# ONLINE MODEL SELECTION

KASSRAIE, PARNIAN

We consider the problem of sequential inference and optimization, when the target function can be queried iteratively, but drawing samples is costly. This setting formalizes applications such as molecular design, personalized mHealth, scheduled clinical trials, and environmental monitoring, to name a few. Sequential decision-making and Bandits address such problems through algorithms that iteratively interact with the environment by drawing samples that are expected to be informative, or yield a high target value. To this end, such algorithms maintain an adaptive estimate of the target function, and use it for choosing the next sample. The statistical modeling of the target function plays a crucial role here; it is not known a priori which model is going to yield the most sample efficient algorithm, and we can only select the right model as we gather empirical evidence. This leads us to ask, can we perform online model selection, while simultaneously optimizing for the target function?

We detail the problem of online model selection and its challenges, e.g., handling non-i.i.d. and non-diverse data. We recover a scenario under which simultaneous model selection and optimzation is possible, and propose an exponential weighting algorithm for probabilistic model aggregation. The algorithm can be stopped at any time with valid regret guarantees, and its regret has an exponentially improved dependence ($\log M$) on the number of models $M$. Our approach utilizes a novel time-uniform analysis of the Lasso and establishes a new connection between online learning and high-dimensional statistics. This result is presented in [1].

**Open Direction (1).** We tackle the problem of online model selection over classes of linear functions, however it remains open for general non-parametric model classes. Iterating back to the open problem of [2], we ask, on which classes can we perform online model selection with a regret of rate $\log M$?

**Open Direction (2).** Model selection seems to inherently rely on diversity of data. This requires us to mix our sampling method with uniform draws, using a mixing ratio that vanishes as more samples are acquired. However, we conjecture that such pure exploration is not required, and there is *just enough diversity* in the data that is collected for online inference as [3] might suggest.

## References

[1] P. Kassraie, N. Emmenegger, A. Krause, and A. Pacchiano (2023). *Anytime Model Selection in Linear Bandits.* Proceedings of Advances in Neural Information Processing Systems.

[2] A. Agarwal, H. Luo, B. Neyshabur, and R. Schapire (2017). *Corralling a band of bandit algorithms.* In Conference on Learning Theory.

[3] D. Banerjee, A. Ghosh, S. Chowdhury, A. Gopalan (2023). *Exploration in Linear Bandits with Rich Action Sets and its Implications for Inference.*