# Solvable model for distribution networks on random graphs

D. Nasiev,[1] J. van Mourik,[1] and R. Kühn[2]

[1]*Information Engineering, Aston University, Aston Triangle, Birmingham B4 7ET, United Kingdom*
[2]*Department of Mathematics, King's College London, The Strand, London WC2R 2LS, United Kingdom*
(Received 19 June 2007; published 12 October 2007)

We propose a simple model that captures the salient properties of distribution networks, and study the possible occurrence of blackouts, i.e., sudden failings of large portions of such networks. The model is defined on a random graph of finite connectivity. The nodes of the graph represent hubs of the network, while the edges of the graph represent the links of the distribution network. Both, the nodes and the edges carry dynamical two state variables representing the functioning or dysfunctional state of the node or link in question. We describe a dynamical process in which the breakdown of a link or node is triggered when the level of maintenance it receives falls below a given threshold. This form of dynamics can lead to situations of catastrophic breakdown, if levels of maintenance are themselves dependent on the functioning of the net, once maintenance levels locally fall below a critical threshold due to fluctuations. We formulate conditions under which such systems can be analyzed in terms of thermodynamic equilibrium techniques, and under these conditions derive a phase diagram characterizing the collective behavior of the system, given its model parameters. The phase diagram is confirmed qualitatively and quantitatively by simulations on explicit realizations of the graph, thus confirming the validity of our approach.

PACS number(s): 02.50.−r, 89.65.−s, 05.45.−a, 05.70.Fh

## I. INTRODUCTION

In recent years, the study of physical properties of networks has received much attention. This is partly a response to the fact that modern societies are increasingly relying on network based technologies, including mobile and land-line telecommunication, the Internet or, lately, grid computing, but also on traditional transport infrastructure, such as rail or road networks. Much of the interest is also fueled by the realization that complex systems in biology [1–3], chemistry [4], sociology [5], and economy [6,7] can be analyzed using various network related techniques and paradigms [8,9]. The issues here are to *understand* the workings of a given complex system in terms of properties of the underlying network in terms of which it might be described, and given some fundamental understanding has been obtained, perhaps go further and characterize the way in which the efficiency of a system would depend on properties of the net.

Network properties of interest include statistical measures characterizing structural and topological aspects globally over a network, such as the degree distribution, clustering coefficients, centrality, percolation thresholds, and more, as well as more specific information concerning, e.g., the presence or absence of certain motifs in (local) connectivity patterns.

Connectivity issues aside, both the nodes and the links in a network may be further characterized by (graded) quality measures. The capacity of a specific directed information channel, the current carrying capacity of a power line, or just a distance between two nodes in a net would constitute examples of graded link properties, whereas the computing power of a server, the capacity of a local water reservoir, the susceptibility of an individual to infection by a virus (or an opinion, or by the desire to acquire a new gadget) would be examples of properties specifically associated with "nodes."

It goes without saying that functionality as well as efficiency of network based operations will depend in various

degrees on the properties of a given network. Points of concern then are robustness of *efficiency*, and—on a more basic level—of *functionality* against unexpected failures of links or nodes, or the resilience of network based operations against directed attacks [10–13]. Clearly these issues have implications in the realm of biological evolution; they would have to influence design decisions in engineering contexts, or political decisions, e.g., when setting up supply infrastructure. On a day-to-day basis, financial institutions are nowadays *required* to set aside capital to cover financial losses incurred by process failures (operational risk); as processes in organizations would normally be set up in a way to mutually support each other, collective effects creating the possibility of extreme events in large process networks are clearly relevant [14,15] and need to be properly quantified in order to set aside the right amount of capital. Erring in either direction would be costly (for different reasons) and affect the competitiveness of the organization in question.

The present investigation is concerned with issues of resilience of network functionality against random failures or directed attacks in supply infrastructure and distribution networks, such as power grids, various forms of information transfer networks (telephone, Internet), traffic systems (logistics), or metabolic networks, where these issues are particularly relevant. Failures of nodes and/or links can threaten the basic functionality of the network and if cascading through the system (or affecting hubs of central importance), even lead to a global breakdown. Examples are major blackouts in power grids, massive gridlock in traffic systems, or death of an organism.

We shall look at a simplified setting in which the breakdown of a link or node is triggered when the level of maintenance or support it receives falls below a given threshold, leaving the case where failures are induced by loads exceeding critical levels to a future study [16]. The simplified setting renders our model a generalization of previously studied

041120-1

models for operational risk [14,15] that includes dynamical properties link variables in the analysis.

Our paper is organized as follows. In Sec. II we introduce our model, describing its coupled link and node dynamics and interpreting the parameters of the model in terms of *a priori* and conditional probabilities for link and node failure. Conditions are identified under which the system can be analyzed using thermodynamic equilibrium methods. The statistical mechanics approach to analyze long term properties of the model under these conditions is briefly outlined in Sec. III. Results in terms of a phase diagram characterizing the collective behavior of the system, given its model parameters are presented and discussed in Sec. IV. The phase diagram is confirmed qualitatively and quantitatively by simulations on explicit realizations of the graph, thus confirming the validity of our approach. Section V concludes with a summary and an outlook on future lines of research.

## II. MODEL

We describe a distribution network as a set of $N$ vertices or nodes that are interconnected by $N_e$ dynamic edges or links. The nodes are thought of as relay stations that organize the (re)distribution of "goods" through the links by which they are connected. With each node $i$ we associate a dynamical variable $v_i$ that can be either 1 or 0, designating whether the node is up and running ($v_i=1$) or in a nonoperational state ($v_i=0$). Similarly, with each pair $(ij)$ of nodes which are connected in the distribution network in question, we associate a dynamic two-state link variable $\ell_{ij}$ that describes whether a link is operational ($\ell_{ij}=1$), or down ($\ell_{ij}=0$).

The distribution network can be thought of as a dynamical structure on a (random) graph. The *structure* of the underlying graph is completely determined by the symmetric connectivity matrix $\mathbf{c}=\{c_{ij}:i\neq j=1,\ldots,N\}$, with $c_{ij}=1$ if nodes $i$ and $j$ are connected, and $c_{ij}=0$ otherwise. In this paper, we restrict ourselves to the ensemble of so-called Erdös-Rényi graphs, which are determined by the following probability distribution:

$$P(\mathbf{c}) = \prod_{i<j}\left[\left(1-\frac{C}{N}\right)\delta_{c_{ij},0}+\frac{C}{N}\delta_{c_{ij},1}\right]\delta_{c_{ij},c_{ji}}, \quad (1)$$

i.e., every possible link $(ij)$ is present with probability $\frac{C}{N}$. This typically results in a random graph where the coordination numbers $L_i$ of the nodes are Poisson distributed with parameter $C$ (the average connectivity): $P_C(L)=\frac{C^L}{L!}e^{-C}$. The total number of links in the graph is then given by $N_e=\frac{CN}{2}$, as each edge connects exactly two nodes. For any finite $C$ (in comparison to $N$), the graph is sparse, the nodes have finite connectivity, and the number of nodes $N$ and the number of edges $N_e$ are of the same order of magnitude.

A dynamical evolution of the node and link variables on the graph can be described and motivated in analogy with previous operational risk modeling [14] as follows. Define a node support as

$$h_i = \sum_{j(\neq i)} c_{ij}\ell_{ij}(a_{ij}v_j + b_{ij}) + \mu_i, \quad (2)$$

and a link support as

$$h_{ij} = \tilde{a}_{ij}v_iv_j + \tilde{b}_{ij}(v_i+v_j) + d_{ij}. \quad (3)$$

In these expressions, the parameter $a_{ij}$ would quantify the contribution of node $j$ to the support of $i$, while $b_{ij}$ could describe external support to $i$ that is funneled through the link $(ij)$, while $\mu_i$ would describe a baseline support independent of the dynamical state of the system itself.

Similarly, $\tilde{a}_{ij}$ would describe a contribution to the maintenance of link $(ij)$ that originates from a collaboration of the nodes connected by it, while $\tilde{b}_{ij}$ would quantify separate contributions from adjacent functioning nodes which would exist without collaboration, and $d_{ij}$ would again characterize a baseline support independent of the dynamical state of the system itself.

We suggest a probabilistic asynchronous dynamics which stipulates that vertices and nodes fail, if they receive less than a critical threshold support. Without loss of generality the thresholds can be taken to be zero (by appropriate definition of baseline support levels). Thus, in formal terms we have

$$v_i(t+\Delta t) = \Theta[h_i(t)+\xi_i(t)], \quad (4)$$

$$\ell_{ij}(t+\Delta t) = \Theta[h_{ij}(t)+\xi_{ij}(t)], \quad (5)$$

where the vertex and edge noises are taken to be independent and white, and either zero mean Gaussians with variances $\sigma$ and $\tilde{\sigma}$, respectively,

$$\langle\xi_i(t)\xi_j(t')\rangle = \sigma^2\delta_{i,j}\delta_{t,t'}, \quad (6)$$

$$\langle\xi_{ij}(t)\xi_{kl}(t')\rangle = \tilde{\sigma}^2\delta_{(ij),(kl)}\delta_{t,t'}, \quad (7)$$

or thermal noises with probability density functions

$$p(\xi_i) = \frac{1}{2}\frac{d}{d\xi_i}\tanh\left(\frac{\beta\xi_i}{2}\right), \quad (8)$$

$$p(\xi_{ij}) = \frac{1}{2}\frac{d}{d\xi_{ij}}\tanh\left(\frac{\tilde{\beta}\xi_{ij}}{2}\right). \quad (9)$$

By integrating over the noises in Eqs. (4) and (5) one obtains

$$P(\ell_{ij}(t+\Delta t) = 1) = \langle\ell_{ij}(t+\Delta t)\rangle = \phi(h_{ij}(t)/\tilde{\sigma}) \quad (10)$$

$$P(v_i(t+\Delta t) = 1) = \langle v_i(t+\Delta t)\rangle = \phi(h_i(t)/\sigma), \quad (11)$$

in the case of Gaussian noises, with $\phi(x)$ denoting the integrated unit-variance normal density, while the corresponding result for the thermal noise model would be

$$P(\ell_{ij}(t+\Delta t) = 1) = \langle\ell_{ij}(t+\Delta t)\rangle = \phi_{\tilde{\beta}}(h_{ij}(t)), \quad (12)$$

$$P(v_i(t+\Delta t) = 1) = \langle v_i(t+\Delta t)\rangle = \phi_{\beta}(h_i(t)), \quad (13)$$

with

$$\phi_{\beta}(x) = \frac{1}{2}\left[1+\tanh\left(\frac{\beta x}{2}\right)\right]. \quad (14)$$

The collective properties of the system do not crucially depend on which noise model is chosen as long as the noise

levels are properly matched. We shall in what follows mostly work with the thermal noise model.

Note that the current formulation of the dynamics is in terms of resources or the lack of them as responsible for the functioning or the failure of nodes and links. In particular, one would sensibly expect the parameters $a_{ij}$ and $\tilde{a}_{ij}$ as well as the $b_{ij}$ and $\tilde{b}_{ij}$ to be positive within this interpretation. The case where failure of links or nodes is triggered by load levels exceeding certain critical values is left to a separate study [16].

The parameters $d_{ij}$ and $\mu_i$ are related to the *a priori* probability that the isolated link ($ij$) or the isolated node $i$, respectively, are up. In a similar vein the $a_{ij}$ and $\tilde{a}_{ij}$ as well as the $b_{ij}$ and $\tilde{b}_{ij}$ can be related to *conditional* probability of node or link failure (or survival) for given configurations of node and link-states, in analogy to operational risk modeling [14,15].

In the present investigation, we will restrict our attention to a case in which the model parameters satisfy a number of symmetry relations, which ensure that the long time behavior of the system can be described by thermodynamic equilibrium theory. These are

$$\forall_{ij}: \quad a_{ij} = \tilde{a}_{ij}, \quad b_{ij} = \tilde{b}_{ij}, \tag{15}$$

and

$$\forall_{i<j}: \quad a_{ij} = a_{ji}, \quad b_{ij} = b_{ji}, \tag{16}$$

as well as $a_{ii}=0$ and $b_{ii}=0$ and equality of noises of the link and node dynamics,

$$\beta = \tilde{\beta}. \tag{17}$$

These conditions together are sufficient to ensure that the stochastic dynamics (13) and (12) satisfies detailed balance with respect to the Gibbs-Boltzmann equilibrium distribution at inverse temperature $\beta$ that is generated by the energy function

$$\mathcal{H} = -\sum_{(ij)} c_{ij}\ell_{ij}[a_{ij}v_i v_j + b_{ij}(v_i + v_j) + d_{ij}] - \sum_i \mu_i v_i \tag{18}$$

(see the Appendix). This feature allows one to study the collective behavior of the distribution networks by the methods of equilibrium statistical mechanics, which greatly simplifies the analysis, and will be the main focus of the remainder of the present paper.

Clearly a full dynamical study is required whenever any of the symmetry assumptions is violated. However, it is well known that thermodynamic equilibrium theory often gives a fair qualitative description, provided the violations are not too drastic.

## III. THEORETICAL ANALYSIS

For the analysis of the model described in the previous section, we use the replica method for sparsely connected systems [17–20]. In contrast to previous models studied with this method, we have the new ingredient that not only the

nodes but also the links have dynamical variables, though we shall find that this introduces only minor complications in the analysis.

The general strategy is to calculate the free energy of the model, from which all relevant thermodynamic quantities can be derived. The free energy is expected to be self-averaging, entailing that, in the large system limit and under very general conditions on the disorder distribution in the model (i.e., **c**, $a_{ij}$, $b_{ij}$, $d_{ij}$, and $\mu_i$), the average of the free energy over all disorder configurations coincides with its typical value: the free energy of any explicit realization of the problem drawn from this ensemble will be equal to the average with probability 1.

Given the Hamiltonian (18), the disorder dependent partition function $\mathcal{Z}$ and the (dimensionless) free energy $\mathcal{F}$ are obtained from

$$\mathcal{Z}(\mathbf{c},\{a,b,d,\mu\}) = \text{Tr}_{\{v\}} \, \text{Tr}'_{\{\ell\}} \exp(-\beta\mathcal{H}),$$

$$\mathcal{F}(\mathbf{c},\{a,b,d,\mu\}) = -\ln(\mathcal{Z}(\mathbf{c},\{a,b,d,\mu\})),$$

where Tr$'$ denotes a trace over those $\ell_{ij}$ for which $c_{ij}=1$. In order to average the free energy over the disorder, we use the replica trick $\langle\ln(\mathcal{Z})\rangle_D = \lim_{n\to 0} n^{-1}\ln\langle\mathcal{Z}^n\rangle_D$, so that we have to calculate the average of the replicated partition function $\langle\mathcal{Z}^n\rangle_D$ as follows:

$$\langle\mathcal{Z}^n\rangle_D = \left\langle \text{Tr}_{\{v^\alpha\}} \, \text{Tr}'_{\{\ell^\alpha\}} \exp\left[\beta\sum_\alpha \left(\sum_{(ij)} c_{ij}\ell_{ij}^\alpha[a_{ij}v_i^\alpha v_j^\alpha \right.\right.\right.$$
$$\left.\left.\left. + b_{ij}(v_i^\alpha + v_j^\alpha) + d_{ij}] + \sum_i \mu_i v_i^\alpha\right)\right]\right\rangle_D. \tag{19}$$

Here $\langle \ \rangle_D$ indicates the average over the disorder. At this point we do not yet need to specify the disorder distribution for the $\{a_{ij}, b_{ij}, d_{ij}, \mu_i\}$. We start by performing the average over the connectivity matrix **c** and the trace over the link variables $\{\ell^\alpha\}$, which can be easily done as the replicated partition sum factorizes over these variables. This is a well-known fact, which has been extensively used in the study of Ising models with annealed bond disorder [21].

After some relatively standard manipulations, which involve introducing the order parameter

$$\rho(\tilde{v}) \equiv \frac{1}{N}\sum_{i=1}^N \delta_{\tilde{v},\tilde{v}_i}, \quad \tilde{v} \equiv \{v^\alpha: \ \alpha = 1, \ldots, n\}, \tag{20}$$

and its conjugate $\hat{\rho}(\tilde{v})$ to enforce the definition, one arrives at the following compact expression for $\langle\mathcal{Z}^n\rangle_D$:

$$\langle\mathcal{Z}^n\rangle_D = \int \mathcal{D}\{\rho(\tilde{v})\hat{\rho}(\tilde{v})\} e^{(N[C/2G_1(\rho)-G_2(\rho,\hat{\rho})+G_3(\hat{\rho})])}. \tag{21}$$

The functionals appearing in Eq. (21) are given by

$$G_1 = \text{Tr}_{\tilde{v},\tilde{w}} \, \rho(\tilde{v})\rho(\tilde{w})\langle\prod_\alpha (1 + e^{\beta[av^\alpha w^\alpha + b(v^\alpha + w^\alpha) + d]})\rangle_{a,b,d} - 1,$$

$$G_2 = \text{Tr}_{\tilde{v}} \, \rho(\tilde{v})\hat{\rho}(\tilde{v}),$$

$$G_3 = \ln\langle\mathrm{Tr}_{\tilde{v}}\exp[\mu|\tilde{v}| + \hat{\rho}(\tilde{v})]\rangle_\mu, \qquad (22)$$

where we have used the shorthand $|\tilde{v}| \equiv \Sigma_\alpha v^\alpha$.

For the analytic continuation $n \to 0$, we need to make an ansatz about the symmetry between the replicas, and assume replica symmetry (RS)

$$\rho(\tilde{v}) = \int_0^\infty d\pi(x)\frac{x^{|\tilde{v}|}}{(1+x)^n}, \qquad (23)$$

$$\hat{\rho}(\tilde{v}) = \lambda_0\int_0^\infty d\hat{\pi}(\hat{x})\frac{\hat{x}^{|\tilde{v}|}}{(1+\hat{x})^n}. \qquad (24)$$

Further introducing

$$C_0 = 1 + e^{\beta d}, \quad C_1 = 1 + e^{\beta(b+d)}, \qquad (25)$$

$$C_2 = 1 + e^{\beta(a+2b+d)}, \quad \{x_l\}_L = \prod_{l=1}^L x_l, \qquad (26)$$

and following standard reasoning, we arrive at the typical RS free energy per node

$$\mathcal{F} = -\frac{1}{\beta}\left(\frac{C}{2}\mathcal{G}_1(\pi) - C\mathcal{G}_2(\pi,\hat{\pi}) + \mathcal{G}_3(\hat{\pi})\right), \qquad (27)$$

where

$$\mathcal{G}_1 = \int_0^\infty \{d\pi(x_l)\}_2\langle\ln[C_0 + C_1(x_1 + x_2) + C_2 x_1 x_2]\rangle_{a,b,d},$$

$$\mathcal{G}_2 = \int_0^\infty d\pi(x)d\hat{\pi}(\hat{x})\ln(1 + x\hat{x}),$$

$$\mathcal{G}_3 = \sum_L P_C(L)\int_0^\infty \{d\hat{\pi}(\hat{x}_l)\}_L\langle\ln(1 + e^{\beta\mu}\{\hat{x}_l\}_L)\rangle_\mu. \quad (28)$$

It is understood that Eq. (27) has to be evaluated in the saddle point with respect to $\pi$ and $\hat{\pi}$. This implies that $\pi$ and $\hat{\pi}$ have to satisfy the self-consistency equations

$$\pi(x) = \sum_L \frac{LP_C(L)}{C}\int_0^\infty \{d\hat{\pi}(\hat{x}_l)\}_{L-1}\langle\delta(x - e^{\beta\mu}\{\hat{x}_l\}_{L-1})\rangle_\mu,$$

$$\hat{\pi}(\hat{x}) = \int_0^\infty d\pi(x)\left\langle\delta\left(\hat{x} - \frac{C_1 + C_2 x}{C_0 + C_1 x}\right)\right\rangle_{a,b,d}. \qquad (29)$$

From the free energy, one can obtain physically relevant quantities, such as the fraction $m_n$ of working nodes,

$$m_n = \sum_L P_C(L)\int_0^\infty \{d\hat{\pi}(\hat{x}_l)_L\}\frac{e^{\beta\mu}\{\hat{x}_l\}_L}{1 + e^{\beta\mu}\{\hat{x}_l\}_L}, \qquad (30)$$

the fraction $m_\ell$ of working links,

$$m_\ell = \int_0^\infty \{d\pi(x_l)\}_2\frac{M_0 + M_1(x_1 + x_2) + M_2 x_1 x_2}{C_0 + C_1(x_1 + x_2) + C_2 x_1 x_2}, \qquad (31)$$

in which $M_i = C_i - 1$, and the internal energy

$$\mathcal{U} = -\frac{C}{2}\int_0^\infty \{d\pi(x_l)\}_2\left\langle\frac{E_0 + E_1(x_1 + x_2) + E_2 x_1 x_2}{C_0 + C_1(x_1 + x_2) + C_2 x_1 x_2}\right\rangle_{a,b,d}$$

$$- \sum_L P_C(L)\int_0^\infty \{d\hat{\pi}(\hat{x}_l)\}_L\left\langle\frac{\mu e^{\beta\mu}\{\hat{x}\}_L}{(1 + e^{\beta\mu}\{\hat{x}\}_L)}\right\rangle_\mu, \qquad (32)$$

with $E_2 = (a + 2b + d)e^{\beta(a+2b+d)}$, $E_1 = (b+d)e^{\beta(b+d)}$, and $E_0 = de^{\beta d}$.

One should note that the inverse temperature $\beta$ can be absorbed into the parameters $\{a, b, d, \mu\}$ and is only used to derive some of the physical quantities. In what follows, we implicitly assume that $\beta = 1$.

At this point one should also note that although the links are dynamical variables, it was not necessary to introduce a separate order parameter to describe the link distribution. It turns out that in any such model the trace over the link variables (irrespective of the type of variable) can be carried out directly as the partition function factorizes over the links.

It should be noted that the saddle-point equations (29) can only be solved numerically using, e.g., a population dynamics algorithm [22]. The analysis of these equations, and the implications different types of solution have concerning the possible occurrence of sudden global system failures are presented in the next section.

## IV. RESULTS

We solve Eqs. (29) numerically for fixed $\beta = 1$ and various connectivities $C$ and distributions of $a$, $b$, $d$, and $\mu$, using a population dynamics algorithm [22].

After extensive numerical experiments we have obtained the full (at least five-dimensional) phase diagram, and we find that the parameter space can be qualitatively divided into four distinct regions: (O) the operational state ($m_n \simeq 1$) is the only stable solution; (N) the nonoperational state ($m_n \simeq 0$) is the only stable solution; (CO) coexistence of the stable operational state with a metastable nonoperational state; and (CN) coexistence of the stable nonoperational state with a metastable operational state.

We have not found any region where more than two locally stable phases coexist. The boundaries between the single-solution and multiple-solution regions are spinodal surfaces (associated with dynamical transitions).

As plots of phase diagrams in more than two dimensions are not easy to interpret quantitatively, we have restricted ourselves to presenting two-dimensional plots in the $(a, \mu)$ plane, for several combinations of the other parameters, thus indicating the major trends and characteristics of the full phase diagram. In all plots, the numerical precision is of the order of the line widths.

With reference to Figs 1–4, we refer to the boundary between the (O) and (CO) region as the *upper* spinodal, and the boundary between the (N) and (CN) region as the *lower* spinodal. The upper spinodal demarcates the locations in parameter space where the nonoperational (N) phase ceases to exist as a metastable phase, while the lower spinodal gives the locations where the operational (O) phase becomes
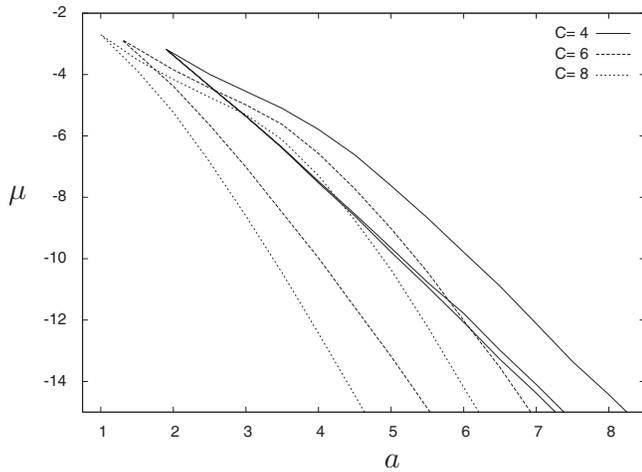
FIG. 1. A section of the phase diagram in the $(a,\mu)$ plane, with $b=d=0$ for three different average connectivities: $C=4$ (full lines), $C=6$ (dashed lines), and $C=8$ (dotted lines). For each, the top line is the upper spinodal [marking the boundary between the (CO) and the (O) region], the middle line is the equilibrium transition, and the bottom line is the lower spinodal [marking the boundary between the (CNO) and the (N) region]. Note that for average connectivities $C=6$ and 8, the transition and the lower spinodal are so close that they cannot be distinguished on the scale of this plot, thus reducing the (CN) region to a very narrow strip.

unstable. The boundary between the (CO) and (CN) regions is given by the surface in parameter space where the free energies of the two solutions coincide, i.e., they mark the thermodynamic equilibrium transition.

The two spinodals and the thermodynamic transition may meet (end) in lines (more precisely, lower dimensional surfaces) of critical points (analogous to critical points of liquid-gas systems), so that it is always possible to go from the (O) to the (N) region continuously, following some path through the phase diagram that avoids any transitions.

Except when explicitly mentioned, all plots shown have
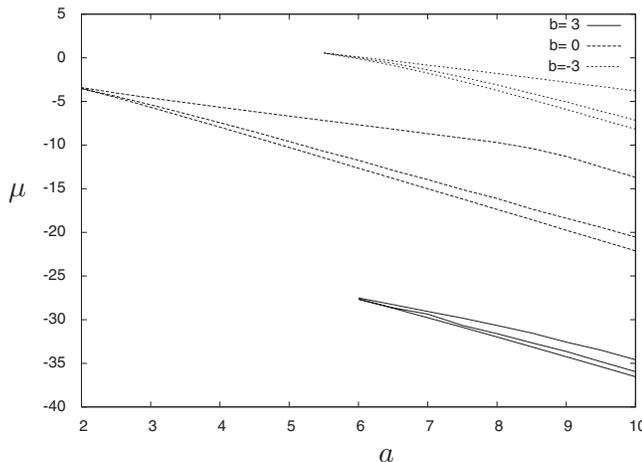


FIG. 2. A section of the phase diagram in the $(a,\mu)$ plane, with $d=0$, $C=4$ for three different values of $b$: $b=3$ (full lines), $b=0$ (dashed lines), and $b=-3$ (dotted lines). For each, the top line is the upper spinodal, the middle line is the transition, and the bottom line is the lower spinodal.
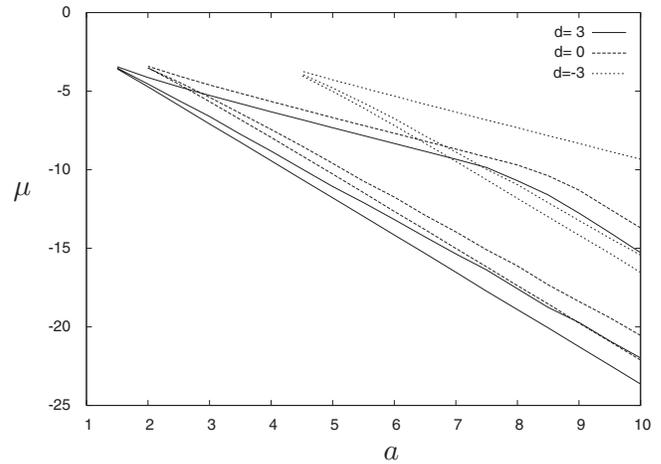


FIG. 3. A section of the phase diagram in the $(a,\mu)$ plane, with $b=0$, $C=4$ for three different values of $d$: $d=3$ (full lines), $d=0$ (dashed lines), and $d=-3$ (dotted lines). For each, the top line is the upper spinodal, the middle line is the transition, and the bottom line is the lower spinodal.

been made for constant parameters $a$, $b$, $d$, and $\mu$. We have also solved the model with varying parameters, e.g., $a_{ij} \sim \mathcal{N}(\langle a\rangle, \sigma)$. The general tendency is that an increase in the variance $\sigma$ leads to a reduction of the (CO) and (CN) regions (pulling the spinodals closer to the transition), while the thermodynamic transition remains virtually unchanged. This implies that an increase in the variability of the resources reduces the region where the working system is *locally* stable, thus increasing the likelihood of sudden system failures.

Since both $m_n$ and $m_\ell$ depend on the same distributions $\pi$ and $\hat{\pi}$, either observable can be used to monitor the dynamical transitions. In general, any discontinuity in $m_n$ corresponds to a discontinuity in the same direction in $m_\ell$, although the amplitude of the discontinuity depends on the relative strength of the various parameters. We note that the operational risk model (without dynamical link variables) as presented in [15] can be recovered by taking the parameter $d \to \infty$.
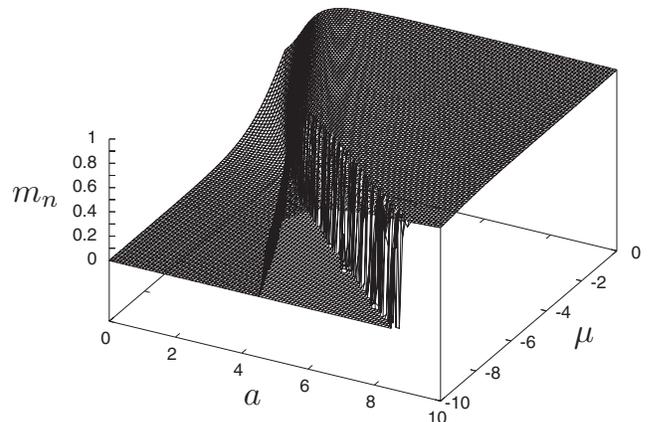


FIG. 4. The fraction of nodes that are 1 $(m_n)$, as a function of $a$ and $\mu$ for $C=4$, $b=d=0$, both starting from the (N) phase and the (O) phase.
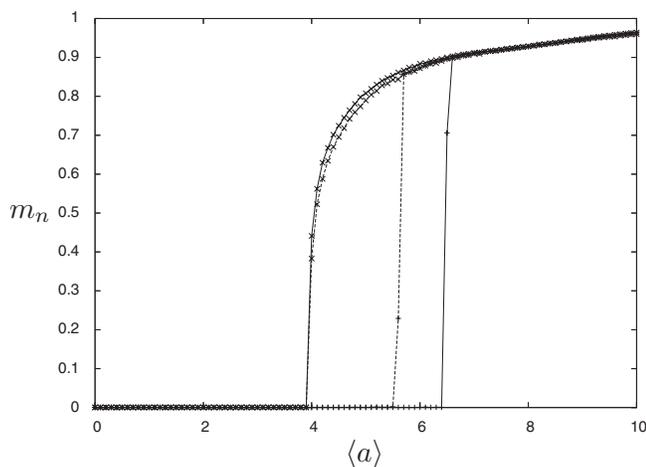
FIG. 5. Monte Carlo simulation of a system of $10^5$ nodes, with parameters $C=4$, $\langle\mu\rangle=-8$, $\langle b\rangle=0$, and $\langle d\rangle=0$, exhibiting hysteresis of $m_n$ as a function of $\langle a\rangle$ across the transition. Connecting solid lines are for uniform parameters ($\sigma=0$), dotted lines are for nonuniform parameters with $\sigma=0.5$. The symbols "$\times$" indicate that we started from the (O) solution, while "$+$"indicates we started from the (N) solution. Note that the fluctuations of $m_n$ over the measuring sweeps are of the order of the line widths. We used $10^2$ equilibration and measuring sweeps for each value of $\langle a\rangle$.

The (CO) and (CN) regions are of particular interest with respect to the possible occurrence of sudden global system failures. Starting from a state where the working system is globally stable, one may, by gradually increasing the average load or by gradually reducing the available resources, push the system over the thermodynamic transition into a region where the operational (O) phase no longer corresponds to the global minimum of the free energy, *without realizing it*, as there are no detectable signatures of such a (first order) equilibrium transition.

If the system is pushed further in this direction, it will certainly collapse once the spinodal is reached. However, as long as the system is still relatively close to the equilibrium transition point, it may continue operating for a very long time until a rare unfavorable fluctuation tips it over the edge, and a global collapse takes place. Both the system size, and the distance from the spinodal will determine the likelihood that rare unfavorable fluctuations actually manage to destabilize the system. The time before the collapse occurs, fluctuates wildly, hence it is impossible to accurately predict the timing of global system failures.

In order to confirm the validity of our analytical solution, we have also performed Monte Carlo simulations of the model using a variant where node and link variables are pooled and random sequentially updated according to the Metropolis algorithm. In general, our simulation results for spinodals perfectly coincide with the theoretical predictions to within numerical precision. Therefore we have not presented separate figures that compare the two. We do, however, include a simulation illustrating hysteresis as one of the parameters is varied across the transition region (Fig. 5) as well as a run on a small system in Fig. 6, exhibiting the
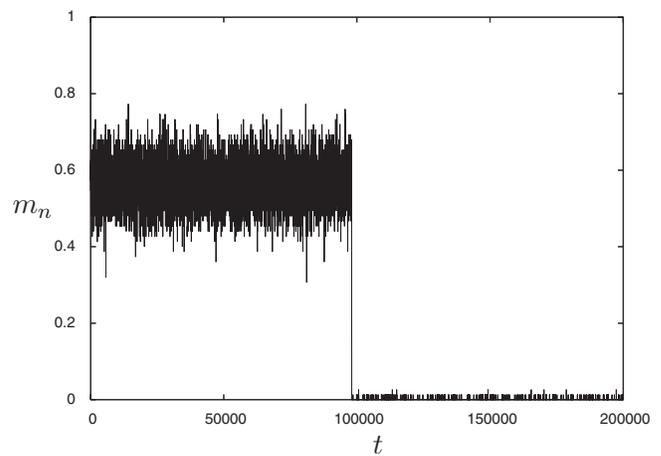


FIG. 6. Simulation of a distribution network of 75 nodes, with constant parameters $C=4$, $\langle\mu\rangle=-8$, $\langle b\rangle=0$, $\langle d\rangle=0$, and $\sigma=0$. Plotted is $m_n$ as a function of time (measured in Monte Carlo steps per degree of freedom). Starting from the metastable (O) phase, the system collapses after about $10^5$ sweeps to the globally stable (N) phase. Note that the large fluctuations of $m_n$ are due to the small system size.

eventual collapse through bubble nucleation of a metastable (O) phase into the globally stable (N) phase.

Incidentally, the agreement between the theoretical (RS) predictions and the corresponding simulation results, is a strong indication that there is no replica symmetry breaking in the studied region of parameter space. There are two main reasons for this. First, the interesting and relevant region in parameter space corresponds to predominantly cooperative relations between node and link variables, so that the degree of frustration in the system is low. Apart from this, effects of frustration are further reduced due to the dynamical nature of the links.

## V. CONCLUSION AND OUTLOOK

We have investigated the performance of distribution networks realized as sets of nodes interacting via dynamical links arranged on a (random) graph. In the present investigation, we restricted ourselves to so-called Erdös-Renyi graphs with Poissonian connectivity distributions. We have formulated the coupled node and link dynamics in such a way that breakdown of a link or node is caused when a level of support it receives falls below a given threshold. We have identified conditions under which the collective behavior of such systems can be analyzed using equilibrium statistical mechanics, and we perform such an analysis using replica and mean-field techniques for finitely connected random systems.

The model generalizes previous models of operational risk, in which link-dynamics is included as an essential new ingredient. Formally the conventional operational risk situation can be recovered by considering a limit in which the links are infinitely resilient against failure, thereby freezing out the dynamics of links.

A phase diagram characterizing the global behavior of the system and its dependence on the system parameters has

been obtained, which confirms and quantifies intuitions one would have about such a system. In the region of parameter space of interest for an interpretation as distribution networks the system can be in one of two possible phases, an operational phase (O) where most nodes and links are up and running, and a nonoperational phase (N) where most nodes and links are down. In certain regions of parameter space, both phases can coexist; these are further divided into subregions where either the operational phase (CO) or the nonoperational phase (CN) are the absolutely stable ones. These subregions are separated by a thermodynamic first order equilibrium transition. Coexistence regions are bounded by spinodal surfaces, which mark locations in parameter space where the metastable phases become unstable (and therefore cease to exist). We find that spinodals can meet in lower-dimensional surfaces of critical second-order transitions (much as in liquid-gas systems).

One observes the following main trends. By increasing average connectivity $C$, or by increasing the resilience of links against spontaneous failure (parametrized by $\langle d \rangle$), or—to a lesser extent—by increasing the resilience of nodes against spontaneous failure (parametrized by $\langle \mu \rangle$) the coexistence region is moved to lower values of the parameter $\langle a \rangle$, which quantifies the individual contributions to node support as well as the cooperative link support. These trends are clearly in line with intuition one would have about systems of this type, concerning the beneficial roles of redundancy of resources or the reliability of nodes and links. There is a concurrent increase of the width of the coexistence region in the $\langle a \rangle$-$\langle \mu \rangle$ plane, though throughout most of the coexistence region the operational (O) phase appears to be the absolutely stable phase (the nonoperational phase the metastable one).

For systems described by our theory, a characterization of phases as stable or metastable could provide a more systematic and more general approach to assess the stability of a system than current versions of so-called "well-being analysis" [23,24] for power grids. Within a well-being analysis, states are classified according to whether they would survive the failure of a link or node chosen from a preestablished contingency list (which would contain important supply units or transmission lines) without shedding load (such states are characterized as "healthy"), or only at the cost of shedding load (in which case states are called "marginal"), while the "at risk" states are those in which load is shed. Thermodynamic stability would not only check the stability of the system against individual failures within a handcrafted list, but also against rare unfavorable fluctuations which include combined events. If desired, thermodynamic stability could indeed be supplemented by checking stability against failure of selected elements, by looking at the collective behavior of the system (evaluating restricted partition functions and free energies) with a selected fraction of nodes and/or links frozen in the nonoperational state.

From the point of view of assessing the risk of failure of the net, the coexistence regions are clearly the most important regions in parameter space. In these regions, an all operational state of the system coexists with a nonoperational state. Parameter changes, which drive the system into the coexistence region, would not be detectable in the collective

behavior of the system and, as already noted in [15], the same is true, in particular, also for parameter changes that exchange the relative stabilities of the operational (O) and the nonoperational (N) phase. As a consequence, parameter changes that would result in making catastrophic breakdown of the system an event that is eventually *bound to occur* under normal operating conditions would go unnoticed.

In the present investigation we have restricted ourselves to situations in which the system parameters satisfy a set of symmetry relations, allowing us to use equilibrium methods to analyze collective behavior. We have checked by simulations that moderate violations of these symmetry conditions do not substantially alter the macroscopic properties of the system. This has also been observed for (somewhat simpler) models of operational risk [15], in which the violation of detailed balance by either changing the noise model or by adopting asymmetric interactions, has not led to qualitative changes in the collective behavior of the system. Nevertheless, proper dynamics techniques and/or numerical simulations would be required to study this issue in greater detail for the present model.

Changing the topology of the net (e.g., introducing scale-free link distributions) as they are often observed in real world networks is easily accomplished and we are currently looking at this case [16].

A more important point concerns the *mechanisms* triggering failures of nodes or links. In many distribution networks the dominant mechanism is related to traffic along certain links or through certain nodes exceeding critical values, which triggers the failure of a given element, and rerouting traffic appears as a crucial ingredient responsible for cascading failures in the system. This aspect is not covered in the present investigation, but it is clearly of great importance. An investigation taking these constraints into account is under way.

### APPENDIX: DETAILED BALANCE

Here, we show that under the conditions stated in Eqs. (15)–(17), the dynamics (12)–(14) satisfies detailed balance with respect to the Gibbs-Boltzmann distribution with Hamiltonian (18).

We first consider a transition $\{v, \ell\} \rightarrow \{v, \ell\}'$ with $v_i = 0$, $v_i' = 1$, and $v_j' = v_j$, $\ell_{ij}' = \ell_{ij}$ $\forall$ $j \neq i$, for which

$$\text{Prob}(\{v, \ell\}') = W(\{v, \ell\}'|\{v, \ell\}) = \phi_\beta(h_i), \quad \text{(A1)}$$

where $W(\{v, \ell\}'|\{v, \ell\})$ is the transition probability, and $h_i = h_i(\{v, \ell\})$ is the node support (2) for node $i$ in the state $\{v, \ell\}$.

For the reverse transition $\{v, \ell\}' \rightarrow \{v, \ell\}$, we have that $W(\{v, \ell\}|\{v, \ell\}') = \phi_\beta(-h_i)$ with the node support $h_i$ unchanged, as it is *independent* of $v_i$. Hence,

$$\frac{W(\{v, \ell\}'|\{v, \ell\})}{W(\{v, \ell\}|\{v, \ell\}')} = \frac{\phi_\beta(h_i)}{\phi_\beta(-h_i)} = e^{\beta h_i} = e^{-\beta[\mathcal{H}(\{v, \ell\}') - \mathcal{H}(\{v, \ell\})]},$$

$$\text{(A2)}$$

i.e., the ratio of the equilibrium distributions of the states $\{v,\ell\}'$ and $\{v,\ell\}$. This establishes detailed balance for all transitions involving node variables.

Next, we consider a transition $\{v,\ell\} \rightarrow \{v,\ell\}'$ with $\ell_{ij}=0$, $\ell'_{ij}=1$, and $v'_i=v_i \quad \forall \; i, \; \ell'_{kl}=\ell_{kl} \quad \forall \; (kl) \neq (ij)$. Using the symmetry conditions (15)–(17), we obtain the corresponding ratio of transition probabilities,

$$\frac{W(\{v,\ell\}'|\{v,\ell\})}{W(\{v,\ell\}|\{v,\ell\}')} = \frac{\phi_\beta(h_{ij})}{\phi_\beta(-h_{ij})} = e^{\beta h_{ij}} = e^{-\beta[\mathcal{H}(\{v,\ell\}')-\mathcal{H}(\{v,\ell\})]},$$

(A3)

where $h_{ij}$ is the link support (3) which is independent of $\ell_{ij}$. This establishes detailed balance for all transitions involving link variables.

[1] R. May, Nature (London) **238**, 413 (1972).

[2] A. Perelson and G. Weisbuch, Rev. Mod. Phys. **69**, 1219 (1997).

[3] R. Solé and J. Montoya, Proc. R.. Soc. London, Ser. B **268**, 2039 (2001).

[4] C. Gadgil, C. H. Lee, and H. Othmer, Bull. Math. Biol. **67**, 901 (2005).

[5] M. E. J. Newman, Phys. Rev. E **64**, 016132 (2001).

[6] A. Aleksiejuk and J. Holyst, Physica A **299**, 198 (2001).

[7] J. Hatchett and R. Kühn, J. Phys. A **39**, 2231 (2006).

[8] R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).

[9] S. Dorogovtsev and J. Mendes, *Evolution of Networks: From Biological Networks to the Internet and WWW* (Oxford University Press, Oxford, 2003).

[10] R. Albert, H. Jeong, and A.-L. Barabási, Nature (London) **406**, 378 (2000).

[11] R. Cohen, K. Erez, D. ben-Avraham, and S. Havlin, Phys. Rev. Lett. **85**, 4626 (2000).

[12] R. Cohen, K. Erez, D. ben-Avraham, and S. Havlin, Phys. Rev. Lett. **86**, 3682 (2001).

[13] P. Holme, B. J. Kim, C. N. Yoon, and S. K. Han, Phys. Rev. E **65**, 056109 (2002).

[14] R. Kühn and P. Neu, Physica A **322**, 650 (2003).

[15] K. Anand and R. Kühn, Phys. Rev. E **75**, 016111 (2007).

[16] D. Nasiev, J. M. van Mourik, and R. Kühn (unpublished).

[17] L. Viana and A. J. Bray, J. Phys. C **18**, 3037 (1985).

[18] K. Y. M. Wong and D. Sherrington, J. Phys. A **20**, L793 (1987).

[19] M. Leone, A. Vázquez, A. Vespignani, and R. Zecchina, Eur. Phys. J. B **28**, 191 (2002).

[20] B. Wemmenhove and A. C. C. Coolen, J. Phys. A **36**, 9617 (2003).

[21] I. Syozi, in *Phase Transitions and Critical Phenomena*, edited by C. Domb and M. S. Green (Academic Press, London, New York, 1972), Vol. 1, 269–329.

[22] M. Mézard and G. Parisi, Eur. Phys. J. B **20**, 217 (2001).

[23] A. M. L. da Silva, L. C. de Resende, L. A. da Fonseca Manso, and R. Billinton, IEEE Trans. Power Syst. **19**, 1763 (2004).

[24] W. Wangdee and R. Billinton, IEEE Trans. Power Syst. **21**, 188 (2006).