Supplementary Material for "The Sample Complexity of Online One-Class Collaborative Filtering"

6. Proof of Theorem 1

Theorem 1 follows immediately from the following result.

Theorem 2 *Suppose that there are at least $\frac{N}{2K}$ users of the same type, for all user types, and assume that at least a fraction ν of all items is likable to a given user, for all users. Moreover, suppose that for some $\gamma \in [0,1)$, all users satisfy condition (1). Pick $\delta > 0$ and suppose that the number of nearest neighbors k, the batch size Q, and the parameter η, are chosen such that $k \leq \frac{9N}{40K}$, $\eta \leq \nu/2$,*

$$\frac{k}{Q} \geq \frac{64\log(8M/\delta)}{p_{\mathrm{f}}\Delta^2}, \tag{7}$$

and

$$Q \geq \frac{10}{\nu}\log(4/\delta). \tag{8}$$

Then the reward accumulated by the User-CF algorithm up to time $T \in [T_{start}, \frac{4}{5}\nu M p_{\mathrm{f}}]$ with

$$T_{start} = \frac{\left(512\max\left(\log\left(\frac{4NQ}{k\Delta}\right), \log\left(\frac{88}{\delta}\right)\right)\right)^{\frac{1}{1-\alpha}}}{(3p_{\mathrm{f}}^2(1-\gamma)^2\nu)^{\frac{1}{1-\alpha}}\left(1-\max\left(\frac{1}{T},\frac{2}{\eta Q}\right)\right)}$$

satisfies

$$\frac{\mathbb{E}\left[\mathrm{reward}(T)\right]}{NT} \geq \left(1 - \frac{T_{start}}{T} - 2^{\alpha}\frac{(T-T_{start})^{1-\alpha}}{T(1-\alpha)} - \max\left(\frac{1}{T},\frac{2}{\eta Q}\right)\right)(1-\delta). \tag{9}$$

Theorem 1 follows by choosing the parameter of the User-CF algorithm as follows:

$$\eta = \frac{\nu}{2}, \quad k = \frac{9}{40}\frac{N}{K}, \quad \text{and} \quad Q = k\frac{p_{\mathrm{f}}\Delta^2}{64\log(8M/\delta)}.$$

To see this, note that by definition, the conditions on k and η and condition (7) on Q are satisfied. By (4), condition (8) holds and $\frac{2}{\eta Q} = \frac{K}{N}\frac{c'\log(M/\delta)}{p_{\mathrm{f}}\Delta^2}$. Moreover, $\max\left(\log\left(\frac{4NQ}{k\Delta}\right), \log\left(\frac{88}{\delta}\right)\right) \leq \tilde{c}\log(N/\delta)$.

6.1. Proof of Theorem 2

Theorem 2 is proven by showing that at time $t \geq T_{\text{start}}$ the following holds for all users u:

i) the neighborhood of u is sufficiently well explored by similarity exploration steps so that most of the nearest neighbors of u are *good*, i.e., are of the same user type as u (similarly, neighbors are called *bad* if they are of a different user type than u),

ii) for $t \geq T_{\text{start}}$, the estimates $\hat{p}_{ui}$, for all $i \in \mathcal{Q}_q, q = 0,\ldots,\frac{t}{\eta Q}-1$ correctly predict whether i is likable by u or not, and

iii) there exist items in the sets $\mathcal{Q}_q, q = 0,\ldots,\frac{t}{\eta Q}-1$ that are likable by u and that have not been rated by u at previous times steps.

Conditions i, ii, and iii guarantee that an exploitation step recommends a likable item.

Formally, we start by defining the following events:

$$\mathcal{G}_\beta(t) = \{\text{At time } t\text{, no more than } \beta k \text{ of the } k\text{-nearest neighbors of } u \text{ are bad}\}, \tag{10}$$

$$\mathcal{L}(t) = \{\text{at time } t\text{, there exists an item } i \in \mathcal{Q}_q,$$
$$q = 0, \ldots, t/(\eta Q) - 1 \text{ that is likable by } u\}, \tag{11}$$

and

$$\mathcal{E}(t) = \bigcup_{q=0,\ldots,\frac{t}{\eta Q}-1} \mathcal{E}_q(t), \tag{12}$$

with

$$\mathcal{E}_q(t) = \{\text{Conditioned on } \mathcal{G}_{\frac{\Delta}{4Q}}(t)\text{, for all } i \in \mathcal{Q}_q,$$
$$\hat{p}_{ui} > p_{\mathrm{f}}/2, \text{ if } p_{ui} > 1/2 + \Delta, \text{ and}$$
$$\hat{p}_{ui} < p_{\mathrm{f}}/2, \text{ if } p_{ui} < 1/2 - \Delta\}. \tag{13}$$

For convenience, we omit in the notion of $\mathcal{L}(t)$, $\mathcal{G}_\beta(t)$, $\mathcal{E}(t)$, and $\mathcal{E}_q(t)$ the dependence on u. The significance of those definitions is that if $\mathcal{L}(t)$, $\mathcal{G}_{\frac{\Delta}{4Q}}(t)$, and $\mathcal{E}(t)$ hold simultaneously, then the recommendation made to user u by an exploitation step at time t is likable. We can therefore lower-bound the reward $\mathbb{E}\left[\mathrm{reward}(T)\right]$ as follows:

$$\frac{\mathbb{E}\left[\mathrm{reward}(T)\right]}{NT} = \frac{1}{NT}\sum_{t=0}^{T-1}\sum_{u=0}^{N-1}\mathbb{P}\left[X_{ui(u,t)} = 1\right]$$
$$\geq \frac{1}{NT}\sum_{u=0}^{N-1}\sum_{t=0, t\notin\{\eta Q q:\, q=0,1,\ldots\}}^{T-1}\mathbb{P}\left[\text{exploitation at } t\right]\mathbb{P}\left[X_{ui(u,t)} = 1 | \text{exploitation at } t\right] \tag{14}$$
$$\geq \frac{1}{N}\sum_{u=0}^{N-1}\left(\frac{1}{T}\sum_{t=0}^{T-1}(1-(2/t)^\alpha)\mathbb{P}\left[X_{ui(u,t)} = 1 | \text{exploitation at } t\right] - \max\left(\frac{1}{T}, \frac{2}{\eta Q}\right)\right) \tag{15}$$
$$\geq \frac{1}{N}\sum_{u=0}^{N-1}\left(\frac{1}{T}\sum_{t=T_{\text{start}}}^{T-1}(1-\delta)(1-(2/t)^\alpha) - \max\left(\frac{1}{T}, \frac{2}{\eta Q}\right)\right) \tag{16}$$
$$\geq (1-\delta)\left(1 - \frac{T_{\text{start}}}{T} - 2^\alpha\frac{(T-T_{\text{start}})^{1-\alpha}}{T(1-\alpha)} - \max\left(\frac{1}{T}, \frac{2}{\eta Q}\right)\right). \tag{17}$$

Here, (14) follows from

$$\mathbb{P}\left[X_{ui(u,t)} = 1 | \text{preference exploration at } t\right] \geq 0 \quad \text{and} \quad \mathbb{P}\left[X_{ui(u,t)} = 1 | \text{similarity exploration at } t\right] \geq 0.$$

For (15) we used, for $t \neq \eta Q q$,

$$\mathbb{P}\left[\text{exploration at } t\right] = 1 - (t - \lfloor t/(\eta Q)\rfloor)^{-\alpha} \geq 1 - (t(1-1/(\eta Q)))^{-\alpha} \geq 1 - (2/t)^\alpha$$

which follows from $\eta Q \geq 2$. Moreover we used for (15) that the fraction of preference exploration steps up to time T is at most $\max(\frac{1}{T}, \frac{2}{\eta Q})$. To see that, note that at $T \in \{\eta Q q, \ldots, \eta Q(q+1)\}$ we have performed $q+1$ preference exploration steps. It follows that, for $q \geq 1$, the fraction of preference exploration steps performed up to T is given by $\frac{q+1}{q\eta Q} \leq \frac{2}{\eta Q}$. Thus, for any $T \geq 1$, the fraction of preference exploration steps is $\leq \max(\frac{1}{T}, \frac{2}{\eta Q})$. Equality (16) follows from

$$\mathbb{P}\left[X_{ui(u,t)} = 1 | \text{exploitation at } t\right] \geq \mathbb{P}\left[\mathcal{E}(t) \cap \mathcal{G}_{\frac{\Delta}{4Q}}(t) \cap \mathcal{L}(t)\right]$$
$$\geq 1 - \delta. \tag{18}$$

Here, inequality (18) holds for $t \geq T_{\text{start}}$ and is established below. Finally, inequality (17) follows from

$$\sum_{t=T_{\text{start}}}^{T-1} t^{-\alpha} \leq \int_{T_{\text{start}}-1}^{T-1} t^{-\alpha} = \frac{1}{1-\alpha}t^{1-\alpha}|_{t=T_{\text{start}}-1}^{T-1}$$
$$= \frac{(T-1)^{1-\alpha} - (T_{\text{start}}-1)^{1-\alpha}}{1-\alpha} \leq \frac{(T-T_{\text{start}})^{1-\alpha}}{1-\alpha}.$$

It remains to establish (18). To this end, define for notational convenience

$$A := \frac{256 \max\left(\log\left(\frac{4NQ}{k\Delta}\right), \log\left(\frac{88}{\delta}\right)\right)}{3p_{\mathrm{f}}^2(1-\gamma)^2\nu},$$

and let T_s be the number of similarity exploration steps executed up to time T. Inequality (18) follows by noting that, for all $t \geq T_{\text{start}}$, by the union bound,

$$\begin{aligned}
\mathbb{P}\left[(\mathcal{E}(t) \cap \mathcal{G}_{\frac{\Delta}{4Q}}(t) \cap \mathcal{L}(t))^c\right] &\leq \mathbb{P}\left[\mathcal{E}^c(t)\right] + \mathbb{P}\left[\mathcal{G}^c_{\frac{\Delta}{4Q}}(t)\right] + \mathbb{P}\left[\mathcal{L}^c(t)\right] \\
&\leq \mathbb{P}\left[\mathcal{E}^c(t)\right] + \mathbb{P}\left[\mathcal{G}^c_{\frac{\Delta}{4Q}}(t) | T_s \geq A\right] + \mathbb{P}\left[T_s \leq A\right] + \mathbb{P}\left[\mathcal{L}^c(t)\right] \quad (19) \\
&\leq \frac{\delta}{4} + \frac{\delta}{4} + \frac{\delta}{4} + \frac{\delta}{4} = \delta. \quad (20)
\end{aligned}$$

Here, inequality (19) follows since for two events C, B we have that

$$\mathbb{P}\left[C\right] = \mathbb{P}\left[C \cap B\right] + \mathbb{P}\left[C \cap B^c\right] = \mathbb{P}\left[C|B\right]\mathbb{P}\left[B\right] + \mathbb{P}\left[C|B^c\right]\mathbb{P}\left[B^c\right] \leq \mathbb{P}\left[C|B\right] + \mathbb{P}\left[B^c\right]. \quad (21)$$

Inequality (20) follows from

$$\mathbb{P}\left[\mathcal{E}^c(t)\right] \leq \delta/4 \quad (22)$$

$$\mathbb{P}\left[\mathcal{G}^c_{\frac{\Delta}{4Q}}(t) | T_s \geq A\right] \leq \delta/4 \quad (23)$$

$$\mathbb{P}\left[T_s \leq A\right] \leq \delta/4 \quad (24)$$

$$\mathbb{P}\left[\mathcal{L}^c(t)\right] \leq \delta/4. \quad (25)$$

In the remainder of this proof, we establish the inequalities (22)-(25). The key ingredient for these bounds are concentration inequalities, in particular a version of Bernstein's inequality (Bardenet and Maillard, 2015).

Proof of (22): By the union bound, we have, for all $t = 0, \ldots, M-1$, that

$$\mathbb{P}\left[\mathcal{E}^c(t)\right] \leq \sum_{q=0}^{M/Q-1} \mathbb{P}\left[\mathcal{E}_q^c(t)\right] \leq \frac{\delta}{4}$$

as desired. Here, we used $\mathbb{P}\left[\mathcal{E}_q^c(t)\right] \leq \frac{\delta Q}{4M}$, which follows from Lemma 1 stated below with $\delta' = \frac{\delta Q}{4M}$ and $T_r = 1$ (note that the assumption (26) of Lemma 1 is implied by the assumption (7) of Theorem 2).

Lemma 1 (Preference exploration) *Suppose we recommend T_r random items to each user, chosen uniformly at random from a set $\mathcal{Q} \subseteq [M]$ of Q items. Suppose that p_{vi} is Δ-bounded away from $1/2$, for all $i \in \mathcal{Q}$ and for all $v \in \mathcal{N}_u$, where $\mathcal{N}_u$ is a set of k users, of which no more than βk, with $\beta \leq \frac{\Delta T_r}{4Q}$, of the users are of a different type than u. Fix $\delta' > 0$. If*

$$T_r \frac{k}{Q} \frac{p_{\mathrm{f}}\Delta^2}{64\log(2Q/\delta')} \geq 1 \quad (26)$$

then, with probability at least $1-\delta'$, for all $i \in \mathcal{Q}$, $\hat{p}_{ui} > \frac{p_{\mathrm{f}}}{2}$ if $p_{ui} \geq 1/2 + \Delta$ and $\hat{p}_{ui} < \frac{p_{\mathrm{f}}}{2}$ if $p_{ui} \leq 1/2 - \Delta$.

Proof of (23): Inequality (23) follows from Lemma 2 below, which ensures that a user has many good and only few bad neighbors.

Lemma 2 (Many good and few bad neighbors) *Let $\mathcal{T}_u$ be the subsets of all users $[N]$ that are of the same type of u and suppose its cardinality satisfies $\geq \frac{N}{2K}$. Suppose that, for some constant $\gamma \in [0,1)$, condition (1) holds, and that the number of nearest neighbors k satisfies $k \leq \frac{9N}{40K}$. Choose $\beta \in (0,1)$, and suppose*

$$T_s \geq \frac{64\log(N/(\beta k))}{3p_{\mathrm{f}}^2(1-\gamma)^2 \frac{1}{M}\min_{v\in\mathcal{T}_u}\langle \mathbf{p}_u, \mathbf{p}_v\rangle} \quad (27)$$

similarity exploration steps have been performed. Then, with probability at least $1 - 11e^{-\frac{3}{64}T_s p_{\mathrm{f}}^2(1-\gamma)^2\frac{1}{M}\min_{v\in\mathcal{T}_u}\langle \mathbf{p}_u, \mathbf{p}_v\rangle}$, the set of nearest neighbors $\mathcal{N}_u$ of user u (defined in Section 3), contains no more than βk bad neighbors.

To see that inequality (23) follows from Lemma 2, we first note that $T_s \geq A$ guarantees that condition (27) of Lemma 2 is satisfied (with $\beta = \frac{\Delta}{4Q}$). To see this, note that since each user likes at least a fraction ν of the items, we have

$$\frac{1}{M} \min_{v \in \mathcal{T}_u} \langle \mathbf{p}_u, \mathbf{p}_v \rangle \geq \nu \left(\frac{1}{2} + \Delta \right)^2 \geq \frac{\nu}{4}. \tag{28}$$

Lemma 2 therefore implies

$$\mathbb{P}\left[\mathcal{G}^c_{\frac{\Delta}{4Q}}(t) | T_s \geq A \right] \leq 11 e^{-\frac{3}{64} T_s p_f^2 (1-\gamma)^2 \frac{1}{M} \min_{v \in \mathcal{T}_u} \langle \mathbf{p}_u, \mathbf{p}_v \rangle} \leq 11 e^{-\log(88/\delta)} = \frac{\delta}{8},$$

as desired. For the second inequality above we used (28) and $T_s \geq A$.

Proof of (24): We next establish the inequality $\mathbb{P}[T_s \leq A] \leq \delta/4$. To this end, recall that a similarity exploration step is carried out at $t = 0, \ldots, T-1, t \neq \eta Q q, q = 0, 1, \ldots$ with probability $1/(t - \lfloor t/(\eta Q) \rfloor)$. Recall from the discussion below inequality (17), that the fraction of time steps up to time T for which $t = \eta Q q$, for some q, is at most $\max(\frac{1}{T}, \frac{2}{\eta Q})$. It follows that the number of similarity exploration steps, T_s, carried out after $t \geq T_{\text{start}}$ steps of the User-CF algorithm, stochastically dominates the random variable $S = \sum_{t=1}^{\tilde{T}} Z_t$, $\tilde{T} = T_{\text{start}}(1 - \max(\frac{1}{T}, \frac{2}{\eta Q}))$, where Z_t is a binary random variable with $\mathbb{P}[Z_t = 1] = 1/t^\alpha$. It follows that

$$\mathbb{P}[T_s \leq A] = \mathbb{P}\left[T_s \leq \tilde{T}^{1-\alpha}/2 \right] \leq e^{-\frac{\tilde{T}^{1-\alpha}}{20}} \leq \delta/4, \tag{29}$$

where the first inequality holds by definition of T_{start}, i.e.,

$$T_{\text{start}} = (2A)^{\frac{1}{1-\alpha}} / \left(1 - \max\left(\frac{1}{T}, \frac{2}{\eta Q} \right) \right),$$

and the second inequality holds by Lemma 3 stated below. Finally, the last inequality in (29) follows from

$$\tilde{T} = (2A)^{\frac{1}{1-\alpha}} \geq \frac{128}{3} \log(44/\delta).$$

The following lemma appears in (Bresler et al., 2014).

Lemma 3 *Let $S = \sum_{t=1}^{\tilde{T}} Z_t$ where Z_t is a binary random variable with $\mathbb{P}[Z_t = 1] = 1/t^\alpha$, $\alpha \in (0, 4/7)$. We have that*

$$\mathbb{P}\left[S_T \leq \tilde{T}^{1-\alpha}/2 \right] \leq e^{-\frac{\tilde{T}^{1-\alpha}}{20}}.$$

Proof of (25): Suppose $t < \eta Q$, consider user u, and let N_0 be the total number of items likable by u in the set $\mathcal{Q}_0$ (recall that $\mathcal{Q}_0$ is choosen uniformly at random from the subset of items $[M]$ of cardinality Q). Note that $N_0 > \eta Q$ implies that at $t < \eta Q$, there exist items that are likable by u in $\mathcal{Q}_0$ that have not been recommended to u yet. Therefore, we can upper bound the probability that no likable items are left to recommend, for $t < \eta Q$, by

$$\mathbb{P}[\mathcal{L}^c(t)] \leq \mathbb{P}[N_0 \leq \eta Q] \leq \mathbb{P}[N_0 \leq Q\nu/2] \leq \mathbb{P}[N_0 \leq \mathbb{E}[N_0] - Q\nu/2] \tag{30}$$

$$\leq e^{-Q \frac{(\nu/2)^2}{2\nu(1-\nu) + \frac{2}{3}\frac{\nu}{2}}} = e^{-Q \frac{\nu/4}{2(1-\nu)+\frac{1}{3}}} \leq e^{-Q \frac{\nu}{10}} \leq \frac{\delta}{4}. \tag{31}$$

Here, the first inequality in (30) follows from $\eta \leq \nu/2$, by assumption; the second inequality in (30) follows from $\mathbb{E}[N_0] \geq \nu Q$ (since at least a fraction of ν of the items is likable by u), the first inequality in (31) follows from Bernstein's inequality (Bardenet and Maillard, 2015), and finally the last inequality in (30) holds by assumption (8). We have established that $\mathbb{P}[\mathcal{L}^c(t)] \leq \delta/4$, for $t < \eta Q$. Using the exact same line of arguments yields the same bound for $t \in [\eta Q, \eta M]$.

It remains to upper bound $\mathbb{P}[\mathcal{L}^c(t)]$ for $t \in [\eta M, \frac{4}{5}\nu M p_f]$. To this end, let $N_u^c(T)$ be the number of (likable) items that have been rated by user u after T time steps, and note that if $N_u^c(T)$ is strictly smaller than the (minimum) number of likable items, then there are likable items left to recommend. Formally,

$$\mathbb{P}[\mathcal{L}^c(t)] \leq \mathbb{P}[N_u^c(T) \geq \nu M] \tag{32}$$

where we used that for each user u, at least νM items are likable. Recall that with probability $p_{ui}p_{\mathrm{f}} \leq p_{\mathrm{f}}$ a likable item i is rated if it is recommended to u. Once rated, an item is not recommended again.

Note that $N_u^c(T)$ is statistically dominated by a sum of independent binary random variables Z_t with $\mathbb{P}\left[Z_t = 1\right] = p_{\mathrm{f}}$. We therefore have that

$$\mathbb{P}\left[N_u^c(T) \geq \nu M\right] \leq \mathbb{P}\left[N_u^c(T) \geq T(p_{\mathrm{f}} + \frac{p_{\mathrm{f}}}{4})\right] \leq e^{-\frac{T p_{\mathrm{f}}^2}{2}} \leq e^{-\frac{T_{\text{start}} p_{\mathrm{f}}^2}{2}} \leq \frac{\delta}{4}. \tag{33}$$

Here, the first inequality holds by the assumption $T \leq \frac{4}{5}\nu M p_{\mathrm{f}}$, the second inequality follows by Hoeffding's inequality, the third inequality follows by $T \geq T_{\text{start}}$, and the last inequality follows from $T_{\text{start}} \geq \frac{2}{p_{\mathrm{f}}^2}\log(4/\delta)$, which holds by definition of T_{start}. Application of (33) on (32) concludes the proof of $\mathbb{P}\left[\mathcal{L}^c(t)\right] \leq \delta/4$.

6.2. Proof of Lemma 2

Recall that $\mathbf{r}_u^{\text{sim}} \in \{0,1\}^M$ is the vector containing the responses R_{ui} of user u to previous *similarity* exploitation steps up to time t, and that we assume in Lemma 2, that T_s similarity exploration steps have been performed up to time t. To establish Lemma 2, we show that there are more than k users v that are of the same user type as u and satisfy $\frac{1}{T_s}\left\langle \mathbf{r}_u^{\text{sim}}, \mathbf{r}_v^{\text{sim}} \right\rangle \geq \theta$, and at the same time, there are fewer than $k\beta$ users of a different user type as u that satisfy $\frac{1}{T_s}\left\langle \mathbf{r}_u^{\text{sim}}, \mathbf{r}_v^{\text{sim}} \right\rangle \geq \theta$ for a certain threshold θ chosen below. This is accomplished by the following two lemmas.

Lemma 4 (Many good neighbors) *Suppose there are at least $\frac{N}{2K}$ users of the type as user u (including u), and suppose that T_s similarity exploration steps have been performed. Then, with probability at least $1 - 10p_{\text{good}}$,*

$$p_{\text{good}} \coloneqq e^{-\frac{3}{16}T_s p_g (1-\theta/p_g)^2}, \quad p_g \coloneqq p_{\mathrm{f}}^2 \frac{1}{M} \min_{v \in \mathcal{T}_u} \left\langle \mathbf{p}_u, \mathbf{p}_v \right\rangle,$$

at least $\frac{9N}{40K}$ users v of the same user type as u obey $\frac{1}{T_s}\left\langle \mathbf{r}_u^{sim}, \mathbf{r}_v^{sim} \right\rangle \geq \theta$.

Lemma 5 (Few bad neighbors) *Suppose that T_s similarity exploration steps have been performed. Then, with probability at least $1 - p_{\text{bad}}$, where*

$$p_{\text{bad}} = e^{-\frac{T_s p_b(\theta/p_b - 1)^2/4}{1 + (\theta/p_b - 1)/3}}, \quad p_b \coloneqq p_{\mathrm{f}}^2 \max_{v \notin \mathcal{T}_u} \frac{1}{M}\left\langle \mathbf{p}_v, \mathbf{p}_u \right\rangle,$$

at most Np_{bad} users v of a different user type than u obey $\frac{1}{T_s}\left\langle \mathbf{r}_u^{sim}, \mathbf{r}_v^{sim} \right\rangle \geq \theta$.

We set

$$\theta = \frac{p_g + p_b}{2}.$$

With this choice, by Lemma 4, there are more than $\frac{9N}{40K} \geq k$ (the inequality holds by assumption) users v of the same type as u that satisfy $\frac{1}{T_s}\left\langle \mathbf{r}_u^{\text{sim}}, \mathbf{r}_v^{\text{sim}} \right\rangle \geq \theta$, with probability at least $1 - 10p_{\text{good}}$. By Lemma 5, there are no more than Np_{bad} users v of a different type as u with $\frac{1}{T_s}\left\langle \mathbf{r}_u^{\text{sim}}, \mathbf{r}_v^{\text{sim}} \right\rangle \geq \theta$. Thus, by the union bound, $\mathcal{N}_u$ contains less than $p_{\text{bad}}N$ bad neighbors with probability at least

$$1 - 10p_{\text{good}} - p_{\text{bad}} \geq 1 - 11e^{-\frac{3}{64}T_s p_g (1-\gamma)^2}.$$

Here, we used

$$p_{\text{good}} = e^{-\frac{3}{64}T_s p_g (1 - p_b/p_g)^2} \leq e^{-\frac{3}{64}T_s p_g (1-\gamma)^2}$$

where the inequality follows by $p_b/p_g \leq \gamma$, by (1). Moreover, we used

$$\begin{aligned} p_{\text{bad}} &= e^{-\frac{T_s p_b(\theta/p_b - 1)^2/4}{1+(\theta/p_b - 1)/3}} = e^{-\frac{T_s p_b (p_g/p_b - 1)^2/16}{1 + (p_g/p_b - 1)/6}} = e^{-\frac{T_s p_g (\sqrt{p_g/p_b} - \sqrt{p_b/p_g})^2/16}{1 + (p_g/p_b - 1)/6}} \leq e^{-\frac{T_s p_g (\sqrt{1/\gamma} - \sqrt{\gamma})^2/16}{1 + (1/\gamma - 1)/6}} \\ &\leq e^{-\frac{T_s p_g (\sqrt{1/\gamma} - \sqrt{\gamma})^2/16}{1 + (1/\gamma - 1)}} = e^{-T_s p_g (1-\gamma)^2/16}. \end{aligned} \tag{34}$$

Here, the first inequality follows from the absolute value of the exponent being decreasing in p_b/p_g, and from the assumption $p_b/p_g \leq \gamma$, by (1).

To conclude the proof, we needed to establish that the maximum number of bad neighbors Np_{bad} satisfies $Np_{\text{bad}} \leq \beta k$. This follows directly by noting that, by assumption (27), the RHS of (34) is upper-bounded by $\frac{\beta k}{N}$.

6.2.1. Proof of Lemma 4

Consider u and assume there are exactly $\frac{N}{2K}$ users from the same user type. There could be more, but it is sufficient to consider $\frac{N}{2K}$. Let v be of the same user type. We start by showing that $\frac{1}{T_s}\left\langle \mathbf{r}_u^{\text{sim}}, \mathbf{r}_v^{\text{sim}} \right\rangle \geq \theta$ with high probability. To this end, note that $\left\langle \mathbf{r}_u^{\text{sim}}, \mathbf{r}_v^{\text{sim}} \right\rangle = \sum_{t=0}^{T_s-1} R_{u\pi(t)} R_{v\pi(t)}$ where π is the random permutation of the item space drawn by the User-CF algorithm at initialization, and $R_{u\pi(t)} R_{v\pi(t)}$ is a binary random variable, independent across t, with success probability $p_{\text{f}}^2 p_{u\pi(t)} p_{v\pi(t)}$. Setting $a \coloneqq p_{\text{f}}^2 \frac{1}{M} \langle \mathbf{p}_u, \mathbf{p}_v \rangle$, for notational convenience, it follows that

$$\mathbb{P}\left[\frac{1}{T_s}\left\langle \mathbf{r}_u^{\text{sim}}, \mathbf{r}_v^{\text{sim}} \right\rangle \leq \theta\right] = \mathbb{P}\left[\frac{1}{T_s}\left\langle \mathbf{r}_u^{\text{sim}}, \mathbf{r}_v^{\text{sim}} \right\rangle \leq a - (a-\theta)\right] \tag{35}$$

$$\leq e^{-\frac{T_s(a-\theta)^2/2}{a+(a-\theta)/3}} \tag{36}$$

$$= e^{-\frac{T_s a(1-\theta/a)^2/2}{1+(1-\theta/a)/3}} \leq e^{-\frac{3}{8}T_s a(1-\theta/a)^2} \tag{37}$$

$$\leq e^{-\frac{3}{8}T_s p_g(1-\theta/p_g)^2} \leq p_{\text{good}}. \tag{38}$$

Here, (36) follows from Bernstein's inequality (Bardenet and Maillard, 2015), and for (38) we used that the RHS of (37) is decreasing in a.

Next, consider the random variable

$$W = \sum_{v \in \mathcal{T}_u} G_v, \quad G_v = \mathbb{1}\left\{\frac{1}{T_s}\left\langle \mathbf{r}_u^{\text{sim}}, \mathbf{r}_v^{\text{sim}} \right\rangle \geq \theta\right\},$$

where $\mathcal{T}_u$ is the subset of all users $[N]$ that are of the same time as user u, as before. By Chebyshev's inequality,

$$\mathbb{P}\left[W - \mathbb{E}\left[W\right] \leq -\frac{\mathbb{E}\left[W\right]}{2}\right] \leq \frac{\text{Var}(W)}{(\mathbb{E}\left[W\right]/2)^2}. \tag{39}$$

Since there are at least $\frac{N}{2K}$ users of the same type, the carnality of $\mathcal{T}_u$ is lower bounded by $\frac{N}{2K} - 1$. It follows with (38) that

$$\mathbb{E}\left[W\right] \geq (1 - p_{\text{good}})\left(\frac{N}{2K} - 1\right).$$

Next, we upper bound the variance of W. We have

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

With $G_v = G_v^2$,

$$\text{Var}(G_v) = \mathbb{E}\left[G_v^2\right] - \mathbb{E}\left[G_v\right]^2 = \mathbb{E}\left[G_v\right](1 - \mathbb{E}\left[G_v\right]) \leq 1 - \mathbb{E}\left[G_v\right] \leq p_{\text{good}}.$$

Similarly,

$$\text{Cov}(G_v, G_w) = \mathbb{E}\left[G_v G_w\right] - \mathbb{E}\left[G_v\right]\mathbb{E}\left[G_w\right] \leq 1 - (1-q)^2 \leq 2p_{\text{good}}.$$

Thus, we obtain

$$\text{Var}(W) \leq \left(\frac{N}{2K} - 1\right) p_{\text{good}} + \left(\frac{N}{2K} - 1\right)\left(\frac{N}{2K} - 2\right) 2p_{\text{good}} \leq \left(\frac{N}{2K} - 1\right)^2 2p_{\text{good}}.$$

Plugging this into (39) yields

$$\mathbb{P}\left[W - \mathbb{E}\left[W\right] \leq -\frac{\mathbb{E}\left[W\right]}{2}\right] \leq \frac{8p_{\text{good}}}{(1 - p_{\text{good}})^2} \leq 10p_{\text{good}},$$

for $p_{\text{good}} \leq 1/10$. It follows that the number of good neighbors is larger than

$$W \geq \mathbb{E}\left[W\right]/2 \geq (1 - p_{\text{good}})\frac{N}{4K} \geq \frac{9N}{40K}$$

with probability at least $1 - 10p_{\text{good}}$.

6.2.2. Proof of Lemma 5

Let u and v be two fixed users of different user types. Similarly as in the proof of Lemma 4, we start by showing that $\frac{1}{T_s}\left\langle \mathbf{r}_u^{\text{sim}}, \mathbf{r}_v^{\text{sim}} \right\rangle \leq \theta$ with high probability. To this end, note that $\left\langle \mathbf{r}_u^{\text{sim}}, \mathbf{r}_v^{\text{sim}} \right\rangle = \sum_{t=0}^{T_s-1} R_{u\pi(t)} R_{v\pi(t)}$ where π is a random permutation of the item space and $R_{u\pi(t)} R_{v\pi(t)}$ is a binary random variable, independent across t, with success probability $p_{\text{f}}^2 p_{u\pi(t)} p_{v\pi(t)}$. Setting $a = p_{\text{f}}^2 \frac{1}{M} \left\langle \mathbf{p}_u, \mathbf{p}_v \right\rangle$, for notational convenience, it follows that

$$
\begin{aligned}
\mathbb{P}\left[\frac{1}{T_s}\left\langle \mathbf{r}_u^{\text{sim}}, \mathbf{r}_v^{\text{sim}} \right\rangle \geq \theta\right] &= \mathbb{P}\left[\frac{1}{T_s}\left\langle \mathbf{r}_u^{\text{sim}}, \mathbf{r}_{v'}^{\text{sim}} \right\rangle \geq a + (\theta - a)\right] \\
&\leq e^{-\frac{T_s(\theta-a)^2/2}{a+(\theta-a)/3}} \qquad (40) \\
&\leq e^{-\frac{T_s p_b(\theta/p_b-1)^2/2}{1+(\theta/p_b-1)/3}} = p_{\text{bad}}^2. \qquad (41)
\end{aligned}
$$

Here, (40) follows from Bernstein's inequality. Specifically, we use that π is a random permutation of the item space as well as that $R_{ui}R_{vi}$ are binary random variables independent across i (note that Bernstein's inequality also applies to sampling without replacement, see e.g., (Bardenet and Maillard, 2015)). Finally, for inequality (41), we used that $a \leq p_b = p_{\text{f}}^2 \max_{v \notin \mathcal{T}_u} \frac{1}{M} \left\langle \mathbf{p}_v, \mathbf{p}_u \right\rangle$.

Set $N_{bad} = \sum_{v \notin \mathcal{T}_u} \mathbb{1}\{u \text{ and } v \text{ are declared neighbors}\}$. By inequality (41), we have $\mathbb{E}\left[N_{\text{bad}}\right] \leq p_{\text{bad}}^2 N$. Thus, by Markov's inequality,

$$
\mathbb{P}\left[N_{\text{bad}} \geq N p_{\text{bad}}\right] \leq \frac{\mathbb{E}\left[N_{\text{bad}}\right]}{N p_{\text{bad}}} \leq \frac{p_{\text{bad}}^2 N}{N p_{\text{bad}}} = p_{\text{bad}},
$$

which concludes the proof.

6.3. Proof of Lemma 1 (preference exploration)

Assume w.l.o.g. that $p_{ui} > 1/2 + \Delta$, for all $i \in \mathcal{Q}$. The case where some of the p_{ui} satisfy $p_{ui} < 1/2 - \Delta$ is treated analogously. To prove Lemma 1, we may further assume that $p_{ui} = \frac{1}{2} + \Delta$, for all $i \in \mathcal{Q}$, since $\mathbb{P}\left[\hat{p}_{ui} > \frac{p_{\text{f}}}{2}\right]$ is increasing in p_{ui}.

Consider a fixed item $i \in \mathcal{Q}$, and let $\mathcal{N}_u^{\text{good}}$ be the subset of $\mathcal{N}_u$ corresponding to users that are of the same type as u and to which additionally an recommendation has been made by drawing T_r items uniformly from $\mathcal{Q}$ for each user u. Let N_g be the cardinality of $\mathcal{N}_u^{\text{good}}$. In order to upper-bound $\mathbb{P}\left[\hat{p}_{ui} \leq \frac{p_{\text{f}}}{2}\right]$, we first note that by (21),

$$
\mathbb{P}\left[\hat{p}_{ui} \leq \frac{p_{\text{f}}}{2}\right] \leq \mathbb{P}\left[\hat{p}_{ui} \leq \frac{p_{\text{f}}}{2} \middle| N_g \geq n_g\right] + \mathbb{P}\left[N_g \leq n_g\right]. \qquad (42)
$$

Here, we defined

$$
n_g := \frac{T_r k}{Q}(1/2 - \beta). \qquad (43)
$$

We next upper bound the probabilities on the RHS of (42). We start with the first probability on the RHS of (42):

$$
\begin{aligned}
\mathbb{P}\left[\hat{p}_{ui} \leq \frac{p_{\text{f}}}{2} \middle| N_g = n_g'\right] &\leq \mathbb{P}\left[\frac{\sum_{v \in \mathcal{N}_u^{\text{good}}} R_{vi}}{n_g' + \beta k} \leq \frac{p_{\text{f}}}{2} \middle| N_g = n_g'\right] \qquad (44) \\
&= \mathbb{P}\left[\frac{1}{n_g'} \sum_{v \in \mathcal{N}_u^{\text{good}}} R_{vi} \leq \frac{p_{\text{f}}}{2} \frac{n_g' + \beta k}{n_g'} \middle| N_g = n_g'\right] \\
&= \mathbb{P}\left[\frac{1}{n_g'} \sum_{v \in \mathcal{N}_u^{\text{good}}} R_{vi} \leq p_{\text{f}}\left(\frac{1}{2} + \Delta\right) - p_{\text{f}}\left(\Delta - \frac{\beta k}{2n_g'}\right) \middle| N_g = n_g'\right] \\
&= \mathbb{P}\left[\sum_{v \in \mathcal{N}_u^{\text{good}}} \left(R_{vi} - p_{\text{f}}\left(\frac{1}{2} + \Delta\right)\right) \leq -n_g' p_{\text{f}}\left(\Delta - \frac{\beta k}{2n_g'}\right) \middle| N_g = n_g'\right] \\
&\leq e^{-\frac{n_g' p_{\text{f}}(\Delta - \beta k/(2n_g'))^2/2}{(1/2+\Delta)+(\Delta-\beta k/(2n_g'))/3}} \qquad (45)
\end{aligned}
$$

where (44) follows from the number of users n_{ui} in $\mathcal{N}_u$ that received recommendation i being upper bounded by $N_g + \beta k$ (recall that βk is the maximum number of bad neighbors in $\mathcal{N}_u$), and by assuming adversarially that all recommendations given to bad neighbors did yield $R_{vi} = 0$. Finally, (45) follows from Bernstein's inequality; to apply Bernstein's inequality, we used that $\mathbb{E}[R_{vi}] = p_{\mathrm{f}}(1/2 + \Delta)$, and that the variance of R_{vi} is upper bounded by $p_{\mathrm{f}}(1/2 + \Delta)$, for $v \in \mathcal{N}_u^{\text{good}}$. Next, note that by Bayes theorem,

$$
\begin{aligned}
\mathbb{P}\left[\hat{p}_{ui} \leq 1/2 | N_g \geq n_g\right] &= \frac{\mathbb{P}\left[\{\hat{p}_{ui} \leq 1/2\} \cap \{N_g \geq n_g\}\right]}{\mathbb{P}\left[N_g \geq n_g\right]} \\
&= \frac{\sum_{n'_g \geq n_g} \mathbb{P}\left[\hat{p}_{ui} \leq 1/2 \middle| N_g \geq n_g\right] \mathbb{P}\left[N_g = n'_g\right]}{\mathbb{P}\left[N_g \geq n_g\right]} \\
&\leq e^{-\frac{n_g p_{\mathrm{f}} (\Delta - \beta k/(2n'_g))^2/2}{(1/2+\Delta) + (\Delta - \beta k/(2n'_g))/3}} && (46) \\
&\leq e^{-\frac{n_g p_{\mathrm{f}} \Delta^2/8}{1/2+\Delta+\Delta/6}} \leq e^{-\frac{n_g p_{\mathrm{f}} \Delta^2}{16}} \leq e^{-\frac{T_r k p_{\mathrm{f}} \Delta^2}{Q64}}. && (47)
\end{aligned}
$$

Here, inequality (46) follows from inequality (45) and using that the RHS of inequality (45) is increasing in n'_g. For inequality (47) we used the definition of n_g in (43), and that

$$
\frac{\beta k}{n_g} = \frac{\beta k}{\frac{T_r k}{Q}(1/2 - \beta)} = \frac{Q}{T_r} \frac{\beta}{1/2 - \beta} \leq \Delta. \tag{48}
$$

Here, the inequality (48) holds by $\beta \leq \frac{\Delta T_r}{4Q}$, by assumption, and $\beta \leq 1/4$, due to $\Delta \leq 1/2$ and $T_r \leq Q$ (since we recommend each item at most once).

We proceed with upper bounding $\mathbb{P}[N_g \leq n_g]$ in (42). Recall that N_g is the number of times item i has been recommended to one of the $\geq (1 - \beta)k$ good neighbors in $\mathcal{N}_u$.

We will only consider the T_r random items recommended to each user; this yields an upper bound on $\mathbb{P}[N_g \leq n_g]$. Recall that those items are chosen from the Q items in $\mathcal{Q}$, and that, by assumption, of the k neighbors at least $(1 - \beta)k$ are good. By Bernstein's inequality,

$$
\begin{aligned}
\mathbb{P}\left[N_g \leq n_g\right] &= \mathbb{P}\left[N_g \leq T_r \frac{(1-\beta)k}{Q} - \frac{T_r k}{2Q}\right] \\
&\leq e^{-\frac{T_r k (\frac{1}{2Q})^2/2}{\frac{1-\beta}{Q}(1 - \frac{1-\beta}{Q}) + \frac{1}{3}\frac{1}{2Q}}} \leq e^{-\frac{T_r k (\frac{1}{2Q})^2/2}{\frac{1-\beta}{Q}(1 - \frac{1-\beta}{Q}) + \frac{1}{3}\frac{1}{2Q}}} \leq e^{-\frac{T_r k \frac{1}{8Q}}{1+1/6}} \leq e^{-\frac{T_r k}{10Q}}. && (49)
\end{aligned}
$$

Application of inequalities (47) and (49) to inequality (42) together with a union bound yields

$$
\mathbb{P}\left[\hat{p}_{ui} \leq 1/2, \text{for one or more } i \in \mathcal{Q}\right] \leq Q\left(e^{-\frac{T_r k p_{\mathrm{f}} \Delta^2}{Q64}} + e^{-\frac{T_r k}{10Q}}\right) \leq 2Q e^{-\frac{T_r k p_{\mathrm{f}} \Delta^2}{Q64}}, \tag{50}
$$

where we used that $p_{\mathrm{f}} \Delta^2 \leq 1$. By (26), the RHS above is smaller than δ'. This concludes the proof.

7. Proof of Proposition 1

Consider a set of users with K user types that are non-overlapping in their preferences, specifically, consider a set of users where every user u belonging to the k-th user type has preference vector

$$
[\mathbf{p}_u]_i = \begin{cases} 1, & \text{if } i \in [k(M-1)/K, \ldots, kM/K] \\ 0, & \text{otherwise.} \end{cases}
$$

Consider a given user u. At time T, the expected number of ratings obtained by u is upper bounded by p_{f}^2. Thus, for all $T \leq \frac{\lambda}{p_{\mathrm{f}}^2}$ in at least a fraction λ of the runs of the algorithm, the algorithm has no information on the user u, and the best it can do is to recommend a random item. For our choice of preference vectors, with probability at most $1/K$, it will recommend a likable item. Therefore, an upper bound on the expected regret is given by $(\lambda + 1/K)NT$.