Simple models of small-world networks with directed links

A. Ramezanpour* and V. Karimipour†
Department of Physics, Sharif University of Technology, P.O. Box 11365-9161, Tehran, Iran
(Received 12 May 2002; published 24 September 2002)

We investigate the effect of directed short- and long-range connections in a simple model of a small-world network. Our model is one in which we can determine many quantities of interest by an exact analytical method. We calculate the function $V(T)$, defined as the number of sites affected up to time T when a naive spreading process starts in the network. As opposed to shortcuts, the presence of unfavorable bonds has a negative effect on this quantity. Hence, the spreading process may not be able to affect all of the network. We define and calculate a quantity identified as the average size of the accessible world in our model. The interplay of shortcuts and unfavorable bonds on the small world properties is studied.

DOI: 10.1103/PhysRevE.66.036128 PACS number(s): 89.75.Hc, 05.40.−a , 05.20.−y

I. INTRODUCTION

Real life networks, whether made by nature (e.g., neural, metabolic, and ecological networks) [1–5], or by humans (e.g., the World Wide Web, power grids, transport networks, and social networks of relations between individuals or institutes) [6–8], have special features that are a blend of those of regular networks, on the one hand, and completely random ones, on the other hand. To study any process in these networks (the spreading of an epidemic in human society, a virus in the internet, or an electrical power failure in a large city, to name only a few), an understanding of their topological and connectivity properties is essential (for a review, see [9] and references therein). Recently obtained data from many real networks show that like random networks [10,11], they have a small diameter, and like regular networks, they have high clustering. Since the pioneering work of Watts and Strogatz [12], these networks have attracted a lot of attention and have been studied from various directions [13–17].

In contrast to most of the models studied so far, many real networks such as the World Wide Web, neural networks, power grids, and metabolic, and ecological networks have directed one-way links [3,18–20]. These types of networks may have significant differences in both their static and dynamic properties with the Watts-Strogatz (WS) model and its variations [19,21,22]. The presence of directed links strongly affects many of the properties of a network. For example, for the same pattern of shortcuts, the average shortest path in a directed network is longer than that in an undirected one, due to the presence of bonds with the wrong directions (blocks) in many paths. So is the spreading time of any dynamic effect on the lattice.

Consider the quantity $V(T)$, defined as the average number of sites that are visited at least once when we start a naive spreading process at a site and continue it for T steps. Note that we mean an average over an ensemble of networks and initially infected sites, and by the naive spreading process we mean that at each step of the spreading process all the neighbors of an infected site are equally infected. The quantity $V(T)$ may be taken as a crude approximation of the number of people who have been infected by a contiguous disease after T time steps has elapsed since the first person has been infected. Clearly this is a simplification of the real phenomena, since in the real world a disease may not affect an immunized person or may not transmit with certainty during contact. However, as a first approximation, $V(T)$ may give a sensible measure of the effect in the whole network. Since in a directed network an effect only spreads to those neighbors into which there are correctly directed links, there will be pronounced differences in this important quantity between a directed and an undirected network. As a concrete example, consider a ring with N sites without any shortcuts, where, to emphasize the absence of shortcuts, we denote $V(T)$ by $V_0(T)$. If all the links have the same direction, we have $V_0(T)=T$, and if all of them are bidirectional, we have $V_0(T)=2T$. In both cases the whole lattice is infected after a finite time. However, if the links are randomly directed then $V_0(T)$ may be much lower and, furthermore, there is a finite probability that only a small fraction of the whole lattice becomes infected.

Adding shortcuts to this ring of course has a positive effect on the spreading. In a sense, we have a chance to see the interplay of two different concepts of small worlds in these networks. The size of the world as a whole may be small due to the ease of communication with the remote points provided by long-range connections; however, the world accessible to an individual may be small due to the absence of properly directed links to connect it to the outside world.

It is therefore natural to ask how the presence of directed links and (or) directed shortcuts quantitatively affect the small-world properties of a network. How can we make a simple model of a small-world network with such random directions? A WS-type model for these networks may be as shown in Fig. 1. However, due to their complexity, these networks should usually be studied by numerical or simulation methods, and they seldom lend themselves to exact analytical treatment.

Aim, structure, and results of this paper

As we will show in this paper, with slight simplification one can introduce simpler models that, while retaining most

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

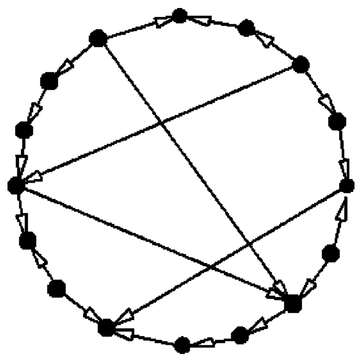

FIG. 1. A WS-type model of a directed network.

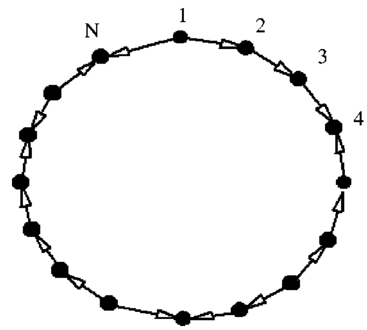

FIG. 3. A regular ring with randomly directed links without shortcuts. You can also see the accessible world of site 1.

of the small world features, are still amenable to analytical treatment. This is what we are trying to do in this paper. In this paper we introduce one such model following our earlier work [23], which was in turn inspired by the work of [24]. The basic simplifying feature of these networks is that all the shortcuts are made via a central site; see Fig. 2. For such a network, many of the small-world quantities can be calculated exactly. In particular, once $V(T)$, defined above, is calculated, many other quantities, such as the average shortest path between two sites, can be obtained. An exact calculation of $V(T)$ is, however, difficult for the case where both the shortcuts and the links have random directions. We therefore proceed in two steps to separate the effects of randomness in the two types of connections. First, in Sec. II, we remove the shortcuts and calculate exactly $V(T)$ for a ring with random links; see Fig. 3. To emphasize the absence of shortcuts we denote this quantity by $V_0(T)$. Note that $V_0(T)$ depends only on the structure of the underlying ring and its short-range connections. Then, in Sec. III, we consider only the effects of randomly directed shortcuts, that is, we let directions of the links on the ring be regular and fixed clockwise, and exactly calculate $V(T)$, where again for emphasis on the shortcuts we denote this quantity $S(T)$.

We then argue, in Sec. IV, that in the scaling limit where the number of sites goes to infinity with the number of shortcuts kept finite, most of the spreading takes place via the links and only from time to time does it propagate to remote points via the shortcuts. In this limit it is plausible to suggest a form for $V(T)$ that takes into account the effect of both the random links and the shortcuts in the form $V(T) = S[V_0(T)]$. This may not be an exact relation but as we will see it will give a fairly good approximation of $V(T)$, as shown by the agreement of our analytical results and the results of simulations. This then means that, in more complicated networks, one can separate the effects of short- and long-range connections and superimpose their effect in a suitable way. We conclude the paper with a discussion.

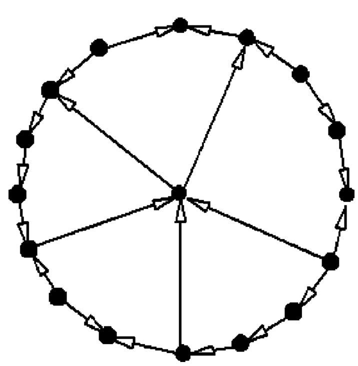

FIG. 2. A simple substitute for the network of Fig. 1.

II. EXACT CALCULATION OF $V_0(T)$ IN A RING WITH RANDOM BONDS

Consider a regular ring of N sites whose bonds are directed randomly. Each link may be directed clockwise with probability r, counterclockwise with probability ℓ, and bidirectionally with probability $1-r-\ell$.

Thus, we have a problem similar to bond percolation in a small-world network. Suppose that at time $T=0$ site number 1 is infected with a virus. We ask the following questions: After T seconds, how many sites have been infected on average? What is the average speed of propagation of this disease in the network? These questions have obvious answers for rings with regularly directed or bidirectional bonds, namely, the number of infected sites are, respectively, T and $2T$, with corresponding speeds of propagation being 1 and 2. In the randomly directed network, the situation is different. For example, if both neighbors of site 1 are directed into this site, this site cannot affect any other site of the network. Such a site, being effectively isolated, has an *accessible world* [18] of zero size (Fig. 3). To proceed with exact calculation, consider the right-hand side of site 1. The probability that exactly $k<T$ extra sites to the right have been infected is $P_+(k):=(1-\ell)^k\ell$, and the probability that exactly T extra sites have been infected is $P_+(T):=(1-\ell)^T$. Therefore, the average number of extra sites infected to the right of the original site is

$$V_0^+(T)=\sum_{k=1}^{T} kP_+(k)=T(1-\ell)^T+\sum_{k=0}^{T-1} k(1-\ell)^k\ell$$
$$=T(1-\ell)^T+\frac{1}{\ell}[1-\ell+(1-\ell)^T(\ell-1-\ell T)]. \tag{1}$$

Going to the large N limit, where

$$N\to\infty, \quad \ell\to 0, \quad \mu:=\ell N, \quad t:=\frac{T}{N}, \quad v_+(t):=\frac{V_0^+(T)}{N}, \tag{2}$$

we find the simple result

$$v_0^+(t)=\frac{1}{\mu}(1-e^{-\mu t}). \tag{3}$$

The same type of reasoning gives the number of sites infected to the left $v_0^-(t)$, and thus the total number of infected sites will be

$$v_0(t)=\frac{1}{\mu}(1-e^{-\mu t})+\frac{1}{\lambda}(1-e^{-\lambda t}), \tag{4}$$

where $\lambda:=rN$. What are the meanings of the scaled variables? The parameter μ is the total number of sparse blocked sites in the way of propagation to the right, with a similar meaning for λ. $v_0(t)$ is the fraction of infected sites up to time t. In a bidirectional lattice, all the sites could be infected after the passage of $T=N/2$ seconds, or at $t=\frac{1}{2}$, and if t passes $\frac{1}{2}$, some of the sites become doubly visited. Therefore, it is plausible, for the sake of comparison, to define a quantity in our ring, namely, the average size of the *accessible* world as $v_0^{acc}:=v_0(\frac{1}{2})$, which turns out to be

$$v_0^{acc}=\frac{1}{\mu}(1-e^{-(\mu/2)})+\frac{1}{\lambda}(1-e^{-(\lambda/2)}). \tag{5}$$

It is seen that the presence of only a small number of blocked bonds causes a significant drop in the average size of this accessible world. For example, a value of $\lambda=\mu=4$ leads to $v_0^{acc}\sim 0.4$. The long-range connections (shortcuts) make the world small with the ease of communication that they provide. However, blockades make the world small in this new sense. The speed of propagation is found from

$$\dot{v}_0(t)=e^{-\mu t}+e^{-\lambda t}. \tag{6}$$

In the symmetric case, where $\lambda=\mu$, Eq. (4) simplifies to

$$v_0(t)=\frac{2}{\mu}(1-e^{-\mu t}), \tag{7}$$

with

$$\dot{v}_0(t)=2e^{-\mu t}. \tag{8}$$

Note that at the early stages of spreading, when $\mu t\ll 1$, and the effects of blocked bonds have not yet been experienced, the infection propagates with a speed equal to 2, as in a regular network. The effect of blocking comes into play when t becomes comparable to $1/\mu$.

As a few shortcuts may enhance the speed of propagation, a few blocked bonds may have the opposite effect. First, the blocks reduce the speed of propagation, as is clear from Eq. (6) and second, and more importantly, they reduce the number of accessible sites, or the size of the accessible world. It will thus be of interest to see how these two effects compete in a random network where there are both shortcuts and blocks. We will study this in the final section of this paper. Toward this end, we first study the effect of directed shortcuts in an otherwise regular ring with no blocks.

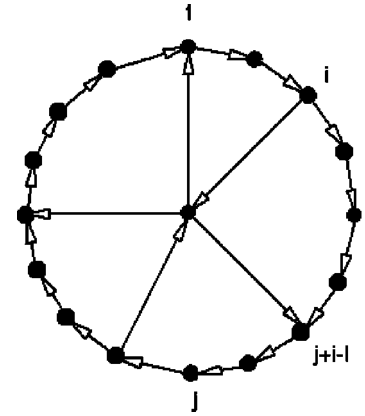

FIG. 4. Randomly directed shortcuts added to a ring with clockwise links.

III. THE LONG-RANGE CONNECTIONS

In this section we consider only the effect of randomly directed shortcuts in the spreading process and obtain exactly the function $S(T)$ for this network; see Fig. 4. Note that this function has the same meaning as $V(T)$, except that for emphasis on the role of shortcuts in it we have adopted a new name for it. We fix a regular clockwise ring. Between a site and the center there is a shortcut going into the center with probability p and out of the center with probability q. The site remains unconnected to the center with probability $1-p-q$. The average number of connections into and out of the center are, respectively, $M_i:=Np$ and $M_o:=Nq$.

Consider sites 1 and j. We want to find the probability that the shortest path between these two sites is of length l, a probability that we denote by $P(1,j;l)$. A typical shortest path of length l connecting these two nodes is shown in Fig. 4, where the first inward connection to the center occurs at site i and the last outward connection from the center occurs at site $j+i-l$. Such a path occurs with probability $(1-p)^{i-1}pq(1-q)^{l-i}$. Summing over all such configurations gives us the probability for the shortest path between sites 1 and j to be of length l. For $l\neq j-1$, we have

$$p(1,j;l\neq j-1)=\sum_{i=1}^{l}(1-p)^{i-1}pq(1-q)^{l-i} = pq\left[\frac{(1-p)^l}{q-p}+\frac{(1-q)^l}{p-q}\right], \tag{9}$$

and $p(1,j;j-1)$ is determined from normalization

$$P(1,j;j-1)=1-\sum_{l=1}^{j-2}P(1,j;l)=\frac{1}{p-q}(p(1-q)^{j-1} - q(1-p)^{j-1}). \tag{10}$$

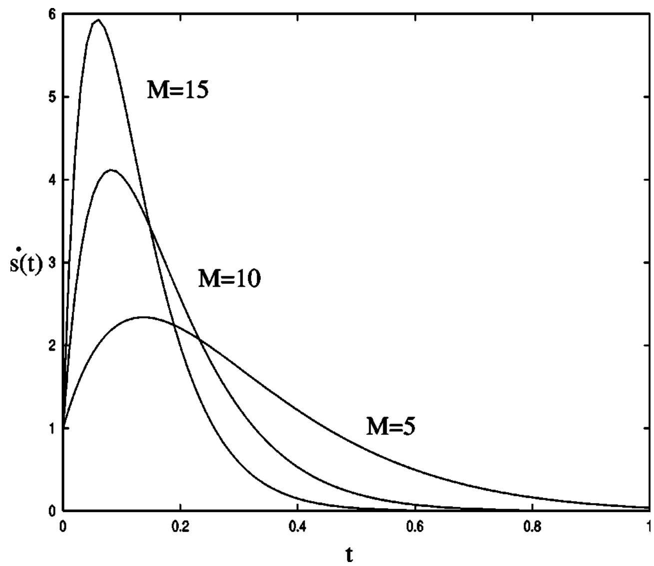

FIG. 5. The speed of propagation in a ring, for several values of randomly directed shortcuts.

Note that $p(1,j;l\neq j-1)$ does not depend on j, a property that is true for standard small-world networks [25].

Now consider a naive spreading process starting at site 1. The number of sites affected up to time T, denoted by $S(T)$, builds up in two ways: via the links on the ring and via the shortcuts. The first way gives a contribution $T+1$ and the second way gives a contribution $(N-T-1)\Sigma_{l=1}^{T}p(1,j;l)$ [25], where $(N-T-1)$ is the number of sites beyond direct reach at time T, which has been multiplied by the probability of any of these sites being a distance shorter than T from site 1 via a shortcut. Putting this together we find

$$S(T)=T+1+(N-T-1)\sum_{l=1}^{T}P(1,j;l)=N+(N-T-1)\times\left[\frac{q}{p-q}(1-p)^{T+1}+\frac{p}{q-p}(1-q)^{T+1}\right]. \qquad (11)$$

In the scaling limit where $N\rightarrow\infty$, $p,q\rightarrow 0$, and where M_i and M_o are kept fixed and $s(t):=S(T)/N$, we find

$$s(t)=1-\frac{1-t}{M_i-M_o}(M_i e^{-M_o t}-M_o e^{-M_i t}). \qquad (12)$$

In the symmetric case, where $M_i=M_o=M$, this equation simplifies to

$$s(t)=1-(1-t)(1+Mt)e^{-Mt} \qquad (13)$$

with the speed of propagation

$$\dot{s}(t)=e^{-Mt}(1+Mt+M^2t-M^2t^2). \qquad (14)$$

Figure (5) shows the speed of propagation as a function of time for several values of M.

IV. THE SPREADING EFFECT IN A DIRECTED SMALL-WORLD NETWORK

We now come to the problem of composing both the blocks and the shortcuts in a model of a small-world network. That is, we consider the ring of Fig. 2 where randomly directed shortcuts are added to a ring with randomly directed links. We cannot obtain exact expressions for this network from first principle probability considerations. However, we can obtain expressions for $v(t)$ in the scaling limit by a heuristic argument and compare our results with those of simulations. Consider Eq. (13). This equation shows how the presence of $2M$ randomly directed shortcuts in a regular clockwise ring affects the spreading effect. On the other hand, we know that the number of sites infected up to time t in the absence of shortcuts has changed to $v_0(t)$. Due to the rarity of shortcuts compared to the regular bonds, most of the spreading takes place via the local bonds. The role of shortcuts is simply to make multiple spreading processes happen in different regions of the network. This role is the same for whatever the underlying lattice is and, therefore, for a general network, at least in the scaling regime, we can assume that Eq. (13) can be elevated to $v(t)=s[v_0(t)]$, i.e.,

$$v(t)=1-[1-v_0(t)][1+Mv_0(t)]e^{-Mv_0(t)}. \qquad (15)$$

For a fully random network where $2M$ randomly directed shortcuts are distributed on a ring with already random links, we assume that this relation holds true with $v_0(t)$ taken from Eq. (4). This suggestion may not provide an exact solution for the network. However, we think it provides a fairly good approximation. In fact, exact solution for the case where all the links on the ring are bidirectional is possible and it confirms the above ansatz, that is, we obtain an exact expression only by setting $v_0(t)=2t$ in the above formula. Moreover, this separation of the effect of short- and long-range connections may also be useful in more complicated networks. Whether or not this assumption is plausible can be checked by comparison with simulations. The results of simulations are compared with those of Eqs. (4) and(12) in Figs. (6) and (7).

V. STATIC PROPERTIES

Once the functions $V(T)$ or $v(t)$ are obtained, the static properties of the network, i.e., the average shortest path between two arbitrary sites and its probability distribution, can be calculated directly.

Since $V(T)$ by definition is the number of sites whose shortest distance to site 1 is less than or equal to T, we find the number of sites whose shortest distance is exactly T to be $V(T)-V(T-1)$. Since site 1 is an arbitrary site, we find the probability distribution of the shortest distance between two arbitrary sites that are accessible to each other as $P(T)=[V(T)-V(T-1)]/V_{acc}$, where V_{acc} is the average size of the accessible world. (There is of course a slight approximation here in that we are taking averages of the denominator and numerator separately.)

For a regular ring with shortcuts, $V_{acc}=N$, since all the sites are accessible. We will discuss the case of random rings

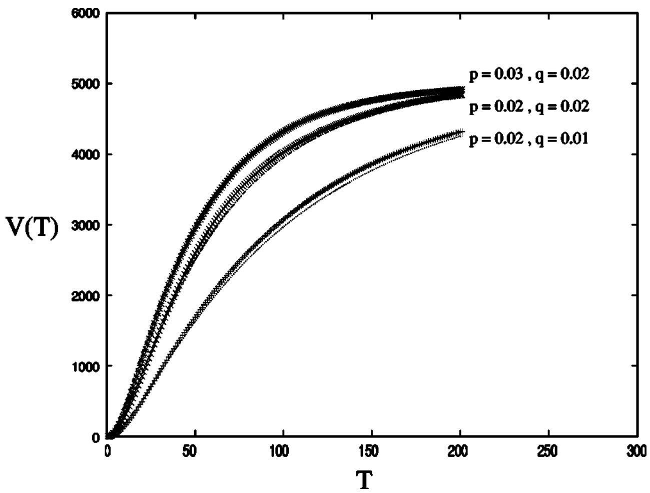

FIG. 6. $V(T)$ for a fully random network in the case $N=5000$, $r=0.02$, $l=0$. Analytic results (lines) vs simulations (symbols), which have been averaged over 1000 realizations of the network.

in a sequel to this paper. In the scaling regime, the above formulas transform to

$$P(t)=\dot{v}(t). \tag{16}$$

Note that $P(t)$ is normalized, i.e., $\int_0^1 P(t)dt=v(1)-v(0)=1$. The average shortest path for the network of Fig. 4 when $M_i=M_o=M$ turns out to be

$$\langle t\rangle \equiv \int_0^1 tP(t)dt=\int_0^1 t\dot{v}(t)=\frac{1}{M^2}[2M-3+(M+3)e^{-M}]. \tag{17}$$

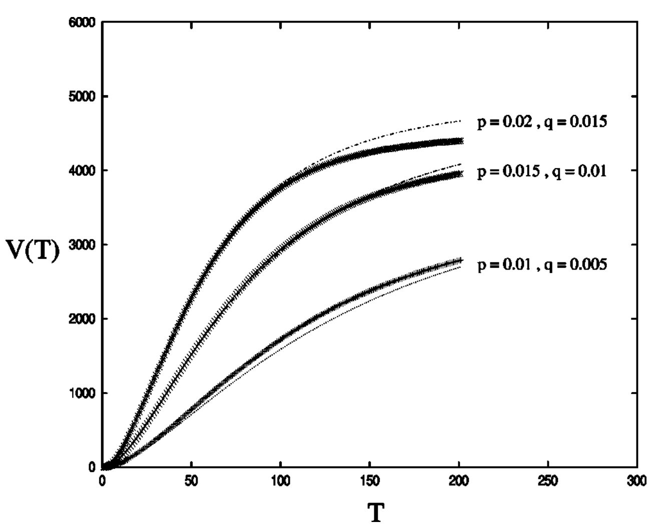

FIG. 7. $V(T)$ for a fully random network in the case $N=5000$, $r=l=0.005$. Analytic results (lines) *vs*. simulations (symbols).

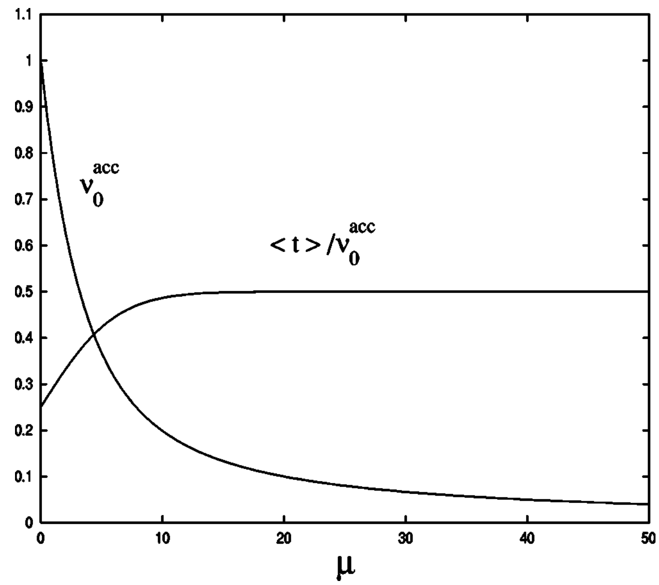

FIG. 8. The average size of the accessible world and the average shortest path for a regular ring with randomly directed bonds without shortcuts.

This is in accord with the result of [24]. This formula shows that the presence of a small number of shortcuts causes a significant drop in the average shortest path from 1 to very small values. In this sense, the world gets smaller by long-range connections.

We now study the static effects of random directed bonds on a ring without shortcuts. The presence of blocks makes the world small in a different sense, namely, for each site the number of accessible sites gets smaller. In fact, the average size of the world accessible to a site is not N anymore, but is given by $V(N/2)$ [see the paragraph leading to Eq. (5)]. Hence, the probability of shortest paths is given by $P(T):=[V(T)-V(T-1)]/V(N/2)$, or in the scaling limit by

$$P(t):=\frac{\dot{v}(t)}{v\left(\frac{1}{2}\right)}. \tag{18}$$

This probability is normalized, i.e., $\int_0^{1/2}P(t)dt=1$. We obtain from Eq. (18)

$$\langle t\rangle=\frac{1}{v\left(\frac{1}{2}\right)}\int_0^{1/2}t\dot{v}(t)dt. \tag{19}$$

However, in order to assess the situation in this network, we should compare the average shortest path with the size of this small world itself, namely, we should calculate $\langle t\rangle/v_0^{acc}$. Inserting Eq. (7) into Eq. (19), we find

$$\frac{\langle t\rangle}{v_0^{acc}}=\frac{2-(\mu+2)e^{-\mu/2}}{4(1-e^{-(\mu/2)})^2}. \tag{20}$$

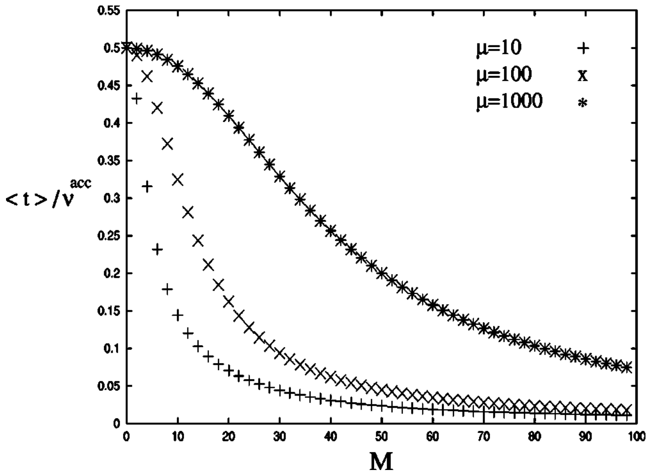

FIG. 9. The average shortest path for a fully random network.

Figure (8) shows both the average size of the accessible world v_0^{acc} and the ratio $\langle t\rangle/v_0^{acc}$ of the average shortest path to the size of the accessible world as a function of the number of blocks μ. It is seen that for $\mu=0$, when there is no block, the size is 1 and the average of the shortest path is $\frac{1}{4}$, as it should be. With a few blocks, the size drops dramatically and the average of the shortest path within the world increases. Note that with increasing μ the average shortest path increases to its maximum value of $\frac{1}{2}$.

For the fully random network, we use Eqs. (15) and (18) to obtain the average of shortest path. The result is shown in Fig. (9) for several values of the parameters.

VI. CONCLUSION

We have studied the effect of directed short- and long-range connections in a simple model of a small-world network. In our models, all the shortcuts pass via a central site in the network. This makes possible an almost exact calculation of many of the properties of the network. We have calculated the function $V(T)$, defined as the number of sites affected up to time T when a naive spreading process starts in the network. As opposed to shortcuts, the presence of unfavorable bonds has a negative effect on this quantity. Hence, the spreading process may be able to affect only a fraction of the total sites of the network. We have defined this fraction to be the average size of the accessible world in our model and have calculated it exactly for our model. We have also studied the interplay of shortcuts and unfavorable bonds on the small-world properties, such as the size of the accessible world, the speed of propagation of a spreading process, and the average shortest path between two arbitrary sites. Our results show that one can separately take into account the effect of randomness in the directions of shortcuts and the short-range connections in the underlying lattice, and, at the end, superimpose the two effects in a suitable way. We expect that this will also hold in more complicated lattices of small-world networks.

[1] L.F. Lago-Fernandez, R. Huerta, F. Corbacho, and J.A. Siguenza, Phys. Rev. Lett. **84**, 2758 (2000).
[2] J.W. Bohland and A.A. Minai, Neurocomputing **38–40**, 489 (2001).
[3] H. Jeong, B. Tombor, R. Albert, Z.N. Oltvai, and A.-L. Barabasi, Nature (London) **407**, 651 (2000).
[4] R.V. Kulkarni, E. Almaas, and D. Stroud, Phys. Rev. E **61**, 4268 (2000).
[5] A.R. Puniyani, R.M. Lukose, and B.A. Huberman, e-print cond-mat/0107212.
[6] R. Albert, H. Jeong, and A.L. Barabasi, Nature (London) **401**, 130 (1999).
[7] M.E.J. Newman, J. Stat. Phys. **101**, 819 (2000).
[8] J. Davidsen, H. Ebel, and S. Bornholdt, Phys. Rev. Lett. **88**, 128701 (2002).
[9] R. Albert and A.-L. Barabasi, Rev. Mod. Phys. **74**, 47 (2002).
[10] B. Bollaoas, *Random Graphs* (Academic, New York, 1985).
[11] L.A.N. Amaral, A. Scala, M. Barthelemy, and H.E. Stanley, Proc. Natl. Acad. Sci. U.S.A. **97**, 11149 (2000).
[12] D.J. Watts and S.H. Strogatz, Nature (London) **393**, 440 (1998).
[13] A. Barrat and M. Weight, Eur. Phys. J. B **13**, 547 (2000).
[14] M.E.J. Newman, C. Moore, and D.J. Watts, Phys. Rev. Lett. **84**, 3201 (2000).
[15] M. Kuperman and G. Abramson, Phys. Rev. Lett. **86**, 2909 (2001).
[16] D.H. Zanette, Phys. Rev. E **64**, 036204 (2001).
[17] S. Jespersen and A. Blumen, Phys. Rev. E **62**, 6270 (2000).
[18] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, and J. Wiener, Comput. Netw. **33**, 309 (2000).
[19] B. Tadic, Physica A **293**, 273 (2001).
[20] O. Sporns, G. Tononi, and G.M. Edelman, Neural Networks **13**, 909 (2000).
[21] A.D. Sanchez, J.M. Lopez, and M.A. Rodriguez, Phys. Rev. Lett. **88**, 048701 (2002).
[22] B. Tadic, e-print cond-mat/0110033.
[23] V. Karimipour and A. Ramzanpour, Phys. Rev. E **65**, 036122 (2002).
[24] S.N. Dorogovtsev and J.F.F. Mendes, Europhys. Lett. **50**, 1 (2000).
[25] E. Almaas, R.V. Kulkarni, and D. Stroud, Phys. Rev. Lett. **88**, 098101 (2002).