Author feedback for "Polynomial Cost of Adaptation $\mathcal{X}$-Armed Bandits"

We thank the reviewers for their overall positive and constructive comments.

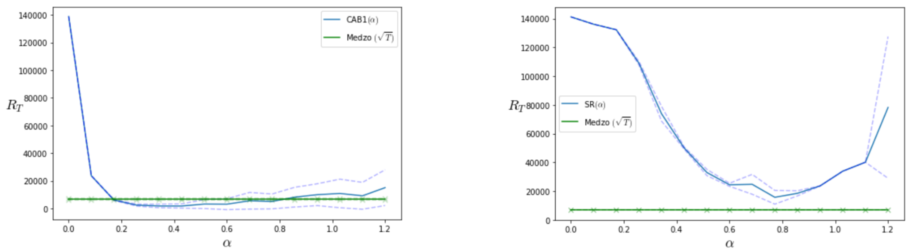
To answer some concerns of Reviewer #1, we would like to reemphasize the significance of the paper. Adapting to the unknown smoothness in X-armed bandits has been an open problem since Bubeck, Munos, Stoltz, and Szepesvári [3], and a few partial answers have been published since then (described the literature review). This paper completes the picture in the minimax Hölder setting.

Moreover, the Hölder assumption (stated under various names) is standard in this line of work, e.g., in Agrawal [1], Kleinberg [4], Auer et al. [2], Bubeck et al. [3] and Locatelli and Carpentier [6]. Hölder regularity is also omnipresent in non-parametric statistics.

Furthermore, as stressed by Reviewer #2, the paper introduces novel algorithmic ideas that could be applied to other settings. The vanilla X-armed bandits model can be extended in many ways, the first one coming to mind being the contextual setting of Krishnamurthy et al. [5].

Therefore, we believe a strong point of the paper lies in the algorithmic idea together with the proof techniques, and this is why we insisted on being complete and thorough in the mathematical steps.

That said, we acknowledge the technicality of the paper. As suggested by Reviewer # 2, we will add an illustrative figure describing how the algorithm behaves and giving some intuition. We will also follow the recommendation of Reviewer #3 and add the following numerical experiments, for illustrative purposes. We recall that it is quite standard (albeit unfortunate), that papers in this area do not include experiments.



(a) CAB1 **(b)** Subroutine

Figure 1: Average regrets of CAB1 and Subroutine (from Locatelli and Carpentier [6]) tuned with varying values of α, and of Medzo, after $T = 300000$ time steps. The algorithms were run 30 times and the error bars are 1.96 times the standard deviation. The problem used is $x \mapsto (1/2)\sin(13x)\sin(27x) + 0.5$, taken from Valko et al. [7]

With no knowledge of the true regularity, Medzo obtains a regret that is almost the same as that of algorithms optimally tuned. Intriguingly, CAB1 performs quite well when the smoothness is overestimated, although the variance becomes quite high. Experiments will be further commented in the final version.

Specific points :

R1: "The horizon T is assumed to be a prior knowledge. This should be stated and commented [...]. " : Indeed, Subsection 3.3 and Appendix B discuss this and describe how we can get rid this requirement. In the final version we will recall that by “anytime” we mean without the knowledge of T.

R2 "Can this algorithmic technique deal with cases in which the function is *spatially inhomogenous*, for instance if the Hölder exponent α varies with the input point x." : This is a good point. Our guarantees hold if the Hölder property is satisfied in a small neighbourhood around the maximum, but the minimal size of this neighbourhood depends on T. Our analysis only requires that in every discretizations, (i.e., in the first), the average payoff of the cell containing the optimal action is close to optimal. This is actually the case in previous papers (Zooming, HOO, Locatelli and Carpentier). We will add this remark in the main paper, with some detail in the appendix.

R3 : "I think I have read some papers that use a similar idea, but this one is the first one I know about using this trick on X-armed bandit model." : We would be happy to read (and cite) these works if you find them.