

# Entropy of Spatial Network Ensembles

Justin P. Coon,<sup>1</sup> Carl P. Dettmann,<sup>2</sup> and Orestis Georgiou<sup>2,3</sup>

<sup>1</sup>*Department of Engineering Science, University of Oxford, Parks Road, Oxford OX1 3PJ, UK*

<sup>2</sup>*School of Mathematics, University of Bristol, University Walk, Bristol BS8 1TW, UK*

<sup>3</sup>*Ultrahaptics, The West Wing, Glass Wharf, Bristol BS2 0EL, UK*

(Dated: April 5, 2018)

We analyze complexity in spatial network ensembles through the lens of graph entropy. Mathematically, we model a spatial network as a *soft* random geometric graph, i.e., a graph with two sources of randomness, namely nodes located randomly in space and links formed independently between pairs of nodes with probability given by a specified function (the “pair connection function”) of their mutual distance. We consider the general case where randomness arises in node positions as well as pairwise connections (i.e., for a given pair distance, the corresponding edge state is a random variable). Classical random geometric graph and exponential graph models can be recovered in certain limits. We derive a simple bound for the entropy of a spatial network ensemble and calculate the conditional entropy of an ensemble given the node location distribution for hard and soft (probabilistic) pair connection functions. Under this formalism, we derive the connection function that yields maximum entropy under general constraints. Finally, we apply our analytical framework to study two practical examples: ad hoc wireless networks and the US flight network. Through the study of these examples, we illustrate that both exhibit properties that are indicative of nearly maximally entropic ensembles.

## I. INTRODUCTION

The topological structure of networks has been studied for many years through the lens of graph entropy [1]. This formalism, which is deeply rooted in statistical physics and information theory, allows one to quantitatively characterize the complexity or inherent information content of systems that can be described by a graphical model [2–7]. Applications of entropy-based methods to the study of networked systems are abundant and include problems related to molecular structure classification [8], social networks [9, 10], data compression [11], quantum entanglement [12, 13], and topological uncertainty in communication networks [14–16]. Of course, this diverse range of applications has led to the definition of numerous entropic measures [17], and the study and unification of such measures continues to garner interest [18].

To date, research on graph entropy has been mostly focused on systems that do not depend on an underlying spatial embedding, or for which this embedding has been abstracted or ignored. In this context, numerous measures of structural entropy have been developed that allow one to quantify the entropy of a single graph [10, 17]. This is typically done by identifying some characteristic of the network that is of interest, and defining a probability distribution on this characteristic by using a frequency interpretation of occurring events. An alternative approach to studying graph entropy is to consider a measure on the entire ensemble, not just a single graph. In this case, a probability distribution is defined on the ensemble, and various measures of entropy (e.g., Gibbs entropy, Shannon entropy, von Neumann entropy, Rényi entropy) can be calculated for the ensemble [6]. The canonical example that one might begin with is the Erdős-Rényi (ER) model, for which each edge exists in-

dependently with probability  $p$ . It is straightforward to observe that the probability of a particular graph  $G$  containing  $n$  nodes and  $k$  edges is just

$$P(G) = p^k(1-p)^{c_n-k} \quad (1)$$

where  $c_n = n(n-1)/2$ . Hence, a probability distribution is well defined on the graph ensemble  $\mathcal{G}$  for  $k = 0, \dots, c_n$  since  $P(G) \geq 0$  for all  $G \in \mathcal{G}$  and  $\sum_{G \in \mathcal{G}} P(G) = 1$ . Suppose we wish to calculate the Shannon entropy of the ensemble  $\mathcal{G}$ , which admits the expression

$$H(\mathcal{G}) = - \sum_{G \in \mathcal{G}} P(G) \ln P(G). \quad (2)$$

By enumerating the different combinations of edges that form the graphs in  $\mathcal{G}$ , it follows that the Shannon entropy of the ER ensemble can be written as

$$H(\mathcal{G}) = -c_n(p \ln p + (1-p) \ln(1-p)). \quad (3)$$

In contrast to the spatially independent models discussed above, geography has been shown to play an important role in many engineered, physical and social networks [19]. For example, an empirical study conducted in Belgium and reported in [20] demonstrated that social connectivity between two individuals decays like  $r^{-2}$  over distances from about 1 km to 100 km. Spatial network models have also been employed to study the spread of epidemics [21], mobile phone viruses [22], and to analyze connectivity in wireless communication networks [23]. Local structural observables that describe these networks, such as the clustering coefficient and the degree distribution, are well understood for simple models. However, while some work has been done to characterize the entropy of spatial networks [24], relatively little research has focused on the entropy of spatial network ensembles. The notable exception to this lack of

available results in the literature is [25] and the works pertaining to wireless communication networks by the first author [14–16].

In this contribution, we characterize the Shannon entropy of spatial network ensembles by employing a *soft* random geometric graph (RGG) framework [26]. Randomness in this context is derived both from the node positions in space as well as from inherent uncertainties associated with the edge states. We begin by providing an upper bound on spatial network ensemble entropy, which is derived from a simple result from information theory. We then discuss entropy in the *conditional* sense, whereby we assume some prior knowledge of the underlying node locations and average the entropy over the spatial distribution. In many physical systems, such knowledge is available; hence, this measure of conditional entropy lends itself nicely to practical interpretations.

Our approach to characterizing network entropy takes a different route than the work reported in [25]. Specifically, [25] considered the case where node positions are static and known *a priori*, which enables the study of systems with link specific constraints and even ensembles of multiplexes. In contrast, we focus on extracting information about systems for which node locations are, themselves, a source of randomness by averaging over all possible spatial node configurations. This is a conventional information theoretic approach to understanding how one random variable yields information about another; hence, our work can be viewed as a step towards a larger information theoretic framework for spatial networks. Indeed, using our formalism facilitates the calculation of the mutual information (more generally, the relative entropy) between the network topology and the underlying point process that statistically characterizes node locations [27]. Under this formalism, we derive the connection function that yields maximum entropy under general constraints. We then apply our analytical methods to study two practical examples: ad hoc wireless networks and the conterminous US flight network. We illustrate the intriguing result that both systems exhibit properties that are indicative of nearly maximally entropic ensembles.

## II. ENTROPY

Consider a set  $\mathcal{V}$  of  $n$  nodes embedded in a space  $\mathcal{K} \subset \mathbb{R}^d$ , which has volume  $K = \text{vol}(\mathcal{K})$ . The locations of the nodes are denoted by  $\mathbf{r}_1, \dots, \mathbf{r}_n$ , and these locations form a point process in  $\mathcal{K}$ . In this work, we will consider a simple, uniform point process, which implies the spatial distribution of nodes can be described by a constant, finite intensity of  $\rho = n/K$  nodes per unit volume. In what follows, it will suffice to assume that  $n$  is fixed. Hence, the triple  $(\mathcal{V}, \mathcal{K}, \rho)$  describes a binomial point process. Similar results to those disclosed herein follow when  $n$  is Poisson distributed.

An (undirected) edge exists between nodes  $i$  and  $j$  in  $\mathcal{V}$

with probability  $p(r_{i,j})$ , where  $r_{i,j}$  denotes the distance between points  $\mathbf{r}_i$  and  $\mathbf{r}_j$ . In this work, we assume the spatial embedding is Euclidean, i.e.,  $r_{i,j} = \|\mathbf{r}_i - \mathbf{r}_j\|$ . In the case where  $p(r_{i,j})$  is an indicator function that is one for  $r < r_0$  and zero otherwise, with  $r_0$  denoting the maximum connection range, we recover the hard disk model used in the classical RGG formalism [28]. For other connection functions  $p$  – e.g., monotonically decreasing functions in the distance argument – we have the so-called *soft* RGG model [26, 29]. As discussed above, we denote the set of all graphs by  $\mathcal{G}$ . For a particular graph  $G \in \mathcal{G}$ , the set of edges is signified by  $\mathcal{E}_G$ . Note that these sets are defined without reference to a particular underlying spatial embedding, or even the spatial distribution of the vertices.

We are interested in quantifying the Shannon entropy of the ensemble  $\mathcal{G}$ . Each graph  $G \in \mathcal{G}$  occurs with probability  $P(G)$ , which depends upon the spatial distribution of the nodes and the pair connection function  $p$ . The definition of the Shannon entropy of  $\mathcal{G}$  was given in eq. (2). In contrast to the example of ER graph ensembles, the existence of edge  $(i, j)$  is a function of the pair distance  $r_{i,j}$ . Viewing a graph as a random variable  $G$  with support  $\mathcal{G}$ , we can easily deduce that the distribution of  $G$  is equivalent to the distribution of the edge set only, since the spatial embedding of the vertices is captured in the edge probabilities. Let  $X_{i,j}$  denote a Bernoulli random variable that models the existence of edge  $(i, j)$ . It follows that we can write the entropy of the network ensemble as the joint entropy of the sequence  $\{X_{i,j}\}$ , i.e.,

$$H(\mathcal{G}) = H(X_{1,2}, X_{1,3}, \dots, X_{n-1,n}). \quad (4)$$

We can now invoke the well known independence bound on joint Shannon entropy to obtain the inequality

$$H(\mathcal{G}) \leq \sum_{i < j} H(X_{i,j}) \quad (5)$$

where equality holds if all  $\{X_{i,j}\}$  are independent. A more thorough treatment of this bound in the context of spatial network entropy is provided in [14].

The random variable  $X_{i,j}$  is physically related to nodes  $i$  and  $j$ , but it should be stressed that  $P(X_{i,j} = 1) \neq p(r_{i,j})$ , since pair distance information is marginalized in eq. (5). More accurately, we can write

$$P(X_{i,j} = 1) = \int_0^D f(r_{i,j}) p(r_{i,j}) dr_{i,j} \quad (6)$$

which is just the probability that edge  $(i, j)$  exists averaged over all pair distances, where  $D = \sup_{\mathbf{r}_i, \mathbf{r}_j \in \mathcal{K}} \|\mathbf{r}_i - \mathbf{r}_j\|$  is the diameter of the domain  $\mathcal{K}$  and  $f(r_{i,j})$  is the pair distance probability density corresponding to nodes  $i$  and  $j$ . By noting that the pair distance density and the pair connection function are, respectively, identical for all  $i \neq j$ , we can simply let  $\bar{p} := P(X_{i,j} = 1)$  and write

$$H(\mathcal{G}) \leq c_n H_2(\bar{p}) \quad (7)$$

where

$$H_2(\bar{p}) = -\bar{p} \ln \bar{p} - (1 - \bar{p}) \ln(1 - \bar{p}) \quad (8)$$

is the binary entropy function. So by assuming pair distances are independent in the soft RGG case, we obtain an ER-like result, but where the pair connection probability is averaged over the pair distance distribution. Clearly, when  $\bar{p} = 1/2$ , the bound is maximized.

### III. CONDITIONAL ENTROPY

To enable us to study the effect of a particular embedding on the entropy of a soft RGG ensemble, we turn to the information theoretic notion of *conditional entropy*. In the context of our problem, the conditional entropy of the graph ensemble given the distribution of vertex locations is defined as

$$H(\mathcal{G}|\mathcal{R}) = \langle H(\mathcal{G}|\mathbf{r}_1, \dots, \mathbf{r}_n) \rangle \quad (9)$$

where the notation  $\langle \cdot \rangle$  denotes the average of a function over the vertex positions, which for uniformly distributed vertices in  $\mathcal{K}$  is given by

$$\langle F(\mathbf{r}_1, \dots, \mathbf{r}_n) \rangle = \frac{1}{K^n} \int_{\mathcal{K}^n} F(\mathbf{r}_1, \dots, \mathbf{r}_n) d\mathbf{r}_1 \cdots d\mathbf{r}_n. \quad (10)$$

This quantity can be more conveniently expressed in terms of an average over the pair distances

$$H(\mathcal{G}|\mathcal{R}) = \langle H(\mathcal{G}|r_{1,2}, r_{1,3}, \dots, r_{n-1,n}) \rangle \quad (11)$$

where the average is defined in the appropriate manner with respect to the joint pair distance density  $f(r_{1,2}, r_{1,3}, \dots, r_{n-1,n})$ . By using a similar argument to that employed in eq. (5), it is possible to show that

$$H(\mathcal{G}|r_{1,2}, \dots, r_{n-1,n}) = \sum_{i < j} H(X_{i,j}|r_{i,j}) \quad (12)$$

where equality follows since, in this case, the edge states  $\{X_{i,j}\}$  are independent conditioned on the pair distances  $\{r_{i,j}\}$ . Averaging the right-hand side of eq. (12) over the density  $f(r_{1,2}, r_{1,3}, \dots, r_{n-1,n})$  leads naturally to

$$H(\mathcal{G}|\mathcal{R}) = c_n \int_0^D f(r) H_2(p(r)) dr. \quad (13)$$

A fundamental result of information theory states that conditioning reduces uncertainty [27]. Combining this notion with eq. (7) manifests in the relation

$$H(\mathcal{G}|\mathcal{R}) \leq H(\mathcal{G}) \leq c_n H_2(\bar{p}). \quad (14)$$

In fact, the lower and upper bounds follow directly from the concavity of Shannon entropy and Jensen's inequality, i.e., the left-hand side is the average of the entropy, whereas the right-hand side is the entropy of the average. The tightness of each bound depends upon the pair

connection function  $p(r)$ . In practical networks, such as social networks or communication networks, the underlying system parameters dictate the form of this function. In general, very soft pair connection functions lead to relatively tight bounds in this context since there is little dependence upon the spatial embedding. On the other hand, if  $p(r) = 1$  for  $0 \leq r < r_0$  and  $p(r) = 0$  otherwise, then we recover the classical RGG formalism and  $H(\mathcal{G}|\mathcal{R}) = 0$ , but the upper bound on  $H(\mathcal{G})$  is still  $\mathcal{O}(n^2)$ .

### IV. ENTROPY MAXIMIZING CONNECTION FUNCTIONS

Understanding how different parameters affect uncertainty in spatial networks is of great importance. Hence, it is of interest to determine the function  $p(r)$  that maximizes the conditional entropy of the network ensemble. Without further information, we can immediately observe that the entropy maximizing function (in the conditional sense) is just  $p(r) = 1/2$ , in which case  $H(\mathcal{G}|\mathcal{R}) = H(\mathcal{G}) = c_n \ln 2$ . This is rather uninteresting, however, since we are back to the ER graph ensemble in which the spatial embedding is irrelevant. To make progress beyond this benchmark, we note that practical spatial networks typically exhibit certain properties, which we may wish to incorporate into the maximization task through constraints. For example, the mean degree of a geometric graph ensemble can be written as

$$\bar{\delta} = (n-1) \int_0^D f(r) p(r) dr. \quad (15)$$

Considering a network with a given mean degree, we can formulate a constrained variational problem, which has Lagrangian

$$\mathcal{L}(p) = c_n f(r) H_2(p(r)) - \mu(n-1) f(r) p(r) \quad (16)$$

with  $\mu$  denoting the required multiplier for the mean degree constraint. Setting  $\mathcal{L}_p = 0$  and using the constraint to solve for the multiplier yields the stationary function  $p = \bar{\delta}/(n-1)$  for all  $r \in [0, D]$ , and thus we again see that the spatial embedding of the graph does not affect the maximum entropy probability. Hence, by using eqs. (6), (7), and (13), we arrive at the entropy relation

$$H(\mathcal{G}|\mathcal{R}) = H(\mathcal{G}) \sim \frac{\bar{\delta}}{2} n \ln n \quad (17)$$

for large  $n$ . Although the previous two examples are independent of the spatial embedding and are, in some sense, uninteresting in the context of the present discussion, it is worth noting the difference in entropies for the dense and the sparse cases:  $H(\mathcal{G}) = \mathcal{O}(n^2)$  for the dense network with  $p = 1/2$  and  $H(\mathcal{G}) = \mathcal{O}(n \ln n)$  for the sparse network with  $p = \mathcal{O}(n^{-1})$ .

Many geometric networks observed in our physical world are characterized by more than just the mean degree. Other local features often play an important role

in the overall structure of the network. For example, in wireless communication networks, pairwise connections often have a well-defined statistical make-up, which is governed by the modulation/demodulation techniques used at each node and the arrangement of scatterers in the environment [30]. From a more general perspective, we may consider a graph ensemble where the pair connection function obeys a set of constraints

$$\int_0^D \theta_\ell(r) p(r) dr = k_\ell, \quad \ell = 1, \dots, L \quad (18)$$

where  $\{\theta_\ell\}$  and  $\{k_\ell\}$  are independent of  $n$ . In such a case, we might seek the maximizing function  $p(r)$  subject to these constraints. We now have a constrained variational problem comprised of eq. (13) and the constraints given in eq. (18). Solving the associated Euler-Lagrange equation yields the stationary function

$$p(r) = \frac{1}{e^{\psi(r)} + 1} \quad (19)$$

where

$$\psi(r) = \frac{1}{c_n f(r)} \sum_{\ell=1}^L \lambda_\ell \theta_\ell(r) \quad (20)$$

and  $\{\lambda_\ell\}$  are the undetermined multipliers. This solution points to explicit dependence on the spatial embedding through the pair distance density  $f$  and the constraints captured by  $\{\theta_\ell\}$ . Note that if  $\theta_\ell(r) = g_\ell(r)f(r)$  for some function  $g_\ell$  and  $\ell = 1, \dots, L$ , the maximizing pair connection function is independent of  $f$ . Of course, the multipliers  $\{\lambda_\ell\}$  still depend upon the geometric properties of the network in this case. It will have occurred to many readers that eq. (19) resembles the classical entropy maximizing distribution pertaining to quantum state occupancy in systems of noninteracting fermions. Park and Newman pointed out a similar relation for exponential random graphs in [31]. Here, we have arrived at a similar result for spatial network ensembles in bounded domains.

For finite network domains,  $f$  is independent of  $n$ . Assuming for large  $n$  that  $\psi = \mathcal{O}(n^{-\epsilon})$  for  $\epsilon > 0$ , we have  $p(r) = 1/2 + \mathcal{O}(n^{-\epsilon})$ . But this yields a degenerate outcome in which the constraints are independent of the pair connection function. Hence, we must have that  $\psi(r) = \mathcal{O}(1)$ , i.e., the multipliers scale like  $n^2$ . The entropy maximizing connection function yields a dense network in the large  $n$  limit, i.e.,  $H(\mathcal{G}) = \mathcal{O}(n^2)$ . The mean degree of this network is asymptotically

$$\bar{\delta} \sim n \int_0^D \frac{f(r)}{e^{\psi(r)} + 1} dr. \quad (21)$$

In finite networks, it is meaningful to consider adding a mean degree constraint to the set of constraints defined in eq. (18). In this case, the entropy maximizing connection function is

$$p(r) = \frac{1}{e^{\psi(r)+2\mu/n} + 1}. \quad (22)$$

Applying a mean degree constraint induces sparsity if the constraint is kept fixed for any  $n$ . However, positive constraints cannot be satisfied in the large  $n$  limit for  $p \sim 1/n$  (see eq. (18)). Rigorously, this follows from a fundamental result from calculus, which states that for all smooth functions  $\xi(r)$  on the interval  $[0, D]$ ,  $\int_0^D \xi(r) p(r) dr = 0$  if and only if  $p(r) = 0$  on the interval. Hence, stationary functions do not exist for some constraint conditions.

## V. EXAMPLES

Various examples of real-world spatial networks exist, and each is governed by particular connection rules and design constraints. Here, we explore two such examples.

### A. Wireless Communication Networks

The first example we will explore is related to ad hoc deployments of electronic devices that have the capability to communicate wirelessly. In this example, a pairwise connection between two nodes separated by a distance  $r$  forms with probability [23]

$$p(r) = e^{-(r/r_0)^\eta}. \quad (23)$$

The parameter  $r_0$  denotes the typical connection range; it depends on physical quantities such as the wavelength of the transmission and antenna gains at the transmitter and receiver. The parameter  $\eta > 0$ , known as the *path loss exponent*, is typically an experimentally determined number that indicates how quickly a transmission is attenuated as it propagates through the wireless medium. Eq. (23) follows from information theoretic arguments, but possesses two convenient properties: (1) it is intuitive in that it exhibits a connection rule that captures a monotonically decreasing probability of pairwise connectivity as the distance between nodes increases; and (2) the parameter  $\eta$  serves to define the degree of pairwise uncertainty experienced in the network, with  $\eta \rightarrow \infty$  signifying a hard connection rule whereby two nodes are connected with probability one if they are separated by less than  $r_0$  and are not connected otherwise, i.e.,  $\eta \rightarrow \infty$  signifies the classical RGG formalism.

Fig. 1 illustrates the effect that certainty in each pairwise connection has on the conditional entropy. This figure relates to a network residing in a sphere of unit radius in three dimensions. As we increase  $\eta$ , the pairwise uncertainty decreases and the conditional entropy of the ensemble correspondingly decreases. Clearly, as  $\eta \rightarrow \infty$ , the conditional entropy tends to zero, although the unconditional entropy  $H(\mathcal{G})$  remains positive.

In wireless networks, it is natural to consider a system in which both the average pair connection function and the correlation between the connection function and

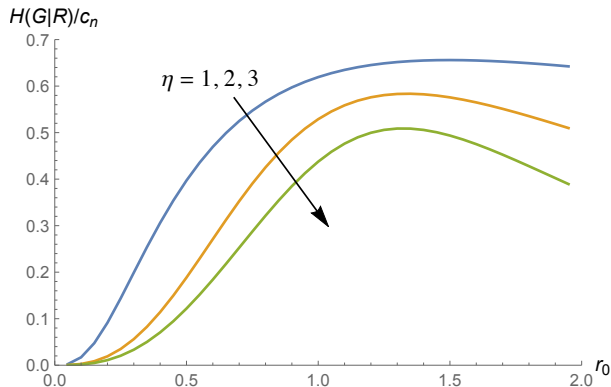


FIG. 1. Conditional entropy normalized with respect to the total number of possible edges plotted against the typical connection range in a wireless communication network. The network domain is a sphere of unit radius in three dimensions;  $f(r) = 3r^2(1 - 3r/4 + r^3/16)$ . Three different exponents  $\eta = 1, 2, 3$  were used to show how an increase in edge certainty (increasing  $\eta$ ) reduces entropy.

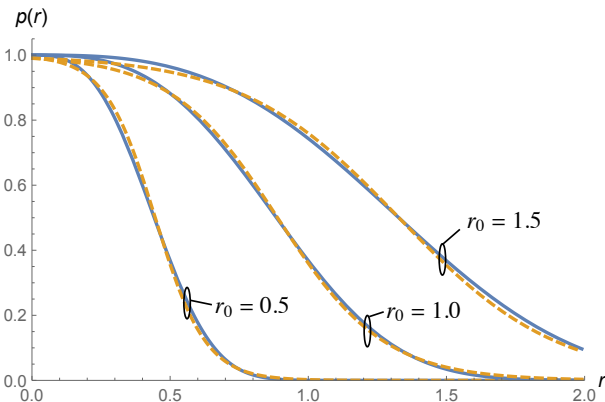


FIG. 2. Test function defined in eq. (23) (solid lines) and maximum conditional entropy function (dashed lines) plotted as a function of node separation for a network of  $n = 10$  nodes. For the test function,  $\eta = 3$  and  $r_0 = 0.5, 1.0, 1.5$ . Average connection probability and correlation constraints were imposed for the maximization.

distance are constrained, i.e.,

$$k_\ell = \int_0^D r^{\ell-1} p(r) f(r) dr, \quad \ell = 1, 2. \quad (24)$$

The correlation constraint somewhat captures the monotonically decreasing nature of the likelihood of pairwise connectivity as node separation increases, whereas the average connection probability constraint gives an indication of average performance in the network. Let us consider the exponential test function given in eq. (23) with  $\eta = 3$ . For each value of  $r_0$ , one can easily calculate the integrals in eq. (24). Here, we consider the case where the network domain is a sphere of unit radius in three dimensions;  $f(r) = 3r^2(1 - 3r/4 + r^3/16)$ . This sets the constraints, which can be used to solve for the unde-

termined multipliers that define the entropy maximizing function (eq. (19)). In this way, we can compare the test function and the maximum entropy connection probability as a function of distance  $r$  (Fig. 2). Our analysis provides an interesting approach to studying complexity in wireless networks by allowing one to compare practical connection functions to the maximum entropy curves. In the example discussed here, it is clear that the practical test function is extremely close to the entropy maximizing function. In fact, the test functions depicted in Fig. 2 yield conditional entropies of over 99.4% of the maximum in each case. This result points to the intriguing conclusion that wireless networks adhering to the model studied here are nearly maximally complex (under the specified constraints).

## B. Flights in the United States

Another example with practical significance can be found in the analysis of airline routes in the United States. For this case, we have analyzed all direct routes between primary airports – i.e., airports listed by the Federal Aviation Administration (FAA) as those that provide scheduled passenger services and which process over 10,000 boardings per year – in the conterminous states. This analysis focused on 336 airports and 2,422 unique routes. Data was obtained from [32, 33]. A histogram of the connection probability, viewed as a function of distance between cities, was constructed, and a polynomial fit was produced from this data. An 8th order polynomial was chosen to ensure an adequate goodness of fit was achieved (sum of squared errors:  $2.36 \times 10^{-3}$ ) while capturing the physical characteristics at the boundaries, i.e., the probability of connectivity should go to zero for short and long node separations. The polynomial was truncated at these boundaries. The average connection probability and correlation between distance and the connection function were calculated to be  $k_1 = 0.043$  and  $k_2 = 28.990$  miles, respectively. These constraints were then applied separately to compute the corresponding maximum entropy connection functions. The maximum entropy function was also computed when both constraints are active. The polynomial fit and the maximum entropy function are depicted in Fig. 3. There is clearly a discrepancy between each of the maximum entropy curves and the interpolated observed connection function. Yet it is interesting to note that the entropy resulting from the observed connection function is greater than 99.5% that of the maximum predicted value when both constraints are applied. However, this high value should be benchmarked. Comparing this to the case where only the mean constraint is applied, we found that the ratio was 95.3%, a high value despite the dissimilar connection functions (Fig. 3). When only the correlation constraint is applied, the ratio is 84.9%.

It is important to note that the idea of ensemble entropy viewed in this context relates to a *typical* flight

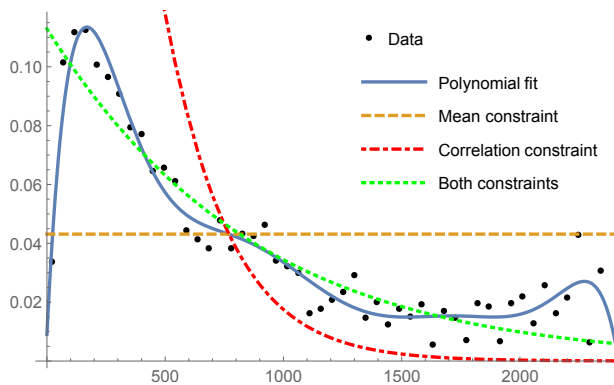


FIG. 3. Pair connection function derived from US flight data (solid line) and maximum conditional entropy functions for different sets of constraints (dashed lines) plotted as a function of the distance between primary airports in the conterminous US. Markers indicate raw histogram data obtained from [32, 33] centered in bins 47.3 miles wide. The solid line is an 8th-order polynomial fit of the marker data truncated at the boundaries.

network derived from an identical connection rule to that observed in this example. To elucidate this point, observe that the pair connection function estimated here is taken from data that describes a single network. So the Shannon entropy of the ensemble associated with this connection function would appear to bear little value. Yet, we know that the entropy of the ensemble is approximately equal to the logarithm of the size of the typical set of networks, and any particular network belonging to the typical set is observed with probability approximately equal to the reciprocal of the size of the set. The US flight network can be viewed as a member of the typical set of networks that arise from the corresponding pair connection function and pair distance distribution. Through this reasoning, we can relate the notion of entropy defined in eq. (2) to our study of a single realization of a practical network.

To explore this example further, degree-preserving edge permutations were carried out on the airline route data [34, 35], and the randomized networks were analyzed in a similar manner. All examples that were studied exhibited very similar behavior. The results for one such study are shown in Fig. 4. Predictably, the randomization procedure had the effect of “softening” the pair connection function, since the permutations were made independent of internode distance. Interestingly, the entropy resulting from the observed connection function in the randomized network is greater than 99.6% of the maximum predicted value when both constraints are applied. The ratio when only the mean constraint is applied is also approximately 99.6%, and the ratio corresponding to just the active correlation constraint is 73.7%. Although these results may first appear to contradict those obtained for the original data, it should be noted that the maximum entropy function in this example is matched to the new constraints, and hence the observations are rea-

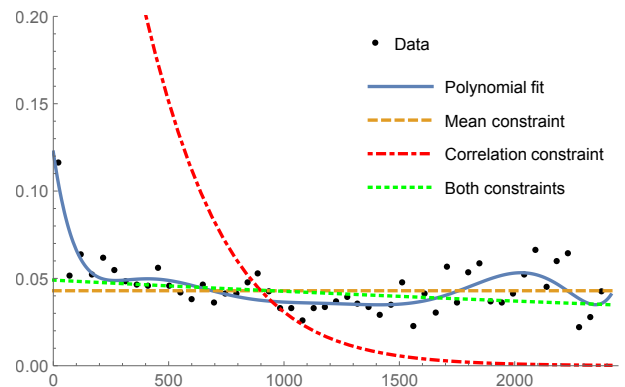


FIG. 4. Pair connection function derived from US flight data by executing a degree-preserving randomization procedure (solid line) and maximum conditional entropy functions for different sets of constraints (dashed lines) plotted as a function of the distance between primary airports in the conterminous US. All parameters are the same as for Fig. 3.

sonable. More importantly, the conclusion that we can draw from these two examples and that of the synthetically generated wireless network is that by limiting the network through both mean and correlation constraints matched to practical systems, we observe nearly maximally entropic behavior.

## VI. CONCLUSIONS

In this paper, we introduced a comprehensive framework for analyzing the Shannon entropy of spatial network ensembles. We first derived a simple upper bound on the (unconditional) entropy of an ensemble. The tightness of the bound relates to the degree of independence observed in the set of edge states. We then defined the conditional entropy of a spatial network ensemble to be the entropy of the graph topology ensemble conditioned on the underlying statistical distribution of the node locations. This formalism was exploited to calculate entropy maximizing pair connection functions conditioned on various constraints. Applying this result to two physical examples illustrated that maximally entropic engineered systems exist in practice.

Numerous extensions and modifications of the framework detailed herein can be explored. For example, characterizing the entropy of temporal network ensembles is a natural task that could shed light on spatial network dynamics. Directed network ensembles can also be studied. Furthermore, one could define different connection functions for different layers in a spatial multiplex and apply the method proposed here to study such ensembles. In general, the formalism adopted in this work can be used to derive a general theory of information related to spatial network structure. Along with such a theory would come operational interpretations of entropy in spatial networks – e.g., minimum description length of the

network topology – and these interpretations will lead to better understanding and more optimal design (where applicable) of physical networks.

## ACKNOWLEDGMENTS

This work was supported by EPSRC grant numbers EP/N002350/1 and EP/N002458/1.

- 
- [1] G. Simonyi, *Combinatorial Optimization* **20**, 399 (1995).
  - [2] A. Mowshowitz, *The bulletin of mathematical biophysics* **30**, 175 (1968).
  - [3] A. Mowshowitz, *Bulletin of Mathematical Biology* **30**, 225 (1968).
  - [4] A. Mowshowitz, *Bulletin of Mathematical Biology* **30**, 387 (1968).
  - [5] A. Mowshowitz, *Bulletin of Mathematical Biology* **30**, 533 (1968).
  - [6] K. Anand and G. Bianconi, *Physical Review E* **80**, 045102 (2009).
  - [7] K. Anand, G. Bianconi, and S. Severini, *Physical Review E* **83**, 036109 (2011).
  - [8] D. Bonchev, *Information theoretic indices for characterization of chemical structures*, 5 (Research Studies Press, 1983).
  - [9] M. G. Everett, *Social Networks* **7**, 353 (1985).
  - [10] M. Dehmer and A. Mowshowitz, *Information Sciences* **181**, 57 (2011).
  - [11] Y. Choi and W. Szpankowski, **58**, 620 (2012).
  - [12] N. de Beaudrap, V. Giovannetti, S. Severini, and R. Wilson, *A Panorama of Mathematics: Pure and Applied* **658**, 227 (2016).
  - [13] D. E. Simmons, J. P. Coon, and A. Datta, arXiv preprint arXiv:1703.01142 (2017).
  - [14] J. P. Coon, in *Global Communications Conference (GLOBECOM), 2016 IEEE* (IEEE, 2016) pp. 1–6.
  - [15] J. P. Coon and P. J. Smith, in *Communications (ICC), 2017 IEEE International Conference on* (IEEE, 2017) pp. 1–7.
  - [16] A. Cika, J. P. Coon, and S. Kim, in *Spatial Stochastic Models for Wireless Networks (SpaSWiN). International Workshop on* (IEEE, 2017) pp. 1–7.
  - [17] M. Dehmer, *Applied Mathematics and Computation* **201**, 82 (2008).
  - [18] S. Fehr and S. Berens, *IEEE Transactions on Information Theory* **60**, 6801 (2014).
  - [19] M. Barthélemy, *Physics Reports* **499**, 1 (2011).
  - [20] R. Lambiotte, V. D. Blondel, C. De Kerchove, E. Huens, C. Prieur, Z. Smoreda, and P. Van Dooren, *Physica A: Statistical Mechanics and its Applications* **387**, 5317 (2008).
  - [21] S. Riley, *Science* **316**, 1298 (2007).
  - [22] P. Wang, M. C. González, C. A. Hidalgo, and A.-L. Barabási, *Science* **324**, 1071 (2009).
  - [23] O. Georgiou, C. P. Dettmann, and J. P. Coon, *Journal of Statistical Physics* **147**, 758 (2012).
  - [24] D. Graff, “Implementation and evaluation of a SWARM-LINDA system,” (2008).
  - [25] A. Halu, S. Mukherjee, and G. Bianconi, *Physical Review E* **89**, 012806 (2014).
  - [26] M. D. Penrose, *The Annals of Applied Probability* **26**, 986 (2016).
  - [27] T. M. Cover and J. A. Thomas, *Elements of information theory* (John Wiley & Sons, 2012).
  - [28] M. Penrose, *Random geometric graphs*, Vol. 5 (Oxford University Press Oxford, 2003).
  - [29] C. P. Dettmann and O. Georgiou, *Physical Review E* **93**, 032313 (2016).
  - [30] J. P. Coon, O. Georgiou, and C. P. Dettmann, *Phys. Rev. E* **85**, 011138 (2012).
  - [31] J. Park and M. E. Newman, *Physical Review E* **70**, 066117 (2004).
  - [32] “OpenFlights Airport database,” <https://raw.githubusercontent.com/jpatokal/openflights/master/data/airports.dat>, accessed: 2017-04-19.
  - [33] “OpenFlights Route database,” <https://raw.githubusercontent.com/jpatokal/openflights/master/data/routes.dat>, accessed: 2017-04-18.
  - [34] S. Maslov and K. Sneppen, *Science* **296**, 910 (2002).
  - [35] “Degree-preserving randomization MATLAB code,” [http://maslov.bioengineering.illinois.edu/sym\\_generate\\_srand.m](http://maslov.bioengineering.illinois.edu/sym_generate_srand.m), accessed: 2018-02-07.