APPENDIX

A AE GENERATION ALGORITHM

The function (6) is used to estimate the gradient from the Natural evolution strategy (NES) Ilyas et al. (2018). The detailed method for zeroth-order AE generation in VFL is presented in Algorithm 2. In step 6, we use antithetic sampling to generate noise for efficiency.

$$\nabla_{\boldsymbol{\eta}_i^t} L(\boldsymbol{\eta}_i^t, \mathcal{C}^t) \approx \frac{1}{\sigma n} \sum_{j=1}^{n} \boldsymbol{\delta}_j L\left(\boldsymbol{\eta}_i^t + \sigma \boldsymbol{\delta}_j, \mathcal{C}^t\right), \tag{6}$$

Algorithm 2 Zeroth-order AE generation in VFL

```
1: Input: Batch $[B^t]$, adversarial embedding $\boldsymbol{h}_{i,a}^t$, benign embedding $\boldsymbol{h}_{i,b}^t, i \in [B^t]$, corruption pattern $\mathcal{C}^t$, learning rate $lr$, the sample size of the Gaussion noise $n$, the perturbation budget $\beta$, query budget $Q$, and the embedding range $[lb_i, ub_i], i \in [B^t]$.
2: Initialization: $\boldsymbol{\eta}_{i,m}^t = \mathbf{0}, m \in [M]$ , $\boldsymbol{\eta}_i^t = [\boldsymbol{\eta}_{i,a_1}^t, \ldots, \boldsymbol{\eta}_{i,a_C}^t]$, counter $s = 0$.
3: for $i \in [B^t]$ do
4:     for $q \in [\frac{Q}{n}]$ do
5:         Clamp the perturbation to $\|\boldsymbol{\eta}_i^t\|_\infty \le \beta(ub_i - lb_i)$.
6:         Make a query to the server with adversarial embedding $\tilde{\boldsymbol{h}}_{i,a}^t = \boldsymbol{h}_{i,a}^t + \boldsymbol{\eta}_i^t$
7:         if the attack is not successful then
8:             Initiate $\frac{n}{2}$ noise vectors $\boldsymbol{\delta}_v \sim \mathcal{N}(0, I), v \in \{1, .., \frac{n}{2}\}$, another $\frac{n}{2}$ noise vectors are $\boldsymbol{\delta}_u = -\boldsymbol{\delta}_v, u \in \{\frac{n}{2}, ..., n\}$.
9:             Clamp the perturbation to $\|\boldsymbol{\eta}_i^t + \boldsymbol{\delta}_j\|_\infty \le \beta(ub_i - lb_i)$, where $j \in [n]$.
10:            Make $n$ queries to the server and estimate the gradient $\hat{\boldsymbol{G}}$ through function (6).
11:            Update the perturbation $\boldsymbol{\eta}_i^t = \boldsymbol{\eta}_i^t - lr * \hat{\boldsymbol{G}}$.
12:        else
13:            Break the loop, store $\boldsymbol{\eta}_i^t$ and $s = s + 1$.
14:        end if
15:    end for
16: end for
17: Clamp $\|\boldsymbol{\eta}_i^t\|_\infty \le \beta(ub_i - lb_i), i \in [B^t]$, return $\boldsymbol{\eta}_i^t$ and the attack success rate $\frac{s}{B^t}$.
```

B PROOFS IN SECTION REGRET ANALYSIS

In this section, we provide detailed proofs of lemmas and the theorem in **Section 6 Regret Analysis** of our paper. We initiate the proof procedure by establishing the definitions for two key events and three supporting facts, intended to streamline the proof process.

Fact 1 (Hoeffding's inequality). *Let $X_1, \ldots, X_n$ be independent i.i.d. random variables bounded in $[a, b]$, then for any $\delta > 0$, we have*

$$\Pr\left(\left|\frac{\sum_{i=1}^n X_i}{n} - \mathbb{E}(X_i)\right| \ge \delta\right) \le 2\exp\left(\frac{-2n\delta^2}{(b-a)^2}\right).$$

Fact 2 (Abramowitz And Stegun 1964). *For a Gaussian distributed random variable Z with mean m and variance σ^2, for any z,*

$$\frac{1}{4\sqrt{\pi}} \cdot e^{-7z^2/2} < \Pr(|Z - m| > z\sigma) \le \frac{1}{2}e^{-z^2/2}$$

Fact 3 (Concentration Bounds). *Let $X_1, \ldots, X_n$ be 0-1-valued random variables. Suppose that there are $0 \leqslant \delta_i \leqslant 1$, for $1 \leqslant i \leqslant n$, such that, for every set $S \subseteq [n]$, $\Pr\left[\wedge_{i \in S} X_i = 1\right] \leqslant \prod_{i \in S} \delta_i$. Let $\delta = (1/n)\sum_{i=1}^n \delta_i$. Then, for any γ such that $\delta \leqslant \gamma \leqslant 1$, we have $\Pr\left[\sum_{i=1}^n X_i \geqslant \gamma n\right] \leqslant e^{-nD(\gamma \| \delta)}$, where $D(a\|b)$ is the cross entropy of a and b.*

Definition 3 (Events $E_1(t)$ and $E_2(t)$). $E_1(t)$ is the event that the optimal arm 1 satisfies $n_1(t) < \frac{(t-1)}{N}, \forall t \in [T - t_0]$. $E_2(t)$ is the event that the optimal arm 1 is not identified in the empirical competitive set $\mathcal{E}^t$ at round t, $t > t_0$.

Based on the above facts and the definition, we then provide the following lemmas.

Lemma 3. *Let* $\gamma = \frac{N-1}{N}$*, and* $\delta = (N-1)(\frac{1}{2}\exp(-\Delta_{\min}^2/16) + 2\exp(-\Delta_{\min}^2/4) - \frac{1}{2}\exp(-5\Delta_{\min}^2/16) + \exp\left(-\frac{(t_0-1)\Delta_{\min}^2}{2N}\right) + \exp(-\Delta_{\min}^2))$*. The probability of event* $E_1(t)$ *is upper bounded by* $\Pr(n_1(t) < \frac{t-1}{N}) \leq \exp\left(-tD(\gamma\|\delta)\right)$.

Proof. Let $X_\tau = 0$ denote the optimal arm 1 is pulled at τ round, and $X_\tau = 1$ denotes that the best arm is not pulled. Considering the probability $\Pr(n_1(t) < \frac{t-1}{N}, t > t_0)$, we assume that each arm is pulled at least two times after the warm-up round t_0. Therefore, we can transform the probability $\Pr(n_1(t) < \frac{t}{N}, t > t_0)$ into $\Pr(\sum_{\tau=t_0+1}^{t} X_\tau > \frac{(N-1)}{N}t + \frac{2N+1}{N}) \leq \Pr(\sum_{\tau=t_0+1}^{t} X_\tau > \frac{(N-1)}{N}t)$.

In our algorithm, for every set $S \subseteq [t - t_0]$, $\Pr\left(\wedge_{\tau\in S} X_\tau = 1\right) = \prod_{\tau\in S}\Pr\left(X_\tau = 1|\mathcal{F}_\tau\right)$, where $\mathcal{F}_\tau$ is the history of pulling the optimal arm 1 until round τ. We first analyze the upper bound of the probability $\Pr\left(X_\tau = 1|\mathcal{F}_\tau\right)$, when $\tau \in [t - t_0]$.

From our algorithm, we can derive that $\Pr\left(X_\tau = 1|\mathcal{F}_\tau\right) \leq \sum_{\ell\in[N]\backslash 1}\Pr\left(\theta_1(\tau) < \theta_\ell(\tau)|\mathcal{F}_\tau\right)$ $+ \sum_{\ell\in[N]\backslash 1}\Pr\left(\hat{\varphi}_1(\tau) < \hat{\mu}_\ell(\tau), n_\ell(t) \geq \frac{(\tau-1)}{N}\right)$, where ℓ is a sub-optimal arm. We then analyze the bound of probablity $\Pr\left(X_\tau = 1|\mathcal{F}_\tau\right)$ as follows:

$$
\begin{aligned}
&\sum_{\ell\in[N]\backslash 1}\Pr\left(\theta_1(\tau) < \theta_\ell(\tau)|\mathcal{F}_\tau\right) + \sum_{\ell\in[N]\backslash 1}\Pr\left(\hat{\varphi}_1(\tau) < \hat{\mu}_\ell(\tau), n_\ell(\tau) \geq \frac{(\tau-1)}{N}\right) \\
&\leq \sum_{\ell\in[N]\backslash 1}\Pr\left(\left(\theta_1(\tau) < \mu_1 - \frac{\Delta_\ell}{2}\right)\bigcup\left(\theta_\ell(\tau) > \mu_1 - \frac{\Delta_\ell}{2}\right)\right) \\
&+ \sum_{\ell\in[N]\backslash 1}\Pr\left(\left(\hat{\varphi}_1(\tau) < \mu_1 - \frac{\Delta_{\min}}{2}\right)\bigcup\left(\hat{\mu}_\ell(\tau) > \mu_1 - \frac{\Delta_{\min}}{2}\right), n_\ell(\tau) \geq \frac{(\tau-1)}{N}\right) \\
&\overset{(a)}{\leq} \sum_{\ell\in[N]\backslash 1}\Pr\left(\theta_1(\tau) < \mu_1 - \frac{\Delta_\ell}{2}\right) + \sum_{\ell\in[N]\backslash 1}\Pr\left(\theta_\ell(\tau) > \mu_\ell + \frac{\Delta_\ell}{2}\right) \\
&+ \sum_{\ell\in[N]\backslash 1}\Pr\left(\hat{\varphi}_1(\tau) < \frac{\sum_{t=1}^{T}\mathbb{E}[r_1^{\max}(t)]}{T} - \frac{\Delta_{min}}{2}\right) \\
&+ \sum_{\ell\in[N]\backslash 1}\Pr\left(\hat{\mu}_\ell(\tau) > \mu_\ell + \frac{\Delta_{min}}{2}, n_\ell(\tau) \geq \frac{(\tau-1)}{N}\right) \\
&\overset{(b)}{\leq} (N-1)((\frac{1}{2}\exp(-\Delta_{\min}^2/16) + 2\exp(-\Delta_{\min}^2/4) \\
&- \frac{1}{2}\exp(-5\Delta_{\min}^2/16) + \exp\left(-\frac{(t_0-1)\Delta_{\min}^2}{2N}\right) + \exp(-\Delta_{\min}^2)),
\end{aligned} \tag{7}
$$

where we have (a) from the union bound. For inequality (b), our objective is to delineate the upper bounds of to derive the upper bound of $\Pr\left(\theta_1(\tau) < \mu_1 - \frac{\Delta_\ell}{2}\right)$ and $\Pr\left(\theta_\ell(\tau) > \mu_\ell + \frac{\Delta_\ell}{2}\right)$. To achieve this, we invert our approach to discuss the lower bounds of $\Pr\left(\theta_1(\tau) \geq \mu_1 - \frac{\Delta_\ell}{2}\right)$ and $\Pr\left(\theta_\ell(\tau) \leq \mu_\ell + \frac{\Delta_\ell}{2}\right)$. We first focus on the probability $\Pr\left(\theta_1(\tau) \geq \mu_1 - \frac{\Delta_\ell}{2}\right)$:

$$\begin{aligned}\Pr\left(\theta_1(\tau) \geq \mu_1 - \frac{\Delta_\ell}{2}\right) &\geq \Pr\left(\theta_1(\tau) \geq \hat{\mu}_1(\tau) - \frac{\Delta_\ell}{4} \geq \mu_1 - \frac{\Delta_\ell}{2}\right) \\ &= \Pr\left(\theta_1(\tau) \geq \hat{\mu}_1(\tau) - \frac{\Delta_\ell}{4}\right) \Pr\left(\hat{\mu}_1(\tau) - \frac{\Delta_\ell}{4} \geq \mu_1 - \frac{\Delta_\ell}{2}\right) \\ &\overset{(c)}{\geq} \left(1 - \frac{1}{4}\exp(-n_1(\tau)\Delta_\ell^2/32)\right)\left(1 - \exp(-n_1(\tau)\Delta_\ell^2/8)\right) \\ &= 1 - \frac{1}{4}\exp(-n_1(\tau)\Delta_\ell^2/32) - \exp(-n_1(\tau)\Delta_\ell^2/8) + \frac{1}{4}\exp(-5n_1(\tau)\Delta_\ell^2/32),\end{aligned} \tag{8}$$

where the inequality (c) is from Fact 1 and 2. Similarly, we can derive

$$\Pr\left(\theta_\ell(\tau) \leq \mu_\ell + \frac{\Delta_\ell}{2}\right) \geq 1 - \frac{1}{4}\exp(-n_\ell(\tau)\Delta_\ell^2/32) - \exp(-n_1(\tau)\Delta_\ell^2/8) + \frac{1}{4}\exp(-5n_\ell(\tau)\Delta_\ell^2/32).$$

Then we can derive (b) using Fact 1 and we have ensured each arm is pulled at least 2 times during the warm-up round t_0.

For every $S \subseteq [t - t_0]$, we have an upper bound value $\delta_{\max}$ for $\Pr(X_\tau = 1 | \mathcal{F}_\tau)$: $\delta_{\max} = (N-1)(\frac{1}{2}\exp(-\Delta_{\min}^2/16) + 2\exp(-\Delta_{\min}^2/4) - \frac{1}{2}\exp(-5\Delta_{\min}^2/16) + \exp\left(-\frac{(t_0-1)\Delta_{\min}^2}{2N}\right) + \exp(-\Delta_{\min}^2))$. Let $\delta = \frac{1}{t-t_0}\sum_{\tau=t_0+1}^{t} \delta_{\max} = \delta_{\max}$ and $\gamma = \frac{(N-1)}{N}$, we can derive the following bound from Fact 3:

$$\Pr(n_1(t) < \frac{t-1}{N}) = \Pr(\sum_{\tau=0}^{t} X_\tau > t\frac{(N-1)}{N}) \leq \exp\left(-tD(\gamma \| \delta)\right). \tag{9}$$

□

Lemma 4. *After the warm-up round t_0, for any sub-optimal arm $k \neq 1, \Delta_k = \mu_1 - \mu_k \geq 0$, the following inequality holds,*

$$\sum_{t=t_0+1}^{T} \Pr\left(k = k^{\text{emp}}(t), n_1(t) \geq \frac{(t-1)}{N}\right) \leq \frac{4N}{\Delta_k^2}$$

Proof. We bound the probability by :

$$\begin{aligned}&= \sum_{t=t_0+1}^{T} \Pr\left(k = k^{\text{emp}}(t), n_1(t) \geq \frac{(t-1)}{N}\right) \\ &\overset{(d)}{=} \sum_{t=t_0+1}^{T} \Pr\left(k = k^{\text{emp}}(t), n_1(t) \geq \frac{(t-1)}{N}, n_k(t) \geq \frac{(t-1)}{N}\right) \\ &\leq \sum_{t=t_0+1}^{T} \Pr\left(\hat{\mu}_k(t) \geq \hat{\mu}_1(t), n_k(t) \geq \frac{(t-1)}{N}, n_1(t) \geq \frac{(t-1)}{N}\right) \\ &\leq \sum_{t=t_0+1}^{T} \Pr\left(\left((\hat{\mu}_1(t) \leq \mu_1 - \frac{\Delta_k}{2}) \bigcup (\hat{\mu}_k(t) \geq \mu_1 - \frac{\Delta_k}{2})\right), n_k(t) \geq \frac{(t-1)}{N}, n_1(t) \geq \frac{(t-1)}{N}\right) \\ &= \sum_{t=t_0+1}^{T} \Pr\left(\left((\hat{\mu}_1(t) \leq \mu_1 - \frac{\Delta_k}{2}) \bigcup (\hat{\mu}_k(t) \geq \mu_k + \frac{\Delta_k}{2})\right), n_k(t) \geq \frac{(t-1)}{N}, n_1(t) \geq \frac{(t-1)}{N}\right) \\ &\overset{(e)}{\leq} \sum_{t=t_0+1}^{T} \Pr\left(\hat{\mu}_1(t) - \mu_1 \leq -\frac{\Delta_k}{2}, n_1(t) \geq \frac{(t-1)}{N}\right) + \sum_{t=t_0+1}^{T} \Pr\left(\hat{\mu}_k(t) - \mu_k \geq \frac{\Delta_k}{2}, n_k(t) \geq \frac{(t-1)}{N}\right) \\ &\overset{(f)}{\leq} \sum_{t=t_0+1}^{T} 2\exp\left(\frac{-(t-1)\Delta_k^2}{2N}\right) \overset{(g)}{\leq} \frac{4N}{\Delta_k^2},\end{aligned} \tag{10}$$

Here, (d) holds because of the truth that the empirical best arm is $k^{emp}(t)$ selected from the set $\mathcal{S}_t = \{k \in [N] : n_k(t) \geq \frac{(t-1)}{N}\}$. Inequality (e) follows the union bound. We have (f) from the truth that $\hat{\mu}_k(t) = \frac{\sum_{\tau=1}^{t} r_k(\tau)\mathbb{1}(k(\tau)=k)}{n_k(t)}, \forall k \in [N]$ and Fact 1. The last inequality (g) uses the fact that $\frac{\Delta_k^2}{2N} > 0$ and the geometric series. □

Proof of Lemma 1. Now, we prove Lemma 1 in the main paper.

Proof. During t_0 warm-up rounds, the maximum pulling times of a non-competitive arm k^{nc} are bound in t_0. We then analyze the expected number of times pulling k^{nc} after round t_0.

$$
\begin{aligned}
&\sum_{t=t_0+1}^{T} \Pr(k(t) = k^{nc}) \\
&= \sum_{t=t_0+1}^{T} \Pr(k(t) = k^{nc}, n_1(t) \geq \frac{(t-1)}{N}) + \sum_{t=t_0+1}^{T} \Pr(k(t) = k^{nc}, n_1(t) < \frac{(t-1)}{N}) \\
&\overset{(h)}{\leq} \sum_{t=t_0+1}^{T} \Pr(k(t) = k^{nc}, k^{nc} = k^{emp}(t), n_1(t) \geq \frac{(t-1)}{N}) \\
&+ \sum_{t=t_0+1}^{T} \Pr\left(k(t) = k^{nc}, k^{nc} \in \mathcal{S}_t \setminus k^{emp}(t), n_1(t) \geq \frac{(t-1)}{N}\right) + \sum_{t=t_0+1}^{T} \Pr(n_1(t) < \frac{(t-1)}{N}) \\
&\overset{(i)}{\leq} \sum_{t=t_0+1}^{T} \Pr\left(\hat{\mu}_1(t) \leq \hat{\varphi}_{k^{nc}}(t), k(t) = k^{nc}, n_1(t) \geq \frac{(t-1)}{N}\right) + \frac{4N}{\Delta_{k^{nc}}^2} + \sum_{t=t_0+1}^{T} \exp\left(-tD(\gamma\|\delta)\right) \\
&\leq \sum_{t=t_0+1}^{T} \Pr\left(\left((\hat{\mu}_1(t) \leq \mu_1 + \frac{\tilde{\Delta}_{k^{nc},1}}{2}) \bigcup (\hat{\varphi}_{k^{nc}}(t) \geq \mu_1 + \frac{\tilde{\Delta}_{k^{nc},1}}{2})\right), k(t) = k^{nc}, n_1(t) \geq \frac{(t-1)}{N}\right) \\
&+ \sum_{t=t_0+1}^{T} \exp\left(-tD(\gamma\|\delta)\right) + \frac{4N}{\Delta_{k^{nc}}^2} \\
&\overset{(j)}{\leq} \sum_{t=t_0+1}^{T} \Pr\left(\hat{\mu}_1(t) \leq \mu_1 + \frac{\tilde{\Delta}_{k^{nc},1}}{2} n_1(t) \geq \frac{(t-1)}{N}\right) \\
&+ \sum_{t=t_0+1}^{T} \Pr\left(\hat{\varphi}_{k^{nc}}(t) \geq \frac{\sum_{t=1}^{T} \mathbb{E}[r_{k^{nc}}^{max}(t)]}{T} - \frac{\tilde{\Delta}_{k^{nc},1}}{2}, k(t) = k^{nc}\right) + \sum_{t=t_0+1}^{T} \exp\left(-tD(\gamma\|\delta)\right) + \frac{4N}{\Delta_{k^{nc}}^2} \\
&\overset{(k)}{\leq} \sum_{t=t_0+1}^{T} \exp\left(\frac{-(t-1)\tilde{\Delta}_{k^{nc},1}^2}{2N}\right) + \sum_{j=1}^{T} \Pr\left(\hat{\varphi}_{k^{nc}}(\tau_j) - \frac{\sum_{t=1}^{T} \mathbb{E}[r_{k^{nc}}^{max}(t)]}{T} \geq -\frac{\tilde{\Delta}_{k^{nc},1}}{2}\right) \\
&+ \sum_{t=t_0+1}^{T} \exp\left(-tD(\gamma\|\delta)\right) + \frac{4N}{\Delta_{k^{nc}}^2} \\
&\overset{(l)}{\leq} \sum_{t=t_0+1}^{T} \exp\left(\frac{-(t-1)\tilde{\Delta}_{k^{nc},1}^2}{2N}\right) + \sum_{j=1}^{T} \exp\left(-\frac{j\tilde{\Delta}_{k^{nc},1}^2}{2}\right) + \sum_{t=t_0+1}^{T} \exp\left(-tD(\gamma\|\delta)\right) + \frac{4N}{\Delta_{k^{nc}}^2} \\
&\overset{(m)}{\leq} \frac{2N}{\tilde{\Delta}_{k^{nc},1}^2} + \frac{2}{\tilde{\Delta}_{k^{nc},1}^2} + \frac{1}{D(\gamma\|\delta)} + \frac{4N}{\Delta_{k^{nc}}^2} = \mathcal{O}(1),
\end{aligned}
\tag{11}
$$

Here, both (h) and (j) are derived using the union bound. We have (i) from the Lemma 4 and Lemma 3. The inequality (k) is obtained from Fact 1, wherein j in (k) explicitly denotes the round index when arm k^{sub} is pulled. Inequality (l) stems from Fact 1. We have (m) because of the truth that $\frac{\tilde{\Delta}_{k^{nc},1}^2}{2N} \geq 0$, $\frac{\tilde{\Delta}_{k^{nc},1}^2}{2} \geq 0$, and $D(\gamma\|\delta) \geq 0$. We also use geometric series in (m).

□

We provide another Lemma to facilitate the proof of Lemma 2 in the main paper.

Lemma 5. *The following inequality holds,*

$$\Pr\left(E_2(t)\right) \leq 4(N-1)t\exp\left(-\frac{(t-1)\Delta_{min}^2}{2N}\right) + \frac{(N-1)}{D(\gamma\|\delta)},$$

where $\Delta_{\min} = \min_k \Delta_k$.

Proof.

$$\begin{aligned}
\Pr\left(E_2(t)\right) &\overset{(n)}{\leq} \sum_{\ell\in[N]\setminus 1} \Pr\left(\hat{\varphi}_1(t) < \hat{\mu}_\ell(t), n_1(t) \geq \frac{(t-1)}{N}, n_\ell(t) \geq \frac{(t-1)}{N}\right) \\
&+ \sum_{\ell\in[N]\setminus 1} \Pr\left(\hat{\varphi}_1(t) < \hat{\mu}_\ell(t), n_1(t) < \frac{(t-1)}{N}, n_\ell(t) \geq \frac{(t-1)}{N}\right) \\
&+ \sum_{\ell\in[N]\setminus 1} \Pr\left(\hat{\mu}_1(t) < \hat{\mu}_\ell(t), n_1(t) \geq \frac{(t-1)}{N}, n_\ell(t) \geq \frac{(t-1)}{N}\right) \\
&\leq \sum_{\ell\in[N]\setminus 1} \Pr(\Big((\hat{\varphi}_1(t) < \mu_1 - \frac{\Delta_{min}}{2})\bigcup(\hat{\mu}_\ell(t) > \mu_1 - \frac{\Delta_{min}}{2})\Big), \\
&n_1(t) \geq \frac{(t-1)}{N}, n_\ell(t) \geq \frac{(t-1)}{N}) \\
&+ \sum_{\ell\in[N]\setminus 1} \Pr(\Big((\hat{\mu}_1(t) < \mu_1 - \frac{\Delta_{min}}{2})\bigcup(\hat{\mu}_\ell(t) > \mu_1 - \frac{\Delta_{min}}{2})\Big), \\
&n_1(t) \geq \frac{(t-1)}{N}, n_\ell(t) \geq \frac{(t-1)}{N}) + \sum_{\ell\in[N]\setminus 1} \Pr\left(n_1(t) < \frac{(t-1)}{N}\right) \\
&\overset{(o)}{\leq} \sum_{\ell\in[N]\setminus 1} \Pr\left((\hat{\varphi}_1(t) < \frac{\sum_{t=1}^T \mathbb{E}[r_1^{\max}(t)]}{T} - \frac{\Delta_{min}}{2}, n_1(t) \geq \frac{(t-1)}{N}\right) \\
&+ \sum_{\ell\in[N]\setminus 1} \Pr\left((\hat{\mu}_1(t) < \mu_1 - \frac{\Delta_{min}}{2}, n_1(t) \geq \frac{(t-1)}{N}\right) \\
&+ 2\sum_{\ell\in[N]\setminus 1} \Pr\left(\hat{\mu}_\ell(t) > \mu_\ell + \frac{\Delta_{min}}{2}, n_\ell(t) \geq \frac{(t-1)}{N}\right) + (N-1)\exp\left(-tD(\gamma\|\delta)\right) \\
&\overset{(p)}{\leq} 4(N-1)\exp\left(-\frac{(t-1)\Delta_{min}^2}{2N}\right) + (N-1)\exp\left(-tD(\gamma\|\delta)\right),
\end{aligned} \tag{12}$$

Inequality (n), using union bound, arises from the observation that when arm 1 is absent from the empirical competitive set $\mathcal{E}^t$ at round t, it is either not selected as the empirical best arm $k^{emp}(t)$ or its $\hat{\varphi}_1(t)$ is less than the estimated mean of the empirical best arm $\hat{\mu}_{k^{emp}(t)}(t)$. The validity of inequality (n) relies on the fact that $\frac{\sum_{t=1}^T \mathbb{E}[r_1^{\max}(t)]}{T} \geq \mu_1$ and Lemma 3. We establish the final inequality (p) by leveraging Fact 1. □

Proof of Lemma 2. Now, we present proof details of Lemma 2 in the main paper.

Proof. We split the analysis of $\sum_{t=1}^T \Pr(k(t) = k^{sub})$ into three parts: the pulls in the warm round; the pulls when the event $E_2(t)$ happens after the warm round; the pulls when the complementary of

$E_2(t)$ happens. We summarize it as follows:

$$\begin{aligned}
\sum_{t=1}^{T}\Pr\left(k(t)=k^{sub}\right) &= \sum_{t=1}^{t_0}\Pr\left(k(t)=k^{sub}\right)+\sum_{t=t_0+1}^{T}\Pr\left(k(t)=k^{sub},E_2(t)\right)\\
&\quad+\sum_{t=t_0+1}^{T}\Pr\left(k(t)=k^{sub},E_2^c(t)\right)\\
&\le \sum_{t=1}^{T}\Pr\left(k(t)=k^{sub}\right)+\sum_{t=t_0+1}^{T}\Pr\left(E_2(t)\right)
\end{aligned} \tag{13}$$

When event $E_2(t)$ does not happen, the analysis of the upper bound of pulling the competitive but sub-optimal arm aligns to plain TS. We apply the result from Agrawal and Goyal (2012), which bounds the number of times a sub-optimal arm $k \neq 1$ is pulled within $\mathcal{O}(\log(T))$. In Lemma 5, when $E_2(t)$ happens, we derive the following bound:

$$\begin{aligned}
\sum_{t=t_0+1}^{T}\Pr(E_2(t)) &\le \sum_{t=t_0+1}^{T}\left(4(N-1)\exp\left(-\frac{(t-1)\Delta_{min}^2}{2N}\right)+(N-1)\exp\left(-tD(\gamma\|\delta)\right)\right)\\
&\le \frac{8N(N-1)}{\Delta_{\min}^2}+\frac{1}{D(\gamma\|\delta)}\\
&=\mathcal{O}(1).
\end{aligned} \tag{14}$$

The proof is completed. □

Proof the Theorem 1.

Proof. We revisit the definition of expected regret, given by:

$$\mathbb{E}[R(T)]=\mathbb{E}\left[\sum_{t=1}^{T}(\mu_1-\mu_{k(t)})\right]=\mathbb{E}\left[\sum_{k=1}^{N}n_k(T)\Delta_k\right].$$

Considering D competitive arms and $(N-D)$ non-competitive arms, the regret of E-TS in T rounds is bounded by:

$$\begin{aligned}
\mathbb{E}[R(T)] &= \sum_{k^{nc}\in[N-D]}\mathbb{E}[n_{k^{nc}}(T)]\Delta_{k^{nc}}+\sum_{k^{sub}\in[D]}\mathbb{E}[n_{k^{sub}}(T)]\Delta_{k^{sub}}\\
&\overset{(1)}{\le}\sum_{k^{nc}\in[N-D]}\Delta_{k^{nc}}\mathcal{O}(1)+\sum_{k^{sub}\in[D]}\Delta_{k^{sub}}\mathcal{O}(\log(T))\\
&\le (N-D)\mathcal{O}(1)+D\mathcal{O}(\log(T)),
\end{aligned} \tag{15}$$

where the inequality (1) is from Lemma 2 and Lemma 1. Thus the proof is finalized. □

C Supplementary experiments and experimental details

C.1 Dataset and model structure

Table 1 provides essential information about each dataset used in our study. We will introduce more details regarding the dataset characteristics and the corresponding model structures.

The **Credit** dataset consists of information regarding default payments, demographic characteristics, credit data, payment history, and credit card bill statements from clients in Taiwan. The dataset is partitioned evenly across six clients, each managing a bottom model with a Linear-BatchNorm-ReLU structure. The server hosts the top model, comprising of two Linear-ReLU-BatchNorm layers followed by a WeightNorm-Linear-Sigmoid layer.

The **Real-sim** dataset is from LIBSVM, which is a library for support vector machines (SVMs). 10 clients equally hold the data features and compute embeddings through a bottom model with 2 Linear-ReLU-BatchNorm layers. The server controls the top model with 3 Linear-ReLU layers.

The **FashionMNIST** dataset consists of 28×28 grayscale images of clothing items. The dataset is equitably distributed across 7 clients, with each holding a data portion of 28×4 dimensions. On the client side, it holds a Linear-BatchNorm-ReLU bottom model. On the server side, the top model comprises eight groups of Conv-BatchNorm-ReLU structures, two MaxPool layers, two Linear-Dropout-ReLU layers, and a final Linear output layer.

The **CIFAR-10** dataset contains 60,000 color images of size 32×32, representing vehicles and animals. We divide each image into 4×32 sub-images and distribute them among 8 clients. Each client's bottom model consists of 2 convolutional layers and 1 max-pooling layer. The server's top model is built with 6 convolutional layers and 3 linear layers.

The **Caltech-7** dataset, a subset of seven classes from the Caltech-101 object recognition collection, is distributed across six clients. Each client is assigned one unique feature view, encompassing the Gabor feature, Wavelet moments (WM), CENTRIST feature, Histogram of Oriented Gradients (HOG) feature, GIST feature, and Local Binary Patterns (LBP) feature, respectively. Every client maintains a bottom model utilizing a Linear-BatchNorm-ReLU structure. At the server level, the top model comprises eight Linear-ReLU layers, two Dropout layers, and a final Linear output layer.

The **IMDB** dataset comprises 50,000 highly polarized movie reviews, each categorized as either positive or negative. For distributed processing across 6 clients, each review is divided into several sentences, and an equal number of these sentences are allocated to each client. Each client utilizes a Bert model without fine-tuning—at the bottom level to obtain an embedding with 512 dimensions. These embeddings are then input to the server's top model, which consists of two Linear-ReLU layers followed by a final Linear output layer.

Table 1: VFL dataset and parameters descriptions.

Task	Tabular		CV		Multi-view	NLP
Dataset name	Credit	Real-sim	FashionMNIST	CIFAR10	Caltech-7	IMDB
Number of samples	30,000	72,309	70,000	60,000	1474	50,000
Feature size	23	20,958	784	1024	3766	-
Number of classes	2	2	10	10	7	2
Number of clients	7	10	7	8	6	6
Batchsize B^t	32	512	128	32	16	64
Warm-up rounds t_0	50	50	80	80	80	40

C.2 Experimental result in ablation study

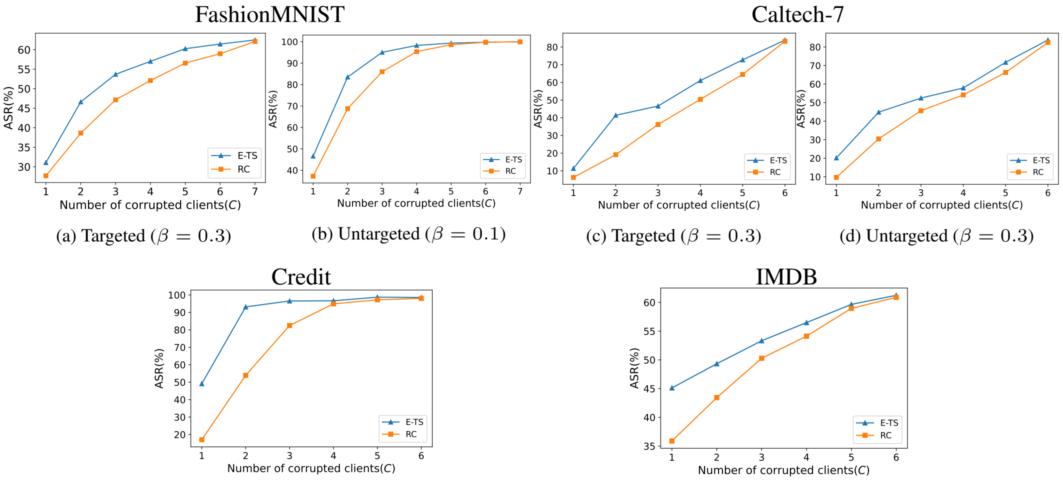
Additional experiments have been conducted across a variety of datasets under diverse corruption constraints, as illustrated in Figure 5.

C.3 Dynamics of arm Selection and empirical competitive set in TS and E-TS

We investigated the arm selection behavior of TS and E-TS during a targeted attack on FashionMNIST, as shown in Figure 6. This study also tracked the variation in the size of E-TS's empirical competitive set, depicted in Figure 6. The parameters for this analysis were consistent with those in the FashionMNIST targeted attack scenario (Figure 1): $t_0 = 80$, $C = 2$, $\beta = 0.15$, $Q = 2000$ and the number of arms $N = \binom{7}{2} = 21$. We list all arms as follow:

[0: (client 1, client 2), 1: (client 1, client 3), 2: (client 1, client 4), 3: (client 1, client 5), 4:(client 1, client 6), 5: (client 1, client 7), 6: (client 2, client 3), 7: (client 2, client 4), 8: (client 2, client 5), 9: (client 2, client 6), 10: (client 2, client 7), 11: (client 3, client 4), 12: (client 3, client 5), 13: (client 3, client 6), 14: (client 3, client 7), 15: (client 4, client 5), 16: (client 4, client 6), 17: (client 4, client 7), 18: (client 5, client 6), 19: (client 5, client 7), 20: (client 6, client 7)].

Analysis of Figure 6(a) reveals that initially, E-TS selected a suboptimal arm. However, after 140 rounds, it consistently chose arm 5 (representing the pair of client 1 and client 7), indicating a stable



(a) Targeted ($\beta = 0.3$)

(b) Untargeted ($\beta = 0.1$)

(c) Targeted ($\beta = 0.3$)

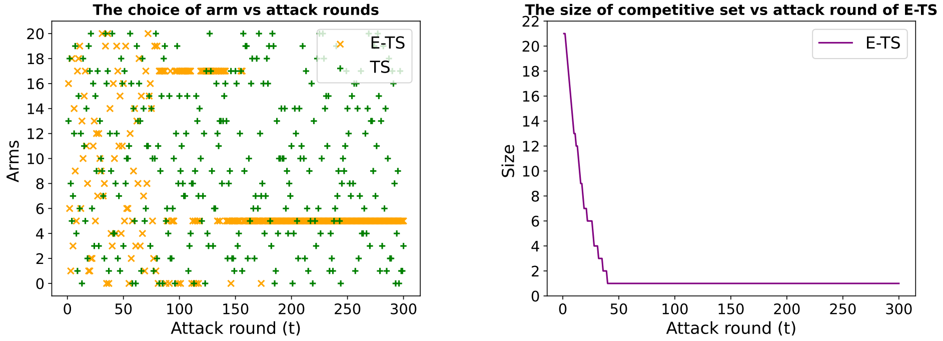
(d) Untargeted ($\beta = 0.3$)

(e) $\beta = 0.3$

(f) $\beta = 0.05$

Figure 5: ASR using different number of corrupted clients.

selection. In contrast, TS continued to explore different arms during this period. Figure 6(b) shows that the empirical competitive set in E-TS reduced to a single arm within the first 40 rounds. Initially, the competitive arm selected by E-TS was not optimal. Nevertheless, E-TS effectively narrowed down its focus to this suboptimal arm, eventually dismissing it as non-competitive and identifying the best arm for selection.



(a) The choice of arm in E-TS and TS.

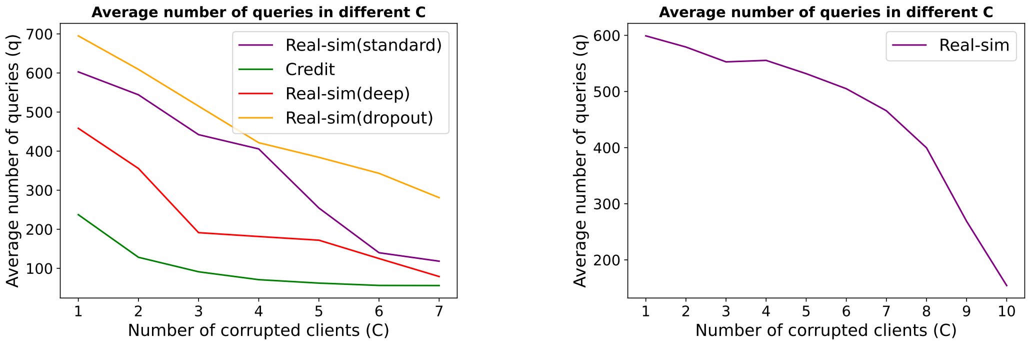
(b) The change of the empirical competitive set's size.

Figure 6: Dynamics of arm selection and competitive set in E-TS and TS.

C.4 Minimum query budget and corruption channels to achieve 50% ASR

To explore how the necessary number of queries and corrupted channels vary across different models, datasets, and systems, we conducted experiments using Credit and Real-sim datasets. We specifically analyzed the average number of queries q required to attain a 50% ASR under various levels of client corruption (corruption constraint C). For this analysis, we applied the proposed attack on both the Credit and Real-sim datasets in a 7-client setting. We varied C from 1 to 7 and recorded the average queries q needed for attacking over 50% of the samples successfully. In addition to assessing the impact of different datasets, we investigated the influence of model complexity by attacking two deeper Real-sim models contrasting it with the standard 3-layer server model. Specifically, the standard 3-layer model Real-sim(standard) has a Dropout layer after the first layer of the server model and achieves 96.1% test accuracy. One deeper server model Real-sim(deep) added an extra three layers to the Real-sim(standard) after the Dropout layer of the server model. Another model

Real-sim(dropout) structure is the same as Real-sim(deep) except that it added another Dropout layer before the penultimate layer of the server model. Both Real-sim(deep) and Real-sim(dropout) have 97% test accuracy. Furthermore, to analyze the system's effect on q and C, we conducted experiments on Real-sim in a 10-client scenario, varying C from 1 to 10 and recording q. Throughout these experiments, we maintained $\beta = 0.8$ and $t_0 = 2N$, where N denotes the number of arms. The results are presented in Figure 7.



(a) 7 clients.

(b) 10 clients.

Figure 7: Average number of queries in different corruption constraint to achieve 50% ASR.

From Figure 7, we observe that the required average number of queries decreases with a looser (or higher) corruption constraint C. The comparison of Real-sim and Credit (Figure 7(a)) reveals that simpler datasets in the same task category (both being tabular datasets) necessitate fewer queries.

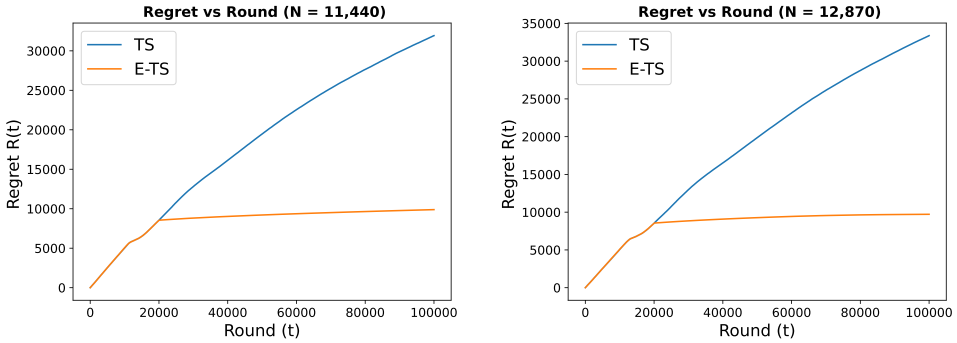
Contrary to our initial assumption, a deeper model does not necessarily require more queries. The results for Real-sim(standard), Real-sim(deep), and Real-sim(norm) from Figure 7(a) suggest that attacking a Real-sim(deep) requires fewer queries. A deeper model with an extra Dropout layer can make the model more robust and needs more quires to achieve 50% ASR. The reason for that is the deeper model will learn a different hidden feature of the sample, thus making the model have different robustness compared to the shallow one. Dropout can enhance robustness by preventing the model from becoming overly reliant on any single feature or input node, encouraging the model to learn more robust and redundant representations.

Comparing Figure 7 (a) and (b), we deduce that systems with more clients demand a greater number of queries to achieve the same ASR at a given C, due to each client possessing fewer features.

In conclusion, to attain a target ASR with the same C, simpler datasets within the same task require fewer queries. Systems with a higher number of clients necessitate more queries. However, the influence of the model's complexity does not simply depend on the scales of model parameters but is affected more by the Dropout layer.

C.5 Discussion on the large exploration spaces

We extend the experiments in Figure 4 to larger exploration spaces, i.e. set the corruption constraint $C = 7$ and $C = 8$, which results in $\binom{16}{7} = 11,440$, $\binom{16}{8} = 12,870$ arms, respectively. However, constrained by the computation power and limited time in the rebuttal period, we compare E-TS and plain TS in large exploration spaces through numerical simulation where ASR is substituted with a sample in Gaussian distribution. For the simulation, we created a list of means starting from 0 up to 0.99, in increments of 0.01, each with a variance of 0.1. This list was extended until it comprised $11,440 - 1$ and $12,870 - 1$ elements, to which we added the best arm, characterized by a mean of 1 and a variance of 0.1. This list represents the underlying mean and variance of the arms. Upon playing an arm, a reward is determined by randomly sampling a value, constrained to the range $[0, 1]$. With knowledge of the underlying mean, we plotted the cumulative regret over rounds, $R(t) = \sum_{\tau=1}^{t}(\mu_1 - \mu_{k(\tau)})$, where μ_1 is the best arm's mean, and $k(\tau)$ is the arm selected in round τ. These results are presented in Figure 8.



(a) $N = 11,440$.

(b) $N = 12,870$.

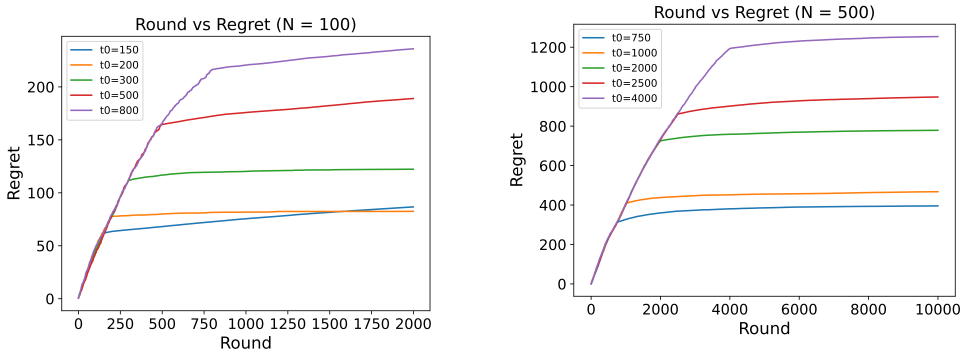
Figure 8: Regret in large exploration spaces.

The results from Figure 8 reveal that in large exploration spaces, TS struggles to locate the best arm within a limited number of rounds. In contrast, E-TS demonstrates more rapid convergence, further confirming the benefits of utilizing an empirical competitive set in large exploration spaces.

C.6 The study of optimal choice on the warm-up round t_0

To ascertain the ideal number of warm-up rounds t_0 for different arm settings, we conducted numerical experiments with $N = 100$ and $N = 500$. For $N = 100$, we experimented with $t_0 = 150$ (less than $2N$), $t_0 = 200, 300, 500$ (within $[2N, 5N]$), and $t_0 = 800$ (greater than $5N$). Similarly, for $N = 500$, the settings were $t_0 = 750$ (less than $2N$), $t_0 = 1000, 2000, 2500$ (within $[2N, 5N]$), and $t_0 = 4000$ (greater than $5N$).

In these experiments, ASR was replaced with Gaussian distribution samples. We initialized 100 arms with means from 0 to 0.99 (in 0.01 increments) and variances of 0.1. The reward for playing an arm was sampled from its Gaussian distribution. The cumulative regret $R(T)$ was computed as $R(T) = \sum_{t=1}^{T}(\mu_1 - \mu_{k(t)})$, where $\mu_1 = 0.99$ and $k(t)$ is the arm selected at round t, as illustrated in Figure 9.



(a) N = 100.

(b) N = 500.

Figure 9: E-TS performance using different warm-up rounds.

Figure 9(a) shows that E-TS converges faster with a smaller t_0, but with $t_0 = 150$, it converges to a sub-optimal arm. Figure 9(b) indicates faster convergence with smaller t_0. Both figures suggest that $t_0 = 2N$ achieves the most stable and rapid convergence, while $t_0 > 5N$ results in the slowest convergence rate. Analyzing the pull frequencies of each arm during t_0, we find that with $t_0 < 2N$, most arms are pulled only once, and some are never explored. Conversely, with $t_0 \in [2N, 5N]$, most arms are pulled at least twice, yielding a more reliable estimation of their prior distributions.

Thus, we recommend setting t_0 to at least N, with the optimal range being $[2N, 5N]$ in practical scenarios. This range ensures that each arm is sampled at least twice using TS, enabling a more accurate initial assessment of each arm's prior distribution. Such preliminary knowledge is vital for E-TS to effectively form an empirical competitive set of arms. If t_0 is too small, there's an increased risk of E-TS prematurely converging on a suboptimal arm due to inadequate initial data, possibly overlooking the best arm in the empirical competitive set.