Thank you for the reviews. We address below specific questions/comments (shortened) made by the reviewers.

**Reviewers 1 and 2**

**Question**: Please provide more motivation for the setting.

**Answer**: Applications of the proposed setting include any case where there is an exploration stage in which rewards are not collected, but payment is still proportional to the quality of the arm. In the paper we give an ad-placement survey application as an example. As another example, consider a sensor field, where the goal is to find the most active sensor, but each activity report has a communication cost. Here, "pulling an arm" is done by activating the sensor's reporting option for a given period of time. Yet another example is that of finding the best crowd worker on a crowd-sourcing platform. Here, an "arm pull" is equivalent to providing the worker with a test task. The workers are paid according to their success rate in the test task, thus a more successful worker will be paid more during the exploration.

**Question**: **Reviewer 1**: It is said that the best arm corresponds to the best expression for an ad, and pulling arms is equivalent to observing the number of clicks when the ad is linked to that expression. But then, the player cannot actively "pull an arm" in any rounds she wants. Also, why is the cost of pulling an arm (which means observing something) equals the reward of the arm? **Reviewer 2**: About the motivation that the cost of arm pull is equal to the expected future reward of that arm, I am not sure this is actually the case used in ad placement.

**Answer**: In the example application, the player can "pull an arm" by linking its ad to the expression represented by this ad for a given period of time, and then observing the number of clicks. The cost of pulling that arm is thus the number of clicks on that ad, assuming the common pay-per-click advertisement model.

**Reviewer 1**

**Question**: Why is your regret bound significantly better than those in [14]?

**Answer**: We give an algorithm with a cost that has no dependence on the number of arms. This is a significant improvement, since it allows any number of arms. For instance, in the ad-placement application, the number of arms could be the number of search expressions, which is huge. Similarly, there could be a huge number of sensors in a field of nano sensors. Moreover, completely removing the dependence on the number of arms is of considerable theoretical significance.

**Question**: Line 85: why do you call these "single" arm-pulls, while these are batched arm pulls?

**Answer**: $n$ is defined to be the total number of times that any arm is pulled, regardless of whether some of the arms are pulled together or not. We will clarify this.

Thank you for the detailed comments, we will fix all of them.

**Reviewer 2**

**Question**: The authors did not show explicitly the constants in the analysis.

**Answer**: This work shows that is is possible to have no dependence on $K$, which is an interesting and non-trivial result. We plan to optimize the constants and derive a practical implementation in future work.

**Question**: The assumption that it is possible to "pull all arms in finite time" for manipulating an unlimited number of arms is not practical.

**Answer**: We explain in lines 141-165 how it can be practical in some applications to pull many or all arms together. For instance, in the ad-placement example, one can link the ad to all search queries, and simply record which search queries actually came up and resulted in an ad-click. In the sensor field example, all arms can be "pulled" at the same time, by activating the reporting option of all sensors at once. In the crowd-worker option, a test task can be uploaded without any filters so that all crowd-workers can try it.

**Question**: It seems that $\theta$ was not used in the proof of Theorem 3.1.

**Answer**: $\theta$ is used in lemma 5.4. The lemma is then used in the proof of theorem 5.6 (the restatement of theorem 3.1).

Thank you for the detailed comments, we will fix all of them.

**Reviewer 3**

**Question**: Naming quantities so their qualitative function is well motivated would help. Please provide more explanations in the analysis.

**Answer**: We will rename the constants and add explanations on their meaning for better readability. We will add further explanations to the technical analysis.