

Exponential Approximation, Method of Types for Empirical Neighbourhood Distributions of Random Graphs by Random Allocations

K. Doku-Amponsah¹

¹ Statistics Department, University of Ghana, Legon, Ghana

Correspondence: K. Doku-Amponsah, Statistics Department, University of Ghana, Box LG115, Legon, Ghana.
Tel: 233-20-516-4254. E-mail: kdoku@ug.edu.gh

Received: January 8, 2014 Accepted: April 8, 2014 Online Published: April 21, 2014

doi:10.5539/ijsp.v3n2p110 URL: <http://dx.doi.org/10.5539/ijsp.v3n2p110>

The research is financed by University of Ghana

Abstract

In this article we find exponential good approximation of the empirical neighbourhood distribution of symbolled random graphs conditioned to a given empirical symbol distribution and empirical pair distribution. Using this approximation we shorten or simplify the proof of (Doku-Amponsah & Morters, 2010, Theorem 2.5); the large deviation principle (LDP) for empirical neighbourhood distribution of symbolled random graphs. We also show that the LDP for the empirical degree measure of the classical Erdős-Rényi graph is a special case of (Doku-Amponsah & Moerters, 2010, Theorem 2.5). From the LDP for the empirical degree measure, we derive an LDP for the the proportion of isolated vertices in the classical Erdős-Rényi graph.

Keywords: concentration inequalities, coupling, empirical occupancy measure, empirical degree measure, sparse random graphs, bins and balls

Mathematics Subject Classification: 60F10, 05C80

1. Introduction

The Erdős-Rényi graph $\mathcal{G}(n, p)$ or $\mathcal{G}(n, nc/2)$ is the simplest imaginable random graph, which arises by taking n vertices, and placing an edge between any two of distinct nodes or vertices with a fixed probability $0 < p < 1$ or inserting a fixed number nc edges at random among the n vertices (See Van Der Hofstad, 2009). Several large deviation (LD) results for this graphs have been found (See, for example, O'Connell, 1998; Biggins & Penman, 2009; Doku-Amponsah & Moerters, 2010; Doku-Amponsah, 2006; Bordenave & Caputo, 2013; Mukherjee, 2013).

O'Connell (1998) proved an LDP for the relative size of the largest connected component and the number of isolated vertices in the random graph $\mathcal{G}(n, p)$ with $p = O(1/n)$. O'Connell (1998) also presented an LDP and a related result for the number of isolated vertices in the random graph $\mathcal{G}(n, c/n)$. i.e. the near-critical or sparse case. An LDP for the proportion of edges to the number of potential vertices of the supercritical case has been found by Biggins and Penman (2009). Doku-Amponsah and Moerters (2010) and Doku-Amponsah (2006, 2012) obtained several LDPs, including the LDP for empirical degree distribution for near-critical or sparse case. Boucheron et al. (2002) attempted to prove from the LDP for the empirical occupancy process an LDP for the degree distribution of the random graph $\mathcal{G}(n, nc)$. But Doku-Amponsah et al. (2010) conjectured that the prove of this LDP does not hold. Recently, Bordenave and Caputo (2013) obtained LDPs for the empirical neighbourhood distribution in $\mathcal{G}(n, c/n)$ and $\mathcal{G}(n, nc/2)$. The LDP for the empirical degree distribution in $\mathcal{G}(n, c/n)$ has been proved in (Mukherjee, 2013).

Our main aim in this article is to obtain an exponential approximation result, see Lemma 4, for the empirical Neighbourhood distribution of symbolled random graphs. Using this result we shorten or simplify the proof of the LDP for empirical Neighbourhood distribution of symbolled random graphs conditioned on a given empirical symbol measure and empirical pair distribution (See example, Doku-Amponsah et al., 2010 or Doku-Amponsah, 2012, Theorem 2.5.1).

Further, we show that the large deviation principle for the empirical degree measure of $\mathcal{G}(n, nc/2)$ is a special case of (Doku-Amponsah, 2012, Theorem 2.5.1). From this result we find an LDP for the proportion of isolated vertices in the graph $\mathcal{G}(n, nc/2)$. Note that the LDP for the proportion of isolated vertices in the graph $\mathcal{G}(n, nc/2)$ is new in the literature (See, O’Connell, 1998 for similar result).

The main technique used in this article is the method of types (see Dembo & Zeitouni, 1998, Section 2.1). The method of types is applied to an exponential approximate model for the symbolled random graph model which we shall obtain by randomly allocating symbolled balls in to symbolled bins.

The symbol random graphs (see Penman, 1998) or Inhomogeneous random graphs (see Van Der Hofstad, 2009), which has Erdős-Rényi graph with one symbol as an example permit a dependence between symbol and connectivity of the nodes. In the next subsection, we review the *symbolled random graph model* as in (Doku-Amponsah et al., 2010).

1.1 The Symbolled Random Graph Model

We begin by fixing the following notations. Let \mathcal{Y} be a symbol or colour set. Let $\mathcal{V} = \{1, \dots, n\}$ be a fixed set of n vertices. Denote by E the edge set, i.e.

$$E \subset \mathcal{E} := \{(e_1, e_2) \in \mathcal{V} \times \mathcal{V} : e_1 < e_2\},$$

where we have used the *formal* ordering of links to simply describe *unoriented* edges.

Let $p_n : \mathcal{Y} \times \mathcal{Y} \rightarrow [0, 1]$ be a symmetric function and $\mu : \mathcal{Y} \rightarrow [0, 1]$ a probability law. We may describe the simply *symbolled random graph* Y with n vertices in the following manner: Any node $v \in \mathcal{V}$ gets symbol $Y(v)$ independently and identically according to the *symbol law* λ . Given the colours, we join any two nodes $u, v \in \mathcal{V}$ with an edge, independently of everything else, with a *edge probability* $p_n(Y(u), Y(v))$ otherwise we keep them disconnected. We always look at $Y = ((Y(v) : v \in \mathcal{V}), E)$ under the combine distribution of graph and symbol. We interpret Y as symbolled random graph and consider $Y(v)$ as the symbol of the node v (See Cannings & Penman, 2003; Penman, 1998).

Our interest in this article is on the symbolled random graph models in the *near-critical* regimes. Thus, we look at cases when the edge probability $p_n(a, b)$ satisfies $np_n(a, b) \rightarrow C(a, b)$, for all $a, b \in \mathcal{Y}$, and $C : \mathcal{Y} \times \mathcal{Y} \rightarrow [0, \infty)$.

By $\mathcal{W}(X)$ we denote the space of probability measures on a finite or countable set X , and by $\tilde{\mathcal{W}}(X)$ we denote the subspace of finite measures defined on X , and we endow both with the weak topology. Further, we denote by $\tilde{\mathcal{W}}_*(X \times X)$ the subspace of symmetric measures. By convection, we write

$$\mathbb{N} = \{0, 1, 2, \dots\}.$$

For any symbolled graph $Y = ((Y(v) : v \in \mathcal{V}), E)$ with n nodes we recall from (Doku-Amponsah & Moerters, 2010), the definition of the *empirical symbol distribution* $L_Y^1 \in \mathcal{W}(\mathcal{Y})$, by

$$L_Y^1(a) := \frac{1}{n} \sum_{u \in \mathcal{V}} \delta_{Y(u)}(a), \quad \text{for } a \in \mathcal{Y},$$

and the *empirical pair distribution* $L_Y^2 \in \tilde{\mathcal{W}}_*(\mathcal{Y} \times \mathcal{Y})$, by

$$L_Y^2(b, a) := \frac{1}{n} \sum_{(e_1, e_2) \in E} [\delta_{(Y(e_2), Y(e_1))} + \delta_{(Y(e_1), Y(e_2))}](b, a), \quad \text{for } a, b \in \mathcal{Y}.$$

We observe that by definition L_Y^2 is finite symmetric measure with total mass $\|L_Y^2\|$ equal to $2|E|/n$. Finally we recall the definition of the *empirical Neighbourhood distribution* $M_Y \in \mathcal{W}(\mathcal{Y} \times \mathbb{N}^{\mathcal{Y}})$, by

$$M_Y(a, \ell) := \frac{1}{n} \sum_{u \in \mathcal{V}} \delta_{(Y(u), \mathcal{L}(u))}(a, \ell), \quad \text{for } (a, \ell) \in \mathcal{Y} \times \mathbb{N}^{\mathcal{Y}},$$

where $\mathcal{L}(u) = (\ell^u(b), b \in \mathcal{Y})$ and $\ell^u(b)$ is the number of nodes of symbol b linked to node v . For any $\mu \in \mathcal{W}(\mathcal{Y} \times \mathbb{N}^{\mathcal{Y}})$ we denote by μ_1 the \mathcal{Y} - marginal of μ and for every $(b, a) \in \mathcal{Y} \times \mathcal{Y}$, let μ_2 be the law of the pair $(a, l(b))$ under the measure μ . Define the measure (finite), $\langle \mu(\cdot, \ell), l(\cdot) \rangle \in \tilde{\mathcal{W}}(\mathcal{Y} \times \mathcal{Y})$ by

$$\Delta_2(\mu)(b, a) := \sum_{l(b) \in \mathbb{N}} \mu_2(a, l(b))l(b), \quad \text{for } a, b \in \mathcal{Y}$$

and write $\Delta_1(\mu) = \mu_1$. We define the function $\Delta: \mathcal{W}(\mathcal{Y} \times \mathbb{N}^{\mathcal{Y}}) \rightarrow \mathcal{W}(\mathcal{Y}) \times \tilde{\mathcal{W}}(\mathcal{Y} \times \mathcal{Y})$ by $\Delta(\mu) = (\Delta_1(\mu), \Delta_2(\mu))$ and note that $\Delta(M_X) = (L_Y^1, L_Y^2)$. Observe that Δ_1 is a continuous function but Δ_2 is *discontinuous* in the weak topology. In particular, in the summation $\sum_{l(b) \in \mathbb{N}} \mu_2(a, l(b))l(b)$ the function $l(b)$ may be unbounded and so the functional $\mu \rightarrow \Delta_2(\mu)$ would not be continuous in the weak topology. We call a pair of measures $(\pi, \mu) \in \tilde{\mathcal{W}}(\mathcal{Y} \times \mathcal{Y}) \times \mathcal{W}(\mathcal{Y} \times \mathbb{N}^{\mathcal{Y}})$ *sub-consistent* if

$$\Delta_2(\mu)(b, a) \leq \pi(b, a), \quad \text{for all } a, b \in \mathcal{Y}, \tag{1}$$

and *consistent* if equality holds in (1). For any $n \in \mathbb{N}$ we define the following sets:

$$\begin{aligned} \mathcal{W}_n(\mathcal{Y}) &:= \{ \nu \in \mathcal{W}(\mathcal{Y}) : n\nu(b) \in \mathbb{N} \text{ for all } b \in \mathcal{Y} \}, \\ \tilde{\mathcal{W}}_n(\mathcal{Y} \times \mathcal{Y}) &:= \{ \pi \in \tilde{\mathcal{W}}_*(\mathcal{Y} \times \mathcal{Y}) : \frac{n}{1+\mathbb{1}(a=b)} \pi(b, a) \in \mathbb{N} \text{ for all } b, a \in \mathcal{Y} \}. \end{aligned}$$

1.2 The Conditional Symbolled Random Graph Models

In the remaining part of this article we may assume that $\nu(a) > 0$ for all $a \in \mathcal{Y}$. Note that the law of the symbolled random graph given the empirical symbol distribution ν_n and empirical pair distribution π_n ,

$$\mathbb{P}_{(\nu_n, \pi_n)} = \mathbb{P}\{ \cdot \mid \Delta(M_Y) = (\nu_n, \pi_n) \},$$

may be described as follows:

- We assign symbols to the nodes by drawing from the pool of n symbols which contains any symbol $a \in \mathcal{Y}$, $n\nu_n(a)$ times without replacement;
- For each unordered pair $\{a, b\}$ of symbols we construct (exactly) $m_n(b, a)$ edges by drawing without replacement from the collection of potential edges linking nodes of symbol a and b , where

$$m_n(b, a) := \begin{cases} n \pi_n(b, a), & \text{if } a \neq b, \\ \frac{n}{2} \pi_n(b, a), & \text{if } a = b. \end{cases}$$

By Y_n we denote the conditional symbolled random graph with empirical symbol measure ν_n and empirical pair measure π_n .

2. Main Results

The main theorem in this section is an LDP for the proportion of isolated nodes in the random graph $\mathcal{G}(n, nc/2)$. We recall from (Doku-Amponsah & Moerters, 2010), the *empirical degree measure* $D_Y \in \mathcal{W}(\mathbb{N})$ of the symbolled random as

$$D_Y(k) = \sum_{b \in \mathcal{Y}} \sum_{l \in \mathbb{N}^{\mathcal{Y}}} \delta_k(\sum_{a \in \mathcal{Y}} l(a)) M_Y(b, l), \quad \text{for } k \in \mathbb{N}.$$

Theorem 1 *Suppose D_Y is the empirical degree measure of the random graph $\mathcal{G}(n, nc/2)$. Then, as $n \rightarrow \infty$, the proportion of isolated nodes $D_Y(0)$, obeys an LDP with good, convex rate function*

$$\eta(x) = \begin{cases} x \log \frac{x}{e^{-c}} + (1-x) \log \frac{(1-x)}{(1-e^{-c})} + c \log \lambda - c \log c, & \text{if } x \geq 1-c, \\ \infty & \text{if } x < 1-c, \end{cases} \tag{2}$$

where $\lambda = \lambda(x, c)$ is the unique root of $\frac{1-e^{-\lambda}}{\lambda} = \frac{1-x}{c}$.

Theorem 2 below and the contraction principle imply the LDP for the proportion of isolated vertices, i.e. Theorem 1. O’Connell (1998) obtained similar large deviation result for the number of isolated nodes in the random graph $\mathcal{G}(n, c/n)$.

Theorem 2 (Doku-Amponsah, 2006, 2012) *Suppose D_Y is the empirical degree measure of the random graph $\mathcal{G}(n, nc/2)$. Then, as $n \rightarrow \infty$, D_Y obeys an LDP in the space $\mathcal{W}(\mathbb{N})$ with good, convex rate function*

$$\delta(d) = \begin{cases} H(d \parallel q_c), & \text{if } \langle d \rangle = c, \\ \infty, & \text{otherwise.} \end{cases} \tag{3}$$

where q_x is a Poisson distribution with parameter x .

Here we remark, that the LDP result of Boucheron et al. (2002, Theorem 7.1) holds and the conjecture that Boucheron et al. (2002, Lemma 7.2) does not hold is false. In fact the coupling argument of Boucheron et al. (2002) and Bennetts inequality (see Bennett, 1962), proves Theorem 7.1 of (Boucheron et al., 2002). Recently, Bordenave et al. (2002, Corollary 1.9) and Mukherjee (2013) confirm 2.

Note, the degree distribution D_Y is a continuous function of M_Y , and so Theorem 3 below and the contraction principle gives the LDP for D_Y . In fact the LDP for D_Y (above) is a special case of Theorem 3 which was first proved in Doku-Amponsah and Moerters (2010) or Doku-Amponsah (2012) by approximating a given symbolled random graph from below by another symbolled random graphs with the degree of each nodes growing in order of $o(n^{1/3})$ (See Doku-Amponsah & Moerters, 2010, Lemma 4.10, pp. 26-29). Note, that in the special case of classical Erdős-Rényi graph $\mathcal{G}(n, nc/2)$ we have that $M = D$, the degree distribution and

$$\langle \Delta(M_Y) \rangle = 2|E|/n = c.$$

Theorem 3 (Doku-Amponsah & Moerters, 2010) *Suppose the sequence (ν_n, π_n) in $\mathcal{W}_n(\mathcal{Y}) \times \tilde{\mathcal{W}}_n(\mathcal{Y} \times \mathcal{Y})$ posses a limit (ν, π) in $\mathcal{W}(\mathcal{Y}) \times \tilde{\mathcal{W}}_*(\mathcal{Y} \times \mathcal{Y})$. Let Y be a symbolled random graph with n nodes conditioned on the event $\{\Delta(M_Y) = (\nu_n, \pi_n)\}$. Then, as $n \rightarrow \infty$, the empirical Neighbourhood distribution M_Y obeys an LDP in $\mathcal{W}(\mathcal{Y} \times \mathbb{N}^{\mathcal{Y}})$ with good rate function*

$$\tilde{J}_{(\nu, \pi)}(\mu) = \begin{cases} H(\mu \| Poi), & \text{if } (\pi, \mu) \text{ is sub-consistent and } \mu_1 = \nu, \\ \infty, & \text{otherwise,} \end{cases} \tag{4}$$

$$Poi(a, l) := \mu_1(a) \prod_{b \in \mathcal{Y}} e^{-\frac{\pi(b,a)}{\mu_1(a)}} \frac{1}{l(b)!} \left(\frac{\pi(b,a)}{\mu_1(a)} \right)^{l(b)}, \quad \text{for } a \in \mathcal{Y}, l \in \mathbb{N}^{\mathcal{Y}}.$$

3. Proof of Main Results

3.1 Proof of Theorem 3: Exponential Approximation by Random Allocation

In order to improve (shorten) the proof of Theorem 3, we pass to a simple *random allocation model*, which turns out to be equivalent. This model is best described in term of symbolled balls being placed randomly into symbolled bins.

Fix $n \geq 1$, a symbol law $\nu_n \in \mathcal{W}_n(\mathcal{Y})$ and an edge law $\pi_n \in \tilde{\mathcal{W}}_n(\mathcal{Y} \times \mathcal{Y})$. The bins $\mathcal{V} = \{1, \dots, n\}$ are now symbolled by drawing without replacement from the pool of symbols, which contains the symbol $a \in \mathcal{Y}$ exactly $n\nu_n(a)$ times. For each ordered pair $(b, a) \in \mathcal{Y} \times \mathcal{Y}$ of symbols, we independently and identically place $n\pi_n(b, a)$ balls of symbol b into the $n\nu_n(a)$ bins of symbol a by drawing without replacement. We denote by $\tilde{\mathbb{P}}_{(\nu_n, \pi_n)}$ the distribution of the random allocation model with symbol law $\nu_n \in \mathcal{W}_n(\mathcal{Y})$ and an edge law $\pi_n \in \tilde{\mathcal{W}}_n(\mathcal{Y} \times \mathcal{Y})$.

In the resulting constellation we denote, for any bin $v \in \{1, \dots, n\}$, by $\tilde{Y}(v)$ its symbol, and by $l^v(b)$ the number of balls of symbol $b \in \mathcal{Y}$ it contains. Now define the *empirical occupancy measure* of the constellation by

$$M_{\tilde{Y}}(b, l) = \frac{1}{n} \sum_{u \in \mathcal{V}} \delta_{(\tilde{Y}(u), \tilde{L}(u))}(b, l), \quad \text{for } (b, l) \in \mathcal{Y} \times \mathbb{N}^{\mathcal{Y}},$$

where $\tilde{L}(u) = (l^u(a), a \in \mathcal{Y})$ is the symbol distribution in bin v . In our first theorem we establish exponential equivalence of the law of the empirical occupancy measure $M_{\tilde{Y}}$ under the random allocation model $\tilde{\mathbb{P}}_{(\nu_n, \pi_n)}$ and the law of the empirical Neighbourhood distribution M under

$$\mathbb{P}_{(\nu_n, \pi_n)} = \mathbb{P}\{\cdot \mid \Delta(M_Y) = (\nu_n, \pi_n)\},$$

the law of the symbolled random graph conditioned to have symbol law ν_n and edge distribution π_n . Recall the definition of exponential equivalence (see Dembo & Zeitouni, 1998, Definition 4.2.10).

Lemma 4 *The law of $M_{\tilde{Y}}$ under $\tilde{\mathbb{P}}_{(\nu_n, \pi_n)}$ is exponentially equivalent to the law of M_Y under $\mathbb{P}_{(\nu_n, \pi_n)}$.*

Proof. Define the metric d of total variation by

$$d(\mu, \tilde{\mu}) = \frac{1}{2} \sum_{(a,l) \in \mathcal{Y} \times \mathbb{N}^{\mathcal{Y}}} |\mu(a, l) - \tilde{\mu}(a, l)|, \quad \text{for } \mu, \tilde{\mu} \in \mathcal{W}(\mathcal{Y} \times \mathbb{N}^{\mathcal{Y}}).$$

As this metric produces the weak topology, the proof of Lemma 4 is equivalent to showing that for every $\varepsilon > 0$,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}\{d(M_{\bar{Y}}, M_Y) \geq \varepsilon\} = -\infty, \tag{5}$$

where \mathbb{P} indicates a suitable coupling between the random allocation model and the symbolled graph.

To begin, denote by $V(a)$ the collection of nodes(bins) which have symbol $a \in \mathcal{Y}$ and observe that

$$\#V(a) = nv_n(a).$$

For every $a, b \in \mathcal{Y}$, begin: At each step $k = 1, \dots, m_n(b, a)$, we randomly pick two nodes $V_1^k \in V(a)$ and $V_2^k \in V(b)$. Drop one ball of symbol b in bin V_1^k and one ball of symbol a in V_2^k , and link V_1^k to V_2^k by an edge unless $V_1^k = V_2^k$ or the two nodes are already linked. If one of these two things happen, then we simply choose an edge randomly from the set of all potential edges linking symbols a and b , which are not yet present in the graph. This completes the construction of a graph with $L_Y^1 = v_n, L_Y^2 = \pi_n$ and

$$d(M_Y, M_{\bar{Y}}) \leq \frac{2}{n} \sum_{a,b \in \mathcal{Y}} B^n(b, a), \tag{6}$$

where $B^n(b, a)$ is the total number of steps $k \in \{1, \dots, m_n(b, a)\}$ at which there is discrepancy between the vertices V_1^k, V_2^k drawn and the nodes which formed the k^{th} edge connecting a and b in the random graph construction.

Given $a, b \in \mathcal{Y}$, the probability that $V_1^k = V_2^k$ or the two nodes are already linked is equal to

$$p_{[k]}(b, a) := \frac{1}{m_n(b, a)} \mathbb{1}_{\{b=a\}} + \left(1 - \frac{1}{m_n(b, a)} \mathbb{1}_{\{b=a\}}\right) \frac{(k-1)}{(m_n(b, a))^2}.$$

$B^n(b, a)$ is a sum of independent Bernoulli random variables $X_1, \dots, X_{m_n(b, a)}$ with ‘success’ probabilities equal to $p_{[1]}(b, a), \dots, p_{[m_n(b, a)]}(b, a)$. Note that $\mathbb{E}[X_k] = p_{[k]}(b, a)$ and

$$\text{Var}[X_k] = p_{[k]}(b, a)(1 - p_{[k]}(b, a)).$$

Now, we have

$$\mathbb{E}B^n(b, a) = \sum_{k=1}^{m_n(b, a)} p_{[k]}(b, a) = \mathbb{1}_{\{b=a\}} + \frac{1}{2} \left(1 - \mathbb{1}_{\{b=a\}} \frac{1}{m_n(b, a)}\right) \left(1 - \frac{1}{m_n(b, a)}\right) \leq \frac{1}{2} + \mathbb{1}_{\{b=a\}}.$$

We write

$$\sigma_n^2(b, a) := \frac{1}{m_n(b, a)} \sum_{k=1}^{m_n(b, a)} \text{Var}[X_k]$$

and observe that

$$\lim_{n \rightarrow \infty} \mathbb{E}(B^n(b, a)) = \lim_{n \rightarrow \infty} \text{Var}(B^n(b, a)) = \lim_{n \rightarrow \infty} m_n(b, a) \sigma_n^2(b, a) = \mathbb{1}_{\{b=a\}} + \frac{1}{2}.$$

We Define $h(x) = (1 + x) \log(1 + x) - x$, for $x \geq 0$ and use Bennett’s inequality (see Bennett, 1962), to obtain, for sufficiently large n

$$\mathbb{P}\left\{\frac{B^n(b, a)}{n} \geq \frac{\mathbb{1}_{\{b=a\}} + \frac{1}{2}}{n} + \delta_1\right\} \leq \exp\left[-m_n(b, a) \sigma_n^2(b, a) h\left(\frac{n\delta_1}{m_n(b, a) \sigma_n^2(b, a)}\right)\right],$$

for any $\delta_1 > 0$. Let $\varepsilon \geq 0$ and choose $\delta_1 = \frac{\varepsilon}{2m^2}$. Suppose that we have $B^n(b, a) \leq \delta$. Then, by (6),

$$d(M_Y, \mu_n) \leq 2\delta_1 m^2 = \varepsilon.$$

Hence,

$$\begin{aligned} \mathbb{P}\{d(M_Y, M_{\bar{Y}}) > \varepsilon\} &\leq \sum_{a,b \in \mathcal{Y}} \mathbb{P}\{B^n(b, a) \geq n\delta_1\} \leq m^2 \sup_{a,b \in \mathcal{Y}} \mathbb{P}\{B^n(b, a) \geq \mathbb{1}_{\{b=a\}} + \frac{1}{2} + (n\delta_1)/2\} \\ &\leq m^2 \sup_{a,b \in \mathcal{Y}} \exp\left[-m_n(b, a) \sigma_n^2(b, a) h\left(\frac{n\delta_1}{m_n(b, a) \sigma_n^2(b, a)}\right)\right]. \end{aligned}$$

Let $0 \leq \delta_2 \leq 1$. Then, for sufficiently large n we have

$$\begin{aligned} \frac{1}{n} \log \mathbb{P}\{d(M_Y, M_{\tilde{Y}}) > \varepsilon\} &\leq -(1 - \delta_2)h\left(\frac{n\delta_1}{2(1+\delta_2)}\right) \\ &= -\left(\mathbb{1}_{\{b=a\}} + \frac{1}{2} - \delta_2\right)\left[\frac{1}{n} + \frac{\delta_1}{2(\mathbb{1}_{\{b=a\}} + \frac{1}{2} + \delta_2)}\right] \log\left(1 + \frac{n\delta_1}{2(\mathbb{1}_{\{b=a\}} + \frac{1}{2} + \delta_2)}\right) - \frac{\delta_1}{2(\mathbb{1}_{\{b=a\}} + \frac{1}{2} + \delta_2)}. \end{aligned} \tag{7}$$

This completes the proof of the lemma. □

3.2 Proof of Theorem 3: Large Deviation Probabilities by the Method of Types

Now Lemma 4 and the large deviation principle for $M_{\tilde{Y}}$ under $\tilde{\mathbb{P}}_{(v_n, \pi_n)}$ implies the same large deviation principle for M_Y under $\mathbb{P}_{(v_n, \pi_n)}$ in the weak topology (See, for example, Dembo & Zeitouni, 1998, Theorem 4.2.13). Consequently, the proof of Theorem 3 is equivalent to showing that for every $\Sigma \subset \mathcal{W}(\mathcal{Y} \times \mathbb{N}^{\mathcal{Y}})$,

Lemma 5

$$\begin{aligned} - \inf_{\mu \in \text{int}(\Sigma)} \tilde{J}_{(v, \pi)}(\mu) &\leq \liminf_{n \rightarrow \infty} \frac{1}{n} \log \tilde{\mathbb{P}}\{M_{\tilde{Y}} \in \Sigma \mid \Delta(M_{\tilde{Y}}) = (v_n, \pi_n)\} \\ &\leq \limsup_{n \rightarrow \infty} \frac{1}{n} \log \tilde{\mathbb{P}}\{M_{\tilde{Y}} \in \Sigma \mid \Delta(M_{\tilde{Y}}) = (v_n, \pi_n)\} \leq - \inf_{\mu \in \text{cl}(\Sigma)} \tilde{J}_{(v, \pi)}(\mu). \end{aligned}$$

We begin the proof of Lemma 5 by recalling the definition of Δ a function on $\mathcal{W}(\mathcal{Y} \times \mathbb{N}^{\mathcal{Y}})$ given by

$$\mu \mapsto (\Delta_1(\mu), \Delta_2(\mu)).$$

Let $(v_n, \pi_n) \rightarrow (v, \pi) \in \mathcal{W}(\mathcal{Y}) \times \mathcal{W}(\mathcal{Y} \times \mathcal{Y})$ and write

$$\mathcal{K}^{(n)}(v_n, \pi_n) = \{\mu_n : \mu_n = M, \Delta(\mu_n) = (v_n, \pi_n), \text{ for some random allocations process on } n \text{ bins}\}.$$

We denote by $\mathcal{S}(\mu)$ the support of μ and write for $\mu_n \in \mathcal{K}^{(n)}(v_n, \pi_n)$,

$$\vartheta_1^{(n)}(\pi_n, \mu_n) = n\alpha_1^{(n)}(v_n, \pi_n) - n\beta_1^{(n)}(\mu_n) - \frac{1}{2n}|\mathcal{S}(\mu_n)| \log 2\pi n,$$

where

$$\begin{aligned} \alpha_1^{(n)}(v_n, \pi_n) &= -\frac{1}{n} \log |\mathcal{K}^{(n)}(v_n, \pi_n)| + \frac{1}{n} \sum_{a,b \in \mathcal{Y}} \log \pi_n(b, a) + \frac{1}{2n}(|\mathcal{Y}| + |\mathcal{Y}|^2) \log 2\pi n \\ &\quad + \frac{1}{n^2} \sum_{a \in \mathcal{Y}} \frac{1}{12v_n(a)+1/n} + \frac{1}{n} \sum_{a \in \mathcal{Y}} \log v_n(a) + \frac{1}{n^2} \sum_{a,b \in \mathcal{Y}} \frac{1}{12\pi_n(b,a)+1/n}, \\ \beta_1^{(n)}(\mu_n) &= \frac{1}{n} \sum_{\substack{(a,l) \in \mathcal{Y} \times \mathbb{N}^{\mathcal{Y}} \\ \mu_n(a,l) > 0}} \log \mu_n(a, l) + \frac{1}{n^2} \sum_{\substack{(a,l) \in \mathcal{Y} \times \mathbb{N}^{\mathcal{Y}} \\ \mu_n(a,l) > 0}} \frac{1}{12\mu_n(a,l)+1/n}. \end{aligned}$$

We write $\vartheta_2^{(n)}(\pi_n, \mu_n) = n\alpha_2^{(n)}(v_n, \pi_n) - n\beta_2^{(n)}(\mu_n)$, where

$$\begin{aligned} \alpha_2^{(n)}(v_n, \pi_n) &= \frac{1}{n} \log |\mathcal{K}^{(n)}(v_n, \pi_n)| + \frac{1}{n^2} \sum_{a,b \in \mathcal{Y}} \frac{1}{12\pi_n(b,a)} + \frac{1}{n^2} \sum_{\substack{a \in \mathcal{Y} \\ v_n(a) > 0}} \frac{1}{12v_n(a)} \\ &\quad + \frac{1}{n} \sum_{a,b \in \mathcal{Y}} \log \pi_n(b, a) + \frac{1}{2n}(|\mathcal{Y}| + |\mathcal{Y}|^2) \log 2\pi n + \frac{1}{n} \sum_{a \in \mathcal{Y}} \log v_n(a), \\ \beta_2^{(n)}(\mu_n) &= \frac{1}{n} \sum_{\substack{(a,l) \in \mathcal{Y} \times \mathbb{N}^{\mathcal{Y}} \\ \mu_n(a,l) > 0}} \log \mu_n(a, l) + \frac{1}{n} \sum_{\substack{(a,l) \in \mathcal{Y} \times \mathbb{N}^{\mathcal{Y}} \\ \mu_n(a,l) > 0}} \frac{1}{12\mu_n(a,l)} \end{aligned}$$

We prove Lemma 5 above from the following lemma which uses the the idea of the method of types (see Dembo & Zeitouni, 1998, Chapter 2).

Lemma 6 For any $\mu_n \in \mathcal{K}^{(n)}(v_n, \pi_n)$,

$$e^{-nH(\mu_n \parallel \text{Poi}_n) + \vartheta_1^{(n)}(\pi_n, \mu_n)} \leq \tilde{\mathbb{P}}\{M_{\tilde{Y}} = \mu_n \mid \Delta(M_{\tilde{Y}}) = (v_n, \pi_n)\} \leq |\mathcal{K}^{(n)}(v_n, \pi_n)|^{-1} e^{-nH(\mu_n \parallel \text{Poi}_n) + \vartheta_2^{(n)}(\pi_n, \mu_n)},$$

where

$$Poi_n(a, l) = v_n(a) \prod_{b \in \mathcal{Y}} \frac{e^{-\pi_n(b,a)/v_n(a)} [\pi_n(b, a)/v_n(a)]^{l(b)}}{l(b)!}, \text{ for } l \in \mathbb{N}^{\mathcal{Y}}$$

$$\lim_{n \rightarrow \infty} \vartheta_2^{(n)}(\pi_n, \mu_n) = \lim_{n \rightarrow \infty} \vartheta_1^{(n)}(\pi_n, \mu_n) = 0$$

Proof. The proof of this lemma uses the idea of the method of types (see Dembo & Zeitouni, 1998, Chapter 2), combinatoric argument and good estimates from refined Stirling’s formula.

We denote by \tilde{Y} the random allocation process and observe that for any $\mu_n \in \mathcal{K}^{(n)}(v_n, \pi_n)$ we have

$$\tilde{\mathbb{P}}\{M_{\tilde{Y}} = \mu_n \mid \Delta(M_{\tilde{Y}}) = (v_n, \pi_n)\} = \frac{\#\{\tilde{Y} : M_{\tilde{Y}} = \mu_n, \Delta(M_{\tilde{Y}}) = (v_n, \pi_n)\}}{\#\{\tilde{Y} : (L_{\tilde{Y}}^1, L_{\tilde{Y}}^2) = (v_n, \pi_n)\}}. \tag{8}$$

Now, the right side of (8) may be evaluated in the following way:

- For a given empirical measure μ_n with $\Delta(\mu_n) = (v_n, \pi_n)$ there are

$$\prod_{a \in \mathcal{Y}} \binom{nv_n(a)}{n\mu_n(a, l), l \in \mathbb{N}^{\mathcal{Y}}} \prod_{a, b \in \mathcal{Y}} \binom{n\pi_n(b, a)}{l_a^{(j)}(b), j = 1, \dots, nv_n(a)}$$

equally likely random allocation processes and

- for every empirical and empirical pair measure $\Delta(\mu_n) = (v_n, \pi_n)$ there are $\prod_{a, b \in \mathcal{Y}} (nv_n(a))^{n\pi_n(b,a)}$ equally likely random allocation processes.

Therefore, (8) is equivalent to

$$\tilde{\mathbb{P}}\{\tilde{M} = \mu_n \mid \Delta(\tilde{M}) = (v_n, \pi_n)\} \tag{9}$$

$$= \prod_{a \in \mathcal{Y}} \binom{nv_n(a)}{n\mu_n(a, l), l \in \mathbb{N}^{\mathcal{Y}}} \prod_{a, b \in \mathcal{Y}} \binom{n\pi_n(b, a)}{l_a^{(j)}(b), j = 1, \dots, nv_n(a)} \left(\frac{1}{nv_n(a)}\right)^{n\pi_n(b,a)}, \tag{10}$$

while $\tilde{\mathbb{P}}\{\tilde{M} = \mu_n \mid \Delta(\tilde{M}) = (v_n, \pi_n)\} = 0$ when $\Delta(\mu_n) \neq (v_n, \pi_n)$ by convention.

Suppose $\pi_n(b, a) = 0$, for some $a, b \in \mathcal{Y}$ then

$$\binom{n\pi_n(b, a)}{l_a^{(j)}(b), j = 1, \dots, nv_n(a)} = 1. \tag{11}$$

Suppose $\pi_n(b, a) > 0$, a good estimate of $(n\pi_n(b, a))!$ can be obtained from the refined Stirling’s approximation, as

$$\begin{aligned} \exp\left(n\pi_n(b, a) \log n\pi_n(b, a) - n\pi_n(b, a) + \frac{1}{2} \log \pi_n(b, a) + \frac{1}{12n\pi_n(b,a)+1} + \frac{1}{2} \log 2n\pi\right) &\leq (n\pi_n(b, a))! \\ &\leq \exp\left(n\pi_n(b, a) \log n\pi_n(b, a) - n\pi_n(b, a) + \frac{1}{2} \log \pi_n(b, a) + \frac{1}{12n\pi_n(b,a)} + \frac{1}{2} \log 2\pi n\right). \end{aligned}$$

Similarly, from the refined Stirling’s approximation (see Feller, 1971, p. 52), we have

$$\begin{aligned} &\exp\left(n \sum_{a \in \mathcal{Y}} v_n(a) \log v_n(a) - n \sum_{(a,l)} \mu_n(a, l) \log \mu_n(a, l) + \frac{1}{n} \sum_{a \in \mathcal{Y}} \frac{1}{12v_n(a)+1/n} + \frac{|\mathcal{Y}|-|S(\mu_n)|}{2} \log 2\pi n\right) \\ &\times \exp\left(-\frac{1}{n} \sum_{\substack{(a,l) \in \mathcal{Y} \times \mathbb{N}^{\mathcal{Y}} \\ \mu_n(a,l) > 0}} \frac{1}{12\mu_n(a,l)+1/n}\right) \leq \prod_{a \in \mathcal{Y}} \binom{nv_n(a)}{n\mu_n(a, l), l \in \mathbb{N}^{\mathcal{Y}}} \leq \exp\left(n \sum_{a \in \mathcal{Y}} v_n(a) \log v_n(a)\right) \\ &\times \exp\left(-n \sum_{(a,l)} \mu_n(a, l) \log \mu_n(a, l) - \frac{1}{n} \sum_{\substack{(a,l) \in \mathcal{Y} \times \mathbb{N}^{\mathcal{Y}} \\ \mu_n(a,l) > 0}} \frac{1}{12\mu_n(a,l)} + \frac{|\mathcal{Y}|-|S(\mu_n)|}{2} \log 2\pi n + \frac{1}{n} \sum_{a \in \mathcal{Y}} \frac{1}{12v_n(a)}\right). \end{aligned} \tag{12}$$

We observe that $\prod_{a,b \in \mathcal{Y}} \prod_{j=1}^{nv_n(a)} l_a^{(j)}(b)! = \prod_{b \in \mathcal{Y}} \exp(n \sum_{(a,l)} (\log l(b)!) \mu_n(b, a))$, and hence

$$\begin{aligned} & \exp\left(\sum_{b \in \mathcal{Y}} \left[n \sum_{a \in \mathcal{Y}} \pi_n(b, a) \log \pi_n(b, a) - n \sum_{a \in \mathcal{Y}} \pi_n(b, a) - n \sum_{a \in \mathcal{Y}} \pi_n(b, a) \log v_n(a)\right] + \frac{1}{2} \sum_{a,b \in \mathcal{Y}} \log \pi_n(b, a)\right) \\ & \times \exp\left(n \sum_{b \in \mathcal{Y}} \sum_{(a,l)} (\log l(b)!) \mu_n(a, l) + \frac{|\mathcal{Y}|^2}{2} \log 2\pi n + \sum_{a,b \in \mathcal{Y}} \frac{1}{12n\pi_n(b,a)+1}\right) \\ & \leq \prod_{a,b \in \mathcal{Y}} \left(l_a^{(j)}(b), j = 1, \dots, nv_n(a)\right) \left(\frac{1}{nv_n(a)}\right)^{n\pi_n(b,a)} \\ & \leq \exp\left(n \sum_{b \in \mathcal{Y}} \sum_{(a,l)} (\log l(b)!) \mu_n(a, l) + \frac{1}{2} \sum_{a,b \in \mathcal{Y}} \log \pi_n(b, a)\right) \\ & \times \exp\left(\sum_{b \in \mathcal{Y}} \left[n \sum_{a \in \mathcal{Y}} \pi_n(b, a) \log \sum_{a \in \mathcal{Y}} \pi_n(b, a) - n \sum_{a \in \mathcal{Y}} \pi_n(b, a) - n \sum_{a \in \mathcal{Y}} \pi_n(b, a) \log v_n(a)\right] + \sum_{a,b \in \mathcal{Y}} \frac{1}{12n\pi_n(b,a)}\right) \\ & \times \exp\left(\frac{|\mathcal{Y}|^2}{2} \log 2\pi n\right). \end{aligned}$$

Putting everything together and choosing $\vartheta_1^{(n)}(\pi_n, \mu_n)$ and $\vartheta_2^{(n)}(\pi_n, \mu_n)$ appropriately, we have that

$$\begin{aligned} & \exp\left(nH(\mu_n) + \sum_{b \in \mathcal{Y}} \left[n \sum_{a \in \mathcal{Y}} \pi_n(b, a) \log \sum_{a \in \mathcal{Y}} \pi_n(b, a) - n \sum_{a \in \mathcal{Y}} \pi_n(b, a) - n \sum_{a \in \mathcal{Y}} \pi_n(b, a) \log v_n(a)\right]\right) \\ & \times \exp\left(-nH(v) - n \sum_{b \in \mathcal{Y}} \sum_{(a,l)} (\log l(b)!) \mu_n(a, l) + \vartheta_1^{(n)}(\pi_n, \mu_n)\right) \leq \tilde{\mathbb{P}}\{M = \mu_n \mid \Delta(M) = (v_n, \pi_n)\} \\ & \leq \exp\left(-nH(v) - \sum_{b \in \mathcal{Y}} \left[n \sum_{(a,l)} (\log l(b)!) \mu_n(a, l) - n \sum_{a \in \mathcal{Y}} \pi_n(b, a) - n \sum_{a \in \mathcal{Y}} \pi_n(b, a) \log v_n(a)\right]\right) \\ & \times \exp\left(nH(\mu_n) + n \sum_{b \in \mathcal{Y}} \sum_{a \in \mathcal{Y}} \pi_n(b, a) \log \sum_{a \in \mathcal{Y}} \pi_n(b, a) - \log |\mathcal{K}^{(n)}(v_n, \pi_n)| + \vartheta_2^{(n)}(\pi_n, \mu_n)\right). \end{aligned}$$

Collecting and rearranging terms properly and using $\Delta(\mu_n) = (v_n, \pi_n)$, we have that

$$\begin{aligned} & H(v_n) - H(\mu_n) - \sum_{b \in \mathcal{Y}} \left[\sum_{(a,l)} l(b) \mu_n(a, l) \log \sum_{(a,l)} l(b) \mu_n(a, l) - \sum_{(a,l)} l(b) \mu_n(a, l) \right. \\ & \left. - \sum_{(a,l)} l(b) \mu_n(a, l) \log v_n(a) - \sum_{(a,l)} (\log l(b)!) \mu_n(a, l) \right] \\ & = \sum_{(a,l)} \mu_n(a, l) \left[\log \mu_n(a, l) - \log v_n(a) - \sum_{b \in \mathcal{Y}} \left(\log \left(\frac{\pi_n(b,a)}{v_n(a)} \right)^{l(b)} - \frac{\pi_n(b,a)}{v_n(a)} - \log \sum_l (\log l(b)!) \right) \right] \\ & = \sum_{(a,l)} \mu_n(a, l) \left[\log \mu_n(a, l) - \log (v_n(a) \prod_{b \in \mathcal{Y}} \frac{(\pi_n(b,a)/v_n(a))^{l(b)} \exp(-\pi_n(b,a)/v_n(a))}{l(b)!}) \right] \\ & = H(\mu_n \parallel Poi_n) \end{aligned}$$

which completes the proof of Lemma 6. □

We prove from Lemma 6 and Doku-Amponsah and Moerters (2010, Lemmas 4.1 and 4.4), upper bounds and lower bounds in the large deviation principle for all finite n . Let $\Sigma \subset \mathcal{W}(\mathcal{Y} \times \mathbb{N}^{\mathcal{Y}})$. Then Lemma 6 gives the upper bound

$$\begin{aligned} \tilde{\mathbb{P}}\{M_{\tilde{Y}} \in \Sigma \mid \Delta(M_{\tilde{Y}}) = (v_n, \pi_n)\} &= \sum_{\mu_n \in \Sigma \cap \mathcal{K}^{(n)}(v_n, \pi_n)} \tilde{\mathbb{P}}\{M_{\tilde{Y}} = \mu_n \mid \Delta(M_{\tilde{Y}}) = (v_n, \pi_n)\} \\ &\leq \sum_{\mu_n \in \Sigma \cap \mathcal{K}^{(n)}(v_n, \pi_n)} |\mathcal{K}^{(n)}(v_n, \pi_n)|^{-1} e^{-nH(\mu_n \parallel Poi_n) + n\vartheta_2^{(n)}(v_n, \pi_n)} \\ &\leq e^{-n \inf_{\mu_n \in \Sigma \cap \mathcal{K}^{(n)}(v_n, \pi_n)} H(\mu_n \parallel Poi_n) + n\vartheta_2^{(n)}(v_n, \pi_n)}. \end{aligned} \tag{13}$$

The corresponding lower bound is

$$\begin{aligned} \tilde{\mathbb{P}}\{M_{\bar{Y}} \in \Sigma \mid \Delta(M_{\bar{Y}}) = (v_n, \pi_n)\} &= \sum_{\mu_n \in \Sigma \cap \mathcal{K}^{(n)}(v_n, \pi_n)} \tilde{\mathbb{P}}\{M_{\bar{Y}} = \mu_n \mid \Delta(M_{\bar{Y}}) = (v_n, \pi_n)\} \\ &\geq e^{-n \inf_{\mu_n \in \Sigma \cap \mathcal{K}^{(n)}(v_n, \pi_n)} H(\mu_n \parallel Poi_n)} \sum_{\mu_n \in \Sigma \cap \mathcal{K}^{(n)}(v_n, \pi_n)} e^{\theta_1^{(n)}(v_n, \pi_n)} \\ &\geq e^{-n \inf_{\mu_n \in \Sigma \cap \mathcal{K}^{(n)}(v_n, \pi_n)} H(\mu_n \parallel Poi_n) + n\theta_1^{(n)}(v_n, \pi_n)}. \end{aligned} \tag{14}$$

Since $\limsup_{n \rightarrow \infty} \frac{1}{n} \theta_2^{(n)}(v_n, \pi_n) = 0$ and $\liminf_{n \rightarrow \infty} \frac{1}{n} \theta_1^{(n)}(v_n, \pi_n) = 0$ (by Doku-Amponsah & Moerters, 2010, Lemmas 4.1 and 4.4), the normalized logarithmic limits of (13) and (14) gives

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \tilde{\mathbb{P}}\{M_{\bar{Y}} \in \Sigma \mid \Delta(M_{\bar{Y}}) = (v_n, \pi_n)\} = - \liminf_{n \rightarrow \infty} \left\{ \inf_{\mu_n \in \Sigma \cap \mathcal{K}^{(n)}(v_n, \pi_n)} H(\mu_n \parallel Poi_n) \right\} \tag{15}$$

and

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \tilde{\mathbb{P}}\{M_{\bar{Y}} \in \Sigma \mid \Delta(M_{\bar{Y}}) = (v_n, \pi_n)\} = - \limsup_{n \rightarrow \infty} \left\{ \inf_{\mu_n \in \Sigma \cap \mathcal{K}^{(n)}(v_n, \pi_n)} H(\mu_n \parallel Poi_n) \right\} \tag{16}$$

The upper bound in Lemma 6 follows from (15), as $\Sigma \cap \mathcal{K}^{(n)}(v_n, \pi_n) \subset \Sigma$ for all n .

Now fix $\mu \in \mathcal{W}(\mathcal{Y} \times \mathbb{N}^{\mathcal{Y}})$. Then, by Doku-Amponsah and Moerters (2010, Lemma 4.9), there exists a sequence $\mu_n \in \Sigma \cap \mathcal{K}^{(n)}(v_n, \pi_n)$ such that $\mu_n \rightarrow \mu$ as $n \rightarrow \infty$. Therefore, by continuity of entropy (see, example Doku-Amponsah & Moerters, 2010, p. 19, Equation (14)), we have that

$$\limsup_{n \rightarrow \infty} \left\{ \inf_{\mu' \in \Sigma \cap \mathcal{K}^{(n)}(v_n, \pi_n)} H(\mu' \parallel Poi_n) \right\} \leq \lim_{n \rightarrow \infty} H(\mu_n \parallel Poi_n) = H(\mu \parallel Poi).$$

Recall that $H(\mu \parallel Q) = \infty$ whenever, for some $(b, l) \in \mathcal{Y} \times \mathbb{N}^{\mathcal{Y}}$, $\mu(b, l) > 0$ while $Poi(b, l) = 0$. Hence, by the preceding inequality we have

$$\limsup_{n \rightarrow \infty} \left\{ \inf_{\mu' \in \Sigma \cap \mathcal{K}^{(n)}(v_n, \pi_n)} H(\mu' \parallel Poi_n) \right\} \leq \inf_{\mu \in \text{int}(\Sigma)} H(\mu \parallel Poi),$$

which gives the lower bound in Lemma 6, for μ satisfying $\Delta(\mu) = (v, \pi)$. To conclude the prove of the lower bound we note that by Doku-Amponsah and Moerters, (2010, Lemma 4.6) for any $\mu \in \Sigma$ with $\mu_1 = v$ and $\Delta_2(\mu) \leq \pi$ there exists $\mu_n \in \mathcal{K}^{(n)}(v_n, \pi_n)$ converging weakly to μ such that $H(\mu_n \parallel Poi_n)$ converges to $H(\mu \parallel Poi)$.

3.3 Proof of Theorem 1

We obtain this corollary from Theorem 2 by the application of the contraction principle, Dembo and Zeitouni (1998, Theorem 4.2.1) to the linear map $F: \mathcal{W}(\mathbb{N}) \rightarrow [0, 1]$ given by $F(d) = d(0)$.

In fact Theorem 2 implies an LDP for random variable $F(D_Y) = D_Y(0)$ with good, convex rate function

$$\eta(x) = \inf \{H(d \parallel q_c) : d \in \mathcal{W}(\mathbb{N}), d(0) = x, \sum_{k=0}^{\infty} kd(k) = c\}.$$

Note that, for a general x the class of distributions satisfying the two constraints might be non empty. Since we have

$$c = \sum_{k=1}^{\infty} kd(k) \geq \sum_{k=1}^{\infty} d(k) = 1 - x,$$

the class is necessarily empty if $c < 1 - x$. If $c \geq 1 - x$, a Lagrangian calculation gives that the minimum is attained at p , defined by $p(0) = x$, $p(k) := Z(x, c)^{-1} \frac{(\lambda(x, c))^k}{k!}$ where $\lambda(x, c)$ is the unique root of

$$\frac{e^\lambda - 1}{\lambda} = \frac{1 - x}{c}$$

and $Z(x, c) := \frac{e^{\lambda-1}}{1-x}$. Therefore we have that

$$\begin{aligned} \eta(x) &= x \log \frac{x}{q_c(0)} + (1 - x) \log \frac{(1 - x)}{1 - q_c(0)} + (1 - x) \sum_{k=1}^{\infty} d_x(k) \log \frac{d_x(k)}{\tilde{q}_c(k)}. \\ &= x \log \frac{x}{q_c(0)} + (1 - x) \log \frac{(1 - x)}{1 - q_c(0)} + c \log \frac{1}{c} \end{aligned} \tag{17}$$

if $c \geq 1 - x$ and ∞ otherwise. In particular if $x = e^{-c}$ then $\lambda(x, c) = c$, which gives $\eta(e^{-c}) = 0$. This completes the proof of the theorem.

Acknowledgements

We are thankful to the referees for their suggestions which have helped improved this article.

References

- Bennett, G. (1962). Probability Inequalities for the Sum of Independent Random Variables. *Journal of the American Statistical Association*, 57(297), 33-5. <http://dx.doi.org/10.2307/2282438>
- Biggins, J. D. (2004). Large deviations for mixtures. *El. Comm. Probab.*, 9, 60-71.
- Biggins, J. D., & Penman, D. B. (2009). Large deviations in randomly coloured random graphs. *Electron. Comm. Probab.*, 14, 290-01. <http://dx.doi.org/10.1214/ECP.v14-1478>
- Bordenave, C., & Caputo, P. (2013). Large deviations of empirical neighborhood distribution in sparse random graphs. arxiv:1308.5725 (2013).
- Boucheron, S., Gamboa, F., & Leonard, C. (2002). Bins and balls: Large deviations of the empirical occupancy process. *Ann. Appl. Probab.*, 12, 607-636. <http://dx.doi.org/10.1214/aoap/1026915618>
- Cannings, C., & Penman, D. B. (2003). Models of random graphs and their applications. In D. N. Shanbhag & C. R. Rao (Eds.), *Handbook of Statistics 21. Stochastic Processes: Modeling and Simulation* (pp. 51-91). Elsevier.
- Doku-Amponsah, K. (2006). Large deviations and basic information theory for hierarchical and networked data structures (PhD Thesis, Bath).
- Doku-Amponsah, K. (2011). Asymptotic equipartition properties for hierarchical and networked structures. *ESAIM: Probability and Statistics*. Cambridge University Press. <http://dx.doi.org/10.1051/ps/2010016>
- Doku-Amponsah, K. (2012). *Basic information theory, thermo-limits for network structures*. LAP Lambert Academic Publishing.
- Doku-Amponsah, K., & Mörters, P. (2010). Large deviation principle for empirical measures of coloured random graphs. *The annals of Applied Probability*, 20(6), 1989-2021. <http://dx.doi.org/10.1214/09-AAP647>
- Dembo, A., Mörters, P., & Sheffield, S. (2005). Large deviations of Markov chains indexed by random trees. *Ann. Inst. Henri Poincaré: Probab. et Stat.*, 41, 971-996. <http://dx.doi.org/10.1016/j.anihpb.2004.09.005>
- Dembo, A., & Zeitouni, O. (1998). *Large deviations techniques and applications*. New York: Springer. <http://dx.doi.org/10.1007/978-1-4612-5320-4>
- Feller, W. (1967). *An introduction to probability theory and its applications* (Vol. I, 3rd ed.). New York: Wiley.
- Mukherjee, S. (2013). Large deviation for the empirical degree distribution of an Erdos-Renyi graph.
- Newman, M. E. (2000). *Random graphs as models of networks*. Retrieved from <http://arxiv.org/abs/cond-mat/0202208>
- O'Connell, N. (1998). Some large deviation results for sparse random graphs. *Probab. Theory Relat. Fields*, 110, 277-285. <http://dx.doi.org/10.1007/s004400050149>
- Penman, D. B. (1998). Random graphs with correlation structure (PhD Thesis, Sheffield). <http://dx.doi.org/10.2140/pjm.1958.8.171>
- Sion, M. (1958). On general Minimax theorems. *Pacific J. Math.*, 8, 171-176.
- Van Der Hofstad, R. (2009). Random Graphs and Complex Networks. Eindhoven University of Technology. Unpublish Manuscript.

Notes

Note 1. $(2\pi)^{\frac{1}{2}} n^{n+\frac{1}{2}} e^{-n+1/(12n+1)} < n! < (2\pi)^{\frac{1}{2}} n^{n+\frac{1}{2}} e^{-n+1/(12n)}$ (see Feller, 1971, p. 52).

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).