter. It is as much an alternative representation as it is a simplification. We can then rewrite the matching coverage as a continuous function of the process describing the algorithm and we can proceed to the study of its properties. The section starts by proving a theorem that bridges the first two sections of chapter II as hinted by the simulations. An equivalence of sorts is established between the behavior of criteria on the finite given graphs and their behavior on the configuration model. The tendencies a given criteria exhibits on defined graphs stays the same when extended to the configuration model. The second section then concludes the chapter with the establishment of a proper scaling of the measure valued representation. It can be seen as a normalization, used to study and predict the behavior of matching criteria based

algorithms when the graphs get arbitrarily large. In the section, we mostly define the normalized version of the measure valued processes in prevision of the next chapter, where we prove its convergence to a deterministic function.

III.1 Notations and Hypotheses

For the rest of this work, the superscript n will be used to represent the size of the empirical distributions/measures. As the number of vertices, it will also be the scaling parameter of our study. Recall that, $\mathcal{M}_F(\mathbb{N})$ is the space of finite valued measures on \mathbb{N} . We shorten it to \mathcal{M}_F . Likewise, we define \mathcal{M}_F^n as the subset of \mathcal{M}_F whose elements are counting measures with total mass $\leq n$. \mathcal{M}_n is then the space or normalised space associated to \mathcal{M}_F^n . We write $\mathcal{M}_n := \frac{1}{n}\mathcal{M}_F^n = \{\frac{1}{n}\mu \mid \mu \in \mathcal{M}_F^n\}$. Finally, \mathcal{M} is the space of \mathbb{N} valued measures with total mass ≤ 1 . Usually, those measure spaces are endowed with their weak topology (resp. *vague* topology):

$$\mu^n \xrightarrow{\mathcal{M}_F} \mu \iff \langle \mu^n, f \rangle := \int f d\mu^n \rightarrow \int f d\mu := \langle \mu, f \rangle \quad (\text{III.1})$$

for all bounded continuous function $f \in \mathcal{C}_b(\mathbb{N})$ (resp. compactly supported functions $f \in \mathcal{C}_c(\mathbb{N})$). Again $\mathcal{C}_b(\mathbb{N})$ (resp. $f \in \mathcal{C}_c(\mathbb{N})$) is endowed with its usual topology, $\|f\| = \sup_{x \in \mathbb{N}} |f(x)|$.

Moreover, we reintroduce some functions and notations of interest :

- $\chi(x) = x$ the **identity function**
- $\nabla f(x) = f(x) - f(x-1)$ the **discrete gradient**
- For two real valued functions f and g , we say that $f(n) = o(g(n))$ if $\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = 0$.
- Likewise $f(n) = O(g(n))$ if $\limsup_{n \rightarrow \infty} \left| \frac{f(n)}{g(n)} \right| < C$ for some positive constant C .

Finally, unless the criterion is specified μ_t^n , $n \in \mathbb{N}$, $t \geq 0$ designates the process affiliated with a generic matching criterion Φ .

III.1.1 Main Hypotheses

Let $\{\bar{\mu}^n; n \in \mathbb{N}^*\}$ be a sequence of $\mathbb{D}[(0, T), \mathcal{M}_F(\mathbb{N})]$ valued processes, let ν in \mathcal{M}_F , $\langle \nu, \chi^p \rangle < \infty$ and $p \leq 3$. We say that $\{\bar{\mu}^n; n \in \mathbb{N}^*\}$ satisfies **Ho** if, for some $\nu \in \mathcal{M}$ the following conditions on the initial states hold :

Hypothese $\mathfrak{H}0$

1. Weak convergence of the initial degree distributions:

$$\langle \bar{\mu}_0^n, f \rangle \xrightarrow{n} \langle \nu, f \rangle < \infty, \quad f \in \mathcal{C}_b(\mathbb{N}) \quad (\text{III.2})$$

2. Convergence of the first three initial moments:

$$\langle \bar{\mu}_0^n, \chi^p \rangle \xrightarrow{n} \langle \nu, \chi^p \rangle < \infty \quad (\text{III.3})$$

for $1 \leq p \leq 3$.

As we will later see, these conditions guarantee a bound on the moments of the empirical degree distribution throughout all the dynamic. However, their interest is not limited to the technical aspect. The average number of second neighbors¹ of a typical node in the configuration model $CM(\nu)$ is shown to be (Newman [2018]; van der Hofstad [2016]; Angel et al. [2017]) $\langle \nu, \chi^2 \rangle - \langle \nu, \chi \rangle$ and each one of those second neighbors have a degree that again follows the size-biased degree distribution. $\mathfrak{H}0$ warrants that this number of neighbors is (almost surely) finite, thus one can define a dynamic depending on second neighbors of a given node. And since second neighbors characterize the neighboring edges of any given edge, this allows us to transpose the construction of local algorithms asymptotically to the configuration model. Moreover, it is a way to ensure that loops and parallel edges do not disturb our construction since their number becomes $O(1)$ (Angel et al. [2017]) meaning their influence on the *proportion of matched nodes* vanishes as n goes to infinity.

Notice that $\mathfrak{H}0$ is an hypothesis on the initial degree distribution and it does not impact the dynamic yet. It does however affect the properties of the CM.

Notes III.1.1. *Those restrictions on the moments of the degree distribution are standard for the configuration model. The whole premise of our work rests on the fact that realizations of our models are close to that of simple graphs (in the sense that multiple edges are asymptotically negligible). On the configuration model, this is guaranteed by a **second moment** on the normalized degree sequence Janson [2013]. As will be seen later, this is also the reason why we have a crucial $o(\sqrt{n})$ estimate for the degree of the first node selected I .*

Now I' is chosen amongst the neighbors of I who all follow a distribution that is very close to a size biased version of the initial degree distribution. But, for the size biased version of a measure to admit a variance, the original version needs to have at least 3

¹Second neighbors are simply neighbors of neighbors

moments, justifying our hypothesis. A finite variance on our dynamic will prove to be essential in establishing most of the useful properties in the present chapter and even later (tightness, maximum degree control and cancellation of the martingale part for example).

In fact, the moment requirement will have to be increased to 3.5 to prove the main theorem for GREEDY.

III.2 Discrete time Markov chain

We can properly instigate the study of the process induced by the matching dynamic on degree sequences. Unless stated otherwise, in this section the graph size is considered constant and the iteration count is the variable.

III.2.1 Measure valued representation and Markov Chain dynamic

Let $\mu_0^n : \sum_i \delta_{d(i)} := \sum_i \delta_{a_0(v_i)}$ be an initial degree sequence and $\mu_k^n, k \in \llbracket 0, n \rrbracket$ be the measure that tracks the availability of nodes over the course of a matching algorithm as set in chapter II (II.2). Going forward, the performance of the algorithms is tracked using measures. Specifically, let Φ be a matching criterion. First, we will initiate a comparison between the local algorithms on defined graphs to their extension when the structural layer of randomness is added by the configuration model. For all $0 \leq k \leq n$:

- We let $\mathring{\mu}_k$ be the empirical degree distribution of all unexplored nodes at k in the remaining graph $\mathring{\mathcal{G}}_k$ associated to the construction in Section II.1, that is

$$\mathring{\mu}_k = \sum_{v \in \mathring{\mathcal{U}}_k} \delta_{d_k(v)}, \quad (\text{III.4})$$

where $d_k(v)$ is the degree of v in $\mathring{\mathcal{G}}_k$. Recall that the $\mathring{\circ}$ notation designates given deterministic graphs.

- Likewise μ_k will be the analogous empirical distribution representing the availabilities of all unexplored nodes at k when the configuration model is used (Section II.2):

$$\mu_k = \sum_{v \in \mathcal{U}_k} \delta_{a_k(v)}. \quad (\text{III.5})$$

The following theorem links the two constructions. It is a formalization of the fact that conditioning by a given structure first before applying the algorithms does not influence

the final law of the dynamic. The matching criteria behave "as expected" when ran on the CM.

Theorem III.2.1. *Let n be a positive integer, $\mathring{\mathcal{G}}$ is a graph of size n , and let \mathfrak{d} its degree sequence. Let $\dot{\mu}_t$ be the measure-valued stochastic process defined by (III.4) with initial value*

$$\dot{\mu}_0 = \sum_{i=1}^n \delta_{d(i)}.$$

Let \mathcal{G} be the resulting multigraph of the second construction. Then for any $t \in \llbracket 0, n \rrbracket$ and any measure $\nu \in \mathcal{M}_{\mathbb{F}}^n(\mathbb{N})$. $\mathcal{M}_{\mathbb{F}}^n(\mathbb{N})$ is the space of finite point measures such that $\nu(i) \in \mathbb{N}$, $\sum_{i \leq n} \nu(i) \leq n$ and $\nu(i) = 0$ if $i \geq n$. We get that

$$\mathbb{P} \left[\mu_t = \nu \mid \mathcal{G} = \mathring{\mathcal{G}} \right] = \mathbb{P} [\dot{\mu}_t = \nu].$$

Proof. Suppose that $\mathcal{G} = \mathring{\mathcal{G}}$, that is, the final result of the second construction produces the graph $\mathring{\mathcal{G}}$. We index the nodes of \mathcal{V} consistently in the two constructions. Take $\dot{\mu}_0 = \mu_0$, the result is obtained by induction on t .

Suppose that, at some time $t \in \llbracket 0, n - 1 \rrbracket$ we have

$$\mathring{\mathcal{G}}_t = \mathcal{U}_t \quad \text{and} \quad a_t(v) = d_t(v) \quad \text{for all } v \text{ in } \dot{\mathcal{U}}_t = \mathcal{U}_t. \quad (\text{III.6})$$

We exhibit a coupling such that (III.6) holds also at time $t + 1$, so that the empirical measures stay the same. Notice first that as $\mathring{\mathcal{G}}_t = \mathcal{G}_t$, at Step $\dot{1}$ and 1 respectively, we can set a common realization of the draw on $\llbracket 1, |\mathcal{U}_t| \rrbracket$, leading to the same values for \dot{I} and I . As described in the first chapter, the reader is reminded that D_{I_l} is the number of edges going from I' to I_l , and that $F_{I'_j}$ is likewise the number of edges going from I to I'_j .

Since $a_t(I) > 0$, at step 1), as we have to produce the same graph, the uniform pairing procedure leads to the same set of neighbors for I as in $\mathring{\mathcal{G}}$, namely $I'_1 = \dot{I}'_1, \dots, I'_u = \dot{I}'_u$ because the same indexes can be picked in the two constructions. Then, as $d_t(\dot{I}) = a_t(I) > 0$, at steps $\dot{2}$) and 2) again we can align the draws so that $\dot{I}' = I'$ in $\dot{\mathcal{U}}_t = \mathcal{U}_t$, using the matching criterion Φ . Finally, as $G = \tilde{G}$ the uniform pairing procedure leads again to the same set of neighbors for I' too as in \tilde{G} , namely $I_1 = \dot{I}_1, \dots, I_\ell = \dot{I}_\ell$, again using the fact that we can use the same indexing. We then follow up with

$$\dot{\mathcal{U}}_{t+1} = \dot{\mathcal{U}}_t \setminus \{\dot{I}, \dot{I}'\} = \mathcal{U}_t \setminus \{I, I'\} = \mathcal{U}_{t+1}.$$

G does not have multiple edges, so $u = d_t(\dot{I}) = K$ and $\ell = d_t(\dot{I}') = K'$. Moreover, for

any $l \in \llbracket 1, k \rrbracket \setminus \{m\}$, at step 1a) we obtain that

$$a_{t+1}(I'_l) = a_t(I'_l) - D_{I_l} = a_t(I'_l) - 1 = d_t(I'_l) - 1 = a_{t+1}(I'_l),$$

And applying the same argument to the neighbors of I' , $l \in \llbracket 1, \ell \rrbracket \setminus \{\sigma(m)\}$,

$$a_{t+1}(I_l) = a_t(I_l) - F_{I_l} = a_t(I_l) - 1 = d_t(I_l) - 1 = a_{t+1}(I_l).$$

Last, for any $v \in \mathcal{U}_{t+1} \setminus \{I_1, \dots, I_\ell, I'_1, \dots, I'_k\}$ we get that

$$a_{t+1}(v) = a_t(v) = d_t(v) = d_{t+1}(v).$$

In all cases, expression (III.6) holds at step $t + 1$.

Thus, we can conclude that (III.6) holds in fact for all $t \in \llbracket 0, n \rrbracket$. (III.4) and (III.5) imply in particular that $\hat{\mu}_t = \mu_t$ for all $t \in \llbracket 0, n \rrbracket$, concluding the proof. \blacksquare

This theorem establishes that whenever the configuration model produces a given graph, the local criterion based matching algorithm simply behaves like an exploration algorithm on the given graph (as given is the first section of chapter II).

We can finally tend to the study of the dynamic induced on the degree distribution.

Proposition III.2.1. *Fix $n \in \mathbb{N}, n > 0$, the matching algorithms on the CM induce a \mathcal{M}_F^n valued **Markov chain**, $(\mu_k^n)_{k \geq 0}$ with transition kernel*

$$\mathfrak{L}^n F(\mu) := \mathbb{E}_\mu \left(F(\mu + \vartheta(\mu)) - F(\mu) \right), \quad (\text{III.7})$$

for $\mu \in \mathcal{M}_F^n$ and F a continuous bounded function $\mathcal{M}_F^n \rightarrow \mathbb{R}$. $\vartheta(\mu_k^n)$ is defined as

$$\vartheta(\mu_k^n) = - \left(\delta_K + \delta_{K'} + \sum_{I'_l \in \mathcal{N}(I)} \left(\delta_{a_k(I'_l)} - \delta_{a_k(I'_l) - D_{I'_l}} \right) + \sum_{I_l \in \mathcal{N}(I')} \left(\delta_{a_k(I_l)} - \delta_{a_k(I_l) - F_{I_l}} \right) \right). \quad (\text{III.8})$$

In particular, for $\Pi_f(\mu) = \langle \mu, f \rangle$, $f \in \mathcal{C}_b(\mathbb{N})$, equation (III.7) is expressed as

$$\mathfrak{L}^n \Pi_f(\mu_k^n) = \mathbb{E}_{\mu_k^n} \left(\langle \vartheta(\mu_k^n), f \rangle \right). \quad (\text{III.9})$$

Moreover, the associated matching coverage can be rewritten as

$$\mathbf{M}_\Phi^n(\mu_0^n) = 1 - \frac{\mathcal{I}_n^n}{n} = 1 - \frac{\mu_n^n(\{0\})}{n}. \quad (\text{III.10})$$

All the quantities used in the definition of ϑ depend on the *current state of the measure*. We do not write the dependance to simplify notations.

Remark III.2.1. *From this point forward, all the results are enunciated only on the **cylin-**
drical functions $\Pi_f(\mu) := \langle \mu, f \rangle$ (Dawson [1993]) for $f : \mathbb{R} \mapsto \mathbb{R}$, in our case in partic-
ular f would mostly be continuous and bounded. Theorem 3.2.6 shows that to characterize
the convergence in law of measure valued process, it is enough to display the convergence
over such functions. They are **convergence determining**.*

Proof. Proposition III.2.1

First, we bring back equation (II.14) for iteration number $k \leq n$:

$$\left\{ \begin{array}{l} \mathcal{A}_{k+1} = \mathcal{A}_k \cup \{(I_1, a_k(I_1) - F_{I_1}), \dots, (I_\ell, a_k(I_\ell) - F_{I_\ell})\} \\ \quad \setminus \{(I, a_k(I)), (I_1, a_k(I_1)), \dots, (I_\ell, a_k(I_\ell))\}; \\ \mathcal{A}_{k+1} = \mathcal{A}_{k+1} \cup \{(I'_1, a_k(I'_1) - D_1), \dots, (I'_u, a_k(I'_u) - D_u)\} \\ \quad \setminus \{(I', a_k(I')), (I'_1, a_k(I'_1)), \dots, (I'_u, a_k(I'_u))\}. \end{array} \right.$$

Now for $k \geq 0$, we recall that μ_k represent the availabilities at iteration k . We can inspect the variations of the availabilities using the operations in equation (II.14). The couples associated to the matched nodes I and I' (or to their degrees K and K') are subtracted from the availability set while all the other couples involved only have theirs updated. Equation (III.8) is only a matter of writing the empirical measures associated to μ_k and μ_{k+1} .

$$\vartheta(\mu_k) = - \left(\delta_K + \delta_{K'} + \sum_{I'_1 \in \mathcal{N}(I)} \left(\delta_{a_k(I'_1)} - \delta_{a_k(I'_1) - D_{I'_1}} \right) + \sum_{I_1 \in \mathcal{N}(I')} \left(\delta_{a_k(I_1)} - \delta_{a_k(I_1) - F_{I_1}} \right) \right). \quad (\text{III.11})$$

If we establish that $\vartheta(\mu_k^n)$ is independent of $(\mu_s^n)_{s \leq k-1}$, it is immediate to write its *kernel* as :

$$\mathfrak{L}^n F(\mu) = \mathbb{E} \left(F(\mu_{k+1}^n) - F(\mu_k^n) \mid \mu_k^n = \mu \right) \quad (\text{III.12})$$

But given the matching criterion Φ , $\Phi(\mu)$ can be chosen to be independent of the process $(\mu_k^n)_{k \geq 0}$ when μ is deterministic, so that the randomness in the choices is independent from that of the process. For example, the independence is achieved if $(\Omega_1, \mathfrak{F}_1)$ is chosen as a sub sigma algebra of (Ω, \mathfrak{F}) so that $\Phi(\mu)$ is \mathfrak{F}_1 measurable and that \mathfrak{F}_1 is independent from $\sigma(\mu_{k,k \geq 0}^n)$. When such a sigma field does not exist, one can for example extend

(Ω, \mathfrak{F}) into $(\Omega \cup \Omega_1, \sigma(\mathfrak{F}, \mathfrak{F}_1))$. This ensures that $\Phi(\mu_k^n)$ only depends on μ_k^n . Moreover, it is immediate that the next steps of the uniform pairing procedure depend only on the current availabilities. Hence, all the operations that define ϑ depend only on the current state of the process, showing that $(\mu_k^n)_{k \geq 0}$ is a Markov chain.

At last, we can revisit the definition of ϑ to obtain the matching coverage as in (III.10). Ending every iteration, the points associated to the degrees of matched nodes are removed from the measure (those are δ_K and $\delta_{K'}$) while their neighbors lose mass. Only matched nodes are removed from the availability set so that at time n , when all the edges are created, the only remaining points in the measure are those affiliated with isolated nodes with no availability (recall that \mathcal{A}_n consists of all the non matched nodes whose availability is now 0). Thus, equation (III.10) is a direct consequence of equation (II.15). ■

Characterizing the matching dynamic as a Markov chain will later allow us to use scaling and the powerful theory of *Markov processes*. The rest of this document will be dedicated to the examination of the large scale behavior of dynamics (Wormald [1999]; Kurtz [1970]) and their concentration around a deterministic function serving as a prediction of their performance. Given a criteria, we will often be able to compute said function, predicting the matching coverage. The convergence will be fully established in chapter IV. For now, the next sections are devoted to tailoring the given Markov chain into a scalable process and to establishing its useful properties.

III.2.2 Moment estimates and properties

For any n , the sequence of measures $(\mu_k^n)_{k \geq 0}$ loses mass (half edges). We formulate this fact in the next proposition:

Proposition III.2.2. *Let $f : \mathbb{N} \rightarrow \mathbb{R}$ be a positive and increasing function, and $k \geq 0$. Then,*

$$\langle \mu_k^n, f \rangle \geq \langle \mu_{k+1}^n, f \rangle. \quad (\text{III.13})$$

Specifically, for all $p > 1$, if $\langle \mu_0^n, \chi^p \rangle < \infty$ then $\langle \mu_k^n, \chi^p \rangle < \infty$ for all k .

Proof. For all such k and f , notice that

$$\langle \mu_{k+1}^n, f \rangle - \langle \mu_k^n, f \rangle = \langle \vartheta(\mu_k^n), f \rangle, \quad (\text{III.14})$$

making it so that (III.13) is directly implied by $\langle \vartheta(\mu_k^n), f \rangle < 0$. But

$$\begin{aligned} \langle \vartheta(\mu_k^n), f \rangle = & - \left(f(K) + f(K') + \sum_{I'_l \in \mathcal{N}(I)} \left(f(a_k(I'_l)) - f(a_k(I'_l) - D_{I'_l}) \right) \right. \\ & \left. + \sum_{I_l \in \mathcal{N}(I')} \left(f(a_k(I_l)) - f(a_k(I_l) - F_{I_l}) \right) \right). \end{aligned}$$

Both $f(K)$ and $f(K')$ are positive because f is so. And for any l , $\left(f(a_k(I'_l)) - f(a_k(I'_l) - D_{I'_l}) \right)$ and $\left(f(a_k(I_l)) - f(a_k(I_l) - F_{I_l}) \right)$ are positive because f is increasing. Hence $\langle \mu_{k+1}^n, f \rangle - \langle \mu_k^n, f \rangle = \langle \vartheta(\mu_k^n), f \rangle < 0$, proving (III.13). Then, applying (III.13) k times yields

$$\infty > \langle \mu_0^n, f \rangle \geq \langle \mu_k^n, f \rangle, \quad (\text{III.15})$$

concluding the proof. ■

The preceding result actually yields an *universal bound over the moments* of the trajectories under $\mathfrak{H}\mathfrak{o}$. Almost all of the results in the next chapter rest on this universal bound.

Corollary III.2.1 (Universal bound). *Suppose that the sequence of Markov chains $(\frac{1}{n}\mu^n)_{n \geq 1}$ satisfies $\mathfrak{H}\mathfrak{o}$. Then,*

$$\sup_{n,k} \frac{1}{n} \langle \mu_k^n, \chi^p \rangle \leq M < \infty \quad (\text{III.16})$$

for some $M > 0$.

Proof. For any k, n , (III.13) brings forth $\frac{1}{n} \langle \mu_k^n, \chi^p \rangle \leq \frac{1}{n} \langle \mu_0^n, \chi^p \rangle$. But since $\frac{1}{n}\mu_0^n$ satisfies $\mathfrak{H}\mathfrak{o}$, $\frac{1}{n} \langle \mu_0^n, \chi^p \rangle$ is convergent and therefore bounded by some finite M . We write

$$\frac{1}{n} \langle \mu_k^n, \chi^p \rangle \leq \frac{1}{n} \langle \mu_k^0, \chi^p \rangle \leq M, \quad (\text{III.17})$$

which concludes the proof. ■

We conclude this section by providing yet another estimate. We assess the scale of the typical degrees in the local matching dynamic. This result of capital importance will be useful all around for bounds on the amplitude of matching criteria². We start with a general preliminary result,

²By amplitude, we mean the number of half-edges created per iteration

Lemma III.2.1. *Let a degree sequence $\{d^n = (d^n(1), \dots, d^n(i)); n \in \mathbb{N}^*\}$ and the corresponding point measures $\{\mu^n; n \in \mathbb{N}^*\} = \{\sum_{1 \leq i \leq n} \delta_{d(i)}; n \in \mathbb{N}^*\}$. Suppose that for all $f \in \mathcal{C}_b(\mathbb{N}) \cup \{\chi^2\}$,*

$$\lim_{n \rightarrow \infty} \sum_i \frac{f(d^n(i))}{n} = \lim_{n \rightarrow \infty} \frac{\langle \mu^n, f \rangle}{n} = \langle \nu, f \rangle < \infty \quad (\text{III.18})$$

for some ν in \mathcal{M}_F , and define the maximal degree

$$d_{max}^n = \max_i d^n(i).$$

Then we have that

$$\lim_{n \rightarrow \infty} \frac{d_{max}^n}{\sqrt{n}} = 0. \quad (\text{III.19})$$

Proof. Take $\epsilon > 0$. Since $\langle \nu, \chi^2 \rangle < \infty$, we gather the existence of M such that

$$\langle \nu, \chi^2 \rangle - \langle \nu, \chi^2 \mathbb{1}_{[0, \sqrt{M}]} \rangle < \epsilon.$$

But as $\chi^2 \mathbb{1}_{[0, \sqrt{M}]} \in \mathcal{C}_b(\mathbb{N})$ we also have that

$$\lim_{n \rightarrow \infty} \frac{\langle \mu^n, \chi^2 \mathbb{1}_{[0, \sqrt{M}]} \rangle}{n} = \langle \nu, \chi^2 \mathbb{1}_{[0, \sqrt{M}]} \rangle.$$

Thus,

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{\langle \mu^n, \chi^2 \mathbb{1}_{[\sqrt{M}+1, \infty)} \rangle}{n} &= \lim_{n \rightarrow \infty} \left(\frac{\langle \mu^n, \chi^2 \rangle}{n} - \langle \mu^n, \chi^2 \mathbb{1}_{[0, \sqrt{M}]} \rangle \right) \\ &= \langle \nu, \chi^2 \rangle - \langle \nu, \chi^2 \mathbb{1}_{[0, \sqrt{M}]} \rangle < \epsilon. \end{aligned} \quad (\text{III.20})$$

Now, assuming that there exists n such that $\frac{(d_{max}^n)^2}{n} \geq \epsilon \vee M/n$ would imply that

$$\frac{\langle \mu^n, \chi^2 \mathbb{1}_{[\sqrt{M}+1, \infty)} \rangle}{n} \geq \epsilon,$$

a contradiction to (III.20). Therefore $\frac{(d_{max}^n)^2}{n}$ is either upper-bounded by ϵ or M/n . In any case we get that

$$\limsup_{n \rightarrow \infty} \frac{(d_{max}^n)^2}{n} < \epsilon,$$

and we conclude the proof by letting ϵ be arbitrarily small. \blacksquare

Proposition III.2.3 (Order for the availability of nodes). *Let $\{\mu^n; n \in \mathbb{N}^*\}$ be a sequence of processes satisfying $\mathfrak{H}\mathfrak{o}$. For all $k \in \llbracket 0, n \rrbracket$, let $K_{\mu_k^n}$ and $K'_{\mu_k^n}$ be the degree choices under a generic choice criterion Φ . Then the following relation holds*

$$\lim_{n \rightarrow \infty} \frac{K_{\mu_k^n}}{\sqrt{n}} = \lim_{n \rightarrow \infty} \frac{K'_{\mu_k^n}}{\sqrt{n}} = 0. \quad (\text{III.21})$$

Proof. The result is proven in two steps. First, we prove that for $k \in \llbracket 0, n \rrbracket$, it is clear that if $M_k^n = \max_i a_k^n(v_i)$ then M_k^n is decreasing in k . The procedure is essentially the same as in the proof of (III.13). By definition, $\forall l > M_k^n, \mu_k(l) > 0$. As a consequence,

$$\begin{aligned} \mu_{k+1}^n(M_k^n) - \mu_k^n(M_k^n) &= \langle \vartheta(\mu_k^n), \mathbb{1}_{M_k^n} \rangle \\ &= - \left(\mathbb{1}_{M_k^n}(K) + \mathbb{1}_{M_k^n}(K') + \sum_{I'_l \in \mathcal{N}(I)} \left(\mathbb{1}_{M_k^n}(a_k(I'_l)) - \mathbb{1}_{M_k^n}(a_k(I'_l) - D_{I'_l}) \right) \right. \\ &\quad \left. + \sum_{I_l \in \mathcal{N}(I')} \left(\mathbb{1}_{M_k^n}(a_k(I_l)) - \mathbb{1}_{M_k^n}(a_k(I_l) - F_{I_l}) \right) \right). \end{aligned} \quad (\text{III.22})$$

But, $\mathbb{1}_{M_k^n}(a_k(I'_l) - D_{I'_l}) = 1$ would mean that $a_k(I'_l) > M_k^n$, which is impossible by definition of M_k^n . Therefore, for all l , $\mathbb{1}_{M_k^n}(a_k(I'_l) - D_{I'_l}) = 0$ and likewise, $\mathbb{1}_{M_k^n}(a_k(I_l) - F_{I_l}) = 0$. Ultimately, $\langle \vartheta(\mu_k^n), \mathbb{1}_{M_k^n} \rangle \leq 0$ and $\mu_k^n(M_k^n)$ is decreasing in k .

But under $\mathfrak{H}\mathfrak{o}$, $\frac{1}{n}\mu_0^n$ satisfies the conditions of Lemma III.2.1 because

$$\lim_{n \rightarrow \infty} \left\langle \frac{1}{n}\mu_0^n, \chi^2 \right\rangle = \langle \nu, \chi^2 \rangle$$

and $\lim_{n \rightarrow \infty} \langle \frac{1}{n}\mu_0^n, f \rangle = \langle \nu, f \rangle$ for $f \in \mathcal{C}_b$. It means that $d_{max}^n(\mu_0^n) = M_0^n$ is $o(\sqrt{n})$. The assertion is proved by noticing that for all k ,

$$K_{\mu_k^n} \leq M_k^n \leq M_0^n. \quad (\text{III.23})$$

Proving the estimate for $K'_{\mu_k^n}$ is more subtle. The neighbors of a vertex do not have μ_k^n as their degree distribution. Since they are discovered following an edge, they follow a *size biased* version of $\mu_k^n - \delta_{K_{\mu_k^n}}$:

$$\left\langle \overline{\mu_k^n - \delta_{K_{\mu_k^n}}}, f \right\rangle := \frac{\left\langle \mu_k^n - \delta_{K_{\mu_k^n}}, \chi f \right\rangle}{\left\langle \mu_k^n - \delta_{K_{\mu_k^n}}, \chi \right\rangle}, \quad (\text{III.24})$$

in other words the chances that some half-edge is linked to a node is proportional to the nodes' number of stubs. After replacing $\langle \mu_k^n, f \rangle$ by $\langle \overline{\mu_k^n - \delta_{K_{\mu_k^n}}}, f \rangle$, the proof is the same. Only, we do have to show that if $\frac{1}{n}\mu_0^n$ satisfies $\mathfrak{H}\mathfrak{O}$, $\frac{1}{n}(\overline{\mu_k^n - \delta_{K_{\mu_k^n}}})$ still verifies the conditions of Lemma III.2.1. And it does, since for all bounded f ,

$$\lim_{n \rightarrow \infty} \frac{\langle \overline{\mu_k^n - \delta_{K_{\mu_k^n}}}, f \rangle}{n} = \lim_{n \rightarrow \infty} \frac{\langle \mu_0^n - \delta_{K_{\mu_0^n}}, \chi f \rangle}{\langle \mu_0^n - \delta_{K_{\mu_0^n}}, \chi \rangle} = \lim_{n \rightarrow \infty} \frac{\langle \check{\mu}_k^n - \delta_{K_{\mu_k^n}}, f \rangle}{n} = \frac{\langle \nu, \chi f \rangle}{\langle \nu, \chi \rangle}$$

by dominated convergence, and again

$$\lim_{n \rightarrow \infty} \frac{\langle \check{\mu}_k^n, \chi^2 \rangle}{n} = \lim_{n \rightarrow \infty} \frac{\langle \mu_0^n, \chi^3 \rangle}{\langle \mu_0^n, \chi \rangle} = \frac{\langle \nu, \chi^3 \rangle}{\langle \nu, \chi \rangle}.$$

■

Remark III.2.2. *For the final part of this proof, we used the fact that the neighbors asymptotically follow a size-biased version of μ_k^n . Otherwise, using exact laws causes small deviations from the perfect size biased version of μ_k^n (for example, in the case of the proof, a term in $\delta_{K_{\mu_k^n}}$ appeared). This deviation makes the exact generator of the matching algorithms inconvenient to use. Thankfully, we will see in the next chapter that the deviation term vanishes when n gets large and that we can approximate the exact generator by that of a much simpler dynamic.*

Notes III.2.1. *It is seen here that the existence of at least a third moment guarantees a convenient order for the law of the maximum degree, further justifying our hypothesis.*

III.3 Normalization and scaling

Finally, we aim to provide a suitable normalization for the Markov chain dynamic. In fact, the scaled process provided here will be the real object of our study. Later, it is used to study the asymptotic properties of the criteria.

III.3.1 Normalization

Fix $n \in \mathbb{N}^*$. Towards that purpose, we recall

$$\mathcal{M}_n := \frac{1}{n} \mathcal{M}_F^n,$$

where \mathcal{M}_F^n is the space of point measures with integer masses and with total mass smaller than n .

The $\bar{}$ notation is used to signal normalized objects.

Naturally, the process μ^n can be extended to a rescaled right-continuous with left-hand limits (*RCLL*) semimartingale. As is often the case, the limiting behavior of the normalized process will be a continuous process (Kurtz [1971]). This is why we extend the Markov chain into a process on $\mathbb{D}([0, 1], \mathcal{M}_F)$, giving us access to the *Skorokhod topology* (Méléard and Roelly [1993]; Joffe and Metivier [1986]) for that space. For $0 \leq t \leq 1$ we denote

$$\bar{\mu}_t^n = \frac{1}{n} \mu_{[nt]}^n,$$

where $[x] = e \in \mathbb{N}, e \leq x < e + 1$. We then accelerate time allowing the matching coverage to conveniently become

$$\mathbf{M}_{\Phi}^n(\mu_0^n) = 1 - \bar{\mu}_1^n(\{0\}). \tag{III.25}$$

This goes to show that the rescaled process converges directly in tune with the matching coverage. One immediate downside is that $\bar{\mu}^n$ is no more a Markov process on $\mathbb{D}([0, 1], \mathcal{M}_F)$. After all, its jump times are deterministic. Moreover, we still have no guarantee on the behavior of the criteria through normalization.

Since choice functions are defined on $(\cup_{n \geq 1} \mathcal{M}_F^n)$, they can naturally be extended to functions on normalized degree sequences: $(\cup_{n \geq 1} \mathcal{M}_n)$. Again for a normalized degree sequence $\bar{\mu}^n \in \mathcal{M}_n$, we will use the superscript $\bar{}$ and we designate the choice functions on normalized sequences as:

$$\bar{\Phi}(\bar{\mu}^n) = \Phi(n\bar{\mu}^n). \tag{III.26}$$

On one hand, since the space $(\cup_{n \geq 1} \mathcal{M}_n)$ is dense in \mathcal{M} (recall that \mathcal{M} is the subspace of \mathcal{M}_F with total masses ≤ 1)³, a natural assumption to make about choice functions would be that, if $\bar{\mu}^n \rightarrow \bar{\mu}$ for some adequate topology stronger than the weak topology, then it should also be true that $\bar{\Phi}(\bar{\mu}^n) \rightarrow \bar{\Phi}(\bar{\mu})$.

On the other hand however, even though some criteria are well defined on \mathcal{M}_n for every n , sometimes they can not trivially be extended to the whole space \mathcal{M} . As an example, take $\bar{\mu}$ to be a measure with unbounded support, the truncations of $\bar{\mu}$ on \mathcal{M}_n would admit a maximal degree while $\text{MAX}(\bar{\mu})$ is obviously not defined. As a consequence, we have to impose some further restrictions (that would be either on the criteria under consideration or on the domains considered).

³For $\bar{\mu} \in \mathcal{M}$, take $\bar{\mu}^n = \sum_{i \leq n} \frac{[n\mu(\{i\})]}{n} \delta_i$

- Definition III.3.1.** 1. A criterion is (asymptotically) **well behaved** on a (closed) subspace \mathcal{M}_Φ of \mathcal{M} , when for all $n \geq 1$, $\bar{\mu}^n \in \mathcal{M}_n$ and $\bar{\mu} \in \mathcal{M}_\Phi$ such that $\bar{\mu}^n \rightarrow \bar{\mu}$, we have $\bar{\Phi}(\bar{\mu}^n) \rightarrow \bar{\Phi}(\bar{\mu})$ for a topology that is stronger than the weak topology.
2. A criterion **preserves the moments** of order up to $p \geq 1$ if, for all $\bar{\mu} \in \mathcal{M}_\Phi$ and $v < p$, $\langle \bar{\mu}, \chi^v \rangle < \infty$ implies $\langle \bar{\Phi}(\bar{\mu}), \chi^v \rangle < \infty$

The order of the moment preservation is the important part. For example, if the base measure μ has moments up to an order $p + 1$, the uniform choice on its size biased version only preserves the moments up to p . This means the properties we can prove on our algorithms should also be true only up to order p . Moreover, take μ as any measure with exponential moments, technically $\Phi(\mu)$ could be any probability law, especially one admitting no moments of higher order thus failing to preserve all the moments of the initial distribution.

Proposition III.3.1. Let $p \geq 1$, $n \geq 1$ and μ^n a sequence of Markov chains under the criterion Φ , such that :

1. Φ is well behaved and it preserves the moments of order p
2. $\frac{1}{n} \langle \mu_0^n, \chi^p \rangle \xrightarrow{n} \langle \bar{\mu}, \chi^p \rangle$ for some $\bar{\mu} \in \mathcal{M}$

Then, there exists $0 < M < \infty$

$$\sup_{n>0, t>0} \langle \bar{\Phi}(\bar{\mu}^n), \chi^p \rangle < M \tag{III.27}$$

Proof. Since Φ preserves moments, this is a direct consequence of III.16. ■

III.3.2 Continuous time Markov chain

We finally undertake the restoration of the Markov property for the normalized continuous processes. To extend the discrete time Markov chain into a continuous one, assume $\theta_1, \theta_2, \dots$ is a sequence of independent identically distributed exponential random variables with parameter 1. Moreover, assume that $(\theta_i)_{i>0}$ is independent from $\bar{\mu}_t^n$. Define the following sequence of random times

$$\tau_0 = 0, \tau_{i+1} = \tau_i + \theta_i/n.$$

The associated **continuous time Markov chain** to $\bar{\mu}^n$ is the following process :

$$t \longmapsto \tilde{\mu}_t^n := \bar{\mu}_{l/n}^n \quad \text{for } \tau_l \leq t < \tau_{l+1}, l \in \mathbb{N}. \tag{III.28}$$

From now on, the tilde~ superscript is used for the objects related to the continuous Markov extension.

Proposition III.3.2. *For all $n \in \mathbb{N}^*$, $\tilde{\mu}^n$ is a continuous time Markov extension in the sense that*

1. $\tilde{\mu}^n$ is a Feller Markov process with generator $\tilde{\mathcal{L}}^n$:

$$\tilde{\mathcal{L}}^n F(\bar{\mu}) := n \mathbb{E}_{\bar{\mu}} \left(F \left(\bar{\mu} + \frac{1}{n} \vartheta(n\bar{\mu}) \right) - F(\bar{\mu}) \right),$$

for $\bar{\mu}$ in $\bar{\mathcal{M}}^n$, in particular,

$$\tilde{\mathcal{L}}^n \Pi_f(\bar{\mu}) := \mathbb{E}_{\bar{\mu}} \langle \vartheta(n\bar{\mu}), f \rangle.$$

2. For $f \in \mathcal{C}_b$,

$$\left| \mathbb{E}_{\tilde{\mu}_0^n} \langle \tilde{\mu}_t^n, f \rangle - \mathbb{E}_{\bar{\mu}_0^n} \langle \bar{\mu}_t^n, f \rangle \right| \leq \|\tilde{\mathcal{L}}^n \Pi_f\| |1/n + (t/n)^{1/2}|.$$

Proof. Fix $n \in \mathbb{N}^*$. Since $\tilde{\mu}_t^n$ is a pure jump markov process with bounded intensities, it is Feller continuous (see Theorem 12.18 of Kallenberg [2002], or Ethier and Kurtz [1986] p. 163). Thus it admits a strongly defined generator: For all $F \in \mathcal{C}_b(\mathcal{M}^n)$ and $h > 0$,

$$\begin{aligned} \mathbb{E}_{\tilde{\mu}_0^n} (F(\tilde{\mu}_h^n) - F(\tilde{\mu}_0^n)) &= \mathbb{P}(\tau_1 > h) \times 0 \\ &\quad + \mathbb{P}(\tau_1 < h, \tau_2 > h) \mathbb{E}_{\tilde{\mu}_0^n} (F(\tilde{\mu}_{\tau_1}^n) - F(\tilde{\mu}_0^n) \mid \tau_1 < h, \tau_2 > h) \\ &\quad + \mathbb{P}(\tau_1 < h, \tau_2 < h) \mathbb{E}_{\tilde{\mu}_0^n} (F(\tilde{\mu}_h^n) - F(\tilde{\mu}_0^n) \mid \tau_1 < h, \tau_2 < h) \end{aligned}$$

Notice that $\mathbb{P}(\tau_1 < h, \tau_2 > h) = \mathbb{P}(\theta_1/n < h, \theta_1/n + \theta_2/n > h) = (hn + o(nh))$ and $\mathbb{P}(\tau_1 < h, \tau_2 < h) = (o(nh))$ because θ_1 and θ_2 are independent exponential r.v.'s. We then write

$$\begin{aligned} \mathbb{E}_{\tilde{\mu}_0^n} (F(\tilde{\mu}_h^n) - F(\tilde{\mu}_0^n)) &= nh \mathbb{E}_{\tilde{\mu}_0^n} (F(\tilde{\mu}_{\tau_1}^n) - F(\tilde{\mu}_0^n) \mid \tau_1 < h, \tau_2 > h) + o(nh) \\ &= nh \mathbb{E}_{\tilde{\mu}_0^n} (F(\bar{\mu}_0^n + (1/n)\vartheta(n\bar{\mu}_0^n)) - F(\bar{\mu}_0^n)) + o(nh), \end{aligned}$$

proving that

$$\lim_{h \rightarrow 0} \frac{1}{h} \mathbb{E}_{\tilde{\mu}_0^n} (F(\tilde{\mu}_h^n) - F(\tilde{\mu}_0^n)) = \tilde{\mathcal{L}}^n F(\tilde{\mu}_0^n).$$

As for the second assertion, we first introduce

$$\gamma_t := \sum_{i=0}^{\lfloor nt \rfloor - 1} \theta_{i+1}/n. \quad (\text{III.29})$$

By definition, for all t we have that $\bar{\mu}_t^n = \tilde{\mu}_{\gamma_t}^n$. Assertion 1 showed that $\tilde{\mu}^n$ is Markov, whence the expression

$$\mathbb{E}_{\bar{\mu}_0^n} \langle \bar{\mu}_t^n, f \rangle = \mathbb{E}_{\bar{\mu}_0^n} \left(\langle \tilde{\mu}_0^n, f \rangle + \int_0^t \mathfrak{L}^n \langle \tilde{\mu}_s^n, f \rangle ds \right), \quad (\text{III.30})$$

Plus, because γ_t is a stopping time

$$\mathbb{E}_{\bar{\mu}_0^n} \langle \bar{\mu}_t^n, f \rangle = \left(\mathbb{E}_{\bar{\mu}_0^n} \langle \tilde{\mu}_0^n, f \rangle + \int_0^{\gamma_t} \mathfrak{L}^n \langle \tilde{\mu}_s^n, f \rangle ds \right) \quad (\text{III.31})$$

Which we follow up with

$$\begin{aligned} |\mathbb{E}_{\bar{\mu}_0^n} \langle \bar{\mu}_t^n, f \rangle - \mathbb{E}_{\bar{\mu}_0^n} \langle \tilde{\mu}_t^n, f \rangle| &= \left| \mathbb{E}_{\bar{\mu}_0^n} \int_{\gamma_t}^t \mathfrak{L}^n \langle \tilde{\mu}_s^n, f \rangle ds \right| \\ &\leq \sup_s |\mathfrak{L}^n \langle \tilde{\mu}_s^n, f \rangle| |\mathbb{E}_{\bar{\mu}_0^n} |\gamma_t - t|| \quad (\text{III.32}) \end{aligned}$$

But

$$\begin{aligned} \mathbb{E}_{\bar{\mu}_0^n} |\gamma_t - t| &= \mathbb{E}_{\bar{\mu}_0^n} \left| \gamma_t - \frac{\lfloor nt \rfloor + nt - \lfloor nt \rfloor}{n} \right| \\ &\leq 1/n + \mathbb{E}_{\bar{\mu}_0^n} \left| \sum_{i=0}^{\lfloor nt \rfloor - 1} \frac{\theta_{i+1} - 1}{n} \right| \leq 1/n + \left(\sum_{i=0}^{\lfloor nt \rfloor - 1} \frac{1}{n^2} \right)^{1/2} \leq 1/n + (t/n)^{1/2} \quad (\text{III.33}) \end{aligned}$$

where we used the fact that $\mathbb{E}\theta_i^2 = 2$. ■

As for now, this means the convergence of $\bar{\mu}^n$ can be estimated by that of a measure valued Markov process. The last key component for outlining the convergence is exhibiting the characteristic martingale. We tend to its study in the final part of chapter III.

III.3.3 Semi-martingale representation

We have just proven that for all n , $\tilde{\mu}^n$ is a Feller process having generator $\tilde{\mathfrak{L}}^n$. For any f and $n \geq 1$, the martingale problem for $\tilde{\mu}^n$ yields that:

Proposition III.3.3. *The process defined for all $t \in [0, 1]$ and $f \in \mathcal{C}_b$ by*

$$\mathcal{M}_t^{f,n} = \langle \tilde{\mu}_t^n, f \rangle - \langle \tilde{\mu}_0^n, f \rangle - \int_0^t \tilde{\mathfrak{L}}^n \Pi_f(\bar{\mu}) ds \quad (\text{III.34})$$

is a square integrable martingale. Moreover, its predictable quadratic variation is equal to :

$$\langle\langle \mathcal{M}^{f,n} \rangle\rangle_t = \frac{1}{n} \int_0^t \mathbb{E}_{\tilde{\mu}_s^n} (\langle \vartheta(\tilde{\mu}_s^n), f \rangle^2) ds. \quad (\text{III.35})$$

For a \mathbb{R} valued semi-martingale X , we designate by $[X]$ its *quadratic variation*. Thus, if $0 = t_0 < t_1 < \dots \leq t_k = t$ ($k \in \mathbb{N}$) is a subdivision of the time interval, $[X]$ is defined by the following expression when it exist:

$$[X]_t = \lim_{\max t_i - t_{i-1} \rightarrow 0, 1 \leq i \leq k} \sum_{1 \leq j \leq k} (X_{t_j} - X_{t_{j-1}})^2,$$

where the limit is taken in probability.

We also designate by $\langle\langle X \rangle\rangle$ the predictable quadratic variation of X , the unique predictable compensator of $[X]$.

Proof. It is a characteristic property of Markov processes that $\mathcal{M}_t^{f,n}$ is martingale (e.g 5.1.2 in Dawson [1993], or the chapter 4 of Ethier and Kurtz [1986]). We simply need to compute the quadratic variation.

The book of Decreusefond and Moyal [2012] gives an explicit form for the quadratic variation of measure valued pure jump processes in its Theorem 7.15. Here we adapt the proof to the particular case of the matching process previously defined. Fix $t \geq 0$ and recall that if $\Pi_f^2(\tilde{\mu}_t^n) = \langle \tilde{\mu}_t^n, f \rangle^2$,

$$\begin{aligned} \mathfrak{L}^n \Pi_f^2(\tilde{\mu}_t^n) &= n \mathbb{E}_{\tilde{\mu}_t^n} \left(\Pi_f^2(\tilde{\mu}_t^n + \frac{1}{n} \vartheta(n\tilde{\mu}_t^n)) - \Pi_f^2(\tilde{\mu}_t^n) \right) \\ &= n \mathbb{E}_{\tilde{\mu}_t^n} \left(\langle \tilde{\mu}_t^n, f \rangle^2 + 2 \frac{1}{n} \langle \vartheta(n\tilde{\mu}_t^n), f \rangle \langle \tilde{\mu}_t^n, f \rangle + \frac{1}{n^2} \langle \vartheta(n\tilde{\mu}_t^n), f \rangle^2 - \langle \tilde{\mu}_t^n, f \rangle^2 \right) \\ &= n \mathbb{E}_{\tilde{\mu}_t^n} \left(2 \frac{1}{n} \langle \vartheta(\tilde{\mu}_t^n), f \rangle \langle \tilde{\mu}_t^n, f \rangle + \frac{1}{n^2} \langle \vartheta(\tilde{\mu}_t^n), f \rangle^2 \right). \end{aligned}$$

On one hand, the martingale problem for Π_f^2 gives that

$$\Pi_f^2(\tilde{\mu}_t^n) - \Pi_f^2(\tilde{\mu}_0^n) - \int_0^t \mathfrak{L}^n \Pi_f^2(\tilde{\mu}_s^n) ds \quad (\text{III.36})$$

is a martingale. On the other side, since $X_t := \langle \tilde{\mu}_t^n, f \rangle$ is a finite variation pure jump

process, it is in particular a semi-martingale. For $g \in \mathcal{C}^2(\mathbb{R})$, where $\mathcal{C}^2(\mathbb{R})$ is the space of doubly differentiable functions with continuous second derivatives, Itô's formula (See Jacod and Shiryaev [2013] theorem 4.57 p57) yields :

$$g(X_t) - g(X_0) = \int_0^t g'(X_{s-})dX_s + \frac{1}{2} \int_0^t g''(X_{s-})d[X_s]^c + \sum_{s \leq t} \left(g(X_s) - g(X_{s-}) - g'(X_{s-})\Delta X_s - \frac{1}{2}g''(X_{s-})(\Delta X_s)^2 \right).$$

In our particular case, taking $g(x) = x^2$ gives

$$\langle \tilde{\mu}_t^n, f \rangle^2 - \langle \tilde{\mu}_0^n, f \rangle^2 = 2 \int_0^t \langle \tilde{\mu}_{s-}^n, f \rangle d \langle \tilde{\mu}_s^n, f \rangle + \sum_{s \leq t} (\langle \tilde{\mu}_s^n, f \rangle - \langle \tilde{\mu}_{s-}^n, f \rangle)^2. \quad (\text{III.37})$$

Noticing that the martingale problem for $\langle \tilde{\mu}_s^n, f \rangle$ is equivalent, in its differential form, to $d \langle \tilde{\mu}_s^n, f \rangle = d\mathcal{M}_t^{f,n} + \mathfrak{L}^n \Pi_f(\tilde{\mu}_s^n) ds$ and also that $\sum_{s \leq t} (\langle \tilde{\mu}_s^n, f \rangle - \langle \tilde{\mu}_{s-}^n, f \rangle)^2 = [\mathcal{M}^{f,n}]_t$ because $\langle \mu_s, f \rangle$ et $\mathcal{M}_s^{f,n}$ have the same jumps. The last expression easily implies that

$$\langle \tilde{\mu}_t^n, f \rangle^2 - \langle \tilde{\mu}_0^n, f \rangle^2 = 2 \int_0^t \langle \tilde{\mu}_{s-}^n, f \rangle d\mathcal{M}_t^{f,n} + 2 \int_0^t \langle \tilde{\mu}_{s-}^n, f \rangle \tilde{\mathfrak{L}}^n \Pi_f(\tilde{\mu}_s^n) ds + [\mathcal{M}^{f,n}]_t \quad (\text{III.38})$$

But a difference of (local) martingales is a (local) martingale, and (III.36) is compared to (III.38) yielding that

$$[\mathcal{M}^{f,n}]_t - \left(\int_0^t \left(\tilde{\mathfrak{L}}^n \Pi_f^2(\tilde{\mu}_s^n) - 2 \langle \tilde{\mu}_{s-}^n, f \rangle \tilde{\mathfrak{L}}^n \Pi_f(\tilde{\mu}_s^n) \right) ds \right) \quad (\text{III.39})$$

is a martingale. By uniqueness of the predictable compensator of $[\mathcal{M}^{f,n}]$, we can conclude that

$$\ll \mathcal{M}^{f,n} \gg_t = \int_0^t \left(\mathfrak{L}^n \Pi_f^2(\tilde{\mu}_s^n) - 2 \langle \tilde{\mu}_{s-}^n, f \rangle \tilde{\mathfrak{L}}^n \Pi_f(\tilde{\mu}_s^n) \right) ds = n \mathbb{E}_{\tilde{\mu}_t^n} \left(\left(\frac{1}{n} \langle \vartheta(\tilde{\mu}_t^n), f \rangle \right)^2 \right).$$

Finally, we prove that $\mathcal{M}^{f,n}$ is square integrable. For all $t \in [0, 1]$, considering that

$f(a_t) \leq \|f\|$ and $a_t < n$, we beget

$$\begin{aligned}
 \ll \mathcal{M}^{f,n} \gg_t &= \frac{1}{n} \mathbb{E}_{\tilde{\mu}_t^n} (\langle \vartheta(\tilde{\mu}_t^n), f \rangle^2) \\
 &= \frac{1}{n} \mathbb{E}_{\tilde{\mu}_t^n} \left(f(K) + f(K') + \sum_{I'_l \in \mathcal{N}(I)} \left(f(a_k(I'_l)) - f(a_k(I'_l) - D_{I'_l}) \right) \right. \\
 &\quad \left. + \sum_{I_l \in \mathcal{N}(I')} \left(f(a_k(I_l)) - f(a_k(I_l) - F_{I_l}) \right) \right)^2 \\
 &\leq \|f\|^2 (2 + 4n)^2 < \infty,
 \end{aligned}$$

which implies that $\mathcal{M}^{f,n}$ is indeed a square integrable martingale. ■

Chapter IV

Performance Prediction and Concentration Results

In this final chapter, we enunciate and we prove the main convergence result. As the number of nodes n becomes large, the matching coverage is concentrated around a deterministic value. In a similar fashion to the law of large numbers, this value will be used as a predictor for the performance of any given criterion. Subsequently, those predictors offer a framework for the comparison of matching criteria. Our work here closely follows that of Bermolen et al. [2017] on the jamming constant and can be seen in the context of the differential equation method (eg. Wormald [1999]). Pioneered by Kurtz [1971], this method consists of approaching random dynamics (especially on networks) by a limiting deterministic function solution of an Ordinary Differential Equation. Our particular procedure is an extension of the method in turn introduced by (Decreusefond et al. [2012]) where the Differential Equation Method was extended to measure valued processes. It is closely tied to Noiry et al. [2021] where the same approach is applied to online bipartite matching, only the degree distribution is viewed in the light of its generating functions instead of using a measure valued approach.

The chapter is comprised of three main sections. In the first section, the main result of our research is established: the stochastic matching dynamic converges to a deterministic continuous function. A direct implication of this result is that the matching coverage also converges to a deterministic value. Evermore, we support this result by updating the simulations given in chapter II with the predicted values. Afterward, the following section is devoted to the proof of the main convergence result. The proof works as a typical convergence result in the Skorokhod topology. First, the existence of several limits is established via tightness, before the limits are characterized as unique with the use of the martingale property. The final section of the chapter is devoted to an application of the main result on the GREEDY criterion.

IV.1 Main theorem

In this first section, we enunciate the main result of the thesis. Under properly defined criteria, the (random) dynamic defined by the algorithms, and by extension their matching coverage, converges to deterministic values.

First, we have to establish the asymptotic behavior of any given algorithm. Remark III.2.2 hinted that the exact law of our processes is inconvenient to deal with. We provide an alternative dynamic whose generator encodes the asymptotic behavior of criteria despite being simpler to use.

Definition IV.1.1 (Alternative dynamic). *Fix a counting measure $\mu \in \mathcal{M}_F(\mathbb{N})$. To each $i \in \mathbb{N}^*$ are associated $\mu(i)$ buckets of i items. (Buckets and items correspond to nodes and half-edges in the original construction.) So there are $b := \langle \mu, \mathbf{1}_{\mathbb{N}^*} \rangle$ buckets and a total of $c = \langle \mu, \chi \rangle$ items. We label the buckets arbitrarily, as $1, \dots, b$. Denote for any $j = 1, \dots, b$, $a(j)$ the cardinality of bucket j . We also label the items from 1 to c as follows: items 1 to $a(1)$ are the elements of bucket 1, labeled arbitrarily, items $a(1) + 1, \dots, a(1) + a(2)$ are the elements of bucket 2, and so on... For any $i = 1, \dots, c$, we let $B(i)$ be the bucket of item i .*

We perform the following random experiment, which mimics the original dynamics when the sampling is performed with replacement:

- (i) *We draw an element, say j , of $\llbracket 1, b \rrbracket$, following $\Phi_K(B(1), \dots, B(b))$ and set $\hat{I} := j$, the corresponding bucket. We denote by $\hat{K} = a(\hat{I})$, its cardinality. To differentiate this choice from the one without replacement, we denote its law $\Phi_{\hat{K}}^1$.*
- (ii) *Then we draw uniformly at random, and **with replacement**, \hat{K} items among c , denoted by $i_1, \dots, i_{\hat{K}}$. For all $\ell \in \llbracket 1, \hat{K} \rrbracket$ we denote by $\hat{B}_\ell = B(i_\ell)$, the bucket i_ℓ belongs to.*
- (iii) *Following $\Phi_{K'}(B_1, \dots, B_\ell)$, one item $i_{\hat{K}}$ is drawn among c . We also denote by $\hat{I}' = B(i_{\hat{K}})$, the match of \hat{I} , and by $\hat{K}' = a(\hat{I}')$, the cardinality of that bucket. Again, to differentiate this choice from the one without replacement, we denote its law $\Phi_{\hat{K}'}$.*
- (iv) *If $\hat{K}' > 1$, we draw uniformly at random, and with replacement, $\hat{K}' - 1$ items among c , denoted by $i'_1, \dots, i'_{\hat{K}'-1}$. Then, for all $\ell \in \llbracket 1, \hat{K}' - 1 \rrbracket$ we denote by $\hat{B}'_\ell = B(i'_\ell)$, the bucket i'_ℓ belongs to.*

¹In reality, K and \hat{K} have the same law

Finally, we can exhibit the measure representation induced by that procedure,

$$\widehat{\vartheta}(\mu) = - \left(\delta_{\widehat{K}} + \delta_{\widehat{K}'} + \sum_{\ell \in [1, \widehat{K}-1]} \left(\delta_{a(\widehat{B}_\ell)} - \delta_{a(\widehat{B}_\ell)-1} \right) + \sum_{\ell \in [1, \widehat{K}'-1]} \left(\delta_{a(\widehat{B}'_\ell)} - \delta_{a(\widehat{B}'_\ell)-1} \right) \right). \quad (\text{IV.1})$$

Thus, the approximated dynamic is written as follows,

Lemma IV.1.1. *For any counting measure μ such that $\langle \mu, \chi \rangle > 0$ and any $f \in \mathcal{C}_b$, we have*

$$\mathbb{E}_\mu \left[\left\langle \widehat{\vartheta}(\mu), f \right\rangle \right] = \mathfrak{L}\Pi_f(\mu),$$

where,

$$\mathfrak{L}\Pi_f(\mu) := - \left(\left\langle \Phi_{\widehat{K}}(\mu), f \right\rangle + \left\langle \Phi'_{\widehat{K}'}(\mu), f \right\rangle + \frac{\langle \mu, \chi \nabla f \rangle}{\langle \mu, \chi \rangle} \left(\left\langle \Phi_{\widehat{K}}(\mu), \chi - 1 \right\rangle + \left\langle \Phi'_{\widehat{K}'}(\mu), \chi - 1 \right\rangle \right) \right), \quad (\text{IV.2})$$

or recalling from equation (III.26) that $\bar{\Phi}$ defines the matching criterion adapted to normalized (instead of counting) measures:

$$\bar{\mathfrak{L}}\Pi_f(\bar{\mu}) := - \left(\left\langle \bar{\Phi}_{\widehat{K}}(\bar{\mu}), f \right\rangle + \left\langle \bar{\Phi}'_{\widehat{K}'}(\bar{\mu}), f \right\rangle + \frac{\langle \bar{\mu}, \chi \nabla f \rangle}{\langle \bar{\mu}, \chi \rangle} \left(\left\langle \bar{\Phi}_{\widehat{K}}(\bar{\mu}), \chi - 1 \right\rangle + \left\langle \bar{\Phi}'_{\widehat{K}'}(\bar{\mu}), \chi - 1 \right\rangle \right) \right), \quad (\text{IV.3})$$

where $\bar{\mu}$ is a normalized measure.

Proof. From its very construction above \widehat{K} follows $\Phi_K(\mu)$:

$$\mathbb{P}_\mu \left(\widehat{K} = k \right) = \Phi_K(\mu)(k). \quad (\text{IV.4})$$

And, as the bucket \widehat{K}' is drawn following $\Phi_{K'}(\mu')$ where $\mu' = \sum_{\ell \in [1, \widehat{K}]} \delta_{a(\widehat{B}_\ell)}$ the size \widehat{K}' of that bucket follows $\Phi_{K'}$ applied to a K sized sample of the size-biased distribution associated to μ . Hence,

$$\Phi'_{K'}(\mu)(k') = \mathbb{E}_\mu \mathbb{P}_{\mu'} \left(\widehat{K}' = k' \right) \quad (\text{IV.5})$$

Likewise, all r.v.'s $a(\widehat{B}_\ell)$, $\ell \in \llbracket 1, \widehat{K} - 1 \rrbracket$ and $a(\widehat{B}'_\ell)$, $\ell \in \llbracket 1, \widehat{K}' - 1 \rrbracket$, follow the size-biased distribution associated to μ : conditionally on $\widehat{K} = k$ and $\widehat{K}' = k' > 0$, for any $l \in \llbracket 1, k \rrbracket$ and any $\ell \in \llbracket 1, k' - 1 \rrbracket$ we clearly have that

$$\begin{aligned} \mathbb{P}_\mu \left(a(\widehat{B}_l) = x \right) &= \frac{x\mu(x)}{\langle \mu, \chi \rangle}, \quad x \in \mathbb{N}; \\ \mathbb{P}_\mu \left(a(\widehat{B}'_\ell) = y \right) &= \frac{y\mu(y)}{\langle \mu, \chi \rangle}, \quad y \in \mathbb{N}. \end{aligned} \tag{IV.6}$$

Giving us the desired expression:

$$\begin{aligned} &\mathbb{E}_\mu \left[\langle \widehat{\vartheta}(\mu), f \rangle \right] \\ &= - \sum_{k=1}^{+\infty} \left\{ f(k) \Phi_{\widehat{K}}(\mu)(k) + (k-1) \Phi_{\widehat{K}}(\mu) \sum_{y=1}^{+\infty} \frac{y\mu(y)}{\langle \mu, \chi \rangle} (f(y) - f(y-1)) \right\} \\ &\quad - \sum_{k'=0}^{+\infty} \left\{ f(k') \Phi'_{\widehat{K}'}(\mu)(k') + (k'-1) \Phi'_{\widehat{K}'}(\mu) \sum_{y=1}^{+\infty} \frac{y\mu(y)}{\langle \mu, \chi \rangle} (f(y) - f(y-1)) \right\} \\ &= - \left\{ \langle \Phi_{\widehat{K}}(\mu), f \rangle + \langle \Phi'_{\widehat{K}'}(\mu), f \rangle + \frac{\langle \mu, \chi \nabla f \rangle}{\langle \mu, \chi \rangle} (\langle \Phi_{\widehat{K}}(\mu), \chi \rangle + \langle \Phi'_{\widehat{K}'}(\mu), \chi \rangle) \right\} \\ &= \mathfrak{L}\Pi_f(\mu). \end{aligned}$$

■

Remark IV.1.1. *We draw attention to the fact the alternative construction allows us to essentially make every choice **with replacement** this time. Hence, every choice function is directly applied either to the base measure or to its perfectly size biased version. Later, we will show that this alternative dynamic stays arbitrarily close to the 'real' dynamic when the graphs get large. For now, it is used to state the main theorem.*

But before that, one last definition is needed.

Definition IV.1.2. *Let \mathcal{M}_Φ a (closed) subspace of \mathcal{M} , \mathfrak{L} a linear operator on $\mathcal{C}_b(\mathcal{M})$, and $\mu : [0, 1] \rightarrow \mathcal{M}$. The operator \mathfrak{L} **characterizes** μ on \mathcal{M}_Φ if, for all $t \in [0, 1]$, $f \in \mathcal{C}_b$ and η such that*

$$\langle \eta_t, f \rangle = \langle \mu_0, f \rangle + \int_0^t \mathfrak{L}\Pi_f(\eta_s) ds, \tag{IV.7}$$

we have $\eta_t = \mu_t$.

The preceding definition is a placeholder for the fact that the limiting system, characterized by the operator \mathfrak{L} , has a unique solution. Oftentimes, a Cauchy-Lipschitz type

condition on the operator \mathfrak{L} is used (See for example Kurtz [1970]). Uniqueness of the limiting process in our case is harder to identify, thus we have to enclose it in a definition and allow it to become criteria dependent.

IV.1.1 Convergence of local matching algorithms

The main convergence result is finally stated.

Theorem IV.1.1. *Assume that $\bar{\Phi} := (\bar{\Phi}_{\mathbf{K}}(\cdot), \bar{\Phi}'_{\mathbf{K}'}(\cdot))^2$ is a well behaved and moment preserving matching criterion that is defined on a subspace \mathcal{M}_{Φ} of \mathcal{M} . Assume that $(\bar{\Phi}_{\hat{\mathbf{K}}}(\cdot), \bar{\Phi}'_{\hat{\mathbf{K}}'}(\cdot))$ is the alternative dynamic associated to $\bar{\Phi}$. Recall the operator*

$$\begin{aligned} \bar{\mathfrak{L}}\Pi_f(\bar{\mu}) := & - \left(\langle \bar{\Phi}_{\hat{\mathbf{K}}}(\bar{\mu}), f \rangle + \langle \bar{\Phi}'_{\hat{\mathbf{K}}'}(\bar{\mu}), f \rangle \right. \\ & \left. + \frac{\langle \bar{\mu}, \chi \nabla f \rangle}{\langle \bar{\mu}, \chi \rangle} (\langle \bar{\Phi}_{\hat{\mathbf{K}}}(\bar{\mu}), \chi - 1 \rangle + \langle \bar{\Phi}'_{\hat{\mathbf{K}}'}(\bar{\mu}), \chi - 1 \rangle) \right). \end{aligned} \quad (\text{IV.8})$$

If the sequence of processes $(\bar{\mu}^n)_{n \geq 1}$ is under $\mathfrak{H}\mathfrak{o}$, and $\bar{\mathfrak{L}}$ characterizes a measure valued function $\bar{\mu}$, then for all $f \in \mathcal{C}_b$ we have the convergence in probability

$$\sup_{t \leq 1} |\langle \bar{\mu}_t^n, f \rangle - \langle \bar{\mu}_t, f \rangle| \xrightarrow{(n, \mathbb{P})} 0. \quad (\text{IV.9})$$

Corollary IV.1.1 (Convergence of the matching coverage). *Under the above assumptions, for $\bar{\mu}_0^n \rightarrow \nu \in \mathcal{M}_{\Phi}$, we get*

$$\mathbf{M}_{\Phi}^n(\bar{\mu}_0^n) \xrightarrow{(n, \mathbb{P})} \mathbf{M}_{\Phi}(\nu) := 1 - \bar{\mu}_1(\{0\}). \quad (\text{IV.10})$$

Proof. Equation (III.25) provides the matching coverage as a function of the normalized matching process :

$$\mathbf{M}_{\Phi}^n(\mu_0^n) = 1 - \frac{\mathcal{I}_n^n}{n} = 1 - \bar{\mu}_1^n(\{0\}).$$

The result is obtained by merely taking the limit in n . ■

The main convergence result can be viewed as a law of large numbers for the matching coverage. At worst, it is an asymptotic result specific to the configuration model restricted to some constrained degree distributions. However, the constraints make it so

²Recall that K and K' are the degrees of the vertices picked as extremities of the edge added to the matching

the configuration model, and the associated algorithms produce matching coverages that are asymptotically equal to that of simple graphs. Hence, we interpret the main result as *a prediction of the matching coverage produced by a criterion on a graph drawn uniformly following a given degree distribution.*

IV.1.2 Further simulations

We finally complement the simulations given in Chapter II. In this section, we go back to 3-regular graphs, this time using the measure valued representation for the degree sequences. In this case, we readily obtain the degree distributions :

$$\bar{\mu}_0^+ = \delta_3 \quad \text{and} \quad \bar{\mu}_0^- = \delta_3. \tag{IV.11}$$

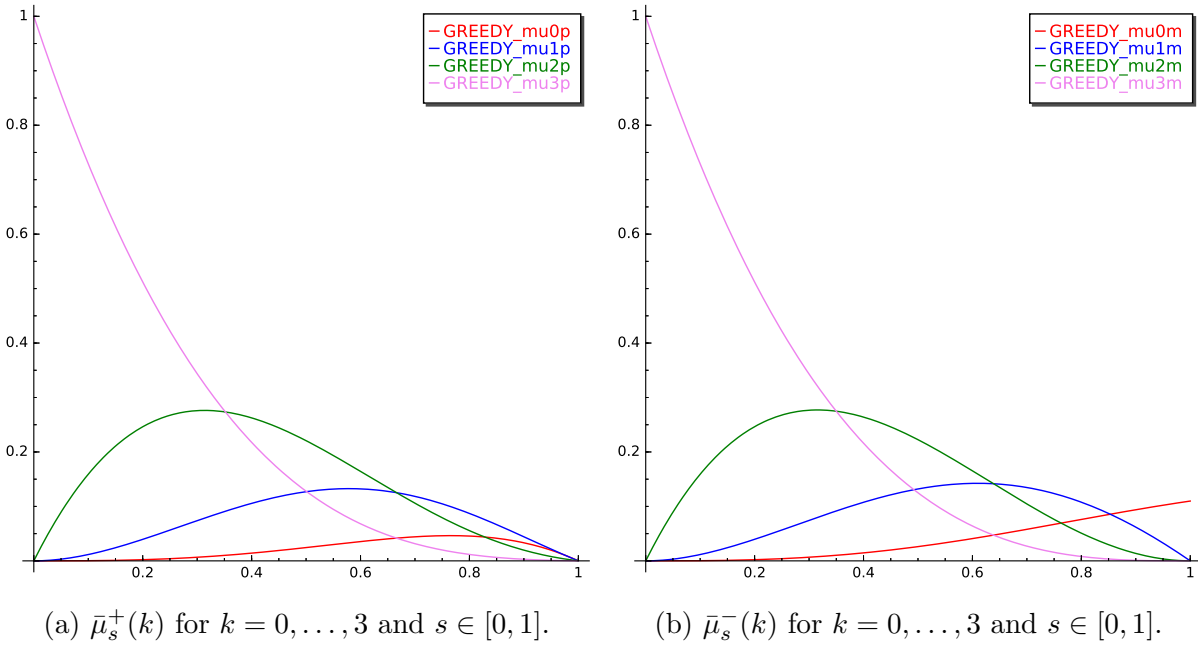


Figure IV.1: GREEDY-EDO results for $\bar{\mu}_0^+ = \bar{\mu}_0^- = \delta_3$

By a numerical resolution, we predict the final value $\bar{\mu}_1^-(0) = 0.1098$ (which is the final value of the red curve in Figure IV.1). From this, we deduce the approximate matching coverage which is a very similar to the one obtained for the greedy algorithm in Noiry et al. [2021].

$$\bar{\mathbf{M}}_{\text{GREEDY}}(\delta_3, \delta_3) = 1 - \bar{\mu}_1^-(0) = 0.8902. \tag{IV.12}$$

Regarding MINRES, similarly to GREEDY we deduce the approximate matching cover-

age

$$\bar{\mathbf{M}}_{\text{MINRES}}(\delta_3, \delta_3) = 1 - \bar{\mu}_1^-(0) = 0.9378. \tag{IV.13}$$

Again, we compile the evolution of the simulated matching coverage as the graph size n grows but this time we also add the matching coverage predicted by (IV.10).

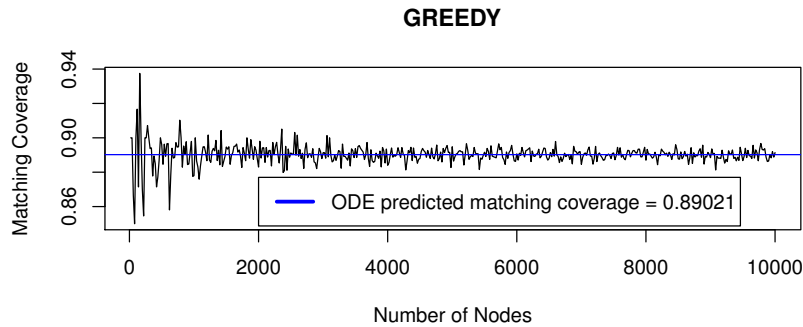


Figure IV.2: Matching coverage of GREEDY as the graph size tends to ∞

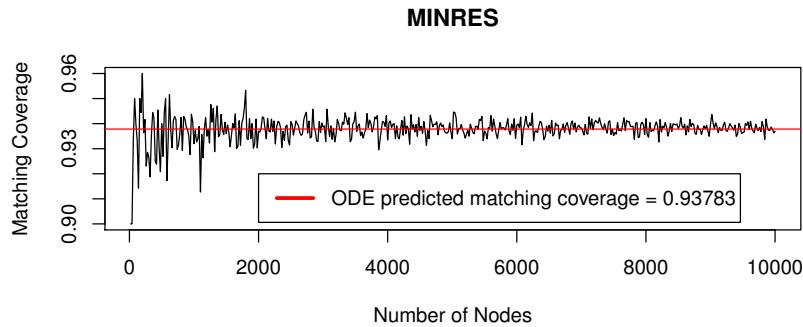


Figure IV.3: Matching coverage of MINRES as the graph size tends to ∞

As the figures show, the prediction given by the ODE is fairly accurate. In fact, the exact numbers are more striking when we update table (II.1).

	Graph Size n	200	500	1000	3000	5000	$\bar{\mathbf{M}}_{\Phi}(\delta_3, \delta_3)$
$\tilde{\mathbf{M}}_{\text{GREEDY}}^n(G)$	Mean	0.8904	0.8916	0.8911	0.8897	0.8898	0.8902
	Std Dev	0.0198	0.0109	0.009	0.0041	0.00311	
$\tilde{\mathbf{M}}_{\text{MINRES}}^n(G)$	Mean	0.9356	0.9365	0.9396	0.9378	0.9385	0.9378
	Std Dev	0.0148	0.0096	0.0052	0.0040	0.0025	

Table IV.1: Recap. of Average Matching Size updated with predictions

Thus, having heuristically confirmed the predictions of the main theorem, we dedicate

the next section to the formal proof of the main theorem.

IV.2 Proof of the main convergence theorem

The proof of the main theorem is a typical proof of convergence for a Markov process: it is shown that the law of the sequence $\bar{\mu}_t^n$ is tight (thus relatively compact) in the Skorokhod topology, then it is shown that all the limits are in fact characterized by the given operator $\bar{\mathfrak{L}}$, and finally it follows that since the ODE resulting from $\bar{\mathfrak{L}}$ identifies a unique function $\bar{\mu}_t$, the sequence of processes converges to $\bar{\mu}_t$. In summary, for any $t \leq 1$ we will establish bounds on each of the terms of

$$\begin{aligned} |\langle \bar{\mu}_t^n, f \rangle - \langle \bar{\mu}_t, f \rangle| &= |\langle \bar{\mu}_t^n, f \rangle - \langle \tilde{\mu}_t^n, f \rangle + \langle \tilde{\mu}_t^n, f \rangle - \langle \bar{\mu}_t, f \rangle| \\ &\leq |\langle \bar{\mu}_t^n, f \rangle - \langle \tilde{\mu}_t^n, f \rangle| + \left| \langle \bar{\mu}_0^n, f \rangle + \mathcal{M}_t^{f,n} + \int_0^t \tilde{\mathfrak{L}}^n \Pi_f(\bar{\mu}_s^n) ds - \langle \bar{\mu}_0, f \rangle - \int_0^t \bar{\mathfrak{L}} \Pi_f(\bar{\mu}_s) ds \right| \\ &\leq |\langle \bar{\mu}_t^n, f \rangle - \langle \tilde{\mu}_t^n, f \rangle| + |\langle \bar{\mu}_0^n, f \rangle - \langle \bar{\mu}_0, f \rangle| + |\mathcal{M}_t^{f,n}| \\ &\quad + \left| \int_0^t \tilde{\mathfrak{L}}^n \Pi_f(\tilde{\mu}_s^n) - \int_0^t \bar{\mathfrak{L}} \Pi_f(\bar{\mu}_s) ds \right|, \end{aligned} \quad (\text{IV.14})$$

where we used the fact that

$$\langle \bar{\mu}_t^n, f \rangle = \langle \bar{\mu}_0^n, f \rangle + \mathcal{M}_t^{f,n} + \int_0^t \tilde{\mathfrak{L}}^n \Pi_f(\bar{\mu}_s^n), \quad (\text{IV.15})$$

and that $\bar{\mu}_t$ is a solution of

$$\langle \bar{\mu}_t, f \rangle = \langle \bar{\mu}_0, f \rangle + \int_0^t \bar{\mathfrak{L}} \Pi_f(\bar{\mu}_s) ds, \quad 0 \leq t \leq 1. \quad (\text{IV.16})$$

The convergence is established if the right-hand term of (IV.14) goes to 0.

Proof. The first step of the proof is to establish the existence of sub-sequential limits for the sequence of processes.

Part 1: Tightness

First, the relative compactness of the laws of the processes $(\tilde{\mu}_t^n)_{n \geq 1}$ is proven in $\mathbb{D}([0, 1], (\mathcal{M}, \tau_v))$, where τ_v denotes here the **vague** topology. The relative compactness is simpler to prove in the vague topology. Recalling that the *tilde* process $\tilde{\mu}_t^n$ is the continuous time Markov approximation (with exponential inbetween jumps), we use **Prohorov's theorem** (Parthasarathy [1967]; Billingsley [1999]) which equates the relative compactness of $(\tilde{\mu}_t^n)_{n \geq 1}$ with its tightness. Thus, we have to prove the tightness,

Lemma IV.2.1. *Under the conditions of IV.1.1, the sequence $(\tilde{\mu}^n)_{n \geq 1}$ is tight in $\mathbb{D}([0, 1], (\mathcal{M}, \tau_v))$.*

Proof. Méleard and Roelly (Méleard and Roelly [1993]) have shown that for $(\tilde{\mu}^n)_{n \geq 1}$ to be tight in $\mathbb{D}([0, 1], (\mathcal{M}, \tau_v))$, it is enough that, for any $f \in \mathcal{C}_c$, the sequence of processes $(\langle \tilde{\mu}^n, f \rangle)_{n \geq 1}$ is tight on $\mathbb{D}([0, 1], \mathbb{R})$. Roelly's tightness criterion (Roelly-Coppoletta [1986]) is used to prove the latter:

1. Fix $t \geq 0$, for all n we write $|\langle \tilde{\mu}_t^n, f \rangle| \leq \|f\| |\langle \tilde{\mu}_t^n, 1 \rangle| \leq \|f\|$. Hence, $(\langle \tilde{\mu}_t^n, f \rangle)_{n \geq 1}$ is a bounded sequence on \mathbb{R} , which means that it is tight (since it is relatively compact).
2. Fix $n \geq 1$. Recalling (III.34), for $t \in [0, 1]$, $\langle \tilde{\mu}_t^n, f \rangle$ is rewritten as

$$\langle \tilde{\mu}_t^n, f \rangle = \mathcal{M}_t^{f,n} + V_t^{f,n}.$$

Let $\delta > 0$ and two stopping times S_n and T_n such that $S_n \leq T_n \leq S_n + \delta$. On the one hand, for $\epsilon > 0$ using the Markov inequality we have

$$\mathbb{P}\left(|\langle \mathcal{M}^{f,n} \rangle_{T_n} - \langle \mathcal{M}^{f,n} \rangle_{S_n}| > \epsilon\right) \leq \frac{1}{\epsilon} \mathbb{E}\left(|\langle \mathcal{M}^{f,n} \rangle_{T_n} - \langle \mathcal{M}^{f,n} \rangle_{S_n}|\right).$$

Remembering the definition IV.34, we write

$$\begin{aligned} |\langle \mathcal{M}^{f,n} \rangle_{T_n} - \langle \mathcal{M}^{f,n} \rangle_{S_n}| &= \left| \frac{1}{n} \int_{S_n}^{T_n} \mathbb{E}_{\tilde{\mu}_s^n} (\langle \vartheta(\tilde{\mu}_s^n), f \rangle^2) ds \right| \\ &= \left| \frac{1}{n} \int_{S_n}^{T_n} \mathbb{E}_{\tilde{\mu}_s^n} \left(\left\{ f(K) + f(K') + \sum_{I'_i \in \mathcal{N}(I)} \left(f(a_k(I'_i)) - f(a_k(I'_i) - D_{I'_i}) \right) \right. \right. \right. \\ &\quad \left. \left. \left. + \sum_{I_i \in \mathcal{N}(I')} \left(f(a_k(I_i)) - f(a_k(I_i) - F_{I_i}) \right) \right\}^2 \right) ds \right| \\ &\leq \frac{\delta}{n} (\|f\| (2 + 4 \times o(\sqrt{n}))^2). \quad (\text{IV.17}) \end{aligned}$$

But, under $\mathfrak{H}\mathfrak{o}$ (III.21) is verified, and since $\max \mathfrak{d}^n = o(\sqrt{n})$ and $T_n - S_n \leq \delta$, it follows that

$$\mathbb{P}\left(|\langle \mathcal{M}^{f,n} \rangle_{T_n} - \langle \mathcal{M}^{f,n} \rangle_{S_n}| > \epsilon\right) \xrightarrow{n} 0. \quad (\text{IV.18})$$

On the other hand, since Φ preserves the first moments and we are under $\mathfrak{H}\mathfrak{o}$,

(III.27) yields that both $\mathbb{E}(K) < M$ and $\mathbb{E}(K') < M$. We follow with

$$|V_{T_n}^{f,n} - V_{S_n}^{f,n}| = \left| \int_{S_n}^{T_n} \mathbb{E}_{\tilde{\mu}_s^n}(\langle \vartheta \tilde{\mu}_s^n, f \rangle) ds \right| \leq \delta \|f\| (2 + 4M). \quad (\text{IV.19})$$

Thus,

$$\mathbb{P}\left(|V_{T_n}^{f,n} - V_{S_n}^{f,n}| > \epsilon\right) \leq \frac{\delta}{\epsilon} \|f\| (2 + 4M), \quad (\text{IV.20})$$

proving the tightness in $\mathbb{D}([0, 1], (\mathcal{M}, \tau_v))$. ■

As said at the start of this part, the vague convergence is not enough. Noticing that the processes characterized by $\bar{\mathfrak{L}}$ are continuous in time and that continuity of the limiting processes is required, we need a further characterization of the eventual sub-sequential limits.

Lemma IV.2.2. *The sequence $(\langle \tilde{\mu}^n, f \rangle)_{n \geq 1}$ is C-tight³.*

Proof. For all n and $t \in [0, 1]$, let $(\Delta \langle \tilde{\mu}_t^n, f \rangle)$ be the jump of $\langle \tilde{\mu}_t^n, f \rangle$ at t . $\Delta \langle \tilde{\mu}_t^n, f \rangle$ is either 0 or $\frac{1}{n} \langle \vartheta(n\tilde{\mu}_t^n), f \rangle$. But

$$\frac{1}{n} |\langle \vartheta(n\tilde{\mu}_t^n), f \rangle| \leq \frac{1}{n} \|f\| (2 + 4o(\sqrt{n})). \quad (\text{IV.21})$$

This goes to show that, $\sup_t \Delta \langle \tilde{\mu}_t^n, f \rangle \xrightarrow{n} 0$, proving that limits are indeed continuous. ■

Hence, we can finally write the tightness in $\mathbb{D}([0, 1], (\mathcal{M}, \tau_w))$ where τ_w denotes the weak topology. We stress that the weak tightness is needed since the matching coverage depends on $\langle \bar{\mu}, \mathbb{1}_{\{0\}} \rangle$ and whence $\mathbb{1}_{\{0\}}$ is a compactly supported function, the associated mass is $\langle \bar{\mu}_t, \mathbb{1}_{\{0\}} \rangle = \langle \bar{\mu}_t, \mathbb{1} \rangle - \langle \bar{\mu}_t, \mathbb{1}_{\{\mathbb{R}_*^+\}} \rangle$ and it depends on the non compactly supported function $\mathbb{1}_{\{\mathbb{R}_*^+\}}$.

Corollary IV.2.1. *The sequence of processes $(\tilde{\mu}^n)_{n \geq 1}$ is tight in $\mathbb{D}([0, 1], (\mathcal{M}, \tau_w))$.*

Proof. Following (Méléard and Roelly [1993]), $(\tilde{\mu}^n)_{n \geq 1}$ is weakly tight if it is vaguely C-tight and the mass process $(\langle \tilde{\mu}^n, 1 \rangle)_{n \geq 1}$ is tight in $\mathbb{D}([0, 1], \mathbb{R})$.

We have yet to prove the tightness of $(\langle \tilde{\mu}^n, 1 \rangle)_{n \geq 1}$ in $\mathbb{D}([0, 1], \mathbb{R})$. But we can notice that the proof is exactly the same as Lemma IV.2.1 for $f \equiv 1$. ■

³C-tight means the sub-sequential limits are continuous in time

So far, we have proven that the sequence of normalized processes admit subsequential limits. We also know that the eventual limits are continuous. Since the processes themselves are described (following Doob) by the equation

$$\langle \bar{\mu}_t^n, f \rangle = \langle \bar{\mu}_0^n, f \rangle + \mathcal{M}_t^{f,n} + \int_0^t \tilde{\mathfrak{L}}^n \Pi_f(\bar{\mu}_s^n),$$

for them to be deterministic, we need the martingale part of the description to cancel itself.

Part 2: Cancellation of the martingale part

It is straightforward to write that for all n and t ,

$$\mathbb{E}\left((\mathcal{M}_t^{f,n})^2\right) = \mathbb{E}\langle\langle \mathcal{M}^{f,n} \rangle\rangle_t \leq \frac{t}{n} (\|f\| (2 + 4 \times o(\sqrt{n})))^2.$$

And, since the rightmost term goes to 0, Doob's martingale inequality then yields that

$$\mathbb{P}(\sup_{t \leq 1} |\mathcal{M}_t^{f,n}| > \epsilon) \leq \frac{1}{\epsilon^2} \frac{t}{n} (\|f\| (2 + 4 \times o(\sqrt{n})))^2$$

goes again to 0. Thus $\mathcal{M}_t^{f,n}$ converges in probability to 0.

Hence, the convergence in probability implies following equation (IV.15) that when n gets large enough, the behavior of the dynamic is close to that of $\langle \bar{\mu}_0^n, f \rangle + \int_0^t \tilde{\mathfrak{L}}^n \Pi_f(\bar{\mu}_s^n)$.

Going forward, the sequence of processes $\bar{\mu}_t^n$ have been shown to have continuous limits and we have just established that those limits are described by a deterministic equation of the *tilde* operators. But, the $\tilde{\mathfrak{L}}^n$ operators do not have an explicit form and, as was shown in section III.3.1 they are not guaranteed to converge themselves. *Henceforth, the goal of the next part of the proof to set up $\bar{\mathfrak{L}}$ operator as the limit of the tilde operators. We prove that well-behaved, moment preserving matching criteria induce convergent operators.*

The identification of the limit is done through a coupling with the alternative dynamic which induces $\bar{\mathfrak{L}}$ as a generator.

Part 3: Identifying the Limit

Remark IV.2.1. *Throughout this section and all the thesis, it is important to remember that the bars $\bar{\cdot}$ are associated with normalized objects and that tildes $\tilde{\cdot}$ are for the continuous time Markov Chain. When there is no accent, we are working directly with counting measures. For example, $\mathfrak{L}\Pi_f(\mu)$ is a generator associated to counting measures and is \mathcal{M}_F valued, while $\tilde{\mathfrak{L}}^n \Pi_f(\bar{\mu}_t^n)$ is the generator of the continuous Markov extension and is thus \mathcal{M}_n valued and $\bar{\mathfrak{L}}\Pi_f(\bar{\mu}_t^n)$ is the limiting generator of both the normalized process and its Markov extension. Technically, it is associated with the alternative construction but since it is the limit, we chose not to add the hat superscript.*

We quantify the error when replacing the real dynamic by the alternative one.

First, let us define the following event in the alternate construction,

$$\widehat{T}(\mu) = \left\{ \text{The buckets } \widehat{I}, \widehat{I}', \widehat{B}_1, \dots, \widehat{B}_{\widehat{K}-1}, \widehat{B}'_1, \dots, \widehat{B}'_{\widehat{K}'-1} \text{ are all distinct} \right\}.$$

The event \widehat{T} simply means the local neighborhood under study is **tree-like**. We deduce the following result :

Lemma IV.2.3. *For all $\mu \in \mathcal{M}_F(\mathbb{N})$ having finite third moment, we have that*

$$\mathbb{P}_\mu \left(\widehat{T}(\mu) \right) \geq 1 - \frac{\langle \mu, \chi^2 \rangle}{\langle \mu, \chi \rangle^2} \left(\langle \Phi_K(\mu), \chi^2 \rangle + 2 \langle \Phi_K(\mu), \chi \rangle + 2 \langle \Phi_{K'}(\mu), \chi^2 \rangle \right). \quad (\text{IV.22})$$

Proof. The event $\widehat{T}(\mu)^c$ is included in

$$\begin{aligned} & \left(\widehat{B}_j = \widehat{B}_l \text{ for some } j \neq l \in \llbracket 1, \widehat{K} \rrbracket \right) \cup \left(\widehat{B}'_j = \widehat{B}'_l \text{ for some } j \neq l \in \llbracket 1, \widehat{K}' - 1 \rrbracket \right) \\ & \cup \left(\widehat{B}'_j \in \{ \widehat{I} \} \cup \{ \widehat{B}_1, \dots, \widehat{B}_{\widehat{K}-1} \} \text{ for some } j \in \llbracket 1, \widehat{K}' - 1 \rrbracket \right) \\ & \cup \left(\widehat{B}'_j \in \{ \widehat{I}' \} \cup \{ \widehat{B}'_1, \dots, \widehat{B}'_{\widehat{K}'-1} \} \text{ for some } j \in \llbracket 1, \widehat{K}' - 1 \rrbracket \right). \end{aligned}$$

Consequently,

$$\begin{aligned} & \mathbb{P}_\mu \left(\widehat{T}(\mu)^c \right) \\ & \leq \mathbb{P}_\mu \left(\widehat{B}_j = \widehat{B}_l \text{ for some } j \neq l \in \llbracket 1, \widehat{K} \rrbracket \right) + \mathbb{P}_\mu \left(\widehat{B}'_j = \widehat{B}'_l \text{ for some } j \neq l \in \llbracket 1, \widehat{K}' - 1 \rrbracket \right) \\ & \quad + \mathbb{P}_\mu \left(\widehat{B}'_j \in \{ \widehat{I} \} \cup \{ \widehat{B}_1, \dots, \widehat{B}_{\widehat{K}-1} \} \text{ for some } j \in \llbracket 1, \widehat{K}' - 1 \rrbracket \right) \\ & = \sum_{k=2}^{\infty} \sum_{j \neq l \in \llbracket 1, k \rrbracket} \mathbb{P}_\mu \left(\widehat{B}_j = \widehat{B}_l \mid \widehat{K} = k \right) \mathbb{P}_\mu \left(\widehat{K} = k \right) \\ & \quad + \sum_{k'=3}^{\infty} \sum_{j \neq l \in \llbracket 1, k'-1 \rrbracket} \mathbb{P}_\mu \left(\widehat{B}'_j = \widehat{B}'_l \mid \widehat{K}' = k' \right) \mathbb{P}_\mu \left(\widehat{K}' = k' \right) \\ & + \sum_{k=2}^{\infty} \sum_{k'=2}^{\infty} \sum_{j=1}^{k'-1} \mathbb{P}_\mu \left(\widehat{B}'_j \in \{ \widehat{I} \} \cup \{ \widehat{B}_1, \dots, \widehat{B}_{\widehat{K}-1} \} \mid \widehat{K} = k, \widehat{K}' = k' \right) \mathbb{P}_\mu \left(\widehat{K} = k, \widehat{K}' = k' \right). \end{aligned}$$

Because the \widehat{B}_j 's have the same law, we deduce that

$$\begin{aligned} \mathbb{P}_\mu \left(\widehat{T}(\mu)^c \right) &\leq \sum_{k=2}^{\infty} \binom{k}{2} \mathbb{P}_\mu \left(\widehat{B}_1 = \widehat{B}_2 \mid \widehat{K} = k \right) \Phi_K(\mu)(k) \\ &\quad + \sum_{k'=3}^{\infty} \binom{k'-1}{2} \mathbb{P}_\mu \left(\widehat{B}'_1 = \widehat{B}'_2 \mid \widehat{K}' = k' \right) \Phi'_{K'}(\mu)(k') \\ &+ \sum_{k=2}^{\infty} \sum_{k'=2}^{\infty} (k'-1) \mathbb{P}_\mu \left(\widehat{B}'_1 \in \{\widehat{I}\} \cup \{\widehat{B}_1, \dots, \widehat{B}_{\widehat{K}-1}\} \mid \widehat{K} = k, \widehat{K}' = k' \right) \Phi_K(\mu)(k) \Phi'_{K'}(\mu)(k'). \end{aligned} \tag{IV.23}$$

But, for the first term, for all $k \geq 2$ conditioning on the size of B_1 , it follows that

$$\begin{aligned} \mathbb{P}_\mu \left(\widehat{B}_1 = \widehat{B}_2 \mid \widehat{K} = k \right) &= \sum_{l=1}^{\infty} \mathbb{P}_\mu \left(\widehat{B}_1 = \widehat{B}_2 \mid \widehat{K} = k, a(\widehat{B}_1) = l \right) \mathbb{P}_\mu \left(a(\widehat{B}_1) = l \mid \widehat{K} = k \right) \\ &= \sum_{l=1}^{\infty} \frac{l}{\langle \mu, \chi \rangle} \frac{l\mu(l)}{\langle \mu, \chi \rangle} = \frac{\langle \mu, \chi^2 \rangle}{\langle \mu, \chi \rangle^2} \end{aligned} \tag{IV.24}$$

Likewise, for the second term, for all $k' \geq 3$ we obtain that

$$\mathbb{P}_\mu \left(\widehat{B}'_1 = \widehat{B}'_2 \mid \widehat{K}' = k' \right) = \frac{\langle \mu, \chi^2 \rangle}{\langle \mu, \chi \rangle^2}. \tag{IV.25}$$

For the final term, observe that

$$\widehat{D} = \sum_{i=1}^{\widehat{K}-1} a(\widehat{B}_i).$$

Then, for all $k \geq 2, k' \geq 2$ and $d \geq 1$ we get that

$$\begin{aligned} &\mathbb{P}_\mu \left(\widehat{B}'_1 \in \{\widehat{I}\} \cup \{\widehat{B}_1, \dots, \widehat{B}_{\widehat{K}-1}\} \mid \widehat{K} = k, \widehat{K}' = k' \right) \\ &= \sum_{d=0}^{\infty} \mathbb{P}_\mu \left(\widehat{B}'_1 \in \{\widehat{I}\} \cup \{\widehat{B}_1, \dots, \widehat{B}_{\widehat{K}-1}\} \mid \widehat{K} = k, \widehat{K}' = k', \widehat{D} = d \right) \\ &\quad \times \mathbb{P}_\mu \left(\widehat{D} = d \mid \widehat{K} = k, \widehat{K}' = k' \right) \\ &\leq \sum_{d=0}^{\infty} \frac{k+k'+d}{\langle \mu, \chi \rangle} \mathbb{P}_\mu \left(\widehat{D} = d \mid \widehat{K} = k, \widehat{K}' = k' \right) \\ &= \frac{k+k'+\mathbb{E}_\mu \left[\widehat{D} \mid \widehat{K} = k, \widehat{K}' = k' \right]}{\langle \mu, \chi \rangle} = \frac{k+k'+(k-1)\frac{\langle \mu, \chi^2 \rangle}{\langle \mu, \chi \rangle}}{\langle \mu, \chi \rangle}, \end{aligned}$$

because \widehat{T} has \widehat{K} neighbors again following the size-biased distribution associated to μ . This, combined with (IV.24) and (IV.25) in (IV.23), yields that

$$\begin{aligned} \mathbb{P}_\mu(\widehat{T}^c) &\leq \frac{\langle \mu, \chi^2 \rangle}{\langle \mu, \chi \rangle^2} \langle \Phi_K(\mu), \chi^2 \rangle + \frac{\langle \mu, \chi^2 \rangle}{\langle \mu, \chi \rangle^2} \langle \Phi_{K'}(\mu), \chi^2 \rangle \\ &\quad + \sum_{k=2}^{\infty} \sum_{k'=2}^{\infty} \frac{k + k' + k \frac{\langle \mu, \chi^2 \rangle}{\langle \mu, \chi \rangle}}{\langle \mu, \chi \rangle} \Phi_K(\mu)(k) (k' \Phi_{K'}(\mu)(k')). \end{aligned}$$

Ultimately noticing that $\frac{\langle \mu, \chi^2 \rangle}{\langle \mu, \chi \rangle} \geq 1$, we obtain that

$$\mathbb{P}_\mu(\widehat{T}^c) \leq \frac{\langle \mu, \chi^2 \rangle}{\langle \mu, \chi \rangle^2} (\langle \Phi_K(\mu), \chi^2 \rangle + 2 \langle \Phi_K(\mu), \chi \rangle + 2 \langle \Phi_{K'}(\mu), \chi^2 \rangle).$$

■

We follow with the intermediary approximation of the generator:

Proposition IV.2.1. *Let μ in \mathcal{M}_F^n , the following is true for some constant $C > 0$:*

$$|\mathfrak{L}\Pi_f(\mu) - \mathfrak{L}^n\Pi_f(\mu)| \leq C\|f\| \frac{\langle \mu, \chi^2 \rangle}{\langle \mu, \chi \rangle^2} (\langle \Phi_K(\mu), \chi^2 \rangle + \langle \Phi_K(\mu), \chi \rangle + \langle \Phi_{K'}(\mu), \chi^2 \rangle).$$

Proof. Fix a measure μ , and denote the event

$$T(\mu) = \left\{ I \text{ and } I' \text{ do not have any self-loops/multiple edges,} \right. \\ \left. \text{given that the degree measure is } \mu \right\}.$$

Throughout this proof, for notational simplicity let us skip the dependence in μ of ϑ , $\widehat{\vartheta}$, T and \widehat{T} . First, as uniform sampling with replacement and conditioned on not drawing twice the same element amounts in law to a uniform sampling without replacement, we get that for any point measure ν ,

$$\mathbb{P}_\mu(\widehat{\vartheta} = \nu | \widehat{T}) = \mathbb{P}_\mu(\vartheta = \nu | T). \tag{IV.26}$$

Also, it is immediate to observe that self-loops and multiple edges occur with a larger probability if draws of half-edges are performed with replacement, with respect to draws

without replacement. Thus we get that $\mathbb{P}_\mu(T) \geq \mathbb{P}_\mu(\widehat{T})$, and set

$$q_\mu := \frac{\mathbb{P}_\mu(T) - \mathbb{P}_\mu(\widehat{T})}{1 - \mathbb{P}_\mu(\widehat{T})} \in (0, 1).$$

Let the $\mathcal{M}_F(\mathbb{N})$ -valued r.v. $\check{\vartheta}$ be drawn from the distribution $\mathbb{P}_\mu(\vartheta = . | T)$ with probability q_μ , and independently, from the distribution $\mathbb{P}_\mu(\vartheta = . | T^c)$ with probability $1 - q_\mu$, and let us set

$$\tilde{\vartheta} = \widehat{\vartheta} \mathbb{1}_{\widehat{T}} + \check{\vartheta} \mathbb{1}_{\widehat{T}^c}.$$

So defined, $\tilde{\vartheta}$ is a $\mathcal{M}_F(\mathbb{N})$ -valued r.v. that coincides with $\widehat{\vartheta}$ on \widehat{T} , and that has the same distribution as ϑ . To see this, observe that for all $v \in \mathcal{M}_p(\mathbb{N})$,

$$\begin{aligned} \mathbb{P}_\mu(\tilde{\vartheta} = v) &= P_\mu(\tilde{\vartheta} = v | \widehat{T}) \mathbb{P}_\mu(\widehat{T}) + P_\mu(\tilde{\vartheta} | \widehat{T}^c) \mathbb{P}_\mu(\widehat{T}^c) \\ &= \mathbb{P}_\mu(\widehat{\vartheta} = v | \widehat{T}) \mathbb{P}_\mu(\widehat{T}) + \left(\mathbb{P}_\mu(\vartheta = v | T) q_\mu + P_\mu(\vartheta = v | T^c) (1 - q_\mu) \right) \mathbb{P}_\mu(\widehat{T}^c) \\ &= \mathbb{P}_\mu(\vartheta = v | T) P_\mu(\widehat{T}) \\ &\quad + \mathbb{P}_\mu(\vartheta = v | T) \left(\mathbb{P}_\mu(T) - \mathbb{P}_\mu(\widehat{T}) \right) + \mathbb{P}_\mu(\vartheta = v | T^c) \mathbb{P}_\mu(T^c) \\ &= \mathbb{P}_\mu(\vartheta = v), \end{aligned}$$

where we used (IV.26) in the third equality. Therefore, as $\widehat{T} \subset \{\widehat{\vartheta} = \widehat{\vartheta}'\}$, we obtain from the definition of the generator and Lemma IV.1.1 that :

$$\begin{aligned} |\mathfrak{L}\Pi_f(\mu) - \mathfrak{L}^n\Pi_f(\mu)| &= \left| \mathbb{E}_\mu[\langle \vartheta, f \rangle] - \mathbb{E}_\mu[\langle \widehat{\vartheta}, f \rangle] \right| \\ &= \left| \mathbb{E}_\mu[\langle \tilde{\vartheta} - \widehat{\vartheta}, f \rangle \mathbb{1}_{\widehat{T}^c}] \right| \\ &\leq \mathbb{E}_\mu \left[\langle \tilde{\vartheta} - \widehat{\vartheta}, f \rangle^2 \right]^{1/2} \mathbb{P}_\mu(\widehat{T}^c)^{1/2} \\ &\leq \left(2\mathbb{E}_\mu[\langle \tilde{\vartheta}, f \rangle^2] + 2\mathbb{E}_\mu[\langle \widehat{\vartheta}, f \rangle^2] \right)^{1/2} \mathbb{P}_\mu(\widehat{T}^c)^{1/2} \\ &= \left(2\mathbb{E}_\mu[\langle \vartheta, f \rangle^2] + 2\mathbb{E}_\mu[\langle \widehat{\vartheta}, f \rangle^2] \right)^{1/2} \mathbb{P}_\mu(\widehat{T}^c)^{1/2}. \quad (\text{IV.27}) \end{aligned}$$

But, first, it follows from (IV.1) that

$$\begin{aligned} \mathbb{E}_\mu \left[\langle \widehat{\vartheta}, f \rangle^2 \right] &\leq \|f\|^2 \mathbb{E}_\mu \left[\langle \widehat{\vartheta}, \mathbf{1} \rangle^2 \right] = \|f\|^2 \mathbb{E}_\mu \left[\left(2 + 2\widehat{K} + 2\widehat{K}' \right)^2 \right] \\ &\leq 12 \|f\|^2 \left(1 + \mathbb{E}_\mu \left[\widehat{K}^2 \right] + \mathbb{E}_\mu \left[(\widehat{K}')^2 \right] \right). \end{aligned}$$

Also, by the same arguments we obtain that

$$\mathbb{E}_\mu \left[\langle \vartheta, f \rangle^2 \right] \leq 12 \|f\|^2 \left(1 + \mathbb{E}_\mu \left[K^2 \right] + \mathbb{E}_\mu \left[(K')^2 \right] \right)$$

which concludes the proof using (IV.27) and Lemma IV.2.3. \blacksquare

At last, this provides the large graph approximation,

Corollary IV.2.2. *Let $\tilde{\mu}^n$ a sequence of \mathcal{M}_n valued processes associated to a well behaved and moment preserving criterion Φ . Moreover, assume that its initial conditions are under $\mathfrak{H}\mathfrak{o}$. For any bounded f , and any $t \leq 1$, we have*

$$\left| \bar{\mathfrak{L}}\Pi_f(\bar{\mu}_t^n) - \tilde{\mathfrak{L}}^n\Pi_f(\tilde{\mu}_t^n) \right| = o\left(\frac{1}{\sqrt{n}}\right). \quad (\text{IV.28})$$

Proof. First, we rewrite the definitions of the scaled counterparts of \mathfrak{L}^n and of the matching criteria :

$$\tilde{\mathfrak{L}}^n\Pi_f(\tilde{\mu}_t^n) := \mathfrak{L}^n\Pi_f(n\tilde{\mu}_t^n) \qquad \bar{\Phi}(\bar{\mu}_t^n) := \Phi(n\bar{\mu}_t^n)$$

Proposition IV.2.1 yields that

$$\left| \bar{\mathfrak{L}}\Pi_f(\bar{\mu}_t^n) - \tilde{\mathfrak{L}}^n\Pi_f(\tilde{\mu}_t^n) \right| \leq C\|f\| \frac{\langle \bar{\mu}_t^n, \chi^2 \rangle}{\langle \bar{\mu}_t^n, \chi \rangle^2} \left(\langle \Phi_K(\bar{\mu}_t^n), \chi^2 \rangle + \langle \Phi_K(\bar{\mu}_t^n), \chi \rangle + \langle \Phi_{K'}(\bar{\mu}_t^n), \chi^2 \rangle \right).$$

Since the initial moments up to $p = 3$ are bounded, and the criterion is moment preserving, it follows easily from proposition III.3.1 that, for some $C' > 0$:

$$\left| \bar{\mathfrak{L}}\Pi_f(\bar{\mu}_t^n) - \tilde{\mathfrak{L}}^n\Pi_f(\tilde{\mu}_t^n) \right| \leq C'\|f\| \left(\frac{\langle \bar{\mu}_t^n, \chi^2 \rangle}{\langle \bar{\mu}_t^n, \chi \rangle^2} \right)^{1/2}.$$

Again using $\mathfrak{H}\mathfrak{o}$ and the fact that the moments are non increasing, we write $\langle \bar{\mu}_t^n, \chi \rangle = O(n)$ and also $\langle \bar{\mu}_t^n, \chi^2 \rangle = O(n)$. Finally, the approximation is complete :

$$\left| \bar{\mathfrak{L}}\Pi_f(\bar{\mu}_t^n) - \tilde{\mathfrak{L}}^n\Pi_f(\tilde{\mu}_t^n) \right| \leq C'\|f\|O\left(\frac{1}{\sqrt{n}}\right) = O\left(\frac{1}{\sqrt{n}}\right). \quad (\text{IV.29})$$

■

In summary, we have just now proven that the operator driving the large graph behavior of the limits get arbitrarily close to $\bar{\mathfrak{L}}$. The only thing left is to link all the eventual limits to the solution of the ODE driven by the *bar* operator. In the final part of the proof, we can finally establish the main theorem.

Part 4: Convergence to the unique solution

To complete the proof, one needs to remember the equation (IV.14) : For all n, t and f as above,

$$|\langle \bar{\mu}_t^n, f \rangle - \langle \bar{\mu}_t, f \rangle| \leq |\langle \bar{\mu}_t^n, f \rangle - \langle \tilde{\mu}_t^n, f \rangle| + |\langle \bar{\mu}_0^n, f \rangle - \langle \bar{\mu}_0, f \rangle| + |\mathcal{M}_t^{f,n}| \\ + \left| \int_0^t \tilde{\mathfrak{L}}^n \Pi_f(\tilde{\mu}_s^n) - \int_0^t \bar{\mathfrak{L}} \Pi_f(\bar{\mu}_s) ds \right|,$$

where $\bar{\mu}$ is characterized by $\bar{\mathfrak{L}}$ in the sense that $\bar{\mu}$ is the unique deterministic function such that for any suitable initial state ν ,

$$\langle \bar{\mu}_t, f \rangle = \langle \nu, f \rangle + \int_0^t \bar{\mathfrak{L}} \Pi_f(\bar{\mu}_s) ds.$$

Ultimately, we can use the results we established earlier to show that every term on the right hand side tends to zero and that every converging sub sequence converges to $\bar{\mu}$.

- First, under our premise, for some positive M :

$$\left| \tilde{\mathfrak{L}}^n \Pi_f(\bar{\mu}_t^n) \right| \leq \|f\| (2 + 4M). \quad (\text{IV.30})$$

Hence, using proposition III.3.2 which characterizes the Markov approximation $|\langle \bar{\mu}_t^n, f \rangle - \langle \tilde{\mu}_t^n, f \rangle| \xrightarrow{n, \mathbb{P}} 0$ and we can use the *tilde* processes.

- The convergence of $|\langle \bar{\mu}_0^n, f \rangle - \langle \nu, f \rangle|$ to 0 is again immediate since it is required in $\mathfrak{H}0$.
- As seen in the part 2 of the proof, the martingale part also converges to 0.
- For the last term, suppose $\tilde{\mu}^{\phi(n)}$ converges in law to $\bar{\mu}$ along a subsequence ϕ . **Skorokhod's representation theorem** gives the existence of a common probability space $(\Omega', \mathfrak{F}', \mathbb{P}')$ where $X^{\phi(n)}$ are $D([0, 1], \mathcal{M}_{\Phi})$ valued random variables, each $X^{\phi(n)}$ has the same law as $\tilde{\mu}^{\phi(n)}$ and $X^{\phi(n)}$ converges a.s to X and X again has the same

law as $\bar{\mu}$. To start off, write

$$\begin{aligned} & \left| \int_0^t \tilde{\mathfrak{L}}^n \Pi_f(\tilde{\mu}_s^n) - \int_0^t \bar{\mathfrak{L}} \Pi_f(\bar{\mu}_s) ds \right| \\ & \leq \int_0^t |\tilde{\mathfrak{L}}^n \Pi_f(\tilde{\mu}_s^n) - \bar{\mathfrak{L}} \Pi_f(\bar{\mu}_s^n)| ds + \left| \int_0^t \bar{\mathfrak{L}} \Pi_f(\bar{\mu}_s) - \bar{\mathfrak{L}} \Pi_f(\bar{\mu}_s^n) ds \right|. \end{aligned}$$

The first term on the RHS goes to zero by the large graph approximation of the semigroup (IV.2.2). For the second term, consider a convergent subsequence $\bar{\mu}^{\psi(n)}$ and let it converge almost surely to some μ^* (by placing ourselves directly in the common probability space). Using the hypothesis that the matching criterion is well behaved and the continuous mapping theorem, we have that $\bar{\mathfrak{L}} \Pi_f(\bar{\mu}_s^n) \xrightarrow{n, \mathbb{P}} \bar{\mathfrak{L}} \Pi_f(\bar{\mu}_s^*)$. Moreover, again by our premises, $\bar{\mathfrak{L}} \Pi_f(\bar{\mu}_s^n)$ is bounded. We use the dominated convergence theorem to next assert that $\int_0^t \bar{\mathfrak{L}} \Pi_f(\bar{\mu}_s^n) ds \xrightarrow{n} \int_0^t \bar{\mathfrak{L}} \Pi_f(\bar{\mu}_s^*) ds$. Scaling (IV.15) alongside $\psi(n)$ yields that

$$\langle \mu_t^*, f \rangle = \langle \nu, f \rangle + \int_0^t \bar{\mathfrak{L}} \Pi_f(\bar{\mu}_s^*) ds.$$

But $\bar{\mathfrak{L}}$ characterizes only $\bar{\mu}$, and we finally conclude that all the subsequential limits of $\bar{\mu}^n$ are in fact equal to $\bar{\mu}$.

We have therefore shown that

$$\sup_{t \in [0,1]} |\langle \bar{\mu}_t^n, f \rangle - \langle \bar{\mu}_t, f \rangle| \xrightarrow{n, \mathbb{P}} 0, \tag{IV.31}$$

effectively proving the convergence of $\bar{\mu}_t^n$ to $\bar{\mu}_t$, the unique solution of the ODE driven by $\bar{\mathfrak{L}}$. ■

Remark IV.2.2. *Our proof uses the same ideas upon which the differential equation method is based (Kurtz [1969] or Ethier and Kurtz [1986]). In both these works, the conditions for a convergence to a deterministic differential equation are given. In our proof, we adapted those arguments and expanded them to abstract matching criteria when required. Similarly to the work of Kurtz [1969] :*

- *The fact that the generator $\bar{\mathfrak{L}}$ is the limit of the generators $\tilde{\mathfrak{L}}^n$ in the sense of Kurtz is ensured by corollary IV.2.2. This corollary depends on the assumption that a chosen criterion is well behaved and moment preserving.*
- *The fact that the limiting generator $\bar{\mathfrak{L}}$ defines a semigroup is ensured by the fact that it is simply the generator associated to an alternative Markov construction. Thus, it has to be under the conditions of the Hille Yosida theorem.*

- Finally, in Kurtz [1969] a condition of Lipschitzity was used to ensure that the deterministic differential equation born from $\bar{\mathcal{L}}$ uniquely defined its solutions. In our case, even though Lipschitzity is sufficient, it will prove difficult to establish, especially on measure spaces. For this reason, we reframed that condition in a more abstract formulation.

At long last, we provide an example for the main result using the GREEDY criterion. It is demonstrated that the GREEDY algorithm verifies all the hypotheses of the main theorem and thus we are allowed to predict its matching coverage by solving the ODE (as is done in the simulations).

IV.3 Illustrative example : greedy

For this section, we recall that GREEDY uses a double uniform criterion, where the uniform choice is made over the positive activity vertices. We also set f as a generic function in \mathcal{C}_b . It is important to remember that the limiting behavior is given by the alternative, simpler dynamic. The GREEDY algorithm makes its first choice straight on the degree distribution:

$$\langle \Phi_{\widehat{K}}^{\text{GREEDY}}(\mu), f \rangle = \frac{\langle \mu, f \mathbf{1}_{\mathbb{N}^*} \rangle}{\langle \mu, \mathbf{1}_{\mathbb{N}^*} \rangle}, \quad (\text{IV.32})$$

whilst the second choice is made over the K neighbors which each follow a size biased distribution. We use the computations from (IV.5) apply :

$$\begin{aligned} \langle \Phi_{\widehat{K}'}^{\text{GREEDY}}(\mu), f \rangle &= \sum_{k'} f(k') \mathbb{P}_{\mu}(\widehat{K}' = k') \\ &= \sum_k \sum_{k'} f(k') \mathbb{P}_{\mu}(\widehat{K}' = k' | K = k) \mathbb{P}_{\mu}(\widehat{K} = k) \\ &= \sum_k \sum_{k'} f(k') \frac{k' \mu(k')}{\langle \mu, \chi \rangle} \frac{k \mathbf{1}_{\mathbb{N}^*}(k) \mu(k)}{\langle \mu, \mathbf{1}_{\mathbb{N}^*} \rangle} \\ &= \frac{\langle \mu, \chi f \rangle}{\langle \mu, \chi \rangle}. \end{aligned} \quad (\text{IV.33})$$

Hence, the operator $\bar{\mathfrak{L}}$ is explicit for GREEDY. Plugging (IV.32) and (IV.33) into (IV.3) we immediately get that

$$\bar{\mathfrak{L}}\Pi_f(\mu) = - \left\{ \frac{\langle \mu, f \mathbf{1}_{\mathbb{N}^*} \rangle}{\langle \mu, \mathbf{1}_{\mathbb{N}^*} \rangle} + \frac{\langle \mu, (\chi - \mathbf{1}_{\mathbb{N}^*}) \rangle}{\langle \mu, \mathbf{1}_{\mathbb{N}^*} \rangle} \frac{\langle \mu, \chi \nabla f \rangle}{\langle \mu, \chi \rangle} + \frac{\langle \mu, \chi f \rangle}{\langle \mu, \chi \rangle} + \frac{\langle \mu, \chi^2 - \chi \rangle}{\langle \mu, \chi \rangle} \frac{\langle \mu, \chi \nabla f \rangle}{\langle \mu, \chi \rangle} \right\}. \quad (\text{IV.34})$$

Ultimately, the main theorem for the GREEDY criterion is expressed:

Theorem IV.3.1. *Assuming that the graphs are under the GREEDY criterion, if the sequence of processes $(\bar{\mu}_t^n)_{n \geq 1}$ is under $\mathfrak{H}\mathfrak{o}$ and $\langle \nu, \xi^{3.5} \rangle \leq M$, then for $f \in \mathcal{C}_b$*

$$\sup_{t \leq 1} |\langle \bar{\mu}_t^n, f \rangle - \langle \bar{\mu}_t, f \rangle| \xrightarrow{(n, \mathbb{P})} 0, \quad (\text{IV.35})$$

where $\bar{\mu}_\cdot$ is characterized by the operator $\bar{\mathfrak{L}}$ defined by (IV.34) when $\langle \mu, \mathbf{1}_{\mathbb{N}^*} \rangle > 0$, and

$$\bar{\mathfrak{L}}\Pi_f(\mu) = 0 \text{ for all } f \in \mathcal{C}_b,$$

whenever $\langle \mu, \mathbf{1}_{\mathbb{N}^*} \rangle = 0$.

The convergence for GREEDY is proven by verifying that both the uniform choice function and the operator it produces verify the conditions of Theorem IV.1.1. To do so, we use the two following sections.

IV.3.1 Proof of Theorem IV.3.1: Well behavedness and moment properties

Lemma IV.3.1 (GREEDY: Well-behavedness and moment preservation). *Let $\mu \in \mathcal{M}$ such that $\langle \mu, \chi \rangle > 0$, then the uniform choice function is well behaved in the sense given by Definition III.3.1. Moreover, the GREEDY criterion preserves moments up to $p - 1$ when $p \geq 2$ and $\langle \mu, \chi^p \rangle \leq M$ for some $M > 0$.*

Proof. First, we prove the well behavedness. By definition, the weak topology is the weakest topology such that the functions $(\Pi_f(\cdot), f \in \mathcal{C}_b)$ are continuous. Notice that $f \mathbf{1}_{\mathbb{N}^*}$ is still in \mathcal{C}_b so that

$$\Phi_{\hat{R}}^{\text{GREEDY}}(\mu) = \frac{\Pi_{f \mathbf{1}_{\mathbb{N}^*}}(\mu)}{\Pi_{\mathbf{1}_{\mathbb{N}^*}}(\mu)}$$

is a continuous function on \mathcal{M} when $\Pi_{\mathbf{1}_{\mathbb{N}^*}}(\mu)$ is non zero : The uniform choice function is well behaved.

For the moment preservation property, assume that $\Pi_f(\mu^n) \rightarrow \Pi_f(\mu)$ and that $\langle \mu, \chi^p \rangle \leq M$, we can write for any positive K

$$|\Pi_{\chi^p}(\mu^n) - \Pi_{\chi^p}(\mu)| \leq |\Pi_{\chi^p \mathbf{1}_{[0,K]}}(\mu^n - \mu)| + |\Pi_{\chi^p \mathbf{1}_{[K,\infty)}}(\mu)| + |\Pi_{\chi^p \mathbf{1}_{[K,\infty)}}(\mu^n)|.$$

Using the weak convergence and the boundedness of $\langle \mu, \chi^p \rangle$, the right hand side can be as small as we wish and we can conclude that the uniform choice function conserves the moments up to order p .

Thus, the first uniform choice $\Phi_K^{\text{GREEDY}}(\mu)$ conserves all the moments of μ . Next, we will apply the same argument to the size biased of version μ . Recall the size biased version is written

$$\langle \tilde{\mu}, f \rangle = \frac{\langle \mu, \chi f \rangle}{\langle \mu, \chi \rangle}.$$

Since the moments of μ are preserved to order p , it follows that $\Phi_{K'}^{\text{GREEDY}}(\mu)$ preserves the moments of the size biased version of μ up to $p - 1$. ■

Before proceeding to the next step, we enounce the asymptotic version of Proposition III.2.2, which means the moments of the GREEDY dynamic stay non increasing.

Corollary IV.3.1 (Boundedness of moments). *Let $\bar{\mu}$ a solution characterized by a generator $\bar{\mathfrak{L}}$ with initial distribution $\nu \in \mathcal{M}$. If for some $p > 1$ and some $M > 0$, $\langle \nu, \chi^p \rangle < M$ then for all $t \leq 1$, $\langle \bar{\mu}_t, \chi^{p-1} \rangle$ is non increasing. In particular, $\langle \bar{\mu}_t, \chi^{p-1} \rangle < M$.*

Proof. The proof is immediate by noticing that the function $t \mapsto \bar{\mathfrak{L}} \Pi_{\chi^{p-1}}(\bar{\mu}_t)$, which is the derivative of the $p - 1$ th moment is well defined and negative. ■

We can now prove that the operator given by the GREEDY criterion characterizes the solution of the induced ODE.

IV.3.2 Proof of Theorem IV.3.1: Characterization of the solutions

For this section, assume $\nu \in \mathcal{M}$ is an initial degree distribution with $\langle \nu, \chi^{3.5} \rangle < M$. Using corollary IV.3.1, for any solution $\bar{\mu}_t$, we observe that

$$\sup_{t \in [0,1]} (m_{\bar{\mu}_t}, M_{\bar{\mu}_t}, V_{\bar{\mu}_t}) < M.$$

Lemma IV.3.2 (Uniqueness of the solution characterized by $\bar{\mathfrak{L}}$). *Let x_t and y_t two solutions under the GREEDY criterion, and characterized by the generator $\bar{\mathfrak{L}}$ with $x_0 = y_0 = \nu$. Then, for all $t \leq 1$ we have that $x_t = y_t$.*

Proof. The proof is split in two parts. We first get rid of the obvious case. When $\langle x_0, \mathbf{1}_{\mathbb{N}^*} \rangle = \langle y_0, \mathbf{1}_{\mathbb{N}^*} \rangle = 0$, then $\bar{\mathfrak{L}}\Pi_f(x_0) = \bar{\mathfrak{L}}\Pi_f(y_0)$. It follows that the solutions are both constant, namely $x_t = y_t = \nu$ for all t .

Else, given $\beta > 0$ small, both $\langle x_t, \mathbf{1}_{\mathbb{N}^*} \rangle$ and $\langle y_t, \mathbf{1}_{\mathbb{N}^*} \rangle$ are non increasing in t , hence we define

$$T_\beta = \sup(t > 0 \mid \langle x_t, \mathbf{1}_{\mathbb{N}^*} \rangle \wedge \langle y_t, \mathbf{1}_{\mathbb{N}^*} \rangle > \beta),$$

and we show that any solution is unique for $t \leq T_\beta$. Then, we set

$$z_t = x_t - y_t.$$

and, for some $\alpha > 3.5$,

$$\Gamma_t = \sum_{i>0} i^\alpha z_t^2(\{i\}).$$

We will show that

$$(\Gamma_t)' \leq (C_1 + C_2)\Gamma_t, \quad 0 \leq t \leq T_\beta. \tag{IV.36}$$

Using Gronwall's lemma, this would imply that

$$\Gamma_t \leq \Gamma_0 e^{(C_1+C_2)t}, \quad 0 \leq t \leq T_\beta.$$

But $\Gamma(0) = 0$ because of the same initial conditions for x and y , so this would imply that $\Gamma_t = 0$ for all $0 \leq t \leq T_\beta$, which would immediately imply in turn that $z_t = 0$ for all such t , and thereby conclude the proof of uniqueness until T_β .

So we are rendered to show (IV.36). For notational simplicity, we drop the dependency in time for all processes. Moreover, for any measure μ , we introduce the following :

- $m_\mu = \langle \mu, \mathbf{1}_{\mathbb{N}^*} \rangle$;
- $M_\mu = \langle \mu, \chi \rangle$;
- $V_\mu = \langle \mu, \chi^2 \rangle$;
- $A_\mu = \langle \mu, \chi f \rangle + \langle \mu, \chi^2 - \chi \rangle \frac{\langle \mu, \chi \Delta f \rangle}{\langle \mu, \chi \rangle}$.

The proof amounts to upper-bounding

$$\Gamma' = 2 \sum i^\alpha z(\{i\}) z'(\{i\}).$$

For this, we start by finding a convenient expression for $z'(\{i\})$. By the very definition of $\bar{\mathfrak{L}}$ any $f \in \mathcal{C}_b$ $\langle z, f \rangle'$ can be rewritten as

$$\begin{aligned} \langle z, f \rangle' &= \bar{\mathfrak{L}}\Pi_f(x) - \bar{\mathfrak{L}}\Pi_f(y) \\ &= \left[\frac{\langle y, f \mathbf{1}_{\mathbb{N}^*} \rangle}{m_y} - \frac{\langle x, f \mathbf{1}_{\mathbb{N}^*} \rangle}{m_x} \right] + \left[(M_y - m_y) \frac{\langle y, \chi \Delta f \rangle}{m_y M_y} - (M_x - m_x) \frac{\langle x, \chi \Delta f \rangle}{m_x M_x} \right] \\ &\quad + \left[\frac{1}{M_y} A_y - \frac{1}{M_x} A_x \right] \\ &=: D_0(x, y, f) + D_1(x, y, f) + D_2(x, y, f). \end{aligned} \quad (\text{IV.37})$$

- First, expanding the expression of $D_1(x, y, f)$ we get that

$$\begin{aligned} D_1(x, y, f) &= \frac{\langle x, \chi \Delta f \rangle}{m_x M_x} \langle y - x, \chi - \mathbf{1}_{\mathbb{N}^*} \rangle + \langle y, \chi - \mathbf{1}_{\mathbb{N}^*} \rangle \left[\frac{\langle y, \chi \Delta f \rangle}{m_y M_y} - \frac{\langle x, \chi \Delta f \rangle}{m_x M_x} \right] \\ &= \frac{\langle x, \chi \Delta f \rangle}{m_x M_x} \langle y - x, \chi - \mathbf{1}_{\mathbb{N}^*} \rangle \\ &\quad + \langle y, \chi - \mathbf{1}_{\mathbb{N}^*} \rangle \left[\frac{\langle y, \chi \Delta f \rangle m_x M_x - \langle x, \chi \Delta f \rangle m_y M_y}{m_y M_y m_x M_x} \right] \\ &= \frac{\langle x, \chi \Delta f \rangle}{m_x M_x} \langle y - x, \chi - \mathbf{1}_{\mathbb{N}^*} \rangle \\ &\quad + \langle y, \chi - \mathbf{1}_{\mathbb{N}^*} \rangle \frac{\langle y - x, \chi \Delta f \rangle m_y M_y - \langle y, \chi \Delta f \rangle (m_y M_y - m_x M_x)}{m_y M_y m_x M_x}. \end{aligned} \quad (\text{IV.38})$$

- The same technique is used for evaluating $D_2(x, y, f) := \left[\frac{m_y}{M_y} A_y - \frac{m_x}{M_x} A_x \right] :$

$$D_2(x, y, f) = A_y \left[\frac{m_y}{M_y} - \frac{m_x}{M_x} \right] + \frac{m_x}{M_x} [A_y - A_x]. \quad (\text{IV.39})$$

First off :

$$\left[\frac{m_y}{M_y} - \frac{m_x}{M_x} \right] = \frac{M_x m_{y-x} + M_{x-y} m_x}{M_x M_y}.$$

Remembering the definition of A , we also have that

$$\begin{aligned} A_y - A_x &= \langle y - x, \chi f \rangle + \langle y - x, \chi^2 - \chi \rangle \frac{\langle y, \chi \Delta f \rangle}{M_y} \\ &\quad + \langle x, \chi^2 - \chi \rangle \frac{\langle y - x, \chi \Delta f \rangle M_y + \langle y, \chi \Delta f \rangle \langle y - x, \chi \rangle}{M_y M_x}. \end{aligned}$$

Henceforth, $D_2(x, y, f)$ is rewritten :

$$\begin{aligned}
 D_2(x, y, f) = & A_y \left[\frac{M_x m_{y-x} + M_{x-y} m_x}{M_x M_y} \right] + \frac{m_x}{M_x} [\\
 & \langle y - x, \chi f \rangle + \langle y - x, \chi^2 - \chi \rangle \frac{\langle y, \chi \Delta f \rangle}{M_y} \\
 & + \langle x, \chi^2 - \chi \rangle \frac{\langle y - x, \chi \Delta f \rangle M_y + \langle y, \chi \Delta f \rangle \langle y - x, \chi \rangle}{M_y M_x}] \quad (IV.40)
 \end{aligned}$$

In particular, using $f = \mathbb{1}_i$, we have $z(\{i\}) := \langle z, \mathbb{1}_i \rangle$ and :

$$\begin{aligned}
 z'(\{i\}) = & (i+1) \left[\frac{M_y - m_y}{M_x} + \frac{m_x(V_x - M_x)}{M_x^2} \right] z(\{i+1\}) \\
 & - \left[1 + i \left(\frac{M_y - m_y}{M_x} + \frac{m_x(V_x - M_x)}{M_x^2} + \frac{m_x}{M_x} \right) \right] - \frac{M_z - m_z}{M_x} (ix(\{i\}) - (i+1)x(\{i+1\})) \\
 & - \frac{M_y - m_y}{M_y M_x} (iy(\{i\}) - (i+1)y(\{i+1\})) M_z + A_{y(\{i\})} \frac{m_x M z - M_x m_z}{M_y M_x} \\
 & - \frac{m_x}{M_x} \frac{V_z - M_z}{M_y} (iy(\{i\}) - (i+1)y(\{i+1\})) \\
 & - \frac{m_x}{M_x} \frac{V_x - M_x}{M_x M_y} (iy(\{i\}) - (i+1)y(\{i+1\})) m_z.
 \end{aligned}$$

One last time, we simplify using $E_{x,y} = \frac{M_y - m_y}{M_x} + \frac{m_x(V_x - M_x)}{M_x^2}$:

$$\begin{aligned}
 dz_i = & \left[(i+1) E_{x,y} z(\{i+1\}) - \left(1 + i \left(E_{x,y} + \frac{m_x}{M_x} \right) \right) z(\{i\}) \right] - \\
 & \frac{M_z - m_z}{M_x} (ix(\{i\}) - (i+1)x(\{i+1\})) - \frac{M_y - m_y}{M_y M_x} (iy(\{i\}) - (i+1)y(\{i+1\})) M_z \\
 & + A_{y(\{i\})} \frac{m_x M z - M_x m_z}{M_y M_x} - \frac{m_x}{M_x} \frac{V_z - M_z}{M_y} (iy(\{i\}) - (i+1)y(\{i+1\})) - \\
 & \frac{m_x}{M_x} \frac{V_x - M_x}{M_x M_y} (iy(\{i\}) - (i+1)y(\{i+1\})) m_z. \quad (IV.41)
 \end{aligned}$$

To start off, we are interested in

$$\Gamma'_1 := \sum_i i^\alpha \left[(i+1) E_{x,y} z(\{i+1\}) z(\{i\}) - \left(1 + i \left(E_{x,y} + \frac{m_x}{M_x} \right) \right) z(\{i\})^2 \right]. \quad (IV.42)$$

Using the fact that $2z(\{i+1\})z(\{i\}) = z(\{i+1\})^2 + z(\{i\})^2 - (z(\{i+1\}) - z(\{i\}))^2$, then

that $2z(\{i+1\})z(\{i\}) \leq z(\{i+1\})^2 + z(\{i\})^2$, we write

$$\begin{aligned}
\Gamma'_1 &\leq \frac{1}{2}E_{x,y} \sum_i i^\alpha(i+1)z(\{i+1\})^2 + \frac{1}{2}E_{x,y} \sum_i i^\alpha(i+1)z(\{i\})^2 \\
&\quad - \sum_i i^\alpha \left(1 + i(E_{x,y} + \frac{m_x}{M_x})\right) z(\{i\})^2 \\
&\leq E_{x,y} \sum_i i^{\alpha+1}z(\{i\})^2 + \frac{1}{2}E_{x,y} \sum_i i^\alpha z(\{i\})^2 - \sum_i i^\alpha z(\{i\})^2 - (E_{x,y} + \frac{m_x}{M_x}) \sum_i i^{\alpha+1}z(\{i\})^2 \\
&\leq \frac{1}{2}E_{x,y} \sum_i i^\alpha z(\{i\})^2 - \sum_i i^\alpha z(\{i\})^2.
\end{aligned} \tag{IV.43}$$

The first majoration is ended considering that $|E_{x,y}| \leq \frac{2M}{\beta} + \frac{M^2}{\beta^2} + \frac{M}{\beta}$:

$$|\Gamma'_1| \leq \left(\frac{M^2 + 3M\beta}{2\beta^2} + 1\right) \Gamma,$$

or

$$|\Gamma'_1| \leq C_1 \Gamma \tag{IV.44}$$

Afterwards, we have to find a similar majoration for $\Gamma'_2 := \frac{1}{2}\Gamma' - \Gamma'_1$. First notice for z that :

- For the mass m_z :

$$m_z \leq \sum_k z_k = \sum_k k^{\alpha/2} z_k k^{-\alpha/2}.$$

Using Cauchy-Schwartz yields

$$m_z \leq \left(\sum_k k^\alpha z_k^2\right)^{1/2} \left(\sum_k k^{-\alpha}\right)^{1/2}.$$

Since $\alpha > 3.5$, it follows that

$$m_z \leq \Gamma^{1/2} C_m. \tag{IV.45}$$

where $C_m > 0$

- In a similar fashion :

$$M_z = \sum_k k z_k = \sum_k k^{\alpha/2} z_k k^{1-\alpha/2}$$

yields that

$$M_z \leq \Gamma^{1/2} C_M. \quad (\text{IV.46})$$

with $\alpha > 1.5$

- Finally,

$$V_z \leq \Gamma^{1/2} C_V. \quad (\text{IV.47})$$

since $\alpha > 3.5$

Thus, we write

$$\begin{aligned} \Gamma'_2 = & \left[\frac{M_z - m_z}{M_x} \sum_i i^\alpha z(\{i\}) (ix(\{i\}) - (i+1)x(\{i+1\})) - \right. \\ & M_z \frac{M_y - m_y}{M_y M_x} \sum_i i^\alpha z(\{i\}) (iy(\{i\}) - (i+1)y(\{i+1\})) + \\ & \frac{m_x M_z - M_x m_z}{M_y M_x} \sum_i i^\alpha z(\{i\}) A_{y(\{i\})} - \\ & \frac{m_x V_z - M_z}{M_x M_y} \sum_i i^\alpha z(\{i\}) (iy(\{i\}) - (i+1)y(\{i+1\})) - \\ & \left. m_z \frac{m_x V_x - M_x}{M_x M_x M_y} \sum_i i^\alpha z(\{i\}) (iy(\{i\}) - (i+1)y(\{i+1\})) \right]. \quad (\text{IV.48}) \end{aligned}$$

Since the terms outside the sums are multiples of the moments of z , they are to be bounded by $\Gamma^{1/2}$ multiplied by some constant, we can also find bounds for the other terms :

- Starting with

$$\sum_i i^\alpha z(\{i\}) (i+1)y(\{i+1\}),$$

Cauchy-Schwartz yields again

$$\begin{aligned} \sum_i i^\alpha z(\{i\}) (i+1)y(\{i+1\}) &= \sum_i i^{\alpha/2} z(\{i\}) (i+1) i^{\alpha/2} y(\{i+1\}) \\ &\leq \Gamma^{1/2} \left(\sum_i (i+1)^2 i^\alpha y(\{i+1\})^2 \right)^{1/2}. \end{aligned}$$

But

$$\sum_i (i+1)^2 i^\alpha y(\{i+1\})^2 \leq \left(\sum_i (i+1) i^{\alpha/2} y_{i+1} \right)^2,$$

So that

$$\sum_i (i+1)^2 i^\alpha y_{i+1}^2 \leq \sum_i i^{1+\alpha/2} y_i \leq M.$$

We obtain

$$\sum_i i^\alpha z_i (i+1) y_{i+1} \leq M\Gamma^{1/2} \quad (\text{IV.49})$$

- The same argument yields

$$\sum_i i^\alpha z_i i y_i \leq M\Gamma^{1/2}. \quad (\text{IV.50})$$

We finally beget :

$$\begin{aligned} |\Gamma'_2| &\leq \frac{C_M + C_m}{\beta} \Gamma^{1/2} 2M\Gamma^{1/2} + C_M \Gamma^{1/2} \frac{2M}{\beta^2} 2M\Gamma^{1/2} + \\ &M \frac{C_M + C_m}{\beta^2} \Gamma^{1/2} \left(C_M + \frac{2M}{\beta} 2M \right) \Gamma^{1/2} + \\ &\frac{M}{\beta} \frac{C_V + C_M}{\beta} \Gamma^{1/2} 2M\Gamma^{1/2} + C_m \Gamma^{1/2} \frac{2M^2}{\beta^3} 2M\Gamma^{1/2}, \end{aligned}$$

or

$$|\Gamma'_2| \leq C_2 \Gamma \quad (\text{IV.51})$$

Together with IV.44, IV.51 is equivalent to IV.36 ending the proof of the uniqueness up to T_β .

Since β is arbitrary and $\Pi_{\mathbf{1}_{\mathbb{N}^*}}(\mu_t)$ is continuous in t . We can define :

$$T_0 = \sup_{\beta > 0} T_\beta \quad (\text{IV.52})$$

and extend the uniqueness to $t \leq T_0$. But, we have already proven the uniqueness after T_0 . Because for any $t > T_0$, the generator is null and the argument at the beginning of this proof applies : $\mu_t = \mu_{T_0}$.

Hence, the generator associated to the greedy criterion globally *characterizes* the solutions of the associated differential equation and the main theorem is true for the GREEDY criterion. ■

Remark IV.3.1. *The well behavedness and the moment preservation are fairly easy to assess for most matching criteria. As an illustrative example, for MINRES the first choice*

function is the same as that of GREEDY :

$$\langle \Phi_{\widehat{K}}^{\text{MINRES}}(\mu), f \rangle = \langle \Phi_{\widehat{K}}^{\text{GREEDY}}(\mu), f \rangle = \frac{\Pi_{f\mathbf{1}_{\mathbb{N}^*}}(\mu)}{\Pi_{\mathbf{1}_{\mathbb{N}^*}}(\mu)}.$$

As for the second choice function $\Phi_{\widehat{K}'}^{\text{MINRES}}(\mu)$, it is obtained by taking the minimum of $\widehat{K} - 1$ iid degrees each following the size biased degree distribution. We write :

$$\Phi_{\widehat{K}'}^{\text{MINRES}}(\mu)(\{i\}) = \mathbb{E}_\mu \mathbb{P}(\widehat{K}' = i | \widehat{K}) = \mathbb{E}_\mu (\mathbb{P}(\widehat{K}' \geq i | \widehat{K}) - \mathbb{P}(\widehat{K}' \geq i + 1 | \widehat{K})),$$

$$\Phi_{\widehat{K}'}^{\text{MINRES}}(\mu)(\{i\}) = \mathbb{E}_\mu \left(\left(\frac{\chi \Pi_{\mathbf{1}_{[i, \infty)}}(\mu)}{\Pi_\chi(\mu)} \right)^{\widehat{K}-1} - \left(\frac{\chi \Pi_{\mathbf{1}_{[i, \infty)}}(\mu)}{\Pi_\chi(\mu)} \right)^{\widehat{K}-1} \right).$$

And then,

$$\Phi_{\widehat{K}'}^{\text{MINRES}}(\mu)(\{i\}) = \sum_{k \geq 1} \frac{\Pi_{\mathbf{1}_k}(\mu)}{\Pi_{\mathbf{1}_{\mathbb{N}^*}}(\mu)} \left(\left(\frac{\chi \Pi_{\mathbf{1}_{[i, \infty)}}(\mu)}{\Pi_\chi(\mu)} \right)^{k-1} - \left(\frac{\chi \Pi_{\mathbf{1}_{[i, \infty)}}(\mu)}{\Pi_\chi(\mu)} \right)^{k-1} \right).$$

Restricting the space to $\mathcal{M}_{\Phi^{\text{MINRES}}}$ makes this sum finite and well behavedness (and moment conservation) ensues by the same continuity argument we used for GREEDY.

The complicated part for most criteria is uniqueness. Although it is suggested by the simulations, we were only able to prove it in trivial cases, or by severely restricting the space of measures under consideration. The authors are currently working on a paper expanding the proof of the uniqueness for other choice functions such as MINRES.

Chapter V

Notes and conclusion

Ultimately, the purpose of our work was to present an alternative framework to the study of online matching algorithms. In a world where large networks are gaining more and more prominence, our goal was to use the relatively recent tools offered by stochastic processes, queuing networks and random graphs to build efficient matching algorithms.

From a practical perspective, this work offers a way *to predict* the performances of specific matching policies given a population distribution, simply by (numerically for example) solving an ODE¹. For example, our results can be used to test how well *individual* matching policies would fare globally into a given population without having to go through extensive surveys. Independently of the population, our work can be extended to the comparison of given policies in a general framework. A relevant question for example would be which policies are better than total randomness? What would be the cost of applying such policies? In any case, The algorithms built in this framework are efficient by themselves. As shown by Aoudi et al. [2022] or alternatively by Noiry et al. [2021]) the present method, when adapted to bipartite networks already produces better results than the prevailing approach for online matching, even for mere greedy algorithms.

From a theoretical perspective, the present work falls within the scope of the methods that use random networks to emulate large networks and their limiting properties. Specifically, we studied exploration/marking processes on the configuration which we transposed to their measure valued representation. The '*choice functions*' used here can be viewed as marking processes as well as *favored directions* guiding the exploration. Henceforth, the techniques provided here can easily be generalized to problems other than matching on graphs. The properties proved on matching policies are easily applicable to several other questions involving marking. In fact, our work itself follows directly from that of Bermolen et al. [2017] on the jamming constant.

¹Our code for general/bipartite matching algorithms is available upon request

As for the extensions of our work, the most important challenges are twofold. First, we still have to establish a consistent framework for the study of the uniqueness in an abstract policy. Even for the *Greedy* criterion, the usual Cauchy-Lipschitz type conditions for the limits were insufficient. Second, thus far the comparison of given policies is only possible numerically. A theoretical comparison between abstract criteria seems very much out of reach. And even though we suspect a coupling argument to be sufficient, our work has not lead us to the adequate coupling and the optimal criteria we found, we only found numerically. To date, *Minres* is the policy that produces the best matching coverage and a legitimate question to ask is that of the global optimality of *Minres* amongst all criteria.

All things considered, we believe that the present work opens a promising line of research.

Appendix A

Appendix

A.1 The Configuration Model and Degree Sequences

The goal of this section is a brief presentation of the configuration model and a survey of its relevant properties to local matching. We include it for the sake of completeness, although we refer the interested reader to Newman [2018] or van der Hofstad [2016] for more thorough presentations of both its theoretical properties and its practical uses.

Let $\mathfrak{d} := (d(1), \dots, d(n))$ be a degree sequence associated to a graph $(\mathcal{V}, \mathcal{E})$ and let $\mu := \sum_1^n \delta_{d(i)}$ be its representation as a measure on $\mathcal{M}_F(\mathbb{N})$. The configuration model is the construction of a (multi-)graph with degree sequence \mathfrak{d} through the *uniform pairing procedure*. Any vertex v is given a number of half-edges (also called stubs) equivalent to its degree $d(v)$. So that

$$m_\mu = \langle \mu, 1 \rangle := \int 1 d\mu = n \tag{A.1}$$

is the number of nodes.

$$M_\mu = \langle \mu, \chi \rangle := \sum d(i) \tag{A.2}$$

is the number of half-edges. The construction of the graph is then done through pairing uniformly stubs into edges.

Definition A.1.1. A **configuration** c is an involution without any fixed points on the set of half edges $\mathfrak{m} := \{1, \dots, M_\mu\}$: $c(i) = j$ and $c(j) = i$ for any stub i . $(i, c(i))$ is called an edge. The **configuration model** $\mathbf{CM}(\mathfrak{d})$ or $\mathbf{CM}(\mu)$ is the uniform drawing of one among any $(M - 1)!! := (M - 1)(M - 3)\dots 3 \cdot 1$ configurations.

It is immediate that any given configuration is also a representation of a (multi-)graph. Thus, the configuration model can in fact also be viewed as an uniform drawing of a given

graph weighted by the number of underlying configurations representing it.

A.1.1 Algorithmic Construction of a Configuration

(For this section, we are thankful to P. Moyal for his course at IECL - Nancy Graphes Aléatoires et Applications)

Let $\mathfrak{d} \in \llbracket 1, n-1 \rrbracket^n$ a degree sequence. Usually, uniformly drawing a full involution proves to be complicated. Instead, configurations are build sequentially using one of the two following equivalent algorithms.

Definition A.1.2. *The configuration model $\text{CM}_n(\mathfrak{d})$ is the realization of a random graph on $\llbracket 1, n \rrbracket$ with degree distribution \mathfrak{d} , obtained by the algorithm 2.*

Algorithm 2 Configuration Model: Focus on Edges

Require: Degree sequence $\mathfrak{d} := (d(1), \dots, d(n))$.

Take n buckets of stubs, whose sizes are $d(1), \dots, d(n)$. At first, the stubs are said to be ‘available’.

while Available stubs remain **do**

 Pick a *first* stub.

 Pair it to another available stub chosen **uniformly at random** and delete both from the available stubs.

end while

Or equivalently,

Algorithm 3 Configuration Model: Focus on Nodes

Require: Degree sequence $\mathfrak{d} := (d(1), \dots, d(n))$.

Take n buckets of stubs, which sizes are $d(1), \dots, d(n)$. At first, the stubs are said to be ‘available’.

$N_0 = n$

while $N_k > 0$ **do**

 Pick a *first* bucket/node i_k between $\llbracket 1, N_k \rrbracket$,

 Pair one by one each of the remaining $d_k(i_k)$ stubs of i_k by sequentially choosing another stub **uniformly at random** between remaining available half-edges.

 Set as $d_{k+1}(j)$ the amount of remaining stubs for the j -th node after pairing, and as N_{k+1} the number of nodes with remaining stubs.

$k + 1 \rightarrow k$.

end while

A.1.2 The CM as an approximation for large simple graphs

Here we provide two propositions that justify the use of the CM as a model for large networks.

Although the underlying configurations are uniformly drawn on the CM, the same can not be said of the resulting graphs. For example, when there are parallel edges, swapping their ending stubs gives rise to different equiprobable configurations while the generated graph is the same. Hence, the same graph would have more weight than any other simple graph (generated by a single configuration). This flaw can be circumvented by noticing that, *conditioned on producing a simple graph*, the configuration model is a uniform drawing (For example Proposition 7.4 in van der Hofstad [2016]).

Proposition A.1.1 (The CM is a uniform drawing when conditioned to produce simple graphs). *Let \mathcal{G} a multi-graph with adjacency matrix A and degree distribution $\mathfrak{d} = (d(1), \dots, d(n))$ (and μ the associated measure). The probability for the CM to produce any given graph*

$$\mathbb{P}(\text{CM}_n(\mathfrak{d}) = \mathcal{G}) = \frac{1}{M_\mu!!} \frac{\prod_{i \in \llbracket 1, n \rrbracket} d(i)!}{\prod_{i \in \llbracket 1, n \rrbracket} 2^{A_{ii}} \prod_{1 \leq i < j \leq n} A_{ij}!}. \quad (\text{A.3})$$

When we restrict the graphs produced to be simple, this probability becomes a uniform drawing on simple graphs \mathcal{G} ;

$$\mathbb{P}[\text{CM}_n(\mathfrak{d}) = \mathcal{G} \mid \text{CM}_n(\mathfrak{d}) \text{ is simple}] = \frac{1}{\text{Card}\{\mathcal{G} \in \mathcal{G}(n) : \mathcal{G} \text{ follows } \mathfrak{d}\}}.$$

This property is powerful on its own, but it remains conditional. To justify the use of the CM as an approximation for real networks, we have to complement it with yet another property that guarantees the happening of the condition.

Proposition A.1.2 (CM and simple graphs). *Let \mathcal{G} a multi-graph with adjacency matrix A and degree distribution $\mathfrak{d} = (d(1), \dots, d(n))$ (and μ the associated measure). Let $\nu = \frac{\langle \mu, \chi^2 - \chi \rangle}{\langle \mu, \chi \rangle}$, we have:*

$$\mathbb{P}[\text{CM}_n(\mathfrak{d}) \text{ produces a simple graph}] \geq \exp\left(-\frac{\nu}{2} - \frac{\nu^2}{4}\right).$$

Both those propositions in conjunction ensure that the CM can be conditioned to mimic simple graphs. Afterwards, the graph produced is uniformly drawn.

A.1.3 The Independence Property

In this section, we explore the central property that allows us to build matching algorithms on the configuration. On the CM, the order in which we explore the edges (for example the *first* edges/nodes in the algorithm) does not influence the final law of the CM.

Proposition A.1.3 (Independence Property). *Recall that \mathfrak{m} is the set of stubs. The order in which the stubs are completed into edges does not influence the final law of the configuration model. In particular, edges can be created in any order regardless of the vertex they come from as long as the uniform pairing procedure is conserved.*

Proof. Let o be an arbitrary reordering of \mathfrak{m} . We show that any configuration c built using the sequential uniform pairing is uniformly drawn.

For this, notice that $c(o(1))$ is still chosen among $M - 1$ other stubs. Then $c(o(2))$ is chosen among $M - 3$ choices and so on. All in all, the process still produces $(M_\mu - 1)!!$ configurations which have the same uniform probability $\frac{1}{M_\mu - 1}$. ■

It is also said the uniform pairing process is *exchangeable*. Exchangeability allows us to build the exact number of edges needed at each step without influencing the macroscopic properties of the CM. Moreover these edges can be built in any order. Effectively, this means that the order in which the graph is built/explored does not bear any influence on its law.

At last, we provide one last macroscopic property of the configuration model which is abundantly used throughout our work

Proposition A.1.4 (Neighboring Degree Distribution). *Pick a random node I , the neighbor degree distribution is the degree distribution of a neighbor I' . We will call this distribution \mathbb{P}_μ :*

$$\mathbb{P}_\mu(k) = \mathbb{P}(d(I') = k) = \frac{k\mu(k)}{\langle \mu, \chi \rangle} \quad (\text{A.4})$$

It is a uniform law on all nodes size-biased by their available stubs oftentimes we simply denote it $\check{\mu}$.

Proof. There are $k\mu(k) - 1_{d(I)=k}$ stubs coming from degree k nodes over $\langle \mu, \chi \rangle - 1$ total stubs to be matched. The distribution above is a reasonable approximation when the number of nodes is large. ■

A.1.4 Graphicality of degree sequences

So far, we have assumed that the sequences of degrees used are *graphic*, meaning they represent the degrees of a graph. One way to have graphic sequences is to directly take the degrees of a given network. For example, this will be the case for the networks under our study or for surveyed networks. However, one legitimate question to ask when simulating networks is how do to produce graphic sequences (for example from a given probability

distribution). In this section, we provide the main characterization result for graphic sequences discovered by Erdos and Gallai which is used to construct graphical sequences.

Theorem A.1.1 (Erdos and Gallai on the characterization of graphic number sequences (1960)). *A sequence of non-negative integers $d(1) \geq \dots \geq d(n)$ can be represented as the degree sequence of a finite simple graph on n vertices if and only if $d(1) + \dots + d(n)$ is even and*

$$\sum_{i=1}^k d(i) \leq k(k-1) + \sum_{i=k+1}^n \min(d(i), k)$$

holds for every $1 \leq k \leq n$.

This test is used on randomly generated degree sequence. Mostly, when the sequence is close to a graphical sequence, it is 'truncated'. Furthermore, this approach to testing can be perfected using a generalization of the Erdos-Gallai theorem.

Theorem A.1.2 (Aigner and Triesch on the characterization of graphic number sequences). *A finite sequences of nonnegative integers $(d(1), \dots, d(n))$ with $d(1) \geq \dots \geq d(n)$ is graphic if $\sum_{i=1}^n d(i)$ is even and there exists a sequence $(c(1), \dots, c(n))$ that is graphic and majorizes $(d(1), \dots, d(n))$.*

In some contexts such as the truncation of degree sequences, this second result is simpler to use. The idea is to generate a random sequence of numbers (following a given distribution) and then to lower degrees until it becomes graphic.

A.2 Tools for the convergence of processes

This section of the annex is dedicated to the tools used in proving the convergence of processes. Most of the results are stated without proof and can be found in Parthasarathy's book [Parthasarathy, 1967]. Again for the sake of completeness, we give structural properties of $\mathcal{M}_F(\mathbb{N})^1$ as a reminder.

A.2.1 Topological Properties

Most of our work pertains to subspaces of $\mathcal{M}_F(\mathbb{N})$. To allege notations, we will only refer to it space as \mathcal{M}_F . First, we introduce the topologies under study.

Definition A.2.1. *Let $\mathcal{C}_b(\mathbb{R})$ be the space of continuous bounded functions on \mathbb{R} and let $\mathcal{C}_c(\mathbb{R})$ of continuous compactly supported functions. All of those spaces are endowed with*

¹ \mathbb{N} is of course endowed with its discrete topology

the supremum norm. The weak topology denoted τ_w is the topology induced on \mathcal{M}_F by its duality with $\mathcal{C}_b(\mathbb{R})$. For $\mu \in \mathcal{M}_F$, it the weakest topology such that the functions $\mu \mapsto \langle \mu, f \rangle, f \in \mathcal{C}_b(\mathbb{R})$ are continuous. The vague topology denoted τ_v is the topology induced by the duality with $\mathcal{C}_c(\mathbb{R})$.

Since $\mathcal{C}_c(\mathbb{R}) \subset \mathcal{C}_b(\mathbb{R})$, it is immediate that the weak topology is stronger than the vague topology.

Proposition A.2.1. *When endowed with the weak topology \mathcal{M}_F is not separable, while (\mathcal{M}_F, τ_v) is a complete separable space.*

This proposition is the reason the vague topology is used. Separable spaces are more convenient to use when proving convergence. Moreover, as will be shown hereafter, the weak convergence can easily follow from a vague convergence under sufficient restrictions. We state yet another important result pertaining to the characterisation of sequences.

Theorem A.2.1 (Prohorov). *Let Π be a family of probability measures on a measured space (S, \mathcal{S}) . If Π is tight, then it is relatively compact*

Usually, Prohorov's theorem is stated as an equivalence. Here we only state one side of the equivalence since the sufficiency is conditional to completeness and separability.

A.2.2 Characterizing Convergence

We remind the reader that the measures are considered on the Borel sigma algebra. Any time a measure is mentioned on some space E , it is considered on $(E, \mathcal{B}(E))$ where $\mathcal{B}(E)$ the sigma algebra generated by the topological open sets. Let $\mathbb{D}((0, T), \mathcal{M}_F)$ (resp. $\mathbb{C}((0, T), \mathcal{M}_F)$) be the space of right continuous, limited on the left functions (resp. continuous) from $(0, T)$ to \mathcal{M}_F endowed with the Skorokhod topology.

The main result of this thesis is a convergence proof for the law of sequence of processes $\mathbb{D}((0, T), \mathcal{M}_F)$. Essentially, it boils down to establishing relative compactness (the existence of several sub sequential limits) followed by the identification of the sub sequential limits as one and only process.

Relative compactness of the laws is proven via tightness using Prohorov's theorem. First, vague tightness is proven then it is extended to the weak topology. The following theorem characterizes weak convergence in terms of vague convergence :

Theorem A.2.2 (Convergence, Méléard and Roelly). *Let μ^n a sequence of processes in $\mathbb{D}((0, T), (\mathcal{M}_F, \tau_w))$ and μ a process in $\mathbb{C}((0, T), (\mathcal{M}_F, \tau_w))$. Then $\mu^n \Rightarrow \mu$ in $\mathbb{D}((0, T), (\mathcal{M}_F, \tau_w))$ if and only if $\mu^n \Rightarrow \mu$ in $\mathbb{D}((0, T), (\mathcal{M}_F, \tau_v))$ and $\langle \mu^n, 1 \rangle \Rightarrow \langle \mu, 1 \rangle$ in $\mathbb{D}((0, T), \mathbb{N})$, where we let the \Rightarrow designate the convergence in law.*

With this convergence result, we can prove the simpler vague tightness and then generalize it. To do so, we are provided with the following tightness criteria:

Theorem A.2.3 (Roelly's tightness criteria [Roelly-Coppoletta, 1986]). *Let μ^n a sequence of processes in $\mathbb{D}((0, T), (\mathcal{M}_F, \tau_v))$. Assume that for any $f \in \mathcal{C}_c(\mathbb{R})$, $\langle \mu^n, f \rangle$ is decomposable in a way that $\langle \mu^n, f \rangle = M^{f,n} + V^{f,n}$ where $M^{f,n}$ is a local martingale and $V^{f,n}$ has a finite variation. If*

1. *For $f \in \mathcal{C}_c(\mathbb{R})$ and for each $t \geq 0$, the law of $\langle \mu^n, f \rangle$ is tight.*
2. *The quadratic variation of $\ll M^{f,n} \gg$ and $V^{f,n}$ satisfy : for each $n \in \mathbb{N}, \epsilon, \eta$ positive, there exists $\delta > 0$ such that for each stopping time T_n bounded by n :*

$$\limsup_n \sup_{0 \leq \theta \leq \delta} \mathbb{P} \left(|V_{T_n+\theta}^{f,n} - V_{T_n}^{f,n}| \right) \leq \epsilon$$

$$\limsup_n \sup_{0 \leq \theta \leq \delta} \mathbb{P} \left(|\ll M^{f,n} \gg_{T_n+\theta} - \ll M^{f,n} \gg_{T_n}| \right) \leq \epsilon$$

Then the laws of μ^n are tight.

This tightness criteria makes use of the separability of $\mathcal{C}_c(\mathbb{R})$ and by extension of that of (\mathcal{M}_F, τ_v) (See [Roelly-Coppoletta, 1986]). Thus, it forces us to characterize vague convergence first.

As for the identification of the limit, most of the techniques used are pretty straightforward and directly detailed in the manuscript.

Bibliography

- Albert, R. and Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Rev. Mod. Phys.*, 74:47–97.
- Angel, O., van der Hofstad, R., and Holmgren, C. (2017). Limit laws for self-loops and multiple edges in the configuration model.
- Aoudi, M. H. A. D., Moyal, P., and Robin, V. (2022). Markovian online matching algorithms on large bipartite random graphs. *Methodology and Computing in Applied Probability*, 24(4):3195–3225.
- Berge, C. (1957). Two theorems in graph theory. *Proceedings of the National Academy of Sciences*, 43(9):842–844.
- Bermolen, P., Jonckheere, M., Larroca, F., and Moyal, P. (2016). Estimating the transmission probability in wireless networks with configuration models. *ACM Transactions on Modeling and Performance Evaluation of Computing Systems*, 1(2):1–23.
- Bermolen, P., Jonckheere, M., Larroca, F., and Saenz, M. (2019). Degree-greedy algorithms on large random graphs. *ACM SIGMETRICS Performance Evaluation Review*, 46(3):27–32.
- Bermolen, P., Jonckheere, M., and Moyal, P. (2017). The jamming constant of uniform random graphs. *Stochastic Processes and their Applications*, 127(7):2138–2178.
- Billingsley, P. (1999). *The Space D*, chapter 3, pages 121–179. John Wiley and Sons, Ltd.
- Birkhoff, G. (1946). Three observations on linear algebra. (spanish). *Revista A*, 5:147–151.
- Bollobás, B. (1980). A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European Journal of Combinatorics*, 1(4):311–316.
- Bollobás, B. and Béla, B. (2001). *Random graphs*. Number 73. Cambridge university press.

- Borgs, C., Chayes, J. T., Cohn, H., and Zhao, Y. (2019). An l^p theory of sparse graph convergence i: Limits, sparse random graph models, and power law distributions. *Transactions of the American Mathematical Society*, 372(5):3019–3062.
- Borodin, A., Karp, R., Tardos, G., and Wigderson, A. (1996). On the power of randomization in online algorithms. *Algorithmica*.
- Borodin, A., MacRury, C., and Rakheja, A. (2020). Bipartite stochastic matching: Online, random order, and i.i.d. models.
- Brightwell, G., Janson, S., and Luczak, M. (2015). The greedy independent set in a random graph with given degrees.
- Broutin, N. (2021). Cours :limites d'échelles de graphes aléatoires. <https://perso.lpsm.paris/~broutinn/teaching/m2graphes.html>. Accessed: 2010-09-30.
- Broutin, N. and Marckert, J.-F. (2014). Asymptotics of trees with a prescribed degree sequence and applications. *Random Structures & Algorithms*, 44(3):290–316.
- Brualdi, R., Harary, F., and Miller, Z. (2006). Bigraphs versus digraphs via matrices. *Journal of Graph Theory*, 4(1):51–73.
- Brubach, B., Grammel, N., Ma, W., and Srinivasan, A. (2021). Follow your star: New frameworks for online stochastic matching with known and unknown patience.
- Champagnat, N., Ferrière, R., and Meleard, S. (2008). From individual stochastic processes to macroscopic models in adaptive evolution. *Stochastic Models - STOCH MODELS*, 24:2–44.
- Chen, J., Huang, Q., Kanj, I., and Xia, G. (2021). Optimal streaming algorithms for graph matching.
- Chen, N., Litvak, N., and Olvera-Cravioto, M. (2017). Generalized pagerank on directed configuration networks. *Random Structures & Algorithms*, 51(2):237–274.
- Chen, N. and Olvera-Cravioto, M. (2013). Directed random graphs with given degree distributions. *Stochastic Systems*, 3(1):147 – 186.
- Dantzig, G. (1951). Maximization of linear function of variables subject to linear inequalities. *Activity Analysis of Production and Allocation. Proceedings of the Conference on Linear Programming*.

- Dawson, D. (1993). Measure-valued markov processes. In Hennequin, P.-L., editor, *Ecole d'Eté de Probabilités de Saint-Flour XXI - 1991*, pages 1–260, Berlin, Heidelberg. Springer Berlin Heidelberg.
- Decreusefond, L., Dhersin, J.-S., Moyal, P., and Tran, V. C. (2012). Large graph limit for an sir process in random network with heterogeneous connectivity. *Ann. Appl. Probab.*, 22(2):541–575.
- Decreusefond, L. and Moyal, P. (2008a). Fluid limit of a heavily loaded edf queue with impatient customers. *Markov Processes and Related Fields*, 14:131–157.
- Decreusefond, L. and Moyal, P. (2008b). A functional central limit theorem for the m/GI/ ∞ queue. *The Annals of Applied Probability*, 18(6):2156–2178.
- Decreusefond, L. and Moyal, P. (2012). *Markov Process*, chapter 7, pages 193–235. John Wiley & Sons, Ltd.
- Donato, D., Laura, L., Leonardi, S., and Millozzi, S. (2004). Large scale properties of the webgraph. *European Physical Journal B*, 38(2):239–.
- Doytchinov, B., Lehoczky, J., and Shreve, S. (2001). Real-time queues in heavy traffic with earliest-deadline-first queue discipline. *The Annals of Applied Probability*, 11(2):332–378.
- Dua, R. and Pettie, S. (2014). Linear-time approximation for maximum weight matching. *J. ACM*, 61(1).
- Durrett, R. (2006). *Random Graph Dynamics*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press.
- Edmonds, J. (1965). Paths, trees, and flowers. *Canadian Journal of Mathematics*, 17:449–467.
- Egerváry, E. (1931). Matrixok kombinatorius tulajdonságairol. *Mat. Fiz. Lapok*, 38.
- Erdős, P. (1965). A problem on independent r-tuples. In *ARTICLE IN PRESS B. Bollobás et al./Journal of Combinatorial Theory, Series A*. Citeseer.
- Erdős, P. and Rényi, A. (2006). *On the evolution of random graphs*, pages 38–82. Princeton University Press, Princeton.
- Ethier, S. N. and Kurtz, T. G. (1986). *Generators and Markov Processes*, chapter 4, pages 155–274. John Wiley & Sons, Ltd.

- Farhadi, A., Hajiaghayi, M., Mai, T., Rao, A., and Rossi, R. A. (2019). Approximate maximum matching in random streams.
- Feldman, J., Mehta, A., Mirrokni, V., and Muthukrishnan, S. (2009). Online stochastic matching: Beating $1-1/e$. In *2009 50th Annual IEEE Symposium on Foundations of Computer Science*, pages 117–126. IEEE.
- Ford, L. R. and Fulkerson, D. R. (1956). Maximal flow through a network. *Canadian Journal of Mathematics*, 8:399–404.
- Ford, L. R. and Fulkerson, D. R. (1957). A simple algorithm for finding maximal network flows and an application to the hitchcock problem. *Canadian Journal of Mathematics*, 9:210–218.
- Ford, L. R. and Fulkerson, D. R. (1958). Network flow and systems of representatives. *Canadian Journal of Mathematics*, 10:78–84.
- Fournier, N. and Méléard, S. (2004). A microscopic probabilistic description of a locally regulated population and macroscopic approximations. *The Annals of Applied Probability*, 14(4).
- Frieze, A. and Kannan, R. (1999). Quick approximation to matrices and applications. *Combinatorica*, 19:175–220.
- Frobenius, G. (1912). *Über Matrizen aus nicht negativen Elementen*. Preussische Akademie der Wissenschaften Berlin: Sitzungsberichte der Preußischen Akademie der Wissenschaften zu Berlin. Reichsdr.
- Frobenius, G. (1917). Über zerlegbare determinanten. *Sitzungsber. Königl. Preuss. Akad. Wiss.*, pages 147–151.
- Gallai, T. (1964). Maximale systeme unabhängiger kanten. *MTA Mat. Kut. Int. Közl.*
- Gamarnik, D. and Goldberg, D. A. (2010). Randomized greedy algorithms for independent sets and matchings in regular graphs: Exact results and finite girth corrections. *Combinatorics, Probability and Computing*, 19(1):61–85.
- Goel, G. and Mehta, A. (2008). Online budgeted matching in random input models with applications to adwords. In *SODA*, volume 8, pages 982–991.
- Gromoll, H. C., Puha, A. L., and Williams, R. J. (2002). The fluid limit of a heavily loaded processor sharing queue. *The Annals of Applied Probability*, 12(3):797–859.

- Grötschel, M., László, L., and Schrijver, A. (1981). The ellipsoid method and its consequences in combinatorial optimization. *Combinatorica*, 1:169–197.
- Hall, M. (1956). An algorithm for distinct representatives. *American Mathematical Monthly*, 63:716.
- Hall, P. (1935). On representatives of subsets. *Journal of The London Mathematical Society-second Series*, pages 26–30.
- Hofstad, R. v. d. (2016). *Random Graphs and Complex Networks*, volume 1 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press.
- Jacod, J. and Shiryaev, A. (2013). *Limit Theorems for Stochastic Processes*. Grundlehren der mathematischen Wissenschaften. Springer Berlin Heidelberg.
- Janson, S. (2013). The probability that a random multigraph is simple, ii.
- Joffe, A. and Metivier, M. (1986). Weak convergence of sequences of semimartingales with applications to multitype branching processes. *Advances in Applied Probability*, 18(1):20–65.
- Kallenberg, O. (2002). *Poisson and Pure Jump-Type Markov Processes*, pages 224–248. Springer New York, New York, NY.
- Kallenberg, O. (2017). *Random Measures, Theory and Applications*. Probability Theory and Stochastic Modelling. Springer International Publishing.
- Karl, M. (1927). Zur allgemeinen kurventheorie. *Fundamenta Mathematicae*, 10(1):96–115.
- Karp, R. M. (1991). An introduction to randomized algorithms. *Discrete Applied Mathematics*, 34(1):165–201.
- Karp, R. M. (1992). On-line algorithms versus off-line algorithms: How much is it worth to know the future? In *IFIP Congress*.
- Karp, R. M., Vazirani, U. V., and Vazirani, V. V. (1990). An optimal algorithm for on-line bipartite matching. In *Proceedings of the Twenty-Second Annual ACM Symposium on Theory of Computing, STOC '90*, pages 352–358, New York, NY, USA. Association for Computing Machinery.
- Kaspi, H. and Ramanan, K. (2011). Law of large numbers limits for many-server queues. *The Annals of Applied Probability*, 21(1):33–114.

- König, D. (1915). Line systems and determinants. *Math. Termész. Ert.*, 33:221–229.
- König, D. (1916). Über graphen und ihre anwendung auf determinantentheorie und mengenlehre. *Mathematische Annalen*, 77(4):453–465.
- König, D. (1931). Graphen und matrizen. *Mat. Fiz. Lapok*, 38.
- König, D. (1936). Theorie der endlichen und unendlichen graphen, akad. *Verlagsgesellschaft, Leipzig*.
- Kuhn, H. W. (1955). The hungarian method for the assignment problem. *Naval Research Logistics Quarterly*, 2(1-2):83–97.
- Kurtz, T. G. (1969). Extensions of trotter’s operator semigroup approximation theorems. *Journal of Functional Analysis*, 3(3):354–375.
- Kurtz, T. G. (1970). Solutions of ordinary differential equations as limits of pure jump markov processes. *Journal of Applied Probability*, 7(1):49–58.
- Kurtz, T. G. (1971). Limit theorems for sequences of jump markov processes approximating ordinary differential processes. *Journal of Applied Probability*, 8(2):344–356.
- Liggett, T. (2012). *Interacting Particle Systems*. Grundlehren der mathematischen Wissenschaften. Springer New York.
- Lovász, L. and Plummer, M. (2009). *Matching Theory*. AMS Chelsea Publishing Series. Akadémiai Kiadó.
- Lovasz, L. and Szegedy, B. (2004). Limits of dense graph sequences.
- Mahdian, M. and Yan, Q. (2011). Online bipartite matching with random arrivals: An approach based on strongly factor-revealing lps. In *Proceedings of the Forty-Third Annual ACM Symposium on Theory of Computing*, STOC ’11, pages 597–606, New York, NY, USA. Association for Computing Machinery.
- Manasse, M. S., McGeoch, L. A., and Sleator, D. D. (1988). Competitive algorithms for on-line problems. In *Symposium on the Theory of Computing*.
- Mehta, A. (2013). Online matching and ad allocation. *Foundations and Trends® in Theoretical Computer Science*, 8(4):265–368.
- Méléard, S. and Roelly, S. (1993). Sur les convergences étroite ou vague de processus à valeurs mesures.

- Metropolis, N. (1987). The beginning of the monte carlo method.
- Metropolis, N. and Ulam, S. (1949). The monte carlo method. *Journal of the American Statistical Association*, 44(247):335–341.
- Mitsuko, W. M. (2010). *The works of KONIG Dénes (1884–1944) in the domain of mathematical recreations and his treatment of recreational problems in his works of graph theory*. Theses, Université Paris-Diderot - Paris VII.
- Molloy, M., Reed, B., Newman, M., Barabási, A.-L., and Watts, D. J. (2011). A critical point for random graphs with a given degree sequence. In *The Structure and Dynamics of Networks*, pages 240–258. Princeton University Press.
- Newman, M. (2018). *Networks*. Oxford university press.
- Noiry, N., Sentenac, F., and Perchet, V. (2021). Online matching in sparse random graphs: Non-asymptotic performances of greedy algorithm. *CoRR*, abs/2107.00995.
- Ore, O. (1955). Graphs and matching theorems. *Duke Mathematical Journal*, 22(4):625–639.
- Parthasarathy, K. (1967). Ii - probability measures in a metric space. In Parthasarathy, K., editor, *Probability Measures on Metric Spaces*, volume 352 of *Probability and Mathematical Statistics: A Series of Monographs and Textbooks*, pages 26–55. Academic Press.
- Petersen, J. (1891). Die Theorie der regulären graphs. *Acta Mathematica*, 15(none):193 – 220.
- Petersen, J. (1898). Sur le théorème de tait. *Le Mathématiciens des Intermédiaire*, 5:225 – 227.
- Rabin, M. O. (1980). Probabilistic algorithm for testing primality. *Journal of Number Theory*, 12(1):128–138.
- R.J., W., W.O., J., and K., L. (1976). *Graph Theory 1736-1936*. Clarendon Press.
- Roelly-Coppoletta, S. (1986). A criterion of convergence of measure-valued processes: application to measure branching processes. *Stochastics*, 17(1-2):43–65.
- Sleator, D. D. and Tarjan, R. E. (1985). Amortized efficiency of list update and paging rules. *Commun. ACM*, 28:202–208.

- Solovay, R. and Strassen, V. (1977). A fast monte-carlo test for primality. *SIAM J. Comput.*, 6:84–85.
- Tait (1880). Xxiii.—note on a theorem in geometry of position. *Transactions of the Royal Society of Edinburgh*, 29(2):657–660.
- Tirodkar, S. (2018). Deterministic Algorithms for Maximum Matching on General Graphs in the Semi-Streaming Model. In Ganguly, S. and Pandya, P., editors, *38th IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2018)*, volume 122 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 39:1–39:16, Dagstuhl, Germany. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.
- Tutte, W. T. (1947). The factorization of linear graphs. *Journal of the London Mathematical Society*, s1-22(2):107–111.
- van der Hofstad, R. (2016). *Random Graphs and Complex Networks: Volume 1*. Cambridge University Press, USA, 1st edition.
- Veitch, V. and Roy, D. M. (2015). The class of random graphs arising from exchangeable random measures.
- von Neumann, J. (1953). 1. *A Certain Zero-sum Two-person Game Equivalent to the Optimal Assignment Problem*, pages 5–12. Princeton University Press, Princeton.
- Watts, D. J. and Strogatz, S. H. (1998). Collective dynamics of ‘small-world’ networks. *Nature*, 393:440–442.
- Wormald, N. (1999). The differential equation method for random graph processes and greedy algorithms. volume 73, page 155.
- Wormald, N. C. (1995). Differential equations for random processes and random graphs. *The Annals of Applied Probability*, 5(4):1217–1235.