Distance distribution in random graphs and application to network exploration

Vincent D. Blondel,* Jean-Loup Guillaume,† Julien M. Hendrickx,‡ and Raphaël M. Jungers§
Department of Mathematical Engineering, Université catholique de Louvain,
4 avenue Georges Lemaitre, B-1348 Louvain-la-Neuve, Belgium

(Received 22 June 2007; revised manuscript received 28 August 2007; published 5 December 2007)

We consider the problem of determining the proportion of edges that are discovered in an Erdős-Rényi graph when one constructs all shortest paths from a given source node to all other nodes. This problem is equivalent to the one of determining the proportion of edges connecting nodes that are at identical distance from the source node. The evolution of this quantity with the probability of existence of the edges exhibits intriguing oscillatory behavior. In order to perform our analysis, we introduce a different way of computing the distribution of distances between nodes. Our method outperforms previous similar analyses and leads to estimates that coincide remarkably well with numerical simulations. It allows us to characterize the phase transitions appearing when the connectivity probability varies.

DOI: 10.1103/PhysRevE.76.066101 PACS number(s): 89.75.Hc, 89.20.Hh, 02.50.−r, 05.50.+q

I. INTRODUCTION

The small-world phenomenon has attracted increasing attention over the last few years [1,2]. In a small-world network, the average distance between two nodes is small as compared to the total number of nodes. In many natural networks, it is typically of the order of $\log(n)$ (n is the total number of nodes) and several models have been proposed to explain this phenomenon (see, e.g., [1,3,4]). In some applications though, one is interested not only in this so-called "average intervertex distance", but in the whole intervertex distance distribution.

Even though this distribution is of much interest, it has not been studied very much in the literature. A theoretical method for the computation of the distances in uncorrelated random networks of infinite size was proposed by Dorogovtsev *et al.* in 2003 [5]. In 2004, Fronczak *et al.* analyzed the distance between nodes for a wide class of random networks of finite size that generalizes the Erdős-Rényi graphs, the so-called uncorrelated random networks with hidden variables [6]. They propose an approximation of the distribution of the distance between nodes that performs well for a certain range of parameter values. Their formula has the advantage of being simple and analytical, but the approximations made in the calculations lead to significant differences from the numerical evidence for some ranges of parameters.

Our work is motivated by the analysis of algorithms that have been recently developed for analyzing networks such as the internet. A typical way of doing that is to use the freeware TRACEROUTE, which provides the user a short path from his computer to any other one in the internet. In the all shortest paths (ASP) model, introduced to model this strategy, one chooses a particular node s of the network, and then constructs all shortest paths from s to all other nodes of the network [7]. Some edges of the network may not belong to any of these shortest paths and so they are left undiscovered. The problem considered in [7] is that of determining the proportion of edges of the network that are discovered. Thus the question is "what is the proportion of edges that are on at least one shortest path starting from the source?" As pointed out in [7], the edges that are not discovered are exactly those connecting nodes that are at identical distances from the source. Indeed, if an edge connects two equidistant nodes, it cannot be on a shortest path from the source, since any path using this edge (say going from v_1 to v_2) can be shortened by going directly to v_2 via the shortest path to v_2. Conversely, if an edge links two nodes that are not at the same distance, then it links a node v_1 at a certain distance d to a node v_2 at a distance $d+1$, and at least one shortest path to v_2 passes through this edge. We are therefore interested in computing the number of edges connecting nodes that are at the same distance from the source. Other models exist for representing network analysis strategies. For instance, [7] introduces the unique shortest path (USP) model. In the USP model one chooses only one shortest path from the source to each node in the graph and so there are possibly more edges that are left undiscovered. Our work is also relevant to the analysis of this model, as it counts the proportion of edges that are never found by any single or multiple USP search.

In [7], massive numerical simulations were performed to analyze the proportion of edges that are on shortest paths in Erdős-Rényi graphs. In such random graphs, edges are all equally likely to be present and the probability of presence is given by some fixed probability p. We do not consider self-loops or multiple edges. So, for constructing an Erdős-Rényi graph, one needs to fix two parameters: the number of nodes n and the probability of existence for every edge p. As shown in Fig. 1, the proportion of edges that are discovered in the ASP model presents an interesting dependence on the parameter p. One can directly explain some characteristics of this curve. When p is very small the graph is highly disconnected and consists in small connected components. Most edges therefore do not belong to any path starting from the source, and the proportion of observed edges is close to zero. Conversely, if p is very high, the graph is almost complete, and every shortest path has length 1. So $n-1$ edges are

*vincent.blondel@uclouvain.be
†jean-loup.guillaume@uclouvain.be
‡julien.hendrickx@uclouvain.be
§raphael.jungers@uclouvain.be

 ©2007 The American Physical Society

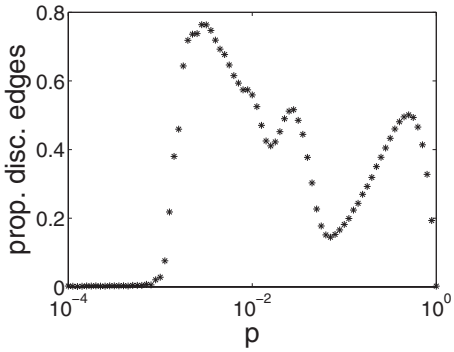

FIG. 1. Evolution with p (logarithmic scale) of the proportion of edges that lie on a shortest path in an Erdős-Rényi random graph with $n=1000$ vertices. Each value is computed by averaging the observations made on 1000 graphs [7].

found, while there are almost $\frac{1}{2}n(n-1)$ edges in the graph, and thus the proportion also vanishes.

The aims of this paper are first to introduce another simple model of intervertex distances in Erdős-Rényi graphs that can be used to compute the curve of Fig. 1 without any numerical experiment, and second to analyze the oscillating behavior of this curve and explain the phase transitions appearing with variations of the graph connectivity. Note that similar oscillating behaviors in random graphs have recently been observed [8], and that these phenomena seem to open challenging questions in random graph theory. This paper proposes a precise analysis of such an oscillating behavior in the simple theoretical framework of Erdős-Rényi graphs. One could imagine exploiting these oscillations to optimize the design of a network or to develop a method for its analysis, although this is beyond the scope of this paper. Besides, such applications of the concepts developed here would probably require some further analysis and extension of our results, because real networks often exhibit nontrivial correlations between nodes that do not occur in Erdős-Rényi graphs. These extensions, however, would most likely not lead to the derivation of simple analytical solutions providing an intuitive understanding of the phenomena as is done here.

The remainder of the paper is organized as follows. In Sec. II, we introduce a recurrence equation allowing us to evaluate the intervertex distance distribution for Erdős-Rényi graphs, and compare our results to those previously published [6]. From this function we derive a theoretical expression for values shown on Fig. 1. In Sec. III we analyze this curve, characterize the phase transitions, and give analytical expressions in different phases (proved in the Appendix). In Sec. IV we conclude and make some remarks on practical applications of the phenomena studied in the paper.

II. APPROXIMATION OF INTERVERTEX DISTANCE DISTRIBUTION

In this section we propose an approximation for the intervertex distance distribution in Erdős-Rényi graphs. We compare our results to those obtained by Fronczak *et al.* [6] in a more generic situation, and show how our results outperform theirs in the particular case of Erdős-Rényi graphs. We also analyze the accuracy of our model and its dependence on the graph connectedness. We then use our intervertex distance distribution to estimate the proportion of equidistant pairs of nodes.

In the following, we consider the distance between a randomly selected node and a fixed but initially randomly selected "source node." Since this source is randomly selected, all results obtained for the distance probability can also be applied to the distance between two randomly selected vertices. Let F_d be the probability for a randomly selected node to be at a distance larger than d from the source, that is, the probability that there is no path of length smaller than or equal to d from the source to this node. The probability f_d for the node to be at a distance exactly d from the source is then given by $f_d=F_{d-1}-F_d$. Obviously, $F_0=1-1/n$. We now derive a recurrence relation allowing the computation of F_d for higher values of d. A node is at a distance larger than d from the source if it is not the source itself, which happens with probability $1-1/n$, and if it is connected to no node at distance less than d from the source, which happens with probability $(1-p)^{n_d}$, n_d being the number of nodes at distance less than d from the source. We have therefore the following simple relation:

$$F_d=\left(1-\frac{1}{n}\right)\sum_{k=1}^{n-1}P[n_d=k](1-p)^k, \tag{1}$$

where $P[n_d=k]$ denotes the probability that $n_d=k$. In order to express the probability F_d, we should thus know the distribution of n_d. We approximate this quantity to be always exactly equal to its expectation $\langle n_d\rangle=(1-F_{d-1})n$. Introducing this approximation in (1), we obtain a recurrence relation for F,

$$F_d=\left(1-\frac{1}{n}\right)(1-p)^{(1-F_{d-1})n}, \tag{2}$$

which allows us to compute f_d for any d. This formula is different from, but provably equivalent to, Eq. (6) in [9], which was derived independently for other purposes.

In [6], Fronczak *et al.* propose an expression for the intervertex distance distribution of any random graph with hidden variables (these graphs are a generalization of Erdős-Rényi graphs). In these graphs, two nodes i and j are connected with a probability $p_{i,j}=h_ih_j/\beta$, where each node v has its own hidden variable h_v, and $\beta=\langle h\rangle n$. So, in a large graph, the hidden variables represent the expected degree of the vertices. In the particular case of Erdős-Rényi graphs, that is, when $h_v=np$ for all $v\in V$, the expression of Fronczak *et al.* for the function F of intervertex distance distribution reduces to

$$F_d=e^{-1/n(np)^d}. \tag{3}$$

This result has a straightforward interpretation as the solution of another recurrence equation in d, although it is not obtained in that way in [6]. A vertex i is at a distance larger than d from the source node if all its neighbors are at dis-

tance larger than $d-1$ from the source. Approximating the number of neighbors by its expectation np and neglecting the dependence effects, one obtains the recurrence $F_d = F_{d-1}^{np}$. The relation (3) is then obtained again by taking $F_0 = e^{-1/n} \simeq 1 - 1/n$ as initial condition. Numerical experiments confirm indeed that taking $e^{-1/n}$ or $1-1/n$ as initial condition has no influence on the results if n is sufficiently large.

In Fig. 2 we compare the predictions from the two models with numerical results. One can see that both models perform very well when the average degree np is significantly larger than 1 and if p is not too big, as in Fig. 2(c). For an average degree $np<1$ which is below the emergence of the giant connected component (see [1]), our results approximately match the experimental observations, while Fronczak *et al.*'s model is not valid as it gives an increasing curve [see Fig. 2(a)]. For values of np larger than but close to 1, both models present significant errors but ours is closer to the experimental observations [see Fig. 2(b)]. Finally, for a large p, one can see in Fig. 2(d) that our results match the experimental data very well while those obtained with the model of [6] are significantly different.

The fact that the model derived in [6] behaves very differently from our model for a certain range of values of p may seem surprising. Our derivation presents indeed various similarities with the interpretation of Fronczak *et al.*'s model as a solution of a recursive equation. Three reasons can, however, explain why a model based on this interpretation gives less accurate results than ours. First, for $np \ll 1$, the possibility for the randomly selected node to be the source cannot be neglected, as very few nodes are in the connected component of the source. When np is larger than but close to 1, the approximation that a node has exactly np neighbors leads to proportionally more important errors. This problem could be solved by considering a binomial distribution for the number of neighbors in our interpretation of Fronczak *et al.*'s model. Finally, for large values of p, the number of neighbors of the randomly selected node is large, so that some independence problems are not negligible. Indeed, the probabilities for two neighbors of the randomly selected node to be the source are not independent, as there is exactly one source in the graph.

The errors of our model, observed for values of np larger than but close to 1 are due to the approximation mentioned above: To obtain the recurrence equation (2) from (1), we suppose that the number n_d of vertices at a distance smaller than d from the source is exactly equal to its expectancy $n(1-F_{d-1})$, instead of considering its probability distribution. In this range of parameters, the distribution is far from being centered because of the existence of a peak around 0 [see Fig. 3(a)]. For these values, indeed, the graph is not totally connected. If the source happens not to be in the giant connected component, almost all nodes are at an infinite distance from it, so that n_d is close to 0 for any d. The weight of the peak thus represents the probability for a randomly selected source not to be in the giant connected component. It is known that, when np grows, this probability tends exponentially to 0 independently of n ([10], Theorem 5.4). This problem therefore appears only when the average degree np is very small (but larger than 1), independently of the size n of the graph. Figure 3(b) shows that the problem is already almost negligible when $np=4$ (for these values, the giant connected component already contains more than 98% of the vertices). Note that for $np<1$ the graph is highly disconnected so that almost no nodes are at a finite distance from the source. The distribution $P[n_d=k]$ consists thus of only one peak around 0 and is therefore centered.

We close this section by explaining how the distance distribution can be used to compute the proportion P_s of edges belonging to shortest paths starting at the source node. As explained in the Introduction, the edges that do not belong to any shortest path are those connecting nodes that are at the same distance from the source, in addition to all edges that are not in the same connected component as the source. Since the expectation of the number of nodes at distance d from the source is equal to nf_d, the expected number of edges connecting these nodes is roughly equal to $\frac{1}{2}p(nf_d)^2$. Taking $\frac{1}{2}pn^2$ as the total number of edges, we obtain the following expression for the proportion of edges that lie on a shortest path in an Erdős-Rényi graph, which we denote by $P_s(n,p)$ in the following:

$$P_s(n,p) = 1 - \frac{\sum_{d=1}^{n} p(nf_d)^2}{n^2 p} = 1 - \sum_{d=1}^{n} f_d^2. \tag{4}$$

Note that this expression implicitly handles the edges that are not in the same connected component as the source if we take $f_n = F_{n-1} \approx F_\infty$. Indeed, this quantity represents those nodes that are not connected to the source, as they are at a distance larger than $n-1$. The evolutions with p of P_s using the two models presented above are represented in Fig. 4 for $n=1000$ and 10 000. One can see that our results match the experiments very well except when np is larger than but close to 1, which is the range of parameters for which our model has already been shown to be less accurate. Moreover, the range of values of np for which our model is less accurate appears not to grow with n.

III. ANALYSIS OF THE CURVE

In this section we analyze the function $P_s(n,p)$ generated with our model and (4). We show the appearance of a sort of phase transition: for some particular values, a weak variation of the probability p may cause abrupt changes in the proportion of discovered edges with the ASP model and dramatically affect the properties of the graph. We give analytical formulas for the asymptotic behavior in several phases.

We begin by analyzing the first transition, starting from small values of p. It is well known [1] that, in an Erdős-Rényi graph, a giant component emerges when p becomes larger than $1/n$. If the average degree np is sufficiently small, the graph is not connected and the only edges that the observer can see are in the (small) connected component of the source. This quantity is negligible in view of the total number of edges, and so the function is approximately zero. Note, however, that such graphs do not contain many cycles, so that most paths starting from the source are shortest paths. Therefore, the observer discovers approximately all edges in

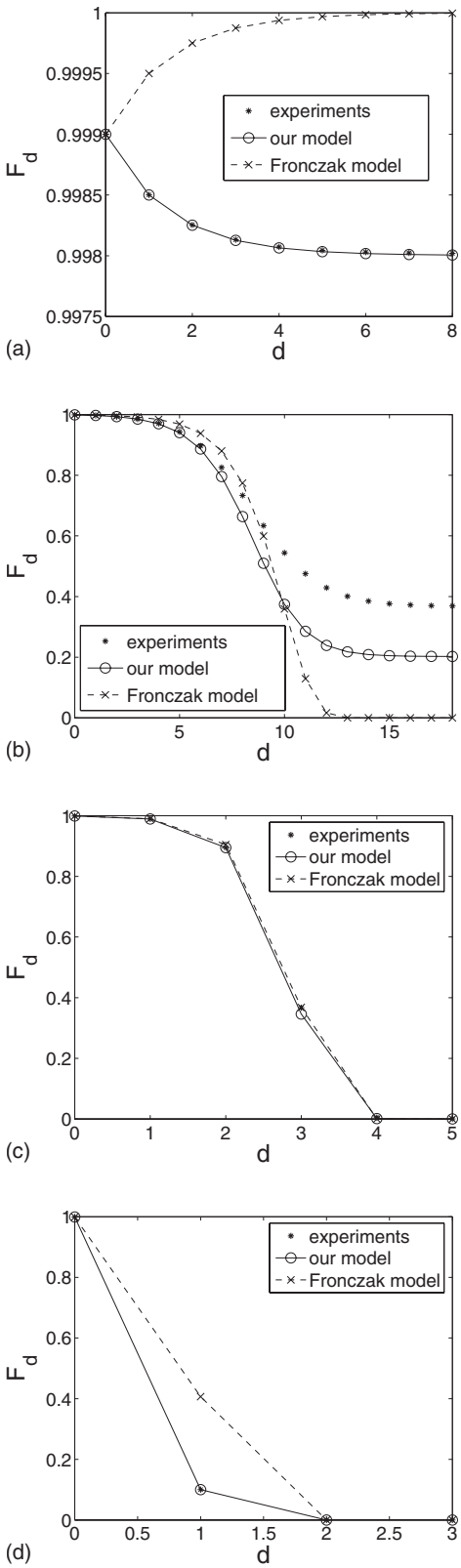

FIG. 2. Evolution of F_d, the probability for a random node to be at a distance larger than d from the source node, for $n=1000$ nodes and for $np=$ (a) 0.5, (b) 2, (c) 10, and (d) 900. The three curves represent the experimental observations (averaged on 500 graphs), our model, and the model of Fronczak *et al.* given in [6].

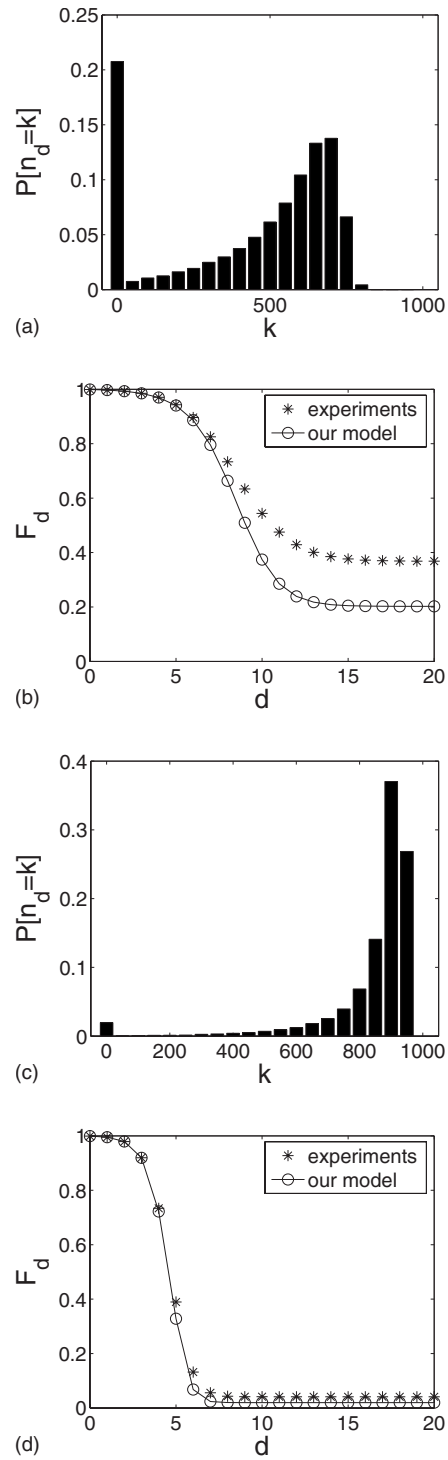

FIG. 3. Representation of $P[n_d=k]$, the probability that there are exactly k nodes at distance less than d from the source, obtained experimentally, compared to F_d, the proportion of nodes at a distance larger than d from the source, obtained experimentally and with our model, for $np=2$ (a,b) and 4 (c,d), with $n=1000$ in both cases. $P[n_d=k]$ is represented for $d=11$ in (a) and for $d=7$ in (c) as typical path lengths are different when $np=2$ or 4. The distribution in (a) is bimodal as it contains a large peak around 0, while the peak in (c) is much smaller. Our approximation of n_d by its average value $n(1-F_{d-1})$ leads thus to larger errors for $np=2$ than for $np=4$.

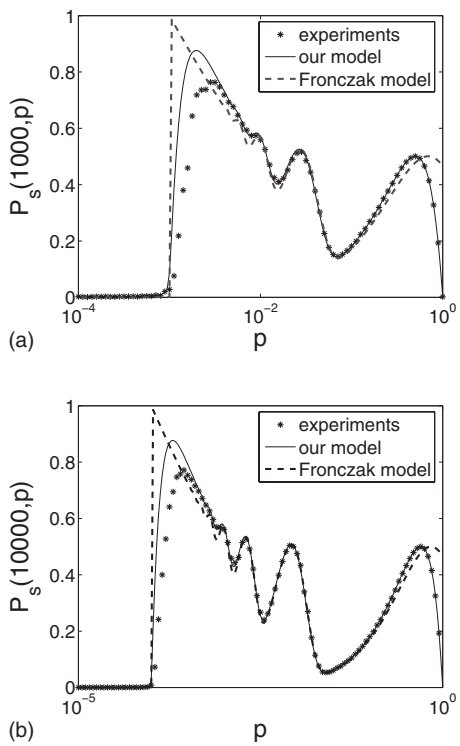

FIG. 4. Comparison of the evolution of $P_s(n,p)$ with $n=1000$ (a) and 10 000 (b) according to numerical experiments [7], to our model, and to Fronczak *et al.*'s model.

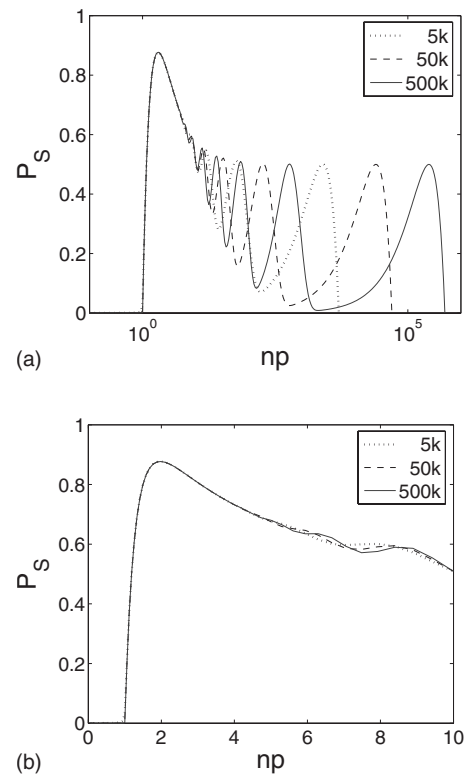

FIG. 5. Evolution of $P_s(n,p)$ with np for different values of n. All curves present a sharp increase between $np=1$ and 2, and a global maximum at $np\simeq 2$. For larger values, the curves present several oscillations, with local maxima tending to 0.5. (b) is a zoomed-in linear-scale version of (a).

its connected component. When p grows the size of the connected components increases, so that more and more edges are discovered. Now, when $np\approx 1$, the giant component emerges very quickly, and the source is in this component with a large probability. Since most of the edges are also in this component, the proportion of discovered edges increases rapidly with np. Simultaneously with the apparition of a giant component, there also appear a non-negligible number of cycles in the graphs, so that not all edges lie on shortest paths any longer. As a result of these two conflicting phenomena a (global) optimum is reached for $np\approx 2$. Experimentally, our model gives an optimum that seems to lie exactly at $np=2$, but we have not been able to prove this, nor to express analytically the values of $P_s(n,p)$ around $np\approx 2$. However, experiments seem to indicate that in this range of parameters $P_s(n,p)$ depends only on np. All this can be seen in Fig. 5, for different values of n. When np becomes larger, one can see that P_s does not depend only on np, and presents an oscillatory behavior. In particular, the successive values of the maxima seem to tend to $\frac{1}{2}$. We explain this phenomenon in the following.

As can be observed in Fig. 6, the shape of $P_s(p)$ tends to the parabola $2p(1-p)$ on any interval $[\epsilon,1]$ when n increases (note that the x axis is in linear scale). This fact can be proved theoretically, based on our model of evolution of F_d. In the following, for the sake of clarity in our analysis, we modify (2), and study the slightly different equation

$$F_d(n,p) = (1-p)^{n[1-F_{d-1}(n,p)]}. \tag{5}$$

This new approximation is justified by the fact that we will consider asymptotic behaviors for $n\to\infty$. Moreover, the results that we derive can be obtained without making this approximation. Observe that $F_1(n,p)=1-p$, so that $F_2(n,p)=(1-p)^{np}$. When n grows, $F_0=1-1/n\to 1$, and if p is bounded from below by an arbitrary positive constant ϵ, $F_2(n,p)=(1-p)^{np}$ tends uniformly to 0. As a consequence the probability $f(d)$ for a node to be at a distance d from the source tends uniformly to 0 for all d except for $d=1,2$, for which $f_1=F_0(n,p)-F_1(n,p)\to p$ and $f_2=F_1(n,p)-F_2(n,p)\to 1-p$. It follows then from (4) that

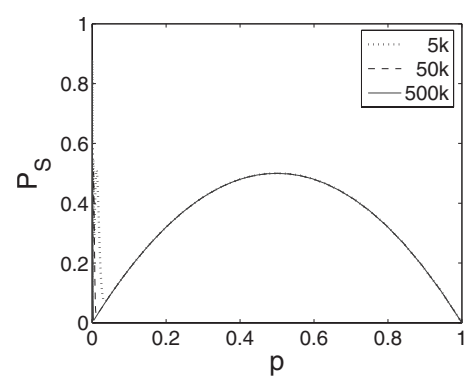

FIG. 6. Evolution of $P_s(n,p)$ with p for different values of n. On any interval $[\epsilon,1]$, P_s tends to the parabola $2p(1-p)$ when n increases.

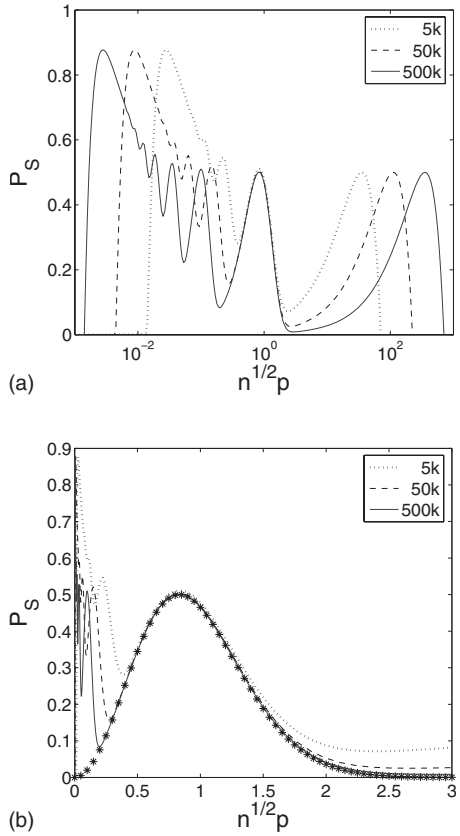

FIG. 7. Evolution with $n^{1/2}p$ of $P_s(n,p)$ for different values of n. (b) is a zoomed-in linear-scale version of (a). Asymptotically, P_c behaves as $2e^{-np^2}(1-e^{-np^2})$ which is represented by $*$ in (b).

$$P_s(n,p) \to 1 - p^2 - (1-p)^2 = 2p(1-p),$$

so that asymptotically the last maximum of P_s is $\frac{1}{2}$ and is reached at $p=\frac{1}{2}$. The asymptotic parabolic character of P_s is thus here due to the fact that almost all nodes tend to be at a distance either 1 or 2 from the source when n grows and p is sufficiently large, as can be observed, for example, in Fig. 2(d).

We now analyze the oscillating behavior between the first and last maxima. One can see in Fig. 7 that, around the second-rightmost maximum, P_s depends only on $n^{1/2}p$, and that P_s asymptotically behaves as

$$P_s \simeq 2e^{-(n^{1/2}p)^2}(1 - e^{-(n^{1/2}p)^2}) \tag{6}$$

around this maximum. The maximum therefore tends to $\frac{1}{2}$ when $n\to\infty$ and is attained for $(n^{1/2}p)^2=\ln 2$. To explain (6), we show in the Appendix that, as above, all nodes are asymptotically at distance either 2 or 3 when $n\to\infty$ with $\epsilon < np^2 < R$, where ϵ and R are arbitrarily positive constants. As in the case of the parabola, this together with (4) implies that P_s then asymptotically behaves as $2(1-F_2)F_2$. We also show that $F_2(n,p)$ tends to e^{-np^2}, which implies (6).

Actually, the previous relations can be generalized inductively: we prove in the Appendix that, when $n\to\infty$ with $\epsilon < n^{d-1}p^d < R$, F_d converges uniformly to $e^{-n^{d-1}p^d}$, while all $F_{d'}$ with $d'<d$ converge uniformly to 1 and all others to 0. This means that in this range of parameters, and when n tends to infinity, almost all nodes are at distance d or $d+1$ from the source. It follows then from (4) that

$$\lim_{\epsilon<n^{d-1}p^d<R} P_s(n,p) = 2(1-e^{-n^{d-1}p^d})e^{-n^{d-1}p^d},$$

which, as for $d=1,2$, is a parabolic curve with respect to F_d. This parabolic curve attains its maximum $\frac{1}{2}$ when $e^{-n^{d-1}p^d}=\frac{1}{2}$. So, when $n\to\infty$, P_s contains an unbounded number of oscillations and local maxima with asymptotic values $\frac{1}{2}$, and these maxima are attained when $n^{d-1}p^d=\ln 2$ for each $d>1$ as can be seen on some additional examples in Fig. 8. Experimentally, all local maxima but the first global one can be explained in that way. Between two maxima, there is a zone where asymptotically $F_d\simeq 1$ and $F_{d+1}\simeq 0$, so that almost all nodes are at distance $d+1$ from the source, and $P_s\simeq 1-1^2=0$. Such behavior is obtained when $n\to\infty$ with either large values of $n^{d-1}p^d$ but still $\epsilon<n^{d-1}p^d<R$, or small values of n^dp^{d+1} but still $\epsilon<n^dp^{d+1}<R$. One can indeed see in Figs. 7 and 8, for example, that the values of the local minima decrease significantly when n increases. Let us mention that an explanation of the oscillatory behavior based on the fact that almost all nodes are at distance d or $d+1$ from the source was suggested without proof in [7].

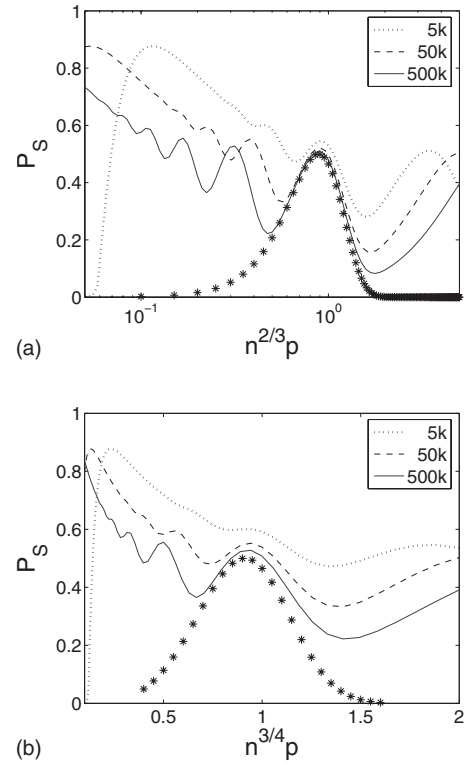

FIG. 8. Evolution of $P_s(n,p)$ with $n^{2/3}p$ (a) and $n^{3/4}p$ (b), for different values of n. Asymptotically, local maxima $1/2$ appear for $n^{2/3}p=\sqrt[3]{\ln 2}$ and $n^{3/4}p=\sqrt[4]{\ln 2}$. The $*$ represent the theoretical asymptotic behavior.

IV. CONCLUSIONS AND FUTURE WORK

The goal of this paper was twofold: First, we have proposed a simple model for the computation of the intervertex distance distribution in a random graph, via a recurrence equation for the probability for two randomly chosen nodes to be at distance apart more than d. Contrary to the model of Fronczak *et al.*, our recurrence equation is not explicitly solvable, but it is more accurate. It has to be noted that, for the range of parameters corresponding to the oscillating behavior analyzed at the end of Sec. III, the two models are equally valid, and that the analysis that we have made for such values could also be made using Fronczak *et al.*'s model. Let us add that the ideas behind the derivation of the formula remain valid for more general graphs such as random graphs with hidden variables. In the particular case of Erdős-Rényi graphs, these ideas lead to a simple recurrence equation, allowing computation of explicit values numerically, and proof of the asymptotic behavior of the curve experimentally obtained in [7]. Nevertheless, a further analysis for more general graphs could be interesting.

Second, following numerical simulations in previous work motivated by practical graph exploration questions [7], we have analyzed the proportion of edges connecting nodes that are equidistant from a certain source node in random graphs. The evolution of this quantity with the parameter p exhibits an intriguing oscillating behavior, which we have been able to explain and reproduce with a great accuracy using our model. We have also characterized precisely the (infinite number of) transitions for this quantity, and the analytical evolution with p in the different phases.

ACKNOWLEDGMENTS

The authors wish to thank Renaud Lambiotte for his useful advice. The research reported here was partially supported by the "Communauté francaise de Belgique—Actions de Recherche Concertées," by the EU HYCON Network of Excellence (Contract No. FP6-IST-511368), and by the Belgian Program on Interuniversity Attraction Poles initiated by the Belgian Federal Science Policy Office. J.H. and R.J. acknowledge support from the FNRS Foundation.

APPENDIX: EXPRESSION OF THE ASYMPTOTIC BEHAVIOR

In this appendix, we provide an analytical expression for F_d when n tends to infinity with $\epsilon < n^{d-1}p^d < R$, and we show that, in this range of parameters, almost all nodes are at distance d or $d+1$. Suppose first that $n \to \infty$ with $0 < np^2 < R$ for an arbitrary R. Then $p \to 0$ so that $F_1 = (1-p) \to 1$ uniformly with np^2. From our recurrence formula (5), we have

$$F_2(n,p) = (1-p)^{np} = [(1-p)^{1/p}]^{np^2},$$

which, together with the classical relation $\lim_{p\to 0}(1-p)^{1/p} = e^{-1}$, implies that

$$\lim_{np^2<R} F_2(n,p) = e^{-np^2} \tag{A1}$$

holds uniformly for $0 < np^2 < R$. We now show that

$$\lim_{\epsilon<np^2<R} F_3(n,p) = \lim_{\epsilon<np^2<R} (1-p)^{n(1-F_2)} = 0$$

for any two arbitrary constants ϵ and R. This implies that almost all nodes are at distance 2 or 3 from the source. It follows from (A1) that $1-F_2$ is uniformly bounded from below by a positive constant when $n \to \infty$ with $\epsilon < np^2 < R$, so that we just need to prove the uniform decay of $(1-p)^n$. The latter expression can be rewritten as

$$(1-p)^n = \left[\left(1 - \frac{np^2}{(np)}\right)^{(np)}\right]^{1/p}. \tag{A2}$$

Since $np \to \infty$ when $n \to \infty$ with $\epsilon < np^2 < R$, it holds uniformly that

$$e^{-R} \le \lim_{\epsilon<np^2<R} \left(1 - \frac{np^2}{(np)}\right)^{np} \le e^{-\epsilon}.$$

And since $1/p \to \infty$, it follows from (A2) that

$$\lim_{\epsilon<np^2<R} (1-p)^n = 0,$$

which implies the desired result.

It remains to prove our assertions about the asymptotic behavior of F_d for any $d > 2$. We first prove by induction that the two following relations hold uniformly for $n^{d-1}p^d < R$ where R is any arbitrary positive constant:

$$\lim_{n^{d-1}p^d<R} F_{d-1}(n,p) = 1, \tag{A3}$$

$$\lim_{n^{d-1}p^d<R} F_d(n,p) = e^{-n^{d-1}p^d}. \tag{A4}$$

These relations hold for $d=2$ as shown above. Let us now assume that they hold for a certain $d-1$ and prove that they then hold for d. Observe first that, when $n \to \infty$ with $n^{d-1}p^d < R$, $n^{d-2}p^{d-1}$ tends uniformly to 0 and is bounded. It follows then from the induction hypothesis that

$$F_{d-1}(n,p) \to e^{-n^{d-2}p^{d-1}} \to 1 - n^{d-2}p^{d-1}$$

uniformly when $n \to \infty$, $n^{d-1}p^d < R$. So Eq. (A3) is proved. By our recurrence relation (5), $F_d(n,p) = (1-p)^{n(1-F_{d-1})}$. Therefore, it holds that

$$\lim_{n^{d-1}p^d<R} F_d(n,p) = (1-p)^{(np)^{d-1}} = [(1-p)^{1/p}]^{n^{d-1}p^d}.$$

Since $n^{d-1}p^d$ is bounded, and since p tends thus uniformly to 0 when $n \to \infty$, the last equation becomes

$$\lim_{n^{d-1}p^d<R} F_d(n,p) = e^{-n^{d-1}p^d}$$

uniformly for $n^{d-1}p^d \in (0,R)$, which proves (A4).

Using the results above we now prove that for any $d > 2$ the following holds uniformly:

$$\lim_{\epsilon<n^{d-1}p^d<R} F_{d+1}(n,p) = 0, \tag{A5}$$

where ϵ and R are two arbitrary positive constants. By (5), we have

$$F_{d+1}(n,p) = (1-p)^{n(1-F_d)}.$$

It follows from the results above that $1-F_d$ is uniformly bounded from below by a positive constant when $n \to \infty$ with $\epsilon < n^{d-1}p^d < R$, so that we just need to prove the uniform decay of $(1-p)^n$. The latter expression can be rewritten as

$$(1-p)^n = \left[\left(1 - \frac{n^{d-1}p^d}{(np)^{d-1}}\right)^{(np)^{d-1}}\right]^{1/n^{d-2}p^{d-1}}. \tag{A6}$$

Since $(np)^{d-1} \to \infty$ when $n \to \infty$ with $\epsilon < n^{d-1}p^d < R$, it uniformly holds that

$$e^{-R} \leq \lim_{\epsilon < n^{d-1}p^d < R} \left(1 - \frac{n^{d-1}p^d}{(np)^{d-1}}\right)^{(np)^{d-1}} \leq e^{-\epsilon}.$$

And since $n^{d-2}p^{d-1} \to 0$ when $n \to \infty$ with $\epsilon < n^{d-1}p^d < R$, it follows from (A6) that

$$\lim_{\epsilon < n^{d-1}p^d < R} (1-p)^n = 0,$$

which implies the desired result (A5).

[1] B. Bollobas, *Random Graphs* (Cambridge University Press, Cambridge, U.K., 2001).
[2] J. Travers and S. Milgram, Sociometry **32**, 425 (1969).
[3] S. N. Dorogovtsev and J. F. F. Mendes. *Evolution of Networks: From Biological Nets to the Internet and WWW* (Oxford University Press, Oxford, 2003).
[4] D. J. Watts, *Small Worlds: The Dynamics of Networks Between Order and Randomness* (Princeton University Press, Princeton, NJ, 2003).
[5] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin, Nucl. Phys. B **653**, 307 (2003).
[6] A. Fronczak, P. Fronczak, and A. Janusz, Phys. Rev. E **70**, 056110 (2004).
[7] J.-L. Guillaume and M. Latapy, Complex Syst. **16**, 83 (2005).
[8] J. Sienkiewicz, P. Fronczak, and J. A. Holyst, e-print arXiv:cond-mat/0608273.
[9] A. Baronchelli and V. Loreto, Phys. Rev. E **73**, 026103 (2006).
[10] S. Janson, T. Luczak, and A. Ruci, *Random Graphs* (Wiley-Interscience, New York, 2000).