High-Accuracy Approximation of Binary-State Dynamics on Networks

James P. Gleeson
MACSI, Department of Mathematics & Statistics, University of Limerick, Ireland
(Received 17 May 2011; published 4 August 2011)

Binary-state dynamics (such as the susceptible-infected-susceptible (SIS) model of disease spread, or Glauber spin dynamics) on random networks are accurately approximated using master equations. Standard mean-field and pairwise theories are shown to result from seeking approximate solutions of the master equations. Applications to the calculation of SIS epidemic thresholds and critical points of nonequilibrium spin models are also demonstrated.

DOI: 10.1103/PhysRevLett.107.068701 PACS numbers: 89.75.Hc, 05.45.−a, 64.60.aq, 89.75.Fb

Dynamical processes running on complex networks are used to model a wide variety of phenomena [1,2]. Examples include spreading of diseases or opinions through a population [3,4], neural activity in the brain [5], and cascading bank defaults in a financial system [6]. The structure of the underlying network (e.g., its degree distribution) may strongly influence the dynamics and determine critical values of parameters (e.g., the critical temperature of the Ising spin model [7], or the epidemic threshold for disease-spread models [8,9]). Accurate prediction of dynamics and critical points on networks of arbitrary degree distribution thus remains an important unsolved problem [1].

Mean-field theories (MF) are relatively simple to derive and can be quite accurate for dynamics on well-connected networks [10]. However, on sparse networks, or close to critical points, MF theories perform poorly (see, for example, Fig. 1 below). Pairwise approximations (PA), which take into account the states of both nodes at the ends of a network edge, improve on MF, but have been derived for fewer dynamical processes (examples are [11,12]). In this Letter we demonstrate a tractable master equation approach for binary-state dynamics, with accuracy exceeding both MF and PA. We show that PA and MF theories may be derived by seeking approximate solutions of the master equations. We write down the explicit PA equations for the general case, thus giving the first derivation of pairwise approximations for a range of dynamical processes. Finally, we use the master equations to calculate critical points such as the epidemic threshold for the susceptible-infected-susceptible (SIS) model (or contact process), and the critical noise level in the majority-vote model [13].

We consider binary-state dynamics on static, undirected, connected networks in the limit of infinite network size. For convenience, we call the two possible states of a node *susceptible* and *infected*, as is common in disease-spread models. However, this approach also applies to other binary-state dynamics, such as spin models [14], where each node may be in the $+1$ (spin-up $=$ infected) or the -1 (spin-down $=$ susceptible) state. The networks have degree distribution P_k and are generated by the configuration model [2]. Dynamics are stochastic, and are defined by infection and recovery probabilities which depend on the degree k of a node, and on the current number m of infected neighbors of the node. Thus $F_{k,m}dt$ is defined as the probability that a k-degree node that is susceptible at time t, with m infected neighbors, changes its state to infected by time $t + dt$, where dt in an infinitesimally small time interval. Similarly, $R_{k,m}dt$ is the probability that a k-degree infected node with m infected neighbor moves to the susceptible state within a time dt. These general infection and recovery probabilities can describe many dynamical processes of interest; see Table I for some examples.

Approximate master equations for dynamics of this type can be derived by generalizing the approach used in [18] for SIS dynamics; see [19]. Let $s_{k,m}(t)$ [$i_{k,m}(t)$] be the fraction of k-degree nodes that are susceptible (infected) at time t, and have m infected neighbors. Then the fraction $\rho_k(t)$ of k-degree nodes that are infected at time t is given by $\rho_k(t) = \sum_{m=0}^{k} i_{k,m} = 1 - \sum_{m=0}^{k} s_{k,m}$, and the fraction of infected nodes in the whole network is found by summing over all k classes: $\rho(t) = \langle \rho_k(t) \rangle \equiv \sum_k P_k \rho_k(t)$.

The master equations for the evolution of $s_{k,m}(t)$ and $i_{k,m}(t)$ are [19]

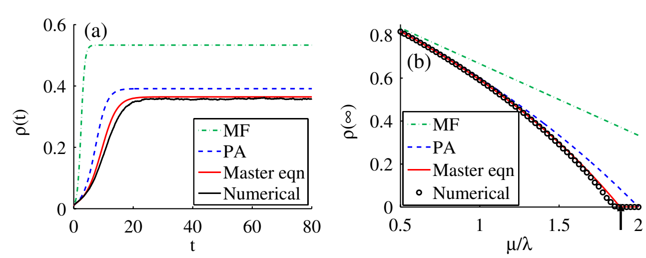

FIG. 1 (color online). (a) Infected fraction $\rho(t)$ in the SIS disease-spread model on 3-regular random graphs, with transmission rate $\lambda = 1$ and recovery rate $\mu = 1.4$. (b) Steady-state fraction of infected nodes as a function of the nondimensional recovery rate μ/λ. The arrow marks the epidemic threshold predicted from the linearized master equations [top row of Table II(a)].

TABLE I. Infection and recovery rates for some examples of binary-state dynamics on networks: k is the node's degree, m is its number of infected neighbors. Parameters λ and μ are SIS transmission and recovery rates; T and J are the temperature and interaction strength for the Ising model; Q is the majority-vote noise parameter. Note $T = 0$ Glauber dynamics are identical to those of the $Q = 0$ majority-vote model.

Process	$F_{k,m}$	$R_{k,m}$
SIS [15]	λm	μ
Voter model [16]	m/k	$1 - F_{k,m}$
Glauber dynamics [17]	$[1 + \exp(\frac{2J}{T}(k - 2m))]^{-1}$	$1 - F_{k,m}$
Majority-vote [13]	$\begin{cases} Q & \text{if } m < k/2 \\ 1/2 & \text{if } m = k/2 \\ 1 - Q & \text{if } m > k/2 \end{cases}$	$1 - F_{k,m}$

$$\frac{d}{dt}s_{k,m} = -F_{k,m}s_{k,m} + R_{k,m}i_{k,m} - \beta^s(k-m)s_{k,m} + \beta^s(k-m+1)s_{k,m-1} - \gamma^s m s_{k,m} + \gamma^s(m+1)s_{k,m+1}, \quad (1)$$

$$\frac{d}{dt}i_{k,m} = -R_{k,m}i_{k,m} + F_{k,m}s_{k,m} - \beta^i(k-m)i_{k,m} + \beta^i(k-m+1)i_{k,m-1} - \gamma^i m i_{k,m} + \gamma^i(m+1)i_{k,m+1}, \quad (2)$$

for each m in the range $0, \ldots, k$, and for each k class in the network. The first two terms on the right-hand side of each equation represent transitions due to infection or recovery of a k-degree node. The remaining four terms account for infection or recovery of a neighbor. The rates β^s, γ^s, β^i, and γ^i are approximated by tracking the number of edges of each type. To calculate β^s, for example, we count the number of S-S edges (i.e., edges between two susceptible nodes) in the network at time t, and then count the number of edges which switch from being S-S edges to S-I edges in the time interval dt; the probability $\beta^s dt$ is given by taking the ratio of the latter to the former, giving $\beta^s = \langle\sum_{m=0}^k(k-m)F_{k,m}s_{k,m}\rangle/\langle\sum_{m=0}^k(k-m)s_{k,m}\rangle$. Similarly, we have $\gamma^s = \langle\sum_{m=0}^k(k-m)R_{k,m}i_{k,m}\rangle/\langle\sum_{m=0}^k(k-m)i_{k,m}\rangle$, $\beta^i = \langle\sum_{m=0}^k mF_{k,m}s_{k,m}\rangle/\langle\sum_{m=0}^k ms_{k,m}\rangle$, and $\gamma^i = \frac{\langle\sum_{m=0}^k mR_{k,m}i_{k,m}\rangle}{\langle\sum_{m=0}^k mi_{k,m}\rangle}$; see [19].

The master Eqs. (1) and (2), with the time-dependent rates β^s, γ^s, β^i and γ^i (defined as nonlinear functions of $s_{k,m}$ and $i_{k,m}$), form a closed system of deterministic equations which can be solved numerically using standard methods. Assuming a randomly-chosen fraction $\rho(0)$ of nodes are initially infected, the initial conditions are $s_{k,m}(0) = (1-\rho(0))B_{k,m}(\rho(0))$, $i_{k,m}(0) = \rho(0)B_{k,m}(\rho(0))$, where $B_{k,m}(q)$ denotes the binomial factor

$$\binom{k}{m}q^m(1-q)^{k-m}.$$

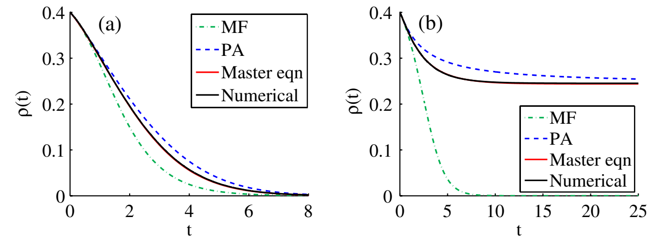

FIG. 2 (color online). Infected fraction $\rho(t)$ (i.e., fraction of $+1$ spins) for zero-temperature Glauber dynamics on (a) networks with truncated power-law degree distribution: $P_k \propto k^{-2.5}$ for $3 \le k \le 20$, and (b) 3-regular random graphs. In each case the initial condition is $\rho(0) = 0.4$.

Note that the evolution equations are completely prescribed by the functions $F_{k,m}$ and $R_{k,m}$, and so this method can be applied to any stochastic dynamical process defined by transition rates of this type. For the SIS model, Eqs. (1) and (2) were derived in [18], but were not analyzed as here.

Figure 1(a) shows the infected fraction $\rho(t)$ of nodes in the SIS model run on a 3-regular random graph (i.e., a Bethe lattice, with $P_k = \delta_{k,3}$). The master Eqs. (1) and (2) clearly give a better approximation to the actual stochastic dynamics than standard methods [here, the mean-field theory of [3] and the pair-approximation method of [11,20]—note these are reproduced by Eqs. (4) and (3) below]. The steady-state infected fraction is plotted as a function of the nondimensional recovery rate μ/λ in Fig. 1(b). The master equation solutions give a significantly better estimate of the epidemic threshold than the standard approximations: we pursue this further below. Figs. 2(a) and 2(b) demonstrate that similar conclusions hold for zero-temperature Glauber dynamics [21] on networks with truncated power-law degree distributions and on 3-regular random graphs. Here the comparison is with the mean-field theory of [22] [see also (4) below], and the pair approximation from Eq. (3) below. Figure 2(b) shows that our approach captures the fact that $T = 0$ Glauber dynamics on networks can freeze in disordered states; this phenomenon is not captured at all by MF [22].

For dynamics on a general network, with nonempty degree classes from $k = 0$ up to a cutoff $k_{\max}$, the number of differential equations in the system (1) and (2) is $(k_{\max} + 1)(k_{\max} + 2)$, and so grows with the square of the largest degree. In certain no-recovery cases (i.e., $R_{k,m} \equiv 0$), such as Watts' threshold model [23], k-core size calculations [24], and bootstrap percolation [25], we can show that an *exact* solution of the master equations is obtained by solving just two differential equations (as given in [26]). For general dynamics, however, some approximation is necessary if it is desirable to reduce the master equations to a lower-dimensional system. One possibility is to consider the parameters $p_k(t)$ [$q_k(t)$], defined as the probability that a randomly-chosen neighbor of a susceptible (infected) k-degree node is infected at time t. Noting that $p_k(t)$ can be expressed in terms of $s_{k,m}$ as $\sum_{m=0}^k m s_{k,m}/\sum_{m=0}^k k s_{k,m}$, an evolution equation for p_k may be derived by multiplying Eq. (1) by m and summing over m. The right-hand side of the resulting equation contains higher moments of $s_{k,m}$, so a closure approximation is needed to proceed. If we make the ansatz that $s_{k,m}$ and $i_{k,m}$ are proportional to binomial distributions: $s_{k,m} \approx (1-\rho_k)B_{k,m}(p_k)$, $i_{k,m} \approx \rho_k B_{k,m}(q_k)$, we obtain the *pair approximation (PA)*, consisting of the $3k_{\max}+1$ differential equations:

$$\frac{d}{dt}\rho_k = -\rho_k \sum_{m=0}^{k} R_{k,m}B_{k,m}(q_k) + (1-\rho_k)\sum_{m=0}^{k} F_{k,m}B_{k,m}(p_k),$$

$$\frac{d}{dt}p_k = \sum_{m=0}^{k}\left[p_k - \frac{m}{k}\right]\left[F_{k,m}B_{k,m}(p_k) - \frac{\rho_k}{1-\rho_k}R_{k,m}B_{k,m}(q_k)\right] + \bar{\beta}^s(1-p_k) - \bar{\gamma}^s p_k,$$

$$\frac{d}{dt}q_k = \sum_{m=0}^{k}\left[q_k - \frac{m}{k}\right]\left[R_{k,m}B_{k,m}(q_k) - \frac{1-\rho_k}{\rho_k}F_{k,m}B_{k,m}(p_k)\right] + \bar{\beta}^i(1-q_k) - \bar{\gamma}^i q_k \quad (3)$$

for each k class. The rates here are given by inserting the binomial ansatz into the general formulas, so that $\bar{\beta}^s$, for example, is $\langle(1-\rho_k)\sum_m (k-m)F_{k,m}B_{k,m}(p_k)\rangle/\langle(1-\rho_k)k(1-p_k)\rangle$; initial conditions are $\rho_k(0) = p_k(0) = q_k(0) = \rho(0)$.

A cruder, *mean-field (MF)*, approximation results from replacing both p_k and q_k with ω: $s_{k,m} \approx (1-\rho_k)B_{k,m}(\omega)$, $i_{k,m} \approx \rho_k B_{k,m}(\omega)$, where $\omega = \langle \frac{k}{z}\rho_k\rangle$ is the probability that one end of a randomly chosen edge is infected. Using this ansatz in the master equations yields a closed system of $k_{\max}+1$ differential equations for the fraction ρ_k of infected k-degree nodes:

$$\frac{d}{dt}\rho_k = -\rho_k \sum_{m=0}^{k} R_{k,m}B_{k,m}(\omega) + (1-\rho_k)\sum_{m=0}^{k} F_{k,m}B_{k,m}(\omega), \quad (4)$$

with $\rho_k(0) = \rho(0)$.

The PA and MF approximations (3) and (4) yield increasingly simpler systems of equations for any process that can be expressed in terms of infection and recovery rates $F_{k,m}$ and $R_{k,m}$. For the SIS model, the PA Eqs. (3) are those of Eames and Keeling [11], while the MF Eqs. (4) are precisely those of Pastor-Satorras and Vespignani [3]. For the voter model [16], the MF Eqs. (4) reduce to those in [27], while the PA Eqs. (3) lie between those of Ref. [12] and Ref. [28] in terms of complexity. The MF Eqs. (4) for zero-temperature Glauber dynamics reproduce the mean-field theory of [22] (in the limit of infinite network size). For this and related nonequilibrium spin models, such as the majority-vote model, steady-state PA equations for the special case of 4-regular graphs (i.e., $P_k = \delta_{k,4}$) are derived in [14]. However, to our knowledge, no PA equations such as (3) have been derived for these dynamics on networks with arbitrary degree distribution P_k. Note also that a coarser type of PA, using the ansatz $s_{k,m} = (1-\rho_k)B_{k,m}(p)$, $i_{k,m} = \rho_k B_{k,m}(q)$ (i.e., with k-independent parameters p and q) gives the equations recently derived in [29] for SIS, and those in [28] for the voter model. This will be examined in future work.

We briefly highlight another important application of the master equations: the calculation of the epidemic threshold for the SIS disease-spread model [8,9]. If the seed fraction of infected nodes $\rho(0)$ is sufficiently small, an appropriate linearization of the master Eqs. (1) and (2) determines whether the infected fraction will grow (to epidemic proportions), or will decay to zero. This reduces the problem to linear stability analysis, and so to the calculation of the largest eigenvalue of a matrix (with dimension of order $k_{\max}^2$). In Table II(a) we show the critical values of the

TABLE II. (a) Critical values of μ/λ for epidemic spread in the SIS model on z-regular graphs. Lower and upper bounds for the critical value of μ/λ for the contact process on a tree (defined as the largest value of μ/λ for which the infection survives forever with positive probability) are from [30]. Note that the largest eigenvalue of the adjacency matrix for these networks is $\lambda_1 = z$, so the method of Prakash *et al.* [31] gives the same (inaccurate) prediction for the critical value as MF theory. (b) Critical value of the noise parameter Q in the majority-vote model on Poisson (Erdös-Rényi) random graphs of mean degree $\langle k\rangle = z$. Numerical values are from [32], other values are determined via stability analysis of Eqs. (1)–(4).

(a) SIS, z regular					(b) Majority-vote, PRG				
z	bounds [30]	Master Eq.	PA [11]	MF [3]	z	num [32]	Master Eq.	PA	MF
3	(1.65, 2)	1.88	2	3	3	0.135	0.137	0.141	0.180
4	(2.56, 3)	2.91	3	4	4	0.181	0.184	0.185	0.214
5	(3.58, 4)	3.93	4	5	6	0.240	0.242	0.242	0.259
10	(8.63, 9)	8.97	9	10	8	0.275	0.277	0.276	0.288

parameter μ/λ for SIS dynamics on z-regular random graphs calculated in this way, and compare with the explicit values predicted by PA [11,20] and MF [3] methods (i.e., $z-1$ and z, respectively). Recently it was argued that SIS infection can persist indefinitely in networks containing nodes of sufficiently high degree, due to recurring reinfections between hub nodes and their neighbors [9]. The master equation formalism does not capture this effect, because the definitions of the rates (β^s, γ^s, etc.) use global counts of edge types, and so wash out structural correlations specific to the immediate neighborhood of hub nodes.

Linear stability analysis may also be applied to spin models with up-down symmetry, which have the property $R_{k,m} = 1 - F_{k,m} = F_{k,k-m}$, and where the magnetization $M(t)$ (the average of all spins in the network) is given by $M = 2\rho - 1$. Stability analysis of the (disordered) fixed point with $\rho = 1/2$ gives the location of critical points marking the transition between disordered and ordered phases. Applying this method to Glauber dynamics reproduces the results of [7] for the critical temperature of the Ising model. It also accurately approximates numerically-determined critical values for nonequilibrium spin models, such as the critical noise Q_c in the majority-vote model; see Table II(b).

In summary, we have derived the master Eqs. (1) and (2)—first introduced for SIS dynamics in [18]—for general binary-state dynamics on networks, and demonstrated that their accuracy supersedes standard MF and PA methods. Mean-field and pairwise theories are derived as approximate solutions of the master equations, and Eqs. (3) explicitly give pair approximations for any dynamics defined by infection and recovery rates $F_{k,m}$ and $R_{k,m}$. Finally, we demonstrated the application of the master equations to calculating epidemic thresholds and critical parameter values via linear stability analysis, improving significantly on existing MF and PA estimates.

This work was funded by Science Foundation Ireland awards 06/IN.1/I366 and MACSI 06/MI/005. Helpful discussions with Sergey Melnik, Rick Durrett, and Claudio Castellano, and participants at the SAMSI Dynamics on Networks workshop are gratefully acknowledged.

[1] A. Barrat, M. Barthélemy, and A. Vespignani, *Dynamical Processes on Complex Networks* (Cambridge University Press, Cambridge, England, 2008).
[2] M. E. J. Newman, *Networks: An Introduction* (Oxford University Press, Oxford, 2010).
[3] R. Pastor-Satorras and A. Vespignani, Phys. Rev. Lett. **86**, 3200 (2001).
[4] C. Castellano, S. Fortunato, and V. Loreto, Rev. Mod. Phys. **81**, 591 (2009).
[5] C. J. Honey *et al.*, Proc. Natl. Acad. Sci. U.S.A. **106**, 2035 (2009); A. V. Goltsev *et al.*, Phys. Rev. E **81**, 061921 (2010).
[6] A. G. Haldane and R. M. May, Nature (London) **469**, 351 (2011); R. M. May and N. Arinaminpathy, J. R. Soc. Interface **7**, 823 (2009).
[7] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, Phys. Rev. E **66**, 016104 (2002); M. Leone *et al.*, Eur. J. Phys. B **28**, 191 (2002).
[8] R. Parshani, S. Carmi, and S. Havlin, Phys. Rev. Lett. **104**, 258701 (2010).
[9] C. Castellano and R. Pastor-Satorras, Phys. Rev. Lett. **105**, 218701 (2010); R. Durrett, Proc. Natl. Acad. Sci. U.S.A. **107**, 4491 (2010).
[10] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, Rev. Mod. Phys. **80**, 1275 (2008).
[11] K. T. D. Eames and M. J. Keeling, Proc. Natl. Acad. Sci. U.S.A. **99**, 13 330 (2002).
[12] E. Pugliese and C. Castellano, Europhys. Lett. **88**, 58 004 (2009).
[13] M. J. de Oliveira, J. Stat. Phys. **66**, 273 (1992).
[14] M. J. de Oliveira, J. F. F. Mendes, and M. A. Santos, J. Phys. A **26**, 2317 (1993).
[15] R. M. Anderson and R. M. May, *Infectious Diseases of Humans: Dynamics and Control* (Oxford University Press, Oxford, 1992); T. E. Harris, Ann. Probab. **2**, 969 (1974).
[16] T. M. Liggett, *Interacting Particle Systems* (Springer, New York, 1985).
[17] R. J. Glauber, J. Math. Phys. (N.Y.) **4**, 294 (1963).
[18] V. Marceau *et al.*, Phys. Rev. E **82**, 036116 (2010).
[19] J. P. Gleeson, arXiv:1104.1537.
[20] S. A. Levin and R. Durrett, Phil. Trans. R. Soc. B **351**, 1615 (1996).
[21] C. Castellano *et al.*, Phys. Rev. E **71**, 066107 (2005).
[22] C. Castellano and R. Pastor-Satorras, J. Stat. Mech. (2006) P05001.
[23] D. J. Watts, Proc. Natl. Acad. Sci. U.S.A. **99**, 5766 (2002); D. Centola, V. M. Eguíluz, and M. W. Macy, Physica (Amsterdam) **374A**, 449 (2007).
[24] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, Phys. Rev. Lett. **96**, 040601 (2006); A. V. Goltsev, S. N. Dorogovtsev, and J. F. F. Mendes, Phys. Rev. E **73**, 056101 (2006).
[25] G. J. Baxter *et al.*, Phys. Rev. E **82**, 011103 (2010).
[26] J. P. Gleeson, Phys. Rev. E **77**, 046117 (2008).
[27] V. Sood and S. Redner, Phys. Rev. Lett. **94**, 178701 (2005).
[28] F. Vazquez and V. M. Eguíluz, New J. Phys. **10**, 063011 (2008).
[29] T. House and M. J. Keeling, J. R. Soc. Interface **8**, 67 (2010).
[30] R. Pemantle, Ann. Probab. **20**, 2089 (1992); T. M. Liggett, Ann. Probab. **24**, 1675 (1996).
[31] B. A. Prakash *et al.*, arXiv:1004.0060.
[32] L. F. C. Pereira and F. G. B. Moreira, Phys. Rev. E **71**, 016123 (2005).