The Network Visibility Problem

KHASHAYAR GATMIRY, Massachusetts Institute of Technology, USA
MANUEL GOMEZ-RODRIGUEZ, Max Planck Institute for Software Systems, Germany

Social media is an attention economy where broadcasters are constantly competing for attention in their followers' feeds. Broadcasters are likely to elicit greater attention from their followers if their posts remain visible at the top of their followers' feeds for a longer period of time. However, this depends on the rate at which their followers receive information in their feeds, which in turn depends on the broadcasters they follow. Motivated by this observation and recent calls for fairness of exposure in social networks, in this article, we look at the task of recommending links from the perspective of visibility optimization. Given a set of candidate links provided by a link recommendation algorithm, our goal is to find a subset of those links that would provide the highest visibility to a set of broadcasters. To this end, we first show that this problem reduces to maximizing a nonsubmodular nondecreasing set function under matroid constraints. Then, we show that the set function satisfies a notion of approximate submodularity that allows the standard greedy algorithm to enjoy theoretical guarantees. Experiments on both synthetic and real data gathered from Twitter show that the greedy algorithm is able to consistently outperform several competitive baselines.

CCS Concepts: • **Computing methodologies** → **Machine learning**; • **Information systems** → **Social networks**;

Additional Key Words and Phrases: Link recommendation, network optimization, fairness of exposure, visibility optimization, submodularity

1 INTRODUCTION

Social media users are eager to gain new followers—to grow their audience—so, whenever they decide to share a new story, it receives a greater amount of views, likes, and shares.At the same time, users actually share quite a portion of their followers and, as a consequence, they are constantly competing with each other for attention [3, 15], which becomes a scarce commodity of great value [11]. In this context, recent empirical studies have shown that stories at the top of a user's feed are more likely to be noticed and consequently liked or shared [18, 19, 27].

The above empirical findings have motivated the recently introduced when-to-post problem [20, 32, 34, 38, 39], which aims to help a user, a *broadcaster*, find the best times to share stories with her

Authors' addresses: K. Gatmiry, Massachusetts Institute of Technology, Dept. of EECS and Computer Science, Cambridge, MA 02139, USA; email: gatmiry@mit.edu; M. Gomez-Rodriguez, Max Planck Institute for Software Systems, Paul-Ehrlich-Strasse, 26, Kaiserslautern, Germany; email: manuelgr@mpi-sws.org.

followers—the times when her stories would enjoy higher visibility and would consequently elicit greater attention from her audience. While this line of work has shown great promise at helping broadcasters increase their visibility, it assumes the links between broadcasters and followers are given. However, these links are of great importance to the broadcasters' visibility—they define their audience. Motivated by this observation and recent calls for fairness of exposure in social networks [5, 30], in this work, we look at the task of recommending links in social networks from the perspective of visibility optimization.

Link recommendation has a rich history in the recommender systems literature [28]. However, link recommendation algorithms have traditionally focused on maximizing the followers' utility— they recommend users to follow broadcasters whose posts may be found more interesting. This uncompromising focus on the utility of the followers, rather than the utility of the broadcasters, has been called into question as social media platforms are increasingly used as news sources.[1,2] In this work, given a set of candidate links provided by a link recommendation algorithm, our goal is to find a subset of those links that maximize the broadcasters' utility, as measured by the visibility their posts would achieve through that subset of links. Here, note that we aim to complement, rather than substitute, traditional link recommendation algorithms maximizing the followers' utility.

Our work. We approach the above problem from the perspective of temporal point processes and nonsubmodular set function maximization, where we measure the visibility a broadcaster achieves with respect to her followers as the number of stories posted by her that lie within the top k positions in her followers' feeds over time. A desirable property of this visibility measure, which is also shared by similar visibility measures used in recent work [20, 34, 38, 39], is that it can be easily extracted from real data without interventions—given any posting strategy for a broadcaster, one can always measure the visibility she would achieve with respect to any set of followers using a separate held-out set of the followers' feeds.

More specifically, we represent users' posts and feeds using the temporal point processes, which characterize the continuous time interval between posts using intensity functions [1, 31]. Given a set of broadcasters, we derive a formula that links the visibility they achieve with respect to their followers to their posting intensities and the followers' feed intensities due to other broadcasters. Based on this formula, we make the following contributions:

(i) We show that the visibility satisfies a notion of approximate submodularity, $\breve{\alpha}$-submodularity [40]. This notion is characterized by a key quantity, the inverse generalized curvature $\breve{\alpha}$ [6, 17, 26], which we are able to bound in our problem.

(ii) We show that the standard greedy algorithm [29] achieves a constant approximation factor of $(1 + 1/(1 - \breve{\alpha}))^{-1}$ at maximizing visibility under constraints on the number of edges per broadcaster. More broadly, these guarantees hold for maximizing any nondecreasing $\breve{\alpha}$-submodular function under matroid constraints.

(iii) We analyze the sample complexity of an empirical estimate of the visibility and show that, given a sufficient number of samples, the greedy algorithm is guaranteed to find a set of edges such that the broadcasters' visibility is at least $0.5\,OPT/(1 + 1/(1 - \breve{\alpha}))$, where OPT is the optimal value.

We believe the above contributions have implications beyond the specific problem we study here because, to the best of our knowledge, the maximization of $\breve{\alpha}$-submodular functions under general matroid constraints has not been studied. Finally, we experiment with both synthetic and real

---

[1]https://www.nytimes.com/2018/09/05/technology/lawmakers-facebook-twitter-foreign-influence-hearing.html.
[2]https://www.economist.com/business/2018/09/06/how-social-media-platforms-dispense-justice.

data gathered from Twitter and show that the greedy algorithm is able to consistently outperform several nontrivial baselines.

Further related work. In addition to previous work on the when-to-post problem, link recommendation, and $\check{\alpha}$-submodularity, our work also relates to set function maximization, network manipulation, and fair ranking.

The problem of maximizing a nondecreasing set function emerges in a wide variety of important real-world applications such as feature selection, sparse modeling, and experimental design, to name a few. If the set function of interest is nondecreasing and satisfies a natural diminishing property called submodularity,[3] then the problem is well understood. For example, under a simple cardinality constraint, it is known that the *standard* greedy algorithm enjoys an approximation factor of $(1-1/e)$ [29, 35]. Moreover, this constant factor has been improved using the curvature [10, 36] of a submodular function. Under a general matroid constraint, a variation of the standard greedy algorithm yields a 1/2-approximation [14] and, more recently, it has been shown that there exist polynomial time algorithms that yield a $(1 - 1/e)$-approximation [7, 13].

However, there are many important applications where the corresponding set function is not submodular. In this context, Bian et al. [4] have shown that, under a cardinality constraint, the standard greedy algorithm enjoys an approximation factor of $\frac{1}{\alpha}(1 - e^{-\gamma\alpha})$, where γ is the submodularity ratio [12] of the set function and α is the curvature of the set function. Very recently, Harshaw et al. [16] have also shown that there is no polynomial algorithm with better guarantees. However, the problem of maximizing a non-submodular nondecreasing set function subject to a general matroid constraint has only been studied very recently by Chen et al. [9], who have shown that a randomized version of the standard greedy algorithm enjoys an approximation factor of $(1+1/\gamma)^{-2}$. In our work, we further advance the state-of-the-art on non-submodular nondecreasing set function by studying the maximization of $\check{\alpha}$-submodular functions under general matroid constraints.

In network manipulation [22–24, 33], the goal is to find a set of edges in a social network whose addition (removal) can maximize (minimize) the spread of a contagion (e.g., a story, a disease). In contrast, our goal is to find a set of edges whose addition can maximize social media users' visibility. Moreover, our algorithm enjoys theoretical guarantees while most algorithms for network manipulation do not, except for one notable exception [21]. Finally, in fair ranking [2, 5, 30, 37], the goal is to provide rankings under fairness constraints in terms of exposure allocation of the information sources. In our work, we do not focus on providing rankings but recommending links between the information sources, i.e., the broadcasters, and the end-users, i.e., the followers.

2 PROBLEM FORMULATION

In this section, we first revisit how to use the theory of temporal point processes [1, 31] to represent broadcasters and feeds in social and information networks [20, 38, 39]. Then, we define our visibility measure and derive a relationship between this visibility measure and the intensity functions of broadcasters and followers. Finally, we conclude with a statement of the network visibility problem.

Representation of broadcasters and feeds. Given a directed network $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, we assume any user $u \in \mathcal{V}$ can be a broadcaster, a follower, or both, each broadcaster can be followed by multiple followers, and each follower can follow multiple broadcasters. Then, we represent the broadcasting times of the users as a collection of counting processes denoted by a vector $\mathbf{N}(t)$, in

---

[3]A set function $F(\cdot)$ is submodular iff it satisfies that $F(\mathcal{A} \cup \{v\}) - F(\mathcal{A}) \geq F(\mathcal{B} \cup \{v\}) - F(\mathcal{B})$ for all $\mathcal{A} \subseteq \mathcal{B} \subset \mathcal{V}$ and $v \in \mathcal{V}$, where $\mathcal{V}$ is the ground set.

which the ith dimension is the number of messages or *stories* broadcasted by user i up to time t, and characterize these counting processes using their corresponding intensities, i.e., $\mathbb{E}[d\boldsymbol{N}(t)] = \boldsymbol{\mu}(t)\,dt$. Moreover, given the adjacency matrix $\boldsymbol{A} \in \{0, 1\}^{n \times n}$, where $A_{ij} = 1$ indicates that user j follows user i, we can represent the times of the stories users receive in their feeds from the broadcasters they follow as a sum of counting processes, $\boldsymbol{A}^T\boldsymbol{N}(t)$, and calculate the corresponding intensities as $\boldsymbol{\gamma}(t) = \boldsymbol{A}^T\boldsymbol{\mu}(t)$.

Finally, from the perspective of set of broadcasters $\mathcal{B}$, it is useful to define $\mathcal{E}_\mathcal{B}$ as the set of their outgoing edges, i.e., who follows them, and the counting processes $\boldsymbol{M}_{\backslash\mathcal{B}}(t) = \boldsymbol{A}^T\boldsymbol{N}(t) - \boldsymbol{A}_\mathcal{B}^T\boldsymbol{N}(t)$, in which the jth dimension, $M_{j\backslash\mathcal{B}}(t)$, represents the times of the stories user j receives due to other broadcasters she follows and $\boldsymbol{A}_\mathcal{B}$ is a matrix composed of the rows of $\boldsymbol{A}$ corresponding to the set of broadcasters $\mathcal{B}$. For each of these counting processes, the intensity is given by $\gamma_{j\backslash\mathcal{B}}(t) = \gamma_j(t) - A_\mathcal{B}^T\boldsymbol{\mu}(t)$.

Definition of visibility. Given a broadcaster i and one of her followers j, we keep track of the number of stories $r_K(t, i, j)$ posted by i that are among the top K positions of j's feed at time t, which clearly depends on the feed ranking mechanism in the corresponding social network. Here, for simplicity, we assume each user's feed ranks stories in inverse chronological order,[4] as in previous work [20, 38, 39]. Then, given an observation time window $[t_0, t_f]$ and a deterministic sequence of broadcasting events, we can define the deterministic top K visibility of broadcaster i with respect to follower j as

$$\mathcal{T}_K(i, j) := \int_{t_0}^{t_f} r_K(t, i, j)dt, \tag{1}$$

which is the number of stories posted by i's that are among the top K positions of j's feed over time. However, since the sequence of broadcasting events are generated from stochastic processes, we will consider the expected value of the top K visibility instead, i.e.,

$$\mathcal{U}_K(i, j) = \mathbb{E}\left[\mathcal{T}_K(i, j)\right] = \int_{t_0}^{t_f} \mathbb{E}\left[r_K(t, i, j)\right] dt = \int_{t_0}^{t_f} \sum_{k=1}^{K} g_k(t, i, j)dt, \tag{2}$$

where $g_k(t, i, j)$ is the probability that a story posted by broadcaster i is at position k of follower j's feed at time t. Finally, given a set of broadcasters $\mathcal{B}$, define their average top K visibility with respect to their followers as

$$\mathcal{U}_K(\mathcal{E}_\mathcal{B}) = \sum_{(i,j)\in\mathcal{E}_\mathcal{B}} \mathcal{U}_K(i, j) := \sum_{j\in\mathcal{V}} \mathcal{U}_K(\mathcal{E}_\mathcal{B}, j), \tag{3}$$

where, with an overload of notation $\mathcal{U}_K(\mathcal{E}_\mathcal{B}, j) = \sum_{i\in\mathcal{B}\,:\,(i,j)\in\mathcal{E}_\mathcal{B}} \mathcal{U}_K(i, j)$ and the argument $\mathcal{E}_\mathcal{B}$ reminds that the visibility is a function of outgoing edges corresponding to the set of broadcasters $\mathcal{B}$, i.e., it is a function of who follows them. Here, note that, by using the linearity of expectation, we can also write $\mathcal{U}_K(\mathcal{E}_\mathcal{B}, j)$ in terms of the number of stories $r_K(t, \mathcal{E}_\mathcal{B}, j)$ posted by the broadcasters that are among the top K positions of user j's feed at time t and the probability $g_k(t, \mathcal{E}_\mathcal{B}, j)$ that a story posted by the broadcasters is at position k of user j's feed at time t, i.e.,

$$\mathcal{U}_K(\mathcal{E}_\mathcal{B}, j) = \int_{t_0}^{t_f} \mathbb{E}\left[r_K(t, \mathcal{E}_\mathcal{B}, j)\right] = \int_{t_0}^{t_f} \sum_{k=1}^{K} g_k(t, \mathcal{E}_\mathcal{B}, j), \tag{4}$$

where we have again overloaded the notation for simplicity.

---

[4]At the time of writing, Twitter, Facebook, and Weibo allow choosing such an ordering.

Computation of visibility. In this section, we aim to find an expression for the average top K visibility of a set of broadcasters $\mathcal{B}$ with respect to their followers, given by Equation (3), in terms of the intensity functions characterizing the broadcasters and the feeds.

To this end, we first compute the probability $g_1(t) = g_1(t, i, j)$ that one story from a broadcaster i with $\mu_i(t) = \mu(t)$ is at the top of a follower j's feed with $\gamma_{j \setminus i} = \gamma(t)$ at time t. More specifically, for $g_1(t)$, we can easily realize that our definition of visibility matches the one used by Karimi et al. [20].[5] Then, following their same line of reasoning, we have that $g_1(t)$ satisfies the following equation:

$$g_1(t + dt) = g_1(t)(1 - \gamma(t)dt) + (1 - g_1(t))\mu(t)dt. \tag{5}$$

Then, by rearranging the terms and letting $dt \to 0$, one finds that the probability satisfies the following differential equation:

$$g_1'(t) = -g_1(t)(\mu(t) + \gamma(t)) + \mu(t).$$

Now, we can proceed with the induction step for $g_k(t) = g_k(t, i, j)$ with $k > 1$. Here, note that our definition of visibility differs from the one used by Karimi et al. for $k > 1$ and thus we cannot directly follow their reasoning. However, by definition, we have that $g_k(t)$ satisfies the following equation:

$$g_k(t + dt) = g_{k-1}(t)(\mu(t) + \gamma(t))dt + g_k(t)(1 - \mu(t) - \gamma(t)dt),$$

where each term relates to one of the two possible situations: (i) the story at position $k - 1$ of the follower's feed was posted by broadcaster i (*w.p.* $g_{k-1}(t)$) and a broadcaster posts a story in $[t, t+dt]$ (*w.p.* $(\mu(t) + \gamma(t))$); (ii) the story at position k of the follower's feed was posted by broadcaster i (*w.p.* $g_k(t)$) and nobody posts a story in $[t, t + dt]$ (*w.p.* $(1 - \mu(t) - \gamma(t))$). Again, by rearranging terms and letting $dt \to 0$, it follows that:

$$g_k'(t) = (g_{k-1}(t) - g_k(t))(\mu(t) + \gamma(t)). \tag{6}$$

Perhaps surprisingly, we can find a closed form expression for $g_k(t)$, given the following Lemma (proven in Appendix 1):

LEMMA 1. *Given a broadcaster with intensity $\mu(t)$ and one of her followers with feed intensity due to other broadcasters $\gamma(t)$, the probability $g_k(t)$ that a story posted by the broadcaster is at position k of the follower's feed at time t is given by*

$$g_k(t) = \int_0^t \frac{J^{k-1}(\mu + \gamma, \tau, t)}{(k - 1)!} e^{-J(\mu + \gamma, \tau, t)} \mu(\tau) d\tau, \tag{7}$$

where $J(\lambda, \tau, t) = \int_\tau^t \lambda(x) dx$.

Next, we plug Equation (7) into $r_K(t) = r_K(t, i, j)$, defined in Equation (2), and obtain

$$\mathbb{E}[r_K(t)] = \int_0^t \left[\sum_{k=1}^K \frac{J^{k-1}(\mu + \gamma, \tau, t)}{(k - 1)!} \right] e^{-J(\mu + \gamma, \tau, t)} \mu(\tau) d\tau = (K-1)! \int_0^t \Gamma(K, J(\mu + \gamma, \tau, t)) \mu(\tau) d\tau,$$

where $\Gamma(K, x)$ is the incomplete gamma function. Now, let $F_K(x) = -\sum_{i=0}^{K-1} (K-i) \frac{x^i}{i!} e^{-x}$ be the anti-derivative of $\Gamma(K, x)$. Then, using that $\frac{dJ(\mu + \gamma, \tau, t)}{d\tau} = -(\mu(\tau) + \gamma(\tau))$, $J(\mu + \gamma, t, t) = 0$, $F_K(0) = -K$

---

[5]In Karimi et al. [20], visibility is defined as the probability that at least one story posted by i is among the K recent stories in j's feed.

and $J(\mu + \gamma, 0, t) = \int_0^t \mu(x) + \gamma(x)dx = \mathbb{E}[(A^T N(t))_j]$, we can simplify the above expression into:

$$\mathbb{E}\left[r_K(t)\right] = K + F_K(\mathbb{E}[(A^T N(t))_j]) - \frac{1}{(K-1)!} \int_0^t \Gamma(K, J(\mu + \gamma, \tau, t))\gamma(\tau)d\tau. \tag{8}$$

Finally, if we plug Equation (8) into Equation (2), then we obtain an expression for the average top K visibility of broadcaster i with respect to follower j in terms of the intensity functions $\mu(t)$ and $\gamma(t)$:

$$\mathcal{U}_K(i, j) = K(t_f - t_0) + \int_{t_0}^{t_f} F_K\left(\mathbb{E}[(A^T N(t))_j]\right) dt - \frac{1}{(K-1)!} \int_{t_0}^{t_f} \int_0^t \Gamma(K, J(\mu + \gamma, \tau, t))\gamma(\tau)d\tau dt. \tag{9}$$

Given a set of broadcasters $\mathcal{B}$, we can proceed similarly as above and show that:

(i) The probability $g_k(t, \mathcal{E}_\mathcal{B}, j)$ that a story posted by the broadcasters is at position k of user j's feed at time t is given by Equation (7) with $\gamma(t) = \gamma_{j \setminus \mathcal{B}}(t)$ and $\mu(t) = \sum_{i \in \mathcal{B} : (i,j) \in \mathcal{E}_\mathcal{B}} \mu_i(t)$.

(ii) The average top K visibility $\mathcal{U}_K(\mathcal{E}_\mathcal{B}, j)$ of the broadcasters with respect to user j is given by Equation (9) with $\gamma(t) = \gamma_{j \setminus \mathcal{B}}(t)$ and $\mu(t) = \sum_{i \in \mathcal{B} : (i,j) \in \mathcal{E}_\mathcal{B}} \mu_i(t)$.

Finally, the above results allow us to write $\mathcal{U}_K(\mathcal{E}_\mathcal{B})$, given by Equation (3), in terms of intensity functions characterizing the broadcasters and the feeds, as we were aiming for.

The network visibility problem. Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a directed network, where each node $i \in \mathcal{V}$ has an intensity $\mu_i(t)$. Then, given a set of candidate edges $\mathcal{W}$ provided by a link recommendation algorithm, with $\mathcal{W} \cap \mathcal{E} = \emptyset$, and a set of broadcasters $\mathcal{B} \subseteq \mathcal{V}$, our goal is to find a subset of those edges $\mathcal{E}_\mathcal{B} \subseteq \mathcal{W}$ that maximizes the average top K visibility of the broadcasters $\mathcal{U}_K(\mathcal{E}_\mathcal{B})$ under a given constraint on the number of edges per broadcaster, i.e.,

$$\begin{aligned} \underset{\mathcal{E}_\mathcal{B} \subseteq \mathcal{W}}{\text{maximize}} \quad & \mathcal{U}_K(\mathcal{E}_\mathcal{B}) \\ \text{subject to} \quad & |\{j \in \mathcal{V} : (i,j) \in \mathcal{E}_\mathcal{B}\}| \leq c_i, \quad \forall i \in \mathcal{B}, \end{aligned} \tag{10}$$

where c_i is the maximum number of edges that broadcaster i can *afford*. Here, we assume that a social media platform will charge broadcasters for each edge recommendation and, without loss of generality, we assume that each edge recommendation has a cost of one unit to the broadcaster and the followers are equally likely to follow the recommended broadcasters.

Finally, note that, in the above expression, we can rewrite the constraints as $|\mathcal{B}|$ partition matroid constraints [25], i.e., $|\mathcal{E}_\mathcal{B} \cap \mathcal{W}_{i,:}| \leq c_i \forall i \in \mathcal{B}$, where $\mathcal{W}_{i,:}$ denotes the ground set of possible outgoing edges from i.

3 ON THE THEORETICAL PROPERTIES OF VISIBILITY

In this section, we will show that, for a large family of intensity functions, the average top K visibility satisfies $\check{\alpha}$-submodularity [40], a notion of approximate submodularity.

Definition 2. A set function $F : \mathcal{W} \to \mathbb{R}$ is called $\check{\alpha}$-submodular if, for any $w \in \mathcal{W}$ and subsets $\mathcal{A} \subseteq \mathcal{B} \subseteq \mathcal{W}$,

$$F(\mathcal{A} \cup \{w\}) - F(\mathcal{A}) \geq (1 - \check{\alpha})\left[F(\mathcal{B} \cup \{w\}) - F(\mathcal{B})\right], \tag{11}$$

where the smallest $\check{\alpha} \in [0, 1]$ such that the above inequality is true is called the inverse generalized curvature [6, 17, 26].

Let $\{\mu_i(t)\}_{i \in \mathcal{B}}$ be the intensities of a set of broadcasters $\mathcal{B}$ and $\{\gamma_{i \setminus \mathcal{B}}(t)\}_{i \in \mathcal{V}}$ be the users' feed intensities due to other broadcasters. Moreover, define $\inf(\lambda) = \inf\{\lambda(t) \mid t \in (t_0, t_f)\}$, $\sup(\lambda) = \sup\{\lambda(t) \mid t \in (t_0, t_f)\}$. Then, we will assume that:

(i) All intensities are bounded above, i.e., $\sup(\mu_i) \leq \lambda_{\max}$, $\forall i \in \mathcal{B}$ and $\sup(\gamma_{i\backslash\mathcal{B}}) \leq \lambda_{\max}$, $\forall i \in \mathcal{V}$ with $0 < \lambda_{\max} < \infty$.

(ii) The users' feed intensities due to other broadcasters are bounded below, i.e., $\inf(\gamma_{i\backslash\mathcal{B}}) \geq \lambda_{\min}$, $\forall i \in \mathcal{V}$ with $0 < \lambda_{\min} \leq \lambda_{\max}$.

Note that these assumptions are natural in most practical scenarios—the first assumption is satisfied if broadcasters post a finite number of stories per unit of time and the second assumption is satisfied if, at any time, there is always a nonzero probability that a user's feed receives a post from the other broadcasters. Under these assumptions, it readily follows that, for any set $\mathcal{E}_{\mathcal{B}}$ of outgoing edges from the broadcasters, the feed intensities $\gamma_i(t)$ are ξ-bounded, i.e., $\sup(\gamma_i) \leq \xi \inf(\gamma_i)$, with $\xi \leq \frac{\lambda_{\max}}{\lambda_{\min}}(1 + |\mathcal{W}_{:,i}|)$, and we can characterize the inverse generalized curvature $\check{\alpha}$ of the average top K visibility using the following Theorem and Corollary (proven in Appendix B):

THEOREM 3. *Suppose that, for any set $\mathcal{E}_{\mathcal{B}}$, the feed intensities $\gamma_i(t)$ are ξ-bounded and*

$$\mathbb{E}[\boldsymbol{M}_{\backslash\mathcal{B}}(t_0)] \geq K - 1 + \zeta\sqrt{K-1}, \tag{12}$$

for some positive real ζ, which simply states that t_0 is large enough so the expected number of stories posted by the other broadcasters by t_0 in each feed is greater than the RHS. Then, the inverse generalized curvature $\check{\alpha}$ of the average top K visibility $\mathcal{U}_K(\mathcal{E}_{\mathcal{B}})$, defined by Equation (3), satisfies that,

$$\check{\alpha} \leq \check{\alpha}^* = 1 - \frac{1}{\xi\left(\frac{6.154e^2}{1-\frac{4e^{\frac{7}{4}}}{\zeta}}\zeta\sqrt{K} + 1\right)}. \tag{13}$$

COROLLARY 4. *The inverse generalized curvature of the average top K visibility is $\check{\alpha} \approx 1 - \Omega\left(\frac{1}{\xi^2\sqrt{K}}\right)$.*

Moreover, we can tighten the upper bound $\check{\alpha}^*$ given by Theorem 3 using the following Theorem (proven in Appendix B) if we assume that, at each time t, the intensity function $\mu_i(t)$ of each broadcaster is lower than a fraction of each of her follower's feed intensities. This is an assumption that is likely to hold in practice, since the stories posted by a single broadcaster are typically a small percentage of the stories her followers receive in their feeds over time.

THEOREM 5. *Suppose the conditions in Theorem 3 hold and, in addition,*

$$\mu_i(t) \leq \frac{\gamma_{j\backslash\mathcal{B}}(t)}{\rho\sqrt{K-1}}, \quad \forall t \in [t_0, t_f], \forall i \in \mathcal{B}, \forall j \in \mathcal{V}, \tag{14}$$

where $\rho > 0$. Then, the inverse generalized curvature $\check{\alpha}$ of the average top K visibility $\mathcal{U}_K(\mathcal{E}_{\mathcal{B}})$ satisfies that

$$\check{\alpha} \leq \check{\alpha}^* = 1 - \frac{1}{\xi\left(\frac{2e^{\frac{7}{4}}}{\left(1-\frac{4e^{\frac{7}{4}}}{\zeta}\right)\min\{\rho,1\}e^{-\frac{1}{\rho^2}}}\xi + 1\right)}. \tag{15}$$

Note that we can always find a constant $\rho > 0$ so Equation (14) is satisfied, however, the term $\min\{\rho, 1\}e^{-\frac{1}{\rho^2}}$ will decrease drastically when $\rho \geq 1$. Finally, by combining Theorems 3 and Theorem 5, we obtain the following Corollary:

COROLLARY 6. *Suppose the conditions in Theorem 5. Then, the inverse generalized curvature of the average top K visibility is*

$$\check{\alpha} \approx 1 - \Omega\left(\frac{1}{\xi^2}\max\left\{\frac{1}{\sqrt{K}}, \rho e^{-\frac{1}{\rho^2}}\right\}\right). \tag{16}$$

Moreover, if $\rho \geq 1$, then $\check{\alpha} \approx 1 - \Omega\left(\frac{1}{\xi^2}\right)$.

ALGORITHM 1: Standard greedy algorithm

Input: Ground set $\mathcal{W}$, partition $\{\mathcal{N}_i\}_{i=1}^n$ of $\mathcal{W}$, matroid collection $\{\mathcal{I}_i\}_{i=1}^n$, set function F
Output: Set of items $\mathcal{S}_n$
1: $S_0 \leftarrow \varnothing$
2: **for** $i = 1, \ldots, n$ **do**
3:     $S_i^0 \leftarrow \varnothing, U_i^0 \leftarrow \varnothing$
4:     **for** $j = 1, \ldots, |\mathcal{N}_i|$ **do**   % Proceed iteratively over partitions
5:         $v^* \leftarrow \mathrm{argmax}_{v \in \mathcal{N}_i \setminus U_i^{j-1}} F(S_{i-1} \cup S_i^{j-1} \cup \{v\}) - F(S_{i-1} \cup S_i^{j-1})$ % Find best item
6:         $U_i^j \leftarrow U_i^{j-1} \cup \{v^*\}$ % Item is considered
7:         **if** $S_i^{j-1} \cup \{v^*\} \in \mathcal{I}_i$ **then**
8:             $S_i^j \leftarrow S_i^{j-1} \cup \{v^*\}$ % Item is selected
9:         **end if**
10:     **end for**
11:     $S_i \leftarrow S_{i-1} \cup S_i^{|\mathcal{N}_i|}$
12: **end for**
13: **return** $\mathcal{S}_n$

The above result implies that the larger the difference between the upper bound on the broadcasters' intensities and the lower bound on the user's feed intensities due to other broadcasters, the greater the inverse generalized curvature of the average top K visibility.

4   SOLVING THE NETWORK VISIBILITY PROBLEM

In this section, we first show that the standard greedy algorithm [29] (Algorithm 1) achieves an approximation factor of $(1 + 1/(1 - \check{\alpha}))^{-1}$ at maximizing nondecreasing $\check{\alpha}$-submodular set functions under n matroid constraints. Finally, we conclude with a theoretical analysis of the complexity and finite-sample robustness of the greedy algorithm.

The standard greedy algorithm. Given a ground set $\mathcal{W}$, a partition $\{\mathcal{N}_i\}_{i=1}^n$ of $\mathcal{W}$, a collection of matroids $\{\mathcal{I}_i\}_{i=1}^n$ with respect to each partition, and a nondecreasing set function $F : \mathcal{W} \to \mathbb{R}$ with inverse generalized curvature $\check{\alpha}$, the greedy algorithm proceeds iteratively over the partitions $\mathcal{N}_i$. At each iteration, it *considers* the best item that provides the highest marginal gain, among the set of items in the corresponding partition that have not yet been considered in previous iterations. Then, it adds this item to the set of selected items if it does not violate the corresponding matroid constraint. Algorithm 1 summarizes the greedy algorithm.

In the network visibility problem, defined by Equation (10), $\mathcal{W}$ is the ground set of possible outgoing edges from the broadcasters in $\mathcal{B}$, each partition $\mathcal{N}_i = \mathcal{W}_{i,:}$ is the set of possible outgoing edges from broadcaster $i \in \mathcal{B}$, the constraint that each broadcaster $i \in \mathcal{B}$ can pick at most c_i edges is a type of matroid constraint—a cardinality constraint—over $\mathcal{N}_i$, and the set function $F = \mathcal{U}_K(\mathcal{E}_{\mathcal{B}})$ is nondecreasing (i.e., adding more edges always provides positive gain) and its inverse generalized curvature $\check{\alpha}$ is given by Equations (13) or (15). Here, note that, in the first iterations, the greedy algorithm may just connect broadcasters to followers with the lowest feed intensity. However, the feed of these followers will increase its intensity as the greedy algorithm connects them to more broadcasters and, as a consequence, in later iterations, the greedy algorithm may not necessarily assign broadcasters to these followers anymore.

Approximation guarantees of the greedy algorithm. Our main result is the following theorem, which shows that the greedy algorithm achieves an approximation factor of $(1 + 1/(1 - \check{\alpha}))^{-1}$ (proven in Appendix D):

THEOREM 7. *Given a ground set $\mathcal{W}$, a partition $\{\mathcal{N}_i\}_{i=1}^{n}$ of $\mathcal{W}$, a collection of matroids $\{\mathcal{I}_i\}_{i=1}^{n}$ with respect to each partition, and a nondecreasing set function $F : \mathcal{W} \to \mathcal{R}$ with inverse generalized curvature $\check{\alpha}$. Then, the greedy algorithm returns a set $\mathcal{S}$ such that*

$$F(\mathcal{S}) \geq \frac{OPT}{1 + 1/(1 - \check{\alpha})},$$

where OPT is the optimal value.

In the network visibility problem, the above theorem implies that the greedy algorithm provides a $O(\xi^2 \sqrt{K})$ approximation factor. Moreover, we would like to highlight that the proof of the above theorem differs significantly from that of Theorem 2.1 in Nemhauser [29], which is much more involved. More specifically, in our proof, we cannot apply Proposition 2.2 in Nemhauser [29], because the decreasing monotonicity of the marginal gains of the added elements fails to hold in the absence of the submodularity condition. As a result, our proof does not resort to linear program duality and it is generalizable for the case of having an intersection of P matroids instead of just one.

In comparison with the previous literature on the maximization of non-submodular nondecreasing set functions under matroid constraints, the result most closely related to ours is by Chen et al. [9], who have recently shown that a randomized version of the standard greedy algorithm enjoys an approximation factor of $(1 + 1/\gamma)^{-2}$, where γ is the submodularity ratio [12]. However, to the best of our knowledge, the maximization of α-submodular functions under matroid constraints has not been analyzed before in the literature.

Computational complexity of the greedy algorithm. In the network visibility problem, defined by Equation (10), the greedy algorithm (Algorithm 1) is computationally efficient. More specifically, the computational complexity is given by the following Theorem (proven in Appendix E):

THEOREM 8. *In the network visibility problem, the greedy algorithm terminates in $O(\kappa|\mathcal{B}||\mathcal{V}|)$, where κ is the maximum time needed to evaluate the average top K visibility $\mathcal{U}_K(\mathcal{E}_\mathcal{B}, j)$ for any $j \in \mathcal{V}$, $\mathcal{B}$ is the set of broadcasters and $\mathcal{V}$ is the set of users.*

Robustness of the greedy algorithm. To solve the network visibility problem with the greedy algorithm, we need to compute the average top K visibility, which depends on a set of *unknown* intensities of the broadcasters and feeds. In practice, one could adopt a specific functional form for these intensities and fit them using historical data, however, that could lead to poor estimates of the visibility and, more importantly, it would be difficult to assess the impact of these empirical estimates on the approximation guarantees of the greedy algorithm. Instead, we will directly estimate the average top K visibility using historical data.

Given a directed network $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, a set of broadcasters $\mathcal{B} \subseteq \mathcal{V}$ with intensities $\{\mu_i(t)\}_{i \in \mathcal{B}}$, edges $\mathcal{E}_\mathcal{B} \subseteq \mathcal{E}$ from the broadcasters to their followers, followers' feed intensities due to other broadcasters $\gamma_{j \setminus \mathcal{B}}(t)$, and n sequences of posts of length $\Delta = t_f - t_0$, our empirical estimate of the average top K visibility $\mathcal{U}(\mathcal{E}_\mathcal{B}, j)$ of the broadcasters in user j's feed is given by

$$\hat{\mathcal{U}}(\mathcal{E}_\mathcal{B}) = \sum_{j \in \mathcal{V}} \hat{\mathcal{U}}(\mathcal{E}_\mathcal{B}, j) = \sum_{j \in \mathcal{V}} \frac{\sum_{i=1}^{K} \sum_{\ell=1}^{n} \Delta_{i,j}^{(\ell)}}{n}, \tag{17}$$

where $\Delta_{i,j}^{(\ell)}$ is the amount of time that a post from the set of broadcasters $\mathcal{B}$ is at the ith position of user j's feed in realization ℓ. Here, note that the empirical estimate does not explicitly depend

on the intensities of the broadcasters and feeds and, given an arbitrary set of edges $\mathcal{E}_{\mathcal{B}}$, one can always measure the visibility they would reach without actual interventions. In contrast, measures based on users' reactions (e.g., number of likes, shares, and replies) are difficult to estimate from real data, due to the presence of other confounding factors such as users' influence, content, and wording.

Further, we perform a formal analysis of the sample complexity of the above empirical estimate. Let $0 \leq y \leq 1$, then we can show that:

$$\mathbb{P}\{\mathcal{U}(\mathcal{E}_{\mathcal{B}}, j) - yK(t_f - t_0) \leq \hat{\mathcal{U}}(\mathcal{E}_{\mathcal{B}}, j) \leq \mathcal{U}(\mathcal{E}_{\mathcal{B}}, j) + yK(t_f - t_0)\} \geq 1 - e^{Z_y - Q_y n}, \qquad (18)$$

where Z_y and Q_y are functions of, e.g., α, $\mathbb{E}[A_j^T(N(t_f) - N(t_0))]$,

$$\beta_j = \inf_{t \in (t_0, t_f)} \frac{\mu(t)}{\mu(t) + \gamma_{j \setminus \mathcal{B}}(t)}, \quad \rho_j = \sup_{t \in (t_0, t_f)} \frac{\mu(t)}{\mu(t) + \gamma_{j \setminus \mathcal{B}}(t)}, \text{ and } \alpha.$$

However, for space constraints, we defer the details of this formal analysis to Appendix F. Finally, given the above error bound, we characterize the approximation factor that the greedy algorithm achieves if it uses these empirical estimates with the following Theorem (proven in Appendix G):

THEOREM 9. *Let the number of realizations* $n \geq \frac{Z_y + \log \frac{\sum_{i \in \mathcal{B}} |W_{i,:}|}{\delta}}{Q_y}$. *Then, with probability at least* $1 - \delta$, *the greedy algorithm returns a set of edges* $\mathcal{E}_{\mathcal{B}}$ *such that*

$$\mathcal{U}_K(\mathcal{E}_{\mathcal{B}}) \geq \frac{OPT}{1 + 1/(1 - \check{\alpha})} - 4yK(t_f - t_0) \sum_{i \in \mathcal{B}} c_i.$$

5 EXPERIMENTS ON SYNTHETIC DATA

In this section, we evaluate the greedy algorithm (Algorithm 1) using synthetic data. To this end, we first compare its performance with that of several competitive baselines and then validate its robustness and scalability.

Experimental setup. Unless stated otherwise, we use (periodic) piece-constant intensities $\mu_i(t) = \sum_{k=0}^{T-1} \mu_{i,k} I(t \in [t_k, t_{k+1}])$ and $\gamma_{j \setminus \mathcal{B}}(t) = \sum_{j=0}^{T-1} \gamma_{j,k} I(t \in [t_k, t_{k+1}])$ for the broadcasters and the feeds, respectively, where $T = 24$ days is the period, $t_{k+1} - t_k = 1$ day is the length of each piece and, each piece, we pick $\mu_{i,k}$ and $\gamma_{j,k}$ uniformly at random. Note that, for piece-constant intensities, we are able to compute $\mathcal{U}_K(\mathcal{E}_{\mathcal{B}}, j)$ analytically. We compare the performance of the greedy algorithm with the same three heuristics as in the main section of the article,[6] in which each broadcaster picks c_i feeds in turn and (i) it picks those with the lowest feed intensity $\int_0^T \gamma(t)dt$ by the time it chooses (CP), (ii) it picks those with the lowest visibility $\mathcal{U}_K$ by the time it chooses (UP), and (iii) it picks those with the lowest value of $\mathcal{U}_K \times \int_0^T \gamma(t)dt$ (CUP). Here, we will run both our greedy algorithm and the baselines using the true intensity values and then report the average (theoretical) value of top K visibility, however, note that all can be run using empirical estimates of the relevant quantities, i.e., $\mathcal{U}_K$ using Equation (17) or $\int_0^T \gamma(t)dt$, using maximum likelihood estimation.

Results. We first compare the performance of the greedy algorithm and all baselines in a setting with 60 broadcasters and 600 feeds. Figure 1 summarizes the results, which show that the greedy algorithm beats the baselines by large margins under different K and ξ values. We did experiment with a wide range of parameter settings (e.g., K, ξ, T or c_i) and found that the greedy algorithm

---

[6]Initially, we also considered a trivial baseline that picks edges uniformly at random, however, its performance was not at all competitive and we decided to omit it.
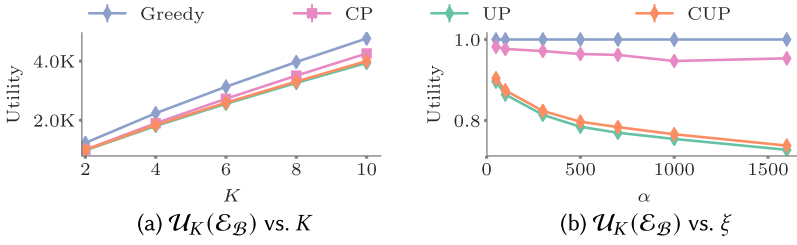
Fig. 1. Average top K visibility achieved by the greedy algorithm (GP) and the three heuristics for a setting with 60 broadcasters and 600 feeds. Panel (a) shows the average top K visibility, $\mathcal{U}_K(\mathcal{E}_\mathcal{B})$, for different K values, where we sampled $\mu_{i,k}$ and $\gamma_{j,k}$ from $U[0.01, 0.1]$ and $U[0.4, 50]$, respectively, and $c_i = 20$. Panel (b) shows the average top K visibility, normalized with respect to the visibility achieved by the greedy algorithm, where we sampled $\mu_{i,k}$ and $\gamma_{j,k}$ from $U[0.01, 0.1]$ and $U[0.05, 0.05] \times \xi$, respectively, $K = 10$ and $c_i = 50$.
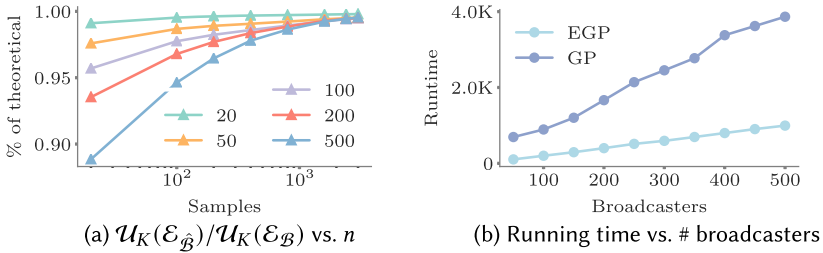

Fig. 2. Robustness and running time of the greedy algorithm. Panel (a) compares the visibility achieved by the solution $\mathcal{E}_\mathcal{B}$ ($\mathcal{E}_{\hat{\mathcal{B}}}$) provided by the greedy algorithm using the theoretical (empirical) visibility, where we sampled $\mu_{i,k}$ and $\gamma_{j,k}$ from $U[0.001, 0.1]$ and $U[0.05, 0.05 \times \xi]$, respectively, $K = 10$ and $c_i = 30$. Panel (b) shows the overall running time of the greedy algorithm using the theoretical (GP) and empirical (GP) visibility, where we sampled $\mu_{i,k}$ and $\gamma_{j,k}$ from $U[0.001, 0.1]$ and $U[0.05, 0.05 \times \xi]$, $K = 10$, and $c_i = 30$.

consistently beats the baselines. Next, we compare the visibility values achieved by the solution $\mathcal{E}_\mathcal{B}$ the greedy algorithm provides using the theoretical visibility, given by Equation (9), against the solution $\mathcal{E}_{\hat{\mathcal{B}}}$ it provides using the empirical visibility, given by Equation (17). Figure 2(a) summarizes the results, which show that, in agreement with Theorem 9, the quality of the solution the greedy algorithm provides using the empirical visibility converges to the one it provides using the theoretical visibility. Finally, we compute the running time of the greedy algorithm against the number of broadcasters. Figure 2(b) summarizes the results, which show that the running time is linear in the number of broadcasters, also in agreement with Theorem 8. In additional experiments, we also found that the running time is linear in the number of walls, superlinear with respect to the number of pieces T and it is independent on the budget per broadcaster, however, for space constraints, we do not include the corresponding plots.

6   EXPERIMENTS ON REAL DATA

In this section, we compare the performance of the greedy algorithm and several competitive baselines using data gathered from Twitter.

Data description and experimental setup. We use data gathered from Twitter as reported in previous work [8], which comprises profiles of 52 million users, 1.9 billion directed follow links among these users, and 1.7 billion public tweets. The follow link information is based on a snapshot taken at the time of data collection, in September 2009. Here, we focus on the tweets posted during
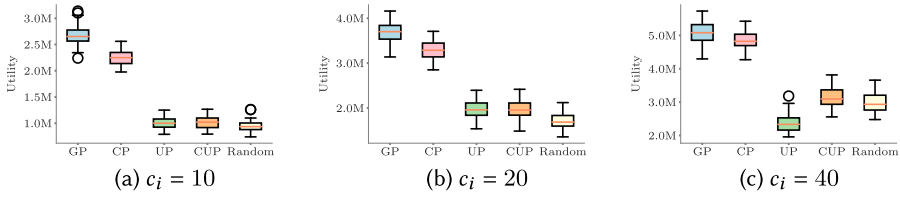
Fig. 3. Average top $K = 10$ visibility achieved by the greedy algorithm, the three heuristics and the trivial baseline using Twitter data. The solid horizontal line shows the median visibility and the box limits correspond to the 25%–75% percentiles.

a two month period, from July 1, 2009, to September 1, 2009, to be able to consider the social graph to be approximately static, sample a set $\mathcal{A}$ of $2,000$ users uniformly at random, record all the tweets they posted. Then, for $c_i \in \{10, 20, 40\}$, we repeat the following procedure 50 times:

1. we pick uniformly at random a set $\mathcal{B} \subseteq \mathcal{A}$ of 80 users as broadcasters;
2. for each broadcaster i, we pick uniformly at random a set $\mathcal{H}_i$ of 20 of their followers;
3. we record all tweets not posted by broadcasters in $\mathcal{B}$ in the feeds of the users in $\mathcal{H} = \cup_i \mathcal{H}_i$; and,
4. we run the greedy algorithm, three heuristic methods, and a trivial baseline that picks edges uniformly at random and record the sets $\mathcal{E}_{\mathcal{B}} = \{(i, j) : i \in \mathcal{B}, j \in \mathcal{H}\}$ each provides.

Under the three heuristics, each broadcaster picks c_i feeds in turn and (i) it picks those with the lowest feed intensity $\int_0^T \gamma(t)dt$ by the time it chooses (CP), (ii) it picks those with the lowest visibility $\mathcal{U}_K$ by the time it chooses (UP), or (iii) it picks those with the lowest value of $\mathcal{U}_K \times \int_0^T \gamma(t)dt$ (CUP).

In each repetition of the above procedure, we use the tweets posted during the first month to compute empirical estimates of the relevant quantities used by all methods, such as $\mathcal{U}_K$ or $\int_0^T \gamma(t)dt$, and we use the tweets posted during the second month to compute empirical estimates of $\mathcal{U}_K$ achieved by each method.

Solution quality. Figure 3 summarizes the results by means of box plots, which show that the greedy algorithm consistently beats all heuristics and the trivial baseline. Moreover, we did experiment with other parameters settings (e.g., $|\mathcal{B}|$, $|\mathcal{H}|$, K, and c_i) and found our method to be consistently superior to alternatives.

7 CONCLUSIONS

In this article, we have introduced the network visibility problem where, given a set of candidate links provided by a link recommendation algorithm, the goal is to find the subset of those links that maximize the average visibility a set of users achieve. Then, we have shown that the corresponding objective function satisfies a notion of approximate submodularity, $\check{\alpha}$-submodularity, which allows the standard greedy algorithm to enjoy theoretical guarantees. Finally, using real data gathered from Twitter, we have demonstrated that the greedy algorithm is able to consistently outperform several baselines in practice.

Our work opens many venues for future work. For example, in this work, we have relied on previous link recommendation algorithms to pick a set of candidate edges that followers may find interesting. A natural next step is to design algorithms that jointly maximize the broadcasters' and followers' utilities. Moreover, we have assumed each user's feed ranks stories in inverse chronological order. It would be interesting to lift this assumption and allow for algorithmic feeds. In

practice, different users may be more or less likely to follow the recommended broadcasters, therefore, it would be worthwhile to augment our algorithm to account for that uncertainty. In addition, as some followers might have greater potential to amplify a broadcaster's posts, it would be interesting to weigh differently each follower's contribution to the broadcasters' utilities. Moreover, it would be interesting to understand whether our approximation guarantees for $\check{\alpha}$-submodular function maximization under matroid constraints are tight. Finally, it would be interesting to find other application domains where $\check{\alpha}$-submodular functions emerge.

APPENDICES

A   PROOF OF LEMMA 1

We prove this lemma by induction on k. For $g_0(t)$, we have the boundary condition $g_1(0) = 0$, trivially, because the expected number of stories at position $k = 1$ is zero at the beginning of time. Moreover, $g_1(t)$ satisfies the first order differential equation $g_1'(t) = -(\mu(t) + \gamma(t))g_1(t) + \mu(t)$ with boundary condition $g_1(0) = 0$, which has the unique solution

$$g_1(t) = \int_0^t e^{-\int_\tau^t \mu(x) + \gamma(x)dx} \mu(\tau)d\tau = \int_0^t e^{-J(\mu+\gamma,\tau,t)} \mu(\tau)d\tau,$$

which proves the base of induction. Now, for $g_k(t)$, we have the following recursive differential equation:

$$g_k'(t) = (g_{k-1}(t) - g_k(t))(\mu(t) + \gamma(t)).$$

Suppose we know the closed form formula

$$g_k(t) = \int_0^t \frac{J^{k-1}(\mu+\gamma,\tau,t)}{(k-1)!} e^{-J(\mu+\gamma,\tau,t)} \mu(\tau)d\tau.$$

Then, $g_k(t)$ is again satisfying a first-order differential equation with boundary condition $g_i(0) = 0$. Thus, it suffices to show the closed form formula satisfies the above recursive differential equation.

According to Leibnitz formula for calculating differentials of integrals with non-constant limits,

$$
\begin{aligned}
g_k'(t) = {} & \frac{J^{k-1}(\mu+\gamma,\tau,t)}{(k-1)!} e^{-J(\mu+\gamma,\tau,t)} \mu(\tau) \Big|_{\tau=0}^{\tau=t} \frac{d}{dt}(t) + \int_0^t \frac{J^{k-2}(\mu+\gamma,\tau,t)}{(k-2)!} e^{-J(\mu+\gamma,\tau,t)} \frac{dJ}{dt} \mu(\tau)d\tau \\
& - \int_0^t \frac{J^{k-1}(\mu+\gamma,\tau,t)}{(k-1)!} e^{-J(\mu+\gamma,\tau,t)} \frac{dJ}{dt} \mu(\tau)d\tau \\
= {} & (g_{k-1}(t) - g_k(t))(\mu(t) + \gamma(t)),
\end{aligned}
$$

where $\frac{dJ}{dt} = \frac{d}{dt}(J(\mu+\gamma,\tau,t))$ and the last line follows from the fact that $\frac{dJ}{dt} = \mu(t) + \gamma(t)$. This completes the step of induction.

B   PROOF SKETCH OF THEOREMS 3 AND 5

To (upper) bound the inverse generalized curvature $\check{\alpha}$ of the average top K visibility $\mathcal{U}_K(\mathcal{E}_\mathcal{B})$, defined by Equation (3), we have to provide a $\check{\alpha}^* \geq \check{\alpha}$ such that, for any $w \in \mathcal{W}$ and subsets $\mathcal{A} \subseteq \mathcal{B} \subseteq \mathcal{W}$, it holds that

$$F(\mathcal{A} \cup \{w\}) - F(\mathcal{A}) \geq (1 - \check{\alpha}^*)[F(\mathcal{B} \cup \{w\}) - F(\mathcal{B})], \tag{19}$$

The following Lemma lets us omit several sums and integrations that Equation (3) depends on while deriving a bound for the inverse generalized curvature.

LEMMA 10. *Let $\{F_\sigma\}_{\sigma \in \wp}$ be a family of set functions $F_\sigma : \mathcal{W} \to \mathbb{R}$ parametrized by σ, such that for each fixed $\sigma \in \wp$, F_σ is a set function with inverse generalized curvature $\check{\alpha}_\sigma \leq \check{\alpha}^*$. Then, the following statements hold:*

— Suppose $\wp = (a, b) \subseteq \mathbb{R}$ and $\tilde{F}(S) = \int_a^b F_\sigma(S) \, d\sigma$ for every subset $S \subseteq \mathcal{V}$ and the integral always exists. Then, $\tilde{F}$ has inverse generalized curvature $\check{\alpha} \leq \check{\alpha}^*$.

— Suppose $\wp$ is a discrete set and $\tilde{F}(S) = \sum_{\sigma \in \wp} F_\sigma(S)$. Then, $\tilde{F}$ has inverse generalized curvature $\check{\alpha} \leq \check{\alpha}^*$.

In particular, for both cases introduced, if for each σ, F_σ is submodular, then F is submodular as well.

PROOF. By the definition of inverse generalized curvature, $\forall v \in \mathcal{V}, A \subseteq B, \forall \sigma$:

$$F_\sigma(A \cup \{v\}) - F_\sigma(A) \geq (1 - \check{\alpha}_\sigma)(F_\sigma(B \cup \{v\}) - F_\sigma(B)) \geq (1 - \check{\alpha}^*)(F_\sigma(B \cup \{v\}) - F_\sigma(B)). \quad (20)$$

By integrating over both sides of Equation (20) over σ we get

$$F(A \cup \{v\}) - F(A) \geq (1 - \check{\alpha}_\sigma)(F(B \cup \{v\}) - F(B)) \geq (1 - \check{\alpha}^*)(F(B \cup \{v\}) - F(B)),$$

which proves that F has inverse generalized curvature $\check{\alpha} \leq \check{\alpha}^*$. The proof for the second part follows the same way by summing over $\sigma \in \wp$. □

More specifically, using the first statement of Lemma 10, to obtain an upper bound for the inverse generalized curvature for $\mathcal{U}_K(\mathcal{E}_\mathcal{B})$, it is sufficient to obtain an upper bound for each $\mathcal{U}_K(\mathcal{E}_\mathcal{B}, j)$, which is a summation of three terms, as given by Equation (9). The first term, $K(t_f - t_0)$, is a constant and does not appear in the marginal gains $\{\rho_e\}_{e \in \mathcal{E}_\mathcal{B}}$, therefore, it does not affect the inverse generalized curvature. Using the second statement of Lemma 10, it is sufficient to obtain an upper bound for the inverse generalized curvature of the second and third term separately.

The second term is the integration of the function F_K in the time interval (t_0, t_f). Therefore, using the first statement of Lemma 10, it is sufficient to obtain an upper bound for the inverse generalized curvature of the function F_K for any $t \in [t_0, t_f]$. Here, note that the function F_K only depends on the edges pointing at follower j, therefore, its ground set $\mathcal{W}_{\cdot, j}$ is the ground set of possible outgoing edges from the broadcasters to j. Next, we will use the following Lemma to show that F_K is a submodular function and thus its inverse generalized curvature is $\check{\alpha} = 1$:

LEMMA 11. Let $F : \mathcal{W} \to \mathbb{R}^+, F(S) = f(c + G(S))$ be a nonnegative set function, where f is a concave function over $\mathbb{R}^+$, G is a nonnegative modular function over $\mathcal{W}$, and c is a nonnegative constant. Then, F is submodular.

PROOF. According to Lemma 15 and by using the property of modular functions, for $A \subseteq B$ and $v \in V$

$$\begin{aligned} F(B \cup \{v\}) - F(B) &= f(c + G(B \cup \{v\})) - f(c + G(B)) \\ &= f(c + G(B \setminus A) + G(A \cup \{v\})) - f(c + G(B \setminus A) + G(A)) \\ &\leq f(c + G(A \cup \{v\})) - f(c + G(A)) = F(A \cup \{v\}) - F(A). \quad\square \end{aligned}$$

By definition,

$$\mathbb{E}[(A^T N(t))_j] = \sum_{i \in \mathcal{B} : (i,j) \in \mathcal{E}_\mathcal{B}} \mathbb{E}[N_i(t)] + \sum_{i \notin \mathcal{B} : (i,j) \in \mathcal{E}} \mathbb{E}[N_i(t)].$$

For a fixed t, the above equation reveals that $\mathbb{E}[(A^T N(t))_j]$ is a sum of a modular function over the groundset $\mathcal{W}_{\cdot, j}$ of possible incoming edges to j and a constant. Moreover, F_K is concave in $\mathbb{R}^+$, since $F_K''(x) = -\frac{x^{K-1}e^{-x}}{(K-1)!} \leq 0$. Therefore, we can use Lemma 11 to conclude that $F_K(\mathbb{E}[(A^T N(t))_j])$ is submodular over the groundset $\mathcal{W}_{\cdot, j}$. This means that an upper bound for the inverse generalized curvature of the third term in Equation (9) will be an upper bound for the inverse generalized curvature of $\mathcal{U}_K(\mathcal{E}_\mathcal{B})$, on the grounds of Lemma 10.

To upper bound the inverse generalized curvature of the third term, using the second statement of Lemma 10, it is sufficient to provide an upper bound on the inverse generalized curvature of

$$- \int_0^t \Gamma(K, J(\mu + \gamma, \tau, t))\gamma(\tau)d\tau := \mathcal{P}_j(\mathcal{E}_{\mathcal{B}}), \qquad (21)$$

where $\mu(\tau) = \sum_{i \in \mathcal{B} : (i,j) \in \mathcal{E}_{\mathcal{B}}} \mu_i(\tau)$ and $\gamma(\tau) = \gamma_{j \setminus \mathcal{B}}(\tau)$. Here, note that the above function depends on the edges $\mathcal{E}_{\mathcal{B}}$ through $\mu(\tau)$. Moreover,

$$J(\mu + \gamma, \tau, t) = \int_\tau^t \sum_{i \in \mathcal{B} : (i,j) \in \mathcal{E}_{\mathcal{B}}} \mu_i(x)dx + \int_\tau^t \gamma_{j \setminus \mathcal{B}}(x)dx,$$

which reveals that $J(\mu+\gamma, \tau, t)$ is the sum of a modular function over the groundset $\mathcal{W}_{\cdot, j}$ of possible incoming edges to j and a constant. Thus, if $-\Gamma(K, x)$ was concave for $x \in \mathbb{R}^+$, then we could combine the second statement of Lemma 10 and Lemma 11 to conclude that $\mathcal{P}_j(\mathcal{E}_{\mathcal{B}})$ is submodular over $\mathcal{W}_{\cdot, j}$. Unfortunately, $-\Gamma$ is convex on $(0, K-1)$ and concave on $(K-1, \infty)$.

With this in mind, we define the time point $0 \leq \tau_0 = \tau_0(\mu+\gamma, t) \leq t$ such that $J(\mu+\gamma, \tau_0, t) = K-1$, which will allow us to analyze the domains $(0, K-1)$ and $(K-1, \infty)$ separately. Here, note that such τ_0 exists because $J(\mu + \gamma, t, t) = 0$,

$$J(\mu + \gamma, 0, t) = \int_0^t \sum_{i \in \mathcal{B} : (i,j) \in \mathcal{E}_{\mathcal{B}}} \mu_i(x) + \gamma_{j \setminus \mathcal{B}}(x)dx \geq \int_0^t \gamma_{j \setminus \mathcal{B}}(x)dx$$

$$= \mathbb{E}[M_{j \setminus \mathcal{B}}(t)] \geq \mathbb{E}[M_{j \setminus \mathcal{B}}(t_0)] \geq K - 1 + \zeta \sqrt{K - 1} \geq K - 1,$$

using Equation (12) in Theorem 3, and $J(\mu + \gamma, \tau, t)$ is a continuous (nonincreasing) function with respect to τ.

The following Lemma (proven in Appendix C) introduces a key inequality to derive an upper bound on the inverse generalized curvature of $\mathcal{P}(\mathcal{E}_{\mathcal{B}})$.

LEMMA 12. *Let $(\mathcal{V}, \mathcal{E})$ be a directed network, $\mathcal{B} \subseteq \mathcal{V}$ a set of broadcasters, $\mathcal{E}_{\mathcal{B}}, \tilde{\mathcal{E}}_{\mathcal{B}}$ two possible sets of outgoing edges for these broadcasters, such that $\mathcal{E}_{\mathcal{B}} \subseteq \tilde{\mathcal{E}}_{\mathcal{B}}$, $j \in \mathcal{V}$ a given user with feed intensity due to other broadcasters $\gamma_{j \setminus \mathcal{B}}(t) = \gamma(t)$,*

$$\mu(t) = \sum_{i \in \mathcal{B}:(i,j) \in \mathcal{E}_{\mathcal{B}}} \mu_i(t) \quad and \quad \tilde{\mu}(t) = \sum_{i \in \mathcal{B}:(i,j) \in \tilde{\mathcal{E}}_{\mathcal{B}}} \mu_i(t)$$

be the intensities due to the broadcasters in $\mathcal{B}$ in user j's feed, and assume the intensities $\gamma(t)$, $\gamma(t)+\mu(t)$ and $\gamma(t) + \tilde{\mu}(t)$ to be ξ-bounded. Consider a broadcaster $i \in \mathcal{B}$ with intensity $\mu_i(t) = \lambda(t)$, such that $(i,j) \notin \tilde{\mathcal{E}}_{\mathcal{B}}$. Then, under the conditions of Theorem 3, it holds that

$$\theta \left[\int_0^t \Gamma(K, J(\mu + \gamma, \tau, t))d\tau - \int_0^t \Gamma(K, J(\mu + \gamma + \lambda, \tau, t))d\tau \right]$$

$$\geq \int_0^t \Gamma(K, J(\tilde{\mu} + \gamma, \tau, t))d\tau - \int_0^t \Gamma(K, J(\tilde{\mu} + \gamma + \lambda, \tau, t))d\tau, \qquad (22)$$

$$where \qquad \theta = \theta_1 = \frac{6.154e^2}{1 - \frac{4e^{\frac{7}{4}}}{\zeta}} \xi \sqrt{K} + 1. \qquad (23)$$

Moreover, under extra condition of Theorem 5, Equation (22) also holds for

$$\theta = \theta_2 = \frac{2e^{\frac{7}{4}}}{\left(1 - \frac{4e^{\frac{7}{4}}}{\zeta}\right) \min\{\rho, 1\}e^{-\frac{1}{\rho^2}}} \xi + 1. \qquad (24)$$

With the above Lemma, we are now ready to derive an upper bound on the inverse generalized curvature of $\mathcal{P}_j(\mathcal{E}_\mathcal{B})$. Consider the same definitions and assumptions as in the above Lemma. Then,

$$
\begin{aligned}
\mathcal{P}_j(\mathcal{E}_\mathcal{B} \cup \{(i,j)\}) - \mathcal{P}_j(\mathcal{E}_\mathcal{B}) &= \int_0^t (\Gamma(K, J(\mu + \gamma, \tau, t)) - \Gamma(K, J(\mu + \gamma + \lambda, \tau, t)))\gamma(\tau)d\tau \\
&\geq \inf(\gamma) \int_0^t (\Gamma(K, J(\mu + \gamma, \tau, t)) - \Gamma(K, J(\mu + \gamma + \lambda, \tau, t)))d\tau \\
&\overset{(25)}{\geq} \frac{\inf(\gamma)}{\theta} \int_0^t (\Gamma(K, J(\tilde{\mu} + \gamma, \tau, t)) - \Gamma(K, J(\tilde{\mu} + \gamma + \lambda, \tau, t)))d\tau \\
&\overset{(26)}{\geq} \frac{\inf(\gamma)}{\theta \sup(\gamma)} \int_0^t (\Gamma(K, J(\tilde{\mu} + \gamma, \tau, t)) - \Gamma(K, J(\tilde{\mu} + \gamma + \lambda, \tau, t)))\gamma(\tau)d\tau \\
&\overset{(27)}{\geq} \frac{1}{\theta \xi} \int_0^t (\Gamma(K, J(\tilde{\mu} + \gamma, \tau, t)) - \Gamma(K, J(\tilde{\mu} + \gamma + \lambda, \tau, t)))\gamma(\tau)d\tau \\
&= (1 - \check{\alpha})(\mathcal{P}_j(\tilde{\mathcal{E}}_\mathcal{B} \cup \{(i,j)\}) - \mathcal{P}_j(\tilde{\mathcal{E}}_\mathcal{B})),
\end{aligned}
$$

where Equation (25) follows from the Lemma and Equation (27) follows from the ξ-boundedness of γ. This result implies that $\check{\alpha}^*$, defined in Equations (13) or (15), is an upper bound on the inverse generalized curvature of $\mathcal{P}_j(\mathcal{E}_\mathcal{B})$, thereby, it is an upper bound bound on that of $\mathcal{U}(\mathcal{E}_\mathcal{B})$ as well.

C PROOF OF LEMMA 12

We start by subtracting the **left-hand side (LHS)** and the **right-hand side (RHS)** of Equation (22) and splitting the integration interval into subintervals $(0, \tau_0)$ and (τ_0, t) with $\tau_0 = \tau_0(\mu + \gamma, t)$ such that $J(\mu + \gamma, \tau_0, t) = K - 1$:

$$
LHS - RHS = \Delta(0, \tau_0) + \Delta(\tau_0, t),
$$

where

$$
\begin{aligned}
\Delta(a, b) = \theta & \left[\int_a^b \Gamma(K, J(\mu + \gamma, \tau, t))d\tau - \int_a^b \Gamma(K, J(\mu + \gamma + \lambda, \tau, t))d\tau \right] \\
& - \left[\int_a^b \Gamma(K, J(\tilde{\mu} + \gamma, \tau, t))d\tau - \int_a^b \Gamma(K, J(\tilde{\mu} + \gamma + \lambda, \tau, t))d\tau \right].
\end{aligned}
$$

For $\tau \leq \tau_0$, we have that $J(\mu + \gamma, \tau, t) \geq K - 1$ using the fact that $J(\mu + \gamma, \tau, t)$ is nonincreasing with respect to τ. Moreover, note that the intensity $\tilde{\mu}(t)$ is the summation of $\mu(t)$ with the intensity due to broadcasters $\tilde{\mathcal{E}}_\mathcal{B} \backslash \mathcal{E}_\mathcal{B}$. Therefore,

$$
J(\tilde{\mu} + \gamma, \tau, t) = \int_\tau^t \tilde{\mu}(x) + \gamma(x)dx \geq \int_\tau^t \mu(x) + \gamma(x)dx = J(\mu + \gamma, \tau, t) \geq K - 1.
$$

In a similar way, we can conclude that $J(\mu + \gamma + \lambda, \tau, t), J(\tilde{\mu} + \gamma + \lambda, \tau, t) \geq K - 1$. Then, using that the composite function $\Gamma(K, J(\mu + \gamma + \lambda, \tau, t))$ is convex in J for $\tau \leq \tau_0$ and, for any convex function f, $x \geq x'$, and $y \geq 0$, $f(x) - f(x + y) \leq f(x') - f(x' + y)$ (refer to Lemma 15), it follows that

$$
\Gamma(K, J(\mu + \gamma, \tau, t)) - \Gamma(K, J(\mu + \gamma + \lambda, \tau, t)) \geq \Gamma(K, J(\tilde{\mu} + \gamma, \tau, t)) - \Gamma(K, J(\tilde{\mu} + \gamma + \lambda, \tau, t)).
$$

Next, we can integrate the above equation and obtain that

$$\int_0^{\tau_0} \Gamma(K, J(\mu + \gamma, \tau, t))d\tau - \int_0^{\tau_0} \Gamma(K, J(\mu + \gamma + \lambda, \tau, t))d\tau$$

$$\geq \int_0^{\tau_0} \Gamma(K, J(\tilde{\mu} + \gamma, \tau, t))d\tau - \int_0^{\tau_0} \Gamma(K, J(\tilde{\mu} + \gamma + \lambda, \tau, t))d\tau,$$

which implies

$$\Delta(0, \tau_0) \geq (\theta - 1) \left(\int_0^{\tau_0} \Gamma(K, J(\mu + \gamma, \tau, t))d\tau - \int_0^{\tau_0} \Gamma(K, J(\mu + \gamma + \lambda, \tau, t))d\tau \right) \geq 0, \quad (28)$$

using that $\Gamma(K, x)$ is nonincreasing with respect to x and $\theta > 1$. Unfortunately, $\Delta(\tau_0, t)$ can be negative. However, in the following, we will show that $\Delta(\tau_0, t) \geq -\Delta(0, \tau_0)$.

Let $d = J(\lambda, \tau_0, t)$. First, note that, for $\tau \leq \tau_0$, $J(\lambda, \tau, t) \geq J(\lambda, \tau_0, t) = d$. Then, starting from Equation (28), we have that

$$\Delta(0, \tau_0) \geq (\theta - 1) \int_{J(\mu + \gamma, 0, t)}^{J(\mu + \gamma, \tau_0, t)} [\Gamma(K, J(\mu + \gamma, \tau, t)) - \Gamma(K, J(\mu + \gamma + \lambda \tau, t))] \frac{d\tau}{dJ(\mu + \gamma, \tau, t)} dJ(\mu + \gamma, \tau, t)$$

$$= (\theta - 1) \int_{K-1}^{J(\mu + \gamma, 0, t)} \frac{1}{\mu(\tau) + \gamma(\tau)} [\Gamma(K, J(\mu + \gamma, \tau, t)) - \Gamma(K, J(\mu + \gamma + \lambda, \tau, t))] dJ(\mu + \gamma, \tau, t)$$

$$\geq \frac{\theta - 1}{\sup(\mu + \gamma)} \int_{K-1}^{J(\mu + \gamma, 0, t)} [\Gamma(K, J(\mu + \gamma, \tau, t)) - \Gamma(K, J(\mu + \gamma, \tau, t) + d)] dJ(\mu + \gamma, \tau, t)$$

$$= \frac{\theta - 1}{\sup(\mu + \gamma)} \int_{K-1}^{J(\mu + \gamma, 0, t)} [\Gamma(K, x) - \Gamma(K, x + d)] dx$$

$$= \frac{\theta - 1}{\sup(\mu + \gamma)} \left[\int_{K-1}^{J(\mu + \gamma, 0, t)} \Gamma(K, x) dx - \int_{K-1+d}^{J(\mu + \gamma, 0, t) + d} \Gamma(K, x) dx \right]$$

$$= \frac{\theta - 1}{\sup(\mu + \gamma)} \left[\int_{K-1}^{K-1+d} \Gamma(K, x) dx - \int_{J_s^t(0)}^{J_s^t(0) + d} \Gamma(K, x) dx \right].$$

Next, we can use the first statement of Lemma 16 (refer to Appendix H) to bound the second integration term above and obtain that

$$\Delta(0, \tau_0) \geq \frac{\theta - 1}{\sup(s)} \left(1 - \frac{4e^{\frac{7}{4}}}{\zeta} \right) \int_{K-1}^{K-1+d} \Gamma(K, x) dx. \quad (29)$$

Second, note that, for $\tau \geq \tau_0$, $J(\lambda, \tau, t) \leq d$. Then, we have that

$$\Delta(\tau_0, t) \geq - \left(\int_{\tau_0}^t \Gamma(K, J(\tilde{\mu} + \gamma, \tau, t))d\tau - \int_{\tau_0}^t \Gamma(K, J(\tilde{\mu} + \gamma + \lambda, \tau, t))d\tau \right)$$

$$= - \int_{J(\tilde{\mu} + \gamma, \tau_0, t)}^{J(\tilde{\mu} + \gamma, t, t)} [\Gamma(K, J(\tilde{\mu} + \gamma, \tau, t)) - \Gamma(K, J(\tilde{\mu} + \gamma + \lambda, \tau, t))] \frac{d\tau}{dJ(\tilde{\mu} + \gamma, \tau, t)} dJ(\tilde{\mu} + \gamma, \tau, t)$$

$$= - \int_0^{J(\tilde{\mu} + \gamma, \tau_0, t)} \frac{1}{\tilde{\mu}(\tau) + \gamma(\tau)} [\Gamma(K, J(\tilde{\mu} + \gamma, \tau, t)) - \Gamma(K, J(\tilde{\mu} + \gamma + \lambda, \tau, t))] dJ(\tilde{\mu} + \gamma, \tau, t)$$

$$\geq - \frac{1}{\inf(\tilde{\mu} + \gamma)} \int_0^{J(\tilde{\mu} + \gamma, \tau_0, t)} (\Gamma(K, x) - \Gamma(K, x + d)) dx$$

$$\geq -\frac{1}{\inf(\tilde{\mu} + \gamma)} \int_0^\infty (\Gamma(K, x) - \Gamma(K, x + d)) dx$$

$$= -\frac{1}{\inf(\tilde{\mu} + \gamma)} \int_0^d \Gamma(K, x) dx \geq -\frac{\xi}{\sup(\tilde{\mu} + \gamma)} \int_0^d \Gamma(K, x) dx, \tag{30}$$

where the last inequality follows from ξ-boundedness of the user's feed intensity. Before we proceed further, note that, under the extra condition of Theorem 5, we can upper bound d as follows:

$$d = J(\lambda, \tau_0, t) \leq \int_{\tau_0}^t \lambda(x) dx \leq \int_{\tau_0}^t \frac{\gamma(x)}{\rho\sqrt{K-1}} dx \leq \int_{\tau_0}^t \frac{\mu(x) + \gamma(x)}{\rho\sqrt{K-1}} dx = \frac{1}{\rho\sqrt{K-1}} J(\mu + \gamma, \tau_0, t) = \frac{\sqrt{K-1}}{\rho},$$

and this enables us to use the third statement of Lemma 16.

Finally, combining Equations (29) and (30), where θ is given by either Equations (23) or (24) together with the second or third statement of Lemma 16, depending on whether we have the extra condition of Theorem 5, it follows that:

$$\Delta(0, \tau_0) \geq \frac{\theta - 1}{\sup(s)} \left(1 - \frac{4e^{\frac{7}{4}}}{\zeta}\right) \int_{K-1}^{K-1+d} \Gamma(K, x) dx \geq \frac{\xi}{\sup(\tilde{s})} \int_0^d \Gamma(K, x) dx \geq -\Delta(\tau_0, t),$$

which completes the proof of Lemma 12.

D   PROOF OF THEOREM 7

Throughout the proof, given a set function F, we define the marginal gain function ρ_Ω of each subset $\Omega \subseteq \mathcal{V}$ as $\rho_\Omega(\mathcal{S}) = F(\mathcal{S} \cup \Omega) - F(\mathcal{S}), \forall \mathcal{S} \subseteq \mathcal{V}$. Whenever $\Omega = \{v\}$ is a singleton, we use the symbol ρ_v instead of $\rho_{\{v\}}$ for simplicity.

Let $\mathcal{S}$ be the greedy solution, $\mathcal{T}$ be the optimal set maximizing F over $\mathcal{W}$. For $\leq i \leq n$ and $1 \leq j \leq |\mathcal{N}_i|$, let $\mathcal{S}_i = \mathcal{S} \cap (\mathcal{N}_1 \cup \cdots \cup \mathcal{N}_i)$, $\mathcal{T}_i = \mathcal{T} \cap (\mathcal{N}_1 \cup \cdots \cup \mathcal{N}_i)$, $\mathcal{U}_i^j$ be the first j elements in $\mathcal{N}_i$ that are considered in the greedy algorithm, and define $\mathcal{T}_i^j = \mathcal{U}_i^j \cap \mathcal{T}$, $\mathcal{S}_i^j = \mathcal{U}_i^j \cap \mathcal{S}$. Moreover, define $\{s_{ik}\}_{k=1}^{p_i}$ and $\{t_{ik}\}_{k=1}^{q_i}$ be the elements in $\mathcal{S}_i - \mathcal{S}_{i-1}$ and $\mathcal{T}_i - \mathcal{T}_{i-1}$, respectively, in order of their consideration in the algorithm.

First, we show that $|\mathcal{T}_i^j| \leq |\mathcal{S}_i^j|$. According to the definition of the greedy algorithm, adding any element from $\mathcal{U}_i^j \backslash \mathcal{S}_i^j$ to $\mathcal{S}_i^j$ violates the matroid constraint $\mathcal{I}_i$ (otherwise, that element should have been picked by the greedy algorithm). Thus, $\mathcal{U}_i^j \subseteq \text{span}(\mathcal{S}_i^j)$, which implies $\text{rank}(\mathcal{U}_i^j) = \text{rank}(\mathcal{S}_i^j)$. Moreover, $\mathcal{T}_i^j \subseteq \mathcal{U}_i^j$ implies $\text{rank}(\mathcal{T}_i^j) \leq \text{rank}(\mathcal{U}_i^j)$, therefore, $\text{rank}(\mathcal{T}_i^j) \leq \text{rank}(\mathcal{S}_i^j)$. However, $\mathcal{S}_i^j$ and $\mathcal{T}_i^j$ are both independent sets of matroid $\mathcal{I}_i$, because $\mathcal{S}$ and $\mathcal{T}$ are both feasible solutions. This implies $\text{rank}(\mathcal{T}_i^j) = |\mathcal{T}_i^j|$, $\text{rank}(\mathcal{S}_i^j) = |\mathcal{S}_i^j|$, proving our claim.

Then, our claim readily implies that, for each $1 \leq i \leq n$, $q_i \leq p_i$ and also, for each $1 \leq k \leq q_i$, s_{ik} is considered in the greedy algorithm at some point before t_{ik}. This means that, at the point that the greedy picks s_{ik}, t_{ik} does not have a higher marginal gain than s_{ik}, i.e.,

$$\rho_{s_{ik}}(Q_{ik}) \geq \rho_{t_{ik}}(Q_{ik}), \quad \forall 1 \leq i \leq n, \ 1 \leq k \leq q_i, \tag{31}$$

where Q_{ik} is the set of elements picked by the greedy until s_{ik} is picked (but not including s_{ik} itself). Hence, we can write

$$F(\mathcal{T}) \overset{(32)}{\leq} F(\mathcal{S} \cup \mathcal{T}) = F(\mathcal{S}) + \sum_{i=1}^{n} \sum_{k=1}^{q_i} \rho_{t_{ik}}(\mathcal{S} \cup \mathcal{T}_{i-1} \cup \{t_{i1}, \ldots, t_{i(k-1)}\})$$

$$\overset{(33)}{\leq} F(\mathcal{S}) + \sum_{i=1}^{n} \sum_{k=1}^{q_i} \frac{1}{1-\check{\alpha}} \rho_{t_{ik}}(Q_{ik}) \leq F(\mathcal{S}) + \frac{1}{1-\check{\alpha}} \sum_{i=1}^{n} \sum_{k=1}^{q_i} \rho_{s_{ik}}(Q_{ik})$$

$$\leq F(\mathcal{S}) + \frac{1}{1-\check{\alpha}} \sum_{i=1}^{n} \sum_{k=1}^{p_i} \rho_{s_{ik}}(Q_{ik}) = \left(1 + \frac{1}{1-\check{\alpha}}\right) F(\mathcal{S}),$$

where, in steps (32) and (33), we have used the monotonicity and $\check{\alpha}$-submodularity property of F, respectively.

E PROOF OF THEOREM 8

For each broadcaster $i \in \mathcal{B}$, we have to add its outgoing edge with the largest marginal gain and repeat this c_i times. However, picking an edge (i, j) only changes the marginal gain of the edges pointing towards user j, since $\mathcal{U}_K(\mathcal{E}_{\mathcal{B}}, j)$ are independent. Hence, we can use the order statistics algorithm to obtain the c_i largest marginal gains, with a computational cost $O(|\mathcal{V}|)$. Therefore, the overall algorithm attains time complexity $O(\kappa |\mathcal{B}||\mathcal{V}|)$.

F SAMPLE COMPLEXITY ANALYSIS

The following theorems provide error bounds for our empirical estimate of the average top K visibility, given by Equation (17):

THEOREM 13. *Suppose that, at the beginning of each realization l, there are at least K posts in user j's feed and the average number of posts published in user j's feed $P = \mathbb{E}[A_j^T(N(t_f) - N(t_0))] \geq e^2 + K$. Moreover, let*

$$\beta_j = \inf_{t \in (t_0, t_f)} \frac{\mu(t)}{\mu(t) + \gamma_{j \backslash \mathcal{B}}(t)} \quad \text{and} \quad \rho_j = \sup_{t \in (t_0, t_f)} \frac{\mu(t)}{\mu(t) + \gamma_{j \backslash \mathcal{B}}(t)},$$

where $\mu(t) = \sum_{i \in \mathcal{B}:(i,j) \in \mathcal{E}_{\mathcal{B}}} \mu_i(t)$ and all intensities be ξ-bounded. Then, it holds that

$$\mathbb{P}\{(1-\delta)\mathcal{U}(\mathcal{E}_{\mathcal{B}}, j) \leq \hat{\mathcal{U}}(\mathcal{E}_{\mathcal{B}}, j) \leq (1+\delta)\mathcal{U}(\mathcal{E}_{\mathcal{B}}, j)\} \geq 1 - e^{Z - Qn},$$

where

$$Q = \beta_j \tilde{c} \left(\sqrt{1 + \frac{\delta}{4\xi} c''} - 1 \right)^2 \quad \text{and} \quad Z = \ln(2 \max\{-\ln(0.1\delta\beta_j), e^2 P\} + 1) + \ln(K),$$

with

$$\epsilon = \min \left\{ \frac{\lfloor \frac{P-1}{2} \rfloor}{\ln(10) - \ln(\beta_j \delta)}, \frac{2}{5e^2} \right\}, \quad \tilde{c} = \max \left\{ 0.1\beta_j \delta, \frac{\epsilon}{2\xi} \min \left\{ 1, \frac{e-2}{8e} \frac{\delta}{\xi} \right\} \right\},$$

$$c'' = e^{-(e^2+2)}(e-2) \frac{\tilde{c}^{\frac{1}{P-1}}}{\max \left\{ \sqrt{\ln\left(\frac{1}{\tilde{c}}\right)}, \sqrt{(e^2+1)P} \right\}}.$$

THEOREM 14. *Suppose the conditions in Theorem 13 and let $0 \leq y \leq 1$. Then, it also holds that:*

$$\mathbb{P}\{\mathcal{U}(\mathcal{E}_{\mathcal{B}}, j) - yK(t_f - t_0) \leq \hat{\mathcal{U}}(\mathcal{E}_{\mathcal{B}}, j) \leq \mathcal{U}(\mathcal{E}_{\mathcal{B}}, j) + yK(t_f - t_0)\} \geq 1 - e^{Z_y - Q_y n},$$

where

$$Q_y = \tilde{c} \left(\sqrt{1 + \frac{y}{4\xi^2} c''} - 1 \right)^2 \quad \text{and} \quad Z_y = \ln(2 \max\{\ln(10\xi) - \ln y, e^2 P\} + 1) + \ln(K),$$

with

$$\epsilon = \min \left\{ \frac{\lfloor \frac{P-1}{2} \rfloor}{\ln(10\xi) - \ln y}, \frac{2}{5e^2} \right\}, \quad \tilde{c} = \max \left\{ 0.1 \frac{y}{\xi}, \frac{\epsilon}{2\xi} \min \left\{ 1, \frac{e-2}{8e} \frac{y}{\xi \rho_j} \right\} \right\},$$

$$c'' = e^{-(e^2+2)} (e-2) \frac{\tilde{c}^{\frac{1}{P-1}}}{\max \left\{ \sqrt{\ln\left(\frac{1}{\tilde{c}}\right)}, \sqrt{(e^2+1)P} \right\}}.$$

Before we proceed to prove the above two theorems, we introduce a set of definitions and notation. First, we rewrite our empirical estimate of the average top K visibility, given by Equation (17), as

$$\hat{\mathcal{U}}(\mathcal{E}_\mathcal{B}, j) = \sum_{i=1}^{K} \frac{\sum_{\ell=1}^{n} \Delta_{i,j}^{(\ell)}}{n} = \sum_{i=1}^{K} \Delta_{i,j},$$

where $\Delta_{i,j} = \frac{\sum_{\ell=1}^{n} \Delta_{i,j}^{(\ell)}}{n}$ and we define

$$\gamma_{min} = \inf_{t \in (t_0, t_f)} \mu(t) + \gamma_{j\setminus\mathcal{B}}(t) \quad \text{and} \quad \gamma_{max} = \sup_{t \in (t_0, t_f)} \mu(t) + \gamma_{j\setminus\mathcal{B}}(t).$$

Then, for $r \in \{1, 2, \ldots\}$, we introduce the following definitions:

(i) $P_{j,i,r}^{(\ell)}$ is the rth story that user j receives in the ith position of her feed in the ℓth realization.

(ii) $Y_{j,i,r}^{(\ell)}$ is the time that story $P_{j,i,r}^{(\ell)}$ remains in the ith position and $Y_{j,i,r} = \frac{\sum_{\ell=1}^{n} Y_{j,i,r}^{(\ell)}}{n}$.

(iii) $S_{j,i,r}^{(\ell)}$ is the amount of time between (t_0, t_f) that story $P_{j,i,r}^{(\ell)}$ remains in the ith position if $P_{j,i,r}^{(\ell)}$ has been shared by a broadcaster in $\mathcal{B}$ and zero otherwise, and $S_{j,i,r} = \frac{\sum_{\ell=1}^{n} S_{j,i,r}^{(\ell)}}{n}$.

(iv) $E_{j,i,r}^{(\ell)}$ indicates whether user j receives story $P_{j,i,r}^{(\ell)}$ sometime during (t_0, t_f), i.e., $E_{j,i,r}^{(\ell)} = 1$ if user j receives the rth story in the ith position of her feed during (t_0, t_f) and $E_{j,i,r}^{(\ell)} = 0$ otherwise.

(v) $W_{j,i,r}^{(\ell)}$ indicates whether story $P_{j,i,r}^{(\ell)}$ is shared by a broadcaster in $\mathcal{B}$, i.e., $W_{j,i,r}^{(\ell)} = 1$ if story $P_{j,i,r}^{(\ell)}$ is shared by a broadcaster in $\mathcal{B}$ and $W_{j,i,r}^{(\ell)} = 0$ otherwise.

Note that, given these definitions, it readily follows that

$$\Delta_{i,j}^{(\ell)} = \sum_{r=1}^{\infty} S_{j,i,r}^{(\ell)}, \quad \Delta_{i,j} = \sum_{r=1}^{\infty} S_{j,i,r}.$$

Moreover, we further define the conditional random variable $M_{j,i,r}^{(\ell)} \sim S_{j,i,r}^{(\ell)} \mid \{E_{j,i,r}^{(\ell)}, W_{j,i,r}^{(\ell)}\}$, the probabilities $q_{j,i,r} = \Pr\{E_{j,i,r}^{(\ell)}\}$ and $p_{j,i,r} = \Pr\{W_{j,i,r}^{(\ell)}|E_{j,i,r}^{(\ell)}\}$, and the average $\mu_{j,i,r} = \mathbb{E}[M_{j,i,r}^{(\ell)}]$. In the remainder, we fix the indices i, j and, for simplicity, we omit them from all variables, including β_j, ρ_j, and γ_j.

To prove Theorems 13 and 14, note that we have to bound the error for all S_i's simultaneously. To do so, we can think of resorting to a Chernoff bound, however, the challenge in doing so is that, as r grows, the expectation of S_r's decrease, as q_r's decrease drastically, and thus the Chernoff bound becomes weaker as r grows. To overcome this, we will partition the indices r into three part, defined by the following two critical indices:

(i) r_1 is the smallest index such that

$$q_{r_1} \leq c' = \frac{\beta}{2\rho} \epsilon \min \left\{ 1, \frac{e-2}{2e} \frac{\lambda_{\min}}{\lambda_{\max}} \frac{\delta}{4} \right\}, \tag{34}$$

where, according to Equation (64), $q_{\frac{P-1}{2}} \geq \frac{1}{2}$. Hence, since $c' < \frac{1}{2}$, we have that $r_1 > \frac{P-1}{2}$.

(ii) r_2 is the smallest index such that $q_{r_2} \leq 0.1\delta\beta$.

Then, for the first part, we will bound the corresponding S_r's using a Chernoff bound. For the second part, we will bound the corresponding S_r's with respect to the order of magnitudes of the expectation of the first part, again using a Chernoff bound. For the third part, we will use standard binomial distribution tail inequalities to bound their error independently. Now, we are ready to state the proofs of both theorems, which rely on several technical lemmas from Appendix H.

F.1 Theorem 13

We distinguish two cases:

— **Case $r_1 < r_2$.** In this case, using Lemmas 24, 25, and 27, with $\delta = \frac{\delta}{4}$ in Lemma 25, we obtain (note that S_r is actually $S_{j,i,r}$ that we have dropped the indexes j, i for simplicity of the notation):

$$\mathbb{P}\{\Delta_{i,j} \geq (1+\delta)\mathbb{E}[\Delta_{i,j}] \text{ or } \Delta_{i,j} \leq (1-\delta)\mathbb{E}[\Delta_{i,j}]\}$$

$$\leq \sum_{r=1}^{r_1-1} \mathbb{P}\left\{ S_r \leq \left(1 - \frac{\delta}{4}\right)\mathbb{E}[S_r] \text{ or } S_r \geq \left(1 + \frac{\delta}{4}\right)\mathbb{E}[S_r]\right\} + \mathbb{P}\left\{ \sum_{r=r_1}^{r_2-1} S_r \geq \frac{\delta}{2}\mathbb{E}[\Delta_{i,j}]\right\}$$

$$+ \mathbb{P}\left\{ \sum_{r=r_2}^{\infty} S_r \geq \frac{\delta}{4}\mathbb{E}[\Delta_{i,j}]\right\}$$

$$\leq (2r_2 - 1)\exp\left(-\beta c'\left(\sqrt{1 + \frac{\lambda_{\min}}{\lambda_{\max}}\frac{\delta}{4}c''} - 1\right)^2 n\right)$$

$$\leq \left(2\max\left\{\ln\left(\frac{1}{0.1\delta\beta}\right), e^2 P\right\} + 1\right)\exp\left(-\beta c'\left(\sqrt{1 + \frac{\lambda_{\min}}{\lambda_{\max}}\frac{\delta}{4}c''} - 1\right)^2 n\right), \tag{35}$$

where Equation (35) follows from Equation (78). In this case,

$$\tilde{c} = \max\left\{0.1\beta_j\delta, \frac{\epsilon}{2}\frac{\beta_j}{\rho_j}\min\left\{1, \frac{e-2}{8e}\frac{\delta}{\xi}\right\}\right\} \leq \max\left\{0.1\beta_j\delta, \frac{\epsilon}{2}\frac{\beta_j}{\rho_j}\min\left\{1, \frac{e-2}{8e}\frac{\delta\lambda_{\min}}{\lambda_{\max}}\right\}\right\}$$

$$= \frac{\epsilon}{2}\frac{\beta_j}{\rho_j}\min\left\{1, \frac{e-2}{8e}\frac{\delta\lambda_{\min}}{\lambda_{\max}}\right\} = c',$$

where we have used the assumption that $r_1 \leq r_2$ and the fact that $\frac{\lambda_{\min}}{\lambda_{\max}} \geq \frac{1}{\xi}$. Moreover, it can be easily checked that the condition $\tilde{c} \leq c'$ implies $c''' \leq c''$. Therefore, using the inequality given by fourth statement of Lemma 23, we can substitute the quantities c' and c'' in Equation (35), by $\tilde{c}$ and c''', respectively, which implies

$$\mathbb{P}\{\Delta_{i,j} \geq (1+\delta)\mathbb{E}[\Delta_{i,j}] \text{ or } \Delta_{i,j} \leq (1-\delta)\mathbb{E}[\Delta_{i,j}]\}$$

$$\leq \left(2\max\left\{\ln\left(\frac{1}{0.1\delta\beta}\right), e^2 P\right\} + 1\right)\exp\left(-\beta\tilde{c}\left(\sqrt{1 + \frac{\lambda_{\min}}{\lambda_{\max}}\frac{\delta}{4}c'''} - 1\right)^2 n\right), \tag{36}$$

where we have restored indexes i and j to emphasize the dependence of S over them. By summing up the above inequalities for $i = 1, \ldots, K$, we obtain the desired result.

— **Case** $r_1 \geq r_2$. In this case, there is no need in using index r_1. Therefore, for $1 \leq r \leq r_2$, we bound the error probability of each S_r using Theorem 22 and, for $r > r_2$, we bound the error probability of each S_r using Lemma 25.

First, note that for $r < r_2$ we have $q_r \geq 0.1\delta\beta$. Thus, according to Theorem 22, for $r < r_2$,

$$
\mathbb{P}\left\{ S_r \leq \left(1 - \frac{\delta}{4}\right)\mathbb{E}[S_r] \ \text{or} \ S_r \geq \left(1 + \frac{\delta}{4}\right)\mathbb{E}[S_r] \right\} \leq 2\exp\left(-0.1\beta^2\delta\left(\sqrt{1 + \frac{\lambda_{\min}}{\lambda_{\max}}\frac{\delta}{4}\bar{c}} - 1\right)^2 n\right).
$$

where $\bar{c} = e^{-(e^2+2)}(e-2)\dfrac{(0.1\delta\beta)^{\frac{1}{P-1}}}{\max\{\sqrt{\ln(\frac{1}{0.1\delta\beta})}, \sqrt{(e^2+1)P}\}}$. Second, note that it does not make sense to pick $\delta > \frac{2}{\beta}$, because according to Equation (67), this would imply that S exceeds its maximum possible value, which has probability zero. Thus, we can assume that $\delta \leq \frac{2}{\beta}$. Moreover, it is easy to check the following inequality for a real positive x:

$$
(\sqrt{1+x} - 1)^2 \leq x.
$$

Hence,

$$
0.1\beta^2\delta\left(\sqrt{1 + \frac{\lambda_{\min}}{\lambda_{\max}}\frac{\delta}{4}\bar{c}} - 1\right)^2 \leq 0.1\beta^2\delta\frac{\lambda_{\min}}{\lambda_{\max}}\frac{\delta}{4}\bar{c} \leq \frac{\delta}{8}\beta,
$$

which implies that

$$
e^{-\beta\frac{\delta}{8}n} \leq \exp\left(-0.1\beta^2\delta\left(\sqrt{1 + \frac{\lambda_{\min}}{\lambda_{\max}}\frac{\delta}{4}\bar{c}} - 1\right)^2 n\right).
$$

Thus,

$$
\mathbb{P}\{S \geq (1+\delta)\mathbb{E}[S] \ \text{or} \ S \leq (1-\delta)\mathbb{E}[S]\}
$$

$$
\leq \sum_{r=1}^{r_2-1} \mathbb{P}\left\{ S_r \leq \left(1 - \frac{\delta}{2}\right)\mathbb{E}[S_r] \ \text{or} \ S_r \geq \left(1 + \frac{\delta}{2}\right)\mathbb{E}[S_r] \right\} + \mathbb{P}\left\{ \sum_{r=r_2}^{\infty} S_r \geq \frac{\delta}{2}\mathbb{E}[S] \right\}
$$

$$
\leq (2r_2 - 1)\exp\left(-0.1\beta^2\delta\left(\sqrt{1 + \frac{\lambda_{\min}}{\lambda_{\max}}\frac{\delta}{4}\bar{c}} - 1\right)^2 n\right)
$$

$$
\leq \left(2\max\left\{\ln\left(\frac{1}{0.1\delta\beta}\right), e^2 P\right\} + 1\right)\exp\left(-0.1\beta^2\delta\left(\sqrt{1 + \frac{\lambda_{\min}}{\lambda_{\max}}\frac{\delta}{4}\bar{c}} - 1\right)^2 n\right). \tag{37}
$$

In addition,

$$
\tilde{c} = \max\left\{0.1\beta_j\delta, \ \frac{\epsilon}{2}\frac{\beta_j}{\rho_j}\min\left\{1, \frac{e-2}{8e}\frac{\delta}{\xi}\right\}\right\} \leq \max\left\{0.1\beta_j\delta, \ \frac{\epsilon}{2}\frac{\beta_j}{\rho_j}\min\left\{1, \frac{e-2}{8e}\frac{\delta\lambda_{\min}}{\lambda_{\max}}\right\}\right\}
$$

$$
= 0.1\beta_j\delta,
$$

which again implies that $c''' \leq \bar{c}$. Therefore, using the inequality given by fourth statement of Lemma 23, we can substitute the quantities $0.1\beta\delta$ and $\bar{c}$ in Equation (37) by $\tilde{c}$ and c''', respectively,

which again implies the inequality given by Equation (36). As in the previous case, summing over $i = 1, \ldots, K$ completes the proof.

F.2 Theorem 14

By setting $\delta = \frac{y}{\rho_j}$ in Theorem 13, $(1 - \delta)\mathcal{U}(\mathcal{E}_\mathcal{B}, j) \leq \hat{\mathcal{U}}(\mathcal{E}_\mathcal{B}, j) \leq (1 + \delta)\mathcal{U}(\mathcal{E}_\mathcal{B}, j)$ implies that $\mathcal{U}(\mathcal{E}_\mathcal{B}, j) - yK(t_f - t_0) \leq \hat{\mathcal{U}}(\mathcal{E}_\mathcal{B}, j) \leq \mathcal{U}(\mathcal{E}_\mathcal{B}, j) + yK(t_f - t_0)$ based on the inequality given by Equation (67). Substituting $\delta = \frac{y}{\rho_j}$ in the bounds of Theorem 13 and using the relation $\frac{\beta_j}{\rho_j} \leq \xi^2$ (because $\mu(t)$ and $\mu(t) + \gamma_{j \setminus \mathcal{B}}(t)$ are both ξ-bounded, their ratio can oscillate at most with ξ^2 order of magnitude), we can write

$$Q_y = Q = \beta_j \tilde{c} \left(\sqrt{1 + \frac{y}{4\xi\rho_j} c'''} - 1 \right)^2 = \tilde{c} \left(\sqrt{\beta_j + \frac{y\beta_j}{4\xi\rho_j} c'''} - \sqrt{\beta_j} \right)^2$$

$$\geq \tilde{c} \left(\sqrt{1 + \frac{y\beta_j}{4\xi\rho_j} c'''} - 1 \right)^2 \geq \tilde{c} \left(\sqrt{1 + \frac{y}{4\xi^3} c'''} - 1 \right)^2,$$

where we have used the fourth statement of Lemma 23. Substituting $\delta = \frac{y}{\rho_j}$ in the other equations of Theorem 13 completes the proof.

G PROOF OF THEOREM 9

We adapt the proof of Theorem 7. Define $S, S_i^j, S_i, s_{ik}, Q_{ik}$ with respect to the greedy with empirical estimates—the empirical greedy. Let $\mathcal{E}_\mathcal{B}^{(i)}$ be the edges picked by the empirical greedy before picking broadcaster i's edges. Let E be the event that $\forall i \in \mathcal{B}, \forall j : (i, j) \in \mathcal{W}_{i,:}$,

$$\left| \hat{\mathcal{U}} \left(\mathcal{E}_\mathcal{B}^{(i)}, j \right) - \mathcal{U} \left(\mathcal{E}_\mathcal{B}^{(i)}, j \right) \right| \leq yK(t_f - t_0). \tag{38}$$

Then, we can apply union bounds and Equation (18) to conclude that

$$\Pr\{E^o\} \leq \left(\sum_{i \in \mathcal{B}} |\mathcal{W}_{i,:}| \right) e^{Z_y - Q_y n} \leq \delta, \tag{39}$$

which implies that E happens with high probability. Moreover, $\forall i \in \mathcal{B}, 1 \leq k \leq c_i, s_{ik}$'s empirical marginal gain (when added to Q_{ik}) is not less than that of t_{ik}, due to the selection rule of the greedy algorithm. Consequently, given that E happens, by rewriting the empirical utilities in terms of the theoretical ones based on the inequality in Equation (79), we can rewrite Equation (31) as $\rho_{s_{ik}}(Q_{ik}) \geq \rho_{t_{ik}}(Q_{ik}) - 4yK(t_f - t_0)$. Then, we plug this expression into the proof of Theorem 7 and obtain:

$$F(\mathcal{T}) \leq F(\mathcal{S}) + \sum_{i \in \mathcal{B}} \sum_{k=1}^{c_i} \frac{1}{1 - \tilde{\alpha}} \rho_{t_{ik}}(Q_{ik})$$

$$\leq F(\mathcal{S}) + \frac{1}{1 - \tilde{\alpha}} \sum_{i \in \mathcal{B}} \sum_{k=1}^{c_i} (\rho_{s_{ik}}(Q_{ik}) + 4yK(t_f - t_0))$$

$$= \left(1 + \frac{1}{1 - \tilde{\alpha}} \right) F(\mathcal{S}) + \frac{4}{\zeta} \left(\sum_{i \in \mathcal{B}} c_i \right) yK(t_f - t_0). \tag{40}$$

To conclude, we divide both sides by $1 + \frac{1}{1 - \tilde{\alpha}}$ and use that $\frac{\frac{1}{\zeta}}{\frac{1}{\zeta} + 1} \leq 1$.

H   ADDITIONAL TECHNICAL LEMMAS

LEMMA 15.   *Let f be a smooth function that is concave over domain (a, b). Let x, y, z be real numbers such that $a \leq x, y, x + z, y + z \leq b$, $x \leq y$, and $z \geq 0$. Then $f(y + z) - f(x + z) \leq f(y) - f(x)$.*

PROOF.   We can write

$$f(y + z) - f(x + z) = \int_{x+z}^{y+z} f'(t)dt \leq \int_{x}^{y} f'(t)dt = f(y) - f(x),$$

where the inequality follows from the fact that f is concave in (a, b), therefore, it has negative derivative.   □

LEMMA 16.   *The following statements about the incomplete gamma function $\Gamma(K, x)$ hold: We have the following inequalities:*

1. *Let $d \geq 0$ and $\mathcal{K} \geq K - 1 + \zeta\sqrt{K - 1}$,*

$$\frac{4e^{\frac{7}{4}}}{\zeta} \int_{K-1}^{K-1+d} \Gamma(K, x)dx \geq \int_{\mathcal{K}}^{\mathcal{K}+d} \Gamma(K, x) \tag{41}$$

2. *Let $d \geq 0$,*

$$\frac{\int_{K}^{K-1+d} \Gamma(K, x)dx}{\int_{0}^{d} \Gamma(K, x)dx} \geq \frac{0.1625}{e^2 \sqrt{K}} \tag{42}$$

3. *Let $0 \leq d \leq \frac{\sqrt{K-1}}{\rho}$,*

$$\frac{\int_{K}^{K-1+d} \Gamma(K, x)dx}{\int_{0}^{d} \Gamma(K, x)dx} \geq \frac{1}{2e^{\frac{7}{4}}} \min\{\rho, 1\} e^{-\frac{1}{\rho^2}} \tag{43}$$

PROOF.   The proof of this lemma relies on four inequalities for the Γ function and its antiderivative F_K that are given by Lemmas 17–20 (refer to the end of this Section).

First, we proceed to prove the first statement of the Lemma. For $0 \leq u \leq d$, we can write

$$\frac{\Gamma(K, K' + u)}{\Gamma(K, K - 1 + u)} = e^{K-1-K'} \frac{\sum_{i=0}^{K-1} \frac{(K'+u)^i}{i!}}{\sum_{i=0}^{K-1} \frac{(K-1+u)^i}{i!}} \leq e^{K-1-K'} \left(\frac{K' + u}{K - 1 + u}\right)^{K-1} = e^{K-1-K'} \left(1 + \frac{K' - K + 1}{K - 1 + u}\right)^{K-1}$$

$$\leq e^{K-1-K'} \left(\left(1 + \frac{K' - K + 1}{K - 1 + u}\right)^{\frac{K-1+u}{K'-K+1}}\right)^{\frac{K-1}{K-1+u}(K'-K+1)}$$

$$\leq e^{K-1-K'} e^{-\frac{K-1}{K-1+u}(K-1-K')} = e^{\frac{u}{K-1+u}(K-1-K')}.$$

For $u \geq \sqrt{K - 1}$, we have $\frac{u}{K-1+u} \geq \frac{1}{2\sqrt{K-1}}$. Thus, for $u \geq \sqrt{K - 1}$,

$$\frac{\Gamma(K, K' + u)}{\Gamma(K, K - 1 + u)} \leq e^{\frac{1}{2\sqrt{K-1}}(K-1-K')} \leq e^{-\frac{\zeta}{2}} \leq \frac{4e^{\frac{7}{4}}}{\zeta}. \tag{44}$$

However, According to Lemmas 19 and 20, for $u \leq \sqrt{K - 1}$, we have

$$\frac{\Gamma(K, K' + u)}{\Gamma(K, K - 1 + u)} \leq \frac{\Gamma(K, K - 1 + \zeta\sqrt{K - 1})}{\Gamma(K, K - 1 + \sqrt{K - 1})} \leq \frac{\frac{2}{e\zeta}}{\frac{1}{2e^{\frac{11}{4}}}} = \frac{4e^{\frac{7}{4}}}{\zeta}. \tag{45}$$

Integrating over Equations (44) and (45) for $0 \leq u \leq d$ completes the proof of the first part of Lemma 16.

Next we prove the second part of Lemma 16. Define

$$G(d) = \frac{F_K(K-1+d) - F_K(K-1)}{F_K(d) - F_K(0)} = \frac{\int_K^{K-1+d} \Gamma(K,x)dx}{\int_0^d \Gamma(K,x)dx}. \tag{46}$$

If the function $g(d) = \frac{\Gamma(K,d)}{\Gamma(K,K-1+d)}$ is increasing in the interval $(0, d_0)$, then G is decreasing in $(0, d_0)$. Based on this fact, we prove that for $d \in (0, \frac{1}{e-1}(K-1))$, $g(d)$ is increasing. To this end, we calculate the derivative of $g(d)$:

$$g'(d) = \frac{-\frac{d^{K-1}e^{-d}}{(K-1)!}\Gamma(K, K-1+d) + \frac{(K-1+d)^{K-1}e^{-(K-1+d)}}{(K-1)!}\Gamma(K,d)}{\Gamma^2(K-1+d)}.$$

Hence, condition $g'(d) \geq 0$ is equivalent to

$$\frac{\Gamma(K, K-1+d)}{\Gamma(K,d)} \leq \left(\frac{\frac{K-1}{d}+1}{e}\right)^{K-1}.$$

However, note that for $d \in (0, \frac{1}{e-1}(K-1))$,

$$\left(\frac{\frac{K-1}{d}+1}{e}\right)^{K-1} \geq \left(\frac{\frac{K-1}{d}+1}{e}\right)^{K-1} \geq 1 \geq \frac{\Gamma(K, K-1+d)}{\Gamma(K,d)}.$$

Therefore, with the definition

$$\tilde{d} = \mathrm{argmin}_d G(d),$$

we conclude $\tilde{d} \geq \frac{1}{e-1}(K-1)$, which implies

$$G(d) \geq G(\tilde{d}) = \frac{F_K(K-1+\tilde{d}) - F_K(K-1)}{F_K(\tilde{d}) - F_K(0)} \geq \frac{F_K(K-1+\frac{1}{e-1}(K-1)) - F_K(K-1)}{F_K(\infty) - F_K(0)}$$

$$= \frac{F_K\left(\frac{e}{e-1}(K-1)\right) - F_K(K-1)}{K} = \frac{-F_K(K-1) - \left(-F_K\left(\frac{e}{e-1}(K-1)\right)\right)}{K} \tag{47}$$

$$\geq \frac{-0.325 F_K(K-1)}{K} \geq \frac{-0.325 \frac{\sqrt{K}}{2e^2}}{K} = \frac{0.1625}{e^2\sqrt{K}} = \Omega\left(\frac{1}{\sqrt{K}}\right), \tag{48}$$

where we have used Lemmas 18 and 17 in Equation (47) and Equation (48), respectively. This completes the proof for the second part.

To prove the last part, note that for $x \geq 0$, $\Gamma(K, x)$ is decreasing, and $0 \leq \frac{\Gamma(K,x)}{(K-1)!} \leq 1$. Thus,

$$G(d) = \frac{\int_{K-1}^{K-1+d} \frac{\Gamma(K,x)}{(K-1)!}dx}{\int_0^d \frac{\Gamma(K,x)dx}{(K-1)!}} \geq \frac{\int_{K-1}^{K-1+d} \frac{\Gamma(K,K-1+d)}{(K-1)!}dx}{\int_0^d \frac{\Gamma(K,0)}{(K-1)!}dx}$$

$$\geq \frac{d\Gamma(K, K-1+d)}{d} = \Gamma(K, K-1+d) \geq \frac{1}{2e^{\frac{7}{4}}}\min\{\rho, 1\}e^{-\frac{1}{\rho^2}},$$

where the last inequality follows from Lemma 20. The proof is complete. □

Lemma 17.

$$-\frac{F_K(K-1)}{(K-1)!} \geq \frac{\sqrt{K}+1}{2e^2}.$$

Proof.

$$-\frac{F_K(K-1)}{(K-1)!} = \sum_{k=0}^{K-1}(K-1-k)\frac{(K-1)^k e^{-(K-1)}}{k!} > \sum_{k=K-\sqrt{K}}^{K-1}(K-1-k)\frac{(K-1)^k e^{-(K-1)}}{k!}$$

$$= \sum_{k=K-\sqrt{K}}^{K-1}(K-1-k)\frac{(K-1)\ldots(k+1)}{(K-1)^{K-1-k}}\frac{(K-1)^{K-1}e^{-(K-1)}}{(K-1)!}$$

$$= \sum_{i=0}^{\sqrt{K}} i\frac{(K-1)\ldots(K-i)}{(K-1)^i}\frac{(K-1)^{K-1}e^{-(K-1)}}{(K-1)!}$$

$$= \frac{(K-1)^{K-1}e^{-(K-1)}}{(K-1)!}\sum_{i=0}^{\sqrt{K}}i\frac{(K-1)\ldots(K-i)}{(K-1)^i} \geq \frac{(K-1)^{K-1}e^{-(K-1)}}{(K-1)!}\sum_{i=0}^{\sqrt{K}}i\frac{(K-i)^i}{(K-1)^i}$$

$$= \frac{(K-1)^{K-1}e^{-(K-1)}}{(K-1)!}\sum_{i=0}^{\sqrt{K}}i\left(1-\frac{i-1}{K-1}\right)^i = \frac{(K-1)^{K-1}e^{-(K-1)}}{(K-1)!}\sum_{i=0}^{\sqrt{K}}i\left[\left(1-\frac{(i-1)}{K-1}\right)^{\frac{K-1}{i-1}-1}\right]^{\frac{i(i-1)}{K-i}}$$

$$\geq \frac{(K-1)^{K-1}e^{-(K-1)}}{(K-1)!}\sum_{i=0}^{\sqrt{K}}ie^{-\frac{i(i-1)}{K-i}} \geq \frac{(K-1)^{K-1}e^{-(K-1)}}{(K-1)!}\sum_{i=1}^{\sqrt{K}}i$$

$$\geq \frac{(K-1)^{K-1}e^{-(K-1)}}{e^2(K-1)^{(K-1)+\frac{1}{2}}e^{-(K-1)}}\frac{\sqrt{K}(\sqrt{K}+1)}{2} \geq \frac{\sqrt{K}+1}{2e^2}. \qquad \square$$

Lemma 18.

$$F_K(K-1) \geq 1.49 F_K\left(\frac{e}{e-1}(K-1)\right).$$

Proof. For the right-hand side, we can write

$$-\frac{F_K\left(\frac{e}{e-1}(K-1)\right)}{(K-1)!} = \sum_{i=0}^{K-1}(K-i)\frac{\left(\frac{e}{e-1}(K-1)\right)^i e^{-\frac{e}{e-1}(K-1)}}{i!}$$

$$\leq \sum_{i=0}^{K-1}(K-i)(K-1)^{K-1-i}\frac{\left(\frac{e}{e-1}(K-1)\right)^i e^{-\frac{e}{e-1}(K-1)}}{(K-1)!}$$

$$\leq \sum_{i=0}^{K-1}(K-i)(K-1)^{K-1-i}\frac{\left(\frac{e}{e-1}(K-1)\right)^i e^{-\frac{e}{e-1}(K-1)}}{\sqrt{2\pi}(K-1)^{K-1+\frac{1}{2}}e^{-(K-1)}}$$

$$= \sum_{i=0}^{K-1}(K-i)\frac{\left(\frac{e}{e-1}\right)^i e^{-\frac{1}{e-1}(K-1)}}{\sqrt{2\pi(K-1)}} = \sum_{i=0}^{K-1}\frac{K-i}{\left(\frac{e}{e-1}\right)^{K-1-i}}\frac{\left(\frac{e}{e-1}\right)^{K-1}e^{-\frac{1}{e-1}(K-1)}}{\sqrt{2\pi(K-1)}}$$

$$= \frac{\left(\frac{e}{e-1}\right)^{K-1}e^{-\frac{1}{e-1}(K-1)}}{\sqrt{2\pi(K-1)}}\sum_{i=0}^{K-1}(K-i)\left(\frac{e-1}{e}\right)^{K-1-i}$$

$$= \frac{\left(\frac{e}{e-1}\right)^{K-1}e^{-\frac{1}{e-1}(K-1)}}{\sqrt{2\pi(K-1)}}\left[\frac{K}{\frac{e-1}{e}-1}\left(\frac{e-1}{e}\right)^K + \frac{1-\left(\frac{e-1}{e}\right)^K}{\left(\frac{e-1}{e}-1\right)^2}\right]$$

$$\leq \frac{\left(\frac{e}{e-1}\right)^{K-1}e^{-\frac{1}{e-1}(K-1)}}{\sqrt{2\pi(K-1)}}e^2 = \frac{e^{\left(\ln\left(\frac{e}{e-1}\right)-\frac{1}{e-1}\right)(K-1)}}{\sqrt{2\pi(K-1)}}e^2 \leq \frac{e^{-0.12(K-1)}}{\sqrt{2\pi(K-1)}}e^2.$$

For $K \leq 20$, one can check the inequality using a computer program. For $K > 20$, according to Lemma 17,

$$\frac{F_K(K-1)}{(K-1)!} \geq \frac{\sqrt{K}+1}{2e^2} \geq \frac{\sqrt{K}}{2e^2} \geq 1.49\frac{e^{-0.12(K-1)}}{\sqrt{2\pi(K-1)}}e^2 \geq 1.49\frac{F_K\left(\frac{e}{1-e}(K-1)\right)}{(K-1)!},$$

where the third inequality can be checked to be true for $K > 20$. Thus, the proof is complete. □

LEMMA 19.

$$\frac{\Gamma(K, K-1+\omega\sqrt{K-1})}{(K-1)!} \leq \frac{2}{e}w^{-1}.$$

PROOF.

$$\frac{\Gamma(K, K-1+\omega\sqrt{K-1})}{(K-1)!} = \sum_{i=0}^{K-1}\frac{(K-1+\omega\sqrt{K-1})^i e^{-(K-1+\omega\sqrt{K-1})}}{i!}$$

$$\leq \sum_{i=0}^{K-1}(K-1)^{K-1-i}\frac{(K-1+\omega\sqrt{K-1})^i e^{-(K-1+\omega\sqrt{K-1})}}{(K-1)!}$$

$$\leq \sum_{i=0}^{K-1}(K-1)^{K-1-i}\frac{(K-1+\omega\sqrt{K-1})^i e^{-(K-1+\omega\sqrt{K-1})}}{e(K-1)^{K-1+\frac{1}{2}}e^{-(K-1)}}$$

$$= \frac{e^{-\omega\sqrt{K-1}}}{e\sqrt{K-1}}\sum_{i=0}^{K-1}\left(\frac{K-1+\omega\sqrt{K-1}}{K-1}\right)^i$$

$$= \frac{e^{-\omega\sqrt{K-1}}}{e\sqrt{K-1}}\sum_{i=0}^{K-1}\left(1+\frac{\omega}{\sqrt{K-1}}\right)^i$$

$$= \frac{e^{-\omega\sqrt{K-1}-1}}{\sqrt{K-1}}\frac{\left(\frac{\omega}{\sqrt{K-1}}+1\right)^K - 1}{\frac{\omega}{\sqrt{K-1}}}$$

$$\leq \frac{e^{-\omega\sqrt{K-1}-1}}{\sqrt{K-1}}\frac{\left(\frac{\omega}{\sqrt{K-1}}+1\right)^{K-1}\left(\frac{\omega}{\sqrt{K-1}}+1\right)}{\frac{\omega}{\sqrt{K-1}}}$$

$$= \frac{e^{-\omega\sqrt{K-1}-1}}{\sqrt{K-1}}\frac{\left[\left(\frac{\omega}{\sqrt{K-1}}+1\right)^{\frac{\sqrt{K-1}}{\omega}}\right]^{w\sqrt{K-1}}\left(\frac{\omega}{\sqrt{K-1}}+1\right)}{\frac{\omega}{\sqrt{K-1}}}$$

$$\leq e^{-\omega\sqrt{K-1}-1}e^{w\sqrt{K-1}}\left(\frac{\omega}{\sqrt{K-1}}+1\right)w^{-1} = \frac{1}{ew}\left(\frac{\omega}{\sqrt{K-1}}+1\right) \leq \frac{2}{e}w^{-1}.$$

□

LEMMA 20. For $t = \omega\sqrt{K-1}$,

$$\frac{\Gamma(K, K-1+t)}{(K-1)!} \geq \frac{1}{2e^{\frac{7}{4}}}\min\left\{\frac{1}{\omega}, 1\right\}e^{-\omega^2}.$$

PROOF. Define $s = \max\{2, 2\omega\}$. Then,

$$\frac{\Gamma(K, K-1+t)}{(K-1)!} = \sum_{k=0}^{K-1} \frac{(K-1+t)^k e^{-(K-1+t)}}{k!} \geq \sum_{k=K-\frac{\sqrt{K}}{s}}^{K-1} \frac{(K-1+t)^k e^{-(K-1+t)}}{k!}$$

$$= \sum_{i=0}^{\frac{\sqrt{K}}{s}} \frac{(K-1+t)^{K-1-i} e^{-(K-1+t)}}{(K-1-i)!} = \sum_{i=0}^{\frac{\sqrt{K}}{s}} \frac{(K-1)...(K-i)}{(K-1+t)^i} \frac{(K-1+t)^{K-1} e^{-(K-1+t)}}{(K-1)!}$$

$$\geq \frac{(K-1+t)^{K-1} e^{-(K-1+t)}}{(K-1)!} \sum_{i=0}^{\frac{\sqrt{K}}{s}} \left(\frac{K-i}{K-1+t}\right)^i$$

$$= \frac{(K-1+t)^{K-1} e^{-(K-1+t)}}{(K-1)!} \sum_{i=0}^{\frac{\sqrt{K}}{s}} \left(1 - \frac{i+t-1}{K-1+t}\right)^i$$

$$= \frac{(K-1+t)^{K-1} e^{-(K-1+t)}}{(K-1)!} \sum_{i=0}^{\frac{\sqrt{K}}{s}} \left(\left(1 - \frac{i+t-1}{K-1+t}\right)^{\frac{K-1+t}{i+t-1}+1}\right)^{\frac{i(i+t-1)}{k+2t+i-2}}$$

$$\geq \frac{(K-1+t)^{K-1} e^{-(K-1+t)}}{(K-1)!} \sum_{i=0}^{\frac{\sqrt{K}}{s}} e^{-\frac{i(i+t-1)}{k+2t+i-2}}.$$

But according to the choice of s, we have

$$i(i+t-1) \leq \frac{\sqrt{K}}{s}\left(\frac{\sqrt{K}}{s} + \omega\sqrt{K} - 1\right) = \frac{K}{s^2} + \frac{\omega K}{s} - \frac{\sqrt{K}}{s}.$$

For the case $\omega \geq 1$, we have $s = 2\omega$,

$$\frac{K}{s^2} + \frac{\omega K}{s} - \frac{\sqrt{K}}{s} = \frac{K}{4\omega^2} + \frac{K}{2} - \frac{\sqrt{K}}{2\omega} \leq \frac{3}{4}K \leq \frac{3}{4}(K+2t+i-2).$$

In the other case where $\omega \leq 1$, we have $s = 2$,

$$\frac{K}{s^2} + \frac{\omega K}{s} - \frac{\sqrt{K}}{s} = \frac{K}{4} + \frac{\omega K}{2} - \frac{\sqrt{K}}{2} \leq \frac{3}{4}K \leq \frac{3}{4}(K+2t+i-2).$$

Hence, in both cases, we obtain

$$\frac{i(i+t-1)}{k+2t+i-2} \leq \frac{3}{4}.$$

Therefore,

$$\frac{\Gamma(K, K-1+t)}{(K-1)!} \geq \frac{(K-1+t)^{K-1} e^{-(K-1+t)}}{(K-1)!} \sum_{i=0}^{\frac{\sqrt{K}}{s}} e^{-\frac{3}{4}} \geq \frac{(K-1+t)^{K-1} e^{-(K-1+t)}}{e\sqrt{K-1}(K-1)^{K-1} e^{-(K-1)}} \frac{\sqrt{K}}{s} e^{-\frac{3}{4}}$$

$$\geq \frac{1}{s}\left(1 + \frac{t}{K-1}\right)^{K-1} e^{-t} e^{-\frac{7}{4}} = \frac{1}{s}\left(\left(1 + \frac{t}{K-1}\right)^{\frac{K-1}{t}+1}\right)^{\frac{t(K-1)}{t+K-1}} e^{-t} e^{-\frac{7}{4}}$$

$$\geq \frac{1}{s} e^{\frac{t(K-1)}{t+K-1}} e^{-t} e^{-\frac{7}{4}} = \frac{e^{-\frac{7}{4}}}{s} e^{-\frac{t^2}{t+K-1}} \geq e^{-\frac{7}{4}} \frac{1}{2} \min\left\{1, \frac{1}{\omega}\right\} e^{-\frac{t^2}{K-1}} = \frac{1}{2e^{\frac{7}{4}}} \min\left\{1, \frac{1}{\omega}\right\} e^{-\omega^2}.$$

$\square$

LEMMA 21. *Under the definitions in Section F, let $P = \int_{t_0}^{t_f} \mu(t) + \gamma_{j \setminus \mathcal{B}}(t) dt$, $f_{M_r^{(\ell)}}$ and $F_{M_r^{(\ell)}}$ be the PDF and CDF functions of $M_r^{(\ell)}$, respectively, and $\tilde{F}_{M_r^{(\ell)}} = 1 - F_{M_r^{(\ell)}}$. Then, $\forall r \geq 1$ and $0 \leq s \leq \lambda_{min}$, the following inequalities hold:*

$$\beta \leq p_r \leq \rho, \tag{49}$$

$$\tilde{F}_{M_r^{(\ell)}}(t) \leq e^{-\lambda_{min} t}, \tag{50}$$

$$\frac{1}{\lambda_{min}} \geq \mu_r \geq \frac{e^{-\frac{i+r-1}{P-1}}}{\sqrt{i+r-1}} \frac{\frac{e-2}{e}}{\lambda_{max}}, \tag{51}$$

$$\mathbb{E}[e^{s M_r^{(\ell)}}] - 1 - s\mu_r \leq \frac{s^2}{(\lambda_{min} - s)\lambda_{min}}, \tag{52}$$

$$\mathbb{E}[e^{-s M_r^{(\ell)}}] - 1 + s\mu_r \leq \frac{s^2}{(\lambda_{min} + s)\lambda_{min}}. \tag{53}$$

PROOF. We prove each statement in turn.

— **First inequality (Equation (49)).** First, we condition $W_r^{(\ell)}$ on the time when $P_r^{(\ell)}$ is received.

$$p_r = \mathbb{P}\{W_r^{(\ell)} \mid E_r^{(\ell)}\} = \int_{t_0}^{t_f} \mathbb{P}\{W_r^{(\ell)} \mid P_r^{(\ell)} \text{ is received at time t}\} \mathbb{P}\{P_r^{(\ell)} \text{ is received at time t} \mid E_r^{(\ell)}\} \, dt$$

$$= \int_{t_0}^{t_f} \frac{\mu(t)}{\mu(t) + \gamma_{j \setminus \mathcal{B}}(t)} \mathbb{P}\{P_r^{(\ell)} \text{ is received at time t} \mid E_r^{(\ell)}\} \, dt.$$

Now the desired inequality follows from the fact that $\beta_j \leq \frac{\mu(t)}{\mu(t) + \gamma_{j \setminus \mathcal{B}}(t)} \leq \rho_j$.

— **Second inequality (Equation (50)).** First, we condition $M_r^{(\ell)}$ on the time when $P_r^{(\ell)}$ is received. Thus, for $t_0 \leq t \leq t_f$, we have

$$M_r^{(\ell)} \mid \{P_r^{(\ell)} \text{ is received at time t}\}$$

$$\sim S_r^{(\ell)} \mid \{P_r^{(\ell)} \text{ is received at time t} , E_r^{(\ell)} , W_r^{(\ell)}\}$$

$$\sim S_r^{(\ell)} \mid \{P_r^{(\ell)} \text{ is received at time t}\} = \tilde{E}_t \leq E_t,$$

where $\tilde{E}_t = \min\{E_t, T_2 - t\}$, E_t is an exponential-like random variable with CDF function

$$F_{E_t}(\tau) = 1 - e^{-\int_0^\tau \gamma(t+x) dx},$$

and the reason why we have an inequality is that we have a limit time t_f and $P_r^{(\ell)}$ cannot occur after t_f. Therefore, the random variable $S_r^{(\ell)} \mid \{P_r^{(\ell)} \text{ is received at time t}\}$ is bounded from above by E_t. Moreover, we have that

$$\tilde{F}_{E_t}(\tau) = e^{-\int_0^\tau \gamma(t+x) dx} \leq e^{-t\lambda_{min}},$$

which implies that

$$\mathbb{P}\left\{M_r^{(\ell)} \geq t \mid P_r^{(\ell)} \text{ is received at time t}\right\} \leq e^{-t\lambda_{min}}.$$

Then, for $M_r^{(\ell)}$

$$\mathbb{P}\left\{M_r^{(\ell)} \geq t\right\} = \int_{t_0}^{t_f} \mathbb{P}\left\{M_r^{(\ell)} \geq t \mid P_r^{(\ell)} \text{ is received at time t}\right\} \mathbb{P}\left\{P_r^{(\ell)} \text{ is received at time t} \mid E_r^{(\ell)}\right\} dt$$

$$\leq \int_{T_1}^{T_2} e^{-t\lambda_{min}} \mathbb{P}\left\{P_r^{(\ell)} \text{ is received at time t} \mid E_r^{(\ell)}\right\} dt = e^{-t\lambda_{min}},$$

which completes the proof.

— **Third inequality (Equation (51)).** We first prove the upper bound as follows:

$$\mu_i = \mathbb{E}[M_r^{(\ell)}] = \int_{t_0}^{t_f} \mathbb{E}[M_r^{(\ell)} \mid P_r^{(\ell)} \text{ is received at t}]\mathbb{P}\{P_r^{(\ell)} \text{ is received at t} \mid E_r^{(\ell)}\}dt$$

$$\leq \int_{t_0}^{t_f} \frac{1}{\lambda_{\min}}\mathbb{P}\{P_r^{(\ell)} \text{ is received at t} \mid E_r^{(\ell)}\}dt = \frac{1}{\lambda_{\min}}.$$

Then, we proceed to prove the lower bound. Define $\mathcal{P}_x$ to be a Poisson random variable with parameter x. According to the definition of q_r'

$$q_r = \mathbb{P}\{\text{Receiving at least } i + r - 1 \text{ posts before time } t_f\},$$

which corresponds to the tail distributions of $\mathcal{P}_P$ ($\mathbb{P}\{\mathcal{P}_P \geq i + r - 1\}$). Moreover, define

$$q_r' = \mathbb{P}\left\{\text{Receiving at least } i + r - 1 \text{ posts before time } t_f - \frac{1}{\lambda_{\max}}\right\}.$$

Thus, if we define

$$P' = \int_{t_0}^{t_f - \frac{1}{\lambda_{\max}}} \lambda(x)dx,$$

then we get

$$q_r' = \mathbb{P}\{\mathcal{P}_{P'} \geq i + r - 1\}.$$

Now we claim that

$$\mu_r \geq \frac{q_r'}{q_r} \frac{\frac{e-2}{e}}{\lambda_{\max}}. \tag{54}$$

To prove this claim, we proceed as follows: First, we write

$$q_r\mu_r = q_r\,\mathbb{E}\left[M_r^{(\ell)}\right] = \mathbb{E}\left[\mathbf{1}\left\{E_r^{(\ell)}\right\}M_r^{(\ell)}\right] \geq \mathbb{E}\left[\mathbf{1}\left\{E_r^{(\ell)}, P_r^{(\ell)} \text{ is received before } t_f - \frac{1}{\lambda_{\max}}\right\}M_r^{(\ell)}\right]$$

$$= \mathbb{P}\left\{E_r^{(\ell)}, P_r^{(\ell)} \text{ is received before } t_f - \frac{1}{\lambda_{\max}}\right\}\mathbb{E}\left[M_r^{(\ell)} \mid E_r^{(\ell)}, P_r^{(\ell)} \text{ is received before } t_f - \frac{1}{\lambda_{\max}}\right] = q_r'\,\mathbb{E}[\widetilde{M}], \tag{55}$$

where we have named the last conditioned random variable as $\widetilde{M}$. Then, we again use the idea of conditioning on the time when $P_r^{(\ell)}$ is received:

$$\widetilde{M} \mid \left\{P_r^{(\ell)} \text{ is received at time t before } t_f - \frac{1}{\lambda_{\max}}\right\} = S_r^{(\ell)} \mid \left\{P_r^{(\ell)} \text{ is received at time t before } t_f - \frac{1}{\lambda_{\max}}\right\} = \tilde{E}_t,$$

where

$$\tilde{E}_t = \min\{E_t, T_2 - t\} \geq \min\left\{E_t, \frac{1}{\lambda_{\max}}\right\}.$$

Therefore, it follows that

$$\mathbb{E}\left[\widetilde{M} \mid \left\{P_r^{(\ell)} \text{ is received at time t before } t_f - \frac{1}{\lambda_{\max}}\right\}\right] \geq \mathbb{E}[\min\{E_t, \frac{1}{\lambda_{\max}}\}]$$

$$\geq \mathbb{E}\left[\mathbf{1}\left\{E_t \leq \frac{1}{\lambda_{\max}}\right\}E_t\right] = \int_0^{\frac{1}{\lambda_{\max}}} \lambda(\tau)e^{-\int_0^\tau \lambda(x)dx}\tau d\tau \geq \int_0^{\frac{1}{\lambda_{\max}}} \lambda_{\max}e^{-\tau\lambda_{\max}}\tau d\tau$$

$$= \left[-\tau e^{\tau\lambda_{\max}}\right]_0^{\frac{1}{\lambda_{\max}}} + \int_0^{\frac{1}{\lambda_{\max}}} e^{-\tau\lambda_{\max}}d\tau = -\frac{1}{\lambda_{\max}}e^{-1} + \frac{1}{\lambda_{\max}} - \frac{1}{\lambda_{\max}}e^{-1} = \frac{e-2}{e}\frac{1}{\lambda_{\max}}.$$

Hence,

$$
\begin{aligned}
\mathbb{E}[\widetilde{M}] &= \int_{t_0}^{t_f - \frac{1}{\lambda_{\max}}} \mathbb{E}\left[\widetilde{M} \mid P_r^{(\ell)} \text{ is received at t}\right] \mathbb{P}\left\{P_r^{(\ell)} \text{ is received at t} \mid P_r^{(\ell)} \text{ is received before } t_f - \frac{1}{\lambda_{\max}}\right\} dt \\
&\geq \int_{t_0}^{t_f - \frac{1}{\lambda_{\max}}} \frac{e-2}{e} \frac{1}{\lambda_{\max}} \mathbb{P}\left\{P_r^{(\ell)} \text{ is received at t} \mid P_r^{(\ell)} \text{ is received before } t_f - \frac{1}{\lambda_{\max}}\right\} dt = \frac{e-2}{e} \frac{1}{\lambda_{\max}}.
\end{aligned}
$$

(56)

Finally, if we apply Equation (56) to Equation (55), then we obtain the inequality given by Equation (54). Next, we aim to prove that

$$
\frac{q_r'}{q_r} \geq \frac{e^{-\frac{i+r-1}{P-1}}}{\sqrt{i+r-1}}.
$$

(57)

To this aim, we first note that

$$
\int_{t_0 - \frac{1}{\lambda_{\max}}}^{t_f} \lambda(x) dx \leq \int_{t_0 - \frac{1}{\lambda_{\max}}}^{t_f} \lambda_{\max} = 1,
$$

which implies $P' \geq P - 1$ and thus

$$
q_r' \geq \mathbb{P}\{\mathcal{P}_{P-1} \geq i + r - 1\}.
$$

Moreover, we have that

$$
\mathbb{P}\{\mathcal{P}_{P-1} \geq i + r - 1\} \geq \mathbb{P}\{\mathcal{P}_{P-1} = i + r - 1\} = \frac{(P-1)^{i+r-1} e^{-(P-1)}}{(i+r-1)!},
$$

and, according to the standard Chernoff bound for the Poisson tail,

$$
q_i = \mathbb{P}\{\mathcal{P}_P \geq i + r - 1\} \leq \frac{e^{-P}(eP)^{i+r-1}}{(i+r-1)^{i+r-1}},
$$

Thus,

$$
\frac{q_r'}{q_r} \geq \frac{\frac{(P-1)^{i+r-1} e^{-(P-1)}}{(i+r-1)!}}{\frac{e^{-P}(eP)^{i+r-1}}{(i+r-1)^{i+r-1}}} \geq \frac{\frac{(P-1)^{i+r-1} e^{-(P-1)}}{e\sqrt{i+r-1}(i+r-1)^{i+r-1} e^{-(i+r-1)}}}{\frac{e^{-P}(eP)^{i+r-1}}{(i+r-1)^{i+r-1}}} = \frac{(1-\frac{1}{P})^{i+r-1}}{\sqrt{i+r-1}} = \frac{\left(\left(1-\frac{1}{P}\right)^{P-1}\right)^{\frac{i+r-1}{P-1}}}{\sqrt{i+r-1}} \geq \frac{e^{-\frac{i+r-1}{P-1}}}{\sqrt{i+r-1}},
$$

which proves Equation (57). Finally, we can combine Equations (54) and (57) to obtain the desired lower bound, given by Equation (50).

— **Fourth inequality (Equation (52)).** By definition, we have

$$\mathbb{E}\left[e^{sM_r^{(\ell)}}\right] - 1 = \int_0^\infty f_{M_r^{(\ell)}}(t)e^{st}dt - 1 = \int_0^\infty f_{M_r^{(\ell)}}(t)(e^{st} - 1)dt$$
$$= \left[-\tilde{F}_{M_r^{(\ell)}}(t)(e^{st} - 1)\right]_0^\infty + \int_0^\infty s\tilde{F}_{M_r^{(\ell)}}(t)e^{st}dt = \int_0^\infty s\tilde{F}_{M_r^{(\ell)}}(t)e^{st}dt.$$

Note that by integration by parts,

$$\mu_r = \mathbb{E}\left[M_r^{(\ell)}\right] = \int_0^\infty tf_{M_r^{(\ell)}}(t)dt = \left[-t\tilde{F}_{M_r^{(\ell)}}(t)\right]_0^\infty - \int_0^\infty -\tilde{F}_{M_r^{(\ell)}}dt = \int_0^\infty \tilde{F}_{M_r^{(\ell)}}dt.$$

Therefore,

$$\mathbb{E}\left[e^{sM_r^{(\ell)}}\right] - 1 - s\mu_i = \int_0^\infty s\tilde{F}_{M_r^{(\ell)}}(t)e^{st}dt - s\int_0^\infty \tilde{F}_{M_r^{(\ell)}}(t)dt = \int_0^\infty s\tilde{F}_{M_r^{(\ell)}}(t)(e^{st} - 1)dt.$$

But note that $e^{st} - 1$ is positive in interval $(0, \infty)$. Therefore, according to Equation (50), it follows that

$$\mathbb{E}\left[e^{sM_r^{(\ell)}}\right] - 1 - s\mu_i \leq \int_0^\infty se^{-\lambda_{\min}t}(e^{st} - 1) = s\left(\int_0^\infty e^{-(\lambda_{\min}-s)t} - \int_0^\infty e^{-\lambda_{\min}t}v\right)$$
$$= s\left(\frac{1}{\lambda_{\min} - s} - \frac{1}{\lambda_{\min}}\right) = \frac{s^2}{(\lambda_{\min} - s)\lambda_{\min}}.$$

— **Fifth inequality (Equation (53)).** By definition, we have

$$\mathbb{E}\left[e^{-sM_r^{(\ell)}}\right] - 1 = \int_0^\infty f_{M_r^{(\ell)}}(t)e^{-st}dt - 1 = \int_0^\infty f_{M_r^{(\ell)}}(t)(e^{-st} - 1)dt$$
$$= \left[\tilde{F}_{M_r^{(\ell)}}(t)(e^{-st} - 1)\right]_0^\infty - \int_0^\infty s\tilde{F}_{M_r^{(\ell)}}(t)e^{-st}dt = -\int_0^\infty s\tilde{F}_{M_r^{(\ell)}}(t)e^{-st}dt.$$

Therefore, similar to the proof of the previous statement,

$$\mathbb{E}\left[e^{-sM_r^{(\ell)}}\right] - 1 + s\mu_i = -\int_0^\infty s\tilde{F}_{M_r^{(\ell)}}(t)e^{-st}dt + s\int_0^\infty \tilde{F}_{M_r^{(\ell)}}(t)dt = \int_0^\infty s\tilde{F}_{M_r^{(\ell)}}(t)(1 - e^{-st})dt.$$

But note that $1 - e^{-st}$ is positive in interval $(0, \infty)$. Therefore, according to Equation (50),

$$\mathbb{E}\left[e^{sM_r^{(\ell)}}\right] - 1 - s\mu_i \leq \int_0^\infty se^{-\lambda_{\min}t}(1 - e^{-st}) = s\left(\int_0^\infty e^{-\lambda_{\min}t} - \int_0^\infty e^{-(\lambda_{\min}+s)t}\right)$$
$$= s\left(\frac{1}{\lambda_{\min}} - \frac{1}{\lambda_{\min} + s}\right) = \frac{s^2}{(\lambda_{\min} + s)\lambda_{\min}}. \qquad \square$$

THEOREM 22. *Under the definitions in Section F, the following inequality holds:*

$$\mathbb{P}\left\{S_r \leq (1 - \delta)\mathbb{E}[S_r] \text{ or } S_r \geq (1 + \delta)\mathbb{E}[S_r]\right\} \leq 2e^{-q_r p_r(\sqrt{\lambda_{\min}\delta\mu_r + 1} - 1)^2 n}.$$

PROOF. First, for $\delta \geq 0$, we prove

$$\mathbb{P}\left\{S_r \geq (1 + \delta)\mathbb{E}[S_r]\right\} \leq e^{-q_r p_r(\sqrt{1 + \lambda_{\min}\delta\mu_r} - 1)^2 n}. \tag{58}$$

The idea is to use the moment generating function of $M_r^{(\ell)}$ and derive a Chernoff bound accordingly. According to the Markov inequality,

$$\mathbb{P}\{S_r \ge (1+\delta)\,\mathbb{E}[S_r]\} \le \frac{\mathbb{E}\left[e^{sS_r}\right]}{e^{\mathbb{E}[s(1+\delta)S_r]}} = \prod_\ell \frac{\mathbb{E}\left[e^{sS_r^{(\ell)}}\right]}{e^{s(1+\delta)\,\mathbb{E}[S_r^{(\ell)}]}} = \left(\frac{\mathbb{E}\left[e^{sS_r^{(\ell)}}\right]}{e^{s(1+\delta)\,\mathbb{E}[S_r^{(\ell)}]}}\right)^n. \tag{59}$$

However,

$$\mathbb{E}\left[e^{sS_r^{(\ell)}}\right] = \left(1 - \mathbb{P}\left\{E_r^{(\ell)}\right\}\mathbb{P}\left\{W_r^{(\ell)} \mid E_r^{(\ell)}\right\}\right)e^0 + \mathbb{P}\left\{E_r^{(\ell)}\right\}\mathbb{P}\left\{W_r^{(\ell)} \mid E_r^{(\ell)}\right\}e^{sM_r^{(\ell)}}$$

$$= 1 + p_r q_r \left(\mathbb{E}\left[e^{sM_r^{(\ell)}}\right] - 1\right) \le e^{p_r q_r \left(\mathbb{E}[e^{sM_r^{(\ell)}}]-1\right)}.$$

Therefore,

$$\mathbb{P}\{S_r \ge (1+\delta)\,\mathbb{E}[S_r]\} \le e^{p_r q_r \left(\mathbb{E}\left[e^{sM_r^{(\ell)}}\right]-1-s(1+\delta)\mu_r\right)n} = e^{\left(p_r q_r \left(\mathbb{E}\left[e^{sM_r^{(\ell)}}\right]-1-s\mu_r\right)-sp_r q_r \delta\mu_r\right)n}.$$

Now, according to the inequality given by Equation (52),

$$\mathbb{P}\{S_r \ge (1+\delta)\,\mathbb{E}[S_r]\} \le e^{nq_r p_r \left(\frac{s^2}{(\lambda_{\min}-s)\lambda_{\min}} - s\delta\mu_r\right)}. \tag{60}$$

Here, we aim to find the optimal s that maximizes the function

$$g(s) = q_r p_r \left(\frac{s^2}{(\lambda_{\min}-s)\lambda_{\min}} - s\delta\mu_r\right).$$

By taking derivatives and finding the roots of a third degree polynomial, the optimal s^* is obtained as

$$s^* = \left(1 - \frac{1}{\sqrt{1+\lambda_{\min}\delta\mu_r}}\right)\lambda,$$

which results in

$$-\left(\sqrt{1+\lambda_{\min}\delta\mu_r} - 1\right)^2.$$

Hence,

$$\mathbb{P}\{S_r \ge (1+\delta)\,\mathbb{E}[S_r]\} \le e^{-nq_r p_r \left(\sqrt{1+\lambda_{\min}\delta\mu_r}-1\right)^2}.$$

Next, for $0 \le \delta \le 1$, we prove

$$\mathbb{P}\{S_r \le (1-\delta)\,\mathbb{E}[S_r]\} \le e^{-nq_r p_r \left(1-\sqrt{1-\lambda_{\min}\delta\mu_r}\right)^2 n}.$$

To do so, we start by writing

$$\mathbb{P}\{S_r \ge (1-\delta)\,\mathbb{E}[S_r]\} \le \frac{\mathbb{E}\left[e^{-sS_r}\right]}{e^{\mathbb{E}[-s(1-\delta)S_r]}} = \prod_\ell \frac{\mathbb{E}\left[e^{-sS_r^{(\ell)}}\right]}{e^{-s(1-\delta)\,\mathbb{E}[S_r^{(\ell)}]}} = \left(\frac{\mathbb{E}\left[e^{-sS_r^{(\ell)}}\right]}{e^{-s(1-\delta)\,\mathbb{E}[S_r^{(\ell)}]}}\right)^n. \tag{61}$$

However,

$$\mathbb{E}\left[e^{-sS_r^{(\ell)}}\right] = 1 + p_r q_r \left(\mathbb{E}\left[e^{-sM_r^{(\ell)}}\right] - 1\right) \le e^{p_r q_r \left(\mathbb{E}\left[e^{-sM_r^{(\ell)}}\right]-1\right)},$$

which gives

$$\mathbb{P}\{S_r \le (1-\delta)\,\mathbb{E}[S_r]\} \le e^{np_r q_r \left(\mathbb{E}\left[e^{-sM_r^{(\ell)}}\right]-1+s(1-\delta)\mu_r\right)} = e^{np_r q_r \left(\mathbb{E}\left[e^{-sM_r^{(\ell)}}\right]-1+s\mu_r\right)-nsp_r q_r \delta\mu_r}.$$

By plugging the inequality given by Equation (53) into the above equation, we obtain

$$\mathbb{P}\{S_r \leq (1-\delta)\,\mathbb{E}[S_r]\} \leq e^{nq_r p_r \left(\frac{s^2}{(\lambda_{\min}+s)\lambda_{\min}} - s\delta\mu_r \right)}.$$

Here, we aim to find the optimal s that maximizes the function

$$g(s) = q_r p_r \left(\frac{s^2}{(\lambda_{\min}+s)\lambda_{\min}} - s\delta\mu_r \right).$$

Proceeding similarly as before, the optimal s^* is

$$s^* = \sqrt{\frac{1}{1 - \lambda_{\min}\delta\mu_r}} - 1,$$

which results in

$$g(s^*) = (1 - \sqrt{1 - \lambda_{\min}\delta\mu_r})^2.$$

Therefore,

$$\mathbb{P}\{S_r \leq (1-\delta)\,\mathbb{E}[S_r]\} \leq e^{-nq_r p_r (1 - \sqrt{1 - \lambda_{\min}\delta\mu_r})^2}.$$

Now, we proceed to prove the main inequality. If $\delta \leq 1$, according to the union bound

$$\mathbb{P}\{S_r \leq (1-\delta)\,\mathbb{E}[S_r] \text{ or } S_r \geq (1+\delta)\,\mathbb{E}[S_r]\} \leq e^{-q_r p_r (\sqrt{\lambda_{\min}\delta\mu_r + 1} - 1)^2 n} + e^{-q_r p_r (1 - \sqrt{1 - \lambda_{\min}\delta\mu_r})^2 n},$$

and, if $\delta \geq 1$, then the second probability becomes zero, which implies that

$$\mathbb{P}\{S_r \leq (1-\delta)\,\mathbb{E}[S_r] \text{ or } S_r \geq (1+\delta)\,\mathbb{E}[S_i]\} \leq e^{-q_r p_r (\sqrt{\lambda_{\min}\delta\mu_r + 1} - 1)^2 n}.$$

But it is easy to check that for an arbitrary real $0 \leq x \leq 1$,

$$\sqrt{x+1} - 1 \leq 1 - \sqrt{1-x}. \tag{62}$$

Also note that according to the inequality given by Equation (51), for $\delta \leq 1$

$$0 \leq \lambda_{\min}\delta\mu_r \leq \delta \leq 1.$$

Therefore, we can use the inequality given by Equation (62) and conclude that for $0 \leq \delta \leq 1$,

$$e^{-q_r p_r (1 - \sqrt{1 - \lambda_{\min}\delta\mu_r})^2} \leq e^{-q_r p_r (\sqrt{\lambda_{\min}\delta\mu_r + 1} - 1)^2},$$

which completes the proof. □

LEMMA 23. *Under the definitions in Section F and the assumptions in Theorem 13, the following statements hold:*

1. *For $c \leq 1$ and $r \geq \max\{\ln(\frac{1}{c}) - P, e^2 P\}$, we have*

$$q_r \leq c. \tag{63}$$

2. *For $j \leq \frac{P-1}{2}$,*

$$q_j \geq \frac{1}{2}, \ \mu_j \geq \frac{e-2}{2e} \frac{1}{\lambda_{max}}. \tag{64}$$

3. *For $0 \leq c \leq 1$, if $q_r > c$, then*

$$\mu_r \geq \frac{e^{-(e^2+2)}(e-2)}{\lambda_{max}} \frac{e^{-\frac{\ln\left(\frac{1}{c}\right)}{P-1}}}{\max\left\{ \sqrt{\ln\left(\frac{1}{c}\right)}, \sqrt{(e^2+1)P} \right\}}. \tag{65}$$

4. *Define the function $f(x, y) = \sqrt{x + y} - \sqrt{x}$ with domain $\mathbb{R}^{+2}$. Then, f is monotone decreasing with respect to x and monotone increasing with respect to y.*

5. *For $\psi, w \leq 1$ such that $\frac{w(1-\psi)}{1-\psi w} \leq 0.43$, we have*

$$D(w || \psi w) \geq w \left(\ln \left(\frac{1}{\psi} \right) - 1.31 \frac{(1-w)(1-\psi)}{1-w\psi} \right), \tag{66}$$

   where w and ψw in LHS represent the Bernoulli distributions $(w, 1-w)$ and $(\psi w, 1 - \psi w)$.

6. *For random variable S,*

$$\frac{\beta}{2}(t_f - t_0) \leq \mathbb{E}[S] \leq \rho(t_f - t_0). \tag{67}$$

PROOF. We prove each statement in turn.

—**First statement (Equation (63)).** According to a standard Chernoff bound for Poisson variables, for $x \geq P$,

$$\mathbb{P}\{X \geq x\} \leq \frac{e^{-P}(eP)^x}{x^x}.$$

Hence,

$$q_r \leq \frac{e^{-P}(eP)^{\max\{\ln\left(\frac{1}{c}\right)-P, e^2 P\}}}{\left(\max \left\{ \ln \left(\frac{1}{c} \right) - P, e^2 P \right\} \right)^{\max\{\ln\left(\frac{1}{c}\right)-P, e^2 P\}}} \leq e^{-P} \left(\frac{1}{e} \right)^{\ln\left(\frac{1}{c}\right)-P} = c.$$

—**Second statement (Equation (64)).** By definition, the number of posts X received in a realization follow a Poisson distribution with parameter P. Then, according to a standard Chernoff bound for Poisson tail, for $j \leq P$,

$$\mathbb{P}\{X \geq j\} \geq 1 - \frac{e^{-P}(eP)^j}{j^j}.$$

But note that, for $f(j) = (\frac{eP}{j})^j$ with $j \leq P$ we have

$$f'(j) = \left(1 + \frac{j^2}{eP} \right) \ln \left(\frac{eP}{j} \right) \geq 0,$$

Thus, f is increasing with respect to j. Therefore, for $j \leq \frac{P-1}{2} \leq \frac{P}{2}$ we get

$$\mathbb{P}\{X \geq j\} \geq 1 - \frac{e^{-P}(eP)^{\frac{P}{2}}}{\left(\frac{P}{2} \right)^{\frac{P}{2}}} = 1 - \left(\frac{2}{e} \right)^{\frac{P}{2}} \geq \frac{1}{2},$$

for $P \geq 5$. Hence, for $j \leq \frac{P}{2}$,

$$q_j \geq \frac{1}{2}.$$

This proves the first part. For the second part, it is enough to combine Equation (51) with the above result for q'_j, noting the fact that $j \leq \frac{P-1}{2} \leq \frac{P'}{2}$, which implies $q'_j \geq \frac{1}{2}$ (the condition $P' \geq 5$ also implies that we should have $P \geq 6$), i.e.,

$$\mu_j \geq \frac{q'_j}{q_j} \frac{\frac{e-2}{e}}{\lambda_{\max}} \geq \frac{e-2}{2e} \frac{1}{\lambda_{\max}},$$

where we have also used the fact that $q_j \leq 1$.

— **Third statement (Equation (65)).** By combining the assumption $e^2 + K \leq P$ and Equation (63), we have that $r \leq \max\{\ln(\frac{1}{c}) - P, e^2 P\}$. Therefore, based on Equation (51), we have

$$\mu_r \geq \frac{e-2}{e} \frac{1}{\lambda_{\max}} \frac{e^{-\frac{i+r-1}{P-1}}}{\sqrt{i+r-1}} \geq \frac{e-2}{e} \frac{1}{\lambda_{\max}} \frac{\min\left\{e^{-\frac{K+\ln\left(\frac{1}{c}\right)-P-1}{P-1}}, e^{-(e^2+1)}\right\}}{\max\left\{\sqrt{K + \ln\left(\frac{1}{c}\right) - P - 1}, \sqrt{(e^2+1)P}\right\}}$$

$$\geq \frac{e-2}{e} \frac{1}{\lambda_{\max}} \frac{\min\left\{e^{-\frac{\ln\left(\frac{1}{c}\right)}{P-1}}, e^{-(e^2+1)}\right\}}{\max\left\{\sqrt{\ln\left(\frac{1}{c}\right)}, \sqrt{(e^2+1)P}\right\}} \geq \frac{e^{-(e^2+1)}(e-2)}{e\lambda_{\max}} \frac{e^{-\frac{\ln\left(\frac{1}{c}\right)}{P-1}}}{\max\left\{\sqrt{\ln\left(\frac{1}{c}\right)}, \sqrt{(e^2+1)P}\right\}}.$$

— **Fourth statement.** For any $(x, y) \in \mathbb{R}^{+2}$,

$$\frac{df}{dx} = \frac{1}{2\sqrt{x+y}} - \frac{1}{2\sqrt{x}} \leq 0,$$

$$\frac{df}{dy} = \frac{1}{2\sqrt{x+y}} \geq 0,$$

which completes the proof.

— **Fifth statement (Equation (66)).** Starting from the definition of cross-entropy, we have that:

$$D(w||\psi w) = w \ln\left(\frac{w}{\psi w}\right) + (1 - w) \ln\left(\frac{1-w}{1-\psi w}\right) = w \ln\left(\frac{1}{\psi}\right) + (1 - w) \ln\left(1 - \frac{w(1-\psi)}{1-\psi w}\right).$$

But, for $0 \leq x \leq 0.43$, it holds that

$$\ln(1 - x) \geq -1.31x.$$

Hence,

$$D(w||\psi w) \geq w \ln\left(\frac{1}{\psi}\right) - 1.31w \frac{(1-w)(1-\psi)}{1-w\psi} w = w \left(\ln\left(\frac{1}{\psi}\right) - 1.31\frac{(1-w)(1-\psi)}{1-w\psi}\right).$$

— **Sixth statement (Equation (67)).** The idea is to calculate the expectation $\mathbb{E}[S]$ as follows:

$$\mathbb{E}[S] = \sum_{r=1}^{\infty} \mathbb{E}\left[\mathbf{1}\left\{E_r^{(\ell)}\right\} \mathbf{1}\left\{W_r^{(\ell)}\right\} S_r\right] = \sum_{r=1}^{\infty} \mathbb{P}\left\{E_r^{(\ell)}\right\} \int_0^{\infty} t\mathbb{P}\left\{S_r = t \mid E_r^{(\ell)}\right\} \mathbb{P}\left\{W_r^{(\ell)} \mid E_r^{(\ell)}, S_r = t\right\}.$$

(68)

But note that

$$\mathbb{P}\left\{W_r^{(\ell)} \mid E_r^{(\ell)}, S_r = t\right\}$$

$$= \int_0^{\infty} \mathbb{P}\left\{W_r^{(\ell)} \mid P_r^{(\ell)} \text{ is received at t}, Y_r = t\right\} \mathbb{P}\left\{P_r^{(\ell)} \text{ is received at t}, W_r^{(\ell)}, S_r = t\right\}$$

$$= \int_0^{\infty} \mathbb{P}\left\{W_r^{(\ell)} \mid P_r^{(\ell)} \text{ is received at t}\right\} \mathbb{P}\left\{P_r^{(\ell)} \text{ is received at t} \mid E_r^{(\ell)}, S_r = t\right\},$$

and using the fact that our intensity function μ is at least a fraction β of the whole intensity given any time, we have that

$$\mathbb{P}\{W_r^{(\ell)} \mid P_r^{(\ell)} \text{ is received at t}\} \geq \beta.$$

Therefore,

$$\mathbb{P}\left\{W_r^{(\ell)} \mid E_r^{(\ell)}, S_r = t\right\} \geq \int_0^\infty \beta \mathbb{P}\left\{P_r^{(\ell)} \text{ is received at t} \mid E_r^{(\ell)}, S_r = t\right\} = \beta. \tag{69}$$

Then, by applying Equation (69) into Equation (68),

$$\mathbb{E}[S] \geq \sum_{r=1}^\infty \mathbb{P}\left\{E_r^{(\ell)}\right\} \int_0^\infty t \mathbb{P}\left\{S_r = t \mid E_r^{(\ell)}\right\} \beta = \sum_{r=1}^\infty q_r \beta \int_0^\infty t \mathbb{P}\left\{S_r = t \mid E_r^{(\ell)}\right\} = \beta \sum_{r=1}^\infty q_r \mu_r.$$

Note that the quantity $\sum_{r=1}^\infty q_r \mu_r$ is exactly the total amount of time that position i is occupied between t_0 and t_f. According to our assumption, we know that with probability 1 all the positions are occupied at time t_0, which implies that $\sum_{r=1}^\infty q_r \mu_r = t_f - t_0 \geq \frac{1}{2}(t_f - t_0)$. Hence,

$$\mathbb{E}[\Delta_{i,j}] \geq \frac{\beta}{2}(t_f - t_0). \tag{70}$$

The upper bound can be proved with exactly the same approach, using the inequality $\sum_{r=1}^\infty q_r \mu_r \leq (t_f - t_0)$ instead. □

LEMMA 24. *For each $r < r_1$,*

$$\mathbb{P}\left\{S_r \leq \left(1 - \frac{\delta}{4}\right)\mathbb{E}[S_r] \text{ or } S_r \geq \left(1 + \frac{\delta}{4}\right)\mathbb{E}[S_r]\right\} \leq 2\exp\left(-\beta c'\left(\sqrt{1 + \frac{\lambda_{min}}{\lambda_{max}}\frac{\delta}{4}c''} - 1\right)^2 n\right),$$

where

$$c'' = e^{-(e^2+2)}(e - 2)\frac{c'^{\frac{1}{P-1}}}{\max\left\{\sqrt{\ln\left(\frac{1}{c'}\right)}, \sqrt{(e^2 + 1)P}\right\}},$$

and c' is defined by Equation (34).

PROOF. Using Theorem 22, we can write

$$\mathbb{P}\left\{S_r \leq \left(1 - \frac{\delta}{4}\right)\mathbb{E}[S_r] \text{ or } S_r \geq \left(1 + \frac{\delta}{4}\right)\mathbb{E}[S_r]\right\} \leq 2\exp\left(-q_r p_r\left(\sqrt{1 + \lambda_{min}\frac{\delta}{4}\mu_r} - 1\right)^2 n\right).$$

However, note that according to Lemma 65, we have $\mu_r \geq \frac{c''}{\lambda_{max}}$. Also, we know that $p_r \geq \beta, q_r \geq c'$. Hence, using the fourth statement of Lemma 23, we obtain our desired result. □

LEMMA 25.

$$\mathbb{P}\left\{\sum_{r=r_2}^\infty S_r \geq \delta \mathbb{E}[S]\right\} \leq e^{-\beta\frac{\delta}{2}n} \leq \exp\left(-\beta c'\left(\sqrt{1 + \frac{\lambda_{min}}{\lambda_{max}}\delta c''} - 1\right)^2 n\right).$$

PROOF. For $1 \leq \ell \leq n$, Let $\mathcal{E}_\ell$ an indicator function with value 1 if at least r_2 stories are received in realization s. Then, we know that $\mathbb{E}[\mathcal{E}_\ell] = q_{r_2} \leq 0.1\delta\beta$. Therefore, if we define the random variable $N(\mathcal{E})$ as the number of times that at least r_2 stories are received, then we get that $N(\mathcal{E})$ is a binomial random variable with success probability q_{r_2}. Hence, according to a standard Chernoff tail bound for binomial distributions, it follows that:

$$\mathbb{P}\{N(\mathcal{E}) \geq n\beta\delta\} \leq e^{-nD\left(\frac{n\beta\delta}{n}\|q_{r_2}\right)} = e^{-nD(\beta\delta\|q_{r_2})} \leq e^{-nD(\beta\delta\|0.1\delta\beta)}.$$

But the exponent can be checked to be greater than $\delta\beta$ according to Equation (66), using parameters $w = \delta\beta$ and $\psi = 0.1$. Therefore,

$$\mathbb{P}\{N(\mathcal{E}) \geq n\beta\delta\} \leq e^{-n\beta\delta}. \tag{71}$$

Now suppose we have $N(\mathcal{E}) \leq n\beta\delta$. Then,

$$\sum_{i=i_2}^{\infty} S_i = \frac{1}{n} \sum_{\ell=1}^{n} \sum_{i=i_2}^{\infty} S_r^{(\ell)} \leq \frac{(t_f - t_0)n\beta\delta}{n} = (t_f - t_0)\beta\delta \leq 2\delta \, \mathbb{E}[S], \tag{72}$$

where the last inequality follows from Equation (67). Substituting δ by $\frac{\delta}{2}$ we conclude that

$$\mathbb{P}\left\{\sum_{r=r_2}^{\infty} S_r \geq \delta \, \mathbb{E}[S]\right\} \leq \mathbb{P}\left\{N(\mathcal{E}) \geq n\beta\frac{\delta}{2}\right\} \leq e^{-n\beta\frac{\delta}{2}}. \tag{73}$$

Finally, note that

$$\beta c'\left(\sqrt{1 + \frac{\lambda_{\min}}{\lambda_{\max}}\delta c''} - 1\right)^2 \leq c'\frac{\lambda_{\min}}{\lambda_{\max}}\delta c'' \leq \frac{\delta}{2}\beta,$$

which completes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

LEMMA 26. *For $r \geq r_1$ and $j \leq \frac{P-1}{2}$,*

$$\mathbb{P}\left\{S_r \geq \frac{\epsilon\delta}{2} \, \mathbb{E}[S_j]\right\} \leq \exp\left(-\beta c'\left(\sqrt{1 + \frac{\lambda_{min}}{\lambda_{max}}\frac{\delta}{4}c''} - 1\right)^2 n\right).$$

PROOF. First, note that combining Equations (49) and (64), we have that $p_j q_j \geq \frac{\beta}{2}$. Therefore,

$$p_r q_r \leq \rho\frac{\beta}{2\rho}\epsilon = \frac{\beta}{2}\epsilon \leq \epsilon p_j q_j. \tag{74}$$

Second, note that according to Equation (64) and the definition of r_1, we have that

$$\frac{\epsilon\delta}{4} \, \mathbb{E}[S_j] = \frac{\epsilon\delta}{4}p_j q_j \, \mathbb{E}[M_j] \geq \frac{\epsilon\delta}{4}\beta\frac{e-2}{4e}\frac{1}{\lambda_{\max}} = \rho\frac{1}{\lambda_{\min}}\epsilon\frac{e-2}{4e}\frac{\beta}{\rho}\frac{\lambda_{\min}}{\lambda_{\max}}\frac{\delta}{4} \geq p_r q_r \, \mathbb{E}[M_r] = \mathbb{E}[S_r]. \tag{75}$$

Thus,

$$\mathbb{P}\left\{S_r \geq \frac{\epsilon\delta}{2} \, \mathbb{E}[S_j]\right\} = \mathbb{P}\left\{S_r \geq \frac{\epsilon\delta}{4} \, \mathbb{E}[S_j] + \frac{\epsilon\delta}{4} \, \mathbb{E}[S_j]\right\} \leq \mathbb{P}\left\{S_r \geq \mathbb{E}[S_r] + \frac{\epsilon\delta}{4} \, \mathbb{E}[S_j]\right\}.$$

Then, according to Equations (58) and (74) and the fourth statement of Lemma 23, it follows that:

$$\mathbb{P}\left\{S_r \geq \frac{\epsilon\delta}{2}\mathbb{E}[S_j]\right\} \leq \exp\left(-p_r q_r \left(\sqrt{1 + \lambda_{\min}\frac{\epsilon\delta}{4}\frac{\mathbb{E}[S_j]}{\mathbb{E}[S_r]}\mu_r} - 1\right)^2 n\right)$$

$$= \exp\left(-\left(\sqrt{p_r q_r + \lambda_{\min}\frac{\epsilon\delta}{4}\frac{\mathbb{E}[S_j]}{\mathbb{E}[S_r]}\mu_r p_r q_r} - \sqrt{p_r q_r}\right)^2 n\right)$$

$$= \exp\left(-\left(\sqrt{p_r q_r + \lambda_{\min}\frac{\epsilon\delta}{4}\mathbb{E}[S_j]} - \sqrt{p_r q_r}\right)^2 n\right)$$

$$\leq \exp\left(-\left(\sqrt{\epsilon p_j q_j + \lambda_{\min}\frac{\epsilon\delta}{4}\mathbb{E}[S_j]} - \sqrt{\epsilon p_j q_j}\right)^2 n\right)$$

$$= \exp\left(-\epsilon p_j q_j \left(\sqrt{1 + \lambda_{\min}\frac{\delta}{4}\mathbb{E}[S_j]} - 1\right)^2 n\right)$$

$$\leq \exp\left(-\epsilon\beta\frac{1}{2}\left(\sqrt{1 + \lambda_{\min}\frac{\delta}{4}\frac{e-2}{2e\lambda_{\max}}} - 1\right)^2 n\right) \tag{76}$$

$$\leq \exp\left(-\beta c'\left(\sqrt{1 + \frac{\lambda_{\min}}{\lambda_{\max}}\frac{\delta}{4}c''} - 1\right)^2 n\right), \tag{77}$$

where Equation (76) follows from Equations (64), and (77) follows from $c'' \leq \frac{e-2}{2e}$, the fourth statement of Lemma 23, and the fact that $c' \leq \frac{1}{2}\epsilon$. □

LEMMA 27.

$$\mathbb{P}\left\{\sum_{r=r_1-1}^{r_2-1} S_r \geq \frac{\delta}{2}\mathbb{E}[S]\right\} \leq (r_2 - r_1)\exp\left(-\beta c'\left(\sqrt{1 + \frac{\lambda_{min}}{\lambda_{max}}\frac{\delta}{4}c''} - 1\right)^2 n\right).$$

PROOF. We divide the indices in the range between r_1 and r_2 into $\lfloor\frac{P-1}{2}\rfloor$ groups $A_1, \ldots, A_{\lfloor\frac{P-1}{2}\rfloor}$, and we assign each group to a particular index $1 \leq j \leq \frac{P-1}{2}$. Then, using Equation (63) for $r_2 - 1$, we get

$$r_2 - 1 \leq \max\left\{\ln\left(\frac{1}{0.1\beta\delta}\right), e^2 P\right\} \tag{78}$$

$$= \left\lfloor\frac{P-1}{2}\right\rfloor\max\left\{\frac{\ln\left(\frac{1}{0.1\beta\delta}\right)}{\lfloor\frac{P-1}{2}\rfloor}, \frac{P}{\lfloor\frac{P-1}{2}\rfloor}e^2\right\} \leq \left\lfloor\frac{P-1}{2}\right\rfloor\max\left\{\frac{\ln\left(\frac{1}{0.1\beta\delta}\right)}{\lfloor\frac{P-1}{2}\rfloor}, 2.5e^2\right\} = \left\lfloor\frac{P-1}{2}\right\rfloor\frac{1}{\epsilon},$$

due to the fact that $P \geq 6$. Moreover, each A_r has cardinality at most $\frac{1}{\epsilon}$ due to the fact that $\frac{r_2 - r_1}{\lfloor \frac{P-1}{2} \rfloor} \leq \frac{r_2 - 1}{\lfloor \frac{P-1}{2} \rfloor} \leq \frac{1}{\epsilon}$. Therefore, we have $\forall 1 \leq j \leq \frac{P-1}{2}$, $\frac{1}{|A_j|} \geq \epsilon$. Hence, we can write

$$\mathbb{P}\left\{ \sum_{r=r_1-1}^{r_2-1} S_r \geq \frac{\delta}{2} \mathbb{E}[S] \right\} \leq \mathbb{P}\left\{ \sum_{r=r_1-1}^{r_2-1} S_r \geq \frac{\delta}{2} \mathbb{E}[\sum_{j=1}^{\frac{P-1}{2}} S_j] \right\} \leq \sum_{j=1}^{\lfloor \frac{P-1}{2} \rfloor} \mathbb{P}\left\{ \sum_{r \in A_j} S_r \geq \frac{\delta}{2} \mathbb{E}[S_j] \right\}$$

$$\leq \sum_{j=1}^{\lfloor \frac{P-1}{2} \rfloor} \sum_{r \in A_j} \mathbb{P}\left\{ S_r \geq \epsilon \frac{\delta}{2} \mathbb{E}[S_j] \right\}$$

$$\leq (r_2 - r_1) \exp\left(-\beta c' \left(\sqrt{1 + \frac{\lambda_{\min}}{\lambda_{\max}} \frac{\delta}{4} c''} - 1 \right)^2 n \right),$$

where the last inequality follows from Lemma 26. □

LEMMA 28. *Let the number of realizations* $n \geq \frac{Z_y + \log \frac{\sum_{i \in \mathcal{B}} |\mathcal{W}_{i,:}|}{\delta}}{Q_y}$. *Then, with probability at least* $1 - \delta$, *the greedy algorithm returns a set of edges* $\mathcal{E}_{\mathcal{B}}$ *such that*

$$\mathcal{U}_K(\mathcal{E}_{\mathcal{B}}) \geq \frac{1}{\frac{1}{1-\check{\alpha}} + 1} OPT - 4yK(t_f - t_0) \sum_{i \in \mathcal{B}} c_i.$$

PROOF. We adapt the proof of Theorem 7. Define $S, S_i^j, S_i, s_{ik}, Q_{ik}$ with respect to the greedy with empirical estimates—the empirical greedy. Let $\mathcal{E}_{\mathcal{B}}^{(i)}$ be the edges picked by the empirical greedy before picking broadcaster i's edges. Let E be the event that $\forall i \in \mathcal{B}, \forall j : (i, j) \in \mathcal{W}_{i,:}$,

$$\left| \hat{\mathcal{U}}\left(\mathcal{E}_{\mathcal{B}}^{(i)}, j \right) - \mathcal{U}\left(\mathcal{E}_{\mathcal{B}}^{(i)}, j \right) \right| \leq yK(t_f - t_0). \tag{79}$$

Then, we can apply union bounds and Equation (18) to conclude that $\Pr\{E^o\} \leq (\sum_{i \in \mathcal{B}} |\mathcal{W}_{i,:}|) e^{Z_y - Q_y n} \leq \delta$, which implies that E happens with high probability. Moreover, $\forall i \in \mathcal{B}, 1 \leq k \leq c_i$, s_{ik}'s empirical marginal gain (when added to Q_{ik}) is not less than that of t_{ik}, due to the selection rule of the greedy algorithm. Consequently, given that E happens, by rewriting the empirical utilities in terms of the theoretical ones based on the inequality in Equation (79), we can rewrite Equation (31) as $\rho_{s_{ik}}(Q_{ik}) \geq \rho_{t_{ik}}(Q_{ik}) - 4yK(t_f - t_0)$. Then, we plug this expression into the proof of Theorem 7 and obtain:

$$F(\mathcal{T}) \leq F(\mathcal{S}) + \sum_{i \in \mathcal{B}} \sum_{k=1}^{c_i} \frac{1}{1 - \check{\alpha}} \rho_{t_{ik}}(Q_{ik}) \leq F(\mathcal{S}) + \frac{1}{1 - \check{\alpha}} \sum_{i \in \mathcal{B}} \sum_{k=1}^{c_i} (\rho_{s_{ik}}(Q_{ik}) + 4yK(t_f - t_0))$$

$$= \left(1 + \frac{1}{1 - \check{\alpha}} \right) F(\mathcal{S}) + \frac{4}{\zeta} \left(\sum_{i \in \mathcal{B}} c_i \right) yK(t_f - t_0).$$

Then, we can divide both sides by $1 + \frac{1}{1-\check{\alpha}}$ and use that $\frac{\frac{1}{\zeta}}{\frac{1}{\zeta} + 1} \leq 1$ to complete the proof. □

REFERENCES

[1] O. Aalen, O. Borgan, and H. Gjessing. 2008. *Survival and Event History Analysis: A Process Point of View*. Springer.

[2] A. Asudehy, H. Jagadishy, J. Stoyanovichz, and G. Das. 2017. Designing fair ranking schemes. *arXiv preprint arXiv:1712.09752* (2017).

[3] L. Backstrom, E. Bakshy, J. M. Kleinberg, T. M. Lento, and I. Rosenn. 2011. Center of attention: How Facebook users allocate attention across friends. In *ICWSM*.

[4] A. Bian, J. Buhmann, A. Krause, and S. Tschiatschek. 2017. Guarantees for greedy maximization of non-submodular functions with applications. *arXiv preprint arXiv:1703.02100* (2017).

[5]   A. Biega, K. Gummadi, and G. Weikum. 2018. Equity of attention: Amortizing individual fairness in rankings. *arXiv preprint arXiv:1805.01788* (2018).

[6]   Ilija Bogunovic, Junyao Zhao, and Volkan Cevher. 2018. In *AISTATS*.

[7]   Gruia Calinescu, Chandra Chekuri, Martin Pál, and Jan Vondrák. 2011. Maximizing a monotone submodular function subject to a matroid constraint. *SIAM J. Comput.* 40, 6 (2011), 1740−1766.

[8]   M. Cha, H. Haddadi, F. Benevenuto, and K. Gummadi. 2010. Measuring user influence in Twitter: The million follower fallacy. In *ICWSM*.

[9]   Lin Chen, Moran Feldman, and Amin Karbasi. 2018. Weakly submodular maximization beyond cardinality constraints: Does randomization help greedy? In *ICML*.

[10] Michele Conforti and Gérard Cornuéjols. 1984. Submodular set functions, matroids and the greedy algorithm: Tight worst-case bounds and some generalizations of the Rado-Edmonds theorem. *Discr. Appl. Math.* 7, 3 (1984), 251−274.

[11] M. Crawford. 2015. *The World beyond Your Head: On Becoming an individual in an Age of Distraction.* Farrar, Straus and Giroux.

[12] A. Das and D. Kempe. 2011. Submodular meets spectral: Greedy algorithms for subset selection, sparse approximation and dictionary selection. In *ICML*.

[13] Yuval Filmus and Justin Ward. 2012. A tight combinatorial algorithm for submodular maximization subject to a matroid constraint. In *FOCS*. IEEE, 659−668.

[14] Marshall L. Fisher, George L. Nemhauser, and Laurence A. Wolsey. 1978. An analysis of approximations for maximizing submodular set functions—II. In *Polyhedral Combinatorics*. Springer, 73−87.

[15] M. Gomez-Rodriguez, K. P. Gummadi, and B. Schoelkopf. 2014. Quantifying information overload in social media and its impact on social contagions. In *ICWSM*.

[16] Christopher Harshaw, Moran Feldman, Justin Ward, and Amin Karbasi. 2019. Submodular maximization beyond non-negativity: Guarantees, fast algorithms, and applications. In *ICML*.

[17] Hamed Hassani, Mahdi Soltanolkotabi, and Amin Karbasi. 2017. Gradient methods for submodular maximization. In *NIPS*.

[18] N. Hodas and K. Lerman. 2012. How visibility and divided attention constrain social contagion. In *SocialCom*.

[19] J. Kang and K. Lerman. 2015. VIP: Incorporating human cognitive biases in a probabilistic model of retweeting. In *ICSC*.

[20] M. Karimi, E. Tavakoli, M. Farajtabar, L. Song, and M. Gomez-Rodriguez. 2016. Smart broadcasting: Do you want to be seen? In *KDD*.

[21] E. Khalil, B. Dilkina, and L. Song. 2014. Scalable diffusion-aware optimization of network topology. In *KDD*.

[22] M. Kimura, K. Saito, and H. Motoda. 2008. Solving the contamination minimization problem on networks for the linear threshold model. In *PRICAI*.

[23] M. Kimura, K. Saito, and H. Motoda. 2009. Blocking links to minimize contamination spread in a social network. *ACM Trans. Knowl. Discov. Data* 3, 2 (2009), 1−23.

[24] C. Kuhlman, G. Tuli, S. Swarup, M. Marathe, and S. Ravi. 2013. Blocking simple and complex contagion by edge removal. In *ICDM*.

[25] E. Lawler. 2001. *Combinatorial Optimization: Networks and Matroids.* Courier Corporation.

[26] Benny Lehmann, Daniel Lehmann, and Noam Nisan. 2006. Combinatorial auctions with decreasing marginal utilities. *Games Econ. Behav.* 55, 2 (2006), 270−296.

[27] K. Lerman and T. Hogg. 2014. Leveraging position bias to improve peer recommendation. *PloS One* 9, 6 (2014), e98914.

[28] L. Lü and T. Zhou. 2011. Link prediction in complex networks: A survey. *Phys. A: Statist. Mechan. Applic.* 390, 6 (2011), 1150−1170.

[29] G. Nemhauser, L. Wolsey, and M. Fisher. 1978. An analysis of approximations for maximizing submodular set functions—I. *Math. Prog.* 14, 1 (1978), 265−294.

[30] A. Singh and T. Joachims. 2018. Fairness of exposure in rankings. In *KDD*.

[31] D. Snyder and M. Miller. 2012. *Random Point Processes in Time and Space.* Springer.

[32] N. Spasojevic, Z. Li, A. Rao, and P. Bhattacharyya. 2015. When-to-post on social networks. In *KDD*.

[33] H. Tong, B. Prakash, T. Eliassi-Rad, M. Faloutsos, and C. Faloutsos. 2012. Gelling, and melting, large graphs by edge manipulation. In *CIKM*.

[34] U. Upadhyay, A. De, and M. Gomez-Rodriguez. 2018. Deep reinforcement learning of marked temporal point processes. In *NIPS*.

[35] Jan Vondrák. 2008. Optimal approximation for the submodular welfare problem in the value oracle model. In *STOC*. ACM, 67−74.

[36] Jan Vondrák. 2010. Submodularity and curvature: The optimal algorithm. *RIMS Kokyuroku Bessatsu* 23 (2010), 253−266.

[37] K. Yang and J. Stoyanovich. 2017. Measuring fairness in ranked outputs. In *ICSSDM*.

[38] A. Zarezade, A. De, U. Upadhyay, H. Rabiee, and M. Gomez-Rodriguez. 2018. Steering social activity: A stochastic optimal control point of view. *J. Mach. Learn. Res.* 18 (2017), 205–1.

[39] A. Zarezade, U. Upadhyay, H. Rabiee, and M. Gomez-Rodriguez. 2017. RedQueen: An online algorithm for smart broadcasting in social networks. In *WSDM*.

[40] Haifeng Zhang and Yevgeniy Vorobeychik. 2016. Submodular optimization with routing constraints. In *AAAI*.