We thank the reviewers for providing detailed and quality reviews. The primary objective of this paper is to provide the first linear bandit algorithm with a sample complexity nearly matching the information-theoretic lower bound.

Reviewer 2: We thank the reviewer for the positive comments and encouraging us to point out the novelty in our analysis. Despite extensive work on the fixed-confidence pure exploration linear bandit problem dating back to 2014, no existing paper has presented a near-optimal non-asymptotic algorithm. In the most recent work on linear bandits [2], obtaining such a result was raised as an existing open problem in the conclusion and our work gives a clear answer. We present the first non-asymptotic algorithm that nearly achieves the information-theoretic lower bound. Obtaining this result necessarily requires insights into the problem and analysis techniques beyond what has appeared in the literature.

The primary novelty in our analysis is quantifying the relationship between the algorithm sample complexity and the lower bound we provide. We characterize this relationship in the second inequality following line 438 and it is proven in the analysis that follows line 439. In this portion of the proof we show that the algorithm sample complexity including the problem-dependent terms matches the lower bound up to logarithmic factors. No analysis of this type has appeared in the literature previously and it is the fundamental component in the proof to show the algorithm we present is nearly optimal. This analysis opens up future research avenues to refine the proof technique in an effort to remove the $\log(1/\Delta_{\min})$ term showing up in the final result. A number of recent results on linear bandits have appeared, yet since they were unable to relate the sample complexity bounds to the lower bound as we do in our work, it has not been clear if any of them have come close to matching the lower bound. As we discuss in the related work section and demonstrate in our numerical experiments, it turns out that several recent algorithms have fundamental flaws that may be challenging to notice since the bounds cannot be compared to the lower bound.

Unlike the upper bound, we agree that the lower bound is obtained using standard techniques and we acknowledge this fact in our manuscript. However, we believe the tight lower bound we give is an important result to be able to characterize optimality. A final novelty in our analysis is obtaining tighter bounds on the problem-dependent terms. In Lemma 1, we show that $\rho(\mathcal{Y}) \leq d/\gamma_{\mathcal{Y}}^2$ where $\gamma_{\mathcal{Y}}$ is the gauge of $\mathcal{Y}$. In previous works [1], the problem-dependent quantity has been naively bounded as $\rho(\mathcal{Y}) \leq 4d$. The example after Lemma 1 shows the bound we provide can be significantly tighter. Proposition 1 is a lower bound, so d being even in the statement is simply giving an example to show there is a problem instance for which a static strategy must incur a factor of d samples more than necessary.

Reviewer 3: Thanks for pointing out the importance of clearly presenting the definition of an arm. In the transductive linear bandit problem, there are two finite sets of vectors $\mathcal{X} \subset \mathbb{R}^d$ and $\mathcal{Z} \subset \mathbb{R}^d$. An arm is a d-dimensional vector in the set $\mathcal{X}$ that can be measured directly, while vectors in the set $\mathcal{Z}$ cannot be measured directly. The objective is to identify the vector $z^* = \arg\max_{z \in \mathcal{Z}} z^\top \theta^*$ while obtaining only measurements from the set of arms of the form $x^\top \theta^* + \eta$.

The transductive experiment at the end of the experiments section is a general transductive example and not a combinatorial bandit problem since the set $\mathcal{Z} \not\subset \{0,1\}^d$. Comparing to existing combinatorial bandit algorithms is an interesting direction of future work. In the experiments the noise was generated from a standard normal distribution; this detail was included in the appendix on line 522, but we plan to move it to the main paper.

We appreciate you pointing out several typos that we are fixing. In Algorithm 1, $n_{\min}$ should be $r(\epsilon)$ and λ^* should be λ_t^* as you pointed out. In general, when $\sum_{x \in \mathcal{X}} \lambda_x x x^\top$ is non-invertible, pseudo-inverse can be applied. We will provide further details on the computation of λ^* in the paper thanks to your suggestion. While the computationally complexity of solving for λ^* is not a primary focus of our paper, empirically we find that it can be obtained efficiently.

Thanks for the detailed comments on the proof of Theorem 1. There is a typo on line 488 and $\nu_{\theta,i}$ should be defined as the reward distribution of the i–th arm in $\mathcal{X}$ so that $\nu_{\theta,i} = \mathcal{N}(x_i^\top \theta, 1)$. The index i is with respect to the n elements in $\mathcal{X}$ and the index j is with respect to the m elements in $\mathcal{Z}$.

Reviewer 4: Thank you for the careful review and giving several useful suggestions. We will make it clear the generalizations pertain only to pure exploration. Exploring insights into problem-dependent regret bounds is an interesting direction for future work. We plan to edit the paper to include pertinent references to the results in the least squares section. Proposition 1 says that there exists a problem instance for which a static strategy must incur a factor of d samples more than the optimal sample complexity. This indicates that it is necessary to devise an adaptive algorithm to obtain a near-optimal sample complexity. We will clear this up in the paper. The version of the algorithm from [1] implemented in the experiments is similar to our algorithm. The primary discrepancies are the greedy rounding procedure, phases being of a random length, and a slightly different condition to discard candidate vectors.

[1] Marta Soare, Alessandro Lazaric, and Rémi Munos. Best-arm identification in linear bandits. In *Advances in Neural Information Processing Systems*, pages 828–836, 2014.

[2] Chao Tao, Saúl Blanco, and Yuan Zhou. Best arm identification in linear bandits with linear dimension dependency. In *International Conference on Machine Learning*, pages 4884–4893, 2018.