Exponential Random Graphs

Tássio Naia dos Santos

Dissertation submitted in
partial fulfillement of the
requirements for the degree of
Master of Sciences in
Computer Science

Instituto de Matemática e Estatística
Universidade de São Paulo

Program: Computer Science
Advisor: Prof. Dr. Yoshiharu Kohayakawa

This work has been financially supported by CAPES.

São Paulo, November 2013

Exponential Random Graphs

This is the dissertation written by the candidate Tássio Naia dos Santos, revised according to suggestions of the Evaluation Comission.

Acknowledgements

I would especially like to thank my wife Sophie and family for the precious suport and patience, Yoshiharu Kohayakawa for the invaluable advice, and Fábio Happ Botler for many, many things. Also, I would like to thank Florência Graciela Leonardi, Daniel Morgato Martin e Rafael Crivellari Saliba Schouery for insightful comments on the text. Finally, I would like to thank CAPES for funding this project.

Abstract

Santos, T. N. **Exponential Random Graphs.** Dissertation — Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2013.

We study the behavior of the edge-triangle family of exponential random graphs (ERG) using the Markov Chain Monte Carlo method. We compare ERG subgraph counts and edge correlations to those of the classic Binomial Random Graph (BRG, also called Erdős–Rényi model).

It is a known theoretical result that for some parameterizations the limit ERG subgraph counts converge to those of BRGs, as the number of vertices grows [BBS11, CD11]. We observe this phenomenon on graphs with few (≈ 20) vertices in our simulations.

Resumo

Santos, T. N. **Grafos Aleatórios Exponenciais.** Dissertação — Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2013.

Estudamos o comportamento da família aresta-triângulo de grafos aleatórios exponenciais (ERG) usando métodos de Monte Carlo baseados em Cadeias de Markov. Comparamos contagens de subgrafos e correlações entre arestas de ERGs às de Grafos Aleatórios Binomiais (BRG, também chamados de Erdős–Rényi).

É um resultado teórico conhecido que para algumas parametrizações os limites das contagens de subgrafos de ERGs convergem para os de BRGs, assintoticamente no número de vértices [BBS11, CD11]. Observamos esse fenômeno em grafos com poucos (≈ 20) vértices em nossas simulações.

Contents

List of Tables xi

List of Figures xiii

1 Introduction 1

2 Preliminaries 3
- 2.1 Asymptotic values 4
- 2.2 Discrete probability 4
  - Conditional probability 4
  - Random variables 5
  - Expected value and variance 5
  - Conditional expectation 6
- 2.3 Graphs 6
  - Clustering in empirical networks 7
- 2.4 Statistics 7

3 Binomial and Exponential Random Graphs 9
- 3.1 Binomial Random Graph 9
  - Properties of $G(n, p)$ 9
- 3.2 Exponential Random Graph 10
  - Dependency structure 12
- 3.3 Markov Graphs 13
- 3.4 The edge-triangle model 13
- 3.5 Edge-triangle model and BRGs 16

4 MCMC 17
- 4.1 Markov Chains 17
- 4.2 Markov Chain Monte Carlo method 18
- 4.3 Mixing time of ERG 19

5 Graph Limits 23

6 Computational Experiments 25
- 6.1 Software for the simulation 26
- 6.2 Experiments 26
- 6.3 Comparison 27

7 Discussion and final comments 29

Bibliography 31

A Tables of computational experiments 35

B Proofs **41**

List of Tables

6.1 Growth of $\binom{n}{2}$ as a function of n. 26
6.2 Some sampled data statistics. 28

A.1 Statistical summary of the edge-triangle ERG samples 35
A.2 Statistical summary of the BRG $G(n, u^\star)$ samples. 37
A.3 Some numerical values for the relation between $u^\star = u^\star(\beta_1, \beta_2)$. 40

List of Figures

2.1 Examples of graphs. 6

6.1 Graphic of $u^\star$ as a function of (β_1, β_2). 27

7.1 Almost-bipartite edge-triangle sample. 30

Chapter 1

Introduction

Many interesting questions may be formulated in terms of a “network” of interactions among some class of objects. This abstraction is very versatile: one may consider physical networks (for instance, train lines connecting cities, neuronal conexions), social networks (collaboration between researchers, friendship relations), conceptual (links among internet pages, gene interaction), among others.

The focus of this dissertation is the *exponential random graph* (ERG), a model used for the study of empirical networks (that is, networks which are *observed* in nature, society *et cetera*). It is a probabilistic model, and its application entails calculations made using a computer. We study one of the most commonly used methods for ERG sampling, called Markov Chain Monte Carlo (MCMC, defined in the section 4.2). In addition, we sampled some exponential random graphs (from the edge-triangle model, described in section 3.4), and compared some of its characteristics to those observed on samples from another distribution (binomial random graph, see section 3.1).

The next chapter presents some important concepts from probability (section 2.2), combinatorics (section 2.3) and statistics (section 2.4).

Chapter 2

Preliminaries

In this chapter we introduce notation and concepts used throughout the text. We use the symbols

$\doteq$ indicates "equality by definition"

$\mathbb{N}$ set of natural numbers $\mathbb{N} \doteq \{1, 2, 3, \ldots\}$

$\mathbb{R}$ set of real numbers

2^A *power set* of A, $2^A \doteq \{B \colon B \subseteq A\}$

$[n]$ "canonical" subset with n elements, $[n] \doteq \{1, \ldots, n\}$

$|A|$ number of elements on the set A, for instance $|2^{[n]}| = 2^n$

$n!$ factorial, $n! \doteq n(n-1)(n-2)\cdots 2 \cdot 1$

$(n)_k$ falling factorial, $(n)_k \doteq n(n-1)(n-2)\cdots(n-k+1)$

$\binom{A}{k}$ family of *k-subsets*, $\binom{A}{k} \doteq \{B \subseteq A \colon |B| = k\}$

$\binom{n}{k}$ number of k-subsets of $[n]$, $\binom{n}{k} \doteq \left|\binom{[n]}{k}\right| = \frac{(n)_k}{k!}$

$f[A]$ image of the function f, restricted to A, $f[A] \doteq \{f(a) \colon a \in A\}$

A^k if A is a set, denotes the cartesian product $A \times A \times \cdots \times A$ (k factors)

p, q except if otherwise noted, $0 \leq p \leq 1$ and $q \doteq 1 - p$

$V(H), E(H), N(v)$ set of the vertices, edges of a graph H; set of neighbors of the vertex v

We adopt the notation "[proprerty]" from Iverson, which means:

$$[\,\text{property}\,] \doteq \begin{cases} 1 & \text{if the property holds, and} \\ 0 & \text{otherwise.} \end{cases}$$

For instance: $\binom{n}{k} = \sum_{A \subseteq [n]} [\,|A| = k\,]$.

2.1 Asymptotic values

Let f and g be sequences of positive numbers. We write $f = o(g)$ if $\lim_{n\to\infty} \frac{f(n)}{g(n)} = 0$; and we write $f = \Omega(g)$ if there exist constants C and n_0 such that $Cf(n) \geq g(n)$ for $n > n_0$. Finally, we write $f = \Theta(g)$ se $f = \Omega(g)$ and $g = \Omega(f)$.

All logarithms are relative to the natural base $e \approx 2.718$.

2.2 Discrete probability

A probability space is a triple $(\Omega, \mathcal{F}, \mathrm{P})$, where Ω is a *countable set* (i.e., there is an injection $f\colon \Omega \to \mathbb{N}$), $\mathcal{F} = 2^{\Omega}$ is the set of all subsets of Ω (the possible *events* of a random experiment) and $\mathrm{P}\colon \mathcal{F} \to [0,1]$ is a function which satisfies the following properties:

1. $0 \leq \mathrm{P} \leq 1$, for all $A \in \mathcal{F}$;
2. $\mathrm{P}(\Omega) = 1$;
3. if $A_1, A_2, \ldots \in \mathcal{F}$ are pairwise disjoint, then $\mathrm{P}(\bigcup_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} \mathrm{P}(A_k)$.

We can motivate these probability axioms interpreting $\mathrm{P}(A)$ as the empirical frequency expected for the occurence of the event A. If we perform n "independent" experiments (that is, such that the outcome of each one does not interfere with the others), and if we count the number n_A of *ocurrences* of A (that is, the number of times the result of one such experiment was an element a from A), then the empirical frequency $f(A) \doteq n_A/n$ should approximate $\mathrm{P}(A)$ when n is "large enough." Note that the function f satisfies the three properties above.

Conditional probability

Many statements about probability have the form "if A happens, then the probability of B is p", where A and B are events and p is a probability. To include such formulations in our formalism, we consider an experiment repeated n times, and two events A and B: we count the number of occurrences $n_A, n_B, n_{A\cap B}$ of the events A, B and $A \cap B$ (simultaneous occurrences of A and B), respectively. Considering only the experiments in which B ocurred, the empirical frequence of A is $n_{A\cap B}/n_B$ (assuming B occurs), and we may write

$$\frac{n_{A\cap B}}{n_B} = \frac{n_{A\cap B}/n}{n_B/n}.$$

These fractions can be seen as probabilities, and motivate the following definition.

Given that B occurs, we know that A occurs if and only if $A \cap B$ occurs. Hence, the conditional probability of A given B, which we denote by $\mathrm{P}(A|B)$, must be proportional to $\mathrm{P}(A \cap B)$. Let $\mathrm{P}(A|B) = \alpha \mathrm{P}(A \cap B)$ for some constant $\alpha = \alpha(B)$. The conditional probability $\mathrm{P}(\Omega|B)$ must be 1, and therefore $\alpha \mathrm{P}(\Omega \cap B) = 1$, thus $\alpha = 1/\mathrm{P}(B)$. We define the conditional probability of some event A given the occurrence of some event B by $\mathrm{P}(A|B) \doteq \mathrm{P}(A \cap B)/\mathrm{P}(B)$. Note that $(\Omega_B, \mathcal{F}_B, \mathrm{P}_B)$, where $\Omega_B \doteq \Omega \cap B$, $\mathcal{F}_B \doteq \{A \cap B \colon A \in \mathcal{F}\}$, and $\mathrm{P}_B(A) \doteq \mathrm{P}(A|B)$ is a probability space.

This definition is a starting point for the notion of "independence:" we say that the events A and B are *independent* if $\mathrm{P}(A \cap B) = \mathrm{P}(A) \cdot \mathrm{P}(B)$. If $\mathrm{P}(B) > 0$, this implies $\mathrm{P}(A|B) = \mathrm{P}(A)$ e $\mathrm{P}(B|A) = \mathrm{P}(B)$.

Finally, we enunciate an important result, the *law of total probability*. Let $\mathcal{F} = \{B_i \colon i \in I\}$ be a partition of Ω, that is, a family of Ω subsets such that $\bigcup_{i\in I} B_i = \Omega$ and, for $i, j \in I$, we have $B_i \cap B_j = \emptyset$ whenever $i \neq j$. Hence, for every event A, we have $\mathrm{P}(A) = \sum_{i\in I} \mathrm{P}(A|B_i) \cdot \mathrm{P}(B_i)$.

Random variables

Let $(\mathcal{F}, \Omega, \mathrm{P})$ be a probability space, as in the previous section. A (real) *random variable*, or RV, is a function $X\colon \Omega \to \mathbb{R}$ such that for all $a \in \mathbb{R}$, we may attribute probability to the event

$$\{X \leq a\} \doteq \{\omega \in \Omega \colon X(\omega) \leq a\}.$$

In other words, X is such that $\{X \leq a\} \in \mathcal{F}$. In particular, a function $X\colon \Omega \to E$, where E is a countable set is called *discrete random variable* if for all $x \in E$ we have $\{X = x\} \in \mathcal{F}$ (where $\{X = x\} \doteq \{\omega \in \Omega \colon X(\omega) = x\}$). An *indicator variable* of an event A is $f_A(\omega) \doteq [\omega \in A]$.

Two discrete random variables X and Y are independent if for all $x \in X[\Omega]$ and $y \in Y[\Omega]$ we have $\mathrm{P}(X = x \text{ and } Y = y) = \mathrm{P}(X = x) \cdot \mathrm{P}(Y = y)$. Furthermore, two discrete random variables X, Y are *conditionally independent*, given variables Z_i, for $i \in I$ (I an index set) if for all values $z_i \in Z_i[\Omega]$, $i \in I$ and all $x \in X[\Omega]$ and $y \in Y[\Omega]$

$$\mathrm{P}(X = x \text{ e } Y = y|\Omega_Z) = \mathrm{P}(X = x|\Omega_Z) \cdot \mathrm{P}(Y = y|\Omega_Z),$$

where $\Omega_Z \doteq \bigcap_{i\in I}\{Z_i = z_i\}$. Note that neither independence implies conditional independence nor the reverse.

Finally, a set of random variables $\{X_\lambda\}_{\lambda\in\Lambda}$ is *independent* if for all subsets of indexes $A \subseteq \Lambda$ and all sets of values $\{x_\lambda\}_{\lambda\in A}$, with $x_\lambda \in X_\lambda[\Omega]$ we have

$$\mathrm{P}(\bigcap_{\lambda\in A}\{X_\lambda = x_\lambda\}) = \prod_{\lambda\in A} \mathrm{P}(X_\lambda = x_\lambda).$$

Expected value and variance

The *expectation* $\mathbb{E}(X)$ of a random variable X is a "weighted average" of the values $X[\Omega]$. If X is a discrete RV,

$$\mathbb{E}(X) \doteq \sum_{\omega\in\Omega} X(\omega)\mathrm{P}(\omega) = \sum_{x\in X[\Omega]} x\mathrm{P}(X = x).$$

Note that for all RVs X, Y, we have $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$, and for every constant c, we have $\mathbb{E}(cX) = c\mathbb{E}(X)$. At last, if X is an indicator variable of the event A, then $\mathbb{E}(X) = \mathrm{P}(A)$. The *variance* $\mathrm{Var}(X)$ de X is defined as

$$\mathrm{Var}(X) \doteq \mathbb{E}((X - \mathbb{E}(X))^2).$$

The two following inequalities justify interpreting the expectation of X as the "expected value" of the variable, and its variance as a measure of *concentration* of X's values on the interval $[\mathbb{E}(X) - k \operatorname{Var}(X), \mathbb{E}(X) + k \operatorname{Var}(x)]$, para $k > 0$.

Theorem 2.1 (Markov's inequality). Let X be a RV which assumes only nonnegative values. For all $t > 0$ we have $\mathrm{P}(X - \mathbb{E}(X) \geq t) \leq \mathbb{E}(X)/t$.

Theorem 2.2 (Chebyshev's inequality). Let X be a RV with finite expectation $\mu \doteq \mathbb{E}(X)$ and finite variance σ^2. For all $k > 0$ we have $\mathrm{P}(|X - \mu| \geq k\sigma) \leq 1/k^2$.

Conditional expectation

The *conditional expectation* of the discrete RV X, given an event B is

$$\mathbb{E}(X|B) \doteq \sum_{\omega \in \Omega} X(\omega)\mathrm{P}(\omega|B) = \sum_{x \in X[\Omega]} x\mathrm{P}(X = x|B)$$

where $B \in \mathcal{F}$ is some event. If $\mathbb{E}(X)$ is limited, we have the *law of total expectation*: $\mathbb{E}(X) = \sum_{i \in I} \mathrm{P}(B_i)\mathbb{E}(X|B_i)$, where $\{B_i : i \in I\}$ is a partition of Ω.

2.3 Graphs

The usual abstract representation of a network, or graph $G = G(V, E)$ consists of a set V of *vertices* and a set of *edges* $E \subseteq \binom{V}{2}$. When $\{x, y\} \in E$, we say the vertices x and y are *connected*, or that they are *neighbors*. A vertex is *isolated* if it has no neighbors. For example, the graph with 3 vertices all connected among themselves is called a *triangle*, and the graph such that all but one vertex have the same (unique) neighbor is called a *star* (see figure 2.1). The graph on n vertices all connected is called *complete graph* and denoted by K_n. A *subgraph* of $G = (V, E)$ is a graph $H = (V', E')$ such that $V' \subseteq V$ and $E' \subseteq E$; we denote "H is a subgraph of G" by $H \subseteq G$. Bollobás [Bol98] and Diestel [Die10] have excellent introductions to the theory of graphs.

Two graphs $G = (V, E)$ and $H = (W, F)$ are *isomorph* if there is a bijection $f : V \to W$ such that $ij \in E$ if and only if $f(i)f(j) \in F$.

In this text, all graphs are *finite*, that is, have a finite set of vertices.

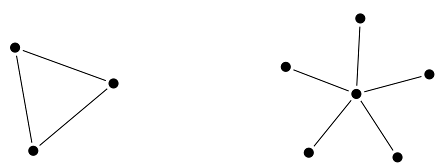

Figure 2.1: Examples of graphs. To the left, a triangle; to the right, a 5-star. Dots represent vertices and the lines represent edges.

Studies conducted in the last decades have highlighted structural characteristics shared by many empirical networks. The interested reader will find many surveys about such characteristics [AB02]. For example, these networks have many vertices, and a number of edges roughly linear on the number of vertices.

Clustering in empirical networks

Another property of interest which was observed in empirical networks is called *clustering*. This notion is an attempt to measure "transitivity" of the relation indicated by the edges of the graph. Informally, we can interpret this in the lines of "friends of my friends are my friends."

We now formalize the notion of clustering. Consider a graph $G = (V, E)$ (whose vertices and edges represent, say, people and their friendship relationships, respectively). Let $N(v)$ be the set of neighbors of some vertex $v \in V$, and let $\mathrm{d}(v) \doteq |N(v)|$. The number $e_v \doteq \left|E \cap \binom{N(v)}{2}\right|$ of edges connecting neighbors of v is a natural number between 0 and $\binom{\mathrm{d}(v)}{2}$. We define the *clustering* $\mathrm{clus}(v) \doteq e_v \binom{\mathrm{d}(v)}{2}^{-1}$. Studies suggest that the average of $\mathrm{clus}(v)$ is "high" in empirical networks [AB02].

We denote the *average clustering* (or simply *clustering*) of a graph G by $\mathrm{clus}(G) \doteq |V|^{-1} \sum_{v \in V} \mathrm{clus}(v)$. Note that each connection between two neighbors u and w of v corresponds to a triangle (u, v, w) in G. Thus, we expect to find a "high" number of triangles in empirical networks.

Naturally, it is important to agree as to what is a high clustering. We adopt the convention of taking as a reference the average value of $\mathrm{clus}(G)$, over (all) graphs wich n vertices. This approach is very common (see section 3.1).

Claim 2.3 (Average clustering of graphs on n vertices). Let $\mathcal{G}_n$ be the set of the graphs whose vertex set is $[n]$. We have

$$\frac{1}{|\mathcal{G}_n|} \sum_{G \in \mathcal{G}_n} \mathrm{clus}(G) = \frac{1}{2}.$$

Proof. See appendix B for a proof based on counting. A shorter, probabilistic argument is presented in the section 3.1. □

In general, we use n to denote the number of vertices of a graph. The number of edges and triangles of a graph G are $e(G)$ and $t(G)$, respectively. Also, if G is a graph with vertex set V and A is some set such that $A \subseteq \binom{V}{2}$, we say $A \subseteq G$ if all edges in A are edges of G. For $i, j \in V$, we write "$ij \in G$" whenever $\{i, j\}$ is an edge of G.

2.4 Statistics

We distinguish *parameters* and *estimators* of probability distributions. A *parameter* is a function of the probability space (for instance, the expectation of a random variable), and an *estimator* is a function of a sample (i.e., the realization of experiment), which we often use to estimate the value of some parameter (for example, the sample mean—see discussion below).

As an example, consider a probability distribution P, uniform, over the set $[n]$. The probability of the event $A \subseteq [n]$ is $\mathrm{P}(A) \doteq |A|/n$. Let $X(i) \doteq i$ be a RV in the probability space $([n], 2^{[n]}, \mathrm{P})$. The expectation of X is a parameter of the model, and has value $\mathbb{E}(X) \doteq \sum_{i \in [n]} i\mathrm{P}(X = i) = \sum_{i=1}^{n} i/n = (n+1)/2$.

Let $X_1, X_2, \ldots, X_k$ be RVs independent and identically distributed to X. We can estimate X by the sample mean $M \doteq \sum_{i=1}^{k} X_k/k$. Note that $\mathbb{E}(M) = \mathbb{E}(X)$.

We adopt the *sample mean* $\overline{X}$ and the sample standard deviation s_X as estimators, respectively, of the expectation $\mathbb{E}(X)$ and the *standard deviation* $\sqrt{\operatorname{Var} X}$.

The *sample mean* of X is denoted by $\overline{X}$, and the *sample standard deviation* by s_X.

$$\overline{X} \doteq \frac{\sum_{i=1}^{N} X_i}{N} \quad \text{and} \quad s_X \doteq \sqrt{\frac{\sum_{i=1}^{N}(X_i - \overline{X})^2}{N-1}},$$

where the samples are indexed from 1 to N. Furthermore, we estimate the covariance $C(x,y) \doteq \mathbb{E}\big((X - \mathbb{E}(X)(Y - \mathbb{E}(Y))\big)$ and the correlation $C(x,y)/\sqrt{\operatorname{Var} X \cdot \operatorname{Var} Y}$ between two RVs X and Y, using their *sample covariance* $\operatorname{Cov}(X,Y)$ and *sample correlation* $\operatorname{Corr}(X,Y)$, respectively:

$$\operatorname{Cov}(X,Y) \doteq \frac{\sum_{i=1}^{N}(X_i - \overline{X})(Y_i - \overline{Y})}{N-1} \quad \text{and} \quad \operatorname{Corr}(X,Y) \doteq \frac{\operatorname{Cov}(X,Y)}{s_X s_Y}.$$

Informally, the covariance is a measure of a linearity relation between RVs, and the correlation is a normalized version of the covariance (since $|\operatorname{Corr}(X,Y)| \le 1$). Note that an equivalent expression for the covariance between X and Y is $\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$, and therefore independent variables have covariation zero.

Chapter 3

Binomial and Exponential Random Graphs

In the following discussions, we use the word *model* somewhat loosely, to indicate a distribution, or family of distributions of probability. The models we describe have *parameters*, such as the number of vertices n, and often we are interested in the behaviour of the model as n tends to infinity.

3.1 Binomial Random Graph

A *random graph* is a random varaible assuming graphs as values. One of the most studied random graph models is the Binomial Random Graph (BRG), denoted by $G(n, p)$. It is a graphs on n labelled vertices, constructed adding each edge independently of the others with probability $p = p(n)$. Therefore, $G(n, 1/2)$ is an uniform distribution over the $2^{\binom{n}{2}}$ labelled graphs. In general, the probability of $G(n, p)$ be a given graph $H = ([n], E)$ is

$$\mathrm{P}(H) = p^{e(H)}(1-p)^{\binom{n}{2}-e(H)}. \tag{3.1}$$

For a more extensive study of random graphs, we refer the reader to Bollobás [Bol01] and to Janson, Łuczak and Ruciński [JŁR00].

Properties of $G(n, p)$

In this section we calculate the expected number of edges and triangles of $G(n, p)$. The calculations are elementary, and illustrate a kind of reasoning very typical of probabilistic combinatorics.

Claim 3.1 Let $G \sim G(n, p)$ be a BRG, and $A \subseteq [n]$ be a subset of the vertices of G. The expected number of edges of G between vertices of A is $\mathbb{E}(e(A)) = \binom{|A|}{2}p$.

Proof. Let $X_e = [\, e \in E(G)\,]$. Thus, $X_e = 1$ if the edge e is present in G, and 0 otherwise. Since X_e is an indicator variable, we have $\mathbb{E}(X_e) = \mathrm{P}(X_e = 1)$. Also,

$e(A) = \sum_{e \in \binom{A}{2}} X_e$, and by linearity of expectation

$$\mathbb{E}(e(A)) = \mathbb{E}\Big(\sum_{e \in \binom{A}{2}} X_e\Big) = \sum_{e \in \binom{A}{2}} \mathbb{E}(X_e) = \binom{|A|}{2} p. \qquad \square$$

Claim 3.2 Let $G \sim G(n,p)$. We have $\mathbb{E}(\text{clus}(G)) = p$.

Proof. By linearity of expectation, we have

$$\mathbb{E}(\text{clus}(G)) = n^{-1} \sum_{v \in [n]} \mathbb{E}(\text{clus}(v)). \tag{3.2}$$

consider the events $\Omega_A \doteq \{N(v) = A\}$, for $A \subseteq [n] - v$. These events form a *partition* of Ω (since $\Omega = \bigcup_{A \subseteq [n]-v} \Omega_A$ and also $\Omega_A \cap \Omega_B = \emptyset$ if $A \neq B$). Therefore,

$$\begin{aligned}
\mathbb{E}(\text{clus}(v)) &= \sum_{A \subseteq [n]-v} \mathrm{P}(\Omega_A)\mathbb{E}(\text{clus}(v)|\Omega_A) \\
&= \sum_{A \subseteq [n]-v} \mathrm{P}(\Omega_A)\mathbb{E}(e(A)|\Omega_A)\tbinom{|A|}{2}^{-1} \\
&= \sum_{A \subseteq [n]-v} \mathrm{P}(\Omega_A)\tbinom{|A|}{2}p\tbinom{|A|}{2}^{-1} \\
&= p,
\end{aligned}$$

Where we used claim 3.1 to obtain $\mathbb{E}(e(A)|\Omega_A). = \binom{|A|}{2}p$. Making the substitution of this value in the equation (3.2), we complete the proof. $\square$

Another proof of the claim 3.2 is providade on the appendix B.

Claim 3.3 Let $G \sim G(n,p)$. We have $\mathbb{E}(t(G)) = \binom{n}{3}p^3$.

Proof. Let $X(\{u,v,w\}) = [\{u,v,w\}$ form a triangle in $G]$. We have

$$\mathbb{E}(t(G)) = \mathbb{E}\Big(\sum_{A \in \binom{[n]}{3}} X(A)\Big) = \sum_{A \in \binom{[n]}{3}} \mathbb{E}(X(A)) = \sum_{A \in \binom{[n]}{3}} p^3 = \binom{n}{3}p^3. \qquad \square$$

3.2 Exponential Random Graph

In spite of being rich in properties, the BRG model is not appropriate to describe empirical networks—as observed by Erdős and Rényi [ER60]. In fact, there exist many other models for networks in the literature [WS98, AB99, CDS10, vdH09], which have been proposed with such goal. Our focus is the model called Exponential Random Graph (ERG), which is used in the social sciences [HL81, SPRH06]. In this model, the probability of a graph G with n vertices is

$$p_\beta(G) \doteq \exp\left(\sum_{i=1}^{k} \beta_i T_i(G) - \psi(\beta)\right) \tag{3.3}$$

where $\beta = (\beta_1, \ldots, \beta_k)$ is a vector of real parameters; $T_1, T_2, \ldots, T_k$ are real functions on the space of graphs (for instance, the number of edges, triangles, stars, circuits,...), and ψ is a *normalizing constant*, so that $\sum_G p_\beta(G) = 1$.

The expression in the exponent is occasionally referred in the literature as *Hamiltonian* (a term from statistical mechanics, not related to the graph theory usage), which is used to weight the probability measure over the graphs, assigning greater mass to graphs with "desirable" properties. For instance, fix parameters $h, \beta > 0$ and, for every graph G with n labelled vertices, $e(G)$ edges and $t(G)$ triangles, define the Hamiltonian of G as

$$H(G) \doteq he(G) + \beta t(G). \tag{3.4}$$

A probability measure on the space of (labelled n-vertex) graphs may then be defined as

$$p_n(G) = \frac{e^{H(G)}}{e^{-\psi}}, \tag{3.5}$$

where ψ is the normalizing constant, occasionally called *partition function* of the model.

We now show that every BRG is a ERG. Consider the ERG with distribution $p_n(G) = \exp(\beta e(G) - \psi)$, where ψ is a function of n and β. Since the sum of probabilities of the graphs with n vertices equals 1, we have

$$1 = \sum_{i=0}^{\binom{n}{2}} \sum_{\substack{G \\ e(G)=i}} \exp(\beta i - \psi) = e^{-\psi} \sum_{i=0}^{\binom{n}{2}} \binom{\binom{n}{2}}{i} e^{\beta i} = e^{-\psi}(1+e^{\beta})^{\binom{n}{2}}, \tag{3.6}$$

where G runs over all labelled graphs with n vertices and i edges. Thus $e^{\psi} = (1+e^{\beta})^{\binom{n}{2}}$ and

$$p_n(G) = e^{e(G)\beta}(1+e^{\beta})^{-\binom{n}{2}} = \left(\frac{e^{\beta}}{1+e^{\beta}}\right)^{e(G)} \left(1 - \frac{e^{\beta}}{1+e^{\beta}}\right)^{\binom{n}{2}-e(G)}, \tag{3.7}$$

which is $G\big(n, e^{\beta}/(1+e^{\beta})\big)$ with $0 < p = e^{\beta}/(1+e^{\beta}) < 1$. In the extreme cases of $p \in \{0,1\}$, the random graph $G(n,p)$ takes on unique values, and may be written in the form of an ERG model using indicator functions (of the empty and complete graph).

However, in general, the probability distributions of ERGs and of BRGs are distinct. Furthermore, the former are hard to compute (the normalizing constant may involve a nontrivial sum over $2^{\binom{n}{2}}$ graphs), rendering practically impossible direct sampling of ERGs. This motivated the search of distributions to approximate ERG models. In particular, one wish to sample from these distributions, an essential step of statistical applications of these models [CD11].

We present a characterization of the probability distributions of random graphs, obtained by Frank e Strauss [FS86], through application of the Hammersley–Clifford Lemma [Bes74, Gri]. It is expressed in terms of conditional dependencies between the indicator variables of the edges.

It will be useful to represent a graph $G = (V,E)$ with n vertices by a vector $\mathbf{x} = (x_e) \in \{0,1\}^{\binom{n}{2}}$, such that $x_e = [e \in E]$. Thus, a random graph is a probability distribution over the vectors in $\{0,1\}^{\binom{n}{2}}$.

Dependency structure

Let $Z_1, \ldots, Z_m$ be discrete random variables. The *dependency graph* D of $Z_1, \ldots, Z_m$ is a graph $D = (V, E)$ with $V = [m]$, and edges consisting of the pairs $\{i, j\} \in \binom{[m]}{2}$ such that Z_i e Z_j are conditionally dependent, given the values of Z_k, for $k \in [m] \setminus \{i, j\}$. That is, $ij \in E$ if

$$\begin{aligned} &\mathrm{P}(Z_i = z_i, Z_j = z_j \,|\, Z_k = z_k,\ k \neq i, j) \\ &= \mathrm{P}(Z_i = z_i \,|\, Z_k = z_k,\ k \neq i, j)\ \mathrm{P}(Z_j = z_j \,|\, Z_k = z_k,\ k \neq i, j), \end{aligned} \tag{3.8}$$

where $z_k \in \{0, 1\}$, for all $k \in [m]$ and $\{i, j, k\} \in \binom{[m]}{3}$. For instance, a sequence of independent random variables has an empty dependency graph (i.e., with no edges), and a Markov Chain $(Z_1, \ldots, Z_m)$ has dependency graph with edges $\{i, i+1\}$ for $i \in [m-1]$ (see section 4.1). A *clique* of the dependency graph D is a nonempty subset A of $[m]$ such that $\binom{A}{2} \subseteq E(D)$.

Theorem 3.4 (Frank–Strauss [FS86]). The probability distribution of a random graph G on the vertex set $[n]$ and dependency structure D can be written as

$$\mathrm{P}(G) = c^{-1} \exp \sum_{A \subseteq G} \alpha_A [\, A \text{ is a clique of } D \,], \tag{3.9}$$

where c is a normalizing constant

$$c = \sum_{G:\, \mathrm{P}(G) > 0} \exp \sum_{A \subseteq G} \alpha_A [\, A \text{ is a clique of } D \,], \tag{3.10}$$

and α_A are arbitrary constants.

As an example, let us consider the BRG $G(n, p)$. The $m = \binom{n}{2}$ indicator variables of edges $X = \big(X(1), \ldots, X(m)\big)$ de $G(n, p)$ are independent, and therefore its dependency graph is empty. Thus its probability distribution may be written as

$$\mathrm{P}(X = \mathbf{x}) = c^{-1} \prod_{x_e = 1} \exp \alpha_e, \tag{3.11}$$

where $\mathbf{x} = (x_1, \ldots, x_m) \in \{0, 1\}^{\binom{n}{2}}$ is a vactor with $\binom{n}{2}$ coordinates, representing a graph on n vertices, and $\alpha_e = \alpha_{\{e\}}$. The normalizing constant is

$$c = \sum_{\mathbf{y} \in \{0,1\}^k} \ \prod_{e \text{ edge of } \mathbf{y}} \exp \alpha_e = \prod_{i<j} (1 + \exp \alpha_e). \tag{3.12}$$

Factoring c according to the presence or absence of each edge e, the equation (3.11) becomes

$$\mathrm{P}(X = \mathbf{x}) = \left[\prod_{e \text{ edge of } \mathbf{x}} \frac{\exp \alpha_e}{1 + \exp \alpha_e} \right] \Bigg/ \left[\prod_{e \text{ is } \textbf{not} \text{ edge of } \mathbf{x}} (1 + \exp \alpha_e) \right]. \tag{3.13}$$

(Compare with equation (3.7).) And then the probability $p_e \doteq \mathrm{P}(X(e) = 1)$ is $p_e = \exp \alpha_e / (1 + \exp \alpha_e)$, or, equivalently $\alpha_e = \log\big(p_e / (1 - p_e)\big)$.

3.3 Markov Graphs

The dependency structure introduced in the last section motivates the definition of yet another class of random graphs, the *Markov graphs* [FS86]. A random graph is a *Markov graph* if disjoint edges are conditionally independent. In symbols, denoting by $X_{ab} \doteq [\{a,b\} \in E(G)]$ the indicator variable of the edge $\{a,b\}$, we have, for all *distinct* vertices a, b, c and d of the graph:

$$\mathrm{P}(X_{ab} = x_{ab} \text{ e } X_{cd} = x_{cd} \mid \Omega') = \mathrm{P}(E(ab) = e_{ab} \mid \Omega') \cdot \mathrm{P}(E(cd) = e_{cd} | \Omega') \qquad (3.14)$$

where $\Omega' \doteq \{X_e = x_e, e \in E(G) \setminus \{ab, cd\}\}$ and $x_{ab}, x_{cd}, x_e \in \{0,1\}$. Hence, for Markov graphs, the cliques of the dependency graph D correspond to sets of edges such that any pair of edges shares a vertex. (The only graphs without isolated vertices satisfying this restriction are triangles and stars—see Lemma 3.5.)

We highlight (again) that independence does not imply conditional independence (nor vice-versa), and note that the Markov graphs are *not*, a generalization (or a subclass) of BRGs. On the other hand, they are a subclass of ERGs (theorems 3.6 and 3.7).

Lemma 3.5 Let $G = (V, E)$ be a graph with at least one edge and no isolated vertices. If all pairs of edges $e, f \in E$ of G have a common vertex (i.e.: $e \cap f \neq \emptyset$), then G is a triangle or a star.

Proof. See appendix B. □

In the same article, Frank and Strauss present general expressions for the probability distributions of Markov Graphs. The following result deals with the particular case in which isomorph graphs have the same probability.

Theorem 3.6 (Frank–Strauss [FS86]). The probability distribution of a Markov Graph may be written as

$$\mathrm{P}(G) = c^{-1} \exp\left(\tau t + \sum_{k=1}^{n-1} \sigma_k s_k\right), \qquad (3.15)$$

where τ and σ_k are arbitrary constants, t and s_k are the number of triangles and k-stars in G, respectively, and c is a normalizing constant.

We demonstrate (theorem 3.7) that, if a random graph has probability distribution which can be written in the form of the equation (3.15), then it is a Markov Graph, thus completing the characterization of these models in terms of their probability distribution.

3.4 The edge-triangle model

The *edge-triangle* family of ERGs consists of the probability distributions over labelled graphs on n vertices of the form

$$p_{n,\beta_1,\beta_2} = p_{\beta_1,\beta_2}(G) = \exp\left(2\beta_1 e(G) + \frac{6\beta_2}{n} t(G) - n^2\psi\right), \qquad (3.16)$$

where $\psi = \psi_n(\beta_1, \beta_2)$, and the normalization ensures the model is non-trivial for large n (otherwise almost all graphs are empty or complete). This model has been the focus of many studies (see, for example [FS86, HJ99]).

The following theorem completes a characterization of the Markov Graphs (see theorem 3.6).

Theorem 3.7 Every edge-triangle ERG is a Markov Graph.

Proof. We prove that if G is a random graph with probability distribution given by (3.16), then, for all edges $e \doteq \{a, b\}$ and $f \doteq \{c, d\}$,

$$\mathrm{P}(X_e = x_e, X_f = x_f | \Omega_{ef}) = \mathrm{P}(X_e = x_e | \Omega_{ef}) \cdot \mathrm{P}(X_f = x_f | \Omega_{ef}), \tag{3.17}$$

where by convenience we write "$X_e = x_e, X_f = x_f$" for "$\{X_e = x_e\} \cap \{X_f = x_f\}$", and $\Omega_{ef} \doteq \bigcap_{g \in \binom{[n]}{2} \setminus \{e,f\}} \{X_g = x_g\}$. As before, for any edge $e' = \{u, v\} \in \binom{[n]}{2}$, we write $x_{e'} \doteq x_{uv} \in \{0, 1\}$, and for any triple of vertices $\{u, v, w\} \in \binom{[n]}{3}$, we set $x_{uvw} \doteq x_{uv} x_{uw} x_{vw}$. We write the probability distribution of G as

$$p(G) \doteq \exp\Bigg(\beta_1 \sum_{\{i,j\} \in \binom{[n]}{2}} X_{ij} \quad + \quad \beta_2 \sum_{\{i,j,k\} \in \binom{[n]}{3}} X_{ijk} \quad - \quad \psi, \Bigg)$$

where $X_{ij} \doteq [\, ij \in G \,]$, $X_{ijk} \doteq X_{ij} X_{ik} X_{jk}$, and $\psi = \psi_n$ is the normalizing constant of the model. Let $\Omega_e \doteq \{X_e = x_e\}$, and $\Omega_f \doteq \{X_f = x_f\}$. We have

$$\mathrm{P}(X_e = x_e | \Omega_{ef}) \cdot \mathrm{P}(X_f = x_f | \Omega_{ef}) \doteq \mathrm{P}(\Omega_e | \Omega_{ef}) \cdot \mathrm{P}(\Omega_f | \Omega_{ef}) \tag{3.18}$$

$$\doteq \frac{\mathrm{P}(\Omega_e \cap \Omega_{ef}) \cdot \mathrm{P}(\Omega_f \cap \Omega_{ef})}{\left(\mathrm{P}(\Omega_{ef})\right)^2}. \tag{3.19}$$

since $\mathrm{P}(\{X_e = x_e, X_f = x_f | \Omega_{ef}) = \mathrm{P}(\Omega_e \cap \Omega_f \cap \Omega_{ef}) / \mathrm{P}(\Omega_{ef})$, it is enought to prove that

$$\mathrm{P}(\Omega_e \cap \Omega_{ef}) \cdot \mathrm{P}(\Omega_f \cap \Omega_{ef}) = \mathrm{P}(\Omega_e \cap \Omega_f \cap \Omega_{ef}) \cdot \mathrm{P}(\Omega_{ef}).$$

Observe that $\mathrm{P}(\Omega_e \cap \Omega_{ef}) = \sum_{i \in \{0,1\}} \mathrm{P}(\{X_e = x_e\} \cap \{X_f = i\} \cap \Omega_{ef})$. Since $\{0, 1\} = \{x_f, 1 - x_f\} = \{x_e, 1 - x_e\}$, we can write

$$\mathrm{P}(\Omega_e \cap \Omega_{ef}) = \sum_{k \in \{x_f, 1 - x_f\}} \exp\Bigg(\beta_1 \Big(x_e + k + \sum_{ij \neq e,f} x_{ij}\Big) + \beta_2 \Big(\sum_{\substack{u,v,w \in [n] \\ ab \notin \{uv,uw,vw\} \\ cd \notin \{uv,uw,vw\}}} x_{uvw} + \sum_{u \neq a,b} x_e x_{ua} x_{ub} + \sum_{v \neq c,d} k x_{vc} x_{vd} \Big) - \psi \Bigg).$$

An analogous expression is valid for $\mathrm{P}(\Omega_f \cap \Omega_{ef})$:

$$\mathrm{P}(\Omega_f \cap \Omega_{ef}) = \sum_{\ell \in \{x_e, 1 - x_e\}} \exp\Bigg(\beta_1 \Big(\ell + x_f + \sum_{ij \neq e,f} x_{ij}\Big) + \beta_2 \Big(\sum_{\substack{u,v,w \in [n] \\ ab \notin \{uv,uw,vw\} \\ cd \notin \{uv,uw,vw\}}} x_{uvw} + \sum_{u \neq a,b} \ell x_{ua} x_{ub} + \sum_{v \neq c,d} x_f x_{vc} x_{vd} \Big) - \psi \Bigg).$$

We can factor the product

$$
\begin{aligned}
P(\Omega_e \cap \Omega_{ef}) \cdot \mathrm{P}(\Omega_f \cap \Omega_{ef}) &= \exp\Big(\sum_{ij \neq e,f} x_{ij} + \sum_{\substack{u,v,w \in [n] \\ ab \notin \{uv,uw,vw\} \\ cd \notin \{uv,uw,vw\}}} x_{uvw} - \psi \Big)^2 \\
&\cdot \Bigg(\sum_{k \in \{x_f, 1-x_f\}} \exp\Big(\beta_1 \Big(x_e + k \Big) + \beta_2 \Big(\sum_{u \neq a,b} x_e x_{ua} x_{ub} + \sum_{v \neq c,d} k x_{vc} x_{vd} \Big) \Big) \Bigg) \\
&\cdot \Bigg(\sum_{\ell \in \{x_e, 1-x_e\}} \exp\Big(\beta_1 \Big(\ell + x_f \Big) + \beta_2 \Big(\sum_{u \neq a,b} \ell x_{ua} x_{ub} + \sum_{v \neq c,d} x_f x_{vc} x_{vd} \Big) \Big) \Bigg) \qquad (3.20)
\end{aligned}
$$

We observe that for all constants C_1, C_2, C_3, C_4,

$$
\begin{aligned}
&\Big(\sum_{k \in \{x_f, 1-x_f\}} \exp(x_e C_1 + k C_2 + x_e C_3 + k C_4) \Big) \cdot \Big(\sum_{\ell \in \{x_e, 1-x_e\}} \exp(\ell C_1 + x_f C_2 + \ell C_3 + x_f C_4) \Big) \\
&= \exp\big(x_e (C_1 + C_3) + x_f (C_2 + C_4) \big) \sum_{\substack{k \in \{0,1\} \\ \ell \in \{0,1\}}} \exp\big(\ell (C_1 + C_3) + k (C_2 + k C_4) \big)
\end{aligned}
$$

Substituting in (3.20), we obtain,

$$
\begin{aligned}
&P(\Omega_e \cap \Omega_{ef}) \cdot \mathrm{P}(\Omega_f \cap \Omega_{ef}) \\
&= \exp\Big(\sum_{ij \neq e,f} x_{ij} + \sum_{\substack{u,v,w \in [n] \\ ab \notin \{uv,uw,vw\} \\ cd \notin \{uv,uw,vw\}}} x_{uvw} - \psi \Big)^2 \\
&\quad \cdot \exp\Big(x_e \Big(\beta_1 + \beta_2 \sum_{u \neq a,b} x_{ua} x_{ub} \Big) + x_f \Big(\beta_1 + \beta_2 \sum_{v \neq c,d} x_{vc} x_{vd} \Big) \Big) \\
&\quad \cdot \sum_{\substack{k \in \{0,1\} \\ \ell \in \{0,1\}}} \exp\Big(\ell \Big(\beta_1 + \beta_2 \sum_{u \neq a,b} x_{ua} x_{ub} \Big) + k \Big(\beta_1 + \beta_2 \sum_{v \neq c,d} x_{vc} x_{vd} \Big) \Big) \\
&= \mathrm{P}(\Omega_e \cap \Omega_f \cap \Omega_{ef}) \sum_{\substack{k \in \{0,1\} \\ \ell \in \{0,1\}}} \mathrm{P}(\{X_e = \ell\} \cap \{X_f = k\} \cap \Omega_{ef}) \\
&= \mathrm{P}(\Omega_e \cap \Omega_f \cap \Omega_{ef}) \sum_{\substack{k \in \{0,1\} \\ \ell \in \{0,1\}}} \mathrm{P}(\Omega_{ef} | X_e = \ell, X_f = k) \mathrm{P}(X_e = \ell, X_f = k) \\
&= \mathrm{P}(\Omega_e \cap \Omega_f \cap \Omega_{ef}) \mathrm{P}(\Omega_{ef}).
\end{aligned}
$$

□

We note that, with small tweaks, the proof above may be generalized, proving the theorem 3.7 for ERGs

$$
p'_n(G) \doteq c^{-1} \exp\Big(\sum_{\{i,j\} \in \binom{[n]}{2}} \beta_{ij} X_{ij} + \sum_{\{i,j,k\} \in \binom{[n]}{3}} \beta_{ijk} X_{ij} X_{ik} X_{jk} \Big),
$$

where β_{ij} and β_{ijk} are constants which may depend on n, and c is a normalizing constant. This formulation, slightly more general than the version we enunciated, is the reciprocal of the theorem presented by Frank and Strauss (Theorem 3 in [FS86]).

3.5 Edge-triangle model and BRGs

If the parameter β_2 of the edge-triangle model is positive, then we can determine the limit of the normalizing constant in ψ_n (equation (3.16)) as $n \to \infty$. This is a result obtained by Chatterjee and Diaconis [CD11], using graph limmits (see chapter 5):

$$\psi_n(\beta_1, \beta_2) \simeq \sup_{0 \le u \le 1} \left(\beta_1 u + \beta_2 u^3 - \frac{1}{2} u \log u - \frac{1}{2}(1-u)\log(1-u) \right). \tag{3.21}$$

Furthermore, the value $u^\star$ which attains the maximum in (3.21) is such that $G(n, u^\star)$ is, in a sense,"close" to $p_{n,\beta_1,\beta_2}(G)$ (see chapter 5).

Chapter 4

Markov Chain Monte Carlo method

Markov chains and Monte Carlo Methods are subject of a large body of mathematical literature. In the followin, we present some aspects of this rich theory. Excellent introductions to the subject have been written by Brémaud [Bré99], Diaconis [Dia08], and Levin, Peres and Wilmer [LPW09].

4.1 Markov Chains

Let $\mathcal{X}$ be a finite set, and $K(x,y)$ be a matrix with lines and columns indexed by $\mathcal{X}$ such that $K(x,y) \geq 0$ for all $x, y \in \mathcal{X}$ and $\sum_{y\in\mathcal{X}} K(x,y) = 1$ for each $x \in \mathcal{X}$. Hence each line of K defines a probability distribution and we can use K to direct a random walk over $\mathcal{X}$: from x, we proceed to y with probability $K(x,y)$. A *Markov Chain* is a sequence of random variables $\{X_i\}_{i\geq 0}$ each one taking on values in $\mathcal{X}$, such that the conditional probability distribution of X_{n+1} given $X_j = x_j$, where $x_j \in \mathcal{X}$ e $0 \leq j \leq n$ is

$$\mathrm{P}(X_{n+1} = x_{n+1} \,|\, X_j = x_j,\, 0 \leq j \leq n) = K(x_n, x_{n+1}). \tag{4.1}$$

Thus $\mathrm{P}(X_{n+2} = z \,|\, X_n = x) = \sum_{y\in\mathcal{X}} K(x,y)K(y,z)$. In general, the k-th power of K has $K^k(x,y) = \mathrm{P}(X_{n+k} = y \,|\, X_n = x)$. A probability distribution π over $\mathcal{X}$ is *stationary* for K if

$$\sum_{x\in\mathcal{X}} \pi(x)K(x,y) = \pi(y), \tag{4.2}$$

that is, if π is a left eigenvector of K with eigenvalue 1. An interpretation of (4.2) is "pick x according to π and follow one step according to $K(x,y)$; the probability of going to y is $\pi(y)$." The following theorem guarantees that under some natural correctness conditions, π is unique and large powers of K converge to the matrix with all lines equal to $\pi(x)$.

Theorem 4.1 Let $\mathcal{X}$ be a finite set and $K(x,y)$ a Markov Chain indexed by $\mathcal{X}$. If there exists $n^\star$ such that $K^{n^\star}(x,y) > 0$ for all $n \geq n^\star$, then K has a unique stationary distribution π and

$$\lim_{n\to\infty} K^n(x,y) = \pi(y) \quad \text{for each } x \text{ e } y \text{ in } \mathcal{X}.$$

A Markov Chain satisfying the theorem conditions is said to be *ergodic*. The probabilistic content of the theorem is that starting from any initial state x, the n-th step of a simulation of the chain has probability close to $\pi(y)$ of being in y if n is large. A key observation is that, tipically, in the application of the method we are about to describe, $|\mathcal{X}|$ is large; it is simple to go from x to y according to $K(x, y)$; and it is hard to sample directly from π [Dia08].

As an example, consider the edge-triangle ERG. The state space $\mathcal{X}$ is the set $\mathcal{G}_n$ of all $2^{\binom{n}{2}}$ labelled graphsgrafos with n vertices, and the normalizing constant of the model is

$$e^{-\psi} = \sum_{G \in \mathcal{G}_n} p_\beta(G),$$

which has a number of terms exponential in n.

4.2 Markov Chain Monte Carlo method

The *Markov Chain Monte Carlo method* (MCMC) is a technique used to sample from some probability distribution π. It has the advantage of not requiring the calculation of the normalizing constant (π needs only to be known up to a multiplicative constant), and its application in this study requires only the computation of the Hamiltonian of a limited number of graphs. The MCMC method consists in simulating a Markov Chain having p_β as stationary distribution. To obtain an approximate sample from p_β, we read the state of the chain after a large enough number of steps is taken.

Let $\mathcal{X}$ be a finite set and $\pi(x)$ a probability distribution over $\mathcal{X}$, known up to a normalizing constant. Let $J(x, y)$ be a transition matrix of a Markov Chain over $\mathcal{X}$ such that $J(x, y) > 0$ if and only if $J(y, x) > 0$, and let $A(x, y) \doteq \pi(y)J(y, x)/\pi(x)J(x, y)$. Note that J does not need to be related to π. We call J the *proposal chain*, and A the *accepting ratio*. The *Metropolis–Hastings* algorithm transforms J in a new Markov Chain $K(x, y)$ with stationary distribution π. The algorithmic description of the transformation is the following: from x, choose y with probability $J(x, y)$; if $A(x, y) \geq 1$, proceed to state y; otherwise, throw a coin with probability of "heads" $A(x, y)$: if the coin falls "heads", proceeed to state y and stay in x otherwise. In symbols,

$$K(x, y) \doteq \begin{cases} J(x, y) & \text{if } x \neq y \text{ and } A(x, y) \geq 1, \\ J(x, y)A(x, y) & \text{if } x \neq y \text{ and } A(x, y) < 1, \\ J(x, y) + \sum_{z:\ A(x,z)<1} J(x, z)(1 - A(x, z)) & \text{if } x = y. \end{cases} \tag{4.3}$$

Note that the normalizing constant is cancelled in the calculations. The accepting ratio is such that the chain K satisfies $\pi(x)K(x, y) = \pi(y)K(y, x)$; this implies K has a stationary distribution π. In fact,

$$\sum_{x \in \mathcal{X}} \pi(x)K(x, y) = \sum_{x \in \mathcal{X}} \pi(y)K(y, x) = \pi(y) \sum_{x \in \mathcal{X}} K(y, x) = \pi(y),$$

and thus K and π satisfy (4.2). When the proposal chain is simmetric (that is, $J(x, y) = J(y, x)$), the algorithm of equation (4.3) is called *Metropolis*.

Given a proposal chain, the essential question is determining the "speed" of convergence to the stationary distribution, that is, how many steps of the simulation are necessary for the current state distribution to be "close" to the stationary distribution. This value is called *mixing time* of the chain.

4.3 Mixing time of ERGs

We can evaluate the distance between the Markov Chain $K(x, y)$ and its stationary distribution π using the *total variation distance*

$$\|K_x^n - \pi\|_{\mathrm{TV}} \doteq \sup_{A \subseteq \mathcal{X}} |K^n(x, A) - \pi(A)|, \tag{4.4}$$

where $K^n(x, A) = \sum_{y \in A} K^n(x, y)$ and $\pi(A) = \sum_{y \in A} \pi(y)$. If $\mathcal{X}$ is finite, we have

$$\|K_x^n - \pi\|_{\mathrm{TV}} = \max_{A \subseteq \mathcal{X}} |K^n(x, A) - \pi(A)| \tag{4.5}$$

$$= \frac{1}{2} \sum_{y \in \mathcal{X}} |K^n(x, y) - \pi(y)|. \tag{4.6}$$

To demonstrate (4.6), consider the set $A^\star$ which attains the maximum

$$|K^n(x, A^\star) - \pi(A^\star)| = \max_{A \subseteq \mathcal{X}} |K^n(x, A) - \pi(A)|,$$

and let $B \doteq \mathcal{X} \setminus A^\star$. We have

$$K^n(x, \mathcal{X}) = K^n(x, A^\star) + K^n(x, B) = \pi(A^\star) + \pi(B) = \pi(\mathcal{X}) = 1$$

$$K^n(x, A^\star) - \pi(A^\star) = -(K^n(x, B) - \pi(B)). \tag{4.7}$$

Also, note that if $a \in A^\star$ then the sign of $K^n(x, a) - \pi(a)$ is the same as the sign of $K^n(x, A^\star) - \pi(A^\star)$, for otherwise, if $A' \doteq A \setminus \{a\}$,

$$|K^n(x, A') - \pi(A')| > |K^n(x, A^\star) - \pi(A^\star)|.$$

The same reasoning shows that there is no $b \in B$ such that the sign of $K^n(x, b) - \pi(b)$ is the same as the sign of $K^n(x, A^\star) - \pi(A^\star)$, for otherwise, if $A' \doteq A \cup \{b\}$,

$$|K^n(x, A') - \pi(A')| > |K^n(x, A^\star) - \pi(A^\star)|.$$

In both cases, the set A' contradict the choice of $A^\star$.

All this shows that we can decide whether a given element $y \in \mathcal{X}$ belongs to $A^\star$ (or to B) by checking the signal of $K^\star(x, y) - \pi(y)$, which amounts to say that

$$\Big|\sum_{y \in B} K^n(x, y) - \pi(y)\Big| = \sum_{y \in B} |K^n(x, y) - \pi(y)|,$$

and also

$$\Big|\sum_{y \in A^\star} K^n(x, y) - \pi(y)\Big| = \sum_{y \in A^\star} |K^n(x, y) - \pi(y)|.$$

Therefore, by (4.7)

$$
\begin{aligned}
|K^n(x, A^\star) - \pi(A^\star)| &= |K^n(x, B) - \pi(B)| \\
2|K^n(x, A^\star) - \pi(A^\star)| &= |K^n(x, A^\star) - \pi(A^\star)| + |K^n(x, B) - \pi(B)| \\
2 \max_{A \subseteq \mathcal{X}} |K^n(x, A) - \pi(A)| &= \sum_{y \in A^\star} |K^n(x, y) - \pi(y)| + \sum_{y \in B} |K^n(x, y) - \pi(y)| \\
\max_{A \subseteq \mathcal{X}} |K^n(x, A) - \pi(A)| &= \frac{1}{2} \sum_{y \in \mathcal{X}} |K^n(x, y) - \pi(y)|.
\end{aligned}
$$

The expression $\|K_x^n - \pi\|_{\mathrm{TV}}$ is a number between 0 and 1, and we are interested in the following problem: given K, π, x and $\epsilon > 0$, how large must be n so that

$$
\|K_x^n - \pi\|_{\mathrm{TV}} < \epsilon. \tag{4.8}
$$

The *mixing time* of K is the smallest time $n^\star$ such that $\max_{x \in \mathcal{X}} \|K_x^{n^\star} - \pi\|_{\mathrm{TV}} < e^{-1}$. Another way of limiting the mixing time is using *coupling of chains* [Dia08]. In that technique, two Markov Chain processes evolve simultaneously, according to the transition operator K, until the point they meet: from then they become "coupled" and proceed together.

Formally, the coupling of the chains X and Y, defined over the state space $\mathcal{X}$ is a process $Z_n = (X_n, Y_n)$ over the state space $\mathcal{X} \times \mathcal{X}$ such that each of the coordinates is marginally distributed as a Markov process K. That is, writing $Q\big((i,j),(i',j')\big) \doteq \mathrm{P}\big((X_{n+1}, Y_{n+1}) = (i',j') \mid (X_n, Y_n) = (i,j)\big)$,

$$
\sum_{j' \in \mathcal{X}} Q\big((i,i'),(j,j')\big) = K(i,j) \quad \text{and} \quad \sum_{j \in \mathcal{X}} Q\big((i,i'),(j,j')\big) = K(i',j').
$$

As an example, consider two chains, one of wich starts from a random state taken according to the stationary distribution and another which starts from some fixed state. Since the stationary chain is stationary at every step, an upper bound to the mixing time can be obtained estimating the number of transitions until coupling. We have the following useful Lemma.

Lemma 4.2 (Mixing time Lemma). For a Markov chain K, suppose that there are two coupled copies, Y and Z, such that each has marginal distribution X and

$$
\max_{y,z} \mathrm{P}(Y_t \neq Z_t \,|\, Y_0 = y, Z_0 = z) \leq (2e)^{-1}.
$$

The mixing time of X is bounded above by t.

That is, if the probability of "non-coupling" at the time t is suficiently low, then t is an upper bound for the mixing time (for a proof and further discussion, see [LPW09]).

We mention here the results of Bhamidi, Bresler and Sly [BBS11], about the mixing time of ERGs from equation (4.9). Their results hold for proposal schemes called *local dinamics*, where the proposal chain only allows transition between graphs which differ in at most the state of at most $o(n)$ edges. If we allow the change of at most one edge at per step of the chain, then the proposal scheme is called *Glauber dynamics*.

These authors study the (subclass of) ERGs whose distribution may be expressed as

$$p_{n,\beta}(X) = \exp\left(\sum_{i=1}^{k} \beta_i \frac{\mathrm{dens}(G_i, X)}{n^{e(G_i)-2}} - \psi_n(\beta)\right), \tag{4.9}$$

where G_i are graphs with $e(G_i)$ edges each, for $i = 1, \ldots, k$, and $\mathrm{dens}(G_i, X)$ is the number of labelled copies of G_i in X (that is the number of edge-preserving injections $V(G_i) \to [n]$ from the set of vertices of G_i to distinct vertices of X).

Bhamidi, Bresler and Sly obtained a characterization of the behaviour of local dinamics: when $\beta_2, \ldots, \beta_k$ are positive (i.e., $\beta = (\beta_1, \beta_2, \ldots, \beta_k) \in \mathbb{R} \times (\mathbb{R}_+)^{k-1}$), if n is large enough, then

- or the model is essentially the same as some BRG, and the mixing time of the Markov chain is $n^2 \log n$;
- or the Markov chain takes an exponential number o steps to mix.

Chapter 5

Graph Limits

Chatterjee and Diaconis [CD11] compare binomial random graphs and exponential random graphs using the theory of graph limits developed in a series of articles by L. Lovász, V.T. Sós, B. Szegedy, C. Borgs, J. Chayes, K. Vesztergombi, A. Schriver and M. Freedman (see [BCL+06, BCL+08, BCL+, LS06, Lov12]). In these studies, large graphs are compared using subgraph counts. In this section we briefly outline some of the ideas involved in the comparison of the models, referring the interested reader to the aforementioned publications.

A *graph homomorphism* is an edge-preserving function $f\colon V(G) \to V(H)$ from vertices of one graph G to the vertex set of another graph H. That is, whenever $ij \in E(G)$, we have $f(i)f(j) \in E(H)$. For any graphs G, H, we denote by $|\hom(G,H)|$ the number of homomorphisms from G to H (that is, the number of functions $V(G) \to V(H)$ between the vertex sets of G and of H such that every pair connected vertices of H is mapped to a pair of connected vertices in G. We also define the *homomorphism density*

$$\mathrm{dens}(H,G) \doteq \frac{|\hom(H,G)|}{|V(G)|^{|V(H)|}} \tag{5.1}$$

which is the probability a function $V(H) \to V(G)$ chosen uniformly at random is a homomorphism.

Let $\{G_n\}_{n\geq 1}$ be a sequence of graphs such that the number of vertices of G_n tends to infinity as $n \to \infty$. Suppose the graphs G_n become more similar as n increases, in the sense $\mathrm{dens}(H, G_n)$ tends to a limit $\mathrm{dens}(H)$ for every graph H. One of the results of the work of Lovász and coworkers is the identification of a limit object to such sequences.

The *graph limit* of the sequence, or *graphon*, is an object from which the values of $\mathrm{dens}(H)$ may be read. Also, it is a fact that every graphon (see definition below) is the limit of some graph sequence [LS06].

The limit objects are functions $h \in \mathcal{W}$, where $\mathcal{W}$ is the space of all measurable functions from $[0,1]^2$ to $[0,1]$ which satisfy $h(x,y) = h(y,x)$, for all x, y.

The graphon determines all subgraph limit deensities: let H be a graph with vertex set $V(H) = [k]$ and

$$\mathrm{dens}(H,h) = \int_{[0,1]^k} \prod_{\{i,j\}\in E(H)} h(x_i,x_j)\,\mathrm{d}x_1 \ldots \mathrm{d}x_k. \tag{5.2}$$

We say that a sequence of graphs $\{G_n\}_{n\geq 1}$ *converges* to h if for every graph H

$$\lim_{n\to\infty} \operatorname{dens}(H, G_n) = \operatorname{dens}(H, h). \tag{5.3}$$

For a fixed graph G, we define

$$f^G(x,y) = \begin{cases} 1 & \text{if } (\lceil nx \rceil, \lceil ny \rceil) \text{ is an edge of } G, \\ 0 & \text{otherwise.} \end{cases}. \tag{5.4}$$

It follows that the limit of the constant sequence $G, G, \ldots$ is f^G, that is $\operatorname{dens}(H, f^G) = \operatorname{dens}(H, G)$ for every graph H.

In addition, there is a metric in the space of graphons (defined through a distance, called *cut distance*) which has the following property: graphons h_1 and h_2 which are close have values $\operatorname{dens}(F, h_1)$ and $\operatorname{dens}(F, h_2)$ similar, for every fixed graph F. A more comprehensive discussion of the subject would escapes the escope of this text, and we refer the interested reader to the aforementioned literature.

Consider two graph sequences $\{F_k\}_{k\geq 1}$ and $\{G_\ell\}_{\ell\geq 1}$, obtained, respectively, according to some BRG model, and some ERG model, with $|V(F_i)| = |V(G_i)| = i$; and let α (resp. β) be the BRG (resp. ERG) model used to generate the sequence $\{F_k\}_{k\geq 1}$ (resp. $\{G_\ell\}_{\ell\geq 1}$). Suppose that the sequences have (graph) limits f and g, respectively. Note that f and g are random variables.

We can compare the models observing the expected distance between f and g. We shall say two models of random graphs α and β are *close*, or *similar* if the cut distance $\operatorname{d}(G_\alpha, G_\beta)$, between the respective graphons is arbitrarily small almost-surely. That is, in symbols símbolos, for all $\varepsilon > 0$,

$$\operatorname{P}\big(\operatorname{d}(G_\alpha, G_\beta) < \varepsilon\big) = 1. \tag{5.5}$$

Extending results of Bhamidi, Bresler and Sly, Chatterjee and Diaconis [CD11], observe that many ERG models are close to some BRG. In the particular case of the edge-triangle model with $\beta_2 > 0$, Chatterjee e Diaconis have determined a means of estimating the parameter $u^\star = u^\star(\beta_1, \beta_2)$ (equation (3.21)) such that $G(n, u^\star)$ is close to the edge-triangle model with n vertices and parameters β_1, β_2 (equation (3.16)).

Chapter 6

Computational Experiments

The similarity of the BRG and ERG models (in the sense of chapter 5) has an asymptotic nature, steeming from its definition in terms of convergence of the homomorphism densities in infinite sequences of graphs. In a finite setting, we expect to find similar homomorphism densities $\operatorname{dens}(H, F) \approx \operatorname{dens}(H, G)$, for graphs F, G with number of vertices n sufficiently large and H with $|V(H)| \ll n$ vertices; where F, G are sampled, respectively, according to the edge-triangle model (equation (3.16)) and the "corresponding" BRG (equation (3.1), with $p = u^\star$ maximizing (3.21)).

In this section we describe an exploratory study of the behavior of $\operatorname{dens}(K_i, G)$, for $i = 2, 3$ and for G sampled from the edge-triangle ERG model and the BRG model. Note that $\operatorname{dens}(K_i, G) = i! \cdot \big|\{K_i \subseteq G\}\big|$ is proporcional to the number of subgraphs K_i of G. Therefore, we can measure the homomorphism density by simply counting the number of copies of K_i in the sampled graphs.

On one hand, the size (number of vertices) of graphs we can sample is limited, given that the mixing time of the Metropolis-Hastings algorithm may be exponential on the number of vertices of the model. On the other hand, small graphs may have very different homomorphism densities, if only because the theorems motivanting our simulations are asymptotic statements.

However, extensive simulation using computers, reported by Mark Handcock and David Hunter, indicates that $n = 20$ vertices already ensure a good approximation of $\psi_n(\beta_1, \beta_2)$ using equation (3.21) (see [CD11]), suggesting that the asymptotic behavior can be observed in models within reach of simulation.

Our simulations were made using the package ergm of statistical tools for analysis of networks [HHB+13, HHB+08] (see section 6.1). We used this implementation to sample "close" ERG and BRG.

We use the following heuristics to choose the number of steps to execute in the simulation: for models with n vertices, we sampled the Markov Chain state every $n^2 \approx 2\binom{n}{2}$ steps. We chose this number since any two n-vertex graphs differ in at most $\binom{n}{2}$ edges, and, in particular, the graphs $([n], E)$ and $([n], \binom{[n]}{2} \setminus E)$ differ in *exatcly* $\binom{n}{2}$ edges. We exhibit the relation between the state space size and the number of transitions between sampling on the table 6.1.

Table 6.1: Growth of $\binom{n}{2}$ as a function of n. The *order* of x is $10^{\lfloor \log_{10} x \rfloor}$.

n	$\binom{n}{2}$	order of $\binom{n}{2}$	order of $2^{\binom{n}{2}}$
5	10	10^1	10^3
10	45	10^1	10^{13}
15	105	10^2	10^{31}
20	190	10^2	10^{57}
25	300	10^2	10^{90}
30	435	10^2	10^{130}
35	595	10^2	10^{179}
40	780	10^2	10^{234}
45	990	10^2	10^{298}
50	1 225	10^3	10^{368}
100	4 950	10^3	10^{1490}

6.1 Software for the simulation

The ergm package for the R suite of statistical software [HHB+08] provides, among other tools, an implementation of the Metropolis–Hastings algorithm for simulation of ERGs via MCMC (see section 4.2). The package allows configuration of important parameters such as which terms constitute the model (for instance: number of edges, stars, triangles, etc.); their respective coefficients (the vector β of equation (3.3)); the number of steps taken before the first sample is taken ("burn-in" steps); and the number of steps to take between samples. The package is extensible, and it is possible to create new terms or proposal chains other then the Glauber dynamics.

6.2 Experiments

As previously discussed, to each parametrization of the edge-triangle model corresponds a unique value $u^\star = u^\star_{\beta_1,\beta_2}$ which attains the maximum on equation (3.21). This value is such that BRG $G(n, u^\star)$ is asymptotically close (in the sense of section 5) to the edge-triangle model p_{n,β_1,β_2} (see section 5 and figure 6.1).

We have sampled 100 graphs from the edge-triangle model p_{n,β_1,β_2}, for every pair

$$(\beta_1, \beta_2), \quad \text{for} \quad \beta_1, \beta_2 \in \{0\,,\, 0.2\,,\, 0.4\,,\, 0.6\,,\, 0.8\,,\, 1\}; \quad \text{and}$$

with $n = 5, 10, 15, \ldots, 50$ and 100 vertices. For each choice of these parameters (n, β_1, β_2), we also sampled 100 BRGs $G(n, u^\star)$. The Table A.3 shows the corresponding values of $u^\star$. The Metropolis–Hastings algorithm was configured for n^2 steps before the first sample, and also n^2 steps between samples. We recall that the size of the state space of the simulated chain is approximately 2^{n^2} (see Table 6.1).

To quantify the proximity of the obtained samples, we chose the number of edges (K_2) and triangles (K_3), as well as the correlation between adjacent and independent edges (the correlation between any edges should be zero for BRGs). The normalized counts of triangles (# triângulos)$/\binom{n}{3}$ have been used to compare samples of graphs with distinct numbers of vertices.

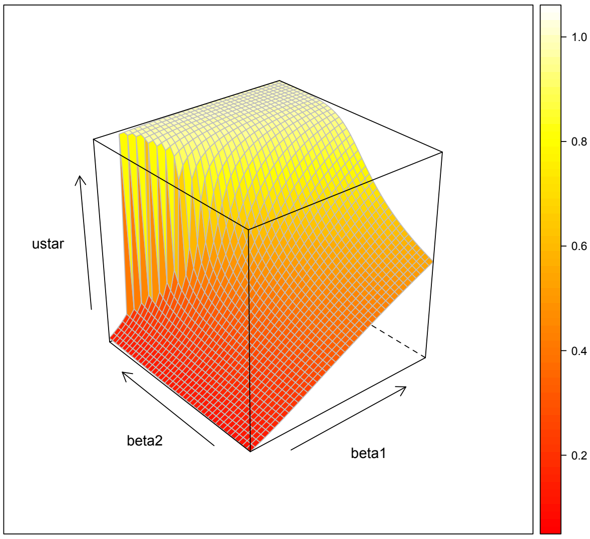

Figure 6.1: Graphic of $u^\star$ as a function of β_1 and β_2, for $\beta_1, \beta_2 \in [0, 1]$. See section 3.4, equation (3.21) and table A.3.

The Tables A.1 and A.2 present a summary of some of the statistics calculated from the samples. We calculate sample averages and sample deviations of the number of triangles and edges of the samples.

6.3 Comparison

Consider the edge-triangle models of parameters $\beta_1 = 0.2$, $\beta_2 = 0.2$. On Table A.3, we see the value of $u^\star$ which attains the maximum on equation (3.21) ($u^\star = 0.743$), and the value of its cube ($u^{\star 3} = 0.4106$). These values are very close to the edge and triangle densities sampled using the edge-triangle model (see Table 6.2). This phenomenon can be observed in most of the simulated parametrizations, and suggests the existence of a similarity between the homomorphism densities exhibited by both models. This, however, does not seem to be the case for edge correlations: although simulated BRG samples exhibit almost zero correlation between edges, the corresponding ERG model does not *seem* to be consistent in this respect. The simulations, however, are not conclusive in this respect.

Table 6.2: Some sampled values: edge density $\overline{e}/\binom{n}{2}$ and triangle density $\overline{t}/\binom{n}{3}$ of the models (BRG and edge-triangle ERG, and their parameters.

					BRG		ERG	
n	β_1	β_2	$u^\star$	$u^{\star 3}$	$\overline{e}/\binom{n}{2}$	$\overline{t}/\binom{n}{3}$	$\overline{e}/\binom{n}{2}$	$\overline{t}/\binom{n}{3}$
20	0.2	0.2	0.743	0.4106	0.7447	0.4133	0.7330	0.3990
25	0.2	0.2	0.743	0.4106	0.7427	0.4100	0.7303	0.3956
30	0.2	0.2	0.743	0.4106	0.7402	0.4057	0.7260	0.3867
100	0.2	0.2	0.743	0.4106	0.7430	0.4102	0.7352	0.3993
20	0.8	0.4	0.980	0.942	0.9805	0.9439	0.9726	0.9237
25	0.8	0.4	0.980	0.942	0.9807	0.9435	0.9707	0.9191
30	0.8	0.4	0.980	0.942	0.9811	0.9441	0.9722	0.9227
100	0.8	0.4	0.980	0.942	0.8020	0.9419	0.9729	0.9252

Chapter 7

Discussion and final comments

We have studied some known results about binomial and exponential random graphs, and performed some simulations of both. The study was exploratory, and the choice of parameterizations was motiated by similarities of subgraph densities the models seem to display [BBS11, CD11].

On another direction, it would be interesting to investigate parameterizations of ERGs which *differ* from BRGs in the same respect [AR13, BHLN13]. For instance, it is known that samples of the edge-triangle model with parameter β_2 sufficiently negative (that is, that "forbid" triangles) have typically smaler odd-length cycles than what would be expected of a BRG of same edge density [CD11]. This suggests an experiment in which we measure how "bipartite" are the sampled graphs: that can be done, for example, by calculating the maximum umber of edges of a bipartite subgraphs, the graph's maximum cut size (see Figure 7.1). This number may be compared to the expected size of a maximum cut of a BRG with same edge density $p = e(G)/\binom{|V(G)|}{2}$ (which is, asymptotically, $n^2p/4$). However, calculating the exact size of a graph's maximum cut is a computationally complex (NP-complete) problem, which needs better estimating than what we have done here. (Observe that, in general, there is a constant c, positive, such that every graph with $2m^2$ edges has a bipartite subgraph with at least $m^2 + m/2 + c\sqrt{m}$ edges; and that a triangle-gree graph with $e > 1$ edges has a bipartite subgraph with at least $e/2 + c'e^{4/5}$, for some positive constant c' [Alo96].)

Finally, we indicate two important topics, related to sampling in general, which unfortunately could not be covered in this study with the necessary detail. The first one is the matter of assessing the quality of the obtained samples, estimating how close we are to the stationary distribution of the ERG model); and the second is the choice of suitable statistical procedures to employ when comparing the densities and correlations observed.

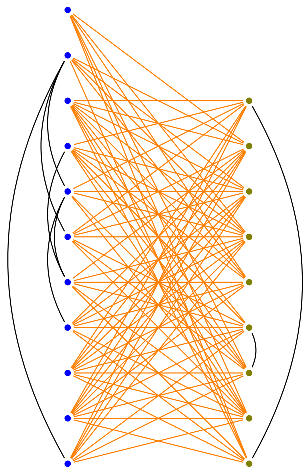

Figure 7.1: Graph generated by the edge-triangle model with parameters $n = 20$, $\beta_1 = 100$, and $\beta_2 = -200$. The 77 edges connecting vertices "of the right" to vertices "of the left" form the largest cut of the graph.

Bibliography

[AB99] Reka Albert and Albert-László. Barabási. Emergence of Scaling in Random Networks. *Science*, 286:509–512, October 1999.

[AB02] Reka Albert and Albert-László. Barabási. Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74:47–97, January 2002.

[Alo96] Noga Alon. Bipartite Subgraphs. *Combinatorica*, 16:301–311, 1996.

[AR13] David Aristoff and Charles Radin. Emergent structures in large networks. *Journal of Applied Probability*, 50(3):883–888, 2013.

[BBS11] Shankar Bhamidi, Guy Bresler, and Allan Sly. Mixing time of exponential random graphs. *Ann. Appl. Stat.*, 21(6):2146–2170, 2011.

[BCL$^+$] Christian Borgs, Jennifer T. Chayes, László Lovász, Vera T. Sós, and Katalin Vesztergombi. Convergent Sequences of Dense Graphs II. Multiway Cuts and Statistical Physics.

[BCL$^+$06] Christian Borgs, Jennifer T. Chayes, László Lovász, Vera T. Sós, and Katalin Vesztergombi. Counting graph homomorphisms. In Martin Klazar, Jan Kratochvíl, Martin Loebl, Jiří Matoušek, Pavel Valtr, and Robin Thomas, editors, *Topics in Discrete Mathematics*, volume 26 of *Algorithms and Combinatorics*, pages 315–371. Springer Berlin Heidelberg, 2006.

[BCL$^+$08] Christian Borgs, Jennifer T. Chayes, László Lovász, Vera T. Sós, and Katalin Vesztergombi. Convergent sequences of dense graphs I: Subgraph frequencies, metric properties and testing. *Advances in Mathematics*, 219:1801–1851, 2008.

[Bes74] Julian Besag. Spatial Interaction and the Statistical Analysis of Lattice Systems. *Journal of the Royal Statistical Society. Series B (Methodological)*, 36(2):192–236, 1974.

[BHLN13] Shankar Bhamidi, Jan Hannig, Chia Ying Lee, and James Nolen. The importance sampling technique for understanding rare events in Erdős–Rényi random graphs. *ArXiv e-prints*, February 2013.

[Bol98] Béla Bollobás. *Modern graph theory*, volume 184 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1998.

[Bol01] Béla Bollobás. *Random Graphs*. Cambridge University Press, 2001.

[Bré99] Pierre Brémaud. *Markov Chains: Gibbs Fields, Monte Carlo Simulation, and Queues*. Texts in Applied Mathematics. Springer, 1999.

[CD11] Sourav Chatterjee and Persi Diaconis. Estimating and Understanding Exponential Random Graph Models. *ArXiv e-prints*, August 2011.

[CDS10] Sourav Chatterjee, Persi Diaconis, and Allan Sly. Random graphs with a given degree sequence. *ArXiv e-prints*, May 2010.

[Dia08] Persi Diaconis. The Markov chain Monte Carlo revolution. *Bulletin of the American Mathematical Society*, 46(2):179–205, November 2008.

[Die10] Reinhard Diestel. Graph Theory. 173:451, 2010.

[ER60] Paul Erdős and Alfréd Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hungary. Acad. Sci.*, 5:17–61, 1960.

[FS86] Ove Frank and David Strauss. Markov Graphs. *Journal of the American Statistical Association*, 81(395):832–842, 1986.

[Gri] Geoffrey Grimmett. Probability on graphs. Lecture Notes on Stochastic Processes on Graphs and Lattices. Statistical Laboratory, University of Cambridge.

[HHB+08] David R. Hunter, Mark S. Handcock, Carter T. Butts, Steven M. Goodreau, Morris, and Martina. ergm: A package to fit, simulate and diagnose exponential-family models for networks. *Journal of Statistical Software*, 24(3), 2008.

[HHB+13] Mark S. Handcock, David R. Hunter, Carter T. Butts, Steven M. Goodreau, Pavel N. Krivitsky, and Martina Morris. *ergm: Fit, Simulate and Diagnose Exponential-Family Models for Networks*. The Statnet Project (http://www.statnet.org), 2013. R package version 3.1-0.

[HJ99] Olle Häggström and Johan Jonasson. Phase transition in the random triangle model. *Journal of Applied Probability*, 36(4):1101–1115, 1999.

[HL81] Paul W. Holland and Samuel Leinhardt. An Exponential Family of Probability Distributions for Directed Graphs. *Journal of The American Statistical Association*, 76:33–50, 1981.

[JŁR00] Svante Janson, Tomasz Łuczak, and Andrzej Ruciński. *Random graphs*. Wiley-Interscience series in discrete mathematics and optimization. John Wiley, 2000.

[Lov12] L. Lovász. *Large Networks and Graph Limits*. American Mathematical Society colloquium publications. American Mathematical Society, 2012.

[LPW09] David A. Levin, Yuval Peres, and Elizabeth L. Wilmer. *Markov Chains and Mixing Times*. American Mathematical Society, 2009.

[LS06] László Lovász and Balázs Szegedy. Limits of dense graph sequences. *J. Combin. Theory Ser. B*, pages 933–957, 2006.

[SPRH06] Tom A. B. Snijders, Philippa E. Pattison, Garry L. Robins, and Mark S. Handcock. New specifications for exponential random graph models. *Sociological Methodology*, 36(1):99–153, 2006.

[vdH09] Remco van der Hofstad. Random graphs and complex networks. Available at http://www.win.tue.nl/~rhofstad/NotesRGCN.pdf, 2009. Lecture Notes.

[WS98] Duncan J. Watts and Steven H. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393(6684):440–442, June 1998.

Appendix A

Tables of computational experiments

The tables A.1 and A.2 present a selection of our simulation results. The table A.3 presents some points of the function $(\beta_1, \beta_2) \mapsto u^\star$ relating the ege-triangle and the similar BRG model.

Table A.1: Statistical summary of the edge-triangle ERG samples. The columns represent the number of vertices n, the parametrs β_1 and β_2 of the model (see equation (3.16)), and, for vertices $v_1, \ldots, v_4$ uniformly and randomly chosen (and fixed for each 100 samples), the correlation and covariance between the pairs of edges e_{12}, e_{13} and e_{12}, e_{34}.

n	β_1	β_2	$\bar{e}$	s_e	$\bar{t}$	s_t	$\bar{e}/\binom{n}{2}$	$\bar{t}/\binom{n}{3}$	Corr e_{12}, e_{13}	Corr e_{12}, e_{34}
20	0	0	0.9559e2	0.706e1	0.1449e3	0.347e2	0.5031	0.1271	0.196	-0.817e-2
20	0	0.2	0.1121e3	0.936e1	0.241e3	0.592e2	0.59	0.2114	-0.241e-1	0.578e-1
20	0	0.4	0.1477e3	0.111e2	0.5512e3	0.11e3	0.7774	0.4835	-0.116	0.714e-1
20	0	0.6	0.1788e3	0.95e1	0.9575e3	0.11e3	0.9411	0.8399	-0.602e-1	0.316
20	0	0.8	0.1852e3	0.101e2	0.1065e4	0.119e3	0.9747	0.9342	-0.292e-1	-0.144e-1
20	0	1	0.1878e3	0.914e1	0.1108e4	0.107e3	0.9884	0.9719	NaN	NaN
20	0.2	0	0.1149e3	0.626e1	0.2533e3	0.43e2	0.6047	0.2222	-0.107	0.179
20	0.2	0.2	0.1393e3	0.795e1	0.4543e3	0.713e2	0.7332	0.3985	-0.118	-0.561e-1
20	0.2	0.4	0.169e3	0.922e1	0.81e3	0.113e3	0.8895	0.7105	-0.111	-0.116
20	0.2	0.6	0.182e3	0.914e1	0.1009e4	0.109e3	0.9579	0.8851	-0.292e-1	-0.292e-1
20	0.2	0.8	0.1871e3	0.882e1	0.1094e4	0.105e3	0.9847	0.9596	NaN	-0.101e-1
20	0.2	1	0.1884e3	0.807e1	0.1117e4	0.988e2	0.9916	0.9798	NaN	NaN
20	0.4	0	0.1308e3	0.821e1	0.3719e3	0.706e2	0.6884	0.3262	0.131	0.202e-1
20	0.4	0.2	0.1536e3	0.849e1	0.6081e3	0.954e2	0.8084	0.5334	0.356e-1	-0.114
20	0.4	0.4	0.1757e3	0.799e1	0.9073e3	0.1e3	0.9247	0.7959	-0.807e-1	-0.111
20	0.4	0.6	0.1853e3	0.839e1	0.1063e4	0.104e3	0.9753	0.9325	-0.251e-1	-0.251e-1
20	0.4	0.8	0.1878e3	0.772e1	0.1106e4	0.969e2	0.9884	0.9702	NaN	NaN
20	0.4	1	0.1886e3	0.68e1	0.1119e4	0.89e2	0.9926	0.9816	NaN	NaN
20	0.6	0	0.1439e3	0.801e1	0.4975e3	0.764e2	0.7574	0.4364	-0.139	0.836e-1
20	0.6	0.2	0.1674e3	0.82e1	0.7852e3	0.943e2	0.8811	0.6888	-0.117	0.309
20	0.6	0.4	0.181e3	0.848e1	0.991e3	0.104e3	0.9526	0.8693	-0.629e-1	0.158
20	0.6	0.6	0.1864e3	0.788e1	0.1081e4	0.988e2	0.9811	0.9482	0.492	-0.101e-1
20	0.6	0.8	0.1878e3	0.851e1	0.1106e4	0.107e3	0.9884	0.9702	NaN	NaN
20	0.6	1	0.1885e3	0.833e1	0.1119e4	0.102e3	0.9921	0.9816	NaN	NaN
20	0.8	0	0.157e3	0.772e1	0.6445e3	0.814e2	0.8263	0.5654	-0.278e-2	0.116
20	0.8	0.2	0.1744e3	0.87e1	0.8883e3	0.105e3	0.9179	0.7792	0.	-0.765e-1
20	0.8	0.4	0.1848e3	0.777e1	0.1053e4	0.997e2	0.9726	0.9237	0.366	-0.101e-1
20	0.8	0.6	0.1871e3	0.679e1	0.1092e4	0.904e2	0.9847	0.9579	-0.205e-1	NaN
20	0.8	0.8	0.1885e3	0.804e1	0.1118e4	0.1e3	0.9921	0.9807	0.438	-0.144e-1
20	0.8	1	0.1889e3	0.644e1	0.1124e4	0.87e2	0.9942	0.986	NaN	NaN
20	1	0	0.1664e3	0.63e1	0.7679e3	0.761e2	0.8758	0.6736	-0.157	0.112e-2
20	1	0.2	0.1799e3	0.719e1	0.9714e3	0.937e2	0.9468	0.8521	0.202	-0.177e-1
20	1	0.4	0.1855e3	0.686e1	0.1065e4	0.942e2	0.9763	0.9342	0.313	NaN
20	1	0.6	0.1878e3	0.75e1	0.1105e4	0.993e2	0.9884	0.9693	NaN	0.221

Table A.1: (continuation)

n	β_1	β_2	$\bar{e}$	s_e	$\bar{t}$	s_t	$\bar{e}/\binom{n}{2}$	$\bar{t}/\binom{n}{3}$	Corr e_{12}, e_{13}	Corr e_{12}, e_{34}
20	1	0.8	0.1885e3	0.786e1	0.1118e4	0.989e2	0.9921	0.9807	NaN	NaN
20	1	1	0.1886e3	0.926e1	0.1121e4	0.11e3	0.9926	0.9833	NaN	0.571
25	0	0	0.1507e3	0.842e1	0.2904e3	0.503e2	0.5023	0.1263	0.705e-1	-0.14
25	0	0.2	0.1786e3	0.101e2	0.4959e3	0.808e2	0.5953	0.2156	-0.311e-1	-0.121e-1
25	0	0.4	0.2396e3	0.138e2	0.1192e4	0.17e3	0.7987	0.5183	-0.123	-0.554e-1
25	0	0.6	0.2828e3	0.159e2	0.1943e4	0.227e3	0.9427	0.8448	0.219	-0.292e-1
25	0	0.8	0.294e3	0.159e2	0.2181e4	0.229e3	0.98	0.9483	NaN	NaN
25	0	1	0.2963e3	0.138e2	0.2229e4	0.214e3	0.9877	0.9691	NaN	NaN
25	0.2	0	0.1775e3	0.101e2	0.4777e3	0.8e2	0.5917	0.2077	0.107	0.225
25	0.2	0.2	0.2192e3	0.12e2	0.9099e3	0.139e3	0.7307	0.3956	0.126	0.212
25	0.2	0.4	0.2652e3	0.148e2	0.1607e4	0.222e3	0.884	0.6987	0.711e-1	-0.109
25	0.2	0.6	0.2903e3	0.142e2	0.2097e4	0.218e3	0.9677	0.9117	NaN	-0.177e-1
25	0.2	0.8	0.2948e3	0.15e2	0.2197e4	0.221e3	0.9827	0.9552	0.579	-0.328e-1
25	0.2	1	0.2969e3	0.14e2	0.2243e4	0.215e3	0.9897	0.9752	0.438	NaN
25	0.4	0	0.2051e3	0.953e1	0.7363e3	0.948e2	0.6837	0.3201	0.122	0.34e-1
25	0.4	0.2	0.2457e3	0.112e2	0.1275e4	0.149e3	0.819	0.5543	0.311	-0.663e-1
25	0.4	0.4	0.2783e3	0.122e2	0.1847e4	0.185e3	0.9277	0.803	0.113	0.565
25	0.4	0.6	0.2918e3	0.134e2	0.2129e4	0.207e3	0.9727	0.9257	NaN	-0.328e-1
25	0.4	0.8	0.2962e3	0.137e2	0.2226e4	0.209e3	0.9873	0.9678	NaN	0.1e1
25	0.4	1	0.2978e3	0.12e2	0.226e4	0.189e3	0.9927	0.9826	NaN	NaN
25	0.6	0	0.2293e3	0.981e1	0.103e4	0.117e3	0.7643	0.4478	0.4e-1	-0.176
25	0.6	0.2	0.2645e3	0.11e2	0.1583e4	0.163e3	0.8817	0.6883	-0.721e-1	-0.58e-1
25	0.6	0.4	0.2867e3	0.123e2	0.2018e4	0.191e3	0.9557	0.8774	0.144	-0.417e-1
25	0.6	0.6	0.2939e3	0.143e2	0.2176e4	0.219e3	0.9797	0.9461	-0.251e-1	0.49
25	0.6	0.8	0.2971e3	0.122e2	0.2245e4	0.195e3	0.9903	0.9761	NaN	NaN
25	0.6	1	0.2978e3	0.112e2	0.2259e4	0.187e3	0.9927	0.9822	NaN	NaN
25	0.8	0	0.2493e3	0.103e2	0.1324e4	0.146e3	0.831	0.5757	0.199	-0.299e-1
25	0.8	0.2	0.2763e3	0.114e2	0.1804e4	0.182e3	0.921	0.7843	-0.444e-1	0.328
25	0.8	0.4	0.2912e3	0.124e2	0.2114e4	0.196e3	0.9707	0.9191	0.163	-0.444e-1
25	0.8	0.6	0.2957e3	0.132e2	0.2215e4	0.208e3	0.9857	0.963	NaN	NaN
25	0.8	0.8	0.2973e3	0.131e2	0.225e4	0.205e3	0.991	0.9783	NaN	NaN
25	0.8	1	0.2981e3	0.116e2	0.2265e4	0.19e3	0.9937	0.9848	NaN	NaN
25	1	0	0.264e3	0.97e1	0.1572e4	0.145e3	0.88	0.6835	0.209e-1	0.112e-2
25	1	0.2	0.2841e3	0.111e2	0.1961e4	0.178e3	0.947	0.8526	0.158	-0.1
25	1	0.4	0.2938e3	0.11e2	0.2168e4	0.179e3	0.9793	0.9426	-0.292e-1	0.421
25	1	0.6	0.2966e3	0.114e2	0.2232e4	0.186e3	0.9887	0.9704	NaN	NaN
25	1	0.8	0.2977e3	0.112e2	0.2256e4	0.188e3	0.9923	0.9809	NaN	NaN
25	1	1	0.298e3	0.127e2	0.2265e4	0.203e3	0.9933	0.9848	NaN	NaN
30	0	0	0.2174e3	0.106e2	0.5087e3	0.737e2	0.4998	0.1253	0.101	-0.202e-1
30	0	0.2	0.2595e3	0.137e2	0.8801e3	0.132e3	0.5966	0.2168	-0.141	0.688e-1
30	0	0.4	0.354e3	0.191e2	0.2224e4	0.289e3	0.8138	0.5478	-0.103	-0.533e-1
30	0	0.6	0.41e3	0.236e2	0.3431e4	0.388e3	0.9425	0.8451	-0.526e-1	0.438
30	0	0.8	0.4263e3	0.193e2	0.3843e4	0.361e3	0.98	0.9466	-0.101e-1	-0.101e-1
30	0	1	0.4299e3	0.212e2	0.3943e4	0.395e3	0.9883	0.9712	NaN	NaN
30	0.2	0	0.26e3	0.114e2	0.8672e3	0.113e3	0.5977	0.2136	0.402e-1	0.135
30	0.2	0.2	0.3158e3	0.136e2	0.157e4	0.192e3	0.726	0.3867	0.144e-1	0.523e-1
30	0.2	0.4	0.3888e3	0.184e2	0.2921e4	0.318e3	0.8938	0.7195	-0.58e-1	-0.887e-1
30	0.2	0.6	0.4198e3	0.217e2	0.3673e4	0.389e3	0.9651	0.9047	-0.144e-1	-0.204e-1
30	0.2	0.8	0.4282e3	0.224e2	0.3899e4	0.398e3	0.9844	0.9603	0.571	0.1e1
30	0.2	1	0.431e3	0.197e2	0.3971e4	0.369e3	0.9908	0.9781	NaN	NaN
30	0.4	0	0.3009e3	0.105e2	0.1344e4	0.13e3	0.6917	0.331	-0.212	0.794e-1
30	0.4	0.2	0.3578e3	0.163e2	0.2275e4	0.26e3	0.8225	0.5603	0.249	-0.192
30	0.4	0.4	0.4073e3	0.176e2	0.3352e4	0.316e3	0.9363	0.8256	0.116e-1	-0.795e-1
30	0.4	0.6	0.4244e3	0.199e2	0.3792e4	0.37e3	0.9756	0.934	-0.204e-1	-0.204e-1
30	0.4	0.8	0.4296e3	0.184e2	0.393e4	0.357e3	0.9876	0.968	-0.292e-1	NaN
30	0.4	1	0.4312e3	0.198e2	0.3976e4	0.37e3	0.9913	0.9793	NaN	0.335
30	0.6	0	0.3332e3	0.131e2	0.1826e4	0.197e3	0.766	0.4498	0.236e-2	-0.441e-1
30	0.6	0.2	0.3848e3	0.151e2	0.2825e4	0.272e3	0.8846	0.6958	-0.127e-1	0.454e-2
30	0.6	0.4	0.4158e3	0.175e2	0.3563e4	0.332e3	0.9559	0.8776	NaN	NaN
30	0.6	0.6	0.4277e3	0.197e2	0.388e4	0.368e3	0.9832	0.9557	NaN	NaN
30	0.6	0.8	0.4306e3	0.189e2	0.3957e4	0.362e3	0.9899	0.9746	NaN	NaN
30	0.6	1	0.4318e3	0.184e2	0.3989e4	0.355e3	0.9926	0.9825	NaN	NaN
30	0.8	0	0.36e3	0.141e2	0.231e4	0.219e3	0.8276	0.569	-0.772e-1	-0.1
30	0.8	0.2	0.4003e3	0.161e2	0.3179e4	0.3e3	0.9202	0.783	0.116e-1	0.765e-1
30	0.8	0.4	0.4229e3	0.166e2	0.3746e4	0.322e3	0.9722	0.9227	-0.359e-1	NaN
30	0.8	0.6	0.4293e3	0.186e2	0.3923e4	0.358e3	0.9869	0.9663	-0.144e-1	-0.144e-1
30	0.8	0.8	0.4316e3	0.179e2	0.3983e4	0.344e3	0.9922	0.981	0.492	0.492
30	0.8	1	0.4322e3	0.167e2	0.3998e4	0.334e3	0.9936	0.9847	NaN	NaN
30	1	0	0.3809e3	0.127e2	0.273e4	0.223e3	0.8756	0.6724	0.148	0.264
30	1	0.2	0.4122e3	0.161e2	0.3469e4	0.3e3	0.9476	0.8544	-0.101e-1	-0.276e-1

Table A.1: (continuation)

n	β_1	β_2	$\overline{e}$	s_e	$\overline{t}$	s_t	$\overline{e}/\binom{n}{2}$	$\overline{t}/\binom{n}{3}$	Corr e_{12}, e_{13}	Corr e_{12}, e_{34}
30	1	0.4	0.4263e3	0.147e2	0.3834e4	0.302e3	0.98	0.9443	-0.101e-1	NaN
30	1	0.6	0.4303e3	0.173e2	0.3948e4	0.336e3	0.9892	0.9724	0.521	0.394
30	1	0.8	0.432e3	0.164e2	0.3992e4	0.333e3	0.9931	0.9833	NaN	NaN
30	1	1	0.4325e3	0.165e2	0.4005e4	0.328e3	0.9943	0.9865	NaN	NaN
100	0	0	0.2463e4	0.501e2	0.1993e5	0.115e4	0.4976	0.1233	0.6e-1	0.796e-1
100	0	0.2	0.299e4	0.816e2	0.3587e5	0.249e4	0.604	0.2218	0.929e-1	0.114
100	0	0.4	0.4119e4	0.2e3	0.9396e5	0.976e4	0.8321	0.5811	0.625e-1	-0.145
100	0	0.6	0.4717e4	0.265e3	0.1411e6	0.157e5	0.9529	0.8726	-0.292e-1	0.492
100	0	0.8	0.4859e4	0.258e3	0.154e6	0.161e5	0.9816	0.9524	NaN	-0.101e-1
100	0	1	0.4899e4	0.242e3	0.1577e6	0.155e5	0.9897	0.9753	NaN	NaN
100	0.2	0	0.296e4	0.634e2	0.346e5	0.197e4	0.598	0.214	0.543e-1	-0.276
100	0.2	0.2	0.3639e4	0.122e3	0.6457e5	0.511e4	0.7352	0.3993	-0.134	-0.442e-1
100	0.2	0.4	0.4483e4	0.217e3	0.1209e6	0.123e5	0.9057	0.7477	-0.602e-1	-0.87e-1
100	0.2	0.6	0.48e4	0.239e3	0.1484e6	0.151e5	0.9697	0.9177	0.398	0.202
100	0.2	0.8	0.4884e4	0.228e3	0.1562e6	0.148e5	0.9867	0.966	NaN	NaN
100	0.2	1	0.4906e4	0.229e3	0.1583e6	0.149e5	0.9911	0.979	0.438	0.1e1
100	0.4	0	0.3403e4	0.843e2	0.5261e5	0.323e4	0.6875	0.3254	-0.18e-1	0.989e-1
100	0.4	0.2	0.4102e4	0.153e3	0.9245e5	0.764e4	0.8287	0.5717	0.108	-0.292e-1
100	0.4	0.4	0.4657e4	0.218e3	0.1355e6	0.132e5	0.9408	0.838	-0.516e-1	-0.516e-1
100	0.4	0.6	0.4841e4	0.221e3	0.1521e6	0.144e5	0.978	0.9406	0.335	NaN
100	0.4	0.8	0.4896e4	0.217e3	0.1573e6	0.145e5	0.9891	0.9728	NaN	-0.177e-1
100	0.4	1	0.4911e4	0.215e3	0.1587e6	0.143e5	0.9921	0.9814	-0.101e-1	0.1e1
100	0.6	0	0.3792e4	0.114e3	0.729e5	0.519e4	0.7661	0.4508	-0.783e-1	0.316e-1
100	0.6	0.2	0.4404e4	0.171e3	0.1144e6	0.973e4	0.8897	0.7075	-0.586e-1	-0.403e-1
100	0.6	0.4	0.4758e4	0.207e3	0.1444e6	0.132e5	0.9612	0.893	-0.328e-1	-0.292e-1
100	0.6	0.6	0.487e4	0.214e3	0.1547e6	0.142e5	0.9838	0.9567	NaN	-0.144e-1
100	0.6	0.8	0.4905e4	0.207e3	0.158e6	0.14e5	0.9909	0.9771	NaN	0.704
100	0.6	1	0.4915e4	0.208e3	0.1591e6	0.141e5	0.9929	0.9839	NaN	NaN
100	0.8	0	0.4104e4	0.138e3	0.9244e5	0.696e4	0.8291	0.5717	0.108	0.327e-1
100	0.8	0.2	0.4585e4	0.178e3	0.129e6	0.109e5	0.9263	0.7978	NaN	-0.127
100	0.8	0.4	0.4816e4	0.201e3	0.1496e6	0.133e5	0.9729	0.9252	NaN	NaN
100	0.8	0.6	0.4889e4	0.198e3	0.1564e6	0.136e5	0.9877	0.9672	NaN	NaN
100	0.8	0.8	0.4911e4	0.204e3	0.1586e6	0.139e5	0.9921	0.9808	NaN	NaN
100	0.8	1	0.4916e4	0.208e3	0.1592e6	0.141e5	0.9931	0.9845	NaN	NaN
100	1	0	0.4338e4	0.144e3	0.1092e6	0.825e4	0.8764	0.6753	0.425e-2	-0.888e-1
100	1	0.2	0.47e4	0.189e3	0.139e6	0.12e5	0.9495	0.8596	-0.276e-1	NaN
100	1	0.4	0.485e4	0.201e3	0.1528e6	0.134e5	0.9798	0.945	-0.144e-1	-0.144e-1
100	1	0.6	0.4902e4	0.199e3	0.1577e6	0.135e5	0.9903	0.9753	NaN	NaN
100	1	0.8	0.4915e4	0.198e3	0.1589e6	0.136e5	0.9929	0.9827	0.704	0.492
100	1	1	0.4919e4	0.199e3	0.1593e6	0.137e5	0.9937	0.9852	NaN	NaN

Table A.2: Statistical summary of the BRG $G(n, u^\star)$ samples. The columns represent the number of vertices n, the parameters β_1 and β_2 of the related edge-triangle model (see equation (3.16)), the corresponding $u^\star$ value (see equation (3.21) and table A.3), and, for vertices $v_1, \ldots, v_4$ uniformly and randomly chosen (and fixed for each 100 samples), the correlation and covariance between the pairs of edges e_{12}, e_{13} and e_{12}, e_{34}.

n	β_1	β_2	$\overline{e}$	s_e	$\overline{t}$	s_t	$\overline{e}/\binom{n}{2}$	$\overline{t}/\binom{n}{3}$	Corr e_{12}, e_{13}	Corr e_{12}, e_{34}
20	0	0	0.9506e2	0.704e1	0.1431e3	0.325e2	0.5003	0.1255	0.9e-1	0.259e-1
20	0	0.2	0.1148e3	0.635e1	0.2525e3	0.42e2	0.6042	0.2215	-0.403e-1	-0.157
20	0	0.4	0.1612e3	0.519e1	0.6969e3	0.656e2	0.8484	0.6113	0.673e-1	0.123
20	0	0.6	0.1833e3	0.251e1	0.1023e4	0.423e2	0.9647	0.8974	-0.421e-1	-0.602e-1
20	0	0.8	0.1883e3	0.12e1	0.111e4	0.214e2	0.9911	0.9737	NaN	-0.101e-1
20	0	1	0.1895e3	0.658	0.113e4	0.119e2	0.9974	0.9912	NaN	-0.144e-1
20	0.2	0	0.1139e3	0.682e1	0.2456e3	0.446e2	0.5995	0.2154	-0.83e-1	0.785e-2
20	0.2	0.2	0.1415e3	0.581e1	0.4712e3	0.584e2	0.7447	0.4133	-0.882e-1	-0.229e-1
20	0.2	0.4	0.174e3	0.363e1	0.8759e3	0.552e2	0.9158	0.7683	-0.983e-1	-0.104
20	0.2	0.6	0.1859e3	0.208e1	0.1067e4	0.359e2	0.9784	0.936	-0.328e-1	-0.328e-1
20	0.2	0.8	0.1889e3	0.103e1	0.1121e4	0.182e2	0.9942	0.9833	NaN	NaN
20	0.2	1	0.1898e3	0.495	0.1136e4	0.887e1	0.9989	0.9965	NaN	NaN
20	0.4	0	0.1315e3	0.666e1	0.3789e3	0.579e2	0.6921	0.3324	-0.265e-1	-0.139
20	0.4	0.2	0.1593e3	0.48e1	0.6721e3	0.614e2	0.8384	0.5896	-0.169	-0.156
20	0.4	0.4	0.1812e3	0.31e1	0.989e3	0.51e2	0.9537	0.8675	0.295	-0.144e-1

Table A.2: (continuation)

n	β_1	β_2	$\overline{e}$	s_e	$\overline{t}$	s_t	$\overline{e}/\binom{n}{2}$	$\overline{t}/\binom{n}{3}$	Corr e_{12}, e_{13}	Corr e_{12}, e_{34}
20	0.4	0.6	0.1876e3	0.165e1	0.1097e4	0.291e2	0.9874	0.9623	-0.144e-1	NaN
20	0.4	0.8	0.1893e3	0.763	0.1127e4	0.137e2	0.9963	0.9886	NaN	NaN
20	0.4	1	0.1898e3	0.426	0.1136e4	0.768e1	0.9989	0.9965	NaN	NaN
20	0.6	0	0.1452e3	0.617e1	0.5106e3	0.653e2	0.7642	0.4479	0.263e-1	0.456e-1
20	0.6	0.2	0.1707e3	0.419e1	0.826e3	0.617e2	0.8984	0.7246	0.745e-1	0.576e-1
20	0.6	0.4	0.1844e3	0.233e1	0.1043e4	0.396e2	0.9705	0.9149	-0.251e-1	-0.177e-1
20	0.6	0.6	0.1885e3	0.107e1	0.1113e4	0.19e2	0.9921	0.9763	-0.101e-1	NaN
20	0.6	0.8	0.1895e3	0.672	0.1132e4	0.12e2	0.9974	0.993	NaN	NaN
20	0.6	1	0.1898e3	0.44	0.1136e4	0.792e1	0.9989	0.9965	NaN	NaN
20	0.8	0	0.1576e3	0.624e1	0.6508e3	0.788e2	0.8295	0.5709	0.297e-1	-0.356e-1
20	0.8	0.2	0.1772e3	0.385e1	0.9244e3	0.594e2	0.9326	0.8109	-0.361e-1	-0.58e-1
20	0.8	0.4	0.1863e3	0.197e1	0.1076e4	0.339e2	0.9805	0.9439	NaN	-0.309e-1
20	0.8	0.6	0.1886e3	0.121e1	0.1114e4	0.214e2	0.9926	0.9772	-0.204e-1	NaN
20	0.8	0.8	0.1897e3	0.51	0.1135e4	0.911e1	0.9984	0.9956	NaN	NaN
20	0.8	1	0.1897e3	0.485	0.1135e4	0.872e1	0.9984	0.9956	NaN	NaN
20	1	0	0.1676e3	0.382e1	0.7817e3	0.536e2	0.8821	0.6857	-0.205e-1	-0.101
20	1	0.2	0.1815e3	0.302e1	0.9931e3	0.492e2	0.9553	0.8711	-0.328e-1	-0.468e-1
20	1	0.4	0.1875e3	0.159e1	0.1097e4	0.277e2	0.9868	0.9623	-0.292e-1	-0.292e-1
20	1	0.6	0.1892e3	0.844	0.1126e4	0.151e2	0.9958	0.9877	NaN	NaN
20	1	0.8	0.1899e3	0.349	0.1137e4	0.628e1	0.9995	0.9974	NaN	NaN
20	1	1	0.1898e3	0.426	0.1136e4	0.768e1	0.9989	0.9965	NaN	NaN
25	0	0	0.1481e3	0.852e1	0.2761e3	0.465e2	0.4937	0.12	-0.147	0.618e-1
25	0	0.2	0.1825e3	0.772e1	0.5189e3	0.664e2	0.6083	0.2256	-0.103	-0.413e-1
25	0	0.4	0.2538e3	0.605e1	0.1393e4	0.995e2	0.846	0.6057	0.141e-1	-0.147
25	0	0.6	0.2898e3	0.334e1	0.2073e4	0.718e2	0.966	0.9013	-0.417e-1	-0.292e-1
25	0	0.8	0.2974e3	0.16e1	0.2241e4	0.361e2	0.9913	0.9743	-0.101e-1	-0.101e-1
25	0	1	0.2993e3	0.856	0.2284e4	0.197e2	0.9977	0.993	NaN	NaN
25	0.2	0	0.1801e3	0.821e1	0.4968e3	0.691e2	0.6003	0.216	-0.299e-1	-0.16
25	0.2	0.2	0.2228e3	0.791e1	0.9429e3	0.986e2	0.7427	0.41	-0.231	-0.737e-1
25	0.2	0.4	0.2756e3	0.46e1	0.1783e4	0.902e2	0.9187	0.7752	0.313	-0.482e-1
25	0.2	0.6	0.2934e3	0.248e1	0.2151e4	0.549e2	0.978	0.9352	-0.144e-1	-0.101e-1
25	0.2	0.8	0.2985e3	0.132e1	0.2265e4	0.3e2	0.995	0.9848	NaN	NaN
25	0.2	1	0.2995e3	0.703	0.2288e4	0.161e2	0.9983	0.9948	NaN	NaN
25	0.4	0	0.2065e3	0.827e1	0.7509e3	0.904e2	0.6883	0.3265	-0.112	0.334e-1
25	0.4	0.2	0.2507e3	0.602e1	0.1343e4	0.976e2	0.8357	0.5839	0.694e-1	-0.512e-1
25	0.4	0.4	0.2859e3	0.32e1	0.199e4	0.674e2	0.953	0.8652	-0.602e-1	0.219
25	0.4	0.6	0.2958e3	0.199e1	0.2205e4	0.442e2	0.986	0.9587	-0.144e-1	-0.204e-1
25	0.4	0.8	0.2988e3	0.11e1	0.2272e4	0.25e2	0.996	0.9878	NaN	NaN
25	0.4	1	0.2996e3	0.565	0.2291e4	0.13e2	0.9987	0.9961	NaN	NaN
25	0.6	0	0.231e3	0.652e1	0.105e4	0.899e2	0.77	0.4565	0.541e-1	0.117e-1
25	0.6	0.2	0.2697e3	0.534e1	0.1671e4	0.995e2	0.899	0.7265	-0.863e-1	0.188
25	0.6	0.4	0.2909e3	0.34e1	0.2097e4	0.727e2	0.9697	0.9117	NaN	NaN
25	0.6	0.6	0.2977e3	0.155e1	0.2247e4	0.351e2	0.9923	0.977	NaN	NaN
25	0.6	0.8	0.2993e3	0.902	0.2284e4	0.207e2	0.9977	0.993	NaN	NaN
25	0.6	1	0.2997e3	0.482	0.2293e4	0.111e2	0.999	0.997	NaN	NaN
25	0.8	0	0.2506e3	0.573e1	0.134e4	0.935e2	0.8353	0.5826	0.78e-2	-0.266e-1
25	0.8	0.2	0.2795e3	0.353e1	0.186e4	0.697e2	0.9317	0.8087	-0.753e-1	0.193e-1
25	0.8	0.4	0.2942e3	0.232e1	0.217e4	0.511e2	0.9807	0.9435	-0.144e-1	-0.177e-1
25	0.8	0.6	0.2987e3	0.103e1	0.227e4	0.233e2	0.9957	0.987	NaN	NaN
25	0.8	0.8	0.2994e3	0.761	0.2286e4	0.174e2	0.998	0.9939	NaN	NaN
25	0.8	1	0.2997e3	0.541	0.2293e4	0.124e2	0.999	0.997	NaN	NaN
25	1	0	0.2649e3	0.622e1	0.1584e4	0.112e3	0.883	0.6887	0.119	0.458e-1
25	1	0.2	0.2869e3	0.394e1	0.2012e4	0.826e2	0.9563	0.8748	0.163	0.163
25	1	0.4	0.296e3	0.179e1	0.221e4	0.4e2	0.9867	0.9609	-0.177e-1	-0.251e-1
25	1	0.6	0.2989e3	0.986	0.2275e4	0.225e2	0.9963	0.9891	NaN	NaN
25	1	0.8	0.2996e3	0.618	0.2291e4	0.142e2	0.9987	0.9961	NaN	NaN
25	1	1	0.2996e3	0.586	0.2291e4	0.135e2	0.9987	0.9961	NaN	NaN
30	0	0	0.2178e3	0.101e2	0.5118e3	0.734e2	0.5007	0.1261	-0.821e-1	0.14
30	0	0.2	0.2644e3	0.11e2	0.9119e3	0.116e3	0.6078	0.2246	0.115	0.114
30	0	0.4	0.3704e3	0.709e1	0.2507e4	0.146e3	0.8515	0.6175	-0.169	0.425e-2
30	0	0.6	0.4205e3	0.361e1	0.3667e4	0.949e2	0.9667	0.9032	0.117	-0.56e-1
30	0	0.8	0.431e3	0.201e1	0.395e4	0.555e2	0.9908	0.9729	-0.144e-1	-0.144e-1
30	0	1	0.4338e3	0.114e1	0.4027e4	0.316e2	0.9972	0.9919	NaN	NaN
30	0.2	0	0.2602e3	0.12e2	0.8713e3	0.125e3	0.5982	0.2146	-0.264e-1	0.729e-1
30	0.2	0.2	0.322e3	0.989e1	0.1647e4	0.152e3	0.7402	0.4057	0.252e-1	-0.279e-1
30	0.2	0.4	0.4e3	0.58e1	0.3156e4	0.137e3	0.9195	0.7773	0.	-0.107e-1
30	0.2	0.6	0.4256e3	0.292e1	0.3802e4	0.783e2	0.9784	0.9365	-0.309e-1	-0.177e-1
30	0.2	0.8	0.4324e3	0.154e1	0.3988e4	0.426e2	0.994	0.9823	NaN	NaN
30	0.2	1	0.4342e3	0.907	0.4037e4	0.254e2	0.9982	0.9943	NaN	NaN

Table A.2: (continuation)

n	β_1	β_2	$\overline{e}$	s_e	$\overline{t}$	s_t	$\overline{e}/\binom{n}{2}$	$\overline{t}/\binom{n}{3}$	Corr e_{12}, e_{13}	Corr e_{12}, e_{34}
30	0.4	0	0.3023e3	0.103e2	0.1363e4	0.144e3	0.6949	0.3357	0.369e-1	-0.4e-1
30	0.4	0.2	0.3653e3	0.716e1	0.2405e4	0.142e3	0.8398	0.5924	0.639e-1	0.106
30	0.4	0.4	0.4131e3	0.424e1	0.3478e4	0.108e3	0.9497	0.8567	-0.745e-1	-0.602e-1
30	0.4	0.6	0.4293e3	0.228e1	0.3901e4	0.622e2	0.9869	0.9608	NaN	NaN
30	0.4	0.8	0.4333e3	0.132e1	0.4014e4	0.368e2	0.9961	0.9887	-0.101e-1	NaN
30	0.4	1	0.4346e3	0.589	0.4048e4	0.165e2	0.9991	0.997	NaN	NaN
30	0.6	0	0.3343e3	0.828e1	0.1841e4	0.14e3	0.7685	0.4534	-0.367e-1	-0.15
30	0.6	0.2	0.3905e3	0.715e1	0.2939e4	0.162e3	0.8977	0.7239	-0.765e-1	-0.677e-1
30	0.6	0.4	0.4215e3	0.387e1	0.3693e4	0.101e3	0.969	0.9096	-0.468e-1	-0.468e-1
30	0.6	0.6	0.4314e3	0.197e1	0.3959e4	0.542e2	0.9917	0.9751	NaN	NaN
30	0.6	0.8	0.4339e3	0.971	0.403e4	0.27e2	0.9975	0.9926	NaN	NaN
30	0.6	1	0.4346e3	0.624	0.4048e4	0.174e2	0.9991	0.997	NaN	NaN
30	0.8	0	0.3614e3	0.848e1	0.2327e4	0.166e3	0.8308	0.5732	0.385e-1	-0.115
30	0.8	0.2	0.4051e3	0.497e1	0.3279e4	0.121e3	0.9313	0.8076	-0.516e-1	-0.56e-1
30	0.8	0.4	0.4268e3	0.275e1	0.3833e4	0.745e2	0.9811	0.9441	NaN	NaN
30	0.8	0.6	0.4327e3	0.14e1	0.3997e4	0.388e2	0.9947	0.9845	NaN	NaN
30	0.8	0.8	0.4342e3	0.899	0.4038e4	0.25e2	0.9982	0.9946	NaN	NaN
30	0.8	1	0.4345e3	0.797	0.4045e4	0.22e2	0.9989	0.9963	NaN	NaN
30	1	0	0.3828e3	0.824e1	0.2767e4	0.18e3	0.88	0.6815	-0.136	-0.512e-1
30	1	0.2	0.4167e3	0.427e1	0.3568e4	0.11e3	0.9579	0.8788	0.164	-0.745e-1
30	1	0.4	0.4289e3	0.258e1	0.3892e4	0.7e2	0.986	0.9586	NaN	-0.177e-1
30	1	0.6	0.4332e3	0.144e1	0.4009e4	0.4e2	0.9959	0.9874	NaN	NaN
30	1	0.8	0.4345e3	0.674	0.4046e4	0.188e2	0.9989	0.9966	NaN	NaN
30	1	1	0.4346e3	0.584	0.4049e4	0.163e2	0.9991	0.9973	NaN	NaN
100	0	0	0.2474e4	0.415e2	0.2019e5	0.104e4	0.4998	0.1249	0.771e-1	-0.355e-1
100	0	0.2	0.302e4	0.3e2	0.3672e5	0.113e4	0.6101	0.2271	0.137	-0.456e-1
100	0	0.4	0.4206e4	0.237e2	0.9918e5	0.168e4	0.8497	0.6134	-0.791e-1	0.625e-1
100	0	0.6	0.4786e4	0.116e2	0.1462e6	0.106e4	0.9669	0.9041	-0.177e-1	-0.309e-1
100	0	0.8	0.4906e4	0.638e1	0.1574e6	0.615e3	0.9911	0.9734	NaN	NaN
100	0	1	0.4937e4	0.351e1	0.1604e6	0.342e3	0.9974	0.992	NaN	NaN
100	0.2	0	0.2965e4	0.32e2	0.3476e5	0.113e4	0.599	0.215	0.434e-1	0.138
100	0.2	0.2	0.3678e4	0.245e2	0.6633e5	0.134e4	0.743	0.4102	-0.915e-1	0.102
100	0.2	0.4	0.4551e4	0.197e2	0.1257e6	0.163e4	0.9194	0.7774	-0.983e-1	0.361e-1
100	0.2	0.6	0.4846e4	0.109e2	0.1517e6	0.103e4	0.979	0.9382	-0.309e-1	-0.251e-1
100	0.2	0.8	0.4921e4	0.507e1	0.1589e6	0.491e3	0.9941	0.9827	NaN	NaN
100	0.2	1	0.4941e4	0.303e1	0.1609e6	0.295e3	0.9982	0.9951	NaN	NaN
100	0.4	0	0.3412e4	0.342e2	0.5297e5	0.161e4	0.6893	0.3276	-0.23	-0.591e-1
100	0.4	0.2	0.4146e4	0.258e2	0.95e5	0.178e4	0.8376	0.5875	-0.143	-0.163e-1
100	0.4	0.4	0.4707e4	0.152e2	0.139e6	0.135e4	0.9509	0.8596	-0.204e-1	-0.421e-1
100	0.4	0.6	0.4885e4	0.835e1	0.1554e6	0.797e3	0.9869	0.961	NaN	NaN
100	0.4	0.8	0.4931e4	0.467e1	0.1598e6	0.454e3	0.9962	0.9882	NaN	NaN
100	0.4	1	0.4944e4	0.252e1	0.1611e6	0.247e3	0.9988	0.9963	NaN	NaN
100	0.6	0	0.3807e4	0.271e2	0.7354e5	0.158e4	0.7691	0.4548	-0.367e-1	0.195
100	0.6	0.2	0.4442e4	0.203e2	0.1169e6	0.161e4	0.8974	0.7229	-0.343e-1	-0.721e-1
100	0.6	0.4	0.4797e4	0.124e2	0.1472e6	0.114e4	0.9691	0.9103	-0.144e-1	-0.144e-1
100	0.6	0.6	0.4907e4	0.615e1	0.1575e6	0.592e3	0.9913	0.974	-0.144e-1	-0.144e-1
100	0.6	0.8	0.4938e4	0.351e1	0.1605e6	0.342e3	0.9976	0.9926	NaN	NaN
100	0.6	1	0.4945e4	0.194e1	0.1612e6	0.19e3	0.999	0.9969	NaN	NaN
100	0.8	0	0.4119e4	0.239e2	0.9317e5	0.163e4	0.8321	0.5762	0.316e-1	0.15
100	0.8	0.2	0.4624e4	0.174e2	0.1318e6	0.149e4	0.9341	0.8151	0.144	-0.56e-1
100	0.8	0.4	0.4852e4	0.102e2	0.1523e6	0.961e3	0.9802	0.9419	NaN	NaN
100	0.8	0.6	0.4921e4	0.498e1	0.1589e6	0.482e3	0.9941	0.9827	NaN	NaN
100	0.8	0.8	0.4941e4	0.316e1	0.1609e6	0.308e3	0.9982	0.9951	NaN	NaN
100	0.8	1	0.4945e4	0.225e1	0.1612e6	0.22e3	0.999	0.9969	NaN	NaN
100	1	0	0.4359e4	0.239e2	0.1104e6	0.181e4	0.8806	0.6827	-0.964e-1	0.476
100	1	0.2	0.4736e4	0.148e2	0.1416e6	0.132e4	0.9568	0.8757	-0.204e-1	-0.204e-1
100	1	0.4	0.4887e4	0.818e1	0.1556e6	0.781e3	0.9873	0.9623	NaN	NaN
100	1	0.6	0.4932e4	0.398e1	0.1599e6	0.387e3	0.9964	0.9889	NaN	NaN
100	1	0.8	0.4945e4	0.251e1	0.1612e6	0.246e3	0.999	0.9969	NaN	NaN
100	1	1	0.4945e4	0.226e1	0.1612e6	0.221e3	0.999	0.9969	NaN	NaN

Table A.3: Some numerical values for the relation between $u^\star$ and the parameters β_1 and β_2 in equation (3.21). Recall that $u^\star$ is the expected edge density of $G(n, u^\star)$, and $(u^\star)^3$ is its expected density of triangles. See figure 6.1 for a plot of $u^\star = u(\beta_1, \beta_2)$.

β_1	β_2	$u^\star$	$(u^\star)^3$
0.0	0.0	0.5	0.125
0.0	0.2	0.610	0.2267
0.0	0.4	0.850	0.6138
0.0	0.6	0.967	0.903
0.0	0.8	0.991	0.9736
0.0	1.0	0.997	0.9923
0.2	0.0	0.599	0.2146
0.2	0.2	0.743	0.4106
0.2	0.4	0.919	0.7756
0.2	0.6	0.979	0.9389
0.2	0.8	0.994	0.9827
0.2	1.0	0.998	0.9949
0.4	0.0	0.690	0.3285
0.4	0.2	0.838	0.5882
0.4	0.4	0.951	0.8609
0.4	0.6	0.987	0.9606
0.4	0.8	0.996	0.9886
0.4	1.0	0.999	0.9966
0.6	0.0	0.768	0.4539
0.6	0.2	0.897	0.722
0.6	0.4	0.969	0.911
0.6	0.6	0.991	0.9741
0.6	0.8	0.997	0.9924
0.6	1.0	0.999	0.9968
0.8	0.0	0.832	0.576
0.8	0.2	0.934	0.8142
0.8	0.4	0.980	0.942
0.8	0.6	0.994	0.9829
0.8	0.8	0.998	0.9949
0.8	1.0	0.999	0.9968
1.0	0.0	0.881	0.6833
1.0	0.2	0.957	0.8761
1.0	0.4	0.987	0.9618
1.0	0.6	0.996	0.9887
1.0	0.8	0.999	0.9967
1.0	1.0	0.999	0.9968

Appendix B

Proofs

Claim 2.3. Let $\mathcal{G}_n$ be the set of all n-vertex graphs. We have

$$\frac{1}{|\mathcal{G}_n|} \sum_{G \in \mathcal{G}_n} \mathrm{clus}(G) = \frac{1}{2}.$$

Proof. Without loss of generality, we consider the graphs $\mathcal{G}_n$ have the same set of vertices $[n] \doteq \{1, 2, \ldots, n\}$. Note that $|\mathcal{G}_n| = 2^{\binom{n}{2}}$.

For each $v \in [n]$, each subset $A \subseteq V \setminus \{v\}$ and each $B \subseteq \binom{A}{2}$, we define the family $\mathcal{F}(v, A, B)$ of graphs $G(V, E)$ such that the neighbors of v in G are exactly the vertices in A, and the edges between vertices of A are exactly those in B.

Note that for every graph $G \in \mathcal{F}(v, A, B)$, we have

$$\mathrm{clus}(v) = |B| \binom{|A|}{2}^{-1}.$$

Therefore, we can make explicit the contribution of each vertex in the average

$$\begin{aligned}
\sum_{G \in \mathcal{G}_n} \mathrm{clus}(G) &= \sum_{G \in \mathcal{G}_n} \sum_{v \in [n]} e_v \binom{\mathrm{d}(v)}{2}^{-1} \\
&= \sum_{v \in [n]} \sum_{G \in \mathcal{G}_n} e_v \binom{\mathrm{d}(v)}{2}^{-1} \\
&= \sum_{v \in [n]} \sum_{A \subseteq [n] \setminus \{v\}} \sum_{\substack{G \in \mathcal{G}_n \\ N(v) = A}} e_v \binom{|A|}{2}^{-1} \\
&= \sum_{v \in [n]} \sum_{A \subseteq [n] \setminus \{v\}} \sum_{B \subseteq \binom{A}{2}} \sum_{\substack{G = (V,E) \in \mathcal{G}_n \\ N(v) = A \\ E \cap \binom{A}{2} = B}} |B| \binom{|A|}{2}^{-1} \\
&= \sum_{v \in [n]} \sum_{A \subseteq [n] \setminus \{v\}} \sum_{B \subseteq \binom{A}{2}} |B| \binom{|A|}{2}^{-1} |\mathcal{F}(v, A, B)|
\end{aligned}$$

Furthermore, we observe that $\left|\mathcal{F}(v, A, B)\right| = \left|\mathcal{F}(v, A, C)\right|$, for all $C \subseteq \binom{A}{2}$. In particular, we can take $C = \binom{A}{2} \setminus B$. If we fix v and A as in the sum above,

$$
\begin{aligned}
&\sum_{\substack{v,A\\ B\subseteq\binom{A}{2}}} \frac{|B|}{\binom{|A|}{2}}|\mathcal{F}(v,A,B)| \\
&\quad= \frac{1}{2}\cdot 2 \sum_{\substack{v,A\\ B\subseteq\binom{A}{2}}} \frac{|B|}{\binom{|A|}{2}}|\mathcal{F}(v,A,B)| \\
&\quad= \frac{1}{2} \sum_{\substack{v,A\\ B\subseteq\binom{A}{2}}} \left(\frac{|B|}{\binom{|A|}{2}}|\mathcal{F}(v,A,B)| + \frac{\binom{|A|}{2}-|B|}{\binom{|A|}{2}}|\mathcal{F}(v,A,\binom{A}{2}\setminus B)|\right) \\
&\quad= \frac{1}{2} \sum_{\substack{v,A\\ B\subseteq\binom{A}{2}}} \left(\frac{|B|}{\binom{|A|}{2}}|\mathcal{F}(v,A,B)| + \frac{\binom{|A|}{2}-|B|}{\binom{|A|}{2}}|\mathcal{F}(v,A,B)|\right) \\
&\quad= \frac{1}{2} \sum_{\substack{v,A\\ B\subseteq\binom{A}{2}}} |\mathcal{F}(v,A,B)|
\end{aligned}
$$

To proceed, we use two auxiliar results.

Claim B.1 Let $v \in [n]$, and let $A, A' \subseteq [n] \setminus \{v\}$, and $B, B' \subset \binom{A}{2}$. If $C = \mathcal{F}(v,A,B) \cap \mathcal{F}(v,A',B')$, then either $C = \mathcal{F}(v,A,B)$ or $C = \emptyset$.

(Proof of Claim B.1.) Observe that for every graph $G \in \mathcal{G}_n$, and $v \in [n]$, there is a unique choice of $C \subseteq [n] \setminus \{v\}$ and of $D \subseteq \binom{C}{2}$ such that $G \in \mathcal{F}(v,C,D)$. Hence, if G is a graph in both sets, then $A = A'$ and $B = B'$, that is, the sets are identical.$\diamond$

Claim B.2 Let $v \in [n]$. We have

$$\bigcup_{\substack{A\subseteq[n]\setminus\{v\}\\ B\subseteq\binom{A}{2}}} \mathcal{F}(v,A,B) = \mathcal{G}_n \quad \text{e ainda} \quad \sum_{\substack{A\subseteq[n]\setminus\{v\}\\ B\subseteq\binom{A}{2}}} |\mathcal{F}(v,A,B)| = |\mathcal{G}_n|.$$

(Proof of Claim B.2). Every graph in the union is an element of $\mathcal{G}_n$, as all of them have n vertices. Also, every graph $G = (V,E) \in \mathcal{G}_n$ is in $\mathcal{F}(v,A,B)$ of the union, taking $A = N(v)$ and $B = E \cap \binom{A}{2}$. Therefore, both sets are the same. The equality of the sums is a corolary of the Claim B.1, since $\mathcal{F}(v,A,B)$ define a partition of $\mathcal{G}_n$.$\diamond$

Fixing v, we have

$$\sum_{\substack{A\subseteq[n]\setminus\{v\}\\ B\subseteq\binom{A}{2}}} |\mathcal{F}(v,A,B)| = \left|\bigcup_{\substack{A\subseteq[n]\setminus\{v\}\\ B\subseteq\binom{A}{2}}} \mathcal{F}(v,A,B)\right| = |\mathcal{G}_n| = 2^{\binom{n}{2}}.$$

And we conclude that

$$\begin{aligned}
|\mathcal{G}_n|^{-1} \sum_{G\in\mathcal{G}_n} \mathrm{clus}(G) = 2^{-\binom{n}{2}} \sum_{G\in\mathcal{G}_n} \mathrm{clus}(G) = \\
&= \frac{1}{n2^{\binom{n}{2}}} \sum_{v\in[n]} \sum_{A\subseteq[n]\setminus\{v\}} \sum_{B\subseteq\binom{A}{2}} |B| \binom{|A|}{2}^{-1} |\mathcal{F}(v,A,B)| \\
&= \frac{1}{2}\cdot\frac{1}{n2^{\binom{n}{2}}} \sum_{v\in[n]} \sum_{A\subseteq[n]\setminus\{v\}} \sum_{B\subseteq\binom{A}{2}} |\mathcal{F}(v,A,B)| \\
&= \frac{1}{2}\cdot\frac{1}{n2^{\binom{n}{2}}} \sum_{v\in[n]} 2^{\binom{n}{2}} \\
&= \frac{1}{2}\cdot\frac{1}{n2^{\binom{n}{2}}} \cdot n2^{\binom{n}{2}} \\
&= \frac{1}{2}
\end{aligned}$$

□

Lemma 3.5. Let $G = (V, E)$ be a graph with at least one edge and with no isolated vertices. If every pair of edges $e, f \in E$ of G has a common vertex (i.e.: $e \cap f \neq \emptyset$), then G is a triangle or a star.

Proof. Since G have no isolated vertex, if $|E| \leq 2$ then by inspection G is a star. We suppose in the rest of the proof that G has at least three distinct edges.

Note that if G has a triangle with vertices $\{a, b, c\} \in V$, then G *is a triangle*. This because every edge $\{d, e\} \in E$ is adjacent (i.e., has a common extreme) to each of the other edges of the triangle, and this is only possible if $|\{d, e\} \cap \{a, b, c\}| > 1$; hence $\{d, e\}$ is one of the edges $\{a, b\}, \{a, c\}$, or $\{b, c\}$.

On the other hand, if there is no triangle in G, then the intersection between any three distinct edges cotains exactly one vertex: this since an edge $\{a, b\}$ cannot intercept two others $\{c, d\}, \{d, e\}$ at different vertices without generating a triangle, and thus $\{a, b\} \cap \{c, d\} = \{a, b\} \cap \{d, e\} = d$. Now, this implies that the vertex $v \in V$ at the intersection between distinct edges is unique, for if (e_1, e_2, e_3) and (e_2, e_3, e_4) are triples of distinct edges ($e_i \in E, e \in \{1, 2, 3, 4\}$), then

$$v = e_1 \cap e_2 \cap e_3 = e_2 \cap e_3 = e_2 \cap e_3 \cap e_4.$$

And, as v is in every edge, G is a star. □

Claim 3.2. Let $G \sim G(n, p)$. We have $\mathbb{E}(\mathrm{clus}(G)) = p$.

Proof. By linearity of expectation, we have

$$\mathbb{E}(\mathrm{clus}(G)) = n^{-1} \sum_{v\in[n]} \mathbb{E}(\mathrm{clus}(v)).$$

Let $E = E(G)$ be the set of edges of G, and $E_v = E \cap \binom{N(v)}{2}$ be the set of edges between neighbors of $v \in V$, with $e_v \doteq |E_v|$. Writing $a \doteq |A|$, $b \doteq |B|$, and $q = 1 - p$, we have, for all $v \in [n]$,

$$\begin{aligned}
\mathbb{E}(\mathrm{clus}(v)) &= \sum_{A \subseteq [n]-v} \sum_{B \subseteq \binom{A}{2}} \frac{b}{\binom{a}{2}} \cdot \mathrm{P}\left(N(v) = A \wedge E_v = B\right) \\
&= \sum_{A \subseteq [n]-v} \sum_{b=0}^{\binom{a}{2}} \binom{\binom{a}{2}}{b} \frac{b}{\binom{a}{2}} \cdot \mathrm{P}\left(N(v) = A \wedge e_v = b\right) \\
&= \sum_{A \subseteq [n]-v} \sum_{b=0}^{\binom{a}{2}} \binom{\binom{a}{2}}{b} \frac{b}{\binom{a}{2}} \cdot p^a q^{n-1-a} p^b q^{\binom{a}{2}-b}
\end{aligned}$$

Fixing $v \in [n]$, the number of graphs $G = (V, E)$ such that $|N(v)| = a$ and $\left|E \cap \binom{N(v)}{2}\right| = b$ is $\binom{n-1}{a}\binom{\binom{a}{2}}{b}$. Therefore,

$$\begin{aligned}
\mathbb{E}(\mathrm{clus}(v)) &= \sum_{A \subseteq [n]-v} \sum_{B \subseteq \binom{A}{2}} \frac{b}{\binom{a}{2}} \cdot \mathrm{P}\left(N(v) = A \wedge E_v = B\right) \\
&= \sum_{a=0}^{n-1} \sum_{b=0}^{\binom{a}{2}} \binom{n-1}{a} \binom{\binom{a}{2}}{b} \frac{b}{\binom{a}{2}} \cdot \mathrm{P}\left(N(v) = A \wedge E_v = B\right) \\
&= \sum_{a=0}^{n-1} \sum_{b=0}^{\binom{a}{2}} \binom{n-1}{a} \binom{\binom{a}{2}}{b} \frac{b}{\binom{a}{2}} p^a q^{n-1-a} p^b q^{\binom{a}{2}-b} \\
&= \sum_{a=0}^{n-1} \binom{n-1}{a} \binom{a}{2}^{-1} p^a q^{n-1-a} \binom{a}{2} p(p+q)^{\binom{a}{2}-1} \\
&= p(p+q)^{n-1} \\
&= p
\end{aligned}$$

□

Where we use the fact that $\sum_{k=0}^{n} \binom{n}{k} k x^k y^{n-k} = nx(x+y)^{n-1}$.

Proof. We simply calculate the derivative (in x) of both sides of Newton's binomial equation.

$$\begin{aligned}
(x+y)^n &= \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} \iff \\
n(x+y)^{n-1} &= \sum_{k=0}^{n} \binom{n}{k} k x^{k-1} y^{n-k} \iff \\
nx(x+y)^{n-1} &= \sum_{k=0}^{n} \binom{n}{k} k x^k y^{n-k}
\end{aligned}$$

□