Internet Mathematics Vol. 5, No. 1–2: 89–109

Directed Random Dot Product Graphs

Stephen J. Young and Edward Scheinerman

Abstract. In this paper we consider three models for random graphs that utilize the inner product as their fundamental object. We analyze the behavior of these models with respect to clustering, the small world property, and degree distribution. These models are motivated by the random dot product graphs developed by Kraetzl, Nickel, and Scheinerman. We extend their results to fully parameterize the conditions under which clustering occurs, characterize the diameter of graphs generated by these models, and describe the behavior of the degree distribution.

1. Introduction

With the ubiquity and importance of the Internet and genetic information in medicine and biology, the study of complex networks relating to the Internet and genetics continues to be an important and vital area of study. This is especially true for networks such as the physical layer of the Internet, the link structure of the world wide web, and protein–protein and protein–gene interaction networks.

Because of the size of these networks [Albert and Barabási 02] and the difficulty of determining complete link information [Achlioptas et al. 05, Lakhina et al. 03], a significant amount of research has gone into finding models that match observed properties of these graphs in order to empirically (via simulation) and theoretically understand and predict properties of these complex networks. There are three models that, together with their variations, are the core models for these complex networks [Bornholdt and Schuster 03, Durrett 07].

1542-7951/08 $0.50 per page

The configurational model and its variants attempt to generate complex networks by specifying the degree sequence and creating edges randomly with respect to that degree sequence. On the other hand, the Barabási–Albert preferential attachment model attempts to model the process by which the network grows. Specifically, it posits that vertices with high degree are more likely to increase in degree when a new vertex is added to the network. In a similar vein, the copying model [Chung et al. 03, Kumar et al. 00] also attempts to model the growth process of a complex network. However, the copying model takes the more distinctly biological viewpoint of replication of existing nodes combined with mutation.

All three of these types of models have had success in reproducing the hallmark features of complex networks, namely a power-law degree distribution, a diameter that grows slowly or is constant with the size of the graph, and one of several clustering properties; see [Bornholdt and Schuster 03, Durrett 07] for a collection of such results.

However, there are many other aspects of complex networks that fail to be captured by these models, for example nonuniform assortativity [Newman 02] and the existence of directed cycles. Thus there is considerable interest in new models for complex networks that exhibit a power-law-like degree sequence, small diameter, and clustering, and are different enough from the three main model classes to exhibit other properties of complex networks that are not exhibited by the current models.

One potential method to create new models is to incorporate geometry into already existing models. Flaxman, Frieze, and Vera [Flaxman et al. 06] used geometry coupled with the preferential attachment model to create a model that generates a random power-law graph that has small separators.

Taking this idea one step further, one can add semantic information to an already existing model. One such model is the random dot product graph model applied by Caldarelli et al. and Azar et al. [Azar et al. 01, Caldarelli et al. 02] and formalized by Kraetzl, Nickel, Scheinerman, and Tucker [Kraetzl et al. 05a, Kraetzl et al. 05b]. In their work they assign to each vertex a vector in $\mathbb{R}^d$, and then any edge is present with probability equal to the dot product of the endpoints. Thus, with the vertices thought of as members of a social network, the vectors together with the dot product encode semantically the idea of differing "interests" and varying levels of "talkativeness."

We discuss the two natural generalizations of the random dot product; specifically, we remove the restrictions on the vectors imposed in earlier work and develop directed generalization. In Section 2 we review the model and results of Kraetzl, Nickel, and Scheinerman, and look at the most natural generalization of their model. Then in Section 3 we examine the directed generalization of the

model in Section 2. We follow by briefly discussing in Section 4 a model with less independence than that of Section 3. Finally, we look at some directions for future research and applications of the generalized models in Section 5.

2. Undirected Graph Model

In their paper, Kraetzl, Nickel, and Scheinerman lay out a general framework for constructing a class of random graphs that incorporate geometry. Each vertex is assigned a vector v randomly from some subset S of $\mathbb{R}^d$. Then each edge $\{i, j\}$ is present independently with probability $f(v_i, v_j)$, where $f\colon S \times S \longrightarrow [0, 1]$. The full generality of this class of models is, however, difficult to analyze, so they consider a specific instance of this model, namely, each coordinate of a vector associated with a vertex is $\mathcal{U}^\alpha[0, 1]/\sqrt{d}$, where α is a parameter of the model and d is the dimension of the vector. Then the function f is the inner product of the two vectors; this has the advantage of being easy to compute and symmetric in the variables, i.e., $f(v_i, v_j) = f(v_j, v_i)$. By considering this restricted model, Kraetzl et al. were able to show that with $d = 1$:

1. $\left(\frac{\alpha+1}{2\alpha+1}\right)^2 = \mathbb{P}\left(u \sim w \mid u \sim v \sim w\right) > \mathbb{P}\left(u \sim w\right) = 1/(\alpha+1)^2$.
2. If $\lambda(k)$ is the random variable indicating the number of vertices of degree k, then $\mathbb{E}\left[\lambda(k)\right] \sim \frac{1}{k!\alpha}(1+\alpha)^{1/\alpha}\Gamma\left(\frac{1}{\alpha}+k\right) n^{(\alpha-1)/\alpha}$ for $k \in \mathbb{Z}^+$ as $n \to \infty$.
3. The giant component has diameter at most 6 as $n \to \infty$.

In order to provide perspective for the rest of this paper, we consider the following natural generalization and corresponding proofs. Consider some distribution $\mathbf{X}$ on a subset of $\mathbb{R}^d$ such that $X_i^T X_j \in (0, 1)$ almost surely, where X_i and X_j are distributed as $\mathbf{X}$. Then generate a graph in the following manner: each vertex v has a random variable X_v associated with it distributed as $\mathbf{X}$, and then each edge $\{i, j\}$ is present independently with probability $X_i^T X_j$. Let $G(\mathbf{X}, n)$ denote such a graph generated on n vertices. For clarity of presentation, we will abuse notation and say that a vertex belongs to the region R if the vector associated with the vertex lies in R. We also denote by $(x_u)_i$ the ith component of the random variable X_u, and by $\mathbf{x}_i$ the distribution of the ith component of the distribution $\mathbf{X}$.

We begin by considering the clustering present in the random graph model $G(\mathbf{X}, n)$. Due to the unknown distribution of $\mathbf{X}$ in the model, we are unable to produce an explicit clustering coefficient as with the Kraetzl, Nickel, Scheinerman model. However, in the following proposition we are able to show that there is

positive clustering unless $G(\mathbf{X}, n)$ is essentially an Erdős–Renyi graph model [Bollobás 98].

Proposition 2.1. *Let $G = G(\mathbf{X}, n)$, where $\mathbf{X}$ is a distribution on $\mathbb{R}^d$ such that for any two random variables X_i and X_j distributed as $\mathbf{X}$, $X_i^T X_j \in (0,1)$ almost surely. Then for any vertices $u, v, w \in V(G)$, we have that $\mathbb{P}(u \sim w \mid u \sim v, v \sim w) \geq \mathbb{P}(u \sim w)$, with equality if and only if $\mathbf{X}$ is almost surely constant.*

Proof. Notice that $\mathbb{P}(u \sim v) = \mathbb{P}(v \sim w) = \mathbb{P}(u \sim w)$ and furthermore, since the vectors X_u and X_v are independent, $\mathbb{P}(u \sim v) = \mathbb{E}[\mathbf{X}]^T \mathbb{E}[\mathbf{X}]$. Similarly we have that $\mathbb{P}(u \sim v, v \sim w) = \mathbb{E}[\mathbf{X}]^T \mathbb{E}\left[\mathbf{X}\mathbf{X}^T\right] \mathbb{E}[\mathbf{X}]$. Observe that in order to show that $\mathbb{P}(u \sim w \mid u \sim v, v \sim w) \geq \mathbb{P}(u \sim w)$, it suffices to show that

$$\mathbb{P}(u \sim w, u \sim v, v \sim w) - \mathbb{P}(u \sim v, v \sim w)\,\mathbb{P}(u \sim w) \geq 0.$$

Thus consider

$$\begin{aligned}
&\mathbb{P}(u \sim w, u \sim v, v \sim w) - \mathbb{P}(u \sim v, v \sim w)\,\mathbb{P}(u \sim w) \\
&\quad = \mathbb{E}\left[X_u^T \mathbb{E}\left[\mathbf{X}\mathbf{X}^T\right]^2 X_u\right] - \mathbb{E}[\mathbf{X}]^T \mathbb{E}\left[\mathbf{X}\mathbf{X}^T\right] \mathbb{E}[\mathbf{X}]\,\mathbb{E}[\mathbf{X}]^T \mathbb{E}[\mathbf{X}] \\
&\quad = \sum_{i,j,k=1}^{d} \mathbb{E}\left[(x_u)_i\,\mathbb{E}[\mathbf{x}_i\mathbf{x}_j]\,\mathbb{E}[\mathbf{x}_j\mathbf{x}_k]\,(x_u)_k\right] - \mathbb{E}[\mathbf{x}_i]\,\mathbb{E}[\mathbf{x}_i\mathbf{x}_j]\,\mathbb{E}[\mathbf{x}_j]\,\mathbb{E}[\mathbf{x}_k]^2 \\
&\quad = \sum_{i,j,k=1}^{d} \mathbb{E}[\mathbf{x}_i\mathbf{x}_j]\left(\mathbb{E}[\mathbf{x}_j\mathbf{x}_k]\,\mathbb{E}[\mathbf{x}_i\mathbf{x}_k] - \mathbb{E}[\mathbf{x}_i]\,\mathbb{E}[\mathbf{x}_j]\,\mathbb{E}[\mathbf{x}_k]^2\right).
\end{aligned}$$

Now note that $\operatorname{cov}(\mathbf{X}) = \mathbb{E}\left[\mathbf{X}\mathbf{X}^T\right] - \mathbb{E}[\mathbf{X}]\,\mathbb{E}[\mathbf{X}]^T$ is a symmetric positive semidefinite matrix, and thus there is an orthonormal matrix Q such that $Q \operatorname{cov}(\mathbf{X})\, Q^T$ is diagonal. But $\operatorname{cov}(QX) = \mathbb{E}\left[Q\mathbf{X}\mathbf{X}^T Q^T\right] - \mathbb{E}[Q\mathbf{X}]\,\mathbb{E}\left[\mathbf{X}^T Q^T\right] = Q \operatorname{cov}(\mathbf{X})\, Q^T$ and $(QX_i)^T(QX_j) = X_i^T Q^T Q X_j = X_i^T X_j$, so we may assume without loss of generality that $\operatorname{cov}(\mathbf{X})$ is diagonal. In particular, if $i \neq j$, then $\mathbb{E}[\mathbf{x}_i\mathbf{x}_j] = \mathbb{E}[\mathbf{x}_i]\,\mathbb{E}[\mathbf{x}_j]$. Thus we have that if $i \neq k$ and $j \neq k$, then

$$\mathbb{E}[\mathbf{x}_i\mathbf{x}_k]\,\mathbb{E}[\mathbf{x}_k\mathbf{x}_j] - \mathbb{E}[\mathbf{x}_i]\,\mathbb{E}[\mathbf{x}_k]^2\,\mathbb{E}[\mathbf{x}_j] = 0.$$

Furthermore if $i = k \neq j$, then

$$\mathbb{E}[\mathbf{x}_i\mathbf{x}_k]\,\mathbb{E}[\mathbf{x}_k\mathbf{x}_j] - \mathbb{E}[\mathbf{x}_i]\,\mathbb{E}[\mathbf{x}_k]^2\,\mathbb{E}[\mathbf{x}_j] = \mathbb{E}[\mathbf{x}_i]\,\mathbb{E}[\mathbf{x}_j]\operatorname{Var}(\mathbf{x}_i),$$

and similarly for $j = k \neq i$. Note as well that if $i = j = k$, then

$$\mathbb{E}[\mathbf{x}_i\mathbf{x}_k]\,\mathbb{E}[\mathbf{x}_k\mathbf{x}_j] - \mathbb{E}[\mathbf{x}_i]\,\mathbb{E}[\mathbf{x}_k]^2\,\mathbb{E}[\mathbf{x}_j] = \mathbb{E}\left[\mathbf{x}_i^2\right]^2 - \mathbb{E}[\mathbf{x}_i]^4.$$

Using these equalities, we have that

$$\begin{aligned}
&\mathbb{P}(u \sim w, u \sim v, v \sim w) - \mathbb{P}(u \sim v, v \sim w)\,\mathbb{P}(u \sim w) \\
&= \sum_{i<j} \mathbb{E}[\mathbf{x}_i \mathbf{x}_j]\,\mathbb{E}[\mathbf{x}_i]\,\mathbb{E}[\mathbf{x}_j]\,(\mathrm{Var}(\mathbf{x}_i) + \mathrm{Var}(\mathbf{x}_j)) + \sum_{k=1}^{d} \mathbb{E}\left[\mathbf{x}_k^2\right]\left(\mathbb{E}\left[\mathbf{x}_k^2\right]^2 - \mathbb{E}[\mathbf{x}_k]^4\right) \\
&= \sum_{i<j} \mathbb{E}[\mathbf{x}_i]^2\,\mathbb{E}[\mathbf{x}_j]^2\,(\mathrm{Var}(\mathbf{x}_i) + \mathrm{Var}(\mathbf{x}_j)) + \sum_{k=1}^{d} \mathbb{E}\left[\mathbf{x}_k^2\right]\mathrm{Var}(\mathbf{x}_k)\left(\mathbb{E}\left[\mathbf{x}_k^2\right] + \mathbb{E}[\mathbf{x}_k]^2\right) \\
&\geq 0.
\end{aligned}$$

Hence $\mathbb{P}(u \sim w \mid u \sim v, v \sim w) \geq \mathbb{P}(u \sim w)$.

Clearly, equality holds if $\mathbf{X}$ is almost surely constant, since then $\mathrm{Var}(\mathbf{x}_i) = 0$ for all i. Now suppose that equality holds. Then for every i,

$$\mathbb{E}\left[\mathbf{x}_i^2\right]\mathrm{Var}(\mathbf{x}_i)\left(\mathbb{E}\left[\mathbf{x}_i^2\right] + \mathbb{E}[\mathbf{x}_i]^2\right) = 0,$$

and so either $\mathrm{Var}(\mathbf{x}_i) = 0$ or $\mathbb{E}\left[\mathbf{x}_i^2\right] = 0$. But in both these cases $\mathbf{x}_i$ is almost surely constant, and thus equality holds if and only if $\mathbf{X}$ is almost surely constant. □

Let $\lambda(k)$ be the number of vertices of degree k in the random graph model. In [Kraetzl et al. 05a], the asymptotics of $\mathbb{E}[\lambda(k)]$ are shown to be

$$\frac{1}{k!\alpha}(1+\alpha)^{1/\alpha}\,\Gamma\left(\frac{1}{\alpha} + k\right) n^{(\alpha-1)/\alpha},$$

where α is the power of the uniform distribution used in the model. We do not explicitly consider $\mathbb{E}[\lambda(k)]$ but rather consider $\mathbb{P}(\deg(v) = k)$. However, we note that $\mathbb{E}[\lambda(k)] = n\mathbb{P}(\deg(v) = k)$, and so these measures are equivalent.

Proposition 2.2. *Let $\mathbf{X}$ be a random variable and let $G = G(\mathbf{X}, n)$. Then the probability that a vertex has degree k in G is*

$$\int \binom{n-1}{k} \left(\mathbb{E}[\mathbf{X}]^T X\right)^k \left(1 - \mathbb{E}[\mathbf{X}]^T X\right)^{n-1-k} d\mathbf{X}.$$

Proof. Let $v \in V(G)$ be such that v has vector X_v. Then

$$\begin{aligned}
\mathbb{P}\left(\deg(v) = k \mid X_v\right) &= \int \mathbb{P}\left(\deg(v) = k \mid X_{V(G)}\right) (d\mathbf{X})^{n-1} \\
&= \int \sum_{\substack{S \subseteq V(G)-\{s\} \\ |S|=k}} \prod_{i \in S} X_i^T X_v \prod_{i \notin S \cup \{v\}} \left(1 - X_i^T X_v\right) (dX)^{n-1} \\
&= \sum_{\substack{S \subseteq V(G)-\{s\} \\ |S|=k}} \int \prod_{i \in S} X_i^T X_v \prod_{i \notin S \cup \{v\}} \left(1 - X_i^T X_v\right) (dX)^{n-1} \\
&= \sum_{\substack{S \subseteq V(G)-\{s\} \\ |S|=k}} \prod_{i \in S} \mathbb{E}\left[\mathbf{X}\right]^T X_v \prod_{i \notin S \cup \{v\}} \left(1 - \mathbb{E}\left[\mathbf{X}\right]^T X_v\right) \\
&= \binom{n-1}{k} \left(\mathbb{E}\left[\mathbf{X}\right]^T X_v\right)^k \left(1 - \mathbb{E}\left[\mathbf{X}\right]^T X_v\right)^{n-1-k}.
\end{aligned}$$

Now integrating to remove the conditioning on X_v, we get

$$\int \mathbb{P}\left(\deg(v) = k \mid X_v\right) d\mathbf{X} = \int \binom{n-1}{k} \left(\mathbb{E}\left[\mathbf{X}\right]^T X\right)^k \left(1 - \mathbb{E}\left[\mathbf{X}\right]^T X\right)^{n-1-k} d\mathbf{X}.$$

□

Corollary 2.3. *Let* $\mathbf{X}$ *be distributed like* $\mathcal{U}^\alpha[0,1]$. *Then if* v *is a vertex in* $G = G(\mathbf{X}, n)$, *then* $\mathbb{P}\left(\deg(v) = k\right)$ *asymptotically is* $\left(\frac{1+\alpha}{n}\right)^{1/\alpha} \frac{\Gamma(k+1+\frac{1}{\alpha})}{k!}$.

Proof. First note that $\mathbb{E}\left[X\right] = 1/(1+\alpha)$. Thus we have that

$$\begin{aligned}
\mathbb{P}\left(\deg(v) = k\right) &= \int \binom{n-1}{k} \left(\mathbb{E}\left[\mathbf{X}\right]^T X\right)^k \left(1 - \mathbb{E}\left[\mathbf{X}\right]^T X\right)^{n-1-k} d\mathbf{X} \\
&= \binom{n-1}{k} \int_0^1 \left(\frac{x}{1+\alpha}\right)^k \left(1 - \frac{x}{1+\alpha}\right)^{n-1-k} \frac{1}{\alpha} x^{\frac{1}{\alpha}-1} dx \\
&= \binom{n-1}{k} \sum_{j=0}^{n-1-k} \binom{n-1-k}{j} \frac{(-1)^j}{\alpha(1+\alpha)^{k+j}} \int_0^1 x^{k+j+\frac{1}{\alpha}-1} dx \\
&= \sum_{j=0}^{n-1-k} \binom{n-1}{k, j, n-1-k-j} \frac{(-1)^j}{(\alpha k + \alpha j + 1)(1+\alpha)^{k+j}} \\
&= \binom{n-1}{k} \frac{{}_2F_1\left(\begin{matrix} k+\frac{1}{\alpha}, k+1-n \\ k+1+\frac{1}{\alpha} \end{matrix}; \frac{1}{1+\alpha}\right)}{(\alpha k + 1)(1+\alpha)^k},
\end{aligned}$$

where ${}_2F_1\left(\begin{smallmatrix}a,b\\c\end{smallmatrix};z\right)$ is Gauss's hypergeometric function. We note that

$$ {}_2F_1\left(\begin{matrix}a,b\\c\end{matrix};z\right) = (1-z)^{-a}\,{}_2F_1\left(\begin{matrix}a,c-b\\c\end{matrix};\frac{z}{z-1}\right) $$

and that

$$ \begin{aligned} {}_2F_1\left(\begin{matrix}a,b\\c\end{matrix};z\right) &= e^{a\pi i}\frac{\Gamma(c)\Gamma(b-c+1)}{\Gamma(a+b-c+1)\Gamma(c-a)}z^{-a}\,{}_2F_1\left(\begin{matrix}a,a-c+1\\a+b-c+1\end{matrix};\frac{z-1}{z}\right)\\ &\quad+\frac{\Gamma(c)\Gamma(b-c+1)}{\Gamma(a)\Gamma(b-a+1)}z^{a-c}(1-z)^{c-a-b}\,{}_2F_1\left(\begin{matrix}1-a,c-a\\b-a+1\end{matrix};\frac{1}{z}\right) \end{aligned} $$

[Abramowitz and Stegun 64, Temme 03]. Then using these relations, we obtain

$$ \begin{aligned} &{}_2F_1\left(\begin{matrix}k+\frac{1}{\alpha},k+1-n\\k+1+\frac{1}{\alpha}\end{matrix};\frac{1}{1+\alpha}\right)\\ &\quad=\left(\frac{1+\alpha}{\alpha}\right)^{k+\frac{1}{\alpha}}\,{}_2F_1\left(\begin{matrix}k+\frac{1}{\alpha},\frac{1}{\alpha}+n\\k+1+\frac{1}{\alpha}\end{matrix};\frac{-1}{\alpha}\right)\\ &\quad=(1+\alpha)^{k+\frac{1}{\alpha}}\frac{\Gamma(k+1+\frac{1}{\alpha})\Gamma(n-k)}{\Gamma(n+\frac{1}{\alpha})}\,{}_2F_1\left(\begin{matrix}k+\frac{1}{\alpha},0\\\frac{1}{\alpha}+n\end{matrix};1+\alpha\right)\\ &\qquad-\frac{\alpha k+1}{\alpha^2(n-k)}\left(\frac{1+\alpha}{\alpha}\right)^{1-n+k}\,{}_2F_1\left(\begin{matrix}1-k-\frac{1}{\alpha},1\\n-k+1\end{matrix};-\alpha\right). \end{aligned} $$

But ${}_2F_1\left(\begin{smallmatrix}k+\frac{1}{\alpha},0\\\frac{1}{\alpha}+n\end{smallmatrix};1+\alpha\right)=1$ by definition. Thus we have that

$$ \begin{aligned} \mathbb{P}\left(\deg(v)=k\right) =&\frac{(1+\alpha)^{\frac{1}{\alpha}}\Gamma(k+1+\frac{1}{\alpha})\Gamma(n)}{k!\Gamma(n+\frac{1}{\alpha})}\\ &-\binom{n-1}{k}\frac{\alpha^{n-1}\,{}_2F_1\left(\begin{matrix}1-k-\frac{1}{\alpha},1\\n-k+1\end{matrix};-\alpha\right)}{\alpha^{k-2}(1+\alpha)^{n-1}(n-k)}. \end{aligned} $$

Now using an integral definition of the hypergeometric function [Temme 03], we have that

$$ \begin{aligned} &1\le {}_2F_1\left(\begin{matrix}1-k-\frac{1}{\alpha},1\\n-k+1\end{matrix};-\alpha\right)\\ &\quad=(n-k)\int_0^\infty\frac{(1+\alpha-\alpha e^{-t})^{k-1+1/\alpha}}{e^{(n-k)t}}dt\le(1+\alpha)^{k-1+1/\alpha}. \end{aligned} $$

Thus as $n\to\infty$, the second term goes to zero in an exponential manner, and hence we have that asymptotically,

$$ \mathbb{P}\left(\deg(v)=k\right)\sim\frac{(1+\alpha)^{1/\alpha}\Gamma(k+1+\frac{1}{\alpha})\Gamma(n)}{k!\Gamma(n+\frac{1}{\alpha})}. $$

However, using Stirling's approximation, we have

$$\frac{\Gamma(n)}{\Gamma(n+\frac{1}{\alpha})} \sim \frac{\sqrt{\frac{2\pi}{n}}\left(\frac{n}{e}\right)^n}{\sqrt{\frac{2\pi}{n+\frac{1}{\alpha}}}\left(\frac{n+\frac{1}{\alpha}}{e}\right)^{n+1/\alpha}} = \sqrt{1+1/\alpha n}\left(1+\frac{1}{\alpha n}\right)^{-n} e^{1/\alpha}\left(n+\frac{1}{\alpha}\right)^{-1/\alpha},$$

which behaves asymptotically like $n^{-1/\alpha}$ as n approaches infinity. Thus the asymptotic behavior of $\mathbb{P}(\deg(v)=k)$ is $\left(\frac{1+\alpha}{n}\right)^{1/\alpha}\frac{\Gamma(k+1+\frac{1}{\alpha})}{k!}$. Furthermore, as k gets large this approaches $(\frac{k(1+\alpha)}{n})^{1/\alpha}$. □

We now move on to consider the asymptotic behavior of the diameter of $G(\mathbf{X},n)$. If $\mathbf{X}$ is constant, $G(\mathbf{X},n)$ reduces to an Erdős–Renyi random graph with parameter $\mathbf{X}^T\mathbf{X}$. Thus it is clear that the asymptotic diameter of $G(\mathbf{X},n)$ depends on the variability of $\mathbf{X}$. Specifically, it is the variability of $\mathbf{X}$ that results in a diameter greater than 2 in the asymptotic limit. Thus, it is apparent that the nature of the probability distribution $\mathbf{X}$ should drive the diameter of $G(\mathbf{X},n)$ and hence drive any statement on the diameter of $G(\mathbf{X},n)$. Thus we will appeal to the geometry of the distribution of $\mathbf{X}$ in order to show that the diameter of an arbitrarily large fraction of $G(\mathbf{X},n)$ is almost surely at most 5. First, however, we need the following preliminary lemma.

Lemma 2.4. *Let $\mathbf{X}$ be a distribution on $\mathbb{R}^d$ such that for random variables X_u and X_v distributed as $\mathbf{X}$, $X_u^T X_v \in (0,1)$ almost surely. Then $\|\mathbf{X}\| \leq 1$ almost surely.*

Proof. Suppose not; then $\mathbb{P}(\|\mathbf{X}\|>1)\neq 0$. We can then partition the support of $\mathbf{X}$ by the shells $A_i = \{x\in\mathbb{R}^d \mid i<\|x\|\leq i+1\}$ for $i\in\mathbb{Z}^+$. Since there are countably many such shells and $\mathbb{P}(\mathbf{X}\in\cup_i A_i)=\mathbb{P}(\|\mathbf{X}\|>1)\neq 0$, there exists some i such that $\mathbb{P}(\mathbf{X}\in A_i)\neq 0$. But then this shell can be partitioned into finitely many angular regions such that any pair of points in a region have angular distance at most $\arccos(1/i^2)$. Furthermore, there is at least one such angular region, call it R, such that $\mathbf{X}$ lies in R with positive probability. Now note that for any two vectors $x,y\in R$, $x^Ty > i^2\cos(\arccos(1/i^2))>1$. Then if X_u and X_v are distributed as $\mathbf{X}$, we have that $\mathbb{P}(X_u,X_v\in R)\neq 0$ and hence $\mathbb{P}(X_u^TX_v\notin(0,1))>0$, a contradiction. Thus $\|X\|$ is almost surely at most 1. □

We note that in a similar fashion, $\|\mathbf{X}\|>0$ almost surely; specifically, $\mathbf{X}$ is almost surely not 0.

Remark 2.5. We denote by $B(c;r)$ (respectively $\overline{B}(c;r)$) the open (respectively closed) ball of radius r centered at c.

This result allows us to generalize naturally in many ways the proof of a bounded-diameter giant component that appears in [Kraetzl et al. 05a]. In particular, in the following theorem we are able to use the compactness of $\mathbf{X}$ to break $G(\mathbf{X}, n)$ into smaller Erdős–Renyi-like subgraphs, which have known asymptotic diameter.

Theorem 2.6. *Let $\mathbf{X}$ be a random variable over $\mathbb{R}^d$ and let $G = G(\mathbf{X}, n)$. Then an arbitrarily large fraction of G almost surely is connected of diameter at most 5.*

Proof. By Lemma 2.4, $\|\mathbf{X}\|$ is almost surely at most 1. Thus we may assume without loss of generality that $\mathbf{X} \in \overline{B}(0;1)$. Let $0 < \delta < \frac{1}{4}$ and choose $\epsilon > 0$ such that $\mathbb{P}(\mathbf{X} \in B(0;\epsilon)) < \delta$. Then let A be the closed annulus $\overline{B}(0;1) - B(0;\epsilon)$. For all $\alpha \in A$, choose

$$r_\alpha \in \left\{ r > 0 \mid \forall x, y \in B(\alpha; r), x^T y > \frac{\epsilon^2}{4} \right\},$$

which is nonempty by the continuity of the inner product. Then $\cup_{\alpha \in A} B(\alpha; r_\alpha)$ is an open cover of A, and since A is compact, there exists a finite cover, say $\{B(\alpha_i; r_{\alpha_i})\}$.

Fix i such that $\mathbb{P}(\mathbf{X} \in B(\alpha_i; r_{\alpha_i})) \neq 0$. Then as $n \to \infty$, there are almost surely infinitely many vertices that lie in $B(\alpha_i; r_{\alpha_i})$. It then follows from a result of Erdős and Renyi, since the probability of every edge is at least $\epsilon^2/4$ and for fixed $\{X_v\}$ each edge is present independently, that the graph induced by $B(\alpha_i; r_{\alpha_i})$ is almost surely of diameter at most 2. Clearly, if $\mathbb{P}(X \in B(\alpha_i; r_{\alpha_i})) = 0$, then there are almost surely no vertices in that region, and moreover, those regions do not affect the diameter of $G(\mathbf{X}, n)$.

Now consider two regions $\mathcal{R}_i = B(\alpha_i; r_{\alpha_i})$ and $\mathcal{R}_j = B(\alpha_j; r_{\alpha_j})$ occurring with positive probability. There is a naturally defined probability measure on $\mathcal{R}_i \times \mathcal{R}_j$. Furthermore, since $\mathbb{P}\left(X_i^T X_j = 0\right) = 0$, there exist $\hat{\epsilon}$, $\hat{\delta} > 0$ such that $\mathbb{P}\left(X_i^T X_j > \hat{\delta} \mid X_i \in \mathcal{R}_i, X_j \in \mathcal{R}_j\right) > \hat{\epsilon}$. But since $\hat{\delta}$ and $\hat{\epsilon}$ are independent of n, and $\mathcal{R}_i$ and $\mathcal{R}_j$ almost surely contain an infinite number of vertices, there is almost surely an edge e_{ij} between the regions.

Now given vertices $u \in \mathcal{R}_i$ and $v \in \mathcal{R}_j$, there is almost surely a path of length 2 from u to the endpoint of e_{ij} in $\mathcal{R}_i$ and similarly for v. Thus combining these two paths of length 2 with the edge e_{ij}, there is almost surely a path of length 5 from u to v. Hence, for any pair of vertices in A, there is almost surely a path of length at most 5 between them. Now since δ can be arbitrarily small, A contains an arbitrarily large fraction of the graph, and hence an arbitrarily large fraction of the graph is connected with diameter 5. □

Thus we have that $G(\mathbf{X}, n)$ exhibits clustering at the local level, has asymptotically constant diameter, and the degree distribution can be controlled explicitly with the variation of $\mathbf{X}$. Furthermore, as Kraetzl, Nickel, and Scheinerman have shown, for well-chosen distributions of $\mathbf{X}$, the degree distribution of $G(\mathbf{X}, n)$ can resemble a power-law degree distribution. Hence, their model not only results in a model for power-law graphs, but is one of a large class of random graph models that yield power-law-type graphs.

3. Directed Inner Product Graph

We now consider a directed generalization of the model above. Specifically, let $\mathbf{X}$ and $\mathbf{Y}$ be distributions on a real inner product space Ω such that if X and Y are random variables, distributed as $\mathbf{X}$ and $\mathbf{Y}$ respectively, then almost surely we have that $\langle X, Y\rangle \in (0, 1)$. Assign to each vertex v a pair (X_v, Y_v) such that X_v and Y_v are independent and (X_v, Y_v) is distributed as $(\mathbf{X}, \mathbf{Y})$. Then each arc $u \to v$ is present independently with probability $\langle X_u, Y_v\rangle$. Denote such a random graph on n vertices by $\overrightarrow{G}(\mathbf{X}, \mathbf{Y}, n)$. In this section we proceed to show that the results regarding $G(\mathbf{X}, n)$ generalize in spirit to the directed model $G(\mathbf{X}, \mathbf{Y}, n)$, beginning first with the result on local clustering, and then proceeding to show that the degree distribution can be controlled in a similar manner as in $G(\mathbf{X}, n)$, and finally that asymptotically, an arbitrarily large fraction of $G(\mathbf{X}, \mathbf{Y}, n)$ is strongly connected with constant diameter.

In order to show the clustering results we need the following convexity result.

Lemma 3.1. *Let Ω be a real inner product space and let $a, b \in \Omega$. Let $D \subseteq \Omega$ be a region such that for all $x \in D$, $\langle a, x\rangle \in (0, 1)$ and $\langle b, x\rangle \in (0, 1)$. Then $u\colon D \longrightarrow \mathbb{R}$ defined by $x \longmapsto \langle a, x\rangle \langle b, x\rangle$ is a convex function of x.*

Proof. Let $F : (0, 1) \times (0, 1) \longrightarrow \mathbb{R}$ be defined by $(x, y) \longmapsto xy$. Then $\nabla^2 F = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$. This matrix, although not positive semidefinite, is positive semidefinite over $[0, 1] \times [0, 1]$, and hence $F(x, y)$ is convex over its domain [Ben-Tal and Nemirovski 01]. Since $\langle a, x\rangle$ is a real inner product, for any $\lambda \in [0, 1]$ and $x, y \in D$, $\langle a, \lambda x + (1 - \lambda)y\rangle = \lambda \langle a, x\rangle + (1 - \lambda) \langle a, y\rangle$. Thus $\langle a, x\rangle$ is a convex function in x and similarly for $\langle b, x\rangle$. Thus $u(x) = F(\langle a, x\rangle , \langle b, x\rangle)$ is the composition of convex functions and is thus convex. □

Theorem 3.2. *Let Ω be a real inner product space and let $X_u, X_v, X_w, Y_u, Y_v, Y_w$ be independent random variables, not necessarily identically distributed, over Ω*

such that $\langle X_i, Y_j\rangle \in (0,1)$ *for all* i, j. *Now consider the random directed graph in which each arc* $i \to j$ *is present, independently, with probability* $\langle X_i, Y_j\rangle$. *Then we have that*

$$\begin{aligned}
\mathbb{P}(u \to w \mid u \to v, v \to w) &\geq \mathbb{P}(u \to w),\\
\mathbb{P}(u \to w \mid u \to v, w \to v) &\geq \mathbb{P}(u \to w),\\
\mathbb{P}(u \to w \mid v \to u, v \to w) &\geq \mathbb{P}(u \to w),\\
\mathbb{P}(u \to w \mid w \to v, v \to u) &= \mathbb{P}(u \to w).
\end{aligned}$$

Proof. First note that by the linearity of the expectation and the inner product, $\mathbb{P}(u \to w) = \langle \mathbb{E}[X_u], \mathbb{E}[Y_w]\rangle$. Now consider

$$\begin{aligned}
\mathbb{P}(u \to w \mid u \to v, v \to w) &= \frac{\mathbb{E}\left[\langle X_u, Y_w\rangle \langle X_u, Y_v\rangle \langle X_v, Y_w\rangle\right]}{\mathbb{E}\left[\langle X_u, Y_v\rangle \langle X_v, Y_w\rangle\right]}\\
&= \frac{\int \langle \mathbb{E}[X_v], Y_w\rangle \left(\int \langle X_u, Y_w\rangle \langle X_u, \mathbb{E}[Y_v]\rangle \, dX_u\right) dY_w}{\langle \mathbb{E}[X_u], \mathbb{E}[Y_v]\rangle \langle \mathbb{E}[X_v], \mathbb{E}[Y_v]\rangle}\\
&\geq \frac{\int \langle \mathbb{E}[X_v], Y_w\rangle \langle \mathbb{E}[X_u], Y_w\rangle \langle \mathbb{E}[X_u], \mathbb{E}[Y_v]\rangle \, dY_w}{\langle \mathbb{E}[X_u], \mathbb{E}[Y_v]\rangle \langle \mathbb{E}[X_v], \mathbb{E}[Y_v]\rangle}\\
&= \frac{\int \langle \mathbb{E}[X_v], Y_w\rangle \langle \mathbb{E}[X_u], Y_w\rangle \, dY_w}{\langle \mathbb{E}[X_v], \mathbb{E}[Y_v]\rangle}\\
&\geq \frac{\langle \mathbb{E}[X_v], \mathbb{E}[Y_w]\rangle \langle \mathbb{E}[X_u], \mathbb{E}[Y_w]\rangle}{\langle \mathbb{E}[X_v], \mathbb{E}[Y_v]\rangle}\\
&= \langle \mathbb{E}[X_u], \mathbb{E}[Y_w]\rangle,
\end{aligned}$$

where the inequalities come from the convexity of $u(x) = \langle a, x\rangle \langle b, x\rangle$ and Jensen's inequality [Grimmett and Stirzaker 01]. In a similar fashion we have that

$$\begin{aligned}
\mathbb{P}(u \to w \mid u \to v, w \to v) &= \frac{\mathbb{E}\left[\langle X_u, Y_w\rangle \langle X_u, Y_v\rangle \langle X_w, Y_v\rangle\right]}{\mathbb{E}\left[\langle X_u, Y_v\rangle \langle X_w, Y_v\rangle\right]}\\
&\geq \frac{\langle \mathbb{E}[X_u], \mathbb{E}[Y_w]\rangle \, \mathbb{E}\left[\langle \mathbb{E}[X_u], Y_v\rangle \langle \mathbb{E}[X_w], Y_v\rangle\right]}{\mathbb{E}\left[\langle \mathbb{E}[X_u], Y_v\rangle \langle \mathbb{E}[X_w], Y_v\rangle\right]}\\
&= \langle \mathbb{E}[X_u], \mathbb{E}[Y_w]\rangle,
\end{aligned}$$

$$\begin{aligned}
\mathbb{P}\left(u \to w \mid v \to u, v \to w\right) &= \frac{\mathbb{E}\left[\langle X_u, Y_w\rangle \langle X_v, Y_u\rangle \langle X_v, Y_w\rangle\right]}{\mathbb{E}\left[\langle X_v, Y_u\rangle \langle X_v, Y_w\rangle\right]} \\
&\geq \frac{\langle \mathbb{E}\left[X_u\right], \mathbb{E}\left[Y_w\right]\rangle \mathbb{E}\left[\langle X_v, \mathbb{E}\left[Y_u\right]\rangle \langle X_v, \mathbb{E}\left[Y_w\right]\rangle\right]}{\mathbb{E}\left[\langle X_v, \mathbb{E}\left[Y_u\right]\rangle \langle X_v, \mathbb{E}\left[Y_w\right]\rangle\right]} \\
&= \langle \mathbb{E}\left[X_u\right], \mathbb{E}\left[Y_w\right]\rangle, \\
\mathbb{P}\left(u \to w \mid v \to u, w \to v\right) &= \frac{\mathbb{E}\left[\langle X_u, Y_w\rangle \langle X_v, Y_u\rangle \langle X_w, Y_v\rangle\right]}{\mathbb{E}\left[\langle X_v, Y_u\rangle \langle X_w, Y_v\rangle\right]} \\
&= \frac{\langle \mathbb{E}\left[X_u\right], \mathbb{E}\left[Y_w\right]\rangle \mathbb{E}\left[\langle X_v, Y_u\rangle \langle X_w, Y_v\rangle\right]}{\mathbb{E}\left[\langle X_v, Y_u\rangle \langle X_w, Y_v\rangle\right]} \\
&= \langle \mathbb{E}\left[X_u\right], \mathbb{E}\left[Y_w\right]\rangle,
\end{aligned}$$

which completes the proof. □

Note that as an immediate corollary, by letting X_u, X_v, X_w (respectively Y_u, Y_v, Y_w) be independent identically distributed as $\mathbf{X}$ (respectively $\mathbf{Y}$), we have that for u, v, $w \in V(\overrightarrow{G}(\mathbf{X}, \mathbf{Y}, n))$,

$$\begin{aligned}
\mathbb{P}\left(u \to w \mid u \to v, v \to w\right) &\geq \mathbb{P}\left(u \to w\right), \\
\mathbb{P}\left(u \to w \mid u \to v, w \to v\right) &\geq \mathbb{P}\left(u \to w\right), \\
\mathbb{P}\left(u \to w \mid v \to u, v \to w\right) &\geq \mathbb{P}\left(u \to w\right), \\
\mathbb{P}\left(u \to w \mid w \to v, v \to u\right) &= \mathbb{P}\left(u \to w\right).
\end{aligned}$$

We observe that since the clustering is a local behavior, we can avoid the restriction that each vertex is identically distributed. However, for the rest of this section, we restrict our attention to the case in which every vertex is independent and identically distributed in order to adequately address the global properties of degree distribution and diameter.

Proposition 3.3. *Let $G = \overrightarrow{G}(\mathbf{X}, \mathbf{Y}, n)$, where $\mathbf{X}$ and $\mathbf{Y}$ are distributions over a real inner product space such that if X_i and Y_j are independent random variables distributed as $\mathbf{X}$ and $\mathbf{Y}$, then $\langle X_i, Y_j\rangle \in (0, 1)$ almost surely. Then for a vertex v,*

$$\begin{aligned}
\deg^-(v) &= \int \binom{n-1}{k} \langle \mathbb{E}\left[\mathbf{Y}\right], X\rangle^k \left(1 - \langle \mathbb{E}\left[\mathbf{Y}\right], X\rangle\right)^{n-1-k} d\mathbf{X}, \\
\deg^+(v) &= \int \binom{n-1}{k} \langle \mathbb{E}\left[\mathbf{X}\right], Y\rangle^k \left(1 - \langle \mathbb{E}\left[\mathbf{X}\right], Y\rangle\right)^{n-1-k} d\mathbf{Y}.
\end{aligned}$$

Proof. Similarly to proposition 3.3, we have that

$$\begin{aligned}
\mathbb{P}\left(\deg^-(v)=k\right) &= \int \mathbb{P}\left(\deg^-(v)=k \mid X_v\right) d\mathbf{X} \\
&= \iint \mathbb{P}\left(\deg^-(v)=k \mid X_v, Y_{V(G)-\{v\}}\right) d\mathbf{X}\,(d\mathbf{Y})^{n-1} \\
&= \iint \sum_{\substack{S\subseteq V(D)-\{v\}\\ |S|=k}} \prod_{i\in S}\langle X_v, Y_i\rangle \prod_{i\notin S\cup\{v\}} (1-\langle X_v, Y_i\rangle) d\mathbf{X}\,(d\mathbf{Y})^{n-1} \\
&= \int \sum_{\substack{S\subseteq V(G)-\{v\}\\ |S|=k}} \prod_{i\in S}\int \langle X_v, Y_i\rangle\, d\mathbf{Y} \prod_{i\notin S\cup\{v\}} \int (1-\langle X_v, Y_i\rangle) d\mathbf{Y}\, d\mathbf{X} \\
&= \int \sum_{\substack{S\subseteq V(D)-\{v\}\\ |S|=k}} \prod_{i\in S}\langle X_v, \mathbb{E}\left[\mathbf{Y}\right]\rangle \prod_{i\notin S\cup\{v\}} \int (1-\langle X_v, \mathbb{E}\left[\mathbf{Y}\right]\rangle) d\mathbf{X} \\
&= \int \binom{n-1}{k} \langle X, \mathbb{E}\left[\mathbf{Y}\right]\rangle^k \left(1-\langle X, \mathbb{E}\left[\mathbf{Y}\right]\rangle\right)^{n-1-k} d\mathbf{X}.
\end{aligned}$$

A symmetric argument applies for the in-degree of v. □

This leads to an immediate result on the density of edges in $\overrightarrow{G}(\mathbf{X}, \mathbf{Y}, n)$.

Corollary 3.4. *Let* $\mathbf{X}$ *and* $\mathbf{Y}$ *be distributions over a real inner product space such that if* X_i *and* Y_j *are independent random variables distributed as* $\mathbf{X}$ *and* $\mathbf{Y}$*, then* $\langle X_i, Y_j\rangle \in (0,1)$ *almost surely, and let* $G = \overrightarrow{G}(\mathbf{X}, \mathbf{Y}, n)$*. Then* $\mathbb{E}\left[|E(G)|\right] = n(n-1)\left\langle \mathbb{E}\left[\mathbf{X}\right], \mathbb{E}\left[\mathbf{Y}\right]\right\rangle$.

Proof.

$$\begin{aligned}
\mathbb{E}\left[|E(G)|\right] &= n\sum_{k=0}^{n-1} k\mathbb{P}\left(\deg^-(v)=k\right) \\
&= n\sum_{k=0}^{n-1} k\int \binom{n-1}{k} \langle X, \mathbb{E}\left[\mathbf{Y}\right]\rangle^k \left(1-\langle X, \mathbb{E}\left[\mathbf{Y}\right]\rangle\right)^{n-1-k} d\mathbf{X} \\
&= n\int \sum_{k=1}^{n-1} k\binom{n-1}{k} \langle X, \mathbb{E}\left[\mathbf{Y}\right]\rangle^k \left(1-\langle X, \mathbb{E}\left[\mathbf{Y}\right]\rangle\right)^{n-1-k} d\mathbf{X}
\end{aligned}$$

$$= n(n-1)\int \sum_{k=1}^{n-1} \binom{n-2}{k-1} \langle X, \mathbb{E}[\mathbf{Y}]\rangle^k (1 - \langle X, \mathbb{E}[\mathbf{Y}]\rangle)^{n-1-k}\, d\mathbf{X}$$
$$= n(n-1)\int \langle \mathbb{E}[\mathbf{Y}], X\rangle (\langle \mathbb{E}[\mathbf{Y}], X\rangle + 1 - \langle \mathbb{E}[\mathbf{Y}], X\rangle)^{n-2} d\mathbf{X}$$
$$= n(n-1)\langle \mathbb{E}[\mathbf{Y}], \mathbb{E}[\mathbf{X}]\rangle.$$

□

This implies that the edge density is $\Theta(n^2)$, contrary to conventional wisdom regarding complex networks. However, we feel that this tradeoff in practice is acceptable for two reasons, the first being that $\langle \mathbb{E}[\mathbf{X}], \mathbb{E}[\mathbf{Y}]\rangle$ is typically small. Furthermore, although the results regarding the diameter of the graph would not hold, one could consider $\mathbf{X}$ and $\mathbf{Y}$ as functions of n and introduce sparsity in that manner. We also note that particularly for the world wide web, gene–protein networks, and the Internet, it is widely accepted that empirical studies are not capturing all the edges present. Combine this fact with recent work showing that the incompleteness can severely skew some statistics of the data [Achlioptas et al. 05, Lakhina et al. 03], and it is plausible that one or more of these networks is not truly sparse. In addition, the recent work of Leskovec, Kleinberg, and Faloutsos [Leskovec et al. 07] has shown that for many social networks the number of edges is becoming superlinear in the number of vertices as these networks evolve.

We now turn to the asymptotic connectivity of $\overrightarrow{G}(\mathbf{X}, \mathbf{Y}, n)$. The result and proof echo the earlier proof for $G(\mathbf{X}, n)$ in Theorem 2.6. However, before we move on to the asymptotic diameter, we require the following generalization of a standard result about Erdős–Renyi random graphs.

Lemma 3.5. *Let D be a directed random graph on n vertices such that each directed edge is present independently with probability at least p. Then D is almost surely strongly connected with directed diameter 2.*

Proof. Consider any pair of vertices, say u and v. The probability that there is not a directed path of length at most 2 from u to v is at most $(1-p^2)^{n-2}(1-p)$. Thus the probability that u and v are not strongly connected by paths of length at most 2 is at most $1-(1-(1-p^2)^{n-2}(1-p))^2 = 2(1-p^2)^{n-2}(1-p)-(1-p^2)^{2n-4}(1-p)^2$. But then, the expected number of such pairs that are not strongly connected by paths of length at most 2 is at most

$$n(n-1)(2(1-p^2)^{n-2}(1-p) - (1-p^2)^{2n-4}(1-p)^2),$$

which approaches 0 as $n \to \infty$. Thus D is almost surely strongly connected with directed diameter at most 2 [Bollobás 98]. □

Theorem 3.6. *Let Ω be a real inner product space and let $\mathbf{X}$ and $\mathbf{Y}$ be distributions over Ω such that for independent random variables X and Y, distributed as $\mathbf{X}$ and $\mathbf{Y}$ respectively, $\langle X, Y\rangle \in (0,1)$ almost surely, and such that there is some compact set K such that $\mathbb{P}(\mathbf{X} \in K) = 1 = \mathbb{P}(\mathbf{Y} \in K)$. Now consider $G = \overrightarrow{G}(\mathbf{X}, \mathbf{Y}, n)$. Then an arbitrarily large fraction of G is almost surely strongly connected of diameter at most 5.*

Proof. Let $0 < \delta < \frac{1}{4}$ and choose some $\epsilon > 0$ such that $\mathbb{P}(\mathbf{X} \in B(0;\epsilon)) < 1 - \sqrt{1-\delta}$ and $\mathbb{P}(\mathbf{Y} \in B(0;\epsilon)) < 1 - \sqrt{1-\delta}$. Such an ϵ exists, since $\langle \mathbf{X}, \mathbf{Y}\rangle > 0$ almost surely, and hence $\|\mathbf{X}\| \neq 0$ and $\|\mathbf{Y}\| \neq 0$. Now let $A_X = K - B(0;\epsilon)$ and $A_Y = K - B(0;\epsilon)$. Then for a vertex v, the probability that the associated pair (X_v, Y_v) is not in $A = A_X \times A_Y$ is at most $1 - (\sqrt{1-\delta})^2 = \delta$. Thus, asymptotically, the fraction of vertices in A is almost surely at least $1 - \delta$.

Note that since A is a product of compact sets, A is compact. Now for each pair $(\alpha_x, \alpha_y) \in A$, choose

$$r_{(\alpha_x,\alpha_y)} \in \left\{ r > 0 \mid \forall (x,y) \in B(\alpha_x; r) \times B(\alpha_y; r), \langle x, y\rangle > \frac{\langle \alpha_x, \alpha_y\rangle}{4} \right\},$$

which is nonempty by the continuity of the inner product. Then

$$\bigcup_{(\alpha_x,\alpha_y)\in A} B\left(\alpha_x; r_{(\alpha_x,\alpha_y)}\right) \times B\left(\alpha_y; r_{(\alpha_x,\alpha_y)}\right)$$

is an open cover of A and hence has some finite subcover, say

$$\left\{ B\left(\alpha_{xi}; r_{(\alpha_{x_i},\alpha_{y_i})}\right) \times B\left(\alpha_{y_i}; r_{(\alpha_{x_i},\alpha_{y_i})}\right) \right\}.$$

Denote $B\left(\alpha_{xi}; r_{(\alpha_{x_i},\alpha_{y_i})}\right) \times B\left(\alpha_{y_i}; r_{(\alpha_{x_i},\alpha_{y_i})}\right)$ by $\mathcal{R}_i$. Now, as above, $\mathcal{R}_i$ almost surely contains either infinitely many vertices or none. If none, then $\mathcal{R}_i$ is irrelevant to the graph. However, if $\mathcal{R}_i$ contains infinitely many vertices, then since the probability of an arc between any two vertices in $\mathcal{R}_i$ is bounded away from zero, the graph induced by $\mathcal{R}_i$ is almost surely strongly connected with diameter at most 2.

Now consider any two regions $\mathcal{R}_i$ and $\mathcal{R}_j$ occurring with positive probability. Again, note that there is a natural probability measure on

$$B\left(\alpha_{xi}; r_{(\alpha_{x_i},\alpha_{y_i})}\right) \times B\left(\alpha_{y_i}; r_{(\alpha_{x_j},\alpha_{y_j})}\right).$$

Then since $\langle \mathbf{X}, \mathbf{Y}\rangle > 0$, there exist some δ' and ϵ' such that

$$\mathbb{P}\left(\langle \mathbf{X}, \mathbf{Y}\rangle > \delta' \mid (\mathbf{X}, \mathbf{Y}) \in B\left(\alpha_{xi}; r_{(\alpha_{x_i},\alpha_{y_i})}\right) \times B\left(\alpha_{y_i}; r_{(\alpha_{x_j},\alpha_{y_j})}\right)\right) > \epsilon'.$$

But then there is almost surely an arc from $\mathcal{R}_i$ to $\mathcal{R}_j$ and similarly for an arc from $\mathcal{R}_j$ to $\mathcal{R}_i$. Furthermore, since i and j were arbitrary, this implies that the graph induced by A is strongly connected with diameter at most 5. Finally, since δ is arbitrary, A represents an arbitrarily large fraction of the graph. □

Thus we have that $\overrightarrow{G}(\mathbf{X}, \mathbf{Y}, n)$, for appropriate choices of $\mathbf{X}$ and $\mathbf{Y}$, can produce directed power-law-like graphs. Indeed, we believe it is likely from computational experimentation that there exist choices of $\mathbf{X}$ and $\mathbf{Y}$ such that the in-degree sequence of $\overrightarrow{G}(\mathbf{X}, \mathbf{Y}, n)$ differs from the out-degree sequence qualitatively as well as quantitatively. Specifically, we believe that there exist $\mathbf{X}$ and $\mathbf{Y}$ such that the in-degree sequence is qualitatively Poisson distributed and the out-degree sequence is distributed in a near-power-law fashion.

Although we believe that $\overrightarrow{G}(\mathbf{X}, \mathbf{Y}, n)$ is a promising model for many real-world problems, there are classes of real-world phenomena for which there is empirical evidence for a directed power-law-type graph, but for which the independence of $\mathbf{X}$ and $\mathbf{Y}$ makes little sense. For instance, consider the world wide web, or the community of bloggers, or even an acquaintance or friendship network. Intuitively, we believe that there is something inherent about the person or web page that governs what it links to (communicates with) and what links to (communicates with) it, and hence those relationships should be coupled in some strong manner. Thus we consider a final model generalizing both $G(\mathbf{X}, n)$ and $\overrightarrow{G}(\mathbf{X}, \mathbf{Y}, n)$ that yields a graph with the in-degree and the out-degree of a vertex explicitly coupled.

4. Spherical Inner Product Graphs

We now consider the following model as a generalization of $\overrightarrow{G}(\mathbf{X}, \mathbf{Y}, n)$ in which $\mathbf{X}$ and $\mathbf{Y}$ are not independent. Each vertex v is assigned a pair $(\mu_v S_v, \rho_v S_v)$, where S_v is chosen uniformly from the surface of the positive orthant of the d-dimensional unit sphere, μ_v is chosen according to $\mathcal{U}^\alpha(0,1)$, and ρ_v is chosen according to $\mathcal{U}^\beta(0,1)$. Then every arc (u, v) is present independently with probability $\mu_v \rho_u {S_u}^T S_v$. Denote such a graph by $\overrightarrow{\mathbb{S}}^d(\alpha, \beta, n)$. We note that many of the results of the previous sections generalize immediately.

Corollary 4.1. *Let μ_u, μ_v, μ_w be independent random variables distributed as $\mathcal{U}^\alpha(0,1)$ and let ρ_u, ρ_v, ρ_w be independent uniform random variables distributed as $\mathcal{U}^\beta(0,1)$. Let S_u, S_v, S_w be independent random variables on the surface of the positive orthant of the d-dimensional unit sphere. Now consider the random directed graph in which each arc $i \to j$ is present, independently, with probability $\mu_i \rho_j S_i^T S_j$.*

Then we have that

$$\begin{aligned}
\mathbb{P}(u \to w \mid u \to v, v \to w) &> \mathbb{P}(u \to w), \\
\mathbb{P}(u \to w \mid u \to v, w \to v) &> \mathbb{P}(u \to w), \\
\mathbb{P}(u \to w \mid v \to u, v \to w) &> \mathbb{P}(u \to w), \\
\mathbb{P}(u \to w \mid w \to v, v \to u) &= \mathbb{P}(u \to w).
\end{aligned}$$

Proof. Note that

$$\begin{aligned}
\mathbb{P}(u \to w \mid u \to v, v \to w) &= \frac{\mathbb{E}\left[\mu_u \rho_w S_u^T S_w \mu_u \rho_v S_u^T S_v \mu_v \rho_w S_v^T S_w\right]}{\mathbb{E}\left[\mu_u \rho_v S_u^T S_v \mu_v \rho_w S_v^T S_w\right]} \\
&= \frac{\mathbb{E}\left[\mu_u^2 \rho_w^2 \mu_u \mu_w S_u^T S_w S_u^T S_v S_v^T S_w\right]}{\mathbb{E}\left[\mu_u \rho_v \mu_v \rho_w S_u^T S_v S_v^T S_w\right]} \\
&= \frac{\mathbb{E}\left[\mu_u^2\right]}{\mathbb{E}\left[\mu_u\right]} \frac{\mathbb{E}\left[\rho_w^2\right]}{\mathbb{E}\left[\rho_w\right]} \frac{\mathbb{E}\left[S_u^T S_w S_u^T S_v S_v^T S_w\right]}{\mathbb{E}\left[S_u^T S_v S_v^T S_w\right]}.
\end{aligned}$$

Now since $\mathrm{Var}(\mu_u) > 0$, we have $\mathbb{E}\left[\mu_u^2\right] > \mathbb{E}\left[\mu_u\right]^2$, and hence $\mathbb{E}\left[\mu_u^2\right]/\mathbb{E}\left[\mu_u\right] > \mathbb{E}\left[\mu_u\right]$. Similarly, $\mathbb{E}\left[\rho_w^2\right]/\mathbb{E}\left[\rho_w\right] > \mathbb{E}\left[\rho_w\right]$. Furthermore, by the result for the undirected case (Theorem 2.1), we have

$$\frac{\mathbb{E}\left[S_u^T S_w S_u^T S_v S_v^T S_w\right]}{\mathbb{E}\left[S_u^T S_v S_v^T S_w\right]} > \mathbb{E}\left[S_u^T S_w\right].$$

Thus we have that

$$\begin{aligned}
\mathbb{P}(u \to w \mid u \to v, v \to w) &= \frac{\mathbb{E}\left[\mu_u^2\right]}{\mathbb{E}\left[\mu_u\right]} \frac{\mathbb{E}\left[\rho_w^2\right]}{\mathbb{E}\left[\rho_w\right]} \frac{\mathbb{E}\left[S_u^T S_w S_U^T S_v S_v^T S_w\right]}{\mathbb{E}\left[S_u^T S_v S_v^T S_w\right]} \\
&> \mathbb{E}\left[\mu_u\right] \mathbb{E}\left[\rho_w\right] \mathbb{E}\left[S_u^T S_w\right] \\
&= \mathbb{P}(u \to w).
\end{aligned}$$

A similar reduction to the undirected case holds for the remaining three cases. □

Specifically, we note that this implies that $\overrightarrow{\mathbb{S}}^d(\alpha, \beta, n)$ exhibits clustering in the same manner as $\overrightarrow{G}(\mathbf{X}, \mathbf{Y}, n)$. Furthermore, we note that in fact, the restrictions on the distributions of the μ and ρ variables can be relaxed to any distribution on $(0, 1)$, and the resulting graph will exhibit clustering in the manner of $\overrightarrow{G}(\mathbf{X}, \mathbf{Y}, n)$.

Corollary 4.2. *Let $G = \overrightarrow{\mathbb{S}}^d(\alpha, \beta, n)$. Then an arbitrarily large fraction of G is almost surely strongly connected with diameter at most 5 as $n \to \infty$.*

Proof. This result follows immediately from Theorem 2.6, by noting that the proof of the theorem relies entirely on the geometry of random variables, and not on any specific properties of the probability distribution on that geometry. Thus the dependence or independence between $\mathbf{X}$ and $\mathbf{Y}$ does not affect the diameter of the graph, yielding the desired result. □

Thus, combining these results with computational experiments, we have that $\overrightarrow{\mathbb{S}}^d(\alpha, \beta, n)$ exhibits power-law-like behavior. In fact, we observe that excepting the clustering results, the independence or dependence of $\mathbf{X}$ and $\mathbf{Y}$ does not affect the behavior of $\overrightarrow{G}(\mathbf{X}, \mathbf{Y}, n)$. Specifically, neither the degree sequence nor the diameter of $\overrightarrow{G}(\mathbf{X}, \mathbf{Y}, n)$ depends on the independence of $\mathbf{X}$ and $\mathbf{Y}$ at a given vertex. Thus with further work, the model of Kraetzl, Nickel, and Scheinerman can be extended to a large class of models with prescribed degree sequences, asymptotically constant diameter, and under appropriate conditions, local clustering.

5. Future Work

Although there is limited computational evidence for a variety of power-law-like degree distributions and an explicit representation of the degree sequence given the parameters of the model, there are obvious practical limitations to the model $G(\mathbf{X}, n)$ and its directed generalizations $\overrightarrow{G}(\mathbf{X}, \mathbf{Y}, n)$ and $\overrightarrow{\mathbb{S}}^d(\alpha, \beta, n)$. Specifically, further work is needed in finding the reverse mapping that would take degree distributions, or pairs thereof, and provide distributions, or pairs thereof, on $\mathbb{R}^d$ that yield those degree distributions or something near those degree distributions. There is some indication that such a reverse procedure exists, in that $\mathbb{P}(\deg(v) = n - 1 \mid v \in G(\mathbf{X}, n))$ is the $(n-1)$th moment of $\mathbb{E}[\mathbf{X}]^T X$ and we can write $\mathbb{P}(\deg(v) = n - 2 \mid v \in G(\mathbf{X}, n))$ as a linear function of the $(n-1)$th and $(n-2)$th moments of $\mathbb{E}[\mathbf{X}]^T X$, and similarly down the line.

Thus, for fixed n and degree probability distributions, it is possible to use back-substitution to find the first through $(n-1)$th moments of $\mathbb{E}[\mathbf{X}]^T X$, which leaves open the question of how those moments can be transformed into statements about $\mathbf{X}$ and furthermore, whether there a more general algorithm that can sidestep the tedious back-substitution.

There is also significant room for improvement in the rapid generation of both $G(\mathbf{X}, n)$ and $\overrightarrow{G}(\mathbf{X}, \mathbf{Y}, n)$. The naive method of generating $\overrightarrow{G}(\mathbf{X}, \mathbf{Y}, n)$, for

instance, involves generating n copies of both $\mathbf{X}$ and $\mathbf{Y}$ to form two matrices in $\mathbb{R}^{d\times n}$ and then taking the matrix product of those matrices, resulting in a matrix in $(0,1)^{n\times n}$ that represents the probability of each arc being present. This results in an algorithm that is approximately $\mathcal{O}\left(n^2 d\right)$ for generating the graph using $\mathcal{O}\left(n^2\right)$ space. For large n, these space and time requirements begin to pose a problem, especially the space requirement. Specifically, despite the fact that there are only $\mathbb{E}\left[\mathbf{X}\right]^T \mathbb{E}\left[\mathbf{Y}\right] n^2$ edges, the calculation process by naive methods requires space of order $2n^2$. Although both space requirements are the same asymptotically, in practice $2/(\mathbb{E}\left[\mathbf{X}\right]^T \mathbb{E}\left[\mathbf{Y}\right])$ is a prohibitively large factor of overuse of space.

We feel that one of the major directions for future work with the directed model $\overrightarrow{G}(\mathbf{X},\mathbf{Y},n)$ is the revisiting of algorithmic and structural properties of complex networks. Specifically, in papers such as [Achlioptas et al. 05, Eubank et al. 04, Gkantsidis et al. 05, Kleinberg 99, Lakhina et al. 03, Mihail et al. 06, Mihail et al. 05], undirected or specially directed models for complex networks are used to analyze structural and algorithmic properties of the network. However, since the model $\overrightarrow{G}(\mathbf{X},\mathbf{Y},n)$ incorporates the direction in a nonspecific manner, it provides an opportunity to revisit these results and confirm that the behavior shown is not a function of being undirected or the specific nature of the directions. We feel that a robust directed model for complex networks is important in bringing the theory of complex networks closer to practical questions about real-world networks that inspired the initial research.

References

[Abramowitz and Stegun 64] Milton Abramowitz and Irene A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, National Bureau of Standards Applied Mathematics Series 55. Washington, DC: U.S. Government Printing Office, 1964.

[Achlioptas et al. 05] Dimitris Achlioptas, David Kempe, Aaron Clasuet, and Christopher Moore. "On the Bias of Traceroute Sampling or, Power-Law Degree Distributions in Regular Graphs." In *Proceedings of the Thirty-Seventh Annual ACM Symposium on Theory of Computing*, pp. 694–703. New York: ACM Press, 2005.

[Albert and Barabási 02] Réka Albert and Albert-László Barabási. "Statistical Mechanics of Complex Networks." *Rev. Modern Phys.* 74:1 (2002), 47–97.

[Azar et al. 01] Yossi Azar, Amos Fiat, Anna Karlin, Frank McSherry, and Jared Saia. "Spectral Analysis of Data." In *Proceedings of the 33rd ACM Symposium on Theory of Computing*, pp. 619–626. New York: ACM Press, 2001.

[Ben-Tal and Nemirovski 01] Aharon Ben-Tal and Arkadi Nemirovski. *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*. Philadelphia: SIAM, 2001.

[Bollobás 98] Béla Bollobás. *Modern Graph Theory*, Graduate Texts in Mathematics 184. New York: Springer, 1998.

[Bornholdt and Schuster 03] Stefan Bornholdt and Heinz Georg Schuster, editors. *Handbook of Graphs and Networks*. Weinheim: Wiley-VCH, 2003.

[Caldarelli et al. 02] G. Caldarelli, A. Capocci, P. de Los Rios, and M. A. Muñoz. "Scale-Free Networks from Varying Vertex Intrinsic Fitness." *Physical Review Letters* 89:25 (2002), 258702.

[Chung et al. 03] Fan Chung, David J. Galas, T. Gregory Dewey, and Lincoln Lu. "Duplication Models for Biological Networks." *Journal of Computational Biology* 10:5 (2003), 677–687.

[Durrett 07] Rick Durrett. *Random Graph Dynamics*, Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge, UK: Cambridge University Press, 2007.

[Eubank et al. 04] Stephen Eubank, V. S. Anil Kumar, Madhav V. Marathe, Aravind Srinivasan, and Nan Wang. "Structural and Algorithmic Aspects of Massive Social Networks." In *Proceedings of the 15th ACM-SIAM Symposium on Discrete Algorithms*, pp. 718–727. Philadelphia: SIAM, 2004.

[Flaxman et al. 06] Abraham D. Flaxman, Alan M. Frieze, and Juan Vera. "A Geometric Preferential Attachment Model of Networks." *Internet Math.* 3:2 (2006), 187–205.

[Gkantsidis et al. 05] Christos Gkantsidis, Milena Mihail, and Amin Saberi. "Hybrid Search Schemes for Unstructured Peer-to-Peer Networks." *INFOCOM2005, Proceedings IEEE* 3:13–17 (2005), 1526–1537.

[Grimmett and Stirzaker 01] Geoffrey R. Grimmett and David R. Stirzaker. *Probability and Random Processes*, third edition. New York: Oxford University Press, 2001.

[Kleinberg 99] Jon M. Kleinberg. "The Small World Phenomenon: An Algorithmic Perspective." In *Proceedings of the Thirty-Second Annual ACM Symposium on the Theory of Computing*, pp. 163–170. New York: ACM Press, 1999.

[Kraetzl et al. 05a] Miro Kraetzl, Christine Nickel, and Edward R. Scheinerman. "Random Dot Product Graphs: A Model for Social Networks." Manuscript, 2005.

[Kraetzl et al. 05b] Miro Kraetzl, Christine Nickel, Edward R. Scheinerman, and Kimberly Tucker. "Random Dot Product Graphs." Available at http://www.ipam.ucla.edu/abstract.aspx?tid=5498, 2005.

[Kumar et al. 00] Ravi Kumar, Prabhakar Raghavan, Sridhar Rajagopalan, D. Sivakumar, Andrew Tompkins, and Eli Upfal. "The Web as a Graph." In *Proceedings of the 19th ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems*, pp. 1–10. New York: ACM Press, 2000.

[Lakhina et al. 03] Anukool Lakhina, John W. Byers, Mark Crovella, and Peng Xie. "Sampling Biases in IP Topology Measurements." In *INFOCOM 2003: Twenty-Second Annual Joint Conference of the IEEE Computer and Communications Societies*, Vol. 1, pp 332–341. Los Alamitos, CA: IEEE Press, 2003.

[Leskovec et al. 07] Jure Leskovec, Jon Kleinberg, and Christos Faloutsos. "Graph Evolution: Densification and Shrinking Diameters." *ACM Trans. Knowl. Discov. Data* 1:1 (2007), Article No. 2.

[Mihail et al. 05] Milena Mihail, Amin Saberi, and Prasad Tetali. "Random Walks with Lookahead in Power Law Random Graphs." *Internet Mathematics* 3:2 (2006), 147–152.

[Mihail et al. 06] Milena Mihail, Christos Papadimitriou, and Amin Saberi. "On Certain Connectivity Properties of the Internet Topology." *J. Comput. System Sci.* 72:2 (2006), 239–251.

[Newman 02] M. E. J. Newman. "Assortative Mixing in Networks." *Physical Review Letters* 89:20 (2002), 208701.

[Temme 03] Nico M. Temme. "Large Parameter Cases of the Gauss Hypergeometric Function." In *Proceedings of the Sixth International Symposium on Orthogonal Polynomials, Special Functions and Their Applications, Journal of Computational and Applied Mathematics* 153:1–5 (2003), 441–462.

Stephen J. Young, School of Mathematics, Georgia Institute of Technology, 686 Cherry Street, Atlanta, GA 30332 (young@math.gatech.edu)

Edward Scheinerman, Whiting School of Engineering, The Johns Hopkins University, 3400 North Charles Street, Baltimore, MD 21218 (ers@jhu.edu)

Received February 14, 2008; accepted May 6, 2008.