

# LARGE DEVIATION PRINCIPLE FOR THE NORM OF THE LAPLACIAN MATRIX OF INHOMOGENEOUS ERDŐS-RÉNYI RANDOM GRAPHS

RAJAT SUBHRA HAZRA, FRANK DEN HOLLANDER, AND MAARTEN MARKERING

ABSTRACT. We consider an inhomogeneous Erdős-Rényi random graph  $G_N$  with vertex set  $[N] = \{1, \dots, N\}$  for which the pair of vertices  $i, j \in [N]$ ,  $i \neq j$ , is connected by an edge with probability  $r_N(\frac{i}{N}, \frac{j}{N})$ , independently of other pairs of vertices. Here,  $r_N: [0, 1]^2 \rightarrow (0, 1)$  is a symmetric function that plays the role of a reference graphon. Let  $\lambda_N$  be the maximal eigenvalue of the Laplacian matrix of  $G_N$ . We show that if  $\lim_{N \rightarrow \infty} \|r_N - r\|_\infty = 0$  for some limiting graphon  $r: [0, 1]^2 \rightarrow (0, 1)$ , then  $\lambda_N/N$  satisfies a downward LDP with rate  $\binom{N}{2}$  and an upward LDP with rate  $N$ . We identify the associated rate functions  $\psi_r$  and  $\widehat{\psi}_r$ , and derive their basic properties.

## 1. INTRODUCTION AND MAIN RESULTS

Section 1.1 provides background. Section 1.2 states the LDP for the empirical graphon associated with inhomogeneous Erdős-Rényi random graphs. Section 1.3 looks at graphon operators, in particular, the Laplacian operator that is the central object in the present paper. Sections 1.4–1.5 state the downward, respectively, upward LDP for the largest eigenvalue of the Laplacian matrix and present basic properties of the associated rate functions. Section 1.6 places the various theorems in their proper context.

**1.1. Background.** Spectra of matrices associated with a graph play a crucial role in understanding the geometry of the graph. Given a finite graph on  $N$  vertices, two important matrices are the *adjacency matrix*  $A_N$  and the *Laplacian matrix*  $L_N = D_N - A_N$ , where  $D_N$  is the diagonal matrix whose elements are the degrees of the vertices. In this paper we focus on the largest eigenvalue of  $L_N$  when the underlying graph is a *dense inhomogeneous* Erdős-Rényi random graph. The largest eigenvalue of  $A_N$  satisfies a large deviation principle (LDP). This fact is an immediate consequence of the LDP for the empirical graphon derived in [16] and [25] in combination with the contraction principle, because the norm of the adjacency graphon operator is bounded and continuous on the space of graphons. The rate function is given in terms of a variational formula involving the rate function of the LDP for the empirical graphon. In [11] we analysed this variational formula, identified the basic properties of the rate function, and identified its scaling behaviour near its unique minimiser and its two boundary points.

The extension of the LDP to  $L_N$  poses new challenges, because  $L_N$  is a more delicate object than  $A_N$ . For one, the upward and the downward large deviations for the largest eigenvalue of  $L_N$  live on *different scales*, and the norm of the Laplacian graphon operator lacks certain continuity properties that hold for the norm of the adjacency graphon operator. Like for  $A_N$ , it is not possible to explicitly solve the variational formulas for the two associated rate functions. Nonetheless, we derive their basic properties and identify their scaling behaviour near their

---

*Date:* July 5, 2023.

*2000 Mathematics Subject Classification.* 05C80, 60B20, 60C05, 60F10.

*Key words and phrases.* Inhomogeneous Erdős-Rényi random graph, Laplacian matrix, largest eigenvalue, graphons, large deviation principle, rate function.

The research in this paper was supported through NWO Gravitation Grant NETWORKS 024.002.003.

unique minimisers. Our proofs require an analysis of the functional analytic properties of the Laplacian operator, in combination with variational tools.

The literature on the eigenvalues of the Laplacian of a random matrix is limited. For a general Wigner matrix with independent centered entries satisfying certain moment conditions, the empirical spectral distribution was derived in [8] and was identified as the free additive convolution of the semicircle law and the Gaussian distribution. For entries with a common mean  $m$  it was shown that the empirical spectral distribution converges to the Dirac measure at  $m$ . Moreover, under a fourth moment condition it was shown that the spectral norm is of order  $\sqrt{2N \log N}$  when the entries are centered, and of order  $N$  when the entries are not centered. Although the moment conditions were rather general, [8] did not cover all relevant cases of Erdős-Rényi random graphs. This was resolved in [20], [18] for the case where the average degree diverges with  $N$ . The results were extended to inhomogeneous Erdős-Rényi graphs in [10], [12]. The empirical distribution in the sparse homogeneous Erdős-Rényi graph was identified in [21]. Later, an alternative proof through local weak convergence was provided in [6]. The results on extreme eigenvalues are limited and challenging. Recently, it was shown in [9] that for the Gaussian orthogonal ensemble the largest eigenvalue of the Laplacian, after appropriate scaling and centering, converges to the Gumbel distribution.

Large deviations for Erdős-Rényi random graphs were explored extensively in various works, including [13], [14], [24] by utilising graphon theory, specifically subgraph densities and maximal eigenvalues. For a comprehensive review of the literature, we refer to [13]. Large deviation theory for random matrices originated in [3], where the focus was on the empirical spectral distribution of  $\beta$ -ensembles with a quadratic potential. The rate was found to be  $N^2$ , and the rate function was identified by using a non-commutative notion of entropy. Subsequently, the maximal eigenvalue of such ensembles was investigated in [4]. In [5], large deviations of the empirical spectral distribution were derived for random matrices with non-Gaussian tails, while the study of the maximal eigenvalue in that context was carried out in [1], [2]. However, the adjacency matrix and the Laplacian matrix of an inhomogeneous Erdős-Rényi random graph fall outside these regimes. The goal of the present paper is to understand the *large deviations of the largest eigenvalue* of an *inhomogeneous Erdős-Rényi graph* under certain density assumptions.

## 1.2. LDP for inhomogeneous Erdős-Rényi random graphs.

1.2.1. *Graphons.* Let

$$(1.1) \quad \mathcal{W} = \{h: [0, 1]^2 \rightarrow [0, 1]: h(x, y) = h(y, x) \forall x, y \in [0, 1]\}$$

denote the set of graphons. Let  $\mathcal{M}$  denote the set of Lebesgue measure-preserving bijective maps  $\phi: [0, 1] \mapsto [0, 1]$ . For two graphons  $h_1, h_2 \in \mathcal{W}$ , the *cut-distance* is defined by

$$(1.2) \quad d_{\square}(h_1, h_2) = \sup_{S, T \subset [0, 1]} \left| \int_{S \times T} dx dy [h_1(x, y) - h_2(x, y)] \right|,$$

and the *cut-metric* by

$$(1.3) \quad \delta_{\square}(h_1, h_2) = \inf_{\phi \in \mathcal{M}} d_{\square}(h_1, h_2^{\phi}),$$

where  $h_2^{\phi}(x, y) = h_2(\phi(x), \phi(y))$ . The cut-metric defines an equivalence relation  $\sim$  on  $\mathcal{W}$  by declaring  $h_1 \sim h_2$  if and only if  $\delta_{\square}(h_1, h_2) = 0$ , and leads to the quotient space  $\widetilde{\mathcal{W}} = \mathcal{W}/\sim$ . For  $h \in \mathcal{W}$ , we write  $\widetilde{h}$  to denote the equivalence class of  $h$  in  $\widetilde{\mathcal{W}}$ . The equivalence classes correspond to relabelings of the vertices of the graph. The pair  $(\widetilde{\mathcal{W}}, \delta_{\square})$  is a compact metric space [23].

1.2.2. *LDP for empirical graphon.* Consider a sequence of symmetric functions  $r_N: [0, 1]^2 \rightarrow (0, 1)$ ,  $N \in \mathbb{N}$ , that are constant on the blocks  $[\frac{i-1}{N}, \frac{i}{N}) \times [\frac{j-1}{N}, \frac{j}{N})$ ,  $1 \leq i, j \leq N$ , such that

$$(1.4) \quad \lim_{N \rightarrow \infty} \|r_N - r\|_\infty = 0$$

for some symmetric function  $r: [0, 1]^2 \rightarrow (0, 1)$  that plays the role of a *reference graphon*. Let  $G_N$  be the inhomogeneous Erdős-Rényi random graph with vertex set  $[N] = \{1, \dots, N\}$  for which the pair of vertices  $i, j \in [N]$ ,  $i \neq j$ , is connected by an edge with probability  $(r_N)_{ij}$ , independently of other pairs of vertices, where  $(r_N)_{ij}$  is the value of  $r_N$  on the block  $[\frac{i-1}{N}, \frac{i}{N}) \times [\frac{j-1}{N}, \frac{j}{N})$ . The function  $r_N$  plays the role of a *block reference graphon* converging to some *reference graphon*  $r$  as  $N \rightarrow \infty$ . We assume that the reference graphon  $r \in \mathcal{W}$  in (1.4) satisfies

$$(1.5) \quad \log r, \log(1 - r) \in L^1([0, 1]^2),$$

and that the analogue of (1.5) holds for  $\log r_N$  and  $\log(1 - r_N)$  as well.

Let  $A_N$  be the *adjacency matrix* of  $G_N$  defined by

$$(1.6) \quad A_N(i, j) = \begin{cases} 1, & \text{if there is an edge between vertex } i \text{ and vertex } j, \\ 0, & \text{otherwise.} \end{cases}$$

Let  $D_N$  be the diagonal degree matrix defined by  $D_N(i, i) = \sum_{j \in [N] \setminus \{i\}} A_N(i, j)$  and  $D_N(i, j) = 0$  for  $i \neq j$ , and put  $L_N = D_N - A_N$ , which is the *Laplacian matrix*. Write  $\mathbb{P}_N$  to denote the law of  $G_N$ . Use the same symbol for the law on  $\mathcal{W}$  induced by the map that associates with the graph  $G_N$  its empirical graphon  $h^{G_N}$ , defined by

$$(1.7) \quad h^{G_N}(x, y) = \begin{cases} 1, & \text{if there is an edge between vertex } \lceil Nx \rceil \text{ and vertex } \lceil Ny \rceil, \\ 0, & \text{otherwise.} \end{cases}$$

Write  $\tilde{\mathbb{P}}_N$  to denote the law of  $\tilde{h}^{G_N}$ .

The following LDP, which is an extension of the celebrated LDP for homogeneous ERRG derived in [14], is proved in [16] and [25]. (For more background on large deviation theory, see, for instance, [15].)

**Theorem 1.1.** [LDP for inhomogeneous ERRG, [16], [25]] *Subject to (1.5), the sequence  $(\tilde{\mathbb{P}}_N)_{N \in \mathbb{N}}$  satisfies the large deviation principle on  $(\tilde{\mathcal{W}}, \delta_\square)$  with rate  $\binom{N}{2}$ , i.e.,*

$$(1.8) \quad \begin{aligned} \limsup_{N \rightarrow \infty} \binom{N}{2}^{-1} \log \tilde{\mathbb{P}}_N(\mathcal{C}) &\leq - \inf_{\tilde{h} \in \mathcal{C}} J_r(\tilde{h}) \quad \forall \mathcal{C} \subset \tilde{\mathcal{W}} \text{ closed,} \\ \liminf_{N \rightarrow \infty} \binom{N}{2}^{-1} \log \tilde{\mathbb{P}}_N(\mathcal{O}) &\geq - \inf_{\tilde{h} \in \mathcal{O}} J_r(\tilde{h}) \quad \forall \mathcal{O} \subset \tilde{\mathcal{W}} \text{ open,} \end{aligned}$$

where the rate function  $J_r: \tilde{\mathcal{W}} \rightarrow \mathbb{R}$  is given by

$$(1.9) \quad J_r(\tilde{h}) = \inf_{\phi \in \mathcal{M}} I_r(h^\phi),$$

where  $h$  is any representative of  $\tilde{h}$  and

$$(1.10) \quad I_r(h) = \int_{[0, 1]^2} dx dy \mathcal{R}(h(x, y) \mid r(x, y)), \quad h \in \mathcal{W},$$

with

$$(1.11) \quad \mathcal{R}(a \mid b) = a \log \frac{a}{b} + (1 - a) \log \frac{1-a}{1-b}$$

the relative entropy of two Bernoulli distributions with success probabilities  $a \in [0, 1]$ ,  $b \in (0, 1)$  (with the convention  $0 \log 0 = 0$ ).

**Remark 1.2.** Theorem 1.1 was proved in [16] under the assumption that  $r$  is bounded away from 0 and 1. In [25] this assumption was relaxed to (1.5), and it was also shown that  $J_r$  is a good rate function, i.e.,  $J_r \not\equiv \infty$  and  $J_r$  has compact level sets. Note that (1.9) differs from the expression in [16], where the rate function is the lower semi-continuous envelope of  $I_r(h)$ . However, as shown in [25], under (1.5) the two rate functions are equivalent, since  $J_r(\tilde{h})$  is lower semi-continuous on  $\widetilde{\mathcal{W}}$ . ♠

As an application of Theorem 1.1, it was shown in [11] that the largest eigenvalue of the adjacency matrix satisfies the LDP with rate  $\binom{N}{2}$ . The rate function was analysed in detail for reference graphons that are rank-1. In the present paper we focus on the LDP for the largest eigenvalue of the Laplacian matrix  $L_N$ .

**1.3. Graphon operators.** For  $h \in \mathcal{W}$ , the graphon operator  $\mathcal{T}_h$  is the integral operator on  $L^2([0, 1])$  defined by

$$(1.12) \quad (\mathcal{T}_h u)(x) = \int_{[0,1]} dy h(x, y)u(y), \quad x \in [0, 1].$$

Note that  $\mathcal{T}_h$  is a compact operator. Define the *degree function* as

$$(1.13) \quad d_h(x) = \int_{[0,1]} dy h(x, y), \quad x \in [0, 1].$$

The *degree operator*  $\mathcal{D}_h$  is the multiplication operator on  $L^2([0, 1])$  defined by

$$(1.14) \quad (\mathcal{D}_h u)(x) = d_h(x)u(x),$$

The *Laplacian operator*  $\mathcal{L}_h$  is the linear integral operator on  $L^2([0, 1])$  defined by

$$(1.15) \quad (\mathcal{L}_h u)(x) = \int_{[0,1]} dy h(x, y)[u(x) - u(y)], \quad x \in [0, 1].$$

Note that

$$(1.16) \quad \mathcal{L}_h = \mathcal{D}_h - \mathcal{T}_h.$$

Recall that, given an operator  $S$  on a Hilbert space, the *spectrum* of  $S$  is defined as

$$(1.17) \quad \sigma(S) = \{\lambda \in \mathbb{C} : S - \lambda I \text{ is not invertible}\}.$$

Let  $\sigma_d(S)$  denote the *discrete spectrum* of  $S$ , which consists of all the isolated eigenvalues with finite algebraic multiplicity. The *essential spectrum* of  $S$  is denoted by

$$(1.18) \quad \sigma_{\text{ess}}(S) = \sigma(S) \setminus \sigma_d(S).$$

The essential spectrum is closed, and the discrete spectrum can only have accumulation points on the boundary of the essential spectrum. Since it is known that compact operators do not affect the essential spectrum, we have

$$(1.19) \quad \sigma_{\text{ess}}(\mathcal{L}_h) = \sigma_{\text{ess}}(\mathcal{D}_h).$$

The operator  $\mathcal{L}_h$  is not as well-behaved as  $\mathcal{T}_h$  with the cut-norm. In fact, even when a sequence of graphons  $(h_n)_{n \in \mathbb{N}}$  converges in cut-norm to a graphon  $h$ , the eigenvalues and eigenvectors of  $\mathcal{L}_{h_n}$  may not converge to those of  $\mathcal{L}_h$ , as was already observed in [17], [27]. Assuming as in [11] that the reference graphon is rank-1 does not help. In fact, if we assume that  $r$  is rank-1 and is continuous, then 0 is the only eigenvalue of  $\mathcal{L}_r$  and  $\sigma_{\text{ess}}(\mathcal{L}_r) = d_r([0, 1])$ , as shown in [17, Proposition 5.11].

If  $h$  is the empirical graphon of a graph  $G$  with  $N$  vertices, then  $N\|\mathcal{T}_h\|$  equals the largest eigenvalue of the adjacency matrix of  $G$ , and  $N\|\mathcal{D}_h\|$  equals the maximum degree of  $G$ . In fact,

for any graphon  $h$  the spectrum of  $\mathcal{D}_h$  equals the range of  $d_h$  and the operator norm of  $\mathcal{D}_h$  equals the supremum norm  $d_h$ , i.e.,

$$(1.20) \quad \|\mathcal{D}_h\| = \|d_h\|_\infty,$$

where  $\|d_h\|_\infty$  is the  $L^\infty$ -norm of the function  $d_h$ . This follows from the fact that  $\mathcal{D}_h$  is a multiplication operator. Let

$$(1.21) \quad \|\mathcal{L}_h\| = \sup_{\substack{u \in L^2([0,1]) \\ \|u\|_2=1}} \|\mathcal{L}_h u\|_2$$

be the operator norm of  $\mathcal{L}_h$ , where  $\|\cdot\|_2$  denotes the  $L^2$ -norm. Since  $\mathcal{L}_h$  is a normal operator with a non-negative spectrum,  $\|\mathcal{L}_h\|$  also equals the supremum of the spectrum of  $\mathcal{L}_h$ .

**Proposition 1.3. [Properties of the Laplacian operator]**

(i) Let  $h$  be a graphon and  $\mathcal{L}_h$  be the Laplacian operator on  $L^2([0,1])$ . Then  $\mathcal{L}_h$  is a bounded operator, and  $h \mapsto \|\mathcal{L}_h\|$  is lower semi-continuous in the cut-metric.

(ii) Let  $G$  be a graph with  $N$  vertices and  $h^G$  be the empirical graphon associated with  $G$  and let  $L_N$  be the Laplacian matrix with spectral norm  $\|L_N\|$ . Then it follows that

$$(1.22) \quad \frac{\|L_N\|}{N} = \|\mathcal{L}_{h^G}\| \quad \forall N.$$

**Remark 1.4.** Note that  $h \mapsto \|\mathcal{L}_h\|$  is not continuous in the cut-metric. For example, consider the sequence of graphons  $(h_N)_{N \in \mathbb{N}}$  such that  $h_N$  is the empirical graphon of the  $N$ -star graph (i.e., 1 vertex connected by an edge to each of the  $N - 1$  other vertices, and no further edges). Then  $h_N \downarrow 0$  as  $N \rightarrow \infty$  in the cut-metric, but  $\|\mathcal{L}_{h_N}\| = 1$  for all  $N \in \mathbb{N}$ . ♠

Proposition 1.3 is proven in Section 2.2.

**1.4. Main theorems: downward large deviations.** Let

$$(1.23) \quad \lambda_N = \|L_N\| = \sup_{\substack{u \in L^2([0,1]) \\ \|u\|_2=1}} \|L_N u\|_2.$$

be the maximal eigenvalue of  $L_N$ , where  $\|\cdot\|_2$  denotes the  $L^2$ -norm. Abbreviate

$$(1.24) \quad C_r = \|\mathcal{L}_r\|.$$

Our goal is to show that  $\lambda_N/N$  satisfies a downward LDP as  $N \rightarrow \infty$ , with rate  $\binom{N}{2}$  and with a rate function that can be analysed in detail. We write  $\mathbb{P}_N^*$  to denote the law of  $\lambda_N$ .

**Theorem 1.5. [Downward LDP]** Subject to (1.5),

$$(1.25) \quad \lim_{N \rightarrow \infty} \binom{N}{2}^{-1} \log \mathbb{P}_N^*(\lambda_N/N \leq \beta) = -\psi_r(\beta), \quad \beta \in [0, C_r],$$

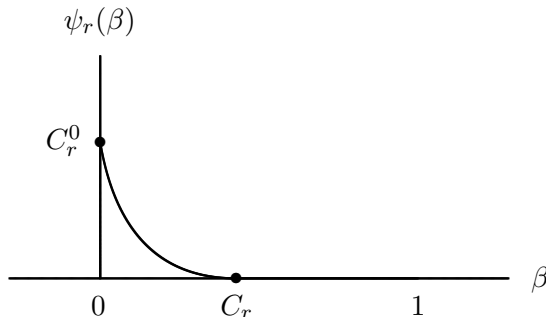
with

$$(1.26) \quad \psi_r(\beta) = \inf_{\substack{\tilde{h} \in \tilde{\mathcal{W}} \\ \|\mathcal{L}_{\tilde{h}}\| \leq \beta}} J_r(\tilde{h}) = \inf_{\substack{h \in \mathcal{W} \\ \|\mathcal{L}_h\| \leq \beta}} I_r(h).$$

The second equality in (1.26) uses that  $\|\mathcal{L}_r\| = \|\mathcal{L}_{r\phi}\|$  for any  $\phi \in \mathcal{M}$ , as is evident after replacing  $u$  by  $u^{\phi^{-1}}$  in (1.21) given by  $u^{\phi^{-1}}(x) = u(\phi^{-1}(x))$ . Since the maximal eigenvalue is invariant under relabelling of the vertices, we need not worry about the equivalence classes.

Let

$$(1.27) \quad C_r^0 = \int_{[0,1]^2} \log \frac{1}{1-r}.$$

FIGURE 1. Graph of  $\beta \mapsto \psi_r(\beta)$ .

When  $\beta = C_r$ , the optimal graphon is the reference graphon  $r$  almost everywhere, for which  $I_r(r) = 0$ , and no large deviation occurs. When  $\beta = 0$ , the optimal graphon is the zero graphon  $\underline{0} \equiv 0$ , for which  $I_r(\underline{0}) = C_r^0$  (see Fig. 1).

**Theorem 1.6. [Properties of the rate function]** *Subject to (1.5):*

- (i)  $\psi_r$  is continuous and strictly decreasing on  $[0, C_r]$ , with  $\psi_r(0) = C_r^0 > 0$  and  $\psi_r(C_r) = 0$ .
- (ii) For every  $\beta \in [0, C_r]$ , the set of minimisers of the variational formula for  $\psi_r(\beta)$  in (1.26) is non-empty and compact in  $\widetilde{\mathcal{W}}$ .

Under the *additional assumptions* that

$$(1.28) \quad \lambda_{\max}(\mathcal{L}_r) < \|d_r\|_{\infty},$$

$$(1.29) \quad r \text{ is bounded away from 0 and 1,}$$

$$(1.30) \quad r \text{ is continuous,}$$

we are able to compute the behaviour of  $\psi_r$  around  $C_r$ . Here  $\lambda_{\max}(\mathcal{L}_r)$  denotes the largest eigenvalue of  $\mathcal{L}_r$ . Note that (1.29) is stronger than (1.5). Recall that  $d_r(x) = \int_{[0,1]} dy r(x, y)$ ,  $x \in [0, 1]$ . It is easy to check that, subject to (1.28),

$$(1.31) \quad C_r = \|d_r\|_{\infty}.$$

**Theorem 1.7. [Scaling of the rate function]** *Subject to (1.28)–(1.29),*

$$(1.32) \quad \psi_r(\beta) \asymp \int_{S_r(\beta)} dx \frac{1}{v_r(x)} (d_r(x) - \beta)^2, \quad \beta \uparrow C_r,$$

where

$$(1.33) \quad S_r(\beta) = \{x \in [0, 1] : d_r(x) \geq \beta\}$$

and

$$(1.34) \quad v_r(x) = \int_{[0,1]} dy r(x, y)[1 - r(x, y)].$$

The proofs of Theorems 1.5–1.7 are given in Sections 3.1–3.3.

**1.5. Main theorems: upward large deviations.** Put

$$(1.35) \quad J_r(x, \beta) = \int_{[0,1]} dy \mathcal{R}(\widehat{r}_{\beta}(x, y) \mid r(x, y)), \quad x \in [0, 1],$$

where

$$(1.36) \quad \widehat{r}_\beta(x, y) = \frac{e^{\theta(x, \beta)} r(x, y)}{e^{\theta(x, \beta)} r(x, y) + [1 - r(x, y)]}$$

is a *Cramér-type transform* of the reference graphon  $r$ , with a Lagrange multiplier function  $\theta(x, \beta)$ ,  $x \in [0, 1]$ , chosen such that  $\int_{[0,1]} dy \widehat{r}_\beta(x, y) = \beta$ ,  $x \in [0, 1]$ . Note that  $\widehat{r}_\beta$  does not need to be symmetric and therefore is not necessarily a graphon. Under the *additional assumption* that

$$(1.37) \quad r \text{ is non-negative definite (i.e., } \mathcal{T}_r \text{ is a non-negative definite operator)}$$

we are able to derive an upward LDP and identify the associated rate function in term of  $J_r$ .

**Theorem 1.8. [Upward LDP]** *Subject to (1.29) and (1.37),*

$$(1.38) \quad \lim_{N \rightarrow \infty} N^{-1} \log \mathbb{P}_N^*(\lambda_N/N \geq \beta) = -\widehat{\psi}_r(\beta), \quad \beta \in [C_r, 1],$$

with

$$(1.39) \quad \widehat{\psi}_r(\beta) = \inf_{x \in [0,1]} J_r(x, \beta).$$

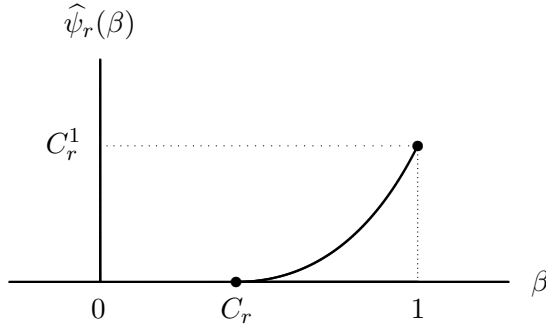


FIGURE 2. Graph of  $\beta \mapsto \widehat{\psi}_r(\beta)$ .

Define

$$(1.40) \quad C_r^1 = \widehat{\psi}_r(1) = \inf_{x \in [0,1]} \int_{[0,1]} dy \log \frac{1}{r(x, y)}.$$

When  $\beta = C_r$ , the Lagrange multiplier in (1.36) is  $\theta(x, C_r) \equiv 0$ , for which  $J_r(x, C_r) \equiv 0$ , and no large deviation occurs. When  $\beta = 1$ , the Lagrange multiplier is  $\theta(x, 1) \equiv \infty$ , for which  $\widehat{\psi}_r(1) = C_r^1$  (see Fig. 2).

**Theorem 1.9. [Properties of the rate function]** *Subject to (1.29),  $\widehat{\psi}_r$  is continuous and strictly increasing on  $[C_r, 1]$ , with  $\widehat{\psi}_r(C_r) = 0$  and  $\widehat{\psi}_r(1) = C_r^1 > 0$ .*

**Theorem 1.10. [Scaling of the rate function]** *Subject to (1.29)–(1.30) and (1.37),*

$$(1.41) \quad \widehat{\psi}_r(\beta) \sim \widehat{K}_r(\beta - C_r)^2, \quad \beta \downarrow C_r,$$

with

$$(1.42) \quad \widehat{K}_r = \frac{1}{2} \inf_{x \in \mathcal{D}_r} \frac{1}{v_r(x)},$$

where  $\mathcal{D}_r = \{x \in [0, 1]: d_r(x) = \|d_r\|_\infty\}$ .

The proofs of Theorems 1.8–1.10 are given in Sections 4.1–4.3.

### 1.6. Discussion.

1. Theorems 1.5–1.7 establish the downward LDP and identify basic properties of the rate function  $\psi_r$ . Note that  $\lim_{\beta \uparrow C_r} S_r(\beta) = \mathcal{D}_r$ . Hence, if  $|\mathcal{D}_r| = 0$ , then the scaling of  $\psi_r$  near  $C_r$  is *faster than quadratic*, which suggests that  $\hat{\lambda}_N/N$  does *not* satisfy a standard central limit theorem. On the other hand, if  $|\mathcal{D}_r| > 0$ , then the scaling is quadratic and a standard central limit is expected to hold. Both questions are open. Several scenarios for the precise scaling are possible depending on how  $d_r$  scales near its maxima.
2. Assumption (1.5) is basic because it underlies Theorem 1.1, which is the jump board for our downward LDP in Theorem 1.5 and the general properties of the downward rate function in Theorem 1.6. Assumption (1.28) is needed for the upper bound in the scaling of the downward rate function in Theorem 1.7. It is the most severe assumption in the present paper, although it is still satisfied for a large class of graphons, for instance, those  $r$  that are rank-1 and continuous (see [17, Proposition 5.11]). Assumption (1.28) guarantees that, for all graphons  $h$  close enough to  $r$  in  $L^2$ -norm,  $\|\mathcal{L}_h\| = \|d_h\|_\infty$ . This allows us to reduce the analysis of the rate function for the LDP of the Laplacian norm to the analysis of the rate function for the maximum degree, which is considerably easier. Assumption (1.29) implies that  $v_r(x) \geq \delta d_r(x)$ ,  $x \in [0, 1]$  for some  $\delta > 0$  (via (1.34)), and ensures that the integral in (1.32) is well-defined.
3. Theorems 1.8–1.10 establish the upward LDP and identify basic properties of the rate function  $\hat{\psi}_r$ . The decay of  $\hat{\psi}_r$  towards zero is *quadratic*, which suggests that  $\hat{\lambda}_N/N$  satisfies a standard central limit theorem. The interpretation of (1.42) is that the curvature of  $\hat{\psi}_r$  at its unique zero  $C_r$  is the *inverse* of the variance of the associated central limit theorem, in line with standard folklore of large deviation theory (see [7]). Since  $v_r(x)$  can be viewed as the variance of the empirical distribution of the degrees of the vertices with label  $\approx xN$  and the large deviations are controlled by  $x \in \mathcal{D}_r$ , the relation in (1.42) is intuitively plausible.
4. Assumption (1.37) is very similar to Assumption (1.28), since it implies  $\lambda_{\max}(\mathcal{L}_r) \leq \|d_r\|_\infty$ . Again, it is needed to make sure that for our upward LDP in Theorem 1.8 *the largest degree is dominant*. In fact, we will see that  $J(x, \beta)$  in (1.35) is the upward rate function for the degrees of the vertices with label  $\approx xN$ . Assumption (1.29) is needed to ensure that the general properties of the upward rate function in Theorem 1.9 hold, Assumptions (1.29)–(1.30) are needed to get the sharp scaling of the upward rate function in Theorem 1.10 as stated in (1.41)–(1.42).
5. Whereas the scaling constant in Theorem 1.10 is sharp, it is not sharp in Theorem 1.7. The proof of Theorem 1.7 shows that it lies between 1 and 2.
6. The fine properties of  $\psi_r$  and  $\hat{\psi}_r$  remain elusive. Neither convexity nor analyticity are obvious.
7. In the proofs we derive an LDP for the maximum degree of  $G_N$  and analyse its rate function. The results we obtain for the maximum degree are analogous to Theorems 1.5–1.10 and are of independent interest. See Remark 3.4.

## 2. PROOF OF PROPOSITION

Section 2.1 shows that the degree operator is lower semi-continuous. Section 2.2 provides the proof of Proposition 1.3.

### 2.1. Lower semi-continuity.

**Lemma 2.1.** *The map  $h \mapsto \|\mathcal{D}_h\|$  is lower semi-continuous in the cut-metric.*



*Proof.* Let  $(\tilde{f}_n)_{n \in \mathbb{N}} \subset \widetilde{\mathcal{W}}$  be a sequence converging to some  $\tilde{f} \in \widetilde{\mathcal{W}}$  in the cut-metric  $\delta_\square$ . Without loss of generality we may assume that it converges in the cut-distance  $d_\square$  as well, i.e.,  $\lim_{n \rightarrow \infty} d_\square(f_n, f) = 0$ . Assume that

$$(2.1) \quad \liminf_{n \rightarrow \infty} \|\mathcal{D}_{f_n}\| < \|\mathcal{D}_f\|.$$

Then there exists a  $\delta > 0$  (independent of  $n$ ) such that  $\|\mathcal{D}_f\| > \|\mathcal{D}_{f_n}\| + \delta$  for infinitely many  $n \in \mathbb{N}$ . By (1.20), there exists a set  $A \subset [0, 1]$  of positive measure such that

$$(2.2) \quad d_f(x) > \|\mathcal{D}_f\| - \frac{\delta}{2} > \|\mathcal{D}_{f_n}\| + \frac{\delta}{2} \geq d_{f_n}(x) + \frac{\delta}{2} \quad \forall x \in A.$$

Hence, by (1.13),

$$(2.3) \quad \int_{A \times [0,1]} dx dy f(x, y) > \int_{A \times [0,1]} dx dy f_n(x, y) + \frac{\delta}{2} \lambda(A)$$

with  $\lambda$  the Lebesgue measure. By the definition of the cut-distance in (1.2),

$$(2.4) \quad d_\square(f_n, f) \geq \int_{A \times [0,1]} dx dy [f(x, y) - f_n(x, y)] > \frac{\delta}{2} \lambda(A).$$

Since  $\lambda(A)$  and  $\delta$  are independent of  $n$ , this inequality is a contradiction with the fact that  $\lim_{n \rightarrow \infty} d_\square(f_n, f) = 0$ . Hence (2.1) is impossible.  $\square$

## 2.2. Proof of Proposition 1.3.

*Proof.* (i) The first part of the statement follows from the fact that  $\mathcal{L}_h$  is the sum of two bounded operators. For the second part of the statement, we need that  $h \mapsto \|\mathcal{T}_h\|$  is continuous and that  $h \mapsto \|\mathcal{D}_h\|$  is lower semi-continuous. The former was shown in [24, Lemma 3.6], the latter was shown in Lemma 2.1.

Let  $g \in L^2([0, 1]^2)$ . Then, by (1.16),

$$(2.5) \quad \begin{aligned} & \liminf_{n \rightarrow \infty} [\|\mathcal{L}_{f_n}(g)\|_2^2 - \|\mathcal{L}_f(g)\|_2^2] \\ & \geq \liminf_{n \rightarrow \infty} [\|\mathcal{D}_{f_n}(g)\|_2^2 - \|\mathcal{D}_f(g)\|_2^2] + \liminf_{n \rightarrow \infty} [\|T_{f_n}(g)\|_2^2 - \|T_f(g)\|_2^2] \\ & \quad - 2 \limsup_{n \rightarrow \infty} [\langle \mathcal{D}_{f_n}(g), T_{f_n}(g) \rangle - \langle \mathcal{D}_f(g), T_f(g) \rangle] \\ & \geq -2 \limsup_{n \rightarrow \infty} [\langle \mathcal{D}_{f_n}(g), T_{f_n}(g) \rangle - \langle \mathcal{D}_f(g), T_f(g) \rangle]. \end{aligned}$$

It remains to show that the last expression equals 0. Since the simple functions are dense in  $L^2([0, 1]^2)$ , we may assume without loss of generality that  $g$  is simple, i.e.,  $g = \sum_{i=1}^k \alpha_i \mathbf{1}_{A_i}$ .

Estimate

$$\begin{aligned}
& \left| \langle \mathcal{D}_{f_n}(g), T_{f_n}(g) \rangle - \langle \mathcal{D}_f(g), T_f(g) \rangle \right| \\
&= \left| \int_{[0,1]^3} dy dy' dx f_n(x, y) f_n(x, y') g(x) g(y) - \int_{[0,1]^3} dy dy' dx f(x, y) f(x, y') g(x) g(y) \right| \\
&= \left| \int_{[0,1]^3} dy dy' dx [f_n(x, y) f_n(x, y') - f(x, y) f(x, y')] g(x) g(y) \right| \\
(2.6) \quad &= \left| \sum_{1 \leq i, j \leq k} \alpha_i \alpha_j \int_{[0,1]^3} dy dy' dx [f_n(x, y) f_n(x, y') - f(x, y) f(x, y')] \mathbf{1}_{A_i}(x) \mathbf{1}_{A_j}(y) \right| \\
&\leq \sum_{1 \leq i, j \leq k} |\alpha_i \alpha_j| \left( \left| \int_{[0,1]^3} dy dx dy' f_n(x, y') [f_n(x, y) - f(x, y)] \mathbf{1}_{A_i}(x) \mathbf{1}_{A_j}(y) \right| \right. \\
&\quad \left. + \left| \int_{[0,1]^3} dy' dx dy f(x, y) [f_n(x, y') - f(x, y')] \mathbf{1}_{A_i}(x) \mathbf{1}_{A_j}(y) \right| \right).
\end{aligned}$$

Note that, for every  $y' \in [0, 1]$ ,  $0 \leq f_n(x, y') \mathbf{1}_{A_i}(x), \mathbf{1}_{A_j}(y) \leq 1$  and, for every  $y \in [0, 1]$ ,  $0 \leq f(x, y) \mathbf{1}_{A_i}(x) \mathbf{1}_{A_j}(y) \leq 1$ . Hence, by the definition of the cut-distance, the above expression is bounded from above by

$$(2.7) \quad \sum_{1 \leq i, j \leq k} |\alpha_i \alpha_j| 2d_{\square}(f_n, f),$$

which tends to zero as  $n \rightarrow \infty$ .

We have thus shown that  $\liminf_{n \rightarrow \infty} \|\mathcal{L}_{f_n}(g)\|_2 \geq \|\mathcal{L}_f(g)\|_2$  for all simple functions  $g \in L^2([0, 1])$ . Now let  $\varepsilon > 0$ . Then there exists a simple  $g' \in L^2([0, 1])$  such that  $\|\mathcal{L}_f\| < \|\mathcal{L}_f(g')\| + \varepsilon$ . By the previous result,

$$(2.8) \quad \liminf_{n \rightarrow \infty} \|\mathcal{L}_{f_n}\| \geq \liminf_{n \rightarrow \infty} \|\mathcal{L}_{f_n}(g')\|_2 \geq \|\mathcal{L}_f(g')\| > \|\mathcal{L}_f\| - \varepsilon.$$

Since this holds for arbitrary  $\varepsilon > 0$ , the proof is complete.

(ii) Let  $G$  be a graph with  $N$  vertices. We prove that  $\frac{1}{N}L_N$  and  $\mathcal{L}_{h^G}$  have the same eigenvalues. First, let  $u = (u_1, \dots, u_N)$  be an eigenvector of  $\frac{1}{N}L_N$  with eigenvalue  $\lambda$ . Let  $u: [0, 1] \rightarrow \mathbb{R}$ , and define  $\bar{u}: [0, 1] \rightarrow \mathbb{R}$  to be the step function that on  $[\frac{i-1}{N}, \frac{i}{N})$ ,  $1 \leq i \leq N$ , equals the average of  $u$  over that interval. Then  $u$  is an eigenfunction of  $\mathcal{L}_{h^G}$  with eigenvalue  $\lambda$  if and only if, for all  $1 \leq j \leq N$  and  $x \in [\frac{j-1}{N}, \frac{j}{N})$ ,

$$(2.9) \quad (\mathcal{L}_{h^G} u)(x) = \frac{1}{N} \sum_{i \in [N] \setminus j} A_N(i, j) u(x) - \frac{1}{N} \sum_{i \in [N] \setminus j} A_N(i, j) \bar{u}_i = \lambda u(x),$$

i.e.,

$$(2.10) \quad u(x) = \left( \frac{1}{N} \sum_{i \in [N] \setminus j} A_N(i, j) - \lambda \right)^{-1} \frac{1}{N} \sum_{i \in [N] \setminus j} A_N(i, j) \bar{u}_i.$$

Similarly,  $v = (v_1, \dots, v_N)$  is an eigenvector of  $\frac{1}{N}L_N$  with eigenvalue  $\lambda$  if and only if

$$(2.11) \quad v_j = \left( \frac{1}{N} \sum_{i \in [N] \setminus j} A_N(i, j) - \lambda \right)^{-1} \frac{1}{N} \sum_{i \in [N] \setminus j} A_N(i, j) v_i.$$

Since every eigenfunction of  $\mathcal{L}_{hG}$  is constant on the blocks  $[\frac{j-1}{N}, \frac{j}{N})$ ,  $1 \leq j \leq N$ , we can view each eigenfunction of  $\mathcal{L}_{hG}$  as an eigenvector of  $\frac{1}{N}L_N$ , and vice versa.

Finally, note that  $\mathcal{L}_{hG}$  is a normal operator and has a non-negative spectrum. So the norm of  $\mathcal{L}_{hG}$  equals the supremum of its spectrum. Furthermore, the essential spectrum of  $\mathcal{L}_{hG}$  equals the range of the degree function  $d_{hG}(x)$  [17, Proposition 5.11]. Since the largest eigenvalue of the Laplacian matrix is bounded from below by the maximum degree [19], it follows that the largest eigenvalue of  $\mathcal{L}_{hG_N}$  equals its norm.  $\square$

### 3. PROOF OF THEOREMS FOR DOWNWARD LDP

Sections 3.1–3.3 provide the proof of Theorems 1.5–1.7, respectively.

#### 3.1. Proof of Theorem 1.5.

3.1.1. *Upper bound.* By Proposition 1.3(i),  $h \mapsto \|\mathcal{L}_h\|$  is lower semi-continuous. Hence the set  $\{\tilde{h} \in \widetilde{\mathcal{W}}: \|\mathcal{L}_{\tilde{h}}\| \leq \beta\}$  is closed in  $\widetilde{\mathcal{W}}$ . By Proposition 1.3(ii),  $\frac{\|L_N\|}{N} = \|\mathcal{L}_{h^{G_N}}\|$ , and so we can use Theorem 1.1 in combination with the contraction principle [15] to get

$$(3.1) \quad \limsup_{N \rightarrow \infty} \binom{N}{2}^{-1} \log \mathbb{P}_N \left( \frac{\|L_N\|}{N} \leq \beta \right) \leq - \inf_{\substack{\tilde{h} \in \widetilde{\mathcal{W}} \\ \|\mathcal{L}_{\tilde{h}}\| \leq \beta}} J_r(\tilde{h}) = - \inf_{\substack{h \in \mathcal{W} \\ \|\mathcal{L}_h\| \leq \beta}} I_r(h).$$

3.1.2. *Lower bound.* The proof proceeds via a change of measure. The key is the following lemma.

**Lemma 3.1.** *Let  $h^{G_N}$  be the empirical graphon corresponding to the inhomogeneous Erdős-Rényi random graph  $G_N$ . Then  $\lim_{N \rightarrow \infty} \|\mathcal{L}_{h^{G_N}}\| = \|\mathcal{L}_r\|$  in probability.*

*Proof.* Since  $\mathcal{L}_h$  is the sum of  $\mathcal{D}_h$  and  $-\mathcal{T}_h$  (recall (1.16)), it suffices to prove the claim for these two operators separately.

To show that  $\mathcal{D}_{h^{G_N}}$  converges to  $\mathcal{D}_r$  in operator norm in probability, let  $d_{r_N}$  be the degree function of the block graphon  $r_N$ . Then

$$(3.2) \quad \|d_{h^{G_N}} - d_r\|_\infty \leq \|d_{h^{G_N}} - d_{r_N}\|_\infty + \|d_{r_N} - d_r\|_\infty.$$

Because  $\|r_N - r\|_\infty \downarrow 0$  as  $N \rightarrow \infty$  by (1.4), the second term vanishes. As for the first term,

$$(3.3) \quad \begin{aligned} & \mathbb{P}_N(\|d_{h^{G_N}} - d_{r_N}\|_\infty \geq t) \\ & \leq \sum_{i \in [N]} \mathbb{P}_N \left( \left| \frac{1}{N} \sum_{i \in [N] \setminus j} A_N(i, j) - \frac{1}{N} \sum_{i \in [N] \setminus j} r_{N, ij} \right| \geq t \right) \leq 2Ne^{-2Nt^2} \downarrow 0 \end{aligned}$$

by a straightforward application of Hoeffding's inequality. We use that  $A_N(i, j)$  has Bernoulli distribution with mean  $(r_N)_{ij}$ . Hence  $\|d_{h^{G_N}} - d_{r_N}\|_\infty \rightarrow 0$  as  $N \rightarrow \infty$  in probability, and so

$$(3.4) \quad \|\mathcal{D}_{h^{G_N}} - \mathcal{D}_r\| = \|d_{h^{G_N}} - d_r\|_\infty \rightarrow 0 \text{ in probability.}$$

To show that  $\mathcal{T}_{h^{G_N}}$  converges to  $\mathcal{T}_r$  in operator norm in probability, it suffices to note that

$$(3.5) \quad \|\mathcal{T}_{h^{G_N}} - \mathcal{T}_r\| \leq \sqrt{2} d_{\square}(h^{G_N}, r) \downarrow 0 \quad \text{as } N \rightarrow \infty \text{ in probability.}$$

The first inequality was shown in [24, Lemma 3.6]. Convergence in the cut-metric was shown in [13, Lemma 5.11].  $\square$

We are now ready to prove the lower bound.

**Proposition 3.2.** *Let  $\beta \in [0, C_r]$ , and let  $h \in \mathcal{W}$  be such that  $\|\mathcal{L}_h\| < \beta$ . Then*

$$(3.6) \quad \liminf_{N \rightarrow \infty} \frac{2}{N^2} \log \mathbb{P}_N(\|\mathcal{L}_{h^{G_N}}\| < \beta) \geq -I_r(h).$$

*Proof.* The proof comes in three steps.

**1.** We begin by giving the proof for the case when  $h$  is a *continuous* graphon. Denote the law of an inhomogeneous ERRG with reference graphon  $r_N$  by  $\mathbb{P}_{N,r_N}$  instead of  $\mathbb{P}_N$ , and the law of an inhomogeneous ERRG with reference graphon  $h_N$  by  $\mathbb{P}_{N,h_N}$ , where  $h_N$  is the block graphon that is obtained by averaging  $h$  over blocks of size  $1/N$ . Let  $x = \|\mathcal{L}_h\|$ , and define

$$(3.7) \quad U_\varepsilon^x = \{f \in \mathcal{W} \mid \|\mathcal{L}_f\| \in (x - \varepsilon, x + \varepsilon)\}.$$

Note that  $h$  satisfies (1.4) by uniform continuity. Hence, by Lemma 3.1,

$$(3.8) \quad \mathbb{P}_{N,h_N}(U_\varepsilon^x) \uparrow 1, \quad N \rightarrow \infty,$$

for all  $\varepsilon > 0$ . Write

$$(3.9) \quad \mathbb{P}_{N,r_N}(U_\varepsilon^x) = \mathbb{P}_{N,h_N}(U_\varepsilon^x) \frac{1}{\mathbb{P}_{N,h_N}(U_\varepsilon^x)} \int_{U_\varepsilon^x} \exp\left(-\log \frac{d\mathbb{P}_{N,h_N}}{d\mathbb{P}_{N,r_N}}\right) d\mathbb{P}_{N,h_N}.$$

By Jensen's inequality,

$$(3.10) \quad \log \mathbb{P}_{N,r_N}(U_\varepsilon^x) \geq \log \mathbb{P}_{N,h_N}(U_\varepsilon^x) - \frac{1}{\mathbb{P}_{N,h_N}(U_\varepsilon^x)} \int_{U_\varepsilon^x} \left(\log \frac{d\mathbb{P}_{N,h_N}}{d\mathbb{P}_{N,r_N}}\right) d\mathbb{P}_{N,h_N}.$$

Using (3.8) and [13, Lemma 5.7], we obtain

$$(3.11) \quad \liminf_{N \rightarrow \infty} \frac{2}{N^2} \log \mathbb{P}_{N,r_N}(U_\varepsilon^x) \geq -\lim_{N \rightarrow \infty} \frac{2}{N^2} \int_{U_\varepsilon^x} \left(\log \frac{d\mathbb{P}_{N,h_N}}{d\mathbb{P}_{N,r_N}}\right) d\mathbb{P}_{N,h_N} = -I_r(h).$$

Note that [13, Lemma 5.7] was stated for the homogeneous ERRG, but the proof may be extended to the inhomogeneous ERRG (see [26, Theorem 6.1]). Because  $U_\varepsilon^x \subseteq \{f \in \mathcal{W} \mid \|\mathcal{L}_f\| < \beta\}$  for  $\varepsilon > 0$  small enough, we conclude that

$$(3.12) \quad \liminf_{N \rightarrow \infty} \frac{2}{N^2} \log \mathbb{P}_{N,r_N}(\|\mathcal{L}_{h^{G_N}}\| < \beta) \geq -I_r(x).$$

**2.** We next extend the proof to the case where  $h$  is a *block graphon* such that  $\|\mathcal{L}_h\| < \beta$ . Assume that  $h$  is constant on the blocks  $[\frac{i-1}{N}, \frac{i}{N}) \times [\frac{j-1}{N}, \frac{j}{N})$ ,  $1 \leq i, j \leq N$ . Then there exists a sequence of continuous graphons  $(h_k)_{k \in \mathbb{N}}$  that converges to  $h$  in  $L^2$  with  $h_k \leq h$ . Note that, for  $f \in \mathcal{W}$ ,

$$(3.13) \quad \begin{aligned} \|\mathcal{L}_f\| &= \sup_{\|u\|_2 \leq 1} \langle g, \mathcal{L}_f(u) \rangle = \sup_{\|u\|_2 \leq 1} \int_{[0,1]^2} dy dx u(x) f(x, y) [u(x) - u(y)] \\ &= \frac{1}{2} \sup_{\|u\|_2 \leq 1} \left[ \int_{[0,1]^2} dy dx u(x) f(x, y) [u(x) - u(y)] + \int_{[0,1]^2} dy dx u(y) f(x, y) [u(y) - u(x)] \right] \\ &= \frac{1}{2} \sup_{\|u\|_2 \leq 1} \int_{[0,1]^2} dx dy f(x, y) [u(x) - u(y)]^2, \end{aligned}$$

where we use that  $f$  is a symmetric function. From the above formula it is immediate that, since  $h_k \leq h$ , we have  $\|\mathcal{L}_{h_k}\| \leq \|\mathcal{L}_h\| < \beta$ . Then, by Lemma 3.2,

$$(3.14) \quad \liminf_{N \rightarrow \infty} \frac{2}{N^2} \log \mathbb{P}_N(\|\mathcal{L}_{h^{G_N}}\| < \beta) \geq -I_r(h_k), \quad k \geq k_0.$$

We conclude by noting that  $I_r$  is continuous in  $L^2([0, 1]^2)$ , implying that  $I_r(h_k)$  converges to  $I_r(h)$  as  $k \rightarrow \infty$ .

**3.** We finally extend the proof to the case when  $h$  is an *arbitrary* graphon such that  $\|\mathcal{L}_h\| < \beta$ . Let  $\bar{h}_N$  be the  $N$ -block approximant of  $h$  for some  $N$ . Then  $\bar{h}_N$  converges to  $h$  in  $L^2$  as  $N \rightarrow \infty$  [13, Proposition 2.6]. Again, it suffices to prove

$$(3.15) \quad \|\mathcal{L}_{\bar{h}_N}\| \leq \|\mathcal{L}_h\|.$$

First, suppose that  $\|\mathcal{L}_{\bar{h}_N}\| = \|d_{\bar{h}_N}\|_\infty$ . For  $x \in [\frac{i-1}{N}, \frac{i}{N})$ ,

$$(3.16) \quad d_{\bar{h}_N}(x) = N \int_{[\frac{i-1}{N}, \frac{i}{N})} dx h(x) \leq \|d_h\|_\infty.$$

It follows that  $\|\mathcal{L}_{\bar{h}_N}\| = \|d_{\bar{h}_N}\|_\infty \leq \|d_h\|_\infty \leq \|\mathcal{L}_h\|$ . Next, suppose that  $\|\mathcal{L}_{\bar{h}_N}\| > \|d_{\bar{h}_N}\|_\infty$ . Then  $\|\mathcal{L}_{\bar{h}_N}\| = \lambda_{\max}(\mathcal{L}_{\bar{h}_N})$  and there exists an eigenfunction  $u$  of  $\mathcal{L}_h$  with  $\|u\|_2 = 1$  such that  $\langle \mathcal{L}_{\bar{h}_N} u, u \rangle = \|\mathcal{L}_{\bar{h}_N}\|$ . In the proof of Proposition 1.3 it was shown that  $u$  is constant on each of the intervals  $[\frac{i-1}{N}, \frac{i}{N})$ ,  $1 \leq i \leq N$ . Let  $u_i$  be the value of  $u$  on  $[\frac{i-1}{N}, \frac{i}{N})$ ,  $1 \leq i \leq N$ . Then

$$(3.17) \quad \begin{aligned} \langle \mathcal{L}_h u, u \rangle &= \int_{[0,1]^2} dx dy h(x, y) [u(x) - u(y)] u(x) \\ &= \sum_{i,j=1}^N \int_{[\frac{i-1}{N}, \frac{i}{N}) \times [\frac{j-1}{N}, \frac{j}{N})} dx dy h(x, y) [u_i - u_j] u_i \\ &= \sum_{i,j=1}^N \int_{[\frac{i-1}{N}, \frac{i}{N}) \times [\frac{j-1}{N}, \frac{j}{N})} dx dy \bar{h}_N(x, y) [u_i - u_j] u_i = \langle \mathcal{L}_{\bar{h}_N} u, u \rangle. \end{aligned}$$

The desired result now follows from the fact that  $\|\mathcal{L}_h\| = \sup_{\|v\|_2=1} \langle \mathcal{L}_h v, v \rangle \geq \langle \mathcal{L}_h u, u \rangle = \langle \mathcal{L}_{\bar{h}_N} u, u \rangle = \|\mathcal{L}_{\bar{h}_N}\|$ .  $\square$

From Proposition 3.2, we obtain the lower bound

$$(3.18) \quad \liminf_{N \rightarrow \infty} \frac{2}{N^2} \log \mathbb{P}_N(\|\mathcal{L}_h^{G_N}\| < \beta) \geq - \inf_{\substack{h \in \mathcal{W} \\ \|\mathcal{L}_h\| < \beta}} I_r(h).$$

So it just remains to show

$$(3.19) \quad \inf_{\substack{h \in \mathcal{W} \\ \|\mathcal{L}_h\| < \beta}} I_r(h) = \inf_{\substack{h \in \mathcal{W} \\ \|\mathcal{L}_h\| \leq \beta}} I_r(h).$$

Let  $h \in \mathcal{W}$  with  $\|\mathcal{L}_h\| = \beta$ . Then  $\|\mathcal{L}_{(1-\varepsilon)h}\| = (1-\varepsilon)\beta < \beta$  for all  $\varepsilon > 0$ . Furthermore,  $(1-\varepsilon)h$  converges to  $h$  in  $L^2$  as  $\varepsilon \downarrow 0$  and  $I_r$  is continuous in  $L^2$ , so  $I_r((1-\varepsilon)h) \rightarrow I_r(h)$  as  $\varepsilon \downarrow 0$ . Hence, the left-hand side of (3.19) is bounded from above by the right-hand side. The converse inequality is trivial, which settles the proof of the lower bound.

### 3.2. Proof of Theorem 1.6.

*Proof.* (i) It was shown in the proof of Proposition 3.2 that if  $h_1 \leq h_2$ , then  $\|\mathcal{L}_{h_1}\| \leq \|\mathcal{L}_{h_2}\|$ . Via the same proof as for [11, Theorem 1.5], it follows that  $\psi_r$  is strictly decreasing on  $[0, C_r]$ .

Let  $t > 0$ , and let  $(t_n)_{n \in \mathbb{N}}$  be a strictly increasing sequence converging to  $t$ . By the compactness of  $\widetilde{\mathcal{W}}$ , the lower semi-continuity of  $I_r$  and  $\tilde{h} \mapsto \|\mathcal{L}_{\tilde{h}}\|$ , and the invariance of  $\tilde{h} \mapsto \|\mathcal{L}_{\tilde{h}}\|$  under measure-preserving bijections, there exists a sequence of graphons  $(\tilde{f}_n)_{n \in \mathbb{N}}$  such that  $\|\mathcal{L}_{\tilde{f}_n}\| \geq t_n$  and  $\psi_r(t_n) = I_r(\tilde{f}_n)$  for all  $n \in \mathbb{N}$ . Again, by the compactness of  $\widetilde{\mathcal{W}}$ , we may assume that the sequence  $(\tilde{f}_n)_{n \in \mathbb{N}}$  converges to some  $\tilde{f} \in \widetilde{\mathcal{W}}$  in the cut-metric. Because  $\tilde{h} \mapsto \|\mathcal{L}_{\tilde{h}}\|$  is lower

semi-continuous, we have  $\|\mathcal{L}_{\tilde{f}}\| \geq t$ . Once again invoking the lower semi-continuity of  $I_r$ , we conclude that

$$(3.20) \quad \liminf_{n \rightarrow \infty} \psi_r(t_n) = \liminf_{n \rightarrow \infty} I_r(\tilde{f}_n) \geq I_r(\tilde{f}) \geq \psi_r(t).$$

Right-continuity of  $\psi_r$  now follows by noting that  $\psi_r$  is strictly decreasing. The proof of the left-continuity of  $\psi_r$  is the same as the proof of (3.19).

(ii) This statement is immediate from the lower semi-continuity of  $h \mapsto \|\mathcal{L}_h\|$  and  $I_r$ , combined with the compactness of  $\widetilde{\mathcal{W}}$ .  $\square$

### 3.3. Proof of Theorem 1.7.

*Proof.* We first prove a large deviation principle for the degree of a single vertex  $\approx xN$  with rate  $N$  and with rate function  $J_r(x, \beta)$  defined in (1.35). We subsequently derive upper and lower bounds for  $\psi_r(\beta)$  in terms of this rate function, and use these to compute the scaling of  $\psi_r$  near its minimum.

3.3.1. *LDP for single degrees.* Let  $d_i$  be the degree of vertex  $i$ . For  $x \in [0, 1]$ , let  $i_x \in \{1, \dots, N\}$  be the index such that  $x \in [\frac{i_x-1}{N}, \frac{i_x}{N})$ . A crucial ingredient is the LDP for the family  $N^{-1}d_{i_x}$ ,  $N \in \mathbb{N}$ , for fixed  $x \in [0, 1]$ , which we prove with the help of the Gärtner-Ellis theorem. To that end, note that  $d_{i_x} = \sum_{j \in [N]} A_{i_x j}$ , with  $A_{i_x j} = 1$  if there is an edge between  $i$  and  $j$  and  $A_{i_x j} = 0$  otherwise. Therefore the cumulant generating function of  $N^{-1}d_{i_x}$  is

$$(3.21) \quad \begin{aligned} \Lambda_N(x, \theta) &= \log \mathbb{E} \left[ \exp \left( \theta \frac{1}{N} \sum_{j \in [N]} A_{i_x j} \right) \right] = \sum_{j \in [N] \setminus i} \log \mathbb{E} \left[ \exp \left( \theta \frac{1}{N} A_{i_x j} \right) \right] \\ &= \sum_{j \in [N]} \log \left( (r_N)_{i_x j} e^{\frac{1}{N}\theta} + [1 - (r_N)_{i_x j}] \right) \\ &= N \int_{[0,1] \setminus [\frac{i_x-1}{N}, \frac{i_x}{N})} dy \log \left( r_N(x, y) e^{\frac{1}{N}\theta} + [1 - r_N(x, y)] \right). \end{aligned}$$

Since  $\|r_N - r\|_\infty \rightarrow 0$ , we have

$$(3.22) \quad \Lambda_r(x, \theta) = \lim_{N \rightarrow \infty} \frac{1}{N} \Lambda_N(x, N\theta) = \int_{[0,1]} dy \log \left( r(x, y) e^\theta + [1 - r(x, y)] \right).$$

Since  $\theta \mapsto \Lambda_r(x, \theta)$  is finite and differentiable on  $\mathbb{R}$ , the Gärtner-Ellis theorem tells us that the family  $N^{-1}d_{i_x}$ ,  $N \in \mathbb{N}$ , satisfies the LDP on  $[0, 1]$  with rate  $N$  and with rate function  $x \mapsto J_r(x, \beta)$  given by

$$(3.23) \quad J_r(x, \beta) = \sup_{\theta \in \mathbb{R}} [\theta \beta - \Lambda_r(x, \theta)].$$

The supremum is attained at the unique  $\theta(x, \beta)$  such that

$$(3.24) \quad \int_{[0,1]} dy \widehat{r}_\beta(x, y) = \beta \quad \forall x \in [0, 1],$$

with

$$(3.25) \quad \widehat{r}_\beta(x, y) = \frac{e^{\theta(x, \beta)} r(x, y)}{e^{\theta(x, \beta)} r(x, y) + [1 - r(x, y)]}.$$

The Lagrange multiplier  $\theta(x, \beta)$  exists because  $r \in (0, 1)$  almost everywhere by (1.5). A simple computation shows that

$$(3.26) \quad J_r(x, \beta) = \int_{[0,1]} dy \mathcal{R}(\widehat{r}_\beta(x, y) \mid r(x, y)) = \sup_u \int_{[0,1]} dy \mathcal{R}(u(y) \mid r(x, y)),$$

where the supremum is taken over all the measurable functions  $u: [0, 1] \rightarrow [0, 1]$  satisfying  $\int_{[0,1]} dy u(y) \leq \beta$ . The map

$$\theta \mapsto \int_{[0,1]} dy \frac{e^\theta r(x, y)}{e^\theta r(x, y) + 1 - r(x, y)}$$

is a continuous bijection from  $\mathbb{R}$  to  $(0, 1)$ , again because  $r \in (0, 1)$  almost everywhere, and so the inverse map  $\beta \mapsto \theta(x, \beta)$  is continuous.

**3.3.2. Properties of  $J_r$ .** We need the following lemma for the derivatives of  $J_r$  with respect to  $\beta$ . Henceforth we write  $J'_r = \frac{\partial J_r}{\partial \beta}$  and  $\theta' = \frac{\partial \theta}{\partial \beta}$ , and use the upper index ( $k$ ) for the  $k$ -th derivative.

**Lemma 3.3.**  $\beta \mapsto J_r(x, \beta)$  is analytic for every  $x \in [0, 1]$ , with

$$(3.27) \quad J_r^{(k)}(x, \beta) = \theta^{(k-1)}(x, \beta).$$

Moreover, subject to (1.29),

$$(3.28) \quad \sup_{\substack{x \in [0,1] \\ \beta \in [\varepsilon, 1-\varepsilon]}} |J_r^{(k)}(x, \beta)| < \infty \quad \forall \varepsilon > 0.$$

*Proof.* Recall from Section 3.3.1 that

$$(3.29) \quad J_r(x, \beta) = \int_{[0,1]} dy \mathcal{R}(\widehat{r}_\beta(x, y) \mid r(x, y)) = \theta(x, \beta)\beta - \int_{[0,1]} dy \log \left( 1 + (e^{\theta(x, \beta)} - 1)r(x, y) \right).$$

Differentiating with respect to  $\beta$  and using that  $d\widehat{r}_\beta(x) = \beta$ , we obtain

$$(3.30) \quad J'_r(x, \beta) = \theta(x, \beta).$$

Implicit differentiation of the equation

$$(3.31) \quad \int_{[0,1]} dy \frac{e^\theta r(x, y)}{e^\theta r(x, y) + (1 - r(x, y))} = \beta$$

gives

$$(3.32) \quad \theta'(x, \beta) = \left( \int_{[0,1]} dy \frac{e^{\theta(x, \beta)} r(x, y) (1 - r(x, y))}{[e^{\theta(x, \beta)} r(x, y) + (1 - r(x, y))]^2} \right)^{-1}.$$

Fix  $\varepsilon > 0$ . Since  $r$  is bounded away from 0 and 1, we see that  $\theta'(x, \beta)$  is bounded for  $x \in [0, 1]$  and  $\beta \in [C_r, 1 - \varepsilon]$  when  $\theta(x, \beta)$  is. Iteratively applying the chain, product and quotient rules of differentiation, we obtain that  $\theta^{(k)}$  is some polynomial of order  $k$  in the variables

$$(3.33) \quad \left( \int_{[0,1]} dy \frac{e^\theta r(x, y) (1 - r(x, y))}{[e^\theta r(x, y) + (1 - r(x, y))]^2} \right)^{-1}, \quad \int_{[0,1]} dy \frac{f(e^\theta, r, \theta', \dots, \theta^{(k-1)})}{[e^\theta r(x, y) + (1 - r(x, y))]^j},$$

with  $f$  a polynomial and  $j \in \mathbb{N}$ . By induction, we obtain that  $f$  is bounded when  $\theta$  is bounded, and hence that  $\theta^{(k)}$  is bounded.

Thus, it remains to show  $\sup_{x \in [0,1], \beta \in [\varepsilon, 1-\varepsilon]} |\theta(x, \beta)| < \infty$ . Let

$$(3.34) \quad r_- = \inf_{(x,y) \in [0,1]^2} r(x, y) > 0, \quad \tilde{\theta}(\beta) = \log \frac{\beta(1 - r_-)}{(1 - \beta)r_-}.$$

Since the map  $r \mapsto \frac{e^{\theta r}}{e^{\theta r} + (1-r)}$  is increasing, we have

$$(3.35) \quad \int_{[0,1]} dy \frac{e^{\tilde{\theta} r(x,y)}}{e^{\tilde{\theta} r(x,y)} + (1-r(x,y))} \geq \frac{e^{\tilde{\theta} r_-}}{e^{\tilde{\theta} r_-} + (1-r_-)} = \beta.$$

Hence,  $\theta(x, \beta) \leq \tilde{\theta}(\beta) < \infty$ . Since  $\tilde{\theta}(\beta)$  is independent of  $x$  and  $\sup_{\beta \in [\varepsilon, 1-\varepsilon]} |\tilde{\theta}(\beta)| < \infty$ , all claims are settled.  $\square$

3.3.3. *Lower bound.* Let  $\bar{h}$  be a minimizer of (1.26). Then  $\int_{[0,1]} dy \bar{h}(x, y) \leq \|\mathcal{L}_{\bar{h}}\| \leq \beta$  for each  $x \in S_r(\beta)$ . Hence, by (3.26),

$$(3.36) \quad \int_{[0,1]} dy \mathcal{R}(\bar{h}(x, y) \mid r(x, y)) \geq J_r(x, \beta),$$

which implies

$$(3.37) \quad \psi_r(\beta) = I_r(\bar{h}) \geq \int_{S_r(\beta)} dx J_r(x, \beta).$$

3.3.4. *Upper bound.* Let

$$(3.38) \quad h_\beta(x, y) = \begin{cases} r(x, y), & x, y \notin S_r(\beta) \\ \hat{r}_\beta(x, y), & x \in S_r(\beta) \text{ or } y \in S_r(\beta), \text{ but } (x, y) \notin S_r(\beta)^2, \\ \min\{\hat{r}_\beta(x, y), \hat{r}_\beta(y, x)\}, & x, y \in S_r(\beta). \end{cases}$$

Then  $h_\beta$  converges to  $r$  in  $L^\infty$  as  $\beta \uparrow C_r$ . We show that  $\|\mathcal{L}_{h_\beta}\| \leq \beta$  for  $\beta$  close enough to  $C_r$ . Note that  $\|d_{h_\beta}\|_\infty \leq \beta$  by construction, so we only need to show  $\lambda_{\max}(\mathcal{L}_{h_\beta}) \leq \beta$  for  $\beta$  large enough. Let  $\lambda_\beta = \lambda_{\max}(\mathcal{L}_{h_\beta})$  and let  $\lambda'$  be the limit of any convergent subsequence of  $(\lambda_\beta)_{k \in \mathbb{N}}$ . We show that  $\lambda'$  is an eigenvalue of  $\mathcal{L}_h$ , so that  $\limsup_{\beta \uparrow C_r} \lambda_\beta \leq \lambda_{\max}(\mathcal{L}_r) \leq \beta$  for  $\beta$  close enough to  $C_r$ . Here we use Assumption (1.28).

Assume, by contradiction, that  $\lambda'$  is not an eigenvalue of  $\mathcal{L}_r$ . Without loss of generality, we may assume that  $\lambda_\beta$  converges to  $\lambda'$ . Consider  $F: L^\infty([0, 1]^2) \times \mathbb{R} \times L^2([0, 1]) \rightarrow L^2([0, 1])$  given by  $F(g, \mu, u) = \mathcal{L}_g u - \mu u$ . This map is bounded and affine in each coordinate, and hence is continuously Fréchet differentiable. The Fréchet derivative of  $F$  at a point  $(g, \mu, u)$  is given by

$$(3.39) \quad ((DF)(g, \mu, u))(f, \nu, w) = F(f, \nu, u) + F(g, \mu, w).$$

Indeed, let  $(f_k, \nu_k, w_k)$  such that  $\|(f_k, \nu_k, w_k)\| = \|f_k\|_\infty + |\nu_k| + \|w_k\|_2 \rightarrow 0$  as  $k \rightarrow \infty$ . Then

$$(3.40) \quad \frac{\|F(g + f_k, \mu + \nu_k, u + w_k) - F(g, \mu, u) - ((DF)(g, \mu, u))(f_k, \nu_k, w_k)\|}{\|(f_k, \nu_k, w_k)\|} \\ = \frac{\|F(f_k, \nu_k, w_k)\|}{\|(f_k, \nu_k, w_k)\|} \leq \frac{\|\mathcal{L}_{f_k} w_k\|_2 + \|\nu_k w_k\|_2}{\|f_k\|_\infty + |\nu_k| + \|w_k\|_2} \leq \frac{2\|f_k\|_\infty \|w_k\|_2 + |\nu_k| \|w_k\|_2}{\|f_k\|_\infty + |\nu_k| + \|w_k\|_2} \rightarrow 0, \quad k \rightarrow \infty.$$

Note that  $F$  is not necessarily Fréchet differentiable as a function  $L^2([0, 1]^2) \times \mathbb{R} \times L^2([0, 1]) \rightarrow L^2([0, 1])$ . Since  $\lambda'$  is not an eigenvalue of  $\mathcal{L}_r$ , the map

$$(3.41) \quad w \mapsto ((DF)(r, \lambda', 0))(0, 0, w) = F(r, \lambda', w)$$

has a trivial kernel and so is an isomorphism on its image space. So, by the implicit function theorem for Banach spaces [22, Theorem I.5.9], there exists a neighbourhood  $U$  of  $(r, \lambda') \in L^\infty([0, 1]^2) \times \mathbb{R}$  and a neighbourhood  $V$  of  $0 \in L^2([0, 1])$  such that  $F(g, \mu, u) = 0$  if and only if  $u = 0$  for all  $(g, \mu, u) \in U \times V$ . Since  $h_\beta \rightarrow r$  in  $L^\infty$  and  $\lambda_\beta \rightarrow \lambda'$  as  $\beta \uparrow C_r$ , we have that  $(h_\beta, \lambda_\beta) \in U$  for  $\beta \geq \beta_0$ . However, since  $\lambda_\beta$  is an eigenvalue of  $\mathcal{L}_{h_\beta}$ , there exists an eigenfunction  $u_\beta \in L^2([0, 1])$  with  $u_\beta \neq 0$  such that  $F(h_\beta, \lambda_\beta, u_\beta) = 0$ . Since  $V$  is a neighbourhood of  $0$ , we



can rescale  $u_\beta$  such that it lies in  $V$ . This yields a contradiction, and so  $\lambda'$  must be an eigenvalue of  $\mathcal{L}_r$ .

Now since  $\|\mathcal{L}_{h_\beta}\| \leq \beta$  for  $\beta$  sufficiently close to  $C_r$ , we have

$$\begin{aligned}
(3.42) \quad \psi_r(\beta) &\leq I_r(h_\beta) \\
&= 2 \int_{S_r(\beta) \times ([0,1] \setminus S_r(\beta))} dx dy \mathcal{R}(\hat{r}_\beta(x, y) \mid r(x, y)) \\
&\quad + \int_{S_r(\beta)^2} dx dy \mathcal{R}(\min\{\hat{r}_\beta(x, y), \hat{r}_\beta(y, x)\} \mid r(x, y)) \\
&\leq 2 \int_{S_r(\beta) \times ([0,1] \setminus S_r(\beta))} dx dy \mathcal{R}(\hat{r}_\beta(x, y) \mid r(x, y)) + \int_{S_r(\beta)^2} dx dy \mathcal{R}(\hat{r}_\beta(x, y) \mid r(x, y)) \\
&\quad + \int_{S_r(\beta)^2} dx dy \mathcal{R}(\hat{r}_\beta(y, x) \mid r(x, y)) \\
&= 2 \int_{S_r(\beta)} dx J_r(x, \beta), \quad \beta \uparrow C_r.
\end{aligned}$$

For the last equality, we use that  $r$  is symmetric and that  $S_r(\beta)^2$  is a symmetric domain.

**Remark 3.4.** The above shows that, up to a constant, the last integral in (3.42) also equals the rate function in the LDP for the maximum degree. The upper bound can also be shown directly by the following computation.

The state space for the edges determining the adjacency matrix  $(A_{ij})$  is  $\{0, 1\}^{\binom{N}{2}}$ , which we endow with the standard partial ordering. The probability distribution of the edges is the product measure  $\prod_{\{i,j\}} \text{BER}(r_N(\frac{i}{N}, \frac{j}{N}))$ , which is log-convex. Note that the events  $\{N^{-1} \max_{1 \leq i \leq N-1} d_i \leq \beta\}$  and  $\{N^{-1} d_N \leq \beta\}$  are non-increasing in the partial ordering. Hence we can use the FKG-inequality iteratively, to get

$$(3.43) \quad \mathbb{P}\left(N^{-1} \max_{i \in [N]} d_i \leq \beta\right) \geq \prod_{i \in [N]} \mathbb{P}(N^{-1} d_i \leq \beta).$$

Consequently,

$$\begin{aligned}
(3.44) \quad &\frac{1}{\binom{N}{2}} \log \mathbb{P}\left(N^{-1} \max_{i \in [N]} d_i \leq \beta\right) \geq \frac{1}{\binom{N}{2}} \sum_{i=1}^N \log \mathbb{P}(N^{-1} d_i \leq \beta) \\
&= \frac{2N}{N-1} \frac{1}{N} \sum_{i \in [N]} \log \mathbb{P}(N^{-1} d_i \leq \beta) = \frac{2N}{N-1} \int_{[0,1]} dx \frac{1}{N} \log \mathbb{P}(N^{-1} d_{i_x} \leq \beta).
\end{aligned}$$

By the law of large numbers,  $\lim_{N \rightarrow \infty} N^{-1} d_{i_x} = \int_{[0,1]} dy r(x, y) = d_r(x)$   $\mathbb{P}$ -a.s., and so in the limit as  $N \rightarrow \infty$  the last integral in (3.44) may be restricted to the set  $S_r(\beta) = \{x \in [0, 1]: d_r(x) \geq \beta\}$ , i.e.,

$$\liminf_{N \rightarrow \infty} \frac{1}{\binom{N}{2}} \log \mathbb{P}\left(N^{-1} \max_{i \in [N]} d_i \leq \beta\right) \geq -2 \int_{S_r(\beta)} dx J_r(x, \beta),$$

where we use that the family  $N^{-1} d_{i_x}$ ,  $N \in \mathbb{N}$ , satisfies the LDP on  $[0, 1]$  with rate  $N$  and with rate function  $x \mapsto J_r(x, \beta)$ , as shown in Section 3.3.1.  $\spadesuit$

3.3.5. *Scaling of  $J_r(x, \beta)$ .* Via Lemma 3.3 it is straightforward to show that  $J_r(x, d_r(x)) = J'_r(x, d_r(x)) = 0$  and  $J''_r(x, d_r(x)) = \frac{1}{v_r(x)}$ . So, by Taylor expansion,

$$(3.45) \quad J_r(x, \beta) = \frac{1}{2v_r(x)}(\beta - d_r(x))^2 + \frac{1}{6}\theta''(x, \beta_*)(\beta - d_r(x))^3$$

for some  $\beta_* = \beta_*(x) \in [\beta, d_r(x)]$ . Inserting this expansion into the lower and upper bounds derived above and using that  $\theta''(x, \beta)$  is bounded uniformly in  $x$  and  $\beta$ , we obtain

$$(3.46) \quad \psi_r(\beta) \asymp \int_{S_r(\beta)} dx J_r(x, \beta) \sim \int_{S_r(\beta)} dx \frac{1}{2v_r(x)}(\beta - d_r(x))^2, \quad \beta \uparrow C_r.$$

□

#### 4. PROOF OF THEOREMS FOR UPWARD LDP

Sections 4.1–4.3 provide the proof of Theorems 1.8–1.10, respectively.

##### 4.1. Proof of Theorem 1.8.

*Proof.* We first prove the LDP for the maximum degree. Afterwards we prove the LDP for the maximal eigenvalue.

4.1.1. *LDP for the maximum degree.* For  $x \in [0, 1]$ , let  $i_x \in \{1, \dots, N\}$  be the index such that  $x \in [\frac{i_x-1}{N}, \frac{i_x}{N})$ . Let  $d_i$  be the degree of vertex  $i$ . First note that we can sandwich

$$(4.1) \quad \sup_{x \in [0,1]} \mathbb{P}(N^{-1}d_{i_x} \geq \beta) \leq \mathbb{P}\left(N^{-1} \max_{i \in [N]} d_i \geq \beta\right) \leq \sum_{i \in [N]} \mathbb{P}(N^{-1}d_i \geq \beta) \leq N \sup_{x \in [0,1]} \mathbb{P}(N^{-1}d_{i_x} \geq \beta).$$

Using the LDP for the single degrees derived in Section 3.3.1, we obtain

$$(4.2) \quad \begin{aligned} \lim_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{P}\left(N^{-1} \max_{i \in [N]} d_i \geq \beta\right) &= \lim_{N \rightarrow \infty} \sup_{x \in [0,1]} \frac{1}{N} \log \mathbb{P}(N^{-1}d_{i_x} \geq \beta) \\ &= \sup_{x \in [0,1]} \lim_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{P}(N^{-1}d_{i_x} \geq \beta) = - \inf_{x \in [0,1]} J_r(x, \beta) = -\widehat{\psi}_r(\beta). \end{aligned}$$

The convergence of  $\Lambda(\theta)$  in (3.22) is uniform in  $x \in [0, 1]$  because  $\|r_N - r\|_\infty \rightarrow 0$  as  $N \rightarrow \infty$ . Hence the convergence in the LDP for  $N^{-1}d_{i_x}$  is uniform in  $x \in [0, 1]$ . This allows us to swap the supremum and the limit in the second equality.

4.1.2. *LDP for  $\lambda_{\max}(L_N)$ .* By Weyl's interlacing inequalities, we have  $\lambda_{\max}(L_N) \leq \lambda_{\max}(D_N) - \lambda_{\min}(A_N)$ . We also know that  $\lambda_{\max}(D_N) \leq \lambda_{\max}(L_N)$ . Hence

$$(4.3) \quad \lambda_{\max}(D_N) \leq \lambda_{\max}(L_N) \leq \lambda_{\max}(D_N) - \lambda_{\min}(A_N).$$

Since  $r$  is non-negative definite, the smallest eigenvalue of  $T_r$  is 0. Now let  $(T_n)_{n \in \mathbb{N}}$  be a sequence of self-adjoint operators acting on  $L^2([0, 1])$  and converging in operator norm to some operator  $T$ . Then

$$(4.4) \quad |\lambda_{\min}(T_n) - \lambda_{\min}(T)| = \left| \inf_{\|f\|_2=1} \langle f, T_n f \rangle - \inf_{\|f\|_2=1} \langle f, T f \rangle \right| \leq \sup_{\|f\|_2=1} |\langle f, T_n f \rangle - \langle f, T f \rangle| \rightarrow 0$$

as  $n \rightarrow \infty$ . Thus, the map  $h \mapsto \lambda_{\min}(\mathcal{T}_h)$  is continuous in the cut norm. Hence, by the contraction principle,  $N^{-1}\lambda_{\min}(A_N) = \lambda_{\min}(T_{h^{G_N}})$  converges to  $\lambda_{\min}(T_r) = 0$  at an exponential rate with coefficient  $N^2$ . By (4.2) and (4.3) we can sandwich, for all  $\varepsilon > 0$ ,

$$\begin{aligned}
(4.5) \quad -\widehat{\psi}_r(\beta) &= -\inf_{x \in [0,1]} J_r(x, \beta) = \lim_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{P} \left( N^{-1} \max_{i \in [N]} d_i \geq \beta \right) \\
&\leq \lim_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{P}(N^{-1} \lambda_{\max}(L_N) \geq \beta) \\
&\leq \lim_{N \rightarrow \infty} \frac{1}{N} \log \left[ \mathbb{P} \left( N^{-1} \max_{i \in [N]} d_i \geq \beta - \varepsilon \right) + \mathbb{P}(N^{-1} \lambda_{\min}(A_N) \leq -\varepsilon) \right] \\
&= \lim_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{P} \left( N^{-1} \max_{i \in [N]} d_i \geq \beta - \varepsilon \right) = -\inf_{x \in [0,1]} J_r(x, \beta - \varepsilon) = -\widehat{\psi}_r(\beta - \varepsilon).
\end{aligned}$$

The result now follows from continuity of  $\widehat{\psi}_r$ , as stated in Theorem 1.9. Note that the proof of Theorem 1.9 does not use Theorem 1.8.  $\square$

## 4.2. Proof of Theorem 1.9.

4.2.1. *Proof that  $\widehat{\psi}_r$  is strictly increasing.* Assume that  $\inf_{x \in [0,1]} \theta(x, \beta) > 0$  for all  $\beta > C_r$ . Then, by Lemma 3.3, for  $\beta_2 > \beta_1 > C_r$ ,

$$\begin{aligned}
(4.6) \quad \widehat{\psi}_r(\beta_2) &= \inf_{x \in [0,1]} J_r(x, \beta_2) \\
&\geq \inf_{x \in [0,1]} J_r(x, \beta_1) + \theta(x, \beta_1)(\beta_2 - \beta_1) \\
&\geq \widehat{\psi}_r(\beta_1) + \inf_{x \in [0,1]} \theta(x, \beta_1)(\beta_2 - \beta_1) > \widehat{\psi}_r(\beta_1).
\end{aligned}$$

For the first inequality we use that  $J'_r(x, \beta) = \theta(x, \beta)$  is increasing in  $\beta$ . Hence, it suffices to show  $\inf_{x \in [0,1]} \theta(x, \beta) > 0$  for all  $\beta > C_r$ .

Note that the map  $r \mapsto \frac{e^{\theta r}}{1+(e^{\theta}-1)r}$  is concave. Let  $\widetilde{\theta}(x, \beta) = \log \frac{\beta(1-d_r(x))}{(1-\beta)d_r(x)}$ . Then, by Jensen's inequality, for all  $x \in [0, 1]$ ,

$$(4.7) \quad \int_{[0,1]} dy \frac{e^{\widetilde{\theta}(x,\beta)} r(x, y)}{1 + (e^{\widetilde{\theta}(x,\beta)} - 1)r(x, y)} \leq \frac{e^{\widetilde{\theta}(x,\beta)} d_r(x)}{1 + (e^{\widetilde{\theta}(x,\beta)} - 1)d_r(x)} = \beta.$$

Recall that  $\theta(\beta, x)$  is chosen such that

$$(4.8) \quad \int_{[0,1]} dy \frac{e^{\theta(x,\beta)} r(x, y)}{1 + (e^{\theta(x,\beta)} - 1)r(x, y)} = \beta.$$

This implies that  $\theta(x, \beta) \geq \widetilde{\theta}(x, \beta) \geq \log \frac{\beta(1-C_r)}{(1-\beta)C_r}$ . We conclude the proof that  $\widehat{\psi}_r$  is strictly increasing by noting that this lower bound is strictly positive for  $\beta > C_r$  and is independent of  $x$ .

4.2.2. *Continuity of  $\widehat{\psi}_r$ .* Right-continuity of  $\widehat{\psi}_r$  follows from the fact that the infimum of continuous functions is right-continuous. By Lemma 3.3,  $\sup_{x \in [0,1]} \theta(x, \beta) < \infty$  for all  $\beta \geq C_r$ .

Thus,

$$\begin{aligned}
(4.9) \quad \widehat{\psi}_r(\beta - \varepsilon) &= \inf_{x \in [0,1]} J_r(x, \beta - \varepsilon) \\
&\geq \inf_{x \in [0,1]} [J_r(x, \beta) - \theta(x, \beta)\varepsilon] \\
&\geq \widehat{\psi}_r(\beta) - \sup_{x \in [0,1]} \theta(x, \beta)\varepsilon \rightarrow \widehat{\psi}_r(\beta), \quad \varepsilon \downarrow 0.
\end{aligned}$$

4.2.3. *Value of  $\widehat{\psi}_r$  at the boundary.* We first show that  $\widehat{\psi}_r(C_r) = 0$ . By Lemma 3.3,

$$\begin{aligned}
(4.10) \quad \widehat{\psi}_r(C_r) &= \inf_{x \in [0,1]} \left[ J_r(x, d_r(x)) + (C_r - d_r(x)) \sup_{\beta \in [d_r(x), C_r]} \theta(x, \beta) \right] \\
&\leq \inf_{x \in [0,1]} (C_r - d_r(x)) \sup_{\beta \in [d_r(x), C_r]} \theta(x, \beta) = 0,
\end{aligned}$$

where we use that  $C_r = \sup_{x \in [0,1]} d_r(x)$  and that  $\theta(x, \beta)$  is bounded uniformly in  $x$  and  $\beta$ . Since  $\hat{r}_1(x, y) \equiv 1$ , we have

$$(4.11) \quad \widehat{\psi}_r(1) = \inf_{x \in [0,1]} J_r(x, 1) = \inf_{x \in [0,1]} \int_{[0,1]} dy \log \frac{1}{r(x, y)}.$$

4.3. **Proof of Theorem 1.10.** Recall from Section 3.3.5 that

$$(4.12) \quad J_r(x, \beta) = \frac{1}{2v_r(x)} (\beta - d_r(x))^2 + \frac{1}{6} \theta''(x, \beta_*) (\beta - d_r(x))^3,$$

for some  $\beta_* = \beta_*(x) \in [d_r(x), \beta]$ . By Lemma 3.3,  $\theta''$  is bounded uniformly in  $x$  and also uniformly in  $\beta$  bounded away from 1, so it immediately follows that

$$(4.13) \quad \widehat{\psi}_r(\beta) \leq \inf_{x \in \mathcal{D}_r} J_r(x, \beta) \leq \inf_{x \in \mathcal{D}_r} \frac{1}{2v_r(x)} (\beta - C_r)^2 [1 + o(1)], \quad \beta \downarrow C_r.$$

Let  $x_* = x_*(\beta) = \arg \min_{x \in [0,1]} J_r(x, \beta)$ . It is clear from the above that  $\beta - C_r \leq \beta - d_r(x_*) = O(\beta - C_r)$ . Hence

$$(4.14) \quad \widehat{\psi}_r(\beta) = J_r(x_*, \beta) \geq \frac{1}{2v_r(x_*)} (\beta - C_r)^2 + O((\beta - C_r)^3), \quad \beta \downarrow C_r.$$

By continuity of  $r$ ,  $\inf_{x \in \mathcal{D}_r} |x - x_*| \downarrow 0$  as  $\beta \downarrow C_r$ . Since also  $v_r$  is continuous, this implies that  $\frac{1}{2v_r(x_*)} \geq \inf_{x \in \mathcal{D}_r} \frac{1}{2v_r(x)} [1 + o(1)]$  as  $\beta \downarrow C_r$ . We conclude that

$$(4.15) \quad \widehat{\psi}_r(\beta) \geq \inf_{x \in \mathcal{D}_r} \frac{1}{2v_r(x)} (\beta - C_r)^2 [1 + o(1)], \quad \beta \downarrow C_r,$$

which settles the claim.

## REFERENCES

- [1] F. Augeri. Large deviations principle for the largest eigenvalue of Wigner matrices without Gaussian tails. *Electron. J. Probab.*, 21:Paper No. 32, 49, 2016.
- [2] F. Augeri, A. Guionnet, and J. Husson. Large deviations for the largest eigenvalue of sub-Gaussian matrices. *Comm. Math. Phys.*, 383:997–1050, 2021.
- [3] G. Ben Arous and A. Guionnet. Large deviations for Wigner’s law and Voiculescu’s non-commutative entropy. *Probab. Theory Related Fields*, 108(4):517–542, 1997.
- [4] G. Ben Arous, A. Dembo, and A. Guionnet. Aging of spherical spin glasses. *Probab. Theory Related Fields*, 120(1):1–67, 2001.

- [5] C. Bordenave and P. Caputo. A large deviation principle for Wigner matrices without Gaussian tails. *Ann. Probab.*, 42(6):2454–2496, 2014.
- [6] C. Bordenave and M. Lelarge. Resolvent of large random graphs. *Random Structures Algorithms*, 37(3):332–352, 2010.
- [7] W. Bryc. A remark on the connection between the large deviation principle and the central limit theorem. *Stat. Probab. Lett.*, 18:253–256, 1993.
- [8] W. Bryc, A. Dembo, and J. Tiefeng. Spectral measure of large random Hankel, Markov and Toeplitz matrices. *Ann. Probab.*, 34(1):1–38, 2006.
- [9] A. Campbell, K. Luh, S. O’Rourke, S. Arenas-Velilla, and V. Pérez-Abreu. Extreme eigenvalues of laplacian random matrices with gaussian entries. *arXiv preprint arXiv:2211.17175*, 2022.
- [10] A. Chakrabarty, R. S. Hazra, F. den Hollander, and M. Sfragara. Spectra of adjacency and Laplacian matrices of inhomogeneous Erdős-Rényi random graphs. *Random Matrices Theory Appl.*, 10(1):Paper No. 2150009, 34, 2021.
- [11] A. Chakrabarty, R. S. Hazra, F. den Hollander, and M. Sfragara. Large deviation principle for the maximal eigenvalue of inhomogeneous Erdős-Rényi random graphs. *J. Theoret. Probab.*, 35(4):2413–2441, 2022.
- [12] A. Chatterjee and R. S. Hazra. Spectral properties for the Laplacian of a generalized Wigner matrix. *Random Matrices Theory Appl.*, 11(3):Paper No. 2250026, 66, 2022.
- [13] S. Chatterjee. *Large Deviations for Random Graphs*, volume 2197 of *Lecture Notes in Mathematics*. Springer, Cham, 2017. Lecture notes from the 45th Probability Summer School held in Saint-Flour, June 2015, École d’Été de Probabilités de Saint-Flour.
- [14] 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.
- [15] F. den Hollander. *Large Deviations*, volume 14 of *Fields Institute Monographs*. American Mathematical Society, Providence, RI, 2000.
- [16] S. Dhara and S. Sen. Large deviation for uniform graphs with given degrees. *Ann. Appl. Probab.*, 32(3):2327–2353, 2022.
- [17] P. Diao, D. Guilloit, A. Khare, and B. Rajaratnam. Model-free consistency of graph partitioning. *arXiv preprint arXiv:1608.03860*, 2016.
- [18] X. Ding and T. Jiang. Spectral distributions of adjacency and Laplacian matrices of random graphs. *Ann. Appl. Probab.*, 20(6):2086–2117, 2010.
- [19] R. Grone and R. Merris. The Laplacian spectrum of a graph ii. *SIAM Journal on discrete mathematics*, 7(2):221–229, 1994.
- [20] T. Jiang. Empirical distributions of Laplacian matrices of large dilute random graphs. *Random Matrices Theory Appl.*, 1(3):1250004, 20, 2012.
- [21] O. Khorunzhy, M. Shcherbina, and V. Vengerovsky. Eigenvalue distribution of large weighted random graphs. *J. Math. Phys.*, 45(4):1648–1672, 2004.
- [22] S. Lang. *Fundamentals of Differential Geometry*, volume 191 of *Graduate Texts in Mathematics*. Springer, Berlin, 1999.
- [23] L. Lovász. *Large Networks and Graph Limits*, volume 60 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2012.
- [24] E. Lubetzky and Y. Zhao. On replica symmetry of large deviations in random graphs. *Random Structures Algorithms*, 47(1):109–146, 2015.
- [25] M. Markering. The large deviation principle for inhomogeneous Erdős-Rényi random graphs. *J. Theoret. Probab.*, 36(4):711–727, 2023.
- [26] M. Markering. The large deviation principle for inhomogeneous Erdős-Rényi random graphs. *Bachelor Thesis, Leiden University*, 2020.

- [27] U. Von Luxburg, M. Belkin, and O. Bousquet. Consistency of spectral clustering. *The Annals of Statistics*, pages 555–586, 2008.

MATHEMATISCH INSTITUUT, UNIVERSITEIT LEIDEN, THE NETHERLANDS

*Email address:* `r.s.hazra@math.leidenuniv.nl`

*Email address:* `denholla@math.leidenuniv.nl`

DPMMS, UNIVERSITY OF CAMBRIDGE, UNITED KINGDOM

*Email address:* `mjrm2@cam.ac.uk`