Structural properties of spatially embedded networks

K. Kosmidis[1(a)], S. Havlin[2] and A. Bunde[1]

[1] *Institut für Theoretische Physik III, Justus-Liebig-Universität Giessen - 35392 Giessen,Germany, EU*
[2] *Minerva Center and Department of Physics, Bar-Ilan University - Ramat-Gan 52900, Israel*

Abstract – We study the effects of spatial constraints on the structural properties of networks embedded in one- or two-dimensional space. When nodes are embedded in space, they have a well-defined Euclidean distance r between any pair. We assume that nodes at distance r have a link with probability $p(r) \sim r^{-\delta}$. We study the mean topological distance l and the clustering coefficient C of these networks and find that they both exhibit phase transitions for some critical value of the control parameter δ depending on the dimensionality d of the embedding space. We have identified three regimes. When $\delta < d$, the networks are not affected at all by the spatial constraints. They are "small-worlds" $l \sim \log N$ with zero clustering at the thermodynamic limit. In the intermediate regime $d < \delta < 2d$, the networks are affected by the space and the distance increases and becomes a power of $\log N$, and have non-zero clustering. When $\delta > 2d$ the networks are "large" worlds $l \sim N^{1/d}$ with high clustering. Our results indicate that spatial constrains have a significant impact on the network properties, a fact that should be taken into account when modeling complex networks.

Introduction. – The interdisciplinary field of complex networks has recently received considerable attention [1–11]. A network is a set of nodes, representing the elementary units of a system. These nodes are connected by edges (bonds) when an interaction between the nodes is present. In a social network, for example, the nodes denote individuals. An edge connects two nodes if there is a relationship between the two individuals. Networks are particularly interesting as theoretical models of complex systems since they provide a straightforward way to describe all sorts of interactions between the elementary units of a system.

Although there are infinitely many types of networks the attention of scientific community is actually focused in four types: regular networks, random graphs, "small-world" networks and scale-free networks.

Regular networks consist of nodes that are identical to each other. The number of edges k emanating from each node (the "degree" of the node) is a constant. Some of the most common constructions used in computational and theoretical physics, such as the square lattice, the Bethe lattice or the complete graph (a graph where each node is connected to every other node) belong to this type of networks.

Random graphs were introduced and studied by Erdös and Rényi [12]. In Erdös-Rényi networks the distribution of the degrees of the nodes is Poissonian, and this fact facilitates the study of several important properties and allows analytical results. These networks have been extensively studied and used as some of the first models of social networks.

"Small-world" networks were initially proposed by Watts and Strogatz as an improved model of social networks [6,7]. These networks show simultaneously the important feature of "clustering" (*i.e.* nodes that have common neighbors have a higher probability of being themselves connected) and the small-world phenomenon meaning that their topological diameter is slowly (logarithmically) increasing with increased system size.

In scale-free networks [1–3] finally, the degree distribution follows a power law $P(k) \sim k^{-\gamma}$, with γ typically between two and three. Scale-free networks have the ultra-small world property as the diameter of the network scales only logarithmically with the logarithm of the number of nodes [9].

In all the above cases, the spatial arrangement of nodes is considered of no importance. It is assumed that

---

(a)E-mail: Kosmas.Kosmidis@theo.physik.uni-giessen.de

spatial constraints can be neglected such that mean-field approaches become applicable. Many real networks are, however, embedded in 2D or 3D space. Examples are the Internet, airline networks, social networks, transportation and wireless communication networks [3,13] etc. If these spatial constraints are important, mean-field theory is no longer valid [14], and new approaches have to be developed instead. Spatially constrained networks are a promising type of complex networks whose importance has been only recently realized [15–18]. Up to now, most efforts deal with the simplest case, *i.e.* how to embed in the normal Euclidean space complex networks with a given degree distribution, while allowing only closest neighbor connections. If one allows additional long-range connections then the resulting networks are more difficult to study but much closer to the real complex networks.

Here we study spatially constrained networks embedded in one- and two-dimensional space. We are interested in the simple case of networks that have a Poissonian degree distribution and where nodes are connected to each other with a probability $p(r) \sim r^{-\delta}$, *i.e.* depending on the Euclidean distance r between the nodes. The choice of a power law for the distance distribution is supported from recent findings on the topology of real complex networks, such as the Internet [19], human travels [20] and in social networks [21].

We propose a simple algorithm to generate such networks and we study properties such as the topological distance l between nodes and the clustering coefficient C. The exponent δ has the role of a control parameter of the properties of the embedded networks and we find that there are three regimes of interest with qualitatively different behavior. Networks with $\delta < d$, where d is the dimensionality of the embedding space, behave more or less like random graphs, networks with $\delta > 2d$ behave much like regular lattices while networks with δ between these d and $2d$ show intermediate power law behavior with critical exponents that depend on δ.

Conceptually, our model is related to the "small-world" networks [6,7,22] and to long-range percolation [23–25]. In both the above cases, an underlying lattice structure is assumed and is "perturbed" by the addition of a fraction of long-range connections. Thus, there is a competition between short- and long-range interactions which determines the properties of the network. In our model, no such underlying structure exists. A perfectly ordered lattice is recovered only in the limit of infinite δ. There is, however, still a "hidden" competition between short and long range, but now the magnitude of both "short" and "long" is controlled by the exponent δ and the system size.

Methods. – To construct the networks, we first arrange the nodes in a one-dimensional $(d=1)$ linear chain or in a two-dimensional $(d=2)$ regular square lattice. Thus, between any pair of nodes there is a well-defined Euclidean distance. Next, we initially choose a node i randomly and assume that it is connected to k_t

other nodes. For each of these nodes, we first choose its distance r from node i with probability $P(r) = cr^{d-1}r^{-\delta}$, where c is determined from the normalization condition $\int_1^L P(r)\mathrm{d}r = 1$, with $L = N^{1/d}$. Then we pick one of the N_r sites that are at distance r within the underlying lattice. We repeat this process for all nodes i in the underlying lattice and then remove multiple connections. The nodes of the final network we obtain this way do not have a fixed degree. Their mean degree $\bar{k}$ depends on both k_t and δ. In other words, for obtaining a certain $\bar{k}$ value for fixed δ, we have to choose an appropriate value for k_t. When the networks are constructed we measure their actual degree probability distribution $p(k)$ to verify that all networks have at least approximately the same mean degree $\bar{k}$ and the actual distance distribution $p(r)$ *i.e.* the fraction of connected nodes that are at distance r from each other, which is affected by the process of the removal of duplicate edges. In what follows we have succeeded in obtaining $\bar{k} \approx 4$ (to be precise $3.3 < \bar{k} \leqslant 4$) using $k_t = 2$ when $\delta \leqslant 2.25$, $k_t = 3$ when $2.25 < \delta \leqslant 3$ and $k_t = 5$ for $\delta = 4$ for the one-dimensional case. For the two-dimensional case $k_t = 2$ produced networks with $3.6 < \bar{k} \leqslant 4$ for all $\delta \leqslant 4.5$. Thus, no particular "fine-tuning" was necessary.

Quantities of interest. – We are interested in the mean topological distance l between the nodes of the network and the mean clustering coefficient C. The topological distance between two nodes a and b is the minimum number of links that a walker has to cross in order to arrive from a to b. The (local) clustering coefficient of a node i in a network is defined as

$$C(k_i) = \frac{t_i}{[k_i(k_i-1)/2]}, \tag{1}$$

where t_i is the number of triangles (loops of length 3) attached to this node divided by the maximum possible number of such loops. For calculating the mean topological distance l and the mean clustering coefficient C we average over all nodes in the network (typically 10^5) and over all network (typically 10^2) realizations.

Results. – First, we show that the generated networks have the desired degree and distance distributions. Figure 1 shows the degree distribution $P(k)$ and the distance distribution $P(r)$ for the generated one-dimensional (fig. 1a, c) and two dimensional (fig. 1b, d) networks, for three characteristic δ values each, $\delta = 0.5$, 1.5, 2.5 $(d=1)$ and $\delta = 2$, 3.5, 5.5 $(d=2)$. Figures 1a, b confirm that the networks have a narrow degree distribution around $\bar{k} \approx 4$. Figures 1c, d confirm that $P(r)$ has the desired distance distribution $P(r) \sim r^{-(\delta-d+1)}$ for both $d=1$ and $d=2$.

Next, we study the mean topological distance l as a function of the network size N. In the limiting case $\delta \to 0$ the spatial constraints are not significant and we expect that the networks will belong to the universality class of Erdös-Rényi random graphs, where l scales as $l \sim \log N$.
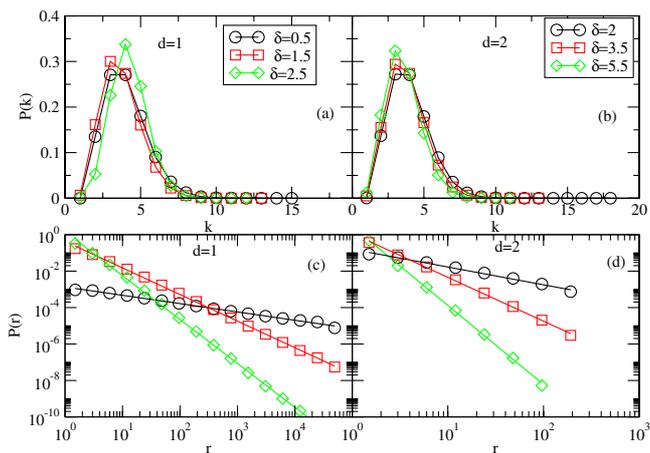
Fig. 1: (a) Probability $P(k)$ that a node has degree k for $d = 1$ networks with $N = 10^5$, $\bar{k} \approx 4$ and for $\delta = 0.5, 1.5$ and 2.5. (b) Probability $P(k)$ for $d = 2$ networks with $N = 5 \times 10^4$, $\bar{k} \approx 4$ and for $\delta = 2, 3.5, 5.5$. (c) Probability $P(r)$ that a node has a connection at distance r for the same $d = 1$ networks as in (a). The points are logarithmically binned simulation data. (d) Probability $P(r)$ that a node has a connection at distance r for the same $d = 2$ networks as in (b). The points are logarithmically binned simulation data. In (c,d) the straight lines have slopes equal to $-(\delta - d + 1)$.

In the opposite case $\delta \to \infty$, the spatial constraints will result in an ordered lattice-like structure that belongs to the universality class of regular lattices, where the mean topological distance l scales as $N^{1/d}$ (d is the dimensionality of the embedding space). We are interested in how the dependence of the topological distance l on N changes when δ changes from zero to infinity.

Our results (figs. 2 and 3) indicate that in both $d = 1$ and $d = 2$, there are three regimes separated by $\delta = d$ and $\delta = 2d$, where the behavior of the networks is qualitatively different. In the first regime $0 \leqslant \delta < d$ (see fig. 2a for $d = 1$ and fig. 2b for $d = 2$), l scales as $\log N$ independent of δ. In the third regime, $\delta > 2d$ (see fig. 2c for $d = 1$ and fig. 2d for $d = 2$) the result suggests that l asymptotically scales as $N^{1/d}$ as for regular lattices. This dependence is well observed in $d = 1$ (fig. 2c) for $\delta = 4$ and in $d = 2$ for $\delta = 5.5$. For δ closer to $2d$, this scaling is only recovered for sufficiently large networks. For $\delta = 2d$, l scales as N^γ, with $\gamma \approx 0.35$ for $d = 1$ and $\gamma \approx 0.3$ for $d = 2$ (see the lowest curves in figs. 2c and d, respectively).

Figure 3a, b shows the mean topological distance l as a function of N in the intermediate regime $d < \delta < 2d$. The plots suggest that for both $d = 1$ and $d = 2$, l scales as $l \sim (\ln(N))^\alpha$. Accordingly, we can summarize the dependence of l on N as follows:

$$l \sim \begin{cases} \log(N), & \delta \leqslant d, \\ (\log(N))^\alpha, & d \leqslant \delta < 2d, \\ N^{1/d}, & \delta > 2d. \end{cases} \quad (2)$$

When $\delta = 2d$ the numerical results are quite well described by a power law N^γ where $\gamma \approx 0.35$ for $d = 1$ and $\gamma \approx 0.3$ for
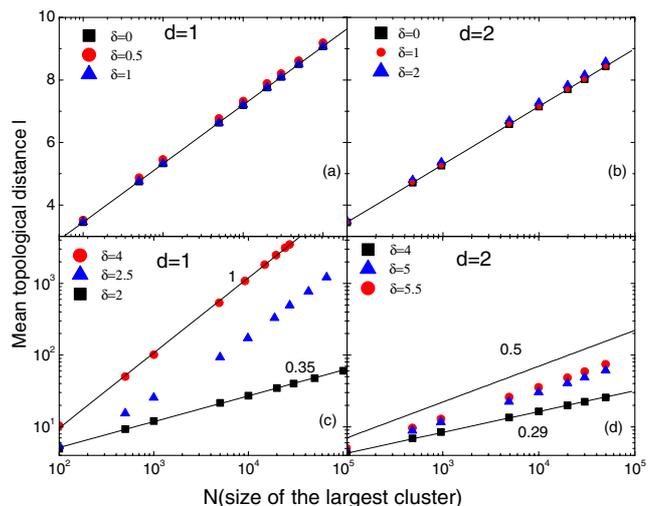

Fig. 2: (a) Mean topological distance l as a function of the networks size N for $d = 1$ networks with $\bar{k} \approx 4$ and for $\delta = 0, 0.5$ and 1, *i.e.* in Regime 1 (LinearLog plot). (b) Mean topological distance l as a function of the networks size N for $d = 2$ networks with $\bar{k} \approx 4$ and for $\delta = 0, 1$ and 2, *i.e.* in Regime 1 (LinearLog plot). The straight line is guide to the eye. (c) Mean topological distance l as a function of the networks size N for $d = 1$ networks with $\bar{k} \approx 4$ and for $\delta = 2, 2.50$ and 4, *i.e.* in Regime 3 (LogLog plot). (d) Mean topological distance l as a function of the networks size N for $d = 2$ networks with $\bar{k} \approx 4$ and for $\delta = 4, 5, 5.5$, *i.e.* in Regime 3 (LogLog plot). The straight line has slope 0.5.

$d = 2$. It is, however, probable that there are logarithmic corrections to this scaling form. The variation of the exponent α with δ is shown in fig. 3c for $d = 1$ and fig. 3d for $d = 2$. The results suggest that approximately

$$\alpha = \begin{cases} \dfrac{1/\delta}{2 - \delta}, & d = 1, \\[2mm] \dfrac{4/\delta}{4 - \delta}, & d = 2. \end{cases} \quad (3)$$

Equation (3) provides a good fit to the results in figs. 3c, d, and although is not a rigorous analytic result it may be useful for practical purposes[1]. Note that eq. (3) ensures a smooth transition between the first and second regime. To match with the power law behavior when $\delta > 2d$, α diverges when approaching $\delta = 2d$ from below. In $d = 2$, also a weaker singularity could fit the data.

We like to note that these transition points arise naturally in the scaling of the mean distance $\bar{r}$ and the maximum distance r_{max} with the system size L. By definition $\bar{r} = \int_1^L r p(r) \mathrm{d}r$, while r_{max} is obtained from the condition $L^d \int_{r_{max}}^L p(r) \mathrm{d}r \approx 1$. It is easy to verify that

$$\bar{r} \sim \begin{cases} L, & \delta \leqslant d, \\ L^{d - \delta + 1} + c_1, & \delta > d, \end{cases} \quad (4)$$

---

[1]Of course, in eq. (2) in the range $d \leqslant \delta < 2d$ a more complex scaling form cannot be excluded.
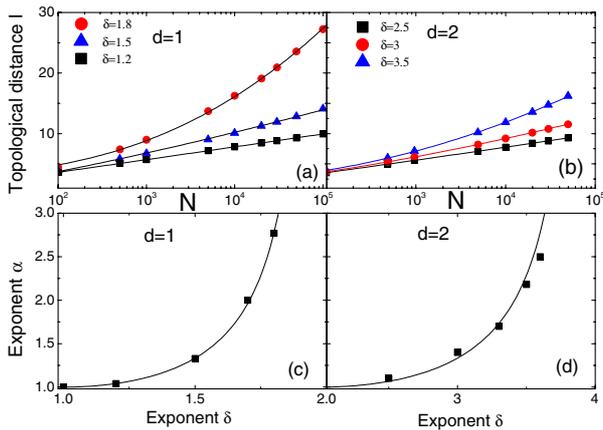
Fig. 3: (a) Mean topological distance l as a function of the networks size N for $d = 1$ networks with $\bar{k} \approx 4$ and for $\delta = 1.2, 1.5$ and 1.8 (bottom to top), *i.e.* in Regime 2 (LinearLog plot). The lines are fits to the simulation data of the scaling assumption (eq. (2)). (b) Mean topological distance l as a function of the networks size N for $d = 2$ networks, $\bar{k} \approx 4$ and for $\delta = 2.5, 3, 3.5$ (bottom to top), *i.e.* in Regime 2 (LinearLog plot). Points are simulation data and lines are fits of eq. (2). (c) Exponent α as a function of δ for $d = 1$ networks in the second regime ($1 < \delta < 2$). (d) Exponent α as a function of δ for $d = 2$ networks in the second regime ($2 < \delta < 4$). Points are best fit estimates to the simulation data. Lines are plots of eq. (3).

where c_1 does not depend on N and

$$
r_{max} \sim \begin{cases} L, & 0 < \delta \leqslant 2d, \\ L^{\frac{d}{\delta-d}}, & \delta \geqslant 2d. \end{cases} \tag{5}
$$

Accordingly, the relative mean distance $\bar{r}/L$ is finite for $\delta \leqslant d$ and 0 for $\delta > d$ in the limit of large L. Similarly, the relative maximum distance r_{max}/L is finite for $\delta \leqslant 2d$ and 0 for $\delta > 2d$ for L going to infinity. The above relations can qualitatively elucidate the existence of different regimes. For simplicity consider the case $d = 1$ and assume one attempts to go from the one end of the "chain" to the other. When the mean length $\bar{r}$ of each link is comparable to the lattice size L, then he can succeed with only one step on average. The lattice is a small world (regime 1). When the the mean length $\bar{r}$ is less than L but r_{max} is of the order of L, then the walker will have to do more steps but he will eventually arrive at a site that has a long edge r_{max} that will allow him to arrive at the end of the lattice. The lattice is not a small world as in the first case but also not a regular lattice since there are links that decrease the topological distance considerably (regime 2). When both $\bar{r}$ and r_{max} are less than L, there are no shortcuts. Renormalizing the graph will result in a regular linear chain (regime 3).

Finally, we study how the clustering coefficient C depends on δ and N (fig. 4). We are particularly interested in the behavior at the transition points $\delta = d$ and $\delta = 2d$. We expect that for $\delta \to 0$, $C \sim 1/N$ similar to the case of Erdös-Rényi random graphs. When $\delta \to \infty$ we expect the
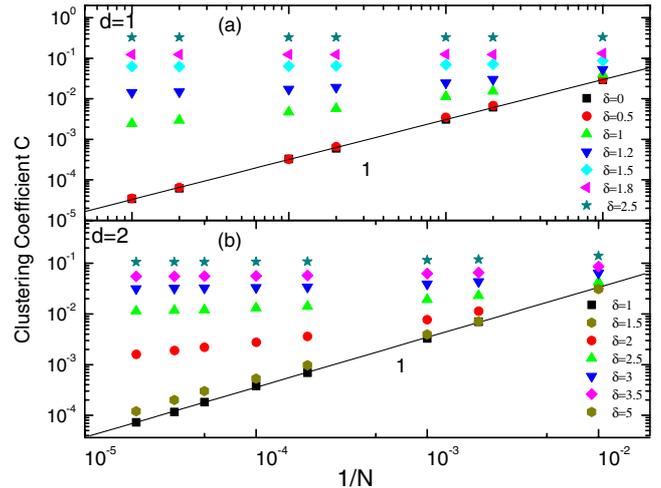

Fig. 4: (a) Clustering coefficient C as a function of $1/N$ for $\delta = 0, 0.5, 1, 1.2, 15, 1.8, 2.5$ and $d = 1$. (b) Clustering coefficient C as a function of $1/N$ for $\delta = 1, 1.5, 2, 2.5, 3, 3.5, 5$ and $d = 2$. Straight lines are best fits with slope 1.

(high) clustering of the lattices. We find that the critical point at $\delta = d$ separates a phase of zero clustering when $\delta < d$ from a phase of finite clustering when $\delta > d$. The second transition point at $\delta = 2d$ does not influence C remarkably. Figures 4a, b show C as a function of $1/N$ for $d = 1$ and $d = 2$, respectively. The clustering coefficient C decreases as $1/N$ for $\delta < d$ in both one and two dimensions, while for $\delta > d$ the slope of the curves in the double logarithmic plot becomes zero, confirming the existence of a regime with non-zero clustering. It is remarkable that as soon as δ becomes larger than d, the spatial effects become very pronounced. The topology not only affects the mean topological distance but also increases simultaneously the clustering coefficient.

Conclusions. – We have examined the mean topological distance l and the clustering coefficient C of networks that are embedded in one- or two-dimensional space. Both exhibit phase transitions for some critical value of the control parameter δ. We have identified three regimes. When $\delta < d$, where d is the dimensionality of the embedding space, the networks are "small-worlds" with zero clustering at the thermodynamic limit. In the intermediate regime $d < \delta < 2d$, the topology changes. The topological distance becomes a power of the logarithm of the system size and the networks have non-zero clustering. When $\delta > 2d$ the networks do not have the small-world property and have high clustering. Our results indicate that spatial constraints have an important influence on the network properties and this fact should be taken into account when modeling real complex networks. A natural extension of this work will be to consider fractal or multifractal substrates instead of regular lattices [13,26] and study their influence on the topological properties of the embeded networks.

REFERENCES

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

[2] Albert R., Jeong H. and Barabási A. L., *Nature (London)*, **401** (1999) 130.

[3] Mendes J. F. F., Dorogovtsev S. N. and Ioffe A. F., *Evolution of Networks: From Biological Nets to the Internet and WWW* (Oxford University Press, Oxford) 2003.

[4] Pastor-Sattoras R. and Vespignani A., *Evolution and Structure of the Internet: A Statistical Physics Approach* (Cambridge University Press, Cambridge) 2004.

[5] Newman M. E. J., Watts D. J. and Strogatz S. H., *Proc. Natl. Acad. Sci. U.S.A.*, **99** (2002) 2566.

[6] Watts D. J. and Strogatz S. H., *Nature (London)*, **393** (1998) 440.

[7] Watts D. J., *Small Worlds* (Princeton University Press, Princeton, NJ) 1999.

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

[9] Cohen R., Havlin S. and ben-Avraham D., *Handbook of Graphs and Networks*, edited by Bornholdt S. and Schuster H. G. (Wiley-VCH) 2002, Chapt. 4.

[10] Cohen R. and Havlin S., *Phys. Rev. Lett.*, **90** (2003) 058701.

[11] Gallos L. K. *et al.*, *Phys. Rev. Lett.*, **94** (2005) 188701.

[12] Erdös P. and Rényi A., *Publ. Math.*, **6** (1959) 290.

[13] Glauche I. *et al.*, *Physica A*, **325** (2003) 577.

[14] Stanley H. E., *Introduction to Phase Transitions and Critical Phenomena* (Oxford University Press, New York) 1987.

[15] Warren C. P., Sander L. M. and Sokolov I. M., *Phys. Rev. E*, **66** (2002) 056105.

[16] Rozenfeld A. *et al.*, *Phys. Rev. Lett.*, **89** (2002) 218701.

[17] Soares D. J. B., Tsallis C., Mariz A. M. and da Silva L. R., *Europhys. Lett.*, **70** (2005) 70.

[18] Xu X., Zhang X. and Mendes J. F. F., arXiv:0710. 3907v1 [physics.soc-ph] (2007).

[19] Yook S.-H., Jeong H. and Barabási A. L., *Proc. Natl. Acad. Sci. U.S.A.*, **99** (2002) 13382.

[20] Brockmann D. *et al.*, *Nature*, **439** (2006) 462.

[21] Liljeros F., private communication (2006).

[22] Kleinberg J. M., *Nature*, **406** (2000) 845.

[23] Moukarzel C. F. and Argollo de Menezes M., *Phys. Rev. E*, **65** (2002) 056709.

[24] Benjamini I. and Berger N., *Random Struct. Algorithms*, **19** (2001) 102.

[25] Coppersmith D., Gamarnik D. and Sviridenko M., *Random Struct. Algorithms*, **21** (2002) 1.

[26] Roberson M. R. and ben-Avraham D., *Phys. Rev. E*, **74** (2006) 017101.