Large deviations in randomly coloured random graphs

J. D. Biggins* and D. B. Penman†

September 24, 2003

Abstract

Models of random graphs where the presence or absence of an edge depends on the random types (colours) of its vertices, so that whether or not edges are present can be dependent, are considered. In particular, large deviations in the number of edges are studied. An application of a result on large deviations for mixtures allows a fairly complete treatment of this question. This is a natural example with two different non-degenerate large deviation regimes, one arising from large deviations in the colourings followed by typical edge placement and the other from large deviation in edge placement.

1 Introduction

This paper considers, in the first instance, models of random graphs where each vertex is independently assigned one of a number of colours, and the probability that an edge arises depends on the colours of its two vertices. In such a model, whether or not edges arise can be correlated. Let α be the overall probability that an edge arises. It is natural to compare these models with the classical model of random graphs $G(n, \alpha)$ (see [1]) where each edge arises independently with probability α, with the aim of understanding how the correlation structure modifies the behaviour of the random graphs. These models with correlation structure were introduced in [6]. Motivation for the construction and areas of application for the model are discussed in [3]. Familiarity with [3] and [6] is neither assumed nor needed here. However, it is worth stressing that the colouring is the mechanism that induces the correlation but is regarded as invisible. It is the characteristics of the resulting random graph that are the focus of attention.

Throughout, graphs will be finite, undirected, without loops or multiple edges, and with vertices labelled by the integers. Given n, the number of vertices, and k, the finite number of colours, each vertex is coloured independently, receiving colour i with probability $s_i > 0$. The probability that an edge between a vertex of colour i and one of colour j arises is p_{ij}; thus $\widetilde{P} = (p_{ij})$ is symmetric. We will consider the family of such models as n varies, with k, $\mathbf{s} = (s_1, s_2, \dots, s_k)$ and $\widetilde{P}$ being fixed. Note that if every entry in $\widetilde{P}$ is equal to α this framework will be equivalent to $G(n, \alpha)$, once the colours are ignored.

Let $\mathcal{E}$ be the number of edges present; if we want to emphasise the number of vertices, we shall write $\mathcal{E}(n)$. The probability of a large deviation in the number of edges $\mathcal{E}$ is considered; the basic form of a large deviation result is explained next.

*The University of Sheffield, U.K. (`j.biggins@sheffield.ac.uk`)
†The University of Essex, U.K. (`dbpenman@essex.ac.uk`)

Definition 1 *The sequence of probability measures (P_n) obeys a large deviation principle (LDP) with constants (a_n), tending to infinity, if there is a lower semicontinuous non-negative function I (a rate function) such that for every closed G and open F*

$$\limsup \frac{\log P_n\{F\}}{a_n} \leq -\inf_{y\in F} I(x) \quad \textit{and} \quad \liminf \frac{\log P_n\{G\}}{a_n} \geq -\inf_{y\in G} I(x).$$

The rate function I is called 'good' if for every finite β the set $\{x : I(x) \leq \beta\}$ is compact.

The basic case of large deviation theory is when $\{X_i\}$ are independent and identically distributed and P_n is the distribution of $S_n/n = \sum_{i=1}^{n} X_i/n$, see [4, Theorem 2.2.3]. That result immediately yields the large deviation behaviour of $\mathcal{E}$ in $G(n, \alpha)$.

In the random graph model considered here there are two sources of variation, the generation of the colours and the subsequent generation of the edges. Thus there are two sources of large deviations, in the colourings and in the numbers of edges given the colouring. The result is two different regimes, one arising from large deviations in both the colouring and the numbers of edges and the second from large deviations in the colourings and typical numbers of edges.

This intuition, and the way the model is described, suggests that its behaviour should be understood through large deviation results for mixtures. This motivated a separate study, extending results in [5], which concerns exchangeable variables, of that general question. A result derived from that study that suffices for this one is recorded in the next section. This approach produces, without difficulty, results for more general models. It is not essential to colour each vertex independently, what matters is that the distributions of the numbers of colours obeys a suitable LDP (which it does in the independent case).

We initially approached large deviations in the two colour case by using directly the Gärtner-Ellis theorem. That approach relies on computing the moment generating function of $\mathcal{E}(n)$ and examining its limiting behaviour, and the properties of the limit, as n goes to infinity. It seemed not to extend to the general case.

The next section describes the general large deviation results used. The third section looks at the large deviation behaviour of the number of edges given that the colouring is according to some fixed proportions asymtotically. These results combine with suitable assumptions about the colouring to give the sought for large deviation principles for the number of edges, which are discussed in the fourth section. Theorem 14 there describes the large deviation behaviour for the number of edges in the model when vertices are coloured independently. In the fifth section the rate functions that arise are described in more detail, allowing more direct comparison with $G(n, \alpha)$. Finally, an extension drawing colours from a larger set and associating a random variable with every edge is discussed briefly in the final section.

2 General LDP results.

The first result, which is a consequence of those described in [2], is given in a rather general setting. Once it has been stated a few remarks are offered to connect the generalities to the particular problem considered here.

Let $\mathcal{Y}$ be a Hausdorff, regular, topological space and Δ be a Polish space. Let Δ_n be closed subsets of Δ. Suppose that for all $\mathbf{x} \in \Delta$ there are $\mathbf{x}_n \in \Delta_n$ with $\mathbf{x}_n \to \mathbf{x}$. Let μ^n be a probability measure on the Borel σ-algebra of Δ concentrated on Δ_n; μ^n

is the mixing distribution. For each $\mathbf{x} \in \Delta_n$, let $P^n_{\mathbf{x}}$ be a probability measure on the Borel σ-algebra of $\mathcal{Y}$, such that, for every measurable $A \subset \mathcal{Y}$, the map $\mathbf{x} \to P^n_{\mathbf{x}}(A)$ is measurable on Δ_n; $P^n_{\mathbf{x}}$ is the conditional distribution, given $\mathbf{x}$. The marginal distribution, P^n, mixing over $\mathbf{x}$ with μ^n, is given by

$$P^n(A) = \int_{\Delta} P^n_{\mathbf{x}}(A) d\mu^n(\mathbf{x}) \quad \left(= \int_{\Delta_n} P^n_{\mathbf{x}}(A) d\mu^n(\mathbf{x}) \right).$$

The general result gives conditions for (P^n) to obey an LDP when (μ^n) and $(P^n_{\mathbf{x}})$ satisfy suitable LDPs.

Theorem 1 *Suppose that:*
(i) μ^n satisfies an LDP with constants (c_n) and good rate function ψ;
(ii) whenever $\mathbf{x}_n \in \Delta_n$ and $\mathbf{x}_n \to \mathbf{x} \in \Delta$, $\{P^n_{\mathbf{x}_n}\}$ satisfies an LDP with constants (c_n) and rate $\lambda_{\mathbf{x}}(x)$;
(iii) $\lambda_{\mathbf{x}}(y)$ is lower semicontinuous in $(y, \mathbf{x})$, jointly.
Then P^n satisfies an LDP with constants (c_n) and rate

$$\lambda(y) = \inf\{\lambda_{\mathbf{x}}(y) + \psi(\mathbf{x}) : \mathbf{x} \in \Delta\}.$$

Furthermore, if for each $\mathbf{x}$, $\lambda_{\mathbf{x}}(y)$ is a good rate function then λ is good.

In this application, Δ will be the set of probability distribution over the colours (which, as a compact subset of $\mathbb{R}^k$, is Polish) and $\mathbf{x}$ is in Δ_n exactly when $n\mathbf{x}$ contains only integer values, corresponding to a possible realization of the proportions of colours on the graph with n vertices. Of course, for $\mathbf{x} \in \Delta_n$, $P^n_{\mathbf{x}}$ is derived from the distribution of the edges given that the numbers of the colours are $n\mathbf{x}$. (More formal definitions are given at the start of the next section.) The (normalized) number of edges takes values in the reals and so $\mathcal{Y}$ will be the real line, which is Hausdorff and regular.

Clearly, information on the large deviation behaviour of $\{P^n_{\mathbf{x}_n}\}$ is needed to apply Theorem 1. The independence of the edge placement given the colouring means this is fairly easy to deal with, using a rather restricted form of the Gärtner-Ellis theorem. However, before stating the required form of that theorem, a little convex analysis is introduced first.

Definition 2 *Let $\phi : \mathbf{R} \to \mathbf{R}$ be a finite convex function. Its Fenchel dual (also a finite convex function) is*

$$\widehat{\phi}(y) = \sup_{\theta}(\theta y - \phi(\theta)).$$

Lemma 2 *Suppose the convex functions ϕ_n converge to ϕ, which is necessarily convex, as $n \to \infty$. Then*

$$\widehat{\phi}(y) \le \liminf \widehat{\phi}_n(y).$$

Proof. For any $\epsilon > 0$, there is a finite θ such that

$$\widehat{\phi}(y) - \epsilon \le \theta y - \phi(\theta) = \theta y - \phi_n(\theta) + (\phi_n(\theta) - \phi(\theta)) \le \widehat{\phi}_n(y) + (\phi_n(\theta) - \phi(\theta)).$$

Hence $\widehat{\phi}(y) - \epsilon \le \liminf_n \widehat{\phi}_n(y)$. □

Theorem 3 *Suppose* (S_n) *is a sequence of random variables, and* (c_n) *is a sequence of constants with* $\lim_{n\to\infty} c_n = \infty$*; let* P_n *be the distribution of* S_n/c_n*. Define* $\phi_n(\theta) = c_n^{-1}\log(\mathbf{E}(e^{\theta S_n}))$*. Assume that*

$$\lim_{n\to\infty} \phi_n(\theta) = \phi(\theta)$$

exists pointwise and is finite and differentiable for all θ*. Then* P_n *satisfies the LDP with constants* (c_n) *and rate function* $\widehat{\phi}$*.*

Proof. See [4, Theorem 2.3.6 and Exercise 2.3.20]. □

The general formulation of Gärtner-Ellis does not assume that the limit ϕ is finite and differentiable for all θ. However the conclusion is also weaker.

3 LDPs conditional on the colouring

At the moment the number of colours is finite and then it would be satisfactory to treat the set of probability distributions on them as a compact subset of Euclidean space. However, to allow for the possibility of a richer selection of colours it is worth being more general. Suppose that the set of available colours, Σ, is a Polish space. Let Δ be the set of probability distributions on Σ (equipped with the Lévy metric, which gives the topology of weak convergence of distributions and makes it a Polish space, [4, Theorem D.8, p319]). Whenever $\mathbf{x}$ or $\mathbf{x}_n$ are used they are members of Δ, even when this is not made explicit. We denote weak convergence of a sequence $(\mathbf{x}_n)$ of measures to another such $\mathbf{x}$ by $\mathbf{x}_n \xrightarrow{w} \mathbf{x}$. When there are only a finite number of possible colours $\mathbf{x}_n \xrightarrow{w} \mathbf{x}$ is identical to the convergence of the vectors in $\mathbb{R}^k$.

Let $\mathbf{N}_n$ be the vector-valued random variable which gives the number of vertices of the colours. From now on, for notational convenience, let $a_n = n(n-1)/2$. Let $\Delta_n \subset \Delta$ be the set of possible values for $n^{-1}\mathbf{N}_n$. Let $P^n_{\mathbf{x}}$ be the distribution of $\mathcal{E}(n)/a_n$ when $\mathbf{x} \in \Delta_n$.

The following lemma is a routine exercise using the independence of the edges given the colouring, that is given $\mathbf{N}_n$.

Lemma 4 *The m.g.f. of* $\mathcal{E}(n)$ *conditional on* $\mathbf{N}_n = \mathbf{j}$*, where* $\sum_r j_r = n$*, is*

$$\prod_{r<s}(p_{rs}e^\theta + 1 - p_{rs})^{j_r j_s} \prod_r (p_{rr}e^\theta + 1 - p_{rr})^{j_r(j_r-1)/2}.$$

Definition 3 *Let*

$$\phi_{n^{-1}\mathbf{j},n}(\theta) = \frac{\log\left(\mathbf{E}\left(e^{\theta\mathcal{E}(n)} \mid \mathbf{N}_n = \mathbf{j}\right)\right)}{a_n},$$

$A(\theta)_{ij} = \log(1 - p_{ij} + p_{ij}e^\theta)$ *and* $\phi_{\mathbf{x}}(\theta) = \mathbf{x}^T A(\theta)\mathbf{x}$*, where* $\mathbf{x}$ *is a distribution on the colours.*

Lemma 5 *If* $\mathbf{x}_n \xrightarrow{w} \mathbf{x}$ *then* $\mathbf{x}_n^T A(\theta)\mathbf{x}_n \to \mathbf{x}^T A(\theta)\mathbf{x}$*.*

Proof. For fixed θ, $A(\theta)$ is a bounded continuous function on $\Sigma \times \Sigma$ and the product measure $\mathbf{x}_n \times \mathbf{x}_n$ on $\Sigma \times \Sigma$ converges weakly to $\mathbf{x} \times \mathbf{x}$. □

Lemma 6 *If $\mathbf{j}_n$ is a possible numbers of colours on n vertices and $n^{-1}\mathbf{j}_n \xrightarrow{w} \mathbf{x}$ then*

$$\phi_{n^{-1}\mathbf{j}_n,n}(\theta) \longrightarrow \mathbf{x}^T A(\theta)\mathbf{x} = \phi_{\mathbf{x}}(\theta),$$

and $\phi_{\mathbf{x}}(\theta)$ is an everywhere differentiable convex function of θ with $\phi_{\mathbf{x}}(0) = 0$.

Proof. Let $\mathbf{d}(\theta)_i = \log(1 - p_{ii} + p_{ii}e^\theta)$. We have

$$\begin{aligned}\mathbf{E}\left(e^{\theta\mathcal{E}(n)} \mid \mathbf{N}_n = \mathbf{j}\right) &= \prod_{1\le s<t\le k}(1 - p_{st} + p_{st}e^\theta)^{j_s j_t} \prod_{1\le s\le k}(1 - p_{ss} + p_{ss}e^\theta)^{j_s(j_s-1)/2} \\ &= \exp(\mathbf{j}^T A(\theta)\mathbf{j}/2 - \mathbf{j}^T\mathbf{d}(\theta)/2).\end{aligned}$$

Hence, using Lemma 5,

$$\phi_{n^{-1}\mathbf{j}_n,n}(\theta) = \frac{\mathbf{j}_n^T A(\theta)\mathbf{j}_n - \mathbf{j}_n^T\mathbf{d}(\theta)}{n(n-1)} \longrightarrow \phi_{\mathbf{x}}(\theta).$$

The limit's differentiability is a consequence of $A'_{ij}(t)$ being bounded uniformly in (i,j) in any neighbourhood of θ and dominated convergence; its convexity follows from being the limit of convex functions; $A(0) = 0$ implies that $\phi_{\mathbf{x}}(0) = 0$ for every $\mathbf{x}$. □

Theorem 7 *When $\mathbf{x}_n \in \Delta_n$ and $\mathbf{x}_n \xrightarrow{w} \mathbf{x}$, $P^n_{\mathbf{x}_n}$ obeys an LDP (with constants (a_n)) and rate function $\widehat{\phi}_{\mathbf{x}}(y)$.*

Proof. This follows immediately from the previous lemma and Theorem 3. □

Lemma 8 *$\widehat{\phi}_{\mathbf{x}}(y)$ is lower semicontinuous in $(y, \mathbf{x})$ and infinite for every $y \notin [0,1]$.*

Proof. Suppose $\mathbf{x}_n \xrightarrow{w} \mathbf{x}$ and $y_n \to y$. Then, for any $\epsilon > 0$, there is a finite θ such that

$$\begin{aligned}\widehat{\phi}_{\mathbf{x}}(y) - \epsilon &\le \theta y - \phi_{\mathbf{x}}(\theta) \\ &= \theta y_n - \phi_{\mathbf{x}_n}(\theta) + \theta(y - y_n) - (\phi_{\mathbf{x}}(\theta) - \phi_{\mathbf{x}_n}(\theta)) \\ &\le \widehat{\phi}_{\mathbf{x}_n}(y_n) + \theta(y - y_n) - (\phi_{\mathbf{x}}(\theta) - \phi_{\mathbf{x}_n}(\theta))\end{aligned}$$

and so, since $\phi_{\mathbf{x}_n}(\theta) \to \phi_{\mathbf{x}}(\theta)$ by Lemma 5,

$$\widehat{\phi}_{\mathbf{x}}(y) - \epsilon \le \liminf_n \widehat{\phi}_{\mathbf{x}_n}(y_n)$$

as required.

Using the explicit form of A

$$\lim_{\theta\to-\infty} \phi'_{\mathbf{x}}(\theta) = \lim_{\theta\to-\infty} \mathbf{x}^T A'(\theta)\mathbf{x} \ge 0$$

and so, by calculus, $\widehat{\phi}_{\mathbf{x}}(y) = \infty$ when $y < 0$. Similarly, letting $\theta \to \infty$, $\widehat{\phi}_{\mathbf{x}}(y) = \infty$ when $y > 1$. □

Lemma 9 *$\widehat{\phi}_{\mathbf{x}}(y) = 0$ exactly when $y = \mathbf{x}^T\widetilde{P}\mathbf{x}$.*

Proof. Note first that $\phi'_{\mathbf{x}}(0) = \mathbf{x}^T\widetilde{P}\mathbf{x}$. Since $(\theta y - \phi_{\mathbf{x}}(\theta))$ is concave in θ, considering its derivative at $\theta = 0$ gives the result. □

Theorem 10 *When $\mathbf{x}_n \in \Delta_n$ and $\mathbf{x}_n \xrightarrow{w} \mathbf{x}$, $P^n_{\mathbf{x}_n}$ obeys an LDP (with constants (n)) and rate function which is zero at $\mathbf{x}^T\widetilde{P}\mathbf{x}$ and infinite elsewhere.*

Proof. This follows from Theorem 7 and Lemma 9. □

4 LDPs for the number of edges

The colouring $n^{-1}\mathbf{N}_n$ has distribution μ^n and the distribution of $\mathcal{E}(n)/a_n$ given $\mathbf{N}_n/n = \mathbf{x}$ is $P^n_{\mathbf{x}}$; P^n is the marginal distribution of $\mathcal{E}(n)/a_n$.

Theorem 11

(i) Suppose the colouring μ^n obeys an LDP with constants (a_n) and rate function that is zero throughout Δ. Then P^n satisfies an LDP with constants (a_n) and rate $\Phi(y) = \inf\{\widehat{\phi}_{\mathbf{x}}(y) : \mathbf{x} \in \Delta\}$.

(ii) Suppose the colouring μ^n obeys an LDP, with constants (b_n), and a good rate function, ψ. Suppose also that $(b_n/a_n) \to 0$. Then P^n satisfies an LDP with constants (b_n) and rate $\Psi(y) = \inf\{\psi(\mathbf{x}) : \mathbf{x}^T\widetilde{P}\mathbf{x} = y\}$.

Proof. The first part is an application of Theorems 1 and 7 and Lemma 8. The second is an application of Theorems 1 and 10. □

The next two results show that independent colourings (with k finite) produce μ^n with the right properties for both parts of Theorem 11.

Lemma 12 *For independent colouring with a finite number of possible colours, (μ^n) obeys an LDP with constants (n) and good rate function (on Δ). Furthermore, the rate function is finite throughout Δ.*

Proof. It is a consequence of Sanov's Theorem, [4, Definition 2.1.5 et seq., and Theorem 2.1.10] that when $\mathbf{N}_n$ has a multinomial distribution, the distributions of $\mathbf{n}_n = n^{-1}\mathbf{N}_n$ satisfy an LDP with constants (n) and rate given by

$$\psi(\mathbf{x}) = \begin{cases} \sum_i x_i \log(x_i/s_i) & \text{for } \mathbf{x} \in \Delta \\ \infty & \text{otherwise} \end{cases}, \tag{1}$$

where $(s_1, s_2, \ldots, s_k)$ are the probabilities of the colours. This rate function is continuous, convex and bounded on Δ. □

The first part of this theorem holds for more general sets of colours; [4, §6.2]. However the finiteness of the rate throughout Δ does not hold more generally.

Lemma 13 *Suppose the colouring μ^n obeys an LDP, with constants (b_n), and a rate function that is finite throughout Δ. Suppose also that $(b_n/a_n) \to 0$. Then μ^n obeys an LDP with constants (a_n) and rate function that is zero throughout Δ.*

Proof. The bounds in the LDP with constants (b_n) are finite and so they they converge to zero when multiplied by b_n/a_n. □

Theorem 14 *Suppose vertices are coloured independently from k colours with probabilities $(s_1, s_2, \ldots, s_k)$.*

(i) P^n satisfies an LDP with constants (a_n) and rate $\Phi(y) = \inf\{\widehat{\phi}_{\mathbf{x}}(y) : \mathbf{x} \in \Delta\}$.

(ii) P^n satisfies an LDP with constants (n) and rate $\Psi(y) = \inf\{\psi(\mathbf{x}) : \mathbf{x}^T\widetilde{P}\mathbf{x} = y\}$, with ψ defined at (1).

Proof. Lemmas 12 and 13 together show that μ^n obeys an LDP with constants (a_n) and rate function that is zero throughout Δ. Now Theorem 11(i) gives part (i) here. Lemma 12 and Theorem 11(ii) give part (ii) here.

□

5 Properties of the rate functions

In this section we investigate the properties of Φ and Ψ. The arguments rely heavily on the fact that Δ is compact when the number of possible colours is finite. The properties derived allow similarities and contrasts to be drawn with $G(n, \alpha)$. The first three Lemmas examine the rate function in Theorem 11(i).

Lemma 15 *The infimum in* $\Phi(y) = \inf_{\mathbf{x}} \widehat{\phi}_{\mathbf{x}}(y)$ *is attained.*

Proof. Let $\mathbf{x}_n$ be such that $\widehat{\phi}_{\mathbf{x}_n}(y) \to \Phi(y)$, with (using the compactness of Δ) $\mathbf{x}_n \xrightarrow{w} \mathbf{x}$. Then $\phi_{\mathbf{x}_n} \to \phi_{\mathbf{x}}$ and so, by Lemma 2

$$\Phi(y) = \liminf_n \widehat{\phi}_{\mathbf{x}_n}(y) \geq \widehat{\phi}_{\mathbf{x}}(y) \geq \Phi(y). \quad \square$$

Definition 4 *Define:*

$$m = \inf\{\mathbf{x}^T \widetilde{P} \mathbf{x} : \mathbf{x} \in \Delta\} \quad \textit{and} \quad M = \sup\{\mathbf{x}^T \widetilde{P} \mathbf{x} : \mathbf{x} \in \Delta\};$$

$$l = \inf\{\mathbf{x}^T I(\widetilde{P}_{ij} = 1)\mathbf{x} : \mathbf{x} \in \Delta\} \quad \textit{and} \quad L = \sup\{\mathbf{x}^T I(\widetilde{P}_{ij} > 0)\mathbf{x} : \mathbf{x} \in \Delta\}.$$

These are attained because Δ is compact, and $l \leq m \leq M \leq L$.

Lemma 16
(i) $\Phi(y) = 0$ if and only if $m \leq y \leq M$.
(ii) $\Phi(y) < \infty$ if and only if $l \leq y \leq L$
(iii) Φ is strictly monotonic in $[l, m]$ and $[M, L]$.

Proof. By Lemma 9, $\widehat{\phi}_{\mathbf{x}}(y) = 0$ exactly when $y = \mathbf{x}^T \widetilde{P} \mathbf{x}$. Thus, since $\mathbf{x}^T \widetilde{P} \mathbf{x}$ takes all values in $[m, M]$ as $\mathbf{x}$ varies, $\Phi(y) = 0$ for all $y \in [m, M]$. On the other hand, by Lemma 15, $\widehat{\phi}_{\mathbf{x}}(y) = 0$ for some $\mathbf{x}$ when $\Phi(y) = 0$, and then $y = \mathbf{x}^T \widetilde{P} \mathbf{x} \in [m, M]$. This proves (i).

Similarly, $\phi'_{\mathbf{x}}(\theta) \to \mathbf{x}^T I(\widetilde{P}_{ij} = 1)\mathbf{x}$, as $\theta \to -\infty$ and so $\widehat{\phi}_{\mathbf{x}}(y) = \infty$ when $y < \mathbf{x}^T I(\widetilde{P}_{ij} = 1)\mathbf{x}$. On the other hand, when $y = l$ direct calculation shows that $\widehat{\phi}_{\mathbf{x}}(l) < \infty$ for any $\mathbf{x}$ that provides the infimum in the definition of l. Considering $\theta \to \infty$ shows that $\Phi(y) = \infty$ for $y > L$ and $\Phi(L) < \infty$. This proves the 'only if' part of (ii).

Take $y < m$. By Lemma 15 we know that, for $\epsilon > 0$ there is a suitable $\mathbf{x}_*$ such that

$$\Phi(y - \epsilon) = \widehat{\phi}_{\mathbf{x}_*}(y - \epsilon) = \sup_\theta [\theta(y - \epsilon) - \mathbf{x}_*^T A(\theta)\mathbf{x}_*].$$

Since $y < m$, both $\widehat{\phi}_{\mathbf{x}_*}(y - \epsilon)$ and $\widehat{\phi}_{\mathbf{x}_*}(y)$ are strictly positive, by part (i). Hence there must be some $\delta > 0$ such that

$$\begin{aligned} \sup_\theta [\theta(y - \epsilon) - \mathbf{x}_*^T A(\theta)\mathbf{x}_*] &= \sup_{\theta \leq -\delta} [\theta(y - \epsilon) - \mathbf{x}_*^T A(\theta)\mathbf{x}_*] \\ &\geq \sup_{\theta \leq -\delta} [\theta y - \mathbf{x}_*^T A(\theta)\mathbf{x}_*] + \delta\epsilon \\ &= \widehat{\phi}_{\mathbf{x}_*}(y) + \delta\epsilon \\ &\geq \Phi(y) + \delta\epsilon. \end{aligned}$$

Hence $\Phi(y - \epsilon) \geq \Phi(y) + \delta\epsilon$ which gives strict monotonicity when $\Phi(y)$ is finite. Furthermore taking $y - \epsilon = l$ and using $\Phi(l) < \infty$ shows that $\Phi(y)$ is finite for $l < y < m$. The range $M < y < L$ is handled similarly. This completes the proof of the rest of (ii) and of (iii). $\square$

The final lemma concerns the rate in Theorem 11(ii).

Lemma 17 *Suppose ψ is convex, finite on Δ and takes the value zero at a single* $\mathbf{s}$. *Then Ψ is strictly monotone on $[m, \alpha]$ and on $[\alpha, M]$. It is infinite outside $[m, M]$.*

Proof. Note first that $\Psi(\alpha) = \psi(\mathbf{s}) = \inf_{\mathbf{x}} \psi(\mathbf{x}) = 0$ and this infimum is attained only at $\mathbf{s}$. Take $y \in [m, \alpha)$ and $\tilde{\mathbf{x}}$ such that $\Psi(y) = \psi(\tilde{\mathbf{x}})$ and $\tilde{\mathbf{x}}^T \widetilde{P} \tilde{\mathbf{x}} = y$. Since ψ is always finite, $\Psi(y) < \infty$. Take $z \in (y, \alpha)$. Then, for suitable $\delta > 0$, $(1-\delta)\tilde{\mathbf{x}} + \delta \mathbf{s} \in \{\mathbf{x} : \mathbf{x}^T \widetilde{P} \mathbf{x} = z\}$ and, by convexity,

$$\psi((1-\delta)\tilde{\mathbf{x}} + \delta \mathbf{s}) \leq (1-\delta)\psi(\tilde{\mathbf{x}}) + \delta\psi(\mathbf{s}) = (1-\delta)\psi(\tilde{\mathbf{x}}) < \psi(\tilde{\mathbf{x}})$$

Hence $\Psi(z) < \Psi(y)$ as required. The range from α to M is similar. □

Without the assumption that ψ is finite on Δ the proof still works to show that Ψ is monotone when finite either side of its minimum. However, it may be infinite for some values within $[m, M]$ which would mean that the two parts of Theorem 11 would leave a range of values (where Ψ is infinite and Φ is zero) where an LDP with constants intermediate between (b_n) and (a_n) might be appropriate.

In $G(n, \alpha)$, $m = \alpha = M$, and so the 'inner' large deviation regime (with constants (n) and rate Ψ) is degenerate.

6 Extensions

Two directions for extension are indicated. The first has already been mentioned; it is allowing a larger set of colours. Specifically, the colours are drawn from a Polish space, Σ. The second is the possibility of associating a random variable with each edge that is more general than the indicator variables for the presence of that edge.

Let $M(\sigma_1, \sigma_2; \theta)$ be the moment generating function of a random variable associated with an edge with vertices of colours σ_1 and σ_2, with mean $m(\sigma_1, \sigma_2)$. Assume, for each θ, that $M(\sigma_1, \sigma_2; \theta)$ is a bounded continuous function on $\Sigma \times \Sigma$, with a derivative that is bounded uniformly in $\Sigma \times \Sigma$ on a neighbourhood of any θ. For $\mathbf{x} \in \Delta$ let s_1 and s_2 be independent colours selected using $\mathbf{x}$ and let $\phi_{\mathbf{x}}(\theta) = EM(s_1, s_2; \theta)$ and $m_{\mathbf{x}} = Em(s_1, s_2)$. Finally. let $\mathcal{E}$ be the sum of the variables over the edges and P^n be the distribution of $\mathcal{E}/a_n$. The next result is obtained by working through the details of the arguments leading to Theorem 11 and checking that nothing has changed. We omit the proof.

Theorem 18

(i) Suppose the colouring μ^n obeys an LDP with constants (a_n) and rate function that is zero throughout Δ. Then P^n satisfies an LDP with constants (a_n) and rate $\Phi(y) = \inf\{\widehat{\phi_{\mathbf{x}}}(y) : \mathbf{x} \in \Delta\}$.

(ii) Suppose the colouring μ^n obeys an LDP, with constants (b_n), and good rate function, ψ, and $(b_n/a_n) \to 0$. Then P^n satisfies an LDP with constants (b_n) and rate $\Psi(y) = \inf\{\psi(\mathbf{x}) : m_{\mathbf{x}} = y\}$.

For independent colourings when there are infinitely many possible colours, the rate function is no longer finite throughout Δ and so any analogue of Theorem 14(i) needs more argument. The basic issue is whether for independent colourings μ^n obeys an LDP with constants n^2, and, if it does, whether the rate function is degenerate. Let $\Delta_{\mathbf{s}}$ be those $\mathbf{x} \in \Delta$ that are absolutely continuous with respect to $\mathbf{s}$. An attractive conjecture, given the general form of the rate function in Sanov's theorem, is that the the appropriate rate function for the analogue of Theorem 14(i) is

$$\Phi(y) = \inf\{\widehat{\phi}_{\mathbf{x}}(y) : \mathbf{x} \in \Delta_{\mathbf{s}}\}.$$

However, we have not made any serious attempt to examine this. In contrast, Theorem 14(ii) generalises in the obvious way.

References

[1] Bollobás, B. *Random graphs.* Academic Press 1985.

[2] Biggins, J.D. Large deviations for mixtures. (2003) Submitted to *Probability Theory and Related Fields.*
preprint available at `http://www.shef.ac.uk/~st1jdb/ldfm.html`

[3] Cannings, C., Penman, D. B. Models of random graphs and their applications *Handbook of Statistics 21. Stochastic Processes: Modeling and Simulation.* Eds: D.N. Shanbhag, and C.R. Rao. Elsevier (2003) 51–91.

[4] Dembo, A., Zeitouni, O. *Large deviation techniques*, Jones and Bartlett, Boston (1993).

[5] Dinwoodie, I. H., Zabell, S. L. Large deviations for exchangeable random vectors. Ann. Prob. **20** (1992) 1147–1166.

[6] Penman, D. B. Random graphs with correlation structure. Ph.D thesis, University of Sheffield, 1998.

J.D. Biggins
Dept of Probability and Statistics,
Hicks Building,
The University of Sheffield,
Sheffield, S3 7RH,
U.K.

D.B. Penman
Dept of Mathematics,
University of Essex,
Wivenhoe Park,
Colchester C04 3SQ,
U.K.