

Learning to maximize global influence from local observations *

Gábor Lugosi

*ICREA & Universitat Pompeu Fabra
Barcelona, Spain*

GABOR.LUGOSI@GMAIL.COM

Gergely Neu

*Universitat Pompeu Fabra
Barcelona, Spain*

GERGELY.NEU@GMAIL.COM

Julia Olkhovskaya

*Universitat Pompeu Fabra
Barcelona, Spain*

JULIA.OLKHOVSKAYA@GMAIL.COM

Editors: ...

Abstract

We study a family online influence maximization problems where at each time $t = 1, \dots, T$, the decision maker selects one from a large number of agents with the goal of maximizing influence. Upon choosing an agent, the decision maker places a piece of information at the agent. The information then spreads in an unobserved network over which the agents communicate. The edges of the network are generated according to some fixed but unknown distribution. The set of influenced nodes is the connected component of the random graph containing the vertex corresponding to the selected agent. The goal of the decision maker is to reach as many nodes as possible. However, the decision maker cannot observe the entire network. The available feedback is only some information about a small neighborhood of the selected vertex. Our results show that such partial local observations can be sufficient for maximizing global influence. We model the underlying random graph as a sparse inhomogeneous Erdős-Rényi graph. We study in detail three specific families of random graph models: stochastic block models, Chung–Lu models and Kronecker random graphs. We show that in these cases one may learn to maximize influence by merely observing the degree of the selected vertex in the generated random graph. We propose sequential learning algorithms that aim at maximizing influence, and provide their theoretical analysis in both the subcritical and supercritical regimes of all considered models.

Keywords: Influence maximization, sequential prediction, multi-armed bandits, stochastic block models

1. Introduction

Finding influential nodes in networks has a long history of study. The problem has been cast in a variety of different ways according to the notion of influence and the information available to a decision maker. We refer the reader to [Kempe, Kleinberg, and Tardos \(2003\)](#); [Chen, Wang, and Wang \(2010\)](#); [Chen, Lakshmanan, and Castillo \(2013a\)](#); [Vaswani, Lakshmanan, and Schmidt \(2015\)](#); [Carpentier and Valko \(2016\)](#); [Wen, Kveton, Valko, and Vaswani \(2017\)](#); [Wang and Chen \(2017\)](#); [Khim, Jog, and Loh \(2019\)](#); [Perrault, Healey, Wen, and Valko \(2020\)](#) and the references therein for recent progress in various directions.

* A part of the material of the paper appeared in the Proceedings of ALT 2019

The most studied influence maximization setup is an offline discrete optimization problem of finding the set of the most influential nodes in a network. This setup assumes that the probability of influencing is known, or at least data is available that allows one to estimate these probabilities. However, such information is often not available or is difficult to obtain. Also, the network over which information spreads is rarely fixed. To avoid such assumptions, we introduce a novel model of influence maximization in a sequential setup, where the underlying network changes every time and the learner has only partial information about the set of influenced nodes.

Specifically, we define and explore a sequential decision-making model in which the goal of a decision maker is to find one among a set of n agents with maximal (expected) influence.

We parametrize the information spreading mechanism by a symmetric $n \times n$ matrix P , whose entries $p_{i,j} \in [0, 1]$ express “affinity” or “probability of communication” between agents i and j . We assume that $p_{i,i} = 0$ for all $i \in [n]$. The matrix P defines an inhomogeneous random graph G in a natural way: an (undirected) edge is present between nodes $i < j$ with probability $p_{i,j}$ and all edges are independent. When two nodes are connected by an edge, information flows between the corresponding agents. Hence, a piece of information placed at a node i spreads to the nodes of the entire connected component of i in G .

In the sequential decision-making process we study, an independent random graph is formed at each time instance $t = 1, \dots, T$ on the vertex set $[n]$. The random graph formed at time t is denoted by G_t . Hence, G_1, \dots, G_T is an independent, identically distributed sequence of random graphs on the vertex set $[n]$, whose distribution is determined by the matrix P .

If the decision maker selects a node $a \in [n]$ at time t , then the information placed at the node spreads to every node of the connected component of a in the graph G_t .

The goal of the decision maker is to spread information as much as possible, that is, to reach as many agents as possible. The *reward* of the decision maker at time t is the number of nodes in the connected component containing the selected node in G_t .

In this paper, we study a setting where the decision maker has no prior knowledge of the distribution P , so she has to learn about this distribution on the fly, while simultaneously attempting to maximize the total reward. This gives rise to a dilemma of *exploration versus exploitation*, commonly studied within the framework of *multi-armed bandit* problems (for a survey, see [Bubeck and Cesa-Bianchi \(2012\)](#)). Indeed, if the decision maker could observe the size of the set of all influenced nodes in every round, the sequential influence maximization problem outlined above could be naturally formulated as a *stochastic multi-armed bandit* problem [Lai and Robbins \(1985\)](#); [Auer, Cesa-Bianchi, and Fischer \(2002a\)](#). However, this direct approach has multiple drawbacks. First of all, in many applications, the number n of nodes is so large that one cannot even hope to maintain individual statistics about each of them, let alone expect any algorithm to identify the most influential node in any reasonable time. More importantly, in most cases of interest, tracking down the set of *all* influenced agents may be difficult or downright impossible due to privacy and computational considerations. This motivates the study of a more restrictive setting where the decision maker has to manage with only partial observations of the set of influenced nodes.

We address this latter challenge by considering a more realistic observation model, where after selecting an agent A_t to be influenced, the learner only observes a local neighbourhood of A_t in the realized random graph G_t , or even only the number of immediate neighbours of A_t (i.e., the degree of vertex A_t in G_t). This model raises the following question: is it possible to maximize global influence while only having access to such local measurements? Our key technical result is answering this question in the positive for some broadly studied random graph models.

The rest of the paper is structured as follows. In Section 1.1 we formalize the sequential influence maximization problem. In Section 1.2 a general model of inhomogeneous random graphs is described and the crucial notions of sub-, and super-criticality are formally introduced. Section 2 is dedicated to the general case when the underlying random graph is an arbitrary inhomogeneous random graph and the learner only knows whether it is in the subcritical or supercritical regime. We show that in both cases online influence maximization is possible by only observing a small “local” neighborhood of the selected node. We provide two separate algorithms and regret bounds for the subcritical and supercritical cases, respectively. In Section 8, we consider the situation when the learner has even less information about the underlying random graph. In particular, we assume that the learner only observes the degree of the selected node in the realized random graph. We study three well-known special cases of inhomogeneous random graphs that are commonly used to model large social networks, namely stochastic block models, the Chung-Lu model, and Kronecker random graphs. We prove that in these three random graph models, degree observations are sufficient to maximize global influence both in the subcritical and supercritical regimes. In Section 4 we provide some discussion and comparison to the previous work. In sections 5, 6, 7, and 8 we present all proofs.

1.1. Problem setup

We now describe our problem and model assumptions formally. We consider the problem of sequential influence maximization on the set of nodes $V = [n]$, formalized as a repeated interaction scheme between a learner and its environment. We assume that node i influences node j with (unknown) probability $p_{i,j}(= p_{j,i})$. At each iteration, a new graph G_t is generated on the vertex set V by independent draws of the edges such that edge (i, j) is present with probability $p_{i,j}$ and all edges are independent. The set of nodes influenced by the chosen node A_t is the connected component of G_t that contains A_t . $C_{i,t}$ denotes the connected component containing vertex i :

$$C_{i,t} = \{v \in V : v \text{ is connected to } i \text{ by a path in } G_t\} .$$

The feedback that the decision maker receives after choosing a node is some “local” information around the chosen vertex A_t in G_t . We consider several feedback models. In the simplest case, the feedback is the degree of vertex A_t in G_t . In another model, the information might consist of the vertices found after a few steps of depth-first exploration of G_t started from vertex A_t . In a general framework, we may define a “local neighborhood” of A_t , denoted by $\widehat{C}_{A_t,t}$, where $\widehat{C}_{A_t,t} \subset C_{A_t,t}$. For each model considered below, we specify later what exactly $\widehat{C}_{A_t,t}$ is. In the general setup, the following steps are repeated for each round $t = 1, 2, \dots$:

1. the learner picks a vertex $A_t \in V$,
2. the environment generates a random graph G_t ,
3. the learner observes the local neighborhood $\widehat{C}_{A_t,t}$,
4. the learner earns the reward $r_{t,A_t} = |C_{A_t,t}|$.

We stress that the learner does *not* observe the reward, only the local neighborhood $\widehat{C}_{A_t,t}$. Define c_i as the expected size of the connected component associated with the node i : $c_i = \mathbb{E}[|C_{i,1}|]$. Ideally, one would like to minimize the *expected regret* defined as

$$R_T = \mathbb{E} \left[\sum_{t=1}^T \left(\max_{i \in V} c_i - c_{A_t} \right) \right] . \tag{1}$$

Since we are interested in settings where the total number of nodes n is very large, even with a fully known random graph model, finding the optimal node maximizing c_i is infeasible. Such computational issues have led to alternative definitions of the regret such as the *approximation regret* Kakade, Kalai, and Ligett (2009); Chen, Wang, and Yuan (2013b); Streeter and Golovin (2009) or the *quantile regret* Chaudhuri, Freund, and Hsu (2009); Chernov and Vovk (2010); Luo and Schapire (2014); Koolen and Van Erven (2015).

In the present paper, we consider the α -quantile regret as our performance measure, which, instead of measuring the learner's performance against the single best decision, uses a near-optimal action as a baseline. For a more technical definition, let i_1, i_2, \dots, i_n be an ordering of the nodes satisfying $c_{i_1} \leq c_{i_2} \leq \dots \leq c_{i_n}$, and denote the α -quantile over the mean rewards as $c_\alpha^* = c_{i_{\lceil(1-\alpha)n\rceil}}$. Then, defining the set $V_\alpha^* = \{i_{\lceil(1-\alpha)n\rceil}, \dots, i_n\}$ as the set of α -near-optimal nodes, we define the α -quantile regret as

$$R_T^\alpha = \mathbb{E} \left[\sum_{t=1}^T \left(\min_{i \in V_\alpha^*} c_i - c_{A_t} \right) \right] = \mathbb{E} \left[\sum_{t=1}^T (c_\alpha^* - c_{A_t}) \right]. \quad (2)$$

1.2. Inhomogeneous Erdős-Rényi random graphs

Next we discuss the random graph models considered in this paper. All belong to the *inhomogeneous Erdős-Rényi model*, that is, edges are present independently of each other, with possibly different probabilities. Moreover, the graphs we consider are *sparse* graphs, that is, the average degree is bounded.

Since we consider large values of n , it is advantageous to formulate this model as in Bollobás, Janson, and Riordan (2007). To this end, let κ be a bounded symmetric non-negative measurable function on $[0, 1] \times [0, 1]$. Each edge (i, j) for $1 \leq i < j \leq n$ is present with probability $p_{i,j} = \min(\kappa(i/n, j/n)/n, 1)$, independently of all other edges. We are interested in random graphs where the average degree is $O(1)$ (as $n \rightarrow \infty$). This assumption makes the problem both more realistic and challenging: denser graphs are connected with high probability, making the problem essentially vacuous. A random graph drawn from the above distribution is denoted by $G(n, \kappa)$.

We consider two fundamentally different regimes of the parameters $G(n, \kappa)$: the *subcritical* case in which the size of the largest connected component is sublinear in n (with high probability), and the *supercritical* case where the largest connected component is at least of size cn for some constant $c > 0$, with high probability. (We say that an event holds *with high probability* if its probability converges to one as $n \rightarrow \infty$.) Such a connected component of linear size is called a *giant component*. These regimes can be formally characterized with the help of the integral operator T_κ , defined by

$$(T_\kappa f)(x) = \int_{(0,1)} \kappa(x, y) f(y) d\mu(y),$$

for any measurable bounded function f , where μ is the Lebesgue measure. We call κ subcritical if $\|T_\kappa\|_2 < 1$ and supercritical if $\|T_\kappa\|_2 > 1$. We use the same expressions for a random graph $G(n, \kappa)$.

It follows from Bollobás, Janson, and Riordan (2007, Theorem 3.1) that, with high probability, $G(n, \kappa)$ has a giant component if it is supercritical, while the number of vertices in the largest component is $o(n)$ with high probability if it is subcritical.

In the analysis for a fixed number of nodes we denote $A_{i,j} = \kappa(i/n, j/n)$ so that $p_{i,j} = \min(A_{i,j}/n, 1)$. This model is sometimes called the *binomial random graph* and was first considered by [Kovalenko \(1971\)](#).

2. Observations of censored component size

First we study a natural feedback model in which the decision maker, unable to explore the entire connected component $C_{i,t}$ of the influenced node i in G_t , resorts to exploring the connected component up to a certain (small) number of nodes. More precisely, we define feedback as the result of counting the number of nodes in $C_{i,t}$ by (say, depth-first search) exploration of the connected component, which stops after revealing K nodes, or before, if $|C_{i,t}| < K$. Here K is a fixed positive integer, independent of the number of nodes n .

The main results of this section show that this type of feedback is sufficient for sequential influence maximization. However, the subcritical and supercritical cases need to be treated separately as they are quite different. In the subcritical case, the expected size of the connected component of any vertex is of constant order while in the supercritical case there exist vertices whose connected component is linear in n . This also means that the rewards – and therefore the per-round regrets – are of different order of magnitude (as a function of n) in the subcritical and supercritical cases.

For simplicity, we assume that the decision maker knows in advance whether the function κ defining the inhomogeneous random graph is subcritical or supercritical, as we propose different algorithms for both cases. We believe that this is a mild assumption. (We also assume that $\|T_\kappa\|_2 \neq 1$, that is, the random graph is not exactly critical.)

2.1. Subcritical case

First we study the subcritical case, that is, we assume that $\|T_\kappa\|_2 < 1$. In this case the proposed influence-maximization algorithm uses the censored size of the connected component of the selected node. That is, for a node $i \in [n]$, we define $u_{i,t}(K)$ as the result of counting the number of nodes in $C_{i,t}$ by exploration of the connected component, which stops after revealing K nodes or before, if $|C_{i,t}| < K$. Hence, the feedback is $u_{i,t}(K) = \min(|C_{i,t}|, K)$.

A key ingredient in our analysis in the subcritical case is an estimate for the lower tail of the size of the connected component containing a fixed vertex. We state it in the following lemma:

Lemma 1 *For any subcritical κ , there exist positive constants $\lambda(\kappa)$, $g(\kappa)$ and $n_0(\kappa)$, such that for any $n \geq n_0$, for any node i in $G(n, \kappa)$, the size of the connected component C_i of a vertex i satisfies*

$$\mathbb{P}[|C_i| > u] \leq e^{-\lambda(\kappa)u} g(\kappa). \quad (3)$$

Unfortunately, there is no closed-form expression for the dependence $\lambda(\kappa)$ and $g(\kappa)$ on κ . The idea of the proof of this lemma relies on the analysis of Theorem 12.5 in [Bollobás, Janson, and Riordan \(2007\)](#). To obtain this concentration result, we show that the size of the connected component in $G(n, \kappa)$ is stochastically dominated by the total progeny of the multitype Poisson branching process with carefully chosen parameters. We introduce branching processes in Section 5 and prove Lemma 1 in Section 6.

Now we are ready to define an estimate of $c_i = \mathbb{E}|C_i|$ in the sequential decision game. For a fixed a constant K , we define the estimate $\hat{u}_{i,t}(K) = (1/t) \sum_{s=1}^t u_{i,s}(K) \mathbb{1}_{\{A_s=i\}}$. Using the

concentration inequality (3), with the choice of the threshold parameter $K = \frac{\log(T)}{\lambda}$ with $\lambda > \lambda(\kappa)$, we get that the bias of $\hat{u}_{i,t}(K)$ is at most $\frac{g(\kappa)}{T}$. We state this result more formally in Lemma 8.

The censored observations are bounded, since $u_{i,t}(K) \in [1, K]$. We use those observations as rewards in our bandit problem and we feed them to an instance of the UCB algorithm (Auer, Cesa-Bianchi, Freund, and Schapire, 2002b). We call the resulting algorithm Local UCB(V_0), defined in Algorithm 1 below.

A minor challenge is that, since we are interested in very large values of n , it is infeasible to use *all* nodes as separate actions in our bandit algorithm. To address this challenge, we propose to *subsample* a set of representative nodes for UCB to play on. The size of the subsampled nodes depends on the quantile α targeted in the regret definition (2) and the time horizon T . Our algorithm uniformly samples a subset V_0 of size

$$|V_0| = \left\lceil \frac{\log T}{\log(1/(1-\alpha))} \right\rceil \quad (4)$$

and plays Local UCB(V_0) for the corresponding regime on the resulting set. Note that the size of V_0 is chosen such that the probability that V_0 does not contain any of the αn nodes with the largest values of c_i is at most $1/T$.

To simplify the presentation, we introduce some more notation. Analogously to the α -optimal reward c_α^* , we define the α -optimal censored component size $u_{*,\alpha}(K) = \min_{i \in V_\alpha^*} u_i(K)$ and we define the corresponding gap parameters $\Delta_{\alpha,i} = (c_\alpha^* - c_i)_+$, $\delta_{\alpha,i}^{sub}(K) = (u_{*,\alpha}(K) - u_i(K))_+$ and $\Delta_{\alpha,\max} = \max_i \Delta_{\alpha,i}$. $N_{i,t} = \sum_{s=1}^t \mathbb{I}_{\{A_s=i\}}$ denotes the number of times node i is selected up to time t .

Algorithm 1 Local UCB(V_0) for subcritical $G(n, \kappa)$.

Parameters: A set of nodes $V_0 \subseteq V$, $K > 0$.

Initialization: Select each node in V_0 once. For each $i \in V_0$, set $N_{i,|V_0|} = 1$ and $\hat{u}_{i,|V_0|} = u_{i,i}(K)$.

For $t = |V_0|, \dots, T$, **repeat**

1. Select any node $A_{t+1} \in \arg \max_i \hat{u}_{i,t}(K) + K \sqrt{\frac{\log t}{N_{i,t}}}$.
 2. Observe $u_{A_{t+1},t+1}(K)$, update $\hat{u}_{i,t+1}$ and $N_{i,t+1}$ for all $i \in [n]$.
-

For the subcritical case, Local UCB(V_0) has the following performance guarantee:

Theorem 2 (Subcritical inhomogeneous random graph) *Assume that κ is subcritical. Let V_0 be a uniform subsample of V with size given in (4) and define the event $\mathcal{E} = \{V_0 \cap V_\alpha^* \neq \emptyset\}$. Then for $G(n, \kappa)$ with $n > n_0(\kappa)$, the expected α -quantile regret of Local UCB(V_0) satisfies*

$$R_T^\alpha \leq \Delta_{\alpha,\max} + \mathbb{E} \left[\sum_{i \in V_0} \Delta_{\alpha,i} \left(\frac{4K^2 \log T}{(\delta_{\alpha,i}^{sub}(K))^2} + 8 \right) \middle| \mathcal{E} \right],$$

where the expectation is taken over the random choice of V_0 . If κ is such that $\lambda(\kappa) > \lambda$, $g(\kappa) < g$, then, taking $K = \frac{\log T}{\lambda}$, we have

$$R_T^\alpha \leq \frac{4 \log T}{\lambda} \sqrt{\frac{T}{\log(1/(1-\alpha))}} + 8\Delta_{\alpha, \max} \left[\frac{\log T}{\log(1/(1-\alpha))} \right] + 2g.$$

We prove Theorem 2 in Section 6. Observe that one may choose the value of K as a constant, regardless of the number n of the nodes. This means that the feedback information is truly ‘‘local’’ in the sense that only a constant number of vertices of the connected component of the selected node need to be explored. How large K needs to be depends on the parameter λ . An undesirable feature of Local UCB(V_0) is that the learner needs to know the parameter λ that depends on the unknown function κ . To resolve this problem we propose a version of a ‘‘doubling trick’’ (see, e.g., Section 2.3 Cesa-Bianchi and Lugosi (2006)). While in our problem it is not possible to control the range of $\lambda(\kappa)$ explicitly, we still can control the frequency with which $|C_i|$ is censored by choosing the range of K . In order to do this, we propose a variation of Local UCB(V_0), such that we split time T into episodes $q = 1, 2, \dots$ in the following way. At the beginning of each episode q , the learner starts a new instance of Local UCB(V_0) with a threshold parameter $K_q = 2^q \log T$ and starts a new time counter t_q . Then, at each time step of the current episode, the learner computes the empirical probability $\hat{p}_q = \frac{1}{t_q} \sum_{\tau=t_{q-1}+1}^{t_q} \mathbb{I}\{|C_{A_\tau, \tau}| > K_q\}$, that is updated each time when the size of connected component of the chosen node exceeds K_q . Once \hat{p}_q gets larger than $\frac{1}{T} + \sqrt{\frac{\ln T}{2(t_q+1)}}$, the episode q finishes and the next episode begins. In this way, the length of each episode and the total number of episodes Q_{\max} are random. We call this algorithm UCB(V_0)-DOUBLE, and show that it has the following performance guarantee:

Algorithm 2 UCB(V_0)-DOUBLE for subcritical $G(n, \kappa)$.

Parameters: A set of nodes $V_0 \subseteq V$, $T > 0$.

Initialization: $K_0 = \log T$, $t = 1$, $q = 0$, $t_q = 0$, $\hat{p}_q = 0$.

While $t \leq T$, **repeat:**

- Select each node in V_0 once. For each $i \in V_0$, set $N_{i,t} = 1$, $\hat{u}_{i,t} = u_{i,t}(K_q)$ and $\hat{p}_q = \frac{1}{|V_0|} \sum_{\tau=t_{q-1}+1}^{t_{q-1}+1+|V_0|} \mathbb{I}\{|C_{A_\tau, \tau}| > K_q\}$.

While $\hat{p}_q \leq \frac{1}{T} + \sqrt{\frac{\ln T}{2(t_q+1)}}$, **repeat:**

1. Select any node $A_{t_q+1} \in \arg \max_i \hat{u}_{i,t_q}(K_q) + K_q \sqrt{\frac{\log T}{N_{i,t_q}}}$.
 2. Observe $u_{A_{t_q+1}, t+1}$,
 3. Update $\hat{u}_{A_{t_q+1}, t_q+1} = \frac{1}{t_q} \sum_{\tau=t_{q-1}+1}^{t_q} u_{i,\tau}(K_q) \mathbb{I}\{A_\tau = A_{t_q+1}\}$, $N_{A_{t_q+1}, t_q+1} = N_{A_{t_q}, t_q} + 1$,
 $\hat{p}_q = \frac{1}{t_q} \sum_{\tau=t_{q-1}+1}^{t_q} \mathbb{I}\{|C_{A_\tau, \tau}| > K_q\}$,
 4. Update $t_q = t_q + 1$ and $t = t + 1$.
- Set $t_{q+1} = 0$, $\hat{p}_{q+1} = 0$, $K_{q+1} = 2K_q$ and $q = q + 1$.
-

Theorem 3 Assume that κ is subcritical and $n > n_0(\kappa)$. Let V_0 be a uniform subsample of V with size given in (4) and define the event $\mathcal{E} = \{V_0 \cap V_\alpha^* \neq \emptyset\}$. Then for $G(n, \kappa)$ with $n > n_0(\kappa)$, the expected α -quantile regret of UCB(V_0)-DOUBLE satisfies

$$R_T^\alpha \leq \Delta_{\alpha, \max} + \frac{64}{3} \mathbb{E} \left[\sum_{i \in V_0} \Delta_{\alpha, i} \left(\frac{\log^3 T}{(\lambda(\kappa) \cdot \min_{q \in [Q_{\max}]} \{\delta_{\alpha, i}^{sub}(2^q \log T)\})^2} + 8 \right) \middle| \mathcal{E} \right],$$

where the expectation is taken over the random choice of V_0 , and

$$R_T^\alpha = \mathcal{O} \left(\frac{\sqrt{T(\log(1/\lambda(\kappa)) + 1)}}{\lambda(\kappa) \sqrt{\ln(1/(1-\alpha))}} \log^2 T \right).$$

The proof of Theorem 3 may be found in Section 6. In Theorem 2 and Theorem 3 we present two types of regret bounds. The first set of these bounds are polylogarithmic¹ in the time horizon T , but show strong dependence on the parameters of the distribution of the graphs G_t . Such bounds are usually called *instance-dependent*, and they are typically interesting in the regime where T grows large. However, these bounds become vacuous for finite T as the gap parameters $\delta_{\alpha, i}^{sup}(\cdot)$ and $\delta_{\alpha, i}^{sub}(\cdot)$ approach zero. This issue is addressed by our second set of guarantees, which offers a bound of $\tilde{O}(\sqrt{|U|T})$ for some set $U \subseteq V$ that holds simultaneously for all problem instances without becoming vacuous in any regime. Such bounds are commonly called *worst-case*, and they are typically more valuable when optimizing performance over a fixed horizon T .

A notable feature of our bounds is that they show no explicit dependence on the number of nodes n . This is enabled by our notion of α -quantile regret, which allows us to work with a small subset of the total nodes as our action set. Instead of n , our bounds depend on the size of some suitably chosen set of nodes U , which is of the order $\text{polylog } T / \log(1/(1-\alpha))$. Notice that this gives rise to an interesting tradeoff: choosing smaller values of α inflates the regret bounds, but, in exchange, makes the baseline of the regret definition stronger (thus strengthening the regret notion itself).

2.2. Supercritical case

Next we address the supercritical case, that is, when $\|T_\kappa\|_2 > 1$. Here the proposed algorithm uses $v_{i,t}(K)$ defined as the indicator whether $|C_i|$ is larger than K , that is, $v_{i,t}(K) = \mathbb{I}_{\{|C_i(G_t)| > K\}}$.

Since the observation is an indicator function, $v_{i,t}(K) \in \{0, 1\}$. Similarly to the subcritical case, we propose a variant of UCB algorithm, Local UCB(V_0), played over a random subsample of nodes of size defined in (4). We define $v_i(K) = \mathbb{E}[v_{i,t}(K)]$ and $v_*(K) = \max_i v_i(K)$. Analogously to the notation introduced for the subcritical regime, we denote $v_{*,\alpha}(K) = \min_{i \in V_\alpha^*} v_i(K)$ and $\delta_{\alpha, i}^{sup}(K) = (v_i(K) - v_{*,\alpha}(K))_+$.

In the supercritical case, the learner receives $v_{i,t}(K)$ as a reward and we design a bandit algorithm based on this form of indicator observations. Note, that $v_{i,t}(K)$ is a Bernoulli random variable with parameter $\mathbb{P}[|C_i(G_t)| > K]$. The following algorithm is a variant of the UCB algorithm of [Auer, Cesa-Bianchi, Freund, and Schapire \(2002b\)](#). Just like before, $N_{i,t}$ denotes the number of times node i is selected up to time t by the algorithm.

1. Upon first glance, the bound of Theorem 2 may appear to be logarithmic, however, notice that the sum involved in the bound has $\Theta(\log T)$ elements, thus technically resulting in a bound of order $\log^2 T$.

Algorithm 3 Local UCB(V_0) for supercritical $G(n, \kappa)$.

Parameters: A set of nodes $V_0 \subseteq V$, $k(n)$.

Initialization: Select each node in V_0 once. For each $i \in V_0$, set $N_{i,|V_0|} = 1$ and $\widehat{v}_{i,|V_0|}(k(n)) = v_{i,i}(k(n))$.

For $t = |V_0|, \dots, T$, **repeat**

1. Select any node $A_{t+1} \in \arg \max_i \widehat{v}_{i,t}(k(n)) + \sqrt{\frac{\log t}{N_{i,t}}}$.
 2. Observe the feedback $v_{i,t}(k(n))$, update $\widehat{v}_{i,t+1}(k(n))$ and $N_{i,t+1}$ for all $i \in [n]$.
-

Local UCB(V_0) for supercritical $G(n, \kappa)$ satisfies the following regret bound:

Theorem 4 *Let V_0 be a uniform subsample of V with size given in (4) and define the event $\mathcal{E} = \{V_0 \cap V_\alpha^* \neq \emptyset\}$. For any $G(n, \kappa)$ with supercritical κ and $n > n_0(\kappa)$, for any function $k : \mathbb{N} \rightarrow \mathbb{N}$ such that $\lim_{n \rightarrow \infty} k(n) = \infty$, we get*

$$\frac{R_T^\alpha}{n} \leq \frac{1}{n} \Delta_{\alpha, \max} + \frac{1}{n} \mathbb{E} \left[\sum_{i \in V_0} \Delta_{\alpha, i} \left(\frac{4 \log T}{(\delta_{\alpha, i}^{\text{sup}}(k(n)))^2} + 8 \right) \middle| \mathcal{E} \right],$$

where the expectation is taken over the random choice of V_0 , and

$$\frac{R_T^\alpha}{n} \leq 9 \left(\frac{\mathbb{E}[|C_1|]}{n} + 1 \right) \left[\frac{\log T}{\log(1/(1-\alpha))} \right] \sqrt{T \log T}.$$

For the proof of Theorem 4, see Section 7. Note that for a supercritical κ , $\mathbb{E}[C_1] = \Theta_n(n)$. Therefore, R_T^α scales linearly with n and hence it is natural to normalize the regret by the number of nodes. Both in the subcritical and supercritical regimes, our bounds scale linearly with the maximal expected reward c^* , which is of $\Theta_n(1)$ in the subcritical case, but is $\Theta_n(n)$ in the supercritical case. The dependence of the obtained bounds on the time horizon T is similar in both regimes. Note that unlike in the subcritical case, the censoring level K is not a constant anymore as we choose it to be $K = k(n)$ for some function k . Hence, strictly speaking, the feedback is not local as the number of vertices that need to be explored is not independent of the number of nodes even if $k(n)$ can grow arbitrarily slowly. Similarly to the subcritical case, a sufficiently large constant value of K would suffice. The value of the constant should be so large that for any vertex i , the conditional probability – conditioned on the event that i is not in the giant component – that the component of i has size larger than K is sufficiently small. Such a constant exists, see (10) below. However, this value depends on the unknown distribution of the underlying random graph. In the subcritical case we solved this problem by applying a doubling trick. This is made possible by the fact that in the subcritical case one observes the “bad” event that a component has size larger than K and therefore censoring occurs. By “trying” increasingly large values of K one eventually finds a value such that the probability of censoring is sufficiently small. However, in the supercritical case, the “bad” event is that even though the selected vertex is not in the giant component, the size of the component is larger than K . Unfortunately, one cannot decide whether the bad event occurs or simply the vertex lies in the giant component. For this reason, we have been unable to apply an analogous doubling trick in the supercritical case. To circumvent this difficulty, we choose K to be growing with n . This guarantees that the bad event occurs with small probability. The price to pay is that the observation is not entirely local in the strict sense.

3. Degree observations

The results of the previous section show that it is possible to learn to maximize influence under very general conditions if the learner has access to the censored size of the connected component, where the size of censoring may be kept much smaller than the size of the entire network. In this section we consider the case when the learner has access to significantly less information. In particular, we study the case when the learner only observes the degree of the selected vertex A_t (i.e., the number of edges adjacent to A_t) in the graph G_t . Under such a restricted feedback, one cannot hope to learn to maximize influence in the full generality of sparse inhomogeneous random graphs as in Section 2. However, we show that in several well-known models of real networks, degree information suffices for influence maximization. In particular, we study three random graph models that have been introduced to replicate properties of large (social) networks appearing in a variety of applications. These are (1) stochastic block models; (2) the Chung-Lu model; and (3) Kronecker random graphs.

3.1. Three random graph models

We start by introducing the three models we study. All of them are special cases of inhomogeneous Erdős-Rényi graphs.

STOCHASTIC BLOCK MODEL.

In the stochastic block model, the probabilities $p_{i,j}$ are defined through the notion of *communities*, defined as elements of a partition H_1, \dots, H_S of the set of vertices V . We refer to the index m of community H_m as the *type* of a vertex belonging to H_m . Each community H_m contains $\alpha_m n$ nodes (assuming without loss of generality that $\alpha_m n$ is an integer). With the help of the community structure, the probabilities $p_{i,j}$ are constructed as follows: if $i \in H_\ell$ and $j \in H_m$, the probability of i and j being connected is given by $p_{i,j} = \frac{K_{\ell,m}}{n}$, where K is a symmetric matrix of size $S \times S$, with positive elements. The random graph from the above distribution is denoted as $G(n, \alpha, K)$.

In the stochastic block model, identifying a node with maximal reward amounts to finding a node from the most influential community. Consequently, it is easy to see that choosing α such that $\alpha > \min_m \alpha_m$, the near-optimal set V_α^* exactly corresponds to the set of optimal nodes, and thus the quantile regret (2) coincides with the regret (1). We consider the stochastic block models satisfying the following simplifying assumptions:

Assumption 1 $K_{l,m} = k > 0$ for all $l \neq m$.

This assumption requires that nodes i, j belonging to different communities are connected with the same probability. Additionally, in our analysis in the supercritical case we make the following natural assumptions:

Assumption 2 For all l , $K_{l,l} > k$.

In plain words, this assumption requires that the density of edges within communities is larger than the density of edges between communities.

CHUNG-LU MODEL

Another thoroughly studied special case of the inhomogeneous Erdős-Rényi model is the so-called *Chung-Lu model* (sometimes referred to as *rank-1 model*) as first defined by [Chung and Lu \(2002\)](#)

(see also [Chung and Lu \(2006\)](#); [Bollobás, Janson, and Riordan \(2007\)](#)). In this model the edge probabilities are defined by a vector $w \in \mathbb{R}^n$ with positive components, representing the “weight” of each vertex. Then the matrix defining the edge probabilities has entries $A_{ij} = w_i w_j$. We assume that the vector w is such that $w_i w_j / n < 1$ for all i, j . In other words, the Chung–Lu model considers rank-1 matrices of the form $A = ww^\top$. The random graph from the Chung–Lu model is denoted by $G(n, w)$. Chung-Lu random graphs replicate some key properties of certain real networks. For instance, if w is a sequence satisfying a power law, then $G(n, w)$ is a power law model, which allows one to model social networks, see [Chung and Lu \(2006\)](#).

KRONECKER GRAPHS

Kronecker random graphs were introduced by [Leskovec, Chakrabarti, Kleinberg, and Faloutsos \(2005\)](#); [Leskovec \(2008\)](#); [Leskovec, Chakrabarti, Kleinberg, Faloutsos, and Ghahramani \(2010\)](#) as models of large networks appearing in various applications, including social networks. The matrix P of the edge probabilities of a Kronecker random graph $G_{n, P^{[k]}}$ is defined recursively. The model is parametrized by the constants $\zeta, \beta, \gamma \in [0, 1]$. Here one assumes that the number of vertices n is a power of 2. Starting from a 2×2 seed matrix,

$$P^{[1]} = \begin{bmatrix} \zeta & \beta \\ \beta & \gamma \end{bmatrix},$$

we define the matrices $P^{[2]}, \dots, P^{[k]}$ such that for each $i = 2, \dots, k$, $P^{[i]}$ is a $2^i \times 2^i$ matrix obtained from $P^{[i-1]}$ by

$$P^{[i]} = \begin{bmatrix} \zeta P^{[i-1]} & \beta P^{[i-1]} \\ \beta P^{[i-1]} & \gamma P^{[i-1]} \end{bmatrix}.$$

Finally $P = P^{[k]}$. Hence, the Kronecker random graph $G_{n, P^{[k]}}$ has $n = 2^k$ vertices, where each vertex i is characterised by a binary string $s_i \in \{0, 1\}^k$, such that the probability of an edge between nodes i and j is equal to $p_{i,j} = \zeta^{\langle s_i, s_j \rangle} \gamma^{\langle \bar{1} - s_i, \bar{1} - s_j \rangle} \beta^{k - \langle s_i, s_j \rangle - \langle \bar{1} - s_i, \bar{1} - s_j \rangle}$, where $\bar{1} = (1, \dots, 1) \in \{0, 1\}^k$ denotes the all-one vector and $\langle \cdot, \cdot \rangle$ is the usual inner product. [Leskovec, Chakrabarti, Kleinberg, and Faloutsos \(2005\)](#) show that a Kronecker graph with properly tuned values of ζ, β, γ replicates properties of real world networks, such as small diameter, clustering, and heavy-tailed degree distribution.

3.2. Learning with degree feedback: stochastic block models and Chung–Lu graphs

In this section we introduce an online influence maximization algorithm that only uses the degree of the selected node as feedback information. The algorithm is a variant of the kl-UCB algorithm, that was proposed and analyzed by [Garivier and Cappé \(2011\)](#); [Maillard, Munos, and Stoltz \(2011\)](#); [Cappé, Garivier, Maillard, Munos, and Stoltz \(2013\)](#); [Lai \(1987\)](#). The main reason why learning is possible based on degree observations only is that nodes with the largest expected degrees μ^* are exactly the ones with the largest influence c^* . This (nontrivial) fact holds in both the stochastic block model (under Assumptions 1 and 2) and the Chung–Lu model, across both the subcritical and supercritical regimes. These facts are proven in Sections 8.1 and 8.2.

We define $X_{t,i}$ as the degree of node i in the realized graph G_t , and define $\mu_i = \mathbb{E}[X_{1,i}]$ as the expected degree of node i . We also define $c^* = \max_i c_i$ and $\mu^* = \max_i \mu_i$.

Algorithm 4 d -UCB(V_0)

Parameters: A set of nodes $V_0 \subseteq V$.**Initialization:** Select each node in V_0 once. Observe the degree $X_{i,i}$ of vertex i in the graph G_i for $i = 1, \dots, |V_0|$. For each $i \in V_0$, set $N_i(|V_0|) = 1$ and $\hat{\mu}_i(|V_0|) = X_{i,i}$.**For** $t = |V_0|, \dots, T$, **repeat**

1. For each node, compute

$$U_i(t) = \sup \left\{ \mu : \mu - \hat{\mu}_i(t) + \hat{\mu}_i(t) \log \left(\frac{\hat{\mu}_i(t)}{\mu} \right) \leq \frac{3 \log(t)}{N_i(t)} \right\}.$$

2. Select any node $A_{t+1} \in \arg \max_i U_i(t)$.
3. Observe degree $X_{t+1, A_{t+1}}$ of node A_{t+1} in G_{t+1} and update

$$\hat{\mu}_{A_{t+1}}(t+1) = \frac{N_{A_{t+1}}(t) \hat{\mu}_{A_{t+1}}(t) + X_{t+1, A_{t+1}}}{N_{A_{t+1}}(t) + 1}.$$

Update $N_{A_{t+1}}(t+1) = N_{A_{t+1}}(t) + 1$.

The learner uses the observed degrees as rewards, and feeds them to an instance of kl-UCB originally designed for Poisson-distributed rewards. A key technical challenge arising in the analysis is that the degree distributions do not actually belong to the Poisson family for finite n . We overcome this difficulty by showing that the degree distributions have a moment generating function bounded by those of Poisson distributions, and that this fact is sufficient for most of the kl-UCB analysis to carry through without changes.

As in the case of the inhomogeneous Erdős-Rényi model, we subsample a set of size (4) of representative nodes for kl-UCB to play on. For clarity of presentation, we first propose a simple algorithm that assumes prior knowledge of T , and then move on to construct a more involved variant that adds new actions on the fly.

We first present our kl-UCB variant for a fixed set of nodes V_0 as Algorithm 4. We refer to this algorithm as d -UCB(V_0) (short for “degree-UCB on V_0 ”). Our two algorithms mentioned above use d -UCB(V_0) as a subroutine: they are both based on uniformly sampling a large enough set V_0 of nodes so that the subsample includes at least one node from the top α -quantile, with high probability.

We define the α -optimal degree $\mu_\alpha^* = \min_{i \in V_\alpha^*} \mu_i$ and the gap parameter $\delta_{\alpha,i} = (\mu_i - \mu_\alpha^*)_+$. We first present a performance guarantee of our simpler algorithm that assumes knowledge of T , so the learner plays d -UCB(V_0) on the uniformly sampled a subset of size (4). This algorithm satisfies the following performance guarantee:

Theorem 5 *Assume that the underlying random graph is either (a) a subcritical stochastic block model satisfying Assumption 1; (b) a supercritical stochastic block model satisfying Assumptions 1 and 2; (c) a subcritical Chung-Lu random graph; or (d) a supercritical Chung-Lu random graph.*

Let V_0 be a uniform subsample of V with size given in Equation (4) and define the event $\mathcal{E} = \{V_0 \cap V_\alpha^ \neq \emptyset\}$. If the number of vertices n is sufficiently large, then the expected α -quantile regret of d -UCB(V_0) simultaneously satisfies*

Algorithm 5 d -UCB-DOUBLE(β)

Parameters: $\beta \geq 2$.

Initialization: $V_0 = \emptyset$.

For $k = 1, 2 \dots$, **repeat**

1. Sample subset of nodes U_k uniformly such that $|U_k| = \left\lceil \frac{\log \beta}{\log(1/(1-\alpha))} \right\rceil$.
 2. Update action set $V_k = V_{k-1} \cup U_k$.
 3. For rounds $t = \beta^{k-1}, \beta^{k-1} + 1, \dots, \beta^k - 1$, run a new instance of d -UCB (V_k).
-

$$R_T^\alpha \leq \mathbb{E} \left[\sum_{i \in V_0} \Delta_{\alpha,i} \left(\frac{\mu_\alpha^* (18 + 27 \log T)}{\delta_{\alpha,i}^2} + 3 \right) \middle| \mathcal{E} \right] + \Delta_{\alpha,\max},$$

where the expectation is taken over the random choice of V_0 , and

$$R_T^\alpha \leq 18c^* \sqrt{\frac{T\mu^* (2 + 3 \log T)^2}{\log(1/(1-\alpha))}} + \left(\frac{3 \log T}{\log(1/(1-\alpha))} + 4 \right) \Delta_{\alpha,\max}.$$

In contrast to the results obtained in the general setting of Section 2, where we have to run different algorithms in the subcritical and supercritical cases, for the models considered in this section the learner can run the Algorithm 5 without prior knowledge of the regime.

For unknown values of T , we propose the d -UCB-DOUBLE(β) algorithm (presented as Algorithm 5) that uses a doubling trick to estimate T . The following theorem gives a performance guarantee for this algorithm:

Theorem 6 *Assume that the underlying random graph is either (a) a subcritical stochastic block model satisfying Assumption 1; (b) a supercritical stochastic block model satisfying Assumptions 1 and 2; (c) a subcritical Chung-Lu random graph; or (d) a supercritical Chung-Lu random graph.*

Fix T , let k_{\max} be the value of k on which d -UCB-DOUBLE(β) terminates, and define the event $\mathcal{E} = \{V_{k_{\max}} \cap V_\alpha^ = \emptyset\}$. If the number of vertices n is sufficiently large, then the α -quantile regret of d -UCB-DOUBLE(β) simultaneously satisfies*

$$R_T^\alpha \leq \mathbb{E} \left[\sum_{i \in V_{k_{\max}}} \Delta_i \left(\left(\frac{18\mu^*}{\delta_{\alpha,i}^2} + 3 \right) (\log_\beta T + 1) + \frac{27 \log \beta (\log_\beta T + 1)^2}{2\delta_{\alpha,i}^2} \right) \middle| \mathcal{E} \right] + \Delta_{\alpha,\max} \log_\beta T,$$

where the expectation is taken over the random choice of the sets V_1, V_2, \dots , and

$$R_T^\alpha \leq 36c^* \sqrt{\frac{T(\mu^* + \log(\beta T)) \log^2 T}{\log(1/(1-\alpha))}} + \left(\frac{3 \log^2 T}{\log(1/(1-\alpha))} + 4 \right) \Delta_{\alpha,\max}.$$

3.3. Learning with degree feedback in Kronecker random graphs

In this section we study influence maximization when the underlying random network is a Kronecker random graph. We set this model apart as the properties of Kronecker random graphs differ significantly from those of the stochastic block model and the Chung-Lu model. At the same time, we show that observing the degree of the selected nodes is enough to maximize the total influence in this graph model as well. In particular, the same algorithm d -UCB(V_0) introduced above achieves a small regret.

We consider only supercritical regime with parameters are such that $(\zeta + \beta)(\beta + \gamma) > 1$. The reason is that subcritical Kronecker random graphs contain only $o(n)$ non-isolated vertices with high probability, and therefore minimizing the α -quantile regret is essentially trivial in this case.

Denote by H the subgraph of $G_{n,P^{[k]}}$, induced by the vertices of weight $l \geq k/2$. We exploit the property that for the graph $G_{n,P^{[k]}}$ with parameters $(\zeta + \beta)(\beta + \gamma) > 1$, there exists a constant $b(P)$ such that a subgraph of $G_{n,P^{[k]}}$, induced by the vertices of H , is connected with probability at least $1 - n^{-b(P)}$, see [Frieze and Karonski \(2015, Theorem 9.10\)](#). This means that on this event, the connected components C_i are the same for all $i \in H$. This allows us to prove the following:

Theorem 7 *Let V_0 be a uniform subsample of V of size $\left\lceil \frac{\log(nT)}{\log(2)} \right\rceil$. Let $G_{n,P^{[k]}}$ be such that $(\zeta + \beta)(\beta + \gamma) > 1$ and $\zeta > \gamma > \beta$. Then there exists a constant $b(P)$ such that the quantile regret of d -UCB(V_0) satisfies*

$$\frac{R_T^\alpha}{n} \leq \left\lceil \frac{\log(nT)}{\log(2)} \right\rceil \left(\frac{\mu^*(2 + 6 \log T)}{\left(1 - \frac{\beta + \gamma}{\zeta + \beta}\right)^2} + 3 + n^{-b(P)} \frac{\mu^*(2 + 6 \log T)(\zeta + \beta)^2}{(\zeta - \gamma)^2} + 3n^{-b(P)} \right) + 1 .$$

4. Discussion

In this section we highlight some features of our results and discuss directions for future work. Our main results show that online influence maximization is possible with only local feedback information. We establish bounds for the quantile regret that are polylogarithmic in T for all considered random graph models. Notably, our bounds hold for both the subcritical and supercritical regimes of the random-graph models considered, and show no explicit dependence on the number of nodes n .

Previous work. Related online influence maximization algorithms consider more general classes of networks, but make more restrictive assumptions about the interplay between rewards and feedback. The line of work explored by [Wen, Kveton, Valko, and Vaswani \(2017\)](#); [Wang and Chen \(2017\)](#) assumes that the algorithm receives *full feedback* on where the information reached in the previous trials (i.e., not only the number of influenced nodes, but their exact identities and influence paths, too). Clearly, such detailed measurements are nearly impossible to obtain in practice, as opposed to the local observations considered in this paper.

Another related setup was considered by [Carpentier and Valko \(2016\)](#), whose algorithm only receives feedback about the nodes that were directly influenced by the chosen node, but the model does not assume that neighbors in the graph share the information to further neighbors and counts the reward only by the nodes directly connected to the selected one. That is, in contrast to our work, this work does not attempt to show any relation between local and global influence maximization. One

downside to all the above works is that they all provide rather conservative performance guarantees: On the one hand, [Wen, Kveton, Valko, and Vaswani \(2017\)](#) and [Carpentier and Valko \(2016\)](#) are concerned with worst-case regret bounds that uniformly hold for all problem instances for a fixed time horizon T . On the other hand, the bounds of [Wang and Chen \(2017\)](#) depend on topological (rather than probabilistic) characteristics of the underlying graph structure, which inevitably leads to conservative results. For example, their bounds instantiated in our graph model lead to a regret bound of order $n^3 \log T$, which is virtually void of meaning in our regime of interest where n is very large (e.g. much larger than T). In contrast, our bounds do not show any explicit dependence on n . In this light, our work can be seen as the first attempt that takes advantage of specific probabilistic characteristics of the mechanism of information spreading to obtain strong instance-dependent global performance guarantees, all while having access to only local observations.

Other related framework is stochastic online learning under partial monitoring [Agarwal, Bartlett, and Dama \(2010\)](#); [Bartók, Zolghadr, and Szepesvári \(2012\)](#); [Komiyama, Honda, and Nakagawa \(2015\)](#). In this setting the loss is not directly observed by the learner, which makes this setting applicable to a wider range of problems. However, the partial monitoring setup is too general to capture the specific relationship of feedback and influenced component size, resulting to regret bounds that scale with n^2 .

Tightness of the regret bounds. In terms of dependence on T , both our instance-dependent and worst-case bounds are near-optimal in their respective settings: even in the simpler stochastic multi-armed bandit problem, the best possible regret bounds are $\Omega_T(\log T)$ and $\Omega_T(\sqrt{T})$ in the respective settings [Auer, Cesa-Bianchi, and Fischer \(2002a\)](#); [Auer, Cesa-Bianchi, Freund, and Schapire \(2002b\)](#); [Bubeck and Cesa-Bianchi \(2012\)](#). The optimality of our bounds with respect to other parameters such as c^* , μ^* and n is less clear, but we believe that these factors cannot be improved substantially for the models that we studied in this paper. As for the subproblem of identifying nodes with the highest degrees, we believe that our bounds on the number of suboptimal draws is essentially tight, closely matching the classical lower bounds by [Lai and Robbins \(1985\)](#).

Our assumptions. One may wonder how far our argument connecting local and global influence maximization can be stretched. Clearly, not every random graph model enables establishing such a strong connection.

5. Multi-type branching processes

One of the most important technical tools for analyzing the component structure of random graphs is the theory of *branching processes*, see [Bollobás, Janson, and Riordan \(2007\)](#); [van der Hofstad \(2016\)](#). Indeed, while the connected components of an inhomogenous random graph $G(n, \kappa)$ have a complicated structure, many of their key properties may be analyzed through the concept of multi-type Galton–Watson processes.

Recall the notation introduced in Section 1.2. Consider a Galton–Watson process, where an individual $x \in (0, 1]$ is replaced in the next generation by a set of particles distributed as a Poisson process on $(0, 1]$ with intensity $\kappa(x, y)d\mu(y)$ and the number of children has a Poisson distribution with mean $\int_{(0,1]} \kappa(x, y)d\mu(y)$. We denote this branching process, started with a single particle x by $W_\kappa(x)$.

[Bollobás, Janson, and Riordan \(2007\)](#) establishes a connection between the sizes of connected components of $G(n, \kappa)$, the survival probability of a branching process $W_\kappa(x)$, and the function

κ . As shown in Bollobás, Janson, and Riordan (2007), the operator Φ_κ can be directly used for characterizing the probability $\rho(x)$ of survival of the process $W_\kappa(x)$ for all $x \in (0, 1]$. By their Theorem 6.2, the function ρ is the maximum fixed point of the non-linear equation $\Phi_\kappa(f) = f$. Furthermore, as was shown in Bollobás, Janson, and Riordan (2007, Lemma 5.8.), if $\|T_\kappa\|_2 < 1$, then $\rho(x) = 0$ for all x and when $\|T_\kappa\|_2 > 1$, $\rho(x) > 0$ for all x .

To analyze the random graph $G(n, \kappa)$, we use Poisson multi-type Galton–Watson branching processes with n types, parametrized by an $n \times n$ matrix A with positive elements. Therefore, each node corresponds to its own type.

The branching process tracks the evolution of a set of *individuals* of various types. Starting in round $n = 0$ from a single individual of type i , each further generation in the Galton–Watson process $W_\kappa(i)$ is generated by each individual of each type i producing $X_{i,j} \sim \text{Poisson}(A_{i,j}/n)$ new individuals of each type j . Therefore, the number of offsprings of the individual of type i is $\sum_{j=1}^n X_{i,j} \sim \text{Poisson}(\sum_{j=1}^n A_{i,j}/n)$.

Our analysis below makes use of the following quantities associated with the multi-type branching process:

1. $Z_n(i)$ is the number of individuals in generation n of $W_\kappa(i)$ (where $Z_0(i) = 1$);
2. $B(i)$ is the *total progeny*, that is, the total number of individuals generated by $W_\kappa(i)$ and its expectation is denoted by $x_i = \mathbb{E}[B(i)]$;
3. $\rho(i)$ is the *probability of survival*, that is, the probability that $B(i)$ is infinite.

6. Proofs of Theorem 2 and 3.

The connected components C_i of an individual i have a complicated structure, but many key properties can be analyzed through the concept of multi-type Galton–Watson branching processes with n types. Fix an arbitrary node i and let $Y_{i,1}, Y_{i,2}, \dots, Y_{i,n}$ be independent Bernoulli random variables with respective parameters $A_{i,j}/n$ for $i, j \in [n]$. Consider a multitype binomial branching process where an individual of type i produces an individual j with probability $A_{i,j}/n$, and let $B_{Ber}(i)$ denote its total progeny when started from an individual i . In the same way, consider a multitype Poisson branching process where an individual of type i produces $X_{i,j} \sim \text{Poisson}(A_{i,j}/n)$ individuals, and let $B(i)$ denote its total progeny when started from an individual i .

We use the concept of *stochastic dominance* between random variables. The random variable X is *stochastically dominated* by the random variable Y when, for every $x \in \mathbb{R}$, $\mathbb{P}[X \leq x] \geq \mathbb{P}[Y \leq x]$. We denote this by $X \preceq Y$.

Proof of Lemma 1. First, we define an upper approximation to κ . We choose an integer m and we partition the interval $(0, 1]$ into m sets $\mathcal{A}_1, \dots, \mathcal{A}_m$, where $\mathcal{A}_k = ((k-1)/m, k/m]$, $k \in [1, m]$. Also we denote by $\mathcal{A}_m(x)$ the set \mathcal{A}_k for which $x \in \mathcal{A}_k$. Then we bound κ from above by

$$\kappa_m^+(x, y) = \sup\{\kappa(x', y') : x' \in \mathcal{A}_m(x), y' \in \mathcal{A}_m(y)\}.$$

As κ is bounded, there exists a sufficiently large m such that $\|T_{\kappa_m^+}\| < 1$:

$$\|T_{\kappa_m^+}\| \leq \|T_\kappa\| + \|T_{\kappa_m^+} - T_\kappa\| \leq \|T_\kappa\| + \left(\int_{(0,1] \times (0,1]} (\kappa_m^+(x, y) - \kappa(x, y))^2 dx dy \right)^{1/2}.$$

Then for any node i in $G(n, \kappa)$, we define a type $k_i = k$ if $(k-1)/m < i/n \leq k/m$ holds. Then, by our definition of κ_m^+ , we have

$$\mathbb{P}[|C_i| > u] \leq \mathbb{P}[B_{Ber}(k_i) > u].$$

For $k, \ell \in [m]$ we define $p_{k,\ell} = \frac{1}{m} \kappa_m^+(k/m, \ell/m)$. Notice, that for random variables $Y \sim Ber(p)$ and $X \sim Poisson(p')$ with $p' = -\log(1-p) > p$, $Y \preceq X$ holds. This follows from the observation that $\mathbb{P}[Y > 0] = p$ and $\mathbb{P}[X > 0] = p$. It follows that $Ber(p_{k,\ell}) \preceq Poisson((1+\varepsilon)p_{k,\ell})$. Then there exists $\varepsilon > 0$ such that the multitype Poisson branching process $\tilde{B}(k)$ with parameters $(1+\varepsilon)p_{k,\ell}$ is such that $\mathbb{P}[B_{Ber}(k) > u] < \mathbb{P}[\tilde{B}(k) > u]$ and it is subcritical. We also define a random variable $\tilde{X}_{k,\ell} \sim Poisson((1+\varepsilon)p_{k,\ell})$.

Since the total number of descendants of individuals in the first generation are independent, we can write the following recursive equation on the number of descendants of type k :

$$|\tilde{B}(k)| = 1 + \sum_{\ell=1}^m \tilde{X}_{k,\ell} |\tilde{B}(\ell)|.$$

For any type k , for $z_k > 1$, the probability generating function of $|\tilde{B}(k)|$ is $g(k) = \mathbb{E}\left[z_k^{|\tilde{B}(k)|}\right]$ and we denote $g = (g(1), \dots, g(m))^T$. Using that for $X \sim Poisson(\gamma)$ for some $\gamma > 0$, $y > 1$ the probability generating function is $\mathbb{E}[y^X] = e^{\gamma(y-1)}$, we have

$$\begin{aligned} g(k) &= \mathbb{E}\left[z_k^{|\tilde{B}(k)|}\right] = z_k \mathbb{E}\left[z_k^{\tilde{X}_{k,1}|\tilde{B}(1)|} \dots z_k^{\tilde{X}_{k,m}|\tilde{B}(m)|}\right] = z_k \prod_{\ell} \mathbb{E}\left[z_k^{\tilde{X}_{k,\ell}|\tilde{B}(\ell)|}\right] \\ &= z_k \prod_{\ell} \mathbb{E}\left[\left(\mathbb{E}\left[z_k^{|\tilde{B}(\ell)|}\right]\right)^{\tilde{X}_{k,\ell}}\right] = z_k \exp\left((1+\varepsilon) \sum_{\ell} p_{k,\ell} (g(\ell) - 1)\right). \end{aligned}$$

Recall that P denotes the $m \times m$ matrix with entries $p_{k,\ell}$. Our next aim is to study the fixed point of the operator G_P , defined as

$$g = G_P g := z \exp\left((1+\varepsilon)P(g - \bar{1})\right). \quad (5)$$

Define the function $F(z, g) = z \exp\left((1+\varepsilon)P(g - \bar{1})\right) - g$. This function is smooth and the entries of the Jacobian matrix are

$$J_{k,\ell}(z, g) := \frac{\partial F_k}{\partial g_\ell} = z_k(1+\varepsilon)p_{k,\ell} \exp\left((1+\varepsilon) \sum_{\ell} p_{k,\ell} (g_\ell - 1)\right) - \mathbb{I}_{\{k=\ell\}}.$$

Let $P'(g, z)$ be the matrix with elements $z_k(1+\varepsilon)p_{k,\ell} \exp((1+\varepsilon) \sum_{\ell} p_{k,\ell} (g_\ell - 1))$. Then, at point $(\bar{1}, \bar{1})$, $P'_{k,\ell}(\bar{1}, \bar{1}) = (1+\varepsilon)p_{k,\ell}$. Since ε is chosen such that the branching process $\tilde{B}(k)$ is subcritical, $P'(\bar{1}, \bar{1})$ is smaller than one. This means, that we can find $z' = 1 + \delta$, $g' > 0$, such that the largest eigenvalue of $P'(g', z')$ is smaller than one as well, and therefore $J(z', g')$ is invertible. Then, by the implicit function theorem there exists an open set $U_z \subset (1, +\infty)^m$ and a function $q : U_z \rightarrow (0, +\infty)^m$ such that $F(z, q(z)) = \bar{0}$.

Finally, the statement of the lemma is obtained by applying the Chernoff bound:

$$\mathbb{P} \left[\tilde{B}(k) > u \right] = \mathbb{P} \left[z_k^{\tilde{B}(k)} > z_k^u \right] \leq \frac{\mathbb{E} \left[z_k^{\tilde{B}(k)} \right]}{z_k^u}.$$

Denote $\lambda_k = \ln(z_k) > 0$. Then,

$$\frac{\mathbb{E} \left[z_k^{\tilde{B}(k)} \right]}{z_k^u} = \frac{g_k}{z_k^u} = \exp(-\lambda_k u) g_k.$$

Then taking any $\lambda(\kappa) = \min_k \lambda_k$, $g(\kappa) = \max_k g_k$, we get the statement of the lemma. \square

Armed with this concentration result, we can see that the typical size $|C_i|$ of the connected component of any vertex i is $O(1)$. Recall that the learning algorithm has only access to a censored value of $|C_i|$, truncated by a constant K . Our main technical result shows that nodes with the largest expected censored observations $u_*(K)$ are exactly the ones with the largest influence c_* . We formally state this result next:

Lemma 8 *For $G(n, \kappa)$ with subcritical κ , and $n > n_0(\kappa)$, for any node i we have $c_* - c_i \leq u_*(K) - u_i(K) + e^{-\lambda(\kappa)K} g$. Then, for $K = \frac{\log T}{\lambda}$, with $\lambda < \lambda(\kappa)$ we have $c_* - c_i \leq u_*(K) - u_i(K) + \frac{g(\kappa)}{T}$.*

Proof

The expected bias of $u_i(K)$ is, using the result of Lemma 1:

$$\begin{aligned} c_i - u_i(K) &= \mathbb{E} [|C_i(G_t)| - u_i(K)] = \mathbb{E} [(|C_i| - K)_+] \\ &\leq \int_0^\infty \mathbb{P} [|C_i| - K > u] du \leq \int_0^\infty e^{-\lambda(u+K)} du \leq e^{-\lambda K} g(\kappa). \end{aligned}$$

Set $K = \frac{\log T}{\lambda}$. Then,

$$c_* - c_i \leq u_*(\log T/\lambda) - u_i(\log T/\lambda) + \frac{g(\kappa)}{T}.$$

■

Proof of Theorem 2. In order not to overload notation we write $\delta_{\alpha,i}^{sub}$ for $\delta_{\alpha,i}^{sub}(K)$. We first note that, with high probability, the size of V_0 guarantees that the subset contains at least one node from the set V_α^* : $\mathbb{P}[\mathcal{E}] \geq 1 - 1/T$. Then, the regret can be bounded as

$$R_T^\alpha \leq \mathbb{P}[\mathcal{E}^c] T \Delta_{\alpha, \max} + \mathbb{E} \left[\sum_{t=1}^T \sum_{i \in V_0} \mathbb{I}[A_t = i] \Delta_{\alpha, i} \middle| \mathcal{E} \right] \mathbb{P}[\mathcal{E}] \quad (6)$$

$$\leq \Delta_{\alpha, \max} + \mathbb{E} \left[\sum_{i \in V_0} \Delta_{\alpha, i} \mathbb{E}[N_{i, T}] \middle| \mathcal{E} \right]. \quad (7)$$

By Hoeffding's inequality,

$$\mathbb{P} \left[A_{t+1} = i \mid N_{i,t} \geq \frac{4K^2 \log t}{(\delta_{\alpha,i}^{sub})^2} \right] \leq \frac{4}{t^2}.$$

Then,

$$\begin{aligned} \mathbb{E}[N_{i,T}] &\leq \frac{4K^2 \log T}{(\delta_{\alpha,i}^{sub})^2} + \sum_{t=|V_0|}^T \mathbb{P} \left[A_{t+1} = i \mid N_{i,t} \geq \frac{4K^2 \log t}{(\delta_{\alpha,i}^{sub})^2} \right] \\ &\leq \frac{4K^2 \log T}{(\delta_{\alpha,i}^{sub})^2} + \sum_{t=|V_0|}^T \frac{4}{t^2} \leq \frac{4K^2 \log T}{(\delta_{\alpha,i}^{sub})^2} + 8. \end{aligned}$$

Now, observing that $\delta_{\alpha,i}^{sub} \leq \max_{j \in V_0} u_j(K) - u_i(K)$ holds under event \mathcal{E} , we obtain

$$R_T^\alpha \leq \Delta_{\alpha,\max} + \mathbb{E} \left[\sum_{i \in V_0} \Delta_{\alpha,i} \left(\frac{4K^2 \log T}{(\delta_{\alpha,i}^{sub})^2} + 8 \right) \middle| \mathcal{E} \right], \quad (8)$$

thus proving the first statement.

Next, we turn to proving the second statement regarding worst-case guarantees. To do this, we appeal to Proposition 8 and take $K = \frac{\log T}{\lambda}$, where λ is any number, satisfying conditions of Lemma 1. To proceed, let us fix an arbitrary $\varepsilon > 0$ and split the set V_0 into two subsets: $U(\varepsilon) = \{a \in V_0 : \delta_{\alpha,i}^{sub} \leq \varepsilon\}$ and $W(\varepsilon) = V_0 \setminus U(\varepsilon)$. Then, under event \mathcal{E} , we have

$$\begin{aligned} \sum_{i \in V_0} \Delta_{\alpha,i} \mathbb{E}[N_{i,T}] &= \sum_{i \in U(\varepsilon)} \Delta_{\alpha,i} \mathbb{E}[N_{i,T}] + \sum_{i \in W(\varepsilon)} \Delta_{\alpha,i} \mathbb{E}[N_{i,T}] \\ &\leq \varepsilon \sum_{i \in U(\varepsilon)} \mathbb{E}[N_{i,T}] + \frac{g}{T} \sum_{i \in U(\varepsilon)} \mathbb{E}[N_{i,T}] \\ &\quad + \sum_{i \in W(\varepsilon)} \delta_{\alpha,i}^{sub} \left(\frac{4 \left(\frac{\log T}{\lambda} \right)^2 \log T}{(\delta_{\alpha,i}^{sub})^2} \right) + \frac{g}{T} \sum_{i \in W(\varepsilon)} \frac{4 \left(\frac{\log T}{\lambda} \right)^2 \log T}{(\delta_{\alpha,i}^{sub})^2} \\ &\quad + 8|W(\varepsilon)|\Delta_{\alpha,\max} \\ &\leq \varepsilon T + g + \sum_{i \in W(\varepsilon)} \left(\frac{4 \left(\frac{\log T}{\lambda} \right)^2 \log T}{\delta_{\alpha,i}^{sub}} \right) + \frac{g}{T} \sum_{i \in W(\varepsilon)} \frac{4 \left(\frac{\log T}{\lambda} \right)^2 \log T}{(\delta_{\alpha,i}^{sub})^2} \\ &\quad + 8|W(\varepsilon)|\Delta_{\alpha,\max} \\ &\leq \varepsilon T + g + |V_0| \frac{4 \left(\frac{\log T}{\lambda} \right)^2 \log T}{\varepsilon} + \frac{g}{T} |V_0| \frac{4 \left(\frac{\log T}{\lambda} \right)^2 \log T}{\varepsilon^2} + 8|V_0|\Delta_{\alpha,\max} \\ &\leq 4 \left(\frac{\log T}{\lambda} \right) \sqrt{|V_0| T \log T} + 2g + 8|V_0|\Delta_{\alpha,\max}. \end{aligned}$$

where the last step uses the choice $\varepsilon = 2 \left(\frac{\log T}{\lambda} \right) \sqrt{|V_0| \log T / T}$. Plugging in the choice of $|V_0|$ concludes the proof. \square

Proof of Theorem 3. To simplify the notation, we use λ instead of $\lambda(\kappa)$. Let T_q be the length of the q -th iterate. The expected regret over each period q can be bounded as an expected regret of Local UCB(V_0) with parameters $\lambda_q = 2^{-q}$ and T_q time steps. Appealing to Theorem 2, we can bound the expected regret as

$$\begin{aligned} R_T^\alpha &\leq \mathbb{P}[\mathcal{E}^c] T \Delta_{\alpha, \max} + \mathbb{E} \left[\sum_{t=1}^T \sum_{i \in V_0} \mathbb{I}_{\{A_t=i\}} \Delta_{\alpha, i} \middle| \mathcal{E} \right] \\ &\leq \Delta_{\alpha, \max} + \mathbb{E} \left[\sum_{q=1}^{Q_{\max}} \sum_{i \in V_0} \Delta_{\alpha, i} \mathbb{E}[N_{i, T_q}] \middle| \mathcal{E} \right] \end{aligned}$$

Following the analysis of Theorem 2, by (8), we get

$$R_T^\alpha \leq \Delta_{\alpha, \max} + \mathbb{E} \left[\sum_{q=1}^{Q_{\max}} \sum_{i \in V_0} \Delta_{\alpha, i} \left(\frac{4K_q^2 \log T}{(\delta_{\alpha, i}^{sub}(K_q))^2} + 8 \right) \middle| \mathcal{E} \right].$$

We have $Q_{\max} = \lceil \log_2(1/\lambda) \rceil \leq \log_2(1/\lambda) + 1$, and

$$\sum_{q=0}^{Q_{\max}} K_q^2 = \log^2 T \sum_{q=0}^{Q_{\max}} 4^q = \log^2 T \frac{4^{Q_{\max}+1} - 1}{3} \leq \frac{16}{3} \frac{1}{\lambda^2} \log^2 T.$$

This gives us

$$R_T^\alpha \leq \Delta_{\alpha, \max} + \frac{64}{3} \mathbb{E} \left[\sum_{i \in V_0} \Delta_{\alpha, i} \left(\frac{\log^3 T}{(\lambda \cdot \min_{q \in [Q_{\max}]} \{\delta_{\alpha, i}^{sub}(K_q)\})^2} + 8 \right) \middle| \mathcal{E} \right].$$

Next, we prove the second statement regarding worst-case guarantees. To proceed, let us take $\varepsilon_q = \frac{\log T}{\lambda} \sqrt{|V_0| \log T / T_q}$ and split the set V_0 into two subsets: $U(\varepsilon_q) = \{a \in V_0 : \delta_{\alpha, i}^{sub}(K_q) \leq \varepsilon_q\}$ and $W(\varepsilon_q) = V_0 \setminus U(\varepsilon_q)$. Then, under event \mathcal{E} , we have

$$\begin{aligned} &\mathbb{E} \left[\sum_{t=1}^T \sum_{i \in V_0} \mathbb{I}_{\{A_t=i\}} \Delta_{\alpha, i} \right] \\ &\leq \mathbb{E} \left[\sum_{q=1}^{Q_{\max}} \sum_{t_q=1}^{T_q} \sum_{i \in V_0} \Delta_{\alpha, i} \mathbb{I}_{\{|C_{A_{t_q}}| \leq K_q\}} \right] + \mathbb{E} \left[\sum_{q=1}^{Q_{\max}} \sum_{t_q=1}^{T_q} \sum_{i \in V_0} \Delta_{\alpha, i} \mathbb{I}_{\{|C_{A_{t_q}}| > K_q\}} \right] \\ &= \underbrace{\mathbb{E} \left[\sum_{q=1}^{Q_{\max}} \sum_{t_q=1}^{T_q} \sum_{i \in U(\varepsilon_q)} \Delta_{\alpha, i} \mathbb{I}_{\{|C_{A_{t_q}}| \leq K_q\}} \right]}_{\text{Term 1}} + \underbrace{\mathbb{E} \left[\sum_{q=1}^{Q_{\max}} \sum_{t_q=1}^{T_q} \sum_{i \in U(\varepsilon_q)} \Delta_{\alpha, i} \mathbb{I}_{\{|C_{A_{t_q}}| > K_q\}} \right]}_{\text{Term 2}} \\ &\quad + \underbrace{\mathbb{E} \left[\sum_{q=1}^{Q_{\max}} \sum_{t_q=1}^{T_q} \sum_{i \in W(\varepsilon_q)} \Delta_{\alpha, i} \mathbb{I}_{\{|C_{A_{t_q}}| \leq K_q\}} \right]}_{\text{Term 3}} + \underbrace{\mathbb{E} \left[\sum_{q=1}^{Q_{\max}} \sum_{t_q=1}^{T_q} \sum_{i \in W(\varepsilon_q)} \Delta_{\alpha, i} \mathbb{I}_{\{|C_{A_{t_q}}| > K_q\}} \right]}_{\text{Term 4}}. \end{aligned}$$

Term 1:

$$\begin{aligned} \mathbb{E} \left[\sum_{q=1}^{Q_{\max}} \sum_{t_q=1}^{T_q} \sum_{i \in U(\varepsilon_q)} \Delta_{\alpha, i} \mathbb{1}_{\{|C_{A_{t_q}}| \leq K_q\}} \right] &\leq |V_0| \frac{\log T}{\lambda} \mathbb{E} \left[\sum_{q=1}^{Q_{\max}} \sqrt{|V_0| T_q \log T} \right] \\ &\leq |V_0|^{3/2} \frac{\log T}{\lambda} \sqrt{(\log_2(1/\lambda) + 1)T}. \end{aligned}$$

Term 2: The expected bias of $\mu_{i,t}^{sub}(K_q)$ is, using the result of Lemma 1:

$$\begin{aligned} \mathbb{E} [(|C_i| - K_q)_+] &\leq \int_0^\infty \mathbb{P} [|C_i| - K_q > u] du \leq \int_0^\infty e^{-\lambda(u+K_q)} du \leq e^{-\lambda K_q} g \\ &= \left(\frac{1}{T}\right)^{\lambda 2^q} g \leq \left(\frac{1}{T}\right)^{2^q - Q_{\max}} g. \end{aligned}$$

Then,

$$c_* - c_i \leq \mu_*^{sub}(K_q) - \mu_i^{sub}(K_q) + \left(\frac{1}{T}\right)^{2^q - Q_{\max}} g. \quad (9)$$

According to the stopping rule, we get

$$\begin{aligned} \mathbb{E} \left[\sum_{q=1}^{Q_{\max}} \sum_{t_q=1}^{T_q} \sum_{i \in U(\varepsilon_q)} \Delta_{\alpha, i} \mathbb{1}_{\{|C_{A_{t_q}}| > K_q\}} \right] &\leq |V_0| \mathbb{E} \left[\sum_{q=1}^{Q_{\max}} \left(\varepsilon_q + g \left(\frac{1}{T}\right)^{2^q - Q_{\max}} \right) \left(\frac{1}{T} + \sqrt{\frac{\log T}{2T_q}} \right) T_q \right] \\ &\leq |V_0| \mathbb{E} \left[\sum_{q=1}^{Q_{\max}} \left(\frac{\log T}{\lambda} \sqrt{|V_0| \log T / T_q} + g \right) \left(\frac{1}{T} + \sqrt{\frac{\log T}{2T_q}} \right) T_q \right] \\ &\leq |V_0| \left(\frac{\log T}{\lambda} \sqrt{\frac{(\log_2(1/\lambda) + 1)|V_0| \log T}{T}} + \frac{g}{T} \right) + |V_0|^{3/2} (\log_2(1/\lambda) + 1) \frac{\log^2 T}{\lambda} \\ &\quad + |V_0| g \left(\sqrt{(\log_2(1/\lambda) + 1)T \log T} \frac{\log T}{\lambda} \right). \end{aligned}$$

Term 3: Following the analysis of Theorem 2 and by (9), we get

$$\begin{aligned} &\mathbb{E} \left[\sum_{q=1}^{Q_{\max}} \sum_{t_q=1}^{T_q} \sum_{i \in W(\varepsilon_q)} \Delta_{\alpha, i} \mathbb{1}_{\{A_t=i, |C_{A_t}| \leq K_q\}} \right] \\ &\leq \mathbb{E} \left[\sum_{i \in W(\varepsilon_q)} \sum_{q=0}^{Q_{\max}} \delta_{\alpha, i}^{sub}(K_q) \left(\frac{4 \left(\frac{\log T_q}{\lambda}\right)^2 \log T}{(\delta_{\alpha, i}^{sub}(K_q))^2} \right) \right] + 8|V_0| \Delta_{\alpha, \max} \\ &\leq 4\sqrt{|V_0|} \mathbb{E} \left[\sum_{q=0}^{Q_{\max}} \left(\sqrt{T_q} \frac{\log^{3/2} T}{\lambda} \right) \right] + 8|V_0| \Delta_{\alpha, \max} \\ &\leq 4\sqrt{|V_0|} \left(\sqrt{(\log(1/\lambda) + 1)T} \frac{\log^{3/2} T}{\lambda} \right) + 8|V_0| \Delta_{\alpha, \max}. \end{aligned}$$

Term 4:

$$\begin{aligned}
& \mathbb{E} \left[\sum_{q=1}^{Q_{\max}} \sum_{t_q=1}^{T_q} \sum_{i \in W(\varepsilon_q)} \Delta_{\alpha, i} \mathbb{1}_{\{A_t=i, |C_{A_t}| > K_q\}} \right] \\
& \leq \mathbb{E} \left[\sum_{i \in W(\varepsilon_q)} \sum_{q=0}^{Q_{\max}} \left(\delta_{\alpha, i}^{sub}(K_q) + g \left(\frac{1}{T} \right)^{2^q - Q_{\max}} \right) \left(\frac{4 \left(\frac{\log T_q}{\lambda} \right)^2 \log T}{(\delta_{\alpha, i}^{sub}(K_q))^2} \right) \left(\frac{1}{T} + \sqrt{\frac{\log T}{2T_q}} \right) \right] \\
& + 8|W(\varepsilon_q)|\Delta_{\alpha, \max} \\
& \leq 4\sqrt{|V_0|}(\log_2(1/\lambda) + 1) \frac{\log^{3/2} T}{\sqrt{T}\lambda} + 4\sqrt{|V_0|}(\log_2(1/\lambda) + 1) \frac{\log^2 T}{\lambda} \\
& + 4g(\log_2(1/\lambda) + 1) \frac{\log T}{\lambda} + 4g \frac{\log^{3/2} T}{\lambda} \sqrt{(\log_2(1/\lambda) + 1)T} + 8|V_0|\Delta_{\alpha, \max}.
\end{aligned}$$

Putting everything together, we conclude that

$$\begin{aligned}
R_t^\alpha & \leq 4 \frac{1}{\sqrt{\ln(1/(1-\alpha))}} \left(\sqrt{(\log(1/\lambda) + 1)T} \frac{\log^2 T}{\lambda} \right) + 16\Delta_{\alpha, \max} \frac{\log T}{\sqrt{\ln(1/(1-\alpha))}} \\
& + 4g(\log_2(1/\lambda) + 1) \frac{\log T}{\lambda} + 4g \frac{\log^{3/2} T}{\lambda} \sqrt{(\log_2(1/\lambda) + 1)T} \\
& + \log T \cdot \sqrt{\frac{\log_2(1/\lambda + 1)}{T \ln(1/(1-\alpha))}} + \frac{g\sqrt{\log T}}{T\sqrt{\ln(1/(1-\alpha))}} + \frac{2(\log_2(1/\lambda) + 1) \log^{5/2} T}{(\ln(1/(1-\alpha)))^{3/2}} \\
& + g\sqrt{\frac{T(\log_2(1/\lambda) + 1)}{\ln(1/(1-\alpha))}} \log T.
\end{aligned}$$

7. Proof of Theorem 4.

The proof relies on some known properties of the largest connected component in $G(n, \kappa)$ for supercritical κ . We denote the largest and second-largest connected components of G_t by $C_1(G_t)$ and $C_2(G_t)$, respectively. The survival probability of the branching process $W_\kappa(x)$ is denoted as $\rho(x)$. The expected size of the connected component containing vertex i can be estimated in terms of $\rho(i/n)$ and $\mathbb{E}[|C_1|]$ as

$$c_i = \rho(i/n)\mathbb{E}[|C_1|] + o_n(n),$$

see [Bollobás, Janson, and Riordan \(2007, Chapter 9\)](#). The following properties are proved by [Bollobás, Janson, and Riordan \(2007\)](#):

- If $G(n, \kappa)$ is supercritical, then, with high probability, $C_1 = \Theta_n(n)$;
- $C_1(G_n) \rightarrow \sum_{i \in V} \rho(i/n)$ in probability;
- $C_2(G_n) = o_n(n)$ with high probability.

Recall from Section 2 that in the supercritical case the feedback $v_{i,t}(K)$ is the indicator whether $|C_i|$ is larger than K . In the following lemma we show that taking $K = k(n)$ for an arbitrary function of n that diverges to infinity, it is enough to control the bias of the estimate of c_i :

Lemma 9 For any supercritical κ , for any node i satisfying $c_i < c_*$ and for any $K = k(n)$, where $k : \mathbb{N} \rightarrow \mathbb{N}$ is an arbitrary positive function satisfying $\lim_{n \rightarrow \infty} k(n) = \infty$, there exist a positive function $f_\kappa : \mathbb{N} \rightarrow \mathbb{R}$, such that $\lim_{n \rightarrow \infty} f_\kappa(n) = 0$ and

$$\frac{c_* - c_i}{n} \leq (v_*(k(n)) - v_i(k(n))) \frac{\mathbb{E}[C_1]}{n} + f_\kappa(n).$$

Proof

Define a kernel $\bar{\kappa}(x, y) = (1 - \rho(y))\kappa(x, y)$, where ρ is defined in Section 5. By Theorem 6.7 in Bollobás et al. (2007), the branching process W_κ conditional on extinction is subcritical and has the same distribution as the branching process with parameters $W_{\bar{\kappa}}$. Then, by Lemma 1,

$$\mathbb{P}[B(i) > K | B(i) < \infty] \leq e^{-\lambda(\bar{\kappa})k(n)} g(\bar{\kappa}). \quad (10)$$

We relate the size of the connected component to the total progeny of branching process. Following the stochastic dominance $C_i \preceq B(i)$,

$$v_i(k(n)) = \mathbb{P}[|C_i| > k(n)] \leq \rho_i + \mathbb{P}[B(i) > k(n) | B(i) < \infty].$$

This implies, for $n > n_0(\bar{\kappa})$,

$$v_i(k(n))\mathbb{E}[|C_1|] - c_i < \mathbb{P}[B(i) > k(n) | B(i) < \infty] \mathbb{E}[|C_1|] + \rho_i \mathbb{E}[|C_1|] - c_i \leq e^{-\lambda(\bar{\kappa})k(n)} g(\bar{\kappa}) \mathbb{E}[|C_1|] + o_n(n).$$

Finally, using that $\rho_* \leq v_*(k(n))$, we get

$$\frac{c_* - c_i}{n} \leq (v_*(k(n)) - v_i(k(n))) \frac{\mathbb{E}[|C_1|]}{n} + e^{-\lambda k(n)} g(\bar{\kappa}) \frac{\mathbb{E}[|C_1|]}{n} + o_n(1) = \delta_i^{sup}(k(n)) \mathbb{E}[|C_1|] + f_\kappa(n). \quad \blacksquare$$

Proof of Theorem 4. First, by (6),

$$R_T^\alpha \leq \Delta_{\alpha, \max} + \mathbb{E} \left[\sum_{i \in V_0} \Delta_{\alpha, i} \mathbb{E}[N_{i, T}] \middle| \mathcal{E} \right].$$

As we mentioned before, with high probability, $C_2(G_n) = o_n(n)$, which means that if $A_t \notin C_1(G_t)$, then $|C_{A_t}(G_t)| = o_n(n)$. Since $G(n, \kappa)$ is supercritical, $\arg \max_a \mu_a = \arg \max_a \rho_a$. Then, we can approximate distribution of rewards of arm a by a Bernoulli distribution with parameter ρ_a . Using the result of Proposition 9, we reduce the initial problem to the analysis of a multi-armed problem with arms $Z_1, \dots, Z_{|V_0|}$, where $Z_i \sim Ber(u_i)$, for p_i defined in Proposition 9.

By Hoeffding's inequality,

$$\mathbb{P} \left[A_{t+1} = i \mid N_{i, t} \geq \frac{4 \log t}{(\delta_{\alpha, i}^{sup}(K))^2} \right] \leq \frac{4}{t^2}.$$

Then

$$\mathbb{E}[N_{i, T}] \leq \frac{4 \log T}{(\delta_{\alpha, i}^{sup}(k(n)))^2} + \sum_{t=|V_0|}^T \mathbb{P} \left[A_{t+1} = i \mid N_{i, t} \geq \frac{4 \log t}{(\delta_{\alpha, i}^{sup}(k(n)))^2} \right] \leq \frac{4 \log T}{(\delta_{\alpha, i}^{sup}(k(n)))^2} + 8.$$

Now, observing that $\delta_{\alpha,i}^{sup}(k(n)) \leq \max_{j \in V_0} v_j - v_i$ holds under the event \mathcal{E} , we obtain

$$R_T^\alpha \leq \Delta_{\alpha,\max} + \mathbb{E} \left[\sum_{i \in V_0} \Delta_{\alpha,i} \left(\frac{4 \log T}{(\delta_{\alpha,i}^{sup}(k(n)))^2} + 8 \right) \middle| \mathcal{E} \right], \quad (11)$$

thus proving the first statement.

Now we fix an arbitrary $\varepsilon > 0$, we split the set V_0 into two subsets: $U(\varepsilon) = \{a \in V_0 : \delta_{\alpha,i}^{sub}(k(n)) \leq \varepsilon\}$ and $W(\varepsilon) = V_0 \setminus U(\varepsilon)$, where we use the choice $\varepsilon = 2\sqrt{|V_0| \mathbb{E}[C_1] \log T/T}$. Lemma 9 shows that $\frac{c_* - c_i}{n} \leq \frac{v_*(k(n)) - v_i(k(n))}{n} \mathbb{E}[|C_1|] + f_\kappa(n)$. Then there exists $n_0(\kappa)$, such that for any $G(n, \kappa)$ with $n > n_0(\kappa)$, $f_\kappa(n) \leq \varepsilon$ holds. Then, under the event \mathcal{E} , we have

$$\begin{aligned} \frac{1}{n} \sum_{i \in V_0} \Delta_{\alpha,i} \mathbb{E}[N_{i,T}] &= \sum_{i \in U(\varepsilon)} \frac{\Delta_{\alpha,i}}{n} \mathbb{E}[N_{i,T}] + \sum_{i \in W(\varepsilon)} \frac{\Delta_{\alpha,i}}{n} \mathbb{E}[N_{i,T}] \\ &\leq \left(\frac{\varepsilon \mathbb{E}[|C_1|]}{n} + \varepsilon \right) \sum_{i \in U(\varepsilon)} \mathbb{E}[N_{i,T}] + \sum_{i \in W(\varepsilon)} \frac{\Delta_{\alpha,i}}{n} \mathbb{E}[N_{i,T}] \\ &\leq \left(\frac{\varepsilon \mathbb{E}[|C_1|]}{n} + \varepsilon \right) |V_0| T + \sum_{i \in W(\varepsilon)} \delta_{\alpha,i}^{sub}(k(n)) \frac{\mathbb{E}[C_1]}{n} \left(\frac{4 \log T}{(\delta_{\alpha,i}^{sub}(k(n)))^2} \right) \\ &\quad + \varepsilon \sum_{i \in W(\varepsilon)} \left(\frac{4 \log T}{(\delta_{\alpha,i}^{sub}(k(n)))^2} \right) \\ &\leq \left(\frac{\varepsilon \mathbb{E}[|C_1|]}{n} + \varepsilon \right) |V_0| T + |V_0| \left(\frac{\mathbb{E}[|C_1|]}{n} + 1 \right) \frac{8 \log T}{\varepsilon n} \\ &\leq 9 \left(\frac{\mathbb{E}[|C_1|]}{n} + 1 \right) |V_0| \sqrt{T \log T}, \end{aligned}$$

where the last step uses the choice $\varepsilon = \sqrt{\log T/T}$. Plugging in the choice of $|V_0|$ concludes the proof. \square

8. Degree observations.

8.1. Subcritical case

Our main technical result is proving that nodes with the largest expected degrees μ^* are exactly the ones with the largest influence c^* , in both the stochastic block model and the Chung–Lu model, across both the subcritical and supercritical regimes. The following lemma states this result for the subcritical case.

Lemma 10 *Suppose that*

1. G is generated from a subcritical $G(n, \alpha, K)$ satisfying Assumption 1, or
2. G is generated from a subcritical $G(n, w)$.

Then, for any i satisfying $\mu_i < \mu^*$, we have $c^* - c_i \leq 2c^*(\mu^* - \mu_i) + O(1/n)$.

Before stating and proving the lemma, we introduce some useful technical tools. Since we suppose that $G(n, \kappa)$ is subcritical, we have $\mathbb{P}[B(i) = \infty] = 0$ and $x_i = \mathbb{E}[B(i)]$ is finite. First observe that the vector x of expected total progenies satisfies the system of linear equations

$$x = e + \frac{1}{n}Ax,$$

where e is the vector with $e_i = 1$ for all i .

For the analysis of the stochastic block model we define the vector $b \in \mathbb{R}^S$ with coordinates $b_l = \mu_l, l = 1, \dots, S$, where by μ_l we define the expected degree of the node from community H_l . Also we define vector $x' \in \mathbb{R}^S$ with coordinates $x'_l = \mathbb{E}[B(l)], l = 1, \dots, S$, where by $B(l)$ we define the total progeny of the individual of type l . We define $x^* = \max_{i \in [n]} x_i$. Armed with this notation, we begin the proof Lemma 10, which consists of the following steps:

- proving that for any $i, j \in V, x_i - x_j \leq 2x^*(\mu_i - \mu_j)$, (Lemma 11, 12),
- proving that for any $i, j \in V, c_i - c_j = x_i - x_j + O(1/n)$ (Lemmas 13, 14).

These facts together lead to Lemma 10, given that n is large enough to suppress the effects of the residual terms. We begin with analysing the relation between b_l and x'_l in a straightforward way:

Lemma 11 (Coordinate order for mean of the total progeny in the SBM) *Assume that $G(n, \alpha, K)$ is subcritical and that $K_{m\ell} = k > 0$ holds for all $m \neq \ell$. If two coordinates of b are such that $b_l > b_m$, then we have $x'_l > x'_m$, and $x'_l - x'_m \leq 2x^*(b_l - b_m)$.*

Proof For the stochastic block model with S blocks, the system of equations $x = e + Ax$ can be equivalently written as $x' = e + Mx'$, for $M = K \text{diag}(\alpha) \in \mathbb{R}^{S \times S}$, and $x' \in \mathbb{R}^S$, with x'_m now standing for the expected total progeny associated with any node of type m . Similarly, we define b'_m as the expected degree of any node of type m . Notice that the system of equations $x' = e + Mx'$ satisfied by x' can be rewritten as $(I - M)x' = e$, where I is the $S \times S$ identity matrix. By exploiting our assumption on the matrix K and defining $\gamma_m = K_{m,m} - k$, this can be further rewritten as

$$\left(\begin{pmatrix} 1 - \alpha_1 \gamma_1 & & & \\ & \ddots & & \\ & & 1 - \alpha_S \gamma_S & \\ & & & \ddots \end{pmatrix} - k \begin{pmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_S \\ \alpha_1 & \alpha_2 & \cdots & \alpha_S \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_1 & \alpha_2 & \cdots & \alpha_S \end{pmatrix} \right) x' = e,$$

which means that for any m, x'_m satisfies

$$x'_m = \frac{1 + k(\alpha^T x')}{1 - \alpha_m \gamma_m}.$$

Also observe that

$$b'_m = k(\alpha^T \mathbf{1}) + \alpha_m \gamma_m,$$

so, for any pair of types m and ℓ , we have

$$x'_m - x'_\ell = \frac{(1 + k(\alpha^T x'))(\alpha_m \gamma_m - \alpha_\ell \gamma_\ell)}{(1 - \alpha_m \gamma_m)(1 - \alpha_\ell \gamma_\ell)},$$

which proves the first statement.

To prove the second statement, observe that for any pair ℓ and m of communities, we have either $\alpha_m \leq \frac{1}{2}$ or $\alpha_\ell \leq \frac{1}{2}$ (otherwise we would have $\alpha_m + \alpha_\ell > 1$). To proceed, let ℓ and m be such that $x'_m \geq x'_\ell$, and let us study the case $\alpha_\ell \leq \frac{1}{2}$ first. Here, we get

$$\begin{aligned} x'_m - x'_\ell &= \frac{(1 + k(\alpha^\top x'))(\alpha_m \gamma_m - \alpha_\ell \gamma_\ell)}{(1 - \alpha_m \gamma_m)(1 - \alpha_\ell \gamma_\ell)} = \frac{(\alpha_m \gamma_m - \alpha_\ell \gamma_\ell)}{(1 - \alpha_\ell \gamma_\ell)} x'_m \\ &\leq \frac{(\alpha_m \gamma_m - \alpha_\ell \gamma_\ell)}{(1 - \gamma_\ell/2)} x'_m \leq 2x'_m (b'_m - b'_\ell). \end{aligned}$$

In the other case where $\alpha_m \leq \frac{1}{2}$, we can similarly obtain

$$x'_m - x'_\ell \leq 2x'_\ell (b'_m - b'_\ell) \leq 2x'_m (b'_m - b'_\ell).$$

This concludes the proof. ■

For the analysis of the Chung–Lu model, we define $\mu \in \mathbb{R}^n$ as the vector of mean degrees. Then we may prove the following.

Lemma 12 (Coordinate order for mean of the total progeny in the Chung–Lu model) *Assume that $G(n, w)$ is subcritical. If two nodes are such that $\mu_i > \mu_j$, then we have $x_i > x_j$ and $x_i - x_j \leq x^*(\mu_i - \mu_j)$.*

Proof From the system of equations $x = e + \frac{1}{n}Ax$, the coordinates x_i have the form

$$x_i = 1 + \frac{1}{n} \cdot w_i \left(\sum_{j=1}^n w_j x_j \right),$$

which implies that $w_i \geq w_j$ holds if and only if $x_i \geq x_j$. This observation implies for $x^* = \max_i x_i$

$$x_i - x_j \leq \frac{1}{n} \cdot (w_i - w_j) \left(\sum_{j=1}^n w_j \right) x^* = (\mu_i - \mu_j) x^*,$$

thus concluding the proof. ■

The next two lemmas establish the relationship between the expected component size c_i of vertex i and the expected total progeny x_i of the multi-type branching process seeded at vertex i .

Lemma 13 *For any i , the mean of the connected component associated with type i is bounded by the mean of the total progeny: $c_i \leq x_i$.*

Proof Now fix an arbitrary $i \in [n]$ and let $Y_{i,1}, Y_{i,2}, \dots, Y_{i,n}$ be independent Bernoulli random variables with respective parameters $(A_{i,1}/n, A_{i,2}/n, \dots, A_{i,i}/n, \dots, A_{i,n}/n)$. Consider a multitype binomial branching process where the individual of type i produces $Y_{i,j}$ individuals of type j , and let $B_{\text{Ber}}(i)$ denote its total progeny when started from an individual of type i . Recalling the Poisson branching process defined in Appendix 5 with offspring-distributions $X_{i,j}$, we can show $B_{\text{Ber}}(i) \leq B(i)$ using the relation $Y_{i,j} \leq X_{i,j}$.

Considering a node a of type i , we can use Theorem 4.2 of [van der Hofstad \(2016\)](#) to bound the size of the connected component C_a as $|C_a| \preceq B_{\text{Ber}}(i)$, which implies by transitivity of \preceq that $|C_{a_i}| \preceq B(i)$. The proof is concluded by appealing to Theorem 2.15 of [van der Hofstad \(2016\)](#) that shows that stochastic domination implies an ordering of the means. \blacksquare

Next we upper bound the excess that appears in the domination by the branching process:

Lemma 14 $x_i - c_i = O(\frac{1}{n})$.

Proof As in Lemma 13, $B_{\text{Ber}}(i)$ denotes the total progeny of a Bernoulli branching process whose set of parameters corresponds to $G(n, \kappa)$. Then we may decompose the difference as

$$x_i - c_i = x_i - \mathbb{E}[B_{\text{Ber}}(i)] + \mathbb{E}[B_{\text{Ber}}(i)] - c_i.$$

Denote the set of edges in the connected component C_a as $E(C_a)$ and the set of edges containing a vertex v as $E(v)$. We call $|\mathcal{S}|$ the *surplus*, which is the number of edges to be deleted from $E(C_a)$ such that the graph C_a becomes a tree. Then, we have $\mathbb{E}[B_{\text{Ber}}(i)] - c_i \leq \mathbb{E}[|\mathcal{S}|]$. The expectation of the surplus may be written as

$$\begin{aligned} \mathbb{E}[|\mathcal{S}|] &= \mathbb{E}\left[\sum_{e \in E(C_a)} \mathbb{I}\{e \in \mathcal{S}\}\right] = \sum_{k=1}^{\infty} \mathbb{P}[|C_a| = k] \sum_{e \in E(C_a)} \mathbb{E}[\mathbb{I}\{e \in \mathcal{S}\} | |C_a| = k] \\ &= \frac{1}{2} \sum_{v \in C_a} \sum_{e \in E(v)} \mathbb{E}[\mathbb{I}\{e \in \mathcal{S}\} | |C_a| = k]. \end{aligned}$$

Define $A_{\max} = \max_{i,j} A_{i,j}$ as the maximal element of the matrix A . Then for an arbitrary vertex, the probability of an edge $e \in E(v)$ being in the surplus can be upper bounded as

$$\sum_{e \in E(v)} \mathbb{E}[\mathbb{I}\{e \in \mathcal{S}\} | |C_a| = k] \leq \frac{A_{\max} k}{n}.$$

Then we may upper bound the sum as

$$\frac{1}{2} \sum_{v \in C_a} \sum_{e \in E(v)} \mathbb{E}[\mathbb{I}\{e \in \mathcal{S}\} | |C_a| = k] \leq \frac{A_{\max} k^2}{n}.$$

Using our expression for $\mathbb{E}[|\mathcal{S}|]$, we get

$$\mathbb{E}[|\mathcal{S}|] \leq \sum_{k=1}^{\infty} \mathbb{P}[|C_a| = k] \frac{A_{\max} k^2}{n} = \frac{A_{\max} \mathbb{E}[|C_a|^2]}{n}.$$

Now we notice that, by Le Cam's theorem, the total variation distance between the sum of independent Bernoulli random variables with parameters $(A_{i,1}/n, \dots, A_{i,n}/n)$ and the Poisson distribution $\text{Poi}(\sum_{j=1}^n A_{i,j}/n)$ is at most $2(\sum_{j=1}^n A_{i,j}^2)/n$. Using this fact and that the moments of the total progeny of a subcritical branching process do not scale with n (cf. Theorem 1 of [Huaming, 2012](#)), we have $x_i - \mathbb{E}[B_{\text{Ber}}(i)] = O(\frac{1}{n})$, thus proving the lemma. \blacksquare

8.2. Supercritical case

Lemma 15 *Suppose that*

1. G is generated from a supercritical $G(n, \alpha, K)$ satisfying Assumptions 1 and 2, or
2. G is generated from a supercritical $G(n, w)$.

Then, for any node i satisfying $\mu_i < \mu^*$, we have $c^* - c_i \leq c^* (\mu^* - \mu_i) + o_n(n)$.

The proof of Lemma 15 follows from the following lemmas for the stochastic block model and the Chung–Lu model and from the following relation between c_i and ρ_i :

$$c_i = \rho_i \mathbb{E} [|C_1|] + o_n(n),$$

see Bollobás, Janson, and Riordan (2007, Chapter 9).

Lemma 16 (Coordinate order preserving in the stochastic block model.) *Assume the conditions of Lemma 15 and let $l_* = \arg \max_l b_l$. Let $a \in \mathbb{R}^S$ be any vector such that $a_l \in [0, a_{l_*}]$ for all l . Then $(\Phi_M(a))_{l_*} \geq (\Phi_M(a))_l$.*

Proof Let us fix two arbitrary indices l and l' . By the definition of Φ_M , we have

$$\begin{aligned} (\Phi_M(a))_l &= 1 - e^{-((\sum_{m \neq l} \alpha_m a_m)k + \alpha_l K_{l,l} a_l)}, \\ (\Phi_M(a))_{l'} &= 1 - e^{-((\sum_{m \neq l'} \alpha_m a_m)k + \alpha_{l'} K_{l',l'} a_{l'})}. \end{aligned}$$

Notice that if l and l' satisfy

$$\left(\sum_{m \neq l} \alpha_m a_m \right) k + \alpha_l K_{l,l} a_l \geq \left(\sum_{m \neq l'} \alpha_m a_m \right) k + \alpha_{l'} K_{l',l'} a_{l'},$$

we have $(\Phi_M(a))_l \geq (\Phi_M(a))_{l'}$. Now, using the facts that

- $\sum_{m \neq l} \alpha_m a_m - \sum_{m \neq l'} \alpha_m a_m = \alpha_{l'} a_{l'} - \alpha_l a_l$,
- $\alpha_l K_{l,l} \geq \alpha_l k$,
- $\alpha_l K_{l,l} + \alpha_{l'} k \geq \alpha_{l'} K_{l',l'} + \alpha_l k$ and
- $a_l - a_{l'} \geq 0$,

we can verify that

$$\begin{aligned} &\alpha_l K_{l,l} a_l + \alpha_{l'} k a_{l'} - \alpha_l k a_l - \alpha_{l'} K_{l',l'} a_{l'} \\ &= (\alpha_l K_{l,l} + \alpha_{l'} k) a_{l'} + (a_l - a_{l'}) \alpha_l K_{l,l} - (\alpha_{l'} K_{l',l'} + \alpha_l k) a_{l'} - (a_l - a_{l'}) \alpha_l k \geq 0, \end{aligned}$$

thus proving the lemma. ■

Lemma 17 (Order of coordinates of eigenvector in the SBM) *Let a be the eigenvector corresponding to the largest eigenvalue λ of the matrix $M = K \text{diag}(\alpha)$. Then if $l_* = \arg \max_l b_l$, we have $a_{l_*} \geq a_l$ for $l \neq l_*$.*

Proof If a is an eigenvector of M , then for coordinates l, l' :

$$\begin{cases} \left(\sum_{m \neq l} \alpha_m a_m \right) k + \alpha_l k_{l,l} a_l = \lambda a_l, \\ \left(\sum_{m \neq m'} \alpha_m a_m \right) k + \alpha_l K_{l',l'} a_{l'} = \lambda a_{l'} \end{cases}$$

By the Perron-Frobenius theorem and our conditions on matrix M , λ is a real number larger than one. Denote $C = k \sum_{m \neq l, m \neq l'} \alpha_m a_m$, $x = a_l$, $y = a_{l'}$, $a = \alpha_l K_{l,l}$, $b = \alpha_{l'} k$, $c = \alpha_l k$, $d = \alpha_{l'} k_{l',l'}$. Then,

$$\begin{cases} C + ax + by = \lambda x, \\ C + cx + dy = \lambda y \end{cases} \quad (12)$$

Let $r = 1 + \epsilon$ be such that $y = rx = (1 + \epsilon)x$. Then

$$\begin{cases} \frac{C}{x} + a + b + b\epsilon = \lambda, \\ \frac{C}{x} + c + d + d\epsilon = \lambda + \lambda\epsilon \end{cases}$$

and therefore

$$\frac{C}{x} + c + d + d\epsilon = \frac{C}{x} + a + b + b\epsilon + \lambda\epsilon.$$

Rearranging the terms and using the fact that $a + b \geq c + d$, we have

$$0 \leq (a + b) - (c + d) = (d - b - \lambda)\epsilon.$$

Since $K_{l,l} \geq k$, we have $\alpha_l k_{l,l} \geq \alpha_l k$ and $a \geq c$.

We consider two cases separately: First, if $b \geq d$, we have $d - b - \lambda < 0$, which implies $\epsilon < 0$ and $y < x$, therefore proving $a_l > a_{l'}$ for this case. In the case when $b < d$, we have $a + b \geq c + d$ and $\frac{d-b}{a-c} \leq 1$. Subtracting the two equalities of the linear system 12, we get

$$\lambda(1 - r) = (a - c) \left(1 - \frac{d - b}{a - c} r \right).$$

Now, since $\frac{d-b}{a-c} \leq 1$, we have $\lambda \geq a - c$, which implies $\lambda \geq d - b$ and $d - b - \lambda \leq 0$, thus leading to $\epsilon \leq 0$ and $y \leq x$, therefore proving $a_l \geq a_{l'}$ for this case. \blacksquare

Lemma 18 (Order of coordinates of eigenvector in the Chung–Lu model) *Let a be the eigenvector corresponding to the largest eigenvalue λ of the matrix A . Then if $i_* = \arg \max_m b_m$, we have $a_{i_*} \geq a_j$ for $j \neq i_*$.*

Proof It is easy to see that the only eigenvector of A corresponding to a non-zero eigenvalue is $a = w$ with $\lambda_{max} = w^\top w / n$:

$$\frac{1}{n} A w = \frac{1}{n} \cdot (w w^\top) w = \frac{w^\top w}{n} \cdot w.$$

The proof is concluded by observing that the maximum coordinate of the vector b corresponds to the maximum coordinate of w , due to the equality

$$b_i = \frac{1}{n} \cdot w_i \sum_{j=1}^n w_j.$$

■

Lemma 19 (Coordinate order preserving in the Chung–Lu model) *Assume the conditions of Lemma 15 and let $i_* = \arg \max_i b_i$. Let $a = (a_1, \dots, a_n)$ be such that $a_j \in [0, a_{i_*}]$ for all j . Then $(\Phi_A(a))_{i_*} \geq (\Phi_A(a))_j$.*

Proof Let us fix two arbitrary indices i and i' . By the definition of Φ_A , we have

$$(\Phi_A(a))_i = 1 - e^{-w_i(\sum_{j=1}^n w_j a_j)}.$$

Then, using the fact that $w = a$, we have $(\Phi_A(a))_{i_*} \geq (\Phi_A(a))_j$, thus proving the lemma. ■

We finally study the maximal fixed point of the operator Φ_A , keeping in mind this fixed point is exactly the survival-probability vector ρ of the multi-type Galton–Watson branching process [Bollobás, Janson, and Riordan \(2007\)](#). By Lemma 5.9 of [Bollobás, Janson, and Riordan \(2007\)](#), this is the unique fixed point satisfying $\rho_i > 0$ for all i . The following lemma shows that ρ_i takes its maximum at $i_* = \arg \max_i b_i$, concluding the proof of Lemma 15.

Lemma 20 (Fixed point coordinate domination) *Let ρ be the unique non-zero fixed point of Φ_A , and let $i_* = \arg \max_i b_i$. Then, $\rho_{i_*} \geq \rho_j$ and $\rho_{i_*} - \rho_j \leq \rho^*(b_{i_*} - b_j)$ holds for all $j \neq i_*$.*

Proof Letting a be the eigenvector of A that corresponds to the largest eigenvalue λ , Lemma 18 and 17 guarantee $a_{i_*} \geq a_j$ for $j \neq i_*$. Let $\epsilon > 0$ be such that $\epsilon \leq \frac{1-1/\lambda}{a^*}$, where $a^* = \max_{i=1, \dots, S} a_i$. Then by Lemma 5.13 of [Bollobás, Janson, and Riordan \(2007\)](#), $\Phi_M(\epsilon a) \geq \epsilon a$ holds elementwise for the two vectors.

Since the coordinates of the vector ϵa are positive, we can appeal to Lemma 5.12 of [Bollobás, Janson, and Riordan \(2007\)](#) to show that iterative application of Φ_A converges to the fixed point ρ : letting Φ_A^m be the operator obtained by iterative application of Φ_A for m times, we have $\lim_{m \rightarrow \infty} \Phi_A^m(\epsilon a) = \rho$, where ρ satisfies $\rho \geq \epsilon a \geq 0$ and $\Phi_A(\rho) = \rho > 0$. By Lemmas 18 and 17 we have $\rho_{i_*} \geq \rho_j$, for $i_* \neq j$ for both the SBM and the Chung–Lu models, proving the first statement.

The second statement can now be proven directly as

$$\begin{aligned} \rho_{i_*} - \rho_i &= e^{-(\frac{1}{n}A\rho)_j} - e^{-(\frac{1}{n}A\rho)_{i_*}} = e^{-\frac{1}{n}\sum_j A_{i_*j}\rho_j} - e^{-\frac{1}{n}\sum_j A_{ij}\rho_j} \\ &= e^{-\frac{1}{n}\sum_j A_{i_*j}\rho_j} (1 - e^{-\frac{1}{n}\sum_j A_{ij}\rho_j - A_{i_*j}\rho_j}) \leq e^{-\frac{1}{n}\sum_j A_{i_*j}\rho_j} \left(\frac{1}{n} \sum_j (A_{i_*j} - A_{ij})\rho_j \right) \\ &\leq \rho^*(b_{i_*} - b_i), \end{aligned}$$

where the first inequality uses the relation $1 - e^{-z} \leq z$ that holds for all $z \in \mathbb{R}$, and the last step uses the fact that $A\rho$ has positive elements. ■

8.3. Proofs of Theorems 5, 6 and 7 .

Having established that, in order to minimize regret in our setting, it is sufficient to design an algorithm that quickly identifies the nodes with the highest degree. It remains to show that our algorithms indeed achieve this goal. We do this below by providing a bound on the expected number of times $\mathbb{E}[N_{T,i}] = \mathbb{E}\left[\sum_{t=1}^T \mathbb{I}_{\{A_t=i\}}\right]$ that the algorithm picks a suboptimal node i such that $c_i < c^*$, and then using this guarantee to bound the regret.

Without loss of generality, we assume that $V_0 = \{1, 2, \dots, |V_0|\}$. The key to our regret bounds is the following guarantee on the number of suboptimal actions taken by d -UCB(V_0).

Theorem 21 (Number of suboptimal node plays in d -UCB) *Define $\eta_i = (\max_{j \in V_0} \mu_j - \mu_i) / 3$. The number of times that any node $i \in \{i : \mu_i < \max_{j \in V_0} \mu_j\}$ is chosen by d -UCB(V_0) satisfies*

$$\mathbb{E}N_{T,i} \leq \frac{\mu^* (2 + 6 \log T)}{\eta_i^2} + 3. \quad (13)$$

The proof is largely based on the analysis of the kl-UCB algorithm due to [Cappé, Garivier, Maillard, Munos, and Stoltz \(2013\)](#), with some additional tools borrowed from [Ménard and Garivier \(2017\)](#), crucially using that the degree distribution of each node is stochastically dominated by an appropriately chosen Poisson distribution. Specifically, letting Z_i be a Poisson random variable with mean $\mathbb{E}[X_{t,i}]$, we have $\mathbb{E}[e^{sX_{t,i}}] \leq \mathbb{E}[e^{sZ_i}]$ for all s . It turns out that this property is sufficient for the kl-UCB analysis to go through in our case, which is an observation that may be of independent interest.

Before delving into the proof, we introduce some useful notation. We start by defining $Y_{i,1}, \dots, Y_{i,n}$ as independent Bernoulli random variables with respective parameters $\mathbb{B} = (A_{i,1}/n, A_{i,2}/n, \dots, A_{i,n}/n)$, and noticing that the degree $X_{t,i}$ can be written as a sum $X_i = \sum_{j \neq i} Y_{i,j}$. The following lemma, used several times in our proofs, relates this quantity to a Poisson distribution with the same mean.

Lemma 22 *Let $i \in [S]$ and let $Y_{i,1}, Y_{i,2}, \dots, Y_{i,n}$ be independent Bernoulli random variables with respective parameters $p_{i,1}, p_{i,2}, \dots, p_{i,n}$, and let Z_i be a Poisson random variable with parameter $\mu_i = \sum_{j \neq i} p_{i,j}$. Defining $X_i = \sum_{j \neq i} Y_{i,j}$, we have $\mathbb{E}[e^{sX_i}] \leq \mathbb{E}[e^{sZ_i}]$ for all $s \in \mathbb{R}$.*

Proof Fix an arbitrary $s \in \mathbb{R}$ and $i \in [n]$. By direct calculations, we obtain

$$\mathbb{E}e^{sX_i} = \prod_{j=1}^n (\mathbb{E}e^{sY_{i,j}}) \leq \prod_{j=1}^n (1 + p_{i,j}(e^s - 1)) \leq \prod_{j=1}^n \exp(p_{i,j} \cdot (e^s - 1)),$$

where the last step follows from the elementary inequality $1 + x \leq e^x$ that holds for all $x \in \mathbb{R}$. The proof is concluded by observing that $\mathbb{E}e^{sZ_i} = \exp(\mu(e^s - 1))$ and using the definition of μ . \blacksquare

For simplicity, we also introduce the notation $\psi_{\mathbb{B}}(s) = \log \mathbb{E}[e^{sX}]$ and $\phi_{\lambda}(s) = \log \mathbb{E}e^{sZ_i} = \lambda(e^s - 1)$. The proof below repeatedly refers to the Fenchel conjugate of ϕ_{λ} defined as

$$\phi_{\lambda}^*(z) = \sup_{s \in \mathbb{R}} \{sz - \phi(s)\} = z \log\left(\frac{z}{\lambda}\right) + \lambda - z$$

for all $z \in \mathbb{R}$. Finally, we define $d(\mu, \mu') = \mu' - \mu + \mu \log\left(\frac{\mu}{\mu'}\right)$ for all $\mu, \mu' > 0$, noting that $\phi_{\lambda}^*(z) = d(z, \lambda)$.

Proof of Theorem 21. The statement is proven in four steps. Within this proof, we refer to nodes as *arms* and use K to denote the size of V_0 . We use the notation $f(t) = 3 \log t$.

Step 1. We begin by rewriting the expected number of draws $\mathbb{E}[N_i]$ for any suboptimal arm i as

$$\mathbb{E}N_i = \mathbb{E} \left[\sum_{t=K}^{T-1} \mathbb{I}\{A_{t+1} = i\} \right] = \sum_{t=K}^{T-1} \mathbb{P}\{A_{t+1} = i\}.$$

By definition of our algorithm, at rounds $t > K$, we have $A_{t+1} = i$ only if $U_i > U_{i^*}$. This leads to the decomposition:

$$\begin{aligned} \{A_{t+1} = a\} &\subseteq \{\mu^* \geq U_{i^*}(t)\} \cup \{\mu^* < U_{i^*}(t) \text{ and } A_{t+1} = a\} \\ &\subseteq \{\mu^* \geq U_{i^*}(t)\} \cup \{\mu^* < U_i(t) \text{ and } A_{t+1} = a\} \end{aligned}$$

Steps 2 and 3 are devoted to bounding the probability of the two events above.

Step 2. Here we aim to upper bound

$$\sum_{t=K}^{T-1} \mathbb{P}[\mu^* \geq U_{i^*}(t)]. \quad (14)$$

Note, that $\{U_{i^*}(t) \leq \mu^*\} = \{\hat{\mu}_{i^*}(t) \leq U_{i^*}(t) \leq \mu^*\}$. Since $d(\mu, \mu') = \mu' - \mu + \mu \log(\frac{\mu}{\mu'})$ is non-decreasing in its second argument on $[\mu, +\infty)$, and by definition of $U_{i^*} = \sup\{\mu : d(\hat{\mu}_{i^*}(t), \mu) \leq \frac{f(t)}{N_{i^*}(t)}\}$ we have

$$\{\mu^* \geq U_{i^*}(t)\} \subseteq \left\{ \hat{\mu}_{i^*}(t) \leq U_{i^*}(t) \leq \mu^* \text{ and } d(\hat{\mu}_{i^*}(t), \mu^*) \geq \frac{f(t)}{N_{i^*}(t)} \right\},$$

Taking a union bound over the possible values of $N_{i^*}(t)$ yields

$$\{\mu^* \geq U_{i^*}(t)\} \subseteq \bigcup_{n=1}^{t-K+1} \left\{ \mu^* \geq \hat{\mu}_{i^*,n} \text{ and } d(\hat{\mu}_{i^*,n}, \mu^*) \geq \frac{f(t)}{n} \right\} = \bigcup_{n=1}^{t-K+1} D_n(t),$$

where the event $D_n(t)$ is defined through the last step. Since $d(\mu, \mu^*)$ is decreasing and continuous in its first argument on $[0, \mu^*)$, either $d(\hat{\mu}_{i^*,n}, \mu^*) < \frac{f(t)}{n}$ on this interval and $D_n(t)$ is the empty set, or there exists a unique $z_n \in [0, \mu^*)$ such that $d(z_n, \mu^*) = \frac{f(t)}{n}$. Thus, we have

$$\bigcup_{n=1}^{t-K+1} D_n(t) \subseteq \bigcup_{n=1}^{t-K+1} \{\hat{\mu}_{i^*,n} \leq z_n\}.$$

For $\lambda < 0$, let us define $\psi(\lambda)$ as the cumulant-generating function of the sum of binomials with parameters \mathbb{B} , and let $\phi(\lambda)$ be the cumulant-generating function of a Poisson random variable with parameter μ^* . With this notation, we have for any $\lambda < 0$ that

$$\begin{aligned} \mathbb{P}[\hat{\mu}_{i^*,n} \leq z_n] &= \mathbb{P}[\exp(\lambda \hat{\mu}_{i^*,n}) \geq \exp(\lambda z_n)] \\ &= \mathbb{P} \left[\exp \left(\lambda \sum_{i=1}^n X_{i^*,i} - n\psi(\lambda) \right) \geq \exp(n\lambda z_n - n\psi(\lambda)) \right] \\ &\leq \left(\frac{\mathbb{E}e^{\lambda X_{i^*,1}}}{e^{\psi(\lambda)}} \right)^n e^{-n(\lambda z_n - \psi(\lambda))} \leq e^{-n(\lambda z_n - \psi(\lambda))}, \end{aligned}$$

where the last step uses the definition of $\psi(\lambda)$. Now fixing $\lambda^* = \arg \max_{\lambda} \{\lambda z_n - \phi(\lambda)\} = \log(z_n/\mu^*) < 0$, we get by Lemma 22 that

$$e^{-n(\lambda^* z_n - \psi(\lambda^*))} \leq e^{-n(\lambda^* z_n - \phi(\lambda^*))} = e^{-n\phi_{\mu^*}^*(z_n)} = e^{-nd(z_n, \mu^*)}.$$

In view of the definition of z_n and $f(t)$, this gives the bound

$$e^{-nd(z_n, \mu^*)} = e^{-f(t)} = \frac{1}{t^3},$$

which leads to

$$\sum_{t=K}^{T-1} \mathbb{P}[\mu^* \geq U_{i^*}(t)] \leq \sum_{t=K}^{T-1} \sum_{n=1}^{t-K+1} \frac{1}{t^3} < 2,$$

thus concluding this step.

Step 3. In this step, we borrow some ideas by [Ménard and Garivier \(2017, Proof of Theorem 2, step 2\)](#) to upper bound the sum

$$B = \sum_{t=K}^{T-1} \mathbb{P}[\mu^* < U_i(t) \text{ and } A_{t+1} = i]. \quad (15)$$

Writing $\eta = \eta_i = \{\mu^* - \mu_i\}/3$ for ease of notation, we have

$$\begin{aligned} \{\mu^* < U_i(t) \text{ and } A_{t+1} = i\} &\subseteq \{\mu^* - \eta < U_i(t) \text{ and } A_{t+1} = i\} \\ &\subseteq \{d(\hat{\mu}_i(t), \mu^* - \eta) \leq f(t)/N_i(t) \text{ and } A_{t+1} = i\}. \end{aligned}$$

Thus, we have

$$\begin{aligned} B &\leq \sum_{t=K}^{T-1} \mathbb{P}[d(\hat{\mu}_i(t), \mu^* - \eta) \leq f(t)/N_i(t) \text{ and } A_{t+1} = i] \\ &\leq \sum_{n=1}^T \mathbb{P}[d(\hat{\mu}_{i,n}, \mu^* - \eta) \leq f(T)/n] \end{aligned}$$

Defining the integer $n(\eta)$ as

$$n(\eta) = \left\lceil \frac{f(T)}{d(\mu_i + \eta, \mu^* - \eta)} \right\rceil,$$

we have $f(T)/n \leq d(\mu_i + \eta, \mu^* - \eta)$ for all $n \geq n(\eta)$. Thus, we may further upper bound B as

$$\begin{aligned} B &\leq n(\eta) - 1 + \sum_{n=n(\eta)}^T \mathbb{P}[d(\hat{\mu}_{i,n}, \mu^* - \eta) \leq f(T)/n] \\ &\leq \frac{f(T)}{d(\mu_i + \eta, \mu^* - \eta)} + \sum_{n=n(\eta)}^T \mathbb{P}[d(\hat{\mu}_{i,n}, \mu^* - \eta) \leq d(\mu_i + \eta, \mu^* - \eta)]. \end{aligned}$$

By definition of η , we have

$$\{\hat{\mu}_{i,n}, \mu^* - \eta\} \leq d(\mu_i + \eta, \mu^* - \eta) \subseteq \{\hat{\mu}_{i,n} \geq \mu_i + \eta\},$$

which implies

$$\sum_{n=n(\eta)}^T \mathbb{P}[d(\hat{\mu}_{i,n}, \mu^* - \eta) \leq d(\mu_i + \eta, \mu^* - \eta)] \leq \sum_{n=n(\eta)}^T \mathbb{P}[\hat{\mu}_{i,n} \geq \mu_i + \eta].$$

By an argument analogous to the one used in the previous step, we get for a well-chosen λ that

$$\begin{aligned} \sum_{n=n(\eta)}^T \mathbb{P}[\hat{\mu}_{i,n} \geq \mu_i + \eta] &\leq \mathbb{P}[\exp(\lambda \hat{\mu}_{i,n}) \geq \exp(\lambda(\mu_i + \eta))] \\ &= \sum_{n=n(\eta)}^T \mathbb{P}\left[\exp\left(\lambda \sum_{j=1}^n X_{i,j} - n\psi(\lambda)\right) \geq \exp(n\lambda(\mu_i + \eta) - n\psi(\lambda))\right] \\ &\leq \sum_{n=n(\eta)}^T \left(\frac{\mathbb{E}[e^{\lambda X_{i,j}}]}{e^{\psi(\lambda)}}\right)^n e^{-n(\lambda(\mu_i + \eta) - \psi(\lambda))} \\ &\leq \sum_{n=n(\eta)}^T e^{-n(\lambda(\mu_i + \eta) - \phi(\lambda))} = \sum_{n=n(\eta)}^T e^{-nd(\mu_i + \eta, \mu_i)} \\ &\leq \sum_{n=n(\eta)}^{\infty} e^{-nd(\mu_i + \eta, \mu_i)} \leq \frac{1}{e^{d(\mu_i + \eta, \mu_i)} - 1} \leq \frac{1}{d(\mu_i + \eta, \mu_i)}, \end{aligned}$$

where the last step uses the elementary inequality $1 + x \leq e^x$ that holds for all $x \in \mathbb{R}$.

Step 4. Putting together the results from the first three steps, we get

$$\mathbb{E}N_i \leq 3 + \frac{1}{d(\mu_i + \eta, \mu_i)} + \frac{3 \log T}{d(\mu_i + \eta, \mu^* - \eta)}.$$

We conclude by taking a second-order Taylor-expansion of $d(\mu_i + \eta, \mu_i)$ in η to obtain for some $\eta' \in [0, \eta]$ that

$$d(\mu_i + \eta, \mu_i) = \frac{\eta^2}{2(\mu_i + \eta')} \geq \frac{\eta^2}{2(\mu_i + \eta)}.$$

Taking into account the definition of η , we get

$$\frac{1}{d(\mu_i + \eta, \mu_i)} \leq \frac{2\mu^*}{\eta^2}.$$

An identical argument can be used to bound $(d(\mu_i + \eta, \mu^* - \eta))^{-1} \leq 2\mu^*/\eta^2$. \square

The remainder of the section uses Theorem 21 to prove Theorem 5. The proof of Theorem 6 follows from similar ideas and some additional technical arguments.

Proof of Theorem 5. First, by (6),

$$R_T^\alpha \leq \Delta_{\alpha, \max} + \mathbb{E} \left[\sum_{i \in V_0} \Delta_{\alpha, i} \mathbb{E} [N_{T, i}] \middle| \mathcal{E} \right].$$

Now, observing that $\delta_{\alpha, i} \leq 3\eta_i$ holds under event \mathcal{E} , we appeal to Theorem 21 to obtain

$$R_T^\alpha \leq \Delta_{\alpha, \max} + \mathbb{E} \left[\sum_{i \in V_0} \Delta_{\alpha, i} \left(\frac{\mu^* (18 + 27 \log T)}{\delta_{i, \alpha}^2} + 3 \right) \middle| \mathcal{E} \right], \quad (16)$$

thus proving the first statement.

Next, we turn to proving the second statement regarding worst-case guarantees. To do this, we appeal to Propositions 10 and 15 that respectively show $\Delta_i \leq 2c^* \delta_i + O(1/n)$ and $\Delta_i \leq c^* \delta_i + o(n)$ for the sub- and supercritical settings, and we use our assumption that n is large enough so that we have $\Delta_i \leq 3c^* \delta_i$ in both settings. Specifically, we observe that $\delta_i = \Theta_n(1)$ by our sparsity assumption and c^* is $\Theta_n(1)$ in the subcritical and $\Theta_n(n)$ supercritical settings, so, for large enough n , the superfluous $O(1/n)$ and $o(n)$ terms can be respectively bounded by $c^* \delta_i$. To proceed, let us fix an arbitrary $\varepsilon > 0$ and split the set V_0 into two subsets: $U(\varepsilon) = \{i \in V_0 : \delta_{\alpha, i} \leq \varepsilon\}$ and $W(\varepsilon) = V_0 \setminus U(\varepsilon)$. Then, under event \mathcal{E} , we have

$$\begin{aligned} \sum_{i \in V_0} \Delta_{\alpha, i} \mathbb{E} [N_{T, i}] &= \sum_{i \in U(\varepsilon)} \Delta_{\alpha, i} \mathbb{E} [N_{T, i}] + \sum_{i \in W(\varepsilon)} \Delta_{\alpha, i} \mathbb{E} [N_{T, i}] \\ &\leq 3c^* \varepsilon \sum_{i \in U(\varepsilon)} \mathbb{E} [N_{T, i}] + 3c^* \sum_{i \in W(\varepsilon)} \delta_{\alpha, i} \left(\frac{\mu^* (18 + 27 \log T)}{\delta_{\alpha, i}^2} \right) + 3|W(\varepsilon)| \Delta_{\alpha, \max} \\ &\hspace{20em} \text{(by Theorem 21)} \\ &\leq 3c^* \varepsilon T + 3c^* \sum_{i \in W(\varepsilon)} \frac{\mu^* (18 + 27 \log T)}{\delta_{\alpha, i}} + 3|V_0| \Delta_{\alpha, \max} \\ &\leq 3c^* \left(\varepsilon T + |V_0| \frac{\mu^* (18 + 27 \log T)}{\varepsilon} \right) + 3|V_0| \Delta_{\alpha, \max} \\ &\leq 6c^* \sqrt{T|V_0| \mu^* (18 + 27 \log T)} + 3|V_0| \Delta_{\alpha, \max}, \end{aligned}$$

where the last step uses the choice $\varepsilon = \sqrt{|V_0| \mu^* (18 + 27 \log T) / T}$. Plugging in the choice of $|V_0|$ concludes the proof. \square

Proof of Theorem 6. We start by assuming that $\alpha < 1/2$. Also notice that for a uniformly sampled set of nodes U , the probability of U not containing a vertex from V_α^* is bounded as

$$\mathbb{P}[U \cap V_\alpha^* = \emptyset] \leq (1 - \alpha)^{|U|}.$$

By the definition of V_k , this gives that the probability of not having sampled a node from V_α^* in period k of the algorithm is bounded as

$$\mathbb{P}[V_k \cap V_\alpha^* = \emptyset] \leq (1 - \alpha)^{|V_k|} \leq \beta^{-k}.$$

For each period k , the expected regret can be bounded as the weighted sum of two terms: the expected regret of d -UCB (V_k) in period k whenever $V_k \cap V_\alpha^*$ is not empty, and the trivial bound $\Delta_{\alpha, \max} \beta^k$ in the complementary case. Using the above bound on the probability of this event and appealing to Theorem 21 to bound the regret of d -UCB (V_k), we can bound the expected regret as

$$\begin{aligned} \mathbb{E}[R_T^\alpha] &\leq \sum_{k=1}^{k_{\max}} \left(\beta^k \frac{1}{\beta^k} \Delta_{\alpha, \max} + \sum_{i \in V_k} \Delta_{\alpha, i} \left(\frac{\mu^* (2 + 3 \log \beta^k)}{\delta_{\alpha, i}^2} + 3 \right) \right) \\ &\leq k_{\max} \Delta_{\alpha, \max} + \sum_{k=1}^{k_{\max}} \left(\sum_{i \in V_k} \Delta_{\alpha, i} \left(\frac{\mu^* (2 + 3k \log \beta)}{\delta_{\alpha, i}^2} + 3 \right) \right) \\ &\leq k_{\max} \Delta_{\alpha, \max} + \sum_{i \in \bar{V}} \Delta_{\alpha, i} \left(\left(3 + \frac{2\mu^*}{\delta_{\alpha, i}^2} \right) (k_{\max} + 1) + \frac{3 \log \beta (k_{\max} + 1)^2}{2\delta_{\alpha, i}^2} \right). \end{aligned}$$

The proof of the first statement is concluded by upper bounding the number of restarts up to time T as $k_{\max} \leq \frac{\log T}{\log \beta}$.

The second statement is proven by an argument analogous to the one used in the proof of Theorem 5, and straightforward calculations. \square

Proof of Theorem 7. For a node i , such that s_i contains l_i ones, the expected degree is

$$\mu_i = (\zeta + \beta)^{l_i} (\beta + \gamma)^{k - l_i}.$$

Since $\zeta > \gamma > \beta$, we get that $\mu_i > \mu_j$ if $l_i > l_j$. By symmetry of the nodes in the Kronecker graph, if two nodes i and j are such that $l_i = l_j$, then $c_i = c_j$. This implies that for any node i , c_i is a function of l_i . Then we may choose nodes i and j such that $s_i \geq s_j$ coordinate-wise. Then, using the condition $\zeta > \gamma > \beta$, it is straightforward to see that for any vertex k , the probability of the edge (i, k) is greater than that of edge (j, k) . This implies that the connected component is a monotone function of the degree.

Theorem 9.10 in Frieze and Karonski (2015) shows that for a graph $G_{n, P[k]}$ there exists $b(P)$ such that a subgraph of $G_{n, P[k]}$ induced by the vertices $i \in H$ of weight $l_i \geq k/2$ is connected with probability at least $1 - n^{-b(P)}$. We denote by \mathcal{H} the event that the subgraph of $G_{n, P[k]}$ induced by the vertices of H of weight $l \geq k/2$ is connected. This implies that under event \mathcal{H} , $|C_i| = |\max_j C_j|$ for all $i \in H$ and $|C_i| \leq |\max_j C_j|$ for all $i \notin H$. Then we get

$$\begin{aligned} c_\alpha^* - c_i &= \mathbb{E} \left[\max_j |C_j| \middle| \mathcal{H} \right] \mathbb{P}[\mathcal{H}] + \mathbb{E} [|C_\alpha^*| | \mathcal{H}^c] \mathbb{P}[\mathcal{H}^c] - \mathbb{E} \left[\max_j |C_j| \middle| \mathcal{H} \right] \mathbb{P}[\mathcal{H}] - \mathbb{E} [|C_i| | \mathcal{H}^c] \mathbb{P}[\mathcal{H}^c] \\ &\leq \mathbb{E} [|C_*| | \mathcal{H}^c] \mathbb{P}[\mathcal{H}^c] \leq n^{1-b(P)}. \end{aligned}$$

$$\text{For all } i \in V_0 \setminus H, \delta_{\alpha, i} \geq ((\zeta + \beta)(\beta + \gamma))^{k/2} - (\zeta + \beta)^{k/2-1} (\beta + \gamma)^{k/2+1} = ((\zeta + \beta)(\beta + \gamma))^{k/2} \left(1 - \frac{\beta + \gamma}{\zeta + \beta} \right).$$

Since we consider the regime, where $(\zeta + \beta)(\beta + \gamma) > 1$, we get that $\delta_{\alpha, i} > \left(1 - \frac{\beta + \gamma}{\zeta + \beta} \right)$. For all $i, j \in H$, $\delta_{\alpha, i} = (\zeta + \beta)^{l_i} (\beta + \gamma)^{k - l_i} \geq (\zeta + \beta)^{k/2-1} (\beta + \gamma)^{k/2} (\zeta - \gamma) \geq (\zeta - \gamma) / (\zeta + \beta)$. In the same way as we analysed the regret of the stochastic block model and Chung-Lu model, we can write

$$R_T^\alpha \leq nT \mathbb{P}[\mathcal{E}^c] + \mathbb{E} \left[\sum_{i \in V_0} \Delta_{\alpha, i} \mathbb{E}[N_{i, T}] \middle| \mathcal{E} \right].$$

Applying Theorem 21, we get

$$\frac{R_T^\alpha}{n} \leq \mathbb{E} \left[\sum_{i \in V_0 \setminus H} \frac{\Delta_{\alpha,i}}{n} \left(\frac{\mu^*(2 + 6 \log T)}{\left(1 - \frac{\beta + \gamma}{\zeta + \beta}\right)^2} + 3 \right) \middle| \mathcal{E} \right] + n^{-b(P)} \left\lceil \frac{\log(nT)}{\log(2)} \right\rceil \left(\frac{\mu^*(2 + 6 \log T)(\zeta + \beta)^2}{(\zeta - \gamma)^2} + 3 \right) + 1.$$

Applying $|V_0 \setminus H| \leq |V_0|$ and $\Delta_{\alpha,i} \leq n$, we get the final bound on the regret. \square

References

- A. Agarwal, P. Bartlett, and M. Dama. Optimal allocation strategies for the dark pool problem. In Y. W. Teh and M. Titterton, editors, *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, volume 9 of *Proceedings of Machine Learning Research*, pages 9–16, Chia Laguna Resort, Sardinia, Italy, 13–15 May 2010. PMLR. URL <http://proceedings.mlr.press/v9/agarwal10a.html>.
- P. Auer, N. Cesa-Bianchi, and P. Fischer. Finite-time analysis of the multiarmed bandit problem. *Machine Learning*, 47(2-3):235–256, 2002a.
- P. Auer, N. Cesa-Bianchi, Y. Freund, and R. E. Schapire. The nonstochastic multiarmed bandit problem. *SIAM J. Comput.*, 32(1):48–77, 2002b. ISSN 0097-5397.
- G. Bartók, N. Zolghadr, and C. Szepesvári. An adaptive algorithm for finite stochastic partial monitoring, 2012.
- B. Bollobás, S. Janson, and O. Riordan. The phase transition in inhomogeneous random graphs. *Random Struct. Algorithms*, 31(1):3–122, Aug. 2007. ISSN 1042-9832.
- S. Bubeck and N. Cesa-Bianchi. *Regret Analysis of Stochastic and Nonstochastic Multi-armed Bandit Problems*. Now Publishers Inc, 2012.
- O. Cappé, A. Garivier, O.-A. Maillard, R. Munos, and G. Stoltz. Kullback-Leibler Upper Confidence Bounds for Optimal Sequential Allocation. *Annals of Statistics*, 41(3):1516–1541, 2013.
- A. Carpentier and M. Valko. Revealing graph bandits for maximizing local influence. In *Artificial Intelligence and Statistics*, pages 10–18, 2016.
- N. Cesa-Bianchi and G. Lugosi. *Prediction, Learning, and Games*. Cambridge University Press, USA, 2006. ISBN 0521841089.
- K. Chaudhuri, Y. Freund, and D. J. Hsu. A parameter-free hedging algorithm. In *Advances in neural information processing systems*, pages 297–305, 2009.
- W. Chen, C. Wang, and Y. Wang. Scalable influence maximization for prevalent viral marketing in large-scale social networks. In *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD ’10, pages 1029–1038, New York, NY, USA, 2010. ACM. ISBN 978-1-4503-0055-1.

- W. Chen, L. V. Lakshmanan, and C. Castillo. Information and influence propagation in social networks. *Synthesis Lectures on Data Management*, 5(4):1–177, 2013a.
- W. Chen, Y. Wang, and Y. Yuan. Combinatorial multi-armed bandit: General framework and applications. In *International Conference on Machine Learning*, pages 151–159, 2013b.
- A. Chernov and V. Vovk. Prediction with advice of unknown number of experts. In *Proceedings of the Twenty-Sixth Conference on Uncertainty in Artificial Intelligence*, pages 117–125. AUAI Press, 2010.
- F. Chung and L. Lu. The average distances in random graphs with given expected degrees. *Proceedings of the National Academy of Sciences*, 99(25):15879–15882, 2002.
- F. Chung and L. Lu. *Complex Graphs and Networks (Cbms Regional Conference Series in Mathematics)*. American Mathematical Society, USA, 2006. ISBN 0821836579.
- A. Frieze and M. Karonski. *Introduction to Random Graphs*. Cambridge University Press, New York, 2015. Hardcover.
- A. Garivier and O. Cappé. The KL-UCB algorithm for bounded stochastic bandits and beyond. In *Proceedings of the 24th annual Conference On Learning Theory*, pages 359–376, 2011.
- W. Huaming. On total progeny of multitype Galton-Watson process and the first passage time of random walk with bounded jumps. *ArXiv e-prints*, Sept. 2012.
- S. M. Kakade, A. T. Kalai, and K. Ligett. Playing games with approximation algorithms. *SIAM Journal on Computing*, 39(3):1088–1106, 2009.
- D. Kempe, J. Kleinberg, and É. 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*, pages 137–146. ACM, 2003.
- J. Khim, V. Jog, and P.-L. Loh. Adversarial influence maximization. In *2019 IEEE International Symposium on Information Theory (ISIT)*, pages 1–5. IEEE, 2019.
- J. Komiyama, J. Honda, and H. Nakagawa. Regret lower bound and optimal algorithm in finite stochastic partial monitoring. In *Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 1, NIPS'15*, page 1792–1800, Cambridge, MA, USA, 2015. MIT Press.
- W. M. Koolen and T. Van Erven. Second-order quantile methods for experts and combinatorial games. In *Conference on Learning Theory*, pages 1155–1175, 2015.
- I. Kovalenko. Theory of random graphs. *Cybernetics*, 7(4):575–579, 1971.
- T. Lai and H. Robbins. Asymptotically efficient adaptive allocation rules. *Advances in Applied Mathematics*, 6(1):4–22, 1985.
- T. L. Lai. Adaptive treatment allocation and the multi-armed bandit problem. *The Annals of Statistics*, pages 1091–1114, 1987.

- J. Leskovec. *Dynamics of large networks*. PhD thesis, Carnegie Mellon University, School of Computer Science, Machine Learning . . . , 2008.
- J. Leskovec, D. Chakrabarti, J. Kleinberg, and C. Faloutsos. Realistic, mathematically tractable graph generation and evolution, using kronecker multiplication. In A. M. Jorge, L. Torgo, P. Brazdil, R. Camacho, and J. Gama, editors, *Knowledge Discovery in Databases: PKDD 2005*, pages 133–145, Berlin, Heidelberg, 2005. Springer Berlin Heidelberg. ISBN 978-3-540-31665-7.
- J. Leskovec, D. Chakrabarti, J. Kleinberg, C. Faloutsos, and Z. Ghahramani. Kronecker graphs: an approach to modeling networks. *Journal of Machine Learning Research*, 11(2), 2010.
- H. Luo and R. E. Schapire. A drifting-games analysis for online learning and applications to boosting. In *Advances in Neural Information Processing Systems*, pages 1368–1376, 2014.
- O.-A. Maillard, R. Munos, and G. Stoltz. A finite-time analysis of multi-armed bandits problems with Kullback–Leibler divergences. In *Proceedings of the 24th Annual Conference on Learning Theory*, pages 497–514, 2011.
- P. Ménard and A. Garivier. A minimax and asymptotically optimal algorithm for stochastic bandits. In *Proceedings of the 28th International Conference on Algorithmic Learning Theory*, pages 223–237, 2017.
- P. Perrault, J. Healey, Z. Wen, and M. Valko. Budgeted online influence maximization. In *ICML2020*, 2020.
- M. Streeter and D. Golovin. An online algorithm for maximizing submodular functions. In *Advances in Neural Information Processing Systems*, pages 1577–1584, 2009.
- R. van der Hofstad. *Random Graphs and Complex Networks*, volume 1 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, 2016. doi: 10.1017/9781316779422.
- S. Vaswani, L. Lakshmanan, and M. Schmidt. Influence maximization with bandits. *arXiv preprint arXiv:1503.00024*, 2015.
- Q. Wang and W. Chen. Improving regret bounds for combinatorial semi-bandits with probabilistically triggered arms and its applications. In *Advances in Neural Information Processing Systems*, pages 1161–1171, 2017.
- Z. Wen, B. Kveton, M. Valko, and S. Vaswani. Online influence maximization under independent cascade model with semi-bandit feedback. In *Advances in Neural Information Processing Systems*, pages 3026–3036, 2017.