

# Data-driven Practical Stabilization of Nonlinear Systems via Chain Policies: Sample Complexity and Incremental Learning

Roy Siegelmann and Enrique Mallada

**Abstract**—We propose a method for data-driven practical stabilization of nonlinear systems with provable guarantees, based on the concept of *Nonparametric Chain Policies (NCPs)*. The approach employs a normalized nearest-neighbor rule to assign, at each state, a finite-duration control signal derived from stored data, after which the process repeats. Unlike recent works that model the system as linear, polynomial, or polynomial fraction, we only assume the system to be locally Lipschitz. Our analysis builds on the framework of Recurrent Lyapunov Functions (RLFs), which enable data-driven certification of (practical) stability using standard norm functions instead of requiring the explicit construction of a classical Lyapunov function. To extend this framework, we introduce the concept of Recurrent Control Lyapunov Functions (RCLFs), which can certify the existence of an NCP that practically stabilizes an arbitrarily small  $c$ -neighborhood of an equilibrium point. We also provide an explicit sample complexity guarantee of  $\mathcal{O}((3/\rho)^d \log(R/c))$  number of trajectories—where  $R$  is the domain radius,  $d$  the state dimension, and  $\rho$  a system-dependent constant. The proposed Chain Policies are nonparametric, thus allowing new verified data to be readily incorporated into the policy to either improve convergence rate or enlarge the certified region. Numerical experiments illustrate and validate these properties.

## I. INTRODUCTION

*Data-driven control methods* offer a novel paradigm for synthesizing controllers directly from trajectory observations, potentially bypassing the need for accurate system models while reducing computational burden and conservativeness of classical control synthesis [1], [2]. Recent years have witnessed significant progress in data-driven control. These approaches, as well as their level of maturity, depend considerably on the underlying system properties. For *linear systems*, the field has substantially matured: LMI-based formulas [3] and convex programs [4], [5] can transform trajectories into stabilizing feedback controllers with robustness [3], [5]–[7], performance [4], [5], [8] and sample complexity guarantees [9], [10].

For *nonlinear systems*, several approaches have been proposed, with methods highly dependent on the implicit assumptions made on the nonlinear system class and the control synthesis methodology. One prolific line of works considers dynamics formed from dictionary-based hypothesis classes—e.g., using polynomials [11]–[13], fractions of polynomials [14], or general nonlinear functions [15]—and formulate semidefinite programs that render policies with a

wide variety of guarantees, including contraction-based stability [16]–[18] or robustness [13]. Other methods employ general learning techniques to learn models or policies and leverage intrinsic system properties to provide different guarantees, e.g., Koopman operator methods that exploit spectral properties [19], sample complexity analysis for stochastic dynamics [20], and conformal prediction approaches for statistical robustness [21].

Despite the effectiveness of these methods in synthesizing controllers with guarantee, many questions remain unanswered. First, sample complexity guarantees are typically technique-dependent and do not provide clear understanding of how data requirements scale with explicit system properties, such as state dimension or attainable performance levels, or the specific hypothesis class considered. Second, computational complexity of optimization-based methods scales poorly with dictionary size and state dimension. Third, incorporating new data necessitates resolving the underlying optimization problem, often requiring complete recomputation and discarding previous work. As a result, there remains a need for flexible data-driven approaches that can adapt to new information without structural assumptions while providing transparent performance-data trade-offs.

To address these challenges, we introduce the concept of *Nonparametric Chain Policies (NCPs)*, a data-driven approach that requires only Lipschitz assumptions on the system dynamics while providing explicit sample complexity guarantees for practical stabilization. NCPs employ a normalized nearest-neighbor rule to assign finite-duration control signals from a stored library of verified trajectories, enabling direct use of data without parametric modeling or optimization re-solving when new data arrives. Our theoretical guarantees build on the framework of Recurrent Lyapunov Functions [22], [23], which we extend here for the control setting by introducing here the notion of Recurrent Control Lyapunov Functions (RCLFs, Section III).

**Contributions.** Our approach offers three key advantages over existing methods:

- 1) **Explicit sample complexity:** NCPs achieve practical exponential stabilization using  $\mathcal{O}((3/\rho)^d \log(R/\delta))$  sample trajectories, with transparent scaling in dimension  $d$ , target radius  $R$ , precision  $\delta$ , and a system-dependent parameter  $\rho$ .
- 2) **Incremental learning:** The nonparametric nature of NCPs allows for new verified data to be seamlessly incorporated to expand a certified region, or improve performance, without discarding previous guarantees or re-solving optimization problems.

Roy Siegelmann is with the Department of Applied Mathematics and Statistics, Johns Hopkins University, 3400 N Charles St., Baltimore, MD 21218, USA. Email: rsiegel15@jhu.edu. Enrique Mallada is with the Department of Electrical and Computer Engineering, Johns Hopkins University, 3400 N Charles St., Baltimore, MD 21218, USA. Email: mallada@jhu.edu.

- 3) **Performance-complexity trade-offs:** The framework explicitly controls the trade-off between sample requirements and performance through a user-specified parameter  $\rho$  that relates best achievable performance and the performance guaranteed by the NCP.

**Organization.** Section II introduces preliminaries. Section III presents Recurrent Control Lyapunov Functions (R-CLFs) and stability guarantees. Section IV defines Nonparametric Chain Policies and establishes sample complexity results. Section V demonstrates the approach on nonlinear benchmarks, and Section VI concludes.

**Notation.**  $\|\cdot\|$  denotes an arbitrary norm on  $\mathbb{R}^n$ . Given  $x \in \mathbb{R}^n$  and  $r > 0$ , we define the closed ball of radius  $r$  centered at  $x$  as  $B_r(x) := \{y \in \mathbb{R}^n \mid \|y - x\| \leq r\}$ . For a scalar  $a \in \mathbb{R}$ , we write  $[a]_+ := \max\{a, 0\}$ . For a set  $S \subseteq \mathbb{R}^n$  and a point  $x \in \mathbb{R}^n$ , the signed distance from  $x$  to  $S$  is defined as

$$\text{sd}(x, S) := \begin{cases} \inf_{y \in \partial S} \|y - x\|, & \text{if } x \notin S, \\ -\inf_{y \in \partial S} \|y - x\|, & \text{if } x \in S. \end{cases}$$

## II. PRELIMINARIES

We consider a nonlinear control system:

$$\dot{x}(t) = f(x(t), u(t)), \quad (1)$$

with state  $x(t) \in \mathbb{R}^n$  and input  $u(t) \in U \subseteq \mathbb{R}^m$ . We define

$$\mathcal{U}^{(a,b]} := \{u : (a, b] \rightarrow U \mid u \text{ measurable}\},$$

as the set of admissible control signals on interval  $(a, b]$ , and set  $\mathcal{U} := \mathcal{U}^{(0,\infty)}$ . Given  $u_0 \in \mathcal{U}^{(0,a]}$  and  $u_1 \in \mathcal{U}^{(0,b]}$ , their concatenation  $u_0 u_1 \in \mathcal{U}^{(0,a+b]}$  is defined by

$$(u_0 u_1)(t) = \begin{cases} u_0(t), & t \in (0, a], \\ u_1(t), & t \in (a, a + b]. \end{cases}$$

More generally, for a sequence of control signals  $u_n \in \mathcal{U}^{(0,\tau_n]}$ , with  $\tau_n > 0$ ,  $\forall n \in \mathbb{N}$ , we further use  $u_{[n]} := u_0 u_1 \dots u_n$ , and  $u_{[\infty]} = \lim_{n \rightarrow \infty} u_{[n]}$ . In some occasions we slightly abuse notation by using  $u$  interchangeably to represent instantaneous inputs in  $U$  and signals in  $\mathcal{U}^{(a,b]}$ ; the intended meaning will always be clear from context.

For an initial state  $x \in \mathbb{R}^n$  and control signal  $u \in \mathcal{U}^{(0,a]}$ , we denote by  $\phi(t, x, u)$  the solution of (1) for  $t \in (0, a]$ . We further assume the following regularity conditions for (1).

**Assumption 1** (Forward Completeness). *The solutions of the control system (1) are **forward complete**. Specifically, for each initial condition  $x \in \mathbb{R}^n$  and every control signal  $u \in \mathcal{U}$ , the trajectory  $\phi(t, x, u)$  exists and remains bounded for all  $t \geq 0$ .*

**Assumption 2** (Uniform Lipschitz Continuity). *The vector field  $f(x, u)$  of system (1) is locally Lipschitz continuous in  $x$ , uniformly with respect to  $u$ . More precisely, for every compact set  $S \subseteq \mathbb{R}^n$ , there exists a constant  $L_S \geq 0$  such that*

$$\|f(y, u) - f(x, u)\| \leq L_S \|y - x\|, \quad \forall x, y \in S, \forall u \in U.$$

### A. Practical Exponential Stabilizability

In this work, we aspire to render an equilibrium point  $x^* \in \mathbb{R}^n$  practical exponentially stable.

**Definition 1** (Equilibrium Point). *A point  $x^* \in \mathbb{R}^n$  is an **equilibrium point** of system (1) if there exists a control input  $u^* \in U$  such that  $f(x^*, u^*) = 0$ .*

**Definition 2** ((Practical) Exponential Stabilizability). *Let  $S \subseteq \mathbb{R}^n$ . The equilibrium  $x^*$  of system (1) is said to be:*

- (i) **Exponentially Stabilizable** on  $S$  if for every  $x \in S$ , there exists a control signal  $u \in \mathcal{U}$  satisfying

$$\|\phi(t, x, u) - x^*\| \leq K e^{-\lambda t} \|x - x^*\|, \quad \forall t \geq 0; \quad (2)$$

- (ii) **Practically Exponentially Stabilizable** on  $S$  if for every  $x \in S$ , there exists a control signal  $u \in \mathcal{U}$  satisfying

$$\|\phi(t, x, u) - x^*\| \leq K e^{-\lambda t} \|x - x^*\| + c, \quad \forall t \geq 0, \quad (3)$$

for constants  $K \geq 1$ ,  $\lambda > 0$ , and  $c \geq 0$ .

It is well known from the topological entropy literature that it is impossible to exponentially stabilize a system, i.e., achieve (2), using a finite number of control signals [24]. We will therefore aim to enforce the weaker notion of practical exponential stability, i.e., (3), which follows the terminology of [25], [26].

### B. Recurrent Lyapunov Functions

To provide guarantees for our data-driven stabilization framework, we build on the theory of *Recurrent Lyapunov Functions* (RLFs) [22], [23]. Unlike classical Lyapunov functions, which require strict decrease along trajectories, RLFs only require a decrease at a sequence of *recurrent times*. This relaxation broadens the class of certificates available, while still ensuring exponential stability.

We begin with the notion of containment times, which we define for general trajectories of the controlled system.

**Definition 3** (Containment Times). *Given a set  $S \subset \mathbb{R}^n$ , an initial state  $x \in \mathbb{R}^n$ , and an input  $u \in \mathcal{U}$ , the set of containment times is*

$$T_S(x, u) := \{t \in \mathbb{R}_{>0} \mid \phi(t, x, u) \in S\}.$$

For constants  $a, b > 0$  we define

$$T_S(x, u; a, b) := T_S(x, u) \cap (a, a + b],$$

and for convenience  $T_S(x, u; b) := T_S(x, u; 0, b)$ .

We now recall the definition of an RLF in the *autonomous* case, where the trajectory is uniquely determined by the initial condition. In this case we write  $\phi(t, x)$  for the flow of the system.

**Definition 4** (Recurrent Lyapunov Function). *Let  $S \subset \mathbb{R}^n$  be a compact set with  $x^* \in \text{int}(S)$ . A continuous function  $V : \mathbb{R}^n \rightarrow \mathbb{R}_{\geq 0}$  is called a Recurrent Lyapunov Function (RLF) over  $S$  with rate  $\alpha > 0$  and horizon  $\tau > 0$  if*

$$\min_{t \in T_S(x; \tau)} e^{\alpha t} V(\phi(t, x)) \leq V(x), \quad \forall x \in S,$$

where  $T_S(x; \tau) := \{t \in (0, \tau] \mid \phi(t, x) \in S\}$ .

It will also be useful to characterize the set of points that can be reached within a finite interval of time.

**Definition 5** (Reachable Tube). *For the control system (1), a constant  $\tau > 0$ , and a set  $S \subset \mathbb{R}^n$ , we denote the  $\tau$ -reachable tube from  $S$  within  $\tau$  units of time by*

$$\mathcal{R}^\tau(S) = \bigcup_{x \in S, u \in \mathcal{U}, t \in [0, \tau]} \{\phi(t, x, u)\}.$$

### III. RECURRENT CONTROL LYAPUNOV FUNCTIONS

As mentioned before, our guarantees rely on the theory of Recurrent Lyapunov Functions (RLFs) from [22]. In this section, we extend this notion to the control setting, introducing Recurrent Control Lyapunov Functions (R-CLFs), and illustrate how they can be used to certify practical stabilizability. Though RLFs and R-CLFs have been shown to certify stability, asymptotic stability and exponential stability [22], [23], our focus here is on practical exponential stability and thus we will use the following definition.

**Definition 6** (Recurrent Control Lyapunov Function (R-CLF)). *Consider the control system (1) with equilibrium  $x^* \in \mathbb{R}^n$ . Let  $S \subseteq \mathbb{R}^n$  be a set satisfying  $x^* \in \text{int}(S)$ . A continuous function  $V : \mathbb{R}^n \rightarrow \mathbb{R}_{\geq 0}$  is a **Recurrent Control Lyapunov Function (R-CLF)** over  $S$  if the following conditions hold:*

(i) **Positive Definiteness and Linear Bounds:** *There exist constants  $a_1, a_2 > 0$  such that*

$$a_1 \|x - x^*\| \leq V(x) \leq a_2 \|x - x^*\|, \quad \forall x \in S. \quad (4)$$

(ii) **Control  $\alpha$ -Exponential  $(\tau, \delta)$ -Recurrence:** *There exist constants  $\tau, \alpha > 0$  and  $\delta \geq 0$  such that for every  $x \in S$ , there exists  $u \in \mathcal{U}^{[0, \tau]}$  satisfying*

$$\min_{t \in T_S(x, u; \tau)} e^{\alpha t} (V(\phi(t, x, u)) - \delta) \leq [V(x) - \delta]_+. \quad (5)$$

The following lemma characterizes the long term behavior of the control system (1) under the controls  $u \in \mathcal{U}$  that are build upon concatenation of controls satisfying property (ii) of Definition 6.

**Lemma 1** (Characterization of R-CLF). *Let assumptions 1 and 2 hold. Consider an equilibrium  $x^*$  of (1) and a compact set  $S$  satisfying  $x^* \in \text{int}(S)$ . A function  $V : \mathbb{R}^n \rightarrow \mathbb{R}_{\geq 0}$  satisfying (4) is a Recurrent Control Lyapunov Function (R-CLF) over  $S$  if and only if there exists parameters  $\alpha, \tau > 0$  and  $\delta \geq 0$  such that for any  $x \in S$  there is a sequence  $\{t_n\}_{n \in \mathbb{N}}$  and  $u \in \mathcal{U}$  satisfying the following conditions:*

$$\lim_{n \rightarrow +\infty} t_n = +\infty, \text{ with } t_{n+1} - t_n \in (0, \tau], \quad (6a)$$

$$\phi(t_n, x, u) \in S, \text{ and} \quad (6b)$$

$$V(\phi(t_n, x, u)) - \delta \leq \begin{cases} e^{-\alpha t_n} (V(x) - \delta), & n \leq \bar{n}, \\ 0, & \text{o.w.}, \end{cases} \quad (6c)$$

for a non necessarily finite  $\bar{n} \in \mathbb{N} \cup \{\infty\}$ .

*Proof of Lemma 1.* We prove each direction separately.

**Necessity ( $\Rightarrow$ ):** Suppose that  $V$  is a Recurrent Control Lyapunov Function (R-CLF) over the compact set  $S$ . By Definition 6, there exist constants  $\alpha, \tau > 0$  and  $\delta \geq 0$  such that, for any  $x \in S$ , there exists  $\bar{u} \in \mathcal{U}^{[0, \tau]}$  satisfying (5).

We will build  $u \in \mathcal{U}$  and the sequence  $\{t_n\}_{n \in \mathbb{N}}$  inductively. Let  $t_0 = 0$ ,  $x_0 := x \in S$ , and define for  $n \geq 0$ ,

$$\tau_n := \max \left\{ \arg \min_{t \in T_S(x_n, \bar{u}_n; \tau)} e^{\alpha t} (V(\phi(t, x_n, \bar{u}_n)) - \delta) \right\}, \quad (7)$$

$$t_{n+1} := t_n + \tau_n, \quad \text{and} \quad x_{n+1} := \phi(\tau_n, x_n, u_n),$$

where  $\bar{u}_n \in \mathcal{U}^{[0, \tau]}$  is a control satisfying (5),  $u_n$  is its restriction to the interval  $(0, \tau_n]$ , and  $x_n \in S \forall n \in \mathbb{N}$ , by definition. Next, let  $u := \lim_{t \rightarrow \infty} u_{[n]}$ . Note that  $\phi(t_0, x, u) = x_0 \in S$ , and whenever for some  $n$ ,  $\phi(t_n, x, u) = x_n \in S$  that by the group property of the flow

$$\begin{aligned} \phi(t_{n+1}, x, u) &= \phi(t_{n+1} - t_n, \phi(t_n, x, u), u_n) \\ &= \phi(\tau_n, x_n, u_n) = x_{n+1} \in S, \end{aligned}$$

which by induction ensures that (6b) holds.

Next, from the recurrence condition (5), it follows that as long as  $V(\phi(t_n, x, u)) \geq \delta$ , then  $e^{\alpha \tau_n} (V(\phi(\tau_n, x_n, \bar{u}_n)) - \delta) \leq V(x_n) - \delta$ , which implies

$$e^{\alpha t_{n+1}} (V(\phi(t_{n+1}, x, u)) - \delta) \leq e^{\alpha t_n} (V(\phi(t_n, x, u)) - \delta), \quad (8)$$

and, in particular,

$$e^{\alpha t_n} (V(\phi(t_n, x, u)) - \delta) \leq V(x) - \delta, \quad \forall n \leq \bar{n},$$

where  $\bar{n}$  is the last instance with  $V(x_{\bar{n}}) \geq \delta$ . It also follows from (5) that, when  $\bar{n} < \infty$ , for all  $n \geq \bar{n} + 1$ ,  $V(\phi(t_n, x, u)) \leq \delta$ , which completes (6c).

To show (6a), we first note that by definition,  $t_{n+1} - t_n = \tau_n \in (0, \tau]$ . Next, we will show that  $t^* = \infty$ . Suppose not, i.e.,  $t^* < \infty$ . By continuity of  $\phi(t, x, u)$  and compactness of  $S$ ,  $\phi(t_n, x, u) \rightarrow \phi(t^*, x, u) \in S$ .

Now, let  $v_n := V(x_n) - \delta$  and  $v^* = V(\phi(t^*, x, u)) - \delta$ . If  $v^* > 0$ , it follows from (8) and the continuity of  $V$  that  $v_n \downarrow v^* := V(\phi(t^*, x, u)) - \delta$ , and  $e^{\alpha t^*} v^* \leq e^{\alpha t_{n+1}} v_{n+1} \leq e^{\alpha t_n} v_n$ . Thus, for large enough  $n$ ,  $t^* \in (t_n, t_n + \tau]$  and  $t^* > t_{n+1}$ , which implies

$$e^{\alpha t^* - t_n} v^* \leq e^{\alpha \tau_n} v_{n+1} \leq v_n,$$

which contradicts  $\tau_n$  being the max in (7). As similar argument holds when  $v^* < 0$ , and thus  $t^* = \infty$ .

**Sufficiency ( $\Leftarrow$ ):** To show that  $V$  is a R-CLF, it is sufficient for any  $x \in S$  to restrict the corresponding  $u \in \mathcal{U}$  that satisfies (6) to the interval  $[0, \tau]$  and choosing  $t_1$  from the sequence (6) to show (5).  $\square$

We will leverage Lemma 1 to prove (practical) exponential convergence of trajectories. To that end, we need to bound how much a trajectory can travel in between instance of exponential convergence (6c). The following lemma provides a mechanism to obtain such bounds. The proof is based on Grönwall's Lemma [27, Lemma A.1] and can be found in [28].

**Lemma 2** (Containment Lemma). *Let Assumption 2 hold. Consider a compact set  $S \subset \mathbb{R}^n$  and a constant  $\tau > 0$ . Then, for any  $x \in S, u \in \mathcal{U}$  the following holds:*

$$\max_{t \in [0, \tau]} d(\phi(t, x, u), S) \leq F_S \tau e^{L\tau}$$

where  $L := L_{\mathcal{R}^\tau(S)}$ .

We are now ready to show that R-CLFs as defined in Definition 6 guarantee exponential stabilizability.

**Theorem 1** (R-CLF Implies (Practical) Exponential Stabilizability). *Consider the control system (1) with equilibrium  $x^* \in \mathbb{R}^n$ , and let  $S$  be a set satisfying  $S \subseteq \mathbb{R}^n$  and  $x^* \in \text{int}(S)$ . Let Assumption 1 and Assumption 2 hold, and  $V : \mathbb{R}^n \rightarrow \mathbb{R}_{\geq 0}$  be a Recurrent Control Lyapunov Function over  $S$ , with constants  $\alpha, \tau > 0$ ,  $\delta \geq 0$ , and linear bound constants  $a_1, a_2 > 0$  from (4).*

*Then, the equilibrium  $x^*$  is (practically) exponentially stabilizable on  $S$  (when  $\delta > 0$ ). In particular, for every initial condition  $x \in S$ , a control signal  $u \in \mathcal{U}$  satisfying (6) for some sequence of times  $\{t_n\}_{n \in \mathbb{N}}$  ensures:*

$$\|\phi(t, x, u) - x^*\| \leq Ke^{-\lambda t} \|x - x^*\| + c, \quad \forall t \geq 0,$$

where

$$\lambda := \alpha, \quad K := \frac{a_2}{a_1} e^{\alpha\tau} (1 + L\tau e^{L\tau}), \quad \text{and} \quad c := \frac{\delta}{a_1} (1 + L\tau e^{L\tau}),$$

with  $L := L_{\mathcal{R}^\tau(S)}$ .

*Proof.* We will use the control  $u \in \mathcal{U}$  from Lemma 1 to prove this theorem. Given  $x \in S$ , by Lemma 1, there exists  $u \in \mathcal{U}$  and a sequence  $t_n, x_n := \phi(t_n, x) \in S$  satisfying (6). There are two cases. First, assume that  $n \leq \bar{n} \in \mathbb{N} \cup \{\infty\}$ . Since  $V(x_n) \leq e^{-\alpha t_n} (V(x) - \delta) + \delta$  and  $V(x_n) \geq a_1 \|x_n - x^*\|$ , it follows that

$$r_n := \|x_n - x^*\| \leq \frac{1}{a_1} V(x_n) \leq \frac{a_2}{a_1} e^{-\alpha t_n} \|x - x^*\| + \frac{\delta}{a_1}.$$

Consider any time  $t \in (t_n, t_{n+1}]$  and  $B_n := B_n(x^*) \cap S$ . By applying the containment lemma on  $B_n$  and using the fact that  $L \geq L_{\mathcal{R}^\tau(B_n)}$ , we get, using the triangle inequality:

$$\begin{aligned} \|\phi(t, x, u) - x^*\| &\leq \|x_n - x^*\| + \|\phi(t, x, u) - x_n\| \\ &\leq r_n + F_{r_n} \tau e^{L\tau} \leq (1 + L\tau e^{L\tau}) r_n \end{aligned}$$

where  $r_n := \|x_n - x^*\|$ . Then for any  $n \leq \bar{n}$  and any  $t \in (t_n, t_{n+1}]$  we have,

$$\begin{aligned} \|\phi(t, x, u) - x^*\| &\leq (1 + L\tau e^{L\tau}) r_n \\ &\leq (1 + L\tau e^{L\tau}) \left( \frac{a_2}{a_1} e^{-\alpha t_n} \|x - x^*\| + \frac{\delta}{a_1} \right) \\ &\leq Ke^{-\alpha\tau} e^{-\alpha t_n} \|x - x^*\| + \frac{\delta}{a_1} (1 + L\tau e^{L\tau}) \\ &\leq Ke^{-\alpha t} \|x - x^*\| + c \end{aligned}$$

where the last step follows, since  $t \leq t_{n+1} \leq t_n + \tau$ , which implies  $-t_n - \tau \leq -t$ , so that  $e^{-\alpha\tau} e^{-\alpha t_n} \leq e^{-\alpha t}$ .

If  $\bar{n} = \infty$  we are done. Otherwise, consider  $n > \bar{n}$ . We have  $r_n = \|x_n - x^*\| \leq \frac{V(x_n)}{a_1} \leq \frac{\delta}{a_1} \forall n > \bar{n}$ , and thus by Lemma 2 again, for all  $t > t_{\bar{n}}$

$$\|\phi(t, x, u) - x^*\| \leq (1 + L\tau e^{L\tau}) r_n \leq (1 + L\tau e^{L\tau}) \frac{\delta}{a_1} = c.$$

Thus, for all  $t \geq 0$ ,

$$\|\phi(t, x, u) - x^*\| \leq Ke^{-\lambda t} \|x - x^*\| + c,$$

as desired.  $\square$

Theorem 1 states that the existence of an R-CLF implies that  $x^*$  can be made practically exponentially stable. At the core of its proof is the fact that one can find a function  $V$  that satisfies the recurrent condition (5). A key observation of [23], is that condition (5) can be met by a norm, provided  $\tau$  and  $\alpha$  are properly chosen (c.f. [23, Theorem 6]). The caveat is, however, that in order to make R-CLFs practically useful, one would need to store, for each  $x \in S$ , a suitable  $u : [0, \tau] \rightarrow \mathcal{U}$  that ensures (5). In the next section we surprisingly show that when  $\delta > 0$ , only a finite number of such signals are needed.

#### IV. NON-PARAMETRIC CHAIN POLICIES

In the previous section, we introduced Recurrent Control Lyapunov Functions (R-CLFs) to characterize exponential stabilizability via carefully selected control signals. In this section, we propose *nonparametric chain policies*, a systematic, data-driven approach for generating these stabilizing signals. The proposed method aligns closely with recent developments in topological entropy—a notion quantifying the minimal complexity required to accomplish various control tasks (see, e.g., [24], [26], [29]). A distinctive feature of our method is that we do not assume the control signals can be generated online; instead, we explicitly store them in a finite set, called a *control alphabet* [28].

**Definition 7** (Control Alphabet). A *control alphabet* is a finite collection of control signals

$$\mathcal{A} := \{v_i : (0, \tau_i] \rightarrow \mathcal{U}\}_{i=0}^M,$$

where each  $v_i$  is piecewise continuous and defined over a duration  $\tau_i > 0$ .

The control alphabet provides a library of candidate signals. To deploy them, we assign specific controls to regions of influence within the state space. To aid this task we define an assignment set.

**Definition 8** (Assignment Set). An *assignment set* is a finite collection of verification triples

$$\mathcal{K} := \{(x_i, r_i, v_i)\}_{i=1}^N \subseteq \mathbb{R}^n \times \mathbb{R}_{>0} \times \mathcal{A},$$

where  $x_i \in \mathbb{R}^n$  is a center point,  $r_i > 0$  is its radius, and  $v_i \in \mathcal{A}$  is the control signal assigned to that region. The support of  $\mathcal{K}$  is

$$\text{Supp}(\mathcal{K}) := \bigcup_{i=1}^N B_{r_i}(x_i).$$

We denote  $N := |\mathcal{K}|$  as the size of the assignment set.

While an assignment set specifies regions of influence, it does not by itself resolve which control to apply when balls overlap, nor what to do when a state lies outside  $\text{Supp}(\mathcal{K})$ . To address this, we introduce a normalized nearest-neighbor rule with a fall-back option:

$$\iota_{\mathcal{K}}(x) := \begin{cases} \arg \min_{i: (x_i, r_i, v_i) \in \mathcal{K}} \frac{\|x - x_i\|}{r_i}, & r_{\mathcal{K}}(x) \leq 1, \\ 0, & \text{otherwise,} \end{cases}$$

where

$$r_{\mathcal{K}}(x) := \min_{(x_i, r_i, v_i) \in \mathcal{K}} \frac{\|x - x_i\|}{r_i},$$

and  $\iota_{\mathcal{K}}(x) = 0$  corresponds to selecting the default control  $v_0$ .

**Remark 1.** We designate  $v_0 \in \mathcal{A}$  as the default control. Unless otherwise stated, we take  $v_0(t) = u^* \in U$ ,  $\forall t \in [0, \tau_0)$ , where  $u^*$  is the equilibrium control of Definition 1.

The index map  $\iota_{\mathcal{K}}$  specifies, for any state  $x$ , which control from the assignment set (or the default control) should be applied. Building on this rule, we can now formalize the induced feedback policy.

**Definition 9** (Nonparametric Chain Policy). Given an assignment set  $\mathcal{K}$  and default control  $v_0$ , the **nonparametric chain policy** is given by the map  $\pi_{\mathcal{K}} : \mathbb{R}^n \rightarrow \mathcal{A}$ :

$$\pi_{\mathcal{K}}(x) := v_{\iota_{\mathcal{K}}(x)}.$$

**Remark 2.** The policy  $\pi_{\mathcal{K}}$  induces an infinite-horizon control signal  $u_{\mathcal{K},x} \in \mathcal{U}$  through concatenation. Starting with  $x_0 = x$  and the empty signal  $u_{[0]} = \emptyset$ , for each  $n \geq 0$  define

$$u_{[n+1]} = u_{[n]} v_{\iota_{\mathcal{K}}(x_n)}, \quad x_{n+1} = \phi(\tau_{\iota_{\mathcal{K}}(x_n)}, x_n, v_{\iota_{\mathcal{K}}(x_n)}). \quad (9)$$

The resulting control is then

$$u_{\mathcal{K},x} := \lim_{n \rightarrow \infty} u_{[n]}.$$

#### A. Convergence Guarantees of NCPs

With the nonparametric chain policy in place, we now turn to its stability properties. The following theorem establishes conditions under which such a policy renders the equilibrium  $x^*$  practically exponentially stable on a prescribed region.

**Theorem 2** (Practical Exponential Stabilization via Chain Policies). Consider an equilibrium point  $x^* \in \mathbb{R}^n$  of (1), and let  $S \subseteq \mathbb{R}^n$  be a set with  $x^* \in \text{int}(S)$ . Let  $\pi_{\mathcal{K}}$  denote a nonparametric chain policy associated with the assignment set  $\mathcal{K} = \{(x_i, r_i, v_i)\}_{i=1}^N$  and a default control  $v_0 \in \mathcal{A}$ , and define  $\tau := \max\{\tau_0, \tau_1, \dots, \tau_N\}$ , and let  $L := L_{\mathcal{R}^\tau(S)}$ . Suppose the following hold:

(i) **Covering.** There exists  $\varepsilon > 0$  such that

$$B_\varepsilon(x^*) \subset B_{\varepsilon(1+L\tau e^{L\tau})}(x^*) \subset \text{int}(S), \quad (10a)$$

$$\text{cl}(S \setminus B_\varepsilon(x^*)) \subseteq \text{Supp}(\mathcal{K}). \quad (10b)$$

(ii) **Verification.** For each  $(x_i, r_i, v_i) \in \mathcal{K}$  with  $\tau_i > 0$ ,

$$e^{\alpha\tau_i} (\|\phi(\tau_i, x_i, v_i) - x^*\| + r_i e^{L\tau_i}) \leq \|x_i - x^*\| - r_i, \quad (11a)$$

$$\text{sd}(\phi(\tau_i, x_i, v_i), S) + r_i e^{L\tau_i} \leq 0, \quad (11b)$$

(iii) **Equilibrium.** For all  $t \in [0, \tau_0)$ ,  $\phi(t, x^*, v_0) = x^*$ .

Then the equilibrium  $x^*$  is practically exponentially stable on  $S$  under the policy  $\pi_{\mathcal{K}}$ , with constants

$$\lambda = \alpha, \quad K = e^{\alpha\tau} (1 + L\tau e^{L\tau}), \quad c = \varepsilon(1 + L\tau e^{L\tau}).$$

*Proof.* We will show that the control  $u_{\mathcal{K},x}$  induced by the nonparametric chain policy  $\pi_{\mathcal{K}}$  admits a sequence of times  $\{t_n\}_{n \in \mathbb{N}}$  that satisfies the conditions of Lemma 1 for the function  $V(x) = \|x - x^*\|$  over the set  $S$ . This establishes

two points: (1)  $V = \|\cdot - x^*\|$  is an R-CLF with rate  $\alpha$ , and (2) the control  $u_{\mathcal{K},x}$  practically stabilizes  $x^*$  with exponential rate  $\alpha$  over  $S$ .

Let  $x_0 = x \in S$ ,  $t_0 = 0$ , and  $u_{[0]} = \emptyset$ . Define the sequences  $\{x_n\}$ ,  $\{t_n\}$ , and  $u_{[n]}$  according to (9), i.e.,

$$x_{n+1} = \phi(\tau_{\iota_{\mathcal{K}}(x_n)}, x_n, v_{\iota_{\mathcal{K}}(x_n)}), \quad u_{[n+1]} = u_{[n]} v_{\iota_{\mathcal{K}}(x_n)},$$

$$u_{\mathcal{K},x} = \lim_{n \rightarrow \infty} u_{[n]}, \quad t_{n+1} := t_n + \tau_{\iota_{\mathcal{K}}(x_n)}, \quad \forall n \geq 0.$$

By construction, for all  $n \geq 0$ ,

$$0 < \min_{i \in \{0, \dots, N\}} \tau_i \leq t_{n+1} - t_n \leq \max_{i \in \{0, \dots, N\}} \tau_i =: \tau,$$

so condition (6a) holds. Moreover, by induction one shows that for all  $n \geq 1$ ,

$$\phi(t_n, x, u_{\mathcal{K},x}) = x_n = \phi(\tau_{\iota_{\mathcal{K}}(x_{n-1})}, x_{n-1}, v_{\iota_{\mathcal{K}}(x_{n-1})}).$$

We claim that if  $x_n \in S$  then  $x_{n+1} \in S$ . Suppose first that  $x_n \in S \setminus B_\varepsilon(x^*)$ . By the covering condition (10b), there exists  $(x_i, r_i, v_i) \in \mathcal{K}$  such that  $x_n \in B_{r_i}(x_i)$  and the verification condition (11) holds. In particular, by (11b),

$$\begin{aligned} \text{sd}(x_{n+1}, S) &= \text{sd}(\phi(\tau_i, x_n, v_i), S) \\ &\leq \text{sd}(\phi(\tau_i, x_i, v_i), S) + r_i e^{L\tau_i} \\ &\leq 0, \end{aligned}$$

which implies  $x_{n+1} \in S$ .

If instead  $x_n \in B_\varepsilon(x^*)$ , then either  $\iota_{\mathcal{K}}(x_n) \neq 0$  and the above argument applies, or  $\iota_{\mathcal{K}}(x_n) = 0$ , in which case we apply  $v_0$  for time  $\tau_0$ . By the containment lemma applied to the ball  $B_\varepsilon(x^*)$ ,

$$\begin{aligned} \|x_{n+1} - x^*\| &= \|\phi(\tau_0, x_n, v_0) - x^*\| \\ &\leq \varepsilon + d(\phi(\tau_0, x_n, v_0), B_\varepsilon(x^*)) \\ &\leq \varepsilon + F_{B_\varepsilon(x^*)} \tau e^{L\tau_0} \\ &\leq \varepsilon(1 + L\tau_0 e^{L\tau_0}), \end{aligned} \quad (12)$$

so by (10a) and  $\tau_0 \leq \tau$  we conclude  $x_{n+1} \in S$ . Thus,  $x_n \in S$  implies  $x_{n+1} \in S$ , i.e., condition (6b) holds.

*Verification of (6c).* Let  $\delta := \varepsilon(1 + L\tau e^{L\tau})$  and  $\bar{n} := \inf\{n : \|x_n - x^*\| \leq \delta\}$ . If  $\iota_{\mathcal{K}}(x_n) = 0$  and  $\|x_n - x^*\| \leq \varepsilon \leq \delta$ , then by (12) we have  $\|x_{n+1} - x^*\| \leq \delta$ .

$$\begin{aligned} \text{If } \iota_{\mathcal{K}}(x_n) = i \neq 0, \text{ then from (11a) and } x_n \in B_{r_i}(x_i), \\ e^{\alpha(t_{n+1} - t_n)} \|x_{n+1} - x^*\| &\leq e^{\alpha\tau_i} (\|\phi(\tau_i, x_i, v_i) - x^*\| + r_i e^{L\tau_i}) \\ &\leq \|x_i - x^*\| - r_i \leq \|x_n - x^*\|, \end{aligned}$$

which implies

$$e^{\alpha(t_{n+1} - t_n)} (\|x_{n+1} - x^*\| - \delta) \leq \|x_n - x^*\| - \delta.$$

If  $\|x_n - x^*\| \leq \delta$  ( $n \geq \bar{n}$ ), then this inequality ensures  $\|x_{n+1} - x^*\| \leq \delta$ . If  $\|x_n - x^*\| > \delta$  ( $n < \bar{n}$ ), iterating yields

$$\|x_n - x^*\| - \delta \leq e^{-\alpha t_n} (\|x - x^*\| - \delta).$$

Hence for  $n \geq \bar{n}$ ,  $\|x_n - x^*\| \leq \delta$ , while for  $n < \bar{n}$  the excess above  $\delta$  decays exponentially. This verifies condition (6c).

By Lemma 1,  $V(x) = \|x - x^*\|$  is an R-CLF with rate  $\alpha$  over  $S$  and parameter  $\delta = \varepsilon(1 + L\tau e^{L\tau})$ . Therefore, Theorem 1 implies that  $x^*$  is practically exponentially stable on  $S$  under  $\pi_K$ , with constants

$$\lambda = \alpha, \quad K = e^{\alpha\tau}(1 + L\tau e^{L\tau}), \quad c = \varepsilon(1 + L\tau e^{L\tau}).$$

□

### B. Existence and Sample Complexity of NCPs

Theorem 2 establishes that nonparametric chain policies can guarantee practical exponential stability of a region around an equilibrium point that is appropriately covered by data points from  $\mathcal{K}$ . However, it is a priori not clear how many data points are needed to construct such policy, or even whether such a policy exists. The next result provides conditions for existence of Chain Policies as well as a bound on the sample complexity of such policies, i.e., the sizes of the assignment set  $\mathcal{K}$  and alphabet  $\mathcal{A}$  required to construct such policy.

**Theorem 3** (Existence and Sample Complexity of Chain Policies). *Consider the control system (1) with equilibrium  $x^* \in \mathbb{R}^n$ , and assume  $x^*$  is  $\lambda$ -exponentially stabilizable on  $\mathbb{R}^n$  with gain  $K > 0$ . Let  $S = B_R(x^*)$  with  $R > 0$ , and choose  $\varepsilon$  s.t.  $R > \varepsilon > 0$ . Fix any  $\alpha \in (0, \lambda)$  and choose*

$$\tau > \frac{\ln K}{\lambda - \alpha}, \quad L := L_{\mathcal{R}^\tau(S)}, \quad \rho := \frac{1 - Ke^{-(\lambda - \alpha)\tau}}{1 + e^{(L + \alpha)\tau}}. \quad (13)$$

*Then there exists a nonparametric chain policy  $\pi_K$  built from a finite assignment set of verification points  $\{(x_i, r_i)\}_{i=1}^N \subset S$  and associated controls  $\{v_i\}_{i=1}^N$  such that:*

(i) **Practical exponential stability.** *For every  $x \in S$ , the induced closed loop satisfies*

$$\|\phi(t, x, u_{K,x}) - x^*\| \leq Ce^{-\alpha t}\|x - x^*\| + c, \quad \forall t \geq 0,$$

*with  $C = e^{\alpha\tau}(1 + L\tau e^{L\tau})$  and  $c = \varepsilon(1 + L\tau e^{L\tau})$ .*

(ii) **Sample complexity.** *The number  $N$  of covering centers and controls satisfies*

$$N = O\left(\left(\frac{3}{\rho}\right)^d \log \frac{R}{c}\right).$$

**Remark 3** (Performance–Complexity Trade-off). *The definition of  $\rho$  in (13) reveals two contrasting regimes. When  $\lambda - \alpha$  is close to the lower bound  $\ln K/\tau$ , the numerator  $1 - Ke^{-(\lambda - \alpha)\tau}$  approaches zero, so  $\rho \approx 0$ . In this regime the guaranteed rate  $\alpha$  is nearly as fast as the best attainable  $\lambda$ , but the sample complexity bound  $O((3/\rho)^d)$  becomes extremely large. At the other extreme, when  $\alpha \ll \lambda$ , the term  $Ke^{-(\lambda - \alpha)\tau}$  vanishes, and  $\rho$  approaches  $1/(1 + e^{L\tau})$ . In this regime, far fewer samples are needed, but the realized performance  $\alpha$  is much slower than the system's intrinsic rate  $\lambda$ .*

*Thus,  $\rho$  quantifies the fundamental trade-off: choosing  $\alpha$  close to  $\lambda$  yields strong performance at the cost of high sample complexity, while smaller  $\alpha$  reduces sample requirements but sacrifices convergence speed.*

*Proof.* Since  $x^*$  is  $\lambda$ -exponentially stabilizable on  $S$  with constant  $K > 0$ , for each grid center  $x_i \in S$  we can select a

constant control  $v_i$  such that

$$\|\phi(t, x_i, v_i) - x^*\| \leq Ke^{-\lambda t}\|x_i - x^*\|, \quad \forall t \geq 0. \quad (14)$$

We construct the nonparametric chain policy  $\pi_{\mathcal{D}}$  by covering  $S$  with finitely many balls  $B_{r_i}(x_i)$  and assigning to each  $x_i$  the control  $v_i$  above. In our construction, we will assume  $\|\cdot\|$  is the infinity norm, i.e.,  $\|x\| = \max_{j=1,\dots,n} |x_j|$ .

**Step 1: Choice of radii.** Fix  $\alpha \in (0, \lambda)$  and  $\tau > \frac{\ln K}{\lambda - \alpha}$ . Let  $L := L_{\mathcal{R}^\tau(S)}$  and set

$$\rho := \frac{1 - Ke^{-(\lambda - \alpha)\tau}}{1 + e^{(L + \alpha)\tau}} > 0, \quad r_i := \rho\|x_i - x^*\|.$$

By construction  $r_i > 0$ . From (14), for  $t = \tau$  we have

$$Ke^{-(\lambda - \alpha)\tau}\|x_i - x^*\| + r_i e^{(L + \alpha)\tau} = \|x_i - x^*\| - r_i,$$

which implies the verification condition

$$\begin{aligned} \min_{t \in (0, \tau]} e^{\alpha t} (\|\phi(t, x_i, v_i) - x^*\| + r_i e^{Lt}) &\leq \\ &\leq e^{\alpha\tau} (\|\phi(\tau, x_i, v_i) - x^*\| + r_i e^{L\tau}) \\ &\leq \|x_i - x^*\| - r_i. \end{aligned}$$

Hence each  $(x_i, r_i)$  satisfies (11a).

**Step 2: Covering number bound.** We cover the annular region  $B_R(x^*) \setminus B_\varepsilon(x^*)$  by  $n$  concentric annuli  $A_1, \dots, A_n$  of thickness  $R_i = 3^{i-1}\varepsilon$ . Since the total radial width is  $R - \varepsilon$ ,

$$\begin{aligned} R - \varepsilon &\leq \sum_{i=1}^n 3^{i-1}\varepsilon = \varepsilon \frac{3^n - 1}{2} \Rightarrow \\ n &= \left\lceil \log_3 \left( \frac{2R}{\varepsilon} - 1 \right) \right\rceil = O\left(\log \frac{R}{\varepsilon}\right). \end{aligned}$$

Each  $A_i$  is initially partitioned into  $3^d - 1$  hypercubes of side  $R_i$ . For any  $x \in A_i$ ,  $\|x - x^*\| \geq 3^{i-1}\varepsilon$  and  $r_i \leq \rho\|x - x^*\|$ . We refine each hypercube by successive splits into  $3^d$  subcubes until the side length is at most  $\rho\|x - x^*\|$ . This requires

$$3^m \geq \frac{1}{\rho} \Rightarrow m = \lceil \log_3(1/\rho) \rceil$$

splits, producing at most

$$N_{\text{annulus}} = (3^d - 1) 3^{dm} = O\left(\left(\frac{3}{\rho}\right)^d\right)$$

points per annulus.

**Step 3: Total number of points.** Multiplying by the number of annuli,

$$N = n \cdot N_{\text{annulus}} = O\left(\left(\frac{3}{\rho}\right)^d \log \frac{R}{\varepsilon}\right).$$

Finally, since  $c = \varepsilon(1 + L\tau e^{L\tau})$ , this yields the claimed sample complexity bound. □

### C. Incremental Learning of NCPs

The performance–complexity trade-off discussed in Remark 3, together with the existence and sample complexity result of Theorem 3, suggests a practical methodology for progressively improving performance. By actively sampling trajectories more finely and refining the covering set  $\mathcal{K}$ , one

can construct NCPs that certify larger rates  $\alpha$  (by reducing the effective radius  $r$ ), thereby reducing the gap between realized performance and the best attainable rate  $\lambda$ . In other words, performance can be systematically enhanced by enriching the assignment set with additional verified points, which readily enables incremental learning.

Beyond improving rates, another key feature of NCPs is their ability to incrementally *expand the certified region*. The next result formalizes this incremental learning property: previously verified assignments can be reused together with new ones to grow the domain over which stability is guaranteed.

**Theorem 4 (Incremental Learning of  $\mathcal{K}$ ).** *Consider an equilibrium point  $x^* \in \mathbb{R}^n$  of (1) and a set  $S \subseteq \mathbb{R}^n$  satisfying  $x^* \in \text{int}(S)$ . Let  $\pi_{\mathcal{K}}$  be a nonparametric policy with assignment set  $\mathcal{K} = \{(x_i, r_i, v_i)\}_{i=1}^N$  and default control  $v_0 \in \mathcal{A}$  satisfying properties (i)–(iii) of Theorem 2 with parameters  $\alpha, \delta, \tau, L, \varepsilon$ . Take  $x_j \in \mathbb{R}^n \setminus S$ ,  $r_j > 0$ , and  $v_j \in \mathcal{U}^{[0, \tau_j]}$  s.t.  $B_{r_j}(x_j) \cup S = \emptyset$ . Define the enlarged set  $S' := S \cup B_{r_j}(x_j)$ , and let  $L_j = L_{\mathcal{R}^{\tau_j}(B_{r_j}(x_j))}$ , and  $L' = \max\{L_j, L\}$ .*

*Whenever the following conditions are satisfied:*

- (1) **Feasibility of  $(x_j, r_j, v_j)$ :** *The 3-tuple  $(x_j, r_j, v_j)$  with  $v_j : [0, \tau_j] \rightarrow U$  satisfy*

$$\text{sd}(\phi(\tau_j, x_j, v_j), S) + r_j e^{L_j \tau_j} \leq 0.$$

- (2) **Either of the following holds:**

- (a) **Direct verification at  $(x_j, r_j, v_j)$ :** *The tuple satisfies decrease condition:*

$$e^{\alpha \tau_j} (\|\phi(\tau_j, x_j, v_j) - x^*\| + r_j e^{L_j \tau_j}) \leq \|x_j - x^*\| - \epsilon_j \quad (15)$$

*Set  $\alpha' = \alpha$ ,  $\tau' = \max\{\tau, \tau_j\}$ ,  $\delta' = \delta$ .*

- (b) **Bootstrapping:** *There is  $\hat{\mathcal{K}} \subseteq \mathcal{K}$ , such that*

- (i)  $B_{r_j e^{L_j \tau_j}}(\phi(\tau_j, x_j, v_j)) \subseteq \text{Supp}(\hat{\mathcal{K}})$

- (ii) *There is  $\alpha' < \alpha$  such that*

$$\max_{(x_i, r_i, v_i) \in \hat{\mathcal{K}}} \frac{e^{-(\alpha - \alpha')\tau_i}}{e^{-\alpha \tau_j}} \frac{\|x_i - x^*\| + r_i}{\|x_j - x^*\| - r_j} \leq 1.$$

*Set  $\tau' = \tau + \tau_j$ ,  $\delta' = \delta$ .*

*Then the augmented assignment set  $\mathcal{K}' := \mathcal{K} \cup \{(x_j, r_j, v_j)\}$  and the default control  $v_0$  induce a policy  $\pi_{\mathcal{K}'}$  that practically exponentially stabilizes  $x^*$  over  $S' = S \cup B_{r_j}(x_j)$  with*

$$\lambda' = \alpha', \quad K' = e^{\alpha' \tau'} (1 + L' \tau e^{L' \tau'}), \quad c' = \varepsilon(1 + L' \tau e^{L' \tau}).$$

*Proof.* We first note that since  $\mathcal{K}$  and  $v_0$  satisfy Theorem 1 (i)–(iii) on  $S$ , and  $S \cap B_{r_j}(x_j) = \emptyset$ , any initial state  $x \in S$  under  $\pi_{\mathcal{K}'}$  will only trigger assignments from  $\mathcal{K}$  or the default control  $v_0$ . By Theorem 2, this ensures that for every  $x \in S$  there exists a sequence of times  $\{t_n\}$  satisfying the conditions of Lemma 1, along which (6) is satisfied for  $V(x) = \|x - x^*\|$ . Hence, whenever  $x \in S \subset S'$ , the trajectory  $\phi(t, x, u_{\mathcal{K}', x})$  is practically exponentially stable with parameters  $\lambda = \alpha$ ,  $K = e^{\alpha \tau} (1 + L \tau e^{L \tau})$ , and  $c = \varepsilon(1 + L \tau e^{L \tau})$ . Moreover, since  $\alpha \leq \alpha'$ ,  $\tau' \geq \tau$ , and  $L' \geq L$ , the same trajectory also satisfies the practical exponential stability bound with the updated constants  $\lambda'$ ,  $K'$ , and  $c'$  as stated in the theorem.

We next consider the case  $x \in S' \setminus S$ . Since  $S \cap B_{r_j}(x_j) = \emptyset$ , it follows that  $\iota_{\mathcal{K}'}(x) = j$ , so the first control applied by  $\pi_{\mathcal{K}'}$

is  $v_j$ . We will use the sequence (9) induced by  $\pi_{\mathcal{K}'}$  to build a sequence of times  $\{t_n\}_{n \in \mathbb{N}}$  satisfying the properties in (6) of Lemma 1 for the control signal  $u_{\mathcal{K}', x}$ . The result then follows from Theorem 1. We will choose  $\delta'$  and  $\alpha'$  in (6) later on. Recall that  $\pi_{\mathcal{K}}$  satisfies (6) for some  $\alpha$  and  $\delta = \varepsilon(1 + L \tau e^{L \tau})$ . To simplify notation, we use  $u$  to refer to  $u_{\mathcal{K}', x}$ .

We choose  $x_0 = x$  and  $t_0 = 0$ . We will select  $t_1$  differently, depending on which clause of condition (2) in the theorem's hypothesis hold.

**Case (2a).** We choose  $t_1 = \tau_j$ , and accordingly  $x_1 = \phi(\tau_j, x, v_j) = \phi(t_1, x, u)$ . By condition (1) of the theorem, we have

$$\text{sd}(\phi(\tau_j, x, v_j), S) \leq \text{sd}(\phi(\tau_j, x_j, v_j), S) + r_j e^{L_j \tau_j} \leq 0,$$

which implies  $\phi(\tau_j, x, v_j) \in S$ . This ensures  $\phi(t_1, x, u) = \phi(\tau_j, x, v_j) \in S \subset S'$ , i.e., (6b) for the set  $S'$ . It further follows from (15) by a similar argument to Theorem 2, that

$$\|\phi(t_1, x, v_j) - x^*\| \leq e^{-\alpha t_1} \|x - x^*\|,$$

so (6c) holds for  $\alpha$  and any value of  $\delta$ .

**Case (2b).** As in case (2a) after choosing  $v_j$ ,  $\phi(\tau_j, x, v_j) \in S$ . In fact, by (2b.i),  $\phi(\tau_j, x, v_j) \in \text{Supp}(\hat{\mathcal{K}}) \subseteq S$ . Let  $(x_i, r_i, v_i) \in \hat{\mathcal{K}}$  s.t.  $i = \iota_{\mathcal{K}}(\phi(\tau_j, x, v_j))$ . We choose

$$t_1 := \tau_j + \tau_i,$$

$$x_1 := \phi(\tau_j + \tau_i, x, v_j v_i) = \phi(\tau_i, \phi(\tau_j, x, v_j), v_i).$$

Since (11) holds for  $(x_i, r_i, v_i) \in \mathcal{K}$ ,  $x_1 \in S$ , and therefore we have  $\phi(t_1, x, u) = x_1 \in S \subset S'$ ; hence (6b) holds for  $t_1$  and  $S'$ .

It remains to ensure the decrease in (6c) at  $t_1$  with some  $\alpha' < \alpha$  and  $\delta' = \delta$ . We thus consider

$$\begin{aligned} e^{\alpha' t_1} \|\phi(t_1, x, u) - x^*\| &= \\ &= e^{\alpha' (\tau_j + \tau_i)} \|\phi(\tau_j + \tau_i, x, v_j v_i) - x^*\| \\ &= e^{\alpha' (\tau_j + \tau_i)} \|\phi(\tau_i, \phi(\tau_j, x, v_j), v_i) - x^*\| \\ &\leq e^{\alpha' (\tau_j + \tau_i)} e^{-\alpha \tau_i} \|\phi(\tau_j, x, v_j) - x^*\| \\ &= \frac{e^{-(\alpha - \alpha') \tau_i}}{e^{-\alpha' \tau_j}} \left( \frac{\|\phi(\tau_j, x, v_j) - x^*\|}{\|x - x^*\|} \right) \|x - x^*\| \\ &\leq \frac{e^{-(\alpha - \alpha') \tau_i}}{e^{-\alpha' \tau_j}} \left( \frac{\|x_i - x^*\| + r_i}{\|x_j - x^*\| - r_j} \right) \|x - x^*\| \\ &\leq \|x - x^*\|, \end{aligned}$$

where step one follows from definition of  $t_1$  and  $u$ , two from the group property of  $\phi$ , three from (11a) on  $(x_i, r_i, v_i)$ , and the final step from the bootstrapping condition (2b.ii).

In both cases, after  $t_1$  the subsequent points are generated by  $\mathcal{K}$ , so the sequence  $\{t_n\}$  continues to satisfy (6b)–(6c) by Theorem 2. Thus  $V(x) = \|x - x^*\|$  is an R-CLF on  $S'$ , and by Theorem 1,  $\pi_{\mathcal{K}'}$  renders  $x^*$  practically exponentially stable on  $S'$  with the claimed constants.  $\square$

## V. NUMERICAL EXPERIMENTS

Using the sufficiency of NCPs derived in Theorem 2 along with the grid construction of Theorem 3, we next introduce an algorithmic methodology to design NCPs to stabilize a given region. The algorithmic flow is as follows:

- i. Given a region  $S \subset \mathbb{R}^n$ , select a desired convergence rate  $\alpha > 0$ , select a  $\tau_{\max}$  to upper-bound  $\tau_i$ , and determine the one-sided Lipschitz constant  $L$  for the underlying dynamics across  $\mathcal{R}^{\tau_{\max}}(S)^1$ .
- ii. Create a grid of points and radii  $G = \{(x_i, r_i)\}$  covering the region  $S$  with increasingly large radii per annulus according to Theorem 3.
- iii. For each  $x_i$ , derive controls  $v_i$  for  $\tau_{\max}$  time (for example, using sampling methods akin to Model-Predictive Path Integrals (MPPI) [30], [31]).
- iv. For each  $(x_i, r_i, v_i)$ , if condition (11a) is satisfied for some  $\tau_i \leq \tau_{\max}$ , store the largest  $\tau_i$  that achieves equality.
- v. Otherwise, split the ball  $B_{r_i}(g_i)$  into  $3^d$  smaller balls.
- vi. Repeat until a control is found for  $\alpha$ -exponential decrease with slack of at least each ball's radius, or until a pre-defined maximum number of splits is achieved.
- vii. For balls which still fail the conditions, apply condition 2b of Theorem 4 to leverage previously derived controls.
- viii. Trim down the verified region only to those trajectories which satisfy (11b), save verified  $\alpha_i \geq \alpha$  for each cell.

With this algorithmic method, we present a number of case studies in different classic control stabilization problems, which will demonstrate useful features of NCPs.

#### A. Unicycle

Consider the unicycle model moving in the plane, parametrized by its  $x$  position,  $y$  position, and angle of vehicle  $\theta$  with respect to the  $x$ -axis, with two control inputs - velocity  $v$  and angular velocity  $\omega$ . The dynamics are given by

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} v \cos(\theta) \\ v \sin(\theta) \\ \omega \end{bmatrix}.$$

We bound  $u \in [0, 2]$  and  $\omega \in [-1, 1]$ , and run the method to derive NCPs for two different norms, being  $V_1 = \max\{|x|, |y|, |\theta|\}$  and  $V_2 = \sqrt{x^2 + y^2 + 0.01\theta^2}$ . For either choice of norm (simpler or tied to the classic reward function), the method quickly stabilized the entire region of  $(x, y) \in [-20, 20]^2, \theta \in (-\pi, \pi]$ , see Figure 1.

To demonstrate the incremental growth capabilities of NCP, we do two stages of learning. After learning a control for the previous region, i.e.,  $(x, y, \theta) \in [-20, 20]^2 \times (-\pi, \pi]$ , we expand the state space to include values in the region  $(x, y, \theta) \in [-20, 20] \times [20, 25] \times (-\pi, \pi]$ . Trajectories fragments starting in the formerly verified region retain the same behavior, while the new behavior (for initial values in the new region) are depicted in Figure 2.

#### B. Inverted Pendulum

We next analyze the utility of the NCPs on the inverted pendulum. The system consists of a mass  $m$  attached at the end of a rigid pendulum of length  $l$ , pivoting freely about a fixed point. The dynamics are governed by the torque around the pivot due to both gravity and an external control

<sup>1</sup>Achieved through estimating the reachable tube by simulating samples along the boundary and adding precision-correction terms, and then sampling points in that region for OSL while adding precision-correction terms again.

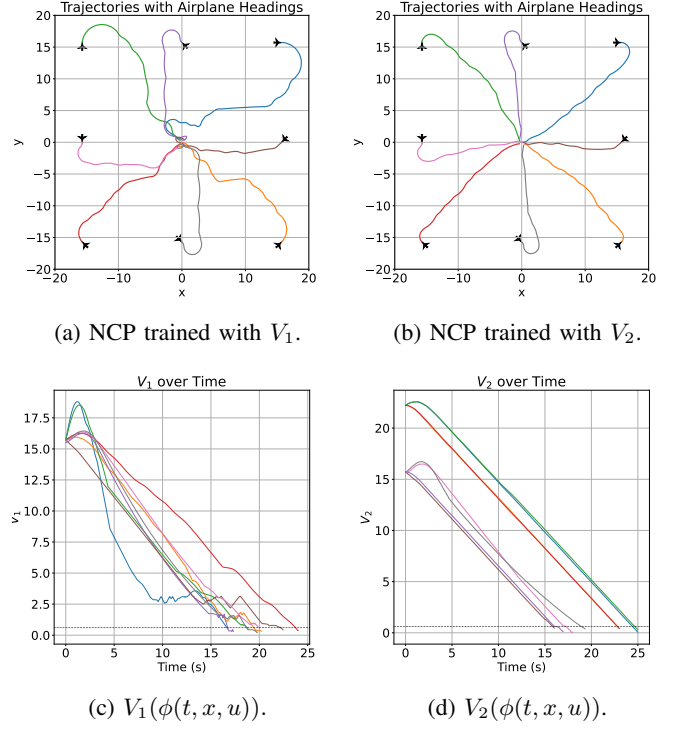


Fig. 1: **Trajectories of Unicycle NCP.** Phase plots of  $(x, y)$  for eight evenly distributed points. The black icons depict the initial facing of the unicycle. Plot (a) contains trajectories from NCP trained to minimize  $V_1$ , which results in sharp turns, while (b) is trained to minimize  $V_2$ , which results in softer turns and smoother overall behavior. Plots (c) and (d) show the development of  $V_1$  and  $V_2$  over time respectively. Both converge exponentially to the equilibrium, with at least  $\alpha = 0.01$ . We have  $\tau_{\max} = 5, \varepsilon = 0.01, L = 1$  and  $c = \varepsilon(1 + L\tau_{\max}e^{L\tau_{\max}}) \simeq 0.613$ , represented by the dotted line.

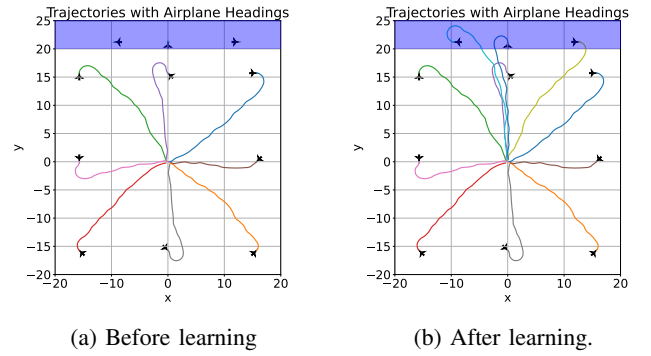


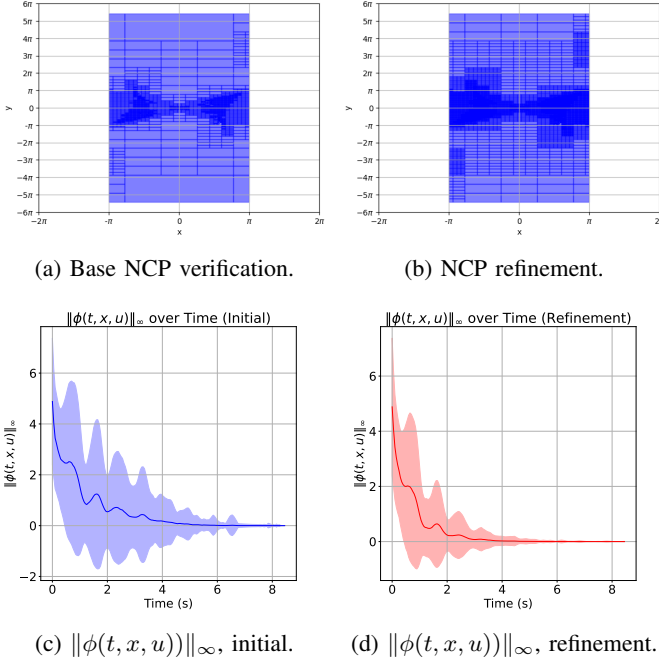
Fig. 2: **Incremental Learning of Unicycle Policy.** Extending the state space from the previously learned region in the  $y$ -direction. Subfigure (a) contains the phase plot before learning, while subfigure (b) contains the phase plot after. The new region is learned without forgetting, such that parts of the trajectory in the old region use previously designed controls.

input. Denoting by  $\theta$  the angle of the pendulum measured from the vertical (with  $\theta = 0$  corresponding to the inverted equilibrium), the equation of motion is given by

$$ml^2\ddot{\theta}(t) = mgl \sin(\theta(t)) + u(t),$$

where  $g$  is the gravitational acceleration, and  $u(t)$  is the external control torque applied at the pivot.

Figure 3 demonstrates the refinement capabilities of NCPs, such that by adding data (simulated by splitting all balls once), the rate of convergence achieved is significantly increased.



**Fig. 3: Additional Data Refinement Facilitates Improved NCP Performance.** Plot (a) contains the balls used to verify the region  $(\theta, \dot{\theta}) \in (-\pi, \pi) \times [-5\pi, 5\pi]$  for the inverted pendulum. Plot (b) is a refinement of plot (a), wherein all balls were split once more and re-verified. The minimum verified rate of convergence for trajectories  $\alpha$  goes from 0.003 to 0.0145, and the average verified  $\alpha$  goes from 1.815 to 3.149. Plot (c) demonstrates the average norm over time of 400 sample trajectories under each schema. We have  $\tau_{\max} = 1.5$ ,  $\varepsilon = 0.01$ ,  $L = 1$ , and  $c = \varepsilon(1 + L\tau_{\max}e^{L\tau_{\max}}) \simeq 0.072$ .

## VI. CONCLUSIONS

In this work we proposed a method for data-driven (practical) stabilization of nonlinear systems using nonparametric Chain Policies. The approach leverages a normalized nearest-neighbor rule to assign, at each state, a finite-duration control signal, after which the process repeats. The method is grounded in the notion of Recurrent Lyapunov Functions (RLFs) as well as their control extension Control-RLFs, which enable certification of stability using standard norm function.

Our analysis establishes that:

- 1) NPC Policies achieve practical exponential convergence to a  $c$ -neighborhood with sample complexity scaling as

$O((3/\rho)^d \log(R/c))$  in terms of the region radius  $R$  and precision  $c$ .

- 2) The framework supports incremental growth: new assignments can be added to expand the verified region while preserving previously established guarantees.
- 3) Controller refinement is monotone: more data only improves convergence rates and enlarges certified region.

These results position Chain Policies as a scalable, data-driven approach to certified stabilization, offering rigorous guarantees together with the ability to expand incrementally as new data becomes available.

## REFERENCES

- [1] K. Zhou, J. Doyle, and K. Glover, “Robust and optimal control,” *Control Engineering Practice*, vol. 4, no. 8, pp. 1189–1190, 1996.
- [2] E. Sontag, *Mathematical Control Theory: Deterministic Finite Dimensional Systems*, ser. Texts in Applied Mathematics. Springer New York, 2013.
- [3] C. De Persis and P. Tesi, “Formulas for data-driven control: Stabilization, optimality, and robustness,” *IEEE Transactions on Automatic Control*, vol. 65, no. 3, pp. 909–924, 2019.
- [4] J. Coulson, J. Lygeros, and F. Dörfler, “Data-enabled predictive control: In the shallows of the deep,” in *2019 18th European control conference (ECC)*. IEEE, 2019, pp. 307–312.
- [5] J. Berberich, J. Köhler, M. A. Müller, and F. Allgöwer, “Data-driven model predictive control with stability and robustness guarantees,” *IEEE transactions on automatic control*, vol. 66, no. 4, pp. 1702–1717, 2020.
- [6] J. Berberich, A. Koch, C. W. Scherer, and F. Allgöwer, “Robust data-driven state-feedback design,” in *2020 American Control Conference (ACC)*. IEEE, 2020, pp. 1532–1538.
- [7] H. J. Van Waarde, M. K. Camlibel, and M. Mesbahi, “From noisy data to feedback controllers: Nonconservative design via a matrix s-lemma,” *IEEE Transactions on Automatic Control*, vol. 67, no. 1, pp. 162–175, 2020.
- [8] C. De Persis and P. Tesi, “Low-complexity learning of linear quadratic regulators from noisy data,” *Automatica*, vol. 128, p. 109548, 2021.
- [9] X. Chen and E. Hazan, “Black-box control for linear dynamical systems,” in *Conference on Learning Theory*. PMLR, 2021, pp. 1114–1143.
- [10] S. W. Werner and B. Peherstorfer, “On the sample complexity of stabilizing linear dynamical systems from data,” *Foundations of Computational Mathematics*, vol. 24, no. 3, pp. 955–987, 2024.
- [11] T. Dai and M. Szaier, “A semi-algebraic optimization approach to data-driven control of continuous-time nonlinear systems,” *IEEE Control Systems Letters*, vol. 5, no. 2, pp. 487–492, 2020.
- [12] M. Guo, C. De Persis, and P. Tesi, “Data-driven stabilization of nonlinear polynomial systems with noisy data,” *IEEE Transactions on Automatic Control*, vol. 67, no. 8, pp. 4210–4217, 2021.
- [13] H. El-Kebir and M. Ornik, “Sum-of-squares data-driven robustly stabilizing and contracting controller synthesis for polynomial nonlinear systems,” *arXiv preprint arXiv:2503.07438*, 2025.
- [14] R. Strässer, J. Berberich, and F. Allgöwer, “Data-driven control of nonlinear systems: Beyond polynomial dynamics,” in *2021 60th IEEE Conference on Decision and Control (CDC)*. IEEE, 2021, pp. 4344–4351.
- [15] N. Monshizadeh, C. De Persis, and P. Tesi, “A versatile framework for data-driven control of nonlinear systems,” *IEEE Transactions on Automatic Control*, 2025.
- [16] A. Oliveira, J. Zheng, and M. Szaier, “Convex data-driven contraction with riemannian metrics,” *IEEE Control Systems Letters*, 2025.
- [17] Z. Hu, C. De Persis, and P. Tesi, “Enforcing contraction via data,” *IEEE Transactions on Automatic Control*, 2025.
- [18] H. Tsukamoto, S.-J. Chung, and J.-J. E. Slotine, “Contraction theory for nonlinear stability analysis and learning-based control: A tutorial overview,” *Annual Reviews in Control*, vol. 52, pp. 135–169, 2021.
- [19] B. Huang, X. Ma, and U. Vaidya, “Data-driven nonlinear stabilization using koopman operator,” in *The Koopman Operator in Systems and Control: Concepts, Methodologies, and Applications*. Springer, 2020, pp. 313–334.
- [20] Y. Chen and U. Vaidya, “Sample complexity for nonlinear stochastic dynamics,” in *2019 American Control Conference (ACC)*. IEEE, 2019, pp. 3526–3531.

- [21] T.-W. Hsu and H. Tsukamoto, "Statistical guarantees in data-driven nonlinear control: Conformal robustness for stability and safety," *IEEE Control Systems Letters*, 2025.
- [22] R. Siegelmann, Y. Shen, F. Paganini, and E. Mallada, "A recurrence-based direct method for stability analysis and gpu-based verification of non-monotonic lyapunov functions," in *62nd IEEE Conference on Decision and Control (CDC)*. IEEE, 12 2023, pp. 6665–6672.
- [23] —, "Stability analysis and data-driven verification via recurrent lyapunov functions," *IEEE Transactions on Automatic Control*, 07 2025.
- [24] F. Colonius, "Minimal bit rates and entropy for exponential stabilization," *SIAM Journal on Control and Optimization*, vol. 50, no. 5, pp. 2988–3010, 2012.
- [25] B. Hamzi and A. J. Krener, "Practical stabilization of systems with a fold control bifurcation," in *New Trends in Nonlinear Dynamics and Control and Their Applications*. Springer, 2004, pp. 37–48.
- [26] F. Colonius and B. Hamzi, "Entropy for practical stabilization," *SIAM Journal on Control and Optimization*, vol. 59, no. 3, pp. 2195–2222, 2021.
- [27] H. K. Khalil, "Nonlinear systems; 3rd ed." 2002.
- [28] H. Sibai and E. Mallada, "Recurrence of nonlinear control systems: Entropy, bit rates, and finite alphabets," pp. 1–16, 02 2025.
- [29] F. Colonius and C. Kawan, "Invariance entropy for control systems," *SIAM J. Control Optim.*, vol. 48, no. 3, pp. 1701–1721, 2009.
- [30] G. Williams, P. Drews, B. Goldfain, J. M. Rehg, and E. A. Theodorou, "Aggressive driving with model predictive path integral control," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*, 2016, pp. 1433–1440.
- [31] —, "Information-theoretic model predictive control: Theory and applications to autonomous driving," *IEEE Transactions on Robotics*, vol. 34, no. 6, pp. 1603–1622, 2018.