Decision Making in Changing Environments: Robustness, Query-Based Learning, and Differential Privacy

Fan Chen FANCHEN@MIT.EDU
Alexander Rakhlin RAKHLIN@MIT.EDU
Massachusetts Institute of Technology

Editors: Nika Haghtalab and Ankur Moitra

The Decision-Estimation Coefficient (DEC) (Foster et al., 2021, 2023) has been recently shown to capture the difficulty of exploration in a wide range of problems in which a learning agent interacts with an unknown environment by making decisions and observing outcomes.[1] Such problems include structured bandits, contextual bandits, and reinforcement learning, among others. The interaction protocol, termed *Decision Making with Structured Observations* (DMSO) in (Foster et al., 2021), assumes that the unknown model is fixed over the length of the interaction. In this paper, we study a setting that interpolates between the stochastic and adversarial DMSO (Foster et al., 2022). This interpolation is achieved by placing constraints on the way the model may change over time. Within the constraint set, the model is allowed to change arbitrarily, and we refer to the setting as that of *constrained adversaries*, or *hybrid DMSO*. In parallel with such constraints on the adversary, we additionally study constraints placed on the information received by the decision-maker, for instance due to privacy requirements or a specific oracle model of computation. The specification of constraints allows us to study—under the same umbrella—decision making with Statistical Queries (SQ) (Kearns, 1998), local differential privacy (LDP) (Kasiviswanathan et al., 2011; Duchi et al., 2013), robustness with respect to model corruption (Huber, 1965; Huber and Ronchetti, 2011), and smooth decision making (Rakhlin et al., 2011).

Our approach begins with the *hybrid DEC* formulation that yields both lower and upper bounds for PAC learning and no-regret learning under hybrid DMSO. Then, by investigating the specific information structures, we derive the corresponding DECs and the statistical guarantees for the aforementioned (and seemingly disparate) settings. As such, the unified viewpoint leads to a systematic "recipe" for analyzing new problems under the hybrid DMSO setting; this is illustrated on numerous examples throughout the paper. What is perhaps even more surprising, all the upper bounds are achieved by only two algorithmic approaches: a generalization of the Exploration-by-Optimization Algorithm (Lattimore and Szepesvári, 2020; Lattimore and Gyorgy, 2021; Foster et al., 2022) and a variant of the Estimation-to-Decision Algorithm (Foster et al., 2021, 2023).

In addition, we show that our framework significantly generalizes other previously studied notions of complexity measures, including SQ dimension (Feldman, 2017) that characterizes the optimal SQ-query complexity of *distribution search problems*, the local-minimax complexity of LDP learning (Duchi and Ruan, 2024), and the representation dimension (Beimel et al., 2013). Further, as a concrete application, our framework provides a near-optimal $\sqrt{T}$-regret for linear contextual bandits with local privacy (without well-conditioned assumptions), settling the open problem of the optimal regret in this setting (Zheng et al., 2020; Han et al., 2021; Li et al., 2024).

[1] Extended abstract. Full version appears as [arXiv 2501.14928, v1].

Acknowledgments

We acknowledge support from ARO through award W911NF-21-1-0328, as well as Simons Foundation and the NSF through awards DMS-2031883 and PHY-2019786.

References

Amos Beimel, Kobbi Nissim, and Uri Stemmer. Characterizing the sample complexity of private learners. In *Proceedings of the 4th conference on Innovations in Theoretical Computer Science*, pages 97–110, 2013.

John C Duchi and Feng Ruan. The right complexity measure in locally private estimation: It is not the fisher information. *The Annals of Statistics*, 52(1):1–51, 2024.

John C Duchi, Michael I Jordan, and Martin J Wainwright. Local privacy and statistical minimax rates. In *2013 IEEE 54th annual symposium on foundations of computer science*, pages 429–438. IEEE, 2013.

Vitaly Feldman. A general characterization of the statistical query complexity. In *Conference on learning theory*, pages 785–830. PMLR, 2017.

Dylan J Foster, Sham M Kakade, Jian Qian, and Alexander Rakhlin. The statistical complexity of interactive decision making. *arXiv preprint arXiv:2112.13487*, 2021.

Dylan J Foster, Alexander Rakhlin, Ayush Sekhari, and Karthik Sridharan. On the complexity of adversarial decision making. *Advances in Neural Information Processing Systems*, 35:35404–35417, 2022.

Dylan J Foster, Noah Golowich, and Yanjun Han. Tight guarantees for interactive decision making with the decision-estimation coefficient. In *The Thirty Sixth Annual Conference on Learning Theory*, pages 3969–4043. PMLR, 2023.

Yuxuan Han, Zhipeng Liang, Yang Wang, and Jiheng Zhang. Generalized linear bandits with local differential privacy. *Advances in Neural Information Processing Systems*, 34:26511–26522, 2021.

Peter J Huber. A robust version of the probability ratio test. *The Annals of Mathematical Statistics*, pages 1753–1758, 1965.

Peter J Huber and Elvezio M Ronchetti. *Robust statistics*. John Wiley & Sons, 2011.

Shiva Prasad Kasiviswanathan, Homin K Lee, Kobbi Nissim, Sofya Raskhodnikova, and Adam Smith. What can we learn privately? *SIAM Journal on Computing*, 40(3):793–826, 2011.

Michael Kearns. Efficient noise-tolerant learning from statistical queries. *Journal of the ACM (JACM)*, 45(6):983–1006, 1998.

Tor Lattimore and Andras Gyorgy. Mirror descent and the information ratio. In *Conference on Learning Theory*, pages 2965–2992. PMLR, 2021.

Tor Lattimore and Csaba Szepesvári. Exploration by optimisation in partial monitoring. In *Conference on Learning Theory*, pages 2488–2515. PMLR, 2020.

Jiachun Li, David Simchi-Levi, and Yining Wang. On the optimal regret of locally private linear contextual bandit. *arXiv preprint arXiv:2404.09413*, 2024.

Alexander Rakhlin, Karthik Sridharan, and Ambuj Tewari. Online learning: Stochastic, constrained, and smoothed adversaries. *Advances in neural information processing systems*, 24, 2011.

Kai Zheng, Tianle Cai, Weiran Huang, Zhenguo Li, and Liwei Wang. Locally differentially private (contextual) bandits learning. *Advances in Neural Information Processing Systems*, 33:12300–12310, 2020.