Damage Spreading and Criticality in Finite Random Dynamical Networks

Thimo Rohlf[1], Natali Gulbahce[2] and Christof Teuscher[3]
[1] *Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, U.S.A*
[2] *Los Alamos National Laboratory, Theoretical Division and Center for Nonlinear Studies, MS B284, Los Alamos, NM 87545, U.S.A.*
[3] *Los Alamos National Laboratory, CCS-3, MS B287, Los Alamos, NM 87545, U.S.A.*

(Dated: August 9, 2007)

We systematically study and compare damage spreading at the sparse percolation (SP) limit for random boolean and threshold networks with perturbations that are independent of the network size N. This limit is relevant to information and damage propagation in many technological and natural networks. Using finite size scaling, we identify a new characteristic connectivity K_s, at which the average number of damaged nodes $\bar{d}$, after a large number of dynamical updates, is independent of N. Based on marginal damage spreading, we determine the critical connectivity $K_c^{sparse}(N)$ for finite N at the SP limit and show that it systematically deviates from K_c, established by the annealed approximation, even for large system sizes. Our findings can potentially explain the results recently obtained for gene regulatory networks and have important implications for the evolution of dynamical networks that solve specific computational or functional tasks.

PACS numbers: 05.45.-a, 05.65.+b, 89.75.-k

Random boolean networks (RBN) were originally introduced as simplified models of gene regulation [1, 2], focusing on a system-wide perspective rather than on the often unknown details of regulatory interactions [3]. In the thermodynamic limit, these disordered dynamical systems exhibit a dynamical order-disorder transition at a sparse critical connectivity K_c [4]; similar observations were made for sparsely connected random threshold (neural) networks (RTN) [5, 6]. For a finite system size N, the dynamics of both systems converge to periodic attractors after a finite number of updates. At K_c, the phase space structure in terms of attractor periods [7], the number of different attractors [8] and the distribution of basins of attraction [9] is complex, showing many properties reminiscent of biological networks [2].

Often, one is interested in the response of dynamical networks to external perturbations; because these signals can disrupt the generic dynamical state (fixed point or periodic attractor) of the network, they are usually referred to as "damage." This type of study has numerous applications, e.g., the spreading of disease through a population [10, 11], the spreading of a computer virus on the internet [12], failure propagation in power grids [13], or the perturbation of gene expression patterns in a cell due to mutations [14]. Mean-field approaches, e.g., the annealed approximation (AA) introduced by Derrida and Pomeau [4], allow for an analytical treatment of damage spreading and exact determination of the critical connectivity K_c under various constraints [15]. It has been shown that local, mean-field-like rewiring rules coupled to order parameters of the dynamics can drive both RBN and RTN to self-organized criticality [16–18].

Mean-field approximations of RBN/RTN dynamics rely on the assumption that $N \to \infty$ and study the rescaled damage $d(t)/N$ (where $d(t)$ is the number of damaged nodes at time t). For an application to real-world problems, these limits are often not very relevant. Going beyond the framework of AA, a number of recent studies focus on the finite-size scaling of (un-)frozen and/or relevant nodes in RBN with respect to N [19, 20]; only few studies, however, consider finite-size scaling of damage spreading in RBN [14, 21]. Here, of particular interest is the "sparse percolation (SP) limit" [21], where the initial perturbation size $d(0)$ does *not* scale up with network size N, i.e., the relative size of perturbations tends to zero for large N. This limit applies to many of the above-mentioned real-world networks (e.g., the spread of a new computer virus on the internet launched from a single computer). In this letter, we systematically study finite-size scaling of damage spreading in the SP limit for both RBN and RTN. We identify a new characteristic point K_s, where the expectation value of the number of damaged nodes after large number of dynamical updates is independent of N. By the definition of marginal damage spreading, we introduce a new approach to estimate the critical connectivity $K_c(N)$ for finite N, and present evidence that, even in the large N limit, the critical connectivity for SP systematically deviates from the predictions of mean-field theory.

First, let us define the dynamics of RBN and RTN. A RBN is a discrete dynamical system composed of N automata. Each automaton is a Boolean variable with two possible states: $\{0, 1\}$, and the dynamics is such that

$$\mathbf{F} : \{0, 1\}^N \mapsto \{0, 1\}^N, \tag{1}$$

where $\mathbf{F} = (f_1, ..., f_i, ..., f_N)$, and each f_i is represented by a look-up table of K_i inputs randomly chosen from the set of N automata. Initially, K_i neighbors and a look-table are assigned to each automaton at random.

An automaton state $\sigma_i^t \in \{0, 1\}$ is updated using its corresponding Boolean function:

$$\sigma_i^{t+1} = f_i(\sigma_{i_1}^t, \sigma_{i_2}^t, ..., \sigma_{i_{K_i}}^t). \quad (2)$$

We randomly initialize the states of the automata (initial condition of the RBN). The automata are updated synchronously using their corresponding Boolean functions. The second type of discrete dynamical system we study is RTN. An RTN consists of N randomly interconnected binary sites (spins) with states $\sigma_i = \pm 1$. For each site i, its state at time $t+1$ is a function of the inputs it receives from other spins at time t:

$$\sigma_i(t+1) = \mathrm{sgn}\left(f_i(t)\right) \quad (3)$$

with

$$f_i(t) = \sum_{j=1}^{N} c_{ij}\sigma_j(t) + h. \quad (4)$$

The N network sites are updated synchronously. In the following discussion the threshold parameter h is set to zero. The interaction weights c_{ij} take discrete values $c_{ij} = +1$ or -1 with equal probability. If i does not receive signals from j, one has $c_{ij} = 0$.

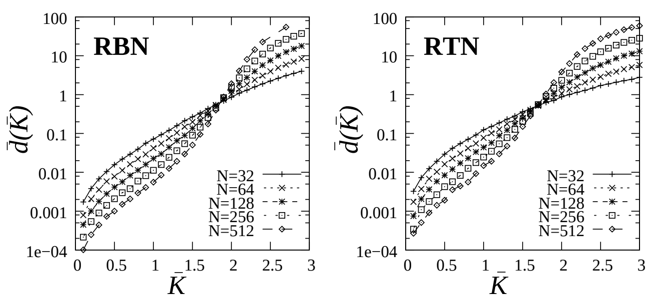

FIG. 1: Average Hamming distance (damage) $\bar{d}$ after 200 system updates, averaged over 10000 randomly generated networks for each value of $\bar{K}$, with 100 different random initial conditions and one-bit perturbed neighbor configurations for each network. For both RBN and RTN, all curves for different N approximately intersect in a characteristic point K_s.

Results. We first study the expectation value $\bar{d}$ of damage, quantified by the Hamming distance of two different system configurations, after a large number T of system updates. Let $\mathcal{N}$ be a randomly sampled set (ensemble) of z_N networks with average degree $\bar{K}$, $\mathcal{I}_n$ a set of z_I random initial conditions tested on network n, and $\mathcal{I}'_n$ a set of z_I random initial conditions differing in one randomly chosen bit from these initial conditions. Then we have

$$\bar{d} = \frac{1}{z_N\, z_I} \sum_{\substack{n=1 \\ \mathcal{N}_n \in \mathcal{N}}}^{z_N} \sum_{\substack{i=1 \\ \vec{\sigma}_i \in \mathcal{I}_n, \vec{\sigma}'_i \in \mathcal{I}'_n}}^{z_I} d_i^n(T), \quad (5)$$

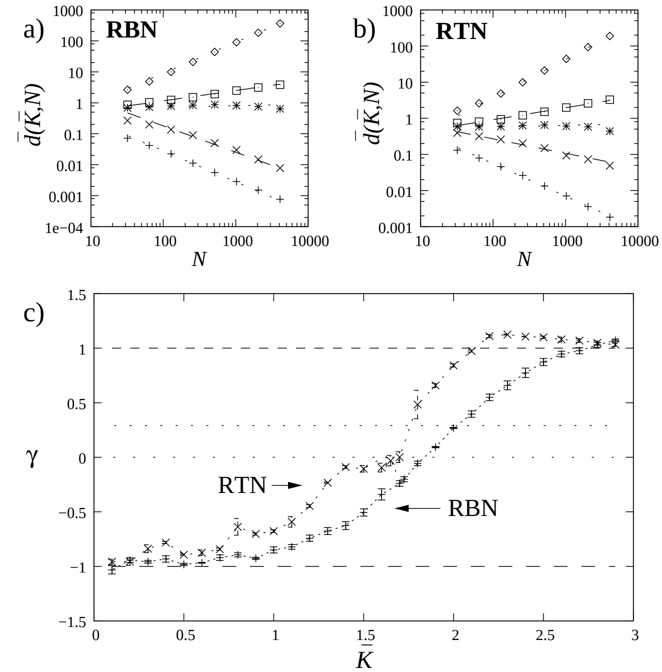

FIG. 2: *Upper panels:* $\bar{d}$ as a function of N, for different $\bar{K}$. a) $\bar{K}$ takes values 2.6, 2.0, 1.87, 1.5 and 1.0 (from top to bottom), b) $\bar{K}$ takes values 2.6, 1.85, 1.725, 1.5 and 1.0 (from top to bottom). *Lower panel*: Scaling exponents $\gamma(\bar{K})$ as a function of $\bar{K}$, as obtained from fits of Eq. (6) for RBN (+) and RTN (x). The straight dashed/dotted lines mark the asymptotes as discussed in the text.

where $d_i^n(T)$ is the measured Hamming distance after T system updates. Fig. 1 shows $\bar{d}$ as a function of the average connectivity $\bar{K}$ for different network sizes N by using a random ensemble for statistics. For both RBN and RTN, the observed functional behavior strongly suggests that the curves approximately intersect at a common point (K_s, d_s), where the observed Hamming distance for large t is independent of the system size N.

To verify this finding, let us now study the finite size scaling behavior of $\bar{d}$ in this (SP) limit. For $\bar{K} \to 0$ and for large $\bar{K}$, it is straightforward to estimate the asymptotic scaling. In the case $\bar{K} \to 0$, non-zero damage can only emerge if the initial perturbation hits a short loop of oscillating nodes (most likely a self-connection or a loop of length two, longer loops can be neglected). The *a priori* probability to generate these loops is $\sim 1/N^2$, and their number is proportional to the total number of links, $\bar{K}N$. Hence, we expect $\bar{d} \sim \bar{K}N/N^2 \propto 1/N$. For large $\bar{K}$, damage percolates through the system, consequently, avalanche sizes are bounded only by the size of the system, and we expect $\bar{d} \sim N$. At criticality, the frozen core of the network always remains undamaged, and asymptotic dynamics is determined completely by the relevant nodes, with a number n_r scaling as $n_r \sim N^{1/3}$ [19]; hence, we expect $\bar{d} \propto N^{1/3}$ at K_c, which is confirmed with high

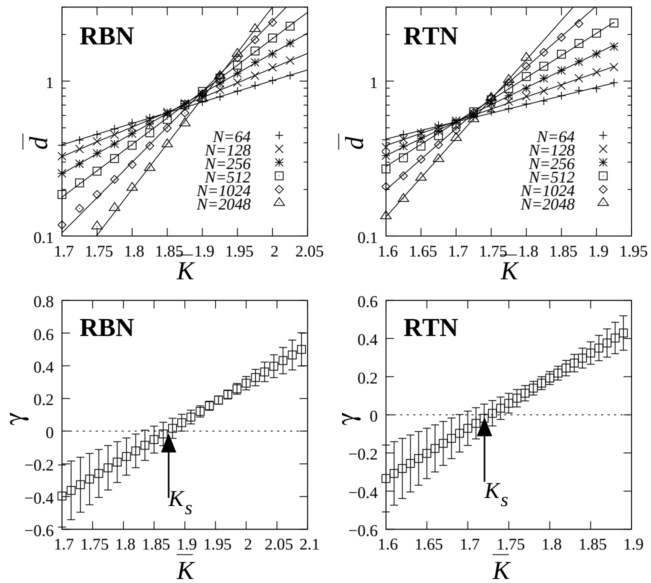

FIG. 3: *Upper panels:* $\bar{d}(\bar{K})$ in semi-log scale, as obtained from high-precision simulations near K_c (ensemble size: 50000 random networks with 100 random initial conditions for each data-point, transient time: 5000 updates). Lines are fits of Eq. (8). *Lower panels:* Scaling exponents γ derived from equating Eq. (6) and (8), with corresponding errorbars. The intersection with $\gamma = 0$ (dashed line) defines K_s.

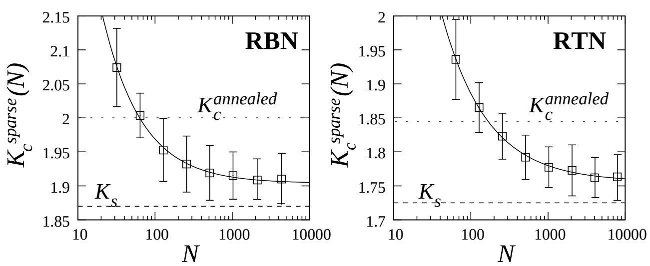

FIG. 4: The critical connectivity $K_c^{sparse}(N)$ in the SP limit as a function of N, calculated from Eq. (12). Curves are power-law fits according to Eq. (13), straight dashed lines mark $K_c^{annealed}$ and K_s for RBN and RTN, respectively.

accuracy by our numerical simulations. For arbitrary $\bar{K}$, we make the scaling ansatz

$$\bar{d}(\bar{K}, N) = a(\bar{K}) \cdot N^{\gamma(\bar{K})} + d_0(\bar{K}),\ -1 \leq \gamma \lesssim 1. \qquad (6)$$

The two upper panels of Fig. 2 illustrate that for both RBN and RTN, the numerically measured finite size scaling for $0 < \bar{K} \leq 3$ obeys this scaling ansatz very well. In both cases, at some K_s slightly below K_c, we find a transition of γ from negative to positive values (Fig. 2, lower panel). The exact determination of the point (K_s, d_s) where $\gamma \approx 0$ is difficult; because of the slow emergence of large damage events near K_c, measurements with finite T can substantially underestimate $\bar{d}$ (in particular for $N \geq 512$). We performed high precision numerical experiments in the interval $1.6 \leq \bar{K} \leq 2.1$, waiting for $T = 5000$ update time steps to let the network dynamics relax after the initial one-bit perturbation; these simulations conclusively show an exponential dependence $\bar{d} \propto \exp(c(N) \cdot \bar{K})$ in this interval, with a constant $c(N)$ depending only on N (Fig. 3, upper panels). This exponential dependence becomes apparent with the following assumptions: an increase $\Delta\bar{d}$ of the average damage is proportional to $\bar{d}$ itself (damage can generate new damage), to an increase $\Delta\bar{K}$ of the average connectivity, and to some function of the system size N. Actually, it cannot be directly proportional to N, because frozen nodes remain undamaged asymptotically. Hence, a rough upper limit is given by the number of nonfrozen nodes, which at K_c scales as $N^{2/3}$, and a lower bound by the number of relevant nodes, that almost certainly propagate damage, i.e. $N^{1/3}$ [19]. To summarize, we approximate

$$\Delta\bar{d} \approx c(N)\, N^{\alpha}\, \bar{d}\ \Delta\bar{K}, \qquad (7)$$

with $1/3 \lesssim \alpha \lesssim 2/3$; replacing $\Delta\bar{d}$ and $\Delta\bar{K}$ with differentials and integrating yields

$$\bar{d}(\bar{K}, N) \approx c_1(N)\, \exp\left[c_2(N)\, N^{\alpha}\, \bar{K}\right]. \qquad (8)$$

In simulations, we find $\alpha \approx 0.42$, which is well within the range we expect from our theoretical considerations as discussed above.

We now apply this dependence to obtain high-accuracy fits of Eq. (6) in the interval $1.6 \leq \bar{K} \leq 2.1$ (Fig. 3, lower panels); these fits yield

$$(K_s^{RBN}, d_s^{RBN}) = (1.875 \pm 0.05, 0.62 \pm 0.05) \qquad (9)$$

for RBN and, correspondingly,

$$(K_s^{RTN}, d_s^{RTN}) = (1.729 \pm 0.045, 0.51 \pm 0.04) \qquad (10)$$

for RTN.

Interestingly, K_s is close to, but distinct from the critical connectivities $K_c^{RBN} = 2$ and $K_c^{RTN} = 1.845$, as predicted by mean-field theory. However, a natural comparison has to consider possible deviations of K_C at the SP limit from these values. An intuitive definition of criticality for finite N can be formulated in terms of *marginal damage spreading*. If at time t one bit is flipped, one requires at time $t+1$ [6**?**]

$$\bar{d}(t+1) = \langle p_s \rangle (K_c) K_c = 1, \qquad (11)$$

where $\langle p_s \rangle (\bar{K})$ is the average damage propagation probability. Naturally, the iteration of this map implies $\bar{d} = 1$ for all t. Note that the relation: $\langle p_s \rangle (K_c) K_c = 1$ is exact only in the framework of the AA. In the SP limit, we instead have to set the right hand side of Eq. (8) to unity; inversion then leads to

$$K_c^{sparse}(N) \approx -\frac{\ln c_1(N)}{c_2(N) N^{\alpha}}. \qquad (12)$$

Fig. 4 shows $K_c^{sparse}(N)$, using the values $c_1(N), c_2(N)$ obtained from numerical fits of Eq. (8) for both RBN and RTN. We find that both systems, in a very good approximation, obey the scaling relationship

$$K_c^{sparse}(N) \approx b \cdot N^{-\delta} + K_c^{\infty} \tag{13}$$

with $b = 3.27 \pm 0.79$, $\delta = 0.85 \pm 0.07$ and $K_c^{\infty} = 1.9082 \pm 0.008$ for RBN and $b = 3.853 \pm 0.76$, $\delta = 0.736 \pm 0.05$ and $K_c^{\infty} = 1.7595 \pm 0.008$ for RTN. Hence, in the limit $N \to \infty$, we can extrapolate

$$K_c^{\infty,RBN} = 1.9082 \pm 0.008 \tag{14}$$

for RBN, and for RTN

$$K_c^{\infty,RTN} = 1.7595 \pm 0.008. \tag{15}$$

Thus, for both RBN and RTN in the SP limit, we make the surprising observation that K_c^{sparse} systematically deviates from $K_c^{annealed}$. While we find $K_c^{sparse}(N) > K_c^{annealed}$ for small $N < 128$, for larger N we observe a monotonic decay that approaches an asymptotic value considerably below $K_c^{annealed}$, suggesting that the observed deviations from the AA also hold in the large N limit.

It is beyond the scope of this letter to discuss possible causes for these deviations in detail (this will be accomplished in a longer paper). In simulations, however, we find that the statistical distributions of damage sizes in the SP limit are highly skewed, with most configurations leading to vanishing damage, and a fat tail of large damage events. These skewed distributions imply that with finite sampling, we always underestimate $\bar{d}$, and hence the true $K_c^{sparse}(N)$ and K_s will deviate even stronger from the AA. Also, local fluctuations in damage propagation cannot be neglected in this limit, as it is assumed in mean-field approaches.

Discussion. We investigated finite size scaling of damage spreading in both RBN and RTN near the sparse percolation (SP) transition. We find that the average damage $\bar{d}$, quantified in terms of the Hamming distance of initially nearby system states, scales $\propto N^{\gamma(\bar{K})}$ over the whole range of sparse connectivities $0 < \bar{K} \leq 3$ studied in this letter. The scaling exponents γ show a cross-over from negative to positive values at characteristic points K_s^{RBN} and K_s^{RTN} below the critical points K_c^{RBN} and K_c^{RTN}. We estimated the critical connectivities $K_c^{sparse}(N)$ based on marginal damage spreading, and found systematic deviations from the annealed approximation. While extrapolations towards large N still require some caution, the network sizes investigated in this study (currently up to $N = 32768$) are highly relevant for biological applications; compare, e.g., the Yeast genome ($N \approx 6000$) and the *E. Coli* metabolic network ($N \approx 1500$) [22]. Interestingly, recent studies suggest that gene regulatory networks appear to be in the ordered regime and reside slightly below the phase transition between order and chaos [14], while theory had proposed the critical line to be an evolutionary attractor [1, 2]. Our study may contribute a possible explanation to these observations: scaling of damage with increasing N (e.g., as a consequence of gene duplications) can be expected to be under strict selective control. Hence, the need for robust systems might drive evolution to K_s rather than to K_c. At the same time, however, K_s is close enough to criticality to enable rich dynamical behavior, as required by biological cells. Currently, this question of where in the dynamical phase space biological networks reside certainly cannot be answered conclusively. Experimental studies of regulatory networks and further studies of in-silico evolutionary processes that adapt dynamical networks, e.g., RBNs or RTNs, to robustly solve specific computational and functional tasks are required.

---

[1] S.A. Kauffman, J. Theor. Biol. **22**, 437 (1969)

[2] S.A. Kauffman, *The Origins of Order: Self-Organization and Selection in Evolution*, Oxford University Press, 1993.

[3] S. Bornholdt, Biol. Chem. **382**, 1289-1299 (2001)

[4] B. Derrida and Y. Pomeau, Europhys. Lett. 1 (1986) 45-49

[5] K.E. Kürten, Phys. Lett. A 129 (1988) 157-160; K.E. Kürten, J. Phys. A 21 (1988) L615-L619.

[6] T. Rohlf and S. Bornholdt, Physica A **310**, 245-259 (2002).

[7] R. Albert and A. L. Barabási, Phys. Rev. Lett. **84**, 5660 (2000)

[8] B. Samuelsson and C. Troein, Phys. Rev. Lett. **90**, 098701 (2003)

[9] U. Bastola and G. Parisi, Physica D **115**, 203 (1998)

[10] R. Pastor-Satorras and A. Vespignani, Phys. Rev. Lett. **86**, 3200 (2001)

[11] M.E.J. Newman, Phys. Rev. E **66**, 016128 (2002)

[12] R. Cohen, S. Havlin and D. ben-Avraham, Phys. Rev. Lett. **91**, 247901 (2003)

[13] M. L. Sachtjen, B. A. Carreras, and V. E. Lynch, Phys. Rev. E **61**, 4877 (2000)

[14] P. Ramö, J. Kesseli and O. Yli-Harja, J. Theor. Biol. **242**, 164 (2006)

[15] R. V. Solé and B. Luque, Phys. Lett. A 196 (1995), 331-334; B. Luque and R. V. Solé, Phys. Rev. E **55** (1996), 257-260

[16] S. Bornholdt and T. Rohlf, Phys. Rev. Lett. **84** (2000) 6114

[17] S. Bornholdt and T. Röhl, Phys. Rev. E **67**, 066118 (2003)

[18] M. Liu and K.E. Bassler, Phys. Rev. E **74**, 041910 (2006)

[19] V. Kaufman, T. Mihaljev and B. Drossel, Phys. Rev. E **72**, 046124 (2005)

[20] T. Mihaljev and B. Drossel, Phys. Rev. E **74**, 046101 (2006)

[21] B. Samuelsson and J. E. S. Socolar, Phys. Rev. E **74**, 036113 (2006)

[22] `http://www.yeastgenome.org` (Yeast); `http://www.ecocyc.org` (*E. Coli*)