Proceedings of Machine Learning Research vol 195:1–2, 2023 36th Annual Conference on Learning Theory

Information-Directed Selection for Top-Two Algorithms

Wei You[*] WEIYOU@UST.HK
The Hong Kong University of Science and Technology
Chao Qin[*] CQ2199@COLUMBIA.EDU
Columbia University
Zihao Wang ZWANGDO@CONNECT.UST.HK
Shuoguang Yang YANGSG@UST.HK
The Hong Kong University of Science and Technology

Editors: Gergely Neu and Lorenzo Rosasco

Abstract

This paper studies the best-k-arm identification problem in a stochastic multi-arm bandit with $K \geq k$ arms in the fixed-confidence setting. The goal is to confidently identify the (sub)set of k arms with the largest mean rewards by sequentially allocating measurement effort.

We focus on the popular top-two algorithm design principle (Russo, 2020), originally mainly proposed for best-arm identification (BAI). It first identifies a pair of top-two candidates and then randomizes to select one from this pair. Various top-two algorithms (Russo, 2020; Qin et al., 2017; Shang et al., 2020) have been proposed for BAI, while the literature predominately focuses on the first step of determining the top-two candidates. The second step of selecting among the top-two candidates is usually simplified by setting a tuning parameter β, serving as the probability of selecting the top candidate. Despite enjoying great empirical performance, top-two algorithms' theoretical analyses are usually tailored to a weaker notion named β-optimality. The literature also proposes adaptive procedures of tuning β, but they require solving the instance complexity optimization problem with plug-in mean estimators, which could be computationally inefficient.

Moreover, although we can extend the top-two algorithms (designed for BAI) to tackle best-k-arm identification, we provide an example showing that for $k > 1$, top-two algorithms can fail to achieve the optimality even if the value of β is set to the *unknown* optimal value, let alone with the potential generalization of the aforementioned adaptive β-tuning procedures for $k > 1$. Indeed, we present a structural analysis of best-k-arm identification and show that $k > 1$ is surprisingly much more complicated than $k = 1$. Consequently, the optimality conditions widely used for designing BAI algorithms, e.g., those in Garivier and Kaufmann (2016); Chen and Ryzhov (2023), are no longer sufficient for $k > 1$, so how to optimally select among the top-two candidates remains open.

Our contributions. We reformulate KKT conditions of the instance complexity optimization problem to characterize asymptotic optimality. The crucial feature of our approach is the inclusion of *dual variables* in the optimality conditions, which allows us to overcome the challenges due to the much more complex optimality conditions for $k > 1$. Based on the *complementary slackness conditions*, we provide a novel interpretation of the top-two design principle, which leads us to extending the existing top-two algorithms and designing new ones. We propose an adaptive selection rule dubbed *information-directed selection* (IDS) that wisely selects among the top-two candidates based on the *stationarity conditions*. We show that integrated with IDS, *top-two Thompson sampling* is optimal for Gaussian BAI, which solves a glaring open problem in Russo (2020). A key feature of IDS is its adaptivity to the proposed top-two candidates. This differs from adaptive β-tuning procedures, which uses the same value of β regardless of the proposed candidates, even though it may be updated over time. As a by-product, we show that top-two algorithms with adaptive β-tuning cannot achieve the notion of β-optimality for $k > 1$. Finally, we demonstrate the superior performance of top-two algorithms with IDS in extensive numerical experiments.[1]

[*] Equal contributions
1. Extended abstract. Full version appears as [https://arxiv.org/abs/2205.12086, v3]

Acknowledgments

Wei You received support from the Hong Kong Research Grants Council [ECS Grant 26212320 and GRF Grant 16212823]. Shuoguang Yang was partially supported by the Early Career Scheme from the Research Grants Council of the Hong Kong SAR, China (Project No. 26209422).

References

Ye Chen and Ilya O Ryzhov. Balancing optimal large deviations in sequential selection. *Management Science*, 69(6):3457–3473, 2023.

Aurélien Garivier and Emilie Kaufmann. Optimal best arm identification with fixed confidence. In *Conference on Learning Theory*, pages 998–1027. PMLR, 2016.

Chao Qin, Diego Klabjan, and Daniel Russo. Improving the expected improvement algorithm. In *Advances in Neural Information Processing Systems*, volume 30, 2017.

Daniel Russo. Simple Bayesian algorithms for best-arm identification. *Operations Research*, 68(6): 1625–1647, 2020.

Xuedong Shang, Rianne de Heide, Pierre Menard, Emilie Kaufmann, and Michal Valko. Fixed-confidence guarantees for Bayesian best-arm identification. In *International Conference on Artificial Intelligence and Statistics*, pages 1823–1832. PMLR, 2020.