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 $\nu$ 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 $\eta$, are chosen such that $k \leq \frac{9N}{40K}, \eta \leq \nu/2$.

$$\frac{k}{Q} \geq \frac{64 \log(8M/\delta)}{p_t \Delta^2}, \quad (7)$$

and

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

Then the reward accumulated by the User-CF algorithm up to time $T \in [T_{\text{start}}, \frac{1}{2} \nu MP_t]$ with

$$T_{\text{start}} = \left( \frac{512 \max \left( \log \left( \frac{4NQ}{M} \right), \log \left( \frac{8N}{\Delta} \right) \right) \frac{1}{\nu - \alpha}}{(3p_t^2(1 - \gamma)^2\nu)^{\frac{1}{\nu - \alpha}} \left( 1 - \max \left( \frac{1}{T}, \frac{2}{\eta Q} \right) \right) } \right)$$

satisfies

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

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

$$\eta = \frac{\nu}{2}, \quad k = \frac{9N}{40K}, \quad \text{and} \quad Q = k \frac{p_t \Delta^2}{64 \log(8M/\delta)}. \quad (10)$$

To see this, note that by definition, the conditions on $k$ and $\eta$ and condition (7) on $Q$ are satisfied. By (4), condition (8) holds and $\frac{2}{\eta Q} = \frac{K}{N} \frac{\nu \log(M/\delta)}{p_t \Delta^2}$. Moreover, $\max \left( \log \left( \frac{4NQ}{M} \right), \log \left( \frac{8N}{\Delta} \right) \right) \leq \frac{1}{2} \log(N/\delta)$.

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 $\hat{\nu}_{ui}$, for all $i \in Q_u, q = 0, \ldots, \frac{1}{\eta Q} - 1$ correctly predict whether $i$ is likable by $u$ or not, and

iii) there exist items in the sets $Q_u, q = 0, \ldots, \frac{1}{\eta Q} - 1$ 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:

$$\mathcal{G}_\beta(t) = \{ \text{At time } t, \text{ no more than } \beta k \text{ of the } k\text{-nearest neighbors of } u \text{ are bad} \}, \quad (10)$$
\( L(t) = \{ \text{at time } t, \text{ there exists an item } i \in Q_q, \} \),
\[ q = 0, \ldots, t/(\eta Q) - 1 \text{ that is likable by } u \}, \quad (11) \]

and
\[ E(t) = \bigcup_{q=0, \ldots, t/(\eta Q) - 1} E_q(t), \quad (12) \]

with
\[ E_q(t) = \{ \text{Conditioned on } G_{\Delta_q}(t), \text{ for all } i \in Q_q, \} \]
\[ \hat{p}_{ui} > p_t/2, \text{ if } p_{ui} > 1/2 + \Delta, \text{ and } \]
\[ \hat{p}_{ui} < p_t/2, \text{ if } p_{ui} < 1/2 - \Delta \}. \quad (13) \]

For convenience, we omit in the notion of \( L(t), G_\beta(t), E(t), \) and \( E_q(t) \) the dependence on \( u \). The significance of those definitions is that if \( L(t), G_\Delta(t), \) and \( E(t) \) 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 \( \mathbb{E}[\text{reward}(T)] \) as follows:
\[
\frac{\mathbb{E}[\text{reward}(T)]}{NT} \geq \frac{1}{NT} \sum_{t=0}^{T-1} \sum_{u=0}^{N-1} \mathbb{P}[X_{ui(u,t)} = 1]
\geq \frac{1}{NT} \sum_{t=0}^{T-1} \sum_{u=0}^{N-1} \mathbb{P}[\text{exploitation at } t] \mathbb{P}[X_{ui(u,t)} = 1|\text{exploitation at } t] \quad (14)
\geq \frac{1}{N} \sum_{u=0}^{N-1} \left( \frac{1}{T} \sum_{t=0}^{T-1} (1 - (2/t)^\alpha) \mathbb{P}[X_{ui(u,t)} = 1|\text{exploitation at } t] - \max \left( \frac{1}{T}, \frac{2}{\eta Q} \right) \right) \quad (15)
\geq \frac{1}{N} \sum_{u=0}^{N-1} \left( \frac{1}{T} \sum_{t=T_{\text{start}}}^{T-1} (1 - \delta)(1 - (2/t)^\alpha) - \max \left( \frac{1}{T}, \frac{2}{\eta Q} \right) \right) \quad (16)
\geq (1 - \delta) \left( 1 - \frac{T_{\text{start}}}{T} - 2\alpha \left( \frac{T - T_{\text{start}}}{T(1 - \alpha)} \right)^{-\alpha} - \max \left( \frac{1}{T}, \frac{2}{\eta Q} \right) \right). \quad (17)
\]

Here, (14) follows from
\[ \mathbb{P}[X_{ui(u,t)} = 1|\text{preference exploration at } t] \geq 0 \quad \text{and} \quad \mathbb{P}[X_{ui(u,t)} = 1|\text{similarity exploration at } t] \geq 0. \]

For (15) we used, for \( t \neq \eta Q \),
\[
\mathbb{P}[\text{exploitation at } t] = 1 - \left( 1 - \left( t - \lfloor t/(\eta Q) \rfloor \right)^{-\alpha} \right) \geq 1 - \left( 1 - 1/(\eta Q) \right)^{-\alpha} \geq 1 - (2/t)^\alpha
\]

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 \(\frac{1}{T}, \frac{2}{\eta Q}\). 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
\[
\mathbb{P}[X_{ui(u,t)} = 1|\text{exploitation at } t] \geq \mathbb{P}[E(t) \cap G_\Delta(t) \cap L(t)] \geq 1 - \delta. \quad (18)
\]

Here, inequality (18) holds for \( t \geq T_{\text{start}} \) and is established below. Finally, inequality (17) follows from
\[
\sum_{t=T_{\text{start}}}^{T-1} t^{-\alpha} \leq \int_{T_{\text{start}}}^{T-1} t^{-\alpha} = \frac{1}{1 - \alpha} t^{1-\alpha} \bigg|_{t=T_{\text{start}}}^{T-1} - \frac{1}{1 - \alpha} t^{1-\alpha} \bigg|_{t=T_{\text{start}}}^{T-1} - \frac{(T - 1)^{1-\alpha} - (T_{\text{start}} - 1)^{1-\alpha}}{1 - \alpha} \leq \frac{(T - T_{\text{start}})^{1-\alpha}}{1 - \alpha}. \]
\[ \]
Online One-Class Collaborative Filtering

It remains to establish (18). To this end, define for notational convenience
\[
A := \frac{256 \max \left( \log \left( \frac{4NQ}{k^2} \right), \log \left( \frac{8k}{T} \right) \right)}{3p(T^2)(1 - \gamma)^2T},
\]
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,
\[
\mathbb{P} \left[ (\mathcal{E}(t) \cap G_{\mathbb{Q}}(t) \cap \mathcal{L}(t))^c \right] \leq \mathbb{P}[E^c(t)] + \mathbb{P}\left[ G^c_{\mathbb{Q}}(t) \right] + \mathbb{P}[L^c(t)]
\]
\[
\leq \mathbb{P}[E^c(t)] + \mathbb{P}\left[ G^c_{\mathbb{Q}}(t) | T_s \geq A \right] \mathbb{P}[T_s \leq A] + \mathbb{P}[L^c(t)]
\]
\[
\leq \frac{\delta}{4} + \frac{\delta}{4} + \frac{\delta}{4} = \delta.
\]

Here, inequality (19) follows since for two events \( C, B \) we have that
\[
\mathbb{P}[C] = \mathbb{P}[C \cap B] + \mathbb{P}[C \cap B^c] = \mathbb{P}[C|B] \mathbb{P}[B] + \mathbb{P}[C|B^c] \mathbb{P}[B^c] \leq \mathbb{P}[C|B] + \mathbb{P}[B^c].
\]

Inequality (20) follows from
\[
\mathbb{P}[E^c(t)] \leq \delta/4
\]
\[
\mathbb{P}\left[ G^c_{\mathbb{Q}}(t) | T_s \geq A \right] \leq \delta/4
\]
\[
\mathbb{P}[T_s \leq A] \leq \delta/4
\]
\[
\mathbb{P}[L^c(t)] \leq \delta/4.
\]

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
\[
\mathbb{P}[E^c(t)] \leq \sum_{q=0}^{M/2} \mathbb{P}[E^c_q(t)] \leq \frac{\delta}{4}
\]
as desired. Here, we used \( \mathbb{P}[E^c_q(t)] \leq \frac{\delta Q}{2^pT^2} \), which follows from Lemma 1 stated below with \( \delta' = \frac{\delta Q}{2^pT} \) 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 \( \mathbb{Q} \subseteq [M] \) of \( Q \) items. Suppose that \( p_{ui} \) is \( \Delta \)-bounded away from 1/2, for all \( i \in \mathbb{Q} \) and for all \( v \in N_u \), where \( N_u \) is a set of \( k \) users, of which no more than \( \beta k \), with \( \beta \leq \frac{N}{4M} \) of the users are of a different type than \( u \). Fix \( \delta' > 0 \). If
\[
T_r \geq \frac{k \Delta^2}{Q \log(2Q/\delta')}
\]
then, with probability at least \( 1 - \delta' \), for all \( i \in \mathbb{Q} \), \( \tilde{p}_{ui} > \frac{p_{ui}}{T_r} \) if \( p_{ui} \geq 1/2 + \Delta \) and \( \tilde{p}_{ui} < \frac{p_{ui}}{T_r} \) 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 \( T_u \) be the subsets of all users \([N]\) that are of the same type of \( u \) and suppose its cardinality satisfies \( \geq \frac{N}{2T_r} \). Suppose that, for some constant \( \gamma \in (0, 1) \), condition (1) holds, and that the number of nearest neighbors \( k \) satisfies \( k \leq \frac{N}{4M} \). Choose \( \beta \in (0, 1) \), and suppose
\[
T_s \geq \frac{64 \log(N/(\beta k))}{3p(T^2)(1 - \gamma)^2} \min_{v \in T_u} \{ p_{uv} \}
\]
similarity exploration steps have been performed. Then, with probability at least \( 1 - 11e^{-\frac{3}{2}T_sT^2(1 - \gamma)^2} \frac{1}{M} \min_{v \in T_u} \{ p_{uv} \} \), the set of nearest neighbors \( N_u \) of user \( u \) (defined in Section 3), contains no more than \( \beta k \) 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 $\beta = \frac{A}{M}$). To see this, note that since each user likes at least a fraction $\nu$ of the items, we have

$$\frac{1}{M} \min_{v \in T_s} \langle p_u, p_v \rangle \geq \nu \left(\frac{1}{2} + \Delta\right)^2 \geq \nu^2. $$

(28)

Lemma 2 therefore implies

$$P \left[ \mathcal{G}_2^c \left( t \right) \mid T_s \geq A \right] \leq 11e^{-\bar{\alpha}T_s \beta(1-\gamma)^2} \cdot \min_{v \in T_s} \langle p_u, p_v \rangle \leq 11e^{-\log(88/\delta)} = \delta^-,$$

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

**Proof of (24):** We next establish the inequality $P \left[ T_s \leq A \right] \leq \delta^-/4$. 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 $S = \sum_{t=1}^{T_s} Z_t$, $T = T_{\text{start}}(1 - \max(\frac{1}{T}, \frac{2}{\eta Q}))$, where $Z_t$ is a binary random variable with $P \left[ Z_t = 1 \right] = 1/t^\alpha$. It follows that

$$P \left[ T_s \leq A \right] = P \left[ T_s \leq \tilde{T}^{1-\alpha}/2 \right] \leq e^{-\frac{1}{2(1-\alpha)}},$$

(29)

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

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

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

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

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

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

$$P \left[ S_T \leq \tilde{T}^{1-\alpha}/2 \right] \leq e^{-\frac{1}{2(1-\alpha)}},$$

**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 $Q_0$ (recall that $Q_0$ is chosen 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 $Q_0$ 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

$$P \left[ L^{c}(t) \right] \leq P \left[ \eta Q \right] \leq P \left[ N_0 \leq \eta Q \right] \leq P \left[ N_0 \leq \eta Q \right] \leq P \left[ \eta Q \right] - Q\nu/2 \leq e^{-Q\nu/2},$$

(30)

Hence, the first inequality in (30) follows from $\eta \leq \nu/2$, by assumption; the second inequality in (30) follows from $E \left[ N_0 \right] \geq \nu Q$ (since at least a fraction of $\nu$ 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

$$P \left[ L^{c}(t) \right] \leq \delta^-/4, \text{ for } t < \eta Q.$$  

Using the exact same line of arguments yields the same bound for $t \in [\eta Q, \eta Q]$. The following lemma states that

$$P \left[ L^{c}(t) \right] \leq \delta^-/4, \text{ for } t < \eta Q.$$  

It remains to upper bound $P \left[ L^{c}(t) \right]$ for $t \in [\eta Q, \frac{4}{3} \nu M]$. To this end, let $N^c_u(T)$ be the number of (likable) items that have been rated by user $u$ after $T$ time steps, and note that if $N^c_u(T)$ is strictly smaller than the (minimum) number of likable items, then there are likable items left to recommend. Formally,

$$P \left[ L^{c}(t) \right] \leq P \left[ N^c_u(T) \geq \nu M \right]$$

(32)
where we used that for each user $u$, at least $\nu M$ items are likable. Recall that with probability $p_{ui}p_t \leq p_t$ 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 $\Pr[Z_t = 1] = p_t$. We therefore have that

$$\Pr[N_u^c(T) \geq \nu M] \leq \Pr[N_u^c(T) \geq T(p_t + \frac{p_t}{4})] \leq e^{-\frac{p_t^2}{2}} \leq e^{-\frac{2\nu M p_t^2}{3}} \leq \delta^4.$$  
(33)

Here, the first inequality holds by the assumption $T \leq \frac{\delta}{4} \nu M p_t$, the second inequality follows by Hoeffding’s inequality, the third inequality follows by $T \geq T_{\text{start}}$, and the last inequality follows from $T_{\text{start}} \geq \frac{\delta}{2} \log(4/\delta)$, which holds by definition of $T_{\text{start}}$. Application of (33) on (32) concludes the proof of $\Pr[C^c(t)] \leq \delta/4$.

6.2. Proof of Lemma 2

Recall that $r_u^v \in \{0, 1\}^M$ 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 $\frac{1}{T_s} \langle r_u^v, r_v^v \rangle \geq \theta$, and at the same time, there are fewer than $k \beta$ users of a different user type as $u$ that satisfy $\frac{1}{T_s} \langle r_u^v, r_v^v \rangle \geq \theta$ for a certain threshold $\theta$ 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 $1 - 10p_{\text{good}}$,

$$p_{\text{good}} := e^{-\frac{1}{3} T_s p_t (1 - \theta/p_t)^2}, \quad p_g := p_t^2 \min_{v \in T_u} \langle p_u, p_v \rangle,$$

at least $\frac{9N}{40T_s}$ users $v$ of the same type as $u$ obey $\frac{1}{T_s} \langle r_u^v, r_v^v \rangle \geq \theta$.

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

$$p_{\text{bad}} = e^{-\frac{1}{3} T_s p_t (1 - \theta/p_t)^2/4}, \quad p_b := p_t^2 \max_{v \in T_u} \langle p_v, p_u \rangle,$$

at most $N p_{\text{bad}}$ users $v$ of a different type as $u$ obey $\frac{1}{T_s} \langle r_u^v, r_v^v \rangle \geq \theta$. Thus, by the union bound, $N_u$ contains less than $p_{\text{bad}} N$ bad neighbors with probability at least $1 - 10p_{\text{good}} - p_{\text{bad}} \geq 1 - 11e^{-\frac{1}{3} T_s p_t (1 - \gamma)^2}$.

Here, we used

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

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

$$p_{\text{bad}} = e^{-\frac{1}{3} T_s p_t (1 - \gamma/p_t)^2/4} \leq e^{-\frac{1}{3} T_s p_t (1 - \gamma/p_t)^2/16} = e^{-\frac{1}{3} T_s p_t (1 - \gamma/p_t)^2/6} \leq e^{-\frac{1}{3} T_s p_t (1 - \gamma/p_t)^2/16} \leq e^{-\frac{1}{3} T_s p_t (1 - \gamma/p_t)^2/16} \leq e^{-\frac{1}{3} T_s p_t (1 - \gamma)^2/16} \leq e^{-\frac{1}{3} T_s p_t (1 - \gamma)^2/16} \leq e^{-\frac{1}{3} T_s p_t (1 - \gamma)^2/16} \leq e^{-\frac{1}{3} T_s p_t (1 - \gamma)^2/16} \leq e^{-\frac{1}{3} T_s p_t (1 - \gamma)^2/16}.$$ 
(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 $N p_{\text{bad}}$ satisfies $N p_{\text{bad}} \leq \beta k$. This follows directly by noting that, by assumption (27), the RHS of (34) is upper-bounded by $\frac{\beta}{N}$. 
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 \( \frac{1}{T_s} \langle r_u^{\text{sim}}, r_v^{\text{sim}} \rangle \geq \theta \) with high probability. To this end, note that \( \langle r_u^{\text{sim}}, r_v^{\text{sim}} \rangle = \sum_{t=0}^{T_s-1} P_{u \pi(t)} R_{u \pi(t)} \) where \( \pi \) is the random permutation of the item space drawn by the User-CF algorithm at initialization, and \( R_{u \pi(t)} \) is a binary random variable, independent across \( t \), with success probability \( p_t^2 p_{u \pi(t)} p_{v \pi(t)} \). Setting \( \alpha := p_t^2 \pi \langle \hat{p}_u, \hat{p}_v \rangle \), for notational convenience, it follows that

\[
P \left[ \frac{1}{T_s} \langle r_u^{\text{sim}}, r_v^{\text{sim}} \rangle \leq \theta \right] = P \left[ \frac{1}{T_s} \langle r_u^{\text{sim}}, r_v^{\text{sim}} \rangle \leq \alpha - (\alpha - \theta) \right] \leq e^{-a(\alpha - \theta)^2/2} \leq e^{-a(\alpha - \theta)/(1-\theta)\alpha^2/2} \leq e^{-a\alpha(1-\theta)/\alpha^2} \leq e^{-\alpha p_t(1-\theta)/\alpha^2} \leq p_{\text{good}}. \tag{38}
\]

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

Next, consider the random variable

\[ W = \sum_{v \in T_u} G_v, \quad G_v = 1 \left\{ \frac{1}{T_s} \langle r_u^{\text{sim}}, r_v^{\text{sim}} \rangle \geq \theta \right\}, \]

where \( T_u \) is the subset of all users \([N]\) that are of the same time as user \( u \), as before. By Chebyshev’s inequality,

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

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

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

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

\[
\text{Var}(W) = \sum_{v \in T_u} \text{Var}(G_v) + \sum_{v, v' \in T_u, v \neq v'} \text{Cov}(G_v, G_{v'}).
\]

With \( G_v = G_v^2 \),

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

Similarly,

\[
\text{Cov}(G_v, G_{v'}) = \mathbb{E}[G_v G_{v'}] - \mathbb{E}[G_v] \mathbb{E}[G_{v'}] \leq 1 - (1-q)^2 \leq 2p_{\text{good}}.
\]

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

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

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

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

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

$$\frac{1}{T_s} \langle r_u^{\text{sim}}, r_v^{\text{sim}} \rangle \leq \theta$$

with high probability. To this end, note that $\langle r_u^{\text{sim}}, r_v^{\text{sim}} \rangle = \sum_{t=0}^{T_s-1} R_{u\pi(t)} R_{v\pi(t)}$ where $\pi$ 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 $p_{uv}(t) P_{uv}(t)$. Setting $a = \frac{p_{uv}}{M^2} (p_u, p_v)$, for notational convenience, it follows that

$$\mathbb{P} \left[ \frac{1}{T_s} \langle r_u^{\text{sim}}, r_v^{\text{sim}} \rangle \geq \theta \right] = \mathbb{P} \left[ \frac{1}{T_s} \langle r_u^{\text{sim}}, r_v^{\text{sim}} \rangle \geq \alpha + (\theta - \alpha) \right] \leq e^{- \frac{\alpha - \theta}{\alpha + (\theta - \alpha)/2}}$$

(40)

(41)

Here, (40) follows from Bernstein’s inequality. Specifically, we use that $\pi$ is a random permutation of the item space as well as that $R_{u\pi(t)} R_{v\pi(t)}$ 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 $a \leq p_{uv} = \frac{p_{uv}}{M^2} \max_{u \in T_s} \frac{1}{M^2} (p_u, p_v)$.

Set $N_{\text{bad}} = \sum_{v \in T_s} \mathbb{1} \{u \text{ and } v \text{ are declared neighbors}\}$. By inequality (41), we have $\mathbb{E}[N_{\text{bad}}] \leq p_{\text{bad}}^2 N$. Thus, by Markov’s inequality,

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

which concludes the proof.

6.3. Proof of Lemma 1 (expectation exploration)

Assume w.l.o.g. that $p_{ui} > 1/2 + \Delta$, for all $i \in Q$. 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 $i \in Q$, since $\mathbb{P}[\hat{p}_{ui} \leq \frac{1}{2}]$ is increasing in $p_{ui}$.

Consider a fixed item $i \in Q$, and let $N_u^{\text{good}}$ be the subset of $N_u$ corresponding to users that are of the same type as $u$ and to which additionally an recommendation has been made by drawing $T_i$ items uniformly from $Q$ for each user $u$. Let $N_g$ be the cardinality of $N_u^{\text{good}}$. In order to upper-bound $\mathbb{P}[\hat{p}_{ui} \leq \frac{1}{2}]$, we first note that by (21),

$$\mathbb{P}[\hat{p}_{ui} \leq \frac{p_{il}}{2}] \leq \mathbb{P} \left[ \hat{p}_{ui} \leq \frac{p_{il}}{2} \middle| N_g \geq n_g \right] + \mathbb{P}[N_g \leq n_g].$$

(42)

Here, we defined

$$n_g := \frac{T_i}{Q} \left( \frac{1}{2} - \beta \right).$$

(43)

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

$$\mathbb{P}[\hat{p}_{ui} \leq \frac{p_{il}}{2} \middle| N_g = n_g'] \leq \mathbb{P} \left[ \sum_{v \in N_u^{\text{good}}} R_{vi} \geq \frac{n_g' + \beta k}{2} \middle| N_g = n_g' \right]$$

(44)

$$= \mathbb{P} \left[ \frac{1}{n_g'} \sum_{v \in N_u^{\text{good}}} R_{vi} \leq \frac{p_{il}}{2} \left( \frac{1}{2} + \frac{\beta k}{n_g'} \right) \middle| N_g = n_g' \right]$$

$$= \mathbb{P} \left[ \sum_{v \in N_u^{\text{good}}} R_{vi} \leq p_{il} \left( \frac{1}{2} + \Delta \right) - p_{il} \left( \Delta - \frac{\beta k}{2n_g'} \right) \middle| N_g = n_g' \right]$$

$$= \mathbb{P} \left[ \sum_{v \in N_u^{\text{good}}} \left( R_{vi} - p_{il} \left( \frac{1}{2} + \Delta \right) \right) \leq -n_g' p_{il} \left( \Delta - \frac{\beta k}{2n_g'} \right) \middle| N_g = n_g' \right]$$

$$\leq e^{- \frac{n_g' p_{il} \left( \Delta - \frac{\beta k}{2n_g'} \right)^2}{\left( \frac{1}{2} + \Delta \right)(\Delta - \frac{\beta k}{2n_g'})}}$$

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

\[
\mathbb{P}[\hat{R}_{ui} \leq 1/2 | N_g \geq n_g] = \frac{\mathbb{P}[\{\hat{R}_{ui} \leq 1/2\} \cap \{N_g \geq n_g\}]}{\mathbb{P}[N_g \geq n_g]} = \sum_{n_g \geq n_g} \mathbb{P}[\hat{R}_{ui} \leq 1/2 | N_g \geq n_g] \mathbb{P}[N_g = n_g] \frac{1}{\mathbb{P}[N_g \geq n_g]}
\]

\[
\leq e^{-\frac{n_g p_i (\Delta - \beta k)(2n_g)^2}{2}}
\]

\[
\leq e^{-\frac{n_g p_i \Delta^2}{8}}
\]

(46)

(47)

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

\[
\frac{\beta k}{n_g} = \frac{\beta k}{T_r (1/2 - \beta)} = \frac{Q}{T_r} \frac{\beta}{1/2 - \beta} \leq \Delta.
\]

(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 \( Q \), and that, by assumption, of the \( k \) neighbors at least \( (1 - \beta)k \) are good. By Bernstein’s inequality,

\[
\mathbb{P}[N_g \leq n_g] = \mathbb{P}[N_g \leq T_r \frac{(1 - \beta)k}{Q} - \frac{T_r k}{2Q}]
\]

\[
\leq e^{-\frac{T_r k (\frac{(1 - \beta)k}{Q})^2}{2}} \leq e^{-\frac{T_r k (\frac{(1 - \beta)k}{Q})^2}{2}} \leq e^{-\frac{T_r k (\frac{(1 - \beta)k}{Q})^2}{2}} \leq e^{-\frac{T_r k (\frac{(1 - \beta)k}{Q})^2}{2}}
\]

(49)

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

\[
\mathbb{P}[\hat{R}_{ui} \leq 1/2, \text{for one or more } i \in Q] \leq Q \left(e^{-\frac{T_r k (\frac{(1 - \beta)k}{Q})^2}{2}} + e^{-\frac{T_r k (\frac{(1 - \beta)k}{Q})^2}{2}}\right) \leq 2Q e^{-\frac{T_r k (\frac{(1 - \beta)k}{Q})^2}{2}}
\]

(50)

where we used that \( p_i \Delta^2 \leq 1 \). By (26), the RHS above is smaller than \( \delta' \). 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

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

Consider a given user \( u \). At time \( T \), the expected number of ratings obtained by \( u \) is upper bounded by \( p_i^2 \). Thus, for all \( T \leq \frac{\Delta}{p_i^2} \) in at least a fraction \( \lambda \) 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 \( (\lambda + 1/K)NT \).