We thank the reviewers for their feedback and detailed comments on the manuscript. We address below the most important issues that will be fixed in the final version of the paper.

Dependence in the dimension d Both Reviewers #2 and #3 noted that the dependence in d would be expected to be of the order of $d^{2/3}$, in light of [12]. This is indeed a relevant remark, and we confirm that by taking $\gamma = 1-(B_T/(dT))^{2/3}$ (as proposed by Reviewer #3) one obtains a regret bound of this order. Regarding the comparison with [5] (Reviewer #2): [5] relates to the non-contextual case, where the regret scales as $K^{1/3}$ only. This can be seen as a special case of our model by choosing $K = d$ with fixed orthogonal contexts. However, in the general contextual case, the $d^{2/3}$ rate appears as a consequence of the need to control deviations in all directions thus adding an additional $\sqrt{d}$ term in the control of the stochastic term. Consequently, the optimal exponent of the dimension term to achieve the equilibrium between the bias term and the stochastic term for the dynamic regret is $1/3$ in the case of a control for fixed directions as in [5], but it is $2/3$ when the control is in all directions as in [12] and in our analysis.

Novelty of the deviation bound We agree with Reviewer #3 that part of the arguments used in the proof of our Theorem 1 are common with those introduced in [1]. However, Theorem 1 is not a mere consequence of the result in [1]. To be more precise: Lemma 1 may indeed be seen as a simple extension of Lemma 8 of [1] to the case of heteroscedastic noises (considering that the noise terms are given by $w_s\eta_s$). The proof of this lemma was included to ease the understanding of the rest of the proof and to make the paper self-contained. Lemma 2 and 3 however differ from Lemma 9 of [1] by the fact that we consider time-dependent regularization parameters (λ_t). As explained in the main text, this is unavoidable when using exponential weights to avoid vanishing effect of the regularization. Technically this implies that $\widetilde{M}_t$ defined in Lemma 2 is no more a supermartingale, although it still holds true that $\mathbb{E}[\widetilde{M}_t] \leq 1$. In the weighted case, we also need to consider mixing constants (μ_t) in the method of mixtures that differ from the regularization parameters λ_t (see Section 2). We will be more explicit regarding these differences with [1].

Related works We thank Reviewers #2 and #3 for their comments that have improved greatly our literature review; we have done our best to include all mentioned references. [Keskin & Zeevi], pointed by Reviewer #2 is an interesting application-oriented paper on non-stationary bandits that has similarities with ours in the particular case where the dimension $d = 2$, where they use a weighted estimator to construct first order optimal policies. It is true that the seminal paper [Auer, 02] already considers shifting environments but they bound a different regret than the dynamic regret in [15]; we'll comment on that. [Luo et al. COLT'18] and [Chen et al. COLT'19] will be discussed in the final version. We also thank Reviewer #2 for correcting the typo in the order of the dynamic regret bound for [12,13] in line 78, where exponents had been exchanged.

Experimental section The objective of Section 4.2 was not to address a real-world application but rather to illustrate the behavior of the algorithms in higher-dimensional problems. To do so, we have used the Criteo dataset only to provide plausible context vectors and to avoid the peculiarities that would have arisen by, for instance, simulating these at random on the hypercube as done in the small-dimensional example of Section 4.1. This experiment is interesting in particular to confirm that the scaling with respect to the various parameters is correct, even if it does not solve a real practical problem. This will be clarified in the text. In addition, we agree that the writing of this section could be improved. We have simplified the preprocessing of the data and also incorporated the correct dependency on d for γ (see above), and the corresponding figure will be included in the final version of the paper.

We provide below answers regarding comments that do not necessitate significant modification of the paper.

- We agree with reviewers #1 and #2 that dependence on B_T is currently a limitation of the proposed approach. As pointed out by Reviewer #2, the method proposed in [12] could also be used with D-LinUCB. We are not sure however that this approach would be practically satisfying for common values of the dimension d and the time-horizon T.
- (Reviewer #3) At line 461, the matrix M refers to $V_{t-1}^{-1}\sum_{s=t-D}^{p}\gamma^{-s}A_sA_s^\top$ and hence is indeed a symmetric matrix.
- (Reviewer #3) At line 483, the problem is the probabilistic nature of the statement: $|x^\top(\hat{\theta}_t - \theta_t)| \leq L$ for all $x \in \mathbb{R}^d$ would indeed imply that $|X^\top(\hat{\theta}_t - \theta_t)| \leq L$ for any $\mathbb{R}^d$-valued random variable X. However, $\mathbb{P}(|x^\top(\hat{\theta}_t - \theta_t)| > L) \leq \delta$ does not imply that $\mathbb{P}(|X^\top(\hat{\theta}_t - \theta_t)| > L) \leq \delta$, this can even be obviously wrong, for instance, when $X = \hat{\theta}_t - \theta_t$. This gap justifies the inclusion of a corrected proof in Appendix C.