A Supplementary Material

Throughout our derivations, if it is clear from context, we omit the argument (t) indexing time, for example writing Y_u instead of $Y_u(t)$.

A.1 Proof of Lemma 1

We reproduce Lemma 1 below for ease of presentation.

Lemma 1. *For user u, after*

$$t \geq \Big(\frac{2\log(10kmn^\alpha/\Delta)}{\Delta^4(1-\gamma)^2}\Big)^{1/(1-\alpha)}$$

time steps,

$$\mathbb{P}(\mathcal{E}_{good}(u,t)) \geq 1 - \exp\Big(-\frac{n}{8k}\Big) - 12\exp\Big(-\frac{\Delta^4(1-\gamma)^2 t^{1-\alpha}}{20}\Big).$$

To derive this lower bound on the probability that the good neighborhood event $\mathcal{E}_{\text{good}}(u,t)$ occurs, we prove four lemmas (Lemmas 4, 5, 6, and 7). Before doing so, we define a constant that will appear several times:

$$\beta \triangleq \exp(-\Delta^4(1-\gamma)^2 t^{1-\alpha}).$$

We begin by ensuring that enough users from each of the k user types are in the system.

Lemma 4. *For a user u,*

$$\mathbb{P}\Big(\textit{user } u\textit{'s type has} \leq \frac{n}{2k} \textit{ users}\Big) \leq \exp\Big(-\frac{n}{8k}\Big).$$

Proof. Let N be the number of users from user u's type. User types are equiprobable, so $N \sim \text{Bin}(n, \frac{1}{k})$. By a Chernoff bound,

$$\mathbb{P}\Big(N \leq \frac{n}{2k}\Big) \leq \exp\Big(-\frac{1}{2}\frac{(\frac{n}{k}-\frac{n}{2k})^2}{\frac{n}{k}}\Big) = \exp\Big(-\frac{n}{8k}\Big). \qquad \square$$

Next, we ensure that sufficiently many items have been jointly explored across all users. This will subsequently be used for bounding both the number of good neighbors and the number of bad neighbors.

Lemma 5. *After t time steps,*

$$\mathbb{P}(\textit{fewer than } t^{1-\alpha}/2 \textit{ jointly explored items}) \leq \exp(-t^{1-\alpha}/20).$$

Proof. Let Z_s be the indicator random variable for the event that the algorithm jointly explores at time s. Thus, the number of jointly explored items up to time t is $\sum_{s=1}^t Z_s$. By our choice for the time-varying joint exploration probability ε_J, we have $\mathbb{P}(Z_s = 1) = \varepsilon_J(s) = \frac{1}{s^\alpha}$ and $\mathbb{P}(Z_s = 0) = 1 - \frac{1}{s^\alpha}$. Note that the centered random variable $\bar{Z}_s = \mathbb{E}[Z_s] - Z_s = \frac{1}{s^\alpha} - Z_s$ has zero mean, and $|\bar{Z}_s| \leq 1$ with probability 1. Then,

$$\begin{aligned}\mathbb{P}\Big(\sum_{s=1}^t Z_s < \frac{1}{2}t^{1-\alpha}\Big) &= \mathbb{P}\Big(\sum_{s=1}^t \bar{Z}_s > \sum_{s=1}^t \mathbb{E}[Z_s] - \frac{1}{2}t^{1-\alpha}\Big) \overset{(i)}{\leq} \mathbb{P}\Big(\sum_{s=1}^t \bar{Z}_s > \frac{1}{2}t^{1-\alpha}\Big)\\ &\overset{(ii)}{\leq} \exp\Big(-\frac{\frac{1}{8}t^{2(1-\alpha)}}{\sum_{s=1}^t \mathbb{E}[\bar{Z}_s^2] + \frac{1}{6}t^{1-\alpha}}\Big) \overset{(iii)}{\leq} \exp\Big(-\frac{\frac{1}{8}t^{2(1-\alpha)}}{\frac{t^{1-\alpha}}{1-\alpha} + \frac{1}{6}t^{1-\alpha}}\Big)\\ &= \exp\Big(-\frac{3(1-\alpha)t^{1-\alpha}}{4(7-\alpha)}\Big) \overset{(iv)}{\leq} \exp(-t^{1-\alpha}/20),\end{aligned}$$

where step (i) uses the fact that $\sum_{s=1}^t \mathbb{E}[Z_s] = \sum_{s=1}^t 1/s^\alpha \geq t/t^\alpha = t^{1-\alpha}$, step (ii) is Bernstein's inequality, step (iii) uses the fact that $\sum_{s=1}^t \mathbb{E}[\bar{Z}_s^2] \leq \sum_{s=1}^t \mathbb{E}[Z_s^2] = \sum_{s=1}^t 1/s^\alpha \leq t^{1-\alpha}/(1-\alpha)$, and step (iv) uses the fact that $\alpha \leq 4/7$. (We remark that the choice of constant $4/7$ isn't special; changing it would simply modify the constant in the decaying exponentially to potentially no longer be $1/20$). $\square$

Assuming that the bad events for the previous two lemmas do not occur, we now provide a lower bound on the number of good neighbors that holds with high probability.

Lemma 6. *Suppose that there are no Δ-ambiguous items, that there are more than $\frac{n}{2k}$ users of user u's type, and that all users have rated at least $t^{1-\alpha}/2$ items as part of joint exploration. For user u, let n_{good} be the number of "good" neighbors of user u. If $\beta \leq \frac{1}{10}$, then*

$$\mathbb{P}\Big(n_{good} \leq (1-\beta)\frac{n}{4k}\Big) \leq 10\beta.$$

We defer the proof of Lemma 6 to Appendix A.1.1.

Finally, we verify that the number of bad neighbors for any user is not too large, again conditioned on there being enough jointly explored items.

Lemma 7. *Suppose that the minimum number of rated items in common between any pair of users is $t^{1-\alpha}/2$ and suppose that γ-incoherence holds for some $\gamma \in [0,1)$. For user u, let n_{bad} be the number of "bad" neighbors of user u. Then*

$$\mathbb{P}(n_{bad} \geq n\sqrt{\beta}) \leq \sqrt{\beta}.$$

We defer the proof of Lemma 7 to Appendix A.1.2.

We now prove Lemma 1, which union bounds over the four bad events of Lemmas 4, 5, 6, and 7. Recall that the good neighborhood event $\mathcal{E}_{\text{good}}(u,t)$ holds if at time t, user u has more than $\frac{n}{5k}$ good neighbors and less than $\frac{\Delta t n^{1-\alpha}}{10km}$ bad neighbors. By assuming that the four bad events don't happen, then Lemma 6 tells us that there are more than $(1-\beta)\frac{n}{4k}$ good neighbors provided that $\beta \leq \frac{1}{10}$. Thus, to ensure that there are more than $\frac{n}{5k}$ good neighbors, it suffices to have $(1-\beta)\frac{n}{4k} \geq \frac{n}{5k}$, which happens when $\beta \leq \frac{1}{5}$, but we already require that $\beta \leq \frac{1}{10}$. Similarly, Lemma 7 tells us that there are fewer than $n\sqrt{\beta}$ bad neighbors, so to ensure that there are fewer than $\frac{\Delta t n^{1-\alpha}}{10km}$ bad neighbors it suffices to have $n\sqrt{\beta} \leq \frac{\Delta t n^{1-\alpha}}{10km}$, which happens when $\beta \leq (\frac{\Delta t}{10kmn^{\alpha}})^2$. We can satisfy all constraints on β by asking that $\beta \leq (\frac{\Delta}{10kmn^{\alpha}})^2$, which is tantamount to asking that

$$t \geq \left(\frac{2\log(10kmn^{\alpha}/\Delta)}{\Delta^4(1-\gamma)^2}\right)^{1/(1-\alpha)}$$

since $\beta = \exp(-\Delta^4(1-\gamma)^2 t^{1-\alpha})$.

Finally, with t satisfying the inequality above, the union bound over the four bad events can be further bounded to complete the proof:

$$\begin{aligned}\mathbb{P}(\mathcal{E}_{\text{good}}(u,t)) &\geq 1 - \exp\Big(-\frac{n}{8k}\Big) - \exp(-t^{1-\alpha}/20) - 10\beta - \sqrt{\beta} \\ &\geq 1 - \exp\Big(-\frac{n}{8k}\Big) - 12\exp\Big(-\frac{\Delta^4(1-\gamma)^2 t^{1-\alpha}}{20}\Big).\end{aligned}$$

A.1.1 Proof of Lemma 6

We begin with a preliminary lemma that upper-bounds the probability of two users of the same type not being declared as neighbors.

Lemma 8. *Suppose that there are no Δ-ambiguous items for any of the user types. Let users u and v be of the same type, and suppose that they have rated at least Γ_0 items in common (explored jointly). Then for $\theta \in (0, 4\Delta^2)$,*

$$\mathbb{P}(\textit{users } u \textit{ and } v \textit{ are not declared as neighbors}) \leq \exp\Big(-\frac{(4\Delta^2-\theta)^2}{2}\Gamma_0\Big).$$

Proof. Let us first suppose that users u and v have rated exactly Γ_0 items in common. The two users are not declared to be neighbors if $\langle \widetilde{Y}_u, \widetilde{Y}_v \rangle < \theta\Gamma_0$. Let $\Omega \subseteq [m]$ such that $|\Omega| = \Gamma_0$. We have

$$\begin{aligned}
\mathbb{E}\big[\langle \widetilde{Y}_u, \widetilde{Y}_v \rangle \big| \text{supp}(\widetilde{Y}_u) \cap \text{supp}(\widetilde{Y}_v) = \Omega\big] &= \sum_{i\in\Omega} \mathbb{E}[\widetilde{Y}_{ui}\widetilde{Y}_{vi} \mid \widetilde{Y}_{ui} \neq 0, \widetilde{Y}_{vi} \neq 0] \\
&= \sum_{i\in\Omega} (p_{ui}^2 + (1-p_{ui})^2 - 2p_{ui}(1-p_{ui})) \\
&= 4\sum_{i\in\Omega} \Big(p_{ui} - \frac{1}{2}\Big)^2. \qquad (2)
\end{aligned}$$

Since $\langle \widetilde{Y}_u, \widetilde{Y}_v \rangle = \sum_{i\in\Omega} \widetilde{Y}_{ui}\widetilde{Y}_{vi}$ is the sum of terms $\{\widetilde{Y}_{ui}\widetilde{Y}_{vi}\}_{i\in\Omega}$ that are each bounded within $[-1, 1]$, Hoeffding's inequality yields

$$\mathbb{P}\big(\langle \widetilde{Y}_u, \widetilde{Y}_v \rangle \le \theta\Gamma_0 \;\big|\; \text{supp}(\widetilde{Y}_u) \cap \text{supp}(\widetilde{Y}_v) = \Omega\big) \le \exp\Bigg(-\frac{\big[\overbrace{4\sum_{i\in\Omega}\big(p_{gi} - \frac{1}{2}\big)^2}^{\text{equation (2)}} - \theta\Gamma_0\big]^2}{2\Gamma_0}\Bigg). \quad (3)$$

As there are no Δ-ambiguous items, $\Delta \le |p_{ui} - 1/2|$ for all users u and items i. Thus, our choice of θ guarantees that

$$4\sum_{i\in\Omega}\Big(p_{ui} - \frac{1}{2}\Big)^2 - \theta\Gamma_0 \ge 4\Gamma_0\Delta^2 - \theta\Gamma_0 = (4\Delta^2 - \theta)\Gamma_0 \ge 0. \quad (4)$$

Combining inequalities (3) and (4), and observing that the above holds for all subsets Ω of cardinality Γ_0, we obtain the desired bound on the probability that users u and v are not declared as neighbors:

$$\mathbb{P}(\langle \widetilde{Y}_u, \widetilde{Y}_v \rangle \le \theta\Gamma_0 \mid |\text{supp}(\widetilde{Y}_u) \cap \text{supp}(\widetilde{Y}_v)| = \Gamma_0) \le \exp\Big(-\frac{(4\Delta^2-\theta)^2}{2}\Gamma_0\Big). \quad (5)$$

Now to handle the case that users u and v have jointly rated more than Γ_0 items, observe that, with shorthand $\Gamma_{uv} \triangleq |\text{supp}(\widetilde{Y}_u) \cap \text{supp}(\widetilde{Y}_v)|$,

$$\begin{aligned}
&\mathbb{P}(u \text{ and } v \text{ not declared neighbors} \mid p_u = p_v, \Gamma_{uv} \ge \Gamma_0) \\
&= \mathbb{P}(\langle \widetilde{Y}_u, \widetilde{Y}_v \rangle < \theta\Gamma_{uv} \mid p_u = p_v,\ \Gamma_{uv} \ge \Gamma_0) \\
&= \frac{\mathbb{P}(\langle \widetilde{Y}_u, \widetilde{Y}_v \rangle \le \theta\Gamma_{uv}, \Gamma_{uv} \ge \Gamma_0 \mid p_u = p_v)}{\mathbb{P}(\Gamma_{uv} \ge \Gamma_0 \mid p_u = p_v)} \\
&= \frac{\sum_{\ell=\Gamma_0}^{m} \mathbb{P}(\langle \widetilde{Y}_u, \widetilde{Y}_v \rangle \le \theta\ell, \Gamma_{uv} = \ell \mid p_u = p_v)}{\mathbb{P}(\Gamma_{uv} \ge \Gamma_0 \mid p_u = p_v)} \\
&= \frac{\sum_{\ell=\Gamma_0}^{m} \begin{array}{l}\big[\mathbb{P}(\Gamma_{uv} = \ell \mid p_u = p_v) \\ \quad \cdot \mathbb{P}(\langle \widetilde{Y}_u, \widetilde{Y}_v \rangle \le \theta\ell \mid p_u = p_v, \Gamma_{uv} = \ell)\big]\end{array}}{\mathbb{P}(\Gamma_{uv} \ge \Gamma_0 \mid p_u = p_v)} \\
&\le \frac{\sum_{\ell=\Gamma_0}^{m} \mathbb{P}(\Gamma_{uv} = \ell \mid p_u = p_v) \exp\big(-\frac{(4\Delta^2-\theta)^2}{2}\Gamma_0\big)}{\mathbb{P}(\Gamma_{uv} \ge \Gamma_0 \mid p_u = p_v)} \\
&\quad \text{by inequality (5)} \\
&= \exp\Big(-\frac{(4\Delta^2-\theta)^2}{2}\Gamma_0\Big).
\end{aligned}$$

□

We now prove Lemma 6.

Suppose that the event in Lemma 4 holds. Let $\mathcal{G}$ be $\frac{n}{2k}$ users from the same user type as user u; there could be more than $\frac{n}{2k}$ such users but it suffices to consider $\frac{n}{2k}$ of them. We define an indicator random variable

$$G_v \triangleq \mathbb{1}\{\text{users } u \text{ and } v \text{ are neighbors}\} = \mathbb{1}\{\langle \widetilde{Y}_u^{(t)}, \widetilde{Y}_v^{(t)} \rangle \ge \theta t^{1-\alpha}/2\}.$$

Thus, the number of good neighbors of user u is lower-bounded by $W = \sum_{v \in \mathcal{G}} G_v$. Note that the G_v's are not independent. To arrive at a lower bound for W that holds with high probability, we use Chebyshev's inequality:

$$\mathbb{P}(W - \mathbb{E}[W] \le -\mathbb{E}[W]/2) \le \frac{4\text{Var}(W)}{(\mathbb{E}[W])^2}. \tag{6}$$

Let $\beta = \exp(-(4\Delta^2 - \theta)^2 \Gamma_0/2)$ be the probability bound from Lemma 8, where by our choice of $\theta = 2\Delta^2(1+\gamma)$ and with $\Gamma_0 = t^{1-\alpha}/2$, we have $\beta = \exp(-\Delta^4(1-\gamma)^2 t^{1-\alpha})$.

Applying Lemma 8, we have $\mathbb{E}[W] \ge (1-\beta)\frac{n}{2k}$, and hence

$$(\mathbb{E}[W])^2 \ge (1 - 2\beta)\frac{n^2}{4k^2}. \tag{7}$$

We now upper-bound

$$\text{Var}(W) = \sum_{v \in \mathcal{G}} \text{Var}(G_v) + \sum_{v \neq w} \text{Cov}(G_v, G_w).$$

Since $G_v = G_v^2$,

$$\text{Var}(G_v) = \mathbb{E}[G_v] - \mathbb{E}[G_v]^2 = \underbrace{\mathbb{E}[G_v]}_{\le 1}(1 - \mathbb{E}[G_v]) \le \beta,$$

where the last step uses Lemma 8.

Meanwhile,

$$\text{Cov}(G_v, G_w) = \mathbb{E}[G_v G_w] - \mathbb{E}[G_v]\mathbb{E}[G_w] \le 1 - (1-\beta)^2 \le 2\beta.$$

Putting together the pieces,

$$\text{Var}(W) \le \frac{n}{2k} \cdot \beta + \frac{n}{2k} \cdot \left(\frac{n}{2k} - 1\right) \cdot 2\beta \le \frac{n^2}{2k^2} \cdot \beta. \tag{8}$$

Plugging (7) and (8) into (6) gives

$$\mathbb{P}(W - \mathbb{E}[W] \le -\mathbb{E}[W]/2) \le \frac{8\beta}{1 - 2\beta} \le 10\beta,$$

provided that $\beta \le \frac{1}{10}$. Thus, $n_{\text{good}} \ge W \ge \mathbb{E}[W]/2 \ge (1-\beta)\frac{n}{4k}$ with probability at least $1 - 10\beta$.

A.1.2 Proof of Lemma 7

We begin with a preliminary lemma that upper-bounds the probability of two users of different types being declared as neighbors.

Lemma 9. *Let users u and v be of different types, and suppose that they have rated at least Γ_0 items in common via joint exploration. Further suppose γ-incoherence is satisfied for $\gamma \in [0, 1)$. If $\theta \ge 4\gamma\Delta^2$, then*

$$\mathbb{P}(\textit{users } u \textit{ and } v \textit{ are declared to be neighbors}) \le \exp\Big(-\frac{(\theta - 4\gamma\Delta^2)^2}{2}\Gamma_0\Big).$$

Proof. As with the proof of Lemma 8, we first analyze the case where users u and v have rated exactly Γ_0 items in common. Users u and v are declared to be neighbors if $\langle \widetilde{Y}_u, \widetilde{Y}_v \rangle \ge \theta\Gamma_0$. We now crucially use the fact that joint exploration chooses these Γ_0 items as a random subset of the m items. For our random permutation σ of m items, we have $\langle \widetilde{Y}_u, \widetilde{Y}_v \rangle = \sum_{i=1}^{\Gamma_0} \widetilde{Y}_{u,\sigma(i)} \widetilde{Y}_{v,\sigma(i)} = \sum_{i=1}^{\Gamma_0} Y_{u,\sigma(i)} Y_{v,\sigma(i)}$, which is the sum of terms $\{Y_{u,\sigma(i)} Y_{v,\sigma(i)}\}_{i=1}^{\Gamma_0}$ that are each bounded within $[-1, 1]$ and drawn without replacement from a population of all possible items. Hoeffding's inequality (which also applies to the current scenario of sampling without replacement [14]) yields

$$\mathbb{P}\big(\langle \widetilde{Y}_u, \widetilde{Y}_v \rangle \ge \theta\Gamma_0 \mid p_u \neq p_v\big) \le \exp\left(-\frac{\big(\theta\Gamma_0 - \mathbb{E}[\langle \widetilde{Y}_u, \widetilde{Y}_v \rangle \mid p_u \neq p_v]\big)^2}{2\Gamma_0}\right). \tag{9}$$

By γ-incoherence and our choice of θ,

$$\theta\Gamma_0 - \mathbb{E}\big[\langle \widetilde{Y}_u, \widetilde{Y}_v\rangle \mid p_u \neq p_v\big] \geq \theta\Gamma_0 - 4\gamma\Delta^2\Gamma_0 = (\theta - 4\gamma\Delta^2)\Gamma_0 \geq 0. \tag{10}$$

Above, we used the fact that Γ_0 randomly explored items are a random subset of m items, and hence

$$\mathbb{E}\big[\tfrac{1}{\Gamma_0}\langle \widetilde{Y}_u, \widetilde{Y}_v\rangle\big] = \mathbb{E}\big[\tfrac{1}{m}\langle Y_u, Y_v\rangle\big],$$

with Y_u, Y_v representing the entire (random) vector of preferences of u and v respectively.

Combining inequalities (9) and (10) yields

$$\mathbb{P}\big(\langle \widetilde{Y}_u, \widetilde{Y}_v\rangle \geq \theta\Gamma_0 \mid p_u \neq p_v\big) \leq \exp\Big(-\frac{(\theta - 4\gamma\Delta^2)^2}{2}\Gamma_0\Big).$$

A similar argument as the ending of Lemma 8's proof establishes that the bound holds even if users u and v have jointly explored more than Γ_0 items. $\square$

We now prove Lemma 7.

Let $\beta = \exp(-(\theta - 4\gamma\Delta^2)^2\Gamma_0/2)$ be the probability bound from Lemma 9, where by our choice of $\theta = 2\Delta^2(1+\gamma)$ and with $\Gamma_0 = t^{1-\alpha}/2$, we have $\beta = \exp(-\Delta^4(1-\gamma)^2 t^{1-\alpha})$.

By Lemma 9, for a pair of users u and v with at least $t^{1-\alpha}/2$ items jointly explored, the probability that they are erroneously declared neighbors is upper-bounded by β.

Denote the set of users of type different from u by $\mathcal{B}$, and write

$$n_{\text{bad}} = \sum_{v \in \mathcal{B}} \mathbb{1}\{u \text{ and } v \text{ are declared to be neighbors}\},$$

whence $\mathbb{E}[n_{\text{bad}}] \leq n\beta$. Markov's inequality gives

$$\mathbb{P}(n_{\text{bad}} \geq n\sqrt{\beta}) \leq \frac{\mathbb{E}[n_{\text{bad}}]}{n\sqrt{\beta}} \leq \frac{n\beta}{n\sqrt{\beta}} = \sqrt{\beta}\,,$$

proving the lemma.

A.2 Proof of Lemma 2

We reproduce Lemma 2 below.

Lemma 2. *For user u at time t, if the good neighborhood event $\mathcal{E}_{good}(u,t)$ holds and $t \leq \mu m$, then*

$$\mathbb{P}(X_{ut} = 1) \geq 1 - 2m\exp\Big(-\frac{\Delta^2 t n^{1-\alpha}}{40km}\Big) - \frac{1}{t^\alpha} - \frac{1}{n^\alpha}\,.$$

We begin by checking that when the good neighborhood event $\mathcal{E}_{\text{good}}(u,t)$ holds for user u, the items have been rated enough times by the good neighbors.

Lemma 10. *For user u at time t, suppose that the good neighborhood event $\mathcal{E}_{good}(u,t)$ holds. Then for a given item i,*

$$\mathbb{P}\Big(\textit{item } i \textit{ has} \leq \frac{tn^{1-\alpha}}{10km} \textit{ ratings from good neighbors of } u\Big) \leq \exp\Big(-\frac{tn^{1-\alpha}}{40km}\Big).$$

Proof. The number of user u's good neighbors who have rated item i stochastically dominates a $\text{Bin}(\frac{n}{5k}, \frac{\varepsilon_R(n)t}{m})$ random variable, where $\frac{\varepsilon_R(n)t}{m} = \frac{t}{mn^\alpha}$ (here, we have critically used the lower bound on the number of good neighbors user u has when the good neighborhood event $\mathcal{E}_{\text{good}}(u,t)$ holds). By a Chernoff bound,

$$\mathbb{P}\Big(\text{Bin}\Big(\frac{n}{5k}, \frac{t}{mn^\alpha}\Big) \leq \frac{tn^{1-\alpha}}{10km}\Big) \leq \exp\Big(-\frac{1}{2}\frac{(\frac{tn^{1-\alpha}}{5km} - \frac{tn^{1-\alpha}}{10km})^2}{\frac{tn^{1-\alpha}}{5km}}\Big) \leq \exp\Big(-\frac{tn^{1-\alpha}}{40km}\Big). \quad \square$$

Next, we show a sufficient condition for which the algorithm correctly classifies every item as likable or unlikable for user u.

Lemma 11. *Suppose that there are no Δ-ambiguous items. For user u at time t, suppose that the good neighborhood event $\mathcal{E}_{good}(u,t)$ holds. Provided that every item $i \in [m]$ has more than $\frac{tn^{1-\alpha}}{10km}$ ratings from good neighbors of user u, then with probability at least $1 - m\exp(-\frac{\Delta^2 tn^{1-\alpha}}{20km})$, we have that for every item $i \in [m]$,*

$$\widetilde{p}_{ui} > \frac{1}{2} \quad \textit{if item } i \textit{ is likable by user } u,$$
$$\widetilde{p}_{ui} < \frac{1}{2} \quad \textit{if item } i \textit{ is unlikable by user } u.$$

Proof. Let A be the number of ratings that good neighbors of user u have provided. Suppose item i is likable by user u. Then when we condition on $A = a_0 \triangleq \lceil \frac{tn^{1-\alpha}}{10km} \rceil$, $\widetilde{p}_{ui}$ stochastically dominates

$$q_{ui} \triangleq \frac{\text{Bin}(a_0, p_{ui})}{a_0 + \Delta a_0} = \frac{\text{Bin}(a_0, p_{ui})}{(1+\Delta)a_0},$$

which is the worst-case variant of $\widetilde{p}_{ui}$ that insists that all Δa_0 bad neighbors provided rating "-1" for likable item i (here, we have critically used the upper bound on the number of bad neighbors user u has when the good neighborhood event $\mathcal{E}_{\text{good}}(u,t)$ holds). Then

$$\begin{aligned}
\mathbb{P}(q_{ui} \le \frac{1}{2} \mid A = a_0) &= \mathbb{P}\Big(\text{Bin}(a_0, p_{ui}) \le \frac{(1+\Delta)a_0}{2} \;\Big|\; A = a_0\Big) \\
&= \mathbb{P}\Big(a_0 p_{ui} - \text{Bin}(a_0, p_{ui}) \ge a_0\Big(p_{ui} - \frac{1}{2} - \frac{\Delta}{2}\Big) \;\Big|\; A = a_0\Big) \\
&\overset{(i)}{\le} \exp\Big(-2a_0\Big(p_{ui} - \frac{1}{2} - \frac{\Delta}{2}\Big)^2\Big) \\
&\overset{(ii)}{\le} \exp\Big(-\frac{1}{2}a_0\Delta^2\Big) \\
&\overset{(iii)}{\le} \exp\Big(-\frac{\Delta^2 tn^{1-\alpha}}{20km}\Big),
\end{aligned}$$

where step (i) is Hoeffding's inequality, step (ii) follows from item i being likable by user u (i.e., $p_{ui} \ge \frac{1}{2} + \Delta$), and step (iii) is by our choice of a_0. Conclude then that

$$\mathbb{P}(\widetilde{p}_{ui} \le \frac{1}{2} \mid A = a_0) \le \exp\Big(-\frac{\Delta^2 tn^{1-\alpha}}{20km}\Big).$$

Finally,

$$\begin{aligned}
\mathbb{P}\Big(\widetilde{p}_{ui} \le \frac{1}{2} \;\Big|\; A \ge \frac{tn^{1-\alpha}}{10km}\Big) &= \frac{\sum_{a=a_0}^{\infty} \mathbb{P}(A=a)\mathbb{P}(\widetilde{p}_{ui} \le \frac{1}{2} \mid A = a)}{\mathbb{P}(A \ge \frac{tn^{1-\alpha}}{10km})} \\
&\le \frac{\sum_{a=a_0}^{\infty} \mathbb{P}(A=a)\exp(-\frac{\Delta^2 tn^{1-\alpha}}{20km})}{\mathbb{P}(A \ge \frac{tn^{1-\alpha}}{10km})} \\
&= \exp\Big(-\frac{\Delta^2 tn^{1-\alpha}}{20km}\Big).
\end{aligned}$$

A similar argument holds for when item i is unlikable. Union-bounding over all m items yields the claim. □

We now prove Lemma 2. First off, provided that $t \le \mu m$, we know that there must still exist an item likable by user u that user u has yet to consume. For user u at time t, supposing that event $\mathcal{E}_{\text{good}}(u,t)$ holds, then every item has been rated more than $\frac{tn^{1-\alpha}}{10km}$ times by the good neighbors of user u with probability at least $1 - m\exp(-\frac{tn^{1-\alpha}}{40km})$. This follows from union-bounding over the m items with Lemma 10. Applying Lemma 11, and noting that we only exploit with probability $1 - \varepsilon_J(t) - \varepsilon_R(n) = 1 - 1/t^\alpha - 1/n^\alpha$, we finish the proof:

$$\begin{aligned}
\mathbb{P}(X_{ut} = 1) &\ge 1 - m\exp\Big(-\frac{tn^{1-\alpha}}{40km}\Big) - m\exp\Big(-\frac{\Delta^2 tn^{1-\alpha}}{20km}\Big) - \frac{1}{t^\alpha} - \frac{1}{n^\alpha} \\
&\ge 1 - 2m\exp\Big(-\frac{\Delta^2 tn^{1-\alpha}}{40km}\Big) - \frac{1}{t^\alpha} - \frac{1}{n^\alpha}.
\end{aligned}$$

A.3 Experimental Results

We demonstrate our algorithm COLLABORATIVE-GREEDY on two datasets, showing that they have comparable performance and that they both outperform two existing recommendation algorithms Popularity Amongst Friends (PAF) [4] and Deshpande and Montanari's method (DM) [12]. At each time step, PAF finds nearest neighbors ("friends") for every user and recommends to a user the "most popular" item, i.e., the one with the most number of $+1$ ratings, among the user's friends. DM doesn't do any collaboration beyond a preprocessing step that computes item feature vectors via matrix completion. Then during online recommendation, DM learns user feature vectors over time with the help of item feature vectors and recommends an item to each user based on whether it aligns well with the user's feature vector.

We simulate an online recommendation system based on movie ratings from the Movielens10m and Netflix datasets, each of which provides a sparsely filled user-by-movie rating matrix with ratings out of 5 stars. Unfortunately, existing collaborative filtering datasets such as the two we consider don't offer the interactivity of a real online recommendation system, nor do they allow us to reveal the rating for an item that a user didn't actually rate. For simulating an online system, the former issue can be dealt with by simply revealing entries in the user-by-item rating matrix over time. We address the latter issue by only considering a dense "top users vs. top items" subset of each dataset. In particular, we consider only the "top" users who have rated the most number of items, and the "top" items that have received the most number of ratings. While this dense part of the dataset is unrepresentative of the rest of the dataset, it does allow us to use actual ratings provided by users without synthesizing any ratings.

An initial question to ask is whether the dense movie ratings matrices we consider could be reasonably explained by our latent source model. We automatically learn the structure of these matrices using the method by Grosse et al. [13] and find Bayesian clustered tensor factorization (BCTF) to accurately model the data. This finding isn't surprising as BCTF has previously been used to model movie ratings data [21]. BCTF effectively clusters both users and movies so that we get structure such as that shown in Figure 1(a) for the Movielens10m "top users vs. top items" matrix. Our latent source model could reasonably model movie ratings data as it only assumes clustering of users.

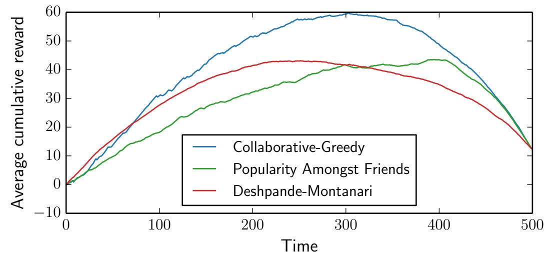
Following the experimental setup of [4], we quantize a rating of 4 stars or more to be $+1$ (likeable), and a rating of 3 stars or less to be -1 (unlikeable). While we look at a dense subset of each dataset, there are still missing entries. If a user u hasn't rated item j in the dataset, then we set the corresponding true rating to 0, meaning that in our simulation, upon recommending item j to user u, we receive 0 reward, but we still mark that user u has consumed item j; thus, item j can no longer be recommended to user u. For both Movielens10m and Netflix datasets, we consider the top $n = 200$ users and the top $m = 500$ movies. For Movielens10m, the resulting user-by-rating matrix has 80.7% nonzero entries. For Netflix, the resulting matrix has 86.0% nonzero entries. For an algorithm that recommends item π_{ut} to user u at time t, we measure the algorithm's average cumulative reward up to time T as

$$\frac{1}{n}\sum_{t=1}^{T}\sum_{u=1}^{n} Y_{u\pi_{ut}}^{(T)},$$

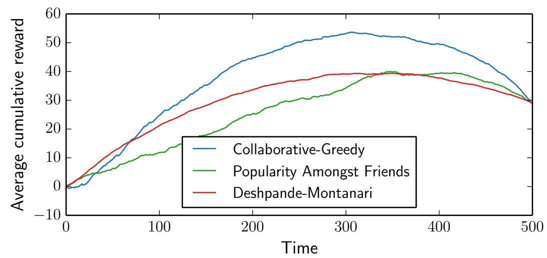
where we average over users.

For all methods, we recommend items until we reach time $T = 500$, i.e., we make movie recommendations until each user has seen all $m = 500$ movies. We disallow the matrix completion step for DM to see the users that we actually test on, but we allow it to see the the same items as what is in the simulated online recommendation system in order to compute these items' feature vectors (using the rest of the users in the dataset). Furthermore, when a rating is revealed, we provide DM both the thresholded rating and the non-thresholded rating, the latter of which DM uses to estimate user feature vectors over time.

Parameters θ and α for and COLLABORATIVE-GREEDY are chosen using training data: We sweep over the two parameters on training data consisting of 200 users that are the "next top" 200 users, i.e., ranked 201 to 400 in number movie ratings they provided. For simplicity, we discretize our search space to $\theta \in \{0.0, 0.1, \ldots, 1.0\}$ and $\alpha \in \{0.1, 0.2, 0.3, 0.4, 0.5\}$. We choose the parameter setting achieving the highest area under the cumulative reward curve. For both Movielens10m and Netflix datasets, this corresponded to setting $\theta = 0.0$ and $\alpha = 0.5$ for COLLABORATIVE-GREEDY.



(a)



(b)

Figure 2: Average cumulative rewards over time: (a) Movielens10m, (b) Netflix.

In contrast, the parameters for PAF and DM are chosen to be the best parameters for the test data among a wide range of parameters. The results are shown in Figure 2. We find that our algorithm COLLABORATIVE-GREEDY outperforms PAF and DM. We remark that the curves are roughly concave, which is expected since once we've finished recommending likeable items (roughly around time step 300), we end up recommending mostly unlikeable items until we've exhausted all the items.