Budgeted online influence maximization

Pierre Perrault, Jennifer Healey, Zheng Wen, Michal Valko

Budgeted Online Influence Maximization

Pierre Perrault [1 2 3] **Jennifer Healey** [1] **Zheng Wen** [4] **Michal Valko** [4 3]

Abstract

We introduce a new budgeted framework for online influence maximization, considering the total cost of an advertising campaign instead of the common cardinality constraint on a chosen influencer set. Our approach models better the real-world setting where the cost of influencers varies and advertizers want to find the best value for their overall social advertising budget. We propose an algorithm assuming an independent cascade diffusion model and edge-level semi-bandit feedback, and provide both theoretical and experimental results. Our analysis is also valid for the cardinality-constraint setting and improves the state of the art regret bound in this case.

1. Introduction

Viral marketing through online social networks now represents a significant part of many digital advertising budgets. In this form of marketing, companies incentivize chosen influencers in social networks (e.g., Facebook, Twitter, YouTube) to feature a product in hopes that their followers will adopt the product and repost the recommendation to their own network of followers. The effectiveness of the chosen set of influencers can be measured by the expected number of users that adopt the product due to their initial recommendation, called the *spread*. *Influence maximization* (IM, Kempe et al., 2003) is the problem of choosing the optimal set of influencers to maximize the spread under a cardinality constraint on the chosen set.

In order to define the spread, we need to specify a diffusion process such as independent cascade (IC) or linear threshold (LT) (Kempe et al., 2003). The parameters of these models are usually *unknown*. Different methods exist to estimate the parameters of the diffusion model from historical data (see section 1.1) however historical data is often difficult to obtain. Another possibility is to consider *online influence maximization* (OIM) (Vaswani et al., 2015; Wen et al., 2017) where an agent actively learns about the network by interacting with it repeatedly, trying to find the best seed influencers. The agent thus faces the dilemma of exploration versus exploitation, allowing us to see it as *multi-armed bandits* problem (Auer et al., 2002). More precisely, the agent faces IM over T rounds. Each round, it selects m seeds (based on *feedback* from prior rounds) and diffusion occurs; then it gains a reward equal to the spread and receives some feedback on the diffusion.

IM and OIM optimize with the constraint of a fixed number of seeds. This reflects a fixed seed cost model, for example, where influencers are incentivized by being given an identical free product. In reality, however, many influencers demand different levels of compensation. Those with a high out-degree (e.g., number of followers) are usually more expensive. Due to these cost variations, marketers usually wish to optimize their seed sets S under a budget $c(S) \leq b$ rather than a cardinality constraint $|S| \leq m$. Optimizing a seed set under a budget has been studied in the *offline* case by Nguyen and Zheng (2013). In the *online* case, Wang et al. (2020) considered the relaxed constraint $\mathbb{E}[c(S)] \leq b$, where the expectation is over the possible randomness of S.[1] We believe however that the constraint of a fixed, equal budget $c(S) \leq b$ at each round does not sufficiently model the willingness to choose a cost-efficient seed set. Indeed, we see that the choice of b is crucial: a b too large translates into a waste of budget (some seeds that are too expensive will be chosen) and a b too small translates into a waste of time (a whole round is used to influence only a few users). To circumvent this issue, instead of a budget per round, in our framework, we allow the agent to choose seed sets of any cost at each round, under an overall budget constraint (equal to $B = bT$ for instance). In summary, we incorporate the OIM framework into a *budgeted bandit* setting. Our setting is more flexible for the agent, and better meets real-world needs.

1.1. Related work on IM

IM can be formally defined as follows. A social network is modeled as a directed graph $G = (V, E)$, with nodes V

[1]Adobe Research, San Jose, CA [2]ENS Paris-Saclay [3]Inria Lille [4]DeepMind. Correspondence to: Pierre Perrault <pierre.perrault@outlook.com>.

Proceedings of the 37^{th} International Conference on Machine Learning, Vienna, Austria, PMLR 119, 2020.

[1]This relaxation is to avoid a computationally costly *partial enumeration* (Krause and Guestrin, 2005; Khuller et al., 1999)

representing users and edges E representing connections. An underlying diffusion model D governs how information spreads in G. More precisely, D is a probability distribution on subgraphs G' of G, and given some seed set S, the spread $\sigma(S)$ is defined as the expected number of S-reachable[2] nodes in $G' \sim D$. IM aims to find S that is a solution to

$$\max_{|S|=m} \sigma(S). \tag{1}$$

Although IM is NP-hard under standard diffusion models — i.e., IC and LT — σ is a monotone *submodular*[3] function (Fujishige, 2005), and given a value oracle access to σ, the standard GREEDY algorithm solves (1) within a $1 - 1/e$ approximation factor (Nemhauser et al., 1978). There have been multiple lines of work for IM, including the development of heuristics, approximation algorithms, as well as alternative diffusion models (Leskovec et al., 2007; Goyal et al., 2011; Tang et al., 2014; 2015). Additionally, there are also results on learning D from data in the case it is not known (Saito et al., 2008; Goyal et al., 2010; Gomez-Rodriguez et al., 2012; Netrapalli and Sanghavi, 2012).

1.2. Related work on OIM

Prior work in OIM has mainly considered either *node level semi-bandit* feedback (Vaswani et al., 2015), where the agent observes all the S-reachable nodes in G', or *edge level semi-bandit* feedback (Wen et al., 2017), where the agent observes the whole S-reachable subgraph (i.e., the subgraph of G' induced by S-reachable nodes). Other, weaker, feedback settings have also been studied including: pairwise influence feedback, where all nodes that would be influenced by a seed set are observed but not the edges connecting them, i.e., $(\{i\}\text{-reachable nodes})_{i\in S}$ is observed (Vaswani et al., 2017); local feedback, where the agent observes a set of out-neighbors of S (Carpentier and Valko, 2016) and immediate neighbor observation where the agent only observes the out-degree of S (Lugosi et al., 2019).

1.3. Our contributions

In this paper, we define the budgeted OIM paradigm and propose a performance metric for an online policy on this problem using the notion of *approximation regret* (Chen et al., 2013). To the best of our knowledge, the both of contributions are new. We then focus our study on the IC model with edge level semi-bandit feedback. We design a CUCB-style algorithm and prove logarithmic regret bounds. We also propose some modifications of this algorithm with improving the regret rates. These gains apply to the non-budgeted setting, giving an improvement over the state-of-the-art analysis of the standard CUCB-approach (Wang and Chen, 2017). Our proof incorporates an approximation guarantee of GREEDY for ratio of submodular and modular functions, which may also be of independent interest.

[2] nodes that are reachable from some node in S.

[3] f is submodular if $f(A\cup\{i\})-f(A)$ is non-increaing with A.

2. Problem definition

In this section, we formulate the problem of budgeted OIM and give a regret definition for evaluating policies in that setting. We also justify our choice for this notion of regret. We typeset vectors in bold and indicate components with indices, e.g., for some set I, $\mathbf{a} = (a_i)_{i\in I} \in \mathbb{R}^I$ is a vector on I. Let $\mathbf{e}_i$ be the i^{th} canonical unit vector of $\mathbb{R}^I$. The incidence vector of any subset $A \subset I$ is $\mathbf{e}_A \triangleq \sum_{i\in A} \mathbf{e}_i$.

We consider a fixed directed network $G = (V, E)$, known to the agent, with $V \triangleq \{1, \dots, |V|\}$. We denote by $ij \in E$ the directed edge from node i to j in G. We assume that G doesn't have self-loops, i.e., for all $ij \in E$, $i \neq j$. For a node $i \in V$, a subset $S \subset V$, and a vector $\mathbf{w} \in \{0,1\}^E$, the predicate $S \overset{\mathbf{w}}{\rightsquigarrow} i$ holds if, in the graph defined by $G_{\mathbf{w}} \triangleq (V, \{ij \in E, w_{ij} = 1\})$, there is a forward path from a node in S to the node i. If it holds, we say that i is influenced by S under $\mathbf{w}$. We define $p_i(S;\mathbf{w}) \triangleq \mathbb{I}\left\{S \overset{\mathbf{w}}{\rightsquigarrow} i\right\}$ and the *spread* as $\sigma(S;\mathbf{w}) \triangleq \left|\left\{i \in V,\ S \overset{\mathbf{w}}{\rightsquigarrow} i\right\}\right|$. Our diffusion process is defined by the random vector $\mathbf{W} \in \{0,1\}^E$, and our cost is defined by the random[4] vector $\mathbf{C} \in [0,1]^{V\cup\{0\}}$ where the added component C_0 represents any fixed costs[5]. Notice, random costs are neither assumed to be mutually independent nor independent from $\mathbf{W}$. We will see that components of $\mathbf{W}$ might however be mutually independent (e.g., for the IC model).

2.1. Budgeted online influence maximization

The agent interacts with the diffusion process across several rounds, using a learning policy. At each round $t \geq 1$, the agent first selects a seed set $S_t \subset V$, based on its past observations. Then, the random vectors for both the diffusion process $\mathbf{W}_t \sim \mathbb{P}_{\mathbf{W}}$ and the costs $\mathbf{C}_t \sim \mathbb{P}_{\mathbf{C}}$ are sampled independently from previous rounds. Then, the agent observes some feedback from both the diffusion process and the costs.

We provide in (2) the expected cumulative rewards F_B defined for some total budget $B > 0$. The goal for the agent is to follow a learning policy π maximizing F_B. In (2), recall that S_t is the seed set selected by π at round t.

[4] Although costs are usually deterministic, we assume randomness for more generality (influencer campaigns may have uncertain surcharges for example).

[5] We provide a toy example where C_0 models a concrete quantity: Assume you want to fill your restaurant. You may pay some seeds and ask them to advertise/influence people. C_0 represents the cost of the food, the staff, the rent, the taxes, ...

$$F_B(\pi) \triangleq \mathbb{E}\left[\sum_{t=1}^{\tau_B-1} \sigma(S_t; \mathbf{W}_t)\right]. \quad (2)$$

τ_B is the random round at which the remaining budget becomes negative: if $B_t \triangleq B - \sum_{t' \leq t}\left(\mathbf{e}_{S_{t'}}^\intercal \mathbf{C}_{t'} + C_{0,t'}\right)$, then $B_{\tau_B-1} \geq 0$ and $B_{\tau_B} < 0$. Notice, quantities B_t and τ_B are usual in budgeted multi-armed bandits (Xia et al., 2016; Ding et al., 2013).

2.2. Performance metric

We restrict ourselves to *efficient* policies, i.e., we consider a complexity constraint on the policy the agent can follow: For a round t, the space and time complexity for computing S_t has to be polynomial in $|V|$, and polylogarithmic in t. To evaluate the performance of a learning policy π, we use the notion of approximation regret (Kakade et al., 2009; Streeter and Golovin, 2009; Chen et al., 2016). The agent wants to follow a learning policy π which minimizes

$$R_{B,\varepsilon}(\pi) \triangleq (1 - 1/e - \varepsilon)F_B^\star - F_B(\pi),$$

where $F_B^\star$ is the best possible value of F_B over all policies (thus leveraging on the knowledge of $\mathbb{P}_{\mathbf{W}}$ and $\mathbb{P}_{\mathbf{C}}$), and where $\varepsilon > 0$ is some parameter the agent can control to determine the tradeoff between computation and accuracy.

Remark 1. *This OIM with a total budget B is different from OIM in previous work, such as Wang and Chen (2017), even when we set all costs to be equal. In our setting, there is only one total budget for all rounds, and the policy is free to choose seed sets of different cost in each round, whereas in the previous work, each round had a fixed budget for the number/cost of seeds selected. Our setting thus avoid the use of a budget per round, which is in practice more difficult to establish than a global budget B. Nevertheless, as we will see in section 6, both types of constraints (global and per round) can be considered simultaneously when the true costs are known.*

2.3. Justification for the approximation regret

In the non-budgeted OIM problem with a cardinality constraint given by $m \in [|V|]$, let us recall that the approximation regret

$$R_{T,\varepsilon}(\pi) \triangleq \sum_{t \leq T} \max_{\substack{S \subset V,\\ |S|=m}} \mathbb{E}[(1-1/e-\varepsilon)\sigma(S; \mathbf{W}) - \sigma(S_t; \mathbf{W})]$$

is standard (Wen et al., 2017; Wang and Chen, 2017). In this notion of regret, the factor $(1 - 1/e - \varepsilon)$ (Feige, 1998; Chen et al., 2010) reflects the difficulty of approximating the following NP-Hard (Kempe et al., 2003) problem in the case the distribution $\mathbb{P}_{\mathbf{W}}$ is described by IC or LT, and is known to the agent:

$$\max_{S \subset V,\ |S|=m} \mathbb{E}[\sigma(S; \mathbf{W})]. \quad (3)$$

For our budgeted setting, at first sight, it is not straightforward to know which approximation factor to choose. Indeed, since the random horizon may be different in $F_B(\pi)$ and in $F_B^\star$, the expected regret $R_{B,\varepsilon}(\pi)$ is not expressed as the expectation of a sum of approximation gaps, so we can't directly reduce the regret level approximability to the gap level approximability. We thus consider a quantity provably close to $R_{B,\varepsilon}(\pi)$ and easier to handle.

Proposition 1. *Define*

$$\lambda^\star \triangleq \max_{S \subset V} \frac{\mathbb{E}[\sigma(S; \mathbf{W})]}{\mathbb{E}\left[\mathbf{e}_{S \cup \{0\}}^\intercal \mathbf{C}\right]}.$$

For all $S \subset V$, define the gap *corresponding to S as*

$$\Delta(S) \triangleq (1 - 1/e - \varepsilon)\lambda^\star \mathbb{E}\left[\mathbf{e}_{S \cup \{0\}}^\intercal \mathbf{C}\right] - \mathbb{E}[\sigma(S; \mathbf{W})].$$

Then, for any policy π selecting S_t at round t,

$$\left| R_{B,\varepsilon}(\pi) - \mathbb{E}\left[\sum_{t=1}^{\tau_B-1} \Delta(S_t)\right] \right| \leq 2|V| + 2\lambda^\star(1 + |V|).$$

From Proposition 1, whose proof can be found in Appendix A, $R_{B,\varepsilon}(\pi)$ and $\mathbb{E}\left[\sum_{t=1}^{\tau_B-1} \Delta(S_t)\right]$ are equivalent in term of regret upper bound rate. Therefore, the factor $(1 - 1/e - \varepsilon)$ should reflect the approximability of

$$\max_{S \subset V} \mathbb{E}[\sigma(S; \mathbf{W})] / \mathbb{E}\left[\mathbf{e}_{S \cup \{0\}}^\intercal \mathbf{C}\right] = \max_{S \subset V} f(S)/c(S). \quad (4)$$

Considering the specific problem where the cost function is of the form $c(S) = c_1|S| + c_0$, for some $(c_0, c_1) \in [0, 1]^2$, we can reduce the approximability of (4) to the approximability of the following problem considered in Wang et al. (2020):

$$\max_{S \subset V} \mathbb{E}[f(S)] \ s.t.\ \mathbb{E}[|S|] \leq m, \quad (5)$$

for some given integer m, where the expectations are with respect to a randomization in the approximation algorithm. Wang et al. (2020) proved that this problem is NP-hard by reducing to the *set cover* problem. We show here that an approximation ratio α better that $1 - 1/e$ yields an approximation for set cover within $(1 - \delta)\log(|V|)$, $\delta > 0$, which is impossible unless NP $\subset$ TIME$\left(n^{\mathcal{O}(\log\log(|V|))}\right)$ (Feige, 1998). Consider the graph where the collection of out-neighborhoods is exactly the collection of sets in the set cover instance. First, trying out all possible values of m, we concentrate on the case in which the optimal m for set cover is tried out. As in Feige (1998), for $k \in \mathbb{N}^*$, we repeatedly apply the algorithm that α-approximate (5). It outputs a set S_k (that can be associated with a set of neighborhoods) and after each application the nodes already covered by previous applications are removed from the graph, giving a sequence of objective functions (f_k) with $f_1 = f$. We thus obtain

$$\mathbb{E}[f_k(S_k)|S_1, \ldots, S_{k-1}] \geq \alpha\left(|V| - \sum_{k'=1}^{k-1} f_{k'}(S_{k'})\right).$$

Noticing that $\mathbb{E}[f(S_1 \cup \cdots \cup S_k)] = \sum_{k'=1}^{k} \mathbb{E}[f_{k'}(S_{k'})]$, we get

$$\mathbb{E}[f(S_1 \cup \cdots \cup S_k)] \geq \big(1-(1-\alpha)^k\big)|V|.$$

After $\ell = \lfloor \log(1/|V|)/\log(1-\alpha) \rfloor < (1-\delta)\log(|V|)$ iterations, we obtain that $S = S_1 \cup \cdots \cup S_\ell$ is a cover, i.e., $f(S) = |V|$. The result follows noticing that in expectation (and so with positive probability), we have $|S| \leq \ell m$.

3. Algorithm for IC with edge level semi-bandit feedback

3.1. Setting

For $\mathbf{w} \in [0,1]^V$, we recall that we can define an IC model by taking $\mathbb{P}_{\mathbf{W}} = \otimes_{ij\in E}\text{Bernoulli}(w_{ij})$. We can extend the two previous functions p_i and σ to $\mathbf{w}$ taking values in $[0,1]^V$ as follows: Let $\mathbf{W} \sim \otimes_{ij\in E}\text{Bernoulli}(w_{ij})$. We define the probability that i is influenced by S under $\mathbf{W}$ as $p_i(S;\mathbf{w}) \triangleq \mathbb{P}\Big[S \overset{\mathbf{W}}{\rightsquigarrow} i\Big]$, and we let the spread be $\sigma(S;\mathbf{w}) \triangleq \mathbb{E}\Big[\Big|\Big\{i \in V,\ S \overset{\mathbf{W}}{\rightsquigarrow} i\Big\}\Big|\Big]$. Another expression for the spread is $\sigma(S;\mathbf{w}) = \sum_{i\in V} p_i(S;\mathbf{w})$. We fix a weight vector on E, $\mathbf{w}^\star \triangleq \big(w^\star_{ij}\big)_{ij\in E} \in [0,1]^E$, a cost vector on $V \cup \{0\}$, $\mathbf{c}^\star \triangleq (c^\star_i)_{i\in V\cup\{0\}} \in [0,1]^{V\cup\{0\}}$, with $c^\star_0 > 0$. These quantities are initially *unknown* to the agent. We assume from now that

$$\mathbb{P}_{\mathbf{W}} \triangleq \otimes_{ij\in E}\text{Bernoulli}\big(w^\star_{ij}\big),$$

and that

$$\mathbb{E}[\mathbf{C}] = \mathbf{c}^\star.$$

We also define $S^\star \in \arg\max_{S\subset V} \sigma(S;\mathbf{w}^\star)/\mathbf{e}_{S\cup\{0\}}^\intercal \mathbf{c}^\star$. We assume that the feedback received by the agent at round t is $\Big\{W_{ij,t},\ ij \in E,\ S_t \overset{\mathbf{W}_t}{\rightsquigarrow} i\Big\}$. The agent also receives semi-bandit feedback from the costs, i.e., $\{C_{i,t},\ i \in V,\ i \in S_t \cup \{0\}\}$ is observed.

3.2. Algorithm design

In this subsection, we present BOIM-CUCB, CUCB for Budgeted OIM problem as Algorithm 1. As we saw in Proposition 1, the policy that, at each round, $(1-1/e-\varepsilon)-$approximately maximize

$$S \mapsto \frac{\sigma(S;\mathbf{w}^\star)}{\mathbf{e}_S^\intercal \mathbf{c}^\star + c^\star_0} \tag{6}$$

has a bounded regret. Thus, BOIM-CUCB shall be based on this objective. Not only there are some estimation concerns due to the unknown parameters $\mathbf{w}^\star, \mathbf{c}^\star$, but in addition to that, we also need to evaluate/optimize our estimates of (6).

We begin by introducing some notations. We define the empirical means for $t \geq 1$ as: For all $i \in V \cup \{0\}$,

$$\overline{c}_{i,t-1} \triangleq \frac{\sum_{t'\in[t-1]} \mathbb{I}\{i \in S_{t'} \cup \{0\}\} C_{i,t'}}{N_{\ominus i,t-1}},$$

Algorithm 1 BOIM-CUCB

Input: $\varepsilon > 0$, $B_0 = B > 0$.
for each round $t \geq 1$ **do**
  If true costs are known, then $\mathbf{c}_t \leftarrow \mathbf{c}^\star$.
  Compute S_t given by Algorithm 2 with input $S \mapsto \sigma(S;\mathbf{w}_t)$, $\mathbf{c}_t$.
  Select seed set S_t, and pay $\mathbf{e}_{S_t\cup\{0\}}^\intercal \mathbf{C}_t$ (i.e., remove this cost from B_{t-1} to get the new budget B_t).
  if $B_t \geq 0$, **then**
    Get the reward $\sigma(S_t;\mathbf{W}_t)$, get the feedback, and update corresponding quantities accordingly.
  else
    The budget is exhausted: leave the for loop.
  end if
end for

and for all $ij \in E$,

$$\overline{w}_{ij,t-1} \triangleq \frac{\sum_{t'\in[t-1]} \mathbb{I}\Big\{S_{t'} \overset{\mathbf{W}_{t'}}{\rightsquigarrow} i\Big\} W_{ij,t'}}{N_{\oplus i,t-1}},$$

where $N_{\ominus i,t-1} \triangleq \sum_{t'=1}^{t-1} \mathbb{I}\{i \in S_{t'}\}$, $N_{\oplus i,t-1} \triangleq \sum_{t'=1}^{t-1} \mathbb{I}\Big\{S_{t'} \overset{\mathbf{W}_{t'}}{\rightsquigarrow} i\Big\}$. Using concentration inequalities, we get confidence intervals for the above estimates. We are then able to use an upper-confidence-bound (UCB) strategy (Auer et al., 2002). More precisely, in the case costs are unknown, we first build the *lower* confidence bound (LCB) on $c^\star_i$ as follows

$$c_{i,t} \triangleq 0 \vee \left(\overline{c}_{i,t-1} - \sqrt{\frac{1.5\log(t)}{N_{\ominus i,t-1}}}\right).$$

We can also define UCBs for $w^\star_{ij}$:

$$w_{ij,t} \triangleq 1 \wedge \left(\overline{w}_{ij,t-1} + \sqrt{\frac{1.5\log(t)}{N_{\oplus i,t-1}}}\right).$$

We use $w_{ij,t} = 1$ (and $c_{i,t} = 0$) when the corresponding counter is equal to 0. Our BOIM-CUCB approach chooses at each round t the seed set S_t given by Algorithm 2 which, as we shall see, approximately maximize $S \mapsto \sigma(S;\mathbf{w}_t)/\Big(\mathbf{e}_{S\cup\{0\}}^\intercal \mathbf{c}_t\Big)$. Indeed, with high probability, this set function is an upper bound on the true ratio (6) (using that σ is non decreasing w.r.t. $\mathbf{w}$). Notice that this approach is followed by Wang and Chen (2017) for the non budgeted setting, i.e., they choose S_t, $|S_t| \leq m$ that approximately maximize $S \mapsto \sigma(S;\mathbf{w}_t)$. To complete the description of our algorithm, we need to describe Algorithm 2. This is the purpose of the following.

3.3. Greedy for ratio maximization

In BOIM-CUCB, one has to approximately maximize the ratio $S \mapsto \sigma(S;\mathbf{w}_t)/\mathbf{e}_{S\cup\{0\}}^\intercal \mathbf{c}_t$, that is a ratio of submodular over modular function. A GREEDY technique can be

Algorithm 2 GREEDY for ratio, Lazy implementation

Input: σ that is an increasing submodular function, $\mathbf{c} \in [0,1]^{V\cup\{0\}}$.
$S_0 \leftarrow \emptyset$.
$\rho \leftarrow [(\infty, i)]_{i\in V}$.
for $k \in [|V|]$ **do**
  checked $\leftarrow S_{k-1}$.
  $(*)$ Remove the first element $\rho[0] = (\sim, i)$ from ρ.
  if $i \notin$ checked **then**
    Insert $((\sigma(\{i\}\cup S_{k-1}) - \sigma(S_{k-1}))/c_i, i)$ in ρ, such that $\rho[:][0]$ is sorted in decreasing order.
    Add i to checked and go back to $(*)$.
  else
    $S_k \leftarrow S_{k-1}\cup\{i\}$.
  end if
end for
$k' \leftarrow \arg\max_{k\in\{0,\dots,|V|\}} \sigma(S_k)/\mathbf{e}_{S_k\cup\{0\}}^\intercal \mathbf{c}$.
Output: $S_{k'}$.

used (see Algorithm 2). Indeed, instead of maximizing the marginal contribution at each time step, as the standard GREEDY algorithm do, the approach is to maximize the so called bang-per-buck, i.e., the marginal contribution divided by the marginal cost. This builds a sequence of increasing subsets, and the final output is the one that maximizes the ratio. We prove in Appendix D the following Proposition 2, giving an approximation factor of $1-1/e$ for Algorithm 2.

Proposition 2. *Algorithm 2 with input $\sigma, \mathbf{c}$ is guaranteed to obtain a solution S such that:*

$$\left(1-e^{-1}\right)\frac{\sigma(S^\star)}{\mathbf{e}_{S^\star\cup\{0\}}^\intercal\mathbf{c}} \le \frac{\sigma(S)}{\mathbf{e}_{S\cup\{0\}}^\intercal\mathbf{c}}.$$

Notice, a similar result as Proposition 2 is stated in Theorem 3.2 of Bai et al. (2016). However, their proof doesn't hold in our case, since their inequality (16) would be true only for a *normalized* cost (i.e. $c_0 = 0$). Actually, $c_0 = 0$ implies that $S^\star$ is a singleton, from subadditivity of σ.

For more efficiency, we use a greedy algorithm with lazy evaluations (Minoux, 1978; Leskovec et al., 2007), leveraging on the *submodularity* of σ. More precisely, in Algorithm 2, instead of taking the $\arg\max$ in the step

$$S_k \leftarrow S_{k-1}\cup\left\{\arg\max_{i\in V\setminus S_{k-1}} \frac{\sigma(\{i\}\cup S_{k-1}) - \sigma(S_{k-1})}{c_i}\right\},$$

we maintain an upper bound ρ (initially ∞) on hhe marginal gain, sorted in decreasing order. In each iteration k, we evaluates the element on top of the list, say i, and updates its upper bound with the marginal gain at S_{k-1}. If after the update the upper bound is greater than the others, submodularity guarantees that i is the element with the largest marginal gain.

Algorithm 2 (and the approximation factor) can't be used directly in the OIM context, since computing the exact spread σ is #P hard (Chen et al., 2010). However, with Monte Carlo (MC) simulations, it can efficiently reach an arbitrarily close ratio of $\alpha = 1-1/e-\varepsilon$, with a high probability $1-1/\left(t\log^2(t)\right)$ (Kempe et al., 2003).

3.4. An alternative to Lazy Greedy: Ratio maximization from sketches

In the previous approach, MC can still be computationally costly, since the marginal contribution have to be re-evaluated each time, using directed reachability computation in each MC instance. There exist efficient alternatives to $\alpha-$approximately maximize the spread with cardinality constraint, such as TIM from Tang et al. (2014) and SKIM from Cohen et al. (2014). Adapting TIM to our ratio maximization context is not straightforward, since it require to know the seed set size in advance, which is not the case in Algorithm 2. SKIM is more promising since it uses the standard GREEDY in a *sketch space*. We provide an adaptation of SKIM for approximately maximize the ratio (see Appendix G). It uses the bottom-k min-hash sketches of Cohen et al. (2014), with a threshold for the length of the sketches that depends on both $k = \left\lfloor \varepsilon^{-2}\log(1/\delta)\right\rfloor$ and the cost, where δ is an upper bound on the probability that the relative error is larger than ε. Exactly as Cohen et al. (2014) proved the approximation ratio for SKIM, this approach reaches a factor of $1-1/e-\varepsilon$, with high probability. More precisely, at round $t \ge 2$, we can actually choose to have the approximation with $\delta = 1/\left(t\log^2(t)\right)$, only adding a $O(\log(t))$ factor in the computational complexity of Algorithm 2 (Cohen et al., 2014), thus remaining efficient.[6]

3.5. Regret bound for Algorithm 1

We provide a gap dependent upper bound on the regret of BOIM-CUCB in Theorem 1. For this, we define, for $i \in V$, the gap

$$\Delta_{i,\min} \triangleq \min_{S\subset V,\ p_i(S;\mathbf{w}^\star)>0,\ \Delta(S)>0} \Delta(S).$$

We also define, with d_k being the out-degree of node k,

$$p_{i,\max} \triangleq \max_{S\subset V,\ p_i(S;\mathbf{w}^\star)>0} \sum_{k\in V} d_k p_k(S;\mathbf{w}^\star).$$

Theorem 1. *If π is the policy described in Algorithm 1, then*

$$R_{B,\varepsilon}(\pi) = \mathcal{O}\left(\log B\left(\sum_{i\in V}|V|\frac{\lambda^\star + d_i p_{i,\max}|V|}{\Delta_{i,\min}}\right)\right).$$

[6] Wang and Chen (2017) uses the notion of approximation regret with a certain fixed probability. Here, we rather fix this probability to 1 and allow for a $O(\log(t))$ factor in the running time.

In addition, if true costs are known, then

$$R_{B,\varepsilon}(\pi) = \mathcal{O}\left(\log B\left(\sum_{i\in V}\frac{d_i p_{i,\max}|V|^2}{\Delta_{i,\min}}\right)\right).$$

A proof of Theorem 1 can be found in Appendix B. Notice that the analysis can be easily used for the non budgeted setting. In this case, it reduces to the state-of-the-art analysis of Wang and Chen (2017), except that we slightly simplify and improve the analysis to replace the factor $\max_{S\subset V}\sum_{k\in V} d_k \mathbb{I}\{p_k(S;\mathbf{w}^\star)>0\}$ by a potentially much lower quantity $p_{i,\max}$. In the case this last quantity is still large, we can further improve it by considering slight modifications to the original Algorithm 1. This is the purpose of the next section.

4. More refined optimistic spreads

We observe that the factor $p_{i,\max}$ in Theorem 1 can be as large as $|E|$ in the worst case. In other word, if $\Delta = \min_i \Delta_{i,\min}$, the rate can be as large as

$$\mathcal{O}\left(\log B\left(\frac{\lambda^\star|V|^2+|E|^2|V|^2}{\Delta}\right)\right).$$

We argue here that we can replace $|E|^2|V|^2$ by $|E||V|^3\log^2(|V|))$. Indeed, leveraging on the mutual independence of random variables W_{ij}, we can hope to get a tighter confidence region for $\mathbf{w}^\star$, and thus a provably tighter regret upper bound (Magureanu et al., 2014; Combes et al., 2015; Degenne and Perchet, 2016). We consider the following confidence region from Degenne and Perchet (2016) (see also Perrault et al. (2019a)) and adapted to our setting.

Fact 1 (Confidence ellipsoid for weights)**.** *For all $t\geq 2$, with probability at least $1-1/\big(t\log^2(t)\big)$,*

$$\sum_{ij\in E} N_{\oplus i,t-1}\big(w^\star_{ij}-\overline{w}_{ij,t-1}\big)^2 \leq \delta(t),$$

where $\delta(t)\triangleq 2\log(t)+2(|E|+2)\log\log(t)+1$.

For OIM (both budgeted and non budgeted), there is a large potential gain in the analysis using the confidence region given by Fact 1 compared to simply using an Hoeffding based one, like in BOIM-CUCB. More precisely, for classical combinatorial semi bandits, Degenne and Perchet (2016) reduced the gap dependent regret upper bound by a factor $\ell/\log^2(\ell)$, where in our case ℓ can be as large as $|E|$. However, there is also a drawback in practice with such confidence region: computing the optimistic spread might be inefficient, even if an oracle for evaluating the spread is available. Indeed, for a fixed $S\subset V$, the problem of maximizing $\mathbf{w}\mapsto\sigma(S;\mathbf{w})$ over $\mathbf{w}$ belonging to some ellipsoid might be hard, since the objective is not necessarily concave. We can overcome this issue using the following Fact 2 (Wen et al., 2017; Wang and Chen, 2017).

Fact 2 (Smoothness property of the spread)**.** *for all $S\subset V$, and all $\mathbf{w},\mathbf{w}'\in[0,1]^E$,*
$\forall k\in V,\ |p_k(S;\mathbf{w})-p_k(S;\mathbf{w}')|\leq\sum_{ij\in E} p_i(S;\mathbf{w})\big|w_{ij}-w'_{ij}\big|$.
In particular,

$$|\sigma(S;\mathbf{w})-\sigma(S;\mathbf{w}')|\leq|V|\sum_{ij\in E} p_i(S;\mathbf{w})\big|w_{ij}-w'_{ij}\big|.$$

For $S\subset V$ and $\mathbf{w}\in\mathbb{R}^E$, we define the confidence "bonus" as follows:

$$\text{Bonus}(S;\mathbf{w})\triangleq|V|\sqrt{\delta(t)\sum_{i\in V,N_{\oplus i,t-1}>0} d_i\frac{p_i(S;\mathbf{w})^2}{N_{\oplus i,t-1}}}.$$

Notice, we don't sum on vertices with a zero counter. We compensate this by using the convention $\overline{w}_{ij,t-1}=1$ when $N_{\oplus i,t-1}=0$. We can successively use Fact 2, Cauchy-Schwartz inequality, and Fact 1 to get, with probability at least $1-1/\big(t\log^2(t)\big)$,

$$\sigma(S;\mathbf{w}^\star)\leq\sigma(S;\overline{\mathbf{w}}_{t-1})+\text{Bonus}(S;\overline{\mathbf{w}}_{t-1}). \tag{7}$$

In the same way, with probability at least $1-1/\big(t\log^2(t)\big)$, we also have (8).

$$\sigma(S;\mathbf{w}^\star)\leq\sigma(S;\overline{\mathbf{w}}_{t-1})+\text{Bonus}(S;\mathbf{w}^\star). \tag{8}$$

Contrary to (7), this "optimistic spread" can't be used directly by the agent since $\mathbf{w}^\star$ is not known.

Although the optimistic spread defined in (7) is now much easier to compute, there is still a major drawback that remains: As a function of $S\subset V$, $\text{Bonus}(S;\overline{\mathbf{w}}_{t-1})$ is not necessarily submodular, so the optimistic spread is itself no longer submodular. This is an issue because submodularity is a crucial property for reaching the approximation ratio $1-1/e-\varepsilon$. We propose here several submodular upper bound to Bonus, defined for $S\subset V$ and $\mathbf{w}\in\mathbb{R}^E$:

- Bonus_1 is actually modular, and simply uses the subadditivity (w.r.t. S) of Bonus:

$$\text{Bonus}_1(S;\mathbf{w})\triangleq|V|\sum_{j\in S}\sqrt{\delta(t)\sum_{i,N_{\oplus i,t-1}>0} d_i\frac{p_i(\{j\};\mathbf{w})^2}{N_{\oplus i,t-1}}}.$$

- Bonus_2 uses the subadditivity of the square root:

$$\text{Bonus}_2(S;\mathbf{w})\triangleq|V|\sum_{i,N_{\oplus i,t-1}>0} p_i(S;\mathbf{w})\sqrt{\frac{\delta(t)d_i}{N_{\oplus i,t-1}}}.$$

- Bonus_3 uses $p_i(S;\mathbf{w})^2\leq p_i(S;\mathbf{w})$, and is submodular as the composition between a non decreasing concave function (the square root) and a monotone submodular function:

$$\text{Bonus}_3(S;\mathbf{w})\triangleq|V|\sqrt{\delta(t)\sum_{i,N_{\oplus i,t-1}>0} d_i\frac{p_i(S;\mathbf{w})}{N_{\oplus i,t-1}}}.$$

- Bonus_4 uses Jensen's inequality, and is submodular as the expectation of the square root of a submodular function.

$$\text{Bonus}_4(S;\mathbf{w}) \triangleq \mathbb{E}\left[|V|\sqrt{\sum_{i\in V, N_{\oplus i,t-1}>0, S\overset{\mathbf{W}}{\rightsquigarrow} i}\frac{\delta(t)d_i}{N_{\oplus i,t-1}}}\right],$$

where $\mathbf{W}\sim\otimes_{ij\in E}\text{Bernoulli}(w_{ij})$.

We can write the following approximation guarantees for the two first bonus:

$$\text{Bonus}(S;\mathbf{w}) \le \text{Bonus}_1(S;\mathbf{w}) \le |S|\text{Bonus}(S;\mathbf{w}), \quad (9)$$

$$\text{Bonus}(S;\mathbf{w}) \le \text{Bonus}_2(S;\mathbf{w}) \le \sqrt{|V|}\text{Bonus}(S;\mathbf{w}).$$

Notice, another approach to get a submodular bonus is to approximate $p_i(S;\overline{\mathbf{w}}_{t-1})$ by the square root of a modular function (Goemans et al., 2009). However, not only this bonus would be much more computationally costly to build than ours, but also, we would get only a $\sqrt{|V|}\log|V|$ approximation factor, which is worst than the one with our Bonus_2. Since increasing the bonus by a factor $\alpha\ge 1$ increases the gap dependent regret upper bound by a factor α^2, we only loose a factor $|V|$ for Bonus_2, compared to the use of Bonus, which is still better than the CUCB approach. Bonus_1 can also be interesting to use when we have some upper bound guarantee on the cardinality of seed sets used (see subsection 5.1). An approximation factor for Bonus_3 or Bonus_4 doesn't seem interesting, because it would involve the inverse of triggering probabilities. We can, however, further upper bound $\text{Bonus}_3(S;\mathbf{w}^\star)$ as follows:

$$\begin{aligned}\text{Bonus}_3(S;\mathbf{w}^\star) &= |V|\sqrt{\delta(t)\sum_{i,N_{\oplus i,t-1}>0} d_i\frac{p_i(S;\mathbf{w}^\star)}{N_{\oplus i,t-1}}}\\ &\le |V|\sqrt{\sum_{j\in S}\sum_{\substack{i\in V,\\ N_{\oplus i,t-1}>0}} d_i\frac{\delta(t)p_i(\{j\};\mathbf{w}^\star)}{N_{\oplus i,t-1}}}\\ &\le |V|\sqrt{\delta(t)\sum_{j\in S}|E|\left(\frac{8}{N_{\ominus j,t-1}}\wedge 1\right)},\end{aligned}$$

where the last inequality only holds under some high probability event, given by the following Proposition 3, involving counters on the costs and counters on the weights.

Proposition 3. *Consider the event defined by* $\mathfrak{P}_t \triangleq \{\forall i\in V, N_{\oplus i,t-1}\ge\delta(t)\}$. *Then, for all* $i,j\in V$,

$$\mathbb{P}\left[\mathfrak{P}_t \text{ and } \frac{\delta(t)p_i(\{j\};\mathbf{w}^\star)}{N_{\oplus i,t-1}} > \frac{8\delta(t)}{N_{\ominus j,t-1}}\right] \le 1/t^2.$$

We thus define for $S\subset V$,

$$\text{Bonus}_5(S) \triangleq |V|\sqrt{\delta(t)\sum_{j\in S}|E|\left(\frac{8}{N_{\ominus j,t-1}}\wedge 1\right)}.$$

This bonus is much more convenient since it does not depend anymore on $\mathbf{w}^\star$, and can thus be computed by the agent. Indeed, although the first four bonuses are likely to be tighter than this last submodular Bonus_5, their dependence in $\mathbf{w}$ forces us to use them for $\mathbf{w}=\overline{\mathbf{w}}_{t-1}$. Even if this doesn't pose any problem in practice, this is more difficult to handle in theory since it would involve optimistic estimates on $p_i(\ \cdot\ ;\overline{\mathbf{w}}_{t-1})$ itself (see the next section for further details). Actually, we will see that the analysis with Bonus_5 is slightly better than the one we would get with Bonus_2, since we loose a factor $|V|\log^2(|V|)/\log^2(|E|)$ compared to the use of Bonus. In addition, it allows a much more interesting constant term in the regret upper bound thanks to the suppression of the dependence in $\mathbf{w}^\star$.

Although we can improve the analysis based on Fact 1 and 2, the inequality in this last fact may be be less rough in practice (we confirm this in section 7). In that case, we suffer from this roughness, since we actually use Fact 2 to design our bonus. In contrast, BOIM-CUCB only uses it in the analysis, and can therefore adapt to a better smoothness inequality. Thus, we consider BOIM-CUCB_5, where we first compute S_t using the BOIM-CUCB approach, and accept it only if

$$\sigma(S_t;\mathbf{w}_t) \le \sigma(S_t;\overline{\mathbf{w}}_{t-1}) + \text{Bonus}_5(S_t), \quad (10)$$

otherwise, we chose S_t maximizing $\sigma(S;\overline{\mathbf{w}}_{t-1}) + \text{Bonus}_5(S)$. For technical reason (due to Proposition 3), we replace S_t by $S_t\cup\{j\}$ if it exists a $j\in V$, such that

$$N_{\oplus j,t-1} < \delta(t). \quad (11)$$

We thus both enjoy the theoretical advantages of Bonus_5 and the practical advantages of BOIM-CUCB. We give the following regret bounds for this approach.

Theorem 2. *If* π *is the policy following* BOIM-CUCB_5, *then*

$$R_{B,\varepsilon}(\pi) = \mathcal{O}\left(\log B\left(\sum_{i\in V}|V|\frac{\lambda^\star + |V||E|\log^2|V|}{\Delta_{i,\min}} + \lambda^\star|V|^2\right)\right).$$

A proof of Theorem 2 can be found in Appendix E. Such analysis also holds in the non-budgeted setting, and maximizing the spread only instead of the ratio, we can build a policy π satisfying the following (with the standard definition of the non-budgeted gaps):

$$R_{T,\varepsilon}(\pi) = \mathcal{O}\left(\log T\sum_{i\in V}\frac{|V|^2|E|\log^2|V|}{\Delta_{i,\min}}\right).$$

The regret rate is thus better than the one from Wang and Chen (2017), gaining a factor $|E|/\big(|V|\log^2(|V|)\big)$.

5. Improvements using Bonus$_1$ and Bonus$_4$

In this section, we show that the use of Bonus$_1$ and Bonus$_4$ leads to a better regret leading term, at the cost of a large second order term. In the following, we propose BOIM-CUCB$_1$ (resp. BOIM-CUCB$_4$), that are the same approach as BOIM-CUCB$_5$ with Bonus$_1(\ \cdot\ ; \overline{\mathbf{w}}_{t-1})$ (resp. Bonus$_4(\ \cdot\ ; \overline{\mathbf{w}}_{t-1})$) instead of Bonus$_5$, and where condition (11) is replaced by

$$\exists j \in V, N_{\oplus j,t-1} \leq |E|\delta(t).$$

5.1. Bonus$_1$ for low cardinality seed sets

In many real world scenarios, maximal cardinality of seed set is small compared to $|V|$. Indeed, in the non-budgeted setting, it is limited by m, and it is usually assumed that m is much smaller than $|V|$. In the budgeted setting, we will see in section 6 how to limit the cost of the chosen seeds, and this is likely to also induce a limit on the cardinality of seeds. Using Bonus$_1$ is more appropriate in this situation, according to the approximation factor (9). We state in Theorem 3 the regret bound for BOIM-CUCB$_1$.

Theorem 3. *If π is the policy* BOIM-CUCB$_1$*, and if all seeds selected have a cardinality bounded by m, then we have*

$$R_{B,\varepsilon}(\pi) = \mathcal{O}\left(\log B\left(\sum_{i\in V} m\frac{\lambda^\star + m|V|^2 d_i \log^2(|E|)}{\Delta_{i,\min}} + \lambda^\star |V|^2|E|\right)\right).$$

A proof can be found in Appendix F. As previously, we can state the following non-budgeted version, with seed set cardinality constrained by m:

$$R_{T,\varepsilon}(\pi) = \mathcal{O}\left(\log T\left(\sum_{i\in V}\frac{m^2|V|^2 d_i \log^2(|E|)}{\Delta_{i,\min}} + |E||V|^2\right)\right).$$

Notice, for both settings, there is an improvement in the main term (the gap dependent one), in the case $m \leq \sqrt{|V|}$. However, there is also a higher gap independent term that appears.

5.2. Bonus$_4$: the same performance as Bonus?

We show here that the regret with Bonus$_4$ is of the same order as what we would have had with Bonus (which is not submodular). However, Bonus$_4$ does not have the calculation guarantees of the other bonuses. We state in Theorem 4 the regret bound for the policy BOIM-CUCB$_4$. Notice that we obtain a bound whose leading term improves by a factor $|E|/\log^2|E|$ that of BOIM-CUCB.

Theorem 4. *If π is the policy* BOIM-CUCB$_4$*, then we have*

$$R_{B,\varepsilon}(\pi) = \mathcal{O}\left(\log B\left(\sum_{i\in V}\frac{|V|\lambda^\star + |V|^2 d_i \log^2(|E|)}{\Delta_{i,\min}} + \lambda^\star|V|^2|E|\right)\right).$$

The proof is given in in Appendix J. As previously, we can state the following non-budgeted version. Notice that the cardinality constrain does not appear in the bound.

$$R_{T,\varepsilon}(\pi) = \mathcal{O}\left(\log T\left(\sum_{i\in V}\frac{|V|^2 d_i \log^2(|E|)}{\Delta_{i,\min}} + |E||V|^2\right)\right).$$

In spite of the superiority in terms of regret of the use of Bonus$_4$, we must point out that, in the worst case, the calculation of this bonus may require a number of sample (and thus a time complexity) polynomial in t, which does not meet the criterion of efficiency that we set ourselves at the beginning of the paper.

6. Knapsack constraint for known costs

In their setting, Wang et al. (2020) considered the relaxed constraint

$$\mathbb{E}\left[\mathbf{e}_{S\cup\{0\}}\mathbf{c}^\star\right] \leq b, \tag{12}$$

instead of ratio maximization, where the expectation is over the possible randomness of S. When true costs are known to the agent, we can actually combine the two settings: a seed set S can be chosen only if it satisfies (12). In this section, we describe modifications this new setting implies. First of all, the regret definition is impacted, and $F_B^\star$ is now maximal for policies respecting the constraint (12) within each round. Naturally, the definitions of $\lambda^\star$ and $S^\star$ are also modified accordingly. Otherwise, apart from Algorithm 2, there is conceptually no change in the approaches that have been described in this paper. We now described the modification needed to make Algorithm 2 works in this setting. The same sequence of set S_k is considered, but instead of choosing the set that maximizes the ratio over all $k \in \{0,\dots,|V|\}$, we restrict the maximization to $k \in \{0,\dots,j\}$, where j is the first index such that $\mathbf{e}_{S_j\cup\{0\}}^\intercal \mathbf{c}^\star > b$. If this maximizer is not S_j, then it satisfies the constraint and is output. Else, we output S_j with probability $(b - \mathbf{e}_{S_{j-1}\cup\{0\}}^\intercal \mathbf{c}^\star)/c_j^\star$ and S_{j-1} with probability $1 - (b - \mathbf{e}_{S_{j-1}\cup\{0\}}^\intercal \mathbf{c}^\star)/c_j^\star$. This way, the expected cost of the output is b. We prove in Appendix K the following Proposition 4, giving an approximation factor of $1 - 1/e$ for the above modification of Algorithm 2.

Proposition 4. *The solution S obtained by the modified Algorithm 2 is such that:*

$$\left(1 - e^{-1}\right)\frac{\mathbb{E}[\sigma(S^\star)]}{\mathbb{E}\left[\mathbf{e}_{S^\star\cup\{0\}}^\intercal \mathbf{c}^\star\right]} \leq \frac{\mathbb{E}[\sigma(S)]}{\mathbb{E}\left[\mathbf{e}_{S\cup\{0\}}^\intercal \mathbf{c}^\star\right]},$$

where the expectation is over the possible randomness of $S, S^\star$.

7. Experiments

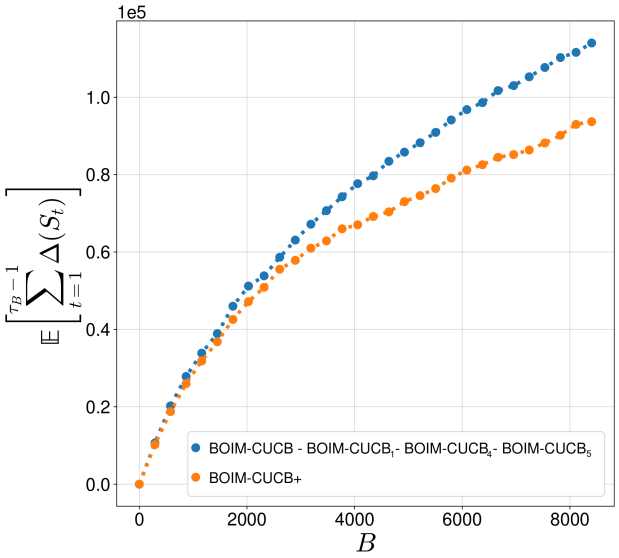

Figure 1. Regret curves with respect to the budget B (expectation computed by averaging over 10 independent simulations).

In this section, we present an experiment for Budgeted OIM. In Figure 1, we plot $\mathbb{E}\left[\sum_{t=1}^{\tau_B-1}\Delta(S_t)\right]$ with respect to the budget B used, running over up to $T = 10000$ rounds. This quantity is a good approximations to the true regret according to Proposition 1. Plotting the true regret would require to compute $F_B^\star$, which is NP-Hard to do. We consider a subgraph of Facebook network (Leskovec and Krevl, 2014), with $|V| = 333$ and $|E| = 5038$, as in Wen et al. (2017). We take $\mathbf{w}^\star \sim U(0,0.1)^{\otimes E}$ and take deterministic, known costs with $c_0^\star = 1$, and $c_i^\star = d_i / \max_{j\in V} d_j$. BOIM-CUCB+ is the same approach as BOIM-CUCB$_5$ with Bonus$(\,\cdot\,;\overline{\mathbf{w}}_{t-1})$ instead of Bonus$_5$, ignoring that Bonus$(\,\cdot\,;\overline{\mathbf{w}}_{t-1})$ is not submodular (it is only sub-additive).

We observed that in BOIM-CUCB$_1$, BOIM-CUCB$_4$, BOIM-CUCB$_5$, Condition 10 (with the correct bonus instead of Bonus$_5$) always holds, meaning that those algorithms coincide with BOIM-CUCB *in practice*, and that the gain only appears through the analysis. We thus plot a single curve for these 4 algorithms in Figure 1. On the other hand, we observe only a slight gain of BOIM-CUCB+ compared to BOIM-CUCB.

Our experiments confirm that Fact 2 is less rough in practice, as we already anticipated. Indeed, our submodular bonuses are not tight enough to compete with BOIM-CUCB, although we gain in the analysis. The slight gain that we have for BOIM-CUCB+ suggests that the issue is not only about the tightness of a submodular upper bound, but rather about the tightness of Fact 2. This is supported by the following observation we made: for the Facebook subnetwork, for 1000 random draws of a seed set and vector pairs in $[0,0.1]^E$, the ratio of the RHS and the LHS in Fact 2 is each time greater than $0.4|V|$.

In Appendix I, we conducted further experiments on a synthetic graph comparing BOIM-CUCB to BOIM-CUCB-REGULARIZED, which greedily maximizes the regularized spread $S \mapsto \sigma(S;\mathbf{w}_t) - \lambda \mathbf{e}_S \mathbf{c}_t$, where λ is a parameter to set. We observed that for an appropriate choice of λ, a performance similar to BOIM-CUCB can be obtained.

8. Discussion and Future work

We introduced a new Budgeted OIM problem, taking both the costs of influencers and fixed costs into account in the seed selection, instead of the usual cardinality constraint. This better represents the current challenges in viral marketing, since top influencers tend to be more and more costly. Our fixed cost can also be seen as the time that a round takes: A null fixed cost would mean that reloading the network to get a new independent instance is *free* and *instantaneous*.[7] Obviously, this is not realistic. We also provided an algorithm for Budgeted OIM under the IC model and the edge level semi-bandit feedback setting.

Interesting future directions of research would be to explore other kinds of feedback or diffusion models for Budgeted OIM. For practical scalability, it would also be good to investigate the incorporation of the linear generalization framework (Wen et al., 2017) into Budgeted OIM. Notice, this extension is *not* straightforward if we want to keep our tighter confidence region. More precisely, we believe that a *linear semi-bandit* approach that is aware of independence between edge observations should be developed (the linear generalization approach of Wen et al. (2017) treats each edge observation as *arbitrary correlated*).

In addition to this, exploring how the use of Fact 2 *in the Algorithm* might be avoided while still using confidence region given by Fact 1 would surely improve the algorithms. One possible way would be to use a Thompson Sampling (TS) approach (Wang and Chen, 2018; Perrault et al., 2020a), where the prior takes into account the mutual independence of weights. However, Wang and Chen (2018) proved in their Theorem 2 that TS gives linear approximation regret for some special approximation algorithms. Thus, we would have to use some specific property of the GREEDY approximation algorithm we use.

[7] In this case, using $|S|$ rounds choosing each time a single different influencer $i \in S$ is better than choosing the whole S in a single round.

Acknowledgements. The research presented was supported by European CHIST-ERA project DELTA, French Ministry of Higher Education and Research, Nord-Pas-de-Calais Regional Council, French National Research Agency project BOLD (ANR19-CE23-0026-04).

References

Auer, P., Cesa-Bianchi, N., and Fischer, P. (2002). Finite-time analysis of the multiarmed bandit problem. *Machine Learning*, 47(2-3):235–256.

Bai, W., Iyer, R., Wei, K., and Bilmes, J. (2016). Algorithms for optimizing the ratio of submodular functions. In *International Conference on Machine Learning*, pages 2751–2759.

Carpentier, A. and Valko, M. (2016). Revealing graph bandits for maximizing local influence. In *International Conference on Artificial Intelligence and Statistics*.

Chen, W., Wang, C., and Wang, Y. (2010). Scalable influence maximization for prevalent viral marketing in large-scale social networks. In *Knowledge Discovery and Data Mining*.

Chen, W., Wang, Y., and Yuan, Y. (2013). Combinatorial multi-armed bandit: General framework and applications. In *Proceedings of the 30th International Conference on Machine Learning*, volume 28 of *Proceedings of Machine Learning Research*, pages 151–159, Atlanta, Georgia, USA. PMLR.

Chen, W., Wang, Y., and Yuan, Y. (2016). Combinatorial multi-armed bandit and its extension to probabilistically triggered arms. *Journal of Machine Learning Research*, 17.

Cohen, E. (1997). Size-estimation framework with applications to transitive closure and reachability. *Journal of Computer and System Sciences*, 55(3):441–453.

Cohen, E. (2016). Minhash sketches.

Cohen, E., Delling, D., Pajor, T., and Werneck, R. F. (2014). Sketch-based influence maximization and computation: Scaling up with guarantees. In *Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management*, pages 629–638. ACM.

Cohen, E. and Kaplan, H. (2007). Summarizing data using bottom-k sketches. In *Proceedings of the twenty-sixth annual ACM symposium on Principles of distributed computing*, pages 225–234. ACM.

Combes, R., Shahi, M. S. T. M., Proutiere, A., and Others (2015). Combinatorial bandits revisited. In *Advances in Neural Information Processing Systems*, pages 2116–2124.

Degenne, R. and Perchet, V. (2016). Combinatorial semi-bandit with known covariance. In *Advances in Neural Information Processing Systems 29*, pages 2972–2980. Curran Associates, Inc.

Ding, W., Qiny, T., Zhang, X.-D., and Liu, T.-Y. (2013). Multi-armed bandit with budget constraint and variable costs.

Feige, U. (1998). A threshold of $\ln n$ for approximating set cover. *Journal of the ACM (JACM)*, 45(4):634–652.

Fujishige, S. (2005). *Submodular functions and optimization*. Annals of discrete mathematics.

Goemans, M. X., Harvey, N. J., Iwata, S., and Mirrokni, V. (2009). Approximating submodular functions everywhere. In *Proceedings of the twentieth annual ACM-SIAM symposium on Discrete algorithms*, pages 535–544. Society for Industrial and Applied Mathematics.

Gomez-Rodriguez, M., Leskovec, J., and Krause, A. (2012). Inferring networks of diffusion and influence. *ACM Trans. Knowl. Discov. Data*, 5(4):21:1–21:37.

Goyal, A., Bonchi, F., and Lakshmanan, L. V. (2010). Learning influence probabilities in social networks. In *Proceedings of the Third ACM International Conference on Web Search and Data Mining*, WSDM '10, pages 241–250, New York, NY, USA. ACM.

Goyal, A., Lu, W., and Lakshmanan, L. V. S. (2011). Simpath: An efficient algorithm for influence maximization under the linear threshold model. In *2011 IEEE 11th International Conference on Data Mining*, pages 211–220.

Kakade, S. M., Kalai, A. T., and Ligett, K. (2009). Playing games with approximation algorithms. *SIAM Journal on Computing*, 39(3):1088–1106.

Kempe, D., Kleinberg, J., and Tardos, É. (2003). Maximizing the spread of influence through a social network. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 137–146. ACM.

Khuller, S., Moss, A., and Naor, J. S. (1999). The budgeted maximum coverage problem. *Information processing letters*, 70(1):39–45.

Krause, A. and Guestrin, C. (2005). A Note on the Budgeted Maximization of Submodular Functions. Technical Report June, CMU.

Leskovec, J., Krause, A., Guestrin, C., Faloutsos, C., Faloutsos, C., VanBriesen, J., and Glance, N. (2007). Cost-effective outbreak detection in networks. In *Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '07, pages 420–429, New York, NY, USA. ACM.

Leskovec, J. and Krevl, A. (2014). SNAP Datasets: Stanford large network dataset collection. http://snap.stanford.edu/data.

Lugosi, G., Neu, G., and Olkhovskaya, J. (2019). Online influence maximization with local observations. In *Proceedings of the 30th International Conference on Algorithmic Learning Theory*, volume 98 of *Proceedings of Machine Learning Research*, pages 557–580, Chicago, Illinois. PMLR.

Magureanu, S., Combes, R., and Proutiere, A. (2014). Lipschitz bandits: Regret lower bound and optimal algorithms. In *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pages 975–999, Barcelona, Spain. PMLR.

Minoux, M. (1978). Accelerated greedy algorithms for maximizing submodular set functions. *Optimization Techniques*, pages 234–243.

Mitzenmacher, M. and Upfal, E. (2017). *Probability and computing: randomization and probabilistic techniques in algorithms and data analysis*. Cambridge university press.

Nemhauser, G. L., Wolsey, L. A., and Fisher, M. L. (1978). An analysis of approximations for maximizing submodular set functions-I. *Mathematical Programming*, 14(1):265–294.

Netrapalli, P. and Sanghavi, S. (2012). Learning the graph of epidemic cascades. *SIGMETRICS Perform. Eval. Rev.*, 40(1):211–222.

Nguyen, H. and Zheng, R. (2013). On budgeted influence maximization in social networks. *IEEE Journal on Selected Areas in Communications*, 31(6):1084–1094.

Perrault, P., Boursier, E., Perchet, V., and Valko, M. (2020a). Statistical efficiency of thompson sampling for combinatorial semi-bandits. *arXiv preprint arXiv:2006.06613*.

Perrault, P., Perchet, V., and Valko, M. (2019a). Exploiting structure of uncertainty for efficient matroid semi-bandits. In *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pages 5123–5132, Long Beach, California, USA. PMLR.

Perrault, P., Perchet, V., and Valko, M. (2019b). Finding the bandit in a graph: Sequential search-and-stop. In *Proceedings of Machine Learning Research*, volume 89 of *Proceedings of Machine Learning Research*, pages 1668–1677. PMLR.

Perrault, P., Perchet, V., and Valko, M. (2020b). Covariance-adapting algorithm for semi-bandits with application to sparse outcomes. *Proceedings of Machine Learning Research*, 125:1–33.

Saito, K., Nakano, R., and Kimura, M. (2008). Prediction of information diffusion probabilities for independent cascade model. In *Knowledge-Based Intelligent Information and Engineering Systems*, pages 67–75, Berlin, Heidelberg. Springer Berlin Heidelberg.

Streeter, M. and Golovin, D. (2009). An online algorithm for maximizing submodular functions. In *Advances in Neural Information Processing Systems*, pages 1577–1584.

Tang, Y., Shi, Y., and Xiao, X. (2015). Influence maximization in near-linear time: A martingale approach. In *Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data*, SIGMOD '15, pages 1539–1554, New York, NY, USA. ACM.

Tang, Y., Xiao, X., and Shi, Y. (2014). Influence maximization: Near-optimal time complexity meets practical efficiency. In *Proceedings of the 2014 ACM SIGMOD international conference on Management of data*, pages 75–86. ACM.

Vaswani, S., Kveton, B., Wen, Z., Ghavamzadeh, M., Lakshmanan, L. V., and Schmidt, M. (2017). Model-independent online learning for influence maximization. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, pages 3530–3539. JMLR. org.

Vaswani, S., Lakshmanan, L. V. S., and Mark Schmidt (2015). Influence maximization with bandits. In *NIPS workshop on Networks in the Social and Information Sciences 2015*.

Wang, Q. and Chen, W. (2017). Improving regret bounds for combinatorial semi-bandits with probabilistically triggered arms and its applications. In *Neural Information Processing Systems*.

Wang, S. and Chen, W. (2018). Thompson Sampling for Combinatorial Semi-Bandits.

Wang, S., Yang, S., Xu, Z., and Truong, V.-A. (2020). Fast thompson sampling algorithm with cumulative oversampling: Application to budgeted influence maximization. *arXiv preprint arXiv:2004.11963*.

Wen, Z., Kveton, B., Valko, M., and Vaswani, S. (2017). Online influence maximization under independent cascade model with semi-bandit feedback. In *Neural Information Processing Systems*.

Xia, Y., Qin, T., Ma, W., Yu, N., and Liu, T.-Y. (2016). Budgeted multi-armed bandits with multiple plays. In *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence*, pages 2210–2216. AAAI Press.

A. Proof of Proposition 1

Proof. Let $\alpha = 1 - 1/e - \varepsilon$. In the proof, we shall consider several policies π one after the other. In each case, we will denote by S_t the seed selected by π at round t, and τ_B the random round where π has exhausted its budget. We denote $(\mathcal{H}_t)_{t\geq 1}$ the filtration corresponding to $(\mathbf{W}_t, \mathbf{C}_t)_{t\geq 1}$. Recall that S_t and B_{t-1} are both measurable with respect to $\mathcal{H}_{t-1}$.

Consider first the policy π that selects $S_t = S^\star \in \arg\max_{S\subset V} \mathbb{E}[\sigma(S;\mathbf{W})]\mathbb{E}\Big[\mathbf{e}_{S\cup\{0\}}^\intercal \mathbf{C}\Big]^{-1}$ at each round $t \geq 1$. We can write

$$\begin{aligned}
F_B^\star + |V| &\geq F_B^\star + \mathbb{E}[\sigma(S^\star;\mathbf{W})] \\
&\geq F_B(\pi) + \mathbb{E}[\sigma(S^\star;\mathbf{W})] && \text{definition of } F_B^\star \\
&= \sum_{t\geq 1} \mathbb{E}[\sigma(S^\star;\mathbf{W}_t)\mathbb{I}\{B_{t-1}\geq 0\}] \\
&= \sum_{t\geq 1} \mathbb{E}[\mathbb{E}[\sigma(S^\star;\mathbf{W})]\mathbb{I}\{B_{t-1}\geq 0\}] && \text{conditioning on } \mathcal{H}_{t-1} \\
&= \lambda^\star \sum_{t\geq 1} \mathbb{E}\Big[\mathbb{E}\Big[\mathbf{e}_{S^\star\cup\{0\}}^\intercal \mathbf{C}\Big]\mathbb{I}\{B_{t-1}\geq 0\}\Big] \\
&= \lambda^\star \sum_{t\geq 1} \mathbb{E}[(C_{0,t} + \mathbf{e}_{S^\star}^\intercal \mathbf{C}_t)\mathbb{I}\{B_{t-1}\geq 0\}] && \text{conditioning on } \mathcal{H}_{t-1} \\
&\geq \lambda^\star B && \text{definition of } \tau_B.
\end{aligned}$$

We can use the inequality

$$F_B^\star + |V| \geq \lambda^\star B \tag{13}$$

with any policy π in the following ways:

- First, we can bound the cost part in the cumulative gap:

$$\begin{aligned}
\alpha\lambda^\star \mathbb{E}\left[\sum_{t=1}^{\tau_B-1} \mathbb{E}\Big[\mathbf{e}_{S_t\cup\{0\}}^\intercal \mathbf{C}\Big]\right] &\leq \alpha\lambda^\star \sum_{t\geq 1} \mathbb{E}\Big[\mathbb{E}\Big[\mathbf{e}_{S_t\cup\{0\}}^\intercal \mathbf{C}\Big]\mathbb{I}\{B_{t-1}\geq 0\}\Big] && B_t \geq 0 \Rightarrow B_{t-1}\geq 0 \\
&= \alpha\lambda^\star \sum_{t\geq 1} \mathbb{E}\big[\big(\mathbf{e}_{S_t}^\intercal \mathbf{C}_t + C_{0,t}\big)\mathbb{I}\{B_{t-1}\geq 0\}\big] && \text{conditioning on } \mathcal{H}_{t-1} \\
&\leq \alpha\lambda^\star B + \alpha\lambda^\star(1+|V|) && \text{definition of } \tau_B, \\
&\leq \alpha F_B^\star + \alpha|V| + \alpha\lambda^\star(1+|V|) && \text{inequality (13)}.
\end{aligned}$$

- Next, we can bound the reward part:

$$\begin{aligned}
\mathbb{E}\left[\sum_{t=1}^{\tau_B-1} \mathbb{E}[\sigma(S_t;\mathbf{W})]\right] &\geq \mathbb{E}\left[\sum_{t=1}^{\tau_B} \mathbb{E}[\sigma(S_t;\mathbf{W})]\right] - |V| \\
&= \sum_{t\geq 1} \mathbb{E}[\mathbb{E}[\sigma(S_t;\mathbf{W})]\mathbb{I}\{B_{t-1}\geq 0\}] - |V| \\
&= \sum_{t\geq 1} \mathbb{E}[\sigma(S_t;\mathbf{W}_t)\mathbb{I}\{B_{t-1}\geq 0\}] - |V| && \text{conditioning on } \mathcal{H}_{t-1} \\
&\geq \sum_{t\geq 1} \mathbb{E}[\sigma(S_t;\mathbf{W}_t)\mathbb{I}\{B_t\geq 0\}] - |V| && B_t \geq 0 \Rightarrow B_{t-1}\geq 0 \\
&= F_B(\pi) - |V|.
\end{aligned}$$

Adding these two inequalities, we get the following upper bound on the cumulative gap:

$$\mathbb{E}\left[\sum_{t=1}^{\tau_B-1} \Delta(S_t)\right] = \alpha\lambda^\star \mathbb{E}\left[\sum_{t=1}^{\tau_B-1} \mathbb{E}\Big[\mathbf{e}_{S_t\cup\{0\}}^\intercal \mathbf{C}\Big]\right] - \mathbb{E}\left[\sum_{t=1}^{\tau_B-1} \mathbb{E}[\sigma(S_t;\mathbf{W})]\right] \leq R_{B,\varepsilon}(\pi) + (\alpha+1)|V| + \alpha\lambda^\star(1+|V|).$$

In the same way, we can derive a lower bound on $\mathbb{E}\Big[\sum_{t=1}^{\tau_B-1}\Delta(S_t)\Big]$, considering first the policy π such that $F_B(\pi)=F_B^\star$:

$$\begin{aligned}
F_B^\star &= \sum_{t\geq 1}\mathbb{E}[\sigma(S_t;\mathbf{W}_t)\mathbb{I}\{B_t\geq 0\}] \\
&\leq \sum_{t\geq 1}\mathbb{E}[\sigma(S_t;\mathbf{W}_t)\mathbb{I}\{B_{t-1}\geq 0\}] && B_t\geq 0\Rightarrow B_{t-1}\geq 0\\
&= \sum_{t\geq 1}\mathbb{E}[\mathbb{E}[\sigma(S_t;\mathbf{W})]\mathbb{I}\{B_{t-1}\geq 0\}] && \text{conditioning on } \mathcal{H}_{t-1}\\
&\leq \sum_{t\geq 1}\mathbb{E}\Big[\lambda^\star\mathbb{E}\Big[\mathbf{e}_{S_t\cup\{0\}}^\intercal\mathbf{C}\Big]\mathbb{I}\{B_{t-1}\geq 0\}\Big] && \text{definition of } \lambda^\star\\
&= \lambda^\star\mathbb{E}\left[\sum_{t=1}^{\tau_B}\big(C_{0,t}+\mathbf{e}_{S_t}^\intercal\mathbf{C}_t\big)\right] && \text{conditioning on } \mathcal{H}_{t-1}\\
&\leq \lambda^\star(B+1+|V|) && \text{definition of } \tau_B.
\end{aligned}$$

i.e.,

$$F_B^\star-\lambda^\star(1+|V|)\leq\lambda^\star B. \tag{14}$$

Considering any policy π:

$$\begin{aligned}
\alpha\lambda^\star\mathbb{E}\left[\sum_{t=1}^{\tau_B-1}\mathbb{E}\Big[\mathbf{e}_{S_t\cup\{0\}}^\intercal\mathbf{C}\Big]\right] &\geq \alpha\lambda^\star\sum_{t\geq 1}\mathbb{E}\Big[\mathbb{E}\Big[\mathbf{e}_{S_t\cup\{0\}}^\intercal\mathbf{C}\Big]\mathbb{I}\{B_{t-1}\geq 1+|V|\}\Big] && B_{t-1}\geq 1+|V|\Rightarrow B_t\geq 0\\
&= \alpha\lambda^\star\sum_{t\geq 1}\mathbb{E}\big[\big(\mathbf{e}_{S_t}^\intercal\mathbf{C}_t+C_{0,t}\big)\mathbb{I}\{B_{t-1}\geq 1+|V|\}\big] && \text{conditioning on } \mathcal{H}_{t-1}\\
&\geq \alpha\lambda^\star B-\alpha\lambda^\star(1+|V|) && \text{definition of } \tau_B.\\
&\geq \alpha F_B^\star-2\alpha\lambda^\star(1+|V|) && \text{inequality (14)},
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E}\left[\sum_{t=1}^{\tau_B-1}\mathbb{E}[\sigma(S_t;\mathbf{W})]\right] &\leq \mathbb{E}\left[\sum_{t=1}^{\tau_B}\mathbb{E}[\sigma(S_t;\mathbf{W})]\right]\\
&= \sum_{t\geq 1}\mathbb{E}[\mathbb{E}[\sigma(S_t;\mathbf{W})]\mathbb{I}\{B_{t-1}\geq 0\}]\\
&= \sum_{t\geq 1}\mathbb{E}[\sigma(S_t;\mathbf{W}_t)\mathbb{I}\{B_{t-1}\geq 0\}] && \text{conditioning on } \mathcal{H}_{t-1}\\
&\leq \sum_{t\geq 1}\mathbb{E}[\sigma(S_t;\mathbf{W}_t)\mathbb{I}\{B_t\geq 0\}]+|V|\\
&= F_B(\pi)+|V|.
\end{aligned}$$

Again adding these two inequalities, we get the desired lower bound:

$$\mathbb{E}\left[\sum_{t=1}^{\tau_B-1}\Delta(S_t)\right]=\alpha\lambda^\star\mathbb{E}\left[\sum_{t=1}^{\tau_B-1}\mathbb{E}\Big[\mathbf{e}_{S_t\cup\{0\}}^\intercal\mathbf{C}\Big]\right]-\mathbb{E}\left[\sum_{t=1}^{\tau_B-1}\mathbb{E}[\sigma(S_t;\mathbf{W})]\right]\geq R_{B,\varepsilon}(\pi)-2\alpha\lambda^\star(1+|V|)-|V|.$$

□

B. Proof of Theorem 1

Proof. Let $\alpha = 1 - 1/e - \varepsilon$, and $t \geq 1$. From Proposition 1, we have to upper bound

$$\mathbb{E}\left[\sum_{t=1}^{\tau_B - 1} \Delta(S_t)\right].$$

Fix $t \geq 1$. We consider the following events:

$$\mathfrak{W}_t \triangleq \left\{\forall ij \in E,\ 0 \leq w^\star_{ij} - w_{ij,t} \leq 2\sqrt{\frac{1.5\log(t)}{N_{\oplus ij,t-1}}}\right\},$$

$$\mathfrak{C}_t \triangleq \left\{\forall i \in V \cup \{0\},\ 0 \leq c^\star_i - c_{i,t} \leq 2\sqrt{\frac{1.5\log(t)}{N_{\ominus i,t-1}}}\right\}.$$

We also consider the event $\mathfrak{A}_t$ under which the $\alpha-$approximation in Algorithm 1 holds. We already saw that

$$\mathbb{P}[\neg\mathfrak{A}_t] \leq \frac{1}{t\log^2(t)}.$$

From Hoeffding inequality, $\mathfrak{C}_t$ doesn't hold with probability bounded by $2(|V|+1)/t^2$, and $\mathfrak{W}_t$ doesn't hold with probability bounded by $2|E|/t^2$. Thus, under the event that either $\mathfrak{C}_t$, $\mathfrak{W}_t$ or $\mathfrak{A}_t$ doesn't hold, then the regret is bounded by a constant (since we have a convergent series).

We thus now assume that $\mathfrak{C}_t$, $\mathfrak{W}_t$ and $\mathfrak{A}_t$ hold. In particular, using $\mathfrak{C}_t$ and $\mathfrak{W}_t$, we can write

$$\lambda^\star = \frac{\sigma(S^\star;\mathbf{w}^\star)}{\mathbf{e}^\intercal_{S^\star\cup\{0\}}\mathbf{c}^\star} \leq \frac{\sigma(S^\star;\mathbf{w}_t)}{\mathbf{e}^\intercal_{S^\star\cup\{0\}}\mathbf{c}_t}.$$

We can use this relation to write

$$\begin{aligned}
\Delta(S_t) &= \lambda^\star\alpha\big(\mathbf{e}^\intercal_{S_t}\mathbf{c}^\star + c^\star_0\big) - \sigma(S_t;\mathbf{w}^\star)\\
&= \lambda^\star\alpha\left(\big(\mathbf{e}^\intercal_{S_t}\mathbf{c}^\star + c^\star_0\big) - \frac{\sigma(S_t;\mathbf{w}_t)}{\alpha\lambda^\star}\right) + (\sigma(S_t;\mathbf{w}_t) - \sigma(S_t;\mathbf{w}^\star))\\
&\leq \lambda^\star\alpha\left(\big(\mathbf{e}^\intercal_{S_t}\mathbf{c}^\star + c^\star_0\big) - \frac{\sigma(S_t;\mathbf{w}_t)}{\alpha\sigma(S^\star;\mathbf{w}_t)}(\mathbf{e}^\intercal_{S^\star}\mathbf{c}_t + c_{0,t})\right) + (\sigma(S_t;\mathbf{w}_t) - \sigma(S_t;\mathbf{w}^\star))\\
&\leq \lambda^\star\alpha\big(\big(\mathbf{e}^\intercal_{S_t}\mathbf{c}^\star + c^\star_0\big) - \big(\mathbf{e}^\intercal_{S_t}\mathbf{c}_t + c_{0,t}\big)\big) + (\sigma(S_t;\mathbf{w}_t) - \sigma(S_t;\mathbf{w}^\star)),
\end{aligned}$$

where the last inequality is from $\mathfrak{A}_t$. Notice here that in the case the costs are known, the first term in this bound disappears, and we can then safely take $\lambda^\star = 0$, explaining why in the final bound the term in front of $\lambda^\star$ disappears. We now use Fact 2, and then $\mathfrak{C}_t$, $\mathfrak{W}_t$ to further get the bound

$$\Delta(S_t) \leq \underbrace{\lambda^\star\alpha\sum_{i\in S_t}\left(1 \wedge 2\sqrt{\frac{1.5\log(t)}{N_{\ominus i,t-1}}}\right)}_{(15)} + \underbrace{|V|\sum_{ij\in E} p_i(S_t;\mathbf{w}^\star)\left(1 \wedge 2\sqrt{\frac{1.5\log(t)}{N_{\oplus ij,t-1}}}\right)}_{(16)}.$$

Then, necessarily either $\Delta(S_t) \leq 2\cdot(15)$ or $\Delta(S_t) \leq 2\cdot(16)$ is true. The first event can be handle exactly as in standard combinatorial semi bandit settings, using the following upper bound on the expectation of the random horizon (Perrault et al., 2019b):

$$\mathbb{E}[\tau_B - 1] \leq (2B/c^\star_0 + 1)^2.$$

This allows us to get a term of order

$$\lambda^\star\log(B/c^\star_0)\sum_{i\in V}\frac{|V|}{\Delta_{i,\min}},$$

in the regret upper bound.

For the second event, the analysis of (Wang and Chen, 2017) uses triggering probability group to deal with $p_i(S_t;\mathbf{w}^\star)$. We propose here another method, simpler, that allows us to transform the factor $|E|$, present in the bound of (Wang and Chen, 2017), into $\max_{S\subset V}\sum_{i\in V} d_i p_i(S;\mathbf{w}^\star)$, a potentially much smaller quantity. More precisely, we first use a reverse amortisation:

$$\begin{aligned}\Delta(S_t) \le -\Delta(S_t) + 4\cdot(16) &= 4|V|\sum_{ij\in E} p_i(S_t;\mathbf{w}^\star)\left(1\wedge 2\sqrt{\frac{1.5\log(t)}{N_{\oplus ij,t-1}}} - \frac{\Delta(S_t)}{4|V|\sum_{k\in V} d_k p_k(S_t;\mathbf{w}^\star)}\right)\\ &\le 4|V|\sum_{ij\in E} p_i(S_t;\mathbf{w}^\star)\mathbb{I}\left\{N_{\oplus ij,t-1}\le 1.5\log(t)\left(\frac{8|V|\sum_{k\in V} d_k p_k(S_t;\mathbf{w}^\star)}{\Delta(S_t)}\right)^2\right\}\left(1\wedge 2\sqrt{\frac{1.5\log(t)}{N_{\oplus ij,t-1}}}\right).\end{aligned}$$

Since we have

$$\mathbb{E}\left[\sum_{t=1}^{\tau_B-1}\Delta(S_t)\right] \le \mathbb{E}\left[\sum_{t=1}^{\tau_B}\mathbb{E}[\Delta(S_t)|\mathcal{H}_{t-1}]\right],$$

and that

$$\mathbb{E}\left[p_i(S_t;\mathbf{w}^\star)\mathbb{I}\left\{N_{\oplus ij,t-1}\le 1.5\log(t)\left(\frac{8|V|\sum_{k\in V} d_k p_k(S_t;\mathbf{w}^\star)}{\Delta(S_t)}\right)^2\right\}\left(1\wedge 2\sqrt{\frac{1.5\log(t)}{N_{\oplus ij,t-1}}}\right)\middle|\mathcal{H}_{t-1}\right]$$

equals

$$\mathbb{E}\left[\mathbb{I}\left\{S_t \overset{\mathbf{w}^\star}{\rightsquigarrow} i\right\}\mathbb{I}\left\{N_{\oplus ij,t-1}\le 1.5\log(t)\left(\frac{8|V|\sum_{k\in V} d_k p_k(S_t;\mathbf{w}^\star)}{\Delta(S_t)}\right)^2\right\}\left(1\wedge 2\sqrt{\frac{1.5\log(t)}{N_{\oplus ij,t-1}}}\right)\middle|\mathcal{H}_{t-1}\right],$$

it is sufficient to bound the quantity

$$\sum_{t=1}^{\tau_B} 4|V|\sum_{ij\in E,\ p_i(S_t;\mathbf{w}^\star)>0}\mathbb{I}\left\{S_t \overset{\mathbf{w}^\star}{\rightsquigarrow} i\right\}\mathbb{I}\left\{N_{\oplus ij,t-1}\le 1.5\log(t)\left(\frac{8|V|\sum_{k\in V} d_k p_k(S_t;\mathbf{w}^\star)}{\Delta(S_t)}\right)^2\right\}\left(1\wedge 2\sqrt{\frac{1.5\log(t)}{N_{\oplus ij,t-1}}}\right).$$

Therefore, counters $N_{\oplus ij,t-1}$ are ensured to increase thanks to the event $\left\{S_t \overset{\mathbf{w}^\star}{\rightsquigarrow} i\right\}$. We can now handle this exactly as in standard combinatorial semi bandit setting, to get a bound of order

$$\log(B/c_0^\star)\sum_{i\in V} d_i \frac{|V|^2\max_{S\subset V,\ p_i(S;\mathbf{w}^\star)>0}\sum_{k\in V} d_k p_k(S;\mathbf{w}^\star)}{\Delta_{i,\min}}.$$

□

Problem-independent bound The problem-independent bound of $\mathcal{O}\big(|V|\sqrt{B\log B\sum_{i\in V} d_i p_{i,\max}}\big)$ is an immediate consequence of our problem-dependent bound, decomposing, classically, the regret in two terms by filtering by whether or not $\Delta(S_t)\le\delta$, and then taking the worst regime for δ.

C. Proof of Proposition 3

Proof. First, notice that we trivially have

$$\mathbb{P}\left[\mathfrak{P}_t \text{ and } p_i(\{j\};\mathbf{w}^\star)\le\frac{8\delta(t)}{N_{\ominus j,t-1}} \text{ and } \frac{\delta(t)p_i(\{j\};\mathbf{w}^\star)}{N_{\oplus i,t-1}} > \frac{8\delta(t)}{N_{\ominus j,t-1}}\right] = 0.$$

Thus, let's prove that

$$\mathbb{P}\left[\mathfrak{P}_t \text{ and } p_i(\{j\};\mathbf{w}^\star) > \frac{8\delta(t)}{N_{\ominus j,t-1}} \text{ and } \frac{\delta(t)p_i(\{j\};\mathbf{w}^\star)}{N_{\oplus i,t-1}} > \frac{8\delta(t)}{N_{\ominus j,t-1}}\right] \le 1/t^2.$$

We define another counter for $(i,j) \in V^2$ as follows:

$$N_{i,j,t-1} \triangleq \sum_{t'=1}^{t-1} \mathbb{I}\Big\{ j \in S_{t'},\ \{j\} \overset{\mathbf{w}_{t'}}{\rightsquigarrow} i \Big\}.$$

Note that we have $N_{i,j,t-1} \leq (N_{\oplus i,t-1} \wedge N_{\ominus j,t-1})$. We can thus remove $\mathfrak{P}_t$ and replace $N_{\oplus i,t-1}$ by $N_{i,j,t-1}$, since this can only increases the probability. By an union bound we have,

$$\mathbb{P}\bigg[p_i(\{j\};\mathbf{w}^\star) > \frac{8\delta(t)}{N_{\ominus j,t-1}} \text{ and } \frac{p_i(\{j\};\mathbf{w}^\star)}{N_{i,j,t-1}} > \frac{8}{N_{\ominus j,t-1}} \bigg] \leq \sum_{t' > \frac{8\delta(t)}{p_i(\{j\};\mathbf{w}^\star)}} \mathbb{P}\bigg[N_{\ominus j,t-1} = t',\ \frac{t' p_i(\{j\};\mathbf{w}^\star)}{8} > N_{i,j,t-1} \bigg].$$

Since the random variables $\mathbb{I}\Big\{ \{j\} \overset{\mathbf{w}_{t''}}{\rightsquigarrow} i \Big\}$ are bernouillies of mean $p_i(\{j\};\mathbf{w}^\star)$, we can apply the Fact 3 to get

$$\mathbb{P}\bigg[N_{\ominus j,t-1} = t',\ \frac{t' p_i(\{j\};\mathbf{w}^\star)}{8} > N_{i,j,t-1} \bigg] \leq \exp\Big(-(7/8)^2 8\delta(t)/2 \Big) < 1/t^3.$$

By taking t' over $\{0,\dots,t-1\}$, the proposition holds.

Fact 3 (Multiplicative Chernoff Bound (Mitzenmacher and Upfal, 2017))**.** *Let $X_1,\dots,X_t$ be Bernoulli random variables, of parameter μ, then for $Y = X_1 + \cdots + X_t$, we have with $\delta \in (0,1)$,*

$$\mathbb{P}[Y \leq (1-\delta)t\mu] \leq e^{-\delta^2 t\mu/2}.$$

□

D. Proof of Proposition 2

Proof. In the proof, we use the notation $\sigma(i|S) \triangleq \sigma(\{i\} \cup S) - \sigma(S)$. For any $k \in [|V|]$,

$$\begin{aligned}
\sigma(S^\star) - \sigma(S_{k-1}) &\leq \sum_{i \in S^\star \setminus S_{k-1}} \sigma(i|S_{k-1}) && \text{Submodularity, monotonicity of } \sigma \\
&\leq \frac{\sigma(i_k|S_{k-1})}{c_{i_k}} \sum_{i \in S^\star \setminus S_{k-1}} c_i && \text{Algorithm 2} \\
&\leq \frac{\sigma(i_k|S_{k-1})}{c_{i_k}} \sum_{i \in S^\star} c_i.
\end{aligned}$$

i.e., for all $k \in [|V|]$ such that $\sigma(S^\star) - \sigma(S_{k-1}) \geq 0$,

$$\frac{c_{i_k}}{\mathbf{e}_{S^\star}^\mathsf{T}\mathbf{c}} \leq \frac{\sigma(i_k|S_{k-1})}{\sigma(S^\star) - \sigma(S_{k-1})}. \tag{17}$$

There must be an index $\ell \in \{0,1,\dots,|V|-1\}$ such that $\mathbf{e}_{S_\ell}^\mathsf{T}\mathbf{c} \leq \mathbf{e}_{S^\star}^\mathsf{T}\mathbf{c} \leq \mathbf{e}_{S_{\ell+1}}^\mathsf{T}\mathbf{c}$. Let $\beta \in [0,1]$ be such that

$$\mathbf{e}_{S^\star}^\mathsf{T}\mathbf{c} = (1-\beta)\mathbf{e}_{S_\ell}^\mathsf{T}\mathbf{c} + \beta \mathbf{e}_{S_{\ell+1}}^\mathsf{T}\mathbf{c}. \tag{18}$$

If $\sigma(S^\star) - (1-\beta)\sigma(S_\ell) - \beta\sigma(S_{\ell+1}) \leq 0$, then we have

$$\big(1 - e^{-1}\big)\frac{\sigma(S^\star)}{\mathbf{e}_{S^\star}^\mathsf{T}\mathbf{c}} \leq \frac{\sigma(S^\star)}{\mathbf{e}_{S^\star}^\mathsf{T}\mathbf{c}} \leq \frac{(1-\beta)\sigma(S_\ell) + \beta\sigma(S_{\ell+1})}{(1-\beta)\mathbf{e}_{S_\ell}^\mathsf{T}\mathbf{c} + \beta\mathbf{e}_{S_{\ell+1}}^\mathsf{T}\mathbf{c}}.$$

Else,

$$\sigma(S^\star) - (1-\beta)\sigma(S_\ell) - \beta\sigma(S_{\ell+1}) > 0 \quad \text{and} \quad \sigma(S^\star) - \sigma(S_k) > 0 \quad \text{for all } k \in [\ell], \tag{19}$$

so we can write the following,

$$\begin{aligned}
\frac{\sigma(S^\star)-(1-\beta)\sigma(S_\ell)-\beta\sigma(S_{\ell+1})}{\sigma(S^\star)} &\le \frac{\sigma(S^\star)-(1-\beta)\sigma(S_\ell)-\beta\sigma(S_{\ell+1})}{\sigma(S^\star|\emptyset)} \\
&= \frac{\sigma(S^\star)-\sigma(S_\ell)-\beta\sigma(i_{\ell+1}|S_\ell)}{\sigma(S^\star)-\sigma(S_\ell)} \prod_{k\in[\ell]} \frac{\sigma(S^\star)-\sigma(S_k)}{\sigma(S^\star)-\sigma(S_{k-1})} \\
&= \left(1-\frac{\beta\sigma(i_{\ell+1}|S_\ell)}{\sigma(S^\star)-\sigma(S_\ell)}\right) \prod_{k\in[\ell]} \left(1-\frac{\sigma(i_k|S_{k-1})}{\sigma(S^\star)-\sigma(S_{k-1})}\right) \\
&\le \left(1-\frac{\beta c_{i_{\ell+1}}}{\mathbf{e}_{S^\star}^\intercal \mathbf{c}}\right) \prod_{k\in[\ell]} \left(1-\frac{c_{i_k}}{\mathbf{e}_{S^\star}^\intercal \mathbf{c}}\right) && \text{(17) and (19)} \\
&\le \exp\left(-\frac{\beta c_{i_{\ell+1}}+\sum_{k\in[\ell]} c(i_k)}{\mathbf{e}_{S^\star}^\intercal \mathbf{c}}\right) && 1-x\le e^{-x} \\
&= \exp\left(-\frac{(1-\beta)\mathbf{e}_{S_\ell}^\intercal \mathbf{c}+\beta\mathbf{e}_{S_{\ell+1}}^\intercal \mathbf{c}}{\mathbf{e}_{S^\star}^\intercal \mathbf{c}}\right) = e^{-1}. && \text{(18)}
\end{aligned}$$

Rearranging the inequality, we obtain the following:

$$\left(1-e^{-1}\right)\sigma(S^\star) \le (1-\beta)\sigma(S_\ell)+\beta\sigma(S_{\ell+1}). \tag{20}$$

i.e.,

$$\left(1-e^{-1}\right)\frac{\sigma(S^\star)}{\mathbf{e}_{S^\star\cup\{0\}}^\intercal \mathbf{c}} \le \frac{(1-\beta)\sigma(S_\ell)+\beta\sigma(S_{\ell+1})}{(1-\beta)\mathbf{e}_{S_\ell\cup\{0\}}^\intercal \mathbf{c}+\beta\mathbf{e}_{S_{\ell+1}\cup\{0\}}^\intercal \mathbf{c}}.$$

The output S of Algorithm 2 maximizes the ratio of $\sigma(S_k)/\mathbf{e}_{S_k\cup\{0\}}^\intercal \mathbf{c}$ over k. Thus,

$$\max_{k\le\ell+1} \frac{\sigma(S_k)}{\mathbf{e}_{S_k\cup\{0\}}^\intercal \mathbf{c}} \le \frac{\sigma(S)}{\mathbf{e}_{S\cup\{0\}}^\intercal \mathbf{c}}.$$

We end the proof remarking that

$$\max_{k\in\{\ell,\ell+1\}} \frac{\sigma(S_k)}{\mathbf{e}_{S_k\cup\{0\}}^\intercal \mathbf{c}} \ge \frac{(1-\beta)\sigma(S_\ell)+\beta\sigma(S_{\ell+1})}{(1-\beta)\mathbf{e}_{S_\ell\cup\{0\}}^\intercal \mathbf{c}+\beta\mathbf{e}_{S_{\ell+1}\cup\{0\}}^\intercal \mathbf{c}}.$$

□

E. Proof of Theorem 2

Proof. Let $\alpha = 1-1/e-\varepsilon$, and $t \ge 1$. From Proposition 1, we have to upper bound

$$\mathbb{E}\left[\sum_{t=1}^{\tau_B-1} \Delta(S_t)\right].$$

In the proof, in addition to $\mathfrak{P}_t \triangleq \{\forall i \in V, N_{\oplus i,t-1} \ge \delta(t)\}$, we consider the following events:

$$\mathfrak{W}_t \triangleq \left\{\sum_{ij\in E} N_{\oplus i,t-1}\left(w_{ij}^\star - \overline{w}_{ij,t-1}\right)^2 \le 2\delta(t)\right\},$$

$$\mathfrak{C}_t \triangleq \left\{\forall i \in V\cup\{0\},\ 0 \le c_i^\star - c_{i,t} \le 1 \wedge 2\sqrt{\frac{1.5\log(t)}{N_{\ominus i,t-1}}}\right\}.$$

$$\mathfrak{B}_t \triangleq \left\{\forall i,j \in V,\ \frac{\delta(t)p_i(\{j\};\mathbf{w}^\star)}{N_{\oplus i,t-1}} \le \frac{8\delta(t)}{N_{\ominus j,t-1}}\right\}.$$

As previously, we have the following upper bound on the expectation of the random horizon:

$$\mathbb{E}[\tau_B - 1] \leq (2B/c_0^\star + 1)^2.$$

For each node $i \in V$, if $N_{\oplus i,t-1} \geq \delta(\tau_B - 1)$, then i will not be intentionally added to the seed set in BOIM-CUCB$_5$. Then, each node is intentionally added for at most $\delta(\tau_B - 1) + 1$ times. Thus, we can write

$$\mathbb{E}\left[\sum_{t=1}^{\tau_B - 1} \Delta(S_t)\mathbb{I}\{\neg\mathfrak{P}_t\}\right] \leq \mathbb{E}[(\delta(\tau_B - 1) + 1)|V|\lambda^\star(|V| + 1)] \leq \Big(\delta\Big((2B/c_0^\star + 1)^2\Big) + 1\Big)|V|\lambda^\star(|V| + 1).$$

We can therefore assume that $\mathfrak{P}_t$ holds. In this case, we have by Proposition 3 that $\mathfrak{B}_t$ doesn't hold with probability bounded by $|V|^2/t^2$. On the other hand, from Fact 1, $\mathfrak{W}_t$ doesn't hold with probability bounded by $1/\big(t\log^2(t)\big)$, and from Hoeffding inequality, $\mathfrak{C}_t$ doesn't hold with probability bounded by $2(|V| + 1)/t^2$. We can consider the event $\mathfrak{A}_t$ under which the $\alpha-$approximation in BOIM-CUCB$_5$ holds. We already saw that

$$\mathbb{P}[\neg\mathfrak{A}_t] \leq \frac{1}{t\log^2(t)}.$$

The regret in the case one of the events $\neg\mathfrak{A}_t, \neg\mathfrak{B}_t, \neg\mathfrak{W}_t, \neg\mathfrak{C}_t$ holds is thus bounded by a constant depending on $|V|$ and $\lambda^\star$. It thus remains to upper bound

$$\mathbb{E}\left[\sum_{t=1}^{\tau_B - 1} \Delta(S_t)\mathbb{I}\{\mathfrak{A}_t, \mathfrak{B}_t, \mathfrak{W}_t, \mathfrak{C}_t\}\right].$$

For this, notice that from $\mathfrak{A}_t$, S_t which is the seed set chosen by our policy at round t, is an α-approximate maximizer of $A \mapsto f(A)/(\mathbf{e}_A^\intercal \mathbf{c}_t + c_{0,t})$, where f is one of the optimistic spreads considered in BOIM-CUCB$_5$. We thus have

$$\frac{f(S_t)}{\mathbf{e}_{S_t}^\intercal \mathbf{c}_t + c_{0,t}} \geq \alpha \frac{f(S^\star)}{\mathbf{e}_{S^\star}^\intercal \mathbf{c}_t + c_{0,t}},$$

where $S^\star \in \arg\max_{S\subset V} \frac{\sigma(S;\mathbf{w}^\star)}{\mathbf{e}_S^\intercal \mathbf{c}^\star + c_0^\star}$. Since under $\mathfrak{W}_t$, $f(S^\star) \geq \sigma(S^\star;\mathbf{w}^\star)$, we can derive the following upper bound on the gap:

$$\begin{aligned}
\Delta(S_t) &= \lambda^\star\alpha\big(\mathbf{e}_{S_t}^\intercal \mathbf{c}^\star + c_0^\star\big) - \sigma(S_t;\mathbf{w}^\star) \\
&= \lambda^\star\alpha\left(\big(\mathbf{e}_{S_t}^\intercal \mathbf{c}^\star + c_0^\star\big) - \frac{f(S_t)}{\alpha\lambda^\star}\right) + (f(S_t) - \sigma(S_t;\mathbf{w}^\star)) \\
&\leq \lambda^\star\alpha\left(\big(\mathbf{e}_{S_t}^\intercal \mathbf{c}^\star + c_0^\star\big) - \frac{f(S_t)}{\alpha f(S^\star)}(\mathbf{e}_{S^\star}^\intercal \mathbf{c}_t + c_{0,t})\right) + (f(S_t) - \sigma(S_t;\mathbf{w}^\star)) && \mathfrak{W}_t, \mathfrak{C}_t \\
&\leq \lambda^\star\alpha\big(\big(\mathbf{e}_{S_t}^\intercal \mathbf{c}^\star + c_0^\star\big) - \big(\mathbf{e}_{S_t}^\intercal \mathbf{c}_t + c_{0,t}\big)\big) + (f(S_t) - \sigma(S_t;\mathbf{w}^\star)). && \mathfrak{A}_t
\end{aligned}$$

From this point, we can use the condition satisfied by f in BOIM-CUCB$_5$:

$$f(S_t) \leq \sigma(S_t;\overline{\mathbf{w}}_{t-1}) + \text{Bonus}_5(S_t).$$

Using Fact 2 with Fact 1, we can further have with Cauchy-Schwartz inequality

$$\sigma(S_t;\overline{\mathbf{w}}_{t-1}) - \sigma(S_t;\mathbf{w}^\star) \leq \text{Bonus}_5(S_t).$$

This allows us to get, using $\mathfrak{C}_t$,

$$\Delta(S_t) \leq \lambda^\star\alpha \sum_{i\in S_t\cup\{0\}} 1 \wedge 2\sqrt{\frac{1.5\log(t)}{N_{\ominus i,t-1}}} + 2\text{Bonus}_5(S_t).$$

Since we have $N_{\ominus 0,t-1} = t$, we can remove $\{0\}$ in $S_t \cup \{0\}$, by looking at the regret under the event $\Big\{2\lambda^\star\alpha\sqrt{1.5\log(t)/t} \leq \Delta(S_t)/2\Big\}$, giving

$$\Delta(S_t) \leq 2\lambda^\star\alpha \sum_{i\in S_t} 1 \wedge 2\sqrt{\frac{1.5\log(t)}{N_{\ominus i,t-1}}} + 4\text{Bonus}_5(S_t). \tag{21}$$

The regret upper bound in the case this event doesn't hold is bounded by a constant depending on the inverse of the squared minimum gap and $\lambda^\star$. The first part in (21) can be handle exactly as in standard combinatorial budgeted semi bandit settings, to get a term of order

$$\lambda^\star \log(B/c_0^\star) \sum_{i \in V} \frac{|V|}{\Delta_{i,\min}},$$

in the regret upper bound. We can use the analysis of Degenne and Perchet (2016) to deal with the second part, to get a term of order

$$\delta(B/c_0^\star)|V|^2|E| \sum_{i \in V} \frac{\log^2(|V|)}{\Delta_{i,\min}},$$

in the regret upper bound. We thus get the desired result. □

F. Proof of Theorem 3

Proof. Let $\alpha = 1 - 1/e - \varepsilon$, and $t \geq 1$. The beginning of the proof is the same as in Theorem 2, except we no longer consider the event $\mathfrak{B}_t$, and we consider a new event:

$$\mathfrak{R}_t \triangleq \{\forall i \in V, N_{\oplus i,t-1} \geq |E|\delta(t)\}.$$

As for Theorem 2:

- The regret in the case $\mathfrak{R}_t$ doesn't hold can be bounded by a term of order

$$\lambda^\star |V|^2 |E| \log(B).$$

- When all the events hold, the same analysis gives

$$\Delta(S_t) \leq 2\lambda^\star \alpha \sum_{i \in S_t} 1 \wedge 2\sqrt{\frac{1.5\log(t)}{N_{\ominus i,t-1}}} + 4\text{Bonus}_1(S_t; \overline{\mathbf{w}}_{t-1}),$$

  and the first term can be handled in the same way.

The second term can be analyzed in the following way: After bounding it by $4m\text{Bonus}(S_t; \overline{\mathbf{w}}_{t-1})$, see that using Fact 2 on the quantity $p_i(S; \overline{\mathbf{w}}_{t-1})$ present in this bonus, we get

$$p_i(S_t; \overline{\mathbf{w}}_{t-1}) \leq p_i(S_t; \mathbf{w}^\star) + \frac{1}{|V|}\text{Bonus}(S_t; \mathbf{w}^\star).$$

By subadditivity, and from $\mathfrak{R}_t$, we have

$$\begin{aligned} 4m\text{Bonus}(S_t; \overline{\mathbf{w}}_{t-1}) &\leq 4m\text{Bonus}(S_t; \mathbf{w}^\star) + 4m|V|\sqrt{\delta(t)\sum_{i \in V} d_i \frac{\text{Bonus}(S_t; \mathbf{w}^\star)^2}{|V|^2 N_{\oplus i,t-1}}} \\ &\leq 4m\text{Bonus}(S_t; \mathbf{w}^\star) + 4m|V|\sqrt{\sum_{i \in V} d_i \frac{\text{Bonus}(S_t; \mathbf{w}^\star)^2}{|V|^2|E|}} = 8m\text{Bonus}(S_t; \mathbf{w}^\star). \end{aligned}$$

We can now use the analysis of Degenne and Perchet (2016), together with the one from Wang and Chen (2017) to deal with probabilistically triggered arms, to get in the regret upper bound a term of order

$$\delta(B/c_0^\star) m^2 |V|^2 \sum_{i \in V} d_i \frac{\log^2(|E|)}{\Delta_{i,\min}}.$$

□

G. SKIM for influence maximization with cost

In this section, we provide an adaptation of SKIM (Cohen et al., 2014) to our ratio maximization setting. Let $(\mathbf{w}, \text{BONUS}) = (\mathbf{w}_t, 0)$ or $(\overline{\mathbf{w}}_{t-1}, \text{Bonus}_5)$ (depending if we want to maximize the CUCB ratio or the Bonus_5 based ratio), SKIM for IC with weights $\mathbf{w}$ can be used (Cohen et al., 2014), but instead of taking $k-1$ as the threshold for the length of the sketch of node i (i.e. i is chosen as soon as $|\text{sketch}[i]| > k-1$), we rather consider

$$kc_i - \frac{\text{BONUS}(\{i\} \cup S) - \text{BONUS}(S)}{|V|}\text{sketch}[i][-1],$$

where $\text{sketch}[i][-1]$ is the last added rank in $\text{sketch}[i]$ (0 if the sketch is empty), and S is the current seed set built so far. This way, we can estimate the (non-optimistic) marginal gain spread of the chosen node i as

$$\frac{k|V|c_i - (\text{BONUS}(\{i\} \cup S) - \text{BONUS}(S))\text{sketch}[i][-1]}{\text{sketch}[i][-1]}.$$

We get the optimistic version by adding $\text{BONUS}(\{i\} \cup S) - \text{BONUS}(S)$, i.e., it is $\frac{k|V|c_i}{\text{sketch}[i][-1]}$. Finally, we get the ratio marginal gain by dived by c_i, giving $\frac{k|V|}{\text{sketch}[i][-1]}$. We do maximize the ratio marginal gain by doing this all procedure, because the ranks are examined in ascending order. Notice we have to normalize costs at the beginning of the loop for finding i, such that each of the threshold are greater than $k-1$, to ensure that the length of the chosen sketch is at least k and that our estimation hold with high probability.

It should be noted that the main difficulty is to go through the ranks in the right order while taking into account the cost. An alternative would have been to draw classical ranks and change them at each update of the sketch of u according to the cost of u. This procedure works for estimating the spread, but we can't guarantee that the ranks are in the descending order of the marginal ratio, so all the marginal ratios must be estimated before selecting the maximizer.

Adapting Lemma 4.2 from Cohen et al. (2014), we obtain that the expected total number of rank insertions at a particular node is $\mathcal{O}\left(k \log\left(|V| k \frac{c_{\max}}{c_{\min}}\right)\right)$, thus giving a global complexity of $\mathcal{O}\left(|E| k \log\left(|V| k \frac{c_{\max}}{c_{\min}}\right)\right)$. We note the dependence in $c_m in$, which although not desired, is only logarithmic. When all the costs are equal, we recover the standard SKIM complexity.

H. Evaluating bonuses with sketches

In this section, we give details on the bonuses evaluation. Notice that the optimistic spread

$$\sigma(S; \overline{\mathbf{w}}_{t-1}) + \text{Bonus}_2(S; \overline{\mathbf{w}}_{t-1}) = \sum_{i \in V} p_i(S; \overline{\mathbf{w}}_{t-1}) \left(1 + |V| \sqrt{\frac{\delta(t) d_i}{N_{\oplus i, t-1}}}\right) \tag{22}$$

is actually a weighted spread, with weights $1 + |V|\sqrt{\frac{\delta(t)d_i}{N_{\oplus i,t-1}}}$. Thus, ranks used in the sketching have to be drawn from a distribution that depends on these weights (Cohen, 2016). This can be done using the exponential or uniform distribution (Cohen, 1997; Cohen and Kaplan, 2007). $\text{Bonus}_3(A)$ is a square root of a weighted spread, and the same as above holds. For $\text{Bonus}_1(A)$, in addition to the above weighted consideration, with weights $\delta(t)d_i/N_{\oplus i,t-1}$, we have to take care of the squared probability $p_i(\{j\}; \overline{\mathbf{w}}_{t-1})^2$. To do so, the graph is replicated in each instance of the sketching, i.e., instead of considering combined reachability sets with a single graph per instance, we consider two independent graphs per instance, and look at node-instance pairs satisfying the reachability on *both* graphs. Notice, we leverage here on the fact that we only evaluate on sets that are singletons $A = \{j\}$. Indeed, in this case, for a node $i \in V$, the probability that A reaches i on one instance, squared, is equal to the probability that some node in A reaches i on both instances.

I. Further experiments

In this section, we present other experiments that we conducted on a complete 10 node graph, with known costs $c_0^\star = 1$, and for all $i \in V$, $c_i^\star$ is randomly drawn in $(0, 1)$. We also chose $\mathbf{w}^\star \sim U(0, 0.1)^{\otimes E}$, as in Section 7. We compare the BOIM-CUCB algorithm to BOIM-CUCB-REGULARIZED, another algorithm that might challenge BOIM-CUCB in our setting. BOIM-CUCB-REGULARIZED is exactly as BOIM-CUCB except that the objective that is optimized is $S \mapsto \sigma(S; \mathbf{w}_t) - \lambda \mathbf{e}_S \mathbf{c}_t$, for λ being an input parameter to the algorithm. We can see that as for BOIM-CUCB, this algorithm have the willingness to

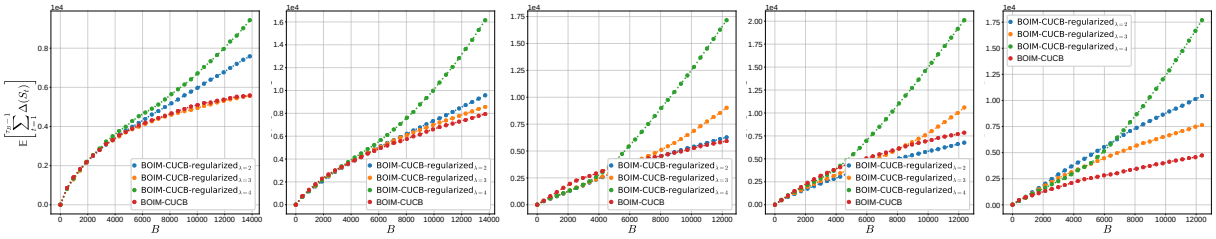

Figure 2. Regret curves on five different problem instances, with respect to the budget B (expectation computed by averaging over 10 independent simulations).

maximize the function σ while minimizing the cost function. The fundamental difference is on the importance given to one or the other function, controled by λ. We use a greedy maximization in BOIM-CUCB-REGULARIZED. A greedy optimization of the objective $S \mapsto \sigma(S; \mathbf{w}_t) - \lambda \mathbf{e}_S \mathbf{c}_t$ is a heuristic which, although not supported in theory, performs well in practice.

We run experiments over up to $T = 10000$ rounds, on five different draws for $\mathbf{w}^\star$ and $\mathbf{c}^\star$, and 3 different values $\lambda = 2, 3, 4$. Results are shown in Figure 2. We observe that BOIM-CUCB is in general better than BOIM-CUCB-REGULARIZED. If the variable λ is properly chosen, performances similar to BOIM-CUCB can be obtained. This is not surprising since BOIM-CUCB aims (but only approximatively) to select $S_t^\star \in \arg\max_{S \subset V} \sigma(S; \mathbf{w}_t)/\mathbf{e}_{S \cup \{0\}}$. If $\lambda = \sigma(S_t^\star; \mathbf{w}_t)/\mathbf{e}_{S_t^\star \cup \{0\}} \mathbf{c}_t$, then we also have that BOIM-CUCB-REGULARIZED aims at choosing $S_t^\star$, since one can notice that $S_t^\star \in \arg\max_{S \subset V} \sigma(S; \mathbf{w}_t) - \lambda \mathbf{e}_S \mathbf{c}_t$.

J. Proof of Theorem 4

J.1. Preliminaries

If we let $p_{S \rightsquigarrow S'}(\mathbf{w}) \triangleq \mathbb{P}\Big[\Big\{i \in V,\ S \overset{\mathbf{w}}{\rightsquigarrow} i\Big\} = S'\Big]$, then another expression is

$$\begin{aligned}\text{Bonus}_4(S; \mathbf{w}) &= \sum_{S' \supset S} p_{S \rightsquigarrow S'}(\mathbf{w}) \underbrace{|V| \sqrt{\delta(t) \sum_{i \in S'} \frac{d_i}{N_{\oplus i, t-1}}}}_{g(S')} \\ &= \sum_{k \geq 0} \big(g(S'_{k+1}) - g(S'_k)\big) \sum_{S' \notin \{S'_0, \ldots, S'_k\}} p_{S \rightsquigarrow S'}(\mathbf{w}).\end{aligned}$$

where $g(S) = g(S'_1) \leq g(S'_2) \leq \ldots$ and $S'_0 = \emptyset$.

Since this bonus shall be used with $\overline{\mathbf{w}}_{t-1}$, we need a smoothness inequality to link $p_{S \rightsquigarrow S'}(\overline{\mathbf{w}}_{t-1})$ to $p_{S \rightsquigarrow S'}(\mathbf{w}^\star)$. We prove here the following such inequality.

Proposition 5. *For all $S \subset V$, all $\mathbf{w}, \mathbf{w}' \in [0,1]^E$ and all collection of subsets of vertices $\mathcal{S}$, we have*

$$\left| \sum_{S' \in \mathcal{S}} \big(p_{S \rightsquigarrow S'}(\mathbf{w}) - p_{S \rightsquigarrow S'}(\mathbf{w}')\big) \right| \leq \sum_{ij \in E} p_i(S; \mathbf{w}) \big| w'_{ij} - w_{ij} \big|.$$

Proof. We assume w.l.o.g. that $\mathbf{w}' \geq \mathbf{w}$. We consider the random graph $G_{\mathbf{W}} = (V, \{ij \in E, W_{ij} = 1\})$, where $\mathbf{W} \sim \otimes_{ij \in E} \text{Bernoulli}(w_{ij})$. We build $G_{\mathbf{W}'}$ from $G_{\mathbf{W}}$ by adding edges ij independently with probability $\frac{w'_{ij} - w_{ij}}{1 - w_{ij}}$ for each ij that is not an edge in $G_{\mathbf{W}}$. Now, see that

$$\sum_{S' \in \mathcal{S}} p_{S \rightsquigarrow S'}(\mathbf{w}) = \mathbb{P}\Big[S \overset{\mathbf{W}}{\rightsquigarrow} S'\Big] - \sum_{S' \notin \mathcal{S}} p_{S \rightsquigarrow S'}(\mathbf{w}),$$

where $S \overset{\mathbf{W}}{\rightsquigarrow} S'$ means $S \overset{\mathbf{W}}{\rightsquigarrow} i$ for all $i \in S'$. Thus,

$$\begin{aligned}
0 \leq p_{S\rightsquigarrow S'}(\mathbf{w}') - p_{S\rightsquigarrow S'}(\mathbf{w}) &= \mathbb{P}\Big[S \overset{\mathbf{W}'}{\rightsquigarrow} S'\Big] - \mathbb{P}\Big[S \overset{\mathbf{W}}{\rightsquigarrow} S'\Big] - \left(\sum_{S'\notin\mathcal{S}} p_{S\rightsquigarrow S'}(\mathbf{w}') - \sum_{S'\notin\mathcal{S}} p_{S\rightsquigarrow S'}(\mathbf{w})\right) \\
&\leq \mathbb{P}\Big[S \overset{\mathbf{W}'}{\rightsquigarrow} S'\Big] - \mathbb{P}\Big[S \overset{\mathbf{W}}{\rightsquigarrow} S'\Big] \\
&= \mathbb{P}\Big[S \overset{\mathbf{W}'}{\rightsquigarrow} S' \text{ but not } S \overset{\mathbf{W}}{\rightsquigarrow} S'\Big] \\
&\leq \mathbb{P}\Big[\text{There is an edge } ij \text{ s.t. } S \overset{\mathbf{W}}{\rightsquigarrow} i,\ W_{ij} = 0, \text{ and } W'_{ij} = 1\Big].
\end{aligned}$$

The last inequality is by noticing that if $S \overset{\mathbf{W}'}{\rightsquigarrow} S'$ but not $S \overset{\mathbf{W}}{\rightsquigarrow} S'$, then there must be a edge ij accessible from S in $G_{\mathbf{W}'}$ such that $W_{ij} = 0$ and $W'_{ij} = 1$. Taking the first such edge ij (watching the contagion spread from S step by step), we see that ij must be accessible from S in $G_{\mathbf{W}}$ as well (since otherwise there's a previous accessible edge $k\ell$ that verifies $W_{k\ell} = 0$ and $W'_{k\ell} = 1$).

We have that $S \overset{\mathbf{W}}{\rightsquigarrow} i$ is independent from W_{ij}, W'_{ij}. Since $\mathbb{P}\big[W_{ij} = 0, \text{ and } W'_{ij} = 1\big] = (1 - w_{ij})\frac{w'_{ij} - w_{ij}}{1 - w_{ij}} = w'_{ij} - w_{ij}$, we have

$$\mathbb{P}\Big[\text{There is an edge } ij \text{ s.t. } S \overset{\mathbf{W}}{\rightsquigarrow} i,\ W_{ij} = 0, \text{ and } W'_{ij} = 1\Big] \leq \sum_{ij\in E} p_i(S;\mathbf{w})\big(w'_{ij} - w_{ij}\big).$$

□

J.2. Main proof of Theorem 4

Proof. We apply a similar analysis as above. When all the events hold, the same analysis gives

$$\Delta(S_t) \leq 2\lambda^\star\alpha \sum_{i\in S_t} 1 \wedge 2\sqrt{\frac{1.5\log(t)}{N_{\ominus i,t-1}}} + 4\text{Bonus}_4(S_t;\overline{\mathbf{w}}_{t-1}),$$

and the first term can be handled in the same way. The second term can be analyzed in the following way: Using Proposition 5 with $\mathcal{S} = \{S' \subset V,\ S' \notin \{S'_0, \dots, S'_k\}\}$, we get

$$\begin{aligned}
4\text{Bonus}_4(S_t;\overline{\mathbf{w}}_{t-1}) &\leq 4\text{Bonus}_4(S_t;\mathbf{w}^\star) + \sum_{k\geq 0}\big(g(S'_{k+1}) - g(S'_k)\big)\frac{1}{|V|}\text{Bonus}_4(S_t;\mathbf{w}^\star) \\
&= \left(4 + \sqrt{\delta(t)\sum_{i\in V}\frac{d_i}{N_{\oplus i,t-1}}}\right)\text{Bonus}_4(S_t;\mathbf{w}^\star) \\
&\leq 5\text{Bonus}_4(S_t;\mathbf{w}^\star),
\end{aligned}$$

where the last inequality uses the event

$$\mathfrak{R}_t \triangleq \{\forall i \in V, N_{\oplus i,t-1} \geq |E|\delta(t)\}.$$

Relying on Theorem 5, we can deal with this last term and obtain a term of order

$$\delta B/c_0^\star \sum_{i\in V} \frac{|V|^2 d_i \log^2(|E|)}{\Delta_{i,\min}}.$$

□

K. Proof of Proposition 4

Proof. There are two possibilities for $S^\star$: either $\mathbb{E}\Big[\mathbf{e}_{S^\star\cup\{0\}}^\intercal \mathbf{c}^\star\Big] < b$, or $\mathbb{E}\Big[\mathbf{e}_{S^\star\cup\{0\}}^\intercal \mathbf{c}^\star\Big] = b$. In the first case, we know that $S^\star$ is not random. Indeed, if it is not the case, then $\frac{\mathbb{E}[\sigma(S^\star)]}{\mathbb{E}\left[\mathbf{e}_{S^\star\cup\{0\}}^\intercal \mathbf{c}^\star\right]}$ is a convex combination of some $\frac{\sigma(S)}{\mathbf{e}_{S\cup\{0\}}^\intercal \mathbf{c}^\star}$ for S in the

support of the distribution of $S^\star$. Necessarily, the maximizer (over S in the support) of the ratio is such that $\mathbf{e}_{S\cup\{0\}}^\intercal \mathbf{c}^\star > b$, since otherwise this maximizer contradicts the definition of $S^\star$. Therefore, increasing the coefficient of this maximizer in the convex combination increases $\mathbb{E}\left[\mathbf{e}_{S^\star\cup\{0\}}^\intercal \mathbf{c}^\star\right]$, which can thus be set to b. Since this also increases $\frac{\mathbb{E}[\sigma(S^\star)]}{\mathbb{E}\left[\mathbf{e}_{S^\star\cup\{0\}}^\intercal \mathbf{c}^\star\right]}$, we get a contradiction since we improved the solution $S^\star$ while still satisfying the constraint.

- Consider the first case. We have as for the proof of Proposition 2, that

$$\left(1-e^{-1}\right)\frac{\sigma(S^\star)}{\mathbf{e}_{S^\star\cup\{0\}}^\intercal \mathbf{c}^\star} \leq \frac{(1-\beta)\sigma(S_\ell)+\beta\sigma(S_{\ell+1})}{(1-\beta)\mathbf{e}_{S_\ell\cup\{0\}}^\intercal \mathbf{c}^\star + \beta\mathbf{e}_{S_{\ell+1}\cup\{0\}}^\intercal \mathbf{c}^\star},$$

  where $\ell \in \{0,1,\dots,|V|-1\}$ is such that $\mathbf{e}_{S_\ell}^\intercal \mathbf{c} \leq \mathbf{e}_{S^\star}^\intercal \mathbf{c} \leq \mathbf{e}_{S_{\ell+1}}^\intercal \mathbf{c}$, and $\mathbf{e}_{S^\star\cup\{0\}}^\intercal \mathbf{c}^\star = (1-\beta)\mathbf{e}_{S_\ell\cup\{0\}}^\intercal \mathbf{c}^\star + \beta\mathbf{e}_{S_{\ell+1}\cup\{0\}}^\intercal \mathbf{c}^\star$. In the case S_ℓ has a greater ratio than $S_{\ell+1}$, it is chosen by our algorithm and has the desired approximation. In the case $S_{\ell+1}$ has the better ratio, it is chosen if its cost is lower than b. If its cost is greater than b, then $\ell+1=j$ and the algorithm chooses S_ℓ with some probability $1-\beta'$ and $S_{\ell+1}$ with probability β'. The goal is to show that the coefficient β' we use for $S_{\ell+1}$ is greater than β. This must be the case since $(1-\beta')\mathbf{e}_{S_\ell\cup\{0\}}^\intercal \mathbf{c}^\star + \beta'\mathbf{e}_{S_{\ell+1}\cup\{0\}}^\intercal \mathbf{c}^\star = b > \mathbf{e}_{S^\star\cup\{0\}}^\intercal \mathbf{c}^\star = (1-\beta)\mathbf{e}_{S_\ell\cup\{0\}}^\intercal \mathbf{c}^\star + \beta\mathbf{e}_{S_{\ell+1}\cup\{0\}}^\intercal \mathbf{c}^\star$.

- For the second case, we let S be the output of the Algorithm 1 considered by Wang et al. (2020). We thus have from their Theorem 1 that $\left(1-e^{-1}\right)\mathbb{E}[\sigma(S^\star)] \leq \mathbb{E}[\sigma(S)]$. Since $\mathbb{E}\left[\mathbf{e}_{S\cup\{0\}}^\intercal \mathbf{c}^\star\right] = \mathbb{E}\left[\mathbf{e}_{S^\star\cup\{0\}}^\intercal \mathbf{c}^\star\right] = b$, we have

$$\left(1-e^{-1}\right)\frac{\mathbb{E}[\sigma(S^\star)]}{\mathbb{E}\left[\mathbf{e}_{S^\star\cup\{0\}}^\intercal \mathbf{c}^\star\right]} \leq \frac{\mathbb{E}[\sigma(S)]}{\mathbb{E}\left[\mathbf{e}_{S\cup\{0\}}^\intercal \mathbf{c}^\star\right]}.$$

  If the expected cost of the output S' of our algorithm is b, then both algorithms coincides and we have the desired result. Else, we have that S' maximizes the ratio over $\{S_0,\dots,S_j\}$, which contains the support of S (that is $\{S_{j-1},S_j\}$), so the ratio evaluated at S' is greater than $\frac{\mathbb{E}[\sigma(S)]}{\mathbb{E}\left[\mathbf{e}_{S\cup\{0\}}^\intercal \mathbf{c}^\star\right]}$, giving again the desired result.

□

L. Generalities on combinatorial multi-armed bandits

In this section, i represent an "arm", i.e., an edge in our OIM context. A_t is the random set of edges that are triggered at round t. Here, the horizon T can be random. Finally, $b_i(S_t)$ is simply some non-negative function (for our Bonus$_4$, this is $|V|$ times the square root of the out-degree) and m_i is the maximum number of edges that can be reached when i is activated. The following theorem is based on Perrault et al. (2020b), Theorem 4.

Theorem 5 (Regret bound for ℓ_2-bonus, with expectation outside the norm)**.** *For all $i\in[n]$, let $(\alpha_i,\beta_{i,T})\in[1/2,1)\times\mathbb{R}_+$. For $t\geq 1$, consider the event*

$$\mathfrak{A}_t \triangleq \left\{\Delta_t \leq \mathbb{E}\left[\left\|\sum_{i\in A_t, N_{i,t-1}>0}\frac{b_i(S_t)\beta_{i,T}^{\alpha_i}\mathbf{e}_i}{N_{i,t-1}^{\alpha_i}}\right\|_2\middle|\mathcal{F}_t\right]\right\},$$

Then, if $\{t\leq T\}\in\mathcal{F}_t$, we have

$$\mathbb{E}\left[\sum_{t=1}^{T}\mathbb{I}\{\mathfrak{A}_t\}\Delta_t\right] \leq \sum_{i\in[n]} 4\log_2(4\sqrt{m_i})\max_{S\in\mathcal{S},\ p_i(S)>0} b_i(S)^{\frac{1}{\alpha_i}}\mathbb{E}[\beta_{i,T}]\eta_i,$$

where

$$\eta_i = \begin{cases} 32\log_2(4\sqrt{m_i})\Delta_{i,\min}^{-1} & \text{if } \alpha_i = 1/2 \\ 2^{\frac{2}{\alpha_i}}\left(\left(1-2^{\frac{1}{\alpha_i}-2}\right)(1-\alpha_i)\Delta_{i,\min}^{\frac{1-\alpha_i}{\alpha_i}}\right)^{-1} & \text{if } 1/2<\alpha_i<1. \end{cases}$$

Proof. Let $t \geq 1$. With a first reverse amortisation, we start by restricting the set of possibles for A_t by only taking those whose error is at least twice as large as Δ_t: assuming that $\mathfrak{A}_t$ holds, we have

$$\begin{aligned}\Delta_t &\leq \mathbb{E}\left[2\left\|\sum_{i\in A_t, N_{i,t-1}>0} \frac{b_i(S_t)\beta_{i,T}^{\alpha_i}\mathbf{e}_i}{N_{i,t-1}^{\alpha_i}}\right\|_2 - \Delta_t \middle| \mathcal{F}_t\right] \\ &\leq \mathbb{E}\left[\mathbb{I}\left\{\left\|\sum_{i\in A_t, N_{i,t-1}>0} \frac{2b_i(S_t)\beta_{i,T}^{\alpha_i}\mathbf{e}_i}{N_{i,t-1}^{\alpha_i}}\right\|_2 \geq \Delta_t\right\}\left\|\sum_{i\in A_t, N_{i,t-1}>0} \frac{2b_i(S_t)\beta_{i,T}^{\alpha_i}\mathbf{e}_i}{N_{i,t-1}^{\alpha_i}}\right\|_2 \middle| \mathcal{F}_t\right]\end{aligned}$$

We now define

$$\Lambda(A_t) \triangleq \left\|\sum_{i\in A_t, N_{i,t-1}>0} \frac{2b_i(S_t)\beta_{i,T}^{\alpha_i}\mathbf{e}_i}{N_{i,t-1}^{\alpha_i}}\right\|_2,$$

and have for any $j \in A_t$ that

$$\Lambda(A_t) \geq \frac{2b_j(S_t)\beta_{j,T}^{\alpha_j}}{N_{j,t}^{\alpha_j}}. \tag{23}$$

Then, we can write:

$$\begin{aligned}\Lambda(A_t) &= -\Lambda(A_t) + \left\|\sum_{i\in A_t, N_{i,t-1}>0} \frac{4b_i(S_t)\beta_{i,T}^{\alpha_i}\mathbf{e}_i}{N_{i,t-1}^{\alpha_i}}\right\|_2 \\ &= -\left\|\sum_{i\in A_t} \frac{\Lambda(A_t)\mathbf{e}_i}{\|\mathbf{e}_{A_t}\|_2}\right\|_2 + \left\|\sum_{i\in A_t, N_{i,t-1}>0} \frac{4b_i(S_t)\beta_{i,T}^{\alpha_i}\mathbf{e}_i}{N_{i,t-1}^{\alpha_i}}\right\|_2 \\ &\leq \left\|\sum_{i\in A_t, N_{i,t-1}>0}\left(\frac{4b_i(S_t)\beta_{i,T}^{\alpha_i}}{N_{i,t-1}^{\alpha_i}} - \frac{\Lambda(A_t)}{\|\mathbf{e}_{A_t}\|_2}\right)^+ \mathbf{e}_i\right\|_2 \\ &= \left\|\sum_{i\in A_t, N_{i,t-1}>0}\left(\frac{4b_i(S_t)\beta_{i,T}^{\alpha_i}}{N_{i,t-1}^{\alpha_i}} - \frac{\Lambda(A_t)}{\|\mathbf{e}_{A_t}\|_2}\right)^+ \mathbb{I}\left\{\Lambda(A_t) \geq \frac{2b_i(S_t)\beta_{i,T}^{\alpha_i}}{N_{i,t-1}^{\alpha_i}}\right\}\mathbf{e}_i\right\|_2 \qquad \text{Using (23)} \\ &\leq \left\|\sum_{i\in A_t, N_{i,t-1}>0} \mathbb{I}\left\{2\Lambda(A_t) \geq \frac{4b_i(S_t)\beta_{i,T}^{\alpha_i}}{N_{i,t-1}^{\alpha_i}} \geq \frac{\Lambda(A_t)}{\|\mathbf{e}_{A_t}\|_2}\right\}\frac{4b_i(S_t)\beta_{i,T}^{\alpha_i}\mathbf{e}_i}{N_{i,t-1}^{\alpha_i}}\right\|_2.\end{aligned}$$

We now consider the following partition of the set of indices:

$$\mathbb{I}\left\{i \in A_t, N_{i,t-1} > 0,\ 2\Lambda(A_t) \geq \frac{4b_i(S_t)\beta_{i,T}^{\alpha_i}}{N_{i,t-1}^{\alpha_i}} \geq \frac{\Lambda(A_t)}{\|\mathbf{e}_{A_t}\|_2}\right\} \subset \bigcup_{k=0}^{\lceil \log_2(\|\mathbf{e}_{A_t}\|_2)\rceil} J_{k,t},$$

where for all integer $1 \leq k \leq \lceil \log_2(\|\mathbf{e}_{A_t}\|_2)\rceil$,

$$J_{k,t} \triangleq \left\{i \in A_t, N_{i,t-1} > 0,\ 2^{1-k}\Lambda(A_t) \geq \frac{4b_i(S_t)\beta_{i,T}^{\alpha_i}}{N_{i,t-1}^{\alpha_i}} \geq 2^{-k}\Lambda(A_t)\right\}.$$

We bound $\Lambda(A_t)^2$ as

$$\begin{aligned}\Lambda(A_t)^2 &\le \left\| \sum_{i\in A_t, N_{i,t-1}>0} \mathbb{I}\left\{2\Lambda(A_t) \ge \frac{4b_i(S_t)\beta_{i,T}^{\alpha_i}}{N_{i,t-1}^{\alpha_i}} \ge \frac{\Lambda(A_t)}{\|\mathbf{e}_{A_t}\|_2}\right\} \frac{4b_i(S_t)\beta_{i,T}^{\alpha_i}\mathbf{e}_i}{N_{i,t-1}^{\alpha_i}} \right\|_2^2 \\ &= \sum_{k=0}^{\lceil \log_2(\|\mathbf{e}_{A_t}\|_2)\rceil} \left\| \sum_{i\in J_{k,t}} \frac{4b_i(S_t)\beta_{i,T}^{\alpha_i}\mathbf{e}_i}{N_{i,t-1}^{\alpha_i}} \right\|_2^2 \\ &\le \sum_{k=0}^{\lceil \log_2(\|\mathbf{e}_{A_t}\|_2)\rceil} 2^{2-2k}\Lambda(A_t)^2 \left\|\mathbf{e}_{J_{k,t}}\right\|_2^2.\end{aligned}$$

So there exists one integer k_t such that $|J_{k_t,t}| = \left\|\mathbf{e}_{J_{k_t,t}}\right\|_2^2 \ge 2^{2k_t-2}(1+\lceil\log_2(\|\mathbf{e}_{A_t}\|_2)\rceil)^{-1}$.

$$\begin{aligned}\sum_{t=1}^T \mathbb{I}\{\mathfrak{A}_t\}\Delta_t &\le \sum_{t=1}^T \mathbb{E}[\mathbb{I}\{\Lambda(A_t)\ge\Delta_t\}\Lambda(A_t)|\mathcal{F}_t] \\ &\le \sum_{t=1}^T \mathbb{E}\left[\sum_{k=0}^{\lceil \log_2(\|\mathbf{e}_{A_t}\|_2)\rceil} \mathbb{I}\{k_t=k,\ \Lambda(A_t)\ge\Delta_t\}\Lambda(A_t)\middle|\mathcal{F}_t\right] \\ &\le \sum_{t=1}^T \mathbb{E}\left[\sum_{k=0}^{\lceil \log_2(\|\mathbf{e}_{A_t}\|_2)\rceil} \mathbb{I}\{k_t=k,\ \Lambda(A_t)\ge\Delta_t\}\frac{\sum_{i\in[n]}\mathbb{I}\{i\in J_{k,t}\}}{2^{2k-2}(1+\lceil\log_2(\|\mathbf{e}_{A_t}\|_2)\rceil)^{-1}}\Lambda(A_t)\middle|\mathcal{F}_t\right] \\ &\le \sum_{t=1}^T\sum_{i=1}^n \mathbb{E}\left[\sum_{k=0}^{\lceil \log_2(\|\mathbf{e}_{A_t}\|_2)\rceil} \frac{\mathbb{I}\left\{i\in A_t,\ 0<N_{i,t-1}^{\alpha_i}\le\frac{2^{k+2}b_i(S_t)\beta_{i,T}^{\alpha_i}}{\Lambda(A_t)},\ \Lambda(A_t)\ge\Delta_t\right\}}{2^{2k-2}(1+\lceil\log_2(\|\mathbf{e}_{A_t}\|_2)\rceil)^{-1}}\Lambda(A_t)\middle|\mathcal{F}_t\right].\end{aligned}$$

Taking the expectation of the above, and using $\{t\le T\}\in\mathcal{F}_t$, we have the bound

$$\begin{aligned}\mathbb{E}\left[\sum_{t=1}^T \mathbb{I}\{\mathfrak{A}_t\}\Delta_t\right] &\le \sum_{i=1}^n \mathbb{E}\left[\sum_{t=1}^T\sum_{k=0}^{\lceil \log_2(\|\mathbf{e}_{A_t}\|_2)\rceil} \frac{\mathbb{I}\left\{i\in A_t,\ 0<N_{i,t-1}^{\alpha_i}\le\frac{2^{k+2}b_i(S_t)\beta_{i,T}^{\alpha_i}}{\Lambda(A_t)},\ \Lambda(A_t)\ge\Delta_t\right\}}{2^{2k-2}(1+\lceil\log_2(\|\mathbf{e}_{A_t}\|_2)\rceil)^{-1}}\Lambda(A_t)\right] \\ &\le \sum_{i=1}^n \sum_{k=0}^{\lceil\log_2(\sqrt{m_i})\rceil} \mathbb{E}\left[\frac{1+\lceil\log_2(\sqrt{m_i})\rceil}{2^{2k-2}} \underbrace{\sum_{t=1}^T \mathbb{I}\left\{i\in A_t,\ 0<N_{i,t-1}^{\alpha_i}\le\frac{2^{k+2}b_i(S_t)\beta_{i,T}^{\alpha_i}}{\Lambda(A_t)},\ \Lambda(A_t)\ge\Delta_t\right\}\Lambda(A_t)}_{(24)_{i,k}}\right].\end{aligned}$$

Applying Proposition 6 gives

$$(24)_{i,k} \le \frac{\max_{S\in\mathcal{S},\ p_i(S)>0} b_i(S)^{\frac{1}{\alpha_i}}\beta_{i,T}2^{\frac{k+2}{\alpha_i}}}{1-\alpha_i}\Delta_{i,\min}^{1-1/\alpha_i},$$

So using $\lceil\log_2(\sqrt{m_i})\rceil + 1 \le \log_2(4\sqrt{m_i})$, we get

$$\mathbb{E}\left[\sum_{t=1}^T \mathbb{I}\{\mathfrak{A}_t\}\Delta_t\right] \le \sum_{i\in[n]} 4\log_2(4\sqrt{m_i}) \max_{S\in\mathcal{S},\ p_i(S)>0} b_i(S)^{\frac{1}{\alpha_i}}\mathbb{E}[\beta_{i,T}]\eta_i,$$

where

$$\eta_i = \begin{cases} 32\log_2(4\sqrt{m_i})\Delta_{i,\min}^{-1} & \text{if } \alpha_i = 1/2 \\ 2^{\frac{2}{\alpha_i}}\left(\left(1-2^{\frac{1}{\alpha_i}-2}\right)(1-\alpha_i)\Delta_{i,\min}^{\frac{1-\alpha_i}{\alpha_i}}\right)^{-1} & \text{if } 1/2 < \alpha_i < 1. \end{cases}$$

□

Proposition 6. *Let $i \in [n]$ and $f_i : \mathbb{R}_+ \to \mathbb{R}_+$ be a non increasing function, integrable on an interval $[\delta_{i,\min}, \delta_{i,\max}] \subset \mathbb{R}_+^\star$. Then for any sequence of real numbers $(\delta_t) \in ([\delta_{i,\min}, \delta_{i,\max}] \cup \{0\})^T$,*

$$\sum_{t=1}^{T} \mathbb{I}\{i \in A_t,\ 1 \le N_{i,t-1} \le f_i(\delta_t)\}\delta_t \le f_i(\delta_{i,\min})\delta_{i,\min} + \int_{\delta_{i,\min}}^{\delta_{i,\max}} f_i(x)\mathrm{d}x.$$

In particular,

- *If $f_i(x) = \beta_{i,T}x^{-1/\alpha_i}$, $\alpha_i \in (0,1)$ and $\beta_{i,T} \ge 0$, then*

$$\begin{aligned}\sum_{t=1}^{T} \mathbb{I}\{i \in A_t,\ 1 \le N_{i,t-1} \le f_i(\delta_t)\}\delta_t &\le \delta_{i,\min}^{1-1/\alpha_i}\frac{\beta_{i,T}}{1-\alpha_i} - \delta_{i,\max}^{1-1/\alpha_i}\frac{\alpha_i\beta_{i,T}}{1-\alpha_i} \\ &\le \delta_{i,\min}^{1-1/\alpha_i}\frac{\beta_{i,T}}{1-\alpha_i}.\end{aligned}$$

- *If $f_i(x) = \beta_{i,T}x^{-1}$, $\beta_{i,T} \ge 0$, then*

$$\sum_{t=1}^{T} \mathbb{I}\{i \in A_t,\ 1 \le N_{i,t-1} \le f_i(\delta_t)\}\delta_t \le \beta_{i,T}\left(1 + \log\left(\frac{\delta_{i,\max}}{\delta_{i,\min}}\right)\right).$$

Proof. Consider $\delta_{i,\max} = \delta_{i,1} \ge \delta_{i,2} \ge \cdots \ge \delta_{i,K_i} = \delta_{i,\min}$ being all possible values for δ_t when $\delta_t \ne 0$. We define a dummy gap $\delta_{i,0} = \infty$ and let $f_i(\delta_{i,0}) = 0$. In (25), we look at times t where $\delta_t \ne 0$ and first break the range $(0, f_i(\delta_t)]$ of the counter $N_{i,t-1}$ into sub intervals:

$$(0, f_i(\delta_t)] = (f_i(\delta_{i,0}), f_i(\delta_{i,1})] \cup \cdots \cup (f_i(\delta_{i,k_t-1}), f_i(\delta_{i,k_t})],$$

where k_t is the index such that $\delta_{i,k_t} = \delta_t$. This index k_t exists by assumption that the subdivision contains all possible values for δ_t when $\delta_t \ne 0$. Notice that in (25), we do not explicitly use k_t, but instead sum over all $k \in [K_i]$ and filter against the event $\{\delta_{i,k} \ge \delta_t\}$, which is equivalent to summing over $k \in [k_t]$.

$$\begin{aligned}&\sum_{t=1}^{T} \mathbb{I}\{i \in A_t,\ N_{i,t-1} \le f_i(\delta_t)\}\delta_t \\ &= \sum_{t=1}^{T}\sum_{k=1}^{K_i} \mathbb{I}\{i \in A_t,\ f_i(\delta_{i,k-1}) < N_{i,t-1} \le f_i(\delta_{i,k}), \delta_{i,k} \ge \delta_t\}\delta_t. \end{aligned} \tag{25}$$

Over each event that $N_{i,t-1}$ belongs to the interval $(f_i(\delta_{i,k-1}), f_i(\delta_{i,k})]$, we upper bound the gap δ_t by $\delta_{i,k}$.

$$(25) \le \sum_{t=1}^{T}\sum_{k=1}^{K_i} \mathbb{I}\{i \in A_t,\ f_i(\delta_{i,k-1}) < N_{i,t-1} \le f_i(\delta_{i,k}), \delta_{i,k} \ge \delta_t\}\delta_{i,k}. \tag{26}$$

Then, we further upper bound the summation by adding events that $N_{i,t-1}$ belongs to the remaining intervals $(f_i(\delta_{i,k-1}), f_i(\delta_{i,k})]$ for $k_t < k \leq K_i$, associating them to a suffered gap $\delta_{i,k}$. This is equivalent to removing the filtering against the event $\{\delta_{i,k} \geq \delta_t\}$.

$$(26) \leq \sum_{t=1}^{T} \sum_{k=1}^{K_i} \mathbb{I}\{i \in A_t,\ f_i(\delta_{i,k-1}) < N_{i,t-1} \leq f_i(\delta_{i,k})\}\delta_{i,k}. \tag{27}$$

Now, we invert the summation over t and the one over k.

$$(27) = \sum_{k=1}^{K_i} \sum_{t=1}^{T} \mathbb{I}\{i \in A_t,\ f_i(\delta_{i,k-1}) < N_{i,t-1} \leq f_i(\delta_{i,k})\}\delta_{i,k}. \tag{28}$$

For each $k \in [K_i]$, the number of times $t \in [T]$ that the counter $N_{i,t-1}$ belongs to $(f_i(\delta_{i,k-1}), f_i(\delta_{i,k})]$ can be upper bounded by the number of integers in this interval. This is due to the event $\{i \in A_t\}$, imposing that $N_{i,t-1}$ is incremented, so $N_{i,t-1}$ cannot be worth the same integer for two different times t satisfying $i \in A_t$. We use the fact that for all $x, y \in \mathbb{R}$, $x \leq y$, the number of integers in the interval $(x, y]$ is exactly $\lfloor y \rfloor - \lfloor x \rfloor$.

$$(28) \leq \sum_{k=1}^{K_i} (\lfloor f_i(\delta_{i,k}) \rfloor - \lfloor f_i(\delta_{i,k-1}) \rfloor)\delta_{i,k}. \tag{29}$$

We then simply expand the summation, and some terms are cancelled (remember that $f_i(\delta_{i,0}) = 0$).

$$(29) = \lfloor f_i(\delta_{i,K_i}) \rfloor \delta_{i,K_i} + \sum_{k=1}^{K_i - 1} \lfloor f_i(\delta_{i,k}) \rfloor (\delta_{i,k} - \delta_{i,k+1}) \tag{30}$$

We use $\lfloor x \rfloor \leq x$ for all $x \in \mathbb{R}$. Finally, we recognize a right Riemann sum, and use the fact that f_i is non increasing to upper bound each $f_i(\delta_{i,k})(\delta_{i,k} - \delta_{i,k+1})$ by $\int_{\delta_{i,k+1}}^{\delta_{i,k}} f_i(x)\mathrm{d}x$, for all $k \in [K_i - 1]$.

$$(30) \leq f_i(\delta_{i,K_i})\delta_{i,K_i} + \sum_{k=1}^{K_i - 1} f_i(\delta_{i,k})(\delta_{i,k} - \delta_{i,k+1}) \tag{31}$$

$$\leq f_i(\delta_{i,K_i})\delta_{i,K_i} + \int_{\delta_{i,K_i}}^{\delta_{i,1}} f_i(x)\mathrm{d}x. \tag{32}$$

□