Cascades on clique-based graphs

Adam Hackett[1] and James P. Gleeson[2]

[1]*Hamilton Institute, National University of Ireland, Maynooth, Co. Kildare, Ireland*
[2]*MACSI, Department of Mathematics & Statistics, University of Limerick, Co. Limerick, Ireland*

We present an analytical approach to determining the expected cascade size in a broad range of dynamical models on the class of highly clustered random graphs introduced by Gleeson [J. P. Gleeson, Phys. Rev. E **80**, 036107 (2009)]. A condition for the existence of global cascades is also derived. Applications of this approach include analyses of percolation, and Watts's model. We show how our techniques can be used to study the effects of in-group bias in cascades on social networks.

I. INTRODUCTION

Emergent behaviors in a complex system depend crucially on the pattern of interactions between its components [1–3]. For example, we observe a cascade when local interactions in the vicinity of an initially isolated effect allow that effect to propagate globally [4,5]. The network substrate of a system represents this pattern in its most abstract and analytically tractable form. This information can be used to construct network models, which provide theoretical insights into the causes of such behaviors. A fundamental problem for the construction of these models is the determination of precisely which structural features are requisite to explain the phenomenon in question and which others are superfluous.

In the configuration model [6,7] an ensemble of random graphs is prescribed by a degree distribution p_k. In each realization drawn from this ensemble, a randomly selected vertex will have k incident edges with probability p_k. This distribution represents the first order of complexity for most network models. From this, a more realistic model can be constructed by including degree-degree correlations [8–11] and/or various forms of clustering [12–15], both of which are explicitly absent from the configuration model. Recently, the study of multiplex networks has introduced a further degree of complexity to this general approach [16–18]. These networks consist of connected layers of networks, where each layer involves interactions of a fundamentally unique kind.

In this paper we focus on random graphs with clustering; specifically, those defined by Gleeson in [15]. Real networks typically contain a large number of short cycles in which a small set of vertices maintain a closed loop of connections. One way to measure the propensity for a vertex to form these types of bonds is through the local clustering coefficient, which is defined as the fraction of pairs of neighbors of a vertex that are also neighbors of each other [19]. The degree-dependent clustering coefficient or clustering spectrum c_k is found by averaging the local clustering coefficient over the class of vertices of degree k [20,21]. A global measure of clustering C_2 can be defined by averaging the local coefficients of all N vertices in the graph. Gleeson [15] has shown how the configuration model can be modified to generate ensembles of highly clustered graphs (see also [12–14]). This is achieved by embedding cliques of connected vertices into an otherwise treelike structure. Each ensemble is prescribed by the joint distribution $\gamma(k,c)$: the probability that in any realization a randomly selected vertex has degree k and is in a clique of c vertices (a c-clique).

Our aim is to provide a generalized analytical approach to determining the expected cascade size on these $\gamma(k,c)$ or clique-based graphs. This goes far beyond the bond percolation process studied in [15] to include a broad class of cascade processes including Watts's threshold model [4], k-core decomposition [22,23], and both site and bond percolation [24,25]. Also of relevance is our earlier work [26] on cascades on edge-triangle graphs [13,14]. Edge-triangle graphs are created by embedding 3-cliques, and only 3-cliques, into an otherwise treelike structure. In each such graph a randomly chosen vertex is incident to s single edges and $2t$ triangle edges with probability $p(s,t)$. In contrast, a $\gamma(k,c)$ graph can contain cliques of many different sizes and may therefore have local clustering levels that are much higher than those in edge-triangle graphs. Furthermore, $\gamma(k,c)$ can be parametrized to match the empirical clustering spectrum c_k and degree distribution p_k of a real-world network [15]. This additional complexity means that a very different analytical approach from that of [26] is required here. Our approach thus provides another significant extension of the methods used by Gleeson and Cahalane [27] and Gleeson [5], who provided analytical results for cascades on configuration-model graphs by introducing a tree-based framework of level-by-level vertex activations. This method was inspired by methods originally developed to study the zero-temperature random-field Ising model on a Bethe lattice [28–30].

The class of cascade dynamics examinable through the tree-based framework consists of those processes that satisfy the following list of properties: (i) Each vertex is assigned a binary value specifying its current state, active (damaged or infected) or inactive (undamaged or susceptible); (ii) the probability of a vertex becoming active (in a synchronous update of all vertices) depends only on its degree k and the number m of its neighbors that are already active and is termed the neighborhood influence response function F_m^k [31,32]; (iii) for any fixed degree k, F_m^k is a nondecreasing function of m; and (iv) once active, a vertex cannot become deactivated [33]. Each of the processes referred to in the preceding paragraph satisfies these constraints and is defined by choosing an appropriate F_m^k, as detailed in [5]. The goal of our analytical approach is the prediction of the expected size of the cascade when a time-dependent process of the type

described here has run to completion. Our analytical results are defined as the fixed point of an iterative process, i.e., the solution of a self-consistent system of equations, but the level-by-level activation approach used in our analysis should not be misunderstood as a time-dependent process in its own right; rather it is a convenient representation of the iteration scheme for solving for the steady-state solution.

The remainder of this paper is structured as follows. In Sec. II we describe in broad outline our generalized approach to cascade dynamics on clique-based graphs. As well as an analytical expression for the expected cascade size, we provide a first-order condition for the existence of cascades whose size scales with the number of vertices N as $N \to \infty$. Section III deals in greater detail with the particulars of clique member activations. We show how to calculate in closed form the number of active vertices in a clique of any size $c \leqslant k + 1$. The analysis of both sections is described in terms of an arbitrary response function. The particular forms that this response takes for various processes are discussed in Sec. IV, where we demonstrate the correspondence between our analytical results and numerical simulations of bond percolation and Watts's model. In Sec. V we present a possible extension of Watts's model in which different weights are assigned to active clique neighbors and active nonclique neighbors. This allows us to vary the influence of a vertex's neighbors on its probability of activation between these two subgroups. We suggest that in future analyses this may provide important insights into the role of group structure and peer influence in processes of social contagion, such as opinion formation [35,36].

II. CASCADE ANALYSIS

As was also the case for edge-triangle graphs [26], in order to extend to clique-based graphs the approach of [5] we must first reconcile the presence of clustering with the locally treelike approximation on which that approach is founded. In considering how best to proceed, let us return briefly to [15] and remind ourselves of the structural properties of the $\gamma(k,c)$ ensemble.

In Fig. 1 we have reproduced Fig. 2 of [15]. This figure shows a portion of an arbitrary $\gamma(k,c)$ graph that has been reconfigured into a treelike formation. The essential characteristics of this reconfiguration can be explained most succinctly by looking at the local edge topology of the randomly chosen vertex A. This vertex, positioned on level $n + 1$ of the tree, has degree $k = 6$ and is a member of a 4-clique. Its six incident edges are made up of $c - 1 = 3$ internal edges, which connect A to its neighboring clique members, and $k - c + 1 = 3$ external edges (emphasized). Of these external edges, one connects A to its parent vertex on the next level up, while the remaining $k - c = 2$ connect A to its external children on level n. The clique neighbors are positioned on an unlabeled intermediate level between A and its grandchildren (circled with dashed line) on level n. This categorization and positioning of vertices is representative of how the tree-based framework operates throughout the graph. Note that any vertex may be treated similarly to A regardless of the size of the clique to which it belongs. For any (k,c) pairing such that $k \geqslant c - 1$ (see [15]), $c - 1$ clique neighbors can always be made to reside in the interspace between a vertex and the level below and one
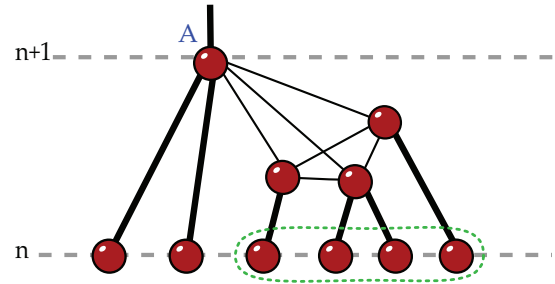

FIG. 1. (Color online) Level-by-level cascade propagation in a $\gamma(k,c)$ graph using the tree approximation. External edges are emphasized.

may also stipulate in general that at most one external edge leads to the parent above. In extreme cases, a vertex with no internal edges is simply a member of a 1-clique and therefore all of its connections will pass directly from one level to the next ($c - 1 = 0$), as in [27]. A vertex with no external edges must reside either at the root of the tree and have no parent if it is part of a clique or it must be entirely isolated and have zero connections in total.

This, then, was the key that allowed Gleeson to calculate the giant connected component (GCC) size S in bond percolation on $\gamma(k,c)$ graphs. Equation (5) of [15] was used to determine the conditional probability that a vertex like A is active (part of the GCC) on each level of the tree and Eq. (6) of [15] then gave S as the probability of activation of the root vertex by using the steady-state value from Eq. (5). The restriction of this theory to bond percolation arises primarily from its reliance on a set of polynomials that were defined and tabulated by Newman in [37]. Crucially, however, those polynomials play no role in the conceptualization described above. Thus our task of extending the theory of [5] amounts to taking this framework and introducing the response function mechanism. Since we shall not apply the polynomials of [37], a straightforward substitution of F_m^k will not suffice. In fact, as we will now show, our approach requires a set of equations entirely different from those of [15].

A. Expected cascade size

With the theoretical foundations in place, we can begin to derive generalized analytical expressions for cascades on $\gamma(k,c)$ graphs. We proceed in the familiar manner by considering the probability q_{n+1} that the randomly selected vertex A in Fig. 1 is active, conditional on its parent vertex being inactive. As is usual for the tree-based approach, we stipulate that the vertex A can become active only due to the influence of the states of the neighboring vertices directly below it in the tree. In this case, however, A has two different types of neighbors: It has $k - c$ external children on level n and $c - 1$ clique neighbors on the intermediate level. Significantly, the ways in which these two types of neighbor can become active in their own right are quite distinct from each other. Thus their contributions to the probability of activation of A must be calculated separately. This is the first problem to be addressed.

Starting with the simpler of the two contributions, let us write down the probability that an arbitrary number, call it

j, of A's external neighbors are active. Since there is no clustering between these vertices, each one is independently activated by its own children on level $n - 1$ with probability q_n. Therefore, the probability that a total of j out of $k - c$ external neighbors are activated in this way is given simply by the binomial probability mass function (PMF)

$$B_j^{k-c}(q_n) = \binom{k - c}{j} q_n{}^j (1 - q_n)^{k-c-j}. \tag{1}$$

For the second contribution to A, matters are made considerably more complicated by the fact that its $c - 1$ clique neighbors are fully connected. This means that the probability that each of these clique neighbors is active depends not only on the states of their children—the four grandchildren of A on level n—but also on the states of one another. Recall from the derivation of our theory for cascades on $p(s,t)$ graphs in [26] that we had to account for the fact that each vertex at the base of a triangle can directly influence the state of the other. We are faced with a similar problem here; however, since we are now dealing with $\gamma(k,c)$ graphs we have a whole spectrum of clique sizes to contend with.

One can appreciate how much more intricate this will make our calculations by imagining that A were part of a very large clique [as it could be, depending on our choice of $\gamma(k,c)$]. For example, if A were in a 10-clique, then $c - 1 = 9$ intermediate vertices would each have a role to play in determining each others' states. The solution in this case would require an extensive list of combinatorial expressions similar to, but extending far beyond, Eqs. (5)–(8) of [26]. Ideally, we would like to avoid tabulating combinatorial terms altogether and instead have a single compact analytical expression that is flexible enough to deal with any clique size. This expression would allow us to feed in the total number of clique neighbors as a variable and would then return the probability that a certain fraction of them are active. Evidently, the derivation of such an expression is not straightforward. We shall therefore postpone this task until later in our presentation.

In the meantime, we continue our analysis of cascade propagation by simply providing the name of this function and taking it for granted that later in Sec. III we will define precisely how it operates. Let us call the relevant function $R_m^{c-1}(q_n)$ and in doing so refer to it as the probability that in a clique of $c - 1$ intermediate vertices a total of m are active, conditional on the top vertex of the c-clique to which they belong (vertex A in Fig. 1) being inactive. The dependence on q_n arises from the fact that each intermediate vertex has its own set of children on level n and each of those children (A's grandchildren in Fig. 1) is active with probability q_n. Summing over all possible values of m gives $\sum_{m=0}^{c-1} R_m^{c-1}(q_n) = 1$.

If we accept the meaning of the label $R_m^{c-1}(q_n)$ and combine it with Eq. (1) above, we now have the necessary terms in which to express the contribution of A's external children and clique neighbors towards its probability of activation q_{n+1}. This takes us very close to defining an iterative equation for q_{n+1} in terms of q_n. The missing ingredient is the probability $\zeta(k,c)$ that the random vertex A, while having degree k and being a member of a c-clique, is also the child of a random vertex on level $n + 2$. This probability plays a role similar to that of the term $(k/z)p_k$ in Eq. (1) of [5], which gives the probability of reaching a

child of degree k by traveling along a randomly chosen edge from its parent in a nonclustered graph (see [38]). Similarly, here $\zeta(k,c)$ closes our iteration by allowing us to average over all vertices on level $n + 1$ in the correct manner. We express this probability as

$$\zeta(k,c) = (k - c + 1)\gamma(k,c)/z_e, \tag{2}$$

where $z_e = \sum_{k,c}(k - c + 1)\gamma(k,c)$ is the average number of external edges per vertex.

Combining all three of our ingredients, we can now write our generalized iterative equation in terms of an arbitrary response function F_{m+j}^k as

$$q_{n+1} = \rho_0 + (1 - \rho_0) \sum_{k,c} \zeta(k,c)\Psi(q_n,k - 1), \tag{3}$$

where

$$\Psi(q_n,x) = \sum_{j=0}^{x-c+1} \sum_{m=0}^{c-1} B_j^{x-c+1}(q_n) R_m^{c-1}(q_n) F_{m+j}^k. \tag{4}$$

Thus we have derived an analytical expression for the probability that a randomly chosen vertex on the next level up, generically called $n + 1$, is active, conditional on its parent being inactive. Referring once again to Fig. 1, Eq. (3) tells us that the vertex A will be found active if it was initially activated as part of the seed fraction ρ_0 or (with probability $1 - \rho_0$) if it subsequently became active in response to the states of the $x = k - 1$ neighbors directly below it in the tree. For the latter, Eq. (4) indicates that there are two distinct contributions from two different sets of neighbors: one from the external children of A and the other from the intermediate clique members. A total of j of the first type of neighbor are active with probability $B_j^{k-c}(q_n)$ and m of the second type with probability $R_m^{c-1}(q_n)$. Whether the sum of j and m is sufficient to activate A is determined by the response function F_{m+j}^k.

In the usual manner, iterating Eq. (3) to the steady state will give us q_∞. This value can then be used in the following expression to determine the probability of activation of the root vertex:

$$\rho = \rho_0 + (1 - \rho_0) \sum_{k,c} \gamma(k,c)\Psi(q_\infty,k). \tag{5}$$

The probability ρ is equivalent to the expected cascade size (see the discussion in [26]). The differences between this equation and Eq. (3) above are attributable to the fact that the root vertex has no parent. This means that all of the root's k edges extend downward to its children, hence $\Psi(q_\infty,k)$. It also means that the correct term for averaging is simply $\gamma(k,c)$.

Taken together, then, Eqs. (3)–(5) constitute the core of our present analytical approach. We can use these equations to investigate various different cascade processes by applying the appropriate definition of the response function F_{m+j}^k in each case. In Sec. IV we will provide the definitions of F_{m+j}^k for bond percolation and Watts's model. Before that, we must also define the function $R_m^{c-1}(q_n)$. This task will occupy all of Sec. III. Next, let us conclude Sec. II by deriving a general first-order cascade condition.

B. Cascade condition

The cascade condition determines whether an infinitesimally small seed fraction ρ_0 of active vertices will generate a nonvanishing mean cascade size as the total number of vertices in the graph diverges ($N \rightarrow \infty$). For this to happen the iteration of Eq. (3) must cause the activation probability q_n to grow from an initial value $q_0 = 0$ to a nonzero steady state q_∞ [5]. If we regard Eq. (3) (with $\rho_0 = 0$) as a nonlinear function of q of the general form $q_{n+1} = H(q_n)$, then this last condition can be expressed, to first order, as $H'(0) > 1$.

To evaluate $H'(0)$ we require the following results for the binomial PMF of Eq. (1):

$$B_j^{k-c}(0) = \delta_{j,0}, \qquad (6)$$

$$\left. \frac{d}{dq} B_j^{k-c}(q) \right|_{q=0} = (k-c)(\delta_{j,1} - \delta_{j,0}). \qquad (7)$$

Using Eqs. (6) and (7) in Eq. (3), we find that the first derivative of $H(q)$, evaluated at $q = 0$, may be expressed as

$$H'(0) = \sum_{k,c} \zeta(k,c) \sum_{m=0}^{c-1} \left[(k-c)\left(F_{m+1}^k - F_m^k\right) \right.$$

$$\left. \times R_m^{c-1}(0) + F_m^k \left. \frac{d}{dq} R_m^{c-1}(q) \right|_{q=0} \right]. \qquad (8)$$

This is the left-hand side of our cascade condition. Note that because this expression depends on $R_m^{c-1}(0)$ and the first derivative of $R_m^{c-1}(q)$ at $q = 0$ it becomes an increasingly arduous task to calculate $H'(0)$ from Eq. (8) as the size of the largest clique in our graph increases. As we shall see in the next section, the evaluation of the function $R_m^{c-1}(q)$ becomes increasingly difficult as the value of c increases. For this reason, in our analysis in Sec. IV we will choose $\gamma(k,c)$ such that the cliques in our graphs are constrained to sizes of $c \leqslant 4$. In addition, we shall make the simplifying assumption that $F_0^k = 0$ (see [26]). This implies that a vertex will never activate if none of its neighbors are active and is a suitable approximation for the calculation of our first-order condition.

III. ACTIVE CLIQUE NEIGHBORS

Backtracking slightly in the flow of our presentation, we will now derive a concise closed-form expression for the probability labeled above as $R_m^{c-1}(q_n)$. Let us begin by recapitulating the meaning of this label. According to our earlier definition, it is the probability that m out of $c - 1$ intermediate-level c-clique vertices are active, given that their own externally linked children are each independently active with probability q_n and that the parent vertex at the top of the c-clique is inactive. In Fig. 1, for example, $R_m^3(q_n)$ is the probability that m of the vertex A's three clique neighbors are active, given that each of the four grandchildren of A (circled) has an activation probability of q_n and that A is itself inactive.

In considering how to calculate $R_m^{c-1}(q_n)$ in general, we see immediately that it is not the states of the external grandchildren that will cause us difficulty, but rather the fact that the state of each intermediate clique member can influence the states of all other members. In our framework, every c-clique has one of its (internally linked) members designated

as the parent and placed on level $n + 1$. This leaves each of the remaining $c - 1$ clique members on the intermediate level with $k - c + 1$ external edges to connect to its own children on level n. The probability that some number j of these children are active is given by the binomial PMF $B_j^{k-c+1}(q_n)$. Thus the probability that an intermediate clique member is activated by its children is quite easy to calculate. In contrast, in order to deal with the influence of the $c - 1$ clique members on one another, we will have to consider carefully the various combinations of states that may exist within the intermediate portion of the clique.

Our first step in tackling this problem is to provide a mechanism for the intermediate clique members to be activated, which combines both internal and external influences. We define

$$G_d^{c-2}(q_n) = \sum_k \frac{\gamma(k,c)}{p_c} \sum_{j=0}^{k-c+1} B_j^{k-c+1}(q_n) F_{d+j}^k, \qquad (9)$$

for $c \geqslant 2$, as the conditional probability that an intermediate c-clique vertex will be activated if d of its $c - 2$ clique neighbors on the same level are active, given its external children are each active with probability q_n and its parent on level $n + 1$ is inactive. The term $\gamma(k,c)/p_c$ is the degree distribution of vertices that belong to a c-clique, where $p_c = \sum_k \gamma(k,c)$. The response function F_{d+j}^k will determine whether d active neighbors plus j active children are enough to cause activation. Defined as such, $G_d^{c-2}(q_n)$ provides a fundamental term in which to express the various possible active configurations, thus permitting us to begin the procedure of counting.

We consider first the simplest nontrivial case, namely, $c = 3$. Suppose we pick from some arbitrary $\gamma(k,c)$ graph a vertex with degree k that is also a member of a 3-clique. If we let this vertex reside on level $n + 1$ of the tree and also position its $c - 1 = 2$ clique neighbors between level $n + 1$ and level n below, our task then is to calculate $R_m^2(q_n)$. To do this, let us refer to Fig. 2 and look at the possible states of these two vertices in isolation from their inactive parent.

Starting with both vertices inactive—the configuration labeled c_0 in Fig. 2—we first count the possible configurations of states after one round ($i = 1$) of synchronous updates. Since we have started from c_0, with both vertices inactive, the probability of either vertex becoming active in this first round is simply G_0^1. Therefore, each possible outcome (c_1, c_2, or c_3 in Fig. 2) is determined by a binomial PMF with probability of success G_0^1. Configuration c_1, in which both vertices have remained inactive, will occur with probability $B_0^2(G_0^1)$. Similarly, configuration c_2, in which one vertex has been activated and the other has remained inactive, will occur with probability $B_1^2(G_0^1)$. Finally, configuration c_3, in which both vertices have been activated, will occur with probability $B_2^2(G_0^1)$. [Note that in each term $G_d^{c-2} \equiv G_d^{c-2}(q_n)$; we will use this abbreviation throughout.]

Having determined the three distinct outcomes of the first round of updates, we will now categorize each configuration into either of two types: terminal or volatile. In a terminal configuration no further changes of state are possible because all vertices have reached their own steady state of either permanent activation or inactivation. In a volatile configuration, however, there exists at least one inactive vertex that is liable
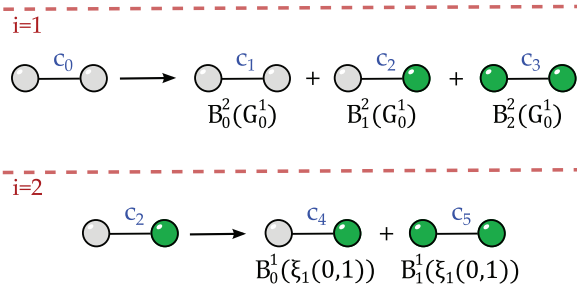
FIG. 2. (Color online) Transition probabilities for a pair ($c - 1 = 2$) of intermediate clique neighbors in a $\gamma(k,c)$ graph. Color indicates vertex state: light gray, inactive; dark gray (green), active.

to become active. Thus, as long as volatile configurations are produced we must continue with another round of updates. The process of updating will reach its end when all configurations are terminal. Categorizing the outcomes of round one tells us whether or not a second round is necessary and also indicates which configurations need to be updated. Configuration c_1 is clearly terminal since the transition from c_0 to c_1 has established that neither vertex can activate while the other remains inactive. Similarly, c_3 is also terminal for the simple reason that we do not allow active vertices to revert to being inactive. Configuration c_2, however, is volatile since the transition from c_0 to c_2 has shown us that one of these vertices can activate without the other first being active, but that the same is not true of this other vertex. That is to say, we know that the inactive vertex in c_2 cannot activate without an active neighbor. What is not clear from c_2 is whether the vertex that did activate in round one is now sufficient to activate the vertex that remained inactive in that round. The only way to determine this is to run a second round ($i = 2$) of updates on c_2.

As was the case in the first round, to begin the second round we must provide an appropriate probability of activation. We want to know if the active vertex in c_2 is enough to activate the inactive vertex in c_2, given that the inactive vertex cannot activate without an active neighbor. This can be decided upon by using the activation probability $\xi_1(0,1)$ defined by the function

$$\xi_{c-2}(a,b) = \frac{G_b^{c-2}(q_n) - G_a^{c-2}(q_n)}{1 - G_a^{c-2}(q_n)}. \tag{10}$$

Equation (10) gives us the conditional probability that in a clique of $c - 1$ intermediate vertices b active vertices are enough to cause the activation of one of their inactive clique neighbors, given that a active vertices are insufficient to do so. The function $\xi_{c-2}(a,b)$ is defined for $0 \leqslant a \leqslant b \leqslant c - 2$ and is non-negative for all such values since by Eq. (9) $G_d^{c-2}(q_n)$ is an increasing function of d. This latter property is true of $G_d^{c-2}(q_n)$ since F_m^k is defined (see Sec. I) to be a nondecreasing function of m [and therefore so is F_{d+j}^k in Eq. (9)]. The configurations produced by updating with this probability are once again given by a binomial PMF. With probability $B_0^1(\xi_1(0,1))$ the inactive vertex will remain inactive, thereby producing configuration c_4. Conversely, with probability $B_1^1(\xi_1(0,1))$ the inactive vertex will activate, thereby producing configuration c_5. Categorizing c_4 and c_5, we find both configurations are terminal and therefore the process of updating may now cease.

With all terminal configurations now achieved, the next step in our derivation of $R_m^2(q_n)$ is to combine the various transition probabilities listed in Fig. 2 and use them to calculate each of $R_0^2(q_n)$, $R_1^2(q_n)$, and $R_2^2(q_n)$. Tracing our way through Fig. 2, we reach a terminal state in which no vertices are active by following the route $c_0 \to c_1$. Similarly, we end with one active vertex by following $c_0 \to c_2 \to c_4$. Finally, a terminal state with two active vertices is given by either of the routes $c_0 \to c_3$ or $c_0 \to c_2 \to c_5$. All of this information can be expressed succinctly using the various transition probabilities associated with each route if we bear in mind that a transition from one configuration to another, symbolized by $\to$, corresponds to the multiplication of probabilities and also that the word *or* corresponds to addition. To summarize, the set of routes described here yields the following set of equations:

$$R_0^2(q_n) = B_0^2(G_0^1), \tag{11}$$

$$R_1^2(q_n) = B_1^2(G_0^1)B_0^1(\xi_1(0,1)), \tag{12}$$

$$R_2^2(q_n) = B_1^2(G_0^1)B_1^1(\xi_1(0,1)) + B_2^2(G_0^1). \tag{13}$$

The final step towards our goal of writing a closed-form expression for $R_m^2(q_n)$ is to find a way of expressing Eqs. (11)–(13) as the outputs of a single function that has been given the inputs $m = 0$, 1, and 2, respectively. There may be a number of different ways of defining such a function, some of which may appear more elegant than others. For our own part, we can offer a particularly concise definition by introducing a new variable and considering how the various combinations of states determined by Eqs. (11)–(13) can be reproduced in a parsimonious manner.

Our new variable is called l_i. We define it as the number of new activations in round i of synchronous updates. In the scheme presented above we had two rounds; therefore, we define the pair $l = (l_1, l_2)$ as the sequence of new activations over both rounds. This allows us to represent all possible routes through the configurations of Fig. 2 as a collection of ordered pairs. For example, $l = (1,0)$ means that there is one activation in round $i = 1$ and no activations in round $i = 2$ and therefore corresponds to the route $c_0 \to c_2 \to c_4$. Similarly, $l = (1,1)$ corresponds to $c_0 \to c_2 \to c_5$. By applying this notation we find that the following equation will reproduce each of the Eqs. (11)–(13) above:

$$R_m^2(q_n) = \sum_{l_1+l_2=m} B_{l_1}^2(G_0^1)B_{l_2}^{2-l_1}(\xi_1(0,l_1)). \tag{14}$$

Note that the summation $\sum_{l_1+l_2=m}$ in Eq. (14) is taken over all pairs $l = (l_1, l_2)$ such that $l_1 + l_2 = m$, where m is the total number of active vertices.

To demonstrate how Eq. (14) operates let us calculate $R_1^2(q_n)$ by setting $m = 1$. The set of all l pairs that add up to this value of m is $l \in \{(0,1),(1,0)\}$. Substituting each of these pairs in turn into the right-hand side of Eq. (14) and then summing gives $R_1^2(q_n) = [0 + 2G_0^1(1 - G_1^1)]$, thereby reproducing Eq. (12) above. The values of $R_0^2(q_n)$ and $R_2^2(q_n)$ are found similarly by using the parameters $m = 0$ and $l = (0,0)$, and $m = 2$ and $l \in \{(0,2),(1,1),(2,0)\}$, respectively.

Thus, in Eq. (14) we have found an expression for $R_m^2(q_n)$, which, we remind ourselves once more, is the conditional

probability that m of the two intermediate vertices in a 3-clique are active, given that each of their own children is active with probability q_n, and that the vertex at the top of the clique is inactive. Recall, however, that our ultimate goal is to provide a general expression for $R_m^{c-1}(q_n)$. Our approach to this problem has been to determine a series of expressions for increasing values of c and then to express each of these as a special case of a single unifying expression. Each individual expression for $R_m^{c-1}(q_n)$, where $c > 3$, can be found by a method similar to the one described above for $R_m^2(q_n)$. The core of this method is the same regardless of the value of c and can be summarized in general as follows.

(i) Simultaneously update the states of all inactive vertices.

(ii) Categorize the resulting configurations of states as either terminal or volatile, removing those that are terminal from further consideration.

(iii) Repeat steps (i) and (ii) until no volatile configurations remain.

Counting the terminal configurations will then provide the various outcomes obtainable in the steady state of the cascade. For example, in determining $R_m^3(q_n)$, the application of these three steps reveals every possible active configuration in a triangle of connected vertices and each associated transition probability. As above, following the different routes towards each terminal configuration indicates the correct sequence of multiplications and additions to employ to calculate the values of $R_m^3(q_n)$ for $0 \leqslant m \leqslant 3$. This procedure yields the following set of equations:

$$R_0^3(q_n) = B_0^3(G_0^2),$$ (15)

$$R_1^3(q_n) = B_1^3(G_0^2)B_0^2(\xi_2(0,1)),$$ (16)

$$R_2^3(q_n) = B_2^3(G_0^2)B_1^2(\xi_2(0,1))B_0^1(\xi_2(1,2))$$
$$+ B_2^3(G_0^2)B_0^1(\xi_2(0,2)),$$ (17)

$$R_3^3(q_n) = B_1^3(G_0^2)B_1^2(\xi_2(0,1))B_1^1(\xi_2(1,2))$$
$$+ B_3^3(G_0^2) + B_2^3(G_0^2)B_1^1(\xi_2(0,2))$$
$$+ B_1^3(G_0^2)B_2^2(\xi_2(0,1)).$$ (18)

Continuing in the same manner as before, an expression for $R_m^3(q_n)$ that contains Eqs. (15)–(18) as special cases can be defined by applying the variable l_i and considering each unique sequence of activations $l = (l_1, l_2, l_3)$. By doing this we have found that the equation

$$R_m^3(q_n) = \sum_{l_1+l_2+l_3=m} B_{l_1}^3(G_0^2)B_{l_2}^{3-l_1}(\xi_2(0,l_1))$$
$$\times B_{l_3}^{3-(l_1+l_2)}(\xi_2(l_1,l_1+l_2))$$ (19)

will reproduce Eqs. (15)–(18).

Observe the similarities between Eqs. (19) and (14). They indicate that to create an expression for $R_m^3(q_n)$ from that for $R_m^2(q_n)$ above all one must do (besides set $c = 4$) is place additional indices l_2 and l_3 in the appropriate positions and include one more multiplicative term, namely, $B_{l_3}^{3-(l_1+l_2)}(\xi_2(l_1,l_1+l_2))$. By running the entire scheme of categorization and route counting over again with $c = 5$ and $l = (l_1, l_2, l_3, l_4)$, we have observed (in calculations not provided here) that a similar relationship also holds between

$R_m^3(q_n)$ and $R_m^4(q_n)$. The pattern of similarities detected in our calculations strongly suggests the following form for a general expression for $R_m^v(q_n)$, where v is an integer $v \geqslant m$:

$$R_m^v(q_n) = \sum_{|l|=m} \prod_{i=1}^v B_{l_i}^{n_{v,i}}(\theta_{v,i}).$$ (20)

Let us unpack this expression. First, note that the variable $n_{v,i}$ in Eq. (20) is defined as $n_{v,i} = v - \sum_{j=1}^{i-1} l_j$ for $i \geqslant 2$ with $n_{v,1} = v$. Next, the variable $\theta_{v,i}$ is defined as $\theta_{v,i} = \xi_{v-1}(\sum_{j=1}^{i-2} l_j, \sum_{j=1}^{i-1} l_j)$ for $i \geqslant 3$ with $\theta_{v,1} = G_0^{v-1}$ and $\theta_{v,2} = \xi_{v-1}(0,l_1)$. Finally, the term $|l|$ in the summation of Eq. (20) is defined in multi-index notation (see, for example, [39]) as $|l| = l_1 + \cdots + l_v$.

By setting $v = c - 1$ in Eq. (20), we have the probability $R_m^{c-1}(q_n)$ expressed in closed form [40]. Applying this definition in Eqs. (3)–(5) (see Sec. II) completes our analytical description of cascades on clique-based graphs and permits us to proceed with the task of verifying our approach. We will provide this verification in the next section by comparing predicted values of the expected cascade size from Eq. (5) against the results of numerical simulations of bond percolation and Watts's model.

It must be noted, however, that as the size of the largest clique of in our graph $c_{\max}$ increases it becomes more and more computationally intensive to evaluate $R_m^{c-1}(q_n)$ using Eq. (20). This is primarily because of the exponentially increasing number of possible combinations for the multi-index l as the number of active clique members to be counted m increases. It can be shown that the number of different choices of l that give nonzero contributions to the sum in Eq. (20) is 2^{m-1}.

IV. SIMULATIONS

To test the theory of the previous two sections we require an appropriate set of definitions for the response function F_{m+j}^k, corresponding to the processes in our familiar broad class (see Sec. I). The function F, however, is the same one that has been used throughout our groups' previous publications [5,15,26]. Gleeson began in [5] by writing it in its simplest generalized form: F_m^k. There it defined the probability that a k-degree vertex in a locally treelike graph may be activated by m active neighbors. In [26], F_m^{s+2t} gave the probability that a k-degree vertex in an edge-triangle graph may be activated by m active neighbors, where $k = s + 2t$. In the current presentation, F_{m+j}^k prescribes the probability that a k-degree vertex in a clique-based graph may be activated by $m + j$ active neighbors, where j and m are the numbers of external and internal neighbors, respectively. Since F has not changed (only its arguments have), the same justifications of our use of the response function mechanism as were given in [26] apply equally here. Therefore, similarly to [26], the definitions of F_{m+j}^k for different processes are found by replacing m with $m + j$ in the definitions of F_m^k given in [5]. With this aspect clarified, we can begin testing our approach against numerical simulations of various processes.
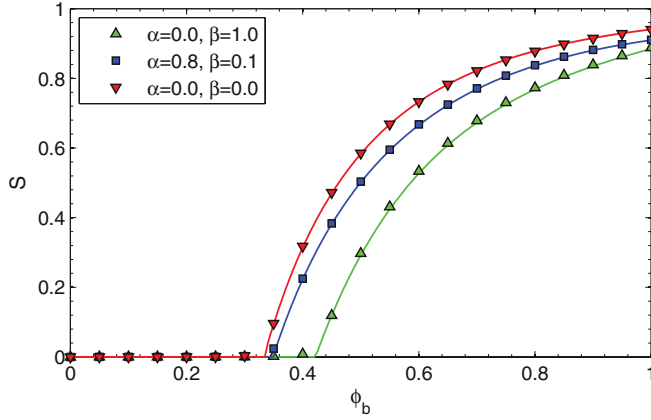
FIG. 3. (Color online) Bond percolation on $\gamma(k,c)$ graphs of $N = 10^5$ vertices and Poisson degree distribution p_k with mean degree $z = 3$. Numerical simulations (symbols) averaged over 100 realizations and theory of Sec. II (lines) on a plot of GCC size S vs. bond occupation probability ϕ_b.

A. Bond percolation

We consider first the process of uniform bond percolation. In this process each edge of the graph (external or internal) is deleted with probability $1 - \phi_b$. The quantity ϕ_b is the bond occupation probability and nondamaged edges are termed occupied. Replacing m with $m + j$ in Eq. (6) of [5] defines F_{m+j}^k for this process:

$$F_{m+j}^k = 1 - (1 - \phi_b)^{m+j}. \tag{21}$$

Applying this definition in the respective $\rho_0 \to 0$ limits of Eqs. (3)–(5) above allows us to use these equations to calculate the expected GCC size S of a clique-based graph, which is nonzero for $\phi_b > \widehat{\phi}_b$. This critical value, $\widehat{\phi}_b$, is known as the bond percolation threshold.

In Fig. 3 we have plotted our calculations of S from Eq. (5) against the results of numerically simulated $\gamma(k,c)$ graphs (see the caption). The parameters chosen for this figure are the same as those used in Fig. 3(a) of [15]. Each graph has a Poisson degree distribution $p_k = z^k e^{-z}/k!$ with mean degree $z = 3$. Following [15], we set $\gamma(k,c) = [(1 - \alpha - \beta)\delta_{c,1} + \alpha\delta_{c,3} + \beta\delta_{c,4}]p_k$ for $k \geqslant 3$, where $\alpha, \beta \in [0,1]$. In this way we create nonzero clustering by assigning a fraction α of k-degree vertices to 3-cliques and a fraction β to 4-cliques. Additionally, since a 2-degree vertex cannot belong to a clique of size $c > 3$, we assign a fraction α of these vertices to 3-cliques using $\gamma(2,c) = [(1 - \alpha)\delta_{c,1} + \alpha\delta_{c,3}]p_2$. We let vertices of degree zero or one belong to 1-cliques: $\gamma(k,c) = p_k\delta_{c,1}$. This choice of $\gamma(k,c)$ limits the largest clique size to $c_{\max} = 4$ and therefore makes the evaluation of $R_m^{c-1}(q_n)$ relatively simple. By varying α and β different levels of clustering can be prescribed. Again following [15], we use three (α, β) pairs: $(0,0)$, $(0.8,0.1)$, and $(0,1)$. Evidently, $(0,0)$ produces a nonclustered graph (downward-pointing triangles). We can use Eq. (2) of [15] to define the global clustering coefficient $C_2 = \sum_k p_k c_k$. From this one may show that $(0.8,0.1)$ produces a clustered graph with $C_2 = 0.31$ (squares) and also that $(0,1)$ gives a graph with $C_2 = 0.35$ (upward-pointing triangles).

The percolation thresholds for each nonzero value of C_2 can be calculated from our cascade condition of Sec. II by

setting $H'(0) = 1$ in Eq. (8) and solving for ϕ_b (see [26]). This of course requires that we first substitute Eq. (21) into Eq. (8). We also require the following results for the function $G_d^{c-2}(q)$ of Eq. (9) in order to evaluate $R_m^{c-1}(0)$ and the first derivative of $R_m^{c-1}(q)$ at $q = 0$:

$$G_d^{c-2}(0) = \sum_k \frac{\gamma(k,c)}{p_c} F_d^k, \tag{22}$$

$$\frac{d}{dq} G_d^{c-2}(q)\bigg|_{q=0} = \sum_k \frac{\gamma(k,c)}{p_c}(k - c + 1)\big(F_{d+1}^k - F_d^k\big). \tag{23}$$

Using Eqs. (21)–(23) in Eq. (8) we calculate the threshold for $C_2 = 0.31$ to be $\widehat{\phi}_b = 0.349$, while for $C_2 = 0.35$ we get $\widehat{\phi}_b = 0.423$. The threshold for $C_2 = 0$ is simply the configuration model value $\widehat{\phi}_b = 1/z$ [41].

The match obtained between theory and numerics in Fig. 3 provides a clear validation of our approach in the case of bond percolation. Furthermore, because we have chosen the same parameters as Fig. 3(a) of [15], the results shown in that figure should correspond exactly with the results shown here in Fig. 3. Comparing these two figures will reveal to the reader that they do indeed match. This illustrates that our approach contains within its scope the ability to produce the same predicted values of S as the theory of [15]. However, as noted earlier at the beginning of Sec. II, Gleeson's equations depend on a set of polynomial functions defined and tabulated in [37]. These polynomials limit the application of his equations to bond percolation. The advantage of our approach over that of [15] is its purported applicability to other processes besides bond percolation. To confirm that it really does possess this flexibility we consider for our second test Watts's model [4].

B. Watts's model

Watts's model provides a simplified description of threshold-dependent cascade dynamics on complex networks. In a sociological setting this model may provide a crude approximation of the processes of contagion that underlie such phenomena as fashions, rumors, or popular opinions. Given, for example, a network of acquaintanceships between a group of people, we can use Watts's model to calculate the steady-state fraction of active vertices in the following binary-state decision process.

We begin by assigning a threshold r_i drawn from the probability distribution $q(r)$ to each vertex $1 \leqslant i \leqslant N$ in the network. At each discrete time step t the state of vertex i is $v_i(t) \in [0,1]$, where $v_i(t) = 1$ indicates the participation of i in the cascade and $v_i(t) = 0$ indicates nonparticipation. The dynamics is instigated by activating a small seed fraction of vertices at $t = 0$. From $t = 1$ until the steady state $\bar{t}$ the state of each vertex is updated synchronously at each t according to the rule

$$v_i(t) = \begin{cases} 1 & \text{if } \frac{1}{k_i}\sum_j a_{ij}v_j(t) > r_i \\ \text{unchanged} & \text{otherwise}, \end{cases} \tag{24}$$

where a_{ij} is the value in position (i,j) of the adjacency matrix of the network and k_i is the degree of vertex i. By this mechanism vertex i will join the cascade if the fraction of its direct neighbors that are active exceeds its threshold, otherwise it will remain inactive. Once active, i will remain in this state.
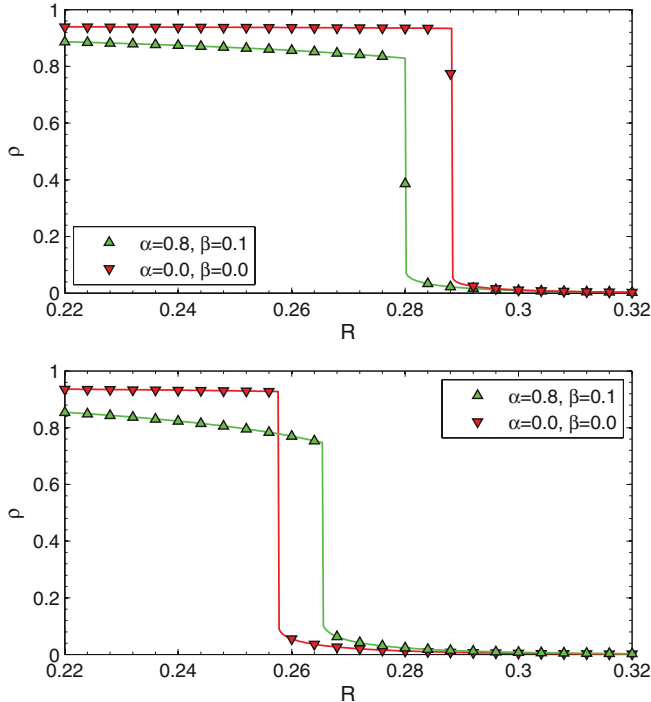
FIG. 4. (Color online) Watts's model on $\gamma(k,c)$ graphs of $N = 10^6$ vertices and Poisson degree distribution p_k with mean degree $z = 3$. Thresholds are drawn from a Gaussian distribution with mean R and standard deviation $\sigma = 0.1$. Numerical simulations (symbols) averaged over 100 realizations and theory of Sec. II (lines) on a plot of cascade size ρ vs. R. In (a) $w_i = 1$ and $w_e = 1$. In (b) $w_i = 1.3$ and $w_e = 0.85$.

In the steady state the final fraction of active vertices is given by $\frac{1}{N} \sum_i v_i(\bar{t})$. By averaging this last value over many individual runs of the model we can determine a numerical evaluation of the expected cascade size ρ.

With the appropriate choice of response function F_{m+j}^k, our Eqs. (3)–(5) provide an analytical match to the numerical results of Watts's model. In Fig. 4 we present values of ρ from Eq. (5) plotted against the results of simulations on $\gamma(k,c)$ graphs. The thresholds in each of these graphs are drawn from a Gaussian distribution: $q(r) = N(R,0.1)$ (see the caption). Therefore, the response function for our equations is defined by replacing m with $m + j$ in Eq. (2) of [5]:

$$F_{m+j}^k = \frac{1}{2}\left[1 + \text{erf}\left(\frac{(m + j)/k - R}{\sigma\sqrt{2}}\right)\right]. \qquad (25)$$

The choice of $q(r) = N(R,0.1)$ means some vertices will be assigned negative thresholds and will therefore automatically activate. This allows us to set $\rho_0 = 0$ in Eqs. (3)–(5). The structural variables used for this figure are the same as those applied previously in Fig. 3. All graphs have Poisson degree distribution p_k with $z = 3$ and $\gamma(k,c)$ is defined by the same three equations as above. We apply two (α,β) pairs, $(0,0)$ and $(0.8,0.1)$, corresponding to $C_2 = 0$ and 0.31, respectively.

Figure 4(a) provides a further validation of our approach and explicitly demonstrates its flexibility. In Fig. 4(b) we investigate a minor modification to Watts's model. The presence of neighbors of two distinct kinds (internal and

external) in clique-based graphs opens up some interesting possibilities for the augmentation of the updating process described by Eq. (24). To take one simple example, consider the following weighting scheme. Let each active internal vertex have weight $w_i \in (0,\infty)$ and each active external vertex have weight $w_e \in (0,\infty)$. The response function for this process is given by multiplying m by w_i and j by w_e in Eq. (25). We propose that such a weighting may provide insights into the role of group structure in determining the outcome of processes of social contagion such as those mentioned above. Problems of this nature have been of interest for quite some time (see [36] and references therein).

When $w_i = w_e = 1$ we have the conventional version of Watts's model in which there is no bias in favor of either type of neighbor [Fig. 4(a)]. However, settings where $w_i > w_e$ or $w_i < w_e$ indicate a respective bias in favor of or against one's clique neighbors over one's external neighbors. If we take the clique as a proxy for a tightly knit social group, then the first setting describes a scenario where the influence of ones peers is favored over influences from outside the immediate peer group. The second setting describes the opposite scenario.

Figure 4(b) demonstrates why this modification of Watts's model is interesting from an analytical perspective. Here we have set $w_i = 1.3$ and $w_e = 0.85$. Comparing this figure to Fig. 4(a), we see that this simple change in weighting can cause a significant change in the expected cascade size ρ. In Fig. 4(a) the nonclustered graph produces a larger value of ρ than the clustered graph at every value of the threshold distribution mean R. However, in Fig. 4(b) this trend is reversed in the region from approximately $R = 0.26$ upward. Based on this observation, we submit that weighted models such as the one provided here may offer new insights into the effects of clustering and decision bias in cascades on social networks [42]. We leave the analysis and modification of this weighted model open to further investigation.

V. CONCLUSION

We have extended Gleeson and Cahalane's [27] analytical approach to modeling cascading phenomena on configuration model graphs to the highly clustered clique-based graphs defined by Gleeson in [15]. An analytical expression for the expected cascade size and a first-order cascade condition have been derived. The use of the generalized response function mechanism in these expressions permits their application to a range of processes that includes site and bond percolation, k-core decomposition, and Watts's threshold model.

We have validated our approach against numerical simulations of bond percolation and Watts's model. In addition, we have proposed a modification of Watts's model that employs the unique structure of clique-based graphs in an investigation of the role of group influence in processes of social contagion. This presents rich ground for further investigation. The analytical framework provided by us here may be useful for such studies.

Perhaps the most significant aspect of our contribution is the derivation of a closed-form expression for the steady-state fraction of active vertices inside a clique of arbitrary size. We anticipate that this expression will find additional applications outside the current setting.

Finally, there are a number of significant challenges that we have yet to address in our broad study of cascades on clustered graphs. We have now provided approaches for a class of monotone binary-state processes on both edge-triangle graphs [26] and clique-based graphs; there are two directions in which we would like to extend this work. First, we would like to modify our techniques to investigate nonmonotone processes. The groundwork for this has been laid in [34]. Second, we would like to investigate cascades on a more sophisticated class of highly clustered graphs than those dealt with so far. Such classes have been described in [44].

APPENDIX: CLUSTERS IN DAMAGED GRAPHS

The results illustrated in Fig. 3 demonstrate the equivalence of the approach to bond percolation provided in [15] and the corresponding approach provided here. By working through the equations of [15] and those of this paper, one may show that the match between the two approaches hinges on the following equation:

$$\sum_{m=1}^{c} P(m|c)x^{m-1} = \sum_{m=0}^{c-1} R_m^{c-1} q^m, \qquad (A1)$$

where $x = 1 - G_0^{c-2}$ and $q = 1 - \phi_b$. On the left-hand side of Eq. (A1) $P(m|c)$ is the probability that in a c-clique that has been damaged by the removal of its edges (each with independent probability q) a connected cluster of m vertices (not necessarily an m-clique) remains.

In [37] the probability $P(m|c)$ was evaluated iteratively using a recursive formula; an explicit formula for $P(m|c)$ was not provided. By making use of Eq. (A1) we can now write an explicit formula for $P(m|c)$.

Applying Eq. (20) allows us to expand the right-hand side of Eq. (A1) and thereby rewrite it as

$$\sum_{m=1}^{c} P(m|c)x^{m-1} = \sum_{|l| \leqslant c-1} \binom{c-1}{l_1} q^{l_1} \prod_{i=2}^{c-1} B_{l_i}^{n_{c-1,i}}(\theta_{c-1,i}) q^{l_i}$$
$$\times \sum_{j=0}^{l_1} \binom{l_1}{j} (-1)^j x^{c-1-l_1+j}. \qquad (A2)$$

To equate coefficients of powers of of x on the left-hand side and right-hand side of Eq. (A2) we simply set $j = m - c + l_1$. This gives us the following expression for the probability $P(m|c)$:

$$P(m|c) = \sum_{|l| \leqslant c-1} \binom{c-1}{l_1} q^{l_1} \prod_{i=2}^{c-1} B_{l_i}^{n_{c-1,i}}(\theta_{c-1,i}) q^{l_i}$$
$$\times \binom{l_1}{m-c+l_1} (-1)^{m-c+l_1}. \qquad (A3)$$

One may easily verify that Eq. (A3) satisfies the normalization condition $\sum_{m=1}^{c} P(m|c) = 1$.

[1] A. Barrat, M. Barthélemy, and A. Vespignani, *Dynamical Processes on Complex Networks* (Cambridge Univeristy Press, Cambridge, 2008).

[2] D. Easley and J. Kleinberg, *Networks, Crowds, and Markets: Reasoning About a Highly Connected World* (Cambridge Univeristy Press, Cambridge, 2010).

[3] M. E. J. Newman, *Networks: An Introduction* (Oxford Univeristy Press, Oxford, 2010).

[4] D. J. Watts, Proc. Natl. Acad. Sci. USA **99**, 5766 (2002).

[5] J. P. Gleeson, Phys. Rev. E **77**, 046117 (2008).

[6] E. A. Bender and E. R. Canfield, J. Comb. Theory Ser. A **24**, 296 (1978).

[7] B. Bollobás, Eur. J. Comb. **1**, 311 (1980).

[8] M. E. J. Newman, S. H. Strogatz, and D. J. Watts, Phys. Rev. E **64**, 026118 (2001).

[9] A. Vázquez and Y. Moreno, Phys. Rev. E **67**, 015101(R) (2003).

[10] A. V. Goltsev, S. N. Dorogovtsev, and J. F. F. Mendes, Phys. Rev. E **78**, 051105 (2008).

[11] J. P. Gleeson, S. Melnik, and A. Hackett, Phys. Rev. E **81**, 066114 (2010).

[12] P. Trapman, Theor. Popul. Biol. **71**, 160 (2007).

[13] J. C. Miller, Phys. Rev. E **80**, 020901(R) (2009).

[14] M. E. J. Newman, Phys. Rev. Lett. **103**, 058701 (2009).

[15] J. P. Gleeson, Phys. Rev. E **80**, 036107 (2009).

[16] J. Gao, S. V. Buldyrev, S. Havlin, and H. E. Stanley, Phys. Rev. Lett. **107**, 195701 (2011).

[17] C. D. Brummitt, K.-M. Lee, and K.-I. Goh, Phys. Rev. E **85**, 045102(R) (2012).

[18] X. Huang, S. Shao, H. Wang, S. V. Buldyrev, H. E. Stanley, and S. Havlin, Europhys. Lett. **101**, 18002 (2013).

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

[20] A. Vázquez, R. Pastor-Satorras, and A. Vespignani, Phys. Rev. E **65**, 066130 (2002).

[21] M. Á. Serrano and M. Boguñá, Phys. Rev. E **74**, 056114 (2006).

[22] A. V. Goltsev, S. N. Dorogovtsev, and J. F. F. Mendes, Phys. Rev. E **73**, 056101 (2006).

[23] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, Phys. Rev. Lett. **96**, 040601 (2006).

[24] S. R. Broadbent and J. M. Hammersley, Proc. Cambridge Philos. Soc. **53**, 629 (1957).

[25] D. Stauffer and A. Aharony, *Introduction to Percolation Theory*, 2nd ed. (Taylor and Francis, London, 1992).

[26] A. Hackett, S. Melnik, and J. P. Gleeson, Phys. Rev. E **83**, 056107 (2011).

[27] J. P. Gleeson and D. J. Cahalane, Phys. Rev. E **75**, 056103 (2007).

[28] J. P. Sethna, K. Dahmen, S. Kartha, J. A. Krumhansl, B. W. Roberts, and J. D. Shore, Phys. Rev. Lett. **70**, 3347 (1993).

[29] D. Dhar, P. Shukla, and J. P. Sethna, J. Phys. A **30**, 5259 (1997).

[30] P. Shukla, Int. J. Mod. Phys. B **17**, 5583 (2003).

[31] D. J. Watts and P. S. Dodds, J. Consum. Res. **34**, 441 (2007).

[32] D. Lopez-Pintado and D. J. Watts, Ration. Soc. **20**, 399 (2008).

[33] An extension of this approach to nonmonotone binary-state dynamics has been provided in [34].

[34] J. P. Gleeson, Phys. Rev. Lett. **107**, 068701 (2011).

[35] N. E. Friedkin, J. Math. Soc. **12**, 103 (1986).

[36] N. E. Friedkin and E. C. Johnsen, *Social Influence Network Theory: A Sociological Examination of Small Group Dynamics* (Cambridge University Press, Cambridge, 2011).

[37] M. E. J. Newman, Phys. Rev. E **68**, 026121 (2003).

[38] M. E. J. Newman, SIAM Rev. **45**, 167 (2003).

[39] M. W. Wong, *An Introduction to Pseudo-Differential Operators*, 2nd ed. (World Scientific, Singapore, 1999).

[40] The condition $\sum_{m=0}^{c-1} R_m^{c-1}(q_n) = 1$ can be verified by a simple algebraic manipulation of the summation.

[41] D. S. Callaway, M. E. J. Newman, S. H. Strogatz, and D. J. Watts, Phys. Rev. Lett. **85**, 5468 (2000).

[42] A model of cascades on weighted multiplex networks has been proposed recently by Yağan and Gligor [43]. The application of this model to clique-based graphs may also provide novel insights.

[43] O. Yağan and V. Gligor, Phys. Rev. E **86**, 036103 (2012).

[44] B. Karrer and M. E. J. Newman, Phys. Rev. E **82**, 066118 (2010).