ASYNCHRONOUS SEMI-ANONYMOUS DYNAMICS OVER LARGE-SCALE NETWORKS

C. RAVAZZI*, G. COMO*†, M. GARETTO‡, E. LEONARDI*§, AND A. TARABLE*

Abstract. We analyze a class of stochastic processes, referred to as *asynchronous* and *semi-anonymous dynamics* (ASD), over directed labeled random networks. These processes are a natural tool to describe general best-response and noisy best-response dynamics in network games where each agent, at random times governed by independent Poisson clocks, can choose among a *finite set of actions*. The payoff is determined by the relative frequency of the different actions among neighbors, while being *independent of the specific identities of neighbors*.

Using a local mean-field approach, we prove rigorously that, under certain conditions on the network and initial node configuration, the evolution of ASD can be approximated, in the large-scale limit, by the solution of a system of non-linear ordinary differential equations. Our framework is very general and applies to a large class of graph ensembles for which the typical random graph is locally tree-like. In particular, we focus on labeled configuration-model random graphs, a generalization of the traditional configuration model which allows different classes of nodes to be mixed together in the network, permitting us, for example, to incorporate a community structure in the system. Our analysis also applies to configuration-model graphs having a power-law degree distribution, an essential feature of many real systems. To demonstrate the power and flexibility of our framework, we consider several examples of dynamics belonging to our class of stochastic processes. Moreover, we illustrate by simulation the applicability of our analysis to realistic scenarios by running our example dynamics over a real social network graph.

Key words. Asynchronous semi-anonymous dynamics, Networks games, Best-response dynamics, Evolutionary game theory, Linear threshold models, Local Weak Convergence.

AMS subject classifications. 68Q25, 68R10, 68U05

1. Introduction.

Many complex systems arising in different domains exhibit cascading phenomena that spread through networks of local interactions. Examples of such cascades include, but are not limited to, infrastructure failures [39], adoption of innovations, conventions and technologies [33, 36, 29], diffusion of beliefs, opinions, fake news [37], memes, and the like [38]. These phenomena can have profound effects on politics [14], social norms [4], financial networks [20], marketing campaigns [21].

The standard mathematical approach to modeling cascading processes is to consider a graph (finite or infinite) in which nodes stand for individuals that can be in one of several (discrete or continuous) states, and edges (directed or undirected, possibly weighted) represent interactions with neighboring nodes. Individuals are supposed to repeatedly update their state over (discrete or continuous) time, depending on the current state of their neighbors [35, 8, 17].

Simple epidemic models in which nodes can change their state as consequence of a single contact with a neighboring node [31] turn out to be too simplistic to describe systems in which individuals tend to react to the joint states of their neighbors. To represent such combined effect, one of the most commonly used models in the literature is the linear threshold model, originally introduced by Granovetter [20] and widely investigated in several variants [40, 15]. The general idea behind such models

*National Research Council of Italy (IEIIT-CNR), c/o Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129, Torino (Italy)

†Department of Mathematical Sciences "Giuseppe Luigi Lagrange", Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129, Torino (Italy)

‡Computer Science Department, University of Turin, Via Verdi 8, 10124 Torino, (Italy)

§Department of Electronics and Telecommunications (DET), Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129, Torino (Italy)

is to assume that a node adopts a given state if the fraction of neighbors (possibly weighted by edges) currently adopting that state exceeds a certain threshold. More in general, researchers have considered so-called networked coordinated games, in which nodes adopt the best-response (according to some payoff matrix) in reaction to the strategies adopted by neighbors [10].

Some fundamental distinctions in this wide class of models are the following. In progressive processes, state transitions are irreversible: once a node joins a given state, it keeps such state indefinitely, irrespective of what happens to neighbors [22, 11, 1, 27, 4]. In non-progressive processes transitions are, instead, reversible, since nodes still remain under the influence of their neighbors after the adoption of a given state [2, 30, 34]. Another crucial distinction concerns the update rule of the nodes: does the future state of a node also depend on the current state of the node, in addition to the states of neighbors, or is it uniquely determined by the neighbors? Indeed the above distinctions, combined to the nature of edges (i.e., directed or undirected), lead to models of widely different nature and analytical tractability.

In this paper we analyze a class of cascading processes, referred to as Asynchronous and Semi-anonymous Dynamics (ASD), with the following characteristics: i) nodes update their (discrete) state over continuous time, according to independent Poisson clocks, following an arbitrary rule that depends on the number (or relative fraction) of neighbors in each possible state, but not on the specific identities of neighbors (which are order-independent); ii) edges are directed; iii) state transitions are reversible (non-progressive model).

In this paper we extends previous work in [34], where authors analyze a two-state, deterministic linear threshold model with synchronous node update over a bounded degree graph, by adopting a mean-field approach. Here, we seek to understand how this approach can be pushed to its greatest generality, extending the class of networks and underlying node dynamics to which it can be rigorously applied. We mention, however, that in [34] authors consider also the progressive variant of their model, while here we focus only on the non-progressive model. The interested reader can refer to [18] for new methods to approximate the dynamics on large networks sampled from a graphon. Although this framework is flexible, it works well on dense network formation models, such as stochastic block models, but not on sparse networks.

One stream of related work [7, 30, 24] analyzes the possible equilibria of a binary-decision game in the case of undirected graphs and synchronous update. Similarly to these works, we study a game where players'payoffs depend on the actions taken by their neighbors in the network but not on the specific identities of these neighbors.

Another huge stream of related work is concerned with the algorithmic aspects of influence maximization [22, 28], where the goal is to find the initial node configuration that maximizes the final size of the cascade. In contrast to such stream of work, here we assume the initial node configuration to be randomly selected according to a given node statistics.

1.1. Overview of main results and paper outline. Our analysis relies on a rigorous proof of a local mean-field approximation result, as we show that the aggregate behavior of ASDs on a large class of locally tree-like random graph ensembles is close to the solution of a finite-dimensional (nonlinear) ODE. Specifically, our results show that the approximation error vanishes in the large-scale limit.

While our results apply over time horizons that remain constant or grow at most logarithmically with the network size, they do have implications to the behavior of the system in the stationary regime. In particular, when combined with, e.g., [6], our

results imply that the weak limit of every converging sequence of stationary probability distributions is an invariant distribution of the ODE. By Poincaré's Recurrence Theorem, the support of all such invariant measures is included in the closure of the recurrent set of the ODE. In particular, when the local mean-field ODE admits a globally attractive equilibrium point x^*, this implies that every sequence of stationary probability distributions for the ODE concentrates on x^* in the large-scale limit.

In Section 2 we formally introduce asynchronous semi-anonymous dynamical processes (ASD) over directed random graphs, presenting some concrete example of ASD. In Section 3 we provide the complete analysis of ASD over a labeled branching process, i.e., an infinite ensemble of labeled graphs with a rooted tree structure. In this case it can be easily shown that the evolution of the fraction of nodes in a given state indeed corresponds to the solution of some ordinary differential equations (see Proposition 2).

In Section 4 we turn our attention to general graph ensembles. The core message is the following: if the graph exhibits a local tree structure, then the analysis on a suitably chosen labeled branching process provides a good approximation of the expected fraction of nodes in a given state (see Proposition 3). A property of Local Weak Convergence is the key feature that provides this link and formalizes the idea that, for large n, the local structure of the graph near a vertex chosen uniformly at random is approximately a branching process. Finally, the analysis of the concentration around the expectation allows us to derive the accuracy of the above approximation.

We will then focus in Section 5 on the labeled configuration-model, a general mix of heterogeneous nodes with class-specific node statistics. This family of graphs is generalization of the traditional configuration model (CM) and allows different classes of nodes to be mixed together in the network, permitting us, for example, to incorporate a community structure in the system. We will explore conditions for Local Weak Convergence (see Theorem 4) and the concentration of the ASD evolution around the expectation (see Theorem 5). In Section 5.3 we will show that the sequence of node degrees is allowed to follow a power-law distribution scaling with the network size. This is particularly important for applications to social network graphs. As a second example, in Section 5.4 we will consider a labeled configuration model with a community structure, which is another fundamental feature found in many real systems. Indeed, by considering as label of a node its membership to a given community, we can represent graphs with a general distribution of in/out degrees among nodes belonging to the same or different communities. This allows us to describe, for example, "assortative" graphs, in which intra-community edges are denser than inter-community edges.

In Section 6 we analytically derive some interesting properties of the ODEs describing the temporal evolution of the system for each of the examples of ASD introduced in Section 2. In particular, we present a detailed analysis on stability of the equilibrium points and discuss the link between the mean-field ODE and the stationary regime. Section 7 presents numerical results of ASD previously introduced in large but finite networks, considering both synthetically generated graphs and real-world social networks. More precisely, we compare the solutions of the differential equations derived through mean field approximation with results obtained by running Monte-Carlo simulations.

Section 8 summarizes our contribution and collects some concluding remarks. Appendix A, B, C, and D, containing intermediate results and results and some of the more technical proofs can be found in the Supplementary materials.

1.2. General notation. Throughout this paper, we use the following notational conventions. Let $\mathbb{N}$, $\mathbb{Z}_+$, $\mathbb{R}$ be the set of natural, non-negative integers and real numbers, respectively. Given $n \in \mathbb{N}$ we use the notation $[n] = \{1, \dots, n\}$.The symbol $|\cdot|$ denotes the absolute value if applied to a scalar value and the cardinality if applied to a set. We denote the indicator function of set A with the notation $\mathbf{1}_A$. Given a finite set $\mathcal{V}$, $\mathbb{R}^{\mathcal{V}}$ denotes the space of real vectors with components labelled by elements of $\mathcal{V}$. If $x \in \mathbb{R}^n$, we denote the ℓ-th entry by x_ℓ and $x_{[\ell]}$ is the projection of x on the sub-space generated by the first ℓ elements, i.e. $x_{[\ell]} = (x_1, x_2, \dots, x_\ell) \in \mathbb{R}^\ell$.

This paper makes frequent use of Landau symbols. The notation "$f(N) = O(g(N))$ when $N \to \infty$" means that positive constants c and N_0 exist, so that $f(N) \leq cg(N)$ for all $N > N_0$. The expression "$f(N) = o(g(N))$ when $N \to \infty$" means that $\lim_{N\to\infty} \frac{f(N)}{g(N)} = 0$.

A labeled directed multigraph is a 6-ple $(\mathcal{V}, \mathcal{E}, \mathcal{A}, \lambda, \sigma, \tau)$, where $\mathcal{V}$, $\mathcal{E}$, and $\mathcal{A}$ are the sets of nodes, links, and labels, respectively, all finite; $\lambda : \mathcal{V} \to \mathcal{A}$ is the map giving the label $\lambda(v)$ of a node $v \in \mathcal{V}$; and $\sigma, \tau : \mathcal{E} \to \mathcal{V}$ are the maps giving the tail node $\sigma(e)$ and head node $\tau(e)$ of a link $e \in \mathcal{E}$ so that e is directed from $\sigma(e)$ to $\tau(e)$. The set of in-neighbors and out-neighbors of a node $v \in \mathcal{V}$ are defined as $\mathcal{N}_v^- = \{w \in \mathcal{V} \setminus \{v\} : (w, v) \in \mathcal{E}\}$ and $\mathcal{N}_v^+ = \{w \in \mathcal{V} \setminus \{v\} : (v, w) \in \mathcal{E}\}$, respectively, and the corresponding in-degree and out-degree as $d_v = |\mathcal{N}_v^-|$ and $k_v = |\mathcal{N}_v^+|$. We define its out-degree vector $\mathbf{k}_v \in \mathbb{Z}_+^{\mathcal{A}}$ as the vector whose component $a \in \mathcal{A}$ represents the number of out-neighbors of v belonging to class a. Similarly, we define for node v the in-degree vector $\mathbf{d}_v \in \mathbb{Z}_+^{\mathcal{A}}$.

A path from a vertex $u \in \mathcal{V}$ to a vertex $v \in \mathcal{V}$ (i.e. a path $u \to v$) is a finite sequence of edges $(u_i, v_i)_{1 \leq i \leq L}$ with $u_1 = u$, $v_L = v$, $v_i = u_{i+1}$. If there is at least a path from u to v, we say that u is connected to v, and the graph distance from u to v is then defined as the minimum length of a path from u to v. If all (ordered) vertex pairs are connected, the graph $\mathcal{G}$ is said strongly connected. If, instead, for any pair (u, v) either a path $u \to v$ or $v \to u$ exists, we say that the graph is weakly connected. We define a simple path as a path along which all vertices are distinct. A directed tree is a weakly connected graph in which no more than one path exists between every pair of vertices (u, v).

2. Asynchronous semi-anonymous dynamics.

2.1. Mathematical model. Let us consider a finite population of n agents interacting in a connected network, which we map onto the nodes of a labeled directed multigraph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{A}, \sigma, \tau, \lambda)$, whereby a link $e \in \mathcal{E}$ represents a direct influence of its head node $\tau(e)$ on its tail node $\sigma(e)$ and each class of nodes $\mathcal{V}_a = \{v \in \mathcal{V} : \lambda(v) = a\}$, for $a \in \mathcal{A}$ may have a different behavior, thus allowing to account for heterogeneity.

Let each agent $v \in \mathcal{V}$ be endowed with a time-varying state $Z_v(t)$ taking values from a finite set $\mathcal{X}$ for every $t \geq 0$. We shall denote the vector of all agents' states by $\mathbf{Z}(t) = (Z_v(t))_{v\in\mathcal{V}}$ and refer to it as the network configuration at time t. We shall consider *asynchronous* and *semi-anonymous dynamics* (ASD) where the state of each agent is updated at random activation times by choosing a new state in response to the state of its out-neighbors, according to a conditional probability distribution that is invariant with respect to permutations of such out-neighbors.

Formally, let $\boldsymbol{Z}(t)$ be a continuous-time Markov chain with finite state space equal to the set of configurations $\mathcal{Z} = \mathcal{X}^{\mathcal{V}}$ and the structure illustrated below.

DEFINITION 1 (Asynchronous semi-anonymous dynamics). *Let* $\mathcal{P} = \{\theta \in \mathbb{R}_+^{\mathcal{X}} :$

$\mathbf{1}'\theta = 1\}$ *be the simplex of probability vectors over* $\mathcal{X}$*. For every label* $a \in \mathcal{A}$*, let*

$$\Theta^{(a)} : \mathbb{Z}_+^{\mathcal{A}\times\mathcal{X}} \to \mathcal{P} \tag{1}$$

be a stochastic kernel, and, for every node $v \in \mathcal{V}$*, let* $\Upsilon_v^{\mathcal{G}} : \mathcal{Z} \to \mathbb{Z}_+^{\mathcal{A}\times\mathcal{X}}$ *be defined by*

$$\big(\Upsilon_v^{\mathcal{G}}(\boldsymbol{z})\big)_{ax} = \big|\big\{e \in \mathcal{E} :\, \sigma(e) = v, \lambda(\tau(e)) = a, z_{\tau(e)} = x\big\}\big| \,, \quad a \in \mathcal{A},\, x \in \mathcal{X}\,.$$

Then, $\boldsymbol{Z}(t)$ *evolves as a continuous-time Markov chain on* $\mathcal{Z}$ *with transition rates:*

$$\Lambda_{\boldsymbol{z},\boldsymbol{z}^+} = \begin{cases} \gamma\, \Theta_{z_v^+}^{(\lambda(v))}(\Upsilon_v^{\mathcal{G}}(\boldsymbol{z})) & \text{if} \quad \boldsymbol{z} \text{ and } \boldsymbol{z}^+ \text{ differ in the } v\text{-th entry only} \\ 0 & \text{if} \quad \boldsymbol{z} \text{ and } \boldsymbol{z}^+ \text{ differ in more than one entry} \end{cases}$$

where γ *denotes the Poisson rate at which node* v *updates its state.*

The formulation above in Definition 1 is very general. Some remarks are in order. Classes can describe heterogeneous nodes in a variety of ways. In our examples, we will consider the following three cases: i) classes describing different update rules of the nodes; ii) classes describing nodes with different degree distributions; iii) classes describing node membership to different 'communities'. In the most general scenario, a class might represent nodes belonging to a specific community, with a given update rule and a particular degree distribution.

We emphasize that the new state of an agent, when it gets updated, does not need to be a deterministic function of its neighborhood. Indeed, we explicitly allow for a stochastic rule of adopting a certain state. This allows us to model noisy or mixed-strategy best-response dynamics in networked games.

Our main interest in this paper is to track the evolution of some macroscopic features, e.g., the evolution of the fraction of nodes belonging to a specific class that are in a given state at time t. We will demonstrate that a mean-field approximation can yield insight into this analysis for a large class of random networks.

2.2. Examples of ASD dynamics. To clarify the general formulation introduced above, we provide three examples of ASD dynamics that will later be studied in more details in our numerical illustration section (see Section 7). Since in our examples the node update rule depends only on the total number of neighbors in a given state, and not on their label, we simplify the general notation introduced before and define:

$$\xi_x(\boldsymbol{z}) = \sum_{a\in\mathcal{A}} \big(\Upsilon_v^{\mathcal{G}}(\boldsymbol{z})\big)_{ax} \,, \quad x \in \mathcal{X}\,. \tag{2}$$

The explicit dependence on the Markov-chain state $\boldsymbol{z}$ will be omitted in the following, whenever possible.

2.2.1. Ternary Linear Threshold Model (TLTM). Let $\mathcal{X} = \{-1, 0, 1\}$ be the set of admissible states and let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{A}, \lambda, \sigma, \tau)$ be a labeled multigraph. The label $a_v = (a_v^+, a_v^-)$ of node v determines two given (in general, asymmetric) thresholds a_v^+, a_v^-, which trigger the transition to state 1 and -1, respectively. Specifically, when activated, the update of node v is given by

$$Z_v(t) = \begin{cases} 1 & \text{if } \sum_{j\in\mathcal{N}_v^+} Z_j(t^-) \geq a_v^+ \\ 0 & \text{if } \sum_{j\in\mathcal{N}_v^+} Z_j(t^-) \in (-a_v^-, a_v^+) \\ -1 & \text{if } \sum_{j\in\mathcal{N}_v^+} Z_j(t^-) \leq -a_v^- \end{cases}$$

where $Z_v(t^-) = \lim_{x\uparrow t} Z_v(x)$. The above rule can be encoded in our general formulation by considering the functions:

$$\Theta_1^{(a)}(\xi_1,\xi_{-1},\xi_0) = \mathbf{1}_{\{\xi_1-\xi_{-1}\geq a^+\}},\ \Theta_0^{(a)}(\xi_1,\xi_{-1},\xi_0) = \mathbf{1}_{\{\xi_1-\xi_{-1}\in(-a^-,a^+)\}}$$
$$\Theta_{-1}^{(a)}(\xi_1,\xi_{-1},\xi_0) = \mathbf{1}_{\{\xi_1-\xi_{-1}\leq -a^-\}}$$

which depend only on the numbers ξ_1, ξ_{-1} and ξ_0 of out-neighbors in state 1, -1 and 0, respectively.

2.2.2. Binary Response with Coordinating and Anti-coordinating agents (BRCA). Inspired by the model in [32], we consider a network game where each agent can choose between two actions in $\mathcal{X} = \{-1, 1\}$. The network consists of two classes of nodes, i.e. $\mathcal{A} = \{+, -\}$ and $\mathcal{V} = \mathcal{V}_+ \cup \mathcal{V}_-$. We assume that $\mathcal{V}_+$ and $\mathcal{V}_-$ represent agents following the majority (i.e., coordinating) or the minority (i.e., anti-coordinating) of their out-neighbors, respectively.

Specifically, an agent is updated according to the following rule: if $v \in \mathcal{V}_+$ then $\forall v \in \mathcal{V}_+$

$$Z_v(t) = \begin{cases} 1 & \text{if } \sum_{j\in\mathcal{N}_v^+} Z_j(t^-) > 0 \\ -1 & \text{if } \sum_{j\in\mathcal{N}_v^+} Z_j(t^-) < 0 \\ \pm 1 & \text{if } \sum_{j\in\mathcal{N}_v^+} Z_j(t^-) = 0 \end{cases},$$

and $\forall v \in \mathcal{V}_-$

$$Z_v(t) = \begin{cases} 1 & \text{if } \sum_{j\in\mathcal{N}_v^+} Z_j(t^-) < 0 \\ -1 & \text{if } \sum_{j\in\mathcal{N}_v^+} Z_j(t^-) > 0 \\ \pm 1 & \text{if } \sum_{j\in\mathcal{N}_v^+} Z_j(t^-) = 0. \end{cases}$$

In essence, when a node is updated, it counts the number of neighbors in state -1 and 1, and adopts the state of the majority of its neighbors if $v \in \mathcal{V}_+$, or it adopts the state of the minority of its neighbors if $v \in \mathcal{V}_-$. In the case of a tie, it chooses uniformly at random between states 1 and -1.

The above rule corresponds in our general framework to the functions

$$\Theta_1^{(+)}(\xi_1,\xi_{-1}) = \mathbf{1}_{\{\xi_1>\xi_{-1}\}} + \frac{1}{2}\mathbf{1}_{\{\xi_1=\xi_{-1}\}},\ \Theta_{-1}^{(+)}(\xi_1,\xi_{-1}) = 1 - \Theta_1^{(+)}(\xi_1,\xi_{-1})$$
$$\Theta_1^{(-)}(\xi_1,\xi_{-1}) = \mathbf{1}_{\{\xi_1<\xi_{-1}\}} + \frac{1}{2}\mathbf{1}_{\{\xi_1=\xi_{-1}\}},\ \Theta_{-1}^{(-)}(\xi_1,\xi_{-1}) = 1 - \Theta_1^{(-)}(\xi_1,\xi_{-1})$$

which depend only on the number ξ_1 and ξ_{-1} of neighboors in state 1 and -1 respectively.

2.2.3. Evolutionary Roshambo Game (ERG). In this example, the best response of an agent follows the same rationale of the popular rock-paper-scissors game. Specifically, we assume that nodes have three possible states, i.e. $\mathcal{X} = \{R, P, S\}$. When an agent is updated, it performs the following computation:

i) for each out-neighbor, it determines its best pairwise response according to the two-player game

φ	R	P	S
R	b	0	c
P	c	b	0
S	0	c	b

with $c > b$, i.e. R wins over S, S wins over P, P wins over R.

ii) The new state of the agent is selected so as to maximize the sum of payoffs provided by pairwise interactions:

$$Z_v(t) = \underset{\omega \in \mathcal{X}}{\operatorname{argmax}} \sum_{j \in \mathcal{N}_v^+} \varphi(\omega, Z_j(t^-)).$$

When the maximizing set is not unique, the new state is selected uniformly at random among the maximizing alternatives.

In this case we have (ignoring ties for simplicity)

$$\Theta^{(R)}(\xi_R, \xi_P, \xi_S) = \mathbf{1}_{\{b\xi_R + c\xi_S > \max\{c\xi_R + b\xi_P, c\xi_P + b\xi_S\}\}}$$
$$\Theta^{(P)}(\xi_R, \xi_P, \xi_S) = \mathbf{1}_{\{c\xi_R + b\xi_P > \max\{b\xi_R + c\xi_S, c\xi_P + b\xi_S\}\}}$$
$$\Theta^{(S)}(\xi_R, \xi_P, \xi_S) = \mathbf{1}_{\{c\xi_P + b\xi_S > \max\{b\xi_R + c\xi_S, c\xi_R + b\xi_P\}\}}.$$

3. ASD on the labeled branching process. In this section we consider a labeled branching process, i.e. a particular ensemble of infinite labeled directed graphs with rooted tree structure, and then analyze ASD on it. As already said, the reason why we introduce this special graph is that the analysis of ASD on it provides fundamental hints for the analysis of ASD on a general locally tree-like ensemble of graphs.

More precisely, we will consider a labeled branching process completely described by probabilities distributions $p_{\mathbf{k},a} = p_{\mathbf{k}|a} p_a$ and $q^b_{\mathbf{k}|a}$. The first is the joint probability distribution that characterizes the root, i.e. the probability that the root has label $a \in \mathcal{A}$ and out-degree vector $\mathbf{k} \in \mathbb{Z}_+^{\mathcal{A}}$. We recall that the component k_b represents the number of out-neighbors belonging to class $b \in \mathcal{A}$. The latter is the vectorial out-degree distribution for a non-root node with label a, whose parent has label b. In next sections we will show that probability distributions $p_{\mathbf{k},a}$ and $q^b_{\mathbf{k}|a}$ specifying the "approximating" labeled branching process, will be chosen so to exactly match statistics' of the network under investigation. For this reason, we inform the reader that the same notation will be adopted to denote statistics on both the network and the associated labelled branching process.

3.1. Labeled branching process. Recall that in our notation $\mathbf{k}_v \in \mathbb{Z}_+^{\mathcal{A}}$ denotes the the out-degree vector of vertex v, whose component $a \in \mathcal{A}$ represents the number of out-neighbors of v belonging to class $a \in \mathcal{A}$.

We will call labeled branching process $\mathcal{T}$ with node set $\mathcal{V} = \{v_0, v_1, ...\}$ and label set $\mathcal{A}$ the rooted tree built through the following procedure:

- Step 0: Start with a root node v_0 and assign to it a random label $A_0 \in \mathcal{A}$ and a random out-degree vector $\mathbf{K}^{(0)} \in \mathbb{Z}_+^{\mathcal{A}}$ with joint probability distribution

$$\mathbb{P}(A_0 = a, \mathbf{K}^{(0)} = \mathbf{k}) = p_{\mathbf{k},a}\,.$$

For every $a \in \mathcal{A}$, add $\mathbf{K}_a^{(0)}$ out-edges with label (A_0, a) to the root v_0 and declare all these edges active. Note that an edge label is defined as the ordered pair of the labels associated to adjacent nodes.

Then, for $h = 1, 2, \ldots$

- Step h: If there are no active edges, stop. Otherwise, take any active edge e, let (a, b) be its label and declare the edge inactive. Assign to edge e a head node $\tau(e) = v_h$ with label $\lambda(v_h) = b$ and generate a random vector $\mathbf{K}^{(h)} = \mathbf{k}$ in $\mathbb{Z}_+^{\mathcal{A}}$ with conditional probability distribution $q^a_{\mathbf{k}|b}$, then for every label $c \in \mathcal{A}$ add $\mathbf{K}_c^{(h)}$ new active outgoing edges to v_h with label (b, c).

3.2. Ordinary differential equations of ASD. Let us now consider the ASD process over the graph $\mathcal{T}$ built above. In the following matrix notation, vectors are meant to be column vectors, unless otherwise specified.

PROPOSITION 2. *Let $\boldsymbol{Z}(t)$, for $t \geq 0$, be the state vector of the ASD on $\mathcal{T}$. Then, for every fixed time $t \geq 0$, the following facts hold*

1. *For every $i \in \mathcal{V}$, the states $\{Z_{\tau(e)}(t) | \, e \in \mathcal{E} : \sigma(e) = i\}$ of the offsprings j of i in $\mathcal{T}$ are independent and identically distributed random variables with $\zeta_{\omega|a,b}(t) = \mathbb{P}(Z_j(t) = \omega \mid A_j = a, A_i = b)$, $\omega \in \mathcal{X}$, $a, b \in \mathcal{A}$ satisfying*

$$\frac{\mathrm{d}\zeta_{\omega|a,b}(t)}{\mathrm{d}t} = \gamma \left(\phi_{\omega|a,b}(\boldsymbol{\zeta}(t)) - \zeta_{\omega|a,b}(t) \right), \ \phi_{\omega|a,b}(\boldsymbol{\zeta}) = \sum_{\boldsymbol{k} \in \mathbb{Z}_+^{\mathcal{A}}} \varphi_\omega^{(\boldsymbol{k},a)}(\boldsymbol{\zeta}) q_{\boldsymbol{k}|a}^b \tag{3}$$

and

$$\varphi_\omega^{(\boldsymbol{k},a)}(\boldsymbol{\zeta}) = \sum_{\substack{\boldsymbol{\xi} \in \mathbb{Z}_+^{\mathcal{A} \times \mathcal{X}}: \\ \boldsymbol{\xi} \mathbf{1} = \boldsymbol{k}}} \Theta_\omega^{(a)}(\boldsymbol{\xi}) \binom{\boldsymbol{k}}{\boldsymbol{\xi}} \prod_{c \in \mathcal{A}} \prod_{g \in \mathcal{X}} [\zeta_{g|c,a}]^{\xi_{cg}}, \quad \binom{\boldsymbol{k}}{\boldsymbol{\xi}} = \prod_c \binom{k_c}{\boldsymbol{\xi}_c},$$

where $\boldsymbol{\xi}_c$ is the c-th row of matrix $\boldsymbol{\xi}$ and ξ_{cg} denotes the (c, g)-th element of matrix $\boldsymbol{\xi}$.

2. *The state $Z_{v_0}(t)$ of the root node v_0 is a random variable with $y_{\omega|a}(t) = \mathbb{P}(Z_{v_0} = \omega \mid A_0 = a)$ satisfying*

$$\frac{\mathrm{d}y_{\omega|a}(t)}{\mathrm{d}t} = \gamma \left(\psi_{\omega|a}(\boldsymbol{\zeta}(t)) - y_{\omega|a}(t) \right), \quad \psi_{\omega|a}(\boldsymbol{\zeta}) = \sum_{\boldsymbol{k} \in \mathbb{Z}_+^{\mathcal{A}}} \varphi_\omega^{(\boldsymbol{k},a)}(\boldsymbol{\zeta}) p_{\boldsymbol{k}|a} \tag{4}$$

Proof. 1. Let v_0 be the root of $\mathcal{T}$. Then, For every $i \in \mathcal{V}$, the states $Z_j(t) : (i, j) \in \mathcal{E}$ of the offsprings of v_i in $\mathcal{T}$ are independent and identically distributed Bernoulli random variables. Define $\zeta_{\omega|a,b}(t) = \mathbb{P}[Z_j(t) = \omega \mid A_j = a, A_i = b]$, $j \in \mathcal{V} \setminus \{v_0\}$, where v_i is the father of v_j, we have

$$\begin{aligned}
&\zeta_{\omega|a,b}(t + \Delta t) \\
&= (\gamma \Delta t + o(\Delta t)) \sum_{\mathbf{k} \in \mathbb{Z}_+^{\mathcal{A}}} \varphi_\omega^{(\mathbf{k},a)}(\boldsymbol{\zeta}(t)) q_{\mathbf{k}|a}^b + (1 - \gamma \Delta t + o(\Delta t)) \, \zeta_{\omega|a,b}(t) + o(\Delta t) \\
&= (\gamma \Delta t + o(\Delta t)) \, \phi_{\omega|a,b}(\boldsymbol{\zeta}(t)) + (1 - \gamma \Delta t + o(\Delta t)) \, \zeta_{\omega|a,b}(t),
\end{aligned}$$

from which we conclude

$$\frac{\mathrm{d}\zeta_{\omega|a,b}(t)}{\mathrm{d}t} = \lim_{\Delta t \to 0} \frac{\zeta_{\omega|a,b}(t + \Delta t) - \zeta_{\omega|a,b}(t)}{\Delta t} = \gamma(\phi_{\omega|a,b}(\boldsymbol{\zeta}(t)) - \zeta_{\omega|a,b}(t)). \ \square$$

2. Define $y(t) = \mathbb{P}[Z_{v_0}(t) = \omega \mid A_0 = a]$, then with the same arguments we have

$$\frac{\mathrm{d}y_{\omega|a}}{\mathrm{d}t} = \gamma(\psi_{\omega|a}(\boldsymbol{\zeta}(t)) - y_{\omega|a}(t)), \quad \psi_{\omega|a}(\boldsymbol{\zeta}(t)) = \sum_{\mathbf{k} \in \mathbb{Z}_+^{\mathcal{A}}} \varphi_\omega^{(\mathbf{k},a)}(\boldsymbol{\zeta}(t)) p_{\mathbf{k}|a}.$$

REMARK 1. *Whenever labels of neighbor nodes are independent, $p_{\boldsymbol{k}|a}$ and $q_{\boldsymbol{k}|a}^b$ depend on a, b only through $k = \sum_i k_i$, and things become simpler. Indeed, if we*

define $\zeta_\omega(t) = \mathbb{P}[Z_j(t) = \omega] = \sum_{a,b\in\mathcal{A}} \zeta_{\omega|a,b}(t) p_a p_b$*, then, it can be easily shown that, similarly to* (3)*, we can derive an ODE for* $\zeta_\omega(t)$ *in the form*

$$\frac{\mathrm{d}\zeta_\omega(t)}{\mathrm{d}t} = \gamma\left(\phi_\omega(\boldsymbol{\zeta}(t)) - \zeta_\omega(t)\right), \tag{5}$$

where

$$\phi_\omega(\boldsymbol{\zeta}(t)) = \sum_{a,b\in\mathcal{A}} \sum_{k\in\mathbb{Z}_+} \varphi_\omega^{(k,a)}(\boldsymbol{\zeta}(t)) q_{k|a,b} p_a p_b, \quad q_{k|a,b} = \sum_{\substack{\boldsymbol{k}\in\mathbb{Z}_+^{\mathcal{A}}:\\ \boldsymbol{k}^T\mathbf{1}=k}} q^b_{\boldsymbol{k}|a},$$

and

$$\varphi_\omega^{(k,a)}(\boldsymbol{\zeta}(t)) = \sum_{\substack{\boldsymbol{\xi}\in\mathbb{Z}_+^{\mathcal{X}}:\\ \boldsymbol{\xi}^T\mathbf{1}=k}} \Theta_\omega^{(a)}(\boldsymbol{\xi}) \binom{k}{\boldsymbol{\xi}} \prod_{g\in\mathcal{X}} [\zeta_g(t)]^{\xi_g}.$$

Analogously, defining $y_\omega(t) = \mathbb{P}[Z_0(t) = \omega] = \sum_{a\in\mathcal{A}} y_{\omega|a}(t) p_a$*, the ODE replacing* (4) *can be written as*

$$\frac{\mathrm{d}y_\omega(t)}{\mathrm{d}t} = \gamma\left(\psi_\omega(\boldsymbol{\zeta}(t)) - y_\omega(t)\right), \tag{6}$$

where

$$\psi_\omega(\boldsymbol{\zeta}(t)) = \sum_{a\in\mathcal{A}} \sum_{k\in\mathbb{Z}_+} \varphi_\omega^{(k,a)}(\boldsymbol{\zeta}(t)) p_{k|a} p_a, \quad p_{k|a} = \sum_{\substack{\boldsymbol{k}\in\mathbb{Z}_+^{\mathcal{A}}:\\ \boldsymbol{k}^T\mathbf{1}=k}} p_{\boldsymbol{k}|a}.$$

REMARK 2. *The above analysis of ASD over the ensemble* $\mathcal{T}$ *can be easily extended to the case in which the rate of activation of a vertex depends on its label* $a \in \mathcal{A}$*. Without loss of generality we will assume* $\gamma = 1$ *in the following.*

4. ASD on labeled random networks. In this section we consider the evolution of ASD process over a multigraph $\mathcal{G}$ taken from a general ensemble of labeled directed graphs $\mathfrak{E}^{(n)}$ of size n. In particular, we show that, under certain conditions on the ensemble and on the initial node configuration, the ASD process over $\mathcal{G}$ can be well approximated by the same process over a labeled branching process $\mathcal{T}$. The ensemble $\mathfrak{E}^{(n)}$ is described by the 'node statistics' $p_{\mathbf{d},\mathbf{k},a,s}$, which provides the probability that a node picked at random has in-degree vector $\mathbf{d}$, out-degree vector $\mathbf{k}$, label $a \in \mathcal{A}$ and initial state $s \in \mathcal{X}$. We shall assume that $p_{\mathbf{d},\mathbf{k},a,s}$ factorizes as $p_{\mathbf{d},\mathbf{k},a,s} = p_{\mathbf{d},\mathbf{k},a} p_{s|a}$. The above node statistics clearly provides all information needed to compute any marginal or conditional distribution we might be interested in. For example, $p_{\mathbf{d},\mathbf{k},a} = \sum_{s\in\mathcal{X}} p_{\mathbf{d},\mathbf{k},a,s}$ is the distribution of in-degree vector, out-degree vector and label of a generic node. As another example, $p_{\mathbf{k},a} = \sum_{\mathbf{d}} p_{\mathbf{d},\mathbf{k},a}$ provides the distribution of out-degree vector and label of a generic node. We denote $p_a = \sum_{\mathbf{k}} p_{\mathbf{k},a}$ the probability for a node to be associated with label a. With intuitive notation, $p_{\mathbf{k}|a} = p_{\mathbf{k},a}/p_a$ denotes the distribution of out-degree vector of a node with label a, and so on.

As it always happens in graphs with heterogeneous degrees, we will need to distinguish the probability law of $\mathbf{k}_v$ for a generic node v picked uniformly at random, and the probability law of $\mathbf{k}_v$ for a node v reached by traversing an edge. This because, in general, we could have correlation between in-degree and out-degree. Moreover, when

we reach a node by following a certain edge, it is also important to distinguish the label of the node originating the traversed edge picked uniformly at random. To account for the above generality, we need to introduce some additional notation. Specifically, we define $q^a_{\mathbf{d},\mathbf{k}|b} := d_a p_{\mathbf{d},\mathbf{k},b} / \sum_{\mathbf{d},\mathbf{k}} d_a p_{\mathbf{d},\mathbf{k},b}$, which is the distribution of in-degree vector $\mathbf{d}$ and out-degree vector $\mathbf{k}$ of a node with label b, reached by traversing an edge from a node with label a. Similarly, $q^a_{\mathbf{k}|b} = \sum_{\mathbf{d}} q^a_{\mathbf{d},\mathbf{k}|b}$ is the marginal distribution of out-degree vector of a node with label b, reached by traversing an edge from a node with label a.

4.1. Relevant neighborhood at time t. We first observe that, since the process evolves through local interactions, the state of a generic node v on a multigraph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{A}, \lambda, \sigma, \tau)$ at time t is determined only by the structure and state of a relatively small neighborhood around v. Given a generic node $v \in \mathcal{V}$, we define the *relevant neighborhood* $\mathcal{N}_t$ of v as the subgraph induced by the set of all nodes in $\mathcal{V}$ having an impact on $Z_v(t)$, i.e., on the state of v at time t. Similarly, we define the *relevant neighborhood* $\mathcal{T}_t$ as the subtree induced by the set of all nodes in $\mathcal{T}$ having an impact on $Z_{v_0}(t)$, where v_0 is the root node of $\mathcal{T}$.

The relevant neighborhood can be built by looking backward in time, identifying dependencies between neighboring nodes. First, observe that the state of v at time t depends on its out-neighbors v' (one-hop away nodes) if and only if v has updated its state in $[0, t]$ at least once, i.e., we can find an update time of v, $\vartheta_v(t) \leq t$. The state of node v depends on a two-hop away node v'', if and only if we can find a common neighbor v' of v and v'', such that $\vartheta_{v'}(t) < \vartheta_v(t) \leq t$. Similarly the state of v depends on a three-hops away node v''' only if we can find two nodes v' and v'' along a directed path from v to v''' such that $\vartheta_{v''}(t) < \vartheta_{v'}(t) < \vartheta_v(t) \leq t$, and so on.

Due to the fact that update times of each node form independent Poisson processes with rate $\gamma = 1$, we can exploit well-known properties of the Poisson process (time-reversibility, memoryless property) to obtain $\mathcal{N}_t$ (or $\mathcal{T}_t$) as the result of a process evolving forward in time, and exploring progressively the neighborhood of v by adding an exponentially distributed delay (of mean 1) on each explored node, up to time t.

More precisely, the relevant neighborhood of v is obtained by the following process. Vertices can be active, neutral or inactive. Initially, the relevant neighborhood is empty.

1. The process starts by activating node v at time $t = 0$. All of the other nodes are set neutral.
2. Upon the activation of a node, a random timer is associated to it, taken from an exponential distribution of mean $\gamma = 1$. Moreover, the node is added to the relevant neighborhood, together with the outgoing edges.
3. Upon expiration of its associated timer: i) an active node is set inactive; ii) all of its neutral out-neighbors are set active and added to the relevant neighborhood, together with outgoing edges.

For $t \geq 0$, we can stop the above exploration process at time t (i.e., we no longer add nodes to the relevant neighborhood after time t), obtaining a truncated version $\mathcal{N}_t$ of $\mathcal{G}$, composed of all the nodes that have been activated. Similarly, we obtain a truncated version $\mathcal{T}_t$ of $\mathcal{T}$.

4.2. Approximation result. Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{A}, \lambda, \sigma, \tau)$ be a multigraph sampled from a given labeled network ensemble $\mathfrak{E}^{(n)}$ of size n. For $t \geq 0$, let $\mathcal{N}_t$ be the relevant neighborhood at time t of a node v chosen uniformly at random from $\mathcal{V}$, and let $\mu_{\mathcal{N}_t}$ be its distribution on the multigraph space. Let $\mathcal{T}_t$ be a labeled branching process as defined in Sec. 3.1, truncated at time t, and let $\mu_{\mathcal{T}_t}$ be its distribution.

Proposition 3 identifies some sufficient conditions to guarantee that the ASD process over a network is well approximated by the solution of the differential equation in (4).

PROPOSITION 3. *For $t \geq 0$, let $Z(t)$ be the state vector of the ASD at time t on $\mathcal{G}$. Let $z_\omega(t) = \frac{1}{n}|\{v \in \mathcal{V} : Z_v(t) = \omega\}|$ be the fraction of state-ω adopters at time t, and $\overline{z}_\omega(t) = \mathbb{E}[z_\omega(t)]$ be its expectation over the ensemble. For any $\epsilon > 0$,*

$$\mathbb{P}(|z_\omega(t) - y_\omega(t)| \geq \epsilon) \leq \mathbb{P}(|z_\omega(t) - \overline{z}_\omega(t)| \geq \epsilon - \|\mu_{\mathcal{N}_t} - \mu_{\mathcal{T}_t}\|_{\mathrm{TV}})$$

where $y_\omega(t)$ is the solution of (4).

Proof. Notice that

$$\mathbb{P}(|z_\omega(t) - y_\omega(t)| \geq \epsilon) \leq \mathbb{P}(|z_\omega(t) - \overline{z}_\omega(t)| + |\overline{z}_\omega(t) - y_\omega(t)| \geq \epsilon) \tag{7}$$

We prove now that $|\overline{z}_\omega(t) - y_\omega(t)| \leq \|\mu_{\mathcal{N}_t} - \mu_{\mathcal{T}_t}\|_{\mathrm{TV}}$ from which we get the result.

Observe that, by definition, the state $Z_v(t)$ of node v depends exclusively on the initial states $Z_j(0) = \sigma_j$ of the agents belonging to the relevant neighborhood $\mathcal{N}_t$ of node v at time t, i.e., $\mathbb{P}(Z_v(t) = \omega) = \chi_\omega(\mathcal{N}_t)$ where χ_ω is a function in the range [0,1].

We thus have

$$\overline{z}_\omega(t) = \mathbb{E}[z_\omega(t)] = \frac{1}{n}\sum_{v\in\mathcal{V}} \mathbb{P}(Z_v(t) = \omega) = \int \chi_\omega(g)\mathrm{d}\mu_{\mathcal{N}_t}(g).$$

On the other hand, considering the state of the root in the labeled branching process $\mathcal{T}$, the output of the ODE (4) satisfies

$$y_\omega(t) = \int \chi_\omega(g)\mathrm{d}\mu_{\mathcal{T}_t}(g).$$

It then follows

$$\begin{aligned}
|\overline{z}_\omega(t) - y_\omega(t)| &\leq \left|\int \chi_\omega(g)\mathrm{d}\mu_{\mathcal{N}_t}(g) - \int \chi_\omega(g)\mathrm{d}\mu_{\mathcal{T}_t}(g)\right| \\
&\leq \left|\int \left(\chi_\omega(g) - \frac{1}{2}\right)\mathrm{d}\mu_{\mathcal{N}_t}(g) - \int \left(\chi_\omega(g) - \frac{1}{2}\right)\mathrm{d}\mu_{\mathcal{T}_t}(g)\right| \\
&\leq \|\mu_{\mathcal{N}_t} - \mu_{\mathcal{T}_t}\|_{\mathrm{TV}}.
\end{aligned}$$

□

From Proposition 3 we deduce that the evolution of the ASD process is well approximated by the solution of the differential equation in (4) for graph ensembles enjoying the following two fundamental properties:

(a) Topological Property: Local Weak Convergence is required, in the sense that $\|\mu_{\mathcal{N}_t} - \mu_{\mathcal{T}_t}\|_{\mathrm{TV}}$ can be made arbitrarily small by increasing the graph size.
(b) Concentration Property: for large graph size, the fraction of state-ω adopters in the ASD process must concentrate around its expectation with probability close to one.

5. ASD over labeled configuration model. Considering the ensable $\mathfrak{E}^{(n)}$ of all labeled networks with given size n and statistics $p_{\mathbf{d},\mathbf{k},a,s}$, we define the corresponding *labeled configuration model* ensamble $\mathfrak{C}_{n,p}$, on which we will restrict our investigation in the rest of the paper. In particular, we provide general bounds for ASD evolution over $\mathfrak{C}_{n,p}$.

Next we consider two specific examples of labeled configuration model, which we believe are particularly interesting, and apply to them the general bounds above, showing asymptotic convergence to the ODE solution as the network size grows large.

5.1. Labeled configuration model. We first explicitly describe the construction of the labeled configuration model $\mathfrak{C}_{n,p}$. For each $v \in \mathcal{V}$, denote with $\boldsymbol{\kappa}_v = (\kappa_v^a)_{a\in\mathcal{A}}$ and $\boldsymbol{\delta}_v = (\delta_v^a)_{a\in\mathcal{A}}$ the out-degree and in-degree vectors, respectively, such that there is exactly a fraction $p_{\mathbf{d},\mathbf{k},a}$ of nodes $v \in \mathcal{V}$ with $(\boldsymbol{\delta}_v, \boldsymbol{\kappa}_v, a_v) = (\mathbf{d}, \mathbf{k}, a)$. Denote with $\mathcal{L}_{a,a'}$ a set of stubs, and define arbitrary maps $\nu_{a,a'}, \gamma_{a,a'} : \mathcal{L}_{a,a'} \to \mathcal{V}$, satisfying the property: $|\nu_{a,a'}^{-1}(v)| = \delta_v^a$ for nodes v with label $\lambda(v) = a'$ and $|\gamma_{a,a'}^{-1}(v)| = \kappa_v^{a'}$ with $\lambda(v) = a$. For all $a, a' \in \mathcal{A}$, let $\pi_{a,a'}$ be chosen uniformly at random among all permutations of $\mathcal{L}_{a,a'}$ and define multigraph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{A}, \lambda, \sigma, \tau)$ with set of nodes $\mathcal{V}$ and $\mathcal{E} = \bigcup_{(a,a')\in\mathcal{A}\times\mathcal{A}} \mathcal{E}_{a,a'}$, where $\mathcal{E}_{a,a'} = (\gamma_{a,a'}(h), \nu_{a,a'}(\pi_{a,a'}(h)) : h \in \mathcal{L}_{a,a'}$, and $\sigma(\gamma_{a,a'}(h), \nu_{a,a'}(\pi_{a,a'}(h))) = \gamma_{a,a'}(h)$ and $\tau(\gamma_{a,a'}(h), \nu_{a,a'}(\pi_{a,a'}(h))) = \nu_{a,a'}(\pi_{a,a'}(h))$.

Denote with $l_{a,a'} = |\mathcal{L}_{a,a'}|$ the total number of edges incoming to nodes with label a', originating from nodes with label a, so that:

$$l_{a,a'} = n \sum_{\mathbf{d},\mathbf{k}} d_a p_{\mathbf{d},\mathbf{k},a'} = n \sum_{\mathbf{d},\mathbf{k}} k_{a'} p_{\mathbf{d},\mathbf{k},a}$$

The total number of edges in the graph is $l = \sum_{a,a'} l_{a,a'}$. The average in-degree of a node, which is equal to the average out-degree, will be denoted by $\bar{d} = l/n$.

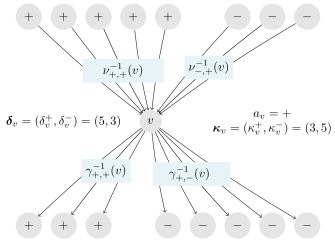
We repeat here for readers' ease the expression of the fraction of nodes with label a', reached from a node with label a, having in-degree vector $\mathbf{d}$ and out-degree vector $\mathbf{k}$:

$$q^a_{\mathbf{d},\mathbf{k}|a'} = \frac{d_a p_{\mathbf{d},\mathbf{k},a'}}{\sum_{\mathbf{d},\mathbf{k}} d_a p_{\mathbf{d},\mathbf{k},a'}} \tag{8}$$

We summarize the notations in Figure 1.(a) and we show an example of notation use for a simple case in which $\mathcal{A} = \{+, -\}$ (see Figure 1.(b)).

$\mathcal{A}$ ordered set of labels
$a_v = \lambda(v)$ label of node v
$\mathcal{V}_a = \{v \in \mathcal{V} : a_v = a\}$
$\boldsymbol{\delta}_v = (\delta_v^a)_{a\in\mathcal{A}}$ in-degree of node v
$\boldsymbol{\kappa}_v = (\kappa_v^a)_{a\in\mathcal{A}}$ out-degree of node v
$\mathcal{L}_{a,a'}$ set of stubs from nodes with label a to nodes with label a'
$l_{a,a'} = |\mathcal{L}_{a,a'}|$ number of edges from nodes with label a to nodes with label a'
$\pi_{a,a'}$ permutation of $\mathcal{L}_{a,a'}$
$\nu_{a,a'} : \mathcal{L}_{a,a'} \to \mathcal{V}$ map with the property $|\nu_{a,a'}^{-1}(v)| = \delta_v^a$ for all $v \in \mathcal{V}_{a'}$
$\gamma_{a,a'} : \mathcal{L}_{a,a'} \to \mathcal{V}$ map with the property $|\gamma_{a,a'}^{-1}(v)| = \kappa_v^{a'}$ for all $v \in \mathcal{V}_a$
$\mathcal{E}_{a,a'} = (\gamma_{a,a'}(h), \nu_{a,a'}(\pi_{a,a'}(h)) : h \in \mathcal{L}_{a,a'}\}$ set of edges from nodes with label a to nodes with label a'
$\mathcal{E} = \bigcup_{a,a'} \mathcal{E}_{a,a'}$

(a)



(b)

FIGURE 1. (a) *Summary of notations* (b) *Labeled configuration model with two classes* $\mathcal{A} = \{+, -\}$

We emphasize that the graph ensemble defined above extends the classical configuration model, which can be recovered as a particular case by setting $|\mathcal{A}| = 1$.

5.2. Bounds to ASD dynamics. In order to apply the general approximation result stated in Proposition 3 to the labeled configuration model defined above, we need to prove both Local Weak Convergence and Concentration Property of ASD. The following theorems actually provide the main results of our paper:

- Theorem 4 is a topological result and is related exclusively to the properties of the labeled configuration model. More precisely, it provides a useful bound on the total variation distance between the relevant neighborhood of a graph drawn uniformly at random from the labeled configuration model ensemble

and the labeled branching process described in Section 3. The proof is rather technical and the interested reader can find the details in Appendix A.1 of the Supplementary Material.

- Theorem 5 is related to the ASD evolution and to the specific properties of the labeled configuration model. It provides a bound on the distance between the fraction of state-ω adopters and its expectation. The proof can be found in Appendix A.2 of the Supplementary Material.

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{A}, \lambda, \sigma, \tau)$ be a multigraph sampled from the ensemble $\mathfrak{C}_{n,p}$. For $t \geq 0$, let $\mathcal{N}_t$ be the relevant neighborhood at time t of a node v chosen uniformly at random from $\mathcal{V}$, and let $\mu_{\mathcal{N}_t}$ be its distribution. Let $\mathcal{T}_t$ be the labeled branching process truncated at time t and let $\mu_{\mathcal{T}_t}$ be its distribution. Moreover, let $W_t^{b,a}$ be the number of edges in $\mathcal{T}_t$ from nodes with label $b \in \mathcal{A}$ to nodes with label $a \in \mathcal{A}$ and let $\mathsf{F}_{W_t^{b,a}}(x_{b,a}) = \mathbb{P}(W_t^{b,a} > x_{b,a})$.

THEOREM 4 (Topological Property). *We have*

$$\begin{aligned} &\|\mu_{\mathcal{T}_t} - \mu_{\mathcal{N}_t}\|_{\mathrm{TV}} \leq \mathbb{P}(\mathcal{T}_t \neq \mathcal{N}_t) \\ &\leq \inf_{\boldsymbol{X} \in (\mathbb{R}^+)^{\mathcal{A}\times\mathcal{A}}} \sum_{a,b\in\mathcal{A}} \left[\mathsf{F}_{W_t^{b,a}}(x_{b,a}) + \frac{x_{b,a}(x_{b,a}+1)}{2} \frac{\sum_{\boldsymbol{d},\boldsymbol{k}} d_b q^b_{\boldsymbol{d},\boldsymbol{k}|a}}{l_{b,a}} \right. \qquad (9) \\ &\left. + \sum_{b' \neq b} x_{b,a} x_{b',a} \frac{\sum_{\boldsymbol{d},\boldsymbol{k}} d_b q^{b'}_{\boldsymbol{d},\boldsymbol{k}|a}}{l_{b,a}} \right] \qquad (10) \end{aligned}$$

EXAMPLE 1 (Topological Property for the classical configuration model). *If $|\mathcal{A}| = 1$ then the model ensemble $\mathfrak{C}_{n,p}$ boils down to the classical configuration model and the bound derived in (9) reduces to*

$$\|\mu_{\mathcal{T}_t} - \mu_{\mathcal{N}_t}\|_{\mathrm{TV}} \leq \mathbb{P}(\mathcal{T}_t \neq \mathcal{N}_t) \leq \inf_{x>0} \left[\mathsf{F}_{\widetilde{W}_t}(x) + \frac{\sum_{d,k} d q_{d,k} x(x+1)}{2n\overline{d}} \right] \qquad (11)$$

where $\widetilde{W}_t$ is the number of nodes in $\mathcal{T}_t$ and $\mathsf{F}_{\widetilde{W}_t}(x) = \mathbb{P}(\widetilde{W}_t > x)$.

We next introduce the concentration result that allows us to estimate to what extent the fraction of state-ω adopters in the ASD process concentrates around its expectation.

THEOREM 5 (Concentration Property). *Let $\mathcal{G}$ be a multigraph sampled from the ensemble $\mathfrak{C}_{n,p}$. We denote with $\mathcal{N}_t^v$ the relevant neighborhood at time t of a node v, sampled with a probability proportional to its in-degree, and with V_t^v the number of nodes in it. For $t \geq 0$, let $Z(t)$ be the state vector of the ASD dynamics on $\mathcal{G}$, $b(t) = |\{v \in \mathcal{V} : Z_v(t) = \omega\}|$ be the number of state-ω adopters at time t conditioned to $\mathcal{G}$. For any $\epsilon > 0, \eta > 0, x > 0, s \geq 1$ we have*

$$\begin{aligned} &\mathbb{P}(|b(t) - \mathbb{E}[b(t)]| > \eta n) \\ &\leq 4\mathrm{e}^{-\frac{\eta^2 n}{1152(1+\epsilon)tx^2}} + \left(1 + \frac{12}{\eta}\right)(1+\epsilon)tn\frac{2^s \mathbb{E}_v\left[|V_t^v|^s\right]}{x^s} + 2\mathrm{e}^{-\frac{nt\epsilon^2}{2(1+\epsilon)}} + 2\mathrm{e}^{-\frac{\eta^2 n}{288(t+\eta/12)}} \\ &+ \left(1 + \frac{4}{\eta}\right)\frac{2^s}{x^s} \sum_{w\in\mathcal{V}} |\boldsymbol{\delta}_w| [\mathbb{E}[|V_t^w|^s] + 2\mathrm{e}^{-\frac{\eta^2 n}{128\overline{d}x^2}} \end{aligned}$$

REMARK 3. *We emphasize that the bounds presented in Theorem 4 and Theorem 5 represent an important step forward with respect to results already known in literature. In particular, Lemma 5 and Proposition 2 in [34] states a similar result for a different, simplified version of our system dynamics in which: i) node updates are synchronized, being triggered by a common discrete time step; ii) the update rule is restricted to be the deterministic linear threshold model; iii) both the maximum in-degree $d_{\max}$ and the maximum out-degree $k_{\max}$ of nodes are supposed to be finite. In contrast to [34], we introduce a much more general result along three different directions:*

1. *We consider asynchronous dynamics (each node is updated by an independent Poisson clock). Hence the neighborhood exploration process in the proof has to take into account this new source of randomness. Specifically, the estimation of the total variation (9) is split into two terms, which are obtained by conditioning on the number of nodes in $\mathcal{T}_t$. The necessity of this refined analysis will be clear in the next section.*
2. *We consider arbitrary semi-anonymous dynamics, with possible random (noisy) response to the state of neighbors. Moreover, we define our dynamics on a much more general ensemble of labelled random graphs, which allows us to differentiate the distribution of incoming/outgoing edges for each pair of classes.*
3. *We allow the maximum in- and out-degree of nodes to possibly scale with n (under some technical constraints). This is crucial for applications to social networks and many other complex systems in which the degree distribution has often been observed to follow a power law. But notice that even in the case of the classic Erdös-Rényi random graph $G(n,p)$, $d_{\max}$ or $k_{\max}$ of course are not independent of n. Note that, in the case of finite $d_{\max}, k_{\max}$, by taking $x = k_{\max}^t$ our bound in (11) leads to the result in [34] (Lemma 5).*

In the next section, we show how the bounds derived above for Local Weak Convergence and concentration property can be used to study asymptotic behavior of ASD on a labeled configuration model with power-law degree distribution. More precisely, we will consider a sequence of labeled graphs with size n and described by distributions $p^{(n)}_{\mathbf{k}|a}, q^{(n)}_{\mathbf{k}|a}$ such that $p^{(n)}_{\mathbf{k}|a} \overset{n\to\infty}{\longrightarrow} p_{\mathbf{k}|a}$, $q^{(n)}_{\mathbf{k}|a} \overset{n\to\infty}{\longrightarrow} q_{\mathbf{k}|a}$. Then we will consider the labeled branching process obtained by the construction above with asymptotic distributions. The following proposition quantifies the distance between the solution corresponding to the differential equation with distribution $p^{(n)}_{\mathbf{k}|a}, q^{(n)}_{\mathbf{k}|a}$ and the solution corresponding to the differential equation with asymptotic distribution $p_{\mathbf{k}|a}, q_{\mathbf{k}|a}$.

PROPOSITION 6. *Let*

- $\boldsymbol{\zeta}^{(n)}(t)$ *be the solution of* (3) *with* $q^{(n)}_{\boldsymbol{k}|a}$ *and initial condition* $\boldsymbol{\zeta}^{(n)}_0$*;*
- $\boldsymbol{\zeta}(t)$ *be the solution of* (3) *with* $q_{\boldsymbol{k}|a}$ *and initial condition* $\boldsymbol{\zeta}_0$*.*
- $\boldsymbol{y}^{(n)}(t)$ *be the solution of* (3) *with* $p^{(n)}_{\boldsymbol{k}|a}$ *and initial condition* $\boldsymbol{y}^{(n)}_0$*;*
- $\boldsymbol{y}(t)$ *be the solution of* (3) *with* $p_{\boldsymbol{k}|a}$ *and initial condition* $\boldsymbol{y}_0$*.*

In addition let $\boldsymbol{\phi}(\boldsymbol{z}) - \boldsymbol{z}$ and $\boldsymbol{\psi}(\boldsymbol{z})$ be Lipschitz continuous in $[0,1]^{|\mathcal{A}|}$, and let L and $M > 0$ be the Lipschitz constants corresponding to infinity norm. Then

$$\sup_{t\in[0,m\Delta]} \|\boldsymbol{\zeta}^{(n)}(t) - \boldsymbol{\zeta}(t)\|_\infty \leq \frac{\|\boldsymbol{\zeta}^{(n)}_0 - \boldsymbol{\zeta}_0\|_\infty}{(1-\Delta L)^m} + \frac{1}{L}\left(\frac{1}{(1-\Delta L)^m} - 1\right) \|q^{(n)}_{\boldsymbol{k}|a} - q_{\boldsymbol{k}|a}\|_{\mathrm{TV}}$$

with $\Delta < 1/L$ *and*

$$\sup_{t\in[0,m\Delta]} \|\boldsymbol{y}^{(n)}(t) - \boldsymbol{y}(t)\|_\infty \leq \frac{\|\boldsymbol{y}_0^{(n)} - \boldsymbol{y}_0\|_\infty}{(1-\Delta)^m} + \left(\frac{1}{(1-\Delta)^m} - 1\right)\left[M \sup_{t\in[0,m\Delta]} \|\boldsymbol{\zeta}^{(n)}(t) - \boldsymbol{\zeta}(t)\|_\infty + \|p_{\boldsymbol{k}|a}^{(n)} - p_{\boldsymbol{k}|a}\|_{\mathrm{TV}}\right].$$

The proof can be found in Appendix B of the Supplementary Material.

In particular, if $p_{\mathbf{k}|a}^{(n)}$ is a truncated version of $p_{\mathbf{k}|a}$, we can apply to the previous bound the following statement of immediate verification:

PROPOSITION 7. *Consider a generic distribution* $p_{\boldsymbol{k}|a}$ *and its truncated version* $p_{\boldsymbol{k}|a}^{(n)}$, *i.e.* $p_{\boldsymbol{k}|a}^{(n)} = \frac{p_{\boldsymbol{k}|a}\mathbf{1}_{\boldsymbol{k}\in\mathcal{B}_n}}{\sum_{\boldsymbol{k}\in\mathcal{B}_n} p_{\boldsymbol{k}|a}}$ *for a generic compact set* $\mathcal{B}_n \in \mathbb{N}^{|\mathcal{A}|}$, *then we have:* $\|p_{\boldsymbol{k}|a}^{(n)} - p_{\boldsymbol{k}|a}\|_{\mathrm{TV}} = 1 - \sum_{\boldsymbol{k}\in\mathcal{B}_n} p_{\boldsymbol{k}|a}$.

5.3. Asymptotic behavior on labeled configuration model with power-law degree distribution. In this section we consider the classical configuration model with a *truncated power-law degree distribution*, which is a particular case of labeled configuration model with $|\mathcal{A}| = 1$. We simplify the notation: let $p_{d,k}^{(n)}$ be the fraction of nodes with in-degree d and out-degree k, where we have highlighted the number of nodes n, and let $\overline{d} = \sum_{d,k} d p_{d,k}^{(n)}$ be the average degree.

ASSUMPTION 1. *Let us assume that* $\sum_{d,k} d q_{d,k}^{(n)} = \Theta(n^\delta)$ *with* $0 \leq \delta < 1/2$. *This means that we allow the average in-degree of a node, reached by an edge selected uniformly at random, to possibly scale with* n. *Moreover, we will assume that* $q_k^{(n)} = O(k^{-\beta})$ *follows a power-law of exponent* $\beta > 2$ *and maximum value* $k_{\max} = \Theta(n^\zeta)$ *with* $\zeta < \min\{(1-\delta)/2, 1/(\beta-1)\}$.

Let μ_s be the s-th moment of $q = \{q_k^{(n)}\}_{k\geq 0}$. From Assumption 1 we have

$$\mu_s = \begin{cases} \Theta(1) & \text{if } \beta > s+1 \\ \Theta(n^{\zeta(s+1-\beta)}) & \text{if } 2 < \beta < s+1. \end{cases} \tag{12}$$

Notice that, being $\beta > 2$, μ_1 is always finite and, therefore, does not scale with n.

In order to guarantee that the ASD over a network drawn uniformly at random from the configuration model ensemble is well approximated by the solution of ODE, it is sufficient that the terms in the upper bounds derived in (11) (see Example 1), in Theorem 5, and in Proposition 6 go to zero when $n \to \infty$. In the following, let $\mathcal{N}^{(n)} = (\mathcal{V}^{(n)}, \mathcal{E}^{(n)})$ be a sequence of networks, each one sampled from the corresponding model ensemble $\mathfrak{C}_{n,p^{(n)}}$, where $\{p^{(n)}\}_n$ is a sequence of truncated versions of a power law distribution of p satisfying Assumption 1. For $t \geq 0$, let $\mathcal{N}_t^{(n)}$ be the relevant neighborhood of a node v chosen uniformly at random from the node set $\mathcal{V}^{(n)}$. Moreover, let $\mathcal{T}_t^{(n)}$ be the sequence of truncated Galton-Watson (GW) processes (see [16]) for which the root offspring follows distribution $p^{(n)}$, while the degree of non-root nodes follow law $q^{(n)}$. Finally, let $p^{(n)} \overset{n\to\infty}{\to} p$ and $q^{(n)} \overset{n\to\infty}{\to} q$.

Before presenting the topological result for the configuration model with power-law degree distribution, we present two technical results, whose proofs are given in Appendix C of the Supplementary Material.

LEMMA 8 (Bound on the number of nodes/edges in $\mathcal{T}_t$). *Consider the GW process $\mathcal{T}$ in which the offspring distribution of the root follows law p, while the degree of remaining nodes follows law q. Let $\mathcal{T}_t$ be the corresponding random tree obtained by truncating $\mathcal{T}$ at time t, and $\widetilde{W}_t$ be the number of nodes in $\mathcal{T}_t$. Let $h_n = c\log n$ for some $c > 0$ and $t = o(h_n)$ as $n \to \infty$, then for any $s > 0$ we have*

$$\mathsf{F}_{\widetilde{W}_t}(x_n) \leq \frac{\mathbb{E}[N_{h_n}^s]}{x_n^s} + o(1/n) \qquad n \to \infty$$

where $\{N_h\}_{h\in\mathbb{N}}$ is the number of nodes in a truncated version of $\mathcal{T}$ with maximal width h.

LEMMA 9. *Let $\{N_h\}_{h\geq 0}$ be a supercritical GW process, in which the offspring distribution of the root follows law p, while the degree of remaining nodes follows law q. We have: $\mathbb{E}[N_h^s] = O(\mu_s \cdot \mu_1^{s(h-1)})$, $\forall \beta > 1$, where μ_j is the j-th moment of q.*

THEOREM 10 (Topological result for configuration model with power-law degree distribution). *With the above definitions, let $\mu_{\mathcal{N}_t^{(n)}}$ and $\mu_{\mathcal{T}_t^{(n)}}$ be the distributions of $\mathcal{N}_t^{(n)}$ and $\mathcal{T}_t^{(n)}$, respectively. Under Assumption 1, for $t = o(\log n)$, we have $\|\mu_{\mathcal{T}_t^{(n)}} - \mu_{\mathcal{N}_t^{(n)}}\|_{\mathrm{TV}} \leq \mathbb{P}(\mathcal{T}_t \neq \mathcal{N}_t) = o(1)$ when $n \to \infty$.*

Proof. From inequality (11), we have

$$\|\mu_{\mathcal{T}_t} - \mu_{\mathcal{N}_t}\|_{\mathrm{TV}} \leq \mathbb{P}(\mathcal{T}_t \neq \mathcal{N}_t) \leq \inf_{x>0}\left[\mathsf{F}_{\widetilde{W}_t}(x) + \frac{\sum_{d,k} dq_{d,k}x(x+1)}{2n\overline{d}}\right] \tag{13}$$

where $\widetilde{W}_t$ is the number of nodes in $\mathcal{T}_t$. Let $x_n = n^{(1-\delta)/2-\gamma}$ for some $\gamma > 0$ and $h_n = c\log n$ for some $c > 0$ then

$$\begin{aligned}\|\mu_{\mathcal{T}_t} - \mu_{\mathcal{N}_t}\|_{\mathrm{TV}} &\leq \frac{\sum_d dq_{d,k}x_n(x_n+1)}{2n\overline{d}} + \mathsf{F}_{\widetilde{W}_t}(x_n) \qquad n \to \infty \\ &\leq \frac{\sum_d dq_{d,k}x_n(x_n+1)}{2n\overline{d}} + \frac{\mathbb{E}[N_{h_n}^s]}{x_n^s} + o(1/n)\end{aligned}$$

where the last inequality holds for any $s > 1$ and $\{N_h\}_{h\in\mathbb{N}}$ is (the number of nodes of) a truncated GW process of maximal width h, in which the offspring distribution of the root follows law p, while the degree of remaining nodes follow law q (see Lemma 8).

Under Assumption 1 we prove that there exists s such that $\frac{\mathbb{E}[N_{h_n}^s]}{x_n^s} = o(1/n)$ as $n \to \infty$ and we conclude that for some $\gamma > 0$

$$\|\mu_{\mathcal{T}_t} - \mu_{\mathcal{N}_t}\|_{\mathrm{TV}} \leq \frac{1}{2\overline{d}n^{2\gamma}} + o(1/n) = o(1) \qquad n \to \infty.$$

To find a suitable value of s, we distinguish two cases.

(i) If $\beta > \lfloor\frac{2}{1-\delta}\rfloor + 2$ we can simply choose $s = \lfloor\frac{2}{1-\delta}\rfloor + 1$. By so doing, we stay in the case $\beta > s+1$, and from (12) and Lemma 9 we get: $\mathbb{E}[N_{h_n}^s] = \Theta(n^{cs\log\mu_1})$ and

$$\frac{\mathbb{E}[N_{h_n}^s]}{x_n^s} = \Theta(n^{-s\left(\frac{1-\delta}{2}-\gamma-c\log\mu_1\right)}) = o(1/n) \qquad n \to \infty$$

(ii) If $\beta \leq \lfloor \frac{2}{1-\delta} \rfloor + 2$, we choose instead a sufficiently large value of s, falling in the case $s > \beta - 1$ in which μ_s scales with n as in (12). In particular, from Lemma 9 we have $\mathbb{E}[N^s_{h_n}] = \Theta(n^{\zeta(s+1-\beta)+cs\log\mu_1})$. Thus

$$\mathbb{E}[N^s_{h_n}]/x^s_n = \Theta\left(n^{\zeta(s+1-\beta)+cs\log\mu_1 - s\left(\frac{1-\delta}{2}-\gamma\right)}\right).$$

We now observe that if there exists $s \in \mathbb{N}$ such that

$$s\left(\frac{1-\delta}{2} - \zeta - c\log\mu_1 - \gamma\right) > 1 - \zeta(\beta-1) \tag{14}$$

then $\mathbb{E}[N^s_{h_n}]/x^s_n = o(1/n)$ for $n \to \infty$. Since $\zeta < 1/(\beta-1)$, the right hand side in (14) is positive, and since $\zeta < \min\{\frac{1-\delta}{2}\}$, we can always find two sufficiently small constants γ and c such that $\left(\frac{1-\delta}{2} - \zeta - c\log\mu_1 - \gamma\right)$ is also positive. Therefore, there exists an integer s large enough such that both $s > \beta - 1$ and (14) are satisfied. ∎

COROLLARY 11. *For $t \geq 0$, let $\boldsymbol{Z}(t)$ be the state vector of the ASD dynamics on $\mathcal{N}^{(n)}$ and $z^{(n)}_\omega(t) = \frac{1}{n}|\{v \in \mathcal{V} : Z_v(t) = \omega\}|$ be the fraction of state-ω adopters at time t. Under Assumptions 1 for $t = o(\log n)$, for any $\eta > 0$*

$$\mathbb{P}(|z^{(n)}_\omega(t) - y_\omega(t)| > \eta) = o(1) \qquad \text{for } n \to \infty$$

where $y_\omega(t)$ is the solution of (4) over a GW tree $\mathcal{T}_t$ with the asymptotic degree statistics p and q.

Proof. Let $\mathcal{T}^{(n)}_t$ be the continuous-time branching process truncated up to time t, $\mu_{\mathcal{T}^{(n)}_t}$ be its distribution. Denote by $\boldsymbol{y}^{(n)}(t)$ the solution of (4) with $p^{(n)}_{\mathbf{k}|a}$ and initial condition $\boldsymbol{y}^{(n)}_0$. We have

$$\mathbb{P}(|z^{(n)}_\omega(t) - y_\omega(t)| \geq \eta) \leq \mathbb{P}(|z^{(n)}_\omega(t) - y^{(n)}_\omega(t)| \geq \eta/2) + \mathbb{P}(|y^{(n)}_\omega(t) - y_\omega(t)| \geq \eta/2) \tag{15}$$

From Theorem 10 we obtain that if $t = o(\log n)$ as $n \to \infty$ we have $\|\mu_{\mathcal{T}^{(n)}_t} - \mu_{\mathcal{N}^{(n)}_t}\|_{\mathrm{TV}} = o(1)$ when $n \to \infty$. We conclude that for any $\eta > 0$ there exists a sufficiently large n_0 such that, if $n \geq n_0$, then $\|\mu_{\mathcal{N}^{(n)}_t} - \mu_{\mathcal{T}^{(n)}_t}\|_{\mathrm{TV}} \leq \mathbb{P}(\mathcal{T}_t \neq \mathcal{N}_t) \leq \eta$. Using Proposition 3, it follows that for any $\eta > 0$ and large enough n:

$$\mathbb{P}(|z^{(n)}_\omega(t) - y_\omega(t)| \geq \eta) \leq \mathbb{P}(|z^{(n)}_\omega(t) - \bar{z}^{(n)}_\omega(t)| \geq \eta/4) + \mathbb{P}(|y^{(n)}_\omega(t) - y_\omega(t)| \geq \eta/2)$$

From Theorem 5, by choosing $x = n^{4/9}$, $s = 3$, and $\epsilon > 0$, we get

$$\mathbb{P}(|z^{(n)}_\omega(t) - y_\omega(t)| \geq \eta) \leq \mathbb{P}(|y^{(n)}_\omega(t) - y_\omega(t)| \geq \eta/2) + o(1)$$

where, we have applying jointly Corollary 27 and Corollary 29 in Appendix C of Supplementary Material to bound the third moment of V^v_t. Finally Propositions 6 and 7 guarantee that $|y^{(n)}_\omega(t) - y_\omega(t)| \to 0$ and $\mathbb{P}(|y^{(n)}_\omega(t) - y_\omega(t)| \geq \eta/2) = 0$ for large enough n. ∎

5.4. Asymptotic behavior on Configuration Block Model. In this section, we apply the bounds derived for Local Weak Convergence and Concentration Property in Section 4 to study the asymptotic ASD on a labeled configuration model with community structure, which is a key feature of many real systems. In particular, we consider a Configuration Block Model (CBM) with K communities of sizes $\{n_i\}_{i=1}^K$, which are mapped into corresponding classes with labels $\{a_i\}_{i=1}^K$.

When the maximum in/out degree of nodes is finite both Local Weak Convergence and Concentration property can be easily proven by taking the simple worst-case in which all nodes have in/out degree equal to the maximum in/out degree, so we will consider here a more challenging case, where in/out degree of nodes is allowed to scale with n.

However, to simplify the analysis, we assume no correlation between in-degree and out-degree of a node. As a consequence, the law of p is the same as the law of q, and $\zeta_{\omega|a,a'}(t) = y_{\omega|a}(t)$ (see Proposition 2).

Moreover, we assume that the number of edges established from a node of community i towards nodes of community j is independent for any pair (i,j), including the special case $i = j$, i.e., intra-community edges. Therefore, $p_{\mathbf{d},\mathbf{k}|a}$ factorizes into:

$$p_{\mathbf{d},\mathbf{k}|a} = \prod_{i\in\mathcal{A}} p^{\mathtt{in}}_{i,a}[d_i] \prod_{j\in\mathcal{A}} p^{\mathtt{out}}_{a,j}[k_j]$$

We will require that in/out degree sequences of the nodes, although possibly dependent on the network size n, generate empirical distributions $p^{\mathtt{in}}_{i,a}[d]$ and $p^{\mathtt{out}}_{a,j}[k]$ with a light tail, for any pair (i,a) or (a,j), thus having finite moments of any order.

In order to guarantee that the ASD over a network drawn uniformly at random from the CBM ensemble is well approximated by the solution of the ODE, it is sufficient that the terms in the upper bound derived in (9) and in Theorem 5 go to zero when $n \to \infty$. In the following, let $\mathcal{N}^{(n)} = (\mathcal{V}^{(n)}, \mathcal{E}^{(n)})$ be a sequence of networks, each one sampled from the corresponding CBM ensemble $G^{(n)} = G(\{n_i\}_{i=1}^K, p^{(n)}_{\mathbf{d},\mathbf{k},a})$, and, for $t \geq 0$, let $\mathcal{N}_t^{(n)}$ be the relevant neighborhood of a node v chosen uniformly at random from the node set $\mathcal{V}^{(n)}$. Moreover, let $\mathcal{T}_t^{(n)}$ be the sequence of GW processes with offspring distribution following law $p^{(n)}$. Finally, let $p^{(n)} \overset{n\to\infty}{\to} p$.

THEOREM 12. *Under above definitions and assumptions on the CBM ensemble, let $\mu_{\mathcal{N}_t^{(n)}}$ and $\mu_{\mathcal{T}_t^{(n)}}$ be the distributions of $\mathcal{N}_t^{(n)}$ and $\mathcal{T}_t^{(n)}$, respectively. For $t = o(\log n)$, we have $\|\mu_{\mathcal{T}_t^{(n)}} - \mu_{\mathcal{N}_t^{(n)}}\|_{\mathrm{TV}} = o(1)$ when $n \to \infty$.*

The proof is provided in Appendix C of the Supplementary Material.

COROLLARY 13. *Under above definitions and assumptions on the CBM ensemble, for $t \geq 0$, let $\boldsymbol{Z}(t)$ be the state vector of the ASD dynamics on $\mathcal{N}^{(n)}$ and $z_\omega^{(n)}(t) = \frac{1}{n}|\{v \in \mathcal{V} : Z_v(t) = \omega\}|$ be the fraction of state-ω adopters at time t. For $t = o(\log n)$ and any $\eta > 0$: $\mathbb{P}(|z_\omega^{(n)}(t) - y_\omega(t)| > \eta) = o(1)$ for $n \to \infty$, where $y_\omega(t)$ is the solution of (4) over a GW tree $\mathcal{T}_t$ with the asymptotic degree statistics p and q.*

The proof follows exactly the same lines of Corollary 11.

REMARK 4. *Note that the approach followed in the proof of Theorem 12 can be extended to a significantly more general class of labeled configuration graphs $\mathfrak{C}_{n,p}$. The key step of the approach pursued in Theorem 12 is to find a distribution meeting the following two constraints: i) it stochastically dominates the out-degree distribution of every class; ii) its tail is not too heavy, so that we can effectively bound the moments on the number of nodes in the corresponding GW truncated tree by exploiting (39) (see*

Appendix C of the Supplementary Material). Under such conditions we can conclude that $\|\mu_{\mathcal{T}_t} - \mu_{\mathcal{N}_t}\|_{\mathrm{TV}} = o(1)$ *when* $n \to \infty$. *In particular, the previous considerations apply every time the dominating distribution is a power law meeting the constraints in Assumption 1.*

6. The missing link between the mean-field ODE and the stationary regime. Our main result implies that the weak limit of every converging sequence of stationary probability distributions of ASDs on a large class of locally tree-like random graph converges, when the size of networks tends to infinity, to a deterministic limit which is the solution of a nonlinear ODE. This convergence is guaranteed over finite time horizons or for time windows that scale logarithmically with the size of the network. However, some form of convergence of the stationary regimes can be guaranteed, by combining the approximation derived in 4.2 with results in [6]. In particular, if the deterministic process has a unique limit point $y^\star$ and the sequence of invariant probabilities is tight, then the sequence of invariant probabilities will converge to the Dirac mass at $y^\star$.

In this section we consider regular random graphs, where all nodes have the same out-degree, and we analytically derive some interesting properties of the ODEs describing the temporal evolution of the system (see Proposition 2), for each of the three examples of ASD introduced in Section 2.2. In particular, we can characterize the equilibrium points of the system and their stability.

6.1. Ternary Linear Threshold Model (TLTM). We assume that all agents have the same out-degree k and symmetric thresholds, i.e., $k_v = k$ and $a_v^{\pm} = r$ for all $v \in \mathcal{V}$. In this case, there is a single class $a = r$ for which $p_{k|r} = q_{k|r} = 1$, and

$$\begin{aligned}
\Theta_1^{(r)}(\xi_1, \xi_{-1}, k - \xi_1 - \xi_{-1}) &= \mathbf{1}_{\{\xi_1 - \xi_{-1} \geq r\}} \\
\Theta_0^{(r)}(\xi_1, \xi_{-1}, k - \xi_1 - \xi_{-1}) &= \mathbf{1}_{\{\xi_1 - \xi_{-1} \in (-r, r)\}} \\
\Theta_{-1}^{(r)}(\xi_1, \xi_{-1}, k - \xi_1 - \xi_{-1}) &= \mathbf{1}_{\{\xi_1 - \xi_{-1} \leq -r\}}
\end{aligned}$$

From Proposition 2 we have that, given an initial distribution $(y_1(0), y_{-1}(0))$, the dynamics over the continuous-time branching process is described by

$$\begin{cases} \frac{\mathrm{d}y_1}{\mathrm{d}t} = \phi_+^{(k,r)}(y_1(t), y_{-1}(t)) - y_1(t), \\ \frac{\mathrm{d}y_{-1}}{\mathrm{d}t} = \phi_-^{(k,r)}(y_1(t), y_{-1}(t)) - y_{-1}(t), \end{cases} \tag{16}$$

and $y_0 = 1 - y_1 - y_{-1}$, where

$$\phi_+^{(k,r)}(x, z) = \sum_{v=0}^{\lfloor \frac{k-r}{2} \rfloor} \sum_{u=v+r}^{k-v} \binom{k}{u} \binom{k-u}{v} x^u z^v (1 - x - z)^{k-u-v} \tag{17}$$

and $\phi_-^{(k,r)}(x, z) = \phi_+^{(k,r)}(z, x)$.

Some analytical properties of the dynamical system can be deduced from the analysis of $\phi_+^{(k,r)}(x, z)$ and $\phi_-^{(k,r)}(x, z)$. In particular, we are interested in finding stationary points, and analysing their stability properties. Before presenting the main result we give some preliminary lemmas.

LEMMA 14. *Let* $\phi_+^{(k,r)}(x, z)$ *be as defined in* (17). *Then*

1. $\phi_+^{(k,r)}(x, z)$ *is non decreasing in* x *and strictly increasing if* $0 < r \leq k$;
2. $\phi_+^{(k,r)}(x, z)$ *is non increasing in* z *and strictly decreasing if* $0 \leq r < k$;

3. $\phi_+^{(k,r)}(0,z) = 0$ *for* $0 < r \leq k$, $\phi_+^{(k,0)}(x,0) = 1$, $\phi_+^{(k,r)}(1,0) = 1$, *for* $0 \leq r \leq k$;
4. $\nabla_x \phi_+^{(k,r)} = \sum_{v=0}^{\lfloor \frac{k-r}{2} \rfloor} \chi(v+r-1,v) x^{v+r-1} z^v (1-x-z)^{k-2v-r}$ *with* $\chi(u,v) = \binom{k}{u}\binom{k-u}{v}(k-u-v)$
5. $\nabla_z \phi_+^{(k,r)} = -\sum_{u=r}^{k} \chi(u,v') x^u z^{v'} (1-x-z)^{k-u-v'-1}$ *with* $v' = (k-u) \wedge (u-r)$.

The proof is trivial and we omit it for brevity.

PROPOSITION 15. *If* $2 \leq r < k$, *then*

1. *the equation* $\phi_+^{(k,r)}(x,0) = x$ *has three solutions* $\{0, x^\star, 1\}$ *with* $x^\star \in (0,1)$;
2. *for every* $x \in (x^\star, \bar{x})$ *with* $0 < x^\star < \bar{x} \leq 1$ *there exists a unique value* $z(x)$ *such that* $\phi_+^{(k,r)}(x, z(x)) = x$.
3. *the equation* $\phi_-^{(k,r)}(0,z) = z$ *has exactly three solutions* $\{0, z^\star, 1\}$ *with* $z^\star \in (0,1)$;
4. *for every* $z \in (z^\star, \bar{z})$ *with* $0 < z^\star < \bar{z} \leq 1$, *there exists a unique value* $x(z)$ *such that* $\phi_-^{(k,r)}(x(z), z) = z$.

Proof. If $2 \leq r < k$, the function $\phi_+^{(k,r)}(x,0)$ has a lazy-S-shaped graph, i.e., it is increasing, with a unique inflection point at $\widetilde{x} = (r-1)/(k-1)$, it is convex on the left-hand side of $\widetilde{x}$ and concave on the right-hand side of $\widetilde{x}$ (see Lemma 4 in [34]). From this fact and the observation that $\phi_+^{(k,r)}(0,0) = 0$ and $\phi_+^{(k,r)}(1,0) = 1$ we get the assertion at Point 1. It can also be proved that $x^\star \in [(r-1)/k, r/k]$. Denoting $F(x,z) = \phi_+^{(k,r)}(x,z) - x$ and observing that $F(x^\star, 0) = \phi_+^{(k,r)}(x^\star, 0) - x^\star = 0$ and $\nabla_z F(x^\star, 0) \neq 0$ (see expression in Point 14 of Lemma 14), the statement in Point 2. is obtained by the implicit function theorem [5]. Point 3. and 4. are straightforward from the relation $\phi_-^{(k,r)}(x,z) = \phi_+^{(k,r)}(z,x)$. ∎

In Figure 2 the functions $z(x)$ such that $\phi_+^{(k,r)}(x, z(x)) = x$ and $x(z)$ such that $\phi_-^{(k,r)}(x(z), z) = z$ are depicted for threshold values $r = 2$ (left) and $r = 3$ (right) and degree $k = 10$. In addition to stationary points $\{(0,0), (x^\star, 0), (1,0), (0,z^\star), (0,1))\}$, as derived in Lemma 14, extra stationary points are placed at intersections between curves $z(x)$ and $x(z)$. In the specific case with $k = 10$ it can be noticed that if $r = 2$ then there are two additional stationary points.

The following proposition gives sufficient conditions guaranteeing that the set of stationary points only contains the trivial points $\{(0,0), (x^\star, 0), (1,0), (0,z^\star), (0,1))\}$.

PROPOSITION 16. *If* $r \geq (k+1)/2$, $\{(0,0), (x^\star, 0), (1,0), (0,z^\star), (0,1))\}$ *are the only fixed points of the system.*

Proof. Notice that $\phi_+^{(k,r)}(x,x) = \phi_-^{(k,r)}(x,x) \leq \sum_{u=r}^{k} \binom{k}{u} x^u (1-x)^{k-u} = \varphi^{(k,r)}(x)$. If $r \geq (k+1)/2$, we have $\varphi^{(k,r)}(x) < x$ for all $x < 1/2$ and, consequently, $\phi_+^{(k,r)}(x,x) < x$ from which we conclude the assertion. ∎

The following corollary can be proved by linearization.

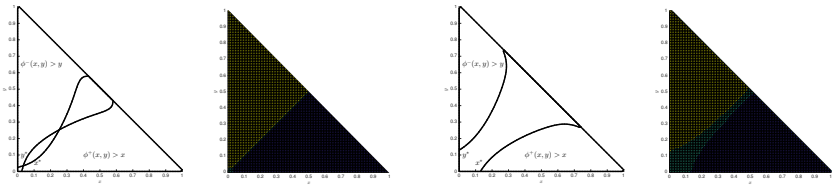
COROLLARY 17. *The following properties hold*

1. $(0,0)$ *is a locally stable stationary point for* $2 \leq r \leq k$, *unstable if* $r = 1$.
2. $(0,1)$ *and* $(1,0)$ *are locally stable stationary points for* $1 \leq r < k$, *unstable otherwise.*
3. $(x^\star, 0)$ *and* $(0, z^\star)$ *are unstable stationary points.*
4. *The set of points* $\{(x,z) \in \mathbb{R}^2 : x = z\}$ *is invariant.*

PROPOSITION 18. *Let us consider the dynamical system in* (16). *Then there are no periodic orbits.*

Proof. The set of points $M_1 = \{(x, z) : x = 0\}$, $M_2 = \{(x, z) : z = 0\}$, and $M_3 = \{(x, z) : x = z\}$ are positively invariant. From Proposition 16 and by applying the Poincaré-Bendixson theorem we conclude that there are no periodic orbits. ∎

The basins of attraction for the ODE in (16) are shown in Figure 2 for degree $k = 10$ and threshold $r \in \{2, 3\}$. Basins are evaluated numerically, by solving the ODE system for a wide set of initial conditions. More specifically, for each initial condition x_0, z_0 the color in the picture represents the asymptotically stable equilibrium point (yellow for (0,1), green for (0,0) and blue for state $(1, 0)$) to which the trajectory tends. As to be expected by Corollary 17, we have that points $(0, 0)$, $(0, 1)$ and $(1, 0)$ are locally stable stationary points.



(a) (b)

FIGURE 2. *Basins of attraction for the ODE in* (16)*:* (a) $k = 10$ *and* $r = 2$ (b) $r = 3$

6.2. Binary Response with Coordinating and Anti-coordinating agents (BRCA). We again consider the homogenous case in which all nodes have out-degree k. For simplicity, in the following, we will suppose that k is odd, in order to avoid ties, although the extension to the case of even k is straightforward. Let α be the fraction of coordinating nodes. From Proposition 2, the evolution of the node states is governed by the following ODE:

$$\frac{\mathrm{d}y_1(t)}{\mathrm{d}t} = \alpha\phi_1\left(y_1(t)\right) + (1-\alpha)\left(1 - \phi_1\left(y_1(t)\right)\right) - y_1(t), \tag{18}$$

where $\phi_1\left(y_1(t)\right)$ is the probability with which a coordinating node enters state 1, given by

$$\phi_1\left(y_1\right) = \sum_{k_1=\lceil k/2 \rceil}^{k} \binom{k}{k_1} y_1^{k_1}(1-y_1)^{k-k_1}. \tag{19}$$

The above ODE derives from the fact that the probability of stepping to state $+1$ for an anti-coordinating node is equal to the probability of stepping to state -1 for a coordinating node.

PROPOSITION 19. *Let us consider the dynamical system in* (18)*. There exists* $\alpha_{\mathrm{th}} \in (0, 1)$ *such that*

- *if* $\alpha \leq \alpha_{\mathrm{th}}$*, then the only stationary point is* $y_1 = 1/2$*, which is stable and has a basin of attraction equal to* $[0, 1]$*;*
- *if* $\alpha > \alpha_{\mathrm{th}}$*, then there are three stationary points:* $z_1 = 1/2$ *is unstable and the other two are symmetric with respect to* $z_1 = 1/2$ *and stable, with basins of attraction* $[0, 1/2)$ *and* $(1/2, 1]$*.*

Proof. Let us define $\ell_1\left(y_1(t)\right)$ the RHS of (18). It is to verify that $\phi_1\left(1/2\right) = 1/2$, $\ell_1\left(1/2\right) = 0$, so that $y_1 = 1/2$ is a stationary point for every α.

The derivative of ℓ_1 with respect to y_1 is given by:

$$\frac{\mathrm{d}\ell_1}{\mathrm{d}y_1} = (2\alpha-1)\frac{\mathrm{d}\phi_1}{\mathrm{d}y_1} - 1 = (2\alpha-1)k\binom{k-1}{\lfloor \frac{k}{2} \rfloor}\left(y_1(1-y_1)\right)^{\lfloor \frac{k}{2} \rfloor} - 1$$

From the above equation, it turns out that we have two distinct regimes, according to whether $\alpha \leq \alpha_{\text{th}}$ or $\alpha > \alpha_{\text{th}}$, with $\alpha_{\text{th}} = \frac{1}{2}\left(1 + \frac{2^{k-1}}{k\binom{k-1}{\lfloor \frac{k}{2} \rfloor}}\right)$

- If $\alpha \leq \alpha_{\text{th}}$, then $\frac{\mathrm{d}\ell_1}{\mathrm{d}y_1} \leq 0$ for $0 < y_1 < 1$. It then turns out that the only stationary point is $y_1 = 1/2$, which is stable and has a basin of attraction equal to $[0, 1]$.
- If $\alpha > \alpha_{\text{th}}$, then $\frac{\mathrm{d}\ell_1}{\mathrm{d}y_1}$ has two zeros, symmetric with respect to $z_1 = 1/2$, in the solutions of the quadratic equation $y_1(1-y_1) = \left((2\alpha-1)k\binom{k-1}{\lfloor \frac{k}{2} \rfloor}\right)^{-\lfloor \frac{k}{2} \rfloor}$ Because of that, there are three stationary points, out of which the one in $z_1 = 1/2$ becomes unstable. The other two are symmetric with respect to $z_1 = 1/2$, i.e., $y^{\pm} = 1/2 \pm \epsilon$, and stable, with basins of attraction $[0, 1/2)$ and $(1/2, 1]$. Figure 3 shows the value of ϵ as a function of k and α. □

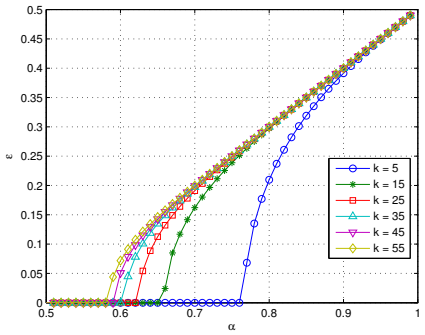


FIGURE 3. *Position of the stable stationary points as a function of the fraction of cooperative nodes and the common node degree.*

6.3. Evolutionary Roshambo Game (ERG). In this section, we consider the ERG dynamics on a regular graph with $k_v = k$ and $b = c/2$.

Let y_ω, $\omega \in \{R, P, S\}$ be the fraction of nodes in state ω. Moreover, let π_ω, $\omega \in \{R, P, S\}$ be the probability that a given node, when activated, switches to state ω thanks to its neighbors' states. More precisely, for $\omega = R$:

$$\begin{aligned}\pi_R =& \mathbb{P}\left\{\xi_S > \frac{k}{3},\ \xi_P < \frac{k}{3}\right\} + \frac{1}{2}\mathbb{P}\left\{\xi_S = \frac{k}{3},\ \xi_P < \frac{k}{3}\right\} \\ &+ \frac{1}{2}\mathbb{P}\left\{\xi_S > \frac{k}{3},\ \xi_P = \frac{k}{3}\right\} + \frac{1}{3}\mathbb{P}\left\{\xi_S = \frac{k}{3},\ \xi_P = \frac{k}{3}\right\}\end{aligned} \tag{20}$$

and analogously for π_S and π_P.

When it is activated, a node in state ω changes state with probability $1 - \pi_\omega$, while a node not in state ω switches to that state with probability π_ω, by definition. We thus have:

$$\frac{dy_\omega}{dt} = (1 - y_\omega)\pi_\omega - y_\omega(1 - \pi_\omega),\ \ \omega \in \{R, P, S\} \tag{21}$$

At an equilibrium point, $\frac{dy_\omega}{dt} = 0$, so that $y_\omega = \pi_\omega$, $\omega \in \{R, P, S\}$.

PROPOSITION 20. *The only equilibrium point for the ERG dynamics on a regular graph is* $(y_R, y_P, y_S) = (1/3, 1/3, 1/3)$. *Moreover this equilibrium point is locally asymptotically stable.*

The technical proof of local stability is given in Appendix D of the Supplementary Material. Unfortunately, even if a numerical analysis of the ODE dynamical system provides some evidence of the global stability of point $(1/3, 1/3, 1/3)$ for moderate values of the degree k, we are not able to analytically prove its global stability in the general case. We can obtain only ad-hoc proofs for few specific and small values of k, but we decided to avoid reporting them, here. However, the dynamics are slightly simplified when we set $b = 0$. We present the detailed analysis in the following section.

Differential equations driving the dynamics of the system are given by (21) and when $b = 0$ the expressions of π_R, π_S and π_P become

(22)
$$\pi_R = \mathbb{P}\{\xi_S > \max(\xi_P, \xi_R)\} + \frac{1}{2}\mathbb{P}\{\xi_S = \max(\xi_P, \xi_R) > \min(\xi_P, \xi_R)\} + \frac{1}{3}\mathbb{P}\{\xi_S = \xi_P = \xi_R\}$$

and analogously for π_S and π_P.

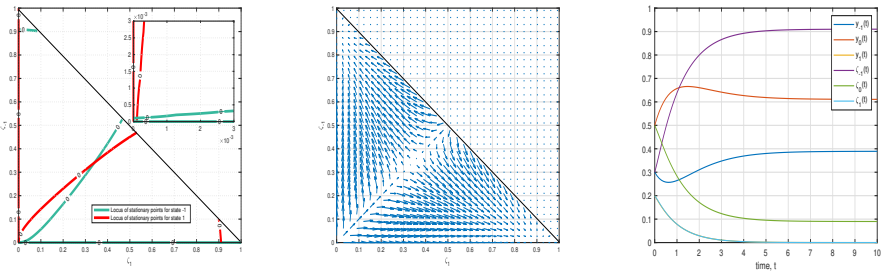
PROPOSITION 21. *The only equilibrium point for the ERG dynamics with* $b = 0$ *on a regular graph is* $(y_R, y_P, y_S) = (1/3, 1/3, 1/3)$. *Furthermore* $(y_R, y_P, y_S) = (1/3, 1/3, 1/3)$ *is globally stable.*

7. Numerical experiments. In this section we present a few interesting cases in which we run our examples of ASD in large but finite networks, considering both synthetically generated graphs and a real-world social network. Results are obtained either by numerically solving the differential equation in (4), or by running detailed Monte-Carlo simulations using an ad-hoc event-driven simulator.

In particular, we will focus on the online social network Epinions, for which a popular snapshot is publicly available at the Stanford Large Network Dataset Collection [25]. The online social network Epinions.com is a consumer review website where the users can review different kind of items with the purpose of rating hundred thousand products and ranking the reviewers to be trusted. The available dataset contains the who-trust-whom relationships of all the members, operating from 1999 until 2014. The network consists of $|\mathcal{V}| = 75879$ nodes and $|\mathcal{E}| = 508837$ directed edges, it is highly connected and contains cycles. The average clustering coefficient is 0.1378. The maximum in-degree is 3035, maximum out-degree is 1801, the average in/out-degree is 6.7. In and out degrees follow an approximate power law distribution with exponent 1.7.

7.1. TLTM on the Epinions graph. We assume that all nodes have symmetric threshold $a_v^{\pm} = 2$. In Figure 4, we show on the left the loci of the stationary solution of (3) in the plane (ζ_{-1}, ζ_1), and on the right an arrow plot representing the gradient of the system of the two ODEs in each possible point for which $\zeta_{-1} + \zeta_1 \leq 1$. As it can be seen from the combinations of the two plots, there are three stable stationary points, which are located at $(\zeta_{-1}, \zeta_1) = (0, 0)$, $(\zeta_{-1}, \zeta_1) = (\bar{\zeta}, 0)$ and $(\zeta_{-1}, \zeta_1) = (0, \bar{\zeta})$, with $\bar{\zeta} \simeq 0.91$ while there are four *unstable* stationary points at values $(\zeta_{-1}, \zeta_1) = (\widetilde{\zeta}, 0)$, $(\zeta_{-1}, \zeta_1) = (0, \widetilde{\zeta})$, $(\zeta_{-1}, \zeta_1) = (\widetilde{\zeta}_1, \widetilde{\zeta}_1)$ and $(\zeta_{-1}, \zeta_1) = (\widetilde{\zeta}_2, \widetilde{\zeta}_2)$, where $\widetilde{\zeta} \simeq 1.02 \times 10^{-4}$, $\widetilde{\zeta}_1 \simeq 1.066 \times 10^{-4}$ and $\widetilde{\zeta}_2 \simeq 0.341$. The arrow plot allows also to verify that the boundary of the two basins of attraction for the stationary points coincides with the line $\zeta_{-1} = \zeta_1$ for $\widetilde{\zeta}_1 \leq \zeta_1 \leq 1/2$.

In Figure 4, we show the evolution over time of the variables $\zeta_\omega(t)$ and $y_\omega(t)$, $\omega \in \{-1, 0, 1\}$, obtained through the numerical solution of ODEs in equations (3)-(4).



(a) (b) (c)

FIGURE 4. *TLTM in the Epinions social network with symmetric thresholds* $r^{\pm} = 2$*.: (a) Stationary solution of ODEs in* (3)*. (b) Gradient of ODE system in* (3) *(c) Transient solution of ODEs in* (3)-(4)*.*

In particular, $y_\omega(t)$ represents the fraction of nodes in state ω at time t, while $\zeta_\omega(t)$ represents the fraction of edges connected to a node in state ω at time t. At time $t = 0$, the fraction of nodes of any degree in states $-1, 0, 1$ is $0.3, 0.5, 0.2$, respectively. As it can be seen, the fraction of nodes in state 1 decreases exponentially with time (the curve of $\zeta_1(t)$ is superimposed on that of $y_1(t)$). Instead, the curve for $\zeta_{-1}(t)$ reaches the fixed point $\overline{\zeta} = 0.91$, as predicted by Figure 4. The fixed point for $y_{-1}(t)$ is lower, at about 0.39, implying that the fraction of nodes with higher degree that asymptotically reach state -1 is larger than for nodes with lower degree.

7.2. TLTM on the Configuration Block Model. In this subsection, we show results for a network with $n = 10^7$ nodes divided into two equal-size communities (classes), with size $n/2 = 5 \cdot 10^6$. We consider the TLTM with symmetric thresholds $a_v^{\pm} = 2$. We put 10^5 seeds in state 1 in community 1, and 10^5 seeds in state -1 in community 2. The out-degree distribution of the first class is given by $p_{(k_1,k_2)|1} = p_{11}(k_1)p_{12}(k_2)$ where

$$p_{11}(k_1) = \binom{n/2-1}{k_1} p_A^{k_1}(1-p_A)^{1-k_1}, \qquad p_{12}(k_2) = \binom{n/2}{k_2} p_B^{k_2}(1-p_B)^{1-k_2}$$

p_A and p_B being chosen so that the average degree toward community 1 is 20, while the average degree toward community 2 is 6. Analogously, the out-degree distribution of the second class is given by $p_{(k_1,k_2)|2} = p_{21}(k_1)p_{22}(k_2)$ where

$$p_{21}(k_1) = \binom{n/2}{k_1} p_C^{k_1}(1-p_C)^{1-k_1}, \qquad p_{22}(k_2) = \binom{n/2-1}{k_2} p_A^{k_2}(1-p_A)^{1-k_2}$$

p_C being chosen so that the average degree toward community 1 is 5. The network shows a slight asymmetry, since nodes in community 1 have slightly more edges directed towards nodes of community 2 than viceversa.

Figure 5 compares the fraction of nodes in state 1 and -1 in either community, averaged across 100 simulation runs, against analytic results. We observe very good agreement between analysis (thin curves) and simulation (thick curves). Interestingly, in the beginning we have two weakly interfering percolation processes in the two communities, producing a significant increase of nodes in state -1 in community 2, and a significant increase of nodes in state 1 in community 1. However, the percolation process in community 2 grows faster, because nodes in community 2 receive less influence from nodes in community 1 than viceversa. As a consequence, the percolation process of nodes in state -1 eventually invades also community 1, while nodes in state 1 vanish to zero throughout the network.

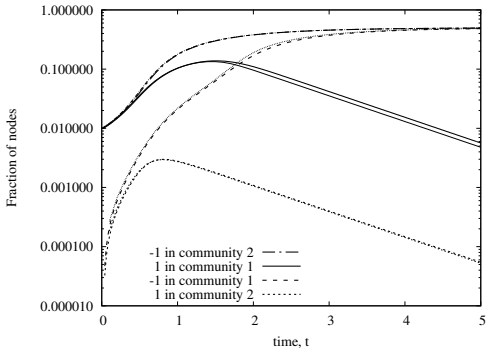


FIGURE 5. *Evolution over time of the fraction of nodes in states 1 and -1 in the two communities, according to analysis (thin curves) and simulation (average of 100 runs) (thick curves).*

7.3. BRCA on the regular graph. Here we consider a simple regular graph where all nodes have fixed out-degree $k = 21$ and fixed in-degree $d = 21$. Note that, since the out-degree is odd, the best response is always deterministic (i.e., there are no ties). We perform a single simulation run with $n = 10^5$, with the following initial configuration: 30,000 coordinating nodes in state 1, 10,000 coordinating nodes in state -1, 40,000 anti-coordinating nodes in state 1, 20,000 anti-coordinating nodes in state -1.

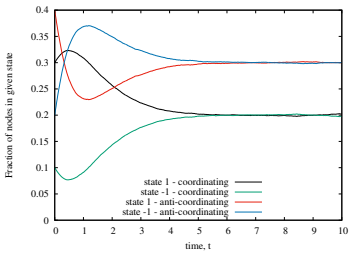


FIGURE 6. *Evolution over time of the fraction of nodes in the ABRD game, according to a single simulation run.*

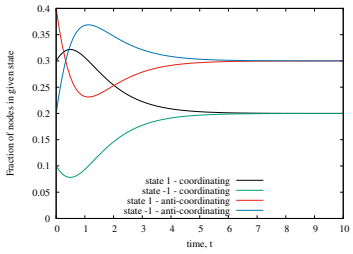


FIGURE 7. *Evolution over time of the fraction of nodes in the ABRD game, according to analysis.*

In Figures 6 and 7 we show the fraction of nodes in each of the possible states as function of time, according to simulation and analysis, respectively. We notice a perfect agreement between analytical prediction and simulation. Small fluctuations around the equlibrium configuration appear on the (single) simulation sample path.

To assess the degree of concentration of the process around its average, we carried out the following experiment: we performed 400 runs of the system, where the variability across runs is due to multiple reasons: i) the network topology generated by the configuration model; ii) the initial selection of nodes in the various states; iii) the temporal dynamics of the process (Poisson clocks). We then sampled the system with time granularity $\Delta t = 0.01$, and at each time instant we evaluated the average, the minimum, and the maximum fraction of nodes in each state, across the 400 runs. In Figure 8 we show the results of the above experiment for the fraction of nodes in state -1 (either coordinating or anti-coordinating), with $n = 10^3$ (top-left), $n = 10^4$

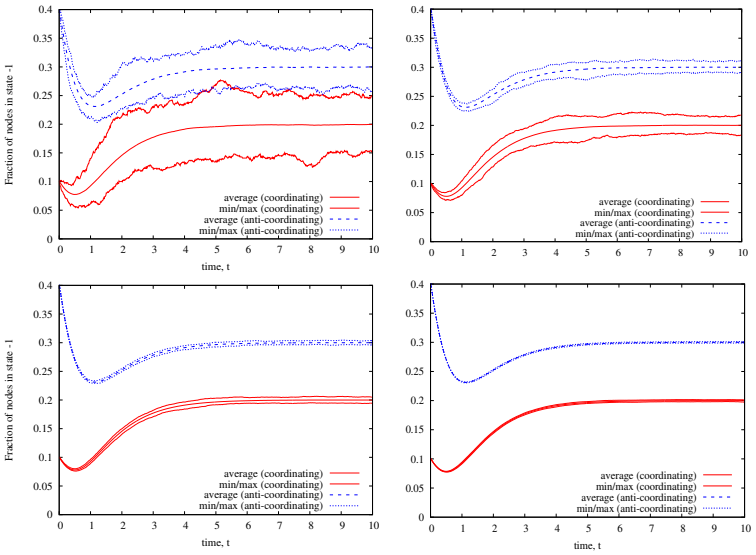


FIGURE 8. *Evolution over time of the fraction of nodes in state -1 in the ABRD game, across 400 simulation runs.*

(top-right), $n = 10^5$ (bottom-left), $n = 10^6$ (bottom-right). Thin curves above and below the thicker line (denoting the average), correspond to maximum and minimum values. Results for the fraction of nodes in state 1 are not shown, and they exhibit a similar variability. We observe that results become more concentrated passing from $n = 10^3$ to $n = 10^6$ nodes.

7.4. BRCA on the Epinions graph. We now investigate BRCA dynamics on the Epinions graph with a fraction of coordinating nodes equal to $\alpha = 0.7$, evenly distributed among nodes of any degree. Our main goal in this section will be to understand better the origin of possible discrepancies between analysis and simulation.

A numerical analysis of (3)-(4) shows that, similarly to the regular case, the stationary points for the Epinions degree distribution are three, out of which the one in $\zeta_1 = 1/2$ is unstable. The other two are positioned at $\zeta_1 \simeq 0.33$ and $\zeta_1 \simeq 0.67$ and are stable. Such stationary points correspond to fractions of nodes in state 1 given by $y_1 \simeq 0.426$ and $y_1 \simeq 0.574$, respectively.

Let us now move to the time evolution of fractions of nodes in a given state. We consider the following initial condition: 42% of coordinating nodes in state 1, 28% of coordinating nodes in state -1, 10% of anti-coordinating nodes in state 1, 20% of anti-coordinating nodes in state -1. The left plot in Figure 9 shows the fraction of nodes in each possible state as function of time, comparing simulation (thick curves) and analysis (thin curves). For simulations, we have plotted the average of 400 runs, where in each run we randomly select a different seed set. We observe that, after a very similar initial behavior, simulations results tend to a different equilibrium point with respect to analysis, as one can see by looking at the fraction of nodes at time $t = 10$. We have identified two main reasons for the observed discrepancies: i) the first one due to the fact that the structure of the Epinions graph is not captured by the configuration model; ii) the second one due to the fact that the network size is not large enough to converge to a unique equilibrium point across different runs, due

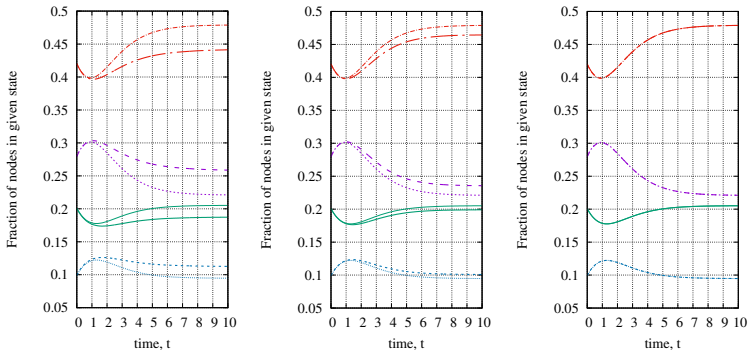


FIGURE 9. *BRCA dynamics: evolution over time of the fraction of nodes in each possible state, according to simulation (thick curves) and analysis (thin curves). From top curve to bottom curve: coordinating nodes in state 1, coordinating nodes in state -1, anti-coordinating nodes in state 1, anti-coordinating nodes in state -1. From left to right plot: original Epinions graph (left); configuration model matched to the Epinions graph (middle); configuration model matched to the Epinions graph, with* $10n$ *nodes (right).*

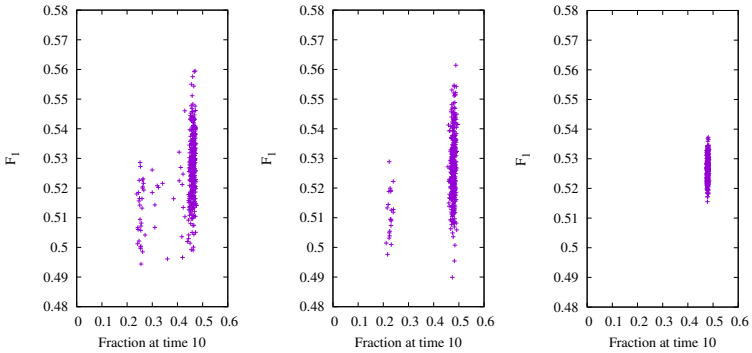


FIGURE 10. F_1 *metric vs fraction of coordinating node in state 1 (sampled at time* $t = 10$*): original Epinions graph (left); configuration model matched to the Epinions graph (middle); configuration model matched to the Epinions graph, with* $10n$ *nodes (right).*

to random effects.

To separate out the impact of the above two reasons, we have performed the following experiments. First, we have run simulations in which, in each run, we randomly reshuffle the edges while maintaining the same node statistics. Note that, by so doing, we generate graph according to the configuration model matched to the Epinions graph. The results of this experiment are shown in the middle plot of Fig. 9. As expected, now we observe a much better agreement between analysis and simulation. Still, there are non negligible discrepancies in the final fraction of nodes in each possible state.

An in-depth inspection of simulation results revealed that about 5% of simulation runs tend to a completely different equilibrium than the remaining 95% of the runs. This fact is illustrated by the middle plot in Figure 10, where we have put a mark for each of the 400 runs, showing on the x axes the fraction of coordinating nodes in state 1, sampled at time $t = 10$. For each run, we also computed the fraction of coordinating nodes that would transit to state 1 if their clock would fire at time $t = 0$.

This fraction, denoted as F_1, is plotted on the y axes, and it is meant to capture a possible initial bias towards entering state 1, due to the initial network condition, which is especially dependent on the random seed allocation.

We can observe that simulations converge to two main equilibria, and that simulations runs where the final fraction of coordinating nodes in state 1 is smaller (around 0.2) have smaller values of the F_1 metric specified above, suggesting that the initial seed allocation is the main responsible for driving the system into a different configuration. We emphasize that a similar behavior can be observed also on the original Epinions graph, for which an analogous investigation of single simulation runs produced the left plot in Fig. 10.

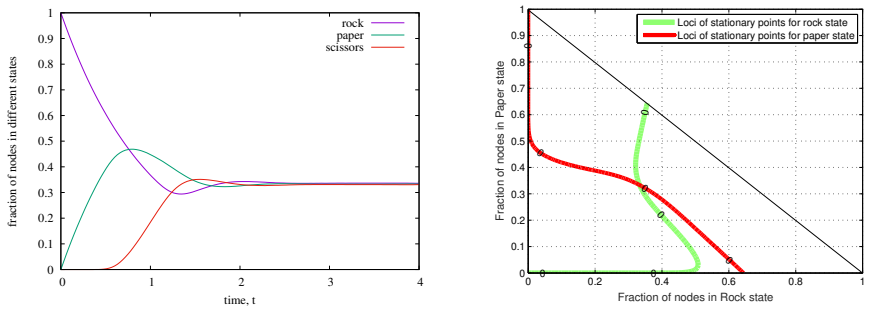
To remove the bi-stable outcome of simulations, we performed the following additional experiment: we considered again the node statistics of the original Epinions graph, but this time we generated graphs of size ten times larger than the original one (i.e., with $n = 758790$ nodes), using the configuration model.

The right plot in Fig. 10 shows that simulation results are now much more concentrated around a unique equilibrium. Moreover, in the right plot of Fig. 9 we observe an almost perfect agreement between analysis and simulation for the evolution over time of the fraction of nodes in each possible state (in this plot thick and thin curves are essentially overlapped and thus indistinguishable).

We conclude that, when our analytic approach is used to predict the behavior of ASD dynamics on realistic (finite) graph, one must be aware of two main sources of errors: one due to the fact that real graphs are not completely described by the configuration model; the other due to the fact that, in finite graphs, randomness can possibly drive the system to different equilibria, especially when initial conditions are close to the border of the attraction basin of the expected equilibrium. However, our experiments with the Epinions graph suggest that our approach has remarkable accuracy even in realistic graphs. Moreover, when the number of nodes exceeds, say, one million, results are sufficiently concentrated around their average to justify our mean-field approach, at least for the types of ASD dynamics that we have examined so far.

7.5. ERG on the regular graph. Here, we show numerical results for the ERG dynamics with $b = c/2$ on a simple regular graph where all nodes have fixed out-degree $k = 50$ and fixed in-degree $d = 50$, and all nodes starts in the same state (the rock state). In simulation, we take 1 million nodes, and perform a single run of the system. In Figure 11 we show the fraction of nodes in each of the three states as function of time, according to simulation. In Figure 11.(a), in the plane representing the fraction of nodes in the paper and rock states, we show the loci of the stationary solutions for (4) with $\omega \in \{R, P\}$. As shown in Section 6.3, the only stationary solution for the Roshambo game is $y_\omega = 1/3$, $\omega \in \{R, P, S\}$, which is the only point lying in the intersection of the curves shown in Figure 11. Starting from $y_R = 1$, as it is shown in Figure 11.(b), there is a rapid increase in the fraction of nodes in the paper state, since it is the best response. When the node populations in the rock and paper states become of equal size, then nodes in the scissors state start to appear (in the plane, the trajectory deviates from the boundary $y_R + y_P = 1$). The three populations then tend to the equilibrium point in which they have the same size.

8. Conclusions. In this paper we have proposed a mathematical framework showing that general semi-anonynoums dynamics in large scale random graphs converge to the solution of ordinary differential equations, allowing fast numerical prediction of the transient behavior of many cascading processes in complex systems and, in



(a) (b)

FIGURE 11. *ERG dynamics: (a) Evolution over time of the fraction of nodes in states R,P,S. (b) Stationary solution of ODEs*

some cases, analytical estimation of their points of equilibrium. With respect to existing literature, we have extended the above mean-field approximation along several directions: i) asynchronous node activation; ii) arbitrary semi-anonymous dynamics, including noisy best-response and class-dependent behavior; iii) general random graph exhibiting a local tree-structure, possibly mixing heterogeneous nodes and unbounded in/out degrees. Our main contribution is a rigorous mathematical proof of convergence, which requires a careful combination of many independent results related to the different framework components. Despite the generality of our approach, we have not considered important variations of semi-anonynoums dynamics such as non-reversible transitions. Moreover, it remains still largely open how to analytically characterize in a tractable way the behavior of ASD in undirected network.

REFERENCES

[1] D. Acemoglu, A. Ozdaglar, and E. Yildiz. Diffusion of innovations in social networks. In *2011 50th IEEE Conference on Decision and Control and European Control Conference*, pages 2329–2334, 2011.

[2] E. M. Adam, M. A. Dahleh, and A. Ozdaglar. On the behavior of threshold models over finite networks. In *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*, pages 2672–2677, 2012.

[3] N. Alon and J.H. Spencer. *The Probabilistic Method*. Wiley Series in Discrete Mathematics and Optimization. Wiley, 2004.

[4] Hamed Amini, Rama Cont, and Andreea Minca. Resilience to contagion in financial networks. *Mathematical Finance*, 26(2):329–365, 2016.

[5] Tom M. Apostol. *Calculus, Vol. 2: Multi-Variable Calculus and Linear Algebra with Applications to Differential Equations and Probability*. J. Wiley, New York, 1967.

[6] M. Benaim and J.-Y. Le Boudec. On Mean Field Convergence and Stationary Regime. *ArXiv e-prints*, November 2011.

[7] Lawrence Blume. The statistical mechanics of best-response strategy revision. *Games and Economic Behavior*, 11(2):111–145, 1995.

[8] Lawrence E. Blume. The statistical mechanics of strategic interaction. *Games and Economic Behavior*, 5(3):387 – 424, 1993.

[9] Charles Bordenave, David D. McDonald, and Alexandre Proutière. A particle system in interaction with a rapidly varying environment: Mean field limits and applications. *Networks Heterog. Media*, 5:31–62, 2010.

[10] Joris Broere, Vincent Buskens, Jeroen Weesie, and Henk Stoof. Network effects on coordination in asymmetric games. *Scientific Reports*, 7, 12 2017.

[11] Damon Centola, Victor M. Eguiluz, and Michael W. Macy. Cascade dynamics of complex propagation. *Physica A: Statistical Mechanics and its Applications*, 374(1):449–456, 2007.

[12] T. K. Chalker, A. P. Godbole, P. Hitczenko, J. Radcliff, and O. G. Ruehr. On the size of a random sphere of influence graph. *Advances in Applied Probability*, 31(3):596?609, 1999.
[13] L. Comtet. *Advanced Combinatorics*. Reidel, Dordrecht, 1974.
[14] Easley David and Kleinberg Jon. *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. Cambridge University Press, USA, 2010.
[15] Y. Desmond Zhong, V. Srivastava, and N. E. Leonard. On the linear threshold model for diffusion of innovations in multiplex social networks. In *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*, pages 2593–2598, Dec 2017.
[16] Rick Durrett. *Probability: Theory and Examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 4 edition, 2010.
[17] Glenn Ellison. Learning, local interaction, and coordination. *Econometrica*, 61(5):1047–71, 1993.
[18] Selman Erol, Francesca Parise, and Alexander Teytelboym. Contagion in graphons. In Péter Biró, Jason Hartline, Michael Ostrovsky, and Ariel D. Procaccia, editors, *EC '20: The 21st ACM Conference on Economics and Computation, Virtual Event, Hungary, July 13-17, 2020*, page 469. ACM, 2020.
[19] William Feller. *An Introduction to Probability Theory and Its Applications*, volume 1. Wiley, January 1968.
[20] Mark Granovetter. Threshold models of collective behavior. *American Journal of Sociology*, 83(6):1420–1443, 1978.
[21] Matthew O. Jackson. *Social and Economic Networks*. Princeton University Press, USA, 2008.
[22] David Kempe, Jon Kleinberg, and Éva Tardos. Maximizing the spread of influence through a social network. In *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '03, pages 137–146, New York, NY, USA, 2003. ACM.
[23] Thomas Kurtz. Solutions of ordinary differential equations as limits of pure jump markov processes. *Journal of Applied Probability*, 7:49–58, 04 1970.
[24] Marc Lelarge. Diffusion and cascading behavior in random networks. *Games and Economic Behavior*, 75(2):752 – 775, 2012.
[25] Jure Leskovec and Andrej Krevl. SNAP Datasets: Stanford large network dataset collection. http://snap.stanford.edu/data, June 2014.
[26] David A. Levin, Yuval Peres, and Elizabeth L. Wilmer. *Markov chains and mixing times*. American Mathematical Society, 2006.
[27] Y. Lim, A. Ozdaglar, and A. Teytelboym. A simple model of cascades in networks. *submitted*, 2016.
[28] Christopher W. Lynn and Daniel D. Lee. Maximizing influence in an ising network: A mean-field optimal solution. In *Proceedings of the 30th International Conference on Neural Information Processing Systems*, NIPS'16, pages 2495–2503, USA, 2016. Curran Associates Inc.
[29] Andrea Montanari and Amin Saberi. The spread of innovations in social networks. *Proceedings of the National Academy of Sciences*, 107(47):20196–20201, 2010.
[30] Stephen Morris. Contagion. *The Review of Economic Studies*, 67(1):57–78, 2000.
[31] M. E. J. Newman. Spread of epidemic disease on networks. *Phys. Rev. E*, 66:016128, Jul 2002.
[32] Pouria Ramazi, James R. Riehl, and Ming Cao. Networks of conforming or nonconforming individuals tend to reach satisfactory decisions. *Proceedings of the National Academy of Sciences*, 113:12985 – 12990, 2016.
[33] Everett M. Rogers. *Diffusion of innovations*. Free Press, New York, NY [u.a.], 5th edition, 08 2003.
[34] W.S. Rossi, G. Como, and F. Fagnani. Threshold models of cascades in large-scale networks. *IEEE Transactions on Network Science and Engineering*, November 2017.
[35] Thomas C. Schelling. *Micromotives and Macrobehavior*. W. W. Norton & Company, October 1978.
[36] Jameson L. Toole, Meeyoung Cha, and Marta C. Gonzlez. Modeling the adoption of innovations in the presence of geographic and media influences. *PLOS ONE*, 7(1):1–9, 01 2012.
[37] Petter Trnberg. Echo chambers and viral misinformation: Modeling fake news as complex contagion. *PLOS ONE*, 13(9):1–21, 09 2018.
[38] Soroush Vosoughi, Deb Roy, and Sinan Aral. The spread of true and false news online. *Science*, 359(6380):1146–1151, 2018.
[39] Shaolin Wang. Cascading model of infrastructure networks based on complex network. *Journal of Networks*, 8:1448–1454, 06 2013.
[40] Duncan J. Watts. A simple model of global cascades on random networks. *Proceedings of the National Academy of Sciences*, 99(9):5766–5771, 2002.