We would like to thank the reviewers for their valuable feedback.

[R1, R3] Asymptotic problem-dependent vs finite-time worst-case regret. We agree that finite-time regret is the performance measure of interest. At the same time, problem-dependent optimality is a stronger notion than worst-case optimality, as it requires an algorithm to perform optimally in *every single* instance. Linear contextual bandits have been studied extensively from a *worst-case* perspective and minimax optimal strategies exist. These strategies are robust to worst-case instances, but in general they fail to adapt to the structure of the problem (e.g., they ignore informative arms) and may perform poorly in practice (see e.g., [7] and our experiment in Fig.1). In recent years, several attempts have been made to design more adaptive algorithms by leveraging *asymptotic lower bounds*, which effectively capture all problem-specific characteristics (e.g., set of arms, possible constraints on the parameters, reward distribution) into the regret bound. While the resulting algorithms directly inherit asymptotic optimality, the question is whether their more problem-adaptive behavior translates to competitive finite-time performance w.r.t. worst-case optimal algorithms. While this is the case for best-arm identification (see e.g., "Explicit Best Arm Identification in Linear Bandits Using No-Regret Learners" [Zaki et al., 2020] and "Gamification of Pure Exploration for Linear Bandits" [Degenne et al., 2020]), it still remained as an open question for regret minimization in linear bandit, where algorithms like OSSB and OAM have several practical limitations and they are rarely preferable over LinUCB or LinTS. We believe our paper is a significant step forward in addressing this question: SOLID resolves most of the issues of existing asymptotically optimal algorithms, significantly improves their finite-time regret guarantees, and it is shown to be empirically better than *practical* versions of LinUCB and LinTS in a variety of settings (including real data, see App. K.3).

[R1, R3] Dependence on contexts. We significantly improved the regret guarantees w.r.t. [14] by removing any dependency on $1/\rho_{\min}$ (which is at least as large as $|\mathcal{X}|$). Yet, we conjecture the dependence on $|\mathcal{X}|$ could be improved further. SOLID optimizes and updates a context-arm exploration strategy and this may suggest a polynomial dependency on the size of the exploration strategy is unavoidable. Nonetheless, we managed to push the dependency on the number of arms to a logarithmic term (greatly improving previous results) and a similar approach could be used for contexts by avoiding concentrating $\hat{\rho}$ to ρ (see e.g., Lemma 11), which is currently the main source of dependency on $|\mathcal{X}|$. This conjecture is also supported by empirical evidence. On the Jester dataset in App. K.3, SOLID's performance is not significantly affected by the number of contexts (almost 200) and it still performs better than LinUCB/LinTS. We will run additional experiments on Jester for different values of $|\mathcal{X}|$ to further investigate the dependency.

[all] Non-contextual case. We would like to bring to the reviewers' attention that while the paper is framed in the general contextual case, our contribution should also be assessed in the simpler and yet significant non-contextual case. Our algorithm and analysis resolve many open questions in this setting, including the dependence on $|\mathcal{A}|$, the derivation of confidence sets over parameters with optimal asymptotic scaling (Thm. 1), and the efficient incremental computation of the lower bound. After the submission, we have also analyzed the *finite-time worst-case* properties of SOLID when $|\mathcal{X}| = 1$. In this case, a simple proof following almost directly from the proof of Thm. 2, we derived an $\tilde{\mathcal{O}}(\sqrt{dn})$ regret bound that holds for *any horizon* n. This implies that SOLID is the first algorithm that is both *finite-time minimax optimal* and *asymptotically problem-dependent optimal* for linear (non-contextual) bandits.

[R1] Experiments. We ran the *practical* version of LinUCB/LinTS, using confidence sets without numerical constants, with log-determinant of the design matrix for LinUCB, and without the *theoretical* oversampling $\sqrt{d}$-factor for LinTS.

[R1] "The leading order term depends inversely on the smallest possible gaps between arms' mean rewards". This is not the case in problems with structure (e.g., linear). Examples like the one of the first experiment (which extends the one in [7]) show that there exist problems in which one can make some arm gap arbitrarily small, yet the optimal regret rate $v^\star(\theta^\star)$ does not scale with it since pulls are allocated to other informative arms. As a result, the algorithm's behavior and performance are not negatively affected by the existence of policies that are extremely similar to the optimal one.

[R2] Sub-Gaussian assumption. The lower-bound remains the same, except that the KL divergence needs to be computed for some distribution in the sub-Gaussian family. The analysis would be almost identical thanks to the Lipschitz property of KL divergences between sub-Gaussian distributions (see, e.g., [15]) and the results would be the same with a distribution-dependent Lipschitz constant in lower order terms.

[R3] "a simple epsilon-greedy algorithm [...] is already asymptotically optimal". An algorithm is asymptotically optimal if its regret scales as $\log(n)$ with a leading problem-dependent constant matching the $v^\star(\theta^\star)$ in the lower bound. This is very important in practice because it certifies that the algorithm effectively adapts to the problem's structure. In this sense, an epsilon-greedy algorithm is far from being asymptotically optimal as it only recovers a $O(\log n)$ regret possibly with a prohibitively large constant (e.g., scaling linearly with the number of context-arms or inverse of the gaps, which can be extremely small). This directly translates into a poor performance in practice.

[R3] "The authors need to take a look at Chu et al.". We cite [4], which refine the original results ofChu et al. [2011].

[R4] Thanks for the supportive review and for the suggested corrections. We have already updated the paper accordingly.