Online Data-Driven Reachability Analysis using Zonotopic Recursive Least Squares

Alireza Naderi Akhormeh ^a, Amr Hegazy ^{a,b}, and Amr Alanwar ^a

^a Technical University of Munich, Germany ^b German University in Cairo, Egypt

Abstract

Reachability analysis is a key formal verification technique for ensuring the safety of modern cyberphysical systems subject to uncertainties in measurements, system models (parameters), and inputs. Classical model-based approaches rely on accurate prior knowledge of system dynamics, which may not always be available or reliable. To address this, we present a data-driven reachability analysis framework that computes over-approximations of reachable sets directly from online state measurements. The method estimates time-varying unknown models using an Exponentially Forgetting Zonotopic Recursive Least Squares (EF-ZRLS) method, which processes data corrupted by bounded noise. Specifically, a time-varying set of models that contains the true model of the system is estimated recursively, and then used to compute the forward reachable sets under process noise and uncertain inputs. Our approach applies to both discrete-time Linear Time-Varying (LTV) and nonlinear Lipschitz systems. Compared to existing techniques, it produces less conservative reachable-set over-approximations, remains robust under slowly varying dynamics, and operates solely on real-time data without requiring any pre-recorded offline experiments. Numerical simulations and real-world experiments validate the effectiveness and practical applicability of the proposed algorithms.

Key words: Data-driven Reachability analysis, robust estimation, set-based parameter estimation, zonotopic recursive least squares, system identification, safety analysis.

1 Introduction

Reachability analysis in cyber-physical systems identifies all bounded reachable sets to formally avoid unsafe sets. Reachability computation is categorized as modelbased when the analytical dynamics of the dynamic system are known, and data-driven when only state or output measurements are available. In model-based analysis, different approaches such as support function [19], level set [20], and zonotope [5] are commonly used. Certain methods, such as level set techniques, can provide an exact representation of reachable sets for different system classes. However, these methods rely on predefined models, which are often derived from first principles. In many real-world dynamic systems, constructing an accurate model is non-trivial, which makes model-based reachability approaches challenging when an exact system model is unavailable.

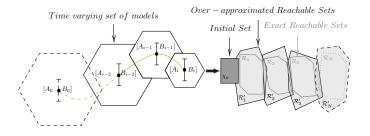


Fig. 1. The reachable sets consistent with noisy inputstate data for a time-varying dynamical system.

Data-driven reachability analysis, unlike model-based approaches, enables the computation of over-approximate reachable sets directly from observed input-output data, even when an explicit system model is unavailable. This paradigm constructs a data-driven representation from recorded or real-time trajectories, which is then used to approximate the systems reachable set. Recent advances leverage deep learning to solve the Hamilton-Jacobi (HJ) Partial Differential Equation (PDE) for high-dimensional systems. For instance, [6] employs a self-supervised PDE loss to learn the value function offline, assuming known system dynamics (the

^{*} Corresponding author Alireza Naderi Akhormeh.

Email addresses: alireza.naderi@tum.de (Alireza Naderi
Akhormeh), amr.hazem@student.guc.edu.eg (Amr Hegazy),
alanwar@tum.de (Amr Alanwar).

Hamiltonian is analytically computed, not learned). In contrast, [11] extends this framework to partially unknown dynamics by incorporating data-driven Hamiltonian approximation. Both methods operate offline but enable real-time reachability queries once trained. While these methods can compute more accurate reachable sets than other approaches, they require substantial training data and struggle with scalability in high-dimensional systems due to computational complexity.

Using the zonotope, which can reduce computational cost in robust estimation, could be beneficial. Zonotopes have been applied in reachability analysis of logical systems [1], hybrid logical and dynamical systems [8,12], as well as state estimation and fault diagnosis [24]. Consequently, zonotope-based approaches can compute more conservative reachable sets with less data and offer better scalability in higher-dimensional systems. In [2], reachable sets of discrete-time linear and nonlinear systems are computed using a zonotope-based approach by estimating a set of models from noisy input-state data. By collecting multiple trajectories, the method identifies a set of consistent models that account for known bounded noise corrupting the measurements. Although the method effectively handles uncertainty and model estimation from collected data using the least squares method and has been successfully applied to provide safety in control of nonlinear systems [14], it still requires pre-collected trajectories to construct the set of datadriven models. This makes it challenging when system models change over time steps or when there is difficulty in collecting offline data. To overcome model estimation challenges, recursive model estimation methods can be used to estimate and update models over time steps with new data. A set-based zonotopic recursive least squares method, introduced in [26], was designed for model estimation in a single-state (single-output regression model) Linear Time-Invariant (LTI) system.

The main idea of the proposed data-driven reachability analysis framework is illustrated in Figure 1. Our proposed method computes data-driven reachable sets using a recursive matrix zonotope formulation. To ensure that the reachable set encloses all possible system trajectories derived from finite noisy data, we construct a time-varying matrix zonotope that captures all models consistent with the noisy measurements, rather than relying on a single potentially inaccurate model. This is achieved through a set-based EF-ZRLS method, which estimates the models of the multi-output, timevarying regression models whose outputs are corrupted by bounded noise. The true system models are guaranteed to lie within these sets. Using the latest identified matrix zonotope that contains the true model, we propagate the initial state set forward to obtain the corresponding over-approximated reachable sets based on the ideas presented in [2] for a discrete-time LTV system. Finally, we extend the framework to address a broader class of discrete-time nonlinear Lipschitz systems.

The main contributions of this paper are summarized as follows:

- (1) A set-based EF-ZRLS method is proposed for estimating the unknown model of a time-varying multioutput regression model with measurements corrupted by bounded noise (Proposition 1). The optimality of the EF-ZRLS is established, and the corresponding optimal model estimation correction gain is derived (Theorem 1). The proposed EF-ZRLS formulation is applicable to both vector and matrix zonotopes.
- (2) Algorithm 1 is introduced to compute overapproximated reachable sets (Theorem 2) for discrete-time LTV systems (25) subject to bounded process noise.
- (3) Algorithm 2 is introduced to compute overapproximated reachable sets (Theorem 3) for a class of discrete-time nonlinear Lipschitz systems (1) subject to bounded process noise.
- (4) A real-world experiment, two numerical examples, and a comparative analysis are presented, contrasting the proposed data-driven reachability analysis with the alternative model-based approach.

To recreate our results, readers can use our publicly available code 1 and a video of experimental results 2 .

The remainder of the paper is organized as follows. Section 2 presents the problem formulation and relevant preliminaries. In the subsequent section, we propose a robust recursive set-based estimation method to estimate the unknown models of a time-varying regression model corrupted by noisy data. In Section 4, we introduce our proposed data-driven reachability analysis method for LTV and nonlinear Lipschitz continuous dynamical systems. Section 5, evaluates the effectiveness of the proposed method through numerical examples and an experimental case study. Finally, Section 6 concludes the paper and discusses potential directions for future work.

2 Problem Formulation and Preliminaries

In this section, we present the problem statement, the preliminary definitions for reachability analysis, and introduce details about the notation used throughout this work.

2.1 Problem Formulation

We consider a discrete-time system

$$x_{k+1} = f(x_k, u_k) + w_k. (1)$$

 $^{^{1}\ \}mathrm{https://github.com/TUM-CPS-HN/ZRLS}$

 $^{^{\}mathbf{2}} \ \, \text{https://www.youtube.com/watch?v=6va1PpCKL9A}$

where $f: \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}^{n_x}$ a differentiable unknown function, $w_k \in \mathcal{Z}_w \subset \mathbb{R}^{n_x}$ denotes the noise bounded by a noise zonotope \mathcal{Z}_w , $u_k \in \mathcal{U}_k$ the input bounded by an input zonotope $\mathcal{U}_k \in \mathbb{R}^{n_u}$, $x_k \in \mathbb{R}^{n_x}$ the measured state, and $x_0 \in \mathcal{X}_0 \subset \mathbb{R}^{n_x}$ the initial state x_0 of the system bounded by the initial set \mathcal{X}_0 .

Reachability analysis computes the set of states x_k that can be reached given a set of uncertain initial states \mathcal{X}_0 and a set of uncertain inputs \mathcal{U}_k . More formally, it can be defined as follows:

Definition 1 (Exact Reachable Set) The exact reachable set \mathcal{R}_N after N time steps subject to inputs $u_k \in \mathcal{U}_k$, $\forall k \in \{0, \dots, N-1\}$, and noise $w_k \in \mathcal{Z}_w$, is the set of all states trajectories starting from initial set \mathcal{X}_0 after N steps:

$$\mathcal{R}_{N} = \left\{ x_{N} \in \mathbb{R}^{n_{x}} \middle| x_{k+1} = f(x_{k}, u_{k}) + w_{k}, x_{0} \in \mathcal{X}_{0}, \\ u_{k} \in \mathcal{U}_{k}, w_{k} \in \mathcal{Z}_{w} : \forall k \in \{0, ..., N-1\} \right\}.$$
 (2)

We consider a sliding window single input-state trajectory with lengths N_D , denoted by $\{u_i\}$, and $\{x_i\}$, for all $i \geq 0$. Let us denote the sliding window using the shifted signals as follows

$$X_{+} = \begin{bmatrix} x_{i-N_D+1}, x_{i-N_D+2}, \dots, x_i \end{bmatrix},$$

$$X_{-} = \begin{bmatrix} x_{i-N_D}, x_{i-N_D+1}, \dots, x_{i-1} \end{bmatrix},$$

$$U_{-} = \begin{bmatrix} u_{i-N_D}, u_{i-N_D+1}, \dots, u_{i-1} \end{bmatrix}.$$

We denote the set of all available data by $D_{-} = (X_{-}, U_{-})$ and $D = (X_{+}, X_{-}, U_{-})$.

We aim to compute an over-approximation of the exact reachable sets when the model of the system in (1) is unknown, but input and noisy states are available at each time step.

2.2 Notation

The set of real and natural numbers are denoted as \mathbb{R} and \mathbb{N} , respectively, and $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. The 2-norm of a vector $\bar{x} \in \mathbb{R}^{\bar{n}}$ is defined as $\|\bar{x}\|_2 := \sqrt{\sum_{i=1}^{\bar{n}} |\bar{x}_i|^2}$. For a matrix $A \in \mathbb{R}^{n_A \times m_A}$, the Frobenius norm is defined as $\|A\|_F := \sqrt{\sum_{i=1}^{n_A} \sum_{j=1}^{m_A} |a_{ij}|^2} = \|\text{vec}(A)\|_2$, and the matrix infinity norm is defined as $\|A\|_{max} := \max_{i,j} |a_{ij}|$. The transpose and MoorePenrose pseudoinverse of a matrix X are denoted as X^{\top} and X^{\dagger} , respectively. We denote the Kronecker product by \otimes , the element at row i and column j of matrix A by $(A)_{i,j}$, and column j of A by $(A)_{i,j}$. The vectorization of a matrix A is defined by $\text{vec}(A) \in \mathbb{R}^{n_A m_A}$, obtained by stacking the columns

of A into a single column vector. Conversely, the inverse operation $\operatorname{unvec}(\bar{x})$ reshapes a vector $\bar{x} \in \mathbb{R}^{n_A m_A}$ back into a matrix in $\mathbb{R}^{n_A \times m_A}$. For a list or vector of elements, we denote the element i of vector or list a by $a^{(i)}$. The element-wise (Hadamard) product of two matrices is denoted by \odot . We use $\mathcal{Z}_1 \ominus \mathcal{Z}_2$ to denote $\mathcal{Z}_1 \oplus (-1)\mathcal{Z}_2$, which is not the Minkowski difference. We also define for N time steps

$$\mathcal{F} = \cup_{k=0}^{N} (\mathcal{R}_k \times \mathcal{U}_k).$$

2.3 Set Representations

We start by defining some set representations that are used in the reachability analysis.

Definition 2 (**Zonotope** [16]) Given a center $c_{\mathcal{Z}} \in \mathbb{R}^{n_{\mathcal{Z}}}$ and $\gamma_{\mathcal{Z}} \in \mathbb{N}$ generator vectors in a generator matrix $G_{\mathcal{Z}} = \left[g_{\mathcal{Z}}^{(1)} \dots g_{\mathcal{Z}}^{(\gamma_{\mathcal{Z}})}\right] \in \mathbb{R}^{n_{\mathcal{Z}} \times \gamma_{\mathcal{Z}}}$, a zonotope is defined as

$$\mathcal{Z} = \left\{ x \in \mathbb{R}^{n_{\mathcal{Z}}} \mid x = c_{\mathcal{Z}} + \sum_{i=1}^{\gamma_{\mathcal{Z}}} \beta^{(i)} g_{\mathcal{Z}}^{(i)}, \left| \beta^{(i)} \right| \leq 1 \right\}.$$
(3)

We use the shorthand notation $\mathcal{Z} = \langle c_{\mathcal{Z}}, G_{\mathcal{Z}} \rangle$ for a zonotope.

Let $L \in \mathbb{R}^{m_L \times n_Z}$ be a linear map. Then $LZ = \langle Lc_Z, LG_Z \rangle$ [4, p.18].

Given two zonotopes $\mathcal{Z}_1 = \langle c_{\mathcal{Z}_1}, G_{\mathcal{Z}_1} \rangle$ and $\mathcal{Z}_2 = \langle c_{\mathcal{Z}_2}, G_{\mathcal{Z}_2} \rangle$, the Minkowski sum $\mathcal{Z}_1 \oplus \mathcal{Z}_2 = \{z_1 + z_2 | z_1 \in \mathcal{Z}_1, z_2 \in \mathcal{Z}_2\}$ can be computed exactly as follows [16]:

$$\mathcal{Z}_1 \oplus \mathcal{Z}_2 = \left\langle c_{\mathcal{Z}_1} + c_{\mathcal{Z}_2}, [G_{\mathcal{Z}_1}, G_{\mathcal{Z}_2}] \right\rangle. \tag{4}$$

We define and compute the Cartesian product of two zonotopes \mathcal{Z}_1 and \mathcal{Z}_2 by

$$\mathcal{Z}_{1} \times \mathcal{Z}_{2} = \left\{ \begin{bmatrix} z_{1} \\ z_{2} \end{bmatrix} \middle| z_{1} \in \mathcal{Z}_{1}, z_{2} \in \mathcal{Z}_{2} \right\} \\
= \left\langle \begin{bmatrix} c_{\mathcal{Z}_{1}} \\ c_{\mathcal{Z}_{2}} \end{bmatrix}, \begin{bmatrix} G_{\mathcal{Z}_{1}} & 0 \\ 0 & G_{\mathcal{Z}_{2}} \end{bmatrix} \right\rangle.$$
(5)

Definition 3 (Matrix Zonotope [4, p.52]) Given a center matrix $C_{\mathcal{M}} \in \mathbb{R}^{n_{\mathcal{M}} \times m_{\mathcal{M}}}$ and $\gamma_{\mathcal{M}} \in \mathbb{N}$ generator matrices $\check{G}_{\mathcal{M}} = \left[G_{\mathcal{M}}^{(1)} \dots G_{\mathcal{M}}^{(\gamma_{\mathcal{M}})}\right] \in \mathbb{R}^{n_{\mathcal{M}} \times (m_{\mathcal{M}} \gamma_{\mathcal{M}})}$, a matrix zonotope is defined as

$$\mathcal{M} = \left\{ X \in \mathbb{R}^{n_{\mathcal{M}} \times m_{\mathcal{M}}} \mid X = C_{\mathcal{M}} + \sum_{i=1}^{\gamma_{\mathcal{M}}} \beta^{(i)} G_{\mathcal{M}}^{(i)}, \left| \beta^{(i)} \right| \leq 1 \right\}.$$
(6)

We use the shorthand notation $\mathcal{M} = \langle C_{\mathcal{M}}, \check{G}_{\mathcal{M}} \rangle$ for a matrix zonotope.

A matrix zonotope can be over-approximated by an interval matrix $I_{\mathcal{M}} = \operatorname{int}(\mathcal{M}) = [\underline{\mathcal{M}}, \overline{\mathcal{M}}]$, where $\underline{\mathcal{M}} \in \mathbb{R}^{n_{\mathcal{M}} \times m_{\mathcal{M}}}$ and $\overline{\mathcal{M}} \in \mathbb{R}^{n_{\mathcal{M}} \times m_{\mathcal{M}}}$ denote its lower bound and upper bound, respectively. The Frobenius norm of the interval matrix $I_{\mathcal{M}}$ is defined as $||I_{\mathcal{M}}|| = ||I_{\mathcal{M}_c}| + \Delta||_F$ with $I_{\mathcal{M}_c} = 0.5(\overline{\mathcal{M}} + \underline{\mathcal{M}})$ and $\Delta = 0.5(\overline{\mathcal{M}} - \mathcal{M})$ [13].

By construction, any matrix zonotope (Definition 3) can be put in one-to-one correspondence with a vector zonotope (Definition 2), via

$$\mathcal{Z} = \{ \operatorname{vec}(M) \mid M \in \mathcal{M} \}, \quad \mathcal{M} = \{ \operatorname{unvec}(Z) \mid Z \in \mathcal{Z} \},$$

with centers and generators related by

$$c_{\mathcal{Z}} = \operatorname{vec}(C_{\mathcal{M}}), \quad g_{\mathcal{Z}_i} = \operatorname{vec}(G_{\mathcal{M}_i}), \quad \forall i = 1, \dots, \gamma_{\mathcal{M}}.$$

3 Robust Adaptive Recursive Parameter Estimation

A robust adaptive parameter (model) estimation method is proposed for a linear regression model with noisy measurements and time-varying parameters. The approach utilizes a recursive, set-based method to iteratively estimate a time-varying set of parameters over consecutive time steps, which has been introduced in [26] for a single-output time-invariant linear regression model. We extend it to the time-varying multi-output regression model.

3.1 Exponentially Forgetting Zonotopic Recursive Least Squares

We consider a discrete-time system in regression form given by

$$y_k = \varphi_k \theta_{\text{tr},k} + v_k, \theta_{\text{tr},k} = \theta_{\text{tr},k-1} + \delta \theta_{k-1}.$$
 (7)

where $y_k \in \mathbb{R}^{p \times m}$ is the measured output, $\varphi_k \in \mathbb{R}^{p \times n}$ is the regressor matrix at time step k, and $\theta_{\mathrm{tr},k} \in \mathbb{R}^{n \times m}$ is the true time-varying (unknown) parameter matrix, and the term $\delta \theta_{k-1} \in \mathbb{R}^{n \times m}$ represents the change in the parameters.

Assumption 1 (Bounded Noise) The additive measurement noise $v_k \in \mathbb{R}^{p \times m}$ is bounded as $||v_k||_{max} \leq \sigma_v$, with $\sigma_v \in \mathbb{R}_{\geq 0}$ a known scalar for all $k \geq 0$. To describe the uncertainty structure of the noise, we define the generator matrix $G_v \in \mathbb{R}^{p \times pm}$ as $G_v = \left[Q_v^{(1)}, Q_v^{(2)}, \dots, Q_v^{(pm)}\right]$, where each $Q_v^{(l)} \in \mathbb{R}^{p \times m}$

contains exactly one nonzero element equal to σ_v , and all other entries are zero. Specifically, for each $l \in \{1, \ldots, pm\}$, there exists a unique index pair (i, j) such that $(Q_v^{(l)})_{i,j} = \sigma_v$, and $(Q_v^{(l)})_{i_l,j_l} = 0 \quad \forall \ (i_l, j_l) \neq (i, j)$. Then, the noise satisfies $v_k \in \langle 0, G_v \rangle$ for all k.

Assumption 2 (Bounded Parameter Variations)

There exists $\sigma_{\theta} \geq 0$ such that, $\|\delta\theta_{k-1}\|_{max} \leq \sigma_{\theta}$. In addition to describe the uncertainty structure of parameter changing, we define the generator matrix $G_{\theta} \in \mathbb{R}^{n \times nm}$ as $G_{\theta} = \left[G_{\theta}^{(1)}, G_{\theta}^{(2)}, \dots, G_{\theta}^{(nm)}\right]$, where each $G_{\theta}^{(l)} \in \mathbb{R}^{n \times m}$ contains exactly one nonzero element equal to σ_{θ} , and all other entries are zero. Specifically, for each $l \in \{1, \dots, nm\}$, there exists a unique index pair (i, j) such that $(G_{\theta}^{(l)})_{i,j} = \sigma_{\theta}$, and $(G_{\theta}^{(l)})_{i_{l},j_{l}} = 0 \quad \forall \ (i_{l}, j_{l}) \neq (i, j)$. Then, the uncertainaty satisfies $\delta_{\theta_{k-1}} \in \langle 0, G_{\theta} \rangle$ for all k.

To estimate unknown parameters, we need the regressor matrix sequence to be sufficiently rich. Consequently, we assume the sequence $\{\varphi_k\}_{k=0}^{\infty}$ to be Persistently Exciting (PE) as defined in Definition 4.

Definition 4 (Persistently Exciting) The sequence $\{\varphi_k\}, k=0,1,2,...\infty$ is said to be PE if for some constant $S \geq 1$ and all j there exist positive constants α and β such that

$$0 < \alpha I_n \le \sum_{i=j}^{j+S-1} \varphi_i^{\top} \varphi_i \le \beta I_n < \infty.$$

Suppose the unknown true parameter matrix $\theta_{\mathrm{tr},k}$ must be conservatively estimated. Under PE condition on φ_k , the goal is to recursively compute a zonotope that overapproximates $\theta_{\mathrm{tr},k}$, such that

$$\theta_{\text{tr},k+1} \in \langle C_{k+1}, G_{\lambda_{k+1}} \rangle.$$
 (8)

where $C_{k+1} \in \mathbb{R}^{n \times m}$ is the center matrix, and $G_{\lambda_{k+1}}^{(i)} \in \mathbb{R}^{n \times m}$ is the *i*-th generator matrix for all $i = 1, \ldots, r$. The number of generators is denoted by $r \in \mathbb{N}_{>0}$. We define recursive estimator as follows

$$\theta_{k+1} = \theta_k + K_k (y_k - \varphi_k \theta_k - v_k). \tag{9}$$

such that $\theta_{k+1} \in \mathbb{R}^{n \times m}$ provides the updated estimate when a new measurement is added. By multiplying the correction gain $K_k \in \mathbb{R}^{n \times p}$ with the output estimation error $(y_k - \varphi_k \theta_k - v_k)$, the previous estimation θ_k is updated at each sample time k. Inspired by [3,18] of using set-propagation for designing diffusion observer and [26] to parameter estimation of a regression model, we propose the Proposition 1 to estimate a recursive set of unknown time-varying parameters in (7).

Proposition 1 (Recursive Zonotopic Estimator)

Consider the deterministic system in (7) with initial parameter set $\theta_0 \in \langle C_0, G_0 \rangle$ where $\operatorname{rank}(G_0) \geq nm$ and choosing K_k such that $\operatorname{rank}(I_n - K_k \varphi_k) = n$. The EFZRLS is updated recursively by (8) for all $k = 0, 1, \ldots \infty$ as follows

$$C_{k+1} = (I_n - K_k \varphi_k) C_k + K_k y_k, \tag{10}$$

$$G_{\lambda_{k+1}}^{(i)} = \lambda^{-\frac{1}{2}} (I_n - K_k \varphi_k) G_{\lambda_k}^{(i)},$$

$$G_{\lambda_{k+1}} = \left[G_{\lambda_{k+1}}^{(1)}, \dots, G_{\lambda_{k+1}}^{(N_G)}, G_{v_k} \right], \ \forall i = 0, 1, \dots, N_G.$$
(11b)

where $N_G \in \mathbb{N}$ is the number of generators, $\lambda \in (0,1]$ is the forgetting factor, and for all $p, m \in \mathbb{N}_{>0}$

$$G_{v_k} = \left[-K_k Q_v^{(1)}, -K_k Q_v^{(2)}, \dots, -K_k Q_v^{(pm)} \right].$$

Proof. We can rewrite the (9) as follows

$$\theta_{k+1} = (I_n - K_k \varphi_k)\theta_k + K_k y_k + K_k v_k. \tag{12}$$

Considering (7), the estimated parameter contained the true value and its deviation as θ_k . Given $\theta_k \in \langle C_k, G_k \rangle$ and according to (12)

$$\theta_{k+1} \in ((I_n - K_k \varphi_k) \odot \langle C_k, G_k \rangle)$$

$$\oplus ((I_n - K_k \varphi_k) \odot \langle \mathbf{0}, G_{\theta} \rangle)$$

$$\oplus (K_k \odot \langle y_k, \mathbf{0} \rangle) \oplus (-K_k \odot \langle \mathbf{0}, G_v \rangle).$$

By applying the linear map and Minkowski sum properties of matrix zonotopes, the above equation leads to the center update equation (10) and the generator update equation as follows

$$G_{k+1} = \left| (I_n - K_k \varphi_k) G_k, (I_n - K_k \varphi_k) G_\theta, G_{v_K} \right|.$$

Given the matrix zonotope Definition 3, we can write

$$\langle C_{k+1}, G_{k+1} \rangle = C_{k+1} + \sum_{j=1}^{N_G} \alpha_j (I_n - K_k \varphi_k) G_k^{(j)}$$
(13)
+
$$\sum_{i=1}^{nm} \beta_i (I_n - K_k \varphi_k) G_{\theta}^{(i)} + \sum_{j=1}^{mp} \xi_j G_{v_k}^{(j)},$$

$$|\alpha_j| \le 1, \ |\beta_i| \le 1, \ |\xi_i| \le 1.$$

With regards to the definition of the proposition, we have $G_{\theta}^{(i)} \in \operatorname{span}(G_k)$ which means $G_{\theta}^{(i)} = \sum_{j=1}^{N_G} \eta_{ij} G_k^{(j)}$

for all i = 1, 2, ..., nm and $\eta_{ij} > 0$. Then substituting it into (13), we have

$$\langle C_{k+1}, G_{k+1} \rangle = C_{k+1} + \sum_{j=1}^{mp} \xi_j G_{v_k}^{(j)}$$

$$+ \sum_{j=1}^{N_G} (I_n - K_k \varphi_k) (\alpha_j + \beta_j \sum_{j=1}^{nm} \eta_{ij}) G_k^{(j)},$$

$$|\alpha_j| \le 1, \ |\beta_i| \le 1, \ |\xi_j| \le 1.$$
(14)

If $\lambda^{\frac{-1}{2}}\gamma_j=\alpha_j+\beta_j\sum_{j=1}^{nm}\eta_{ij}$ for any $\lambda>0$. Then, using the triangle inequality, we have

$$\lambda^{\frac{-1}{2}} |\gamma_j| \le |\alpha_j| + |\beta_j| \sum_{j=1}^{nm} |\eta_{ij}|.$$

Using the fact that $|\alpha_j| \leq 1$ and $|\beta_j| \leq 1$, we have

$$|\gamma_j| \le (1 + \sum_{j=1}^{nm} |\eta_{ij}|) \lambda^{\frac{1}{2}}.$$

Then by choosing

$$0 < \lambda \le \frac{1}{(1 + \sum_{j=1}^{nm} |\eta_{ij}|)^2} \le 1. \tag{15}$$

leads to $|\gamma_j| \leq 1$. Therefore we can rewrite (13) as follows

$$\langle C_{k+1}, G_{\lambda_{k+1}} \rangle = C_{k+1} + \sum_{j=1}^{N_G} \lambda^{\frac{-1}{2}} (I_n - K_k \varphi_k) \gamma_j G_k^{(j)}$$

$$+ \sum_{j=1}^{mp} \xi_j G_{v_k}^{(j)}, \quad |\gamma_j| \le 1.$$
(16)

Using matrix zonotope Definition 3 and zonotope containment condition in [25, Theorem 3] leads to $\langle C_{k+1}, G_{k+1} \rangle \subseteq \langle C_{k+1}, G_{\lambda_{k+1}} \rangle$. Therefore, the generator update (16) is written as follows

$$G_{\lambda_{k+1}}^{(i)} = [\lambda^{\frac{-1}{2}} (I_n - K_k \varphi_k) G_{\lambda_k}^{(i)}, G_{v_k}].$$

While in the Proposition 1, we proposed the center and generators update equations to find a set containing true parameters $\theta_{\text{tr},k}$, the correction gain K_k is still unknown in (10) and (11a). Theorem 1 proposes an optimal gain that minimizes the estimated parameter set.

Theorem 1 (Optimal Estimator Gain) Given the generator matrix $G_{\lambda_{k+1}}$ as in (11), and under PE

condition of φ_k as specified in Definition 4, the optimal estimator-gain $K_k^{\star} = \arg\min_{K_k} \|G_{\lambda_{k+1}}\|_{F,W}^2$ that minimizes the trace of the covariance matrix $J_W = Tr(WP_{k+1})$, is computed by

$$K_k^{\star} = P_k \varphi_k^{\top} \Lambda^{-1}, \tag{17}$$

$$P_{k+1} = \lambda^{-1} (I_n - K_k^{\star} \varphi_k) P_k \tag{18}$$

where

$$\Lambda = \varphi_k P_k \varphi_k^{\perp} + \lambda Q,$$

with $Q = m\sigma_v^2 \mathbf{1} \in \mathbb{R}^{p \times p}$, and K_k^{\star} is independent of the weighting matrix $W \in \mathbb{R}^{n \times n}$.

Proof. Given the covariance Definition 5, we substitute (11) into the covariance matrix $P_{k+1} = G_{\lambda_{k+1}} G_{\lambda_{k+1}}^{\dagger}$ as follows

$$P_{k+1} = \lambda^{-1} (I_n - K_k \varphi_k) G_{\lambda_k} G_{\lambda_k}^{\top} (I_n - K_k \varphi_k)^{\top}$$
 (19)
+ $K_k \sum_{j=1}^{pm} Q_v^{(j)} Q_v^{(j)^{\top}} K_k^{\top}.$

Given $P_k = G_{\lambda_k} G_{\lambda_k}^{\top}$ and define $Q = \sum_{j=1}^{pm} Q_v^{(j)} Q_v^{(j)}$, (19) is written as follows

$$P_{k+1} = \lambda^{-1} (I_n - K_k \varphi_k) P_k (I_n - K_k \varphi_k)^{\top} + K_k Q K_k^{\top}.$$
(20)

With replacing P_{k+1} in J_W , it derives in

$$J_W = \operatorname{Tr}(\lambda^{-1}W(I_n - K_k\varphi_k)P_k(I_n - K_k\varphi_k)^\top + WK_kQK_k^\top).$$
(21)

Afterwards, performing appropriate manipulations, and thanks to the symmetry of P_k and W is obtained

$$J_W = \operatorname{Tr}(\lambda^{-1}W P_k - 2\lambda^{-1}W P_k \varphi_k^{\top} K_k^{\top} + \lambda^{-1}W K_k \Lambda K_k^{\top}), \tag{22}$$

with $\Lambda = \varphi_k P_k \varphi_k^{\top} + \lambda Q$. Finally, determining $\frac{\partial J_W^{\vee}}{\partial K_k} = 0$ through Property 1 and solving for K_k it results in $2\lambda^{-1}WP_k\varphi_k^{\top} = 2\lambda^{-1}WK_k\Lambda$ which yields to (17)

Then, substituting (17) into (20) and do some manipulations, we have

$$\begin{split} P_{k+1} &= \lambda^{-1} P_k - \lambda^{-1} P_k \varphi_k^\top \Lambda^{-1} \varphi_k P_k \\ &+ \lambda^{-1} P_k \varphi_k^\top \Lambda^{-1} \Lambda \Lambda^{-1} \varphi_k P_k^\top. \end{split}$$

Then, applying $\Lambda^{-1}\Lambda = I_p$, using the symmetric property of P_k and making some manipulations, the following expression is achieved:

$$P_{k+1} = \lambda^{-1} (I_n - P_k \varphi_k^{\top} \Lambda^{-1} \varphi_k) P_k.$$
 (23)

Then replacing $P_k \varphi_k^{\top} \Lambda^{-1}$ by K_k^{\star} it yields to (18).

Through the Corollary 1, the optimal correction gain (17) satisfies the definition conditions in Proposition 1.

Corollary 1 Given PE condition of φ_k as specified in Definition (4), using correction gain (17) and covariance update matrix (18), then $\operatorname{rank}(I_n - K_k \varphi_k) = n$.

Proof. Given the covariance update matrix 18, we can write $(I_n - K_k^* \varphi_k) = \lambda P_{k+1} P_k^{-1}$. Since the P_k is a positive definite matrix and it is of full rank under the PE condition, $(I_n - K_k^{\star} \varphi_k)$ is of full rank.

Remark 1 The λ in (11a) scales up the generators to cover the uncertainties and serves as the forgetting factor that exponentially decreases the influence of past data in the covariance update in (18). When there is no uncertainty in parameter $G_{\theta} = 0$, choosing $\lambda = 1$, which means the generators are updated only by the noise effect, and all data have the same weight in updating the covariance matrix. However, when $G_{\theta} \neq \mathbf{0}$, the forgetting factor λ is chosen less than 1, which means the generators should be scaled up to cover these uncertainties in computing the boundary of the parameter set, prioritizing recent measurements in computing the covariance matrix.

Remark 2 The convergence of the updating center (10) using the optimal gain (17) has been studied in [10]. It has been established that under the PE condition of φ_k , the parameter estimation error $\tilde{\theta}_k = \theta_k - \theta_{tr,k}$ globally uniformly and ultimately converges to a bounded set.

Additionally, Proposition 1 proved that for a sufficiently small λ , the updated zonotope in (8) can cover all deviations in time-varying parameters. However, if λ is chosen too small, the covariance matrix update in (18) may cause the covariance matrix to diverge. Therefore, it is recommended in the literature to choose λ in the range of 0.95 to 1 for normal recursive least squares with forgetting factor. With this selection, a dedicated method is required to estimate the system parameters, which are shown to be slow time-varying (e.g., with small δ_{θ}).

Remark 3 The most straightforward way to initialize the center in (10) is to initialize the $C_0 = 0$ and initialize the initial covariance matrix in (18) with a large enough identity matrix $P_0 = \tau I_n$ for $\tau \in \mathbb{R}_{>0}$ which means the initial center C_0 is not so accurate. On the other hand, in case of availability of collecting offline data, first collect at least N_0 snapshots (more precisely, enough snapshots so that $\Phi = [\varphi_0, \varphi_1, ..., \varphi_{N_0}]^{\top}$ has rank n+m and $Y = [y_1, y_2, ..., y_{N_0}]^{\top}$), and then compute the initial $C_0 = \Phi^{\dagger} Y$ and $P_0 = \Phi^{\dagger}$ or $P_0 = G_0 G_0^{\top}$.

Remark 4 Without loss of generality, to estimate the unknown parameters in (7), we can use the vectorpermutation form of this regression as

$$\bar{y}_k = \bar{\varphi}_k \bar{\theta}_{tr,k} + \bar{v}_k,
\bar{\theta}_{tr,k} = \bar{\theta}_{tr,k-1} + \delta \bar{\theta}_{k-1}.$$
(24)

Where $\bar{y}_k = vec(y_k)$, $\bar{\theta}_{tr,k} = vec(\bar{\theta}_{tr,k})$, $\bar{v}_k = vec(v_k)$, and $\bar{\varphi}_k = I_m \otimes \varphi_k$. Then all matrix zonotopes will be converted to the (vector) zonotope, and the center update equation in (10) will be a vector $\bar{C}_k = vec(C_k) \in \mathbb{R}^{nm}$ and following the dimension of covariance matrix in (18) increased to $\bar{P}_k \in \mathbb{R}^{(nm \times nm)}$. Given the study in [17] using vector-permutation makes more computational cost since the Kronecker product creates numerous zero terms in the regressor, increasing computational cost. In addition, using vector-permutation is beneficial when the noise distribution is different for each output, while in our proposed method, no distribution is considered for noise.

Remark 5 Using the generator updating equation in (11b) can result in a large generator matrix after several time steps, potentially causing computational issues. To reduce this, one can employ the reduced-order generator $\downarrow_q(G_{\lambda_{k+1}})$ as described in Proposition 2.

4 Data-Driven Reachability Analysis

In this section, we compute over-approximations of discrete-time LTV and nonlinear Lipschitz dynamic systems over a predefined time horizon, without assuming any prior knowledge of the underlying dynamic models. The available data are assumed to be corrupted by process noise.

4.1 Linear Time Varying Systems with Process Noise

Consider $f(x_k, u_k)$ in (1) is a discrete-time linear timevarying function. Then, the system (1) can be written as follows

$$x_{k+1} = A_{\text{tr},k} x_k + B_{\text{tr},k} u_k + w_k, A_{\text{tr},k} = A_{\text{tr},k-1} + \delta A_{k-1}, B_{\text{tr},k} = B_{\text{tr},k-1} + \delta B_{k-1}.$$
(25)

where $A_{\mathrm{tr},k} \in \mathbb{R}^{n_x \times n_x}$, and $B_{\mathrm{tr},k} \in \mathbb{R}^{n_x \times n_u}$ are the true unknown time-varying system model, and $\delta A_k \in \mathbb{R}^{n_x \times n_x}$, and $\delta B_k \in \mathbb{R}^{n_x \times n_u}$ are change in the system model which $\left\| \left[\delta A_k \ \delta B_k \right] \right\|_{max} \leq \sigma_{AB}$ and the noise bounded by $\|w_k\|_{max} \leq \sigma_w$ for $k = 0, 1, ..., \infty$.

We define the unknown bounded noise set with Assumption 3.

Assumption 3 (Noise Zonotope) The noise w_k is contained in a known zonotope for all $k \in \mathbb{Z}_{\geq 0}$, i.e., $w_k \in \mathcal{Z}_w = \langle c_{\mathcal{Z}_w}, G_{\mathcal{Z}_w} \rangle$, where $c_{\mathcal{Z}_w} \in \mathbb{R}^{n_{\mathcal{Z}_w}}$ is the center, $G_{\mathcal{Z}_w} \in \mathbb{R}^{n_{\mathcal{Z}_w} \times \gamma_{\mathcal{Z}_w}}$ is the generator matrix, and the zonotope \mathcal{Z}_w contains the origin.

To guarantee an over-approximation of the reachable sets for the unknown system in (25), we estimate a time-varying set of models $\begin{bmatrix} A_k & B_k \end{bmatrix}$ which contains the true system model. By applying the Proposition 1 and Theorem 1, we estimate this time-varying set of models. Let $\hat{\mathcal{R}}_k$ denotes the reachable set computed based on the noisy data using matrix zonotopes. We propose Algorithm 1 to compute $\hat{\mathcal{R}}_k$ as an over-approximation of the exact reachable set \mathcal{R}_k . The following theorem proves that $\hat{\mathcal{R}}_k \supseteq \mathcal{R}_k$.

Theorem 2 (Reachability Analysis of the LTV Systems) Given a minimum enough dense data on the set D with length N_D which the PE condition in Definition 4 is satisfied, and choosing a large enough G_{λ_0} such that $[A_{tr,0}, B_{tr,0}] \in \tilde{\mathcal{M}}_0 = \langle [A_0, B_0], G_{\lambda_0} \rangle$ where $[A_0, B_0]$ is initial model estimation, then the reachable sets computed in Algorithm 1 over-approximates the exact reachable sets, i.e., $\hat{\mathcal{R}}_k \supseteq \mathcal{R}_k$ for $k = 0, 1, \ldots N - 1$.

Proof. Considering $w_k \in \mathcal{Z}_w$, we find \mathcal{M}_k such that $[A_{\operatorname{tr},k}, B_{\operatorname{tr},k}] \in \mathcal{M}_k$. Therefore we show for all $\begin{bmatrix} x_k \\ u_k \end{bmatrix} \in \mathcal{F}$, we have

$$x_{k+1} \in \mathcal{M}_k \begin{bmatrix} x_k \\ u_k \end{bmatrix} \oplus \mathcal{Z}_w \oplus \mathcal{Z}_\epsilon.$$

for all $k=0,1,\ldots,N-1$. Since the variations in the system model in (25) are bounded by σ_{AB} , Assumption 2 guarantees that the perturbations satisfy $[\delta_{A_i},\delta_{B_i}]\in G_{AB}$. By Proposition 1 and Theorem 1, given $[A_{\mathrm{tr},0},B_{\mathrm{tr},0}]\in \tilde{\mathcal{M}}_0$, there exists a scalar $\lambda\in(0,1]$ such that $[A_{\mathrm{tr},i},B_{\mathrm{tr},i}]\in \tilde{\mathcal{M}}_i$ for all $i\geq 0$.

Since no new measured data is available when computing the reachable sets for steps k = 0, 1, ..., N-1, we assume that the maximum deviation in the system model is bounded by σ_{AB} . Applying Lemma 1, the zonotope

$$\mathcal{M}_k = \langle C_{\tilde{\mathcal{M}}_i}, [G_{\tilde{\mathcal{M}}_i}, G_{\hat{\mathcal{M}}_k}] \rangle$$

where

$$\hat{\mathcal{M}}_k = \langle \mathbf{0}, (k\sigma_{AB}E_1, k\sigma_{AB}E_2, \dots, k\sigma_{AB}E_{n_x n_{xx}}) \rangle$$

contains all possible system models at time step k.

To account for the spatial generalization error of the true system dynamics for any state-input pair $z_k = \begin{bmatrix} x_k^\top & u_k^\top \end{bmatrix}^\top \in \mathcal{F}$ at time step k. Since \mathcal{F} is compact, we assume that all points $z_j \in D_-$ for $j=1,2,...,N_D$ are dense in \mathcal{F} . In other words, there exists some $\delta \geq 0$ such

Algorithm 1 LTV-Reachability

Input: Input-state trajectory D_- , initial set \mathcal{X}_0 , process noise zonotope \mathcal{Z}_w and matrix zonotope G_{AB} , and input zonotope U_k , $\forall k = 0, \dots, N-1$

Output: Over approximate reachable sets $\hat{\mathcal{R}}_k, \forall k =$ $0, \ldots, N-1$

- 1: $\hat{\mathcal{R}}_0 = \mathcal{X}_0$,
- 2: Update the Center of model set, ⊳ Use 10
- 3: Update the generators of model set, ▶ Use 11
- 4: Update the identified model set $\tilde{\mathcal{M}}_i$, ⊳ Use 8
- 5: Update convenience matrix, ⊳ Use 18
- ⊳ Use 26, 6: Compute δ ,
- $\mathcal{Z}_{\epsilon} = \langle \mathbf{0}, I_{\mathcal{M}_{max}} \hat{\delta}/2, \dots, I_{\mathcal{M}_{max}} \hat{\delta}/2) \rangle,$ for $\hat{k} = 0: N-1$ do
- $\hat{\mathcal{M}}_{k} = \langle \mathbf{0}, [k\sigma_{AB}E_{1}, k\sigma_{AB}E_{2}, \dots, k\sigma_{AB}E_{n_{x}n_{u}}] \rangle, \\ \mathcal{M}_{k} = \langle C_{\tilde{\mathcal{M}}_{i}}, [G_{\tilde{\mathcal{M}}_{i}}, G_{\hat{\mathcal{M}}_{k}}] \rangle$
- $\hat{\mathcal{R}}_{k+1} = \mathcal{M}_k \; (\hat{\mathcal{R}}_k \times \mathcal{U}_k) + \mathcal{Z}_{\epsilon} + \mathcal{Z}_w.$
- 12: end for

that for every $z_k \in \mathcal{F}$, there exists a $z_i \in D_-$ such that

$$\|\mathcal{M}_k z_k - \mathcal{M}_k z_j\| \le I_{\mathcal{M}_{max}} \delta.$$

Where $I_{\mathcal{M}_k} = \operatorname{int}(\mathcal{M}_k)$ and $I_{\mathcal{M}_{max}} = \sup_k \| I_{\mathcal{M}_k} \|$. This spatial generalization error is captured by the zonotope [15]

$$\mathcal{Z}_{\epsilon} = \langle \mathbf{0}, I_{\mathcal{M}_{max}} \delta/2, \dots, I_{\mathcal{M}_{max}} \delta/2) \rangle.$$

By compactness of \mathcal{U}_k , \mathcal{R}_k for $k = 0, \dots, N$, the product space \mathcal{F} is compact. Therefore, for any $\delta > 0$, there exists a finite dataset $\{z_j\}_{j=1}^{N_D}$

$$\mathcal{F}$$
, where $z_j = \begin{bmatrix} (X_-)_{.,j} \\ (U_-)_{.,j} \end{bmatrix}$, that is δ -dense in \mathcal{F} , i.e., $\forall z \in \mathcal{F}, \ \exists z_j \in D_- \text{ such that } \|z_k - z_j\| \leq \delta$.

Remark 6 ([2]) Note that computing δ is non-trivial in practice. If we assume that the data is evenly spread out in the compact input set, then the following can be a good approximation of the δ

$$\hat{\delta} = \max_{z_i \in D_-} \min_{z_j \in D_-, j \neq i} ||z_i - z_j||. \tag{26}$$

The $\hat{\delta}$ may be estimated either from existing recorded data or from an updated sliding window D_{-} , which results in $a\ more\ conservative\ estimate.$

Remark 7 Practical approaches exist for estimating σ_{AB} in order to define $\hat{\mathcal{M}}_k$ in Algorithm (1). Moreover, the time-varying part of (25) can be expressed as

$$A_{tr,k+N} = A_{tr,k} + \sum_{j=0}^{N-1} \delta A_{k+j},$$

$$B_{tr,k+N} = B_{tr,k} + \sum_{j=0}^{N-1} \delta B_{k+j}.$$

Without loss of generality, for slowly time-varying systems over a short horizon of N steps, it can be assumed that $\sum_{j=0}^{N-1} \delta A_{k+j} \simeq 0$ and $\sum_{j=0}^{N-1} \delta B_{k+j} \simeq 0$. This approximation ensures that the over-approximation of the reachable sets remains valid. In fact, for most slowly varying linear systems, model variations become significant only over longer horizons, i.e., when $N \gg k$.

4.2 Lipschitz Nonlinear Systems

We consider a discrete-time Lipschitz nonlinear control system in (1). A local linearization of (1) is performed by a Taylor series expansion around the time varying

linearization point [2] as
$$z_k^* = \begin{bmatrix} x_k^* \\ u_k^* \end{bmatrix}$$

$$f(z_k) = f(z_k^{\star}) + \frac{\partial f(z)}{\partial z_k} \Big|_{z_k = z_k^{\star}} (z_k - z_k^{\star}) + \dots$$

The infinite Taylor series [7] can be represented by a first-order Taylor series and a Lagrange remainder term $L(z_k)$ [4, p.65], that depends on the model, as follows.

$$f(z_k) = f(z_k^*) + \frac{\partial f(z_k)}{\partial z_k} \Big|_{z_k = z_k^*} (z_k - z_k^*) + L(z_k).$$
 (27)

Since the model is assumed to be unknown, we aim to over-approximate L(z) from data. We rewrite (27) as follows

$$f(x_k, u_k) = f(x_k^*, u_k^*) - \tilde{A}_{tr} x_k^* - \tilde{B}_{tr} u_k^*$$

$$+ \tilde{A}_{tr} x_k + \tilde{B}_{tr} u_k + L(x_k, u_k)$$
(28)

$$\tilde{A}_{\mathrm{tr},k} = \frac{\partial f(x_k, u_k)}{\partial x_k} \Big|_{x = x_k^{\star}, u = u_k^{\star}}, \tilde{B}_{\mathrm{tr},k} = \frac{\partial f(x_k, u_k)}{\partial u_k} \Big|_{x_k = x_k^{\star}, u_k = u_k^{\star}}$$

By substituting $\Delta f_{\mathrm{tr},k}^{\star} = f(x_k^{\star}, u_k^{\star}) - \tilde{A}_{\mathrm{tr},k} x^{\star} - \tilde{B}_{\mathrm{tr},k} u_k^{\star}$ in (28), we obtain:

$$f(x_k, u_k) = \left[\Delta f_{\text{tr},k}^{\star} \ \tilde{A}_{\text{tr},k} \ \tilde{B}_{\text{tr},k} \right] \begin{bmatrix} \mathbf{1} \\ x_k \\ u_k \end{bmatrix} + L(x_k, u_k).$$
(29)

where $\Delta f_{\mathrm{tr},k}^{\star}$, $A_{\mathrm{tr},k}$ and $B_{\mathrm{tr},k}$ are unknown linearized system model which we use EF-ZRLS to estimate timevarying sets which contains these matrices. To overapproximate the remainder term $L(z_k)$ from data, we need to assume that f is Lipschitz continuous for all z in the reachable set \mathcal{F} . The continuity of nonlinear systems can be defined using the Lipschitz condition.

Assumption 4 It holds that $f: \mathcal{F} \to \mathbb{R}^{n_x}$ is Lipschitz continuous, i.e., that there is some $L^* \geq 0$ such that $||f(z) - f(z')||_2 \leq L^* ||z - z'||_2$ holds for all $z, z' \in \mathcal{F}$.

For data-driven methods applied to nonlinear systems, Lipschitz continuity is a common assumption (e.g., [21,22]). Our goal is to approximate the Lipschitz nonlinear system by a LTV system over discrete time steps. By invoking Proposition 1 and Theorem 1 and applying them to (29) without the Lipschitz term L(x,u), we obtain a time-varying set containing $\left[\Delta f_{\mathrm{tr},k}^{\star} \tilde{A}_{\mathrm{tr},k} \tilde{B}_{\mathrm{tr},k}\right]$. Then, by employing a sliding window of length N_D that stores the most recent N_D data points, we compute an over-approximation of $L(x_k,u_k)$. The following theorem proves the over-approximation of the reachable sets \mathcal{R}_k' out of Algorithm 2 for the exact reachable sets \mathcal{R}_k .

Theorem 3 (Reachability Analysis of Lipschitz Systems) Given sufficiently dense data on the set D of length N_D , satisfying the PE condition in Definition 4, and choosing a large enough zonotope radius G_{λ_0} such that the linearization model

$$\left[\Delta f_{\mathrm{tr},0}^{\star} \ \tilde{A}_{\mathrm{tr},0} \ \tilde{B}_{\mathrm{tr},0}\right] \in \tilde{\mathcal{M}}_{f,0} = \langle \left[\Delta f_{0}^{\star} \ \tilde{A}_{0} \ \tilde{B}_{0}\right], G_{\lambda_{0}} \rangle,$$

where $\left[\Delta f_0^{\star} \ \tilde{A}_0 \ \tilde{B}_0\right]$ is the initial model estimation, then the reachable sets computed in Algorithm 2 over-approximates the exact reachable sets, i.e., $\mathcal{R}_k \subseteq \mathcal{R}_k'$, $\forall k = 0, 1, \dots, N-1$.

Proof. We use the proof of [2, Theorem 5] and extend it to this work. We know from (29) that

$$f(z_k) = M_{\text{tr}_k} \begin{bmatrix} \mathbf{1} \\ z_k \end{bmatrix} + L(z_k),$$

where $M_{\text{tr}_k} = \left[\Delta f_{\text{tr},k}^{\star} \ \tilde{A}_{\text{tr},k} \ \tilde{B}_{\text{tr},k} \right]$. Hence, we need to show that $\mathcal{Z}_L + \mathcal{Z}_{\bar{\epsilon}}$ over-approximates the modeling mismatch and the term L(z)

$$f(z_k) \in \mathcal{M}_k \begin{bmatrix} \mathbf{1} \\ z_k \end{bmatrix} \oplus \mathcal{Z}_L \oplus \mathcal{Z}_{\bar{\epsilon}}.$$

where $M_{\text{tr}_k} \in \mathcal{M}_k$ for all $z_k \in \mathcal{F}$. Since $f(z_k)$ is Lipschitz, we can conclude that $M_{\text{tr},k}$ is bounded. Therefore, we can write

$$M_{\text{tr},i} = M_{\text{tr},i-1} + \delta M_{i-1}.$$

where δM_{i-1} denotes the variation at step i-1. If we can show that $\|\delta M_{i-1}\|_{max} \leq \sigma_M$, then, similar to Assumption 2, we can define a set G_M that over-approximates the system model changes, i.e., $\delta M_{i-1} \in G_M$.

Given that $w_i \in \mathcal{Z}_w$ and by invoking Proposition 1 and Theorem 1, we can find a $\lambda \in (0,1]$ such that

$$M_{\mathrm{tr},i} \in \tilde{\mathcal{M}}_{f,i}, \quad \forall i \geq 0.$$

while $M_{\text{tr},0} \in \tilde{\mathcal{M}}_{f,0}$. By defining $\bar{i} = i - N_D + j$ To over-approximate $L(z_j)$, we know that for all $z_j \in D_-$, we can write

$$(X_{+})_{\cdot,j} - (W_{-})_{\cdot,j} = (C_{\tilde{\mathcal{M}}_{f,\bar{i}}} + \Delta M_{\bar{i}}) \begin{bmatrix} \mathbf{1} \\ z_{j} \end{bmatrix} + L(z_{j}), \forall j = 1, 2, \dots, N_{D}.$$

where $\Delta M_{\bar{i}} = M_{\mathrm{tr},\bar{i}} - C_{\tilde{\mathcal{M}}_{f,\bar{i}}}$. This implies

$$(X_{+})_{\cdot,j} - C_{\tilde{\mathcal{M}}_{f,\bar{i}}} \begin{bmatrix} \mathbf{1} \\ z_{j} \end{bmatrix} \in \Delta \tilde{\mathcal{M}}_{f,\bar{i}} + L(z_{j})$$

$$+ \mathcal{Z}_{\cdots}.$$
(30)

where $\Delta \tilde{\mathcal{M}}_{f,\bar{i}} = \langle \mathbf{0}, G_{\tilde{\mathcal{M}}_{f,\bar{i}}} \rangle$. Next, we aim to find one zonotope \mathcal{Z}_L that over-approximates $L(z_j)$ for all the data points, i.e., $\forall z_j \in D_-$

$$(X_+)_{\cdot,j} - C_{\tilde{\mathcal{M}}_{f,\bar{i}}} \begin{bmatrix} \mathbf{1} \\ z_j \end{bmatrix} \in \Delta \tilde{\mathcal{M}}_{f,\bar{i}} \oplus \mathcal{Z}_L \oplus \mathcal{Z}_w.$$

This can be found by the upper bound (\bar{l} in line 6) and lower bound (\underline{l} in line 7) from data, and hence \mathcal{Z}_L in line 8 in Algorithm 2. Thus, we can over-approximate the model mismatch and the nonlinearity term for all data points $z_i \in D_-$, by

$$f(z_j) \in \tilde{\mathcal{M}}_{f,\bar{i}} \begin{bmatrix} \mathbf{1} \\ z_j \end{bmatrix} + \mathcal{Z}_L.$$

Given the covering radius δ of our system together with Assumption 4, we know that for every $z \in \mathcal{F}$, there exists a $z_j \in D_-$ such that $||f(z_k) - f(z_j)|| \le L^*||z_k - z_j|| \le L^*\delta$. This yields $\mathcal{Z}_{\bar{\epsilon}} = \langle \mathbf{0}, \operatorname{diag}(L^{\star^{(1)}}\delta/2, \ldots, L^{\star^{(n_x)}}\delta/2) \rangle$.

To complete the proof, similar to Theorem 2, we need to use $\hat{\mathcal{M}}_k$ to overapproximate the reachable sets when no new measurement is available by using Lemma 1. Then using

$$\mathcal{M}_{f,k} = \langle C_{\tilde{\mathcal{M}}_{f,i}}, [G_{\tilde{\mathcal{M}}_{f,i}}, G_{\hat{\mathcal{M}}_k}] \rangle$$

Algorithm 2 Nonlinear Lipschitz-Reachability

Input: Input-state trajectory D, initial set \mathcal{X}_0 , process noise zonotope \mathcal{Z}_w and matrix zonotope G_M , and input

Output: Over approximate reachable sets \mathcal{R}'_k , $\forall k =$ $0,\ldots,N-1$

- 1: $\mathcal{R}_0' = \mathcal{X}_0$, 2: Update the Center of model set, ⊳ Use 10
- 3: Update the generators of model set,
- 4: Update the identified model set, $\tilde{\mathcal{M}}_{f,i}$ ⊳ Use 8
- 5: Update Convenience Matrix,

⊳ Use 18

6:
$$\bar{l} = \max_{j} \left((X_{+})_{.,j} - C_{\tilde{\mathcal{M}}_{f,\bar{i}}} \begin{bmatrix} \mathbf{1} \\ (X_{-})_{.,j} \\ (U_{-})_{.,j} \end{bmatrix} \right)$$
7: $\underline{l} = \min_{j} \left((X_{+})_{.,j} - C_{\tilde{\mathcal{M}}_{f,\bar{i}}} \begin{bmatrix} \mathbf{1} \\ (X_{-})_{.,j} \\ (U_{-})_{.,j} \end{bmatrix} \right)$

- 8: $\mathcal{Z}_{L} = \operatorname{zonotope}(\underline{l}, \overline{l}),$ 9: Compute \hat{L}^{\star} and $\hat{\delta},$ \triangleright Use 31 and 26 10: $\mathcal{Z}_{\bar{\epsilon}} = \langle \mathbf{0}, \operatorname{diag}(\hat{L}^{\star^{(1)}} \hat{\delta}/2, \dots, \hat{L}^{\star^{(n_{x})}} \hat{\delta}/2) \rangle,$ 11: for k = 0 : N 1 do

- $\hat{\mathcal{M}}_{k} = \langle \mathbf{0}, [k\sigma_{AB}E_{1}, k\sigma_{AB}E_{2}, \dots, k\sigma_{AB}E_{n_{x}n_{u}}] \rangle, \\
 \mathcal{M}_{f,k} = \langle C_{\tilde{\mathcal{M}}_{f,i}}, [G_{\tilde{\mathcal{M}}_{f,i}}, G_{\hat{\mathcal{M}}_{k}}] \rangle,$
- $\mathcal{R}'_{k+1} = \mathcal{M}_{f,k} \left(1 \times \mathcal{R}'_k \times \mathcal{U}_k \right) + \mathcal{Z}_L + \mathcal{Z}_{\bar{\epsilon}} + \mathcal{Z}_w.$
- 15: end for

yields

$$f(z_k) \in \mathcal{M}_{f,k} \begin{bmatrix} \mathbf{1} \\ z_k \end{bmatrix} + \mathcal{Z}_L + \mathcal{Z}_{\overline{\epsilon}}.$$

For an infinite amount of data, i.e., $\delta \to 0$, we can see that $\mathcal{Z}_{\bar{\epsilon}} \to 0$, i.e., the formal \mathcal{Z}_L then fully captures the modeling mismatch and the Lagrange remainder.

Remark 8 ([2]) Note that determining L^* is nontrivial in practic. If we assume that the data is evenly spread out in the compact input set of f, then the following can be a good approximation of the upper bound on L^* for each dimension o:

$$\hat{L}^{\star^{(o)}} = \max_{z_i, z_j \in D_-, i \neq j} \frac{\|f^{(o)}(z_i) - f^{(o)}(z_j)\|}{\|z_i - z_j\|}.$$
 (31)

Computing the Lipschitz constant for each dimension decreases the conservatism, especially when the data has a different scale for each dimension. Furthermore, note that from the proof of the above theorem, we see that for every reachability step, k = 1, ..., N, local information on L^* in the set $\mathcal{R}_k \times \mathcal{U}_k$ can be used. Since in each time step a new sample date is added to D_{-} , the L^{\star} approximation is updated.

Remark 9 Initially, when the dataset D is sparse, the over-approximation guarantees are not theoretically ensured. However, as data accumulates and D becomes denser in \mathcal{F} , the reachability results become provably con-

Evaluation

In this section, we apply the computational approaches to over-approximate the exact reachable sets from data. We collect online data, where the system model is unknown. We compare results with the method proposed in [2] to compute the reachable sets, which uses a batch Least Squares (LS) method to estimate a set of models given a rich enough offline data set to estimate models.

5.1 Example 1: A LTV System

To demonstrate the usefulness of the presented approach, we consider the reachability analysis of a fivedimensional system $A_k \in \mathbb{R}^{5 \times 5}$ and $B_k \in \mathbb{R}^5$ with a single input, which is a discretization of the system used in [4, p.39] with sampling time 0.1 sec. The initial set is chosen to be $\mathcal{X}_0 = \langle \mathbf{1}, 0.1I_5 \rangle$. The input set is $\mathcal{U}_k = \langle 10, 2.25 \rangle$. We consider computing the reachable set when there is a random noise sampled from the zonotope $\mathcal{Z}_w = \langle \mathbf{0}, 0.005 I_5 \rangle$.

To evaluate the Algorithm 1, we simulate the following senarious,

- (1) A trajectory of length $N_D = 30$ is used as input data on D, which is rich enough to estimate parameters with the LS method. There is no change in parameters (e.g. $\sigma_{AB} = 0$).
- (2) A trajectory of length $N_D = 30$ is used as input data on D, which is rich enough to estimate parameters. The parameters change over time-steps as $\delta A_k =$ $-0.0001\mathbf{1}_{5\times 5}$ and $\delta B_k = 0.0003\mathbf{1}_{5\times 5}$ and $\sigma_{AB} =$

For both scenarios, we consider the initial center $C_0 = \mathbf{0}$ and the initial convenience matrix $P_0 = 10^7 I_5$, and the initial generator $G_0 = 1.5E_i$ for $i = 1, 2, ..., n_x n_u$ where E_i are the standard basis matrices with a single non-zero element of 1.

The reachable set computed by Algorithm 1 converges, after a number of iterations, to the reachable set obtained using the LS method [2, Algorithm 1], as illustrated in Figure 2a. In fact, the model set estimated by the proposed ZRLS (equivalent to EF-ZRLS with $\lambda = 1$) eventually converges to the optimal model set computed

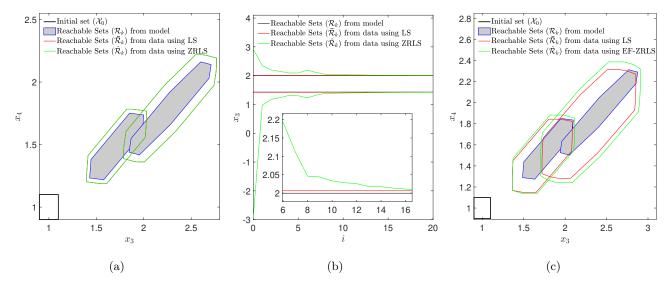


Fig. 2. The projection of the reachable sets of the LTV system in (25) computed via Algorithm 1 ($\hat{\mathcal{R}}_k$), the LS method ($\bar{\mathcal{R}}_k$) computed in [2, Algorithm 1], from noisy input-state data. There is no change in system matrices (e.g. $\sigma_{AB} = 0$)in (a) and (b), and $\sigma_{AB} = 0.0003$ in (c).

by the LS method, which relies on an offline batch of data stored in D. To ensure an over-approximation of the reachable sets at the initial step, it is necessary to choose a sufficiently large initial generator so that the true system model is guaranteed to lie within the initial model set. Once this condition is satisfied, the estimated model set converges to the optimal set. The convergence rate is affected by the choice of the initial covariance matrix, the generator matrices, and the lower bound of the PE condition specified in Definition 4. Figure 2b illustrates the convergence of the over-approximated x_2 trajectory over time steps.

As shown in Figure 2c, when the system model changes over time steps, the reachable sets computed by the LS method fail to over-approximate the exact reachable sets, while the reachable sets computed by Algorithm 1 successfully over-approximate exact reachable sets. The EF-ZRLS with $\lambda=0.92$, tries to minimize the covariance matrix affected by recent data, while the LS method considers the same weight of estimation error for all data on D. Decreasing λ leads to having more conservative reachable states since the estimated model set is enlarged.

5.2 Example 2: Lipschitz Nonlinear Systems

We apply the proposed data-driven reachability analysis to a continuous stirred tank reactor (CSTR) simulation model [9]. The initial set is a zonotope $\mathcal{X}_0 = \langle [1.35, 10.9]^{\top}, \operatorname{diag}([0.005, 0.3]) \rangle$. The input set $\mathcal{U}_k = \langle [1.1, -1.3]^{\top}, \operatorname{diag}([0.1, 0.2]) \rangle$ and the noise set $\mathcal{Z}_w = \langle 0, 0.003I_2 \rangle$. We define a set \bar{D} that contains 35 trajectories with length $N_{\bar{D}}$ collected offline to be used in the LS method in [2] to compute reachability sets. The length of the moving window D is $N_D = 5$, which

contains the latest data. The $\lambda=0.96$ is selected in EF-ZRLS, and the initial covariance matrix, initial set, and generators are the same as example 1. In addition, model variations are assumed to be negligible in computing the four short horizons N=4.

In Figure 3a, the reachable sets computed by Algorithm 2 are shown over four steps ahead. As illustrated, these reachable sets are less conservative than those obtained using the LS method proposed in [2], since the Lagrangian remainder that captures nonlinearities is computed globally over the entire dataset \bar{D} , and all data in \bar{D} are used to estimate the model for system linearization around the current point. In contrast, Algorithm 2 computes reachable sets in a more local manner: the linearized system model is updated recursively based on the latest data, and the set \mathcal{Z}_L , which over-approximates the Lagrangian remainder, is constructed from a moving window of data D. Increasing the length of stored data in D results in more conservative reachable sets, while also enabling the capture of nonlinearities over a longer prediction horizon N. It is worth noting that the data corresponding to the prediction horizon are already contained in D.

5.3 Example 3: Autonomous Vehicle

We utilized the JetRacer ROS AI Kit, an autonomous racing robot with an NVIDIA Jetson Nano Developer Kit (4GB RAM) and Raspberry Pi RP2040 dual-core microcontroller. The inputs to the vehicle are the linear and angular velocity, and the outputs are the position and heading of the vehicle.

We consider process noise bounded by the zonotope $\mathcal{Z}_w = \langle \mathbf{0}_{3\times 3}, 0.0005I_3 \rangle$. The forgetting factor $\lambda = 0.99$

with horizon step N=3 and $N_D=5$ is considered for implementing Algorithm 2. The initial covariance matrix is set $P_0=100I_3$, and the initial generator and center are the same as the previous examples. We consider two experimental scenarios. In the first scenario, the car and its environment remain unchanged over time. In the second one, a $2 \ kg$ payload is added on top of the car.

The over-approximated reachable sets computed using Algorithm 2 are illustrated in Figure 3b. The larger sets observed at the beginning of the experiment result from the use of a large initial covariance matrix and initial generator, reflecting the uncertain initial value of the center of the model set due to the absence of the true initial model. In addition, our proposed method is less conservative compared to the LS method in [2], while we do not need to have pre-collected data for identifying model.

In the second scenario, a payload is added on top of the car to evaluate the performance of our method under model changes. As shown in Figure 3c, the proposed method remains robust to model variations, whereas the LS method fails to over-approximate the reachable sets.

6 Conclusion

We consider the problem of computing the reachable set using noisy data for two classes of time-discrete systems without requiring a mathematical model of the system or any pre-collected data. To do this, we introduced EFZRLS, which is an optimal set-based recursive estimation method to estimate a time-varying set of models given noisy input-state data for the multi-output regression models. Then, we over-approximated the reachable sets of LTV and Lipschitz nonlinear systems, which are a superset of the exact reachable sets.

In future works, we are going to extend the proposed EF-ZRLS to use adaptive λ in [10] to decrease the conservatism of the estimating set of models.

Appendix A: Definitions and Proposition

Definition 5 (Matrix Zonotope Covariance [26]) Given the matrix zonotope of the Definition 3, the convenience of a matrix zonotope is defined as

$$\mathrm{cov}(\mathcal{M}) = \sum_{j=1}^{\gamma_{\mathcal{M}}} \tilde{G}_{\mathcal{M}}^{(j)} \tilde{G}_{\mathcal{M}}^{(j)^{\top}} = \tilde{G}_{\mathcal{M}} \tilde{G}_{\mathcal{M}}^{\top}.$$

Definition 6 (Interval Hull [26]) Given a zonotope of Definition 2, the interval hull $b(G_{\mathcal{Z}}) \subset \mathbb{R}^{n_x \times n_x}$ is defined as the smallest axis-aligned box containing \mathcal{Z} , i.e.,

$$\mathcal{Z} \subseteq \langle c_{\mathcal{Z}}, b(G_{\mathcal{Z}}) \rangle.$$

where $b(G_{\mathcal{Z}})$ is a diagonal matrix with entries

$$b(G_{\mathcal{Z}})_{ii} = \sum_{j=1}^{\gamma_{\mathcal{Z}}} \left| \left(G_{\mathcal{Z}} \right)_{ij} \right|.$$

Definition 7 (Zonotope Reduction Operator [26])Given a zonotope of Definition 2, the reduction operator

$$\downarrow_q: \mathbb{R}^{n_{\mathcal{Z}} \times \gamma_{\mathcal{Z}}} \to \mathbb{R}^{n_x \times q}$$

with $q < \gamma_{\mathcal{Z}}$, produces a reduced generator matrix such that the resulting reduced zonotope satisfies

$$\langle c_{\mathcal{Z}}, G_{\mathcal{Z}} \rangle \subseteq \langle c_{\mathcal{Z}}, \downarrow_q (G_{\mathcal{Z}}) \rangle.$$

Proposition 2 (Matrix Zonotope Reduction) Let a matrix zonotope \mathcal{M} of Definition 3 and the zonotope \mathcal{Z} of Definition 2 to be the corresponding equivalent vector zonotope, where $G_{\mathcal{Z}} = \left[\operatorname{vec}(G_{\mathcal{M}}^{(1)}), \dots, \operatorname{vec}(G_{\mathcal{M}}^{\gamma_{\mathcal{M}}}) \right]$. If \downarrow_q denotes the zonotope order-reduction operator from Definition 7, then a reduced generator matrix for \mathcal{M} is given by

$$\downarrow_q (\tilde{G}_{\mathcal{M}}) = \operatorname{unvec}(\downarrow_q (G_{\mathcal{Z}})),$$

and the resulting reduced matrix zonotope satisfies

$$\mathcal{M} \subseteq \langle C_{\mathcal{M}}, \downarrow_q (\tilde{G}_{\mathcal{M}}) \rangle.$$

Proof. Let $\mathcal{Z} = \langle c_{\mathcal{Z}}, G_{\mathcal{Z}} \rangle$ be the vector zonotope equivalent to \mathcal{M} . From Definition 7, we know that applying the reduction operator to the generator matrix $G_{\mathcal{Z}}$ results in the containment relation:

$$\langle c_{\mathcal{Z}}, G_{\mathcal{Z}} \rangle \subseteq \langle c_{\mathcal{Z}}, \downarrow_q (G_{\mathcal{Z}}) \rangle.$$

We can apply the linear and bijective map $unvec(\cdot)$ to both sides, which preserves the containment relation:

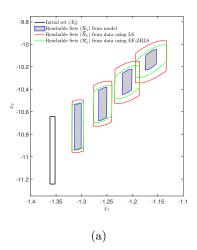
$$\operatorname{unvec}(\langle c_{\mathcal{Z}}, G_{\mathcal{Z}} \rangle) \subseteq \operatorname{unvec}(\langle c_{\mathcal{Z}}, \downarrow_q (G_{\mathcal{Z}}) \rangle).$$

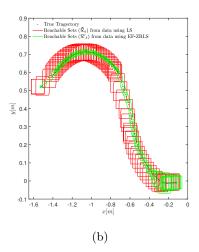
By the properties of the unvec(.) operator and the definition of a matrix zonotope, we have $\mathcal{M} = \text{unvec}(\langle c_z, G_z \rangle)$ and

$$\langle C_{\mathcal{M}}, \operatorname{unvec}(\downarrow_q (G_{\mathcal{Z}})) \rangle = \operatorname{unvec}(\langle c_{\mathcal{Z}}, \downarrow_q (G_{\mathcal{Z}}) \rangle).$$

Defining the reduced generator matrix for \mathcal{M} as $\downarrow_q (\tilde{G}_{\mathcal{M}}) = \text{unvec}(\downarrow_q (G_{\mathcal{Z}}))$, the above containment simplifies to the desired result

$$\mathcal{M} \subseteq \langle C_{\mathcal{M}}, \downarrow_q (\tilde{G}_{\mathcal{M}}) \rangle.$$





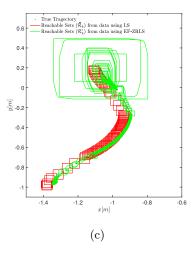


Fig. 3. The projection of the reachable sets of the Lipschitz nonlinear system in (1) computed via Algorithm 2 (\mathcal{R}'_k), the LS method ($\bar{\mathcal{R}}_k$) in [2, Algorithm 6], from noisy input-state data. (a): Nonlinear system of Example 2. (b): Real experiment of autonomous vehicle without payload shown in Figure 4a. (c): Real experiment of autonomous vehicle with payload shown in Figure 4b.

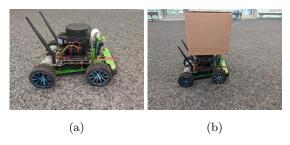


Fig. 4. The Jet Racer ROS AI Kit. (a): without payload. (b): with 2 kg payload.

Property 1 (Matrix Derivative [23]) For given matrices A, B, C, X with conformable dimensions, the following derivative rules hold:

Appendix B: Lemmas and Proofs

Lemma 1 (Matrix Perturbation) Let $C_{\mathcal{M}_p,0} \in \mathbb{R}^{n_{\mathcal{M}_p} \times m_{\mathcal{M}_p}}$ be an initial nominal matrix. Consider a sequence of matrices $\{C_{\mathcal{M}_p,j}\}_{j=0}^k$ where $C_{\mathcal{M}_p,j+1} = C_{\mathcal{M}_p,j} + \delta C_{\mathcal{M}_p,j}$. By defining a perturbation matrix $\delta C_{\mathcal{M}_p,j} \in \mathbb{R}^{n_{\mathcal{M}_p} \times m_{\mathcal{M}_p}}$ which is bounded by a constant $\mu > 0$ with respect to the element-wise infinity norm, i.e., $\|\delta C_{\mathcal{M}_p,j}\|_{max} \leq \mu$.

Then, the matrix $C_{\mathcal{M}_p,k}$ (after k steps) is contained within the matrix zonotope

$$C_{\mathcal{M}_n,k} \in \mathcal{M}_{p_k}$$
.

where

$$\mathcal{M}_{p_k} = \langle C_{\mathcal{M}_p,0}, \check{G}_{\mathcal{M}_{p_k}} \rangle.$$

with the generator set $\check{G}_{\mathcal{M}_{p_k}} = \left\{ G_{\mathcal{M}_{p_k}}^{(1)} \dots G_{\mathcal{M}_{p_k}}^{(\gamma_{\mathcal{M}_p})} \right\}$, where $\gamma_{\mathcal{M}_p} := n_{\mathcal{M}_p} \cdot m_{\mathcal{M}_p}$ and the generators are defined as

$$G_{\mathcal{M}_{p_k}}^{(s)} := k\mu E_s.$$

and E_s are the standard basis matrices of size $n_{\mathcal{M}_p} \times m_{\mathcal{M}_p}$ with a single non-zero element of 1. Here, the index s corresponds to the flattened index of the matrix elements.

Proof. The matrix $C_{\mathcal{M}_p,k}$ is given by the sum of successive perturbations

$$C_{\mathcal{M}_p,k} = C_{\mathcal{M}_p,0} + \sum_{j=0}^{k-1} \delta C_{\mathcal{M}_p,j}.$$

The perturbation matrix $\delta C_{\mathcal{M}_p,j}$ is bounded by $\|\delta C_{\mathcal{M}_p,j}\|_{\infty} \leq \mu$, which directly implies that for every element (i,l) of the matrix, $|(\delta C_{\mathcal{M}_p,j})_{(i,l)}| \leq \mu$.

Now, for any element (i, l)

$$\left(\sum_{j=0}^{k-1} \delta C_{\mathcal{M}_p, j}\right)_{(i,l)} = \sum_{j=0}^{k-1} (\delta C_{\mathcal{M}_p, j})_{(i,l)}.$$

Since $|(\delta C_{\mathcal{M}_p,j})_{(i,l)}| \leq \mu$ for each term in the sum, the sum of these elements is bounded as:

$$\left| \sum_{j=0}^{k-1} (\delta C_{\mathcal{M}_p,j})_{(i,l)} \right| \le \sum_{j=0}^{k-1} |(\delta C_{\mathcal{M}_p,j})_{(i,l)}| \le \sum_{j=0}^{k-1} \mu = k\mu.$$

Thus, the sum of perturbations is a matrix whose elements are bounded by $k\mu$. Let's call the sum of perturbations $\Delta C_{\mathcal{M}_n}$. Each element $(\Delta C_{\mathcal{M}_n})_{(i,l)}$ satisfies

 $|(\Delta C_{\mathcal{M}_p})_{(i,l)}| \leq k\mu$. We can express $\Delta C_{\mathcal{M}_p}$ as a linear combination of standard basis matrices:

$$\Delta C_{\mathcal{M}_p} = \sum_{i=1}^{n_{\mathcal{M}_p}} \sum_{l=1}^{m_{\mathcal{M}_p}} (\Delta C_{\mathcal{M}_p})_{(i,l)} E_{i,l}.$$

By using the flattened index s for the pair (i, l), we can rewrite this sum as

$$\Delta C_{\mathcal{M}_p} = \sum_{s=1}^{\gamma_{\mathcal{M}_p}} c_s E_s,$$

where E_s is a standard basis matrix and $|c_s| \leq k\mu$. We can define a set of coefficients $|\beta| \leq 1$ such that $c_s = k\mu\beta$. Substituting this back into the sum, we get:

$$\Delta C_{\mathcal{M}_p} = \sum_{s=1}^{\gamma_{\mathcal{M}_p}} \beta(k\mu E_s).$$

This is precisely the definition of the matrix zonotope with center zero and generators $k\mu E_s$. Since $C_{\mathcal{M}_p,k} = C_{\mathcal{M}_p,0} + \Delta C_{\mathcal{M}_p}$, the matrix $C_{\mathcal{M}_p,k}$ is contained within the matrix zonotope \mathcal{M}_{p_k} with center $C_{\mathcal{M}_p,0}$ and the defined generators.

References

- Amr Alanwar, Frank J. Jiang, and Karl H. Johansson. Polynomial logical zonotope: A set representation for reachability analysis of logical systems. *Automatica*, 171:111896, 2025.
- [2] Amr Alanwar, Anne Koch, Frank Allgöwer, and Karl Henrik Johansson. Data-driven reachability analysis from noisy data. IEEE Transactions on Automatic Control, 68(5):3054–3069, 2023.
- [3] Amr Alanwar, Jagat Jyoti Rath, Hazem Said, Karl Henrik Johansson, and Matthias Althoff. Distributed set-based observers using diffusion strategies. *Journal of the Franklin Institute*, 360(10):6976–6993, 2023.
- [4] Matthias Althoff. Reachability analysis and its application to the safety assessment of autonomous cars. PhD thesis, Technische Universität München, 2010.
- [5] Matthias Althoff, Olaf Stursberg, and Martin Buss. Reachability analysis of nonlinear systems with uncertain parameters using conservative linearization. In 2008 47th IEEE Conference on Decision and Control, pages 4042–4048. IEEE, 2008.
- [6] Somil Bansal and Claire J Tomlin. Deepreach: A deep learning approach to high-dimensional reachability. In 2021 IEEE International Conference on Robotics and Automation (ICRA), pages 1817–1824. IEEE, 2021.
- [7] Martin Berz and Georg Hoffstätter. Computation and application of Taylor polynomials with interval remainder bounds. *Reliable Computing*, 4(1):83–97, 1998.
- [8] Trevor J. Bird, Herschel C. Pangborn, Neera Jain, and Justin P. Koeln. Hybrid zonotopes: A new set representation for reachability analysis of mixed logical dynamical systems. *Automatica*, 154:111107, 2023.
- [9] José Manuel Bravo, Teodoro Alamo, and Eduardo F Camacho. Robust MPC of constrained discrete-time nonlinear systems based on approximated reachable sets. Automatica, 42(10):1745–1751, 2006.

- [10] Sven Brüggemann and Robert R Bitmead. Exponential convergence of recursive least squares with forgetting factor for multiple-output systems. *Automatica*, 124:109389, 2021.
- [11] Vamsi Krishna Chilakamarri, Zeyuan Feng, and Somil Bansal. Reachability analysis for black-box dynamical systems. arXiv preprint arXiv:2410.07796, 2024.
- [12] Christophe Combastel. Functional sets with typed symbols: Mixed zonotopes and polynotopes for hybrid nonlinear reachability and filtering. Automatica, 143:110457, 2022.
- [13] Raena Farhadsefat, Jiri Rohn, and Taher Lotfi. Norms of interval matrices. Academy of Sciences of the Czech Republic, Institute of Computer Science, Tech. Rep, 2011.
- [14] Mahsa Farjadnia, Amr Alanwar, Muhammad Umar B Niazi, Marco Molinari, and Karl Henrik Johansson. Robust datadriven predictive control of unknown nonlinear systems using reachability analysis. European Journal of Control, 74:100878, 2023.
- [15] Mahsa Farjadnia, Angela Fontan, Amr Alanwar, Marco Molinari, and Karl Henrik Johansson. Robust data-driven tube-based zonotopic predictive control with closed-loop guarantees. In 2024 IEEE 63rd Conference on Decision and Control (CDC), pages 6837–6843. IEEE, 2024.
- [16] Wolfgang Kühn. Rigorously computed orbits of dynamical systems without the wrapping effect. Computing, 61(1):47– 67, 1998.
- [17] Brian Lai and Dennis S Bernstein. Efficient batch and recursive least squares for matrix parameter estimation. IEEE Control Systems Letters, 2024.
- [18] Vu Tuan Hieu Le, Cristina Stoica, Teodoro Alamo, Eduardo F Camacho, and Didier Dumur. Zonotope-based setmembership estimation for multi-output uncertain systems. In 2013 IEEE International Symposium on Intelligent Control (ISIC), pages 212–217. IEEE, 2013.
- [19] Colas Le Guernic and Antoine Girard. Reachability analysis of linear systems using support functions. Nonlinear Analysis: Hybrid Systems, 4(2):250–262, 2010.
- [20] Ian Mitchell and Claire J Tomlin. Level set methods for computation in hybrid systems. In *International workshop* on hybrid systems: Computation and control, pages 310–323. Springer, 2000.
- [21] Jan Maximilian Montenbruck and Frank Allgöwer. Some problems arising in controller design from big data via input-output methods. In 55th Conference on Decision and Control, pages 6525–6530. IEEE, 2016.
- [22] Carlo Novara, Lorenzo Fagiano, and Mario Milanese. Direct feedback control design for nonlinear systems. Automatica, 49(4):849–860, 2013.
- [23] Kaare Brandt Petersen, Michael Syskind Pedersen, et al. The matrix cookbook. *Technical University of Denmark*, 7(15):510, 2008.
- [24] Brenner S. Rego, Davide M. Raimondo, and Guilherme V. Raffo. Line zonotopes: A tool for state estimation and fault diagnosis of unbounded and descriptor systems. *Automatica*, 179:112380, 2025.
- [25] Sadra Sadraddini and Russ Tedrake. Linear encodings for polytope containment problems. In 2019 IEEE 58th conference on decision and control (CDC), pages 4367–4372. IEEE, 2019.
- [26] Sergio E Samada, Vicenç Puig, and Fatiha Nejjari. Zonotopic recursive least-squares parameter estimation: Application to fault detection. *International Journal of Adaptive Control* and Signal Processing, 37(4):993–1014, 2023.