

**Early appraisal of the fixation probability in directed networks**Valmir C. Barbosa,<sup>1</sup> Raul Donangelo,<sup>2,3</sup> and Sergio R. Souza<sup>2,4</sup><sup>1</sup>*Programa de Engenharia de Sistemas e Computação, COPPE, Universidade Federal do Rio de Janeiro, Caixa Postal 68511, 21941-972 Rio de Janeiro, RJ, Brazil*<sup>2</sup>*Instituto de Física, Universidade Federal do Rio de Janeiro, Caixa Postal 68528, 21941-972 Rio de Janeiro, RJ, Brazil*<sup>3</sup>*Instituto de Física, Facultad de Ingeniería, Universidad de la República, Julio Herrera y Reissig 565, 11.300 Montevideo, Uruguay*<sup>4</sup>*Instituto de Física, Universidade Federal do Rio Grande do Sul, Caixa Postal 15051, 91501-970 Porto Alegre, RS, Brazil*

(Received 18 May 2010; revised manuscript received 11 August 2010; published 22 October 2010)

In evolutionary dynamics, the probability that a mutation spreads through the whole population, having arisen from a single individual, is known as the fixation probability. In general, it is not possible to find the fixation probability analytically given the mutant's fitness and the topological constraints that govern the spread of the mutation, so one resorts to simulations instead. Depending on the topology in use, a great number of evolutionary steps may be needed in each of the simulation events, particularly in those that end with the population containing mutants only. We introduce two techniques to accelerate the determination of the fixation probability. The first one skips all evolutionary steps in which the number of mutants does not change and thereby reduces the number of steps per simulation event considerably. This technique is computationally advantageous for some of the so-called layered networks. The second technique, which is not restricted to layered networks, consists of aborting any simulation event in which the number of mutants has grown beyond a certain threshold value and counting that event as having led to a total spread of the mutation. For advantageous mutations in large populations and regardless of the network's topology, we demonstrate, both analytically and by means of simulations, that using a threshold of about  $[N/(r-1)]^{1/4}$  mutants, where  $N$  is the number of simulation events and  $r$  is the ratio of the mutants' fitness to that of the remainder of the population, leads to an estimate of the fixation probability that deviates in no significant way from that obtained from the full-fledged simulations. We have observed speedups of two orders of magnitude for layered networks with 10 000 nodes.

DOI: [10.1103/PhysRevE.82.046114](https://doi.org/10.1103/PhysRevE.82.046114)

PACS number(s): 89.75.Fb, 87.23.Kg, 02.10.Ox, 02.50.-r

**I. INTRODUCTION**

We consider directed networks in which each node is inhabited by a single individual of a population and whose edges represent the possibilities for an individual's offspring to replace some other individual. Such networks provide the substrate on which the evolution of the population can be studied given the constraints imposed by their structures. In this modality of evolutionary dynamics, known as evolutionary graph dynamics since its introduction in [1], the population evolves in discrete-time steps, each of which involves the fitness-based selection of an individual for reproduction and the use of its offspring to replace one of its out-neighbors in the network. The chief quantity one targets in such studies is the probability that a mutation arising at a randomly chosen individual, henceforth called a mutant, of the otherwise homogeneous population eventually spreads through all the population. This probability is known as the fixation probability.

In the past decade, the study of several other phenomena has been approached from a similar perspective of interacting agents. Such phenomena have included differently constrained forms of the dynamics of evolution [2,3], the spread of epidemics through populations [4], the emergence of cooperation in biological and social systems [5–8], and various others [9–11]. In most cases, what the interacting agents do, driven by either competition or the goal of promoting cooperation, is spread information through the network in order to attempt to influence the states of other agents. In general,

network structure is a major player in affecting the global outcome of such interactions, and this holds to the extent that subtle structural changes can have relevant consequences [1,12–17]. The importance of network structure, in fact, is also central in several other areas, as for example those discussed in [18–20].

The fixation probability is very heavily influenced by the structure of the underlying network as well. In rare cases it is possible to calculate it analytically from both structure and the relative fitnesses of the individuals [1,21–24], but in general one has to resort to simulations of the evolutionary steps. Given the network and the mutant's fitness, the simulation is conducted as a number of independent events, each of which starts by placing the mutant at a randomly chosen node and then carries out the evolutionary steps until either fixation (all nodes contain mutants) or extinction (no node contains a mutant) occurs. The fraction of events ending in fixation is an estimate of the fixation probability. This simulation-based approach to obtaining the fixation probability can be very time consuming not only because many independent events are needed but also because each event can require a significantly large number of steps to converge. This is illustrated in Fig. 1, which suggests two properties of the simulation process. The first is that events ending in extinction usually require substantially fewer steps to conclude than those ending in fixation. The second is that the number of steps required for fixation varies widely with network topology.

The various network topologies used in Fig. 1 recur throughout the paper, so we pause momentarily to introduce them. We emphasize, though, that our purpose in the present

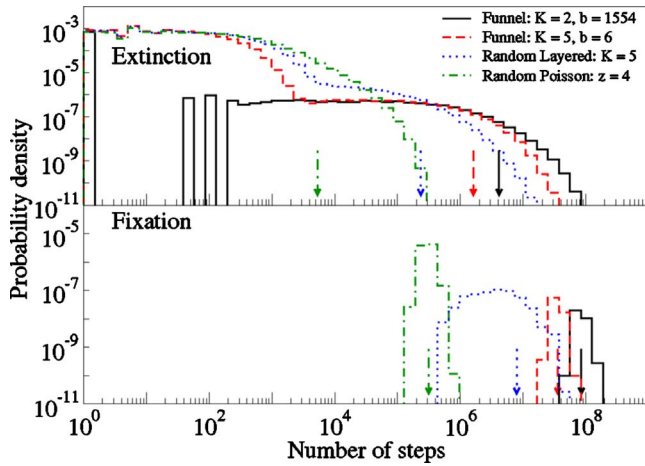


FIG. 1. (Color online) Probability densities associated with the number of steps required for extinction and fixation of a mutant 10% fitter than the remainder of the population. Data are log-binned to the base 1.5 and represent averages over  $10^6$  events for each  $K$ -funnel,  $10^7$  events for each of the two random networks ( $10^4$  events for each of  $10^3$  graphs in each case; in the case of the random network with a Poisson-distributed number of out-neighbors, each of these graphs has at least 95% of the nodes in the GSCC). Arrows indicate the means. All the networks have 1555 nodes.

work is not to dwell on the properties of these topologies but rather on how to accelerate the determination of the fixation probability for any given network. The topologies in question are the  $K$ -funnel [1], a random generalization thereof [25], and the directed variant of the Erdős-Rényi random graphs [26] discussed in [27,28]:

(a) For  $b, K > 1$  integers, the  $K$ -funnel of base  $b$  (a branching parameter) has node set partitioned into  $K$  subsets, called layers and numbered  $k=0, 1, \dots, K-1$ , such that there are  $b^k$  nodes in layer  $k$ . The  $K$ -funnel, therefore, has  $(b^K - 1)/(b - 1)$  nodes. For  $k=1, 2, \dots, K-1$ , an edge exists directed from each node in layer  $k$  to each node in layer  $k-1$ . Additionally, an edge exists directed from the single node in layer 0 to each of the nodes in layer  $K-1$ . An illustration is given in Fig. 2 for  $K=b=3$ .

(b) The  $K$ -funnel is generalized by any network in which nodes occupy layers and edges exist only from all nodes in a layer to all nodes in the next according to some cyclic arrangement of the layers. The generalization we use [25], labeled random layered in Fig. 1, is constructed randomly by first placing one node at each layer and then adding one extra node at a time to a layer that is chosen with probability proportional to  $(K-d)^a$ . Here,  $d$  is the distance to the layer in question from the layer whose number of nodes is the first peak one finds while moving upstream from it, and the strength parameter  $a \geq 1$  regulates the probabilistic bias.

(c) The other random networks we use [27,28], labeled random Poisson in Fig. 1, depart from the layered structure altogether and are built as follows. Given  $n$  nodes and the desired expected number of out-neighbors, denoted by  $z$ , every ordered pair of distinct nodes is connected by an edge directed from the first node to the second in the pair with probability  $z/(n-1)$ . It follows that a randomly chosen node has a number of out-neighbors given by the Poisson distri-

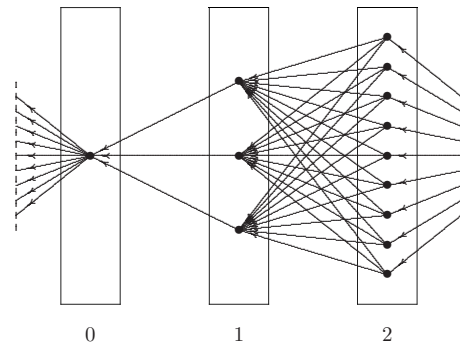


FIG. 2. The 3-funnel for branching parameter  $b=3$ . Each rectangular box corresponds to a layer. Edges connect back to the nodes in layer 2 from the single node in layer 0.

bution of mean  $z$  and that, if  $z > 1$ , a giant strongly connected component (GSCC) exists with high probability. The simulation of the evolutionary dynamics on such a network is confined to its GSCC.

We also remark, before proceeding, that some of our own recent work related to the fixation probability has been strongly influenced by the computational difficulties associated with estimating it. For example, in [25] we set out to create a randomized growing mechanism for layered networks that would result in fitness amplifiers in the sense in [1]. The mechanism we created gives rise to the  $K$ -layer random networks introduced above and, indeed, we were able to demonstrate the desired amplification effect for 1000 nodes. However, we fell short of demonstrating to full completion that the same holds for significantly larger networks, due mainly to the very large times required to compute the fixation probabilities of the thousands of candidate networks.

Our central concern in this paper is the devising of simulation strategies that can make the calculation of the fixation probability substantially faster while maintaining accuracy or reducing it only imperceptibly. We proceed along two tracks. The first is targeted at eliminating the evolutionary steps at which no change is effected to the number of mutants in the network. This occurs whenever the simulation prescribes that a mutant's offspring is to replace another mutant or that a nonmutant's offspring is to replace another nonmutant. Our results in this track are better suited to the layered networks introduced above; we demonstrate them for the  $K$ -funnel. The second track we pursue builds on the realization that, if in general it takes a lot more steps for fixation to occur than for extinction, then detecting early in the course of a simulation event that fixation is highly likely to occur can be used as a surrogate to the eventual detection of fixation and thereby reduce the number of necessary steps. We have found out that, nearly regardless of the network for a large number of nodes, there exists a threshold number of mutants beyond which fixation is practically guaranteed to take place. We give results for a wide variety of networks.

We organize the remainder of the paper in the following manner. In Sec. II we briefly review the key notions related to the fixation probability. Then, in Secs. III and IV we pursue the two tracks outlined above for computing the fixation

probability more efficiently for layered networks and for unrestricted networks, respectively. We conclude in Sec. V.

**II. FIXATION PROBABILITY**

Let  $V$  be a population of  $n$  individuals and, for  $i \in V$ , let  $O_i \subseteq V \setminus \{i\}$  be the set of individuals that an offspring of  $i$  can replace during the evolution of  $V$ . Let also  $I_i$  be the set of individuals whose offspring can replace  $i$ . These give rise to a directed graph, called  $D$ , whose set of nodes is  $V$  and whose set of edges, denoted by  $E$ , contains the edge  $(i, j)$  if and only if  $j \in O_i$  (equivalently,  $i \in I_j$ ). Each individual  $i$  has a fitness  $f_i > 0$  associated with it, and similarly to each edge  $(i, j)$  there corresponds a probability  $w_{ij}$  such that  $\sum_{j \in O_i} w_{ij} = 1$  for all  $i \in V$ . The dynamics of evolution that we consider occurs in a sequence of steps. At each step, an individual  $i$  is chosen with probability proportional to  $f_i$ , then another individual  $j \in O_i$  is chosen with probability  $w_{ij}$ , and finally  $j$  is replaced by an offspring of  $i$  having the same fitness  $f_i$ .

The fixation probability of graph  $D$ , denoted by  $\rho$ , is the probability that a mutation spreads through all of population  $V$  given that it arises from one single individual and that all individuals in the remainder of the population have the same fitness. The value of  $\rho$  depends on the structure of  $D$  and on the ratio  $r$  of the mutant's fitness to that of the other individuals. Moreover, it is the relationship between  $\rho$  and  $r$  that determines whether evolution is driven primarily by natural selection or by random drift: essentially, natural selection predominates when  $\rho$  and  $r$  are highly correlated, random drift otherwise. Note in this context that, if  $D$  is not strongly connected (i.e., there exist nodes  $i$  and  $j$  such that no directed path leads from  $i$  to  $j$ ), then  $\rho > 0$  if and only if there exists an individual from which all others are reachable. This may cause random drift to be the main driver of evolution, so henceforth we assume that  $D$  is strongly connected (thus  $\rho > 0$  necessarily).

This type of evolutionary dynamics can be described by a discrete-time Markov chain of states  $0, 1, \dots, n$ , each representing a possible number of mutants in graph  $D$ . In this chain, states 0 and  $n$  are absorbing and all others are transient. If  $s$  is a transient state, then from  $s$  it is possible to move to state  $s+1$  (with probability  $p_s$ ), to state  $s-1$  (with probability  $q_s$ ), or to remain at state  $s$  (with probability  $1-p_s-q_s$ ). The use of this chain to describe the dynamics stems directly from the fact that, at each step, the number of mutants in  $D$  changes by at most 1. In principle, however, it is possible to zoom in on any transient state  $s$  and replace it with a collection of  $\binom{n}{s}$  states, each representing one of the possibilities for the location of the  $s$  mutants in  $D$ . This would be of no use in what follows, so we contemplate this possibility no further.

If we denote by  $P_{n+1}(n|s)$  the probability that, having started at state  $s$ , this  $(n+1)$ -state system eventually enters state  $n$ , then it is well known that

$$P_{n+1}(n|s) = \frac{1 + \sum_{u=1}^{s-1} \prod_{v=1}^u q_v/p_v}{1 + \sum_{u=1}^{n-1} \prod_{v=1}^u q_v/p_v} \tag{1}$$

(cf. [29,30]). Note, with regard to this expression, that the subscript  $n+1$ , which functions as an indicator of system size, is somewhat redundant at this point. The reason why we adopt it is that in Sec. IV the expression is used to combine systems of different sizes and then the subscript becomes essential for disambiguation. Note also that, in a context wider than the one we are dealing with in this paper, Eq. (1) can be interpreted from the perspective of a one-dimensional random walk on the real line. This random walk is restricted to the integers in the interval  $[0, n]$  and, having started at the integer  $s > 0$  in this interval, eventually ends at  $n$  with the probability given by the equation. In our context of evolutionary dynamics, it is a consequence of Eq. (1) that the fixation probability  $\rho$ , which is given by  $P_{n+1}(n|1)$ , is such that

$$\rho = \frac{1}{1 + \sum_{u=1}^{n-1} \prod_{v=1}^u q_v/p_v} \tag{2}$$

When the probabilities  $w_{ij}$  are such that  $\sum_{j \in I_i} w_{ji} = 1$  for all  $i \in V$  (i.e., not only do the probabilities associated with the outgoing edges of  $i$  sum up to 1 but also those of the incoming edges), and only then, the isothermal theorem in [1,31] establishes that  $p_v/q_v = r$  for all  $v \in \{1, 2, \dots, n-1\}$ . In this case, and assuming that  $r \neq 1$  (i.e., the mutation is either advantageous or disadvantageous but not neutral), it follows from Eq. (2) that the fixation probability, now denoted by  $\rho_1$ , is

$$\rho_1 = \frac{1 - 1/r}{1 - 1/r^n} \tag{3}$$

This includes the Moran process [2], in which graph  $D$  is such that  $O_i = V \setminus \{i\}$  for all  $i \in V$  and  $w_{ij} = 1/|O_i|$  for all  $(i, j) \in E$  (we use  $|X|$  to denote the cardinality of set  $X$ ). It also includes the more general case in which the  $w_{ij}$ 's are thus constrained and, moreover, all nodes have as many incoming edges as outgoing edges, provided this number is the same for all nodes.

Other noteworthy cases for which  $w_{ij} = 1/|O_i|$  for all  $(i, j) \in E$  are the directed graphs that in [1] are shown to be fitness amplifiers with respect to the expression for the fixation probability  $\rho_1$ . The isothermal theorem does not in general apply to these graphs but, for them, substituting  $r^K$  for  $r$  in Eq. (3) yields  $\rho = \rho_K$  as  $n \rightarrow \infty$ , where

$$\rho_K = \frac{1 - 1/r^K}{1 - 1/r^{Kn}} \tag{4}$$

This is the case, for example, of the  $K$ -funnel. So, even though it has an artificially contrived structure, the  $K$ -funnel remains a centrally important graph because of its ability to amplify fitness arbitrarily with respect to the Moran process.

This property stems essentially from the role played by the branching parameter  $b$  of the  $K$ -funnel. Even at its smallest value ( $b=2$ ), it ensures that layer  $K-1$  has more nodes than all the others put together. Consequently, the first mutant arising in the network is likely to appear in this layer and is also likely to be long lived (only a nonmutant in layer 0 can destroy it, but this is unlikely) and perpetuate the mutation through its offspring.

Moreover, as we explain in Sec. IV, the  $K$ -funnel has provided an indispensable reference against which to detect which of the randomly grown layered networks in [25] have the potential to be fitness amplifiers as well. The latter are the only grown networks we know of that have the fitness-amplification property. In our opinion, the fact that they emerge out of random growth decisions makes them (and also the  $K$ -funnel that guided their discovery) remarkable structures.

### III. FIXATION IN LAYERED NETWORKS

We henceforth use  $w_{ij}=1/|O_i|$  for all  $(i,j) \in E$  exclusively and first present material that applies to any graph  $D$ , layered or otherwise. For  $t \geq 0$  an integer, let  $E(t) \subset E$  be the set of edges  $(i,j)$  such that, at the  $t$ th evolutionary step, node  $i$  is a mutant but node  $j$  is not. If  $P^+(t)$  denotes the probability that, at step  $t+1$ , the number of mutants in  $D$  increases (necessarily by 1), then it can be calculated by summing up the probabilities that this happens at each of the edges in  $E(t)$ . For edge  $(i,j) \in E(t)$ , the replacement of nonmutant  $j$  by an offspring of mutant  $i$  occurs if and only if  $i$  is chosen when fitness-based selection is performed and, moreover,  $j$  is chosen given that  $i$  has been. The former happens with probability  $r/[M(t)r+n-M(t)]$ , where  $M(t)$  is the number of mutants at step  $t$  [so  $M(0)=1$ ], the latter with probability  $w_{ij}$ . We then have

$$P^+(t) = \sum_{(i,j) \in E(t)} \frac{r|O_i|}{n + M(t)(r-1)}. \quad (5)$$

Letting  $m_i(t)$  be an indicator of whether node  $i$  is a mutant at step  $t$  [i.e.,  $m_i(t)=1$  in the affirmative case and 0 otherwise], this expression can be rewritten as a sum over all edges in  $E$ . To do this, we simply recognize that

$$m_i(t)[1 - m_j(t)] = \begin{cases} 0 & \text{if } (i,j) \in E \setminus E(t) \\ 1 & \text{if } (i,j) \in E(t), \end{cases} \quad (6)$$

and then

$$\begin{aligned} P^+(t) &= \frac{r}{n + M(t)(r-1)} \sum_{(i,j) \in E} \frac{m_i(t)[1 - m_j(t)]}{|O_i|} \\ &= \frac{r}{n + M(t)(r-1)} \left[ \sum_{(i,j) \in E} \frac{m_i(t)}{|O_i|} - \sum_{(i,j) \in E} \frac{m_i(t)m_j(t)}{|O_i|} \right] \\ &= \frac{r}{n + M(t)(r-1)} \left[ \sum_{i \in V} m_i(t) \sum_{j \in O_i} \frac{1}{|O_i|} \right. \\ &\quad \left. - \sum_{(i,j) \in E} \frac{m_i(t)m_j(t)}{|O_i|} \right]. \end{aligned} \quad (7)$$

Using  $\sum_{j \in O_i} (1/|O_i|) = 1$  for all  $i \in V$  and  $\sum_{i \in V} m_i(t) = M(t)$  yields, finally,

$$P^+(t) = \frac{r}{n + M(t)(r-1)} \left[ M(t) - \sum_{(i,j) \in E} \frac{m_i(t)m_j(t)}{|O_i|} \right]. \quad (8)$$

Similarly, the probability  $P^-(t)$  that the number of mutants decreases (necessarily by 1) at step  $t+1$  is

$$P^-(t) = \frac{1}{n + M(t)(r-1)} \left[ \sum_{(i,j) \in E} \frac{m_j(t)}{|O_i|} - \sum_{(i,j) \in E} \frac{m_i(t)m_j(t)}{|O_i|} \right]. \quad (9)$$

The number of mutants remains unchanged at step  $t+1$  with probability  $1 - P^+(t) - P^-(t)$ .

These expressions for  $P^+(t)$  and  $P^-(t)$  hold regardless of whether the isothermal theorem applies to graph  $D$  or not. When it does apply, the two expressions allow the evolutionary dynamics to be observed from an interesting perspective, since in this case  $D$  is such that  $\sum_{(i,j) \in E} m_j(t)/|O_i| = \sum_{j \in V} m_j(t) \sum_{i \in I_j} 1/|O_i| = \sum_{j \in V} m_j(t) = M(t)$ , and therefore  $P^+(t)/P^-(t) = r = p_s/q_s$  for state  $s \in \{1, 2, \dots, n-1\}$  such that the number of mutants  $M(t) = s$ . This, essentially, is the argument behind the isothermal theorem.

Moreover, there is insight to be gained from the two expressions even if the isothermal theorem does not hold. For fitness ratio  $r$  sufficiently close to 1, what makes  $P^+(t)$  and  $P^-(t)$  differ from each other is the balance between  $M(t)$  and  $\sum_{(i,j) \in E} m_j(t)/|O_i|$ , where the former does not depend on the topology of  $D$  (given  $t$ ) while the latter does. Thus, for example, if we consider the  $K$ -funnel and let  $M_k(t)$  be the number of mutants in layer  $k$  at step  $t$ , we have

$$\begin{aligned} \sum_{(i,j) \in E} \frac{m_j(t)}{|O_i|} &= \sum_{k=0}^{K-2} \frac{M_k(t)b^{k+1}}{b^k} + \frac{M_{K-1}(t)}{b^{K-1}} \\ &= b[M(t) - M_{K-1}(t)] + \frac{M_{K-1}(t)}{b^{K-1}}. \end{aligned} \quad (10)$$

Readily, maintaining a relatively high value for the ratio  $P^+(t)/P^-(t)$  (a strong forward bias) in the case of the  $K$ -funnel depends crucially on how close  $M(t)$  and  $M_{K-1}(t)$  are to each other, i.e., on how close the largest layer is to containing a significant fraction of the  $M(t)$  mutants. The fact that  $P^+(t)/P^-(t) \rightarrow r^K$  as  $n \rightarrow \infty$  indicates that the topology of the  $K$ -funnel is in fact successful at maintaining the necessary distribution of mutants.

Equations (8) and (9) are also useful in that they provide an alternative mechanism for simulating the evolutionary dynamics. Instead of repeatedly choosing  $i$ , then  $j \in O_i$  to receive  $i$ 's offspring, until the mutation either spreads through the whole of graph  $D$  or dies out, we use the two equations to decide, at each step  $t$ , whether the number of mutants will increase, decrease, or remain the same at step  $t+1$ . In the former two cases, and only in these cases, we choose the nodes to be involved, create or destroy a mutant as the case may be, then compute  $P^+(t+1)$  and  $P^-(t+1)$ . By doing so, all steps in which no mutant is created or destroyed are skipped. Of course, in order for this alternative to be computationally



attractive the suppression of these steps has to compensate for the additional effort to calculate the probabilities at every step that is not suppressed.

While in the general case this will not be so, for layered networks like the  $K$ -funnel, in which all nodes in the same layer are topologically identical to one another (they all have the same in- and out-neighbors), the simulation can be conducted by keeping track only of the number of mutants in each layer and the alternative becomes attractive. We then consider the generalization of the  $K$ -funnel obtained by letting layer  $k$  have any number  $n_k > 0$  of nodes. In this case, we can rewrite Eqs. (8) and (9) by decomposing each of  $P^+(t)$  and  $P^-(t)$  into  $K$  summands, each related to a pair of subsequent layers. Thus, we obtain  $P^+(t) = \sum_{k=0}^{K-1} P_k^+(t)$  and  $P^-(t) = \sum_{k=0}^{K-1} P_k^-(t)$ , with

$$P_k^+(t) = \frac{rM_{k+1}(t)[n_k - M_k(t)]}{n_k[n + M(t)(r-1)]} \quad (11)$$

and

$$P_k^-(t) = \frac{M_k(t)[n_{k+1} - M_{k+1}(t)]}{n_k[n + M(t)(r-1)]}, \quad (12)$$

where  $M_k(t)$  is as for the  $K$ -funnel. In these expressions and henceforth, layer numbers are incremented or decremented modulo  $K$  (the number of layers).

For layered networks such as these, once it has been decided that a mutant is to be created or destroyed at step  $t+1$ , the layer at which this is to happen is  $k$  with probability proportional to  $P_k^+(t)$ , in the case of creation, or  $P_k^-(t)$ , in the case of destruction [consistently with this, notice that  $P_k^+(t)=0$  if  $M_{k+1}(t)=0$  or  $M_k(t)=n_k$  and that  $P_k^-(t)=0$  if  $M_k(t)=0$  or  $M_{k+1}(t)=n_{k+1}$ ]. Once  $M_k(t+1)$  has been updated from  $M_k(t)$ , it follows from Eqs. (11) and (12) that only  $P_k^+(t+1)$  and  $P_{k-1}^+(t+1)$ , or  $P_k^-(t+1)$  and  $P_{k-1}^-(t+1)$ , need to be calculated. This is so because, although  $M(t+1)$  is updated from  $M(t)$  as well, the only effect this has is to alter the factor in the denominator of Eqs. (11) and (12) that is common to all layers. As a consequence, the probabilities corresponding to any layer other than  $k$  or  $k-1$  need not be calculated.

Computational results on the  $K$ -funnel are given in Fig. 3 for a variety of  $n$  and  $K$  values, where we show the speedup afforded by the use of Eqs. (11) and (12) to compute the fixation probability. This speedup is defined as the ratio of how much processor time is needed to determine the fixation probability when the two equations are not used to the processor time that is needed when they are. It indicates, in all cases examined, a reduction to less than half the time required by the simulation that goes through all the evolutionary steps. We note, however, that the approach we discuss in Sec. IV allows for much more significant speedups.

#### IV. FIXATION IN ARBITRARY NETWORKS

In the absence of the computational facilitation provided by layered networks, which, as we have seen, allows the fixation probability to be computed more efficiently by skipping all steps of the simulation in which no mutant is created

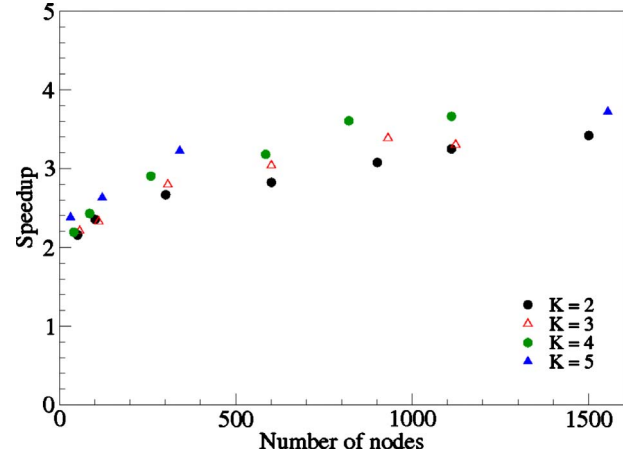


FIG. 3. (Color online) Speedups resulting from the use of Eqs. (11) and (12) on the  $K$ -funnel with fitness ratio  $r=1.1$ . Data are averages over at least  $10^4$  events.

or destroyed, for an unrestricted topology we turn to the alternative strategy of attempting an early stop of each simulation event based on how many mutants there are. The central question is whether there exists a threshold number of mutants which, once crossed from below, ensures that fixation is bound to occur with probability as close to 1 as one wishes. We provide an affirmative answer in what follows.

The probability that the mutation eventually dies out, given that  $s$  mutants are originally present, equals  $1 - Q_{n+1}(0|s)$ . We denote it by  $Q_{n+1}(0|s)$ , and it follows from Eq. (1) that

$$Q_{n+1}(0|s) = \frac{\sum_{u=s}^{n-1} \prod_{v=1}^u q_v/p_v}{1 + \sum_{u=1}^{n-1} \prod_{v=1}^u q_v/p_v}. \quad (13)$$

We are interested in the probability that, conditioned on the fact that extinction does actually occur, the number of mutants eventually grows from the initial  $s$  to some fixed value  $M \in \{s, s+1, \dots, n-1\}$  but does not surpass it. We denote this probability by  $Q_{n+1}^M(0|s)$  and remark that, should its dependency with  $M$  be known, we would immediately be able to discover the desired threshold for the number of mutants by adopting  $s=1$  and specifying a lower bound on the probability. In other words, we would discover the threshold, call it  $M^*$ , by specifying  $Q^*$  such that  $Q_{n+1}^M(0|1) \geq Q^*$  for all  $M \leq M^*$ . This follows from the intuitive expectation that  $Q_{n+1}^M(0|s)$  is to decrease as  $M$  grows for sufficiently large  $M$ .

From its definition as a conditional probability,  $Q_{n+1}^M(0|s)$  is given by  $AB/Q_{n+1}(0|s)$ , where  $A$  is the probability that the number of mutants in the network eventually increases from  $s$  to  $M$  and  $B$  is the probability that, given that it has  $M$  mutants, the system eventually returns to state 0 without ever increasing its number of mutants beyond  $M$ . We calculate the values of  $A$  and  $B$  by resorting to discrete-time Markov chains entirely analogous to the one we have been using but now having reduced numbers of states. The first of these

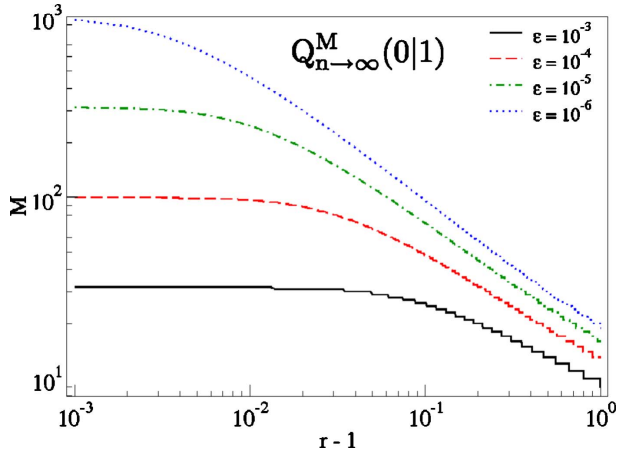


FIG. 4. (Color online) Contour plots of the probability that the population acquires a maximum of  $M$  mutants, given that it eventually dies out [Eq. (15)], as a function of  $M$  and the fitness ratio  $r$  for one single initial mutant ( $s=1$ ) and a number of nodes  $n \rightarrow \infty$ . Contour levels are denoted by  $\epsilon$ .

chains has states  $0, 1, \dots, M$ , of which 0 and  $M$  are absorbing while all else remains unchanged. We set  $A$  to the probability that the system gets absorbed into state  $M$  having started at state  $s$ , that is,  $A = P_{M+1}(M|s)$ . The second chain has states  $0, 1, \dots, M+1$ , with 0 and  $M+1$  as the absorbing states and everything else unchanged. We set  $B$  to the probability that absorption occurs at state 0 once the system is started at state  $M$ , that is,  $B = Q_{M+2}(0|M)$ . We then obtain

$$Q_{n+1}^M(0|s) = \frac{P_{M+1}(M|s)Q_{M+2}(0|M)}{Q_{n+1}(0|s)}. \quad (14)$$

Closed-form expressions are, in general, not known for Eq. (1) or (13), so we assume that it suffices to consider the case in which  $p_v/q_v=r$  for all transient states  $v$ . That is, we proceed with deriving an exact expression for the isothermal case and use the same expression as an approximation in all other cases. In the isothermal case and for  $r \neq 1$ , Eq. (14) yields

$$Q_{n+1}^M(0|s) = \frac{(1-1/r^s)(1-1/r)}{(1-1/r^M)(1-1/r^{M+1})r^M} \frac{1-1/r^n}{1/r-1/r^n}. \quad (15)$$

Readily,  $Q_{n+1}^M(0|s)$  becomes independent of the number of nodes  $n$  as  $n$  grows, regardless of whether the fitness ratio  $r > 1$  or  $r < 1$  (but note that the two limits differ). For  $n \rightarrow \infty$ , in Fig. 4 we show plots of  $Q_{n+1}^M(0|1) = \epsilon$  for different values of  $\epsilon$  and  $1 < r \leq 2$ . As expected, for fixed  $\epsilon$  the value of  $M$  increases with decreasing  $r$  in this range, and somewhat counterintuitively we see that the rate of increase is ever smaller as  $r$  approaches 1. However, this is easily confirmed as we realize that, as in the figure, the limit  $Q$  of  $Q_{n+1}^M(0|1)$  as  $r \rightarrow 1$  is such that  $M \approx \sqrt{1/Q}$ .

Computational results on the distribution of  $M$ , the maximum number of mutants achieved when the dynamics ends in extinction, are given in Figs. 5–7 for a variety of topologies. Figure 5 refers to the  $K$ -funnel; Fig. 6 refers to random networks with a Poisson-distributed number of out-neighbors, and Fig. 7 refers to a selection of networks that

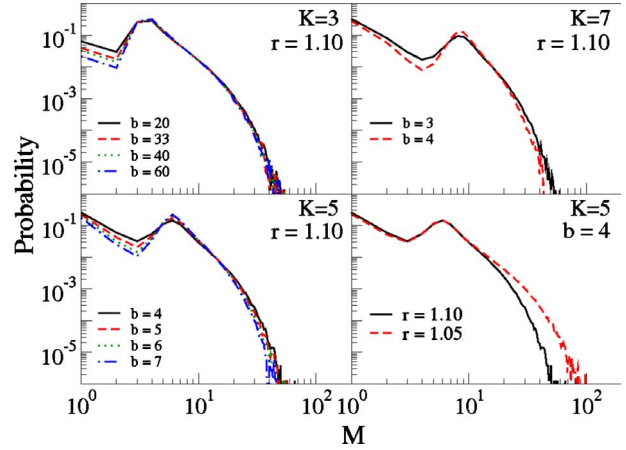


FIG. 5. (Color online) Probability distributions of the maximum number of mutants that occur when the mutation eventually dies out ( $M$ ) for the  $K$ -funnel with assorted values for  $K$ , branching parameter  $b$ , and fitness ratio  $r$ . Data are averages over  $10^6$  events.

includes an instance of each of these two types and also the unidirectional ring. The latter figure also includes plots of the analytical prediction given in Eq. (15). In this respect, notice that the prediction that corresponds to the value used in the simulations for the fitness ratio  $r$  matches the data for the unidirectional ring perfectly. This, of course, is consistent with the fact that, in this case, the requirements of the isothermal theorem are satisfied and therefore the prediction by Eq. (15) is exact. In all other cases the prediction is only approximate. Nevertheless, in the inset of Fig. 7 we demonstrate that, for the random network in use, the prediction of Eq. (15) with a slightly lower value of  $r$  is also a perfect match.

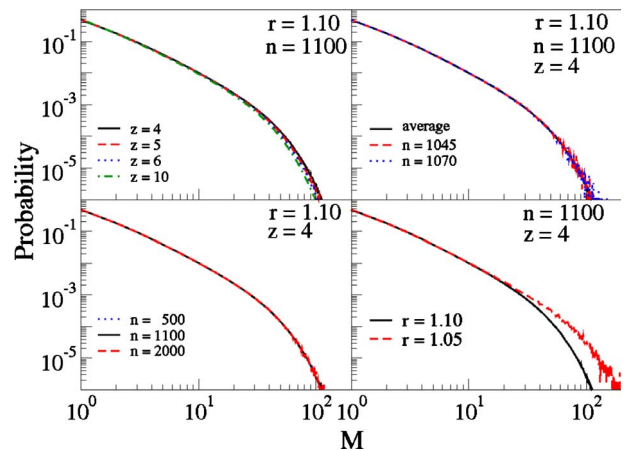


FIG. 6. (Color online) Probability distributions of the maximum number of mutants that occur when the mutation eventually dies out ( $M$ ) for random networks with a Poisson-distributed number of out-neighbors. Data are averages over  $10^7$  events ( $10^4$  events for each of  $10^3$  graphs with at least  $0.95n$  nodes in the GSCC). The plot labeled average refers to all data for fitness ratio  $r=1.1$ , number of nodes  $n=1100$ , and expected number of out-neighbors  $z=4$  in all panels. This is compared in the same panel to the results corresponding to a single graph whose GSCC has 1045 or 1070 nodes, for which  $10^6$  events have been run.

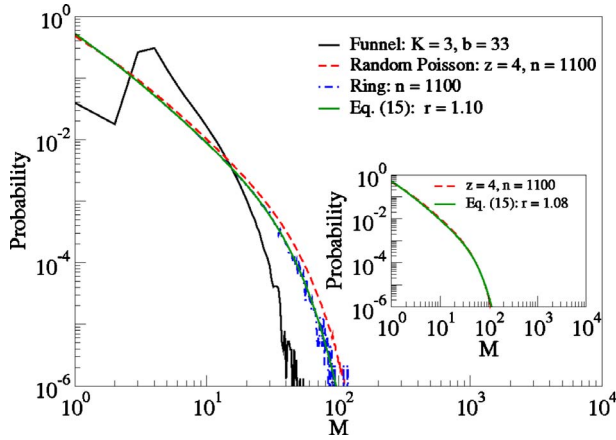


FIG. 7. (Color online) Probability distributions of the maximum number of mutants that occur when the mutation eventually dies out ( $M$ ) for assorted topologies with fitness ratio  $r=1.1$ . Data are averages over  $10^6$  events for the  $K$ -funnel and the unidirectional ring,  $10^7$  events for the random network ( $10^4$  events for each of  $10^3$  graphs with at least  $0.95n$  nodes in the GSCC).

Notice also, in all three figures, that for  $M=100$  no probability is above  $10^{-5}$ . This means that, during the simulations, of all events that ended in extinction, no more than one out of  $10^5$  achieved more than 100 mutants. As a consequence, should we use  $M^*=100$  (corresponding roughly to  $Q^*=10^{-6}$  given that there were at least  $10^6$  events) and count as an event ending in fixation any event achieving more than  $M^*$  mutants, we would be introducing a deviation of no more than  $10^{-5}$  with respect to the actual value of the fixation probability. But the fixation probability obtained by full simulations of the evolutionary dynamics is itself subject to the so-called standard error that is inherent to any Monte Carlo simulation. If  $\hat{\rho}$  denotes the fixation probability calculated after  $N$  events, then the standard error is the standard deviation of the 0's (extinctions) and 1's (fixations) accumulated along the events divided by  $\sqrt{N}$ , that is,  $\sqrt{\hat{\rho}(1-\hat{\rho})/N}$ . This function of  $\hat{\rho}$  is plotted in Fig. 8 for different values of  $N$ , along with a flat line for the constant  $10^{-5}$ . Clearly, the

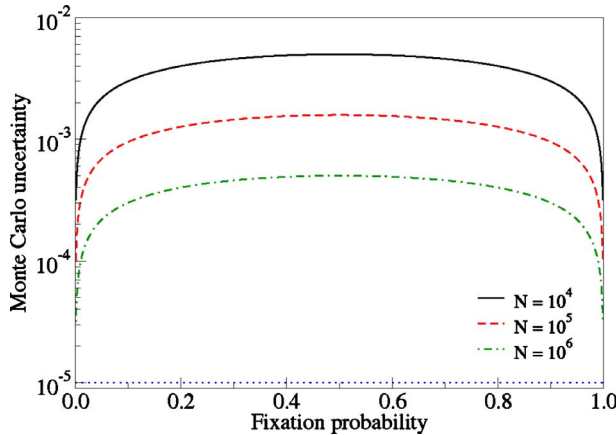


FIG. 8. (Color online) Inherent uncertainty associated with the Monte Carlo computations of the fixation probability through  $N$  simulation events, as given by the standard error of the fixation probability estimate  $\hat{\rho}$ .

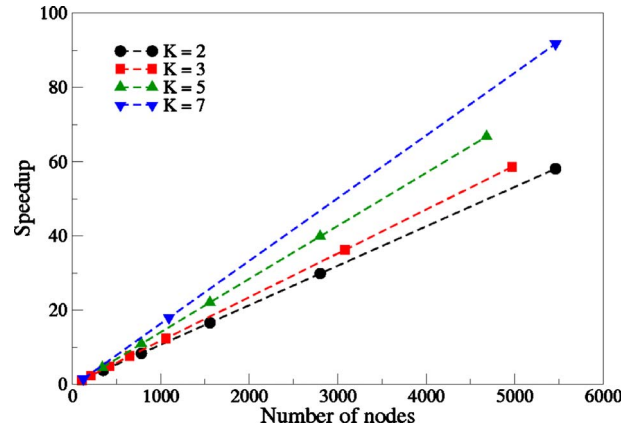


FIG. 9. (Color online) Speedups resulting from the use of the  $M^*=100$  threshold on the  $K$ -funnel with fitness ratio  $r=1.1$ . Data are averages over  $10^6$  events.

additional deviation introduced by the use of  $M^*=100$  is negligible when compared to the standard error of a Monte Carlo simulation.

The above expression for the standard error can also be used to quickly estimate an upper bound on  $M^*$ , in the isothermal case, given the fitness ratio  $r > 1$  and  $N$ , the number of events in the Monte Carlo simulation. To do this, we recall from Fig. 4 that the error incurred when deciding that an event reaching a number  $M$  of mutants will fixate is  $Q \leq 1/M^2$  (nearing equality as  $r$  approaches 1). Since there is not much to be gained in making this error significantly smaller than the one implicit in a Monte Carlo simulation, we can determine an upper bound on  $M^*$  from the condition  $1/(M^*)^2 \geq \sqrt{\hat{\rho}(1-\hat{\rho})}/N$ . For a large population size  $n$ , we have  $\hat{\rho} \approx (1-1/r)/(1-1/r^n) \approx 1-1/r$ , which leads to the simple estimate  $M^* \leq [Nr^2/(r-1)]^{1/4}$ .

Speedup figures resulting from the use of early fixation detection for  $M^*=100$  are shown in Figs. 9 and 10 for the  $K$ -funnel (with several  $n$  and  $K$  values) and the  $K$ -layer random networks in [25] (with two values of  $n$ ,  $K=5$ , and several values of the  $a$  parameter that, as explained in Sec. I, is used in [25] to control the layer-selection mechanism as the network is grown by the addition of new nodes), respectively. Speedup is now defined as the ratio of how much processor time is needed to determine the fixation probability without early detection to how much processor time is needed when early detection is used. Plots in the latter figure are given against  $S(X, Y)$ , which in [25] is used to obtain, for networks that are close to having an exponentially growing number of nodes per layer as one moves upstream through the layers, the basis of such exponential. As explained in [25], this basis approaches  $e^{S(X, Y)}$  and equals the branching parameter  $b$  for the  $K$ -funnel itself. [The  $X$  and  $Y$  appearing in the notation  $S(X, Y)$  are length- $K$  numerical sequences related to the layers in the  $K$ -layer random network that was grown ( $X$ ) and in the  $K$ -funnel ( $Y$ ). The value of  $S(X, Y)$  is the slope of the least-squares linear approximation of  $Y$  as a function of  $X$ .]

Clearly, speedups are very significant, particularly for the  $K$ -funnel with the largest values of  $n$  and  $K$  and all random networks of 10 000 nodes. We also remark that, in Fig. 9, the



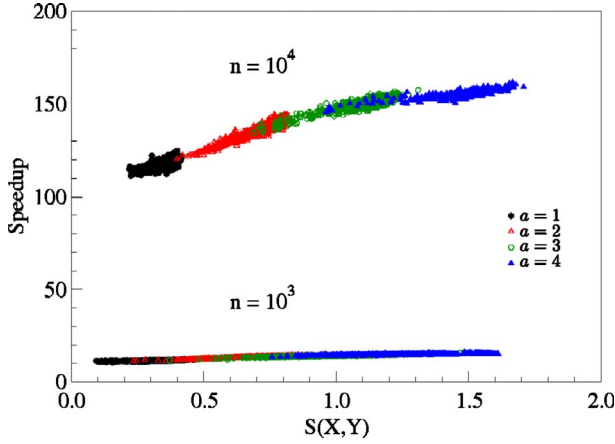


FIG. 10. (Color online) Speedups resulting from the use of the  $M^*=100$  threshold on the random layered networks in [25] with  $K=5$  layers and fitness ratio  $r=1.1$  for different values of the strength parameter  $a$  defined in Sec. I. Each graph whose layer populations are close by more than 0.9 (in the sense of the Pearson correlation coefficient) to being an exponentially growing sequence is shown as a point and represented by its  $S(X,Y)$  value and speedup. Data are averages over  $10^4$  events for each of  $10^3$  graphs, except for the case of  $n=10\,000$  nodes with  $a=4$ , in which  $10^4$  events are used for each of 500 graphs.

fact that slopes increase with the number of layers  $K$  is closely related to Eq. (4): as the fraction of events that lead to fixation grows with  $K$ , we expect from Fig. 1 that the average number of steps behaves likewise; qualitatively, this explains why speedups increase with  $K$ .

### V. FINAL REMARKS

One useful heuristic that has emerged from our study in Sec. IV is that, when performing Monte Carlo simulations to compute a large graph's fixation probability for  $r > 1$ , particularly if  $r$  is only slightly above 1, it suffices to choose the value of  $M^*$  to be about  $[N/(r-1)]^{1/4}$ , where  $N$  is the number of events in the simulation. Proceeding in this way ensures that the error incurred is roughly the standard error that is inherent to the simulations, given approximately by no more than  $\sqrt{(r-1)/N}$ . We derived this value of  $M^*$  for the isothermal case but, empirically, demonstrated that it can also be used with a much wider variety of topologies.

In Sec. I we mentioned that, in [25], we were unable to extend to larger values of the number of nodes  $n$  our conclusions regarding the fitness-amplification properties of the  $K$ -layer random networks we used in some of this paper's experiments. Provisioned with the technique of Sec. IV, we can now bypass the computational difficulties that hampered our progress in that occasion by employing early detection of fixation. Doing this for  $M^*=100$  has resulted in the data shown in Fig. 11, from which it is finally clear that, also for  $n=10\,000$ , it is possible to grow layered networks that achieve significant fitness amplification. As we see in the figure, for  $K=5$  many grown networks have values of the fixation probability  $\rho$  between  $\rho_2$  and  $\rho_3$ . We note, however, that this is still an easier scenario than that in [25], where

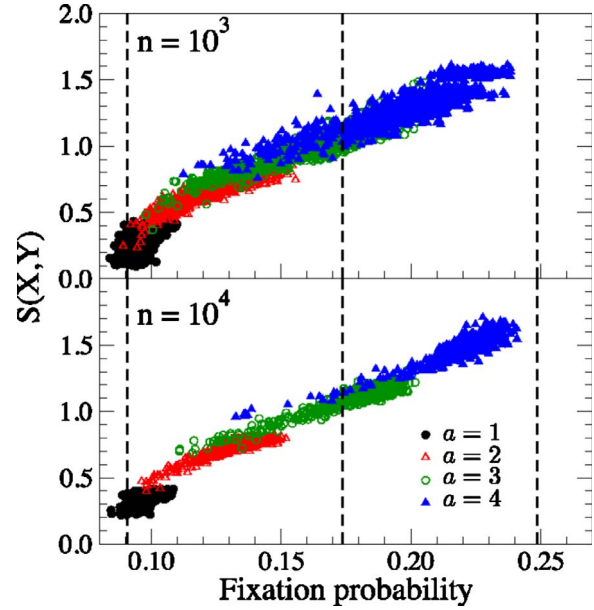


FIG. 11. (Color online) Results for the random layered networks in [25] with  $K=5$  layers and fitness ratio  $r=1.1$  for different values of the strength parameter  $a$ . Each graph whose layer populations are close by more than 0.9 (in the sense of the Pearson correlation coefficient) to being an exponentially growing sequence is shown as a point and represented by its fixation probability and  $S(X,Y)$  value. Data are averages over  $10^4$  events for each of  $10^3$  graphs, except for the case of  $n=10\,000$  nodes with  $a=4$ , in which  $10^4$  events are used for each of 500 graphs. Dashed lines indicate the fixation probabilities  $\rho_1-\rho_3$ .

$K=10$  was used for  $n=10\,000$ , since in the present case we had to calculate the speedups given in Fig. 10 and these required that simulation events be carried out to completion.

Another interesting by-product of our use of the threshold number of mutants  $M^*$  is that, should the dynamics be started with  $s > M^*$  randomly placed mutants of equal fitness, then fixation would occur almost surely. This is so because the probability of there being so many mutants in a dynamics that is bound to extinction is as small as allowed by the choice of the probability threshold  $Q^*$ . However, note that Eq. (15) is of no immediate help in quantifying the ‘‘almost surely,’’ since it is conditioned on the dynamics ending in extinction and therefore does not apply to those cases in which the number of mutants becomes large enough that extinction is unlikely. Nevertheless, whenever the isothermal theorem holds, Eq. (1) implies that fixation from the initial  $s$  mutants occurs with probability  $(1-1/r^s)/(1-1/r^n)$  (see also [1]), which is asymptotically equal to 1 for  $r > 1$  as both  $s$  and  $n$  grow.

It is also worth mentioning that, because this study has been targeted at directed graphs, Eqs. (8) and (9) are also applicable to the special case of an undirected graph and can lead to useful insight also in this case. Specifically, suppose we take any strongly connected graph  $D$  without antiparallel edges and make it functionally undirected by adding to it the antiparallel counterpart of every one of its edges. If  $E$  and the  $O_i$  sets continue to refer to the original  $D$ , then the contribution of each mutant  $l$  to  $\sum_{(i,j) \in E} m_j(t)/|O_i|$  in Eq. (9) jumps



from  $\sum_{(i,l) \in E} 1/|O_i|$  to  $\sum_{(i,l) \in E} 1/|O_i| + \sum_{(l,i) \in E} 1/|O_l|$ , therefore leading to a smaller  $P^+(t)/P^-(t)$  ratio. On the other hand, if  $D$  already has antiparallel edges, then one curious special case is that of the 2-superstar in [1], which has a central node with  $n-1$  peripheral nodes that connect to it through antiparallel edge pairs. This graph is already functionally undirected and, for the number of nodes  $n \rightarrow \infty$ , the fixation probability  $\rho \rightarrow \rho_2$ . So the 2-superstar is somewhat of an exception with regard to the  $P^+(t)/P^-(t)$  ratio for undirected graphs. In fact, thus far we have been unable to find any other undirected graphs for which  $\rho \geq \rho_2$ , but to the best of our knowledge the question of whether any exist remains unsettled.

We also note that fixation for the 2-superstar, when it happens, is bound to take a considerable number of steps to occur (this can be seen in Fig. 1 since the 2-superstar and the 2-funnel with the same number of nodes are the same graph). For the 2-superstar, the sum from Eq. (9) mentioned above is  $M_p(t)/(n-1) + m_c(t)(n-1)$ , where  $M_p(t)$  is the number of peripheral mutants at step  $t$  and  $m_c(t)$  is either 1 or 0, indicating whether or not a central mutant exists at that step, respectively. Clearly, for fitness ratio  $r$  only slightly above 1, obtaining  $P^+(t) > P^-(t)$  depends crucially on the existence of the central mutant. This, in turn, can easily change from step to step until fixation is eventually approached. It is precisely in cases such as this that the early estimates of the fixation probability introduced in Sec. IV are most useful.

We remark, finally, that the rule for individual selection and offspring placement we have used throughout this work

is only one of various possibilities. Normally such a rule is expressed in terms of an edge  $(i, j)$  of graph  $D$ . In the present case it is a rule of the so-called birth-death type with selection at the birth since first  $i$  is selected based on the individuals' fitnesses and only then is  $j$  chosen for replacement by the offspring of  $i$ . Another birth-death alternative would be to pick  $i$  independently of the fitnesses and do fitness-based selection to pick  $j$  instead. Should the order in which  $i$  and  $j$  are picked be reversed we would have the so-called death-birth rules, again with two possibilities regarding which of the two to select based on fitness. These rules and still others, such as those that pick the edge  $(i, j)$  as a single entity, are described in [32] and references therein. Since none of these alternative rules is mathematically equivalent to the one we have used, our results cannot be expected to carry over directly. However, we find it reasonable to expect that similar criteria to accelerate the determination of the fixation probability also exist in those cases. Studying them remains open to further research.

## ACKNOWLEDGMENTS

We acknowledge partial support from CNPq, CAPES, a FAPERJ BBP grant, the joint PRONEX initiative of CNPq and FAPERJ under Grant No. 26.171.528.2006, and CNPq-PROSUL.

- 
- [1] E. Lieberman, C. Hauert, and M. A. Nowak, *Nature (London)* **433**, 312 (2005).
- [2] P. A. P. Moran, *Proc. Cambridge Philos. Soc.* **54**, 60 (1958).
- [3] A. Valleriani and T. Meene, *Ecol. Modell.* **208**, 159 (2007).
- [4] M. Barthélemy, A. Barrat, R. Pastor-Satorras, and A. Vespignani, *Phys. Rev. Lett.* **92**, 178701 (2004).
- [5] H. Ohtsuki, C. Hauert, E. Lieberman, and M. A. Nowak, *Nature (London)* **441**, 502 (2006).
- [6] F. C. Santos and J. M. Pacheco, *Phys. Rev. Lett.* **95**, 098104 (2005).
- [7] P. D. Taylor, T. Day, and G. Wild, *Nature (London)* **447**, 469 (2007).
- [8] F. Fu, L. Wang, M. A. Nowak, and C. Hauert, *Phys. Rev. E* **79**, 046707 (2009).
- [9] A. Grönlund and P. Holme, *Adv. Complex Syst.* **8**, 261 (2005).
- [10] E. Pugliese and C. Castellano, *EPL* **88**, 58004 (2009).
- [11] G. Szabó and G. Fáth, *Phys. Rep.* **446**, 97 (2007).
- [12] J.-Y. Guan, Z.-X. Wu, Z.-G. Huang, X.-J. Xu, and Y.-H. Wang, *EPL* **76**, 1214 (2006).
- [13] B. J. Kim, A. Trusina, P. Holme, P. Minnhagen, J. S. Chung, and M. Y. Choi, *Phys. Rev. E* **66**, 021907 (2002).
- [14] H. Ohtsuki, M. A. Nowak, and J. M. Pacheco, *Phys. Rev. Lett.* **98**, 108106 (2007).
- [15] H. Ohtsuki, J. M. Pacheco, and M. A. Nowak, *J. Theor. Biol.* **246**, 681 (2007).
- [16] A. Szolnoki and G. Szabó, *EPL* **77**, 30004 (2007).
- [17] Z.-X. Wu, X.-J. Xu, and Y.-H. Wang, *Chin. Phys. Lett.* **23**, 531 (2006).
- [18] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, *Phys. Rep.* **424**, 175 (2006).
- [19] *Handbook of Graphs and Networks*, edited by S. Bornholdt and H. G. Schuster (Wiley-VCH, Weinheim, Germany, 2003).
- [20] *The Structure and Dynamics of Networks*, edited by M. Newman, A.-L. Barabási, and D. J. Watts (Princeton University Press, Princeton, NJ, 2006).
- [21] T. Antal, S. Redner, and V. Sood, *Phys. Rev. Lett.* **96**, 188104 (2006).
- [22] M. Broom and J. Rychtář, *Proc. R. Soc. London, Ser. A* **464**, 2609 (2008).
- [23] N. Masuda and H. Ohtsuki, *New J. Phys.* **11**, 033012 (2009).
- [24] V. Sood, T. Antal, and S. Redner, *Phys. Rev. E* **77**, 041121 (2008).
- [25] V. C. Barbosa, R. Donangelo, and S. R. Souza, *Phys. Rev. E* **80**, 026115 (2009).
- [26] P. Erdős and A. Rényi, *Publ. Math. (Debrecen)* **6**, 290 (1959).
- [27] R. M. Karp, *Random Struct. Algorithms* **1**, 73 (1990).
- [28] V. C. Barbosa, R. Donangelo, and S. R. Souza, *Physica A* **321**, 381 (2003).
- [29] M. A. El-Shehawey, *J. Phys. A* **33**, 9005 (2000).
- [30] S. Karlin and H. M. Taylor, *A First Course in Stochastic Processes*, 2nd ed. (Academic, New York, 1975).
- [31] M. A. Nowak, *Evolutionary Dynamics* (Harvard University Press, Cambridge, MA, 2006).
- [32] N. Masuda, *J. Theor. Biol.* **258**, 323 (2009).