A Novel General Framework for Sharp Lower Bounds in Succinct Stochastic Bandits

Guo Zeng[1] Jean Honorio[1,2]

[1]School of Computing and Information Systems, The University of Melbourne

[2]ARC Training Centre in Optimisation Technologies, Integrated Methodologies, and Applications

Nov 2025

Problem Setting

At each round t

The agent executes an action A_t from a fixed action set $\mathcal{A}$. Then, the bandit machine generates a reward

$$y_t = \langle A_t, \theta \rangle + \eta_t \quad \text{where } \eta_t \sim \mathcal{N}(0,1)$$

- The parameter θ is unknown to the agent and has a succinct structure,
  - e.g. entry-sparse vector, group-sparse matrix, or low-rank matrix
- The agent follows a policy π to select actions based on past interactions

The **regret** of a policy π on a bandit instance $(\mathcal{A}, \theta)$ up to n-rounds, relative to an omniscient policy that always select the best actions:

$$R_n^\pi(\mathcal{A}, \theta) := \mathbb{E}_{\pi,\theta}\left[\sum_{t=1}^{n} \max_{X\in\mathcal{A}} \langle X, \theta \rangle - \sum_{t=1}^{n} y_t\right] \quad \text{where } \max_{X\in\mathcal{A}} |\langle X, \theta \rangle| \le 1$$

Single-round regret is **bounded**. Regret grows at best $O(1)$, at worst $O(n)$

Regret Upper and Lower Bounds

- Upper bound limits the regret growth of **a specific policy** uniformly across a class of bandit instances, called **the assumption class**
- Lower bound upholds the best achievable regret rate of **any policy** over the assumption class

Succinct structure	Prior upper bounds	Prior lower bounds
Entry-sparse vector $\theta \in \mathbb{R}^d$ At most s non-zero entries	$O(\sqrt{sdn})$ in [1] $O(s\sqrt{n}\log(dn))$ in [5] $O(\sqrt{sn\log(dn)})$ in [8] $O(C_{\min}^{-2/3}s^{2/3}n^{2/3})$ in [2] $O(s^{1/3}n^{2/3}\sqrt{\log(dn)})$ in [7]	$\Omega(\min(C_{\min}^{-1/3}s^{1/3}n^{2/3}, \sqrt{dn}))$ in [2] $\Omega(\sqrt{sdn})$ in [6]
Group-sparse matrix $\theta \in \mathbb{R}^{d_1 \times d_2}$ At most s non-zero rows	$O(\sqrt{sd_2d_1n})$ in [3] $O(s^{1/3}n^{2/3}(\sqrt{d_2} + \sqrt{\log d_1}))$ in [7]	—
Low-rank matrix $\theta \in \mathbb{R}^{d_1 \times d_2}$ Rank limited to s	$O((d_1 + d_2)^{3/2}\sqrt{sn})$ in [4] $O((d_1 + d_2)^{3/2}\sqrt{sn})$ in [3] $O(s^{1/3}n^{2/3}\log(d_1 + d_2))$ in [7]	—

Contribution

- **A succinctness model** in general vector space, along with lemmas that may be of independent interest.
- **A general framework** for deriving minimax lower bounds of succinct linear bandits in both data-rich ($n >> d$) and data-poor ($n << d$) regimes
  - Revolving around: *information-regret trade-off* & *succinctness support*.
- **Improved and novel lower bounds** from applying this framework to three stochastic linear bandit problems that exhibit *succinct structure* and permit *well-conditioned exploration*.

Succinctness Model

When a general vector $X \in \mathbb{V}$ can be called "succinct" and how succinct it is?

The succinct unit set $\mathcal{U} \subseteq \mathbb{V}$ exists, containing "succinct units"

- pre-defined "1-succinct" vectors of unit length, i.e. $\|X\| = \langle X, X \rangle = 1$

A set of d succinct units $\{E_i\}_{i=1}^d \subseteq \mathcal{U}$ forms **a succinct support** if and only if

$$\sup_{E \in \mathcal{U}} \sum_{i=1}^{d} |\langle E, E_i \rangle| = 1$$

$X \in \mathbb{V}$ is s-succinct if and only if

- $X = \sum_{i=1}^s a_i E_i$ for some support $\{E_i\}_{i=1}^s$ and scalar coefficients $\{a_i\}_{i=1}^s$

Two semi-norms $Q(\cdot)$ and $R(\cdot)$

$$\forall X \in \mathbb{V}: \quad Q(X) := \sup_{E \in \mathcal{U}} \langle X, E \rangle, \quad R(X) := \sup_{Q(Y) \leq 1} \langle X, Y \rangle$$

- If X is s-succinct, $|\langle X, Y \rangle| \leq \min\left(Q(X)R(Y), R(X)Q(Y)\right)$ for any $Y \in \mathbb{V}$

General Lower Bound

Consider all actions $X \in \mathbb{V}$ and parameters $\theta \in \mathbb{V}$ satisfy $Q(X) \leq 1$ and $R(\theta) \leq 1$

Assume the existence of

- a parameter $\theta_0 \in \mathbb{V}$ decomposable to some support $\{E_1, E_2, ..., E_k\}$
- an action set $\mathcal{H} \subseteq \mathbb{V}$ satisfying $\max_{X \in \mathcal{H}} \langle X, \theta_0 \rangle \leq -C_0$ for some constant $C_0 > 0$
- $s - k$ groups of succinct units, $\mathcal{G}_1, ..., \mathcal{G}_{s-k}$, where any $(E'_1, E'_2, ..., E'_{s-k}) \in \prod_{i=1}^{s-k} \mathcal{G}_i$ can form a support of cardinality $s \geq 6k$ together with $\{E_1, E_2, ..., E_k\}$

Then we can construct an action set $\mathcal{A}$ such that, given any policy π, there exists an s-succinct parameter θ to incur regret

$$R_n^{\pi}(\mathcal{A}, \theta) \geq \frac{\min(C_0, e^{-4}/8)}{3} \cdot \min(s^{\frac{2}{3}} n^{\frac{2}{3}} q^{\frac{1}{3}}, s\sqrt{pn})$$

- $q \geq 1$ captures the geometric interplay between $\mathcal{H}$ and each group $\mathcal{G}_i$
- $p \geq 1$ is related to the geometry within each group $\mathcal{G}_i$ itself

Application

For the entry-sparse setting, the assumption class: $\theta \in \mathbb{R}^d$ has at most s non-zero entries, $\mathcal{A} \subseteq \mathbb{R}^d$ satisfies $C_{\min}(\mathcal{A}) > 0$ where

$$C_{\min}(\mathcal{A}) := \max_{\mu \in \Pr(\mathcal{A})} \min_{\|\beta\|=1} \mathbb{E}_{X\sim\mu}[\langle X, \beta\rangle^2]$$

- the largest possible minimum eigenvalue of the population covariance matrix
- imposed to bypass the linear regret in data-poor regime
- readily extendable to the group-sparse & low-rank settings where $\mathcal{A} \subseteq \mathbb{R}^{d_1 \times d_2}$

For each setting, we define $\mathcal{U}$ and identify a valid combination of θ_0, $\mathcal{H}$ and $\{\mathcal{G}_i\}$

Setting	q	p	**Lower bound**	
entry-sparse	$C_{\min}^{-1}$	d/s	$\Omega\big(\min\big(C_{\min}^{-\frac{1}{3}} s^{\frac{2}{3}} n^{\frac{2}{3}}, \sqrt{dsn}\big)\big)$	improved in data-poor
group-sparse	$C_{\min}^{-1}$	$d_1 d_2/s$	$\Omega\big(\min\big(C_{\min}^{-\frac{1}{3}} s^{\frac{2}{3}} n^{\frac{2}{3}}, \sqrt{d_1 d_2 sn}\big)\big)$	novel
low-rank	$C_{\min}^{-1}/s$	$d_1 d_2/s^2$	$\Omega\big(\min\big(C_{\min}^{-\frac{1}{3}} s^{\frac{1}{3}} n^{\frac{2}{3}}, \sqrt{d_1 d_2 n}\big)\big)$	novel

References

[1] Yasin Abbasi-Yadkori, David Pal, and Csaba Szepesvari. Online-to-confidence-set conversions and application to sparse stochastic bandits. In *Artificial Intelligence and Statistics*, pages 1–9. PMLR, 2012.

[2] Botao Hao, Tor Lattimore, and Mengdi Wang. High-dimensional sparse linear bandits. *Advances in Neural Information Processing Systems*, 33:10753–10763, 2020.

[3] Nicholas Johnson, Vidyashankar Sivakumar, and Arindam Banerjee. Structured stochastic linear bandits. *arXiv preprint arXiv:1606.05693*, 2016.

[4] Kwang-Sung Jun, Rebecca Willett, Stephen Wright, and Robert Nowak. Bilinear bandits with low-rank structure. In *International Conference on Machine Learning*, pages 3163–3172. PMLR, 2019.

[5] Gi-Soo Kim and Myunghee Cho Paik. Doubly-robust lasso bandit. *Advances in Neural Information Processing Systems*, 32, 2019.

[6] Tor Lattimore and Csaba Szepesvári. *Bandit algorithms*. Cambridge University Press, 2020.

[7] Wenjie Li, Adarsh Barik, and Jean Honorio. A simple unified framework for high dimensional bandit problems. In *International Conference on Machine Learning*, pages 12619–12655. PMLR, 2022.

[8] Min-hwan Oh, Garud Iyengar, and Assaf Zeevi. Sparsity-agnostic lasso bandit. In *International Conference on Machine Learning*, pages 8271–8280. PMLR, 2021.

The End