MITSUBISHI ELECTRIC RESEARCH LABORATORIES
http://www.merl.com

Data-Driven Control Policies for Partially Known Systems via Kernelized Lipschitz Learning

Chakrabarty, A.; Jha, D.; Wang, Y.

TR2019-047 June 29, 2019

Abstract

Generating initial stabilizing control policies that satisfy operational constraints in the absence of full model information remains an open but critical challenge. In this paper, we propose a systematic framework for constructing constraint enforcing initializing control policies for a class of nonlinear systems based on archival data. Specifically, we study systems for which we have linear components that are modeled and nonlinear components that are unmodeled, but satisfy a local Lipschitz condition. We employ kernel density estimation (KDE) to learn a local Lipschitz constant from data (with high probability), and compute a constraint enforcing control policy via matrix multipliers that utilizes the learned Lipschitz constant. We demonstrate the potential of our proposed methodology on a nonlinear system with an unmodeled local Lipschitz nonlinearity.

American Control Conference (ACC)

Data-Driven Control Policies for Partially Known Systems via Kernelized Lipschitz Learning

Ankush Chakrabarty*, Devesh K. Jha, Yebin Wang

***Abstract*—Generating initial stabilizing control policies that satisfy operational constraints in the absence of full model information remains an open but critical challenge. In this paper, we propose a systematic framework for constructing constraint enforcing initializing control policies for a class of nonlinear systems based on archival data. Specifically, we study systems for which we have linear components that are modeled and nonlinear components that are unmodeled, but satisfy a local Lipschitz condition. We employ kernel density estimation (KDE) to learn a local Lipschitz constant from data (with high probability), and compute a constraint enforcing control policy via matrix multipliers that utilizes the learned Lipschitz constant. We demonstrate the potential of our proposed methodology on a nonlinear system with an unmodeled local Lipschitz nonlinearity.**

***Index Terms*—Machine learning; Lipschitz constant estimation; numerical computing; constrained control; stability guarantees; linear matrix inequalities.**

I. Introduction

Nonlinearities are ubiquitous in real-world applications. In many applications, these nonlinearities are unmodeled, or part of high-fidelity modeling software, rendering them too complicated for controller design. Recent efforts tackle this issue by learning from operational (on-line) or archival data (off-line). Since the exact structure of the nonlinearity may be unknown or not amenable for analysis, researchers have proposed 'indirect' data-driven controllers that employ non-parametric learning methods such as Gaussian processes to construct models from operational data [1] to improve control policies on-line [2], [3]. These approaches generally require an initial control policy that is stabilizing and robust to unmodeled dynamics. Depending on the sensitivity of the dynamics to the unmodeled components, designing such an initial control policy is difficult.

Conversely, 'direct' methods, such as those proposed in [4]–[6], directly compute policies using a combination of archival/legacy and operational input-output data without constructing an intermediate model. For example, in [7], a human expert was introduced into the control loop to conduct initial experiments to ensure safety while generating archival data. Although the aforementioned references provide excellent methods for utilizing archival data, general design of initializing control policies for multivariate systems (especially nonlinear systems) in a computationally tractable manner remains an open challenge.

*All authors are affiliated with Mitsubishi Electric Research Laboratories, Cambridge, MA, USA. Corresponding author: A. Chakrabarty. Phone: +1 (617) 758-6175. Email: chakrabarty@merl.com.

Our key insight is that information regarding the structure of an unmodeled nonlinearity may be encapsulated using only a few parameters. Therefore, it may not be necessary to model the unknown component itself in order to compute a *safe* (stabilizing and constraint satisfying) control policy. Instead, one can exploit structural information. For instance, the class of Lipschitz nonlinear functions (which constitute a large share of nonlinearities observed in applications) can be described using only one parameter: the Lipschitz constant. Recent work has investigated the utility of Lipschitz properties in constructing controllers when an oracle is available [8] or in designing models for prediction [9] with on-line data used for controller refinement [10], assuming the model error is bounded, and the bound is known. In this paper, we construct control policies that respect constraints and certify stability (with high probability) for applications where only off-line data is available, and no oracle is present. We do so through the systematic use of multiplier matrices that enable the representation of nonlinear dynamics through quadratic constraints [11], [12] without requiring knowledge of the underlying nonlinearity. The construction of these multiplier matrices for Lipschitz systems require the Lipschitz constant, which is not always available, and therefore, must be estimated: we refer to this as *Lipschitz learning*. Historically, methods that estimate the Lipschitz constant [13]–[15] do not provide certificates on the quality of the estimate. Herein, we provide conditions that, if satisfied, enable us to estimate the Lipschitz constant of an unknown locally Lipschitz nonlinearity with high probability. To this end, we employ kernel density estimation (KDE): a non-parametric data-driven method that employs kernels to approximate smooth probability density functions to arbitrarily high accuracy. We refer to our proposed KDE-based Lipschitz constant estimation algorithm as *kernelized Lipschitz learning*.

The **contributions** of this paper include: (i) the formulation of an algorithm to construct stabilizing and constraint satisfying policies for nonlinear systems without knowing the exact form of the nonlinearity; (ii) a kernelized Lipschitz learning mechanism to estimate the Lipschitz constant with high probability; and, (iii) a multiplier-matrix based controller design based on Lipschitz learning from legacy data that forces exponential stability on the closed-loop dynamics (with the same probability as the kernelized Lipschitz learner).

II. Notation

We denote by $\mathbb{R}$ the set of real numbers, $\mathbb{R}_+$ as the set of positive reals, and $\mathbb{N}$ as the set of natural numbers.

The Hausdorff distance between two subsets A and B of a metric space $\mathbb{R}^n$ equipped with the metric ρ is given by $\rho_H(A,B) = \max\{\sup_{x\in B}\rho(x,A), \sup_{x\in A}\rho(x,B)\}$. We define a ball $\mathcal{B}_\epsilon(x) := \{y : \rho(x,y) \le \epsilon\}$ and the sum $A \oplus \epsilon := \bigcup_{x\in A}\mathcal{B}_\epsilon(x)$. The complement of a set A is denoted by A^c. For every $v \in \mathbb{R}^n$, we denote $\|v\| = \sqrt{v^\top v}$, where $v^\top$ is the transpose of v. The sup-norm or ∞-norm is defined as $\|v\|_\infty \triangleq \sup_{t\in\mathbb{R}}\|v(t)\|$. We denote by $\lambda_{\min}(P)$ and $\lambda_{\max}(P)$ as the smallest and largest eigenvalue of a square, symmetric matrix P. The symbol $\succ$ $(\prec)$ indicates positive (negative) definiteness and $A \succ B$ implies $A - B \succ 0$ for A, B of appropriate dimensions. Similarly, $\succeq$ $(\preceq)$ implies positive (negative) semi-definiteness. The operator norm is denoted $\|P\|$ and is defined as the maximum singular value of P. For a symmetric matrix, we use the $\star$ notation to imply symmetric terms, that is, $\left[\begin{smallmatrix} a & b \\ b^\top & c\end{smallmatrix}\right] \equiv \left[\begin{smallmatrix} a & b \\ \star & c\end{smallmatrix}\right]$. The $\mathrm{diag}(\cdot)$ operator converts a set of block matrices into a matrix whose diagonal blocks are the block matrices. The symbol $\mathbf{Pr}$ denotes the probability measure.

III. Motivation

We consider nonlinear systems of the form

$$x^+ = Ax + Bu + G\phi(q), \tag{1a}$$

$$q = C_q x + D_q u, \tag{1b}$$

where $x, x^+ \in \mathbb{X} \subset \mathbb{R}^{n_x}$ denotes the state of the system and its update*, respectively. The state and control inputs are available for measurement. The control input is denoted $u \in \mathbb{U} \subset \mathbb{R}^{n_u}$ and the nonlinearity is denoted by $\phi \in \mathbb{R}^{n_\phi}$ with the argument $q \in \mathbb{D}_q \subset \mathbb{R}^{n_q}$ that can be represented as a linear combination of the state and control input. The system matrices A, B, G, C_q and D_q have appropriate dimensions. The admissible state and input spaces are denoted $\mathbb{X}$ and $\mathbb{U}$, respectively, and these spaces are compact, convex, and contain the origin in their interior. Since $\mathbb{X}$ and $\mathbb{U}$ are bounded, so is $\mathbb{D}_q$.

We make the following assumptions on our system and constraints.

Assumption 1. *The matrices A and B are known. The matrix G has full column rank and is one-hot. That is, only the non-zero element locations are known; its exact elements are unknown. The matrices C_q and D_q are completely unknown.*

We require the following definition to describe the class of nonlinearities considered in this paper.

Definition 1. *A function $f : \mathcal{D}_f \to \mathbb{R}^n$ is Lipschitz continuous in the domain $\mathcal{D}_f$ if*

$$\|f(d_1) - f(d_2)\| \le \mathfrak{L}_f\|d_1 - d_2\| \tag{2}$$

for some $\mathfrak{L}_f > 0$ and all $d_1, d_2 \in \mathcal{D}_f$. We define the scalar

$$\mathfrak{L}_f^* = \inf_{\mathbb{R}_+}\{\mathfrak{L}_f : \text{condition (2) holds}\} \tag{3}$$

as the Lipschitz constant of f.

*For continuous-time systems, $x^+ = \dot{x}$ and for discrete-time systems, $x^+ = x_{t+1}$. For both, we denote the initial time as t_0 and initial state x_0.

Assumption 2. *The nonlinearity ϕ is Lipschitz continuous in the domain $\mathbb{D}_q$ with unknown Lipschitz constant. That is,*

$$\|\phi(q_1) - \phi(q_2)\| \le \mathfrak{L}_\phi^*\|q_1 - q_2\| \tag{4}$$

for any $q_1, q_2 \in \mathbb{D}_q$, and $\mathfrak{L}_\phi^*$ is unknown. Also, $\phi(0) = 0$.

Assumptions 1 and 2 imply that the linear component of the true system (1) is known, but the rest is unknown. However, we do know the vector space through which the nonlinearity enters the dynamics of (1), since the non-zero locations of G are flagged.

Example 1. Consider the following nonlinear system:

$$\dot{x}_1 = -2x_1 + 3x_2 \tag{5}$$

$$\dot{x}_2 = 3x_1 + x_2 + u + 1.5\sin(x_1). \tag{6}$$

In accordance with our assumptions, our model knowledge is:

$$\dot{x} = \begin{bmatrix} -2 & 3 \\ 3 & 1 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \phi(q). \tag{7}$$

The nonlinearity $\phi(q) = 1.5\sin(q)$ is completely unknown, and so are $C_q = \begin{bmatrix} 1 & 0 \end{bmatrix}$ and $D_q = 0$; G is one-hot. ■

Remark 1. Note that Assumption 1 can be relaxed. Instead of exactly knowing a linear system described by A and B, one could possess knowledge of some other system matrices $\tilde{A}$, $\tilde{B}$ of appropriate dimensions. For instance, one could have no prior knowledge, in which case $\tilde{A} = 0$ and $\tilde{B} = 0$. When $\tilde{A} \neq A$, $\tilde{B} \neq B$, the resulting nonlinearity is

$$\tilde{\phi}(\tilde{q}) = G\phi(q) + (A - \tilde{A})x + (B - \tilde{B})u, \tag{8}$$

where $\tilde{q} = [q, x, u]^\top$. Using the triangle and Cauchy-Schwarz inequalities and denoting the Lipschitz constant of $\tilde{\phi}$ to be $\mathfrak{L}_{\tilde{\phi}}^*$, we get

$$\mathfrak{L}_{\tilde{\phi}}^* = \mathfrak{L}_\phi^*\|G\| + \|A - \tilde{A}\| + \|B - \tilde{B}\| > \mathfrak{L}_\phi^*.$$

The more knowledge we have of A and B, the smaller $\|A - \tilde{A}\| + \|B - \tilde{B}\|$ will be, and hence, the controller will be less conservative. To simplify the ensuing discussion, we continue with the stronger assumption that $\tilde{A} = A$ and $\tilde{B} = B$. ■

Remark 2. Similarly, one could relax the assumption on G and take it to be the identity matrix. Again, for simplicity, we assume G is one-hot, which implies that we know how the system is actuated, but we do not know the gains of the actuating channels. ■

The following assumption is made on the class of constraints considered.

Assumption 3. *The constraint sets $\mathbb{X}$ and $\mathbb{U}$ are described by the matrix inequalities*

$$\mathcal{X}' = \left\{ \begin{bmatrix} x \\ u \end{bmatrix} \in \mathbb{R}^{n_x+n_u} : c_i^\top x + d_i^\top u \le 1 \right\}, \tag{9}$$

for $i = 1, \dots, n_c$, where n_c is the total number of state and input constraints and $c_i \in \mathbb{R}^{n_x}$ and $d_i \in \mathbb{R}^{n_u}$.

Remark 3. The matrix inequality (9) defines a polytopic admissible state and input constraint set. Note that $c_i = 0$ implies that the ith constraint is an input constraint, and $d_i = 0$ implies that it is a state constraint. ■

Finally, we make the following assumption (akin to [5], [6]) on the availability of legacy or archival data generated by the system during prior experiments.

Assumption 4. *At design time, we have a sufficiently rich dataset $\mathcal{D}$ consisting of unique state-input pairs from within the admissible state and input space, and corresponding state update information. Specifically, $\mathcal{D} = \{x_j, u_j, x_j^+\}_{j=1}^N$, where $x_j, x_j^+ \in \mathbb{X}$ and $u_j \in \mathbb{U}$.*

By 'sufficiently rich', we mean that the dataset should have samples that are well-dispersed on $\mathbb{X}$ and $\mathbb{U}$.

Remark 4. We iterate that the states and inputs in $\mathcal{D}$ needs to be within $\mathbb{X}$ and $\mathbb{U}$, respectively, since the function ϕ may be locally Lipschitz and not globally Lipschitz. For example, if $\phi(x) = x^3$ on $\mathbb{X} = \{x : |x| \le 1\}$, then a local Lipschitz constant is $\mathfrak{L}_\phi^* = 3$, but using data collected from outside the region of interest $\mathbb{X}$ will result in larger estimates of $\mathfrak{L}_\phi^*$ since ϕ is not globally Lipschitz on $\mathbb{R}$. ■

Our **objective** in this paper is to leverage the dataset $\mathcal{D}$ to design a control policy $u = Kx$ such that the closed-loop system

$$x^+ = (A + BK)x + G\phi\left((C_q + D_q K)x\right) \tag{10}$$

is stabilized to the origin while satisfying state and input constraints in spite of unmodeled dynamics†. This is a direct data-driven controller because no model of ϕ is identified in the controller design step. For brevity, in the ensuing discussion, we will focus only on discrete-time systems but the results hold for continuous-time systems with slight modifications.

†Since the control policy is obtained using randomly generated data, certificates of closed-loop stability are provided with high probability rather than with certainty.

IV. Kernelized Lipschitz Learning

In this section, we provide a brief overview of kernel density estimation and provide a methodology for estimating Lipschitz constants from data with high-probability.

A. Computation of Lipschitz estimates

For each $\{x_j, u_j, x_j^+\} \in \mathcal{D}$, we estimate the nonlinear term using (1). That is,

$$\phi(q_j) = G^\dagger\left(x_j^+ - Ax_j - Bu_j\right),$$

where $G^\dagger$ exists by Assumption 1. If $n_\phi > 1$, the following procedure will be repeated for each component of ϕ. Therefore, this algorithm will yield n_ϕ Lipschitz constant estimates, one for each dimension of ϕ. To avoid notational complications, we proceed (w.l.o.g) with $n_\phi = 1$.

As a prerequisite to estimating the Lipschitz constant for the unknown non-linear function, we need to estimate the matrix C_q and D_q (see (1)). While estimating the exact elements of these matrices is non-trivial, we can estimate the non-zero elements in the matrices, which is enough to design control policies, because the exact elements of C_q and D_q will be subsumed within the Lipschitz constant. This problem is analogous to the problem of feature selection and sparse learning, known as automatic relevance determination (ARD) [16]. The basic idea in ARD is to give feature weights independent some parametric prior densities; these densities are subsequently refined by maximizing the likelihood of the data [16], [17]. Once the relevant q_j is identified, we can use the dataset $\{\phi(q_j), q_j\}_{j=1}^N$ to obtain N Lipschitz underestimates using the estimator

$$\hat{L}_j = \max\{\hat{\ell}_j\}, \tag{11}$$

where the kth element of $\hat{\ell}_j$ is given by

$$\hat{\ell}_{jk} = \frac{|\phi(q_j) - \phi(q_k)|}{\|q_j - q_k\|}, \tag{12}$$

with $k \in \{1, \dots, N\} \setminus j$. This estimator has been widely used in the literature to construct algorithms for determining Lipschitz constants, see for example: [13], [14], [18]. The sequence $\{\ell_j\}_{j=1}^N$ are empirical samples drawn from an underlying univariate distribution $\mathbb{L}_\phi$. Clearly, the true distribution $\mathbb{L}_\phi$ has finite support; indeed, its left-hand support is zero and its right-hand support is $\mathfrak{L}_\phi^*$. This leads us to the key idea of our approach: *determining the true Lipschitz constant is tantamount to estimating the support of the distribution* $\mathbb{L}_\phi$. Actually, we need an overestimate of $\mathfrak{L}_\phi^*$ (that is, our estimate should be larger than $\mathfrak{L}_\phi^*$) so that it $\hat{\mathfrak{L}}_\phi > \mathfrak{L}_\phi^*$ is a valid Lipschitz constant estimate. Common methods of tackling the support estimation is by assuming prior knowledge about the density shape of $\mathbb{L}_\phi$ or using Strongin overestimates of the Lipschitz constant. However, we avoid these overestimators because they are provably unreliable even for globally Lipschitz functions [15, Theorem 3.1]. Instead we try to fit the density directly from local estimates and the data in a non-parametric manner using KDE level sets.

B. Lipschitz Constant Estimation using KDE Level Sets

Let $X_1, X_2, \dots, X_n$ be an independent, identically distributed (i.i.d.) sample from an unknown density function $\mathbb{P}$. We define a support for the density $\mathbb{P}$ by the set $\Omega := \{X : \mathbb{P}(X) > 0\}$. An empirical density $\hat{\mathbb{P}}_N$ is expressed using the KDE

$$\hat{\mathbb{P}}_N(X) = \frac{1}{Nh^d}\sum_{j=1}^N \mathcal{K}\left(\frac{X - X_j}{h}\right), \tag{13}$$

where $\mathcal{K} : \mathbb{R}^d \to \mathbb{R}$ is a smooth function called the kernel function and $h > 0$ is the kernel bandwidth. The consistency of KDE has been proven in literature where uniform asymptotic convergence and convergence rate (under appropriate assumptions) have been provided [19], [20]. We make the following assumptions on the class of density functions and kernels.

Assumption 5. *The density $\mathbb{P}$ and kernel $\mathcal{K}$ satisfy assumptions (G), (K1), and (K2) in [20]. Furthermore, there exists a* *scalar $\lambda_0 \in (0,1)$ such that for any $\lambda \le \lambda_0$, there exists a $X_0 \in \Omega$ such that $\mathbb{P}(X) \le \lambda$ for all $X \ge X_0$.*

The first three assumptions of Assumption 5 are standard in the KDE literature [20]–[23] and are not repeated for brevity. Roughly speaking, they ensure that the true density belongs to the space of smooth functions, and therefore can be approximated sufficiently well by kernels that exhibit a sufficiently small covering number for this function space. The last part of Assumption 5 is not uncommon in Lipschitz estimation. For example, the authors in [13] have shown that estimating Lipschitz constants from data using (11) results in a reversed Weibull distribution that satisfies this assumption.

Since Ω is bounded, the following result is a direct consequence of Assumption 5.

Lemma 1. *If Assumption 5 holds, then for any $\delta > 0$ there exists an $X' \in \Omega$ such that $\int_{X'}^{\infty} \mathbb{P}(\mu)\, d\mu < \delta$.*

We need the following definition from [20] for the ensuing discussion.

Definition 2. *Recall N is the number of data points, and let $\lambda > 0$. A λ-density level set for $\mathbb{P}(X)$ is given by $D_\lambda = \{X : \mathbb{P}(X) = \lambda\}$. A set $S_{N,1-\beta}$ is asymptotically valid for D_λ if $\mathbf{Pr}(S_{N,1-\beta} \supset D_\lambda) = 1 - \beta + \mathcal{O}(r_N)$, where $\mathcal{O}(r_N) \to 0$ as $N \to \infty$ and $1-\beta \in (0,1)$ is a confidence-level.*

The intuition behind our Lipschitz constant estimation is as follows. Suppose we estimate an asymptotically valid set $S_{N,1-\beta}$ for D_λ where $\lambda \in (0, \epsilon_0)$ is made sufficiently small to ensure that the probability induced by the sub-level set $\{X : \mathbb{P}(X) < \lambda\}$ is small, which is true if Assumption 5 holds. Then, the maximum of $S_{N,1-\beta}$ will be a high-probability estimate of $\mathfrak{L}_\phi^*$.

We use the quantile-based bootstrap method proposed in [20, Section 4.1] to estimate an asymptotically valid set $S_{N,1-\beta}$. Let $\hat{D}_\lambda$ denote the λ-density level set estimate obtained using the KDE $\hat{\mathbb{P}}_N$ and let $W_\beta^\star$ denote the maximum of $\hat{D}_\lambda$. Clearly, $W_\beta^\star$ is the closest point to the right-hand support of $\mathbb{P}$ for a fixed λ. Let $W_N := \rho_H(D_\lambda, \hat{D}_\lambda)$ and $w_\beta^\star := \mathcal{F}_{W_N}^{-1}(1-\beta)$, where $\mathcal{F}_{W_N}$ denotes the cumulative distribution of the random variable W_N. If we could compute $\hat{D}_\lambda \oplus w_\beta^\star$, we would obtain an asymptotically valid set for D_λ. However, computing $w_\beta^\star$ is impossible without knowing D_λ, so we resort to a bootstrap method for estimating it. Concretely, N samples are drawn from $\{\hat{L}_k\}$ with replacement and the λ-density level set $\hat{D}_\lambda^k$ is computed for each of these samples, and the procedure is repeated N_b times. Let $w_k = \rho_H\left(\hat{D}_\lambda^k, \hat{D}_\lambda\right)$ denote the Hausdorff distance computed for each bootstrap iteration. These w_k generate a distribution from which the confidence level $w_\beta^\star = \mathcal{F}_{w_{N_b}}^{-1}(1-\beta)$ can be computed. Then,

$$S_{N,1-\beta} := \hat{D}_\lambda \oplus w_\beta^\star \tag{14}$$

is our proposed bootstrap confidence set. The following result establishes the asymptotic validity of $S_{N,1-\beta}$.

Theorem 1. *Assumption 5 holds. For any $0 < \lambda \ll 1$ such that $\lambda = \mathbb{P}(X')$ and $\forall\ X > X'$, $\mathbb{P}(X) < \lambda$, we get*

$$\mathbf{Pr}(\mathfrak{L}_\phi^* \notin S_{N,1-\beta}) < \beta + \delta + \mathcal{O}(r_N)$$

where the positive scalar $\delta = \int_{X'}^{\infty} \mathbb{P}(X) dX \ll 1$.

Proof. (Sketch) The desired probability can be written in terms of its complement $\mathbf{Pr}(\mathfrak{L}_\phi^* \notin S_{N,1-\beta}) = 1 - \mathbf{Pr}(\mathfrak{L}_\phi^* \in S_{N,1-\beta})$. Since, $S_{N,1-\beta}$ is an asymptotically valid set of D_λ, we can estimate a lower bound on the complement term using the joint probability $\mathbf{Pr}(\mathfrak{L}_\phi^* \in D_\lambda, D_\lambda \subset S_{N,1-\beta})$. That is, $\mathbf{Pr}(\mathfrak{L}_\phi^* \in S_{N,1-\beta}) \ge \mathbf{Pr}(\mathfrak{L}_\phi^* \in D_\lambda, D_\lambda \subset S_{N,1-\beta}) = \mathbf{Pr}(\mathfrak{L}_\phi^* \in D_\lambda)\mathbf{Pr}(D_\lambda \subset S_{N,1-\beta}) = (1-\delta)(1-\beta+\mathcal{O}(r_N))$, from Lemma 1 and [20, Theorem 4]. Simplifying these terms and neglecting the higher-order terms concludes the proof. ☐

In practice, our proposed Lipschitz constant estimate for a given $\beta > 0$ is given by

$$\hat{\mathfrak{L}}_\phi = \max\{S_{N,1-\beta}\} = |W_\beta^\star| + |w_\beta^\star|. \tag{15}$$

The pseudocode for our proposed Lipschitz learner is provided in Algorithm 1.

Algorithm 1 Kernelized Lipschitz Learning Algorithm

Require: Initial dataset, $\{x_k, \phi(C_q x_k)\}_{k=1}^N$
Require: Confidence parameter, $0 < \beta \ll 1$
Require: Number of bootstraps, N_b
$\{q_k, \phi(q_k)\} \leftarrow$ Estimate C_q via ARD
for k in $1, \dots, N$ **do**
  for j in $\{1, \dots, N\} \setminus k$ **do**
    $\hat{\ell}_{jk} \leftarrow$ compute using (12)
  $\hat{L}_k \leftarrow \max\{\ell_{jk}\}$
$\hat{\mathbb{L}}_\phi \leftarrow$ KDE with cross-validated $\mathcal{K}$ and h using $\hat{L}_k$
$\lambda \leftarrow$ by inspection of KDE
$\hat{D}_\lambda \leftarrow$ compute λ-density level set using $\hat{\mathbb{L}}_\phi$
$W_\beta^\star \leftarrow \max \hat{D}_\lambda$
$w \leftarrow \emptyset$
for k in $1, \dots, N_b$ **do**
  Resample N times with replacement from $\hat{\mathbb{L}}_\phi$
  $\hat{D}_\lambda^k \leftarrow \lambda$-density level set using bootstrapped samples
  $w_k \leftarrow \rho_H(\hat{D}_k, \hat{D})$
$w_\beta^\star \leftarrow$ bootstrap $(1-\beta)$ confidence interval of $\{w_k\}$
$\hat{\mathfrak{L}}_\phi \leftarrow |W_\beta^\star| + |w_\beta^\star|$

V. Initial Control Policy Design

With the estimate $\hat{\mathfrak{L}}_\phi$ obtained using KDE methods, we are now ready to construct the initial control policy K. Note that with any control policy K, the constraint set described in (9) is equivalent to the set

$$\mathcal{X} = \left\{x \in \mathbb{R}^{n_x} : (c_i + d_i K)^\top x \le 1\right\}, \tag{16}$$

for $i = 1, \dots, r$. Before we state the main design theorem, we require the following result from [24, pp. 69].

Lemma 2. *The ellipsoid*

$$\mathcal{E}_P = \{x \in \mathbb{R}^{n_x} : x^\top P x \leq 1\} \tag{17}$$

is a subset of $\mathcal{X}$ *if and only if*

$$(c_i + d_i K) P^{-1} (c_i + d_i K)^\top \leq 1 \tag{18}$$

for $i = 1, \ldots, r$.

We also need the following stability definition.

Definition 3. *The equilibrium point* $x = 0$ *of the closed-loop system* (10) *is locally exponentially stable with a decay rate* α *and a domain of attraction* $\mathcal{E}_P$ *if there exist scalars* $C_0 > 0$ *and* $\alpha \in (0,1)$ *such that* $\|x_t\| \leq C_0 \alpha^{(t-t_0)} \|x_0\|$ *for any* $x_0 \in \mathcal{E}_P$.

Lemma 3. *Let* $V : [0, \infty) \times \mathcal{E}_P \to \mathbb{R}$ *be a continuously differentiable function such that*

$$\gamma_1 \|x\|^2 \leq V(t, x_t) \leq \gamma_2 \|x\|^2 \tag{19a}$$

$$V(t, x_{t+1}) - V(t, x_t) \leq -(1 - \alpha^2) V(t, x_t), \tag{19b}$$

for any $t \geq t_0$ *and* $x \in \mathcal{E}_P$ *along the trajectories of the system*

$$x^+ = \varphi(x), \tag{20}$$

where γ_1, γ_2, *and* α *are positive scalars, and* φ *is a nonlinear function. Then the equilibrium point* $x = 0$ *for the system* (20) *is locally exponentially stable with a decay rate* α *and a domain of attraction* $\mathcal{E}_P$.

The following design theorem provides a method to construct a stabilizing policy such that the origin is a locally exponentially stable equilibrium of the closed-loop system and constraint satisfaction is guaranteed within a prescribed ellipsoid of attraction $\mathcal{E}_P \subset \mathcal{X}$ without knowing the nonlinearity ϕ.

Theorem 2. *Assumptions 1–3 hold. If there exist matrices* $P = P^\top \succ 0 \in \mathbb{R}^{n_x \times n_x}$, $K \in \mathbb{R}^{n_u \times n_x}$, *and scalars* $\alpha \in (0,1)$, $\nu > 0$ *such that*

$$\begin{bmatrix} 1 & c_i + d_i K \\ \star & P \end{bmatrix} \succeq 0 \tag{21a}$$

$$\Psi + \Gamma^\top \mathcal{M} \Gamma \preceq 0 \tag{21b}$$

are satisfied for all $i = 1, 2, \ldots, r$ *with*

$$\Psi = \begin{bmatrix} (A+BK)^\top P (A+BK) - \alpha^2 P & (A+BK)^\top P G \\ G^\top P (A+BK) & G^\top P G \end{bmatrix},$$

$$\Gamma = \begin{bmatrix} C_q + D_q K & 0 \\ 0 & I \end{bmatrix},$$

$$\mathcal{M} = \begin{bmatrix} \nu^{-1} (\mathfrak{L}_\phi^*)^2 I & 0 \\ 0 & -\nu^{-1} I \end{bmatrix},$$

then the equilibrium point $x = 0$ *of the closed-loop system* (10) *is locally exponentially stable with a decay rate* α *and a domain of attraction* $\mathcal{E}_P$ *defined in* (17). *Furthermore, if the initial state* $x_0 \in \mathcal{E}_P$, *then the closed-loop states and inputs* x_t *and* u_t *satisfy the constraints* (9) *for all* $t \geq t_0$.

Note that we do not need to know ϕ to satisfy conditions (21). Instead, Theorem 2 provides conditions that leverage matrix multipliers similar to those described in [11].

We now provide LMI-based conditions for computing K, P and ν via convex programming.

Theorem 3. *Fix* $\alpha \in (0,1)$ *and* $\hat{\mathfrak{L}}_\phi$. *If there exist matrices* $S = S^\top \succ 0$, Y, *and a scalar* $\nu > 0$ *such that the LMI conditions*

$$\begin{bmatrix} 1 & c_i S + d_i Y \\ \star & S \end{bmatrix} \succeq 0 \tag{23a}$$

$$\begin{bmatrix} -\alpha^2 S & \star & \star & \star \\ 0 & -\nu I & \star & \star \\ AS + BY & \nu G & -S & \star \\ \hat{\mathfrak{L}}_\phi (C_q S + D_q Y) & 0 & 0 & -\nu I \end{bmatrix} \preceq 0 \tag{23b}$$

are satisfied, then the matrices $K = YS^{-1}$, $P = S^{-1}$ *and the scalar* ν *satisfy the conditions* (21) *with the same* α *and* $\hat{\mathfrak{L}}_\phi$.

A benefit of overestimating $\hat{\mathfrak{L}}_\phi$ is that safety is ensured. This is demonstrated in the following result.

Theorem 4. *Let* (P, K, ν, α) *be a feasible solution to the conditions* (21) *with an overestimate of the Lipschitz constant* $\hat{\mathfrak{L}}_\phi > \mathfrak{L}_\phi^*$ *with high probability. Then* (P, K, ν, α) *is a feasible solution to the conditions* (21) *with the true Lipschitz constant* $\mathfrak{L}_\phi^*$ *with high probability.*

Theorem 4 indicates that if our learned $\hat{\mathfrak{L}}_\phi$ is an overestimate of $\mathfrak{L}_\phi^*$, and we use $\hat{\mathfrak{L}}_\phi$ to obtain a safe stabilizing control policy, then this is also a safe stabilizing control policy for the true system (1).

Remark 5. Having a feasible solution to (21) with an underestimator of $\mathfrak{L}_\phi^*$ is not sufficient to guarantee a feasible solution for the true Lipschitz constant, because $\delta\mathcal{M}$ may not be negative semi-definite in that case. Of course, extremely conservative overestimates of $\hat{\mathfrak{L}}_\phi$ will result in conservative control policies or result in infeasibility. In our proposed approach, we have observed that the confidence parameter β dictates the conservativeness of the overestimate. ■

VI. Numerical Example

Consider the problem of estimating an invariant set for the nonlinear system

$$x_{k+1} = x_k + \tau \left(\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ -1 & 1.5 & -2 \end{bmatrix} x_k + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} u_k - \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \phi(q_k) \right),$$

where $\phi(q) = -0.7q^3$ is the unknown nonlinearity and $q = x_3$ is determined automatically via Bayesian ARD. The sampling time $\tau = 0.1$; the continuous-time version of this system was investigated in [25] assuming full model knowledge. The state and input constraints are given by $|x_1| \leq 0.5$, $|x_2| \leq 0.75$, $|x_3| \leq 0.75$, and $-1 \leq u \leq 1$. Clearly, the true Lipschitz constant in this constrained state space is $\mathfrak{L}_\phi^* = 0.7 \times 3 \times 0.75^2 = 1.1812$.

For the purposes of archival data, we simulate the system from the equilibrium with persistent excitation satisfying the control constraints and archive 50 data points from each run; therefore, $N = 500$. We fix $\beta = 10^{-4}$, $\lambda = 10^{-4}$, and

$N_b = 100$ and perform kernelized Lipschitz learning to obtain a Lipschitz constant estimate $\hat{\mathfrak{L}}_\phi = 1.21$ with 50% data for training and validation for KDE. The optimal kernel obtained by cross-validation is the Gaussian kernel and the optimal bandwidth is obtained to be $h = 0.0183$. Subsequently, this value of $\hat{\mathfrak{L}}_\phi$ is used to solve the LMIs (23) with α fixed at 0.5. We obtain a stabilizing control gain with associated invariant set $\mathcal{E}_P$ shown in Fig. 1[A]. As we can see, the invariant ellipsoid (blue) consumes a large percentage of the volume of $\mathbb{X}$: it is not conservative. Furthermore, the trajectories in Fig. 1[B] indicate that state and input constraints are satisfied using this stabilizing policy for 100 initial conditions starting from within $\mathcal{E}_P$. We also check that choosing lazily based on

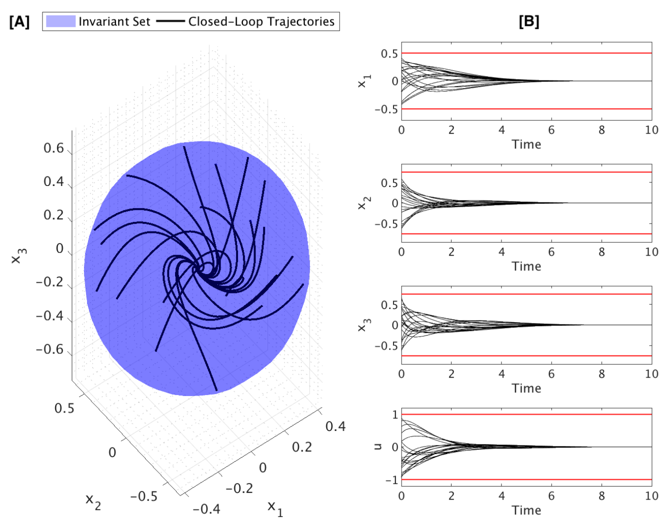

Fig. 1. Illustration of state and input (black lines) constraint satisfaction using Lipschitz learning for invariant set estimation. [A] The invariant set $\mathcal{E}_P$ with a few trajectories shown in state-space. [B] Temporal variation of states and inputs. The constraints are shown using red lines.

the maximum of $\hat{\mathfrak{L}}_\phi$ (11) and then adding a positive scalar (for example, 1), leads to a conservative estimate of the invariant set, whereas ignoring the nonlinearity and designing based on just the known linear component (A, B) results in constraint violation and instability.

VII. Conclusions

Our proposed approach provides a systematic method for constructing stabilizing control policies for systems with unknown dynamics using archival data. The method is applicable for systems where legacy data is available and dynamics are not completely known: for example, smart factories, biomedical systems, and robotic systems. We consign stronger theoretical guarantees on our Lipschitz learning algorithm and generalization of the approach to wider classes of nonlinearities to future work.

References

[1] D. Romeres, M. Zorzi, R. Camoriano, and A. Chiuso, "Online semi-parametric learning for inverse dynamics modeling," in *Proc. of the IEEE Conf. Dec. and Ctrl*, 2016, pp. 2945–2950.

[2] F. Berkenkamp, R. Moriconi, A. P. Schoellig, and A. Krause, "Safe learning of regions of attraction for uncertain, nonlinear systems with Gaussian processes," *Proc. of the IEEE Conf. Decision and Control*, pp. 4661–4666, 2016.

[3] L. Hewing and M. N. Zeilinger, "Cautious model predictive control using gaussian process regression," *CoRR*, vol. abs/1705.10702, 2017. [Online]. Available: http://arxiv.org/abs/1705.10702

[4] Y. Jiang, Y. Wang, S. A. Bortoff, and Z.-P. Jiang, "Optimal co-design of nonlinear control systems based on a modified policy iteration method," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 26, no. 2, pp. 409–414, 2015.

[5] M. Tanaskovic, L. Fagiano, C. Novara, and M. Morari, "Data-driven control of nonlinear systems: An on-line direct approach," *Automatica*, vol. 75, pp. 1–10, 2017.

[6] D. Piga, S. Formentin, and A. Bemporad, "Direct data-driven control of constrained systems," *IEEE Transactions on Control Systems Technology*, vol. 26, no. 4, pp. 1422–1429, 2018.

[7] L. Fagiano and C. Novara, "Automatic crosswind flight of tethered wings for airborne wind energy: a direct data-driven approach," *IFAC Proceedings Volumes*, vol. 47, no. 3, pp. 4927–4932, 2014.

[8] A. Chakrabarty, V. C. Dinh, M. J. Corless, A. E. Rundell, S. H. Zak, G. T. Buzzard *et al.*, "Support vector machine informed explicit nonlinear model predictive control using low-discrepancy sequences." *IEEE Trans. Automat. Contr.*, vol. 62, no. 1, pp. 135–148, 2017.

[9] J.-P. Calliess, "Conservative decision-making and inference in uncertain dynamical systems," Ph.D. dissertation, PhD thesis, University of Oxford, 2014.

[10] D. Limon, J. Calliess, and J. M. Maciejowski, "Learning-based nonlinear model predictive control," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 7769–7776, 2017.

[11] A. Chakrabarty, M. J. Corless, G. T. Buzzard, S. H. Żak, and A. E. Rundell, "State and unknown input observers for nonlinear systems with bounded exogenous inputs," *IEEE Transactions on Automatic Control*, vol. 62, no. 11, pp. 5497–5510, 2017.

[12] X. Xu, B. Açıkmeşe, M. Corless, and H. Sartipizadeh, "Observer-based output feedback control design for systems with incrementally conic nonlinearities," in *Proc. of the American Control Conference (ACC)*, 2018, pp. 1364–1369.

[13] G. Wood and B. Zhang, "Estimation of the Lipschitz constant of a function," *Journal of Global Optimization*, vol. 8, no. 1, pp. 91–103, 1996.

[14] R. G. Strongin, "On the convergence of an algorithm for finding a global extremum," *Engineering Cybernetics*, vol. 11, pp. 549–555, 1973.

[15] P. Hansen, B. Jaumard, and S.-H. Lu, "On using estimates of Lipschitz constants in global optimization," *Journal of Optimization Theory and Applications*, vol. 75, no. 1, pp. 195–200, 1992.

[16] M. E. Tipping, "Sparse bayesian learning and the relevance vector machine," *Journal of machine learning research*, vol. 1, no. Jun, pp. 211–244, 2001.

[17] W. Chu and Z. Ghahramani, "Preference learning with gaussian processes," in *Proceedings of the 22nd international conference on Machine learning*. ACM, 2005, pp. 137–144.

[18] J.-P. Calliess, "Lipschitz optimisation for Lipschitz interpolation," in *American Control Conference (ACC), 2017*, 2017, pp. 3141–3146.

[19] V. Vapnik, *The nature of statistical learning theory*. Springer science & business media, 2013.

[20] Y.-C. Chen, C. R. Genovese, and L. Wasserman, "Density level sets: Asymptotics, inference, and visualization," *Journal of the American Statistical Association*, vol. 112, no. 520, pp. 1684–1696, 2017.

[21] E. Giné and A. Guillou, "Rates of strong uniform consistency for multivariate kernel density estimators," in *Annales de l'Institut Henri Poincare (B) Probability and Statistics*, vol. 38, no. 6. Elsevier, 2002, pp. 907–921.

[22] B. Cadre, "Kernel estimation of density level sets," *Journal of multivariate analysis*, vol. 97, no. 4, pp. 999–1023, 2006.

[23] E. Mammen and W. Polonik, "Confidence regions for level sets," *Journal of Multivariate Analysis*, vol. 122, pp. 202–214, 2013.

[24] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear matrix inequalities in system and control theory*. SIAM, 1994, vol. 15.

[25] C. Bohm, T. Raff, R. Findeisen, and F. Allgower, "Calculating the terminal region of NMPC for Lure systems via LMIs," in *American Control Conference, 2008*, 2008, pp. 1127–1132.