

# Pair-matching: Link prediction with adaptive queries

Christophe Giraud, Yann Issartel, Luc Lehericy, and Matthieu Lerasle

**Abstract.** The pair-matching problem appears in many applications where one wants to discover matches between pairs of entities or individuals. Formally, the set of individuals is represented by the nodes of a graph where the edges, unobserved at first, represent the matches. The algorithm queries pairs of nodes and observes the presence/absence of edges. Its goal is to discover as many edges as possible with a fixed budget of queries. Pair-matching is a particular instance of multi-armed bandit problem in which the arms are pairs of individuals and the rewards are edges linking these pairs. This bandit problem is non-standard though, as each arm can only be played once.

Given this last constraint, sublinear regret can be expected only if the graph presents some underlying structure. This paper shows that sublinear regret is achievable in the case where the graph is generated according to a stochastic block model (SBM) with two communities. Optimal regret bounds are computed for this pair-matching problem. They exhibit a phase transition related to the Kesten–Stigum threshold for community detection in SBM. The pair-matching problem is considered in the case where each node is constrained to be sampled less than a given amount of times, for example for ensuring individual fairness. We show how optimal regret rates depend on this constraint. The paper is concluded by a conjecture regarding the optimal regret when the number of communities is larger than 2. Contrary to the two communities case, we argue that a statistical-computational gap would appear in this problem.

## 1. Introduction

### 1.1. Motivation

Many real world data can be represented as a graph of pairwise relationships. Examples include social networks connections, metabolic networks, protein-protein interaction networks, citations network, recommendations and so on. Matchmaking algorithms and link prediction algorithms are routinely used in many practical situations to discover biochemical interactions, new contacts, hidden connections between criminals, or to match players in online multiplayer video games and sport tournaments. As testing a link in biological networks, or discovering connections between criminals

---

*Mathematics Subject Classification 2020:* 62H30 (primary); 68T05 (secondary).

*Keywords:* adaptive link prediction, stochastic block model, sequential matching.

can be expansive, link prediction algorithms are useful to focus on the most relevant links. In social networks or online video games, they can help in finding relevant partners.

## 1.2. Problem

These applications raise the following mathematical problem that this paper intends to study. Suppose that there exists a graph whose nodes represent a set of entities or individuals and whose edges represent matches between entities or individuals. The nodes are known to the statistician while the edges are typically hidden at first. Matchmaking algorithms make queries on pairs of individuals, trying to discover as many edges as possible. For biological networks like protein-protein interaction networks, the individuals are proteins, an edge is an interaction between the two proteins and a query is an experiment to test whether the interaction exists. The goal of matchmaking algorithms is to discover as many edges of the graph as possible while minimizing the number of mismatches. To stress that the focus lies on discovering graph structures, the problem at hand is called hereafter pair-matching rather than matchmaking.

In this paper, pair-matching algorithms are constrained to explore the graph as they cannot make queries on edges that have already been observed. To learn interesting features on unobserved edges from previous observations, it is necessary to make assumptions on the structure of the hidden graph. This paper considers the arguably simplest situation where the graph has been generated according to an assortative conditional stochastic block model (SBM) [23] with two balanced communities; see Section 2.2 for a formal presentation. In this model, individuals are grouped into two (unobserved) communities and the probability of a match (edge) between two individuals is larger if they belong to the same community than to different ones. In this context, the set of pairs is partitioned into good and bad ones, good pairs contain two individuals from the same community and bad pairs two individuals from different communities. A pair-matching algorithm samples pairs and should sample as many good pairs as possible. Of course, the partition into good and bad pairs is unknown.

When the graph is fully observed, communities are recovered using clustering algorithms, which have been extensively studied over the past few years; see, for example, the recent overviews in [1, 30, 37]. A key parameter in the analysis of clustering algorithms, called here *scaling parameter*  $s$ , is the ratio

$$s = \frac{(p - q)^2}{p + q},$$

where  $p$  is the probability of connection within a community and  $q$  the probability of connection between communities. This parameter measures the difficulty of clustering; see Section 2.2 for details. The quality of a pair-matching algorithm is

evaluated by the expected number of discovered edges after  $T$  queries. Equivalently, the performance can be measured by the expected number of pairs sampled that do not contain edges, which should be as small as possible; see Section 2.4 for details. This last quantity is proportional to the expected number of bad pairs sampled, which is called *sampling regret* in this paper. Besides constraining the algorithms to sample each pair only once, we also force algorithms to sample each individual less than a certain amount of times  $B_T$  before  $T$  queries have been made. This constraint corresponds to practical situations where each individual may not be solicited too many times. For example, for fairness policy, the algorithm may be required to sample a similar amount of time a large fraction of the  $n$  individuals. Such fairness constraint is then implemented by forcing to have at most  $B_T = cT/n$  queries per individual, where  $c \geq 1$ .

### 1.3. Contribution

Our objective is to understand precisely the order of magnitude of the optimal regret in this online learning problem. This task is not trivial due to the originality of our setting, where pairs and nodes cannot be sampled as much as desired, but where the hidden structure of the graph should help to learn useful information. Our main contribution is that the sampling regret of any strategy that cannot sample pairs more than once, that is invariant to nodes labelling and which satisfies the above constraint (see Assumptions (NR), (IL) and (SpS) in Section 2.3 for details) is larger than

$$T \wedge \frac{\sqrt{T} \vee (T/B_T)}{s}, \tag{1.1}$$

up to multiplicative constants. Moreover, a polynomial-time algorithm with sampling regret bounded from above by a constant times  $T \wedge (\sqrt{T} \vee (T/B_T))/s$  is described and analyzed; see Theorem 2. It then follows that  $T \wedge r(T, s)$ , where  $r(T, s) = (\sqrt{T} \vee (T/B_T))/s$  is the order of magnitude of the optimal regret we were looking for and that this rate can be achieved in polynomial time. As a consequence, when  $T = O(1/s^2)$  pairs have been sampled, the linear sampling regret is unavoidable since

$$r(T, s) \geq \frac{\sqrt{T}}{s} \geq T.$$

Likewise,

$$r(T, s) \geq \frac{T}{(sB_T)} \gtrsim T$$

when  $B_T = O(1/s)$  and linear sampling regret is also unavoidable when the constraint is too strong. On the other hand, when  $B_T \gg 1/s$ , and  $T \gg 1/s^2$ , our algorithm reaches the optimal sub-linear sampling regret  $r(T, s)$ . A particular interesting arises

when  $T = n^\alpha$  for some  $\alpha \in (1, 2)$  and when, for fairness requirement, the algorithm is required to sample at most  $cT/n$  time each individual. In this situation actually, the optimal sampling regret becomes of order

$$T \wedge \frac{T^{1/\alpha}}{s}.$$

This result illustrates the price to pay for the fairness constraint: The optimal unconstrained rate  $\sqrt{T}/s$  is replaced by the larger rate  $T^{1/\alpha}/s$ .

#### 1.4. Related literature

The following problem, related to matchmaking, has recently attracted attention, in particular in Bradley–Terry models; see [6, 56]. The task is to infer, from the observation of pairs, a vector of parameters characterizing the strength of players. Most results considered the case where all the graph is observed; see [8, 24]. Recent contributions dealing in particular with ranking issues also consider the case of partially observed graphs; see [26, 45, 46] and the references therein. In all cases, the list of observed pairs is given as input to the algorithm evaluating the strength of all players. The choice of a relevant list of successive observed pairs, independent of the observation of the edges is sometimes called a scheduling problem; see [13]. Scheduling problems are different from matchmaking problems considered here where the algorithm should choose the observed pairs and can use preliminary observations to make its choice. For online video games, classical algorithms used to evaluate strength of players are ELO or TRUESKILL; see [22, 36]. Matchmaking algorithms such as EOMM (see [12]), which is used with TRUESKILL (see [36]), are then used to pair players, taking as inputs these estimated strengths. In this approach, the number of mismatches during the learning phase is not controlled. It is an important conceptual difference with this paper where the matchmaking problem is considered together with the problem of discovering the strength (communities here). Here, pair-matching algorithms have to simultaneously explore the graph to evaluate the strength and sample as many good pairs as possible to optimize the number of matches. Closer to our setting is the active ranking literature, where the goal is to discover adaptively the rank or strength of players with a minimal amount of queries; see [21, 25, 49]. Contrary to our problem, only the exploration matters in adaptive ranking and no notion of regret is investigated.

Pair-matching algorithms take sequential decisions to explore new pairs exploiting previous observations. This kind of exploration and exploitation dilemma is typical in multi-armed bandit problems; see [7, 29, 42, 50]. In stochastic multi-armed bandit problems, a set of actions, called *arms* is proposed to a player who chooses one of these actions at each time step and receives a payoff. The payoffs are independent random variables with unknown distribution. For any arm, payoffs are identi-

cally distributed. The player wants to maximize its total payoff after  $T$  queries. The pair-matching problem introduced above can be seen as a non-standard instance of stochastic multi-armed bandit problems. In this interpretation, each pair of nodes is an arm and the associated payoff is 1 if an edge links these nodes and 0 otherwise. The payoffs hence follow a Bernoulli distribution with parameter  $p$  for good pairs and parameter  $q$  for bad pairs. The unusual feature is that each arm can only be played once, so the pair-matcher must choose a new arm at each time step. For this reason, optimal strategies differ in spirit from classical strategies in bandit problems; see Section 2.1 for more details. While the proofs of lower bounds involve useful inequalities borrowed from the classical bandit literature, in particular the data processing inequality from [18, 28], they require a non-trivial adaptation to work in our setting. The lemmas involved in this adaptation may be of use in a wider class of problems with pairwise observations.

Forgetting the constraint that a node cannot be sampled more than  $B_T$  times, the pair-matching bandit problem could be seen as an extreme version of mortal or rotting bandit problems (see [9, 31, 43]), where every arm would systematically die or have zero pay-off after the first sampling. Without additional assumptions, the regret would be inexorably linear in the querying budget  $T$ . Here, an important difference with classical mortal or rotting multi-armed bandits is that payoffs are structured by the underlying stochastic block model (SBM). While pair-matching can easily be formalized as a bandit problem, and while we explore the parallel with some specific bandit problems like  $k$  out of  $m$  bandit, it turns out that neither the algorithms nor the results or their proof can be used in our framework. In particular, our strategy mixes clustering, iterative rounds of screening and exploitation steps in order to identify a sufficiently large set of nodes in a single community, thus focusing on eliminating sub-optimal arms (at the price of also eliminating optimal ones) rather than identifying good ones as bandit algorithms usually do. Notice though that the paper [15], which is posterior to our work, considered related questions in a different setting.

Stochastic block models have attracted a lot of attention in the recent years, with a focus on the determination of optimal strategies for clustering and for parameter estimation; see [1, 37]. In this prolific literature, the graph is fully observed and the question is to identify precisely the weakest separation between the probabilities of connection necessary to perfectly or partially recover the communities, or to estimate the parameters of the SBM. Closer to our setting, the paper [55] investigates the question of recovering communities from a minimal number of observed pairs, sampled sequentially. In this problem, the question is to assign a community to all nodes after a minimal number  $T$  of time steps and try to minimize the number of misclassified nodes. This is quite different from the minimization of the sampling regret considered here, where we seek to find on a budget as many good pairs as possible and not to

classify all nodes. As discussed in Section 3.3, applying the algorithm of [55] would lead to a suboptimal regret in our problem.

The formalization of the pair-matching problem considered in this paper may be restrictive in some applications. Section 5 presents some conjectures that seem reasonable for  $K$  classes SBMs. Other graph structures would also be interesting, such as Bradley–Terry models of [6, 56], which have been used for sport tournaments in [47], chess ranking in [27] and predictions of animal behaviors in [53]. Various constraints dealing with first discoveries for example may be interesting depending on the applications: the first match of a node is the most important in some situations, and, for the search of a life partner, discovering a match with a node already connected in the observed graph is (for most nodes at least) less interesting than a match with an isolated node. These constraints naturally induce different versions of the pair-matching problem and raise mathematical questions of interest. Multiplayer video games suggest the extension to hypergraphs of the pair-matching problem. Indeed, the value of a player could be evaluated as part of a team and with respect to a possible team of opponents rather than simply as part of a pair. Finally, in many practical situations, additional information on individuals is available and could be used to improve pair-matching algorithms. It is clear from our first results that this information is necessary to avoid linear regret in applications such as life partner research. These extensions are postponed to follow-up works. This paper should be seen as a first step to formalize and study the important sequential pair-matching problem. It focuses on a toy example but opens several interesting questions that arise when dealing with natural constraints in practical applications of interest.

### 1.5. Organization and notation

The remainder of the paper is decomposed as follows. Section 2 introduces the formal setting and objectives. As a warm-up, Section 3 focuses on the case where the algorithms are not constrained to sample nodes more than a certain amount of times. Section 4 presents the main results where the algorithm are constrained. Section 5 gives conjectures for  $K$ -classes SBMs. Finally, in Section 6, we assess the behavior of the algorithms on synthetic data, and we explore the estimation of the scaling parameter as well as the robustness of our result to slight model misspecification. The proofs of the main results are postponed to the appendix.

**Notation.** We write  $x_n \lesssim y_n$  and  $x_n = O(y_n)$ , if there exist numerical constants such that  $x_n \leq C y_n$  for all  $n \geq n_0$ ; and we write  $x_n \asymp y_n$  and  $x_n = \Theta(y_n)$ , if  $x_n = O(y_n)$  and  $y_n = O(x_n)$  that is, if there exist numerical constants  $c, c' > 0$  and  $n_0$  such that  $c x_n \leq y_n \leq c' x_n$  for all  $n \geq n_0$ . We denote by  $\lceil x \rceil$  (respectively,  $\lfloor x \rfloor$ ) the upper (respectively, lower) integer part of  $x$ ; by  $|A|$  the cardinal of a set  $A$ ; and by  $A \Delta B$  the symmetric difference between two sets  $A$  and  $B$ .

## 2. Setting and problem formalization

### 2.1. A special bandit problem

In the pair-matching problem described above, the data-scientist uncovers sequentially a random graph, whose nodes are clustered into two communities. The probability to have an edge between two nodes within a community is  $p$ , while this probability is  $q$  for two nodes belonging to different communities. The problem to discover sequentially as many edges as possible in  $T$  steps, without querying more than  $B_T$  times a node, can be interpreted as a non-standard multi-armed bandit problem. Actually, each pair  $\{a, b\}$  of nodes can be seen as an arm and the discovery of match as a payoff. The payoff of the arm  $\{a, b\}$  follows a Bernoulli distribution with parameter  $p$  if  $a$  and  $b$  belongs to the same communities, and with parameter  $q$  if they are in different communities. This bandit problem is non-standard, as

- (1) the arms are sampled at most once,
- (2) at most  $B_T$  arms involving a given node can be sampled up to time  $T$ ,
- (3) the distribution of the payoffs have a hidden structure inherited from the SBM set-up.

Compared with the standard multi-armed bandit problem, points (1) and (2) make this problem harder, while point (3) is a strong structural property that gives hope to find regimes with sub-linear regret.

These special features make this problem quite different from classical bandit problems. In classical bandit problems, optimal strategies have to identify the best arm (or some of the best arms) and each arm is played many times to reach this goal. Here, half the arms are “optimal” but one cannot play an arm more than once. Therefore, instead of identifying one of these, optimal strategies should avoid bad arms, possibly disregarding a non-negligible proportion of good arms in the process.

The constraint (2) also induces a specific exploration / exploitation trade-off. When the community of a node is identified, we wish to query it with a maximum of nodes of the same community in order to maximize the rewards (exploitation). Yet, we also need to pair this node to some new nodes in order identify the community of new nodes (exploration). Since a node can be queried at most  $B_T$  times, we need to trade-off between these two strategies.

Due to these unusual features of the problem, the classical bandit literature is of little help in order to design some optimal sampling algorithm. It is yet useful to establish our lower bounds, which involve inequalities from [18, 28].

In the remainder of this section, the problem, the assumptions and the objectives are described more formally.

### 2.2. Two-classes SBM

The  $n$  individuals are represented by the set  $V = \{1, \dots, n\}$ . Matches are represented by a set of edges  $E$  between nodes in  $V$ : there is a match between  $a$  and  $b$  in  $V$  if and only if the pair  $\{a, b\}$  belongs to  $E$ . Hereafter, a set of two distinct elements in  $V$  is called a *pair* and an element of  $E$  is called an *edge*. The graph  $(V, E)$  is conveniently represented by its adjacency matrix  $A \in \mathbb{R}^{n \times n}$ , with entries  $A_{ab} = 1$  if  $\{a, b\} \in E$  and  $A_{ab} = 0$  otherwise. In the following, any graph  $(V, E)$  is identified with its adjacency matrix  $(A_{ab})_{a,b \in V}$ . For any pair  $e = \{a, b\}$ , the notations  $A_e$  and  $A_{ab}$  are used indifferently. Since the graph is undirected, the adjacency matrix  $A$  is symmetric, and since there is no self-matching (no self-loop in the graph), the diagonal of  $A$  is equal to zero.

Individuals are grouped into two (unknown) communities according to their affinity. To model this situation, the graph  $(V, E)$  is random and distributed as a two-classes conditional stochastic block model. Let  $0 < q, p < 1$ , and let  $n_1$  denote an integer  $n_1 \geq n - n_1 \geq 1$ . The collection  $\text{cSBM}(n_1, n - n_1, p, q)$  of two-classes conditional stochastic block model distributions on graphs is defined as follows. Let  $G = \{G_1, G_2\}$  be a partition of  $\{1, \dots, n\}$  into two groups, with  $|G_1| = n_1$  and  $|G_2| = n - n_1$ . The partition  $G$  represents the communities of individuals. Let  $\mu_G$  denotes the distribution on graphs with nodes  $\{1, \dots, n\}$ , such that the adjacency matrix is symmetric, null on the diagonal and with lower diagonal entries  $(A_{ab})_{a < b}$  sampled as independent Bernoulli random variables with  $\mu_G(A_{ab} = 1) = p$  when  $a$  and  $b$  belong to the same group  $G_i$ , and  $\mu_G(A_{ab} = 1) = q$  when  $a$  and  $b$  belong to different groups. In other words, two individuals are matched with probability  $p$  if they belong to the same community, and with probability  $q$  otherwise. The class  $\text{cSBM}(n_1, n - n_1, p, q)$  is defined as the set of all distributions  $\mu_G$  defined above, where  $G = \{G_1, G_2\}$  describes the set of partitions of  $\{1, \dots, n\}$  satisfying  $|G_1| = n_1$  and  $|G_2| = n - n_1$ :

$$\begin{aligned} \text{cSBM}(n_1, n - n_1, p, q) = \{ \mu_G : G = \{G_1, G_2\} \text{ is a partition} \\ \text{satisfying } |G_1| = n_1, |G_2| = n - n_1 \}. \end{aligned}$$

In the following, the communities are balanced and matches happen with higher probability if individuals belong to the same community. Formally,  $n$  is even and the graph  $(V, E)$  has been generated according to a distribution  $\mu$  in  $\text{cSBM}(n/2, n/2, p, q)$ , for some unknown parameters  $p$  and  $q$  such that  $0 < q < p \leq 1/2$ . As  $q < p$ , the distribution of  $(V, E)$  is called an *assortative*  $\text{cSBM}(n/2, n/2, p, q)$ . All along the paper, the ratio  $p/q$  is also assumed bounded from above. To sum up,  $p$  and  $q$  are smaller than  $1/2$  and satisfy

$$1 < p/q \leq \rho^*. \tag{2.1}$$

The rationale for the upper bound in (2.1) is provided in the remark below Theorem 1. Given  $p$  and  $q$ , the following scaling parameter plays a central role

$$s = \frac{(p - q)^2}{p + q}. \tag{2.2}$$

This parameter appears in various results in the literature on SBM. The following property, proved for example in [2, 10, 16, 17, 19, 32, 54], will be used repeatedly in the paper. When the graph  $(V, E) \sim \text{cSBM}(n_1, n - n_1, p, q)$ , there exist polynomial-time clustering algorithms that return a partition of  $\{1, \dots, n\}$  such that, with large probability, the proportion of misclassified nodes decreases exponentially:

$$\text{proportion of misclassified nodes} \leq \exp(-cns) \text{ when } ns \geq c',$$

where  $c, c' > 0$  are numerical constants. The rate  $ns$  of exponential decay in this result is optimal (up to a constant) when (2.1) is met. Hence, the scaling parameter  $s$  drives the difficulty of clustering. To stress the importance of  $s$ , the following parametrization will be used henceforth:

$$p = \frac{s(\alpha + \sqrt{\alpha})}{2}, \quad q = \frac{s(\alpha - \sqrt{\alpha})}{2},$$

with  $\alpha = (p + q)^2 / (p - q)^2$ . In this parametrization, assumption (2.1) is met if and only if  $\alpha$  is bounded from below by  $(\rho^* + 1)^2 / (\rho^* - 1)^2$ . Another useful property is that there exist numerical constants  $c_1, c_2 > 0$  such that non-trivial community recovery is possible as soon as  $s \geq c_1/n$  (see [2, 5, 10, 14, 16, 19, 35, 40]) and perfect community recovery is possible as soon as  $s \geq c_2 \log(n)/n$ ; see [2, 11, 39].

The reader familiar with SBM literature may be more comfortable with the parametrization  $p = a_n/n$  and  $q = b_n/n$  for a SBM distribution with two communities. For a comfortable translation of the results, the following relations between  $s, \alpha$  and  $a_n, b_n$  are provided:

$$\begin{aligned} s &= \frac{(a_n - b_n)^2}{n(a_n + b_n)}, & \alpha &= \frac{(a_n + b_n)^2}{(a_n - b_n)^2}, \\ \frac{a_n}{b_n} &= \frac{\alpha + \sqrt{\alpha}}{\alpha - \sqrt{\alpha}}, & a_n + b_n &= n\alpha s. \end{aligned}$$

With these notations, the optimal sampling regret (1.1) can be rewritten as

$$T \wedge \frac{\sqrt{T} \vee (T/B_T)}{s} = T \wedge \left( (\sqrt{T} \vee (T/B_T)) \frac{n(a_n + b_n)}{(a_n - b_n)^2} \right),$$

which is smaller than  $T$  as soon as

$$B_T \wedge \sqrt{T} \geq \frac{n(a_n + b_n)}{(a_n - b_n)^2}.$$

In particular, in the sparse regime where  $a_n$  and  $b_n$  are constants,  $\sqrt{T}$  has to be of the order of  $n$ , while  $s$  has to be of the order of  $1/n$ . The regret, which is lower bounded by  $T \wedge \sqrt{T}/s$ , is then at least of the order of  $T \wedge n^2$ , and thus linear.

### 2.3. Sequential matching strategies

Denote by  $\mathcal{E}$  the set of all pairs of nodes, that is the set of all subsets of  $V$  containing two distinct elements. Heuristically, a sequential matching strategy samples at each time  $t$  a new pair  $\hat{e}_t \in \mathcal{E}$ , using only past observations  $(\hat{e}_1, \dots, \hat{e}_{t-1}, A_{\hat{e}_1}, \dots, A_{\hat{e}_{t-1}})$  and an internal randomness of the algorithm.

Formally, let  $U_0, U_1, \dots$  be i.i.d. uniform random variables in  $[0, 1]$ , independent of  $A$  and representing the sequence of internal randomness for the algorithm. A sequential matching strategy  $\psi$  on  $\mathcal{E}$  (shortened *strategy* in the following) is a sequence

$$\psi = (\psi_t)_{0 \leq t \leq \binom{n}{2}-1}$$

of measurable functions  $\psi_t: \mathcal{E}^t \times \{0, 1\}^t \times [0, 1]^{t+1} \rightarrow \mathcal{E}$ . Any sequential matching strategy  $\psi$  defines a matching algorithm as follows. The first pair is sampled as  $\hat{e}_1 = \psi_0(U_0)$ . Then, at each time  $t \geq 0$ , the pair  $\hat{e}_{t+1}$  is defined by

$$\hat{e}_{t+1} = \psi_t(\hat{\mathcal{E}}_t, (A_e)_{e \in \hat{\mathcal{E}}_t}, U_0, \dots, U_t) \quad \text{with } \hat{\mathcal{E}}_t = \{\hat{e}_1, \dots, \hat{e}_t\}.$$

The strategy takes as input the observed graph  $(A_e)_{e \in \hat{\mathcal{E}}_t}$  and possibly an internal independent randomness  $U_t$  to output the new observed pair  $\hat{e}_{t+1}$ .

In the following, strategies are assumed to satisfy the following constraints: a pair can only be sampled once and strategies are invariant to labelling of the nodes. These constraints can be formalized as follows.

**Non-redundancy (NR).** *The strategy  $\psi$  samples any pair at most once, that is, for any  $0 \leq t \leq \binom{n}{2} - 1$  and  $e_1, \dots, e_t \in \mathcal{E}$ , the map  $\psi_t(\{e_1, \dots, e_t\}, \dots) \notin \{e_1, \dots, e_t\}$ .*

Invariance to labelling requires some notation. For any pair  $e \in \mathcal{E}$  and any strategy  $\psi$ , let

$$N_e(\psi, t) := \mathbf{1}_{e \in \hat{\mathcal{E}}_t} \tag{2.3}$$

indicate if the pair  $e$  has been sampled or not before time  $t$  by the strategy  $\psi$ . For any non-redundant strategy  $\psi$  (i.e., satisfying (NR)), pairs are sampled at most once and the observation of  $\{N_e(\psi, t) : e \in \mathcal{E}\}$  is equivalent to that of  $\hat{\mathcal{E}}_t$ .

Let  $\mu$  be a distribution in  $\text{cSBM}(n/2, n/2, p, q)$  and  $\sigma$  be a permutation of  $V$ . For any pair  $\{a, b\} \in \mathcal{E}$ , let  $\sigma(\{a, b\}) := \{\sigma(a), \sigma(b)\}$ . Let  $\mu^\sigma$  denote the distribution of  $(A_{\sigma(e)})_{e \in \mathcal{E}}$ , where  $(A_e)_{e \in \mathcal{E}}$  is distributed according to  $\mu$ .

**Invariance to labelling (IL).** *The distribution of the outcomes of the strategy  $\psi$  is invariant by permutations of the nodes labels: For any  $\mu \in \text{cSBM}(n/2, n/2, p, q)$  and any permutation  $\sigma$  on  $V$ , the distribution of  $(N_e(\psi, t) : e \in \mathcal{E}, 1 \leq t \leq \binom{n}{2})$  under  $\mu^\sigma$  is the same as the distribution of  $(N_{\sigma(e)}(\psi, t) : e \in \mathcal{E}, 1 \leq t \leq \binom{n}{2})$  under  $\mu$ .*

**Remark 1.** Any algorithm can be made invariant to labelling by simply relabelling all the vertices at random before applying the algorithm.

Besides (NR) and (IL), we consider strategies that do not sample a node more than  $B$  times before time  $T$ . This constraint appears naturally in practical situations. For example, if the algorithm matches biological entities or individuals, one may not want to query too many times each individual for logistic or acceptability reasons. To stress that the constraint  $B$  typically grows with the time horizon  $T$ , it is denoted  $B_T$ . Formally, for any  $a \in V$ , let

$$N_a(\psi, t) = \sum_{b \in V: b \neq a} N_{\{a,b\}}(\psi, t) \quad (2.4)$$

denote the number of times the node  $a$  has been sampled in a pair  $\{a, b\}$  after  $t$  queries.

**Sparse sampling (SpS).** *Let  $T$  and  $B_T$  denote two integers. The strategy  $\psi$  is called  $B_T$ -sparse up to time  $T$  if it satisfies*

$$\forall a \in V, \quad N_a(\psi, T) \leq B_T.$$

Since  $N_a(\psi, T) \leq (n-1) \wedge T$  for all nodes  $a$ , choosing  $B_T \geq (n-1) \wedge T$  corresponds to the unconstrained case.

## 2.4. Objectives of the pair-matcher

Let  $\mu \in \text{cSBM}(n/2, n/2, p, q)$  be the distribution of an assortative conditional stochastic block model with associated partition  $G = \{G_1, G_2\}$ . Pairs within a community have a larger probability to lead to a match, than pairs between two different communities. Accordingly, pairs  $\{a, b\}$  with  $a$  and  $b$  from the same community are called *good pairs*, and  $\mathcal{E}^{\text{good}}(\mu)$  (or simply  $\mathcal{E}^{\text{good}}$ ) denotes the set of such pairs. Similarly,  $\mathcal{E}^{\text{bad}}(\mu)$  (or simply  $\mathcal{E}^{\text{bad}}$ ) denotes the set of pairs  $\{a, b\}$  with  $a$  and  $b$  from two different communities.

The objective of the pair-matcher is to discover as many edges (i.e., matches between individuals) as possible with  $T$  queries, in expectation or with high-probability. For simplicity, we focus henceforth on the maximization of the expected number of edges discovered. The strategy  $\psi$  of the pair-matcher should then maximize the number of discovered edges, in expectation with respect to the randomness of the SBM

and the strategy. Optimal strategies should therefore sample as many pairs in  $\mathcal{E}^{\text{good}}$  as possible. Formally, consider a time horizon  $T$  smaller than

$$|\mathcal{E}^{\text{good}}| = 2 \binom{n/2}{2} \sim n^2/4,$$

and denote by

$$N^{\text{bad}}(\psi, T) = \sum_{e \in \mathcal{E}^{\text{bad}}} N_e(\psi, T),$$

respectively,

$$N^{\text{good}}(\psi, T) = \sum_{e \in \mathcal{E}^{\text{good}}} N_e(\psi, T),$$

the number of pairs in  $\mathcal{E}^{\text{bad}}$  (respectively,  $\mathcal{E}^{\text{good}}$ ) sampled up to time  $T$ . Applying Wald's lemma at the second line, we can compute the expected number of discoveries for any strategy  $\psi$  as follows:

$$\begin{aligned} \mathbb{E}_\mu \left[ \sum_{t=1}^T A_{\hat{e}_t} \right] &= \sum_{e \in \mathcal{E}^{\text{good}}(\mu)} \mathbb{E}_\mu[\mathbf{1}_{e \in \hat{\mathcal{E}}_t} A_e] + \sum_{e \in \mathcal{E}^{\text{bad}}(\mu)} \mathbb{E}_\mu[\mathbf{1}_{e \in \hat{\mathcal{E}}_t} A_e] \\ &= \sum_{e \in \mathcal{E}^{\text{good}}(\mu)} p \mathbb{E}_\mu[\mathbf{1}_{e \in \hat{\mathcal{E}}_t}] + \sum_{e \in \mathcal{E}^{\text{bad}}(\mu)} q \mathbb{E}_\mu[\mathbf{1}_{e \in \hat{\mathcal{E}}_t}] \\ &= p \mathbb{E}_\mu[N^{\text{good}}(\psi, T)] + q \mathbb{E}_\mu[N^{\text{bad}}(\psi, T)] \\ &= pT - (p - q) \mathbb{E}_\mu[N^{\text{bad}}(\psi, T)], \end{aligned}$$

where the last line follows from  $N^{\text{good}}(\psi, T) + N^{\text{bad}}(\psi, T) = T$ . Since  $p > q$ , the maximal expected value of discoveries is achieved by any oracle strategy  $\psi^*$  sampling only edges in  $\mathcal{E}^{\text{good}}$ . In that case,  $N^{\text{bad}}(\psi^*, T) = 0$  and the maximal expected number of discoveries is equal to  $pT$ . The regret of the strategy  $\psi$  is defined as the difference between  $pT$  and its expected number of discoveries:

$$R_T(\psi) = pT - \mathbb{E}_\mu \left[ \sum_{t=1}^T A_{\hat{e}_t} \right] = (p - q) \mathbb{E}_\mu[N^{\text{bad}}(\psi, T)].$$

As long as  $T \leq |\mathcal{E}^{\text{good}}|$ , the regret is proportional to the expected number of sampled between-group pairs  $\mathbb{E}_\mu[N^{\text{bad}}(\psi, T)]$ . Therefore, the main results analyze this last quantity rather than the regret. The expected number of bad sampled pairs  $\mathbb{E}_\mu[N^{\text{bad}}(\psi, T)]$  is called hereafter *sampling-regret*.

**Remark 2.** Without assumption on  $\psi$ , the distribution of  $N^{\text{bad}}(\psi, T)$  may depend on the distribution  $\mu$  of the cSBM. On the other hand, when the strategy  $\psi$  fulfils (II), the distribution of  $N^{\text{bad}}(\psi, T)$  does not depend on the distribution  $\mu$  in cSBM( $n/2, n/2, p, q$ ). Indeed, let  $\mu, \mu'$  be two distributions in cSBM( $n/2, n/2, p, q$ ). By definition,

there exists a permutation  $\sigma$  on  $\{1, \dots, n\}$  such that  $\mu' = \mu^\sigma$ , where  $\mu^\sigma$  has been defined in Section 2.3. Since  $\mathcal{E}^{\text{bad}}(\mu^\sigma) = \sigma^{-1}(\mathcal{E}^{\text{bad}}(\mu))$ , it follows from (IL) that the distribution under  $\mu^\sigma$  of  $\sum_{e \in \mathcal{E}^{\text{bad}}(\mu^\sigma)} N_e(\psi, T)$  is the same as the distribution under  $\mu$  of  $\sum_{e \in \mathcal{E}^{\text{bad}}(\mu)} N_e(\psi, T)$ .

### 3. Warm-up: Unconstrained optimal pair-matching

#### 3.1. Optimal rates for unconstrained pair-matching

As a warm-up, we focus first on the simplest case, where  $B_T = +\infty$ , which amounts to remove the constraint (SpS). Let  $\Psi_\infty$  denote the set of strategies  $\psi$  fulfilling (NR) and (IL). The first main result describes the best sampling-regret that can be achieved by a strategy in  $\Psi_\infty$ , as a function of  $s$  and  $T$ .

**Theorem 1.** *Let  $T$  and  $n$  be positive integers with  $T \leq |\mathcal{E}^{\text{good}}| = 2\binom{n/2}{2}$ . Let  $p, q \in [0, 1/2]$  be two parameters fulfilling (2.1) and such that*

$$s \leq \frac{1}{32(1 + \rho^*)},$$

where the scaling parameter  $s$  is defined in (2.2). For any  $\mu \in \text{cSBM}(n/2, n/2, p, q)$ , we then have

$$\inf_{\psi \in \Psi_\infty} \mathbb{E}_\mu[N^{\text{bad}}(\psi, T)] \geq \frac{1}{32} \left[ \frac{\sqrt{T}}{32(1 + \rho^*)s} \wedge T \right]. \tag{3.1}$$

Moreover, there exist two numerical constants  $c_1, c_2 > 0$ , and a strategy  $\psi \in \Psi_\infty$  corresponding to a polynomial-time algorithm described in Section 3.2, taking  $s$  as input, such that, for any  $p, q$  satisfying (2.1), any  $\mu \in \text{cSBM}(n/2, n/2, p, q)$  and any time horizon  $1 \leq T \leq c_2 n^2$ ,

$$\mathbb{E}_\mu[N^{\text{bad}}(\psi, T)] \leq c_1 \left[ \frac{\sqrt{T}}{s} \wedge T \right].$$

The proof of Theorem 1 is provided in the appendix. The lower bound is proved in Section A and the upper bound in Section B. The upper bound derives from a stronger result showing that similar bounds hold with high probability; see Theorem 7 for a precise statement. Theorem 1 provides only the upper bound in expectation for clarity.

**Remark 3.** The parameter  $\rho^*$  only appears in the lower bound. In fact, the SNR showing up in the proof of the lower bound is  $\tilde{s} := kl(p, q) \vee kl(q, p) \leq 1/16$ , where  $kl(p, q) = p \log(p/q) + (1 - p) \log((1 - p)/(1 - q))$  is the Kullback–Leibler

divergence between two Bernoulli distributions with parameters  $p$  and  $q$ . Under the condition  $p/q \leq \rho^*$ , we have  $s \leq \tilde{s} \leq 2(1 + \rho^*)s$ , so  $s$  and  $\tilde{s}$  are equivalent; see Lemma 19. Thus, the quantity  $\rho^*$  appears when writing the condition with  $s$  instead of  $\tilde{s}$  as SNR.

We have chosen to use  $s$  instead of  $\tilde{s}$  for convenience, as it is a classical SNR in the SBM literature, and it allows us to use existing results and clustering routines straightforwardly. We stress that  $s$  can strongly differ from  $\tilde{s}$  when  $\rho^*$  is large, but this difference is large only in the ‘easiest’ setting where  $p$  and  $q$  are markedly different. On the other hand, the regime where  $\rho^*$  is small is more challenging. In the upper bound, the constants  $c_1$  and  $c_2$  are numerical constants that do not depend on  $\rho^*$ .

Theorem 1 states that, when (2.1) holds, for any  $\mu \in \text{cSBM}(n/2, n/2, p, q)$  and any time horizon  $1 \leq T \leq c_2 n^2$ , the optimal sampling-regret

$$\inf_{\psi \in \Psi_\infty} \mathbb{E}_\mu [N^{\text{bad}}(\psi, T)] \asymp \frac{\sqrt{T}}{s} \wedge T,$$

grows linearly with  $T$  as long as  $T \lesssim 1/s^2$  and becomes sub-linear, of order  $\sqrt{T}/s$ , when  $T \gtrsim 1/s^2$ .

**Remark 4.** For the convenience of the reader familiar with the SBM literature, the conclusion of Theorem 1 in terms of the parametrization  $p = a_n/n$  and  $q = b_n/n$  (as in Section 2.2) is

$$\inf_{\psi \in \Psi_\infty} \mathbb{E}_\mu [N^{\text{bad}}(\psi, T)] \asymp \frac{n(a_n + b_n)}{(a_n - b_n)^2} \sqrt{T} \wedge T,$$

since  $s = (a_n - b_n)^2/n(a_n + b_n)$ .

This result can be understood intuitively. As long as communities cannot be recovered better than random, there is no hope of getting better sampling-regret than with purely random sampling of the pairs. In this regime, the sampling-regret grows linearly with  $T$ . To identify when this occurs, consider the situation where pairs are sampled at random among  $N$  nodes and  $T = \beta N^2/2$  (with  $\beta \leq 1$ ). Then the *observed* edges at time  $T$  are approximately distributed as in a SBM with  $N$  nodes, within-group connection probability  $p_\beta = \beta p$ , and between-group connection probability  $q_\beta = \beta q$ . It follows from [5, 14, 35, 38] that weak recovery of the communities is possible if and only if  $N(p_\beta - q_\beta)^2 \geq 2(p_\beta + q_\beta)$ , which is equivalent to  $\sqrt{\beta T} s \geq \sqrt{2}$  or  $T \geq 2/(\beta s^2)$ . Since  $\beta \leq 1$  by definition, no information about the communities can be recovered when  $T \leq 2/s^2$ . Hence, the sampling-regret is expected to grow linearly with  $T$  for  $T = O(1/s^2)$ . This intuition is confirmed by equation (3.1).

When  $T \gg 1/s^2$ , the situation is different. Classical results, such as those in [2, 10, 16, 19, 32, 39, 54] among others, ensure that the communities of  $N$  nodes can

be recovered almost perfectly if  $N \gg 1/s$  and all edges between these nodes are observed. Therefore, when  $1/s \ll N = (\sqrt{T}/s)^{1/2} \ll \sqrt{T}$ , one can sample all the edges between  $N$  nodes and recover almost perfectly their community with a sampling regret smaller than  $N^2 = \sqrt{T}/s$ .

A recipe in order to get a sublinear regret is the following. If we are able to find a community of  $\Theta(\sqrt{T})$  nodes, then we can spend a budget of  $T$  queries without further regret by sampling pairs among these  $\Theta(\sqrt{T})$  nodes. To do so, we need to identify a community of  $\Theta(\sqrt{T})$  nodes from the  $N$  clustered nodes, with a regret smaller than  $\sqrt{T}/s$ . Given the  $N$  clustered nodes, a carefully designed screening strategy (see Step 2 of Algorithm 1), can identify the community of a new node with a sampling regret of order  $O(1/s)$ . Proceeding recursively,  $\Theta(\sqrt{T})$  new nodes can be identified with a sampling-regret of order  $O(\sqrt{T}/s)$ . The remaining budget of  $T$  queries can then be spent by sampling pairs among these  $\Theta(\sqrt{T})$  nodes without further regret if there were no errors in the community assignment. This informal reasoning suggests that the optimal sampling-regret grows like  $\sqrt{T}/s$  when  $T \gg 1/s^2$ . Again, this intuition is confirmed by equation (3.1). An algorithm achieving the optimal upper bound in Theorem 1 and taking as input  $s$  and the time horizon  $T$  is provided in Section 3.2. It essentially proceeds as in the informal strategy outlined above, even if some steps have to be refined. In particular, the identification of a community of  $\Theta(\sqrt{T})$  nodes has to be conducted with care in order to balance the regret and the community assignment errors. The dependency of the algorithm of Section 3.2 on the time horizon  $T$ , can be easily dropped out with a classical doubling trick; see Section B in the appendix.

To sum up the discussion: in the early stage where  $T = O(1/s^2)$ , one cannot do better than random guessing, up to multiplicative constant factors. In the second stage where  $T \geq 1/s^2$ , the rate  $\sqrt{T}/s$  can be interpreted as follows. A total of  $\Theta(\sqrt{T})$  nodes are involved at time  $T$  and, for each of them,  $\Theta(1/s)$  observations are necessary to obtain an educated guess of their community.

Finally, Theorem 1 can be equivalently stated in terms of the regret  $R_T(\psi)$ : for any time horizon  $1 \leq T \leq c_2 n^2$ , the minimal regret satisfies

$$\inf_{\psi \in \Psi_\infty} R_T(\psi) \asymp \sqrt{\alpha}(\sqrt{T} \wedge (sT)),$$

when the assumptions of Theorem 1 are met.

### 3.2. Algorithm with specified horizon $T$

This section presents an algorithm achieving the upper bound in Theorem 1. This algorithm takes as input the scaling parameter  $s$  and the time horizon  $T$ . This dependency on the time-horizon can be avoided with the classical doubling trick; see Section B in the appendix. We discuss in Section 5.1 a heuristic for the preliminary estimation of  $s$  involving less than  $O(1/s^2)$  edges.

When the horizon  $T$  is  $O(1/s^2)$ , any strategy achieves a regret of order  $O(T)$ . Hence, without loss of generality, it is assumed in the remaining of the section that  $T \geq c_{th}/s^2$  for some numerical constant  $c_{th}$ . Moreover, as Theorem 1 holds for  $T \leq c_2 n^2$ , it is also assumed that this condition is fulfilled for a sufficiently small constant  $c_2$ .

The algorithm proceeds in three steps. In the first step, a core-set  $\mathcal{N}$  of  $|\mathcal{N}| = \Theta(\sqrt{T}/\log(s\sqrt{T}))$  vertices is chosen uniformly at random and each pair within this core-set is sampled with probability  $\Theta((\log(s\sqrt{T}))^2/(s\sqrt{T}))$ . Hence, an average of  $\Theta(\sqrt{T}/s)$  pairs are sampled within this core-set. A community recovery algorithm is run on this observed graph that outputs two estimated communities with a fraction of misclassified nodes vanishing as  $O(\log(s\sqrt{T})/(s\sqrt{T}))$  with high probability.

The second step identifies with high probability  $\Theta(\sqrt{T})$  vertices from the same community, say Community 1. To do so, it picks uniformly at random a set  $\mathcal{A}_0$  of  $8\sqrt{2T}$  vertices outside of the core-set  $\mathcal{N}$  (this is possible thanks to the condition  $T \leq c_2 n^2$ ) and samples pairs between this set and the estimated Community 1 of the core-set. This set of edges is used to estimate the connectivity between these vertices and Community 1. Vertices with low connectivity, that seem to belong to Community 2, are removed online to keep the sampling regret under control. The goal of this screening is not to classify perfectly the  $8\sqrt{2T}$  picked vertices, but instead to sift out vertices of Community 2 with a low sampling regret. In particular, a price to pay to achieve this goal is to possibly remove a non-negligible proportion of vertices of Community 1 from the  $8\sqrt{2T}$  picked vertices. This second step of the algorithm is crucial for getting the optimal regret rate  $O(\sqrt{T}/s)$ . A simplified version of this second step can be connected to a particular  $k$  out of  $m$  best arms identification problem. This connection is discussed in Section 3.3 below.

The third step samples all pairs  $\{a, b\}$  such that  $a$  and  $b$  belong to the  $\Theta(\sqrt{T})$  vertices isolated in the second step of the algorithm, until the remaining budget of  $T$  queries is expended.

The pair-matching algorithm calls an external clustering algorithm (generically denoted by GOODCLUST in what follows). GOODCLUST takes as input a graph  $(V, E)$  and outputs a partition  $\hat{G} = (\hat{G}_1, \hat{G}_2)$ . We require that GOODCLUST fulfils the following recovery property: There exist numerical constants  $c_1^{\text{GC}}, c_1^{\text{GC}} > 0$  such that, for all  $N = N_1 + N_2$  and all  $\tilde{p}, \tilde{q} \in [0, 1]$ , if  $(V, E) \sim \text{cSBM}(N_1, N_2, \tilde{p}, \tilde{q})$ , the proportion of misclassified nodes

$$\varepsilon_N = \frac{|\hat{G}_1 \Delta G_1| + |\hat{G}_2 \Delta G_2|}{2N},$$

with  $\Delta$  the symmetric difference, satisfies

$$\varepsilon_N \leq \exp\left(-c_1^{\text{GC}} N \frac{(\tilde{p} - \tilde{q})^2}{\tilde{p}}\right), \tag{3.2}$$

**Algorithm 1: Unconstrained algorithm**

**Inputs:**  $s$  scaling parameter,  $T$  time horizon,  $V$  set of nodes.

**Internal constants:**  $c_{\theta_0} = 2 \vee (1/c_1^{\text{GC}})$ ,  $C_k = 2200$  and  $C_I = 4$ .

**Step 1: finding communities in a core-set**

- (1) Sample uniformly at random a set  $\mathcal{N} \subset V$  of  $N = \lceil \sqrt{T} / \log(s\sqrt{T}) \rceil$  nodes.
- (2) Sample each pair of  $\mathcal{N}$  with probability  $c_{\theta_0}(\sqrt{T}/s \binom{N}{2})$ , call  $\mathcal{O}_0 \subset \mathcal{E}$  the output.
- (3) Estimate global connectivity  $\tau = (p + q)/2$  by  $\hat{\tau} = \frac{1}{|\mathcal{O}_0|} \sum_{e \in \mathcal{O}_0} A_e$ .
- (4) Run GOODCLUST on the graph with nodes set  $\mathcal{N}$  and edges present in  $\mathcal{O}_0$ .  
Output, for any  $x \in \mathcal{N}$ ,  $\hat{Z}_x$  the estimated community of  $x$ . Choose the label  $\hat{Z} = 1$  for the largest estimated community.

**Step 2: expanding the communities**

- (5) Sample uniformly at random a set  $\mathcal{A}_0$  of  $|\mathcal{A}_0| = \lceil 8\sqrt{2T} \rceil$  nodes in  $V \setminus \mathcal{N}$ .
- (6) Set  $k = \lceil C_k/s \rceil$  and  $I = \lceil C_I \log(s\sqrt{T}) \rceil$ .
- (7) **For**  $i = 1, \dots, I$ , **do**
  - (a) **For**  $x \in \mathcal{A}_{i-1}$ , sample  $k$  nodes  $(y_{k(i-1)+a}^x)_{a=1, \dots, k}$  uniformly at random in  $\mathcal{N} \cap \{\hat{Z} = 1\} \setminus \{y_a^x\}_{a=1, \dots, k(i-1)}$ .
  - (b) Sample the pairs  $(\{x, y_{k(i-1)+a}^x\})_{a=1, \dots, k}$  and let  $\hat{p}_{x,i} = \frac{1}{ki} \sum_{a=1}^{ki} A_{xy_a^x}$ .
  - (c) Select  $\mathcal{A}_i = \{x \in \mathcal{A}_{i-1} : \hat{p}_{x,i} \geq \hat{\tau}\}$ .
  - (d) **In case**<sup>1</sup>  $\mathcal{A}_i = \emptyset$ , **then** set  $\mathcal{A}_I = \emptyset$  and **BREAK**.

**Step 3: sampling pairs within estimated communities**

- (8) Sample uniformly at random pairs within the set  $\mathcal{A}_I$  until  $T$  pairs have been sampled overall. If the number of sampled pairs is smaller than  $T$  after all pairs in  $\mathcal{A}_I$  have been sampled, then sample the remaining pairs at random.

**Output:**  $T$  pairs sampled at points (2), (7.b) and (8) of the algorithm.

with probability at least  $1 - c^{\text{GC}}/N^3$ . Algorithms achieving this proportion of misclassification can be found, e.g., in [19]; see also [2, 10, 16, 17, 32, 54] for similar results.

<sup>1</sup>with high probability, this undesirable case does not happen

### 3.3. Community expansion versus $k$ out of $m$ best arm identification

As proved in Lemma 8 in Appendix B, after Step 1, with high probability, we end up with a set of  $N$  classified nodes, where at most  $O(1/s)$  of them are misclassified, and the empirical connectivity  $\hat{\tau}$  does not deviate from the population one  $\tau = (p + q)/2$  by more than  $(p - q)/4$ . The goal of Step 2 is then to identify  $\sqrt{2T}$  new nodes of Community 1, with at most  $O(1/s)$  misclassified nodes and a regret at most  $O(\sqrt{T}/s)$ . Let us connect this problem to a  $k$  out of  $m$  best arms identification problem.

Let us consider a simplified version of the problem of Step 2. Assume that we have identified  $N_1 = N/2$  nodes of Community 1 with no error, that we have access to the population connectivity  $\tau$  and that among the  $M = 8\sqrt{2T}$  nodes in  $\mathcal{A}_0$ , half of them are of Community 1. Then, each node  $a \in \mathcal{A}_0$  can be seen as an arm, and pulling the arm  $a$  amounts to query a pair  $\{a, b\}$  with  $b$  one of the  $N_1$  nodes of Community 1 identified at Step 1. The mean reward of the arm  $a$  is  $p$  if it belongs to Community 1, and  $q$  otherwise. Hence, a simplified version of the problem in Step 2 amounts to identify  $k = \sqrt{2T}$  out of  $m = M/2 = 4\sqrt{2T}$  best arms, with at most  $O(1/s)$  errors, and a cumulated regret  $O(k/s)$ . We have the additional constraint that an arm can be pulled at most  $N_1$  times, but we will forget this additional feature in this discussion, for simplicity of the comparison.

The problem of identifying  $k$  out of  $m$  best arms with a tolerance  $\epsilon$  has been investigated in [20, 41]. The focus on these papers is on the minimal sample size needed to identify  $k$  arms whose expected reward is larger than the  $m$ th largest expected reward minus  $\epsilon$ . The main result of [41] states that, with probability at least  $1 - k^{-2}$  and with a sample size of

$$O\left(\frac{1}{(p - q)^2} \left( M \log\left(\frac{m + 1}{m + 1 - k}\right) + k \log(k) \right)\right),$$

the algorithm AL-Q-FK can recover  $k$  out of the  $m$  best arms with a tolerance  $\epsilon = (p - q)/2$ . The sampling regret is not considered and it can be as large as the sample size. In the same setting, the screening algorithm of Step 2 achieves the following performance. For  $m \geq ck \geq c'/s$ , with probability at least  $1 - c''k^{-2}$  a budget of at most  $O(ks^{-1} \log(sk))$  queries, and a sampling regret at most  $O(k/s)$ , the algorithm identifies a set of arms with at least  $k$  out of  $m$  best arms and at most  $O(1/s)$  arms not in the  $m$  best ones. As  $s = (p - q)^2/(p + q)$ , the sampling regret achieved by the screening algorithm of Step 2 is at least  $(p + q)/\log(k)$  times smaller. We can explain this gain by several reasons. The  $p + q$  improvement comes from the fact that we explicitly take into account the fact that the rewards have a Bernoulli distribution. The  $1/\log(k)$  improvement is obtained by a careful design of the algorithm to keep the regret low, at the price of possibly  $O(1/s)$  identification errors.

Specified to the simplified version of the problem in Step 2 depicted above, the AL-Q-FK algorithm would return  $k = \sqrt{2T}$  nodes out of the  $m = M/2 = 4\sqrt{2T}$  nodes of Community 1 with a sampling regret

$$O\left(\frac{\sqrt{T}}{(p+q)s} \log(\sqrt{T})\right).$$

This sampling regret is larger than the  $O(\sqrt{T}/s)$  regret needed for our Step 2, so the AL-Q-FK cannot be used as a black-box for Step 2.

We emphasize also that the expansion of the communities in Step 2 is somewhat more complex than the simplified version described above: at Step 1, up to  $O(1/s)$  nodes are misclassified, we only have access to the empirical connectivity  $\hat{\tau}$ , an arm can only be pulled  $N_1$  times and the number of best arms is random.

We also emphasize that we cannot use the algorithm of [55] as a black-box to identify  $\sqrt{2T}$  nodes of Community 1 within  $\mathcal{A}_0$  with at most  $1/s$  errors and with a sampling regret  $O(\sqrt{T}/s)$ . Indeed, if we take  $\Theta(\sqrt{T})$  nodes and apply the procedure of [55] to classify them with a sampling regret at most  $O(\sqrt{T}/s)$ , then a *fixed* proportion of the nodes are misclassified and pairing them together at Step 3 would generate a final regret of order  $\Theta(T)$ . In addition, from the lower bounds in [55], we observe that the above phenomenon occurs, whatever the algorithm, if we try to classify *all* the nodes in  $\mathcal{A}_0$ . To overcome this issue, the algorithm of Step 2 recovers the class for a *fraction* only of the nodes in  $\mathcal{A}_0$  with a sampling regret at most  $O(\sqrt{T}/s)$  and at most  $1/s$  errors. When recovering the class of  $\sqrt{2T}$  nodes within  $\mathcal{A}_0$ , we do not sample pairs at random, but we carefully select them in order to avoid as much as possible the sampling of bad pairs.

## 4. Constrained optimal pair-matching

### 4.1. Main results

Let us now consider the general problem, where sparse sampling (SpS) is enforced. The algorithm described in Section 3.2 for unconstrained pairs-matching uses extensively the opportunity to make “localized” queries: At time  $T$ , a small number of  $\Theta(\sqrt{T})$  nodes has been queried a large number of  $\Theta(\sqrt{T})$  times, while other nodes have been queried less than  $O(\log(s\sqrt{T})^2/s)$  times. So, the strategy has to be adapted to fulfil (SpS).

For a sparsity bound  $B_T$ , denote by  $\Psi_{B_T, T}$  the set of strategies  $\psi$  fulfilling the Non-redundancy (NR), Invariance to labelling (IL) and Sparse sampling (SpS) properties at time  $T$ .

**Theorem 2.** *Let  $T$  and  $n$  be positive integers with  $T \leq |\mathcal{E}^{\text{good}}| = 2\binom{n}{2}$ . Let  $p, q \in [0, 1/2]$  be two parameters fulfilling (2.1) and such that the parameter  $s$ , defined in (2.2), fulfils*

$$s \leq \frac{1}{32(1 + \rho^*)}.$$

*Then, for any  $\mu \in \text{cSBM}(n/2, n/2, p, q)$ ,*

$$\inf_{\psi \in \Psi_{B_T, T}} \mathbb{E}_\mu[N^{\text{bad}}(\psi, T)] \geq \frac{1}{32} \left[ \frac{\sqrt{T} \vee (T/B_T)}{32(1 + \rho^*)s} \wedge T \right].$$

*Conversely, there exist two numerical constants  $c_1, c_2 > 0$  such that, for any time horizon  $T$  and constraint  $B_T$  satisfying  $1 \leq T \leq c_1 n (B_T \wedge n)$ , there exists a strategy  $\psi \in \Psi_{B_T, T}$  corresponding to a polynomial-time algorithm, described in Section 4.2, such that*

$$\mathbb{E}_\mu[N^{\text{bad}}(\psi, T)] \leq c_2 \left[ \frac{\sqrt{T} \vee (T/B_T)}{s} \wedge T \right]. \tag{4.1}$$

We refer to the appendix for a proof of this theorem. The lower bound is proved in Section A and the upper bound in Section C.

Compared with Theorem 1, Theorem 2 shows that the sparse sampling constraint (SpS) amounts to replace  $\sqrt{T}$  by  $\sqrt{T} \vee (T/B_T)$  in the optimal sampling-regret. In particular, the sparse sampling constraint downgrades optimal rates only when  $B_T$  is smaller than  $\sqrt{T}$ . Actually, a close look at the unconstrained algorithm (Algorithm 1) reveals that, by construction, it satisfies assumption (SpS) with  $B_T = 17\sqrt{T}$ . So, in the regime where  $B_T \geq 17\sqrt{T}$ , the lower bound cannot be worse than the upper-bound of the unconstrained setting of Theorem 1.

When  $B_T \lesssim \sqrt{T}$ , the optimal sampling-regret is of order  $(T/(B_T s)) \wedge T$ . This rate can be understood as follows. If  $B_T \leq 1/s$ , there is not enough observations per node to infer their community better than at random, which induces an unavoidable linear regret. When  $B_T \gg 1/s$ , to proceed as in Step 3 of the constrained case, one needs to identify a sufficiently large set of nodes of the same community, among which one can sample up to  $T$  pairs without adding regret. As each node can now be paired with at most  $B_T$  others, this set should be of size  $\Theta(T/B_T)$  instead of  $\Theta(\sqrt{T})$  in the unconstrained case. As the identification of the community of a node requires at least  $\Theta(1/s)$  queries, the sampling-regret expected to identify this large set of nodes is  $\Theta(T/(B_T s))$ .

The previous informal discussion suggests to extend the algorithm described in Section 3.2 for the unconstrained case. This extension, fully described and commented in Section 4.2, still proceeds in three steps and goes as follows. The first step of the constrained algorithm is essentially the same as the first step of the unconstrained algorithm, with  $\sqrt{T}$  replaced by  $B = (B_T \wedge \sqrt{T})/2$ . In this first step, all pairs are

sampled among a set of  $B/\log(sB) \leq B_T$  nodes, so the constraint cannot be violated. Then, to keep the sampling-regret under control while not violating the (SpS) constraint, the trick is to apply recursively a variant of the screening algorithm in Step 2 and repeat these screenings until a total number of  $\Theta(T/(B_T \wedge \sqrt{T}))$  nodes are correctly classified, with a small proportion of error. Finally, one can sample at most  $B_T \wedge \sqrt{T}$  pairs for each of these nodes in Step 3 with a controlled regret. The resulting algorithm extends the unconstrained one of Section 3.2 where  $B_T \wedge \sqrt{T} = \sqrt{T}$  and where the screening step is only applied once. This extension is fully described in Section 4.2.

To illustrate the theorem, one can discuss the results with the constraint  $B_T = T^\gamma$ , where  $0 < \gamma \leq 1/2$ . As mentioned in the introduction, this situation arises with  $\gamma = 1 - 1/\alpha$  when  $T = n^\alpha$ , and when, for fairness reasons, the algorithm is required to sample at most  $B_T = cT/n = cT^{1-1/\alpha}$  times each node. In this case, the optimal sampling-regret is of order  $T \wedge (T^{1-\gamma}/s)$  which becomes, in the example discussed in introduction the mentioned rate  $T \wedge T^{1/\alpha}/s$ . It follows that any pair-matching algorithm that is  $T^\gamma$ -sparse up to time  $T$  (besides satisfying (NR) and (IL)) has linear sampling-regret up to time  $s^{-1/\gamma}$ . On the other hand, there exist strategies with optimal sampling-regret of order  $T^{1-\gamma}/s$  after time  $s^{-1/\gamma}$ .

Notice that the sparse sampling property  $N_a(\psi, T) \leq B_T$  only constrains the algorithm at the time horizon  $T$ . This time horizon has therefore to be specified beforehand for this constraint to be defined. In many practical situations, this specification is not reasonable and a more realistic constraint takes the form:  $N_a(\psi, t) \leq B_t$  at any time  $t \in \{1, \dots, T\}$ . In the case where  $B_t = \Theta(t^\gamma/(\log t)^\tau)$ , the constraint can be enforced using a doubling trick, without enlarging the regret by more than a multiplicative numerical constant. This doubling trick is discussed in detail in Section 4.4.

## 4.2. Algorithm with sparse sampling

Algorithm 1 achieves optimal regret in the unconstrained case. It identifies first a set of  $\Theta(\sqrt{T})$  nodes from one community with  $O(1/s)$  misclassified nodes and a regret of order  $O(\sqrt{T}/s)$  in Steps 1 and 2. Then, it pairs these nodes together in Step 3 with a  $O(\sqrt{T}/s)$  regret (due to the misclassified nodes).

The algorithm described in this section follows essentially the same steps. It identifies first a set of nodes from a single community (with small error) and then samples pairs among them. It has to be adapted to fulfil the (SpS) constraint. As the unconstrained algorithm fulfils the (SpS) constraint for any  $B_T \geq 17\sqrt{T}$ , it is assumed in the remaining of this section that  $B_T = O(\sqrt{T})$ . Moreover, as the result holds for  $T \leq c_1 n(B_T \wedge n)$ , this assumption is granted in the remaining of the section.

To respect the constraint (SpS), no node may be sampled in more than  $B_T$  pairs. Hence, to perform the last step, the algorithm has to identify  $\Theta(T/B_T)$  nodes from

one community. It should achieve this identification with a sampling-regret smaller than  $O(T/(sB_T))$  while respecting the (SpS) constraint. To respect the (SpS) constraint in the first step of the algorithm, a core-set  $\mathcal{N}_{\text{init}}$  of cardinality smaller than  $B_T$  is chosen. Formally, in points (1) and (2) of Step 1 in Algorithm 1,  $\sqrt{T}$  is replaced by  $(B_T \wedge \sqrt{T})/2$ . Then, as in the unconstrained case, Step 2 expands the communities in order to identify, with high probability and up to a small error,  $\Theta(T/B_T)$  nodes from one community. The main difference with the unconstrained case is that this expansion cannot be achieved in a single step of screening. Actually,

- (i)  $\Theta(N/s)$  pairs are required to identify the community of  $\Theta(N)$  new nodes;
- (ii) any node from the core-set  $\mathcal{N}_{\text{init}}$  cannot be sampled more than  $B_T$  times.

By (ii), one cannot sample more than  $O(|\mathcal{N}_{\text{init}}|B_T)$  pairs and by (i), it follows that at most  $O(|\mathcal{N}_{\text{init}}|B_T s) = O(B_T^2 s)$  nodes can be classified with a single screening step based on  $\mathcal{N}_{\text{init}}$ . The main idea of the new algorithm is to iterate the screening step, expanding progressively the communities. Along these iterations, to satisfy the (SpS) constraint, the screening has to be conducted with more care than in Step 2 of the unconstrained algorithm (Algorithm 1). The trick is to apply the SCREENING function (Algorithm 3), which compartmentalizes the nodes in order to enforce the condition (SpS). This iterative process outputs a set of  $\Theta(T/B_T)$  nodes from a single community (with a small proportion of error with high probability). The algorithm finally pairs nodes among this subset while respecting the (SpS) constraint in Step 3 of the algorithm.

### 4.3. Screening versus $k$ out of $m$ best arms identification

Similarly to Section 3.3, let us compare the screening step to a  $k$  out of  $m$  best arms identification problem. The main additional feature compared to the situation discussed in Section 3.3, is that an arm  $a$  cannot be sampled more than  $B$  times. Hence, a simplified version of the screening problem amounts to identify  $k$  out of  $m$  best arms with tolerance  $\epsilon = (p - q)/2$ , with the constraint that each arm cannot be sampled more than  $B$  times. In these simplified setting, the screening function achieves the following performance. Assume that  $M \geq ck$  and  $k, B \geq c'/s$ . With probability  $1 - c(sk)^{-1}$ , with a budget of  $O(k s^{-1} \log(s(B \wedge k)))$  queries, and with a sampling regret at most  $O(k/s)$ , the screening function identifies at least  $k$  arms of Community 1 with at most  $O((k(sB)^{-1}) \vee s^{-1})$  errors.

The situation handled by the screening function is actually somewhat more complex than the stylized bandit problem depicted above. Actually, among the initial set of  $N$  classified nodes, we have up to  $cN/(sB)$  misclassified nodes. At the same time, we cannot query more than  $B$  times any of these classified nodes. Hence, we need a careful querying policy in order to avoid the misclassified nodes to generate errors,

while keeping the (SpS) condition enforced. Fulfilling together these two conditions is the main hurdle in the design and analysis of the screening function.

### Algorithm 2: Constrained algorithm

**Inputs:**  $s$  scaling parameter,  $T$  time horizon,  $V_{\text{init}}$  the set of the  $n$  nodes of the whole graph,  $B_T$  constraint.

**Internal constants:** set  $c_{\theta_0} = 8 \vee (1/c_1^{\text{GC}})$  and  $B = (B_T \wedge \sqrt{T})/2$ .

#### Step 1: finding communities in a core-set

- (1) Sample uniformly at random an initial set  $\mathcal{N}_{\text{init}} \subset V_{\text{init}}$  of  $N_{\text{init}} = \lceil \frac{B}{\log(sB)} \rceil$  nodes.
- (2) Sample each pair of  $\mathcal{N}_{\text{init}}$  with probability  $c_{\theta_0} \frac{B}{s} / \binom{N_{\text{init}}}{2}$ , call  $\mathcal{O}_0 \subset \mathcal{E}$  the output.
- (3) Estimate mean connectivity  $\tau = (p + q)/2$  by  $\hat{\tau} = \frac{1}{|\mathcal{O}_0|} \sum_{(x,x') \in \mathcal{O}_0} A_{x,x'}$ .
- (4) Run GOODCLUST on the graph  $(\mathcal{N}_{\text{init}}, \mathcal{O}_0)$  and output, for any  $x \in \mathcal{N}_{\text{init}}$ ,  $\hat{Z}_x$  the estimated community of  $x$  (with the convention that the largest estimated community is labelled by 1).

#### Step 2: iteratively expanding the communities

**Internal constants:** set  $N^{(0)} = \lceil N_{\text{init}}/2 \rceil$ ,

$$t_f = \left\lceil \frac{\log(\lceil 2T/B \rceil / N^{(0)})}{\log \lfloor \log(sB) \rfloor} \right\rceil$$

and for all  $t \in \{0, \dots, t_f\}$ ,

$$N^{(t)} = N^{(0)} \lfloor \log(sB) \rfloor^t \wedge \left\lceil \frac{2T}{B} \right\rceil.$$

- (5) Let  $\mathcal{N}^{(0)}$  be a set of  $N^{(0)}$  nodes in  $\mathcal{N}_{\text{init}} \cap \{\hat{Z} = 1\}$  sampled uniformly at random, and let  $V^{(0)} = V_{\text{init}} \setminus \mathcal{N}_{\text{init}}$ .
- (6) **For**  $t = 1, \dots, t_f$ , **set**

$$(\mathcal{N}^{(t)}, V^{(t)}) = \text{SCREENING}(\mathcal{N}^{(t-1)}, N^{(t)}, B, \hat{\tau}, V^{(t-1)}).$$

#### Step 3: sampling pairs within estimated communities

- (7) Sample pairs within the set  $\mathcal{N}^{(t_f)}$  while respecting the constraint (SpS) with  $B_T$ , until  $T$  pairs have been sampled overall (the sampling method does not matter).

**Algorithm 3: Function** SCREENING( $\mathcal{N}, N', B, \nu, V$ ) = ( $\mathcal{N}', V'$ )

**Inputs:** a reference core-set  $\mathcal{N}$  of cardinality  $N$ , a target number of nodes  $N'$ , a constraint  $B \in \mathbb{R}_+$ , a threshold  $\nu \in [0, 1]$ , a set of “new” nodes  $V$ .

**Output:** a set of nodes  $\mathcal{N}' \subset V$  of cardinality at most  $N'$  and the set of nodes  $V' \subset V$  that are still “new” after running SCREENING. (Most of the nodes of  $\mathcal{N}'$  will belong to the most represented community in  $\mathcal{N}$ .)

**Internal constants:** a number of pairs per step  $k = \lceil C_k/s \rceil$  and a number of steps  $I = \lceil C_I \log(sB) \rceil$ , with  $C_k = 2500$ ,  $C_I = 1026$ .

- (1) Sample uniformly at random a set  $\mathcal{A}_0$  of  $|\mathcal{A}_0| = 4N'$  nodes in  $V$ .
- (2) Let  $m = \lfloor N/(kI) \rfloor$ . Take a uniform partition of  $\mathcal{N}$  into  $m$  sets  $(\mathcal{V}_j)_{1 \leq j \leq m}$  of cardinality  $kI$  and one set of cardinality smaller than  $kI$ . Likewise, take a uniform partition of  $\mathcal{A}_0$  into  $m$  sets  $(\mathcal{A}_0^{(j)})_{1 \leq j \leq m}$  with cardinality in  $\{\lfloor 4N'/m \rfloor, \lceil 4N'/m \rceil\}$ .

(3) **For**  $j = 1, \dots, m$  **and**  $i = 1, \dots, I$ , **do**

**For each**  $x \in \mathcal{A}_{i-1}^{(j)}$ , **do**

- (i) Sample  $k$  nodes  $(y_{k(i-1)+a}^x)_{a=1, \dots, k}$  uniformly at random in  $\mathcal{V}_j \setminus \{y_a^x\}_{a=1, \dots, k(i-1)}$ .
- (ii) Sample pairs  $(\{x, y_{k(i-1)+a}^x\})_{a=1, \dots, k}$  and compute

$$\hat{p}_{x,i} = \frac{1}{ki} \sum_{a=1}^{ki} A_{xy_a^x}.$$

(iii) Select  $\mathcal{A}_i^{(j)} = \{x \in \mathcal{A}_{i-1}^{(j)} : \hat{p}_{x,i} \geq \nu\}$ .

(4) Set  $\mathcal{N}'$  a set of  $N'$  nodes sampled uniformly at random from  $\bigcup_{1 \leq j \leq m} \mathcal{A}_I^{(j)}$ .

**In case**<sup>2</sup>  $|\bigcup_{1 \leq j \leq m} \mathcal{A}_I^{(j)}| < N'$ , **then** sample at random  $N'$  nodes in  $\mathcal{A}_0$ .

(5) Set  $V' = V \setminus \mathcal{A}_0$ .

**Return** ( $\mathcal{N}', V'$ ).

#### 4.4. Pathwise sparse sampling algorithm

The algorithm presented above fulfils the sparse sampling condition (SpS) at time horizon  $T$ . In many practical situations, it is more natural to consider condition (SpS)

<sup>2</sup>with high probability, this undesirable case does not happen

at all times  $t = 1, 2, \dots$  rather than only at a predefined time horizon  $t = T$ . Formally, condition (SpS) would be replaced by  $N_a(\psi, t) \leq B_t$ , for all  $t = 1, 2, \dots$ . It is possible to modify the previous algorithm to build a strategy  $\psi$  such that, when  $B_t = \Theta(t^\gamma \log^{-\tau}(t))$ , the sampling regret  $\mathbb{E}_\mu[N^{\text{bad}}(\psi, t)]$  fulfils

$$\mathbb{E}_\mu[N^{\text{bad}}(\psi, t)] = O\left(\frac{\sqrt{t} \vee (t/B_t)}{s} \wedge t\right) \quad \text{for } t = 1, 2, \dots$$

Assume that there exist  $\gamma \in (0, 1/2]$  and  $\tau \in [0, +\infty)$  such that  $B_t = t^\gamma / (\log t)^\tau$ , so  $\sqrt{t} \vee (t/B_t) = t^{1-\gamma} \log^\tau(t)$ . In this case, a pathwise sampling condition can be enforced using the simple doubling trick. For any positive integer  $l$ , let  $t_l = 2^l$ . At each time  $t_l$ , the new algorithm discards all nodes and pairs previously sampled and starts the algorithm of Section 4.2 with the remaining nodes, time horizon  $T = t_{l+1} - t_l$  and terminal sparse sampling constraint  $N_a(\psi, t_{l+1} - t_l) \leq \min_{t_l \leq t \leq t_{l+1}} B_t$ . The resulting strategy does not depend on any time horizon and it fulfils the condition  $N_a(\psi, t) \leq B_t$  for all  $t = 1, 2, \dots$ .

Moreover, for any  $l$  such that  $t_l \geq e^{\tau/\gamma}$ ,  $\min_{t_l \leq t \leq t_{l+1}} B_t = B_{t_l}$ . Hence, for any  $l$  such that  $t_l < c_1 n (B_t \wedge n)$  and for any  $t$  such that  $t_{l-1} \leq t \leq t_l < c_1 n (B_t \wedge n)$ ,

$$\begin{aligned} \mathbb{E}[N^{\text{bad}}(\psi, t)] &= O\left(1 + \sum_{k=1}^l \frac{(t_k - t_{k-1})^{1-\gamma} \log^\tau(t_k - t_{k-1})}{s} \wedge (t_k - t_{k-1})\right) \\ &= O\left(\left(\frac{1}{s} \sum_{r=0}^{l-1} 2^{r(1-\gamma)} (r \log(2))^\tau\right) \wedge t_l\right) \\ &= O\left(\frac{t_l^{1-\gamma} \log^\tau(t_l)}{s} \wedge t_l\right) = O\left(\frac{t^{1-\gamma} \log^\tau(t)}{s} \wedge t\right). \end{aligned}$$

According to Theorem 2, the sampling-regret of the algorithm derived from the doubling trick is then rate optimal.

### 5. Discussion

The present paper provides the optimal sampling-regret for pair-matching in the case where  $G = (E, V)$  is a conditional SBM with a number of groups  $K = 2$ , where the groups have  $n/K$  elements, with intra-class probability of connection  $p$  and inter-class  $q$ . Algorithm 1 in Section 3.2 runs in polynomial time and has optimal sampling-regret given in Theorem 1, up to a multiplicative constant. Let us discuss the two following questions: How can we estimate the scaling parameter  $s$ ? How does the rates depend on the number  $K$  of groups?

### 5.1. A heuristic to estimate the scaling parameter $s$

Algorithms 1 and 2 take the scaling parameter  $s$  as input. This parameter is typically unknown in practice and an estimated value  $\hat{s}$  has to be plugged in the algorithm. To guarantee a sampling-regret smaller than  $O(T \wedge (\sqrt{T}/s))$ , the estimator  $\hat{s}$  should use at most  $O(1/s^2)$  edges and satisfy  $\hat{s} \asymp s$  with high probability. The following heuristic builds a possible estimator  $\hat{s}$ .

Pick uniformly at random  $N$  nodes in  $V$  and sample all  $N(N - 1)/2$  pairs between these  $N$  nodes. When  $Ns > 2$ ,  $p = a/N$  and  $q = b/N$ , the results in [38] ensure that, as  $N \rightarrow \infty$ ,  $a$  and  $b$  can be consistently estimated. Therefore,  $Ns = (a - b)^2/(a + b)$  can also be consistently estimated from these  $T = N(N - 1)/2 = O(1/s^2)$  observations. Yet, this estimator requires the knowledge  $Ns$  larger than 2 and cannot therefore be used directly when  $s$  is unknown.

However, when  $p = a/N$  and  $q = b/N$  and  $N \rightarrow \infty$ , it is theoretically possible to detect whether  $Ns = (a - b)^2/(a + b)$  is smaller or larger than 2. To proceed, denote by  $\mathcal{B}$  the non-backtracking matrix associated to the graph; see [5] for a definition of the non-backtracking matrix. Let  $\lambda_1, \lambda_2, \dots$  be the eigenvalues of  $\mathcal{B}$  ranked in decreasing order of their moduli. The main result of [5] shows that, when  $p = a/N$  and  $q = b/N$ , with  $a, b > 0$  fixed, except on an event of vanishing probability as  $N \rightarrow \infty$ ,

$$\begin{aligned} |\lambda_2|^2 &< \lambda_1 && \text{when } Ns < 2, \\ |\lambda_2|^2 &> \lambda_1 && \text{when } Ns > 2. \end{aligned}$$

Hence, in this asymptotic setting where  $s = \Theta(1/N)$ , it is possible to detect if  $Ns > 2$  by looking at the ratio  $|\lambda_2|^2 > \lambda_1$ . In addition, when  $Ns > 2$ , the ratio  $2|\lambda_2|^2/\lambda_1$  consistently estimates  $(a - b)^2/(a + b)$ .

This result suggests the following recursive algorithm to estimate  $s$ : fix some  $\epsilon > 0$  and start with a set  $V_1$  of 2 nodes  $i$  and  $j$  picked uniformly at random in  $V$ . Query the pair  $\{i, j\}$  and let  $E_1$  denote the set of edges in  $E \cap \{i, j\}$ . At each step  $k \geq 2$ , pick at random a set  $V_k$  of  $2^k$  nodes in  $V \setminus \bigcup_{\ell \leq k-1} V_\ell$ . Sample all pairs in  $V_k$ , and denote by  $E_k$  the set of edges among these pairs. Build the non-backtracking matrix  $\mathcal{B}_k$  of the graph  $(V_k, E_k)$  and compute  $\lambda_1^{(k)}$  and  $\lambda_2^{(k)}$  the eigenvalues of this matrix with largest moduli. If  $|\lambda_2^{(k)}|^2 < (1 + \epsilon)\lambda_1^{(k)}$  iterate. If  $|\lambda_2^{(k)}|^2 > (1 + \epsilon)\lambda_1^{(k)}$  stop, denote by  $\hat{k}$  the stopping iteration time and  $\hat{N} = 2^{\hat{k}}$  the number of nodes sampled in the last graph  $(V_{\hat{k}}, E_{\hat{k}})$ . Output  $\hat{s} = 2|\lambda_2^{(\hat{k})}|^2/(\hat{N}\lambda_1^{(\hat{k})})$ .

Assume that  $p = a/N$  and  $q = b/N$  with  $a, b \in \mathbb{R}^+$  fulfilling

$$\frac{(a - b)^2}{a + b} > 2.$$

Let  $\Omega_N$  denote the event where simultaneously  $2 \leq \hat{N}s \leq 8(1 + \epsilon)$  and  $s/2 \leq \hat{s} \leq 2s$ . Then the results of [5] suggest that the event  $\Omega_N$  holds with probability tending to 1 as  $N \rightarrow \infty$ . In addition, the total number of sampled edges is

$$\sum_{k=1}^{\hat{k}} \binom{2^k}{2} = O(\hat{N}^2) = O(1/s^2)$$

on this event. While the results of [5] suggest that the procedure should work for vanishingly small  $s$  (large  $N$  limit), we emphasize that they only hold in a setting where  $p = a/N$ ,  $q = b/N$ , with  $a, b$  fixed and  $N \rightarrow \infty$ , and we cannot turn them into a theoretical guarantee that  $\Omega_N$  holds with probability close to 1. We evaluate the performance of this heuristic numerically in Section 6.3.

### 5.2. Case with $K > 2$ groups

Let us discuss the case where the number of groups  $K$  is larger than 2, still assuming that all the groups have  $n/K$  elements, with intra-class probability of connection  $p$  and inter-class  $q$ . Contrary to  $K = 2$ , we expect in this case an information-computation gap and conjecture the following optimal rates for pair-matching.

**Conjecture 1.** Define  $\Psi_\infty^{\text{poly}}$  as the intersection of  $\Psi_\infty$  defined in Section 3.1, with polynomial-time algorithms. Let

$$s_K = \frac{(p - q)^2}{q + (p - q)/K}.$$

Under Assumption (2.1) and  $s_K \leq 1$ , without computational constraint:

$$\inf_{\psi \in \Psi_\infty} \mathbb{E}[N^{\text{bad}}(\psi, T)] \asymp \left( \left( \frac{K \log(K)}{s_K} \right)^2 \vee \frac{K\sqrt{T}}{s_K} \right) \wedge T. \tag{5.1}$$

With polynomial time constraint:

$$\inf_{\psi \in \Psi_\infty^{\text{poly}}} \mathbb{E}[N^{\text{bad}}(\psi, T)] \asymp \left( \left( \frac{K^2}{s_K} \right)^2 \vee \frac{K\sqrt{T}}{s_K} \right) \wedge T. \tag{5.2}$$

Let us explain the heuristics leading to these rates.

For  $K = 2$ , a central tool to design the rate-optimal polynomial-time algorithm (Algorithm 1) is the existence of polynomial-time algorithms (called GOODCLUST) achieving non-trivial classification for a cSBM( $N/2, N/2, p, q$ ) when  $Ns$  is larger than some constant. When  $K > 2$  and the number of nodes  $N \rightarrow \infty$ , for  $p, q$  scaling as  $1/N$ , the papers [3, 5, 48] provide polynomial-time algorithms GOODCLUST $_K^{\text{poly}}$  achieving a non-trivial classification for

$$Ns_K > K^2 =: \lambda_K^{\text{poly}}.$$

Furthermore, it is conjectured in [14] that there does not exist any polynomial-time algorithm achieving non-trivial classification when  $Ns_K < K^2$ . The threshold  $\lambda_K^{\text{poly}}$  is known as the Kesten–Stigum (KS) threshold. While the conjecture of [14] is relative to the case where  $K$  is fixed,  $p = a/N$ ,  $q = b/N$ , and  $N$  goes to infinity, this conjecture has been recently supported non-asymptotically by a Low-Degree polynomial lower bound [33,44] whenever  $K^2 \leq N$ . This supports the conjecture that, when  $s_K \leq 1$ , non-trivial clustering is possible only when  $N = \Theta(K^2/s_K)$ .

The information theoretic threshold  $\lambda_K^{\text{inf}}$  for non-trivial classification is below  $\lambda_K^{\text{poly}}$  for  $K \geq 5$ . Actually, the paper [4] proved that

$$\lambda_K^{\text{inf}} \asymp K \log(K) \quad \text{and} \quad \lambda_K^{\text{inf}} < \lambda_K^{\text{poly}}$$

for  $K \geq 5$ , so, if the conjecture of [14] holds, there is an information-computation gap for  $K \geq 5$ . A consequence of the result of [4] is that there exist algorithms  $\text{GOODCLUST}_K^{\text{inf}}$ , with exponential complexity, achieving non-trivial classification for  $Ns_K = O(K \log(K))$ .

Theorem 1 requires that  $\text{GOODCLUST}$  has more than non-trivial classification, it should have vanishing classification error. Several papers have established, under assumption (2.1), the existence of algorithms  $\text{GOODCLUST}_K^{\text{poly}}$  and  $\text{GOODCLUST}_K^{\text{inf}}$  with misclassification proportion smaller than  $\exp(-cNs_K/K)$ , for some positive constant  $c$ . This result is obtained for  $Ns_K \geq c'\lambda_K^{\text{poly}}$  for  $\text{GOODCLUST}_K^{\text{poly}}$ ; see, e.g., [10, 16, 17, 19] and for  $Ns_K \gg \lambda_K^{\text{inf}}$  for  $\text{GOODCLUST}_K^{\text{inf}}$ ; see [57].

As a consequence, without computational constraint, a linear sampling regret is expected for any algorithm as long as the time horizon satisfies  $\sqrt{2T}s_K < \lambda_K^{\text{inf}}$ , or equivalently

$$T < 0.5(\lambda_K^{\text{inf}}/s_K)^2 = 0.5(K \log K/s_K)^2.$$

On the other hand, when  $T \gg (K(\log K)^2/s_K)^2$ , one can choose  $N$  fulfilling

$$\lambda_K^{\text{inf}}/s_K \ll N \leq (K\sqrt{T}/s_K)^{1/2} \ll \sqrt{T}.$$

Selecting  $N$  nodes uniformly at random and observing all pairs of these  $N$  nodes,  $\text{GOODCLUST}_K^{\text{inf}}$  classifies correctly the  $N$  nodes, but a proportion of them no greater than  $\exp(-cNs_K/K)$ . The sampling-regret for this step does not exceed the number  $O(N^2) = O(K\sqrt{T}/s_K)$  of pairs sampled. Since  $Ns_K/K \gg \log(K)$ , the proportion of misclassified nodes among these  $N$  nodes is small and a screening procedure as in Step 2 of Algorithm 1 can be applied in order to classify correctly  $\sqrt{T}$  nodes. As an average of  $K/s_K$  queries is necessary to classify one new node, this step will have a regret scaling as  $K\sqrt{T}/s_K$ . Then, we can pair all nodes of the same group until the budget of  $T$  queries is spent. Hence, in the regime where  $T \gg (K(\log K)^2/s_K)^2$ , the final regret should be proportional to  $N^2 + K\sqrt{T}/s_K \asymp K\sqrt{T}/s_K$ . To sum-up

the discussion, without computational constraints, one can expect a sampling-regret of order

$$((K \log(K)/s_K)^2 \vee K \sqrt{T}/s_K) \wedge T,$$

which is the conjectured rate (5.1).

Using polynomial time algorithms for clustering, the information-theoretic threshold  $\lambda_K^{\text{inf}}$  should be replaced by the KS-threshold  $\lambda_K^{\text{poly}}$ . Following the same reasoning as before, linear regret is expected as long as

$$T < 0.5(\lambda_K^{\text{poly}}/s_K)^2 = 0.5(K^2/s_K)^2.$$

On the other hand, when  $\sqrt{T} \gg K^3/s_K$ , one can pick  $N$  nodes at random with  $N$  fulfilling

$$\lambda_K^{\text{poly}}/s_K \ll N \leq (K \sqrt{T}/s_K)^{1/2} \ll \sqrt{T}.$$

A polynomial time algorithm  $\text{GOODCLUST}_K^{\text{poly}}$  run with all pairs based on these nodes classifies correctly these  $N$  nodes, except for a proportion at most  $\exp(-cNs_K/K)$  of them. The sampling-regret associated to this classification step is smaller than  $N^2 \leq K \sqrt{T}/s_K$ . The screening step classifies correctly  $\sqrt{T}$  nodes with a regret  $K \sqrt{T}/s_K$ . The remaining budget until sampling  $T$  pairs is spent by pairing together nodes in a same estimated group. Ultimately, taking into account the computational constraint, one can expect a sampling-regret of order  $((K^2/s_K)^2 \vee K \sqrt{T}/s_K) \wedge T$ , which is the conjectured rate (5.2).

### 5.3. Unbalanced partitions and pairwise dependent probabilities

Although we use a simple random graph model for the ease of exposition, our analysis can be extended to a more general set-up which relaxes the following assumptions:

- (i) the graph has only two groups;
- (ii) the two groups have the same number  $n/2$  of nodes;
- (iii) the probabilities  $p$  and  $q$  of (intra- and inter-group) connection are constants.

For the relaxation of (i), we refer the reader to the Section 5.2, where we discuss conjectures in a SBM with  $K$  groups, for any integer  $K \geq 2$ .

We can relax assumption (ii) by assuming that  $|G_1| = cn$  and  $|G_2| = (1-c)n$  with  $c$  a numerical constant in  $(0, 1)$ . Among the ten points in Lemmas 8 and 9, one can readily check that only point (5) – about the estimation of  $(p+q)/2$  – does not hold anymore. Indeed, for  $c \neq 1/2$ , the global connectivity of the graph is not equal to  $(p+q)/2$ , and thus our estimator  $\hat{\tau}$  of the global connectivity is not a consistent estimator of  $(p+q)/2$ . A solution is to replace  $\hat{\tau}$  by  $\hat{\tau}' = (\hat{p} + \hat{q})/2$ , where the estimators  $\hat{p}$  and  $\hat{q}$  of  $p$  and  $q$  are obtained from the following two-step procedure.

One run GOODCLUST on a first set of sampled pairs in the core-set; then conditionally to the estimated group labels  $\hat{Z}_x$ , one estimate  $p$  by sampling new intra-group pairs (i.e., pairs  $(x, y)$  satisfying  $\hat{Z}_x = \hat{Z}_y$ ) and one estimate  $q$  by sampling new inter-group pairs ( $\hat{Z}_x \neq \hat{Z}_y$ ).

Instead of the model assumption (iii), let us consider the situation where the probability  $p_{ij}$  of connection between the nodes  $i$  and  $j$  belongs to  $[p_0, p]$  if  $i$  and  $j$  are in the same group, while  $p_{ij} \in [q, q_0]$  otherwise, with  $q_0 < p_0$ . Turnkey clustering algorithms work in such situations; see [57], for example. Then, still with the idea of having two distinct groups of nodes, assume that the intra-group variations  $\delta := (p - p_0) \vee (q_0 - q)$  are smaller than the inter group separation  $\Delta := p_0 - q_0$ . Under this assumption, one can readily check that our estimator  $\hat{\tau}$  of the global connectivity satisfies

$$|\hat{\tau} - (p_0 + q_0)/2| \leq \Delta/4$$

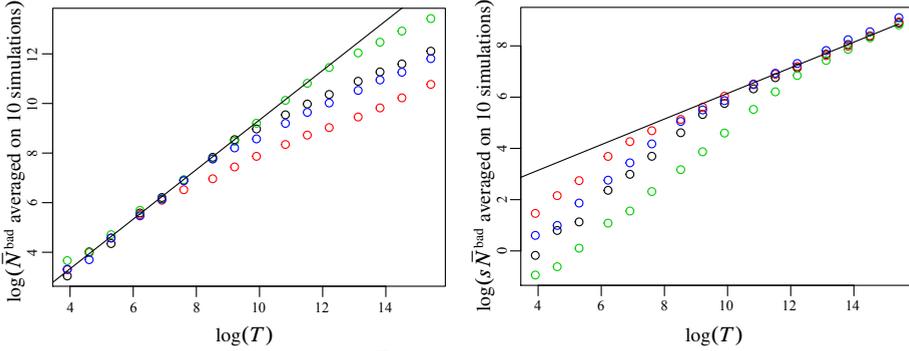
(i.e., Lemma 8(5)), and that all the other points in Lemmas 8 and 9 still hold. In this situation of non-constant probabilities of connection, note that the link between the regret and the sampling regret (seen in Section 2.4) does not hold anymore, and we only control the sampling regret.

## 6. Numerical experiments

### 6.1. Unconstrained setting

In this section, we empirically assess the sharpness of our unconstrained algorithm from Section 3.2 on the following parameters:

- the budget  $T$  is taken in  $\{50, 100, 200, 500, 1\,000, 2\,000, 5\,000, 10\,000, 20\,000, 50\,000, 100\,000, 200\,000, 500\,000, 1\,000\,000, 2\,000\,000, 5\,000\,000\}$ ;
- the connection probabilities are taken as  $(p, q) = (0.6, 0.4), (0.7, 0.3), (0.55, 0.45)$  or  $(0.4, 0.2)$ , which corresponds to  $s = 0.04, 0.16, 0.01$ , or  $0.0666\dots$  respectively;
- the groups are either balanced, or one group contains 20% of the individuals while the other contains 80%;
- the pool of available individuals is assumed infinite (i.e.,  $n$  infinite), which means that one may freely add new individuals and decide their class independently of the classes of previously sampled individuals;
- for each choice of parameter, ten experiments are performed to obtain an averaged regret.

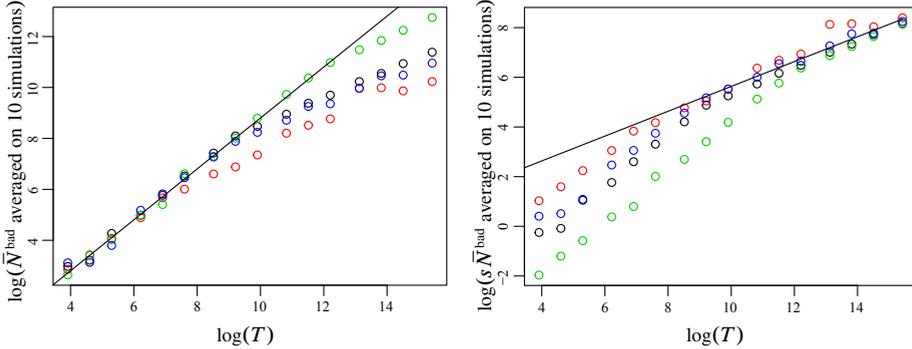


**Figure 1.** Average number of errors  $\bar{N}^{\text{bad}}$  over ten simulations in the unconstrained case with balanced communities. The graphs show  $\log(\bar{N}^{\text{bad}})$  (left) and  $\log(s\bar{N}^{\text{bad}})$  (right) as a function of  $\log(T)$ , confirming the two regimes (linear in  $T$  and proportional to  $\sqrt{T}/s$ ). The lines have slope 1 and 1/2, respectively. Green:  $s = 0.01$ , black:  $s = 0.04$ , blue:  $s = 0.06666\dots$ , red:  $s = 0.16$ .

A few adaptations are required for the implementation to work efficiently and for all values of  $T$ . For the unconstrained algorithm,

- if  $\log(s\sqrt{T}) < 1$ , it is replaced by 1 in the formulas of the size  $N$  of the core-set and of the number  $I$  of steps in Step 2;
- if the size of the core-set  $N$  is smaller than the size  $2kI$  required to have at least one estimated class large enough to perform Step 2,  $N$  is replaced by  $2kI$  (more generally  $KkI$  when there are  $K$  groups);
- GOODCLUST classifies the individuals by applying the Lloyd algorithm to the first two eigenvectors of the trimmed-adjacency matrix of the graph (more generally to the first  $K$  eigenvectors), see [10];
- the values of the constants are fixed to  $c_0 = 2$ ,  $C_I = 10$ ,  $C_k = 0.1$ , and the number of individuals in the core-set is taken as  $N = \lceil c_{\mathcal{N}} \sqrt{T} / \log(s\sqrt{T}) \rceil$  with  $c_{\mathcal{N}} = 2$ ;
- instead of sampling a set  $\mathcal{A}_0$  in Step 2 and pruning the individuals with low connectivity, we instead proceed individual by individual, discarding them if their connectivity dips below the threshold at one point and adding them to  $\mathcal{A}_I$  if not.

These modifications are intended to make the algorithm work more efficiently and for smaller values of  $T$  than the ones presented in the theorems above, without changing its theoretical properties.



**Figure 2.** Average number of errors  $\bar{N}^{\text{bad}}$  over ten simulations in the unconstrained case with a 80%–20% split between communities. The graphs show  $\log(\bar{N}^{\text{bad}})$  (left) and  $\log(s\bar{N}^{\text{bad}})$  (right) as a function of  $\log(T)$ . The lines have slope 1 and 1/2 respectively. Green:  $s = 0.01$ , black:  $s = 0.04$ , blue:  $s = 0.06666\dots$ , red:  $s = 0.16$ .

The results for the unconstrained algorithm are displayed in Figure 1 (balanced communities) and Figure 2 (unbalanced communities). In both cases, as predicted by Theorem 1, we observe an initial linear regime for the regret for  $T = O(1/s^2)$ , followed by a  $\sqrt{T}/s$  regime for  $T$  large compared to  $1/s^2$ .

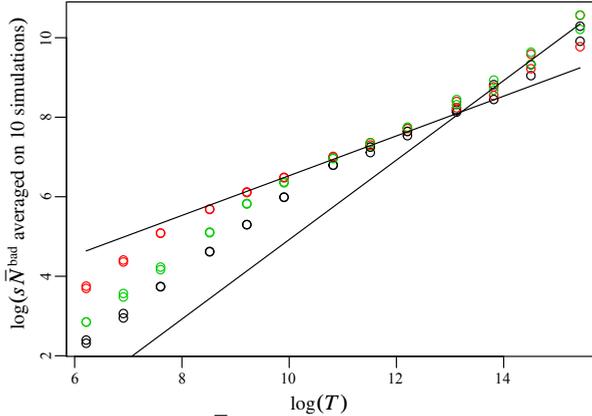
**6.2. Constrained setting**

In this section, the constrained algorithm from Section 4.2 is assessed, the results are displayed in Figure 3. The parameters are the same than in the previous section, with the difference that we only consider balanced communities and budgets  $T$  greater or equal to 500. The constraint  $B_T$  is taken in  $\{500, 1\ 000\}$ .

In addition to the adaptations from the previous section, we modified the algorithm as follows:

- if  $\log(s\sqrt{T}) < 1$ , since this corresponds to the linear case, the edges are sampled at random;
- the values of the constants are fixed to  $c_0 = 8$ ,  $C_I = 10$ ,  $C_k = 0.1$ , and the number of individuals in the core-set is taken as  $N = \lceil c_{\mathcal{N}}\sqrt{T}/\log(s\sqrt{T}) \rceil$  with  $c_{\mathcal{N}} = 4$ ;
- if the size of the core-set is smaller than  $5/2 \cdot c_0/s = 20/s$ , which ensures that the constraint is satisfied with high probability,  $N$  is replaced by  $20/s$ .

Again, we observe empirically the three regimes from Theorem 2: the sampling regret is first linear in  $T$ , then proportional to  $\sqrt{T}/s$  and then linear again as  $T/(Bs)$ .



**Figure 3.** Average number of errors  $\bar{N}^{\text{bad}}$  over ten simulations in the constrained case with balanced communities. The graph shows  $\log(s\bar{N}^{\text{bad}})$  as a function of  $\log(T)$ . The lines have slope  $1/2$  and  $1$ , respectively. This confirms the three regimes of Theorem 2: linear in  $T$  for small and large  $T$  and proportional to  $\sqrt{T}/s$  in between. Black:  $s = 0.04$ , green:  $s = 0.06666\dots$ , red:  $s = 0.16$ .

### 6.3. Estimation of the scaling parameter $s$

In this section, we investigate the empirical performance of the heuristic procedure from Section 5.1 for estimating  $s$ . As  $s \rightarrow 0$ , the two key properties that we are looking for, are

$$2 < \hat{N}s = O(1) \quad \text{and} \quad \hat{s} = \Theta(s).$$

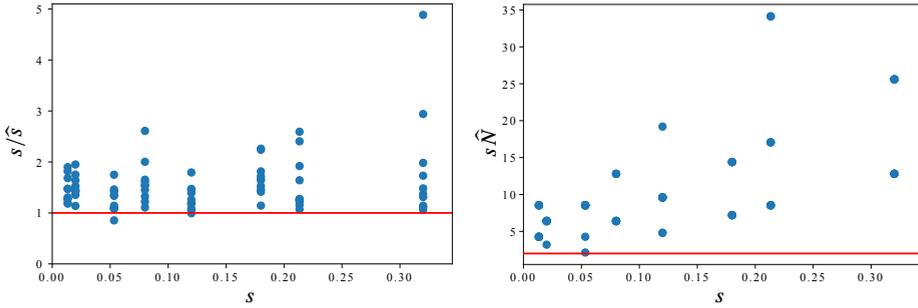
We focus on the empirical evaluation of these two properties.

For non-vanishing values of  $s$ , random fluctuations of the eigenvalues of the non-backtracking matrix can blur the asymptotic results of [5]. We evaluate the noise level with the modulus of the third eigenvalue  $|\lambda_3^{(k)}|$ , whose asymptotic value remains smaller than  $\sqrt{\lambda_1^{(k)}}$  when  $s$  is vanishingly small. We changed the stopping criterion accordingly: the algorithm stops when

$$|\lambda_3^{(k)}|^2 < (1 + \epsilon)\lambda_1^{(k)} \quad \text{and} \quad |\lambda_2^{(k)}|^2 \geq (1 + \epsilon)\lambda_1^{(k)},$$

with  $\epsilon = 0.1$ . The additional condition  $|\lambda_3^{(k)}|^2 < (1 + \epsilon)\lambda_1^{(k)}$  is meant to ensure that the stopping condition  $|\lambda_2^{(k)}|^2 \geq (1 + \epsilon)\lambda_1^{(k)}$  is not due to pure random fluctuations. This modification does not change the asymptotic properties of the estimator or the asymptotic of the number of edges sampled, that is  $O(1/s^2)$  when  $s \rightarrow 0$ .

The parameters taken were balanced groups with  $(p, q) = (0.45, 0.05), (0.4, 0.1), (0.35, 0.15), (0.3, 0.2), (0.3, 0.1/3), (0.8/3, 0.2/3), (0.7/3, 0.1),$  or  $(0.2, 0.4/3)$ . For each value of  $(p, q)$ , the ratio  $s/\hat{s}$  and the product  $\hat{N}s$  are displayed in Figure 4 for ten simulations.



**Figure 4.** On the left, ratio of the true scaling parameter  $s$  to the estimated scaling parameter  $\hat{s}$  as defined in Section 5.1. The red line corresponds to  $s/\hat{s} = 1$ . On the right, product  $s\hat{N}$ , where  $\hat{N}$  is the number of nodes at the stopping time. The red line corresponds to  $s\hat{N} = 2$ .

We used the SUNBEAM package from [52], available at [51], to compute the eigenvalues of the non-backtracking matrix. The results are displayed in Figure 4. The left-hand side figure shows that the estimator  $\hat{s}$  is always comparable to  $s$ , and almost always smaller than  $s$ , which makes it a good choice to replace  $s$  by  $\hat{s}$  in the pair-matching algorithm. Finally, the right-hand side figure shows that the condition  $s\hat{N} > 2$  is always satisfied, thus ensuring that  $\hat{s}$  consistently estimates  $s$ .

### 6.4. Resilience to model misspecification

The above simulations use data generated by a cSBM model. We now investigate how the algorithm’s performance is affected when the data does not exactly follow a cSBM distribution.

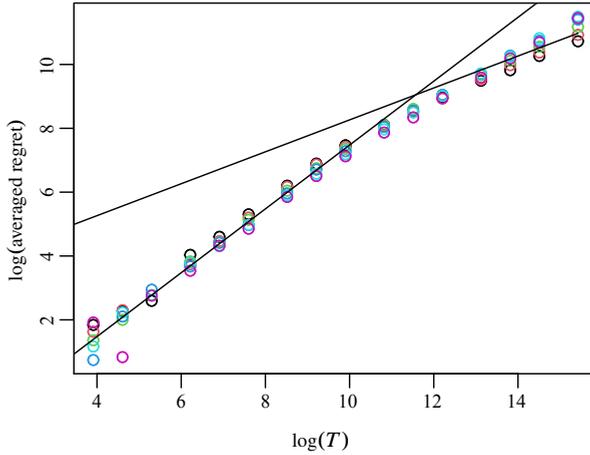
We consider the following alternative model. Let  $0 < p, q < 1$  be connectivity parameters as above, as well as a misspecification parameter  $\sigma \in [0, 1/2)$ . When sampling a new node  $i$ , it is affected the label  $Z_i = 0$  or  $1$  with probability  $1/2$ . In addition to its label, it is assigned a hidden state  $S_i \sim \text{Unif}([Z_i - \sigma, Z_i + \sigma])$ . The probability that two nodes  $i$  and  $j$  are connected is given by

$$\mathbb{P}(A_{ij} = 1 \mid Z_i, Z_j, S_i, S_j) = p + (q - p) \left( |S_i - S_j| - \frac{2\sigma}{3} \right).$$

The case  $\sigma = 0$  is the usual, well-specified, case. In general, for any  $\sigma$ ,

$$\mathbb{P}(A_{ij} = 1 \mid Z_i = Z_j) = p \quad \text{and} \quad \mathbb{P}(A_{ij} = 1 \mid Z_i \neq Z_j) = q,$$

that is, the expected connectivity between members of the same class (respectively, between members of different classes) is the same as in the well-specified setting. Where the distribution differs is in the joint distribution of more than two nodes: if



**Figure 5.** Logarithm of the average regret  $\frac{1}{10} \sum_{k=1}^{10} (pT - \sum_{t=1}^T A_{\hat{e}_t}^{(k)})$  in the misspecified case, as a function of  $\log(T)$ . The lines have slope  $1/2$  and  $1$  respectively. Black:  $\sigma = 0$ , red:  $\sigma = 0.1$ , green:  $\sigma = 0.2$ , blue:  $\sigma = 0.3$ , cyan:  $\sigma = 0.4$ , violet:  $\sigma = 0.5$ . The two regimes from the well-specified case (linear then proportional to  $\sqrt{T}$ ) are still visible.

two nodes are connected, chances are good their hidden states are closer than average, thus making it more likely that if a new node connects to one of them, then it will connect to the other.

We take the same possible values of  $T$  as in Section 6.1, with balanced communities,  $(p, q) = (0.6, 0.4)$ , and an infinite pool of individuals and misspecification values  $\sigma \in \{0, 0.1, 0.2, 0.3, 0.4\}$ . Ten experiments are performed for each choice of parameters, and the average regrets computed over them, are displayed in Figure 5. The two regimes from the well-specified case (linear then proportional to  $\sqrt{T}$ ) are still visible.

## A. Proof of the lower bounds

### A.1. Distributional properties under assumption (II)

Recall that  $\mathcal{E}$  denotes the set of all pairs in  $\{1, \dots, n\}$ . The invariance to labelling property enforces some invariances on the distribution of the  $(N_e(\psi, T) : e \in \mathcal{E})$ , with  $N_e(\psi, T)$  defined by (2.3) and on the distribution of the  $(N_a(\psi, T) : a = 1, \dots, n)$  with  $N_a(\psi, T)$  defined by (2.4).

Let  $\mu$  be a distribution in  $\text{cSBM}(n/2, n/2, p, q)$  associated to a partition  $G = \{G_1, G_2\}$  of  $\{1, \dots, n\}$ . Consider a permutation  $\sigma$  which leaves the partition  $G$  invariant, that is such that, either  $\sigma(G_1) = G_1$  and hence  $\sigma(G_2) = G_2$ , or  $\sigma(G_1) = G_2$  and thus  $\sigma(G_2) = G_1$ . Then, the distribution  $\mu^\sigma$  defined in Section 2.3 is equal to

the distribution  $\mu$ . Hence the invariance to labelling property ensures that for any permutation  $\sigma$  leaving  $G$  invariant, the vectors  $(N_e(\psi, t) : e \in \mathcal{E}; t = 1, \dots, \binom{n}{2})$  and  $(N_{\sigma(e)}(\psi, T) : e \in \mathcal{E}; t = 1, \dots, \binom{n}{2})$  have the same distribution. As a consequence, the following property holds.

**Lemma 3.** *When the strategy  $\psi$  fulfils the invariance to labelling property, then the random variables  $(N_e(\psi, T) : e \in \mathcal{E}^{\text{good}})$  are pairwise exchangeable. The same property holds for  $(N_e(\psi, T) : e \in \mathcal{E}^{\text{bad}})$  and  $(N_a(\psi, T) : a = 1, \dots, n)$ .*

*Proof.* Let  $\{a, b\}, \{a', b'\}$  denote two pairs in  $\mathcal{E}^{\text{good}}$  and let  $\sigma$  be a  $G$ -invariant permutation such that  $\sigma(\{a, b\}) = \{a', b'\}$  and  $\sigma(\{a', b'\}) = \{a, b\}$ . Since  $\mu = \mu^\sigma$  and  $\psi$  is invariant to labelling, the random variables  $(N_{\{a,b\}}, N_{\{a',b'\}})$  and  $(N_{\{a',b'\}}, N_{\{a,b\}})$  have the same distribution. The same reasoning applies for pairs in  $\mathcal{E}^{\text{bad}}$ .

Consider now two nodes  $a, b \in \{1, \dots, n\}$ . Let  $\sigma$  be a  $G$ -invariant permutation on  $\{1, \dots, n\}$  such that  $\sigma(a) = b$  and  $\sigma(b) = a$ . Since  $\mu = \mu^\sigma$  and  $\psi$  is invariant to labelling, the random variables  $(N_a(\psi, T), N_b(\psi, T))$  and  $(N_b(\psi, T), N_a(\psi, T))$  have the same distribution. ■

**A.2. Proof of the lower bound in Theorems 2 and 1**

This section contains the proof of the first part of Theorem 2. The first part of Theorem 1 follows by taking  $B_T = T$ .

Let  $kl(p, q) = p \log(p/q) + (1 - p) \log((1 - p)/(1 - q))$  be the Kullback–Leibler divergence between two Bernoulli distributions with parameters  $p$  and  $q$ . We actually prove the following stronger lower bound: when  $\tilde{s} = kl(p, q) \vee kl(q, p)$  satisfies  $\tilde{s} \leq 1/16$ , for any  $\mu \in \text{cSBM}(n/2, n/2, p, q)$ ,

$$\inf_{\psi \in \Psi_{B_T, T}} \mathbb{E}_\mu [N^{\text{bad}}(\psi, T)] \geq \frac{1}{32} \left[ \frac{\sqrt{T} \vee (T/B_T)}{16\tilde{s}} \wedge T \right]. \tag{A.1}$$

The first part of Theorem 2 follows from this bound and from Lemma 19 which ensures that  $s \leq \tilde{s} \leq 2(1 + \rho^*)s$  when (2.1) holds.

Recall that  $N_a(\psi, T)$  denotes the number of pairs involving the node  $a$  sampled by the strategy  $\psi$  up to time  $T$ . Let  $N_a^{\text{bad}}(\psi, T)$  be the number of pairs  $\{a, b\}$  with  $b$  not in the community of  $a$  sampled up to time  $T$ . Hereafter in the proof, the strategy  $\psi$  is fixed and, to simplify notations, the dependency of  $N_a$  and  $N_a^{\text{bad}}$  on  $\psi$  is dropped out:  $N_a^{\text{bad}}(\psi, T)$  is denoted  $N_a(T)$  and  $N_a(\psi, T)$  is denoted  $N_a^{\text{bad}}(T)$ . Let also

$$N_a^{\text{good}}(T) = N_a(T) - N_a^{\text{bad}}(T).$$

The number of between-group sampled pairs is

$$N^{\text{bad}}(T) = \frac{1}{2} \sum_{a=1}^n N_a^{\text{bad}}(T).$$

Let us also recall that  $N_{\{a,b\}}(\psi, T) \in \{0, 1\}$  (denoted  $N_{\{a,b\}}(T)$ ), is the number of times the pair  $\{a, b\}$  has been sampled before time  $T$ . Likewise, let

$$N_{aB}(T) = \sum_{b \in B} N_{\{a,b\}}(T)$$

be the number of times a pair between node  $a$  and the set of nodes  $B$  has been sampled before time  $T$ . For  $t \geq 0$ , let  $\mathcal{F}_t$  be the  $\sigma$ -algebra gathering information available up to time  $t$ :  $\mathcal{F}_t$  is the  $\sigma$ -algebra generated by  $(\hat{\mathcal{E}}_t, (A_e)_{e \in \hat{\mathcal{E}}_t}, U_0, \dots, U_t)$ .

The main tools for proving equation (A.1) are given by the next two lemmas. The first lemma is directly adapted from [18, 28]. It is a derivative of the data processing inequality, and it merely states that the Kullback–Leibler divergence between two distributions of observation stopped at some stopping time is larger than the Kullback–Leibler divergence of processed versions of these distributions.

**Lemma 4.** *Let  $\tilde{T}$  be a stopping time with respect to the filtration  $(\mathcal{F}_t)_{t \geq 0}$ . Let  $\mu, \mu' \in \text{cSBM}(n/2, n/2, p, q)$  and let  $v = (v_{ab})_{a < b}$  and  $v' = (v'_{ab})_{a < b}$  denote their connection probabilities, that is  $v_{ab} = \mu(\{a, b\} \in E)$  and  $v'_{ab} = \mu'(\{a, b\} \in E)$  for all  $a, b \in V$ . If  $\tilde{T} \leq T$  a.s., then for any  $\mathcal{F}_{\tilde{T}}$ -measurable random variable  $Z$  taking values in  $[0, 1]$ ,*

$$\sum_{a < b} \mathbb{E}_\mu [N_{\{a,b\}}(\tilde{T})] kl(v_{ab}, v'_{ab}) \geq kl(\mathbb{E}_\mu[Z], \mathbb{E}_{\mu'}[Z]), \tag{A.2}$$

where  $kl(p, q) = p \log(p/q) + (1 - p) \log((1 - p)/(1 - q))$  is the Kullback–Leibler divergence between two Bernoulli distributions with parameters  $p$  and  $q$ .

*Proof.* The lemma follows directly from [28, Lemma 1] and [18, Lemma 1]. As discussed in Section 2.1, the pair-matching problem can be seen as a bandit problem with restrictions on the set of admissible strategies. Since [28, Lemma 1] and [18, Lemma 1] hold for any strategy, inequality (A.2) holds in particular for any strategy  $\psi$  satisfying the constraints  $\psi_t(\hat{\mathcal{E}}_t, \dots) \notin \hat{\mathcal{E}}_t$  and  $N_a(t) \leq B_T$ . ■

While the previous lemma is only based on the bandit nature of the problem, the next lemma is based on the constraint that arms can only be sampled once.

**Lemma 5.** *Let  $M$  be a positive real number and consider  $T \geq 1$ . Then*

$$\sum_{a=1}^n (N_a(T) \wedge M) \geq \left( (M \sqrt{T}) \vee \frac{MT}{B_T} \right) \wedge \frac{T}{2}.$$

*Proof.* Let  $S_1 = \{a : N_a(T) \leq M\}$  and  $S_2 = \{a : N_a(T) > M\}$ . If

$$\sum_{a \in S_1} N_a(T) \geq T/2,$$

then

$$\sum_{a=1}^n (N_a(T) \wedge M) \geq \sum_{a \in S_1} N_a(T) \geq T/2.$$

Assume now that  $\sum_{a \in S_1} N_a(T) < T/2$ . Since  $2T = \sum_{a=1}^n N_a(T)$ , we have

$$\begin{aligned} 2T &\leq T/2 + \sum_{a \in S_2} N_a(T) \\ &= T/2 + \sum_{a \in S_2} N_{aS_1}(T) + \sum_{a \in S_2} N_{aS_2}(T) \\ &= T/2 + \sum_{a \in S_1} N_{aS_2}(T) + \sum_{a \in S_2} N_{aS_2}(T) \\ &\leq T + |S_2|(B_T \wedge |S_2|). \end{aligned}$$

Hence,  $|S_2| \geq \sqrt{T} \vee (T/B_T)$  and

$$\sum_{a=1}^n (N_a(T) \wedge M) \geq |S_2|M \geq (M\sqrt{T}) \vee (MT/B_T).$$

The proof is complete. ■

With these two lemmas, the core inequality of the proof can be established. This inequality shows that if  $N_a(t) = O(1/\tilde{s})$ , then  $N_a^{\text{bad}}(t)$  is of the same order of magnitude than  $N_a(t)$ .

Let  $G = (G_1, G_2)$  be a partition of  $\{1, \dots, n\}$  with  $G_1 = \{1, \dots, n/2\}$  and  $G_2 = \{n/2 + 1, \dots, n\}$ . Let  $\mu \in \text{cSBM}(n/2, n/2, p, q)$  be the distribution of a conditional SBM with classes  $G_1$  and  $G_2$ , within-group connection probability  $p$  and between-group connection probability  $q$ . Unless specified,  $\mathbb{E} = \mathbb{E}_\mu$  in the following.

**Lemma 6.** *Let  $M$  be a positive integer such that  $16M\tilde{s} \leq 1$  and define the stopping time*

$$\tilde{T} = T \wedge \inf\{t : \max(N_1(t), N_n(t)) \geq M\}.$$

*Setting  $N_{1+n}(T) = N_1(T) + N_n(T)$  and  $N_{1+n}^{\text{bad}}(T) = N_{1G_2}(T) + N_{nG_1}(T)$ , we have*

$$\mathbb{E}[N_{1+n}^{\text{bad}}(\tilde{T})] \geq \frac{1}{4} \mathbb{E}[N_{1+n}(\tilde{T})] \geq \frac{1}{4} \mathbb{E}[N_1(T) \wedge M]. \tag{A.3}$$

*Proof.* Some arguments of this proof are inspired by [18]. The last inequality in (A.3) follows directly from

$$N_{1+n}(\tilde{T}) \geq N_1(T)\mathbf{1}_{\tilde{T}=T} + M\mathbf{1}_{\tilde{T}<T} \geq N_1(T) \wedge M.$$

It remains to show the first inequality. Consider the transposition  $\sigma = (1, n)$  of 1 and  $n$  which switches the labels 1 and  $n$  while keeping other nodes unchanged. Let  $\mu^\sigma$  be the distribution of  $(A_{\sigma(a),\sigma(b)})_{ab}$ . The partition  $G^\sigma = \{G_1^\sigma, G_2^\sigma\}$  associated to  $\mu^\sigma$ , corresponds to  $G$  with 1 and  $n$  switched, that is

$$G_1^\sigma = \{n, 2, \dots, n/2\} \quad \text{and} \quad G_2^\sigma = \{n/2 + 1, \dots, n - 1, 1\}.$$

Let  $M$  be a positive integer and set

$$\mathcal{Z} = \frac{N_{1G_2}(\tilde{T}) + N_{nG_1}(\tilde{T})}{2(M \wedge B_T)} \in [0, 1].$$

By invariance to labelling,

$$\begin{aligned} \mathbb{E}_{\mu^\sigma} [N_{1G_2}(\tilde{T}) + N_{nG_1}(\tilde{T})] &= \mathbb{E}_{\mu^\sigma} [N_{1G_2^\sigma}(\tilde{T}) + N_{nG_1^\sigma}(\tilde{T}) + 2N_{\{1,n\}}(\tilde{T})] \\ &= \mathbb{E}_\mu [N_{1G_1}(\tilde{T}) + N_{nG_2}(\tilde{T}) + 2N_{\{1,n\}}(\tilde{T})]. \end{aligned}$$

Hence, setting  $\tilde{M} = M \wedge B_T$ , Lemma 4 ensures that

$$\begin{aligned} &(kl(p, q) \vee kl(q, p)) \mathbb{E}_\mu [N_1(\tilde{T}) + N_n(\tilde{T})] \\ &\geq kl(\mathbb{E}_\mu [N_{1G_2}(\tilde{T}) + N_{nG_1}(\tilde{T})]/(2\tilde{M}), \mathbb{E}_{\mu^\sigma} [N_{1G_2}(\tilde{T}) + N_{nG_1}(\tilde{T})]/(2\tilde{M})) \\ &= kl(\mathbb{E}_\mu [N_{1G_2}(\tilde{T}) + N_{nG_1}(\tilde{T})]/(2\tilde{M}), \\ &\quad \mathbb{E}_\mu [N_{1G_1}(\tilde{T}) + N_{nG_2}(\tilde{T}) + 2N_{\{1,n\}}(\tilde{T})]/(2\tilde{M})) \\ &\geq \frac{1}{2(M \wedge B_T)} \\ &\quad \times \frac{(\mathbb{E}_\mu [N_{1G_2}(\tilde{T}) + N_{nG_1}(\tilde{T})] - \mathbb{E}_\mu [N_{1G_1}(\tilde{T}) + N_{nG_2}(\tilde{T}) + 2N_{\{1,n\}}(\tilde{T})])^2}{\mathbb{E}_\mu [N_{1G_2}(\tilde{T}) + N_{nG_1}(\tilde{T})] \vee \mathbb{E}_\mu [N_{1G_1}(\tilde{T}) + N_{nG_2}(\tilde{T}) + 2N_{\{1,n\}}(\tilde{T})]}, \end{aligned}$$

where the last line follows from Lemma 19. Setting

$$N_{1+n}^{\text{good}}(T) = N_{1G_1}(T) + N_{nG_2}(T),$$

the last inequality can be written as

$$\begin{aligned} &2(M \wedge B_T)\tilde{\mathcal{S}} \mathbb{E}[N_{1+n}(\tilde{T})] (\mathbb{E}[N_{1+n}^{\text{good}}(\tilde{T}) + 2N_{\{1,n\}}(\tilde{T})] \vee \mathbb{E}[N_{1+n}^{\text{bad}}(\tilde{T})]) \\ &\geq (\mathbb{E}[N_{1+n}^{\text{good}}(\tilde{T}) + 2N_{\{1,n\}}(\tilde{T})] - \mathbb{E}[N_{1+n}^{\text{bad}}(\tilde{T})])^2. \end{aligned} \quad (\text{A.4})$$

If  $\mathbb{E}[N_{1+n}^{\text{good}}(\tilde{T})] \leq \mathbb{E}[N_{1+n}^{\text{bad}}(\tilde{T})]$ , then

$$2 \mathbb{E}[N_{1+n}^{\text{bad}}(\tilde{T})] \geq \mathbb{E}[N_{1+n}^{\text{good}}(\tilde{T})] + \mathbb{E}[N_{1+n}^{\text{bad}}(\tilde{T})] = \mathbb{E}[N_{1+n}(\tilde{T})]$$

and Lemma 6 follows.

Assume therefore that  $\mathbb{E}[N_{1+n}^{\text{good}}(\tilde{T})] \geq \mathbb{E}[N_{1+n}^{\text{bad}}(\tilde{T})]$ . It follows that

$$2 \mathbb{E}[N_{1+n}^{\text{good}}(\tilde{T})] \geq \mathbb{E}[N_{1+n}^{\text{good}}(\tilde{T})] + \mathbb{E}[N_{1+n}^{\text{bad}}(\tilde{T})] = \mathbb{E}[N_{1+n}(\tilde{T})],$$

so inequality (A.4) implies

$$\begin{aligned} 4(M \wedge B_T)\tilde{s} \mathbb{E}[N_{1+n}^{\text{good}}(\tilde{T}) + 2N_{\{1,n\}}(\tilde{T})]^2 \\ \geq (\mathbb{E}[N_{1+n}^{\text{good}}(\tilde{T}) + 2N_{\{1,n\}}(\tilde{T})] - \mathbb{E}[N_{1+n}^{\text{bad}}(\tilde{T})])^2. \end{aligned}$$

Rearranging the expression gives

$$\begin{aligned} \mathbb{E}[N_{1+n}^{\text{bad}}(\tilde{T})] &\geq \mathbb{E}[N_{1+n}^{\text{good}}(\tilde{T}) + 2N_{\{1,n\}}(\tilde{T})](1 - \sqrt{4(M \wedge B_T)\tilde{s}}) \\ &\geq \frac{1}{2} \mathbb{E}[N_{1+n}^{\text{good}}(\tilde{T})] \geq \frac{1}{4} \mathbb{E}[N_{1+n}(\tilde{T})], \end{aligned}$$

since  $M \wedge B_T \leq 1/(16\tilde{s})$  by assumption. The proof is complete.  $\blacksquare$

The lower bound in Theorem 2 can now be proved. Recall that for any strategy  $\psi \in \Psi_{B_T, T}$ , Assumption (II) implies that the sampling-regret  $\mathbb{E}_\mu[N^{\text{bad}}(\psi, T)]$  does not depend on  $\mu \in \text{cSBM}(n/2, n/2, p, q)$ ; see Remark 2. Therefore, it is sufficient to prove (A.1) for any strategy  $\psi$  invariant by labelling, with the distribution  $\mu$  defined above Lemma 6.

Let  $M$  be a positive integer such that

$$1 \leq M \wedge B_T \leq \frac{1}{16\tilde{s}}.$$

First, Lemma 3 ensures that, for any pair  $\{a, b\} \in \mathcal{E}^{\text{bad}}$ ,  $\mathbb{E}[N_{\{a,b\}}(T)] = \mathbb{E}[N_{\{1,n\}}(T)]$ , and hence

$$\mathbb{E}[N^{\text{bad}}(T)] = \frac{n^2}{4} \mathbb{E}[N_{\{1,n\}}(T)] = \frac{n}{4} \mathbb{E}[N_{1+n}^{\text{bad}}(T)].$$

Lemma 3 also ensures that  $\mathbb{E}[N_a(T) \wedge M] = \mathbb{E}[N_1(T) \wedge M]$  for all  $a \in \{1, \dots, n\}$ . By Lemma 6, it follows that

$$\begin{aligned} 16 \mathbb{E}[N^{\text{bad}}(T)] &= 4n \mathbb{E}[N_{1+n}^{\text{bad}}(T)] \geq 4n \mathbb{E}[N_{1+n}^{\text{bad}}(\tilde{T})] \\ &\geq n \mathbb{E}[N_1(T) \wedge M] = \sum_{a=1}^n \mathbb{E}[N_a(T) \wedge M]. \end{aligned}$$

Hence, by Lemma 5

$$16 \mathbb{E}[N^{\text{bad}}(T)] \geq \left( (M\sqrt{T}) \vee \frac{MT}{B_T} \right) \wedge \frac{T}{2}.$$

For  $\tilde{s} \leq 1/16$ , taking  $M$  equal to the integer part of  $1/(16\tilde{s})$  gives

$$16 \mathbb{E}[N^{\text{bad}}(T)] \geq \left( \frac{\sqrt{T}}{32\tilde{s}} \vee \frac{T}{32\tilde{s}B_T} \right) \wedge \frac{T}{2}.$$

Since the sampling-regret does not depend on the choice of  $\mu$ , the proof is complete.

## B. Proof of the unconstrained upper bound

This section proves the following result, from which follows the upper bound of Theorem 1, as explained below Theorem 7.

**Theorem 7.** *There exist numerical constants  $c_1, c_2 > 0$ , such that, for any  $T \leq c_2 n^2$ , with probability at least  $1 - 13/T$ , the algorithm described in Section 3.2 fulfils*

$$N^{\text{bad}}(\psi, T) \leq c_1 \left( T \wedge \frac{\sqrt{T}}{s} \right).$$

Let us explain how the upper bound of Theorem 1 follows from Theorem 7. First, let us note that the upper bound of Theorem 7 also holds in expectation. Indeed, since  $N^{\text{bad}}(\psi, T) \leq T$ , the algorithm described in Section 3.2 fulfils

$$\mathbb{E}[N^{\text{bad}}(\psi, T)] \leq c_1 \left( T \wedge \frac{\sqrt{T}}{s} \right) + 13 \leq c'_1 \left( T \wedge \frac{\sqrt{T}}{s} \right). \quad (\text{B.1})$$

Second, we can get an horizon free algorithm by applying a doubling trick. For any integer  $l$ , let  $t_l = 2^l$ . At each time  $t_l$ , discard all nodes and pairs involved in the previous iterations of the algorithm and restart the algorithm described in Section 3.2 with time horizon  $t_{l+1} - t_l$ . The resulting strategy does not depend on any time horizon. Let us prove that this horizon-free algorithm also has a  $O(T \wedge (\sqrt{T}/s))$  sampling regret. The argument for this proof is classical: according to the upper bound (B.1), for any  $t_{l-1} \leq T \leq t_l < c_2 n^2$ ,

$$\begin{aligned} & \mathbb{E}[N^{\text{bad}}(\psi, T)] \\ & \leq c_1 \left( \frac{\sqrt{t_0}}{s} \wedge t_0 + \frac{\sqrt{t_1 - t_0}}{s} \wedge (t_1 - t_0) + \dots + \frac{\sqrt{t_l - t_{l-1}}}{s} \wedge (t_l - t_{l-1}) \right) \\ & \leq c_1 \left( \frac{1}{s} + \frac{1}{s} \sum_{r=0}^{l-1} 2^{r/2} \right) \wedge t_l \leq c_1 \left( \frac{\sqrt{t_l}}{(\sqrt{2}-1)s} \wedge t_l \right) \leq 4c_1 \left( \frac{\sqrt{T}}{s} \wedge T \right). \end{aligned}$$

Hence, we have proved that the upper bound of Theorem 1 is a consequence of Theorem 7. ■

The proof of Theorem 7 is quite lengthy. To help the reader to understand the organization of this demonstration, the section starts with a sketch of proof.

**B.1. Outline of the proof of Theorem 7**

As any strategy has at most linear regret, it is sufficient to prove that there exist two positive numerical constants  $c_{\text{thresh}}$  and  $c_1$  such that, for any  $T \geq c_{\text{thresh}}/s^2$ , the number  $N^{\text{bad}}(\psi, T)$  of pairs sampled among  $\mathcal{E}^{\text{bad}}$  by the strategy  $\psi$  described in Algorithm 1 in Section 3.2 is smaller than  $c_1 \sqrt{T}/s$  with probability at least  $1 - 13/T$ . As a consequence, in the proof, without loss of generality, it is assumed that  $T \geq c_{\text{thresh}}/s^2$ , for a sufficiently large constant  $c_{\text{thresh}}$ . To prove the theorem, it is sufficient to show that neither Steps 1, 2 nor 3 of the algorithm sample more than  $O(\sqrt{T}/s)$  “bad” pairs, where a bad pair involves one node from Community 1 and one from Community 2.

*Step 1.* In the first step, the algorithm samples at random a core-set  $\mathcal{N}$  of  $N = \lceil \sqrt{T}/\log(s\sqrt{T}) \rceil$  nodes (point (1) in the algorithm). In this core-set, with large probability, at least  $\lceil N/4 \rceil$  nodes from each community are sampled. This result follows from Hoeffding’s concentration inequality for hypergeometric random variables, it is rigorously established in Lemma 8 (2).

Each pair of the core-set is sampled with probability proportional to  $\sqrt{T}/(s\binom{N}{2})$  (point (2) of the algorithm). With high probability, the set of sampled pairs  $\mathcal{O}_0$  has cardinality  $|\mathcal{O}_0| \asymp \sqrt{T}/s$ ; see Lemma 8 (3). At this point, the observed graph follows a cSBM with connection probabilities

$$\tilde{p} \asymp p\sqrt{T}/\left(s\binom{N}{2}\right) \quad \text{and} \quad \tilde{q} \asymp q\sqrt{T}/\left(s\binom{N}{2}\right).$$

By (3.2), setting  $\tilde{s} = (\tilde{p} - \tilde{q})^2/(\tilde{p} + \tilde{q})$  the proportion of misclassified nodes by GOODCLUST is upper bounded by

$$\exp(-c_1^{\text{GC}} N \tilde{s}) = \exp\left(-c \frac{\sqrt{T}/s}{N} s\right) = \exp(-\log(s\sqrt{T})) \leq \frac{1}{Ns},$$

with probability at least  $1 - c_2^{\text{GC}}/N^3$ . In particular, at most  $1/s$  nodes of the core-set are misclassified. A rigorous proof of this last statement is provided in Lemma 8 (4).

Let us comment briefly the choice of the cardinalities  $N$  of the core-set and  $|\mathcal{O}_0|$  of the sampled pairs in this first step of the algorithm. These are chosen to guarantee the following properties:

- (1.i)  $N$  is sufficiently large to make the probability  $c^{\text{GC}}/N^3$  small and, on the other hand,  $N$  is sufficiently small so that one can classify a large proportion of  $\mathcal{N}$  with less than  $|\mathcal{O}_0| = O(\sqrt{T}/s)$  observed pairs;
- (1.ii)  $|\mathcal{O}_0|$  is large enough to ensure that the proportion of misclassified nodes in  $\mathcal{N}$  satisfies  $\exp(-c_1^{\text{GC}} \mathbb{E}[|\mathcal{O}_0|]s/N) \leq 1/(Ns)$ ;
- (1.iii) on the other hand,  $|\mathcal{O}_0|$  is small enough, namely  $|\mathcal{O}_0| = O(\sqrt{T}/s)$ , to ensure a regret  $O(\sqrt{T}/s)$  in this exploratory phase of the algorithm.

Before moving to the screening step (Step 2) of the algorithm, the estimator

$$\hat{\tau} = \frac{1}{|\mathcal{O}_0|} \sum_{\{x,x'\} \in \mathcal{O}_0} A_{\{x,x'\}}$$

of  $\tau = (p + q)/2$  is shown to satisfy, with large probability,

$$|\hat{\tau} - p| \wedge |\hat{\tau} - q| \geq |\hat{\tau} - \tau|.$$

This property is obtained by a careful application of Bernstein inequality for hypergeometric random variables in Lemma 8 (5). This estimation of  $\tau$  is sufficient for the screening step.

*Step 2.* The second step of the algorithm samples uniformly at random a set  $\mathcal{A}_0$  of  $\lceil 8\sqrt{2T} \rceil$  nodes. These nodes are screened with the following objectives:

- (2.i) a set of at least  $\lceil \sqrt{2T} \rceil$  nodes among  $\mathcal{A}_0$  are selected containing at most  $1/s$  members of Community 2;
- (2.ii) a set of at most  $O(\sqrt{T}/s)$  bad pairs is sampled during this screening.

Claims (2.i) and (2.ii) are formally established in Lemma 9 (8), (10) and Lemma 9 (9), respectively.

The main tool for proving these two properties is Lemma 10. It ensures that the probability that a node from Community 2 is not removed after  $i$  steps of screening decreases exponentially fast with  $i$ . Therefore, after  $I \asymp \log(s\sqrt{T})$  screening steps, each node from Community 2 remains with probability at most  $e^{-c \log(s\sqrt{T})}$ . Since there are  $O(\sqrt{T})$  nodes in  $\mathcal{A}_0$ , the expected number of remaining nodes from Community 2 is upper bounded, when  $T \gtrsim 1/s^2$ , by

$$O(\sqrt{T} e^{-c \log(s\sqrt{T})}) \lesssim \frac{1}{s}.$$

The same bound holds with high probability. Similar arguments are used to obtain that, with large probability, less than  $\lceil 8\sqrt{2T} \rceil - \lceil \sqrt{2T} \rceil$  nodes are removed during the screening step, which shows property (2.i).

The proof of property (2.ii) is more involved. At point (7.b) of the algorithm, a bad pair is sampled when it involves either:

- (2.ii.a) a node of Community 2 and a well-classified node of the core-set;
- (2.ii.b) a node of Community 1 and a misclassified node of the core-set.

The number of pairs in the case (2.ii.a) is simply bounded from above by  $|\mathcal{A}_0| = O(\sqrt{T})$  multiplied by the number of misclassified nodes in the core-set. We have checked in Step 1, that the number of misclassified nodes in the core-set is bounded from above by  $O(1/s)$ . So, on this event, the number of such bad pairs is at most  $O(\sqrt{T} \cdot 1/s)$ .

The number of pairs in the case (2.ii.b) is bounded from above as follows. During each screening step (point (7)), a node is queried  $k = O(1/s)$  times. Thus, the number of queries of a node from Community 2 during this screening step is  $k$  times the number of screening steps before it is removed. Recall that, from Lemma 10, the probability that a node of Community 2 remains after  $i$  screening steps decreases exponentially fast with  $i$ . Hence, the expected number of queries of a node from Community 2 is bounded from above by

$$k \sum_{i \geq 1} e^{-ci} = O(k) = O(1/s).$$

The number of sampled pairs in case (2.ii.b) is smaller than the total number of queries on nodes from Community 2 in  $\mathcal{A}_0$ , which is smaller than  $O(|\mathcal{A}_0|k) = O(\sqrt{T}/s)$ . This bound also holds with high probability, which proves property (2.ii.b).

*Step 3.* During Step 3 of the algorithm, pairs within  $\mathcal{A}_I$  are sampled until  $T$  pairs have been sampled overall. On the event where  $|\mathcal{A}_I|$  is larger than  $\sqrt{2T}$ , this sampling is possible. In addition, on the event where the number of nodes from Community 2 in  $\mathcal{A}_I$  is upper bounded by  $1/s$ , the number of bad pairs in  $\mathcal{A}_I$  is smaller than

$$O(|\mathcal{A}_I|/s) = O(|\mathcal{A}_0|/s) = O(\sqrt{T}/s).$$

## B.2. Proof of Theorem 7

All we need is to prove that there exists a numerical constant  $c_{\text{thresh}} \geq 1$ , such that, for any  $T \geq c_{\text{thresh}}/s^2$ , the upper bound  $N^{\text{bad}}(\psi, T) \leq c_1 \sqrt{T}/s$  holds with probability at least  $1 - 13/T$ . We focus then on the case where  $T \geq c_{\text{thresh}}/s^2$ .

Denote by  $\hat{G} = \{\hat{G}_1, \hat{G}_2\}$  the partition of  $\mathcal{N}$  output by the GOODCLUST algorithm and by  $S \Delta S'$  the symmetric difference between two sets  $S, S'$ . Define the community labelling vectors  $Z$  and  $\hat{Z}$  by  $Z_x = j$  for all  $x \in G_j$  and  $\hat{Z}_x = j$  for all  $x \in \hat{G}_j$ . The following lemma controls the first step of the algorithm.

**Lemma 8.** *There exists numerical constants  $c_{\text{thresh}} \geq e$  and  $T_0 \geq 1$  such that, if  $T_0 \leq T \leq n^2/16$  and  $s\sqrt{T} \geq c_{\text{thresh}}$ , then with probability at least  $1 - 9/T$ :*

- (1) *only a small part of the nodes has been sampled:  $N \leq n/4$ ;*

- (2) the two communities of the sampled nodes are approximately balanced, that is  $N_1 \wedge N_2 \geq N/4$ , where  $N_j := |\{Z = j\} \cap \mathcal{N}|$  is the number of nodes from community  $j$  in  $\mathcal{N}$ ;
- (3) the cardinality of the sample pairs fulfils  $c_{\mathcal{O}_0} \sqrt{T}/2s \leq |\mathcal{O}_0| \leq 3c_{\mathcal{O}_0} \sqrt{T}/2s$ ;
- (4) the fraction of misclassified nodes is upper bounded by

$$\varepsilon_N = \inf_{\pi \text{ permutation on } \{1,2\}} \frac{1}{2N} \sum_{k=1}^2 |\{Z = k\} \Delta \{\widehat{Z} = \pi(k)\}| \leq \frac{1}{sN};$$

- (5)  $|\widehat{\tau} - (p + q)/2| \leq (p - q)/4$ .

We refer to Section B.3.1 for a proof of this lemma.

At the end of the first step,  $|\mathcal{O}_0| = O(\sqrt{T}/s)$  pairs have been sampled according to Lemma 8 (3), thus resulting in a number of sampled bad pairs  $O(\sqrt{T}/s)$ . Let us now turn to the second step of the algorithm.

Assume without loss of generality that the community labelling  $\widehat{Z}$  of the nodes in  $\mathcal{N}$  is mostly in agreement with  $Z$ , i.e., the infimum in the definition of  $\varepsilon_N$  is achieved for the identity permutation:

$$\varepsilon_N = \frac{1}{2N} \sum_{k=1}^2 |\{Z = k\} \Delta \{\widehat{Z} = k\}|.$$

If it is not the case, the remaining of the proof still holds but with  $\{Z = 1\}$  replaced by  $\{Z = 2\}$ .

For each  $x \in \mathcal{A}_0$ , the (distinct) nodes  $\{y_1^x, \dots, y_{kI}^x\}$  are sampled uniformly at random in  $\mathcal{N} \cap \{\widehat{Z} = 1\}$ . Let  $\mathcal{V}_{x,0} = \emptyset$  for all  $x \in \mathcal{A}_0$  and  $\mathcal{V}_{x,i} = \{y_1^x, \dots, y_{ki}^x\}$  for  $i = 1, \dots, I$ . Note that  $|\mathcal{V}_{x,j}| = kj$  for all  $x \in \mathcal{A}_0$ . By induction, construct the sequences of sets  $(\mathcal{A}_i)_{0 \leq i \leq I}$ , which contain the “active” nodes remaining at each iteration, and  $(\mathcal{O}_i)_{0 \leq i \leq I}$ , which contain the sampled pairs.

More formally, for  $i \geq 1$  and all  $x \in \mathcal{A}_{i-1}$ , the pairs  $\{(x, y_{(i-1)k+a}^x), 1 \leq a \leq k\}$  are observed at iteration  $i$ , so that

$$\mathcal{O}_i = \mathcal{O}_{i-1} \cup \bigcup_{x \in \mathcal{A}_{i-1}} \{(x, y_{(i-1)k+a}^x), 1 \leq a \leq k\}.$$

We remind the reader that we estimate the connectivity between  $x$  and Community 1 by

$$\widehat{p}_{x,i} = \frac{1}{ki} \sum_{y \in \mathcal{V}_{x,i}} A_{x,y},$$

and only keep the nodes whose estimated connectivity is large enough in the active set:

$$\mathcal{A}_i = \{x \in \mathcal{A}_{i-1} : \widehat{p}_{x,i} \geq \widehat{\tau}\}.$$

After  $I$  iterations, the total number of sampled pairs is

$$|\mathcal{O}_I| = |\mathcal{O}_0| + k \sum_{i=0}^{I-1} |\mathcal{A}_i|$$

and the number of sampled bad pairs from this step is upper bounded by

$$k \sum_{i=0}^{I-1} |\mathcal{A}_i \cap \{Z \neq 1\}| + |\mathcal{A}_0 \cap \{Z = 1\}| \times |\mathcal{N} \cap \{\hat{Z} \neq Z\}|, \quad (\text{B.2})$$

where the first term comes from the pairs connecting Community 2 to the core-set and the second term comes from the pairs connecting Community 1 to a misclassified vertex of the core-set.

The following lemma controls this screening step.

**Lemma 9.** *There exists numerical constants  $T'_0, c'_{\text{thresh}}$  larger than 1 such that if  $T'_0 \leq T \leq (3n/64\sqrt{2})^2$  and  $s\sqrt{T} \geq c'_{\text{thresh}}$ , then with probability at least  $1 - 13/T$ , Lemma 8 holds and*

- (6) *the algorithm does not run out of connections with the core-set of the first step:  $kI \leq N$ ;*
- (7) *it is possible to take  $|\mathcal{A}_0|$  new vertices:  $|\mathcal{A}_0| \leq 3n/4 \leq n - N$ ;*
- (8) *few vertices from the wrong community remain:  $|\mathcal{A}_I \cap \{Z \neq 1\}| \leq 1/s$ ;*
- (9) *the number of sampled bad pairs from nodes in the wrong community is controlled:*

$$k \sum_{i=0}^{I-1} |\mathcal{A}_i \cap \{Z \neq 1\}| \leq C_{\text{fail}} \frac{\sqrt{T}}{s}$$

*for a numerical constant  $C_{\text{fail}}$ ;*

- (10) *enough vertices from Community 1 remain for the next step:*

$$|\mathcal{A}_I \cap \{Z = 1\}| \geq \sqrt{2T}.$$

We refer to Section B.3.2 for a proof of this lemma.

Equation (B.2) together with Lemma 9 (9) and Lemma 8 (4) entail that the number of sampled bad pairs during the screening step is again  $O(\sqrt{T}/s)$ .

Finally, during the last step, the algorithm uses the remaining budget to observe pairs uniformly at random between vertices of  $\mathcal{A}_I$ . Lemma 9 (10) ensures that the number of possible pairs is larger than  $T - \sqrt{T/2}$ , which allows to spend the whole budget (since at least  $\lceil \sqrt{T/2} \rceil$  pairs have been observed in the previous steps), and Lemma 9 (8) ensures that the number of sampled bad pairs of this step is again  $O(\sqrt{T}/s)$ . Hence, the total number of bad pairs sampled during the whole process is  $O(\sqrt{T}/s)$ .

### B.3. Proofs of the technical lemmas

**B.3.1. Proof of Lemma 8.** (1) The proof of Lemma 8(1) is straightforward: since  $\sqrt{T} \leq n/4$  by assumption, the condition  $N \leq n/4$  holds as soon as  $N \leq \sqrt{T}$ , that is

$$\left\lceil \frac{\sqrt{T}}{\log(s\sqrt{T})} \right\rceil \leq \sqrt{T}$$

by definition of  $N$ . Therefore point (1) holds true as soon as  $s\sqrt{T} \geq c_{\text{thresh}}$  for some numerical constant  $c_{\text{thresh}}$ .

(2) There are only two communities, so it is enough to consider the first one. Since the communities are balanced, the number  $N_1$  of nodes from Community 1 in the core-set follows an hypergeometric distribution with parameters  $(N, 1/2, n)$ . Therefore,

$$\mathbb{P}\left(\left|N_1 - \frac{N}{2}\right| \geq \sqrt{2N \log N}\right) \leq \frac{2}{N^4}$$

using equation (D.1). Since  $N = \lceil \sqrt{T} / \log(s\sqrt{T}) \rceil$ , we have

$$\frac{2}{N^4} \leq \frac{2 \log(s\sqrt{T})^4}{T^2} \leq \frac{(\log T)^4}{8T^2}$$

using  $s \leq 1$ , which is upper bounded by  $1/T$  for all  $T \geq 1$ . Assuming  $s\sqrt{T} \geq c_{\text{thresh}}$  for some numerical constant  $c_{\text{thresh}} \geq e$ , one has  $\sqrt{T} / \log \sqrt{T} \leq N \leq \sqrt{T}$ , so that

$$\frac{\sqrt{2N \log N}}{N/4} \leq 4\sqrt{2} \sqrt{\frac{\log(\sqrt{T})}{\sqrt{T} / \log \sqrt{T}}} \leq 4\sqrt{2} \sqrt{\frac{\log(\sqrt{T})^2}{\sqrt{T}}}.$$

Therefore, it is smaller than 1 as soon as  $T \geq T_{0,2}$  for some numerical constant  $T_{0,2}$ , which entails

$$\mathbb{P}\left(\left|N_1 - \frac{N}{2}\right| \geq \frac{N}{4}\right) \leq \frac{1}{T},$$

and the same for  $N_2$ .

(3) The number  $|\mathcal{O}_0|$  of sampled pairs in the core-set  $\mathcal{N}$  follows a binomial distribution with parameters  $\left(\binom{N}{2}, c_{\mathcal{O}_0} \sqrt{T} / s \binom{N}{2}\right)$ . Therefore,

$$\mathbb{P}\left(\left||\mathcal{O}_0| - c_{\mathcal{O}_0} \frac{\sqrt{T}}{s}\right| \geq \sqrt{2c_{\mathcal{O}_0} \frac{\sqrt{T}}{s} \log(2T) + \log(2T)}\right) \leq \frac{1}{T}$$

using Bernstein's inequality (D.3). This implies that

$$\frac{1}{2}c_{\mathcal{O}_0} \frac{\sqrt{T}}{s} \leq |\mathcal{O}_0| \leq \frac{3}{2}c_{\mathcal{O}_0} \frac{\sqrt{T}}{s} \tag{B.3}$$

as soon as  $T \geq T_{0,3}$  for some numerical constant  $T_{0,3}$ .

Let us check that the probability parameter of the binomial distribution is well defined, that is, the condition  $c_{\theta_0} \sqrt{T}/s \binom{N}{2} \in [0, 1]$  is satisfied. One can show that  $N \geq 8$  as soon as  $T \geq T_{0,3}$  for some numerical constant  $T_{0,3}$ . Then

$$\binom{N}{2} \geq \frac{N^2}{4},$$

so that the condition holds as soon as  $c_{\theta_0} \sqrt{T}/s \leq N^2/4$ , which is implied by

$$c_{\theta_0} \sqrt{T}/s \leq \frac{1}{4} \frac{T}{(\log(s\sqrt{T}))^2},$$

or equivalently

$$\frac{s\sqrt{T}}{(\log(s\sqrt{T}))^2} \geq 4c_{\theta_0}.$$

Hence, the condition  $c_{\theta_0} \sqrt{T}/s \binom{N}{2} \in [0, 1]$  holds as soon as  $s\sqrt{T} \geq c_{\text{thresh}}$  for some numerical  $c_{\text{thresh}}$ . This, together with (B.3), concludes the proof of Lemma 8 (3).

(4) Since each pair of  $\mathcal{N}$  is sampled with probability  $c_{\theta_0} \sqrt{T}/s \binom{N}{2}$ , the matrix  $\tilde{A}$  defined by  $\tilde{A}_{x,x'} = A_{x,x'}$  if the pair  $\{x, x'\}$  has been sampled and zero otherwise has the same distribution as the adjacency matrix of a fully observed SBM with connection probabilities  $\tilde{p} = p c_{\theta_0} \sqrt{T}/s \binom{N}{2}$  and  $\tilde{q} = q c_{\theta_0} \sqrt{T}/s \binom{N}{2}$ . Therefore, the proportion  $\varepsilon_N$  of misclassified nodes in  $\mathcal{N}$  by the GOODCLUST algorithm is upper bounded by

$$\varepsilon_N \leq \exp\left(-c_1^{\text{GC}} N \frac{(\tilde{p} - \tilde{q})^2}{\tilde{p} + \tilde{q}}\right)$$

with probability at least  $1 - c^{\text{GC}}/N^3$ . Hence, with probability at least  $1 - 1/T$ , we have

$$\varepsilon_N \leq \exp\left(-2c_1^{\text{GC}} c_{\theta_0} \frac{\log(s\sqrt{T})}{2}\right)$$

using  $N := \lceil \sqrt{T}/\log(s\sqrt{T}) \rceil \leq 2\sqrt{T}/\log(s\sqrt{T})$  as soon as  $T \geq T_{0,4}$  for some numerical constant  $T_{0,4}$ . Hence, by taking  $c_{\theta_0} \geq 1/(c_1^{\text{GC}})$ , with probability at least  $1 - 1/T$ , one has

$$\varepsilon_N \leq \exp(-\log(s\sqrt{T})) = \frac{1}{s\sqrt{T}},$$

so that

$$\varepsilon_N \leq \frac{1}{sN}$$

as soon as  $N \leq \sqrt{T}$ , which holds true when  $s\sqrt{T} \geq c_{\text{thresh}}$  for some numerical constant  $c_{\text{thresh}}$ .

(5) Let  $\mathcal{O}_{\text{within}} := \mathcal{O}_0 \cap \mathcal{E}^{\text{good}}$  be the subset of within-group pairs, and  $\mathcal{O}_{\text{out}} := \mathcal{O}_0 \setminus \mathcal{O}_{\text{within}}$  the subset of pairs between two different communities. Then

$$\hat{\tau} = \frac{|\mathcal{O}_{\text{within}}|}{|\mathcal{O}_0|} \frac{1}{|\mathcal{O}_{\text{within}}|} \sum_{(x,x') \in \mathcal{O}_{\text{within}}} A_{x,x'} + \frac{|\mathcal{O}_{\text{out}}|}{|\mathcal{O}_0|} \frac{1}{|\mathcal{O}_{\text{out}}|} \sum_{(x,x') \in \mathcal{O}_{\text{out}}} A_{x,x'}.$$

Conditionally to the number of sampled pairs  $|\mathcal{O}_0|$  and the number of within-group pairs  $|\mathcal{O}_{\text{within}}|$ , the sum  $\sum_{(x,x') \in \mathcal{O}_{\text{within}}} A_{x,x'}$  (respectively,  $\sum_{(x,x') \in \mathcal{O}_{\text{out}}} A_{x,x'}$ ) is independent of  $\mathcal{O}_0$ , and is a sum of i.i.d. Bernoulli random variables with parameter  $p$  (respectively,  $q$ ). Therefore, Bernstein's inequality (D.2) ensures that with probability at least  $1 - 4/T$ , we have

$$\begin{aligned} \left| \frac{|\mathcal{O}_{\text{within}}|}{|\mathcal{O}_0|} \left| \frac{1}{|\mathcal{O}_{\text{within}}|} \sum_{(x,x') \in \mathcal{O}_{\text{within}}} A_{x,x'} - p \right| \right| &\leq \sqrt{2p \frac{\log T}{|\mathcal{O}_0|}} + \frac{\log T}{|\mathcal{O}_0|}, \\ \left| \frac{|\mathcal{O}_{\text{out}}|}{|\mathcal{O}_0|} \left| \frac{1}{|\mathcal{O}_{\text{out}}|} \sum_{(x,x') \in \mathcal{O}_{\text{out}}} A_{x,x'} - q \right| \right| &\leq \sqrt{2q \frac{\log T}{|\mathcal{O}_0|}} + \frac{\log T}{|\mathcal{O}_0|}. \end{aligned}$$

Using point (3), one has  $|\mathcal{O}_0| \geq c_{\mathcal{O}_0} \sqrt{T}/(2s)$  with probability at least  $1 - 1/T$ , so that

$$\begin{aligned} \left| \hat{\tau} - \left( \frac{|\mathcal{O}_{\text{within}}|}{|\mathcal{O}_0|} p + \frac{|\mathcal{O}_{\text{out}}|}{|\mathcal{O}_0|} q \right) \right| &\leq 2 \sqrt{2ps \frac{\log T}{c_{\mathcal{O}_0} \sqrt{T}/2}} + 2s \frac{\log T}{c_{\mathcal{O}_0} \sqrt{T}/2} \\ &\leq 2(p-q) \sqrt{2 \frac{\log T}{c_{\mathcal{O}_0} \sqrt{T}/2}} + 2(p-q) \frac{\log T}{c_{\mathcal{O}_0} \sqrt{T}/2} \end{aligned}$$

with probability at least  $1 - 5/T$ , using  $s = (p-q)^2/p \leq p-q$ . Finally, since  $c_{\mathcal{O}_0}/2 \geq 1$  and  $|\mathcal{O}_{\text{within}}| = |\mathcal{O}_0| - |\mathcal{O}_{\text{out}}|$ ,

$$\begin{aligned} \left| \hat{\tau} - \frac{p+q}{2} \right| &\leq \left| \frac{|\mathcal{O}_{\text{within}}|}{|\mathcal{O}_0|} p + \frac{|\mathcal{O}_{\text{out}}|}{|\mathcal{O}_0|} q - \frac{p+q}{2} \right| + 2(p-q) \sqrt{2 \frac{\log T}{\sqrt{T}}} + 2(p-q) \frac{\log T}{\sqrt{T}} \\ &\leq \left| \left( 2 \frac{|\mathcal{O}_{\text{out}}|}{|\mathcal{O}_0|} - 1 \right) \frac{p-q}{2} \right| + \frac{|p-q|}{16} \end{aligned} \tag{B.4}$$

as soon as  $T \geq T_{0,4}$  for some numerical constant  $T_{0,4}$ .

Conditionally to the number of pairs  $|\mathcal{O}_0|$  and the sizes  $N_1$  and  $N_2$  of the two communities sampled in  $\mathcal{N}$ , the number  $|\mathcal{O}_{\text{out}}|$  of between group pairs follows a hypergeometric distribution with parameters  $(|\mathcal{O}_0|, r, \binom{N}{2})$  with  $r = N_1 N_2 / \binom{N}{2}$ . Conditionally to  $|\mathcal{O}_0|$  and the event  $3/8 \leq r \leq 5/8$ , the random variable  $|\mathcal{O}_{\text{out}}|$  dominates

stochastically an hypergeometric random variable with parameters  $(|\mathcal{O}_0|, 3/8, \binom{N}{2})$  and it is stochastically dominated by an hypergeometric random variable with parameters  $(|\mathcal{O}_0|, 5/8, \binom{N}{2})$ . There exists a real  $\gamma > 0$  such that  $N_1 = \gamma N$  and  $N_2 = (1 - \gamma)N$ , so that

$$\begin{aligned} r &= \frac{\gamma N(1 - \gamma)N}{N(N - 1)/2} = 2\gamma(1 - \gamma)\left(1 + \frac{1}{N - 1}\right) \\ &= 2\gamma(1 - \gamma)\left(1 + \frac{1}{\sqrt{T}/\log(s\sqrt{T}) - 1}\right). \end{aligned}$$

Using point (2), one has with probability at least  $1 - 1/T$  that  $\frac{1}{4} \leq \gamma \leq \frac{3}{4}$  which entails  $\frac{3}{8} \leq r \leq \frac{5}{8}$  as soon as  $T \geq T_{0,5}$  for some numerical constant  $T_{0,5}$ . Therefore,

$$\mathbb{P}\left(|\mathcal{O}_{\text{out}}| \leq \frac{3|\mathcal{O}_0|}{8} - \sqrt{\frac{|\mathcal{O}_0| \log T}{2}}\right) \leq \frac{1}{T} + \frac{1}{T}$$

using equation (D.1), and similarly

$$\mathbb{P}\left(|\mathcal{O}_{\text{out}}| \geq \frac{5|\mathcal{O}_0|}{8} + \sqrt{\frac{|\mathcal{O}_0| \log T}{2}}\right) \leq \frac{1}{T} + \frac{1}{T}.$$

Using point (3), one has with probability at least  $1 - 1/T$  that  $|\mathcal{O}_0| \geq c_{\mathcal{O}_0} \sqrt{T}/(2s)$ , which entails that

$$\sqrt{\frac{|\mathcal{O}_0| \log T}{2}} \leq \frac{|\mathcal{O}_0|}{16},$$

as soon as  $T \geq T_{0,6}$  for some numerical constant  $T_{0,6}$ . Hence,

$$\frac{5}{16} \leq \frac{|\mathcal{O}_{\text{out}}|}{|\mathcal{O}_0|} \leq \frac{11}{16}$$

with probability  $1 - 5/T$ . This, together with equation (B.4), concludes the proof of Lemma 8 (5) (which holds with probability  $1 - 8/T$ ). ■

**B.3.2. Proof of Lemma 9.** (6, 7) There exists a constant  $c'_{\text{thresh}}$  such that

$$4C_I C_k \leq \frac{s\sqrt{T}}{(\log(s\sqrt{T}))^2}$$

as soon as  $s\sqrt{T} \geq c'_{\text{thresh}}$ . It follows that  $kI \leq N$ .

(7) Lemma 9 (7) follows by straightforward algebra.

(8) For all  $x \in \mathcal{A}_0$ , denote by  $T_x = \max\{i : x \in \mathcal{A}_i\}$  the index of the last iteration where the vertex  $x$  was in the active set. Let us first show that if  $x$  is not in the first community, then  $T_x$  has sub-exponential tails.

**Lemma 10.** *Set  $\rho' = 1/2000$ . If  $C_k \geq (\log 3)/\rho'$ , then*

$$\forall i \in \mathbb{N}^*, \quad \mathbb{P}(T_x \geq i) \leq e^{-\rho' C_k i}. \quad (\text{B.5})$$

We refer to Section B.4 for a proof of this lemma.

*Proof of Lemma 9, cont.* Let us now finish the proof of point (8). Let  $T^{(1)} = |\mathcal{O}_0|$  and  $V_x = \mathbf{1}_{T_x \geq I}$ . Conditionally on  $\mathcal{F}_{T^{(1)}}$ , the variables  $(V_x)_{x \in \mathcal{A}_0 \cap \{Z \neq 1\}}$  are i.i.d. Bernoulli random variables with parameter  $r \leq e^{-\rho' C_k I}$  by equation (B.5). Therefore, for all  $i \in \mathbb{N}$ ,

$$\mathbb{P}\left(\sum_{x \in \mathcal{A}_0 \cap \{Z \neq 1\}} V_x = i\right) \leq \frac{|\mathcal{A}_0|^i}{i!} r^i,$$

so that

$$\begin{aligned} \mathbb{P}(|\mathcal{A}_I \cap \{Z \neq 1\}| \geq i) &\leq \sum_{j \geq i} \frac{|\mathcal{A}_0|^j}{j!} r^j \\ &\leq \frac{(|\mathcal{A}_0| r)^i}{i!} \sum_{j \geq 0} \left(\frac{|\mathcal{A}_0| r}{i}\right)^j \\ &\leq 2 \frac{(|\mathcal{A}_0| r)^i}{i!} \end{aligned}$$

as soon as  $i \geq 2|\mathcal{A}_0| r$ . For  $i = \lceil 1/s \rceil$ , this condition holds if  $16\sqrt{2} \leq (s\sqrt{T})^{\rho' C_k C_I - 1}$ , which holds when  $C_I C_k \geq 4/\rho'$  and  $s\sqrt{T} \geq c'_{th}$ .

Taking  $i = \lceil 1/s \rceil$  and using that  $i! \geq (i/e)^i$  for all  $i \geq 1$ , it follows that

$$\mathbb{P}\left(|\mathcal{A}_I \cap \{Z \neq 1\}| \geq \frac{1}{s}\right) \leq 2 \left(\frac{e|\mathcal{A}_0| r}{\lceil 1/s \rceil}\right)^{\lceil 1/s \rceil} \leq 2(se|\mathcal{A}_0| r)^{1/s}$$

as soon as  $se|\mathcal{A}_0| r \leq 1$ .

We want to take  $r$  small enough such that  $2(se|\mathcal{A}_0| r)^{1/s} \leq 1/T$ , that is

$$\log(se|\mathcal{A}_0|) + s \log(2T) \leq (-\log r),$$

which holds as soon as

$$\rho' C_k I \geq \log(s\sqrt{T}) + \log(32e\sqrt{2}) + s \log T.$$

using  $|\mathcal{A}_0| \leq 16\sqrt{2T}$ .

Note that

$$\frac{s \log T}{\log(s\sqrt{T})} = 2 \frac{s\sqrt{T}}{\log(s\sqrt{T})} \frac{\log \sqrt{T}}{\sqrt{T}} \leq 2,$$

since  $\log(x)/x$  is decreasing for  $x > e$  and  $s\sqrt{T} \leq \sqrt{T}$ , so that there exists a numerical constant  $c'_{\text{thresh}}$  such that if  $s\sqrt{T} \geq c'_{\text{thresh}}$ , then point (8) is implied by

$$\rho' C_k I \geq 4 \log(s\sqrt{T}),$$

which holds when  $C_I C_k \geq 4/\rho'$ .

(9) Note that

$$k \sum_{i=0}^{I-1} |\mathcal{A}_i \cap \{Z \neq 1\}| = k \sum_{x \in \mathcal{A}_0 \cap \{Z \neq 1\}} T_x.$$

Conditionally on  $\mathcal{A}_0$ , the random variables  $(T_x)_{x \in \mathcal{A}_0 \cap \{Z \neq 1\}}$  are i.i.d. random variables which are stochastically dominated by random variables  $Y_x \sim \mathcal{E}(\rho' C_k)$  by equation (B.5). These exponential random variables satisfy

$$\mathbb{E} \left( Y_x - \frac{1}{\rho' C_k} \right)^2 \leq \frac{1}{(\rho' C_k)^2}$$

and for all  $a \in \mathbb{N}$  such that  $a \geq 3$ , we have

$$\mathbb{E} \left( Y_x - \frac{1}{\rho' C_k} \right)_+^a \leq \frac{a!}{(\rho' C_k)^a}.$$

It follows that, by Bernstein’s inequality (see, for instance, [34, Proposition 2.9]), for all  $t > 0$ , we have

$$\mathbb{P} \left( \sum_{x \in \mathcal{A}_0} Y_x - \frac{|\mathcal{A}_0|}{\rho' C_k} \geq \frac{2\sqrt{|\mathcal{A}_0|t}}{\rho' C_k} + \frac{t}{\rho' C_k} \right) \leq e^{-t},$$

and therefore by taking  $t = \log T$ , with probability at least  $1 - 1/T$ , we obtain

$$\begin{aligned} \sum_{x \in \mathcal{A}_0 \cap \{Z \neq 1\}} T_x &\leq \frac{16\sqrt{2T}}{\rho' C_k} + \frac{2\sqrt{16\sqrt{2T} \log T}}{\rho' C_k} + \frac{\log T}{\rho' C_k} \\ &\leq \frac{32\sqrt{T}}{\rho' C_k} \end{aligned}$$

as soon as  $T \geq T'_0$  for some numerical constant  $T'_0$ . Hence, with probability at least  $1 - 1/T$ ,

$$k \sum_{i=0}^{I-1} |\mathcal{A}_i \cap \{Z \neq 1\}| \leq \frac{64\sqrt{T}}{\rho' s},$$

using  $k \leq 2C_K/s$ .

(10) The same proof as that of equation (B.6) shows that for all  $x \in \mathcal{A}_0 \cap \{Z = 1\}$ , for all  $i \geq 1$  and for all  $t > 0$ ,

$$\mathbb{P}\left(\hat{p}_{x,i} < p - \frac{|p - q|}{8} - |p - q| \sqrt{\frac{t}{2ki}} - 2\sqrt{2p \frac{t}{ki}} - 2\frac{t}{ki}\right) \leq 3e^{-t},$$

so that by union bound and the inequality  $k \geq C_k$ , we have

$$\mathbb{P}\left(\exists i \geq 1, \hat{p}_{x,i} < \frac{7p + q}{8} - |p - q| \left[ \sqrt{\frac{\log(2\pi^2 i^2)}{C_k i}} \left(\frac{1}{\sqrt{2}} + 2\sqrt{2}\right) + 2\frac{\log(2\pi^2 i^2)}{C_k i} \right] \right) \leq \frac{1}{4}.$$

Therefore, if  $C_k$  is larger than a numerical constant,

$$\mathbb{P}\left(\exists i \geq 1, \hat{p}_{x,i} < \frac{3p + q}{4}\right) \leq \frac{1}{4},$$

which, combined with point (5) of Theorem 8, implies

$$\mathbb{P}(\exists i \geq 1, \hat{p}_{x,i} < \hat{v}) \leq \frac{1}{4}.$$

Let  $V_x = \mathbf{1}_{x \in \mathcal{A}_I}$  for all  $x \in \mathcal{A}_0 \cap \{Z = 1\}$ . The above inequality ensures that, conditionally on  $\mathcal{A}_0$ , the  $(V_x)_{x \in \mathcal{A}_0 \cap \{Z = 1\}}$  are i.i.d. Bernoulli random variable with parameter  $r \geq 3/4$ . Therefore, Hoeffding's inequality entails

$$\mathbb{P}\left(|\mathcal{A}_I \cap \{Z = 1\}| \leq \frac{3|\mathcal{A}_0 \cap \{Z = 1\}|}{4} - \sqrt{|\mathcal{A}_0| \frac{\log T}{2}}\right) \leq \frac{1}{T}.$$

Let us assume for now that  $|\mathcal{A}_0 \cap \{Z = 1\}| \geq 2\sqrt{2T}$  with probability  $1 - 1/T$ . Then this ensures that for  $T$  larger than some numerical constant,

$$\mathbb{P}(|\mathcal{A}_I \cap \{Z = 1\}| \leq \sqrt{2T}) \leq \frac{1}{T} + \frac{1}{T}.$$

To conclude, note that conditionally on  $\mathcal{N}$ , the random variable  $|\mathcal{A}_0 \cap \{Z = 1\}|$  is an hypergeometric random variable with parameters  $(\lceil 8\sqrt{2T} \rceil, r', n - N)$ , where

$$r' = \frac{n/2 - |\mathcal{N} \cap \{Z = 1\}|}{n - N} \geq \frac{n/2 - 3/4 \cdot n/4}{n} \geq \frac{5}{16}$$

by points (1) and (2) of Theorem 8. Therefore, equation (D.1) implies that

$$\mathbb{P}\left(|\mathcal{A}_0 \cap \{Z = 1\}| \leq \frac{5}{16} 8\sqrt{2T} - \sqrt{\frac{16\sqrt{2T} \log T}{2}}\right) \leq \frac{1}{T},$$

so that for  $T$  larger than a numerical constant

$$\mathbb{P}(|\mathcal{A}_0 \cap \{Z = 1\}| \leq 2\sqrt{2T}) \leq \frac{1}{T}. \quad \blacksquare$$

**B.4. Proof of Lemma 10**

Let  $x \in \mathcal{A}_0 \cap \{Z \neq 1\}$  and assume that we are in the event of probability at least  $1 - 9/T$  where Theorem 8 holds. For all  $i \in \mathbb{N}^*$ ,

$$\begin{aligned} \mathbb{P}(T_x \geq i) &= \mathbb{P}(\forall j \in \{1, \dots, i\}, \hat{p}_{x,j} \geq \hat{\tau}) \\ &\leq \mathbb{P}(\hat{p}_{x,i} \geq \hat{\tau}) \\ &\leq \mathbb{P}\left(\hat{p}_{x,i} \geq \frac{p + 3q}{4}\right), \end{aligned}$$

using Lemma 8 (5).

Following the same proof as in Theorem 8 (5), one can show that for all  $x \in \mathcal{A}_0 \cap \{Z \neq 1\}$ , all  $i \geq 1$  and all  $t > 0$ ,

$$\mathbb{P}\left(\hat{p}_{x,i} \geq q + \frac{|\mathcal{V}_{x,i}^-|}{|\mathcal{V}_{x,i}|} |p - q| + 2\sqrt{2p \frac{t}{ki}} + 2\frac{t}{ki}\right) \leq 2e^{-t},$$

where  $\mathcal{V}_{x,i}^- := \mathcal{V}_{x,i} \cap \{Z \neq 1\}$ .

For  $s\sqrt{T} \geq c'_{\text{thresh}}$ , with  $c'_{\text{thresh}}$  such that  $\log(s\sqrt{T})/s\sqrt{T} \leq 1/64$ , one has  $1/s \leq N/64$ . Then, points (2) and (4) of Lemma 8 imply that

$$|\mathcal{N} \cap \{\hat{Z} = 1\} \cap \{Z \neq 1\}| \leq \frac{N}{64} \quad \text{and} \quad \hat{N}_1 := |\mathcal{N} \cap \{\hat{Z} = 1\}| \geq \frac{N}{8}.$$

Therefore, the proportion of misclassified vertices in  $\mathcal{N} \cap \{\hat{Z} = 1\}$  is at most  $1/8$ , so that conditionally on  $\hat{N}_1$  and the event of Lemma 8,  $|\mathcal{V}_{x,i}^-|$  is stochastically dominated by an hypergeometric distribution with parameters  $(ki, 1/8, \hat{N}_1)$ . Hence, equation (D.1) entails

$$\mathbb{P}\left(\frac{|\mathcal{V}_{x,i}^-|}{|\mathcal{V}_{x,i}|} \geq \frac{1}{8} + \sqrt{\frac{t}{2ki}}\right) \leq e^{-t},$$

so that for all  $i \geq 1$  and  $t > 0$ ,

$$\mathbb{P}\left(\hat{p}_{x,i} \geq q + \frac{|p - q|}{8} + |p - q| \sqrt{\frac{t}{2ki}} + 2\sqrt{2p \frac{t}{ki}} + 2\frac{t}{ki}\right) \leq 3e^{-t}. \tag{B.6}$$

Note that

$$\begin{aligned} &\frac{p + 3q}{4} - \left(q + \frac{|p - q|}{8} + |p - q| \sqrt{\frac{t}{2ki}} + 2\sqrt{2p \frac{t}{ki}} + 2\frac{t}{ki}\right) \\ &\geq \frac{|p - q|}{8} - \sqrt{\frac{t}{Cki}} \left(\frac{|p - q|\sqrt{s}}{\sqrt{2}} + 2\sqrt{2ps}\right) - 2\frac{ts}{Cki} \\ &\geq |p - q| \left(\frac{1}{8} - \sqrt{\frac{t}{Cki}} \left(\frac{1}{\sqrt{2}} + 2\sqrt{2}\right) - 2\frac{t}{Cki}\right), \end{aligned}$$

since  $s = (p - q)^2/p \leq 1$ .

Thus, there exists a numerical constant  $\rho = 10^{-3}$  such that by taking  $t = \rho C_k i$ ,

$$\frac{p + 3q}{4} \geq q + \frac{|p - q|}{8} + |p - q| \sqrt{\frac{t}{2ki}} + 2\sqrt{2p \frac{t}{ki}} + 2\frac{t}{ki},$$

so that

$$\mathbb{P}\left(\hat{p}_{x,i} \geq \frac{p + 3q}{4}\right) \leq 3e^{-\rho C_k i},$$

and finally by letting  $\rho' = \rho/2$  and if  $C_k \geq (\log 3)/\rho'$ , then

$$\forall i \in \mathbb{N}^*, \quad \mathbb{P}(T_x \geq i) \leq e^{-\rho' C_k i}.$$

### C. Proof of the constrained upper bound

This section proves the upper bound in Theorem 2. Recall that  $B = (B_T \wedge \sqrt{T})/2$  in the constrained algorithm (Algorithm 2).

It is enough to prove the upper bound in Theorem 2 in the case where  $sB \geq c_{\text{thresh}}$  for some numerical constant  $c_{\text{thresh}} \geq 1$ . Indeed, if  $sB \leq c_{\text{thresh}}$ , equation (4.1) automatically holds with  $c_2 \geq c_{\text{thresh}}$ . Hereafter, it is then assumed that  $sB \geq c_{\text{thresh}}$ .

The first step of the constrained algorithm (Algorithm 2) is almost identical to that of the unconstrained algorithm after replacing  $\sqrt{T}$  by  $B = (B_T \wedge \sqrt{T})/2$  in the cardinality of the core-set. The following lemma is a slight variant of Lemma 8 in this setting. The proof is omitted.

**Lemma 11.** *There exist numerical constants  $c_{\text{thresh}} \geq e$  and  $B_0 \geq 1$ , such that, if  $B \geq B_0$  and  $sB \geq c_{\text{thresh}}$  and  $T/B \leq n/136$ , then with probability at least  $1 - 9/(sB)$ :*

(1) *the number  $N_{\text{init}}$  of sampled nodes satisfies*

$$N_{\text{init}} \leq \frac{n}{8} - 4 \sum_{t=1}^{t_f-1} N^{(t)};$$

(2) *at least  $N_{\text{init}}/4$  nodes of each community have been sampled, that is*

$$|\{Z = j\} \cap \mathcal{N}_{\text{init}}| \geq N_{\text{init}}/4$$

*for each  $j \in \{1, 2\}$ ;*

(3) *the proportion  $\varepsilon_{N_{\text{init}}}$  of misclassified nodes satisfies*

$$\varepsilon_{N_{\text{init}}} = \inf_{\pi \text{ permutation on } \{1,2\}} \frac{1}{2N_{\text{init}}} \sum_{k=1}^2 |\{Z = k\} \Delta \{\hat{Z} = \pi(k)\}| \leq \frac{4}{512^2} \frac{1}{sB};$$

(4)  $|\hat{\tau} - (p + q)/2| \leq (p - q)/4$ .

At the end of the first step,  $|\mathcal{O}_0|$  pairs have been sampled and the sampling-regret therefore does not exceed  $\mathbb{E}[|\mathcal{O}_0|] = c_{\mathcal{O}_0} B/s \leq c_{\mathcal{O}_0} T/(sB)$  since, by definition of  $B$ ,  $T \geq B^2$ .

Let us proceed with the second step. To show that the sampling regret in the second step does not exceed  $O(T/(sB))$ , it is sufficient to prove that there exist two numerical constants  $c_{\text{proba}}$  and  $c_{\text{regret}}$  such that for any  $(T, B)$  satisfying  $sB \geq c_{\text{thresh}}$  and  $T/B \leq n/136$ , the number of bad pairs sampled during the second step is bounded from above by  $c_{\text{regret}} T/(sB)$  with probability at least  $1 - c_{\text{proba}}/(sB)$ . Indeed, since the number of bad pairs sampled in the second step  $N_{\text{Step 2}}^{\text{bad}}(\psi, T)$  cannot be larger than  $T$ , it directly follows that the sampling-regret during the second step is upper bounded by

$$\mathbb{E}[N_{\text{Step 2}}^{\text{bad}}(\psi, T)] \leq c_{\text{regret}} \frac{T}{sB_T} + T \frac{c_{\text{proba}}}{sB_T} \leq c' \frac{T}{sB_T}.$$

The following lemma provides such a control of the number of bad pairs accumulated in Step 2, as well as an upper bound on the number of misclassified nodes. It is a counterpart to Lemma 9 of the unconstrained case.

**Lemma 12.** *There exist two numerical constants  $B'_0 \geq 1$  and  $c'_{\text{thresh}} \geq e$  such that if  $B \geq B'_0$ ,  $sB \geq c'_{\text{thresh}}$  and  $T/B \leq n/136$ , then with probability at least  $1 - 63/(sB)$ , Lemma 11 holds. In addition, for all iterations of SCREENING in point (6) of the constrained algorithm,*

- (5) *it is always possible to sample  $|\mathcal{A}_0|$  new vertices:*

$$|V^{(0)}| \geq \dots \geq |V^{(t_f-1)}| \geq \frac{7n}{8};$$

- (6) *no node from  $\mathcal{A}_0$  has more than  $2B$  adjacent pairs sampled during the whole execution of the constrained algorithm;*  
 (7) *the algorithm does not run out of connections with the reference core-set:*

$$kI \leq N^{(0)} \leq \dots \leq N^{(t_f-1)};$$

and further there exists a numerical constant  $C_{\text{fail}}$  such that, for all  $t \in \{1, \dots, t_f\}$ , during the call  $\text{SCREENING}(\mathcal{N}^{(t-1)}, N^{(t)}, B, \hat{\tau}, V^{(t-1)})$ ,

- (8) *the number of bad pairs sampled during the  $t$ -th call to SCREENING from nodes in  $\mathcal{A}_0$  is controlled:*

$$\sum_{x \in \mathcal{A}_0} |\{y_a^x : (x, y_a^x) \text{ sampled and } Z_{y_a^x} \neq Z_x\}| \leq C_{\text{fail}} \frac{N^{(t)}}{s};$$

- (9) *few vertices from the wrong community remain:*

$$|\mathcal{N}^{(t)} \cap \{Z \neq 1\}| \leq 8N^{(t)}/(sB);$$

(10) enough vertices from Community 1 remain for the construction of the core-set  $\mathcal{N}^{(t)}$  of  $N^{(t)}$  nodes:

$$\sum_{j=1}^m |\mathcal{A}_I^{(j)} \cap \{Z = 1\}| \geq N^{(t)}.$$

As a consequence, the total number of bad pairs sampled during the second step is upper bounded by

$$C_{\text{fail}} \sum_{t=1}^{t_f} \frac{N^{(t)}}{s} \leq 2C_{\text{fail}} \frac{N_{t_f}}{s} \leq 4C_{\text{fail}} \frac{T}{sB},$$

with probability larger than  $1 - 63/(sB)$ .

We refer to Section C.1 for a proof of Lemma 12.

Let us now conclude the proof of the upper bound of Theorem 2. In the third step, the core-set  $\mathcal{N}^{(t_f)}$  has  $\lceil T/B \rceil \leq 2T/B$  nodes and a proportion of misclassified nodes smaller than  $8/(sB)$  with probability larger than  $1 - 63/(sB)$  by Lemma 12 (9). Since each node of  $\mathcal{N}^{(t_f)}$  is sampled at most  $B$  times, the number of bad pairs sampled during the third step is smaller than  $16T/(sB)$  with probability at least  $1 - 63/(sB)$ , and smaller than  $T$  otherwise.

Hence, using again that we always have  $N^{\text{bad}}(\psi, T) \leq T$ , the total sampling-regret  $\mathbb{E}[N^{\text{bad}}(\psi, T)]$  during the whole process is  $O(T/(sB))$ . The proof of the upper bound of Theorem 2 is complete.

### C.1. Proof of Lemma 12

Lemma 12 simultaneously controls all the iterations of SCREENING. To prove this, we use the following lemma, which controls each iteration.

**Lemma 13.** *There exists a numerical constant  $c'_{\text{thresh}} \geq e$  such that the following holds: Let  $\mathcal{N} \subset V_{\text{init}}$ ,  $N' \in \mathbb{N}$ ,  $B > 0$ ,  $\nu \in [0, 1]$  and  $V \subset V_{\text{init}}$ , and*

$$(\mathcal{N}', V') = \text{SCREENING}(\mathcal{N}, N', B, \nu, V).$$

Write  $N = |\mathcal{N}|$ . Assume that  $sB \geq c'_{\text{thresh}}$ , that  $B \leq 4N' \leq 4N \log(sB)$ , that the proportion of misclassified nodes  $|\mathcal{N} \cap \{Z \neq 1\}|/|\mathcal{N}|$  is upper bounded by  $c_{\text{misclas}}/(sB)$  for some constant  $c_{\text{misclas}} \in [8/512^2, 8]$ , that  $\nu \in [(p + 3q)/4, (3p + q)/4]$ , that  $|V| \geq 7n/8$ , and that no node in  $V$  is adjacent to a pair sampled before this call to SCREENING. Then with probability at least  $1 - 6/(sN')$ ,

(1) *the proportion  $|\mathcal{N}' \cap \{Z \neq 1\}|/|\mathcal{N}'|$  of misclassified nodes after SCREENING is upper bounded by  $c_{\text{misclas}}^{\text{after}}/(sB)$ , where*

$$\begin{aligned} c_{\text{misclas}}^{\text{after}} &= c_{\text{misclas}} \vee 8 && \text{if } N' \geq B \log(sB)^{3/2}, \text{ and} \\ c_{\text{misclas}}^{\text{after}} &= 512c_{\text{misclas}} && \text{otherwise.} \end{aligned}$$

(2) *the number of sampled bad pairs is controlled: there exists a numerical constant  $C_{\text{fail}}$  (for instance,  $C_{\text{fail}} = 26C_k + 2 = 65002$ ) such that*

$$\sum_{x \in \mathcal{A}_0} |\{y_a^x : (x, y_a^x) \text{ was sampled and } Z_{y_a^x} \neq Z_x\}| \leq C_{\text{fail}} \frac{N'}{s};$$

(3) *no node in  $\mathcal{N}'$  or  $V$  has more than  $B$  adjacent pairs sampled during this call to SCREENING;*

(4)  $|V'| \geq |V| - 4N'$ ;

(5) *it is possible to construct the core-set  $\mathcal{N}'$  with  $N'$  nodes after point (3):*

$$\sum_{j=1}^m |\mathcal{A}_I^{(j)} \cap \{Z = 1\}| \geq N'.$$

(6) *no node in  $V'$  is adjacent to a pair sampled before or during this call to SCREENING.*

Lemma 13 is proved in Section C.2.

To prove Lemma 12, we control the  $t_f$  screening calls at the second step of the constrained algorithm (Algorithm 2) as follows. For the first step, denote by  $E_0$  the event of probability  $1 - 9/(sB)$  where all the points of Lemma 11 are true. For each  $t \in \{1, \dots, t_f\}$ , denote by  $E_t$  the event where all the points of Lemma 13 are satisfied by the output of SCREENING at the  $t$ -th call, which is

$$(\mathcal{N}^{(t)}, V^{(t)}) = \text{SCREENING}(\mathcal{N}^{(t-1)}, N^{(t)}, B, \hat{\tau}, V^{(t-1)}).$$

On the event  $\bigcap_{0 \leq t \leq t_f} E_t$ , all the points of Lemma 12 can be easily derived; see Section C.1.1 for a detailed proof.

Therefore, Lemma 12 holds with a probability at least  $\mathbb{P}(\bigcap_{0 \leq t \leq t_f} E_t)$ . To prove that  $\bigcap_{0 \leq t \leq t_f} E_t$  holds with high probability, we proceed by induction. First, the event  $E_0$  holds with probability at least  $1 - 9/(sB)$  by Lemma 11. Next, for any  $t \in \{1, \dots, t_f\}$ , we check in Section C.1.2, that, on the event  $E_0 \cap \dots \cap E_{t-1}$ , the assumptions of Lemma 13 holds at the  $t$ -th call of the SCREENING routine. Hence, according to Lemma 13, conditionally on the event  $E_0 \cap \dots \cap E_{t-1}$ , the event  $E_t$

holds with probability at least  $1 - 6/(sN^{(t)})$ . By induction, we thus have

$$\begin{aligned} \mathbb{P}\left(\bigcap_{0 \leq t \leq t_f} E_t\right) &= \mathbb{P}(E_0)\mathbb{P}(E_1 | E_0) \dots \mathbb{P}(E_{t_f} | E_{t_f-1}, \dots, E_0) \\ &\geq \left(1 - \frac{9}{sB}\right) \prod_{t=0}^{t_f} \left(1 - \frac{6}{sN^{(t)}}\right), \end{aligned}$$

which is larger than

$$\begin{aligned} 1 - \frac{9}{sB} - \sum_{t=1}^{t_f} \frac{6}{sN^{(t)}} &= 1 - \frac{9}{sB} - \frac{6}{sN^{(0)}} \sum_{t=1}^{t_f-1} \lfloor \log(sB) \rfloor^{-t} - \frac{6}{s\lceil T/B \rceil} \\ &\geq 1 - \frac{9}{sB} - \frac{12 \log(sB)}{sB} \cdot \frac{\lfloor \log(sB) \rfloor^{-1}}{1 - \lfloor \log(sB) \rfloor^{-1}} - \frac{6}{s(T/B)} \\ &\geq 1 - \frac{9}{sB} - \frac{48}{sB} - \frac{6}{sB} = 1 - \frac{63}{sB}, \end{aligned}$$

using for the last inequality that  $B \leq \sqrt{T}/2$  and  $sB \geq c'_{\text{thresh}}$  for some numerical constant  $c'_{\text{thresh}} > 0$ .

To conclude, Lemma 12 holds with probability at least  $1 - 63/(sB)$ , provided that the conclusions of Lemma 12 hold on the event  $\bigcap_{0 \leq t \leq t_f} E_t$ , and that the assumptions of Lemma 13 are satisfied at each call of SCREENING. These two points are proved in the next two subsections.

**C.1.1. The conclusions of Lemma 12 hold on  $\bigcap_{0 \leq t \leq t_f} E_t$ .** Assume that the event  $\bigcap_{0 \leq t \leq t_f} E_t$  holds, and let us show that all the points of Lemma 12 are fulfilled.

(7, 8, 10) Points (8) and (10) of Lemma 12 follow directly from Lemma 13 (2), (5). As for point (7), it is satisfied when

$$4C_k C_I \frac{\log(sB)}{s} \leq \frac{B}{2 \log(sB)},$$

which holds as soon as  $sB \geq c'_{\text{thresh}}$  for some numerical constant  $c'_{\text{thresh}}$ .

(9) In the initial core-set, the proportion of misclassified nodes is upper bounded by Lemma 11 as follows:

$$|\mathcal{N}^{(0)} \cap \{Z \neq 1\}|/|\mathcal{N}^{(0)}| \leq 2\varepsilon_N \leq \frac{8}{512^2} \cdot \frac{1}{sB}.$$

For the next core-set  $\mathcal{N}^{(1)}$ , it implies that

$$|\mathcal{N}^{(1)} \cap \{Z \neq 1\}|/|\mathcal{N}^{(1)}| \leq 512 \frac{8}{512^2} \cdot \frac{1}{sB} = \frac{8}{512} \cdot \frac{1}{sB},$$

using  $c_{\text{misclas}} = 8/512^2$  in Lemma 13 (1). For the subsequent core-sets, the proportion of misclassified nodes is upper bounded as above, updating the value of  $c_{\text{misclas}}$  at each step. We thus have

$$|\mathcal{N}^{(2)} \cap \{Z \neq 1\}|/|\mathcal{N}^{(2)}| \leq 512 \frac{8}{512} \cdot \frac{1}{sB} = \frac{8}{sB},$$

and for all  $t \geq 3$ ,

$$|\mathcal{N}^{(t)} \cap \{Z \neq 1\}|/|\mathcal{N}^{(t)}| \leq \frac{8}{sB},$$

since  $N^{(t)} \geq B \log(sB)^{3/2}$  as soon as  $t \geq 3$  and  $sB \geq c'_{\text{thresh}}$  for some numerical constant  $c'_{\text{thresh}}$ .

(5) At the  $t$ -th call to SCREENING, the output of “new” nodes  $V^{(t)}$  satisfies the recursive inequality  $|V^{(t)}| \geq |V^{(t-1)}| - 4N^{(t)}$  by construction of the algorithm. The sequence of inequalities telescopes, leaving

$$|V^{(t)}| \geq |V^{(0)}| - \sum_{s=1}^t 4N^{(s)},$$

which is larger than  $7n/8$  since  $|V^{(0)}| = n - N_{\text{init}}$  and

$$N_{\text{init}} \leq \frac{n}{8} - \sum_{s=1}^{t_f-1} 4N^{(s)}$$

by Lemma 11 (1).

(6) A node can fall into four categories:

(i) it is never used;

(ii) it is used in Step 1 and possibly in the first iteration of SCREENING. Then the number of adjacent sampled pairs is at most  $N_{\text{init}} + B$  by construction of Step 1 and Lemma 13 (3), which is smaller than  $2B$  as soon as  $sB \geq c_{\text{thresh}}$  for some numerical constant  $c_{\text{thresh}}$ ;

(iii) it is used in (at most) two consecutive iterations of SCREENING (and nowhere else). Then the number of adjacent sampled pairs is at most  $2B$  by Lemma 13;

(iv) it is used in the last iteration of SCREENING and (possibly) in Step 3. Then the number of adjacent sampled pairs is at most  $B + B$  by Lemma 13 and by construction of Step 3.

**C.1.2. Checking the assumptions of Lemma 13.** Assume that the events  $E_0, \dots, E_{t-1}$  hold together, and let us check the assumptions of Lemma 13.

First, the condition  $sB \geq c'_{\text{thresh}}$  comes from Lemma 12. Then, following the proof of point (9), we can check that

$$|\mathcal{N}^{(t-1)} \cap \{Z \neq 1\}|/|\mathcal{N}^{(t-1)}| \leq c_{\text{misclas}}/(sB)$$

for  $c_{\text{misclas}} \in [8/512^2, 8]$ . For the threshold  $\hat{\tau}$  taking a value in  $[(p+3q)/4, (3p+q)/4]$ , it is stated in Lemma 11. The input of “new” nodes  $V^{(t-1)}$  satisfies

$$|V^{(t-1)}| \geq \frac{7n}{8},$$

as seen above in the proof of point (5). Finally, the inequality

$$B \leq 4N^{(t)} \leq 4N^{(t-1)} \log(sB)$$

is satisfied by construction of the algorithm, as soon as  $sB \geq c'_{\text{thresh}}$  for some numerical constant  $c'_{\text{thresh}}$ .

## C.2. Proof of Lemma 13: Control of SCREENING

In this section, we work conditionally to  $\mathcal{F}_{T_{\text{start}}}$  where  $T_{\text{start}}$  is the number of pairs sampled before the current call to SCREENING.

Let us state the two main technical results that allow to prove Lemma 13. Write  $\mathcal{V}(x) := \mathcal{V}_j$  for each  $j \in \{1, \dots, m\}$  and  $x \in \mathcal{A}_0^{(j)}$ . The first one controls the properties of the sets  $(\mathcal{V}(x))_{x \in \mathcal{A}_0}$ . Given a subset of nodes  $S$ , denote by  $\text{misclas}(S)$  the set of misclassified nodes in  $S$ , that is the set of all  $x \in S$  such that  $Z_x \neq 1$  in SCREENING.

**Lemma 14.** *The sets  $(\mathcal{V}(x))_{x \in \mathcal{A}_0}$  satisfy the following:*

- (1) for all  $y \in \mathcal{N}$ ,  $|\{x \in \mathcal{A}_0 : y \in \mathcal{V}(x)\}| \leq B$ ;
- (2)  $\mathbb{P}(|\{x \in \mathcal{A}_0 : |\text{misclas}(\mathcal{V}(x))| \geq kI/16\}| \geq c_{\text{misclas}}^{\text{after}}/2 \cdot N'/sB) \leq 2/sN'$ ,  
where  $c_{\text{misclas}}^{\text{after}}$  is defined as in Lemma 13;
- (3)  $\sum_{x \in \mathcal{A}_0} |\text{misclas}(\mathcal{V}(x))| \leq N'/s$ .

The proof of the above lemma is postponed to Section C.3. The next lemma allows one to control the effectiveness of point (3) of SCREENING. Its proof follows the same lines as those of Lemma 10 (for proving (C.1)) and Lemma 9 (10) (for proving (C.2)), it is therefore omitted.

**Lemma 15.** *Conditionally to the choice of the set  $\mathcal{A}_0$  and  $(\mathcal{V}(x))_{x \in \mathcal{A}_0}$ , the variables  $(T_x)_{x \in \mathcal{A}_0}$  are independent and for all  $x \in \mathcal{A}_0$  and all  $i \in \{1, \dots, I\}$ ,*

$$\mathbb{P}\left(T_x \geq i \mid Z_x \neq 1 \text{ and } |\text{misclas}(\mathcal{V}(x))| \leq \frac{kI}{16}\right) \leq e^{-i}, \quad (\text{C.1})$$

$$\mathbb{P}\left(T_x \geq I \mid Z_x = 1 \text{ and } |\text{misclas}(\mathcal{V}(x))| \leq \frac{kI}{16}\right) \geq \frac{3}{4}. \quad (\text{C.2})$$

Let us now prove Lemma 13.

*Proof of Lemma 13.* Note that points (4) and (6) follow from the construction of the algorithm and that point (3) follows straightforwardly from Lemma 14 (1) (for the nodes from  $\mathcal{N}$ ) and from the construction of the algorithm (for the nodes from  $\mathcal{A}_0$ ).

(1) We first observe that

$$\begin{aligned} |\mathcal{N}' \cap \{Z \neq 1\}| &\leq \sum_{j=1}^m |\text{misclas}(\mathcal{A}_I^{(j)})| \\ &= \sum_{j=1}^m \sum_{\substack{x \in \mathcal{A}_0 \text{ s.t.} \\ |\text{misclas}(\mathcal{V}(x))| > kI/16}} \mathbf{1}_{x \in \mathcal{A}_I^{(j)} \text{ and } Z_x \neq 1} + \sum_{j=1}^m \sum_{\substack{x \in \mathcal{A}_0 \text{ s.t.} \\ |\text{misclas}(\mathcal{V}(x))| \leq kI/16}} \mathbf{1}_{x \in \mathcal{A}_I^{(j)} \text{ and } Z_x \neq 1} \\ &\leq \sum_{x \in \mathcal{A}_0} \mathbf{1}_{|\text{misclas}(\mathcal{V}(x))| \geq kI/16} + \sum_{\substack{x \in \mathcal{A}_0 \text{ s.t.} \\ |\text{misclas}(\mathcal{V}(x))| \leq kI/16}} \mathbf{1}_{T_x \geq I \text{ and } Z_x \neq 1} \\ &\leq \frac{c_{\text{misclas}}^{\text{after}}}{2} \frac{N'}{sB} + \sum_{\substack{x \in \mathcal{A}_0 \text{ s.t.} \\ |\text{misclas}(\mathcal{V}(x))| \leq kI/16}} \mathbf{1}_{T_x \geq I \text{ and } Z_x \neq 1} \end{aligned}$$

with probability at least  $1 - 2/(sN')$  by Lemma 14 (2).

The second term is dominated by a binomial random variable with parameters  $(|\mathcal{A}_0|, e^{-I})$  by Lemma 15, so it is dominated by a binomial random variable  $X$  with parameters  $(4N', 1/(sB)^{1026})$  since  $I \geq 1026 \log(sB)$ . Equation (D.5) implies that for  $sB \geq 512$  (which is implied by  $c'_{\text{thresh}} \geq 512$ ),

$$\begin{aligned} \mathbb{P}\left(\frac{1}{4N'} X \geq \frac{1}{512sB}\right) &\leq \exp\left(-\frac{1}{2} 4N' \frac{1}{512sB} \log\left(\frac{(sB)^{1025}}{512}\right)\right) \\ &\leq \exp\left(-4N' \frac{\log(sB)}{sB}\right). \end{aligned}$$

We want this probability to be smaller than  $1/(sN')$ , that is

$$\frac{\log(sB)}{sB} \geq s \frac{\log(sN')}{4sN'},$$

which is true since  $s \leq 1$ , and the function  $x \mapsto (\log x)/x$  is non-increasing for  $x \geq e$ , and  $4sN' \geq sB \geq e$  by assumption. Therefore,

$$\mathbb{P}\left(|\mathcal{N}' \cap \{Z \neq 1\}| \geq \frac{c_{\text{misclas}}^{\text{after}} + (8/512)}{2} \cdot \frac{N'}{sB}\right) \leq \frac{3}{sN'},$$

which implies point (1) (since  $c_{\text{misclas}}^{\text{after}} \geq 8/512$  by definition).

(2) Given a subset  $S$  of  $\mathcal{A}_0$ , denote by  $\text{bad}(S)$  the number of sampled bad pairs coming from nodes in  $S$  during SCREENING, that is

$$\text{bad}(S) = \sum_{x \in S} |\{y_i^x : i \leq k((T_x + 1) \wedge I) \text{ and } Z_{y_i^x} \neq Z_x\}|.$$

The total number of bad pairs sampled during SCREENING can be decomposed into

$$\begin{aligned} \text{bad}(\mathcal{A}_0) &= \sum_{x \in \mathcal{A}_0} \text{bad}(\{x\}) \mathbf{1}_{|\text{misclas}(\mathcal{V}(x))| > kI/16} \\ &\quad + \sum_{x \in \mathcal{A}_0 \cap \{Z=1\}} \text{bad}(\{x\}) \mathbf{1}_{|\text{misclas}(\mathcal{V}(x))| \leq kI/16} \\ &\quad + \sum_{x \in \mathcal{A}_0 \cap \{Z \neq 1\}} \text{bad}(\{x\}) \mathbf{1}_{|\text{misclas}(\mathcal{V}(x))| \leq kI/16} \\ &\leq kI \sum_{x \in \mathcal{A}_0} \mathbf{1}_{|\text{misclas}(\mathcal{V}(x))| \geq kI/16} \end{aligned} \quad (\text{C.3})$$

$$+ \sum_{x \in \mathcal{A}_0 \cap \{Z=1\}} |\text{misclas}(\mathcal{V}(x))| \quad (\text{C.4})$$

$$+ \sum_{x \in \mathcal{A}_0 \cap \{Z \neq 1\}} k(T_x + 1) \mathbf{1}_{|\text{misclas}(\mathcal{V}(x))| \leq kI/16} \quad (\text{C.5})$$

The first sum is controlled by Lemma 14(2):

$$\mathbb{P}\left((\text{C.3}) \geq \frac{c_{\text{misclas}}^{\text{after}}}{2} \frac{N'kI}{sB}\right) \leq \frac{2}{sN'}.$$

Thus, since  $kI/B \leq 4C_k C_I \log(sB)/(sB)$  by definition, there exists a constant  $c'_{\text{thresh}}$  such that  $kI/B \leq 2/c_{\text{misclas}}^{\text{after}}$  as soon as  $sB \geq c'_{\text{thresh}}$ , so that

$$\mathbb{P}\left((\text{C.3}) \geq \frac{N'}{s}\right) \leq \frac{2}{sN'}.$$

Likewise, by Lemma 14(3),

$$(\text{C.4}) \leq \frac{N'}{s}.$$

By Lemma 15, the variables  $T_x$  in the third sum are stochastically dominated by i.i.d. exponential random variables with parameter 1. Therefore, using the inequality  $k \leq 2C_k/s$ , the term (C.5) is stochastically dominated by

$$8C_k \frac{N'}{s} + 2 \frac{C_k}{s} \sum_{i=1}^{4N'} Y_i,$$

where  $(Y_i)_{i \in \mathbb{N}^*}$  are i.i.d. exponential random variables with parameter 1. These exponential random variables satisfy

$$\mathbb{E}(Y_i - 1)^2 \leq 1,$$

and for all  $a \in \mathbb{N}$  such that  $a \geq 3$ , we have

$$\mathbb{E}(Y_i - 1)_+^a \leq a!.$$

It follows that Bernstein’s inequality (see, for instance, [34, Proposition 2.9]) entails for all  $t > 0$ ,

$$\mathbb{P} \left( \sum_{i=1}^{4N'} Y_i - 4N' \geq 4\sqrt{N't} + t \right) \leq e^{-t}.$$

Thus, by taking  $t = N'$ , with probability at least  $1 - e^{-N'} \geq 1 - 1/N' \geq 1 - 1/(sN')$ , we have

$$\sum_{i=1}^{4N'} Y_i \leq 9N'.$$

Hence, with probability at least  $1 - 3/(sN')$ ,

$$\text{bad}(\mathcal{A}_0) \leq (26C_k + 2) \frac{N'}{s}.$$

(5) Write  $\mathcal{A}_I = \bigcup_{j=1}^m \mathcal{A}_I^{(j)}$  and for each  $x \in \mathcal{A}_0$ , let  $V_x = \mathbf{1}_{x \in \mathcal{A}_I}$  indicate whether  $x$  has been kept until the end of point (3) of SCREENING. Lemma 15 ensures that the random variables  $(V_x)_{x \in \mathcal{A}_0 \cap \{Z=1\}}$  s.t.  $|\text{misclas}(\mathcal{V}(x))| \leq kI/16$  dominate i.i.d. Bernoulli random variables with parameter  $3/4$ . Therefore, Hoeffding’s inequality (D.1) entails

$$\mathbb{P} \left( |\mathcal{A}_I \cap \{Z = 1\}| \leq \frac{3|\mathcal{A}_0 \cap \{Z = 1\} \cap \{x : |\text{misclas}(\mathcal{V}(x))| \leq kI/16\}}{4} - \sqrt{|\mathcal{A}_0| \frac{\log(sN')}{2}} \right) \leq \frac{1}{sN'}.$$

Note that

$$\left| \left\{ x \in \mathcal{A}_0 \text{ s.t. } |\text{misclas}(\mathcal{V}(x))| > \frac{kI}{16} \right\} \right| \leq \frac{c_{\text{misclas}}^{\text{after}}}{2} \cdot \frac{N'}{(sB)}$$

with probability at least  $1 - 2/(sN')$  by Lemma 14. Since  $|\mathcal{A}_0| = 4N'$ , the previous equation entails

$$\mathbb{P} \left( |\mathcal{A}_I \cap \{Z = 1\}| \leq \frac{3|\mathcal{A}_0 \cap \{Z = 1\}}{4} - \frac{3c_{\text{misclas}}^{\text{after}} N'}{8sB} - \sqrt{4N' \frac{\log(sN')}{2}} \right) \leq \frac{3}{sN'}.$$

Let us assume for now that  $|\mathcal{A}_0 \cap \{Z = 1\}| \geq (11/7)N'$  with probability at least  $1 - 1/(sN')$ . Then this ensures that for  $N'$  and  $sB$  larger than some numerical constants (which is guaranteed by  $B \geq B_0$  and  $sB \geq c'_{\text{thresh}}$ ),

$$\mathbb{P}\left(|\mathcal{A}_I \cap \{Z = 1\}| \leq N'\right) \leq \frac{4}{sN'},$$

which gives point (5), provided that  $|\mathcal{A}_0 \cap \{Z = 1\}| \geq (11/7)N'$ .

The random variable  $|\mathcal{A}_0 \cap \{Z = 1\}|$  is an hypergeometric random variable with number of draws  $4N'$  and initial probability of a winning draw  $r' \in [3/7, 4/7]$  because the number of nodes that have not been sampled at the start of SCREENING is bigger than  $7n/8$  by assumption and because the true communities are balanced.

Therefore, Hoeffding's inequality (D.1) implies

$$\mathbb{P}\left(|\mathcal{A}_0 \cap \{Z = 1\}| \leq \frac{3}{7}4N' - \sqrt{\frac{4N' \log(sN')}{2}}\right) \leq \frac{1}{sN'},$$

so that for  $N'$  large enough (which is implied by  $B \geq B_0$  for some numerical constant  $B_0$ ),

$$\mathbb{P}\left(|\mathcal{A}_0 \cap \{Z = 1\}| \leq \frac{11}{7}N'\right) \leq \frac{1}{sN'}. \quad \blacksquare$$

### C.3. Proof of Lemma 14

(1, 3) To check point (1), it suffices to check that  $\lceil 4N'/m \rceil \leq B$ . For  $sB \geq c'_{\text{thresh}}$  with a numerical constant  $c'_{\text{thresh}}$  large enough, one has

$$m = \left\lfloor \frac{N}{kI} \right\rfloor \geq \frac{N}{2kI}$$

and

$$\left\lceil \frac{4N'}{m} \right\rceil \leq \frac{16N'kI}{N} \leq 64C_k C_I \frac{(\log(sB))^2}{s} \leq B. \quad (\text{C.6})$$

For point (3), note that

$$\begin{aligned} \sum_{x \in \mathcal{A}_0} |\text{misclas}(\mathcal{V}(x))| &\leq \left\lceil \frac{4N'}{m} \right\rceil \sum_{j=1}^m |\text{misclas}(\mathcal{V}_j)| \leq 16kI \frac{N'}{N} |\text{misclas}(\mathcal{N})| \\ &\leq 64C_k C_I \frac{\log(sB)}{s} c_{\text{misclas}} \frac{N'}{sB} \leq 64C_k C_I \frac{\log(sB)}{sB} c_{\text{misclas}}^{\text{after}} \frac{N'}{s} \\ &\leq \frac{N'}{s} \end{aligned}$$

by assumption on the number of misclassified nodes in  $\mathcal{N}$ , and as soon as  $sB \geq c'_{\text{thresh}}$  for some numerical constant  $c'_{\text{thresh}}$ .

(2) *Small core-sets.* In this section, we assume  $N' < B \log(sB)^{3/2}$ . By equation (C.6), we have

$$\begin{aligned} \left| \left\{ x \in \mathcal{A}_0 : |\text{misclas}(\mathcal{V}(x))| \geq \frac{kI}{16} \right\} \right| &\leq \left\lceil \frac{4N'}{m} \right\rceil \sum_{j=1}^m \mathbf{1}_{|\text{misclas}(\mathcal{V}_j)| \geq kI/16} \\ &\leq 16 \frac{N'}{N} kI \sum_{j=1}^m \mathbf{1}_{|\text{misclas}(\mathcal{V}_j)| \geq kI/16}. \end{aligned}$$

Note that

$$\frac{kI}{16} \sum_{j=1}^m \mathbf{1}_{|\text{misclas}(\mathcal{V}_j)| \geq kI/16} \leq \sum_{j=1}^m |\text{misclas}(\mathcal{V}_j)| = |\text{misclas}(\mathcal{N})| \leq \frac{c_{\text{misclas}} N}{sB}$$

by assumption, so that

$$\begin{aligned} \left| \left\{ x \in \mathcal{A}_0 : |\text{misclas}(\mathcal{V}(x))| \geq \frac{kI}{16} \right\} \right| &\leq 16 \frac{N'}{N} kI \cdot \frac{16}{kI} \cdot \frac{c_{\text{misclas}} N}{sB} \\ &= 256 c_{\text{misclas}} \frac{N'}{sB} = \frac{c_{\text{misclas}}^{\text{after}}}{2} \cdot \frac{N'}{sB}. \end{aligned}$$

This bound is not random, it holds with probability 1.

(2) *Large core-sets.* In this section, we assume  $N' \geq B \log(sB)^{3/2}$ .

The number of misclassified nodes in each  $\mathcal{V}_j$  can be controlled more easily by introducing a coupling with i.i.d. Bernoulli random variables. Note that this coupling is a theoretical tool and does not appear in the algorithm.

**Lemma 16.** *Let  $K$  be a random variable taking values in  $\{0, \dots, N\}$ . Let  $(X_x)_{x \in \mathcal{N}}$  be a vector of random variables taking values in  $\{0, 1\}$  such that*

- $\sum_{x \in \mathcal{N}} X_x = K$ ;
- *the distribution of  $(X_x)_{x \in \tilde{\mathcal{N}}}$  is invariant under permutation of  $\mathcal{N}$ .*

*Note that these two points together with the distribution of  $K$  characterize the distribution of  $(X_x)_{x \in \mathcal{N}}$ . Then for all  $u > 0$ , there exists a coupling with i.i.d. Bernoulli random variables  $(Y_x)_{x \in \mathcal{N}}$  with parameter  $u$  such that by writing  $M = \sum_{x \in \mathcal{N}} Y_x$ ,  $M$  is independent of  $(X_x)_{x \in \mathcal{N}}$  and*

$$M \geq K \implies (\forall x \in \mathcal{N}, X_x \leq Y_x). \tag{C.7}$$

*Proof.* Let  $M$  be a binomial random variable with parameters  $(N, u)$  such that  $M$  and  $K$  are independent. Let  $(\tilde{X}_i)_{1 \leq i \leq N}$  and  $(\tilde{Y}_i)_{1 \leq i \leq N}$  be random variables such that conditionally to  $M$  and  $K$  and for all  $1 \leq i \leq N$ ,

$$\tilde{X}_i = \begin{cases} 1 & \text{if } i \leq K, \\ 0 & \text{otherwise,} \end{cases} \quad \tilde{Y}_i = \begin{cases} 1 & \text{if } i \leq M, \\ 0 & \text{otherwise.} \end{cases}$$

Let  $\sigma$  be a uniform random variable in the set of bijections from  $\{1, \dots, N\}$  to  $\mathcal{N}$  that is independent of  $K, M, (\tilde{X}_i)_i$  and  $(\tilde{Y}_i)_i$ , and define  $X'_x = \tilde{X}_{\sigma^{-1}(x)}$  and  $Y_x = \tilde{Y}_{\sigma^{-1}(x)}$  for all  $x \in \mathcal{N}$ .

Then the random vector  $(X'_x)_{x \in \mathcal{N}}$  has the same distribution as the random vector  $(X_x)_{x \in \mathcal{N}}$ , the random variables  $(Y_x)_{x \in \mathcal{N}}$  are i.i.d. Bernoulli random variables with parameter  $u$ , and equation (C.7) holds for these two vectors. ■

Let  $M$  and  $(Y_x)_{x \in \mathcal{N}}$  be the random variables given by Lemma 16 applied to

$$(X_x)_{x \in \mathcal{N}} = (\mathbf{1}_{\hat{Z}_x \neq Z_x})_{x \in \mathcal{N}}, \quad K = |\text{misclas}(\mathcal{N})|, \quad u = \frac{2c_{\text{misclas}}^{\text{after}}}{(sB)} + \frac{4 \log(sB)^2}{B}.$$

Note that the algorithm is invariant by permutation of the nodes of  $\mathcal{N}$ , so that we may assume without loss of generality that the distribution of these  $(X_x)_{x \in \mathcal{N}}$  is invariant by permutation of  $\mathcal{N}$ .

By Assumption of Lemma 13, we have  $K \leq c_{\text{misclas}}/(sB)$ . Let us show that  $M \geq c_{\text{misclas}}/(sB)$  with probability at least  $1 - 1/(sN')$ , which implies  $M \geq K$  with probability at least  $1 - 1/(sN')$ . Since  $M$  is a binomial random variable with parameters  $(N, u)$ , Bernstein's inequality (D.2) entails

$$\mathbb{P}(M \leq Nu - \sqrt{2Nut} - t) \leq e^{-t}.$$

Since  $\sqrt{2ab} \leq a/2 + b$  for all  $a, b > 0$ , it holds with probability at least  $1 - 1/(sN')$  that

$$\begin{aligned} M &\geq Nu - \frac{Nu}{2} - \log(sN') - \log(sN') \\ &\geq \frac{c_{\text{misclas}}N}{sB} + 2N \frac{\log(sB)^2}{B} - 2 \log(sN'). \end{aligned}$$

Note that

$$\frac{2 \log(sN')}{2N \log(sB)^2/B} \leq \frac{\log(sN \log(sB))/N}{\log(sB)^2/B} = \frac{\log(sN \log(sB))/sN \log(sB)}{\log(sB)/sB} \leq 1,$$

as soon as  $sB \geq e$  since the application  $x \mapsto (\log x)/x$  is non-increasing for  $x \geq e$  and  $sN \log(sB) \geq sB \geq e$  (the second to last inequality comes from the assumption  $N' = N \lfloor \log(sB) \rfloor \geq B \log(sB)^{3/2}$  made at the beginning of the current subsection). Therefore,

$$\mathbb{P}\left(M \leq \frac{c_{\text{misclas}}N}{sB}\right) \leq \frac{1}{sN'},$$

and finally, according to Lemma 16,

$$\mathbb{P}(\forall x \in \mathcal{N}, \mathbf{1}_{\hat{Z}_x \neq Z_x} \leq Y_x) \geq 1 - \frac{1}{sN'}. \tag{C.8}$$

We can now proceed to the conclusion of the proof of point (2) when  $N' \geq B \log(sB)^{3/2}$ . We have

$$\begin{aligned} \left| \left\{ x \in \mathcal{A}_0 : |\text{misclas}(\mathcal{V}(x))| \geq \frac{kI}{16} \right\} \right| &\leq \left\lceil \frac{4N'}{m} \right\rceil \sum_{j=1}^m \mathbf{1}_{|\text{misclas}(\mathcal{V}_j)| \geq kI/16} \\ &\leq 16 \frac{N'}{N} kI \sum_{j=1}^m \mathbf{1}_{\sum_{x \in \mathcal{V}_j} Y_x \geq kI/16} \end{aligned}$$

with probability at least  $1 - 1/(sN')$  by equations (C.6) and (C.8).

Note that the random variable  $\sum_{j=1}^m \mathbf{1}_{\sum_{x \in \mathcal{V}_j} Y_x \geq kI/16}$  is a binomial random variable with parameters

$$\left( m, \mathbb{P} \left( \sum_{x \in \mathcal{V}_j} Y_x \geq \frac{kI}{16} \right) \right),$$

and that  $\sum_{x \in \mathcal{V}_j} Y_x$  is a binomial random variable with parameters  $(kI, u)$  with

$$u = \frac{2c_{\text{misclas}}^{\text{after}}}{(sB)} + \frac{4 \log(sB)^2}{B}.$$

Since  $5u \leq 1/16$  for  $sB \geq c'_{\text{thresh}}$ , we can apply equation (D.5) to obtain

$$\mathbb{P} \left( \sum_{x \in \mathcal{V}_j} Y_x \geq \frac{kI}{16} \right) \leq \exp \left( -\frac{kI}{32} \log \frac{1}{16u} \right). \tag{C.9}$$

Note that

$$\begin{aligned} \log \frac{1}{16u} &\geq \log \frac{sB}{256(1 \vee (s \log(sB)^2))} \\ &= \log(sB) - \log 256 - 0 \vee \log(s \log(sB)^2) \\ &\geq \frac{2}{3} \log(sB) - 0 \vee \log((sB)^{1/3}) \\ &\geq \frac{2}{3} \log(sB) - \frac{1}{3} \log(sB) = \frac{1}{3} \log(sB), \end{aligned}$$

when  $sB \geq c'_{\text{thresh}}$  for  $c'_{\text{thresh}}$  large enough. Therefore, equation (C.9) implies

$$\mathbb{P} \left( \sum_{x \in \mathcal{V}_j} Y_x \geq \frac{kI}{16} \right) \leq \exp \left( -\frac{kI}{96} \log(sB) \right).$$

It remains to control the probability that a binomial random variable with parameters

$$\left( m, \exp \left( -\frac{kI}{96} \log(sB) \right) \right)$$

exceeds

$$\frac{c_{\text{misclas}}^{\text{after}}}{2} \cdot \frac{N'/(sB)}{16kIN'/N}.$$

To apply equation (C.9), check that

$$\begin{aligned} & \frac{c_{\text{misclas}}^{\text{after}}/2 \cdot ((N'/(sB))/(16kIN'/N))}{m \exp(-(kI/96) \log(sB))} \\ &= \frac{c_{\text{misclas}}^{\text{after}}}{2} \cdot \frac{N'/m}{16kIN'/N} \cdot \frac{1}{sB} \exp\left(\frac{kI}{96} \log(sB)\right) \\ &\geq \frac{c_{\text{misclas}}^{\text{after}}}{2} \cdot \frac{1}{16sB} \exp\left(\frac{kI}{96} \log(sB)\right) \quad [\text{since } m \leq N/kI] \\ &\geq \exp\left(\frac{kI}{200} \log(sB)\right) \geq 5, \end{aligned}$$

for  $sB \geq c'_{\text{thresh}}$ . Thus, equation (D.5) and  $c_{\text{misclas}}^{\text{after}} = c_{\text{misclas}} \vee 8$  imply

$$\begin{aligned} & \mathbb{P}\left(16 \frac{N'}{N} kI \sum_{j=1}^m \mathbf{1}_{\sum_{x \in \mathcal{V}_j} Y_x \geq kI/16} \geq \frac{c_{\text{misclas}} \vee 8}{2} \cdot \frac{N'}{sB}\right) \\ &\leq \exp\left(-\frac{c_{\text{misclas}} \vee 8}{4} \cdot \frac{N'/(sB)}{16kIN'/N} \cdot \frac{kI}{200} \log(sB)\right) \\ &\leq \exp\left(-\frac{N \log(sB)}{1600sB}\right) \\ &\leq \exp\left(-\frac{N'}{1600sB}\right) \quad [\text{since } N \log(sB) \geq N']. \end{aligned}$$

We want this probability to be smaller than  $1/(sN')$ , that is

$$\frac{N'}{1600sB} \geq \log(sN'),$$

which holds as soon as  $N' \geq [2 \cdot 1600sB \log(1600s^2B)]$ , which is implied by the assumption  $N' \geq B \log(sB)^{3/2}$  for  $sB \geq c'_{\text{thresh}}$ . Thus,

$$\mathbb{P}\left(\left|\left\{x \in \mathcal{A}_0 : |\text{misclas}(\mathcal{V}(x))| \geq \frac{kI}{16}\right\}\right| \geq \frac{c_{\text{misclas}}^{\text{after}}}{2} \cdot \frac{N'}{sB}\right) \leq \frac{2}{sN'}.$$

The proof is complete. ■

## D. Probabilistic inequalities

We recall Bernstein and Hoeffding inequalities for binomial and hypergeometric distributions.

**Lemma 17.** For  $n \geq 1$ ,  $p \in [0, 1]$  and  $N \geq n$ , let  $X$  be either a binomial random variable with parameters  $(n, p)$  or a sum of  $m$  i.i.d. hypergeometric random variables with parameters  $(n/m, p, N)$ . Then, for all  $t > 0$ ,

$$\mathbb{P}\left(X - np \geq \sqrt{\frac{nt}{2}}\right) \leq e^{-t} \quad \text{and} \quad \mathbb{P}\left(|X - np| \geq \sqrt{\frac{nt}{2}}\right) \leq 2e^{-t}, \quad (\text{D.1})$$

and

$$\mathbb{P}(X - np \geq \sqrt{2npt} + t) \leq e^{-t}, \quad (\text{D.2})$$

$$\mathbb{P}(|X - np| \geq \sqrt{2npt} + t) \leq 2e^{-t}. \quad (\text{D.3})$$

The following lemma allows to control large deviations of binomial and hypergeometric random variables.

**Lemma 18.** Let  $X$  be either a binomial random variable with parameters  $(n, p)$  or a sum of  $m$  i.i.d. hypergeometric random variables with parameters  $(n/m, p, N)$ . Then for all  $c \in [p, 1]$ ,

$$\mathbb{P}(X \geq nc) \leq e^{-n \cdot kl(c,p)}, \quad (\text{D.4})$$

where  $kl(c, p) = c \log(c/p) + (1 - c) \log((1 - c)/(1 - p))$ . In particular, if  $c \geq 5p$ ,

$$\mathbb{P}(X \geq nc) \leq e^{-\frac{1}{2}nc \log(c/p)}. \quad (\text{D.5})$$

*Proof of Lemma 18.* The large deviation inequality (D.4) is derived by the classical Cramèr–Chernoff’s method; see, for instance, [34, Chapter 2].

For inequality (D.5), note that for all  $0 < \alpha < 1/p$ ,

$$\begin{aligned} kl(\alpha p, p) &= \frac{\alpha p}{2} \log \alpha + \left[ \frac{\alpha p}{2} \log \alpha - (1 - \alpha p) \log \frac{1 - p}{1 - \alpha p} \right] \\ &= \frac{\alpha p}{2} \log \alpha + \left[ \frac{\alpha p}{2} \log \alpha - (1 - \alpha p) \log \left( 1 + p \frac{\alpha - 1}{1 - \alpha p} \right) \right] \\ &\geq \frac{\alpha p}{2} \log \alpha + \left[ \frac{\alpha p}{2} \log \alpha - p(\alpha - 1) \right] \\ &\geq \frac{\alpha p}{2} \log \alpha + p \left[ \frac{\alpha \log \alpha}{2} + 1 - \alpha \right], \end{aligned}$$

and the term inside the square brackets is positive as soon as  $\alpha \geq 5$ . ■

We also recall some classical controls on the Kullback–Leibler divergence between two Bernoulli distribution.

**Lemma 19.** For any  $p_1, p_2 \in [0, 1]$ ,

$$\frac{(p_1 - p_2)^2}{p_1 \vee p_2} \leq kl(p_1, p_2) \leq \frac{(p_1 - p_2)^2}{p_1(1 - p_1) \wedge p_2(1 - p_2)}.$$

In particular, for any  $q \leq p \leq 1/2$ ,

$$\begin{aligned} s &= \frac{(p-q)^2}{p+q} \leq \frac{(p-q)^2}{p} \\ &\leq kl(p, q) \vee kl(q, p) \leq \frac{2(p-q)^2}{q} = 2\left(1 + \frac{p}{q}\right)s. \end{aligned}$$

**Miscellaneous inequalities.** The following inequality is used repeatedly in the proofs.

**Lemma 20.** For all  $x > 0$  and  $y \geq 0$ ,

$$x \geq (2y \log y) \vee e \implies \frac{x}{\log x} \geq y.$$

**Funding.** Christophe Giraud acknowledges partial support by grant ANR-19-CHIA-0021-01 (BiSCottE, ANR), and by grant ANR-21-CE23-0035 (ASCAI, ANR and DFG). The work of Yann Issartel was supported by an AMX scholarship, via École Polytechnique.

## References

- [1] E. Abbe, Community detection and stochastic block models: Recent developments. *J. Mach. Learn. Res.* **18** (2017), article no. 177 Zbl 1403.62110 MR 3827065
- [2] E. Abbe and C. Sandon, [Community detection in general stochastic block models: Fundamental limits and efficient algorithms for recovery](#). In *2015 IEEE 56th Annual Symposium on Foundations of Computer Science—FOCS 2015*, pp. 670–688, IEEE Computer Soc., Los Alamitos, CA, 2015 MR 3473334
- [3] E. Abbe and C. Sandon, Detection in the stochastic block model with multiple clusters: Proof of the achievability conjectures, acyclic BP, and the information-computation gap. [v1] 2015, [v4] 2016, arXiv:1512.09080
- [4] J. Banks, C. Moore, J. Neeman, and P. Netrapalli, Information-theoretic thresholds for community detection in sparse networks. In *29th Annual Conference on Learning Theory*, pp. 383–416, Proceedings of Machine Learning Research 49, PMLR, 2016.
- [5] C. Bordenave, M. Lelarge, and L. Massoulié, [Nonbacktracking spectrum of random graphs: Community detection and nonregular Ramanujan graphs](#). *Ann. Probab.* **46** (2018), no. 1, 1–71 Zbl 1386.05174 MR 3758726
- [6] R. A. Bradley and M. E. Terry, [Rank analysis of incomplete block designs. I. The method of paired comparisons](#). *Biometrika* **39** (1952), 324–345 Zbl 0047.12903 MR 0070925
- [7] A. N. Burnetas and M. N. Katehakis, [Optimal adaptive policies for sequential allocation problems](#). *Adv. in Appl. Math.* **17** (1996), no. 2, 122–142 Zbl 0854.60032 MR 1390571
- [8] F. Caron and A. Doucet, [Efficient Bayesian inference for generalized Bradley–Terry models](#). *J. Comput. Graph. Statist.* **21** (2012), no. 1, 174–196 MR 2913362

- [9] D. Chakrabarti, R. Kumar, F. Radlinski, and E. Upfal, Mortal multi-armed bandits. In *Advances in Neural Information Processing Systems 21*, pp. 273–280, Curran Associates, Red Hook, NY, 2008
- [10] P. Chin, A. Rao, and V. Vu, Stochastic block model and community detection in sparse graphs: A spectral algorithm with optimal rate of recovery. In *Proceedings of The 28th Conference on Learning Theory*, pp. 391–423, Proceedings of Machine Learning Research 40, PMLR, 2015
- [11] Y. Chen and J. Xu, Statistical-computational tradeoffs in planted problems and submatrix localization with a growing number of clusters and submatrices. *J. Mach. Learn. Res.* **17** (2016), article no. 27 Zbl [1360.62320](#) MR [3491121](#)
- [12] Z. Chen, S. Xue, J. Kolen, N. Aghdaie, K. A. Zaman, Y. Sun, and M. S. El-Nasr, [EOMM: An engagement optimized matchmaking framework](#). In *Proceedings of the 26th International Conference on World Wide Web*, pp. 1143–1150, International World Wide Web Conferences Steering Committee, Geneva, Switzerland, 2017
- [13] S. L. Corff, M. Lerasle, and E. Vernet, A Bayesian nonparametric approach for generalized Bradley–Terry models in random environment. 2018, arXiv:[1808.08104v1](#)
- [14] A. Decelle, F. Krzakala, C. Moore, and L. Zdeborová, [Asymptotic analysis of the stochastic block model for modular networks and its algorithmic applications](#). *Phys. Rev. E* **84** (2011), article no. 066106
- [15] R. de Heide, J. Cheshire, P. Ménard, and A. Carpentier, Bandits with many optimal arms. In *Advances in Neural Information Processing Systems 34*, pp. 22457–22469, Curran Associates, Red Hook, NY, 2021
- [16] Y. Fei and Y. Chen, [Exponential error rates of SDP for block models: Beyond Grothendieck’s inequality](#). *IEEE Trans. Inform. Theory* **65** (2019), no. 1, 551–571 Zbl [1432.90102](#) MR [3901009](#)
- [17] C. Gao, Z. Ma, A. Y. Zhang, and H. H. Zhou, Achieving optimal misclassification proportion in stochastic block models. *J. Mach. Learn. Res.* **18** (2017), article no. 60 Zbl [1440.62244](#) MR [3687603](#)
- [18] A. Garivier, P. Ménard, and G. Stoltz, [Explore first, exploit next: The true shape of regret in bandit problems](#). *Math. Oper. Res.* **44** (2019), no. 2, 377–399 Zbl [1435.62308](#) MR [3959077](#)
- [19] C. Giraud and N. Verzelen, [Partial recovery bounds for clustering with the relaxed  \$K\$ -means](#). *Math. Stat. Learn.* **1** (2018), no. 3-4, 317–374 Zbl [1426.62186](#) MR [4059724](#)
- [20] S. Goschin, A. Weinstein, M. L. Littman, and E. Chastain, Planning in reward-rich domains via pac bandits. In *Proceedings of the Tenth European Workshop on Reinforcement Learning*, pp. 25–42, Proceedings of Machine Learning Research 24, PMLR, 2013
- [21] R. Heckel, N. B. Shah, K. Ramchandran, and M. J. Wainwright, [Active ranking from pairwise comparisons and when parametric assumptions do not help](#). *Ann. Statist.* **47** (2019), no. 6, 3099–3126 Zbl [1435.62293](#) MR [4025736](#)
- [22] R. Herbrich, T. Minka, and T. Graepel, [TrueSkill™: A Bayesian skill rating system](#). In *Advances in Neural Information Processing Systems 19: Proceedings of the 2006 Conference*, pp. 569–576, MIT Press, Cambridge, MA, 2007

- [23] P. W. Holland, K. B. Laskey, and S. Leinhardt, [Stochastic blockmodels: First steps](#). *Social Networks* **5** (1983), no. 2, 109–137 MR [0718088](#)
- [24] D. R. Hunter, [MM algorithms for generalized Bradley–Terry models](#). *Ann. Statist.* **32** (2004), no. 1, 384–406 Zbl [1105.62359](#) MR [2051012](#)
- [25] K. G. Jamieson and R. Nowak, Active ranking using pairwise comparisons. In *Advances in Neural Information Processing Systems 24*, pp. 2240–2248, Curran Associates, Red Hook, NY, 2011
- [26] M. Jang, S. Kim, C. Suh, and S. Oh, Top-k ranking from pairwise comparisons: When spectral ranking is optimal. 2016, arXiv:[1603.04153](#)
- [27] H. Joe, [Extended use of paired comparison models, with application to chess rankings](#). *J. Roy. Statist. Soc. Ser. C* **39** (1990), no. 1, 85–93 Zbl [0707.62149](#) MR [1038891](#)
- [28] E. Kaufmann, O. Cappé, and A. Garivier, On the complexity of best-arm identification in multi-armed bandit models. *J. Mach. Learn. Res.* **17** (2016), article no. 1 Zbl [1360.62433](#) MR [3482921](#)
- [29] T. L. Lai and H. Robbins, [Asymptotically efficient adaptive allocation rules](#). *Adv. in Appl. Math.* **6** (1985), no. 1, 4–22 Zbl [0568.62074](#) MR [0776826](#)
- [30] C. M. Le, E. Levina, and R. Vershynin, [Concentration of random graphs and application to community detection](#). In *Proceedings of the International Congress of Mathematicians—Rio de Janeiro 2018. Vol. IV. Invited lectures*, pp. 2925–2943, World Sci. Publ., Hackensack, NJ, 2018 Zbl [1448.05179](#) MR [3966517](#)
- [31] N. Levine, K. Crammer, and S. Mannor, Rotting bandits. In *Advances in Neural Information Processing Systems 30*, pp. 3074–3083, Curran Associates, Red Hook, NY, 2017
- [32] Y. Lu and H. H. Zhou, Statistical and computational guarantees of Lloyd’s algorithm and its variants. 2016, arXiv:[1612.02099](#)
- [33] Y. Luo and C. Gao, Computational lower bounds for graphon estimation via low-degree polynomials. [v1] 2023, [v3] 2024, arXiv:[2308.15728](#)
- [34] P. Massart, [Concentration inequalities and model selection](#). Lecture Notes in Math. 1896, Springer, Berlin, 2007 Zbl [1170.60006](#) MR [2319879](#)
- [35] L. Massoulié, [Community detection thresholds and the weak Ramanujan property](#). In *STOC’14—Proceedings of the 2014 ACM Symposium on Theory of Computing*, pp. 694–703, ACM, New York, 2014 Zbl [1315.68210](#) MR [3238997](#)
- [36] T. Minka, R. Cleven, and Y. Zaykov, TrueSkill 2: An improved Bayesian skill rating system. MSR-TR-2018-8, Microsoft, 2018
- [37] C. Moore, The computer science and physics of community detection: Landscapes, phase transitions, and hardness. *Bull. Eur. Assoc. Theor. Comput. Sci. EATCS* (2017), no. 121, 26–61 Zbl [1409.68220](#) MR [3699594](#)
- [38] E. Mossel, J. Neeman, and A. Sly, [Reconstruction and estimation in the planted partition model](#). *Probab. Theory Related Fields* **162** (2015), no. 3–4, 431–461 Zbl [1320.05113](#) MR [3383334](#)
- [39] E. Mossel, J. Neeman, and A. Sly, [Consistency thresholds for the planted bisection model](#). *Electron. J. Probab.* **21** (2016), article no. 21 Zbl [1336.05117](#) MR [3485363](#)
- [40] E. Mossel, J. Neeman, and A. Sly, [A proof of the block model threshold conjecture](#). *Combinatorica* **38** (2018), no. 3, 665–708 Zbl [1424.05272](#) MR [3876880](#)

- [41] W. Ren, J. Liu, and N. B. Shroff, Exploring  $k$  out of top  $\rho$  fraction of arms in stochastic bandits. In *Proceedings of the Twenty-Second International Conference on Artificial Intelligence and Statistics*, pp. 2820–2828, Proceedings of Machine Learning Research 89, PMLR, 2019
- [42] H. Robbins, [Some aspects of the sequential design of experiments](#). *Bull. Amer. Math. Soc.* **58** (1952), 527–535 Zbl [0049.37009](#) MR [0050246](#)
- [43] J. Seznec, A. Locatelli, A. Carpentier, A. Lazaric, and M. Valko, Rotting bandits are no harder than stochastic ones. In *Proceedings of the Twenty-Second International Conference on Artificial Intelligence and Statistics*, pp. 2564–2572, Proceedings of Machine Learning Research 89, PMLR, 2019
- [44] T. Schramm and A. S. Wein, [Computational barriers to estimation from low-degree polynomials](#). *Ann. Statist.* **50** (2022), no. 3, 1833–1858 Zbl [07547952](#) MR [4441142](#)
- [45] N. B. Shah, S. Balakrishnan, J. Bradley, A. Parekh, K. Ramchandran, and M. J. Wainwright, Estimation from pairwise comparisons: sharp minimax bounds with topology dependence. *J. Mach. Learn. Res.* **17** (2016), article no. 58 Zbl [1360.62409](#) MR [3504618](#)
- [46] N. B. Shah and M. J. Wainwright, Simple, robust and optimal ranking from pairwise comparisons. *J. Mach. Learn. Res.* **18** (2017), article no. 199 MR [3827087](#)
- [47] C. Sire and S. Redner, [Understanding baseball team standings and streaks](#). *Eur. Phys. J. B* **67** (2009), 473–481
- [48] L. Stephan and L. Massoulié, Robustness of spectral methods for community detection. [v1] 2018, [v2] 2019, arXiv:[1811.05808v2](#)
- [49] B. Szörényi, R. Busa-Fekete, A. Paul, and E. Hüllermeier, Online rank elicitation for Plackett–Luce: A dueling bandits approach. In *Advances in Neural Information Processing Systems 28*, pp. 604–612, Curran Associates, Red Hook, NY, 2015
- [50] W. R. Thompson, [On the likelihood that one unknown probability exceeds another in view of the evidence of two samples](#). *Biometrika* **25** (1933), no. 3-4, 285–294 Zbl [59.1159.03](#)
- [51] L. Torres, SuNBEaM: Spectral non-backtracking embedding and pseudo-metric, 2018, <https://github.com/leotrs/sunbeam> visited on 30 July 2024
- [52] L. Torres, P. Suárez-Serrato, and T. Eliassi-Rad, [Non-backtracking cycles: Length spectrum theory and graph mining applications](#). *Applied Network Science* **4** (2019), no. 1, article no. 41
- [53] M. J. Whiting, D. M. Stuart-Fox, D. O’Conner, D. Firth, N. C. Bennett, and S. P. Bloomberg, [Ultraviolet signals ultra-aggression in a lizard](#). *Animal behaviour* **72** (2006), 353–363
- [54] S.-Y. Yun and A. Proutière, Accurate community detection in the stochastic block model via spectral algorithms. 2014, arXiv:[1412.7335](#)
- [55] S.-Y. Yun and A. Proutière, Community detection via random and adaptive sampling. In *Proceedings of The 27th Conference on Learning Theory*, pp. 138–175, Proceedings of Machine Learning Research 35, PMLR, 2014
- [56] E. Zermelo, [Die Berechnung der Turnier-Ergebnisse als ein Maximumproblem der Wahrscheinlichkeitsrechnung](#). *Math. Z.* **29** (1929), no. 1, 436–460 Zbl [54.0543.01](#) MR [1545015](#)

- [57] A. Y. Zhang and H. H. Zhou, [Minimax rates of community detection in stochastic block models](#). *Ann. Statist.* **44** (2016), no. 5, 2252–2280 Zbl [1355.60125](#) MR [3546450](#)

Received 10 December 2022; revised 27 January 2024.

**Christophe Giraud**

Laboratoire de Mathématiques d’Orsay, Université Paris-Saclay, CNRS, Bâtiment 307, rue Michel Magat, 91405 Orsay, France; [christophe.giraud@universite-paris-saclay.fr](mailto:christophe.giraud@universite-paris-saclay.fr)

**Yann Issartel**

Laboratoire de Mathématiques d’Orsay, Université Paris-Saclay, CNRS, Bâtiment 307, rue Michel Magat, 91405 Orsay; Laboratoire Traitement et Communication de l’Information (LTCI), Institut Polytechnique de Paris, Télécom Paris, 19 Pl. Marguerite Perey, 91120 Palaiseau, France; [yann.issartel@telecom-paris.fr](mailto:yann.issartel@telecom-paris.fr)

**Luc Lehericy**

Laboratoire de Mathématiques d’Orsay, Université Paris-Saclay, CNRS, Bâtiment 307, rue Michel Magat, 91405 Orsay; Laboratoire J.A. Dieudonné, Université Côte d’Azur, Parc Valrose, 06108 Nice, France; [luc.lehericy@univ-cotedazur.fr](mailto:luc.lehericy@univ-cotedazur.fr)

**Matthieu Lerasle**

Centre de Recherche en Économie et de Statistiques (CREST), CNRS, École polytechnique, GENES, ENSAE Paris, Institut Polytechnique de Paris, 5 avenue Henri Le Chatelier, 91120 Palaiseau, France; [matthieu.lerasle@ensae.fr](mailto:matthieu.lerasle@ensae.fr)