

On the Lossy Compression of Spatial Networks

Praneeth Kumar Vippathalla, Martin Wachiye Wafula, Mihai-Alin Badiu and Justin P. Coon

Department of Engineering Science

University of Oxford

OX1 3PJ Oxford, U.K.

Email: {praneeth.vippathalla, martin.wafula, mihai.badiu, justin.coon}@eng.ox.ac.uk

Abstract—In this paper, we address the lossy compression of spatial networks, namely random geometric graphs, where two nodes are connected by an edge with a probability that depends on the distance between the nodes. We carry out this study by considering the n th order information-distortion function, which quantifies the complexity of a random graph under a distortion criterion. Our main result is a partial characterization of the information-distortion function for a random geometric graph with the Hamming distortion measure.

I. INTRODUCTION

As graph-structured data of enormous sizes are ubiquitous nowadays, there is a need for studying the fundamental limits of graph compression. The main objectives of graph compression are to store a graph efficiently by utilizing less memory or to speed up an algorithm by running it on the compressed version. Recently, there have been some works [1]–[12] addressing various aspects of graph compression. In the case of lossless compression of a graph, the entropy of a graph governs the asymptotically optimal behavior. For instance, the works [1], [6]–[9] studied the entropy of a stochastic block graph, and the structural entropy of an Erdős-Rényi (ER) random graph and spatial networks such as random geometric graphs.

Because of storage limitations, it is also of interest to compress a large graph in a lossy way. We can allow a graph feature to be distorted within a certain level. A natural framework to study this problem is the classical rate-distortion theory. As graph sources are arbitrary and the distortion functions could be complicated, the study of lossy compression is a challenging task. For instance, [5] considered the lossy compression problem of an ER random graph with a distortion measure on the degree of the nodes. They showed lower and upper bounds on the asymptotic rate of compression. As it is difficult to extend their techniques to the other graph models such as spatial networks, which is the focus of this paper, we study the n th order information-distortion function. This is a natural analogue of entropy in the context of lossy compression.

This research was funded in whole or in part by the U. S. Army Research Laboratory and the U. S. Army Research Office (W911NF-22-1-0070), and EPSRC (EP/T02612X/1). For the purpose of Open Access, the authors have applied a CC BY public copyright licence to any Author Accepted Manuscript (AAM) version arising from this submission.

A. Information-Distortion Function of a Random Graph

Let $G_n = (V, E)$ be a graph with the vertex set $V = [n] \triangleq \{1, 2, \dots, n\}$ and the edge set E . We denote by \mathcal{G}_n the set of all graphs on n vertices. For simplicity, we restrict ourselves to undirected and simple graphs, where simple means that there are no self loops or multiple edges. For a random graph $G_n \in \mathcal{G}_n$, P_{G_n} denotes its probability distribution.

Given a distortion function $d_n : \mathcal{G}_n \times \mathcal{G}_n \rightarrow [0, \infty)$ and a distortion level D_n , which specifies the cost of representing a graph with another graph, the n th order *information-distortion function* of a random graph G_n is defined as

$$I_{G_n}(D_n) \triangleq \min_{P_{\hat{G}_n|G_n} : \mathbb{E}[d_n(G_n, \hat{G}_n)] \leq D_n} I(G_n; \hat{G}_n). \quad (1)$$

The information-distortion function is useful for identifying the fundamental limits of lossy compression. It is well-known from the classical rate-distortion theory [13, Ch. 13] that for an i.i.d. source $\{X^n\}_{n \geq 1}$ and a per-letter distortion function d , the minimum rate of compression $R(D)$ at an (average) distortion level D is given by

$$R(D) = \min_{P_{\hat{X}|X} : \mathbb{E}[d(X, \hat{X})] \leq D} I(X; \hat{X}). \quad (2)$$

The right-hand side of (2) is the first order information-distortion function of the i.i.d. source $\{X^n\}_{n \geq 1}$. On the other hand, the distribution P_{G_n} of a random graph need not have the desired independence structure. For such arbitrary sources [14, Ch. 5], the n th order information-distortion function can be used to study the lossy compression.

Consider the following variable-length lossy compression of random graphs. An encoder $\phi_n : \mathcal{G}_n \rightarrow \{0, 1\}^*$ maps a graph to a codeword, which is a binary string of finite (but arbitrary) length. Here $\{0, 1\}^*$ denotes the set of binary strings of finite length. We also assume that the encoder satisfies the prefix condition. A decoder $\psi : \{0, 1\}^* \rightarrow \mathcal{G}_n$ maps a codeword to a graph. The encoder and decoder pair is designed such that they satisfy the distortion constraint $\mathbb{E}[d_n(G_n, \psi(\phi(G_n)))] \leq D_n$ for a distortion level D_n . Let $l_n^*(D_n)$ denote the least average length of a codeword $\mathbb{E}[l(\phi(G_n))]$. We can easily show that

$$I_{G_n}(D_n) \leq l_n^*(D_n). \quad (3)$$

This follows from the observation that for a prefix code [13, Thm. 5.4.1]

$$l_n^*(D_n) \geq H(\phi(G_n)), \quad (4)$$

which is further lower bounded by $H(\psi(\phi(G_n))) \geq I(\psi(\phi(G_n)); G_n)$, where the reconstructed graph satisfies the average distortion condition, yielding (3). Though an upper bound on the minimum optimal length $l_n^*(D_n)$ in terms of $I_{G_n}(D_n)$ and D_n is not known for a fixed n , asymptotically optimal results for the rate are known.

Define the asymptotic rate of the compression scheme for the distortion level $D_n = \binom{n}{2}D$ to be

$$R(D) = \limsup_{n \rightarrow \infty} \frac{1}{\binom{n}{2}} l_n^* \left(\binom{n}{2} D \right). \quad (5)$$

It was shown in [14, Thm. 5.7.1 and Remark 5.7.3] that

$$R(D) = I(D) \triangleq \limsup_{n \rightarrow \infty} \frac{1}{\binom{n}{2}} I_{G_n} \left(\binom{n}{2} D \right), \quad (6)$$

under some reasonable assumptions on the distortion function and by allowing a stochastic encoder instead of deterministic encoder. It was also noted in [14, Cor. 5.7.1] that the stochastic encoder can be implemented without loss of optimality by randomizing between two deterministic encoders.

The n th order information-distortion function is also useful in the regimes where the order of the distortion levels $\{D_n\}_{n \geq 1}$ is not the same as that of the source entropy. For example, in the result (6), a notion of rate is defined with respect to the scaling $\binom{n}{2}$ with $D_n = \binom{n}{2}D$. But we can consider other regimes that are of interest, say $H(G_n) = \Theta(n^2)$ and $D_n = n \log n D$. For such cases, where we may not be able to define a useful notion of rate, we can still work with $I_{G_n}(D_n)$ owing to the inequality (3). The effect of distortion on the information-distortion function can be made explicit through the Lagrangian dual formulation

$$I_{G_n}(D_n) = \sup_{\lambda \geq 0} \inf_{P_{G_n|G_n}} I(G_n; \hat{G}_n) + \lambda \left[\mathbb{E}[d_n(G_n, \hat{G}_n)] - D_n \right] \quad (7)$$

For instance, in the case of $H(G_n) = \Theta(n^2)$ and $D_n = n \log n D$, the distortion constraint contributes only to the $n \log n$ term of the expression $H(G_n) - H(G_n|\hat{G}_n) + \lambda \left[\mathbb{E}[d_n(G_n, \hat{G}_n)] - D_n \right]$ with the corresponding coefficient quantifying the effect of distortion on compression. With the above motivation, we study the information-distortion function of spatial networks in this paper.

II. INFORMATION-DISTORTION FUNCTION OF SPATIAL NETWORKS

Random spatial networks are those random graphs with an underlying hidden spatial embedding of the nodes. One important example of a random spatial network is a random geometric graph. In a random geometric graph, vertices are in a d -dimensional Euclidean space and the edge connectivity between a pair of nodes is governed by the distance between them.

Formally, the vertices are spatially embedded in a bounded set $\mathcal{K} \subset \mathbb{R}^d$ with strictly positive volume with respect to the Lebesgue measure. Let $K \triangleq \sup_{x, y \in \mathcal{K}} \|x - y\|_2$ be the diameter of the domain \mathcal{K} . For a given connectivity function

$p : [0, K] \rightarrow [0, 1]$ and a fixed positive integer n , a *random geometric graph* (RGG) is obtained as follows. We first draw n points $\mathbf{X} = (X_i : 1 \leq i \leq n)$ uniformly at random from \mathcal{K} and independently of each other. The distance between two nodes i and j is denoted by $R_{i,j} \triangleq \|X_i - X_j\|_2$. We then form a graph G_n on the vertex set $V = [n]$ by drawing an edge between a pair of vertices i and j with probability $p(R_{i,j})$ for $1 \leq i < j \leq n$. Let P_{G_n} denote the induced probability distribution on the space of all graphs \mathcal{G}_n with n vertices.

The presence or absence of an edge between vertices i and j in a graph is denoted by a binary variable $E_{i,j}$. This allows us to view a graph G_n in terms of its adjacency matrix representation $(E_{i,j})_{i,j \in [n]}$, where $E_{i,j} = 1$ if i and j are adjacent in G_n , and $E_{i,j} = 0$ otherwise. As the graphs are simple and undirected, this matrix is symmetric with zeros on the diagonal. It means that G_n can equivalently be represented by $\{E_{i,j} : 1 \leq i < j \leq n\}$. Hence, the probability of a random geometric graph G_n is completely specified by the joint distribution of the collection $\{E_{i,j} : 1 \leq i < j \leq n\}$, and vice versa.

A. Hamming Distortion Measure

While there are various distortion measures that one can work with depending on the context, in this preliminary work, we consider the Hamming distortion measure d_n between two graphs G_n and \hat{G}_n that counts the number of pairs of vertices for which the values of edge random variables differ in G_n and \hat{G}_n , i.e.,

$$d_n(G_n, \hat{G}_n) = \sum_{i < j} \mathbb{1}(E_{i,j} \neq \hat{E}_{i,j}). \quad (8)$$

The Hamming distortion penalizes a reconstructed graph \hat{G}_n that has many pairs of vertices with mismatched edge random variables from the original graph G_n .

B. Main Results

In this section, we will present our result on the characterization of the n th order information-distortion function (1) of a random geometric graph with the Hamming distortion measure. For establishing the result, we first study the information-distortion function with conditioning on the node locations \mathbf{X} , which is defined as

$$I_{G_n|\mathbf{X}}(D_n) \triangleq \inf_{P_{\hat{G}_n|G_n, \mathbf{X}} : \mathbb{E}[d_n(G_n, \hat{G}_n)] \leq D_n} I(G_n; \hat{G}_n | \mathbf{X}). \quad (9)$$

This quantity arises when \mathbf{X} is available as side information to both the encoder and decoder. It is worth noting that the information-distortion functions with and without conditioning satisfy the following inequality.

Lemma 1 ([15]). *For a pair (\mathbf{X}, G_n) and $D_n \geq 0$,*

$$I_{G_n|\mathbf{X}}(D_n) \leq I_{G_n}(D_n) \leq I_{G_n|\mathbf{X}}(D_n) + I(G_n; \mathbf{X}). \quad (10)$$

Before we present our result, let us define a few quantities, which will simplify the notation later on. We define $R \triangleq \|X - Y\|_2$, where X and Y are independent and uniformly distributed random variables over \mathcal{K} . For $D_n \geq 0$ and a

connectivity function $p : [0, K] \rightarrow [0, 1]$, define a function $q^* : [0, K] \rightarrow [0, 1]$ as follows:

$$q^*(r) = \min\{p(r), 1 - p(r), \mu\} \quad (11)$$

with μ chosen to satisfy the condition

$$\binom{n}{2} \mathbb{E}[q^*(R)] = D_n \text{ if } D_n \leq \binom{n}{2} \mathbb{E}[\min\{p(R), 1 - p(R)\}],$$

and set $\mu = \frac{1}{2}$ otherwise.

Theorem 1 (Characterization of $I_{G_n|\mathbf{X}}(D_n)$). *Let G_n be a random geometric graph with \mathbf{X} being the node locations and $p : [0, K] \rightarrow [0, 1]$ being the connectivity function. The n th order information-distortion function of G_n conditioned on node locations at the Hamming distortion D_n is given by*

$$I_{G_n|\mathbf{X}}(D_n) = \binom{n}{2} [\mathbb{E}[h_2(p(R))] - \mathbb{E}[h_2(q^*(R))]], \quad (12)$$

where $q^* : [0, K] \rightarrow [0, 1]$ is defined as in (11).

Proof. See Section IV. \square

The above theorem together with Lemma 1 can be used to obtain a characterization of $I_{G_n}(D_n)$. To state this result, we need the following concept of the largest convex lower bound. The *convex envelope*, $\text{Conv}[f]$, of a function $f : [0, \infty) \rightarrow \mathbb{R}$ is defined as

$$\text{Conv}[f] \triangleq \sup\{g : g \text{ is convex, } g(x) \leq f(x) \forall x \in [0, \infty)\},$$

where the supremum is pointwise.

Theorem 2 (Bounds on $I_{G_n}(D_n)$). *Let G_n be a random geometric graph with \mathbf{X} being the node locations and $p : [0, K] \rightarrow [0, 1]$ being the connectivity function. The n th order information-distortion function of G_n at the Hamming distortion D_n is bounded as below*

$$\binom{n}{2} [\mathbb{E}[h_2(p(R))] - \mathbb{E}[h_2(q^*(R))]] \leq I_{G_n}(D_n), \quad (13)$$

and

$$I_{G_n}(D_n) \leq \text{Conv}[J_{G_n}](D_n), \quad (14)$$

where $q^* : [0, K] \rightarrow [0, 1]$ is defined as in (11), and the function J_{G_n} is defined as

$$J_{G_n}(D_n) = \binom{n}{2} [h_2(\mathbb{E}[p(R)]) - \mathbb{E}[h_2(q^*(R))]]$$

if $D_n \leq \binom{n}{2} \min\{\mathbb{E}[p(R)], 1 - \mathbb{E}[p(R)]\}$, and $J_{G_n}(D_n) = 0$, otherwise.

Proof. The lower bound (13) follows readily by combining the result of Theorem 1 with the lower bound in Lemma 1.

For the upper bound (14), first note that

$$I_{G_n}(D_n) = 0 \text{ if } \binom{n}{2} \min\{\mathbb{E}[p(R)], 1 - \mathbb{E}[p(R)]\} \leq D_n.$$

To see this, consider, in (1), the conditional distribution $P_{\hat{G}_n|G_n} = P_{\hat{G}_n}$ such that $\mathbb{P}(\hat{E}_{ij} = 1) = \mathbb{1}\{\mathbb{P}(E_{ij} =$

$0) < \mathbb{P}(E_{ij} = 1)\}$ for $i < j$. The average distortion for this distribution is $\mathbb{E}[d_n(G_n, \hat{G}_n)] = \sum_{i < j} \mathbb{P}(E_{i,j} \oplus \hat{E}_{i,j} = 1) = \sum_{i < j} \min\{\mathbb{P}(E_{ij} = 0), \mathbb{P}(E_{ij} = 1)\} = \binom{n}{2} \min\{\mathbb{E}[p(R)], 1 - \mathbb{E}[p(R)]\}$. Under the restriction on D_n , we have $\mathbb{E}[d_n(G_n, \hat{G}_n)] \leq D_n$, which implies that $0 \leq I_{G_n}(D_n) \leq I(G_n; \hat{G}_n) = 0$.

For $D_n \leq \binom{n}{2} \min\{\mathbb{E}[p(R)], 1 - \mathbb{E}[p(R)]\}$, we will combine the result of Theorem 1 with the upper bound in Lemma 1. Consider the mutual information $I(G_n; \mathbf{X})$ term, which can be bounded using the same approach as in [8].

$$\begin{aligned} I(G_n; \mathbf{X}) &= H(G_n) - H(G_n|\mathbf{X}) \\ &\leq \sum_{i < j} H(E_{i,j}) - H(G_n|\mathbf{X}) \\ &\stackrel{(a)}{=} \sum_{i < j} H(E_{i,j}) - \sum_{i < j} H(E_{i,j}|X_i, X_j) \\ &\stackrel{(b)}{=} \binom{n}{2} h_2(\mathbb{E}[p(R)]) - \binom{n}{2} \mathbb{E}[h_2(p(R))], \end{aligned} \quad (15)$$

where (a) follows by applying the chain rule of entropy and noting that $E_{i,j}$ is independent of the rest of the random variables given X_i and X_j for all $i < j$, and in (b), we use the fact that for $i < j$, $\mathbb{P}(E_{i,j} = 1) = \mathbb{E}_R[\mathbb{P}(E_{i,j} = 1|R_{i,j} = R)] = \mathbb{E}[p(R)]$ and $H(E_{i,j}|X_i, X_j) = \mathbb{E}_R[H(E_{i,j}|R_{i,j} = R)] = \mathbb{E}[h_2(p(R))]$.

By setting

$$\begin{aligned} J_{G_n}(D_n) &\triangleq I_{G_n|\mathbf{X}}(D_n) + I(G_n; \mathbf{X}) \\ &= \binom{n}{2} [h_2(\mathbb{E}[p(R)]) - \mathbb{E}[h_2(q^*(R))]] \end{aligned}$$

for $D_n \leq \binom{n}{2} \min\{\mathbb{E}[p(R)], 1 - \mathbb{E}[p(R)]\}$, and $J_{G_n}(D_n) \triangleq 0$, otherwise, we see that $I_{G_n}(D_n) \leq J_{G_n}(D_n)$. As $I_{G_n}(D_n)$ is a convex function in D_n , (14) follows from the definition of the convex envelope. \square

III. NUMERICAL EVALUATION

In this section, we will explore an example and numerically evaluate the bounds on the information-distortion function. A connectivity function that arises in spatial networks, namely wireless networks [8], is

$$p(r) = e^{-(r/r_0)^\eta},$$

where the parameter r_0 denotes the typical connectivity range and η is called the path loss exponent. We consider two spatial domains: a unit square and a circle of unit area. In Fig. 1 and 2, we plot the upper and lower bounds (13) and (14) on the normalized information distortion function

$$\bar{I}(\bar{D}) \triangleq \frac{1}{\binom{n}{2}} I_{G_n} \left(\binom{n}{2} \bar{D} \right).$$

The quantities are normalized so that the effect of the connectivity function on the information-distortion function can be seen more clearly. For instance, it can be seen in Fig. 1 that if the typical connectivity range r_0 increases, the information distortion function increases up to a certain r_0 and it decreases

later on. This behavior is closely related to the complexity of the random graph. On the other hand, Fig. 2 plots the bounds as functions of D_n .

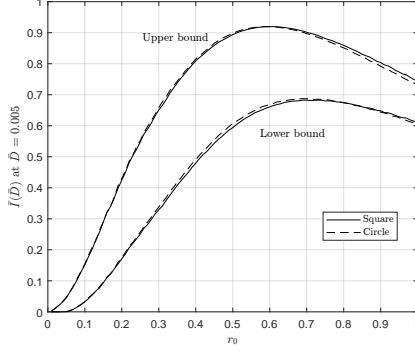


Fig. 1. Upper and lower bounds on the (normalized) information-distortion function as functions of the connectivity parameter r_0 on two spatial domains: a) a unit square b) a circle of unit area. For this plot, the (normalized) distortion $\bar{D} = 0.005$ and $\eta = 2$.

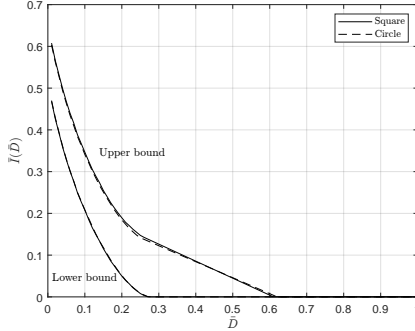


Fig. 2. Upper and lower bounds on the (normalized) information-distortion function as functions of (normalized) distortion level on two spatial domains: a) a unit square b) a circle of unit area. For this plot, $r_0 = 0.75$ and $\eta = 2$.

IV. PROOF OF THEOREM 1

By conditioning on the node locations, the random geometric graph G_n becomes an inhomogeneous Erdős-Rényi graph, meaning that the conditional edge random variables are independent with different edge probabilities. With this observation, we can follow the approach used in [13, Thm. 10.3.1] to characterize $I_{G_n|\mathbf{X}}(D_n)$.

A. A Lower Bound on $I_{G_n|\mathbf{X}}(D_n)$

Consider a conditional joint distribution $P_{\hat{G}_n|G_n,\mathbf{X}}$ satisfying the constraint $\mathbb{E}[d_n(G_n, \hat{G}_n)] \leq D_n$. For this distribution, we have

$$\begin{aligned} I(G_n; \hat{G}_n | \mathbf{X}) &= H(G_n | \mathbf{X}) - H(G_n | \hat{G}_n, \mathbf{X}) \\ &\stackrel{(a)}{=} \sum_{i < j} H(E_{i,j} | X_i, X_j) - H(G_n | \hat{G}_n, \mathbf{X}) \\ &\geq \sum_{i < j} H(E_{i,j} | X_i, X_j) - \sum_{i < j} H(E_{i,j} | \hat{E}_{i,j}, X_i, X_j) \\ &= \sum_{i < j} \mathbb{E}_R [H(E_{i,j} | R_{ij} = R) - H(E_{i,j} \oplus \hat{E}_{i,j} | \hat{E}_{i,j}, R_{ij} = R)] \end{aligned}$$

$$\begin{aligned} &\stackrel{(b)}{\geq} \sum_{i < j} \mathbb{E}_R [\max \{H(E_{i,j} | R_{ij} = R) \\ &\quad - H(E_{i,j} \oplus \hat{E}_{i,j} | R_{ij} = R), 0\}] \\ &\stackrel{(c)}{=} \sum_{i < j} \mathbb{E} [\max \{h_2(p(R)) - h_2(q_{ij}(R)), 0\}] \\ &\geq \mathbb{E} [\max \{ \binom{n}{2} h_2(p(R)) - \sum_{i < j} h_2(q_{ij}(R)), 0 \}] \\ &\stackrel{(d)}{\geq} \binom{n}{2} \mathbb{E} [h_2(p(R))] - \binom{n}{2} \mathbb{E} [\min \{h_2(q(R)), h_2(p(R))\}] \end{aligned}$$

In (a), we use the conditional independence of the edge random variables. In the inequality (b), the bound $H(E_{i,j} | R_{ij} = r) - H(E_{i,j} \oplus \hat{E}_{i,j} | R_{ij} = r)$ is useful only when it is greater than 0 because we already have $H(E_{i,j} | R_{ij} = r) - H(E_{i,j} \oplus \hat{E}_{i,j} | \hat{E}_{i,j}, R_{ij} = r) = I(E_{i,j}; \hat{E}_{i,j} | R_{ij} = r) \geq 0$. In (c), we define

$$q_{i,j}(r) \triangleq \min \left\{ \mathbb{P}(E_{i,j} \oplus \hat{E}_{i,j} = 1 | R_{ij} = r), \right. \\ \left. \mathbb{P}(E_{i,j} \oplus \hat{E}_{i,j} = 0 | R_{ij} = r) \right\}$$

for $r \in [0, K]$ and $i < j$. The inequality (d), where we set $q(r) \triangleq \sum_{i < j} \frac{q_{i,j}(r)}{\binom{n}{2}}$ for $r \in [0, K]$, follows immediately from Jensen's inequality and the concavity of the binary entropy function:

$$\sum_{i < j} h_2(q_{i,j}(r)) \leq \binom{n}{2} h_2(q(r)).$$

It immediately follows from the distortion condition that $q(r)$ satisfies the condition, $\binom{n}{2} \mathbb{E}[q(R)] \leq D_n$.

By minimizing the lower bound in (d) over all functions $q(r)$ satisfying the condition $\binom{n}{2} \mathbb{E}[q(R)] \leq D_n$, we get

$$\begin{aligned} I_{G_n|\mathbf{X}}(D_n) &\geq \binom{n}{2} \left[\mathbb{E} [h_2(p(R))] \right. \\ &\quad \left. - \max_{\substack{q: [0, K] \rightarrow [0, 1]: \\ \binom{n}{2} \mathbb{E}[q(R)] \leq D_n}} \mathbb{E} [\min \{h_2(q(R)), h_2(p(R))\}] \right]. \end{aligned} \quad (16)$$

The optimization problem in (16) is equivalent to solving the following problem:

$$A^* \triangleq \max_{\substack{q: [0, K] \rightarrow [0, 1]: \\ q(r) \leq \bar{p}(r), \binom{n}{2} \mathbb{E}[q(R)] \leq D_n}} \mathbb{E} [h_2(q(R))], \quad (17)$$

where we use the notation $\bar{f}(r)$ to denote $\min\{f(r), 1 - f(r)\}$, for a function $f : [0, K] \rightarrow [0, 1]$. Let us consider two separate cases to solve this optimization problem by finding an optimizer q^* . In the case of $\binom{n}{2} \mathbb{E}[\bar{p}(R)] \leq D_n$,

$$A^* = \mathbb{E} [h_2(p(R))]$$

with $q^*(r) = \bar{p}(r)$ because $\mathbb{E} [h_2(q(R))] \leq \mathbb{E} [h_2(p(R))]$ for every $q(r) \leq \bar{p}(r)$. Here, $q^*(r) = \min\{\bar{p}(r), \mu\}$ with $\mu = \frac{1}{2}$.

In the case of $D_n \leq \binom{n}{2} \mathbb{E}[\bar{p}(R)]$. Consider the function $q^*(r) = \min\{\bar{p}(r), \mu\}$ with μ chosen such that $\binom{n}{2} \mathbb{E}[q^*(R)] =$

D_n . It is possible to choose such a μ depending on the value of D_n because $\binom{n}{2}\mathbb{E}[\min\{\bar{p}(r), \mu\}]$ is a continuous¹ and non-decreasing function of $\mu \in [0, \frac{1}{2}]$ with the maximum value being $\binom{n}{2}\mathbb{E}[\bar{p}(r)]$, which is greater than D_n .

We will now show that $A^* = \mathbb{E}[h_2(q^*(R))]$, i.e., $\mathbb{E}[h_2(q(R))] \leq \mathbb{E}[h_2(q^*(R))]$ for any q that satisfies the constraints in the optimization problem (17). Assume that $D_n > 0$ (which implies that $\mu > 0$). On the other hand, if $D_n = 0$ then $A^* = 0$ because $\mathbb{E}[h_2(q(R))] \leq h_2(\mathbb{E}[q(R)]) \leq h_2(D_n/\binom{n}{2}) = 0$.

As a consequence of the concavity of the binary function, we have the following relation:

$$h_2(q(r)) - h_2(q^*(r)) \leq [q(r) - q^*(r)]h'_2(q^*(r)),$$

for any $q(r) \in [0, 1]$ and $q^*(r) \in (0, 1)$ with $h'_2(q^*(r)) = \log_2\left(\frac{1-q^*(r)}{q^*(r)}\right)$. If $q^*(r) = 0$ then $\bar{p}(r) = 0$ (because $\mu > 0$), which means that $q(r) \leq \bar{p}(r) = 0$. By using these observations, we can write

$$\begin{aligned} & \mathbb{E}[h_2(q(R)) - h_2(q^*(R))] \\ & \leq \mathbb{E}\left[[q(R) - q^*(R)] \log_2\left(\frac{1-q^*(R)}{q^*(R)}\right) 1_{q^*(R)>0}\right]. \end{aligned} \quad (18)$$

Let $\mathcal{A} = \{r : \mu \leq \bar{p}(r)\}$. On \mathcal{A} , $q^*(r) = \mu$, and on \mathcal{A}^c , $q^*(r) = \bar{p}(r)$. Since $\binom{n}{2}\mathbb{E}[q(R)] \leq D_n = \binom{n}{2}\mathbb{E}[q^*(R)]$, we have

$$\begin{aligned} 0 & \geq \mathbb{E}[q(R) - q^*(R)] \\ & = \mathbb{E}[[q(R) - q^*(R)] 1_{q^*(R)>0}] \\ & = \mathbb{E}[[q(R) - q^*(R)] 1_{q^*(R)>0} 1_{\mathcal{A}}] \\ & \quad + \mathbb{E}[[q(R) - q^*(R)] 1_{q^*(R)>0} 1_{\mathcal{A}^c}]. \end{aligned} \quad (19)$$

Using the above observations, we can bound (18) as follows:

$$\begin{aligned} & \mathbb{E}[h_2(q(R)) - h_2(q^*(R))] \\ & \leq \mathbb{E}\left[[q(R) - q^*(R)] \log_2\left(\frac{1-q^*(R)}{q^*(R)}\right) 1_{q^*(R)>0}\right] \\ & = \mathbb{E}\left[[q(R) - q^*(R)] \log_2\left(\frac{1-q^*(R)}{q^*(R)}\right) 1_{q^*(R)>0} 1_{\mathcal{A}}\right] \\ & \quad + \mathbb{E}\left[[q(R) - q^*(R)] \log_2\left(\frac{1-q^*(R)}{q^*(R)}\right) 1_{q^*(R)>0} 1_{\mathcal{A}^c}\right] \\ & = \mathbb{E}\left[[q(R) - q^*(R)] \log_2\left(\frac{1-\mu}{\mu}\right) 1_{q^*(R)>0} 1_{\mathcal{A}}\right] \\ & \quad + \mathbb{E}\left[[q(R) - q^*(R)] \log_2\left(\frac{1-\bar{p}(R)}{\bar{p}(R)}\right) 1_{q^*(R)>0} 1_{\mathcal{A}^c}\right] \\ & = \log_2\left(\frac{1-\mu}{\mu}\right) \mathbb{E}[[q(R) - q^*(R)] 1_{q^*(R)>0} 1_{\mathcal{A}}] \\ & \quad + \mathbb{E}\left[[q(R) - q^*(R)] \log_2\left(\frac{1-\bar{p}(R)}{\bar{p}(R)}\right) 1_{q^*(R)>0} 1_{\mathcal{A}^c}\right] \\ & \stackrel{(e)}{\leq} -\log_2\left(\frac{1-\mu}{\mu}\right) \mathbb{E}[[q(R) - q^*(R)] 1_{q^*(R)>0} 1_{\mathcal{A}^c}] \end{aligned}$$

¹As $\min\{\bar{p}(r), \mu\}$ is bounded from above by $\bar{p}(r)$, which is an integrable function, the continuity immediately follows by applying the dominated convergence theorem (DCT) for any convergent sequence $\mu_n \rightarrow \mu$.

$$\begin{aligned} & + \mathbb{E}\left[[q(R) - q^*(R)] \log_2\left(\frac{1-\bar{p}(R)}{\bar{p}(R)}\right) 1_{q^*(R)>0} 1_{\mathcal{A}^c}\right] \\ & = \mathbb{E}\left[[q(R) - q^*(R)] \left[\log_2\left(\frac{1-\bar{p}(R)}{\bar{p}(R)}\right) - \log_2\left(\frac{1-\mu}{\mu}\right)\right] 1_{q^*(R)>0} 1_{\mathcal{A}^c}\right] \\ & = \mathbb{E}\left[[q(R) - \bar{p}(R)] \left[\log_2\left(\frac{1-\bar{p}(R)}{\bar{p}(R)}\right) - \log_2\left(\frac{1-\mu}{\mu}\right)\right] 1_{q^*(R)>0} 1_{\mathcal{A}^c}\right] \\ & \leq 0, \end{aligned}$$

where (e) follows from (19), and the last inequality is due to the fact that $q(r) \leq \bar{p}(r)$ and the fact that on the event \mathcal{A}^c , $\bar{p}(r) \leq \mu$, which implies that $\log_2\left(\frac{1-\bar{p}(r)}{\bar{p}(r)}\right) \geq \log_2\left(\frac{1-\mu}{\mu}\right)$ as the derivative of binary entropy function is non-increasing. This shows that $A^* = \mathbb{E}[h_2(q^*(R))]$, with $q^*(r)$ being an optimizer. Hence we have

$$I_{G_n|\mathbf{X}}(D_n) \geq \binom{n}{2} [\mathbb{E}[h_2(p(R))] - \mathbb{E}[h_2(q^*(R))]], \quad (20)$$

where $q^* : [0, K] \rightarrow [0, 1]$ and $q^*(r)$ is defined as above.

B. An Upper Bound on $I_{G_n|\mathbf{X}}(D_n)$:

Let $q^*(r)$ be the optimizer in (20). Consider the conditional probability distribution $P_{\hat{G}_n|G_n, \mathbf{X}}$ (or equivalently, $P_{\{\hat{E}_{i,j}\}_{i<j}|\{E_{i,j}\}_{i<j}, \mathbf{X}}$) of the form

$$P_{\{\hat{E}_{i,j}\}_{i<j}|\{E_{i,j}\}_{i<j}, \mathbf{X}} = \prod_{i<j} \frac{P_{E_{ij}|\hat{E}_{i,j}, X_i, X_j} P_{\hat{E}_{i,j}|X_i, X_j}}{P_{E_{ij}|X_i, X_j}},$$

where $P_{E_{ij}|\hat{E}_{i,j}, X_i, X_j}(\cdot|x_i, x_j)$ is a binary symmetric channel with the crossover probability $q^*(r_{ij})$, $P_{\hat{E}_{i,j}|X_i, X_j}(1|x_i, x_j) = \frac{p(r_{ij}) - q^*(r_{ij})}{1 - 2q^*(r_{ij})}$ and $P_{\hat{E}_{i,j}|X_i, X_j}(0|x_i, x_j) = \frac{1 - p(r_{ij}) - q^*(r_{ij})}{1 - 2q^*(r_{ij})}$, which are non-negative by virtue of the way the $q^*(r_{ij})$'s are defined. We can easily see that this distribution satisfies the distortion criterion:

$$\mathbb{E}[d_n(G_n, \hat{G}_n)] = \sum_{i<j} \mathbb{P}(E_{i,j} \oplus \hat{E}_{i,j} = 1) = \sum_{i<j} \mathbb{E}[q^*(R)] = D_n,$$

which implies that

$$\begin{aligned} I_{G_n|\mathbf{X}}(D) & \leq I(G_n; \hat{G}_n|\mathbf{X}) \\ & = \binom{n}{2} [\mathbb{E}[h_2(p(R))] - \mathbb{E}[h_2(q^*(R))]]. \end{aligned} \quad (21)$$

By combining (21) and (20), we prove Theorem 1. \square

V. FUTURE WORK

In this work, we obtain upper and lower bounds on the n th order information-distortion function of random geometric graph under the Hamming distortion criterion. We intend to extend the current work in future by considering various distortion measure that take into account more practically relevant features such as clustering, transitivity and so on.

REFERENCES

- [1] Y. Choi and W. Szpankowski, "Compression of graphical structures: Fundamental limits, algorithms, and experiments," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 620–638, Feb. 2012.
- [2] P. Delgosha and V. Anantharam, "Universal lossless compression of graphical data," *IEEE Trans. Inf. Theory*, vol. 66, no. 11, pp. 6962–6976, Nov. 2020.
- [3] H. Nikpey, S. Sarkar, and S. S. Bidokhti, "Compression with unlabeled graph side information," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Taipei, Taiwan, Jun. 2023, pp. 713–718.
- [4] M. W. Wafula, P. K. Vippathalla, J. Coon, and M.-A. Badiu, "Rate-distortion function of the stochastic block model," in *Proc. 57th Asilomar Conf. Signals Syst. Comput.*, Pacific Grove, CA, USA, Oct. 2023, pp. 699–703.
- [5] R. Bustin and O. Shayevitz, "On lossy compression of directed graphs," *IEEE Trans. Inf. Theory*, vol. 68, no. 4, pp. 2101–2122, Apr. 2022.
- [6] M.-A. Badiu and J. P. Coon, "Structural complexity of one-dimensional random geometric graphs," *IEEE Trans. Inf. Theory*, vol. 69, no. 2, pp. 794–812, Sep. 2023.
- [7] E. Abbe, "Graph compression: The effect of clusters," in *Proc. 54th Annu. Allerton Conf. Commun. Contr. Comput.*, Monticello, IL, USA, Sep. 2016, pp. 1–8.
- [8] J. P. Coon, C. P. Dettmann, and O. Georgiou, "Entropy of spatial network ensembles," *Phys. Rev. E*, vol. 97, 042319, Apr. 2018.
- [9] M.-A. Badiu and J. P. Coon, "On the distribution of random geometric graphs," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Vail, CO, USA, Jun. 2018, pp. 2137–2141.
- [10] T. Łuczak, A. Magnier, and W. Szpankowski, "Compression of preferential attachment graphs," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Paris, France, Jul. 2019, pp. 1697–1701.
- [11] K. Turowski, A. Magnier, and W. Szpankowski, "Compression of dynamic graphs generated by a duplication model," *Algorithmica*, vol. 82, no. 9, pp. 2687–2707, Apr. 2020.
- [12] T. Łuczak, A. Magnier, and W. Szpankowski, "Asymmetry and structural information in preferential attachment graphs," *Random Struct. Alg.*, vol. 55, no. 3, pp. 696–718, Mar. 2019.
- [13] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. USA: Wiley-Interscience, 2006.
- [14] T. S. Han, *Information-spectrum methods in information theory*. Berlin, Germany: Springer-Verlag, 2003.
- [15] R. Gray, "A new class of lower bounds to information rates of stationary sources via conditional rate-distortion functions," *IEEE Trans. Inf. Theory*, vol. 19, no. 4, pp. 480–489, Jul. 1973.