We thank the review team for the extensive set of comments and suggestions. Here, we detail our responses to your comments. Due to space limitations, the reviewers' comments are summarized in *italic*. We also adopt the same numbers for the references as the submission.

Reviewer 2. *Bayesian vs frequentist regret*. This is a great point. We actually establish our Bayesian regret bound by first deriving frequentist upper bounds on regret and then taking expectation with respect to the prior (see lines 532-537 on page 17). We used Bayesian notion throughout the paper which allows for an easier exposition of the ideas/results. We would be happy to add a discussion on the Bayesian vs frequentist regret in our final version.

Reviewer 3.

- *SS-Greedy does not establish universal rate optimality.* This is correct, but as mentioned in the abstract, this is one of the cases that the universal optimality does not tell the whole story and our empirical observations suggest that SS-Greedy performs very well in practice. We leave improved analysis of regret rates for SS-Greedy for future work.
- *Experimental results for subgaussian and uniformly upward-looking not included.* Due to space limitations, we pushed most of our simulations to the appendices. In fact, we include simulations over a wide range of beta priors with both Bernoulli and Gaussian reward (uniformly upward-looking as shown in Appendix C) in Appendix E.
- *Extension of results to $k > T$.* Thanks for bringing this to our attention. All our results can be extended to the case that $k > T$ (here the sub-sampling is inevitable). We will add a remark on this point in our final version.
- *Theoretical results and comparison with prior work.* Due to space limitations, we decided to discuss the most related works and refer the reader to the recent monographs by [18] and [23]. However, for improving the presentation of the paper, we will expand the related work section to address your point.
- *Summary of contributions, notations and typos.* Thanks for this suggestion and pointing out the typos. We will add a summary of contributions, fix all the typos and define the notations that have not been defined in the final version.

Reviewer 4.

- *Assumption 1 and its role in theoretical analysis.* You are absolutely right that Assumption 1 is central in our analysis. As shown in Section 7, our results slightly change for more general β-regular priors. This definition puts a constraint on $\mathbb{P}[\mu > 1 - \epsilon]$, which quantifies how many arms are ϵ-optimal. The larger number of ϵ-optimal arm means it is more likely that Greedy concentrates on an ϵ-optimal arm which is one of main components of our theoretical analysis. It would be interesting to remove all the assumptions on the prior, but it seems that we should not hope to get any result better than the well-known worst-case regret of $\Omega(\sqrt{kT})$ for arbitrary priors as all mass can be put on difficult problems (see Section 35.1 of [18]). Alternatively, one can replace it with some other assumption, however, we decided to use β-regular priors adapted from the literature on infinitely many-armed bandits [10, 25].
- *Simplicity of algorithmic idea and regret improvement.* A key (and practical) benefit that the Greedy algorithm offers is its simplicity. In our analysis, we observed that most subgaussian rewards are uniformly upward-looking and that is the main reason that we proved tighter regret bounds for this family. It is an interesting future direction to improve the regret bounds for subgaussian rewards. The linear regret for small values of k is an inherent property of greedy algorithms as the lack of active exploration leads to a linear regret with some probability. We prove that if k is large, this probability is small (and discuss how it decays depending on k for different rewards) and hence its contribution to total regret is negligible.
- *Comparison of assumption in Theorem 2 and Assumption 1.* Thanks for bringing this to our attention. In the small k regime, the lower bound on $\mathbb{P}[\mu > 1 - \epsilon]$ in Assumption 1 is not needed for establishing regret bounds. Hence, the assumption that density $g(x) \leq D_0$ for all $x \in [0, 1]$ implies that $\mathbb{P}[\mu \geq 1 - \epsilon] \leq D_0\epsilon$, implying the upper bound in Assumption 1. We will revise this sentence in the final version to address this.
- *Difficulties in analysis of Greedy & discussion of regret for different families.* The main difficulty in analyzing Greedy algorithms is that it is possible that all the good arms get poor observations at the beginning and hence the algorithm would stick to a sub-optimal arm, leading to a linear regret. We can be hopeful that the probability of this event is small if number of arms is large and is indeed a key component in our analysis of the Greedy algorithm. In particular, the quantity $q_\theta(\mu)$ (defined in Eq. (1)) captures how likely it is for a good arm to have its sample mean drop below θ (or get poor observations). For the exact same reason, the shape of $q_\theta(\mu)$ dictates the final regret bound and is the main reason for getting different rates for Bernoulli, subgaussian, and uniformly upward-looking distributions.
- *Informative priors.* This is an interesting point. If all arms have the same prior, the Bayesian regret should not depend on sampling strategy. But you are right that for different arm priors, uniform sampling can be sub-optimal and this would be an interesting future direction.
- *Comments on simulations.* Our intent in including simulations with both $d = 2$ and $d = 6$ is precisely to shed light on the relative impact of context diversity versus the presence of many arms, on the (good) performance of the greedy algorithm. Informally, we expect that for $k = 300$ and $d = 2$, context diversity is insufficient to ensure adequate free exploration; indeed, while greedy performs reasonably on average, the variance is high – the opposite of what we would expect if context diversity were sufficient (see, e.g., [6]). On the other hand, for $d = 6$, we can likely attribute the lower variance in this case to the context diversity, providing an additional source of free exploration that also contributes to the better performance of greedy. As we mainly consider Bayesian regret in our paper, we compare the expected performance, but we agree that in some applications it would be beneficial to consider other metrics.