Distributed Compression of Graphical Data

Payam Delgosha and Venkat Anantharam

Department of Electrical Engineering and Computer Sciences
University of California, Berkeley
{pdelgosha, ananth} @ berkeley.edu

February 22, 2018

Abstract

In contrast to time series, graphical data is data indexed by the nodes and edges of a graph. Modern applications such as the internet, social networks, genomics and proteomics generate graphical data, often at large scale. The large scale argues for the need to compress such data for storage and subsequent processing. Since this data might have several components available in different locations, it is also important to study distributed compression of graphical data. In this paper, we derive a rate region for this problem which is a counterpart of the Slepian–Wolf Theorem. We characterize the rate region when the statistical description of the distributed graphical data is one of two types – a marked sparse Erdős–Rényi ensemble or a marked configuration model. Our results are in terms of a generalization of the notion of entropy introduced by Bordenave and Caputo in the study of local weak limits of sparse graphs.

1 Introduction

Nowadays, storing combinatorically structured data is of great importance in many applications such as the internet, social networks and biological data. For instance, a social network could be presented as a graph where each node models an individual and each edge stands for a friendship. Also, vertices and edges can carry marks, e.g. the mark of a vertex represents its type, and the mark of an edge represents its shared information. Due to the sheer amount of such data, compressing it has drawn attention, see e.g. [CS12], [Abb16], [DA17]. As the data is not always available in one location, it is important to also consider distributed compression of graphical data.

Traditionally, distributed lossless compression is modeled using two (or more) correlated stationary and ergodic processes representing the components of the data at the individual locations. In this case, the rate region is given by the Slepian–Wolf Theorem [CT12]. We adopt an analogous framework, namely that two correlated marked random graphs on the same vertex set are presented to two encoders which then individually compress their data such that a third party can recover both realizations from the two compressed representations, with a vanishing probability of error in the asymptotic limit of data size.

We characterize the compression rate region for two scenarios, namely, a marked sparse Erdős–Rényi ensemble and a marked configuration model. We employ the framework of local weak convergence, also called the objective method, as a counterpart for marked graphs of the notion of stochastic processes [BS01, AS04, AL07]. Our characterization is best understood in terms of a generalization of a measure of entropy introduced by Bordenave and Caputo, which we call the BC entropy [BC14]. It turns out that the BC entropy captures the per–vertex growth rate of the Shannon entropy for the ensembles we study in this paper. This motivates it as a natural measure governing the asymptotic compression bounds.

The paper is organized as follows. In Section 2 we introduce the notation and formally state the problem. Sections 3 and 4 give a brief introduction to the objective method and the BC entropy, mostly specialized for the examples we study. Finally, in Section 5, we characterize the rate region for the scenarios we present in Section 2.

2 Notations and Problem Statement

The set of real numbers is denoted by $\mathbb{R}$. For an integer n, $[n]$ denotes the set $\{1, 2, \dots, n\}$. For a probability distribution P, $H(P)$ denotes its Shannon entropy. Also, for a random variable X, we denote by $H(X)$ its Shannon entropy. For a positive integer N and a sequence of positive integers $\{a_i\}_{1 \le i \le k}$ such that $\sum a_i \le N$, we define

$$\binom{N}{\{a_i\}_{1 \le i \le k}} := \frac{N!}{a_1! \dots a_k! (N - a_1 - \dots - a_k)!}.$$

For sequences of reals a_n and b_n we write $a_n = O(b_n)$ if, for some constant $C \ge 0$, we have $|a_n| \le C|b_n|$ for n large enough. Furthermore, we write $a_n = o(b_n)$ if $a_n/b_n \to 0$ as $n \to \infty$. We denote by $\mathbb{1}[A]$ the indicator of the event A. For a probability distribution P , $X \sim P$ denotes that the random variable X has law P. Throughout the paper logarithms are to the natural base.

A marked graph with edge mark set Ξ and vertex mark set Θ is a graph where each edge carries a mark in Ξ and each vertex carries a mark in Θ. We assume that all graphs are simple unless otherwise stated. Also, we assume that all edge and vertex mark sets are finite. For two vertices v and w in a graph G, $v \sim_G w$ denotes that v and w are adjacent in G.

Let G be a marked graph on a finite vertex set with edges and vertices carrying marks in the sets Ξ and Θ, respectively. We denote the edge mark count vector of G by $\vec{m}_G = \{m_G(x)\}_{x \in \Xi}$, where $m_G(x)$ is the number of edges in G carrying mark x. Furthermore, we denote the vertex mark count vector of G by $\vec{u}_G = \{u_G(\theta)\}_{\theta \in \Theta}$, where $u_G(\theta)$ denotes the number of vertices in G with mark θ. Additionally, for a graph G on the vertex set $[n]$, we denote the degree sequence of G by $\overrightarrow{\mathrm{dg}}_G = \{\mathrm{dg}_G(1), \dots, \mathrm{dg}_G(n)\}$ where $\mathrm{dg}_G(i)$ denotes the degree of vertex i. For a degree sequence $\vec{d} = (d(1), \dots, d(n))$ and an integer k, we define

$$c_k(\vec{d}) := |\{1 \le i \le n : d(i) = k\}. \tag{1}$$

Also, for two degree sequences $\vec{d} = (d(1), \dots, d(n))$ and $\vec{d'} = (d'(1), \dots, d'(n))$, and two integers k and l, we define

$$c_{k,l}(\vec{d}, \vec{d'}) := |\{1 \le i \le n : d(i) = k, d'(i) = l\}|. \tag{2}$$

Given a degree sequence $\vec{d} = (d(1), \dots, d(n))$, we let $\mathcal{G}^{(n)}_{\vec{d}}$ denote the set of simple unmarked graphs G on the vertex set $[n]$ such that $\mathrm{dg}_G(i) = d(i)$ for $1 \le i \le n$.

Throughout this paper, we assume that Ξ_1 and Ξ_2 are two fixed and finite sets of edge marks. Moreover, Θ_1 and Θ_2 are two fixed and finite vertex mark sets. For $i \in \{1, 2\}$ and an integer n, let $\mathcal{G}^{(n)}_i$ be the set of marked graphs on the vertex set $[n]$ with edge and vertex mark sets Ξ_i and Θ_i, respectively. For two graphs $G_1 \in \mathcal{G}^{(n)}_1$ and $G_2 \in \mathcal{G}^{(n)}_2$, $G_1 \oplus G_2$ denotes the superposition of G_1 and G_2 which is a marked graph defined as follows: a vertex $1 \le v \le n$ in $G_1 \oplus G_2$ carries the mark (θ_1, θ_2) where θ_i is the mark of v in G_i. Furthermore, we place an edge in $G_1 \oplus G_2$ between vertices v and w if there is an edge between them in at least one of G_1 of G_2, and mark this edge (x_1, x_2), where, for $1 \le i \le 2$, x_i is the mark of the edge (v, w) in G_i if it exists and $\circ_i$ otherwise. Here $\circ_1$ and $\circ_2$ are auxiliary marks not present in $\Xi_1 \cup \Xi_2$. Note that $G_1 \oplus G_2$ is a marked graph with edge and vertex mark sets $\Xi_{1,2} := (\Xi_1 \cup \{\circ_1\}) \times (\Xi_2 \cup \{\circ_2\}) \setminus \{(\circ_1, \circ_2)\}$ and $\Theta_{1,2} := \Theta_1 \times \Theta_2$, respectively. We use the terminology *jointly marked graph* to refer to a marked graph with edge and vertex makr sets $\Xi_{1,2}$ and $\Theta_{1,2}$ respectively. With this, let $\mathcal{G}_{1,2}^{(n)}$ denote the set of jointly marked graphs on the vertex set $[n]$. Moreover, for $i \in \{1,2\}$, we say that a graph is in the i–th domain if edge and vertex marks come from Ξ_i and Θ_i, respectively. For a jointly marked graph $G_{1,2}$ and $1 \le i \le 2$, the i–th marginal of $G_{1,2}$, denoted by G_i, is the marked graph in the i–th domain obtained by projecting all vertex and edge marks onto Ξ_i and Θ_i, respectively, followed by removing edges with mark $\circ_i$. Note that any jointly marked graph $G_{1,2}$ is uniquely determined by its marginals G_1 and G_2, because $G_{1,2} = G_1 \oplus G_2$. Given an edge mark count $\vec{m} = \{m(x)\}_{x \in \Xi_{1,2}}$, for $x_1 \in \Xi_1 \cup \{\circ_1\}$, with an abuse of notation we define

$$m(x_1) := \sum_{(x_1', x_2') \in \Xi_{1,2} \,:\, x_1' = x_1} m((x_1', x_2')). \tag{3}$$

In a similar fashion, we define $m(x_2)$ for $x_2 \in \Xi_2 \cup \{\circ_2\}$. Likewise, given a vertex mark count vector $\vec{u} = \{u(\theta)\}_{\theta \in \Theta_{1,2}}$, and for $\theta_1 \in \Theta_1$ and $\theta_2 \in \Theta_2$, we define

$$u(\theta_1) = \sum_{\theta_2' \in \Theta_2} u((\theta_1, \theta_2')) \qquad u(\theta_2) = \sum_{\theta_1' \in \Theta_1} u((\theta_1', \theta_2)). \tag{4}$$

Assume that we have a sequence of random graphs $G_{1,2}^{(n)} \in \mathcal{G}_{1,2}^{(n)}$, drawn according to some ensemble distribution. Additionally, assume that there are two encoders who want to compress realizations of such jointly marked graphs in a distributed fashion. Namely, the i–th encoder, $1 \le i \le 2$, has only access to the i–th marginal $G_i^{(n)}$. We assume that the encoders know the distribution of $G_{1,2}^{(n)}$.

Definition 1. *An $\langle n, L_1^{(n)}, L_2^{(n)} \rangle$ code is a tuple of functions $(f_1^{(n)}, f_2^{(n)}, g^{(n)})$ for each n such that*

$$f_i^{(n)} : \mathcal{G}_i^{(n)} \to [L_i^{(n)}] \qquad i \in \{1,2\},$$

and

$$g^{(n)} : [L_1^{(n)}] \times [L_2^{(n)}] \to \mathcal{G}_{1,2}^{(n)}.$$

The probability of error for this code corresponding to the ensemble of $G_{1,2}^{(n)}$, which is denoted by $P_e^{(n)}$, is defined as

$$P_e^{(n)} := \mathbb{P}\left(g^{(n)}(f_1^{(n)}(G_1^{(n)}), f_2^{(n)}(G_2^{(n)})) \neq G_{1,2}^{(n)} \right).$$

Now we define our achievability criterion.

Definition 2. *A rate tuple $(\alpha_1, R_1, \alpha_2, R_2) \in \mathbb{R}^4$ is said to be achievable for the above scenario if there is a sequence of $\langle n, L_1^{(n)}, L_2^{(n)} \rangle$ codes such that*

$$\limsup_{n \to \infty} \frac{\log L_i^{(n)} - (\alpha_i n \log n + R_i n)}{n} \le 0 \qquad i \in \{1,2\}, \tag{5}$$

and also $P_e^{(n)} \to 0$. The rate region $\mathcal{R} \in \mathbb{R}^4$ is defined as follows: for fixed α_1 and α_2, if there are sequences $R_1^{(m)}$ and $R_2^{(m)}$ with limit points R_1 and R_2 in $\mathbb{R}$, respectively, such that for each m, the rate tuple $(\alpha_1, R_1^{(m)}, \alpha_2, R_2^{(m)})$ is achievable, then we include $(\alpha_1, R_1, \alpha_2, R_2)$ in the set $\mathcal{R}$.

In this paper, we characterize the above rate region for the following two sequences of ensembles:

The Erdős–Rényi ensemble: Assume that nonnegative real numbers $\vec{p} = \{p_x\}_{x \in \Xi_{1,2}}$ together with a probability distribution $\vec{q} = \{q_\theta\}_{\theta \in \Theta_{1,2}}$ are given such that for all $x_1 \in \Xi_1 \cup \{\circ_1\}$ and $x_2 \in \Xi_2 \cup \{\circ_2\}$, we have

$$\sum_{\substack{(x'_1,x'_2) \in \Xi_{1,2} \\ x'_1 = x_1}} p_{(x'_1,x'_2)} > 0 \qquad \text{and} \qquad \sum_{\substack{(x'_1,x'_2) \in \Xi_{1,2} \\ x'_2 = x_2}} p_{(x'_1,x'_2)} > 0. \tag{6}$$

For an integer n large enough, we define the probability distribution $\mathcal{G}(n; \vec{p}, \vec{q})$ on $\mathcal{G}_{1,2}^{(n)}$ as follows: for each pair of vertices $1 \le i < j \le n$, the edge (i,j) is present in the graph and has mark $x \in \Xi_{1,2}$ with probability p_x/n, and is not present with probability $1 - \sum_{x \in \Xi_{1,2}} p_x/n$. Furthermore, each vertex in the graph is given a mark $\theta \in \Theta_{1,2}$ with probability q_θ. The choice of edge and vertex marks is done independently.

The configuration model ensemble: Assume that a fixed integer $\Delta > 0$ and a probability distribution $\vec{r} = \{r_k\}_{k=0}^{\Delta}$ supported on the set $\{0, \dots, \Delta\}$ are given, such that $r_0 < 1$. Moreover, assume that probability distributions $\vec{\gamma} = \{\gamma_x\}_{x \in \Xi_{1,2}}$ and $\vec{q} = \{q_\theta\}_{\theta \in \Theta_{1,2}}$ on the sets $\Xi_{1,2}$ and $\Theta_{1,2}$, respectively, are given. We assume that for all $x_1 \in \Xi_1 \cup \{\circ_1\}$ and $x_2 \in \Xi_2 \cup \{\circ_2\}$, we have

$$\sum_{\substack{(x'_1,x'_2) \in \Xi_{1,2} \\ x'_1 = x_1}} \gamma_{(x'_1,x'_2)} > 0 \qquad \text{and} \qquad \sum_{\substack{(x'_1,x'_2) \in \Xi_{1,2} \\ x'_2 = x_2}} \gamma_{(x'_1,x'_2)} > 0. \tag{7}$$

Furthermore, let $\vec{d}^{(n)} = \{d^{(n)}(1), \dots, d^{(n)}(n)\}$ be a sequence of degree sequences such that for all n and $1 \le i \le n$, we have $d^{(n)}(i) \le \Delta$ and also $\sum_{i=1}^n d^{(n)}(i)$ is even. Let $m_n := (\sum_{i=1}^n d^{(n)}(i))/2$. Additionally, if for $0 \le k \le \Delta$, $c_k(\vec{d}^{(n)})$ denotes the number of $1 \le i \le n$ such that $d^{(n)}(i) = k$, we assume that for some constant $K > 0$,

$$\sum_{k=0}^{\Delta} |c_k(\vec{d}^{(n)}) - n r_k| \le K n^{1/2}. \tag{8}$$

Now, we define the law $\mathcal{G}(n; \vec{d}^{(n)}, \vec{\gamma}, \vec{q})$ on $\mathcal{G}_{1,2}^{(n)}$ for n large enough as follows. First, we pick an unmarked graph on the vertex set $[n]$ uniformly at random among the set of graphs G with maximum degree Δ such that for each $0 \le k \le \Delta$, $c_k(\overrightarrow{\mathrm{dg}}_G) = c_k(\vec{d}^{(n)})$.[1] Then, we assign i.i.d. marks with law $\vec{\gamma}$ on the edges and i.i.d. marks with law $\vec{q}$ on the vertices.

As we will discuss in Section 3 below, the sequence of Erdős–Rényi ensembles defined above converges in the local weak sense to a marked Poisson Galton Watson tree. Moreover, the configuration model ensemble converges in the same sense to a marked Galton Watson process with degree distribution $\vec{r}$. We will characterize the achievability rate regions in Section 5 in terms of these limiting objects for the above two sequences of ensembles. This will turn out to be best understood in terms of a measure of entropy discussed in Section 4 below.

3 The framework of Local Weak Convergence

In this section, we discuss the framework of local weak convergence mainly in the context of the Erdős–Rényi and configuration model ensembles discussed in Section 2. For a general discussion, the reader is referred to [BS01, AS04, AL07].

[1] The fact that each degree is bounded to Δ, $r_0 < 1$ and the sum of degrees is even implies that $\vec{d}^{(n)}$ is a graphic sequence for n large enough. This is, for instance, a consequence of Theorem 4.5 in [BC14].

Let Ξ and Θ be finite mark sets. A marked graph G with edge and vertex mark sets Ξ and Θ respectively together with a distinguished vertex o, is called a rooted marked graph and is denoted by (G, o). For a rooted marked graph (G, o) and integer $h \geq 1$, $(G, o)_h$ denotes the h neighborhood of o, i.e. the subgraph consisting of vertices with distance no more than h from o. Note that $(G, o)_h$ is connected by definition. Two connected rooted marked graphs (G_1, o_1) and (G_2, o_2) are said to be isomorphic if there is a vertex bijection between the two graphs that maps o_1 to o_2, preserves adjacencies and also preserves vertex and edge marks. With this, we denote the isomorphism class corresponding to a rooted marked graph (G, o) by $[G, o]$. We simply use $[G, o]_h$ as a shorthand for $[(G, o)_h]$. Let $\mathcal{G}_*(\Xi, \Theta)$ denote the set of isomorphism classes $[G, o]$ of connected rooted marked graphs on a countable vertex set with edge and vertex marks coming from the sets Ξ and Θ, respectively. It can be shown that $\mathcal{G}_*(\Xi, \Theta)$ can be turned into a separable and complete metric space [AL07]. For a probability distribution μ on $\mathcal{G}_*(\Xi, \Theta)$, let $\deg(\mu)$ denote the expected degree at the root in μ.

For a finite marked graph G and a vertex v in G, let $G(v)$ denote the connected component of v. With this, if v is a vertex chosen uniformly at random in G, we define $U(G)$ be the law of $[G(v), v]$, which is a probability distribution on $\mathcal{G}_*(\Xi, \Theta)$.

Let $G_{1,2}^{(n)}$ be a random jointly marked graph with law $\mathcal{G}(n; \vec{p}, \vec{q})$ and let v_n be a vertex chosen uniformly at random in the set $[n]$. A simple Poisson approximation implies that $D_x(v_n)$, the number of edges adjacent to v_n with mark $x \in \Xi_{1,2}$, converges in distribution to a Poisson random variables with mean p_x as n goes to infinity. Moreover, $\{D_x(v_n)\}_{x \in \Xi_{1,2}}$ are asymptotically mutually independent. A similar argument can be repeated for any other vertex in the neighborhood of v_n. Also, it can be shown that the probability of having cycles converges to zero. In fact, the structure of $(G_{1,2}^{(n)}, v_n)_h$ converges in distribution to a rooted marked Poisson Galton Watson tree with depth h.

More precisely, let $(T_{1,2}^{\mathrm{ER}}, o)$ be a rooted jointly marked tree defined as follows. First, the mark of the root is chosen from distribution $\vec{q}$. Then, for $x \in \Xi_{1,2}$, we independently generate D_x with law Poisson(p_x). We then add D_x many edges with mark x to the root o. For each offspring, we repeat the same procedure independently, i.e. choose its mark and edges with each mark from the corresponding Poisson distribution. Recursively repeating this, we get a connected jointly marked tree $T_{1,2}^{\mathrm{ER}}$ rooted at o, which has possibly countably infinitely many vertices. Let $\mu_{1,2}^{\mathrm{ER}}$ denote the law of the isomorphism class $[T_{1,2}^{\mathrm{ER}}, o]$. Note that $\mu_{1,2}^{\mathrm{ER}}$ is a probability distribution on $\mathcal{G}_*(\Xi_{1,2}, \Theta_{1,2})$. The above discussion implies that $[G_{1,2}^{(n)}, v_n]_h$ converges in distribution to $[T_{1,2}^{\mathrm{ER}}, o]_h$. In fact, even a stronger statement can be proved. More precisely, if we consider the sequence of random graphs $G_{1,2}^{(n)}$ independently on a joint probability space, $U(G_{1,2}^{(n)})$ converges weakly to $\mu_{1,2}^{\mathrm{ER}}$ with probability one. With this, we say that, almost surely, $\mu_{1,2}^{\mathrm{ER}}$ is the *local weak limit* of the sequence $G_{1,2}^{(n)}$, where the term "local" stands for looking at a fixed depth neighborhood of a typical node.

With the above construction, for $1 \leq i \leq 2$, let T_i^{ER} be the i–th marginal of $T_{1,2}^{\mathrm{ER}}$. Moreover, let μ_i^{ER} be the law of $[T_i^{\mathrm{ER}}(o), o]$. Therefore, μ_i^{ER} is a probability distribution on $\mathcal{G}_*(\Xi_i, \Theta_i)$. Similarly, one can see that, almost surely, μ_i^{ER} is the local weak limit of the sequence $G_i^{(n)}$.

A similar picture also holds for the configuration model. More precisely, let $(T_{1,2}^{\mathrm{CM}}, o)$ be a rooted jointly marked random tree constructed as follows. First, we generate the degree of the root with law $\vec{r}$. Then, for each offspring w of o, we independently generate the offspring count of w with law $\vec{r'} = \{r'_k\}_{k=0}^{\Delta-1}$ defined as

$$r'_k = \frac{(k+1) r_{k+1}}{\mathbb{E}[X]}, \qquad 0 \leq k \leq \Delta - 1,$$

where X has law $\vec{r}$. We continue this process recursively, i.e. for each vertex other than the root, we independently generate its offspring count with law r'. The distribution $\vec{r}'$ is called the *sized biased* distribution, and takes into account the fact that each node other than the root has an extra edge on top of it, and hence its degree should be biased in order to get the correct degree distribution $\vec{r}$. Then, for each vertex and edge existing in the graph $T_{1,2}^{\mathrm{CM}}$, we generate marks independently with laws $\vec{q}$ and $\vec{\gamma}$, respectively. Let $\mu_{1,2}^{\mathrm{CM}}$ be the law of $[T_{1,2}^{\mathrm{CM}}, o]$. Moreover, for $1 \le i \le 2$, let μ_i^{CM} be the law of $[T_i^{\mathrm{CM}}(o), o]$. It can be shown that if $G_{1,2}^{(n)}$ has law $\mathcal{G}(n; \vec{d}^{(n)}, \vec{\gamma}, \vec{q})$ then, almost surely, $\mu_{1,2}^{\mathrm{CM}}$ is the local weak limit of $G_{1,2}^{(n)}$, and μ_i^{CM} is the local weak limit of $G_i^{(n)}$, for $1 \le i \le 2$.

Given Ξ and Θ, not all probability distributions on $\mathcal{G}_*(\Xi, \Theta)$ can appear as the local weak limit of a sequence. In fact, the condition that all vertices have the same chance of being chosen as the root for a finite graph manifests itself as a certain stationarity condition at the limit called *unimodularity* [AL07].

4 The BC entropy

In this section, we discuss a notion of entropy for probability distributions on a space of rooted marked graphs. This is a marked version of the entropy defined by Bordenave and Caputo in [BC14], which was defined for probability distributions on the space of rooted (unmarked) graphs. To distinguish it from the Shannon entropy, we call this notion of entropy the BC entropy.

Let Ξ and Θ be finite mark sets and let μ be a probability distribution on $\mathcal{G}_*(\Xi, \Theta)$. Moreover, let $\vec{m}^{(n)}$ and $\vec{u}^{(n)}$ be sequences of edge and vertex mark counts, respectively, such that for all $x \in \Xi$, $2m^{(n)}(x)/n$ converges to the expected number of edges with mark x connected to the root in μ, and for all $\theta \in \Theta$, $u^{(n)}(\theta)/n$ converges to the probability of the mark of the root in μ being θ. Let $\mathcal{G}^{(n)}_{\vec{m}^{(n)},\vec{u}^{(n)}}(\mu, \epsilon)$ be the set of graphs G on the vertex set $[n]$ with edge and vertex marks in Ξ and Θ, respectively, such that $\vec{m}_G = \vec{m}^{(n)}$, $\vec{u}_G = \vec{u}^{(n)}$, and $U(G)$ is in the ball around μ with radius ϵ with respect to the Lévy–Prokhorov distance [Bil13].

Definition 3. *If $a_n := \sum_{x \in \Xi} m^{(n)}(x)$, define*

$$\overline{\Sigma}(\mu, \epsilon) := \limsup_{n \to \infty} \frac{\log |\mathcal{G}^{(n)}_{\vec{m}^{(n)},\vec{u}^{(n)}}(\mu, \epsilon)| - a_n \log n}{n}$$
$$\underline{\Sigma}(\mu, \epsilon) := \liminf_{n \to \infty} \frac{\log |\mathcal{G}^{(n)}_{\vec{m}^{(n)},\vec{u}^{(n)}}(\mu, \epsilon)| - a_n \log n}{n}$$

Note that both $\overline{\Sigma}(\mu, \epsilon)$ and $\underline{\Sigma}(\mu, \epsilon)$ decrease as ϵ decreases. Therefore, we may define the upper and lower BC entropies as $\overline{\Sigma}(\mu) := \lim_{\epsilon \downarrow 0} \overline{\Sigma}(\mu, \epsilon)$ and $\underline{\Sigma}(\mu) := \lim_{\epsilon \downarrow 0} \underline{\Sigma}(\mu, \epsilon)$. If $\overline{\Sigma}(\mu) = \underline{\Sigma}(\mu)$, we denote the common value by $\Sigma(\mu)$ and call it the BC entropy of μ.

Using similar techniques as in the proof of Theorem 1.2 in [BC14], one can show that $\overline{\Sigma}(\mu)$ and $\underline{\Sigma}(\mu)$ do not depend on the specific choice of the sequences $\vec{m}^{(n)}$ and $\vec{u}^{(n)}$, and $\overline{\Sigma}(\mu) = \underline{\Sigma}(\mu)$ for all μ with positive expected degree of the root.

Now, we connect the asymptotic behavior of the entropy of the ensembles defined in Section 2 to the BC entropy of their local weak limits. Assume that $G_{1,2}^{(n)}$ has law $\mathcal{G}(n; \vec{p}, \vec{q})$. Let $d_{1,2}^{\mathrm{ER}} := \deg(\mu_{1,2}^{\mathrm{ER}}) = \sum_{x \in \Xi_{1,2}} p_x$. Moreover, we use the following notational conventions for $x_i \in \Xi_i$ and $\theta_i \in \Theta_i$, $1 \le i \le 2$.

$$\begin{aligned} p_{x_1} &:= \sum_{x'_2 \in \Xi_2 \cup \{\circ_2\}} p_{(x_1, x'_2)} \quad & p_{x_2} &:= \sum_{x'_1 \in \Xi_1 \cup \{\circ_1\}} p_{(x'_1, x_2)} \\ q_{\theta_1} &:= \sum_{\theta'_2 \in \Theta_2} q_{(\theta_1, \theta'_2)} \quad & q_{\theta_2} &:= \sum_{\theta'_1 \in \Theta_1} q_{(\theta'_1, \theta_2)} \end{aligned} \tag{9}$$

For $1 \le i \le 2$, let $d_i^{\text{ER}} := \deg(\mu_i^{\text{ER}}) = \sum_{x_i \in \Xi_i} p_{x_i}$. If $Q = (Q_1, Q_2)$ has law $\vec{q}$, it can be easily verified that with $s(x)$ defined as $\frac{x}{2} - \frac{x}{2}\log x$ for $x > 0$ and zero for $x = 0$, we have

$$H(G_{1,2}^{(n)}) = \frac{d_{1,2}^{\text{ER}}}{2} n \log n + n \left(H(Q) + \sum_{x \in \Xi_{1,2}} s(p_x) \right) + o(n) \tag{10a}$$

$$H(G_1^{(n)}) = \frac{d_1^{\text{ER}}}{2} n \log n + n \left(H(Q_1) + \sum_{x_1 \in \Xi_1} s(p_{x_1}) \right) + o(n) \tag{10b}$$

$$H(G_2^{(n)}) = \frac{d_2^{\text{ER}}}{2} n \log n + n \left(H(Q_2) + \sum_{x_2 \in \Xi_2} s(p_{x_2}) \right) + o(n) \tag{10c}$$

Using a generalization of Theorem 1.3 in [BC14], it can be seen that the coefficient of n in the above 3 equations are $\Sigma(\mu_{1,2}^{\text{ER}})$, $\Sigma(\mu_1^{\text{ER}})$ and $\Sigma(\mu_2^{\text{ER}})$, respectively.

Similarly, for the configuration model, let $G_{1,2}^{(n)}$ be distributed according to $\mathcal{G}(n; \vec{d}^{(n)}, \vec{\gamma}, \vec{q})$ and let X be a random variable with law $\vec{r}$. Moreover, let $\Gamma^k = (\Gamma_1^k, \Gamma_2^k)$, $1 \le k \le \Delta$, be an i.i.d. sequence distributed according to $\vec{\gamma}$. With this, let

$$X_1 := \sum_{k=1}^{X} \mathbb{1}\left[\Gamma_1^k \neq \circ_1\right] \qquad X_2 := \sum_{k=1}^{X} \mathbb{1}\left[\Gamma_2^k \neq \circ_2\right]. \tag{11}$$

Note that for $1 \le i \le 2$, X_i is basically the distribution of the degree of the root in μ_i^{CM}. If $d_{1,2}^{\text{CM}} := \deg(\mu_{1,2}^{\text{CM}})$ and, for $1 \le i \le 2$, $d_i^{\text{CM}} := \deg(\mu_i^{\text{CM}})$, it can be seen that (see Appendix A for the details)

$$\begin{aligned} H(G_{1,2}^{(n)}) = \frac{d_{1,2}^{\text{CM}}}{2} n \log n + n\Big(&- s(d_{1,2}^{\text{CM}}) + H(X) - \mathbb{E}\left[\log X!\right] \\ &+ H(Q) + \frac{d_{1,2}^{\text{CM}}}{2} H(\Gamma)\Big) + o(n) \end{aligned} \tag{12a}$$

$$\begin{aligned} H(G_1^{(n)}) = \frac{d_1^{\text{CM}}}{2} n \log n + n\Big(&- s(d_1^{\text{CM}}) + H(X_1) - \mathbb{E}\left[\log X_1!\right] \\ &+ H(Q_1) + \frac{d_1^{\text{CM}}}{2} H(\Gamma_1 | \Gamma_1 \neq \circ_1)\Big) + o(n) \end{aligned} \tag{12b}$$

$$\begin{aligned} H(G_2^{(n)}) = \frac{d_2^{\text{CM}}}{2} n \log n + n\Big(&- s(d_2^{\text{CM}}) + H(X_2) - \mathbb{E}\left[\log X_2!\right] \\ &+ H(Q_2) + \frac{d_2^{\text{CM}}}{2} H(\Gamma_2 | \Gamma_2 \neq \circ_2)\Big) + o(n) \end{aligned} \tag{12c}$$

Also, it can be seen that the coefficients of n in the above equations are $\Sigma(\mu_{1,2}^{\text{CM}})$, $\Sigma(\mu_1^{\text{CM}})$ and $\Sigma(\mu_2^{\text{CM}})$, respectively.

If $\mu_{1,2}$ is any of the two distributions $\mu_{1,2}^{\text{ER}}$ or $\mu_{1,2}^{\text{CM}}$, and μ_1 and μ_2 are its marginals, we define the conditional BC entropies as $\Sigma(\mu_2|\mu_1) := \Sigma(\mu_{1,2}) - \Sigma(\mu_1)$ and $\Sigma(\mu_1|\mu_2) := \Sigma(\mu_{1,2}) - \Sigma(\mu_2)$.

5 Main Results

Now, we are ready to state our main result, which is to characterize the rate region in Definition 2. In the following, for pairs of reals (α, R) and (α', R'), we write $(\alpha, R) \succ (\alpha', R')$ if either $\alpha > \alpha'$, or $\alpha = \alpha'$ and $R > R'$. We also write $(\alpha, R) \succeq (\alpha', R')$ if either $(\alpha, R) \succ (\alpha', R')$ or $(\alpha, R) = (\alpha', R')$.

Theorem 1. *Assume $\mu_{1,2}$ is either of the two distributions $\mu_{1,2}^{ER}$ or $\mu_{1,2}^{CM}$ defined in Section 4. Then, if $\mathcal{R}$ is the rate region for the sequence of ensembles corresponding to $\mu_{1,2}$ defined in Section 2, a rate tuple $(\alpha_1, R_1, \alpha_2, R_2) \in \mathcal{R}$ if and only if*

$$(\alpha_1, R_1) \succeq ((d_{1,2} - d_2)/2, \Sigma(\mu_1|\mu_2)) \tag{13a}$$

$$(\alpha_2, R_2) \succeq ((d_{1,2} - d_1)/2, \Sigma(\mu_2|\mu_1)) \tag{13b}$$

$$(\alpha_1 + \alpha_2, R_1 + R_2) \succeq (d_{1,2}/2, \Sigma(\mu_{1,2})) \tag{13c}$$

where $d_{1,2} = \deg(\mu_{1,2})$, $d_1 = \deg(\mu_1)$ and $d_2 = \deg(\mu_2)$.

We prove the achievability for the Erdős–Rényi case and the configuration model in Sections 5.1 and 5.2, respectively. Afterwards, we prove the converse for the two cases in Sections 5.3 and 5.4, respectively. Before this, we state the following general lemma used in the proofs, whose proof is straightforward using Stirling's approximation.

Lemma 1. *Assume that a positive integer k and sequences of integers a_n and $b_1^n, \dots, b_k^n$ are given.*

1. *If $a_n/n \to a > 0$ and for each $1 \le i \le k$, $b_i^n/n \to b_i \ge 0$ where $a = \sum_{i=1}^k b_i$, we have*

$$\lim_{n\to\infty} \frac{1}{n} \log \binom{a_n}{\{b_i^n\}_{1\le i\le k}} = aH\left(\left\{\frac{b_i}{a}\right\}_{1\le i\le k}\right).$$

2. *If $a_n/\binom{n}{2} \to 1$ and $b_i^n/n \to b_i \ge 0$, we have*

$$\lim_{n\to\infty} \frac{\log \binom{a_n}{\{b_i^n\}_{1\le i\le k}} - \left(\sum_{i=1}^k b_i^n\right)\log n}{n} = \sum_{i=1}^k s(2b_i),$$

where $s(x)$ is defined to be $\frac{x}{2} - \frac{x}{2}\log x$ for $x > 0$ and 0 if $x = 0$.

5.1 Proof of Achievability for the Erdős–Rényi case

Here we show that a rate tuple $(\alpha_1, R_1, \alpha_2, R_2)$ is achievable for the Erdős–Rényi ensemble if it satisfies the following

$$(\alpha_1, R_1) \succ ((d_{1,2}^{\text{ER}} - d_2^{\text{ER}})/2, \Sigma(\mu_1^{\text{ER}}|\mu_2^{\text{ER}})) \tag{14a}$$

$$(\alpha_2, R_2) \succ ((d_{1,2}^{\text{ER}} - d_1^{\text{ER}})/2, \Sigma(\mu_2^{\text{ER}}|\mu_1^{\text{ER}})) \tag{14b}$$

$$(\alpha_1 + \alpha_2, R_1 + R_2) \succ (d_{1,2}^{\text{ER}}/2, \Sigma(\mu_{1,2}^{\text{ER}})) \tag{14c}$$

Note that if a rate tuple $(\alpha_1', R_1', \alpha_2', R_2')$ satisfies the weak inequalities (13a)–(13c) then, for any $\epsilon > 0$, $(\alpha_1', R_1' + \epsilon, \alpha_2', R_2' + \epsilon)$ satisfies the above strict inequalities. As we show below, this implies that $(\alpha_1', R_1' + \epsilon, \alpha_2', R_2' + \epsilon)$ is achievable. Hence, after sending $\epsilon \to 0$, we get $(\alpha_1', R_1', \alpha_2', R_2') \in \mathcal{R}$.

We show that any $(\alpha_1, R_1, \alpha_2, R_2)$ satisfying (14a)–(14c) is achievable by employing a random binning method. More precisely, for $i \in \{1, 2\}$, we set $L_i^{(n)} = \lfloor \exp(\alpha_i n \log n + R_i n) \rfloor$ and for each $G_i \in \mathcal{G}_i^{(n)}$, we assign $f_i^{(n)}(G_i)$ uniformly at random in the set $[L_i^{(n)}]$ and independent of everything else.

To describe our decoding scheme, we first need to define some notation. Let $\mathcal{M}^{(n)}$ denote the set of edge count vectors $\vec{m} = \{m(x)\}_{x\in\Xi_{1,2}}$ such that

$$\sum_{x\in\Xi_{1,2}} |m(x) - np_x/2| \le n^{2/3}.$$

Moreover, let $\mathcal{U}^{(n)}$ denote the set of vertex mark count vectors $\vec{u} = \{u(\theta)\}_{\theta \in \Theta_{1,2}}$ such that

$$\sum_{\theta \in \Theta_{1,2}} |u(\theta) - nq_\theta| \leq n^{2/3}.$$

Furthermore, we define $\mathcal{G}^{(n)}_{\vec{p},\vec{q}}$ to be the set of graphs $H^{(n)}_{1,2} \in \mathcal{G}^{(n)}_{1,2}$ such that $\vec{m}_{H^{(n)}_{1,2}} \in \mathcal{M}^{(n)}$ and $\vec{u}_{H^{(n)}_{1,2}} \in \mathcal{U}^{(n)}$. Upon receiving $(i,j) \in [L^{(n)}_1] \times [L^{(n)}_2]$, we form the set of graphs $H^{(n)}_{1,2} \in \mathcal{G}^{(n)}_{\vec{p},\vec{q}}$ such that $f^{(n)}_1(H^{(n)}_1) = i$ and $f^{(n)}_2(H^{(n)}_2) = j$, where $H^{(n)}_1$ and $H^{(n)}_2$ are the marginals of $H^{(n)}_{1,2}$. If this set has only one element, we output this element as the decoded graph; otherwise, we report an error.

In what follows, assume that $G^{(n)}_{1,2}$ is a random graph with law $\mathcal{G}(n;\vec{p},\vec{q})$. We consider the following four error events corresponding to the above scheme

$$\begin{aligned}
\mathcal{E}^{(n)}_1 &:= \{G^{(n)}_{1,2} \notin \mathcal{G}^{(n)}_{\vec{p},\vec{q}}\},\\
\mathcal{E}^{(n)}_2 &:= \{\exists H^{(n)}_{1,2} \neq G^{(n)}_{1,2} : f^{(n)}_i(H^{(n)}_i) = f^{(n)}_i(G^{(n)}_i), i \in \{1,2\}\},\\
\mathcal{E}^{(n)}_3 &:= \{\exists H^{(n)}_2 \neq G^{(n)}_2 : G^{(n)}_1 \oplus H^{(n)}_2 \in \mathcal{G}^{(n)}_{\vec{p},\vec{q}}, f^{(n)}_2(H^{(n)}_2) = f^{(n)}_2(G^{(n)}_2)\},\\
\mathcal{E}^{(n)}_4 &:= \{\exists H^{(n)}_1 \neq G^{(n)}_1 : H^{(n)}_1 \oplus G^{(n)}_2 \in \mathcal{G}^{(n)}_{\vec{p},\vec{q}}, f^{(n)}_1(H^{(n)}_1) = f^{(n)}_1(G^{(n)}_1)\}.
\end{aligned}$$

Note that outside the above four events, the decoder successfully decodes the input graph $G^{(n)}_{1,2}$.

Using Chebyshev's inequality, for some $\kappa > 0$ we have $\mathbb{P}(\mathcal{E}^{(n)}_1) \leq \kappa n^{-1/3}$, which converges to zero as n goes to infinity. Moreover, using the union bound, we have

$$\mathbb{P}\left(\mathcal{E}^{(n)}_2\right) \leq \frac{|\mathcal{G}^{(n)}_{\vec{p},\vec{q}}|}{L^{(n)}_1 L^{(n)}_2}. \tag{15}$$

Note that for each graph $H^{(n)}_{1,2} \in \mathcal{G}^{(n)}_{\vec{p},\vec{q}}$, the mark count vectors $\vec{m}_{H^{(n)}_{1,2}}$ and $\vec{u}_{H^{(n)}_{1,2}}$ are in the sets $\mathcal{M}^{(n)}$ and $\mathcal{U}^{(n)}$, respectively. Additionally, we have $|\mathcal{M}^{(n)}| \leq (2n^{2/3})^{|\Xi_{1,2}|}$ and $|\mathcal{U}^{(n)}| \leq (2n^{2/3})^{|\Theta_{1,2}|}$. Therefore,

$$|\mathcal{G}^{(n)}_{\vec{p},\vec{q}}| \leq (2n^{2/3})^{(|\Xi_{1,2}|+|\Theta_{1,2}|)} \max_{\substack{\vec{m}\in\mathcal{M}^{(n)}\\ \vec{u}\in\mathcal{U}^{(n)}}} A_1(\vec{m},\vec{u}), \tag{16}$$

where

$$A_1(\vec{m},\vec{u}) := \binom{n}{\{u(\theta)\}_{\theta\in\Theta_{1,2}}}\binom{\binom{n}{2}}{\{m(x)\}_{x\in\Xi_{1,2}}}.$$

Now, let $\vec{m}^{(n)}$ and $\vec{u}^{(n)}$ be sequences in $\mathcal{M}^{(n)}$ and $\mathcal{U}^{(n)}$, respectively. Then, for all $x \in \Xi_{1,2}$ and $\theta \in \Theta_{1,2}$, we have $m^{(n)}(x)/n \to p_x/2$ and $u^{(n)}(\theta)/n \to q_\theta$. Thereby, using Lemma 1, we have

$$\begin{aligned}
\lim_{n\to\infty} &\frac{\log A_1(\vec{m}^{(n)},\vec{u}^{(n)}) - (\sum_{x\in\Xi_{1,2}} m^{(n)}(x)) \log n}{n}\\
&= H(\vec{q}) + \sum_{x\in\Xi_{1,2}} s(p_x) = \Sigma(\mu^{\mathrm{ER}}_{1,2}).
\end{aligned}$$

Substituting this into (16) and using the fact that $|m^{(n)}(x) - np_x/2| \leq n^{2/3}$, we have

$$\lim_{n\to\infty} \frac{\log|\mathcal{G}^{(n)}_{\vec{p},\vec{q}}| - n\frac{d^{\mathrm{ER}}_{1,2}}{2}\log n}{n} = \Sigma(\mu^{\mathrm{ER}}_{1,2}). \tag{17}$$

Substituting this into (15), we have

$$\begin{aligned}
&\limsup \frac{1}{n} \log \mathbb{P}\left(\mathcal{E}_2^{(n)}\right) \\
&\le \limsup \frac{\log |\mathcal{G}_{\vec{p},\vec{q}}^{(n)}| - n\frac{d_{1,2}^{\mathrm{ER}}}{2}\log n - n\Sigma(\mu_{1,2}^{\mathrm{ER}})}{n} \\
&\quad + \limsup \frac{n(\frac{d_{1,2}^{\mathrm{ER}}}{2} - \alpha_1 - \alpha_2)\log n + n(\Sigma(\mu_{1,2}^{\mathrm{ER}}) - R_1 - R_2)}{n} \\
&\quad + \limsup \frac{n(\alpha_1+\alpha_2)\log n + n(R_1+R_2) - \log L_1^{(n)} L_2^{(n)}}{n}.
\end{aligned}$$

The first term is nonpositive due to (17), the second term is strictly negative due to the assumption (14c), and the third term is nonpositive due to our choice of $L_1^{(n)}$ and $L_2^{(n)}$. Consequently, the RHS is strictly negative, which implies that $\mathbb{P}(\mathcal{E}_2^{(n)}) \to 0$.

Now, we show that $\mathbb{P}(\mathcal{E}_3^{(n)} \setminus \mathcal{E}_1^{(n)})$ vanishes. In order to do so, for $H_1^{(n)} \in \mathcal{G}_1^{(n)}$, define $S_2^{(n)}(H_1^{(n)}) := \{H_2^{(n)} \in \mathcal{G}_2^{(n)} : H_1^{(n)} \oplus H_2^{(n)} \in \mathcal{G}_{\vec{p},\vec{q}}^{(n)}\}$. Using the union bound, we have

$$\begin{aligned}
\mathbb{P}\left(\mathcal{E}_3^{(n)} \setminus \mathcal{E}_1^{(n)}\right) &\le \sum_{H_{1,2}^{(n)} \in \mathcal{G}_{\vec{p},\vec{q}}^{(n)}} \mathbb{P}(G_{1,2}^{(n)} = H_{1,2}^{(n)}) \frac{|S_2^{(n)}(H_1^{(n)})|}{L_2^{(n)}} \\
&\le \frac{1}{L_2^{(n)}} \max_{H_{1,2}^{(n)} \in \mathcal{G}_{\vec{p},\vec{q}}^{(n)}} |S_2^{(n)}(H_1^{(n)})|.
\end{aligned} \tag{18}$$

It can be shown that (See Appendix B)

$$\limsup_{n\to\infty} \frac{\max_{H_{1,2}^{(n)} \in \mathcal{G}_{\vec{p},\vec{q}}^{(n)}} \log |S_2^{(n)}(H_1^{(n)})| - n\frac{d_{1,2}^{\mathrm{ER}} - d_1^{\mathrm{ER}}}{2}\log n}{n} \le \Sigma(\mu_2^{\mathrm{ER}} | \mu_1^{\mathrm{ER}}), \tag{19}$$

where $H_1^{(n)}$ is the first marginal of $H_{1,2}^{(n)}$. Substituting in (18), we get

$$\begin{aligned}
\limsup \frac{1}{n}\log \mathbb{P}\left(\mathcal{E}_3^{(n)} \setminus \mathcal{E}_1^{(n)}\right) &\le \limsup \frac{n\frac{d_{1,2}^{\mathrm{ER}} - d_1^{\mathrm{ER}}}{2}\log n + n\Sigma(\mu_2^{\mathrm{ER}}|\mu_1^{\mathrm{ER}}) - \log L_2^{(n)}}{n} \\
&\le \limsup \frac{n(\frac{d_{1,2}^{\mathrm{ER}} - d_1^{\mathrm{ER}}}{2} - \alpha_2)\log n + n(\Sigma(\mu_2^{\mathrm{ER}}|\mu_1^{\mathrm{ER}}) - R_2)}{n} \\
&\quad + \limsup \frac{n\alpha_2 \log n + nR_2 - \log L_2^{(n)}}{n}.
\end{aligned} \tag{20}$$

Note that the first term is strictly negative due to the assumption (14b), while the second term is nonpositive due to our way of choosing $L_2^{(n)}$. This means that $\mathbb{P}(\mathcal{E}_3^{(n)} \setminus \mathcal{E}_1^{(n)})$ goes to zero as n goes to infinity. Similarly, $\mathbb{P}(\mathcal{E}_4^{(n)} \setminus \mathcal{E}_1^{(n)})$ converges to zero. This means that there exists a sequence of deterministic codebooks with vanishing probability of error, which completes the proof of achievability.

5.2 Proof of Achievability for the Configuration model

Our achievability proof for this case is very similar in nature to that for the Erdős–Rényi case, with the modifications discussed below.

Let $\mathcal{D}^{(n)}$ be the set of degree sequences $\vec{d}$ with entries bounded by Δ such that $c_k(\vec{d}) = c_k(\vec{d}^{(n)})$ for all $0 \leq k \leq \Delta$. Moreover, redefine $\mathcal{M}^{(n)}$ to be the set of mark count vectors $\vec{m}$ such that $\sum_{x\in\Xi_{1,2}} m(x) = m_n$ and $\sum_{x\in\Xi_{1,2}} |m(x) - m_n\gamma_x| \leq n^{2/3}$, where recall that $m_n = (\sum_{i=1}^n d^{(n)}(i))/2$. We use the same definition for $\mathcal{U}^{(n)}$ as in the previous section, i.e. the set of vertex mark count vectors $\vec{u}$ such that $\sum_{\theta\in\Theta_{1,2}} |u(\theta) - nq_\theta| \leq n^{2/3}$.

In what follows, let X be a random variable with law $\vec{r}$, X_1 and X_2 defined as in (11) and $\Gamma = (\Gamma_1, \Gamma_2)$ a random variable with law $\vec{\gamma}$.

We define $\mathcal{W}^{(n)}$ to be the set of graphs $H^{(n)}_{1,2} \in \mathcal{G}^{(n)}_{1,2}$ such that: *(i)* $\overrightarrow{\mathrm{dg}}_{H^{(n)}_{1,2}} \in \mathcal{D}^{(n)}$, *(ii)* $\vec{m}_{H^{(n)}_{1,2}} \in \mathcal{M}^{(n)}$, *(iii)* $\vec{u}_{H^{(n)}_{1,2}} \in \mathcal{U}^{(n)}$, *(iv)* for all $0 \leq l \leq k \leq \Delta$, recalling the notation in (2), we have

$$|c_{k,l}(\overrightarrow{\mathrm{dg}}_{H^{(n)}_{1,2}}, \overrightarrow{\mathrm{dg}}_{H^{(n)}_1}) - n\mathbb{P}\left(X = k, X_1 = l\right)| \leq n^{2/3}, \tag{21}$$

and *(v)*, for all $0 \leq l \leq k \leq \Delta$, we have

$$|c_{k,l}(\overrightarrow{\mathrm{dg}}_{H^{(n)}_{1,2}}, \overrightarrow{\mathrm{dg}}_{H^{(n)}_2}) - n\mathbb{P}\left(X = k, X_2 = l\right)| \leq n^{2/3}. \tag{22}$$

We employ a similar random binning framework as in Section 5.1. For decoding, upon receiving a pair (i,j), we form the set of graphs $H^{(n)}_{1,2} \in \mathcal{W}^{(n)}$ such that $f^{(n)}_1(H^{(n)}_1) = i$ and $f^{(n)}_2(H^{(n)}_2) = j$. If this set has only one element, we output it as the source graph; otherwise, we output an indication of error. In order to prove the achievability, we consider the four error events $\mathcal{E}^{(n)}_i$, $1 \leq i \leq 4$, exactly as those in the previous section, with $\mathcal{G}^{(n)}_{\vec{p},\vec{q}}$ being replaced with $\mathcal{W}^{(n)}$.

It can be shown that if $G^{(n)}_{1,2} \sim \mathcal{G}(n; \vec{d}^{(n)}, \vec{\gamma}, \vec{q})$, the probability of $G^{(n)}_{1,2} \in \mathcal{W}^{(n)}$ goes to one as n goes to infinity (see Lemma 4 in Appendix A). Therefore, $\mathbb{P}(\mathcal{E}^{(n)}_1)$ goes to zero.

To show that $\mathbb{P}(\mathcal{E}^{(n)}_2)$ vanishes, similar to the analysis in Section 5.1, we find an asymptotic upper bound for $\log|\mathcal{W}^{(n)}|$. By only considering the conditions *(i)*, *(ii)* and *(iii)* in the definition of $\mathcal{W}^{(n)}$, we have

$$\begin{aligned}\log|\mathcal{W}^{(n)}| \leq \log\binom{n}{\{c_k(\vec{d}^{(n)})\}_{k=0}^{\Delta}} + \log|\mathcal{G}^{(n)}_{\vec{d}^{(n)}}| \\ + \log\left((2n^{2/3})^{|\Xi_{1,2}|}\max_{\vec{m}\in\mathcal{M}^{(n)}}\binom{m_n}{\{m(x)\}_{x\in\Xi_{1,2}}}\right) \\ + \log\left((2n^{2/3})^{|\Theta_{1,2}|}\max_{\vec{u}\in\mathcal{U}^{(n)}}\binom{n}{\{u(\theta)\}_{\theta\in\Theta_{1,2}}}\right).\end{aligned} \tag{23}$$

By assumption, we have $r_0 < 1$, hence $d^{\mathrm{CM}}_{1,2} > 0$. The condition (8) together with Lemma 3 then implies that

$$\lim_{n\to\infty}\frac{\log|\mathcal{G}^{(n)}_{\vec{d}^{(n)}}| - n\frac{d^{\mathrm{CM}}_{1,2}}{2}\log n}{n} = -s(d^{\mathrm{CM}}_{1,2}) - \mathbb{E}\left[\log X!\right]. \tag{24}$$

Using this together with Lemma 1 for the other terms in (23), we have

$$\begin{aligned}\limsup_{n\to\infty}\frac{\log|\mathcal{W}^{(n)}| - n\frac{d^{\mathrm{CM}}_{1,2}}{2}\log n}{n} \leq -s(d^{\mathrm{CM}}_{1,2}) + H(X) \\ + \frac{d^{\mathrm{CM}}_{1,2}}{2}H(\Gamma) + H(Q) - \mathbb{E}\left[\log X!\right] = \Sigma(\mu^{\mathrm{CM}}_{1,2}),\end{aligned}$$

where Γ and Q are random variables with law $\vec{\gamma}$ and $\vec{q}$, respectively.

Now, in order to show that $\mathbb{P}(\mathcal{E}_3^{(n)} \setminus \mathcal{E}_1^{(n)})$ vanishes, we prove a counterpart for (19). For $H_1^{(n)} \in \mathcal{G}_1^{(n)}$, we define $S_2^{(n)}(H_1^{(n)})$ to be the set of graphs $H_2^{(n)} \in \mathcal{G}_2^{(n)}$ such that $H_1^{(n)} \oplus H_2^{(n)} \in \mathcal{W}^{(n)}$. Then, it can be shown that (see Appendix C)

$$\limsup_{n\to\infty} \frac{\max_{H_{1,2}^{(n)} \in \mathcal{W}^{(n)}} \log |S_2^{(n)}(H_1^{(n)})| - n\frac{d_{1,2}^{\mathrm{CM}} - d_1^{\mathrm{CM}}}{2} \log n}{n} \leq \Sigma(\mu_2^{\mathrm{CM}} | \mu_1^{\mathrm{CM}}). \tag{25}$$

Then, similar to (20), this shows that $\mathbb{P}(\mathcal{E}_3^{(n)} \setminus \mathcal{E}_1^{(n)})$ vanishes. The proof for $\mathbb{P}(\mathcal{E}_4^{(n)} \setminus \mathcal{E}_1^{(n)})$ is similar. This completes the proof of achievability.

5.3 Proof of the Converse for the Erdős–Rényi case

In this section, we show that every rate tuple $(\alpha_1, R_1, \alpha_2, R_2) \in \mathcal{R}$ for the Erdős–Rényi scenario must satisfy the conditions (13a)–(13c). By definition, for a rate tuple $(\alpha_1, R_1, \alpha_2, R_2) \in \mathcal{R}$, there exist sequences $R_1^{(m)}$ and $R_2^{(m)}$ such that for each m, $(\alpha_1, R_1^{(m)}, \alpha_2, R_2^{(m)})$ is achievable, and besides we have $R_1^{(m)} \to R_1$ and $R_2^{(m)} \to R_2$. If we show that $(\alpha_1, R_1^{(m)}, \alpha_2, R_2^{(m)})$ satisfies (13a)–(13c) for each m, it is easy to see that $(\alpha_1, R_1, \alpha_2, R_2)$ must also satisfy the same inequalities. Therefore, it suffices to show that any achievable rate tuple satisfies (13a)–(13c).

For this, take an achievable rate tuple $(\alpha_1, R_1, \alpha_2, R_2)$ together with a corresponding sequence of $\langle n, L_1^{(n)}, L_2^{(n)} \rangle$ codes $(f_1^{(n)}, f_2^{(n)}, g^{(n)})$. By definition, we have

$$\limsup_{n\to\infty} \frac{\log L_i^{(n)} - (\alpha_i n \log n + R_i n)}{n} \leq 0 \qquad i \in \{1,2\}, \tag{26}$$

and also the error probability $P_e^{(n)}$ goes to zero as n goes to infinity. Now, we define the set $\mathcal{A}^{(n)} \subseteq \mathcal{G}_{1,2}^{(n)}$ as

$$\mathcal{A}^{(n)} := \mathcal{G}_{\vec{p},\vec{q}}^{(n)} \cap \{H_{1,2}^{(n)} \in \mathcal{G}_{1,2}^{(n)} : g^{(n)}(f_1^{(n)}(H_1^{(n)}), f_2^{(n)}(H_2^{(n)})) = H_{1,2}^{(n)}\}, \tag{27}$$

where $\mathcal{G}_{\vec{p},\vec{q}}^{(n)}$ was defined in Section 5.1. In fact, $\mathcal{A}^{(n)}$ is the set of "typical" graphs with respect to the Erdős–Rényi model that are successfully decoded by the code $(f_1^{(n)}, f_2^{(n)}, g^{(n)})$. In the following, let $G_{1,2}^{(n)} \sim \mathcal{G}^{(n)}(n; \vec{p}, \vec{q})$ be distributed according to the Erdős–Rényi model. Moreover, let $P_{\mathrm{ER}}^{(n)}$ be the law of $G_{1,2}^{(n)}$, i.e. for $H_{1,2}^{(n)} \in \mathcal{G}_{1,2}^{(n)}$, $P_{\mathrm{ER}}^{(n)}(H_{1,2}^{(n)}) := \mathbb{P}(G_{1,2}^{(n)} = H_{1,2}^{(n)})$. With this, we define a random variable $\tilde{G}_{1,2}^{(n)}$ whose distribution is the conditional distribution of $G_{1,2}^{(n)}$, conditioned on lying in $\mathcal{A}^{(n)}$, i.e.

$$\mathbb{P}\left(\tilde{G}_{1,2}^{(n)} = H_{1,2}^{(n)}\right) = \begin{cases} P_{\mathrm{ER}}^{(n)}(H_{1,2}^{(n)})/\pi_n & H_{1,2}^{(n)} \in \mathcal{A}^{(n)} \\ 0 & \text{o.t.w.} \end{cases} \tag{28}$$

where $\pi_n := \mathbb{P}\left(G_{1,2}^{(n)} \in \mathcal{A}^{(n)}\right)$ is the normalizing factor. Additionally, let $\tilde{P}_{\mathrm{ER}}^{(n)}$ be the law of $\tilde{G}_{1,2}^{(n)}$. If, for $i \in \{1,2\}$, $\tilde{M}_i^{(n)}$ denotes $f_i^{(n)}(\tilde{G}_i^{(n)})$, we have

$$\begin{aligned} \log L_1^{(n)} + \log L_2^{(n)} &\geq H(\tilde{M}_1^{(n)}) + H(\tilde{M}_2^{(n)}) \geq H(\tilde{M}_1^{(n)}, \tilde{M}_2^{(n)}) \\ &= H(\tilde{G}_{1,2}^{(n)}), \end{aligned} \tag{29}$$

where the last equality follows from the fact that, by definition, $\tilde{G}_{1,2}^{(n)}$ takes values among the graphs that are successfully decoded, and hence is uniquely identified given $\tilde{M}_1^{(n)}$ and $\tilde{M}_2^{(n)}$.

Now, we find a lower bound for $H(\tilde{G}_{1,2}^{(n)})$. For doing so, note that for $H_{1,2}^{(n)} \in \mathcal{G}_{1,2}^{(n)}$ and n large enough, we have

$$
\begin{aligned}
-\log P_{\mathrm{ER}}^{(n)}(H_{1,2}^{(n)}) = &- \sum_{x \in \Xi_{1,2}} m_{H_{1,2}^{(n)}}(x) \log \frac{p_x}{n} - \left[\binom{n}{2} - \sum_{x \in \Xi_{1,2}} m_{H_{1,2}^{(n)}}(x) \right] \log \left(1 - \frac{\sum_{x \in \Xi_{1,2}} p_x}{n} \right) \\
&- \sum_{\theta \in \Theta_{1,2}} u_{H_{1,2}^{(n)}}(\theta) \log q_\theta .
\end{aligned} \tag{30}
$$

On the other hand, due to the definition of $\mathcal{G}_{\vec{p},\vec{q}}^{(n)}$, if $H_{1,2}^{(n)} \in \mathcal{G}_{\vec{p},\vec{q}}^{(n)}$ then, for all $x \in \Xi_{1,2}$ and $\theta \in \Theta_{1,2}$, we have

$$
\begin{aligned}
n\frac{p_x}{2} - n^{2/3} \le m_{H_{1,2}^{(n)}}(x) \le n\frac{p_x}{2} + n^{2/3}, \text{ and} \\
nq_\theta - n^{2/3} \le u_{H_{1,2}^{(n)}}(\theta) \le nq_\theta + n^{2/3}.
\end{aligned}
$$

Substituting these in (30) and using the inequality $\log(1-x) \le -x$ which holds for $x \in (0,1)$, for n large enough, we have

$$
\begin{aligned}
-\log P_{\mathrm{ER}}^{(n)}(H_{1,2}^{(n)}) \ge &\sum_{x \in \Xi_{1,2}} \left(n\frac{p_x}{2} - n^{2/3} \right) (\log n - \log p_x) + \left[\binom{n}{2} - \sum_{x \in \Xi_{1,2}} \left(n\frac{p_x}{2} + n^{2/3} \right) \right] \frac{\sum_{x \in \Xi_{1,2}} p_x}{n} \\
&- \sum_{\theta \in \Theta_{1,2}} (nq_\theta - n^{2/3}) \log q_\theta .
\end{aligned}
$$

Using $\sum_{x \in \Xi_{1,2}} p_x = d_{1,2}^{\mathrm{ER}}$ and simplifying the above, we realize that there exists a constant $c > 0$ that does not depend on n or $H_{1,2}^{(n)}$, such that for all $H_{1,2}^{(n)} \in \mathcal{A}^{(n)}$, we have

$$
\begin{aligned}
-\log P_{\mathrm{ER}}^{(n)}(H_{1,2}^{(n)}) &\ge n\frac{d_{1,2}^{\mathrm{ER}}}{2} \log n - n \sum_{x \in \Xi_{1,2}} \frac{p_x}{2} \log p_x + n \sum_{x \in \Xi_{1,2}} \frac{p_x}{2} - n \sum_{\theta \in \Theta_{1,2}} q_\theta \log q_\theta - cn^{2/3} \log n \\
&= n\frac{d_{1,2}^{\mathrm{ER}}}{2} \log n + n\Sigma(\mu_{1,2}^{\mathrm{ER}}) - cn^{2/3} \log n.
\end{aligned} \tag{31}
$$

Now, if $\tilde{G}_{1,2}^{(n)}$ is the random variable defined in (28), we have

$$
\begin{aligned}
H(\tilde{G}_{1,2}^{(n)}) &= - \sum_{H_{1,2}^{(n)} \in \mathcal{A}^{(n)}} \tilde{P}_{\mathrm{ER}}^{(n)}(H_{1,2}^{(n)}) \log \tilde{P}_{\mathrm{ER}}^{(n)}(H_{1,2}^{(n)}) \\
&= \log \pi_n - \frac{1}{\pi_n} \sum_{H_{1,2}^{(n)} \in \mathcal{A}^{(n)}} P_{\mathrm{ER}}^{(n)}(H_{1,2}^{(n)}) \log P_{\mathrm{ER}}^{(n)}(H_{1,2}^{(n)}).
\end{aligned}
$$

Note that since the probability of error of the above code vanishes, i.e. $P_e^{(n)} \to 0$, and $\mathbb{P}\left(G_{1,2}^{(n)} \in \mathcal{G}_{\vec{p},\vec{q}}^{(n)} \right) \to 1$, we have $\pi_n \to 1$ as $n \to \infty$. On the other hand, with probability one, we have $\tilde{G}_{1,2}^{(n)} \in \mathcal{G}_{\vec{p},\vec{q}}^{(n)}$. Also, by the definition of π_n, we have $\sum_{H_{1,2}^{(n)} \in \mathcal{A}^{(n)}} P_{\mathrm{ER}}^{(n)}(H_{1,2}^{(n)}) = \pi_n$. Thereby, employing the bound (31), we have

$$
\liminf_{n \to \infty} \frac{H(\tilde{G}_{1,2}^{(n)}) - n\frac{d_{1,2}^{\mathrm{ER}}}{2} \log n}{n} \ge \Sigma(\mu_{1,2}^{\mathrm{ER}}). \tag{32}
$$

Now, using the assumption (26) together with the bound (29), we have

$$
\begin{aligned}
0 &\geq \limsup_{n\to\infty} \frac{\log L_1^{(n)} + \log L_2^{(n)} - (\alpha_1+\alpha_2) n \log n - n(R_1+R_2)}{n} \\
&\geq \liminf_{n\to\infty} \frac{H(\tilde{G}_{1,2}^{(n)}) - n\frac{d_{1,2}^{\mathrm{ER}}}{2}\log n - n\Sigma(\mu_{1,2}^{\mathrm{ER}})}{n} + \liminf_{n\to\infty} \frac{n\frac{d_{1,2}^{\mathrm{ER}}}{2}\log n + n\Sigma(\mu_{1,2}^{\mathrm{ER}}) - (\alpha_1+\alpha_2)n\log n - n(R_1+R_2)}{n}.
\end{aligned} \tag{33}
$$

The first term is nonnegative due to (32). Consequently,

$$
0 \geq \liminf_{n\to\infty} \frac{n\left(\frac{d_{1,2}^{\mathrm{ER}}}{2} - \alpha_1 - \alpha_2\right)\log n + n(\Sigma(\mu_{1,2}^{\mathrm{ER}}) - R_1 - R_2)}{n}. \tag{34}
$$

Note that this is impossible unless $\alpha_1 + \alpha_2 \geq d_{1,2}^{\mathrm{ER}}/2$. Furthermore, if $\alpha_1 + \alpha_2 = d_{1,2}^{\mathrm{ER}}$, it must be the case that $R_1 + R_2 \geq \Sigma(\mu_{1,2}^{\mathrm{ER}})$. But this is precisely (13c) for $\mu_{1,2} = \mu_{1,2}^{\mathrm{ER}}$.

Now, we turn to showing (13a). We have

$$
\begin{aligned}
\log L_1^{(n)} &\geq H(\tilde{M}_1^{(n)}) \geq H(\tilde{M}_1^{(n)}|\tilde{M}_2^{(n)}) \\
&= H(\tilde{G}_1^{(n)}, \tilde{M}_1^{(n)}|\tilde{M}_2^{(n)}) - H(\tilde{G}_1^{(n)}|\tilde{M}_1^{(n)}, \tilde{M}_2^{(n)}) \\
&\overset{(a)}{=} H(\tilde{G}_1^{(n)}|\tilde{M}_2^{(n)}) \\
&\overset{(b)}{\geq} H(\tilde{G}_1^{(n)}|\tilde{G}_2^{(n)}) \\
&= H(\tilde{G}_{1,2}^{(n)}) - H(\tilde{G}_2^{(n)}).
\end{aligned} \tag{35}
$$

where (a) uses the facts that $\tilde{M}_1^{(n)}$ is a function of $\tilde{G}_1^{(n)}$ and also, since $\tilde{G}_{1,2}^{(n)} \in \mathcal{A}^{(n)}$, given $\tilde{M}_1^{(n)}$ and $\tilde{M}_2^{(n)}$ we can unambiguously determine $\tilde{G}_{1,2}^{(n)}$ and hence $\tilde{G}_1^{(n)}$. Also, (b) uses data processing inequality. Now, we find an upper bound for $H(\tilde{G}_2^{(n)})$. Note that since $\tilde{G}_{1,2}^{(n)} \in \mathcal{A}^{(n)}$ with probability one, we have

$$
H(\tilde{G}_2^{(n)}) \leq \log |\mathcal{A}_2^{(n)}|, \tag{36}
$$

where

$$
\mathcal{A}_2^{(n)} := \{H_2^{(n)} \in \mathcal{G}_2^{(n)} : H_1^{(n)} \oplus H_2^{(n)} \in \mathcal{A}^{(n)} \text{ for some } H_1^{(n)} \in \mathcal{G}_1^{(n)}\}.
$$

But for $H_{1,2}^{(n)} := H_1^{(n)} \oplus H_2^{(n)} \in \mathcal{A}_2^{(n)}$, since $\mathcal{A}_2^{(n)} \subseteq \mathcal{G}_{\vec{p},\vec{q}}^{(n)}$, by definition we have that for all $x \in \Xi_{1,2}$ and all $\theta \in \Theta_{1,2}$,

$$
\sum_{x\in\Xi_{1,2}} |m_{H_{1,2}^{(n)}}(x) - np_x/2| \leq n^{2/3} \text{ and } \sum_{\theta\in\Theta_{1,2}} |u_{H_{1,2}^{(n)}}(\theta) - nq_\theta| \leq n^{2/3}.
$$

Moreover, for $x_2 \in \Xi_2$ and $\theta_2 \in \Theta_2$, we have $m_{H_2^{(n)}}(x_2) = \sum_{x_1 \in \Xi_1 \cup \{\circ_1\}} m_{H_{1,2}^{(n)}}((x_1, x_2))$ and $u_{H_2^{(n)}}(\theta_2) = \sum_{\theta_1\in\Theta_1} m_{H_{1,2}^{(n)}}((\theta_1, \theta_2))$. Using this in the above and using triangle inequality, we realize that for $H_2^{(n)} \in \mathcal{A}_2^{(n)}$, we have $\vec{m}_{H_2^{(n)}} \in \mathcal{M}_2^{(n)}$ and $\vec{u}_{H_2^{(n)}} \in \mathcal{U}_2^{(n)}$, where $\mathcal{M}_2^{(n)}$ is the set of edge mark count vectors $\vec{m}$ such that $\sum_{x_2\in\Xi_2} |m(x_2) - np_{x_2}/2| \leq n^{2/3}$, and $\mathcal{U}_2^{(n)}$ is the set of vertex mark count vectors $\vec{u}$ such that $\sum_{\theta_2\in\Theta_2} |u(\theta_2) - nq_{\theta_2}| \leq n^{2/3}$. Consequently,

$$
|\mathcal{A}_2^{(n)}| \leq (2n^{2/3})^{(|\Xi_2|+|\Theta_2|)} \max_{\vec{m}\in\mathcal{M}_2^{(n)}} \binom{\binom{n}{2}}{\{m(x_2)\}_{x_2\in\Xi_2}} \max_{\vec{u}\in\mathcal{U}_2^{(n)}} \binom{n}{\{u(\theta_2)\}_{\theta_2\in\Theta_2}}.
$$

Using Lemma 1 and the definition of $\mathcal{M}_2^{(n)}$ and $\mathcal{U}_2^{(n)}$ above, with $Q = (Q_1, Q_2) \sim \vec{q}$, we get

$$\limsup_{n\to\infty} \frac{\log |\mathcal{A}_2^{(n)}| - n\frac{d_2^{\mathrm{ER}}}{2}\log n}{n} \le H(Q_2) + \sum_{x_2\in\Xi_2} s(p_{x_2}) = \Sigma(\mu_2^{\mathrm{ER}}).$$

Substituting into (36), we get

$$\limsup_{n\to\infty} \frac{\log H(\tilde{G}_2^{(n)}) - n\frac{d_2^{\mathrm{ER}}}{2}\log n}{n} \le \Sigma(\mu_2^{\mathrm{ER}}).$$

Using this together with (32) and substituting into (35) we get

$$\liminf_{n\to\infty} \frac{\log L_1^{(n)} - n\frac{d_{1,2}^{\mathrm{ER}} - d_2^{\mathrm{ER}}}{2}\log n}{n} \ge \Sigma(\mu_{1,2}^{\mathrm{ER}}) - \Sigma(\mu_2^{\mathrm{ER}}) = \Sigma(\mu_1^{\mathrm{ER}}|\mu_2^{\mathrm{ER}}).$$

Using a similar method as in (33) and (34), this implies (13a). The proof of (13b) is similar. This completes the proof of the converse for the Erdős–Rényi case.

5.4 Proof of the Converse for the configuration model

The proof of the converse for the configuration model is similar to that for the Erdős–Rényi model presented in the previous section. Take an achievable rate tuple $(\alpha_1, R_1, \alpha_2, R_2)$ together with a sequence of $\langle n, L_1^{(n)}, L_2^{(n)}\rangle$ codes $(f_1^{(n)}, f_2^{(n)}, g^{(n)})$ achieving this rate tuple. Moreover, redefine the set $\mathcal{A}^{(n)}$ to be

$$\mathcal{A}^{(n)} := \mathcal{W}^{(n)} \cap \{H_{1,2}^{(n)} \in \mathcal{G}_{1,2}^{(n)} : g^{(n)}(f_1^{(n)}(H_1^{(n)}), f_2^{(n)}(H_2^{(n)})) = H_{1,2}^{(n)}\}, \tag{37}$$

where the set $\mathcal{W}^{(n)}$ was defined in Section 5.2. Now, let $G_{1,2}^{(n)} \sim \mathcal{G}(n; \vec{d}^{(n)}, \vec{\gamma}, \vec{q})$ be distributed according to the configuration model ensemble, and let $\tilde{G}_{1,2}^{(n)} \in \mathcal{A}^{(n)}$ have the distribution obtained from that of $G_{1,2}^{(n)}$ by conditioning on it lying in the set $\mathcal{A}^{(n)}$. Note that the normalizing constant $\pi_n := \mathbb{P}(G_{1,2}^{(n)} \in \mathcal{A}^{(n)})$ goes to 1 as $n \to \infty$ since $\mathbb{P}(G_{1,2}^{(n)} \in \mathcal{W}^{(n)}) \to 1$ and the error probability of the code, $P_e^{(n)}$, vanishes. Moreover, let $P_{\mathrm{CM}}^{(n)}$ and $\tilde{P}_{\mathrm{CM}}^{(n)}$ be the laws of $G_{1,2}^{(n)}$ and $\tilde{G}_{1,2}^{(n)}$, respectively. In the following, we show that

$$\liminf_{n\to\infty} \frac{H(\tilde{G}_{1,2}^{(n)}) - n\frac{d_{1,2}^{\mathrm{CM}}}{2}\log n}{n} \ge \Sigma(\mu_{1,2}^{\mathrm{CM}}), \tag{38}$$

and

$$\limsup_{n\to\infty} \frac{H(\tilde{G}_2^{(n)}) - n\frac{d_2^{\mathrm{CM}}}{2}\log n}{n} \le \Sigma(\mu_2^{\mathrm{CM}}). \tag{39}$$

The rest of the proof is then identical to that of the previous section, so we only focus on proving the above two statements.

For (38), note that for $H_{1,2}^{(n)} \in \mathcal{G}_{1,2}^{(n)}$ such that $\overrightarrow{\mathrm{dg}}_{H_{1,2}^{(n)}} \in \mathcal{D}^{(n)}$, where $\mathcal{D}^{(n)}$ was defined in Section 5.2, we have

$$-\log P_{\mathrm{CM}}^{(n)}(H_{1,2}^{(n)}) = \log \binom{n}{\{c_k(\vec{d}^{(n)})\}_{k=0}^{\Delta}} + \log |\mathcal{G}_{\vec{d}^{(n)}}^{(n)}| - \sum_{x\in\Xi_{1,2}} m_{H_{1,2}^{(n)}}(x)\log\gamma_x - \sum_{\theta\in\Theta_{1,2}} u_{H_{1,2}^{(n)}}(\theta)\log q_\theta.$$

Now, if $H_{1,2}^{(n)} \in \mathcal{W}^{(n)}$, using the definition of $\mathcal{W}^{(n)}$ we realize that there exists a constant $c > 0$ such that

$$-\log P_{\text{CM}}^{(n)}(H_{1,2}^{(n)}) \geq \log \binom{n}{\{c_k(\vec{d}^{(n)})\}_{k=0}^{\Delta}} + \log |\mathcal{G}_{\vec{d}^{(n)}}^{(n)}| - \sum_{x \in \Xi_{1,2}} m_n \gamma_x \log \gamma_x - \sum_{\theta \in \Theta_{1,2}} n q_\theta \log q_\theta - cn^{2/3} =: K_n.$$

Note that the right hand side is a constant independent of $H_{1,2}^{(n)}$ and is denoted by K_n. Since $\tilde{G}_{1,2}^{(n)}$ falls in $\mathcal{W}^{(n)}$ with probability one, this means that $H(\tilde{G}_{1,2}^{(n)}) \geq \log \pi_n + K_n$. But $\pi_n \to 1$ as $n \to \infty$. Therefore, using the assumption (8) together with (24) from Section 5.2 and also the fact that $m_n/n \to d_{1,2}^{\text{CM}}/2$, we realize that

$$\liminf_{n \to \infty} \frac{H(\tilde{G}_{1,2}^{(n)}) - n \frac{d_{1,2}^{\text{CM}}}{2} \log n}{n} \geq H(X) - s(d_{1,2}^{\text{CM}}) - \mathbb{E}\left[\log X!\right] + \frac{d_{1,2}^{\text{CM}}}{2} H(\Gamma) + H(Q),$$

where $X \sim \vec{r}$, $\Gamma \sim \vec{\gamma}$ and $Q \sim \vec{q}$. Note that the right hand side is precisely $\Sigma(\mu_{1,2}^{\text{CM}})$, hence we have proved (38).

In order to show (39), note that $H(\tilde{G}_2^{(n)}) \leq \log |\mathcal{A}_2^{(n)}|$ where $\mathcal{A}_2^{(n)}$ consists of graphs $H_2^{(n)} \in \mathcal{G}_2^{(n)}$ such that for some $H_1^{(n)} \in \mathcal{G}_1^{(n)}$, we have $H_1^{(n)} \oplus H_2^{(n)} \in \mathcal{A}^{(n)}$. Since $\mathcal{A}^{(n)} \subseteq \mathcal{W}^{(n)}$, for all $H_2^{(n)} \in \mathcal{A}_2^{(n)}$ we have

$$\sum_{x_2 \in \Xi_2} |m_{H_2^{(n)}}(x_2) - m_n \gamma_{x_2}| \leq n^{2/3} \text{ and } \sum_{\theta_2 \in \Theta_2} |u_{H_2^{(n)}}(\theta_2) - n q_{\theta_2}| \leq n^{2/3}. \tag{40}$$

On the other hand, the condition (22) implies that $\overrightarrow{\text{dg}}_{H_2^{(n)}} \in \mathcal{D}_2^{(n)}$ where $\mathcal{D}_2^{(n)}$ denotes the set of degree sequences $\vec{d}$ of size n with elements bounded by Δ such that

$$|c_k(\vec{d}) - n\mathbb{P}\left(X_2 = k\right)| \leq (\Delta + 1) n^{2/3} \qquad \forall 0 \leq k \leq \Delta, \tag{41}$$

where X_2 is the random variable defined in (11). Consequently, we have

$$\begin{aligned} \log |\mathcal{A}_2^{(n)}| \leq \log |\mathcal{D}_2^{(n)}| + \max_{\vec{d} \in \mathcal{D}_2^{(n)}} \log |\mathcal{G}_{\vec{d}}^{(n)}| + \max_{H_2^{(n)} \in \mathcal{A}_2^{(n)}} \log \binom{\sum_{x_2 \in \Xi_2} m_{H_2^{(n)}}(x_2)}{\{m_{H_2^{(n)}}(x_2)\}_{x_2 \in \Xi_2}} \\ + \max_{H_2^{(n)} \in \mathcal{A}_2^{(n)}} \log \binom{n}{\{u_{H_2^{(n)}}(\theta_2)\}_{\theta_2 \in \Theta_2}}. \end{aligned} \tag{42}$$

Note that (41) implies that $|\mathcal{D}_2^{(n)}| \leq (2(\Delta+1)n^{2/3})^{\Delta+1} \max_{\vec{d} \in \mathcal{D}_2^{(n)}} \binom{n}{\{c_k(\vec{d})\}_{k=0}^{\Delta}}$. Therefore, Lemma 1 implies

$$\limsup_{n \to \infty} \frac{1}{n} \log |\mathcal{D}_2^{(n)}| \leq H(X_2). \tag{43}$$

On the other hand, the assumptions $r_0 < 1$ and (7) imply that $d_2^{\text{CM}} > 0$. Hence, using Lemma 3 in Appendix A we have

$$\limsup_{n \to \infty} \frac{\max_{\vec{d} \in \mathcal{D}_2^{(n)}} \log |\mathcal{G}_{\vec{d}}^{(n)}| - n \frac{d_2^{\text{CM}}}{2} \log n}{n} \leq -s(d_2^{\text{CM}}) - \mathbb{E}\left[\log X_2!\right]. \tag{44}$$

Moreover, if $H_2^{(n)}$ is a sequence in $\mathcal{A}_2^{(n)}$, from (40), for all $x_2 \in \Xi_2$, we have

$$\lim_{n \to \infty} \frac{m_{H_2^{(n)}}(x_2)}{\sum_{x_2' \in \Xi_2} m_{H_2^{(n)}}(x_2')} = \frac{\gamma_{x_2}}{\sum_{x_2' \in \Xi_2} \gamma_{x_2'}} = \mathbb{P}\left(\Gamma_2 = x_2 | \Gamma_2 \neq \circ_2\right),$$

where $\Gamma = (\Gamma_1, \Gamma_2)$ has law $\vec{\gamma}$. Additionally, we have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{x_2 \in \Xi_2} m_{H_2^{(n)}}(x_2) = \frac{d_2^{\mathrm{CM}}}{2}.$$

Thereby, from Lemma 1, we have

$$\limsup_{n\to\infty} \frac{1}{n} \max_{H_2^{(n)} \in \mathcal{A}_2^{(n)}} \log \binom{\sum_{x_2\in\Xi_2} m_{H_2^{(n)}}(x_2)}{\{m_{H_2^{(n)}}(x_2)\}_{x_2\in\Xi_2}} \le \frac{d_2^{\mathrm{CM}}}{2} H(\Gamma_2 | \Gamma_2 \neq \circ_2). \tag{45}$$

Finally, as we have $u_{H_2^{(n)}}(\theta_2)/n \to q_{\theta_2}$ for all $\theta_2 \in \Theta_2$, another usage of Lemma 1 implies that

$$\limsup_{n\to\infty} \frac{1}{n} \max_{H_2^{(n)} \in \mathcal{A}_2^{(n)}} \log \binom{n}{\{u_{H_2^{(n)}}(\theta_2)\}_{\theta_2\in\Theta_2}} \le H(Q_2), \tag{46}$$

where $Q = (Q_1, Q_2)$ has law $\vec{q}$. Now, combining (43), (44), (45) and (46) and substituting into (42), and also using the bound $H(\tilde{G}_2^{(n)}) \le \log |\mathcal{A}_2^{(n)}|$, we realize that

$$\limsup_{n\to\infty} \frac{H(\tilde{G}_2^{(n)}) - n\frac{d_2^{\mathrm{CM}}}{2} \log n}{n} \le H(X_2) - s(d_2^{\mathrm{CM}}) - \mathbb{E}\left[\log X_2!\right] + \frac{d_2^{\mathrm{CM}}}{2} H(\Gamma_2|\Gamma_2 \neq \circ_2) + H(Q_2).$$

But the right hand side is precisely $\Sigma(\mu_2^{\mathrm{CM}})$. This completes the proof of (39). As was mentioned before, the rest of the proof is identical to that in the previous section.

6 Conclusion

We gave a counterpart of the Slepian–Wolf Theorem for graphical data, employing the framework of local weak convergence. We derived the rate region for two graph ensembles, namely an Erdős–Rényi model and a configuration model.

A Asymptotic behavior of the entropy of the configuration model

Here, we prove (12a)–(12c). Before this, we set some notation and state some general lemmas. In what follows, we employ the definitions of the sets $\mathcal{W}^{(n)}$ and $\mathcal{D}^{(n)}$ from Section 5.2. Moreover, $\Gamma = (\Gamma_1, \Gamma_2)$ and $Q = (Q_1, Q_2)$ be random variables with laws $\vec{\gamma}$ and $\vec{q}$, respectively. Let $\beta_1 := \mathbb{P}(\Gamma_1 \neq \circ_1)$ and $\tilde{\Gamma}_1$ be a random variable on Ξ_1 with the law of Γ_1 conditioned on $\Gamma_1 \neq \circ_1$. Let $F_{1,2}^{(n)}$ be a simple unmarked graph chosen uniformly at random from the set $\cup_{\vec{d}\in\mathcal{D}^{(n)}} \mathcal{G}_{\vec{d}}^{(n)}$. By definition, $G_{1,2}^{(n)} \sim \mathcal{G}(n; \vec{d}^{(n)}, \vec{\gamma}, \vec{q})$ is obtained from $F_{1,2}^{(n)}$ by adding independent edge and vertex marks according to the laws of $\vec{\gamma}$ and $\vec{q}$ respectively. Let $F_1^{(n)}$ be obtained from $F_{1,2}^{(n)}$ by independently removing each edge with probability $1 - \beta_1$. Then, $G_1^{(n)}$ can be thought of as obtained from $F_1^{(n)}$ by adding independent vertex and edge marks with the laws of Q_1 and $\tilde{\Gamma}_1$, respectively. Hence, we may consider $G_{1,2}^{(n)}$, $F_{1,2}^{(n)}$, $G_1^{(n)}$ and $F_1^{(n)}$ on a joint probability space.

It is straightforward to see the following.

Lemma 2. *Assume X is an integer valued random variable taking value in $\{0,\dots,\Delta\}$ and $0 \le \epsilon \le 1$. Let $\{Y_i\}_{i\ge 1}$ be a sequence of i.i.d. Bernoulli random variables with $\mathbb{P}(Y_i = 1) = \epsilon$. Define the random variable X_1 to be $\sum_{i=1}^{X} Y_i$ and $X_2 := X - X_1$. Then, we have*

$$H(X_1, X_2) = H(X) + \mathbb{E}[X] H(Y_1) - \mathbb{E}\left[\log \binom{X}{X_1}\right].$$

The following lemma which is a direct consequence of Theorem 4.5 and Corollary 4.6 in [BC14], is useful in the asymptotic analysis of the count of the graphs with a given degree sequence.

Lemma 3. *Given an integer Δ, assume that Y is an integer random variable bounded by Δ such that $d := \mathbb{E}[Y] > 0$. Moreover, assume that for each n, $\vec{a}^{(n)} = (a^{(n)}(1),\dots,a^{(n)}(n))$ is a degree sequence of length n with entries bounded by Δ such that $b_n := \sum_{i=1}^n a^{(n)}(i)$ is even and, for $0 \le k \le \Delta$, $c_k(\vec{a}^{(n)})/n \to \mathbb{P}(Y = k)$. Then, we have*

$$\lim_{n\to\infty} \frac{\log |\mathcal{G}^{(n)}_{\vec{a}^{(n)}}| - \frac{b_n}{2}\log n}{n} = -s(d) - \mathbb{E}[\log Y!],$$

where $s(d) := \frac{d}{2} - \frac{d}{2}\log d$.

Lemma 4. *If $G^{(n)}_{1,2} \sim \mathcal{G}(n; \vec{d}^{(n)}, \vec{\gamma}, \vec{q})$, we have $\mathbb{P}(G^{(n)}_{1,2} \notin \mathcal{W}^{(n)}) \le \kappa n^{-1/3}$ for some constant $\kappa > 0$.*

Proof. For this, we note the following. Condition (i) in the definition of $\mathcal{W}^{(n)}$ always holds for a realization $G^{(n)}_{1,2}$. Chebyshev's inequality implies that conditions (ii) and (iii) hold with probability at least $1 - \kappa_1 n^{-1/3}$, for some $\kappa_1 > 0$. To show (iv), fix k and l and for $1 \le i \le n$, let Y_i be the indicator of $\mathrm{dg}_{G^{(n)}_{1,2}}(i) = k$ and $\mathrm{dg}_{G^{(n)}_1}(i) = l$. With $Y := \sum_{i=1}^n Y_i$, we have $c_{k,l}(\overrightarrow{\mathrm{dg}}_{G^{(n)}_{1,2}}, \overrightarrow{\mathrm{dg}}_{G^{(n)}_1}) = Y$. Note that an edge of $G^{(n)}_{1,2}$ exists in $G^{(n)}_1$ if its mark is not of the form $(\circ_1, x_2)$, which happens with probability β_1. Therefore,

$$\mathbb{E}\left[Y_i | F^{(n)}_{1,2}\right] = \mathbb{1}\left[\mathrm{dg}_{F^{(n)}_{1,2}}(i) = k\right] \binom{\mathrm{dg}_{F^{(n)}_{1,2}}(i)}{l} \beta_1^l (1-\beta_1)^{k-l}.$$

Consequently,

$$\mathbb{E}\left[Y | F^{(n)}_{1,2}\right] = c_k(\vec{d}^{(n)}) \binom{k}{l} \beta_1^l (1-\beta_1)^{k-l}.$$

Since this is a constant, it is also equal to $\mathbb{E}[Y]$. Now, if $s_{k,l} := \mathbb{P}(X = k, X_1 = l)$, we have $s_{k,l} = r_k \binom{k}{l} \beta_1^l (1-\beta_1)^{k-l}$. Hence, the assumption (8) implies that

$$|\mathbb{E}[Y] - n s_{k,l}| \le n^{1/2} \binom{k}{l} \beta_1^l (1-\beta_1)^{k-l}. \tag{47}$$

Furthermore, since edge marks are chosen independently, conditioned on $F^{(n)}_{1,2}$, if i and j are nonadjacent vertices in $F^{(n)}_{1,2}$, then Y_i are Y_j are independent, conditioned on $F^{(n)}_{1,2}$. As a result, if $\mathcal{I}$ denotes the set of (i,j) with $1 \leq i \neq j \leq n$ such that i and j are not adjacent in $F_{1,2}^{(n)}$, we have

$$\begin{aligned}
\mathbb{E}\left[Y^2|F_{1,2}^{(n)}\right] &= \sum_{i=1}^{n} \mathbb{E}\left[Y_i^2|F_{1,2}^{(n)}\right] + \sum_{1\leq i\neq j\leq n} \mathbb{E}\left[Y_iY_j|F_{1,2}^{(n)}\right] \\
&\leq n + \sum_{(i,j)\notin\mathcal{I}} \mathbb{E}\left[Y_iY_j|F_{1,2}^{(n)}\right] + \sum_{(i,j)\in\mathcal{I}} \mathbb{E}\left[Y_iY_j|F_{1,2}^{(n)}\right] \\
&\leq n + 2m_n + \sum_{(i,j)\in\mathcal{I}} \mathbb{E}\left[Y_iY_j|F_{1,2}^{(n)}\right] \\
&\overset{(a)}{=} n + 2m_n + \sum_{(i,j)\in\mathcal{I}} \mathbb{E}\left[Y_i|F_{1,2}^{(n)}\right] \mathbb{E}\left[Y_j|F_{1,2}^{(n)}\right] \\
&\leq n + 2m_n + \sum_{1\leq i\neq j\leq n} \mathbb{E}\left[Y_i|F_{1,2}^{(n)}\right] \mathbb{E}\left[Y_j|F_{1,2}^{(n)}\right] \\
&= n + 2m_n + \mathbb{E}\left[Y|F_{1,2}^{(n)}\right]^2,
\end{aligned}$$

where (a) uses the fact that conditioned on $F_{1,2}^{(n)}$, the random variables Y_i and Y_j are independent for $(i,j) \in \mathcal{I}$. From (8), we have $|m_n - nd_{1,2}/2| \leq \kappa_2 n^{1/2}$ and $\kappa_2 := \Delta(\Delta+1)/4$ is a constant. As we saw above, $\mathbb{E}\left[Y|F_{1,2}^{(n)}\right] = \mathbb{E}[Y]$. Therefore, we have $\operatorname{Var} Y \leq \kappa_3 n$ for some $\kappa_3 > 0$. This together with (47) and Chebyshev's inequality implies that the condition (iv) holds with probability at least $1-\kappa_4 n^{-1/3}$, for some $\kappa_4 > 0$. Similarly, the same statement holds for the condition (v). ☐

Now we show (12a)–(12c). In the following, with $X \sim \vec{r}$ and X_1 and X_2 defined as in (11), let $B_{1,2}^{(n)}$ be the set of pairs of degree sequences $\vec{d}$ and $\vec{\delta}$ with n elements bounded by Δ such that for all $0 \leq k,l \leq \Delta$, $|c_{k,l}(\vec{d},\vec{\delta}) - n\mathbb{P}(X_1 = k, X_1' = l)| \leq n^{2/3}$, where $X_1' := X - X_1$. Moreover, let $B_1^{(n)}$ be the set of $\vec{d}$ such that for some $\vec{\delta}$, we have $(\vec{d},\vec{\delta}) \in B_{1,2}^{(n)}$. For $\vec{d} \in B_1^{(n)}$, let $B_{2|1}^{(n)}(\vec{d})$ be the set of degree sequences $\vec{\delta}$ such that $(\vec{d},\vec{\delta}) \in B_{1,2}^{(n)}$.

Now, we show (12a). Since $G_{1,2}^{(n)}$ is formed by adding independent vertex and edge marks to $F_{1,2}^{(n)}$, we have

$$H(G_{1,2}^{(n)}) = \log|\mathcal{D}^{(n)}| + \log|\mathcal{G}_{\vec{d}^{(n)}}^{(n)}| + m_n H(\Gamma) + nH(Q).$$

From (8), we have $|m_n - nd_{1,2}^{\mathrm{CM}}/2| \leq \frac{\Delta K}{2} n^{1/2}$. Moreover, we have $\mathbb{E}[X] > 0$. Consequently, using Lemma 3 and the fact that $\frac{1}{n}\log|\mathcal{D}^{(n)}| \to H(X)$, we get (12a).

We now turn to showing (12b). Since the expected number of the edges in $F_1^{(n)}$ is $nd_1^{\mathrm{CM}}/2$, we have

$$H(G_1^{(n)}) = H(F_1^{(n)}) + n\frac{d_1^{\mathrm{CM}}}{2} H(\Gamma_1|\Gamma_1 \neq \circ_1) + nH(Q_1). \tag{48}$$

With this, we focus on $H(F_1^{(n)})$. With E_n being the indicator of $G_{1,2}^{(n)} \notin \mathcal{W}^{(n)}$, we have

$$\begin{aligned}
H(F_1^{(n)}) = H(F_1^{(n)}|E_n = 0)\mathbb{P}(E_n = 0) \\
+H(F_1^{(n)}|E_n = 1)\mathbb{P}(E_n = 1).
\end{aligned}$$

From Lemma 4, we have

$$H(F_1^{(n)}|E_n = 1)\mathbb{P}(E_n = 1) \leq (H(F_{1,2}^{(n)}) + m_n H(\beta_1))\kappa n^{-1/3}, \tag{49}$$

where $H(\beta_1)$ denotes the binary entropy of β_1. Note that as we discussed above, $H(F_1^{(n)}) = O(n \log n)$. Thereby, the RHS of the above is $o(n)$. On the other hand, by the definition of $\mathcal{W}^{(n)}$, if $E = 0$, $\overrightarrow{\mathrm{dg}}_{F_1^{(n)}} \in B_1^{(n)}$. Therefore, $H(F_1^{(n)}|E=0) \le \log B_1^{(n)} + \max_{\vec{d} \in B_1^{(n)}} \log |\mathcal{G}_{\vec{d}}^{(n)}|$. The assumption $r_0 < 0$ together with (7) imply that $d_1^{\mathrm{CM}} > 0$. Hence, using Lemma 3 together with (49),

$$\limsup \frac{H(F_1^{(n)}) - n \frac{d_1^{\mathrm{CM}}}{2} \log n}{n} \le -s(d_1^{\mathrm{CM}}) + H(X_1) - \mathbb{E}\left[\log X_1!\right]. \tag{50}$$

Now, let $\tilde{F}_1^{(n)}$ be the unmarked graph consisting of the edges removed from $F_{1,2}^{(n)}$ to obtain $F_1^{(n)}$ and note that

$$\begin{aligned} H(F_1^{(n)}) &= H(F_1^{(n)}, \tilde{F}_1^{(n)}) - H(\tilde{F}_1^{(n)}|F_2^{(n)}) \\ &= H(F_{1,2}^{(n)}) + m_n H(\beta_1) - H(\tilde{F}_1^{(n)}|F_1^{(n)}) \end{aligned} \tag{51}$$

Note that, conditioned on $E = 0$, we have $\overrightarrow{\mathrm{dg}}_{\tilde{F}_1^{(n)}} \in B_{2|1}^{(n)}(\overrightarrow{\mathrm{dg}}_{F_1^{(n)}})$. Moreover, the assumption (7) together with $r_0 < 0$ imply that $d_{1,2}^{\mathrm{CM}} - d_1^{\mathrm{CM}} > 0$. Hence, using a similar method in proving (50), we have

$$\begin{aligned} \limsup \frac{H(\tilde{F}_1^{(n)}|F_1^{(n)}) - n \frac{d_{1,2}^{\mathrm{CM}} - d_1^{\mathrm{CM}}}{2} \log n}{n} &\le -s(d_{1,2}^{\mathrm{CM}} - d_1^{\mathrm{CM}}) \\ &\quad + H(X_1'|X_1) - \mathbb{E}\left[\log X_1'!\right]. \end{aligned}$$

Using this together with the asymptotic of $H(F_{1,2}^{(n)})$ which was derived above in showing (12a) and substituting into (51), followed by a simplification using Lemma 2, we get

$$\liminf \frac{H(F_1^{(n)}) - n \frac{d_1^{\mathrm{CM}}}{2} \log n}{n} \ge -s(d_1^{\mathrm{CM}}) + H(X_1) - \mathbb{E}\left[\log X_1!\right]. \tag{52}$$

This together with (50) and (48) completes the proof of (12b). The proof of (12c) is similar.

B Bounding $|S_2^{(n)}(H_1^{(n)})|$ for the Erdős–Rényi case

Note that for $G_{1,2}^{(n)} \in \mathcal{G}_{1,2}^{(n)}$ and $H_2^{(n)} \in \mathcal{G}_2^{(n)}$, if $G_1^{(n)} \oplus H_2^{(n)} \in \mathcal{G}_{\vec{p},\vec{q}}^{(n)}$, we have $\vec{m}_{G_1^{(n)} \oplus H_2^{(n)}} \in \mathcal{M}^{(n)}$ and $\vec{u}_{G_1^{(n)} \oplus H_2^{(n)}} \in \mathcal{U}^{(n)}$. On the other hand, for fixed $\vec{m} \in \mathcal{M}^{(n)}$ and $\vec{u} \in \mathcal{U}^{(n)}$, if for all $x_1 \in \Xi_1$ we have $m(x_1) = m_{G_1^{(n)}}(x_1)$ and for all $\theta_1 \in \Theta_1$ we have $u(\theta_1) = u_{G_1^{(n)}}(\theta_1)$, then the number of $H_2^{(n)}$ such that $\vec{m}_{G_1^{(n)} \oplus H_2^{(n)}} = \vec{m}$ and $\vec{u}_{G_1^{(n)} \oplus H_2^{(n)}} = \vec{u}$ is at most

$$A_2(\vec{m}, \vec{u}) := \left(\prod_{x_1 \in \Xi_1} \binom{m(x_1)}{\{m(x_1, x_2)\}_{x_2 \in \Xi_2 \cup \{\circ_2\}}} \right) \times \binom{\binom{n}{2} - \sum_{x_1 \in \Xi_1} m(x_1)}{\{m(\circ_1, x_2)\}_{x_2 \in \Xi_2}} \times \left(\prod_{\theta_1 \in \Theta_1} \binom{u(\theta_1)}{\{u(\theta_1, \theta_2)\}_{\theta_2 \in \Theta_2}} \right),$$

where we have used the notational conventions in (3) and (4). Consequently, we have

$$\begin{aligned} \max_{G_{1,2}^{(n)} \in \mathcal{G}_{\vec{p},\vec{q}}^{(n)}} |S_2^{(n)}(G_1^{(n)})| &\le |\mathcal{M}^{(n)}||\mathcal{U}^{(n)}| \max_{\substack{\vec{m} \in \mathcal{M}^{(n)} \\ \vec{u} \in \mathcal{U}^{(n)}}} A_2(\vec{m}, \vec{u}) \\ &\le (2n^{2/3})^{(|\Xi_{1,2}| + |\Theta_{1,2}|)} \max_{\substack{\vec{m} \in \mathcal{M}^{(n)} \\ \vec{u} \in \mathcal{U}^{(n)}}} A_2(\vec{m}, \vec{u}). \end{aligned} \tag{53}$$

Now, if $\vec{m}^{(n)}$ and $\vec{u}^{(n)}$ are sequences in $\mathcal{M}^{(n)}$ and $\mathcal{U}^{(n)}$, respectively, for all $x \in \Xi_{1,2}$ we have $m^{(n)}(x)/n \to p_x/2$. Furthermore, for all $x_1 \in \Xi_1$ and $\theta_1 \in \Theta_1$, we have $m^{(n)}(x_1)/n \to p_{x_1}/2$ and $u^{(n)}(\theta_1)/n \to q_{\theta_1}$. As a result, using Lemma 1, for any such sequences $\vec{m}^{(n)}$ and $\vec{u}^{(n)}$, with $Q = (Q_1, Q_2)$ having law $\vec{q}$ we have

$$\begin{aligned}
\lim_{n\to\infty} & \frac{\log A_2(\vec{m}^{(n)}, \vec{u}^{(n)}) - (\sum_{x_2 \in \Xi_2} m^{(n)}(\circ_1, x_2)) \log n}{n} \\
&= \sum_{x_2\in\Xi_2} s(p_{\circ_1,x_2}) + \sum_{x_1 \in \Xi_1} \frac{p_{x_1}}{2} H\left(\left\{ \frac{p_{(x_1,x_2)}}{p_{x_1}} \right\}_{x_2 \in \Xi_2 \cup \{\circ_2\}} \right) \\
&\quad + \sum_{\theta_1\in\Theta_1} q_{\theta_1} H\left(\left\{ \frac{q_{\theta_1,\theta_2}}{q_{\theta_1}} \right\}_{\theta_2\in\Theta_2} \right) \\
&= H(Q_2|Q_1) + \sum_{x\in\Xi_{1,2}} s(p_x) - \sum_{x_1\in\Xi_1} s(p_{x_1}) \\
&= \Sigma(\mu_2^{\text{ER}} | \mu_1^{\text{ER}}),
\end{aligned}$$

where the second inequality follows by rearranging the terms and using the definition of $s(.)$. Using the fact that $|m^{(n)}(\circ_1, x_2) - np_{\circ_1,x_2}/2| \leq n^{2/3}$,

$$\lim_{n\to\infty} \frac{\log A_n(\vec{m}^{(n)}, \vec{u}^{(n)}) - n\frac{d_{1,2}^{\text{ER}} - d_1^{\text{ER}}}{2} \log n}{n} = \Sigma(\mu_2^{\text{ER}} | \mu_1^{\text{ER}}).$$

This together with (53) implies (19).

C Bounding $|S_2^{(n)}(H_1^{(n)})|$ for the configuration model

Here, we find an upper bound for $\max_{H_{1,2}^{(n)}} |S_2^{(n)}(H_1^{(n)})|$ and use it to show (25). Fix $G_{1,2}^{(n)} \in \mathcal{W}^{(n)}$ and assume $H_2^{(n)} \in S_2^{(n)}(G_1^{(n)})$. With $H_{1,2}^{(n)} := G_1^{(n)} \oplus H_2^{(n)}$, let $\tilde{H}_2^{(n)}$ be the subgraph of $H_{1,2}^{(n)}$ consisting of edges not present in $G_1^{(n)}$. Employing the notation of Appendix A, by the definition of the set $\mathcal{W}^{(n)}$, we have $\vec{\text{dg}}_{\tilde{H}_2^{(n)}} \in B_{2|1}^{(n)}(\vec{\text{dg}}_{G_1^{(n)}})$. Therefore, we can think of $H_{1,2}^{(n)}$ as constructed from $G_1^{(n)}$ by adding a graph to $G_1^{(n)}$ with degree sequence $\vec{\text{dg}}_{\tilde{H}_2^{(n)}}$, markings its edges, adding second domain marks to edges in $G_1^{(n)}$, and also adding second domain marks to vertices. Consequently,

$$\begin{aligned}
\max_{H_{1,2}^{(n)}\in\mathcal{W}^{(n)}} \log |S_2^{(n)}(H_1^{(n)})| &\leq \max_{G_{1,2}^{(n)}\in\mathcal{W}^{(n)}} \log |B_{2|1}^{(n)}(\vec{\text{dg}}_{G_1^{(n)}})| + \max_{G_{1,2}^{(n)}\in\mathcal{W}^{(n)}, \vec{\delta}\in B_{2|1}^{(n)}(\vec{\text{dg}}_{G_1^{(n)}})} \log |\mathcal{G}_{\vec{\delta}}^{(n)}| \\
&\quad + \max_{\vec{m}\in\mathcal{M}^{(n)}} \log \binom{m_n - \sum_{x_1\in\Xi_1} m(x_1)}{\{m(\circ_1, x_2)\}_{x_2\in\Xi_2}} \prod_{x_1\in\Xi_1} \binom{m(x_1)}{\{m(x_1,x_2)\}_{x_2\in\Xi_2}} \\
&\quad + \max_{\vec{u}\in\mathcal{U}^{(n)}} \log \prod_{\theta_1\in\Theta_1} \binom{u(\theta_1)}{\{u(\theta_1,\theta_2)\}_{\theta_2\in\Theta_2}}
\end{aligned} \tag{54}$$

The definition of $B_{2|1}^{(n)}$ implies that

$$\lim_{n\to\infty} \frac{1}{n} \max_{H_{1,2}^{(n)}\in\mathcal{W}^{(n)}} \log |B_{2|1}^{(n)}(\vec{\text{dg}}_{H_1^{(n)}})| = H(X - X_1|X_1).$$

Note that the assumption (7) together with $r_0 < 1$ implies that $d_{1,2}^{\mathrm{CM}} - d_1^{\mathrm{CM}} > 0$. Therefore, Lemma 3 in Appendix A implies that

$$\limsup_{n\to\infty} \max_{\substack{G_{1,2}^{(n)} \in \mathcal{W}^{(n)} \\ \vec{\delta} \in B(\overrightarrow{\mathrm{dg}}_{G_1^{(n)}})}} \frac{\log |\mathcal{G}_{\vec{\delta}}^{(n)}| - n \frac{d_{1,2}^{\mathrm{CM}} - d_1^{\mathrm{CM}}}{2} \log n}{n} \leq -s(d_{1,2}^{\mathrm{CM}} - d_1^{\mathrm{CM}}) - \mathbb{E}\left[\log(X - X_1)!\right].$$

Furthermore, a usage of Lemma 1 implies that the third and the fourth terms in (54) divided by n converge to $\frac{d_{1,2}^{\mathrm{CM}}}{2} H(\Gamma_2|\Gamma_1)$ and $H(Q)$, respectively, where $\Gamma = (\Gamma_1, \Gamma_2)$ has law $\vec{\gamma}$ and Q has law $\vec{q}$. Putting these together, we have

$$\begin{aligned}\lim_{n\to\infty} \frac{\max_{G_{1,2}^{(n)} \in \mathcal{W}^{(n)}} \log |S_2^{(n)}(G_1^{(n)})| - n \frac{d_{1,2}^{\mathrm{CM}} - d_1^{\mathrm{CM}}}{2} \log n}{n} &= -s(d_{1,2}^{\mathrm{CM}} - d_1^{\mathrm{CM}}) + H(X - X_1|X_1) \\ &\quad - \mathbb{E}\left[\log(X - X_1)!\right] + \frac{d_{1,2}^{\mathrm{CM}}}{2} H(\Gamma_2|\Gamma_1) + H(Q_2|Q_1).\end{aligned}$$

Using Lemma 2 and rearranging, this is precisely equal to $\Sigma(\mu_2^{\mathrm{CM}}|\mu_1^{\mathrm{CM}})$, which completes the proof of (25).

Acknowledgments

The authors acknowledge support from the NSF grants ECCS-1343398, CNS- 1527846, CCF-1618145, the NSF Science & Technology Center grant CCF- 0939370 (Science of Information), and the William and Flora Hewlett Foundation supported Center for Long Term Cybersecurity at Berkeley.

References

[Abb16] Emmanuel Abbe. Graph compression: The effect of clusters. In *Communication, Control, and Computing (Allerton), 2016 54th Annual Allerton Conference on*, pages 1–8. IEEE, 2016.

[AL07] David Aldous and Russell Lyons. Processes on unimodular random networks. *Electron. J. Probab*, 12(54):1454–1508, 2007.

[AS04] David Aldous and J Michael Steele. The objective method: probabilistic combinatorial optimization and local weak convergence. In *Probability on discrete structures*, pages 1–72. Springer, 2004.

[BC14] Charles Bordenave and Pietro Caputo. Large deviations of empirical neighborhood distribution in sparse random graphs. *Probability Theory and Related Fields*, pages 1–73, 2014.

[Bil13] Patrick Billingsley. *Convergence of probability measures*. John Wiley & Sons, 2013.

[BS01] Itai Benjamini and Oded Schramm. Recurrence of distributional limits of finite planar graphs. *Electron. J. Probab.*, 6:no. 23, 13 pp. (electronic), 2001.

[CS12] Yongwook Choi and Wojciech Szpankowski. Compression of graphical structures: Fundamental limits, algorithms, and experiments. *IEEE Transactions on Information Theory*, 58(2):620–638, 2012.

[CT12] Thomas M Cover and Joy A Thomas. *Elements of information theory.* John Wiley & Sons, 2012.

[DA17] Payam Delgosha and Venkat Anantharam. Universal lossless compression of graphical data. In *Information Theory (ISIT), 2017 IEEE International Symposium on*, pages 1578–1582. IEEE, 2017.