We thank all the reviewers for their helpful and constructive comments. Here we respond to the major concerns. We will fix the minor issues in the new version of the paper.

On Remark 2 and the statement about optimality (R1 and R2). We agree with R2 that λ_i's depend on α, and Remark 2 only says that the LSA algorithm is instance-wise asymptotically optimal when α is chosen properly, e.g. $1/20$. We did not prove the optimality for $\alpha > 1/20$, due to the factor $1/10$ in the upper bound of Theorem 1, which arises from several places in the proof details, and is also responsible for the large constants in Remark 2. However, we believe that this is only because of the limitation of our current proof techniques, and the algorithm might be asymptotically optimal with much smaller constants for larger α, which we leave as a future work. Empirical evaluation suggests that the algorithm performs very well when $\alpha \in [1.15, 1.55]$. We will add this clarification and revise the exposition of the statements following R2's suggestion.

On comparison of LSA and APT (R1, R2, and R3). $\mathrm{APT}(\epsilon = 0)$ may be intrigued by a hard arm (as it has to correctly label all arms to achieve a small simple regret) and waste many samples, and cannot achieve the optimal sample efficiency for the aggregate regret. On Line 60, we also argued that APT with $\epsilon > 0$ cannot achieve the optimal aggregate regret. Indeed, in our experiments, $\mathrm{APT}(\epsilon = 0)$ performs consistently worse than LSA in all settings.

On the other hand, a simple corollary of our main theorem shows that, for the simple regret *with no* indifference zone ($\epsilon = 0$), our LSA achieves optimality up to a $\ln K$ factor in the budget T. Note that the optimality definition in (Locatelli et al., 2016) also allows logarithmic slacks of T. In contrast, in our paper, we aim for the strict asymptotic optimality up to only constant factors. We leave proving the strict asymptotic optimality and generalizing LSA to achieve the optimal simple regret *with* indifference zone as an open direction for future research. We attach new experimental results (Figure 1) to show that LSA performs better than $\mathrm{APT}(0)$ for the simple regret with no indifference zone.

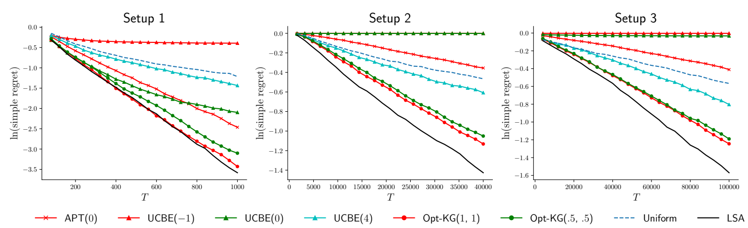

Figure 1: Simple regret on a logarithmic scale for different settings.

Other comments by R1. Regarding the motivation of the problem. Indeed, the aggregate regret is more practically relevant than the previously used simple regret, as the aggregate regret corresponds to the accuracy which practitioners care about, while the simple regret is not empirically meaningful. For example, we use the binary labeling task in crowdsourcing as a motivating example for the thresholding bandit problem with aggregate regret. The same problem (under the Bayesian setting) and its application to crowdsourcing has been extensively studied in (Chen et al., 2015). The reviewer may refer to (Chen et al., 2015) for more detailed information on applying our problem to crowdsourcing tasks.

Regarding the definition of the simple regret and the relationship between the two regrets. The simple regret is defined On Line 25. By Markov's inequality, we know that an algorithm with δ aggregate regret has at most δ simple regret, and an algorithm achieving δ simple regret has at most $K\delta$ aggregate regret.

Regarding "parameter-free" in APT. The APT algorithm is named by the authors of (Locatelli et al., 2016) to be "Anytime Parameter-free Thresholding". In their setting, the ϵ indifference zone is given as a problem input and not considered as a tuning parameter. However, ϵ has to be set by the user.

Regarding the uniform sampling method, being worse than the optimal algorithm by a factor of $\Omega(K)$ is very bad. It means that the algorithm cannot distinguish the arms when making adaptive query decisions (which is indeed the case for uniform sampling), and therefore is a trivial performance guarantee.

Regarding the technical contribution, we think our algorithm design and analysis technique to avoid the $\ln K$ caused by union bounds is highly non-trivial. In addition, the variable confidence level technique is also an important technical contribution. We believe that the overall technical contribution is significant, as also mentioned by both R2 and R3.

Regarding the constants. Since most of the contributions of this paper are theoretical, we do not focus on optimizing the constants, which is usual in most theoretical work. Please refer to the first paragraph for the setting of α in Remark 2.

Regarding Lemma 19, we set $a = \lambda_i \Delta_i^2/5 - \varkappa/2$ and $b = 1/(4\alpha)$ to prove Lemma 16. We agree with the reviewer that when $b \gg 1$, the confidence interval becomes longer while the error probability does not change much. In our proof, we only use the lemma when $b = O(1)$. It is possible that the lemma could be further improved when $b \gg 1$.

Regarding the analysis. On Line 267, we note that our Lemma 19 is a generalization of Hoeffding's maximal inequality, which is used in the analysis of MOSS. Our Lemma 19 can also replace the usage of Hoeffding's maximal inequality in the analysis of MOSS, and may find other applications. However, the analysis of our algorithm is very different from MOSS, despite them sharing the high-level intuition. We will make this clear in the next version of the paper.

Regarding the statistical significance. In Appendix F.1, we use T-tests to show that the confidence intervals of the performances of LSA and other algorithms do not overlap in most settings.