Online Adversarial Stabilization of Unknown Linear Time-Varying Systems

Jing Yu, Varun Gupta, and Adam Wierman

Abstract— This paper studies the problem of online stabilization of an unknown discrete-time linear time-varying (LTV) system under bounded non-stochastic (potentially adversarial) disturbances. We propose a novel algorithm based on convex body chasing (CBC). Under the assumption of infrequently changing or slowly drifting dynamics, the algorithm guarantees bounded-input-bounded-output stability in the closed loop. Our approach avoids system identification and applies, with minimal disturbance assumptions, to a variety of LTV systems of practical importance. We demonstrate the algorithm numerically on examples of LTV systems including Markov linear jump systems with finitely many jumps.

I. INTRODUCTION

Learning-based control of linear-time invariant (LTI) systems in the context of linear quadratic regulators (LQR) has seen considerable progress. However, many real-world systems are time-varying in nature. For example, the grid topology in power systems can change over time due to manual operations or unpredictable line failures [1]. Therefore, there is increasing recent interest in extending learning-based control of LTI systems to the linear time-varying (LTV) setting [2], [3], [4], [5], [6].

LTV systems are widely used to approximate and model real-world dynamical systems such as robotics [7] and autonomous vehicles [8]. In this paper, we consider LTV systems with dynamics of the following form:

$$x_{t+1} = A_t x_t + B_t u_t + w_t, \tag{1}$$

where $x_t \in \mathbb{R}^n$, $u_t \in \mathbb{R}^m$ and w_t denotes the state, the control input, and the bounded and potentially adversarial disturbance, respectively. We use $\theta_t = [A_t \ B_t]$ to succinctly denote the system matrices at time step t.

On the one hand, **offline control** design for LTV systems is well-established in the setting where the underlying LTV model is *known* [9], [10], [11], [12], [13]. Additionally, recent work has started focusing on regret analysis and non-stochastic disturbances for known LTV systems [2], [14].

On the other hand, **online control** design for LTV systems where the model is *unknown* is more challenging. Historically, there is a rich body of work on adaptive control design for LTV systems [15], [16], [17]. Also related is the system identification literature for LTV systems [18], [19], [20],

which estimates the (generally assumed to be stable) system to allow the application of the offline techniques.

In recent years, the potential to leverage modern data-driven techniques for controller design of unknown linear systems has led to a resurgence of work in both the LTI and LTV settings. There is a growing literature on "learning to control" unknown LTI systems under stochastic or no noise [21], [22], [23]. Learning under bounded and potentially adversarial noises poses additional challenges, but online stabilization [24] and regret [25] results have been obtained.

In comparison, there is much less work on learning-based control design for unknown LTV systems. One typical approach, exemplified by [3], [26], [27], derives stabilizing controllers under the assumption that *offline* data representing the input-output behavior of (1) is available and therefore an *offline* stabilizing controller can be pre-computed. Similar *finite-horizon* settings where the algorithm has access to offline data [28], or can *iteratively* collect data [29] were also considered. In the context of *online* stabilization, i.e., when offline data is not available, work has derived stabilizing controllers for LTV systems through the use of predictions of θ_t, e.g., [30]. Finally, another line of work focuses on designing regret-optimal controllers for LTV systems [31], [6], [4], [5], [32]. However, with the exception of [30], existing work on *online* control of unknown LTV systems share the common assumption of either of open-loop stability or knowledge of an offline stabilizing controller. Moreover, the disturbances are generally assumed to be zero or stochastic noise independent of the states and inputs.

In this paper, we propose an online algorithm for stabilizing unknown LTV systems under bounded, potentially adversarial disturbances. Our approach uses convex body chasing (CBC), which is an online learning problem where one must choose a sequence of points within sequentially presented convex sets with the aim of minimizing the sum of distances between the chosen points [33], [34]. CBC has emerged as a promising tool in online control, with most work making connections to a special case called *nested* convex body chasing (NCBC), where the convex sets are sequentially nested within the previous set [35], [36]. In particular, [37] first explored the use of NCBC for learning-based control of time-invariant nonlinear systems. NCBC was also used in combination with System Level Synthesis to design a distributed controller for networked systems [24] and in combination with model predictive control [38] for LTI system control as a promising alternative to system identification based methods. However, this line of work depends fundamentally on the time invariance of the system,

which results in nested convex sets. LTV systems do not yield nested sets and therefore represent a significant challenge.

This work addresses this challenge and presents a novel online control scheme (Algorithm 1) based on CBC (non-nested) techniques that guarantees bounded-input-bounded-output (BIBO) stability as a function of the total model variation $\sum_{t=1}^{\infty} \|\theta_t - \theta_{t-1}\|$, without predictions or offline data under bounded and potentially adversarial disturbances for unknown LTV systems (Theorem 1). This result implies that when the total model variation is finite or growing sublinearly, BIBO stability of the closed loop is guaranteed (Corollaries 1 and 2). In particular, our result depends on a refined analysis of the CBC technique (Lemma 1) and is based on the perturbation analysis of the Lyapunov equation. This contrasts with previous NCBC-based works for time-invariant systems, where the competitive ratio guarantee of NCBC directly applies and the main technical tool is the robustness of the model-based controller, which is a proven using a Lipschitz bound of a quadratic program in [24] and is directly assumed to exist in [37].

We illustrate the proposed algorithm via numerical examples in Section IV to corroborate the stability guarantees. We demonstrate how the proposed algorithm can be used for data collection and complement data-driven methods like [27], [3], [28]. Further, the numerics highlight that the proposed algorithm can be efficiently implemented by leveraging the linearity of (1) despite the computational complexity of CBC algorithms in general (see Section III-B for details).

Notation. We use $\mathbb{S}^{n-1}$ to denote the unit sphere in $\mathbb{R}^n$ and $\mathbb{N}_+$ for positive integers. For $t, s \in \mathbb{N}_+$, we use $[t : s]$ as shorthand for the set of integers $\{t, t+1, \ldots, s\}$ and $[t]$ for $\{1, 2, \ldots, t\}$. Unless otherwise specified, $\|\cdot\|$ is the operator norm. We use $\rho(\cdot)$ for the spectral radius of a matrix.

II. PRELIMINARIES

In this section, we state the model assumptions underlying our work and review key results for convex body chasing, which we leverage in our algorithm design and analysis.

A. Stability and model assumptions

We study the dynamics in (1) and make the following standard assumptions about the dynamics.

Assumption 1 *The disturbances are bounded: $\|w_t\|_\infty \leq W$ for all $t \geq 0$.*

Assumption 2 *The unknown time-varying system matrices $\{\theta_t\}_{t=1}^{\infty}$ belong to a known (potentially large) polytope Θ such that $\theta_t \in \Theta$ for all t. Moreover, there exists $\kappa > 0$ such that $\|\theta\| \leq \kappa$ and θ is stabilizable for all $\theta \in \Theta$.*

Bounded and non-stochastic (potentially adversarial) disturbances is a common model both in the online learning and control problems [39], [40]. Since we make no assumptions on how large the bound W is, Assumption 1 models a variety of scenarios, such as bounded and/or correlated stochastic noise, state-dependent disturbances, e.g., the linearization

and discretization error for nonlinear continuous-time dynamics, and potentially adversarial disturbances. Assumption 2 is standard in learning-based control, e.g. [41], [42].

We additionally assume there is a quadratic known cost function of the state and control input at every time step t to be minimized, e.g. $x_t^\top Q x_t + u_t^\top R u_t$, with $Q, R \succ 0$. For a given LTI system model $\theta = [A \; B]$ and cost matrices Q, R, we denote $K = \text{LQR}(\theta; Q, R)$ as the optimal feedback gain for the corresponding infinite-horizon LQR problem.

Remark 1 *Representing model uncertainty as convex compact parameter sets where every model is stabilizable is not always possible. In particular, if a parameter set Θ has a few singular points where (A, B) loses stabilizability such as when $B = 0$, a simple heuristic is to ignore these points in the algorithm since we assume the underlying true system matrices θ_t must be stabilizable.*

B. Convex body chasing

Convex Body Chasing (CBC) is a well-studied online learning problem [35], [36]. At every round $t \in \mathbb{N}_+$, the player is presented a convex body/set $\mathcal{K}_t \subset \mathbb{R}^n$. The player selects a point $q_t \in \mathcal{K}_t$ with the objective of minimizing the cost defined as the total path length of the selection for T rounds, e.g., $\sum_{t=1}^{T} \|q_t - q_{t-1}\|$ for a given initial condition $q_0 \notin \mathcal{K}_1$. There are many known algorithms for the CBC problem with a *competitive ratio* guarantee such that the cost incurred by the algorithm is at most a constant factor from the total path length incurred by the offline optimal algorithm which has the knowledge of the entire sequence of the bodies. We will use CBC to select θ_t's that are consistent with observed data.

1) The nested case: A special case of CBC is the *nested convex body chasing* (NCBC) problem, where $\mathcal{K}_t \subseteq \mathcal{K}_{t-1}$. A known algorithm for NCBC is to select the *Steiner point* of $\mathcal{K}_t$ at t [36]. The Steiner point of a convex set $\mathcal{K}$ can be interpreted as the average of the extreme points of $\mathcal{K}$ and is defined as $\text{st}(\mathcal{K}) := \mathbb{E}_{v:\|v\| \leq 1}[g_{\mathcal{K}}(v)]$, where $g_{\mathcal{K}}(v) := \text{argmax}_{x \in \mathcal{K}} v^\top x$ and the expectation is taken with respect to the uniform distribution over the unit ball. The intuition is that Steiner point remains "deep" inside of the (nested) feasible region so that when this point becomes infeasible due to a new convex set, this convex set must shrink considerably, which indicates that the offline optimal must have moved a lot. Given the initial condition $q_0 \notin \mathcal{K}_1$, the Steiner point selector achieves competitive ratio of $O(n)$ against the offline optimal such that for all $T \in \mathbb{N}_+$, $\sum_{t=1}^{T} \|\text{st}(\mathcal{K}_t) - \text{st}(\mathcal{K}_{t-1})\| \leq O(n) \cdot \text{OPT}$, where OPT is the offline optimal total path length. There are many works that combine the Steiner point algorithm for NCBC with existing control methods to perform learning-based online control for LTI systems, e.g., [24], [37], [38].

2) General CBC: For general CBC problems, we can no longer take advantage of the nested property of the convex bodies. One may consider naively applying NCBC algorithms when the convex bodies happen to be nested and restarting the NCBC algorithm when they are not. However, due to the myopic nature of NCBC algorithms, which try

to remain deep inside of each convex set, they no longer guarantee a competitive ratio when used this way. Instead, [33] generalizes ideas from NCBC and proposes an algorithm that selects the *functional Steiner point* of the *work function*.

Definition 1 (Functional Steiner point) *For a convex function $f : \mathbb{R}^n \to \mathbb{R}$, the functional Steiner point of f is*

$$\mathsf{st}(f) = -n \cdot \fint_{v:\|v\|=1} f^*(v) \, v \, dv, \tag{2}$$

where $\fint_{x \in \mathcal{S}} f(x) dx$ denotes the normalized value $\frac{\int_{x \in \mathcal{S}} f(x) dx}{\int_{x \in \mathcal{S}} 1 dx}$ of $f(x)$ on the set $\mathcal{S}$, and

$$f^*(v) := \inf_{x \in \mathbb{R}^n} f(x) - \langle x, v \rangle \tag{3}$$

is the Fenchel conjugate of f.

The CBC algorithm selects the functional Steiner point of the *work function*, which records the smallest cost required to satisfy a sequence of requests while ending in a given state, thereby encapsulating information about the offline-optimal cost for the CBC problem.

Definition 2 (Work function) *Given an initial point $q_0 \in \mathbb{R}^n$, and convex sets $\mathcal{K}_1, \ldots, \mathcal{K}_t \subset \mathbb{R}^n$, the work function at time step t evaluated at a point $x \in \mathbb{R}^n$ is given by:*

$$\omega_t(x) = \min_{q_s \in \mathcal{K}_s} \|x - q_t\| + \sum_{s=1}^{t} \|q_s - q_{s-1}\|. \tag{4}$$

Importantly, it is shown that the functional Steiner points of the work functions are valid, i.e., $\mathsf{st}(\omega_t) \in \mathcal{K}_t$ for all t [33]. On a high level, selecting the functional Steiner point of the work function helps the algorithm stay competitive against the currently estimated offline optimal cost via the work function, resulting in a competitive ratio of n against the offline optimal cost (OPT) for general CBC problems,

$$\sum_{t=1}^{T} \|\mathsf{st}(\omega_t) - \mathsf{st}(\omega_{t+1})\| \le n \cdot \text{OPT}. \tag{5}$$

Given the non-convex nature of (2) and (4), we note that, in general, it is challenging to compute the functional Steiner point of the work function. However, in the proposed algorithm, we are able to leverage the linearity of the LTV systems and numerically approximate both objects with efficient computation in Section III-B.

III. MAIN RESULTS

We present our proposed online control algorithm to stabilize the unknown LTV system (1) under bounded and potentially adversarial disturbances in Algorithm 1. After observing the latest transition from x_t, u_t to x_{t+1} at $t+1$ according to (1) (line 2), the algorithm constructs the set of all feasible models $\widehat{\theta}_t$'s (line 3) such that the model is *consistent* with the observation, i.e., there exists an admissible disturbance $\widehat{w}_t$ satisfying Assumption 1 such that the state transition from x_t, u_t to x_{t+1} can be explained by the tuple $(\widehat{\theta}_t, \widehat{w}_t)$. We call this set the *consistent model set* $\mathcal{P}_t$

and we note that the unknown true dynamics $\theta_t = [A_t \ B_t]$ belongs to $\mathcal{P}_t$. The algorithm then selects a *hypothesis* model out of the consistent model set $\mathcal{P}_t$ using the CBC algorithm by computing the functional Steiner point (2) of the work function (4) with respect to the history of the consistent parameter sets $\mathcal{P}_1, \ldots, \mathcal{P}_t$ (line 4). In particular, we present an efficient implementation of the functional Steiner point chasing algorithm in Section III-B by taking advantage of the fact that $\mathcal{P}_t$'s are polytopes that can be described by intersection of half-spaces. The implementation is summarized in Algorithm 2. Based on the selected hypothesis model $\widehat{\theta}_t$, a certainty-equivalent LQR controller is synthesized (line 5) and the state-feedback control action is computed (line 6).

Note that, by construction, at time step $t \in \mathbb{N}_+$ we perform certainty-equivalent control $\widehat{K}_{t-1}$ based on a hypothesis model $\widehat{\theta}_{t-1}$ computed using retrospective data, even though the control action ($u_t = \widehat{K}_{t-1} x_t$) is applied to the dynamics (θ_t) that we do not yet have any information about. In order to guarantee stability, we would like for $\widehat{K}_{t-1}$ to be stabilizing the "future" dynamics (θ_t). This is the main motivation behind our choice of the CBC technique instead of regression-based techniques for model selection. Thanks to the competitive ratio guarantee (5) of the functional Steiner point selector, when the true model variation is "small," our previously selected hypothesis model will stay "consistent" in the sense that $\widehat{K}_{t-1}$ can be stabilizing for θ_t despite the potentially adversarial or state-dependent disturbances. On the other hand, when the true model variation is "large," $\widehat{K}_{t-1}$ does not stabilize θ_t, and we see growth in the state norm. Therefore, our final state bound is in terms of the total variation of the true model.

We show in the next section that, by drawing connections between the stability of the closed-loop system and the path length cost of the selected hypothesis model via CBC, we are able to stabilize the unknown LTV system without any identification requirements, e.g., the selected hypothesis models in Algorithm 1 need not be close to the true models. It is observed that even in the LTI setting, system identification can result in large-norm transient behaviors with numerical stability issues if the underlying unknown system is open-loop unstable or under non-stochastic disturbances; thus motivating the development of NCBC-based online control methods [25], [24], [37]. In the LTV setting, it is not sufficient to use NCBC ideas due to the time-variation of the model; however, the intuition for the use of CBC is similar. In fact, it can be additionally beneficial to bypass identification in settings where the true model is a moving target, thus making identification more challenging. We illustrate this numerically in Section IV.

A. Stability Analysis

The main result of this paper is the BIBO stability guarantee for Algorithm 1 in terms of the true model variation and the disturbance bound. We sketch the proof in this section and refer interested reader to the full version of this paper [43] for the formal proof. This result depends on a refined analysis of the competitive ratio for the functional Steiner

Algorithm 1: UNKNOWN LTV STABILIZATION

Input: $W > 0$, $\Theta \subset \mathbb{R}^{n \times (n+m)}$
Initialize: $u_0 = 0$, $\widehat{\theta}_0 \in \Theta$

1 **for** $t + 1 = 1, 2, \ldots$ **do**
2   Observe x_{t+1}
3   Construct consistent set
      $\mathcal{P}_t := \{\theta = [A, B] : \|x_{t+1} - Ax_t - Bu_t\|_\infty \leq W\} \cap \Theta$
4   Select hypothesis model $\widehat{\theta}_t \leftarrow \text{CBC}(\{\mathcal{P}_s\}_{s=1}^t; \widehat{\theta}_0)$
5   Synthesize controller $\widehat{K}_t \leftarrow \text{LQR}\left(\widehat{\theta}_t; Q, R\right)$
6   Compute feedback control input $u_{t+1} = \widehat{K}_t x_{t+1}$
7 **end**

Algorithm 2: CBC

Input: $\mathcal{P}_1, \ldots, \mathcal{P}_t, \widehat{\theta}_0, N$
Output: $\widehat{\theta}_t$

1 **for** $k = 0, 1, \ldots N$ **do**
2   Sample v_i uniformly from $\mathbb{S}^{n-1}$
3     $h_i \leftarrow (12)$
4 **end**
5 $\widehat{\theta}_t \leftarrow \text{proj}_{\Theta \cap \mathcal{P}_t}\left(-\frac{n}{N}\sum_{i=1}^N h_i v_i\right)$

point chasing algorithm introduced in [33], stated as follows.

Lemma 1 (Partial-path competitive ratio) *For $t \in \mathbb{N}_+$, let $s, e \in [t]$ and $s < e$, and let $\Theta \subset \mathbb{R}^n$ be a convex compact set. Denote $\widehat{\Delta}_{[s,e]} := \sum_{\tau=s+1}^e \|\text{st}(\omega_\tau) - \text{st}(\omega_{\tau-1})\|_F$ as the partial-path cost of the functional Steiner point selector during interval $[s, e]$ and $\{\text{OPT}_\tau\}_{\tau=1}^t$ as the (overall) offline optimal selection for $\mathcal{K}_1, \ldots, \mathcal{K}_t \subset \Theta$. The functional Steiner point chasing algorithm has the following competitive ratio,*

$$\widehat{\Delta}_{[s,e]} \leq n\left(\text{dia}(\Theta) + 2\kappa + \sum_{\tau=s+1}^e \|\text{OPT}_\tau - \text{OPT}_{\tau-1}\|_F\right).$$

on interval $[s, e]$, where $\text{dia}(\Theta) := \max_{\theta_1, \theta_2 \in \Theta} \|\theta_1 - \theta_2\|_F$ denotes the diameter of Θ and $\kappa := \max_{\theta \in \Theta} \|\theta\|_F$.

    Proof: See Appendix A. ∎

Theorem 1 (BIBO Stability) *Under Assumption 1 and 2, the closed loop of (1) under Algorithm 1 is BIBO stable such that for all $t \geq 0$,*

$$\|x_t\| \leq W \cdot c_1 \sum_{s=0}^{t-2} c_2^{\Delta_{[s,t-1]}} \rho_L^{t-s}$$

where $\Delta_{[s,t-1]} := \sum_{\tau=s+1}^{t-1} \|\theta_\tau - \theta_{\tau-1}\|_F$ is the true model variation, W is the disturbance bound, and $c_1, c_2 > 0, \rho_L \in (0,1)$ are constants that depend on the system-theoretical quantities of the worst-case model in the parameter set Θ.

    Proof Sketch: At a high level, the structure of our proof is as follows. We first use the fact that our time-varying feedback gain $\widehat{K}_t$ is computed according to a hypothesis model

from the *consistent* model set. Therefore, we can characterize the closed-loop dynamics in terms of the consistent models $\widehat{\theta}_t$ and $\widehat{K}_t$. Specifically, consider a time step t where we take the action $u_t = \widehat{K}_{t-1} x_t$ after observing x_t. Then, we observe $x_{t+1} = A_t x_t + B_t u_t + w_t$ and select a new hypothesis model $\widehat{\theta}_t = [\widehat{A}_t \ \widehat{B}_t]$ that is consistent with this new observation. Since we have selected a consistent hypothesis model, there is some admissible disturbance $\widehat{w}_t$ satisfying Assumption 1 such that

$$x_{t+1} = \left(A_t + B_t \widehat{K}_{t-1}\right) x_t + w_t = \left(\widehat{A}_t + \widehat{B}_t \widehat{K}_{t-1}\right) x_t + \widehat{w}_t.$$

Without loss of generality, we assume initial condition $x_0 = 0$. We therefore have

$$x_t = \widehat{w}_{t-1} + \sum_{s=0}^{t-2} \prod_{\tau \in [t-1:s+1]} \left(\widehat{A}_\tau + \widehat{B}_\tau \widehat{K}_{\tau-1}\right) \widehat{w}_s. \quad (6)$$

We have two main challenges in bounding $\|x_t\|$ in (6):
1) $\widehat{K}_t$ is computed using $\widehat{\theta}_t$ in Algorithm 1, but is applied to the next time step $\widehat{\theta}_{t+1}$. While we know $\rho(\widehat{A}_t + \widehat{B}_t \widehat{K}_t) < 1$, in (6) we have $\widehat{K}_{t-1}$ instead of $\widehat{K}_t$.
2) Naively applying submultiplicativity of the operator norm for (6) results in bounding $\left\|\left(\widehat{A}_\tau + \widehat{B}_\tau \widehat{K}_{\tau-1}\right)\right\|$. However, even if $\widehat{K}_{t-1}$ satisfies $\rho(\widehat{A}_t + \widehat{B}_t \widehat{K}_t) < 1$, in general the operator norm can be greater than 1.

To address the first challenge, our key insight is that by selecting hypothesis models via CBC technique, in any interval where the true model variation is small, our selected hypothesis model also vary little. Specifically, by Lemma 1, we can bound the partial-path variation of the selected hypothesis models with the true model partial-path variation $\Delta_{[s,e]}$ as follows.

$$\widehat{\Delta}_{[s,e]} \leq n\left(\text{dia}(\Theta) + 2\kappa + \sum_{\tau=s}^{e-1} \|\text{OPT}_{\tau+1} - \text{OPT}_\tau\|_F\right)$$
$$\leq n\left(\text{dia}(\Theta) + 2\kappa + \Delta_{[s,e]}\right). \quad (7)$$

where Θ and κ are from Assumption 2. A consequence of (7) is that, during intervals where the true model variation is small, we have $\left(\widehat{A}_t + \widehat{B}_t \widehat{K}_{t-1}\right) \approx \left(\widehat{A}_t + \widehat{B}_t \widehat{K}_t\right)$.

    For the second challenge, we leverage the concept of sequential strong stability [44], which allows bounding $\left\|\prod_{\tau \in [t-1:s+1]} \left(\widehat{A}_\tau + \widehat{B}_\tau \widehat{K}_{\tau-1}\right)\right\|$ approximately with $\prod_{\tau \in [t-1:s+1]} \rho\left(\widehat{A}_\tau + \widehat{B}_\tau \widehat{K}_\tau\right)$ times $O\left(\exp(\Delta_{[s,t-1]})\right)$.

    We now sketch the proof. The helper lemmas and the formal proof can be found in [43]. Consider $L_t, H_t \in \mathbb{R}^{n \times n}$ with $H_t \succ 0$ such that

$$\widehat{A}_t + \widehat{B}_t \widehat{K}_{t-1} := H_t^{1/2} L_t H_t^{-1/2}.$$

We use I_s as shorthand for the interval $[t-1 : s+1]$. Then each summand in (6) can be bounded as

$$\left\|\prod_{\tau \in I_s} \left(\widehat{A}_\tau + \widehat{B}_\tau \widehat{K}_{\tau-1}\right)\right\|$$

$$\leq \underbrace{\left\|H_{t-1}^{1/2}\right\|\left\|H_{s+1}^{-1/2}\right\|}_{(a)} \underbrace{\prod_{k\in I_{s+1}}\left\|H_k^{-1/2}H_{k-1}^{1/2}\right\|}_{(b)} \underbrace{\prod_{\tau\in I_s}\|L_\tau\|}_{(c)} \qquad (8)$$

Therefore showing BIBO stability comes down to bounding individual terms in (8). In particular we will show that by selecting appropriate H_t and L_t, term (a) is bounded by a constant C_H that depends on system theoretical properties of the worst-case parameter in Θ. For (b) and (c), we isolate the instances when

$$\left\|\widehat{\theta}_t - \widehat{\theta}_{t-1}\right\|_F \leq \epsilon \qquad (9)$$

for some chosen $\epsilon > 0$. For instances where (9) holds, we use the perturbation analysis of the Lyapunov equation involving the matrix $\widehat{A}_t + \widehat{B}_t\widehat{K}_{t-1}$ (Lemma 6 from the extended version [43] for (b) and [45, Theorem 8] for (c)) to bound (b) and (c) in terms of the partial-path movement of the selected parameters $\widehat{\Delta}_{[s,e]} := \sum_{\tau=s+1}^{e}\|\mathsf{st}(\omega_{\tau+1}) - \mathsf{st}(\omega_\tau)\|_F$. Specifically, Lemma 6 in [43] implies

$$\left\|H_t^{-1/2}H_{t-1}^{1/2}\right\| \leq \begin{cases} e^{\frac{\beta\left\|\widehat{\theta}_t-\widehat{\theta}_{t-1}\right\|_F}{2}}, & \text{if (9) holds} \\ \bar{H} & \text{otherwise,} \end{cases} \qquad (10)$$

where $\beta, \bar{H} > 1$ are constants. We also show that from [45, Theorem 8] ,

$$\|L_t\| \leq \begin{cases} \rho_L & \text{if (9) holds} \\ \bar{L} & \text{otherwise,} \end{cases} \qquad (11)$$

for $\rho_L \in (0,1)$ and $\bar{L} > 1$ a constant. We now plug (10) and (11) into (8). Denote by $n_{[s,t]}$ the number of pairs $(\tau, \tau-1)$ with $s+1 \leq \tau \leq t-1$ where (9) fails to hold. Let $\Delta_{[s,e]} := \sum_{\tau=s+1}^{e}\|\theta_\tau - \theta_{\tau-1}\|_F$ be the true model partial-path variation. Then (8) can be bounded as

$$\left\|\prod_{\tau\in[t-1:s+1]}\left(\widehat{A}_\tau + \widehat{B}_\tau\widehat{K}_{\tau-1}\right)\right\|$$

$$\leq C_H \cdot \bar{H}^{n_{[s,t]}} \cdot e^{\frac{\beta\widehat{\Delta}_{[s+1,t-1]}}{2}} \cdot \bar{L}^{n_{[s,t]}} \cdot \rho_L^{t-s-\widehat{n}_{[s,t]}-1}$$

$$\leq C_H\left(\frac{\bar{L}\bar{H}}{\rho_L}\right)^{\frac{\widehat{\Delta}_{[s,t-1]}}{\epsilon_*}} e^{\frac{\beta\widehat{\Delta}_{[s+1,t-1]}}{2}} \cdot \rho_L^{t-s-1}$$

$$\leq C_H\left(\frac{\bar{L}\bar{H}}{\rho_L}\right)^{\frac{\bar{n}\left(\text{dia}(\Theta)+2\kappa+\Delta_{[s,t-1]}\right)}{\epsilon_*}} e^{\frac{\beta\bar{n}\left(\text{dia}(\Theta)+2\kappa+\Delta_{[s+1,t-1]}\right)}{2}} \cdot \rho_L^{t-s-1}$$

$$=: c \cdot c_2^{\Delta_{[s,t-1]}}\rho_L^{t-s},$$

for constants c, c_2 and $\bar{n} := n(n+m)$ for the dimension of the parameter space for A_t, B_t. In the second inequality, we used the observation that $n_{[s,t]} \leq \frac{\widehat{\Delta}_{[s,t-1]}}{\epsilon}$ and in the last inequality we used Lemma 1. Combined with (6) and Assumption 1, this proves the desired bound. ∎

An immediate consequence of Theorem 1 is that when the model variation in (1) is bounded or sublinear, Algorithm 1 guarantees BIBO stability. This is summarized below.

Corollary 1 (Bounded variation) *Suppose* (1) *has model variation* $\Delta_{[0,t]} \leq M$ *for a constant* M. *Then,*

$$\sup_t \|x_t\| \leq \frac{W \cdot c_1 \cdot c_2^M}{1-\rho_L}.$$

Corollary 2 (Unbounded but sublinear variation) *Let* $\alpha \in (0,1)$ *and* $t \in \mathbb{N}_+$. *Suppose* (1) *is such that for each* $k \leq t$, $\Delta_{[k,k+1]} \leq \delta_t := 1/t^{(1-\alpha)}$, *implying a total model variation* $\Delta_{[0,t]} = O(t^\alpha)$. *Then for large enough* t, $\rho_L c_2^{\delta_t} \leq \frac{1+\rho_L}{2}$, *and therefore*

$$\|x_k\| \leq W \cdot c_1 \sum_{i=0}^{k}\left(\rho_L c_2^{\delta_t}\right)^i \leq \frac{2c_1}{1-\rho_L}.$$

Corollary 1 can be useful for scenarios where the mode of operation of the system changes infrequently and for systems such that $\theta(t) \to \theta^\star$ as $t \to \infty$ [46]. As an example, consider power systems where a prescribed set of lines can potentially become disconnected from the grid and thus change the grid topology. Corollary 2 applies to slowly drifting systems [47].

B. Efficient implementation of CBC

In general, implementation of the functional Steiner point of the work function may be computationally inefficient. However, by taking advantage of the LTV structure, we can design an efficient implementation in our setting. The key observation here is that for each t, $\mathcal{P}_t$ (Algorithm 1, line 3) can be described by the intersection of half-spaces because the ambient parameter space Θ is assumed to be a polytope and the observed online transition data from x_t, u_t to x_{t+1} specifies two half-space constraints at each time step due to linearity of (1). Our approach to approximate the functional Steiner point for chasing the consistent model sets is inspired by [34] where second-order cone programs (SOCPs) are used to approximate the (nested set) Steiner point of the sublevel set of the work functions for chasing half-spaces.

Denote $\{(a_i, b_i)\}_{i=1}^{p_t}$ as the collection of p_t half-space constraints describing $\mathcal{P}_t$, i.e., $a_i^\top\theta \leq b_i$. To approximate the integral for the functional Steiner point (2) of ω_t, we sample N number of random directions $v \in \mathbb{S}^{n-1}$, evaluate the Fenchel conjugate of the work function ω_t^* at each v with an SOCP, and take the empirical average. Finally we project the estimated functional Steiner point back to the set of consistent model $\mathcal{P}_t \cap \Theta$. Even though the analytical functional Steiner point (2) is guaranteed to be a member of the consistent model set, the projection step is necessary because we are integrating numerically, which may result in an approximation that ends up outside of the set. We summarize this procedure in Algorithm 2. Specifically, given a direction $v \in \mathbb{S}^{n-1}$, the Fenchel conjugate of the work function at time step t is

$$\omega_t^*(v) = \inf_{x\in\mathbb{R}^n} \omega_t(x) - \langle x, v\rangle$$

$$= \min_{\substack{x\in\mathbb{R}^n \\ q_s\in\mathcal{K}_s}} \sum_{s=1}^{t}\|q_s - q_{s-1}\| + \|x - q_t\| - \langle x, v\rangle.$$

This can be equivalently expressed as the following SOCP with decision variables $x, q_1, \ldots, q_t, \lambda, \lambda_1, \ldots, \lambda_t$:

$$\min_{\substack{x, q_1, \ldots, q_t \\ \lambda, \lambda_1, \ldots, \lambda_t}} \quad \lambda + \sum_{s=1}^{t} \lambda_s - \langle v, x \rangle$$

$$\text{s.t.} \quad \|q_s - q_{s-1}\| \leq \lambda_s, \quad \text{for } s \in [t] \qquad (12)$$
$$\|x - q_t\| \leq \lambda$$
$$a_i^\top q_s \leq b_i, \quad \text{for } i \in [p_s], s \in [t]$$

Another potential implementation challenge is that the number of constraints in the SOCP (12) grows linearly with time due to the construction of the work function (4). This is a common drawback of online control methods based on CBC and NCBC techniques and can be overcome through truncation or over-approximation of the work functions in practice. Additionally, if the LTV system is periodic with a known period, then we can leverage Algorithm 1 during the initial data collection phase. Once representative (persistently exciting) data is available, one could employ methods like [3] to generate a stabilizing controller for the unknown LTV system. In Section IV, we show that data collection via Algorithm 1 results in a significantly smaller state norm than random noise injection when the system is unstable.

IV. SIMULATION

In this section, we demonstrate Algorithm 1 in two LTV systems. Both of the systems we consider are open-loop unstable, thus the algorithms must work to stabilize them. We use the same algorithm parameters for both, with $\Theta = [-2, 3]^2$, LQR cost matrices $Q = I$ and $R = 1$.

A. Example 1: Markov linear jump system

We consider the following Markov linear jump system (MLJS) model from [48], with

$$A_1 = \begin{bmatrix} 1.5 & 1 \\ 0 & 0.5 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0.6 & 0 \\ 0.1 & 1.2 \end{bmatrix}, \quad B_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

$$B_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \Pi = \begin{bmatrix} 0.8 & 0.2 \\ 0.1 & 0.9 \end{bmatrix}$$

where Π is the transition probability matrix from θ_1 to θ_2 and vice versa. We inject uniformly random disturbances such that $w_t \in \{-10\mathbb{1}, -3\mathbb{1}, 3\mathbb{1}\}$ where $\mathbb{1}$ is the all-one vector. We set the disturbances to be zero for the last 10 time steps to make explicit the stability of the closed loop. We implement certainty-equivalent control based on online least squares (OLS) with different sliding window sizes $L = 5, 10, 20$ and a exponential forgetting factor of 0.95 [49] as the baselines.

We show two different MLJS models generated from 2 random seeds and show the results in Figure 1. For both systems, the open loop is unstable. In Figure 1a the OLS-based algorithms fail to stabilize the system for window size of $L = 20$, while stabilizing the system but incurring larger state norm than the proposed algorithm for $L = 5, 10$. On the other hand, in Figure 1b, OLS with $L = 5$ results in unstable closed loop. This example highlights the challenge of OLS-based methods, where the choice of window size is crucial for the performance. Since the underlying LTV

system is unknown and our goal is to control the system *online*, it is unclear how to select appropriate window size to guarantee stability for OLS-based methods a priori. In contrast, Algorithm 1 does not require any parameter tuning.

We note that while advanced least-squares based identification techniques that incorporate sliding window with *variable* length exist, e.g. [4], [49], due to the unknown system parameters, it is unclear how to choose the various algorithm parameters such as thresholds for system change detection. Therefore, we only compare Algorithm 1 against fixed-length sliding window OLS methods as baselines.
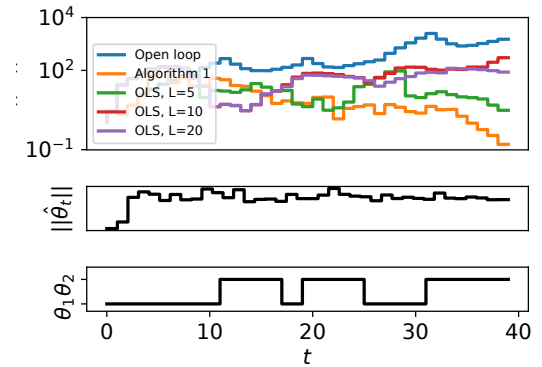
(a) closed loop of the system generated with seed # 1
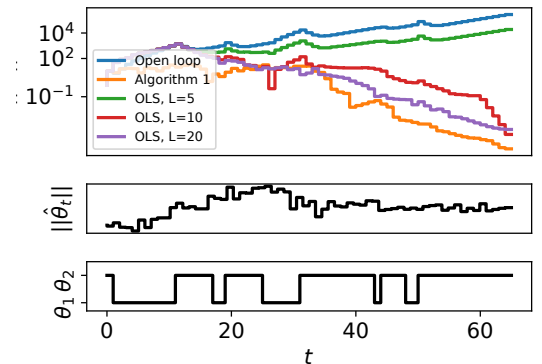
(b) closed loop of the system generated with seed # 2

Fig. 1: Markov linear jump system for two different random seeds. For each seed: **Top** plot shows the state norm trajectories of the proposed algorithm, certainty-equivalent control based on online least squares (OLS) with different sliding window sizes, and the open loop. **Middle** plot shows the norm of the selected hypothesis model via Algorithm 2. **Bottom** plot shows the true model switches.

B. Example 2: LTV system

Our second example highlights that Algorithm 1 is a useful data-collection alternative to open-loop random noise injection. We consider the LTV system from [3], [28], with

$$A(k) = \begin{bmatrix} 1.5 & 0.0025k \\ -0.1\cos(0.3k) & 1 + 0.05^{3/2}\sin(0.5k)\sqrt{k} \end{bmatrix},$$

$$B(k) = 0.05 \begin{bmatrix} 1 \\ \frac{0.1k+2}{0.1k+3} \end{bmatrix}.$$

where we modified $A(1,1)$ from 1 to 1.5 to increase the instability of the open loop in the beginning; thus making it
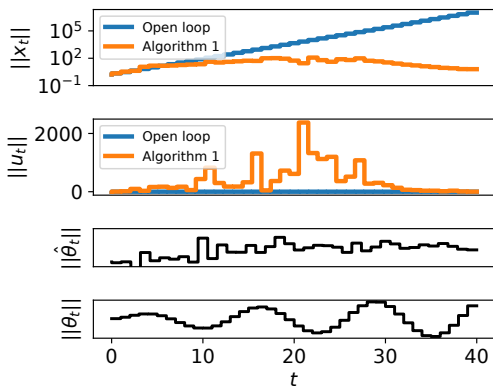
Fig. 2: Simulation result for the LTV system in example 2. Here we plot the the state and control norm, as well as the selected hypothesis model via CBC $\widehat{\theta}_t$ and true models θ_t.

more challenging to stabilize. We consider no disturbances here, which is a common setting in direct data-driven control, e.g., [3], [26], [27]. In particular, we compare the proposed algorithm against randomly generated bounded inputs from $\mathsf{UNIF}[-1,1]$. We also modify the control inputs from Algorithm 1 to be $u_t = \widehat{K}_{t-1} x_t + \eta_t \cdot \mathbb{1}$ with $\eta_t \sim \mathsf{UNIF}[-1,1]$ so that we can collect rich data in the closed loop. This is motivated by the growing body of data-driven control methods such as [3], [27], [28] that leverage sufficiently rich offline data to perform control design for unknown LTV systems. However, most of these works directly inject random inputs for data collection. It is evident in Figure 2 that when the open-loop system is unstable it may be undesirable to run the system without any feedback control. Therefore, Algorithm 1 complements existing data-driven methods by allowing safe data collection with significantly better transient behavior.

V. CONCLUDING REMARKS

In this paper, we propose a model-based approach for stabilizing an unknown LTV system under arbitrary non-stochastic disturbances in the sense of bounded input bounded output under the assumption of infrequently changing or slowly drifting dynamics. Our approach uses ideas from convex body chasing (CBC), which is an online problem where an agent must choose a sequence of points within sequentially presented convex sets with the aim of minimizing the sum of distances between the chosen points. The algorithm requires minimal tuning compared to online least squares based control with sliding window. Future work includes sharpening the stability analysis to go beyond the BIBO guarantee, which will require controlling the difference between the estimated disturbances and true disturbances. Another direction is to extend the current results to the networked case, similar to [24].

REFERENCES

[1] D. Deka, S. Backhaus, and M. Chertkov, "Structure learning in power distribution networks," *IEEE Transactions on Control of Network Systems*, vol. 5, no. 3, pp. 1061–1074, Sep. 2018.

[2] P. Gradu, E. Hazan, and E. Minasyan, "Adaptive regret for control of time-varying dynamics," *arXiv preprint arXiv:2007.04393*, 2020.

[3] B. Nortmann and T. Mylvaganam, "Data-driven control of linear time-varying systems," in *2020 59th IEEE Conference on Decision and Control (CDC)*. IEEE, 2020, pp. 3939–3944.

[4] Y. Luo, V. Gupta, and M. Kolar, "Dynamic regret minimization for control of non-stationary linear dynamical systems," *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, vol. 6, no. 1, pp. 1–72, 2022.

[5] Y. Lin, J. Preiss, E. Anand, Y. Li, Y. Yue, and A. Wierman, "Online adaptive controller selection in time-varying systems: No-regret via contractive perturbations," *arXiv preprint arXiv:2210.12320*, 2022.

[6] E. Minasyan, P. Gradu, M. Simchowitz, and E. Hazan, "Online control of unknown time-varying dynamical systems," *Advances in Neural Information Processing Systems*, vol. 34, pp. 15 934–15 945, 2021.

[7] R. Tedrake, "Underactuated robotics: Learning, planning, and control for efficient and agile machines course notes for mit 6.832," *Working draft edition*, vol. 3, p. 4, 2009.

[8] P. Falcone, F. Borrelli, H. E. Tseng, J. Asgari, and D. Hrovat, "Linear time-varying model predictive control and its application to active steering systems: Stability analysis and experimental validation," *International Journal of Robust and Nonlinear Control: IFAC-Affiliated Journal*, vol. 18, no. 8, pp. 862–875, 2008.

[9] K. S. Tsakalis and P. A. Ioannou, *Linear time-varying systems: control and adaptation*. Prentice-Hall, Inc., 1993.

[10] R. Tóth, *Modeling and identification of linear parameter-varying systems*. Springer, 2010, vol. 403.

[11] J. Mohammadpour and C. W. Scherer, *Control of linear parameter varying systems with applications*. Springer Science & Business Media, 2012.

[12] W. Zhang, Q.-L. Han, Y. Tang, and Y. Liu, "Sampled-data control for a class of linear time-varying systems," *Automatica*, vol. 103, pp. 126–134, 2019.

[13] R. Mojgani and M. Balajewicz, "Stabilization of linear time-varying reduced-order models: A feedback controller approach," *International Journal for Numerical Methods in Engineering*, vol. 121, no. 24, pp. 5490–5510, 2020.

[14] G. Goel and B. Hassibi, "Regret-optimal estimation and control," *IEEE Transactions on Automatic Control*, 2023.

[15] K. Tsakalis and P. Ioannou, "Adaptive control of linear time-varying plants," *Automatica*, vol. 23, no. 4, pp. 459–468, 1987.

[16] J.-J. Slotine and J. Coetsee, "Adaptive sliding controller synthesis for non-linear systems," *International Journal of Control*, vol. 43, no. 6, pp. 1631–1651, 1986.

[17] R. Marino and P. Tomei, "Adaptive control of linear time-varying systems," *Automatica*, vol. 39, no. 4, pp. 651–659, 2003.

[18] M. Verhaegen and X. Yu, "A class of subspace model identification algorithms to identify periodically and arbitrarily time-varying systems," *Automatica*, vol. 31, no. 2, pp. 201–216, 1995.

[19] B. Bamieh and L. Giarre, "Identification of linear parameter varying models," *International Journal of Robust and Nonlinear Control: IFAC-Affiliated Journal*, vol. 12, no. 9, pp. 841–853, 2002.

[20] T. Sarkar, A. Rakhlin, and M. Dahleh, "Nonparametric system identification of stochastic switched linear systems," in *2019 IEEE 58th Conference on Decision and Control (CDC)*, 2019.

[21] S. Dean, S. Tu, N. Matni, and B. Recht, "Safely learning to control the constrained linear quadratic regulator," in *2019 American Control Conference (ACC)*. IEEE, 2019, pp. 5582–5588.

[22] S. Talebi, S. Alemzadeh, N. Rahimi, and M. Mesbahi, "On regularizability and its application to online control of unstable lti systems," *IEEE Transactions on Automatic Control*, 2021.

[23] S. Lale, K. Azizzadenesheli, B. Hassibi, and A. Anandkumar, "Reinforcement learning with fast stabilization in linear dynamical systems," in *International Conference on Artificial Intelligence and Statistics*. PMLR, 2022, pp. 5354–5390.

[24] J. Yu, D. Ho, and A. Wierman, "Online adversarial stabilization of unknown networked systems," *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, vol. 7, no. 1, pp. 1–43, 2023.

[25] X. Chen and E. Hazan, "Black-box control for linear dynamical systems," in *Conference on Learning Theory*. PMLR, 2021.

[26] M. Rotulo, C. De Persis, and P. Tesi, "Online learning of data-driven controllers for unknown switched linear systems," *Automatica*, vol. 145, p. 110519, 2022.

[27] S. Baros, C.-Y. Chang, G. E. Colon-Reyes, and A. Bernstein, "Online data-enabled predictive control," *Automatica*, vol. 138, p. 109926, 2022.

[28] B. Pang, T. Bian, and Z.-P. Jiang, "Data-driven finite-horizon optimal control for linear time-varying discrete-time systems," in *2018 IEEE Conference on Decision and Control (CDC)*, 2018, pp. 861–866.

[29] S.-J. Liu, M. Krstic, and T. Başar, "Batch-to-batch finite-horizon lq control for unknown discrete-time linear systems via stochastic extremum seeking," *IEEE Transactions on Automatic Control*, vol. 62, no. 8, pp. 4116–4123, 2016.

[30] G. Qu, Y. Shi, S. Lale, A. Anandkumar, and A. Wierman, "Stable online control of linear time-varying systems," in *Learning for Dynamics and Control*. PMLR, 2021, pp. 742–753.

[31] Y. Ouyang, M. Gagrani, and R. Jain, "Learning-based control of unknown linear systems with thompson sampling," *arXiv preprint arXiv:1709.04047*, 2017.

[32] Y. Han, R. Solozabal, J. Dong, X. Zhou, M. Takac, and B. Gu, "Learning to control under time-varying environment," *arXiv preprint arXiv:2206.02507*, 2022.

[33] M. Sellke, "Chasing convex bodies optimally," in *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms*. SIAM, 2020, pp. 1509–1518.

[34] C. Argue, A. Gupta, Z. Tang, and G. Guruganesh, "Chasing convex bodies with linear competitive ratio," *Journal of the ACM (JACM)*, vol. 68, no. 5, pp. 1–10, 2021.

[35] N. Bansa, M. Böhm, M. Eliáš, G. Koumoutsos, and S. W. Umboh, "Nested convex bodies are chaseable," in *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*. SIAM, 2018, pp. 1253–1260.

[36] S. Bubeck, B. Klartag, Y. T. Lee, Y. Li, and M. Sellke, "Chasing nested convex bodies nearly optimally," in *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms*. SIAM, 2020.

[37] D. Ho, H. Le, J. Doyle, and Y. Yue, "Online robust control of nonlinear systems with large uncertainty," in *International Conference on Artificial Intelligence and Statistics*. PMLR, 2021, pp. 3475–3483.

[38] C. Yeh, J. Yu, Y. Shi, and A. Wierman, "Robust online voltage control with an unknown grid topology," in *Proceedings of the Thirteenth ACM International Conference on Future Energy Systems*, 2022, pp. 240–250.

[39] B. Ramasubramanian, B. Xiao, L. Bushnell, and R. Poovendran, "Safety-critical online control with adversarial disturbances," in *2020 59th IEEE Conference on Decision and Control (CDC)*. IEEE, 2020.

[40] E.-W. Bai, R. Tempo, and H. Cho, "Membership set estimators: size, optimal inputs, complexity and relations with least squares," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 42, no. 5, pp. 266–277, 1995.

[41] A. Cohen, T. Koren, and Y. Mansour, "Learning linear-quadratic regulators efficiently with only sqrt(t) regret," in *International Conference on Machine Learning*. PMLR, 2019, pp. 1300–1309.

[42] N. Agarwal, B. Bullins, E. Hazan, S. Kakade, and K. Singh, "Online control with adversarial disturbances," in *International Conference on Machine Learning*. PMLR, 2019, pp. 111–119.

[43] J. Yu, V. Gupta, and A. Wierman, "Online stabilization of unknown linear time-varying systems," *arXiv preprint arXiv:2304.02878*, 2023.

[44] A. Cohen, A. Hasidim, T. Koren, N. Lazic, Y. Mansour, and K. Talwar, "Online linear quadratic control," in *International Conference on Machine Learning*. PMLR, 2018, pp. 1029–1038.

[45] M. Simchowitz and D. Foster, "Naive exploration is optimal for online lqr," in *International Conference on Machine Learning*. PMLR, 2020.

[46] W. Hahn *et al.*, *Stability of motion*. Springer, 1967, vol. 138.

[47] F. Amato, G. Celentano, and F. Garofalo, "New sufficient conditions for the stability of slowly varying linear systems," *IEEE Transactions on Automatic Control*, vol. 38, no. 9, pp. 1409–1411, 1993.

[48] J. Xiong and J. Lam, "Stabilization of discrete-time markovian jump linear systems via time-delayed controllers," *Automatica*, vol. 42, no. 5, pp. 747–753, 2006.

[49] J. Jiang and Y. Zhang, "A revisit to block and recursive least squares for parameter estimation," *Computers & Electrical Engineering*, vol. 30, no. 5, pp. 403–416, 2004.

APPENDIX

A. Proof of Lemma 1

We have

$$\sum_{\tau=s+1}^{e} \left\| \widehat{\theta}_\tau - \widehat{\theta}_{\tau-1} \right\|_F = \sum_{\tau=s+1}^{e} \left\| \mathsf{st}(\omega_\tau) - \mathsf{st}(\omega_{\tau-1}) \right\|_F$$

$$\overset{(a)}{\le} n \int_v \left(\sum_{\tau=s+1}^{e} \left| \omega_\tau^*(v) - \omega_{\tau-1}^*(v) \right| \right) v \, dv$$

$$\overset{(b)}{=} n \int_v \left(\sum_{\tau=s+1}^{e} \omega_\tau^*(v) - \omega_{\tau-1}^*(v) \right) v \, dv$$

$$= n \int_v \left(\omega_e^*(v) - \omega_s^*(v) \right) v \, dv$$

$$\overset{(c)}{\le} n \cdot \left(\min_x \omega_e(x) - \min_y \omega_s(y) + 2\kappa \right) \quad (13)$$

where (a) is due to the definition (2). For (b), we used the observation that $\omega_t^*(v)$ is non-decreasing in time. For (c), by definition of the Fenchel conjugate (3), we have that $\omega_e^*(v) = \inf_x \omega_e(x) - \langle x, v \rangle$. Denote $(x^\star, q_1^\star, \ldots, q_e^\star)$ as the optimal solution to the problem $\min_x \omega_e(x)$. It is clear that $\omega_e^*(v) \le \omega_e(x^\star) - \langle x^\star, v \rangle \le \min_x \omega_e(x) + \kappa$ where in the last inequality we used Cauchy-Shwarz and $\kappa := \max_{\theta \in \Theta} \|\theta\|_F$. Similarly, we also have $\omega_s^*(v) \ge \inf_y \omega_s(y) - \kappa$.

Denote $\mathrm{OPT}_{[0,e]}$ as the minimizing trajectory $(\mathrm{OPT}_0, \ldots, \mathrm{OPT}_e)$ to $\min_x \omega_e(x)$ where $\mathrm{argmin}_x \omega_e(x) = \mathrm{OPT}_e$. This last equality is by the observation that if $x^\star := \mathrm{argmin}_x \omega_e(x) \ne \mathrm{OPT}_e$, then $\omega_e(\mathrm{OPT}_e) \le \omega_e(x^\star)$ by definition (4), thus contradicting that $x^\star$ is defined to be the minimizer of ω_e. We also denote $\mathrm{INT}_{[0,s]}$ as the minimizing trajectory to $\min_y \omega_s(y)$. To reduce notation, we denote $\Delta_{[s,e]}^{\mathrm{OPT}} := \sum_{\tau=s+1}^{e} \|\mathrm{OPT}_\tau - \mathrm{OPT}_{\tau-1}\|_F$ and $\Delta_{[s,e]}^{\mathrm{INT}} := \sum_{\tau=s+1}^{e} \|\mathrm{INT}_\tau - \mathrm{INT}_{\tau-1}\|_F$. Then we have

$$(13) = n \cdot \left(\Delta_{[0,e]}^{\mathrm{OPT}} - \Delta_{[0,s]}^{\mathrm{INT}} + 2\kappa \right)$$

$$\overset{(c)}{\le} n \cdot \left(\Delta_{[0,e]}^{\mathrm{OPT}} - \Delta_{[0,s]}^{\mathrm{OPT}} + \mathsf{dia}(\Theta) + 2\kappa \right)$$

$$= n \cdot \left(\Delta_{[s,e]}^{\mathrm{OPT}} + \mathsf{dia}(\Theta) + 2\kappa \right).$$

where (c) holds because if $\sum_{\tau=1}^{s} \|\mathrm{OPT}_\tau - \mathrm{OPT}_{\tau-1}\|_F > \sum_{\tau=1}^{s} \|\mathrm{INT}_\tau - \mathrm{INT}_{\tau-1}\|_F + \mathsf{dia}(\Theta)$ and $\mathrm{OPT}_{[0,s]} \ne \mathrm{INT}_{[0,s]}$, then we can replace the $[0,s]$ portion of the optimal trajectory $\mathrm{OPT}_{[0,e]}$ with $\mathrm{INT}_{[0,s]}$ and achieve a lower cost for $\omega_e(\mathrm{OPT}_e)$, thus contradicting the optimality of $\mathrm{OPT}_{[0,e]}$. To see why the fictitious trajectory $(\mathrm{INT}_{[0,s]}, \mathrm{OPT}_{[s+1,e]})$ achieves lower cost than $\mathrm{OPT}_{[0,e]}$, we compare the total movement cost during the interval $[0, s+1]$,

$$\sum_{\tau=1}^{s} \|\mathrm{INT}_\tau - \mathrm{INT}_{\tau-1}\|_F + \|\mathrm{OPT}_{s+1} - \mathrm{INT}_s\|_F$$

$$\le \sum_{\tau=1}^{s} \|\mathrm{INT}_\tau - \mathrm{INT}_{\tau-1}\|_F + \|\mathrm{OPT}_{s+1} - \mathrm{OPT}_s\|_F$$
$$+ \|\mathrm{OPT}_s - \mathrm{INT}_s\|_F$$

$$\le \sum_{\tau=1}^{s} \|\mathrm{INT}_\tau - \mathrm{INT}_{\tau-1}\|_F + \|\mathrm{OPT}_{s+1} - \mathrm{OPT}_s\|_F + \mathsf{dia}(\Theta)$$

$$< \sum_{\tau=1}^{s} \|\mathrm{OPT}_\tau - \mathrm{OPT}_{\tau-1}\|_F + \|\mathrm{OPT}_{s+1} - \mathrm{OPT}_s\|_F,$$

which means the fictitious trajectory achieves lower overall cost. Therefore (c) must hold. ■