Letters in Mathematical Physics (2006) 78:235–262
DOI 10.1007/s11005-006-0123-1

Complex Networks: from Graph Theory to Biology

ANNICK LESNE[1,2]
[1] *Université Pierre et Marie Curie-Paris 6, UMR 7600, Paris, 75005 France.*
[2] *Institut des Hautes Études Scientifiques, Le Bois-Marie, 35 route de Chartres, 91440, Bures-sur-Yvette, France. e-mail:lesne@ihes.fr*

Received: 31 January 2006; revised version: 12 October 2006
Published online: 22 November 2006

Abstract. The aim of this text is to show the central role played by networks in complex system science. A remarkable feature of network studies is to lie at the crossroads of different disciplines, from mathematics (graph theory, combinatorics, probability theory) to physics (statistical physics of networks) to computer science (network generating algorithms, combinatorial optimization) to biological issues (regulatory networks). New paradigms recently appeared, like that of 'scale-free networks' providing an alternative to the random graph model introduced long ago by Erdös and Renyi. With the notion of statistical ensemble and methods originally introduced for percolation networks, statistical physics is of high relevance to get a deep account of topological and statistical properties of a network. Then their consequences on the dynamics taking place in the network should be investigated. Impact of network theory is huge in all natural sciences, especially in biology with gene networks, metabolic networks, neural networks or food webs. I illustrate this brief overview with a recent work on the influence of network topology on the dynamics of coupled excitable units, and the insights it provides about network emerging features, robustness of network behaviors, and the notion of static or dynamic motif.

Mathematics Subject Classification (2000). 05C50, 60K35, 90B10, 90B15, 90C35.

Keywords. complex systems, excitable dynamics, graph theory, Markov chains, motifs, networks, percolation, random graphs, scale-free networks, statistical ensembles.

1. Introduction

The name *network* first belongs to common language; as such, it refers to many familiar and concrete structures: man-made networks like railways or roads, airlines, electric supply networks, or the Internet, social networks, biological networks (metabolic reactions, neural networks, blood circulation, food webs) and tree-like networks (hydrographical networks) with simpler properties.

Formalization follows quite straightforwardly: a network is a set V of N *vertices* pairwise connected by a subset E of *edges*; these edges can be oriented, weighted, signed, or not. The network *configuration* is defined as the set $\bar{s}(t)=[s_i(t); i\in V]$ describing the instantaneous state $s_i(t)$ of each node i (either discrete, e.g. a Boolean variable $s_i = 1$ or 0, or taking continuous values). In a mathematical context, one rather speaks of a *graph* $G=(V, E)$ with E a given subset of $V \times V$. The edges are in this case unweighted (either present or absent) but possibly

oriented if some (i, j) and (j, i) do not belong jointly to E. A network structure can be associated to any pair property or interaction within a group V of elements: two elements $i, j \in V$ sharing this property will be connected, that writes $(i, j) \in E$. What might appear as a mere re-formulation of an ensemble relationship actually proves to be very fruitful, since it makes available the geometrical concepts and results of graph theory, on the mathematical side, and the operational methods of complex network theory, on the statistical physics side.

It is worth noticing that a graph may be associated to a real space, or not. For instance, V can be a discrete set of points in the plane $\mathbf{R}^2$ or space $\mathbf{R}^3$ (think of railways or roads); in this case, the network inherits the natural distance of the underlying space, and this distance associates a length r_{ij} to each pair (i, j). But in the general case, graph definition involves abstract relationships: V is an arbitrary set of points, with a priori neither linear nor topological structure (think for instance of a social network in a given place, say a school or a company, or to a network as Internet in which the connection lengths play almost no role in practice). Edges then have no intrinsic length and the sole topology is that associated with network connectedness: the distance between any two elements is defined as the *minimal number of elementary steps in a path connecting them* (this amounts to set the distance between neighbors nodes i and j to $r_{ij} = 1$). Here shows up the interest of the notion of network in providing a novel topological structure, super-imposed to the natural topology of the system, if any, and closely related to the functional meaning of the connections. Networks exhibiting the most complex behaviors are those where this intrinsic topology of the connections and the natural topology of the underlying space coexist, because of competition, frustration or various selection effects that might ensue from this super-imposition; current examples are food webs and neural networks.

The graph is represented by its *adjacency matrix* $A = (A_{ij})_{i,j \in V}$, defined by

$A_{ij} = 1$ if there is an edge from i to j,
elsewise $A_{ij} = 0$.

A is a $N \times N$ matrix where N is the number of nodes in the graph. It is symmetric when edges are not oriented. Its diagonal elements vanish when there is no loops, i.e. no edges relating a node to itself. Within a straightforward extension, its components might take other values than 0 or 1 in order to account for weighted edges: the weight A_{ij} of the edge (i, j) might be its length in case of a network embedded in the plane, the coupling strength in case of an interaction network, or a kinetic rate in case of a reaction network. If not explicitly mentioned, I will consider only symmetric and unweighted edges.

The *degree* of a node i is the number $k_i = \sum_{j=1}^{N} A_{ij}$ of direct connections it establishes;[1] its neighborhood $\mathcal{V}_i$ and neighbors $j \in \mathcal{V}_i$ are the k_i nodes directly

[1] In case of oriented edges, one defines the *in-degree* $k_i^{\text{in}} = \sum_{j=1}^{N} A_{ji}$ (number of edges incoming in i) and *out-degree* $k_i^{\text{out}} = \sum_{j=1}^{N} A_{ij}$ (number of edges coming out of i). The elementary steps in a path have now to be consistently oriented, what is ensured by considering the non-vanishing contributions $A_{ii_1} A_{i_1 i_2}, \ldots A_{i_{n-1} j}$ in the powers A^n of the (now non-symmetric) adjacency matrix A.

connected to i. Powers of A describe paths: the component A^n_{ij} of the n-th power A^n gives the number of paths connecting i to j in n steps, where each non-vanishing contribution $A_{ii_1}A_{i_1i_2},\cdots A_{i_{n-1}j}$ (equal to 1) describes one of the possible n-steps paths $[i, i_1, i_2, \ldots, i_{n-1}, j]$ from i to j. Several other topological features can be defined: the number of minimal paths through a given node or a given edge, the fraction of triangles (i.e. the fraction of edges present among all possible links relating neighbors of a given site) and many other meaningful ones. I present and discuss the main ones in Section 4.3.

Network studies and their framework, as just sketched, let aside the trivial situation where a few specific nodes or motifs, of special nature, control the network and phenomena involving it (think of the electric plant in the power supply network or to a pace-maker neuron ruling certain neural networks); this case is akin to organization ruled by a chief, far less rich and surprising than self-organization and its emergent properties. In contrast, network theory considers the situation where nodes are identical and play equivalent roles. In this case, the basic origin of observed behavior, hence the essential controlling factor is the connection pattern, namely the very existence of a network structure. One of the challenges is to understand the impact of local properties and events on the network global emergent behavior, with an obvious interest for predicting and controlling the phenomena occurring in the network. As we shall see, this impact varies according to the considered phenomena (propagation, epidemic spreading, coupled dynamics or response to external inputs, to quote but a few).

2. Graph Theory and Markov Chains

Graph theory is too wide-ranging a domain of mathematics to give even a brief overview within the scope of this presentation. I will focus on a specific a specific point, namely the inter-relations between graph theory and Markov chains describing random and memoryless transitions between discrete states: each of these states corresponds to a node of the graph, and transitions correspond to its edges. We shall see that this mapping is very fruitful, on the one hand to use graph theory framework and methods to establish Markov chain properties, on the other hand to explore some of the topological properties of a given graph, for instance its community structure, by means of a random walk generated by a suitable Markov chain.

2.1. CYCLES AND MARKOV CHAINS

An important notion, qualitatively but also technically, is that of *cycle*, namely a closed path on the network (it is a sub-graph of $n \geqslant 3$ nodes and the same number n of edges), also called a *loop* in physical literature.A cycle $(i_0, i_1, \ldots, i_n = i_0)$ is said to be *elementary* if it is not possible to make a closed short-circuit, i.e. if it does not exist a link $(i_q, i_{q'}) \in E$ with $q' > q + 1$ providing a short-cut reducing the size n of the cycle to $n + q + 1 - q'$ steps. A closely related notion is that of *maximal spanning tree* i.e. a connected subgraph $\mathcal{T}$ of N nodes without any cycle; it is maximal

insofar as adding any of the edges let apart creates a cycle: a maximal spanning tree thus contains $N-1$ edges. The edges of a spanning tree $\mathcal{T}$ can be oriented unambiguously by choosing a root, namely a site i towards which they are consistently directed, yielding an oriented tree $\mathcal{T}_i$ for any of the N possible choices for i. Any spanning tree $\mathcal{T}$ generates a family $[\mathcal{C}_\alpha]_\alpha$ of elementary cycles, where each cycle $\mathcal{C}_\alpha$ is obtained by adding one of the edges l_α not belonging to the tree $\mathcal{T}$ and providing the additional edge required to close a path in $\mathcal{T}$. These cycles have varying lengths n ranging between the obvious bounds $n=3$ and $n=N$. Such a family of $M-N+1$ elementary cycles provided a basis in the sense that [34]:

THEOREM 1. *Any cycle $\mathcal{C}$ can be decomposed into a linear superposition $\sum_{\alpha=1}^{M-N+1} a_\alpha(\mathcal{C})\mathcal{C}_\alpha$ of elementary cycles with $a_\alpha(\mathcal{C})=0, -1$ or 1. This expression means that the contribution associated to a given edge (i,j) comes from cycles $\mathcal{C}_\alpha$ containing (i,j) or (j,i), with $\sum_{\{\alpha \mid (i,j)\in\mathcal{C}_\alpha \text{ or } (j,i)\in\mathcal{C}_\alpha\}} a_\alpha(\mathcal{C})$ equals 1 if (i,j) belongs to $\mathcal{C}$, -1 if (j,i) belongs to $\mathcal{C}$, 0 if neither (i,j) nor (j,i) belongs to $\mathcal{C}$.*

A transition matrix $R=[R_{i\to j}]_{i,j=1,\ldots,N}$ on V (i.e. any square matrix with $R_{ij}\geqslant 0$ and $\sum_i R_{ij}=1$) defines a graph according to $A_{ij}=1$ if $R_{ij}>0$, else $A_{ij}=0$ (if $R_{ij}=0$). This graph has symmetric edges as soon as R_{ij} and R_{ji} are jointly vanishing, or jointly non-vanishing. Following a method basically introduced by Kirchhoff in 1847, the stationary distribution of the Markov chain generated by R (with states $i\in V$) can be related to the features of this graph, namely to the whole set[2] $[\mathcal{T}^{(\mu)}]_\mu$ of maximal spanning trees:

THEOREM 2. *For any node $i=1,\ldots,N$, let $[\mathcal{T}_i^{(\mu)}]_\mu$ be the family of oriented maximal spanning trees derived from the set $[\mathcal{T}^{(\mu)}]_\mu$ by fixing a root i. Each tree $\mathcal{T}_i^{(\mu)}$ is associated with an algebraic value $b(\mathcal{T}_i^{(\mu)})$ defined as the product of all transitions probabilities R_{jl} such that the edge (j,l) is present in the oriented graph $\mathcal{T}_i^{(\mu)}$. Then the stationary distribution of the Markov chain generated by R writes $p_i^{(\text{stat})}=S_i/S$ where $S_i=\sum_\mu b_\mu(\mathcal{T}_i^{(\mu)})$ and $S=\sum_{i=1}^N S_i$.*

This expression straightforwardly implies the following fundamental result [34]:

THEOREM 3. *Let R be a transition matrix ($R_{ij}\geqslant 0$, $\sum_j R_{ij}=1$) assumed to be*

(i) *irreducible: any state i can be related to any other j in a finite number of steps (i.e. there exists an integer n_{ij} such that $R_{ij}^{n_{ij}}>0$) or equivalently the eigenvalue $\lambda_0=1$ is simple.*

(ii) *aperiodic, namely all eigenvalues except $\lambda_0=1$ have a modulus strictly lower than 1;*

Then the Markov chain generated by R possesses a unique stationary distribution $p^{(\text{stat})}$ with strictly positive components: $p_i^{(\text{stat})}>0$ for all $i=1,\ldots,N$.

[2]The number $\mu_{\max}$ of maximal spanning trees depends on the number N of nodes and M of edges of the graph, but in a way varying with its topology, so that no general formula is available.

These now quite old and acknowledged theorems and their proofs illustrate the deep and operational relation between graph theory and Markov chain theory, the former providing demonstrative and constructive tools to the latter.

2.2. RANDOM WALKS TO EXPLORE A GRAPH

Conversely, a given graph (V, E) and its adjacency matrix A define a random walk in V, with elementary steps in E, prescribed by the transition matrix[3] $R_{i\to j} \equiv R_{ij}$

$$R_{ij} = \frac{A_{ij}}{k_i} \quad \text{with} \quad k_i = \sum_{j=1}^{N} A_{ij} \quad \text{being the out-degree of } i. \tag{1}$$

By construction $R_{ij} \geqslant 0$ and $\sum_{j=1}^{N} R_{ij} = 1$. It thus defines a bona fide transition matrix: R_{ij} describes the conditional probability of a step from i to j knowing that the walker is in i before doing the step.The benefit of introducing this transition matrix R follows from the fact that the exploration of V by the associated random walk evidences some of the graph topological properties, for instance the presence of densely connected regions. Indeed, one can show that such regions, called *communities*, correspond to meta-stable domains in which the random walk remains trapped for long [16,38]. Moreover, these meta-stable features appear as a *transient ergodicity breaking* and accordingly reflect in the spectrum of R: their spectral signature is the existence of eigenvalues λ_a very close to 1. We state below (Section 2.3) in more details the consequences on the transient and asymptotic dynamics of such a spectral feature. In the present context, it is enough to know that the right eigenvectors[4] associated with these quasi-degenerate eigenvalues λ_a are quasi-piece-wise constant. The regions in V where they take almost the same value are invariant upon the evolution law up to times $t \ll \tau_a \sim (\log 1/\lambda_a)^{-1}$ (in other words, one observes mostly internal transitions up to time τ_a) leading, according to the very definition of R, to identify these regions with communities. *The random walk generated by R thus appears as an exploratory tool allowing to probe quantitatively the graph topology.*

By construction, R gives the same weight $1/k_i$ to the k_i links originating from the node i; in the case when A is symmetric, the asymptotic probability of the node i (equivalently the frequency of visit in the stationary regime) is proportional to its degree: $p_i^{(\text{stat})} = k_i / \sum_j k_j = k_i/2M$ where M is the total number of edges. The iterates of R provide the *relative weights* of the different paths (recall that, by contrast, the iterates of A provide the *number* of different paths). This approach

[3]Both notations $R_{j\leftarrow i}$ and $R_{i\to j}$ are encountered for the conditional probability of a transition from state i to state j. The quality of the notation $R_{i\to j}$ and its iterates R^n_{ij} is to exhibit the followed path in our usual way of reading, from left to right, the advantage of the reverse notation is to lead to the usual matrix-theoretic notation $p_t = R^t p_0$ for the evolution of a probability distribution p_0 on V. I use here the first one $R_{ij} \equiv R_{i\to j}$, besides the one currently used for adjacency matrix $A_{ij} \equiv A_{i\to j}$.

[4]Respectively, left eigenvectors when adopting the notation $R_{j\leftarrow i}$ as in [16].

can be generalized to weighted edges, defining now the transition matrix as $R_{ij} = A_{ij}/\sum_l A_{il}$ then carrying the analysis as in the unweighted case, within the general framework of Markov chains.

2.3. SPECTRAL ANALYSES

Given a graph, *three different spectral analyses can be performed*. Without entering technicalities, I will thus compare the spectral analysis of

- the adjacency matrix A (defined in Section 2.1);
- the matrix $\Delta = A - K$ where $K_{ij} = k_i \delta_{ij}$ is the diagonal matrix of degrees;
- the transition matrix $R = K^{-1}A$ (previously mentioned in Section 2.2).

The spectra of the matrices A, Δ and R are related in a simple way *only if the graph is regular*, that is, when all nodes have the same degree $k_i \equiv k$; their eigenvalues then satisfy

$$\lambda_a(A) = k\,\lambda_a(R) = \lambda_a(\Delta) + k. \tag{2}$$

In the general case, these three spectra differ and provide different informations on the graph, all relevant both for the determination of its statistical properties and the choice of the best-suited class of models ('statistical ensemble') in which considering the genericity of its properties.

It follows from the normalization $\sum_j R_{ij} = 1$ that R possesses an eigenvalue $\lambda_0(R) = 1$. It corresponds to the eigenvalue $\lambda_0(\Delta) = 0$ of Δ (whose existence directly follows from the relations $\sum_j \Delta_{ij} = 0$). In the special case of a regular graph, it implies that A possesses an eigenvalue $\lambda_0(A) = k$ giving the common value k of the degrees.

Δ is called the *graph Laplacian matrix* because it corresponds to a discrete version of the Laplacian on the graph. To explain this point, let us consider an array $x_i = ai$ where each node has two neighbors $j = i \pm 1$; it comes $\sum_j \Delta_{ij}\phi(x_j) = \phi(x_{i+1} + \phi(x_{i-1} - 2\phi(x_i) \approx a^2\phi''_{xx}(x_i)$ since here $k_i = 2$. On a more general graph, the analog of the Laplace diffusion equation writes[5]

$$\phi_i(t+1) = \frac{1}{k_i}\sum_j A_{ij}\phi_j(t) \quad \text{i.e.} \quad \phi(t+1) = R\phi(t). \tag{3}$$

Not surprisingly, we recover the Markov process generated by the transition matrix $R = K^{-1}A$. The above graph diffusion equation (3) thus describes the relaxation of any initial distribution $\phi(0)$ to the stationary state $p^{(\text{stat})}$ of the Markov

[5]The graph analog of Laplace diffusion equation only makes sense in finite time due to the discreteness of the jumps from a site i to its neighbors j. It appears as a space-time discretization of the diffusion equation $\partial\phi = (R-1)\phi$ where $R - 1 = K^{-1}\Delta$.
But the solution of this equation would exhibit a discretization effect, originating from the gap between $e^{t(R-1)}$ and R^t: the proper relaxation equation is (3). The influence of discretization and specificities of the discrete Laplacian have been thoroughly investigated on regular grids, see e.g. [8].

chain, corresponding to the eigenvector with eigenvalue $\lambda_0(R)=1$ of R and eigenvalue $\lambda_0(\Delta)=0$ of Δ.The spectral analysis of R provides many other direct informations, both on the asymptotics and the transients of the random walk it generates:

- as previously mentioned, if the eigenvalue $\lambda_0(R)=1$ is simple,then R is *irreducible*, namely each element of V can be related to any other one by a finite path. There is accordingly a *unique* stationary distribution $p^{(\mathrm{stat})}$. Irreducibility thus appears as the *stochastic analog of an ergodicity property* for the evolution generated by R;
- if all non-trivial eigenvalues have a modulus strictly lower than 1 ($|\lambda_a(R)|<1$), R is said to be *aperiodic*. Joint irreducibility and aperiodicity of R ensure that any initial distribution of probability p_0 on V *converges* to the unique stationary distribution $p^{(\mathrm{stat})}=\lim_{t\to\infty} R^t p_0$;
- when A is symmetric, the occupancy of a node i in the stationary regime is directly proportional to its degree k_i, namely $p_i^{(\mathrm{stat})}=k_i/\sum_j k_j$. It is straightforward to see that R then satisfies the *detailed balance property*:[6] $p_i^{(\mathrm{stat})}R_{ij}=p_j^{(\mathrm{stat})}R_{ji}$ for any pair of sites (i,j). It means that once the stationary regime is reached, there are as many transitions from i to j than from j to i; in consequence, this stationary regime is in fact an *equilibrium state* (with no probability currents);
- having sorted the eigenvalues $[\lambda_a(R)]_{a\geqslant 1}$ of R according to decreasing modulus, the times $[\tau_a]_a$ defined by $\lambda_a=\mathrm{e}^{-1/\tau_a}$ yield the *relaxation times* towards the asymptotic distribution $p^{(\mathrm{stat})}$; for a typical initial distribution, this relaxation is controlled by the largest time τ_1;
- the characteristic time τ_1 also quantifies the *correlation time* of typical observables A (that is, whose projection $\langle A|q_1\rangle$ on the eigenvector q_1 associated to λ_1 does not vanish). Intermediary times τ_a with $a\geqslant 2$ describe the correlation times of non typical observables A (such that $\langle A|q_1\rangle=0$).

More can be said in the case when some (say, n) eigenvalues of R are close to 1, hence being associated with slow modes: $\tau_a\gg 1$ for $a=1,\ldots,n$. We have shown that such a quasi-degeneracy of R induces a multi-scale dynamic structure in the state space V, all the more when a definite *scale separation* $\tau_1\gg\tau_2\gg\cdots\gg\tau_n\gg 1$ leads to clear-cut hierarchical features [16]. The proof relies on the additional assumption, satisfied in all generic situations, that R admits a *spectral decomposition*:

$$R_{ij}=\sum_{a=0}^{N-1}\lambda_a B_i^{(a)} q_j^{(a)} \quad \text{hence} \quad (R^t)_{ij}=\sum_{a=0}^{N-1}\lambda_a^t B_i^{(a)} q_j^{(a)} \quad (i,j=1,\ldots,N) \qquad (4)$$

[6]This property is also known as *microreversibility* since it ensures the invariance of the Markov chain upon time-reversal.

where $[B^{(a)}]_a$ and $[q^{(a)}]_a$ are, respectively, the right and left eigenvectors of R, with $q^{(0)} = p^{(\mathrm{stat})}$, $B_i^{(0)} \equiv 1$ and $\sum_{i=1}^{N} B_i^{(a)} q_i^{(a')} = \delta_{aa'}$. This expansion clearly shows that the contribution of the mode a to the dynamics decreases in time as e^{-t/τ_a}.

The key preliminary result is the piecewise constancy, at leading order[7] with respect to the scale separation factor, of the n right eigenvectors associated with the n quasi-degenerate eigenvalues. More precisely, we can partition the state space in regions

$$V^{(n)}_{b_1,\ldots,b_n} = \{i \in V \mid B_i^{(a)} \approx b_a \text{ for } a = 1, \ldots, n\} \tag{5}$$

(possibly leaving aside some 'dust') and also consider the coarse-grained versions:

$$V^{(n-1)}_{b_1,\ldots,b_{n-1}} = \cup_{b_n} V^{(n)}_{b_1,\ldots,b_n} \quad \text{up to} \quad V^{(1)}_{b_1,\ldots,b_{n-1}} = \cup_{b_2,\ldots,b_n} V^{(n)}_{b_1,\ldots,b_n}. \tag{6}$$

These nested partitions are adapted to the dynamics in the following sense:

THEOREM 4. [16] *For times $t \ll \tau_a$, the regions $V^{(a)}_{b_1,\ldots,b_a}$ are quasi-invariant: the trajectory of a random walker remains in the region where it starts (the index a here varies in $\{1, \ldots, n\}$). τ_a appears as the characteristic time of the transitions from a region $V^{(a)}_{b_1,\ldots,b_{a-1},b_a}$ to another one $V^{(a)}_{b_1,\ldots,b_{a-1},b'_a}$ within one and the same $V^{(a-1)}_{b_1,\ldots,b_{a-1}}$. For $\tau_{a+1} \ll t \ll \tau_a$, the probability distribution has reached a quasi-stationary state within each $V^{(a)}_{b_1,\ldots,b_a}$, following the relaxation to 0 of the modes with $a' > a$ since then $\mathrm{e}^{-t/\tau_{a'}} \ll 1$.*

The spectral analysis of R thus provides a whole hierarchical picture of the state space and its exploration by the random walk generated by R. To conclude, I underline a general insight from this study: the coincidence of the correlation times, the relaxation times and the characteristic times of transition (inverse transition rates between meta-stable regions), moreover all directly related to the eigenvalues of R.

2.4. LANDSCAPE EXPLORATION

Given a discrete set V of elements, belonging or not to a real space, a *landscape* is defined as a function Φ on V with real positive values. In this context, the edges $(i, j) \in E$ of a graph (V, E) can be seen as a set of elementary moves on the landscape. As explained in Section 2.2, the graph defines a random walk in V, whose transition matrix R is obtained through the proper normalization of the graph adjacency matrix. By restricting the set of points j that can be reached

[7] All the given statements involve some level of approximation, as indicated in using the adverb 'quasi' and $\approx$ instead of $=$ in the formulas. The associated tolerance, controlled by the fast modes $a > n$ with $\tau_a = \mathcal{O}(1) \ll \tau_n$, could be written explicitly and bounded. We here let apart these (quite cumbersome) technicalities; it is enough to know that this tolerance tends to 0 when the scale separation factor between the fast and slow modes increases to ∞.

from i, the definition of E allows to account for various constraints (supplementing the energetic constraints associated with the landscape) for instance conservation laws, steric hindrance or topological invariants. As detailed below, it also allows to monitor the exploration, e.g. enforcing large steps or directed motion. I shall here present how such an 'a priori' exploration of V provides insights about the landscape features.

The *correlation time* $t_{\mathrm{corr}}(\Phi, R)$ is defined as the characteristic range of the correlations between the values taken by Φ along a trajectory of the random walk generated by R ('a priori motion'): it is roughly the number of steps that the random walker needs to reach a region whose altitude is no longer correlated with the starting point altitude. It is thus a joint characteristic of the landscape Φ and the graph (V, E) (or equivalently R). To compute this time from the knowledge of a recorded or simulated trajectory $(i_t)_{t=1,\ldots,T}$ with T steps, one exploits an ergodicity argument, allowing to estimate statistical averages[8] by identification with temporal averages along a typical realization of the random walk:

$$C_{\Phi,R}(\theta) \equiv \langle \Phi(i_{t+\theta})\Phi(i_t)\rangle - \langle\Phi\rangle^2 \quad \text{(independent of } t\text{)}$$
$$\approx \frac{1}{T-\theta}\left[\sum_{t=1}^{T-\theta}\Phi(i_t)\Phi(i_{t+\theta})\right] - \left[\frac{1}{T}\sum_{t=1}^{T}\Phi(i_t)\right]^2 \sim \mathrm{e}^{-\theta/t_{\mathrm{corr}}(\Phi,R)} \qquad (7)$$

(such an identification being exact in the limit as $T \to \infty$ and provided R is irreducible). The *landscape correlation length* $l_{\mathrm{corr}}(\Phi)$ is then defined as the mean-square distance traveled during $t_{\mathrm{corr}}(\Phi)$ steps [43]. Whereas t_{corr} depends on R, $l_{\mathrm{corr}}(\Phi)$ is an intrinsic characteristic of the landscape Φ and it gives an estimate of the size of its wells [15].

The set E of elementary steps associated with a graph (V, E) might be exploited as a library of *test-moves* ('a priori motion') in a Monte Carlo[9] simulation. In such a simulation, the landscape super-imposes an *energetic acceptance criterion*, for instance the Metropolis criterion: if $\Phi_j \leqslant \Phi_i$, the test-move $i \to j$ is accepted with probability 1; if $\Phi_j \geqslant \Phi_i$, it is accepted with probability $\mathrm{e}^{-\beta(\Phi_j-\Phi_i)}$, ensuring that V is sampled with a weight proportional to $\mathrm{e}^{-\beta\Phi}$ where β is a tunable parameter (the inverse temperature in statistical mechanics). The analysis of the transition matrix R of the graph provides clues on the test-motion properties for instance its ergodicity if R is irreducible.

In this context, we have proposed *a novel Monte Carlo algorithm to improve the landscape exploration* in simulations aiming at exploring or sampling V. Indeed,

[8] Here, the statistical average would be an average over V weighted by the stationary distribution $p^{(\mathrm{stat})}$ of R.

[9] Conversely, any dynamic Monte Carlo method in a state space V defines a graph whose nodes are the elements of V and whose links correspond to the test-moves of the simulation.

steps of size far larger than $l_{\mathrm{corr}}(\Phi)$ presumably[10] carry the trajectory from a well to another one, preventing it to remain trapped in a local minimum of the landscape: this is an invaluable property when looking for the ground state or for sampling the whole space V. When looking for an absolute minimum of Φ, it is not enough to visit each well; it is also necessary to explore the well enough precisely to determine the bottom-value of Φ. Exploring each well is also required for sampling purposes. In consequence, what is needed is an *alternation of local steps, inside the wells, and steps of long range* $l > l_{\mathrm{corr}}(\Phi)$, achieving transition from a well of Φ to another.The weight reflecting the difficulty to climb the landscape hill thus super-imposes, in a multiplicative fashion, to the basic weight of the test-moves (inversely proportional to the degree of the site they come from). The ensuing exploratory mechanism is still a Markov chain, and the hierarchical space-time structure of the landscape reflects in the spectrum of the associated transition matrix (a product of R with the transition matrix associated with the energetic acceptance criterion) as exposed in Section 2.3. Such a 'nested Monte Carlo' algorithm is specially well-suited in situations exhibiting strong geometric constraints (e.g. steric hindrance, as in granular media and polymer systems, or topological invariants, as with elastic rods or their microscopic analogs like DNA or actin filaments) [21].

3. Random Graphs

3.1. ERDÖS–RENYI MODEL

Given a set V of N points, the simplest model of random graph, developed by Erdös and Renyi [10] amounts to establish a link with a probability p between any two points of V, independently for all the different pairs of points. The adjacency matrix A has thus random coefficients, statistically independent and equal to 1 with probability p; in other words, they are *independent random Boolean variables*. Here graph theory meets random matrix theory.[11]The topology is fully prescribed by the set of edges E: *there is no longer an underlying real space* and the neighborhood of a site i is composed of all the sites directly connected to it. The distance between two sites is given by the number of steps of the minimal path connecting them. The network diameter is the average distance between pair of sites.

A variant of this random graph model is rather parametrized by the number M of links, placed at random; in other words, it amounts to choose M pairs among the $N(N-1)/2$ possible ones. This variant coincides with the first model in the

[10]More precisely, the 'correlation length conjecture' (rigorously proved in some special situations) states that a region of linear size $l_{\mathrm{corr}}(\Phi)$ typically contains one minimum of the landscape Φ [40].

[11]This theory, with more physical motivations than graph theory, recently developed mainly in the context of quantum mechanics and in particular of quantum chaos, dealing with the quantum behavior of systems whose classical analogs are chaotic [19].

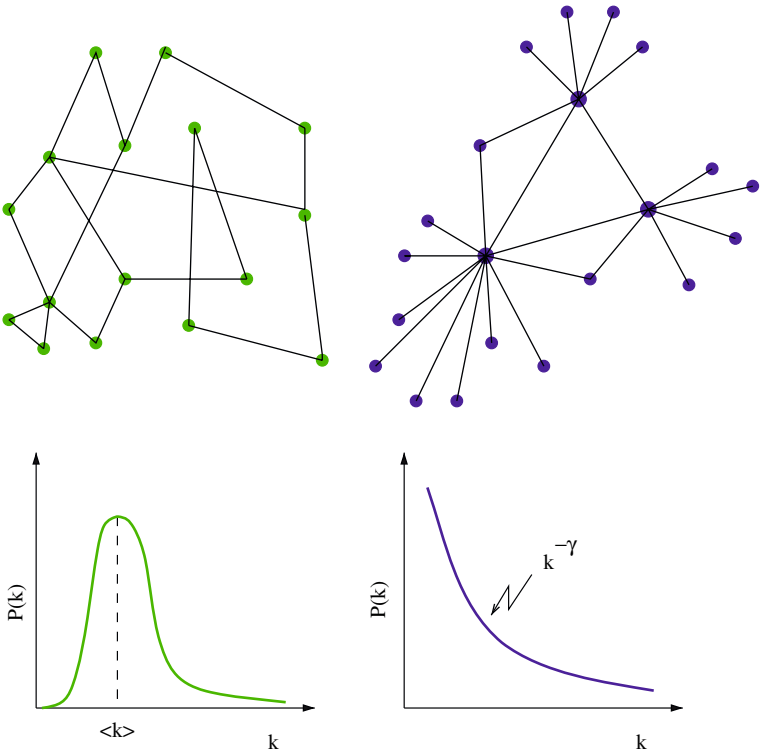

Figure 1. Comparison of random graphs (*left*) and scale-free networks (*right*). The first ones are homogeneous, with a degree distribution $P(k)$ peaked around $\langle k \rangle$ and decreasing faster than an exponential law (as $e^{-k \log k}$), while the second ones have a highly heterogeneous structure, exhibiting high-degree nodes and also more low-degree nodes (e.g. $k = 1$), with a power-law degree distribution $P(k) \sim k^{-\gamma}$ with $\gamma > 1$ (currently $2 < \gamma < 3$ in real complex networks, corresponding to a finite average degree but an infinite variance).

limit as $N \to \infty$ *at fixed* $p = 2M/(N-1)$. In both cases, the average degree $\langle k \rangle = 2M/N = Np$ corresponds to the typical degree value.

An important network characteristic is the *degree distribution* $P(k)$, describing the probability for a node to be directly connected to k other ones. In the limit $N \to \infty$ *now at fixed* $\langle k \rangle = Np$, the Erdös–Renyi model is shown to exhibit a Poisson degree distribution peaked around the average degree $\langle k \rangle$ and decreasing at large k *faster than an exponential*[12] (Figure 1):

$$P(k) = \frac{\lambda^k e^{-\lambda}}{k!} \quad \text{with} \quad \lambda = \langle k \rangle \tag{8}$$

[12]Recall that the Poisson distribution $P_\lambda(k)$ approaches a Gaussian for large λ; more precisely, the variable $(k-\lambda)/\sqrt{\lambda}$ tends in law to a normal distribution (centered of variance 1) when $\lambda \to \infty$.

3.2. CAYLEY TREES AND INFINITE-DIMENSIONAL GRAPHS

One can extend plane ($d = 2$, e.g. a square grid) or spatial ($d = 3$, e.g. a cubic mesh) lattices and consider abstract graphs, whose connections are not embedded in a real space.It is for instance the case of a *Cayley tree* (or *Bethe lattice* for the physicists), originating from a central point from which radiate z edges, then built recursively: each node is related to $z - 1$ new nodes ignoring each another, and in particular not connected. The degree of each node is thus z (leaving apart the newly added nodes of degree 1) and all nodes are equivalent in the infinite-size limit $N \to \infty$. This seemingly artificial model plays an important role: due to the absence of coupling between nodes not directly related along the tree, it lends itself to exact analytical approaches of several issues (for instance the determination of percolation threshold, see below Section 3.3). In graphs with closed paths, termed cycles, these approaches can be traveled only at the price of a *mean-field approximation*, that consists precisely in ignoring the presence of cycles, thus reducing the original graph to a Cayley tree. The latter appears as a *universal model* approximating all graphs where one expects that correlations following from cycles play a negligible role (of course a point to be checked quantitatively, at least a posteriori).

An equivalent viewpoint is to consider a Cayley tree as a lattice of connectivity z in *infinite dimension*:[13] general theoretical arguments, following from the theory of critical phenomena and renormalization group, invoke this infinite value $d = \infty$ of the dimension in support of the mean-field approach validity [24].

3.3. PERCOLATION NETWORKS

Percolation networks were introduced as simple (actually the simplest) models of *disordered media*. They involve regular lattices (e.g. square, hexagonal, or triangular grids in $d = 2$, cubic grids in $d = 3$). Random holes are introduced by breaking the bonds of the regular grid with a probability $1 - p$ (or replacing their nodes with holes with a probability $1 - p$, in case of site percolation). As in the Erdös-Renyi model, the presence or absence of a link is independent of the status of the other links. This kind of spatially discrete model with independent Boolean local state variables is specially well suited for numerical studies. One then observes remarkable statistical properties, namely a *percolation transition* when varying the control parameter p, corresponding to the appearance of a connected component spanning the whole network.

[13] In finite dimension d, a regular lattice of connectivity z exhibits cycles, for instance of length 4 on a square lattice ($z = 4$) or of length 3 on a triangular lattice ($z = 6$). One shows that the fraction of cycles among the paths coming out a given node decreases as the space dimension d increases, tending to 0 as $d \to \infty$.

In the limit[14] $N \to \infty$, for $p \geqslant p_c$, this spanning cluster contains an infinite number of nodes and is currently named 'the infinite cluster'. Right at the *percolation threshold* $p = p_c$ (and of course below, for $p < p_c$) it contains a vanishing fraction $P(p_c) = 0$ of links; this fraction $P(p)$ becomes strictly positive for $p > p_c$, providing a relevant order parameter and endowing the percolation transition with a continuous, *second-order-like transition* nature. Accordingly, the transition is described by scaling laws, e.g.

$$P(p) \sim (p - p_c)^{\beta} \quad (p \geqslant p_c) \tag{9}$$

and the infinite cluster exhibits a fractal nature at the percolation threshold. Several analytical approaches have been developed to quantitatively describe the percolation transition, mainly its threshold p_c and the scaling exponents: phenomenological scaling theories, mean-field approximation, renormalization methods [24]. They exemplify the fruitfulness of tools, concepts and methods of statistical mechanics, beyond their original application to molecular systems within the framework established by Boltzmann and Gibbs.

The Erdös–Renyi random graph model can be seen as an infinite-dimensional extension of bond-percolation models: all sites are a priori accessible by a direct link from a given node (not only the $z = 2d$ nearest neighbors as on a hypercubic lattice in dimension d) but only a fraction p of links are actually present. One shows that the percolation threshold is reached[15] when the average degree is $\langle k \rangle = 1$. The relation $\langle k \rangle = Np$ then gives the finite-size percolation threshold $p_c = 1/N$ of the Erdös-Renyi model, vanishing in the infinite-size limit, and the infinite cluster reduces to a spanning tree. The random graph percolation threshold can be recovered with a qualitative argument, identifying the percolation threshold of a N-site random graph with that, equal to $p_c = 1/(z-1)$, of a Cayley tree where the number of neighbors of a given site when all bonds are present is $z = N$ [24].

3.4. SUCCESSES AND FLAWS OF RANDOM GRAPH MODELS

The Erdös–Renyi model offered the first model of random graph in which the real space where the sites lie (if any) plays no role. Its topology is exclusively prescribed by the connection pattern. As such, it was adopted as a paradigm, having allowed to develop concepts (degree distribution, diameter) and several methods to investigate random networks [6]. It has proven to be successful in several applications, for instance (the list is by far only a sample) in devising optimization algorithms for communications or resources allocation, in pattern recognition, in

[14]Note that a bona fide transition occurs only in this limit, as for any other phase transition; otherwise the change is smooth and the threshold $p_c(N)$ is still random, what is referred as 'finite-size effects'.

[15]The general condition writes $\langle k^2 \rangle \geqslant 2\langle k \rangle$, yielding $\langle k \rangle_c = 1$ in a graph with a Poisson degree distribution [32].

clustering and classification methods, or in probabilistic methods used to prove graphs properties like e.g. the minimal number of edge crossings over all drawings of a simple graph on the plane.

Nevertheless, the relevance of the Erdös-Renyi model as a formal and ideal representation of *real* complex networks has been deeply questioned for a short decade [1,12]. A real network of N nodes and M links can be confronted to a random graph having on the average the same number of links, i.e. a random graph with N nodes defined by the probability $p = 2M/N(N-1)$ or more directly the variant model defined by its number M of links. One then sees that the *random graph model does not account, by far, for the properties of most real complex networks*, for instance metabolic networks, food webs, social networks or the Internet. In particular, real networks have a degree distribution $P(k)$ decreasing far more slowly at large k values, typically according a *power law* (see Figure 1):

$$P(k) \sim k^{-\gamma} \quad (\text{with } \gamma > 1). \tag{10}$$

This power-law dependence reflects the absence of a typical value for the degree, and such networks are currently termed *scale-free networks*; their main properties will be presented in Section 4. Notwithstanding this gap, the Erdös-Renyi model remains a reference point with thoroughly known statistical and topological properties. Observing different properties in a real network evidences by contrast that it essentially differs from a random graph and that other ingredients should be taken into account in its modeling.

4. Modeling Real Complex Networks

4.1. STATISTICAL MECHANICS OF NETWORKS

The determination of 'macroscopic' collective behavior of a large number of 'microscopic' elements lies at the very core of statistical mechanics. In the systems usually considered in statistical mechanics, interactions between these elements are always simple and homogeneous, even in critical phenomena. Let us quote for instance ferromagnetic interactions between spins, excluded-volume interactions between the monomers of a polymer chain in solution, short-range (diffusive coupling where each element is coupled to its nearest neighbors) or long-range (global coupling where each element is coupled to any other one) interactions between oscillators in coupled map lattice models. Interaction networks are then simple and regular, being either square or cubic lattices (or the more isotropic triangular or hexagonal variants), either infinite-dimensional graphs where each node is connected to each other. Often, they simply result from a discretization of the real space where the considered phenomenon occurs.

Interactions are not always so simple in real systems, and they can form networks with irregular and complicated architecture. In biology, for instance, interactions can be relayed by various means, possibly being non-local, oriented,

involving delays, or even with different spatial ranges in the same system. Current examples are gene networks, metabolic networks [22], neural networks [20], food webs [33], communication networks as the World Wide Web or the Internet [32] and social networks [41].

Statistical mechanics of networks is a way of studying such complex systems, focusing on the connection pattern and its topological and statistical properties. It extends the spirit of percolation models to more complex topologies, often self-organized, and in which real space and associated natural distance might be absent or of little relevance. A first benefit of this formalization is to provide indices allowing to quantify the network complexity. But the main interest is to sort real networks into classes, defined by a set of typical properties and providing the relevant *statistical ensembles*. In this context, a statistical ensemble gathers all graphs sharing some prescribed ingredients, e.g. the degree distribution or a growth mechanism [32]. The theoretical issue is then to determine the properties ensuing from these basic ingredients, either for prediction purposes in situations where the relevance of the statistical ensemble is well assessed, either to validate its relevance by confronting the predictions and the available experimental data; a discrepancy would evidence the need of including additional features in the definition of the statistical ensemble to get an adequate model of the real network under investigation.

4.2. SCALE-FREE NETWORKS

Following from extensive experimental data analysis, it recently appeared that the relevant paradigm to describe real networks is rather that of *scale-free networks*. In such networks, the average number of links per node does not at all represent the typical degree: actually, there is *no typical node hence no typical degree*. Most often, the empirical degree distribution is consistent with the simplest generic analytical form, namely a power-law[16] dependence $P(k) \sim k^{-\gamma}$ with $\gamma > 1$. The power-law decay of $P(k)$ indicates a non-negligible probability to observe highly connected nodes, and also a higher probability to observe weakly connected nodes ($k=1$ or $k=2$) than in a random graph with $\langle k \rangle \gg 1$.

When links are oriented, two different distributions $P^{\text{in}}(k)$ (probability that k links enter a given node) and $P^{\text{out}}(k)$ (probability that k links come out a given node) are to be distinguished, associated with two exponents γ^{in} and γ^{out}. For instance, such a dual power law is observed in metabolic networks with $\gamma^{\text{in}} = 2.4$ and $\gamma^{\text{out}} = 2.0$ [22]. It is also the case of the Internet and the World Wide Web, with here also different exponents γ^{in} and γ^{out}, and of several biological networks [13,30] In other situations, for instance gene networks (where the nodes are the

[16] A value $\gamma > 2$ is required to have a finite average degree; smaller values of γ can be considered, including values $\gamma < 1$ (a priori precluded to get a normalized $P(k)$) provided the distribution $P(k)$ is truncated at large k values.

genes and a connection (i, j) reflects that the protein coded by gene i regulates the gene j), data are better fitted with an hybrid behavior: a power-law out-degree distribution P^{out} (a protein coded by a given gene can regulate an arbitrary large number of other genes) and an exponentially decaying in-degree distribution P^{in} having a finite characteristic scale imposed by physical reasons,[17] for instance steric hindrance preventing an arbitrarily large number of proteins to be involved in the neighborhood of a given gene promoter [18]. Nevertheless, experimental data are too scarce to faithfully determine the actual nature of the distribution. In most cases, several models can be supported by the same experimental data.

It is indeed to be underlined that it is a highly delicate matter to assess with some certainty the power-law dependence of an experimental degree distribution (see Figure 2). First, it is recommended to estimate the *cumulative degree distribution function* $F(k) = \sum_{k'<k} P(k')$ rather than the distribution $P(k)$ itself: it yields smoother plots (statistical under- and over-representation that might arise in the sample due to finite-size fluctuations average out) and avoids the arbitrary windowing choice required in drawing the histogram when degree values have to be boxed to get a sufficient statistics. Given the empirical degrees $(k_i)_{i=1,\ldots,N}$, one estimates

$$1 - \widehat{F}_N(k) = \frac{1}{N} \text{Card}\{i, k_i \geqslant k\} \quad (\widehat{F}_N \approx F \text{ when } N \to \infty). \tag{11}$$

In case when $P(k) \sim k^{-\gamma}$ at large k, then $\widehat{F}_N(k)$ behaves as $k^{1-\gamma}$. But even with this procedure, data over several decades are required, that correspond here to degrees ranging between 1 and 1,000 at least. In many experimental situations, such degree values are simply not encountered, and the support of the empirical distribution is far narrower. The limitation in estimating the degree distribution might come from the experimental setting but also (as it is frequently the case in biology) from the system itself. In this case, the very issue of the distribution nature (Poisson law vs power law) is somehow ill-posed and meaningless: the distribution reduces to the recorded data and can often be fit by either of the two functional forms; additional arguments should be invoked to justify the modeling choice. To summarize, a power-law shape appears at best as the simplest fit consistent with the data, being more accurate and more faithful than an exponential fit. It is in fact the *simplest generic dependence when there is no characteristic scale*.

Moreover, observing a power-law degree distribution, even with accuracy and certainty, does not represent in itself an explanation of its origin (see for instance the recent essay by Fox Keller, providing an historical perspective and a critical

[17] Let us assume that all factors (transcription factors, co-activators,...) involved around the promoter of a given gene i in order to regulate the expression of this gene have roughly the same concentrations, say c. Since each of these factors corresponds to a connection (ending in i) in the gene network, the probability $P(k)$ for the gene i to have an in-degree equal to k scales as $P(k) \sim c^k$. This scaling estimate supports an exponential nature for the in-degree distribution, with a characteristic degree $k_c = 1/\log(1/c)$. This formula might be read in the reverse way, giving the typical concentration c required for a gene network with typical in-degree k_c to operate properly.

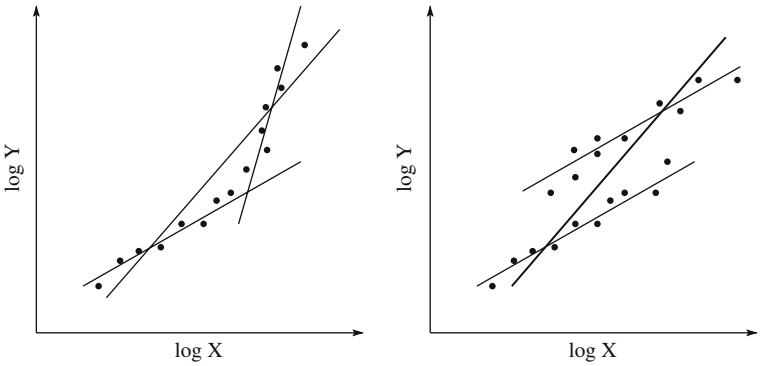

Figure 2. Misleading estimate of the exponent α of a power law $Y \sim X^{\alpha}$ (slope of the bold curve) from experimental data (the points in this sketch) in case of an undetected crossover (*left*) or in case of two sub-populations verifying the same scaling law with different prefactors (*right*).

assessment of this highly fashionable notion of scale-free network [14]). The only unquestionable conclusion that can be drawn is the high heterogeneity among the nodes, reflecting in the absence of characteristic degree and long tails in the degree distribution, all the more dispersed than γ is small. A power-law degree distribution might accompany various architectures, either clustered or not, either modular or not, either hierarchical or not, for instance: further specification of the statistical ensemble might be necessary. Only a modeling approach, based on additional knowledge of the system, can suggest the relevant ensemble and provide a real understanding of the network properties and mechanisms at work in the observed behaviors.

The implicit reference in the name 'scale-free networks' to *scale invariance* as encountered in the context of critical phenomena and fractal structures (both notions involving power-law relationships) should be considered with caution: if not totally misleading, it needs at least to be argued and established beyond a simple metaphor. The scale-invariance of complex networks, if stated without further specification, is indeed an ill-defined feature. More precisely, the degree distribution is trivially scale invariant, but the scale invariance of the network configuration (connection pattern) is another feature, far less obvious: it has to be defined as a statistical property associated with some scale transformation. On the experimental side, quantities to evidence and if possible measure network self-similarity have to be introduced and algorithms to compute them from data have to be designed. Actually, it remains an open field of investigations to describe the behavior of complex networks upon changing the observation scale, to determine what is the relevant concept(s) of self-similarity for such networks and to devise the appropriate coarse-graining procedures allowing to evidence and exploit their self-similarity, if any [26].

The differences observed between scale-free networks and random graphs suggest that real complex networks do not emerge by chance[18] but according to some organizing principles and submitted to various selection mechanisms [9].What realistic mechanism might account of the absence of characteristic degree and observed shape for $P(k)$? Many! Numerous models, more or less plausible, simple and explanatory flourished [32]. The simplest one is based on a *joint principle of growth and preferential attachment*:[19] the new links are added towards nodes already highly connected: 'rich get richer'; this design principle produces a power-law degree distribution with $\gamma = 3$ [4].

Totally different explanations might also be mentioned, all leading to a power-law degree distribution but various architectures, in particular as regards the clustering and spatial distribution of links. We have investigated for instance the situation where each node i connects at random with all nodes of a neighborhood of a size n_i depending on i, with a probability p for each link to actually be present, as in the Erdös-Renyi model. Its out-degree is thus a random variable with mean $\langle k_i^{\text{out}} \rangle = n_i p$. Model consistency requires to orient the links and to distinguish links arriving at a node and links coming out. Out-degree distribution reflects the size distribution $(n_i)_i$, and the issue is now to explain the features of this size distribution, e.g. the absence of characteristic size, and their origin (the question is simply shifted). A robust feature associated with this design, that can be confronted to data, concerns the in-degree distribution: it is always a Poisson law (with parameter the average degree[20] $\langle k^{\text{in}} \rangle = \langle k^{\text{out}} \rangle = \langle k \rangle$), whatever the size distribution $(n_i)_i$ is [25]. Numerous other mechanisms can be invoked to account for the power-law decay of complex network degree distribution. This is besides a general fact when investigating scaling laws: they are delicate to assess experimentally (see Figure 2) and still more delicate to interpret, since numerous alternative explanations can account for the appearance of a power law (see for instance [24] and [39]).

4.3. DIAMETER, CLUSTERING COEFFICIENT AND DEGREE DISTRIBUTION

At this stage, we have come across the network degree distribution. Two other important features are its *diameter*, defined as the average distance between two nodes (i.e., the average over all pairs of nodes of the number of steps of the minimal path connecting them) and the *clustering coefficient*, defined on Figure 3. It is meaningful to compare the four most current network models as regards these three features: regular grids, Erdös-Renyi random graphs, small-worlds obtained

[18] Chance being the common word for 'at random with statistical independent and equi-probable events'.

[19] A quite similar derivation of power-law distributions—termed skewed distributions at that time—has been proposed by Simon [38] in the stream of its investigations about the general organization of complex systems and their hierarchical structure; but it was not directly applied to networks hence remained ignored for long in this context.

[20] It is indeed to note that the average in-degree and the average out-degree always coincide, whatever the degree distributions are, since $\sum_{i,j} A_{ij} = \sum_j k_j^{\text{in}} = N \langle k^{\text{in}} \rangle = \sum_i k_i^{\text{out}} = N \langle k^{\text{out}} \rangle$.

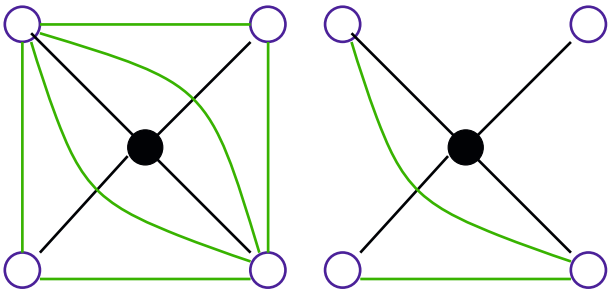

Figure 3. Clustering coefficient, defined as the fraction of edges actually present compared to the maximal number of edges bewteen neighbors (○) of a given node (●).

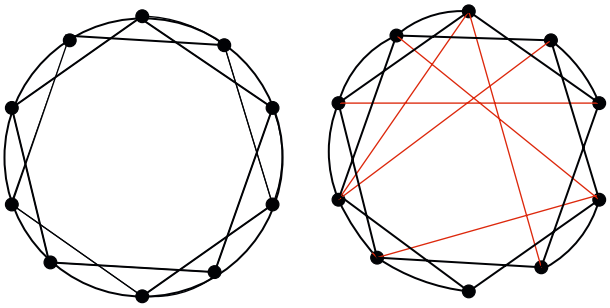

Figure 4. Watts and Strogatz model. Edges of a regular network are rewired at random (independently) with a given probability p, yielding a 'small world' of diameter increasing logarithmically instead of linearly with the number of nodes.

after a re-wiring (presented on Figures 4, 5 [42]) and scale-free networks obtained by a growth according the preferential attachment principle [4]. Table I summarizes their (in)ability to account for the real complex network features.

4.4. COMMUNITY STRUCTURE

An important issue, for instance in the context of social networks (where besides it has been introduced) is the detection of communities, i.e. more densely connected regions.

A first approach is inspired from classification methods developed for general sets of points and based on a quantitative estimate of their similarity, currently by means of a distance. Let us quote for instance the *K-means method* and the *hierarchical clustering*. In the first method, the number K of classes is chosen a priori (this method is thus termed a 'supervised method'); the procedure is then recursive, starting from K points $x_1^{(0)}, \ldots, x_K^{(0)}$ chosen as centers: each point i is associated to the closest center $x_{\alpha(i)}^{(0)}$. Once this sorting has been done for all points, the center of gravity of each sub-group is computed, and the resulting points are then

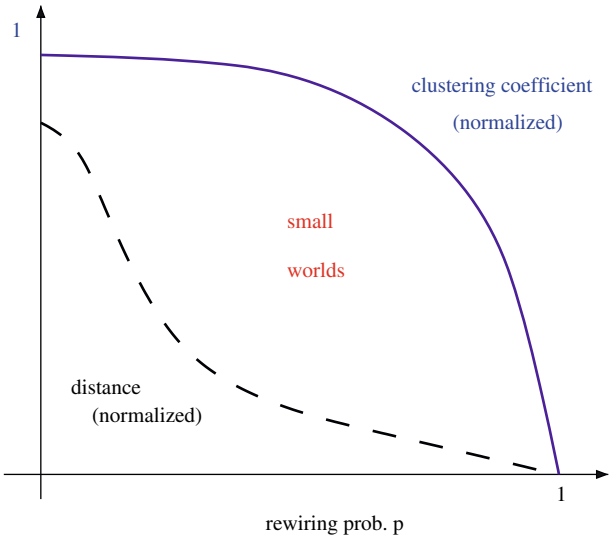

Figure 5. Diameter and clustering coefficient of the Watts and Strogatz model, as a function of the rewiring probability. For intermediary values of p, the network exhibits a small diameter and at the same time a high clustering coefficient, namely the signature of a '*small world*'.

Table I. Comparison of the main network features observed in the most currently used models, compared to real complex networks exhibiting a diameter increasing slower than a power of the number N of nodes, a finite clustering coefficient (i.e. non-vanishing in the limit as $N \to \infty$) and a power-law degree distribution

	Regular lattices	Random graph (Erdös–Renyi)	Small world (Watts–Strogatz)	Scale–free (Barabasi–Albert)	Real complex networks
small diameter	NO	×	×	×	×
clustering	×	NO	×	NO	×
$P(k) \sim k^{-\gamma}$	NO	NO	NO	×	×

used to update the position of the centers into $x_1^{(1)}, \ldots, x_K^{(1)}$. The procedure is iterated using as reference points these new centers, and so forth till the partition does not evolve anymore. The second method consists in determining a tree by grouping the closest points together, in successive steps. Site closeness is quantified by means of a similarity distance, to be defined in each specific instance. This hierarchical method yields a family of partitions more or less grained according to the considered level in the tree, ranging from the elementary level where each group reduces to a single point up to the tree root where all points form a single class.

The second approach was already presented in Section 2.2. It also provides a hierarchical picture, but according a far different procedure, based on the *graph exploration by a random walk with transition matrix* $R_{i\to j} = A_{i\to j}/k_i$; communities are then identified with quasi-invariant regions in which the walk remains trapped for long. This method recovers a hierarchical clustering provided one introduces a dynamic distance between the sites, reflecting the time required for the walk to connect them [17].

4.5. BOOLEAN NETWORKS

It would be of high value for the analysis and modeling of real complex networks to have the theoretical landmarks and guidelines offered by reference models with well-known and controlled behavior. Such a picture is available for instance for dynamical systems (ordinary differential equations of low dimension) following the determination of all the possible asymptotic behaviors, the description of their bifurcations, and the evidence of 'normal forms', that is, simple universal evolution laws to which smooth and generic dynamics reduce by conjugation. This is the spirit in which design and investigations of *random Boolean networks* have been conducted. These models involve networks (V, E) where the state of each site $i \in V$ is described by a Boolean variable x_i (i.e. taking only the values 0 and 1). The evolution, in discrete time, is prescribed by defining for each site i a random map associating to each of the 2^{k_i} possible configurations the k_i neighbors of i at time t a value 0 or 1 for x_i at time $t+1$. These 2^{k_i} choices, defining a 'library' at site i, are in general performed at random and independently each from the other.

The interest of such a random model, that is not based on an explicit mechanism of interaction between a site and its neighbors, is to *account in an effective way for the presence of both inhibitory and excitatory connections, and non-linear interactions*, that would be difficult if not impossible to describe in detail. The fit between the model and the properties of the actual dynamics is performed *directly at the level of statistical features* of the interactions, for instance a bias towards activation if more configurations yields $x_i = 1$ than $x_i = 0$. It is then possible to achieve a systematic investigation of the attractors, their basin of attraction and transients duration, and to determine the dependence of these features with respect to the network size. Several asymptotic dynamic regimes, with different levels of complexity, can be observed [3]. This kind of models have been introduced in the context of gene networks [23] where they are still developed.[21] They prove to be fruitful in numerous other contexts, and provide among the best charted dynamic network models.

[21] See e.g. [36], where the classification provided by the generic dynamic behaviors of random Boolean networks is used as a reference point to analyze the dynamic nature of the transcriptional regulation in eukaryotic cells.

5. Local Features Versus Emerging Properties

I will here give a short overview of the kind of issues that a statistical modeling of networks allows to undertake, with a focus on the articulation between their local and global features.

5.1. NETWORK ROBUSTNESS

A remarkable feature of scale-free networks is the presence of 'hubs', namely highly connected nodes. As one of their effects is to reduce the network diameter, they notably improve the efficiency of communications and network properties will be much affected by the removal of such nodes. Scale-free networks (and in particular real networks exhibiting a long-tail degree distribution) are thus highly sensitive to attacks, namely a targeted destruction of the nodes following decreasing degree. In contrast, these networks are robust with respect to the random removal of nodes: a large fraction of nodes can be removed without notably hampering communications within the network,[22] differing thereby markedly with random graphs. One here recovers the observed robustness of real complex networks. This discussion gives another illustration of the expected differences between, on the one hand real complex networks, on the other hand percolation networks and random graphs used for long to model them, motivating to devise alternative models.

5.2. PROPAGATION PHENOMENA

Static statistical modeling of networks, of interest for its own sake to sort real networks into universality classes and to evidence relevant observables and essential parameters, is also a preliminary step towards dynamic studies. This perspective has been developed extensively for percolation networks. The percolation model design is based on static (topological and statistical) properties of the real network; it then allows a *quantitative* study of dynamical phenomena involving the network. Numerous dynamic studies, intending to predict, control or optimize the dynamic behavior of the system, have been conducted in this spirit, e.g. for forest fire propagation or liquid flowing inside a porous medium [24]. It is in fact in observing propagation phenomena that the existence of a percolation threshold appears the most obviously, as a threshold separating situations where propagation stops from those where it develops and crosses the whole network: the fire either stops or reaches the other side of the forest, the coffee seeps into the cup

[22] It can be shown that in a network where the nodes have a minimal degree m and a power-law degree distribution with exponent $\gamma > 3$, a fraction $g \leqslant g_c = 1 - [m(\gamma-2)/(\gamma-3) - 1]^{-1}$ of nodes can be removed without breaking off the network connectedness. If $\gamma \leqslant 3$, the formula for the threshold $g_c(N)$ is a bit more complicated and still depends on the network size N; it gives $g_c = 0.9$ for realistic values of the networks parameters, for instance those derived from experimental measures in the Internet [32].

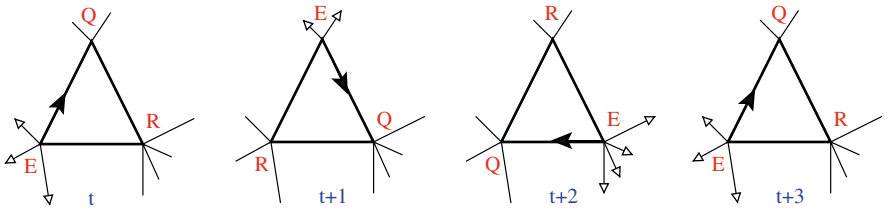

Figure 6. Dynamical motif in a coupled excitable dynamics. Each node can be in either of the three states E (excitable), Q (quiescent) or R (refractory). Its evolution, in discrete time, follows $E \to R \to Q$ and $Q \to E$ if a neighbor is excited, else the node remains quiescent. This *triangle motif* has the remarkable property to be insensitive to the state and dynamics of the surrounding network, keeping its periodic dynamics. Such a motif acts as a 'pace-maker', generically imposing its period 3 to the whole network on the long term.

or not. The special topology of real complex networks requires to undertake investigations of propagation phenomena on more appropriate network models. Applications are numerous, ranging from the propagation of information or viruses in the Internet, the spreading of a contagious disease in a population structured by social relations or moving along a network of airlines, or the fate of drugs in the metabolic network. The remarkable result is the *absence of percolation threshold in scale-free networks*. A quite foreseeable conclusion is that the knowledge of the network topology is essential to predict and describe quantitatively (and already qualitatively) propagation phenomena occurring in the network [32].

5.3. INFLUENCE OF THE NETWORK TOPOLOGY ONTO ITS DYNAMICS

In the previous subsection Section 5.2, we have mentioned some results about the dynamics *on* the network; in Section 5.1, the issue concerned some properties *of* the network itself. We here turn to a third class of questions, concerning *the dynamics of the network nodes*, i.e. the evolution of their individual states $s_i(t)$ when the edges mediate interactions between the nodes and their states. Given a set of coupled elements, the influence of the interaction network topology on their dynamics can be investigated within the frame provided by cellular automata [11]. Like in Boolean networks, state variables are discrete, but taking possibly more than two different values. The main difference lies in the dynamic rules: here the discrete-time updating rules involve the state of neighboring states (in general their expression, either deterministic or stochastic, is spatially homogeneous). The evolution of the state $x_i(t)$ of site i is thus determined by the states $x_j(t)$ of all neighbors j of i, for instance through a threshold condition $\sum_j x_j(t) \geqslant \theta$ where θ is fixed.

In a recent work [7], we studied the influence of network topology, more specifically its degree distribution (either a Poisson law or a power law) on an excitable dynamics, intending to model in a minimal though realistic way a neural

network:[23] each site can be in either of three states E (excitable), Q (quiescent) or R (refractory). Its discrete time evolution writes $E \to R \to Q$ and $Q \to E$ if one of its neighbors is excited, otherwise it remains quiescent. Quite surprisingly compared to the results stated in Sections 5.1 and 5.2, the influence of topology happens to be rather weak and in any case only quantitative: the generic behavior is only more robust and more coherent in a power-law network. Explanation of this low sensitivity, detailed on Figure 6, lies in the presence of dynamic motifs imposing their intrinsic dynamics to the whole network, whatever its topology is. This result leads us to discuss in more detail the role of local properties and motifs in the overall network behavior.

As previously underlined in Section 4.2 in a general context, reducing the topology to the degree distribution might well be an over-simplification, and other features could be more discriminating for the resulting dynamic behavior, for instance the clustering coefficient or the distribution of cycle size (length of the minimal nontrivial path connecting a node to itself, if any). Preliminary studies have thus to be conducted to determine what are the relevant statistical ensembles. For instance, in the example presented above [7], one of the results of our study is precisely to suggest what cannot be guessed on intuitive grounds, namely that the statistics of triangles is likely to be a key ingredient and should be explicitly taken into account when investigating an excitable dynamics on a network.

5.4. MOTIFS AND BIOLOGICAL REGULATORY NETWORKS

A *motif* is any over-represented or under-represented ('anti-motif') subgraph in the considered network, compared to the predictions of a given statistical model. This notion rises a huge interest in the context of biological regulatory networks, for instance metabolic networks and gene networks. Looking for motifs is today considered to be a privileged approach, operational and straightforwardly meaningful, to unravel biological functions and their actors (proteins or enzymatic reactions, for instance) [35]. Comparison of motifs statistics in regulatory networks of different species [5] has been proposed as a pathway towards the evolutionary history of these networks. Other remarkable local structures can be investigated, for instance communities (previously discussed in Sections 2.2 and 4.4), corresponding to densely connected regions and interpreted as functional modules. But these approaches should be confronted to fundamental issues regarding the behavior of a network, the relevant notion of causality and emergent properties, observed at the level of the whole network but impossible to foresee by simply looking at the detailed properties of the elementary ingredients (nodes, edges, local structures). Some caveats ensue, and some caution is required in the analysis to avoid an

[23] Other exemplary studies, similarly focusing on the dynamic consequences of the absence of characteristic degree, are [2] (for Boolean networks) [28] (for two-state cellular automata) and [27] (for population evolutionary dynamics).

over-interpretation of the motifs and their role in the overall behavior of the network and its regulation.

A motif is a local property, and it is thus necessary to consider its embedding in the whole network and to investigate the emergent consequences of this local statistical feature [31]. The importance of a motif should be thus evaluated in terms of its influence on some global property of the network, for instance propagation or dynamics properties. Otherwise, the presence of motifs can simply result, as an auxiliary and non-functional fact in the present network, of the growth process or, in a biological context, to reflect some remnant of its evolutionary history, in other words a mere statistical bias reflecting some specificity of the network genesis [29]. In a similar way, the community structure is not in itself a significant network property unless links have an intrinsic meaning, like connections in a social network. In the general case where connections only mediate interactions or indicate the possibility of some pairing, communities do not necessarily reflect some essential organizational or functional property. We thus underline the importance to put the motif interpretation back in a definite context: whether it is mere topological property, a property following from flux analysis, or a dynamic feature as in Figure 6; the presence of over-represented motifs will not have the same origin and the same meaning in each case.

An unquestionable interest for motif analysis and community detection in a real network is to confront these statistical properties with those of possible models, especially when several concurrent ones are available. This viewpoint has been largely developed in the context of percolation networks, where the cluster statistics provides a quantitative criterion to appreciate the adequacy of a model to the real system, before exploiting this model, e.g. to predict dynamic behavior of the system. Another interest for motif determination is to put constraints on analytical studies, for instance the requirement to take into account local correlations (reflecting in the existence of under- or over-represented motifs). The notion of motif is thus relevant when it is included among the other local statistical features (degree distribution, clustering coefficient, for instance) taken into account in the network analysis, with no disguised hope to avoid this necessary study.

To summarize, detecting a motif is not a directly meaningful result. In a similar spirit, it is improper to identify more densely connected regions with modules, in the sense of functionally independent blocks. This kind of interpretation rather reflect the wish to escape the network specificity as regards causal relationships and our lack of tools to grasp them as a whole: it would be simpler to investigate networks if they could be reduced to an array of independent building blocks with well-defined and context-free properties (e.g. invariant response functions). It is rarely the case, and network complexity precisely originates in the fact that the properties and even the potentialities of the elements are strongly affected by the very phenomena occurring in it. For instance, the outcome of the competition between several possible pathways is highly dependent of the surrounding state and evolution. Even local manifestations, i.e. events observed locally (excitation of

a given node, for instance) follow from the collective dynamics of the network. Networks thus rise a huge challenge to the theoreticians, demanding to develop systematic methods to describe and predict the consequences of numerous concerted or competitive interactions and feedbacks between local and global network properties.

6. Conclusions and Perspectives

I have presented here only the simplest situation of networks with symmetric and unweighted links. More realistic models should further account for the oriented, valued, or even signed character of the connections (either excitatory or inhibitory), the possible hybrid nature of the nodes (hence of the connections), their variation in time (non-stationarity) and in space (inhomogeneity). The first step is then to discriminate the inessential ingredients, leading only to a quantitative change in qualitatively similar observed behaviors, from the essential ones, determining different behaviors and requiring to elaborate novel minimal models to investigate their generic properties. Other perspectives open up, tackling network growth, their appearance in the course of Evolution (for biological, ecological or social networks), their robustness with respect to failures or targeted modifications of their local structure. It is a whole world, with entries in almost all scientific disciplines, that asks for exploration, both on the theoretical and experimental sides.

Acknowledgements

I am grateful to Sylvie Paycha and Alessandra Frabetti for inviting me to contribute to this special issue. I warmly acknowledge A. Benecke, A. Gerbaud, M. Latapy, X. Manière, L.S. Schulman and M. Vergassola for fruitful discussions.

References

1. Albert, R., Barabasi, A.L.: Statistical mechanics of complex networks. Rev. Mod. Phys. **74**, 47–97 (2001)
2. Aldana, M.: Dynamics of Boolean networks with scale free topology. Physica D **185**, 45–66 (2003)
3. Aldana, M., Coppersmith, S., Kadanoff, L.P.: Boolean dynamics with random couplings. In: Kaplan, E., Marsden, J.E., Sreenivasan, K.R. (eds.) Perspectives and Problems in Nonlinear Science. A celebratory volume in honor of Lawrence Sirovich, Springer, Berlin Heidelberg New york (2003)
4. Barabasi, A.L., Albert, R.: Emergence of scaling in random networks. Science **286**, 509–512 (1999)
5. Berg, J., Lassig, M.: Local graph alignment and motif search in biological networks. Proc. Natl. Acad. Sci. USA **101**, 14689–14694 (2004).
6. Bollobas B.: Random Graphs. Cambridge University Press, Cambridge (2001)
7. Carvunis, A.R., Latapy, M., Lesne, A., Magnien, C., Pezard, L.: Dynamics of three-state excitable units on random vs power-law networks. Physica A (2005 in press). Available at http://www.liafa.jussieu.fr/~latapy/index.php?item=publis&lang=fr

8. Ciliberti, S., Caldarelli, G., De Los Rios, P., Pietronero, L., Zhang, Y.C.: Discretized diffusion processes. Phys. Rev. Lett. **85**, 4848–4851 (2000)
9. Dorogovtsev, S.N., Mendes J.F.F.: Evolution of Networks: from Biological Nets to the Internet and WWW.Oxford University Press, Oxford (2003)
10. Erdös, P., Rényi, A.: On the evolution of random graphs. Publ. Math. Inst. Hung. Acad. Sci. **5**, 17–61 (1960)
11. Ermentrout, G.B., Edelstein-Keshet, L.: Cellular automata approaches to biological modelling. J. Theor. Biol. **160**, 97–133 (1993)
12. Faloutsos, M., Faloutsos, P., Faloutsos, C.: On power-law relationships of the Internet topology. Comput. Commun. Rev. **29**, 251–262 (1999)
13. Farkas, I., Derényi, I., Jeong, H., Neda, Z., Oltvai, Z.N., Ravasz, E., Schrubert, A., Barabasi, A.L.: Networks in life: scaling properties and eigenvalues spectra, Physica A **314** 25–34 (2002)
14. Fox Keller, E.: Revisiting "scale-free" networks. BioEssays **27**, 1060–1068 (2005)
15. García-Pelayo, R., Stadler, P.F.: Correlation length, isotropy and meta-stable states. Physica D **107**, 240–254 (1997)
16. Gaveau, B., Lesne, A., Schulman, L.S.: Spectral signatures of hierarchical relaxation. Phys Lett A **258** 222–228 (1999)
17. Gaveau, B., Schulman, L.S.: Dynamical distance: coarse grains, pattern recognition and network analysis. Bull. Sci. Math. **129**, 631–642 (2005)
18. Guelzim, N., Bottani, S., Bourgine, P., Képès, F.: Topological and causal structure of the yeast transcriptional regulatory network. Nat. Genet. **31**, 60–63 (2002)
19. Guhr, T., Müller-Groeling, A., Weidenmüller, H.A.: Random-matrix theories in quantum physics: common concepts. Phys. Rep. **299**, 189–425 (1998)
20. Hopfield, J.J.: Neural networks and physical systems with emergent collective computational abilities. Proc. Natl. Acad. Sci. USA **79**, 2554–2558 (1982)
21. Imbert, J.B., Lesne, A.: Improved ergodicity in constrained Monte Carlo sampling (2006, in preparation)
22. Jeong, H., Tombor, B., Albert, R., Oltvai, Z.N., Barabasi, A.L.: The large-scale organization of metabolic networks. Nature **407**, 651–654 (2000)
23. Kauffman, S.A.: The Origins of Order: Self-organization and Selection in Evolution. Oxford University Press, Oxford (1993)
24. Lesne, A.: Renormalization Methods. Wiley, New York (1998)
25. Lesne, A.: Disordered random graph models: scale-free out with a typical in-degree (2006, submitted)
26. Lesne, A.: Scale-invariant behavior of scale-free graphs (2006, submitted)
27. Lieberman, E., Hauert, C., Nowak, M.A.: Evolutionary dynamics on graphs. Nature **433**, 312–316 (2005)
28. Marr, C., Hütt, M.T.: Topology regulates pattern formation capacity of binary cellular automata on graphs. Physica A **354**, 641–662 (2005)
29. Mazurie, A., Bottani, S., Vergassola, M.: An evolutionary and functional assessment of regulatory network motifs. Genome Biol. **6**, R35 (2005)
30. Newman, M.E.J.: The structure and function of complex networks. SIAM Rev. **45**, 167–256 (2003)
31. Papin, J.A., Reed, J.L., Palsson, B.: Hierarchical thinking in network biology: the unbiased modularization of biochemical networks. Trends Biochem Sci. **29**, 641–647 (2004)
32. Pastor-Satorras, R., Vespignani, A.: Evolution and Structure of the Internet. Cambridge University Press, Cambridge (2004)
33. Pimm, S.L.: Food Webs. Chapman & Hall, London (1982)

34. Schnakenberg, J.: Network theory of microscopic and macroscopic behavior of solutions of master equations. Rev. Mod. Phys. **48**, 571–585 (1976)
35. Shen-Orr, S., Milo, R., Mangan, S., Alon, U.: Network motifs in the transcriptional regulation network of Escherichia coli. Nat. Genet. **31**, 64–68 (2002)
36. Shmulevich, I., Kauffman, S.A., Aldana, M.: Eukaryotic cells are dynamically ordered or critical but not chaotic. Proc. Natl. Acad. Sci. USA **102**, 13439–13444 (2005)
37. Simon, H.A.: On a class of skew distribution functions. Biometrika **42**, 425–440 (1955)
38. Simonsen, I: Diffusion and networks: a powerful combination! Physica A **357**, 317–330 (2005)
39. Sornette, D.: Critical Phenomena in Natural Sciences. Springer, Berlin heidelberg Newyork New York, (2000)
40. Stadler, P.F.: Landscapes and their correlation functions. J. Math. Chem. **20**, 1–45 (1996)
41. Strogatz, S.H.: Exploring complex networks. Nature **410**, 268–276 (2001)
42. Watts, D.J., Strogatz, S.H.: Collective dynamics of "small-world" networks.[Nature **393**, 440–442 (1998)
43. Weinberger, E.D.: Correlated and uncorrelated fitness landscapes and how to tell the difference. Biol. Cybern. **63**, 325–336 (1990)