Estrada, Ernesto and Sheerin, Matthew James (2015) Random rectangular graphs. [Report] (Unpublished) ,

Random Rectangular Graphs

Ernesto Estrada and Matthew Sheerin

Department of Mathematics and Statistics, University of Strathclyde,

26 Richmond Street, Glasgow, G1 1XH, U.K.

A generalization of the random geometric graph (RGG) model is proposed by considering a set of points uniformly and independently distributed on a rectangle of unit area instead of on a unit square $[0,1]^2$. The topological properties, such as connectivity, average degree, average path length and clustering, of the *random rectangular graphs* (RRGs) generated by this model are then studied as a function of the rectangle sides lengths a and $b = 1/a$, and the radius r used to connect the nodes. When $a = 1$ we recover the RGG, and when $a \to \infty$ the very elongated rectangle generated resembles a one-dimensional RGG. We provided computational and analytical evidence that the topological properties of the RRG differ significantly from those of the RGG. The connectivity of the RRG depends not only on the number of nodes as in the case of the RGG, but also on the side length of the rectangle. As the rectangle is more elongated the critical radius for connectivity increases following first a power-law and then a linear trend. Also, as the rectangle becomes more elongated the average distance between the nodes of the graphs increases, but the local cliquishness of the graphs also increases thus producing graphs which are relatively long and highly locally connected. Finally, we found the analytic expression for the average degree in the RRG as a function of the rectangle side lengths and the radius. For different values of the side length, the expected and the observed values of the average degree display excellent correlation, with correlation coefficients larger than 0.9999.

PACS: 89.75.-k; 02.10.Ox

I. INTRODUCTION

The use of graphs for representing physical systems is becoming ubiquitous in many areas of theoretical and applied physics [1]. We can mention the use of graphs in statistical mechanics and condensed matter physics, for solving Feynmann integrals as well as in the study of quantum phenomena [1, 2]. More recently, the use of graphs has been very broaded by their application in the analysis of complex systems [3–5]. In this case, those graphs receive the name of complex networks, due to the fact that they represent the skeleton of complex interconnected systems. In this case, networks are used to study a variety of physical scenarios, ranging from social and infrastructural, to biological and ecological ones. Here, we will use the terms graphs and networks interchangeably. When graphs are used to represent real-world physical systems it is necessary to have at our disposal some null model that allows us to evaluate which properties of the system have arisen from their connectivity pattern. In this sense, the common election is the use of random graphs. That are graphs with the same number of nodes and edges as the one under study, but in which the connection between the nodes is made randomly and independently [6]. There are several of these random models of great usability in current network theory, such as the Erdös-Rényi [8], the Barabási-Albert [9] or the Watts-Strogatz [10] model to mention just three.

In many real-world scenarios the networks emerge under certain geometrical constraints. This is the case of the so-called spatial networks [11], which include infrastructural networks such as road networks, airport transportation networks, etc., [11] and certain biological networks such as brain networks or the networks representing the proximity of cells in a biological tissue (see [3]). The list also includes the networks of patches and corridors in a landscape [12], the networks of galleries in animal nests [13, 14], and the networks of fractures in rocks [15], among others. The classical election of a random graph used to represent these systems are the so-called random geometric graphs [16, 17]. Here the term *random geometric graph* (RGG) is reserved for the case in which the nodes of the graph are distributed randomly and independently in a unit square and two nodes are connected if they are inside a disk of a given radius. Other graphs in which the edges are constructed by using different geometric rules will be named here generically as *random proximity graphs*.

RGGs have found important applications in the area of wireless communication devices [18–20], such as mobile phones, wireless computing systems, wireless sensor networks, etc. This was indeed the first application in mind when Gilbert proposed the very first RGG model [21]. RGGs have also found applications in areas such as modelling of epidemic spreading in spatial populations, which may include cases such the spreading of worms in a computer network, viruses in a human population, or rumors in a social network [22–26]. RGGs have been used to describe how cities have been evolving under the geometric constraints imposed by their geographic locations [27]. For a wider perspective on the applications of spatial graphs the reader is referred to the review [11].

In all the previously mentioned real-world scenarios, the shape of the location in which the nodes of the graph are distributed may play a fundamental role in the topological and dynamical properties of the resulting graphs. That is, it is intuitive to think that the connectivity, distance, clustering and other fundamental topological properties of the graphs are affected if we, for instance, elongate the unit square in which the points are distributed.

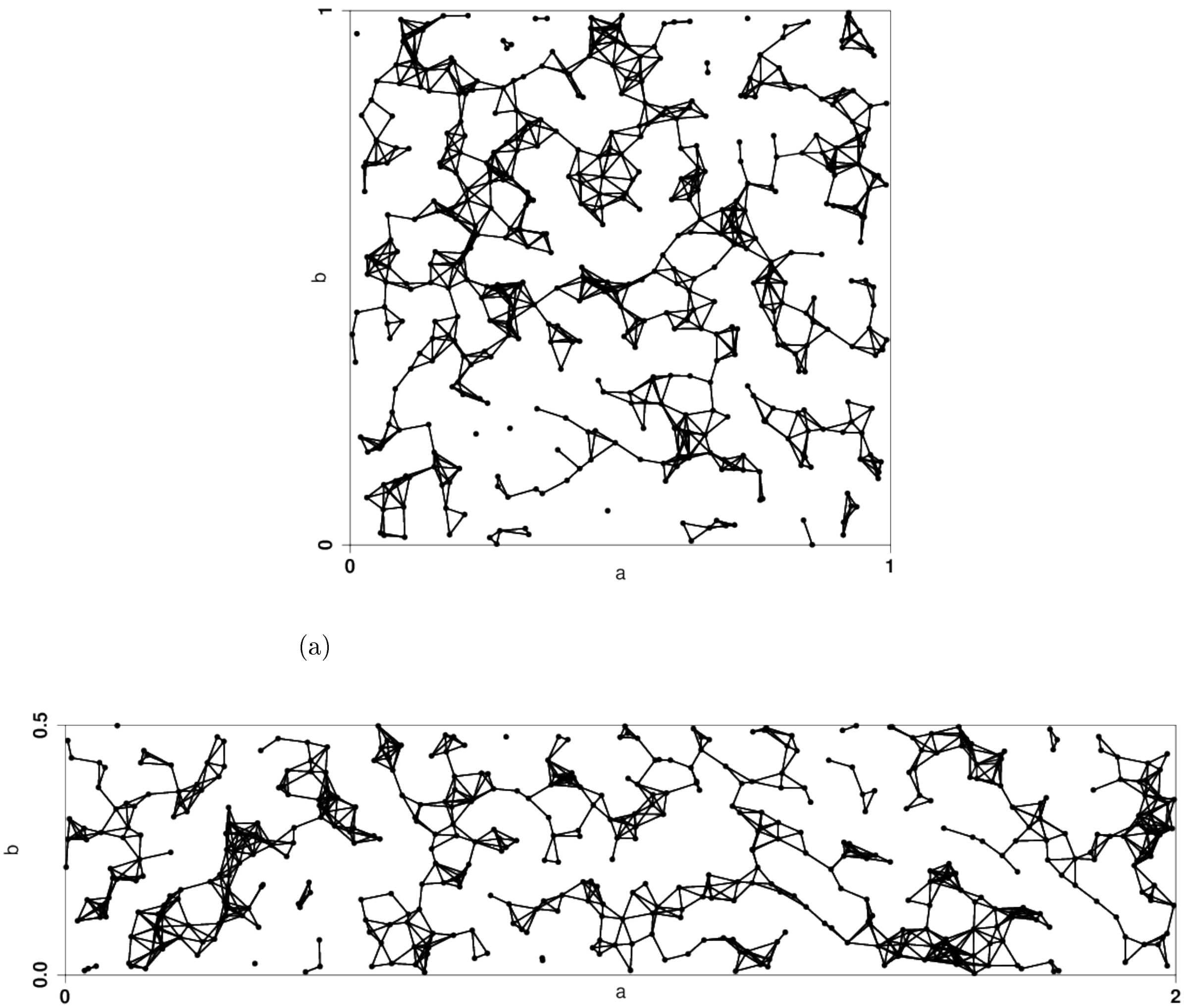
Here, we develop a new model that generalizes the RGG by allowing the embedding of the nodes in a unit rectangle instead of a unit square. Our main goal is to investigate how the elongation of a unit square influences the topological properties of the graphs generated by the model. This generalized graphs will be named here the *random rectangular graphs* (RRGs). In this work we concentrate on the influence of the ratio of the lengths of the two sides of the rectangle on the topological properties of the graphs emerging on them, such as their connectivity, average degree, average path length and clustering coefficient. In particular, we find the analytical expression of the *average degree*. The average degree is a simple but highly important property of graphs, which is related to several dynamical processes.

II. DEFINITION OF THE MODEL

The RGG is defined in general by distributing uniformly and independently n points in the unit d-dimensional cube $[0,1]^d$ [16]. Then, two points are connected by an edge if their Euclidean distance is at most r, which is a given fixed number known as the radius.

Let us now define a unit hyperrectangle as the Cartesian product $[a_1,b_1]\times[a_2,b_2]\times\cdots\times[a_d,b_d]$ where $a_i,b_i\in\mathbb{R}, a_i\leq b_i$, and $1\leq i\leq d$. Hereafter we will restrict ourselves to the 2-dimensional case, which corresponds to a rectangle of unit area, that we will call the unit rectangle. Now, the RRG is defined by distributing uniformly and independently n points in the unit rectangle $[a,b]$ and then connecting two points by an edge if their Euclidean distance is at most r. It is evident that the only change we have introduced here is to consider a rectangle of unit area instead of the analogous square. The rest of the construction process remains the same as for the RGG. This means that $RRG\rightarrow RGG$ as $(a/b)\rightarrow 1$. In this sense we can say that the RRG is a generalization of the RGG. In Fig. 1 we illustrate an RGG and an RRG constructed with the same number of nodes and edges.

An interesting question is what happen at the other extreme, when $a\rightarrow\infty$. In this case we have that $b\rightarrow 0$, which means that the n points are uniformly and independently distributed on the straight line. Let us now consider a disk of radius $r>0$ centered at each of these points and let us connect every point to the other points which lie inside its disk. For very small values of r each node can only be connected to its nearest neighbors in such a way that the whole graph is a path or a collection of paths of different lengths. As $r\rightarrow\infty$, a node



(a)

(b)

Figure 1. Illustration of two random rectangular graphs with $a = 1$ (top), which corresponds to a random geometric graph on a unit square and with $a = 2$ ($b = 0.5$) (bottom). Both graphs are built with 500 nodes and 1750 edges.

is connected not only to its nearest neighbors but to second, third, and so forth, forming a one-dimensional random graph. Thus, the resulting graph resembles a one-dimensional RGG, that is a graph created by placing the n points uniformly and independently on the interval $[0, 1]$ and then connecting pairs of nodes if they are at a Euclidean distance smaller or equal than certain radius r (see for instance [28–30]).

III. COMPUTATIONAL ANALYSIS OF RRG

In this section we study computationally a few topological properties of the RRGs. In the following we will consider RRGs with $a = 1/b$ and consequently we will report only the value of a. For instance, $a = 1$ represents a unit square and the RRG is identical to the classical RGG. For $a = 5$ we have a very elongated rectangle with sides $a = 5$ and $b = 0.2$. We study here some important structural parameters of networks, such as the connectivity, average degree, average path length and the average clustering coefficient.

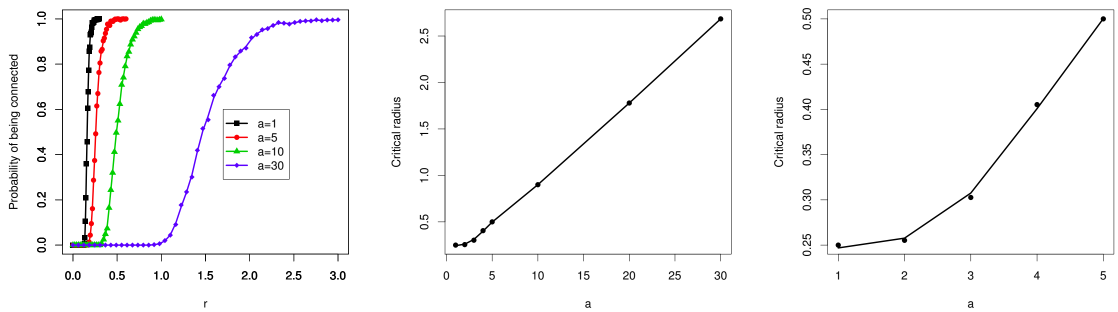
3.1 Connectivity and average node degrees

In the case of the RGGs it is a well known result that increasing the radius of the disks centered at each point produces a phase transition from a disconnected to a connected graph at certain critical radius. That is, for

$$\pi r^2 = \frac{\log n + \gamma_n}{n}, \tag{1}$$

the RGG is connected if $n \to \infty$ and $\gamma_n \to \infty$ and disconnected if $\gamma_n \to -\infty$ [16]. In the Fig. 2(a) we illustrate this result for a RGG with $n = 100$ nodes, i.e., $a = 1$, where it can be seen that the critical radius is about 0.25, which corresponds to a value of $\gamma_n \approx 15$. As the square is elongated the critical radius increases with the value of a. For instance, for $a = 5$ the critical radius is about 0.5, and for $a = 30$ it is about 3. The main reason for this increase in the critical radius is that as we elongate the rectangle the points have to cover a longer region of the rectangle and as so their separation increases. As a consequence, we need to increase the radius in order to guarantee the connectivity of the network. As can be seen in the Fig. 2 (b) there is a linear trend between the length of the side of the rectangle and the critical radius of the RRGs for values of $a \gtrsim 5$. For the values $1 \leq a \lesssim 5$ the relation between the critical radius and the side length of the rectangle is a power-law of the form $r_c \sim a^{3.72}$.

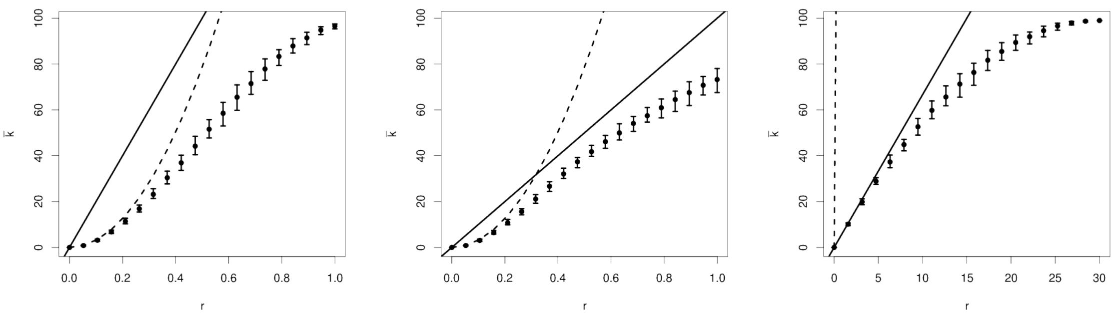
We now analyze the average degree of the nodes in the RRGs. A well-known result in the theory of RGGs, i.e, when $a = b = 1$, is that the average degree is approximated by $\langle k \rangle \simeq \pi r^2 n$ [16], such that $\langle k \rangle \sim r^2$. This relation is only true for very small values of r because for larger radii it is known that the border effects play a fundamental role in the



(a) (b) (c)

Figure 2. (color online) Probability that the RRG is connected as a function of the radius for graphs with $n = 100$ (a). Dependence of the critical radius for connectivity with the side length of the rectangle for general (b) and small values of a (c). Every point is the average of 1000 random realizations

deviations from this scaling. For larger values of r it is expected that $\langle k \rangle \simeq n - 1$ due to the higher density of the resulting graph. Consequently, the change of $\langle k \rangle$ with r is expected to be quadratic for small r and then change its behavior for larger values of the radius as a consequence of the increase in the border effects. In the case of the RRGs the border effects along the longest edge of the rectangle are much bigger than for the unit square. Thus, we would expect that this transition from the quadratic to the non-quadratic behavior is more dramatic in the RRG than in the RGG. This is illustrated in the Fig. 3, where it can be seen that in the RGG the quadratic approximation is indeed very good for values of the radius $0 < r \leq 0.3$. In a RRG with $a = 2$ this quadratic approximation is still good for small values of the radius. However, as it can be seen, the linear approximation is better for the RRG with $a = 2$ than for the RGG. The straight line shown in both plots corresponds to the expected linear relation between $\langle k \rangle$ and r if we consider that the circle of small enough radius r around a typical node can be well-approximated by a very thin rectangle of length $2r$ and width b. In this case $\langle k \rangle \simeq (2b)\, r$, where obviously $2b$ is a constant for a given rectangle and a linear trend instead of a quadratic one is expected. If we elongate more the rectangle, e.g., by taking $a = 30$ (see Fig. 3 (right)), this linear approximation remains for a wider range of the radii $0 < r \leq 5$, which indicates that indeed it corresponds to the approximation of the circle by a thin rectangle when the RRG is very narrow. In this case the quadratic approximation is very bad and only valid for a very narrow region of values of the radius.



(a) (b) (c)

Figure 3. Illustration of the change in the average degree with the radius for (a) a random geometric graph in a unit square, (b) a random rectangular graph with $a = 3$, and (c) with $a = 30$. All graphs have $n = 100$. The dotted line corresponds to the quadratic approximation of $\langle k \rangle$ with r (see text) and the solid line corresponds to its linear approximation.

An important observation extracted from the plots of the average degree versus the radius is the existence of three different regimes in these plots. Due to their sigmoid shapes we observe that the dependance of the average degree with the radius is different for the regions $0 \leq r \leq b$, $b \leq r \leq a$ and $a \leq r \leq \sqrt{a^2 + b^2}$. This is important because we will use these three regimes for the analytic calculation of the average degree in general RRGs.

3.3 Average path length and clustering

Let $\Gamma = (V, E)$ be a simple connected graph. A *path* of length k in Γ is a set of nodes $i_1, i_2, \ldots, i_k, i_{k+1}$ such that for all $1 \leq l \leq k$, $(i_l, i_{l+1}) \in E$ with no repeated nodes. The *shortest-path* or *geodesic distance* between two nodes $u, v \in V$ is defined as the length of the shortest path connecting these nodes. We will write $d(u, v)$ to denote the distance between u and v. Here we will call, as usually in network theory, average path length to the following quantity:

$$\langle l \rangle = \frac{2}{n\,(n-1)} \sum_{u<v} d\,(u, v)\,. \tag{2}$$

On the other hand, the *local clustering coefficient* of a node u, which quantifies the degree of transitivity of local relations in a network is defined as [10]:

$$C_u = \frac{2|\{(v,w) : v, w \in N_u; (v,w) \in E\}|}{k_u(k_u - 1)}, \tag{3}$$

where $N_u = \{v : (u,v) \in E\}$ and k_u is the degree of the node u. Taking the mean of these values as u varies among the nodes in Γ, one gets the *average clustering coefficient* of the network: $\langle C \rangle = \frac{1}{n}\sum_{u=1}^{n} C_u$.

We study here graphs with 1000 nodes and 7500 edges. For every value of a we report the average of 10 random realizations. In the Fig. 4 we illustrate the variation of the average path length and average clustering coefficients for these graphs. The plot of $\langle l \rangle$ versus a agrees with our intuition that as we elongate the rectangle there are nodes which are more far apart from each other and as a result the average path length of the whole graph increases. There is an almost linear increase of $\langle l \rangle$ for values of $1 \le a \lesssim 15$ after which the dependence is very flat. In this region we have that $a \to \infty$, which corresponds to a good approximation to a one-dimensional RGG. For $a = 1$ it is known that the average path length depends on the inverse of the radius, $\langle l_e \rangle = \Theta(1/r)$ [31]. The actual radius used for the plot in Fig. 4 is $r = 0.0713$ which gives an estimate of the average path length of 14.02, which is not too far from the observed value in the plot for $a = 1$. In the case of $a = 30$ we are in the presence of a very elongated rectangle, which is very similar to a one-dimensional RGG. A crude estimate of the average path length in this case would be $\langle l_e \rangle = n/\langle k \rangle$, which in the current case will give $\langle l_e \rangle \approx 66.6$, which is relatively close to the observed value of $\langle l \rangle \approx 50$ for $a = 30$.

It is interesting to note that the average clustering coefficient also increases as the rectangle becomes more elongated. This is a consequence of the fact that we are now compressing the nodes into a narrower region, which allow them to be locally closer to each other and create more triangles. However, as soon as $a \ge 15$ for these graphs (this value will depend on the number of nodes of the graph) the dependence of the average path length and clustering coefficient with the side length is very flat. This is due to the fact that for large enough values of a the graphs behave as a one-dimensional random geometric graph.

Let $r^2 = \dfrac{\log n + \gamma_n}{n\pi}$ as in (1), then if $n \to \infty$ and $\gamma_n \to \infty$, it is known that the average clustering coefficient is given by [17]

$$\langle C_d \rangle = \begin{cases} 1 - H_d(1) & d \text{ } even \\ \frac{3}{2} H_d(1/2) & d \text{ } odd, \end{cases} \tag{4}$$

where d is the dimension of the hypercube in which the nodes are embedded and

$$H_d(x) = \frac{1}{\sqrt{\pi}} \sum_{i=x}^{d/2} \frac{\Gamma(i)}{\Gamma\left(i+\frac{1}{2}\right)} \left(\frac{3}{4}\right)^{i+\frac{1}{2}}, \tag{5}$$

where $\Gamma(i)$is the Gamma function. Thus, for $d = 2$, $\langle C_2 \rangle = 1 - \frac{3\sqrt{3}}{4\pi} \approx 0.5865$ and for $d = 1$, $\langle C_1 \rangle = 3/4 = 0.75$.

As can be seen in the Fig. 4 for $a = 1$ the average clustering coefficient is $\langle C \rangle \approx 0.61$, which is very close to the expected value for the 2-dimensional RGG. When $a = 30$ the average clustering coefficient is $\langle C \rangle \approx 0.75$, which coincides with the exact value expected for the one-dimensional RGG. Consequently, the RRG generalizes the values of the clustering coefficient of both, the one- and two-dimensional RGG, for $a = 1$ and $a \rightarrow \infty$, respectively. In addition, it provides a series of intermediate values of the clustering coefficient for intermediate values of the side length of the rectangle.

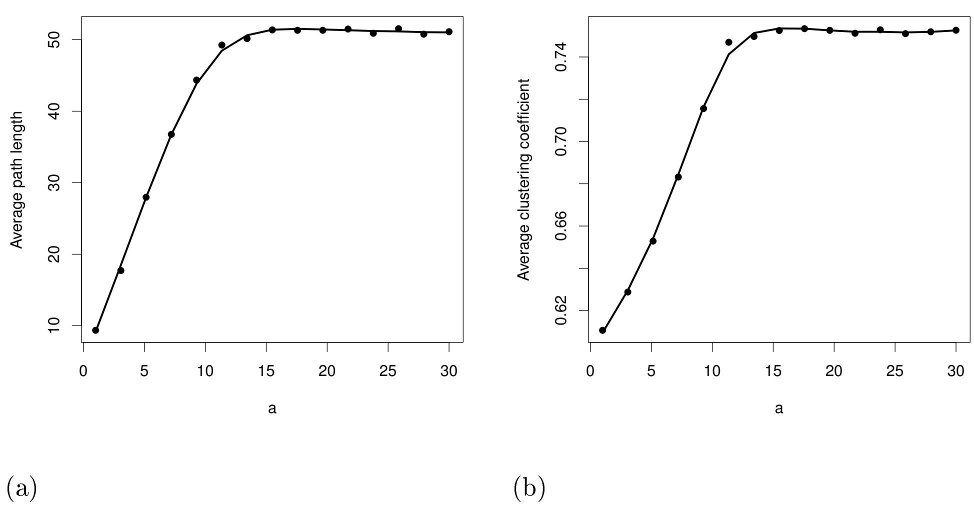
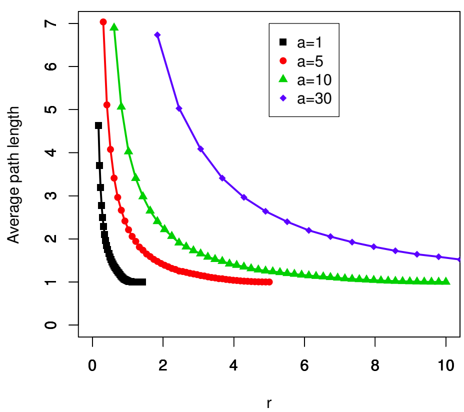
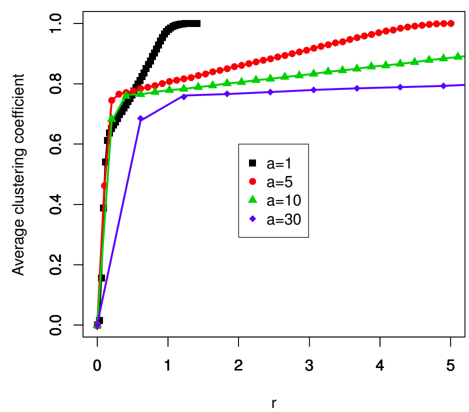


Figure 4. Change of the average path length (left) and average clustering coefficient (right) in a RRG with the systematic variation of the side length a of the rectangle for graphs having $n = 1000$ and $\langle k \rangle = 15$. Every point represents the average of 10 realizations. The standard deviations of each point are not illustrated for the sake of clarity.

We now further explore the relation between the radius r and the average path length and clustering for RRGs with different side lengths. We consider graphs with $n = 100$ nodes and the extreme cases $a = 1$ (RGG) and $a = 30$. As the radius increases the graph is becoming more and more dense, which is reflected in the exponential decay of the average path length to the value $\langle l \rangle = 1$, which corresponds to that of a complete graph. There is not substantial differences in the decay of the average path length with the increase of the radius for $a = 1$ and $a = 30$. For the average clustering coefficient the results for both cases are very similar and they are characterized by an abrupt increase in the clustering at the beginning of the plot and then a linear increase until the value of $\langle C \rangle = 1$ is reached for the complete graph.



(a)



(b)

Figure 5. (color online) Variation of the average path length with the radius (a), as well as the variation of the average clustering coefficient with the radius (b) for graphs with $a = 1, 5, 10, 30$. Every point is the average of 100 random realizations.

4 ANALYTICAL RESULTS FOR $\langle k \rangle$

Given a node, there are $n-1$ nodes distributed in the rest of the rectangle. Define A_p to be the area within radius r of a point p which lies within the rectangle. Since nodes are uniformly and independently distributed, the expected degree of a node v_i is $\mathbf{E}(k_i) = (n-1)A_i/(ab)$, where A_i is taken for the point where node v_i is located. This is because dividing the nodes between the area within distance r and the rest of the rectangle gives rise to the Binomial distribution $Bin(n-1, A_i/(ab))$ as it can be considered like a partition of a Poisson process. Averaging this over all possible node locations (i.e., the points in the rectangle) gives

$$\mathbf{E}\langle k\rangle = \frac{\int_p \{(n-1)A_p/(ab)\}}{ab} = \frac{(n-1)\int_p A_p}{(ab)^2} \tag{6}$$

Let $f(a,b,r)$ to be the area within radius r of a point which lies in the rectangle, integrated over all points, i.e., $f(a,b,r) = \int_p A_p$. Based on the computational results obtained for the average degree we consider here the previously detected regions: $0 \leq r \leq b$, $b \leq r \leq a$ and $a \leq r \leq \sqrt{a^2+b^2}$, recalling that $a \geq b$. We call these cases $i = 1, 2, 3$, respectively. Thus, the function $f(a,b,r)$ takes different forms f_i for each case i. This means that we can write $\mathbf{E}\langle k\rangle = \dfrac{(n-1)f_i}{(ab)^2}$ with

$$f_i = \begin{cases} f_1 & 0 \leq r \leq b \\ f_2 & b \leq r \leq a \\ f_3 & a \leq r \leq \sqrt{a^2+b^2} \end{cases} \tag{7}$$

and our task is now to find the analytical expressions for f_i for these three cases separately.

Case 1

This case corresponds to the covering of each point in the rectangle by a circle of small radius, $0 \leq r \leq b$. Let us fix a value of the radius to r. The area of this circle is $A_{\bigcirc} = \pi r^2$. Thus, if we consider intersecting circles covering the whole rectangle of area $A_{\square} = ab$, the total area covered is:

$$A_{\boxdot} = A_{\bigcirc} A_{\square} = \pi r^2 ab.$$

The problem is that many of these circles have segments outside the rectangle. Thus, the question is to calculate the area coming from the contribution of those circles which are not entirely inside the rectangle. In order to obtain this contribution we start by considering a circle with radius $0 \leq r \leq b$ located at the center of the rectangle. We now displace the circle to the edge of length b of the rectangle and allow that segments of this edge define chords of the circle. We stop the displacement when there is a semicircle outside (and another inside) the rectangle. We then have an infinite collection of segments of the circle (see Fig. 6). We proceed by stacking each segment of the circle over the other, starting from the semicircle, in such a way that they define a section of a cylinder that has been intersected by a plane as illustrated in the Fig. 6. The sum of the areas of all these segments of the circle equals the contribution of one circle to the area outside the rectangle. This total area is easy to calculate by simply considering it equal to the volume of the section of the cylinder:

$$V = \int_{-r}^{r} \left(r^2 - x^2\right) dx = \frac{4}{3} r^3. \tag{8}$$

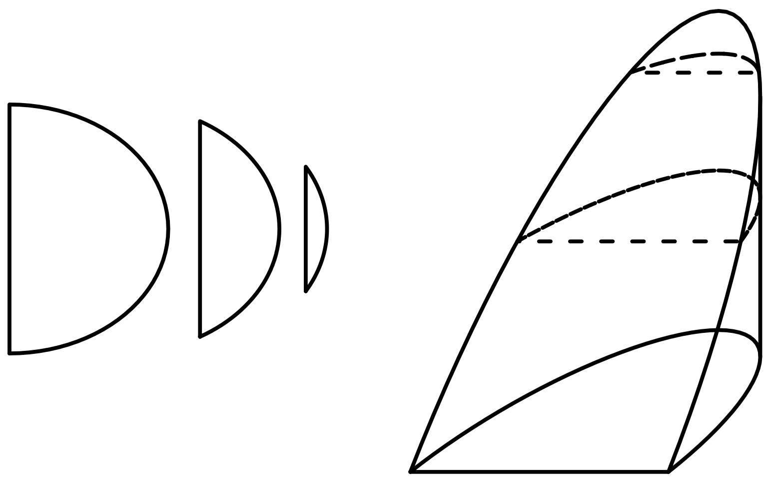


Figure 6. Illustration of the stacking of the segments of the circle which lie out the rectangle (left), and the section of the cylinder formed by all the stacked segments (right).

Using this volume, which corresponds to the area of a circle outside the rectangle, we can calculate the total area coming from all circles moving in the direction right-left (R-L) as well as those moving in the direction top-bottom (T-B). That is, the first area is given by Vb and the second is given by Va. The problem is that we are counting twice the area for some points which are in the square with area r^2 which is located at each of the four corners of the rectangle (see Fig. 7). In order to account for this area we consider a quarter of a circle (a pie) moving in the R-L direction and don't count the contribution for the area of the square at the corner. That is, we obtain the total areas in the R-L and T-B directions as:

$$A_{R-L} = \frac{1}{2} V(b - r), \tag{9}$$

$$A_{T-B} = \frac{1}{2} Va. \tag{10}$$

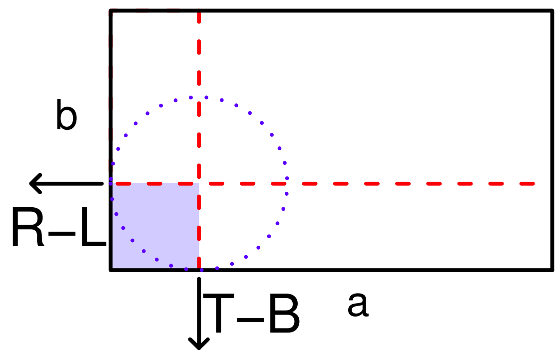


Figure 7. (color online) Illustration of a circle centered at a point inside the rectangle with sides a and b and separated from the edges of the rectangle by a distance equal to the radius of the circle r. The arrows R-L and T-B indicate the directions of displacement of the circle used to calculate the areas outside the rectangle. The square at the corner which has sides of length equal to r is shadowed. the section of the circle which is inside this square corresponds exactly to a quarter of the circle.

For the area of the square at the corner we have already its contribution in the T-B direction. Now for contribution in the R-L direction we must consider that some quarter circles protrude both below and to the left of the rectangle. We then calculate the R-L contribution in such a way that we do not double-count anything

$$A_{\square} = \int_0^r \int_0^t \frac{1}{2}\left(r^2 - x^2\right) dx dt = \frac{5}{24} r^4. \tag{11}$$

We are now in condition to calculate the total area of the circles covering only the space inside the rectangle when $0 \leq r \leq b$, which is

$$\begin{aligned} f_1 =& A_{\boxdot} - 4\left(A_{R-L} + A_{T-B} + A_{\square}\right) & (12) \\ =& \pi r^2 ab - \frac{4}{3}(a+b) r^3 + \frac{1}{2} r^4. & (13) \end{aligned}$$

Notice that we have multiplied the parenthesis in (12) by 4 because we have previously considered the areas of quarter circles.

Case 2

In this case every point is covered by a circle of radius, $b \leq r \leq a$. We take a similar approach to the one used in Case 1 with the following adaptation. Taking the bottom-left quadrant of the circle as before, there is always part of the quarter circle protruding from the bottom of the rectangle. Equivalently, every circle now protrudes from both the top and bottom of the rectangle. This makes certain geometric arguments used in Case 1 invalid, such that the one of being able to fit a square of length r into the rectangle. For points at distance t from the top edge of the rectangle, the area of this protrusion is $\int_t^r \sqrt{r^2 - x^2}dx$, and integrating over t gives

$$\begin{aligned} V' &= \int_0^b \int_t^r \sqrt{r^2 - x^2}\, dx\, dt \\ &= \frac{1}{4}\pi r^2 b - (\frac{1}{3}r^2 + \frac{1}{6}b^2)\sqrt{r^2 - b^2} - \frac{1}{2}r^2 b \arcsin(\frac{b}{r}) \\ &\quad + \frac{1}{3}r^3 \end{aligned} \tag{14}$$

This considers all the points in a vertical line through the rectangle, so we multiply by the length a to get

$$A_{T-B,2} = aV' \tag{15}$$

In this case we can no longer fit a square of length r inside the rectangle, so we modify $A_{\square}$ accordingly and obtain

$$\begin{aligned} A_{\square,2} &= \int_0^b \int_0^t \frac{1}{2}\left(r^2 - x^2\right) dxdt \\ &= \frac{1}{4}r^2 b^2 - \frac{1}{24}b^4 \end{aligned} \tag{16}$$

We now obtain the total area f_2 in a similar way as for f_1

$$\begin{aligned} f_2 &= A_{\boxdot} - 4(A_{T-B,2} + A_{\square,2}) \\ &= -\frac{4}{3}ar^3 - r^2b^2 + \frac{1}{6}b^4 + a(\frac{4}{3}r^2 + \frac{2}{3}b^2)\sqrt{r^2 - b^2} + 2r^2 ab \arcsin(\frac{b}{r}). \end{aligned} \tag{17}$$

Case 3

In this case the circles have radius $a \leq r \leq \sqrt{a^2+b^2}$. Here we consider a slightly different approach because the geometry of the system involved changes in relation to the previous cases. In this case all of the quarter circles protrude from both the left and bottom edges of the rectangle, and thus all circles extend beyond all sides of the rectangle. This means that for many points, the overlap between the (bottom-left quadrant) quarter circle and the rectangle corresponds to a smaller rectangle, though for some points near the top-right this is not true. We assume first that this overlap is always a rectangle and correct this later. For a point of distance $0 \leq x \leq a$ from the left of the rectangle and distance $0 \leq y \leq b$ from the bottom the area is xy, and we integrate these rectangular areas over all points to obtain

$$A_R = \int_0^b \int_0^a xy\, dx\, dy = \frac{1}{4}a^2b^2. \tag{18}$$

The quarter circles for some points in the top-right do not fully cover the bottom-left of the rectangle, so we calculate what we must subtract to account for these interior areas. We consider an affected point on the top edge of the rectangle, and displace this in the $T-B$ direction. This produces shapes such as in Fig. 8 (left), which may be stacked in a similar way to the circular segments of Case 1 to produce a solid with the shape illustrated in the Fig. 8(right).

We now calculate the volume of a point at distance $\sqrt{r^2-b^2} \leq t \leq a$ from the top-left corner of the rectangle, and make use of the fact that its cross-sections in one axis are right triangles

$$V'' = \int_{\sqrt{r^2-b^2}}^{t} \frac{1}{2}(b-y)^2\, dx \tag{19}$$

We now integrate this over t, to find the value A_I which we subtract to account for the interior areas

$$\begin{aligned} A_I = \int_{\sqrt{r^2-b^2}}^{a} V''dt = &\frac{1}{4}(a^2b^2 + a^2r^2 + b^2r^2) - \frac{1}{24}(a^4+b^4) + \frac{1}{8}r^4 \\ &- b(\frac{1}{3}r^2 + \frac{1}{6}a^2)\sqrt{r^2-a^2} - a(\frac{1}{3}r^2 + \frac{1}{6}b^2)\sqrt{r^2-b^2} \\ &+ \frac{1}{2}abr^2(\arccos(\frac{b}{r}) - \arcsin(\frac{a}{r})). \end{aligned} \tag{20}$$

We now have everything we need to write f_3, which has the following form

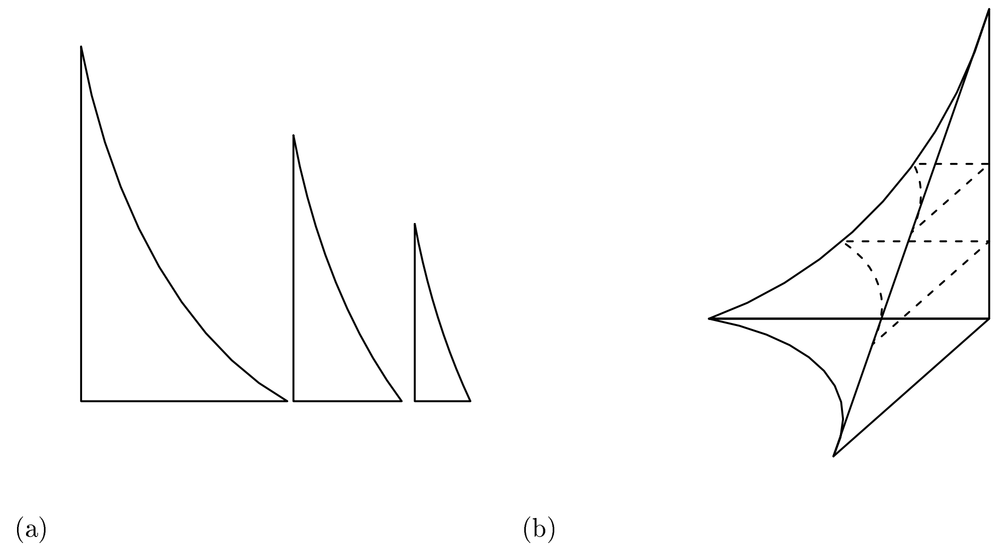


Figure 8. Illustration of the interior areas which are not covered by the quarter circles, which get smaller as the quarter circle is displaced downwards (left), and the solid formed by all the stacked areas (right). Note that the cross-sections along the horizontal axis are right triangles

$$
\begin{aligned}
f_3 &= 4(A_R - A_I) \\
&= -r^2(a^2+b^2) + \frac{1}{6}(a^4+b^4) - \frac{1}{2} \\
&\quad +b(\frac{4}{3}r^2 + \frac{2}{3}a^2)\sqrt{r^2-a^2} + a(\frac{4}{3}r^2 + \frac{2}{3}b^2)\sqrt{r^2-b^2} \\
&\quad -2abr^2(\arccos(\frac{b}{r}) - \arcsin(\frac{a}{r}))
\end{aligned} \tag{21}
$$

In the next section of this work we compare these analytical results with the average degree observed for different RRGs.

5 ANALYTICAL VS. COMPUTATIONAL RESULTS FOR $\langle k \rangle$

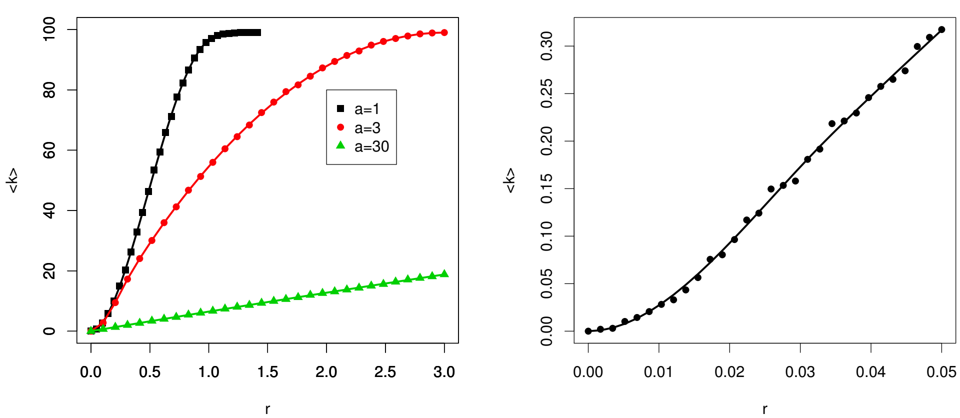
Here we analyze the goodness of fit of the values of the average degree observed in RRGs as a function of the radius for different values of the sides of the rectangle. We recall that the expected average degree of a RRG is given by

$$\mathbf{E}\langle k \rangle = \frac{(n-1)f_i}{(ab)^2}, \tag{22}$$

where

$$f_i = \begin{cases} 0 \le r \le b & \pi r^2 ab - \frac{4}{3}(a+b)r^3 + \frac{1}{2}r^4 \\ b \le r \le a & -\frac{4}{3}ar^3 - r^2b^2 + \frac{1}{6}b^4 + a(\frac{4}{3}r^2 + \frac{2}{3}b^2)\sqrt{r^2 - b^2} \\ & +2r^2ab\arcsin(\frac{b}{r}) \\ a \le r \le \sqrt{a^2+b^2} & +b(\frac{4}{3}r^2 + \frac{2}{3}a^2)\sqrt{r^2-a^2} + a(\frac{4}{3}r^2 + \frac{2}{3}b^2)\sqrt{r^2-b^2} \\ & -2abr^2(\arccos(\frac{b}{r}) - \arcsin(\frac{a}{r})). \end{cases} \quad (23)$$

In the Fig. 9 we illustrate the results for three RRGs with $n = 100$ nodes and values of $a = 1, 3, 30$, respectively. The solid circles represent the observed values of the average degree for the corresponding graphs averaged over 100 random realizations. The solid line is the expected values according to the expressions (22) and (23). The Pearson correlation coefficients for the linear regression between the observed and expected values is larger than 0.9999 in the three cases. We enlarge the region of small radii for the case $a = 30$ (see Fig. 9) where it can be seen that it is a perfect fit also for this region, here the Pearson correlation coefficient is 0.994.



(a) (b)

Figure 9. (color online) Illustration of the fit between the observed (black circles) and expected (solid line) values of the average degree for RRGs with different side lengths of the rectangle. (a) An RGG with $a = 1, 3, 30$, (b) $a = 30$ for small radii.

6 CONCLUSIONS AND FUTURE OUTLOOK

We have introduced here a generalization of the RGG in which we embed the points into a unit rectangle instead of on a unit square. We consider a rectangle with sides of lengths a and $b = 1/a$, such as when $a = 1$ we have the particular case of the 'classical' random geometric graph embedded in a unit square. Also, when $a \to \infty$ we have a very elongated rectangle which resembles a one-dimensional RGG. We have provided computational and analytical evidence that reaffirm the fact that the topological properties of the RRG differ significantly from those of the RGG. We have seen that the connectivity of the RRG depends not only on the number of nodes as in the case of the RGG, but also on the length of the side of the rectangle. In particular, as the length of the side of the rectangle increases, i.e., the rectangle is more elongated, the critical radius for connectivity increases following first a power-law and then a linear trend. In other words, by keeping the number of nodes and the radius constant, the connectivity increases as the area in which the points are located is more regular, i.e., more squared.

The analysis of the average path length and clustering coefficient indicate that as the rectangle becomes more elongated the average distance between the nodes of the graphs increases due to the fact that the nodes have to cover a longer region in the rectangle than in the square. However, the graphs are also locally more connected as $a \to \infty$ as reflected by the increase in the average clustering coefficient. Then, the elongation of the rectangle makes graphs which are relatively long and highly locally connected.

We also found the analytic expression for the average degree in the RRG. In this case we have discovered that there are three regimes for the values of the radius in terms of the length of the sides of the rectangle. The expected value of the average degree is then expressed as functions of the lengths of the rectangle and the radius. We have shown that for different values of the side length, the expected and the observed values of the average degree display excellent correlation, with correlation coefficients larger than 0.99.

The introduction of the RRGs open new possibilities for studying spatially embedded random graphs. There are many open questions that derive from this work, such as the search for analytical expressions for the average path length, clustering coefficient and other topological properties as a function of the side length of the rectangle. Also the study of dynamical processes taking place on the nodes and edges of these graphs is of great interest to explore how the shape constraints influence the dynamics on the RRGs. The analysis rectangular proximity graphs, such as the rectangular Gabriel graphs and random rectangular neighborhood graphs on is also interesting for many of the practical applications of these graphs as mentioned in the introduction. The generalization of the RRG model to higher dimensions is also of both theoretical and practical interest. In closing, the current work is expected to open new horizons for the study of random spatial graphs and its applications in physics and beyond.

IV. ACKNOWLEDGMENT

EE thanks the Royal Society for a Wolfson Research Merit Award. MS thanks Weir Advanced Research Centre at Strathclyde and EPRSC for partial financial support of this work.

---

[1] E. Estrada, *Graphs and Networks*, in *Mathematical Tools for Physicists*, edited by M. Grinfeld, (John Wiley & Sons, 2014).

[2] G. Berkolaiko, Amer. Math. Soc. **415** (2006).

[3] E. Estrada, *The Structure of Complex Networks: Theory and Applications*, (Oxford University Press, 2011).

[4] M. E. J. Newman, SIAM Rev. **45**, 167 (2003).

[5] L. d. F. Costa, O. Oliveira, G. Travieso, F. A. Rodrigues, P. Villas Boas, L. Antiqueira, M. Viana, and L. Correa Rocha, Adv. Phys. **60**, 329 (2011).

[6] M. J. E. Newman, Preprint arXiv:cond-mat/0202208 (2002).

[7] B. Bollobás, *Random Graphs*, (Academic Press, New York, 1985).

[8] P. Erdös and A. Rényi, *On the evolution of random graphs*, Selected Papers of Alfréd Rényi, Vol. 2, 482-525 (1976).

[9] A. L. Barabási and R. Albert, Science **286**, 509-512 (1999).

[10] D. J. Watts and S. H. Strogatz, Nature **393**, 440-442 (1998).

[11] M. Barthélémy, Physics Reports **499**, 1-101 (2011).

[12] D. Urban and T. Keitt, Ecology, **82**, 1205 (2001).

[13] A. Perna, S. Valverde, J. Gautrais, C. Jost, R. Solé, P. Kuntz and G. Theraulaz, Physica A, 387:6235-6244, (2008).

[14] J. Buhl, J. Gautrais, R.V. Solé, P. Kuntz, S. Valverde, J.L. Deneubourg, and G. Theraulaz, Eur. Phys. J. **B42**, 123 (2004).

[15] E. Santiago, J. X. Velasco-Hernández, and M. Romero-Salcedo, Expert Systems with Applications 41(3):811-820, (2014).

[16] M. Penrose, *Random geometric graphs* (Oxford University Press, 2003).

[17] J. Dall, and M. Christensen, Phys. Rev. E **66** (2002).

[18] P. Gupta and P.R. Kumar, *Critical Power for asymptotic connectivity in wireless networks*, in *Stochastic analysis, control, optimization and applications* (Birkhäuser Boston, 1999).

[19] G. J. Pottie and W. J. Kaiser, Communications of the ACM **43** 51-58, 5 (2000).

[20] D. Estrin, R. Govindan, J. Heidemann and S. Kumar, *Next century challenges: Scalable coordination in sensor networks*, in Proceedings of the ACM/IEEE International Conference on Mobile Computing and Networking (Seattle, Washington, USA, August 1999), p. 263-270.

[21] E. N. Gilbert, Ann. Math. Stat. **30** 1141 (1959) 1141-1144.

[22] P. Wang and M. C. González, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences **367.1901** 3321-3329 (2009).

[23] A. Díaz-Guilera, J. Gómez-Gardeñes, Y. Moreno, and M. Nekovee, Int. J. Bif. Chaos **19**, 687 (2009).

[24] M. Nekovee, New J. Phys. **9(6)**, 189 (2007).

[25] V. Isham, J. Kaczmarska, and M. Nekovee, Phys. Rev. E **83(4)** (2011).

[26] Z. Toroczkai, and H. Guclu, Physica A **378(1)**, 68-75 (2007).

[27] D. Watanabe, *A study on analyzing the grid road network: patterns using relative neighborhood graph*, The Ninth International Symposium on Operations Research and Its Applications (ISORA'10), Chengdu-Jiuzhaigou, China, ORSC & APORC. (2010).

[28] [3] M. Desai and D. Manjunath, Communications Letters, IEEE **6(10)**, 437-439 (2002).

[29] C.H. Foh, G. Liu, B. S. Lee, B. C. Seet, K. J. Wong and C.P. Fu, Communications Letters, IEEE **9(1)**, 31-33 (2005).

[30] E. Godehardt and J. Jaworski, Random Structures and Algorithms **9**, 137-161 (1996).

[31] G. Ercal, *Small worlds and rapid mixing with a little more randomness on random geometric graphs*, in NETWORKING (Springer Berlin Heidelberg, 2011), p. 281-293.