Dear Reviewers,

Thank you for your time and effort. We are happy to see the positive feedback, especially on the appeal of our game-theoretic equilibrium technique.

Reviewer 1 The main message concerning our algorithm in the single-answer case is that it is very close to Track-and-Stop, and has similar performance.

The reviewer asks about the lower bound in (Garivier and Kaufmann, 2016). That lower bound deals with single-answer problems exclusively, and the paper does not mention ϵ-best arm. We believe the relevant paper is the earlier (Kaufmann, Cappé and Garivier, 2015), which does have a lower bound for ϵ-best arm in Remark 5.

As we show in our paper, the optimal rate at which the sample complexity grows with the confidence $\ln \frac{1}{\delta}$ is

$$\frac{1}{\min_{i \in i^*(\mu)} \max_{w \in \triangle} \inf_{\lambda \in \neg i} \sum_k w_k d(\mu_k, \lambda_k)}.$$

We can indeed obtain the bound from ours as follows. (Kaufmann, Cappé and Garivier, 2015) aim to get the best lower bound obtainable from using alternative bandit models $\lambda \in \neg i$ which differ in only one arm, i.e. $\#\{k \mid \mu_k \neq \lambda_k\} = 1$ (see Figure 1(a)). We search over all of $\neg i$. For ϵ-Best Arm Identification specifically, it turns out that it suffices to move only two arms (see Figure 1(b)).

Unfortunately, neither approach when taken to its extreme results in a closed form expression for the rate. For Gaussian arms, it is known that the advantage of two-arm movements is at most a factor 2, see (Garivier and Kaufmann, 2016, page 6). However, for Bernoulli bandits the advantage can be arbitrary.

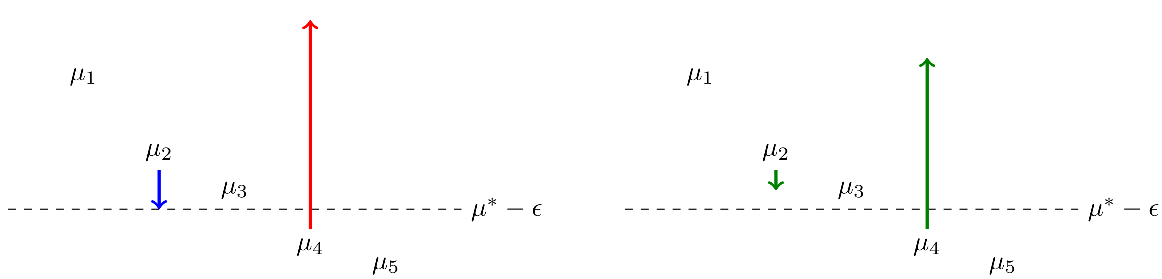
Reviewer 2 We can indeed cite `https://arxiv.org/abs/1905.08165` (note however that it was on arxiv for less than a week before the submission deadline). It studies the general fixed confidence pure exploration problem with a single answer, while the main point of interest of our paper is the multiple-answer case.

Thank you for including the typos you spotted. In contrast with the other reviewers, you found the paper dense. To help us improve the paper, would you perhaps be able to update your review with line numbers of particularly opaque paragraphs?

Reviewer 3 Thanks you for your reference [1]. We will add it to the discussion as it is indeed relevant to our paper for two reasons. It extends the Track-and-Stop approach to the general identification case (beyond Best Arm). Even though the technical contribution in the paper are formulated for the Gaussian case only, the ideas naturally extend to other families. In addition, the paper gives non-asymptotic terms that are hinting at problem complexity dependencies beyond confidence. We are not currently making any claims on this front.

The paper matches the asymptotic rate up to a multiplicative constant. In our opinion though, this is a nice step on the way to, but still a far cry from, matching the lower bound rate *exactly*. The latter is possible for Best Arm and, as we show now, many other "general sampling" problems.

We have some ideas for and are are working on techniques for obtaining efficient algorithms non-asmptotic bounds that matching the asymptotic optimal rate. It will be very interesting to compare lower-order terms once we are there. However, this is a separate project beyond the current paper.



(a) Two single-arm motions that make arm $i = 2$ incorrect. The blue moves arm i to $\lambda_2 = \mu_1 - \epsilon$. The red moves any other arm $j \neq i$ to $\lambda_j = \mu_i + \epsilon$. All other arms k remain fixed at $\lambda_k = \mu_k$ in either case.

(b) A single two-arm motion that makes arm $i = 2$ incorrect. We move both arm i to some level $\lambda_i = \zeta - \epsilon/2$ and some other arm $j \neq i$ to $\lambda_j = \zeta + \epsilon/2$. We choose the level ζ to optimise the bound.