Bose-Einstein Condensation in Complex Networks

Ginestra Bianconi[1] and Albert-László Barabási[1,2]

[1]*Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556*

[2]*Institute for Advanced Studies, Collegium Budapest, Szentháromság utca 2, H-1014 Budapest, Hungary*

(Received 16 October 2000; revised manuscript received 26 January 2001)

The evolution of many complex systems, including the World Wide Web, business, and citation networks, is encoded in the dynamic web describing the interactions between the system's constituents. Despite their irreversible and nonequilibrium nature these networks follow Bose statistics and can undergo Bose-Einstein condensation. Addressing the dynamical properties of these nonequilibrium systems within the framework of equilibrium quantum gases predicts that the "first-mover-advantage," "fit-get-rich," and "winner-takes-all" phenomena observed in competitive systems are thermodynamically distinct phases of the underlying evolving networks.

DOI: 10.1103/PhysRevLett.86.5632 PACS numbers: 89.75.Hc, 03.75.Fi, 05.65.+b, 87.23.Ge

Competition for links is a common feature of complex systems: on the World Wide Web (www) the sites compete for URLs to enhance their visibility [1], in the business world companies compete for links to consumers [2], and in the scientific community scientists and publications compete for citations, a measure of their impact on the field [3]. A common feature of these systems is that the nodes self-organize into a complex network, whose topology and evolution closely reflect the dynamics and outcome of this competition [1,3–6]. Here we show that, despite their nonequilibrium and irreversible nature, evolving networks can be mapped into an equilibrium Bose gas [7], nodes corresponding to energy levels, and links representing particles. This mapping predicts that the common epithets used to characterize competitive systems, such as "winner takes all," "fit get rich" (FGR), or "first mover advantage," emerge naturally as thermodynamically and topologically distinct phases of the underlying complex evolving network. In particular, we predict that such networks can undergo Bose-Einstein (BE) condensation, in which a single node captures a macroscopic fraction of links.

Fitness model [8].—Consider a network that grows through the addition of new nodes such as the creation of new webpages, the emergence of new companies, or the publication of new papers. At each time step we add a new node, connecting it with m links to the nodes already present in the system. The rate at which nodes acquire links can vary widely as supported by measurements on the www [4], and by empirical evidence in citation [3] and economic networks. To incorporate the different ability of the nodes to compete for links we assign a fitness parameter to each node η, chosen from a distribution $\rho(\eta)$, accounting for the differences in the content of webpages, the quality of products and marketing of companies, or the importance of the findings reported in a publication. The probability Π_i that a new node connects one of its m links to a node i already present in the network depends on the number of links k_i and on the fitness η_i of node i, such that

$$\Pi_i = \frac{\eta_i k_i}{\sum_\ell \eta_\ell k_\ell}. \tag{1}$$

Equation (1) incorporates in the simplest possible way the fact that new nodes link preferentially to nodes with higher k [9] (i.e., connecting to more visible websites, favoring more established companies, or citing more cited papers) and with larger fitness (i.e., websites with better content, companies with better products and sales practice, or papers with novel results). Thus fitness (η_i) and the number of links (k_i) jointly determine the attractiveness and evolution of a node.

Mapping to a Bose gas.—We assign an *energy* ϵ_i to each node, determined by its fitness η_i through the relation

$$\epsilon_i = -\frac{1}{\beta}\log\eta_i, \tag{2}$$

where β is a parameter playing the role of inverse temperature, $\beta = 1/T$, whose relevance to real networks will be discussed later. A link between two nodes i and j with energies ϵ_i and ϵ_j (e.g., fitnesses η_i and η_j) corresponds to two noninteracting *particles* on the energy levels ϵ_i and ϵ_j (Fig. 1). Adding a new node to the network corresponds to adding a new energy level ϵ_i and $2m$ particles to the system. Of these $2m$ particles, some m are deposited on the level ϵ_i (corresponding to the m outgoing link that node i processes), while the other m particles are distributed between the other energy levels (representing the links pointing to m nodes present in the system), the probability that a particle lands on level i being given by (1). Deposited particles are inert, i.e., they are not allowed to jump to other energy levels.

Each node (energy level) added to the system at time t_i with energy ϵ_i is characterized by the occupation number $k_i(\epsilon_i, t, t_i)$, denoting the number of links (particles) that the node (energy level) has at time t. The rate at which level ϵ_i acquires new particles is

 0031-9007/01/86(24)/5632(4)$15.00

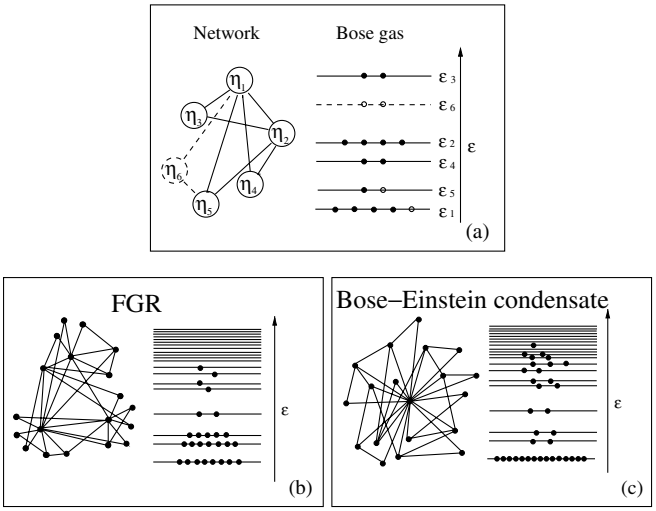

FIG. 1. Schematic illustration of the mapping between the network model and the Bose gas. (a) In the network of five nodes (continuous circles and lines) each node is characterized by a fitness η_i chosen from a distribution $\rho(\eta)$. Equation (2) assigns an energy ϵ_i to each η_i, generating a system of random energy levels (right). A link from node i to node j corresponds to a particle at level ϵ_i and one at ϵ_j. The network evolves by adding a new node (dashed circle, η_6) at each time step which connects to other $m = 2$ nodes (dashed lines), chosen randomly following (1). In the gas this results in the addition of a new energy level (ϵ_6, dashed line), populated by $m = 2$ particles, and the deposition of $m = 2$ other particles at energy levels to which η_6 is connected (ϵ_2 and ϵ_5). The number of energy levels and particles increases linearly with time, as t and $2mt$, respectively. (b) In the FGR phase we have a continuous connectivity distribution, the several highly connected nodes linking the numerous small nodes together. In the energy diagram this corresponds to a decreasing occupation level with increasing energy. (c) In the Bose-Einstein condensate the fittest node attracts a finite fraction of all links, corresponding to a highly populated energy level and only sparsely populated higher energies. In (b) and (c) the diagram shows only incoming links, ignoring the default $m = 2$ particles on each energy level corresponding to the outgoing links.

$$\frac{\partial k_i(\epsilon_i, t, t_i)}{\partial t} = m \frac{e^{-\beta\epsilon_i} k_i(\epsilon_i, t, t_i)}{Z_t}, \tag{3}$$

where Z_t is the partition function, defined as

$$Z_t = \sum_{j=1}^{t} e^{-\beta\epsilon_j} k_j(\epsilon_j, t, t_j). \tag{4}$$

We assume that each node increases its connectivity following a power law

$$k_i(\epsilon_i, t, t_i) = m\left(\frac{t}{t_i}\right)^{f(\epsilon_i)}, \tag{5}$$

where $f(\epsilon)$ is the energy dependent dynamic exponent. Since η is chosen randomly from the distribution $\rho(\eta)$, the energy levels are chosen from the distribution $g(\epsilon) = \beta\rho(e^{-\beta\epsilon})e^{-\beta\epsilon}$. We can now determine Z_t by averaging over $g(\epsilon)$, i.e.,

$$\langle Z_t \rangle = \int d\epsilon\, g(\epsilon) \int_1^t dt_0\, e^{-\beta\epsilon_i} k(\epsilon, t, t_0) = mz^{-1}t[1 + O(t^{-\alpha})], \tag{6}$$

where

$$\frac{1}{z} = \int d\epsilon\, g(\epsilon) \frac{e^{-\beta\epsilon}}{1 - f(\epsilon)} \tag{7}$$

is the inverse fugacity and $\alpha = \min_\epsilon[1 - f(\epsilon)] > 0$. Since z is positive for any $\beta \neq 0$ we introduce the chemical potential μ as $z = e^{\beta\mu}$, which allows us to write (6) and (7) as

$$e^{-\beta\mu} = \lim_{t\to\infty} \frac{\langle Z_t \rangle}{mt}. \tag{8}$$

By using (8) we can solve the continuum equation (3), finding in a self-consistent way the solutions of form (5), where the dynamic exponent is

$$f(\epsilon) = e^{-\beta(\epsilon-\mu)}. \tag{9}$$

By combining (7) and (9), we find that the chemical potential is the solution of the equation

$$I(\beta, \mu) = \int d\epsilon\, g(\epsilon) \frac{1}{e^{\beta(\epsilon-\mu)} - 1} = 1. \tag{10}$$

The system defined above has a number of properties that make it an unlikely candidate for an equilibrium Bose gas [7]. First, the inertness of the particles is a nonequilibrium feature, in contrast with the ability of particles in a quantum gas to jump between energy levels, leading to a temperature driven equilibration. Second, both the number of eligible energy levels (nodes) and the particles populating them (links) increase linearly in time, in contrast with the fixed system size employed in quantum systems. Despite these apparent differences, Eq. (10) indicates that in the thermodynamic limit ($t \to \infty$) the fitness model maps into a Bose gas. Indeed, since in an ideal gas of volume $v = 1$, we have [7]

$$\int d\epsilon\, g(\epsilon) n(\epsilon) = 1, \tag{11}$$

where $n(\epsilon)$ is the occupation number of a level with energy ϵ. Equation (10) indicates that for the inert gas inspired by the fitness model the occupation number follows the familiar Bose statistics [7]

$$n(\epsilon) = \frac{1}{e^{\beta(\epsilon-\mu)} - 1}, \tag{12}$$

i.e., the evolving network maps into a Bose gas. Thus the irreversibility and the inertness of the network are resolved by the stationarity of the asymptotic distribution, allowing the occupation number to follow Bose statistics in the thermodynamic limit $t \to \infty$.

Bose-Einstein condensation.—The solutions (5), (6), and (9) exist only when there is a μ that satisfies Eq. (10).

However, $I(\beta,\mu)$ defined in (10) takes its maximum at $\mu = 0$, thus when $I(\beta,0) < 1$ for a given β and $g(\epsilon)$, Eq. (10) has no solution. The absence of a solution is a well-known signature of Bose-Einstein condensation [7], indicating that a finite $n_0(\beta)$ fraction of particles condensate on the lowest energy level. Indeed, due to mass conservation at time t, we have t energy levels populated by $2mt$ particles, i.e.,

$$2mt = \sum_{t_0=1}^{t} k(\epsilon_{t_0}, t, t_0) = mt + mtI(\beta,\mu). \quad (13)$$

When $I(\beta,0) < 1$, Eq. (13) has to be replaced with

$$2mt = mt + mtI(\beta,\mu) + n_0(\beta), \quad (14)$$

where $n_0(\beta)$ is given by [7]

$$\frac{n_0(\beta)}{mt} = 1 - I(\beta,0). \quad (15)$$

The occupancy of the lowest energy level corresponds to the number of links the node with the largest fitness has. Thus the emergence of a nonzero $n_0(\beta)$, a signature of Bose-Einstein condensation in quantum gases, represents a "winner-takes-all" phenomenon for networks, the fittest node acquiring a finite fraction of the links, independent of the size of the network.

The mapping to a Bose gas and the possibility of Bose-Einstein condensation in random networks predict the existence of three distinct phases characterizing the dynamical properties of evolving networks: (a) a scale-free phase, (b) a fit-get-rich phase, and (c) a Bose-Einstein condensate. Next we discuss each of these possible phases separately.

(a) *Scale-free phase.*—When all nodes have the same fitness, i.e., $\rho(\eta) = \delta(\eta - 1)$, $[g(\epsilon) = \delta(\epsilon)]$, the model reduces to the scale-free model [9], introduced to account for the power-law connectivity distribution observed in diverse systems, such as the www [5,10], coauthorship networks [11], Internet [12], or citation networks [3]. The model describes a "first-mover-wins" behavior, in which the oldest nodes acquire most links. Indeed, (9) predicts that $f(\epsilon) = 1/2$; i.e., according to (5) all nodes increase their connectivity as $t^{1/2}$, the older nodes with smaller t_i having larger k_i. However, the oldest and "richest" node is not an absolute winner, since its share of links, $k_{\max}(t)/(mt)$, decays to zero as $t^{-1/2}$ in the thermodynamic limit. Thus a continuous hierarchy of large nodes coexists, such that the degree distribution $P(k)$, giving the probability to have a node with k links, follows a power law $P(k) \sim k^{-3}$ [9,13]. Rewiring, aging, and other local processes can modify the scaling exponents or introduce exponential cutoffs in $P(k)$ [10,13–15] while leaving the thermodynamic character of the phase unchanged.

(b) *Fit-get-rich phase.*—This phase emerges in systems for which nodes have different fitnesses and Eq. (10) has a solution [i.e., $I(\beta,\mu) = 1$]. Equation (5) indicates that each node increases its connectivity in time, but the dynamic exponent is larger for nodes with higher fitness [8]. This allows fitter nodes to join the system at a later time and to surpass the less fit but older nodes by acquiring links at higher rates [4]. Consequently, this phase describes a "fit-get-rich" phenomenon, in which, with time, the fitter prevails. But, while there is a clear winner, similar to the scale-free phase the fittest node's share of all links decreases to zero in the thermodynamic limit. Indeed, since $f(\epsilon) < 1$, the relative connectivity of the fittest node decreases as $k(\epsilon_{\min},t)/(mt) \sim t^{f(\epsilon_{\min})-1}$. This competition again leads to the emergence of a hierarchy of a few large "hubs" accompanied by many less connected nodes, $P(k)$ following $P(k) \sim k^{-\gamma}$, where γ can be calculated analytically if $\rho(\eta)$ is known [8].

(c) *Bose-Einstein condensate.*—Bose-Einstein condensation appears when (10) has no solution, at which point (5), (9), and (10) break down. In the competition for links the node with the largest fitness emerges as a clear winner, a finite fraction of particles $[n_0(\beta)]$ landing on this energy level. Thus BE condensation predicts a real winner-takes-all phenomenon, in which the fittest node is not only the largest but, despite the continuous emergence of new nodes that compete for links, it also always acquires a finite fraction of links [Eq. (15)].

To demonstrate the existence of a phase transition from the FGR phase to a BE condensate, we assume that the energy (fitness) distribution follows

$$g(\epsilon) = C\epsilon^{\theta}, \quad (16)$$

where θ is a free parameter and the energies are chosen from $\epsilon \in (0,\epsilon_{\max})$, the normalization giving $C = (\theta + 1)/\epsilon_{\max}^{\theta+1}$. For this class of distributions the condition for a Bose condensation is

$$\frac{\theta + 1}{(\beta\epsilon_{\max})^{\theta+1}} \int_{\beta\epsilon_{\min}(t)}^{\beta\epsilon_{\max}} dx \, \frac{x^{\theta}}{e^{x} - 1} < 1, \quad (17)$$

where $\epsilon_{\min}(t)$ corresponds to the lowest energy (fittest) node present in the system at time t. Extending the limits of integration to zero and infinity, we find the lower bound for the critical temperature $T_{\mathrm{BE}} = 1/\beta_{\mathrm{BE}}$

$$T_{\mathrm{BE}} > \epsilon_{\max}[\zeta(\theta + 1)\Gamma(\theta + 2)]^{-1/(\theta+1)}. \quad (18)$$

We simulated numerically the discrete network model described above, using the energy distribution (16). The chemical potential μ, measured numerically, indicates a sharp transition from a positive to a negative value (Fig. 2a), corresponding to the predicted phase transition between the BE and the FGR phases. The difference between the two phases is seen in the relative occupation number of the most connected node for different temperatures (Fig. 2b). We find that the ratio $k_{\max}(t)/mt$ is independent of time in the BE phase, indicating that the largest node maintains a finite fraction of the total number of links even as the network continues to expand, a signature of BE condensation. In contrast, for $T > T_{\mathrm{BE}}$,

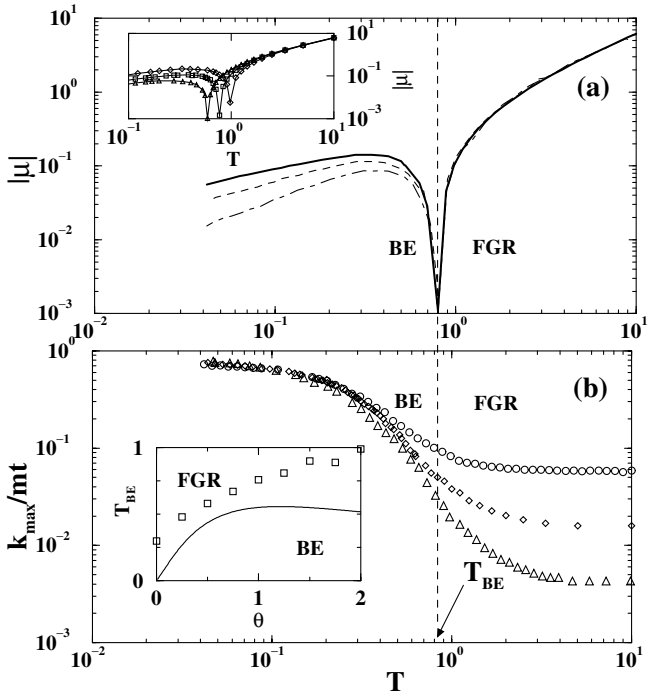

FIG. 2. Numerical evidence for Bose-Einstein condensation in a network model. (a) By choosing the energies from the distribution (16) with $\epsilon_{\max} = 1$, we calculated the chemical potential μ numerically as the network evolved in time, using (8), plotting $|\mu|$ on a logarithmic scale. The temperature at which μ changes sign corresponds to the sharp drop in $|\mu|$, and identifies the critical temperature $T_{\rm BE}$ for Bose-Einstein condensation. The data are shown for $\theta = 1$ and for different times (i.e., system sizes) $t = 10^3$ (continuous line), 10^4 (dashed line), and 10^5 (long-dashed line), averaged over 500, 100, and 30 runs, respectively. The inset shows the chemical potential for different values of the exponent θ in (16), i.e., $\theta = 0.5$, 1.0, 2.0, indicating the θ dependence of $T_{\rm BE}$. (b) Fraction of the total number of links connected to the most connected ("winner") node, $k_{\max}/(mt)$, plotted as a function of T, shown for $m = 2$ and $\theta = 1$. The three curves recorded at $t = 10^3$, 10^4, 10^5 indicate the difference between the two phases: in the BE phase (left) the fittest node maintains a finite fraction of links even as the system expands, while in the FGR phase (right) the fraction of links connected to the most connected node decreases with time. The inset shows the $(\theta, T_{\rm BE})$ phase diagram, the continuous line corresponding to the lower bound predicted by Eq. (18), while the symbols represent the numerically measured $T_{\rm BE}$ as indicated by the position of the peaks in inset (a).

the most connected node gradually loses its share of links, $k_{\max}(t)/mt$ decreasing continuously with time. The numerically determined phase diagram (Fig. 2b) confirms that the analytical prediction (18) offers a lower bound for $T_{\rm BE}$.

Since real networks have a T independent $\rho(\eta)$ fitness distribution, whether they are in the BE or the FGR phase is independent of T. Indeed, for example, choosing $\rho(\eta) = (\lambda + 1)(1 - \eta)^{\lambda}$ the network undergoes a BE condensation for $\lambda > \lambda_{\rm BE} = 1$, and T vanishes from all topologically relevant quantities. Thus T plays the role of a simple control parameter in the model, allowing one to tune it across the transition from the FGR to the BE phase. The presence of $T_{\rm BE}$ in the numerically studied model (Fig. 2) is rooted in our technically simpler choice of defining $g(\epsilon)$ to be independent of T. However, as the inset in Fig. 2b shows, by changing θ the phase transition emerges for fixed T as well, thus T is not necessary for such a transition [16].

[1] S. Lawrence and C. L. Giles, Nature (London) **400**, 107 (1999).

[2] A. Kirman, J. Evol. Econ. **4**, 339 (1997).

[3] S. Redner, Euro. Phys. J. B **4**, 131 (1998).

[4] L. A. Adamic and B. A. Huberman, Science **287**, 2115 (2000).

[5] R. Albert, H. Jeong, and A.-L. Barabási, Nature (London) **401**, 130 (1999).

[6] D. J. Watts and S. H. Strogatz, Nature (London) **393**, 440 (1998).

[7] K. Huang, *Statistical Mechanics* (Wiley, Singapore, 1987).

[8] G. Bianconi and A.-L. Barabási, Europhys. Lett. (to be published).

[9] A.-L. Barabási and R. Albert, Science **286**, 509 (1999); A.-L. Barabási, R. Albert, and H. Jeong, Physica (Amsterdam) **281A**, 69 (2000).

[10] R. Albert and A.-L. Barabási, Phys. Rev. Lett. **85**, 5234 (2000).

[11] M. E. J. Newman, cond-mat/0011144; A.-L. Barabási, H. Jeong, Z. Neda, E. Ravasz, A. Schubert, and T. Vicsek, cond-mat/0104162.

[12] M. Faloutsos, P. Faloutsos, and C. Faloutsos, Comput. Commun. Rev. **29**, 251 (1999); R. Albert, H. Jeong, and A.-L. Barabási, Nature (London) **406**, 378 (2000).

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

[14] L. A. N. Amaral, A. Scala, M. Barthèlèmy, and H. E. Stanley, Proc. Natl. Acad. Sci. U.S.A. **97**, 11 149 (2000).

[15] P. L. Krapivsky, S. Redner, and F. Leyvraz, Phys. Rev. Lett. **85**, 4629 (2000); P. L. Krapivsky and S. Redner, cond-mat/0011094.

[16] Recently Krapivsky *et al.* [15] predicted a gelation phenomenon for nonlinear preferential attachment, $\Pi(k) \sim k^{\nu}$, and $\nu > 1$. While both gelation and the BE condensation predict the capture of the links by a single node, BE condensation can appear only in the presence of fitness.