On the Sample Complexity of Stabilizing LTI Systems on a Single Trajectory
(Extended Abstract)

Yang Hu
SEAS
Cambridge, MA, USA
yanghu@g.harvard.edu

Adam Wierman
CMS
California Institute of Technology
Pasadena, CA, USA
adamw@caltech.edu

Guannan Qu
ECE
Carnegie Mellon University
Pittsburgh, PA, USA
gqu@andrew.cmu.edu

I. EXTENDED ABSTRACT

This is an extended abstract. The full version of the paper is in NeurIPS 2022 and can be found in [1].

Linear Time-Invariant (LTI) systems, namely $x_{t+1} = Ax_t + Bu_t$, where $x_t \in \mathbb{R}^n$ is the state and $u_t \in \mathbb{R}^m$ is the control input, are one of the most fundamental dynamical systems in control theory, and have wide applications across engineering, economics and societal domains. For systems with known dynamical matrices $(A, B)$, there is a well-developed theory for designing feedback controllers with guaranteed stability, robustness, and performance [2], [3]. However, these tools cannot be directly applied when $(A, B)$ is unknown.

Driven by the success of machine learning [4], [5], there has been significant interest in learning-based (adaptive) control, where the learner does not know the underlying system dynamics and learns to control the system in an online manner, usually with the goal of achieving low regret [6]–[14].

Despite the progress, an important limitation in this line of work is a common assumption that the learner has a priori access to a known stabilizing controller. This assumption simplifies the learning task, since it ensures a bounded state trajectory in the learning stage, and thus enables the learner to learn with low regret. However, assuming a known stabilizing controller is not practical, as stabilization itself is nontrivial and considered equally important as any other performance guarantee.

To overcome this limitation, in this paper we consider the learn-to-stabilize problem, i.e., learning to stabilize an unknown dynamical system without prior knowledge of any stabilizing controller. Understanding the learn-to-stabilize problem is of great importance to the learning-based control literature, as it serves as a precursor to any learning-based control algorithms that assume knowledge of a stabilizing controller.

The learn-to-stabilize problem has attracted extensive attention recently. For example, [15] and [16] adopt a model-based approach that first excites the open-loop system to learn dynamical matrices $(A, B)$, and then designs a stabilizing controller, with a sample complexity scaling linearly in $n$, the state dimension. However, a linearly-scaling sample complexity could be unsatisfactory for some specific instances, since the state trajectory still blows up exponentially when the open-loop system is unstable, incurring a $2^\Theta(n)$ state norm, and hence a $2^\Theta(n)$ regret (in LQR settings, for example). Another recent work [17] proposes a policy-gradient-based discount annealing method that solves a series of discounted LQR problems with increasing discount factors, and shows that the control policy converges to a near-optimal policy. However, this model-free approach only guarantees a $(n)$ sample complexity. In fact, to the best of our knowledge, state-of-the-art learn-to-stabilize algorithms with theoretical guarantees always incur state norms exponential in $n$.

It has been shown in [16] that all general-purpose control algorithms are doomed to suffer a worst-case regret of $2^\Omega(n)$. This result is intuitive, since from an information-theoretic perspective, a complete recovery of $A$ should take $\Theta(n)$ samples since $A$ itself involves $n^2$ parameters. However, this does not rule out the possibility that we can achieve better regret in specific systems. Our work is motivated by the observation that it is not always necessary to learn the whole matrix $A$ to stabilize an LTI system. For example, if the system is open-loop stable, we do not need to learn anything to stabilize it. For general LTI systems, it is still intuitive that open-loop stable “modes” exist and need not be learned for the learn-to-stabilize problem. So, we focus on learning a controller that stabilizes only the unstable “modes”, making it possible to learn a stabilizing controller without exponentially exploding state norms. The central question of this paper is: Can we exploit instance-specific properties of an LTI system to learn to stabilize it on a single trajectory, without incurring a state norm exponentially large in $n$?

**Contribution.** In this paper, we answer the above question by designing an algorithm that stabilizes an LTI system with only $O(k)$ state samples along a single trajectory, where $k$ is the instability index of the open-loop system and is defined as the number of unstable “modes” (i.e., eigenvalues with moduli larger than $1$) of matrix $A$. Our result is significant in the sense
that $k$ can be considerably smaller than $n$ for practical systems and, in such cases, our algorithm stabilizes the system using asymptotically fewer samples than prior work; specifically, it only incurs a state norm (and regret) in the order of $2^{\tilde{O}(k)}$, which is much smaller than $2^{O(n)}$ of prior state of the art when $k \ll n$.

REFERENCES