REPORT DOCUMENTATION PAGE	Form Approved OMB NO. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
26-02-2021	Final Report	3-Jan-2019 - 30-Nov-2020

4. TITLE AND SUBTITLE	
Final Report: Topology Discovery in Wireless Networks via Spatial Graph Entropy	5a. CONTRACT NUMBER W911NF-19-1-0048
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER 611102
6. AUTHORS	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES	8. PERFORMING ORGANIZATION REPORT NUMBER
University of Oxford University Offices, Wellington Square -	

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES)	10. SPONSOR/MONITOR'S ACRONYM(S) ARO
U.S. Army Research Office P.O. Box 12211 Research Triangle Park, NC 27709-2211	11. SPONSOR/MONITOR'S REPORT NUMBER(S) 74179-NS.6

12. DISTRIBUTION AVAILIBILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Justin Coon
UU	UU	UU	UU		19b. TELEPHONE NUMBER 441-865-2833

Standard Form 298 (Rev 8/98)
Prescribed by ANSI Std. Z39.18

RPPR Final Report
as of 26-Apr-2021

Agency Code: 21XD

Proposal Number: 74179NS
Agreement Number: W911NF-19-1-0048

INVESTIGATOR(S):

Name: Justin Coon
Email: justin.coon@eng.ox.ac.uk
Phone Number: 441865283393
Principal: Y

Organization: **University of Oxford**
Address: University Offices, Wellington Square, Oxford OX1 2JD,
Country: GBR

DUNS Number: 226694883
EIN:

Report Date: 28-Feb-2021
Date Received: 26-Feb-2021

Final Report for Period Beginning 03-Jan-2019 and Ending 30-Nov-2020

Title: Topology Discovery in Wireless Networks via Spatial Graph Entropy

Begin Performance Period: 03-Jan-2019
End Performance Period: 30-Nov-2020

Report Term: 0-Other

Submitted By: Justin Coon
Email: justin.coon@eng.ox.ac.uk
Phone: (441) 865-283393

Distribution Statement: 1-Approved for public release; distribution is unlimited.

STEM Degrees: 1 **STEM Participants:** 3

Major Goals: 1. Develop the theoretical foundations for analyzing the entropy of directed spatial graphs [May 31, 2019 target/actual - 100% complete];

2. Apply the theoretical formulation to study the entropy scaling properties of wireless networks with and without topology constraints (e.g., ad hoc/mesh as well as tree topologies) [Aug 31, 2019 target/October 31, 2019 actual - 100% complete];

NB: A no-cost extension to the project was granted to facilitate the leveraging of other funding for the postdoc involved. As a result, the postdoc worked toward objectives 3 and 4 in a part-time manner. Unfortunately, Covid-19 caused some disruption, since the postdoc and PI were on lockdown and unable to access some written mathematical proofs in their offices. This issue was overcome during the past few months and work progressed to a satisfactory conclusion.

3. Compute the overhead and acquisition time of leading conventional TD algorithms, and benchmark these against lower bounds based on the entropy formalism [Dec 31, 2019 target/June 30, 2020 actual - 100% complete];

4. Investigate novel methods of performing TD that draw on knowledge of entropy scaling [Feb 28, 2020 target/Aug 01, 2020 actual - 100% complete].

Accomplishments: All primary objectives were met. The primary contributions generated within this project are summarized in new entropy bounds for directed and asymptotically connected graphs. Crucially, the bounds on the structural entropy in one dimensional networks point to a very simple compression scheme with encoding that is linear in the number of nodes. For (quasi-) one-dimensional networks, such as convoys, one may use this encoding scheme to propagate information about the topological structure of the network quickly and efficiently to all devices in a time period that grows linearly with the size of the network.

More specifically, in relation to objectives 1 and 2: Entropy scaling for directed networks, asymptotically connected spatial networks, and linear (1-D) networks was analyzed and quantified.

In relation to objectives 3 and 4: A compression scheme that achieves the bound for asymptotically connected networks was determined. The time scaling for TD methods that utilize this scheme is proportional to the complexity scaling. Similar results were found for 1-D networks where structures are required to be encoded.

All details can be found in the attached report.

A secondary line of enquiry was also followed in the project. This broad thread focused more on the physical aspects of communications in wireless networks. It was recognized that, for tactical and other mobile ad hoc networks, one must know something about the signal-to-noise ratio (SNR) distribution in order to ascertain how often TD would need to be carried out (due to potential changes in the network topology owing to mobility). To this end, the SNR corresponding to a link between two devices that undergo mobility according to an Ornstein-Uhlenbeck (OU) model was statistically characterized, which enabled us to further quantify the stochastic properties of the link state. Ultimately, this information was used to study how signal fading and mobility affect the network topology.

In a vein similar to the SNR study, the link-level performance of so-called intelligent reflective surface (IRS) systems was analyzed in order to understand how IRS-enabled links may affect network structure. IRS technology is currently receiving a huge amount of attention owing to its relatively simplistic design and ability to improve the communication channel state. The statistical properties of the SNR for such systems were explored, and it was shown that the SNR follows a Nakagami model when the phase alignment at the reflective surface is imperfect. Furthermore, it was shown that the model is reasonably invariant to the actual phase error distribution.

Finally, attention was given to the idea that tactical and other ad hoc networks are often used to take measurements of physical parameters and propagate these through the network to a central aggregation or processing point. Such measurements are effected by noise or other deleterious events introduced in the channel, and it is known that outdated measurements are less useful than up-to-date data. In this context, a study was undertaken to characterize the value of information (VoI) in hidden Markov models (HMMs). The purpose of concentrating on HMMs is to enable one to study, in a rigorous information-theoretic framework, how noisy channels affect VoI. To make progress and to gain initial insight, it was assumed the measurement of interest relates to the position of a device, which undergoes OU mobility, i.e., it is anchored to a point, but deviates from this point in a manner akin to Brownian motion. The VoI was defined as the mutual information between the actual position of the device at a given time and a sequence of observed measurements of that position, which were attained at previous instants. It should be noted that the aim would be to next apply this VoI formalism to the TD problem, where the topology or structure of the network would replace the device position as the observable of interest. Unfortunately, the project concluded before this topic could be explored further.

Of fundamental interest, but which could not be investigated in the present project, is the consideration of higher-dimensional networks -- especially 2-D -- and the entropy scaling properties of such systems. Additionally, it should be noted that all results reported herein and in the attached report were obtained under the assumption that lossless compression of the network state information is required. A natural modification to this assumption would be to treat the lossy case, whereby a rate-distortion theory of graph compression would need to be developed. This is a rich area of research worthy of further investigation.

Training Opportunities: Nothing to Report

Results Dissemination: Due to COVID-19, little could be done to disseminate the work through lectures and other in-person engagements. Prof. Coon is still scheduled to deliver the Bradley Distinguished Lecture at VA Tech in 2022, or possibly 2023 if the university is still closed next year. He has also been invited to the École Nationale Supérieure de l'Électronique et de ses Applications (ENSEA) near Paris as a CY Advanced Scholar for a three-week period this summer. In both of these engagements, he intends to speak about the work undertaken in the project at a fundamental level.

Papers to report during the past few months are as follows:

Wang, Z., Badiu, M. A., & Coon, J. P. (2021). A Framework for Characterizing the Value of Information in Hidden Markov Models. arXiv preprint arXiv:2102.08841 (submitted to IEEE Trans. Info. Theory).

Badiu, M. A., & Coon, J. P. (2021). Structural Complexity of One-Dimensional Random Geometric Graphs. In preparation (to be submitted to IEEE Trans. Info. Theory).

Honors and Awards: Prof. Justin Coon was the recipient of a CY Advanced Scholar grant for his work on graph entropy. The award will support a visit to ENSEA in France, where he will lecture on this work.

RPPR Final Report
as of 26-Apr-2021

Protocol Activity Status:

Technology Transfer: Nothing to Report

PARTICIPANTS:

Participant Type: PD/PI
Participant: Justin Porter Coon
Person Months Worked: 1.00 **Funding Support:**
Project Contribution:
National Academy Member: N

Participant Type: Postdoctoral (scholar, fellow or other postdoctoral position)
Participant: Mihai-Alin Badiu
Person Months Worked: 12.00 **Funding Support:**
Project Contribution:
National Academy Member: N

International Collaboration:

GBR

GBR

ARTICLES:

Publication Type: Journal Article Peer Reviewed: Y **Publication Status:** 4-Under Review
Journal: IEEE Transactions on Information Theory
Publication Identifier Type: Other Publication Identifier: arXiv:2102.08841
Volume: Issue: First Page #:
Date Submitted: 2/26/21 12:00AM Date Published:
Publication Location:
Article Title: A Framework for Characterising the Value of Information in Hidden Markov Models
Authors: Zijing Wang, Mihai-Alin Badiu, Justin P. Coon
Keywords: Value of information, age of information, hidden Markov models, Ornstein-Uhlenbeck process
Abstract: In this paper, a general framework is formalised to characterise the value of information (VoI) in hidden Markov models. Specifically, the VoI is defined as the mutual information between the current, unobserved status at the source and a sequence of observed measurements at the receiver, which can be interpreted as the reduction in the uncertainty of the current status given that we have noisy past observations of a hidden Markov process. We explore the VoI in the context of the noisy Ornstein-Uhlenbeck process and derive its closed-form expressions. Moreover, we study the effect of different sampling policies on VoI, deriving simplified expressions in different noise regimes and analysing the statistical properties of the VoI in the worst case. In simulations, the validity of theoretical results is verified, and the performance of VoI in the Markov and hidden Markov models is also analysed.
Distribution Statement: 2-Distribution Limited to U.S. Government agencies only; report contains proprietary info
Acknowledged Federal Support: **Y**

RPPR Final Report
as of 26-Apr-2021

CONFERENCE PAPERS:

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: IEEE Global Communications Conference
Date Received: Conference Date: 07-Dec-2020 Date Published:
Conference Location: Taipei
Paper Title: Study of Intelligent Reflective Surface Assisted Communications with One-bit Phase Adjustments
Authors: Tianxiong Wang, Gaojie Chen, Justin P. Coon, Mihai-Alin Badiu
Acknowledged Federal Support: **Y**

Partners

,

I certify that the information in the report is complete and accurate:
Signature:
Signature Date:

Topology Discovery in Wireless Networks via Spatial Graph Entropy

Grant Number: W911NF-19-1-0048
Final Report

Prof. Justin P. Coon and Dr. Mihai-Alin Badiu
Department of Engineering Science
University of Oxford
e-mail: justin.coon@eng.ox.ac.uk

February 26, 2021

Contents

1 Introduction and Summary of Contributions

1.1 Background and Motivation

Uncertainty is pervasive in modern wireless networks. The sources of this uncertainty range from the humans that interact with the networks and the geographic locations of the nodes down to the transmission protocols and the underlying scattering processes that affect signal propagation. Conventionally, knowledge of these characteristics has been used to construct a model of a wireless system, which can be optimized to achieve a stated objective, such as scheduling traffic or attaining a specified rate of communication at a prescribed error rate. These goals are almost always focused on local system features, e.g., individual device performance.

Fundamental questions regarding the *global* implications of uncertainty on network structure – for example, topology, degree distribution, and complexity – have only recently been asked in the context of spatial networks [15, 29]. Yet, the sheer complexity of many modern networks, including wireless sensor networks and multi-tier tactical wireless networks, points to the need to better understand how local uncertainty affects global network performance. For example, it is easy to surmise that the average end-to-end transmission delay observed in a network comprised of two equally sized subnetworks joined by a single bridging link of intermittent reliability will vary significantly over time. This notion extends to operations such as topology discovery and control, routing protocols, and traffic management, all of which are particularly important in tactical and military networks where nodes are mobile and fixed base stations generally do not exist [24].

The structural uncertainty of networks has been quantified in the physics and complex networks communities by invoking various definitions of *graph entropy*. This formalism allows one to quantitatively characterize the complexity or inherent information content of systems that can be described by a graphical model. With regard to communication networks, entropy measures have been exploited to quantify node and route stability [26] with the aim of improving link prediction [47] and routing protocols [2, 9]. Topological uncertainty in dynamic mobile ad hoc networks was investigated in [42] from a network layer perspective, and [37] treated self-organization in networks using a basic graph entropy framework. Crucially, those investigations did not facilitate a quantitative analysis of wireless systems that experience fading and interference, nor did they account for the spatial characteristics of the network (e.g., dependence on node positions).

Recently, a more formal analysis of the scaling properties of wireless ad hoc network entropy was conducted; see [4, 13, 15–18]. These studies were the first to leverage point process models, stochastic geometry, and graph entropy – which come together to create the new formalism of **spatial graph entropy** – to describe a quantitative method of analyzing and controlling complexity in large-scale wireless networks. The entropy of a network represents the minimum number of bits required to describe the network topology, i.e., the lossless compression bound is the entropy of the network. This idea extends to graph structures, i.e., unlabelled graphs. Graph entropy can also be thought of as the uncertainty that one has with regard to the topology (or structural characteristics) of a network. Hence, intuitively, if pairwise connections occur with very high (or very low) probability, the entropy of the network in question is quite low, since the network will almost always be nearly completely connected (or completely disconnected). Hence, by controlling the entropy through connection range scaling, it is possible to limit the amount of memory required to store the network state, or indeed the amount of bandwidth needed to convey the network topology from one device to another.

This project was based around the central hypothesis that knowledge of the structural uncertainty of a wireless network – quantified appropriately using the framework of spatial graph entropy developed in [4, 13, 15–18] – can be used to improve topology discovery (TD) in wireless systems. The development and analysis of TD algorithms is an important sub-discipline of the field of network engineering known as *network tomography*. This larger field is more broadly focused on the inference of internal network characteristics based only on end-point data. Here, we are concerned more with fundamental limits of TD as viewed through the lens of graph entropy.

1.2 Summary of Contributions

The main objectives of the project were set to be the following:

O1 Develop the theoretical foundations for analyzing the entropy of directed spatial graphs;

O2 Apply the theoretical formulation to study the entropy scaling properties of wireless networks with and without topology constraints (e.g., ad hoc/mesh as well as tree topologies);

O3 Compute the overhead and acquisition time of leading conventional TD algorithms, and benchmark these against lower bounds based on the entropy formalism;

O4 Investigate novel methods of performing TD that draw on knowledge of entropy scaling.

The primary methodology that was adopted to work toward these objectives was rooted in the calculation of fundamental limits of graph compression. Indeed, the first two objectives were achieved by directly adapting the theory developed in [15, 17]. For the third objective, entropy bounds were achieved for various network configurations and assumptions. Little in the way of formal results pertaining to conventional TD methods are available or useful in the context of a comparison to these bounds. Although some TD methods were considered, it was found that these were heuristic, and hence it was not fruitful to study them in the formal framework developed in this project. Instead, focus was shifted to the fourth objective, where interesting practical graph compression methods were identified, which in turn can be used in TD algorithms to efficiently convey topological or structural data from one point to another in a network. Full details of these contributions are included herein, with the culmination of the project appearing in final brief, but important, paragraphs of section 4. Some of these details are also included in a draft of a journal paper, which is currently in preparation [6].

A secondary line of enquiry was also followed in the project. This broad thread focused more on the physical aspects of communications in wireless networks. It was recognized that, for tactical and other mobile ad hoc networks, one must know something about the signal-to-noise ratio (SNR) distribution in order to ascertain how often TD would need to be carried out (due to potential changes in the network topology owing to mobility). To this end, the SNR corresponding to a link between two devices that undergo mobility according to an Ornstein-Uhlenbeck (OU) model was statistically characterized, which enabled us to further quantify the stochastic properties of the link state. Ultimately, this information was used to study how signal fading and mobility affect the network topology. Complete details are reported in [14].

In a vein similar to the SNR study, the link-level performance of so-called intelligent reflective surface (IRS) systems was analyzed in order to understand how IRS-enabled links may affect network structure. IRS technology is currently receiving a huge amount of attention owing to its relatively simplistic design and ability to improve the communication channel state. The statistical properties of the SNR for such systems were explored, and it was shown that the SNR follows a Nakagami model when the phase alignment at the reflective surface is imperfect. Furthermore, it was shown that the model is reasonably invariant to the actual phase error distribution. Full details can be found in [5].

Finally, attention was given to the idea that tactical and other ad hoc networks are often used to take measurements of physical parameters and propagate these through the network to a central aggregation or processing point. Such measurements are effected by noise or other deleterious events introduced in the channel, and it is known that outdated measurements are less useful than up-to-date data. In this context, a study was undertaken to characterize the value of information (VoI) in hidden Markov models (HMMs). The purpose of concentrating on HMMs is to enable one to study, in a rigorous information-theoretic framework, how noisy channels affect VoI. To make progress and to gain initial insight, it was assumed the measurement of interest relates to the position of a device, which undergoes OU mobility, i.e., it is anchored to a point, but deviates from this point in a manner akin to Brownian motion. The VoI was defined as the mutual information between the actual position of the device at a given time and a sequence of observed measurements of that position, which were attained at previous instants. Full details can be found in [44, 45]. It should be noted that the aim would be to next apply this VoI formalism to the TD problem, where the topology or structure of the network would replace the device position as the observable of interest. Unfortunately, the project concluded before this topic could be explored further.

The remainder of this report will focus on the contributions made in relation to the primary methodology discussed above.

2 Compression of Erdős-Rényi Graphs

To gain an understanding of the compressibility of network topologies and structures, it is prudent to begin with a simple model. In this case, we turn our attention to the well-known Erdős-Rényi (ER) model where the edges are independent and identically distributed (i.i.d.). Although this is not a spatial graph model, we can obtain some useful insights into the compression problem, which influenced our research on structural compression later in the project (see section 4).

Consider an ER graph G with n nodes and edge probability p. We are concerned with the following question: how compressible is G? It is well known that the lower bound on the average description length of a random source (variable) using a prefix code is given by the entropy of the source. When we concatenate multiple observations from the source to create an *extension* of the code, the average code length bound remains unchanged. If the source observations are not i.i.d., then the lower bound on the average code length is the entropy rate of the source (rather than the entropy).

These classical results persist for ER graphs. Indeed, we can define G is a measurable map from a discrete, finite, measurable space $(\Omega, \mathcal{A})$ equipped with a probability measure P (i.e., a probability space) to an index set $[2^{\binom{n}{2}}] := \{1, \ldots, 2^{\binom{n}{2}}\}$. Hence, we can easily construct a D-ary encoder $C : [2^{\binom{n}{2}}] \to \mathcal{D}$ for G, which maps each index in $[2^{\binom{n}{2}}]$ to a string in the set of all finite-length D-ary strings $\mathcal{D}$. Denoting the length of the code corresponding to a particular graph g as $\ell(g)$, it follows from classical results on data compression that

$$H(G) \leq \mathsf{E}[\ell(G)] \leq H(G) + 1$$

is achievable. Furthermore, if we know the distribution of G, we can use a Huffman code to efficiently encode the graph.

In this section, we first recall the distribution and entropy of G, and then briefly discuss the optimal Huffman encoding approach. We then explore three alternative graph representations. The first is the standard approach of mapping a graph $\omega \in \Omega$ to a binary string of length $\binom{n}{2}$, each bit corresponding to the state of a unique edge in ω. We refer to this representation as the *adjacency construction*. In the second representation, a graph is characterized as a collection of subsets of connected vertices. We refer to this representation as the *vertex construction*. In the third representation, all $\binom{n}{2}$ possible edges in a graph are indexed, and a graph is characterized as an ordered sequence of edge indices. We refer to this representation as the *edge construction*. For each construction, we compute the distribution of the associated random variable that defines it and use this to evaluate the entropy of the representation. We also comment on optimal and suboptimal compression methods for each representation.

2.1 Distribution and Entropy of G

In the following, we refer to a graph by its index $g \in [2^{\binom{n}{2}}]$ and simply call it g. For an ER graph, recall that an edge between nodes i and j exists with probability p

independent of i and j. For a graph g with k edges, it follows that[1]

$$\mathsf{P}(g) = p^k (1-p)^{\binom{n}{2}-k}.$$

Let $w(g)$ denote the number of edges in g. To compute the entropy of G, we take the following steps:

$$\begin{aligned}
H(G) &= \sum_{g} \mathsf{P}(g) \log \frac{1}{\mathsf{P}(g)} \\
&= \sum_{g\colon w(g)=0} \mathsf{P}(g) \log \frac{1}{\mathsf{P}(g)} + \cdots + \sum_{g\colon w(g)=\binom{n}{2}} \mathsf{P}(g) \log \frac{1}{\mathsf{P}(g)} \\
&= \sum_{g\colon w(g)=0} (1-p)^{\binom{n}{2}} \log \frac{1}{(1-p)^{\binom{n}{2}}} + \cdots + \sum_{g\colon w(g)=\binom{n}{2}} p^{\binom{n}{2}} \log \frac{1}{p^{\binom{n}{2}}} \\
&= -\sum_{k=0}^{\binom{n}{2}} \binom{\binom{n}{2}}{k} p^k (1-p)^{\binom{n}{2}-k} \log \left(p^k (1-p)^{\binom{n}{2}-k} \right) \\
&= -\sum_{k=0}^{\binom{n}{2}} \mathsf{P}(w(G)=k) k \log p - \sum_{k=0}^{\binom{n}{2}} \mathsf{P}(w(G)=k) \left(\binom{n}{2} - k \right) \log(1-p) \\
&= -\binom{n}{2} p \log p - \binom{n}{2} (1-p) \log(1-p) \\
&= \binom{n}{2} H(p).
\end{aligned}$$

Note that no specific graph representation was used to arrive at this result. We simply mapped a graph to one of the indices in $[2^{\binom{n}{2}}]$.

Since we know the distribution of each graph (index), we can compress G by using the Huffman algorithm. This would require us to directly encode all $2^{\binom{n}{2}}$ indices, which is prohibitively complex even for relatively small values of n. Even for $n = 7$, the algorithm would have to map sequences to over two million indices. In general, the runtime of Huffman's algorithm scales like $O(m \log m)$, with m denoting the number of symbols in the source alphabet. Applied to the problem of ER graph compression, the scaling becomes $O(n^2 2^{\binom{n}{2}})$. Clearly, this would be impractical for TD applications. It is, therefore, wise to consider other graph representations in the hope that a more computationally friendly solution could be found.

[1] For completeness, it is important to note that the measure P acts on the σ-algebra $\mathcal{A}$. Here, $g \in [2^{\binom{n}{2}}]$, and we use the notation $\mathsf{P}(g)$ to represent the measure of the preimage $\mathsf{P} \circ G^{-1}(g)$ in the usual way.

2.2 Adjacency Construction

Let $G_{\mathrm{a}} : \Omega \to \{0,1\}^{\binom{n}{2}}$ be a random variable that maps a graph to a binary sequence of length $\binom{n}{2}$. Each element of the sequence corresponds to an edge, which may or may not exist, in the graph. Since each edge exists with probability p independent of other edges, it follows that

$$G_{\mathrm{a}} \sim \mathsf{Ber}_p^{\binom{n}{2}}$$

i.e., G_{a} has a multivariate Bernoulli distribution. Thus, the entropy of G_{a} is

$$H(G_{\mathrm{a}}) = H\left(\mathsf{Ber}_p^{\binom{n}{2}}\right) = \binom{n}{2} H(\mathsf{Ber}_p) = \binom{n}{2} H(p).$$

So, the adjacency construction preserves compression optimality.

Again, since we know the distribution of G_{a}, we can apply the Huffman algorithm to obtain an optimal code. We know that the indexing method does not scale (see above). Instead, we can use the adjacency representation to construct a simple, but optimal, encoding scheme. First, consider the naive encoding suggested by the representation: write the full length-$\binom{n}{2}$ binary string. If $p = 1/2$, this approach would be optimal, since the entropy bound would be achieved. But, for general p, we can do better.

2.2.1 Partitioning Based on $w(G_{\mathbf{a}})$

Consider the following straightforward algorithm:

1. Divide the total set of graphs into subsets according to the number of edges in each graph.

2. Within a given subset, graphs are uniformly distributed. So, roughly $\log \binom{\binom{n}{2}}{k}$ bits will be needed to represent a graph, given that we know it has k edges.

3. Use the Huffman algorithm to encode the number of edges in the graph and prefix this code to the length $\log \binom{\binom{n}{2}}{k}$ sequence selected in the previous step.

The code generated by this algorithm is a prefix code. To see this, note that the Huffman code is a prefix code, and it is first decoded to identify the number of bits used for the second, fixed-length part of the code.

To study the efficiency of the code, we derive the following upper bound on the average code length.

$$\begin{aligned}
\mathsf{E}[\ell(G_{\mathrm{a}})] &= \sum_g \ell(g)\mathsf{P}(g) \\
&\leq \sum_{k=0}^{\binom{n}{2}} \mathsf{P}(w(G_{\mathrm{a}}) = k)\left(\ell(k) + \log\binom{\binom{n}{2}}{k} + 1\right) \\
&\leq H(w(G_{\mathrm{a}})) + \mathsf{E}\left[\log\binom{\binom{n}{2}}{w(G_{\mathrm{a}})}\right] + 2 \\
&= -\sum_{k=0}^{\binom{n}{2}} \mathsf{P}(w(G_{\mathrm{a}}) = k)\log\left(p^k(1-p)^{\binom{n}{2}-k}\right) + 2 \\
&= \binom{n}{2}H(p) + 2.
\end{aligned}$$

Hence, the code is efficient. Also, the runtime of the Huffman algorithm scales like $O(n^2 \log n)$. However, an indexing operation is required for each subset of graphs. For the subset corresponding to graphs with k edges, it may be possible to use lexicographic ordering of constant-weight binary sequences of length $\binom{n}{2}$ to index the graphs. The runtimes of naive algorithms appear to scale like $O\left(\binom{n}{2}\binom{\binom{n}{2}}{k}\right)$. Summing over all k gives $O(n^2 2^{\binom{n}{2}})$, so we do not gain much by using this approach.

2.2.2 Partitioning Evenly and Indexing

Since the runtime of the Huffman algorithm scales like $O(m \log m)$, it makes sense to run the algorithm on subsequences of the length-$\binom{n}{2}$ sequence. If we divide each full sequence into sequences of (roughly) length $\log n$, we can index the $2^{\log n} = n$ possible sequences and run the Huffman algorithm (roughly) $\binom{n}{2}/\log n$ times. The total runtime of this approach scales like $O\left(\binom{n}{2}/\log n \cdot n \log n\right) = O(n^3)$, which is much more palatable than the other approaches we have discussed.

Each code (for each subsequence) is efficient, i.e., the average length of the code is close to $H(p)\log n$. Since all subsequences are independent, the average length of the extension is close to $\binom{n}{2}H(p)$.

The drawback to this approach is that it is not based on graph structure, but rather relies on arbitrary indexing. This leads to rather costly storage requirements. It may be interesting to explore this issue further, as there is most likely a trade-off between storage requirements and compression algorithm runtime.

2.3 Vertex Construction

Let $V = [n]$ denote the index set corresponding to the set of vertices. Let $V' = 2^V \setminus (V \cup \emptyset)$ be the collection of all unordered subsets of V of size two or greater. Let

$G_{\mathrm{v}} : \Omega \to \mathcal{V}$, where $\mathcal{V} = 2^{V'}$ is the set of all subsets of V'. For a given graph $\omega \in \Omega$, the interpretation of a subset $S \in V'$ is related to subgraphs of the full graph, but it is not unique. For example, S may represent a clique in ω, or it may represent a star subgraph in ω. Thus, in general, the random variable G_{v} is not well defined. However, we include the general exposition here for completeness.

Consider the following restriction to the previous definition: let V' be the collection of all unordered subsets of V of size two. There are $\binom{n}{2}$ elements in V', which can be interpreted as the set of all possible edges in an undirected graph with n nodes. Thus, $\mathcal{V} = 2^{V'}$ contains $2^{\binom{n}{2}}$ elements, and for each element $g \in \mathcal{V}$, the preimage $G_{\mathrm{v}}^{-1}(g) \in \Omega$ is a unique (measurable) graph.

Let $g \in \mathcal{V}$ be a graph described according to a (restricted) vertex construction. We can equivalently write

$$g = \{\{v_{11}, v_{12}\}, \{v_{21}, v_{22}\}, \ldots\}$$

where $v_{ij} \in ([n] \cup \emptyset)$ is a vertex index for $i = 1, \ldots, w(g)$ and $j = 1, 2$. To encode g, we will need to encode v_{ij} efficiently. The length of $C(g)$ is

$$\ell(g) = \sum_{i=1}^{w(g)} (\ell(v_{i1}) + \ell(v_{i2}))$$

where we define $\ell(g) := \ell(\emptyset)$ when the sum is empty. Since edges occur independently with probability p, for a graph with weight $w(g)$, we have no *a priori* knowledge of how likely any vertex will appear in this representation of the graph. As a result, we have that

$$\ell(g) = \begin{cases} 2w(g)\ell(v), & w(g) > 0 \\ \ell(\emptyset), & w(g) = 0 \end{cases}$$

where v denotes an arbitrary vertex.

Note that $\mathsf{P}(G_{\mathrm{v}} = \emptyset) = (1-p)^{\binom{n}{2}}$. Consequently, it is more likely than not that a graph will be nonempty for all nontrivial configurations. Indeed, a quick analysis reveals that $p < 1 - e^{-1/n} \simeq 1/n$ is required for $\mathsf{P}(w(G_{\mathrm{v}}) = 0) > \mathsf{P}(w(G_{\mathrm{v}}) > 0)$, which implies the mean degree is less than one. We can represent the event $\{G_{\mathrm{v}} = \emptyset\}$ with a single bit, say "0". For other graphs, we simply prefix the encoding of the graph with a "1" and use roughly $\log n$ bits to represent each vertex index. Using this simple encoding strategy, let us calculate the average code length for the restricted vertex construction:

$$\begin{aligned}
\mathsf{E}[\ell(G_{\mathrm{v}})] &= \sum_{g} \ell(g)\mathsf{P}(g) \\
&= \sum_{k=0}^{\binom{n}{2}} \sum_{g:\, w(g)=k} \ell(g)\mathsf{P}(g) \\
&\le 1 \cdot \mathsf{P}(w(G_{\mathrm{v}})=0) + \sum_{k=1}^{\binom{n}{2}} \mathsf{P}(w(G_{\mathrm{v}})=k) \cdot 2k(\log n + 2) \\
&= (1-p)^{\binom{n}{2}} + 2p\binom{n}{2}(\log n + 2).
\end{aligned}$$

This upper bound is very close to $H(G)$ for small p. However, it is important to note that vertex construction does not yield a prefix code. Hence, a termination sequence must be appended to each code, which reduces the efficiency of the code.

2.4 Edge Construction

Let $E = [\binom{n}{2}]$ denote an index set corresponding to the set of edges. Let $G_{\mathrm{e}} : \Omega \to \mathcal{E}$, where $\mathcal{E} = 2^E$ is the set of all subsets of E. For a given graph $\omega \in \Omega$, the interpretation of a set $g \in \mathcal{E}$ is that it uniquely represents ω since g specifies the edges that exist in the graph.

Let $g \in \mathcal{E}$ be a graph described according to the edge construction. We can equivalently write

$$g = \{e_1, e_2, \ldots\}$$

where $e_i \in ([\binom{n}{2}] \cup \emptyset)$ is an edge index for $i = 1, \ldots, w(g)$. To encode g, we will need to encode e_i efficiently. The length of $C(g)$ is

$$\ell(g) = \sum_{i=1}^{w(g)} \ell(e_i)$$

where, again, we define $\ell(g) := \ell(\emptyset)$ when the sum is empty. Edges occur independently with probability p. Hence, we have no *a priori* knowledge of how likely any edge will appear in this graph construction. As a result, we have that

$$\ell(g) = \begin{cases} w(g)\ell(e), & w(g) > 0 \\ \ell(\emptyset), & w(g) = 0 \end{cases}$$

where e denotes an arbitrary vertex.

In a similar manner as discussed above, we select a single "0" to represent $\{G_{\mathrm{e}} = \emptyset\}$. For other graphs, we prefix the encoding of the graph with a "1" and use roughly $\log \binom{n}{2}$

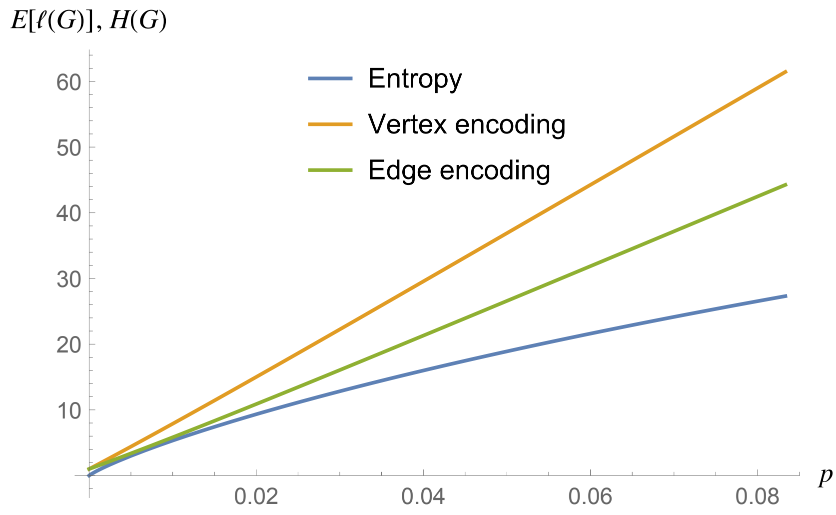

Figure 1: The entropy and average code lengths for different constructions plotted against the edge probability p. The number of nodes is $n = 12$.

bits to represent each edge index. The average code length for this construction is given by

$$\begin{aligned}
\mathsf{E}[\ell(G_{\mathrm{e}})] &= \sum_{g} \ell(g)\mathsf{P}(g) \\
&= \sum_{k=0}^{\binom{n}{2}} \sum_{g:\, w(g)=k} \ell(g)\mathsf{P}(g) \\
&\leq 1 \cdot \mathsf{P}(w(G_{\mathrm{v}}) = 0) + \sum_{k=1}^{\binom{n}{2}} \mathsf{P}(w(G_{\mathrm{v}}) = k) \cdot k \left(\log \binom{n}{2} + 2 \right) \\
&= (1-p)^{\binom{n}{2}} + p\binom{n}{2} \left(\log \binom{n}{2} + 2 \right) \\
&= (1-p)^{\binom{n}{2}} + p\binom{n}{2} (\log n + \log(n-1) + 1).
\end{aligned}$$

As with the vertex construction, this upper bound is very close to $H(G)$ for small p, but the edge construction does not yield a prefix code. The code can be modified easily to give a prefix code, which will reduce the efficiency. The reduction will be minimal for large n.

2.5 Summary and Discussion

In this initial work, a fundamental view of graph compression was observed using an ER model. The purpose of this work was to explore how different graph representations yield different limits on compressibility. These discrepancies manifested clearly and suggested that it may be possible in some cases to represent graphs using their substructures (i.e., motifs). This basic idea will be exploited in section 4 to determine the compression bounds for (quasi-) one-dimensional spatial networks.

3 Compression of Random Geometric Graphs

In this section, we turn our attention to the lossless compression limits of random geometric graphs (RGGs). To this end, we compute bounds on the entropy of directed RGGs and analyze the scaling behavior of these bounds in the limit of large numbers of nodes (or high node density).

3.1 Preliminaries

Consider a set of n nodes $\mathcal{V}$ embedded in a compact space $\mathcal{K}_d \subset \mathbb{R}^d$. The (random) locations of the nodes are denoted by the random variables $Z_1, \ldots, Z_n$, which are i.i.d. with density f_Z defined on $\mathcal{K}_d$. The (Euclidean) distance between two nodes, say i and j, is captured by the random variable $R_{ij} = |Z_i - Z_j|$. We can collect all pairwise distances into the length-$n(n-1)/2$ vector $\boldsymbol{R}_n = (R_{ij})_{i<j}$. We assume here that $\boldsymbol{R}_n$ has a well defined density $f_{\boldsymbol{R}_n}$. Note that the marginal density $f_{R_{ij}}$ is defined on the interval $\mathcal{I} = [0, \mathrm{dia}(\mathcal{K}_d)]$.

We are interested in the distribution of the random graph G_n that is formed by connections between nodes in $\mathcal{V}$. The existence of pairwise connections is determined by node proximity as well as node properties. Let $p_{ij} : \mathcal{I} \to [0,1]$ denote a mapping that indicates the probability that a connection exists from i to j, i.e., there is a directed edge (i,j). Note that for undirected graphs, we must have $p_{ij} = p_{ji}$. We will be concerned with directed graphs unless stated otherwise. Under this formalism, we define a set of $n(n-1)$ Bernoulli random variables $\{X_{ij}\}_{i \neq j}$ such that $X_{ij} = 1$ indicates that directed edge (i,j) exists, and $X_{ij} = 0$ indicates its absence. We collect these variables into the vector $\boldsymbol{X}_n = (X_{ij})_{i \neq j}$. It follows that that $\mathsf{P}(G_n) \equiv \mathsf{P}(\boldsymbol{X}_n)$.

3.2 Entropy of Directed RGGs

The entropy of G_n is defined as

$$H(G_n) \coloneqq \mathsf{E}[-\log \mathsf{P}(G_n)] = \mathsf{E}[-\log \mathsf{P}(\boldsymbol{X}_n)] =: H(\boldsymbol{X}_n). \tag{1}$$

The $\{X_{ij}\}$ variables are correlated, so the direct computation of this quanitity does not appear to be possible in general. However, we can immediately write the following classical upper bound.

$$H(G_n) \leq \sum_{i \neq j} H(X_{ij}). \tag{2}$$

If we assume $p_{ij} = p$ for all i and j, we have the bound

$$H(G_n) \leq n(n-1)H(X) = n(n-1)H(\bar{p}) \tag{3}$$

where

$$\bar{p} = \mathsf{E}[p(R)] = \int_{\mathcal{I}} f_R(r) p(r) \,\mathrm{d}r. \tag{4}$$

This is the directed analogue to the results presented in [15]. The only difference is the prefactor is doubled in the directed case. Thus, all results reported in [15] can be extended trivially to directed graphs. Furthermore, if we assume (i, j) exists if and only if (j, i) exists, we recover the undirected model:

$$H(G_n) \leq \frac{n(n-1)}{2} H(\bar{p}). \tag{5}$$

A more interesting bound can be obtained by slightly reducing our assumption of independence and writing

$$H(G_n) \leq \sum_{i<j} H(X_{ij}, X_{ji}). \tag{6}$$

Here, we have effectively assumed that forward and reverse links for different pairs of nodes are independent, but dependence exists for the links associated with the same pair of nodes. If we suppose that $p_{ij} = p$ for all i and j, we can write

$$H(G_n) \leq \frac{n(n-1)}{2} H(X, Y) \tag{7}$$

where X and Y are Bernoulli random variables that represent the link states in the forward direction (X) and the reverse direction (Y). Moreover, if the forward and reverse links are statistically independent when conditioned on the distance between the nodes, we have that

$$\begin{aligned} \mathsf{P}(x, y) &= \mathsf{E}[\mathsf{P}(X, Y \mid R)] \\ &= \int_{\mathcal{I}} f_R(r) p(r)^{x+y} (1 - p(r))^{2-x-y} \, \mathrm{d}r \end{aligned}$$

from which $H(X, Y)$ can be evaluated easily.

Although bounds such as those given in (5) and (7) are rather simple, there is evidence to suggest that they provide useful estimates for the minimum description length of graph topologies [15]. The $O(n^2)$ scaling of these bounds (for fixed link entropy) is predictable. But a number of key questions arise regarding the effect that local connectivity has on this scaling. For example, how does entropy scale for asymptotically connected graphs? This question is answered below.

3.3 Entropy of Asymptotically Connected Graphs

A graph G_n is said to be *asymptotically connected* if $\lim_{n\to\infty} \mathsf{P}(\{G_n \text{ is connected}\}) = 1$. Gupta and Kumar [27] showed that when G_n resides in the circle of unit area in $\mathbb{R}^2$ and the deterministic connection range satisfies $\pi r_0(n)^2 = n^{-1}(\log n + c(n))$, then G_n is asymptotically connected if and only if $c(n) \to +\infty$. This result has been generalized to RGGs contained in the d-dimensional unit box $\mathcal{K}_d = [0,1]^d$, for which the critical pairwise connection distance threshold for asymptotic connectivity is

$$r_c(n) = \left(\frac{\log n}{nV_d}\right)^{\frac{1}{d}} \tag{8}$$

where V_d is the volume of a unit ball in d dimensions.

It is of interest to analyze the entropy of RGGs under the assumption that the pairwise connection range is on the critical threshold. Such analysis would logically yield information about the compressibility of (asymptotically) connected graphs. Before considering a more general case, we begin with a brief illustration for a path graph (a.k.a., a linear graph). This example is, in essence, non-geometric. Nevertheless, it will serve as a useful point of reference when studying RGGs. A path Π_n consisting of n nodes can be written as the node index sequence $\Pi_n = k_1 k_2 \cdots k_n$ where $k_i \in \mathcal{V}$ for $i = 1, \ldots, n$. There are clearly $n!$ ways to arrange the sequence if we assume it is a directed path; for undirected paths, there are $n!/2$ unique paths. Suppose that each path appears with equal probability, e.g., for undirected graphs, $\mathsf{P}(\Pi_n = \pi_i) = 2/n!$ for all i. Then the entropy of the n-node random path graph is

$$H(\Pi_n) = \log n! + O(1) \sim n \log n \tag{9}$$

where the $O(1)$ term is included to capture the reduction of $\log 2$ for the case where Π_n is undirected. This relation is completely intuitive, since one can imagine a "compression" scheme whereby the nodes in the path are simply listed in order[2].

3.3.1 Undirected Graphs

Now, let us consider an undirected RGG constructed in $\mathcal{K}_d = [0,1]^d$. We treat the canonical model here, where the connection function is defined as

$$p(r) = \begin{cases} 1, & r \le r_0 \\ 0, & r > r_0. \end{cases} \tag{10}$$

Since we are interested in asymptotically connected graphs, we choose the connection range to be the critical threshold $r_0(n) = r_c(n)$. To compute the simple entropy bound given in (5), we must evaluate (4). For small arguments, the pair distance density admits the expansion

$$f_R(r) = \frac{2\pi^{d/2} r^{d-1}}{\Gamma(d/2)}(1 - \beta r + O(r^2)) \tag{11}$$

where β is a term (independent of n) that represents the volume of the projection of the space $\mathcal{K}_d$ onto a $(d-1)$-dimensional hyperplane averaged over all directions on a $(d-1)$-dimensional sphere (referenced to the origin). Hence, we have

$$\bar{p} = \frac{2\pi^{d/2} r_0^d}{d\Gamma(d/2)} + O(r_0^{d+1}) = \frac{\log n}{n} + o\left(\frac{\log n}{n}\right). \tag{12}$$

[2]Roughly $\log n$ bits are needed to specify a node index, and this must be done n times.

The binary entropy function for small arguments evaluates to

$$H_2(x) = (1 - \log x)x + O(x^2). \tag{13}$$

Combining these results with (5) gives us an asymptotic equivalence for the entropy bound for asymptotically connected graphs:

$$H(G_n) \leq \frac{n \log n}{2} (\log n - \log \log n + 1) + o(n \log n). \tag{14}$$

Comparing to the path graph example, we see a modest increase in the entropy by a factor of $\frac{1}{2} \log n$ in the limit of large n. This is not completely a result of the fact that we study an upper bound here. Indeed, an asymptotically connected RGG will exhibit clustering, which does not depend on the system size (only the dimension) [21]. Hence, more information is captured in the topology of these graphs than for path graphs, and thus the former are not as compressible as the latter.

Using this analysis, we can consider a simple TD scheme in which each node of the network communicates a list of its neighbors through the network. For n nodes, $\log n$ bits are required to identify a node. In the asymptotically connected regime, each node has a number of neighbors that is Poisson distributed with mean $\log n$. By the law of large numbers, the number of bits collected from all of the n nodes to identify the network topology is thus $n(\log n)^2$ bits. Note that in this scheme each link is counted twice, so a slightly refined scheme would reduce the number of required bits by a factor of two. Interestingly, this corresponds to the dominating term in the upper bound given above.

3.3.2 Directed Graphs

To obtain a simple upper bound for asymptotically connected directed graphs, we use (3). The only difference in the result relative to (14) is a factor of two.

3.4 Random Number of Nodes

It is easy to generalize the results above to the case where the number of nodes is random. Specifically, let $N : \Omega_N \to \Omega'_N$, where $\Omega'_N \subset \mathbb{N}_0$, be a square integrable random variable that defines the number of nodes in $\mathcal{K}_d$. For now, we refrain from describing N by a particular distribution; suffice to say it has probability measure μ_N. We are now concerned with the random graph G, the distribution of which contains information about N as well as the topology. We can, in general, write

$$H(G, N) = H(N \mid G) + H(G) = H(G \mid N) + H(N). \tag{15}$$

But since G encodes the value of N, we have that $H(N \mid G) = 0$. Thus,

$$H(G) = H(G \mid N) + H(N) = \mathsf{E}_{\mu_N}[H(G_N)] + H(N) \tag{16}$$

where E_{μ_N} signifies that the expectation is taken with respect to the measure μ_N. Applying the simple (loose) bound for a directed graph given above yields

$$H(G) \leq H(N) + \mathsf{E}_{\mu_N}[N(N-1)]H_2(\bar{p}). \tag{17}$$

Application of the tighter bound gives

$$H(G) \leq H(N) + \frac{1}{2}\mathsf{E}_{\mu_N}[N(N-1)]H(X,Y). \tag{18}$$

We may consider some examples in order to better understand the rate of growth of the bounds given here. Suppose N is uniformly distributed on $\Omega'_N = \{0,\ldots,m\}$. Then we have that $H(N) \leq \log(m+1)$ and $\mathsf{E}_{\mu_N}[N(N-1)] = m(m-1)/3$. Hence, we see that the entropy associated with the graph topology dominates. Now suppose N is Poisson distributed with density λ (note that $\Omega'_N = \mathbb{N}_0$). In this case, we have that $\mathsf{E}_{\mu_N}[N(N-1)] = \lambda^2$ and $H(N) \sim \frac{1}{2}\log(2\pi\mathsf{e}\lambda)$ for large λ. Interestingly, the square-law dominance appears once more in the term related to the graph topology. This simple analysis reveals that the main task that must be considered in TD algorithms is the compression of the network topology; encoding the number of devices that are active in the network is relatively straightforward by comparison.

4 Structural Entropy of One-Dimensional Random Geometric Graphs

In wireless communications, one-dimensional (1-D) random geometric graphs (RGGs) can model a vehicular network, sensors deployed along a river, a military convoy, train backbone network, etc. Many of the graph properties (e.g., connectivity, average degree) depend on the random structures in the ensemble, and not on the labels of the nodes. Therefore, it is relevant to characterize the structures and associated properties induced by the assumed graph model when the nodes are indistinguishable (i.e., their labels are irrelevant). In particular, it is interesting to study how rich (complex) the ensemble of structures is, how the complexity scales with the number of nodes and how it is influenced by the connection range. Below, we provide details of our investigations into these questions, which conclude with simple rules on how to compress 1-D network structures.

4.1 Preliminaries

4.1.1 Graph Model

The one-dimensional random geometric graph $G_{n,r}$ is constructed by placing n points randomly on a line segment and connecting by an edge every pair of nodes that are separated by a distance smaller than a threshold r. More formally, assume the n points have the random locations $X_1, \ldots, X_n$ in the interval $[0, 1]$. For now, we refrain from defining the (joint) distribution of $X_1, \ldots, X_n$, since some results shown below are universal. However, we will consider the specific case where $X_1, \ldots, X_n$ are i.u.d.. The corresponding graph $G_{n,r} = (V, E)$ has the set of vertices $V = \{1, \ldots, n\}$ and the set of undirected edges $E = \{(i, j) \in V^2 \mid |X_i - X_j| \leq r,\ i < j\}$, for a fixed connection range $r > 0$.

The graph $G_{n,r}$ depends deterministically on the node locations. However, it does not retain the locations but only the connectivity among the nodes. Also, an infinite number of configurations of locations may be mapped to the same graph. We denote by $\mathcal{G}_{n,r}$ the set of all possible graphs that are produced by this model, i.e., the support of $G_{n,r}$. Fig. 2 illustrates two such graphs. The probability distribution of $(X_1, \ldots, X_n)$ induces a distribution of graphs in $\mathcal{G}_{n,r}$, such that a graph $G \in \mathcal{G}_{n,r}$ occurs with probability $P(G)$.

4.1.2 Connectivity

A graph is said to be connected if there exists at least one path between every pair of distinct vertices. The connectivity of one-dimensional random geometric graphs has been studied extensively, e.g., see [30] and the references therein. Trivially, $G_{n,r}$ is connected for $r \geq 1$, as $G_{n,r}$ is deterministically a complete graph in that case. For $r \in (0, 1)$ and finite n, the graph is connected with a probability smaller than one

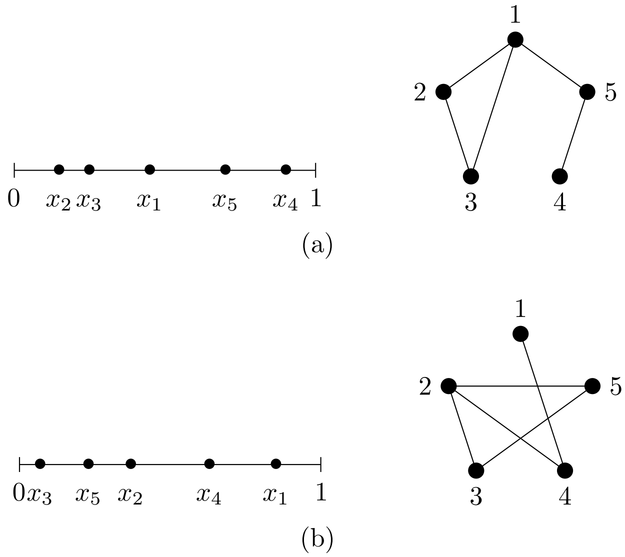

Figure 2: Two outcomes of $G_{n,r}$ for $n = 5$ and $r = 1/3$; the points are i.u.d. on $[0, 1]$ (left); the resulting graph (right).

in general (e.g., when the random locations have a non-vanishing density on $[0, 1]$); for $X_1, \dots, X_n$ independent and uniformly distributed, the probability is available in a closed form expression [22, 25].

Most of the works consider the connectivity of $G_{n,r}$ for large n and study how the connection range r should scale with n to achieve connectivity, i.e., what scaling functions $r : \mathbb{N} \to \mathbb{R}_+$, $n \to r_n$, are appropriate. For i.u.d. point locations, the connectivity exhibits a typical behaviour in the sense that there exists a critical range function $r_n^* = \frac{\log n}{n}$ such that, as n becomes large, the graph is connected (or disconnected) with high probability depending on how the scaling r that is being used deviates from the critical scaling r^*. Specifically, for a connection range function in the form $r_n = r_n^* + \frac{\alpha_n}{n}$ for some $\alpha : \mathbb{N} \to \mathbb{R}$, it holds that [30]

$$\lim_{n\to\infty} \mathsf{P}(G_{n,r} \text{ is connected}) = \begin{cases} 0, & \text{if } \lim_{n\to\infty} \alpha_n = -\infty, \\ 1, & \text{if } \lim_{n\to\infty} \alpha_n = \infty. \end{cases} \tag{19}$$

For example, a deviation function $\alpha_n = \pm \log \log n$ determines $G_{n,r}$ to be connected or disconnected with probability one depending on the sign, in the limit of large n. More generally, it is shown in [31] that graph connectivity also obeys a strong zero-one law when the distribution of the point locations has a non-vanishing density.

4.1.3 Graphical Structures

The graphical structures produced by the model of Sec. 4.1.1 can be defined formally based on the notion of graph isomorphism.

Definition 1 (Structure). *Two graphs $G_1 = (V, E_1) \in \mathcal{G}_{n,r}$ and $G_2 = (V, E_2) \in \mathcal{G}_{n,r}$ have the same structure, which is denoted as $G_1 \simeq G_2$, if and only if there exists a graph isomorphism between them, i.e., if and only if there exists an adjacency-preserving permutation of the vertices $\pi : V \to V$ such that $(i, j) \in E_1 \iff (\pi(i), \pi(j)) \in E_2$, for all $i, j \in V$.*

For example, one can verify that the two graphs in Fig. 2 have the same structure, because the permutation $\pi(1) = 2$, $\pi(2) = 3$, $\pi(3) = 5$, $\pi(4) = 1$ and $\pi(5) = 4$ (applied to the vertices in Fig. 2a) is edge preserving.

The relation '$\simeq$' is an equivalence relation, which partitions the set $\mathcal{G}_{n,r}$ into disjoint equivalence classes. All graphs in an equivalence class have the same structure, while graphs in different classes are structurally different. Thus, one can identify a structure by its defining equivalence class or by a representative member of that class. In the following, we denote by $\mathcal{S}_{n,r}$ the set of all structures of $\mathcal{G}_{n,r}$.

4.1.4 Structural Entropy

The probability distribution of $G_{n,r}$ with pmf $P(G)$, $G \in \mathcal{G}_{n,r}$, induces a probability distribution over the set of structures $\mathcal{S}_{n,r}$. Let $f : \mathcal{G}_{n,r} \to \mathcal{S}_{n,r}$ be a deterministic surjection that maps each graph in $\mathcal{G}_{n,r}$ to its corresponding structure in $\mathcal{S}_{n,r}$. Accordingly, the preimage $f^{-1}(S)$ of a structure $S \in \mathcal{S}_{n,r}$ is the set of all graphs in $\mathcal{G}_{n,r}$ that belong to the equivalent class represented by S. Based on the distribution of graphs, the probability of a structure $S \in \mathcal{S}_{n,r}$ can thus be expressed as

$$\mathsf{P}(S) = \sum_{G \in f^{-1}(S)} P(G). \tag{20}$$

The entropy of $G_{n,r}$ is

$$H(G_{n,r}) = -\sum_{G \in \mathcal{G}_{n,r}} P(G) \log_2 P(G) \tag{21}$$

whereas the structural entropy is defined by [12]

$$H(S_{n,r}) = -\sum_{S \in \mathcal{S}_{n,r}} P(S) \log_2 P(S) \tag{22}$$

4.2 Structures of $\mathcal{G}_{n,r}$

4.2.1 Graph Ordering

A one-dimensional random geometric graph is naturally visualized by depicting its nodes in the order of their locations. For example, by doing so, we realize that the graphs in Fig. 2a and 2b have the same structure, which is illustrated in Fig. 3. In the following such a representation of structure is formally justified. For connected structures, we

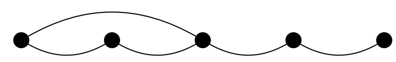

Figure 3: The common structure of the graphs depicted in Fig. 2. The nodes are displayed in the order of the point locations and their labels are discarded.

show that each equivalence class can be represented by a particular graph whose nodes are indexed according to the order of the underlying point locations. This special graph, which we call "ordered graph", is then designated as the structure representative. Disconnected structures can be treated as a collection of smaller, connected structures.

Definition 2 (Ordered graph)**.** *Let $X_1, \dots, X_n \in [0, 1]$ be the random locations of n points and let $X_{(1)}, \dots, X_{(n)}$ be the order statistics of the n locations, i.e., $X_{(k)}$ is the kth smallest of the $X_1, \dots, X_n$. The random graph $G_{n,r}$ is constructed as in Sec. 4.1.1. We define the ordered graph $\hat{G}_{n,r}$ as the graph with nodes $V = \{1, \dots, n\}$ and edges $\hat{E} = \{(i, j) \in V \times V \mid |X_{(i)} - X_{(j)}| \leq r, i < j\}$.*

In the following, $\hat{\mathcal{G}}_{n,r}$ denotes the support of the random ordered graph $\hat{G}_{n,r}$.

Lemma 1. *The graphs $G_{n,r}$ and $\hat{G}_{n,r}$ are isomorphic.*

Proof. Let $\sigma : \{1, \dots, n\} \to \{1, \dots, n\}$ be the bijection that reindexes the points in ascending order of their location, such that $X_{\sigma^{-1}(1)} \leq X_{\sigma^{-1}(2)} \leq \dots \leq X_{\sigma^{-1}(n)}$, i.e., $X_{\sigma^{-1}(i)} \equiv X_{(i)}$ is the ith order statistic of the random locations. From the definitions of $G_{n,r}$ and $\hat{G}_{n,r}$, it can be verified that σ is an edge-preserving permutation and therefore an isomorphism. □

For example, the sets of locations in Fig. 2a and Fig. 2b generate the same ordered graph, which is depicted without the vertex labels in Fig. 3.

4.2.2 Maximal Cliques

In the following, an ordered geometric graph is represented in terms of its constituent maximal cliques. A clique of a graph is a subset of the vertices of the graph, such that any two vertices are adjacent, i.e., the subgraph induced by the clique is complete. A clique is maximal if it is not a subset of a larger clique, i.e., no adjacent vertex can be included to extend the clique. The number of maximal cliques of a graph with n vertices can take any value between one and n, the extremes corresponding to the complete graph and respectively the empty graph (in which case each maximal clique consists of one isolated node).

Ordered graphs have the following basic property.

Lemma 2. *If nodes i and j, $i < j$, of $\hat{G}_{n,r}$ are connected by an edge, then the set of vertices $\{i, i + 1, \dots, j\}$ is a clique.*

Proof. Given that (i, j) is an edge, $|X_{(j)} - X_{(i)}| \leq r$. Furthermore, for the order statistic $(X_{(1)}, \ldots, X_{(n)})$, it holds that $|X_{(l)} - X_{(k)}| \leq |X_{(j)} - X_{(i)}| \leq r$, for all $k, l \in \{i, i+1, \ldots, j\}$. It follows that (k, l) is an edge, for every $k, l \in \{i, i+1, \ldots, j\}$, and therefore $\{i, i+1, \ldots, j\}$ is a clique. □

Thus, a maximal clique of $\hat{G}_{n,r}$ can be identified by its end-nodes only (i.e., those with lowest and highest indices), so that we denote a maximal clique $\{a, a+1, \ldots, b\}$, $a \leq b$, by $[a : b]$. In this way, any ordered graph having k maximal cliques ($1 \leq k \leq n$) can be represented by the set $\{[a_1 : b_1], \ldots, [a_k : b_k]\}$, where the indices of the end-nodes of the maximal cliques need to satisfy

$$\begin{aligned} & 1 = a_1 < a_2 < \ldots < a_k \leq n, \\ & 1 \leq b_1 < b_2 < \ldots < b_k = n, \\ & a_{i+1} \leq b_i + 1, \; i = 1, \ldots, k-1. \end{aligned} \tag{23}$$

The above conditions ensure that the maximal cliques cannot include one another, although they may be overlapping, and each vertex belongs to at least one maximal clique. When $a_{i+1} \leq b_i$, the ith and $(i+1)$th maximal cliques are overlapping, whereas when $a_{i+1} = b_i + 1$ there is a break in the ordered graph between vertices b_i and $b_i + 1$. Also, it is possible that $a_i = b_i$ for some i, which occurs when node a_i is isolated and thus constitutes, itself, the ith maximal clique. The example in Fig. 4 illustrates the decomposition of an ordered graph into its constituent maximal cliques.

4.2.3 Ordered Graphs and Structures

Lemma 3. *Any isomorphism π between two connected ordered graphs in $\hat{\mathcal{G}}_{n,r}$ is either the identity $\pi(i) = i$, for all $i \in V$, or the "backward identity" $\pi(i) = n - i + 1$, for all $i \in V$.*

Proof. Let G and G' be two outcomes of $\hat{G}_{n,r}$ that are connected and isomorphic. We consider their maximal-clique representations $G \equiv \{[a_1 : b_1], \ldots, [a_k : b_k]\}$ and $G' \equiv \{[a'_1 : b'_1], \ldots, [a'_{k'} : b'_{k'}]\}$, where the end-nodes satisfy the conditions (23) with the stronger restriction that the strict inequalities $a_{i+1} < b_i + 1$ and $a'_{j+1} < b'_j + 1$ must hold, because the graphs are connected and therefore consecutive maximal cliques must overlap. Note that the isomorphism π maps each maximal clique of G onto a unique maximal clique of G'. Therefore, the two graphs have the same number of maximal cliques, i.e., $k' = k$. Moreover, since consecutive maximal cliques are overlapping, any two successive maximal cliques of G are mapped by π onto two consecutive maximal cliques of G'. In particular, noting that the first (i.e., leftmost) maximal clique $[1 : b_1]$ of G does not have a preceding maximal clique, it follows that $[1 : b_1]$ is mapped onto either $[1 : b'_1]$, in which case $b'_1 = b_1$ and $\pi(i) = i$, $i = 1, \ldots, b_1$, or $[a'_k, n]$, in which case $a'_k = n - b_1 + 1$ and and $\pi(i) = n - i + 1$, $i = 1, \ldots, b_1$. Then, by considering all the maximal cliques in succession, it results that π can only be the identity or the backward identity. □

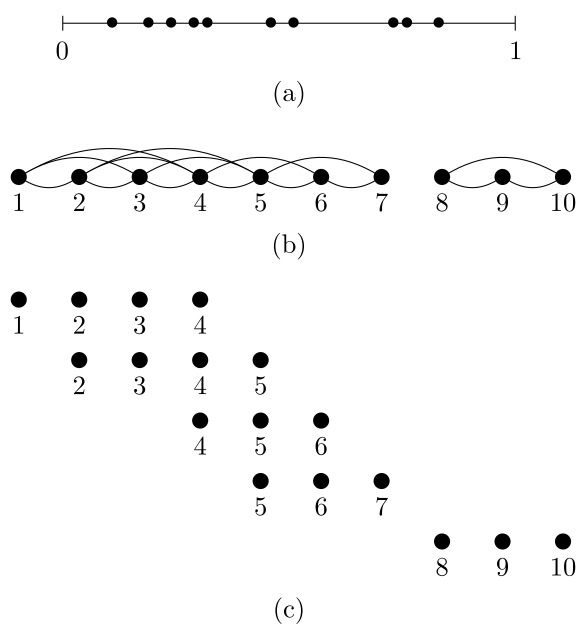

Figure 4: Illustration of the decomposition of an ordered graph into maximal cliques: (a) the locations of $n = 10$ points (b) the ordered graph determined by the set of locations and connection range $r = 0.2$; (c) the ordered graph consists of five maximal cliques: $[1 : 4]$, $[2 : 5]$, $[4 : 6]$, $[5 : 7]$, and $[8 : 10]$; each maximal clique is a complete subgraph.

Proposition 1. *The connected graphs in $\mathcal{G}_{n,r}$ that have the same structure have corresponding ordered graphs that are identical up to left-right reversal.*

Proof. The result follows from Lemma 1 and Lemma 3, together with the fact that composition of isomorphisms is an isomorphism. □

Similarly to the surjection f introduced earlier, which maps each graph of $\mathcal{G}_{n,r}$ to its corresponding structure in $\mathcal{S}_{n,r}$, we now define the surjective map $\hat{f} : \hat{\mathcal{G}}_{n,r} \to \mathcal{S}_{n,r}$, which associates each ordered graph with its structure.

Proposition 1 enables us to study connected ordered graphs of $\hat{\mathcal{G}}_{n,r}$ as representatives of the connected structures in $\mathcal{S}_{n,r}$, with the understanding that an ordered graph and its image under left-right reversal represent the same structure. Accordingly, for each $S \in \mathcal{S}_{n,r}$ that is connected, it holds that either $|\hat{f}^{-1}(S)| = 1$, when the ordered graph in $\hat{f}^{-1}(S)$ is left-right symmetrical, or $|\hat{f}^{-1}(S)| = 2$, otherwise.

Disconnected structures can be treated similarly. In this case, permutations of connected components of an ordered graph should also be considered identical, in addition to left-right reversal. Thus, in general, the number of ordered graphs that have a common structure S with N_S connected components satisfies

$$1 \leq |\hat{f}^{-1}(S)| \leq 2^{N_S} N_S! \tag{24}$$

4.3 Counting Structures

In this section, we assess the number of distinct structures of the graphs in $\mathcal{G}_{n,r}$, i.e., the cardinality of $\mathcal{S}_{n,r}$, for any number of nodes n, connection range $r > 0$, and probability distribution of the node locations. Specifically, we exploit the connection between structure and ordered graphs made in the previous section to find an upper bound on $|\mathcal{S}_{n,r}|$, which finally gives the bound on the structural entropy presented in Theorem 1.

4.3.1 Total Number of Structures

Based on the surjection f defined in Sec. 4.1, the number of structures can be expressed as

$$|\mathcal{S}_{n,r}| = \sum_{G \in \mathcal{G}_{n,r}} \frac{1}{|f^{-1}(f(G))|}, \tag{25}$$

where the denominator gives the size of the equivalence class the graph $G \in \mathcal{G}_{n,r}$ belongs to. We will however consider the following alternative expression based on the ordered graphs

$$|\mathcal{S}_{n,r}| = \sum_{G \in \hat{\mathcal{G}}_{n,r}} \frac{1}{|\hat{f}^{-1}(\hat{f}(G))|} \leq |\hat{\mathcal{G}}_{n,r}|, \tag{26}$$

because we are able to find an upper limit for the number of ordered graphs. Moreover, the set $\hat{\mathcal{G}}_{n,r}$ is smaller than $\mathcal{G}_{n,r}$. By doing so, we obtain the following upper bound on the number of structures.

Proposition 2. *For any number of nodes n, spatial distribution of the nodes, and connection range $r > 0$, the number of structures in $\mathcal{S}_{n,r}$ is upper bounded by the nth Catalan number, i.e.,*

$$|\mathcal{S}_{n,r}| \leq c_n = \frac{1}{n+1}\binom{2n}{n} < \frac{4^n}{n^{3/2}\sqrt{\pi}}. \tag{27}$$

It is well known that the Catalan number c_n and $\frac{4^n}{n^{3/2}\sqrt{\pi}}$ are asymptotically equivalent, which means that the last inequality in (27) is tight for large n.

Given that the entropy of any distribution is upper bounded by the cardinality of the support, we have $H(\mathcal{S}_{n,r}) < \log_2 |\mathcal{S}_{n,r}|$, such that Proposition 2 immediately gives the following result.

Theorem 1. *For any number n of points, any spatial distribution of the points, and any connection range $r > 0$, the structural entropy is upper bounded by*

$$H(S_{n,r}) < 2n - \frac{3}{2}\log_2 n - \frac{1}{2}\log_2 \pi. \tag{28}$$

4.3.2 Number of Connected Structures

Considering now only connected graphs, let $\hat{\mathcal{G}}^{\mathrm{c}}_{n,r}$ and $\mathcal{S}^{\mathrm{c}}_{n,r}$ be the subsets that include all the connected graphs/structures of $\hat{\mathcal{G}}_{n,r}$ and $\mathcal{S}_{n,r}$, respectively. In the asymptotically connected regime (i.e., for an appropriate connection range r such that $G_{n,r}$ is connected with probability one as n grows large, as defined in Sec. 4.1.2), the graph $G_{n,r}$ is connected with probability one, such that $\hat{\mathcal{G}}^{\mathrm{c}}_{n,r} = \hat{\mathcal{G}}_{n,r}$ and $\mathcal{S}^{\mathrm{c}}_{n,r} = \mathcal{S}_{n,r}$ (the sets $\hat{\mathcal{G}}_{n,r}$ and $\mathcal{S}_{n,r}$ include only connected graphs and structures).

The following bounds on the number of connected structures follow from the fact that $\hat{f}^{-1}(S)$ has at most two elements for connected S:

$$\frac{1}{2}|\hat{\mathcal{G}}^{\mathrm{c}}_{n,r}| \leq |\mathcal{S}^{\mathrm{c}}_{n,r}| \leq |\hat{\mathcal{G}}^{\mathrm{c}}_{n,r}|. \tag{29}$$

By finding an upper limit for $|\hat{\mathcal{G}}^{\mathrm{c}}_{n,r}|$, we obtain the following upper bound for $|\mathcal{S}^{\mathrm{c}}_{n,r}|$.

Proposition 3. *For any number of nodes n, spatial distribution of the nodes, and connection range $r > 0$, the number of connected structures in $\mathcal{S}_{n,r}$ is upper bounded by*

$$|\mathcal{S}^{c}_{n,r}| \leq c_{n-1} = \frac{1}{n}\binom{2n-2}{n-1}. \tag{30}$$

4.4 Probabilistic characterization of maximal cliques

Theorem 2 (Inner maximal cliques)**.** *In an ordered graph with n nodes and connection range $r \in (0,1)$, the probability that $[i,j]$, with $1 < i \leq j < n$, is a maximal clique is*

$$\mathsf{P}([i,j]) = \begin{cases} \sum_{a=j-i}^{n} (-1)^{a-j+i}\binom{n}{a} r^a (1-r)^{n-a}, & \textit{if } r \leq \frac{1}{2} \textit{ and } j \geq i \\ \sum_{a=0}^{j-i-1} (-1)^{j-i-1-a}\binom{n}{a} r^a (1-r)^{n-a}, & \textit{if } r > \frac{1}{2} \textit{ and } j > i. \end{cases} \tag{31}$$

Proof. The conditions for $[i,j]$ to be a maximal are $X_{(j)} - X_{(i)} < r$, $X_{(j)} - X_{(i-1)} > r$ and $X_{(j+1)} - X_{(i)} > r$. We first consider the probability that $[i,j]$ is a maximal clique, given the coordinates $X_{(i)} = x$ and $X_{(j)} = y$, where $y - x < r$, as depicted in Fig. 5. In this case, the distributions of $\left(X_{(1)}, \ldots, X_{(i-1)}\right)$ and $\left(X_{(j+1)}, \ldots, X_{(n)}\right)$ are conditionally independent, each representing the order statistics of $i-1$ and respectively $n-j$ points uniformly distributed on $[0,x]$ and $[y,1]$, respectively.

Figure 5: Illustration of the maximal clique $[i,j]$, given that $X_{(i)} = x$ and $X_{(j)} = y$ with $y - x < r$; the conditions are $X_{(i-1)} < y - r$ and $X_{(j+1)} > x + r$.

Thus, we have

$$
\begin{aligned}
&\mathsf{P}\left([i,j] \mid X_{(i)} = x, X_{(j)} = y\right) \\
&= \mathsf{P}\left(y - X_{(i-1)} > r \mid X_{(i)} = x, X_{(j)} = y\right) \mathsf{P}\left(X_{(j+1)} - x > r \mid X_{(i)} = x, X_{(j)} = y\right) \\
&= \mathsf{P}\left(x - X_{(i-1)} > r - (y-x) \mid X_{(i)} = x, X_{(j)} = y\right) \\
&\quad \times \mathsf{P}\left(X_{(j+1)} - y > r - (y-x) \mid X_{(i)} = x, X_{(j)} = y\right) \\
&= \left(1 - \frac{r-(y-x)}{x}\right)^{i-1} \left(1 - \frac{r-(y-x)}{1-y}\right)^{n-j}
\end{aligned}
$$

where in the last line we have used the expression for the gap distribution [], scaled by the length of the corresponding interval. The joint pdf of $X_{(i)}$and $X_{(j)}$ is []

$$
f_{X_{(i)},X_{(j)}}(x,y) = \frac{n!\, x^{i-1}(y-x)^{j-i-1}y^{n-j}}{(i-1)!\,(j-i-1)!\,(n-j)!}. \tag{32}
$$

Denote $\mathcal{C} = \left\{(x,y) \in [0,1]^2 \mid y - x < r, y - r > 0, x + r < 1\right\}$. We have

$$
\begin{aligned}
\mathsf{P}\left([i,j]\right) &= \iint_{(x,y)\in\mathcal{C}} \mathsf{P}\left([i,j] \mid X_{(i)} = x, X_{(j)} = y\right) f_{X_{(i)},X_{(j)}}(x,y)\,\mathrm{d}x\,\mathrm{d}y \\
&= \iint_{(x,y)\in\mathcal{C}} \frac{n!\,(y-r)^{i-1}(y-x)^{j-i-1}(1-r-x)^{n-j}}{(i-1)!\,(j-i-1)!\,(n-j)!}\,\mathrm{d}x\,\mathrm{d}y
\end{aligned} \tag{33}
$$

Let $u = y - x$. For $r \leq 1/2$, we write (33) as

$$
\mathsf{P}\left([i,j]\right) = n! \int_0^r \mathrm{d}u \int_{r-u}^{1-r} \mathrm{d}x \frac{(x+u-r)^{i-1}u^{j-i-1}(1-r-x)^{n-j}}{(i-1)!\,(j-i-1)!\,(n-j)!} \tag{34}
$$

We make the transformation $x \to (r-u)x + (1-r)(1-x)$ and obtain

$$
\mathsf{P}\left([i,j]\right) = n! \int_0^r \mathrm{d}u \frac{u^{j-i-1}(1-2r+u)^{n-j+i}}{(i-1)!\,(j-i-1)!\,(n-j)!} \int_0^1 x^{i-1}(1-x)^{n-j}\,\mathrm{d}t \tag{35}
$$

$$
= n! \int_0^r \frac{u^{j-i-1}(1-2r+u)^{n-j+i}}{(j-i-1)!\,(n-j+i)!}\,\mathrm{d}u \tag{36}
$$

Next, we expand

$$(1-2r+u)^{n-j+i} = \sum_{a=0}^{n-j+i} (-1)^a \binom{n-j+i}{a} (r-u)^a (1-r)^{n-j+i-a} \tag{37}$$

and make the change $u \to ru$, such that we obtain

$$\begin{aligned}
\mathsf{P}([i,j]) &= \sum_{a=0}^{n-j+i} (-1)^a \frac{n!\, r^{a+j-i}(1-r)^{n-j+i-a}}{(n-j+i-a)!} \int_0^1 \frac{u^{j-i-1}(1-u)^a}{(j-i-1)!\, a!} \mathrm{d}u & (38)\\
&= \sum_{a=0}^{n-j+i} (-1)^a \frac{n!\, r^{a+j-i}(1-r)^{n-j+i-a}}{(n-j+i-a)!\,(j-i+a)!} & (39)\\
&= \sum_{a=j-i}^{n} (-1)^{a-j+i} \binom{n}{a} r^a (1-r)^{n-a} & (40)
\end{aligned}$$

When $r > 1/2$, we have different integration limits in (34), i.e.,

$$\mathsf{P}([i,j]) = n! \int_{2r-1}^{r} \mathrm{d}u \int_{r-u}^{1-r} \mathrm{d}x \frac{(x+u-r)^{i-1} u^{j-i-1} (1-r-x)^{n-j}}{(i-1)!\,(j-i-1)!\,(n-j)!}. \tag{41}$$

Similar calculations yield the result in (31). □

Remarkably, the probability (31) depends on the size of the maximal clique only and not on the absolute values of the indices of the left and right nodes.

Corollary 1 (Recursive identity)**.** *The maximal-clique probabilities satisfy*

$$\mathsf{P}([i,j+1]) = -\mathsf{P}([i,j]) + \binom{n}{j-i} r^{j-i}(1-r)^{n-j+i}, \tag{42}$$

for any $i \geq 2$, $i \leq j \leq n-2$ *and* $r \in (0,1)$.

Proof. The identity can be verified from (31). □

Because of the effect of the margins, the probabilities are different in the case when $i = 1$ and/or $j = n$, as shown next.

Proposition 4 (Boundary effects)**.** *In an ordered graph with* n *nodes and connection range* $r \in (0,1)$*, the probability of observing a maximal clique containing the leftmost and/or rightmost nodes is*

$$\mathsf{P}([1,j]) = \mathsf{P}([n-j+1,n]) = \binom{n}{j-1} r^{j-1}(1-r)^{n-j+1}, \quad 1 \leq j < n, \tag{43}$$

and

$$\mathsf{P}([1,n]) = n\, r^{n-1} - (n-1) r^n. \tag{44}$$

Proof. The proof follows the steps of the previous one, just that there are fewer conditions for the maximal cliques at the boundaries. The conditions for $[1, j]$ to be a maximal are that $X_{(j)} - X_{(1)} < r$ and $X_{(j+1)} - X_{(1)} > r$, which gives

$$\mathsf{P}\left([1, j] \mid X_{(1)} = x, X_{(j)} = y\right) = \left(1 - \frac{r - (y - x)}{1 - y}\right)^{n-j} \tag{45}$$

Multiplying by the joint pdf (32) (with $i = 1$) and marginalizing, we obtain

$$\mathsf{P}\left([1, j]\right) = n! \int_0^r \mathrm{d}u \int_0^{1-r} \mathrm{d}x \frac{u^{j-2}(1 - r - x)^{n-j}}{(j-2)!\,(n-j)!} \tag{46}$$

$$= \binom{n}{j-1} r^{j-1}(1 - r)^{n-j+1}. \tag{47}$$

For $[1, n]$ to be maximal clique, the condition is $X_{(n)} - X_{(1)} < r$, which gives

$$\mathsf{P}\left([1, n]\right) = n! \int_0^r \mathrm{d}u \int_0^{1-u} \mathrm{d}x \frac{u^{n-2}}{(n-2)!} \tag{48}$$

$$= n(n-1) \int_0^r u^{n-2}(1 - u)\,\mathrm{d}u = n\,r^{n-1} - (n-1)r^n. \tag{49}$$

□

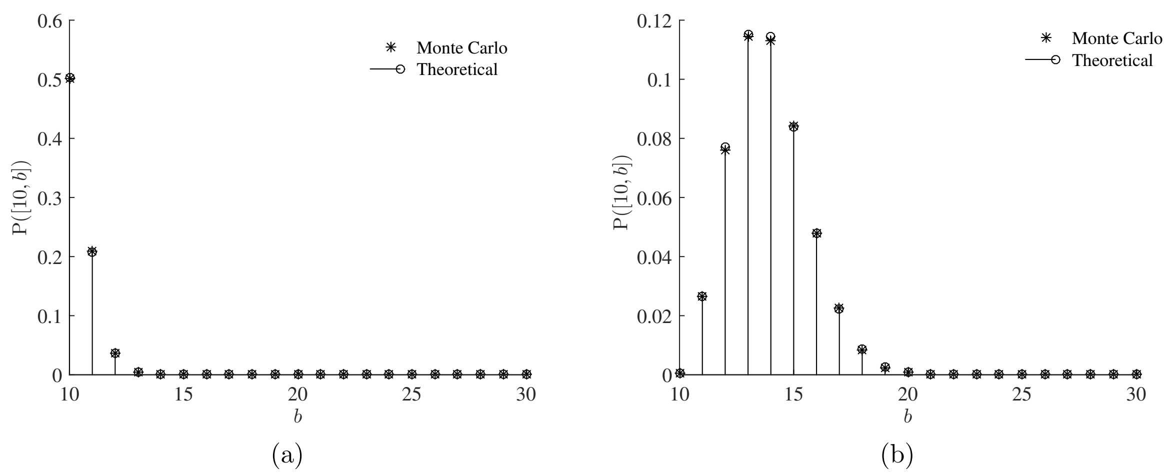

Figure 6: The probability of observing a maximal clique with leftmost node index $a = 10$ vs. the index b of the rightmost node; $n = 30$ nodes and (left) $r = 0.1 \times \frac{\ln n}{n}$ and (right) $r = \frac{\ln n}{n}$.

Corollary 2 (Endpoint of a maximal clique)**.** *When $r \ll 1$, node i ($i \geq 2$) is the left endpoint of a maximal clique with probability* $\mathsf{P}([i, \cdot]) = \frac{1}{2} + \frac{1}{2}(1 - 2r)^n$.

Proof. The probability that node i ($i \geq 2$) is the left endpoint of a maximal clique is $\mathsf{P}([i,\cdot]) = \sum_{j=i}^{n} \mathsf{P}[(i,j)]$. In the identity (42) we sum over all j's and get

$$-\mathsf{P}([i,i]) + 2\sum_{j=i}^{n} \mathsf{P}([i,j]) = \sum_{j=i}^{n} \binom{n}{j-i} r^{j-i}(1-r)^{n-j+i}. \tag{50}$$

Thus,

$$\mathsf{P}([i,\cdot]) = \frac{1}{2}(1-2r)^n + \frac{1}{2}\sum_{a=0}^{n-i} \binom{n}{a} r^a (1-r)^{n-a}. \tag{51}$$

Since $r \ll 1$ and $\mathsf{P}([i,j])$ is exponentially decaying with j, we extend the summation limit up to n, as those extra terms are insignificant, such that

$$\mathsf{P}([i,\cdot]) = \frac{1}{2} + \frac{1}{2}(1-2r)^n. \tag{52}$$

Note that we ignored the margin effect (i.e., when i is close to n), which influences the last $O(nr)$ nodes. □

Corollary 3 (Average number of maximal cliques)**.** *When $r \ll 1$, the expected number of maximal cliques is $\frac{n}{2} + \frac{n}{2}(1-2r)^n$.*

Proof. Let $A_i \in \{0,1\}$ be a variable that indicates whether node i is the left endpoint of a maximal clique. The number of maximal cliques is $K = \sum_i A_i$ and its expected value is thus

$$\mathsf{E}[K] = \sum_i \mathsf{E}[A_i] = \sum_i \mathsf{P}([i,\cdot]) = \frac{n}{2} + \frac{n}{2}(1-2r)^n. \tag{53}$$

□

Corollary 4 (Size of a maximal clique)**.** *When n is large and $\frac{\ln n}{n} < r \ll 1$, the size of a maximal clique starting at node i has size $S = j - i + 1$ with distribution*

$$\mathsf{P}(S_i = s) = 2\sum_{k=s-1}^{n} \binom{n}{k} (-1)^{s-1+k} r^k (1-r)^{n-k} \tag{54}$$

Corollary 5 (Average size of a maximal clique)**.** *When n is large and $\frac{\ln n}{n} < r \ll 1$, the average size of a maximal clique is $nr + \frac{3}{2}$.*

Theorem 3. *The entropy of an ordered graph is bounded by*

$$H(\hat{G}_{n,r}) < \mathsf{E}[K]\, H(S) + H(K) \tag{55}$$

Proof. By viewing the ordered graph as a collection of K maximal cliques $C_1, \ldots, C_K$, where $C_k = [a_k, b_k]$ is a maximal clique with endpoints a_k and b_k, we express

$$H(\hat{G}_{n,r}) = H(C_1, \ldots, C_K \mid K) + H(K) - H(K \mid \hat{G}_{n,r}). \tag{56}$$

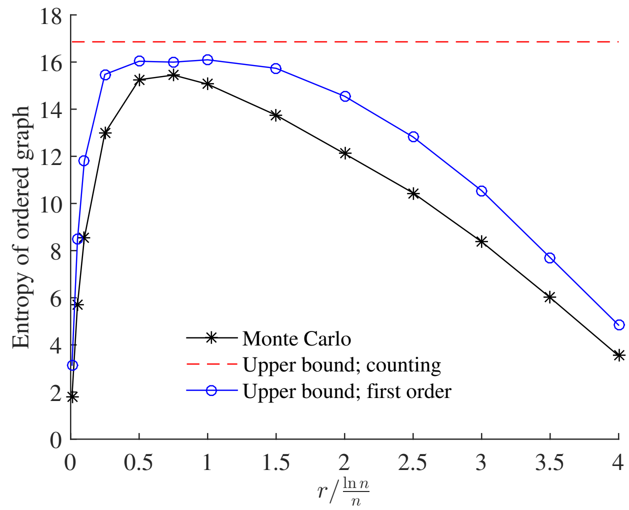

Figure 7: The entropy of random ordered graphs with $n = 12$ nodes vs. the connection range and the upper bound in Theorem 1 and Theorem 3.

Given $\hat{G}_{n,r}$, the maximal-clique decomposition is unique and thus deterministic, which implies $K \mid \hat{G}_{n,r}$ is deterministic, such that $H(K \mid \hat{G}_{n,r}) = 0$. Using the sub-additivity property of joint entropy and then the fact that conditioning reduces entropy, we obtain

$$H(\hat{G}_{n,r}) < \sum_k \mathsf{P}(K = k) \sum_{l=1}^{k} H(C_l \mid K = k) + H(K) \tag{57}$$

$$< \sum_k \mathsf{P}(K = k) \sum_{l=1}^{k} H(C_l) + H(K) \tag{58}$$

Ignoring the effect of the boundary, the maximal cliques have identical marginal distribution, which is actually the distribution of the size, we finally obtain

$$H(\hat{G}_{n,r}) < \mathsf{E}[K]\, H(S) + H(K) \tag{59}$$

□

4.5 Bounds on the Structural Entropy

In the previous section, we characterized the structural complexity of the considered random geometric graph model by studying the number of possible structures in $\mathcal{S}_{n,r}$. That is, all the different structures were equally accounted for without taking into account the distribution of the node locations. However, the distribution of the locations dictates how prevalent the different structures are in the ensemble of graphs.

In this section, we focus on the structural entropy $H(S_{n,r})$ given by (22) and thus explicitly consider the distribution of the locations. We further exploit the connection between the random structure $S_{n,r}$ and the ordered graph $\hat{G}_{n,r}$ to obtain upper and lower bounds on $H(S_{n,r})$.

Lemma 4. *For any number of nodes n, distribution of the locations, and connection range r,*

$$H(S_{n,r}) \leq H(\hat{G}_{n,r}). \tag{60}$$

Proof. According to the map $\hat{f}$ introduced in Sec. 4.2.3,

$$H(S_{n,r}) = H(\hat{f}(\hat{G}_{n,r})) \leq H(\hat{G}_{n,r}), \tag{61}$$

where the inequality holds because the entropy of deterministic functions of a random variable is no larger than the entropy of the variable [20]. □

When considering only the connected graphs, we find an upper bound and a lower bound that determine the structural entropy within one bit.

Lemma 5. *Let $S^{\mathrm{c}}_{n,r}$ denote the random structure conditioned on it being connected. The connected ordered graph $\hat{G}^{c}_{n,r}$ is similarly defined. Then, for any number of nodes n, distribution of the locations, and connection range r,*

$$H(\hat{G}^{c}_{n,r}) - 1 \leq H(S^{c}_{n,r}) \leq H(\hat{G}^{c}_{n,r}). \tag{62}$$

Proof. As established earlier, the preimage $\hat{f}^{-1}(S)$ for connected structures contains only one element when S is left-right symmetrical, and two elements, otherwise. Thus, for any $S \in \mathcal{S}^{\mathrm{c}}_{n,r}$, (20) gives

$$P(S) = \begin{cases} P(\hat{G}), & \text{if } S \text{ is symm.}, S = \hat{f}(\hat{G}), \\ P(\hat{G}_1) + P(\hat{G}_2), & \text{if } S \text{ is asymm.}, \hat{G}_1, \hat{G}_2 \in \hat{f}^{-1}(S). \end{cases} \tag{63}$$

We now write

$$\begin{aligned} H(\hat{G}^{\mathrm{c}}_{n,r}) &= -\sum_{\hat{G} \in \hat{\mathcal{G}}^{\mathrm{c}}_{n,r}} P(\hat{G}) \log_2 P(\hat{G}) \\ &= -\sum_{S \in \mathcal{S}^{\mathrm{c}}_{n,r}} \sum_{\hat{G} \in f^{-1}(S)} P(\hat{G}) \log_2 P(\hat{G}). \end{aligned} \tag{64}$$

For any asymmetrical structure S, define $q_S = \frac{1}{P(S)} \min(P(\hat{G}_1), P(\hat{G}_2))$, where $\{\hat{G}_1, \hat{G}_2\} \equiv \hat{f}^{-1}(S)$. From (63), $q_S < 1$. Using (63), we further express (64) as

$$H(\hat{G}^{\mathrm{c}}_{n,r}) = -\sum_{S \in \mathcal{S}^{\mathrm{c}}_{n,r}} P(S) \log_2 P(S) - \sum_{\substack{S \in \mathcal{S}^{\mathrm{c}}_{n,r} \\ S \text{ asymm.}}} P(S) \left[q_S \log_2 q_S + (1 - q_S) \log_2 (1 - q_S) \right] \tag{65}$$

$$= H(S^{\mathrm{c}}_{n,r}) + \sum_{\substack{S \in \mathcal{S}^{\mathrm{c}}_{n,r} \\ S \text{ asymm.}}} P(S) H_{\mathrm{b}}(q_S), \tag{66}$$

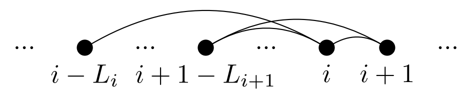

Figure 8: In this illustration, node $i-L_i$ is the leftmost neighbor of node i, i.e., node i is connected by an edge to nodes $i-L_i, i-L_i+1, \dots, i-1$ but not to nodes $1, \dots, i-L_i-1$. Similarly, node $i+1$ has L_{i+1} neighbors to the left.

where $H_{\mathrm{b}}(\cdot)$ is the binary entropy function. Given that H_{b} takes on values smaller than one and $P(S)$ is a probability distribution, the sum term above is between zero and one, and the desired result follows. □

Lemmas 4 and 5 give upper and lower bounds on the structural entropy in terms of the entropy of the ordered graph, $H(\hat{G}_{n,r})$. In the following, we find bounds on $H(\hat{G}_{n,r})$ and study them for i.u.d. point locations and different scaling functions r_n, results which ultimately transfer to bounds for the structural entropy.

4.5.1 Upper bound on $H(\hat{G}_{n,r})$

We represent $\hat{G}_{n,r}$ in terms of the number of leftward neighbors the nodes have, as follows. Recall that the vertices of $\hat{G}_{n,r}$ are indexed in increasing order from left to right, according to the order of their locations. For all $i \in \{2, \dots, n\}$, let L_i be the number of neighbors that node i has to the left, which satisfies $0 \le L_i \le i-1$. It can be inspected (e.g., see Fig. 8) that $0 \le L_{i+1} \le L_i + 1$.

As there is a one-to-one relationship between $\hat{G}_{n,r}$ and $(L_2, \dots, L_n)$, we write

$$\begin{aligned} H(\hat{G}_{n,r}) &= H(L_2, \dots, L_n) \\ &\le H(L_2) + \sum_{i=2}^{n-1} H(L_{i+1} \mid L_i), \end{aligned} \tag{67}$$

where the inequality is obtained by applying the chain rule for entropy and then discarding some of the conditioning variables together with the fact that conditioning reduces the entropy.

To evaluate the upper bound (67), we next obtain the joint probability $\mathsf{P}(L_i = a, L_{i+1} = b)$ for independently and uniformly distributed points, where $0 \le a \le i-1$ and $0 \le b \le a+1$.

When $0 \le a \le i-2$ and $1 \le b \le a+1$: For nodes i and $i+1$ to have L_i and respectively L_{i+1} leftward neighbors, the points must be spatially located as follows: $i-L_i-1$ points in $[0, X_{(i)}-r)$; $L_i-L_{i+1}+1$ points in $[X_{(i)}-r, X_{(i+1)}-r)$; $L_{i+1}-1$ points in $[X_{(i+1)}-r, X_{(i)})$; 0 points in $(X_{(i)}, X_{(i+1)})$; and, finally, $n-i-1$ points in $(X_{(i+1)}, 1]$. When the n points are independent and uniformly distributed on $[0,1]$, we obtain the joint probability by integrating the multinomial distribution over the feasible values of the locations x and y of nodes i and $i+1$, respectively. That is,

$$\mathsf{P}(L_i = a, L_{i+1} = b) = \iint\limits_{\mathcal{D}} \frac{n!\,(x-r)^{i-a-1}(y-x)^{a-b+1}(r+x-y)^{b-1}(1-y)^{n-i-1}}{(i-a-1)!\,(a-b+1)!\,(b-1)!\,(n-i-1)!}\mathrm{d}x\,\mathrm{d}y, \tag{68}$$

where the integration domain is $\mathcal{D} = \{(x,y) \in (0,1)^2 \mid x > r, y > x, y < x + r, y < 1\}$. Assuming $r < 1/2$, we make the change of variable $y - x = u$, such that $0 < u < r$ and $r < x < 1 - u$, and obtain

$$\mathsf{P}(L_i = a, L_{i+1} = b) = \int_0^r \frac{n!\,u^{a-b+1}(r-u)^{b-1}}{(a-b+1)!\,(b-1)!} \int_r^{1-u} \frac{(x-r)^{i-a-1}(1-u-x)^{n-i-1}}{(i-a-1)!\,(n-i-1)!}\mathrm{d}x\,\mathrm{d}u \tag{69}$$

$$= \int_0^r \frac{n!\,u^{a-b+1}(r-u)^{b-1}(1-r-u)^{n-a-1}}{(a-b+1)!\,(b-1)!\,(n-a-1)!}\mathrm{d}u \tag{70}$$

$$= \int_0^1 \frac{n!\,r^{a+1}\,t^{a-b+1}(1-t)^{b-1}(1-r-rt)^{n-a-1}}{(a-b+1)!\,(b-1)!\,(n-a-1)!}\mathrm{d}t, \tag{71}$$

for all $0 \le a \le i-2$ and $1 \le b \le a+1$.

When $0 \le a \le i-2$ and $b = 0$: In this case, assuming $r < 1/2$, the following must hold: $X_{(i+1)} - X_{(i)} > r$; $i - L_i - 1$ points in $[0, X_{(i)} - r)$; L_i points in $[X_{(i)} - r, X_{(i)})$; 0 points in $(X_{(i)}, X_{(i+1)})$; and, finally, $n - i - 1$ points in $(X_{(i+1)}, 1)$. Thus, we write

$$\begin{aligned}\mathsf{P}(L_i = a, L_{i+1} = 0) &= \int_r^{1-r}\int_{x+r}^1 \frac{n!\,(x-r)^{i-a-1}r^a(1-y)^{n-i-1}}{(i-a-1)!\,a!\,(n-i-1)!}\,\mathrm{d}x\,\mathrm{d}y \\ &= \binom{n}{a} r^a(1-2r)^{n-a}.\end{aligned} \tag{72}$$

From the two cases above, we find that

$$\begin{aligned}\mathsf{P}(L_i = a) &= \sum_{b=0}^{a+1} \mathsf{P}(L_i = a, L_{i+1} = b) \\ &= \binom{n}{a} r^a(1-r)^{n-a},\end{aligned} \tag{73}$$

for all $a = 0, \ldots, i-2$, and therefore

$$\begin{aligned}\mathsf{P}(L_i = i-1) &= 1 - \sum_{a=0}^{i-2} \mathsf{P}(L_i = a) \\ &= \sum_{a=i-1}^{n} \binom{n}{a} r^a(1-r)^{n-a}\end{aligned}$$

Thus, L_i has a truncated binomial distribution, as a consequence of the left margin. The margin effect however becomes insignificant for large n and small r, as $i \gg 1$ for

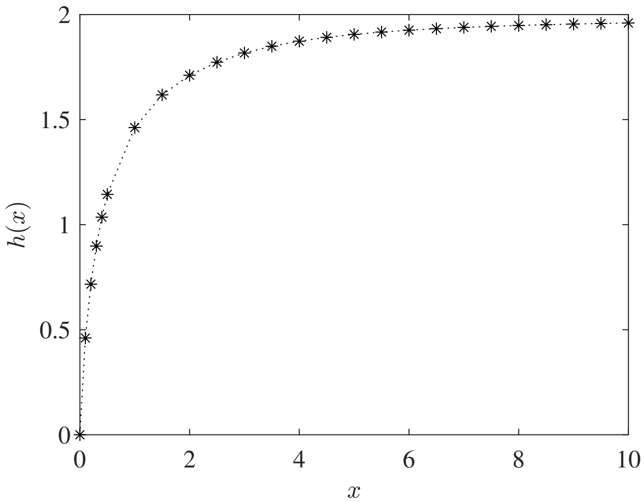

Figure 9: The upper bound on the structural entropy per node for large n and $r_n = \frac{x}{n}$.

most of the terms in (67) and $\mathsf{P}(L_i = i-1)$ corresponds to the tail of the binomial distribution.

Based on (67), we obtain the following result.

Proposition 5. *When the point locations $X_1, \dots, X_n$ are i.u.d. and the connection range scales with n such that $\lim_{n\to\infty} nr_n^2 = 0$, then*

$$\lim_{n\to\infty} \frac{1}{n} H(\hat{G}_{n,r}) = \begin{cases} h(x), & \text{if } \lim_{n\to\infty} nr_n = x < \infty, \\ 2, & \text{if } \lim_{n\to\infty} nr_n = \infty, \end{cases} \tag{74}$$

where the function $h : [0,\infty) \to \mathbb{R}_+$ is given by

$$h(x) = \frac{x\, e^{-x}}{\ln 2} - \sum_{a=0}^{\infty} \frac{x^a e^{-x}}{a!} \sum_{k=0}^{a} \frac{x}{a+1} M(k+1, a+2, -x) \log_2 \left(\frac{x}{a+1} M(k+1, a+2, -x) \right) \tag{75}$$

with M being Kummer's confluent hypergeometric function.

For example, when the connection range function is $r_n = \frac{x}{n}$, for some finite $x > 0$, the structural entropy per node is upper bounded by the curve in Fig. 9 when n grows large; the upper bound asymptotically approaches 2 bits in this case. When $r_n = \frac{x \ln n}{n}$, the upper bound equals 2 bits for any finite x (i.e., it is constant), when $n \to \infty$.

4.5.2 Lower bound on $H(\hat{G}_{n,r})$

Similarly to (67), we represent the ordered graph in terms of the numbers of leftward nodes, such that

$$\begin{aligned}
H(\hat{G}_{n,r}) &= H(L_2, \dots, L_n) \\
&= H(L_2) + \sum_{i=2}^{n-1} H(L_{i+1} \mid L_1, \dots, L_i) && (76) \\
&\geq H(L_2) + \sum_{i=2}^{n-1} H(L_{i+1} \mid L_1, \dots, L_i, X_1, \dots, X_i) && (77) \\
&= H(L_2) + \sum_{i=2}^{n-1} H(L_{i+1} \mid X_1, \dots, X_i) && (78)
\end{aligned}$$

By determining the conditional distribution of the number of neighbors to the left, given the locations of those neighbors, we then obtain the following result.

Proposition 6. *When the point locations $X_1, \dots, X_n$ are i.u.d. and the connection range r_n is proportional to $\frac{\ln n}{n}$ (i.e., in the asymptotically connected regime)*

$$\lim_{n\to\infty} \frac{1}{n} H(\hat{G}_{n,r}) \geq \frac{5}{4\ln 2} \approx 1.8. \tag{79}$$

4.6 Compression Scheme

Propositions 5 and 6 tell us that, in the asymptotically connected regime (which is relevant for connected yet power efficient wireless networks), we need somewhere between 1.8 and 2 bits per node to represent large 1D networks. Based on the graph representation in terms of its maximal cliques, we devise the following compression scheme that requires 2 bits per node, i.e., it asymptotically achieves the 2 bits per node upper bound obtained previously.

Let $\mathbf{a}$ and $\mathbf{b}$ be two binary strings of length n. The ones in a and b indicate the end-nodes of the maximal cliques. Specifically, the position of the kth one in $\mathbf{a}$ ($\mathbf{b}$) gives the index of the leftmost (rightmost) node of the kth maximal clique. For example, the graph in Fig. 4 is encoded as $\mathbf{a} = 1101100100$ and $\mathbf{b} = 0001111001$.

5 Conclusions

The primary contributions generated within this project are summarized in the entropy bounds for directed and asymptotically connected graphs given in this report. Crucially, the bounds on the structural entropy in one dimensional networks point to a very simple compression scheme with encoding that is linear in the number of nodes. For (quasi-) one-dimensional networks, such as convoys, one may use this encoding scheme to propagate information about the topological structure of the network quickly and efficiently to all devices in a time period that grows linearly with the size of the network.

While not fully explored in the project, the secondary contributions associated with physical layer performance and the timeliness of network state information point to further refinements of the theory presented herein. Such refinements could consider the effect of node mobility and transmission delays on TD performance.

Of fundamental interest is the consideration of higher-dimensional networks – especially 2-D – and the entropy scaling properties of such systems. Additionally, it should be noted that all results reported herein were obtained under the assumption that lossless compression of the network state information is required. A natural modification to this assumption would be to treat the lossy case, whereby a rate-distortion theory of graph compression would need to be developed. This is a rich area of research worthy of further investigation.

References

[1] Abbe, E. (2016). 54th IEEE Allerton Conf. Commun., Cont., Comp.

[2] An, B., & Papavassiliou, S. (2002). IEEE Commun. Lett., 6(8).

[3] Asadi, A. R., Abbe, E., & Verdú, S. (2017). IEEE Int. Symp. Info. Theory (ISIT).

[4] Badiu, M.-A., & Coon, J. P. (2018). IEEE Int. Symp. Info. Theory (ISIT).

[5] Badiu, M. A., & Coon, J. P. (2019). IEEE Wirel. Commun. Lett., 9(2).

[6] Badiu, M.-A., & Coon, J. P. (2021). IEEE Trans. Info. Theory, in preparation.

[7] Barthélemy, M. (2011). Phys. Rep., 499(1-3).

[8] Bollobás, B., & Béla, B. (2001). Random graphs (No. 73). CUP.

[9] Boushaba, M., et al. (2017). J. Netw. and Comp. App., 93.

[10] Chen, G., Coon, J. P., & Di Renzo, M. (2017). IEEE Trans. Info. Forens. & Secur., 12(5).

[11] Choi, Y., & Szpankowski, W. (2009). IEEE Int. Symp. Info. Theory (ISIT).

[12] Choi, Y., & Szpankowski, W. (2012). IEEE Trans. Info. Theory, 58(2).

[13] Cika, A., Coon, J. P., & Kim, S. (2017). 15th Int. Symp. Mod. Opt. Mob., Ad Hoc, Wirel. Netw. (WiOpt). **Invited paper**.

[14] Cika, A., Badiu, M. A., & Coon, J. P. (2020). IEEE Trans. Mob. Comp., to appear.

[15] Coon, J. P. (2016). IEEE Global Commun. Conf. (GLOBECOM).

[16] Coon, J. P., & Smith, P. J. (2017). IEEE Int. Conf. Commun. (ICC).

[17] Coon, J. P., Dettmann, C. P., & Georgiou, O. (2018). Phys. Rev. E, 97(4), 042319.

[18] Coon, J. P., Badiu, M.-A.,& Gündüz, D. (2018). 16th Int. Symp. Mod. Opt. Mob., Ad Hoc, and Wirel. Netw. (WiOpt). **Invited paper**.

[19] Cooper, C., Radzik, T., & Siantos, Y. (2014). Soc. Netw. Anal. Min., 4(1).

[20] Cover, T. M., & Thomas, J. A. (2012). Elements of Information Theory. John Wiley & Sons.

[21] Dall, J., & Christensen, M. (2002). Phys. Rev. E, 66(1), 016121.

[22] Desai, M., & Manjunath, D. (2002). IEEE Commun. Lett., 6(10).

[23] Elkies, N. D. Online available: `https://mathoverflow.net/q/133752`.

[24] Elmasry, G. F. (2010). IEEE Commun. Mag., 48(10).

[25] Ghasemi, A., & Nader-Esfahani, S. (2006). IEEE Commun. Lett., 10(4).

[26] Guo, J. L., et al. (2011). Trans. Tech. Pub. Adv. Eng. Forum, vol. 1, pp. 288-292.

[27] Gupta, P., & Kumar, P. R. (1999). Stoch. Anal., Cont., Opt. App., Birkhäuser, Boston, MA.

[28] Haenggi, M., et al. (2009). IEEE J. Sel. Areas in Commun., 27(7).

[29] Halu, A., et al. (2014). Phys. Rev. E, 89(1), 012806.

[30] Han, G., & Makowski, A. M. (2007). IEEE Commun. Lett,, 11(1).

[31] Han, G., & Makowski, A. M. (2009). IEEE Trans. Info. Theory, 55(12).

[32] Høst-Madsen, A., & Zhang, J. (2018). IEEE Int. Symp. Info. Theory (ISIT).

[33] IMT Vision – Framework and overall objectives of the future development of IMT for 2020 and beyond. (2015) Rec. ITU-R M.2083-0.

[34] Jawhar, et al. (2017). IEEE Conf. Comp. Commun. Workshops (INFOCOM WKSHPS).

[35] Kieffer, J. C., Yang, E. H., & Szpankowski, W. (2009). IEEE Int. Symp. Info. Theory (ISIT).

[36] Liu, Y., et al. (2016). IEEE Trans. Intel. Transp. Sys., 17(8).

[37] Lu, J. L., et al. (2008). IEEE Commun. Lett., 12(3).

[38] Łuczak, T., Magner, A., & Szpankowski, W. (2019). IEEE Int. Symp. Info. Theory (ISIT).

[39] Naor, M. (1990). Disc. App. Math., 28(3).

[40] Newman, M. (2010). Networks: An Introduction. OUP.

[41] Penrose, M. (2003). Random geometric graphs (Vol. 5). OUP.

[42] Timo, R., et al. (2005). 6th IEEE Aus. Commun. Theory Workshop.

[43] Turán, G. (1984). Disc. App. Math., 8(3).

[44] Wang, Z., Badiu, M. A., & Coon, J. P. (2020). IEEE Global Commun. Conf. (GLOBECOM).

[45] Wang, Z., Badiu, M. A., & Coon, J. P. (2021). arXiv preprint arXiv:2102.08841.

[46] Yu, T., Wang, X., Jin, J., & McIsaac, K. (2018). IEEE Trans. Indust. Inform., 14(5).

[47] Zayani, M. H. (2012). Link prediction in dynamic and human-centered mobile wireless networks (Doctoral dissertation, Institut National des Télécommunications).