Tree networks with causal structure

P. Bialas,[1] Z. Burda,[2] J. Jurkiewicz,[2] and A. Krzywicki[3]
[1]*Institute of Computer Science, Jagellonian University, ul. Nawojki 11, 30-072 Kraków, Poland*
[2]*M. Smoluchowski Institute of Physics, Jagellonian University, ul. Reymonta 4, 30-059 Kraków, Poland*
[3]*Laboratoire de Physique Théorique, Bâtiment 210, Université Paris–Sud, 91405 Orsay, France*
(Received 26 November 2002; published 17 June 2003)

A geometry of networks endowed with a causal structure is discussed using the conventional framework of the equilibrium statistical mechanics. The popular growing network models appear as particular causal models. We focus on a class of tree graphs, an analytically solvable case. General formulas are derived, describing the degree distribution, the ancestor-descendant correlation, and the probability that a randomly chosen node lives at a given geodesic distance from the root. It is shown that the Hausdorff dimension d_H of the causal networks is generically infinite, in contrast to the maximally random trees where it is generically finite.

DOI: 10.1103/PhysRevE.67.066106 PACS number(s): 02.50.Cw, 05.40.−a, 05.50.+q, 87.18.Sn

I. INTRODUCTION

A. Preamble

The network study is an old field of research, which has recently become particularly active (see Ref. [1]). This is largely due to the opening of an access to rich data on natural systems: the worldwide web, the Internet, the various biological networks (gene transcription, cell metabolism, protein interactions), the sociological networks (citation index, collaborations, phone calls), etc. Interesting empirical regularities have been observed, such as the small-world property of networks or the frequently observed scale-free nature of degree distributions. On the theory side, the natural conceptual framework for network research is the graph theory; the nodes and the links of a graph represent the active agents and the interactions, respectively. However, it was soon realized that the classical graph theory [2] is inadequate and has to be generalized, in order to cover the new reality. This observation has triggered, in turn, an intense theoretical activity, which led to the construction of a number of insightful models. It is clear, however, that much remains to be done.

As recalled in Ref. [3], in studying the complex systems, one can adopt one of the two complementary approaches: the diachronic and the synchronic one. In the former approach, one focuses on the time evolution of the system, which helps discovering the dynamics at work. In the latter approach, one considers an ensemble of dynamically similar systems at a fixed large time, which helps in identifying the generic structural traits.

Most of the recently constructed models adopt the diachronic approach and deal with growing networks [4–20]. This led, among others, to the remarkable discovery of the preferential attachment rule and of its crucial role in network evolution [4,5]. The mathematics of the diachronic approach rests largely on master equations and related manipulations. The synchronic approach has also been advocated, and various static statistical ensembles have been constructed and studied [3,21–30]. Here, the mathematics is that of the equilibrium statistical mechanics and probability theory. One of the aims of this paper is to help establishing a bridge between the two sets of results. It will be seen that the widely accepted distinction between growing and equilibrium networks is not really correct. The truly distinctive property of the growing networks, their causal structure, can be usefully incorporated in a static model. We limit ourselves to tree graphs, because in this case the use of the methods of analytic combinatorics enables one to get exact solutions.

We shall discuss at length the geometry of connected tree graphs endowed with a causal structure. Our discussion will cover, in particular, some of the recently popular growing network models [4,6], but we shall proceed considering static statistical ensembles of trees, employing a formalism which proved to be useful in deriving generic results for other random geometries. For example, we shall show that the causal trees have the ''small world'' property: their Hausdorff dimension is infinite. The advantage of our approach is that it enables one to show at once that this property holds in a wide class of models.

The plan of the paper is as follows. In the next two sections, we recall some basic concepts and define the observables to be calculated later on. In Sec. II, we begin by deriving a number of general results that hold for causal trees. The central role is played by a recursion relation satisfied by the partition function (Sec. II A). This recursion relation enables one to find closed expressions for the degree distribution (Sec. II B), the ancestor-descendant correlation (Sec. II C), and the two-point function measuring the distribution of shortest path lengths between pairs of nodes (Sec. II D). In Sec. II E, we establish contact between our formalism and the growing network models. This is used in Sec. II F, where we calculate explicitly a set of observables for some simple growing networks. In Sec. III, we compare the causal model with the maximally random one. In Sec. III A, we collect all general formulas. In Sec. III B we calculate the observables, assuming the same microstate weights as in Sec. II F, and find dramatically different results. This illustrates the importance of graph symmetries. In Sec. III C, we show how to construct a maximally random model with the same degree distribution as a given causal model. We stress, however, that the identity of degree distributions does not imply that geometries are similar. In particular, the average shortest path lengths are very different: generically they scale like a power of the number of nodes N in maximally random and like $\ln N$

in causal models, respectively. We conclude briefly in Sec. IV.

B. Basic concepts

Let us recall some definitions. A *rooted* tree is a tree with one marked node. A *planted* tree is a rooted one with an extra link attached to the root, so that the degree of the root is increased by unity. The other end of this extra link is not counted as a node, in a sense this end remains ''free.'' The different tree ensembles are simply related and choosing to work with one of them is a matter of convenience. In this paper, we deal with planted tree graphs [31].

By an assumption, a label is attached to each node, and two graphs with identical topology but labeled differently are considered different. We say a tree is endowed with a *causal structure* when the labels always appear in growing numerical order as one moves along the tree from the root towards an arbitrary node. These are the tree graphs we are interested in.

We introduce a statistical ensemble of these trees. Let us denote by T a given topology and by $L(T)$ the number of distinct causal labelings of T. We attach the same weight,

$$\rho(T)=w(n_1,n_2,\ldots,n_N), \tag{1}$$

to each acceptable labeling. Here, N is the total number of nodes, n_i denotes the degree of the node i, and w is some appropriate positive-definite function. A model is defined by choosing a particular form of w.

It will be seen that the presence of a causal structure generates nontrivial observable internode correlations. Hence, it is of interest to discuss the models where these specific correlations do not interfere with correlations of a different origin. With this motivation, we assume in this paper that $\rho(T)$ factorizes [32] as

$$\rho(T)=\prod_{i=1}^{N} q_{n_i}. \tag{2}$$

On the other hand, we keep q_n as general as possible.

The partition functions of canonical and grand-canonical ensembles are defined in the usual way by summing the weights of all possible microstates. Thus, by definition

$$z_N=\frac{1}{N!}\sum_T L(T)\rho(T) \tag{3}$$

and

$$Z(\mu)=\sum_N z_N e^{-\mu N}. \tag{4}$$

The prefactor $1/N!$ in Eq. (3) is compensated by the number of terms in the summand. Indeed, the number of labeled causal trees with N nodes is $(N-1)!$ [33]. It is not difficult to convince oneself that in all cases of physical interest, z_N grows exponentially with N, up to a power prefactor:

$$z_N\sim e^{\bar{\mu}N} \quad \text{for} \quad N\to\infty \tag{5}$$

and $Z(\mu)$ develops a singularity at $\mu=\bar{\mu}$ (see Sec. II A for a more rigorous argument). One is primarily interested in the regime controlled by the singular part of $Z(\mu)$. Indeed, as $\delta\mu\equiv\mu-\bar{\mu}$ tends to zero, one becomes increasingly sensitive to the behavior of trees with an arbitrarily large N.

Of course, Eq. (4) can be inverted by the Laplace transformation

$$z_N=\frac{e^{\bar{\mu}N}}{2\pi i}\int_{-i\infty}^{i\infty} d\,\delta\mu Z(\delta\mu)e^{\delta\mu N} \tag{6}$$

with the integration contour passing on the right of the singularity at $\delta\mu=0$. It will often be convenient to work in the grand-canonical ensemble and to Laplace transform the result to the physically more interesting canonical ensemble at the very end. In practice, we shall always assume that $N\to\infty$ and keep the leading term only.

Note, that there is an analogy between the labelings of graphs and the positions of a system in a (discrete) phase space. Hence, $\rho(T)$ and $L(T)$ are the analogs of the weight of a microstate and of the corresponding phase-space volume, respectively.

Little is specific to causal trees in the content of this section. One could repeat *verbatim* the above definitions and keep the same microstate weights in the context of a different tree ensemble. However, in this new ensemble, the trees would have different symmetries and, therefore, $L(T)$ would be, in general, different. Consequently, the physics would be different too. We shall see later, comparing causal and maximally random trees, that introducing a new symmetry can change dramatically the geometry of generic graphs.

C. Observables

Let us define the observables we shall calculate in this paper. The most popular observable is the *degree distribution*. When the weight of a microstate has the factorized form, as in Eq. (2), the degree distribution π_n is given by a simple and obvious formula

$$\pi_n=N^{-1}q_n\frac{\partial \ln z_N}{\partial q_n}. \tag{7}$$

The factor N^{-1} above is included to have the distribution normalized to unity.

In the thermodynamic limit $N\to\infty$,

$$\pi_n=q_n\frac{\partial\bar{\mu}}{\partial q_n}. \tag{8}$$

The next observable is the *correlation* between node degrees, say the probability that a node has degree k when its neighbor's degree is l. In a causal tree, one of these nodes is an *ancestor* and the other a *descendant*.

A very interesting observable is the *Hausdorff dimension* d_H controlling the scaling with N of the linear size of a typical tree:

$$\langle r\rangle_N\sim N^{1/d_H}. \tag{9}$$

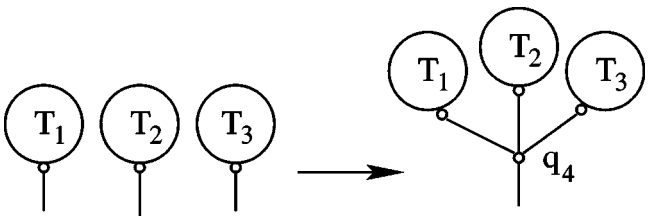

FIG. 1. Construction of the compound tree $T_1 \oplus T_2 \oplus T_3$. The three old roots are connected to the new root with degree four. Notice that all the graphs are planted trees. The trees are not planar and therefore, the relative position of branches in the compound tree is irrelevant.

One usually takes for r the distance between an arbitrary pair of nodes. In rooted trees, it is more natural to consider the distance separating a randomly chosen node from the root. One first calculates a specific two-point function $C(r,\mu)$, the grand-canonical weight of all the trees with a node separated from the root by r steps. Using $C(r,\mu)$, one finds

$$\langle r \rangle_\mu = \frac{\int_0^\infty dr r C(r,\mu)}{\int_0^\infty dr C(r,\mu)}. \tag{10}$$

This quantity usually diverges when $\delta\mu \to 0$. The behavior on the rhs (right-hand side) of Eq. (9) is determined by observing that $\delta\mu$ scales like N^{-1} [see Eq. (6)].

II. GEOMETRY OF TREES ENDOWED WITH A CAUSAL STRUCTURE

A. Recursion relation

We start by deriving a recursion relation for the partition function z_N. To this end, we construct a new planted tree by attaching the k ''free link ends'' of the planted trees $T_1, \ldots, T_k$ to a new root. We denote the resulting compound tree by $T = T_1 \oplus \cdots \oplus T_k$ (this is illustrated in Fig. 1 for $k=3$).

The number of ordered labelings of the compound tree is

$$L(T_1 \oplus \cdots \oplus T_k) = \frac{N!}{N_1! \cdots N_k!} \frac{1}{k!} L(T_1) \cdots L(T_k). \tag{11}$$

Here, N_i denotes the total number of nodes in the tree T_i and $N = \Sigma_i N_i$. One has to give $N+1$ labels to the nodes of the compound tree. However, the smallest label must be attached to the root. The remaining N labels are arbitrarily distributed among the trees. This is the origin of the multinomial factor. Permuting the trees T_i does not change the compound tree [34]. This explains the presence of the factor $1/k!$. It will be seen that the *composition rule* (11) is a very basic property of the model.

Obviously, the weight of the new tree factorizes as

$$\rho(T_1 \oplus \cdots \oplus T_k) = q_{k+1} \rho(T_1) \cdots \rho(T_k). \tag{12}$$

The partition function z_{N+1} can be constructed by summing the trees of size smaller or equal to N:

$$z_{N+1} = \frac{1}{(N+1)!} \sum_{k=1}^{\infty} \sum_{T_1, \ldots, T_k} \delta_{N_1 + \cdots + N_k, N} \times L(T_1 \oplus \cdots \oplus T_k) \rho(T_1 \oplus \cdots \oplus T_k). \tag{13}$$

Inserting Eq. (11) into Eq. (13) and rearranging the terms in the sum, we obtain after trivial algebra

$$z_{N+1} = \frac{1}{N+1} \sum_{k=1} \frac{q_{k+1}}{k!} \sum_{N_1, \ldots, N_k} \delta_{N_1 + \cdots + N_k, N} \prod_{i=1}^{k} z_{N_i}. \tag{14}$$

Adding $z_1 = q_1$ and summing both sides of Eq. (14), we get

$$\sum_N N z_N e^{-N\mu} = e^{-\mu} \left(\sum_{k=0}^{\infty} \frac{q_{k+1}}{k!} Z(\mu)^k \right) \tag{15}$$

or finally,

$$Z'(\mu) = -e^{-\mu} F(Z), \tag{16}$$

where

$$F(Z) = \sum_{k=1}^{\infty} \frac{q_k}{(k-1)!} Z^{k-1}. \tag{17}$$

Equation (16) can be integrated to give

$$e^{-\mu(Z)} = G(Z) \equiv \int_0^Z \frac{dx}{F(x)}. \tag{18}$$

The function $G(Z)$ is a positive monotonically growing function of Z, bounded from above (one can ignore the trivial case where all q_n except q_1 and q_2 are zero). Hence, μ is bounded from below. $Z(\mu)$ has a singularity at some $\mu = \bar{\mu}$. Denote by $\bar{x}$ the radius of convergence of the series $F(Z)$. The critical value of μ is given by

$$\bar{\mu} = -\ln G(\bar{x}). \tag{19}$$

This formula holds also when the radius of convergence $\bar{x}$ is infinite, since all terms in the series (17) are positive and the integral in Eq. (19) is convergent in all cases of interest: $G(\infty) < \infty$.

B. Degree distribution

The degree distribution calculated using Eq. (8) is

$$\pi_n = \frac{1}{G(\bar{x})} \frac{q_n}{(n-1)!} \int_0^{\bar{x}} \frac{dx}{F(x)^2} x^{n-1}. \tag{20}$$

Again, this formula is also valid when $\bar{x} = \infty$.

Summing over n and using the definitions of F and G, one easily checks that π_n is normalized to unity, as it should. One further finds

$$\sum_n n\pi_n = 2 - \frac{\overline{x}}{F(\overline{x})G(\overline{x})}. \tag{21}$$

On a tree, the rhs should be equal to 2 . This is the case when $F(x)$ diverges at $x=\overline{x}$. Otherwise one encounters a pathology, which looks similar to that appearing in some maximally random tree models (and in the so-called balls-in-boxes model, see Refs. [3,35]), where working in the large N limit, one misses singular node(s) contributing term(s) of the type $N^{-1}\delta(n-cN)$. In this limit such nonuniformly behaving terms do not contribute to the normalization, but do contribute to the rhs of Eq. (21). It will be shown later that when $F(\overline{x})<\infty$, the average distance between nodes is finite.

This means that singular node(s)—with unbounded connectivity—are indeed expected to show up. In deriving Eq. (20), the large N limit has been implicitly used and it is an educated guess that one again misses the singular node(s).

C. Ancestor-descendant correlation

Now, we turn to the calculation of the ancestor-descendant degree correlation. It is obvious that an ancestor plays the role of the root of the subgraph involving all its descendants. One can read from Eq. (14) the degree distribution of the root:

$$z_l(N) = \frac{1}{N}\frac{q_l}{(l-1)!}\sum_{N_1,\ldots,N_{l-1}} \delta_{N_1+\cdots+N_{l-1},N-1}\prod_1^{l-1} Z_{N_i}. \tag{22}$$

Going over to the grand-canonical ensemble, one finds

$$\frac{dZ_l(\mu)}{d\mu} = -e^{-\mu}\frac{q_l}{(l-1)!}Z^{l-1}(\mu) \tag{23}$$

which, taking Eq. (16) into account and after integration, yields

$$Z_l(\mu(Z)) = \frac{q_l}{(l-1)!}\int_0^Z dx\frac{x^{l-1}}{F(x)}. \tag{24}$$

Using similar arguments, one writes the weight of graphs, where the root has the degree l and its daughter the degree k as

$$z_{kl}(N) = \frac{q_l}{N(l-2)!}\sum_{N_1,\ldots,N_{l-1}} \delta_{N_1+\ldots+N_{l-1},N-1} \times\prod_{i=1}^{l-2} Z_{N_i} z_k(N_{l-1}). \tag{25}$$

Hence,

$$\frac{dZ_{kl}(\mu)}{d\mu} = -e^{-\mu}\frac{q_l}{(l-2)!}Z^{l-2}(\mu)Z_k(\mu). \tag{26}$$

Integrating the above equation, one finally obtains

$$Z_{kl}(\mu(Z)) = \frac{q_l}{(l-2)!}\frac{q_k}{(k-1)!}\int_0^Z dx_2\frac{x_2^{l-2}}{F(x_2)}\int_0^{x_2} dx_1\frac{x_1^{k-1}}{F(x_1)}, \tag{27}$$

which is the conditional probability, up to normalization, that a descendant has the degree k when the ancestor's degree is l. The normalization is determined summing over k on the rhs above, with the result $(l-1)Z_l(\mu)$. It is a slightly different measure of neighbor correlation than that proposed in Ref. [6], but it carries similar information. Because of the integration, the dependence on k and l does not factorize, in general.

Equation (27) holds in the grand-canonical ensemble. One would like to have an expression valid in the canonical ensemble, where the graph has a well-defined number of nodes. This requires Laplace transforming the rhs of Eq. (27) (one should transform first and normalize next). One can argue that in the thermodynamical limit and in the regime $\delta\mu\to 0$ this often amounts to just replace $Z\to\overline{x}$ in Eq. (27). Indeed, let $Z-\overline{x}\sim\delta\mu^\alpha$. Then,

$$\int d\delta\mu e^{\delta\mu N} Z_{kl}(\mu) \approx Z_{kl}(\overline{\mu})\int d\delta\mu e^{\delta\mu N - c\delta\mu^\alpha}. \tag{28}$$

Evaluating the integral by the saddle-point method, one gets an exponential of a quantity scaling like $N^{\alpha/(\alpha-1)}$. This gives a factor of unity in the limit $N\to\infty$, provided $0<\alpha<1$, a condition often met in applications.

D. Fractal dimension

Repeating over and over the iteration process leading to Eq. (27), one gets

$$Z_{k_1,k_2,\ldots,k_r}(\mu(Z)) = \prod_{j=2}^{r}\frac{q_{k_j}}{(k_j-2)!}\frac{q_{k_1}}{(k_1-1)!}\int_0^Z dx_r\frac{x_r^{k_r-2}}{F(x_r)} \times\int_0^{x_r} dx_{r-1}\frac{x_{r-1}^{k_{r-1}-2}}{F(x_{r-1})}\cdots \times\int_0^{x_2} dx_1\frac{x_1^{k_1-1}}{F(x_1)}. \tag{29}$$

Summing over node degrees $k_1,k_2,\ldots,k_r$, one obtains the weight of all graphs with a point separated by r links from the root, i.e., the two-point correlation function $C(r,\mu)$ introduced in Sec. I C:

$$C(r,\mu(Z)) = \int_0^Z dx_r\frac{F'(x_r)}{F(x_r)}\int_0^{x_r} dx_{r-1}\frac{F'(x_{r-1})}{F(x_{r-1})}\cdots \times\int_0^{x_3} dx_2\frac{F'(x_2)}{F(x_2)}\int_0^{x_2} dx_1. \tag{30}$$

For finite $\overline{x}$, replacing the upper limit of integration over x_1 by $\overline{x}$ and performing all the integrations, one gets

$$C(r,\mu(Z)) \leqslant x \frac{(\ln F(Z))^{r-1}}{(r-1)!}. \tag{31}$$

Hence, the tail of $C(r,\mu)$ falls at least as fast as a Poissonian. Consequently, $\langle r \rangle_\mu$ grows at most like $\ln F(Z)$. Assuming that $F(z)$ has at most a power singularity at $z=\overline{x}$, one concludes that

$$\langle r \rangle_\mu \leqslant \text{const} \ln \frac{1}{\delta\mu} \tag{32}$$

and therefore,

$$\langle r \rangle_N \leqslant \text{const} \ln N, \tag{33}$$

since $\delta\mu$ scales like N^{-1}. The argument is rather heuristic, but suggestive (see also the examples in the following section). It appears that generically the causal trees have the *small-world* property $d_H=\infty$, contrary to the maximum entropy trees whose generic fractal dimension is finite [3,36,37]. This phenomenon is easy to understand intuitively: the causal structure suppresses long branches. This can be seen by noting that along a branch from the root to the leaf no label permutations are possible, hence a tree with a few long branches admits much less causal labelings than a ''short fat'' one.

E. Synchronic view of growing networks

So far, our discussion was very general. Let us now establish a bridge to the popular growing network models. Of course, here we consider only those models where one constructs tree graphs. Successive nodes are attached, one at a time, the attachment probability being a function of the degree of the target node:

$$\text{Prob}(k|t) = \frac{A_k}{A(t)} \quad \text{with} \quad A(t) = \sum_k N_k(t) A_k, \tag{34}$$

where $N_k(t)$ is the number of vertices with degree k in the tree at time t. It is obvious that the tree constructed that way has a causal structure; nodes are labeled by the attachment time.

Note, that from the perspective of a model builder, the concepts of causal and growing networks are complementary rather than equivalent. A causal network is defined by specifying the microstate weights. Every growing network is causal, of course, but the weights corresponding to a given growth process can have a very complicated, nonlocal structure. And conversely, given a set of weights it is, in general, not evident what is the corresponding growth process. Only for a class of models there exists a stationary attachment kernel A_k.

For linear or shifted linear attachment kernels A_k, the normalization factor $1/A(t)$ depends only on the size of the previous configuration and is therefore the same for all trees of the same size [38]. Hence, the preferential attachment recipe is compatible with the factorization of $\rho(T)$; working in the canonical ensemble, at fixed time, we can drop the normalization factor altogether without any loss of generality and set [39]

$$q_n = q_1 \prod_{k=1}^{n-1} A_k, \quad n>1, \tag{35}$$

where q_1 is some positive constant (eventually set to 1 in explicit calculations). It is instructive to check by inspection that the graph weights, produced by the recursion relation (14), coincide with those generated by the growing network model recipe.

F. Examples

1. Barabasi-Albert model

In this model [4], $A_k=k$. Therefore, $q_n=(n-1)!$, $F(x)=(1-x)^{-1}$, and $\overline{x}=1$. Thus, $G(x)=x-\frac{1}{2}x^2$ and the solution to the Eq. (18) is

$$Z(\mu) = 1-\sqrt{1-2e^{-\mu}} \approx 1-\sqrt{\delta\mu}. \tag{36}$$

The degree distribution is found from (20):

$$\pi_n = 2\int_0^1 dx (1-x)^2 x^{n-1} = \frac{4}{n(n+1)(n+2)}, \tag{37}$$

which coincides with the solution given in Refs. [6,40].

The near-neighbor correlation is readily found from Eq. (27):

$$Z_{kl} = (l-1)\int_0^1 dx_2 (1-x_2) x_2^{l-2} \int_0^{x_2} dx_1 (1-x_1) x_1^{k-1}$$

$$= \frac{l-1}{k(k+1)(k+l-1)(k+l)}. \tag{38}$$

This is very similar to the result of Ref. [6], where a slightly different quantity has been calculated. The physical content is the same: the causal structure has induced correlations between the node degrees.

Since in this example

$$\int_0^{x_3} dx_2 \frac{F'(x_2)}{F(x_2)} \int_0^{x_2} dx_1 = \int_0^{x_3} dx_2 \left[\frac{F'(x_2)}{F(x_2)} - 1 \right] \tag{39}$$

one has

$$C(r,\mu(Z)) = \frac{[\ln F(Z)]^{r-1}}{(r-1)!} - C(r-1,\mu(Z)). \tag{40}$$

Hence

$$\langle r \rangle_\mu \sim \ln F(Z) \sim \ln \frac{1}{\sqrt{\delta\mu}}, \tag{41}$$

which implies the following scaling law:

$$\langle r \rangle_N \sim \frac{1}{2} \ln N. \tag{42}$$

2. Krapivsky-Redner model

Now [6] $A_k=k+w$, $q_n=\Gamma(n+w)$, and

$$F(x)=\Gamma(w+1)(1-x)^{-w-1}. \tag{43}$$

Of course, $\overline{x}=1$ and

$$G(1)=\frac{1}{(2+w)\Gamma(w+1)}. \tag{44}$$

Using the above and evaluating the Euler integral that appears in the present case on the rhs of Eq. (20), one obtains

$$\pi_n=\frac{(2+w)\Gamma(3+2w)}{\Gamma(1+w)}\frac{\Gamma(n+w)}{\Gamma(3+n+2w)}. \tag{45}$$

Again this reproduces exactly the result of Ref. [6]. The reader can easily calculate the correlation Z_{kl}. We skip this calculation here, because the result is not particularly instructive. The remarkable fact is the very existence of the correlation, not its particular form, which in this particular case is rather cumbersome.

The calculation of $C(r,\mu)$ is identical to that carried out for the Barabasi-Albert model, except that

$$Z=1-\delta\mu^{\frac{1}{w+2}}, \quad \delta\mu\to 0 \tag{46}$$

which implies

$$\langle r\rangle_N\sim\frac{w+1}{w+2}\ln N. \tag{47}$$

3. Constant attachment kernel

For $A_k=1$,

$$F(x)=\sum_{k=1}^{\infty}\frac{1}{\Gamma(k)}x^{k-1}=e^x. \tag{48}$$

The interest of this example is in the infinite radius of convergence of the above series: $\overline{x}=\infty$. The degree distribution is

$$\pi_n=\frac{1}{\Gamma(n)}\int_0^{\infty}dx e^{-2x}x^{n-1}=2^{-n}, \tag{49}$$

a result found by several people, including the authors of Ref. [6]. We again skip the calculation of Z_{kl}.

Since $d\ln F(x)=dx$, one finds

$$C(r,\mu(Z))=\frac{Z^r}{r!}; \tag{50}$$

and using

$$Z=\ln\frac{1}{\delta\mu}, \quad \delta\mu\to 0, \tag{51}$$

one derives

$$\langle r\rangle_N\sim\ln N. \tag{52}$$

III. CAUSAL VERSUS MAXIMALLY RANDOM TREES

A. An important tiny difference

It is instructive to consider also the well-known case of maximum entropy trees (cf. Refs. [3,36,37,41]). This will help putting the results of the preceding section in proper perspective. As in the preceding section, we start with the composition rule for trees, which in the present case, when the causality constraint is lifted, reads as

$$L(T_1\oplus\cdots\oplus T_k)=\frac{(N+1)!}{N_1!\cdots N_k!}\frac{1}{k!}L(T_1)\cdots L(T_k). \tag{53}$$

Compared to Eq. (11) the difference may seem tiny; one just has $(N+1)!$ instead of $N!$ in the numerator on the rhs. It is so because there is no causality constraint and therefore, all $N+1$ labels can be arbitrarily distributed among the trees. However, this tiny difference has rather dramatic consequences. Indeed, repeating the steps, which led from Eq. (11) to Eq. (16), one obtains

$$Z(\mu)=e^{-\mu}F(Z). \tag{54}$$

No derivative appears on the left-hand side. Instead of Eq. (18), one has

$$e^{-\mu}=H(Z)\equiv\frac{Z}{F(Z)}, \tag{55}$$

where the function $H(Z)$ plays the role analogous to that of $G(Z)$ in the preceding section. Equation (55) implies that the vertex degree distribution is

$$\pi_n=\frac{1}{F(x_*)}\frac{q_n x_*^{n-1}}{(n-1)!}. \tag{56}$$

Here, $x_*=\min(x_{max},\overline{x})$ and x_{max} is the position of the maximum of the function $H(x)$: $H'(x_{max})=0$, $H''(x_{max})<0$.

Since no derivatives appear as one goes over to the grand-canonical ensemble, one does not integrate either. Therefore, iterating the recursion relation, one finds that $Z_{kl}(\mu)$ and $C(r,\mu)$ factorize. The latter equals

$$C(r,\mu)\propto[e^{-\mu}F'(Z)]^r \tag{57}$$

as first derived by Ambjørn *et al.* [36] using diagrammatic arguments.

Let us now consider three examples, with q_n chosen as in Secs. II F.1–II F 3, respectively.

B. Examples

1. First example: $q_n=(n-1)!$

With this choice, that of Sec. II F 1, the function $H(x)$ reads as $H(x)=x-x^2$ and has its maximum at $x=1/2$. Hence, $x_*=1/2$. Since $F(x_*)=2$, the degree distribution calculated from Eq. (56) is now

$$\pi_n=2^{-n}, \tag{58}$$

as in Sec. II F 3.

It is easy to check that for small $\delta\mu$, one has

$$e^{-\mu}F'(Z(\mu))=1-\sqrt{\delta\mu} \tag{59}$$

and therefore,

$$C(r,\mu)\propto e^{-\sqrt{\delta\mu}r}, \tag{60}$$

which implies

$$\langle r\rangle_N\sim N^{1/2}. \tag{61}$$

Hence, $d_H=2$, the generic value [36].

2. Second example: $q_n=\Gamma(n+w)$

With the choice of Sec. II F 2, the function $H(x)=\Gamma(w+1)x(1-x)^{w+1}$ has its maximum at $x_{max}=1/(2+w)$, which is smaller than the radius of convergence $\bar{x}=1$ of the series $F(x)$ as long as $w>-1$. Thus, $x_*=1/(2+w)$ and

$$\pi_n=\frac{2+w}{\Gamma(1+w)}\left(\frac{1+w}{2+w}\right)^{1+w}\frac{\Gamma(n+w)}{\Gamma(n)}\left(\frac{1}{2+w}\right)^n. \tag{62}$$

The fall of the degree distribution is again exponential. One easily checks that again $d_H=2$.

3. Third example: $q_n=1$

With the choice of Sec. II F 3, the function $H(x)=xe^{-x}$ has its maximum at $x_{max}=1$. Thus, $x_*=1$ and correspondingly,

$$\pi_n=\frac{e^{-1}}{(n-1)!}. \tag{63}$$

One can easily check that this result holds for a more general family of weights $q_n=\alpha\beta^{n-1}$ independent of the values of α and β, as long as they are positive. Again $d_H=2$.

C. How to get identical degree distribution in the two models

It is easy to adjust the input parameters of the two models to obtain identical degree distributions. Suppose that in the causal model, the degree distribution is π_n. For a connected tree, one necessarily has

$$\sum_n n\pi_n=2. \tag{64}$$

Set in the maximally random model

$$q_n=(n-1)!\,\pi_n. \tag{65}$$

Obviously,

$$[H^{-1}]'(1)\equiv\sum_n (n-2)\pi_n=0 \tag{66}$$

and $[H^{-1}]''(1)>0$. Thus, $H(x)$ has a maximum at $x=1$. One easily convinces oneself that this is the only maximum of this function. Furthermore, by assumption [see Eq. (64)], the radius of convergence in $F(x)$ is $\geqslant 1$. Hence, $x_*=1$ and

$$\pi_n^{\text{random}}=\pi_n \tag{67}$$

by virtue of Eq. (56). However, the coincidence of the degree distributions does not imply that geometries are similar. On the contrary, as already emphasized, the graph linear sizes have generically a very different scaling behavior in the two models (see also some computer simulation results in Ref. [42]).

IV. SUMMARY AND CONCLUSION

The results of this paper are another illustration of the claim that the opposition between diachrony and synchrony is to a large extent an illusion, except if one is interested in very specific phenomena, such as aging, intrinsically reflecting the running of time. We have discussed the geometry of networks endowed with a causal structure using the conventional framework of equilibrium statistical mechanics. Hence, models that are usually described by specific master equations and static, maximum entropy models have been treated alike. We focused on tree graphs, because only for trees we are able to proceed analytically. We have derived general formulas describing the degree distribution, the ancestor-descendant correlation, and the probability that a node lives at a given geodesic distance from the root. Using these last results, we have shown that our causal networks have generically the small-world property, i.e., their Hausdorff dimension is infinite.

We have also compared the causal model with the maximally random one, assuming the same microstate weights. Because of different symmetry properties of the graphs—in the causal model only a subclass of labelings is allowed—the geometries are dramatically different. In the causal model the degree distribution is qualitatively different, internode correlations are induced, and the Hausdorff dimension becomes infinite instead of being finite.

ACKNOWLEDGMENTS

This work was partially supported by the EC IHP network under Grant No. HPRN-CT-1999-000161 and by Grant No. 2 P03B 096 22 of the Polish Research Foundation (KBN) in years 2002–2004. Laboratoire de Physique Théorique is Unité Mixte du CNRS UMR 8627.

[1] Recent advances in network studies are reviewed in R. Albert and A.-L. Barabasi, Rev. Mod. Phys. **74**, 47 (2002); S.N. Dorogovtsev and J.F.F. Mendes, Adv. Phys. **51**, 1079 (2002); see also S.N. Dorogovtsev and J.F.F. Mendes, *Evolution of Networks: From Biological Nets to the Internet and WWW* (Oxford University Press, New York, 2003), and references below.

[2] B. Bollobas, *Random Graphs* (Academic Press, London, 1985).

[3] Z. Burda, J.D. Correia, and A. Krzywicki, Phys. Rev. E **64**, 046118 (2001).

[4] L. Barabasi and R. Albert, Science **286**, 509 (1999).

[5] L. Barabasi, R. Albert, and H. Jeong, Physica A **272**, 173 (1999).

[6] P.L. Krapivsky and S. Redner, Phys. Rev. E **63**, 066123 (2001).

[7] S.N. Dorogovtsev, J.F.F. Mendes, and A.N. Samukhin, Phys. Rev. Lett. **85**, 4633 (2000).

[8] S.N. Dorogovtsev and J.F.F. Mendes, Phys. Rev. E **62**, 1842 (2000).

[9] S.N. Dorogovtsev and J.F.F. Mendes, Europhys. Lett. **52**, 33 (2000).

[10] S. Bornholdt and H. Ebel, Phys. Rev. E **64**, 035104 (2001).

[11] P.L. Krapivsky and S. Redner, J. Phys. A **35**, 9517 (2002).

[12] K. Klemm and V.M. Eguiluz, Phys. Rev. E **65**, 036123 (2002).

[13] A. Vazquez, M. Boguna, Y. Moreno, R. Pastor-Satorras, and A. Vespignani, e-print cond-mat/0209183.

[14] A. Vazquez, e-print cond-mat/0211528.

[15] A. Vazquez, A. Flammini, A. Maritan, and A. Vespignani, e-print cond-mat/0108043.

[16] G. Bianconi and A. Capocci, e-print cond-mat/0212028.

[17] S. Coulomb and M. Bauer, e-print cond-mat/212371.

[18] J. Kim, P.L. Krapivsky, B. Kahng, and S. Redner, Phys. Rev. E **66**, 055101 (2002).

[19] H. Ebel, J. Davidsen, and S. Bornholdt, e-print cond-mat/0301260.

[20] J. Berg, M. Lässig, and A. Wagner, e-print cond-mat/0207711.

[21] M.E.J. Newman, S.H. Strogatz, and D.J. Watts, Phys. Rev. E **64**, 026118 (2001).

[22] Z. Burda and A. Krzywicki, e-print cond-mat/0207020.

[23] M. Bauer and D. Bernard, e-print cond-mat/0206150.

[24] S.N. Dorogovtsev, J.F.F. Mendes, and A.N. Samukhin, e-print cond-mat/0204111.

[25] S.N. Dorogovtsev, J.F.F. Mendes, and A.N. Samukhin, e-print cond-mat/0206131.

[26] J. Berg and M. Lässig, Phys. Rev. Lett. **89**, 228701 (2002).

[27] R. Cohen, D. ben-Avraham, and S. Havlin, Phys. Rev. E **66**, 036113 (2002).

[28] S.N. Dorogovtsev, J.F.F. Mendes, and A.N. Samukhin, e-print cond-mat/0206467.

[29] J.H. Kim, K.I. Goh, B. Kahng, and D. Kim, e-print cond-mat/0212271.

[30] A. Fronczak, P. Fronczak, and J.A. Holyst, e-print cond-mat/0212230.

[31] A classical reference to tree graphs is G. Polya and R.C. Read, *Combinatorial Enumeration of Groups, Graphs, and Chemical Compounds* (Springer, New York, 1987); A modern and very accessible starting point for a study of analytic combinatorics is the book project by P. Flajolet and R. Sedgewick, http://algo.inria.fr/flajolet/Publications/books.html

[32] Compare this definition to that of Ref. [3], where graphs are identified with the Feynman diagrams of a *minifield theory*: Individual microstates are given the same weight in both models, provided one sets $q_n = n! p_n$, where p_n are the minifield theory couplings. However, in Ref. [3] all labelings were accepted (maximum entropy model), while in this paper we impose the causality constraint satisfied within a subclass of labelings only.

[33] The structure of a causal tree is determined if one gives the labels of the parents of each node. The root has label 1 and no parent. The node with label 2 has only one parent, the root. There are two candidates for being the parent of the node with label 3: the root and the node with label 2. There are three candidates for being the parent of the node with label 3, etc. Pursuing the argument and counting the number of distinct trees, one obtains the result quoted in the text.

[34] In comparing our formulas to those appearing in the literature, one should remember that our graphs are not assumed to be drawn in a plane. This is the origin of $1/k!$ in Eq. (11), $1/(n-1)!$ in Eq. (17) etc. The physics is unchanged when going from unconstrained to planar trees one changes the node weights: $q_n/(n-1)! \to q_n$.

[35] P. Bialas, Z. Burda, and J. Johnston, Nucl. Phys. **B542**, 413 (1999).

[36] J. Ambjørn, B. Durhuus, and T. Jònnson, Phys. Lett. B **244**, 403 (1990).

[37] J. Jurkiewicz and A. Krzywicki, Phys. Lett. B **392**, 291 (1997); erratum: last five lines of Sec. 3.2 erroneous.

[38] For other kernels, one could perhaps make the approximation done by the authors of Ref. [6], i.e., assume that $A(t) = \mu N$.

[39] Those factors may be important if one wants to consider the thermodynamic limit $N \to \infty$.

[40] S.N. Dorogovtsev, J.F.F. Mendes, and A.N. Samukhin, Phys. Rev. Lett. **85**, 4633 (2000).

[41] P. Bialas and Z. Burda, Phys. Lett. B **384**, 75 (1996).

[42] A. Krzywicki, e-print cond-mat/0110574.