

Modeling of Dynamical Systems via Successive Graph Approximations*

Siddharth H. Nair* Monimoy Bujarbaruah*
Francesco Borrelli*

* *University of California Berkeley, Berkeley, CA 94701 USA (e-mails:
{siddharth_nair, monimoyb, fborrelli}@berkeley.edu).*

Abstract: A non-parametric technique for modeling of systems with unknown nonlinear Lipschitz dynamics is presented. The key idea is to successively utilize measurements to approximate the *graph* of the state-update function of the system dynamics using envelopes described by quadratic constraints. The proposed approach is then used for computing outer approximations of the state-update function using convex optimization. We highlight the efficacy of the proposed approach via a detailed numerical example.

Keywords: Discrete-time systems, Nonparametric Methods, Convex Optimization, Recursive Identification, Autonomous Systems.

1. INTRODUCTION

As data-driven decision making and control becomes ubiquitous (Recht (2019); Rosolia et al. (2018)), system identification methods are being integrated with control algorithms for control of uncertain dynamical systems. In the control systems literature, if the actual model of a system is unknown, adaptive control (Sastry and Bodson (2011)) strategies have been applied for simultaneous system identification and control. Techniques for system modelling and identification have been traditionally rooted in statistics and data sciences. Statistical models that describe observed data, can be classified into parametric, non-parametric and semi-parametric (Friedman et al. (2001)). Parametric models assume a model structure a priori, based on the application and domain expertise of the designer. In almost all of classical adaptive control methods, parametric models are learned from data in terms of point estimates, and asymptotic convergence of such estimates are proven under persistence of excitation conditions. The concept of online model learning and adaptation has been extended to systems under constraints as well (Bujarbaruah et al. (2019)), after obtaining a set or a confidence interval containing possible realizations of the system model.

Parametric models are restricted only to specified forms of function classes. In order to widen the richness of model estimates, non-parametric models are increasingly being utilized, whereby the model structure is also inferred from data. For non-parametric modeling of systems, Gaussian Process (GP) regression (Rasmussen (2003)) has been one of the most widely used tools in control theory literature. GP regression keeps track of a Gaussian distribution over infinite dimensional function spaces, in terms of a mean function and a covariance kernel, which are updated with data. Given any system state, GP regression returns the

mean function value at that state, along with a confidence interval. Estimates obtained using GP regression often come with confidence intervals instead of sets containing all possible realizations of the system, which is a critical drawback for robust control. Contrary to GP regression, the kinky inference method in Calliess (2014) and non-linear set membership in Milanese and Novara (2004) use Lipschitz interpolation for function value estimates with tight bounds. The focus of this paper is to propose a simple non-parametric approach for modelling the *unknown* dynamics of a discrete time autonomous system. Similar to Calliess (2014); Milanese and Novara (2004), our proposed approach applies to unknown nonlinear systems with dynamics described by a state-update function which is globally Lipschitz over a bounded domain, with known Lipschitz constant. However instead of identifying the state-update function itself, we identify its graph- the set of all state and corresponding state-update function value pairs. This is presented in the first part of the paper by computing envelopes of the state-update function, which are sets that contain its graph. These envelopes are built by using historical data of state trajectories. In the second part, we provide a method of obtaining convex outer approximations of the unknown state-update function using the s-procedure (Pólik and Terlaky (2007)).

2. NOTATION

$\|\cdot\|$ denotes the Euclidean norm in \mathbb{R}^n unless explicitly stated otherwise. $Df(x) \in \mathbb{R}^{m \times n}$ denotes the Jacobian of $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ at x . $\mathcal{B}_r(x) \subseteq \mathbb{R}^n$ is the open ball of radius r centered at x . The Big O notation $O(\cdot)$ is used to compare functions $F(\cdot)$, $G(\cdot)$ as $F(\epsilon) = O(G(\epsilon))$ if

$$\exists C, \delta > 0 : \|F(\epsilon)\| \leq C\|G(\epsilon)\| \quad \forall \|\epsilon\| \leq \delta.$$

The Minkowski sum of two sets A and B is

$$A \oplus B = \{a + b \mid a \in A, b \in B\}.$$

We use $\text{ell}(c, R)$ to denote an ellipse that is centered at point c and has a shape matrix $R = R^\top \succ 0$.

* This work was partially funded by Office of Naval Research grant ONR-N00014-18-1-2833.

3. PROBLEM FORMULATION

Consider the discrete time autonomous, time invariant system

$$x_{k+1} = f(x_k), \quad (1)$$

where the state-update function $f(\cdot) : \mathcal{X} \rightarrow \mathcal{X}$ describes the system dynamics and is defined over the state space $\mathcal{X} \subseteq \mathbb{R}^n$.

Assumption 1. The function $f(\cdot) : \mathcal{X} \rightarrow \mathcal{X}$ is continuous and differentiable on convex and closed state space $\mathcal{X} \subseteq \mathbb{R}^n$ with $\|Df(x)\| \leq L$ for all $x \in \mathcal{X}$ and some $L > 0$.

Proposition 1. (Rudin et al. (1964)). Let Assumption 1 hold. Then $\|f(x) - f(y)\| \leq L\|x - y\|$ for all $x, y \in \mathcal{X}$, i.e., $f(\cdot)$ is L -Lipschitz in the domain \mathcal{X} .

Now suppose that the function $f(\cdot)$ is unknown. The objective of this work is to compute a set containing $f(x)$ for any state x in the state space \mathcal{X} using trajectory data $\{x_0, x_1, x_2, \dots\}$ and the Lipschitz property of the unknown function $f(\cdot)$.

Assumption 2. The Lipschitz constant L is known.

Remark 1. The problem of characterizing L -Lipschitz un-modelled dynamics $d(\cdot)$ in

$$x_{k+1} = \underbrace{\bar{f}(x_k)}_{\text{assumed model}} + \underbrace{d(x_k)}_{\text{un-modelled dynamics}}$$

can also be cast into a problem of the form (1). In this case, we use the trajectory data $\{x_0, x_1, x_2, \dots\}$ to construct $\{x_1 - \bar{f}(x_0), x_2 - \bar{f}(x_1), \dots\}$ which is then used for computing a set containing $d(x)$ at $x \in \mathcal{X}$.

4. PROPOSED APPROACH

We will make use of the following definitions.

Definition 1. (Graph). The graph of function $f(\cdot) : \mathcal{X} \rightarrow \mathcal{X}$ is the set

$$G(f) = \{(x, f(x)) \in \mathbb{R}^n \times \mathbb{R}^n \mid \forall x \in \mathcal{X}\}. \quad (2)$$

Definition 2. (Envelope). An envelope of function $f(\cdot) : \mathcal{X} \rightarrow \mathcal{X}$ is a set $\mathbf{E}^f \subseteq \mathbb{R}^n \times \mathbb{R}^n$, with the property

$$G(f) \subseteq \mathbf{E}^f. \quad (3)$$

We use trajectory data $\{x_0, x_1, x_2, \dots\}$ of the system dynamics (1) to construct an envelope of the system dynamics $f(\cdot)$. Observe that the trajectory data can be used to construct tuples $(x_k, f(x_k)) = (x_k, x_{k+1})$. In particular, at every time instant N , we have access to measurements $(x_k, f(x_k))$, for all $k = 0, 1, \dots, N-1$. These measurements are utilized to construct envelopes recursively. Our two step approach for envelope construction is summarized next:

(S1) At time N , compute an envelope $\mathcal{E}(x_{N-1})$ using the tuple $(x_{N-1}, f(x_{N-1}))$ and the L -Lipschitz property of $f(\cdot)$.

(S2) Compute a refined envelope \mathbf{E}_N^f by intersecting the refined envelope \mathbf{E}_{N-1}^f from time $N-1$ with the envelope $\mathcal{E}(x_{N-1})$ computed in step (S1), i.e.,

$$\mathbf{E}_N^f = \mathbf{E}_{N-1}^f \cap \mathcal{E}(x_{N-1}).$$

For (S1), the envelope is obtained as the sublevel set of a quadratic function. Afterwards, the envelope in (S2) is

obtained by using the set membership approach (Bujarbaruah et al. (2018); Tanaskovic et al. (2014)). Finally, we use the computed envelope for obtaining a set containing the value of $f(x)$ at any $x \in \mathcal{X}$, using the notion of a *slice* of an envelope defined below.

Definition 3. (Envelope Slice). The slice of an envelope $\mathbf{E}^f \subseteq \mathbb{R}^n \times \mathbb{R}^n$ at a given $\bar{x} \in \mathcal{X}$ is the set defined as

$$\mathbf{E}^f \Big|_{x=\bar{x}} = \{(x, y) \in \mathbf{E}^f \subseteq \mathbb{R}^n \times \mathbb{R}^n \mid x = \bar{x}\}. \quad (4)$$

Fig. 1 shows a typical realization of the proposed approach along with the associated set definitions which are detailed next.

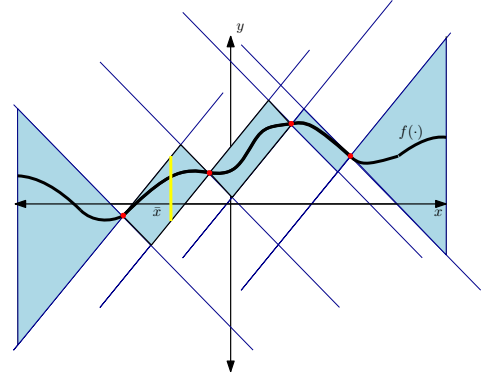


Fig. 1. Construction of an envelope for a one dimensional system to approximate the graph $G(f)$ (black curve) of its state-update function $f(\cdot)$. Tuples $(x, f(x))$ (red points) obtained from trajectory data are used for constructing the envelope (blue set) and its slice (yellow set) at $x = \bar{x}$.

4.1 Envelope Construction

Inspired by Fazlyab et al. (2019); Megretski and Rantzer (1997), we use quadratic constraints (QCs) as our main tool to approximate the graph of a function. A definition appropriate for our purposes is presented below.

Definition 4. (QC Satisfaction). A set $\mathbb{X} \subset \mathbb{R}^n$ satisfies the quadratic constraint specified by a symmetric matrix Q if

$$\begin{bmatrix} x \\ 1 \end{bmatrix}^\top Q \begin{bmatrix} x \\ 1 \end{bmatrix} \leq 0, \quad \forall x \in \mathbb{X}. \quad (5)$$

The following proposition uses a QC to characterize a coarse approximation of the graph of an L -Lipschitz function.

Proposition 2. The graph $G(f)$ of an L -Lipschitz function $f(\cdot)$ satisfies the QC specified by the matrix

$$Q_L^f(x_k) = \begin{bmatrix} -L^2 \mathbb{I}_n & \mathbf{0}_{n \times n} & L^2 x_k \\ \mathbf{0}_{n \times n} & \mathbb{I}_n & -f(x_k) \\ L^2 x_k^\top & -f^\top(x_k) & -L^2 x_k^\top x_k + f^\top(x_k) f(x_k) \end{bmatrix}, \quad (6)$$

for any $(x_k, f(x_k)) \in G(f)$.

Proof. Using the definition of the L -Lipschitz property of $f(\cdot)$ (from Proposition 1), at any $(x_k, f(x_k)) \in G(f)$, we have

$$\begin{aligned} & \|f(x) - f(x_k)\| \leq L\|x - x_k\| \quad \forall (x, f(x)) \in G(f), \\ \Leftrightarrow & (f(x) - f(x_k))^\top (f(x) - f(x_k)) \leq L^2(x - x_k)^\top (x - x_k) \\ & \quad \forall (x, f(x)) \in G(f), \\ \Leftrightarrow & \begin{bmatrix} x \\ f(x) \\ 1 \end{bmatrix}^\top Q_L^f(x_k) \begin{bmatrix} x \\ f(x) \\ 1 \end{bmatrix} \leq 0, \quad \forall (x, f(x)) \in G(f). \end{aligned}$$

Therefore $G(f)$ satisfies the QC specified by $Q_L^f(x_k)$. ■

The following corollary then gives us the definition of the envelope $\mathcal{E}(x_k)$.

Corollary 1. The set defined by

$$\mathcal{E}(x_k) = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^n \mid \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}^\top Q_L^f(x_k) \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \leq 0\} \quad (7)$$

is an envelope for all L -Lipschitz functions that pass through $(x_k, f(x_k))$.

Proof. Follows from Proposition 2. ■

Remark 2. The proposed formulation can also be extended to accommodate bounded noise in the measurements of x_k in (1). Suppose that the measurement model is given by

$$z_k = x_k + w_k,$$

where w_k belongs to a compact set \mathcal{W} . Then the envelope that is guaranteed to contain $G(f)$ is given by $\mathcal{E}(z_k) \oplus (\mathcal{W} \times \mathcal{W})$ where $Q_L^f(\cdot)$ is now constructed using (z_k, z_{k+1}) .

4.2 Successive Graph Approximation

At time N , the envelope $\mathcal{E}(x_{N-1})$ constructed in (7) using the tuple $(x_{N-1}, f(x_{N-1}))$ can now be used to recursively compute a new envelope \mathbf{E}_N^f by refining the envelope \mathbf{E}_{N-1}^f from time $N-1$ via set intersection-

$$\mathbf{E}_N^f = \mathbf{E}_{N-1}^f \cap \mathcal{E}(x_{N-1}) \quad (8)$$

In the following lemma we show that the sets computed in this fashion are indeed envelopes.

Lemma 1. For $N \geq 1$, given a sequence $\{x_k\}_{k=0}^{N-1}$ obtained under the dynamics (1), we have

$$G(f) \subseteq \mathbf{E}_N^f = \mathbf{E}_{N-1}^f \cap \mathcal{E}(x_{N-1}) = \bigcap_{k=0}^{N-1} \mathcal{E}(x_k). \quad (9)$$

Proof. See Appendix. ■

The recursion is initialized with the trivial envelope $\mathbf{E}_0^f = \mathbb{R}^n \times \mathbb{R}^n$. The procedure is described in Algorithm 1.

Algorithm 1 Recursive Envelope Refinement

Initialization: $\mathbf{E}_0^f = \mathbb{R}^n \times \mathbb{R}^n$

Input: At time N , $(x_{N-1}, f(x_{N-1}))$ and \mathbf{E}_{N-1}^f

Output: Approximation of $G(f)$ at time N : \mathbf{E}_N^f

- 1) Compute $Q_L^f(x_{N-1})$ (from (6))
 - 2) Compute $\mathcal{E}(x_{N-1})$ using $Q_L^f(x_{N-1})$ (from (7))
 - 3) Set $\mathbf{E}_N^f = \mathbf{E}_{N-1}^f \cap \mathcal{E}(x_{N-1})$
-

Note that since the envelope at any time N is computed by intersecting with the envelope at time $N-1$, they are getting successively refined, i.e.,

$$\mathbf{E}_N^f \subseteq \mathbf{E}_{N-1}^f \subseteq \mathbf{E}_{N-2}^f \cdots \subseteq \mathbf{E}_0^f \quad (10)$$

Now we provide a condition under which the shrinking sets generated by recursion (8) stop shrinking in finite time or in the limit i.e., recursion (8) attains a fixed point. Intuitively, we would expect this to happen when the incoming tuples $(x, f(x))$ constructed from trajectory data have already been seen previously (or get closer and closer to what already seen). The following definition formalises the notion of such trajectories.

Definition 5. (Zhou (2003)). A p -periodic orbit of the discrete dynamical system (1) is the set of states obtained under dynamics $x_{k+1} = f(x_k)$ with the property that $x_k = x_{k+p}$ for some finite $p \geq 1$ and for all $k \geq 0$, i.e.,

$$\mathcal{O}_p(x_k) = \{x_l \in \mathcal{X} \mid x_{l+1} = f(x_l), l = k, \dots, k+p-1\}. \quad (11)$$

Note that the set $\mathcal{O}_p(\bar{x}_{\text{eq}}) = \{\bar{x}_{\text{eq}}\}$ for all $p \geq 1$ where \bar{x}_{eq} is the fixed point $\bar{x}_{\text{eq}} = f(\bar{x}_{\text{eq}})$ of system (1). Associated to each fixed point, one can define the set of states that converge to it as follows.

Definition 6. (Ortega (1973)). The domain of attraction of fixed point \bar{x}_{eq} is defined as the set

$$\mathbf{D}(\bar{x}_{\text{eq}}) = \{x \in \mathcal{X} \mid x_{k+1} = f(x_k), \lim_{k \rightarrow \infty} x_k = \bar{x}_{\text{eq}}, x_0 = x\}.$$

The following proposition uses Definition 5 and Definition 6 to identify sufficient conditions on system trajectories for termination of the recursion (8).

Proposition 3. Given a system trajectory denoted by the set $\{x_0, x_1, x_2, \dots\}$, the recursion (8) has a fixed point if either of the following conditions hold:

- (1) $\mathcal{O}_p(x_k) \subseteq \{x_0, x_1, x_2, \dots\}$ for some finite $p \geq 1$ and some $k \geq 0$.
- (2) $x_0 \in \mathbf{D}(\bar{x}_{\text{eq}})$ for some fixed point \bar{x}_{eq} .

Proof. See Appendix.

Next we present how the envelope slice is derived from the constructed envelopes for obtaining a set-valued estimate of $f(x)$ at any $x \in \mathcal{X}$.

4.3 Envelope Slice Computation

The set of possible values of function $f(\bar{x})$ at any $\bar{x} \in \mathcal{X}$ can be obtained using (7) from the function values $f(x_k)$ collected at $k = \{0, 1, 2, \dots, N-1\}$. From any k -th measurement, we can obtain a coarse estimate of the set of possible values of $f(\bar{x})$, by constructing the slice of envelope $\mathcal{E}(x_k)$ at $x = \bar{x}$, from Definition 3. We denote this slice with the set $\mathcal{S}(x_k, \bar{x})$ as

$$\begin{aligned} \mathcal{S}(x_k, \bar{x}) &= \mathcal{E}(x_k) \Big|_{x=\bar{x}}, \\ &= \{y \in \mathcal{X} \mid \begin{bmatrix} \bar{x} \\ y \\ 1 \end{bmatrix}^\top Q_L^f(x_k) \begin{bmatrix} \bar{x} \\ y \\ 1 \end{bmatrix} \leq 0\} \\ &= \{y \in \mathcal{X} \mid \begin{bmatrix} y \\ 1 \end{bmatrix}^\top \bar{A}(k, \bar{x}) \begin{bmatrix} y \\ 1 \end{bmatrix} \leq 0\}, \quad (12) \end{aligned}$$

where we have denoted $\bar{A}(k, \bar{x}) = MQ_L^f(x_k)M^\top - L^2 \begin{bmatrix} 0 & 0 \\ 0 & (\bar{x} - 2x_k)^\top \bar{x} \end{bmatrix}$, for any $k = \{0, 1, 2, \dots, N-1\}$, with matrix $M = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. Corollary 1 then implies $f(\bar{x}) \in \mathcal{S}(x_k, \bar{x})$ at any $x \in \mathcal{X}$.

Proposition 4. At any $\bar{x} \in \mathcal{X}$, $\mathcal{S}(x_k, \bar{x})$ is a closed norm ball of radius $L\|\bar{x} - x_k\|$, centered at $f(x_k)$ for each $k = \{0, 1, \dots, N-1\}$.

Proof. Expanding out (12) gives us

$$\|f(\bar{x}) - f(x_k)\|^2 \leq L^2 \|\bar{x} - x_k\|^2, \quad (13)$$

for each $k = \{0, 1, \dots, N-1\}$ and thus proves the claim. ■

As we successively collect data points $(x_k, f(x_k))$ for $k = \{0, 1, 2, \dots, N-1\}$ under dynamics (1), the set of possible values of $f(\bar{x})$ at any $\bar{x} \in \mathcal{X}$ is refined as

$$\mathcal{F}_N(\bar{x}) = \bigcap_{k=0}^{N-1} \mathcal{S}(x_k, \bar{x}) = \bigcap_{k=0}^{N-1} \mathcal{E}(x_k) \Big|_{x=\bar{x}} = \mathbf{E}_N^f \Big|_{x=\bar{x}}, \quad (14)$$

with the guarantee $f(\bar{x}) \in \mathcal{F}_N(\bar{x})$ at any given time $N \geq 1$. Notice that $\mathcal{F}_N(\bar{x})$ is a slice of envelope \mathbf{E}_N^f at $x = \bar{x}$, as per Definition 3. We further note that $\mathcal{F}_N(\bar{x})$ in (14) is convex and compact, as it is an intersection of convex and compact sets (13).

We now wish to quantify the error in approximating the system state as the envelope slice. In the following theorem, we show that if a trajectory starts in the domain of attraction $\mathbf{D}(\bar{x}_{\text{eq}})$ of a fixed point \bar{x}_{eq} of (1), then the error in approximation of $G(f)$ by \mathbf{E}_N^f at points arbitrarily close to \bar{x}_{eq} (measured by the diameter of the envelope slice $\mathcal{F}_N(x)$ at any $x \in \mathcal{X}$), gets arbitrarily small for large enough N .

Theorem 1. Suppose we are given a system trajectory denoted by the set $\{x_0, x_1, x_2, \dots, x_N\}$ where $x_0 \in \mathbf{D}(\bar{x}_{\text{eq}})$. Then for states x arbitrarily close to \bar{x}_{eq} , the diameter of $\mathcal{F}_N(x)$ is arbitrarily small for large enough N , i.e.,

$$\begin{aligned} \forall \epsilon > 0, \exists \bar{N}(\epsilon) : \\ \max_{y \in \mathcal{F}_N(x)} \|y - f(x)\| &= O(\epsilon), \\ \forall x \in \mathcal{B}_\epsilon(\bar{x}_{\text{eq}}), \forall N \geq \bar{N}(\epsilon). \end{aligned}$$

Proof. See Appendix. ■

4.4 Ellipsoidal Outer Approximation of $\mathcal{F}_N(\bar{x})$

In order to design computationally tractable robust optimization algorithms for all realizations of $f(x)$ at any $x \in \mathcal{X}$ and $N \geq 1$, one must have a “nice” geometric representation of the envelope slice $\mathcal{F}_N(\bar{x}) = \mathbf{E}_N^f \Big|_{x=\bar{x}}$, for all $N \geq 1$. We hereby propose an approach to obtain an ellipsoidal outer approximation to $\mathcal{F}_N(\bar{x})$ for any $N \geq 1$ using the s-procedure (Calafiore and El Ghaoui, 2014, Section 11.4), having collected measurements at $k = 0, 1, \dots, N-1$,

Let us parametrize an *ellipsoidal* outer approximation of $\mathcal{F}_N(x_N)$, which we denote by $\text{ell}(c(\bar{x}), R(\bar{x}))$ as

$$\begin{aligned} \text{ell}(c(\bar{x}), R(\bar{x})) = \\ \{y \in \mathbb{R}^n \mid (y - c(\bar{x}))^\top R^{-1}(\bar{x})(y - c(\bar{x})) \leq 1\}, \end{aligned}$$

where vector $c(\bar{x})$ and matrix $R(\bar{x})$ are the decision variables, and are functions of \bar{x} . We seek the smallest ellipsoidal set such that

$$\mathcal{F}_N(\bar{x}) = \bigcap_{k=0}^{N-1} \mathcal{S}_k(x_k, \bar{x}) \subseteq \text{ell}(c(\bar{x}), R(\bar{x})).$$

From s-procedure we know that the above holds true, if there exists scalars $\{\tau_0, \tau_1, \dots, \tau_{N-1}\} \geq 0$ such that

$$\begin{aligned} \begin{bmatrix} R^{-1}(\bar{x}) & -R^{-1}(\bar{x})c(\bar{x}) \\ -c^\top(\bar{x})R^{-1}(\bar{x}) & c^\top(\bar{x})R^{-1}(\bar{x})c(\bar{x}) - 1 \end{bmatrix} \\ - \sum_{k=0}^{N-1} \tau_k \bar{A}^s(k, \bar{x}) \preceq 0. \end{aligned} \quad (15)$$

We reformulate the above feasibility problem (15) as a semi-definite program (SDP) in (Nair et al. (2019)).

5. NUMERICAL EXAMPLE

In this section we demonstrate the approach proposed in Section 4 for characterizing the un-modelled dynamics of a pendulum.

5.1 Pendulum Model

The continuous time model of the considered pendulum is given by

$$ml^2\ddot{\theta} + mgl \sin \theta + \tilde{d}(\theta, \dot{\theta}) = \mathcal{T}, \quad (16)$$

where m is the mass, l is the length, θ is the angle the pendulum makes with the vertical axis, $\tilde{d}(\theta, \dot{\theta})$ is an un-modelled damping force with known Lipschitz constant L_d and \mathcal{T} is a known external torque. In this work, we simulate the system with the damping force $\tilde{d}(\theta, \dot{\theta}) = -L_d\dot{\theta}$ and characterize state-dependent bounds for the same. We write the pendulum dynamics (16) in state-space form as

$$\begin{bmatrix} \dot{\theta} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} 0 & I \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \theta \\ \dot{\theta} \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{\mathcal{T}}{ml^2} - \frac{g}{l} \sin \theta - \frac{\tilde{d}(\theta, \dot{\theta})}{ml^2} \end{bmatrix}, \quad (17)$$

where $x = [\theta \ \dot{\theta}]^\top$ is the state of the pendulum. We consider a torque \mathcal{T} that stabilizes the pendulum’s state when it’s upright, i.e., when $\bar{x}_{\text{eq}} = [\pi \ 0]^\top$. We discretize system (17) and write it in the form of (1) as $x_{k+1} = f(x_k)$. We then simulate the system forward in time with a variational integrator for mechanical systems, as in Nair and Banavar (2019). The simulation parameters are: $m = 2\text{kg}$, $l = 2\text{m}$ and $L_d = 0.2\text{N}$.

5.2 Envelope Construction for Damping Force

The discrete time model $x_{k+1} = f(x_k)$ is decomposed as

$$x_{k+1} = \underbrace{\bar{f}(x_k)}_{\text{assumed model}} + \underbrace{d(x_k)}_{\text{un-modelled damping}}, \quad (18)$$

where $x_k = [\theta_k \ \dot{\theta}_k]^\top$, $d(\cdot)$ is the unknown damping in discrete time with Lipschitz constant $\tilde{L}_d = \frac{L_d T_S}{ml^2}$ and $T_S = 0.005\text{s}$ is the sampling period. Our experiment is succinctly described below:

- Trajectories up to a specified time instant N , starting from four different initial conditions $x_0 = \{[\frac{5\pi}{6} \ 0]^\top, [\frac{5\pi}{3} - 0.5]^\top, [\frac{\pi}{6} \ 0]^\top, [\frac{5\pi}{4} - 0.2]^\top\}$ are simulated (solid lines in Fig. 2) and stored.
- Realizations of the un-modelled dynamics $d(x_k)$ are recorded via $d(x_k) = x_{k+1} - \bar{F}(x_k)$.

- Having recorded the measurements $(x_k, d(x_k))$ for $k = 0, 1, \dots, N - 1$ along all four trajectories, we construct the estimate $\mathcal{D}_N(\bar{x})$ (defined as in (14)) of $d(\bar{x})$ at six different points (\star and \circ in Fig. 2) using Algorithm 1.

Table 1. Uncertainty range (up to three significant digits). Symbol $[\cdot, \cdot]$ denotes an interval

Point \bar{x}	$d(\bar{x})/10^{-4}$	$\mathcal{D}_{100}(\bar{x})/10^{-4}$	$\mathcal{D}_{4000}(\bar{x})/10^{-4}$
$[2.12 \ 0.45]^\top$	0.563	$[-0.837, 0.759]$	$[-0.001, 0.759]$
$[3.11 \ 0.84]^\top$	-1.04	$[-1.22, 0.73]$	$[-1.04, -1.04]$
$[1.40 \ 0.34]^\top$	-0.43	$[-1.58, 0.79]$	$[-0.43, -0.43]$
$[3.05 \ -0.37]^\top$	0.46	$[-0.708, 0.486]$	$[0.46, 0.46]$
$[4.21 \ 0.38]^\top$	-0.47	$[-2.05, 0.74]$	$[-0.56, 0.16]$
$[5.60 \ 0.22]^\top$	-0.28	$[-3.73, 2.46]$	$[-0.28, -0.28]$

From Table 1, we observe that the range of un-modelled dynamics $\mathcal{D}_N(\bar{x})$ shrinks at all points \bar{x} , as more data is collected. This is a direct consequence of the fact that as shown in (14), $\mathcal{D}_N(\bar{x})$ is obtained with successive intersection operations upon gathering new measurements. More-

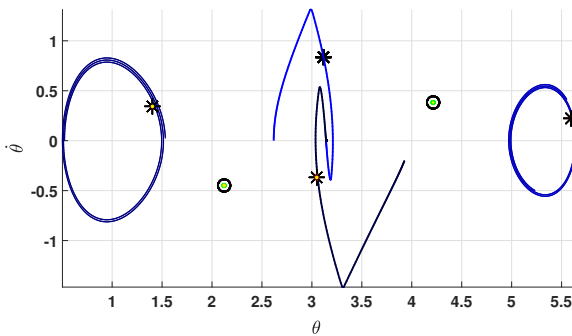


Fig. 2. Data collection trajectories (solid lines) and \bar{x} points (\star and \circ) in state-space. \star is used for \bar{x} close to recorded trajectories while \circ is used for \bar{x} far from recorded trajectories.

over, the learned dynamics are more accurate for points near recorded trajectories than for query points far away, as shown in Fig. 2. For example, at star(\star) points : $\bar{x} = \{[1.40 \ 0.34]^\top, [3.05 \ -0.37]^\top, [3.11 \ 0.84]^\top, [5.60 \ 0.22]^\top\}$, we see around 100% decrease in the uncertainty range estimate as N increases from 100 to 4000. The corresponding percentages at circle(\circ) points : $\bar{x} = \{[4.21 \ 0.38]^\top, [2.12 \ -0.45]^\top\}$ are just around 73% and 34% respectively.

6. CONCLUSIONS

We presented a non-parametric technique for online modeling of systems with nonlinear Lipschitz dynamics. The key idea is to successively use measurements to approximate the *graph* of the function using envelopes described by quadratic constraints. Using techniques from convex optimization, we also computed a set valued estimate of the range of the unknown function at any given point in its domain. We further highlighted the efficacy of the proposed methodology via a detailed numerical example.

REFERENCES

- Bujarbaruah, M., Zhang, X., and Borrelli, F. (2018). Adaptive MPC with chance constraints for FIR systems. In *2018 Annual American Control Conference (ACC)*, 2312–2317. doi:10.23919/ACC.2018.8431586.
- Bujarbaruah, M., Zhang, X., Tanaskovic, M., and Borrelli, F. (2019). Adaptive MPC under time varying uncertainty: Robust and Stochastic. *arXiv preprint arXiv:1909.13473*.
- Calafiore, G.C. and El Ghaoui, L. (2014). *Optimization models*. Cambridge university press.
- Calliess, J.P. (2014). *Conservative decision-making and inference in uncertain dynamical systems*. Ph.D. thesis, University of Oxford.
- Fazlyab, M., Morari, M., and Pappas, G.J. (2019). Safety verification and robustness analysis of neural networks via quadratic constraints and semidefinite programming. *arXiv preprint arXiv:1903.01287*.
- Friedman, J., Hastie, T., and Tibshirani, R. (2001). *The elements of statistical learning*, volume 1. Springer series in statistics New York.
- Megretski, A. and Rantzer, A. (1997). System analysis via integral quadratic constraints. *IEEE Transactions on Automatic Control*, 42(6), 819–830.
- Milanese, M. and Novara, C. (2004). Set membership identification of nonlinear systems. *Automatica*, 40(6), 957–975.
- Nair, S.H. and Banavar, R.N. (2019). Discrete optimal control of interconnected mechanical systems. *11th IFAC Symposium on Nonlinear Control Systems (NOLCOS)*.
- Nair, S.H., Bujarbaruah, M., and Borrelli, F. (2019). Modeling of dynamical systems via successive graph approximations. *arXiv preprint arXiv:1910.03719*.
- Ortega, J.M. (1973). Stability of difference equations and convergence of iterative processes. *SIAM Journal on Numerical Analysis*, 10(2), 268–282.
- Pólik, I. and Terlaky, T. (2007). A survey of the s-lemma. *SIAM review*, 49(3), 371–418.
- Rasmussen, C.E. (2003). Gaussian processes in machine learning. In *Summer School on Machine Learning*, 63–71. Springer.
- Recht, B. (2019). A tour of reinforcement learning: The view from continuous control. *Annual Review of Control, Robotics, and Autonomous Systems*, 2, 253–279.
- Rosolia, U., Zhang, X., and Borrelli, F. (2018). Data-driven predictive control for autonomous systems. *Annual Review of Control, Robotics, and Autonomous Systems*, 1, 259–286.
- Rudin, W. et al. (1964). *Principles of mathematical analysis*, volume 3. McGraw-hill New York.
- Sastry, S. and Bodson, M. (2011). *Adaptive control: Stability, convergence and robustness*. Courier Corporation.
- Tanaskovic, M., Fagiano, L., Smith, R., and Morari, M. (2014). Adaptive receding horizon control for constrained MIMO systems. *Automatica*, 50(12), 3019–3029.
- Zhou, Z. (2003). Periodic orbits on discrete dynamical systems. *Computers & Mathematics with Applications*, 45(6-9), 1155–1161.

7. APPENDIX

7.1 Proof of Lemma 1

For any $(x, f(x)) \in G(f)$, we have from the Lipschitz inequality,

$$\|f(x) - f(y)\| \leq L\|x - y\|, \quad \forall y \in \mathcal{X},$$

and choosing $y = x_k$ for $k = 0, 1, \dots, N-1$ in the above inequality, in view of Corollary 1 yields,

$$\begin{aligned} (x, f(x)) &\in \mathcal{E}(x_k), \quad \forall k = 0, 1, \dots, N-1, \\ &\Rightarrow (x, f(x)) \in \bigcap_{k=0}^{N-1} \mathcal{E}(x_k). \end{aligned}$$

Note the fact that f is globally Lipschitz ensures that the intersections are non-empty. Since this was shown for any $(x, f(x)) \in G(f)$, we can thus conclude that

$$G(f) = \bigcup_{x \in \mathcal{X}} (x, f(x)) \subseteq \bigcap_{k=0}^{N-1} \mathcal{E}(x_k).$$

The other equalities follow from (8). ■

7.2 Proof of Proposition 3

We first prove the implication for when condition (1) holds. Let k^* be the time at which the system enters the p -periodic orbit, i.e., $\mathcal{O}_p(x_{k^*}) \subseteq \{x_0, x_1, x_2, \dots\}$. From Lemma 1 we have $\mathbf{E}_N^f = \bigcap_{k=0}^{N-1} \mathcal{E}(x_k)$ at any time N . For $N \geq k^* + p$, consider the set

$$\begin{aligned} \mathbf{E}_{N+1}^f &= \bigcap_{k=0}^N \mathcal{E}(x_k) \\ &= \left(\bigcap_{k=0}^{k^*-1} \mathcal{E}(x_k) \right) \cap \left(\bigcap_{k=k^*}^{k^*+p-1} \mathcal{E}(x_k) \right) \cap \left(\bigcap_{k=k^*+p}^N \mathcal{E}(x_k) \right), \\ &= \mathbf{E}_{k^*+p}^f \cap \left(\bigcap_{k=k^*+p}^N \mathcal{E}(x_k) \right). \end{aligned}$$

From Definition 5, we have that $x_k \in \mathcal{O}_p(x_{k^*})$ for all $k = k^* + p, \dots, N$. Using (9) and the fact that $f(\cdot)$ is globally Lipschitz on \mathcal{X} , we thus have $\mathbf{E}_{k^*+p}^f \subseteq \mathcal{E}(x_k)$, for all $k = k^* + p, \dots, N$. Combining this implication with the definition of \mathbf{E}_{N+1}^f above yields

$$\mathbf{E}_{N+1}^f = \mathbf{E}_{k^*+p}^f, \quad \forall N \geq k^* + p,$$

and so $\mathbf{E}_{k^*+p}^f$ is a fixed point for recursion (8).

Now we prove the implication for when condition (2) holds, i.e., $\lim_{k \rightarrow \infty} x_k = \bar{x}_{\text{eq}}$. Since the sets $\mathbf{E}_N^f = \bigcap_{k=0}^{N-1} \mathcal{E}(x_k)$ are non-increasing in the sense of (10),

$$\begin{aligned} \mathbf{E}_*^f &= \lim_{N \rightarrow \infty} \mathbf{E}_N^f = \lim_{N \rightarrow \infty} \bigcap_{k=0}^{N-1} \mathcal{E}(x_k), \\ &= \lim_{N \rightarrow \infty} \left[\left(\bigcap_{k=0}^{N-2} \mathcal{E}(x_k) \right) \cap \mathcal{E}(x_{N-1}) \right], \\ &= \left[\lim_{N \rightarrow \infty} \bigcap_{k=0}^{N-2} \mathcal{E}(x_k) \right] \cap \left[\mathcal{E} \left(\lim_{N \rightarrow \infty} x_{N-1} \right) \right]. \end{aligned}$$

The last equality follows from the property of product of convergent sequences. Computing the limits then gives $\mathbf{E}_*^f = \mathbf{E}_*^f \cap \mathcal{E}(\bar{x}_{\text{eq}})$. Thus \mathbf{E}_*^f is a fixed point for (8). ■

7.3 Proof of Theorem 1

From the definition x_0 in the theorem, we have that the sequence $\{x_k\}_{k=0}^{\infty}$ converges to the fixed point \bar{x}_{eq} of (1).

From the definition of the convergence of a sequence, we have that for every $\epsilon > 0$, there exists a $\bar{N}(\epsilon)$, such that

$$\|x_k - \bar{x}_{\text{eq}}\| \leq \epsilon, \quad \forall k \geq \bar{N}(\epsilon).$$

The convergent sequence is a Cauchy sequence satisfying with the same ϵ and $\bar{N}(\epsilon)$ as above. That is,

$$\|x_{k_1} - x_{k_2}\| \leq 2\epsilon, \quad \forall k_2, k_1 \geq \bar{N}(\epsilon). \quad (19)$$

Consider a point \bar{x} at most ϵ distance away from \bar{x}_{eq} . We further have from the Lipschitz inequality,

$$\|f(x_{k_1}) - f(\bar{x})\| \leq L\|x_{k_1} - \bar{x}\|. \quad (20)$$

From Proposition 4, we know that the possible values of $f(\bar{x})$ lie within a sphere of radius $L\|x_{k_1} - \bar{x}\|$ centered at $f(x_{k_1})$. The diameter of the above sphere bounds the maximum error in the estimate of $f(\bar{x})$, i.e.,

$$\|y - f(\bar{x})\| \leq 2L\|x_{k_1} - \bar{x}\|, \quad \forall y \in \mathcal{S}(x_{k_1}, \bar{x}).$$

For k_1 chosen as in (19), the above inequality can be written as

$$\|y - f(\bar{x})\| \leq 4L\epsilon, \quad \forall y \in \mathcal{S}(x_{k_1}, \bar{x}).$$

Now for another k_2 chosen as in (19) such that $f(x_{k_1}) \neq f(x_{k_2})$, we have

$$\|f(x_{k_2}) - f(\bar{x})\| \leq L\|x_{k_2} - \bar{x}\|. \quad (21)$$

The intersections of the envelopes constructed from (20) and (21) is depicted in Fig. 3. We thus obtain a tighter

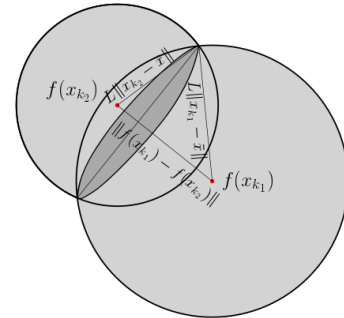


Fig. 3. Intersection of set of possible values of $f(\bar{x})$, given x_{k_1}, x_{k_2}

bound on the error in the estimate of $f(\bar{x})$ via the diameter of the $n-2$ dimensional sphere obtained at the intersection of $n-1$ dimensional spheres, as given by

$$\begin{aligned} &\forall y \in \mathcal{S}(x_{k_1}, \bar{x}) \cap \mathcal{S}(x_{k_2}, \bar{x}) : \|y - f(\bar{x})\| \leq \\ &2\sqrt{\frac{L^2(\|x_{k_1} - \bar{x}\|^2 + \|x_{k_2} - \bar{x}\|^2) - \|f(x_{k_2}) - f(x_{k_1})\|^2}{2}} \\ &\Rightarrow \max_{y \in \mathcal{S}(x_{k_1}, \bar{x}) \cap \mathcal{S}(x_{k_2}, \bar{x})} \|y - f(\bar{x})\| \leq \\ &2\sqrt{\frac{L^2(\|x_{k_1} - \bar{x}\|^2 + \|x_{k_2} - \bar{x}\|^2) - \|f(x_{k_2}) - f(x_{k_1})\|^2}{2}}, \\ &\leq \min(2L\|x_{k_2} - \bar{x}\|, 2L\|x_{k_1} - \bar{x}\|) \leq 4L\epsilon \end{aligned}$$

Taking intersections using all the envelopes collected (which are non-empty due to Lipschitz property of $f(\cdot)$ on \mathcal{X}) further shrinks the possible error and hence yields the desired result. ■