Correlation effects in a simple model of a small-world network

V. Karimipour* and A. Ramzanpour†
Department of Physics, Sharif University of Technology, P.O. Box 11365-9161, Tehran, Iran
(Received 24 September 2001; published 21 February 2002)

We analyze the effect of correlations in a simple model of a small-world network by obtaining exact analytical expressions for the distribution of shortest paths in the network. We enter correlations into a simple model with a distinguished site, by taking the random connections to this site from an Ising distribution. Our method shows how the transfer-matrix technique can be used in the new context of small-world networks.

DOI: 10.1103/PhysRevE.65.036122 PACS number(s): 05.40.−a, 05.20.−y, 89.75.Hc

I. INTRODUCTION

Real networks, such as social networks, neural networks, power grids, and documents in the World Wide Web [2–5], can be modeled neither by totally random networks nor by regular ones (see [4,6,7], and references therein for review). While locally they are clustered as in regular networks, remote sites have often the chance of being connected via shortcuts, as in random graphs, hence reducing the average distance between sites in the network.

In a regular networks with N vertices, the average shortest path between two vertices $\langle l\rangle$ and the clustering coefficient C scale, respectively, as $\langle l\rangle \sim N$ and $C \sim 1$. The clustering coefficient C is defined as the average ratio of the number of existing connections between neighbors of a vertex to the total possible connections among them. In random networks, however we have $\langle l\rangle \sim \ln N$ and $C \sim 1/N$ [8,9].

The properties of many real networks are a hybrid of these two extremes, that is, in these networks one has $\langle l\rangle \sim \ln N$ and $C \sim 1$. These two effects, collectively called the small-world effect, are attributed, respectively, to the presence of shortcuts and the many interconnections that usually exist between the neighboring nodes of such networks [10–12].

In 1998, Watts and Strogatz [13] introduced a simple model of networks showing the small-world behavior, which since then has been investigated as a model of interconnections in many different contexts, ranging from epidemiology [14–16], to polymer physics [17–19], and evolution and navigation [20–22]. The original model of Watts and Strogatz contained a free parameter p, varying which one could interpolate between random and irregular networks. Their model, called hereafter the WS model, consists of a ring of N sites in which each site is connected to its $2k$ nearest neighbors, hence making a regular network. After this stage, each bond is rewired with probability p to another randomly chosen site. The value of p tunes the amount of randomness introduced into the network. Since there is a finite probability of disconnecting the whole network in this way, Newman and Watts [23] modified the model by replacing the rewiring stage by just addition of shortcuts between randomly chosen sites on the ring. Since then many more variants and generalizations of small-world networks and their different characteristics (e.g., their topology, the properties of random walks on them, etc.) have been studied. Of particular interest are three classes of studies. The first class, in which the static properties of small-world networks have been investigated [24–27], the second class, where dynamical aspects have been studied [28,29] and the third class, in which evolving networks are considered [30,31] in order to generate small-world networks with various connectivity distributions, including scale-free distributions.

In this paper we want to consider another variant of the small-world network, one in which correlation of neighboring nodes in making connections to remote sites is taken into account (i.e., the presence of a shortcut between two sites affects other shortcuts in the neighborhood). For example, a node need not make a shortcut to a remote site if there is such a connection in its neighborhood. In such networks then, correlations play an important role. However to perform such a study by exact non-mean-field methods requires a simplification in the original model. We assume that all shortcuts are made via a distinguished site at the center of the ring. More than being a simplification, this type of network has practical relevance in many situations where a central distinguished site governs all the remote interconnections. We note in passing that such central sites accommodating a large number of connections, may exist either in the architecture of the original networks or else may appear dynamically in evolving networks [32]. In this way we assume that contrary to the original model [1], the two configurations in Fig. 1, both with five shortcuts are not equiprobable.

II. THE MODEL AND SMALL-WORLD QUANTITIES

We consider a circular network of N vertices with a distinguished central site, Fig. 2. The links on the ring have unit length. Each shortcut connecting any two sites on the ring is also of unit length. We assign a random variable $s_i \in \{0,1\}$ to each site i of the ring. This random variable is 1 or 0 according to whether the site is connected to the center or not, Fig. 2. Any configuration of these spin variables corresponds to one and only one configuration of connections to the center. For example in Fig. 1 if each bond is independently connected to the center with probability p, then the probabilities of both configurations are equal and proportional to $p^5(1-p)^{11}$. In general, and in the absence of correlations we will have

*Email address: vahid@sharif.edu
†Email address: ramzanpour@mehr.sharif.ac.ir

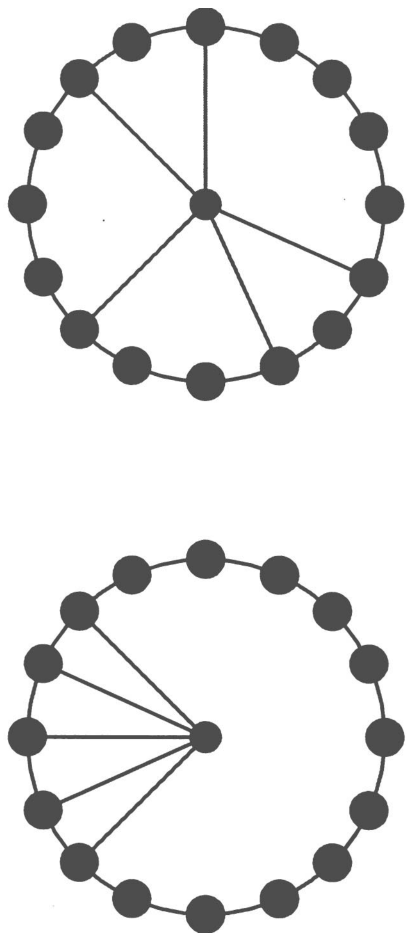

FIG. 1. Two configurations that have different weights in our calculations but equal weights in [1].

$$P=\frac{1}{Z}\left(\frac{p}{1-p}\right)^{s_1+s_2+\cdots+s_N}, \tag{1}$$

where Z is a normalization constant. To consider correlations we generalize the above distribution to an Ising-type distribution, namely, to

$$P\{s_i\}=\frac{1}{Z}\left(\prod_{i=1}^{N} r^{s_i}\zeta^{s_i s_{i+1}}\right). \tag{2}$$

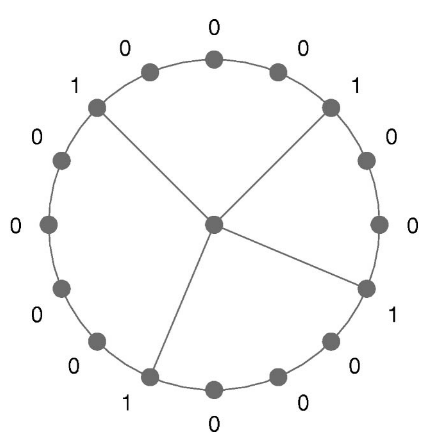

FIG. 2. A simple model of small-world networks.

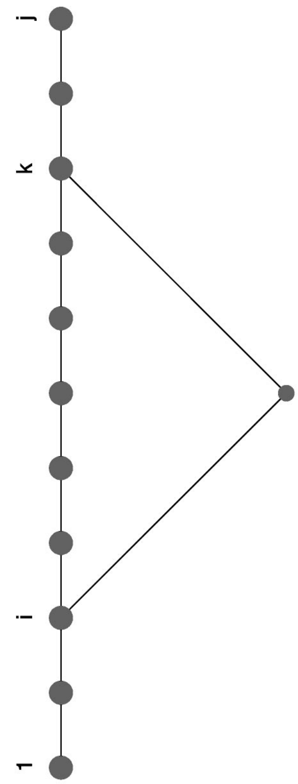

FIG. 3. A typical configuration that decreases l_{1j}.

For $\zeta=1$ and $r=p/(1-p)$ we obtain the original model of [1]. The value of ζ controls the correlations.

First let us consider the directed model, i.e., the links on the circle are directed, say, clockwise. We consider a typical configuration such as the one shown in Fig. 3, in which the nearest shortcuts to sites 1 and j are connected at sites i and k. This configuration reduces the distance between sites 1 and j by an amount $k-i-1$. Note that the sites between i and k may or may not be connected to the center. In any such configuration the quantity $X_{i,k}(1,j)$ defined as

$$(1-s_1)\cdots(1-s_{i-1})s_i s_k(1-s_{k+1})\cdots(1-s_j) \tag{3}$$

takes the value 1. The average of this quantity gives the probability of such a configuration. In order to find the probability of the shortest path between sites 1 and j to be equal to l, we have to sum over all those configurations that give such a shortest path. For $l\neq j-1$ the above probability is given by

$$p(1,j;l)=\sum_{i=1}^{l}\langle X_{i,j+i-l}(1,j)\rangle, \tag{4}$$

where we have used $\langle\cdots\rangle$ for averaging over configurations.

Normalization determines $P(1,j;j-1)$ via

$$p(1,j;j-1)=1-\sum_{l=1}^{j-2}p(1,j;l). \tag{5}$$

The probability that the shortest path between two arbitrary vertices be of length l is obtained from

$$p(l)=1/N\sum_{j=l+1}^{N}p(1,j;l). \tag{6}$$

Now the average shortest path between two randomly chosen sites is

$$\langle l\rangle=\sum_{l=1}^{N-1}lp(l). \tag{7}$$

All the above quantities can be calculated by the transfer-matrix method, in which we write the unnormalized distribution (2) as a product of matrix elements of a matrix T,

$$T=\begin{pmatrix}1 & \sqrt{r}\\ \sqrt{r} & r\zeta\end{pmatrix}, \tag{8}$$

with eigenvalues

$$\lambda_{\pm}=\frac{1}{2}[1+r\zeta\pm\sqrt{(1-r\zeta)^2+4r}]. \tag{9}$$

The partition function is $Z=\lambda_+^N+\lambda_-^N$ and the number of connections per site is given by

$$p:=\frac{r}{N}\frac{\partial}{\partial r}\ln Z. \tag{10}$$

We now consider the continuum limit of the lattice, where the number of vertices goes to infinity and the lattice constant goes to zero as $1/N$ so that the periphery of the lattice is kept constant at 1. We then set $j/N\to x, k/N\to s, i/N\to t$, $l/N\to z$, and $NX_{i,k}(1,j)\to X(t,s)(x)$, where the explicit form of the function $X(t,s)(x)$ will be determined later. We will then have $0\leqslant x,t,s,z\leqslant 1$. Here x is the distance along the ring.

Furthermore, we take $Np(1,j,l)\to\mathcal{Q}(x,z)$, therefore

$$\mathcal{Q}(x,z)=\int_0^z X(t,x+t-z)(x)dt,$$

$$\mathcal{Q}(z)=\int_z^1\mathcal{Q}(x;z)dx, \tag{11}$$

where $\mathcal{Q}(x,z)dz$ is the probability that two points whose distance along the ring is x have a shortest distance between z and $z+dz$. Then $\mathcal{Q}(z)dz$ is the probability that the shortest path between any two points be between z and $z+dz$. So $\int_0^1 dzQ(z)=1$ and finally

$$\langle z\rangle=\int_0^1 z\mathcal{Q}(z)dz. \tag{12}$$

The scaling limit

Intuitively we expect that in the scaling limit, when $N\to\infty$ and $r=r_0/N$, if we keep ζ finite, then the number of connections to the center remains finite and in an infinite lattice the configurations of these connections become quite sparse and hence correlations cannot play a role, at least to leading order. Exact calculation also verifies this expectation. Here we will consider a different scaling limit where $N\to\infty$ and $r=r_0/N$ while $\zeta=\zeta_0 N$. This means that the tendency of an individual, one of whose neighbors has been connected to the center, depends also on the total population. This assumption is not far from reality, specially in cases where the center approves a limited amount of connections and the applicants, competing for connections, are aware of this restriction. It turns out that the model shows three distinct behaviors according to the value of the parameter $r_0\zeta_0$.

For $r_0\zeta_0>1$ we will have

$$\lambda_+=r_0\zeta_0+\frac{r_0}{N(r_0\zeta_0-1)},$$

$$\lambda_-=1-\frac{r_0}{N(r_0\zeta_0-1)}, \tag{13}$$

and from (10) we find

$$p=1-O\left(\frac{1}{N}\right), \tag{14}$$

which means that the whole lattice is filled with connections. Also for $r_0\zeta_0=1$ we obtain $p=1/2$, which is also far from the small-world regime. To be in the small-world regime, we should keep $r_0\zeta_0<1$, which is the case that we will study in detail.

In this case we have

$$\lambda_+=1+\frac{r_0}{N(1-r_0\zeta_0)},$$

$$\lambda_-=r_0\zeta_0-\frac{r_0}{N(1-r_0\zeta_0)}. \tag{15}$$

Also from Eq. (10) we find the total number of connections to be the finite value

$$M_0:=Np=\frac{r_0}{(1-r_0\zeta_0)^2}. \tag{16}$$

To calculate $X_{i,k}(1,j)$ we note that since $p\to 0$ as $1/N$, the values of these quantities where either or both of i and k take the extreme values 1 or j are suppressed. Using the transfer-matrix technique, we obtain from Eqs. (2) and (3) that for $1<i<k<j$,

$$\begin{aligned}\langle X_{i,k}(1,j)\rangle&=T_{00}^{i-2}T_{01}(T^{k-i})_{11}T_{10}T_{00}^{j-k-1}(T^{N-j+1})_{00}\\&=r(T^{k-i})_{11}(T^{N-j+1})_{00},\end{aligned} \tag{17}$$

where $T_{ij}^m=\langle i|T|j\rangle^m$ and $(T^m)_{ij}=\langle i|T^m|j\rangle$. Diagonalizing T, using Eqs. (8) and (9), and taking the continuum limit, we find after some algebra,

$$X(s-t)(x)=M^2e^{-Mx}e^{M(s-t)}, \tag{18}$$

where $M := r_0/(1 - r_0\zeta_0)$. Inserting this value in Eq. (11) and integrating we find

$$\mathcal{Q}(x;z) = zM^2 e^{-Mz}, \tag{19}$$

$$\mathcal{Q}(z) = z(1-z)M^2 e^{-Mz} + (1+zM)e^{-Mz}. \tag{20}$$

Turning to Eq. (12) we obtain

$$\langle z \rangle = \frac{1}{M}(2 + e^{-M}) - \frac{3}{M^2}(1 - e^{-M}). \tag{21}$$

As expressed in [1], these relations already hint at the emergence of a type of small-world behavior, i.e., with connecting only ten sites the average shortest path is reduced from $\frac{1}{2}$ to 0.17; connecting ten more sites reduces this value to 0.09.

We see that as far as $\zeta_0 < 1/r_0$, the effect of correlations is only to modify the relations of [1] by replacing M_0, the actual number of connections, with an effective one M. Expressing M in terms of M_0 and ζ_0 alone, we find $M_0 = M(1 + M\zeta_0)$, which means that for low values of ζ_0, $M \sim M_0$ while for large values of ζ_0 the effective number of connections scales as the square root of the actual number of shortcuts, $M \sim \sqrt{M_0/\zeta_0}$. This effect reflects the tendency of the shortcuts to get clustered under the influence of correlations. Hence correlations tend to decrease the small-world effect, since the connections tend to bunch into clusters.

III. UNDIRECTED AND CLUSTERED NETWORKS

As far as we have $N \to \infty$ and $M_0 =$ finite, we can generalize our results to the cases where (a) the network has no preferred direction and (b) each site of the ring is connected to $2k$ of its neighbors. In this limit, in going from one site of the ring to another one, one travels mostly along the ring. Thus denoting the average shortest paths for the above cases, respectively, by $\langle\langle z \rangle\rangle_a$ and $\langle\langle z \rangle\rangle_b$ we have

$$\langle\langle z \rangle\rangle_a = \frac{1}{2}\langle\langle z \rangle\rangle, \quad \langle\langle z \rangle\rangle_b = \frac{1}{k}\langle\langle z \rangle\rangle, \tag{22}$$

from which we obtain

$$\mathcal{Q}_a(z) = 2\,\mathcal{Q}(2z), \quad 0 \leqslant z \leqslant \frac{1}{2} \tag{23}$$

and

$$\mathcal{Q}_b(z) = k\,\mathcal{Q}(kz), \quad 0 \leqslant z \leqslant \frac{1}{k}. \tag{24}$$

And finally, for the clustered undirected model, one will have

$$\mathcal{Q}_{ab}(z) = 2k\,\mathcal{Q}(2kz), \quad 0 \leqslant z \leqslant \frac{1}{2k}. \tag{25}$$

IV. CONCLUSION

We have considered the effect of correlations in a simple model of a small-world network, and shown that they generally decrease the small-world effect, since under this condition the connections tend to bunch into clusters. More concretely, in our simple model the effect of correlations, which are controlled by a parameter ζ_0, is to reduce (for large ζ_0) the actual number of shortcuts M_0 to an effective one $M \sim \sqrt{M_0/\zeta_0}$, indicating a clustering of connections to bunches in the lattice.

Therefore it seems that the optimal way of designing a small-world network would be with equidistant long-range connections, and in order to see the small-world effect and lower the average shortest path, one would rather use algorithms that anticorrelate the connections. We have derived our results by exact analytical methods and have shown how the transfer-matrix technique can be used for obtaining such properties as average shortest path or the distribution of shortest paths in a model of a small-world network. For all this we have been forced to study a restricted class of models. No doubt by doing computer simulations, one can study these effects in a much broader class of models.

[1] S.N. Dorogovtsev and J.F.F. Mendes, Europhys. Lett. **50**, 1 (2000).

[2] S. Milgram, Psychology Today **2**, 60 (1967).

[3] L.F.L. Fernandez, R. Huerta, F. Corbacho, and J.A. Siguenza, Phys. Rev. Lett. **84**, 2758 (2000).

[4] M.E.J. Newman, J. Stat. Phys. **101**, 819 (2000).

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

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

[7] R. Albert and A.L. Barabasi, e-print cond-mat/0106096.

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

[9] B. Bollaoas, *Random Graphs* (Academic Press, New York, 1985).

[10] R. Kasturirangan, e-print cond-mat/9904055.

[11] N. Mathias and V. Gopal, Phys. Rev. E **63**, 021117 (2001).

[12] V. Latora and M. Marchiori, Phys. Rev. Lett. (to be published).

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

[14] M. Kuperman and G. Abramson, Phys. Rev. Lett. **86**, 2909 (2001).

[15] D.H. Zanette, e-print cond-mat/0105596.

[16] S.C. Manrubia, Jordi Delgado, and Bartolo Luque, Europhys. Lett. **53**, 693 (2001).

[17] P. Sen and B.K. Chakrabarti, e-print cond-mat/0105346.

[18] A. Scala, L.A.N. Amaral, and M. Barthelemy, Europhys. Lett. **55**, 594 (2001).

[19] S. Jespersen and A. Blumen, Phys. Rev. E **62**, 6270 (2000).

[20] F. Bagnoli and M. Bezzi, Phys. Rev. E **64**, 021914 (2001).

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

[22] A.R. Puniyani, R.M. Lukose, and B.A. Huberman, e-print cond-mat/0107212.
[23] M.E.J. Newman and D.J. Watts, Phys. Rev. E **60**, 7332 (1999).
[24] A. Barrat and M. Weigt Eur. Phys. J. B **13**, 547 (2000).
[25] C. Moore and M.E.J. Newman, Phys. Rev. E **62**, 7059 (2000).
[26] R.V. Kulkarni, E. Almaas, and D. Stroud, Phys. Rev. E **61**, 4268 (2000).
[27] M. Kuperman and G. Abramson, Phys. Rev. E **64**, 047103 (2001).
[28] R. Monasson, Eur. Phys. J. B **12**, 525 (1999).
[29] S. Jespersen, I.M. Sokolov, and A. Blumen, Phys. Rev. E **62**, 4405 (2000).
[30] J. Davidsen, H. Ebel, and S. Bornholdt, e-print cond-mat/0108302.
[31] K. Klemm and V.M. Eguiluz, e-print cond-mat/0107607.
[32] G. Bianconi and A.L. Barabasi, Phys. Rev. Lett. **86**, 5632 (2001).