## Asymptotically Efficient Distributed Experimentation

## ILAN LOBEL ANKUR MANI, JOSH REED

## **ACM Reference Format:**

Sequential decision making by a large set of myopic agents has gained significant attention over the past decade. In such settings, even a little amount of experimentation from a few agents would benefit all others but obtaining such experimentation could be challenging for a central planner. The academic literature has focused on mechanisms for promoting experimentation through monetary incentives and persuasion through careful information disclosure. In this paper, we study a simple control that the central planner can use to coordinate experimentation. We consider a set of myopic agents that observe their own histories but not the histories of other agents. In a continuous-time stochastic multi-armed bandit model, the agents pick arms myopically and receive instantaneous rewards. Meanwhile, the central planner can observe the history of all agents. We consider a class of policies where the central planner is allowed to irrevocably remove arms. We show that an appropriately chosen policy within this class can generate the needed experimentation and match the regret bounds for a centralized problem thus mitigating the cost of decentralization. We also quantify the minimum number of agents that are needed for such a policy to be asymptotically optimal and the impact of the number of agents on the speed of learning.