We thank all three reviewers for their careful readings, valuable questions and constructive suggestions. We will revise the paper thoroughly and incorporate all the comments.

[reviewer 1] Thanks for the suggestion on adding a section of experiments. We agree that numerical results can enrich our paper and make it more convincing from an empirical perspective. And we are certainly happy to run some simulations to show the performance of our algorithms and to compare DUCB and DTS algorithms.

[reviewer 2] Thanks for raising the questions on infinite arms and frequentist regret bound for DTS. These are all great directions to explore and to further improve our paper. Thanks for the related references and we will add them in the revision.

(1) **DUCB versus DTS**: Thank you for raising this point, which we did not make clear in the initial submission! The goal of our paper is not to compare the frequentist regret bound for UCB with bayesian regret bound for DTS, but rather to characterize the regret performances in these two regimes for the two algorithms. Yes we agree that the frequentist bound for DTS can be derived by adapting [2, 1] for which it may need more modifications. This is an interesting direction to explore and we will definitely keep it in mind. The frequentist bound for DTS is stronger than the bayesian regret bound. We agree with the reviewer that this is exactly the reason why the frequentist bound for DUCB looks more complicated than the bayesian regret bound for DTS. Thanks for raising this point and we will add a remark on it. In addition, we are certainly happy to run some simulations to show the performance of our algorithms and to compare DUCB and DTS algorithms.

(2) **Infinite arms**: Our analysis does not depend on the assumption that the arms are finite. It also works with infinite arms. Note that the regret bound in our paper is not tight in terms of the feature dimension d. Our goal is to provide some clean and insightful analysis in simple settings. We can further tighten the regret bound from d to $\sqrt{d}$ by assuming the number of arms is finite and possibly changing and by using the baselinUCB/suplinUCB decomposition. (See [3] for the case without delays. Also, when the number of arms is both finite and fixed, the $O(\sqrt{dT})$ bound can be achieved with a simpler analysis.) We will add a remark on this point in the revision.

(3) **How to set** τ: In Algorithm 1, the warm-up period τ is to ensure that $\mathbb{E}[\sum_{t=1}^{\tau} X_t X_t'] \geq 1$. By Proposition 1 (equation (7)),

$$\tau = \left(\frac{C_1\sqrt{d} + C_2\sqrt{\log\left(\frac{1}{\delta}\right)}}{\lambda_{\min}(\Sigma)} \right)^2 + \frac{2B}{\lambda_{\min}(\Sigma)} + 2(\mu_D + M_D) + \sigma_G \sqrt{2\log\left(\frac{1}{\delta}\right)}. \tag{1}$$

The RHS of (1) is determined by the model parameters and can be directly calculated. Note that $\tau = O\left(d + \log\left(\frac{1}{\delta}\right)\right)$, i.e. the larger the feature dimension d is, the longer the warm-up period d is. This is not surprising since the algorithm needs a long time to gather information when the feature dimension is high. Also, τ increases as δ decreasing, which implies that it takes a longer warm-up period when the confidence level $1-\delta$ is larger. τ in Theorem 5 can be set similarly. We will add a remark on this point in the revision.

(4) **Numbering issue**: Thanks for catching it. We will fix the numbering issue in the revision.

(5) **References**: Thanks for the references ([4, 1]). We will add them in the revision.

(6) **Typos**: Thanks for catching the typo. Yes it should be q instead of p in line 207. We will change it.

[reviewer 3] Thanks for raising the questions on the lower bound and frequentist regret bound for DTS. These are all great directions to explore and to further improve our paper.

(1) **Frequentist regret bound for DTS**: The frequentist regret bound for DTS is stronger than the bayesian regret bound. We agree with the reviewer that the frequentist bound can be derived by adapting [2]. This is a great future direction to explore and we will definitely keep it in mind.

(2) **Lower bound**: The lower bound is certainly worth trying. It seems require some extra efforts and we will explore it in the future.

References

[1] Marc Abeille, Alessandro Lazaric, et al. Linear thompson sampling revisited. *Electronic Journal of Statistics*, 11(2):5165–5197, 2017.

[2] Shipra Agrawal and Navin Goyal. Thompson sampling for contextual bandits with linear payoffs. In *International Conference on Machine Learning*, pages 127–135, 2013.

[3] Lihong Li, Yu Lu, and Dengyong Zhou. Provably optimal algorithms for generalized linear contextual bandits. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, pages 2071–2080. JMLR. org, 2017.

[4] Claire Vernade, Olivier Cappé, and Vianney Perchet. Stochastic bandit models for delayed conversions. *arXiv preprint arXiv:1706.09186*, 2017.