A Proof of Theorem 1

Proof. For a candidate selection vector $\boldsymbol{o}$, we have $\hat{\boldsymbol{\theta}}(\boldsymbol{o}) = \arg\min_{\boldsymbol{\theta}} \frac{1}{Z}\sum_v \sum_{k=1}^{|\mathcal{S}^v|} o_k^v l_f(s_k^v, \boldsymbol{\theta})$, where if $o_k^v = 0$, then s_k^v is not leveraged to train the model. Suppose we define:

$$\hat{\boldsymbol{\theta}}(\boldsymbol{o}, \epsilon) = \arg\min_{\boldsymbol{\theta}} [\frac{1}{Z}\sum_v \sum_{k=1}^{|\mathcal{S}^v|} \widetilde{o}_k^v l_f(s_k^v, \boldsymbol{\theta}) + \epsilon \sum_v \sum_{k=1}^{|\mathcal{S}^v|} (\widetilde{o}_k^v - o_k^v) l_f(s_k^v, \boldsymbol{\theta})]. \tag{1}$$

Then we have $\hat{\boldsymbol{\theta}}(\boldsymbol{o}) = \hat{\boldsymbol{\theta}}(\boldsymbol{o}, -\frac{1}{Z})$ and $\widetilde{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}}(\boldsymbol{o}, 0)$.

Since $\hat{\boldsymbol{\theta}}(\boldsymbol{o}, \epsilon)$ is the optimal solution of (1), then

$$\begin{aligned} 0 &\approx \triangledown [\frac{1}{Z}\sum_v \sum_{k=1}^{|\mathcal{S}^v|} \widetilde{o}_k^v l_f(s_k^v, \hat{\boldsymbol{\theta}}(\boldsymbol{o}^u, \epsilon)) + \epsilon \sum_v \sum_{k=1}^{|\mathcal{S}^v|} (\widetilde{o}_k^v - o_k^v) l_f(s_k^v, \hat{\boldsymbol{\theta}}(\boldsymbol{o}^u, \epsilon))] \\ &= \frac{1}{Z}\sum_u \sum_{k=1}^{|\mathcal{S}^u|} \widetilde{o}_k^v \triangledown l_f(s_k^v, \hat{\boldsymbol{\theta}}(\boldsymbol{o}^u, \epsilon)) + \epsilon \sum_v \sum_{k=1}^{|\mathcal{S}^v|} (\widetilde{o}_k^v - o_k^v) \triangledown l_f(s_k^u, \hat{\boldsymbol{\theta}}(\boldsymbol{o}^u, \epsilon))]. \end{aligned} \tag{2}$$

We regard $\frac{1}{Z}\sum_v \sum_{k=1}^{|\mathcal{S}^v|} \widetilde{o}_k^v \triangledown l_f(s_k^v, \hat{\boldsymbol{\theta}}(\boldsymbol{o}, \epsilon)) + \epsilon \sum_v \sum_{k=1}^{|\mathcal{S}^v|} (\widetilde{o}_k^v - o_k^v) \triangledown l_f(s_k^u, \hat{\boldsymbol{\theta}}(\boldsymbol{o}, \epsilon))]$ as a function of $\hat{\boldsymbol{\theta}}(\boldsymbol{o}, \epsilon)$. If $\epsilon \to 0$, then $\hat{\boldsymbol{\theta}}(\boldsymbol{o}, \epsilon) \to \widetilde{\boldsymbol{\theta}}$. According to the Taylor expansion, we have:

$$\begin{aligned} 0 &\approx \frac{1}{Z}\sum_u \sum_{k=1}^{|\mathcal{S}^u|} \widetilde{o}_k^v \triangledown l_f(s_k^u, \widetilde{\boldsymbol{\theta}}) + \epsilon \sum_v \sum_{k=1}^{|\mathcal{S}^v|} (\widetilde{o}_k^v - o_k^v) \triangledown l_f(s_k^v, \widetilde{\boldsymbol{\theta}}) \\ &+ [\frac{1}{Z}\sum_v \sum_{k=1}^{|\mathcal{S}^v|} \widetilde{o}_k^v \triangledown^2 l_f(s_k^v, \widetilde{\boldsymbol{\theta}}) + \epsilon \sum_v \sum_{k=1}^{|\mathcal{S}^v|} (\widetilde{o}_k^v - o_k^v) \triangledown^2 l_f(s_k^v, \widetilde{\boldsymbol{\theta}})](\hat{\boldsymbol{\theta}}(\boldsymbol{o}, \epsilon) - \widetilde{\boldsymbol{\theta}}) \end{aligned} \tag{3}$$

Let $\triangle_\epsilon = \hat{\boldsymbol{\theta}}(\boldsymbol{o}, \epsilon) - \widetilde{\boldsymbol{\theta}}$, then

$$\begin{aligned} 0 &\approx \frac{1}{Z}\sum_v \sum_{k=1}^{|\mathcal{S}^v|} \widetilde{o}_k^v \triangledown l_f(s_k^v, \widetilde{\boldsymbol{\theta}}) + \epsilon \sum_v \sum_{k=1}^{|\mathcal{S}^v|} (\widetilde{o}_k^v - o_k^v) \triangledown l_f(s_k^v, \widetilde{\boldsymbol{\theta}}) \\ &+ [\frac{1}{Z}\sum_v \sum_{k=1}^{|\mathcal{S}^v|} \widetilde{o}_k^v \triangledown^2 l_f(s_k^v, \widetilde{\boldsymbol{\theta}}) + \epsilon \sum_v \sum_{k=1}^{|\mathcal{S}^v|} (\widetilde{o}_k^v - o_k^v) \triangledown^2 l_f(s_k^v, \widetilde{\boldsymbol{\theta}})]\triangle_\epsilon \\ \triangle_\epsilon &\approx -[\frac{1}{Z}\sum_v \sum_{k=1}^{|\mathcal{S}^v|} \widetilde{o}_k^v \triangledown^2 l_f(s_k^v, \widetilde{\boldsymbol{\theta}}) + \epsilon \sum_v \sum_{k=1}^{|\mathcal{S}^v|} (\widetilde{o}_k^v - o_k^v) \triangledown^2 l_f(s_k^v, \widetilde{\boldsymbol{\theta}})]^{-1} \\ &[\frac{1}{Z}\sum_v \sum_{k=1}^{|\mathcal{S}^v|} \widetilde{o}_k^v \triangledown l_f(s_k^v, \widetilde{\boldsymbol{\theta}}) + \epsilon \sum_v \sum_{k=1}^{|\mathcal{S}^v|} (\widetilde{o}_k^v - o_k^v) \triangledown l_f(s_k^v, \widetilde{\boldsymbol{\theta}})] \end{aligned} \tag{4}$$

Since $| \triangledown^2 l_f(s_k^v, \widetilde{\boldsymbol{\theta}})| \le B$ and $\sum_v \sum_{k=1}^{|\mathcal{S}^v|} (\widetilde{o}_k^v - o_k^v) B$ is a small value, we can ignore the term $\epsilon \sum_v \sum_{k=1}^{|\mathcal{S}^v|} (\widetilde{o}_k^v - o_k^v) \triangledown^2 l_f(s_k^v, \widetilde{\boldsymbol{\theta}})$, which lead to the following equation:

$$\triangle_\epsilon \approx -[\frac{1}{Z}\sum_v \sum_{k=1}^{|\mathcal{S}^v|} \widetilde{o}_k^v \triangledown^2 l_f(s_k^v, \widetilde{\boldsymbol{\theta}})]^{-1} [\frac{1}{Z}\sum_v \sum_{k=1}^{|\mathcal{S}^v|} \widetilde{o}_k^v \triangledown l_f(s_k^v, \widetilde{\boldsymbol{\theta}}) + \epsilon \sum_v \sum_{k=1}^{|\mathcal{S}^v|} (\widetilde{o}_k^v - o_k^v) \triangledown l_f(s_k^v, \widetilde{\boldsymbol{\theta}})] \tag{5}$$

If Z is a large value, then $\frac{d\hat{\boldsymbol{\theta}}(\boldsymbol{o}, \epsilon)}{d\epsilon}|_{\epsilon \to 0} \approx \frac{\triangle_{-\frac{1}{Z}}}{-\frac{1}{Z}}$. According to (5),

$$\begin{aligned} \triangle_{-\frac{1}{Z}} &\approx -[\frac{1}{Z}\sum_v \sum_{k=1}^{|\mathcal{S}^v|} \widetilde{o}_k^v \triangledown^2 l_f(s_k^v, \widetilde{\boldsymbol{\theta}})]^{-1} [\frac{1}{Z}\sum_v \sum_{k=1}^{|\mathcal{S}^v|} \widetilde{o}_k^v \triangledown l_f(s_k^v, \widetilde{\boldsymbol{\theta}}) - \frac{1}{Z} \sum_v \sum_{k=1}^{|\mathcal{S}^v|} (\widetilde{o}_k^v - o_k^v) \triangledown l_f(s_k^v, \widetilde{\boldsymbol{\theta}})] \\ &= -[\frac{1}{Z}\sum_v \sum_{k=1}^{|\mathcal{S}^v|} \widetilde{o}_k^v \triangledown^2 l_f(s_k^v, \widetilde{\boldsymbol{\theta}})]^{-1} [\frac{1}{Z}\sum_v \sum_{k=1}^{|\mathcal{S}^v|} o_k^v \triangledown l_f(s_k^v, \widetilde{\boldsymbol{\theta}})] \end{aligned} \tag{6}$$

Thus, $\frac{d\hat{\boldsymbol{\theta}}(\boldsymbol{o},\epsilon)}{d\epsilon}|_{\epsilon\to 0} \approx \frac{\triangle_{-\frac{1}{Z}}}{-\frac{1}{Z}} = [\frac{1}{Z}\sum_v \sum_{k=1}^{|\mathcal{S}^v|} \widetilde{o}_k^v \bigtriangledown^2 l_f(s_k^v, \widetilde{\boldsymbol{\theta}})]^{-1}[\sum_v \sum_{k=1}^{|\mathcal{S}^v|} o_k^v \bigtriangledown l_f(s_k^v, \widetilde{\boldsymbol{\theta}})]$

Because

$$
\begin{aligned}
\frac{L_f(\mathcal{T}^u, \hat{\boldsymbol{\theta}}(\boldsymbol{o})) - L_f(\mathcal{T}^u, \widetilde{\boldsymbol{\theta}})}{-\frac{1}{Z}} = \frac{L_f(\mathcal{T}^u, \hat{\boldsymbol{\theta}}(\boldsymbol{o}, -\frac{1}{Z})) - L_f(\mathcal{T}^u, \widetilde{\boldsymbol{\theta}})}{-\frac{1}{Z}} &\approx \frac{dL_f(\mathcal{T}^u, \hat{\boldsymbol{\theta}}(\boldsymbol{o},\epsilon))}{d\epsilon}|_{\epsilon\to 0} \\
&= \sum_{y\in\mathcal{T}^u} \bigtriangledown l_f(y, \widetilde{\boldsymbol{\theta}}) \times \frac{d\hat{\boldsymbol{\theta}}(\boldsymbol{o},\epsilon)}{d\epsilon}|_{\epsilon\to 0} \\
&\approx \sum_{y\in\mathcal{T}^u} \bigtriangledown l_f(y, \widetilde{\boldsymbol{\theta}}) H_{\widetilde{\boldsymbol{\theta}}}^{-1} [\sum_v \sum_{k=1}^{|\mathcal{S}^v|} o_k^v \bigtriangledown l_f(s_k^v, \widetilde{\boldsymbol{\theta}})]
\end{aligned}
\tag{7}
$$

where $H_{\widetilde{\boldsymbol{\theta}}} = \frac{1}{Z}\sum_v \sum_{k=1}^{|\mathcal{S}^v|} \widetilde{o}_k^v \bigtriangledown^2 l_f(s_k^v, \widetilde{\boldsymbol{\theta}})$, then we have:

$$
\begin{aligned}
L_f(\mathcal{T}^u, \hat{\boldsymbol{\theta}}(\boldsymbol{o})) - L_f(\mathcal{T}^u, \widetilde{\boldsymbol{\theta}}) &\approx -\frac{1}{Z} \sum_{y\in\mathcal{T}^u} \bigtriangledown l_f(y, \widetilde{\boldsymbol{\theta}}) H_{\widetilde{\boldsymbol{\theta}}}^{-1} [\sum_v \sum_{k=1}^{|\mathcal{S}^v|} o_k^v \bigtriangledown l_f(s_k^v, \widetilde{\boldsymbol{\theta}})] \\
L_f(\mathcal{T}^u, \hat{\boldsymbol{\theta}}(\boldsymbol{o})) &\approx L_f(\mathcal{T}^u, \widetilde{\boldsymbol{\theta}}) - \frac{1}{Z} \sum_{y\in\mathcal{T}^u} \sum_v \sum_{k=1}^{|\mathcal{S}^v|} o_k^v \bigtriangledown l_f(y, \widetilde{\boldsymbol{\theta}}) H_{\widetilde{\boldsymbol{\theta}}}^{-1} \bigtriangledown l_f(s_k^v, \widetilde{\boldsymbol{\theta}})
\end{aligned}
\tag{8}
$$

□

B Proof of Theorem 2

Proof. By bringing the result of theorem 1 into $\overline{z}_u(\boldsymbol{o}^u, \boldsymbol{o}^{-u})$, we have

$$
\overline{z}_u(\boldsymbol{o}^u, \boldsymbol{o}^{-u}) = -L_f(\mathcal{T}^u, \widetilde{\boldsymbol{\theta}}) + \frac{1}{Z} \sum_{y\in\mathcal{T}^u} \sum_{v\in\mathcal{U}} \sum_{k=1}^{|\mathcal{S}^v|} o_k^v \bigtriangledown l_f(y, \widetilde{\boldsymbol{\theta}}) H_{\widetilde{\boldsymbol{\theta}}}^{-1} \bigtriangledown l_f(s_k^v, \widetilde{\boldsymbol{\theta}}) - \lambda \sum_{k=1}^{|\mathcal{S}^u|} o_k^u \beta_k^u.
\tag{9}
$$

Recall that $\boldsymbol{g}_y^v = [g(s_1^v, y), g(s_2^v, y), ..., g(s_{|\mathcal{S}^v|}^v, y)]$, where $g(s_k^v, y) = \bigtriangledown_\theta l_f(y, \widetilde{\boldsymbol{\theta}})^T H_{\widetilde{\boldsymbol{\theta}}}^{-1} \bigtriangledown_\theta l_f(s_k^v, \widetilde{\boldsymbol{\theta}})$, then

$$
\begin{aligned}
& z_u(\boldsymbol{\alpha}^u, \boldsymbol{\alpha}^{-u}) = E_{\boldsymbol{o}}[\overline{z}_u(\boldsymbol{o}^u, \boldsymbol{o}^{-u})] \\
&= -L_f(\mathcal{T}^u, \widetilde{\boldsymbol{\theta}}) + E_{\boldsymbol{o}}[\frac{1}{Z} \sum_{y\in\mathcal{T}^u} \sum_{v\in\mathcal{U}} \sum_{k=1}^{|\mathcal{S}^v|} o_k^v g(s_k^{u'}, y) - \lambda \sum_{k=1}^{|\mathcal{S}^u|} o_k^u \beta_k^u] \\
&= -L_f(\mathcal{T}^u, \widetilde{\boldsymbol{\theta}}) + E_{\boldsymbol{o}}[\frac{1}{Z} \sum_{y\in\mathcal{T}^u} \sum_{v\in\mathcal{U}} (\boldsymbol{o}^v)^T \boldsymbol{g}_y^v - \lambda (\boldsymbol{o}^u)^T \boldsymbol{\beta}^u] \\
&= -L_f(\mathcal{T}^u, \widetilde{\boldsymbol{\theta}}) + E_{\boldsymbol{o}}[\sum_{v\in\mathcal{U}} (\boldsymbol{o}^v)^T \sum_{y\in\mathcal{T}^u} \frac{\boldsymbol{g}_y^v}{Z} - \lambda (\boldsymbol{o}^u)^T \boldsymbol{\beta}^u] \\
&= -L_f(\mathcal{T}^u, \widetilde{\boldsymbol{\theta}}) + \sum_{v\neq u} E_{\boldsymbol{o}}[(\boldsymbol{o}^v)^T \sum_{y\in\mathcal{T}^u} \frac{\boldsymbol{g}_y^v}{Z}] + E_{\boldsymbol{o}}[(\boldsymbol{o}^u)^T \sum_{y\in\mathcal{T}^u} \frac{\boldsymbol{g}_y^u}{Z}] - \lambda E_{\boldsymbol{o}}[(\boldsymbol{o}^u)^T \boldsymbol{\beta}^u] \\
&= -L_f(\mathcal{T}^u, \widetilde{\boldsymbol{\theta}}) + \sum_{v\neq u} E_{\boldsymbol{o}}[(\boldsymbol{o}^v)^T \sum_{y\in\mathcal{T}^u} \frac{\boldsymbol{g}_y^v}{Z}] + E_{\boldsymbol{o}^u}[(\boldsymbol{o}^u)^T (\sum_{y\in\mathcal{T}^u} \frac{\boldsymbol{g}_y^u}{Z} - \lambda \boldsymbol{\beta}^u)]
\end{aligned}
\tag{10}
$$

□

C Proof of Theorem 3

Proof. The proof of this Theorem is similar to that of theorem 1. We define:

$$
\hat{\boldsymbol{\theta}}(\boldsymbol{o}, \epsilon) = \arg\min_{\boldsymbol{\theta}} [\frac{1}{Z} \sum_v \sum_{k=1}^{|\mathcal{S}^v|} \widetilde{\boldsymbol{o}}_k^{t,v} l_f(s_k^v, \boldsymbol{\theta}) + \epsilon \sum_v \sum_{k=1}^{|\mathcal{S}^v|} (\widetilde{\boldsymbol{o}}_k^{t,v} - o_k^v) l_f(s_k^v, \boldsymbol{\theta})].
\tag{11}
$$

Then,

$$
\begin{aligned}
0 &\approx \bigtriangledown[\frac{1}{Z}\sum_{v}\sum_{k=1}^{|\mathcal{S}^v|}\widetilde{o}_k^{t,v} l_f(s_k^v, \hat{\boldsymbol{\theta}}(\boldsymbol{o},\epsilon)) + \epsilon\sum_{v}\sum_{k=1}^{|\mathcal{S}^v|}(\widetilde{o}_k^{t,v} - o_k^v) l_f(s_k^v, \hat{\boldsymbol{\theta}}(\boldsymbol{o},\epsilon))] \\
&= \frac{1}{Z}\sum_{u}\sum_{k=1}^{|\mathcal{S}^u|}\widetilde{o}_k^{t,v} \bigtriangledown l_f(s_k^v, \hat{\boldsymbol{\theta}}(\boldsymbol{o},\epsilon)) + \epsilon\sum_{v}\sum_{k=1}^{|\mathcal{S}^v|}(\widetilde{o}_k^{t,v} - o_k^v) \bigtriangledown l_f(s_k^u, \hat{\boldsymbol{\theta}}(\boldsymbol{o},\epsilon))].
\end{aligned}
\tag{12}
$$

According to Taylor expansion at point $\widetilde{\boldsymbol{\theta}}^t$, we have:

$$
\begin{aligned}
0 &\approx \frac{1}{Z}\sum_{u}\sum_{k=1}^{|\mathcal{S}^u|}\widetilde{o}_k^{t,v} \bigtriangledown l_f(s_k^u, \widetilde{\boldsymbol{\theta}}^t) + \epsilon\sum_{v}\sum_{k=1}^{|\mathcal{S}^v|}(\widetilde{o}_k^{t,v} - o_k^v) \bigtriangledown l_f(s_k^v, \widetilde{\boldsymbol{\theta}}^t) \\
&+[\frac{1}{Z}\sum_{v}\sum_{k=1}^{|\mathcal{S}^v|}\widetilde{o}_k^{t,v} \bigtriangledown^2 l_f(s_k^v, \widetilde{\boldsymbol{\theta}}^t) + \epsilon\sum_{v}\sum_{k=1}^{|\mathcal{S}^v|}(\widetilde{o}_k^{t,v} - o_k^v) \bigtriangledown^2 l_f(s_k^v, \widetilde{\boldsymbol{\theta}}^t)](\hat{\boldsymbol{\theta}}(\boldsymbol{o},\epsilon) - \widetilde{\boldsymbol{\theta}}^t)
\end{aligned}
\tag{13}
$$

Let $\triangle_\epsilon = \hat{\boldsymbol{\theta}}(\boldsymbol{o},\epsilon) - \widetilde{\boldsymbol{\theta}}^t$, then:

$$
\begin{aligned}
0 &\approx \frac{1}{Z}\sum_{v}\sum_{k=1}^{|\mathcal{S}^v|}\widetilde{o}_k^{t,v} \bigtriangledown l_f(s_k^v, \widetilde{\boldsymbol{\theta}}^t) + \epsilon\sum_{v}\sum_{k=1}^{|\mathcal{S}^v|}(\widetilde{o}_k^{t,v} - o_k^v) \bigtriangledown l_f(s_k^v, \widetilde{\boldsymbol{\theta}}^t) \\
&+[\frac{1}{Z}\sum_{v}\sum_{k=1}^{|\mathcal{S}^v|}\widetilde{o}_k^{t,v} \bigtriangledown^2 l_f(s_k^v, \widetilde{\boldsymbol{\theta}}^t) + \epsilon\sum_{v}\sum_{k=1}^{|\mathcal{S}^v|}(\widetilde{o}_k^{t,v} - o_k^v) \bigtriangledown^2 l_f(s_k^v, \widetilde{\boldsymbol{\theta}}^t)]\triangle_\epsilon \\
\triangle_\epsilon &\approx -[\frac{1}{Z}\sum_{v}\sum_{k=1}^{|\mathcal{S}^v|}\widetilde{o}_k^{t,v} \bigtriangledown^2 l_f(s_k^v, \widetilde{\boldsymbol{\theta}}^t) + \epsilon\sum_{v}\sum_{k=1}^{|\mathcal{S}^v|}(\widetilde{o}_k^{t,v} - o_k^v) \bigtriangledown^2 l_f(s_k^v, \widetilde{\boldsymbol{\theta}}^t)]^{-1} \\
&[\frac{1}{Z}\sum_{v}\sum_{k=1}^{|\mathcal{S}^v|}\widetilde{o}_k^{t,v} \bigtriangledown l_f(s_k^v, \widetilde{\boldsymbol{\theta}}^t) + \epsilon\sum_{v}\sum_{k=1}^{|\mathcal{S}^v|}(\widetilde{o}_k^{t,v} - o_k^v) \bigtriangledown l_f(s_k^v, \widetilde{\boldsymbol{\theta}}^t)]
\end{aligned}
\tag{14}
$$

By ignoring $\epsilon\sum_v\sum_{k=1}^{|\mathcal{S}^v|}(\widetilde{o}_k^{t,v} - o_k^v) \bigtriangledown^2 l_f(s_k^v, \widetilde{\boldsymbol{\theta}})$, we have:

$$
\triangle_\epsilon \approx -[\frac{1}{Z}\sum_{v}\sum_{k=1}^{|\mathcal{S}^v|}\widetilde{o}_k^{t,v} \bigtriangledown^2 l_f(s_k^v, \widetilde{\boldsymbol{\theta}}^t)]^{-1}[\frac{1}{Z}\sum_{v}\sum_{k=1}^{|\mathcal{S}^v|}\widetilde{o}_k^{t,v} \bigtriangledown l_f(s_k^v, \widetilde{\boldsymbol{\theta}}^t) + \epsilon\sum_{v}\sum_{k=1}^{|\mathcal{S}^v|}(\widetilde{o}_k^{t,v} - o_k^v) \bigtriangledown l_f(s_k^v, \widetilde{\boldsymbol{\theta}}^t)]
\tag{15}
$$

Then

$$
\begin{aligned}
\triangle_{-\frac{1}{Z}} &\approx -[\frac{1}{Z}\sum_{v}\sum_{k=1}^{|\mathcal{S}^v|}\widetilde{o}_k^{t,v} \bigtriangledown^2 l_f(s_k^v, \widetilde{\boldsymbol{\theta}}^t)]^{-1}[\frac{1}{Z}\sum_{v}\sum_{k=1}^{|\mathcal{S}^v|}\widetilde{o}_k^{t,v} \bigtriangledown l_f(s_k^v, \widetilde{\boldsymbol{\theta}}^t) - \frac{1}{Z}\sum_{v}\sum_{k=1}^{|\mathcal{S}^v|}(\widetilde{o}_k^{t,v} - o_k^v) \bigtriangledown l_f(s_k^v, \widetilde{\boldsymbol{\theta}}^t)] \\
&= -[\frac{1}{Z}\sum_{v}\sum_{k=1}^{|\mathcal{S}^v|}\widetilde{o}_k^{t,v} \bigtriangledown^2 l_f(s_k^v, \widetilde{\boldsymbol{\theta}}^t)]^{-1}[\frac{1}{Z}\sum_{v}\sum_{k=1}^{|\mathcal{S}^v|}o_k^v \bigtriangledown l_f(s_k^v, \widetilde{\boldsymbol{\theta}}^t)]
\end{aligned}
\tag{16}
$$

Thus, $\frac{d\hat{\boldsymbol{\theta}}(\boldsymbol{o},\epsilon)}{d\epsilon}|_{\epsilon\to 0} \approx \frac{\triangle_{-\frac{1}{Z}}}{-\frac{1}{Z}} = [\frac{1}{Z}\sum_v\sum_{k=1}^{|\mathcal{S}^v|}\widetilde{o}_k^{t,v} \bigtriangledown^2 l_f(s_k^v, \widetilde{\boldsymbol{\theta}}^t)]^{-1}[\sum_v\sum_{k=1}^{|\mathcal{S}^v|}o_k^v \bigtriangledown l_f(s_k^v, \widetilde{\boldsymbol{\theta}}^t)]$, and we have

$$
\begin{aligned}
\frac{L_f(\mathcal{T}^u, \hat{\boldsymbol{\theta}}(\boldsymbol{o})) - L_f(\mathcal{T}^u, \widetilde{\boldsymbol{\theta}}^t)}{-\frac{1}{Z}} &= \frac{L_f(\mathcal{T}^u, \hat{\boldsymbol{\theta}}(\boldsymbol{o}, -\frac{1}{Z})) - L_f(\mathcal{T}^u, \widetilde{\boldsymbol{\theta}}^t)}{-\frac{1}{Z}} \approx \frac{dL_f(\mathcal{T}^u, \hat{\boldsymbol{\theta}}(\boldsymbol{o},\epsilon))}{d\epsilon}|_{\epsilon\to 0} \\
&= \sum_{y\in\mathcal{T}^u} \bigtriangledown l_f(y, \widetilde{\boldsymbol{\theta}}^t) \times \frac{d\hat{\boldsymbol{\theta}}(\boldsymbol{o},\epsilon)}{d\epsilon}|_{\epsilon\to 0} \\
&\approx \sum_{y\in\mathcal{T}^u} \bigtriangledown l_f(y, \widetilde{\boldsymbol{\theta}}^t) H_{\widetilde{\boldsymbol{\theta}}^t}^{-1}[\sum_{v}\sum_{k=1}^{|\mathcal{S}^v|}o_k^v \bigtriangledown l_f(s_k^v, \widetilde{\boldsymbol{\theta}}^t)]
\end{aligned}
\tag{17}
$$

where $H_{\widetilde{\boldsymbol{\theta}}^t} = \frac{1}{Z}\sum_v \sum_{k=1}^{|\mathcal{S}^v|} \widetilde{o}_k^{t,v} \bigtriangledown^2 l_f(s_k^v, \widetilde{\boldsymbol{\theta}}^t)$. At last,

$$
\begin{aligned}
L_f(\mathcal{T}^u, \hat{\boldsymbol{\theta}}(\boldsymbol{o})) - L_f(\mathcal{T}^u, \widetilde{\boldsymbol{\theta}}^t) &\approx -\frac{1}{Z}\sum_{y\in\mathcal{T}^u} \bigtriangledown l_f(y, \widetilde{\boldsymbol{\theta}}^t) H_{\widetilde{\boldsymbol{\theta}}^t}^{-1}[\sum_v \sum_{k=1}^{|\mathcal{S}^v|} o_k^v \bigtriangledown l_f(s_k^v, \widetilde{\boldsymbol{\theta}}^t)] \\
L_f(\mathcal{T}^u, \hat{\boldsymbol{\theta}}(\boldsymbol{o})) &\approx L_f(\mathcal{T}^u, \widetilde{\boldsymbol{\theta}}^t) - \frac{1}{Z}\sum_{y\in\mathcal{T}^u}\sum_v \sum_{k=1}^{|\mathcal{S}^v|} o_k^v \bigtriangledown l_f(y, \widetilde{\boldsymbol{\theta}}^t) H_{\widetilde{\boldsymbol{\theta}}^t}^{-1} \bigtriangledown l_f(s_k^v, \widetilde{\boldsymbol{\theta}}^t)
\end{aligned}
\tag{18}
$$

□

D Proof of Theorem 4

Proof.

$$
\begin{aligned}
& z_u(\boldsymbol{\alpha}^u, \boldsymbol{\alpha}^{-u}) = E_{\boldsymbol{o}}[\overline{z}_u(\boldsymbol{o}^u, \boldsymbol{o}^{-u})] = E_{\boldsymbol{o}}[\sum_{t=1}^{T} \mathbf{1}(\boldsymbol{o}\in A_t)\overline{z}_u(\boldsymbol{o}^u, \boldsymbol{o}^{-u}, t)] \\
=& \sum_{t=1}^{T}\sum_{\boldsymbol{o}} \mathbf{1}(\boldsymbol{o}\in A_t)\alpha^u(\boldsymbol{o}^u)\boldsymbol{\alpha}^{-u}(\boldsymbol{o}^{-u})\{-L_f(\mathcal{T}^u, \widetilde{\boldsymbol{\theta}}^t) + \frac{1}{Z}\sum_{y\in\mathcal{T}^u}\sum_{v\in\mathcal{U}}\sum_{k=1}^{|\mathcal{S}^v|} o_k^v g(s_k^v, y, t) - \lambda \sum_{k=1}^{|\mathcal{S}^u|} o_k^u \beta_k^u\} \\
=& \sum_{t=1}^{T}\sum_{\boldsymbol{o}} \mathbf{1}(\boldsymbol{o}\in A_t)\alpha^u(\boldsymbol{o}^u)\boldsymbol{\alpha}^{-u}(\boldsymbol{o}^{-u})\{-L_f(\mathcal{T}^u, \widetilde{\boldsymbol{\theta}}^t) + \frac{1}{Z}\sum_{y\in\mathcal{T}^u}\sum_{v\in\mathcal{U}}(\boldsymbol{o}^v)^T \boldsymbol{g}_y^{t,v} - \lambda(\boldsymbol{o}^u)^T\boldsymbol{\beta}^u\} \\
=& \sum_{t=1}^{T}\sum_{\boldsymbol{o}} \mathbf{1}(\boldsymbol{o}\in A_t)\alpha^u(\boldsymbol{o}^u)\boldsymbol{\alpha}^{-u}(\boldsymbol{o}^{-u})\{-L_f(\mathcal{T}^u, \widetilde{\boldsymbol{\theta}}^t) + \frac{1}{Z}\sum_{v\in\mathcal{U}}(\boldsymbol{o}^v)^T \boldsymbol{g}^{t,v} - \lambda(\boldsymbol{o}^u)^T\boldsymbol{\beta}^u\} \\
=& \sum_{t=1}^{T}\sum_{\boldsymbol{o}} \mathbf{1}(\boldsymbol{o}\in A_t)\alpha^u(\boldsymbol{o}^u)\boldsymbol{\alpha}^{-u}(\boldsymbol{o}^{-u})\{-L_f(\mathcal{T}^u, \widetilde{\boldsymbol{\theta}}^t) + \frac{1}{Z}\sum_{v\neq u}(\boldsymbol{o}^v)^T \boldsymbol{g}^{t,v} + \frac{1}{Z}(\boldsymbol{o}^u)^T\boldsymbol{g}^{t,u} - \lambda(\boldsymbol{o}^u)^T\boldsymbol{\beta}^u\}
\end{aligned}
\tag{19}
$$

□

E Proof of Theorem 5

We rewrite the objective as follows:

$$
\max_{\boldsymbol{\alpha}^u\in\triangle} \sum_{\boldsymbol{o}^u} \alpha_{\boldsymbol{o}^u}^u [\sum_{\boldsymbol{o}^{-u}} \alpha_{\boldsymbol{o}^{-u}}^{-u} \sum_{t=1}^{T} \mathbf{1}(\boldsymbol{o}\in A_t) B(\boldsymbol{o}^u, \boldsymbol{o}^{-u}, t)] \tag{20}
$$

Suppose the optimal solution for (20) is $\boldsymbol{\alpha}^u$, and the output of the lth iteration is $\boldsymbol{\alpha}_l^u$. Recall that $\boldsymbol{g} = \sum_{\boldsymbol{o}^{-u}} \alpha_{\boldsymbol{o}^{-u}}^{-u} \sum_{t=1}^{T} \mathbf{1}(\boldsymbol{o}\in A_t) B(\boldsymbol{o}^u, \boldsymbol{o}^{-u}, t)$, then we have:

$$
\begin{aligned}
E[\boldsymbol{\alpha}^u\boldsymbol{g} - \boldsymbol{\alpha}_l^u\boldsymbol{g}] &= E[(\boldsymbol{\alpha}^u - \boldsymbol{\alpha}_l^u)\hat{\boldsymbol{g}}] \\
&= E[\frac{1}{2\gamma}(||\boldsymbol{\alpha}^u - \boldsymbol{\alpha}_l^u||_2^2 + \gamma^2||\hat{\boldsymbol{g}}||_2^2 - ||\boldsymbol{\alpha}^u - (\boldsymbol{\alpha}_l^u + \gamma\hat{\boldsymbol{g}}||_2^2)] \\
&\leq E[\frac{1}{2\gamma}(||\boldsymbol{\alpha}^u - \boldsymbol{\alpha}_l^u||_2^2 + \gamma^2||\hat{\boldsymbol{g}}||_2^2 - ||\boldsymbol{\alpha}^u - \boldsymbol{\alpha}_{l+1}^u||_2^2)] \\
&\leq E[\frac{1}{2\gamma}(||\boldsymbol{\alpha}^u - \boldsymbol{\alpha}_l^u||_2^2 - ||\boldsymbol{\alpha}^u - \boldsymbol{\alpha}_{l+1}^u||_2^2 + \gamma^2 G^2)]
\end{aligned}
\tag{21}
$$

where the third line hold because $\boldsymbol{\alpha}_{l+1}^u = \Pi_\triangle[\boldsymbol{\alpha}_l^u + \gamma\hat{\boldsymbol{g}}(\alpha^u)]$.

In the next,

$$\begin{aligned}\sum_{l=1}^{L} E[\boldsymbol{\alpha}^u \boldsymbol{y} - \boldsymbol{\alpha}_l^u \boldsymbol{y}] &\leq E[\frac{1}{2\gamma}(||\boldsymbol{\alpha}^u - \boldsymbol{\alpha}_1^u||_2^2 - ||\boldsymbol{\alpha}^u - \boldsymbol{\alpha}_{L+1}^u||_2^2 + L\gamma^2 G^2)] \\ &\leq E[\frac{1}{2\gamma}(||\boldsymbol{\alpha}^u - \boldsymbol{\alpha}_1^u||_2^2 + L\gamma^2 G^2)] \\ &\leq \frac{1}{\gamma} + L\gamma^2 G^2\end{aligned} \tag{22}$$

where the third line hold because $\boldsymbol{\alpha}^u$ and $\boldsymbol{\alpha}_1^u$ are both simplex.

At last,

$$\begin{aligned}\frac{1}{L}\sum_{l=1}^{L} E[\boldsymbol{\alpha}^u \boldsymbol{g} - \boldsymbol{\alpha}_l^u \boldsymbol{g}] &\leq \frac{1}{L\gamma} + \gamma^2 G^2 \\ E[\boldsymbol{\alpha}^u \boldsymbol{g}] - \frac{1}{L}\sum_{l=1}^{L} E[\boldsymbol{\alpha}_l^u \boldsymbol{g}] &\leq \frac{1}{L\gamma} + \gamma^2 G^2 \\ \frac{1}{L}\sum_{l=1}^{L} E[\boldsymbol{\alpha}_l^u \boldsymbol{g}] &\geq E[\boldsymbol{\alpha}^u \boldsymbol{g}] - (\frac{1}{L\gamma} + \gamma^2 G^2) \\ \frac{1}{L}\sum_{l=1}^{L} E[\boldsymbol{\alpha}_l^u \boldsymbol{g}] &\geq \max_{\boldsymbol{\alpha}^u} E[\boldsymbol{\alpha}^u \boldsymbol{g}] - (\frac{1}{L\gamma} + \gamma^2 G^2) \\ E[z_u(\hat{\boldsymbol{\alpha}}^u, \boldsymbol{\alpha}^{-u})] = E[\hat{\boldsymbol{\alpha}}^u \boldsymbol{g}] &\geq \max_{\boldsymbol{\alpha}^u} E[\boldsymbol{\alpha}^u \boldsymbol{g}] - (\frac{1}{L\gamma} + \gamma^2 G^2) \\ &= \max_{\boldsymbol{\alpha}^u} E[z_u(\boldsymbol{\alpha}^u, \boldsymbol{\alpha}^{-u})] - (\frac{1}{L\gamma} + \gamma^2 G^2)\end{aligned} \tag{23}$$

F Proof of Theorem 6

In equation (4), the approximation error comes from ignoring the term $\sum_v \sum_{k=1}^{|\mathcal{S}^v|} (\widetilde{o}_k^v - o_k^v) \bigtriangledown^2 l_f(s_k^v, \widetilde{\boldsymbol{\theta}})$, where $\widetilde{o}^v = \{\widetilde{o}_1^v, ...\widetilde{o}_{|\mathcal{S}^v|}^v\}$ is the anchor selection vector of user v. Let D be the hamming distance counting the number of different bits between two vectors, then we consider the value $\sum_v \sum_{k=1}^{|\mathcal{S}^v|} |\widetilde{o}_k^v - o_k^v| B = \sum_v D(\widetilde{\boldsymbol{o}}^v, \boldsymbol{o}^v) B$, which upper bounds the approximation error, and are interested in whether the multi-anchor proposal can reduce this value.

In specific, for a given selection vector $\boldsymbol{o}$, suppose $\widetilde{\boldsymbol{o}}^{t_1}$ is the nearest anchor vector to $\boldsymbol{o}$ in $\boldsymbol{P}$. Since $\boldsymbol{P} \subseteq \boldsymbol{Q}$, we have $\widetilde{\boldsymbol{o}}^{t_1} \in \boldsymbol{Q}$. Suppose $\widetilde{\boldsymbol{o}}^{t_2}$ is the nearest anchor vector to $\boldsymbol{o}$ in $\boldsymbol{Q}$, then according the definition of $\boldsymbol{Q}$, we have $\sum_{v=1}^{N} D(\widetilde{\boldsymbol{o}}^{t_2,v}, \boldsymbol{o}^v) \leq \sum_{v=1}^{N} D(\widetilde{\boldsymbol{o}}^{t_1,v}, \boldsymbol{o}^v)$. Since $B > 0$, we have $\sum_{v=1}^{N} D(\widetilde{\boldsymbol{o}}^{t_2,v}, \boldsymbol{o}^v) B \leq \sum_{v=1}^{N} D(\widetilde{\boldsymbol{o}}^{t_1,v}, \boldsymbol{o}^v) B$. By summing all the candidate selection vectors, and grouping them according to A_t^P and A_t^Q, respectively, we have:

$$\sum_{t=1}^{T_P} \sum_{\boldsymbol{o} \in A_t^P} [\sum_v D(\widetilde{\boldsymbol{o}}^{t,v}, \boldsymbol{o}^v)] B \geq \sum_{t=1}^{T_Q} \sum_{\boldsymbol{o} \in A_t^Q} [\sum_v D(\widetilde{\boldsymbol{o}}^{t,v}, \boldsymbol{o}^v)] B, \tag{24}$$

G More Implementation Details

For the simulation dataset, the threshold η and (a_1, a_2, a_3) are initially set as 0.5 and $(0.5, 1, 1)$, respectively. And then, we tune them in the experiments to study the influence of different dataset sparsities and user willingness characters. For the real world datasets, **Diginetica** and **Amazon Video** are e-commerce datasets, where we are provided with the user-item purchasing records. **Steam** is a game dataset, which includes the interactions (*e.g.*, reviewing behaviors) between the users and games. To evaluate our model efficiently, we remove the users who have interacted with more than

Table 1: Statistics of the datasets

Dataset	# User	# Item	# Interaction	Sparsity
Simulation	1000	1000	6148	99.39%
Diginetica	2852	10739	17073	99.94%
Steam	11942	6955	86595	99.89%
Amazon Video	2790	12435	18703	99.95%

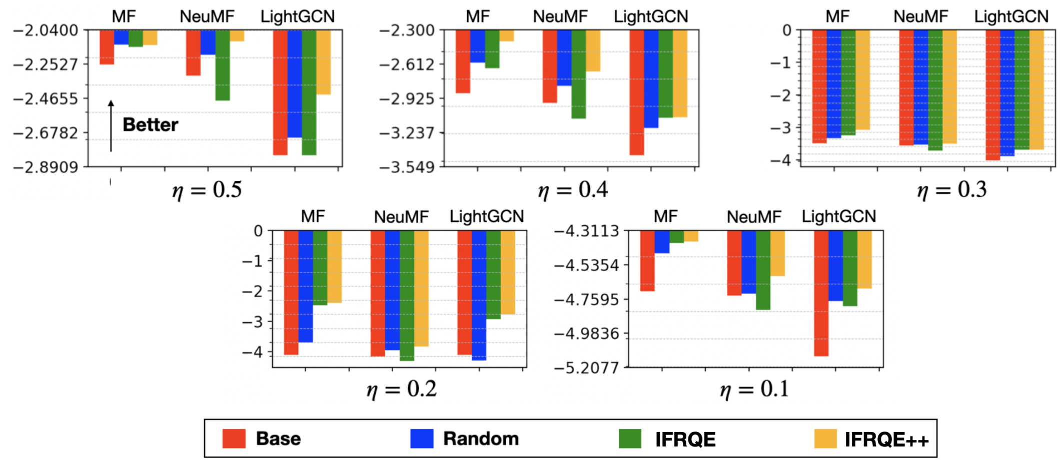

Figure 1: Performance comparison on the dataset with different sparsities.

10 items. Since we do not know the real user disclosing willingness, we simulate it by randomly assigning the willingness vector for each user, and repeat the experiments for ten times to make sure that the experiment results are not from the randomness. The statistics of the above datasets are concluded in Table 1.

In IFRQE++, considering that the space of $\boldsymbol{\alpha}$ can be extremely large, it is less efficient to initialize $\boldsymbol{\alpha}$ completely at random, and blindly learn it in the optimization process. To solve this problem, we initialize $\boldsymbol{\alpha}$ with a prior, assuming that most of the items should be leveraged to train the model for achieving acceptable recommendation performance. In specific, for each $\boldsymbol{\alpha}^u \in \boldsymbol{\alpha}$, we initialize it with a Binomial distribution $p(k,n) = C_n^k s^k (1-s)^{n-k}$, where k is the number of disclosed items, and n is the total number of items in the training set (*i.e.*, $|\mathcal{S}^u|$). Notably, we do not discriminate the item differences in the initialization process of $\boldsymbol{\alpha}^u$. For example, suppose there are three items, then $\boldsymbol{\alpha}^u_{\{1,1,0\}} = \boldsymbol{\alpha}^u_{\{1,0,1\}} = \boldsymbol{\alpha}^u_{\{0,1,1\}}$. In the experiment, we set $s = 0.9$, which means, in the beginning, about $0.9 * |\mathcal{S}^u|$ items will be involved into the model training process.

In order to efficiently compute the inverse of the Hessian matrix, we use the stochastic estimation method discussed in [2]. In specific, according to the Taylor expansion, we can express H^{-1} by $\sum_{i=0}^{\infty}(I-H)^i$. Let $H_j^{-1} = \sum_{i=0}^{j}(I-H)^i$, then we have $H_j^{-1} = I + (I-H)H_{j-1}^{-1}$. To compute $H_{\widetilde{\boldsymbol{\theta}}}^{-1} \bigtriangledown l_f(s_k^v, \widetilde{\boldsymbol{\theta}})$, we uniformly sample a training data s, and approximate H by $\bigtriangledown^2 l_f(s, \widetilde{\boldsymbol{\theta}})$. Then we have the following recursive equation:

$$H_j^{-1} \bigtriangledown l_f(s_k^v, \widetilde{\boldsymbol{\theta}}) = \bigtriangledown l_f(s_k^v, \widetilde{\boldsymbol{\theta}}) + (I - \bigtriangledown^2 l_f(s, \widetilde{\boldsymbol{\theta}}))H_{j-1}^{-1} \bigtriangledown l_f(s_k^v, \widetilde{\boldsymbol{\theta}}) \tag{25}$$

Obviously, when $j \to \infty$, we have $H_j^{-1} \bigtriangledown l_f(s_k^v, \widetilde{\boldsymbol{\theta}}) \to H_{\widetilde{\boldsymbol{\theta}}}^{-1} \bigtriangledown l_f(s_k^v, \widetilde{\boldsymbol{\theta}})$. In the experiment, we resample s for each iteration, and the total number of iterations N_J is tuned to better effectiveness-efficiency trade-off.

For the model parameters, we determine them by grid search. For example, the number of anchor selection vectors is searched in $[1, 2, 3, 4, 5, 6, 7, 8, 9, 10]$. The learning rate and batch size are determined in the ranges of $[0.001, 0.01, 0.05]$ and $[1024, 2048, 4096]$, respectively. The anchor selection vectors are sampled from the Binomial distribution, where, similar to $\boldsymbol{\alpha}^u$, we set the mean as 0.9.

Table 2: Statistics of the simulation datasets with different η's, where the number of users and items are both 1000.

η	0.1	0.2	0.3	0.4	0.5
#Interaction	11296	10068	8828	7507	6184
Sparsity	98.87%	99.00%	99.12%	99.25%	99.39%

Table 3: Parameter settings in the experiments.

Parameter	Tuning range	Simulation	Diginetica	Steam	Amazon Video
Learning rate	$[0.001, 0.01, 0.05]$	0.01	0.01	0.01	0.01
Batch size	$[1024, 2048, 4096]$	2048	2048	2048	2048
Embedding size	$[64, 128, 256]$	64	64	64	64
Drop ratio	$[0.01, 0.1, 0.2]$	0.1	0.1	0.1	0.1
λ	$[0.1, 0.5, 1]$	1	1	1	1
Iteration number M	$[1, 3, 5, 10]$	10	10	10	10
Training epochs	$[50, 100, 150]$	50	50	150	100
L	$[500, 1000, 2000]$	1000	1000	500	1000
T	$[1, 2, 3, 4, 5, 6, 7, 8, 9, 10]$	2	8	4	6
N_J for computing H^{-1}	$[10, 20, 30]$	30	10	20	20

The final parameters used in our experiments are concluded in Table 3. Our project has been released at https://ifqre.github.io/IFRQE/.

H More Experiments

In this section, we present more experiments to evaluate and analyze our proposed models.

H.1 Influence of the Data Sparsity

In real-world scenarios, recommender systems can be applied in different applications, where the dataset sparsity may vary a lot. In this section, we would like to study whether our methods are consistently competitive for the datasets with different sparsities. In order to flexibly control the sparsity, we conduct this experiment based on the simulation dataset. Since the threshold η controls the hardness of generating the user-item interactions, we tune η in the range of $\{0.1, 0.2, 0.3, 0.4, 0.5\}$ to build the datasets with different densities, where larger η can lead more sparse dataset. The statistics of the generated datasets are presented in Table 2. In Figure 1, we report the performance of different models based on the reward, where we can see: the performance of the base model is not satisfied in most cases. IFRQE usually outperforms the random method, although there are a few exceptions. IFRQE++ can always achieve the best performance, which is consistent on all the base models and datasets with different sparsities. These results demonstrate the robustness of our model, and suggests that it can be potentially applied to a wide range of real-world applications.

H.2 Influence of the balancing parameter λ

In the reward function, λ balances the importance of the recommendation quality and user disclosing willingness. To study whether our model can adaptively trade-off the above two aspects, we specify λ with different values, and observe whether our model can always achieve better performance than the baselines. In specific, we set λ as 0.1, 0.5, 1.0 and 2.0 respectively, and the results of comparing our models with the baselines are presented in Table 4. We can see: on different datasets, because the base model completely ignores the user disclosing willingness, the overall reward is the worst comparing with the other methods. Blindly integrating the user disclosing willingness is also suboptimal, which is evidenced by the lower performance of the random method. By designing a principled model to optimize the overall reward, IFRQE can achieve better performance than the base and random models in most cases. As expected, by leveraging more anchor selection

Table 4: Comparison between different models with different λ's. We use bold fonts to label the best performance for each dataset, evaluation metric and base model. "()" indicates the standard error.

Dataset	Simulation	Diginetica	Steam	Amazon Video
$\lambda = 0.1$				
MF	-0.34$_{(.002)}$	-0.29$_{(.009)}$	-0.49$_{(.005)}$	-0.32$_{(.007)}$
w/ Random	-0.46$_{(.005)}$	-0.28$_{(.017)}$	-0.48$_{(.019)}$	-0.31$_{(.011)}$
w/ Threshold	-0.48$_{(.015)}$	-0.48$_{(.022)}$	-0.50$_{(.017)}$	-0.30$_{(.018)}$
w/ IFRQE	-0.34$_{(.007)}$	-0.23$_{(.011)}$	-0.49$_{(.013)}$	-0.30$_{(.008)}$
w/ IFRQE++	**-0.32**$_{(.006)}$	**-0.22**$_{(.005)}$	**-0.47**$_{(.006)}$	**-0.28**$_{(.002)}$
NeuMF	-0.40$_{(.001)}$	-0.30$_{(.002)}$	-0.37$_{(.009)}$	-0.33$_{(.007)}$
w/ Random	-0.39$_{(.012)}$	-0.31$_{(.010)}$	-0.36$_{(.015)}$	-0.34$_{(.006)}$
w/ Threshold	-0.38$_{(.019)}$	-0.38$_{(.012)}$	-0.35$_{(.007)}$	-0.31$_{(.014)}$
w/ IFRQE	-0.37$_{(.005)}$	-0.34$_{(.002)}$	-0.25$_{(.014)}$	-0.37$_{(.012)}$
w/ IFRQE++	**-0.34**$_{(.007)}$	**-0.28**$_{(.003)}$	**-0.24**$_{(.005)}$	**-0.32**$_{(.002)}$
LightGCN	-0.32$_{(.011)}$	-0.28$_{(.016)}$	-0.36$_{(.013)}$	-0.49$_{(.011)}$
w/ Random	-0.30$_{(.013)}$	-0.27$_{(.022)}$	-0.39$_{(.019)}$	-0.47$_{(.016)}$
w/ Threshold	-0.36$_{(.021)}$	-0.33$_{(.030)}$	-0.35$_{(.026)}$	-0.34$_{(.011)}$
w/ IFRQE	-0.26$_{(.005)}$	-0.27$_{(.014)}$	-0.34$_{(.011)}$	-0.25$_{(.003)}$
w/ IFRQE++	**-0.25**$_{(.006)}$	**-0.25**$_{(.016)}$	**-0.33**$_{(.008)}$	**-0.24**$_{(.004)}$
$\lambda = 0.5$				
MF	-1.19$_{(.006)}$	-1.10$_{(.012)}$	-1.54$_{(.011)}$	-1.27$_{(.016)}$
w/ Random	-1.28$_{(.016)}$	-1.08$_{(.007)}$	-1.52$_{(.019)}$	-1.21$_{(.018)}$
w/ Threshold	-1.32$_{(.021)}$	-1.44$_{(.030)}$	-1.373$_{(.026)}$	-1.214$_{(.011)}$
w/ IFRQE	-0.97$_{(.003)}$	-0.85$_{(.007)}$	-0.77$_{(.006)}$	-1.04$_{(.008)}$
w/ IFRQE++	**-0.95**$_{(.001)}$	**-0.81**$_{(.006)}$	**-0.76**$_{(.001)}$	**-1.03**$_{(.004)}$
NeuMF	-1.11$_{(.009)}$	-1.11$_{(.002)}$	-1.42$_{(.006)}$	-1.27$_{(.009)}$
w/ Random	-1.22$_{(.006)}$	-1.10$_{(.010)}$	-1.40$_{(.013)}$	-1.22$_{(.017)}$
w/ Threshold	-1.32$_{(.021)}$	-1.44$_{(.030)}$	-$_{(.026)}$	-1.23$_{(.011)}$
w/ IFRQE	-0.93$_{(.011)}$	-1.01$_{(.002)}$	-1.30$_{(.005)}$	-1.26$_{(.004)}$
w/ IFRQE++	**-0.91**$_{(.006)}$	**-0.98**$_{(.003)}$	**-1.28**$_{(.011)}$	**-1.20**$_{(.007)}$
LightGCN	-1.17$_{(.011)}$	-1.09$_{(.016)}$	-1.41$_{(.009)}$	-1.48$_{(.006)}$
w/ Random	-1.11$_{(.016)}$	-1.06$_{(.022)}$	-1.43$_{(.012)}$	-1.38$_{(.019)}$
w/ Threshold	-1.32$_{(.021)}$	-1.44$_{(.030)}$	-1.74$_{(.026)}$	-1.75$_{(.011)}$
w/ IFRQE	-1.04$_{(.007)}$	-0.97$_{(.014)}$	-1.41$_{(.003)}$	-1.33$_{(.006)}$
w/ IFRQE++	**-1.02**$_{(.001)}$	**-0.96**$_{(.016)}$	**-1.40**$_{(.008)}$	**-1.29**$_{(.005)}$
$\lambda = 1.0$				
MF	-2.25$_{(.032)}$	-2.65$_{(.072)}$	-2.99$_{(.093)}$	-2.43$_{(.022)}$
w/ Random	-2.13$_{(.014)}$	-2.04$_{(.006)}$	-2.82$_{(.015)}$	-2.30$_{(.023)}$
w/ Threshold	-2.20$_{(.011)}$	-2.06$_{(.013)}$	-2.57$_{(.023)}$	-2.20$_{(.015)}$
w/ IFRQE	-2.09$_{(.017)}$	-2.18$_{(.006)}$	-2.43$_{(.015)}$	-2.10$_{(.005)}$
w/ IFRQE++	**-1.92**$_{(.012)}$	**-1.92**$_{(.022)}$	**-2.42**$_{(.014)}$	**-1.98**$_{(.056)}$
NeuMF	-2.32$_{(.011)}$	-2.11$_{(.003)}$	-2.76$_{(.007)}$	-2.47$_{(.010)}$
w/ Random	-2.19$_{(.011)}$	-2.11$_{(.016)}$	-2.59$_{(.011)}$	-2.41$_{(.010)}$
w/ Threshold	-2.18$_{(.012)}$	-2.08$_{(.013)}$	-2.86$_{(.011)}$	-2.21$_{(.025)}$
w/ IFRQE	-2.48$_{(.013)}$	-2.17$_{(.019)}$	-2.45$_{(.027)}$	-2.24$_{(.012)}$
w/ IFRQE++	**-2.11**$_{(.011)}$	**-2.08**$_{(.023)}$	**-2.37**$_{(.027)}$	**-2.17**$_{(.011)}$
LightGCN	-2.82$_{(.012)}$	-2.71$_{(.009)}$	-3.13$_{(.014)}$	-3.05$_{(.007)}$
w/ Random	-2.71$_{(.022)}$	-2.64$_{(.007)}$	-3.07$_{(.019)}$	-2.93$_{(.015)}$
w/ Threshold	-2.70$_{(.008)}$	-2.60$_{(.031)}$	-2.82$_{(.022)}$	-2.30$_{(.021)}$
w/ IFRQE	-2.82$_{(.014)}$	-2.65$_{(.013)}$	-2.82$_{(.008)}$	-2.13$_{(.013)}$
w/ IFRQE++	**-2.44**$_{(.016)}$	**-2.55**$_{(.009)}$	**-2.80**$_{(.005)}$	**-2.03**$_{(.023)}$
$\lambda = 2.0$				
MF	-4.53$_{(.011)}$	-4.23$_{(.052)}$	-5.46$_{(.013)}$	-4.56$_{(.012)}$
w/ Random	-4.30$_{(.024)}$	-4.01$_{(.016)}$	-5.18$_{(.025)}$	-4.64$_{(.023)}$
w/ Threshold	-4.34$_{(.009)}$	-3.72$_{(.027)}$	-4.92$_{(.023)}$	-4.32$_{(.025)}$
w/ IFRQE	-4.12$_{(.017)}$	-3.24$_{(.026)}$	-4.05$_{(.015)}$	-3.71$_{(.043)}$
w/ IFRQE++	**-4.06**$_{(.010)}$	**-3.16**$_{(.032)}$	**-4.02**$_{(.014)}$	**-3.37**$_{(.056)}$
NeuMF	-4.44$_{(.012)}$	-4.12$_{(.043)}$	-5.34$_{(.027)}$	-4.81$_{(.019)}$
w/ Random	-4.21$_{(.012)}$	-4.05$_{(.035)}$	-5.07$_{(.021)}$	-4.58$_{(.031)}$
w/ Threshold	-4.38$_{(.012)}$	-3.74$_{(.013)}$	-4.81$_{(.011)}$	-4.33$_{(.035)}$
w/ IFRQE	-4.56$_{(.013)}$	-4.00$_{(.019)}$	-3.90$_{(.028)}$	-3.06$_{(.022)}$
w/ IFRQE++	**-4.00**$_{(.012)}$	**-3.67**$_{(.021)}$	**-3.40**$_{(.019)}$	**-2.74**$_{(.012)}$
LightGCN	-4.94$_{(.021)}$	-4.10$_{(.009)}$	-5.75$_{(.024)}$	-5.20$_{(.057)}$
w/ Random	-4.70$_{(.032)}$	-3.90$_{(.007)}$	-5.07$_{(.029)}$	-4.73$_{(.033)}$
w/ Threshold	-4.36$_{(.017)}$	-4.32$_{(.031)}$	-5.21$_{(.022)}$	-4.93$_{(.041)}$
w/ IFRQE	-3.89$_{(.013)}$	-2.54$_{(.013)}$	-3.90$_{(.018)}$	-3.71$_{(.023)}$
w/ IFRQE++	**-3.55**$_{(.026)}$	**-2.52**$_{(.009)}$	**-3.40**$_{(.025)}$	**-3.69**$_{(.033)}$

vectors to simulate the validation loss, the final model IFRQE++ achieves the best performance. The above observations are consistent for different λ's, which manifests that our model is robust to the predefined relative importance between the recommendation quality and user disclosing willingness.

H.3 Complete results for section 5.3 and 5.4 in the main paper

To begin with, we present the complete results of the experiments in section 5.3 of the main paper. From the results shown in Figure 2, we can see: similar to the results in the main paper, the validation loss can be in general well approximated in most cases. IFRQE++ can achieve better approximation

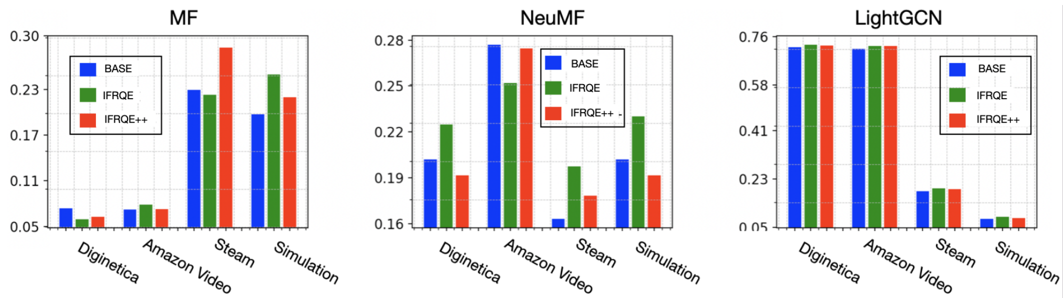

Figure 2: Approximation error on the validation loss for all the datasets and base models.

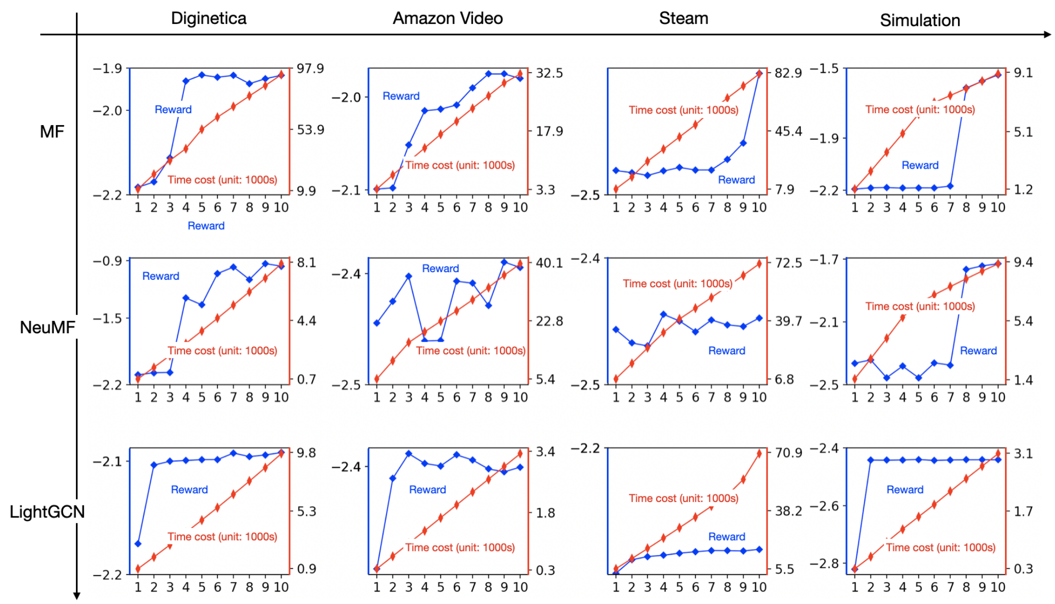

Figure 3: Influence of T for all the datasets and base models.

accuracy than IFRQE, which demonstrate the effectiveness of using more anchor vectors for computing the validation loss. In Figure 3, we show the complete results of the experiments in section 5.4. We can see: the reward changing patterns seem to be quite diverse as more anchor selection vectors are leveraged in our model. For example, in the case of MF + Diginetica, the reward has a performance jump from $T = 3$ to $T = 4$. Similar performance jumping patterns can also be observed in the settings of NeuMF + Diginetica, LightGCN + Diginetica, LightGCN + Amazon Video and the simulation dataset. However, in the case of NeuMF + Amazon Video, the reward changes irregularly as T becomes larger. While different combinations between the base model and dataset may lead to various performance change patterns, a common phenomenon is that, in most cases, the performance tends to be better as more anchor selection vectors. Simultaneously, the time cost is increased almost linearly as more anchor vectors are deployed to achieve better performance. These observations are aligned with the conclusions in the main paper.

H.4 Comparison results with more performance evaluation metrics and base models

To begin with, we augment Table 1 in the main paper by reporting the recommendation performance based on Precision, NDCG and MRR. From the results shown in Table 5, we can draw similar conclusions as Table 1, that is, there are many cases that, although we have removed some items due to the user willingness, the recommendation performances are not lowered.

To further demonstrate the generality of our framework, we additionally conduct experiments with the following base models: (1) DIN [4] is a sequential recommender model, where the user behavior

Table 5: Additional metrics for evaluating the recommendation performance. We use "P" to represent the precision. All the results are percentage values with "%" omitted.

Dataset	Simulation			Diginetica			Steam			Amazon Video		
Metric	P↑	NDCG↑	MRR↑	P↑	NDCG↑	MRR↑	P↑	NDCG↑	MRR↑	P↑	NDCG↑	MRR↑
MF	0.39$_{(.022)}$	1.02$_{(.017)}$	0.72$_{(.023)}$	2.83$_{(.012)}$	11.3$_{(.017)}$	10.3$_{(.023)}$	6.21$_{(.043)}$	19.5$_{(.014)}$	15.8$_{(.093)}$	**1.03**$_{(.012)}$	**3.52**$_{(.021)}$	**2.99**$_{(.022)}$
w/ Random	0.41$_{(.031)}$	1.11$_{(.056)}$	0.80$_{(.024)}$	2.51$_{(.031)}$	9.88$_{(.006)}$	9.00$_{(.014)}$	5.64$_{(.019)}$	17.7$_{(.017)}$	14.3$_{(.015)}$	0.92$_{(.011)}$	3.10$_{(.022)}$	2.61$_{(.023)}$
w/ Threshold	**1.02**$_{(.036)}$	**4.30**$_{(.022)}$	**4.04**$_{(.011)}$	2.80$_{(.022)}$	11.0$_{(.013)}$	9.98$_{(.042)}$	5.55$_{(.028)}$	17.7$_{(.033)}$	14.4$_{(.027)}$	0.90$_{(.026)}$	2.86$_{(.037)}$	2.33$_{(.041)}$
w/ IFRQE	0.37$_{(.025)}$	1.11$_{(.033)}$	0.87$_{(.017)}$	**3.14**$_{(.025)}$	**12.5**$_{(.033)}$	**11.5**$_{(.017)}$	**6.23**$_{(.012)}$	**19.8**$_{(.012)}$	**16.1**$_{(.015)}$	0.34$_{(.015)}$	1.28$_{(.016)}$	1.14$_{(.005)}$
w/ IFRQE++	0.33$_{(.027)}$	1.02$_{(.013)}$	0.82$_{(.012)}$	2.41$_{(.007)}$	9.46$_{(.013)}$	8.21$_{(.012)}$	5.90$_{(.017)}$	18.9$_{(.032)}$	15.4$_{(.014)}$	**1.03**$_{(.012)}$	3.37$_{(.032)}$	2.78$_{(.056)}$
NeuMF	0.52$_{(.037)}$	1.40$_{(.003)}$	1.02$_{(.011)}$	**2.53**$_{(.037)}$	**10.5**$_{(.003)}$	**9.83**$_{(.011)}$	**6.12**$_{(.007)}$	**19.5**$_{(.012)}$	**15.9**$_{(.007)}$	0.99$_{(.014)}$	3.38$_{(.013)}$	2.87$_{(.010)}$
w/ Random	0.54$_{(.019)}$	1.71$_{(.015)}$	1.39$_{(.012)}$	1.43$_{(.019)}$	5.04$_{(.015)}$	4.34$_{(.012)}$	4.64$_{(.005)}$	14.6$_{(.019)}$	11.5$_{(.011)}$	1.08$_{(.014)}$	3.63$_{(.006)}$	3.04$_{(.010)}$
w/ Threshold	0.90$_{(.022)}$	4.02$_{(.028)}$	**3.86**$_{(.011)}$	2.43$_{(.011)}$	9.74$_{(.044)}$	8.94$_{(.016)}$	4.50$_{(.029)}$	14.2$_{(.022)}$	11.5$_{(.011)}$	0.93$_{(.012)}$	3.03$_{(.028)}$	2.49$_{(.017)}$
w/ IFRQE	0.84$_{(.014)}$	2.73$_{(.023)}$	2.26$_{(.033)}$	1.75$_{(.004)}$	7.27$_{(.023)}$	2.99$_{(.033)}$	4.90$_{(.023)}$	14.6$_{(.019)}$	11.49$_{(.027)}$	**1.10**$_{(.015)}$	**3.77**$_{(.014)}$	**3.39**$_{(.012)}$
w/ IFRQE++	**1.80**$_{(.015)}$	**6.34**$_{(.003)}$	3.77$_{(.021)}$	1.80$_{(.015)}$	6.34$_{(.003)}$	3.77$_{(.021)}$	5.32$_{(.021)}$	16.7$_{(.014)}$	13.4$_{(.027)}$	0.72$_{(.031)}$	2.65$_{(.014)}$	2.48$_{(.011)}$
LightGCN	0.54$_{(.025)}$	1.41$_{(.008)}$	1.01$_{(.042)}$	**4.51**$_{(.019)}$	**20.2**$_{(.010)}$	**19.5** $_{(.009)}$	6.04$_{(.010)}$	18.8$_{(.009)}$	15.1$_{(.014)}$	**1.29**$_{(.021)}$	**4.90**$_{(.008)}$	**4.40**$_{(.007)}$
w/ Random	0.49$_{(.013)}$	1.27$_{(.013)}$	0.89$_{(.022)}$	4.50$_{(.013)}$	20.0$_{(.013)}$	19.2$_{(.022)}$	6.03$_{(.027)}$	18.8$_{(.026)}$	15.1$_{(.019)}$	1.05$_{(.003)}$	3.84$_{(.009)}$	3.37$_{(.015)}$
w/ Threshold	**1.16**$_{(.038)}$	**4.51**$_{(.018)}$	**4.09**$_{(.013)}$	2.43$_{(.027)}$	9.74$_{(.041)}$	8.94$_{(.021)}$	5.79$_{(.049)}$	18.1$_{(.027)}$	14.6$_{(.036)}$	0.97$_{(.013)}$	3.28$_{(.023)}$	2.78$_{(.048)}$
w/ IFRQE	0.58$_{(.008)}$	1.60$_{(.010)}$	1.18$_{(.014)}$	3.59$_{(.008)}$	14.9$_{(.010)}$	13.9$_{(.014)}$	**6.21**$_{(.016)}$	**19.8**$_{(.025)}$	**16.1**$_{(.008)}$	1.14$_{(.007)}$	3.98$_{(.018)}$	3.42$_{(.013)}$
w/ IFRQE++	0.45$_{(.022)}$	1.21$_{(.008)}$	0.87$_{(.016)}$	3.70$_{(.022)}$	14.9$_{(.008)}$	13.7$_{(.016)}$	6.15$_{(.018)}$	19.7$_{(.018)}$	16.0$_{(.005)}$	0.97$_{(.013)}$	3.44$_{(.009)}$	2.98$_{(.023)}$
DIN	**0.86**$_{(.022)}$	**2.76**$_{(.017)}$	**2.25**$_{(.023)}$	2.48$_{(.023)}$	8.30$_{(.005)}$	6.95$_{(.031)}$	7.01$_{(.019)}$	23.5$_{(.048)}$	19.7$_{(.005)}$	**1.38**$_{(.038)}$	**5.38**$_{(.012)}$	**4.39**$_{(.026)}$
w/ Random	0.74$_{(.031)}$	2.30$_{(.016)}$	1.84$_{(.024)}$	**2.60**$_{(.016)}$	**8.65**$_{(.042)}$	**7.22**$_{(.024)}$	6.54$_{(.030)}$	21.4$_{(.042)}$	17.7$_{(.014)}$	1.27$_{(.018)}$	3.78$_{(.023)}$	2.95$_{(.015)}$
w/ Threshold	0.82$_{(.004)}$	2.30$_{(.022)}$	1.73$_{(.029)}$	2.24$_{(.010)}$	7.45$_{(.032)}$	6.22$_{(.036)}$	5.79$_{(.036)}$	24.9$_{(.045)}$	20.8$_{(.029)}$	1.26$_{(.009)}$	3.84$_{(.036)}$	3.03$_{(.022)}$
w/ IFRQE	0.72$_{(.027)}$	2.48$_{(.013)}$	2.11$_{(.012)}$	2.35$_{(.017)}$	8.04$_{(.047)}$	6.82$_{(.022)}$	**7.93**$_{(.021)}$	**26.6**$_{(.030)}$	**22.3**$_{(.041)}$	0.74$_{(.034)}$	2.47$_{(.019)}$	2.04$_{(.043)}$
w/ IFRQE++	0.62$_{(.027)}$	1.73$_{(.013)}$	1.27$_{(.012)}$	2.10$_{(.014)}$	6.78$_{(.042)}$	5.56$_{(.030)}$	5.72$_{(.030)}$	17.5$_{(.016)}$	13.9$_{(.045)}$	1.09$_{(.021)}$	3.47$_{(.047)}$	2.83$_{(.038)}$
CDAE	**0.74**$_{(.010)}$	2.73$_{(.015)}$	2.41$_{(.031)}$	**0.89**$_{(.009)}$	2.65$_{(.016)}$	2.06$_{(.042)}$	6.91$_{(.008)}$	21.8$_{(.029)}$	17.6$_{(.014)}$	0.67$_{(.011)}$	2.01$_{(.027)}$	1.58$_{(.045)}$
w/ Random	**0.74**$_{(.013)}$	**2.87**$_{(.013)}$	**2.60**$_{(.022)}$	0.83$_{(.014)}$	2.63$_{(.042)}$	**2.13**$_{(.030)}$	**6.96**$_{(.021)}$	**22.0**$_{(.003)}$	**17.8**$_{(.041)}$	0.73$_{(.011)}$	2.27$_{(.015)}$	1.82$_{(.031)}$
w/ Threshold	0.36$_{(.021)}$	1.05$_{(.003)}$	0.81$_{(.041)}$	0.83$_{(.030)}$	2.61$_{(.026)}$	2.11$_{(.006)}$	5.34$_{(.045)}$	16.9$_{(.044)}$	13.8$_{(.010)}$	0.73$_{(.016)}$	2.27$_{(.040)}$	1.82$_{(.015)}$
w/ IFRQE	0.56$_{(.011)}$	0.87$_{(.015)}$	0.69$_{(.027)}$	**0.89**$_{(.011)}$	**2.68**$_{(.032)}$	2.11$_{(.013)}$	6.77$_{(.008)}$	21.2$_{(.025)}$	17.1$_{(.023)}$	1.08$_{(.027)}$	3.62$_{(.029)}$	**3.03**$_{(.046)}$
w/ IFRQE++	0.52$_{(.037)}$	0.91$_{(.005)}$	0.79$_{(.012)}$	0.86$_{(.007)}$	2.58$_{(.019)}$	2.03$_{(.042)}$	6.85$_{(.018)}$	21.4$_{(.046)}$	17.2$_{(.032)}$	**1.10**$_{(.027)}$	**3.63**$_{(.029)}$	3.00$_{(.046)}$

Table 6: Experiment results with more base models. We use bold fonts to label the best performance for each dataset, evaluation metric and base model. "()" indicates the standard error. The results of F_1 are percentage values with "%" omitted. For the metrics, ↑ means the larger the better, while ↓ means the lower the better. The performance improvements of our model against the baselines are significant under paired t-test.

Dataset	Simulation			Diginetica			Steam			Amazon Video		
Metric	F_1 ↑	wv↓	reward↑	F_1 ↑	wv↓	reward↑	F_1 ↑	wv↓	reward↑	F_1 ↑	wv↓	reward↑
DIN	**1.43**$_{(.047)}$	2.13$_{(.001)}$	-2.56$_{(.044)}$	4.13$_{(.023)}$	2.01$_{(.005)}$	-2.22$_{(.031)}$	11.6$_{(.019)}$	2.62$_{(.048)}$	-2.91$_{(.005)}$	**2.37**$_{(.038)}$	2.36$_{(.012)}$	-2.47$_{(.026)}$
w/ Random	1.23$_{(.024)}$	2.02$_{(.037)}$	-2.42$_{(.037)}$	**4.33**$_{(.016)}$	1.90$_{(.042)}$	-2.10$_{(.024)}$	10.8$_{(.030)}$	2.48$_{(.042)}$	-2.78$_{(.014)}$	2.11$_{(.008)}$	2.24$_{(.023)}$	-2.45$_{(.015)}$
w/ Threshold	1.37$_{(.004)}$	2.01$_{(.022)}$	-2.13$_{(.029)}$	4.32$_{(.010)}$	1.63$_{(.032)}$	-1.70$_{(.036)}$	10.0$_{(.036)}$	2.35$_{(.045)}$	-2.86$_{(.029)}$	2.10$_{(.009)}$	2.12$_{(.006)}$	-2.24$_{(.022)}$
w/ IFRQE	1.36$_{(.030)}$	1.10$_{(.007)}$	-1.36$_{(.016)}$	3.92$_{(.019)}$	1.46$_{(.047)}$	-1.62$_{(.012)}$	**13.2**$_{(.021)}$	2.57$_{(.030)}$	-2.89$_{(.041)}$	**2.37**$_{(.007)}$	2.31$_{(.019)}$	-2.47$_{(.043)}$
w/ IFRQE++	1.03$_{(.016)}$	**1.09**$_{(.020)}$	**-1.33**$_{(.045)}$	3.50$_{(.014)}$	**1.24**$_{(.042)}$	**-1.38**$_{(.031)}$	9.53$_{(.030)}$	**2.02**$_{(.016)}$	**-2.17**$_{(.045)}$	1.82$_{(.021)}$	**1.46**$_{(.047)}$	**-1.66**$_{(.038)}$
CDAE	**1.23**$_{(.010)}$	2.13$_{(.015)}$	-2.30$_{(.031)}$	**1.48**$_{(.009)}$	2.01$_{(.036)}$	-2.08$_{(.012)}$	11.5$_{(.008)}$	2.62$_{(.029)}$	-2.68$_{(.014)}$	1.11$_{(.011)}$	2.36$_{(.027)}$	-2.44$_{(.045)}$
w/ Random	1.22$_{(.013)}$	1.90$_{(.013)}$	-2.71$_{(.022)}$	1.38$_{(.044)}$	1.99$_{(.022)}$	-2.06$_{(.031)}$	**11.6**$_{(.021)}$	2.49$_{(.003)}$	-2.55$_{(.041)}$	1.21$_{(.001)}$	2.24$_{(.015)}$	-2.31$_{(.031)}$
w/ Threshold	0.60$_{(.021)}$	2.01$_{(.003)}$	-2.08$_{(.041)}$	1.38$_{(.030)}$	1.81$_{(.026)}$	-1.88$_{(.016)}$	9.01$_{(.045)}$	2.35$_{(.044)}$	-2.54$_{(.010)}$	1.22$_{(.016)}$	2.12$_{(.040)}$	-2.19$_{(.035)}$
w/ IFRQE	0.92$_{(.011)}$	1.65$_{(.015)}$	-2.07$_{(.027)}$	**1.48**$_{(.011)}$	1.61$_{(.032)}$	-1.68$_{(.013)}$	11.3$_{(.008)}$	2.23$_{(.025)}$	-2.30$_{(.023)}$	1.80$_{(.027)}$	1.64$_{(.029)}$	-1.73$_{(.046)}$
w/ IFRQE++	0.86$_{(.037)}$	**1.45**$_{(.005)}$	**-1.88**$_{(.012)}$	1.43$_{(.007)}$	**1.45**$_{(.019)}$	**-1.52**$_{(.042)}$	11.4$_{(.018)}$	**1.64**$_{(.046)}$	**-1.70**$_{(.032)}$	**1.83**$_{(.049)}$	**1.52**$_{(.009)}$	**-1.61**$_{(.020)}$

importance is discriminated by the attention mechanism. (2) CDAE [3] is a recommender model based on auto-encoder. From the results shown in Table 6, we can observe similar conclusions as reported in the main paper. In general, our framework can always achieve the largest rewards across different datasets, which verifies the effectiveness of our models.

H.5 Comparison between our framework and [1]

In this section, we compare our model with method proposed in [1]. We follow the experiment settings in section 5.1 in the main paper. The comparison results are presented in Table 7. We can see our framework can achieve better performance than [1], and the improvements are consistent

Table 7: Comparison between our framework and [1]. "()" indicates the standard error. The results of metrics are percentage values with "%" omitted. For the metrics, ↑ means the larger the better, while ↓ means the lower the better. The performance improvements of our model against the baselines are significant under paired t-test.

Dataset	Diginetica						
Metric	precision↑	NDCG↑	MRR↑	F_1 ↑	wv↓	rewarde↑	time↓
w/ [1]	0.79$_{(.009)}$	2.27$_{(.027)}$	1.72$_{(.023)}$	1.32$_{(.021)}$	1.30$_{(.013)}$	-1.53$_{(.021)}$	5781$_{(.023)}$
w/ IFRQE++	0.89$_{(.011)}$	2.85$_{(.015)}$	2.34$_{(.022)}$	1.48$_{(.038)}$	1.19$_{(.036)}$	-1.35$_{(.032)}$	637$_{(.026)}$

Table 8: Comparison between our framework and "training from scratch". "()" indicates the standard error. The metrics for evaluating the recommendation performance are percentage values with "%" omitted. For the metrics, ↑ means the larger the better, while ↓ means the lower the better.

Dataset	Diginetica						
Metric	precision↑	NDCG↑	MRR↑	F_1 ↑	wv↓	rewarde↑	time↓
w/ SCR	0.89$_{(.014)}$	2.88$_{(.017)}$	2.37$_{(.023)}$	1.48$_{(.021)}$	1.56$_{(.013)}$	-2.27$_{(.031)}$	3558$_{(.023)}$
w/ IFRQE++	1.48$_{(.018)}$	4.68$_{(.010)}$	3.79$_{(.012)}$	2.47$_{(.028)}$	2.00$_{(.036)}$	-2.29$_{(.032)}$	124$_{(.026)}$
Dataset	video						
Metric	precision↑	NDCG↑	MRR↑	F_1 ↑	wv↓	rewarde↑	time↓
w/ SCR	0.89$_{(.014)}$	2.88$_{(.017)}$	2.37$_{(.023)}$	1.48$_{(.021)}$	1.56$_{(.013)}$	-2.27$_{(.031)}$	3558$_{(.023)}$
w/ IFRQE++	1.48$_{(.018)}$	4.68$_{(.010)}$	3.79$_{(.012)}$	2.47$_{(.028)}$	2.00$_{(.036)}$	-2.29$_{(.032)}$	124$_{(.026)}$

across all the evaluation metrics. An important advantage of our framework is that we can complete the optimization process within a much shorter time.

H.6 Comparison between our framework and the method of "training from scratch"

In this section, we compare our framework with the method of "training from scratch" (we call it as **SCR**), where we drop the influence function, and for each action exploration, we retrain the recommender model. We remain the other model components of this method the same as our framework. The comparison results are presented in Table 8. We can see, the reward of our framework is lower than SCR. This is understandable, since SCR uses the true loss, and our framework only leverages the approximated values. However, we find that the reward gap is not large, which may suggest that our designed influence function can well approximate the true loss, and help to achieve satisfied reward. An important superiority of our framework is the efficiency. As can be seen in the last column of Table 8, we can improve the training efficiency by about 28.7 times. This superiority is very important for the recommender system, which is an on-line service, and has to make quick responses the user feedback.

References

[1] Ziqian Chen, Fei Sun, Yifan Tang, Haokun Chen, Jinyang Gao, and Bolin Ding. Proactively control privacy in recommender systems. *arXiv preprint arXiv:2204.00279*, 2022.

[2] Pang Wei Koh and Percy Liang. Understanding black-box predictions via influence functions. In *International conference on machine learning*, pages 1885–1894. PMLR, 2017.

[3] Yao Wu, Christopher DuBois, Alice X Zheng, and Martin Ester. Collaborative denoising auto-encoders for top-n recommender systems. In *Proceedings of the ninth ACM international conference on web search and data mining*, pages 153–162, 2016.

[4] Guorui Zhou, Xiaoqiang Zhu, Chenru Song, Ying Fan, Han Zhu, Xiao Ma, Yanghui Yan, Junqi Jin, Han Li, and Kun Gai. Deep interest network for click-through rate prediction. In *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining*, pages 1059–1068, 2018.