A LARGE DEVIATION PRINCIPLE FOR THE ERDŐS–RÉNYI UNIFORM RANDOM GRAPH

AMIR DEMBO AND EYAL LUBETZKY

ABSTRACT. Starting with the large deviation principle (LDP) for the Erdős–Rényi binomial random graph $\mathcal{G}(n,p)$ (edge indicators are i.i.d.), due to Chatterjee and Varadhan (2011), we derive the LDP for the uniform random graph $\mathcal{G}(n,m)$ (the uniform distribution over graphs with n vertices and m edges), at suitable $m = m_n$. Applying the latter LDP we find that tail decays for subgraph counts in $\mathcal{G}(n,m_n)$ are controlled by variational problems, which up to a constant shift, coincide with those studied by Kenyon et al. and Radin et al. in the context of constrained random graphs, e.g., the edge/triangle model.

1. INTRODUCTION

The Erdős–Rényi *binomial* random graph model $\mathcal{G}(n,p)$ is the graph on n vertices where each edge is present independently with probability p; the *uniform* random graph $\mathcal{G}(n,m)$ is the uniform distribution over graphs with n vertices and exactly m edges. Chatterjee and Varadhan [5] established a large deviation principle (LDP) for $\mathcal{G}(n,p)$ for p fixed, with the notable application of estimating the probabilities that the number of copies t_H of a fixed subgraph H is abnormally far from its mean. Even for variables in $\mathcal{G}(n,m)$ whose typical behavior mirrors that of their analogs in $\mathcal{G}(n,p)$ for $p = m/\binom{n}{2}$, rare events in the two models may be triggered by different phenomena: for instance, while t_H has asymptotically the same *mean* in $\mathcal{G}(n,m)$ and $\mathcal{G}(n,p = m/\binom{n}{2})$, its variance can in fact feature a different exponent of n (cf. [10, Example 6.55, p. 174]).

Our goal is to build on the work of [5] and show that the variational problem to which it reduced the LDP in $\mathcal{G}(n,p)$—an optimization problem over *graphons* (defined next)—captures the analogous problem in $\mathcal{G}(n,m)$ for $p = m/\binom{n}{2}$ once we further restrict the graphons to have edge density p (see (1.5)). As we later show, the resulting variational problem has already been studied in the context of constrained random graphs.

Let $\mathcal{W}$ be the space of all bounded measurable functions $f : [0,1]^2 \to \mathbb{R}$ that are symmetric ($f(x,y) = f(y,x)$ for all $x,y \in [0,1]$). Let $\mathcal{W}_0 \subset \mathcal{W}$ denote all *graphons*, that is, symmetric measurable functions $[0,1]^2 \to [0,1]$ (these generalize finite graphs; see (1.2)). The cut-norm of $W \in \mathcal{W}$ is given by

$$\|W\|_\square := \sup_{S,T\subset[0,1]} \left| \int_{S\times T} W(x,y)\ dxdy \right| = \sup_{u,v\colon [0,1]\to[0,1]} \left| \int_{[0,1]^2} W(x,y)u(x)v(y)\ dxdy \right| ,$$

(by linearity of the integral it suffices to consider $\{0,1\}$-valued u,v, hence the equality). For any measure-preserving map $\sigma\colon [0,1] \to [0,1]$ and $W \in \mathcal{W}$, let $W^\sigma \in \mathcal{W}$ denote the graphon $W^\sigma(x,y) = W(\sigma(x),\sigma(y))$. The cut-distance on $\mathcal{W}$ is then defined as

$$\delta_\square(W_1,W_2) := \inf_\sigma \|W_1 - W_2^\sigma\|_\square\ ,$$

with the infimum taken over all measure-preserving bijections σ on $[0,1]$. It yields the pseudo-metric space $(\mathcal{W}_0,\delta_\square)$, which is elevated into a genuine metric space $(\widetilde{\mathcal{W}}_0,\delta_\square)$

2010 *Mathematics Subject Classification.* 05C80, 60F10.

Key words and phrases. Large deviations, Erdős–Rényi graphs, constrained random graphs.

upon taking the quotient w.r.t. the equivalence relation $W_1 \sim W_2$ iff $\delta_\square(W_1, W_2) = 0$. In what may be viewed as a topological version of Szemerédi's regularity lemma, Lovász and Szegedy [15] showed that the metric space $(\widetilde{\mathcal{W}}_0, \delta_\square)$ is compact. For a finite simple graph $H = (V(H), E(H))$ with $V(H) = \{1, \ldots, k\}$, its subgraph density in $W \in \mathcal{W}_0$ is

$$t_H(W) := \int_{[0,1]^k} \prod_{(i,j)\in E(H)} W(x_i, x_j)\; dx_1 \cdots dx_k\,,$$

with the map $W \mapsto t_H(W)$ being Lipschitz-continuous in $(\widetilde{\mathcal{W}}_0, \delta_\square)$ (see [3, Thm 3.7]).

Define $I_p \colon [0,1] \to \mathbb{R}$ by

$$I_p(x) := \frac{x}{2}\log\frac{x}{p} + \frac{1-x}{2}\log\frac{1-x}{1-p} \qquad \text{for } p \in (0,1) \text{ and } x \in [0,1]\,, \tag{1.1}$$

and extend I_p to $\mathcal{W}_0$ via $I_p(W) := \int_{[0,1]^2} I_p(W(x,y))\; dxdy$ for $W \in \mathcal{W}_0$. As I_p is convex on $[0,1]$, it is lower-semicontinuous (LSC) on $\widetilde{\mathcal{W}}_0$ w.r.t. the cut-metric topology ([5, Lem. 2.1]).

In the context of the space of graphons $\widetilde{\mathcal{W}}_0$, a simple graph G with vertices $\{1, \ldots, n\}$ can be represented by

$$W_G(x,y) = \begin{cases} 1 & \text{if } (\lceil nx\rceil, \lceil ny\rceil) \text{ is an edge of } G, \\ 0 & \text{otherwise.} \end{cases} \tag{1.2}$$

For two graphs G and H let $\hom(H, G)$ count the number of homomorphisms from H to G (i.e., maps $V(H) \to V(G)$ that carry edges to edges). Let

$$t_H(G) := |V(G)|^{-|V(H)|} |\hom(H,G)| = t_H(W_G)\,.$$

A sequence of graphs $\{G_n\}_{n\geq 1}$ is said to converge if the sequence of subgraph densities $t_H(G_n)$ converges for every fixed finite simple graph H. It was shown in [15] that for any such convergent graph sequence there is a limit object $W \in \widetilde{\mathcal{W}}_0$ such that $t_H(G_n) \to t_H(W)$ for every fixed H. Conversely, any $W \in \widetilde{\mathcal{W}}_0$ arises as a limit of a convergent graph sequence. It was shown in [3] that a sequence of graphs $\{G_n\}_{n\geq 1}$ converges if and only if the sequence of graphons $W_{G_n} \in \mathcal{W}_0$ converges in $\mathcal{W}_0$ w.r.t. $\delta_\square$

A random graph $G_n \sim \mathcal{G}(n,p)$ corresponds to a random point $W_{G_n} \in \widetilde{\mathcal{W}}_0$—inducing a probability distribution $\mathbb{P}(G_n \in \cdot)$ on $\widetilde{\mathcal{W}}_0$ supported on a finite set of points (n-vertex graphs)—and $G_n \to W$ for the constant graphon $W \equiv p$ a.s. for every fixed $0 < p < 1$. Chatterjee and Varadhan [5] showed that, for $0 < p < 1$ fixed, the random graph $\mathcal{G}(n,p)$ obeys an LDP in $(\widetilde{\mathcal{W}}_0, \delta_\square)$ with the rate function $I_p(\cdot)$. Denoting $\|W\|_1 = \int |W(x,y)|\, dxdy$, and considering the restricted spaces

$$\mathcal{W}_0^{(p)} := \left\{W \in \mathcal{W}_0 \;:\; \|W\|_1 = p\right\} \quad \text{and} \quad \widetilde{\mathcal{W}}_0^{(p)} = \left\{W \in \widetilde{\mathcal{W}}_0 \;:\; \|W\|_1 = p\right\},$$

here we deduce the analogous statement for the random graph $\mathcal{G}(n,m)$, the uniform distribution over all graphs with n vertices and exactly m edges, with a rate function $J_p(\cdot)$ restricted to $\widetilde{\mathcal{W}}_0^{(p)}$. As we later conclude, the variational formulas of this LDP for $\mathcal{G}(n,m)$, addressing such random graph structure conditioned on a large deviation, coincide with those studied earlier by Kenyon et al. and Radin et al. (cf. [11–13]).

Theorem 1.1. *Fix $0<p<1$ and let $m_n \in \mathbb{N}$ be such that $m_n/\binom{n}{2} \to p$ as $n\to\infty$. Let $G_n \sim \mathcal{G}(n,m_n)$. Then the sequence $\mathbb{P}(G_n \in \cdot)$ obeys the* LDP *in the space $(\widetilde{\mathcal{W}}_0, \delta_\square)$ with the good rate function J_p (namely, with $\{W : J_p(W)\le\alpha\}$ compact for α finite), where $J_p(W)=I_p(W)$ if $W\in\widetilde{\mathcal{W}}_0^{(p)}$ and is ∞ otherwise; that is, for any closed set $F\subseteq\widetilde{\mathcal{W}}_0$,*

$$\limsup_{n\to\infty} n^{-2}\log\mathbb{P}(G_n\in F) \le -\inf_{W\in F} J_p(W)\,,$$

and for any open $U\subseteq\widetilde{\mathcal{W}}_0$,

$$\liminf_{n\to\infty} n^{-2}\log\mathbb{P}(G_n\in U) \ge -\inf_{W\in U} J_p(W)\,.$$

Define

$$\phi_H(p,r) := \inf\Big\{ I_p(W)\colon W\in\widetilde{\mathcal{W}}_0\ ,\ t_H(W)\ge r\Big\} \tag{1.3}$$

and further let

$$\psi_H(p,r) := \inf\Big\{ I_p(W)\colon W\in\widetilde{\mathcal{W}}_0^{(p)}\ ,\ t_H(W)\ge r\Big\} \tag{1.4}$$

(with I_p having compact level sets in $(\widetilde{\mathcal{W}}_0,\delta_\square)$ and $t_H(\cdot)$ continuous on $(\widetilde{\mathcal{W}}_0,\delta_\square)$, the infimums in (1.3),(1.4) are attained whenever the relevant set of graphons is nonempty). For any $r\ge t_H(p)$ we relate the equivalent form of (1.4) (see Corollary 1.2), given by

$$\psi_H(p,r) = \inf\Big\{ I_p(W)\colon W\in\widetilde{\mathcal{W}}_0^{(p)}\ ,\ t_H(W)= r\Big\}\ , \tag{1.5}$$

to the following variational problem that has been extensively studied (e.g., [11–13,18]) in constrained random graphs such as the edge/triangle model (where H is a triangle):

$$F_H(p,r) := \sup\Big\{ h_e(W)\colon W\in\widetilde{\mathcal{W}}_0^{(p)}\ ,\ t_H(W)= r\Big\}\ , \tag{1.6}$$

where $h_e(x)=-\frac{1}{2}(x\log x+(1-x)\log(1-x))$ is the (natural base) entropy function. As $I_p(x)=-h_e(x)-\frac{x}{2}\log p-\frac{1-x}{2}\log(1-p)$ and $\|W\|_1=p$ throughout $\widetilde{\mathcal{W}}_0^{(p)}$, we see that both variational problems for F_H and $-\psi_H$ have the same set of optimizers, and

$$F_H(p,r) = -\psi_H(p,r)+h_e(p)\,.$$

As a main application of their LDP, Chatterjee and Varadhan [5] showed that the large deviation rate function for subgraph counts in $\mathcal{G}(n,p)$ for any fixed $0<p<1$ and graph H reduces to the variational problem (1.3). Namely, if $G_n\sim\mathcal{G}(n,p)$ then

$$\lim_{n\to\infty} n^{-2}\log\mathbb{P}\left(t_H(G_n)\ge r\right) = -\phi_H(p,r) \qquad \text{for every fixed } p,r\in(0,1) \text{ and } H\,,$$

and, on the event $\{t_H(G_n)\ge r\}$, the graph G_n is typically close to a minimizer of (1.3). Theorem 1.1 implies the analogous statement for the random graph $\mathcal{G}(n,m_n)$ w.r.t. the variational problem (1.4) (similar statements hold for lower tails of subgraph counts both in case of $\mathcal{G}(n,p)$ and that of $\mathcal{G}(n,m_n)$).

Corollary 1.2. *Fixing a subgraph H and $0 < p < 1$, let $r_H \in (t_H(p), 1]$ denote the largest r for which the collection of graphons in (1.4) is nonempty.*

(a) The LSC *function $r \mapsto \psi_H(p,r)$ is zero on $[0, t_H(p)]$ and finite, strictly increasing on $[t_H(p), r_H]$. The nonempty set $F_\star$ of minimizers of (1.4) is a single point $W_\star \equiv p$ for $r \le t_H(p)$ and $F_\star$ coincides for any $r \in [t_H(p), r_H]$ with the minimizers of (1.5).*

(b) For any $m_n \in \mathbb{N}$ such that $m_n/\binom{n}{2} \to p$ as $n \to \infty$ and any right-continuity point $r \in [0, r_H)$ of $t \mapsto \psi_H(p,t)$, the random graph $G_n \sim \mathcal{G}(n, m_n)$ satisfies

$$\lim_{n\to\infty} n^{-2} \log \mathbb{P}\left(t_H(G_n) \ge r\right) = -\psi_H(p,r)\,. \tag{1.7}$$

(c) For any (p,r) as in part (b), and every $\varepsilon > 0$ there is $C = C(H, \varepsilon, p, r) > 0$ so that for all n large enough

$$\mathbb{P}\left(\delta_\square(G_n, F_\star) \ge \varepsilon \;\middle|\; t_H(G_n) \ge r\right) \le e^{-Cn^2}\,. \tag{1.8}$$

Remark 1.3. Since the function $r \mapsto \psi_H(p,r)$ is monotone, it is continuous a.e.; however, the identity (1.7) may fail when $\psi_H(p,\cdot)$ is discontinuous at r. For example, at $r = r_H$ the LHS of (1.7) equals $-\infty$ whenever $m_n/\binom{n}{2} \uparrow p$ slowly enough.

Remark 1.4. The analog of (1.7) in the sparse regime (with edge density $p_n = o(1)$) has been established in [4] in terms of a discrete variational problem in lieu of (1.3), valid when $n^{-c_H} \ll p_n \ll 1$ for some $c_H > 0$ (see also [8], improving the range of p_n, and [2,16,17,19] for analyses of these variational problems in the sparse/dense regimes); In contrast with the delicate regime $p_n = n^{-c}$, such results in the range $p_n \gg (\log n)^{-c}$ of $\mathcal{G}(n,p)$ are a straightforward consequence of the weak regularity lemma (cf. [17, §5]), and further extend to $\mathcal{G}(n, m_n)$, where the discrete variational problem features an extra constraint on the number of edges (see Proposition 3.3).

Consider (p,r) in the setting of Corollary 1.2. The studies of the variational problem for F_H given in (1.6) were motivated by the question of estimating the number of graphs with prescribed edge and H-densities, via the following relation:

$$F_H(p,r) = \lim_{\delta\downarrow 0} \lim_{n\to\infty} \frac{1}{n^2} \log |\mathscr{H}^\delta_{n,p,r}| \text{ where } \mathscr{H}^\delta_{n,p,r} = \left\{ G_n \;:\; \begin{array}{l} \big||E(G_n)|/\binom{n}{2} - p\big| \le \delta\,, \\ |t_H(G_n) - r| \le \delta \end{array} \right\}.$$

(This follows by general principles from the LDP of [5] for $\mathcal{G}(n,p)$; see Proposition 2.1(a), or [18, Thm 3.1] for the derivation in the special case of the edge/triangle model). Corollary 1.2 allows us, roughly speaking, to interchange the order of these two limits; for instance, for any right-continuity point $r \ge t_H(p)$ of $t \mapsto \psi_H(p,t)$ (which holds a.e.), the same variational problem in (1.6) also satisfies

$$F_H(p,r) = \lim_{n\to\infty} \frac{1}{n^2} \log |\mathscr{H}_{n,m_n,r}| \text{ where } \mathscr{H}_{n,m,r} = \left\{ G_n \;:\; \begin{array}{l} |E(G_n)| = m\,, \\ t_H(G_n) \ge r \end{array} \right\}. \tag{1.9}$$

(Indeed, $-\psi_H(p,r) = \lim_{n\to\infty} n^{-2} \log \mathbb{P}\left(t_H(\mathcal{G}(n,m_n)) \ge r\right)$, and this log-probability is then translated to $\log |\mathscr{H}_{n,m_n,r}|$ by adding $n^{-2} \log \binom{\binom{n}{2}}{m_n} \to h_e(p) = F_H(p,r) + \psi_H(p,r)$.)

For the various results (as well as numerical simulations for the many problems related to (1.6) that remain open), the reader is referred to [11–13] and the references therein. We also note a potential connection with certain exponential random graph models $\mathbb{P}_{\theta_n}(\cdot)$, considered by [7, Definition 2.1] as equivalent to the uniformly chosen graph in $\mathscr{H}_{n,m_n,r}$ whenever $n^{-2}\log \mathbb{P}_{\theta_n}(\mathscr{H}_{n,m_n,r}) \to -F_H(p,r)$.

Recall that the law of $\mathcal{G}(n,m_n)$ can be represented as that of a random graph G_n from the model $\mathcal{G}(n,p)$, conditional on $|E(G_n)| = m_n$. While our choice of m_n in Theorem 1.1 is rather typical for $\mathcal{G}(n,p)$ when $n \gg 1$, any LDP and in particular the LDP of [5], deals only with open and closed sets. The challenge in deriving Theorem 1.1 is thus in handling the point conditioning. To this end, we provide in Section 2 a general result (Proposition 2.1) for deriving a conditional LDP, which we then combine in §3.1 with a combinatorial coupling, and thereby prove Theorem 1.1. Building on the latter, §3.2 provides the proof of Corollary 1.2, whereas §3.3 is devoted to the analog of (1.7) for $\mathcal{G}(n,m_n)$ in the range $m_n \gg n^2(\log n)^{-c_H}$ (see Proposition 3.3).

We note that, equipped with the coupling of §3.1, one can prove Corollary 1.2 directly, bypassing much of the general terminology of §2. However, such a direct proof seems neither different, nor shorter than the general approach we take here, with the payoff for mastering the relevant terminology of Proposition 2.1, being its applicability to general point conditioning in large deviations, way beyond random graphs.

2. Conditional LDP

The LDP for $\mathcal{G}(n,m)$ is obtained by the next general relation between a given LDP for measures $\{\mu_n\}$ and the LDP for laws ν_n induced by point-conditioning μ_n.

Proposition 2.1. *Suppose Borel probability measures $\{\mu_n\}$ on a metric space $(\mathcal{X},d)$ satisfy the* LDP *with rate $a_n \to 0$ and good rate function $I(\cdot)$. Fix a metric space $(\mathbb{S},\rho)$, a continuous map $f:(\mathcal{X},d)\to(\mathbb{S},\rho)$ and $s\in\mathbb{S}$. For every $\eta>0$, let Z_n^η denote random variables of the law*

$$\nu_n^\eta := \mu_n\left(\cdot \mid \mathcal{B}^o_{f,s,\eta}\right) , \tag{2.1}$$

where

$$\mathcal{B}_{f,s,\eta} := \{x\in\mathcal{X} : \rho(s,f(x)) \le \eta\} , \qquad \mathcal{B}^o_{f,s,\eta} := \{x\in\mathcal{X} : \rho(s,f(x)) < \eta\} .$$

(a) If

$$\lim_{n\to\infty} a_n \log \mu_n(\mathcal{B}^o_{f,s,\eta}) = 0 \qquad \text{for every } \eta>0 \text{ fixed}, \tag{2.2}$$

then for the good rate function

$$J_0(x) := \begin{cases} I(x), & f(x)=s \\ \infty, & \text{otherwise} \end{cases}$$

and any open $U \subset \mathcal{X}$ and closed $F \subset \mathcal{X}$,

$$\liminf_{\eta\to 0}\liminf_{n\to\infty} a_n \log \nu_n^\eta(U) \ge -\inf_{x\in U} J_0(x)\,, \tag{2.3}$$

$$\limsup_{\eta\to 0}\limsup_{n\to\infty} a_n \log \nu_n^\eta(F) \le -\inf_{x\in F} J_0(x)\,. \tag{2.4}$$

(b) *Suppose (2.2) holds and that $\{Z_n^\eta\}$ form an exponentially good approximation of variables $Z_n \sim \nu_n$, as in [6, Definition 4.2.14]; i.e., for any $\delta > 0$, there exist couplings $\mathbb{P}_{n,\eta}$ of (Z_n, Z_n^η) so that*

$$\lim_{\eta\downarrow 0}\limsup_{n\to\infty} a_n \log \mathbb{P}_{n,\eta}(d(Z_n, Z_n^\eta) > \delta) = -\infty\,. \tag{2.5}$$

Then $\{\nu_n\}$ satisfy the LDP *with rate $a_n \to 0$ and the good rate function $J_0(\cdot)$.*

Proof. We first deduce from (2.2) that for every $\eta > 0$, open $U \subset \mathcal{X}$ and closed $F \subset \mathcal{X}$,

$$\liminf_{n\to\infty} a_n \log \nu_n^\eta(U) \ge -\inf_{x\in U} J_\eta^o(x)\,, \tag{2.6}$$

$$\limsup_{n\to\infty} a_n \log \nu_n^\eta(F) \le -\inf_{x\in F} J_\eta(x)\,, \tag{2.7}$$

where

$$J_\eta(x) := \begin{cases} I(x), & x \in \mathcal{B}_{f,s,\eta} \\ \infty, & \text{otherwise} \end{cases}\,, \qquad J_\eta^o(x) := \begin{cases} I(x), & x \in \mathcal{B}_{f,s,\eta}^o \\ \infty, & \text{otherwise.} \end{cases}$$

Indeed, for any Borel set A and $\eta > 0$,

$$\mu_n(A \cap \mathcal{B}_{f,s,\eta}^o) \le \nu_n^\eta(A) \le \frac{\mu_n(A \cap \mathcal{B}_{f,s,\eta})}{\mu_n(\mathcal{B}_{f,s,\eta}^o)}\,.$$

Hence, for any open set U, we deduce from the LDP for $\{\mu_n\}$ that

$$\liminf_{n\to\infty} a_n \log \nu_n^\eta(U) \ge \liminf_{n\to\infty} a_n \log \mu_n(U \cap \mathcal{B}_{f,s,\eta}^o) \ge -\inf_{x\in U\cap\mathcal{B}_{f,s,\eta}^o} I(x) = -\inf_{x\in U} J_\eta^o(x)\,.$$

Similarly, for any closed set F it follows from (2.2) that

$$\limsup_{n\to\infty} a_n \log \nu_n^\eta(F) \le \limsup_{n\to\infty} a_n \log \mu_n(F \cap \mathcal{B}_{f,s,\eta}) \le -\inf_{x\in F\cap\mathcal{B}_{f,s,\eta}} I(x) = -\inf_{x\in F} J_\eta(x)\,.$$

(a) In the lower bound (2.6) one obviously can use $J_0(\cdot) \ge J_\eta^o(\cdot)$, yielding (2.3). Moreover, we get the bound (2.4) out of (2.7), upon showing that for any closed $F \subseteq \mathcal{X}$,

$$\inf_{y\in F}\{J_0(y)\} \le \liminf_{\eta\downarrow 0}\inf_{y\in F}\{J_\eta(y)\} := \alpha\,. \tag{2.8}$$

To this end, it suffices to consider only $\alpha < \infty$, in which case $J_{\eta_\ell}(y_\ell) \le \alpha + \ell^{-1}$ for some $\eta_\ell \downarrow 0$ and $y_\ell \in F$. As $\{y_\ell\}$ is contained in the compact level set $\{x : I(x) \le \alpha + 1\}$, it has a limit point $y_\star \in F$. Since $J_{\eta_\ell}(y_\ell) = I(y_\ell) \to \alpha$ it follows from the LSC of $x \mapsto I(x)$ that $I(y_\star) \le \alpha$. Passing to the convergent sub-sequence $\rho(f(y_\ell), f(y_\star)) \to 0$. Further, recall that $\rho(s, f(y_\ell)) \le \eta_\ell \downarrow 0$, hence by the triangle inequality $\rho(s, f(y_\star)) = 0$. Consequently, $J_0(y_\star) = I(y_\star) \le \alpha$ yielding (2.8) and completing the proof of part (a).

(b) Clearly, J_η is a good rate function (namely, of compact level sets $\{x : J_\eta(x) \le \alpha\} = \{x : I(x) \le \alpha\} \cap \mathcal{B}_{f,s,\eta}$), and $J_\eta \le J^o_\eta \uparrow J_0$. If $J^o_\eta \equiv J_\eta$ then (2.7)–(2.6) form the LDP for $\{\nu^\eta_n\}$ with the good rate function J_η. While in general this may not be the case, assuming hereafter that (2.5) holds and proceeding as in [6, (4.2.20)], we get from (2.6) that $\{\nu_n\}$ satisfies the LDP lower bound with the rate function

$$\underline{J}(y) := \sup_{\delta>0} \liminf_{\eta\downarrow 0} \inf_{z\in B_{y,\delta}} \{J^o_\eta(z)\}\,,$$

where $B_{y,\delta} = \{z \in \mathcal{X} : d(y,z) < \delta\}$ (see [6, (4.2.17)], noting that no LDP upper bound for ν^η_n is needed here). Since $y \in B_{y,\delta}$ for any $\delta > 0$, we have that

$$J_0(y) = \lim_{\eta\downarrow 0} J^o_\eta(y) \ge \underline{J}(y)$$

and consequently $\{\nu_n\}$ trivially satisfies the LDP lower bound also with respect to the good rate function J_0. Now, precisely as in the proof of [6, Theorem 4.2.16(b)], we get from (2.5) and (2.7) that the corresponding LDP upper bound holds for $\{\nu_n\}$, thanks to (2.8) (see [6, (4.2.18)]), thereby completing the proof of part (b) of Prop. 2.1. ■

3. LDP for the uniform random graph

3.1. **Proof of Theorem 1.1.** Let μ_n be the law of $\mathcal{G}(n,p)$, which obeys the LDP with good rate function $I_p(\cdot)$ on $(\widetilde{\mathcal{W}}_0, \delta_\square)$ and speed n^2, and let ν_n denote the law of $\mathcal{G}(n,m_n)$. We shall apply Proposition 2.1(b) for $\mathbb{S} = \mathbb{R}$ and $s = p$, with f denoting the L^1-norm on graphons (edge density):

$$f(W) := \|W\|_1 = \iint W(x,y)\, dxdy\,.$$

With these choices, the role of Z_n will be assumed by $G_n \sim \mathcal{G}(n,m_n)$, whereas those of the random variables Z^η_n will be assumed by the binomial random graph $\mathcal{G}(n,p)$ conditioned on having between $\frac{1}{2}(p-\eta)n^2$ and $\frac{1}{2}(p+\eta)n^2$ edges:

$$G^\eta_n \sim \big(\mathcal{G}(n,p) \bigm| B^o_{p,\eta}\big)\ , \quad \text{where} \quad B^o_{p,\eta} = \Big\{G : \tfrac{2|E(G)|}{n^2} \in (p-\eta, p+\eta)\Big\} \tag{3.1}$$

Note that $p_n := 2m_n/n^2 \in (p-\eta,p+\eta)$ for all $n \ge n_0(\eta)$. We couple (G_n, G^η_n) so that for such n, deterministically,

$$|E(G_n) \triangle E(G^\eta_n)| < \eta n^2 \tag{3.2}$$

(here $S \triangle T$ denotes symmetric difference). This is achieved by the following procedure:

(i) Draw $G_n \sim \mathcal{G}(n,m_n)$.

(ii) Independently of G_n draw $E_n \sim \mathrm{Bin}(\binom{n}{2}, p)$ and $M_n \sim \big(E_n \mid |2E_n/n^2 - p| < \eta\big)$. Let $D_n = M_n - m_n$ and obtain G^η_n from G_n as follows:
- [shortage] if $D_n \ge 0$: add a uniformly chosen subset of D_n edges missing from G_n.
- [surplus] if $D_n \le 0$: delete a uniformly chosen subset of D_n edges from G_n.

Since $|D_n| < \eta n^2$ this guarantees (3.2) and has $G_n \sim \nu_n$; the additional fact that $G_n^\eta \sim \nu_n^\eta$ is seen by noting that, if $G \sim \mathcal{G}(n,p)$ then $|E(G)| \sim \mathrm{Bin}(\binom{n}{2}, p)$, and on the event that G has M edges, these are uniformly distributed (i.e., the conditional distribution is $\mathcal{G}(n,M)$).

We proceed to show that such $\{G_n^\eta\}$ form an exponentially good approximation of G_n. Indeed, note that if graphs G, G' on n vertices satisfy $|E(G) \triangle E(G')| < \eta n^2$ then $\delta_\square(G,G') < \eta$ (recall (1.2)). In particular, from (3.2) we deduce that for any $\eta \le \delta$ and all $n \ge n_0(\eta)$,

$$\mathbb{P}\left(\delta_\square(G_n, G_n^\eta) > \delta\right) = 0$$

holds under the above coupling of (G_n, G_n^η), thereby implying (2.5).

Finally, Noting that $B_{p,\eta}^o$ of (3.1) is the event $|2E_n/n^2 - p| < \eta$ (with $E_n \sim \mathrm{Bin}(\binom{n}{2}, p)$ under μ_n), we deduce from the LLN that $\mu_n(B_{p,\eta}^o) \to 1$. In particular, for any $\eta > 0$ one has that $n^{-2} \log \mu_n(B_{p,\eta}^o) \to 0$, thereby verifying (2.2) for the case at hand. ∎

3.2. **Proof of Corollary 1.2.** (a) Recalling that $J_p(W) = I_p(W)$ on $\widetilde{\mathcal{W}}_0^{(p)}$ and otherwise $J_p(W) = \infty$, we express (1.4) as

$$\psi_H(p,r) = \inf_{W \in \Gamma_{\ge r}} \{J_p(W)\},$$

for the closed set of graphons

$$\Gamma_{\ge r} := \left\{ W \in \widetilde{\mathcal{W}}_0 \ : \ t_H(W) \ge r \right\}, \tag{3.3}$$

denoting by $\Gamma_{=r}$ the closed subset of graphons with $t_H(W) = r$. The unique global minimizer of $J_p(\cdot)$ over $\widetilde{\mathcal{W}}_0$ is $W_\star \equiv p$. With $W_\star \in \Gamma_{=t_H(p)}$, it follows that $\psi_H(p,r) = 0$ on $[0, t_H(p)]$. Next, for any $r \in (t_H(p), r_H]$, the good rate function $J_p(\cdot)$ is finite on the nonempty set $\Gamma_{\ge r} \cap \widetilde{\mathcal{W}}_0^{(p)}$, hence $\psi_H(p,r) = \alpha$ is finite and positive, with the infimum in (1.4) attained at the nonempty compact set

$$F_\star = \Gamma_{\ge r} \cap \{W \in \widetilde{\mathcal{W}}_0 \ : \ J_p(W) \le \alpha\}. \tag{3.4}$$

Fixing such r and $W_r \in F_\star$, consider the map $W_r(\lambda) := \lambda W_r + (1-\lambda) W_\star$ from $[0,1]$ to $\widetilde{\mathcal{W}}_0^{(p)}$. Thanks to the continuity of $\lambda \mapsto t_H(W_r(\lambda))$ on $[0,1]$, there exists for any $r' \in [t_H(p), t_H(W_r))$ some $\lambda' = \lambda'(r') \in [0,1)$ such that $t_H(W_r(\lambda')) = r'$. Hence, due to the convexity of $J_p(\cdot)$,

$$\psi_H(p,r') \le J_p(W_r(\lambda')) \le \lambda' J_p(W_r)) = \lambda' \alpha < \alpha := \psi_H(p,r).$$

We have shown that $\psi_H(p,r') < \psi_H(p,r)$ for all $r' \in [t_H(p), t_H(W_r))$. Recalling that $t_H(W_r) \ge r$, it follows that $\psi_H(p,\cdot)$ is strictly increasing on $[t_H(p), r_H]$ and further, that necessarily $t_H(W_r) = r$ for any $W_r \in F_\star$. That is, the collection $F_\star$ of minimizers of (1.4) then consists of only the minimizers of (1.5).

Next, if $\psi_H(p,r') \le \alpha < \infty$ for all $r' < r$ then there exist a pre-compact collection $\{W_{r'}, r' < r\}$ in $(\delta_\square, \widetilde{\mathcal{W}}_0)$, with $J_p(W_{r'}) \le \alpha$ and $t_H(W_{r'}) \ge r'$. By the continuity of $t_H(\cdot)$ and the LSC of $J_p(\cdot)$, it follows that $t_H(W_r) \ge r$ and $J_p(W_r) \le \alpha$ for any limit

point W_r of $W_{r'}$ as $r' \uparrow r$. Consequently $\psi_H(p,r) \le \alpha$ as well, establishing the stated left-continuity of $\psi_H(p,\cdot)$ on $[0,r_H]$. Finally, recall that an increasing function, finite on $[0,r_H]$ and infinite otherwise, is LSC iff it is left continuous on $[0,r_H]$.
(b) Considering the LDP bounds of Theorem 1.1 for the closed set $\Gamma_{\ge r}$ and its open subset $\Gamma_{>r} := \Gamma_{\ge r} \setminus \Gamma_{=r}$ we deduce that

$$-\lim_{r'\downarrow r}\{\psi_H(p,r')\} = -\inf_{W\in\Gamma_{>r}}\{J_p(W)\} \le \liminf_{n\to\infty} n^{-2}\log \mathbb{P}\left(t_H(G_n) > r\right) \\ \le \limsup_{n\to\infty} n^{-2}\log \mathbb{P}\left(t_H(G_n) \ge r\right) \le -\psi_H(p,r)\,.$$

By the assumed right-continuity of $t \mapsto \psi_H(p,t)$ at $r \in [0,r_H)$, the preceding inequalities must all hold with equality, resulting with (1.7).
(c) Proceeding to prove (1.8), we fix (p,r) as in part (b). Further fixing $\varepsilon > 0$, let

$$B_{W',\varepsilon} := \left\{W \in \widetilde{\mathcal{W}}_0 \,:\, \delta_\square(W,W') < \varepsilon\right\}$$

denote open cut-metric balls and consider the closed subset of $\Gamma_{\ge r}$,

$$\Gamma_{\ge r,\varepsilon} := \Gamma_{\ge r} \bigcap_{W'\in F_\star} (B_{W',\varepsilon})^c\,. \tag{3.5}$$

In view of (1.7) and the fact that

$$\{\delta_\square(G_n,F_\star) \ge \varepsilon,\; t_H(G_n) \ge r\} = \{W_{G_n} \in \Gamma_{\ge r,\varepsilon}\}\,,$$

it suffices for (1.8) to show that

$$\limsup_{n\to\infty} n^{-2}\log \mathbb{P}\left(W_{G_n} \in \Gamma_{\ge r,\varepsilon}\right) < -\alpha\,.$$

By the LDP upper-bound of Theorem 1.1, this in turn follows upon showing that

$$\inf_{W\in\Gamma_{\ge r,\varepsilon}}\{J_p(W)\} \le \alpha \tag{3.6}$$

contradicts the definition of $F_\star$. Indeed, $J_p(\cdot)$ has compact level sets, so if (3.6) holds then $J_p(W_r) \le \alpha$ for some $W_r \in \Gamma_{\ge r,\varepsilon}$. Recall (3.4) that in particular $W_r \in F_\star$, hence (3.5) implies that $\delta_\square(W_r,W_r) \ge \varepsilon > 0$, yielding the desired contradiction. ∎

3.3. **Sparse uniform random graphs.** In this section we show that, as was the case in $\mathcal{G}(n,p)$, the analog of (1.7), giving the asymptotic rate function for $\mathcal{G}(n,m)$ in the sparse regime $m_n = n^2/\log^c n$ for a suitably small $c > 0$, can be derived in a straightforward manner from the weak regularity lemma. Indeed, the proof below follows essentially the same short argument used for $\mathcal{G}(n,p)$ in [17, Prop. 5.1].

Definition 3.1 (Discrete variational problem for upper tails)**.** Let H be a graph with κ edges, and let $b > 1$. Denote the set of weighted undirected graphs on n vertices by

$$\widehat{\mathscr{G}}_n = \{(a_{ij})_{1\le i\le j\le n} : 0 \le a_{ij} \le 1\,,\, a_{ij} = a_{ji}\,,\, a_{ii} = 0 \text{ for all } i,j\}\,,$$

and extend the definition of the graphon $W_{\widehat{G}}$ in (1.2) to a weighted graph $\widehat{G} \in \widehat{\mathscr{G}_n}$ by replacing the weight 1 corresponding to an edge $(\lceil nx \rceil, \lceil ny \rceil)$ by the weight $a_{\lceil nx \rceil, \lceil ny \rceil}$. Further, for $p \in [0,1]$ we consider the subset of p-average weighted graphs

$$\widehat{\mathscr{G}_n}(p) = \Big\{ \widehat{G} \in \widehat{\mathscr{G}_n} : \sum_{i<j} a_{ij} = p\tbinom{n}{2} \Big\} .$$

Taking $m_n \leq \binom{n}{2}$ and $p_n = m_n / \binom{n}{2}$, the variational problem for $G_n \sim \mathcal{G}(n, m_n)$ is $\widehat{\psi}_H(n, p_n, b)$, where

$$\widehat{\psi}_H(s,p,b) := \inf \Big\{ I_p(W_{\widehat{G}}) \,:\, \widehat{G} \in \widehat{\mathscr{G}_s}(p)\,,\, t_H(W_{\widehat{G}}) \geq b\, p^{\kappa} \Big\} .$$

Recall that for any $n \geq s$, there exists $k \in \mathbb{N}$ be such that $s' = ks \in (n-s, n]$, and any $\widehat{G}' \in \widehat{\mathscr{G}_s}(p)$ yields $\widehat{G} \in \widehat{\mathscr{G}_n}(p)$ by duplicating each non-diagonal entry of $\widehat{G}'$ to fill the corresponding k-dimensional minor of the left-top s'-dimensional sub-matrix, setting to p all other off-diagonal entries of $\widehat{G}$. It further yields $I_p(W_{\widehat{G}}) \leq I_p(W_{\widehat{G}'})$ and $t_H(W_{\widehat{G}}) \geq (1 - s/n)^{V(H)} t_H(W_{\widehat{G}'})$. Consequently,

$$1 - \varepsilon' \leq (1 - s/n)^{V(H)} \quad \Longrightarrow \quad \widehat{\psi}_H(n, p, (1-\varepsilon')b) \leq \widehat{\psi}_H(s, p, b) . \tag{3.7}$$

Further, denoting by $\Delta \geq 2$ the maximal degree in H, we get upon replacing $\widehat{\mathscr{G}_n}(p)$ by $\widehat{\mathscr{G}_n}$ that for some $c(b) > 0$

$$\widehat{\psi}_H(n, p, b) \geq c(b) p^{\Delta} \log(1/p) \tag{3.8}$$

provided n large is enough and $p \gg n^{-1/\Delta}$ (see [2, Theorem 1.5]).

Remark 3.2. When $p_n \to p$ for a fixed $0 < p < 1$, and $r = bp^{\kappa} \in [p^{\kappa}, r_H]$ is a right-continuity point of $t \mapsto \psi_H(p, t)$ (whence (1.7) holds), one has $\psi_H(p, r) = \lim_{n\to\infty} \widehat{\psi}_H(n, p_n, b)$ (e.g., rescale a sequence $\widehat{G}_n$ of minimizers for $\widehat{\psi}_H(n, p_n, b + \varepsilon)$ by p/p_n; conversely, for a minimizer W for $\psi_H(p, r)$, take a sequence G_n with $W_{G_n} \to W$).

Proposition 3.3. *Let H be a fixed graph with κ edges, $b > 1$ and for $m_n \in \mathbb{N}$ let $G_n \sim \mathcal{G}(n, m_n)$ and $p_n = m_n / \binom{n}{2}$. For every $\varepsilon > 0$ there exists some $K < \infty$ such that, if $p_n (\log n)^{1/(2\kappa)} \geq K$ and n is sufficiently large then*

$$-(1+\varepsilon)\widehat{\psi}_H(n, p_n, (1+\varepsilon)b) \leq \frac{1}{n^2} \log \mathbb{P}(t_H(G_n) \geq b\, p_n^{\kappa}) \leq -(1-\varepsilon)\widehat{\psi}_H(n, p_n, (1-\varepsilon)b) .$$

In particular, if $p_n (\log n)^{1/(2\kappa)} \to \infty$ and $\widehat{\psi}_H(n, p_n, b') / \widehat{\psi}_H(n, p_n, b) \to 1$ for $n \to \infty$ followed by $b' \to b$, then

$$\lim_{n\to\infty} \frac{\log \mathbb{P}\left(t_H(G_n) \geq b\, p_n^{\kappa}\right)}{n^2 \widehat{\psi}_H(n, p_n, b)} = -1 .$$

The following simple lemma, whose analog for upper tails in $\mathcal{G}(n,p)$ (addressing only the event $\mathcal{E}_1(G)$ below) was phrased in [17, Lemma 5.2] for triangle counts in $\mathcal{G}(n,p)$, is an immediate consequence of the independence of distinct edges and Cramér's Theorem.

Lemma 3.4. *For $m \ge n \ge s \in \mathbb{N}$, set $p = m/\binom{n}{2}$ and $\widehat{G} = (a_{ij}) \in \widehat{\mathscr{G}}_s(p)$. For an equitable partition $V_1, \ldots, V_s$ of $\{1, \ldots, n\}$ (i.e., $|V_i| - |V_j| \le 1$ for all i, j), define*

$$\mathcal{E}_1(G) = \bigcap_{\substack{i,j \\ a_{ij} > p}} \{d_G(V_i, V_j) \ge a_{ij}\} \,, \qquad \mathcal{E}_2(G) = \bigcap_{\substack{i,j \\ a_{ij} \le p}} \{d_G(V_i, V_j) \le a_{ij}\} \,,$$

where $d_G(X, Y) = \frac{\#\left\{(x,y) \in X \times Y : xy \in E(G)\right\}}{|X||Y|}$. Then, $G_n \sim \mathcal{G}(n, m)$ has

$$\frac{1}{n^2} \log \mathbb{P}\left(\mathcal{E}_1(G) \cap \mathcal{E}_2(G)\right) \le -I_p(W_{\widehat{G}})\big(1 - \frac{3s}{n}\big) + \zeta_n \,, \tag{3.9}$$

where $\zeta_n = \frac{1}{2} n^{-2} \log(cm)$ for some $c < \infty$.

Proof. Let $G'_n \sim \mathcal{G}(n, p)$ and $v_{ij} = |V_i||V_j|$ when $i < j$, $v_{ii} = \binom{|V_i|}{2}$. Recall that $d_{G'_n}(V_i, V_j) \sim v_{ij}^{-1} \operatorname{Bin}(v_{ij}, p)$ for $i \le j$ are mutually independent. Thus, considering for each ij the optimal Cramér's bound, we find that

$$\frac{1}{n^2} \log \mathbb{P}\left(\mathcal{E}_1(G'_n) \cap \mathcal{E}_2(G'_n)\right) \le -\frac{1}{n^2} \sum_{i \le j} 2 v_{ij} I_p(a_{ij})$$

(e.g. see [6, (2.12),(2.1.13)]). Since $\min_{i \ne j}(v_{ij}, 2v_{ii}) \ge (n/s)^2 (1 - 3s/n)$ for any equitable partition, also

$$-\frac{1}{n^2} \sum_{i \le j} 2 v_{ij} I_p(a_{ij}) \le -I_p(W_{\widehat{G}})\big(1 - \frac{3s}{n}\big) \,.$$

Next, since $\mathbb{P}(G_n \in \cdot) = \mathbb{P}(G'_n \in \cdot \mid |E(G'_n)| = m)$, it follows that

$$\log \mathbb{P}\left(\mathcal{E}_1(G_n) \cap \mathcal{E}_2(G_n)\right) - \log \mathbb{P}\left(\mathcal{E}_1(G'_n) \cap \mathcal{E}_2(G'_n)\right) \le -\log \mathbb{P}(\operatorname{Bin}(\tbinom{n}{2}, p) = m) \,.$$

Since $p = m/N$ for $N = \binom{n}{2}$, we complete the proof of (3.9) upon recalling that $\sqrt{cm}\, \mathbb{P}(\operatorname{Bin}(N, m/N) = m) \ge 1$ for some $c < \infty$ and all $N \ge m \ge 0$. ■

Combining the weak regularity lemma (see, e.g., [14, Lemma 9.3]) with the counting lemma for graphons (cf., e.g., [14, Lemma 10.23]) implies the following.

Lemma 3.5. *Let $\eta > 0$ and set $M = 4^{4/\eta^2}$. For every graph G there is an equitable partition $V_1, \ldots, V_s$ of its vertices, for some $s \in [M/2, M]$, such that the weighted graph $\widehat{G} \in \widehat{\mathscr{G}}_s$ with $a_{ij} = d_G(V_i, V_j)$ has $\left|t_H(G) - t_H(\widehat{G})\right| \le \kappa\eta$ for every graph H with κ edges.*

Proof of Proposition 3.3. Fixing $\xi = 1/r$ for $r \in \mathbb{N}$, consider Lemma 3.5 for $\eta_n = \frac{\xi}{\kappa} b p_n^\kappa$. Since $p_n \ge K(\log n)^{-1/(2\kappa)}$ we have for some $\delta = \delta(K, r, \kappa, b) \to 0$ as $K \to \infty$ and $M = M_n \le n^\delta$, that if G_n has $t_H(G_n) \ge b p_n^\kappa$ and $|E(G_n)| = m_n$ then there exists an equitable partition $V_1, \ldots, V_s$ of its vertices, for some $s \in [M/2, M]$, such that the corresponding weighted graph $\widehat{G} \in \widehat{\mathscr{G}}_s(p'_n)$ for some $p'_n = p_n(1 + O(n^{-\delta}))$, satisfies $t_H(W_{\widehat{G}}) \ge b p_n^\kappa (1 - \xi)$. We may round each of the weights a_{ij} of $\widehat{G}$ up to the nearest multiple of $\xi p_n/2$ (only increasing t_H), with the effect of yielding an off-diagonal average which is an integer multiple of $\xi p_n/(s(s-1))$ in the range $[p_n, p_n(1+\xi)]$. Setting $e(\xi) = (1-\xi)/(1+\xi)^\kappa$ we arrive by rescaling at $\widehat{G}' \in \widehat{\mathscr{G}}_s(p_n)$ such that

$$t_H(W_{\widehat{G}'}) \geq e(\xi) b p_n^\kappa .$$

In this process we apply at most $2s^2$ scaling factors on at most $(2(r+1)s^2)^{s^2}$ arrays $\{a_{ij}\}$, thereby considering at most $\exp(CM^2 \log M)$ possible weighted graphs $\widehat{G}'$. Proceeding to establish the stated upper bound, WLOG we have that $1-\varepsilon := (1-\varepsilon')^2 > 1/b$. Then, choose $r \in \mathbb{N}$ large so $e(\xi) \geq 1-\varepsilon'$, hence for $n \geq n_0(\varepsilon')$, $s \leq M$, both $3s/n \leq \varepsilon'$ and

$$\widehat{\psi}_H(n, p_n, (1-\varepsilon)b) \leq \widehat{\psi}_H(s, p_n, e(\xi)b)$$

due to (3.7). We then apply the uniform bound of Lemma 3.4 for all such $\widehat{G}' \in \widehat{\mathscr{G}}_s(p_n)$ and the corresponding equitable partitions for them. The union bound over all those $\widehat{G}'$ and the M^n possible partitions for each, entails at most $\exp(O(n \log n))$ terms (once $\delta < 1/2$). For our choice of p_n the latter factor is negligible since

$$\lim_{n\to\infty} \frac{n}{\log n} \widehat{\psi}_H(n, p_n, (1-\varepsilon)b) = \infty \tag{3.10}$$

(see (3.8)). For the same reason, the uniform additive error ζ_n in (3.9) is also at most $\varepsilon' \widehat{\psi}_H(n, p_n, (1-\varepsilon)b)$, so both can be embedded within the allowed tolerance.

Turning to prove the stated lower bound, fix $\varepsilon < 1$ and $\widehat{G} = (a_{ij}) \in \widehat{\mathscr{G}}_n(p_n)$ such that

$$2\sum_{i<j} I_{p_n}(a_{ij}) \leq (1+\varepsilon') n^2 \widehat{\psi}_H(n, p_n, b') , \qquad t_H(W_{\widehat{G}}) \geq b'\, p_n^\kappa , \tag{3.11}$$

where $b' = (1+\varepsilon)b = (1+\varepsilon')^2\, b$. Then, construct the law $\mathbb{P}$ of $G_n \sim \mathcal{G}(n, m_n)$ out of the law $\mathbb{Q}$ where edge $x_{ij} = 1$ is present with probability a_{ij}, independently of all other edges. Specifically, the log-likelihood between the law $\mathbb{P}'$ of $G_n' \sim \mathcal{G}(n, p_n)$ and $\mathbb{Q}$ is

$$\log \frac{d\mathbb{P}'}{d\mathbb{Q}}(\underline{x}) := -h(\underline{x}) = -\sum_{i<j} \Big[x_{ij} \log \frac{a_{ij}}{p_n} + (1-x_{ij}) \log \frac{1-a_{ij}}{1-p_n} \Big] , \tag{3.12}$$

hence

$$\mathbb{P}(t_H(G_n) \geq b\, p_n^\kappa) \geq \mathbb{E}_{\mathbb{Q}} \Big[e^{-h(\underline{x})} \; ; \; t_H(\underline{x}) \geq b\, p_n^\kappa, \quad \sum_{i<j} x_{ij} = m_n \Big] .$$

Further, since $\mathbb{E}_{\mathbb{Q}}[x_{ij}] = a_{ij}$ whose sum is m_n and $t_H(\cdot)$ is multi-linear in (a_{ij}), we know from (3.11) that

$$\mathbb{E}_{\mathbb{Q}}[h] \leq (1+\varepsilon') n^2 \widehat{\psi}_H(n, p_n, b') , \qquad \mathbb{E}_{\mathbb{Q}}[t_H] \geq b' p_n^\kappa . \tag{3.13}$$

Recall (3.10) that $n^2 \widehat{\psi}_H(n, p_n, b') \gg \log n$, hence it suffices to show that

$$n\, \mathbb{Q}\Big(\mathcal{E}_1 \cap \mathcal{E}_2 \cap \mathcal{E}_3\Big) \geq n \Big[\mathbb{Q}(\mathcal{E}_1) - \mathbb{Q}(\mathcal{E}_2^c) - \mathbb{Q}(\mathcal{E}_3^c) \Big] \tag{3.14}$$

is bounded away from zero when $n \to \infty$, where

$$\mathcal{E}_1 := \Big\{ \underline{x} : \quad \textstyle\sum_{i<j} (x_{ij} - a_{ij}) = 0 \Big\} \,,$$
$$\mathcal{E}_2 := \Big\{ \underline{x} : \ \big| h(\underline{x}) - \mathbb{E}_{\mathbb{Q}}[h] \big| \le \varepsilon' n^2 \widehat{\psi}_H(n, p_n, b') \Big\} \,,$$
$$\mathcal{E}_3 := \Big\{ \underline{x} : \ \big| t_H(\underline{x}) - \mathbb{E}_Q[t_H] \big| \le \frac{\varepsilon}{2} \mathbb{E}_Q[t_H] \Big\} \,.$$

To this end, recall [9, Theorem 5] that for all $n \ge n_0$

$$n\mathbb{Q}(\mathcal{E}_1) \ge n\mathbb{P}'(\mathcal{E}_1) \ge \frac{1}{\sqrt{\pi}} \,. \tag{3.15}$$

Turning to deal with $\mathcal{E}_2^c$, note that $h(\underline{x}) - \mathbb{E}_{\mathbb{Q}}[h] = \sum_{i<j} (x_{ij} - a_{ij}) c_{p_n}(a_{ij})$ with

$$\mathrm{Var}_{\mathbb{Q}}(h) = \sum_{i<j} v_{p_n}(a_{ij}) \,, \quad v_p(a) := a(1-a) c_p(a)^2 \,, \quad c_p(a) := \log\Big(\frac{a}{p}\Big) - \log\Big(\frac{1-a}{1-p}\Big) \,.$$

Let $f_p(a) := v_p(a) - g_p I_p(a)$ for $g_p := 2 + \log[1/(p(1-p))]$. Since $a(1-a)c'_p(a) = 1$ and $I'_p(a) = c_p(a)$, it follows that

$$\frac{f'_p(a)}{c_p(a)} = 2 + (1-2a) c_p(a) - g_p \le 0 \,.$$

Further, $f_p(p) = 0$ whereas $c_p(a) \ge 0$ iff $a \ge p$, hence $v_p(a) \le g_p I_p(a)$ for all $p, a \in [0,1]$. Consequently, by Markov's inequality and the LHS of (3.11)

$$\mathbb{Q}(\mathcal{E}_2^c) = \mathbb{Q}\Big(\big| h(\underline{x}) - \mathbb{E}_{\mathbb{Q}}[h] \big| > \varepsilon' n^2 \widehat{\psi}_H(n, p_n, b') \Big) \le \frac{(1+\varepsilon') g_{p_n}}{2(\varepsilon' n)^2 \widehat{\psi}_H(n, p_n, b')} \,.$$

In view of (3.8) and the assumed lower bound on p_n, we have from the latter bound that for some $c = c(\kappa, \varepsilon', b') < \infty$, as $n \to \infty$,

$$n\, \mathbb{Q}(\mathcal{E}_2^c) \le c (\log n)^c n^{-1} \to 0 \,. \tag{3.16}$$

Finally, up to scaling, $Y(n) := t_H(\underline{x})$ of mean $\mu(n) := \mathbb{E}_{\mathbb{Q}}[t_H]$, is the polynomial

$$\sum_{e' \in \mathsf{E}'} \prod_{ij \in e'} x_{ij} \,,$$

in $\{x_{ij}, i < j\}$, with hyper-edges of $\mathsf{G}' = (\mathsf{K}_n, \mathsf{E}')$ enumerating copies of H within the complete graph K_n. Forcing even one fixed $(ij) \in H$ determines two labels $i < j$ of vertices from $e' \in \mathsf{E}'$, thereby reducing the cardinality of the collection of such e' to at most $O(n^{-2})$ of $|\mathsf{E}'|$. Recall that $n^2 \mu(n) (\log n)^{-2\kappa} \to \infty$ thanks to the RHS of (3.13) and the assumed lower bound on p_n. Thus, from the polynomial concentration of Kim–Vu (see, e.g., [1, Theorem 7.8.1, p. 122] with $\lambda = 2\kappa \log n$), we have for all $n \ge n_0(\kappa, \varepsilon, b')$ and any choice of (a_{ij}), that

$$\mathbb{Q}(\mathcal{E}_3^c) = \mathbb{Q}\Big(\, |Y(n) - \mu(n)| > \frac{\varepsilon}{2} \mu(n) \Big) \le n^{-2} \,. \tag{3.17}$$

Combining (3.15)–(3.17) establishes the claimed uniform bound away from zero in (3.14), thereby completing the proof of the stated lower bound. ∎

Acknowledgment. A.D. was supported in part by NSF grant DMS-1613091 and E.L. was supported in part by NSF grants DMS-1513403 and DMS-1812095.

References

[1] N. Alon and J. H. Spencer. *The probabilistic method*. Wiley Series in Discrete Mathematics and Optimization. John Wiley & Sons, Inc., Hoboken, NJ, fourth edition, 2016.

[2] B. B. Bhattacharya, S. Ganguly, E. Lubetzky, and Y. Zhao. Upper tails and independence polynomials in random graphs. *Adv. Math.*, 319:313–347, 2017.

[3] C. Borgs, J. Chayes, L. Lovász, V. Sós, and K. Vesztergombi. Convergent sequences of dense graphs. I. Subgraph frequencies, metric properties and testing. *Adv. Math.*, 219:1801–1851, 2008.

[4] S. Chatterjee and A. Dembo. Nonlinear large deviations. *Adv. Math.*, 299:396–450, 2016.

[5] S. Chatterjee and S. R. S. Varadhan. The large deviation principle for the Erdős-Rényi random graph. *European J. Combin.*, 32(7):1000–1017, 2011.

[6] A. Dembo and O. Zeitouni. *Large deviations techniques and applications*, volume 38 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2010.

[7] F. den Hollander, M. Mandjes, A. Roccaverde, and N. J. Starreveld. Ensemble equivalence for dense graphs. *Electron. J. Probab.*, 23:Paper No. 12, 26 pp., 2018.

[8] R. Eldan. Gaussian-width gradient complexity, reverse log-Sobolev inequalities and nonlinear large deviations. *Geom. Funct. Anal.* To appear.

[9] W. Hoeffding. On the distribution of the number of successes in independent trials. *Ann. Math. Statist.*, 27:713–721, 1956.

[10] S. Janson, T. Łuczak, and A. Rucinski. *Random graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000.

[11] R. Kenyon, C. Radin, K. Ren, and L. Sadun. Multipodal structure and phase transitions in large constrained graphs. *J. Stat. Phys.*, 168(2):233–258, 2017.

[12] R. Kenyon, C. Radin, K. Ren, and L. Sadun. The phases of large networks with edge and triangle constraints. *J. Phys. A*, 50(43):435001, 22, 2017.

[13] R. Kenyon, C. Radin, K. Ren, and L. Sadun. Bipodal structure in oversaturated random graphs. *Int. Math. Res. Not. (IMRN)*, 2018(4):1009–1044, 2018.

[14] L. Lovász. *Large networks and graph limits*, volume 60 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2012.

[15] L. Lovász and B. Szegedy. Limits of dense graph sequences. *J. Combin. Theory Ser. B*, 96(6):933–957, 2006.

[16] E. Lubetzky and Y. Zhao. On replica symmetry of large deviations in random graphs. *Random Structures Algorithms*, 47(1):109–146, 2015.

[17] E. Lubetzky and Y. Zhao. On the variational problem for upper tails in sparse random graphs. *Random Structures Algorithms*, 50(3):420–436, 2017.

[18] C. Radin and L. Sadun. Phase transitions in a complex network. *J. Phys. A*, 46(30):305002, 2013.

[19] Y. Zhao. On the lower tail variational problem for random graphs. *Combin. Probab. Comput.*, 26(2):301–320, 2017.

Amir Dembo

Department of Mathematics, Stanford University, Sloan Hall, Stanford, CA 94305, USA.

Email address: amir@math.stanford.edu

Eyal Lubetzky

Courant Institute, New York University, 251 Mercer Street, New York, NY 10012, USA.

Email address: eyal@courant.nyu.edu