A. Organization of the Appendix

Appendix B defines the greedy versions of the non-submodularity parameters.

Appendix C provides omitted proofs from Section 3.

Appendix D defines the BinarySearchPivot procedure and omitted proofs from Section 4.1.

Appendix E provides omitted proofs from Section 4.2.

Appendix F defines the Independent Cascade model, proves that classical IM and boosting are subproblems of our IM model, and provides the proof of Theorem 3 from Section 5.

Appendix G provides additional experimental results characterizing the parameters of FastGreedy.

Appendix H presents details of our GIM implementation.

B. Greedy Versions of Non-Submodularity Parameters

We define various greedy versions of the non-submodularity parameters in this section. In this work, these are referred to as FastGreedy weak DR ratio, *etc.*, where the instance is clear from the context.

ThresholdGreedy DR ratio.

Definition 5 (ThresholdGreedy DR ratio)**.** Let an instance $\mathcal{I}$ of Problem MCC be given, with budget constraint k. Let $\mathbf{g}^1, \ldots, \mathbf{g}^T$ be the sequence of values $\mathbf{g}$ takes during execution of ThresholdGreedy on $\mathcal{I}$. The *ThresholdGreedy version of the DR ratio on* $\mathcal{I}$ $\gamma_d^{\text{TG},\mathcal{I}}(f) \in [0,1]$, is the maximum value such that for any $i \in \{1, \ldots, T\}$, for any $s \in S$, if $\mathbf{g}^{i,s}$ is the value of the greedy vector immediately after s was considered during the inner **for** loop of the threshold directly preceding the one in which $\mathbf{g}^i$ was considered ($\mathbf{g}^{i,s} = \mathbf{0}$ if $\mathbf{g}^i$ was considered during the first threshold),

$$\gamma_d^{\text{TG},\mathcal{I}} \delta_{\mathbf{s}}(\mathbf{g}^i) \leq \delta_{\mathbf{s}}(\mathbf{g}^{i,s}).$$

Greedy versions of weak DR ratio.

Definition 6. Let $\mathcal{A} \in \{\text{StandardGreedy}, \text{ThresholdGreedy}\}$, and let an instance $\mathcal{I}$ of Problem MCC be given, with budget constraint k. Let $\mathbf{g}^1, \ldots, \mathbf{g}^T$ be the sequence of values $\mathbf{g}$ takes during execution of $\mathcal{A}$ on $\mathcal{I}$. The $\mathcal{A}$ *version of the weak DR ratio on* $\mathcal{I}$ $\gamma_s^{\mathcal{A},\mathcal{I}} \in [0,1]$, is the maximum value such that for any $s \in S$, for any $i \in \{1, \ldots, T\}$, for any $\mathbf{w}$ such that $\mathbf{g}^i \leq \mathbf{w}$ and $\|\mathbf{w} - \mathbf{g}^i\|_1 \leq k$ and $\mathbf{w} \leq \mathbf{b}$,

$$\gamma_s^{\mathcal{A},\mathcal{I}} \left[f(\mathbf{w}) - f(\mathbf{g}^i) \right] \leq \sum_{s \in \{\mathbf{w} - \mathbf{g}^i\}} \delta_{\mathbf{s}}(\mathbf{g}^i).$$

The *FastGreedy weak DR ratio* differs from the above two only in that the sequence of vectors $\mathbf{g}^1, \ldots, \mathbf{g}^T$ are the value of the greedy vector $\mathbf{g}$ at the beginning of each iteration of the outer **while** loop, instead of all values of $\mathbf{g}$ during execution of the algorithm.

C. Proofs for Section 3

Proof of Proposition 1. Suppose $\mathbf{v} \leq \mathbf{w} \in \mathbb{N}^S$. Let $\{\mathbf{w} - \mathbf{v}\} = \{s_1, ..., s_l\}$. Then,

$$\begin{aligned}
\gamma_d(f(\mathbf{w}) - f(\mathbf{v})) &= \gamma_d \sum_{i=1}^{l} [f(\mathbf{v} + \mathbf{s}_1 + ... + \mathbf{s}_i) - f(\mathbf{v} + \mathbf{s}_1 + ... + \mathbf{s}_{i-1})] \\
&= \gamma_d \sum_{i=1}^{l} \delta_{\mathbf{s}_i}(\mathbf{v} + \mathbf{s}_1 + ... + \mathbf{s}_{i-1}) \\
&\leq \sum_{i=1}^{l} \delta_{\mathbf{s}_i}(\mathbf{v})
\end{aligned}$$

Algorithm 4 BinarySearchPivot$(f, \mathbf{g}, \mathbf{b}, s, k, \tau)$

```
1: Input: f ∈ F_b, g ∈ N^S, b ∈ N^S, s ∈ S, k ∈ N, τ ∈ R^+
2: Output: l ∈ N
3: l_s ← 1, l_t ← min{b_s − g_s, k − ||g||_1},
4: if δ_{l_t s}(g) ≥ l_t τ then
5:     return l_t
6: if δ_s(g) < τ then
7:     return 0
8: while l_t ≠ l_s + 1 do
9:     m = ⌊(l_t + l_s)/2⌋
10:    if δ_{ms}(g) ≥ mτ then
11:        l_s = m
12:    else
13:        l_t = m
14: return l_s
```

Therefore, $\gamma_d \leq \gamma_s$, since γ_s is the maximum number satisfying the above inequality. The second statement follows from reduction of the lattice to sets, the well-known characteriziation of submodularity of a set function in terms of diminishing returns, and application of results on the weak DR ratio from Bian et al. (2017b). □

D. BinarySearchPivot and Proofs for Section 4.1 (ThresholdGreedy)

BinarySearchPivot. The routine BinarySearchPivot (Alg. 4) efficiently finds a pivot for each $s \in S$. BinarySearchPivot uses a modified binary-search procedure that maintains $l_s < l_t$ such that both

$$\delta_{l_s \mathbf{s}}(\mathbf{g}) \geq l_s \tau, \tag{1}$$

$$\delta_{l_t \mathbf{s}}(\mathbf{g}) < l_t \tau. \tag{2}$$

Initially, l_s and l_t do satisfy (1), (2), or else we have already found a valid pivot (lines 4, 4). The midpoint m of the interval $[l_s, l_t]$ is tested to determine if l_s or l_t should be updated to maintain (1), (2); this process continues until $l_t = l_s + 1$.

Lemma 2. BinarySearchPivot finds a valid pivot $l \in \{0, \ldots, l_{max}\}$ in $O(\log l_{max})$ queries of f, where $l_{max} = \min\{\mathbf{b}_s - \mathbf{g}_s, k - \|\mathbf{g}\|_1\}$, $\mathbf{b}_{max} = \max_{s \in S} \mathbf{b}_s$.

Proof of Lemma 2. The routine BinarySearchPivot maintains inequalities (1), (2), it is enough to show that given (1), (2), there exists a $l \in \{l_s, ..., l_t - 1\}$ such that l is a pivot. Consider $l_j = l_t - j$, for $j \in \mathbb{N}$; there must be a smallest $j \geq 1$ such that l_j satisfies property (P1) of being a pivot, since $l_s < l_t$ satisfies property (P1). If property (P2) is unsatisfied, then

$$\begin{aligned} \delta_{(l_j+1)e_i}(s) &= \delta_i(s + l_j) + \delta_{l_j e_i}(s) \\ &\geq \tau + l_j \tau = (l_j + 1)\tau, \end{aligned}$$

contradicting the choice of j since $l_j + 1 = l_{j-1}$. The query complexity follows from a constant number of queries per iteration of the while loop and the fact that each iteration reduces the distance from l_s to l_t by a factor of 2; initially, this distance was l_{max}. □

Omitted proofs from Section 4.1.

Proof that Property 1 holds. Let $\mathbf{g}^{\tau,s}$ be the value of $\mathbf{g}$ immediately after s is considered during the iteration corresponding to τ; then property (P2) of pivot was satisfied: $\delta_{\mathbf{s}}(\mathbf{g}^{\tau,s}) < \tau$. □

Proof of Claim 1. Suppose $\gamma_d \geq \varepsilon$. Suppose $\|\mathbf{g}\|_1 < k$, and let $\mathbf{g}'$ be the solution returned by a modified ThresholdGreedy that continues updating the threshold until $\|\mathbf{g}'\|_1 = k$. Order $\{\mathbf{g}'\} \setminus \{\mathbf{g}\} = \{\mathbf{s}_1, \ldots, \mathbf{s}_\ell\}$, and let $\mathbf{g}'_i = \mathbf{g}'_{i-1} + \mathbf{s}_i$,

$i = 1, \ldots, \ell$, with $\mathbf{g}'_0 = \mathbf{g}$, so that $\mathbf{g}'_\ell = \mathbf{g}'$. Also, let $\mathbf{g}_{i-1} \le \mathbf{g}$ be the vector guaranteed for $\mathbf{s}_i$ by Lemma 1 with the last threshold value τ of ThresholdGreedy. Then

$$\begin{aligned} f(\mathbf{g}') - f(\mathbf{g}) &= \sum_{i=1}^{\ell} \delta_{\mathbf{s}_i}(\mathbf{g}'_{i-1}) \\ &\le \frac{1}{\gamma_d} \sum_{i=1}^{\ell} \delta_{\mathbf{s}_i}(\mathbf{g}_{i-1}) \\ &\le \frac{\ell}{\varepsilon} \frac{\varepsilon^2 M}{k} \le \varepsilon M \le \varepsilon f(\mathbf{\Omega}). \end{aligned}$$

Hence, for any $\Phi > \varepsilon$, if

$$f(\mathbf{g}') \ge \Phi f(\mathbf{\Omega}),$$

then

$$f(\mathbf{g}) \ge (\Phi - \varepsilon) f(\mathbf{\Omega}).$$

□

From proof of Theorem 1:

"If $\gamma_d < \varepsilon$, the ratio holds trivially". If $\gamma_d < \varepsilon$, the ratio holds trivially from the inequality $1 - e^{-x} \le x$, for real $x > 0$, since

$$1 - e^{-\gamma_d \gamma_s \kappa} \le \gamma_d \gamma_s \kappa < \varepsilon.$$

"from which the hypothesis of Claim 1 follows". Since $(1-x) \le e^{-x}$ and $\sum_t l^t = k$, we have $\prod_{t=1}^{T} (1 - l^t \gamma_d \gamma_s \kappa / k) \le \prod_{t=1}^{T} \exp((-l^t \gamma_d \gamma_s \kappa / k)) = \exp(-\gamma_d \gamma_s \kappa)$.

Proof of Corollary 1. As in proof of Theorem 1, suppose $\gamma_d \ge \varepsilon$. Claim 1 still holds as before. Now, let $\mathbf{g}^t$ be the value of $\mathbf{g}$ at the beginning of the tth iteration of the outer **for** loop with threshold value τ_t. Since the inner **for** loop is conducted in parallel, all marginal gains in iteration t are considered with respect to $\mathbf{g}^t$. Order the vectors added in this iteration $l_1 \mathbf{s}_1, \ldots, l_\ell \mathbf{s}_\ell$; because each l_i is a pivot, we know $\delta_{l_i \mathbf{s}_i}(\mathbf{g}^t) \ge l_i \tau_t$ and $\delta_{\mathbf{s}_i}(\mathbf{g}^t + l_i \mathbf{s}_i) < \tau_t$.

Let $\mathbf{g}^t_i = \mathbf{g}^t_{i-1} + l_i \mathbf{s}_i$, so $\mathbf{g}^t_0 = \mathbf{g}^t$ and $\mathbf{g}^t_\ell = \mathbf{g}^{t+1}$. Now for each i and for each $s \in S$, there exists a vector $\mathbf{h}^s_i \le \mathbf{g}^t$ such that $\delta_{\mathbf{s}}(\mathbf{h}^s_i) < \tau_t / \kappa$ (namely $\mathbf{h}^s_i = \mathbf{g}^{t-1} + l^* \mathbf{s}_i$, from when $\mathbf{s}_i$ was considered during the previous iteration $t-1$, or $\mathbf{h}^s_i = \mathbf{0}$ if $t = 1$ is the first iteration). Furthermore $\mathbf{g}^t \le \mathbf{g}^t_i$ and $\delta_{\mathbf{l}_i s_i}(\mathbf{g}^t) \ge l_i \tau_t$. Hence

$$\delta_{l_i \mathbf{s}_i}(\mathbf{g}^t_i) \ge (1-\alpha) \delta_{l_i \mathbf{s}_i}(\mathbf{g}^t) \ge (1-\alpha) l_i \tau_t \ge \kappa (1-\alpha) l_i \delta_{\mathbf{s}}(\mathbf{h}^s_i) \ge \kappa (1-\alpha) l_i \gamma_d \delta_{\mathbf{s}}(\mathbf{g}^t_i),$$

for any $s \in S$. The preceding argument proves an analogue of Claim 2, and the argument from here is exactly analogous to the proof of Theorem 1. □

E. Proofs for Section 4.2

Proof of Theorem 2. Since we have included $\gamma_d \ge \varepsilon$ as a hypothesis, we have the following claim, analogous to Claim 1.

Claim 4. If g is produced by the modified version of FastGreedy that continues until $\|\mathbf{g}\|_1 = k$, and $f(\mathbf{g}) \ge (1 - e^{-\kappa \beta^* \gamma_s}) f(\mathbf{\Omega})$, then the Theorem is proved.

Proof. Suppose $\|\mathbf{g}\|_1 < k$, and let $\mathbf{g}'$ be the solution returned by a FastGreedy* which continues updating the threshold until $\|\mathbf{g}'\|_1 = k$. Order $\{\mathbf{g}'\} \setminus \{\mathbf{g}\} = \{\mathbf{s}_1, \ldots, \mathbf{s}_\ell\}$, and let $\mathbf{g}'_i = \mathbf{g}'_{i-1} + \mathbf{s}_i$, $i = 1, \ldots, \ell$, with $\mathbf{g}'_0 = \mathbf{g}$, so that $\mathbf{g}'_\ell = \mathbf{g}'$.

Then

$$
\begin{aligned}
f(\mathbf{g}') - f(\mathbf{g}) &= \sum_{i=1}^{\ell} \delta_{\mathbf{s}_i}(\mathbf{g}'_{i-1}) \\
&\leq \frac{1}{\gamma_d} \sum_{i=1}^{\ell} \delta_{\mathbf{s}_i}(\mathbf{g}) \\
&\leq \frac{\ell}{\varepsilon} \frac{\varepsilon^2 M}{k} \leq \varepsilon M \leq \varepsilon f(\mathbf{\Omega}).
\end{aligned}
$$

Hence, for any $\Phi > \varepsilon$, if

$$f(\mathbf{g}') \geq \Phi f(\mathbf{\Omega}),$$

then

$$f(\mathbf{g}) \geq (\Phi - \varepsilon) f(\mathbf{\Omega}).$$

□

Thus, for the rest of the proof, let $\mathbf{g}$ be produced by the modified version of FastGreedy as in the hypothesis of Claim 4. Let $\mathbf{s}^t \in S$, $\mathbf{g}^t$ be the direction maximizing the marginal gain on line 3, the solution $\mathbf{g}$ immediately after the tth iteration of the **while** loop, respectively. By the choice of $\mathbf{s}^t$, for each $s \in \{\mathbf{\Omega}\}$, we have $\delta_{\mathbf{s}}(\mathbf{g}^{t-1}) \leq \delta_{\mathbf{s}^t}(\mathbf{g}^{t-1})$. Let $l_1\mathbf{s}'_1, \ldots, l_\ell \mathbf{s}'_\ell$ be the additions on line 3 to the solution $\mathbf{g}$ during iteration t, with each $l_m > 0$ for $m = 1, \ldots, \ell$. Let $\mathbf{g}_0^{t-1} = \mathbf{g}^{t-1}$ and $\mathbf{g}_m^{t-1} = \mathbf{g}_{m-1}^{t-1} + l_m \mathbf{s}'_m$. Let $L_t = \sum_{m=1}^{\ell} l_m$. Now, l_m was chosen by BinarySearchPivot and hence satisfies $\delta_{l_m \mathbf{s}'_m}(\mathbf{g}_{m-1}^{t-1}) \geq l_m \beta\kappa \delta_{\mathbf{s}^t}(\mathbf{g}^{t-1})$ by property (P1) of pivot and the choice of threshold $\tau = \beta\kappa m$. So

$$
\begin{aligned}
f(\mathbf{g}^t) - f(\mathbf{g}^{t-1}) &= \sum_{m=1}^{\ell} f(\mathbf{g}_m^{t-1}) - f(\mathbf{g}_{m-1}^{t-1}) \\
&\geq \sum_{m=1}^{\ell} l_m \beta\kappa \delta_{\mathbf{s}^t}(\mathbf{g}^{t-1}) \\
&= L_t \beta\kappa \delta_{\mathbf{s}^t}(\mathbf{g}^{t-1}) \\
&\geq \frac{L_t \beta^* \kappa}{k} \sum_{s \in \{\mathbf{\Omega} - (\mathbf{g}^{t-1} \wedge \mathbf{\Omega})\}} \delta_s(\mathbf{g}^{t-1}) \\
&\geq \frac{L_t \beta^* \gamma_s \kappa}{k} \left(f(\mathbf{\Omega}) - f(\mathbf{g}^{t-1}) \right),
\end{aligned}
$$

where the first inequality is by definition of $\mathbf{g}_m^t$, the first inequality is by the preceding paragraph, the second equality is by definition of L_t, the second inequality is by the selection of $\mathbf{s}^t$ and that fact $\|\mathbf{\Omega}\|_1 \leq k$, and the third inequality is by the definition of weak DR ratio and the lattice identity $\mathbf{v} \vee \mathbf{w} - \mathbf{v} = \mathbf{w} - \mathbf{v} \wedge \mathbf{w}$. From here,

$$f(\mathbf{g}) \geq \left(1 - \prod_{t=1}^{T} \left(1 - \frac{L_t \beta^* \gamma_s \kappa}{k} \right) \right) f(\mathbf{\Omega}),$$

from which the hypothesis of Claim 4 follows: since $(1-x) \leq e^{-x}$ and $\sum_t L^t = k$, we have $\prod_{t=1}^{T}(1 - L^t \beta^* \gamma_s \kappa / k) \leq \prod_{t=1}^{T} \exp((-L^t \beta^* \gamma_s \kappa / k)) = \exp(-\beta^* \gamma_s \kappa)$. □

Proof of Claim 3. For any i, $m'_i \leq m'_{i-1}/\gamma_d$: to see this, observe $m'_i = \max_{s \in S} \delta_{\mathbf{s}}(\mathbf{g}^i)$, $m'_{i-1} = \max_{s \in S} \delta_{\mathbf{s}}(\mathbf{g}^{i-1})$, for some $\mathbf{g}^{i-1} \leq \mathbf{g}^i$. For each $s \in S$, $\delta_{\mathbf{s}}(\mathbf{g}^{i-1}) \leq m'_{i-1}$, so $\delta_{\mathbf{s}}(\mathbf{g}^i) \leq m'_{i-1}/\gamma_d$, and hence so is m'_i. Since j_ℓ is the last uptick in the sequence before the deletion, we know for every $i > j_\ell$, $m'_i \leq \kappa m'_{i-1}$. Hence the deleted sequence proceeds from $m'_{j_\ell+1} \leq m'_\ell/\gamma_d$ down to $m'_{k_\ell - 1} \geq \kappa m'_{j_\ell}$ by definition of m'_{k_ℓ}, with each term decreasing by a factor of at least κ. □

F. Influence Maximization: A General Framework

"explicit formula for $p^{\mathbf{x}}(H,T)$". $p^{\mathbf{x}}(H,T) = Pr(H|\mathbf{x})Pr(T|\mathbf{x})$, with

$$Pr\left(H|\mathbf{x}\right) = \prod_{(x,y)\in E} p(x,y,\mathbf{x}_y)^{\mathbf{I}((x,y)\in H)}(1-p(x,y,\mathbf{x}_y))^{\mathbf{I}((x,y)\notin H)},$$
$$Pr\left(T|\mathbf{x}\right) = \prod_{y\in V} p(y,\mathbf{x}_y)^{\mathbf{I}(y\in T)}(1-p(y,\mathbf{x}_y))^{\mathbf{I}(y\notin T)}.$$

The Independent Cascade (IC) Model. The IC model is defined as follows. Given a graph $G=(V,E)$, with probabilities $p(e)$ associated to each edge $e \in E$. Let H be a realization of G, where each edge e is included in H with probability $p(e)$. Then, from an initial seed set T of activated users, a user is activated if it is reachable in H from T. Intuitively, the weight on edge (u,v) represents the probability that u activates v (*i.e* user u convinces v to adopt the product). For more information, we refer the reader to Kempe et al. (2003).

Proposition 2. *There is a natural one-to-one correspondence between instances of the IM problem under IC model and a subclass of instances of GIM.*

Proof of Proposition 2. With exactly two levels, our GIM can encapsulate the classical IM problem with the IC model (Kempe et al., 2003). Let weighted social network $G=(V,E)$ and budget k be given, as an instance of the IM problem. This instance corresponds to one of GIM with the same network and budget, as follows. For each edge $(u,v)\in E$, let $w(u,v)$ be its weight. Then we assign $p(u,v,i) = w(u,v)$ for each $i \in \{0,1\}$. Each incentive vector $\mathbf{x}$ is a binary vector, indicating which users are present in the seed set; *i.e.* $p(u,0)=0$ and $p(u,1)=1$, for all $u\in V$. This mapping is injective and hence invertible. □

Proposition 3. *There is a natural one-to-one correspondence between instances of the boosting problem and a subclass of instances of GIM.*

Proof of Proposition 3. Let social network $G=(V,E)$, seed set S, and $k\in\mathbb{N}$ be given as an instance of the boosting problem, where edge $(u,v)\in E$ has weight $p(u,v)$ if v is not boosted, and weight $p'(u,v)$ if v is boosted. The corresponding instance of GIM has two levels. Set $p(y,0)=p(y,1)=1$ for all $y\in S$ and set $p(y,0)=p(y,1)=0$ for all $y\notin S$. For each edge $(u,v)\in E$, set $p(u,v,0)=p(u,v)$, $p(u,v,1)=p'(u,v)$. Hence, spending budget to incentivize a node from level 0 to level 1 does not affect the initial seed set, which is always S. But this incentive does work in exactly the same way as the boosting of a node by changing its incoming edge probabilities; hence, the objective values are the same. This mapping is injective and hence invertible. □

Proof of Theorem 3.

Claim 5. Suppose $\gamma\delta_{\mathbf{s}}(\mathbf{w}) \le \delta_{\mathbf{s}}(\mathbf{v})$, where $\gamma = c_e^{-k\Delta}c_n^{-k}$, and $\mathbf{v},\mathbf{w}$ are any vectors satisfying $\mathbf{v}\le\mathbf{w}$, $\|\mathbf{v}\|_1\le k$, and $\|\mathbf{w}-\mathbf{v}\|_1 \le k$. Then the result of Theorem 3 follows.

Proof. Suppose the hypothesis of the claim holds.

Greedy DR ratios. We will show $\gamma \le \gamma_s^{\mathcal{A},\mathcal{I}}$, where $\gamma_s^{\mathcal{A},\mathcal{I}}$ is the FastGreedy weak DR ratio on instance $\mathcal{I}$. The proofs for the other greedy DR ratios are exactly analogous.

Let $\mathbf{g}^1,\ldots,\mathbf{g}^T$ be the greedy vectors in the definition of FastGreedy weak DR ratio. Let $i\in\{1,\ldots,T\}$, and let $\mathbf{v}=\mathbf{g}^i$. Let $\mathbf{w}\ge\mathbf{v}\in\mathbb{N}^S$ such that $\|\mathbf{w}-\mathbf{v}\|_1\le k$. Let $\{\mathbf{w}-\mathbf{v}\}=\{s_1,...,s_l\}$. Then $\mathbf{v}+\sum_{j=1}^{l}\mathbf{s}_j=\mathbf{w}$. In addition, for every

$m \leq l$, $\mathbf{v} + \sum_{j=1}^{m} \mathbf{s}_j = \mathbf{v}_m$ where $\mathbf{v}_m \in \mathbb{N}^S$, $\mathbf{v}_m \leq \mathbf{w}$. Then,

$$\begin{aligned} \gamma(f(\mathbf{w}) - f(\mathbf{v})) &= \gamma \sum_{j=1}^{l} f(\mathbf{v} + \mathbf{s}_1 + ... + \mathbf{s}_j) - f(\mathbf{v} + \mathbf{s}_1 + ... + \mathbf{s}_{j-1}) \\ &= \gamma \sum_{j=1}^{l} \delta_{\mathbf{s}_j}(\mathbf{v} + \mathbf{s}_1 + ... + \mathbf{s}_{j-1}) \\ &= \gamma \sum_{j=1}^{l} \delta_{\mathbf{s}_j}(\mathbf{v}_{j-1}) \\ &\leq \sum_{j=1}^{l} \delta_{\mathbf{s}_j}(\mathbf{v}), \end{aligned}$$

by the hypothesis of the claim. Therefore $\gamma \leq \gamma_s^{\mathcal{A},\mathcal{I}}$, since the latter is the maximum number satisfying the above inequality for each $\mathbf{w}, \mathbf{g}^i$ as above.

FastGreedy DR ratio β^*. Initally, $\beta = 1$; it decreases by a factor of $\delta \in (0, 1)$ at most once per iteration of the **while** loop of FastGreedy. Suppose $\beta \leq \gamma$ for some iteration i of the **while** loop, and let $\mathbf{g}$ have the value assigned immediately after iteration i, m have the value assigned after line 3 of iteration i. Then Since a valid pivot was found for each $s \in S$ during iteration i, by property (P2) there exists $\mathbf{g}^s \leq \mathbf{g}$, $\delta_{\mathbf{s}}(\mathbf{g}^s) < \beta\kappa m \leq \gamma\kappa m$. Hence $\delta_{\mathbf{s}}(\mathbf{g}) \leq \kappa m$, since $\mathbf{g}$, $\mathbf{g}^s$ are vectors satisfying the conditions on γ in the hypothesis of the claim. In iteration $i+1$, m' has the value of m from iteration i, so the value of m computed during iteration $i + 1$ is at most $\kappa m'$, and β does not decrease during iteration i. It follows that $\beta^* \geq \gamma\delta$. □

Let $\mathbf{v} \leq \mathbf{w}$, $\|\mathbf{v} - \mathbf{w}\|_1 \leq k$. We will consider graph realizations H that have the status of all edges determined; and seed sets $T \subseteq V$.

Then,

$$p^{\mathbf{w}+\mathbf{s}}(H, T) = K_1(H, T)p^{\mathbf{v}+\mathbf{s}}(H, T), \text{ and} \tag{3}$$

$$p^{\mathbf{w}}(H, T) = K_2(H, T)p^{\mathbf{v}}(H, T), \tag{4}$$

where $K_1(H, T) = K_1(H)K_1(T)$, $K_2(H, T) = K_2(H)K_2(T)$ with

$$K_2(H) = \prod_{(x,y)\in E} \left(\frac{p(x, y, \mathbf{w}_y)}{p(x, y, \mathbf{v}_y)}\right)^{\mathbf{I}((x,y)\in H)} \left(\frac{1 - p(x, y, \mathbf{w}_y)}{1 - p(x, y, \mathbf{v}_y)}\right)^{\mathbf{I}((x,y)\notin H)},$$

$$K_2(T) = \prod_{x\in V} \left(\frac{p(x, \mathbf{w}_x)}{p(x, \mathbf{v}_x)}\right)^{\mathbf{I}(x\in T)} \left(\frac{1 - p(x, \mathbf{w}_x)}{1 - p(x, \mathbf{v}_x)}\right)^{\mathbf{I}(x\notin T)},$$

and the definitions of $K_1(H), K_1(T)$ are analogous to the above with vectors $\mathbf{v} + \mathbf{s}, \mathbf{w} + \mathbf{s}$ in place of $\mathbf{v}, \mathbf{w}$.

Lemma 3. Let Δ be the maximum in-degree in G.

$$K_1(H, T) \leq K_2(H, T) \leq c_e^{k\Delta} c_n^k$$

Proof. *(a)* $K_1(T) \leq K_2(T)$: if $s \in T$,

$$\begin{aligned} K_1(T) &= K_2(T) \cdot \frac{p(s, \mathbf{v}_s)}{p(s, \mathbf{w}_s)} \cdot \frac{p(s, \mathbf{w}_s + 1)}{p(s, \mathbf{v}_s + 1)} \\ &= K_2(T) \cdot \frac{z}{z'} \cdot \frac{z' + \alpha'}{z + \alpha} \leq K_2(T), \end{aligned}$$

where $\alpha' \leq \alpha$ by DR-submodularity of $i \mapsto p(s,i)$, and $p(s,\mathbf{v}_s) = z \leq z' = p(s,\mathbf{w}_s)$ by monotonicity of the same mapping. Otherwise, if $s \notin T$,

$$K_1(T) = K_2(T) \cdot \frac{1 - p(s,\mathbf{v}_s)}{1 - p(s,\mathbf{w}_s)} \cdot \frac{1 - p(s,\mathbf{w}_s + 1)}{1 - p(s,\mathbf{v}_s + 1)}$$
$$= K_2(T) \cdot \frac{z}{z'} \cdot \frac{z' - \alpha'}{z - \alpha} \leq K_2(T),$$

where as before $\alpha' \leq \alpha$ by DR-submodularity, but $1 - p(s,\mathbf{v}_s) = z \geq z' = 1 - p(s,\mathbf{w}_s)$.

(b) $K_1(H) \leq K_2(H)$: if $(u,s) \in H$,

$$K_1(H) = K_2(H) \cdot \frac{p(u,s,\mathbf{v}_s)}{p(u,s,\mathbf{w}_s)} \cdot \frac{p(u,s,\mathbf{w}_s + 1)}{p(u,s,\mathbf{v}_s + 1)}$$
$$= K_2(H) \cdot \frac{z}{z'} \cdot \frac{z' + \alpha'}{z + \alpha} \leq K_2(H),$$

where $\alpha' \leq \alpha$ by DR-submodularity of $i \mapsto p(u,s,i)$, and $p(u,s,\mathbf{v}_s) = z \leq z' = p(u,s,\mathbf{w}_s)$ by monotonicity of the same mapping. Otherwise, if $(u,s) \notin H$,

$$K_1(H) = K_2(H) \cdot \frac{1 - p(u,s,\mathbf{v}_s)}{1 - p(u,s,\mathbf{w}_s)} \cdot \frac{1 - p(u,s,\mathbf{w}_s + 1)}{1 - p(u,s,\mathbf{v}_s + 1)}$$
$$= K_2(H) \cdot \frac{z}{z'} \cdot \frac{z' - \alpha'}{z - \alpha} \leq K_2(H),$$

where as before $\alpha' \leq \alpha$ by DR-submodularity, but $1 - p(u,s,\mathbf{v}_s) = z \geq z' = 1 - p(u,s,\mathbf{w}_s)$.

(c) $K_2(H,T) \leq c_e^{k\Delta} c_n^k$:

$$K_2(H,T) = K_2(H)K_2(T) = \prod_{(x,y) \in E} \left(\frac{p(x,y,\mathbf{w}_y)}{p(x,y,\mathbf{v}_y)} \right)^{\mathbf{I}((x,y) \in H)} \left(\frac{1 - p(x,y,\mathbf{w}_y)}{1 - p(x,y,\mathbf{v}_y)} \right)^{\mathbf{I}((x,y) \notin H)}$$
$$\prod_{x \in V} \left(\frac{p(x,\mathbf{w}_x)}{p(x,\mathbf{v}_x)} \right)^{\mathbf{I}(x \in T)} \left(\frac{1 - p(x,\mathbf{w}_x)}{1 - p(x,\mathbf{v}_x)} \right)^{\mathbf{I}(x \notin T)}$$

Each of the fractions in the above product is of the form $\xi(\mathbf{w}_y)/\xi(\mathbf{v}_y)$, where $\mathbf{w}_y \geq \mathbf{v}_y$ and hence can be written

$$\frac{\xi(\mathbf{w}_y)}{\xi(\mathbf{v}_y)} = \prod_{i=1}^{\mathbf{w}_y - \mathbf{v}_y} \frac{\xi(\mathbf{v}_y + i + 1)}{\xi(\mathbf{v}_y + i)} \leq (\mathbf{w}_y - \mathbf{v}_y) \max_j \frac{\xi(j+1)}{\xi(j)}.$$

Hence, by the fact that $\|\mathbf{w} - \mathbf{v}\|_1 \leq k$ and the definitions of c_e, c_n, and the maximum in-degree Δ in G, we have

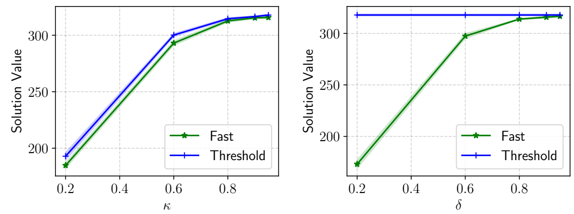
$$K_2(H,T) \leq c_e^{k\Delta} c_n^k.$$

□

Finally, by Lemma 3, we have

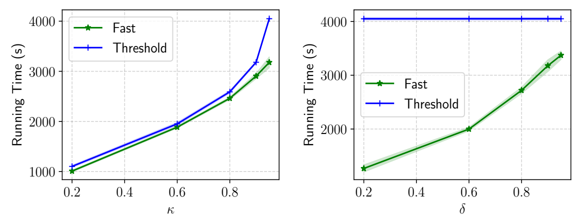
$$\mathbb{A}(\mathbf{w} + \mathbf{s}) - \mathbb{A}(\mathbf{w}) = \sum_{H,T} \left(p^{\mathbf{w}+\mathbf{s}}(H,T) - p^{\mathbf{w}}(H,T) \right) R(H,T)$$
$$= \sum_{H,T} \left(K_1(H,T) p^{\mathbf{v}+\mathbf{s}}(H,T) - K_2(H,T) p^{\mathbf{v}}(H,T) \right) R(H,T)$$
$$\leq \sum_{H,T} K_2(H,T) \left(p^{\mathbf{v}+\mathbf{s}}(H,T) - p^{\mathbf{v}}(H,T) \right) R(H,T)$$
$$\leq {c_e}^{k\Delta} c_n^k \left(\mathbb{A}(\mathbf{v} + \mathbf{s}) - \mathbb{A}(\mathbf{v}) \right).$$

Therefore, the hypothesis of Claim 5 is satisfied, and the result follows. □



(a) κ (b) δ

Figure 5. Performance as δ and κ are varied. Note that ThresholdGreedy does not use δ.



(a) κ (b) δ

Figure 6. Running time as δ and κ are varied. Note that ThresholdGreedy does not use δ.

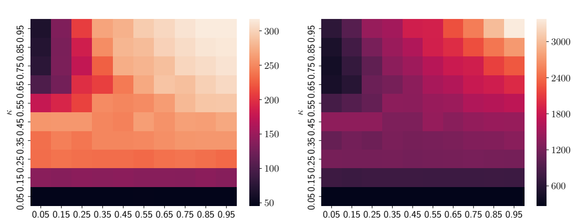
G. Additional Experimental Results

G.1. Characterizing the Parameters of FastGreedy

In this section, we evaluate the impact of varying the parameters of FastGreedy: ε, δ, and κ. We note that ε only impacts the running time and performance if it is used as a stopping condition. However, in our experiments this did not occur: the cardinality constraint was reached first. Therefore, in our experiments FastGreedy = FastGreedy* and we are free to set $\varepsilon = 0$ without changing any of our results.

Figures 5 & 6 show the impact of δ & κ on performance and running time. Note that performance remains similar until κ or δ drops below 0.6. However, the running time plummets to nearly half of what it is at 0.95 in each case, resulting in a similar-quality solution in significantly less time. A natural follow-up question from this figures is: what happens when both parameters are varied at once? Fig. 7 details the answer. In particular, we note that the steep drop in running time remains present, and there is a reasonable gain to be had by dropping both parameters at once – up to a point.

We further notice some interesting patterns in these heat-maps. When κ is near 0, the selection of δ does not appear to matter. This is likely due to the role each parameter plays: κ plays a critical role in identifying a DR-violation, at which point δ is the rate by which β is reduced to compensate. When κ is small, it takes commensurately larger violations for δ to apply. These larger violations are not seen in our simulations, and thus δ has no impact when κ is very small.



(a) Solution Quality (b) Running Time

Figure 7. Performance and running time of FastGreedy as κ and δ are simultaneously varied.

H. Implementation Notes

The `facebook` dataset originally is undirected; we replace each edge $u \leftrightarrow v$ with two edges $u \to v$ and $v \to u$.

We calculated the FastGreedy ratio using the values $\kappa = 0.95$, $\beta^* = 0.9$, $\gamma_s = 0.69857$, $\varepsilon = 0$. The values of κ and ε are taken from the parameters used to run the algorithm (ε can be 0 since the algorithm always returned $\mathbf{g}$ with $\|\mathbf{g}\|_1 = k$), and the values of β^* and γ_s are the minimum over all instances where FastGreedy Submodularity Ratio was computed.

H.1. Evaluating $\delta_{\mathbf{s}}(\mathbf{g})$

As mentioned in Section 6, we evaluate the objective on a set of 10,000 Monte Carlo samples. However, for performance reasons we do not evaluate the marginal gain $\delta_{\mathbf{s}}(\mathbf{g})$ directly by computing each of $f(\mathbf{g}+\mathbf{s})$ and $f(\mathbf{g})$ and then subtracting. Instead, we compute the expected number of activations across the sample set were $\mathbf{s}$ to be added. We accomplish this as follows.

First, we mantain a state associated with the vector $\mathbf{g}$. This state contains a number of variables for logging purposes, in addition to two of note for our discussion here: `samples` and `active`. The former is the list of sampled graphs, each represented as a pair of vectors of floating point numbers corresponding to random thresholds assigned to each node and edge. The latter is a list of sets of nodes currently active in each sampled graph under solution vector $\mathbf{g}$. These active sets are computed only when a new element (or elements) are added to the solution vector and are computed directly by (a) computing the set of externally activated nodes by checking if the random threshold is sufficient to activate the node; then (b) propagating across any active edges according to their thresholds. The code for this is contained in the `active_nodes` function of `src/bin/inf.rs` in the code distribution.

Given this representation, to estimate the marginal gain of ℓ copies of a node s *on each sample* S_i as follows:

1. Check if the node is already active on S_i. If so, return 0 for this sample.
2. Check if the node would be externally activated if added to the solution. If not, return 0.
3. Check if the node would be activated by neighboring nodes if added to the solution. If not, return 0.
4. Compute the set of nodes that would be newly activated if s becomes active via breadth-first-search from s. Return this number c.

Each would-be-activated check is accomplished by comparing the activation probability of the node or edge to the random threshold associated with it in the sample. Then the expected marginal gain is the average result across all samples S_i. When an element (with multiplicity) is chosen to be inserted into the solution, we add it to the solution vector, discard and recompute all samples, and then recompute the `active` set on each sample from scratch. This is in line with prior Monte-Carlo-based solutions for IM (Kempe et al., 2003).

While in theory it is possible to consider the marginal gain of an arbitrary vector $\mathbf{v}$, in our implementation we restrict the values it can take to $\mathbf{v} = \ell\mathbf{s}$, where $\mathbf{s}$ is the unit vector for node s. This simplifies each of the steps above. The implementation of the above is contained in the functions `delta` and `scaled_delta` in `src/bin/inf.rs`.