We thank the reviewers for their valuable and extensive feedback. We briefly recap our contributions. In our work we provide improved lower bounds for the linear-bandit pure exploration problem, an algorithm that acheives this as an upper bound, computationally efficient algorithms in the combinatorial bandit setting, a novel fixed-budget algorithm for linear bandits, and empirical validation of our methods. We emphasize that our paper focuses on using methods from empirical process theory, such as the TIS inequality and Gaussian width, to develop a novel geometric notion of sample complexity, γ^*. Importantly γ^* captures the *moderate confidence* sample complexity in the regime where where δ is a constant such as $\delta = 0.05$, the typical case in practice. By optimizing γ^*, we design a *single algorithm* whose sample complexity is the first to essentially be within a $\log(d)$ factor of every other moderate confidence sample complexity result in the pure exploration MAB literature.

Regarding comments on the lower bound by Reviewers 2 and 4: Reviewers 2 and 4 point out there is an additive d factor in our upper bounds. Theorem 8 in the supplementary material shows that there is a broad class of problems where d samples are necessary. For example, if θ is unconstrained in the sense that the learner has no apriori knowledge about θ prior to the beginning of the game, then in combinatorial bandits at least d samples are necessary. We will expand this discussion in the final manuscript.

We thank reviewers 2 and 4 for their comments on the the lower bound, and will clarify the result in the final manuscript. The purpose of the Theorem is to shed light on the following question: if an agent chooses a distribution over $\mathcal{X}$, samples from it, and then terminates using the MLE (a very natural estimator), what is the best possible sample complexity? We agree with reviewers 2 and 4 that of course given the true θ, the natural approach is to compute z_*, but this estimator tells us nothing about the sample complexity of the problem. By restricting the estimator to output the empirical maximizer of the MLE, we shed light on what an algorithm without knowledge of θ could realistically achieve.

We also note that the lower bound can be strengthened quite easily to state: *if the non-interactive MLE takes fewer than $c(\gamma^* + \log(1/\delta)\rho^*)$ samples on a given problem $(\mathcal{X}, \mathcal{Z}, \theta)$, then it makes a mistake with probability at least δ.* The argument proceeds similarly to the proof in the paper by showing that if fewer than $c(\gamma^* + \log(1/\delta)\rho^*)$ samples are taken, then anti-concentration bounds of a normal random variable (e.g., Proposition 2.1.2 of [30]) and the TIS inequality imply that with probability at least δ, the algorithm makes a mistake.

Reviewers 1 and 3: Reviewer 1 is correct that the extra $\log(1/\delta)$ comes from only optimizing γ^*. We did this for computational reasons and will include motivation and intuition in the final manuscript. We agree with reviewer 3 regarding the importance of extensions to an approximate oracle and of the noise assumption, and will discuss these in the final manuscript. Our algorithms can be extended to the Sub-Gaussian noise case (see the supplementary material).

Reviewer 2: The suggested ϵ-net would give an extremely loose sample complexity: $d\log(1/\epsilon)\rho^*$ whereas our algorithm attains $c\log(d/\delta)\rho^*$ in the problem Top-K. The point of our paper is to find a data-efficient union bound in a generic and computationally efficient manner. In addition, constructing an ϵ-net directly is computationally burdensome.

"For ρ^* and γ^*, can you compute it...": ρ^* is always convex and γ^* is convex in the combinatorial bandit setting; we conjecture that γ^* is convex in general. A linear maximization oracle can be used to compute γ^*.

"when the oracle is efficient, does it mean that its structure can also guarantee a small union bound?": There is no clear-cut relationship. Proposition 2 gives an example where the union bound is large and the oracle is efficient.

"When you say Theorem 4 has an actual factor of d, did you mean the last term?": We are referring to the additive term $+d$ in our sample complexity, which is essentially $\rho^* \log(1/\delta) + \gamma^* + d$.

Reviewer 4: “What is the intuition behind the “Linear bandits” experimental setup?...” The pure exploration linear bandit problem is well established, see [12,27]. As we describe, it is also sufficiently general to encapsulate many problems in MAB, such as combinatorial bandits. Our sample complexity results are essentially within a $\log(d)$ factor of every result in pure exploration MAB.

“It's always really helpful for the reader to use words as well as mathematical notation to explain concepts...” Thank you for the feedback. We will add some clarifying sentences such as "at round k, Algorithm 1 finds a design λ_k that makes the application of the TIS inequality as tight as possible when applied to the remaining items in $\mathcal{Z}$. Then, it takes enough samples to guarantee that all items z with gap larger than $2^{-k}\Gamma$ are eliminated." We will provide a similar description of our other algorithms.

"The notation P_θ and E_θ are used in several places..." θ is is not random; $\mathbb{P}_\theta$ denotes the probability law induced by θ and the sampling rule of an algorithm (see [23] for similar usage). We will clarify its meaning in the final manuscript.

"The definitions of ρ^* and γ^* are said to be based off..." If $c\rho^* \log(|\mathcal{Z}|/\delta)$ samples are taken, then manipulating the inequality (2) shows that each gap is estimated well enough to identify z_* (see [27] Section 3 for a helpful discussion).