A. AUER-N

In this section, we present the proof of the regret guarantee for AUER-N. For the sake of this analysis, we are in fact assuming that the awake sets A^t are generated *arbitrarily* before learning starts. This implies the claimed result when the A^t's are generated i.i.d. according to an arbitrary distribution over $\mathcal{A}$, independent of the losses. We use $\mathbb{I}\{\cdot\}$ to denote the indicator function.

We start off with the following technical lemma.

Lemma 1 *Assume the following ordering for μ_j, $j \in [K]$: $\mu_1 < \mu_2 < \cdots < \mu_K$ and, for $i < j$, let $\Delta_{i,j} = \mu_j - \mu_i$. Then, for any $F_j > 0$, $j \in [K]$, the following inequality holds:*

$$\sum_{1 \le i < j \le K} F_j \frac{\Delta_{i,i+1}}{\Delta_{i,j}^2} \le 2 \sum_{j=2}^{K} \frac{F_j}{\Delta_{j-1,j}} \, .$$

Proof. The result and the proof are extensions of Lemma 3 of Kleinberg et al. (2008) to inequalities augmented with factors F_j. We will use the following equality which, by definition of the Lebesgue integral, holds for any non-negative function f: $\mathbb{E}[f(X)] = \int_0^{+\infty} \mathbb{P}[f(x) \ge t] \, dt$. Thus, considering, in particular, the uniform probability over $\{2, \ldots, K\}$, we can write:

$$\begin{aligned}
\frac{1}{K-1} \sum_{j=2}^{K} \frac{\mathbb{I}\{j > i\} F_j}{\Delta_{i,j}^2} &= \int_0^{+\infty} \mathbb{P}\left[\frac{\mathbb{I}\{j > i\} \, F_j}{\Delta_{i,j}^2} \ge t\right] dt = \int_0^{+\infty} \mathbb{P}\left[\frac{\mathbb{I}\{j > i\} \, F_j^{\frac{1}{2}}}{\Delta_{i,j}} \ge \sqrt{t}\right] dt \\
&= 2 \int_0^{+\infty} \mathbb{P}\left[\frac{\mathbb{I}\{j > i\} \, F_j^{\frac{1}{2}}}{\Delta_{i,j}} \ge u\right] u \, du \qquad (u = \sqrt{t}) \\
&= \frac{2}{K-1} \int_0^{+\infty} \sum_{j=2}^{K} \mathbb{I}\{\frac{\mathbb{I}\{j > i\} \, F_j^{\frac{1}{2}}}{\Delta_{i,j}} \ge t\} \, t \, dt.
\end{aligned}$$

In view of the above, we have

$$\begin{aligned}
\sum_{1 \le i < j \le K} F_j \frac{\Delta_{i,i+1}}{\Delta_{i,j}^2} &= \sum_{i=1}^{K-1} \Delta_{i,i+1} \sum_{j \colon j > i} \frac{F_j}{\Delta_{i,j}^2} \\
&= 2 \sum_{i=1}^{K-1} \Delta_{i,i+1} \int_0^{+\infty} \sum_{j=2}^{K} \mathbb{I}\{\frac{\mathbb{I}\{j > i\} \, F_j^{\frac{1}{2}}}{\Delta_{i,j}} \ge t\} \, t \, dt \\
&= 2 \int_0^{+\infty} \sum_{i=1}^{K-1} \Delta_{i,i+1} \sum_{j=2}^{K} \mathbb{I}\{\frac{\mathbb{I}\{j > i\} \, F_j^{\frac{1}{2}}}{\Delta_{i,j}} \ge t\} \, t \, dt \\
&= 2 \int_0^{+\infty} \sum_{1 \le i < j \le K} \Delta_{i,i+1} \, \mathbb{I}\{\frac{F_j^{\frac{1}{2}}}{\Delta_{i,j}} \ge t\} t \, dt.
\end{aligned}$$

Now, for any $j \in \{2, \ldots, K\}$ and $t > 0$, define $i_t(j)$ by

$$i_t(j) = \operatorname{argmin} \left\{ i \in [K] \colon i \le j, \Delta_{i,j} \le \frac{F_j^{\frac{1}{2}}}{t} \right\}.$$

The index $i_t(j)$ is well defined since for $i = j$, $\Delta_{j,j} = 0$ is upper bounded by $\frac{F_j^{\frac{1}{2}}}{t}$. By definition of $i_t(j)$, we can write

$$\begin{aligned}
\sum_{1\leq i<j\leq K} F_j \frac{\Delta_{i,i+1}}{\Delta_{i,j}^2} = 2\int_0^{+\infty} \sum_{j=2}^{K}\sum_{i=i_t(j)}^{j-1} \Delta_{i,i+1}\, t\, dt &= 2\int_0^{+\infty} \sum_{j=2}^{K} \Delta_{i_t(j),j}\, t\, dt \\
&= 2\sum_{j=2}^{K}\int_0^{+\infty} \Delta_{i_t(j),j}\, t\, dt \\
&= 2\sum_{j=2}^{K}\int_0^{\frac{F_j^{\frac{1}{2}}}{\Delta_{j-1,j}}} \Delta_{i_t(j),j}\, t\, dt \qquad \left(\text{for } t \geq \frac{F_j^{\frac{1}{2}}}{\Delta_{j-1,j}},\, i_t(j) = j\right) \\
&\leq 2\sum_{j=2}^{K}\int_0^{\frac{F_j^{\frac{1}{2}}}{\Delta_{j-1,j}}} F_j^{\frac{1}{2}}\, dt \\
&= 2\sum_{j=2}^{K} \frac{F_j}{\Delta_{j-1,j}} \qquad \left(\text{by def. } \Delta_{i_t(j),j} \leq \frac{F_j^{\frac{1}{2}}}{t}\right),
\end{aligned}$$

which completes the proof. □

With the above lemma handy, we are ready to prove Theorem 1.

Proof. [Theorem 1] For any $i, j \in [K]$, $i < j$, let $M_{i,j}$ denote the number of times expert ξ_j is selected by the algorithm, while some expert ξ_k with $k \in [i]$ could have been selected (because it was awake), where $[i] = \{1, \ldots, i\}$. By definition, $(M_{i,j} - M_{i-1,j})$ is then the number of times expert ξ_j,is selected by the algorithm, while expert ξ_i, $i < j$, could have been selected. Then, using the convention $\Delta_{j,j} = 0$ and $M_{0,j} = 0$ for any $j \in [K]$, the sleeping regret of the algorithm can be expressed as follows:

$$\mathrm{r}_T^{\text{SLEEP}}(\text{AUER-N}) = \mathbb{E}\left[\sum_{1\leq i<j\leq K} (M_{i,j} - M_{i-1,j})\Delta_{i,j}\right] = \sum_{j=2}^{K}\sum_{i=1}^{j-1} \mathbb{E}[M_{i,j}][\Delta_{i,j} - \Delta_{i+1,j}]\,. \tag{1}$$

Thus, to bound the sleeping regret, it suffices to bound $\mathbb{E}[M_{i,j}]$ for $1 \leq i < j \leq K$. This expectation can be rewritten as follows

$$\mathbb{E}[M_{i,j}] = \sum_{t=1}^{T} \mathbb{E}\left[\mathbb{I}\{I_t = j\}\,\mathbb{I}\{A^t \cap [i] \neq \emptyset\}\right]\,, \tag{2}$$

where A^t denotes set of experts awake at time t. For any random variable $\sigma_{i,j} \in [T]$, the above expression can be split into two sums:

$$(2) = \sum_{t=1}^{T} \mathbb{E}\left[\mathbb{I}\{I_t = j\}\,\mathbb{I}\{A^t \cap [i] \neq \emptyset\}\,\mathbb{I}\{T_j(t-1) < \sigma_{i,j}\}\right] + \sum_{t=1}^{T} \mathbb{E}\left[\mathbb{I}\{I_t = j\}\,\mathbb{I}\{A^t \cap [i] \neq \emptyset\}\,\mathbb{I}\{T_j(t-1) \geq \sigma_{i,j}\}\right]\,.$$

Now, define $T^* = \max\left\{t \in [T] \colon \mathbb{I}\{T_j(t-1) < \sigma_{i,j}\} \neq 0\right\}$. Then, by definition, we have

$$\begin{aligned}
\sum_{t=1}^{T} \mathbb{I}\{I_t = j\}\mathbb{I}\{A^t \cap [i] \neq \emptyset\}\mathbb{I}\{T_j(t-1) < \sigma_{i,j}\} &= \sum_{t=1}^{T^*} \mathbb{I}\{I_t = j\}\mathbb{I}\{A^t \cap [i] \neq \emptyset\}\mathbb{I}\{T_j(t-1) < \sigma_{i,j}\} \\
&\leq \sum_{t=1}^{T^*} \mathbb{I}\{I_t = j\} \\
&= T_j(T^*) \leq T_j(T^* - 1) + 1 \leq \sigma_{i,j}\,.
\end{aligned}$$

This shows that, for any $\sigma_{i,j} \in [T]$,

$$(2) \leq \mathbb{E}\left[\sigma_{i,j} + \sum_{t=\sigma_{i,j}+1}^{T} \mathbb{I}\{I_t = j\}\mathbb{I}\{A^t \cap [i] \neq \emptyset\}\mathbb{I}\{T_j(t-1) \geq \sigma_{i,j}\}\right]\,.$$

If expert j is selected at time t, that is $I_t = j$, then it must have the lowest confidence bound: $\widehat{\mu}_j(t-1) - \mathcal{S}_j(t-1) \leq \widehat{\mu}_k(t-1) - \mathcal{S}_k(t-1)$ for all $k \in A^t$. Let $k^* = \operatorname{argmin}_{k \in A^t \cap [i]} \widehat{\mu}_k(t-1) - \mathcal{S}_k(t-1)$, then

$$\begin{aligned}
&\mathbb{E}[\mathbb{I}\{I_t = j\}\mathbb{I}\{A^t \cap [i] \neq \emptyset\}\mathbb{I}\{T_j(t-1) \geq \sigma_{i,j}\}] \\
&\quad \leq \mathbb{P}\left(\widehat{\mu}_j(t-1) - \mathcal{S}_j(t-1) \leq \widehat{\mu}_{k^*}(t-1) - \mathcal{S}_{k^*}(t-1), A^t \cap [i] \neq \emptyset, T_j(t-1) \geq \sigma_{i,j}\right) \ ,
\end{aligned} \tag{3}$$

Next, the first event in the probability can be expressed as follows:

$$\begin{aligned}
&\widehat{\mu}_{k^*}(t-1) - \mathcal{S}_{k^*}(t-1) - \widehat{\mu}_j(t-1) + \mathcal{S}_j(t-1) \geq 0 \\
\Leftrightarrow\ &\widehat{\mu}_{k^*}(t-1) - \mathcal{S}_{k^*}(t-1) - \widehat{\mu}_j(t-1) + \mathcal{S}_j(t-1) - \mathcal{S}_j(t-1) - \mu_{k^*} + \mu_{k^*} - \mu_j + \mu_j - \mu_i + \mu_i + \mathcal{S}_j(t-1) \geq 0 \\
\Leftrightarrow\ &\Big[-\mu_{k^*} + \widehat{\mu}_{k^*}(t-1) - \mathcal{S}_{k^*}(t-1)\Big] + \Big[\mu_j - \widehat{\mu}_j(t-1) - \mathcal{S}_j(t-1)\Big] + \Big[(\mu_{k^*} - \mu_i) - (\mu_j - \mu_i) + 2\mathcal{S}_j(t-1)\Big] \geq 0 \,.
\end{aligned}$$

Thus, for that event to hold, at least one of these three terms must be non-negative. Moreover, if one is non-positive, at least one of the other two is non-negative.

Choose random variable $\sigma_{i,j}$ as follows: $\sigma_{i,j} = \frac{20 \log T}{(\mu_j - \mu_i)^2} \max_{t \in [T]} \frac{T_j(t-1)}{Q_j(t-1)}$. Then, the second event in the probability, $T_j(t-1) \geq \sigma_{i,j}$, implies

$$T_j(t-1) \geq \frac{20 \log T}{(\mu_j - \mu_i)^2} \frac{T_j(t-1)}{Q_j(t-1)} \Rightarrow (\mu_j - \mu_i)^2 \geq \frac{20 \log t}{Q_j(t-1)} \Leftrightarrow \mu_j - \mu_i \geq \sqrt{\frac{20 \log t}{Q_j(t-1)}} \,.$$

In view of that, when the second event $T_j(t-1) \geq \sigma_{i,j}$ holds, we have

$$\begin{aligned}
(\mu_{k^*} - \mu_i) - (\mu_j - \mu_i) + 2\mathcal{S}_j(t-1) &\leq -(\mu_j - \mu_i) + 2\mathcal{S}_{j,t-1} && \text{(def. of } \mu_{k^*}) \\
&= -(\mu_j - \mu_i) + \sqrt{\frac{20 \log t}{Q_j(t-1)}} \leq 0 \,.
\end{aligned}$$

This shows that the third term above is then non-positive and that at least one of the first two is non-negative. Thus, under the above choice of $\sigma_{i,j}$, the following inequality holds:

$$\begin{aligned}
&\mathbb{P}\left(\widehat{\mu}_j(t-1) - \mathcal{S}_j(t-1) \leq \widehat{\mu}_{k^*}(t-1) - \mathcal{S}_{k^*}(t-1), A^t \cap [i] \neq \emptyset, T_j(t-1) \geq \sigma_{i,j}\right) \\
&\quad \leq \mathbb{P}\Big(-\mu_{k^*} + \widehat{\mu}_{k^*}(t-1) - \mathcal{S}_{k^*}(t-1) \geq 0\Big) + \mathbb{P}\Big(\mu_j - \widehat{\mu}_j(t-1) - \mathcal{S}_j(t-1) \geq 0\Big).
\end{aligned} \tag{4}$$

Now, since both the feedback graph G^t and the algorithm's action I_t only depend on information up to time $(t-1)$, it is straightforward to see that, for any $j \in [K]$, the sequence of random variables $L(\xi_j, z_{s_1}), L(\xi_j, z_{s_2}), \ldots$, are i.i.d., and distributed as $L(\xi_j, z_1)$, where s_k is the stopping time $s_k = \min\{t \colon Q_j(t) = k\}$. Using a standard Hoeffding bound, this allows us to bound the second probability in (4) as follows:

$$\begin{aligned}
&\mathbb{P}\Big(\mu_j - \widehat{\mu}_j(t-1) - \mathcal{S}_j(t-1) \geq 0\Big) \\
&\quad = \mathbb{P}\left(\mu_j - \frac{1}{Q(t)} \sum_{s=1}^{t} L(\xi_j, z_s)\mathbb{I}\{j \in N_{I_s}^s\} - \sqrt{\frac{5 \log t}{Q_j(t)}} \geq 0\right) \\
&\quad \leq \sum_{n=1}^{t} \mathbb{P}\left(\mu_j - \frac{1}{n} \sum_{s=1}^{t} L(\xi_j, z_s)\mathbb{I}\{j \in N_{I_s}^t\} - \sqrt{\frac{5 \log t}{n}} \geq 0 \wedge Q_j(t) = n\right) \\
&\quad = \sum_{n=1}^{t} \mathbb{P}\left(\mu_j - \frac{1}{n} \sum_{i=1}^{n} L(\xi_j, z_s) - \sqrt{\frac{5 \log t}{n}} \geq 0\right) \\
&\quad \leq \sum_{n=1}^{t} \frac{1}{t^5} = \frac{1}{t^4} \,.
\end{aligned}$$

The other probability in (4), i.e., $\mathbb{P}(-\mu_{k^*} + \widehat{\mu}_{k^*}(t-1) - \mathcal{S}_{k^*}(t-1) \geq 0)$, can be bounded in a similar way, thereby resulting in the following upper bound:

$$\mathbb{E}[M_{i,j}] \leq \frac{20 \log T}{(\mu_j - \mu_i)^2} \, \mathbb{E}\left[\max_{t \in [T]} \frac{T_j(t-1)}{Q_j(t-1)}\right] + 2 \sum_{t=1}^{T} \frac{1}{t^4} \leq \frac{20 \log T}{(\mu_j - \mu_i)^2} \, \mathbb{E}\left[\max_{t \in [T]} \frac{T_j(t)}{Q_j(t)}\right] + 4 \, .$$

Plugging in the right-hand side of this inequality in (1) to upper-bound $\mathbb{E}[M_{i,j}]$, and using Lemma 1 with $F_j = \mathbb{E}\left[\max_{t\in[T]} \frac{T_j(t)}{Q_j(t)}\right]$ completes the proof. □

B. Lower bound on sleeping regret

This section provides a proof of the lower bound in Theorem 2. The proof of this result follows from extending the arguments in Kleinberg et al. (2008).

Proof. [Theorem 2] We first restate Lemma 11 in Kleinberg et al. (2008).

Lemma 2 (Lemma 11, Kleinberg et al. (2008)) *Suppose we are given two numbers $\mu_1 > \mu_2$, both lying in an interval $[a, b]$ such that $0 < a < b < 1$, and suppose we are given any online algorithm ϕ for the multi-armed bandit problem with two experts which never picks the worse expert more than $o(T^\alpha)$ times for every $\alpha > 0$. Then there is an input instance in the stochastic rewards model, with two experts L and R whose payoff distributions are Bernoulli random variables with means μ_1 and μ_2 or vice-versa, such that for large enough T depending on a, b, μ_1, and μ_2, the regret of algorithm ϕ is $\Omega\left(\frac{\log(T)(\mu_1-\mu_2)}{KL(\mu_2||\mu_1)}\right)$, where the constant inside the $\Omega(\cdot)$ is at least $\frac{1}{2}$.*

Assume that the losses are Bernoulli random variables and that the means $\{\mu_j\}_{j=1}^K$ are bounded away from 0 and 1. Let $A^t \subset [K]$ be the awake set at time t, and suppose that $A^t \sim U(\{2j-1, 2j\}_{j=1}^{K/2})$ independent of the distribution of the losses. For each awake set A and $t \in [T]$, let $s(A, t) \in [T]$ be the time step in which the awake set A occurred for the t-th time. Then we can write for the expected sleeping regret of any algorithm:

$$\begin{aligned}
\mathbb{E}\left[\sum_{t=1}^{T} \mu_{I_t} - \mu_{\sigma(A^t)}\right] &= \mathbb{E}\left[\sum_{t=1}^{T}\sum_{j=1}^{K/2} \mathbb{I}\{A^t = \{2j-1, 2j\}\}(\mu_{I_t} - \mu_{\sigma(A^t)})\right] \\
&= \sum_{j=1}^{K/2} \mathbb{E}\left[\sum_{t=1}^{T} \mathbb{I}\{A^t = \{2j-1, 2j\}\}(\mu_{I_t} - \mu_{\sigma(A^t)})\right] \\
&= \sum_{j=1}^{K/2} \mathbb{E}\left[\sum_{t=1}^{\mathbb{E}[\sum_{t=1}^{T}\mathbb{I}\{A^t=\{2j-1,2j\}\}]} (\mu_{I_{s(\{2j-1,2j\},t)}} - \mu_{\sigma(A^t)})\right] \\
&= \sum_{j=1}^{K/2} \mathbb{E}\left[\sum_{t=1}^{\frac{2T}{K}} (\mu_{I_{s(\{2j-1,2j\},t)}} - \mu_{\sigma(A^t)})\right] \\
&\geq \sum_{j=1}^{K/2} \Omega\left(\frac{\log(2T/K)(\mu_{2j-1} - \mu_{2j})}{\text{KL}(\mu_{2j-1}||\mu_{2j})}\right),
\end{aligned}$$

where the second to last equality follows from Wald's equation, and the inequality follows from applying Lemma 2 to each awake set which is effectively a separate two-armed bandit problem.

Now, since we assume that the means are bounded between a and b, we can upper bound the KL divergence terms as follows:

$$\text{KL}(\mu_{2j-1}||\mu_{2j}) \leq \frac{(\mu_{2j-1} - \mu_{2j})^2}{\mu_{2j}(1 - \mu_{2j})} \leq \frac{(\mu_{2j-1} - \mu_{2j})^2}{\min_{i=1}^K \mu_i(1 - \mu_i)}.$$

Thus, we can write

$$\sum_{j=1}^{K/2} \Omega\left(\frac{\log(2T/K)(\mu_{2j-1} - \mu_{2j})}{\text{KL}(\mu_{2j-1}||\mu_{2j})}\right) \geq \sum_{j=1}^{K/2} \Omega\left(\frac{\log(2T/K)}{\mu_{2j-1} - \mu_{2j}}\right).$$

Similarly, if we consider $A^t \sim U(\{2j, 2j+1\}_{j=1}^{K/2-1})$, then the expected sleeping regret of any algorithm is lower bounded by:

$$\mathbb{E}\left[\sum_{t=1}^{T} \mu_{I_t} - \mu_{\sigma(A^t)}\right] \geq \sum_{j=1}^{K/2-1} \Omega\left(\frac{\log(2T/K)}{\mu_{2j} - \mu_{2j+1}}\right).$$

Thus, if consider an awake set distribution $A^t \sim U(\{2j-1, 2j\}_{j=1}^{K/2})$ and $A^t \sim U(\{2j, 2j+1\}_{j=1}^{K/2-1})$ each with probability 1/2, then the expected sleeping regret of any algorithm is lower bounded by:

$$\mathbb{E}\left[\sum_{t=1}^{T} \mu_{I_t} - \mu_{\sigma(A^t)}\right] \geq \sum_{j=2}^{K} \Omega\left(\frac{\log(2T/K)}{\mu_{j-1} - \mu_j}\right).$$

□

C. UCB-SLG

In this section, we prove Theorem 3.

Proof. [Theorem 3] To simplify the notation, throughout this proof, we replace $L(\cdot, z_t)$ by $L_t(\cdot)$, $\mathbb{E}[\cdot|A_t = \mathtt{A}_k]$ by $\mathbb{E}[\cdot|\mathtt{A}_k]$ and $\nu_{k,i^*(k)}$ by $\nu_{i^*(k)}$. We first decompose the regret in terms of the awake sets $\mathtt{A}_1, \ldots, \mathtt{A}_p$:

$$\mathrm{R}_T^{\text{SLEEP}}(\text{UCB-SLG}) = \sum_{t=1}^{T} \sum_{k=1}^{p} p_k \, \mathbb{E}\left[(L_t(\xi_{I_t}) - L_t(\xi_{i^*(k)}))|\mathtt{A}_k\right] = \sum_{k=1}^{p} p_k \mathrm{R}_{T,k} \ ,$$

where $\mathrm{R}_{T,k} = \sum_{t=1}^{T} \mathbb{E}[(L_t(\xi_{I_t}) - L_t(\xi_{i^*(k)}))|\mathtt{A}_k]$ can be interpreted as the regret for region k at time T. Thus, we can focus on bounding $\mathrm{R}_{T,k}$ for each $k \in [p]$.

Fix $k \in [p]$. Observe that we can disregard any term in $\mathrm{R}_{T,k}$ where the conditional expectation of the chosen expert is less than that of the best expert, $\nu_{k,I_t} \leq \nu_{i^*(k)}$, and bound that by zero:

$$\begin{aligned}
\sum_{t=1}^{T} \mathbb{E}\left[\mathbb{I}\{\nu_{k,I_t} \leq \nu_{i^*(k)}\}(L_t(\xi_{I_t}) - L_t(\xi_{i^*(k)}))|\mathtt{A}_k\right] &\leq \sum_{t=1}^{T} \sum_{i=1}^{K} \mathbb{E}\left[\mathbb{I}\{I_t = i\}\mathbb{I}\{\nu_{k,i} \leq \nu_{i^*(k)}\}(L_t(\xi_i) - L_t(\xi_{i^*(k)}))|\mathtt{A}_k\right] \\
&\leq \sum_{t=1}^{T} \sum_{i=1}^{K} \mathbb{E}\left[\mathbb{I}\{I_t = i\}\mathbb{I}\{\nu_{k,i} \leq \nu_{i^*(k)}\}|\mathtt{A}_k\right](\nu_{k,i} - \nu_{i^*(k)}) \leq 0 \ ,
\end{aligned}$$

where in the second to last inequality, we used the fact that $(L_t(\xi_i) - L_t(\xi_{i^*(k)}))$ and $\mathbb{I}\{I_t = i\}$ are conditionally independent given $\mathtt{A}_k$. Thus, $\mathrm{R}_{T,k}$ can be upper bounded by terms where the conditional expectation of the chosen expert is greater than that of the best expert, $\nu_{k,I_t} > \nu_{i^*(k)}$:

$$\begin{aligned}
\mathrm{R}_{T,k} &\leq \sum_{t=1}^{T} \mathbb{E}\left[\mathbb{I}\{\nu_{k,I_t} > \nu_{i^*(k)}\}(L_t(\xi_{I_t}) - L_t(\xi_{i^*(k)}))|\mathtt{A}_k\right] \\
&= \sum_{t=1}^{T} \sum_{i \in \mathtt{A}_k} \mathbb{E}\left[\mathbb{I}\{I_t = i\}\mathbb{I}\{\nu_{k,i} > \nu_{i^*(k)}\}(L_t(\xi_i) - L_t(\xi_{i^*(k)}))|\mathtt{A}_k\right] && (I_t \text{ must be in } \mathtt{A}_k) \\
&= \sum_{t=1}^{T} \sum_{i \in \mathtt{A}_k} \mathbb{E}\left[\mathbb{I}\{I_t = i\}\mathbb{I}\{\nu_{k,i} > \nu_{i^*(k)}\}|\mathtt{A}_k\right](\nu_{k,i} - \nu_{i^*(k)}), && \text{(cond. indep.)} \\
&= \sum_{t=1}^{T} \sum_{i \in \mathtt{A}_k \setminus \mathtt{B}_k} \mathbb{E}\left[\mathbb{I}\{I_t = i\}\mathbb{I}\{\nu_{k,i} > \nu_{i^*(k)}\}|\mathtt{A}_k\right]\overline{\Delta}_{k,i} && (\text{def. of } \mathtt{B}_k \text{ and } \overline{\Delta}_{k,i}) \\
&= \sum_{i \in \mathtt{A}_k \setminus \mathtt{B}_k} \overline{\Delta}_{k,i}(r^1_{T,k,i} + r^2_{T,k,i}) \ ,
\end{aligned}$$

with $r^1_{T,k,i}$ and $r^2_{T,k,i}$ defined by

$$\begin{aligned}
r^1_{T,k,i} &= \sum_{t=1}^{T} \mathbb{E}\left[\mathbb{I}\{I_t = i\}\mathbb{I}\{\nu_{k,i} > \nu_{i^*(k)}\}\mathbb{I}\{O_{k,i}(t-1) < s_i\}|\mathtt{A}_k\right] \\
r^2_{T,k,i} &= \sum_{t=1}^{T} \mathbb{E}\left[\mathbb{I}\{I_t = i\}\mathbb{I}\{\nu_{k,i} > \nu_{i^*(k)}\}\mathbb{I}\{O_{k,i}(t-1) \geq s_i\}|\mathtt{A}_k\right] ,
\end{aligned}$$

where s_i is a parameter whose value will be selected later. Since the event $I_t = i$ implies in particular the inequality $\widehat{\nu}_{k,i}(t-1) - S_{k,i}(t-1) \leq \widehat{\nu}_{k,i^*(k)}(t-1) - S_{k,i^*(k)}(t-1)$, we have

$$r^2_{T,k,i} \leq \sum_{t=1}^{T} \mathbb{P}\left[\widehat{\nu}_{k,i}(t-1) - S_{k,i}(t-1) \leq \widehat{\nu}_{k,i^*(k)}(t-1) - S_{k,i^*(k)}(t-1) \wedge \nu_{k,i} > \nu_{i^*(k)} \wedge O_{k,i}(t-1) \geq s_i \middle| \mathtt{A}_k\right] .$$

The inequality defining the first event in this probability can be decomposed as follows:

$$\begin{aligned} &\widehat{\nu}_{k,i}(t-1) - S_{k,i}(t-1) \leq \widehat{\nu}_{k,i^*(k)}(t-1) - S_{k,i^*(k)}(t-1) \\ \Leftrightarrow\ & 0 \leq \Big[-\nu_{i^*(k)} + \widehat{\nu}_{k,i^*(k)}(t-1) - S_{k,i^*(k)}(t-1) \Big] + \Big[\nu_{k,i} - \widehat{\nu}_{k,i}(t-1) - S_{k,i}(t-1) \Big] \\ &\quad + \Big[\nu_{i^*(k)} - \nu_{k,i} + 2S_{k,i}(t-1) \Big]. \end{aligned}$$

Thus, if we choose s_i such that the third term be non-positive, this will imply that one of the first two terms at least is non-negative.

Let s_i be defined by $s_i = \frac{20\log(T)}{\Delta_{k,i}^2}$. Then, $O_{k,i}(t-1) \geq s_i$ implies $\nu_{i^*(k)} - \nu_{k,i} + 2S_{k,i}(t-1) \leq 0$, that is the non-positivity of the third term. Thus, with this choice of s_i, if the inequality defining the first event in the probability holds, at least one of the first two terms above must be non-negative. In view of that, by the union bound and Hoeffding's inequality applied to the probability of each event, the following holds:

$$\begin{aligned} r_{T,k,i}^2 &\leq \sum_{t=1}^{T} \mathbb{P}\left[0 \leq -\nu_{i^*(k)} + \widehat{\nu}_{k,i^*(k)}(t-1) - S_{k,i^*(k)}(t-1) | \mathtt{A}_k \right] + \sum_{t=1}^{T} \mathbb{P}\left[0 \leq \nu_{k,i} - \widehat{\nu}_{k,i}(t-1) - S_{k,i}(t-1) \Big| \mathtt{A}_k \right] \\ &\leq 2\sum_{t=1}^{T} \frac{1}{t^4} \leq 4. \end{aligned}$$

Thus, this implies the inequality $\sum_{i\in \mathtt{A}_k \setminus \mathtt{B}_k} r_{T,k,i}^2 \leq 4\big|\mathtt{A}_k \setminus \mathtt{B}_k\big|$. To bound $\sum_{i\in \mathtt{A}_k \setminus \mathtt{B}_k} r_{T,k,i}^1$, we will use the clique covering $\mathcal{C}_k$ defined in Section 4. Since $\mathcal{C}_k$ is a cover of the graph G_k, we can decompose the expression involving $r_{T,k,i}^1$ in terms of the components of the clique cover and write

$$\sum_{i\in \mathtt{A}_k\setminus \mathtt{B}_k} \overline{\Delta}_{k,i}\, r_{T,k,i}^1 \leq \sum_{t=1}^{T} \sum_{C\in\mathcal{C}_k} \sum_{i\in C\setminus \mathtt{B}_k} \mathbb{E}[\overline{\Delta}_{k,i} \mathbb{I}\{I_t = i\} \mathbb{I}\{\nu_{k,i} > \nu_{i^*(k)}\} \mathbb{I}\{O_{k,i}(t-1) < s_i\} | \mathtt{A}_k].$$

Let $O_{k,C}(t-1)$ denote the number of times any expert in clique C has been played up to time $t-1$. Since experts in the same clique are observed together, $O_{k,C}(t-1)$ is less than or equal to the number of times an expert $i \in C$ is observed: $O_{k,C}(t-1) \leq O_{k,i}(t-1)$. Thus, we can upper bound the expression above by replacing $O_{k,i}(t-1)$ with $O_{k,C}(t-1)$ as follows:

$$\begin{aligned} \sum_{i\in \mathtt{A}_k\setminus \mathtt{B}_k} \overline{\Delta}_{k,i}\, r_{T,k,i}^1 &\leq \sum_{t=1}^{T} \sum_{C\in\mathcal{C}_k} \sum_{i\in C\setminus \mathtt{B}_k} \mathbb{E}\left[\overline{\Delta}_{k,i} \mathbb{I}\{I_t = i\} \mathbb{I}\{\nu_{k,i} > \nu_{i^*(k)}\} \mathbb{I}\Big\{O_{k,C}(t-1) < s_i\Big\} \Big| \mathtt{A}_k \right] \\ &\leq \sum_{t=1}^{T} \sum_{C\in\mathcal{C}_k} \sum_{i\in C\setminus \mathtt{B}_k} \mathbb{E}\left[\overline{\Delta}_{k,i} \mathbb{I}\{I_t = i\} \mathbb{I}\{\nu_{k,i} > \nu_{i^*(k)}\} \mathbb{I}\Big\{O_{k,C}(t-1) < \max_{i\in C\setminus \mathtt{B}_k} s_i\Big\} \Big| \mathtt{A}_k \right] \\ &\leq \sum_{C\in\mathcal{C}_k} \Big(\max_{i\in C\setminus \mathtt{B}_k} \overline{\Delta}_{k,i} \Big) \sum_{t=1}^{T} \sum_{i\in C\setminus \mathtt{B}_k} \mathbb{E}\left[\mathbb{I}\{I_t = i\} \mathbb{I}\Big\{\nu_{k,i} > \nu_{i^*(k)}\Big\} \mathbb{I}\Big\{O_{k,C}(t-1) < \max_{i\in C\setminus \mathtt{B}_k} s_i\Big\} \Big| \mathtt{A}_k \right] \\ &\leq \sum_{C\in\mathcal{C}_k} \Big(\max_{i\in C\setminus \mathtt{B}_k} \overline{\Delta}_{k,i} \Big) \sum_{t=1}^{T} \sum_{i\in C\setminus \mathtt{B}_k} \mathbb{E}\left[\mathbb{I}\{I_t = i\} \mathbb{I}\Big\{O_{k,C}(t-1) < \max_{i\in C\setminus \mathtt{B}_k} s_i\Big\} \Big| \mathtt{A}_k \right] \\ &\leq \sum_{C\in\mathcal{C}_k} \Big(\max_{i\in C\setminus \mathtt{B}_k} \overline{\Delta}_{k,i} \Big) \sum_{t=1}^{T} \mathbb{E}\left[\mathbb{I}\{I_t \in C\} \mathbb{I}\Big\{O_{k,C}(t-1) < \max_{i\in C\setminus \mathtt{B}_k} s_i\Big\} \Big| \mathtt{A}_k \right]. \end{aligned}$$

Define s and t^* by $s = \max_{i\in C\setminus \mathtt{B}_k} s_i$ and $t^* = \max\Big\{t \leq T \colon \mathbb{I}\{O_{k,C}(t-1) < s\} \neq 0\Big\}$. Then, we have

$$\sum_{t=1}^{T} \mathbb{I}\{I_t \in C\} \mathbb{I}\Big\{O_{k,C}(t-1) < s\Big\} = \sum_{t=1}^{t^*} \mathbb{I}\{I_t \in C\} \mathbb{I}\Big\{O_{k,C}(t-1) < s\Big\} \leq s,$$

where the last inequality holds since, by definition of t^*, the number of non-zero terms in the last sum it at most s. Thus, we have $\sum_{t=1}^{T} \mathbb{E}[\mathbb{I}\{I_t \in C\}\mathbb{I}\{O_{k,C}(t-1) < s\} | \mathtt{A}_k] \leq s$ and

$$\sum_{i \in \mathtt{A}_k \setminus \mathtt{B}_k} \overline{\Delta}_{k,i}\, r^1_{T,k,i} \leq \sum_{C \in \mathcal{C}_k} \Big(\max_{i \in C \setminus \mathtt{B}_k} \overline{\Delta}_{k,i} \Big)\Big(\max_{i \in C \setminus \mathtt{B}_k} s_i \Big) = 20 \sum_{C \in \mathcal{C}_k} \frac{\max_{i \in C \setminus \mathtt{B}_k} \overline{\Delta}_{k,i}}{\min_{i \in C \setminus \mathtt{B}_k} \overline{\Delta}^2_{k,i}} \log(T),$$

for any clique covering $\mathcal{C}_k$. Combining this inequality with the one for $r^2_{T,k,i}$ gives:

$$\begin{aligned} \mathrm{R}^{\mathrm{SLEEP}}_T(\mathtt{UCB\text{-}SLG}) = \sum_{k=1}^{p} p_k \mathrm{R}_{T,k} &\leq \sum_{k=1}^{p} p_k \sum_{i \in \mathtt{A}_k \setminus \mathtt{B}_k} \overline{\Delta}_{k,i}(r^1_{T,k,i} + r^2_{T,k,i}) \\ &\leq \sum_{k=1}^{p} p_k \left(20 \sum_{C \in \mathcal{C}_k} \frac{\max_{i \in C \setminus \mathtt{B}_k} \overline{\Delta}_{k,i}}{\min_{i \in C \setminus \mathtt{B}_k} \overline{\Delta}^2_{k,i}} \log(T) + 4|\mathtt{A}_k \setminus \mathtt{B}_k| \right). \end{aligned}$$

Taking the minimum of the right-hand side over all possible clique covering $\mathcal{C}_k$ completes the proof. □

D. UCB-ABS

In this section, we prove Theorem 4.

Proof. [Theorem 4] To alleviate notation, throughout this proof we replace $L(\cdot, z_t)$ by $L_t(\cdot)$, $\mathbb{E}[\cdot|A^t = \mathtt{A}_k]$ by $\mathbb{E}[\cdot|\mathtt{A}_k]$, and $\nu_{k,i^*(k)}$ by $\nu_{i^*(k)}$. By the same reasoning as in the proof of Theorem 3, for each $k \in [p]$ the following holds:

$$\mathrm{R}_{T,k} \leq \sum_{t=1}^{T} \mathbb{E}[\mathbb{I}\{\nu_{k,I_t} > \nu_{i^*(k)}\}(L_t(\xi_{I_t}) - L_t(\xi_{i^*(k)}))|\mathtt{A}_k] \; , \tag{5}$$

where $\mathrm{R}_{T,k}$ is defined as in that proof.

Next, in order to bound (5), we split the rounds $t \in [T]$ into three cases that need to be dealt with separately:

1. $\nu_{i^*(k)} \neq c$ and round t is such that $I_t \neq 0$;
2. $\nu_{i^*(k)} = c$ and round t is such that $I_t \neq 0$;
3. Round t is such that $I_t = 0$.

In Case 1, the algorithm will pick an expert that is not $i^*(k)$ if there exists an expert $i \neq 0$ that satisfies $\widehat{\nu}_{k,i}(t-1) \leq \widehat{\nu}_{i^*(k)}(t-1)$. We will use a Follow-The-Leader type argument based on Lemma 1 of Caron et al. (2012). On the other hand, in Case 2, the algorithm will pick an expert that is not $i^*(k)$ if there exists an expert $i \neq 0$ that satisfies $\widehat{\nu}_{k,i}(t-1) - S_k(t-1) \leq c$. We will use a UCB-type argument. Finally, for Case 3, it must be that $c \leq \widehat{\nu}_{i^*(k)}(t-1) - S_k(t-1)$, and we will show that the overall contribution to the regret can be upper bounded by a constant, independent of time horizon T.

Case 1. Since $L_t(\xi_i) - L_t(\xi_{i^*(k)})$ and $\mathbb{I}\{I_t = i\}$ are conditionally independent given[3] $\mathtt{A}_k$, we can decompose the expectation in (5):

$$\sum_{t=1}^{T} \sum_{i \in [K] \setminus (\{0\} \cup \mathtt{B}_k)} \mathbb{E}\Big[\mathbb{I}\{I_t = i\}\mathbb{I}\{\nu_{k,i} > \nu_{i^*(k)}\}|\mathtt{A}_k\Big](\nu_{k,i} - \nu_{i^*(k)})$$

and focus on bounding the number of times each arm $i \in [K] \setminus (\{0\} \cup \mathtt{B}_k)$ was pulled. Similarly to the proof of Theorem 3, we introduce the conditions $O_k(t-1) > s_i$ and $O_k(t-1) < s_i$ for some s_i to be chosen later:

$$\begin{aligned}
&\sum_{t=1}^{T} \mathbb{E}\Big[\mathbb{I}\{I_t = i\}\mathbb{I}\{i \neq \{0\}\}\mathbb{I}\{\nu_{k,i} > \nu_{i^*(k)}\}|\mathtt{A}_k\Big] \\
&= \sum_{t=1}^{T} \mathbb{E}\Big[\mathbb{I}\{I_t = i\}\mathbb{I}\{i \neq \{0\}\}\mathbb{I}\{O_k(t-1) < s_i\}\mathbb{I}\{\nu_{k,i} > \nu_{i^*(k)}\}|\mathtt{A}_k\Big] && (6) \\
&+ \sum_{t=1}^{T} \mathbb{E}\Big[\mathbb{I}\{I_t = i\}\mathbb{I}\{i \neq \{0\}\}\mathbb{I}\{O_k(t-1) > s_i\}\mathbb{I}\{\nu_{k,i} > \nu_{i^*(k)}\}|\mathtt{A}_k\Big] && (7)
\end{aligned}$$

and we bound (7) by a constant that is independent of T. If expert $I_t = i$ where $i \neq \{0\}$ is chosen, then it must be that $\widehat{\nu}_{k,i}(t-1) \leq \widehat{\nu}_{i^*(k)}(t-1)$. Hence,

$$(7) \leq \sum_{t=1}^{T} \mathbb{P}[\widehat{\nu}_{k,i}(t-1) \leq \widehat{\nu}_{i^*(k)}(t-1), i \neq \{0\}, O_k(t-1) > s_i, \nu_{k,i} > \nu_{i^*(k)}|\mathtt{A}_k\Big] \; .$$

We then use Lemma 1 of Caron et al. (2012) to show that the empirical mean of the chosen expert cannot be less than that of the best expert $i^*(k)$ too often. This gives us

$$(7) \leq \sum_{t=1}^{T} 2e^{-s_i(\nu_{k,i} - \nu_{i^*(k)})^2/2} \; .$$

[3] Recall that $\mathtt{A}_k$ is determined by x_t.

Note that this lemma applies here because the loss observations are i.i.d. given $A^t = \mathtt{A}_k$ and since $O_k(t-1) > s_i$, we saw at least s_i observations of the losses. We then choose $s_i = \frac{2\log(T\Delta_{k,i}^2)}{\Delta_{k,i}^2}$, so that $\sum_{t=1}^T 2e^{-s_i(\nu_{k,i}-\nu_{i^*(k)})^2/2} = O(1)$. Lastly, the bound on (6) follows by the same covering argument as in the proof of Theorem 3.

Case 2. By a similar reasoning as in Case 1, the regret is bounded as follows :

$$\sum_{t=1}^{T} \sum_{i\in[K]\setminus(\{0\}\cup\mathtt{B}_k)} \mathbb{E}\Big[\mathbb{I}\{I_t = i\}\mathbb{I}\{\nu_{k,i} > c\}|\mathtt{A}_k\Big](\nu_{k,i} - c) .$$

Again, for each expert $i \in [K]\setminus(\{0\}\cup\mathtt{B}_k)$, we split this sum according to the conditions $O_k(t-1) > s_i$ and $O_k(t-1) < s_i$ for some s_i to be chosen later. That is,

$$\begin{aligned}
&\sum_{t=1}^{T}\mathbb{E}\Big[\mathbb{I}\{I_t = i\}\mathbb{I}\{i \neq \{0\}\}\mathbb{I}\{\nu_{k,i} > c\}|\mathtt{A}_k\Big] \\
&= \sum_{t=1}^{T}\mathbb{E}\Big[\mathbb{I}\{I_t = i\}\mathbb{I}\{i \neq \{0\}\}\mathbb{I}\{O_k(t-1) < s_i\}\mathbb{I}\{\nu_{k,i} > c\}|\mathtt{A}_k\Big] &(8)\\
&+ \sum_{t=1}^{T}\mathbb{E}\Big[\mathbb{I}\{I_t = i\}\mathbb{I}\{i \neq \{0\}\}\mathbb{I}\{O_k(t-1) > s_i\}\mathbb{I}\{\nu_{k,i} > c\}|\mathtt{A}_k\Big] . &(9)
\end{aligned}$$

To bound term (9), since $\nu_{i^*(k)} = c$, if $I_t = i$ was the chosen expert, then it must be that $\widehat{\nu}_{k,i}(t-1) - S_k(t-1) \leq c$. Thus,

$$\begin{aligned}
(9) &\leq \sum_{t=1}^{T}\mathbb{P}\left(\widehat{\nu}_{k,i}(t-1) - S_k(t-1) \leq c, c < \nu_{k,i}, O_k(t-1) > s_i \Big| \mathtt{A}_k\right) \\
&= \sum_{t=1}^{T}\mathbb{P}\left(0 < -\widehat{\nu}_{k,i}(t-1) + S_k(t-1) + c + \nu_{k,i} - \nu_{k,i} + 2S_k(t-1) - 2S_k(t-1), c < \nu_{k,i}, O_k(t-1) > s_i \Big| \mathtt{A}_k\right) \\
&= \sum_{t=1}^{T}\mathbb{P}\left(0 < \nu_{k,i} - \widehat{\nu}_{k,i}(t-1) - S_k(t-1) + c - \nu_{k,i} + 2S_k(t-1), c < \nu_{k,i}, O_k(t-1) > s_i \Big| \mathtt{A}_k\right) ,
\end{aligned}$$

where as in the proof of Theorem 3, we introduced the terms $\nu_{k,i}$ and $S_k(t-1)$. By choosing $s_i = \frac{20\log T}{(\nu_{k,i}-c)^2}$, then the condition $O_k(t-1) > s_i$ implies that $c - \nu_{k,i} + 2S_k(t-1) \leq 0$. This in turn implies that $0 < \nu_{k,i} - \widehat{\nu}_{k,i}(t-1) - S_k(t-1)$, and we bound the probability of this latter event by using a union bound and Hoeffding's inequality:

$$\sum_{t=1}^{T}\mathbb{P}[0 < \nu_{k,i} - \widehat{\nu}_{k,i}(t-1) - S_k(t-1)] \leq \sum_{t=1}^{T}\sum_{s=1}^{t}\frac{1}{t^5} \leq \sum_{t=1}^{T}\frac{1}{t^4} \leq 2 .$$

Again, the bound on (8) follows directly by the covering argument in the proof of Theorem 3.

Case 3. Consider the rounds t where the chosen expert is the all-abstain expert, $(I_t = 0)$ (and where $I_t \notin \mathtt{B}_k$). By the same reasoning as in the previous two cases, the regret in this case can be bounded as follows:

$$(5) \leq \sum_{t=1}^{T}\mathbb{E}[\mathbb{I}\{\nu_{k,I_t} > \nu_{i^*(k)}\}\mathbb{I}\{I_t = 0\}|\mathtt{A}_k](c - \nu_{i^*(k)}) .$$

If the all-abstain expert is chosen at time t, then it must be that $c \leq \widehat{\nu}_{i^*(k)}(t-1) - S_k(t-1)$. Hence,

$$\begin{aligned}
&\sum_{t=1}^{T}\mathbb{E}\Big[\mathbb{I}\{\nu_{k,I_t} > \nu_{i^*(k)}\}\mathbb{I}\{I_t = 0\}|\mathtt{A}_k\Big] \\
&\quad \leq \sum_{t=1}^{T}\mathbb{P}\Big(c \leq \widehat{\nu}_{k,i^*(k)}(t-1) - S_k(t-1), c > \nu_{i^*(k)}\Big|\mathtt{A}_k\Big) . &(10)
\end{aligned}$$

By following a similar logic as in proof of Theorem 3, we then introduce $\nu_{i^*(k)}$ and use the fact that $c > \nu_{i^*(k)}$:

$$\begin{aligned}(10) &\leq \sum_{t=1}^{T} \mathbb{P}\Big(0 \leq -\nu_{i^*(k)} + \widehat{\nu}_{k,i^*(k)}(t-1) - S_k(t-1) + \nu_{i^*(k)} - c, c > \nu_{i^*(k)} \Big| \mathtt{A}_k\Big) \\ &\leq \sum_{t=1}^{T} \mathbb{P}\Big(0 \leq -\nu_{i^*(k)} + \widehat{\nu}_{k,i^*(k)}(t-1) - S_k(t-1) \Big| \mathtt{A}_k\Big) \\ &\leq \sum_{t=1}^{T}\sum_{s=1}^{t} \frac{1}{t^5} \leq \sum_{t=1}^{T} \frac{1}{t^4} \leq 2\,,\end{aligned}$$

where in the third-last inequality we used a union bound in conjunction with Hoeffding's inequality.

Combining the inequalities corresponding to three cases above completes the proof. □

E. Further Experimental Results

In this section, we present further experimental results testing different aspects of our problem. The first set of figures present the experimental results of all datasets using the same experimental setup as Cortes et al. (2018). Figure 4 and Figure 5 show the results for all abstention costs for two datasets eye and HIGGS. These results show that UCB-ABS outperform UCB, UCB-NT, and UCB-GT on most datasets, and approaches the performance of FS. Even though the experiments were carried out for all abstention costs, to simplify exposition, we show the results for the rest of the datasets for abstention costs in $\{0.05, 0.5, 0.95\}$ in Figure 6 and Figure 7. Figure 8, Figure 9, and Figure 10 show the fraction of abstained points for each algorithm for different abstention costs. As expected, all algorithms tend to abstain more often when the cost of abstention is smaller. Lastly, we increased the number of predictions functions from 100 to 200 hyperplanes and increased the number of abstention regions from 21 to 41. We find that the performance of all algorithms improves slightly on some datasets. Figure 11 shows the results of these new experiments for the same set of datasets and abstention costs as in the main part of the paper.

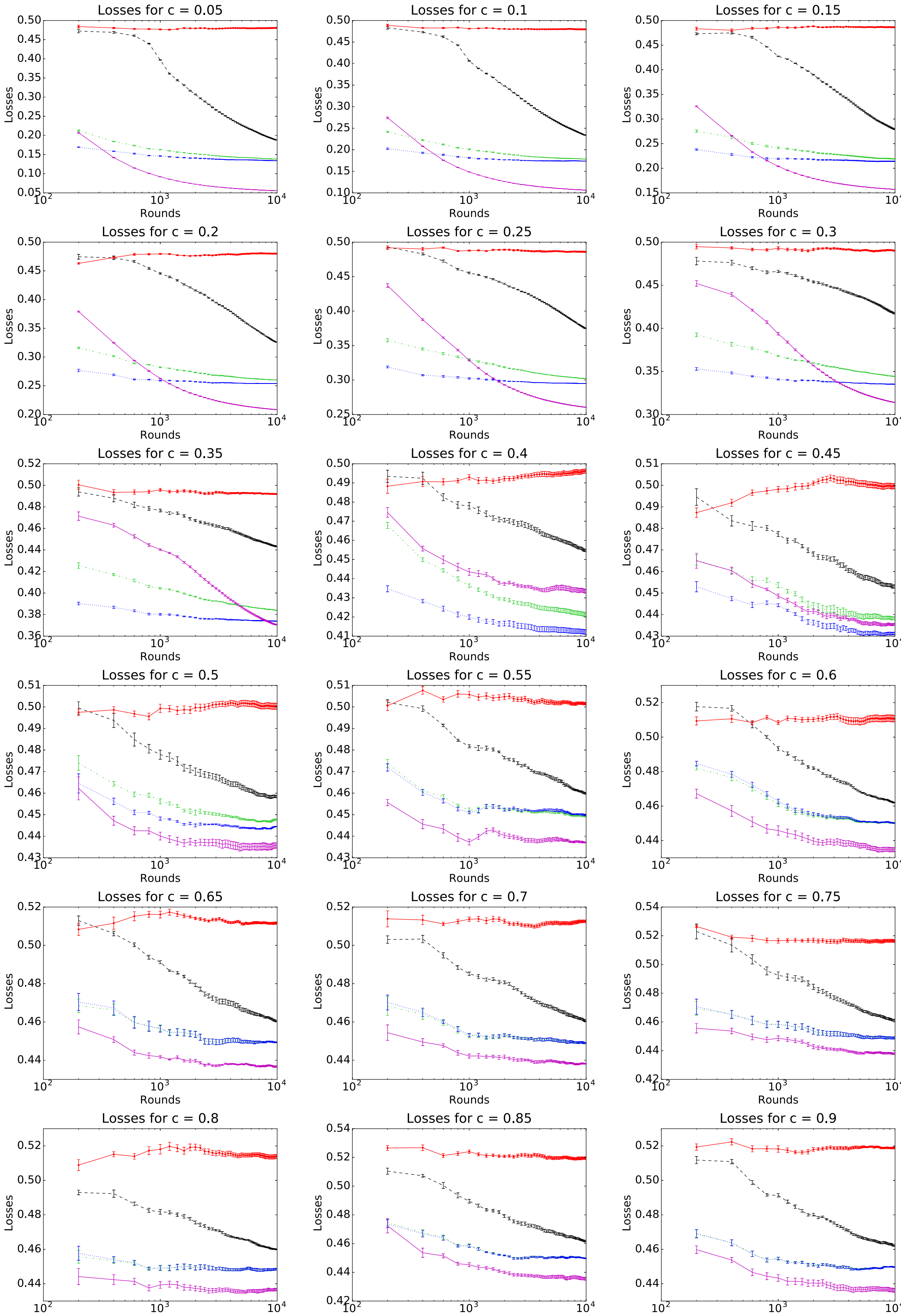

Figure 4: A graph of the averaged loss with standard deviations as a function of t (log scale) for UCB-ABS, UCB-GT, UCB-NT, UCB, and FS. The results are for the `eye` dataset.

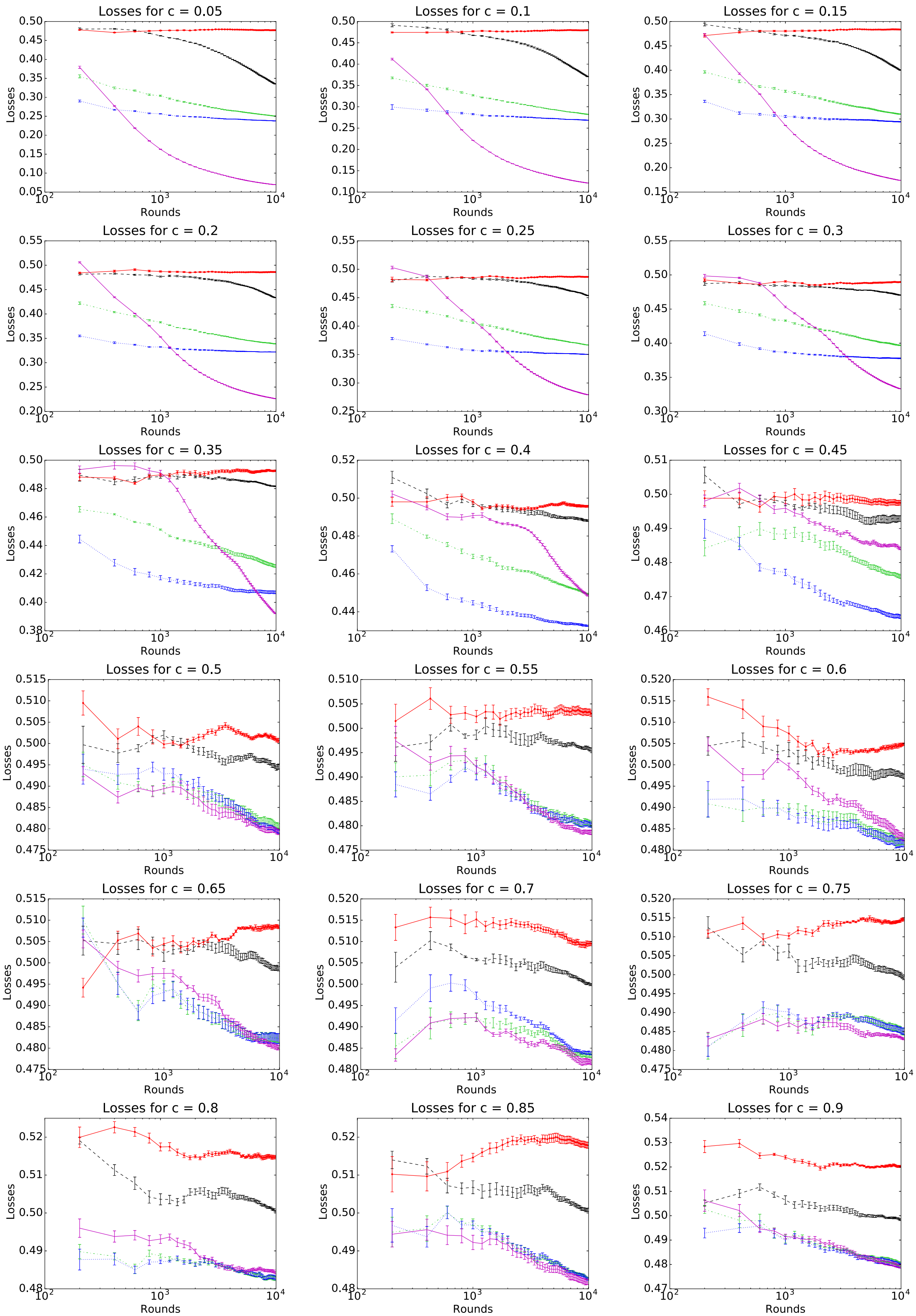

Figure 5: A graph of the averaged loss with standard deviations as a function of t (log scale) for UCB-ABS, UCB-GT, UCB-NT, UCB, and FS. The results are for the HIGGS dataset.

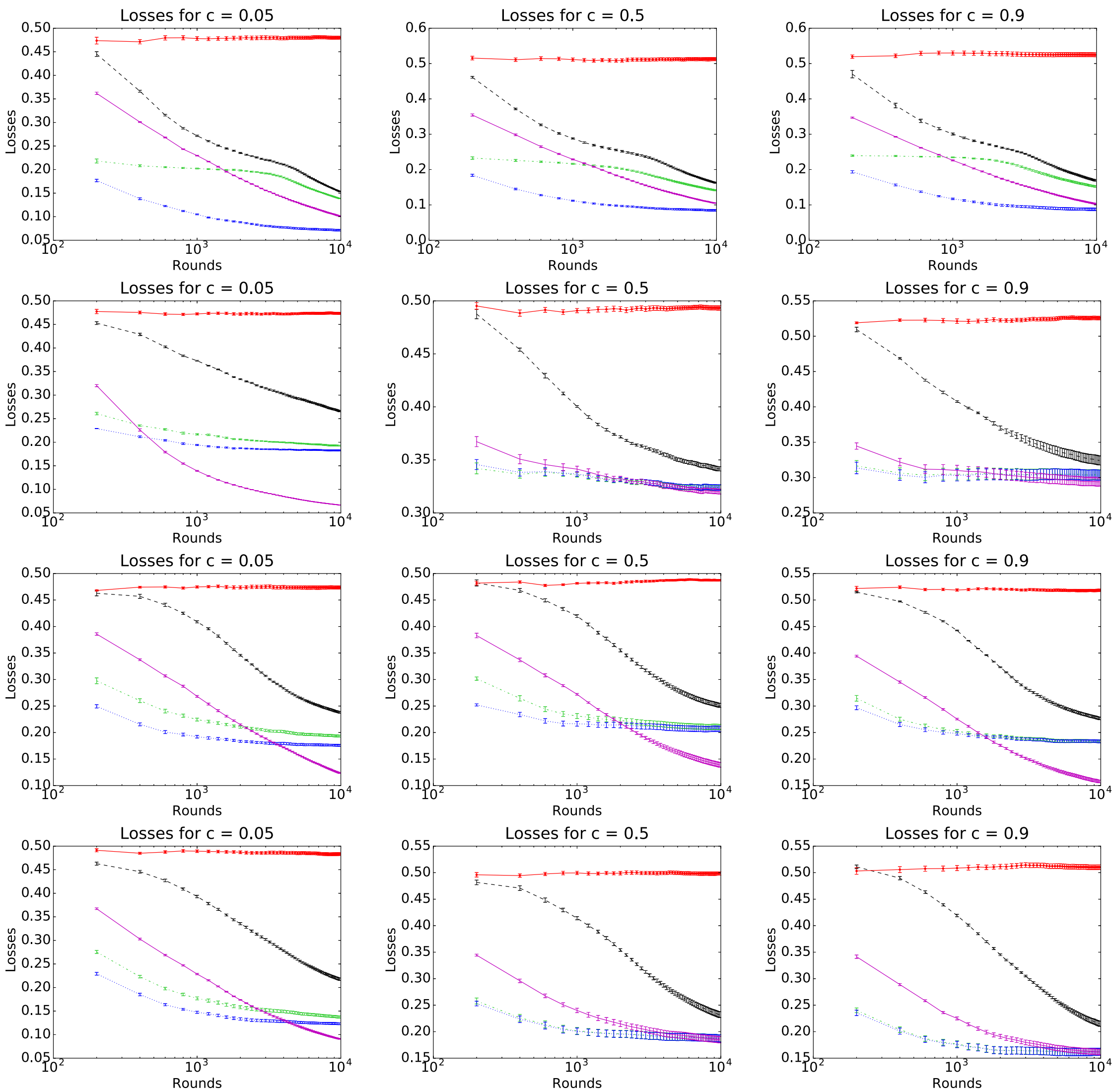

Figure 6: A graph of the averaged loss with standard deviations as a function of t (log scale) for UCB-ABS, UCB-GT, UCB-NT, UCB, and FS. The results are for the skin, cod-rna, guide, ijcnn dataset.

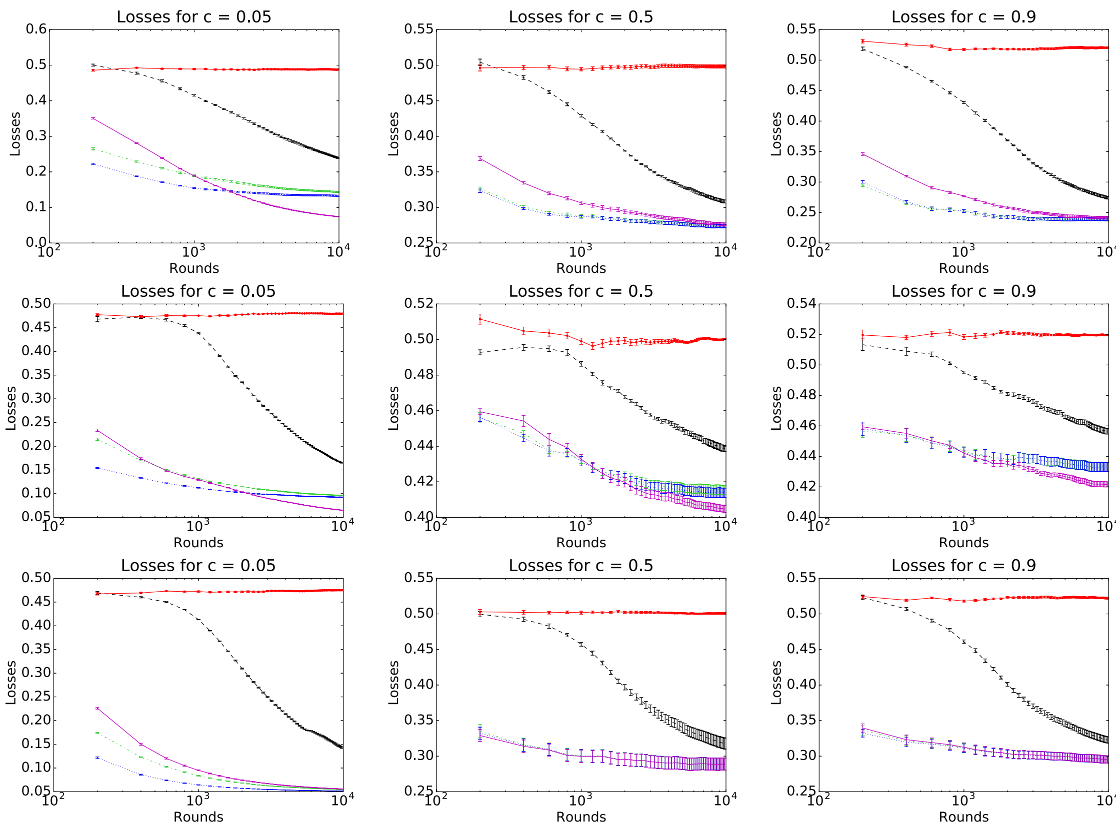

Figure 7: A graph of the averaged loss with standard deviations as a function of t (log scale) for UCB-ABS, UCB-GT, UCB-NT, UCB, and FS. The results are for the CIFAR, covtype, and phish dataset.

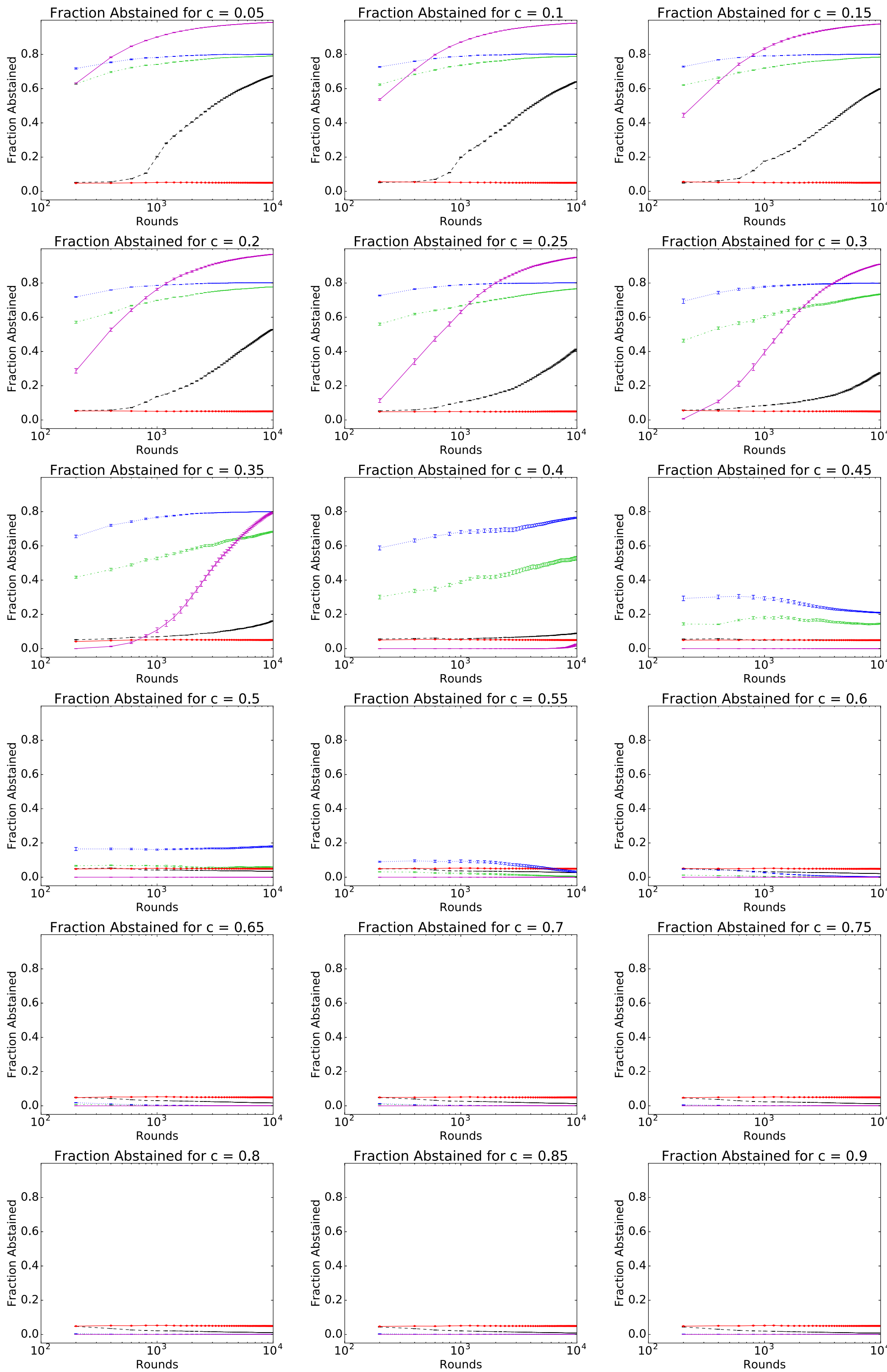

Figure 8: A graph of the fraction abstained with standard deviations as a function of t (log scale) for UCB-ABS, UCB-GT, UCB-NT, UCB, and FS. The results are for the eye dataset.

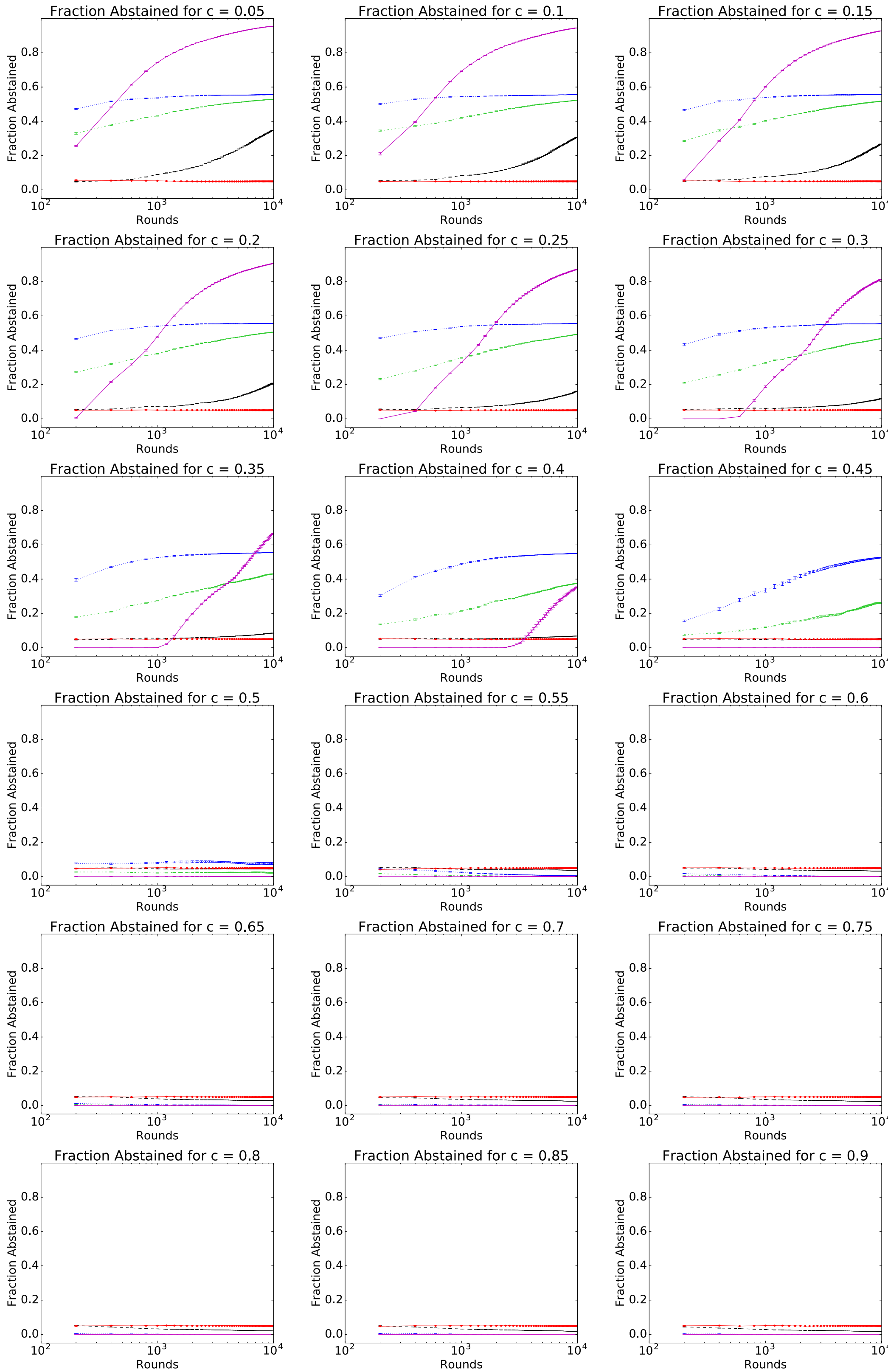

Figure 9: A graph of the fraction abstained with standard deviations as a function of t (log scale) for UCB-ABS, UCB-GT, UCB-NT, UCB, and FS. The results are for the HIGGS dataset.

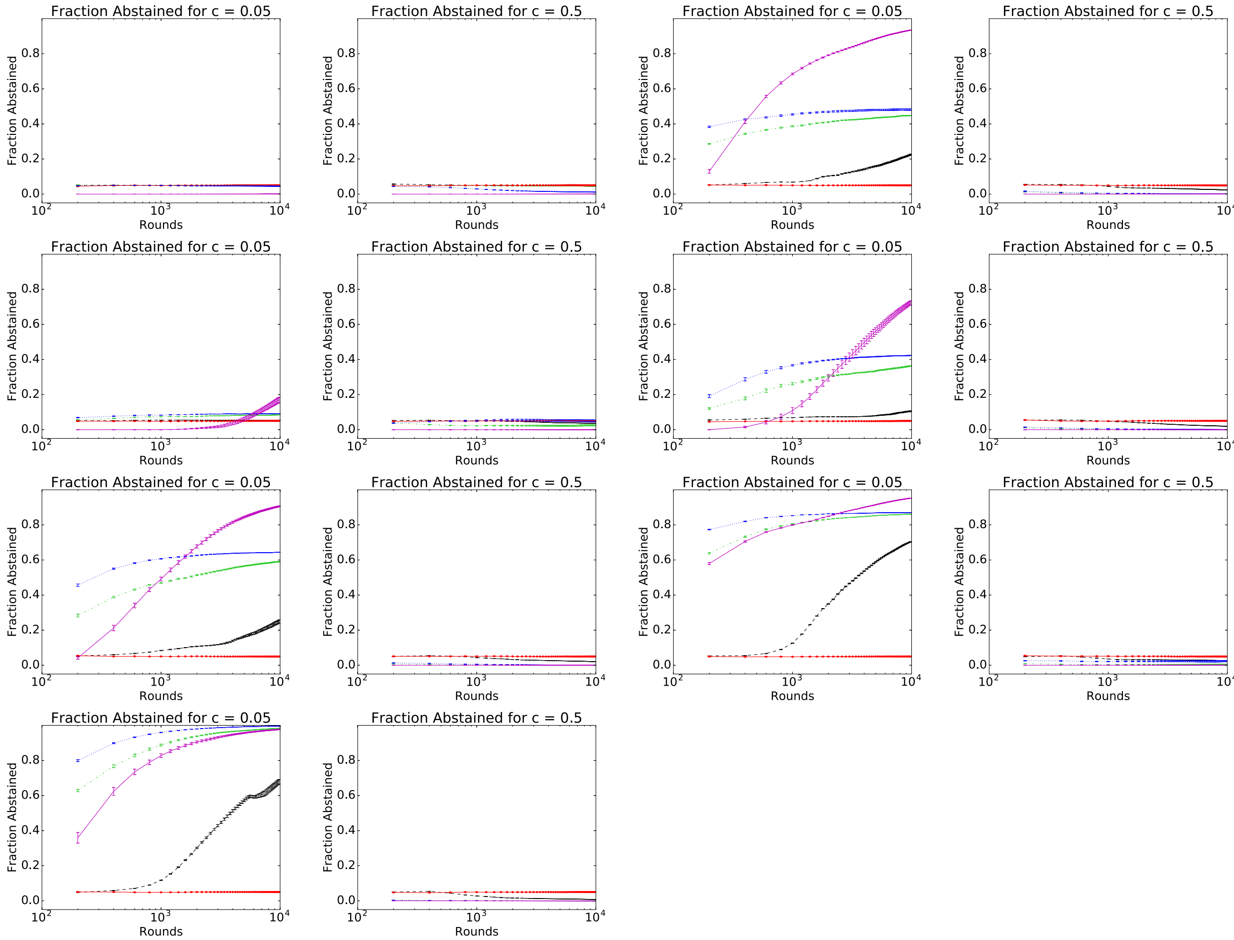

Figure 10: A graph of the fraction abstained with standard deviations as a function of t (log scale) for UCB-ABS, UCB-GT, UCB-NT, UCB, and FS. We show the results for two abstention costs for each dataset. Starting from the top left, the plots are for the skin , cod-rna, guide, ijcnn, CIFAR, covtype and phish datasets.

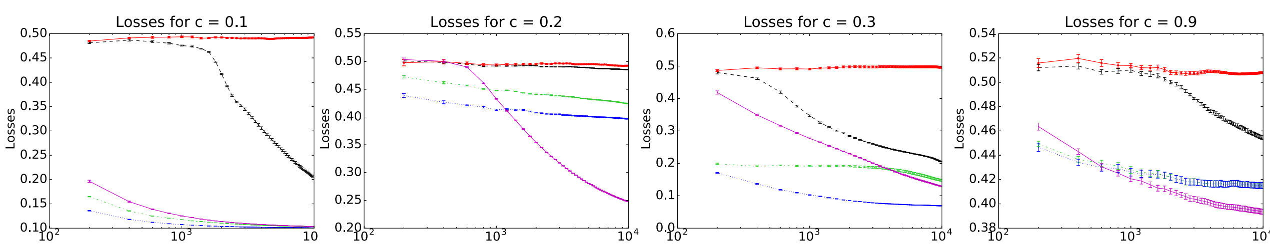

Figure 11: In this set of experiments, we increased the number of abstention functions from 21 to 41 and the number of hyperplanes from 100 to 200. The figures shows a graph of the averaged loss with standard deviations as a function of t (log scale). The algorithms we tested include UCB-ABS, UCB-GT, UCB-NT, UCB, and FS. Starting from the left, the datasets are as follows: eye, HIGGS , skin, and covtype.