Full Connectivity: Corners, edges and faces

Justin Coon,[1] Carl P. Dettmann and Orestis Georgiou[3*]

[1] Toshiba Telecommunications Research Laboratory

32 Queen Square, Bristol BS1 4ND, United Kingdom

[2] University of Bristol School of Mathematics

University Walk, Bristol BS8 1TW, United Kingdom

[3] Max-Planck-Institute for the Physics of Complex Systems

Nöthnitzer Str. 38, 01187 Dresden, Germany

*To whom correspondence should be addressed.

Abstract

We develop a cluster expansion for the probability of *full connectivity* of high density random networks in confined geometries. In contrast to percolation phenomena at lower densities, boundary effects, which have previously been largely neglected, are not only relevant but dominant. We derive general analytical formulas that show a persistence of universality in a different form to percolation theory, and provide numerical confirmation. We also demonstrate the simplicity of our approach in three simple but instructive examples and discuss the practical benefits of its application to different models.

1 Introduction

Percolation is a phase-transition phenomenon in large random networks whereby at a critical value ρ_c of a parameter such as density that controls the connection probabilities, the largest connected component (cluster) of the system experiences a sudden change from being independent of system size (microscopic) to being proportional to it (macroscopic). For example in infinite systems with densities $\rho < \rho_c$ (sub-critical), all clusters are finite

almost surely while for $\rho > \rho_c$ (super-critical), any given node is in an infinite cluster with positive probability. As with other statistical phase transitions, in the thermodynamic limit of large system size, the percolation density ρ_c is very much independent of the size and shape of the system and in general of the microscopic details of the model leading to the phenomenon of universality. Percolation theory does not, however, address the question of when finite systems are fully connected, that is, the probability of the entire system comprising a single cluster. Here we develop a theory of the latter, showing that boundary effects are crucial for understanding this problem, and demonstrate a general formalism for calculating their contributions.

The well developed theory of percolation [1] finds its historic roots in the early 1930s in studies by physicists and chemists concerned with nucleation and condensation of gases into liquids as well as clustering of interacting particles and colloids. The theory of percolation was then initiated in 1957 in order to study random physical processes such as fluid flow through disordered porous media [2]. A more general mathematical approach was formulated soon thereafter in 1959 in the form of random graphs [3] using the famous Kolmogorov zero-one law in probability theory to state that given a infinite graph with connections chosen randomly and independently, the existence probability of an infinite cluster is either zero or one. A wealth of different models and approaches as well as out-standing open problems can be found in [4]. The exact statistical mechanical formalism of continuum percolation on a lattice-free basis was eventually formulated in the late 1970s [5] and has been extensively applied in a wide variety of settings as the theory provides useful information on cluster statistics (for a review article see [6]).

The triumph of percolation theory and statistical physics [7] in describing particle clustering in liquids [8], gases [9, 10] and colloids [11], was also successfully adopted in studies of electrical conductance in disordered media. These have included investigations of transport in carbon nano-tube networks [12] and metallic insulation in composite materials [13]. The concept of quantum percolation and its connections with the quantum Hall effect [14] as well as phenomena such as spontaneous magnetization [15] are but a few of the many recent theoretical advancements. Moreover, percolation theory has gone beyond physical sciences to topics such as network modeling. These include interpreting (self)-organization processes of complex networks, such as those displayed by functionally related proteins [16]; the spread of forest fires [17, 18], epidemics [19, 20, 21], computer [22] and phone [23]
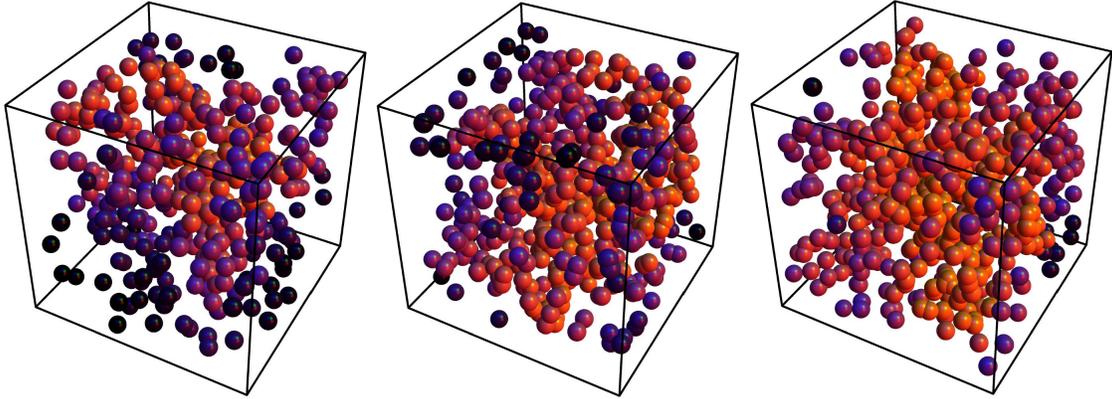
2

Figure 1: (Color online) A single high density realization is shown with $N = 450, 500$ and 600 nodes (represented as balls) randomly placed inside a cubed domain of side $L = 10$. Lighter colors indicate a higher probability of being in the largest connected component; note that isolated nodes (darker colors) are concentrated at the edges and corners at higher densities.

viruses; industrial and economic sectors [24]; and social groups of people [25, 26, 27].

The current work is primarily motivated by the recent adaptation of percolation theory in wireless communications and multihop relay networks. These consist of a number of communication devices (nodes) which can pass messages to each other without the need of a central router. Multihop relay networks can therefore achieve good and reliable coverage and connectivity over a large area even when some nodes are moved or deactivated. Most models of such systems are closely related or derived from studies by probabilists under the topic of random geometric graphs or networks [28]. Percolation theory, both on a lattice and in the continuum has thus been previously applied to random networks by engineers to identify and analyze power management techniques [29, 30], network resilience [31], efficient relay placement [32], coverage and connectivity in wireless sensor networks [33] (a direct adaptation of [10]), and the information theoretic capacity of networks [34].

A major weakness of the theory which is immediately evident in the current context is that many networks, although large, are not of infinite size and are often confined within a finite region. Therefore infinite (or 'giant') clusters and critical percolation densities no longer make sense in the usual thermodynamic limit. Instead, the more suited notion of full connectivity (also called connectivity threshold, originally posed in [3]), is addressed, typically in some asymptotic regime.

Here we consider a fixed connectivity scale r_0 so that nodes placed randomly in a
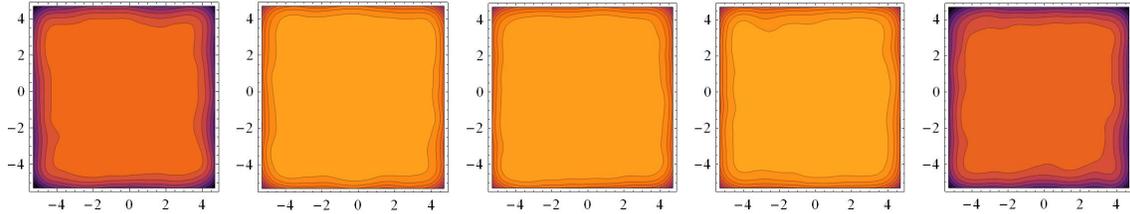
Figure 2: (Color online) Five equally spaced cross sections of the average connectivity (averaged over 100 different realizations) inside a cubed domain of side $L = 10$ with node density $\rho = 0.5$. The outer plots correspond to opposite faces of the cube. The same color scheme as in Figure 1 is used with light and dark colors indicating good and bad connectivity respectively. It is clear that corners, edges and faces of the cube constitute critical regions where full connectivity is lost with maximum probability.

domain $\mathcal{V} \subset \mathbb{R}^d$ are connected with probability given by a function $H(r/r_0)$ of the distance r between them; see the next section for more details. The number of nodes N and domain volume V (*i.e.* area for $d = 2$), and hence the density $\rho = N/V$ are allowed to vary while keeping the overall shape fixed. Other conventions such as fixing V while varying ρ and r_0 (dense network model), and fixing ρ while varying V and r_0 (extended network model) are used in the literature, but are equivalent in the absence of other scale-dependent physical effects [35]. Since r_0 is fixed in our work we often set it equal to unity.

Asymptotic results for full connectivity consider $\rho \to \infty$, with V constant or scaled as a function of ρ. Since the probability of any given node being disconnected decays exponentially with ρ, a limit where V grows exponentially with ρ is almost invariably used in the literature. This has the effect that boundaries become largely insignificant, with a few exceptions such as for higher order connectivity (k-connectivity) or higher dimensions [36]. Here we point out that realistic networks require a very high connection probability (hence large ρ) but do not have exponentially large volumes. So in practice finite size and in particular boundary effects are extremely important. Here we allow V to remain constant or to scale as a power of ρ.

The importance of boundary effects is illustrated in Figure 1, which shows a single realization of (from left to right), 450, 500 and 600 nodes (represented as overlapping balls) placed randomly inside a cubed domain, with each ball colored according to the total probability of it connecting to any of the other $N - 1$ nodes surrounding it. We assume here an exponentially decaying (with distance) connection probability function (see

equation (19)). Notice that isolated, and hence hard to connect, nodes (black) are in this case near corners, edges or faces of the cubed domain. Averaging over 100 such realizations and producing density plots of cross-sections of the cubed domain (see Figure 2) clearly demonstrates that these are critical regions where full connectivity is lost with maximum probability, even though there are many more nodes in the bulk than on the boundaries. To leading order in the high density limit, the probability of obtaining a fully connected network is controlled by the probability of a single isolated node. Lower density corrections to this are due to two, three and larger isolated clusters of nodes. This observation has formed the basis of our recent work [37], where we presented a new cluster-expansion approach to analyzing connectivity in confined geometries. This approach allowed for a consistent analysis and inclusion of the local geometric boundary effects leading to a general closed-form analytical formula for calculating the probability of full connectivity P_{fc}.

The current paper complements our previous results [37] by providing additional details of our approach in a more general setting, extends them with further calculations of second order correction terms, and also demonstrates the inaccuracy of conventional bulk models in the current setting through three simple, yet instructive example network confinements (disk, square and wedge) which are numerically confirmed for both probabilistic and sharp (unit-disk type) pair connectedness functions. The structure of the paper is as follows: In section 2, we derive a high density cluster approach which we then expand up to second order in subsections 2.1, 2.2 and 2.3 assuming a homogeneous system. In subsection 2.2 we also explain our scaling regime in more detail and in 2.3 we identify a key difference between probabilistic and sharp pair connectedness functions which is only apparent in higher order terms involving small isolated clusters of 2 or more nodes. We then compute the overall probability of a fully connected network for a given exponentially decaying connectivity function in subsection 2.4. In section 3, we lift the homogeneity assumption and consider boundary effects explicitly illustrating their importance for any connectivity probability function $H(r)$ that decays suitably fast. We then work through three examples (subsections 3.2, 3.3 and 3.4) for the previously given connectivity function and provide plots of comparisons with numerical simulations. These examples lead to the construction of a general formula for the overall probability of a fully connected network in subsection 3.5. Finally in section 4 we summarize and conclude with a short discussion.

2 A cluster expansion approach

In this section we develop a high density connectivity model based on a cluster expansion approach which is able to produce closed-form analytical results in some generality. We assume a homogeneous network and perform the expansion up to second order and comment on the third and higher order corrections. First however we need to set the stage and define some useful quantities.

Consider N randomly distributed nodes with locations $\mathbf{r}_i \in \mathcal{V}$, a convex subset of $\mathbb{R}^d$, where $i = 1, 2, \ldots N$, according to a uniform density $\rho = N/V$, where $V = |\mathcal{V}|$ and $|\cdot|$ denotes the size of the set using the Lebesgue measure of the appropriate dimension or the cardinality of a finite set. We assume that nodes i and j are directly connected with probability $H(r_{ij})$, often written just H_{ij} where $r_{ij} = |\mathbf{r}_j - \mathbf{r}_i|$ is the distance between i and j.

We may also consider a more general setting, as follows: $\mathbb{R}^d$ could be replaced by a Riemannian manifold $\mathcal{M}^d$ with volume element derived from its metric tensor. The direct connection distance function r_{ij} may or may not be given by geodesic distance derived from the metric tensor. The convexity assumption is then replaced by the statement that this distance function is a metric in the mathematical sense, *i.e.* symmetric, non-negative, zero only when the nodes coincide, and satisfying the triangle inequality.[1] One example of a non-Euclidean manifold is the surface of a sphere and is discussed along with its curvature effects in section 2.4. Otherwise we will use only Euclidean space $\mathbb{R}^d$.

We define the average of an observable O over all possible configurations as

$$\langle O \rangle = \frac{1}{V^N} \int_{\mathcal{V}^N} O(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N) \mathrm{d}\mathbf{r}_1 \mathrm{d}\mathbf{r}_2 \ldots \mathrm{d}\mathbf{r}_N. \tag{1}$$

We also need notation to define the relevant graphs. Let $S = \{1, 2, 3, \ldots, N\}$. A graph $g = (A, L)$ consists of a set $A \subseteq S$ of nodes, together with a collection $L \subseteq \{(i, j) \in A : i < j\}$ of direct links, that is unordered distinct pairs of nodes. As a slight abuse of notation we write $(i, j) \in g$ to denote that (i, j) is an element of the set of links L associated with the graph g. We write G^A for the set of graphs with nodes in A, and G_j^A for the set with nodes

---

[1]If the convexity assumption is relaxed it is a semi-metric function, *i.e.* the triangle inequality may not be satisfied. For example, r_{ij} is effectively infinite if there is an obstacle blocking the path of a wireless signal, but adding a node k that avoids the obstacle may permit indirect connection between i and j, thus in this case $r_{ij} > r_{ik} + r_{kj}$. We defer the treatment of obstacles to a future paper.

in A and largest connected component (cluster) of size j with $1 \leq j \leq |A|$.

The probability that two nodes are connected or not leads to the trivial identity:

$$1 \equiv H_{ij} + (1 - H_{ij}). \tag{2}$$

Multiplying over all links with nodes in a set A expresses the probability of all possible combinations. This can be written as

$$1 = \prod_{i,j \in A; i<j} [H_{ij} + (1 - H_{ij})] = \sum_{g \in G^A} \mathcal{H}_g, \tag{3}$$

where

$$\mathcal{H}_g = \prod_{(i,j) \in g} H_{ij} \prod_{(i,j) \notin g} (1 - H_{ij}). \tag{4}$$

The sum in equation (3) contains $2^{|A|(|A|-1)/2}$ separate terms. Setting $A = S$, this can be expressed as collections of terms determined by their largest cluster:

$$1 = \underbrace{\sum_{g \in G_N^S} \mathcal{H}_g}_{P_{fc}(\mathbf{r}_1, \dots \mathbf{r}_N)} + \sum_{g \in G_{N-1}^S} \mathcal{H}_g + \dots + \underbrace{\sum_{g \in G_1^S} \mathcal{H}_g}_{\prod_{i<j}(1-H_{ij})} . \tag{5}$$

For a given configuration of node positions $\mathbf{r}_i \in \mathcal{V}$, assuming that the nodes are pairwise connected with independent probabilities H_{ij}, the first term in equation (5) is the probability of being fully connected $P_{fc}(\mathbf{r}_1, \dots \mathbf{r}_N)$. The average of this quantity over all possible configurations $P_{fc} = \langle P_{fc}(\mathbf{r}_1, \dots \mathbf{r}_N) \rangle$ is the overall probability of obtaining a fully connected network and is our desired quantity of interest. Hence, rearranging equation (5) allows us to obtain expressions for P_{fc} in a consistent way while keeping track of correction terms.

2.1   Zeroth order approximation

In the very high density limit of $\rho \to \infty$, the right hand side of equation (5) is dominated by the first term

$$P_{fc} \approx 1, \tag{6}$$

and hence the network is fully connected with probability one. The approximation symbol is used here and from now on to indicate that higher order terms are being ignored.

2.2 First order approximation

The first order approximation is obtained when the second term in equation (5) is expanded out explicitly. This takes into account all the ways of having an $N-1$ cluster of nodes. Thus the overall probability of a fully connected network is

$$
\begin{aligned}
P_{fc} &\approx 1 - \langle \sum_{g \in G_{N-1}^S} \mathcal{H}_g \rangle \\
&= 1 - \langle \left(\sum_{\wp=1}^N \prod_{j \neq \wp} (1 - H_{j\wp}) \right) \underbrace{\left(\sum_{g \in G_{N-1}^{S \setminus \{\wp\}}} \mathcal{H}_g \right)}_{\approx 1} \rangle \\
&= 1 - N \langle \prod_{j=1}^{N-1} (1 - H_{jN}) \rangle \\
&= 1 - \frac{N}{V^N} \int_{\mathcal{V}^N} \prod_{j=1}^{N-1} (1 - H(\mathbf{r}_{jN})) \mathrm{d}\mathbf{r}_1 \ldots \mathrm{d}\mathbf{r}_N \\
&= 1 - \frac{N}{V} \int_{\mathcal{V}} \left(1 - \frac{1}{V} \int_{\mathcal{V}} H(\mathbf{r}_{1N}) \mathrm{d}\mathbf{r}_1 \right)^{N-1} \mathrm{d}\mathbf{r}_N,
\end{aligned}
\tag{7}
$$

since all nodes are identical and therefore the sum over $\wp$ can be factored out.

Assuming that the network is homogeneous (*i.e.* there are no boundary effects and therefore the system is symmetric under translational transformations) allows for the change in variables $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_N$ so the integrals decouple and we get

$$
\begin{aligned}
P_{fc} &\approx 1 - N \left(1 - \frac{1}{V} \int_{\mathcal{V}} H(\mathbf{r}) \mathrm{d}\mathbf{r} \right)^{N-1} \\
&= 1 - N e^{-\rho \int_{\mathcal{V}} H(\mathbf{r}) \mathrm{d}\mathbf{r}} \left[1 + \frac{1}{N} \left(\rho \int_{\mathcal{V}} H(\mathbf{r}) \mathrm{d}\mathbf{r} - \frac{(\rho \int_{\mathcal{V}} H(\mathbf{r}) \mathrm{d}\mathbf{r})^2}{2} \right) + \mathcal{O} \left(\frac{\rho^4}{N^2} \right) \right],
\end{aligned}
\tag{8}
$$

for N large.

At this point we are in a position to discuss the various scaling limits and approximations appearing in the theory. Until now we have assumed only that the probability of full connectivity is high, which is reasonable for typical applications such as wireless networks. This means that ρV_H is large where $V_H = \int_{\mathcal{V}} H(\mathbf{r}) \mathrm{d}\mathbf{r}$ is the effective connectivity volume associated with $H(r)$. Increasing N and V at constant ρ will however decrease P_{fc} due to the factor of N in front of the exponential. We see that at fixed P_{fc}, system size scales exponentially with density. While this is a popular scaling in the literature, it has the effect

of hiding boundary effects, which in the current scaling regime as we will see are indeed very important. Alternatively we can increase ρ while keeping V fixed. For this purpose, and for any scaling in which V/ρ decreases, we must use the first line in equation (8) as the second line clearly does not converge. However it is also advantageous to assume V/V_H is large, for example to separate effects due to the range of direct connections which is proportional to $V_H^{1/d}$, and effects due to the finite extent of the whole system. In this case, so for any scaling where V is between linear and exponential in ρ, we can use the simpler second line of equation (8). We will take the latter approach unless otherwise stated.

2.3 Second order approximation

The second order approximation involves enumerating the $N-2$ cluster terms, together with the first approximation of the $G_{N-1}^{S\backslash\{\wp\}}$ term in equation (7) such that

$$P_{fc} \approx 1 - \langle \sum_{g \in G_{N-1}^S} \mathcal{H}_g \rangle - \langle \sum_{g \in G_{N-2}^S} \mathcal{H}_g \rangle. \tag{9}$$

We first examine the first approximation of the G_{N-1}^S term thus including all possible ways of getting an $N-2$ cluster in a $N-1$ node network:

$$\langle \sum_{g \in G_{N-1}^S} \mathcal{H}_g \rangle = \langle \sum_{\wp=1}^{N} \prod_{j \neq \wp} (1 - H_{j\wp}) \underbrace{\sum_{g \in G_{N-1}^{S\backslash\{\wp\}}} \mathcal{H}_g}_{\approx 1 - \sum_{g \in G_{N-2}^{S\backslash\{\wp\}}} \mathcal{H}_g} \rangle, \tag{10}$$

and

$$\langle \sum_{g \in G_{N-2}^{S\backslash\{\wp\}}} \mathcal{H}_g \rangle = \frac{1}{2} \langle \sum_{\ell=1}^{N-1} \prod_{j \neq \ell} (1 - H_{j\ell}) \underbrace{\sum_{g \in G_{N-2}^{S\backslash\{\wp,\ell\}}} \mathcal{H}_g}_{\approx 1} \rangle. \tag{11}$$

In this way we avoid double counting of different decompositions of subgraphs. Expanding equation (10) out we have

$$\langle \sum_{g \in G_{N-1}^S} \mathcal{H}_g \rangle = \langle \sum_{\wp=1}^{N} \prod_{j \neq \wp} (1 - H_{j\wp}) - \frac{1}{2} \left(\sum_{\wp=1}^{N} \overbrace{\prod_{j \neq \wp} (1 - H_{j\wp})}^{(N-1)\text{terms}} \sum_{\ell \neq \wp} \overbrace{\prod_{j \neq \ell, \wp} (1 - H_{j\ell})}^{(N-2)\text{terms}} \right) \rangle$$

$$= N \langle \prod_{j=2}^{N} (1 - H_{1j}) - \frac{N(N-1)}{2} \left((1 - H_{12}) \prod_{j=3}^{N} (1 - H_{1j})(1 - H_{2j}) \right) \rangle, \tag{12}$$

9

since all nodes are identical.

We now consider the last term in (9). There are two possible ways of getting an $N - 2$ cluster. These correspond to having two isolated nodes and one large $N - 2$ cluster, or having a small 2-node cluster and a large $N - 2$ cluster. Hence, adding these two (disjoint) possibilities together we have that

$$
\begin{aligned}
\langle \sum_{g \in G_{N-2}^S} \mathcal{H}_g \rangle = {} & \left\langle \left(\sum_{\wp < \ell} \prod_{j \neq \wp} (1 - H_{j\wp}) \prod_{j \neq \wp, \ell} (1 - H_{j\ell}) \right) \underbrace{\left(\sum_{g \in G_{N-2}^{S \setminus \{\wp, \ell\}}} \mathcal{H}_g \right)}_{\approx 1} \right\rangle \\
& + \left\langle \left(\sum_{\wp < \ell} H_{\wp\ell} \prod_{j \neq \wp, \ell} (1 - H_{j\wp})(1 - H_{j\ell}) \right) \underbrace{\left(\sum_{g \in G_{N-2}^{S \setminus \{\wp, \ell\}}} \mathcal{H}_g \right)}_{\approx 1} \right\rangle, \\
= {} & \frac{N(N-1)}{2} \langle (1 - H_{12}) \prod_{j=3}^N (1 - H_{1j})(1 - H_{2j}) \rangle \\
& + \frac{N(N-1)}{2} \langle H_{12} \prod_{j=3}^N (1 - H_{1j})(1 - H_{2j}) \rangle, \\
= {} & \frac{N(N-1)}{2} \langle \prod_{j=3}^N (1 - H_{1j})(1 - H_{2j}) \rangle,
\end{aligned}
\tag{13}
$$

due to the identity of equation (2).

Putting it together we find a clear sum of zeroth, first and second order terms:

$$
P_{fc} \approx 1 - N \langle \prod_{j=2}^N (1 - H_{1j}) \rangle - \frac{N(N-1)}{2} \langle H_{12} \prod_{j=3}^N (1 - H_{1j})(1 - H_{2j}) \rangle
\tag{14}
$$

This is our most general result, allowing any dimension, region, and suitably decaying function $H(r)$. Assuming that the network is homogeneous as before gives

$$
P_{fc} \approx 1 - N \left(1 - \frac{1}{V} \int_{\mathcal{V}} H(\mathbf{r}) d\mathbf{r} \right)^{N-1} - \frac{N(N-1)}{2V} \int H(\mathbf{r}) \left(1 - \frac{K(\mathbf{r})}{V} \right)^{N-2} d\mathbf{r}
\tag{15}
$$

where

$$
K(\mathbf{r}_1) = V - \int_{\mathcal{V}} (1 - H(\mathbf{r}_2))(1 - H(\mathbf{r}_{12})) d\mathbf{r}_2
\tag{16}
$$

is a quantity that remains bounded as V increases. For sufficiently large V (see (8) above) we can approximate this as

$$
P_{fc} \approx 1 - N e^{-\rho \int_{\mathcal{V}} H(\mathbf{r}) d\mathbf{r}} - \frac{\rho N}{2} \int_{\mathcal{V}} H(\mathbf{r}) e^{-\rho K(\mathbf{r})} d\mathbf{r}.
\tag{17}
$$

10

2.4 The $H(\mathbf{r})$ function and connectivity on a sphere

Translational invariance is unlikely to be a good assumption in practice, except if the nodes are situated on the surface of some n-sphere $S^n \subset \mathbb{R}^{n+1}$ with $n \geq 1$ such that the system is symmetric under rotations about the origin. Here, we consider the realistic case of $n = 2$ for a specific connectivity function related to wireless communication networks, which is exponentially decaying with distance.

The information outage probability P_{out} for a single-input single-output (SISO) Rayleigh fading link model[2], is given by [38]

$$
\begin{aligned}
P_{out} &= \Pr\left(\log_2\left(1 + \text{SNR} \times |h|^2\right) < x\right) \\
&= \Pr\left(|h|^2 < \frac{2^x - 1}{\text{SNR}}\right),
\end{aligned}
\tag{18}
$$

where h is the channel transfer coefficient, x is the minimum outage rate threshold, and the *signal-to-noise ratio* (SNR) $\propto r^{-\eta}$ where r is the dimensionless distance between connected nodes (relative to the signal wavelength), and η is an environment dependent decay parameter[3]. The random variable $|h|^2$ is typically drawn from a *standard exponential* distribution. Therefore the connectivity probability function can be written as

$$
H(\mathbf{r}) = 1 - P_{out} = e^{-\beta r^{\eta}},
\tag{19}
$$

with β typically a small dimensionless constant, and is plotted in Figure 3 for $\beta = 0.01$ and different values of η. Notice that as $\eta \to \infty$ the connectivity is no longer probabilistic and converges to the interesting case of the popular unit disk model where connections have a fixed range of $r_0 = \beta^{-\frac{1}{\eta}}$.

The surface element on a sphere of radius R is $R^2 \sin\theta \mathrm{d}\theta d\phi$ in the usual spherical coordinates. First we use the Euclidean distance in $\mathbb{R}^3$ between two nodes, given in terms of the angle they subtend at the center: $d(\mathbf{r}_1, \mathbf{r}_2) = |\mathbf{r}_{12}| = \sqrt{2}R\sqrt{1 - \cos\theta_{12}}$. Setting $\eta = 2$ (corresponding to propagation in free space) and locating one node at the pole $\theta = 0$ we

---

[2]Other link models can also be considered. These may include single-input multiple-output (SIMO), multiple-input single-output (MISO), and multiple-input multiple-output (MIMO) (see also [39]).

[3]Typically $\eta = 2$ corresponds to propagation in free space but for practical reasons it is often modeled as $\eta > 2$ for cluttered environments.
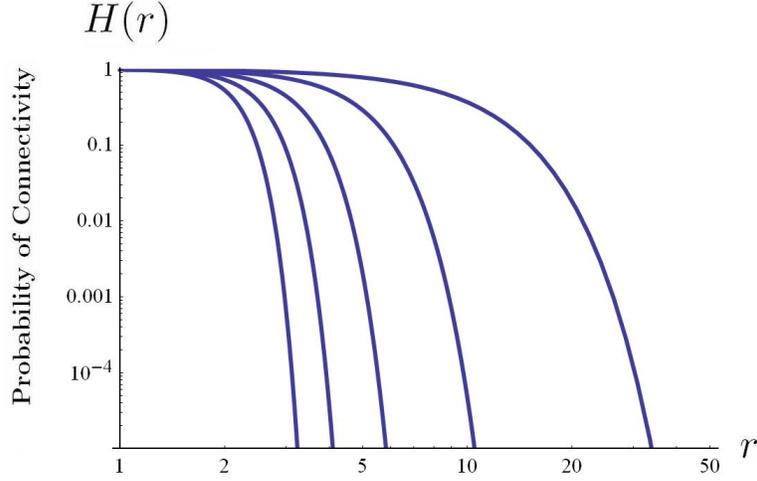
Figure 3: The probability of connecting a distance r from a node is plotted for $\eta = 2, 3, 4, 5, 6$, form right to left, using $\beta = 0.01$.

therefore have that

$$
\begin{aligned}
\int_{S^2} H(\mathbf{r})\mathrm{d}\mathbf{r} &= \int_0^{2\pi} \int_0^{\pi} e^{-2\beta R^2(1-\cos\theta)} R^2 \sin\theta \mathrm{d}\theta \mathrm{d}\phi \\
&= \frac{\pi}{\beta}\left[1 - e^{-4\beta R^2}\right] \approx \frac{\pi}{\beta},
\end{aligned}
\tag{20}
$$

when βR^2 (proportional to V/V_H above) is large.

If instead geodesic (great-circle) distance is considered rather than Euclidean we have $d(\mathbf{r}_1, \mathbf{r}_2) = R\theta$ and hence

$$
\begin{aligned}
\int_{S^2} H(\mathbf{r})\mathrm{d}\mathbf{r} &= \int_0^{2\pi} \int_0^{\pi} e^{-\beta(R\theta)^2} R^2 \sin\theta \mathrm{d}\theta \mathrm{d}\phi \\
&= \frac{\pi}{\beta}\left[1 - \frac{1}{\beta R^2}\left(\frac{1}{6} + \frac{e^{-\beta\pi^2 R^2}}{2\pi^2}\right) + \mathcal{O}\left(\frac{1}{\beta^2 R^4}\right)\right] \approx \frac{\pi}{\beta},
\end{aligned}
\tag{21}
$$

again when βR^2 is large. In this limit we see that the curvature (R-dependent) effects become small for both distance functions. It is thus possible to approximate the neighborhood of each node by a flat Euclidean metric, more generally for other smooth manifolds $\mathcal{M}^d$, assuming a sufficiently fast decay of $H(r)$.

Since the dominant contribution of $\int_{\mathcal{V}} H(\mathbf{r}_1)H(\mathbf{r}_{12})\mathrm{d}\mathbf{r}_1$ is when nodes 1 and 2 are close to each other (*i.e.* when $d(\mathbf{r}_1, \mathbf{r}_2) \ll R$), we adopt the above approximation in order calculate

$K(\mathbf{r}_2)$ using polar coordinates

$$
\begin{aligned}
K(\mathbf{r}_2) &= \int_{S^2} H(\mathbf{r}_1) + H(\mathbf{r}_{12}) - H(\mathbf{r}_1)H(\mathbf{r}_{12})\mathrm{d}\mathbf{r}_1 \\
&\approx 2\frac{\pi}{\beta} - \int_{\mathbb{R}^2} H(\mathbf{r}_1)H(\mathbf{r}_{12})\mathrm{d}\mathbf{r}_1 \\
&= 2\frac{\pi}{\beta} - \int_0^\infty \int_0^{2\pi} r_1 e^{-\beta(2r_1^2 + r_2^2 - 2r_1 r_2 \cos\theta)}\mathrm{d}\theta\mathrm{d}r_1 \\
&= 2\frac{\pi}{\beta} - \int_0^\infty 2\pi r_1 e^{-\beta(2r_1^2 + r_2^2)} I_0(2\beta r_1 r_2)\mathrm{d}r_1 \\
&= 2\frac{\pi}{\beta} - \frac{\pi}{2\beta}e^{-\beta\frac{r_2^2}{2}},
\end{aligned}
$$

(22)

where $I_0(x)$ is the modified Bessel function of the first kind, r_1 and r_2 are the distances of the two nodes from the origin and θ the angle between them. Finally, we may evaluate

$$
\begin{aligned}
\int_{S^2} H(\mathbf{r})e^{-\rho K(\mathbf{r})}\mathrm{d}\mathbf{r} &\approx \int_{\mathbb{R}^2} H(\mathbf{r})e^{-\rho K(\mathbf{r})}\mathrm{d}\mathbf{r} \\
&= e^{-\frac{3\pi\rho}{2\beta}}\left(\frac{2}{\pi\rho} + \mathcal{O}(\rho^{-2})\right),
\end{aligned}
$$

(23)

by expanding the exponential in equation (22) around $r_2 = 0$ as this is near the main contribution to the integral. Altogether, according to equation (17) this gives

$$
P_{fc} \approx 1 - Ne^{-\rho\frac{\pi}{\beta}} - \frac{N}{\pi}e^{-\frac{3\pi\rho}{2\beta}},
$$

(24)

to leading order.

Note that the above calculation can of course be performed for general $\eta > 0$ and n-sphere. The general case is excluded for the sake of brevity, however one case that provides some insight is the "unit-disk" model, corresponding to the limit $\eta \to \infty$ with β scaled so that $\beta^{-1/\eta}$ approaches a limit, r_0. We then have

$$
H(r) = \begin{cases} 1 & r \le r_0 \\ 0 & r > r_0 \end{cases}
$$

(25)

leading to

$$
\int_{S^2} H(\mathbf{r})\mathrm{d}\mathbf{r} \approx \int_{\mathbb{R}^2} H(\mathbf{r})\mathrm{d}\mathbf{r} = 2\pi r_0^2
$$

(26)

for either metric (actually it is exact for the Euclidean metric). We then calculate on $\mathbb{R}^2$ for $r_2 < 2r_0$

$$
\begin{aligned}
K(\mathbf{r}_2) &= 2\pi r_0^2 + \frac{r_2}{2}\sqrt{4r_0^2 - r_2^2} - 2r_0^2 \arccos\frac{r_2}{2r_0} \\
&= \pi r_0^2 + 2r_0 r_2 - \frac{r_2^3}{12r_0} + \dots
\end{aligned}
$$

(27)

13

expanding for small r_2 as this leads to the main contribution to P_{fc}. The latter comes to

$$P_{fc} = 1 - Ne^{-\pi r_0^2 \rho} \left[1 + \frac{\pi}{4r_0^2 \rho} + \ldots \right] \tag{28}$$

where the last term is the second order contribution.

We now make an important observation. From equation (24) we notice that in the high density limit the second order term in our expansion of P_{fc} has exponent of order $\sim -\frac{3\pi\rho}{2\beta}$ and hence is exponentially smaller than the first order term. This is due to the probabilistic nature of the connections, two nodes need more excluded volume on average to be isolated than a single node. For the unit disk model, however, the two nodes can be very close with a probability that is an algebraic function of their mutual distance, requiring roughly the same excluded volume, and leading to a second order effect which differs from the first only by an algebraic function. Hence, the probability that two nearby nodes are isolated from the rest is much less for a probabilistic pair connectedness function than for a sharp one. We expect that this key difference also applies for small isolated clusters consisting of 3 or more nodes.

This means that higher order effects appear to be important mostly for the unit disk model. For the probabilistic model which is more realistic for applications such as wireless networks, we are justified to use only the first order results for the discussion of boundary effects in confined geometries; to this we now turn.

3   Boundary effects

In this section, we lift the homogeneity assumption and consider the first and second order approximations of P_{fc} only to discover that contrary to popular belief and practice, boundary effects matter. This is a central observation of the current work, which we detail and discuss in the following subsections.

3.1   Inhomogeneous first and second order approximations

Returning to the first order equation (7) but no longer assuming homogeneity gives

$$\begin{aligned}
P_{fc} &\approx 1 - \frac{N}{V} \int_{\mathcal{V}} \left(1 - \frac{1}{V} \int_{\mathcal{V}} H(\mathbf{r}_{12}) \mathrm{d}\mathbf{r}_1 \right)^{N-1} \mathrm{d}\mathbf{r}_2 \\
&= 1 - \rho \int_{\mathcal{V}} e^{-\rho \int_{\mathcal{V}} H(\mathbf{r}_{12}) \mathrm{d}\mathbf{r}_1} \left(1 + \mathcal{O}(N^{-1}) \right) \mathrm{d}\mathbf{r}_2,
\end{aligned} \tag{29}$$

14

for large N. This equation was recently given in [35] (equation (8)) with V scaled exponentially with ρ thus ignoring any boundary effects. Recall that for this approximation we only require that $V \gg \rho$ or equivalently $V \gg \sqrt{N}$. This is a key difference from previous percolation approaches.

Similarly, from equations (12) and (13), one can show that the second order approximation of P_{fc} becomes

$$
\begin{aligned}
P_{fc} &= 1 - \rho \int_{\mathcal{V}} e^{-\rho \int_{\mathcal{V}} H(\mathbf{r}_{12}) \mathrm{d}\mathbf{r}_1} \mathrm{d}\mathbf{r}_2 - \frac{N(N-1)}{2V^2} \int \int_{\mathcal{V}} H(\mathbf{r}_{12}) \left(1 - \frac{\hat{K}(\mathbf{r}_1, \mathbf{r}_2)}{V} \right)^{N-2} \mathrm{d}\mathbf{r}_1 \mathrm{d}\mathbf{r}_2 \\
&\approx 1 - \rho \int_{\mathcal{V}} e^{-\rho \int_{\mathcal{V}} H(\mathbf{r}_{12}) \mathrm{d}\mathbf{r}_1} \mathrm{d}\mathbf{r}_2 - \frac{\rho^2}{2} \int \int_{\mathcal{V}} H(\mathbf{r}_{12}) e^{-\rho \hat{K}(\mathbf{r}_1, \mathbf{r}_2)} \mathrm{d}\mathbf{r}_1 \mathrm{d}\mathbf{r}_2,
\end{aligned}
$$

$$(30)$$

where

$$
\hat{K}(\mathbf{r}_1, \mathbf{r}_2) = V - \int_{\mathcal{V}} (1 - H(\mathbf{r}_{13}))(1 - H(\mathbf{r}_{23})) \mathrm{d}\mathbf{r}_3, \tag{31}
$$

and in the last step of (30) we have ignored terms of order $\sim N^{-1}$ since N is also assumed to be large.

Equation (29) suggests that in the high density limit, the probability of having a single $N - 1$ connected cluster is dominated by nodes which are situated in "*hard to connect*" regions of the available domain $\mathcal{V}$. This is because the outer integral in (29) is dominated by contributions where the integral in the exponential is small, for example at corners, edges and faces. The second order corrections in (30), as expected from our calculations in section 2.4, are of secondary importance and do not offer further insight. It is important to note that our approach here contradicts the usual universality scenario found in statistical mechanics and emphasizes the dominant importance of boundaries (and in particular corners). We stress here that this observation does not depend on using Euclidean distance and is valid in any geometry and any dimension where the lack of connectivity is due to a situation involving a single disconnected node and an $N - 1$ cluster.. In such a case, the outage probability will be dominated by situations where that node has a small volume in range.

Returning to the general Rayleigh fading link model of equation (19) we now demonstrate the above observation in a straight forward way. Suppose that $\mathbf{r}_2$ is situated somewhere on the boundary of the network domain $\mathcal{V} \subset \mathbb{R}^d$. Since direct connectivity is expo-

nentially decaying, we may separate the $d\mathbf{r}_1$ integral and extract its leading order behavior

$$\int_{\mathcal{V}} H(\mathbf{r}_{12})d\mathbf{r}_1 \approx \left(\int_0^\infty r^{d-1}e^{-\beta r^\eta}dr\right)\left(\int d\Omega\right) = \frac{\Gamma\left(\frac{d}{\eta}\right)}{\eta\beta^{\frac{d}{\eta}}}\omega, \tag{32}$$

where $\Omega = \frac{2\pi^{\frac{d}{2}}}{\Gamma\left(\frac{d}{2}\right)}$ is the full solid angle in d dimensions and $\omega \in (0, \Omega)$ is the solid angle available from $\mathbf{r}_2$. In the case of $\eta = 2$ the gamma functions cancel and equation (32) simplifies to $\left(\frac{\pi}{\beta}\right)^{\frac{d}{2}}\frac{\omega}{\Omega}$. As disused before, scaling $\beta = r_0^{-\eta}$ we recover the popular unit-disk model in the limit of $\eta \to \infty$, allowing for direct comparison with the probabilistic case of (19) and other earlier results. For example, if $\eta = d$ then equation (32) gives r_0^d/d which is also what the unit-disk model gives.

In general we expect that the outer integral in (29) will have contributions from boundary regions as in (32), with a term of the order of $\exp(-\rho\omega\int H(r)r^{d-1}dr)$ for the smallest ω, for example the pointiest corner, dominating P_{fc} at high density. It is clear from this argument that connectivity is indeed dominated from regions which are hard to connect and is controlled by the solid angle available to them.

The above observation has brought forward a radically different understanding of connectivity in confined geometries, namely that full connectivity is dominated by the critical, hard to connect areas such as corners, edges and faces. Furthermore, this suggests the decomposition of the probability of a fully connected network P_{fc} into a sum of contributions due to boundary objects with different solid angles. In order to obtain a more in-depth understanding of this novel idea we now embark into a more detailed investigation by considering a few example domains. Through these examples, the boundary effects will become clear thus leading to the main result of this section; a general formula (see equation (46)) for P_{fc}.

3.2 Example 1. Disk ($\eta = 2$)

We start with a circular domain $d_R \subset \mathbb{R}^2$ of radius R. Using a Euclidean metric, the distance between two nodes is given by $d(\mathbf{r}_1, \mathbf{r}_2) = |\mathbf{r}_{12}| = \sqrt{|\mathbf{r}_1|^2 + |\mathbf{r}_2|^2 - 2|\mathbf{r}_1||\mathbf{r}_2|\cos\theta}$. Hence, for the case of $\eta = 2$ we must first evaluate

$$\begin{aligned}
\int_{d_R} H(\mathbf{r}_{12})d\mathbf{r}_1 &= \int_0^R \int_0^{2\pi} \left(r_1 e^{-\beta\left(r_1^2 + r_2^2 - 2r_1 r_2 \cos\theta\right)}\right) d\theta dr_1 \\
&= 2\pi \int_0^R \left(r_1 I_0(2r_1 r_2 \beta)e^{-\beta(r_1^2 + r_2^2)}\right) dr_1,
\end{aligned} \tag{33}$$

where $I_0(x)$ is the modified Bessel function of the first kind. Unlike the ordinary Bessel function, which is oscillatory for real arguments, $I_0(x)$ is exponentially growing. When $\mathbf{r}_2$ is close to the center (i.e. $r_2 \approx 0$) we have that $e^{-\beta r_2} I_0(2r_1 r_2 \beta) = 1 + \mathcal{O}(r_2^2)$ so that the integral becomes:

$$2\pi \int_0^R \left(r_1 I_0(2r_1 r_2 \beta) e^{-\beta(r_1^2 + r_2^2)} \right) \mathrm{d}r_1 = \frac{\pi}{\beta} \left(1 - e^{-\beta R^2} \right) + \mathcal{O}(r_2^2) \approx \frac{\pi}{\beta}, \tag{34}$$

where we have assumed that $\beta R^2 \gg 1$. Notice that this is the same assumption made in equations (20) and (21) and will be consistently used throughout our calculations. Otherwise, if $\mathbf{r}_2$ is not near the center of c_R, we have the asymptotic relation $I_0(x) = \frac{e^x}{\sqrt{2\pi x}} \left(1 + \mathcal{O}(x^{-1}) \right)$ so that the integral becomes:

$$
\begin{aligned}
2\pi \int_0^R \left(r_1 I_0(2r_1 r_2 \beta) e^{-\beta(r_1^2 + r_2^2)} \right) \mathrm{d}r_1 &= 2\pi \int_0^R \left(r_1 \frac{e^{2r_1 r_2 \beta}}{\sqrt{4\pi r_1 r_2 \beta}} e^{-\beta(r_1^2 + r_2^2)} \right) \mathrm{d}r_1 \\
&\approx \frac{\sqrt{\pi}}{\sqrt{\beta}} \int_0^R e^{-\beta(r_1 - r_2)^2} \mathrm{d}r_1 \\
&= \frac{\pi}{2\beta} \left(\mathrm{erf}\left[\sqrt{\beta}(R - r_2) \right] + \mathrm{erf}\left[\sqrt{\beta} r_2 \right] \right),
\end{aligned}
\tag{35}
$$

since the main contribution of this integral comes from $r_1 \approx r_2$. Matching the two solutions for small and large r_2 we obtain a nice approximation to equation (33)

$$
\begin{aligned}
\int_{c_R} H(\mathbf{r}_{12}) \mathrm{d}\mathbf{r}_1 &\approx \frac{\pi}{2\beta} \left(\mathrm{erf}\left[\sqrt{\beta}(R - r_2) \right] + 1 \right) \\
&= \frac{\pi}{2\beta} f(r_2).
\end{aligned}
\tag{36}
$$

In order to make progress, we approximate $f(r)$ by

$$\tilde{f}(r) = \begin{cases} c_1, & \text{for} \quad 0 < r < a, \\ c_2 - m(r - R), & \text{for} \quad a \le r < R, \end{cases} \tag{37}$$

where $c_1 = 2\mathrm{erf}\left[\sqrt{\beta}\frac{R}{2} \right] \approx 2$, $c_2 = \mathrm{erf}\left[\sqrt{\beta}R \right] \approx 1$, $m = \frac{2\sqrt{\beta}}{\sqrt{\pi}} \left(1 - e^{-\beta R^2} \right) \approx \frac{2\sqrt{\beta}}{\sqrt{\pi}}$ when $\sqrt{\beta}R \gg 1$ and a is their intersection approximating the location of the effective turning point of $f(r)$.
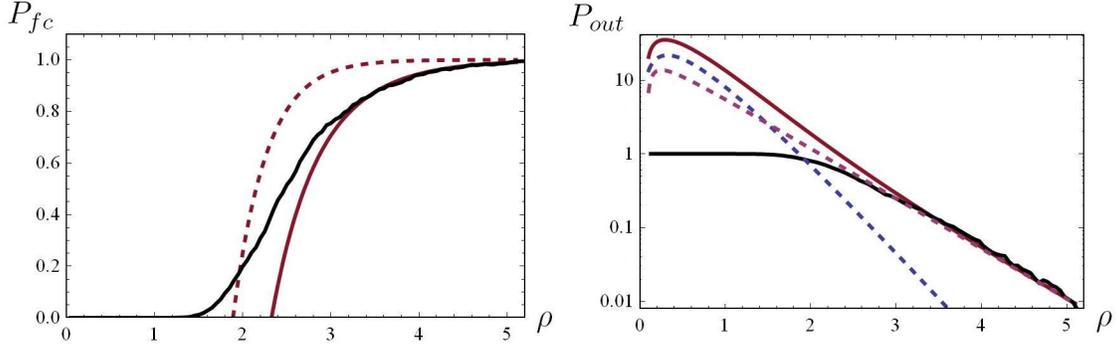
Figure 4: *Left:* (Color online) A plot of $P_{fc}(\rho)$ in a disk d_R of area $A = 200$ ($\beta = 1$) (solid red) compared with the prediction of equation (8) (dashed) and a numerical simulation (black). *Right:* The dashed curves show the contributions from the bulk (blue) and boundary (purple) term in equation (38) on a log-linear scale. The full curves are plots of $P_{out}(\rho)$ with dark red and black corresponding to theory and numerical simulation.

The outer integral in equation (29) can now be approximated to give

$$
\begin{aligned}
P_{fc} &= 1 - \rho \int_{d_R} e^{-\rho \int_{c_R} H(\mathbf{r}_{12})\mathrm{d}\mathbf{r}_1} \mathrm{d}\mathbf{r}_2 \\
&\approx 1 - 2\pi\rho \int_0^R r e^{-\rho \frac{\pi}{2\beta} \tilde{f}(r)} \mathrm{d}r \\
&= 1 - \pi R^2 \rho e^{-\rho \frac{\pi}{\beta}} \left(1 - \frac{\sqrt{\pi}}{\sqrt{\beta}R} - \frac{2\sqrt{\beta}}{\rho\sqrt{\pi}R} + \mathcal{O}(R^{-2}) \right) - 2\pi R \sqrt{\frac{\beta}{\pi}} e^{-\frac{\pi}{2\beta}\rho} \left(1 - \frac{\sqrt{\beta}}{\rho\sqrt{\pi}R} + \mathcal{O}(R^{-2}) \right).
\end{aligned}
\tag{38}
$$

It is clear from equation (38) that P_{fc} is composed of a bulk term and a boundary term with coefficients characteristic of the area and perimeter of c_R, with corrections due to the curvature of the boundary which vanish as $R \to \infty$. Moreover, P_{fc} is dominated by the boundary term in the high density limit. This can be seen on the right panel of Figure 4, and that this transition occurs at $\rho_t \sim \frac{2\beta}{\pi} \ln \frac{\sqrt{\pi}}{2\sqrt{\beta}} R$ as $R \to \infty$.

The left panel of Figure 4 shows an encouraging comparison with a numerical Monte Carlo simulation. Nodes are randomly placed within a circular disc of area $A = 200$ and are connected with probability $H(\mathbf{r}_{ij})$ ($\beta = 1$). The error at lower densities is of course expected as the theory developed above is only a first order approximation to P_{fc}. Nevertheless, the agreement with simulations substantially improves at high densities but is also good in mid-densities. Hence, equation (29) provides a useful path which one may take to predict, control, or even optimize and set benchmarks for achieving full network connectivity.

18

3.3 Example 2. Square $(\eta = 2)$

We now consider a square domain s_L of side L. The integral in the exponent of (29) can be evaluated exactly in Cartesian coordinates for $\eta = 2$ to give

$$\int_{s_L} H(\mathbf{r}_{12})\mathrm{d}\mathbf{r}_1 = \int_{-\frac{L}{2}}^{\frac{L}{2}} \int_{-\frac{L}{2}}^{\frac{L}{2}} \left(e^{-\beta\left((x_1-x_2)^2+(y_1-y_2)^2\right)} \right) \mathrm{d}x_1\mathrm{d}y_1$$
$$= \frac{\pi}{4\beta} h(x_2)h(y_2), \tag{39}$$

where $h(x) = \left(\mathrm{erf}\left[\sqrt{\beta}\frac{(L-2x)}{2} \right] + \mathrm{erf}\left[\sqrt{\beta}\frac{(L+2x)}{2} \right] \right)$. Due to symmetry, we need only consider the positive quadrant of s_L *i.e.* when $x_2, y_2 \in [0, \frac{L}{2}]$. Hence $h(x) \approx \left(\mathrm{erf}\left[\sqrt{\beta}\frac{(L-2x)}{2} \right] + 1 \right)$ and may be further approximated by $\tilde{f}(x)$ with R replaced by $L/2$ (see equation (37)) such that

$$\int_0^{\frac{L}{2}} \int_0^{\frac{L}{2}} e^{-\rho\frac{\pi}{4\beta}h(x)h(y)}\mathrm{d}x\mathrm{d}y \approx \int_0^{\frac{L}{2}} \int_0^{\frac{L}{2}} e^{-\rho\frac{\pi}{4\beta}\tilde{f}(x)\tilde{f}(y)}\mathrm{d}x\mathrm{d}y. \tag{40}$$

Note however that it is crucial that the quadratic $x_2 y_2$ term is not included in the expansion of $h(x_2)h(y_2)$ as it is of higher order. Performing the integrals and multiplying by four we obtain a result

$$P_{fc} = 1 - L^2\rho e^{-\frac{\pi}{\beta}\rho} \left(1 - \mathcal{O}(L^{-1})\right) - 4L\sqrt{\frac{\beta}{\pi}}e^{-\frac{\pi}{2\beta}\rho} \left(1 - \mathcal{O}(L^{-1})\right) - \frac{16\beta}{\rho\pi}e^{-\frac{\pi}{4\beta}\rho}, \tag{41}$$

assuming that $\sqrt{\beta}\frac{L}{2} \gg 1$. Note that the above calculation can naturally be extended in a straight forward way to include for general $\eta > 0$ and higher dimensional $d > 0$ orthotopes (hyper-rectangles) of non-equal sides.

As expected, P_{fc} in equation (41) is composed of a bulk term, a boundary and a corner term with coefficients characteristic of the corresponding volume of the relevant object (*e.g.* area, perimeter, number of corners) and is dominated by the corner term in the high density limit. This is illustrated on the right panel of Figure 5 while the left panel of Figure 5 shows another encouraging comparison with numerical simulations. We also notice that there exists a parameter window $\rho_1 \leq \rho \leq \rho_2$ with $\rho_1 \sim \frac{2\beta}{\pi}\ln\frac{\sqrt{\pi}}{4\sqrt{\beta}}L$ and $\rho_2 \sim \frac{4\beta}{\pi}\ln\frac{\sqrt{\pi}}{4\sqrt{\beta}}L$, where the edges of the square provide the dominant contribution to P_{fc}. Note however that ρ_1 and ρ_2 are only rough estimates of the transition points as they require $L \to \infty$ but only depend logarithmically on L. Finally, the 16 in the last coefficient arises from the number of corners (4) multiplied by a factor. We turn to the calculation of this factor in more generality in the next subsection.
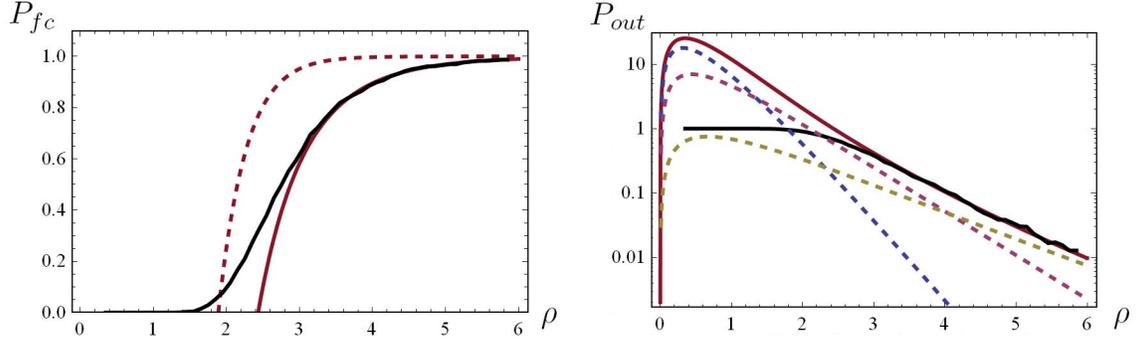
Figure 5: (Color online) *Left:* A plot of $P_{fc}(\rho)$ in the square domain s_L of area $A = 200$ ($\beta = 1$) (solid red) compared with the prediction of equation (8) (dashed) and a numerical simulation (black). *Right:* The dashed curves show the contributions from the bulk (blue), the boundary (purple) and the corner (yellow) term in equation (41) on a log-linear scale. The full curves are plots of $P_{out}(\rho)$ with dark red and black corresponding to theory and numerical simulation.

3.4   Example 3. Wedge: General angle ($\eta > 0$)

In order to understand corner effects better, we now consider for general $\eta > 0$ a wedge $w_\vartheta \subset \mathbb{R}^2$ of angle $\vartheta \in (0, \pi)$. We are concerned here only with the region close to the origin. At distances much larger than the connectivity range the wedge could be truncated to form a circular sector. We ignore the bulk, boundary and two right angled corner contributions to P_{fc}, and concentrate on the middle corner term which we denote by $\mathcal{C}_d(\vartheta)$ for $d = 2$. Note that this approach applies for general geometries containing non-right angles.

The distance between two nodes is now given by $d(\mathbf{r}_1, \mathbf{r}_2) = \sqrt{|\mathbf{r}_1|^2 + |\mathbf{r}_2|^2 - 2|\mathbf{r}_1||\mathbf{r}_2| \cos(\theta_1 - \theta_2)}$ with angular positions $\theta_{1,2} \in (0, \vartheta)$. Following the arguments of section 3.1, we expect that to leading order $\mathcal{C}_2(\vartheta)$ is when $\mathbf{r}_2$ is situated near the central corner *i.e.* when $r_2 \approx 0$. Hence we expand $H(\mathbf{r}_{12})$ about the corner in polar coordinates such that to leading order

$$
\begin{aligned}
\int_{w_\vartheta} H(\mathbf{r}_{12}) \mathrm{d}\mathbf{r}_1 &= \lim_{R \to \infty} \int_0^\vartheta \int_0^R r_1 e^{-\beta r_1^\eta} \left(1 + \eta \beta r_1^{\eta-1} r_2 \cos(\theta_1 - \theta_2) + \mathcal{O}(r_2^2) \right) \mathrm{d}r_1 \mathrm{d}\theta_1 \\
&= \frac{\vartheta \Gamma\left(\frac{2}{\eta}\right) + r_2 \Gamma\left(\frac{1}{\eta}\right) (\sin \theta_2 - \sin(\theta_2 - \vartheta))}{\eta \beta^{\frac{2}{\eta}}} + \mathcal{O}(r_2^2),
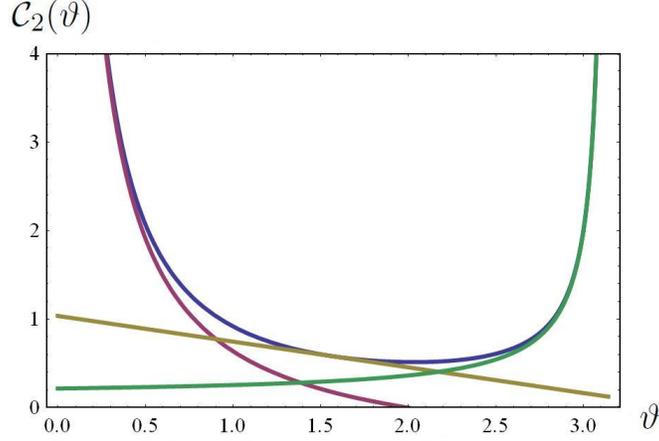\end{aligned}
\tag{42}
$$

Figure 6: (Color online) The contribution from a corner of angle ϑ (blue) for $\rho = \beta = 1$ and $\eta = 2$. Also shown are the three asymptotic regimes (see equation (44)) of $\vartheta \to 0, \frac{\pi}{2}, \pi$ shown in purple, yellow and green respectively.

and hence

$$
\begin{aligned}
\mathcal{C}_2(\vartheta) &= \lim_{R\to\infty} \int_0^\vartheta \rho \int_0^R r_2 e^{-\rho\left(\frac{\vartheta\Gamma\left(\frac{2}{\eta}\right)+r_2\Gamma\left(\frac{1}{\eta}\right)(\sin\theta_2-\sin(\theta_2-\vartheta))}{\eta\beta^{\frac{2}{\eta}}}+\mathcal{O}(r_2^2)\right)} \mathrm{d}r_2\mathrm{d}\theta_2 \\
&= \frac{\beta^{\frac{2}{\eta}} e^{-\rho\frac{\Gamma\left(\frac{2}{\eta}\right)}{\eta\beta^{\frac{2}{\eta}}}\vartheta}}{\rho\sin\vartheta\,\Gamma\left(1+\eta^{-1}\right)^2}.
\end{aligned}
$$

(43)

Reassuringly in the case of a square with $\eta = 2$ we have that $4\mathcal{C}_2(\frac{\pi}{2}) = \frac{16\beta}{\rho\pi} e^{-\frac{\pi}{4\beta}\rho}$ as in equation (41). Furthermore, from (43) we identify three main regimes as follows

$$
\mathcal{C}_2(\vartheta) = \begin{cases}
\frac{4\beta}{\pi\rho\vartheta} - \frac{2}{\pi} + \mathcal{O}(\vartheta), & \text{for} \quad \vartheta \ll 1, \\
\frac{4\beta}{\pi\rho} e^{\frac{-\pi\rho}{4\beta}} - \frac{2}{\pi} e^{\frac{-\pi\rho}{4\beta}}\left(\vartheta - \frac{\pi}{2}\right) + \mathcal{O}((\vartheta - \frac{\pi}{2})^2), & \text{for} \quad \vartheta \approx \frac{\pi}{2}, \\
\frac{4\beta e^{\frac{-\pi\rho}{2\beta}}}{\pi\rho(\pi-\vartheta)} + \frac{2e^{\frac{-\pi\rho}{2\beta}}}{\pi} + \mathcal{O}(\pi - \vartheta), & \text{for} \quad \pi - \vartheta \ll 1,
\end{cases}
$$

(44)

for the case of $\eta = 2$ and are illustrated in Figure 6. Finally, $\mathcal{C}_2(\vartheta)$ has a minimum at

$$
\vartheta_{min} = \pi - \arctan\frac{2\beta}{\rho}
$$

(45)

Physically, the picture shown in Figure 6 makes sense in the following way: the contribution to P_{fc} from a ϑ angled corner is maximum for $\vartheta \ll 1$ as expected due to the high

21

probability of an isolated node. This contribution decreases monotonically as ϑ is increased until $\vartheta > \vartheta_{min}$. At this point our approximations become invalid since the leading exponent now tends to $\sim -\frac{\pi\rho}{2\beta}$ as $\vartheta \to \pi$ which is the same as that of the perimeter term in equation (41), *i.e.* the corner turns into an edge of infinite length and thus $\mathcal{C}_2(\vartheta)$ diverges. Notice however that at very high densities $\vartheta_{min} \to \pi$.

3.5 A General formula

In high-dimensional domains $\mathcal{V} \subset \mathbb{R}^d$ ($d \geq 3$), boundary effects are due to intersections of hyperplanes which the above examples and techniques can generalize to account for. As a result P_{fc} can be expressed as a sum of contributions due to objects of different codimension $i = 0, 1, \ldots d$ (with $i = 0$ corresponding to the bulk/volume term)

$$P_{fc} \approx 1 - \sum_{i=0}^{d} \sum_{j_i} \rho^{1-i} \mathbf{G}_{j_i} \mathbf{V}_{j_i} e^{-\rho\omega_{j_i}\mathcal{H}}, \tag{46}$$

where $d \in \mathbb{N}$ is the space dimension, $\mathbf{G}_{j_i}$ is a geometrical factor for each object j of codimension i while $\mathbf{V}_{j_i}$ is the corresponding $d-i$ dimensional volume of the object with solid angle $\omega_{j_i} \in (0, \Omega)$ and $\mathcal{H} = \int_0^\infty r^{d-1} H(r) \mathrm{d}r$. It is clear that for $i = 0$ and 1, the second sum in (46) contains only one term. The remaining one dimensional integral is easily computed for typical connectivity functions H_{ij} (also see [39]). Corrections due to curved hyper-surfaces are expected to give algebraic corrections as in (38). Note however that $\mathbf{G}_{j_i}$ is H_{ij} dependent while $\mathbf{V}_{j_i}$ is not. In the case of $H(r) = e^{-\beta r^2}$ we find for example for d-dimensional hyper-rectangles where all hyperplanes meet at right angles $\mathbf{G}_{j_i} = 2^{i(i-1)} \left(\frac{\beta}{\pi}\right)^{\frac{i(d-1)}{2}}$ while for the unit-disk model with range r_0 we find $\mathbf{G}_{j_i} = 2^{i(i-1)}/V_{d-1}^i$, where $V_n = \frac{\pi^{n/2} r_0^n}{\Gamma(n/2+1)}$ is the volume of an n-dimensional sphere of radius r_0. These two cases are compared in Figure 7 for a cube of side $L = 7$.

The above formula is a high density expansion giving the probability of achieving a fully connected network of nodes in confined geometries. For all practical purposes ($d \leq 3$) equation (46) can account for most interesting (from an engineering point of view) domains in which a network may be deployed. Furthermore, its closed and simple format emphasizes the logical decomposition of the domain into objects of different full connectivity importance. This is illustrated in Figure 8 comparing two circular sectors ($d = 2$) of angles $\vartheta_1 = \pi/4$ and $\vartheta_2 = 3\pi/4$ but of equal perimeters $\mathbf{V}_{1_1} = 70$. Notice that the density at
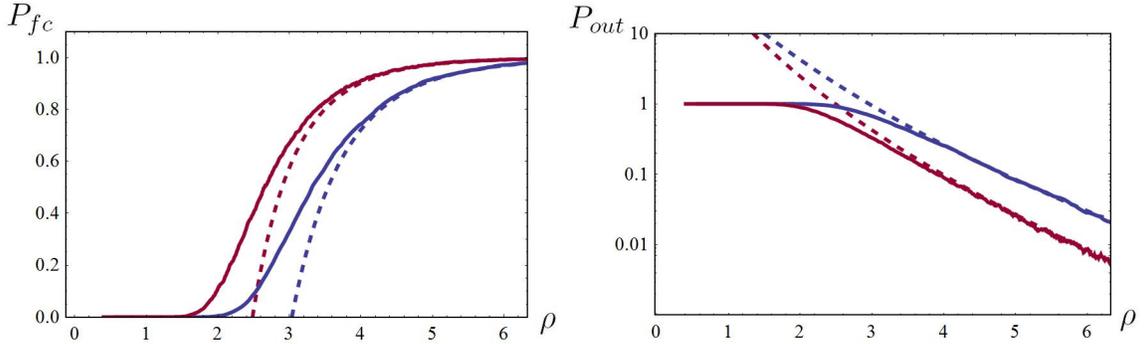
Figure 7: *Left:* (Color online) Comparison of $P_{fc}(\rho)$ from theoretical predictions of (46) (dashed curves) with direct numerical simulation (jagged line) of two random graphs inside a cube of side $L = 7$ for the unit-disk connectivity function with $r_0 = 1$ (blue) and the probabilistic Rayleigh fading link model with $\eta = 2$ (red). *Right:* The corresponding $P_{out}(\rho)$ is plotted on a log-linear scale emphasizing the good agreement between theory and simulations as well as the difference between the two models.

which the numerical simulation results become distinct ($\rho \approx 3$) in the left panel of Figure 8 is also where the corner term of c_{R_1, ϑ_1} overtakes the common perimeter term (shown in black) in the right panel of Figure 8.

It is useful to compare the various boundary contributions for a given geometry. For example, in the case of the square, equation (41), the terms for bulk, edges and corners differ in their exponential, so that in the limit of high density, they could only be comparable if the size of the system (L or V) is growing exponentially with density. Setting the corner and edge terms equal, we find for example

$$V = \frac{16\beta}{\pi \rho^2} e^{\frac{\pi \rho}{2\beta}} \tag{47}$$

which quickly becomes very large. At smaller volumes for a given density, the corners dominate the full connection probability.

4   Conclusion

In this paper we have attempted to understand how boundary effects can influence full connectivity in confined geometries. Based on the fact that at very high densities the probability of full connectivity P_{fc} is simply the complement of the probability of a single isolated node, we have developed a cluster expansion approach which classifies terms with
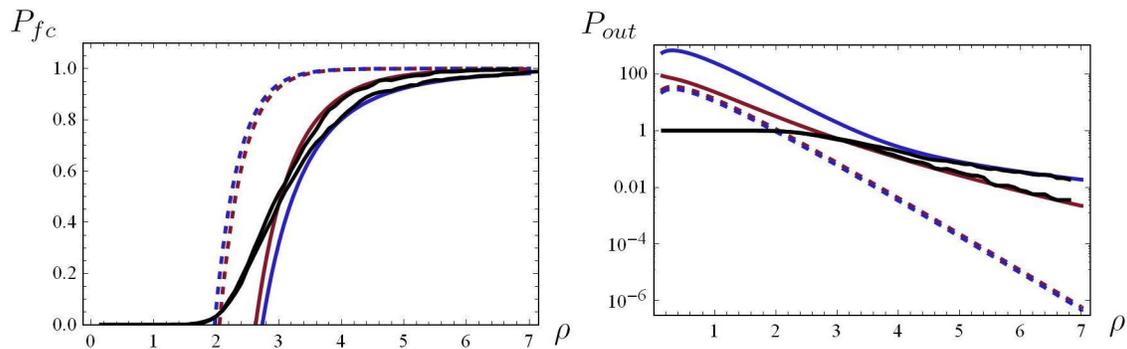
23

Figure 8: (Color online) *Left:* Comparison of $P_{fc}(\rho)$ from theoretical predictions of (46) (solid curves) with direct numerical simulation (black) of two random graphs inside circular sectors c_{R_1,ϑ_1} (blue) and c_{R_2,ϑ_2} (red) using $\beta = 1$ and $\eta = 2$. The sectors are of equal perimeters $\mathbf{V}_{1_1} = 70$ but with $\vartheta_1 = \pi/4$ and $\vartheta_2 = 3\pi/4$. The dashed curves correspond to conventional wisdom (*i.e.* the bulk contribution of equation (8)). *Right:* The corresponding $P_{out}(\rho)$ is plotted on a log-linear scale emphasizing the good agreement between theory and simulations.

respect to their largest connected component. Assuming a homogeneous system and averaging over the available space of the domain, we extract the leading and next-to-leading order behavior of P_{fc} as is given in equations (8) and (17) respectively. Although our approach can be consistently expanded to obtain even higher order corrections to P_{fc}, using an exponentially decaying connectivity test-function[4] (equation (19)), we showed that correction terms are exponentially small. This is because nodes are 'connected in probability' (in this case exponentially) rather than in an all-or-nothing way.

The first main result of this paper follows from lifting the homogeneity assumption to obtain equations (29) and (30), which confirm our initial observation, namely, that in the high density limit, P_{fc} is dominated by nodes which are situated in "*hard to connect*" regions, characterized by their solid angle. Using (19) we then instructively worked through three example domains (disk, square and wedge) in order to reach our second main result, a general formula for P_{fc} (equation (46)). The closed and simple format of (46) emphasizes the logical decomposition of the domain into objects of different full connectivity importance, according to their codimension. Finally, a comparison of our theoretical predictions with direct numerical simulations, in a variety of settings, has demonstrated that conventional percolation models are not sufficient for an accurate description of full network

---

[4]This was taken from a standard Rayleigh fading link model (also see [39]).

connectivity at high node densities.

In summary, full connectivity at high node densities is dominated by local geometric boundary effects such as corners, edges and faces. These contributions have universal properties, distinct but complementary to those of previous percolation approaches, which are typically limited to bulk effects or use a scaling that removes boundary terms (see [35]). As a result, our findings bring forward a radically different understanding of connectivity in confined geometries while also providing useful formulas for the probability of full connectivity, crucial for the design of reliable wireless mesh networks [40]. Moreover, sufficient quantitative detail is available for analyzing and determining system parameters in order to mitigate boundary effects (see [39]). Conversely, boundary effects can be harnessed to avoid full connectivity. Such approaches can be useful in physically important models such as the spread of forest fires [18], epidemics [21], or mobile phone viruses [23], given details of a specific model for connectivity. Finally, one may also consider the inverse problem of "connecting the shape of a drum" [41] to characterize unknown domains containing random networks. More generally, boundary effects are of major interest in the statistical mechanics of systems with very small structures such as in conducting carbon nanotubes [42] or very large and highly connected social and financial networks [26, 27] where the geometry is likely to be dynamic.

Acknowledgments

References

[1] B. Bollobás and O. Riordan, *Percolation*, Cambridge University Press, (2006).

[2] S.R. Broadbent and J.M. Hammersley, Percolation processes, Proc. Cambridge Phil. Soc. **5**, 629, (1957).

[3] P. Erdős. and A. Rényi, On Random Graphs I., in Publ. Math. Debrecen **6**, 290, (1959).

[4] B. Bollobás, *Random graphs*, Cambridge University Press, (2001).

[5] A. Coniglio, U.De. Angelis and A. Forlani, Pair connectedbess and cluster size, J. Phys. A: Math. Gen., **10** 1123, (1977).

[6] G. Stell, Continueem theory of percolation, J. Phys: Condens. Matter, **8** A1, (1996).

[7] T. L. Hill, *Statistical Mechanics*, McGraw-Hill, (1956).

[8] T. DeSimone, S. Demoulini and R. M. Stratt, A theory of percolation in liquids, J. Chem. Phys. **85**, 391, (1986).

[9] T. L. Hill, Molecular clusters in imperfect gases, J. Chem. Phys., **23**, 617, (1955).

[10] Y. Chiew and E. Glandt, Percolation behaviour of permeable and of adhesive spheres, J. Phys. A: Math. Gen., **16**, 2599, (1983).

[11] S. A. Safran, I. Webman, and G.S. Grest, Percolation in interacting colloids, Phys. Rev. A **32**, 506, (1985).

[12] L. Hu, D. S. Hecht, and G. Grüner, Percolation in Transparent and Conducting Carbon Nanotube Networks, Nano Lett., **4** (12), 2513, (2004).

[13] D. Toker, D. Azulay, N. Shimoni, I. Balberg, and O. Millo, Tunneling and percolation in metal-insulator composite materials, Phys. Rev. B **68**, 041403(R), (2003).

[14] D.-H. Lee, Z. Wang, and S. Kivelson, Quantum percolation and plateau transitions in the quantum Hall effect, Phys. Rev. Lett. **70**, 4130, (1993).

[15] A. Kaminski and S. Das Sarma , Polaron Percolation in Diluted Magnetic Semiconductors, Phys. Rev. Lett. **88**, 247202, (2002).

[16] J. Kim, P.L. Krapivsky, B. Kahng and S. Redner, Infinite-order percolation and giant fluctuations in a protein interaction network, Phys. Rev. E **66**, 055101(R), (2002).

[17] B. Drossel and F. Schwabl, Self-organized critical forest-fire model, Phys. Rev. Lett. **69**, 1629, (1992).

[18] S. Pueyo, P.M.L. de Alencastro Graca, R.I. Barbosa, R. Cots, E. Cardona and P.M. Fearnside, Testing for criticality in ecosystem dynamics: The case of Amazonian rainforest and savanna fire, Ecol. Lett. 13, 793, (2010).

[19] C. Moore and M.E.J. Newman, Epidemics and percolation in small-world networks, Phys. Rev. E, **61**, 5678. (2000).

[20] J.C. Miller, Spread of infectious disease through clustered populations, J. Roy. Soc. Interf. **6**, 1121, (2009).

[21] L. Danon *et al*, Networks and the epidemiology of infectious disease, Interdisc. Persp. Infect. Diseases 2011, 284909, (2011).

[22] R. Cohen, S. Havlin and D. ben-Avraham, Efficient Immunization Strategies for Computer Networks and Populations, Phys. Rev. Lett. **91**, 247901, (2003).

[23] P.Wang, M.C. González, C.A. Hidalgo and A.L. Barabási, Understanding the spreading patterns of mobile phone viruses, Science, **324**, 1071, (2009).

[24] D. Duffie and G. Manso, Information percolation in large markets, Amer. Econ. Rev., **97**, 203, (2007).

[25] G. Palla, I. Derényi, I. Farkas, and T. Vicsek, Uncovering the overlapping community structure of complex networks in nature and society, Nature, **435**, 814, (2005).

[26] G. Palla, A.-L. Barabási and T. Vicsek, Quantifying social group evolution, Nature, **446**, 664, (2007).

[27] R. Parshani, S. Buldyrev and S. Havlin, Critical effect of dependency groups on the function of networks, Proc. Natl. Acad. Sci. **108**, 1007, (2011).

[28] M. Penrose, *Random Geometric Graphs*, Oxford University Press, (2003).

[29] I. Glauche, W. Krause, R. Sollacher, and M. Greiner, Continuum percolation of wireless ad hoc communication networks, Physica A., **325**, 577, (2003).

[30] G. Paschos, P. Mannersalo, and S. Stanczak, Extending the percolation threshold using power control, IEEE WCNC, 1, (2009).

[31] D. Callaway, M. Newman, S. Strogatz, and D. Watts, Network robustness and fragility: Percolation on random graphs, Phys. Rev. Lett., **85**, 5468, (2000).

[32] J. Li, L. Andrew, C. Foh, M. Zukerman and H. Chen, Connectivity, coverage and placement in wireless sensor networks, Sensors, **9**, 7664, (2009).

[33] H. Ammari and S. Das, Critical density for coverage and connectivity in three-dimensional wireless sensor networks using continuum percolation, IEEE Trans. Parallel Distrib. Syst., **20**, 872, (2008).

[34] M. Franceschetti, O. Dousse, D. Tse, and P. Thiran, Closing the gap in the capacity of wireless networks via percolation theory, IEEE Trans. Inf. Theory, **53**, 1009, (2007).

[35] G. Mao and B. Anderson, Towards a better understanding of large scale network models, IEEE/ACM Trans. Netw, **20**, 408, (2012).

[36] M. Penrose, On k-connectivity for a geometric random graph, Random Structures & Algorithms, v.15 n.2, 145-164, (1999).

[37] J.P. Coon, C.P. Dettmann and O. Georgiou, Impact of Boundaries on Fully Connected Random Geometric Networks, Phys. Rev. E., **85**, 011138, (2012).

[38] D. Tse and P. Viswanath, *Fundamentals of wireless communication*, Cambridge University Press, (2005).

[39] J.P. Coon, C.P. Dettmann and O. Georgiou, Connectivity of Confined Dense Networks: Boundary Effects and Scaling Laws, preprint arXiv:1201.4013, (2012).

[40] M. Haenggi, J. G. Andrews, F. Baccelli, O. Dousse and M. Franceschetti, Stochastic geometry and random graphs for the design of wireless networks, IEEE J. Select. Area. Commun. **27**, 1029, (2009).

[41] M. Kac, Can one hear the shape of a drum?, Am. Math. Mon, **73**, 1, (1966).

[42] A. Kyrylyuk *et al*, Controlling electrical percolation in multicomponent carbon nanotube dispersions, Nature Nanotech. **6**, 364, (2011).