DOES LINUCB NEED VALID CONFIDENCE SETS?

Anonymous authors
Paper under double-blind review

ABSTRACT

We show that replacing LinUCB's usual confidence width β by any fixed $\alpha > 0$ preserves sublinear regret when the action set is M-smooth and m-strongly convex. Our high-probability regret bound for this generalisation, α-LinUCB, establishes that valid confidence sets are not necessary for the success of LinUCB. We identify the mechanism behind this result: action-set curvature drives α-LinUCB to explore every direction, even when optimism is not guaranteed. To analyse this mechanism, we introduce a logarithmically capped spectral potential that, unlike the standard log-determinant potential, caps contributions from well-explored directions. This yields lower bounds on the smallest eigenvalue of the design matrix. At the standard confidence width $\alpha = \beta \asymp \sqrt{d \log T}$, our T-round regret bound in d dimensions is $\mathcal{O}\big(\frac{M}{m} d\sqrt{T \log T}\big)$, up to iterated-logarithmic factors, improving the horizon dependence of the classical LinUCB guarantee by a factor of $\sqrt{\log T}$.

1 INTRODUCTION

A learner interacting with a bandit environment chooses actions sequentially to earn rewards while gathering the information needed to improve future decisions. We consider this exploration–exploitation trade-off in the stochastic linear bandit setting. Fix an unknown parameter $\theta^\star \in \mathbb{R}^d$, a closed action set $\mathcal{X}$ contained in the unit ball of $\mathbb{R}^d$, and an integer horizon $T \geq 1$. At each round $t \in \{1, \dots, T\}$, the learner selects an action $x_t \in \mathcal{X}$ and receives a reward $y_t = \langle \theta^\star, x_t \rangle + \eta_t$, where the noise η_t is conditionally centred and sub-Gaussian. We consider the T-round (pseudo-)regret,

$$\mathcal{R}_T := \sum_{t=1}^{T} \langle \theta^\star, x^\star - x_t \rangle, \qquad x^\star \in \arg\max_{x \in \mathcal{X}} \langle \theta^\star, x \rangle.$$

LinUCB addresses this trade-off using upper confidence bounds on rewards (Dani et al., 2008; Li et al., 2010; Abbasi-Yadkori et al., 2011). We study α-LinUCB (Algorithm 1), a generalisation that treats the confidence width as a free parameter. Given fixed $\alpha, \lambda_0 > 0$, at each round t, the algorithm selects x_t through

$$(x_t, \theta_t) \in \arg\max_{(x,\theta) \in \mathcal{X} \times \mathcal{C}_t(\alpha)} \langle \theta, x \rangle, \qquad \mathcal{C}_t(\alpha) := \left\{ \theta \in \mathbb{R}^d : \|\theta - \hat{\theta}_t\|_{V_t} \leq \alpha \right\},$$

where $\hat{\theta}_t = V_t^{-1} \sum_{s=1}^{t-1} x_s y_s$ is the ridge estimate of $\theta^\star$ with design $V_t := \lambda_0 \mathbf{I} + \sum_{s=1}^{t-1} x_s x_s^\top$. Equivalently, the action satisfies

$$x_t \in \arg\max_{x \in \mathcal{X}} \left\{ \langle \hat{\theta}_t, x \rangle + \alpha \|x\|_{V_t^{-1}} \right\}.$$

Thus, actions are selected through a combination of the mean and an α-weighted exploration bonus, with $\|x\|_{V_t^{-1}}$ being large for x aligned with directions poorly represented by previous actions.

Standard LinUCB chooses a confidence width $\beta \asymp \sqrt{d \log T}$ so that $\theta^\star \in \bigcap_{t=1}^{T} \mathcal{C}_t(\beta)$ with high probability. On this event, the selected pair (x_t, θ_t) satisfies the optimism inequality $\langle \theta_t, x_t \rangle \geq \langle \theta^\star, x^\star \rangle$, which underlies its usual regret guarantee (Abbasi-Yadkori et al., 2011). For arbitrary $\alpha > 0$, α-LinUCB retains an exploration bonus, but its ellipsoids need not be valid confidence sets. The optimism inequality is then no longer guaranteed, so the usual regret argument does not extend to arbitrary choices of α. This raises the question: *Does LinUCB Need Valid Confidence Sets?*

Under an additional curvature condition on the action set, we answer this question in the negative:

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

- *Regret without valid confidence sets.* When the squared support function of the action set is M-smooth and m-strongly convex, we prove a high-probability regret bound of[1]
$$\mathcal{R}_T = \mathcal{O}\Big(\frac{M}{m}\frac{\alpha^2+\beta^2}{\alpha}\sqrt{dT}\Big),$$
up to iterated-logarithmic factors (Theorem 4.2). The bound is sublinear for every fixed $\alpha > 0$, including widths for which the exploration ellipsoids are not valid confidence sets. The geometric condition, stated in Assumption (A2), includes Euclidean balls, ellipsoids, and convex bodies with uniformly positive and bounded curvature.
- *A geometric exploration mechanism.* We identify the mechanism through which α-LinUCB works without optimism: action-set curvature drives α-LinUCB to explore every direction, even when optimism is not guaranteed. Our spectral theorem makes this precise by controlling the growth of the design matrix (Theorem 4.3). On the unit ball, it gives a lower bound proportional to $\sqrt{t}$ on $\lambda_{\min}(V_t)$, holding with high probability at every round after a burn-in period; the general result controls the corresponding width in the action set geometry.
- *A logarithmically capped potential.* To analyse this exploration mechanism, we introduce a logarithmically capped spectral potential that, unlike the standard log-determinant potential, caps contributions from well-explored directions. The resulting spectral bounds yield regret control without the usual elliptical-potential argument. At the standard width $\alpha = \beta$, this gives $\mathcal{O}\big(\frac{M}{m}d\sqrt{T\log T}\big)$ regret, improving the horizon dependence of the classical $\mathcal{O}(d\sqrt{T}\log T)$ guarantee by a factor of $\sqrt{\log T}$, up to iterated-logarithmic factors.

Moreover, we construct a counterexample confirming that assumptions on the action set are necessary: on a two-dimensional polygon, α-LinUCB incurs linear regret with constant probability, for every sufficiently small $\alpha > 0$ (Proposition 6.1). We support our theoretical findings with experiments: on the unit ball, the observed regret and smallest-eigenvalue growth are consistent with the predicted $\sqrt{t}$ rates, even when the ellipsoids exclude the true parameter in most rounds.

Notation. We write $[T] := \{1, \dots, T\}$, $\mathbf{I}$ for the identity matrix in $\mathbb{R}^{d\times d}$, and $\lambda_{\min}(V)$ for the smallest eigenvalue of a symmetric matrix V. For $u, v \in \mathbb{R}^d$, let $\langle u, v\rangle$ denote their inner product, and define $\|u\|_2 := \sqrt{\langle u, u\rangle}$ and $\|u\|_V := \sqrt{\langle u, Vu\rangle}$ for $V \succ 0$. We write $\mathbb{B}_2^d(r)$ for the d-dimensional closed Euclidean ball of radius r centred at the origin.

2 Related work

Our work is closest to that of Abeille et al. (2025), who use the same geometric condition to analyse Thompson sampling without forced optimism or posterior inflation. Banerjee et al. (2023) also connect curvature with exploration, deriving growth of the smallest eigenvalue of the expected design matrix from an assumed $\mathcal{O}(\sqrt{T})$ regret guarantee. We discuss these and further connections below.

Optimism and confidence sets in linear bandits. The classical $\mathcal{O}(d\sqrt{T}\log(T))$ regret analysis of LinUCB/OFUL combines self-normalised concentration with the elliptical-potential lemma (Li et al., 2010; Abbasi-Yadkori et al., 2011). Each contributes a $\sqrt{\log(T)}$ factor, leaving a logarithmic gap to the $\Omega(d\sqrt{T})$ minimax lower bound (Dani et al., 2008; Rusmevichientong & Tsitsiklis, 2010; Lattimore & Szepesvári, 2020). Lattimore (2023) proves that the continuous-time counterpart of the self-normalised concentration bound cannot be improved for general adaptive covariates. Flynn et al. (2023) construct tighter confidence sequences using adaptive martingale mixtures, improving empirical performance while retaining $\mathcal{O}(d\sqrt{T}\log(T))$ worst-case regret. Our analysis uses the standard confidence bound but with a logarithmically capped potential.

Classical frequentist analyses of linear Thompson sampling also rely on optimism: posterior inflation ensures a constant probability of an optimistic draw, at the cost of an additional $\sqrt{d}$ factor in the regret bound (Agrawal & Goyal, 2013; Abeille & Lazaric, 2017). For finite-armed linear bandits, Lattimore & Szepesvári (2017) show that standard optimism-based algorithms can be suboptimal. Subsequent asymptotically optimal algorithms explicitly optimise exploration allocations to account for the information provided by suboptimal actions (Hao et al., 2020; Tirinzoni et al., 2020).

[1] Throughout, we use the standard asymptotic notation $\mathcal{O}(\cdot)$ and $\Omega(\cdot)$, up to iterated-logarithm factors.

Action-set curvature and exploration. When the action set is smooth and strongly convex, the played arm $x_t = \nabla J_{\mathcal{X}}(\theta_t)$ moves continuously with the parameter, and perturbations of the estimate translate into exploration. The benefit of curvature is classical in online learning, where follow-the-leader enjoys fast rates on strongly convex constraint sets for much the same reason (Huang et al., 2016). In stochastic bandits, Banerjee et al. (2023) quantify this: for action sets with bounded curvature around the optimal arm, any algorithm with $\mathcal{O}(\sqrt{T})$ expected regret has a design matrix whose minimum eigenvalue grows as $\Omega(\sqrt{T})$ in expectation; in contrast well-separated arms (Lattimore & Szepesvári, 2017) force eventual logarithmic growth. Abeille et al. (2025) exploit the same geometry, under the assumption we adopt here, to analyse randomised exploration without forced optimism or posterior inflation, and show that Thompson sampling attains $\mathcal{O}(d\sqrt{T}\log(T))$ regret.

Taking $\alpha = 0$ in our decision rule gives the greedy policy. Greedy methods are studied in contextual bandits, but only known to succeed under diversity of contexts, either assumed (Bastani et al., 2021) or induced by smoothing (Kannan et al., 2018; Raghavan et al., 2023). On a fixed action set no such diversity is available and, accordingly, our bound degrades as $\alpha \to 0$.

Inference under adaptive sampling. Conditions on the growth of the design matrix play a central role in inference from adaptively collected data, beginning with Lai & Wei (1982). Ordinary least squares can have non-Gaussian asymptotic distributions and inflated estimation error under adaptive designs (Khamaru et al., 2025; Lin et al., 2023), motivating debiasing procedures that recover asymptotic normality under conditions on the data-collection process (Deshpande et al., 2018; Khamaru et al., 2025; Ying et al., 2023; Lin et al., 2023). A complementary approach establishes stability of the design after deterministic normalisation; Praharaj & Khamaru (2026) use this to obtain asymptotically valid Wald intervals for a regularised EXP4 algorithm in linear contextual bandits.

3 Assumptions

We use the following two assumptions, placed on the noise and the action set, respectively.

Stochasticity. Let $\mathcal{F}_t$ be the σ-field generated by the history up to round t together with the next action, so that x_{t+1} is $\mathcal{F}_t$-measurable; namely, $\mathcal{F}_t := \sigma(x_1, y_1, \ldots, x_t, y_t, x_{t+1})$.

(A1) (Sub-Gaussian Noise). *The noise variables are conditionally μ-sub-Gaussian: for every $t \in [T]$,*

$$\mathbb{E}[\eta_t \mid \mathcal{F}_{t-1}] = 0 \quad \textit{and} \quad \mathbb{E}\left[\exp\left(\upsilon\, \eta_t\right) \mid \mathcal{F}_{t-1}\right] \leq \exp\left(\frac{\mu^2 \upsilon^2}{2}\right) \quad \textit{for all } \upsilon \in \mathbb{R}.$$

The sub-Gaussian property ensures the concentration of self-normalized martingales, which we shall leverage to control the cumulative effect of noise under adaptive design.

The geometry of the action set. We consider a generic action set $\mathcal{X}$ and impose the same assumptions as in Abeille et al. (2025), which are expressed in terms of the support function of $\mathcal{X}$,

$$J_{\mathcal{X}}(\theta) := \max_{x \in \mathcal{X}} \langle \theta, x \rangle.$$

The optimal expected reward is denoted by $J_\star := J_{\mathcal{X}}(\theta^\star)$. The support function plays a central role in the analysis: whenever it is differentiable at θ_t, the action selected by the algorithm is $x_t = \nabla J_{\mathcal{X}}(\theta_t)$, so the instantaneous regret can be written entirely in terms of $J_{\mathcal{X}}$,

$$r_t = J_{\mathcal{X}}(\theta^\star) - \langle \theta^\star, x_t \rangle = J_{\mathcal{X}}(\theta^\star) - J_{\mathcal{X}}(\theta_t) - \langle \nabla J_{\mathcal{X}}(\theta_t), \theta^\star - \theta_t \rangle = D_{J_{\mathcal{X}}}(\theta^\star, \theta_t),$$

where $D_f(u, v) := f(u) - f(v) - \langle \nabla f(v), u - v \rangle$ denotes the Bregman divergence of a differentiable convex function f.

(A2) (Smooth, Strongly Convex Action Set). *There exist a norm $\|\cdot\|_*$ on $\mathbb{R}^d$ and constants $0 < m \le M$ such that $J_{\mathcal{X}}^2$ is differentiable on $\mathbb{R}^d$ and, for all $\theta, \theta' \in \mathbb{R}^d$,*

$$\frac{m}{2}\|\theta - \theta'\|_*^2 \le D_{J_{\mathcal{X}}^2}(\theta, \theta') \le \frac{M}{2}\|\theta - \theta'\|_*^2; \tag{1}$$

moreover, $\mathcal{X}$ is absorbing, i.e., it contains a neighborhood of the origin.

Assumption (A2) admits a simple geometric interpretation: the boundary of $\mathcal{X}$ must be uniformly curved, with neither flat faces nor corners. Polytopes are thus excluded.

Strong convexity and smoothness are imposed on the squared support function $J_{\mathcal{X}}^2$ rather than on $J_{\mathcal{X}}$ itself because its gradient, $\nabla J_{\mathcal{X}}^2(\theta) = 2J_{\mathcal{X}}(\theta)\nabla J_{\mathcal{X}}(\theta)$, remains well behaved as $\theta \to 0$, where it vanishes, whereas $\nabla J_{\mathcal{X}}$ admits no limit at the origin. The absorbing property ensures that the multiplier $J_{\mathcal{X}}(\theta)$ is positive for every $\theta \neq 0$. Assumption (A2) covers the following geometries:

- *Unit ball.* For $\mathcal{X} = \mathbb{B}_2^d(1)$, the support function is $J_{\mathcal{X}}(\theta) = \|\theta\|_2$ and $D_{J_{\mathcal{X}}^2}(\theta, \theta') = \|\theta - \theta'\|_2^2$; hence Assumption (A2) holds with $\|\cdot\|_* = \|\cdot\|_2$ and $m = M = 2$.
- *Ellipsoids.* For $\mathcal{X} = \{x : x^\top \mathbf{A}^{-1} x \le 1\} = \mathbf{A}^{1/2}\mathbb{B}_2^d(1)$ with a positive definite matrix $\mathbf{A}$, the support function is $J_{\mathcal{X}}(\theta) = \|\theta\|_{\mathbf{A}}$. Under the adapted norm $\|u\|_* := \|\mathbf{A}^{1/2}u\|_2$, $D_{J_{\mathcal{X}}^2}(\theta, \theta') = \|\theta - \theta'\|_{\mathbf{A}}^2 = \|\theta - \theta'\|_*^2$, so that, once again, $m = M = 2$.
- *C^2 bodies.* More generally, the assumption holds for $\mathcal{X} \subseteq \mathbb{B}_2^d(1)$ containing the origin in its interior with a C^2 boundary whose principal radii are bounded away from zero.

See Section B for more details, including the parameters M, m for the C^2 body example.

4 Main Results

We now present our regret upper bound for α-LinUCB. Its statement uses the standard concentration width β and the concentration event $\mathcal{E}$ of Abbasi-Yadkori et al. (2011), which we first reproduce.

Proposition 4.1 (Confidence event; Abbasi-Yadkori et al., 2011)**.** *Let Assumption (A1) hold, let $\|\theta^\star\|_2 \le S^\star$ and $\mathcal{X} \subseteq \mathbb{B}_2^d(1)$, and let $\hat{\theta}_t$ be the ridge estimate with regularisation $\lambda_0 > 0$. Fix a horizon T and a confidence level $\delta \in (0, 1)$, and define*

$$\beta := \mu\sqrt{2\log\left(\frac{1}{\delta}\right) + d\log\left(\left(1 + \frac{T}{d\lambda_0}\right)\right)} + \sqrt{\lambda_0}\, S^\star,$$
$$\mathcal{E} := \left\{ \|\hat{\theta}_t - \theta^\star\|_{V_t} \le \beta \text{ for all } t \in [T] \right\}.$$

Then $\mathbb{P}(\mathcal{E}) \ge 1 - \delta$.

The following regret bound holds for all $\alpha > 0$; the confidence width β and the event $\mathcal{E}$ are used in the analysis only.

Theorem 4.2 (α-LinUCB regret)**.** *Let Assumptions (A1) and (A2) hold. On the event $\mathcal{E}$, hence with probability at least $1 - \delta$,*

$$\mathcal{R}_T = \mathcal{O}\Big(\frac{M}{m}\,\frac{\alpha^2 + \beta^2}{\alpha}\,\sqrt{dT}\Big).$$

Theorem 4.2 shows that α-LinUCB has sublinear regret for every $\alpha > 0$, including values for which $\mathcal{C}_t(\alpha)$ is not a valid confidence set. The bound depends on the geometry of the action set through the ratio M/m, similarly to the bound of Abeille et al. (2025) for randomised exploration algorithms.

As α decreases, the set $\mathcal{C}_t(\alpha)$ shrinks towards $\{\hat{\theta}_t\}$ and the algorithm approaches the greedy policy; accordingly, the guarantee degrades. Likewise, the bound deteriorates as $\alpha \to \infty$, when the algorithm fails to exploit known information. Optimising the bound over α gives $\alpha = \beta$, and one recovers standard LinUCB, but with a regret bound of $\mathcal{O}((M/m)\, d\sqrt{T\log(T)})$, a $\sqrt{\log(T)}$ improvement over the $\mathcal{O}(d\sqrt{T}\log(T))$ upper bound of Abbasi-Yadkori et al. (2011).

The proof of Theorem 4.2 does not rely on the optimism inequality $\langle \theta_t, x_t\rangle \geq \langle \theta^\star, x^\star\rangle$, which is no longer guaranteed when $\theta^\star \notin \mathcal{C}_t(\alpha)$ but instead examines the exploration dynamics of α-LinUCB. The following bound shows that, in a precise sense, the algorithm explores every direction for every $\alpha > 0$. It controls the squared $\|\cdot\|_*$-width of the unit V_t-ellipsoid,

$$\omega_t := \max_{u \in \mathbb{B}_2^d(1)} \|V_t^{-1/2} u\|_*^2 = \max_{\|g\|_*^\circ \leq 1} \|g\|_{V_t^{-1}}^2,$$

where $\|\cdot\|_*^\circ := \sup_{\|v\|_* \leq 1} \langle \cdot, v\rangle$ denotes the dual norm of $\|\cdot\|_*$. When $\mathcal{X}$ is the unit ball and $\|\cdot\|_* = \|\cdot\|_2$, one has $\omega_t = \max_{u \in \mathbb{B}_2^d(1)} \|V_t^{-1/2} u\|_2^2 = \lambda_{\min}(V_t)^{-1}$, so that upper bounding ω_t amounts to lower bounding the smallest eigenvalue of the design, i.e. the growth of the least-explored direction.

Theorem 4.3 (Spectral Growth). *Let Assumptions (A1) and (A2) hold, let $\mathcal{E}$ and β be as in Proposition 4.1, and let $T_B = \mathcal{O}(\mathrm{poly}(\log T))$ be the burn-in threshold. Then, on the event $\mathcal{E}$ (hence with probability at least $1-\delta$), for all $T_B \leq t \leq T$,*

$$\omega_t = \max_{u \in \mathbb{B}_2^d(1)} \|V_t^{-1/2} u\|_*^2 \leq \frac{1}{c_0\sqrt{t}}, \qquad \textit{where} \qquad c_0 = \Omega\Big(\frac{m\alpha}{J_\star\sqrt{d}}\Big).$$

The bound is *deterministic on* $\mathcal{E}$ and holds *at every round* after the burn-in, not on average or asymptotically; it is stated in the geometry of the assumption, with a constant proportional to $m\alpha/(J_\star\sqrt{d})$; and it holds for *every* $\alpha > 0$, so the confidence set need not contain $\theta^\star$ for the algorithm to explore.

On $\mathcal{E}$, the spectral growth bound shows immediately the following deviation bounds:

$$\|\theta_t - \hat{\theta}_t\|_* \leq \alpha\sqrt{\omega_t} \leq \frac{\alpha}{\sqrt{c_0}\, t^{1/4}} \qquad \text{and} \qquad \|\hat{\theta}_t - \theta^\star\|_* \leq \beta\sqrt{\omega_t} \leq \frac{\beta}{\sqrt{c_0}\, t^{1/4}},$$

so that the $\|\cdot\|_*$-radius of the confidence set and the estimation error both decay at the rate $t^{-1/4}$, for all $\alpha > 0$. In the unit-ball case, the bound reads $\lambda_{\min}(V_t) = \Omega\big(\frac{\alpha}{\sqrt{d}}\sqrt{t}\big)$. For $\alpha \asymp \beta \asymp \sqrt{d\log T}$, this improves on the $\Omega\big(\sqrt{t}/(d\log t)\big)$ lower bound for LinUCB implied by Banerjee et al. (2023).

5 Proof Sketch

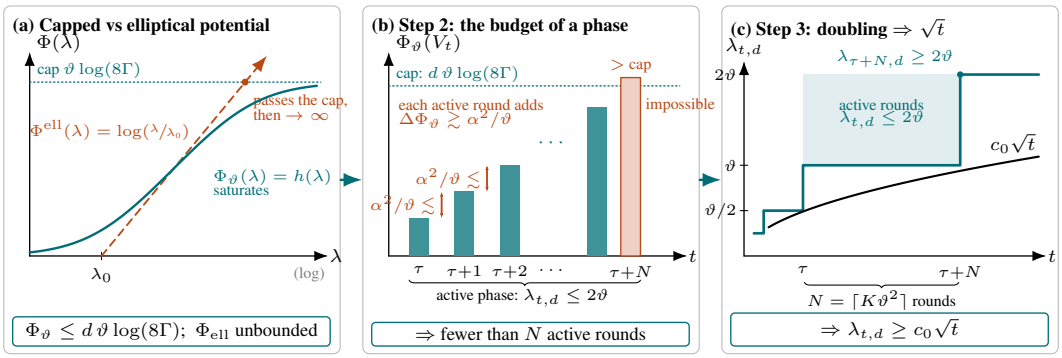


Figure 1: The three steps of the proof of Theorem 4.3 in the unit-ball case.

The detailed proofs of Theorems 4.2 and 4.3 are given in Sections D and E. The argument has two parts: an energy argument, which shows that optimizing over $\mathcal{C}_t(\alpha)$ forces the design matrix to grow in every direction, and a regret decomposition, which uses this growth in place of the optimism inequality of the standard analysis. We now sketch the two proofs, starting with spectral growth.

5.1 SPECTRAL GROWTH

For ease of presentation, we sketch the proof in the unit-ball case, i.e. $\mathcal{X} = \mathbb{B}_2^d(1)$. Take $\|\cdot\|_* = \|\cdot\|_2$, so that $\omega_t = \lambda_{t,d}^{-1}$. Assume that $\theta^\star \neq 0$. The proof follows an energy argument, which tracks the amount of energy that α-LinUCB injects into the least-explored directions. Let $\lambda_{t,1} \geq \cdots \geq \lambda_{t,d}$ denote the eigenvalues of V_t, and $v_{t,1}, \ldots, v_{t,d}$ the eigenvectors. We define the energy of an action x_t in the ith direction, $i \in [d]$, as $x_{t,i}^2/\lambda_{t,i}$, where $x_{t,i} := \langle x_t, v_{t,i}\rangle$. The argument has three steps, summarized in Figure 1.

Step 1. We first define a *logarithmically capped potential*, rather than the usual elliptical potential (denoted by Φ_{ell} in Figure 1). Fix a reference scale $\vartheta > 0$ and let $\Gamma \geq 1$ (with a specific value of Γ fixed in Section D.3).

Definition 5.1 (Capped potential)**.** *For the scale ϑ, set $L := 8\Gamma\vartheta$ and define, for every $V \succeq 0$,*

$$\Phi_\vartheta(V) := \operatorname{tr} h(V), \qquad h(\lambda) := \vartheta \log\left(\frac{L\,(\lambda+\vartheta)}{\vartheta\,(\lambda+L)}\right), \qquad \lambda \geq 0. \tag{2}$$

The map Φ_ϑ sends positive semidefinite matrices to $\mathbb{R}$ and is a rescaled trace in which each eigenvalue is counted through h. It can alternatively be expressed as a difference of ϑ- and L-regularised log-determinants,

$$\Phi_\vartheta(V) = \vartheta\big(\log\det(V+\vartheta\mathbf{I}) - \log\det(V+8\Gamma\vartheta\mathbf{I}) + d\log(8\Gamma)\big),$$

where the additive constant $d\log(8\Gamma)$ ensures that $\Phi_\vartheta(0) = 0$. The growths of the two log-determinant terms along directions of V that significantly exceed $8\Gamma\vartheta$ essentially cancel out, so that the potential is capped; indeed

$$0 \;\leq\; \Phi_\vartheta(V) \;\leq\; d\,\vartheta\log(8\Gamma) \qquad \text{for every } V \succeq 0.$$

Step 2. After a burn-in of polylogarithmic length T_B (see Definition D.10), on the event $\mathcal{E}$, the whole set $\mathcal{C}_t(\alpha)$ lies in the neighbourhood $\{\theta : \|\theta-\theta^\star\| \leq \|\theta^\star\|/2\}$ (Proposition D.8). In particular, $\frac{1}{2}\|\theta^\star\| \leq \|\theta\| \leq \frac{3}{2}\|\theta^\star\|$ for all $\theta \in \mathcal{C}_t(\alpha)$, and the played arms satisfy $\langle x_t, x^\star\rangle \geq 1/2$ for every $t \geq T_B$. Since $x_t \in \mathbb{B}_2^d(1)$,

$$\|\theta_t\| = \langle\hat\theta_t, x_t\rangle + \alpha\|x_t\|_{V_t^{-1}} \leq \|\hat\theta_t\| + \alpha\|x_t\|_{V_t^{-1}} \implies \alpha\|x_t\|_{V_t^{-1}} \geq \|\theta_t\| - \|\hat\theta_t\|,$$

where the equality follows from the definition of (x_t, θ_t) together with $x_t = \theta_t/\|\theta_t\|$. Next, by definition, $\hat\theta_t \pm \alpha v_{t,d}/\sqrt{\lambda_{t,d}} \in \mathcal{C}_t(\alpha)$, and since θ_t maximises the norm over $\mathcal{C}_t(\alpha)$, averaging over the two signs gives

$$\|\theta_t\|^2 \geq \frac{1}{2}\|\hat\theta_t + \frac{\alpha}{\sqrt{\lambda_{t,d}}}v_{t,d}\|^2 + \frac{1}{2}\|\hat\theta_t - \frac{\alpha}{\sqrt{\lambda_{t,d}}}v_{t,d}\|^2 = \|\hat\theta_t\|^2 + \frac{\alpha^2}{\lambda_{t,d}}.$$

Combining the above with the previous display, we have

$$\alpha\|x_t\|_{V_t^{-1}} \geq \|\theta_t\| - \|\hat\theta_t\| = \frac{\|\theta_t\|^2 - \|\hat\theta_t\|^2}{\|\theta_t\| + \|\hat\theta_t\|} \geq \frac{\alpha^2}{3\|\theta^\star\|\lambda_{t,d}},$$

where the last inequality uses that $\theta_t, \hat\theta_t \in \mathcal{C}_t(\alpha)$, so that $\|\theta_t\| + \|\hat\theta_t\| \leq 3\|\theta^\star\|$ by assumption.

Fix a time-step τ and let $\vartheta := \lambda_{\tau,d}$, the smallest eigenvalue at round τ, and let τ' be the last time-step for which $\lambda_{t,d} \leq 2\vartheta$. Call $t \in [\tau, \tau']$ the active time-steps.

We show in Lemma D.11 that for every round $t \geq \tau$,

$$\Phi_\vartheta(V_{t+1}) - \Phi_\vartheta(V_t) \;\geq\; \frac{\vartheta}{4}\sum_{\substack{i\geq 2\\ \lambda_{t,i}<\Gamma\vartheta}} \frac{x_{t,i}^2}{\lambda_{t,i}},$$

a straightforward algebraic bound. Now, using our inequality on $\alpha\|x_t\|_{V_t^{-1}}$,

$$\frac{x_{t,1}^2}{\lambda_{t,1}} + \sum_{\substack{i\geq 2 \\ \lambda_{t,i}\geq \Gamma\vartheta}} \frac{x_{t,i}^2}{\lambda_{t,i}} + \sum_{\substack{i\geq 2 \\ \lambda_{t,i}< \Gamma\vartheta}} \frac{x_{t,i}^2}{\lambda_{t,i}} = \|x_t\|_{V_t^{-1}}^2 \;\geq\; \frac{\alpha^2}{9\|\theta^\star\|^2\lambda_{t,d}^2} \tag{3}$$

If we show that the first two terms are small relative to the lower bound, we can conclude that

$$\Phi_\vartheta(V_{t+1}) - \Phi_\vartheta(V_t) \gtrsim \frac{\vartheta\alpha^2}{\|\theta^\star\|^2\lambda_{t,d}^2}, \tag{4}$$

where $\gtrsim$ denotes inequality up to absolute constants; this is precisely what is established in Lemma D.12 and Corollary D.13. Indeed, let T_B denote the first time-step after the burn-in, which ensures that $\langle x_t, x^\star\rangle \geq 1/2$ for every $t \geq T_B$, and assume $t \geq 2T_B$. Then

$$\lambda_{t,1} \geq \|x^\star\|_{V_t}^2 \geq \sum_{s=T_B}^{t-1} \langle x_s, x^\star\rangle^2 \geq \frac{t-T_B}{4} \geq \frac{t}{8} \implies \frac{x_{t,1}^2}{\lambda_{t,1}} \leq \frac{1}{\lambda_{t,1}} \leq \frac{8}{t}.$$

The upper bound $8/t$ is negligible relative to the lower bound of equation 3 as long as $\lambda_{t,d} = o(\sqrt{t})$; and once that is no longer true, the conclusion of the theorem already holds at round t. For the remaining term, a less immediate argument (step (b) of the proof of Lemma D.12) yields $\sum_{i\geq 2} x_{t,i}^2 \leq C_T/\lambda_{t,d}$, where $C_T > 0$ is polylogarithmic in T, and so

$$\sum_{\substack{i\geq 2 \\ \lambda_{t,i}\geq \Gamma\vartheta}} \frac{x_{t,i}^2}{\lambda_{t,i}} \leq \frac{2C_T}{\Gamma\lambda_{t,d}^2}.$$

The choice of Γ is precisely what makes this upper bound small relative to the lower bound of (3).

Step 3. Set $N := \lceil K\vartheta^2\rceil$ with $K \gtrsim d\log(8\Gamma)/\alpha^2$, and suppose that all rounds $\tau, \ldots, \tau+N-1$ are active. By equation 4 with $\lambda_{t,d} \leq 2\vartheta$, each of them increases the potential by $\gtrsim \alpha^2/\vartheta$ (absorbing $\|\theta^\star\|$ into the constant), hence

$$\Phi_\vartheta(V_{\tau+N}) - \Phi_\vartheta(V_\tau) \;\gtrsim\; N\,\frac{\alpha^2}{\vartheta} \;>\; \frac{d\log(8\Gamma)}{\alpha^2}\,\vartheta^2 \cdot \frac{\alpha^2}{\vartheta} \;=\; d\,\vartheta\log(8\Gamma),$$

which, for a suitable choice of K, exceeds the cap $d\vartheta\log(8\Gamma)$ of the potential: a contradiction. Hence the active phase ends before round $\tau+N$, i.e. the smallest eigenvalue doubles within $N = O(d\,\vartheta^2\log(8\Gamma)/\alpha^2)$ rounds: $\lambda_{\tau+N,d} \;\geq\; 2\lambda_{\tau,d}$. Finally, a backward induction on the first round t at which $\lambda_{t,d} < c_0\sqrt{t}$ would fail turns this doubling into $\lambda_{t,d} \geq c_0\sqrt{t}$ (see Proposition D.16).

While we chose to illustrate the proof sketch in the unit-ball case, the appendix presents the proof in the general case, in the geometry of $\|\cdot\|_*$. The key is a change of variables in the spirit of John's theorem (see Section D.2): the Euclidean structure needed for the spectral bookkeeping is chosen *adapted* to $\|\cdot\|_*$, so that the two norms are compared only through the distortion $D_* \leq \min\{\sqrt{d}, \sqrt{M/m}\}$ of the unit ball of $\|\cdot\|_*$, which enters the bounds only inside logarithms.

5.2 From spectral growth to regret

The argument has two steps. First, after the polylogarithmic burn-in phase T_B(see Definition D.10), the parameters θ_t lie in a neighbourhood of $\theta^\star$ on which $J_\mathcal{X}$ is comparable to $J_\star$ (Proposition D.8), and the instantaneous regret is controlled by the Bregman divergence of $J_\mathcal{X}^2$ (Lemma E.1):

$$r_t \;\leq\; \frac{D_{J_\mathcal{X}^2}(\theta^\star, \theta_t)}{J_\star}.$$

This upper bound does not require $\theta^\star \in \mathcal{C}_t(\alpha)$. Second, the divergence is bounded by the width ω_t, which Theorem 4.3 controls; summing over rounds then bounds the regret directly, and in particular the elliptical potential lemma, with the additional $\sqrt{\log(T)}$ factor it introduces, is never used to sum the widths.

In detail, the horizon is split at the burn-in time T_B. Before T_B, the instantaneous regret r_t is bounded by a constant, owing to the boundedness of $\theta^\star$ and of the action set. From T_B on, the

smoothness half of equation 1 and the definition of the confidence sets reduce the divergence to a multiple of ω_t, and Theorem 4.3 bounds the sum of the widths by $2\sqrt{T}/c_0$. On the event $\mathcal{E}$,

$$\begin{aligned}
\mathcal{R}_T &\le \mathcal{O}(T_B) + \sum_{t=T_B}^{T} \frac{D_{J_{\mathcal{X}}^2}(\theta^\star, \theta_t)}{J_\star} && \text{(Lemma E.1)} \\
&\le \mathcal{O}(T_B) + \frac{M}{2J_\star} \sum_{t=T_B}^{T} \|\theta^\star - \theta_t\|_*^2 && \text{(Assumption (A2))} \\
&\le \mathcal{O}(T_B) + \frac{M}{J_\star} \sum_{t=T_B}^{T} \Big(\|\theta^\star - \hat{\theta}_t\|_*^2 + \|\hat{\theta}_t - \theta_t\|_*^2 \Big) && \big((a+b)^2 \le 2a^2 + 2b^2\big) \\
&\le \mathcal{O}(T_B) + \frac{M(\alpha^2+\beta^2)}{J_\star} \sum_{t=T_B}^{T} \omega_t && \big(\theta^\star \in \mathcal{C}_t(\beta),\ \theta_t \in \mathcal{C}_t(\alpha)\big) \\
&\le \mathcal{O}(T_B) + \frac{M(\alpha^2+\beta^2)}{J_\star\, c_0} \sum_{t=T_B}^{T} \frac{1}{\sqrt{t}} && \text{(Theorem 4.3)} \\
&\le \mathcal{O}(T_B) + \frac{2M(\alpha^2+\beta^2)}{J_\star\, c_0} \sqrt{T} = \mathcal{O}\Big(\frac{M}{m} \frac{\alpha^2+\beta^2}{\alpha} \sqrt{dT} \Big),
\end{aligned}$$

where the last step uses $c_0 = \Omega(m\alpha/(J_\star\sqrt{d}))$ and $T_B = \mathcal{O}(\mathrm{poly}(\log T))$.

6 Curvature matters: a failure case

Theorems 4.2 and 4.3 hold for *every* radius $\alpha > 0$, and Section 5 explained what the curvature imposed by Assumption (A2) does for the algorithm: a perturbation of the optimistic parameter is transmitted to the played arm. On a body with a corner this transmission stops, and the next proposition shows that the conclusion of Theorems 4.2 and 4.3 then fails for every small radius: α-LinUCB locks on a suboptimal vertex with probability bounded away from zero, the design matrix does not grow, and the regret is linear. The instance is the simplest possible (presented in Figure 2): in $\mathbb{R}^2$, with canonical basis (e_1, e_2), for $a \in (0,1)$, let

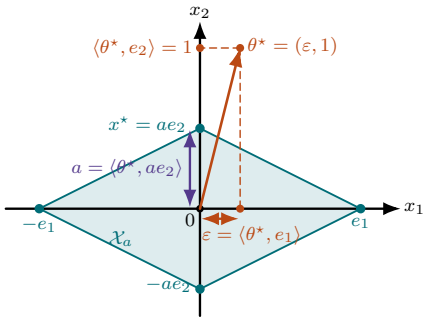


Figure 2: Counter-example.

$$\mathcal{X}_a := \mathrm{conv}\{e_1, -e_1, ae_2, -ae_2\}, \qquad \theta^\star := (\varepsilon, 1) \quad \text{with } 0 < \varepsilon < a, \qquad \eta_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \mu^2),$$

where conv denotes the convex hull; that is, $\mathcal{X}_a$ is the diamond with vertices $\pm e_1$ and $\pm ae_2$.

The set $\mathcal{X}_a$ is a convex body containing the origin in its interior and contained in $\mathbb{B}_2^2(1)$, whose support function is $J_{\mathcal{X}_a}(\theta) = \max\{|\theta_1|, a|\theta_2|\}$ for $\theta = (\theta_1, \theta_2)$; it violates Assumption (A2) at its four corners and four flat faces (Figure 4(b)). The optimal arm is $x^\star = ae_2$, with $J_\star = a$, and the gap of the two vertices $\pm e_1$ is $\Delta := a - \varepsilon > 0$.

Proposition 6.1 (Linear regret on a polytope). *In the instance above, for every $\lambda_0 > 0$, every horizon T and every radius*

$$0 < \alpha \le \alpha_0 := \frac{(1-a^2)\sqrt{\lambda_0}}{2a}\,\varepsilon,$$

with probability at least

$$p_0 := \frac{1}{2}\Big(1 - \exp\Big(-\frac{(1-a^2)\lambda_0\varepsilon^2}{2a^2\mu^2}\Big)\Big) > 0,$$

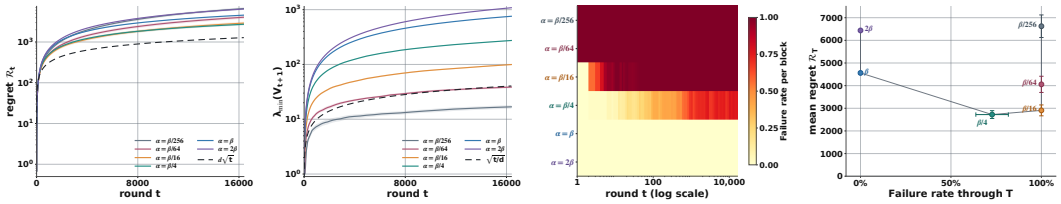
α-LinUCB plays $x_t \in \{e_1, -e_1\}$ at every round $t \in [T]$. On this event $\lambda_{\min}(V_t) = \lambda_0$ for all t and $\mathcal{R}_T \ge \Delta T$; in particular $\mathbb{E}[\mathcal{R}_T] \ge p_0 \Delta T$.

The proof is given in Section B.1. For instance, $a = 1/2$, $\varepsilon = 1/4$, $\lambda_0 = 1$ and Gaussian noise with $\mu = 1/8$ give $p_0 > 0.49$: for every $\alpha \leq 3/16$, half of the runs never play the optimal arm and suffer $\mathcal{R}_T \geq T/4$. The curvature imposed by Assumption (A2) is precisely what excludes such lock-ins, however small the radius; we do not claim that it is necessary, only sufficient for our analysis.

7 Numerical Simulations

While the contribution of this paper is mainly theoretical, we corroborate our results with simulations. We run α-LinUCB on the unit ball for $d = 10$, $\theta^\star = \mathbf{1}_d/\sqrt{d}$, Gaussian noise, $\beta = 20$ and $T = 16{,}300$, with $\alpha \in \{\beta/256, \beta/64, \beta/16, \beta/4, \beta, 2\beta\}$. For each α we run 48 independent replications with the same noise streams, and report in Figure 3 the mean of the pseudo-regret $\mathcal{R}_t$, of $\lambda_{\min}(V_{t+1})$ and of the failure frequency of the confidence set, $\theta^\star \notin \mathcal{C}_t(\alpha)$, with pointwise approximate 95% confidence intervals shaded.

In accordance with Theorem 4.2, the regret scales as $\sqrt{t}$ for every α (panel (a)), and the lowest regret is attained at $\alpha = \beta/4$, not at the valid radius β: the latter minimises the bound, not the regret. In accordance with Theorem 4.3, $\lambda_{\min}(V_{t+1})$ grows like $\sqrt{t}$ for every α, with a constant increasing in α (panel (b)). Panel (c) confirms that the sets are invalid for small radii: for $\alpha \leq \beta/16$ the confidence set fails at almost every round, whereas it never fails for $\alpha \in \{\beta, 2\beta\}$. Finally, panel (d) shows that the trade-off between regret and validity is best at $\alpha = \beta/4$: its set fails in about three quarters of the rounds, yet it achieves $\approx 40\%$ less regret than $\alpha = \beta$.



(a) Regret (b) Minimum eigenvalue (c) Confidence failures (d) Failure frequency vs. regret

Figure 3: Effect of the confidence radius on the Euclidean unit ball for $d = 10$, $\beta = 20$, and $T = 16{,}300$, with $\alpha \in \{\beta/256, \beta/64, \beta/16, \beta/4, \beta, 2\beta\}$. Curves average 48 independent runs; shaded bands are pointwise 95% confidence intervals for their means. The dashed guides in (a) and (b) show $d\sqrt{t}$ and $\sqrt{t/d}$ without fitted constants; they are not finite-time bounds. In (c), a failure means $\theta^\star \notin C_t(\alpha)$, and color shows its frequency in each time block. In (d), bars show pointwise 95% confidence intervals for run means.

Conclusion

This paper studied the α-LinUCB family of algorithms for stochastic linear bandits, in which the exploration radius α is treated as a free parameter rather than as a statistical confidence width. Under a curvature assumption on the action set, we established a sublinear regret bound that holds for every $\alpha > 0$, irrespective of whether $\mathcal{C}_t(\alpha)$ is a valid confidence set. The bound rests on a fine-grained analysis of the spectral growth of the design matrix, which shows that α-LinUCB explores every direction at a rate that we quantify explicitly. This analysis serves two purposes. It substitutes for the optimism inequality, which is no longer guaranteed once $\theta^\star \notin \mathcal{C}_t(\alpha)$, and thereby extends the guarantee to arbitrary radii; and it circumvents the elliptical potential lemma, thereby improving by a factor $\sqrt{\log(T)}$ the best known regret bound for standard LinUCB.

A single factor $\sqrt{\log(T)}$ nonetheless separates our bound from the $\Omega(d\sqrt{T})$ lower bound, and it is now inherited from the concentration inequality for the adaptive design alone. While Lattimore (2023) rules out the removal of this factor in the general adaptive case, it may still be possible to dispense with it within the family of algorithms considered here, now that the dynamics of the algorithm, and the rate at which each direction of the design matrix grows, are better understood. This question is both interesting and far from straightforward, and we leave it for future work.

AI USE STATEMENT

In this work, we used generative AI tools to edit the manuscript for readability and to check spelling, grammar, and the consistency of notation and citations. We have not used generative AI tools to develop theoretical models or conceptual frameworks, formulate mathematical claims, provide ingredients for or write proofs, propose or refine hypotheses, design or provide feedback on the methodology or experiments, implement methods, or interpret results; the remaining tasks with required disclosure (generating synthetic data, translation, dataset cleaning, and qualitative data analysis) are not applicable to this work. We have reviewed all AI-assisted work: every suggested edit was checked by the authors against the original text and the mathematical content. We take responsibility for the final content of this work, including text, claims, or artifacts produced with the aid of generative AI.

REPRODUCIBILITY STATEMENT

This work is primarily theoretical. All assumptions required to derive our results are precisely stated and discussed in the main text (Assumptions (A1) and (A2)). Proof sketches are provided in the main text, and complete proofs are included in the appendices. The full protocol of the simulations (instance, parameters, number of replications) is specified in Section 7.

REFERENCES

Yasin Abbasi-Yadkori, Dávid Pál, and Csaba Szepesvári. Improved algorithms for linear stochastic bandits. In *Advances in Neural Information Processing Systems*, 2011.

Marc Abeille and Alessandro Lazaric. Linear Thompson sampling revisited. In *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*, 2017.

Marc Abeille, David Janz, and Ciara Pike-Burke. When and why randomised exploration works (in linear bandits). In *Proceedings of the 36th International Conference on Algorithmic Learning Theory*, 2025.

Shipra Agrawal and Navin Goyal. Thompson sampling for contextual bandits with linear payoffs. In *Proceedings of the 30th International Conference on Machine Learning*, 2013.

Debangshu Banerjee, Avishek Ghosh, Sayak Ray Chowdhury, and Aditya Gopalan. Exploration in linear bandits with rich action sets and its implications for inference. In *Proceedings of the 26th International Conference on Artificial Intelligence and Statistics*, 2023.

Hamsa Bastani, Mohsen Bayati, and Khashayar Khosravi. Mostly exploration-free algorithms for contextual bandits. *Management Science*, 67(3):1329–1349, 2021.

Amir Beck. *First-Order Methods in Optimization*. Society for Industrial and Applied Mathematics, 2017.

Varsha Dani, Thomas P. Hayes, and Sham M. Kakade. Stochastic linear optimization under bandit feedback. In *Proceedings of the 21st Annual Conference on Learning Theory*, 2008.

Yash Deshpande, Lester Mackey, Vasilis Syrgkanis, and Matt Taddy. Accurate inference for adaptive linear models. In *Proceedings of the 35th International Conference on Machine Learning*, 2018.

Hamish Flynn, David Reeb, Melih Kandemir, and Jan R. Peters. Improved algorithms for stochastic linear bandits using tail bounds for martingale mixtures. In *Advances in Neural Information Processing Systems*, 2023.

Botao Hao, Tor Lattimore, and Csaba Szepesvári. Adaptive exploration in linear contextual bandit. In *Proceedings of the 23rd International Conference on Artificial Intelligence and Statistics*, 2020.

Ruitong Huang, Tor Lattimore, András György, and Csaba Szepesvári. Following the leader and fast rates in linear prediction: Curved constraint sets and other regularities. In *Advances in Neural Information Processing Systems*, 2016.

Sampath Kannan, Jamie Morgenstern, Aaron Roth, Bo Waggoner, and Zhiwei Steven Wu. A smoothed analysis of the greedy algorithm for the linear contextual bandit problem. In *Advances in Neural Information Processing Systems*, 2018.

Koulik Khamaru, Yash Deshpande, Tor Lattimore, Lester Mackey, and Martin J. Wainwright. Near-optimal inference in adaptive linear regression. *The Annals of Statistics*, 53(6):2329–2355, 2025.

Tze Leung Lai and Ching Zong Wei. Least squares estimates in stochastic regression models with applications to identification and control of dynamic systems. *The Annals of Statistics*, 10(1): 154–166, 1982.

Tor Lattimore. A lower bound for linear and kernel regression with adaptive covariates. In *Proceedings of the 36th Conference on Learning Theory*, 2023.

Tor Lattimore and Csaba Szepesvári. The end of optimism? an asymptotic analysis of finite-armed linear bandits. In *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*, 2017.

Tor Lattimore and Csaba Szepesvári. *Bandit Algorithms*. Cambridge University Press, 2020.

Lihong Li, Wei Chu, John Langford, and Robert E. Schapire. A contextual-bandit approach to personalized news article recommendation. In *Proceedings of the 19th International Conference on World Wide Web*, 2010.

Licong Lin, Mufang Ying, Suvrojit Ghosh, Koulik Khamaru, and Cun-Hui Zhang. Statistical limits of adaptive linear models: Low-dimensional estimation and inference. In *Advances in Neural Information Processing Systems*, 2023.

Samya Praharaj and Koulik Khamaru. Avoiding the price of adaptivity: Inference in linear contextual bandits via stability. arXiv:2512.20368v2, 2026.

Manish Raghavan, Aleksandrs Slivkins, Jennifer Wortman Vaughan, and Zhiwei Steven Wu. Greedy algorithm almost dominates in smoothed contextual bandits. *SIAM Journal on Computing*, 52(2): 487–524, 2023.

R. Tyrrell Rockafellar and Roger J.-B. Wets. *Variational Analysis*. Springer, 1998.

Paat Rusmevichientong and John N. Tsitsiklis. Linearly parameterized bandits. *Mathematics of Operations Research*, 35(2):395–411, 2010.

Rolf Schneider. *Convex Bodies: The Brunn–Minkowski Theory*. Cambridge University Press, 2 edition, 2014.

Andrea Tirinzoni, Matteo Pirotta, Marcello Restelli, and Alessandro Lazaric. An asymptotically optimal primal-dual incremental algorithm for contextual linear bandits. In *Advances in Neural Information Processing Systems*, 2020.

Jean Ville. *Étude critique de la notion de collectif*. PhD thesis, Université de Paris, 1939.

Mufang Ying, Koulik Khamaru, and Cun-Hui Zhang. Adaptive linear estimating equations. In *Advances in Neural Information Processing Systems*, 2023.

APPENDIX CONTENTS

A Self-normalised bound for adaptive designs 13
B Further properties of the action set 13
B.1 Failure case 15
C Pseudo-code of α-LinUCB 16
D Spectral growth 17
D.1 Preliminary results 17
D.2 Geometric distortion and adapted coordinates 19
D.3 Definition of the constants 22
D.4 Exploration engine, burn-in and crude bound 22
D.5 Proof of Theorem 4.3 26
E Regret of α-LinUCB 31
E.1 Regret decomposition 31
E.2 Proof of Theorem 4.2 31

A SELF-NORMALISED BOUND FOR ADAPTIVE DESIGNS

Proposition 4.1 (Confidence event; Abbasi-Yadkori et al., 2011)**.** *Let Assumption (A1) hold, let $\|\theta^\star\|_2 \le S^\star$ and $\mathcal{X} \subseteq \mathbb{B}_2^d(1)$, and let $\hat\theta_t$ be the ridge estimate with regularisation $\lambda_0 > 0$. Fix a horizon T and a confidence level $\delta \in (0,1)$, and define*

$$\beta \;:=\; \mu\sqrt{2\log\!\left(\frac{1}{\delta}\right) + d\log\!\left(\left(1+\frac{T}{d\lambda_0}\right)\right)} \;+\; \sqrt{\lambda_0}\,S^\star,$$
$$\mathcal{E} \;:=\; \Big\{\, \|\hat\theta_t - \theta^\star\|_{V_t} \le \beta \;\textit{for all } t \in [T] \,\Big\}.$$

Then $\mathbb{P}(\mathcal{E}) \ge 1-\delta$.

Proposition 4.1 is Theorem 2 of Abbasi-Yadkori et al. (2011), applied with noise parameter μ, $\|x_s\|_2 \le 1$, regularisation λ_0 and $\|\theta^\star\|_2 \le S^\star$; the radius there is $\mu\sqrt{2\log((1/\delta)) + \log((\det(V_t)\,/\,\det(\lambda_0\mathbf{I})))} + \sqrt{\lambda_0}S^\star$, and $\det(V_t) \le (\lambda_0 + t/d)^d$ (since $\operatorname{tr} V_t \le d\lambda_0 + t$) gives the horizon-dependent radius above. We comment on the role of this result in our analysis.

Adaptive design. The covariates $x_1, x_2, \dots$ are not fixed in advance: x_t is chosen by the algorithm as a function of the past rewards, so that $(x_s)_{s\le t}$ and $(\eta_s)_{s\le t}$ are dependent and the classical confidence sets of fixed-design least squares, of radius $\asymp \mu\sqrt{d}$ with high probability, are not available. The bound of Abbasi-Yadkori et al. (2011) handles the sequential design through a self-normalised martingale inequality (the method of mixtures): the quantity $\|\sum_{s\le t} x_s\eta_s\|_{V_t^{-1}}$ is controlled simultaneously for all t, whatever the dependence of x_s on the past, at the price of the $\log\det V_t \le d\log((1 + t/(d\lambda_0)))$ term, i.e. of the factor $\sqrt{\log(T)}$ in $\beta \asymp \sqrt{d\log(T)}$. This is the *price of adaptivity*: it cannot be removed in general, since Lattimore (2023) shows that the bound is unimprovable for adversarially adaptive designs, and the same phenomenon governs estimation and inference from adaptively collected data (Khamaru et al., 2025; Lin et al., 2023). Sharper radii are known only under additional structure on the design, such as the stability condition of Lai & Wei (1982), i.e. convergence of the normalised covariance to a deterministic limit; we do not use such refinements, and the radius β of Proposition 4.1 is the one that enters Theorems 4.2 and 4.3.

How the event is used. The event $\mathcal{E}$ is the only probabilistic input of the paper: every statement of Sections D and E is deterministic on $\mathcal{E}$, and the algorithm itself never uses β. Its role is purely analytical, to locate $\theta^\star$ within $\|\cdot\|_{V_t}$-distance β of the estimate, which the width ω_t converts into the $\|\cdot\|_*$-distance $\beta\,\omega_t^{1/2}$ (Proposition D.2); the set $\mathcal{C}_t(\alpha)$ used by α-LinUCB may be smaller or larger than $\mathcal{C}_t(\beta)$. We take β horizon-dependent for simplicity; the anytime radius β_t with $\log((\det(V_t)\,/\,\det(\lambda_0\mathbf{I})))$ in place of $d\log((1 + T/(d\lambda_0)))$ can be used instead, at the cost of carrying a nondecreasing β_t through the burn-in constants.

B FURTHER PROPERTIES OF THE ACTION SET

Consequences of Assumption (A2). Setting $\theta' = 0$ in equation 1 yields $\frac{m}{2}\|v\|_*^2 \le J_{\mathcal{X}}(v)^2 \le \frac{M}{2}\|v\|_*^2$. Consequently, for $\theta^\star \ne 0$, the scalars

$$\begin{aligned} &\ell_* := \sup_{x\in\mathcal{X}} \|x\|_*^\circ, \quad J_\star := J_{\mathcal{X}}(\theta^\star), \quad \varrho_* := J_\star/\|\theta^\star\|_* \\ &\text{satisfy} \;\; J_\star > 0, \quad \sqrt{m/2} \le \varrho_* \le \|x^\star\|_*^\circ \le \ell_* \le \sqrt{M/2}, \end{aligned} \tag{5}$$

where $x^\star := \nabla J_{\mathcal{X}}(\theta^\star)$. In particular, $\mathcal{X}$ is compact and $J_{\mathcal{X}}$ is ℓ_*-Lipschitz with respect to $\|\cdot\|_*$. These quantities enter the analysis only; the algorithm does not require them.

Unit ball and ellipsoids. For $\mathcal{X} = \mathbb{B}_2^d(1)$, one has $\ell_* = \varrho_* = 1$ and $J_\star = \|\theta^\star\|_2$. For $\mathcal{X} = \mathbf{A}^{1/2}\mathbb{B}_2^d(1)$, under the adapted norm $\|u\|_* := \|\mathbf{A}^{1/2}u\|_2$, one likewise has $\ell_* = \varrho_* = 1$.

C^2 bodies. A convex body $\mathcal{X} \subseteq \mathbb{B}_2^d(1)$ containing the origin in its interior is a C^2 *body* if its boundary is a C^2 hypersurface whose principal radii of curvature lie in $[r_-, r_+]$ for some $0 < r_- \le r_+ < \infty$ (class C_+^2 in Schneider (2014, Sect. 2.5)). Its support function is C^2 on $\mathbb{R}^d \setminus \{0\}$. For a unit vector u, the Hessian $\nabla^2 J_{\mathcal{X}}(u)$ vanishes on u (Euler's relation) and coincides on $u^\perp$ with the reverse Weingarten map of $\partial\mathcal{X}$ at the point $x = \nabla J_{\mathcal{X}}(u)$ with outer normal u; its eigenvalues are the principal radii of curvature of $\partial\mathcal{X}$ at x. Since $\nabla^2 J_{\mathcal{X}}^2 = 2J_{\mathcal{X}}\nabla^2 J_{\mathcal{X}} + 2\nabla J_{\mathcal{X}} \nabla J_{\mathcal{X}}^\top$ with both terms homogeneous of degree zero, writing $u = \theta/\|\theta\|_2$ and $v = au + w$ with $w \perp u$ gives

$$v^\top \nabla^2 J_{\mathcal{X}}^2(\theta)\, v = 2\, J_{\mathcal{X}}(u)\, w^\top \nabla^2 J_{\mathcal{X}}(u)\, w + 2\,\langle x, v\rangle^2, \qquad r_-\|w\|_2^2 \le w^\top \nabla^2 J_{\mathcal{X}}(u)\, w \le r_+\|w\|_2^2. \tag{6}$$

The first (tangential) term captures the curvature of $\partial\mathcal{X}$; the second (radial) term captures the position of the origin, since $\langle x, u\rangle = J_{\mathcal{X}}(u)$ is the distance from the origin to the supporting hyperplane with normal u. Let

$$\rho_{\min} := \min_{\|u\|_2=1} J_{\mathcal{X}}(u) \in (0, 1]$$

denote the distance from the origin to $\partial\mathcal{X}$, positive since $\mathcal{X}$ is absorbing and at most one since $\mathcal{X} \subseteq \mathbb{B}_2^d(1)$. With $\|\cdot\|_* = \|\cdot\|_2$, which is self-dual, $\ell_* \le 1$ and $\varrho_* \ge \rho_{\min}$.

Upper bound. Using $J_{\mathcal{X}}(u) \le 1$ and $\langle x, v\rangle^2 \le \|v\|_2^2$,

$$v^\top \nabla^2 J_{\mathcal{X}}^2(\theta)\, v \;\le\; 2r_+\|w\|_2^2 + 2\|v\|_2^2 \;\le\; 2(r_+ + 1)\|v\|_2^2 =: M\|v\|_2^2.$$

Lower bound. Using $J_{\mathcal{X}}(u) \ge \rho_{\min}$ and $|\langle x, v\rangle| \ge \rho_{\min}|a| - \|w\|_2$,

$$v^\top \nabla^2 J_{\mathcal{X}}^2(\theta)\, v \;\ge\; 2\rho_{\min} r_-\|w\|_2^2 + 2\big(\rho_{\min}|a| - \|w\|_2\big)_+^2.$$

If $\|w\|_2 \ge \rho_{\min}|a|/2$, then $\|v\|_2^2 \le 5\|w\|_2^2/\rho_{\min}^2$ and the first term is at least $\frac{2}{5}\rho_{\min}^3 r_-\|v\|_2^2$. Otherwise, $\|v\|_2^2 \le \frac{5}{4}a^2$ and the second term is at least $\frac{2}{5}\rho_{\min}^2\|v\|_2^2$. Hence

$$v^\top \nabla^2 J_{\mathcal{X}}^2(\theta)\, v \;\ge\; \frac{2\rho_{\min}^2}{5}\min\{1, \rho_{\min} r_-\}\|v\|_2^2 \;\ge\; \frac{\rho_{\min}^2}{9}\min\left\{1, \frac{\rho_{\min} r_-}{3}\right\}\|v\|_2^2 =: m\|v\|_2^2,$$

. The two-sided bound of Assumption (A2) follows by integrating twice along the segment from θ' to θ; the segment may pass through the origin, since $J_{\mathcal{X}}^2$ is C^1 on $\mathbb{R}^d$ (its gradient $2J_{\mathcal{X}}\nabla J_{\mathcal{X}}$ is bounded by $2\|\theta\|_2$ and extends continuously by 0 at the origin).

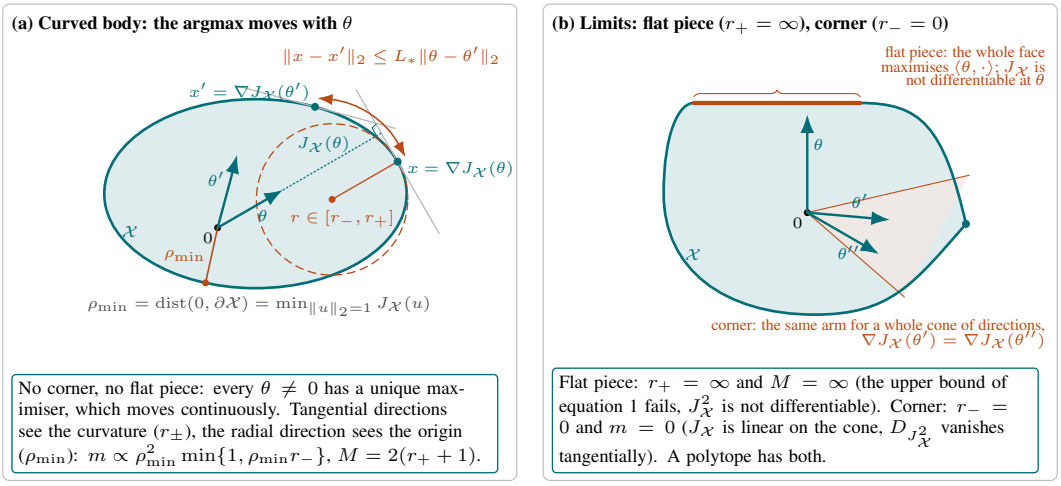


Figure 4: Assumption (A2) illustrated (parameter and action spaces superimposed). (a) A C^2 body: each direction θ selects a unique arm $x = \nabla J_{\mathcal{X}}(\theta)$, the point of $\partial\mathcal{X}$ with outer normal θ, which moves Lipschitz-continuously with θ. The radius of curvature $r \in [r_-, r_+]$ at x governs the tangential term of equation 6, and $\rho_{\min}$ its radial term. (b) The two degenerate limits: a flat face ($r_+ = \infty$, $M = \infty$) and a corner ($r_- = 0$, $m = 0$).

Role of curvature and excluded geometries. In tangential directions, the Hessian of $J_{\mathcal{X}}^2$ equals $2J_{\mathcal{X}}(u)$ times the reverse Weingarten map: flat regions (r_+ large) inflate M, and sharp regions (r_- small) deflate m. In the radial direction, it equals $2J_{\mathcal{X}}(u)^2 \geq 2\rho_{\min}^2$, which deflates m as the origin approaches the boundary, irrespective of curvature (Figure 4). The two limits are the excluded cases. A flat face is a direction θ with a whole face of maximisers: $J_{\mathcal{X}}$ is not differentiable at θ, the arm jumps across the face under arbitrarily small perturbations of θ, and $M = \infty$. A corner x_k is the common maximiser of a whole cone of directions: on this cone, $J_{\mathcal{X}}(\theta) = \langle \theta, x_k \rangle$ is linear, $\nabla^2 J_{\mathcal{X}}^2 = 2x_k x_k^\top$ has rank one, $D_{J_{\mathcal{X}}^2}(\theta, \theta') = \langle \theta - \theta', x_k \rangle^2$ vanishes for $\theta - \theta' \perp x_k$, and $m = 0$. Polytopes and finite arm sets exhibit both. On a C^2 body, by contrast, every $\theta \neq 0$ has a unique maximiser, which moves Lipschitz-continuously with θ (Proposition D.3); this allows a perturbation of the optimistic parameter to reach the arm, and is what the energy argument of the proof exploits. The constants are not sharp: for the unit ball, $r_\pm = \rho_{\min} = 1$ and equation 6 equals $2\|v\|_2^2$, so that $m = M = 2$, whereas the bounds above give $m = 2/5$ and $M = 4$.

B.1 Failure case

Proposition 6.1 (Linear regret on a polytope)**.** *In the instance above, for every $\lambda_0 > 0$, every horizon T and every radius*

$$0 < \alpha \leq \alpha_0 := \frac{(1-a^2)\sqrt{\lambda_0}}{2a}\,\varepsilon,$$

with probability at least

$$p_0 := \frac{1}{2}\Big(1 - \exp\Big(-\frac{(1-a^2)\lambda_0 \varepsilon^2}{2a^2\mu^2}\Big)\Big) > 0,$$

α-LinUCB plays $x_t \in \{e_1, -e_1\}$ at every round $t \in [T]$. On this event $\lambda_{\min}(V_t) = \lambda_0$ for all t and $\mathcal{R}_T \geq \Delta T$; in particular $\mathbb{E}[\mathcal{R}_T] \geq p_0 \Delta T$.

Proof of Proposition 6.1. Step 1: the algorithm plays vertices. For $x \in \mathcal{X}_a$, $\max_{\theta \in \mathcal{C}_t(\alpha)} \langle x, \theta \rangle = \langle x, \hat{\theta}_t \rangle + \alpha \|x\|_{V_t^{-1}} =: g_t(x)$, so $x_t \in \arg\max_{\mathcal{X}_a} g_t$. The function g_t is convex, and strictly convex along every segment not contained in a line through the origin. Its maximum over the polytope is therefore attained at vertices only: a convex function with a maximum at the midpoint of a segment is constant on that segment, and g_t is constant on no segment, so no maximiser lies in the interior of $\mathcal{X}_a$ or in the relative interior of an edge. Since the four vertices lie on the coordinate axes, $V_t = \mathrm{diag}(\lambda_0 + n_{t,1}, \lambda_0 + a^2 n_{t,2})$, where $n_{t,i}$ is the number of rounds $s < t$ with x_s on the i-th axis, $\hat{\theta}_{t,1} = (\lambda_0 + n_{t,1})^{-1} \sum_{s<t,\, x_s = \pm e_1} \langle x_s, e_1 \rangle\, y_s$, and the four indices are

$$g_t(\pm e_1) = \pm\hat{\theta}_{t,1} + \frac{\alpha}{\sqrt{\lambda_0 + n_{t,1}}}, \qquad g_t(\pm a e_2) = \pm a\,\hat{\theta}_{t,2} + \frac{a\alpha}{\sqrt{\lambda_0 + a^2 n_{t,2}}}.$$

Step 2: a sufficient condition for locking on the first axis. Let $\tilde{\eta}_s := \langle x_s, e_1 \rangle\, \eta_s$ if $x_s \in \{\pm e_1\}$ and $\tilde{\eta}_s := \eta_s$ otherwise, and $W_n := \sum_{s \leq n} \tilde{\eta}_s$. As $\langle x_s, e_1 \rangle \in \{\pm 1\}$ is $\mathcal{F}_{s-1}$-measurable and $\eta_s \sim \mathcal{N}(0, \mu^2)$ is independent of $\mathcal{F}_{s-1}$, the conditional law of $\tilde{\eta}_s$ given $\mathcal{F}_{s-1}$ is $\mathcal{N}(0, \mu^2)$: $(\tilde{\eta}_s)_s$ is i.i.d. $\mathcal{N}(0, \mu^2)$. Suppose that $x_s \in \{\pm e_1\}$ for all $s < t$ and put $n := t - 1$. Then $n_{t,1} = n$, $n_{t,2} = 0$, $\hat{\theta}_{t,2} = 0$ and, since $\langle x_s, e_1 \rangle\, y_s = \varepsilon + \tilde{\eta}_s$,

$$\hat{\theta}_{t,1} = \frac{n\varepsilon + W_n}{\lambda_0 + n},$$

so that $x_t \in \{\pm e_1\}$ as soon as

$$|\hat{\theta}_{t,1}| + \frac{\alpha}{\sqrt{\lambda_0 + n}} \;>\; \frac{a\alpha}{\sqrt{\lambda_0}}, \tag{7}$$

the left-hand side being the larger of $g_t(\pm e_1)$ and the right-hand side the common value of $g_t(\pm a e_2)$. For $n < n_0 := \lambda_0(a^{-2} - 1)$, equation 7 holds whatever the rewards, because $\alpha/\sqrt{\lambda_0 + n} >$

$a\alpha/\sqrt{\lambda_0}$: the bonus of the long vertices alone exceeds the frozen bonus of the short ones. For $n \geq N_0 := \lceil n_0 \rceil$, equation 7 holds as soon as $\hat{\theta}_{t,1} > a\alpha/\sqrt{\lambda_0}$, i.e.

$$W_n + \delta n > c_\alpha, \qquad \delta := \varepsilon - \frac{a\alpha}{\sqrt{\lambda_0}}, \qquad c_\alpha := a\alpha\sqrt{\lambda_0}.$$

Since $x_1 \in \{\pm e_1\}$ (at $t = 1$, $g_1(\pm e_1) = \alpha/\sqrt{\lambda_0} > a\alpha/\sqrt{\lambda_0} = g_1(\pm a e_2)$), an induction on t gives

$$\big\{W_n + \delta n > c_\alpha \ \text{ for all } n \geq N_0\big\} \ \subseteq \ \big\{x_t \in \{\pm e_1\} \ \text{ for all } t \geq 1\big\}.$$

Step 3: the probability of the left-hand event. For $\alpha \leq \alpha_0$ we have $a\alpha/\sqrt{\lambda_0} \leq (1-a^2)\varepsilon/2$, hence $\delta \geq \varepsilon/2$ and, using $(1-a^2)(\lambda_0 + N_0) \leq N_0$ (which is $N_0 \geq n_0$),

$$h := \delta N_0 - c_\alpha = \varepsilon N_0 - \frac{a\alpha}{\sqrt{\lambda_0}}(\lambda_0 + N_0) \ \geq \ \varepsilon N_0 - \frac{(1-a^2)\varepsilon}{2}(\lambda_0 + N_0) \ \geq \ \frac{\varepsilon N_0}{2}.$$

Now $\{W_n + \delta n > c_\alpha \ \forall n \geq N_0\} \supseteq \{W_{N_0} \geq 0\} \cap \{W_{N_0+k} - W_{N_0} + \delta k > -h \ \forall k \geq 1\}$: on the intersection, $W_{N_0} + \delta N_0 \geq c_\alpha + h > c_\alpha$ and $W_{N_0+k} + \delta(N_0 + k) > (c_\alpha + h) - h$. The first event is $\mathcal{F}_{N_0}$-measurable and has probability $1/2$, as $W_{N_0} \sim \mathcal{N}(0, \mu^2 N_0)$. For the second, let $\lambda := 2\delta/\mu^2$; conditionally on $\mathcal{F}_{N_0}$, $M_k := \exp\big(-\lambda(W_{N_0+k} - W_{N_0}) - \lambda^2\mu^2 k/2\big)$ is a nonnegative martingale with $M_0 = 1$, and $W_{N_0+k} - W_{N_0} + \delta k \leq -h$ implies $\log M_k \geq \lambda h + k(\lambda\delta - \lambda^2\mu^2/2) = \lambda h$. Ville's maximal inequality (Ville, 1939) therefore gives

$$\mathbb{P}\big(\exists k \geq 1 : W_{N_0+k} - W_{N_0} + \delta k \leq -h \,\big|\, \mathcal{F}_{N_0}\big) \ \leq \ e^{-\lambda h} = e^{-2\delta h/\mu^2} \ \leq \ e^{-\varepsilon^2 N_0/(2\mu^2)} \ \leq \ \exp\Big(-\frac{(1-a^2)\lambda_0\varepsilon^2}{2a^2\mu^2}\Big),$$

using $N_0 \geq n_0 = (1-a^2)\lambda_0/a^2$. Multiplying by $\mathbb{P}(W_{N_0} \geq 0) = 1/2$ gives the probability p_0.

Step 4: consequences. On the locking event, $x_s x_s^\top = e_1 e_1^\top$ for every s, so $V_t = \mathrm{diag}(\lambda_0 + t - 1, \lambda_0)$ and $\lambda_{\min}(V_t) = \lambda_0$; and $\langle \theta^\star, x^\star - x_t \rangle = a \mp \varepsilon \geq \Delta$ at every round, so $\mathcal{R}_T \geq \Delta T$. □

C PSEUDO-CODE OF α-LINUCB

Algorithm 1 summarises the algorithm analysed in this paper; we briefly comment on its steps, complementing the description given in Section 1.

Algorithm 1: α-LinUCB

1 **Input:** Action set $\mathcal{X}$, horizon T, regularization parameter $\lambda_0 > 0$, and exploration radius $\alpha > 0$
2 **Initialize:** $V_1 \leftarrow \lambda_0 \mathbf{I}$
3 **for** $t = 1, \ldots, T$ **do**
4 Compute the ridge estimate $\hat{\theta}_t$ ▷ Estimation
5 Select an action–parameter: $(x_t, \theta_t) \in \underset{(x,\theta) \in \mathcal{X} \times \mathcal{C}_t(\alpha)}{\arg\max} \langle x, \theta \rangle$ ▷ Decision rule
6 Play x_t and observe the reward y_t ▷ Feedback
7 Update $V_{t+1} \leftarrow V_t + x_t x_t^\top$ ▷ Design-matrix update

Estimation (line 4). The ridge estimate is $\hat{\theta}_t = V_t^{-1} \sum_{s=1}^{t-1} x_s y_s$ with design matrix $V_t = \lambda_0 \mathbf{I} + \sum_{s=1}^{t-1} x_s x_s^\top$; both admit rank-one updates, so each round costs $\mathcal{O}(d^2)$ beyond the decision rule.

Decision rule (line 5). The pair (x_t, θ_t) maximises $\langle x, \theta \rangle$ jointly over $\mathcal{X} \times \mathcal{C}_t(\alpha)$, where $\mathcal{C}_t(\alpha) = \{\theta : \|\theta - \hat{\theta}_t\|_{V_t} \leq \alpha\}$. Optimising first over θ at fixed x gives the equivalent bonus form

$$x_t \in \underset{x \in \mathcal{X}}{\arg\max} \Big\{ \langle \hat{\theta}_t, x \rangle + \alpha \|x\|_{V_t^{-1}} \Big\},$$

the familiar upper-confidence-bound rule, with $\|x\|_{V_t^{-1}}$ largest along the directions least represented in the design; conversely, once x_t is chosen, $\theta_t = \hat{\theta}_t + \alpha V_t^{-1} x_t / \|x_t\|_{V_t^{-1}}$ attains the inner maximum.

The radius α. The only departure from standard LinUCB is that the radius α is an arbitrary input, held fixed over the run, rather than the concentration width β of Proposition 4.1: for $\alpha = \beta$ the algorithm is exactly the optimistic algorithm of Abbasi-Yadkori et al. (2011), while for $\alpha < \beta$ the set $\mathcal{C}_t(\alpha)$ need not contain $\theta^\star$ at any round, and as $\alpha \to 0$ the rule approaches the greedy policy. Neither β nor any property of the noise beyond the observations themselves is used by the algorithm; they appear in the analysis only.

D SPECTRAL GROWTH

Let $\beta_+ := \alpha + \beta$, where $\alpha > 0$ is the radius of the set $\mathcal{C}_t(\alpha)$ used by α-LinUCB, no restriction on α is needed, and β is the radius of the valid confidence set $\mathcal{C}_t(\beta)$ of Proposition 4.1; the confidence event is

$$\mathcal{E} := \Big\{\theta^\star \in \bigcap_{t\in[T]} \mathcal{C}_t(\beta)\Big\} = \big\{\|\hat\theta_t - \theta^\star\|_{V_t} \le \beta \ \text{ for all } t \in [T]\big\}, \qquad \mathbb{P}(\mathcal{E}) \ge 1-\delta.$$

D.1 PRELIMINARY RESULTS

Localisation. We introduce the set $\mathcal{N}_*$ of parameters lying within a controlled $\|\cdot\|_*$-distance of $\theta^\star$; on this set, both $J_\mathcal{X}$ and $\|\cdot\|_*$ remain comparable to their values at $\theta^\star$:

$$\begin{gathered}\mathcal{N}_* := \Big\{\theta \in \mathbb{R}^d : \|\theta - \theta^\star\|_* \le \frac{J_\star}{2\ell_*}\Big\}, \\ \tfrac{1}{2}J_\star \le J_\mathcal{X}(\theta) \le \tfrac{3}{2}J_\star \quad \text{and} \quad \tfrac{1}{2}\|\theta^\star\|_* \le \|\theta\|_* \le \tfrac{3}{2}\|\theta^\star\|_* \qquad (\theta \in \mathcal{N}_*),\end{gathered} \tag{8}$$

where the two-sided bounds follow from the ℓ_*-Lipschitz property of $J_\mathcal{X}$ and from $J_\star \le \ell_*\|\theta^\star\|_*$ (Proposition D.1(ii) below); in particular $\mathcal{N}_*$ is a convex set not containing the origin.

Elementary propositions. The proofs rest on the following elementary propositions about the support function and about the width $\omega_t = \max_{u\in\mathbb{B}_2^d(1)}\|V_t^{-1/2}u\|_*^2$; none of them involves a Euclidean structure beyond the one implicit in the formula for ω_t, which Proposition D.2(i) removes.

Proposition D.1 (Consequences of the assumption)**.** *Under Assumption (A2), with $\mathcal{X}$ closed:*

(i) $J_\mathcal{X}$ is convex and positively homogeneous of degree one, $J_\mathcal{X}^2$ is convex on $\mathbb{R}^d$, $J_\mathcal{X}(\theta) > 0$ for every $\theta \neq 0$, and $J_\mathcal{X} = (J_\mathcal{X}^2)^{1/2}$ is differentiable on $\mathbb{R}^d \setminus \{0\}$.

(ii) $\frac{m}{2}\|v\|_*^2 \le J_\mathcal{X}(v)^2 \le \frac{M}{2}\|v\|_*^2$ for all $v \in \mathbb{R}^d$. Consequently $\ell_* \le \sqrt{M/2}$, $\mathcal{X}$ is compact, $|J_\mathcal{X}(\theta) - J_\mathcal{X}(\theta')| \le \ell_*\|\theta - \theta'\|_*$ for all θ, θ' (in particular $J_\mathcal{X}(\theta) \le \ell_*\|\theta\|_*$), and $J_\star \ge \sqrt{m/2}\,\|\theta^\star\|_* > 0$.

(iii) For every $\theta \neq 0$, $x \mapsto \langle\theta, x\rangle$ has a unique maximiser on $\mathcal{X}$, namely $\nabla J_\mathcal{X}(\theta)$; moreover $\|\nabla J_\mathcal{X}(\theta)\|_*^\circ \le \ell_*$ and $\nabla J_\mathcal{X}^2(\theta) = 2J_\mathcal{X}(\theta)\nabla J_\mathcal{X}(\theta)$. In particular $x^\star = \nabla J_\mathcal{X}(\theta^\star)$ is the unique optimal arm and equation 5 holds.

(iv) the gradient of $J_\mathcal{X}^2$ is M-Lipschitz from $(\mathbb{R}^d, \|\cdot\|_*)$ to $(\mathbb{R}^d, \|\cdot\|_*^\circ)$:

$$\|\nabla J_\mathcal{X}^2(\theta) - \nabla J_\mathcal{X}^2(\theta')\|_*^\circ \ \le\ M\|\theta - \theta'\|_* \qquad (\theta, \theta' \in \mathbb{R}^d). \tag{9}$$

(v) At every round t (recall $V_t \succeq \lambda_0\mathbf{I} \succ 0$), the optimistic parameter satisfies $\theta_t \neq 0$ and the played arm is $x_t = \nabla J_\mathcal{X}(\theta_t)$.

Proof of Proposition D.1. (i) $J_\mathcal{X}$ is a supremum of linear functions, hence convex and positively homogeneous, with $J_\mathcal{X}(0) = 0$. Since $\mathcal{X}$ contains a neighbourhood of the origin and all norms on $\mathbb{R}^d$ are equivalent, there is $\varepsilon > 0$ with $\{x : \|x\|_*^\circ \le \varepsilon\} \subseteq \mathcal{X}$, so $J_\mathcal{X}(\theta) \ge \sup_{\|x\|_*^\circ\le\varepsilon}\langle\theta, x\rangle = \varepsilon\|\theta\|_* > 0$ for $\theta \neq 0$. As $J_\mathcal{X} \ge 0$ and $s \mapsto s^2$ is convex and nondecreasing on $[0,\infty)$, $J_\mathcal{X}^2$ is convex; as $J_\mathcal{X}^2$ is differentiable and positive off the origin, so is $J_\mathcal{X} = (J_\mathcal{X}^2)^{1/2}$.

(ii) $J_{\mathcal{X}}^2(0) = 0$ and 0 minimises the differentiable function $J_{\mathcal{X}}^2 \geq 0$, so $\nabla J_{\mathcal{X}}^2(0) = 0$; hence $D_{J_{\mathcal{X}}^2}(v, 0) = J_{\mathcal{X}}(v)^2$, and equation 1 at $\theta' = 0$ is the sandwich. For $x \in \mathcal{X}$, $\|x\|_*^\circ = \sup_{\|v\|_* \leq 1} \langle x, v\rangle \leq \sup_{\|v\|_* \leq 1} J_{\mathcal{X}}(v) \leq \sqrt{M/2}$, so $\ell_* \leq \sqrt{M/2}$; thus $\mathcal{X}$ is bounded and, being closed, compact. If $x \in \mathcal{X}$ maximises $\langle \theta, \cdot\rangle$ (a maximiser exists by compactness), then $J_{\mathcal{X}}(\theta) - J_{\mathcal{X}}(\theta') \leq \langle \theta, x\rangle - \langle \theta', x\rangle \leq \|x\|_*^\circ \|\theta - \theta'\|_* \leq \ell_* \|\theta - \theta'\|_*$, because $J_{\mathcal{X}}(\theta') \geq \langle \theta', x\rangle$; exchange the roles of θ, θ', and take $\theta' = 0$ for the bound $J_{\mathcal{X}}(\theta) \leq \ell_* \|\theta\|_*$. Finally, the sandwich at $v = \theta^\star \neq 0$ gives $J_\star \geq \sqrt{m/2}\, \|\theta^\star\|_* > 0$.

(iii) A maximiser $x \in \mathcal{X}$ exists by compactness. For every θ', $J_{\mathcal{X}}(\theta') \geq \langle \theta', x\rangle = J_{\mathcal{X}}(\theta) + \langle \theta' - \theta, x\rangle$, i.e. x is a subgradient of the convex function $J_{\mathcal{X}}$ at θ; as $J_{\mathcal{X}}$ is differentiable at $\theta \neq 0$ by (i), its subdifferential there is $\{\nabla J_{\mathcal{X}}(\theta)\}$ (Rockafellar & Wets, 1998, Thm. 25.1), so every maximiser equals $\nabla J_{\mathcal{X}}(\theta) \in \mathcal{X}$, whence $\|\nabla J_{\mathcal{X}}(\theta)\|_*^\circ \leq \ell_*$; the chain rule gives $\nabla J_{\mathcal{X}}^2 = 2 J_{\mathcal{X}} \nabla J_{\mathcal{X}}$ off the origin. At $\theta = \theta^\star$, $J_\star = \langle \theta^\star, x^\star\rangle \leq \|x^\star\|_*^\circ \|\theta^\star\|_*$ gives $\varrho_* \leq \|x^\star\|_*^\circ \leq \ell_*$, which together with (ii) is equation 5.

(iv) For a convex function that is differentiable on $\mathbb{R}^d$, the upper bound in equation 1 is equivalent to M-Lipschitz continuity of the gradient from $(\mathbb{R}^d, \|\cdot\|_*)$ to $(\mathbb{R}^d, \|\cdot\|_*^\circ)$ (Beck, 2017, Thm. 5.8); apply this to $J_{\mathcal{X}}^2$, which is convex by (i).

(v) Since $V_t \succ 0$, the confidence set $\mathcal{C}_t(\alpha)$ is compact with nonempty interior, so it contains some $\theta \neq 0$, and $J_{\mathcal{X}}(\theta) > 0$ by (i). By definition of the optimistic pair,

$$\langle \theta_t, x_t\rangle = \max_{\theta \in \mathcal{C}_t(\alpha)} \max_{x \in \mathcal{X}} \langle \theta, x\rangle = \max_{\theta \in \mathcal{C}_t(\alpha)} J_{\mathcal{X}}(\theta) \;>\; 0 = J_{\mathcal{X}}(0),$$

so $\theta_t \neq 0$; and x_t maximises $\langle \theta_t, \cdot\rangle$ on $\mathcal{X}$, so $x_t = \nabla J_{\mathcal{X}}(\theta_t)$ by (iii). □

Proposition D.2 (The width in the $\|\cdot\|_*$ geometry)**.** *For every t, let $\omega_t = \max_{u \in \mathbb{B}_2^d(1)} \|V_t^{-1/2} u\|_*^2$ be the width. Then:*

(i) $\omega_t = \max\{\|v\|_*^2 : \|v\|_{V_t} \leq 1\} = \max_{\|g\|_*^\circ \leq 1} \|g\|_{V_t^{-1}}^2$, *all maxima being attained;*

(ii) for every $r > 0$, $\max\{\|v\|_* : \|v\|_{V_t} \leq r\} = r\, \omega_t^{1/2}$*; in particular*

$$\max_{\theta \in \mathcal{C}_t(\alpha)} \|\theta - \hat{\theta}_t\|_* = \alpha\, \omega_t^{1/2} \qquad \text{and, on } \mathcal{E}, \qquad \|\hat{\theta}_t - \theta^\star\|_* \leq \beta\, \omega_t^{1/2};$$

(iii) $\|g\|_{V_t^{-1}} \leq \|g\|_*^\circ\, \omega_t^{1/2}$ *for every* $g \in \mathbb{R}^d$*; in particular* $x^\top V_t^{-1} x \leq \ell_*^2\, \omega_t$ *for every* $x \in \mathcal{X}$*;*

(iv) if $\|\cdot\|_{\mathrm{E}}$ *is a Euclidean norm on* $\mathbb{R}^d$ *with* $\kappa' \|u\|_{\mathrm{E}} \leq \|u\|_* \leq \kappa \|u\|_{\mathrm{E}}$*, and* $\lambda_{\min}(V_t)$ *denotes the smallest eigenvalue of* V_t *in* $\|\cdot\|_{\mathrm{E}}$*-orthonormal coordinates, then* $\kappa'^2 / \lambda_{\min}(V_t) \leq \omega_t \leq \kappa^2 / \lambda_{\min}(V_t)$*;*

(v) $t \mapsto \omega_t$ *is nonincreasing.*

Proof of Proposition D.2. (i) Substituting $v = V_t^{-1/2} u$ in the definition of ω_t gives $\omega_t = \max\{\|v\|_*^2 : \|v\|_{V_t} \leq 1\}$. Since $\|v\|_* = \sup_{\|g\|_*^\circ \leq 1} \langle g, v\rangle$ (duality of norms) and $\sup_{\|v\|_{V_t} \leq 1} \langle g, v\rangle = \|g\|_{V_t^{-1}}$ (duality of the quadratic norms $\|\cdot\|_{V_t}$ and $\|\cdot\|_{V_t^{-1}}$), exchanging the two suprema gives $\max\{\|v\|_* : \|v\|_{V_t} \leq 1\} = \max_{\|g\|_*^\circ \leq 1} \|g\|_{V_t^{-1}}$; square. All are suprema of continuous functions over compact sets. (ii) $\{v : \|v\|_{V_t} \leq r\} = r\{v : \|v\|_{V_t} \leq 1\}$ and $\|\cdot\|_*$ is homogeneous, so the maximum is $r\, \omega_t^{1/2}$ by (i); apply this to $v = \theta - \hat{\theta}_t$ with $r = \alpha$, and to $v = \hat{\theta}_t - \theta^\star$ with $r = \beta$ on $\mathcal{E}$. (iii) $\|g\|_{V_t^{-1}} = \sup_{\|v\|_{V_t} \leq 1} \langle g, v\rangle \leq \sup_{\|v\|_{V_t} \leq 1} \|g\|_*^\circ \|v\|_* = \|g\|_*^\circ\, \omega_t^{1/2}$, and $\|x\|_*^\circ \leq \ell_*$ for $x \in \mathcal{X}$. (iv) In $\|\cdot\|_{\mathrm{E}}$-orthonormal coordinates, $\omega_t = \sup_{\|u\|_{\mathrm{E}} \leq 1} \|V_t^{-1/2} u\|_*^2$, where

$$\kappa'^2 \|V_t^{-1/2} u\|_{\mathrm{E}}^2 \leq \|V_t^{-1/2} u\|_*^2 \leq \kappa^2 \|V_t^{-1/2} u\|_{\mathrm{E}}^2 \qquad \text{and} \qquad \sup_{\|u\|_{\mathrm{E}} \leq 1} \|V_t^{-1/2} u\|_{\mathrm{E}}^2 = \frac{1}{\lambda_{\min}(V_t)}.$$

(v) For each fixed g, $g^\top V_t^{-1} g$ is nonincreasing in t because $V_{t+1}^{-1} \preceq V_t^{-1}$; take the maximum over $\|g\|_*^\circ \le 1$ in (i). □

Proposition D.3 (Lipschitz argmax)**.** *Under Assumption (A2), for all $\theta, \theta' \in \mathcal{N}_*$,*

$$\|\nabla J_{\mathcal{X}}(\theta) - \nabla J_{\mathcal{X}}(\theta')\|_*^\circ \;\le\; L_* \, \|\theta - \theta'\|_*, \qquad L_* = \frac{M + 2\ell_*^2}{J_\star}. \tag{10}$$

Lipschitz continuity of the argmax map is the condition under which Rusmevichientong and Tsitsiklis (2010) call an arm set strongly convex; here it is derived, in the geometry of the assumption, from the smoothness of $J_{\mathcal{X}}^2$ and the localisation $J_{\mathcal{X}} \ge J_\star/2$ on $\mathcal{N}_*$.

Proof of Proposition D.3. Since $0 \notin \mathcal{N}_*$, $\nabla J_{\mathcal{X}} = \nabla J_{\mathcal{X}}^2/(2J_{\mathcal{X}})$ on $\mathcal{N}_*$ (Proposition D.1(iii)), so

$$\begin{aligned}\|\nabla J_{\mathcal{X}}(\theta) - \nabla J_{\mathcal{X}}(\theta')\|_*^\circ &\le \frac{\|\nabla J_{\mathcal{X}}^2(\theta) - \nabla J_{\mathcal{X}}^2(\theta')\|_*^\circ}{2J_{\mathcal{X}}(\theta)} + \frac{\|\nabla J_{\mathcal{X}}^2(\theta')\|_*^\circ \, |J_{\mathcal{X}}(\theta') - J_{\mathcal{X}}(\theta)|}{2J_{\mathcal{X}}(\theta) J_{\mathcal{X}}(\theta')} \\ &\le \frac{M\|\theta - \theta'\|_*}{J_\star} + \frac{2\ell_*^2 \|\theta - \theta'\|_*}{J_\star},\end{aligned}$$

where we used equation 9, $J_{\mathcal{X}}(\theta) \ge J_\star/2$ from equation 8, $\|\nabla J_{\mathcal{X}}^2(\theta')\|_*^\circ = 2J_{\mathcal{X}}(\theta')\|\nabla J_{\mathcal{X}}(\theta')\|_*^\circ \le 2\ell_* J_{\mathcal{X}}(\theta')$ (Proposition D.1(iii)), and the ℓ_*-Lipschitz property of $J_{\mathcal{X}}$. □

D.2 Geometric distortion and adapted coordinates

Adapted Euclidean structure. The proofs perform spectral bookkeeping (eigenvalues and eigenvectors of V_t, a top direction, a band of small eigenvalues, a determinant), which requires a Euclidean structure. Since α-LinUCB is equivariant under linear reparametrisations and ω_t is intrinsic, this structure is a proof device, and it can be chosen adapted to $\|\cdot\|_*$. Call $D_* \ge 1$ an *admissible distortion* if there is $\mathbf{P} \succ 0$ with

$$\|u\|_* \;\le\; \|\mathbf{P}u\|_2 \;\le\; D_* \|u\|_* \qquad (u \in \mathbb{R}^d), \tag{11}$$

that is, the Euclidean unit ball E of $u \mapsto \|\mathbf{P}u\|_2$ satisfies $D_*^{-1} B_* \subseteq E \subseteq B_*$; the smallest admissible value is the Banach–Mazur distance of B_* to the Euclidean ball. By John's theorem $D_* = \sqrt{d}$ is admissible, and under Assumption (A2) so is $D_* = \sqrt{M/m}$ (Lemma D.4, from the Hessian of $J_{\mathcal{X}}^2$ at a point of twice differentiability); when B_* is an ellipsoid, $D_* = 1$. We fix once and for all an admissible pair $(D_*, \mathbf{P})$ with

$$D_* \;\le\; \min\big\{\sqrt{d},\; \sqrt{M/m}\big\}, \tag{12}$$

and use it only inside the proofs (Lemma D.5). The distortion D_* enters the results through the burn-in threshold, through $T_\star$, and through $\log(8\Gamma)$ (where $4\log(D_*) \le 2\log(d)$); it never multiplies the rate constant.

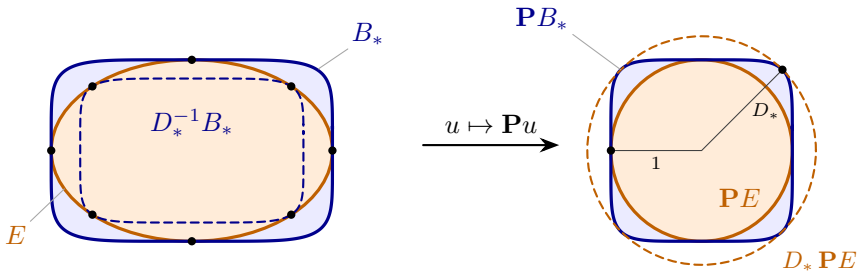
Lemma D.4 (Ellipsoidal distortion of B_* under the assumption)**.** *Under Assumption (A2), there is a symmetric matrix $\mathbf{H} \succ 0$ with*

$$m\|u\|_*^2 \;\le\; u^\top \mathbf{H} u \;\le\; M\|u\|_*^2 \qquad (u \in \mathbb{R}^d).$$

Consequently $\mathbf{P} := (\mathbf{H}/m)^{1/2}$ satisfies equation 11 with $D_* = \sqrt{M/m}$: the assumption itself forbids B_* from being far from an ellipsoid, and John's $\sqrt{d}$ can always be replaced by $\min\{\sqrt{d}, \sqrt{M/m}\}$.

Proof of Lemma D.4. $J_{\mathcal{X}}^2$ is convex and finite on $\mathbb{R}^d$ (Proposition D.1(i)). By Alexandrov's theorem (A. D. Alexandrov, 1939; see Rockafellar and Wets, *Variational Analysis*, Thm. 13.51, or Evans and Gariepy, *Measure Theory and Fine Properties of Functions*, Thm. 6.9), it is twice differentiable at almost every $\theta_0 \in \mathbb{R}^d$: there is a symmetric $\mathbf{H}$ with

$$D_{J_{\mathcal{X}}^2}(\theta_0 + v, \theta_0) = J_{\mathcal{X}}^2(\theta_0 + v) - J_{\mathcal{X}}^2(\theta_0) - \langle \nabla J_{\mathcal{X}}^2(\theta_0), v\rangle = \tfrac{1}{2} v^\top \mathbf{H} v + o(\|v\|_*^2).$$



(a) original coordinates
$D_*^{-1}B_* \subseteq E \subseteq B_*$

(b) adapted coordinates
$\mathbf{P}E \subseteq \mathbf{P}B_* \subseteq D_*\,\mathbf{P}E$

Figure 5: The adapted Euclidean structure of Section D.2. *(a)* The unit ball B_* of $\|\cdot\|_*$ (blue), the Euclidean unit ball E of $u \mapsto \|\mathbf{P}u\|_2$ (orange) inscribed in it, and the rescaled copy $D_*^{-1}B_*$ (dashed); dots mark tangencies. *(b)* The reparametrisation $u \mapsto \mathbf{P}u$ makes E round and sandwiches B_* between two Euclidean balls whose radii differ exactly by the factor D_*; this is equation 11, where all spectral bookkeeping takes place.

Fix such a θ_0. For $u \in \mathbb{R}^d$ and $\varepsilon > 0$, equation 1 with $\theta = \theta_0 + \varepsilon u$ and $\theta' = \theta_0$ gives

$$\frac{m}{2}\,\varepsilon^2\|u\|_*^2 \;\le\; \frac{\varepsilon^2}{2}\,u^\top \mathbf{H} u + o(\varepsilon^2) \;\le\; \frac{M}{2}\,\varepsilon^2\|u\|_*^2.$$

Dividing by $\varepsilon^2/2$ and letting $\varepsilon \to 0$ yields $m\|u\|_*^2 \le u^\top \mathbf{H} u \le M\|u\|_*^2$; in particular $\mathbf{H} \succ 0$. With $\mathbf{P} = (\mathbf{H}/m)^{1/2}$, $\|\mathbf{P}u\|_2^2 = u^\top \mathbf{H} u/m \in [\|u\|_*^2, (M/m)\|u\|_*^2]$, which is equation 11 with $D_* = \sqrt{M/m}$. □

Lemma D.5 (Adapted coordinates)**.** *Let $(D_*, \mathbf{P})$ be an admissible pair equation 11 and define the linear change of variables*

$$\tilde{x} := \frac{\mathbf{P}^{-1}x}{\ell_*}, \qquad \tilde{\theta} := \ell_*\,\mathbf{P}\,\theta, \qquad \textit{so that} \qquad \langle \tilde{\theta}, \tilde{x} \rangle = \langle \theta, x \rangle,$$

and let $\tilde{\mathcal{X}} := \{\tilde{x} : x \in \mathcal{X}\}$ and $\tilde{\theta}^\star := \ell_*\mathbf{P}\theta^\star$. The norms $\|\cdot\|_*$ and $\|\cdot\|_*^\circ$ are not changed: they are simply read through the change of variables, i.e. a point described by $\tilde{\theta}$ has $\|\theta\|_* = \|\mathbf{P}^{-1}\tilde{\theta}\|_*/\ell_*$ and an arm described by $\tilde{x}$ has $\|x\|_*^\circ = \|\ell_*\mathbf{P}\tilde{x}\|_*^\circ$.

(i) $\|x\|_*^\circ/(D_*\ell_*) \le \|\tilde{x}\|_2 \le \|x\|_*^\circ/\ell_*$ *for every x; in particular $\tilde{\mathcal{X}} \subseteq \mathbb{B}_2^d(1)$ and $\|\tilde{x}^\star\|_2 \ge \varrho_*/(D_*\ell_*)$.*

(ii) $\|\tilde{\theta}\|_2/(D_*\ell_*) \le \|\theta\|_* \le \|\tilde{\theta}\|_2/\ell_*$ *for every θ.*

(iii) Every intrinsic quantity is unchanged: $J_{\tilde{\mathcal{X}}}(\tilde{\theta}) = J_{\mathcal{X}}(\theta)$, the gradient $\nabla J_{\tilde{\mathcal{X}}}(\tilde{\theta}) = \mathbf{P}^{-1}\nabla J_{\mathcal{X}}(\theta)/\ell_*$ is the image of $\nabla J_{\mathcal{X}}(\theta)$ (gradients transform like arms), Bregman divergences of J^2, distances $\|\theta - \theta'\|_*$ and $\|x - x'\|_*^\circ$, the scalars ℓ_*, ϱ_*, $J_\star$ and the ball $\mathcal{N}_*$ (mapped onto $\ell_*\mathbf{P}\mathcal{N}_*$) are the same; so Assumption (A2) holds for the problem $(\tilde{\mathcal{X}}, \tilde{\theta}^\star)$ with the same m, M. Moreover α-LinUCB is equivariant: with

$$\tilde{V}_t := \ell_*^{-2}\mathbf{P}^{-1}V_t\mathbf{P}^{-1} = \tilde{\mathbf{R}} + \sum_{s<t}\tilde{x}_s\tilde{x}_s^\top, \qquad \tilde{\mathbf{R}} := \lambda_0\,\ell_*^{-2}\mathbf{P}^{-2},$$

the transported design matrix (the ridge penalty $\lambda_0\|\theta\|_2^2 = \tilde{\theta}^\top\tilde{\mathbf{R}}\,\tilde{\theta}$ is no longer isotropic), and $\hat{\theta}_t = V_t^{-1}\sum_{s<t} x_s y_s$ the ridge estimate,

$$\tilde{V}_t^{-1}\sum_{s<t}\tilde{x}_s y_s = \ell_*\mathbf{P}\,\hat{\theta}_t, \qquad \|\tilde{\theta} - \tilde{\theta}'\|_{\tilde{V}_t} = \|\theta - \theta'\|_{V_t}, \qquad \|\tilde{g}\|_{\tilde{V}_t^{-1}} = \|g\|_{V_t^{-1}},$$

so the confidence ellipsoid, the optimistic pair and the event $\mathcal{E}$ are the images of the original ones, $(\tilde{x}_t, \tilde{\theta}_t)_t$ is a trajectory of α-LinUCB on $(\tilde{\mathcal{X}}, \tilde{\theta}^\star)$ with the same rewards, and the width is unchanged: $\tilde{\omega}_t = \omega_t$.

(iv) In the new coordinates the smallest eigenvalue satisfies $\tilde{\lambda}_{t,d} \geq \omega_t^{-1}/(D_*^2\ell_*^2)$, and $x^\top V_t^{-1} x = \tilde{x}^\top \tilde{V}_t^{-1}\tilde{x} \leq \min\{1/\tilde{\lambda}_{t,d},\ \ell_*^2\,\omega_t\}$ for every $x \in \mathcal{X}$; in particular, at every round with $\omega_t^{-1} \geq 2D_*^2\ell_*^2$, one has $\tilde{\lambda}_{t,d} \geq 2$ and $x^\top V_t^{-1} x \leq 1/2$.

Proof of Lemma D.5. (i) $\|\mathbf{P}^{-1}x\|_2 = \sup_{\|v\|_2\leq 1}\langle x, \mathbf{P}^{-1}v\rangle = \sup_{\|\mathbf{P}u\|_2\leq 1}\langle x, u\rangle$. By equation 11 the constraint set contains $\{u : \|u\|_* \leq 1/D_*\}$ and is contained in $\{u : \|u\|_* \leq 1\}$, so $\|x\|_*^\circ/D_* \leq \|\mathbf{P}^{-1}x\|_2 \leq \|x\|_*^\circ$; divide by ℓ_* and use $\|x^\star\|_*^\circ \geq \varrho_*$. (ii) is equation 11 multiplied by ℓ_*, since $\|\tilde{\theta}\|_2 = \ell_*\|\mathbf{P}\theta\|_2$. (iii) The pairing is invariant, so $J_{\tilde{\mathcal{X}}}(\tilde{\theta}) = \sup_{x\in\mathcal{X}}\langle\theta, x\rangle = J_{\mathcal{X}}(\theta)$; the chain rule gives $\nabla_{\tilde{\theta}} J_{\tilde{\mathcal{X}}} = \ell_*^{-1}\mathbf{P}^{-1}\nabla_\theta J_{\mathcal{X}}$, i.e. gradients transform like arms, hence $\langle\nabla J^2_{\tilde{\mathcal{X}}}(\tilde{\theta}'), \tilde{\theta} - \tilde{\theta}'\rangle = \langle\nabla J^2_{\mathcal{X}}(\theta'), \theta - \theta'\rangle$ and Bregman divergences are unchanged. The norms $\|\cdot\|_*$, $\|\cdot\|_*^\circ$ are by definition evaluated on the original points, so distances, $\ell_* = \sup_{\mathcal{X}}\|\cdot\|_*^\circ$, $\varrho_* = J_\star/\|\theta^\star\|_*$, $J_\star$ and $\mathcal{N}_*$ are unchanged, and equation 1 and equation 9 hold verbatim for the transported problem. For the algorithm: $\tilde{V}_t = \ell_*^{-2}\mathbf{P}^{-1}(\lambda_0\mathbf{I} + \sum_{s<t} x_s x_s^\top)\mathbf{P}^{-1} = \tilde{\mathbf{R}} + \sum_{s<t}\tilde{x}_s\tilde{x}_s^\top$, so $\tilde{V}_t^{-1} = \ell_*^2\mathbf{P}V_t^{-1}\mathbf{P}$ and $\tilde{V}_t^{-1}\sum_s \tilde{x}_s y_s = \ell_*\mathbf{P}V_t^{-1}\sum_s x_s y_s = \ell_*\mathbf{P}\hat{\theta}_t$; the two norm identities follow by direct substitution. Hence $\tilde{\mathcal{C}}_t := \{\tilde{\theta} : \theta \in \mathcal{C}_t(\alpha)\}$, the maximisers of the invariant pairing over $\tilde{\mathcal{X}} \times \tilde{\mathcal{C}}_t$ are the images of those over $\mathcal{X} \times \mathcal{C}_t(\alpha)$ (for each $\theta \neq 0$ the maximising arm is unique, Proposition D.1(iii), and the analysis below never uses a particular selection of θ_t among maximising parameters), and the event $\mathcal{E}$ is the same event. Finally $\tilde{\omega}_t = \max\{\|\theta\|_*^2 : \|\tilde{\theta}\|_{\tilde{V}_t} \leq 1\} = \max\{\|\theta\|_*^2 : \|\theta\|_{V_t} \leq 1\} = \omega_t$ by Proposition D.2(i) and the second identity. (iv) By (ii) and Proposition D.2(iv) with $\kappa' = 1/(D_*\ell_*)$, $\tilde{\lambda}_{t,d} \geq \omega_t^{-1}/(D_*^2\ell_*^2)$. The energy bound is Proposition D.2(iii) together with $\|\tilde{x}\|_2 \leq 1$. □

Convention. Throughout the proofs we keep the tildes of Lemma D.5: $\tilde{x}$, $\tilde{\theta}$, $\tilde{V}_t$ are the adapted coordinates of x, θ, V_t, and every Euclidean object — the norm $\|\cdot\|_2$, eigenvalues, eigenvectors, traces, determinants — is computed on tilded quantities. Untilded symbols denote the original objects, and the norms $\|\cdot\|_*$, $\|\cdot\|_*^\circ$ are always applied to original objects; a bound such as $\|\tilde{x}\|_2 \leq \|x\|_*^\circ/\ell_*$ therefore compares the Euclidean length of the adapted coordinate vector of a point with the intrinsic dual norm of that point. In this notation, Lemma D.5(i), (ii) and (iv) read: $\tilde{\mathcal{X}} \subseteq \mathbb{B}_2^d(1)$, $\|\tilde{x}_s\|_2 \leq 1$, $\operatorname{tr}\tilde{V}_t \leq t - 1 + r_{\lambda_0}$, where

$$r_{\lambda_0} := \operatorname{tr}\tilde{\mathbf{R}} = \lambda_0\,\ell_*^{-2}\operatorname{tr}\mathbf{P}^{-2} \;\leq\; \frac{\lambda_0\, d}{\ell_*^2\,\kappa'^2}, \qquad \kappa' := \min_{\|u\|_2=1}\|u\|_*, \tag{13}$$

is the trace of the transported regulariser (by equation 11, $\lambda_{\min}(\mathbf{P}) \geq \kappa'$; $r_{\lambda_0} = \lambda_0 d$ for the unit ball and $r_{\lambda_0} = \lambda_0 \operatorname{tr}\mathbf{A}^{-1}$ for the ellipsoid $\mathbf{A}^{1/2}\mathbb{B}_2^d(1)$ with its adapted norm), and

$$\begin{gathered}\frac{\|\tilde{\theta}\|_2}{D_*\ell_*} \leq \|\theta\|_* \leq \frac{\|\tilde{\theta}\|_2}{\ell_*}, \qquad \frac{\|x\|_*^\circ}{D_*\ell_*} \leq \|\tilde{x}\|_2 \leq \frac{\|x\|_*^\circ}{\ell_*}, \qquad \|\tilde{x}^\star\|_2 \geq \frac{\varrho_*}{D_*\ell_*},\\ \tilde{\lambda}_{t,d} \geq \frac{\omega_t^{-1}}{D_*^2\ell_*^2}, \qquad \text{hence } \tilde{\lambda}_{t,d} \geq 2 \text{ at every round with } \omega_t^{-1} \geq 2D_*^2\ell_*^2,\end{gathered} \tag{14}$$

where $\tilde{\lambda}_{t,1} \geq \cdots \geq \tilde{\lambda}_{t,d} > 0$ are the eigenvalues of $\tilde{V}_t$ with orthonormal eigenvectors $\tilde{v}_{t,1}, \ldots, \tilde{v}_{t,d}$; each eigenvalue is nondecreasing in t (because $\tilde{V}_{t+1} \succeq \tilde{V}_t$), Weyl's inequality gives $\tilde{\lambda}_{t+1,i} \leq \tilde{\lambda}_{t,i}+1$, and for a vector u of the adapted coordinates we write $u_i := \langle u, \tilde{v}_{t,i}\rangle$ when the round is clear from context; in particular $\tilde{x}_{t,i} := \langle\tilde{x}_t, \tilde{v}_{t,i}\rangle$. We also write $\tilde{e}^\star := \tilde{x}^\star/\|\tilde{x}^\star\|_2$ for the unit vector of the optimal arm, $\tilde{P}^\perp := \mathbf{I} - \tilde{e}^\star(\tilde{e}^\star)^\top$ and $\tilde{\phi}_t := \angle(\tilde{v}_{t,1}, \tilde{e}^\star)$, the sign of $\tilde{v}_{t,1}$ being chosen so that $\cos\tilde{\phi}_t \geq 0$. The scalars ω_t, $\|\theta - \theta'\|_*$, $\|x - x'\|_*^\circ$, $\|\theta - \theta'\|_{V_t} = \|\tilde{\theta} - \tilde{\theta}'\|_{\tilde{V}_t}$, $x^\top V_t^{-1}x = \tilde{x}^\top\tilde{V}_t^{-1}\tilde{x}$, $J_{\mathcal{X}}$, m, M, ℓ_*, ϱ_*, $J_\star$, α, β are the same in both descriptions (Lemma D.5(iii)), so every statement about them proved below is a statement about the original problem. Proposition 4.1 is applied in the original coordinates; the event $\mathcal{E}$ is invariant under the change of variables (Lemma D.5(iii)), and only $\mathbb{P}(\mathcal{E}) \geq 1 - \delta$ is used.

D.3 Definition of the constants

Definition D.6 (Constants). *All constants of this section are fixed here once, in dependency order. Two geometric inputs come first,*

$$c_{\mathrm{eng}} := \frac{m}{6\,J_\star} \qquad \textit{and} \qquad L_* := \frac{M + 2\ell_*^2}{J_\star}, \tag{15}$$

the exploration-engine constant and the Lipschitz constant of the argmax map $\theta \mapsto \nabla J_{\mathcal{X}}(\theta)$ from $(\mathcal{N}_*, \|\cdot\|_*)$ to $(\mathcal{X}, \|\cdot\|_*^\circ)$ (Proposition D.3); then

$$\begin{aligned} c_1 &:= \frac{c_{\mathrm{eng}}\,\alpha}{2\sqrt{d}}, \qquad C_2 := \frac{3\,L_*^2\,\beta_+^2}{\ell_*^2\,c_1}, \qquad \Gamma := \frac{5\times 10^4\,d\,D_*^4\,L_*^4\,\beta_+^4\,\log(T)}{c_{\mathrm{eng}}^4\,\alpha^4}, \\ K &:= \frac{40\,d\,\log(8\Gamma)}{c_{\mathrm{eng}}^2\,\alpha^2}, \qquad c_0 := \min\Big(\frac{1}{4\sqrt{K}},\ \frac{c_1}{\sqrt{\log(9\,T_\star)}}\Big), \end{aligned} \tag{16}$$

where $T_\star$ is the $\mathrm{poly}(\log T)$ explicit burn-in threshold of Definition D.10 below.

Two facts about Γ are used repeatedly: since

$$\frac{L_*\beta_+}{c_{\mathrm{eng}}\alpha} = \frac{6(M+2\ell_*^2)\beta_+}{m\alpha} \;\geq\; 6 \tag{17}$$

(as $M \geq m$ and $\beta_+ \geq \alpha$), $d \geq 2$, $D_* \geq 1$ and $\log(T) \geq 1$,

$$\Gamma \;\geq\; \frac{80\,D_*^2\,L_*^2\,\beta_+^2}{c_{\mathrm{eng}}^2\,\alpha^2} \qquad \text{and} \qquad \Gamma \;\geq\; \frac{1280\,D_*^4\,\ell_*^4\,C_2^2\,\log(T)}{c_{\mathrm{eng}}^2\,\alpha^2} = \frac{46080\,d\,D_*^4\,L_*^4\,\beta_+^4\,\log(T)}{c_{\mathrm{eng}}^4\,\alpha^4}, \tag{18}$$

where $C_2^2 = 9L_*^4\beta_+^4/(\ell_*^4 c_1^2) = 36\,d\,L_*^4\beta_+^4/(\ell_*^4 c_{\mathrm{eng}}^2\alpha^2)$; in particular $\Gamma \geq 1$. Note that $J_\star$ cancels in Γ, that Γ depends on the geometry only through the invariant ratio equation 17 and D_*, and that by equation 12 $4\log(D_*) \leq 2\log(d)$, so $\log(8\Gamma) = O\big(\log(ed) + \log\big(e + (M+2\ell_*^2)\beta_+/(m\alpha)\big) + \log(e\log(eT))\big)$; hence

$$c_0 \;\asymp\; \frac{m\,\alpha}{J_\star\sqrt{d}\,\sqrt{\log(8\Gamma)+\log(9T_\star)}}.$$

D.4 Exploration engine, burn-in and crude bound

The first step of the analysis is a *crude* bound on the width, $\omega_t^{-1} \gtrsim \sqrt{t/\log(t)}$ (Lemma D.9), obtained from an exploration inequality forced by optimism (the *engine*, Lemma D.7) and an elliptic-potential argument. The engine holds at every round, deterministically, with a constant that involves $J_{\mathcal{X}}(\hat\theta_t)$; on $\mathcal{E}$ it forces the width to decay from the first round on, which shows that the threshold Λ_0^* of the burn-in is reached at an explicit round t_0 (Proposition D.8(i)). From t_0 on, the estimate, the optimistic parameter and the whole set $\mathcal{C}_t(\alpha)$ lie in $\mathcal{N}_*$, where the argmax map is L_*-Lipschitz (Proposition D.3), the played arms are well conditioned in the adapted coordinates (Proposition D.8(ii)–(v)), and the engine takes its final form with the constant c_{eng}.

Lemma D.7 (Exploration engine). *Recall $c_{\mathrm{eng}} = m/(6J_\star)$ and $\beta_+ = \alpha + \beta$.*

(i) At every round $t \geq 1$, deterministically, $x_t \neq 0$, the optimistic parameter has the closed form

$$\theta_t \;=\; \hat\theta_t + \alpha\,\frac{V_t^{-1}x_t}{\|x_t\|_{V_t^{-1}}}, \tag{19}$$

and

$$\|x_t\|_{V_t^{-1}} \;\geq\; \frac{m\,\alpha\,\omega_t}{4\,J_{\mathcal{X}}(\hat\theta_t) + 2\,\ell_*\,\alpha\,\omega_t^{1/2}}. \tag{20}$$

(ii) On $\mathcal{E}$, at every round $t \in [T]$,

$$x_t^\top V_t^{-1} x_t \;\geq\; \frac{m^2\alpha^2}{64}\,\min\Big\{\frac{\omega_t^2}{J_\star^2},\;\frac{\omega_t}{\ell_*^2\beta_+^2}\Big\}. \tag{21}$$

(iii) On $\mathcal{E}$, at every round $t \in [T]$ with $\ell_*\beta_+\,\omega_t^{1/2} \leq J_\star/2$ (the exploration engine*),*

$$\|x_t\|_{V_t^{-1}} \;\geq\; c_{\mathrm{eng}}\,\alpha\,\omega_t, \qquad \textit{hence} \qquad x_t^\top V_t^{-1} x_t \;\geq\; c_{\mathrm{eng}}^2\,\alpha^2\,\omega_t^2. \tag{22}$$

The statements and their proof involve neither the burn-in nor a Euclidean structure.

Proof of Lemma D.7. (i) *Reduction to the confidence set.* By definition of the optimistic pair and Proposition D.1(v),

$$\langle \theta_t, x_t\rangle = \max_{\theta\in\mathcal{C}_t(\alpha)}\max_{x\in\mathcal{X}}\langle\theta, x\rangle = \max_{\theta\in\mathcal{C}_t(\alpha)} J_{\mathcal{X}}(\theta), \qquad x_t = \nabla J_{\mathcal{X}}(\theta_t), \qquad \langle\theta_t, x_t\rangle = J_{\mathcal{X}}(\theta_t) > 0,$$

so $\theta_t \in \arg\max_{\theta\in\mathcal{C}_t(\alpha)} J_{\mathcal{X}}(\theta)$ and $x_t \neq 0$.

Stationarity identity. For $x \neq 0$ and $\theta \in \mathcal{C}_t(\alpha)$, the Cauchy–Schwarz inequality in the V_t-geometry gives $\langle \theta - \hat\theta_t, x\rangle \leq \|x\|_{V_t^{-1}}\|\theta - \hat\theta_t\|_{V_t} \leq \alpha\,\|x\|_{V_t^{-1}}$, with equality if and only if $\theta = \hat\theta_t + \alpha\, V_t^{-1}x/\|x\|_{V_t^{-1}}$. Since θ_t maximises $\langle\cdot, x_t\rangle$ over $\mathcal{C}_t(\alpha)$, it is this point:

$$\theta_t - \hat\theta_t = \frac{\alpha}{\|x_t\|_{V_t^{-1}}} V_t^{-1} x_t, \qquad \|\theta_t - \hat\theta_t\|_{V_t} = \alpha, \tag{23}$$

which is the closed form in (i). It also gives the *optimism inequality*

$$J_{\mathcal{X}}(\theta_t) = \langle\hat\theta_t, x_t\rangle + \langle\theta_t - \hat\theta_t, x_t\rangle \;\leq\; J_{\mathcal{X}}(\hat\theta_t) + \alpha\,\|x_t\|_{V_t^{-1}}. \tag{24}$$

The multiplier, via strong convexity of $J_{\mathcal{X}}^2$. Let v° attain the maximum defining ω_t in Proposition D.2(i), so $\|v^\circ\|_{V_t} \leq 1$ and $\|v^\circ\|_* = \omega_t^{1/2}$, and set $\zeta := \alpha\, v^\circ$, so that

$$\|\zeta\|_{V_t} \leq \alpha, \qquad \|\zeta\|_* = \alpha\,\omega_t^{1/2}:$$

the two points $\hat\theta_t \pm \zeta$ lie in $\mathcal{C}_t(\alpha)$, in the least-explored direction of the $\|\cdot\|_*$ geometry. Since θ_t maximises $J_{\mathcal{X}}$ over $\mathcal{C}_t(\alpha)$, equation 24 gives

$$\alpha\,\|x_t\|_{V_t^{-1}} \;\geq\; G := \max\big\{J_{\mathcal{X}}(\hat\theta_t + \zeta),\, J_{\mathcal{X}}(\hat\theta_t - \zeta)\big\} - J_{\mathcal{X}}(\hat\theta_t) \;\geq\; 0, \tag{25}$$

the last inequality by convexity ($G \geq \frac{1}{2}[J_{\mathcal{X}}(\hat\theta_t+\zeta) + J_{\mathcal{X}}(\hat\theta_t-\zeta)] - J_{\mathcal{X}}(\hat\theta_t)$). Now apply the lower bound in equation 1 at $\theta' = \hat\theta_t$ with $\theta = \hat\theta_t \pm \zeta$ and add: the linear terms $\pm\langle\zeta, \nabla J_{\mathcal{X}}^2(\hat\theta_t)\rangle$ cancel, and

$$J_{\mathcal{X}}^2(\hat\theta_t+\zeta) + J_{\mathcal{X}}^2(\hat\theta_t-\zeta) - 2J_{\mathcal{X}}^2(\hat\theta_t) = D_{J_{\mathcal{X}}^2}(\hat\theta_t+\zeta, \hat\theta_t) + D_{J_{\mathcal{X}}^2}(\hat\theta_t-\zeta, \hat\theta_t) \;\geq\; m\|\zeta\|_*^2 = m\,\alpha^2\,\omega_t. \tag{26}$$

On the other hand, for each sign,

$$J_{\mathcal{X}}^2(\hat\theta_t \pm \zeta) - J_{\mathcal{X}}^2(\hat\theta_t) = \big[J_{\mathcal{X}}(\hat\theta_t\pm\zeta) - J_{\mathcal{X}}(\hat\theta_t)\big]\big[J_{\mathcal{X}}(\hat\theta_t\pm\zeta) + J_{\mathcal{X}}(\hat\theta_t)\big] \;\leq\; G\,\big(2J_{\mathcal{X}}(\hat\theta_t) + G\big),$$

because the product is nonpositive when the first factor is, and otherwise the first factor is at most G and the second at most $2J_{\mathcal{X}}(\hat\theta_t) + G$. Summing over the two signs and comparing with equation 26, $2G\,(2J_{\mathcal{X}}(\hat\theta_t) + G) \geq m\alpha^2\,\omega_t$. Finally $G \leq \alpha\,\|x_t\|_{V_t^{-1}} \leq \ell_*\alpha\,\omega_t^{1/2}$ by equation 25 and Proposition D.2(iii), so

$$\|x_t\|_{V_t^{-1}} \;\geq\; \frac{G}{\alpha} \;\geq\; \frac{m\,\alpha\,\omega_t}{4J_{\mathcal{X}}(\hat\theta_t) + 2G} \;\geq\; \frac{m\,\alpha\,\omega_t}{4J_{\mathcal{X}}(\hat\theta_t) + 2\ell_*\alpha\,\omega_t^{1/2}}, \tag{27}$$

which is equation 20. The perturbation is chosen in the $\|\cdot\|_*$ geometry and strong convexity is used in that same geometry. (The radial direction, along which $J_{\mathcal{X}}$ is linear, needs no separate treatment:

it is controlled by the rank-one term $2\nabla J_{\mathcal{X}} \nabla J_{\mathcal{X}}^\top$ of $\nabla^2 J_{\mathcal{X}}^2$, which is what strong convexity of the *square* buys.)

(ii) On $\mathcal{E}$, $\|\hat{\theta}_t - \theta^\star\|_* \le \beta\, \omega_t^{1/2}$ (Proposition D.2(ii)), so the ℓ_*-Lipschitz property of $J_{\mathcal{X}}$ gives $J_{\mathcal{X}}(\hat{\theta}_t) \le J_\star + \ell_* \beta\, \omega_t^{1/2}$, and the denominator in equation 20 is at most

$$4J_\star + 4\ell_* \beta_+\, \omega_t^{1/2} \;\le\; 8 \max\{J_\star,\; \ell_* \beta_+\, \omega_t^{1/2}\}.$$

Hence $\|x_t\|_{V_t^{-1}} \ge \frac{m\alpha}{8} \min\{\omega_t/J_\star,\; \omega_t^{1/2}/(\ell_* \beta_+)\}$, and equation 21 follows by squaring.

(iii) If $\ell_* \beta_+ \omega_t^{1/2} \le J_\star/2$, the same denominator is at most $6J_\star$, and equation 20 gives $\|x_t\|_{V_t^{-1}} \ge m\alpha\, \omega_t/(6J_\star) = c_{\mathrm{eng}}\, \alpha\, \omega_t$. □

Proposition D.8 (Burn-in). *Recall $\beta_+ = \alpha + \beta$ and $L_* = (M + 2\ell_*^2)/J_\star$, and define the* burn-in threshold *and the* burn-in round

$$\begin{aligned} \Lambda_0^* &:= \max\left\{2D_*^2 \ell_*^2,\; \frac{16\, \ell_*^2 \beta_+^2}{J_\star^2},\; \frac{4\, D_*^2 L_*^2 \beta_+^2}{\varrho_*^2}\right\}, \\ t_0 &:= \min\left\{t \ge 1 : \frac{t}{\log(1 + t/(d\lambda_0))} > \max\left\{2d,\; \frac{128\, d\, J_\star^2\, {\Lambda_0^*}^2}{m^2 \alpha^2}\right\}\right\}. \end{aligned} \tag{28}$$

The round t_0 is finite and explicit, since $t \mapsto t/\log(1 + t/(d\lambda_0))$ increases to infinity. On $\mathcal{E}$, for every $t_0 \le t \le T$:

- *(i)* $\omega_t^{-1} \ge \Lambda_0^*$;
- *(ii)* $\|\hat{\theta}_t - \theta^\star\|_* \le \beta\, \omega_t^{1/2} \le J_\star/(4\ell_*)$; *in particular* $\hat{\theta}_t \in \mathcal{N}_*$;
- *(iii)* $\|\theta_t - \hat{\theta}_t\|_* \le \alpha\, \omega_t^{1/2} \le J_\star/(4\ell_*)$ *and* $\|\theta_t - \theta^\star\|_* \le \beta_+\, \omega_t^{1/2} \le J_\star/(4\ell_*)$; *hence* $\theta_t \in \mathcal{N}_*$, *and* $\hat{\theta}_t + \zeta \in \mathcal{N}_*$ *for every* ζ *with* $\|\zeta\|_{V_t} \le \alpha$ *(the whole confidence ellipsoid lies in* $\mathcal{N}_*$*);*
- *(iv)* $x_t = \nabla J_{\mathcal{X}}(\theta_t)$, $\|x_t - x^\star\|_*^\circ \le L_* \beta_+\, \omega_t^{1/2}$, *and* $\frac{1}{2} J_\star \le J_{\mathcal{X}}(\theta_t) \le \frac{3}{2} J_\star$;
- *(v) in the adapted coordinates,* $\|\tilde{x}_t - \tilde{x}^\star\|_2 \le L_* \beta_+\, \omega_t^{1/2}/\ell_* \le \|\tilde{x}^\star\|_2/2$; *therefore*

$$\tfrac{1}{2}\|\tilde{x}^\star\|_2 \le \|\tilde{x}_t\|_2 \le \tfrac{3}{2}\|\tilde{x}^\star\|_2 \qquad \textit{and} \qquad \langle \tilde{e}^\star, \tilde{x}_t \rangle \ge \tfrac{1}{2}\|\tilde{x}^\star\|_2. \tag{29}$$

Nothing is assumed about the trajectory: (i) is derived from the engine equation 21 and an elliptic-potential argument run from the first round, and (ii)–(v) follow from (i).

Proof of Proposition D.8. (i) Fix $T' \in [T]$. In the original coordinates $V_1 = \lambda_0 \mathbf{I}$ and $\|x_t\|_2 \le 1$ (as $\mathcal{X} \subseteq \mathbb{B}_2^d(1)$), so by the matrix determinant lemma, telescoping, and the arithmetic–geometric mean inequality ($\mathrm{tr}\, V_{T'+1} \le d\lambda_0 + T'$),

$$\sum_{t=1}^{T'} \log\bigl(1 + x_t^\top V_t^{-1} x_t\bigr) = \log\left(\frac{\det(V_{T'+1})}{\det(V_1)}\right) \;\le\; d \log\left(1 + \frac{T'}{d\lambda_0}\right).$$

Since $\min\{1, y\} \le 2\log(1 + y)$ for $y \ge 0$, the left side is at least $\frac{1}{2}\sum_{t \le T'} \min\{1, x_t^\top V_t^{-1} x_t\}$. On $\mathcal{E}$, insert equation 21: its right-hand side is $g(\omega_t)$ with $g(\omega) := \frac{m^2\alpha^2}{64} \min\{\omega^2/J_\star^2,\; \omega/(\ell_*^2 \beta_+^2)\}$ nondecreasing in ω, while $\omega_t \ge \omega_{T'}$ for $t \le T'$ (Proposition D.2(v)); hence

$$T' \min\{1, g(\omega_{T'})\} \;\le\; \sum_{t=1}^{T'} \min\{1, x_t^\top V_t^{-1} x_t\} \;\le\; 2d \log\left(1 + \frac{T'}{d\lambda_0}\right).$$

Let $T' \geq t_0$. Since $t \mapsto t/\log(1 + t/(d\lambda_0))$ is increasing, the defining inequality of t_0 holds at T': $T' > 2d\log(1 + T'/(d\lambda_0))$, so the minimum is $g(\omega_{T'}) < 1$, and

$$g(\omega_{T'}) \;\leq\; \frac{2d\,\log(1 + T'/(d\lambda_0))}{T'} \;<\; \frac{m^2\alpha^2}{64\,J_\star^2\Lambda_0^{*2}}.$$

If $\omega_{T'} > 1/\Lambda_0^*$, then by monotonicity of g and the second entry of Λ_0^* ($\ell_*^2\beta_+^2 \leq J_\star^2\Lambda_0^*$, so that at $\omega = 1/\Lambda_0^*$ the first term of the minimum is the smaller one),

$$g(\omega_{T'}) \;\geq\; g(1/\Lambda_0^*) = \frac{m^2\alpha^2}{64}\,\min\Big\{\frac{1}{J_\star^2\Lambda_0^{*2}},\;\frac{1}{\ell_*^2\beta_+^2\Lambda_0^*}\Big\} = \frac{m^2\alpha^2}{64\,J_\star^2\Lambda_0^{*2}},$$

a contradiction. Hence $\omega_{T'} \leq 1/\Lambda_0^*$.

For (ii)–(v), fix $t_0 \leq t \leq T$; throughout, $\omega_t^{-1} \geq \Lambda_0^*$ by (i), and $\alpha, \beta \leq \beta_+$. (ii) By Proposition D.2(ii), on $\mathcal{E}$, $\|\hat\theta_t - \theta^\star\|_* \leq \beta\,\omega_t^{1/2} \leq \beta_+/\sqrt{\Lambda_0^*} \leq J_\star/(4\ell_*)$ by the second entry of Λ_0^*. (iii) $\theta_t \in \mathcal{C}_t(\alpha)$ gives $\|\theta_t - \hat\theta_t\|_* \leq \alpha\,\omega_t^{1/2} \leq J_\star/(4\ell_*)$ in the same way; the triangle inequality gives $\|\theta_t - \theta^\star\|_* \leq (\alpha + \beta)\,\omega_t^{1/2} = \beta_+\,\omega_t^{1/2} \leq J_\star/(4\ell_*)$; and if $\|\zeta\|_{V_t} \leq \alpha$ then $\|\zeta\|_* \leq \alpha\,\omega_t^{1/2} \leq J_\star/(4\ell_*)$, so $\|\hat\theta_t + \zeta - \theta^\star\|_* \leq J_\star/(2\ell_*)$. (iv) $x_t = \nabla J_{\mathcal{X}}(\theta_t)$ is Proposition D.1(v). By equation 10 with $\theta_t, \theta^\star \in \mathcal{N}_*$ and by (iii), $\|x_t - x^\star\|_*^\circ \leq L_*\|\theta_t - \theta^\star\|_* \leq L_*\beta_+\,\omega_t^{1/2}$; the bounds on $J_{\mathcal{X}}(\theta_t)$ are equation 8. (v) By equation 14, $\|\tilde x_t - \tilde x^\star\|_2 \leq \|x_t - x^\star\|_*^\circ/\ell_* \leq L_*\beta_+\,\omega_t^{1/2}/\ell_* \leq L_*\beta_+/(\ell_*\sqrt{\Lambda_0^*}) \leq \varrho_*/(2D_*\ell_*) \leq \|\tilde x^\star\|_2/2$, using the third entry of Λ_0^* and $\|\tilde x^\star\|_2 \geq \varrho_*/(D_*\ell_*)$. The bounds equation 29 follow from the triangle inequality and from $\langle \tilde e^\star, \tilde x_t\rangle = \|\tilde x^\star\|_2 + \langle \tilde e^\star, \tilde x_t - \tilde x^\star\rangle \geq \|\tilde x^\star\|_2 - \|\tilde x_t - \tilde x^\star\|_2$. □

Size of the burn-in. By equation 5 and equation 12 ($\ell_*^2 \leq M/2$, $\varrho_*^{-2} \leq 2/m$, $D_*^2 \leq M/m$, $L_* \leq 2M/J_\star$),

$$\Lambda_0^* \;\leq\; \max\Big\{\frac{M^2}{m},\;\frac{8M\beta_+^2}{J_\star^2},\;\frac{32M^3\beta_+^2}{m^2J_\star^2}\Big\};$$

for the unit ball and for every ellipsoid with its adapted norm, $\Lambda_0^* = \max\{2,\,64\beta_+^2/J_\star^2\}$, whatever the anisotropy. Consequently $t_0 = \widetilde{O}\big(d + d\,J_\star^2\Lambda_0^{*2}/(m^2\alpha^2)\big)$, i.e.

$$t_0 \;=\; \widetilde{O}\Big(d + \frac{d\,M^4J_\star^2}{m^4\alpha^2} + \frac{d\,M^6\beta_+^4}{m^6J_\star^2\alpha^2}\Big),$$

with a logarithmic dependence on $1/(d\lambda_0)$.

Lemma D.9 (Crude bound)**.** *Recall $c_1 = c_{\mathrm{eng}}\alpha/(2\sqrt{d}) = m\alpha/(12J_\star\sqrt{d})$, and let $T_1 := \max(2t_0,\,3d)$. On $\mathcal{E}$, the engine equation 22 holds at every round $t_0 \leq t \leq T$, and*

$$\omega_t^{-1} \;\geq\; c_1\sqrt{\frac{t}{\log(t)}} \qquad \textit{for all } T_1 \leq t \leq T. \tag{30}$$

Proof of Lemma D.9. For $t \geq t_0$, $\ell_*\beta_+\,\omega_t^{1/2} \leq \ell_*\beta_+/\sqrt{\Lambda_0^*} \leq J_\star/4$ by Proposition D.8(i) and the second entry of Λ_0^*, so Lemma D.7(iii) applies at every $t \geq t_0$. For the crude bound, an elliptic-potential argument converts equation 22 into equation 30. We use the adapted coordinates. By the matrix determinant lemma, $\det\left(\tilde V_{t+1}\right) = \det\left(\tilde V_t\right)\left(1 + x_t^\top V_t^{-1}x_t\right)$, so telescoping over $t = t_0, \ldots, T'$ for any $t_0 \leq T' \leq T$,

$$\sum_{t=t_0}^{T'}\log\big(1 + x_t^\top V_t^{-1}x_t\big) = \log\left(\frac{\det\left(\tilde V_{T'+1}\right)}{\det\left(\tilde V_{t_0}\right)}\right) \;\leq\; d\,\log(T'),$$

since $\tilde V_{T'+1} = \tilde V_{t_0} + \sum_{t=t_0}^{T'}\tilde x_t\tilde x_t^\top \preceq \tilde V_{t_0} + (T' - t_0 + 1)\mathbf{I}$ (as $\|\tilde x_s\|_2 \leq 1$) and $\tilde\lambda_{t_0,d} \geq 2$ by equation 14 and Proposition D.8(i), so that $\det\left(\tilde V_{T'+1}\right)/\det\left(\tilde V_{t_0}\right) =$

$\det\left(\mathbf{I} + \tilde{V}_{t_0}^{-1/2}(\tilde{V}_{T'+1} - \tilde{V}_{t_0})\tilde{V}_{t_0}^{-1/2}\right) \le (1 + T'/2)^d \le T'^d$ for $T' \ge 2$. For $t \ge t_0$, $x_t^\top V_t^{-1} x_t \le 1/2$ (Lemma D.5(iv), as $\omega_t^{-1} \ge \Lambda_0^* \ge 2D_*^2\ell_*^2$), so the elementary inequality $\log(1+y) \ge y/2$ on $[0,1]$ applies; inserting the engine equation 22 (Lemma D.7(iii), whose hypothesis $\ell_*\beta_+\omega_t^{1/2} \le J_\star/4$ holds for $t \ge t_0$ by the second entry of Λ_0^*) and the monotonicity of $\omega_\cdot$,

$$(T' - t_0 + 1)\,\frac{c_{\mathrm{eng}}^2\alpha^2\,\omega_{T'}^2}{2} \;\le\; \frac{1}{2}\sum_{t=t_0}^{T'} x_t^\top V_t^{-1} x_t \;\le\; d\,\log(T').$$

For $T' \ge T_1 = \max(2t_0, 3d)$ we have $T' - t_0 + 1 \ge T'/2$, whence $\omega_{T'}^{-2} \ge c_{\mathrm{eng}}^2\alpha^2\,T'/(4d\log(T')) = c_1^2\,T'/\log(T')$; renaming T' as t gives equation 30. □

THE BURN-IN LADDER

All burn-in thresholds are collected in one explicit chain $t_0(in\ Proposition\ D.8) \to T_1 := \max(2t_0,\,3d) \to T_\star \to T_B$.

Definition D.10 (Burn-in ladder)**.** *With $T_1 = \max(2t_0, 3d)$, $c_1 = m\alpha/(12J_\star\sqrt{d})$ and $C_2 = 3L_*^2\beta_+^2/(\ell_*^2 c_1)$ as in equation 16, define*

$$T_\star := \min\Big\{\, t \ge \max(T_1,\,20)\ :\ \forall\, s \ge t,\\ 4T_1 + r_{\lambda_0} \le \frac{L_*^2\,\beta_+^2}{\ell_*^2\,c_1}\sqrt{s\log(s)}\ \text{ and }\ C_2\sqrt{s\log(s)} \le \frac{\varrho_*^2\,s}{16\,D_*^2\ell_*^2}\,\Big\},$$

$$T_B := T_\star.$$

The first inequality makes the pre-T_1 chords and the regulariser negligible in (P1) below; the second is the spectral gap of (P2). The threshold is finite and explicit: writing

$$A := \frac{256\,D_*^4\ell_*^4\,C_2^2}{\varrho_*^4}, \qquad B := \frac{(4T_1 + r_{\lambda_0})^2\,c_1^2\,\ell_*^4}{L_*^4\,\beta_+^4} = \frac{(4T_1 + r_{\lambda_0})^2\,c_{\mathrm{eng}}^2\,\alpha^2\,\ell_*^4}{4\,d\,L_*^4\,\beta_+^4},$$

the second inequality reads $s/\log(s) \ge A$ and the first $s\log(s) \ge B$; both, once true, stay true, and they hold for every $s \ge \max\{3A'\log(A'),\ B,\ 3\}$ with $A' := \max\{A, 3\}$. Hence

$$T_\star \;\le\; \max\{T_1,\ 20,\ 3A'\log(A'),\ B\} \tag{31}$$

Dependence on the horizon. *$T_\star$ depends on T only through $\beta_+ = \alpha + \beta$. The terms of equation 31 scales as β_+^4/α^2 times a polynomial in $(d, M/m, J_\star^{-1})$. Hence, up to such polynomial factors,*

$$T_B = T_\star = \mathcal{O}(\mathrm{poly}(\log T)).$$

D.5 PROOF OF THEOREM 4.3

The engine equation 22 is a lower bound on $x_t^\top V_t^{-1} x_t$, while the crude bound equation 30 is an upper bound on the sum of these quantities. To go from $\sqrt{t/\log(t)}$ to $\sqrt{t}$ we replace the elliptic potential by a logarithmically capped potential whose capacity does not grow with t, and record how much of the engine's energy it captures. The structure of the argument is a two-scale separation: a *doubling lemma* shows that ω_t^{-1} doubles within $O(\vartheta^2)$ rounds of the moment it equals ϑ, and a backward induction turns this into $\omega_t^{-1} \ge c_0\sqrt{t}$ from the initialisation window onwards.

EPOCHS, BAND, AND THE CAPPED POTENTIAL

Epochs are measured on the inverse width ω_t^{-1}, while the potential that records the exploration is a spectral function of $\tilde{V}$ in the adapted coordinates. Fix an *epoch* start $\tau \ge T_\star$ with reference scale $\vartheta := \omega_\tau^{-1}$; a round $t \ge \tau$ is *active* if $\omega_t^{-1} \le 2\vartheta$ (active rounds form an initial segment, by

monotonicity of $\omega_\cdot$, and $\omega_t^{-1} \geq \vartheta$ for all $t \geq \tau$). Set

$$\bar{\vartheta} := \frac{\vartheta}{D_*^2 \ell_*^2}, \qquad \text{so that} \qquad 0 < \bar{\vartheta} \leq \tilde{\lambda}_{t,d} \quad \text{for all } t \geq \tau, \tag{32}$$

because $\tilde{\lambda}_{t,d} \geq \tilde{\lambda}_{\tau,d} \geq \omega_\tau^{-1}/(D_*^2\ell_*^2)$ by equation 14; $\bar{\vartheta}$ is the Euclidean shadow of the epoch scale, and it will cancel from every constant. With $\Gamma \geq 1$ the band width of equation 16, define the *band*, the *mid-scale set*, and the band energy

$$\mathcal{B}_t := \{\, i \geq 2 : \tilde{\lambda}_{t,i} < \Gamma\bar{\vartheta} \,\}, \qquad \mathcal{J}_t := \{\, i \geq 2 : \tilde{\lambda}_{t,i} \geq \Gamma\bar{\vartheta} \,\}, \qquad E_t^{\mathcal{B}} := \sum_{i \in \mathcal{B}_t} \frac{\tilde{x}_{t,i}^2}{\tilde{\lambda}_{t,i}},$$

and, with $L := 8\Gamma\bar{\vartheta}$ (not to be confused with the Lipschitz constant L_*), the *logarithmically capped potential*

$$\Phi_{\bar{\vartheta}}(\mathbf{W}) := \operatorname{tr} h(\mathbf{W}) = \sum_{i=1}^{d} h(\lambda_i(\mathbf{W})), \qquad h(\lambda) := \bar{\vartheta} \log\!\left(\frac{L\,(\lambda + \bar{\vartheta})}{\bar{\vartheta}\,(\lambda + L)}\right), \qquad \mathbf{W} \succeq 0, \tag{33}$$

where $\mathbf{W}$ is any symmetric matrix of the adapted coordinates (it will be $\tilde{V}_t$). The function h increases from $h(0) = 0$ to $h(\infty) = \bar{\vartheta}\log\big(L/\bar{\vartheta}\big) = \bar{\vartheta}\log(8\Gamma)$: it counts $\bar{\vartheta}$-units of growth on the logarithmic scale, but only between $\bar{\vartheta}$ and L, so that its total capacity is bounded independently of the horizon.

Lemma D.11 (Capacity and per-round gain of the capped potential)**.** *Let $\tau \geq T_\star$, $\vartheta := \omega_\tau^{-1}$, $\bar{\vartheta} = \vartheta/(D_*^2\ell_*^2)$ and $L = 8\Gamma\bar{\vartheta}$ be as above.*

(i) For every $\mathbf{W} \succeq 0$, $0 \leq \Phi_{\bar{\vartheta}}(\mathbf{W}) \leq d\,\bar{\vartheta}\log(8\Gamma)$.

(ii) For every $t \geq \tau$, $\Phi_{\bar{\vartheta}}(\tilde{V}_{t+1}) - \Phi_{\bar{\vartheta}}(\tilde{V}_t) \geq \frac{\bar{\vartheta}}{4}\, E_t^{\mathcal{B}} \geq 0$.

Proof of Lemma D.11. Step 0: a determinant representation. Since

$$h(\lambda) = \bar{\vartheta}\big[\log\big(\lambda + \bar{\vartheta}\big) - \log(\lambda + L)\big] + \bar{\vartheta}\log\big(L/\bar{\vartheta}\big),$$

we have, for $\mathbf{W} \succeq 0$,

$$\Phi_{\bar{\vartheta}}(\mathbf{W}) = \bar{\vartheta}\Big[\log\big(\det\big(\mathbf{W} + \bar{\vartheta}\mathbf{I}\big)\big) - \log(\det\left(\mathbf{W} + L\mathbf{I}\right))\Big] + d\,\bar{\vartheta}\log\big(L/\bar{\vartheta}\big).$$

Proof of (i). h is nondecreasing on $[0,\infty)$ with $h(0) = 0$ (because $\bar{\vartheta} < L$), and $h(\lambda) \leq \bar{\vartheta}\log\big(L/\bar{\vartheta}\big) = \bar{\vartheta}\log(8\Gamma)$ for every $\lambda \geq 0$; sum over the d eigenvalues.

Proof of (ii). Fix $t \geq \tau$ and abbreviate $\tilde{V} := \tilde{V}_t$, $\tilde{x} := \tilde{x}_t$, $\tilde{\lambda}_i := \tilde{\lambda}_{t,i}$, $\tilde{x}_i := \tilde{x}_{t,i}$, so that $\tilde{V}_{t+1} = \tilde{V} + \tilde{x}\tilde{x}^\top$; recall $\bar{\vartheta} \leq \tilde{\lambda}_d$ by equation 32 and $\tilde{\lambda}_d \geq 2$ by equation 14, since $\omega_t^{-1} \geq \Lambda_0^* \geq 2D_*^2\ell_*^2$ for $t \geq t_0$ (Proposition D.8(i)). Set

$$a := \tilde{x}^\top(\tilde{V} + \bar{\vartheta}\mathbf{I})^{-1}\tilde{x} = \sum_{i=1}^{d} \frac{\tilde{x}_i^2}{\tilde{\lambda}_i + \bar{\vartheta}}, \qquad b := \tilde{x}^\top(\tilde{V} + L\mathbf{I})^{-1}\tilde{x} = \sum_{i=1}^{d} \frac{\tilde{x}_i^2}{\tilde{\lambda}_i + L}.$$

(a) Determinant lemma. By the matrix determinant lemma applied to $\tilde{V} + \bar{\vartheta}\mathbf{I}$ and to $\tilde{V} + L\mathbf{I}$, the representation of Step 0 yields $\Phi_{\bar{\vartheta}}(\tilde{V} + \tilde{x}\tilde{x}^\top) - \Phi_{\bar{\vartheta}}(\tilde{V}) = \bar{\vartheta}\log((1+a)/(1+b))$, with $0 \leq b \leq a$ (because $L > \bar{\vartheta}$) and $a \leq \tilde{x}^\top\tilde{V}^{-1}\tilde{x} \leq 1/\tilde{\lambda}_d \leq 1/2$, since $(\tilde{V} + \bar{\vartheta}\mathbf{I})^{-1} \preceq \tilde{V}^{-1}$, $\|\tilde{x}\|_2 \leq 1$ and $\tilde{\lambda}_d \geq 2$.

(b) Logarithm. For $0 \leq b \leq a \leq 1/2$, $\log((1+a)/(1+b)) = \log\Big(1 + \frac{a-b}{1+b}\Big) \geq \frac{a-b}{(1+b)(1+\frac{a-b}{1+b})} = \frac{a-b}{1+a} \geq \frac{2}{3}(a-b)$, using $\log(1+y) \geq y/(1+y)$ for $y \geq 0$.

(c) Term-by-term difference. Each summand of $a - b$ is nonnegative:

$$a - b = \sum_{i=1}^{d} \tilde{x}_i^2\Big(\frac{1}{\tilde{\lambda}_i + \bar{\vartheta}} - \frac{1}{\tilde{\lambda}_i + L}\Big) = \sum_{i=1}^{d} \frac{(L - \bar{\vartheta})\,\tilde{x}_i^2}{(\tilde{\lambda}_i + \bar{\vartheta})(\tilde{\lambda}_i + L)} \;\geq\; \sum_{i \in \mathcal{B}_t} \frac{(L - \bar{\vartheta})\,\tilde{x}_i^2}{(\tilde{\lambda}_i + \bar{\vartheta})(\tilde{\lambda}_i + L)}.$$

(d) Band coordinates. For $i \in \mathcal{B}_t$, $\bar{\vartheta} \le \tilde{\lambda}_d \le \tilde{\lambda}_i < \Gamma\bar{\vartheta}$, hence $\tilde{\lambda}_i + \bar{\vartheta} \le 2\tilde{\lambda}_i$, $\tilde{\lambda}_i + L < (\Gamma + 8\Gamma)\bar{\vartheta} = 9\Gamma\bar{\vartheta}$ and $L - \bar{\vartheta} \ge 7\Gamma\bar{\vartheta}$, so

$$\frac{(L-\bar{\vartheta})\,\tilde{x}_i^2}{(\tilde{\lambda}_i+\bar{\vartheta})(\tilde{\lambda}_i+L)} \;\ge\; \frac{7\Gamma\bar{\vartheta}\,\tilde{x}_i^2}{2\tilde{\lambda}_i\cdot 9\Gamma\bar{\vartheta}} = \frac{7}{18}\,\frac{\tilde{x}_i^2}{\tilde{\lambda}_i}.$$

(e) Conclusion. Combining (a)–(d), $\Phi_{\bar{\vartheta}}(\tilde{V}_{t+1}) - \Phi_{\bar{\vartheta}}(\tilde{V}_t) \ge \bar{\vartheta}\cdot\frac{2}{3}\cdot\frac{7}{18}\sum_{i\in\mathcal{B}_t}\tilde{x}_i^2/\tilde{\lambda}_i = \frac{7}{27}\bar{\vartheta}E_t^{\mathcal{B}} \ge \frac{\bar{\vartheta}}{4}E_t^{\mathcal{B}}$. □

LEAKAGE DECOMPOSITION

Lemma D.12 (Leakage decomposition)**.** *Let $\tau \ge T_\star$ be an epoch start with $\vartheta = \omega_\tau^{-1}$, $\bar{\vartheta} = \vartheta/(D_*^2\ell_*^2)$, band $\mathcal{B}_t$ and band energy $E_t^{\mathcal{B}}$ as above. Let $\max(\tau, T_\star) \le t \le T$ be active (i.e. $\vartheta \le \omega_t^{-1} \le 2\vartheta$). Then, on $\mathcal{E}$,*

$$x_t^\top V_t^{-1} x_t \;\le\; E_t^{\mathcal{B}} \;+\; \frac{24}{t} \;+\; \frac{c_{\mathrm{eng}}^2\,\alpha^2\,\omega_t^2}{10}. \tag{34}$$

Proof of Lemma D.12. Step 0: spectral preliminaries. We first establish, for every $T_\star \le t \le T$ on $\mathcal{E}$, the three facts

$$\text{(P1)}\;\; \tilde{Q}_t + r_{\lambda_0} \le C_2\sqrt{t\log(t)}, \qquad \text{where } \tilde{Q}_t := \sum_{s=1}^{t-1}\|\tilde{x}_s - \tilde{x}^\star\|_2^2,$$

$$\text{(P2)}\;\; \tilde{\lambda}_{t,1} \ge \frac{\|\tilde{x}^\star\|_2^2\,t}{8},\;\; \tilde{\lambda}_{t,1} - \tilde{\lambda}_{t,2} \ge \frac{\|\tilde{x}^\star\|_2^2\,t}{16}, \qquad \text{(P3)}\;\; \|\tilde{x}^\star\|_2^2\,\sin^2\tilde{\phi}_t\cdot\tilde{\lambda}_{t,d} \le 16\,C_2^2\,\log(t). \tag{35}$$

Note that the chord sum $\tilde{Q}_t$ dominates the tangential energy $\sum_{s<t}\|\tilde{P}^\perp\tilde{x}_s\|_2^2$, since $\|\tilde{P}^\perp\tilde{x}_s\|_2 = \|\tilde{P}^\perp(\tilde{x}_s - \tilde{x}^\star)\|_2 \le \|\tilde{x}_s - \tilde{x}^\star\|_2$.

(P1): a per-round chord bound, summed against the crude bound. For $t_0 \le s \le T$, on $\mathcal{E}$, Proposition D.8(v) gives $\|\tilde{x}_s - \tilde{x}^\star\|_2^2 \le L_*^2\beta_+^2\,\omega_s/\ell_*^2$. Splitting $\tilde{Q}_t$ at T_1, using the trivial bound $\|\tilde{x}_s - \tilde{x}^\star\|_2^2 \le 4$ before T_1 (arms lie in $\mathbb{B}_2^d(1)$), the crude bound equation 30 on ω_s after, and $\sum_{s=T_1}^{t-1}\sqrt{\log(s)/s} \le 2\sqrt{t\log(t)}$,

$$\tilde{Q}_t + r_{\lambda_0} \;\le\; 4T_1 + r_{\lambda_0} + \frac{L_*^2\,\beta_+^2}{\ell_*^2\,c_1}\sum_{s=T_1}^{t-1}\sqrt{\frac{\log(s)}{s}} \;\le\; 4T_1 + r_{\lambda_0} + \frac{2\,L_*^2\,\beta_+^2}{\ell_*^2\,c_1}\sqrt{t\log(t)} \;\le\; C_2\sqrt{t\log(t)},$$

the last step by the first defining inequality of $T_\star$, since $t \ge T_\star$.

(P2): Rayleigh quotient of $\tilde{e}^\star$ and the gap. Let $\tilde{B}_t := \tilde{\mathbf{U}}^\top\tilde{V}_t\tilde{\mathbf{U}}$ be the compression of $\tilde{V}_t$ to $(\tilde{e}^\star)^\perp$ ($\tilde{\mathbf{U}}$ an orthonormal basis of $(\tilde{e}^\star)^\perp$); then $\operatorname{tr}\tilde{B}_t = \operatorname{tr}(\tilde{\mathbf{U}}^\top\tilde{\mathbf{R}}\,\tilde{\mathbf{U}}) + \sum_{s<t}\|\tilde{P}^\perp\tilde{x}_s\|_2^2 \le r_{\lambda_0} + \tilde{Q}_t$, and by Cauchy interlacing $\tilde{\lambda}_{t,2} \le \lambda_{\max}(\tilde{B}_t) \le \operatorname{tr}\tilde{B}_t \le \tilde{Q}_t + r_{\lambda_0} \le C_2\sqrt{t\log(t)}$. On the other hand, by equation 29, $\langle\tilde{e}^\star, \tilde{x}_s\rangle \ge \|\tilde{x}^\star\|_2/2$ for every $s \ge t_0$, so the Rayleigh quotient of the unit vector $\tilde{e}^\star$ gives

$$\tilde{\lambda}_{t,1} \;\ge\; (\tilde{e}^\star)^\top\tilde{V}_t\tilde{e}^\star \;\ge\; \sum_{s<t}\langle\tilde{e}^\star, \tilde{x}_s\rangle^2 \;\ge\; \frac{\|\tilde{x}^\star\|_2^2(t-t_0)}{4} \;\ge\; \frac{\|\tilde{x}^\star\|_2^2\,t}{8},$$

as $\tilde{\mathbf{R}} \succeq 0$ and $t \ge T_1 \ge 2t_0$. Hence $\tilde{\lambda}_{t,1} - \tilde{\lambda}_{t,2} \ge \|\tilde{x}^\star\|_2^2 t/8 - C_2\sqrt{t\log(t)} \ge \|\tilde{x}^\star\|_2^2 t/16$, by the second defining inequality of $T_\star$ and $\|\tilde{x}^\star\|_2^2 \ge \varrho_*^2/(D_*^2\ell_*^2)$ equation 14.

(P3): misalignment of the top eigenvector. For the unit vector $\tilde{e}^\star$, $(\tilde{e}^\star)^\top\tilde{V}_t\tilde{e}^\star = \operatorname{tr}\tilde{V}_t - \operatorname{tr}\tilde{B}_t$ (both sides equal $(\tilde{e}^\star)^\top\tilde{\mathbf{R}}\,\tilde{e}^\star + \sum_{s<t}\langle\tilde{e}^\star, \tilde{x}_s\rangle^2$, as $\|\tilde{x}_s\|_2^2 = \langle\tilde{e}^\star, \tilde{x}_s\rangle^2 + \|\tilde{P}^\perp\tilde{x}_s\|_2^2$), so that $(\tilde{e}^\star)^\top\tilde{V}_t\tilde{e}^\star \ge \tilde{\lambda}_{t,1} - \tilde{Q}_t - r_{\lambda_0}$, while expanding $\tilde{e}^\star$ in the eigenbasis, $(\tilde{e}^\star)^\top\tilde{V}_t\tilde{e}^\star \le \cos^2\tilde{\phi}_t\,\tilde{\lambda}_{t,1} + \sin^2\tilde{\phi}_t\,\tilde{\lambda}_{t,2}$. Comparing,

$$\sin^2\tilde{\phi}_t\big(\tilde{\lambda}_{t,1} - \tilde{\lambda}_{t,2}\big) \;\le\; \tilde{Q}_t + r_{\lambda_0}, \qquad \text{so} \qquad \sin^2\tilde{\phi}_t \;\le\; \frac{16\,(\tilde{Q}_t + r_{\lambda_0})}{\|\tilde{x}^\star\|_2^2\,t} \;\le\; \frac{16\,C_2}{\|\tilde{x}^\star\|_2^2}\sqrt{\frac{\log(t)}{t}}$$

by (P1)–(P2). Multiplying by $\tilde{\lambda}_{t,d} \le \tilde{\lambda}_{t,2} \le C_2\sqrt{t\log(t)}$ (here $d \ge 2$ is used) proves (P3).

Step 1: the decomposition. Split the energy according to $\{1\} \cup \mathcal{J}_t \cup \mathcal{B}_t$ (the scalar $x_t^\top V_t^{-1} x_t$ is invariant, and we expand it in the adapted eigenbasis):

$$x_t^\top V_t^{-1} x_t = \tilde{x}_t^\top \tilde{V}_t^{-1} \tilde{x}_t = \frac{\tilde{x}_{t,1}^2}{\tilde{\lambda}_{t,1}} + \sum_{i\in\mathcal{J}_t} \frac{\tilde{x}_{t,i}^2}{\tilde{\lambda}_{t,i}} + E_t^{\mathcal{B}}.$$

(a) Top direction. By equation 29 and (P2),

$$\frac{\tilde{x}_{t,1}^2}{\tilde{\lambda}_{t,1}} \;\le\; \frac{\|\tilde{x}_t\|_2^2}{\tilde{\lambda}_{t,1}} \;\le\; \frac{9\|\tilde{x}^\star\|_2^2/4}{\|\tilde{x}^\star\|_2^2\, t/8} = \frac{18}{t} \;\le\; \frac{24}{t}.$$

(b) Mid-scale coordinates. Write $\tilde{x}_{t,i} = \langle \tilde{x}_t - \tilde{x}^\star, \tilde{v}_{t,i}\rangle + \langle \tilde{x}^\star, \tilde{v}_{t,i}\rangle$ for $i \in \mathcal{J}_t$ and use $(u+v)^2 \le 2u^2 + 2v^2$:

$$q_t^{\mathcal{J}} := \sum_{i\in\mathcal{J}_t} \tilde{x}_{t,i}^2 \;\le\; 2\|\tilde{x}_t - \tilde{x}^\star\|_2^2 + 2\sum_{i\ge 2} \langle \tilde{x}^\star, \tilde{v}_{t,i}\rangle^2 = 2\|\tilde{x}_t - \tilde{x}^\star\|_2^2 + 2\|\tilde{x}^\star\|_2^2 \sin^2\tilde{\phi}_t,$$

since $\sum_{i\ge 2} \langle \tilde{x}^\star, \tilde{v}_{t,i}\rangle^2 = \|\tilde{x}^\star\|_2^2 - \langle \tilde{x}^\star, \tilde{v}_{t,1}\rangle^2 = \|\tilde{x}^\star\|_2^2 \sin^2\tilde{\phi}_t$. Now $\min_{i\in\mathcal{J}_t} \tilde{\lambda}_{t,i} \ge \Gamma\bar{\vartheta} = \Gamma\vartheta/(D_*^2\ell_*^2) \ge \Gamma/(2D_*^2\ell_*^2\,\omega_t)$ by activity ($\vartheta \ge \omega_t^{-1}/2$); dividing,

$$\sum_{i\in\mathcal{J}_t} \frac{\tilde{x}_{t,i}^2}{\tilde{\lambda}_{t,i}} \;\le\; \frac{2D_*^2\ell_*^2\,\omega_t\, q_t^{\mathcal{J}}}{\Gamma} \;\le\; \frac{4D_*^2\ell_*^2\,\omega_t\|\tilde{x}_t - \tilde{x}^\star\|_2^2}{\Gamma} + \frac{4D_*^2\ell_*^2\,\omega_t\|\tilde{x}^\star\|_2^2 \sin^2\tilde{\phi}_t}{\Gamma}.$$

Estimation part. By Proposition D.8(v) and the first fact in equation 18,

$$\frac{4D_*^2\ell_*^2\,\omega_t\|\tilde{x}_t - \tilde{x}^\star\|_2^2}{\Gamma} \;\le\; \frac{4D_*^2\,L_*^2\,\beta_+^2\,\omega_t^2}{\Gamma} \;\le\; \frac{c_{\mathrm{eng}}^2\,\alpha^2\,\omega_t^2}{20}.$$

Misalignment part. By (P3), $\log(t) \le \log(T)$, $\tilde{\lambda}_{t,d} \ge \omega_t^{-1}/(D_*^2\ell_*^2)$ and the second fact in equation 18,

$$\frac{4D_*^2\ell_*^2\,\omega_t\|\tilde{x}^\star\|_2^2\sin^2\tilde{\phi}_t}{\Gamma} \;\le\; \frac{64\,D_*^2\ell_*^2\,C_2^2\,\log(T)\,\omega_t}{\Gamma\,\tilde{\lambda}_{t,d}} \;\le\; \frac{64\,D_*^4\ell_*^4\,C_2^2\,\log(T)\,\omega_t^2}{\Gamma} \;\le\; \frac{c_{\mathrm{eng}}^2\,\alpha^2\,\omega_t^2}{20}.$$

Conclusion. Adding (a) and the two parts of (b) to $E_t^{\mathcal{B}}$ gives equation 34, with $\frac{1}{20} + \frac{1}{20} = \frac{1}{10}$. □

Corollary D.13 (Band energy capture)**.** *Under the hypotheses of Lemma D.12, combining equation 34 with the engine equation 22,*

$$E_t^{\mathcal{B}} \;\ge\; \Big(1 - \frac{1}{10}\Big)\, c_{\mathrm{eng}}^2\,\alpha^2\,\omega_t^2 - \frac{24}{t} \;\ge\; \frac{9\,c_{\mathrm{eng}}^2\,\alpha^2}{40\,\vartheta^2} - \frac{24}{t},$$

the last step since $\omega_t^{-1} \le 2\vartheta$ on active rounds.

DOUBLING

Lemma D.14 (Doubling)**.** *Let $\tau \ge T_\star$ with $\vartheta := \omega_\tau^{-1}$, let $K = 40\,d\log(8\Gamma)/(c_{\mathrm{eng}}^2\alpha^2)$ as in equation 16, and suppose $\tau + \lceil K\vartheta^2\rceil \le T$. Then, on $\mathcal{E}$,*

$$\omega_{\tau+\lceil K\vartheta^2\rceil}^{-1} \;\ge\; \min\Big(2\vartheta,\; \frac{c_{\mathrm{eng}}\,\alpha\sqrt{\tau}}{32}\Big).$$

Proof of Lemma D.14. Set $N := \lceil K\vartheta^2\rceil$ and suppose, for contradiction, that $\omega_{\tau+N}^{-1} < \min(2\vartheta, c_{\mathrm{eng}}\alpha\sqrt{\tau}/32)$. Then all rounds $t \in [\tau, \tau+N)$ are active (by monotonicity, $\omega_t^{-1} \le \omega_{\tau+N}^{-1} <$

2ϑ), and $\tau > 1024\vartheta^2/(c_{\mathrm{eng}}^2\alpha^2)$, so that $24/t \le 24/\tau < 24\,c_{\mathrm{eng}}^2\alpha^2/(1024\vartheta^2) \le c_{\mathrm{eng}}^2\alpha^2/(40\vartheta^2)$ for every $t \ge \tau$. By Corollary D.13, for every $t \in [\tau, \tau+N)$,

$$E_t^{\mathcal{B}} \;\ge\; \frac{9\,c_{\mathrm{eng}}^2\alpha^2}{40\vartheta^2} - \frac{c_{\mathrm{eng}}^2\alpha^2}{40\vartheta^2} = \frac{c_{\mathrm{eng}}^2\,\alpha^2}{5\,\vartheta^2}.$$

By Lemma D.11(ii), each such round deposits at least $\frac{\bar\vartheta}{4}\cdot\frac{c_{\mathrm{eng}}^2\alpha^2}{5\vartheta^2} = \frac{\bar\vartheta\,c_{\mathrm{eng}}^2\alpha^2}{20\vartheta^2}$ into the potential, and the potential never decreases; so

$$\Phi_{\bar\vartheta}(\tilde V_{\tau+N}) - \Phi_{\bar\vartheta}(\tilde V_\tau) \;\ge\; N\,\frac{\bar\vartheta\,c_{\mathrm{eng}}^2\alpha^2}{20\vartheta^2} \;\ge\; K\vartheta^2\,\frac{\bar\vartheta\,c_{\mathrm{eng}}^2\alpha^2}{20\vartheta^2} = 2\,d\,\bar\vartheta\log(8\Gamma).$$

But by Lemma D.11(i) the left side is at most $d\,\bar\vartheta\log(8\Gamma)$: contradiction. The Euclidean scale $\bar\vartheta$ appears on both sides of the budget count and cancels, which is why K depends neither on D_* nor on ℓ_*; its factor d is the capacity $d\,\bar\vartheta\log(8\Gamma)$ of the potential, the engine contributing α^2 only. □

PROOF OF THE MAIN THEOREM

Lemma D.15 (Initialisation window). *On $\mathcal{E}$, $\omega_t^{-1} \ge c_0\sqrt{t}$ for all $t \in [T_B, \min(9T_B, T)]$.*

Proof of Lemma D.15. For $t \ge T_B = T_\star \ge T_1$, the crude bound gives $\omega_t^{-1} \ge c_1\sqrt{t/\log(t)}$; and for $t \le 9T_\star$, $\sqrt{t/\log(t)} \ge \sqrt{t}/\sqrt{\log(9T_\star)}$, so $\omega_t^{-1} \ge c_1\sqrt{t}/\sqrt{\log(9T_\star)} \ge c_0\sqrt{t}$ by the definition of c_0. □

Proposition D.16 (Induction). *On $\mathcal{E}$, $\omega_t^{-1} \ge c_0\sqrt{t}$ for all $T_B \le t \le T$.*

Proof of Proposition D.16. Suppose not, and let $t^\dagger := \min\{t \in [T_B, T] : \omega_t^{-1} < c_0\sqrt{t}\}$ be the first failure. By Lemma D.15, $t^\dagger > 9T_B$. Put $\vartheta := \omega_{t^\dagger}^{-1} < c_0\sqrt{t^\dagger}$ and $N := \lceil K\vartheta^2\rceil$; then $K\vartheta^2 < Kc_0^2t^\dagger \le t^\dagger/16$ since $c_0 \le 1/(4\sqrt{K})$, so $N \le t^\dagger/16 + 1 \le t^\dagger/8$ (as $t^\dagger \ge 20 \ge 16$). Let $\tau := t^\dagger - N \ge \frac{7}{8}t^\dagger > 7T_B \ge T_\star$; by minimality of $t^\dagger$, $\omega_\tau^{-1} \ge c_0\sqrt{\tau}$, while by monotonicity $\omega_\tau^{-1} \le \omega_{t^\dagger}^{-1} = \vartheta$. Apply Lemma D.14 at τ with reference scale $\vartheta' := \omega_\tau^{-1} \le \vartheta$: since $\lceil K\vartheta'^2\rceil \le N$ and $\tau + \lceil K\vartheta'^2\rceil \le t^\dagger \le T$, monotonicity gives

$$\vartheta = \omega_{t^\dagger}^{-1} \;\ge\; \omega_{\tau+\lceil K\vartheta'^2\rceil}^{-1} \;\ge\; \min\Big(2\vartheta',\; \frac{c_{\mathrm{eng}}\,\alpha\sqrt{\tau}}{32}\Big).$$

Two cases. If the minimum is $2\vartheta'$, then $\vartheta \ge 2\vartheta' \ge 2c_0\sqrt{\tau} \ge 2c_0\sqrt{7t^\dagger/8} > c_0\sqrt{t^\dagger}$, contradicting $\vartheta < c_0\sqrt{t^\dagger}$. If the minimum is $c_{\mathrm{eng}}\alpha\sqrt{\tau}/32$, then $\vartheta \ge \frac{c_{\mathrm{eng}}\alpha}{32}\sqrt{7t^\dagger/8} \ge c_0\sqrt{t^\dagger}$, because $c_0 \le \frac{1}{4\sqrt{K}} = \frac{c_{\mathrm{eng}}\alpha}{4\sqrt{40\,d\log(8\Gamma)}} \le \frac{c_{\mathrm{eng}}\alpha}{4\sqrt{160}} < \frac{c_{\mathrm{eng}}\alpha}{32}\sqrt{\frac{7}{8}}$ (using $\log(8\Gamma) \ge \log(8) \ge 2$, $d \ge 2$ and $4\sqrt{160} > 32\sqrt{8/7}$), again contradicting the definition of ϑ. □

With this in hand, the proof of Theorem 4.3 is straightforward.

Theorem 4.3 (Spectral Growth). *Let Assumptions (A1) and (A2) hold, let $\mathcal{E}$ and β be as in Proposition 4.1, and let $T_B = \mathcal{O}(\mathrm{poly}(\log T))$ be the burn-in threshold. Then, on the event $\mathcal{E}$ (hence with probability at least $1-\delta$), for all $T_B \le t \le T$,*

$$\omega_t = \max_{u\in\mathbb{B}_2^d(1)} \|V_t^{-1/2}u\|_*^2 \;\le\; \frac{1}{c_0\sqrt{t}}, \qquad \textit{where} \qquad c_0 = \Omega\Big(\frac{m\alpha}{J_\star\sqrt{d}}\Big).$$

Proof of Theorem 4.3. By Proposition D.16, $\omega_t^{-1} \geq c_0\sqrt{t}$ on $\mathcal{E}$ for $T_B \leq t \leq T$; since ω_t is invariant under the change of coordinates of Lemma D.5, this is a statement about the original problem, and it is the first display by Proposition D.2(ii). The three consequences are Proposition D.8(ii)–(iv) with $\omega_t^{-1} \geq c_0\sqrt{t}$. The event $\mathcal{E}$ has probability at least $1-\delta$ by Proposition 4.1. For the constant, $\frac{1}{4\sqrt{K}} = \frac{c_{\mathrm{eng}}\alpha}{\sqrt{640\,d\log(8\Gamma)}}$ and $\frac{c_1}{\sqrt{\log(9T_\star)}} = \frac{c_{\mathrm{eng}}\alpha}{\sqrt{4\,d\log(9T_\star)}}$, so

$$c_0 \;\geq\; \frac{c_{\mathrm{eng}}\,\alpha}{\sqrt{640\,d\,(\log(8\Gamma)+\log(9T_\star))}} \;\geq\; \frac{m\,\alpha}{152\,J_\star\sqrt{d}\,\sqrt{\log(8\Gamma)+\log(9T_\star)}};$$

where, by Definitions D.6 and D.10, the factor $\log(8\Gamma)+\log(9T_\star) = \mathcal{O}\big(\log\log T\big)$ up to terms logarithmic in the problem parameters, i.e. it contributes only an iterated-logarithm factor in T. □

E Regret of α-LinUCB

E.1 Regret decomposition

Lemma E.1 (Per-round regret)**.** *On the event $\mathcal{E}$, for every $T_B \leq t \leq T$, the instantaneous regret $r_t := J_\star - \langle\theta^\star, x_t\rangle$ satisfies*

$$r_t \;\leq\; \frac{D_{J_{\mathcal{X}}^2}(\theta^\star,\theta_t)}{J_\star}.$$

Proof of Lemma E.1. Fix $T_B \leq t \leq T$. By Proposition D.1(iii), $\nabla J_{\mathcal{X}}^2(\theta_t) = 2J_{\mathcal{X}}(\theta_t)\,x_t$ with $\langle\theta_t, x_t\rangle = J_{\mathcal{X}}(\theta_t)$; by Proposition D.8(iv), $J_{\mathcal{X}}(\theta_t) \geq \frac{1}{2}J_\star$, as $t \geq T_B \geq t_0$; and $\langle\theta^\star, x_t\rangle \leq J_\star$ since $x_t \in \mathcal{X}$. Hence

$$\begin{aligned}
\frac{D_{J_{\mathcal{X}}^2}(\theta^\star,\theta_t)}{J_\star} - r_t &= J_\star - \frac{J_{\mathcal{X}}^2(\theta_t)}{J_\star} - \frac{2J_{\mathcal{X}}(\theta_t)}{J_\star}\langle x_t, \theta^\star-\theta_t\rangle - J_\star + \langle\theta^\star, x_t\rangle \\
&= \frac{J_{\mathcal{X}}^2(\theta_t)}{J_\star} - \Big(\frac{2J_{\mathcal{X}}(\theta_t)}{J_\star} - 1\Big)\langle\theta^\star, x_t\rangle && (\langle\theta_t, x_t\rangle = J_{\mathcal{X}}(\theta_t)) \\
&\geq \frac{J_{\mathcal{X}}^2(\theta_t)}{J_\star} - \Big(\frac{2J_{\mathcal{X}}(\theta_t)}{J_\star} - 1\Big)J_\star && (J_{\mathcal{X}}(\theta_t) \geq \tfrac{1}{2}J_\star,\ \langle\theta^\star, x_t\rangle \leq J_\star) \\
&= \frac{\big(J_{\mathcal{X}}(\theta_t) - J_\star\big)^2}{J_\star} \;\geq\; 0.
\end{aligned}$$

□

E.2 Proof of Theorem 4.2

Theorem 4.2 (α-LinUCB regret)**.** *Let Assumptions (A1) and (A2) hold. On the event $\mathcal{E}$, hence with probability at least $1-\delta$,*

$$\mathcal{R}_T \;=\; \mathcal{O}\Big(\frac{M}{m}\,\frac{\alpha^2+\beta^2}{\alpha}\,\sqrt{dT}\Big).$$

Proof of Theorem 4.2. We have $\mathcal{R}_T = \sum_{t=1}^T r_t$ with $r_t = J_\star - \langle\theta^\star, x_t\rangle \geq 0$. For the rounds $t < T_B$ we use the crude bound

$$r_t \;=\; \langle\theta^\star, x^\star - x_t\rangle \;\leq\; \|\theta^\star\|_*\|x^\star - x_t\|_*^\circ \;\leq\; 2\ell_*\|\theta^\star\|_* \;\leq\; \frac{2\ell_* J_\star}{\varrho_*},$$

since $\|x\|_*^\circ \le \ell_*$ on $\mathcal{X}$ and $\varrho_*\|\theta^\star\|_* \le J_{\mathcal{X}}(\theta^\star) = J_\star$ by equation 5. On $\mathcal{E}$, for every $T_B \le t \le T$,

$$\begin{aligned} r_t &\le \frac{D_{J_{\mathcal{X}}^2}(\theta^\star, \theta_t)}{J_\star} \;\le\; \frac{M}{2J_\star}\|\theta^\star - \theta_t\|_*^2 && \text{(Lemma E.1 and equation 1)} \\ &\le \frac{M}{J_\star}\Big(\|\theta^\star - \hat\theta_t\|_*^2 + \|\hat\theta_t - \theta_t\|_*^2\Big) && ((a+b)^2 \le 2a^2 + 2b^2) \\ &\le \frac{M}{J_\star}\big(\alpha^2 + \beta^2\big)\,\omega_t && \text{(Proposition D.2(ii))}, \end{aligned}$$

where the last step uses $\theta^\star \in \mathcal{C}_t(\beta)$ on $\mathcal{E}$ and $\theta_t \in \mathcal{C}_t(\alpha)$ (see Proposition D.2(ii)). Summing,

$$\begin{aligned} \mathcal{R}_T &\le \frac{2\ell_* J_\star}{\varrho_*}\,T_B + \frac{M(\alpha^2+\beta^2)}{J_\star}\sum_{t=T_B}^{T}\omega_t \\ &\le \frac{2\ell_* J_\star}{\varrho_*}\,T_B + \frac{M(\alpha^2+\beta^2)}{J_\star\, c_0}\sum_{t=T_B}^{T}\frac{1}{\sqrt{t}} && \text{(Theorem 4.3, } t \ge T_B) \\ &\le \frac{2\ell_* J_\star}{\varrho_*}\,T_B + \frac{2M(\alpha^2+\beta^2)}{J_\star\, c_0}\sqrt{T} && \Big(\textstyle\sum_{t\le T} t^{-1/2} \le 2\sqrt{T}\Big), \\ &= \mathcal{O}\Big(\frac{M}{m}\,\sqrt{d}\,\frac{\alpha^2+\beta^2}{\alpha}\,\sqrt{T}\Big). \end{aligned}$$

The last step uses $c_0 \gtrsim m\alpha/(J_\star\sqrt{d}\,)$ (Theorem 4.3), so that the second term is at most $\mathcal{O}(\,\frac{M}{m}\,\sqrt{d}\,\frac{\alpha^2+\beta^2}{\alpha}\,\sqrt{T})$, and $T_B = \mathcal{O}(\mathrm{poly}(\log T))$ (Definition D.10). □