EPJ B
Condensed Matter and Complex Systems
EPJ.org your physics journal

Eur. Phys. J. B (2015) 88: 201 DOI: 10.1140/epjb/e2015-60263-6

Analysis of information diffusion for threshold models on arbitrary networks

Sungsu Lim, Inwoo Jung, Seulki Lee and Kyomin Jung

edp sciences
Società Italiana di Fisica
Springer

Eur. Phys. J. B (2015) 88: 201
DOI: 10.1140/epjb/e2015-60263-6

THE EUROPEAN PHYSICAL JOURNAL B

Regular Article

Analysis of information diffusion for threshold models on arbitrary networks

Sungsu Lim[1], Inwoo Jung[2], Seulki Lee[3], and Kyomin Jung[2,a]

[1] Department of Knowledge Service Engineering, KAIST, Daejeon 34141, Korea
[2] Department of Electrical and Computer Engineering, Seoul National University, Seoul 08826, Korea
[3] SoHoBricks Corp., Irvine, CA 92614, USA

Received 1st April 2015 / Received in final form 23 June 2015
Published online 10 August 2015 – © EDP Sciences, Società Italiana di Fisica, Springer-Verlag 2015

Abstract. Diffusion of information via networks has been extensively studied for decades. We study the general threshold model that embraces most of the existing models for information diffusion. In this paper, we first analyze diffusion processes under the linear threshold model, then generalize it into the general threshold model. We give a closed formula for estimating the final cascade size for those models and prove that the actual final cascade size is concentrated around the estimated value, for any network structure with node degrees $\omega(\log n)$, where n is the number of nodes. Our analysis analytically explains the tipping point phenomenon that is commonly observed in information diffusion processes. Based on the formula, we devise an efficient algorithm for estimating the cascade size for general threshold models on any network with any given initial adopter set. Our algorithm can be employed as a subroutine for numerous algorithms for diffusion analysis such as influence maximization problem. Through experiments on real-world and synthetic networks, we confirm that the actual cascade size is very close to the value computed by our formula and by our algorithm, even when the degrees of the nodes are not so large.

1 Introduction

How information, ideas, and rumors diffuse through a social network has been a fundamental question for many decades in various disciplines including computer science, social science, and economics [1–5]. Analysis of information diffusion patterns is crucial in understanding and predicting information flows in applications ranging from epidemiology to viral marketing. For instance, suppose that a small fraction of individuals are attracted by a marketing strategy (i.e., by providing a sample or a discounted price) and that they purchase a product. Then, either actively or inactively, their purchases influence their neighbors by some form of word-of-mouth effect, triggering an information cascade in the network.

There is vast literature on information diffusion analysis in social and information networks [6–11]. Various models of information spreading have been studied. In particular, the independent cascade model (or the stochastic SIR model), the linear threshold model, and their generalization proposed by Kempe et al. [12], called the *general threshold model*, are established based on the common assumption that the neighbors play significant roles for the spread of information [6]. The general threshold model can also be used to describe wide variety of other social network diffusion models (see Appendix A).

In the analysis of diffusion process, the computation of the final cascade size is of the most interest. One essential drawback of most of previous studies on the cascade analysis is that they are not typically applicable to general classes of network structure, since they have focused on limited conditions such as complete graphs [6] and locally tree-like networks [9,10,13]. However, the structure of interactions among individuals in real societies is often highly clustered, meaning that the network has many triangles and short cycles.

In this paper, we first develop a new technique to estimate the final cascade size under the linear threshold model, then generalize it into the general threshold model. We give a closed formula that estimates the final cascade size and prove its closeness to the actual cascade size. Surprisingly, the accuracy of our estimation is proved to be dependent only on the size of network. This implies that the final cascade size is asymptotically independent of network structure if the network is *slightly dense*. Moreover, unlike many previous studies, our result applies to any *threshold distribution* of the general threshold model. Additionally, our method analytically explains the *tipping point phenomenon*, in which the slight change in initial condition can result in drastic change in the final cascade size.

1.1 The general threshold model

For a given graph $G = (V, E)$, each node $v \in V$ can be in one of two states; active (state 1) or inactive (state 0).

[a] e-mail: kjung@snu.ac.kr

In the general threshold model, each node v has its own threshold function f_v:$\{0,1\}^{d_v} \to [0,1]$ which depends only on the states of its in-neighbors where d_v is the in-degree of v. The f_v is assumed to be monotonically increasing pseudo-boolean function. At the beginning of diffusion process, a threshold value $\theta_v \in [0,1]$ is drawn from a probability distribution μ in an i.i.d. manner, for each node $v \in V$. We call μ the *threshold distribution*.

The diffusion process then proceeds in discrete time steps $t = 0, 1, 2 \ldots$; let $Z_v(t) = 1$ if v has adopted the information at iteration t and $Z_v(t) = 0$ otherwise. At time $t = 0$, each node $v \in V$ becomes a initial adopter independently with some probability x_v. Then for each iteration at $t = 1, 2, \ldots$, each node v adopts the information at t if $f_v(\mathcal{N}_v(t-1)) \geq \theta_v$, where $\mathcal{N}_v(t) = (Z_u(t))_{uv \in E}$. This step-wise process repeats itself until no additional adoption is possible. The final cascade size $\mathcal{Z}$ is defined as the number of nodes that have adopted the information at the end of the process.

One of the most widely studied general threshold models is the *linear threshold model*, in which the threshold functions f_v are linear [6,14,15], i.e., each edge has a weight, and a node v becomes active if the weighted sum of its activated neighbors exceeds θ_v. For instance, $f_v(\mathcal{N}_v(t)) = \sum_{u:uv \in E} w_{uv} Z_u(t)$, where $w_{uv} = 1/d_v$, has been widely studied [9]. Another important example of the general threshold model is the *SIR (Susceptible-Infected-Recovered) model*. In this model, each active node has a single chance to influence each of its neighbors independently with some edge activation probabilities. More examples of the general threshold model are given in Appendix A.

1.2 Related work

Although information diffusion processes in social networks have been researched extensively, estimating the final cascade size has been limited to the networks with unrealistic assumptions, which are far from actual properties of real-world social networks. Also, the threshold model has rarely been analyzed for arbitrary graphs and arbitrary threshold distributions because such analysis is very difficult [16]. Here we review related works on estimating the cascade size for some class of diffusion models in networks, along with their assumptions.

(i) *Assumptions on network structure:* Granovetter's pioneering work on the linear threshold modeldemonstrates the information diffusion patterns and conditions for the occurrence of a tipping point on complete graphs [6]. Watts provided how a global cascade is triggered for a sparse, random network without short cycles, under the linear threshold model [9].
Whitney's work states that in Erdös-Rényi random graphs where the threshold values of all nodes are the same, the average cascade size and the set size of initial adopters that triggers a global cascade can be computed [10]. Likewise, most of previous works assumed locally tree-like network structure, while real-world social networks usually have a lot of triangles and short cycles.

(ii) *Assumptions on threshold distributions:* In some previous studies under the threshold models, it is assumed that threshold values θ_v are drawn from the uniform distribution μ on [0,1], or θ_v are assumed to be constant.

Another general assumption for estimating the cascade size is that the final cascade size $\mathcal{Z}$ as a function of a given set of initial adopters is submodular. However, assuming that $\mathcal{Z}$ is submodular usually imposes strong condition on threshold distributions. For example, one can prove that under the linear threshold model, $\mathcal{Z}$ is submodular if and only if the threshold distribution μ is non-increasing. However, it is known that μ follows neither non-increasing nor uniform distribution in most real-world scenarios, but follows some bell-shaped distributions, such as the normal distribution [4,17].

1.3 Our contribution

(1) We provide a closed formula that estimates the final cascade size under the general threshold model. Our result only requires that all nodes $v \in V$ have degrees $\omega(\log n)$, where $n = |V|$, and that the initial adopters are chosen independently at random with some identical probability (i.e., $x_v = x$ for all $v \in V$). Also, the threshold distribution μ can be *any continuous probability distribution*. In Section 3, we apply additional technique so that the x_v's can be distinct for each $v \in V$.
From our result, it follows that the final cascade size is *asymptotically independent* of network structure. Moreover, our analysis displays the *tipping point phenomenon* that is frequently observed in real-world information diffusion scenario, where a slight change in initial adoption pattern can affect greatly on the final cascade size. Therefore, our analysis can comprehensively be applied to many diffusion processes in social networks, for example, caused by the public and targeted advertising with word-of-mouth communications.

(2) Based on our result, we also devise an heuristic algorithm Influence Estimate for estimating the final cascade size. Our algorithm is much more rapid than the traditional Monte-Carlo simulation.
Influence Estimate can be employed as a subroutine of many algorithms for diffusion size analysis such as influence maximization problem [12,18–20], which is interesting since the influence maximization under the threshold models has not been studied widely [21,22], unlike the independent cascade model. Influence Estimate is also applicable for measuring the resilience and the vulnerability of networks [16,23]. The goal of this problem is to find robust structures of networks in the presence of

cascading failures or attacks, such as a financial crisis [24], an electrical blackout [25], the failure of Internet routers [26], and the spread of a disease [27].

(3) A set of experiments to confirm correctness of our closed formula and Influence Estimate were conducted on real-world and synthetic network topologies. Interestingly, although the computation in Influence Estimate is formally based on the extended result for the general threshold model case, not only does it successfully compute the cascade sizes for the cases that strictly follow our assumptions, but also the result is applicable for many cases that do not fully satisfy the conditions required for our analysis for the general threshold model (e.g. the degree of each node is less than $\log n$).

1.4 Organization

The rest of the paper is organized as follows. In Section 2, we prove the result of our technique on the linear threshold model. In Section 3, we generalize our result to an arbitrary general threshold model, which gives rise to Influence Estimate algorithm. Then, Section 4 shows experimental results, and Section 5 describes directions for the future work.

2 Estimation of final cascade size in linear threshold model

In this section, we state and prove a closed formula that estimates the value of the final cascade size $\mathcal{Z}$ in the linear threshold model. First, consider a simple case with a complete graph topology with n nodes where $w_{uv} = \frac{1}{n-1}$, as described by Granovetter [6]. Let F_0 be the cumulative distribution of μ. Let α_t be the proportion of influenced nodes at iteration t, i.e., $\alpha_t = \frac{\sum_{v \in V} Z_v(t)}{n}$. Note that α_0 is the fraction of initial adopters. At iteration 1, any nodes with thresholds at most α_0 will adopt the information. Likewise, at iteration $(t+1)$ with α_t given, any nodes with thresholds at most α_t will adopt the information. Since $Pr(\theta_v \leq \alpha_t) = F_0(\alpha_t)$, we get $\alpha_{t+1} \approx F_0(\alpha_t)$ for $t = 1, 2, 3 \ldots$ by standard concentration theorems such as the Chernoff bound. Let α^* be the minimum solution of $F_0(\alpha) = \alpha$. By repeating this process for $t = 0, 1, 2, \ldots$, we obtain that $\mathcal{Z} \approx n\alpha^*$ as in Figure 1.

In the work, we extend this observation into more general setup. Here we suppose that all individuals become initial adopters with an identical initial adoption probability x, i.e., $x_v = x$ for all $v \in V$. The threshold values θ_v are normalized to be in $[0, 1]$. Consider an arbitrary threshold distribution μ, and assume that $\sum_{u:uv \in E} w_{uv} = 1$ for all $v \in V$[1].

[1] In the linear threshold model, $\sum_{u:uv \in E} w_{uv} \leq 1$, and by normalizing w_{uv}'s and θ_v's, we may assume $\sum_{u:uv \in E} w_{uv} = 1$ without loss of generality.

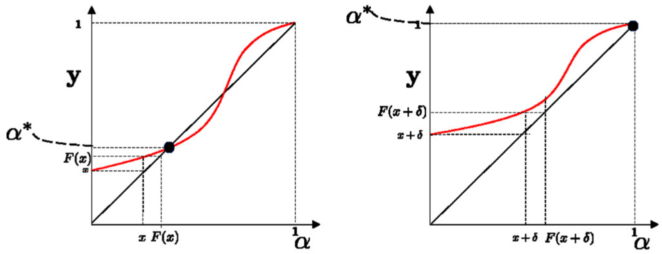


Fig. 1. These two plots describe where the convergence point α^* occurs for F_0 and F. In F where $F(0)$ is set to be $x + \delta$, the convergence point jumps to far greater value, resulting in the tipping point phenomenon. This can represent, for example, the effect of initial marketing in real-world diffusion scenarios.

In order to simplify our analysis, note that the linear threshold model with a new cumulative threshold distribution $F = x + (1-x)F_0$ and zero initial adoption probability is equivalent to the linear threshold model with F_0 and initial adoption probability x. In other words, under threshold distribution F, every node has constant probability x for its threshold θ_v to be exact 0, which means that these nodes will inevitably adopt information at iteration 0. Thus, we consider the general cumulative threshold distribution F with $x = 0$ for the ease of explanation.

Let $\alpha_0 = x$ and $\alpha_t = F(\alpha_{t-1})$ for $t \in \mathbf{N}$. Let $\alpha^* = \lim_{t\to\infty} \alpha_t$. Note that the definition of α_t and α^* are slightly different from those in the example at the beginning of this section, but one can check that $\alpha^* = \min\{\alpha | F(\alpha) = \alpha\} \leq 1$. We will essentially prove that the final cascade size is very close to $\alpha^* n$ with high probability and such property is asymptotically independent of the network structure.

Theorem 1. *Suppose that μ is continuous and $y = F(\alpha)$ is not tangential to $y = \alpha$ at α^*. If $w_{uv} = O\left(\frac{1}{\log n}\right)$ for every $uv \in E$, then for any $\epsilon > 0$ and $\delta > 0$,*

$$Pr[|\mathcal{Z} - \alpha^* n| > n\epsilon] = o(n^{-\delta}).$$

Note that α^* is independent of the network structure, and only dependent on μ and x. Also note that the threshold distribution μ should only satisfy a mild condition that the cumulative distribution $y = F(\alpha)$ is not tangential to the line of $y = \alpha$ at α^*. This is to guarantee α's convergence from both upper and lower sides. The last condition $w_{uv} = O(\frac{1}{\log n})$ is satisfied when for example, the network is slightly dense so that $d_v = \omega(\log n)$ and a certain edge does not have dominant influence so that $w_{uv} = O(\frac{1}{\log n})$. Note that we assumed $\sum_{u:uv \in E} w_{uv} = 1$.

In the proof of Theorem 1, we show that the cascade size after t iteration is very close to $\alpha_t n$ for all t for any network satisfying the conditions of Theorem 1. Specifically, it is proven that any node with threshold $\theta_v < \alpha_t$ adopts the information with high probability until the time t, while any node with threshold $\theta_v > \alpha_t$ does not adopt the information until the time t. This argument is proven by induction on t, and applying Azuma's inequality over each step. The detailed proof of Theorem 1 is in Appendix B.

In Figure 1, we can observe that $\alpha^\star$, the expected fraction of the final cascade size, becomes discontinuous function of the initial adoption probability x if F_0 is given. This implies that a slight change in initial adoption probability of information (e.g., by promotion in real-world) can result in huge difference on the value of $\alpha^\star$. This is commonly referred as the *tipping point phenomenon*. Therefore, when the initial adopters are chosen uniformly at random and independently, we can compute the exact values of initial adoption probability x that triggers the tipping point phenomenon, using the threshold distribution μ.

Examples. There are various examples where assumptions in Theorem 1 are satisfied.

Example 1. Blume et al. [16] studied the cascading failures on d-regular networks. Cascading failure refers to phenomenon where failure of single or several network node results in propagation of failure on other nodes. Here the failure of a node can act as an instance information that is being diffused. The threshold model in this work is assumed to have the threshold value θ_v for each node v that follows a distribution μ and f_v is determined by the number of influenced neighbors of v. This is a linear threshold model with a threshold distribution μ with thresholds being scaled down into $[0,1]$ and $w_{uv} = 1/d$. Our result shows that the effect of the cascading failure on d-regular network is asymptotically independent of the network structure if $d = \omega(\log n)$.

Example 2. Theorem 1 can be extended to the bipartite graph case. Suppose that G is a bipartite graph with a partition (P_1, P_2), in which all nodes in P_i choose their thresholds from the same threshold distribution μ_i, and all nodes in P_i have the same probability x_i of being an initial adopter for each $i = 1$ and 2. Then only if all nodes are of degree $\omega(\log n)$, we can derive a closed formula for the expected final cascade size, and prove concentration inequalities of the final cascade size for each P_i, by the same argument as in Theorem 1. Common examples include the transmission of Sexually Transmitted Disease (STD) explained by information diffusion on bipartite graphs [28].

Example 3. In many social networks, individuals participate in more than one community, and they can engage in more than one type of interaction. The multiplex network model [29,30] allows this realistic scenario. Consider that each node v has J types of interactions with its neighbors, and the threshold function $f_{v,i}$ and the threshold distribution μ_i are given for each type of interaction $i = 1, \ldots, J$, separately. Then, v adopts the information if at least one of J rules is satisfied with respect to each $f_{v,i}$ value. Our analysis can be extended to this case to compute the final cascade size on an arbitrary network under the same condition on node degrees.

3 Extension to general threshold model

In this section, we generalize our results to the general threshold model in which the threshold functions f_v's are arbitrary pseudo-boolean functions. We also allow x_v to be distinct for each $v \in V$, i.e., each individual v becomes an initial adopter independently with its own probability x_v. It is assumed that the threshold distribution μ and its c.d.f. F satisfy the same set of conditions as in Section 2. In the proof of Theorem 1, we consider that F is Lipschitz continuous, and the Lipschitz constant of F is referred to as λ_F in this section.

Similar to Section 2, we normalize the threshold values θ_v's to be in $[0,1]$. Also, for each threshold function $f_v\big([Z_u(t)]_{u\in N_v}\big)$, where N_v is the neighborhood of v, we apply linear transformations so that $f_v(\mathbf{0}) = F^{-1}(x_v)$ and $f_v(\mathbf{1}) = 1$ for each $v \in V$. Then, $Pr(\theta_v \leq f_v(\mathbf{0})) = x_v$ so each v becomes an initial adopter with probability x_v at $t = 0$. Then we set the initial adoption probability to be 0 for all $v \in V$. Under this equivalent model, we can derive a generalized result from Section 2 and apply it to more general set of networks even if the x_v's are distinct for each v.

Remind that $Z_t(v)$ is an indicator variable such that $Z_t(v) = 1$ if a node v is influenced at the tth iteration (adopted the information at the tth iteration or earlier) and $Z_t(v) = 0$ otherwise. Similarly, let $Z(v)$ be the indicator variables representing whether a node v is influenced or not at the end of the process.

Let $f_v^* : [0,1]^{d_v} \to [0,1]$ be the function such that $f_v^*(E[\boldsymbol{\nu}]) = E[f_v(\boldsymbol{\nu})]$ for any $\boldsymbol{\nu} = (\nu_1, \nu_2, \ldots, \nu_{d_v})$ where $\nu_1, \ldots, \nu_{d_v}$ are independent Bernoulli random variables. Let $\alpha_{-1}(v) = 0$ for all v. For $t \in \mathbf{Z}^+$, let

$$\alpha_t(v) = F\big(f_v^*(\alpha_{t-1}(u_1), \alpha_{t-1}(u_2), \ldots, \alpha_{t-1}(u_{d_v}))\big). \quad (1)$$

Note that $\alpha_0(v) = x_v$ for all $v \in V$. Then, one can check that $\alpha_0(v) \leq \alpha_1(v) \leq \ldots \leq \alpha^*(v)$ for all $t \in \mathbf{Z}^+$ and for all $v \in V$, since f_v's and F are monotonically increasing. Since $\alpha_t(v) \leq 1$, $[\alpha_t(v)]_{v\in V} \in [0,1]^n$ converges to some $[\alpha^*(v)]_{v\in V}$. Let $\bar{\sigma}_t = \sum_{v\in V} \alpha_t(v)$ and $\bar{\sigma}^* = \sum_{v\in V} \alpha^*(v)$.[2]

Similar to Section 2, $\bar{\sigma}^*$ and $\bar{\sigma}_t$ become the estimated cascade size, at the end of the process and at time t, respectively. Furthermore, we prove that each $\alpha^*(v)$ and $\alpha_t(v)$ become good estimates for $Z(v)$ and $Z_t(v)$, respectively, whose values are between 0 and 1. We have the following theorem for the general threshold model, which is a generalization of Theorem 1 in Section 2.

Theorem 2. *Let $\mathcal{Z}_t$ be a cascade size at iteration $t \in \mathbf{Z}^+$. Suppose the followings for each $v \in V$.*

(i) *There exists $\lambda_f \geq 0$ such that*

$$|f_v^*(\mathbf{a}) - f_v^*(\mathbf{b})| \leq \lambda_f \|\mathbf{a} - \mathbf{b}\|_\infty$$

for any $\mathbf{a}, \mathbf{b} \in [0,1]^{d_v}$ and all $v \in V$.

(ii) *Let $\Psi_v = \{\boldsymbol{\nu},\ \boldsymbol{\nu}' \in \{0,1\}^{d_v} | \nu(i) \neq \nu'(i)$ for only one $i\}$. Then, suppose*

$$\max_{\Psi_v} |f_v(\boldsymbol{\nu}) - f_v(\boldsymbol{\nu}')| = o\left(\sqrt{\frac{1}{d_v \log n}}\right). \quad (2)$$

[2] Note that $\bar{\sigma}_t$ and $\bar{\sigma}^*$ are the counterparts of $\alpha_t n$ and $\alpha^* n$ in Section 2, respectively.

(iii) *Let* $\nabla f_v(\alpha^*(u_1), \ldots, \alpha^*(u_{d_v}))$ *be a gradient of* f_v *at* $(\alpha^*(u_1), \ldots, \alpha^*(u_{d_v}))$ *where* $u_1, \ldots, u_{d_v}$ *are neighbors of* v. *Suppose that all coordinates of* $\nabla f_v(\alpha^*(u_1), \ldots, \alpha^*(u_{d_v}))$ *are different from* 1. *Then for any constants* $m > 0$, $\epsilon > 0$, *and* $\delta > 0$,

$$\Pr[|\mathcal{Z}_m - \bar{\sigma}_m| < n\epsilon] = 1 - o(n^{-\delta}),$$
$$\Pr[\mathcal{Z} \leq \bar{\sigma}^* + n\epsilon] = 1 - o(n^{-\delta}).$$

Also we have that for all $v \in V$,

$$|E[Z_m(v)] - \alpha_m(v)| = o(1),$$
$$E[Z(v)] \geq \alpha^*(v) + o(1).$$

Let $\lambda = \lambda_F \lambda_f$. *If* $\lambda < 1$,

$$\Pr[|\mathcal{Z} - \bar{\sigma}^*| < n\epsilon] = 1 - o(n^{-\delta}),$$
$$|E[Z(v)] - \alpha^*(v)| = o(1)$$

for all $v \in V$.

The proof of Theorem 2 is a generalization of the proof of Theorem 1 and is given in Appendix C. The three conditions of Theorem 2 are multi-dimensional versions of the conditions of Theorem 1.

The first condition means that f_v^* is Lipschitz continuous, and the condition (2) is satisfied when the graph is slightly dense so that $d_v = \omega(\log n)$ and a certain edge does not have dominant influence so that the condition is satisfied, since f_v's are normalized to be in $[0, 1]$.

The last condition means that for each node v, all d_v partial derivatives of f_v are not equal to 1. To explain the meaning of the condition, consider a set of standard two-dimensional plots where the x-axis stands for a single element of the Euclidean input vector of $f_v^\star$ with other elements being fixed, then the evaluation of $f_v^\star(x)$ is plotted along the y-axis. The condition simply means that $y = f_v^\star(x)$ is never tangential to $y = x$ at $\alpha^*(v)$ for every such plot possible. In short, it is a generalization of the requirement that $y = F(x)$ is not tangential to $y = x$ in Section 2.

Examples. There are numerous examples where the conditions of Theorem 2 are satisfied.

Example 1 *(the SIR model)*. For each $v \in V$, let $N_v(t) = \{Z_t(u_1), \ldots, Z_t(u_{d_v})\}$, where $u_1, \ldots, u_{d_v}$ are neighbors of v. This model can be regarded as the general threshold model with threshold functions $f_v(N_v(t)) = x_v + (1 - x_v)(1 - \prod_{i=1}^{d_v}(x - p_{u_i v} Z_t(u_i)))$, where $p_{u_i v}$ is the probability that a node u_i influences node v. In this case, if $d_v = \omega(\log n)$ and $p_{u_i v} \leq \frac{c_{u_i v}}{d_v}$ so that $c_{u_i v} = O(1)$, then the conditions of Theorem 2 are satisfied hence the final cascade size is concentrated around α^*. Moreover, when $c_{u_i v} < 1$, $\lambda < 1$ it holds that $|E[Z(v)] - \alpha(v)| = o(1)$.

Example 2 *(collaborative behaviors)*. Suppose that an individual is influenced by the collaborative behaviors of K disjoint communities on a network. When $K = w(\log n)$ and $f_v(N_v(t)) = \sum_{k=1}^{K} \frac{1}{K}(\vee_{u_i \in \text{community } k, u_i v \in E} Z_t(u_i))$, it satisfies the conditions of Theorem 2. The function f_v indicates that an individual v is influenced according to the number of communities having at least one adopter.

Algorithm 1 Influence Estimate algorithm

1: $t := -1$
2: **for each** $v \in V$ **do**
3: $\quad \alpha_t(v) := 0$
4: **end for**
5: **while** (stopping criterion not met) **do**
6: $\quad t := t + 1$
7: $\quad$ **for each** $v \in V$ **do**
8: $\quad\quad \alpha_t(v) := F\big(f_v^*(\alpha_{t-1}(u_1), \alpha_{t-1}(u_2), \ldots, \alpha_{t-1}(u_{d_v}))\big)$
9: **return** $\bar{\sigma}^* = \sum_{v \in V} \alpha_t(v)$

As a result of Theorem 2, we propose an efficient heuristic algorithm Influence Estimate that estimates $E[Z(v)]$ for each $v \in V$ and $E[\mathcal{Z}]$ under any general threshold model for any network structure with any initial adopter set. The algorithm essentially computes $\alpha^*(v)$ by an iterative equation (1).

The pseudo-code of Influence Estimate is shown above. Note that if f_v^* can be computed in polynomial time, then Influence Estimate works in polynomial time. If the f_v is from any algebraic functions and some transcendental functions, f_v^* can be evaluated very fast. The threshold functions of the linear threshold model are the most representative example.

The running time of Influence Estimate is $O(m \cdot \sum_{v \in V} T(f_v^*))$ where m is the total number of iterations and $T(f_v^*)$ is the time taken for evaluating f_v^*. This is essentially the same as the time required for one trial of a Monte Carlo experiment. Thus, Influence Estimate is much faster than the Monte Carlo estimation, and it can be used as a subroutine for numerous algorithms for diffusion analysis.

If we consider the situation where each individual becomes an initial adopter independently with a given probability x, as in Section 2, then the $\alpha_0(v)$'s have the same values of α_0 for all $v \in V$, and Step 7 to Step 9 of the algorithm can be reduced to a single step $\alpha_t = F(f_v^*(\alpha_{t-1}))$. Furthermore, in this case, the sequence $\{\alpha_0, \ldots, \alpha_m\}$ for some constant $m > 0$ is independent of the network structure. It means that we can efficiently estimate the influence over time, even if we do not know the specific structure of given large-scale networks.

4 Experiments

We conducted a set of experiments with our algorithm Influence Estimate on real-world social networks, including Epinions networks [31] and political blogs [32], and synthetic networks including PA small-world networks and WS scale-free networks to validate the accuracy of Influence Estimate on networks having nodes with small degrees.

Table 1. Dataset statistics.

Dataset	Nodes	Edges	Average degree
Epinions	75 877	508 K	13.4
PolBlogs	1222	16 K	27.4
WS	10 000	500 K	100
PA	10 000	500 K	100

4.1 Network datasets

- PolBlogs: A directed network of hyperlinks between blogs on US politics [32], which was collected at around the time of the 2004 presidential election.
- Epinions: The Who-trusts-whom online social network of the general consumer review site Epinions.com [31], where nodes are members of the site and a directed edge from u to v represents v trusts u.
- WS: The Watts and Strogatz small-world network of 10,000 nodes with average degree 100. It was generated by uniform and random rewiring process of a ring lattice with probability with 0.5.
- PA: A graph with power-law degree distribution generated by the Barabási-Albert preferential attachment process. We varied the number of nodes and average degrees to observe the effect of network topologies on our theoretical results.

All networks were considered as undirected for higher average degree, but our results are applicable to directed graph as well. We removed nodes and edges that are not included in the largest connected component. Some basic statistics on these networks are given in Table 1.

In particular, the experiments were aimed at examining whether (1) the threshold values of individuals who adopt the information are actually concentrated at each iteration, (2) a phase transition appears at the point Influence Estimate computed, (3) adoption probability computed by Influence Estimate algorithm for each node is consistent with the Monte-Carlo simulation result, and (4) how network sizes and average degrees affect the accuracy of Influence Estimate.

In all of our experiments, we compare the expected value of adoption estimated by our proposed algorithm and the adoption probability from the Monte-Carlo simulation. We use a synthetically generated network or a given network, as an underying network structure. Then, we select a set of initial adopters using the threshold distribution and determine the adopted nodes from the initial adopters using an influence model. Repeating this process until the predefined number of times, and we estimate the adoption probability by taking an average for each node.

4.2 Influence model

We consider two different influence weight conditions for the linear threshold model. In both cases, we select each edge weight independently at random in [0,1]. For the first case, the sum of weights is normalized so that $\sum_{i=1}^{d_v} w_{uv_i} = 1$. For the second case, we adjusted the weights so that $\sum_{i=1}^{d_v} w_{uv_i}$ is selected uniformly at random from [0.5,1] for each node v. We call the latter case as the adjusted influenced weights. A threshold value is given to each node independently from a normal distribution with a specific mean and a standard deviation.

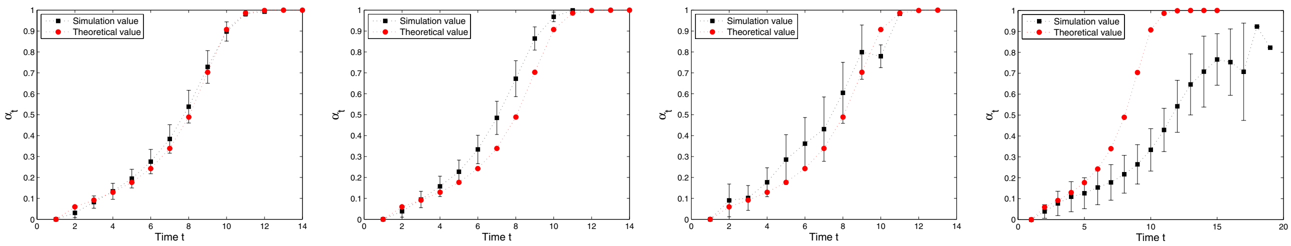
4.3 Results

First, we verified at each iteration of the diffusion process that threshold values of individuals who are influenced at that time are actually concentrated around the mean, as suggested in Theorem 1. We consider the normalized influence weight. The threshold distribution is $N(0.3,0.3)$ and the initial adopter set is selected with $x_v = 0.3$ for all nodes v, respectively.

Figure 2 shows the adopters' average threshold values with standard deviation error bars from 200 simulations. Note that in these simulations, a global cascade appears independently of the network structures. For synthetic network structures, the values show high concentrations around the means and consistency with theoretical value α_t's. However, there is a little deviation for PolBlogs and Epinions data, since they include many nodes with small degrees. Figure 3 represents the results when the threshold distribution is $N(0.5,0.3)$ with $x_v = 0.5$. In this case, the results also shows high concentrations around the mean, and a global cascade does not appear.

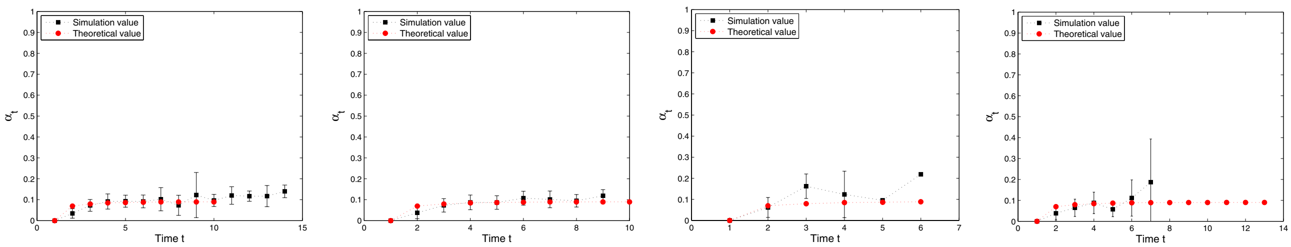
Theorem 2 allowed us to estimate the final cascade size under the adjusted influence weights with three threshold distributions: $N(0.4, 0.2)$, $N(0.3, 0.2)$, and $N(0.2, 0.2)$. We observed that phase transitions almost always appear at a certain tipping point as we increase the fraction x of initial adopters. Figure 4 represents tipping point distributions obtained from 1000 independent Monte-Carlo simulations and the theoretical values as well. For each case, the peak point is consistent with the theoretical value. The experimental results show that a tipping point appears at the predicted point even though the degrees of individuals are not so large.

We also compared $\alpha^*(v)$ with the fraction of simulation cases over 10 000 Monte-Carlo simulations that node v actually adopts the information when the diffusion process is over. We consider the adjusted influence weights with two threshold distributions: $N(0.3, 0.2)$ and $N(0.4, 0.2)$. The initial adopter sets are selected with $x_v = 0.3$ for all node v. For a fixed initial seed set, we conducted 10 000 simulations to assigning threshold θ_v differently. In PolBlogs, almost all values lies near the $y = x$ line, that is, theoretical results show strong consistency with simulation outcomes. In case of Epinions a number of data lies over the $y = x$ line. However, if we filter the data whose degree is less than $\log n$, it shows strong consistency as in Figure 5c. In order to visualize the results more clearly, we randomly select 1000 nodes and draw them in each of Figure 5. Specifically, the fraction of nodes whose estimated adoption probabilities from our algorithm are different with the Monte-Carlo estimate by 0.1 is only 0.036, 0.049, and 0.015 for Figures 5a–5c. That is, more than 95% nodes are predicted well for all the networks we used although the degree constraint is not satisfied.



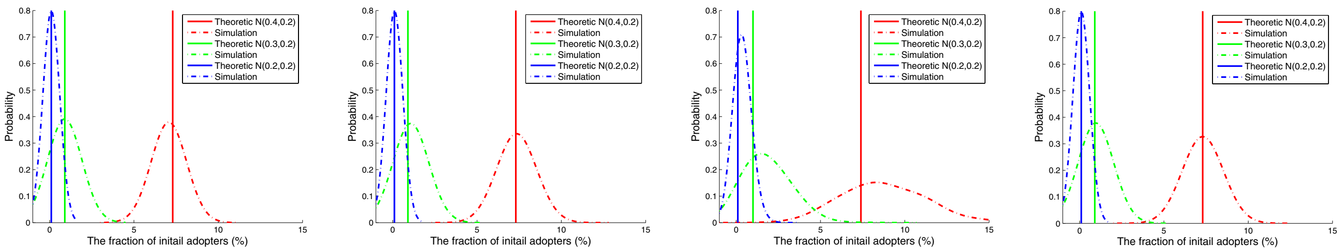
(a) WS (b) PA (c) PolBlogs (d) Epinions

Fig. 2. Concentration of adopters' average threshold values at each iterations with threshold distribution $N(0.3,0.3)$, $x_v = 0.3$ for all node v.



(a) WS (b) PA (c) PolBlogs (d) Epinions

Fig. 3. Concentration of adopters' average threshold values at each iterations with threshold distribution $N(0.5,0.3)$, $x_v = 0.5$ for all node v.



(a) WS (b) PA (c) PolBlogs (d) Epinions

Fig. 4. The distribution of the fraction of initial adopter set when a phase transition appears, obtained from 1000 simulations under the adjusted influenced weights with three different normal distributions whose mean varies from 0.2 to 0.4 and standard deviation is fixed at 0.2. Vertical lines denote theoretical values provided by our algorithm Influence Estimate.

Finally, we compared $\alpha^*(v)$ with the fraction of simulation cases that node v actually adopts the information by varying the network sizes and the average degrees of PA graphs. We consider the adjusted influence weights and the threshold distribution $N(0.3,0.2)$. Figure 6 represents the accuracy changes according to the average degrees of PA graphs, while the network size is fixed to 10 000. Figure 7 represents the accuracy changes according to the network sizes of PA graphs, while the average degree is fixed to 100. For the variation of network size and degrees, our algorithm Influence Estimate closely predicts the adoption probabilities.

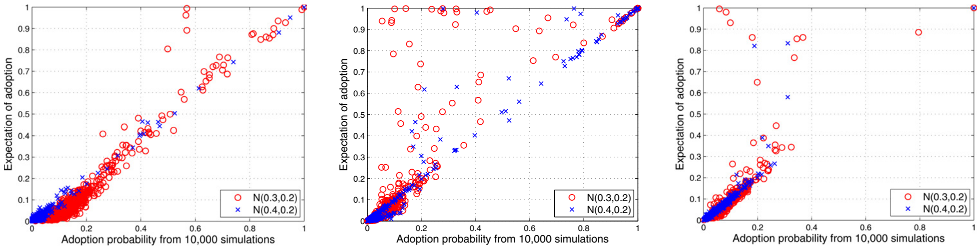
5 Discussion

The novelty of our analysis lies in its broad applicability to a general class of real-world information diffusion models and network topologies.

We believe that our technique can be applied to show concentration results for a broader class of problems in networks where the state of each node is locally determined by general Markov process where state-change rule for each node depends on its neighbors' current states. Examples of such problems include the network voter model, and the popular Gossip protocol.

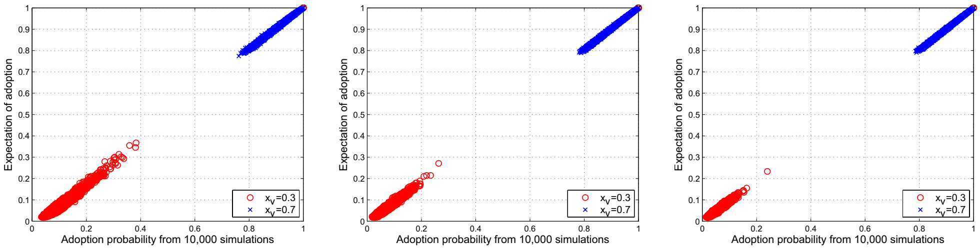
Another merit of our result is that our result can be viewed as a generalization of the *mean-field approximation* [33–38] that is widely used in analysis of information diffusion. In the usual mean-field approximation, the state change rules are approximately described as system of ordinary differential equations (ODEs) and the state ratio converges to the solution of such ODEs for complete graphs. Theorems 1 and 2 guarantee that the same ODEs can be applied for *any slightly dense* network.

As a future work, we plan to work on a generalization of our results into more general network setups. It would be also our future direction to extend down our results to milder conditions than the current requirement of $\omega(\log n)$ degrees.



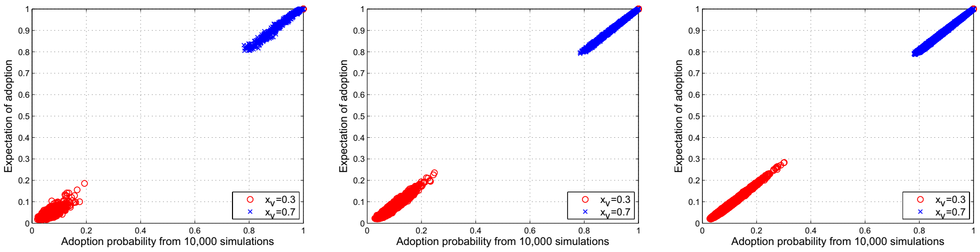
(a) PolBlogs (b) Epinions (c) Epinions representing v such that $d_v \geq \log n$

Fig. 5. Comparison of $\alpha^*(v)$ with the fraction of 10 000 simulation cases that node v actually adopts the information. Conducted under the adjusted influence weights, where $x_v = 0.3$ for all nodes v.



(a) Average degree 50 (b) Average degree 100 (c) Average degree 200

Fig. 6. Comparison of $\alpha^*(v)$ with the fraction of 10 000 simulation cases that similar to Figure 5 when the average degree of network varies. Conducted under the normalized influence weights with PA graph and threshold distribution $N(0.3{,}0.2)$, while $n = 10\,000$ for all three cases.



(a) $n = 2^{10}$ (b) $n = 2^{13}$ (c) $n = 2^{16}$

Fig. 7. Comparison of $\alpha^*(v)$ with the fraction of 10 000 simulation cases that similar to Figure 5 when the size of network varies. Conducted under the adjusted influence weights with PA graph and threshold distribution $N(0.3{,}0.2)$, while the average degrees are 100 for all three cases.

K. Jung is with ASRI, Seoul National University, and he is funded by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (2012R1A1A1014965).

Appendix A: Examples of threshold models

The threshold models are naturally applicable to numerous scenarios that involve basic notions of accumulative effects of neighbors represented by threshold functions.

Example 1 *(network game with selfish agents).* If an individual v decides to adopt (or not to adopt) the information and a neighbor of hers makes the same decision, she gets a payoff P_1 (or P_2 respectively). Otherwise, she gets no payoff. If N_v is the number of neighbors of v adopting the information among the d_v neighbors of v, she adopts the information if $N_vP_1 > (d_v - N_v)P_2$, that is, if $\frac{N_v}{d_v} > \frac{P_2}{P_1+P_2} \equiv \theta_v$. Therefore, this problem can be modeled by a linear threshold model and θ_v plays the role of a "node-specific threshold" for node v to adopt the information. The higher an individual's threshold is more conservative the individual is in adopting the information.

Example 2 *(network market).* In a network market, each potential buyer v has a *reservation price* r_v which denotes the highest price v is willing to pay for a unit of a good. The reservation price r_v is drawn from some probability distribution in an i.i.d. manner. Then, $g_v(z_v)$ represents the benefit of purchasing the product where $g_v : [0,1] \to [0,1]$ is an increasing function and z_v is a fraction of neighbors of v that have already purchased it. An individual v will buy the product if $r_v g_v(z_v) \geq p^*$, where p^* is the market price. The condition can be rewritten as $\frac{1}{d_v}\sum_{u:uv\in E} z_u \geq g^{-1}(\frac{p^*}{r_v})$. Then $g^{-1}(\frac{p^*}{r_v})$ plays a role analogous to the threshold value θ_v of the linear threshold model. Note that this is a generalization of the popular Katz-Shapiro pricing model [39] in which they assume that z_v is a shared information among all buyers and identical to all buyers, which is equivalent to the complete graph case.

Example 3 *(the SIR model).* The SIR model has been widely used in the analysis of epidemic spreading and information propagation in various networks. In this model, every time a node u becomes infected, every uninfected neighbor v of u has a single chance of becoming infected, with probability p_{uv}. The SIR model can be regarded as a special case of the general threshold model that has the threshold functions $f_v(N_v(t)) = x_v + (1-x_v)(1-\prod_{u:uv\in E}(1-p_{uv}Z_t(u)))$ where x_v is the probability that v becomes an initial adopter.

Appendix B: Proof of Theorem 1

Instead of proving Theorem 1 directly, we prove the following lemma, which directly implies Theorem 1.

Lemma 1. *Let $\epsilon > 0$ be given. Let F be a Lipschitz continuous function. Suppose that $y = F(\alpha)$ is not tangential to the line of $y = \alpha$ at α^*, and that for all $uv \in E$, $w_{uv} = o(\sqrt{\frac{1}{d_v \log n}})$. Then,*

- *for any $\delta > 0$,*

$$\Pr[|\mathcal{Z} - \alpha^* n| > n\epsilon] < o(n^{-\delta}), \tag{B.1}$$

- *and the influence can be calculated as*

$$\alpha = (\alpha^* \pm o(1))n. \tag{B.2}$$

Proof of Lemma 1. Let $\epsilon > 0$ be given. We will prove that there exists $\epsilon^* > 0$, dependent only on F and ϵ, such that

$$\Pr[|\mathcal{Z} - \alpha^* n| > n\epsilon] < 2n^{\left(2-\frac{(\epsilon^*)^2}{2C}\right)}, \tag{B.3}$$

where $C = \min_{v\in V, uv\in E}(d_v \log n (w_{uv})^2)$. Then, since $w_{uv} = o(\sqrt{\frac{1}{d_v \log n}})$ for every $uv \in E$, and $\lim_{n\to\infty} C = 0$, we prove (B.3) using the following (B.4) and (B.5):

$$\Pr\left[\mathcal{Z} < (\alpha_m - \frac{\epsilon}{2})n\right] < mn^{\left(1-\frac{(\epsilon^-)^2}{2C}\right)}, \tag{B.4}$$

$$\Pr[\mathcal{Z} > (\alpha^* + \epsilon)n] < n^{\left(2-\frac{(\epsilon^+)^2}{2C}\right)}, \tag{B.5}$$

where ϵ^+, ϵ^- are fixed values that are dependent on a given ϵ, m is a positive integer such that $|\alpha_m - \alpha^*| < \frac{\epsilon}{2}$. The existence of such m is guaranteed by the following lemma.

Lemma 2. *Let $\alpha_0 = F(0) = x$ and $\alpha_t = F(\alpha_{t-1})$ for $t \in \mathbf{N}$. Then, $\alpha_t \leq \alpha^*$ for all $t \in \mathbf{Z}^+$, and for any $\epsilon > 0$, there exists m (independent of n) such that $|\alpha_m - \alpha^*| < \frac{\epsilon}{2}$.*

Proof of Lemma 2. Note that $\alpha^* \geq x = F(0)$ because F is monotonically increasing. First, we show $\alpha_t \leq \alpha^*$ for all $t \in \mathbf{Z}^+$. Suppose not. Since $\alpha_0 = x \leq \alpha^*$, there exists $T \in \mathbf{N}$ such that $\alpha_T \leq \alpha^*$ but $\alpha_{T+1} = F(\alpha_T) > \alpha^* = F(\alpha^*)$. This is a contradiction because F is monotonically increasing. Note that $F(\alpha) > \alpha$ for all $\alpha \in [0, \alpha^* - \frac{\epsilon}{2}]$. Let $\Delta = \min_{\alpha\in[0,\alpha^*-\frac{\epsilon}{2}]}(F(\alpha) - \alpha) > 0$. Then, $\alpha_{t+1} \geq \Delta + \alpha_t$ for all $t \in \mathbf{Z}^+$. Hence, there exists m such that $|\alpha_m - \alpha^*| < \frac{\epsilon}{2}$ for any $\epsilon > 0$.

We will see later that (B.2) implies that any node v with $\theta_v \leq \alpha_t$ adopts the information with high probability until t iterations, while (B.3) implies that any node v with $\theta_v > \alpha_t$ does not adopt the information with high probability until t iterations.

Note that (B.4) and (B.5) prove not only (B.3) but also (B.2) of Lemma 1. Note that in general, if $\Pr[\mathcal{Z} < a] < B$, then $E[\mathcal{Z}] > 0 \cdot B + a(1-B)$ and that if $\Pr[\mathcal{Z} > c] < D$, then $E[\mathcal{Z}] < c(1-D) + nD$. Applying these on (B.2) and (B.3), we have

$$\left(1 - mn^{\left(1-\frac{(\epsilon^-)^2}{2C}\right)}\right)(\alpha^* - \epsilon) < \frac{\alpha}{n} < n^{\left(2-\frac{(\epsilon^+)^2}{2C}\right)} \cdot 1 + (\alpha^* + \epsilon).$$

Hence, for any given constant $\delta > 0$, we can find $N > 0$ such that, for all $n \geq N$, $|\alpha - \alpha^* n| \leq \delta n$.

Let $m, \alpha_0, \ldots, \alpha_m$ be obtained from Lemma 2 for a given ϵ. Let $\lambda \geq 0$ be a Lipschitz constant of F (i.e., $|F(a) - F(b)| \leq \lambda|a - b|$ for all $a, b \in [0, 1]$). Let $\epsilon^- = \frac{\epsilon}{2(\lambda^m + \ldots + \lambda^0)}$.

We first prove (B.4). Let $\epsilon_m = \frac{\epsilon}{2}$ and we generate $\{\epsilon_0, \ldots, \epsilon_{m-1}\}$ by the recurrence relation $\epsilon_{t-1} = \frac{\epsilon_t - \epsilon^-}{\lambda}$. Let $\epsilon_0 = \epsilon^-$. Consequently, $\epsilon_t = (\lambda^t + \lambda^{t-1} + \ldots + \lambda^0)\epsilon^-$.

Then, one can check that

$$\alpha_t - \epsilon_t \leq F(\alpha_{t-1} - \epsilon_{t-1}) + \lambda\epsilon_{t-1} - \epsilon_t = F(\alpha_{t-1} - \epsilon_{t-1}) - \epsilon^- \tag{B.6}$$

using the *mean value theorem.*

Let $\mathcal{Z}_t$ be the actual spread size after the tth iteration. We denote all neighbors of v as $u_1, \ldots, u_{d_v}$ where d_v is the degree of v, define $\{Z_t(v)\}$ as a set of indicator variables representing whether v is influenced until the tth iteration. Note that $Z_t(v) = 1$ if and only if $\sum_{i=1}^{d_v} w_{u_i v} Z_{t-1}(u_i) \geq \theta_v$ and that $\sum_{v\in V} Z_t(v) = \mathcal{Z}_t$. Let A_m be the set of nodes that have threshold less than $\alpha_{m-1} - \epsilon_{m-1}$. Then,

$$\begin{aligned}\Pr\left[\mathcal{Z}_m \geq (\alpha_m - \frac{\epsilon}{2})n\right] &= \Pr[\mathcal{Z}_m \geq (\alpha_m - \epsilon_m)n] \\ &\geq \Pr[\mathcal{Z}_m \geq F(\alpha_{m-1} - \epsilon_{m-1})n - \epsilon^- n]\end{aligned}$$

by (B.6). Noting that $F(\alpha_{m-1}-\epsilon_{m-1})n = E[|A_m|]$,

$$\begin{aligned}\Pr[\mathcal{Z}_m &\geq F(\alpha_{m-1}-\epsilon_{m-1})n-\epsilon^- n] \\ &= \Pr[\mathcal{Z}_m \geq E[|A_m|]-\epsilon^- n] \\ &\geq_{(a)} \Pr[\mathcal{Z}_m \geq |A_m|] \\ &\geq_{(b)} \Pr[\text{All nodes whose thresholds are less than} \\ &\qquad (\alpha_{m-1}-\epsilon_{m-1})\text{are influenced until the } m\text{th} \\ &\qquad \text{iteration}] \\ &= 1-\text{Error}^-.\end{aligned}$$

The inequality (a) demands $|A_m| \geq E[|A_m|]-\epsilon^- \cdot n$ to hold. Since $|A_m|$ is a sum of i.i.d. Bernoulli random variables in $[0,1]$ with probability $F(\alpha_{m-1}-\epsilon_{m-1})$, we get

$$\Pr\left[|A_m| \geq E[|A_m|]-\epsilon^- \cdot n\right] \geq 1-\exp\left(-2n\left(\epsilon^-\right)^2\right)$$

by Hoeffding's inequality. Therefore, for sufficiently large n, we can safely say (a) holds. The inequality (b) holds because the event that $\mathcal{Z}_m \geq |A_m|$ contains the event that $Z_t(v) = 1$ for all $v \in A_m$. The Error$^-$ is defined as Error$^-$ = Pr[There is some node whose threshold is less than $(\alpha_{m-1}-\epsilon_{m-1})$but which is not influenced by the mth iteration].

For $t = 0,\ldots,m-1$, let $\mathcal{E}_t$ be the event that there is a node whose threshold is less than $\alpha_t-\epsilon_t$ but is not influenced by the $(t+1)$th iteration. Note that Error$^-$ = $\Pr[\mathcal{E}_{m-1}]$. Observe that

$$\begin{aligned}\text{Error}^- &\leq \Pr[\mathcal{E}_0 \cup \mathcal{E}_1 \cup \ldots \cup \mathcal{E}_{m-1}] \\ &= \Pr[\mathcal{E}_0] + \sum_{t=1}^{m-1} \Pr[\mathcal{E}_t \cap (\cap_{t'=0}^{t-1} \mathcal{E}_{t'}^c)] \\ &\leq \Pr[\mathcal{E}_0] + \sum_{t=1}^{m-1} \Pr[\mathcal{E}_t \cap \mathcal{E}_{t-1}^c].\end{aligned}$$

Now we define $X_t^-(v)$'s as the indicator variables such that $X_t^-(v) = 1$ if $\theta_v < \alpha_{t-1}-\epsilon_{t-1}$ and $X_t^-(v) = 0$ otherwise. Note that $\mathcal{E}_t$ is the event that there exists a node v such that $\theta_v < \alpha_t - \epsilon_t$, and $\sum_{i=1}^{d_v} w_{u_i v} Z_t(u_i) < \theta_v$. Also note that $\mathcal{E}_{t-1}^c$ is contained in the event $Z_t(v) \geq X_t^-(v)$, for all $v \in V$. Thus, for $t = 1,\ldots,m-1$,

$$\begin{aligned}&\Pr[\mathcal{E}_t \cap \mathcal{E}_{t-1}^c] \\ &\leq \sum_{v\in V} \Pr\Big[\theta_v < \alpha_t-\epsilon_t,\ \sum_{i=1}^{d_v} w_{u_i v} Z_t(u_i) < \theta_v, \\ &\qquad Z_t(v) \geq X_t^-(v)\text{for all } v \in V\Big] \\ &\leq \sum_{v\in V} \Pr\left[\sum_{i=1}^{d_v} w_{u_i v} X_{t-1}^-(u_i) < \alpha_t-\epsilon_t\right] \\ &\leq \sum_{v\in V} \Pr\Big[\sum_{i=1}^{d_v} w_{u_i v} X_{t-1}^-(u_i) < F(\alpha_{t-1}-\epsilon_{t-1})-\epsilon^-\Big] \\ &\equiv \beta_t.\end{aligned}$$

Note that we used (B.6) above. From the following Lemma 3, we obtain $\beta_t < n\exp(-\frac{(\epsilon^-)^2\log n}{2C})$.

Lemma 3. *Let $X_1,\ldots,X_{d_v}$ be i.i.d. Bernoulli random variables with $E[X_1] = E[X_2] = \ldots = E[X_{d_v}] = B$ and $Y(v) = \sum_{i=1}^{d_v} w_{iv}X_i$. If $\max_i w_{iv} \leq \sqrt{\frac{C}{d_v \log n}}$ for some $C > 0$, for any $\tilde{\epsilon} > 0$,*

$$P(Y(v) \leq B-\tilde{\epsilon}) < \exp\left(-\frac{\tilde{\epsilon}^2\log n}{2C}\right).$$

Proof of Lemma 3. Note that $\sum_{j=1}^{d_v} w_{u_j v} E[X_{t-1}^-(u_j)] = F(\alpha_{t-1}-\epsilon_{t-1}) \cdot \sum_{j=1}^{d_v} w_{u_j v} = F(\alpha_{t-1}-\epsilon_{t-1})$ by our assumption.

To begin with, we define a sequence $\{Y_0(v),\ \ldots,\ Y_{d_v}(v)\}$, which is a martingale as follows.

$$\begin{aligned}Y_0(v) &= E\left[\sum_{j=1}^{d_v} w_{jv}X_j\right] = B \\ &\vdots \\ Y_i(v) &= E\left[\sum_{j=1}^{d_v} w_{jv}X_j \,\middle|\, X_1,\ldots,X_i\right] \\ &\vdots \\ Y_{d_v}(v) &= \sum_{j=1}^{d_v} w_{jv}X_j = Y(v).\end{aligned}$$

Since $|Y_{i+1}(v)-Y_i(v)|^2 \leq \frac{C}{d_v \mathrm{log} n}$, by applying Azuma's inequality, the inequality holds.

Similarly, by Lemma 3,

$$\begin{aligned}\Pr[\mathcal{E}_0] &\leq \sum_{v\in V} \Pr\left[\theta_v < \alpha_0-\epsilon_0, \sum_{i=1}^{d_v} w_{u_i v} Z_0(u_i) < \theta_v\right] \\ &\leq \Pr\left[\sum_{i=1}^{d_v} w_{u_i v} Z_0(u_i) < \alpha_0-\epsilon^-\right] \\ &\equiv \beta_0 < \exp\left(-\frac{(\epsilon^-)^2\log n}{2C}\right).\end{aligned}$$

Therefore, Error$^-$ $\leq \sum_{t=0}^{m-1}\beta_t < mn\exp(-\frac{(\epsilon^-)^2\log n}{2C})$. Hence, (B.4) holds.

We prove (B.5) in a similar way. Let γ be a positive real number such that $F(\alpha^*+\Delta) < \alpha^*+\Delta$ for all $\Delta \in (0,\gamma]$. The existence of such γ is guaranteed by assumption that $F(\alpha)$ is not tangential to the line of $y = \alpha$ at α^*. Let $\epsilon^+ = \alpha^*+\epsilon'-F(\alpha^*+\epsilon')$ where $\epsilon' = \min(\frac{\epsilon}{\lambda},\gamma)$. Note that,

$$\begin{aligned}\alpha^*+\epsilon = F(\alpha^*)+\epsilon &= F(\alpha^*)-F(\alpha^*+\epsilon')+F(\alpha^*+\epsilon')+\epsilon \\ &\geq F(\alpha^*+\epsilon')+\epsilon-\lambda\epsilon' \geq F(\alpha^*+\epsilon').\end{aligned}$$

The diffusion process stops within n iterations since at each iteration, at least one node is influenced. Therefore, for $1 \leq t \leq n$, we define $\mathcal{E}_t'(v)$ as the event that a node v with $\theta_v > \alpha^*+\epsilon'$ becomes $Z_t(v) = 1$ at the tth iteration and $X^+(v)$ as the indicator variables such that $X^+(v) = 1$

if $\theta_v < \alpha^* + \epsilon'$ and $X^+(v) = 0$ otherwise. Then, $\mathcal{E}'_t(v)$ is the event that $\theta_v > \alpha^* + \epsilon'$ and $\sum_{i=1}^{d_v} w_{u_i v} Z_{t-1}(u_i) \geq \theta_v$. Also, $\mathcal{E}'_{t-1}(v)^c$ is the event $Z_{t-1}(v) \leq X^+(v)$ for a node $v \in V$. Then, $\mathcal{E}'_0(v) = \phi$. Similar to the proof of (B.4),

$$\begin{aligned}&\Pr[\mathcal{Z}(x) \leq \alpha^* n + n\epsilon]\\ &\quad \geq \Pr[\text{All nodes whose thresholds are greater than } \alpha^*+\epsilon' \text{ are not influenced}]\\ &\quad \geq 1 - \text{Error}^+,\end{aligned}$$

where Error^+ is as follows.

$$\begin{aligned}&\Pr[\exists \text{ a node } v \text{ such that } \theta_v > \alpha^* + \epsilon' \text{and is influenced}]\\
&\leq \sum_{v\in V} \Pr[\text{A node } v \text{ becomes the first node that is influenced among nodes whose thresholds are greater than } (\alpha^*+\epsilon')]\\
&\leq \sum_{v\in V}\sum_{t=1}^{n} \Pr\left[\mathcal{E}'_t(v) \cap (\cap_{u\in V} \mathcal{E}'_{t-1}(u)^c)\right]\\
&\leq \sum_{v\in V}\sum_{t=1}^{n} \Pr\left[\mathcal{E}'_t(v) \cap (\cap_{i=1}^{d_v} \mathcal{E}'_{t-1}(u_i)^c)\right]\\
&\leq \sum_{v\in V}\sum_{t=1}^{n} \Pr\Bigg[\theta_v > \alpha^* + \epsilon', \sum_{i=1}^{d_v} w_{u_i v} Z_{t-1}(u_i) \geq \theta_v,\\
&\qquad Z_{t-1}(u_i) \leq X^+(u_i), i = 1,\ldots,d_v\Bigg]\\
&\leq \sum_{v\in V}\sum_{t=1}^{n} \Pr\left[\sum_{i=1}^{d_v} w_{u_i v} X^+(u_i) > \alpha^* + \epsilon'\right]\\
&= \sum_{v\in V}\sum_{t=1}^{n} \Pr\left[\sum_{i=1}^{d_v} w_{u_i v} X^+(u_i) > F(\alpha^* + \epsilon') + \epsilon^+\right]\\
&< n^2 \exp\left(-\frac{(\epsilon^+)^2 \log n}{2C}\right).\end{aligned}$$

The last inequality holds by Lemma 4, whose proof is almost the same as the proof of Lemma 3.

Lemma 4. *Let $X_1, \ldots, X_{d_v}$ be i.i.d. Bernoulli random variables with $E[X_1] = E[X_2] = \ldots = E[X_{d_v}] = B$ and $Y(v) = \sum_{i=1}^{d_v} w_{iv} X_i$. If $\max_i w_{iv} \leq \sqrt{\frac{C}{d_v logn}}$ for some $C > 0$, for any $\tilde{\epsilon} > 0$,*

$$P(Y(v) \geq B + \tilde{\epsilon}) < \exp\left(-\frac{\tilde{\epsilon}^2 \log n}{2C}\right).$$

Hence, we prove (B.5), which completes the proofs of Lemma 1 and Theorem 1.

Appendix C: Proof of Theorem 2

Proof of Theorem 2. Let m be any finite integer such that m and arbitrarily small $\epsilon > 0$ be given. Remind that our goal is to prove the following equations (C.1)–(C.5): For any $\delta > 0$,

$$\Pr[|\mathcal{Z}_m - \bar{\sigma}_m| < n\epsilon] = 1 - o(n^{-\delta}), \tag{C.1}$$
$$\Pr[\mathcal{Z} \leq \bar{\sigma}^* + n\epsilon] = 1 - o(n^{-\delta}). \tag{C.2}$$

Moreover, for all $v \in V$,

$$|E[Z_m(v)] - \alpha_m(v)| = o(1), \tag{C.3}$$
$$E[Z(v)] \geq \alpha^*(v) + o(1). \tag{C.4}$$

If $\lambda = \lambda_F \lambda_f < 1$,

$$\Pr[|\mathcal{Z} - \bar{\sigma}^*| < n\epsilon] = 1 - o(n^{-\delta}), \tag{C.5}$$
$$|E[Z(v)] - \alpha^*(v)| = o(1). \tag{C.6}$$

We will show that, for some ϵ^-, ϵ^+ and some constant $C > 0$,

$$\Pr[|\mathcal{Z}_m - \bar{\sigma}_m| > n\epsilon] < 2mn^{\left(1-\frac{(\epsilon^-)^2}{2C}\right)}, \tag{C.7}$$
$$\Pr[\mathcal{Z} > \bar{\sigma}^* + n\epsilon] < n^{\left(2-\frac{(\epsilon^+)^2}{2C}\right)}. \tag{C.8}$$

We now explain that (C.7) and (C.8) directly imply (C.1), (C.2) and (C.5). When m is a constant, it suffices to show that $\frac{(\epsilon^*)^2}{2C} - 2 \geq \delta$ holds for any $\delta > 0$ where $\epsilon^* = min\{\epsilon^-, \epsilon^+\}$ or $\epsilon^* = (1-\lambda)\epsilon$ if $\lambda < 1$. Then, $\frac{(\epsilon^*)^2}{2C} - 2 \geq \delta$ is equivalent to $C \leq \frac{(\epsilon^*)^2}{2(\delta+2)}$. In the course of the proof, we will show that C can be any value such that $\max_{\Psi_v} |f_v(\boldsymbol{\nu}) - f_v(\boldsymbol{\nu}')| = \sqrt{\frac{C}{d_v logn}} = o(\sqrt{\frac{1}{d_v logn}})$ for any $\Psi_v = \{\boldsymbol{\nu}, \boldsymbol{\nu}' \in \{0,1\}^{d_v} | \nu(i) \neq \nu'(i) \text{for only one } i\}$. Then, C tends to be 0 for sufficiently large n. Thus, we can choose $C \leq \frac{(\epsilon^*)^2}{2(\delta+2)}$ for any $\delta > 0$, in the limit of $n \to \infty$. Furthermore, in the proof of (C.7) and (C.8), we prove that (C.3), (C.4), and (C.6) hold.

Proof of (C.7). In order to show (C.7), we calculate $\Pr[\mathcal{Z}_m < \bar{\sigma}_m - n\epsilon]$ and then $\Pr[\mathcal{Z}_m > \bar{\sigma}_m + n\epsilon]$ in a similar way. Analogous to the proofs of Lemma 1 and Theorem 1, we define recursively $\{\epsilon_0, \epsilon_1, \ldots, \epsilon_{m-2}\}$ for a given $\epsilon = \epsilon_{m-1}$ and $\epsilon^- = \frac{\epsilon_{m-1}}{\lambda^{m-1}+\lambda^{m-2}+\ldots+\lambda^0}$ by $\epsilon_{t-1} = \frac{\epsilon_t - \epsilon^-}{\lambda}$ for $t = m-1, m-2, \ldots, 0$. Then, $\epsilon_t = (\lambda^t + \lambda^{t-1} + \ldots + \lambda^0)\epsilon^-$ and $\epsilon_0 = \epsilon^-$.

Let $Z_t(v)$ be the indicator variable representing whether a node v has adopted the information by the tth step. Let $X_t^-(v)$ be the indicator variable such that $X_t^-(v) = 1$ if $\theta_v > \alpha_t(v) - \epsilon_t$ and $X_t^-(v) = 0$ otherwise. Let $\delta_t(v) = E[F\left(f_v(X_{t-1}^-(u_1), \ldots, X_{t-1}^-(u_{d_v}))\right)]$. Then,

$$\begin{aligned}\alpha_t(v) - \delta_t(v) &= F(f_v^*(\alpha_{t-1}(u_1), \alpha_{t-1}(u_2), \ldots, \alpha_{t-1}(u_{d_v})))\\ &\quad - F(f_v^*(\alpha_{t-1}(u_1) - \epsilon_{t-1}, \ldots, \alpha_{t-1}(u_{d_v})\\ &\quad - \epsilon_{t-1}))\\ &\leq \lambda\epsilon_{t-1}\end{aligned}$$

by the definitions. Then we get

$$\begin{aligned}\alpha_t(v) - \epsilon_t &= (\alpha_t(v) - \delta_t(v)) - (\epsilon_t - \delta_t(v))\\ &\leq \delta_t(v) - (\epsilon_t - \lambda\epsilon_{t-1}) \leq \delta_t(v) - \epsilon^-.\end{aligned} \tag{C.9}$$

Similarly to the proof of Lemma 1,

$$
\begin{aligned}
&\Pr[\mathcal{Z}_m \geq \bar{\sigma}_m - n\epsilon] \\
&\quad \geq_{(c)} \Pr[\mathcal{Z}_m \geq \sum_{v\in V} \delta_t(v) - n\epsilon^-] \\
&\quad \geq_{(d)} \Pr[\mathcal{Z}_m \geq \sum_{v\in V} F(f_v(X^-_{t-1}(u_1), \ldots, X^-_{t-1}(u_{d_v})))] \\
&\quad \geq_{(e)} \Pr[\mathcal{Z}_m \geq \sum_{v\in V} X^-_{t-1}(v)] \\
&\quad \geq_{(f)} \Pr[\text{All nodes whose thresholds are less than } \\
&\quad (\alpha_{m-1}-\epsilon_{m-1}) \text{ are influenced until the } m\text{th iteration}]
\end{aligned}
$$

where (c) holds by (C.9). We can say (d) holds for sufficiently large n by standard concentration theorems such as Chernoff bounds. (e) is satisfied since the f_v's are normalized to have maximum value of 1. (f) holds because the event that $\mathcal{Z}_m \geq \sum_{v\in V} X^-_{t-1}(v)$ contains the event that $Z_t(v) = 1$ for all v's such that $X^-_{t-1}(v) = 1$.

We define $\mathcal{E}_t(v)$ as the event that a node v has a threshold $\theta_v < \alpha_t(v) - \epsilon_t$ but does not adopt the information at the $(t+1)$th step. Then $\mathcal{E}_t(v)$ is the event that $\theta_v < \alpha_t(v) - \epsilon_t$ and $f_v(Z_{t-1}(u_1), \ldots, Z_{t-1}(u_{d_v})) < \theta_v$, while $\mathcal{E}_t(v)^c$ is the event that $Z_t(v) \geq X^-_t(v)$. Now, using the union bound of probability,

$1 - \Pr$ [All nodes whose thresholds are less than $(\alpha_{m-1} - \epsilon_{m-1})$ are influenced until the mth iteration]

$$
\begin{aligned}
&= \Pr[\text{There exists a node whose threshold is less than } (\alpha_{m-1}-\epsilon_{m-1}) \text{ but which is not influenced until the } m\text{th iteration}] \\
&\leq \sum_{v\in V} \Pr[\text{A node } v \text{ has a threshold less than } (\alpha_{m-1}(v) - \epsilon_{m-1}) \text{ but which is not influenced until the } m\text{th iteration}] \\
&= \sum_{v\in V} \Pr[\mathcal{E}_{m-1}(v)] \\
&\leq \sum_{v\in V} \Pr[\mathcal{E}_0(v) \cup \mathcal{E}_1(v) \cup \ldots \cup \mathcal{E}_{m-1}(v)] \\
&= \sum_{v\in V} \left(\sum_{t=1}^{m-1} \Pr[\mathcal{E}_t(v) \cap (\cap_{u\in V} \mathcal{E}_{t-1}(u)^c)] + \Pr[\mathcal{E}_0(v)] \right) \\
&\leq \sum_{v\in V} \left(\sum_{t=1}^{m-1} \Pr[\mathcal{E}_t(v) \cap (\cap_{i=1}^{d_v} \mathcal{E}_{t-1}(u_i)^c)] + \Pr[\mathcal{E}_0(v)] \right).
\end{aligned}
$$

We then show that $\Pr[\mathcal{Z}_m > \bar{\sigma}_m + n\epsilon]$ is less than $mn\exp(-\frac{(\epsilon^-)^2 \log n}{2C})$ by proving $\Pr[\mathcal{E}_t(v) \cap (\cap_{i=1}^{d_v} \mathcal{E}_{t-1}(u_i)^c)]$ for all $t = 1, \ldots, m-1$ and $\Pr[\mathcal{E}_0(v)]$ are bounded above by $\exp(-\frac{(\epsilon^-)^2 \log n}{2C})$ by the following lemma.

Lemma 5. *Let d_v be the degree of node v. Let $X_1, \ldots, X_{d_v}$ be any Bernoulli random variables, Let $f_v(X_1, \ldots, X_{d_v})$ be the threshold function of node v where $f_v : \{0,1\}^{d_v} \to [0,1]$. When $\max_{\Psi_v} |f_v(\boldsymbol{\nu}) - f_v(\boldsymbol{\nu}')| = \sqrt{\frac{C}{d_v \, logn}}$, for $\Psi_v = \{\boldsymbol{\nu}, \boldsymbol{\nu}' \in \{0,1\}^{d_v} | \nu(i) \neq \nu'(i)$ for only one $i\}$, some $C > 0$ and all $i = 1, \ldots, d_v$, then for any given $\epsilon^-, \epsilon^+ > 0$,*

$$
\begin{aligned}
&\Pr[F(f_v(X_1, \ldots, X_{d_v})) \leq E[F(f_v(X_1, \ldots, X_{d_v}))] - \epsilon^-] \\
&\quad < \exp\left(-\frac{(\epsilon^-)^2 \log n}{2C}\right), \text{ and} \\
&\Pr[F(f_v(X_1, \ldots, X_{d_v})) \geq E[F(f_v(X_1, \ldots, X_{d_v}))] + \epsilon^+] \\
&\quad < \exp\left(-\frac{(\epsilon^+)^2 \log n}{2C}\right).
\end{aligned}
$$

Proof of Lemma 5. The proof is similar to Lemma 3. We define a sequence $Y_0(v), \ldots, Y_{d_v}(v)$, which is a martingale as follows.

$$
\begin{aligned}
Y_0(v) &= E[F(f_v(X_1, X_2, \ldots, X_{d_v}))] \\
&\vdots \\
Y_i(v) &= E[F(f_v(X_1, X_2, \ldots, X_{d_v}) | X_1, \ldots, X_i)] \\
&\vdots \\
Y_{d_v}(v) &= F(f_v(X_1, X_2, \ldots, X_{d_v})).
\end{aligned}
$$

Then, we obtain that $E[f_v(X_1, \ldots, X_{d_v}) | X_1, \ldots, X_{i+1}] - E[f_v(X_1, \ldots, X_{d_v}) | X_1, \ldots, X_i] < \max_{\Psi_v} |f_v(\boldsymbol{\nu}) - f_v(\boldsymbol{\nu}')|$ for $i = 0, \ldots, d_v - 1$. By applying Azuma's inequality, the above lemma is proved.

Applying Lemma 5, we obtain the desired upper bound for $\Pr[\mathcal{E}_t(v) \cap (\cap_{i=1}^{d_v} \mathcal{E}_{t-1}(u_i)^c)]$ for all $t = 1, \ldots, m-1$ and $\Pr[\mathcal{E}_0(v)]$ as follows. For each $t = 1, \ldots, m$,

$$
\begin{aligned}
&\Pr[\mathcal{E}_t(v) \cap (\cap_{i=1}^{d_v} \mathcal{E}_{t-1}(u_i)^c)] \\
&\quad = \Pr[\theta_v < \alpha_t(v) - \epsilon_t, F(f_v(Z_{t-1}(u_1), \ldots, Z_{t-1}(u_{d_v}))) < \theta_v, \\
&\qquad Z_{t-1}(u_i) \geq X^-_{t-1}(u_i), i = 1, \ldots, d_v] \\
&\quad \leq \Pr[F(f_v(X^-_{t-1}(u_1), \ldots, X^-_{t-1}(u_{d_v}))) < \alpha_t(v) - \epsilon_t] \\
&\quad \leq \Pr[F(f_v(X^-_{t-1}(u_1), \ldots, X^-_{t-1}(u_{d_v}))) < E[f_v(X^-_{t-1}(u_1), \\
&\qquad \ldots, X^-_{t-1}(u_{d_v}))] - \epsilon^-] \\
&\quad < \exp\left(-\frac{(\epsilon^-)^2 \log n}{2C}\right).
\end{aligned}
$$

Also we have that

$$
\begin{aligned}
&\Pr[\mathcal{E}_0(v)] \\
&\quad = \Pr[\theta_v < \alpha_0(v) - \epsilon_0, F(f_v(Z_0(u_1), \ldots, Z_0(u_{d_v}))) < \theta_v] \\
&\quad \leq \Pr[F(f_v(Z_0(u_1), \ldots, Z_0(u_{d_v}))) < \alpha_0(v) - \epsilon^-] \\
&\quad < \exp\left(-\frac{(\epsilon^-)^2 \log n}{2C}\right).
\end{aligned}
$$

Therefore, $\Pr[\mathcal{Z}_m < \bar{\sigma}_m - n\epsilon] < mn\exp(-\frac{(\epsilon^-)^2 \log n}{2C})$. Similarly, $\Pr[\mathcal{Z}_m > \bar{\sigma}_m + n\epsilon] < mn\exp(-\frac{(\epsilon^-)^2 \log n}{2C})$. Thus, $\Pr(|\mathcal{Z}_m - \bar{\sigma}_m| > n\epsilon) < 2mn^{(1-\frac{(\epsilon^-)^2}{2C})}$ and (C.7) is proved. As mentioned earlier, it implies (C.1).

Now we can say that for any finite integer m, and any $\delta > 0$,

$$\sum_{v\in V} \Pr[Z_m(v) \le \alpha_m(v) - \epsilon_m] = o\left(n^{-\delta}\right)$$

and

$$\sum_{v\in V} \Pr[Z_m(v) \ge \alpha_m(v) + \epsilon_m] = o\left(n^{-\delta}\right).$$

Then, one can check that the expectation of $Z_m(v)$ satisfies that

$$E[Z_m(v)] \ge (\alpha_m(v) - \epsilon_m)\left(1 - o\left(n^{-\delta}\right)\right) + o\left(n^{-\delta}\right)$$

and

$$E[Z_m(v)] \le (\alpha_m(v) + \epsilon_m)\left(1 - o\left(n^{-\delta}\right)\right) + o\left(n^{-\delta}\right)$$

similarly to that of Appendix B. Thus, $|E[Z_m(v)] - \alpha_m(v)| \le \epsilon_m$. Since ϵ_m is a constant fraction of ϵ and ϵ is any given positive number, $|E[Z_m(v)] - \alpha_m(v)|$ approaches zero in the limit of $\epsilon \to 0$. It means that $|E[Z_m(v)] - \alpha_m(v)| = o(1)$, as stated in (C.3).

Moreover, when $\lambda < 1$, f_v^* is a contraction map so that $\bar{\sigma}_m$ converges to $\bar{\sigma}^*$ in the limit of $m \to \infty$. Since $E[Z_m(v)]$ converges to $E[Z(v)]$, $|E[Z_m(v)] - \alpha_m(v)|$ converges to $|E[Z(v)] - \alpha^*(v)|$ as $m \to \infty$. Therefore, if $\lambda < 1$, then (C.5) and (C.6) hold by the above argument.

Proof of (C.8). We compute $\Pr[\mathcal{Z} > \bar{\sigma}^* + n\epsilon]$ to show (C.8), which proves (C.2). Let $Z(v)$ be the indicator variable representing whether a node v has adopted the information at the final step. Let $X^+(v)$ be the indicator variable such that $X^+(v) = 1$ if $\theta_v < \alpha^*(v) + \gamma$ and $X^+(v) = 0$ otherwise.

Let $\eta = \|\nabla f_v(\alpha^*(u_1), \ldots, \alpha^*(u_{d_v}))\|_\infty$, then one can check that $0 \le \eta < 1$ by the definition of the $\alpha^*(v)'s$ and since all coordinates of $\nabla f_v(\alpha^*(u_1), \ldots, \alpha^*(u_{d_v}))$ are not equal to 1. Assume the contrary, i.e., there exists a coordinate i such that the partial derivative of $f_v(\alpha^*(u_1), \ldots, \alpha^*(u_{d_v}))$ with respect to $\alpha^*(u_{d_i})$ is larger than 1. Consider that all the coordinates except i are held fixed, then we can apply the same argument used in the linear threshold model that leads to the contradiction so that the partial derivatives are less than 1. Then, we can take $\gamma \in [0, \epsilon)$ such that $|F(f_v(\alpha^*(u_1) + \Delta_1, \ldots, \alpha^*(u_{d_v}) + \Delta_{d_v})) - F(f_v(\alpha^*(u_1), \ldots, \alpha^*(u_{d_v})))| \le \eta\|\max(\Delta_1, \Delta_2, \ldots, \Delta_{d_v})\|$ when $0 \le \Delta_i \le \gamma$ for all $i = 1, \ldots, d_v$.

Let $\delta(v) = \alpha^*(v) - E[f_v(X^+(u_1), \ldots, X^+(u_{d_v}))]$. Let $\epsilon^+ = (1-\eta)\gamma$. Then,

$$\begin{aligned}\delta(v) = F(f_v^*(\alpha^*(u_1) + \gamma, \alpha^*(u_2) + \gamma, \ldots, \alpha^*(u_{d_v}) + \gamma))\\ -F(f_v^*(\alpha^*(u_1), \alpha^*(u_2), \ldots, \alpha^*(u_{d_v}))) \le \eta\gamma\end{aligned}$$

so that

$$\begin{aligned}\alpha^*(v) + \epsilon &\ge \alpha^*(v) + \gamma\\ &= (\alpha^*(v) + \delta(v)) + (\gamma - \delta(v))\\ &\ge (\alpha^*(v) + \delta(v)) + (1-\eta)\gamma\\ &= E[F(f_v(X^+(u_1), \ldots, X^+(u_{d_v})))] + \epsilon^+.\end{aligned}$$

We define $\mathcal{E}_t^+(v)$ as the event that $\theta_v > \alpha^*(v) + \gamma$ and $Z_t(v) = 1$, which is the event that $\theta_v > \alpha^*(v) + \gamma$ and $F(f_v(Z_{t-1}(u_1), \ldots, Z_{t-1}(u_{d_v}))) > \theta_v$. Then, $\mathcal{E}_t^+(v)^c$ is the event that $Z_v(t) \le X^+(v)$. Similar as before,

$\Pr[\mathcal{Z} > \bar{\sigma}^* + n\epsilon] \le \Pr$ [Any node whose thresholds are greater than $(\alpha_{m-1} - \epsilon_{m-1})$ are not influenced until the mth iteration].

Then, we have

$1 - \Pr$ [Any node whose thresholds are greater than $(\alpha_{m-1} - \epsilon_{m-1})$ are not influenced until the mth iteration]

$$\begin{aligned}&\le \sum_{v\in V} \Pr[\text{A node } v \text{ has a threshold greater than}\\ &\quad (\alpha^*(v) + \gamma) \text{ but adopts the information}]\\ &= \sum_{v\in V}\left(\sum_{t=1}^{m-1} \Pr[\mathcal{E}_t^+(v) \cap (\cap_{u\in V}\mathcal{E}_{t-1}^+(u)^c)] + \Pr[\mathcal{E}_0^+(v)]\right)\\ &\le \sum_{v\in V}\left(\sum_{t=1}^{m-1} \Pr[\mathcal{E}_t^+(v) \cap (\cap_{i=1}^{d_v}\mathcal{E}_{t-1}^+(u_i)^c)] + \Pr[\mathcal{E}_0^+(v)]\right).\end{aligned} \tag{C.10}$$

Using Lemma 5, we actually bound the probability of events described inside the summation in (C.10). First, we bound $\Pr[\mathcal{E}_t^+(v) \cap (\cap_{i=1}^{d_v}\mathcal{E}_{t-1}^+(u_i)^c)]$ for each v.

$$\begin{aligned}&\Pr[\mathcal{E}_t^+(v) \cap (\cap_{i=1}^{d_v}\mathcal{E}_{t-1}^+(u_i)^c)]\\ &\quad= \Pr[\theta_v > \alpha^*(v) + \gamma,\\ &\qquad F(f_v(Z_{t-1}(u_1), \ldots, Z_{t-1}(u_{d_v}))) > \theta_v,\\ &\qquad Z_{t-1}(u_i) \le X^+(u_i), i = 1, \ldots, d_v]\\ &\quad\le \Pr[F(f_v(X^+(u_1), \ldots, X^+(u_{d_v}))) < \alpha^*(v) + \gamma]\\ &\quad\le \Pr[F(f_v(X^+(u_1), \ldots, X^+(u_{d_v}))) <\\ &\qquad E[F(f_v(X^+(u_1), \ldots, X^+(u_{d_v})))] + \epsilon^+]\\ &\quad< \exp\left(-\frac{(\epsilon^+)^2 \log n}{2C}\right).\end{aligned}$$

In the similar way, we also bound $\Pr[\mathcal{E}_0(v)]$ for each v so that we finally bound (C.10).

$$\begin{aligned}\Pr[\mathcal{E}_0(v)] &= \Pr[\theta_v > \alpha^*(v) + \gamma, F(f_v(Z_0(u_1), \ldots,\\ &\qquad Z_0(u_{d_v}))) > \theta_v]\\ &\le \Pr[F(f_v(Z_0(u_1), Z_0(u_2), \ldots, Z_0(u_{d_v})))\\ &\qquad > \alpha^*(v) + \gamma]\\ &< \exp\left(-\frac{(\epsilon^+)^2 \log n}{2C}\right).\end{aligned}$$

Therefore, $\Pr[\mathcal{Z} > \bar{\sigma}^* + n\epsilon] < n^2 \exp\left(-\frac{(\epsilon^+)^2 \log n}{2C}\right)$. That is, (C.8) holds. As a result, (C.8) implies (C.2).

In the proof of (C.8), we proved that $\sum_{v\in V} \Pr[Z(v) > \alpha^*(v) + \gamma] = o(n^{-\delta})$ for any $\delta > 0$. Then, the expectation of $Z(v)$ satisfies that $E[Z(v)] \geq (\alpha^*(v) - \gamma)(1 - o(n^{-\delta})) + o(n^{-\delta})$. Hence, $E[Z_m(v)] \geq \alpha^*(v) + \gamma$. Since γ is less than ϵ and ϵ is any given positive number, γ tends to 0 in the limit of $\epsilon \to 0$. It implies that $E[Z(v)] \geq \alpha^*(v) + o(1)$, as stated in (C.4). Hence, Theorem 2 is proved.

References

1. J.A. Dodson., E. Muller, Management Science **24**, 1568 (1976)
2. J. Coleman, E.M. Katz, H. Menzel, Sociometry **20**, 253 (1957)
3. D.H.S. Bikhchandani, I. Welch, J. Polit. Econ. **100**, 992 (1992)
4. T.W. Valente, *Network Models of the Diffusion of Innovations* (Hampton Press, 1995)
5. F. Chierichetti, S. Lattanzi, A. Panconesi, in *Proceedings of SODA, 2010*, pp. 1657–1663
6. M. Granovetter, Am. J. Sociol. **83**, 1420 (1978)
7. R. Cohen, K. Erez, D. ben Avraham, S. Havlin, Phys. Rev. Lett. **85**, 4626 (2000)
8. R. Pastor-Satorras, A. Vespignani, Phys. Rev. Lett. **86**, 3200 (2000)
9. D.J. Watts, Proc. Natl. Acad. Sci. USA **99**, 5766 (2002)
10. D.E. Whitney, Phys. Rev. E **82**, 066110 (2010)
11. E. Valdano, L. Ferreri, C. Poletto, V. Colizza, Phys. Rev. X **5**, 021005 (2015)
12. D. Kempe, J. Kleinberg, E. Tardos, in *Proceedings of KDD, 2003*, pp. 137–146
13. M.E.J. Newman, Phys. Rev. E **66**, 016128 (2002)
14. T.C. Schelling, J. Conflict Resolut. **17**, 381 (1973)
15. P. Domingos, M. Richardson, Mining Knowledge-Sharing Sites for Viral Marketing, in *Proceedings of KDD, 2002*, pp. 61–70
16. L. Blume, D. Easley, J. Kleinberg, R. Kleinberg, E. Tardos, in *Proceedings of FOCS, 2011*, pp. 393–402
17. D.M. Romero, Btranscript. Meeder, J. Kleinberg, in *Proceedings of WWW, 2011*, pp. 695–704
18. J. Leskovec, A. Krause, C. Guestrin, C. Faloutsos, J. VanBriesen, N. Glance, in *Proceedings of KDD, 2007*, pp. 420–429
19. W. Chen, C. Wang, Y. Wang, in *Proceedings of KDD, 2010*, pp. 1029–1038
20. A. Goyal, W. Lu, L.V.S. Lakshmanan, in *Proceedings of WWW, 2011*, pp. 47–48
21. W. Chen, Y. Yuan, L. Zhang, in *Proceedings of ICDM, 2010*, pp. 88–97
22. A. Goyal, W. Lu, L.V.S. Lakshmanan, in *Proceedings of ICDM, 2011*, pp. 211–220
23. P. Holme, B.J. Kim, C.N. Yoon, S.K. Han, Phys. Rev. E **65**, 056109 (2002)
24. P. Gai, S. Kapadia, Proc. R. Soc. A **466**, 2401 (2010)
25. S.V. Buldyrev, R. Parshani, G. Paul, H.E. Stanley, S. Havlin, Nature **464**, 1025 (2010)
26. D.S. Callaway, M.E.J. Newman, S.H. Strogatz, D.J. Watts, Phys. Rev. Lett. **85**, 5468 (2000)
27. A. Ganesh, L. Massoulie, D. Towsley, in *Proceedings of INFOCOM, 2005*, pp. 1455–1466
28. G. Ergun, Physica A **308**, 483 (2002)
29. S. Wasserman, K. Faust, *Social Network Analysis* (Cambridge University Press, 1994)
30. M. Szell, R. Lambiotte, S. Thurner, Proc. Natl. Acad. Sci. USA **107**, 13636 (2010)
31. M. Richardson, R. Agrawal, P. Domingos, in *Proceedings of ISWC, 2003*, pp. 351–368
32. L.A. Adamic, N. Glance, in *Proceedings of LinkKDD, 2005*, pp. 36–43
33. K.T.D. Eames, M.J. Keeling, Proc. Natl. Acad. Sci. USA **99**, 13330 (2002)
34. H. Zhou, R. Lipowsky, Proc. Natl. Acad. Sci. USA **102**, 10052 (2005)
35. J.P. Gleeson, Phys. Rev. Lett. **107**, 068701 (2011)
36. J.P. Gleeson, Phys. Rev. X **3**, 021004 (2013)
37. S. Lim, K. Jung, J.C.S. Lui, ACM SIGMETRICS Perform. Eval. Rev. **41**, 31 (2013)
38. Y. Lin, J.C.S. Lui, K. Jung, S. Lim, J. Complex Networks **2**, 431 (2014)
39. M.L. Katz, C. Shapiro, Am. Econ. Rev. **75**, 424 (1985)