



| | |
|-------------------------------------|--|
| Title | Generalized Random Dot Product graph |
| Authors(s) | Ng, Tin Lok James, Murphy, Thomas Brendan |
| Publication date | 2019-05 |
| Publication information | Ng, Tin Lok James, and Thomas Brendan Murphy. "Generalized Random Dot Product Graph." Elsevier, May 2019. https://doi.org/10.1016/j.spl.2019.01.011 . |
| Publisher | Elsevier |
| Item record/more information | http://hdl.handle.net/10197/10709 |
| Publisher's statement | This is the author's version of a work that was accepted for publication in Statistics and Probability Letters. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in Statistics and Probability Letters (148, (2019)) https://doi.org/10.1016/j.spl.2019.01.011 |
| Publisher's version (DOI) | 10.1016/j.spl.2019.01.011 |

Downloaded 2026-05-01 23:35:04

The UCD community has made this article openly available. Please share how this access benefits you. Your story matters! (@ucd_oa)



© Some rights reserved. For more information

Generalized Random Dot Product Graph

Tin Lok James Ng, Thomas Brendan Murphy

School of Mathematics and Statistics, University College Dublin

Abstract

The Random Dot Product model for social network was proposed in Nickel (2007), where asymptotic results such as degree distribution, clustering and diameter on both dense and sparse cases were derived. Young and Scheinerman (2007) explored two generalizations of the model in the dense case and obtained similar asymptotic results. In this paper, we consider a generalization of the Random Dot Product model and derive its theoretical properties under the dense, sparse and intermediate cases. In particular, properties such as the size of the largest component and connectivity can be derived by applying recent results on inhomogeneous random graphs (Bollobás et al., 2007, Devroye and Fraiman, 2014).

1. Introduction

The Random Dot Product model was introduced in Nickel (2007), where each vertex v in the network is associated with a latent variable x_v , drawn from $U^a[0, 1]$, the a^{th} power ($a > 1$) of the uniform distribution on $[0, 1]$. They assume that the probability two vertices v and u are connected is given by

$$P(u \sim v) = \frac{x_u x_v}{n^{2-s}},$$

where $s \leq 2$ and where $s = 2$ corresponds to the dense case.

We generalize their model by assuming that for each vertex v , the latent variable x_v is drawn from a Beta(α, β) distribution. It is easy to see that the a^{th} power of uniform distribution is a special case of the Beta(α, β) distribution by letting $\alpha = 1/a$ and $\beta = 1$. Further, the general form of connection probability

$$P(u \sim v) = \frac{c(\log(n))^b x_u x_v}{n^{2-s}},$$

with $b \in \{0, 1\}$, $0 \leq s \leq 2$, and $c > 0$ is considered in this paper. We note that the case $s < 0$ is uninteresting since the expected number of edges is $O(n^s) \rightarrow 0$ as $n \rightarrow \infty$.

6 As will be shown in the next section, real networks can be better characterized
7 due to the flexibility of Beta distribution. Modeling the latent variable with
8 other distributions may be studied in the future. An alternative extension of
9 the random dot product graphs is proposed in O'Connor et al. (2015) where
10 the identity link is replaced by the logistic link in the connection probability.
11 Theoretical properties of the the model proposed in (O'Connor et al., 2015) may
12 be studied using similar methods considered in this paper.

13 In Section 2, we consider the dense case with $P(u \sim v) = x_u x_v$ and derive the
14 clustering coefficient and the expected number of vertices of a certain degree.
15 We then consider the connectivity under the intermediate case with $P(u \sim v) =$
16 $c \log(n) x_u x_v / n$ in Section 3. The sparse case with $P(u \sim v) = c x_u x_v / n^{2-s}$ with
17 $0 \leq s \leq 1$ is considered in Section 4. In particular, we derive properties related
18 to giant component and degree distribution for the special case $s = 1$. Further,
19 we show that under this regime, the generalized random dot product graph
20 is asymptotically equivalent to the stochastic version of the model explored in
21 Chung and Lu (2003). We summarize the cases that we will consider in this
22 paper in Table 1.

Table 1: Regimes and Connection Probabilities

| Case | Parameters | Edge Probability $P(u \sim v)$ |
|-----------------------|--|-----------------------------------|
| Dense | $s = 2, b = 0, c = 1$ | $x_u x_v$ |
| Intermediate | $s = 1, b = 1, c \in \mathbb{R}$ | $\frac{c \log(n) x_u x_v}{n}$ |
| Sparse | $0 \leq s \leq 1, b = 0, c \in \mathbb{R}$ | $\frac{c x_u x_v}{n^{2-s}}$ |
| Sparse - Special Case | $s = 1, b = 0, c \in \mathbb{R}$ | $\frac{c x_u x_v}{n}$ |

23 2. Dense Case

24 Let $G(n)$ be a graph on n vertices $V(G)$. Each vertex $v \in V(G)$ is associated
25 with a beta random variable $x_v \sim \text{Beta}(\alpha, \beta)$. We assume that conditional on
26 the latent variables x_u and x_v , the probability that two vertices u and v are
27 connected is given by $P(u \sim v | x_u, x_v) = x_u x_v$, where \sim denotes that u and
28 v are connected. We first have the following result concerning the connection
29 probability between two nodes.

30
31 **Proposition 2.1.** *The probability that any two vertices u, v are connected is*
32 *given by*

$$P(u \sim v) = \frac{\alpha^2}{(\alpha + \beta)^2}.$$

Proof.

$$\begin{aligned}
P(u \sim v) &= \left(\frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \right)^2 \int_0^1 \int_0^1 x_u x_v x_u^{\alpha-1} (1-x_u)^{\beta-1} x_v^{\alpha-1} (1-x_v)^{\beta-1} dx_u dx_v \\
&= \left(\frac{\Gamma(\alpha + 1)\Gamma(\beta)}{\Gamma(\alpha + \beta + 1)} \right)^2 \left(\frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \right)^2.
\end{aligned}$$

33

□

34 We next assume that n is sufficiently large, and consider the expected number
35 of isolated vertices in the graph $G(n)$. Let $\chi(k)$ be the number of vertices in
36 $G(n)$ with degree k , the following result can be derived.

37

38 **Proposition 2.2.** *For sufficiently large n , the expected number of vertices in*
39 *$G(n)$ with degree 0 is given by*

$$E(\chi(0)) \sim C \frac{\Gamma(\alpha + \beta)}{n^{1-\alpha}\Gamma(\beta)}$$

40 where $1 \leq C \leq \left(\frac{\alpha}{\alpha+\beta}\right)^{-\alpha}$.

41 *Proof.* Let $d(v)$ be the degree of vertex v in $G(n)$, and let $g(x)$ denote the density
42 of Beta(α, β). We have

$$\begin{aligned}
P(d(v) = 0) &= \int_0^1 \cdots \int_0^1 (1-x_1 y) \cdots (1-x_{n-1} y) \\
&\quad g(x_1) \cdots g(x_{n-1}) g(y) dx_1 \cdots dx_{n-1} dy \\
&= \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 \left(1 - \frac{\alpha}{\alpha + \beta} y\right)^{n-1} y^{\alpha-1} (1-y)^{\beta-1} dy.
\end{aligned}$$

43 To derive an upper bound, we have

$$\begin{aligned}
P(d(v) = 0) &\leq \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 \left(1 - \frac{\alpha}{\alpha + \beta} y\right)^{n-1} y^{\alpha-1} \left(1 - \frac{\alpha}{\alpha + \beta} y\right)^{\beta-1} dy \\
&= \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 \left(1 - \frac{\alpha}{\alpha + \beta} y\right)^{n+\beta-2} y^{\alpha-1} dy \\
&\sim \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^\infty \exp\left(-\frac{\alpha}{\alpha + \beta} t^{\frac{1}{\alpha}}\right) \frac{dt}{\alpha n^\alpha} \\
&= \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{1}{n^\alpha} \left(\frac{\alpha}{\alpha + \beta}\right)^{-\alpha} \Gamma(\alpha)
\end{aligned}$$

44 where \sim denotes asymptotically equal. In the other direction, we have

$$\begin{aligned}
P(d(v) = 0) &\geq \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 (1-y)^{n-1} y^{\alpha-1} (1-y)^{\beta-1} dy \\
&\sim \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^\infty \exp(-t^{\frac{1}{\alpha}}) \frac{dt}{\alpha n^\alpha} \\
&= \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha)}{n^\alpha}.
\end{aligned}$$

45 The result is proved by the property $E(\chi(0)) = (n-1)P(d(v) = 0)$. \square

46 As in Proposition 2.2.3 of Nickel (2007), the result above can be generalized to
47 degree k with $k \ll n$.

48
49 **Proposition 2.3.** *For sufficiently large n , the expected number of vertices in*
50 *$G(n)$ with degree k is given by*

$$E(\chi(k)) \sim C \frac{1}{k!} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \Gamma(\alpha + k) \frac{1}{n^{1-\alpha}}$$

51 where $1 \leq C \leq (\frac{\alpha}{\alpha+\beta})^{-\alpha}$.

52 We now explore the clustering property of the generalized random dot product
53 graph. The clustering coefficient of a vertex v introduced by Watts and Strogatz
54 (1998) is defined as the ratio between the number of triangles connected to v
55 and the number of triples centered on vertex v .

56
57 **Proposition 2.4.** *Conditional on $u \sim v$, and $v \sim w$, the probability that uvw*
58 *forms a triangle is given by*

$$P(u \sim w | u \sim v, v \sim w) = \frac{(\alpha + 1)^2}{(\alpha + \beta + 1)^2}.$$

59 *Proof.* We can see that

$$\begin{aligned}
P(u \sim v, v \sim w, w \sim u) &= \int_0^1 \int_0^1 \int_0^1 x_u^2 x_v^2 x_w^2 g(x_u) g(x_v) g(x_w) dx_u dx_v dx_w \\
&= \left(\frac{(\alpha + 1)\alpha}{(\alpha + \beta + 1)(\alpha + \beta)} \right)^3
\end{aligned}$$

60 and

$$\begin{aligned}
P(u \sim v, v \sim w) &= \int_0^1 \int_0^1 x_u x_v^2 x_w g(x_u) g(x_v) g(x_w) dx_u dx_v dx_w \\
&= \frac{\alpha^3(\alpha + 1)}{(\alpha + \beta + 1)(\alpha + \beta)^3}.
\end{aligned}$$

61 Thus,

$$\begin{aligned} P(u \sim w | u \sim v, v \sim w) &= \frac{P(u \sim v, v \sim w, w \sim u)}{P(u \sim v, v \sim w)} \\ &= \frac{(\alpha + 1)^2}{(\alpha + \beta + 1)^2}. \end{aligned}$$

62

□

We note that the unconditional probability

$$P(u \sim w) = \frac{\alpha^2}{(\alpha + \beta)^2} < \frac{(\alpha + 1)^2}{(\alpha + \beta + 1)^2} = P(u \sim w | u \sim v, v \sim w).$$

In the special case where $\alpha = 1/a$ and $\beta = 1$, Nickel (2007) shows that

$$P(u \sim w | u \sim v, v \sim w) = \left(\frac{a + 1}{2a + 1} \right)^2.$$

63 The assumption that $a \geq 1$ made in Nickel (2007) implies that the quantity
 64 is bounded below by $1/4$ and above by $4/9$. However, by allowing both α
 65 and β to vary, $P(u \sim w | u \sim v, v \sim w)$ can take any value between 0 and
 66 1; thus the generalized model offers greater modeling flexibility. The observed
 67 average clustering coefficient in social networks is often above 0.5 (McAuley and
 68 Leskovec, 2012). For example, Facebook and Twitter have an average clustering
 69 coefficient of 0.6055 and 0.5653, respectively. However, the Wikipedia Talk
 70 network has an average clustering coefficient of 0.0526 (Leskovec et al., 2010a,b).
 71 Thus, the generalized model achieves a range of values for the average clustering
 72 coefficient which are comparable those found in real work networks.

73 3. Intermediate Case

74 We briefly consider the intermediate case where two vertices u and v in $V(G)$
 75 are connected with probability

$$P(u \sim v | x_u, x_v) = \frac{cx_u x_v \log(n)}{n} = \frac{\kappa(x_u, x_v) \log(n)}{n},$$

76 where $\kappa(x, y) = cxy$ with $c \in \mathbb{R}$. In particular, we explore the connectivity of
 77 the generalized random dot product graph.

78 Define $\lambda(x) = \int_S \kappa(x, y) d\mu(y)$ and $\lambda_2(x) = \left(\int_S \kappa(x, y)^2 d\mu(y) \right)^{1/2}$ where μ is
 79 the probability measure for the Beta(α, β) distribution on $S = [0, 1]$, and let
 80 $\lambda_* = \text{ess inf}_x \{\lambda(x)\}$. The following result is due to Devroye and Fraiman (2014)

81

Lemma 3.1. (Devroye and Fraiman (2014) Theorem 1)
 If κ is irreducible, continuous a.e. and $\lambda_2 \in L^\infty(S, \mu)$ then

$$\lim_{n \rightarrow \infty} P(G(n, \kappa) \text{ is connected}) = \begin{cases} 0 & \text{if } \lambda_* < 1 \\ 1 & \text{if } \lambda_* > 1 \end{cases}.$$

82 By Lemma 3.1, we have the following result concerning the connectivity of the
 83 generalized random dot product graph.

84
 85 **Proposition 3.2.** For all fixed c , α and β , the generalized random dot product
 86 graph is disconnected a.s.

87 *Proof.* For the generalized random dot product graph we find that

$$\begin{aligned} \lambda_* &= \operatorname{ess\,inf}_x \{\lambda(x)\} \\ &= \operatorname{ess\,inf}_x \left\{ cx \frac{\alpha}{\alpha + \beta} \right\} \\ &= 0 \end{aligned}$$

88 and the result follows from Lemma 3.1. □

89 4. Sparse Case

90 In the sparse case, we assume that the probability that two vertices u and v in
 91 $G(n)$ are connected is given by

$$P(u \sim v | x_u, x_v) = \frac{cx_u x_v}{n^{2-s}}, \quad (1)$$

92 where $0 \leq s \leq 1$ and $c \in \mathbb{R}$.

93 4.1. Special case: $s = 1$

94 We first consider the case that the probability two nodes are connected is given
 95 by

$$P(u \sim v | x_u, x_v) = \frac{cx_u x_v}{n} = \frac{\kappa(x_u, x_v)}{n},$$

96 for some fixed $c > 0$.

97 The general model $P(u \sim v | x_u, x_v) = \kappa(x_u, x_v)/n$ for some symmetric non-
 98 negative Borel measurable function $\kappa(x_u, x_v)$ is considered in Bollobás et al.
 99 (2007) where many theoretical properties were derived.

100 Let T_κ be the integral operator defined by

$$(T_\kappa f)(x) = \int_0^1 \kappa(x, y) f(y) d\mu(y)$$

101 for any measurable function f such that the integral above is finite or $+\infty$ for
102 a.e. x .

103 The norm of the operator T_κ is defined by

$$\|T_\kappa\| := \sup\{\|T_\kappa f\|_2 : f \geq 0, \|f\|_2 \leq 1\}.$$

104 We have that $\|T_\kappa\| \leq \|\kappa\|_{L^2} = (\int_0^1 \int_0^1 \kappa(x, y)^2 d\mu(x) d\mu(y))^{\frac{1}{2}}$, the Hilbert-Schmidt
105 norm.

106 In the case of generalized random dot product graph, we derive the following
107 properties for $\kappa(x_u, x_v) = cx_u x_v$.

108

109 **Lemma 4.1.** *Under the generalized random dot product graph, the operator*
110 *norm of T_κ can be expressed as*

$$\|T_\kappa\| = c \left(\frac{\alpha}{\alpha + \beta} \right) \left(\frac{\alpha + 1}{\alpha + \beta + 1} \right).$$

111 *Proof.* By definition,

$$\begin{aligned} \|T_\kappa\| &\leq \|\kappa\|_{L^2} \\ &= \int_0^1 \int_0^1 c^2 x^2 y^2 d\mu(x) d\mu(y) \\ &= c(E(X^2)E(Y^2))^{\frac{1}{2}} \\ &= cE(X^2). \end{aligned}$$

By letting

$$f(y) = \left(\frac{\alpha + \beta}{\alpha} \right)^{\frac{1}{2}} \left(\frac{\alpha + \beta + 1}{\alpha + 1} \right)^{\frac{1}{2}} y,$$

112 we first verify that it has unit norm.

$$\begin{aligned} \int_0^1 f(y)^2 d\mu(y) &= \int_0^1 \left\{ \left(\frac{\alpha + \beta}{\alpha} \right)^{\frac{1}{2}} \left(\frac{\alpha + \beta + 1}{\alpha + 1} \right)^{\frac{1}{2}} y \right\}^2 d\mu(y) \\ &= \left(\frac{\alpha + \beta}{\alpha} \right) \left(\frac{\alpha + \beta + 1}{\alpha + 1} \right) E(Y^2) \\ &= 1. \end{aligned}$$

113 We have

$$\begin{aligned}
\|T_\kappa f\|_2^2 &= \int_0^1 \left\{ \int_0^1 cxy \left(\frac{\alpha+\beta}{\alpha}\right)^{\frac{1}{2}} \left(\frac{\alpha+\beta+1}{\alpha+1}\right)^{\frac{1}{2}} y d\mu(y) \right\}^2 d\mu(x) \\
&= c^2 \int_0^1 x^2 \left(\frac{\alpha+\beta}{\alpha}\right) \left(\frac{\alpha+\beta+1}{\alpha+1}\right) \left(\int_0^1 y^2 d\mu(y)\right)^2 d\mu(x) \\
&= c^2 \left(\frac{\alpha+\beta}{\alpha}\right) \left(\frac{\alpha+\beta+1}{\alpha+1}\right) E(X^2)E(Y^2)^2 \\
&= c^2 E(Y^2)^2.
\end{aligned}$$

Hence, with this particular choice of f , we have

$$\|T_\kappa f\|_2 = \|\kappa\|_{L^2} = cE(X^2) = c \left(\frac{\alpha}{\alpha+\beta}\right) \left(\frac{\alpha+1}{\alpha+\beta+1}\right).$$

114

□

115 We now consider the phase transition of the appearance of the giant component.
116 Formally, define $C_l(G_n)$ to be the l -th largest component in the random graph
117 G_n . The following results are due to Bollobás et al. (2007).

118 **Lemma 4.2.** (Bollobás et al. (2007) Theorem 3.1)

119 *If $\|T_\kappa\| \leq 1$, then $C_1(G_n) = o_p(n)$, while if $\|T_\kappa\| > 1$, then $C_1(G) = \theta(n)$ a.s.*

120 **Lemma 4.3.** (Bollobás et al. (2007) Theorem 3.2)

121 *If $\|T_\kappa\| < 1$, and $\sup_{x,y} \kappa(x,y) < \infty$, then $C_1(G_n) = O(\log(n))$ a.s. If $\|T_\kappa\| >$
122 1 , and either $\inf_{x,y} \kappa(x,y) > 0$ or $\sup_{x,y} \kappa(x,y) < \infty$, then $C_2(G_n) = O(\log(n))$
123 a.s.*

124 By application of Lemma 4.2 and Lemma 4.3, we have the following results
125 concerning the phase transition of the generalized random dot product graph.

126 **Proposition 4.4.** *If $c \leq \frac{(\alpha+\beta)(\alpha+\beta+1)}{(\alpha+1)\alpha}$, $C_1(G_n) = o_p(n)$, while if $c > \frac{(\alpha+\beta)(\alpha+\beta+1)}{(\alpha+1)\alpha}$,
127 $C_1(G) = \theta(n)$ a.s.*

128 **Proposition 4.5.** *If $c \leq \frac{(\alpha+\beta)(\alpha+\beta+1)}{(\alpha+1)\alpha}$, $C_1(G_n) = O(\log(n))$ a.s., while if
129 $c > \frac{(\alpha+\beta)(\alpha+\beta+1)}{(\alpha+1)\alpha}$, $C_2(G_n) = O(\log(n))$ a.s.*

131 We now investigate the degree of vertices in generalized random dot product
132 graph. We let $\chi(k)$ be the number of vertices with degree k , and define $\lambda(x) =$
133 $\int_S \kappa(x,y) d\mu(y)$.

134 The following result is due to Bollobás et al. (2007).

135

Lemma 4.6. (Bollobás et al. (2007) Theorem 3.13)

We have for any fixed k ,

$$\frac{\chi(k)}{n} \rightarrow \int_S \frac{\lambda(x)^k}{k!} e^{-\lambda(x)} d\mu(x)$$

136 *in probability.*

137 By the Lemma 4.6, we have the following result for the generalized random dot
138 product graph.

139

Proposition 4.7.

$$\frac{\chi(k)}{n} \rightarrow \int_0^1 \frac{(cx \frac{\alpha}{\alpha+\beta})^k}{k!} e^{-cx \frac{\alpha}{\alpha+\beta}} x^{\alpha-1} (1-x)^{\beta-1} dx$$

140 *in probability.*

141 *Proof.* We have

$$\begin{aligned} \lambda(x) &= \int_0^1 cxy d\mu(y) \\ &= cx \frac{\alpha}{\alpha+\beta} \end{aligned}$$

142 and the result follows by Lemma 4.6. □

143 4.2. Asymptotic Equivalence of Random Graphs

144 In this section, we show that when the connection probability between two
145 nodes u and v is given by $cx_u x_v / n^{2-s}$ with $s < 1$, the generalized random
146 dot product graph is asymptotically equivalent to the stochastic version of
147 the Chung-Lu model (Chung and Lu, 2003). Janson (2010) investigated the
148 asymptotic behavior between two random graphs, and proved conditions under
149 which the asymptotic equivalence holds. Formally, two random graphs
150 $G(n, p(u \sim v))$ and $G(n, p'(u \sim v))$ defined on the same probability space are
151 said to be asymptotically equivalent, i.e., $G(n, p(u \sim v)) \cong G(n, p'(u \sim v))$, if
152 $P(G(n, p(u \sim v)) \neq G(n, p'(u \sim v))) \rightarrow 0$, as $n \rightarrow \infty$. The following result is
153 due to Janson (2010).

154

Lemma 4.8. (Janson (2010) Corollary 2.12)

155 *Let, for each n , $p = \{p_{uv}\}$ and $p' = \{p'_{uv}\}$ be random matrices of edge proba-*
156 *bilities. If*

$$\sum_{u < v} \frac{(p_{uv} - p'_{uv})^2}{p_{uv}} = o_p(1),$$

158 *then $G(n, p) \cong G(n, p')$*

We consider the stochastic version of the Chung-Lu model investigated in Chung and Lu (2003), and show that it is asymptotically equivalent to the generalized random dot product graph model under certain constructions. Chung-Lu model

assumes a given expected degree sequence $\mathbf{w} = (w_1, w_2, \dots, w_n)$ in a random graph $G(n)$, and the probability p_{uv} of having an edge between u and v is

$$p_{uv} = P(u \sim v) = \frac{w_u w_v}{\sum_k w_k}.$$

159 Instead of assuming the expected degree sequence \mathbf{w} is given, as in the case of
 160 Chung-Lu model, we assume that for each vertex v in $G(n)$, $x_v \sim \text{Beta}(\alpha, \beta)$,
 161 and let

$$w_v = \frac{x_v}{n^{1-s}} \frac{\alpha}{\alpha + \beta}.$$

162 Under the construction above, the probability two vertices u and v are connected
 163 is given by

$$\begin{aligned} p'_{uv} = P'(u \sim v) &= \frac{w_u w_v}{\sum_k w_k} \\ &= \frac{x_u n^{s-1} \frac{\alpha}{\alpha + \beta} x_v n^{s-1} \frac{\alpha}{\alpha + \beta}}{\sum_k x_k n^{s-1} \frac{\alpha}{\alpha + \beta}} \\ &= \frac{x_u x_v \frac{\alpha}{\alpha + \beta}}{n^{2-s} \frac{1}{n} \sum_k x_k} \end{aligned}$$

164 To show that the stochastic version of the Chung-Lu model is asymptotically
 165 equivalent to the generalized random dot product graph, we need the following
 166 result from Durrett (2010).

167 **Lemma 4.9** (Durrett (2010) Theorem 2.5.8). *Let X_1, X_2, \dots be i.i.d. with*
 168 *$E|X_k|^p < \infty$, where $1 < p < 2$. We have*

$$\frac{1}{n} \sum_k x_k = E(x_1) + o(n^{\frac{1}{p}-1})$$

169 *a.s.*

170 By Lemma 4.8 and Lemma 4.9, we have the following result concerning asymp-
 171 totic equivalence between the generalized random dot product graph and the
 172 stochastic Chung-Lu model.

173
 174 **Proposition 4.10.** *If $s < 1$, the generalized random dot product graph $G(n, p(u \sim$
 175 $v))$ is asymptotically equivalent to the stochastic Chung-Lu model $G(n, p(u \sim$
 176 $v))$.*

177 *Proof.* Let $\epsilon > 0$, we have

$$\begin{aligned}
 P \left\{ \sum_{u < v} \frac{(p_{uv} - p'_{uv})^2}{p_{uv}} > \epsilon \right\} &= P \left\{ \sum_{u < v} \frac{x_u x_v}{n^{2-s}} \left(1 - \frac{\frac{\alpha}{\alpha+\beta}}{\frac{1}{n} \sum_k x_k} \right)^2 > \epsilon \right\} \\
 &\leq \frac{E \left\{ \sum_{u < v} \frac{x_u x_v}{n^{2-s}} \left(1 - \frac{\frac{\alpha}{\alpha+\beta}}{\frac{1}{n} \sum_k x_k} \right)^2 \right\}}{\epsilon} \\
 &\leq \frac{n^s}{2\epsilon} E \left\{ x_u x_v \left(\frac{o(n^{\frac{1}{p}-1})}{\frac{\alpha}{\alpha+\beta} + o(n^{\frac{1}{p}-1})} \right)^2 \right\} \\
 &\rightarrow 0
 \end{aligned}$$

178 if $s < 2 - 2/p$ where $p \in (1, 2)$. Hence, the result follows. \square

179 5. Conclusion

180 In this paper, we have generalized the random dot product graph proposed in
 181 Nickel (2007), derived various theoretical properties of the model under different
 182 sparsity regimes, and proved that the proposed model is asymptotically equiva-
 183 lent to a stochastic version of the Chung-Lu model (Chung and Lu, 2003).

184 Acknowledgements

185 This work was supported by the Science Foundation Ireland funded Insight
 186 Research Centre (SFI/12/RC/2289)

187 References

- 188 Bollobás, B., Janson, S., Riordan, O.. The phase transition in inhomogeneous
 189 random graphs. *Random Structures & Algorithms* 2007;31(1):3–122.
- 190 Chung, F., Lu, L.. The average distance in a random graph with given expected
 191 degrees. *Internet Mathematics* 2003;1(1):91–113.
- 192 Devroye, L., Fraiman, N.. Connectivity of inhomogeneous random graphs.
 193 *Random Structures & Algorithms* 2014;45(3):408–420.
- 194 Durrett, R.. *Probability: Theory and examples*. 4th ed. Cambridge University
 195 Press, Cambridge, 2010.
- 196 Janson, S.. Asymptotic equivalence and contiguity of some random graphs.
 197 *Random Structures & Algorithms* 2010;36(1):26–45.

- 198 Leskovec, J., Huttenlocher, D., Kleinberg, J.. Signed Networks in Social Me-
199 dia. Proceedings of the SIGCHI Conference on Human Factors in Computing
200 Systems. ACM, 1361-1370, 2010.
- 201 Leskovec, J., Huttenlocher, D., Kleinberg, J.. Predicting Positive and Nega-
202 tive Links in Online Social Networks. Proceedings of the 19th International
203 Conference on World Wide Web. ACM, 641-650, 2010.
- 204 McAuley, J., Leskovec, J.. Discovering Social Circles in Ego Networks. Pro-
205 ceedings of the 25th International Conference on Neural Information Process-
206 ing Systems 2014;8(1):4:1-4:28.
- 207 Nickel, C.L.M.. Random dot product graphs: A model for social networks.
208 Ph.D. thesis; Department of Applied Mathematics & Statistics, The Johns
209 Hopkins University; 2007.
- 210 O'Connor, L., Médard, M., Feizi, S.. Maximum Likelihood Latent Space
211 Embedding of Logistic Random Dot Product Graphs. ArXiv e-prints 2015.
- 212 Young, S.J., Scheinerman, E.R.. Random dot product graph models for social
213 networks. In: Bonato, A., Chung, F.R.K., editors. Algorithms and models for
214 the web-graph. Springer, Berlin; volume 4863 of *Lecture Notes in Computer
215 Science*; 2007. p. 138-149.
- 216 Watts, D.J., Strogatz, S.H.. Collective dynamics of 'small-world' networks.
217 Nature 1998;393(1):440-442.