ISSN 0037–4466, Siberian Mathematical Journal, 2020, Vol. 61, No. 4, pp. 697–714.

EXPONENTIAL CHEBYSHEV INEQUALITIES FOR RANDOM GRAPHONS AND THEIR APPLICATIONS

A. V. Logachov and A. A. Mogulskii UDC 519.2

Abstract: We prove some exponential Chebyshev inequality and derive the large deviation principle and the law of large numbers for the graphons constructed from a sequence of Erdős–Rényi random graphs with weights. Also, we obtain a new version of the large deviation principle for the number of triangles included in an Erdős–Rényi graph.

DOI: 10.1134/S0037446620040114

Keywords: Erdős–Rényi graph, graphon, large deviation principle, law of large numbers

1. Introduction, Main Notation, and Definitions

Consider a collection of independent random variables X_{ij}, where $1 \le i < j \le n$, distributed identically with a random variable X. Assume that

$$\mathbf{P}(X \in [0,1]) = 1, \tag{1.1}$$

and

$$\mathbf{P}(X \in [0,\varepsilon]) > 0, \quad \mathbf{P}(X \in [1-\varepsilon,1]) > 0 \tag{1.2}$$

for every $\varepsilon > 0$.

Given n vertices, associate an edge with weight X_{ij} to each pair (i,j) of vertices for $1 \le i < j \le n$. The resulting random graph is called an *Erdős–Rényi graph with weights* and denoted by Γ_n. Observe that if

$$\mathbf{P}(X = 0) = 1-p, \quad \mathbf{P}(X = 1) = p, \tag{1.3}$$

then we have an ordinary Erdős–Rényi graph; see [1, 2].

Let us define a *graphon*. Consider a triangle in the plane $\mathbb{R}^2$; i.e.,

$$\triangle := \{(x,y) \in \mathbb{R}^2 :\ x, y \in [0,1],\ y \ge x\}.$$

Refer as the *graphon space* $\mathscr{W}$ to the set of all nonnegative measurable functions $f(x,y)$ mapping $\triangle$ to the segment $[0,1]$.

We will introduce the main metric spaces. We can interpret each graphon $f \in \mathscr{W}$ as the density of some measure on $\triangle$. Then to each $f \in \mathscr{W}$ we assign its "distribution function"

$$F(u,v) := \int_0^u \left(\int_x^v f(x,y)\,dy \right) dx, \quad (u,v) \in \triangle.$$

Define the metric $\rho_W = \rho_W(f,g)$ on $\mathscr{W}$ by putting

$$\rho_W(f,g) := \sup_{(u,v)\in\triangle} |F(u,v) - G(u,v)|,$$

The authors were partially supported by the Russian Science Foundation (Grant 18–11–00129).

Original article submitted November 19, 2019; revised April 28, 2020; accepted June 17, 2020.

where $|F(u,v) - G(u,v)|$ is the distance between the distribution functions F and G for two graphons $f, g \in \mathscr{W}$. The metric space $(\mathscr{W}, \rho_W)$ is complete and separable. Moreover, $\mathscr{W}$ is compact in this space.

Define the metric $d = d(f,g)$ on $\mathscr{W}$, see [3, 4], by putting

$$d(f,g) := \sup_{a,b \in \mathscr{H}} \left| \int_0^1 a(x) \left(\int_x^1 b(y)(f(x,y) - g(x,y)\, dy \right) dx \right|,$$

where $\mathscr{H}$ is the set of all measurable functions from the segment $[0,1]$ to $[-1,1]$. The metric space $(\mathscr{W}, d)$ is complete and separable. In this space $\mathscr{W}$ is not compact.

It is obvious that $\rho_W(f,g) \le d(f,g)$ for all $f, g \in \mathscr{W}$; i.e., the metric d is stronger than ρ_W.

Given $f \in \mathscr{W}$, define the rate function

$$I(f) := \int_{(x,y) \in \Delta} \Lambda(f(x,y))\, dxdy,$$

where $\Lambda(\alpha) := \sup_{\lambda \in \mathbb{R}} \{\alpha\lambda - \log \mathbf{E} e^{\lambda X}\}$ is the deviation function of a random variable X.

Observe that if conditions (1.1) and (1.2) are met then $\Lambda(\alpha)$ is a convex function, bounded on $[\varepsilon, 1-\varepsilon]$ for every $0 < \varepsilon < 1$, and equal to ∞ for $\alpha \notin [0,1]$ and zero at the unique point $a := \mathbf{E}X$.

In the case of ordinary Erdős–Rényi graphs, i.e. in case that (1.3) holds, the deviation function is of the form

$$\Lambda_p(\alpha) := \begin{cases} \log\left(\frac{1}{1-p}\right) & \text{if } \alpha = 0, \\ (1-\alpha)\log\left(\frac{1-\alpha}{1-p}\right) + \alpha \log\left(\frac{\alpha}{p}\right) & \text{if } \alpha \in (0,1), \\ \log\left(\frac{1}{p}\right) & \text{if } \alpha = 1, \\ \infty & \text{if } \alpha \notin [0,1]. \end{cases}$$

Given an integer $n \ge 1$, define $s_n = s_n(x,y) \in \mathscr{W}$ by putting $s_n(x,y) = X_{i,j}$ for $(x,y) \in \square_{i,j}$ and $s_n(x,y) = a$ for $(x,y) \in \triangle'_n$, where

(i) $\square_{i,j} = \square_{i,j,n} := \left(\frac{i-1}{n}, \frac{i}{n}\right) \times \left(\frac{j-1}{n}, \frac{j}{n}\right)$ is some open square of side length $\frac{1}{n}$ and the upper right vertex at $\left(\frac{i}{n}, \frac{j}{n}\right)$ which lies entirely in $\triangle$; it is not difficult to see that there are exactly $k_n := \frac{1}{2}n(n-1)$ of these squares;

(ii) $\triangle'_n$ consists of all points of $\triangle$ beyond the union of $\square_{i,j,n}$; i.e.,

$$\triangle'_n := \triangle \setminus \bigcup_{1 \le i < j \le n} \square_{i,j,n}.$$

The function $s_n \in \mathscr{W}$ is the graphon corresponding to Γ_n. For fixed n to each Erdős–Rényi graph with weights there corresponds the appropriate graphon.

We are interested in the large deviation principle (LDP) and the law of large numbers for the sequence of graphons s_n. These problems for ordinary Erdős–Rényi graphs are solved; see [3, 4] and a survey therein. We will compare our result with others in more detail in Section 2 after the main statements.

Given a metric ρ, denote by $\mathfrak{B}_\rho(\mathscr{W})$ the Borel σ-algebra of subsets of the metric space $(\mathscr{W}, \rho)$. Granted $B \in \mathfrak{B}_\rho(\mathscr{W})$, denote its interior and closure in $(\mathscr{W}, \rho)$ by $(B)_\rho$ and $[B]_\rho$.

Recall the definition of LDP.

Definition 1.1. A family of random functions s_n *satisfies the LDP* in a metric space $(\mathscr{W}, \rho)$ with the rate function $I = I(f) : \mathscr{W} \to [0, \infty]$ and the normalization function such that $\psi(n) : \lim_{n\to\infty} \psi(n) = \infty$ whenever $\{f \in \mathscr{W} : I(f) \le c\}$ is compact in $(\mathscr{W}, \rho)$ for each $c \ge 0$ and every $B \in \mathfrak{B}_\rho(\mathscr{W})$ satisfies

$$\limsup_{n\to\infty} \frac{1}{\psi(n)} \log \mathbf{P}(s_n \in B) \le -I([B]_\rho), \quad \liminf_{n\to\infty} \frac{1}{\psi(n)} \log \mathbf{P}(s_n \in B) \ge -I((B)_\rho),$$

where $I(B) = \inf_{y \in B} I(y)$ and $I(\varnothing) = \infty$.

The rest of this article has the following structure: Section 2 contains the main statements, and by way of example we obtain the LDP for the number of triangles in ordinary Erdős–Rényi graphs and compare our results with those available. In Section 3 we prove the main statements. Section 4 is devoted to auxiliary results.

2. The Main Results

In many problems, in particular that of establishing an LDP, we need to find upper bounds for the probability that a sequence of random elements belongs to some set. The exponential inequalities of Chebyshev-type are a convenient tool for these estimates. The following theorem provides exponential Chebyshev-type inequalities for a sequence of graphons s_n in $(\mathscr{W}, \rho_W)$.

Theorem 2.1. *Let $B \in \mathfrak{B}_{\rho_W}(\mathscr{W})$ be convex and satisfy one of the following:*
(1) *B is open in $(\mathscr{W}, \rho_W)$;*
(2) *B is closed in $(\mathscr{W}, \rho_W)$;*
(3) *$I((B)_{\rho_W}) < \infty$.*
Then $\mathbf{P}(s_n \in B) \le e^{-n^2 I(B)}$.

Remark 2.2. Similar inequalities for the sums of random vectors and trajectories of random walks were previously obtained in [5; 6, Section 4.3].

Put $(g)_{\varepsilon,d} := \{f \in \mathscr{W} : d(f,g) < \varepsilon\}$. The following lemma contains a lower bound in a local large deviation principle for the sequence of graphons s_n. This property was obtained in [3, 4] for ordinary Erdős–Rényi graphs. We propose some analog for Erdős–Rényi graphs with weights.

Lemma 2.3. *Given $\varepsilon > 0$ and $g \in \mathscr{W}$, we have*

$$\liminf_{n\to\infty} \frac{1}{n^2} \log \mathbf{P}(s_n \in (g)_{\varepsilon,d}) \ge -I(g).$$

Remark 2.4. The stronger the metric, the more complicated it is to prove lower bounds in the local LDP. The metric d is “strongest” among those for which this is possible. If instead of d we had considered the stronger metric

$$\rho_L = \rho_L(f,g) := \int_0^1 \left(\int_x^1 |(f(x,y) - g(x,y)|\, dy \right) dx \quad (f, g \in \mathscr{W}) \tag{2.1}$$

then the claim of Lemma 2.3 would be false.

The next result is the LDP for the sequence of graphons s_n.

Theorem 2.5. *Suppose that $\mathscr{W}$ is equipped with a metric ρ and the following hold:*
(i) *$\rho_W(f,g) \le \rho(f,g) \le d(f,g)$ for all $f, g \in \mathscr{W}$;*
(ii) *$\mathscr{W}$ is compact in $(\mathscr{W}, \rho)$.*
Then the sequence of random functions s_n satisfies the LDP in $(\mathscr{W}, \rho)$ with the normalization function $\psi(n) = n^2$ and the rate function $I(f)$.

We also prove the law of large numbers for the sequence of graphons s_n.

Theorem 2.6. *The sequence of random functions s_n satisfies the following law of large numbers:*

$$\lim_{n\to\infty} \mathbf{P}(s_n \in (g_a)_{\varepsilon,d}) = 1$$

for every $\varepsilon > 0$, where $g_a(x,y) \equiv a$ and $(x,y) \in \triangle$.

Graphons make a convenient tool for proving limit theorems for the elements of the structure of random graphs. We were exhibit an example of application of Theorem 2.5.

Consider an ordinary Erdős–Rényi graph, meaning that (1.3) gives the distribution of X. Given $v \in [0,1)$, denote by $T_{n,v}$ the number of triangles included in the graph Γ_n and lacking vertices with index less than vn. Consider the sequence of random variables

$$t_{n,v} := \frac{1}{n^3} T_{n,v}.$$

Let us state a corollary to Theorem 2.5 which is the LDP for the sequence of $t_{n,v}$.

Corollary 2.7. *The sequence of random variables $t_{n,v}$ satisfies the LDP in the metric space $\mathbb{R}$ with normalization function $\psi(n) = n^2$ and rate function*

$$I_{\triangle,v}(u) := \inf_{f \in \mathscr{W} : A_v f = u} \int\limits_{(x,y)\in\Delta} \Lambda_p(f(x,y))\, dxdy,$$

where

$$A_v f := \int\limits_v^1 \int\limits_x^1 \int\limits_y^1 f(x,y) f(y,z) f(x,z)\, dxdydz.$$

Remark 2.8. The case $v = 0$, meaning the LDP for the number of triangles included in the graph, was already considered; see [4] for instance.

We will compare our results with those obtained previously. Several other definitions of a graphon are available [3, 4, 7–9]; see also a detailed survey in [4]: The function $f(x,y)$ is extended by symmetry across the diagonal $y = x$ to the whole square $\square := [0,1]^2$. Denote the space of these extended functions by $\mathscr{W}_\square$.

In [3, 4, 7–9], the *cut-metric* is defined on $\mathscr{W}_\square$ as

$$d_\square(f,g) := \sup_{S,T \subseteq [0,1]} \left| \int\limits_S \int\limits_T f(x,y)\, dxdy - \int\limits_S \int\limits_T g(x,y)\, dxdy \right|.$$

Using the symmetry of $f, g \in \mathscr{W}_\square$ and the definition of $d_\square$, we can show that

$$\rho_W(f,g) \le \frac{1}{2} d_\square(f,g) = d(f,g).$$

Therefore, if $\mathscr{W}_\square$ were compact in $(\mathscr{W}_\square, d_\square)$ then the LDP in this space would follow from Theorem 2.5. However, we can show that $\mathscr{W}_\square$ is not compact in $(\mathscr{W}_\square, d_\square)$. If we consider an ordinary Erdős–Rényi graph then the rate function is finite at each function in $\mathscr{W}_\square$. Consequently, the compactness condition for this set is necessary for the LDP to hold in the sense of Definition 1.1. This problem is solved in [3, 4] by considering the weaker metric

$$\delta_\square(f,g) := \inf_\sigma \sup_{S,T \subseteq [0,1]} \left| \int\limits_S \int\limits_T f(\sigma(x), \sigma(y))\, dxdy - \int\limits_S \int\limits_T g(x,y)\, dxdy \right|,$$

where the infimum is taken over all measure-preserving bijections from $[0,1]$ to $[0,1]$. Furthermore, $\mathscr{W}_\square$ is "enlarged" by considering the space $\widetilde{\mathscr{W}}_\square$ constructed from the cosets of $\delta_\square(f,g)$.

The "enlargement" is insensitive to the vertex enumeration in the graph, which prevents us from obtaining limit theorems in the cases that this is important. For instance, knowing the LDP for a sequence of graphons s_n in $(\widetilde{\mathscr{W}}_\square, \delta_\square)$, we cannot deduce Corollary 2.7 when $v \ne 0$. Moreover, this enlargement complicates the study of moderate deviations for random graphs.

Summarizing, we highlight the main differences of our result from the others.

(1) We study more general graphs; i.e., the Erdős–Rényi graphs with weights.

(2) The metrics we consider enjoy a series of convenient properties: The set $\mathscr{W}$ is compact in $(\mathscr{W}, \rho)$, the set $(f)_{\varepsilon,\rho}$ is convex, and the functional $I(f)$ is lower semicontinuous on $(\mathscr{W}, \rho)$.

(3) We obtain the exponential Chebyshev inequality, a convenient tool for various upper bounds.

(4) The metrics we use enable us to naturally study moderate deviations for random graphs (graphons) and obtain Gaussian approximation in the domain of normal deviations; the authors are planning to address these problems in the next article.

3. Proofs of the Main Results

Proof of Theorem 2.1. Given an integer $n \geq 1$, denote by $\mathbb{S}_n$ the class of functions in $\mathscr{W}$ with $f(x, y) = x_{i,j} \in [0, 1]$ for $(x, y) \in \square_{i,j}$ and $f(x, y) = a$ for $(x, y) \in \triangle'_n$.

Consider the mapping $H : \mathbb{S}_n \to \mathbb{R}^{\frac{1}{2}n(n-1)}$ that assigns to each $f \in \mathbb{S}_n$ the vector

$$Hf = \vec{x} := (x_{1,2}, \dots, x_{1,n}, x_{2,3}, \dots, x_{2,n}, \dots, x_{n-2,n-1}, x_{n-2,n}, x_{n-1,n}).$$

The mapping H is clearly bijective and continuous with respect to the metric ρ_W on $\mathbb{S}_n$ and the Euclidean norm on $\mathbb{R}^{\frac{1}{2}n(n-1)}$. This mapping assigns to a random graphon $s_n(x, y)$ the random vector

$$Hs_n = \mathbf{X} := (X_{1,2}, \dots, X_{1,n}, X_{2,3}, \dots, X_{2,n}, \dots, X_{n-2,n-1}, X_{n-2,n}, X_{n-1,n}).$$

Recall that $\{X_{i,j}\}$ are independent and identically distributed with a random variable X.

Assume that condition (1) is met. Observe that for every open convex set $B \subseteq \mathscr{W}$ the image $H(B \cap \mathbb{S}_n)$ is open and convex in $\mathbb{R}^{\frac{1}{2}n(n-1)}$. Therefore, Theorem 4.8 of Section 4 yields

$$\mathbf{P}(s_n \in B) = \mathbf{P}(s_n \in B \cap \mathbb{S}_n) = \mathbf{P}(\mathbf{X} \in H(B \cap \mathbb{S}_n)) \leq e^{-\Lambda_{\mathbf{X}}(H(B \cap \mathbb{S}_n))}, \tag{3.1}$$

where for $A \subseteq \mathbb{R}^{\frac{1}{2}n(n-1)}$ by definition

$$\Lambda_{\mathbf{X}}(A) := \inf_{\vec{x} \in A} \Lambda_{\mathbf{X}}(\vec{x}).$$

Here $\Lambda_{\mathbf{X}}(\vec{x})$ is the deviation function for $\mathbf{X}$.

Since the coordinates of $\mathbf{X}$ are independent and identically distributed with X, it follows that $\Lambda_{\mathbf{X}}(\vec{x})$ is of the form

$$\Lambda_{\mathbf{X}}(\vec{x}) = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \Lambda(x_{i,j}) = n^2 \int\limits_{(x,y) \in \triangle \setminus \triangle'_n} \Lambda(f(x, y))\, dxdy, \tag{3.2}$$

where $f = H^{-1}\vec{x}$.

Observe that for every function $f \in \mathbb{S}_n$ we have $\Lambda(f(x, y)) = 0$ for $(x, y) \in \triangle'_n$; hence,

$$\int\limits_{(x,y) \in \triangle \setminus \triangle'_n} \Lambda(f(x, y))\, dxdy = \int\limits_{(x,y) \in \triangle} \Lambda(f(x, y))\, dxdy = I(f). \tag{3.3}$$

Therefore, $\mathbf{P}(s_n \in B) \leq e^{-n^2 I(B)}$ by (3.1)–(3.3).

Assume that condition (2) is met. Put

$$B_\varepsilon := \{f \in \mathscr{W} : \inf_{g \in B} \rho_W(f, g) < \varepsilon\}.$$

The set $\mathscr{W}$ is compact in $(\mathscr{W}, \rho_W)$; therefore, so is the closed set B. Lemma 4.3 of Section 4 implies that the functional $I(f)$ is lower semicontinuous. Since B is compact, we obtain

$$\mathbf{P}(s_n \in B) \leq \liminf_{\varepsilon \downarrow 0} \mathbf{P}(s_n \in B_\varepsilon) \leq \liminf_{\varepsilon \downarrow 0} e^{-n^2 I(B_\varepsilon)} \leq e^{-n^2 I(B)}.$$

Assume that condition (3) is met. The set $[B]_{\rho_W}$ then satisfies condition (2). Hence, Lemma 4.5 of Section 4 yields

$$\mathbf{P}(s_n \in B) \leq \mathbf{P}(s_n \in [B]_{\rho_W}) \leq e^{-n^2 I([B]_{\rho_W})} = e^{-n^2 I(B)}. \quad \square$$

Proof of Lemma 2.3. Firstly consider the particular case that

$$g(x, y) \in (\delta,\ 1 - \delta) \quad \text{for all } (x, y) \in \triangle \tag{3.4}$$

for some $\delta > 0$. Since $d(g,f) \le \rho_L(g,f)$, see the definition of $\rho_L(g,f)$ in (2.1), by Lemma 4.6, see Section 4, for all $f,g \in \mathscr{W}$ there is some integer $m_0 = m_0(g,\varepsilon) < \infty$ such that

$$d(g, g_m) \le \frac{\varepsilon}{2}$$

for all $m \ge m_0$. See Section 4 for the definition of g_m. The triangle inequality implies that

$$d(s_n, g) \le d(s_n, g_m) + d(g_m, g) \le d(s_n, g_m) + \frac{\varepsilon}{2},$$

and so

$$\mathbf{P}(s_n \in (g)_{\varepsilon,d}) \ge \mathbf{P}(s_n \in (g_m)_{\frac{\varepsilon}{2},d}). \tag{3.5}$$

Therefore, we start with verifying the inequality

$$\liminf_{n\to\infty} \frac{1}{n^2} \log \mathbf{P}(s_n \in (g_m)_{\frac{\varepsilon}{2},d}) \ge -I(g_m)$$

in the particular case (3.4).

By Lemma 5.6 of [4],

$$d(s_n, g_m) = \sup_{a,b\in\mathscr{H}'} \left| \int_{(x,y)\in\triangle} a(x)b(y)(s_n(x,y) - g_m(x,y))\, dxdy \right|, \tag{3.6}$$

where $\mathscr{H}' := \mathscr{H}_n'$ is the class of step functions $c \in \mathscr{H}$ constant on each interval $\left(\frac{i-1}{n}, \frac{i}{n}\right) \in [0,1]$. Put

$$\triangle_{m,n} := \bigcup_{\square_{r,l,m}\in\triangle} \bigcup_{\square_{i,j,n}\in\square_{r,l,m}} \square_{i,j,n}, \quad \overline{\triangle}_{m,n} := \triangle \setminus \triangle_{m,n}.$$

We can obviously choose $m_1 \ge m_0$ so that

$$\limsup_{n\to\infty} \mu(\overline{\triangle}_{m,n}) \le \frac{\varepsilon}{8} \tag{3.7}$$

for every $m \ge m_1$, where $\mu(A)$ is the Lebesgue measure of A. By (3.6) and (3.7),

$$d(s_n, g_m) \le \sup_{a,b\in\mathscr{H}'} \left| \int_{\triangle_m} a(x)b(y)(s_n(x,y) - g_m(x,y))\, dxdy \right| + 2\mu(\overline{\triangle}_{m,n}) \le \sum_{r,l} Y_{r,l} + \frac{\varepsilon}{4} \tag{3.8}$$

with

$$Y_{r,l} := \mathbf{I}_{\triangle}(\square_{r,l,m}) \frac{1}{n^2} \sup_{a,b\in\mathscr{H}'} \sum_{i,j} a_i b_j (X_{i,j} - g_{r,l,m}) \mathbf{I}_{\square_{r,l,m}}(\square_{i,j,n}),$$

where $a_i b_j$ are the values of $a(x)b(y)$ on the squares $\square_{i,j,n}$ and $g_{r,l,m}$ are the values of $g_m(x,y)$ on the squares $\square_{r,l,m}$, while

$$\mathbf{I}_A(B) := \begin{cases} 1 & \text{if } B \subseteq A, \\ 0 & \text{otherwise.} \end{cases}$$

Noticing that (3.8) involves exactly $\frac{m(m-1)}{2} \le m^2$ terms $Y_{r,l}$, we see that

$$\mathbf{P}\left(d(s_n, g_m) < \frac{\varepsilon}{2}\right) \ge \mathbf{P}\left(\sum_{r,l} Y_{r,l} < \frac{\varepsilon}{4}\right) \ge \mathbf{P}\left(\bigcap_{r,l} \left\{ Y_{r,l} < \frac{\varepsilon}{4m^2} \right\}\right). \tag{3.9}$$

Since $\{Y_{r,l} : \square_{r,l,m} \subset \triangle\}$ is obviously a collection of independent random variables, (3.9) yields

$$\mathbf{P}\left(d(s_n, g_m) < \frac{\varepsilon}{2}\right) \ge \prod_{r,l} \mathbf{P}\left(Y_{r,l} < \frac{\varepsilon}{4m^2}\right). \tag{3.10}$$

Bound each factor on the right-hand side of (3.10) from below by

$$\mathbf{P}_{r,l,m}\left(\frac{\varepsilon}{4m^2N}\right) := \mathbf{P}\left(C_{r,l,m}\left(\frac{\varepsilon}{4m^2N}\right)\right), \tag{3.11}$$

where the parameters r and l satisfying $\square_{r,l,m} \subseteq \triangle$, the parameter $N \geq 1$ is chosen below, and the event $C_{r,l,m}\left(\frac{\varepsilon}{4m^2N}\right)$ is defined as

$$C_{r,l,m}\left(\frac{\varepsilon}{4m^2N}\right) := \left\{\sup_{a,b\in\mathscr{H}'} \frac{1}{n^2}\Big|\sum_{i,j} a_i b_j (X_{i,j} - g_{r,l,m})\mathbf{I}_{\square_{r,l,m}}(\square_{i,j,n})\Big| < \frac{\varepsilon}{4m^2N}\right\}.$$

Prior to bounding the terms of (3.11) from below, we mention some useful properties of the deviation function $\Lambda(\alpha)$. Put $A(\lambda) := \log \mathbf{E}e^{\lambda X}$.

It follows from (3.4) that the constant $\hat{g} := g_{r,l,m}$ satisfies $\hat{g} \in (\delta, 1-\delta)$. Since $\Lambda(\alpha)$ is an analytic function on the interval $(0,1)$, the constant

$$\hat{\lambda} := \lambda(\hat{g}), \quad \text{where } \lambda(\alpha) := \Lambda'(\alpha),$$

is determined. Since

$$\Lambda(\alpha) = \lambda(\alpha)\alpha - A(\lambda(\alpha)), \quad A'(\lambda(\alpha)) = \alpha, \quad \mathbf{E}e^{\lambda(\alpha)X}X = e^{A(\lambda(\alpha))}\alpha$$

for all $\alpha \in (0,1)$, we have

$$\Lambda(\hat{g}) = \hat{\lambda}\hat{g} - A(\hat{\lambda}), \quad \mathbf{E}e^{\hat{\lambda}X}X = e^{A(\hat{\lambda})}\hat{g}.$$

Let us bound (3.11) from below. The event $C_{r,l,m}\left(\frac{\varepsilon}{4m^2N}\right)$ obviously includes the event

$$\left\{\frac{1}{n^2}\Big|\sum_{i,j}(X_{i,j} - \hat{g})\mathbf{I}_{\square_{r,l,m}}(\square_{i,j,n})\Big| < \frac{\varepsilon}{4m^2N}\right\}.$$

Consequently, for all sufficiently large n the relations

$$\left|\frac{1}{\hat{k}}S_{\hat{k}} - \hat{g}\right| < \frac{\varepsilon n^2}{4m^2N\hat{k}}, \quad -\hat{\lambda}S_{\hat{k}} \geq -\hat{k}\hat{\lambda}\hat{g} - \hat{k}|\hat{\lambda}|\frac{\varepsilon}{4N} \tag{3.12}$$

hold of this event, where

$$\hat{k} := \sum_{i,j}\mathbf{I}_{\square_{r,l,m}}(\square_{i,j,n}) \sim \frac{n^2}{m^2} \quad \text{as } n \to \infty, \quad S_{\hat{k}} := \sum_{i,j} X_{i,j}\mathbf{I}_{\square_{r,l,m}}(\square_{i,j,n}).$$

By (3.12),

$$\begin{aligned}\mathbf{P}_{r,l,m}\left(\frac{\varepsilon}{4m^2N}\right) &= \mathbf{E}\left(e^{\hat{\lambda}S_{\hat{k}} - \hat{\lambda}S_{\hat{k}}};\ C_{r,l,m}\left(\frac{\varepsilon}{4m^2N}\right)\right)e^{\hat{k}A(\hat{\lambda}) - \hat{k}A(\hat{\lambda})} \\ = e^{\hat{k}A(\hat{\lambda})}\widehat{\mathbf{E}}\left(e^{-\hat{\lambda}S_{\hat{k}}};\ C_{r,l,m}\left(\frac{\varepsilon}{4m^2N}\right)\right) &\geq e^{\hat{k}A(\hat{\lambda}) - \hat{k}\hat{\lambda}\hat{g} - \hat{k}|\hat{\lambda}|\frac{\varepsilon}{4N}}\widehat{\mathbf{P}}\left(C_{r,l,m}\left(\frac{\varepsilon}{4m^2N}\right)\right) \\ &= e^{-\hat{k}\left(\Lambda(\hat{g}) + |\hat{\lambda}|\frac{\varepsilon}{4N}\right)}\widehat{\mathbf{P}}\left(C_{r,l,m}\left(\frac{\varepsilon}{4m^2N}\right)\right),\end{aligned} \tag{3.13}$$

where $\widehat{\mathbf{E}}$ is the expectation of the distribution $\widehat{\mathbf{P}}$ defined as $\widehat{\mathbf{P}}(B) := e^{-\hat{k}A(\hat{\lambda})}\mathbf{E}(e^{\hat{\lambda}S_{\hat{k}}};\ B)$ using the parameter $\hat{\lambda}$, and B is an arbitrary measurable set.

To continue the proof, we need the following auxiliary assertion:

Lemma 3.1. *We have*

$$\lim_{n\to\infty} \widehat{\mathbf{P}}(C_{r,l,m}(\gamma)) = 1$$

for every $\gamma > 0$.

Proof. By Lemmas 5.9 and 5.10 of [4], for every $\delta > 0$ there exists a set of functions $\mathscr{K}' \subset \mathscr{H}'$ such that

$$\sup_{a,b\in\mathscr{H}'} \frac{1}{n^2}\Big|\sum_{i,j} a_i b_j (X_{i,j} - \hat{g})\mathbf{I}_{\square_{r,l,m}}(\square_{i,j,n})\Big|$$
$$\le \frac{1}{1-2\delta} \sup_{a,b\in\mathscr{K}'} \frac{1}{n^2}\Big|\sum_{i,j} a_i b_j (X_{i,j} - \hat{g})\mathbf{I}_{\square_{r,l,m}}(\square_{i,j,n})\Big| \tag{3.14}$$

and $|\mathscr{K}'| \le (3/\delta)^n$, where $|\mathscr{K}'|$ is the cardinality of $\mathscr{K}'$.

Choosing $\delta = 1/4$ and applying (3.14), we find that $|\mathscr{K}'| \le 12^n$ and

$$\widehat{\mathbf{P}}(C_{r,l,m}(\gamma)) \ge 1 - \widehat{\mathbf{P}}\bigg(\sup_{a,b\in\mathscr{K}'}\bigg|\frac{1}{n^2}\sum_{i,j} a_i b_j (X_{i,j} - \hat{g})\mathbf{I}_{\square_{r,l,m}}(\square_{i,j,n})\bigg| \ge \frac{\gamma}{2}\bigg)$$
$$= 1 - \widehat{\mathbf{P}}\bigg(\bigcup_{a,b\in\mathscr{K}'}\bigg\{\bigg|\frac{1}{n^2}\sum_{i,j} a_i b_j (X_{i,j} - \hat{g})\mathbf{I}_{\square_{r,l,m}}(\square_{i,j,n})\bigg| \ge \frac{\gamma}{2}\bigg\}\bigg)$$
$$\ge 1 - \sum_{a,b\in\mathscr{K}'} \widehat{\mathbf{P}}\bigg(\bigg|\frac{1}{n^2}\sum_{i,j} a_i b_j (X_{i,j} - \hat{g})\mathbf{I}_{\square_{r,l,m}}(\square_{i,j,n})\bigg| \ge \frac{\gamma}{2}\bigg). \tag{3.15}$$

Since $\widehat{\mathbf{E}}X = \hat{g}$, using Theorem 4.2 of [4], for all $\varepsilon > 0$ and $a, b \in \mathscr{K}'$ we obtain

$$\widehat{\mathbf{P}}\bigg(\bigg|\frac{1}{n^2}\sum_{i,j} a_i b_j (X_{i,j} - \hat{g})\mathbf{I}_{\square_{r,l,m}}(\square_{i,j,n})\bigg| \ge \frac{\gamma}{2}\bigg) \le 2e^{-\frac{n^2\gamma^2}{8m^2}}. \tag{3.16}$$

Now (3.15) and (3.16) yield

$$\lim_{n\to\infty} \widehat{\mathbf{P}}(C_{r,l,m}(\gamma)) \ge 1 - \lim_{n\to\infty} 2\cdot 12^{2n} e^{-\frac{n^2\gamma^2}{8m^2}} = 1. \quad \square$$

Resume proving Lemma 2.3. Using (3.10), (3.13), and Lemma 3.1, we infer that

$$\liminf_{n\to\infty} \frac{1}{n^2}\log \mathbf{P}\Big(d(s_n, g_m) < \frac{\varepsilon}{2}\Big) \ge -\frac{1}{m^2}\sum_{r,l}\Lambda(g_{r,l,m}) - \frac{|\hat{\lambda}|\varepsilon}{N}.$$

Choosing N arbitrarily large, we can remove the second term on the right-hand side. Therefore, (3.5), (3.9), and the convexity of $\Lambda(\alpha)$ yield

$$\liminf_{n\to\infty} \frac{1}{n^2}\log \mathbf{P}(s_n \in (g)_{\varepsilon,d}) \ge \liminf_{n\to\infty} \frac{1}{n^2}\log \mathbf{P}\Big(d(s_n, g_m) < \frac{\varepsilon}{2}\Big)$$
$$\ge -\frac{1}{m^2}\sum_{r,l}\Lambda(g_{r,l,m}) = \sum_{r,l}\int_{\square_{r,l}} \Lambda(g_m(x,y))\,dxdy$$
$$= -\sum_{r,l}\mu(\square_{r,l})\Lambda\bigg(\frac{1}{\mu(\square_{r,l})}\int_{\square_{r,l}} g(x,y)\,dxdy\bigg)$$
$$\ge -\int_{\bigcup_{r,l}\square_{r,l}} \Lambda(g(x,y))\,dxdy \ge -\int_{(x,y)\in\triangle} \Lambda(g(x,y))\,dxdy = -I(g).$$

The claim of Lemma 2.3 is now established with the additional assumption (3.4).

Assume now that (3.4) is not satisfied. Construct the graphon

$$g^{(\delta)}(x,y) := \min\{\max\{g(x,y),\delta\},1-\delta\},$$

where $\delta > 0$ is sufficiently small so that

$$\delta \le \min\left\{\frac{\varepsilon}{4},\ \frac{a}{4},\ \frac{1-a}{4}\right\}.$$

Since

$$d(g,g^{(d)}) \le \sup_{(x,y)\in\triangle} |g^{(\delta)}(x,y)-g(x,y)| \le \delta,$$

it follows that

$$\mathbf{P}(d(s_n,g)<\varepsilon) \ge \mathbf{P}(d(s_n,g^{(\delta)})<\varepsilon/2).$$

Applying the already established claim to the graphon $g^{(\delta)}$, we obtain

$$\liminf_{n\to\infty}\frac{1}{n^2}\log\mathbf{P}(d(s_n,g)<\varepsilon) \ge -I(g^{(\delta)}).$$

It remains to observe that the deviation function $\Lambda(\alpha)$ increases for $\alpha > a$ and decreases for $\alpha < a$; therefore, $-I(g^{(\delta)}) \ge -I(g)$. Thus, in the last inequality we can replace the right-hand side by $-I(g)$. $\square$

Proof of Theorem 2.5. By Lemma 4.1.23 of [10], for the sequence of measures generated by s_n it suffices to establish the two claims:

the local LDP: *every function $g \in \mathscr{W}$ satisfies*

$$\lim_{\varepsilon\to0}\limsup_{n\to\infty}\frac{1}{n^2}\log\mathbf{P}(\rho(s_n,g)<\varepsilon) = \lim_{\varepsilon\to0}\liminf_{n\to\infty}\frac{1}{n^2}\log\mathbf{P}(\rho(s_n,g)<\varepsilon) = -I(g);$$

the exponential tightness: *for every $c\in\mathbb{R}$ there exists a compact set $K_c \subseteq \mathscr{W}$ such that*

$$\limsup_{n\to\infty}\frac{1}{n^2}\log\mathbf{P}(s_n\notin K_c) \le -c.$$

Let us prove the local LDP. Since every ε-neighborhood is convex in $(\mathscr{W},\rho_W)$ and $I(g)$ is lower semicontinuous, Theorem 2.1 and condition (i) imply that

$$\begin{aligned}\lim_{\varepsilon\to0}\limsup_{n\to\infty}\frac{1}{n^2}\log\mathbf{P}(\rho(s_n,g)<\varepsilon) &\le \lim_{\varepsilon\to0}\limsup_{n\to\infty}\frac{1}{n^2}\log\mathbf{P}(\rho_W(s_n,g)<\varepsilon)\\ &\le \lim_{\varepsilon\to0}\limsup_{n\to\infty}\frac{1}{n^2}\log e^{-n^2I((g)_{\varepsilon,W})} \le -I(g), \end{aligned}\tag{3.17}$$

where $(g)_{\varepsilon,W} := \{f\in\mathscr{W} : \rho_W(f,g)<\varepsilon\}$.

By Lemma 2.3 and condition (i),

$$\liminf_{n\to\infty}\frac{1}{n^2}\log\mathbf{P}(\rho(s_n,g)<\varepsilon) \ge \liminf_{n\to\infty}\frac{1}{n^2}\log\mathbf{P}(d(s_n,g)<\varepsilon) \ge -I(g) \tag{3.18}$$

for every $\varepsilon > 0$. From (3.17) and (3.18) we obtain the local LDP for s_n.

The exponential tightness of the sequence of measures generated by s_n follows from condition (ii). Combined with Lemma 4.3 of Section 4, this condition also implies that $\{f\in\mathscr{W} : I(f)\le c\}$ is compact for every $c \ge 0$. $\square$

Proof of Theorem 2.6. From (3.5) and (3.10) we obtain

$$\mathbf{P}(s_n\in(g_a)_{\varepsilon,d}) \ge \prod_{r,l}\mathbf{P}\left(\frac{1}{n^2}\sup_{a,b\in\mathscr{H}'}\left|\sum_{i,j}a_ib_j(X_{i,j}-a)\mathbf{I}_{\square_{r,l,m}}(\square_{i,j,n})\right| < \frac{\varepsilon}{4m^2}\right). \tag{3.19}$$

If $\hat{g} = a$ then $\mathbf{P}$ and $\widehat{\mathbf{P}}$ coincide; hence, Lemma 3.1 shows that

$$\lim_{n\to\infty} \mathbf{P}\biggl(\sup_{a,b\in\mathscr{H}'} \biggl|\frac{1}{n^2}\sum_{i,j} a_i b_j (X_{i,j}-a)\mathbf{I}_{\square_{r,l,m}}(\square_{i,j,n})\biggr| \le \frac{\varepsilon}{4m^2}\biggr) = 1 \tag{3.20}$$

for every $\varepsilon > 0$.

Now (3.19) and (3.20) yield $\lim_{n\to\infty} \mathbf{P}(s_n \in (g)_{\varepsilon,d}) = 1$. $\square$

Proof of Corollary 2.7. Put $c_{n,v} := \min(c \in \mathbb{N} : c \ge vn)$. Since the random variable X has the Bernoulli distribution; therefore,

$$t_{n,v} = \frac{1}{n^3}\sum_{i=c_{n,v}}^{n-2}\sum_{j=i+1}^{n-1}\sum_{k=j+1}^{n} X_{ij}X_{jk}X_{ik}$$

$$= \frac{1}{n^3}\sum_{i=c_{n,v}}^{n-2}\sum_{j=i+1}^{n-1}\sum_{k=j+1}^{n} s_n(x,y)s_n(y,z)s_n(x,z)\mathbf{I}_{(x,y)}(\square_{i,j})\mathbf{I}_{(y,z)}(\square_{j,k})\mathbf{I}_{(x,z)}(\square_{i,k}). \tag{3.21}$$

We have

$$\sum_{i=c_{n,v}}^{n-2}\sum_{j=i+1}^{n-1}\sum_{k=j+1}^{n}\frac{1}{n^3} = \frac{(n-c_{n,v}+1)(n-c_{n,v})(n-c_{n,v}-1)}{6n^3}$$

$$= \int_{\frac{c_{n,v}}{n}}^{1}\int_{x}^{1}\int_{y}^{1} dxdydz - \frac{n-c_{n,v}}{6n^3}. \tag{3.22}$$

Since $s_n(x,y)$ belongs to $[0,1]$ and is constant for $(x,y) \in \square_{i,k}$, we infer from (3.21) and (3.22) that

$$\int_{v}^{1}\int_{x}^{1}\int_{y}^{1} s_n(x,y)s_n(y,z)s_n(x,z)\,dxdydz - \int_{v}^{\frac{c_{n,v}}{n}}\int_{x}^{1}\int_{y}^{1} dxdydz - \frac{n-c_{n,v}}{6n^3}$$

$$\le t_{n,v} \le \int_{v}^{1}\int_{x}^{1}\int_{y}^{1} s_n(x,y)s_n(y,z)s_n(x,z)\,dxdydz =: \tilde{t}_{n,v}.$$

Consequently, for every $\varepsilon > 0$ we have

$$\limsup_{n\to\infty}\frac{1}{n^2}\log\mathbf{P}(|t_{n,v}-\tilde{t}_{n,v}| > \varepsilon)$$

$$\le \limsup_{n\to\infty}\frac{1}{n^2}\log\mathbf{P}\biggl(\int_{v}^{\frac{c_{n,v}}{n}}\int_{x}^{1}\int_{y}^{1} dxdydz + \frac{n-c_{n,v}}{6n^3} > \varepsilon\biggr) = -\infty.$$

So, Lemma 4.2.13 of [10] implies that if a sequence $\tilde{t}_{n,v}$ satisfies the LDP then $t_{n,v}$ satisfies the same LDP. Thus, it suffices to obtain the LDP for $\tilde{t}_{n,v}$.

Consider the metric

$$\rho_{A_v}(f,g) := \frac{1}{3}|A_v f - A_v g|, \quad f,g \in \mathscr{W}.$$

Furthermore, equip $\mathscr{W}$ with the metric $\rho_{\triangle,v}(f,g) := \max(\rho_W(f,g), \rho_{A_v}(f,g))$. It is easy to see that $\mathscr{W}$ is compact in the space $(\mathscr{W}, \rho_{\triangle,v})$. Lemma 4.7 of Section 4 implies that $\rho_{A_v}(f,g) \le d(f,g)$. Consequently, the metric space $(\mathscr{W}, \rho_{\triangle,v})$ satisfies the hypotheses of Theorem 2.5. Thus, the sequence s_n satisfies the LPD in this space with the normalization function $\psi(n) = n^2$ and the rate function $I(f)$.

Observe that $A_v s_n = \tilde{t}_{n,v}$. Also it is obvious that the operator A from $(\mathscr{W}, \rho_{\triangle,v})$ to $\mathbb{R}$ is continuous. Therefore, according to Theorem 2.5 and the "contraction principle" (see [11, Theorem 3.1] for instance) the sequence $\tilde{t}_{n,v}$ satisfies the LPD with the normalization function $\psi(n) = n^2$ and the rate function

$$I_{\triangle,v}(u) = \inf_{f\in\mathscr{W}:A_v f=u} \int\limits_{(x,y)\in\Delta} \Lambda_p(f(x,y))\,dxdy. \quad \square$$

4. Auxiliary Results

Lemma 4.1. *The functional $I(f)$ is convex.*

PROOF. Take $r, s \in [0,1]$ with $r + s = 1$. For all $f, g \in \mathscr{W}$ we have

$$rI(f) + sI(g) = \int\limits_{(x,y)\in\triangle} (r\Lambda(f(x,y)) + s\Lambda(g(x,y)))\,dxdy$$

$$\geq \int\limits_{(x,y)\in\triangle} (\Lambda(rf(x,y) + sg(x,y))\,dxdy = I(rf + sg),$$

where the inequality $\geq$ follows from the convexity of the deviation function $\Lambda(\alpha)$. $\square$

Lemma 4.2. *If* $\lim_{n\to\infty} \rho_W(f_n, f) = 0$ *then for every* $\varepsilon > 0$ *and every measurable set* $A \subseteq \triangle$ *we have*

$$\left|\int\limits_A f_n(x,y)\,dxdy - \int\limits_A f(x,y)\,dxdy\right| \leq \varepsilon\mu(A)$$

for all sufficiently large n.

PROOF. If $\mu(A) = 0$ then the inequality is obvious. Assume that $\mu(A) > 0$.

Since A is measurable and $\mu(A) > 0$, by [12, Chapter 5, § 3, Definition 2] there are an integer $m(A,\varepsilon)$ and a collection of disjoint rectangles B_k, for $1 \leq k \leq m(A,\varepsilon)$, such that

$$\mu\biggl(A \ominus \bigcup_{k=1}^{m(A,\varepsilon)} B_k\biggr) \leq \frac{\varepsilon}{4}\mu(A),$$

where $\ominus$ stands for symmetric difference.

Since $\lim_{n\to\infty} \rho_W(f_n, f) = 0$ and $\mu(A) > 0$ for n sufficiently large, we have

$$\sum_{k=1}^{m(A,\varepsilon)} \left|\int\limits_{B_k} f_n(x,y)\,dxdy - \int\limits_{B_k} f(x,y)\,dxdy\right| \leq \frac{\varepsilon}{2}\mu(A).$$

Hence,

$$\left|\int\limits_A f_n(x,y)\,dxdy - \int\limits_A f(x,y)\,dxdy\right| \leq \sum_{k=1}^{m(A,\varepsilon)} \left|\int\limits_{B_k} f_n(x,y)\,dxdy - \int\limits_{B_k} f(x,y)\,dxdy\right|$$

$$+2\max(\sup_{(x,y)\in\triangle} f_n(x,y), \sup_{(x,y)\in\triangle} f(x,y))\mu\biggl(A \ominus \bigcup_{k=1}^{m(A,\varepsilon)} B_k\biggr) \leq \varepsilon\mu(A). \quad \square$$

Lemma 4.3. *Consider a function $\Lambda(\alpha) : \mathbb{R} \to [0, \infty]$ satisfying the following:*
(1) $\Lambda(\alpha)$ is bounded and convex on $[\varepsilon, 1-\varepsilon]$ for every $\varepsilon > 0$;
(2) there exist possibly infinite limits

$$\Lambda(0) := \lim_{\varepsilon \downarrow 0} \Lambda(\varepsilon) \quad \text{and} \quad \Lambda(1) := \lim_{\varepsilon \downarrow 0} \Lambda(1-\varepsilon).$$

If $\lim_{n\to\infty} \rho_W(f_n, f) = 0$ for $f \in \mathscr{W}$ and $f_n \in \mathscr{W}$ then

$$\liminf_{n\to\infty} I(f_n) \ge I(f),$$

where

$$I(g) := \int\limits_{(x,y)\in\triangle} \Lambda(g(x,y))\, dxdy, \quad g \in \mathscr{W}.$$

Proof. Consider the case that $\lim_{\varepsilon\downarrow 0} \Lambda(\varepsilon) = \lim_{\varepsilon\downarrow 0} \Lambda(1-\varepsilon) = \infty$.
Put

$$A_\varepsilon := \{(x,y) \in \triangle : \varepsilon \le f(x,y) \le 1-\varepsilon\}, \quad A_0 := \{(x,y) \in \triangle : f(x,y) = 0\},$$
$$A_1 := \{(x,y) \in \triangle : f(x,y) = 1\}.$$

By Theorem 16.1 of [13],

$$\int\limits_{(x,y)\in\triangle} \Lambda(f(x,y))\, dxdy = \int\limits_{\triangle\setminus(A_0\cup A_1)} \Lambda(f(x,y))\, dxdy + \mu(A_0)\Lambda(0) + \mu(A_1)\Lambda(1)$$
$$= \lim_{\varepsilon\downarrow 0} \int\limits_{\triangle\setminus(A_0\cup A_1)} \Lambda(f(x,y))\mathbf{I}_{f(x,y)}([\varepsilon, 1-\varepsilon])\, dxdy + \mu(A_0)\Lambda(0) + \mu(A_1)\Lambda(1)$$
$$= \lim_{\varepsilon\downarrow 0} \int\limits_{A_\varepsilon} \Lambda(f(x,y))\, dxdy + \mu(A_0)\Lambda(0) + \mu(A_1)\Lambda(1),$$

where we put $0 \cdot \infty = 0$.

Applying again Theorem 16.1 of [13], we obtain

$$\int\limits_{(x,y)\in\triangle} \Lambda(f_n(x,y))\, dxdy = \lim_{\varepsilon\downarrow 0} \int\limits_{A_\varepsilon} \Lambda(f_n(x,y))\, dxdy + \int\limits_{A_0} \Lambda(f_n(x,y))\, dxdy$$
$$+ \int\limits_{A_1} \Lambda(f_n(x,y))\, dxdy =: \lim_{\varepsilon\downarrow 0} I_{n,\varepsilon} + I_{n,0} + I_{n,1}. \tag{4.1}$$

Consider the first term on the right-hand side. Since $\Lambda(\alpha)$ is a bounded convex function on $[\varepsilon, 1-\varepsilon]$, it is also continuous there. Since the function $f(x,y)$ is measurable and bounded, there is a sequence of simple functions $\hat{f}_m(x,y)$ with

$$\lim_{m\to\infty} \sup_{(x,y)\in\triangle} |f(x,y) - \hat{f}_m(x,y)| = 0,$$

$$\lim_{m\to\infty} \int\limits_{A_\varepsilon} \Lambda(\hat{f}_m(x,y))\, dxdy = \int\limits_{A_\varepsilon} \Lambda(f(x,y))\, dxdy.$$

Hence, for every $\gamma > 0$ there are an integer m_γ, measurable sets $A_{k,\gamma}$, and constants $\hat{f}_{k,\gamma}$, for $1 \le k \le m_\gamma$, such that $A_{k,\gamma} \cap A_{j,\gamma} = \varnothing$ for $k \ne j$ with $\bigcup_{k=1}^{m_\gamma} A_{k,\gamma} = A_\varepsilon$ and $\hat{f}_m(x,y) \equiv \hat{f}_{k,\gamma}$ for $(x,y) \in A_{k,\gamma}$, and

$$\sup_{(x,y)\in A_{k,\gamma}} |f(x,y) - \hat{f}_{k,\gamma}| < \gamma \tag{4.2}$$

for all $1 \le k \le m_\gamma$. By Lemma 4.2, for all $\gamma > 0$ and $1 \le k \le m_\gamma$ we have

$$\left| \int\limits_{A_{k,\gamma}} f_n(x,y)\,dxdy - \int\limits_{A_{k,\gamma}} f(x,y)\,dxdy \right| < \gamma\mu(A_{k,\gamma}) \tag{4.3}$$

for n sufficiently large. Since $\Lambda(\alpha)$ is uniformly continuous on $[\varepsilon/2, 1-\varepsilon/2]$, for every $\delta > 0$ there exists $\gamma := \gamma(\delta) \in (0, \varepsilon/2)$ such that

$$\sup_{s\in[-\gamma,\gamma]} |\Lambda(\alpha+s) - \Lambda(\alpha)| < \delta \tag{4.4}$$

for all $\alpha \in [\varepsilon, 1-\varepsilon]$.

Using Jensen's inequality, (4.3) and (4.4), for $1 \le k \le m_\gamma$ and n sufficiently large we obtain

$$\begin{aligned} \int\limits_{A_{k,\gamma}} \Lambda(f_n(x,y))\,dxdy &\ge \mu(A_{k,\gamma})\Lambda\Bigg(\frac{1}{\mu(A_{k,\gamma})} \int\limits_{A_{k,\gamma}} f_n(x,y)\,dxdy\Bigg) \\ &\ge \mu(A_{k,\gamma}) \inf_{s\in[-\gamma,\gamma]} \Lambda\Bigg(\frac{1}{\mu(A_{k,\gamma})} \int\limits_{A_{k,\gamma}} f(x,y)\,dxdy + s\Bigg) \\ &\ge \mu(A_{k,\gamma})\Bigg(\Lambda\Bigg(\frac{1}{\mu(A_{k,\gamma})} \int\limits_{A_{k,\gamma}} f(x,y)\,dxdy\Bigg) - \delta\Bigg). \end{aligned} \tag{4.5}$$

From (4.2), (4.4), and (4.5) we infer that

$$\begin{aligned} \int\limits_{A_{k,\gamma}} \Lambda(f_n(x,y))\,dxdy &\ge \mu(A_{k,\gamma}) \inf_{s\in[-\gamma,\gamma]} \Bigg(\Lambda\Bigg(\frac{1}{\mu(A_{k,\gamma})} \int\limits_{A_{k,\gamma}} \hat{f}_{k,\gamma}\,dxdy + s\Bigg) - \delta\Bigg) \\ &\ge \mu(A_{k,\gamma})\Bigg(\Lambda\Bigg(\frac{1}{\mu(A_{k,\gamma})} \int\limits_{A_{k,\gamma}} \hat{f}_{k,\gamma}\,dxdy\Bigg) - 2\delta\Bigg) = \mu(A_{k,\gamma})(\Lambda(\hat{f}_{k,\gamma}) - 2\delta) \\ &= \int\limits_{A_{k,\gamma}} \Lambda(\hat{f}_{k,\gamma})\,dxdy - 2\delta\mu(A_{k,\gamma}). \end{aligned} \tag{4.6}$$

Then (4.2), (4.4), and (4.6) yield

$$\begin{aligned} \int\limits_{A_{k,\gamma}} \Lambda(f_n(x,y))\,dxdy &\ge \int\limits_{A_{k,\gamma}} \inf_{s\in[-\gamma,\gamma]} \Lambda(f(x,y)+s)\,dxdy - 2\delta\mu(A_{k,\gamma}) \\ &\ge \int\limits_{A_{k,\gamma}} \Lambda(f(x,y))\,dxdy - 3\delta\mu(A_{k,\gamma}). \end{aligned} \tag{4.7}$$

By (4.7), for all $\varepsilon > 0$, $\delta > 0$, and sufficiently large n we have

$$\begin{aligned} I_{n,\varepsilon} &= \int\limits_{A_\varepsilon} \Lambda(f_n(x,y))\,dxdy = \sum_{k=1}^{m_\gamma} \int\limits_{A_{k,\gamma}} \Lambda(f_n(x,y))\,dxdy \\ &\ge \sum_{k=1}^{m_\gamma} \Bigg(\int\limits_{A_{k,\gamma}} \Lambda(f(x,y))\,dxdy - 3\delta\mu(A_{k,\gamma})\Bigg) \ge \int\limits_{A_\varepsilon} \Lambda(f(x,y))\,dxdy - 3\delta. \end{aligned} \tag{4.8}$$

Let us estimate $I_{n,0}$ from below. By Lemma 4.2, for every $\varepsilon > 0$ and sufficiently large n we have

$$0 \le \int\limits_{A_0} f_n(x,y)\,dxdy < \varepsilon\mu(A_0). \tag{4.9}$$

Put $A_{0,\varepsilon,n} := \{(x,y) \in A_0 : f_n(x,y) > 2\varepsilon\}$. From (4.9) we find that

$$2\varepsilon\mu(A_{0,\varepsilon,n}) \le \int\limits_{A_0} f_n(x,y)\,dxdy < \varepsilon\mu(A_0).$$

Thus,

$$\mu(A_0 \setminus A_{0,\varepsilon,n}) \ge \frac{\mu(A_0)}{2}.$$

Hence,

$$I_{n,0} = \int\limits_{A_0} \Lambda(f_n(x,y))\,dxdy \ge \int\limits_{A_0\setminus A_{0,\varepsilon,n}} \Lambda(f_n(x,y))\,dxdy \ge \frac{\mu(A_0)}{2} \inf_{\alpha\in[0,2\varepsilon]} \Lambda(\alpha) \tag{4.10}$$

for n sufficiently large.

Similar arguments show that

$$I_{n,1} = \int\limits_{A_1} \Lambda(f_n(x,y))\,dxdy \ge \frac{\mu(A_1)}{2} \inf_{\alpha\in[1-2\varepsilon,1]} \Lambda(\alpha). \tag{4.11}$$

If $\max(\mu(A_0),\mu(A_1)) = 0$ then (4.1) and (4.8) show that for all $\varepsilon > 0$ and $\delta > 0$ we have

$$\begin{aligned}&\liminf_{n\to\infty} \int\limits_{(x,y)\in\triangle} \Lambda(f_n(x,y))\,dxdy \\ &\ge \liminf_{n\to\infty} \int\limits_{A_\varepsilon} \Lambda(f_n(x,y))\,dxdy \ge \int\limits_{A_\varepsilon} \Lambda(f(x,y))\,dxdy - 3\delta.\end{aligned}$$

Passing to the limit as $\delta \to 0$, for every $\varepsilon > 0$ we obtain

$$\liminf_{n\to\infty} \int\limits_{(x,y)\in\triangle} \Lambda(f_n(x,y))\,dxdy \ge \int\limits_{A_\varepsilon} \Lambda(f(x,y))\,dxdy.$$

This implies that

$$\liminf_{n\to\infty} \int\limits_{(x,y)\in\triangle} \Lambda(f_n(x,y))\,dxdy \ge \lim_{\varepsilon\downarrow 0} \int\limits_{A_\varepsilon} \Lambda(f(x,y))\,dxdy = \int\limits_{(x,y)\in\triangle} \Lambda(f(x,y))\,dxdy.$$

If $\max(\mu(A_0),\mu(A_1)) > 0$ then (4.10) and (4.11) yield

$$\begin{aligned}&\liminf_{n\to\infty} \int\limits_{(x,y)\in\triangle} \Lambda(f_n(x,y))\,dxdy \\ &\ge \lim_{\varepsilon\downarrow 0} \max\left(\frac{\mu(A_0)}{2} \inf_{\alpha\in[0,2\varepsilon],} \Lambda(\alpha), \frac{\mu(A_1)}{2} \inf_{\alpha\in[1-2\varepsilon,1]} \Lambda(\alpha)\right) = \infty.\end{aligned}$$

Thus, Lemma 4.3 is established in the case that $\lim_{\varepsilon\downarrow 0} \Lambda(\varepsilon) = \lim_{\varepsilon\downarrow 0} \Lambda(1-\varepsilon) = \infty$. The remaining cases are similar. □

Lemma 4.4. *If $f, g \in \mathscr{W}$ with $I(g) < \infty$ then*

$$\lim_{\varepsilon\downarrow 0} I((f)_{\varepsilon,W}) = I(f), \tag{4.12}$$

$$\lim_{r\uparrow 1} I((1-r)g + rf) = I(f). \tag{4.13}$$

Proof. Lemma 4.3 shows that $I(f)$ is lower semicontinuous, and so (4.12) holds.

Lemma 4.1 implies that $I(f)$ is convex. Using the inequality $I(g) < \infty$, we infer that

$$\lim_{r\uparrow 1} I((1-r)g + rf) \le \lim_{r\uparrow 1}((1-r)I(g) + rI(f)) = I(f). \tag{4.14}$$

On the other hand, lower semicontinuity yields

$$\lim_{r\uparrow 1} I((1-r)g + rf) \ge I(f). \tag{4.15}$$

Finally, (4.14) and (4.15) force (4.13). $\square$

Lemma 4.5. *If $B \in \mathfrak{B}_{\rho_W}(\mathscr{W})$ is a convex set with $I((B)_{\rho_W}) < \infty$ then*

$$I((B)_{\rho_W}) = I(B) = I([B]_{\rho_W}). \tag{4.16}$$

Proof. Let us prove firstly that if B is a convex set then so are the following sets:

(i) $(B)_{\rho_W}$;

(ii) $[B]_{\rho_W}$;

(iii) $[B]_{\rho_W} \cap I_{<\infty}$, where $I_{<\infty} := \{f : I(f) < \infty\}$.

(i) To verify that $(B)_{\rho_W}$ is convex, take $f_1, f_2 \in (B)_{\rho_W}$. Then $(f_1)_{\delta,W},\ (f_2)_{\delta,W} \subset B$ for some $\delta > 0$, and since B is convex, we have

$r(f_1)_{\delta,W} + s(f_2)_{\delta,W} \subset B$, where $r, s \in [0,1]$ with $r + s = 1$. Since

$$(rf_1 + sf_2)_{\delta,W} \subset r(f_1)_{\delta,W} + s(f_2)_{\delta,W},$$

it follows that $(rf_1 + sf_2)_{\delta,W} \subset B$. This means that $f = rf_1 + sf_2$ lies in B together with some neighborhood; i.e., f belongs to $(B)_{\rho_W}$. The convexity of $(B)_{\rho_W}$ is established.

(ii) To verify that $[B]_{\rho_W}$ is convex, take $f, g \in [B]_{\rho_W}$. This means that there exist two sequences $f_n \in B$ and $g_n \in B$ converging respectively to f and g. Since B is convex; $rf_n + sg_n \in B$, where $r, s \in [0,1]$ with $r + s = 1$. Since

$$\rho_W(rf_n + sg_n, rf + sg) \le r\rho_W(f_n, f) + s\rho_W(g_n, g) \to 0,$$

it follows that $rf_n + sg_n \to rf + sg$. This means that $rf + sg$ lies in the closure $[B]_{\rho_W}$ of B. The convexity of $[B]_{\rho_W}$ is established.

(iii) Verify that $[B]_{\rho_W} \cap I_{<\infty}$ is convex. By Lemma 4.1 the functional $I(f)$ is convex, and so $I_{<\infty}$ is convex. Hence, so is $[B]_{\rho_W} \cap I_{<\infty}$ as the intersection of two convex sets.

By assumption $I((B)_{\rho_W}) < \infty$; hence, $(B)_{\rho_W} \cap I_{<\infty}$ is nonempty. Choose $f_0 \in (B)_{\rho_W} \cap I_{<\infty}$ and $f_1 \in [B]_{\rho_W} \cap I_{<\infty}$ and consider the half-open interval

$$[f_0,\ f_1) := \{f_r := rf_1 + (1-r)f_0 :\ r \in [0,1)\}.$$

Let us show that

$$[f_0,\ f_1) \subset (B)_{\rho_W}. \tag{4.17}$$

Both points f_0 and f_1 lie in the convex set $[B]_{\rho_W} \cap I_{<\infty}$, and so f_r lies in $[B]_{\rho_W} \cap I_{<\infty}$ for every $r \in [0,1]$. To justify (4.17), we must exclude the possibility that $f_r \in \partial B$ for all $r \in [0,1)$. However, if $f_r \in \partial B$ then there exists a sequence $g_n \notin [B]_{\rho_W}$ with $g_n \to f_r$. In this case

$$h_n := \frac{g_n - rf_1}{1-r} \to f_0 \quad \text{as } n \to \infty.$$

Since $f_0 \in (B)_{\rho_W}$, it follows that $h_n \in (B)_{\rho_W}$ for all sufficiently large n. Since $g_n = rf_1+(1-r)h_n$ and $f_1 \in [B]_{\rho_W}$, it follows that $g_n \in [B]_{\rho_W}$ for n sufficiently large. The resulting contradiction establishes (4.17).

From (4.17) we infer that

$$I((B)_{\rho_W}) \le I(rf_1 + (1-r)f_0), \quad 0 \le r < 1. \tag{4.18}$$

Since $f_0, f_1 \in I_{<\infty}$, by claim (4.13) of Lemma 4.4 the right-hand side of (4.18) converges to $I(f_1)$ as $r \uparrow 1$. Thus, $I((B)_{\rho_W}) \le I(f_1)$. Since $f_1 \in [B]_{\rho_W} \cap I_{<\infty}$ is arbitrary, we arrive at

$$I((B)_{\rho_W}) \le I([B]_{\rho_W} \cap I_{<\infty}) = I([B]_{\rho_W}).$$

This inequality together with the obvious inequalities

$$I((B)_{\rho_W}) \ge I(B) \ge I([B]_{\rho_W})$$

justifies claim (4.16) of Lemma 4.5. $\square$

Recall that $\square_{r,l} = \square_{r,l,m} := \left(\frac{r-1}{m}, \frac{r}{m}\right) \times \left(\frac{l-1}{m}, \frac{l}{m}\right)$ are open squares of side length $\frac{1}{m}$, for $m \in \mathbb{N}$, with the upper right vertices at $\left(\frac{r}{m}, \frac{l}{m}\right)$, which lie entirely in $\triangle$.

Given $g \in \mathscr{W}$, put

$$g_m(x,y) := \begin{cases} \frac{1}{\mu(\square_{r,l})} \int\limits_{\square_{r,l}} g(x,y)\,dxdy, (x,y) \in \square_{r,l} & \text{if } (x,y) \in \square_{r,l}, \\ a & \text{if } (x,y) \in \overline{\triangle}_m, \end{cases}$$

where $\triangle_m := \bigcup_{r,l} \square_{r,l}$ and $\overline{\triangle}_m := \triangle \setminus \triangle_m$.

Lemma 4.6. *For all $g \in \mathscr{W}$ and $\varepsilon > 0$ there is an integer $m = m(\varepsilon, g)$ such that $\rho_L(g, g_m) \le \varepsilon$.*

Proof. By Luzin's Theorem (the version for $\mathbb{R}^d$, see [14, Theorem 3.6.1]) there are a continuous function $\tilde{g} = \tilde{g}(x,y) : \triangle \to [0,1]$ and a measurable set $A \subseteq \triangle$ such that

$$\mu(\overline{A}) \le \frac{\varepsilon}{6}, \quad \overline{A} := \triangle \setminus A, \tag{4.19}$$

and for all $(x,y) \in A$ we have

$$g(x,y) = \tilde{g}(x,y). \tag{4.20}$$

Since $\tilde{g}(x,y)$ is a continuous function on $\triangle$, there is an integer $m = m(g, \varepsilon)$ such that

$$\sup_{(x,y)\in\triangle_m} |\tilde{g}(x,y) - \tilde{g}_m(x,y)| \le \frac{2\varepsilon}{3} \quad \text{and} \quad \mu(\overline{\triangle}_m) \le \frac{\varepsilon}{6}. \tag{4.21}$$

The triangle inequality yields

$$\rho_L(g, g_m) \le \rho_L(g, \tilde{g}) + \rho_L(\tilde{g}, \tilde{g}_m) + \rho_L(\tilde{g}_m, g_m), \tag{4.22}$$

and it suffices to estimate each term on the right-hand side of (4.22).

Estimate the first term. From (4.19) and (4.20), using the inequality $\sup_{(x,y)\in\triangle} |g(x,y) - \tilde{g}(x,y)| \le 1$, we infer that

$$\rho_L(g, \tilde{g}) = \int\limits_A |g(x,y) - \tilde{g}(x,y)|\,dxdy + \int\limits_{\overline{A}} |g(x,y) - \tilde{g}(x,y)|\,dxdy \le \mu(\overline{A}) \le \frac{\varepsilon}{6}. \tag{4.23}$$

Estimate the second term in (4.22). By (4.21), taking into account the inequality

$$\sup_{(x,y)\in\triangle} |\tilde{g}(x,y) - \tilde{g}_m(x,y)| \le 1,$$

we conclude that

$$\rho_L(\tilde g,\tilde g_m)=\int\limits_{\triangle_m}|\tilde g(x,y)-\tilde g_m(x,y)|\,dxdy+\int\limits_{\overline{\triangle}_m}|\tilde g(x,y)-\tilde g_m(x,y)|\,dxdy$$

$$\le\frac{2\varepsilon}{3}\mu(\triangle_m)+\mu(\overline{\triangle}_m)\le\frac{\varepsilon}{2}. \tag{4.24}$$

To estimate the third term in (4.22), inspect firstly for $(x,y)\in\square_{r,l}$ the variable $|\tilde g_m(x,y)-g_m(x,y)|$. Since $\sup_{(x,y)\in\triangle}|\tilde g(x,y)-g(x,y)|\le 1$, we have

$$|\tilde g_m(x,y)-g_m(x,y)|=\left|\frac{1}{\mu(\square_{r,l})}\int\limits_{\square_{r,l}}(\tilde g(u,v)-g(u,v))\,dudv\right|$$

$$\le\frac{1}{\mu(\square_{r,l})}\int\limits_{\square_{r,l}\cap A}|\tilde g(u,v)-g(u,v)|\,dudv+\frac{1}{\mu(\square_{r,l})}\mu(\square_{r,l}\cap\overline{A})$$

$$=\frac{1}{\mu(\square_{r,l})}\mu(\square_{r,l}\cap\overline{A}). \tag{4.25}$$

Using (4.19), (4.21), and (4.25), as well as $\sup\limits_{(x,y)\in\triangle}|\tilde g_m(x,y)-\tilde g_m(x,y)|\le 1$, we obtain

$$\rho_L(\tilde g_m,g_m)\le\int\limits_{\triangle_m}|\tilde g_m(x,y)-g_m(x,y)|\,dxdy+\int\limits_{\overline{\triangle}_m}|\tilde g_m(x,y)-g_m(x,y)|\,dxdy$$

$$\le\sum_{r,l}\mu(\square_{r,l}\cap\overline{A})+\frac{\varepsilon}{6}\le\mu(\overline{A})+\frac{\varepsilon}{6}\le\frac{\varepsilon}{3}. \tag{4.26}$$

Finally, (4.22)–(4.24) and (4.26) show that $\rho_L(g,g_m)\le\varepsilon$. $\square$

Lemma 4.7. *$\rho_{A_v}(f,g)\le d(f,g)$ for all $f,g\in\mathscr{W}$.*

Proof. Put $h(x,y):=f(x,y)-g(x,y)$. It is easy to see that

$$3\rho_{A_v}(f,g)\le I_1+I_2+I_3,$$

where

$$I_1:=\left|\int\limits_v^1\left(\int\limits_x^1\left(\int\limits_y^1 H_1(x,y,z)\,dz\right)dy\right)dx\right|,\quad H_1(x,y,z):=f(x,y)h(y,z)f(x,z);$$

$$I_2:=\left|\int\limits_v^1\left(\int\limits_v^z\left(\int\limits_v^y H_2(x,y,z)\,dx\right)dy\right)dz\right|,\quad H_2(x,y,z):=h(x,y)g(y,z)f(x,z);$$

$$I_3:=\left|\int\limits_v^1\left(\int\limits_v^y\left(\int\limits_y^1 H_3(x,y,z)\,dz\right)dx\right)dy\right|,\quad H_3(x,y,z):=g(x,y)g(y,z)h(x,z).$$

To estimate I_1, observe that

$$I_1\le\int\limits_v^1\overline{I}_1(x)\,dx,\quad \overline{I}_1(x):=\left|\int\limits_x^1\left(\int\limits_y^1 H_1(x,y,z)\,dz\right)dy\right|.$$

Put $\mathscr{H}_x:=\{a\in\mathscr{H}:a(y)\equiv 0\text{ for }y\in[0,x]\}$.

Since

$$\overline{I}_1(x) \le \sup_{a,b\in\mathscr{H}} \left| \int_x^1 a(y) \left(\int_y^1 b(z)(f(y,z) - g(y,z))\, dz \right) dy \right|$$
$$= \sup_{a\in\mathscr{H}_x, b\in\mathscr{H}} \left| \int_0^1 a(y) \left(\int_y^1 b(z)(f(y,z) - g(y,z))\, dz \right) dy \right|$$
$$\le \sup_{a,b\in\mathscr{H}} \left| \int_0^1 a(y) \left(\int_y^1 b(z)(f(y,z) - g(y,z))\, dz \right) dy \right| = d(f,g)$$

for every fixed $x \in [v, 1]$, we see that

$$I_1 \le \int_v^1 d(f,g)\, dx \le d(f,g).$$

The inequalities $I_2 \le d(f,g)$ and $I_3 \le d(f,g)$ are established similarly. $\square$

For the reader's convenience, we state Theorem 1.3.1 of [6].

Theorem 4.8. *Suppose that a random vector* $\mathbf{X} \in \mathbb{R}^m$, *with* $m \ge 1$, *satisfies Cramér's condition* $[\mathbf{C}_0]$: *There exists* $\lambda > 0$ *such that* $\mathbf{E}e^{\lambda|\mathbf{X}|} < \infty$. *Then for every convex open set* $B \subseteq \mathbb{R}^m$ *we have*

$$\mathbf{P}(\mathbf{X} \in B) \le e^{-\Lambda_{\mathbf{X}}(B)}.$$

References

1. Erdős P. and Rényi A., "On the evolution of random graphs," Publ. Math. Inst. Hungar. Acad. Sci., vol. 5, 17–61 (1960).
2. Bollobás B., *Random Graphs*, Cambridge Univ. Press, London (2001).
3. Chatterjee S. and Varadhan S. R. S., "The large deviation principle for the Erdős–Rényi random graph," European J. Combin., vol. 32, 1000–1017 (2011).
4. Chatterjee S., *Large Deviations for Random Graphs*, Springer, Stanford (2017).
5. Borovkov A. A. and Mogulskii A. A., "Chebyshev type exponential inequalities for sums of random vectors and random walk trajectories," Theory Probab. Appl., vol. 56, no. 1, 21–43 (2012).
6. Borovkov A. A., *Asymptotic Analysis of Random Walks* [Russian], Fizmatlit, Moscow (2013).
7. Borgs C., Chayes J. T., Lovász L., Sós V. T., and Vesztergombi K., "Convergent sequences of dense graphs. I: Subgraph frequencies, metric properties and testing," Adv. Math., vol. 219, no. 6, 1801–1851 (2008).
8. Borgs C., Chayes J. T., Lovász L., Sós V. T., and Vesztergombi K., "Convergent sequences of dense graphs. II: Multiway cuts and statistical physics," Ann. Math., vol. 176, no. 2, 151–219 (2012).
9. Bollobás B. and Riordan O., "Metrics for sparse graphs," in: *Surveys in Combinatorics 2009*. Papers from the 22nd British combinatorial conference, St. Andrews, UK, July 5–10, 2009 (2009), 211–287.
10. Dembo A. and Zeitouni O., *Large Deviations Techniques and Applications*, Springer-Verlag, New York (1998).
11. Wentzell A. D. and Freidlin M. I., *Fluctuations in Dynamical Systems Under the Action of Small Random Perturbations* [Russian], Nauka, Moscow (1979).
12. Kolmogorov A. N. and Fomin S. V., *Elements of the Theory of Functions and Functional Analysis*, Dover Publications, Mineola (2009).
13. Dyachenko M. I. and Ulyanov P. L., *Measure and Integral* [Russian], Faktorial, Moscow (1998).
14. Heil C., *Introduction to Real Analysis*, Springer-Verlag, New York (2019).

A. V. Logachov
Sobolev Institute of Mathematics, Novosibirsk State University
Novosibirsk State University of Economics and Management
Siberian State University of Geosystems and Technology, Novosibirsk, Russia
E-mail address: `omboldovskaya@mail.ru`

A. A. Mogulskii
Sobolev Institute of Mathematics
Novosibirsk State University, Novosibirsk, Russia
E-mail address: `mogul@math.nsc.ru`