Bootstrap percolation on the stochastic block model

Submitted to Bernoulli

Bootstrap percolation on the stochastic block model

GIOVANNI LUCA TORRISI[1], MICHELE GARETTO[2] and EMILIO LEONARDI[3]

[1]*CNR-IAC, Roma, Italy. E-mail:* giovanniluca.torrisi@cnr.it
[2]*Università di Torino, Torino, Italy. E-mail:* michele.garetto@unito.it
[3]*Politecnico di Torino, Torino, Italy. E-mail:* emilio.leonardi@polito.it

We analyze the bootstrap percolation process on the stochastic block model (SBM), a natural extension of the Erdős–Rényi random graph that incorporates the community structure observed in many real systems. In the SBM, nodes are partitioned into two subsets, which represent different communities, and pairs of nodes are independently connected with a probability that depends on the communities they belong to. Under mild assumptions on the system parameters, we prove the existence of a sharp phase transition for the final number of active nodes and characterize the sub-critical and the super-critical regimes in terms of the number of initially active nodes, which are selected uniformly at random in each community.

Keywords: Bootstrap Percolation; Random Graphs; Stochastic Block Model.

1. Introduction

Bootstrap percolation on a graph is a simple activation process that starts with a given number of initially active nodes (called seeds) and evolves as follows. Every inactive node that has at least $r \geq 2$ active neighbors is activated, and remains so forever. The process stops when no more nodes can be activated. There are two main cases of interest: one in which the seeds are selected uniformly at random among the nodes, and one in which the seeds are arbitrarily chosen. In both cases, the main question concerns the final size of the set of active nodes. Bootstrap percolation was introduced in [16] on a Bethe lattice, and successively investigated on regular grids and trees [9, 10]. More recently, bootstrap percolation has been studied on random graphs and random trees [3, 5, 6, 7, 8, 11, 12, 14, 20, 24, 25, 35], motivated by the increasing interest in large-scale complex systems such as technological, biological and social networks. For example, in the case of social networks, bootstrap percolation may serve as a primitive model for the spread of ideas, rumors and trends among individuals. Indeed, in this context one can assume that a person will adopt an idea after receiving sufficient influence by friends who have already adopted it [27, 32, 36].

In more detail, bootstrap percolation has been studied on random regular graphs [11], on random graphs with given vertex degrees [5], on Galton–Watson random trees [12], on random geometric graphs [14], on Chung–Lu random graphs [6, 7, 20] (which notably permit considering the case of power-law node degree distribution), on small-world random graphs [25, 35] and on Barabasi–Albert random graphs [3]. Particularly relevant to our work is the paper by Janson et al. [24], where the authors have provided a detailed analysis of the bootstrap percolation process on the Erdős–Rényi random graph. We emphasize that in [24] the seeds are chosen uniformly at random among the nodes, however, as proved in [18], the critical number of seeds triggering percolation can be significantly reduced if the selection of seeds is optimized.

Over the years, several variants of the bootstrap percolation have been considered. In majority bootstrap percolation, a node becomes active if at least half of its neighbors are active. In jigsaw percolation, introduced in [15], there are two types of edges, one representing “social links" and one representing “compatibility of ideas". Two clusters of nodes merge together if there exists at least one edge of each type between them. Majority and jigsaw bootstrap percolation have been analyzed on the Erdős–Rényi random graph in [23] and [13], respectively.

Community structure is an important characteristic of many real-world graphs. This feature, however, is not present in any of the graphs on which bootstrap percolation (or its variants) have been studied so far. Informally, one says that a graph has a community structure if nodes are partitioned into clusters in such a way that many edges join nodes of the same cluster and comparatively fewer edges join nodes of different clusters [21]. Many methods have been proposed for community detection in real networks (see the review article [19]).

Through the development of the theoretical foundations of community detection, the so-called stochastic block model (SBM) has arisen naturally, and attracted considerable attention. The SBM is essentially the superposition of Erdős–Rényi graphs, and is perhaps the simplest interesting case of a random graph with community structure. In particular, detection of two symmetric communities has been studied in [28], while partial or exact recovery of the community membership has been investigated in [1, 2].

In this paper we study classical bootstrap percolation on the SBM with two (in general asymmetric) communities, assuming that seeds are selected uniformly at random within each community and allowing a different number of seeds for different communities. We prove the existence of a sharp phase transition for the number of eventually active nodes, identifying a sub-critical regime, in which the evolution of the bootstrap percolation process is very limited (in the sense that the final size of active nodes is of the same order as the number of seeds), and a super-critical regime, in which the activation process percolates almost completely (in the sense that the vast majority of nodes will be activated). Although our results generalize some of the main achievements in [24], we emphasize that our techniques significantly differ from those employed in [24]. In particular, we devise a suitable extension of the classical *binomial chain* construction originally proposed in [31] (and also used in [24]), adapting it to the SBM. Furthermore, as opposed to [24], where Doob’s martingale inequality is employed, we use deviation inequalities for the binomial distribution to prove that bootstrap percolation on the SBM concentrates around its average. Our approach provides exponential bounds on the related tail probabilities, which allow us to strengthen the convergence in probability for the final size of active nodes (as obtained in [24]) to the level of almost sure convergence.

To better understand the main difficulties in the analysis of the bootstrap percolation process on the SBM, we recall that in the classical binomial chain construction a (virtual) discrete time is introduced: at each time step a single active node is *explored* by revealing its neighbors. Nodes become active as soon as the number of their explored neighbors reaches the percolation threshold r. In the SBM the stochastic properties of the set of active nodes at time step t heavily depend on the number of nodes that have been explored in each community up to time t, and this makes the analysis of the bootstrap percolation process on the SBM significantly more complex. In particular, it requires the identification of an appropriate *strategy* to select the community in which a new node is explored at every time step.

Although considerably flexible and mathematically tractable, the SBM does not accurately describe most real-world networks. For instance, it does not allow for heterogeneity of nodes within communities. Different variants of the SBM have been proposed to better fit real network data, such as letting nodes follow a given degree sequence [17, 26] or considering overlapping communities and mixed membership models [4, 22]. We acknowledge that analyzing bootstrap percolation on the SBM is only a first step towards a better understanding of this process on more sophisticated community-based models.

The paper is organized as follows. In Section 2 we introduce the model and our assumptions on its parameters. The main results of the paper are stated in Section 3, together with some numerical illustrations. In Section 4 we provide an overview of our analysis, by first introducing the extension to the SBM of the classical binomial chain representation of the bootstrap percolation process, and then by giving a high-level description of our proofs. The detailed proofs are reported in Section 5. Lastly, in the Supplementary Material (SM), we report the proof of some ancillary results.

2. The stochastic block model

2.1. Model description

The SBM $G = G(n_1, n_2, p_1, p_2, q)$, with number of nodes $n = n_1 + n_2$ and parameters $p_1, p_2, q \in [0, 1)$, is a random graph formed by the union of two disjoint Erdős–Rényi random graphs $G_i = G(n_i, p_i)$, $i = 1, 2$, called hereafter communities, where edges joining nodes in different communities G_1 and G_2 are independently added with probability q. In the following we will refer to edges between nodes in the same community as “intra-community" edges and to edges joining nodes in different communities as “inter-community" edges.

Bootstrap percolation on the SBM is an activation process that obeys to the following rules:

- At the beginning, an arbitrary number a_i ($a_i \leq n_i$) of nodes, called seeds, are chosen uniformly at random among the nodes of G_i. Seeds are declared to be active, while nodes not belonging to the set of seeds are initially inactive.
- An inactive node becomes active as soon as at least $r \geq 2$ of its neighbors are active, and then remain active forever, so that the set of active nodes grows monotonically.
- The process stops when no more nodes can be activated.

The bootstrap percolation process naturally evolves through generations of nodes that are sequentially activated. The initial generation $\mathcal{G}_0$ is the set of seeds; the first generation $\mathcal{G}_1$ is composed by all those nodes that are neighbors of at least r seeds; the second generation $\mathcal{G}_2$ is composed by all the nodes that are neighbors of at least r nodes in $\mathcal{G}_0 \cup \mathcal{G}_1$, and so on. The bootstrap percolation process stops when either an empty generation is obtained or all the nodes are active. The final set of active nodes is clearly given by

$$\mathcal{G} \equiv \bigcup_{k \geq 0} \mathcal{G}_k.$$

We conclude this subsection introducing some notation and terminology. Given two functions f_1 and f_2 we write $f_1(m) \ll f_2(m)$ (or equivalently $f_1(m) = o(f_2(m))$), $f_1(m) \sim f_2(m)$, and $f_1(m) = O(f_2(m))$ if, as $m \to \infty$, $f_1(m)/f_2(m) \to 0$, $f_1(m)/f_2(m) \to 1$ and $\limsup_{m\to\infty} |f_1(m)/f_2(m)| < \infty$. Letting $|\mathcal{X}|$ denote the cardinality of a set $\mathcal{X}$, we say that the bootstrap percolation process *percolates* whenever $|\mathcal{G}| = n - o(n)$, that is, whenever almost all the nodes are activated.

2.2. Model assumptions

In the following we consider a sequence of SBMs with a growing number of nodes n. We warn the reader that, unless explicitly written, all the limits in this paper are taken as $n \to \infty$.

We assume that the communities G_1 and G_2 have sizes that are asymptotically of the same order, i.e.,

$$n_1 \sim \nu n_2, \quad \text{for some } \nu \in \mathbb{R}_+ := (0, \infty), \tag{1}$$

and that the inter-community and the intra-community edge probabilities are asymptotically of the same order too, i.e.,

$$q \sim \gamma p_1, \quad \text{for some } \gamma \in \mathbb{R}_+, \quad p_1 \sim \mu p_2, \quad \text{for some } \mu \in \mathbb{R}_+. \tag{2}$$

Note that since $\gamma > 0$ the communities are never isolated. Similarly to [24], we assume

$$1/n_i \ll p_i \ll 1/(n_i^{1/r}), \quad i = 1, 2 \tag{3}$$

and we define the critical number of seeds, in correspondence of which the bootstrap percolation process exhibits a phase transition in the Erdős–Rényi random graph $G(n_i, p_i)$, by

$$g_i := \left(1 - \frac{1}{r}\right) \left(\frac{(r-1)!}{n_i p_i^r}\right)^{\frac{1}{r-1}}, \quad i = 1, 2.$$

As proved in [24], under (3), we have

$$g_i \to \infty, \quad g_i/n_i \to 0, \quad p_i g_i \to 0, \quad i = 1, 2. \tag{4}$$

Note that by (1) and (2) it follows that g_1 and g_2 are asymptotically comparable. Furthermore, similarly to [24], we assume

$$a_i/g_i \to \alpha_i \geq 0, \quad i = 1, 2, \quad \text{with } \max\{\alpha_1, \alpha_2\} > 0. \tag{5}$$

Without loss of generality, we suppose

$$\alpha_1 \geq \alpha_2 \quad \text{with } \alpha_1 > 0. \tag{6}$$

Inspired by some literature on the subject (see e.g. [29]) we say that the SBM is *assortative* if the intra-community edge probabilities exceed the inter-community edge probability. Specifically, a SBM is said assortative if $q^2 < p_1 p_2$. Since $q^2/(p_1 p_2) \to \gamma^2 \mu$ (see (2)), by setting

$$\chi_{ii} = 1, \quad i = 1, 2, \quad \chi_{12} := \gamma(\nu\mu^r)^{\frac{1}{r-1}}, \quad \chi_{21} := \gamma(\nu\mu)^{-1/(r-1)} \quad \text{and} \quad \boldsymbol{\chi} = (\chi_{ij})_{i,j=1,2},$$

the *assortative* condition can be reformulated as $\det\boldsymbol{\chi} > 0$. Therefore, in the following we will refer to *assortative* SBM when $\det\boldsymbol{\chi} > 0$, *dis-assortative* SBM when $\det\boldsymbol{\chi} < 0$ and *neutral* SBM when $\det\boldsymbol{\chi} = 0$. Although these notions do not play any role in our main results (i.e., Theorems 3.2 and 3.3), they do have an impact on the definition of the *critical curve* for the system (see Proposition 3.5).

Finally, we remark once again that, within each community, the seeds must be selected uniformly at random and in such a way that the number of seeds satisfies the constraint (5).

2.3. Bootstrap percolation on the Erdős–Rényi random graph: a quick review

To better position our results with respect to the existing literature, we briefly recall the main achievements in [24]. Note that the Erdős–Rényi random graph corresponds to a SBM with a single community, (i.e., $i = 1$). It has been proved in [24] (see Theorem 3.1(ii)). that:
(i) If (3) and (5) hold (with $i = 1$) and $\alpha_1 < 1$, then

$$|\mathcal{G}|/g_1 \to \frac{r\varphi(\alpha_1)}{(r-1)\alpha_1}, \quad \text{in probability}$$

where $\varphi(\alpha_1)$ is the unique solution in $[0,1]$ of equation $rx - x^r = (r-1)\alpha_1$ with unknown x (see Theorem 3.1(i) in [24]).
(ii) If (3) and (5) hold (with $i = 1$) and $\alpha_1 > 1$, then

$$|\mathcal{G}|/n \to 1, \quad \text{in probability}.$$

3. Main results

The bootstrap percolation process on the Erdős–Rényi random graph exhibits a sharp phase transition, see [24]. The reader may be wondering whether more complex phenomena, such as selective percolation of communities, can be observed on the SBM. We will show that this is not the case. Indeed, under the assumptions described in Subsection 2.2, the bootstrap percolation process either stops with high probability when $O(g_1)$ vertices have been activated (sub-critical case) or percolates (super-critical case). A selective percolation of the communities may be instead observed when $\gamma = 0$ (i.e., $q = o(p_1)$), where the bootstrap percolation process may behave in each community as if they were isolated.

To state our main results we need some additional notation. For $\mathbf{x} = (x_1, x_2) \in [0, \infty)^2$, we define the following functions:

$$\rho_i(\mathbf{x}) := \alpha_i - x_i + r^{-1}(1 - r^{-1})^{r-1}(x_i + \chi_{ij}x_j)^r, \quad i \neq j \in \{1, 2\}$$

and the following sets:

$$\mathcal{D} := \left\{ \mathbf{x} \in [0, r/(r-1)]^2 : \ x_1 + \chi_{12}x_2 \leq \frac{r}{r-1}, \ x_2 + \chi_{21}x_1 \leq \frac{r}{r-1} \right\},$$

$$\mathcal{E}_1 := \{\mathbf{x} \in \mathcal{D} : \ \rho_1(\mathbf{x}) \leq 0\}, \quad \mathcal{E}_2 := \{\mathbf{x} \in \mathcal{D} : \ \rho_2(\mathbf{x}) \leq 0\},$$

$$\widetilde{\mathcal{E}}_1 := \{\mathbf{x} \in \mathcal{D} : \ \rho_1(\mathbf{x}) = 0\}, \quad \widetilde{\mathcal{E}}_2 := \{\mathbf{x} \in \mathcal{D} : \ \rho_2(\mathbf{x}) = 0\}.$$

For a set $\mathcal{H} \subset \mathcal{D}$, we denote by $\overset{\circ}{\mathcal{H}}$ its interior (with respect to the Euclidean topology on $\mathbb{R}^2$ restricted to $\mathcal{D}$). Throughout this paper, we consider the following three disjoint and exhaustive conditions:
$(\mathbf{Sub})$: $\overset{\circ}{\mathcal{E}}_1 \cap \overset{\circ}{\mathcal{E}}_2 \neq \emptyset$, $\qquad (\mathbf{Crit})$: $\overset{\circ}{\mathcal{E}}_1 \cap \overset{\circ}{\mathcal{E}}_2 = \emptyset$, $\widetilde{\mathcal{E}}_1 \cap \widetilde{\mathcal{E}}_2 \neq \emptyset$, $\qquad (\mathbf{Sup})$: $\mathcal{E}_1 \cap \mathcal{E}_2 = \emptyset$.
Hereafter, we refer to such conditions as sub-critical, critical and super-critical regimes, respectively. A graphical representation of these regimes is given in Figures 1, 2 and 3, where the blue curves represent $\widetilde{\mathcal{E}}_1$ and $\widetilde{\mathcal{E}}_2$ (additional notation appearing on the plots will be introduced later on).

***Remark* 3.1.** Let $i \in \{1, 2\}$ be fixed. A straightforward computation shows that if $\alpha_i > 1$, then $\min_{\mathbf{x} \in \mathcal{D}} \rho_i(\mathbf{x}) > 0$, therefore $\mathcal{E}_i = \emptyset$ and so $(\mathbf{Sup})$ holds. Consequently, conditions $(\mathbf{Sub})$ and $(\mathbf{Crit})$ imply $\alpha_i \leq 1$ for any $i = 1, 2$.

3.1. Phase transition on the SBM model

Next theorems provide the main results of the paper. Hereon, for ease of notation, we denote by (C) the set of conditions: (1), (2), (3), (5) and (6).

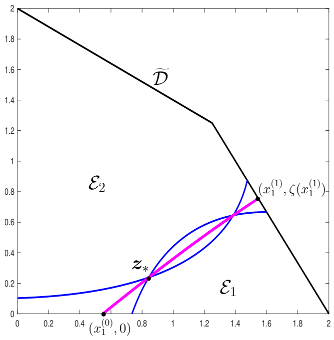

Figure 1. Sub-critical regime, $r = 2$, $\chi_{12} = \chi_{21} = 0.6$, $\alpha_1 = 0.56$, $\alpha_2 = 0.1$.

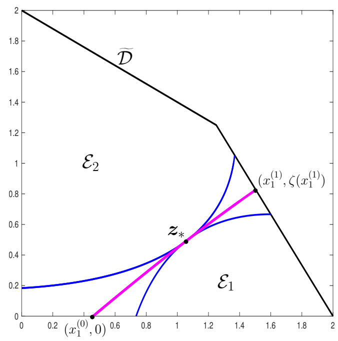

Figure 2. Critical regime, $r = 2$, $\chi_{12} = \chi_{21} = 0.6$, $\alpha_1 = 0.6$, $\alpha_2 = 0.175$.

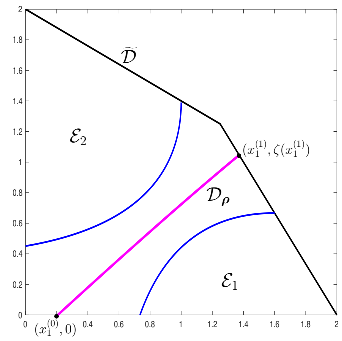

Figure 3. Super-critical regime, $r = 2$, $\chi_{12} = \chi_{21} = 0.6$, $\alpha_1 = 0.6$, $\alpha_2 = 0.4$.

Theorem 3.2. *Assume* (C) *and* $(\mathbf{Sub})$. *Then, for any* $\varepsilon > 0$ *there exists* $c(\varepsilon) \in \mathbb{R}_+$ *such that*

$$P\left(\left|\frac{|\mathcal{G}|}{g_1} - x_*\right| > \varepsilon\right) = O(\mathrm{e}^{-c(\varepsilon)g_1}), \tag{7}$$

where the explicit expression of the positive constant $x_* > 0$ *is given in* (25).

Theorem 3.3. *Assume* (C) *and* $(\mathbf{Sup})$. *Then, for any* $\varepsilon > 0$ *there exists* $c(\varepsilon) \in \mathbb{R}_+$ *such that*

$$P\left(\left|\frac{|\mathcal{G}|}{n} - 1\right| > \varepsilon\right) = O(\mathrm{e}^{-c(\varepsilon)g_1}). \tag{8}$$

Roughly speaking, the above results can be rephrased as follows:
(i) under (C) and $(\mathbf{Sub})$, the bootstrap percolation process on the SBM reaches, as $n \to \infty$, a final size of active nodes which is of the same order as $a_1 + a_2$ (indeed, by (2), the definition of g_i, (5) and (6), it easily follows that $a_1 + a_2 \sim (\alpha_1 + \alpha_2(\nu\mu^r)^{1/(r-1)})g_1)$,
(ii) under (C) and $(\mathbf{Sup})$, the bootstrap percolation process on the SBM percolates, as $n \to \infty$.

***Remark* 3.4.** Replacing the assumption (3) with the (slightly) stronger condition:

For any $i = 1, 2$, $1/n_i \ll p_i$ and either $p_i \ll 1/(n_i)^{\frac{1}{r'}}$ or $p_i \sim c/(n_i)^{\frac{1}{r'}}$, for some $c > 0$ and $r' \in (r, \infty)$,

by a standard application of the Borel-Cantelli lemma, the claims (7) and (8) can be strengthened, respectively, as $\frac{|\mathcal{G}|}{g_1} \to x_*$ and $\frac{|\mathcal{G}|}{n} \to 1$ almost surely.

3.2. The critical curve and the sub-critical and super-critical regions

To complement the results of Theorems 3.2 and 3.3, in this subsection we determine the sub-critical and the super-critical regions of the system, i.e., the set of $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)$ for which either the sub-critical or the super-critical behavior is observed. We restrict our investigation to $\boldsymbol{\alpha} \in [0,1]^2$ since, as already noticed in Remark 3.1, necessarily $(\mathbf{Sup})$ holds whenever $\alpha_1 > 1$ and/or $\alpha_2 > 1$. We write $\overset{\circ}{\mathcal{E}}_i(\alpha_i)$,

$\widetilde{\mathcal{E}}_i(\alpha_i)$ and $\mathcal{E}_i(\alpha_i)$ in place of $\mathring{\mathcal{E}}_i$, $\widetilde{\mathcal{E}}_i$ and $\mathcal{E}_i$, respectively, to make the dependence on α_i explicit. We define the regions

$$\mathcal{R}_{\mathbf{Sub}} := \left\{ \boldsymbol{\alpha} \in [0,1]^2 : \; \mathring{\mathcal{E}}_1(\alpha_1) \cap \mathring{\mathcal{E}}_2(\alpha_2) \neq \emptyset \right\}, \quad \mathcal{R}_{\mathbf{Sup}} := \left\{ \boldsymbol{\alpha} \in [0,1]^2 : \; \mathcal{E}_1(\alpha_1) \cap \mathcal{E}_2(\alpha_2) = \emptyset \right\}, \tag{9}$$

and the curve

$$\mathcal{R}_{\mathbf{Crit}} := \left\{ \boldsymbol{\alpha} \in [0,1]^2 : \; \mathring{\mathcal{E}}_1(\alpha_1) \cap \mathring{\mathcal{E}}_2(\alpha_2) = \emptyset, \widetilde{\mathcal{E}}_1(\alpha_1) \cap \widetilde{\mathcal{E}}_2(\alpha_2) \neq \emptyset \right\}, \tag{10}$$

to which we refer as the sub-critical and the super-critical regions, and the critical curve, respectively.

By exploiting the convexity of the functions $\rho_i(\cdot)$ and by imposing the tangency condition between the curves $\widetilde{\mathcal{E}}_1(\alpha_1)$ and $\widetilde{\mathcal{E}}_2(\alpha_2)$, one can show the following Proposition 3.5, whose proof is elementary, and therefore omitted. From here on, we denote by $\mathcal{M}^t$ the transpose of the matrix $\mathcal{M}$.

Proposition 3.5. *The following claims hold:*
(i) *Under* (C) *and* $\det\boldsymbol{\chi} \neq 0$*, we have*

$$\mathcal{R}_{\mathbf{Crit}} = \left\{ (y_1, y_2)(\boldsymbol{\chi}^{-1})^t - r^{-1}(1-r^{-1})^{r-1}(y_1^r, y_2^r) \in [0,1]^2 : \; 0 \leq y_1 \leq r/(r-1), \right. \\ \left. y_2 = (1-r^{-1}) \left(\frac{1-(1-r^{-1})^{r-1} y_1^{r-1}}{1-(1-r^{-1})^{r-1} y_1^{r-1} \det\boldsymbol{\chi}} \right)^{1/(r-1)} \right\}.$$

(ii) *Under* (C) *and* $\det\boldsymbol{\chi} = 0$*, we have*

$$\mathcal{R}_{\mathbf{Crit}} = \left\{ (y_1, y_2) - r^{-1}(1-r^{-1})^{r-1}((y_1 + y_2\chi_{12})^r, (y_1\chi_{21} + y_2)^r) \in [0,1]^2 : \; 0 \leq y_1 \leq r/(r-1), \right. \\ \left. y_2 = \chi_{21} \left[\frac{r}{r-1} \left(\frac{1}{1+\chi_{21}^{r-1}} \right)^{1/(r-1)} - y_1 \right] \right\}.$$

(iii) *Under* (C)*,* $\mathcal{R}_{\mathbf{Sub}}$ *is the convex set delimited by the curve* $\mathcal{R}_{\mathbf{Crit}}$ *and the coordinate axes.*

Note that $\mathcal{R}_{\mathbf{Crit}}$ depends only on the asymptotic properties of the SBM, which are expressed in terms of the parameters r, γ, μ and ν. In other words, two (sequences of) SBMs with the same parameters r, γ, μ and ν lead to the same critical curve $\mathcal{R}_{\mathbf{Crit}}$, and therefore to the same sub-critical and super-critical regions. Hereafter, we illustrate numerically Proposition 3.5, taking (sequences of) SBMs with parameters $r = 2$, $\gamma = 0.25$, and $\nu = \mu = 1$ as baseline case.

We start by investigating the impact of the various parameters on the sub-critical and the super-critical regions. To this aim, we vary a parameter at a time, keeping fixed all the others, and determine the critical curve.

In Figure 4 we vary the parameter γ, which characterizes the strength of the inter-community connectivity with respect to the intra-community connectivity. When $\gamma < 1$ ($r = 2$, $\mu = \nu = 1$) SBMs are assortative, whereas when $\gamma > 1$ SBMs are disassortative. Finally, in the special case when $\gamma = 1$ the SBMs are neutral (i.e., $\det\boldsymbol{\chi} = 0$) and exhibit the same $\mathcal{R}_{\mathbf{Crit}}$ of Erdős–Rényi random graphs. In this

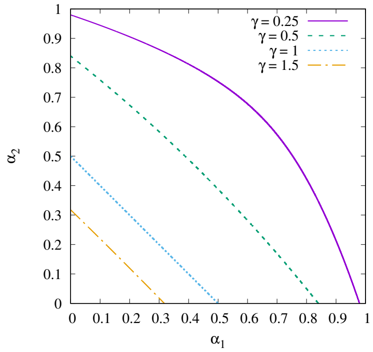

Figure 4. Critical curves $\mathcal{R}_{\mathbf{Crit}}$ for different values of γ.

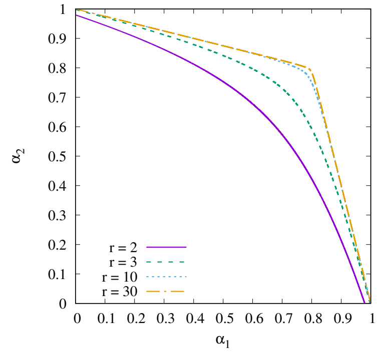

Figure 5. Critical curves $\mathcal{R}_{\mathbf{Crit}}$ for different values of r.

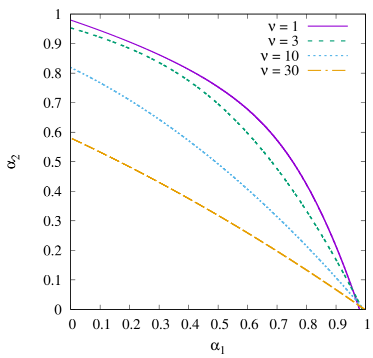

Figure 6. Critical curves $\mathcal{R}_{\mathbf{Crit}}$ for different values of ν.

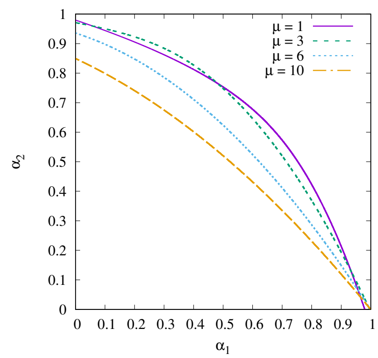

Figure 7. Critical curves $\mathcal{R}_{\mathbf{Crit}}$ for different values of μ.

special case the critical curve corresponds to the segment where $\alpha_1 + \alpha_2 = 0.5$, indeed a straightforward computation gives $g_1 = \frac{1}{2n_1 p_1^2}$. We further note that, as $\gamma \downarrow 0$, the sub-critical region approaches the whole square (because, as $\gamma \to 0$, the fraction of edges connecting the two communities tends to vanish, and therefore the activation process spreads in the two communities as if they were isolated). Finally, since the sub-critical region is convex for any γ, in a SBM with $\mu = \nu = 1$ (i.e., symmetric), we have that the critical number of seeds is minimized when all the seeds are placed in the same community (i.e., either $\alpha_1 = 0$ or $\alpha_2 = 0$). Instead, the critical number of seeds is maximized when the seeds are equally partitioned between the communities (which approximately occurs, notably, when the seeds are chosen uniformly at random among the nodes). Interestingly, in the latter case (i.e., when the seeds are equally partitioned between the communities), a simple computation shows that the critical threshold in a SBM with $\mu = \nu = 1$ is asymptotically equal to the critical threshold in an Erdős–Rényi random graph having the same average degree.

In Figure 5 we vary the threshold parameter r. Note that, as $r \uparrow \infty$, the sub-critical region approaches the domain

$$\{(\alpha_1, \alpha_2) \in [0,1]^2 : \ \alpha_1 + \gamma\,\alpha_2 < 1, \ \alpha_2 + \gamma\,\alpha_1 < 1\}$$

(this property holds for any γ in the symmetric SBM with $\mu = \nu = 1$).

Next, we explore what happens in SBMs with $\mu \neq \nu$ (i.e., asymmetric) by changing either ν or μ. In Figure 6 we fix $r = 2$, $\gamma = 0.25$, $\mu = 1$ and increase the parameter ν, making the first community increasingly larger than the second community. Interestingly, we observe a significant reduction in the (normalized) critical value of α_2 for increasing values of ν when we put all the seeds in the community G_2 (i.e., $\alpha_1 = 0$): this means that fewer and fewer seeds are needed in community G_2 to trigger percolation, as the community G_1 becomes larger and larger. This because the epidemic transfers into the community G_1, where it propagates more easily thanks to the larger number of available nodes. However, note that, to minimize the (un-normalized) critical number of seeds, all the seeds must be placed in the larger community G_1, as a consequence of the fact that g_i, $i = 1, 2$, are different.

Finally, in Figure 7 we fix $r = 2$, $\gamma = 0.25$, $\nu = 1$ and increase the parameter μ, thus increasing the intra-community probability in G_1. For large values of μ, considerations similar to Figure 6 apply.

4. Overview of the methodology

4.1. The extension of the binomial chain construction

We introduce a discrete time $t \geq 0$ and we assign a marks counter $M_v(t)$, $M_v(0) := 0$, to every node v which is not a seed. Seeds are activated at time $t = 0$. We set $\mathcal{U}_i(0) := \emptyset$ and denote by $\mathcal{A}_i(0)$ the set of seeds in the community G_i. The process, then, evolves according to the following recursive procedure. At time $t \in \mathbb{N} := \{1, 2, \ldots\}$:

- We arbitrarily select a community G_j provided that $\mathcal{A}_j(t-1) \setminus \mathcal{U}_j(t-1) \neq \emptyset$.
- From the selected community G_j, we choose, uniformly at random, a node $v \in \mathcal{A}_j(t-1) \setminus \mathcal{U}_j(t-1)$.
- We use the chosen node v, i.e., we explore the node v by revealing its neighbors and by adding a mark to each of them.
- We set $\mathcal{U}_j(t) := \mathcal{U}_j(t-1) \cup \{v\}$ and $\mathcal{U}_i(t) := \mathcal{U}_i(t-1)$, for $i \neq j$. We also set $\mathcal{A}_i(t) := \mathcal{A}_i(t-1) \cup \Delta\mathcal{A}_i(t)$, where $\Delta\mathcal{A}_i(t)$ is the set of nodes in the community G_i that become active exactly at time t, i.e., the set of nodes in G_i that have received the r-th mark exactly at time t. Note that $\Delta\mathcal{A}_i(t) = \emptyset$ for $t < r$, since no other nodes are activated until at least r seeds are used.
- The process terminates as soon as there are no active and still unused nodes, i.e., at time step:

$$T := \min\{t \in \mathbb{N} :\ \mathcal{A}_i(t) = \mathcal{U}_i(t),\ \forall i = 1, 2\}. \tag{11}$$

Note that, since only one node is used at each time step, for any $t \leq T$, $|\mathcal{U}(t)| = t$, where $\mathcal{U}(t) := \mathcal{U}_1(t) \cup \mathcal{U}_2(t)$. Let $\mathcal{A}(t) := \mathcal{A}_1(t) \cup \mathcal{A}_2(t)$ denote the set of active nodes at time $t \leq T$. We clearly have

$$v \in \mathcal{A}(t) \setminus \mathcal{A}(t-1) \quad \text{if and only if} \quad M_v(t) = r, \quad 1 \leq t \leq T \tag{12}$$

where

$$M_v(t) = \sum_{i=1}^{2} \sum_{s=1}^{U_i(t)} I_v^{(i)}(s), \quad \forall\, v \notin \mathcal{A}(t-1) \tag{13}$$

$U_i(t) := |\mathcal{U}_i(t)|$ and the random variables $\{I_v^{(i)}(s)\}_{v \notin \mathcal{A}(t), 1 \leq i \leq 2, 1 \leq s \leq T}$ are independent, with $I_v^{(i)}(s)$ distributed as $\mathrm{Be}(p_i)$ [1] if v is a node of the community G_i, and distributed as $\mathrm{Be}(q)$ if v is a node of the community G_j, $j \neq i$.

[1] Here $\mathrm{Be}(p)$ denotes a Bernoulli distributed random variable with mean $p \in [0, 1]$.

The next proposition guarantees that the order in which active nodes are used does not have any impact on the final set of active vertices $\mathcal{G}$.

Proposition 4.1. *We have* $\mathcal{G} \equiv \mathcal{A}(T)$.

Although Proposition 4.1 may appear rather obvious, it plays a crucial role in our proofs. Therefore, for completeness, we report its proof in the SM.

***Remark* 4.2.** In the description of the binomial chain representation of the bootstrap percolation process, we did not fully specify the rule according to which a community is selected at every time step $t \in \mathbb{N}$. Indeed, we limited ourselves just to mention a general guideline for the selection of the community: at time $t \in \mathbb{N}$, we choose a community G_i which has active and unused nodes. Clearly, this choice can be made in many different ways. Throughout this paper, we refer to such different ways as “strategies". Remarkably, Proposition 4.1 applies to any strategy. It will become clear later on that the opportunity to “arbitrarily" define the strategy for the community selection, provides a fundamental degree of flexibility that comes in handy when we analyze the bootstrap percolation process on the SBM (see the proofs of Theorems 3.2 and 3.3).

Hereon, we put $[n] := \{1, 2, \cdots, n\}$ and let $i \in \{1, 2\}$ be fixed. We have defined the random marks $I_v^{(i)}(s)$ for $v \notin \mathcal{A}(t)$ and $1 \le s \le T$, but, similarly to [24], see Section 2 therein, it is possible to introduce additional, redundant random marks, which are independent and Bernoulli distributed with mean p_i if v is a node of the community G_i and with mean q if v is a node of the community G_j, $j \neq i$, in such a way that $I_v^{(i)}(s)$ is defined for all $v \in G$ and $s \in [n]$. Such additional random marks are added, for any $1 \le s \le T$, to already active nodes and so they have no effect on the underlying bootstrap percolation process. This corresponds to artificially extending the chain construction beyond T, by selecting and exploring at every time $T \le t \le n$ a potentially non-active node. Clearly such extension has no effect on the dynamics of the bootstrap percolation process up to time T, and it is just instrumental. Throughout this paper, we denote by $\mathrm{Bin}(u,p)$, $u \in \mathbb{N}$, $p \in [0,1]$, a random variable following the binomial distribution with parameters (u,p).

Note that, conditionally on $U_1(t)$ and $U_2(t)$, the random variable $M_v(t)$ is the sum of independent random variables with the binomial distribution, i.e., for fixed $i \in \{1, 2\}$ and $t \in [n] \cup \{0\}$ we have

$$M_v(t) \,|\, \{\mathbf{U}(t) = \mathbf{u}(t)\} \stackrel{\mathcal{L}}{=} \mathrm{Bin}(u_i(t), p_i) + \mathrm{Bin}(u_j(t), q), \quad v \in G_i,\, j \neq i \tag{14}$$

where $\mathbf{U}(t) := (U_1(t), U_2(t))$, $\mathbf{u}(t) := (u_1(t), u_2(t))$, the symbol $\stackrel{\mathcal{L}}{=}$ denotes the equality in law and the random variables $\mathrm{Bin}(u_i(t), p_i)$ and $\mathrm{Bin}(u_j(t), q)$ are independent. The number of active nodes in the community G_i at time $t \in [n] \cup \{0\}$ is given by

$$A_i(t) := |\mathcal{A}_i(t)| = a_i + S_i(t), \tag{15}$$

where

$$S_i(t) := \sum_{v \in G_i \setminus \mathcal{A}_i(0)} \mathbf{1}\{Y_v \le t\}, \qquad Y_v := \min\{s \in \mathbb{N} :\ M_v(s) \ge r\}. \tag{16}$$

Since the random variables $\{M_v(t) \,|\, \{\mathbf{U}(t) = \mathbf{u}(t)\}\}_{v \in G_i}$ are independent and identically distributed with law specified by (14), we have

$$S_i(t) \,|\, \{\mathbf{U}(t) = \mathbf{u}(t)\} \stackrel{\mathcal{L}}{=} \mathrm{Bin}(n_i - a_i, b_i(\mathbf{u}(t))), \tag{17}$$

where

$$b_i(\mathbf{u}(t)) := P\left(\text{Bin}(u_i(t), p_i) + \text{Bin}(u_j(t), q) \geq r\right), \quad i \in \{1,2\}, j \neq i. \tag{18}$$

Hereafter, we denote by $A(t) := |\mathcal{A}(t)| = \sum_{i=1}^{2} A_i(t)$, the number of active nodes in the SBM at time t. Note that $|\mathcal{G}| = A(T) = T - 1$.

***Remark* 4.3.** The analysis of the bootstrap percolation process is significantly more complex on the SBM than on the Erdős–Rényi random graph. Indeed, on the SBM, for any $t < T$, the random variables $\{A_i(t)\}_{1 \leq i \leq 2}$ depend on the quantities $\{U_i(t)\}_{1 \leq i \leq 2}$, and so on the chosen strategy. In turn, the choice of a strategy is constrained by the availability of active and unused nodes in the different communities. As a result, $S_i(t)$ is binomial only given the event $\{\mathbf{U}(t) = \mathbf{u}(t)\}$. In contrast, on the Erdős–Rényi random graph the number of used nodes at time t is equal to t, and therefore the law of the number of active and unused nodes at time t is (unconditionally) binomial.

4.2. High level description of the proofs

In broad terms, the proofs of Theorems 3.2 and 3.3 adopt the following approach. First, note that since $T - 1 = |\mathcal{G}|$, we can reduce the computation of the tail probabilities of $|\mathcal{G}|$ to the computation of the tail probabilities of T. Then, exploiting the definition of T given in (11), we aim to upper-bound the tail probabilities of T with a combination of probabilities associated to the events $\{A_i(t) - U_i(t) < 0\}$, $i \in \{1,2\}$, for different time instant t. However, in doing so, the following difficulty arises. $\mathbf{A}(t)$ depends on $\mathbf{U}(t)$, which itself depends on the selected strategy and on the past trajectory $\mathbf{A}(\tau) - \mathbf{U}(\tau)$ for $\tau < t$. This because, as already mentioned in Remark 4.3, whatever strategy is considered, we can choose a node in the community G_i at time τ only if $A_i(\tau - 1) - U_i(\tau - 1) > 0$. We refer to this constraint as *feasibility* constraint.

By (4.1), we can choose whatever strategy is convenient (among those that are feasible, i.e., satisfy the feasibility constraint), indeed the choice of a strategy has no impact on the final number of active nodes. A first crucial step in our proofs consists in identifying such a strategy. In the attempt to balance the number of active and unused nodes in the two communities, a possible candidate is the *max-strategy*, according to which, at time step $1 \leq t \leq T$, one chooses the community with the maximum number of active and unused nodes $A_i(t-1) - U_i(t-1)$. The main drawback of this strategy is that the analysis of the corresponding processes $\mathbf{A}(t)$, $\mathbf{U}(t)$, $t \leq T$, appears prohibitive due to its complex correlation structure. To circumvent this difficulty, we introduce a *hybrid* variant of the max-strategy defined above, according to which, at time t, the community G_1 is selected if and only if

$$\lim_{n \to \infty} \frac{E[A_1(t) - U_1(t) \mid \mathbf{U}(t) = \mathbf{u}(t)]}{g_1} \geq \lim_{n \to \infty} \frac{E[A_2(t) - U_2(t) \mid \mathbf{U}(t) = \mathbf{u}(t)]}{g_2},$$

i.e., at time t we select the community with the largest asymptotic normalized expected number of active and unused nodes.

We go on selecting communities according to this rule up to a random time T', $T' \leq T$, defined as the first time at which the feasibility constraint prevents us from further using our deterministic policy. For every time $t \in (T', T]$, instead, we select communities according to an arbitrary *feasible* strategy, such as the max-strategy. The reason why the *hybrid max-strategy* simplifies the analysis of the bootstrap percolation process is that up to time $t \leq T'$, the process $\mathbf{U}(t)$ is deterministic, with the mapping $t \mapsto (U_1(t)/g_1, U_2(t)/g_2)$ describing a particular well determined curve in $\mathcal{D}$. As a result, the characterization of $P(A_i(t) - U_i(t) < 0)$ becomes extremely simple, since it can be reduced to the tail probability of binomial random variables. Then we can easily bound from above the probability

$P(A_i(t) - U_i(t) < 0)$ by using the concentration inequalities reported in the SM, provided that we are able to characterize the average asymptotic dynamics of $E[A_i(t) - U_i(t)]$. We emphasize that, by so doing, we obtain exponential bounds. Moreover we wish to point out that the asymptotic analysis of the average dynamics of the hybrid max-strategy permits us to identify three regimes, which are shown to be equivalent to $(\mathbf{Sub})$, $(\mathbf{Sup})$ and $(\mathbf{Crit})$.

At last we recall that the interested reader can find the extension to the case of SBMs with $k > 2$ communities in [34]. While the stochastic analysis can be carried out following the same lines as for the case $k = 2$, the identification of a suitable deterministic strategy is not straightforward. We report in the SM a brief discussion of the main issues arising in the case of $k > 2$ communities.

5. Proofs

5.1. Preliminaries

We start by introducing the asymptotic normalized mean number of active and unused nodes. For $t \in [n] \cup \{0\}$ and $i \in \{1, 2\}$, we set

$$R_i(\mathbf{u}(t)) := E[A_i(t) - U_i(t) \,|\, \mathbf{U}(t) = \mathbf{u}(t)] = a_i + (n_i - a_i) b_i(\mathbf{u}(t)) - u_i(t). \tag{19}$$

Hereon, for $\mathbf{x} := (x_1, x_2) \in [0, \infty)^2$, we set

$$\lfloor \mathbf{x} g \rfloor := (\lfloor x_1 g_1 \rfloor, \lfloor x_2 g_2 \rfloor),$$

where $\lfloor x \rfloor$ denotes the greatest integer less than or equal to $x \in \mathbb{R}$. The following lemmas hold.

Lemma 5.1. *Assume (1), (2), (3), (5) and let $i \in \{1, 2\}$ be fixed. Then*

$$\lim_{n \to \infty} \frac{R_i(\lfloor \mathbf{x} g \rfloor)}{g_i} = \rho_i(\mathbf{x}), \quad \forall\, \mathbf{x} \in [0, \infty)^2 \tag{20}$$

Lemma 5.2. *Assume (1), (2), (3), (5), and let $\mathcal{W}$ be a compact subset of $(0, \infty)^2$. Then*

$$\sup_{\mathbf{x} \in \mathcal{W}} \left| \frac{R_i(\lfloor \mathbf{x} g \rfloor)}{g_i} - \rho_i(\mathbf{x}) \right| \to 0, \quad \forall\, i = 1, 2.$$

Lemma 5.3. *Assume (1), (2) and (3). Then, for any $\mathbf{x} \in [0, \infty)^2$, $i \in \{1, 2\}$ and $j \in \{1, 2\} \setminus \{i\}$,*

$$b_i(\lfloor \mathbf{x} g \rfloor) = \Big(1 + O\Big(\mathbf{1}\{x_i > 0\}(\lfloor x_i g_i \rfloor p_i + (\lfloor x_i g_i \rfloor)^{-1}) + \mathbf{1}\{x_j > 0\}(\lfloor x_j g_j \rfloor q + (\lfloor x_j g_j \rfloor)^{-1})\Big)\Big) \times \left(\lfloor x_i g_i \rfloor p_i + \lfloor x_j g_j \rfloor q\right)^r / r!.$$

We postpone the proofs of Lemmas 5.1, 5.2 and 5.3, which are technical, but conceptually rather straightforward, to the SM.

5.2. Equivalent formulations of $(\mathbf{Sub})$, $(\mathbf{Crit})$ and $(\mathbf{Sup})$

Throughout this subsection we assume (C) and (6) with $\alpha_1 \le 1$. We consider the curve

$$\mathcal{D}_{\boldsymbol{\rho}} := \{\mathbf{x} \in \mathcal{D} : \ \rho_1(\mathbf{x}) = \rho_2(\mathbf{x})\} \tag{21}$$

and the conditions:
$(\mathcal{S}ub)$: $\min_{\mathbf{x}\in\mathcal{D}_{\boldsymbol{\rho}}} \rho_1(\mathbf{x}) < 0$, $(\mathcal{C}rit)$: $\min_{\mathbf{x}\in\mathcal{D}_{\boldsymbol{\rho}}} \rho_1(\mathbf{x}) = 0$, $(\mathcal{S}up)$: $\min_{\mathbf{x}\in\mathcal{D}_{\boldsymbol{\rho}}} \rho_1(\mathbf{x}) > 0$.

Note that $\mathcal{D}_{\boldsymbol{\rho}}$ is graphically represented by the purple curve in Figures 1, 2 and 3. The following proposition holds.

Proposition 5.4. *Under the assumption (C) with $\alpha_1 \le 1$, we have that the conditions $(\mathbf{Sub})$, $(\mathbf{Crit})$ and $(\mathbf{Sup})$ are equivalent to $(\mathcal{S}ub)$, $(\mathcal{C}rit)$ and $(\mathcal{S}up)$, respectively.*

The proof of this proposition exploits the following lemma.

Lemma 5.5. *Assume (C) with $\alpha_1 \le 1$. Then:*
(i) $\mathcal{D}_{\boldsymbol{\rho}}$ is the graph of a strictly increasing function of class C^1, say $\zeta(\cdot)$, with domain $[x_1^{(0)}, x_1^{(1)}]$, where $x_1^{(0)}$ is the unique solution of the equation

$$\rho_1(x_1, 0) - \rho_2(x_1, 0) = 0, \quad x_1 \in (0, r/(r-1))$$

and $x_1^{(1)}$ is the unique point in $(0, r/(r-1))$ such that

$$\{(x_1^{(1)}, \zeta(x_1^{(1)}))\} = \widetilde{\mathcal{D}} \cap \mathcal{D}_{\boldsymbol{\rho}},$$

where

$$\widetilde{\mathcal{D}} := \left\{ \mathbf{x} \in \mathcal{D} : \ \max\{x_1 + \chi_{12} x_2, x_2 + \chi_{21} x_1\} = \frac{r}{r-1} \right\}.$$

(ii) $\rho_1(\mathbf{x}^{(0)}) = \rho_2(\mathbf{x}^{(0)}) > 0$, where $\mathbf{x}^{(0)} := (x_1^{(0)}, 0)$.
(iii) $\widetilde{\mathcal{E}}_1$ is the graph of a strictly increasing and strictly concave function of class C^2, say $\zeta_1(\cdot)$, with domain $[y_1^{(0)}, y_1^{(1)}]$ and $\zeta_1(y_1^{(0)}) = 0$. Here $y_1^{(0)}$ is the smallest positive solution of the equation $\alpha_1 - x_1 + r^{-1}(1-r^{-1})^{r-1} x_1^r = 0$, and $y_1^{(1)}$ is the unique point on $(0, r/(r-1))$ such that

$$\{(y_1^{(1)}, \zeta_1(y_1^{(1)}))\} = \widetilde{\mathcal{D}} \cap \widetilde{\mathcal{E}}_1.$$

(iv) $\widetilde{\mathcal{E}}_2$ is the graph of a strictly increasing and strictly convex function of class C^2, say $\zeta_2(\cdot)$, with domain $[0, y_1^{(2)}]$ and $\zeta_2(0) = y_2^{(0)}$. Here $y_2^{(0)}$ is the smallest positive solution of the equation $\alpha_2 - x_2 + r^{-1}(1-r^{-1})^{r-1} x_2^r = 0$, and $y_1^{(2)}$ is the unique point on $(0, r/(r-1))$ such that

$$\{(y_1^{(2)}, \zeta_2(y_1^{(2)}))\} = \widetilde{\mathcal{D}} \cap \widetilde{\mathcal{E}}_2.$$

Having established the above lemma, we define

$$\mathcal{Z} := \widetilde{\mathcal{E}}_1 \cap \widetilde{\mathcal{E}}_2,$$

i.e., $\mathcal{Z}$ is the set of the zeros of both $\rho_1(\cdot,\cdot)$ and $\rho_2(\cdot,\cdot)$, which necessarily lie in $\mathcal{D}_\rho$. Under the assumption (C), by Lemma 5.5 (parts (iii) and (iv)) we have that:

$$\mathcal{Z} = \widetilde{\mathcal{E}}_1 \cap \widetilde{\mathcal{E}}_2 = \emptyset \Leftrightarrow \zeta_1(x) < \zeta_2(x),\ \forall x \in [y_1^{(0)}, y_1^{(1)}] \cap [0, y_2^{(1)}] \Rightarrow \mathcal{E}_1 \cap \mathcal{E}_2 = \emptyset \Leftrightarrow (\mathbf{Sup}). \tag{22}$$

Hence

$$(\mathbf{Sub}) \Leftrightarrow \overset{\circ}{\mathcal{E}}_1 \cap \overset{\circ}{\mathcal{E}}_2 \neq \emptyset \Rightarrow \mathcal{E}_1 \cap \mathcal{E}_2 \neq \emptyset \Rightarrow \mathcal{Z} = \widetilde{\mathcal{E}}_1 \cap \widetilde{\mathcal{E}}_2 \neq \emptyset. \tag{23}$$

Let

$$\mathbf{z}_* = (z_*, \zeta(z_*)) \in \widetilde{\mathcal{E}}_1 \cap \widetilde{\mathcal{E}}_2 \tag{24}$$

denote the zero of $\rho_1(\cdot,\cdot)$ and $\rho_2(\cdot,\cdot)$ in $\mathcal{D}_{\boldsymbol{\rho}}$ with the smallest first coordinate (which is obviously strictly positive), and set

$$x_* := z_* + \zeta(z_*)(\nu\mu^r)^{1/(r-1)} > 0. \tag{25}$$

Here, $\zeta(\cdot)$ is the function whose graph is $\mathcal{D}_{\boldsymbol{\rho}}$ (see Lemma 5.5(i)).

For later purposes, it is important to note that, as immediate consequence of Lemma 5.5 (parts (iii) and (iv)) we have that, under the assumptions (C) and $(\mathbf{Sub})$, there exists a right neighborhood of $z_* > 0$, say $I^+_{z_*}$, such that $\zeta_2(x_1) > \zeta_1(x_1)$ for any $x_1 \in I^+_{z_*}$, i.e.,

$$\rho_1(x_1, \zeta(x_1)) = \rho_2(x_1, \zeta(x_1)) < 0, \quad \forall\, x_1 \in I^+_{z_*}. \tag{26}$$

The proofs of Lemma 5.5 and Proposition 5.4 are reported in the SM.

5.3. Proof of Theorem 3.2

By Remark 3.1 we necessarily have $\alpha_1 \leq 1$. Let $x_1^{(0)}, x_1^{(1)}$ be the extreme points of the domain of $\zeta(\cdot)$ (see Lemma 5.5(i)), consider the segment

$$\mathcal{S} := \{(x_1, 0):\ x_1 \in [0, x_1^{(0)}]\}$$

and denote by $\overline{\zeta}(\cdot)$ the function whose graph is given by $\mathcal{C} := \mathcal{S} \cup \mathcal{D}_{\boldsymbol{\rho}}$, i.e.,

$$\overline{\zeta}(x_1) := \mathbf{1}_{[x_1^{(0)}, x_1^{(1)}]}(x_1)\zeta(x_1), \quad x_1 \in [0, x_1^{(1)}].$$

We recall that, in our terminology, a strategy is a rule according to which at every time step $t \in [n]$ a community is selected, see Remark 4.2.

We proceed by dividing the proof in four steps. Usually, throughout the proof, for ease of notation, we denote by $c > 0$ a generic positive constant, by $c(\varepsilon)$ if it depends on $\varepsilon > 0$.

Step 1: Identification of a suitable strategy

In this section we are going to formally define the *hybrid* variant of the max-strategy, which has been introduced informally in Sect.ion 4.2. We start from its initial deterministic component, in correspondence of which the trajectory of the normalized number of used nodes in each community follows the curve $(x_1, \zeta(x_1))$, as it can be observed by combining (19), Lemma 5.1, (21) and Lemma 5.5 (i). Therefore, our first goal is to define the corresponding un-normalized 'trajectory' $(w_1(t), w_2(t))$ of the actual number of nodes to be used by time t in the community G_i. With this in mind, we first establish a map between the discrete parameter t and the quantity x_1 that parametrizes the curve $(x_1, \zeta(x_1))$. In particular, we set

$$v(x_1) := \lfloor x_1 g_1 \rfloor + \lfloor \overline{\zeta}(x_1) g_2 \rfloor, \quad x_1 \in [0, x_1^{(1)}]. \tag{27}$$

Note that $v([0,x_1^{(1)}])$ is a subset of $[n] \cup \{0\}$, say $v([0,x_1^{(1)}]) = \{t_0, t_1, \ldots, t_{m+1}\}$. Without loss of generality, we assume $t_0 := v(0) = 0 < t_1 < \ldots < t_m < t_{m+1} := v(x_1^{(1)})$. We consider the right-continuous generalized inverse function of v:

$$v^{-1}(t_s) := \inf\{x_1 \in [0, x_1^{(1)}] : \ v(x_1) \geq t_s\}, \quad s = 0, \ldots, m+1. \tag{28}$$

Finally, we set

$$w_1(t_s) := \lfloor v^{-1}(t_s) g_1 \rfloor, \quad w_2(t_s) := \lfloor \overline{\zeta}(v^{-1}(t_s)) g_2 \rfloor, \quad s = 0, \ldots, m+1., \tag{29}$$

Now, to conclude our construction, we extend the definition of $w_i(\cdot)$, $i = 1,2$, to the set $(0, v(x_1^{(1)})) \cap (\mathbb{N} \cup \{0\})$, by interpolating their values in $v([0,x_1^{(1)}])$ as follows. We note that by construction

$$w_i(t_{s+1}) - w_i(t_s) \in \{0,1\}, \quad i = 1,2, \, s = 0, \ldots, m$$

and

$$\sum_{i=1}^{2} w_i(t_s) = t_s, \quad s = 0, \ldots, m+1.$$

So, for any $s = 0, \ldots, m+1$, $t_{s+1} - t_s \in \{1,2\}$. Consequently, for any $t \in ((0, v(x_1^{(1)})) \cap \mathbb{N}) \setminus v([0,x_1^{(1)}])$, there exists $t_s \in v([0,x_1^{(1)}])$, for some $s \in \{0, \ldots, m\}$, such that $t = t_s + 1$ and $t + 1 = t_{s+1}$. For such a t, we define

$$w_1(t) := \lfloor v^{-1}(t_s) g_1 \rfloor + 1 \quad \text{and} \quad w_2(t) := \lfloor \overline{\zeta}(v^{-1}(t_s)) g_2 \rfloor. \tag{30}$$

Note that, by construction,

$$\sum_{i=1}^{2} w_i(t) = t, \quad \forall\, t \in \{0, \ldots, v(x_1^{(1)})\}.$$

Finally, we need to determine the conditions under which the deterministic strategy we are defining can be successfully employed. To this purpose we define the stopping time

$$T' := \min\{1 \leq t \leq v(x_1^{(1)}) : \ A_i(t-1) < w_i(t), \quad \text{for some } 1 \leq i \leq 2\}. \tag{31}$$

At time step $t \geq 0$, we choose the community G_i, $i = 1,2$, if and only if $C_i(t) = 1$, where

$$C_i(0) := 0, \quad C_i(t) := w_i(t) - w_i(t-1), \quad 1 \leq t < T' \tag{32}$$

and

$$C_1(t) = \mathbf{1}\{A_1(t-1) - U_1(t-1) \geq A_2(t-1) - U_2(t-1)\}, \quad C_2(t) = 1 - C_1(t), \quad T' \leq t \leq T \tag{33}$$

In words, the chosen strategy is deterministic and equal to (32) as long as possible. Indeed, T' is the first time at which the deterministic strategy (32) can not be employed because of the lack of usable and active nodes. Note that setting $\boldsymbol{w}(t) := (w_1(t), w_2(t))$, we have

$$\mathbf{U}(t) = \boldsymbol{w}(t),\ \forall\, t < T' \tag{34}$$

The above strategy is well-defined, indeed, by construction, at each time step $t \leq T'$, there exists only one index $i \in \{1, 2\}$ such that $C_i(t) = 1$ ($C_j(t) = 0$ for $j \neq i$), and by (34) we have $T \geq T'$ with $T - 1 = |\mathcal{G}|$. As already mentioned, we extend the process for $T \leq t \leq n$ by adopting an arbitrary "unfeasible" strategy. The choice of the strategy employed for $t \geq T'$ has no impact on T'.

Step 2: outline of the proof.

It can be easily seen that

$$\lim_{\delta\to 0}\lim_{n\to\infty} v(z_* \pm \delta)/v(z_*) = 1.$$

Therefore, for any $\varepsilon > 0$, there exist $\delta_\varepsilon > 0$ and $n_\varepsilon \in \mathbb{N}$ such that for any $n \geq n_\varepsilon$, it holds $v(z_* + \delta_\varepsilon) < (1+\varepsilon)v(z_*)$ and $v(z_* - \delta_\varepsilon) > (1-\varepsilon)v(z_*)$. So, for an arbitrarily fixed $\varepsilon > 0$ and any $n \geq n_\varepsilon$

$$\{|\mathcal{G}|/v(z_*) - 1| > \varepsilon\} = \{|\mathcal{G}| > (1+\varepsilon)v(z_*))\} \cup \{|\mathcal{G}| < (1-\varepsilon)v(z_*)\}$$
$$\subseteq \{|\mathcal{G}| > v(z_* + \delta_\varepsilon)\} \cup \{|\mathcal{G}| < v(z_* - \delta_\varepsilon) \subseteq \{|\mathcal{G}| \geq v(z_* + \delta_\varepsilon)\} \cup \{T' \leq v(z_* - \delta_\varepsilon)\}.$$

Since $v(z_*)/g_1 \to x_*$, the claim then follows if we prove that, for any $\delta > 0$ small enough there exists a positive constant $c(\delta) > 0$ such that

$$P(T' \leq v(z_* - \delta)) = O(\mathrm{e}^{-c(\delta)g_1}) \tag{35}$$

$$P(|\mathcal{G}| \geq v(z_* + \delta)) = O(\mathrm{e}^{-c(\delta)g_1}). \tag{36}$$

Step 3: proof of (35)

We divide the proof of (35) in three parts. In Step 3.1 we prove the inequality

$$P(T' \leq v(z_* - \delta)) < \mathfrak{T}_1 + \mathfrak{T}_2, \tag{37}$$

where

$$\mathfrak{T}_1 := \sum_{s=0}^{v(x_1^{(0)})-1} P(\mathrm{Bin}(n_1 - a_1, b_1((w_1(s), 0))) < w_1(s+1) - a_1) \tag{38}$$

$$\mathfrak{T}_2 := \sum_{s=v(x_1^{(0)})}^{v(z_*-\delta)} \sum_{i=1}^{2} P(\mathrm{Bin}(n_i - a_i, b_i(\boldsymbol{w}(s))) < w_i(s+1) - a_i). \tag{39}$$

In Step 3.2 we prove

$$\mathfrak{T}_1 = O(\mathrm{e}^{-cg_1}) \tag{40}$$

by applying the concentration inequalities reported in the SM to every addend in (38). In Step 3.3 we prove

$$\mathfrak{T}_2 = O(\mathrm{e}^{-c(\delta)g_1}) \tag{41}$$

again by applying the concentration inequalities reported in the SM to every addend in (39).
Step 3.1: *proof of* (37).
Hereon, we set $\mathbf{A}(t) := (A_1(t), A_2(t))$ and, for two vectors $\mathbf{y} = (y_1, y_2)$ and $\mathbf{y}' = (y_1', y_2')$, we write $\mathbf{y} \geq \mathbf{y}'$ if $y_i \geq y_i'$, $i = 1, 2$. From (31) and (34) we get

$$\{T' > t\} = \{\mathbf{A}(s) \geq \boldsymbol{w}(s+1)\ \forall\, 0 \leq s \leq t-1\} \subseteq \{\mathbf{U}(s) = \boldsymbol{w}(s)\ \forall\, 0 \leq s \leq t\}, \tag{42}$$

which yields

$$\begin{aligned} \{T' = t\} &= \{\mathbf{A}(t-1) < \boldsymbol{w}(t),\, \mathbf{A}(s) \geq \boldsymbol{w}(s+1)\ \forall\, 0 \leq s \leq t-2\} && (43)\\ &\subseteq \{\mathbf{A}(t-1) < \boldsymbol{w}(t),\, \mathbf{U}(s) = \boldsymbol{w}(s)\ \forall\, 0 \leq s \leq t-1\}. && (44) \end{aligned}$$

Therefore

$$\begin{aligned} P(T' \leq t) = P\left(\bigcup_{1 \leq s \leq t} \{T' = s\}\right) &= \sum_{s=1}^{t} P(T' = s) \\ &\leq \sum_{s=0}^{t-1} P(\mathbf{A}(s) < \boldsymbol{w}(s+1), \mathbf{U}(h) = \boldsymbol{w}(h)\ \forall 0 \leq h \leq s) && (45)\\ &\leq \sum_{s=0}^{t-1} P(\mathbf{A}(s) < \boldsymbol{w}(s+1), \mathbf{U}(s) = \boldsymbol{w}(s)) && (46) \end{aligned}$$

Consequently,

$$\begin{aligned} P(T' \leq v(z_* - \delta)) \leq &\sum_{s=0}^{v(z_*-\delta)} P(\mathbf{A}(s) < \boldsymbol{w}(s+1) \mid \mathbf{U}(s) = \boldsymbol{w}(s)) \\ = &\sum_{s=0}^{v(x_1^{(0)})-1} \Big(P(S_1(s) + a_1 - w_1(s+1) < 0 \mid \mathbf{U}(s) = (w_1(s), 0)) \Big) \\ &+ \sum_{s=v(x_1^{(0)})}^{v(z_*-\delta)-1} \sum_{i=1}^{2} P(S_i(s) + a_i - w_i(s+1) < 0 \mid \mathbf{U}(s) = \boldsymbol{w}(s)), && (47) \end{aligned}$$

where we used the fact that $w_2(s) = 0$ for $s = 1, \ldots, v(x_1^{(0)})$. The inequality (37) follows from (47), noticing that (17) yields

$$\sum_{s=0}^{v(x_1^{(0)})-1} P(S_1(s) + a_1 - w_1(s+1) < 0 \mid \mathbf{U}(s) = (w_1(s), 0)) = \mathfrak{T}_1$$

and

$$\sum_{s=v(x_1^{(0)})}^{v(z_*-\delta)-1}\sum_{i=1}^{2}P(S_i(s)+a_i-w_i(s+1)<0 \mid \mathbf{U}(s)=\boldsymbol{w}(s))=\mathfrak{T}_2.$$

Step 3.2: *proof of* (40).
We first note that, since $w_2(s)=0$ and therefore $w_1(s)=s$ for $s<v(x_1^{(0)})$, we have

$$\begin{aligned}\mathfrak{T}_1 = &\sum_{s=0}^{\min(a_1,v(x_1^{(0)}))-1} P\Big(\mathrm{Bin}(n_1-a_1,b_1(s,0))<s+1-a_1\Big)\\ &+\sum_{s=\min(a_1,v(x_1^{(0)}))}^{v(x_1^{(0)})-1} P\Big(\mathrm{Bin}(n_1-a_1,b_1(s,0))<s+1-a_1\Big),\end{aligned}$$

with the convention that the second addend on the right hand side is null when $\min(a_1,v(x_1^{(0)}))=v(x_1^{(0)})$. Now, note that by construction

$$\sum_{s=1}^{\min(a_1,v(x_1^{(0)}))-1} P\Big(\mathrm{Bin}(n_1-a_1,b_1(s,0))<s+1-a_1\Big)=0.$$

Therefore, $\mathfrak{T}_1$ is not null only if $a_1<v(x_1^{(0)})$, and so

$$\mathfrak{T}_1 \le \sum_{s=a_1}^{v(x_1^{(0)})-1} P\Big(\mathrm{Bin}(n_1-a_1,b_1(s,0))<s+1-a_1\Big).$$

Now we are going to bound each addend of the sum in the right-hand side by using the inequality (81) in the SM. For any $s\in\{a_1,\dots,v(x_1^{(0)})-1\}$, we have $v^{-1}(s)=s/g_1=w_1(s)/g_1$. Moreover observe that, since $a_1/g_1\to\alpha_1<x_1^{(0)}$, for n sufficiently large $a_1/g_1\in[\alpha_1/2,x_1^{(0)}]$. Similarly, since $v^{-1}(v(x_1^{(0)})-1)=(v(x_1^{(0)})-1)/g_1=(\lfloor x_1^{(0)}g_1\rfloor-1)/g_1\uparrow x_1^{(0)}$, for sufficiently large n we have $v^{-1}(v(x_1^{(0)})-1)\in[\alpha_1/2,x_1^{(0)}]$. Then, as an immediate consequence of the monotonicity of the involved functions, for n sufficiently large, let us say $n>n'$, we have that $v^{-1}(s)\in[\alpha_1/2,x_1^{(0)}]$ for any $s\in\{a_1,\dots,v(x_1^{(0)})-1\}$. Hence we can apply Lemma 5.2 and conclude that, for $n>n''$ (n'' not depending on s and not smaller than n'):

$$\begin{aligned}R_1((s,0))&=R_1((w_1(s),0))=R_1((\lfloor v^{-1}(s)g_1\rfloor,0))=R_1((v^{-1}(s)g_1,0))\\ &>\frac{1}{2}(\rho_1(v^{-1}(s),0))g_1\ge\inf_{x\in[\alpha_1/2,x_1^{(0)}]}\frac{1}{2}(\rho_1(x,0))g_1=\frac{1}{2}\rho_1(\mathbf{x}^{(0)})g_1,\end{aligned}$$

where for the latter equation we have used the property that $\rho_1(\cdot,0)$ is strictly decreasing on $(0,r/(r-1))$. In conclusion, we have, for all $n>n''$:

$$E[\mathrm{Bin}(n_1-a_1,b_1(s,0)))]=R_1((s,0))-a_1+s\geq\frac{1}{2}\rho_1(\mathbf{x}^{(0)})g_1+s-a_1>s+1-a_1.$$

Therefore, we can apply inequality (81) in the SM. Note that, since the mapping $x\mapsto x/(x+y)$, for fixed $y>0$, is strictly increasing on $(0,\infty)$, and the function H defined by (79) is decreasing on $(0,1)$, for every $n>n''$, we have

$$\begin{aligned}P&\left(\mathrm{Bin}(n_1-a_1,b_1((s,0))<s+1-a_1\right)=P\left(\mathrm{Bin}(n_1-a_1,b_1((s,0))\leq s-a_1\right)\\&\leq\exp\left(-\frac{1}{2}\rho_1(\mathbf{x}^{(0)})g_1H\left(\frac{s-a_1}{\frac{1}{2}\rho_1(\mathbf{x}^{(0)})g_1+s-a_1}\right)\right)\leq\exp\left(-\frac{1}{2}\rho_1(\mathbf{x}^{(0)})g_1H\left(\frac{x_1^{(0)}}{\frac{1}{2}\rho_1(\mathbf{x}^{(0)})+x_1^{(0)}}\right)\right).\end{aligned}$$

In conclusion, defined $c':=\frac{1}{2}\rho_1(\mathbf{x}^{(0)})H\left(\frac{x_1^{(0)}}{\frac{1}{2}\rho_1(\mathbf{x}^{(0)})+x_1^{(0)}}\right)$ and $c=\frac{1}{2}c'$, for every $n\geq n''$ we have

$$\mathfrak{T}_1<x_1^{(0)}g_1\mathrm{e}^{-c'g_1}=O(\mathrm{e}^{-cg_1}),$$

which yields (40).

Step 3.3: *proof of* (41).
To prove (41) we follow the same lines as in the proof of (40). As first step we show that for a sufficiently large n (independently from s) $(w_1(s)/g_1,w_2(s)/g_2)$ is contained in a properly compact set $\mathcal{C}_{z_*-\delta,\varepsilon_0}$ satisfying the following property:

$$\mathcal{C}_{z_*-\delta,\varepsilon_0}\subset\{\mathbf{x}\in\mathcal{D}:\ \rho_1(\mathbf{x})>0,\rho_2(\mathbf{x})>0\}.$$

Then we bound $\mathfrak{T}_2$ as follows:

$$\begin{aligned}\mathfrak{T}_2&\leq\sum_{s=v(x_1^{(0)})}^{v(z_*-\delta)-1}\sum_{i=1}^{2}P(\mathrm{Bin}(n_i-a_i,b_i(\boldsymbol{w}(s))/g_1<(w_i(s)-a_i+1)/g_i)\\&\leq v(z_*-\delta)(\mathfrak{s}_1(\varepsilon)+\mathfrak{s}_2(\varepsilon)),\end{aligned}\tag{48}$$

for any $\varepsilon>0$, where

$$\mathfrak{s}_i(\varepsilon):=\sup_{\mathbf{x}\in\mathcal{C}_{z_*-\delta,\varepsilon_0}}P\left(\mathrm{Bin}(n_i-a_i,b_i(\lfloor\mathbf{x}g\rfloor))/g_i<x_i-\alpha_i+\varepsilon\right),\quad i=1,2.\tag{49}$$

Then, exploiting Lemma 5.2 and the inequality (81) in the SM, we are going to show that there exists $\varepsilon=\varepsilon(\delta)$ such that

$$\mathfrak{s}_i(\varepsilon(\delta))=O(\mathrm{e}^{-c(\delta)g_1}),\quad i=1,2.\tag{50}$$

Then (41) immediately follows.

Step 3.3.1: Definition of $\mathcal{C}_{z_*-\delta,\varepsilon_0}$.

Let $\mathcal{C}_x$ be the graph of the function $\overline{\zeta}(\cdot)$ restricted to $[x_1^{(0)}, x]$, for an arbitrary $x \in (x_1^{(0)}, x_1^{(1)}]$. Clearly, for any x, $\mathcal{C}_x$ is a compact set of $\mathbb{R}^2$. Using Lemma 5.5, it is easily seen that, for any δ,

$$\mathcal{C}_{z_*-\delta} \subset \mathcal{C}_{z_*} \cap \{\mathbf{x} \in \mathcal{D} : \ \rho_1(\mathbf{x}) > 0, \rho_2(\mathbf{x}) > 0\}.$$

For $\varepsilon > 0$, let $\mathcal{C}_{z_*-\delta,\varepsilon}$ be the ε-thickening of $\mathcal{C}_{z_*-\delta}$, i.e.,

$$\mathcal{C}_{z_*-\delta,\varepsilon} := \{\mathbf{x} \in \mathbb{R}^2 : \ \mathrm{dist}(\mathbf{x}, \mathcal{C}_{z_*-\delta}) \leq \varepsilon\},$$

where, for $\mathcal{B} \subset \mathbb{R}^2$,

$$\mathrm{dist}(\mathbf{x}, \mathcal{B}) := \inf\{\|\mathbf{x} - \mathbf{y}\| : \ \mathbf{y} \in \mathcal{B}\}$$

and $\|\cdot\|$ is the Euclidean norm. By the regularity properties of the functions ρ_i, $i = 1, 2$, easily follows that there exists $\varepsilon_0 > 0$ small enough such that

$$\mathcal{C}_{z_*-\delta,\varepsilon_0} \subset \{\mathbf{x} \in \mathcal{D} : \ \rho_1(\mathbf{x}) > 0, \rho_2(\mathbf{x}) > 0\}.$$

Step 3.3.2: proof of the relation $(w_1(s)/g_1, w_2(s)/g_2) \in \mathcal{C}_{z_*-\delta,\varepsilon_0}$.

We are going to show that there exists a positive integer n_{ε_0} (not depending on s) such that $(w_1(s)/g_1, w_2(s)/g_2) \in \mathcal{C}_{z_*-\delta,\varepsilon_0}$ for any $n > n_{\varepsilon_0}$ and $v(x_1^{(0)}) \leq s \leq v(z_* - \delta) - 1$. Indeed, given an arbitrary n, for any $v(x_1^{(0)}) \leq s \leq v(z_* - \delta) - 1$, we have

$$\frac{\lfloor v^{-1}(s) g_1 \rfloor}{g_1} \leq \frac{w_1(s)}{g_1} \leq \frac{\lfloor v^{-1}(s) g_1 \rfloor + 1}{g_1}$$

and

$$\frac{\lfloor \overline{\zeta}(v^{-1}(s)) g_2 \rfloor}{g_2} \leq \frac{w_2(s)}{g_2} \leq \frac{\lfloor \overline{\zeta}(v^{-1}(s)) g_2 \rfloor + 1}{g_2}$$

These relations imply

$$\Big| \frac{w_1(s)}{g_1} - v^{-1}(s) \Big| \leq 1/g_1$$

and

$$\Big| \frac{w_2(s)}{g_2} - \overline{\zeta}(v^{-1}(s)) \Big| \leq 1/g_2$$

for any $v(x_1^{(0)}) \leq s \leq v(z_* - \delta) - 1$. Therefore we can select n_{ε_0} such that

$$\|(w_1(s)/g_1, w_2(s)/g_2) - (v^{-1}(s), \overline{\zeta}(v^{-1}(s)))\| \leq \varepsilon_0,$$

and since $(v^{-1}(s), \overline{\zeta}(v^{-1}(s)) \in \mathcal{C}_{z_*-\delta}$ for any $v(x_1^{(0)}) \leq s \leq v(z_* - \delta)$, we deduce that $(w_1(s)/g_1, w_2(s)/g_2) \in \mathcal{C}_{z_*-\delta,\varepsilon_0}$, for any $n > n_{\varepsilon_0}$.

Step 3.3.3: proof of (50).
We shall show (50) for $i=1$, indeed the case $i=2$ can be proved similarly. Setting $\epsilon(\delta):=\min_{\mathbf{x}\in\mathcal{C}_{z_*-\delta,\varepsilon_0}}\rho_1(\mathbf{x})>0$ and $\varepsilon(\delta):=\frac{1}{4}\epsilon(\delta)$, we have

$$\frac{(n_1-a_1)b_1(\lfloor\mathbf{x}g\rfloor)}{g_1}>r^{-1}(1-r^{-1})^{r-1}(x_1+\chi_{12}x_2)^r-\frac{1}{4}\epsilon(\delta)=x_1-\alpha_1+\rho_1(\mathbf{x})-\frac{1}{4}\epsilon(\delta)$$
$$\geq x_1-\alpha_1+\epsilon(\delta)-\frac{1}{4}\epsilon(\delta)>x_1-\alpha_1+\frac{3}{4}\epsilon(\delta)\qquad\text{for all }\mathbf{x}\in\mathcal{C}_{z_*-\delta,\varepsilon_0}. \tag{51}$$

Therefore, by concentration inequality (81) in the SM for all n large enough, we have

$$\sup_{\mathbf{x}\in\mathcal{C}_{z_*-\delta,\varepsilon_0}}P\left(\mathrm{Bin}(n_1-a_1,b_1(\lfloor\mathbf{x}g\rfloor))\leq\left(x_1-\alpha_1+\frac{1}{4}\epsilon(\delta)\right)g_1\right)$$
$$\leq\sup_{\mathbf{x}\in\mathcal{C}_{z_*-\delta,\varepsilon_0}}\exp\left(-(x_1-\alpha_1+\frac{3}{4}\epsilon(\delta))g_1H\left(\frac{x_1-\alpha_1+\frac{1}{4}\epsilon(\delta)}{x_1-\alpha_1+\frac{3}{4}\epsilon(\delta)}\right)\right) \tag{52}$$
$$=O(\mathrm{e}^{-c(\epsilon(\delta))g_1})=O(\mathrm{e}^{-c(\delta)g_1}),$$

where in (52) we used (51) and the fact that H decreases on $(0,1)$.

Step 4: proof of (36)**.**

For $\delta>0$, define the random time

$$Q(\delta):=\max\{t:\ U_1(t)\leq\mathfrak{z}_1,\ U_2(t)\leq\mathfrak{z}_2\}$$

where $\mathfrak{z}_1:=\lfloor(z_*+\delta)g_1\rfloor$ and $\mathfrak{z}_2:=\lfloor\overline{\zeta}(z_*+\delta)g_2\rfloor$. Note that by construction

$$\text{either }U_1(Q(\delta))=\mathfrak{z}_1\qquad\text{and}\qquad U_2(Q(\delta))\leq\mathfrak{z}_2,$$
$$\text{or }U_1(Q(\delta))<\mathfrak{z}_1\qquad\text{and}\qquad U_2(Q(\delta))=\mathfrak{z}_2\quad\text{almost surely.}$$

In other words, defining, for $\mathbf{v}\in\mathbb{N}^2$, the sets

$$\mathcal{F}_{\mathbf{v}}:=\mathcal{F}_{\mathbf{v}}^{(1)}\cup\mathcal{F}_{\mathbf{v}}^{(2)},$$

$$\mathcal{F}_{\mathbf{v}}^{(1)}:=\{(w_1,w_2)\in\mathbb{N}^2:\ w_1=v_1,\ w_2\leq v_2\},\qquad\mathcal{F}_{\mathbf{v}}^{(2)}:=\{(w_1,w_2)\in\mathbb{N}^2:\ w_2=v_2,\ w_1\leq v_1\},$$

the random vector $\mathbf{U}(Q(\delta))$ (whose components are $U_i(Q(\delta))$, $i=1,2$) almost surely satisfies

$$\mathbf{U}(Q(\delta))\in\mathcal{F}_{(\mathfrak{z}_1,\mathfrak{z}_2)}\quad\text{with}\quad|\mathcal{F}_{(\mathfrak{z}_1,\mathfrak{z}_2)}|=\mathfrak{z}_1+\mathfrak{z}_2+1.$$

As immediate consequence, we have that almost surely

$$Q(\delta)=U_1(Q(\delta))+U_2(Q(\delta))\leq\mathfrak{z}_1+\mathfrak{z}_2=v(z_*+\delta).$$

Therefore

$$\{|\mathcal{G}|\geq v(z_*+\delta)\}\subseteq\bigcap_{i=1}^{2}\bigcap_{t\leq v(z_*+\delta)}\{S_i(t)+a_i-U_i(t)\geq0\}$$

$$\subseteq \bigcap_{i=1}^{2} \{S_i(Q(\delta)) + a_i - U_i(Q(\delta)) \geq 0\}$$

$$= \bigcup_{\mathbf{u} \in \mathcal{F}_{(\mathfrak{z}_1,\mathfrak{z}_2)}} \bigcap_{i=1}^{2} \{S_i(Q(\delta)) + a_i - U_i(Q(\delta)) \geq 0, \mathbf{U}(Q(\delta)) = \mathbf{u}\},$$

and so

$$P(|\mathcal{G}| \geq v(z_* + \delta)) \leq \sum_{\mathbf{u} \in \mathcal{F}_{(\mathfrak{z}_1,\mathfrak{z}_2)}} P\left(\bigcap_{i=1}^{2} \{S_i(Q(\delta)) + a_i - U_i(Q(\delta)) \geq 0\} \,\middle|\, \mathbf{U}(Q(\delta)) = \mathbf{u} \right)$$

$$\leq (\mathfrak{z}_1 + \mathfrak{z}_2 + 1) \max_{\mathbf{u} \in \mathcal{F}_{(\mathfrak{z}_1,\mathfrak{z}_2)}} P\left(\bigcap_{i=1}^{2} \{S_i(u_1 + u_2) + a_i - u_i \geq 0\} \,\middle|\, \mathbf{U}(Q(\delta)) = \mathbf{u} \right) \tag{53}$$

$$\leq (\mathfrak{z}_1 + \mathfrak{z}_2 + 1) \max_{1 \leq j \leq 2} \max_{\mathbf{u} \in \mathcal{F}^{(j)}_{(\mathfrak{z}_1,\mathfrak{z}_2)}} P\left(\bigcap_{i=1}^{2} \{S_i(u_1 + u_2) + a_i - u_i \geq 0\} \,\middle|\, \mathbf{U}(Q(\delta)) = \mathbf{u} \right). \tag{54}$$

Note that, for fixed $j \in \{1, 2\}$ and $\mathbf{u} \in \mathcal{F}^{(j)}_{(\mathfrak{z}_1,\mathfrak{z}_2)}$,

$$P\left(\bigcap_{i=1}^{2} \{S_i(u_1 + u_2) + a_i - u_i \geq 0\} \,\middle|\, \mathbf{U}(Q(\delta)) = \mathbf{u} \right) \leq P\left(S_j(u_1 + u_2) + a_j \geq \mathfrak{z}_j \,\middle|\, \mathbf{U}(Q(\delta)) = \mathbf{u} \right)$$
$$= P(\mathrm{Bin}(n_j - a_j, b_j(\mathbf{u})) \geq \mathfrak{z}_j - a_j) \leq P(\mathrm{Bin}(n_j - a_j, b_j((\mathfrak{z}_1, \mathfrak{z}_2)) \geq \mathfrak{z}_j - a_j),$$

where the latter inequality follows from the stochastic ordering properties of the binomial distribution with respect to its arguments. Note, indeed, that $b_j(\mathbf{u})$ (as defined in (18)) is increasing with respect to the components of $\mathbf{u}$. Combining this inequality with (54) we have

$$P(|\mathcal{G}| \geq v(z_* + \delta)) \leq (\mathfrak{z}_1 + \mathfrak{z}_2 + 1) \max_{1 \leq j \leq 2} P(\mathrm{Bin}(n_j - a_j, b_j(\mathfrak{z}_1, \mathfrak{z}_2) \geq \mathfrak{z}_j - a_j).$$

Since

$$\mathfrak{z}_1 + \mathfrak{z}_2 + 1 \sim v(z_* + \delta)$$

the claim then follows if we prove that, for an arbitrarily fixed $i \in \{1, 2\}$, the quantity

$$P(\mathrm{Bin}(n_i - a_i, b_i(\mathfrak{z}_1, \mathfrak{z}_2)) \geq \mathfrak{z}_i - a_i)$$

goes to zero exponentially fast with respect to g_1. For this we employ again the concentration inequality (80) in the SM. Since ideas and computations are similar to those in the proof of (35), we skip some details. By Lemma 5.1 we have

$$(n_1 - a_1) b_1(\mathfrak{z}_1, \mathfrak{z}_2) \sim ((z_* + \delta) - \alpha_1) g_1 + \rho_1(z_* + \delta, \overline{\zeta}(z_* + \delta)) g_1. \tag{55}$$

and

$$(n_2 - a_2) b_2(\mathfrak{z}_1, \mathfrak{z}_2) \sim (\overline{\zeta}(z_* + \delta) - \alpha_2) g_2 + \rho_2(z_* + \delta, \overline{\zeta}(z_* + \delta)) g_2. \tag{56}$$

Therefore

$$\frac{\mathfrak{z}_1 - a_1}{(n_1 - a_1)b_1(\mathfrak{z}_1, \mathfrak{z}_2)} \to \frac{(z_* + \delta) - \alpha_1}{z_* + \delta - \alpha_1 + \rho_1(z_* + \delta, \overline{\zeta}(z_* + \delta))} \tag{57}$$

and

$$\frac{\mathfrak{z}_2 - a_2}{(n_2 - a_2)b_2(\mathfrak{z}_1, \mathfrak{z}_2)} \to \frac{\overline{\zeta}(z_* + \delta) - \alpha_2}{\overline{\zeta}(z_* + \delta) - \alpha_2 + \rho_2(z_* + \delta, \overline{\zeta}(z_* + \delta))}. \tag{58}$$

By Lemma 5.5, we have that there exists $\delta_0 > 0$ such that

$$\max_{1 \le i \le 2} \rho_i((z_* + \delta), \overline{\zeta}(z_* + \delta)) = \epsilon(\delta) < 0, \quad \text{for any } 0 < \delta \le \delta_0.$$

Therefore, by (80) in the SM, for all n large enough, we have

$$P\left(\mathrm{Bin}(n_1 - a_1, b_1(\mathfrak{z}_1, \mathfrak{z}_2)) \ge \mathfrak{z}_1 - a_1\right) \le \exp\left(-(n_1 - a_1)b_1(\mathfrak{z}_1, \mathfrak{z}_2) H\left(\frac{\mathfrak{z}_1 - a_1}{(n_1 - a_1)b_1(\mathfrak{z}_1, \mathfrak{z}_2)}\right)\right),$$

where $H(x) := 1 - x + x \log x, \quad x > 0, \quad H(0) = 1$. The exponential decay of

$$P(\mathrm{Bin}(n_1 - a_1, b_1(\mathfrak{z}_1, \mathfrak{z}_2)) \ge \mathfrak{z}_1 - a_1)$$

easily follows combining this latter inequality with (55) and (57), and using that H increases on $(1, \infty)$. Reasoning in the same way, but using (56) and (58) in place of (55) and (57), respectively, one proves the exponential decay of $P\left(\mathrm{Bin}(n_2 - a_2, b_2(\mathfrak{z}_1, \mathfrak{z}_2)) \ge \mathfrak{z}_2 - a_2\right)$.

5.4. Proof of Theorem 3.3

We give the detailed proof in the case $\alpha_1 \le 1$. The case $\alpha_1 > 1$ follows along similar computations and it is briefly outlined in in the SM.

We denote by $\overline{\zeta}_{\mathrm{ext}}(\cdot)$ the function whose graph is

$$\mathcal{C}_{\mathrm{ext}} := \mathcal{C} \cup \mathcal{R}_{\theta_0},$$

where $\mathcal{C}$ is defined at the beginning of the proof of Theorem 3.2 and $\mathcal{R}_{\theta_0}$, $\theta_0 > 0$ arbitrarily fixed, is the straight line

$$\mathcal{R}_{\theta_0} := \{\mathbf{x} \in \mathbb{R}^2 \setminus \mathcal{D} : \ \mathbf{x} = (x_1, \theta_0(x_1 - x_1^{(1)}) + \overline{\zeta}(x_1^{(1)})), \ x_1 \ge x_1^{(1)}\},$$

i.e.,

$$\overline{\zeta}_{\mathrm{ext}}(x_1) := \mathbf{1}_{[0, x_1^{(1)}]}(x_1)\overline{\zeta}(x_1) + \mathbf{1}_{(x_1^{(1)}, \infty)}(x_1)(\theta_0(x_1 - x_1^{(1)}) + \overline{\zeta}(x_1^{(1)})).$$

Similarly to the proof of Theorem 3.2 (see (27)), we set

$$v(x_1) := \lfloor x_1 g_1 \rfloor + \lfloor \overline{\zeta}_{\mathrm{ext}}(x_1) g_2 \rfloor, \quad x_1 \ge 0$$

and note that $v([0,\infty)) = \{t_s\}_{s\in\mathbb{N}\cup\{0\}}$, for some $t_0 := 0 < t_1 < \ldots < t_m < \ldots$. We define $v^{-1}(t_s)$, $s \in \mathbb{N}\cup\{0\}$, similarly to (28), with obvious changes (i.e., with $[0,\infty)$ in place of $[0, x_1^{(1)}]$ and with $\mathbb{N}\cup\{0\}$ in place of $\{0,\ldots,m+1\}$), $w_i(t_s)$, $i = 1,2$, $s \in \mathbb{N}\cup\{0\}$, similarly to (29), and we extend the definition of $w_i(\cdot)$ to any $t \in \mathbb{N}\setminus v([0,\infty))$ similarly to (30). We define T' as in (31) (with $\{1,\ldots,v(x_1^{(1)})\}$ replaced by $[n]$) and, similarly to the proof of Theorem 3.2, for $t < T'$, the strategy $\{C_i(t)\}$ defined by (32) is adopted. For $T' \le t < T$, we assume that the system switches to the strategy defined by (33). However we wish to emphasize that the choice of the strategy employed when $t \ge T'$ is completely irrelevant for the proof, as it will become clear in the next subsection.

We proceed by giving an outline of the proof and then by dividing the proof itself in five steps. Hereon, for ease of notation, we denote by $c > 0$ a generic positive constant, by $c(\varepsilon)$ if it depends on $\varepsilon > 0$.

Outline of the proof

Let $\varepsilon \in (0,1)$ be small. Since $|\mathcal{G}| \ge T'$, we have

$$P(n - |\mathcal{G}| > \varepsilon n) \le 1 - P(T' \ge \lceil (1-\varepsilon)n \rceil),$$

where $\lceil x \rceil$ denotes the smallest integer greater than or equal to $x \in \mathbb{R}$, and therefore it suffices to show that

$$P(T' < \lceil (1-\varepsilon)n \rceil) \ge O(\mathrm{e}^{-c(\varepsilon)g_1}),$$

for some positive constant $c(\varepsilon) > 0$. We have

$$P(T' < \lceil (1-\varepsilon)n \rceil) = P(T' < kv(x_1^{(1)})) + P(kv(x_1^{(1)}) \le T' < \lfloor p_1^{-1} \rfloor) + P(\lfloor p_1^{-1} \rfloor \le T' < \lceil (1-\epsilon)n \rceil)$$

for some constant $k \in \mathbb{N}$. Therefore the proof is completed if we show that every term on the right-hand side vanishes exponentially fast for sufficiently large n. To this aim, as first step, we give a preliminary bound on $P(T' \in [t_a, t_b))$ for some $t_a, t_b \in [n]$ with $t_a < t_b$.

Step 1: A useful preliminary bound

Note that by (44) we have

$$\begin{aligned}\{T' \in [t_a, t_b)\} &= \cup_{t=t_a}^{t_b-1} \{\mathbf{A}(t-1) < \boldsymbol{w}(t),\, \mathbf{A}(s) \ge \boldsymbol{w}(s+1)\ \forall\, 0 \le s \le t-2\} \\ &\subseteq \cup_{t=t_a}^{t_b-1} \{\mathbf{A}(t-1) < \boldsymbol{w}(t),\, \mathbf{U}(s) = \boldsymbol{w}(s)\ \forall\, 0 \le s \le t-1\}.\end{aligned}$$

Therefore

$$\{T' \in [t_a, t_b)\} \subseteq \{\mathbf{U}(t_a - 1) = \boldsymbol{w}(t_a - 1)\}. \tag{59}$$

Since the paths $A_i(\cdot)$ and the functions $w_i(\cdot)$ are non-decreasing, we have

$$\{\mathbf{A}(t_a - 1) \ge \boldsymbol{w}(t_b), \mathbf{U}(t_a - 1) = \boldsymbol{w}(t_a - 1)\} \cap \{T' \in [t_a, t_b)\} = \emptyset. \tag{60}$$

Combining (59) and (60), we have

$$\begin{aligned}&P(\mathbf{A}(t_a - 1) \ge \boldsymbol{w}(t_b), \mathbf{U}(t_a - 1) = \boldsymbol{w}(t_a - 1)) + P(T' \in [t_a, t_b)) \\ &= P(\mathbf{A}(t_a - 1) \ge \boldsymbol{w}(b), \mathbf{U}(t_a - 1) = \boldsymbol{w}(t_a - 1)) + P(T' \in [t_a, t_b), \mathbf{U}(t_a - 1) = \boldsymbol{w}(t_a - 1)) \\ &\le P(\mathbf{U}(t_a - 1) = \boldsymbol{w}(t_a - 1)),\end{aligned}$$

which yields

$$P(T' \in [t_a, t_b)) \leq P(T' \in [t_a, t_b) \mid \mathbf{U}(t_a - 1) = \boldsymbol{w}(t_a - 1))$$
$$\leq 1 - P(\mathbf{A}(t_a - 1) \geq \boldsymbol{w}(t_b) \mid \mathbf{U}(t_a - 1) = \boldsymbol{w}(t_a - 1)) \leq \sum_{i=1}^{2} P(A_i(t_a - 1) < w_i(t_b) \mid \mathbf{U}(t_a - 1) = \boldsymbol{w}(t_a - 1)). \tag{61}$$

Step 2: Bounding $P(T' < kv(x_1^{(1)}))$

By (46) we have

$$P(T' < kv(x_1^{(1)})) \leq \sum_{s=1}^{kv(x_1^{(1)})-1} P(\mathbf{A}(s) < \boldsymbol{w}(s+1) \mid \mathbf{U}(s) = \boldsymbol{w}(s))$$
$$= \mathfrak{T}_1 + \mathfrak{T}_2 \tag{62}$$

with

$$\mathfrak{T}_1 := \sum_{s=1}^{v(x_1^{(0)})-1} P(\mathbf{A}(s) < \boldsymbol{w}(s+1) \mid \mathbf{U}(s) = \boldsymbol{w}(s)), \quad \mathfrak{T}_2 := \sum_{s=v(x_1^{(0)})}^{kv(x_1^{(1)})-1} P(\mathbf{A}(s) < \boldsymbol{w}(s+1) \mid \mathbf{U}(s) = \boldsymbol{w}(s)).$$

Now, following the same lines as in the proof of (40), we can easily show that

$$\mathfrak{T}_1 = \sum_{s=\min(a_1, v(x_1^{(0)}))}^{v(x_1^{(0)})} P\left(S_1(s) < s + 1 - a_1 \mid \mathbf{U}(s) = (s, 0)\right) = O(\mathrm{e}^{-cg_1}). \tag{63}$$

with $c := \frac{1}{4}\rho_1(\mathbf{x}^{(0)}) H\left(\frac{x_1^{(0)}}{\frac{1}{2}\rho_1(\mathbf{x}^{(0)}) + x_1^{(0)}}\right)$.

Instead, to prove that

$$\mathfrak{T}_2 = O(\mathrm{e}^{-cg_1}),$$

we can follow the same approach as in the proof of (41). Hereon, we skip many details and highlight the main differences. Let $\mathcal{C}_x$ be the graph of the function $\overline{\zeta}_{\text{ext}}(\cdot)$ restricted to $(x_1^{(0)}, x)$, $x > x_1^{(0)}$, and, for $\varepsilon > 0$, let $\mathcal{C}_{x,\varepsilon}$ be the ε-thickening of $\mathcal{C}_x$. As in the proof of Theorem 3.2, one has that there exists $\varepsilon_0 \in (0, 1)$ small enough so that

$$\mathcal{C}_{kv(x_1^{(1)}),\varepsilon_0} \subset \{\mathbf{x} \in \mathcal{D} : \ \rho_1(\mathbf{x}) > 0, \rho_2(\mathbf{x}) > 0\}$$

and it can be shown that there exists n_{ε_0} (not depending on s) such that $(w_1(s)/g_1, w_2(s)/g_2) \in \mathcal{C}_{kv(x_1^{(1)}),\varepsilon_0}$ for any $n > n_{\varepsilon_0}$ and any $v(x_1^{(0)}) \leq s \leq kv(x_1^{(1)}) - 1$. By the assumption $(\mathcal{S}up)$ it follows that $\min_{\mathbf{x} \in \mathcal{C}_{kv(x_1^{(1)}),\varepsilon_0}} \rho_1(\mathbf{x}) =: \epsilon > 0$. Then, proceeding exactly as in the proof of Theorem 3.2, we can show that for n large enough $\mathfrak{T}_2 \leq \mathfrak{s}^{c(\epsilon) g_1}$, where $\mathfrak{s}$ is defined as in (49), with $\mathcal{C}_{kv(x_1^{(1)}),\varepsilon_0}$ in place of $\mathcal{C}_{z_* - \delta, \varepsilon_0}$.

Step 3: Bounding $P(kv(x_1^{(1)}) \le T' < \lfloor p_1^{-1} \rfloor)$

For n so large that so that $\lfloor p_1^{-1} \rfloor - 1 > kv(x_1^{(1)})$, define

$$l := \min\{\ell \ge k : \ p_1 m_\ell \ge 1\}, \quad \text{where } m_\ell := k^{\ell/k} v(x_1^{(1)})$$

Since $m_l \ge \lfloor p_1^{-1} \rfloor$, we have

$$[kv(x_1^{(1)}), \lfloor p_1^{-1} \rfloor] \cap \mathbb{N} \subseteq \bigcup_{\ell=k}^{l-1} [m_\ell, m_{\ell+1}] \cap \mathbb{N}.$$

Now

$$\begin{aligned} P(k\, v(x_1^{(1)}) \le T' < \lfloor p_1^{-1} \rfloor) &\le \sum_{\ell=k}^{l-1} P(m_\ell \le T' < m_{\ell+1}) \\ &\le \sum_{i=1}^{2} \sum_{\ell=k}^{l-1} P(A_i(m_\ell - 1) < w_i(m_{\ell+1}) \mid \mathbf{U}(m_\ell - 1) = \boldsymbol{w}(m_\ell - 1)) \end{aligned}$$

where in the latter inequality we have employed (61). Moreover

$$P(A_i(m_\ell - 1) < w_i(m_{\ell+1}) \mid \mathbf{U}(m_\ell - 1) = \boldsymbol{w}(m_\ell - 1)) = P\left(\text{Bin}(n_i - a_i, b_i(\boldsymbol{w}(m_\ell - 1)) < w_i(m_{\ell+1}) - a_i)\right). \tag{64}$$

Therefore, choosing k large enough and arguing as in the proof of relation (59) in [33], for any $i \in \{1,2\}$, any $\ell \in \{k,\dots,l-1\}$ and all n large enough, we get

$$\begin{aligned} &P(\text{Bin}(n_i - a_i, b_i(\boldsymbol{w}(m_\ell + 1))) < w_i(m_{\ell+1} + 1) - a_i) \\ &\quad \le P(\text{Bin}(n_i, b_i(\boldsymbol{w}(m_\ell + 1))) < w_i(m_{\ell+1} + 1)) \le \mathrm{e}^{-c_1 g_1} \mathrm{e}^{-(\ell - \lceil c \rceil) c_2 g_1}, \end{aligned} \tag{65}$$

for some positive constants $c_1, c_2 > 0$. Finally, by (64), for all n large enough, we have

$$P(k\, v(x_1^{(1)}) \le T' < \lfloor p_1^{-1} \rfloor) \le \sum_{i=1}^{2} \sum_{\ell=\lceil c \rceil}^{l-1} P(\text{Bin}(n_i - a_i, b_i(\boldsymbol{w}(m_\ell - 1))) < w_i(m_{\ell+1}) - a_i) \le c_3 \mathrm{e}^{-c_1 g_1},$$

for some positive constant $c_3 > 0$.

Step 4: Bounding $P(\lfloor p_1^{-1} \rfloor \le T' < \lceil (1-\epsilon) n \rceil)$

Let $c \in (0,1)$ be a small positive constant such that, for all n large enough $P(\text{Bin}(\lfloor p_1^{-1}, p_1) \ge r) \ge 2c$ (see e.g. the proof of Lemma 8.2 Case 3 p. 26 in [24]). For all n large enough we have

$$P(\lfloor p_1^{-1} \rfloor \le T' < \lceil (1-\epsilon) n \rceil) = P(\lfloor p_1^{-1} \rfloor < T' < \lceil cn \rceil) + P(\lceil cn \rceil \le T' < \lceil (1-\epsilon) n \rceil). \tag{66}$$

From (61), we have

$$P(\lfloor p_1^{-1} \rfloor < T' < \lceil cn \rceil) \le 1 - P(\mathbf{A}(\lfloor p_1^{-1} \rfloor - 1) \ge \boldsymbol{w}(\lceil cn \rceil) \mid \mathbf{U}(\lfloor p_1^{-1} \rfloor - 1) = \boldsymbol{w}(\lfloor p_1^{-1} \rfloor - 1))$$

$$\leq \sum_{i=1}^{2} P(\mathrm{Bin}(n_i - a_i, b_i(\boldsymbol{w}(\lfloor p_1^{-1} \rfloor - 1))) < w_i(\lceil cn \rceil) - a_i). \tag{67}$$

Similarly, we get

$$P(\lceil cn \rceil \leq T' < \lceil (1-\epsilon)n \rceil) \leq \sum_{i=1}^{2} P(\mathrm{Bin}(n_i - a_i, b_i(\boldsymbol{w}(\lceil cn \rceil - 1))) < w_i(\lceil (1-\varepsilon)n \rceil) - a_i). \tag{68}$$

The following inequalities are proved in the Step 4 of the proof of Proposition 4.1 in [33] and hold for any $i \in \{1,2\}$ and all n large enough:

$$P(\mathrm{Bin}\Big(n_i - a_i, b_i(\boldsymbol{w}(\lfloor p_1^{-1} \rfloor - 1))) < w_i(\lceil cn \rceil) - a_i\Big) \leq c_1 \mathrm{e}^{-c_2 n}, \quad \text{for some constants } c_1, c_2 > 0$$

$$P\Big(\mathrm{Bin}(n_i - a_i, b_i(\boldsymbol{w}(\lceil cn \rceil - 1))) < w_i(\lceil (1-\varepsilon)n \rceil) - a_i\Big) \leq \mathrm{e}^{-c'(\varepsilon) g_1}, \quad \text{for some constant } c'(\varepsilon) > 0.$$

Therefore $P(\lfloor p_1^{-1} \rfloor \leq T' < \lceil (1-\epsilon)n \rceil) \leq \mathrm{e}^{-c'(\varepsilon) g_1} + c_1 \mathrm{e}^{-c_2 n}$.

References

[1] E. Abbe and C. Sandon. Community detection in general stochastic block models: fundamental limits and efficient algorithms for recovery. *Proceedings of the IEEE 56th Annual Symposium on Foundations of Computer Science*, 670–688, 2015.

[2] E. Abbe, A.S. Bandeira and G. Hall. Exact recovery in the stochastic block model. *IEEE Transactions on Information Theory*, 62: 471–487, 2016.

[3] M. Abdullah and N. Fountoulakis. A phase transition in the evolution of bootstrap percolation processes on preferential attachment graphs. *Random Structures and Algorithms*, 52: 379-418, 2018.

[4] E.M. Airoldi, D.M. Blei, S.E. Fienberg and E.P. Xing. Mixed membership stochastic block models. *Journal of Machine Learning Research*, 9: 1981–2014, 2008.

[5] H. Amini. Bootstrap percolation and diffusion in random graphs with given vertex degrees. *Electronic Journal of Combinatorics*, 17: 1–20, 2010.

[6] H. Amini and N. Fountoulakis. Bootstrap percolation in power-law random graphs. *Journal of Statistical Physics*, 155: 72–92, 2014.

[7] H. Amini, N. Fountoulakis, and K. Panagiotou. Bootstrap percolation in inhomogeneous random graphs. *arXiv preprint arXiv:1402.2815, 2014*.

[8] O. Angel and B. Kolesnik. Large deviations for subcritical bootstrap percolation on the random graph. *ArXiv*: 1705.06815v2, 2018.

[9] J. Balogh and B. Bollobás. Bootstrap percolation on the hypercube. *Probability Theory and Related Fields*, 134: 624–648, 2006.

[10] J. Balogh, Y. Peres and G. Pete. Bootstrap percolation on infinite trees and non-amenable groups. *Combinatorics, Probability and Computing*, 15: 715–730, 2006.

[11] J. Balogh and B.G. Pittel. Bootstrap percolation on the random regular graph. *Random Structures and Algorithms*, 30: 257–286, 2007.

[12] B. Bollobás, K. Gunderson, C. Holmgren, S. Janson and M. Przykucki. Bootstrap percolation on Galton-Watson trees. *Electronic Journal of Probability*, 19: 1–27, 2014.

[13] B. Bollobás, O. Riordan, E. Slivken and P. Smith. The threshold for jigsaw percolation on random graphs. *Electronic Journal of Combinatorics*, 24, Paper #P2.36, 2017.

[14] M. Bradonjić and I. Saniee. Bootstrap percolation on random geometric graphs. *Probability in the Engineering and Informational Sciences,* 28: 169–181, 2014.

[15] C.D. Brummitt, S. Chatterjee, P.S. Dey and D. Sivakoff. Jigsaw percolation: what social networks can collaboratively solve a puzzle?. *The Annals of Applied Probability*, 25: 2013–2038, 2015.

[16] J. Chalupa, P.L. Leath and G.R. Reich. Bootstrap percolation on a Bethe lattice. *Journal of Physics C*, 12: 31–35, 1979.

[17] A. Coja-Oghlan and A. Lanka. Finding planted partitions in random graphs with general degree distributions. *SIAM Journal on Discrete Mathematics*, 23: 1682–1714, 2010.

[18] U. Feige, M. Krivelevich and D. Reichman. Contagious sets in random graphs. *The Annals of Applied Probability*, 27: 2675–2697, 2016.

[19] S. Fortunato. Community detection in graphs. *Physics Reports*, 486: 75–174, 2010.

[20] N. Fountoulakis, M. Kang, C. Koch, T. Makai. A phase transition regarding the evolution of bootstrap processes in inhomogeneous random graphs. *The Annals of Applied Probability*, 28: 990–1051, 2018.

[21] M. Girvan and M.E.J. Newman. Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, 99: 7821–7826, 2002.

[22] P.K. Gopalan and D.M. Blei. Efficient discovery of overlapping communities in massive networks. *Proceedings of the National Academy of Sciences*, 110: 14534–14539, 2013.

[23] C. Holmgren, T. Juškevičius, N. Kettle. Majority bootstrap percolation on $G_{n,p}$. *Electronic Journal of Combinatorics*, 24, Paper #P1.1, 2017.

[24] S. Janson, T. Luczak, T. Turova and T. Vallier. Bootstrap percolation on the random graph $G_{n,p}$. *The Annals of Applied Probability*, 22: 1989–2047, 2012.

[25] S. Janson, R. Kozma, M. Róbert, M. Ruszinkó and Y. Sokolov. Bootstrap percolation on a random graph coupled with a lattice. *Electronic Journal of Combinatorics*, in press.

[26] B Karrer, M.E.J. Newman. Stochastic block models and community structure in networks. *Physical Review E*, 83: 016107, 2011.

[27] D. Kempe, J. Kleinberg and E. Tardos. Maximizing the spread of influence through a social network. *Proceedings of the 9th ACM SIGKDD international conference on knowledge discovery and data mining*, 10 pp., 2003.

[28] L. Massoulié. Community detection thresholds and the weak Ramanujan property. *Proceedings of the 46th annual ACM symposium on theory of computing*, 694–703, 2014.

[29] M.E.J. Newman. Assortative mixing in networks. *Physical Review Letters*, 89: 208701, 2002.

[30] M.D. Penrose. Random geometric graphs. *Oxford University Press*, New York, USA, 2003.

[31] G.P. Scalia–Tomba. Asymptotic final-size distribution for some chain-binomial processes. *Advances in Applied Probability*, 17: 477–495, 1985.

[32] M. Shrestha and C. Moore. Message-passing approach for threshold models of behavior in networks. *Physical Review E*, 89: 022805, 2014.

[33] G.L. Torrisi, M. Garetto and E. Leonardi. A large deviation approach to super-critical bootstrap percolation on the random graph $G_{n,p}$. *Stochastic Processes and their Applications*, 129: 1873–1902, 2019.

[34] G.L. Torrisi, M. Garetto and E. Leonardi. Boostrap percolation on the stochastic block model with k communities. Available at http://arxiv.org/abs/1812.09107, 2020.

[35] T.S. Turova and T. Vallier. Bootstrap percolation on a graph with random and local connections. *Journal of Statistical Physics*, 160: 1249–1276, 2015.

[36] D. Watts. A simple model of global cascades in random networks. *Proceedings of the National Academy of Sciences*, 99: 5766–5771, 2002.