Spatially embedded random networks

L. Barnett* and E. Di Paolo†
Centre for Computational Neuroscience and Robotics, Department of Informatics, School of Science and Technology, University of Sussex, Falmer, Brighton BN1 9QH, United Kingdom

S. Bullock‡
School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, United Kingdom
(Received 31 July 2007; published 20 November 2007)

Many real-world networks analyzed in modern network theory have a natural spatial element; e.g., the Internet, social networks, neural networks, etc. Yet, aside from a comparatively small number of somewhat specialized and domain-specific studies, the spatial element is mostly ignored and, in particular, its relation to network structure disregarded. In this paper we introduce a model framework to analyze the mediation of network structure by spatial embedding; specifically, we model connectivity as dependent on the distance between network nodes. Our spatially embedded random networks construction is not primarily intended as an accurate model of any specific class of real-world networks, but rather to gain intuition for the effects of spatial embedding on network structure; nevertheless we are able to demonstrate, in a quite general setting, some constraints of spatial embedding on connectivity such as the effects of spatial symmetry, conditions for scale free degree distributions and the existence of small-world spatial networks. We also derive some standard structural statistics for spatially embedded networks and illustrate the application of our model framework with concrete examples.

DOI: 10.1103/PhysRevE.76.056115 PACS number(s): 89.75.Hc, 05.10.Ln, 64.60.Ak, 89.75.Da

I. INTRODUCTION

Within the last decade or so advances in computational power have revealed two distinctive characteristics of real-world networks that are clearly not explicable in terms of the standard (Erdös-Rényi) random graph model [1]: namely, the *small world* effect [2], where shortest connected paths between network nodes appear to be surprisingly small in the context of network size (and clustering) and *scale free* degree distribution [3], where the number of connections to a randomly selected node follows a power law distribution. The seminal studies [2,3] supplied putative mechanisms for modeling these seemingly ubiquitous properties and subsequently initiated an explosion of interest and research in the field of complex networks.

Many of the networks studied with respect to these twin paradigms have a *spatial* element, be it that network nodes (and frequently connections) reside in a "real" (Euclidean) space, or at least in some abstract space with a natural notion of "distance" between nodes. A common feature of many of these networks is that connectivity will be in some way related to spatial embedding. For instance, communications networks may involve range-dependent links; social networks may involve distance-limited interactions between agents existing in some (possibly abstract) space; transport networks have an obvious spatial embedding while, more generally, technological/commercial networks frequently feature some cost-per-distance constraint on connectivity. We remark that, apart from the influence of spatial embedding on network *structure*, we might also expect space to have important consequences for the dynamics of *processes* associated with networks (e.g., the flow of information across a communication networks); although we do not address this issue here, we bear it in mind as a cogent motivation for the study of spatial embedding and as a worthwhile direction for future study.

In the context of the myriad mechanisms in the burgeoning complex networks literature purporting to account for various network structural aspects, the spatial element has, arguably, been somewhat neglected. Existing studies of spatially embedded networks have tended to be domain specific and aimed at modeling (more or less realistically) some spatial aspect of network formation [4–12] or, at least, to address somewhat restrictive spatial embeddings [13–16]. It thus seems timely to inquire on a more explicit and inclusive level into the constraints on network structure implied by spatial embedding.

Our starting point in this study is that there exists some relationship between internode spatial distance and connection probability; more specifically, the model presented here assumes that connectivity is *purely* a function of internode distance. Of course, in real-world networks this is unlikely to be quite so literally the case; we might expect connectivity to be *influenced* by, rather than completely *prescribed* by distance considerations. It is also essentially a *static* model, whereas many (if not most) real-world complex networks evolve according to some dynamic growth process. Our model should thus, perhaps, be viewed as a tool or *framework* for garnering insight into the effects of spacial embedding on network structure—a "random graphs in space," if you like—rather than as a realistic model for any specific class of real-world spatial networks. On a practical note, we encounter the inevitable balancing act between generality and analytic tractability. Hopefully our model is extensible in

*lionelb@sussex.ac.uk
†ezequiel@sussex.ac.uk
‡sgb@ecs.soton.ac.uk

both directions; that is, it may be easily specialized to more tractable (and quite likely more domain-specific) scenarios while at the same time some of its more restrictive limitations might be relaxed.

II. THE SERN MODEL

The idea of our spatially embedded random networks (SERN) model is that nodes are placed randomly and independently according to a specified distribution on some metric space. An edge is then assigned independently to each pair of distinct nodes with probability depending only on the distance between the nodes. In the interests of simplicity and analytical convenience we explicitly preclude both *self-connection* and *multiple connections*; most results presented here extend reasonably straightforwardly to the case where either or both are admitted.

Definition 1. A SERN ensemble of size N (*i.e., of N nodes*) *is specified by the following.*

(1) *A metric* embedding space $\mathcal{S}$ *with metric* $d:\mathcal{S}\times\mathcal{S}\rightarrow\mathbb{R}^+$; $\mathcal{S}$ *is to represent the space in which network nodes reside.*

(2) *A connectivity decay function* $\gamma:\mathbb{R}^+\rightarrow[0,1]$; $\gamma(s)$ *is to represent the probability of assigning an edge to a pair of nodes distance s apart.*

(3) *A* node distribution *random variable (rv) X taking values in* $\mathcal{S}$; *X is to represent the location in* $\mathcal{S}$ *of a randomly situated node.*

(4) *A number* $N\in\{1,2,3,\ldots\}$ *of nodes.*

As regards the embedding space, although the construction may be applied to any metric space, the principal situation we have in mind—and to which we restrict ourselves in this study—is that where $\mathcal{S}$ is a Riemannian manifold, possibly with boundary. In this case we assume that the measure on $\mathcal{S}$ under which the node distribution random variable X is defined is the volume element associated with the Riemannian metric and that the (global) metric be *compatible* with the Riemannian metric [17]. Note that we do not demand that the metric be identified with *geodesic* distance as we would like to admit the case where the global metric is inherited from an embedding space; for example we might take distance between points on a circle as Euclidean distance in a plane in which the circle is embedded, rather than as distance around the perimeter of the circle. We note too that while we refer to $\gamma(s)$ as a "decay" function, it need not necessarily decrease monotonically with distance; nevertheless, we might expect it generally to do so in any realistic scenario, e.g. on grounds of per-distance connection cost for technological networks or of increasing density of "obstacles" to connection with distance [6,12].

To instantiate a network from a SERN ensemble of size N, we sample X independently N times to yield nodes at $x_1,x_2,\ldots,x_N\in\mathcal{S}$. We then assign connections independently with probability $\gamma(d(x_i,x_j))$ to each of the $\frac{1}{2}N(N-1)$ potential edges. Though essentially *static*, we might also view this construction as a *dynamic growth* model [18,19], where nodes—and their connections—are assigned sequentially. In general we shall be interested in properties of *large* networks; i.e., for $N\gg 1$. How we approach this is described in detail in Sec. III A.

We might at this point identify an apparent limitation of the SERN model: sampling node locations independently and identically from a single distribution X arguably precludes many plausible node distributions, in particular those exhibiting certain kinds of spatial *clustering*. A reasonable generalization would be to replace independent sampling from a single random variable with a more general *spatial point process* [20]; indeed, our distribution X might be viewed as a (heterogeneous) Poisson point process. We intend to address this generalization in a later paper. Nevertheless, we justify our model as a baseline for a theory of spatial networks; it might be taken to stand towards more realistic spatial network models as classical (Erdös-Rényi) *random graphs* stand towards the many more sophisticated nonspatial network models.

A. Relation to existing models

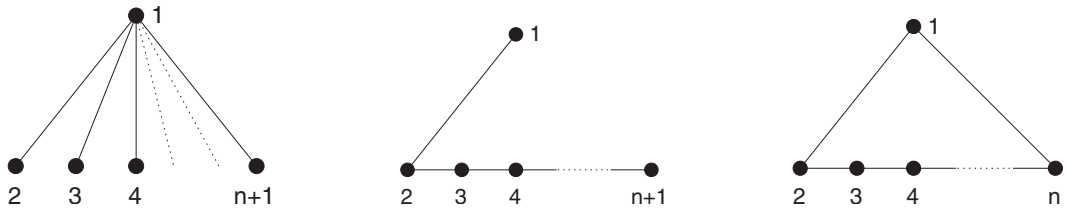
Antecedents for our SERN construction include the Watts-Strogatz small world model [2] (which might already be said to have a spatial element insofar as nodes are preferentially connected to "nearby" nodes) and more pertinently random geometric graph (RGG) models [13,14]. Indeed RGGs may be considered a specialization of our SERN model, where the embedding space is Euclidean, the node distribution uniform, and the connectivity decay function a simple cutoff at a characteristic distance. In Sec. V we examine a generalization of RGGs to more general spaces and node distributions. We remark here that uniformity of the node distribution in traditional RGGs is indeed a significant restriction; as we shall see below, some intriguing aspects of our generalization are direct consequences of *non*uniformity.

Other relevant work includes Ref. [4], which models connectivity in the Internet as dependent on spatial distance with exponential decay and Ref. [6], where a connectivity decay mechanism similar to our $\gamma(s)$ is induced by a density of "obstacles" in a communication network model situated in a two-dimensional arena containing uniformly randomly distributed nodes.

The principal defining property as regards *connectivity* in the SERN model is the *connection probability*:

$$c(x,y)\equiv\gamma(d(x,y)) \tag{1}$$

for $x,y\in\mathcal{S}$, which specifies the probability that nodes located at x,y be connected. Indeed, *all* (nonspatial) properties of our model are specified by a symmetric function $c:\mathcal{S}\times\mathcal{S}\rightarrow[0,1]$. Simply specifying $c(x,y)$ rather than $d(x,y)$ and $\gamma(s)$ yields a model akin to the related "fitness," "type," or "hidden variable" models proposed in various forms in [21–24]. In these models each node is independently assigned a (real or integer-valued) hidden variable drawn independently from a specified distribution and connections are independently assigned according to a symmetric function of the hidden variables associated with the endpoints. The chief difference with our model is that our "hidden" variables have an explicit interpretation as *spatial locations* of nodes rather than mere abstract numerical quantities, while the symmetric



(a) Fan (b) Chain (c) Loop

FIG. 1. Network motifs corresponding to connectivity moments [Eqs. (17a)–(17c)].

function depends solely on *spatial distance* between nodes and is thus constrained by the geometry of the underlying space. Indeed, while in this paper we address only connectivity statistics—those structural properties such as mean node degree, component size, path length, etc.—we nonetheless maintain the distinction between metric and decay on physical grounds, and in particular with regard to scaling in the large network limit (Sec. III). Inherently spatial statistics (such as mean *physical* path length) which depend on the spatial embedding of nodes warrant a study of their own and might be expected to be of particular significance to the analysis of *processes* on spatially embedded networks.

B. Fundamental properties

We remark that a SERN is essentially an "ensemble-of-ensembles"; that is, there are two levels of randomness: node placement and—given node placement—edge assignment. Generally all expectations and moments are, unless indicated to the contrary, taken to be averaged over the full ensemble-of-ensembles. If we wish to consider expectation of a property P, say, *conditional on node placement* $\mathbf{X} \equiv (X_1, X_2, \ldots, X_N)$ we use the language of conditional expectation and write $\mathbf{E}(P|\mathbf{X})$ (note that this is a random variable [25]).

The *mean connectivity* is defined to be

$$\mu \equiv \mathbf{E}(c(X,Y)), \tag{2}$$

where Y is independent and identically distributed (iid) as X [the expectation is guaranteed to exist since $c(X,Y)$ is bounded]. Note that μ corresponds to the connection probability p in a standard Erdös-Rényi random graph, which may indeed be considered a special case of our model with connection decay probability $\gamma(s)=p$ independently of distance s. We also define

$$\sigma^2 \equiv \mathrm{var}(c(X,Y)), \tag{3}$$

$$\rho \equiv \mathrm{corr}(c(X,Y),c(X,Z)), \tag{4}$$

where Y,Z are iid as X, as the *connectivity variance* and *connectivity correlation*, respectively.

Suppose we have a SERN ensemble of size N. Let the random vector $\mathbf{X} \equiv (X_1, X_2, \ldots, X_N)$, where the X_i are iid as X, specify a node distribution. We define the (random) *connectivity matrix* $\mathbf{C}$ by

$$C_{ij} \equiv \begin{cases} c(X_i,X_j), & i \neq j \\ 0, & i = j. \end{cases} \tag{5}$$

The C_{ij} are thus jointly distributed random variables and C_{ij} is just the probability that nodes i,j are connected [26]. For distinct indices i,j,k,l we have

$$\mathbf{E}(C_{ij}) = \mu, \tag{6}$$

$$\mathrm{var}(C_{ij}) = \sigma^2, \tag{7}$$

$$\mathrm{cov}(C_{ij},C_{ik}) = \rho\sigma^2, \tag{8}$$

$$\mathrm{cov}(C_{ij},C_{kl}) = 0. \tag{9}$$

For $x,y \in S$ let $A(x,y)$ denote a Bernoulli trial [27] with probability $c(x,y)$; i.e.,

$$\mathbf{P}(A(x,y) = 1) \equiv c(x,y). \tag{10}$$

The (random) *adjacency matrix* $\mathbf{A}$ is then defined by

$$A_{ij} \equiv \begin{cases} A(X_i,X_j), & i \neq j \\ 0, & i = j. \end{cases} \tag{11}$$

The A_{ij} are thus jointly distributed Bernoulli trials with $A_{ij}=1$ iff nodes i,j are connected and zero otherwise. For distinct indices i,j,k,l:

$$\mathbf{E}(A_{ij}) = \mathbf{P}(A_{ij} = 1) = \mu, \tag{12}$$

$$\mathrm{var}(A_{ij}) = \mu(1-\mu), \tag{13}$$

$$\mathrm{cov}(A_{ij},A_{ik}) = \rho\sigma^2, \tag{14}$$

$$\mathrm{cov}(A_{ij},A_{kl}) = 0. \tag{15}$$

An essential feature of the SERN model is that edge assignment is "conditionally independent" given node placement $\mathbf{X}$; that is, the $A_{ij}|\mathbf{X}$ are independent. More explicitly, given $\mathbf{x}=(x_1,\ldots,x_N)$ say, the random variables $A(x_i,x_j)$ are independently—but not, of course, identically—distributed. We note that $\mathbf{C}$ and $\mathbf{A}$ are related by

$$\mathbf{E}(A_{ij}|\mathbf{X}) = C_{ij}. \tag{16}$$

It will also turn out to be convenient to define the following sequences of higher-order *connectivity moments*, which we may associate with the network structural "motifs" [19] suggested by their respective descriptions (Fig. 1):

$$\textit{Fan moments:}\quad \mu_n \equiv \mathbf{E}(C_{12}C_{13}\dots C_{1,n+1}),\quad n=1,2,\dots,N-1; \tag{17a}$$

$$\textit{Chain moments:}\quad \nu_n \equiv \mathbf{E}(C_{12}C_{23}\dots C_{n,n+1}),\quad n=1,2,\dots,N-1; \tag{17b}$$

$$\textit{Loop moments:}\quad \omega_n \equiv \mathbf{E}(C_{12}C_{23}\dots C_{n,1}),\quad n=3,4,\dots,N. \tag{17c}$$

Since the C_{ij} are bounded all expectations are guaranteed to exist. Note too that the labeling of nodes $1,2,\dots$ is arbitrary (and, since the X_i are iid, irrelevant). This construction might be generalized to arbitrary motifs as follows: if $\mathcal{G}$ is any graph, then we could define $\mu(\mathcal{G})$ to be the appropriate expectation over all subgraphs of our network isomorphic to $\mathcal{G}$.

We have the identities

$$\mu_1 \equiv \nu_1 \equiv \mu, \tag{18}$$

$$\mu_2 \equiv \nu_2 = \rho\sigma^2 + \mu^2. \tag{19}$$

III. STATISTICAL PROPERTIES OF SERN ENSEMBLES

A. The large network limit

As mentioned in Sec. I, we shall be concerned chiefly with the connectivity statistics of "large" networks. More precisely, we consider limiting properties of *sequences* of SERN ensembles of size N as $N\to\infty$. For such sequences we restrict ourselves to the case where both the underlying metric space $\mathcal{S},d(\cdot,\cdot)$ and node distribution X are held fixed [28] and only the decay function $\gamma(s)$ is scaled with network size.

In the passage to the limit we follow common (but not exclusive) practice in network theory and keep the *mean degree* of nodes fixed. We shall see in the next section that for a SERN ensemble of N nodes this is just $(N-1)\mu$; we thus introduce the *mean degree* parameter:

$$\kappa \equiv (N-1)\mu. \tag{20}$$

Now, unlike Erdös-Rényi random graphs where we have a single real-valued scale parameter p, scaling a SERN ensemble involves scaling a *function* $\gamma(s)$ with potentially infinite degrees of freedom. There is therefore no unique "generic" scaling mode; rather, for given mean degree κ, scaling of connectivity with increasing network size requires a *sequence* of decay functions $\gamma_N:\mathbb{R}^+\to[0,1]$ under the constraint that κ be held constant. By abuse of terminology, in referring to "a large SERN ensemble" (cf. Sec. I), it is to be understood that such a scaling sequence is implicit, although in the interests of notational brevity we generally suppress the subscript N.

We return to the topic of scaling later; we shall see that network structure in the large network limit may depend crucially on the precise scaling model (cf. Sec. IV).

B. Degree distribution

Let the rv

$$K_i \equiv \sum_j A_{ij} \tag{21}$$

$(i=1,\dots,N)$ denote the degree of node i. The K_i are identically (but not independently) distributed as K, say. We thus find immediately that the mean degree of a randomly selected node is

$$\mathbf{E}(K) = (N-1)\mu = \kappa. \tag{22}$$

More generally, the joint probability generating function (pgf) [29] of the K_i is

$$G(z_1,\dots,z_N) = \mathbf{E}(z_1^{K_1}\dots z_N^{K_N}) = \mathbf{E}\left(\prod_{i<j}(z_iz_j)^{A_{ij}}\right) = \mathbf{E}\left(\prod_{i<j}[1-(1-z_iz_j)C_{ij}]\right). \tag{23}$$

The pgf for the degree of an arbitrary node (node 1, say) may be calculated as follows:

$$G(z) = \mathbf{E}\left(\prod_{1<j}[1-(1-z)C_{1j}]\right) = \mathbf{E}\left(\mathbf{E}\left(\prod_{1<j}[1-(1-z)C_{1j}]|X_1\right)\right) \tag{24}$$

$$= \mathbf{E}\left(\prod_{1<j}[1-(1-z)\mathbf{E}(C_{1j}|X_1)]\right) \tag{25}$$

$$= \mathbf{E}([1-(1-z)\mathbf{E}(c(X,Y)|X)]^{N-1}), \tag{26}$$

with Y iid as X. At step (24) we condition on X_1; at step (25) we use the fact that the C_{1j} given X_1 are independent and at step (26) that the $\mathbf{E}(C_{1j}|X_1)$ are all equal to $\mathbf{E}(c(X,Y)|X)$. It is now convenient to introduce the function:

$$\varphi(x) \equiv \mathbf{E}(c(x,Y)), \tag{27}$$

which represents the mean connectivity for a node at $x\in\mathcal{S}$ and we define the *conditional mean degree* to be the random variable:

$$\Gamma \equiv (N-1)\varphi(X) = (N-1)\mathbf{E}(c(X,Y)|X). \tag{28}$$

We may then write, Eq. (26) as

$$G(z) = \mathbf{E}\left(\left[1-\frac{1}{N-1}(1-z)\Gamma\right]^{N-1}\right). \tag{29}$$

Γ completely determines the degree distribution. Intuitively, Γ describes how mean node degree varies with location of a node in $\mathcal{S}$. In particular [cf. Eq. (25) below] its *variance* may be considered a measure of the "spatial inhomogeneity" of connectivity of a SERN ensemble. The nth moment of Γ is given by

$$\mathbf{E}(\Gamma^n) = (N-1)^n\mu_n = \kappa^n\frac{\mu_n}{\mu^n}, \tag{30}$$

where μ_n is the nth *fan moment* (17a). In particular,

$$\mathbf{E}(\Gamma) = (N-1)\mu = \kappa,$$

$$\mathrm{var}(\Gamma) = (N-1)^2\rho\sigma^2 = \kappa^2\left(\frac{\mu_2}{\mu^2} - 1\right) \tag{31}$$

and we see immediately that if the connectivity correlation ρ vanishes then Γ has zero variance and is (almost surely) *constant*. The degree distribution is then binomially distributed and we might expect such networks to behave somewhat like random graphs; but note that statistical properties (such as clustering) which do not derive purely from the degree distribution may not be assumed to be as for random graphs. A particular situation where connectivity correlation vanishes is described in the next section.

In the large N limit, provided $\lim_{N\to\infty}\Gamma$ exists,

$$G(z) \to \mathbf{E}(e^{-(1-z)\Gamma}), \tag{32}$$

so that the *degree distribution* of a SERN ensemble is given by

$$\mathbf{P}(K=k) \to \frac{1}{k!}\mathbf{E}(\Gamma^k e^{-\Gamma}) \tag{33}$$

for $k=0,1,2,\ldots$.

Two further technical results are stated here as follows:

Proposition 1. In the large network limit:

$$\sum_{j\neq i} C_{ij} \to \Gamma \tag{34}$$

in distribution for any i.

Proof. We have

$$G(z) = \mathbf{E}(e^{\Sigma_{j\neq i}\ln(1-(1-z)C_{ij})}), \tag{35}$$

$$\to \mathbf{E}(e^{-(1-z)\Sigma_{j\neq i}C_{ij}}), \tag{36}$$

for large N, since the individual $C_{ij}\to 0$. Comparing with (32) and noting that $G(z)$ is just $M_\Gamma(z-1)$, where $M_\Gamma(t)$ is the moment generating function for Γ, the result follows from the continuity theorem [29]. ∎

Proposition 2. Let the rv Λ be the largest eigenvalue of the connectivity matrix $\mathbf{C}$. *Then in the large network limit, $\Lambda\to\Gamma$ in distribution.*

Proof. Since $\mathbf{C}$ is real, symmetric and presumably nonzero, the Perron-Frobenius theorem tells us that such a real-valued Λ exists and is associated with a real, positive (random) eigenvector $\mathbf{V}=(\mathbf{V}_1,\ldots,\mathbf{V}_N)$, say. Then $\Sigma_j C_{ij}V_j = \Lambda V_i$ and summing over i we have $(\Sigma_i V_i)\Lambda = \Sigma_{i,j}C_{ij}V_j$. But from Proposition 1 we have $\Sigma_i C_{ij}\to\Gamma$ and the result follows. ∎

As an example, if the decay function $\gamma(s)$ is const$=\mu$, so that our ensemble is a simple random graph with $C_{ij}=\mu$ for all $i\neq j$, then the characteristic equation for $\mathbf{C}$ is $(\Lambda+\mu)^{N-1}[\Lambda-(N-1)\mu]=0$ so that $\Lambda=\kappa=\Gamma$ in the large network limit.

C. Special SERN ensembles

At this point we introduce some subclasses of SERN ensembles with specialized spatial/statistical properties.

1. Poisson ensembles

From Eqs. (31) and (32) we see that if $\lim_{N\to\infty}\Gamma$ exists and $\mathrm{var}(\Gamma)\to 0$ (equivalently $\frac{\mu_2}{\mu^2}\to 1$) as $N\to\infty$ then the degree distribution tends towards a Poisson distribution with parameter κ. We shall call such an ensemble *Poisson*. Note that it is not sufficient merely that $\rho\to 0$ (cf. Sec. V).

2. Uniform ensembles

We describe a SERN ensemble as *uniform* if the node distribution X is uniform; it follows that the underlying space $\mathcal{S}$ must then have finite measure.

3. Spatially homogeneous ensembles

We describe a SERN ensemble as (spatially) *homogeneous* if $\forall x,y\in\mathcal{S}$ there exists an *isometry* $\phi:\mathcal{S}\to\mathcal{S}$ of $\mathcal{S}$ (i.e., metric-preserving $1-1$ mapping of $\mathcal{S}$ onto itself) that maps x to y and such that $\phi(X)$ has the same distribution as X. Since an isometry preserves the volume element on $\mathcal{S}$, it is necessary (but not sufficient) that the ensemble be uniform.

Proposition 3. Let Y be iid as X. Then if the network is homogeneous X and $d(X,Y)$ are independent.

Proof. Pick some fixed $x_0\in\mathcal{S}$. Then for any $x\in\mathcal{S}$ we can find an isomorphism ϕ such that $\phi(x)=x_0$. For any $a\in\mathbb{R}^+$ we have

$$\mathbf{P}(d(x,Y)\leq a) = \mathbf{P}(d(x_0,\phi(Y))\leq a) = \mathbf{P}(d(x_0,Y)\leq a).$$

Thus the distribution of $d(x,Y)$ does not depend on x and the result follows. ∎

Proposition 4. Let Y,Z be iid X. Then if the network is homogeneous $d(X,Y)$ and $d(X,Z)$ are independent.

Proof. Let $dV(x)$ be the volume element on $\mathcal{S}$. For any $a,b\in\mathbb{R}^+$ we have

$$\mathbf{P}(d(X,Y)\leq a, d(X,Z)\leq b) = \int_{\mathcal{S}} \mathbf{P}(d(x,Y)\leq a, d(x,Z)\leq b)\,dV(x)$$

$$= \int_{\mathcal{S}} \mathbf{P}(d(x,Y)\leq a)\mathbf{P}(d(x,Z)\leq b)\,dV(x)$$

by independence of Y, Z. But from Proposition 3, $\mathbf{P}(d(x,Y)\leq a)=\mathbf{P}(d(X,Y)\leq a)\,\forall x$ and the result follows. ∎

Corollary. If a SERN ensemble is homogeneous then the connectivity correlation ρ vanishes.

In particular, a homogeneous ensemble is Poisson provided that $\lim_{N\to\infty}\Gamma$ exists.

We thus see (cf. Sec. III C 5 below) that *spatial symmetry* imposes a rather severe constraint on network structure.

4. Uniformly continuous ensembles

Suppose that the density of the node distribution X is *uniformly continuous* [30]; we then describe the corresponding SERN ensemble as uniformly continuous. The uniformly continuity condition says roughly that small changes in location $x\in\mathcal{S}$ effect small changes in the density of X (continuity) and of a magnitude depending only on the size of the change in x and not on x itself (uniformity). As an example, the function $f(x)=e^{-x^2/2}$ on the real line is uniformly continuous, while the function $1/x$ on the interval $(0,1]$ is continu-

ous, but not uniformly so. We note that any *uniform* (and in particular any *homogeneous*) ensemble is uniformly continuous (see also Sec. V A).

5. Scale free ensembles

Following the argument in [15] we find that the degree distribution of a SERN ensemble will be scale free iff there is some $m \geq 1$ such that $\mathbf{E}(\Gamma^n)$ diverges for all $n > m$ in the large network limit $N \to \infty$. From Eqs. (30) and (20) we thus have:

Proposition 5. A SERN ensemble is scale free iff $\exists m \geq 1$ *such that for all* $n > m$, $\frac{\mu_n}{\mu^n} \to \infty$ *as* $N \to \infty$.

Clearly a Poisson ensemble cannot be scale free. In [15] it is demonstrated that scale free spatial networks do indeed exist; see also our Example 2 of Sec. V C.

6. Generalized random geometric graphs

Spatial network models with *truncation decay*:

$$\gamma(s) \equiv \begin{cases} 1, & 0 \leq s < r \\ 0, & r \leq s \end{cases} \tag{37}$$

so that $r \to 0$ as $N \to \infty$ have been quite widely studied in the literature—albeit exclusively for flat manifolds with uniform node distribution—as *random geometric graphs* (RGGs) [13,14]. Frequently, the emphasis has been on analysis of *thresholds* for the appearance of various structural properties, often using sophisticated techniques from continuum percolation theory. We generalize RGGs to *arbitrary* SERN ensembles with truncation decay; that is, $\mathcal{S}$ may be any Riemannian manifold and the node distribution X is not necessarily uniform. *Generalized random geometric graphs* (GRGGs) offer the possibility of being far more tractable to analysis than SERN ensembles in full generality; this is largely down to the observation that as $N \to \infty$ we always average quantities increasingly locally. We shall examine GRGGs in more detail in Sec. V.

D. Clustering coefficient

Our definition of clustering coefficient $\mathcal{C}$ will be "the probability that a random triplet of distinct nodes form a triangle, given that it form an 'elbow' "—this corresponds to the (more or less) standard version of

$$\mathcal{C} \equiv \frac{3 \times \text{no. of triangles}}{\text{no. of "elbows"}}$$

—but, we note there are two feasible ways to average this quantity over a SERN ensemble: (i) for a given node distribution $\mathbf{X}$ we construct the clustering coefficient for the (sub)ensemble *with node distribution* $\mathbf{X}$—then we average the coefficients over $\mathbf{X}$. (ii) node triplets are sampled from ensembles with *different* node distributions. Here we choose the second definition as being by far the easier to calculate, both analytically and in Monte Carlo simulation. Without loss of generality we label the nodes 1, 2, and 3 (with the "elbow" at node 1) so that

$$\mathcal{C} \equiv \mathbf{P}(A_{23} = 1 | A_{12} = 1, A_{13} = 1) = \frac{\omega_3}{\mu_2}, \tag{38}$$

that is, the third loop moment divided by the second fan moment.

E. Degree correlation

Firstly, we calculate the *conditional mean degree* of an arbitrary node; i.e., for $k = 0, 1, 2, \ldots$ we calculate the mean degree $\bar{\kappa}(k)$ of an arbitrary node conditional on it being connected to a node of degree k. Without loss of generality we take the nodes to be labeled 1 and 2. We then have

$$\bar{\kappa}(k) \equiv \mathbf{E}(K_2 | A_{12} = 1, K_1 = k). \tag{39}$$

Let us define the joint conditional generating functions (cf. [24]):

$$G(z_1, z_2) \equiv \sum_{k_1, k_2} \mathbf{P}(K_1 = k_1, K_2 = k_2 | A_{12} = 1) z_1^{k_1} z_2^{k_2}. \tag{40}$$

Analogously to Eq. (27), we define

$$\psi(x, y) \equiv \mathbf{E}(c(x, Z) c(y, Z)) \tag{41}$$

for $x, y \in \mathcal{S}$, where Z iid as X. Using the shorthand notation

$$C \equiv c(X, Y), \tag{42}$$

$$\Phi_1 \equiv \varphi(X), \tag{43}$$

$$\Phi_2 \equiv \varphi(Y), \tag{44}$$

$$\Psi \equiv \psi(X, Y), \tag{45}$$

we may then calculate

$$\mu G(z_1, z_2) = z_1 z_2 \mathbf{E}(C[1 - \Phi_1(1 - z_1) - \Phi_2(1 - z_2) + \Psi(1 - z_1)(1 - z_2)]^{N-2}). \tag{46}$$

Setting

$$g_k(z) \equiv \left. \frac{1}{k!} \frac{\partial^k G(z_1, z)}{\partial z_1^k} \right|_{z_1 = 0}, \tag{47}$$

we may now verify that

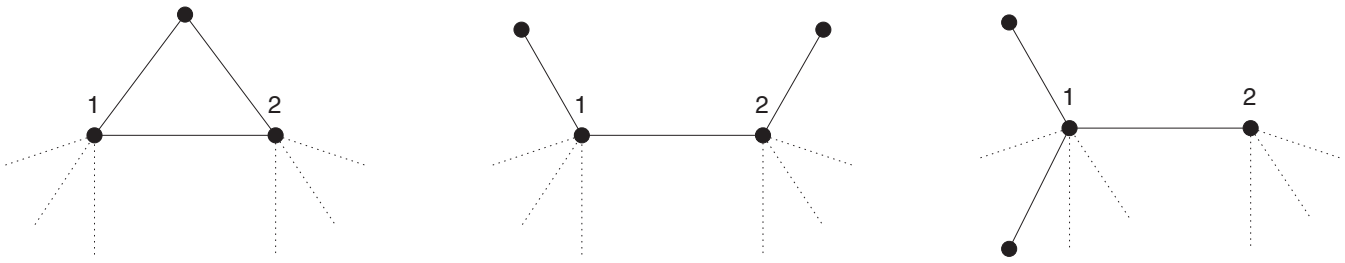
$$\bar{\kappa}(k) = \frac{g_k'(1)}{g_k(1)}, \tag{48}$$

and we have

$$\mu g_k(z) = \binom{N-2}{k-1} z \mathbf{E}(C[\Phi_1 - \Psi(1 - z)]^{k-1} \times [1 - \Phi_1 - (\Phi_2 - \Psi)(1 - z)]^{N-k-1}), \tag{49}$$

so that

$$\mu g_k(1) = \binom{N-2}{k-1} \mathbf{E}(C \Phi_1^{k-1} (1 - \Phi_1)^{N-k-1}), \tag{50}$$



(a) $\mathcal{C}(loop, correlates)$ (b) $\eta_1(chain, correlates)$ (c) $\eta_2(fan, decorrelates)$

FIG. 2. Motifs corresponding to degree correlation factors [Eqs. (56), (38), (57), and (58)].

$$\mu g_k'(1) = \mu g_k(1) + \binom{N-2}{k-1}(k-1)\mathbf{E}(C\Psi\Phi_1^{k-2}(1-\Phi_1)^{N-k-1}) + \binom{N-2}{k-1}(N-k-1)\mathbf{E}(C(\Phi_2-\Psi) \times \Phi_1^{k-1}(1-\Phi_1)^{N-k-2}). \quad (51)$$

Let us now define

$$P_k(x) \equiv \mathbf{P}(K=k|X=x) = \frac{1}{k!}[(N-1)\varphi(x)]^k e^{-(N-1)\varphi(x)} \quad (52)$$

for $x \in \mathcal{S}$ and set

$$\nu_2[k] \equiv \mathbf{E}(c(Y,Z)c(X,Y)P_{k-1}(X)), \quad (53)$$

$$\omega_3[k] \equiv \mathbf{E}(c(Y,Z)c(X,Y)c(X,Z)P_{k-2}(X)). \quad (54)$$

We may think of these as "degree conditional" versions of the corresponding connectivity moments. In the large N limit (with $k \ll N$) we then find that

$$\bar{\kappa}(k) \approx 1 + N^2 \frac{\nu_2[k] + \omega_3[k] - \omega_3[k+1]}{k\mathbf{P}(K=k)}. \quad (55)$$

We may express the (*conditional*) *degree correlation* in terms of partial derivatives of the generating function (40) at $z_1=z_2=1$. It may then be calculated from Eq. (46) to yield (approximately for large N)

$$\mathrm{corr}(K_1,K_2|A_{12}=1) \approx \frac{\eta_1\kappa + \mathcal{C}}{\eta_2\kappa + 1}, \quad (56)$$

where we define

$$\eta_1 \equiv \frac{\mu\nu_3 - \mu_2^2}{\mu^2\mu_2}, \quad (57)$$

$$\eta_2 \equiv \frac{\mu\mu_3 - \mu_2^2}{\mu^2\mu_2}. \quad (58)$$

We note firstly that the clustering coefficient $\mathcal{C}$ is always ≥ 0; $\eta_2 \geq 0$ follows from the Cauchy-Schwartz inequality [29], noting that $\mu_n = \mathbf{E}(\varphi(X)^n)$. In the general case, η_1 might be negative. Figure 2 illustrates the effects of the factors $\mathcal{C}, \eta_1, \eta_2$ entering into the expression (56) for degree correlation. Clustering is associated with increased degree correlation via shared edges creating a loop; η_2 corresponds to small fans emanating from one of the connected nodes, decreasing degree correlation. η_1 corresponds to short chains passing through the connected nodes, affecting degree correlation according to the sign of η_1. We note that in the Poisson case $\eta_2 \equiv 0$, so that the conditional degree correlation is just $\eta_1\kappa + \mathcal{C}$. In general, we might say intuitively that while clustering exerts a "homogeneous" correlating effect on degree, "inhomogeneity" (in the sense of spatial variation of connectivity) exerts an effect on degree correlation via thecoefficients η_1, η_2 (see in particular Sec. V A below). We note that a necessary condition that degree correlation be *dis*assortative is that $\eta_1 < 0$. It is not clear under what conditions this might occur; in Sec. V A we shall see that for uniformly continuous GRGGs (and, we conjecture, in fact for *any* GRGG) we have $\nu_n = \mu_n$ for all n, so that $\eta_1 = \eta_2$ and degree correlation is always assortative.

F. Giant component

For their "configuration model" Newman *et al.* [18,31] developed a powerful generating function formalism to derive—amongst other things—approximations for the phase transition to the appearance of a giant component and the size of the giant component beyond the phase transition. We cannot apply this formalism *mutatis mutandis* for two reasons: firstly—unlike the configuration model—node degrees (conditional on node distribution) are neither identically nor independently distributed for SERN ensembles. Secondly, the formalism depends on neglecting the probability of loops in small components; since SERN ensembles may well feature significant *clustering* this assumption is likely to be nonviable.

As an example, in [31] the following argument is used to derive the size of the giant component for the configuration model in the large network limit: assume there is a giant component and define u to be the probability that an arbitrary node does *not* belong to the giant component. Then none of that node's neighbors must belong to the giant component, leading to the "consistency relation":

$$u = G(u), \quad (59)$$

where $G(z)$ is the generating function for node degree (29). We can then—at least in principle—solve Eq. (59) for u (apart from the trivial solution $u=1$). The problem in our case is that in deploying this procedure naively we disregard

the fact that given a particular node distribution $\mathbf{X}$ the probability that a node belong to the giant component will vary with the choice of node. We should not be surprised, then, that Eq. (59) yields in general a poor approximation for u. In the large network limit Eq. (59) becomes

$$u = \mathbf{E}(e^{-(1-u)\Gamma}). \tag{60}$$

For Poisson ensembles in particular, we have $\Gamma = \text{const} = \kappa$ and Eq. (60) becomes in the large network limit:

$$u = e^{-\kappa(1-u)} \tag{61}$$

which may be solved in terms of (the principal branch of) Lambert's W function as

$$u = -\frac{1}{\kappa} W(-\kappa e^{-\kappa}) \approx \frac{1}{\kappa} e^{1-\kappa}, \tag{62}$$

which suggests, among other things, a phase transition at $\kappa = 1$ as for Erdös-Rényi random graphs; however, as demonstrated in [14] for uniform random geometric graphs (see Sec. V) on Euclidean space there is indeed a phase transition to the appearance of a giant component, but the critical connectivity in fact varies with *spatial dimension*, with the transition at $\kappa = 1$ the limiting value for large dimension. In effect, clustering—which decreases with increasing dimension [14]—renders the procedure inaccurate.

A more exact approach to the "consistency relation" argument runs as follows (see also [22] for a more precise derivation for a comparable model): given a node distribution $\mathbf{X}$ let U_i—now jointly distributed with $\mathbf{X}$—be the probability that node i does not belong to the giant component. The consistency relation then becomes

$$U_i = e^{-\Sigma_j C_{ij}(1-U_j)}. \tag{63}$$

If we could extract a (possibly approximate) *algebraic* solution of Eq. (63) considered as a set of simultaneous equations for the U_i then we could in principle calculate $u = \mathbf{E}(U_i)$; this would appear to be difficult, however, and we have not succeeded in doing so.

We in fact conjecture that, contrary to common assumption, for general SERN ensembles there is a phase transition *only in the case of uniform ensembles*; for all other ensembles, the fraction of the network occupied by the largest component increases smoothly with increasing mean connectivity κ. Intuitively, inhomogeneity in the spatial distribution of nodes induces a clustering of connectivity around regions of higher node density, which "smears out" the transition in the following sense: for a *fixed* node distribution (63) may yield a critical mean connectivity value, but this value will depend on the particular node distribution so that criticality is obliterated by the averaging process. This intuition is confirmed in all simulations we have performed with nonuniform ensembles [32] (see, e.g., Sec. V and Fig. 5).

A further subtlety arises in that simulations suggest that if mean connectivity κ is below the phase transition, or if there is no phase transition, then (holding κ fixed) the fraction of the network occupied by the largest component actually *shrinks* as $N \to \infty$; effectively, the network "breaks up" as network size increases.

G. Mean path length

We need to consider carefully what exactly we intend by "mean path length"; we would like to define the *path length* random variable L to be the minimum number of edges we need to traverse to connect two (distinct) randomly selected nodes in a random instantiation of our ensemble. But note that there may be *no* path connecting the nodes (specifically when the nodes lie in different components); in this case it is customary, by abuse of terminology, to refer to the path length as "infinite," so that L takes values in $\{1, 2, \ldots, N-1, \infty\}$. The most common approach is to take the conditional arithmetic mean $\mathbf{E}(L | L < \infty)$; that is, we take the mean over those pairs of distinct nodes for which a connecting path exists. An alternative (and arguably better) approach is to consider instead the *harmonic mean* $1/\mathbf{E}(1/L)$, which is well defined if we define $1/L \equiv 0$ where there is no connecting path.

We note that

$$\mathbf{P}(L < \infty) \geq (1-u)^2, \tag{64}$$

where u is as in the previous section; i.e., $1-u$ is the fraction of the network occupied by the largest component. We note that, as for the fraction of the network occupied by the largest component (see previous section) $\mathbf{P}(L < \infty)$ actually shrinks as $N \to \infty$; nonetheless, in all cases we have examined in simulation mean path length (both conditional arithmetic or harmonic) still appears to increase with network size.

Following [33] we may calculate the probability:

$$\mathbf{P}(L = \ell) = F_\ell - F_{\ell-1} \tag{65}$$

for $\ell = 1, 2, \ldots, N-1$ in the large network limit, where

$$F_\ell \equiv 1 - \mathbf{E}(\exp(-[\mathbf{C}^\ell]_{12})) \tag{66}$$

(and $F_0 \equiv 0$) yielding

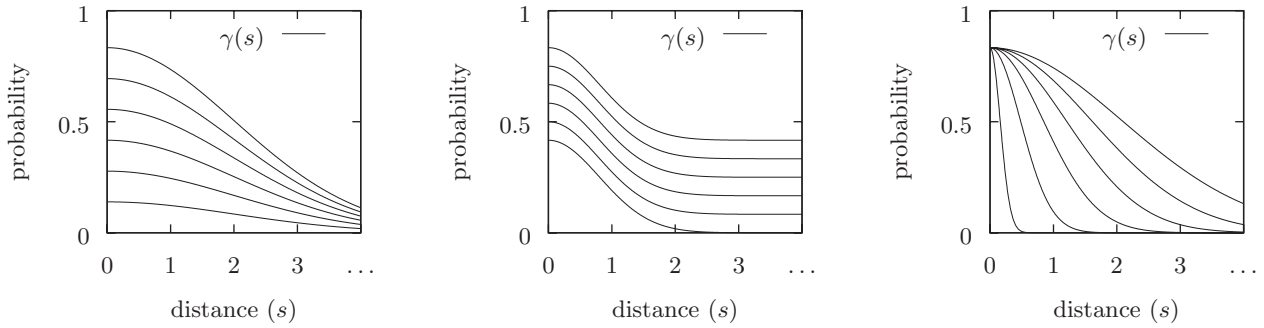
$$\mathbf{P}(L < \infty) = F_{N-1}, \tag{67}$$

$$\mathbf{E}(L | L < \infty) = N - \sum_{\ell=1}^{N-1} F_\ell / F_{N-1}. \tag{68}$$

Now in [33] a specialization of the "hidden variable" model introduced in [24] is analyzed in which the equivalent of our C_{ij} factorizes into the product of independent random variables. This allows the summation in Eq. (68) to be calculated via the Poisson summation formula [30]; in our case the nonseparability of the C_{ij} would seem to preclude this approach and it is not clear how we should calculate either Eq. (68) or indeed Eq. (67).

IV. SMALL WORLD SERN ENSEMBLES

To investigate whether a SERN ensemble may exhibit small world behavior it will become necessary to analyze more closely the scaling mode of the decay function (Sec. III A) in the large network limit in more detail. Consider given a "base" decay function $\hat{\gamma}(s)$, which defines the "shape" of the decay of connection probability with distance. Some plausible base decay functions might be *truncation*



(a) Mode 1: Eq. (69) (b) Mode 2: Eq. (70) (c) Mode 3: Eq. (71)

FIG. 3. Connection probability scaling modes (distance scale on x axis is arbitrary).

decay, where nodes are connected iff they lie within a given distance (cf. Sec. V), *linear decay*, *exponential decay*, *power law decay*, etc. Some intuitively "natural" ways to derive a sequence of decay functions from a base decay function in order to define a large network limit are (Fig. 3):

(1) Reduce the connection probability by a constant factor:

$$\gamma_N(s) \equiv \lambda_N \hat{\gamma}(s) \tag{69}$$

and let $\lambda_N \to 0$ as $N \to \infty$ so as to maintain $\kappa = \text{const}$.

(2) Reduce the connection probability by a constant amount:

$$\gamma_N(s) \equiv \hat{\gamma}(s) - (\hat{\gamma}_{min} - c_N) \tag{70}$$

[where $\hat{\gamma}_{min}$ is the minimum value of $\hat{\gamma}(s)$ over its domain] and let $c_N \to 0$ as $N \to \infty$ so as to maintain $\kappa = \text{const}$; but note that this imposes a limit on the largest possible N.

(3) Let characteristic decay length tend to zero (i.e., shrink the distance at which connections appear with given probability):

$$\gamma_N(s) \equiv \hat{\gamma}\left(\frac{s}{r_N}\right) \tag{71}$$

and let $r_N \to 0$ as $N \to \infty$, again while holding κ constant.

Now Watts and Strogatz originally defined "small world" to mean not just sublogarithmic scaling of mean path length with network size, but also nonvanishing of the clustering coefficient in the large network limit [2]. It is immediately evident from Eq. (69) that scaling according to mode 1 above will result in a vanishing clustering coefficient, since any connectivity moment of order n clearly scales as λ_N^n. Mode 2 does not allow us to take N to arbitrarily large size. For mode 3 we might well expect (and simulation bears this out) a nonzero clustering coefficient since "triangles" of nodes that fall entirely within the connectivity radius will be fully connected. However, the following heuristic argument suggests that scaling according to mode 3 might also be expected to lead to power law rather than logarithmic decay of mean path length.

Suppose then that a decay function $\gamma(s)$ has the property that $\gamma(s) = 0$ for $s > r$, say; that is, there will be no connections for nodes further apart than (spatial) distance r. Let the random variable $D \equiv d(X,Y)$, jointly distributed with mean path length L, be the distance between random nodes X, Y. We attempt to derive a lower bound on mean path length based on the observation that for any r and $\ell = 1, 2, \ldots, N-1$:

$$D > \ell\, r \Rightarrow L > \ell\,, \tag{72}$$

i.e., if two nodes are a greater distance than ℓr apart, then it takes at least $\ell + 1$ edges to link them. From this we deduce that for the conditional arithmetic mean

$$\mathbf{E}(L | L < \infty) \geq \frac{1}{r}\mathbf{E}(D | L < \infty) \tag{73}$$

if the conditional mean $\mathbf{E}(D | L < \infty)$ exists. This is, perhaps, not terribly useful except for the case where mean connectivity is beyond a phase transition to formation of a giant component, in which case we may treat the term $\mathbf{E}(D | L < \infty)$ as constant. Similarly we find for the harmonic mean

$$\frac{1}{\mathbf{E}(1/L)} \geq \frac{1}{r}\frac{1}{\mathbf{E}(1/D)} \tag{74}$$

again provided $\mathbf{E}(1/D)$ exists [34].

In either case, within the (admittedly rather restrictive) provisos outlined above, we see that path length scales at least as $\frac{1}{r}$. Suppose now that a base decay function $\hat{\gamma}(s)$ satisfies $\hat{\gamma}(s) = 0$ for $s > 1$, say. We then have

$$\mu = r \int_0^1 \hat{\gamma}(s) \varrho(rs) ds, \tag{75}$$

where $\varrho(s)$ is the density of $d(X,Y)$, so that if $\varrho(s)$ is reasonably well behaved we have approximately, for small r:

$$\mu \propto r^{\beta} \tag{76}$$

for some $\beta > 0$ and, since we hold $\kappa \equiv (N-1)\mu$ fixed, we see from Eq. (73) [respectively, (74)] that conditional arithmetic (respectively, harmonic) path length scales at least as $N^{1/\beta}$ as $N \to \infty$; in particular, it grows faster than logarithmically, so that our network cannot be small world. We conjecture that this is in fact always the case if connectivity decays according to mode 3; simulations of various SERN ensembles (including severely singular node distributions) under this scaling mode bear out this conclusion.

To summarize: reducing connection probability by a constant factor may well result in mean path length scaling sublogarithmically with network size, but implies a vanishing clustering coefficient. Reducing connection probability by a constant amount does not allow passage to a large network limit. Shrinking the connection radius may well yield a nonzero clustering coefficient, but is likely to lead to superlogarithmic scaling of mean path length.

There is, however, a suggestion in [14] that we may achieve small world behaviour in the sense of Watts and Strogatz by deploying a *combination* of scaling modes (see also [12,35]). We demonstrate a construction along these lines in the following section.

Construction of small world ensembles

Consider the behavior of the clustering coefficient for a decay function sequence:

$$\tilde{\gamma}_N(s) \equiv (1-q_N)\gamma_N(s) + q_N, \tag{77}$$

where $\gamma_N(s)$ is a given decay function sequence yielding a nonzero clustering coefficient $\mathcal{C}_N \to \mathcal{C} > 0$ in the limit and $q_N \to 0$ as $N \to \infty$. Indeed, if the $\gamma_N(s)$ are truncation decay (cf. Sec. V), the resulting ensemble may be thought of as a continuous analog of the original Watts-Strogatz model: a small decay radius r leads to latticelike local connectivity, while a uniform connection probability q corresponds to random distance-independent (hence potentially long range) connections. We wish to determine firstly *how* we may take $q_N \to 0$ in order that the clustering coefficient $\tilde{\mathcal{C}}$ tend to a nonzero limit. We wish to establish in addition whether appropriate q_N scaling can yield sublogarithmically scaling (i.e., small world) mean path length.

Denoting quantities corresponding to the decay function (77) by a tilde (from here forward we drop the subscript N for notational clarity), we have

$$\tilde{\varphi}(x) = (1-q)\varphi(x) + q \tag{78}$$

for $x \in \mathsf{S}$, and

$$\tilde{\mu} = (1-q)\mu + q, \tag{79}$$

$$\tilde{\mu}_2 = (1-q)^2\mu_2 + 2q(1-q)\mu + q^2, \tag{80}$$

$$\tilde{\omega}_3 = (1-q)^3\omega_3 + 3q(1-q)^2\mu_2 + 3q^2(1-q)\mu + q^3. \tag{81}$$

Setting $\lambda \equiv \mu_2/\mu^2$ we may calculate

$$\tilde{\mathcal{C}} - \tilde{\mu} = \frac{(1-q)\lambda(\mathcal{C}-\mu) + 2q(\lambda-1)}{\lambda - 1 + (1+\frac{q/\mu}{1-q}2)}. \tag{82}$$

In the large network limit and assuming $\mathcal{C} \to \text{const} > 0$ we have

$$\tilde{\mathcal{C}} \to \frac{\lambda\mathcal{C} + 2q(\lambda-1)}{\lambda - 1 + (1+q/\mu^2)} \tag{83}$$

and we may state

Proposition 6. In the large network limit $N \to \infty$, *if* $\mathcal{C} \to \text{const} > 0$ *then* $\tilde{\mathcal{C}} \to \text{const} > 0$ *iff*:

$$\text{I.} \quad \lambda \to \text{const} > 0 \quad \textit{and} \quad \frac{q}{\mu} \to c$$

$$\textit{so that} \quad \tilde{\mathcal{C}} \to \frac{\lambda}{\lambda - 1 + (1+c)^2}\mathcal{C} \tag{84}$$

$$\textit{or} \quad \text{II.} \quad \lambda \to \infty \quad \textit{and} \quad \frac{q}{\sqrt{\mu_2}} \to c$$

$$\textit{so that} \quad \tilde{\mathcal{C}} \to \frac{1}{(1+c)^2}\mathcal{C} \tag{85}$$

for some constant $c \geq 0$.

It follows that if $\mathcal{C} \to \text{const} > 0$ we may always scale q with μ so that $\tilde{\mathcal{C}} \to \text{const} > 0$ and $\tilde{\mu} \to 0$ as required. We note that Case I covers Poisson ensembles, where $\lambda \equiv 1$.

We also note from Eq. (78) that if our original ensemble is Poisson, then so is the derived small world ensemble. We may calculate for Case I that

$$\frac{\tilde{\mu}_n}{\tilde{\mu}^n} \to (1+c)^{-n}\sum_{k=0}^{n}\binom{n}{k}c^{n-k}\frac{\mu_k}{\mu^k}, \tag{86}$$

so that the derived small world ensemble is scale free iff the original ensemble is scale free, while for Case II:

$$\frac{\tilde{\mu}_n}{\tilde{\mu}^n} \to \sum_{k=0}^{n}\binom{n}{k}c^{-k}\frac{\mu_k}{(\sqrt{\mu_2})^k}, \tag{87}$$

so that the derived small world ensemble is scale free iff $\exists n_0$ such that $\mu_n/(\sqrt{\mu_2})^n$ diverges for all $n > n_0$.

We now turn to mean path length scaling as $N \to \infty$, holding mean degree fixed. Now it is clear from Eq. (79) that, given fixed mean degree $\tilde{\kappa} \equiv N\tilde{\mu}$ the mean path length for the new ensemble satisfies

$$\mathbf{E}(L(\tilde{\kappa},N)) \leq \mathbf{E}(L_{RG}(Nq,N)), \tag{88}$$

where L_{RG} denotes path length for an Erdös-Rényi *random graph* [36], since the additional connection probability $(1-q)\mu$ in Eq. (79) can only *decrease* mean path length. Now for Case I, $q \to c\mu$ for some constant c, so that in the large network limit we find that $q \approx \frac{c}{1+c}\tilde{\mu}$, so that

$$\mathbf{E}(L(\tilde{\kappa},N)) \leq \mathbf{E}\left(L_{RG}\left(\frac{c}{1+c}\tilde{\kappa},N\right)\right). \tag{89}$$

Since we know that mean path length for a random graph scales logarithmically with network size, it follows that so too does $\mathbf{E}(L(\tilde{\kappa},N))$ and the new ensemble is indeed small world. For Case II we find that $q \approx \tilde{\mu}$ for large N so that

$$\mathbf{E}(L(\tilde{\kappa},N)) \leq \mathbf{E}(L_{RG}(\tilde{\kappa},N)), \tag{90}$$

and the new ensemble is again seen to be small world.

V. GENERALIZED RANDOM GEOMETRIC GRAPHS

In this section we specialize the previously developed formalism to generalized random geometric graphs (GRGGs)

(Sec. III C 6), illustrating our analysis with some examples.

We note here that for GRGGs we always have

$$\sigma^2 \equiv \mu(1-\mu) \approx \mu \tag{91}$$

in the large network limit.

A. Uniformly continuous GRGGs

Suppose we have a uniformly continuous GRGG (Secs. III C 4 and III C 6) on a closed, oriented m-dimensional Riemannian manifold $\mathcal{S}$. Since as $N \to \infty$ we always average quantities increasingly locally we may convince ourselves that to lowest order in r we may approximate our space as *Euclidean*; that is, to a first approximation *curvature* of $\mathcal{S}$ may be neglected in the large network limit (this may be demonstrated more rigorously by choosing a Riemann normal coordinate system [37] locally). It is less obvious that if $\mathcal{S}$ has a sufficiently well-behaved boundary, then the boundary makes a negligible contribution to the connectivity moments in the large network limit.

We first consider the boundary-less case. By uniform continuity we have for any $\varepsilon > 0$ an $r > 0$ such that $d(x,y) < r \Rightarrow |p(y) - p(x)| < \varepsilon$. Thus,

$$\varphi(x) \equiv \int_{d(x,y) \le r} p(y) dV(y), \tag{92}$$

where $dV(y)$ is the volume element in our chosen coordinate system, satisfies

$$\left| \frac{\varphi(x)}{V(r;x)} - p(x) \right| < \varepsilon, \tag{93}$$

where $V(r;x)$ represents the volume of a ball of radius r around x. But if r is chosen small enough, then to lowest order in r, $V(r;x) \approx V_m(r)$, where

$$V_m(r) = \frac{\pi^{m/2}}{\Gamma(\frac{m}{2}+1)} r^m \tag{94}$$

is the volume of a ball of radius r in a Euclidean space of dimension m, and we have

$$\varphi(x) \approx V_m(r) p(x) \tag{95}$$

to lowest order in r.

Now suppose $\mathcal{S}$ has a boundary $\partial\mathcal{S}$. Let us denote by $\partial\mathcal{S}(r)$ the set of points in $\mathcal{S}$ of distance at most r from $\partial\mathcal{S}$. From (95) we have

$$\mu \equiv \int_{\mathcal{S}} \varphi(x) p(x) dV(x), \tag{96}$$

$$\approx V_m(r) \mathbf{E}(p(X)) - \int_{\partial\mathcal{S}(r)} [V_m(r) - V(r;x)] p(x)^2 dV(x). \tag{97}$$

But

$$\int_{\partial\mathcal{S}(r)} [V_m(r) - V(r;x)] p(x)^2 dV(x) \le V_m(r) \int_{\partial\mathcal{S}(r)} p(x)^2 dV(x), \tag{98}$$

and if the boundary is piecewise smooth and we choose r sufficiently small, then

$$\int_{\partial\mathcal{S}(r)} p(x)^2 dV(x) \approx r \int_{\partial\mathcal{S}} p(x)^2 dS(x) \tag{99}$$

where $dS(x)$ is the volume element on the boundary $\partial\mathcal{S}$ in our chosen coordinate system, so that to lowest order in r:

$$\mu \approx V_m(r) \mathbf{E}(p(X)). \tag{100}$$

In the limit of small r, then, nodes near the boundary make a negligible contribution to μ. The argument extends to higher connectivity moments so that:

Proposition 7. For a uniformly continuous GRGG:

$$\mu_n \approx V_m(r)^n \mathbf{E}(p(X)^n) \tag{101}$$

for $n = 1, 2, \ldots$ *in the limit* $N \to \infty$. ∎

We note firstly that

$$\frac{\mu_n}{\mu^n} \to \frac{\mathbf{E}(p(X)^n)}{\mathbf{E}(p(X)^n)} \tag{102}$$

in the large network limit, so that

Corollary (1). A uniformly continuous GRGG can never be scale free. ∎

and secondly, that

$$\frac{\mu_2}{\mu^2} - 1 \to \frac{\mathrm{var}(p(X))}{\mathbf{E}(p(X))^2} \tag{103}$$

in the large network limit, giving:

Corollary (2). A uniformly continuous GRGG is Poisson iff it is uniform. ∎

It is straightforward to show also that $\nu_n \approx \mu_n$ to lowest order in r for $n = 1, 2, \ldots$ (we conjecture that this is, in fact, the case for *any* GRGG). Thus to calculate the conditional degree correlation (56) we have $\eta_1 \equiv \eta_2 \equiv \eta$, say, for the coefficients (57) and (58), which may then be calculated from Eq. (102). As we noted in Sec. III E that $\eta_2 \ge 0$ always, we have

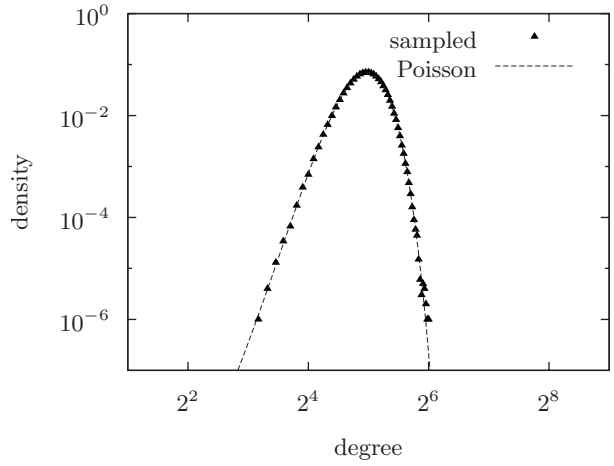
Proposition 8. A uniformly continuous GRGG always has assortative degree correlation. ∎

Furthermore, we see that as η increases from zero (the Poisson or, equivalently, uniform case) so the conditional degree correlation (56) increases from $\mathcal{C}$ to 1; we might interpret this as saying that "inhomogeneity induces additional degree correlation beyond that associated with clustering."

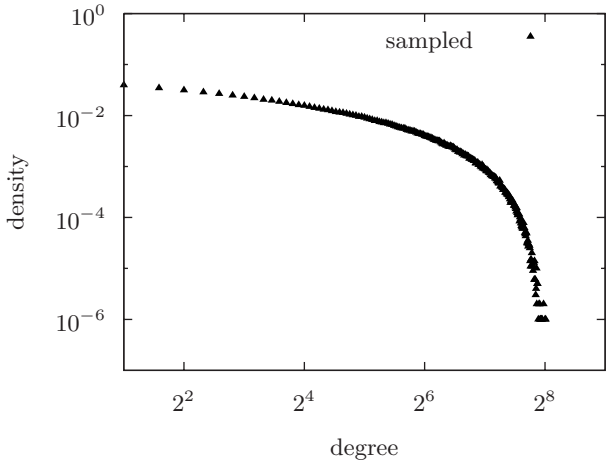
From Eq. (91) we also see that

$$\rho = \frac{\mu_2 - \mu^2}{\mu(1-\mu)} \to 0 \tag{104}$$

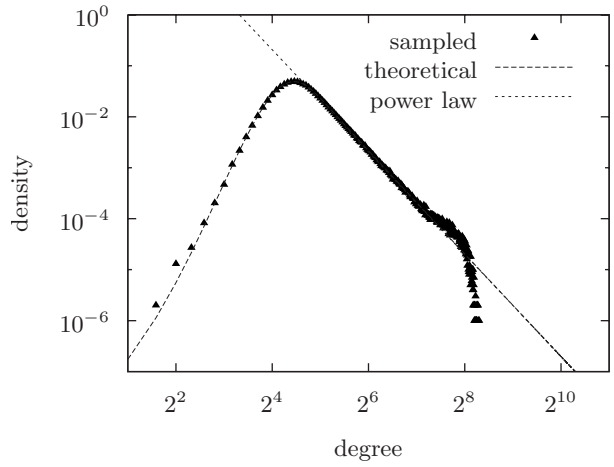
in the large network limit. If the GRGG is not uniform we thus have an example of a *non*-Poisson SERN ensemble for which the connectivity correlation tends to zero in the large network limit (cf. Sec. III C 1).



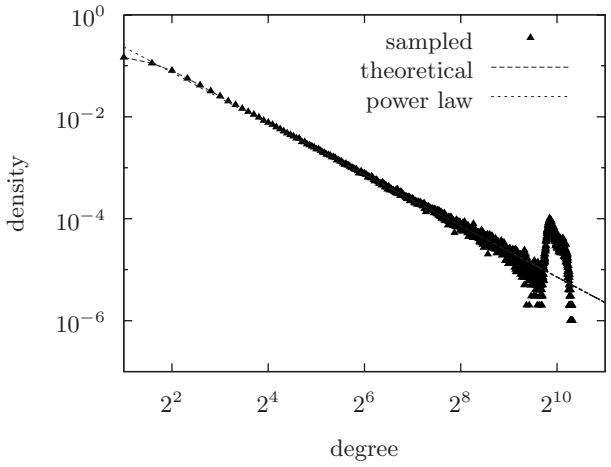
(a) Uniform hypersphere, dim = 6



(b) Gaussian (Section V B), dim = 6



(c) Scale free (Section V C), $\beta = 0.3$



(d) Scale free (Section V C), $\beta = 0.6$

FIG. 4. Degree distribution plotted on a log-log scale for several GRGGs (Sec. V) estimated in sample. Mean connectivity $\kappa=32$, network size $N=2^{16}$, and sample size$=10^6$ for all figures.

The clustering coefficient for RGGs has been computed analytically in [14]; there they note that $\mathcal{C}$ is independent of network size and (mean) connectivity. For our GRGGs their result still holds as a limiting case as $N\to\infty$, although it need no longer be constant. They find that $\mathcal{C}$ is always >0 and for large dimension m:

$$\mathcal{C} \approx 3\sqrt{\frac{2}{\pi m}}\left(\frac{3}{4}\right)^{(m+1)/2}. \tag{105}$$

We note here that simulations indicate (see, e.g., Fig. 6) that for GRGGs, mean path length—both conditional arithmetic and harmonic—appears to scale as a power law in network size as suggested by the discussion in Sec. IV and that the exponent is reasonably well predicted by the lower bounds given by Eqs. (73) and (74). We note too that we may construct a small world ensemble from a uniformly continuous GRGG according to the procedure described in Sec. IV; from Eqs. (102) with $n=2$ we see that uniformly continuous GRGGs always fall under Case I of Proposition 6.

B. Example 1: A nonuniform uniformly continuous GRGG

We take the space $\mathbb{R}^m$ with Euclidean metric and node distribution with Gaussian density:

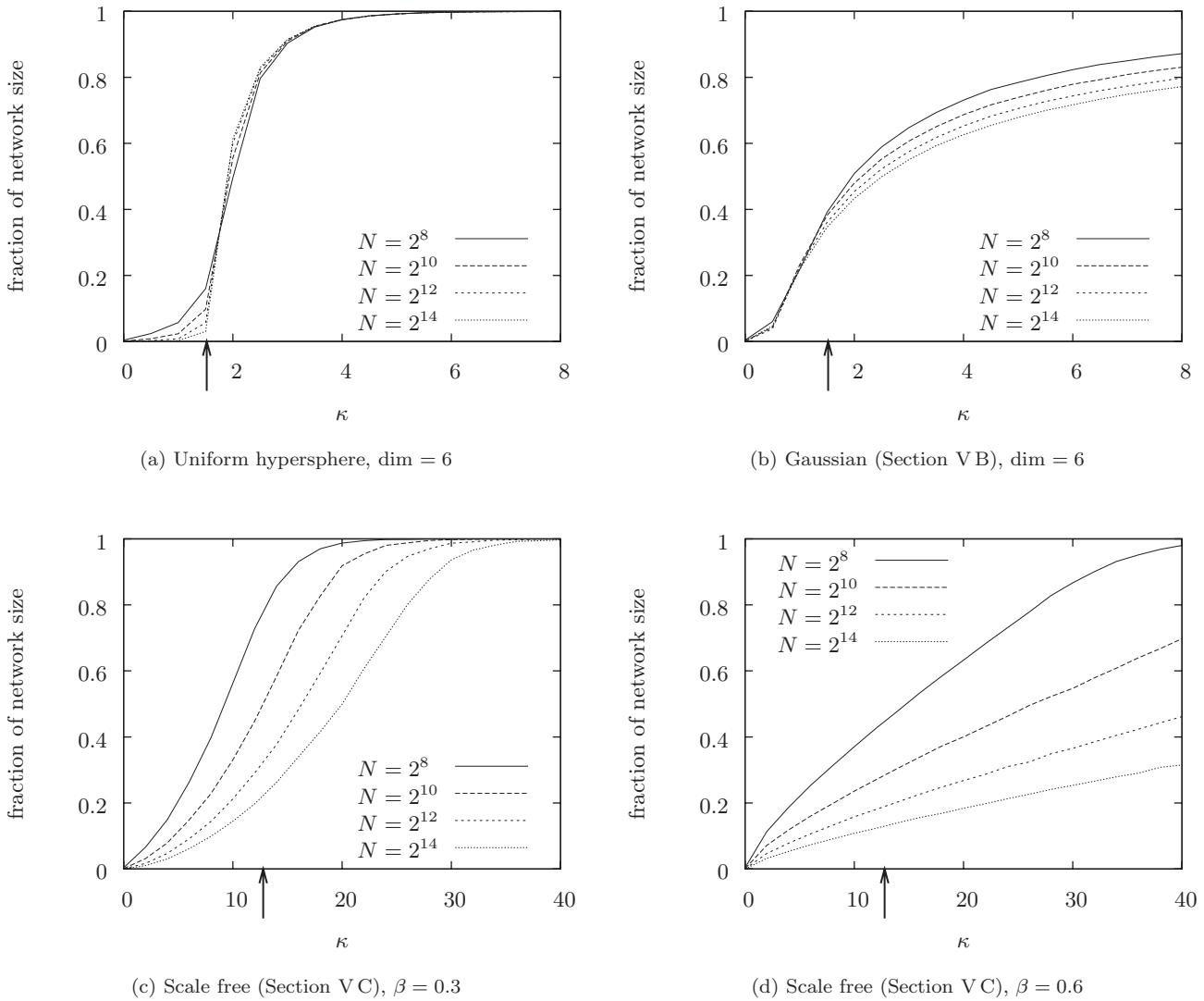
$$p(x_1, \ldots, x_n) = (2\pi)^{-m/2} e^{-(x_1^2+\cdots+x_m^2)/2}. \tag{106}$$

This node distribution is uniformly continuous, but nonuniform and hence non-Poisson. We may calculate:

$$\mu = P\left(\frac{m}{2}, \frac{r^2}{4}\right), \tag{107}$$

$$\approx \frac{2^{-m}}{\Gamma(\frac{m}{2}+1)} r^m, \tag{108}$$

for small r [where $P(a,z) \equiv 1-\Gamma(a,z)/\Gamma(a)$ is the complementary normalized incomplete Gamma function] and in general:



(a) Uniform hypersphere, dim = 6

(b) Gaussian (Section V B), dim = 6

(c) Scale free (Section V C), $\beta = 0.3$

(d) Scale free (Section V C), $\beta = 0.6$

FIG. 5. Fraction of network occupied by largest component (Sec. III F) plotted against mean connectivity with increasing network size, for several GRGGs (Sec. V), estimated in sample. The small arrow marks the value of κ estimated in [14] [Eq. (9)] for the phase transition to the appearance of a giant component for a uniform GRGG of corresponding dimension. Sample sizes are large enough that standard error bars are insignificant.

$$\frac{\mu_n}{\mu^n} \approx \left(\frac{2^n}{n+1}\right)^{m/2} \tag{109}$$

in the large network limit. Setting $n=2$ gives $\mu_2/\mu^2 = (4/3)^{m/2}$; in the sense that var($\Gamma$) measures inhomogeneity of node degree [Sec. III B and Eq. (31)] this suggests (somewhat counter-intuitively) that the ensemble becomes increasingly *inhomogeneous* with increasing dimension. We have also

$$\rho \approx \left[\left(\frac{4}{3}\right)^{m/2} - 1\right]\mu. \tag{110}$$

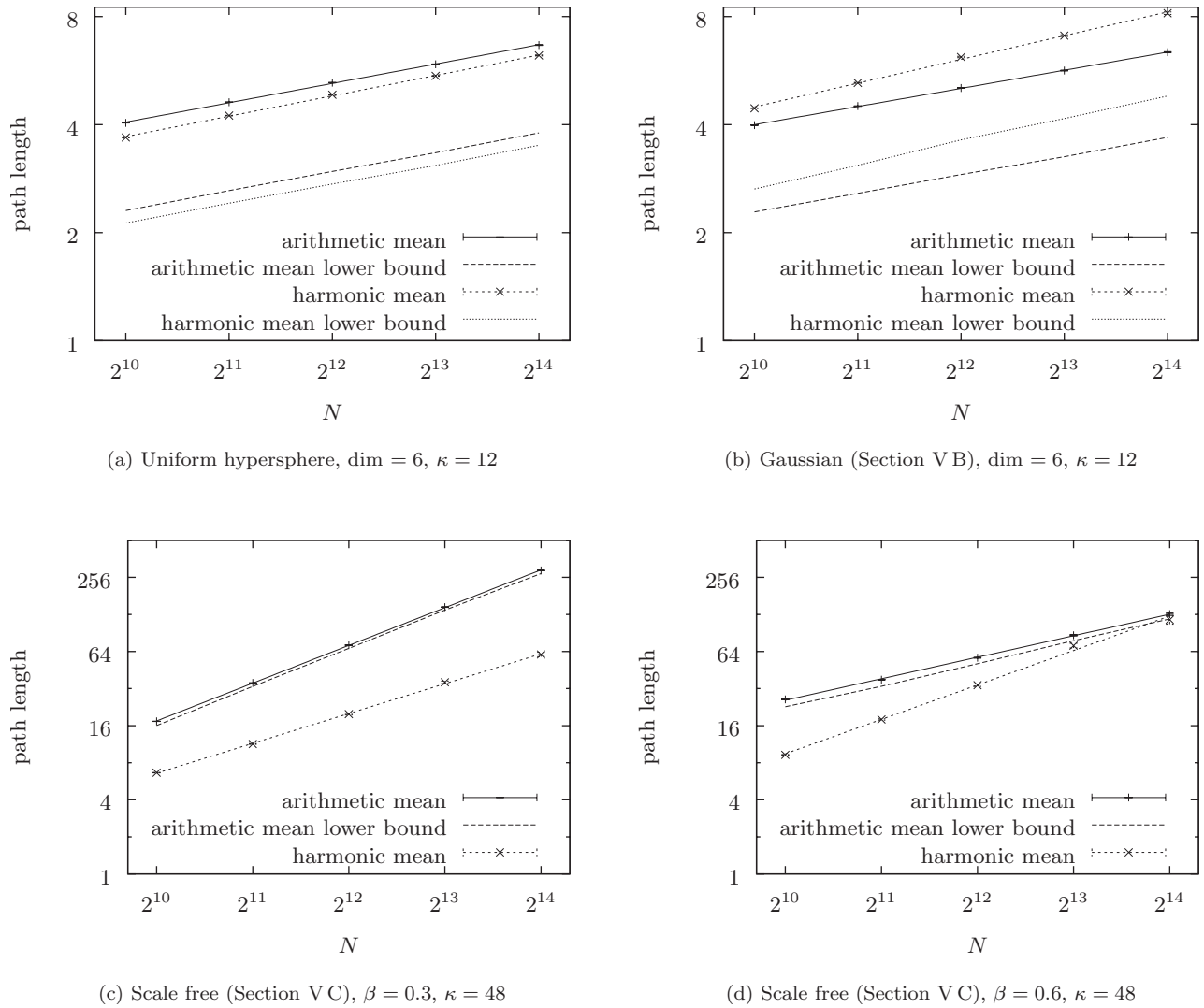
The clustering coefficient is as calculated in [14] for a (Euclidean, uniform) RGG of dimension m. The coefficient $\eta_1 = \eta_2 = \eta$ for the conditional degree correlation is given by

$$\eta = \left(\frac{3}{2}\right)^{m/2} - \left(\frac{4}{3}\right)^{m/2}. \tag{111}$$

Degree correlation is thus always assortative, in accordance with Prop. 8 above. We note that η *increases* with increasing dimension while clustering *decreases* [Eq. (105)]. The overall effect is that degree correlation increases, tending towards 1 in the limit; i.e., with increasing dimension, the effect of inhomogeneity on degree correlation dominates the contribution from clustering.

Figure 4(b) plots the sampled degree distribution for $\kappa = 32$, $N = 2^{16}$ for dimension $m=6$. The corresponding plot for a uniform hypersphere [Fig. 4(a)] is plotted alongside; the deviation from a Poisson distribution is clear.

In Fig. 5(b) the estimated fraction of the network occupied by the largest component is plotted against mean connectivity with increasing network size for dimension $m=6$. Compared with the corresponding plot for a uniform hyper-



(a) Uniform hypersphere, dim = 6, $\kappa = 12$

(b) Gaussian (Section V B), dim = 6, $\kappa = 12$

(c) Scale free (Section V C), $\beta = 0.3$, $\kappa = 48$

(d) Scale free (Section V C), $\beta = 0.6$, $\kappa = 48$

FIG. 6. Conditional arithmetic and harmonic mean path lengths plotted against network size on a log-log scale for several GRGGs (Sec. V) estimated in sample with mean connectivity κ held fixed. The arithmetic (respectively, harmonic) mean lower bound is as calculated from Eq. (73) [respectively, (74)] [for the scale free ensembles the harmonic mean lower bound is omitted since $\mathbf{E}(1/D)$ does not exist in this case].

sphere [Fig. 5(a)], where a phase transition can clearly be seen, this supports our conjecture (Sec. III F) that there should be no phase transition for a nonuniform GRGG and that the size of the largest component increases smoothly with increasing connectivity.

In Fig. 6(b) conditional arithmetic and harmonic mean path lengths are plotted against network size for dimension $m=6$ alongside the corresponding plots for a uniform hypersphere of the same dimension [Fig. 6(a)], for mean connectivity $\kappa=12$. As mentioned earlier, the mean scaling with network size is predicted reasonably accurately by the lower bounds as calculated in Eqs. (73) and (74). In Figs. 7(b) and 7(a) corresponding path lengths are plotted for small world ensembles constructed according to the method described in Sec. IV, with $c=1$. The plots labeled "random graph" show arithmetic mean path lengths for an Erdös-Rényi random graph with the same mean connectivity $\tilde{\kappa}$ as the corresponding small world ensemble; upper bounds are arithmetic mean path lengths for random graphs with mean connectivity Nq as in Eq. (88). In this case the scaling of arithmetic mean path length with network size appears to be reasonably well approximated by that of the corresponding random graph; there is, however, probably insufficient data to be certain.

C. Example 2: A scale free GRGG

Here we choose the underlying space to be the line segment $(0,1]$ with standard metric $d(x,y)=|x-y|$ and degree distribution X with density:

$$p(x) = (1-\beta)x^{-\beta} \quad (\beta < 1). \tag{112}$$

Note that this density (and hence the GRGG) is *not* uniformly continuous. This is basically the situation covered in [15] (Sec. III A), but note that scaling in the large network limit in that paper is *not* the same as ours; while we hold the mean degree κ fixed, there it is assumed that r scales as N^{-1} (in particular we note that under their scheme mean node degree blows up with N for $\beta \geq \frac{1}{2}$).

We have

$$\varphi(x) = \begin{cases} (x+r)^{1-\beta}, & 0 < x \le r \\ (x+r)^{1-\beta} - (x-r)^{1-\beta}, & r < x \le 1-r \\ 1-(x-r)^{1-\beta}, & 1-r < x \le 1. \end{cases} \quad (113)$$

We may calculate

$$\mu = 1-(1-r)^{1-\beta} + (1-\beta)r^{2(1-\beta)}B(1-\beta, 2\beta-1; 1-r), \quad (114)$$

where $B(a,b;z)$ denotes the incomplete Beta function and that, up to leading order for small r:

$$\mu \approx \begin{cases} 2\dfrac{(1-\beta)^2}{1-2\beta}r, & \beta < \dfrac{1}{2} \\ -\dfrac{1}{2}r\ln r, & \beta = \frac{1}{2} \\ (1-\beta)B(1-\beta, 2\beta-1)r^{2(1-\beta)}, & \beta > \dfrac{1}{2}, \end{cases} \quad (115)$$

where $B(a,b)$ is the Beta function. We may calculate in particular that

$$\mu_2 \approx \begin{cases} \dfrac{4(1-\beta)^3}{1-3\beta}r^2, & \beta < \dfrac{1}{3} \\ -\dfrac{32}{27}r^2\ln r, & \beta = \dfrac{1}{3} \\ (1-\beta)[B(1-\beta, -3(1-\beta)) + B(3-2\beta, -3(1-\beta)) - B(2-\beta, -\frac{3}{2}(1-\beta))]r^{3(1-\beta)}, & \beta > \dfrac{1}{3} \end{cases} \quad (116)$$

and in general, for $n \ge 1$ we find that to leading order in r for small r:

$$\mu_n \propto \begin{cases} r^n, & \beta < \dfrac{1}{n+1} \\ -r^n \ln r, & \beta = \dfrac{1}{n+1} \\ r^{(n+1)(1-\beta)}, & \beta > \dfrac{1}{n+1}. \end{cases} \quad (117)$$

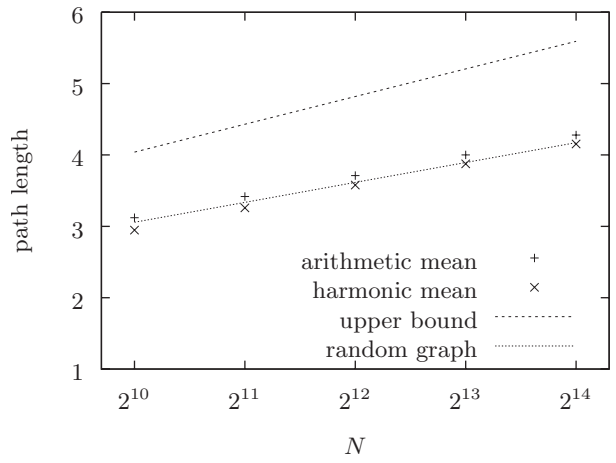
We conclude that for $\beta < \frac{1}{2}$, μ_n/μ^n blows up iff $n > \frac{1}{\beta} - 1$, while for $\beta \ge \frac{1}{2}$, μ_n/μ^n always blows up as $r \to 0$. The ensemble is thus always scale free. Simulations also indicate (as postulated earlier) that $\nu_n = \mu_n$ in the large network limit. We may calculate the first loop moment to be

$$\omega_3 \approx \begin{cases} \dfrac{3(1-\beta)^3}{1-3\beta}r^2, & \beta < \dfrac{1}{3} \\ -\dfrac{8}{9}r^2\ln r, & \beta = \dfrac{1}{3} \\ 2(1-\beta)[B(1-\beta, 3\beta-2) - B(2(1-\beta), 3\beta-2)]r^{3(1-\beta)}, & \beta > \dfrac{1}{3}. \end{cases} \quad (118)$$

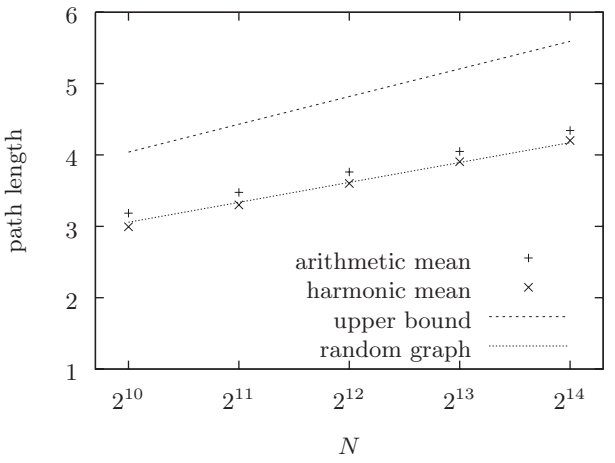
from which we may derive the clustering coefficient $\mathcal{C} \equiv \omega_3/\mu_2$. We find that in the large network limit $\mathcal{C} = \frac{3}{4}$ (the uniformly continuous GRGG value for dimension 1) for $0 < \beta \le \frac{1}{3}$ and thereafter increases smoothly to 1 as β approaches 1. We note that in general, for $\beta < \frac{1}{2}$, our scale free GRGG behaves rather like a *uniformly continuous* GRGG with regard to connectivity moments of order $n < \frac{1}{\beta} - 1$.

Given mean degree κ and network size N, let us set, as in [15], $\alpha \equiv 2Nr$ where r is implicitly defined by solving Eq. (114) for r with $\mu = \kappa/N$. For the actual degree distribution, then, the formula derived in [15] [their Eq. (24)] still furnishes a good approximation and for large k yields the power law $\mathbf{P}(K=k) \propto k^{-1/\beta}$. Figure 4(c) [respectively, 4(d)] plots the sampled degree distribution for $\kappa = 32$, $N = 2^{16}$ for $\beta = 0.3$ (respectively, 0.6). The "squiggles" for large degree are due to finite network size effects.

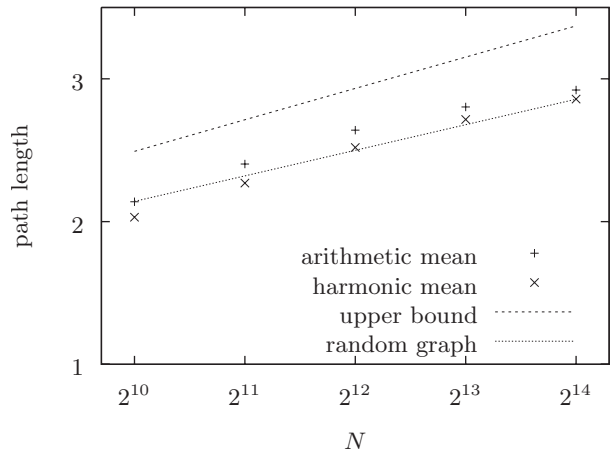
In Fig. 5(c) [respectively, 5(d)] the estimated fraction of the network occupied by the largest component is plotted against mean connectivity with increasing network size for



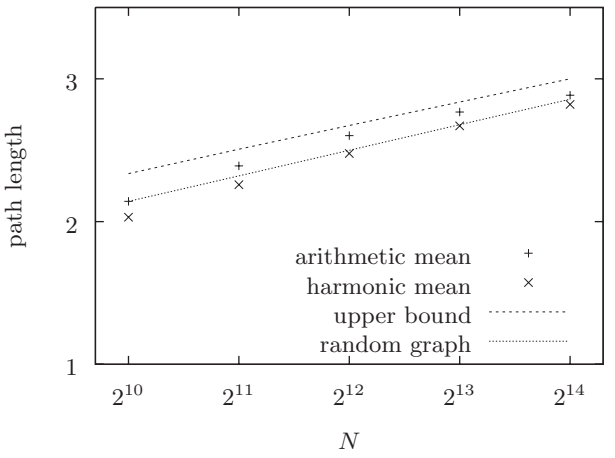
(a) Small world uniform hypersphere, dim = 6, $\tilde{\kappa} = 12$



(b) Small world Gaussian (Section V B), dim = 6, $\tilde{\kappa} = 12$



(c) Small world scale free (Section V C), $\beta = 0.3$, $\tilde{\kappa} = 48$



(d) Small world scale free (Section V C), $\beta = 0.6$, $\tilde{\kappa} = 48$

FIG. 7. Conditional arithmetic and harmonic mean path lengths plotted against network size on a log scale for several small world SERN ensembles constructed from GRGGs (Sec. V) according to the method described in Sec. IV, with $c=1$. The plots labeled "random graph" show arithmetic mean path lengths for a random graph with the same mean connectivity $\tilde{\kappa}$ as the corresponding small world ensemble; upper bounds are arithmetic mean path lengths for random graphs with mean connectivity Nq as in Eq. (88).

$\beta=0.3$ (respectively, 0.6). Again, we see no phase transition (Sec. III F) and the size of the largest component increases smoothly with increasing connectivity. Note too the difference in speed and shape of the approach to 1 between the two cases—it appears that there may be a phase transition (parametrized by β) of scaling behavior of the large component.

In Fig. 6(c) [respectively, 6(d)] conditional arithmetic and harmonic mean path lengths are plotted against network size for $\beta=0.3$ (respectively, 0.6), with mean connectivity fixed at $\kappa=48$. As remarked earlier, the scaling with network size of the arithmetic mean [38] is predicted reasonably accurately by the lower bound as calculated in Eq. (73). Again we may construct a small world ensemble according to the procedure in Sec. IV [Figs. 7(c) and 7(d)]. We find that our scale free GRGG falls into Case I of Proposition 6 precisely when $\beta < \frac{1}{3}$, in which case $\mu_2/\mu^2 \to (1-2\beta)^2/[(1-\beta)(1-3\beta)]$; otherwise μ_2/μ^2 blows up and we have Case II. In either case the resulting small world ensemble is also scale free. In contrast to the uniformly continuous GRGG of the previous section, the form of scaling of path length with network size is not clear.

VI. SUMMARY

A major aim of this study was to set up a viable framework for analysis of the effects exerted by spatial embedding on network structure. In particular, we hope that the formalism of connectivity moments μ_n, ν_n, ω_n (Sec. II B) along with the mean conditional degree Γ (Sec. III B) will prove to furnish useful tools for the analysis of the relationship between spatial embedding and network structure. We note, for example:

(A) Structural properties such as clustering and degree correlations may be expressed directly in terms of the μ_n, ν_n, ω_n.

(B) The degree distribution is completely characterized by the conditional mean degree rv Γ.

(C) Necessary and sufficient conditions for a scale free degree distribution may be expressed succinctly in terms of limiting properties of the μ_n.

Other points of significance we have addressed include the following:

(D) A basic classification of spatial networks into subtypes (Sec. III C).

(E) The severe constraining effects of spatial symmetry on degree distribution (Sec. III C 3).

(F) The effects of inhomogeneity of the spatial node distribution on network structure; for example on degree correlation (Secs. III E and V B) and on the (non)existence of phase transitions (Sec. III F).

(G) A generalization of the Watts-Strogatz construction for deriving a small world network from an arbitrary spatially embedded network with nontrivial clustering, which preserves various structural features of the original network (Sec. IV). Our analysis also illustrates the importance of the connectivity scaling mode in the large network limit and develops some heuristics for calculating bounds on mean path length (Sec. IV).

(H) A generalization of random geometric graphs to more general spatial node distributions (Sec. III C 6). For a particularly well-behaved subclass of such networks (the uniformly continuous GRGGs) we show that uniformity of the spatial node distribution is equivalent to a Poisson degree distribution. We show further that degree correlation is assortative and that nonuniformity in the node distribution contributes positively to degree correlation (Sec. V A). We also demonstrate the analytic tractability of our approach through worked examples of GRGGs (Secs. V B and V C), including a scale free example.

There are, inevitably, open questions and areas requiring clarification and further research. Among these:

(I) We need to develop viable mathematical methods to analyze the formation of a giant component. Of particular importance is to prove (or find counterexamples to) our conjecture that there is no phase transition if the node distribution is nonuniform. Similarly, we currently lack the mathematical tools to derive more accurate estimates of the scaling of mean path length with network size.

(J) We demonstrated that for uniformly continuous GRGGs, degree correlation is always assortative. It seems likely that this may in fact hold for all GRGGs; specifically, if our conjecture that $\mu_n = \nu_n$ for arbitrary GRGGs is correct. Are there, in fact, *any* SERN ensembles with *dis*assortative degree correlation? At present we have no counterexample.

(K) With regard to the classification of subclasses of SERN ensembles, we showed that for uniformly continuous GRGGs uniformity is equivalent to Poisson. It seems reasonable to conjecture that this might hold within a broader class of SERN ensembles. Do there in fact exist SERN ensembles which are Poisson but not uniform, or uniform but not Poisson? Again, we have no example of either. With regard to GRGGs themselves, it also seems quite likely that one might be able to broaden the definition to encompass *any* scaling of the form $\gamma(s) \equiv \hat{\gamma}(s/r)$ [perhaps with some minimal restrictions on the form of the decay function $\hat{\gamma}(s/r)$] while retaining the results of Sec. V.

(L) Regarding scale free SERN ensembles, the only examples we have involve a node distribution which is itself of power law form. Can we find examples of scale free SERN ensembles which are not, in this sense, a case of "power law in, power law out"? We know at least that, if so, they cannot be uniformly continuous GRGGs.

Finally, as regards extending our research, preliminary results suggest that we may indeed broaden the SERN model to include nonindependently distributed nodes (and hence more realistic spatial node distributions), while still retaining a large degree of analytic tractability. We should also like to develop insights into the role of spatial embedding in the dynamics of *processes* taking place on networks and—not least—to investigate the application of our approach to the understanding of real-world spatial networks.

ACKNOWLEDGMENT

This study has been supported by the Spatially Embedded Complex Systems Engineering (SECSE) project, EPSRC Grant No. EP/C51632X/1.

[1] B. Bollobás, *Random Graphs* (Academic, New York, 1985).
[2] D. J. Watts and S. H. Strogatz, Nature (London) **393**, 440 (1998).
[3] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).
[4] B. Waxman, IEEE J. Sel. Areas Commun. **6**, 1617 (1988).
[5] M. Barthélemy, Europhys. Lett. **63**, 915 (2003).
[6] G. Németh and G. Vattay, Phys. Rev. E **67**, 036110 (2003).
[7] S.-H. Yook, H. Jeong, and A.-L. Barabási, Proc. Natl. Acad. Sci. U.S.A. **99**, 13382 (2002).
[8] S. S. Manna and P. Sen, Phys. Rev. E **66**, 066114 (2002).
[9] S. S. Manna, G. Mukherjee, and P. Sen, Phys. Rev. E **69**, 017102 (2004).
[10] S. Muthukrishnan and G. Pandurangan, *Proceedings of the ACM-SIAM Symposium on Discrete Algorithms (SODA)* (Association for Computing Machinaery, New York, 2005).
[11] A. Scirè, I. Tuval, and V. M. Eguíluz, Europhys. Lett. **71**, 318 (2005).
[12] Q. Lu, G. Korniss, and B. K. Szymanski, *Interaction and Emergent Phenomena in Societies of Agents*, Proceedings of the 2006 American Association for Artificial Intelligence Fall Symposium Series (AAAI, Menlo Park, CA, 2006), p. 148.
[13] M. Penrose, *Random Geometric Graphs* (Oxford University Press, Oxford, U.K., 2003).
[14] J. Dall and M. Christensen, Phys. Rev. E **66**, 016121 (2002).
[15] C. Herrmann, M. Barthélemy, and P. Provero, Phys. Rev. E **68**, 026128 (2003).
[16] A. F. Rozenfeld, R. Cohen, D. ben-Avraham, and S. Havlin, Phys. Rev. Lett. **89**, 218701 (2002).
[17] I.e., in the sense that if the Riemannian metric is given by $ds^2 = g_{ij} dx^i dx^j$ in some coordinate system, then $d(x, x+dx) \to ds^2$ uniformly as $|dx| \to 0$.
[18] M. E. J. Newman, SIAM Rev. **45** (2), 167 (2003).

[19] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, Phys. Rep. **424**, 175 (2006).
[20] B. D. Ripley, J. R. Stat. Soc. Ser. B (Methodol.) **39**, 172 (1977).
[21] K.-I. Goh, B. Kahng, and D. Kim, Phys. Rev. Lett. **87**, 278701 (2001).
[22] G. Caldarelli, A. Capocci, P. DeLosRios, and M. A. Muñoz, Phys. Rev. Lett. **89**, 258702 (2002).
[23] B. Söderberg, Phys. Rev. E **66**, 066121 (2002).
[24] M. Boguñá and R. Pastor-Satorras, Phys. Rev. E **68**, 036112 (2003).
[25] If A,B are jointly distributed random variables, then given some b the expectation $\mathbf{E}(A|B=b)$ may be considered a function of b; thus the conditional expectation $\mathbf{E}(A|B)$ is a *random variable*. See, e.g., [29] for a more detailed explanation.
[26] It should be noted that this is distinct from the *adjacency* matrix defined below.
[27] A Bernoulli trial is a random variable taking only the values 0 or 1; it may be thought of as a "biased coin toss."
[28] In principle there is nothing to stop us from scaling either metric or node distribution in the large network limit; indeed, there may be physically compelling grounds for doing so in some application of the model.
[29] D. Stirzaker, *Elementary Probability* (Cambridge University Press, Cambridge, U.K., 1995).
[30] T. M. Apostol, *Mathematical Analysis* (Addison-Wesley, Reading, MA, 1974).
[31] M. E. J. Newman, S. H. Strogatz, and D. J. Watts, Phys. Rev. E **64**, 026118 (2001).
[32] Although it is, admittedly, frequently difficult to establish in simulation whether we are observing a phase transition or not, given inevitable limitations of network size enforced by computational resources. We note that phase transitions are generally more sharply defined in higher dimensions.
[33] A. Fronczak, P. Fronczak, and J. A. Holyst, Phys. Rev. E **70**, 056110 (2004).
[34] Which it may well fail to do; e.g., for the scale-free GRGG of Sec. V C.
[35] S. Jespersen and A. Blumen, Phys. Rev. E **62**, 6270 (2000).
[36] The (arithmetic) mean path length for Erdös-Rényi random graphs may be calculated from the exact formula derived in [33] [Eq. (19)].
[37] J. M. Lee, *Introduction to Smooth Manifolds* (Springer, Berlin, 2003).
[38] There is no corresponding *harmonic* mean lower bound in this case, as $\mathbf{E}(1/D)$ [see Eq. (74), Sec. IV] fails to converge.