Impact of boundaries on fully connected random geometric networks

Justin Coon,[1] Carl P. Dettmann,[2] 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 Strasse 38, D-01187 Dresden, Germany*
(Received 21 October 2011; revised manuscript received 16 November 2011; published 20 January 2012)

Many complex networks exhibit a percolation transition involving a macroscopic connected component, with universal features largely independent of the microscopic model and the macroscopic domain geometry. In contrast, we show that the transition to full connectivity is strongly influenced by details of the boundary, but observe an alternative form of universality. Our approach correctly distinguishes connectivity properties of networks in domains with equal bulk contributions. It also facilitates system design to promote or avoid full connectivity for diverse geometries in arbitrary dimension.

DOI: 10.1103/PhysRevE.85.011138 PACS number(s): 05.70.Np, 89.75.Hc, 84.40.Ua

I. INTRODUCTION

Random geometric network models [1,2] comprise a collection of entities called nodes embedded in a region of typically two or three dimensions, together with connecting links between pairs of nodes that exist with a probability related to the node locations. They appear in numerous complex systems including in nanoscience [3], epidemiology [4,5], forest fires [6], social networks [7,8], and wireless communications [9–11]. Such networks exhibit a general phenomenon called *percolation* [12,13], where at a critical connection probability (controlled by the node density), the largest connected component (cluster) of the network jumps abruptly from being independent of system size (microscopic) to being proportional to system size (macroscopic).

Percolation phenomena are closely related to thermodynamic phase transitions where the number of nodes N goes to infinity and the critical percolation density ρ_c is largely independent of the system size, shape, and of the microscopic details of the model; the phenomenon of universality. At the critical point, conformal invariance in two-dimensional networks leads to detailed expressions for the probability of a connection across general regions [14] and more general connections with conformal field theory [15] and Schramm-Loewner evolution [16]. Here, we take a different approach and are concerned with finite networks and with questions related to percolation, but fundamentally different: What node density ensures a specified probability P_{fc} that the entire network is a single connected component (cluster), that is, *fully connected*? How is this probability affected by the shape of the network domain?

These questions are crucial for many applications, including, for example, the design of reliable wireless mesh networks. These consist of communication devices (the nodes) that pass messages to each other via other nodes rather than a central router. This allows the network to operate seamlessly over a large area, even when nodes are moved or deactivated. A fully connected network means that every node can communicate with every other node through direct or indirect connections. Mesh networks have been developed for many communication systems, including laptops, power distribution ("smart grid") technologies, vehicles for road safety or environmental monitoring, and robots in hazardous locations such as factories, mines, and disaster areas [10].

For many applications of random geometric networks including those above, the *direct connection* between two nodes i and j can be well described by a probability $H_{ij} = H(r_{ij})$, a given function of the distance between the nodes $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$. Often, the nodes are mobile or otherwise not located in advance, hence we assume N uniformly distributed nodes confined in a specified d-dimensional region $\mathcal{V}$ with area ($d = 2$) or volume ($d = 3$) denoted by V. The node density is then defined as $\rho = N/V$. For reference, we will later take $H(r_{ij}) = \exp[-(r_{ij}/r_0)^\eta]$, where r_0 is a relevant length scale, and η determines the sharpness of the cutoff. Note that when $\eta \to \infty$ a step function corresponding to the popular *unit disk* deterministic model [17] is obtained, where connections have a fixed range r_0. Our derivation, however, is completely general and only requires that H_{ij} is sufficiently short-ranged compared to system size. Using this as a basis, we find that, contrary to common belief and practice, the geometrical details of the confined space boundaries (*corners, edges*, and *faces*) dominate the properties of the percolation transition. Moreover, the short-range nature of H_{ij} allows us to separate individual boundary components and obtain analytic expressions for P_{fc} at high densities as a sum over their contributions. We confirm this through computer simulations and argue that the substantial improvement offered by our main result Eq. (7) can be used to predict, control, optimize, or even set benchmarks for achieving full network connectivity in a wide variety of suitable models and applications involving finite size geometries.

II. FULL CONNECTION PROBABILITY

As in conventional continuum percolation theory [18], we start by utilizing a cluster expansion approach [19] to derive a systematic perturbative method for determining the full connection probability P_{fc} as a function of density ρ. Formulation of the expansion can be summarized as follows. The probability of two nodes being connected (or not) leads to the trivial identity $1 \equiv H_{ij} + (1 - H_{ij})$. Multiplying over all links expresses the probabilities $\mathcal{H}_g$ of all $2^{N(N-1)/2}$ possible graphs g,

$$1 = \prod_{i<j} [H_{ij} + (1 - H_{ij})] \equiv \sum_g \mathcal{H}_g. \tag{1}$$

Collecting the terms according to largest cluster size we get

$$1 = \sum_{g\in\mathcal{G}_N} \mathcal{H}_g + \sum_{g\in\mathcal{G}_{N-1}} \mathcal{H}_g + \ldots + \sum_{g\in\mathcal{G}_1} \mathcal{H}_g, \tag{2}$$

where $\mathcal{G}_n$ is the set of all possible graphs with the largest cluster of size $n \in \{1,\ldots,N\}$. The first term on the right-hand side is the probability of being fully connected given a specific configuration of nodes. The average over all random configurations $\langle\rangle \equiv V^{-N}\int_{\mathcal{V}} d^N\mathbf{r}$ of this quantity is thus the overall probability of obtaining a fully connected network P_{fc}. Moreover, the main idea conveyed by Eq. (2) is that at high densities, full connectivity is most likely to be broken by a single isolated node (the $\mathcal{G}_{N-1}$ term); this is sufficient detail for most applications. Further corrections incorporate the probability of several isolated single nodes and smaller clusters of nodes, for which a systematic expansion can be developed [20].

Averaging Eq. (2) over all configurations and noting that to leading order the $N-1$ cluster is fully connected, and that all nodes are identical, the first-order approximation becomes

$$\begin{aligned} P_{fc} &\approx 1 - \left\langle \sum_{g\in\mathcal{G}_{N-1}} \mathcal{H}_g \right\rangle \\ &= 1 - N\left\langle \prod_{j=2}^{N}(1-H_{1j}) \right\rangle \\ &= 1 - \rho\int_{\mathcal{V}} \left(1 - \frac{M(\mathbf{r}_1)}{V}\right)^{N-1} d\mathbf{r}_1, \end{aligned} \tag{3}$$

where the "connectivity mass" accessible from a node placed at $\mathbf{r}_1$ is given by

$$M(\mathbf{r}_1) = \int_{\mathcal{V}} H(r_{12}) d\mathbf{r}_2. \tag{4}$$

Assuming that the volume $V \gg \rho M(\mathbf{r}_1)^2$ for any $\mathbf{r}_1$, which is reasonable if the system is significantly larger than r_0 at moderate densities and that the number of nodes N is large, Eq. (3) simplifies to

$$P_{fc} \approx 1 - \rho\int_{\mathcal{V}} e^{-\rho M(\mathbf{r}_1)} d\mathbf{r}_1. \tag{5}$$

This equation is equivalent to Eq. (8) in Mao and Anderson [21], which was derived for the specific case of a square domain. Following numerous studies by probabilists and engineers [1,2], these authors, however, assumed an exponential scaling of system size V with ρ which essentially renders boundary effects negligible. Scaling the system in such a way is a common approach as it corresponds to the limit of infinite density at fixed connection probability, however, in practice this limit is approached only for unphysically large volumes. In contrast, we do not assume the exponential growth of V, and also consider far more general geometries in any dimension $d \geqslant 1$.

Without an exponentially growing volume V, the behavior of the full connection probability at high densities is qualitatively different: It is controlled by the exponential in Eq. (5), and hence node positions $\mathbf{r}_1$ where the connectivity mass is small, that is, near the boundary of $\mathcal{V}$. Thus in contrast to the usual situation in statistical mechanics, the boundaries (and in particular corners) are important, and we will see they in fact dominate. We illustrate this in Fig. 1 where nodes are placed randomly inside a cube and an average over a large number of possible graphs gives the connectivity of each node. Notice that isolated and hard-to-connect nodes shown as dark balls concentrate near the boundaries of the domain and particularly near corners at higher densities. This observation forms the basis of our work, and has led to a radically different understanding of connectivity in confined geometries which we now detail further.

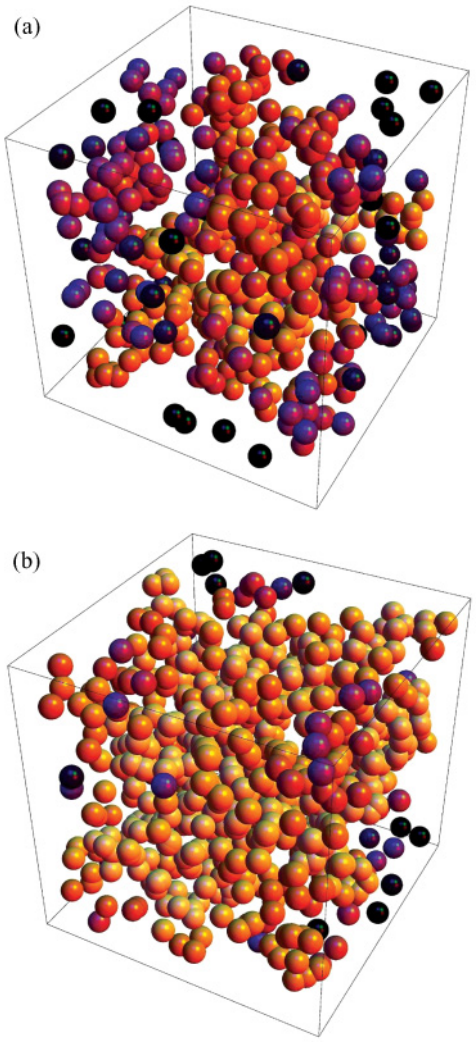

FIG. 1. (Color online) Isolated nodes shown as black balls concentrate at the boundaries of the domain and particularly near corners at higher densities. Nodes are placed randomly in a cube, with lighter colors indicating a higher probability of being in the largest connected component. We use $\eta = 2$, while the side length of the cube is $L = 10r_0$. There are 500 nodes in (a) and 700 nodes in (b).

III. BOUNDARY EFFECTS

The contributions to the integrals in Eq. (5) come from $\mathbf{r}_1$ at *boundary components* $B \subset \mathcal{V}$ of dimension d_B, for example, the bulk, the faces, and right angled edges and corners of a cube, with $d_B = 3,2,1$, and 0, respectively. The short-range nature of H_{ij} allows us to isolate each boundary

component, while to leading order the connectivity mass splits into independent radial and angular integrals, depending only on the local geometry of B and hence

$$M_B = M(\mathbf{r}_B) = \omega_B \int_0^\infty H(r) r^{d-1} dr, \tag{6}$$

where ω_B is the angle ($d = 2$) or solid angle ($d = 3$) subtended by B. For example, if $\mathbf{r}_B$ is near a corner of the cube then $\omega_B = (4\pi)/8$, while near an edge $\omega_B = (4\pi)/4$, near faces $\omega_B = (4\pi)/2$ and $\omega_B = (4\pi)$ for the bulk interior. Hence, from Eq. (5) we see that corner contributions to P_{fc} as a function of ρ are exponentially larger than edge contributions which are themselves exponentially larger than face contributions and so on. This simple argument shows that the dominant contribution to P_{fc} at high densities comes from the "pointiest" corners.

Expanding $H(r_{12})$ about $\mathbf{r}_2$ near the corresponding boundary component we obtain a next-to-leading-order expansion for $M(\mathbf{r}_B)$ which we can then use to approximately evaluate the integral in Eq. (5). Ignoring exponentially smaller correction terms and combining all boundary contributions we arrive at our main result

$$P_{fc} \approx 1 - \rho \sum_B G_B V_B e^{-\rho M_B}, \tag{7}$$

where V_B is the d_B-dimensional "volume" of each component (equal to 1 in the case of a zero-dimensional corner and V when $d_B = d$), G_B is a geometrical factor depending on B and implicitly on H and M_B is as in Eq. (6); see the examples below. Notice that Eq. (7) is completely general as we have only assumed a short-ranged H_{ij} and not used its specific form. Moreover, it also does not depend on using Euclidean distance and holds in any dimension $d \geqslant 1$ and geometry where the lack of connectivity is dominated by a situation involving an $N - 1$ cluster and a single disconnected node. Hence Eq. (7) is a powerful and useful multipurpose tool for analyzing full network connectivity at high densities in a wide variety of suitable models and applications involving finite size geometries.

For example, in the context of single input single output (SISO) wireless communication channels and a Rayleigh fading model [22], information theory predicts $H(r_{ij}) = \exp[-(r_{ij}/r_0)^\eta]$ with η an environment and wavelength-dependent decay parameter equal to 2 for free propagation, increasing to $\eta \approx 4$ for a cluttered environment, while r_0 depends on the minimum outage rate threshold. For nodes confined to a cube of side length L and $\eta = 2$ we find $V_B = L^{d_B}$, $G_B = (2^{3-d_B-1}/\pi\rho r_0^2)^{3-d_B}$, and $M_B = (r_0\sqrt{\pi})^3 2^{d_B-3}$ with contributions from each of the 8 corners, 12 edges, 6 faces, and bulk. However, the derivation is general: Once G_B and M_B have been evaluated for these boundary components (right angled edges, etc.) by standard asymptotic analysis of the relevant integrals, they apply to any geometry with these features and length scales significantly larger than r_0. This independence on the large-scale geometry follows from the short-range nature of H_{ij} and is a type of universality allowing for the calculation of P_{fc} in complex high-dimensional geometries without increased difficulty.

The substantial improvement offered by Eq. (7) becomes clear when compared with the "bulk" contribution corresponding to the current conventional wisdom shown in Fig. 2(a) for a network confined to a cube. Figure 2(b) further demonstrates the inaccuracy of the bulk model as well as the benefits of including boundary effects when analyzing network connectivity in confined geometries.

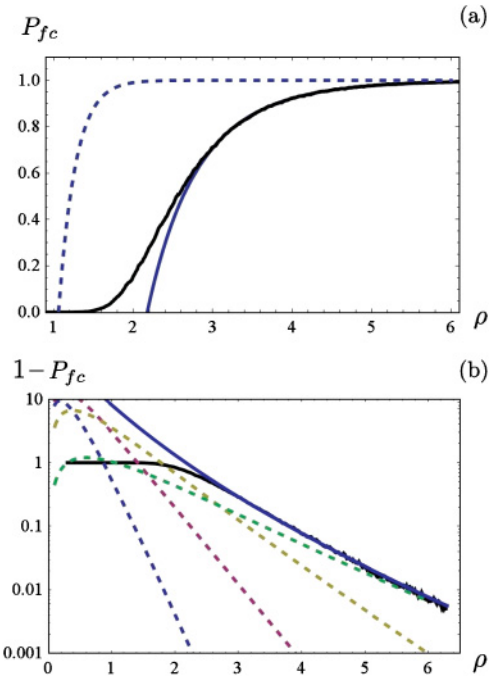

FIG. 2. (Color online) (a) Comparison of the full analytic prediction of Eq. (7) (solid curve) with direct numerical simulation of the random network in a cube of side length $7r_0$ (jagged curve). The dashed line corresponds to the bulk contribution (previous theory). (b) Contributions from the bulk (dotted blue line, left), faces (red line), edges (yellow line), and corners (green line, right), together with the total (solid blue line) and numerical simulation (black jagged curve), showing the dominance of the corners at the highest densities and good agreement between the theory and simulation at moderate to high densities. Here it is convenient to plot the outage probability $P_{\text{out}} = 1 - P_{fc}$.

We can go beyond simple geometries restricted to right-angled corners. Consider the case of a two-dimensional triangle with general angles $0 < \omega_B < \pi$. The relevant integrals for this case come to $M_B = r_0^2\omega_B/2$, with $G_B = 4/\pi\rho^2 r_0^2 \sin\omega_B$ for the corners and $G_B = (2^{2-d_B-1}/\pi\rho r_0^2)^{2-d_B}$ for the edges and bulk and can be generalized easily to higher dimensions. Figure 3 shows two triangles chosen to have identical perimeters and areas; the connectivity at a given density differs only due to the corner angles and agrees perfectly with the full theory of Eq. (7). A bulk theory, even supplemented with edge contributions, is clearly incapable of explaining the difference between the connectivities of networks in these two triangles. Moreover, such a situation motivates inverse problems, similar to "hearing the shape of a drum" [23] by attempting to determine the size and shape details of an unknown domain containing a random network.

IV. DISCUSSION

An important aspect of the theory presented here is how it affects the design of real life random geometric networks.

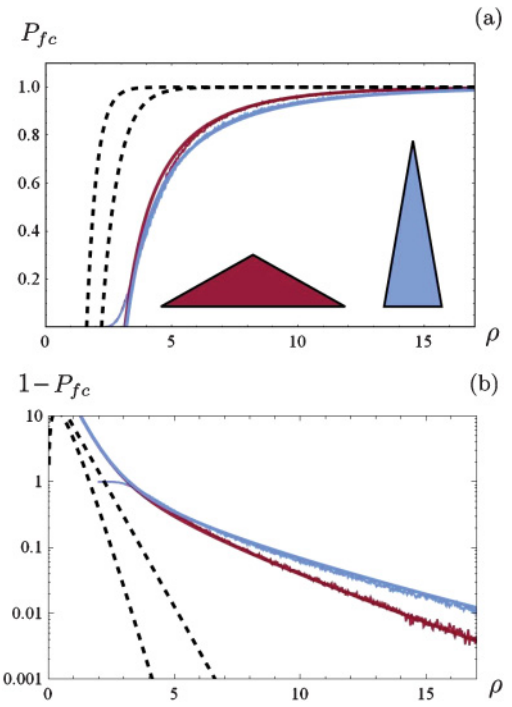

FIG. 3. (Color online) Corner contributions in triangles with equal area and perimeter: Comparison of theory with direct simulation, as in Fig. 2. The red triangle has side lengths of 26.88, 15.44, and 15.44 in units of the connectivity length scale r_0, while the blue triangle has side lengths of 8.40, 24.68, and 24.68. The black dashed lines correspond to the equal bulk (left curve) and bulk + edge (right curve) contributions while neglecting corner contributions. The colored curves give the total (including crucial corner) contributions for each triangle. Both theory and simulation are plotted, showing excellent agreement with the numerical simulations (jagged curves) which cover them completely for $\rho > 4$.

For wireless mesh networks, the lack of connectivity near the boundaries can be mitigated by increasing the signal power, the number of spatial channels, or by constructing a hybrid network with a regular array of fixed nodes along the boundaries as well as randomly placed nodes in the interior. In each of these cases, the design can now be analyzed given information about the cost and connectivity function $H(r)$ and of course the desired connectivity region. Conversely, boundary effects can be harnessed to avoid full connectivity if desired. For example, in the case of forest fires [6] we have a prediction for the number of unburnt regions as a function of the geometric landscape and environment parameters (for example, angles between fire lanes and/or natural boundaries), again given a specific model for connectivity that depends on the type of vegetation, temperature, moisture content, and so on. Similar models could be devised for the spread of epidemics [4] or mobile phone viruses [11] where boundaries are embedded in a more complex (possibly non-Euclidean) space yet H_{ij} is still short-ranged.

We examined connectivity in confined geometries and illustrated the importance of the often neglected boundary effects. We then derived a general high density expansion Eq. (7) for the probability of full connectivity P_{fc} assuming only a short-ranged connectivity function relative to system size and showed that it displays universal features allowing for its easy calculation in complex geometries. This we have confirmed through computer simulations and argued that our approach is well placed to facilitate efficiency in design in a variety of physical applications ranging from wireless networks to forest fire-lanes. Appropriate modifications of our theory can aid the understanding of small boundary-dominated systems such as, for example, the electrical conduction through carbon nanotubes in a polymer matrix [3], but possibly larger systems such as highly connected social and financial networks [7,8].

ACKNOWLEDGMENTS

The authors thank the directors of the Toshiba Telecommunications Research Laboratory for their support, and Charo Del Genio, Jon Keating, and Mark Walters for helpful discussions.

[1] M. Penrose, *Random Geometric Graphs* (Oxford University Press, New York, 2003).
[2] M. Franceschetti and R. Meester, *Random Networks for Communication* (Cambridge University Press, Cambridge, England, 2007).
[3] A. V. Kyrylyuk, M. C. Hermant, T. Schilling, B. Klumperman, C. E. Koning, and P. van der Schoot, Nature Nanotech. **6**, 364 (2011).
[4] J. C. Miller, J. Roy. Soc. Interf. **6**, 1121 (2009).
[5] L. Danon, A. P. Ford, T. House, C. P. Jewell, M. J. Keeling, G. O. Roberts, J. V. Ross, and M. C. Vernon, Interdisc. Persp. Infect. Diseases **2011**, 284909 (2011).
[6] S. Pueyo, P. M. L. D. A. Graça, R. I. Barbosa, R. Cots, E. Cardona, and P. M. Fearnside, Ecol. Lett. **13**, 793 (2010).
[7] G. Palla, A.-L. Barabási, and T. Vicsek, Nature (London) **446**, 664 (2007).
[8] R. Parshani, S. Buldyrev, and S. Havlin, Proc. Natl. Acad. Sci. **108**, 1007 (2011).
[9] M. Haenggi, J. G. Andrews, F. Baccelli, O. Dousse, and M. Franceschetti, IEEE J. Select. Area. Commun. **27**, 1029 (2009).
[10] J. Li, L. Andrew, C. Foh, M. Zukerman, and H. Chen, Sensors **9**, 7664 (2009).
[11] P. Wang, M. C. González, C. A. Hidalgo, and A. L. Barabási, Science **324**, 1071 (2009).
[12] D. S. Callaway, M. E. J. Newman, S. H. Strogatz, and D. J. Watts, Phys. Rev. Lett. **85**, 5468 (2000).
[13] B. Bollobás and O. Riordan, *Percolation* (Cambridge University Press, Cambridge, England, 2006).
[14] J. L. Cardy, J. Phys. A: Math. Gen. **25**, L201 (1992).

[15] J. Fuchs, I. Runkel, and C. Schweigert, J. Math. Phys. **51**, 015210 (2010).
[16] Y. Saint-Aubin, P. A. Pearce, and J. Rasmussen, J. Stat. Mech. (2009) P02028.
[17] S. Durocher, K. R. Jampani, A. Lubiw, and L. Narayanan, Comput. Geom.: Theor. Appl. **44**, 286 (2011).
[18] G. Stell, J. Phys. Condens. Matter **8**, A1 (1996).
[19] T. L. Hill, *Statistical Mechanics* (McGraw-Hill, New York, 1956).
[20] J. P. Coon, C. P. Dettmann, and O. Georgiou, e-print arXiv:1201.3123.
[21] G. Mao and B. Anderson, in *Proceedings of the IEEE Conference on Computer Communications, (Infocom 2011), Shanghai, China* (IEEE, New York, 2011), p. 631.
[22] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication* (Cambridge University Press, Cambridge, England, 2005).
[23] M. Kac, Amer. Math. Mon. **73**, 1 (1966).