Congestion-gradient driven transport on complex networks

Bogdan Danila,[1,*] Yong Yu,[1] Samuel Earl,[2] John A. Marsh,[2] Zoltán Toroczkai,[3,†] and Kevin E. Bassler[1,‡]
[1]*Department of Physics, The University of Houston, Houston, Texas 77004, USA*
[2]*SI International, Beeches Professional Campus, Rome, New York 13440, USA*
[3]*Los Alamos National Laboratory, MS B258, Los Alamos, New Mexico 87545, USA*
(Received 31 March 2006; revised manuscript received 4 August 2006; published 19 October 2006)

We present a study of transport on complex networks with routing based on local information. Particles hop from one node of the network to another according to a set of routing rules with different degrees of congestion awareness, ranging from random diffusion to rigid congestion-gradient driven flow. Each node can be either source or destination for particles and all nodes have the same routing capacity, which are features of *ad hoc* wireless networks. It is shown that the transport capacity increases when a small amount of congestion awareness is present in the routing rules, and that it then decreases as the routing rules become too rigid when the flow becomes strictly congestion-gradient driven. Therefore, an optimum value of the congestion awareness exists in the routing rules. It is also shown that, in the limit of a large number of nodes, networks using routing based on local information jam at any nonzero load. Finally, we study the correlation between congestion at node level and a betweenness centrality measure.

DOI: 10.1103/PhysRevE.74.046114 PACS number(s): 89.75.Hc, 05.60.−k, 89.20.Hh

I. INTRODUCTION

Network transport has been a topic of intensive research in recent years [1–9] due to the wide variety of network systems of practical interest, both natural and human made. These include [1,10,11] the internet, the worldwide web, wireless communication networks, networks of biological processes, social networks, and various distribution networks that make up the infrastructure of technological society. Research topics include the robustness and efficiency of computer and distribution networks, the spread of disease, and the study of the interplay between various biological or economic processes.

A major result obtained recently [1,10–12] is the realization that both natural and man-made networks often display scale-free topology, i.e., they are characterized by a power-law distribution of the node degrees. The explanation for this observation, namely, the connection between form and functionality in networks, remains open. There exist a good number of stochastic models for network growth and evolution that lead to scale-free graphs, such as the preferential attachment model of Barabási and Albert [13]. However, they are not directly related to the transport function of the network. In this paper we explore the possibility of a more direct connection between topology and transport function in complex networks.

Specifically, we explore the dependence of transport capacity on the structure of the network, as well as on the routing rules. We consider a simple transport model, motivated in part by the conditions applicable to a wireless *ad hoc* network [14], where any node may be either the source or the destination of information, and in which all nodes have equal routing capacity. We consider a dynamics in which "particles" (representing information, energy, or goods) are routed sequentially from node to node along undirected network edges until they reach their randomly preassigned destination. The routing choices are made based on local information only. Particles are added to the network at a constant rate, with equal probability for every node, and removed upon reaching their destination. The network load is quantified by the particle generation rate. Each individual network is characterized by a critical value of the load [9,15] beyond which it enters a jammed state, i.e., the average time to destination diverges and consequently the number of particles on the network increases indefinitely. For networks characterized by a given set of structure and routing parameters, we define the average jamming fraction as the ratio of the number of network realizations that end up jamming divided by the total number of realizations. The critical value of the load at which the average jamming fraction reaches 50% is used to quantify the network transport capacity from a statistical point of view.

This study is also motivated in part by the work presented in Refs. [16,17], where the notion of gradient flow networks is introduced. Gradient networks are directed subnetworks of an undirected "substrate" network in which each node has an associated scalar potential and one out-link that points to the node with the smallest (or largest) potential in its neighborhood, defined as the reunion of itself and its nearest neighbors on the substrate network. Here we use a congestion-aware routing rule characterized by a parameter that defines a continuous change from random diffusion to higher degrees of congestion awareness. Increasing congestion awareness results in improved network load balancing. The limiting case of congestion-aware routing is that of rigid congestion-gradient driven flow, when particles are always routed toward the least congested neighbor. In this case, transport takes place along a gradient network, but one that changes dynamically in a manner that is correlated with the degrees of congestion of the nodes.

*Electronic address: dbogdan@mail.uh.edu
†Present address: Department of Physics, University of Notre Dame, Notre Dame, IN 46556.
‡Electronic address: bassler@uh.edu

We present results for two network models characterized by different topologies. The first model is that of random (also known as Erdős-Rényi) networks, which are characterized [18] by the number of nodes and by a constant probability p for the connection between any given pair of nodes. The second model is Barabási-Albert networks, which are grown by preferential attachment [13] and, in the limit of large number of nodes, exhibit a scale-free topology. When results for the two types of networks are compared in this paper, the number of nodes and the average degree are the same. Recently, a detailed analysis of jamming in gradient networks [19] suggested the existence of a critical value of the average degree. For values of the average degree below this critical value, large scale-free networks are somewhat more prone to jamming than random networks with the same number of nodes and average degree while the opposite is true above the critical value.

Our main result is the existence of an optimum value of the congestion awareness parameter, both for random and for Barabási-Albert networks. Below this optimum value, the transport capacity of the network increases with the degree of congestion awareness due to the particles being more likely to avoid waiting in the queues of the busiest nodes. However, above the optimum value the average transport capacity decreases with increasing congestion awareness in spite of the shorter waiting times. The decrease is mainly due to the formation of transport traps, which prevent particle flow between parts of the network. Our results also show that, for a value of the average connectivity below the critical value suggested in [19], Barabási-Albert networks are indeed more prone to jamming than random networks regardless of the degree of congestion awareness.

The outline of the paper is as follows. In Sec. II we give a detailed description of our model and of the criteria that were used to detect jamming. In Sec. III we present results for the jamming fraction as a function of load for different values of the number of nodes and of the congestion awareness parameter. Also in Sec. III we present results for the critical load as a function of both the number of nodes and the congestion awareness parameter. In Sec. IV we present an in-depth analysis of the transport behavior. Section V summarizes our results and conclusions.

II. MODEL

Our model network consists of a set of N identical nodes linked together in such a way that they form a connected, undirected graph characterized by a symmetric adjacency matrix A, defined by

$$A_{ij} = \begin{cases} 1 & \text{if } i \text{ and } j \text{ are connected,} \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

The degree d of a node is defined as the number of links (edges) that connect it to other nodes. We have studied both the case of Erdős-Rényi [18] (random) networks, characterized by a binomial distribution of the node degrees, and the case of Barabási-Albert [13] networks, for which the node degrees obey a power-law distribution. In the case of a random network, the average degree is given [18] by $\langle d\rangle=(N-1)p$, where p is the probability for a link to exist between any given pair of nodes. We discard random networks that turn out to be disconnected. Networks grown according to the Barabási-Albert algorithm have an average degree given by

$$\langle d\rangle(N) = 2m\frac{N - N_0 + L_0/m}{N}, \tag{2}$$

where N_0 is the initial number of nodes, $L_0=N_0(N_0-1)/2$ is the initial number of links, and m is the number of links created with every new node. These networks are always connected. The condition for the average degree to be independent of N is $N_0=2m+1$. All results for Barabási-Albert networks presented in this paper have been obtained for values of $m=3$ and $N_0=7$, and consequently $\langle d\rangle=6$. To facilitate comparison, the probability p that characterizes the random networks has been adjusted to $p=6/(N-1)$. Varying the average degree does not lead to any qualitative changes in the results. However, the transport capacity shows an overall increase with increasing average degree, while its maximum at the optimum value of the congestion awareness parameter becomes less prominent.

Information packets (or any other entities) transported along the network are represented by particles that hop from one node to the next along the graph edges according to rules that will be discussed in the following paragraph. Each node has a particle queue, thus being capable of holding more than one particle at a time. Particles waiting in a queue are processed according to the "first-in, first-out" rule. Updating of the network is done sequentially, so that only one particle is moving or being created at any given instant. Each time step consists of a random sequence of updating events, in the course of which N nodes are randomly selected one at a time to forward a particle, and a random number of new particles are added to the network one at a time, with an average rate of R new particles per time step. All nodes are equally likely to receive a new particle. Every new particle generated is assigned a destination, again with equal probability from among all nodes, and is placed at the end of the queue of its origin node. Upon reaching their destination, particles are removed from the network.

The particle hopping dynamics is characterized by the routing rule used to choose the next location of a particle sitting on a given node from among its neighbors. If no record is kept of the possible paths toward a destination, the routing rule will necessarily be based on local information only, such as the degrees of the neighboring nodes or any other parameters that characterize their congestion status. If, on the other hand, knowledge of some or all possible routes to the destination is used in making the decision, the rule is said to be based on global information. Among the routing rules based on local information we distinguish between random diffusion, in which case the next location is chosen with uniform probability distribution from among the neighbors of the current node, and congestion-aware rules that take into account the degrees of congestion of the neighbors. The extreme case of congestion aware routing is that of rigid congestion-gradient driven flow, when particles are trans-

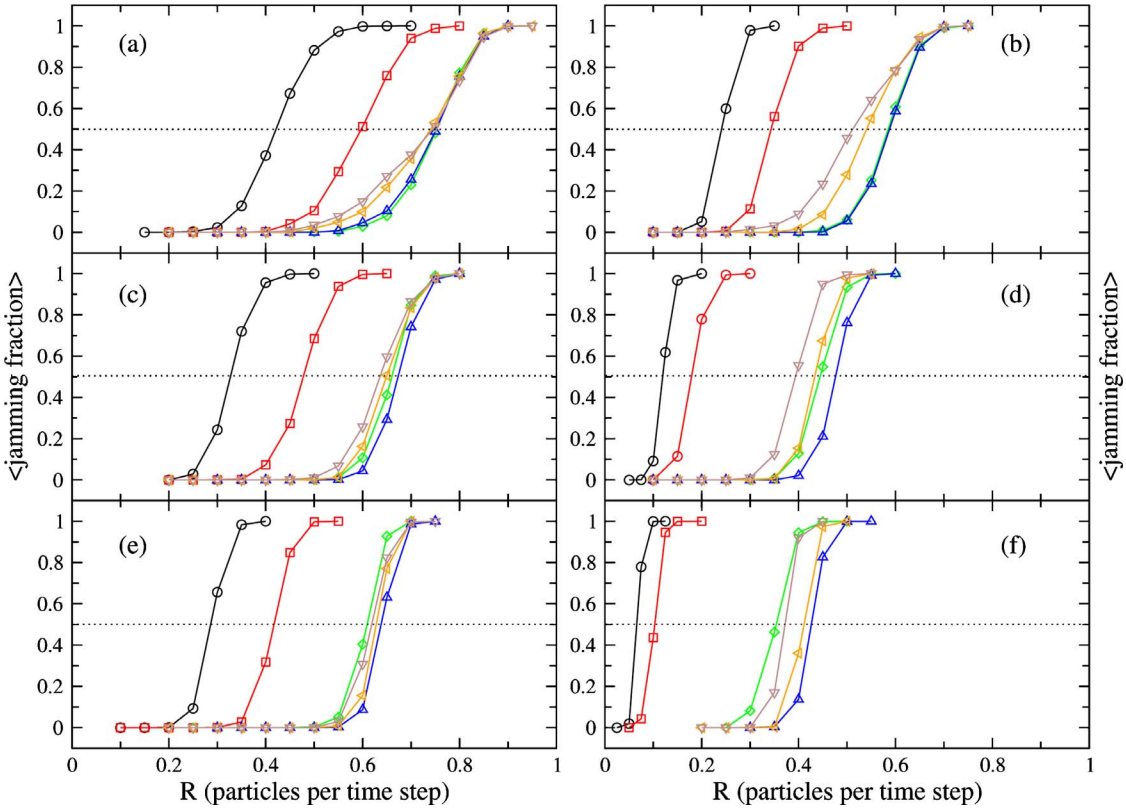

FIG. 1. (Color online) Average jamming fraction as a function of load for $\beta=0$ (circles), 0.1 (squares), 0.5 (diamonds), 1 (up triangles), 3 (left triangles), and 10 (down triangles). Results are for random (a), (c), (e) and Barabási-Albert (b), (d), (f) networks with $N=30$ (a), (b), 100 (c), (d), and 300 (e), (f) nodes.

ported only in the direction of the gradient network generated by using a congestion parameter as scalar potential [16]. We have chosen a congestion-aware routing rule that uses local information, namely, the queue lengths of the neighboring nodes. According to this rule, for a particle currently at node i the probability of hopping to node j is given by

$$P_{ji} = \frac{A_{ji}(q_j+1)^{-\beta}}{\sum_{k=1}^{N} A_{ki}(q_k+1)^{-\beta}}, \tag{3}$$

where q_k is the queue length of node k and β is an adjustable parameter. Varying β from 0 to ∞ allows a smooth transition between the case of random diffusion and that of transport along the gradient network generated by using the queue length as scalar potential. In practice, a value of $\beta \approx 10$ is sufficiently high to lead to essentially rigid congestion aware routing.

For a given realization of the network, simulations were run for a maximum of 10^5 time steps. The simulations were interrupted and jamming declared at any time if either the length of any queue exceeded the maximum value of $20N$, or the total number of particles n on the network exceeded $50N$. In addition, a linear regression for $n(t)$ was attempted every 2×10^4 time steps. Jamming was declared if the result of the regression was a positive slope with a Pearson correlation coefficient $r^2>0.9$ over the last interval of 2×10^4 time steps.

III. RESULTS FOR THE TRANSPORT CAPACITY

In this section we present results for the average jamming fraction f_j computed as a function of the load R as well as for the average transport capacity, quantified by the critical load R^* at which the jamming fraction reaches 50%. The results for the jamming fraction were obtained by averaging over 1000 realizations of the network for a given set of parameters R, β, and N. The critical load is obtained by a linear fit between two points on the $f_j(R)$ curve, one below $f_j=0.5$ and another one above.

Figure 1 shows plots of the average jamming fraction as a function of R for both random [Figs. 1(a), 1(c), and 1(e)] and Barabási-Albert [Figs. 1(b), 1(d), and 1(f)] networks and for values of $N=30$, 100, and 300. Each plot has six curves, corresponding to values of $\beta=0$, 0.1, 0.5, 1, 3, and 10. Jamming of the random networks is seen to occur at higher values of the load than in the case of Barabási-Albert networks with the same number of nodes. This is true regardless of the number of nodes and for all routing regimes, which range from random diffusion ($\beta=0$) to essentially deterministic congestion-aware routing ($\beta=10$). The importance of even a slight degree of congestion awareness is also apparent. In most plots, the critical load corresponding to $\beta=0.1$ is almost halfway between the value corresponding to $\beta=0$ and the highest value. An important empirical result is that, regardless of network topology or routing regime, the critical load decreases as the number of nodes increases. Presumably, this

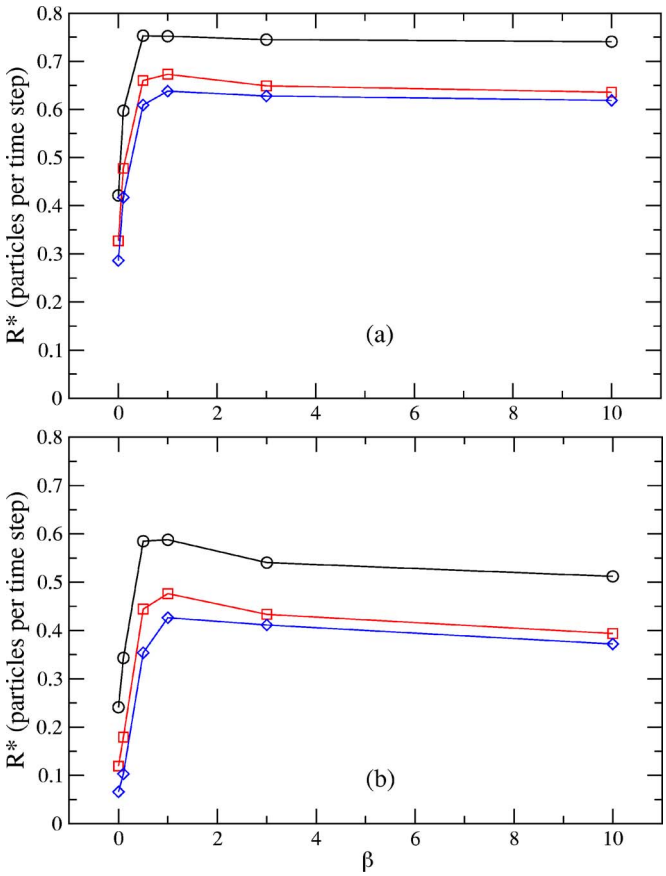

FIG. 2. (Color online) Critical load of the network as a function of the congestion awareness parameter for $N=30$ (circles), 100 (squares), and 300 (diamonds). Results are for random (a) and Barabási-Albert (b) networks with $\langle d\rangle=6$.

is due to the fact that the average length of the walk between two nodes $\langle L\rangle$ increases slightly faster than N^2. This result is even more significant in view of the fact that the most meaningful quantity is, arguably, the average number of new particles per node and time step R/N, which does not depend on the number of nodes as the network size is varied. In the limit of large number of nodes, this quantity approaches zero. The significance of this is that large transport networks with routing based on local information only will jam if each node injects the quantity to be transported at any nonzero average rate.

Perhaps the most important result is the existence of an optimum value of the congestion awareness parameter β. This can be seen more clearly in Figs. 2(a) and 2(b), where the critical load R^* is plotted against β for random and for Barabási-Albert networks, respectively. Regardless of network topology or number of nodes, the critical load increases as the congestion awareness parameter increases until it reaches a maximum. For higher values of β, the critical load decreases monotonically and seems to approach a finite value as $\beta\rightarrow\infty$. This seemingly counterintuitive result is explained in Sec. IV, where we show that it is due to the onset of transport traps. The optimum value of the congestion awareness parameter increases with network size, starting from about 0.5 for very small ($N=30$) networks, and appears to approach an asymptotic value of about 1 for networks with a few hundred nodes. To facilitate comparison between random and Barabási-Albert networks, we present here results for $\langle d\rangle=6$. However, for reasons that will be explained in Sec. IV, the maximum of the critical load is more pronounced at lower average degrees. We also note that, as the number of nodes on the network increases, the maximum of the transport capacity becomes less prominent but does not disappear.

IV. DETAILED ANALYSIS

A. Betweenness definition and calculation

We begin this section by defining a measure that characterizes the nodes of the network from a purely topological point of view. Various definitions of centrality measures have been proposed, each having its scope of applications [3,9,20–24]. Of these, betweenness centrality measures are most appropriate for characterizing transport on networks because they are proportional to the average number of times a path passes through a given node. The differences between the various betweenness definitions are in the types of paths that are considered and the way the number of passes is counted. The measure that we use in this paper is a slightly modified version of the betweenness centrality defined by Guimerà *et al.* [9]. Random walks along the network are considered for particles with a given source node s and destination node t. The betweenness b_i^{st} of a node i with respect to s and t is defined as the average number of times a random walk between s and t passes through i. Finally, the global betweenness of node i is defined as the average over all s and t of b_i^{st}.

In principle, a calculation of the betweenness can be done by starting from the transition probability matrix P, whose elements P_{ij} give the probability for a particle to arrive at node i coming from node j. However, the computation of P is straightforward only in the case of random diffusion. Nevertheless, the results presented in the remainder of this section show that, even in the case of congestion aware routing, there exists a strong correlation between the random walk betweenness of a node computed in the case of random diffusion and the average number of particles it receives per time step. Assuming at first that particles are never removed from the network, we have

$$P_{ij}=\frac{A_{ij}}{d_j}, \tag{4}$$

where d_j is the degree of node j. For particles with destination node t, we can account for their removal once they reach their destination by replacing all elements in column t by 0. The matrix thus obtained will be denoted by P^t. It is easy to verify that $(P^t)^n_{ij}$, where $(P^t)^n$ is the nth power of P^t, gives the probability of arrival at node i coming from node j in n steps for a particle with destination t. Then the betweenness of node i with respect to a source node s and a destination node t is defined as

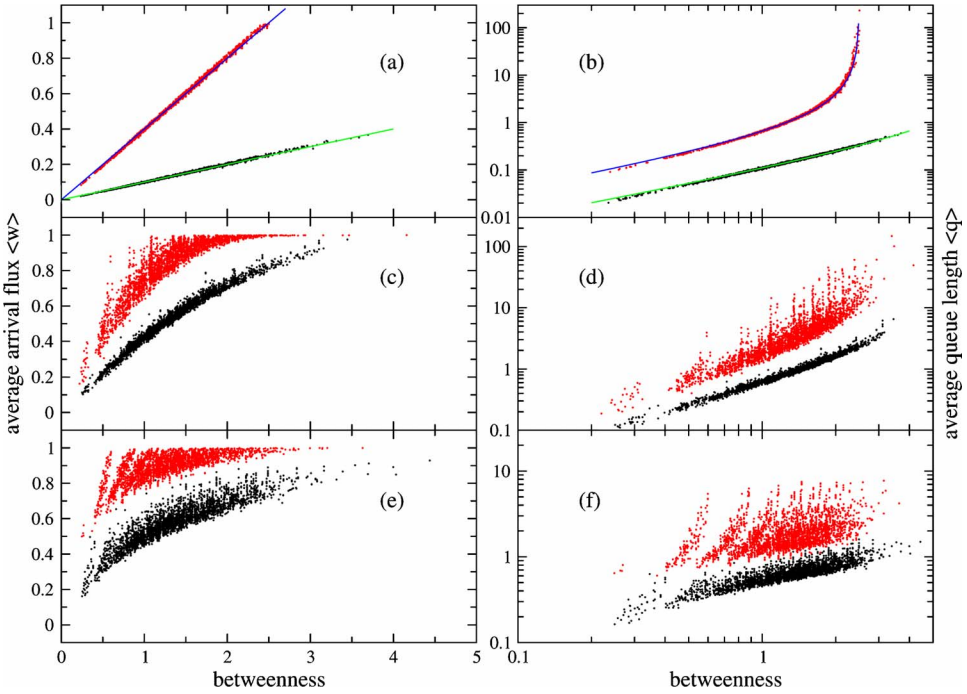

FIG. 3. (Color online) Correlation plots of the average particle flux (left) and average queue length (right) versus betweenness for random networks. The values of β are 0 (a), (b), 0.5 (c), (d), and 10 (e), (f). Lower (black) sets of dots are for $R=0.1$ (a), (b) and $R=0.4$ (c)–(f) while upper (red) sets of dots are for $R=0.4$ (a), (b) and 0.65 (c)–(f).

$$b_i^{st} = \sum_{n=0}^{\infty} (P^t)_{is}^n, \tag{5}$$

where $(P^t)^0$ is taken to be the unit matrix I. Equation (5) can be rewritten as

$$b_i^{st} = (I - P^t)_{is}^{-1}. \tag{6}$$

Finally, the global betweenness of node i is given by

$$b_i = \frac{1}{N^2} \sum_{s,t=1}^{N} (I - P^t)_{is}^{-1}. \tag{7}$$

If this quantity is multiplied by the rate at which new particles are added to the network R, we obtain the average total number of particles (new or from other nodes) reaching node i in the course of a time step.

In its most straightforward form, the algorithm requires N matrix inversions, the number of flops thus being $O(N^4)$. However, since the transformation from $(I-P^t)$ to $(I-P^{t'})$ is a rank-2 perturbation of $(I-P^t)$, one can use the Sherman-Morrison-Woodbury formula [25] to compute $N-1$ of the inverses at a much lower cost once a first inverse has been computed. Specifically, the computation can be done as follows. First, compute $Q=(I-P^t)^{-1}$ for a given t, for example $t=1$. Then for every $t' \neq t$ we can write

$$P^{t'} = P^t + [P_{:t} | - P_{:t'}] \begin{bmatrix} e_t^* \\ \hline e_{t'}^* \end{bmatrix}, \tag{8}$$

where $P_{:s}$ denotes the s column of matrix P and e_s^* the s row of the unit matrix. The vertical and horizontal bars within the two bracketed expressions in (8) are column and row delimiters, respectively. By an application of the Sherman-Morrison-Woodbury formula to $(I-P^{t'})^{-1}$ using (8), and after a little bit of algebra we find

$$(I - P^{t'})^{-1} = Q + [QP_{:t} | e_{t'} - Q_{:t'}] \times \begin{bmatrix} 1 - (QP)_{tt} & Q_{tt'} \\ - (QP)_{t't} & Q_{t't'} \end{bmatrix}^{-1} \begin{bmatrix} Q_{t:} \\ \hline Q_{t':} \end{bmatrix}. \tag{9}$$

Here $e_{t'}$ is the t' column of the unit matrix, while $Q_{:s}$ and $Q_{s:}$ are the s column and row, respectively, of matrix Q. Note that, once Q and the product $QP_{:t}$ have been computed, all coefficients in (9) are known for every t'. By an inspection of (9) we find that the total number of flops necessary for the computation of the N inverses becomes $O(5N^3)$.

B. Correlations with congestion parameters

Next we investigate the correlation of the betweenness measure b with two parameters that can be used to characterize the congestion status of a node. Specifically, we look at the correlation with the time averages of the queue length q and particle flux w, which is defined as the number of particles received by the node in a time step. The results presented in this subsection pertain only to cases of steady state transport, when the network does not jam. Jamming can be characterized in terms of the particle flux as a situation in which its average exceeds the maximum value of $\langle w \rangle = 1$ for at least one node (all nodes can process on average at most one particle per time step). In addition, since the results are similar regardless of the number of nodes, we only show results for random and Barabási-Albert networks with $N=30$ nodes.

Figures 3 and 4 show correlation plots of the average particle flux and average queue length against node betweenness in the case of random and Barabási-Albert networks, respectively, for values of $\beta=0$, 0.5, and 10. The average particle flux is plotted in Figs. 3(a), 3(c), and 3(e) and 4(a), 4(c), and 4(e), while Figs. 3(b), 3(d), and 3(f) and 4(b), 4(d),

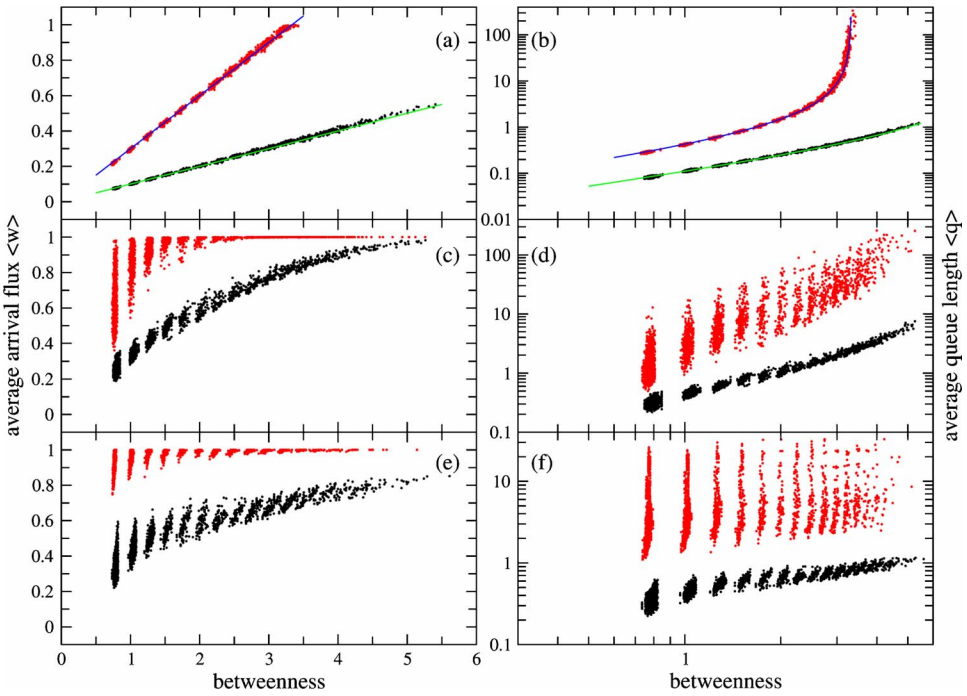

FIG. 4. (Color online) Correlation plots of the average particle flux (left) and average queue length (right) versus betweenness for Barabási-Albert networks. The values of β are 0 (a), (b), 0.5 (c), (d), and 10 (e), (f). Lower (black) sets of dots are for $R=0.1$ (a), (b) and $R=0.3$ (c)–(f) while upper (red) sets of dots are for $R=0.3$ (a), (b) and $R=0.65$ (c)–(f).

and 4(f) show the average queue length. The averages are computed over 10^5 time steps. Each plot contains points corresponding to a value of R that is well below R^* (lower set of dots, colored black online) and to a value close to R^* (upper set of dots, colored red online). Each set contains 3000 dots, corresponding to 100 network realizations with a given set of parameters.

As one would expect based on the way the betweenness measure is defined, it provides an essentially exact description of the statistics of the particle transport in the case of random diffusion ($\beta=0$). There is a very strong linear correlation between the average particle flux through a node and its betweenness, regardless of the value of the load R. If the points corresponding to a given load are fitted with a straight line the value of the slope is, to within the statistical uncertainty, equal to R [straight lines in Figs. 3(a) and 4(a)]. We thus see that the probability for a network to jam at a given load under a random diffusion routing regime is the same as the probability of having at least one node with betweenness higher than the inverse of the load. This also explains the uneven jamming that characterizes networks when the random diffusion routing rule is used, since only the nodes for which $bR>1$ are jamming. As for the average queue length, its correlation with the betweenness measure is approximately linear only when the network load is low, well below its critical value R^*. At higher loads, the correlation is still strong but nonlinear. This is due to the fact that the functional dependence of the average queue length on the average particle flux is nonlinear. One can show that, in the case of random diffusion, the number of particles in a queue is distributed according to an exponential distribution law [9,26] given by

$$p(q) = (1 - \langle w \rangle)\langle w \rangle^q, \tag{10}$$

where $p(q)$ is the probability of having q particles in the queue of a node characterized by an average particle flux $\langle w \rangle$. From (10) it follows that the average queue length is related to the average particle flux by [9,26]

$$\langle q \rangle = \frac{\langle w \rangle}{1 - \langle w \rangle}. \tag{11}$$

The last formula, with $\langle w \rangle = bR$, has been used to generate the curves in Figs. 3(b) and 4(b). As can be seen, these curves provide an excellent fit for the data points obtained from simulations.

Things are more complicated in the case of congestion-aware routing. Neither the average particle flux nor the average queue length correlate linearly with the random walk betweenness except at very low loads. The correlations also become weaker as the load increases. The strength of the correlation is nevertheless surprising, especially in the case of the average flux, and the plots provide some useful insight into the behavior of the network transport. The most meaningful comparison is between the $b-\langle w \rangle$ correlation plots [Figs. 3(a), 3(c), and 3(e) and 4(a), 4(c), and 4(e)]. By comparing these figures we see that the way a congestion-aware routing rule is able to cope with higher loads than random diffusion is by diverting part of the traffic toward lower betweenness nodes. High betweenness nodes that would otherwise have exceeded the maximum particle flux now are kept just below $\langle w \rangle = 1$, while lower betweenness nodes experience higher traffic. The rigid straight-line correlation between node betweenness and average particle flux is bent, which also explains why networks with a congestion-aware routing rule jam more uniformly. The load balancing becomes more pronounced as the load increases. At a given load, the balancing is stronger for higher values of the congestion awareness parameter.

A simple explanation for the decrease in the average transport capacity of the network at large β is that the routing

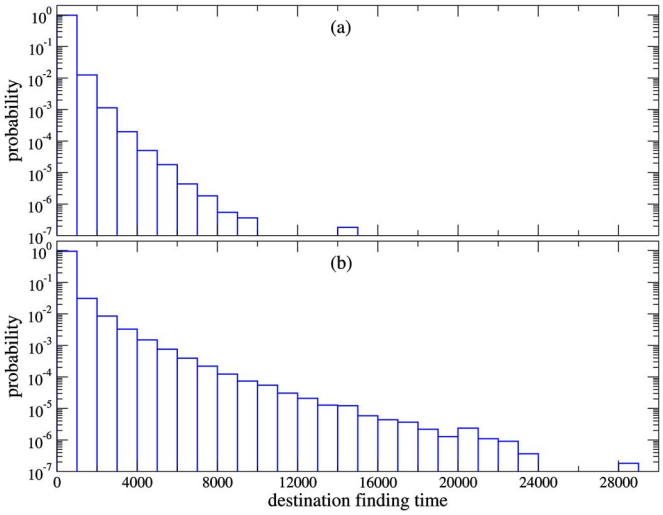

FIG. 5. (Color online) Histograms of the destination finding times at $\beta=1$ (a) and 10 (b) for Barabási-Albert networks with $N=30$ nodes at $R=0.55$.

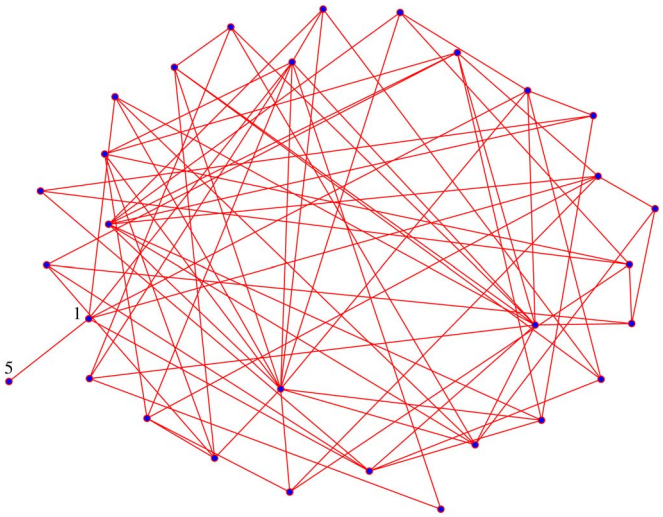

FIG. 6. (Color online) A random network with $N=30$ and $\langle d\rangle=6$. Nodes 1 and 5 form a "trap" as described in Sec. IV C.

rule "overshoots" its goal of balancing the network load. Since large values of β lead to more rigorous load balancing, the particle flux through the high betweenness nodes is lower than in the case of optimum β. But high betweenness nodes tend to have high degrees as well and are in general better positioned to route a particle between two weakly connected parts of the network. This leads to situations where some particles wander about for longer periods of time before finding their destination and consequently to increased likelihood for the average particle flux through some nodes to exceed its maximum allowable value. To verify this, we looked at the statistics of the destination finding times.

In Fig. 5 we present histograms of the destination finding times at $\beta=1$ and $\beta=10$ for Barabási-Albert networks with $N=30$ nodes. The network load is $R=0.55$ (close to R^* for both values of β) and the statistics is based on 100 nonjamming realizations. It is apparent from Fig. 5 that, even in the case of steady state transport, there are particles with much longer destination finding times at $\beta=10$ than at $\beta=1$. This is an indication that a routing rule too rigidly based on congestion awareness may fail to achieve its goal of minimizing the finding times even in the case of jamming networks. The average time a particle sits in a queue is shorter, but this may be compensated by the lengthening experienced by the particle routes. However, the generalization of this explanation to situations of unsteady transport is not straightforward and, as we will see in the next subsection, it does not provide a complete picture of the jamming mechanisms.

C. Jamming scenarios

A more detailed analysis of jamming reveals that the main cause of the decrease in transport capacity at large values of β is the onset of transport traps which destroy the connectivity of the network. Consequently, particles generated on one connected subnetwork whose destination happens to be on another one will not be able to find their destination. A network may develop one or more traps, each containing a small number of nodes, while the majority of the nodes remain part of a "dynamic giant component." There are also cases when, due to statistical fluctuations, the subnetworks are not disjoint all the time. However, jamming occurs if the time intervals when they can exchange particles are on average too short to allow all particles to find their destination.

We analyzed the behavior of 50 random networks with $N=30$ nodes which are jamming at a load $R=0.65$. Of these, 46 networks exhibited between one and three simple two-node traps involving an "outer" node with degree 1 linked to another "inner" node with a higher degree. The way such a structure leads to trapping is explained in detail in the following paragraph. Five of the 46 networks also had more complicated traps, consisting of a node of degree 2 linked to two higher degree nodes. Finally, one network had a single trap consisting of two nodes of degree 1 connected to a hub, which works basically the same way as the simple trap mentioned above, and three networks were jamming uniformly, without the formation of traps.

A random network with $N=30$ and $\langle d\rangle=6$ which exhibits a simple two-node trap is shown in Fig. 6. The trap is formed by nodes 1 and 5. Even from the beginning, when the queues are short, node 5 receives less traffic than any other nodes, thus having a shorter queue on average. However, this does not help node 1 since it will be more likely to send its particles toward node 5 only to receive them back. This "ping-pong effect" leads to a longer than average queue at node 1. When the queue length of node 1 is larger than the network average, it will be extremely unlikely to receive particles from the rest of the network until its queue length decreases. Thus we have periods of time when the transport between the two parts of the network is interrupted. As long as the queue lengths are small and their fluctuations large compared to their averages, these periods are finite. But as particles start accumulating, the trap becomes essentially permanent. The outer node consistently maintains a shorter queue than any node in the dynamic giant component of the network, while the inner node maintains a longer queue. And since the routing rule (almost) rigorously points toward the neighbor

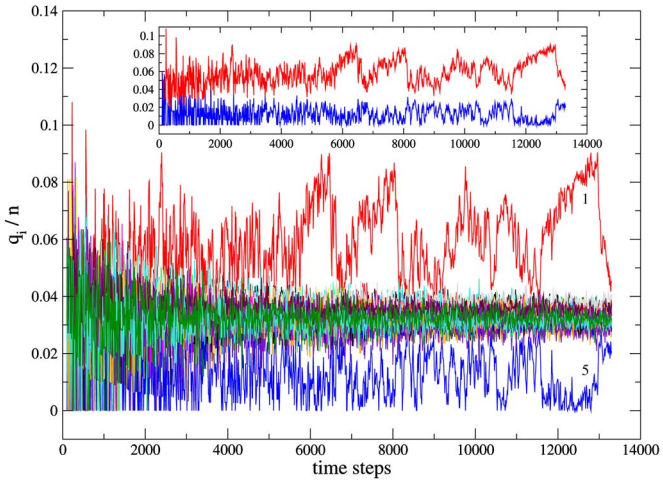

FIG. 7. (Color online) Plot of the fraction of the number of particles q_i/n sitting in the queue of node i versus time for the network in Fig. 6. Each curve has a different color and corresponds to a different queue.

with the shortest queue, this effectively blocks the exchange of particles between the two nodes and the rest of the network since the nodes in the main part of the network to which the inner one is connected will always find other neighbors with shorter queues and send the particles their way. On both sides, particles that cannot find their destination keep accumulating. This eventually increases the destination finding time even for particles that can find their destination, due to the longer waiting periods in queues.

The scenario is illustrated in Figs. 7 and 8, which both pertain to the network in Fig. 6. Figure 7 shows plots of the fractions of the total number of particles sitting in the queues of the various nodes q_i/n as functions of time. As can be seen, the queue length of node 5 is consistently shorter than the queue lengths of all other nodes, while the queue length of node 1 is consistently longer. Plots of the number of particles having a given node as destination s_i versus time are

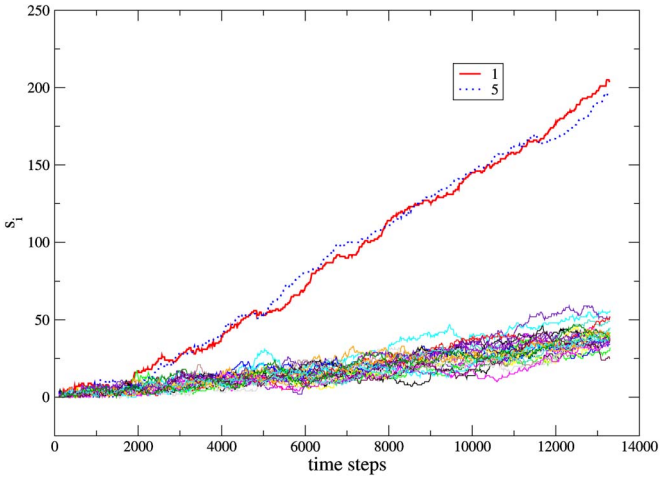

FIG. 8. (Color online) Plot of the number of particles on the network s_i that have a given node i as destination versus time for the network in Fig. 6. Each curve has a different color and corresponds to a different destination node.

shown in Fig. 8. Here we see that as soon as the number of particles on the network increases sufficiently to make the ordering of the queue lengths stable with respect to fluctuations, the number of particles seeking nodes 1 and 5 starts to increase sharply. The number of particles with other destinations increases at a much slower rate since fewer such particles (only those created on nodes 1 or 5) are prevented from reaching their destination. In addition, particles with both origin and destination on the main network experience a lengthening of the destination finding times due to particles seeking nodes 1 or 5 clogging the queues.

As for the five three-node traps, the middle node with degree 2 has a queue shorter than the network average, while both higher degree nodes have longer than average queues. Thus, transport to and from the rest of the network is cut at both ends. Four of these five traps can actually be explained in terms of simple traps, since one of the high degree nodes has a long queue due to being part of a separate two-node trap. This effectively cuts one of the links of the node with degree 2, leading to the formation of another trap through the two-node mechanism mentioned above.

One can imagine other types of structures prone to trap formation, but they are much less likely to occur and were not observed in our simulations. These include traps consisting of more than two nodes of degree 1 connected to the same hub (more likely to occur at lower values of the average degree), as well as traps involving nodes of degrees higher than 1 (more likely to occur at higher average degree). Examples of the latter include a triangle connected only at one vertex with the rest of the network, a chain between two high degree nodes, or a node of degree 3 connected to nodes of much higher degree.

The mechanism of trap formation is different in the case of Barabási-Albert networks. These networks generally consist of three categories of nodes. The first large category includes the hubs, which have high degrees and are very likely to be connected to each other and to lower degree nodes. The second large category includes low degree nodes connected to the hubs but not to each other. Finally, a third smaller category consists of nodes of low or intermediate degree connected to the hubs and to each other. Even with load balancing, the hubs will have longer queues than the nodes with low or intermediate degree. As long as the latter are not connected to each other, particles will be passed back and forth between the two parts of the network: hubs will (almost) always send a particle to one of their low degree neighbors, while these will send it to one of the hubs to which they are connected, not necessarily the one where the particle came from. Thus, transport between any two nodes is possible and the average queues of the various nodes, while not equal, remain relatively close to each other. On the other hand, if two or more low or average degree nodes are connected to each other, from the early stages of jamming they will tend to exchange particles between themselves instead of sending them to the hubs. Thus, their queues will be on average longer than those of the other low or intermediate degree nodes. As particles accumulate on the network and the relative amplitude of the queue length fluctuations

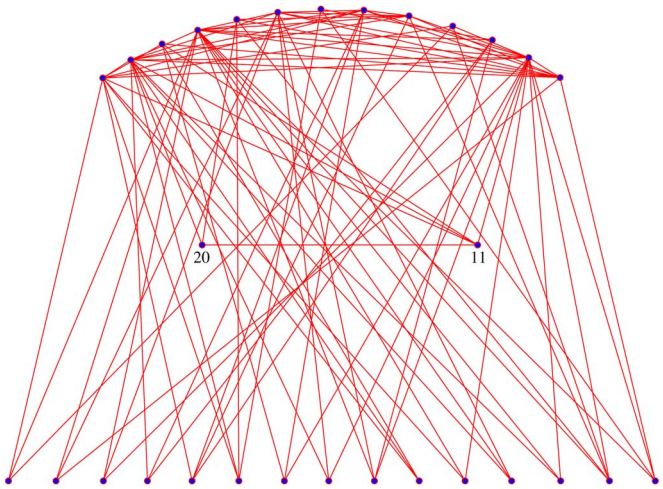

FIG. 9. (Color online) A Barabási-Albert network with $N=30$ and $\langle d\rangle=6$. Nodes 11 and 20 form a "trap" as described in Sec. IV C.

decreases, the network may become locked in a state in which nodes from the third category only exchange particles between themselves while never receiving particles from the hubs. Consequently, both inside and outside the trap, particles that cannot find their destination keep accumulating, which also slows down the particles that can find their destination. An analysis of 50 Barabási-Albert networks with $N=30$ nodes jamming at $R=0.65$ shows that 16 of them are jamming due to trap formation, while 34 are jamming uniformly. A network with $N=30$ and $\langle d\rangle=6$ exhibiting a two-node trap is shown in Fig. 9. Here the trap is formed by nodes 11 and 20. The jamming is illustrated in Fig. 10, which shows plots of the fractions of the total number of particles sitting in the queues of the various nodes q_i/n as functions of time. The time dependence of the number of particles having a given node as destination s_i is very similar to the one in Fig. 8 and is not shown.

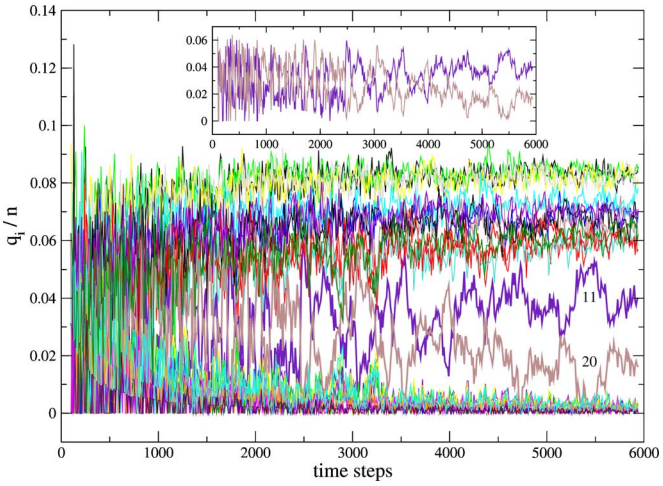

FIG. 10. (Color online) Plot of the fraction of the number of particles q_i/n sitting in the queue of node i versus time for the network in Fig. 9. Each curve has a different color and corresponds to a different queue.

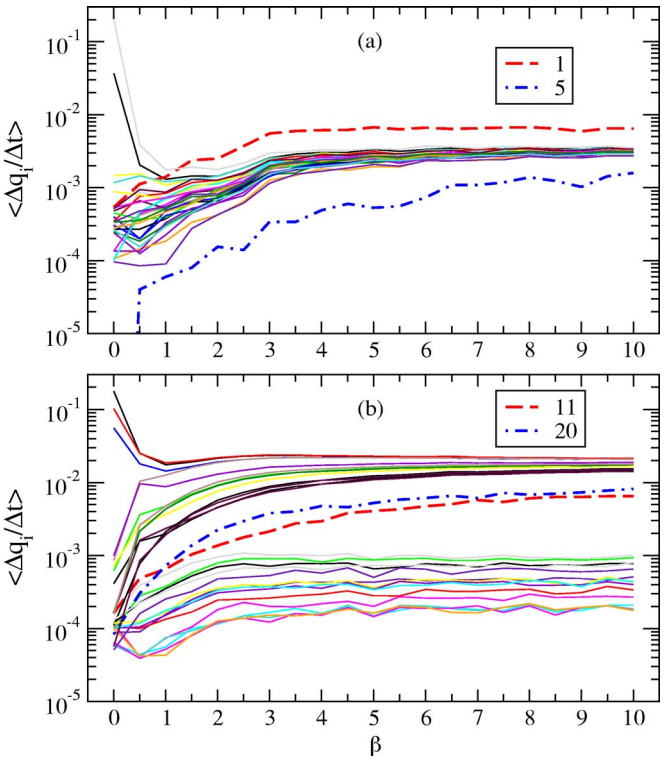

FIG. 11. (Color online) Plot of the average rate of growth of q_i as a function of β for the networks in Figs. 6 (a) and 7 (b).

Now the main reason for the existence of an optimum value of β becomes apparent. By retaining some stochasticity in the routing rule, the onset of the transport traps can be avoided since particles are sufficiently likely to be sent even towards neighbors that have longer queues than others. This is similar to the way simulated annealing or Metropolis Monte Carlo algorithms avoid trapping of the system in metastable states [27,28]. The trapping mechanisms described above also explain why, especially in the case of random networks, the maximum or the transport capacity at optimum β becomes less prominent as the average degree increases since networks become less likely to contain low degree nodes, which are essential for the formation of traps.

Finally, we studied the dependence on the congestion awareness parameter of the average rates of increase of the queue lengths q_i and number of particles seeking a given node s_i for 50 random and 50 Barabási-Albert networks with $N=30$ which jam at any value of R. The averages were taken over 100 transport sessions on each network. We found no evidence of change in the trapping mechanism for a given network as the congestion awareness parameter is varied. Figure 11 shows plots of $\langle \Delta q_i/\Delta t\rangle$ versus β for the networks in Figs. 6 (a) and 7 (b). In the case of the random network, one can see that at low values of β there are other nodes whose queues increase much faster that of node 1. These are the nodes with the highest random walk betweenness on the network. However, at a value of $\beta\approx 1$, which is slightly higher than the critical value, the rate of increase of the queue of node 1 becomes the highest and remains like that as

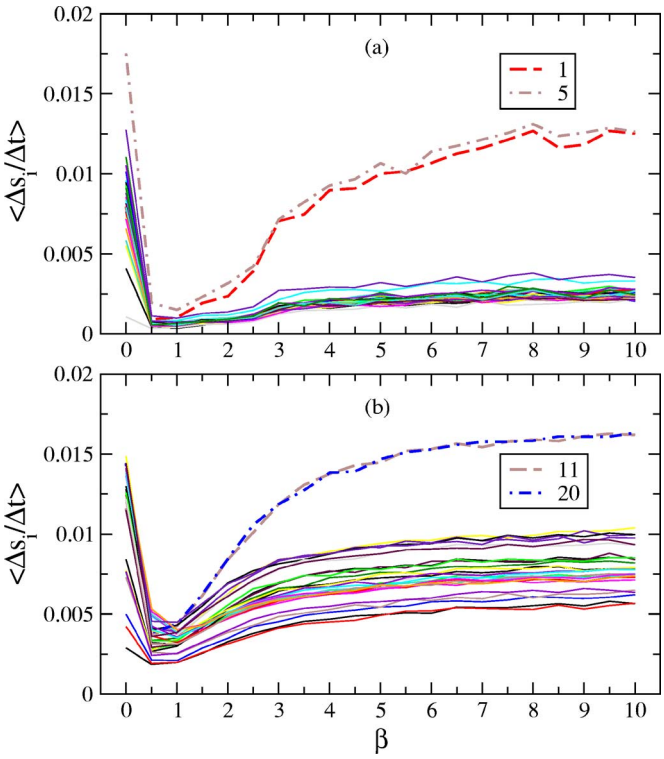

FIG. 12. (Color online) Plot of the average rate of growth of s_i as a function of β for the networks in Figs. 6 (a) and 7 (b).

β increases. Node 5 experiences the lowest rate of increase of the queue length regardless of the value of β, as expected since it is the only node of degree 1 on the network. In the case of the Barabási-Albert network, the two nodes that form the trap have intermediate rates of increase of the queue length regardless of the value of β. The plots of $\langle \Delta s_i/\Delta t \rangle$ versus β presented in Fig. 12 show that, as soon as the critical value of β is exceeded, the nodes that form the traps become more difficult to find than all other nodes and the discrepancy between the rates of increase of the number of particles seeking nodes inside and outside the trap increases significantly with β. Finally, we note that in Fig. 12(a) the rate of increase $\langle \Delta s_5/\Delta t \rangle$ for the outer node is the highest even below the critical value of β. However, this is simply due to its low betweenness and not to the existence of a trap.

Depending on the topology of the network around the critical nodes, traps can occur virtually from the beginning, without the network ever reaching a steady transport state, or they can be triggered by a large statistical fluctuation at a later time. Indeed, the possibility of jamming after an initial period of steady transport is a characteristic of networks operating under (almost) deterministic congestion-aware routing. In all other cases, networks either reach a steady state that lasts indefinitely, or start jamming from the beginning. The connection between jamming start time and the details of network structure seems to be extremely complicated. Furthermore, since simulations can only be run for finite periods of time, we cannot identify all networks that jam due to the formation of transport traps. However, from the point of view of rigid congestion-aware routing it seems safe to distinguish between two types of networks: those that are "structurally fit" for it and able to bear at least as high loads as at intermediate degrees of congestion awareness, and those that are prone to the formation of traps (due to the presence of structural features like those discussed earlier in this section) and end up jamming under lower loads than in the case of a less rigid routing.

V. CONCLUSIONS

We use a simple model to study the behavior of network transport in the case of routing based on local information with various degrees of congestion awareness, ranging from random diffusion to the extreme case of rigid congestion-gradient driven flow. The degree of congestion awareness is controlled by a single scalar parameter. The average transport capacity for networks with a given set of topology and transport parameters is characterized by the critical load under which half of these networks are jamming. At the average connectivity we have used, random networks are more robust against jamming than Barabási-Albert networks with the same number of nodes and average degree. However, in light of the results presented in [19], it is possible that Barabási-Albert (and in general scale-free) networks become more robust once a certain critical value of the average connectivity is exceeded. Regardless of topology, the critical load decreases as the number of nodes increases, which means that the critical load per node approaches zero in the limit of large number of nodes. Consequently, jamming seems to be unavoidable in sufficiently large networks with transport based on local information only.

A somewhat counterintuitive result is the existence of an optimum value of the congestion awareness parameter. Below this value, transport capacity increases with the degree of congestion awareness, and reaches its maximum at the optimum value. The increase is due to the fact that particles are more likely to avoid sitting in the queues of the busiest nodes ("hubs") for extended periods of time. A high degree of congestion awareness does, however, have some unwanted effects which eventually lead to a decrease in transport capacity. We show that this decrease is mainly due to the occurrence of transport traps, which prevent particle exchange between parts of the network. In the case of the random networks, these traps are formed by sets of low degree nodes ("stubs") connected to nodes with higher degrees. The optimum value of the congestion awareness parameter arises from the interplay between the effects that tend to increase the transport capacity and those that tend to decrease it. The overall lower robustness of the Barabási-Albert networks is due to the fact that a power-law (or emergent power-law) distribution of the node degrees allows both more high degree nodes (which are responsible for jamming in the case of low congestion awareness), and more low degree nodes (responsible for jamming in the case of high congestion awareness) compared to the binomial distribution that characterizes the random networks.

The existence of an optimum value of the congestion awareness parameter can be viewed as a result of stochastic interactions, similar to the way that simulated annealing and Metropolis Monte Carlo methods optimize the performance

of computer simulations. In these algorithms, stochastic dynamics allows the system to escape metastable states, whereas in our transport simulations it allows particles to escape traps.

We have also shown that a betweenness centrality measure similar to the one defined in [9] provides an essentially exact description of the statistics of network transport in the case of random diffusion. In addition, we describe an algorithm for the fast computation of this betweenness measure. Furthermore, the betweenness proves to be a useful tool in the analysis of network transport under congestion-aware routing as well.

ACKNOWLEDGMENTS

The authors gratefully acknowledge support from the Air Force Research Laboratory, Information Directorate, under Contract No. FA8750-04-C-0258. B.D., Y.Y., and K.E.B. also acknowledge support from the NSF through Grant No. DMR-0427538. Z.T. was supported by DOE Contract No. W-7405-ENG-36.

[1] M. E. J. Newman, SIAM Rev. **45**, 167 (2003).
[2] D. J. Watts and S. H. Strogatz, Nature (London) **393**, 440 (1998).
[3] P. Holme, Adv. Complex Syst. **6**, 163 (2003).
[4] P. Echenique, J. Gómez-Gardeñes, and Y. Moreno, Phys. Rev. E **70**, 056105 (2004).
[5] P. Echenique, J. Gómez-Gardeñes, and Y. Moreno, Europhys. Lett. **71**, 325 (2005).
[6] M. Anghel, Z. Toroczkai, K. E. Bassler, and G. Korniss, Phys. Rev. Lett. **92**, 058701 (2004).
[7] E. Lopez, S. V. Buldyrev, S. Havlin, and H. E. Stanley, Phys. Rev. Lett. **94**, 248701 (2005).
[8] G. Korniss *et al.*, Phys. Lett. A **350**, 324 (2006).
[9] R. Guimerà, A. Díaz-Guilera, F. Vega-Redondo, A. Cabrales, and A. Arenas, Phys. Rev. Lett. **89**, 248701 (2002).
[10] R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).
[11] A.-L. Barabási, *Linked* (Perseus Publishing, New York, 2002).
[12] M. E. J. Newman, Phys. Rev. E **64**, 016132 (2001).
[13] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).
[14] *Wireless Ad Hoc Networks*, edited by Z. J. Haas *et al.*, special issue of IEEE J. Sel. Areas Commun. **17**(8) (1999).
[15] L. Zhao, Y.-C. Lai, K. Park, and N. Ye, Phys. Rev. E **71**, 026125 (2005).
[16] Z. Toroczkai and K. E. Bassler, Nature (London) **428**, 716 (2004).
[17] Z. Toroczkai, B. Kozma, K. E. Bassler, N. W. Hengartner, and G. Korniss, e-print cond-mat/0408262.
[18] P. Erdős and A. Rényi, Publ. Math. (Debrecen) **6**, 290 (1959); Bull. Internat. Statist. Inst. **38**, 343 (1961).
[19] K. Park, Y.-C. Lai, L. Zhao, and N. Ye, Phys. Rev. E **71**, 065105(R) (2005).
[20] M. E. J. Newman, e-print cond-mat/0309045.
[21] J. D. Noh and H. Rieger, Phys. Rev. Lett. **92**, 118701 (2004).
[22] E. Estrada and J. A. Rodríguez-Velázquez, Phys. Rev. E **71**, 056103 (2005).
[23] V. Latora and M. Marchiori, e-print cond-mat/0402050.
[24] B. K. Singh and N. Gupte, Phys. Rev. E **71**, 055103(R) (2005).
[25] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. (Johns Hopkins University Press, Baltimore, 1996).
[26] O. Allen, *Probability, Statistics and Queuing Theory with Computer Science Application*, 2nd ed. (Academic Press, New York, 1990).
[27] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, Science **220**, 671 (1983).
[28] K. Binder and D. P. Landau, Phys. Rev. B **37**, 1745 (1988).