This article has been accepted for publication in IEEE Control Systems Letters. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/LCSYS.2024.3421194

Inner Approximations of Reachable Sets for Nonlinear Systems using the Minkowski Difference

Mark Wetzlinger, Adrian Kulmburg, and Matthias Althoff, *Member, IEEE*

***Abstract*— Reachability analysis is a formal method that rigorously proves whether a dynamical system can reach certain states. Inner approximations of the exact reachable set contain only states that are definitely reachable and are therefore used to falsify specifications. While the majority of state-of-the-art approaches for nonlinear systems obtain an inner approximation via first computing an outer approximation of the reachable set, we directly obtain sound inner approximations by using the Minkowski difference in a reachability algorithm for nonlinear systems. Our implementation uses a combination of polytopes and constrained zonotopes as set representations, resulting in a low polynomial time complexity in the state dimension. A comparison with state-of-the-art approaches on several benchmarks demonstrates the advantages of our approach.**

***Index Terms*— Formal verification, falsification, reachability analysis, nonlinear systems, set-based computing.**

I. Introduction

FORMAL methods can determine whether an uncertain dynamical system meets a given specification. Reachability analysis computes all states that are reachable under the given uncertainties, and, thus, can be used for formal verification [1]. However, the exact reachable set cannot be computed except for special system classes [2]; if a computed outer approximation cannot verify safety, an inner approximation may falsify safety, instead. Moreover, inner approximations are essential to many control tasks, where we want to find all states for which one can guarantee reaching a goal set through an inner approximation of the maximal backward reachable set [3].

Most approaches for computing inner approximations for nonlinear systems are based on outer approximations: Given an outer approximation for the state and the Jacobian of the flow, one can use extended interval arithmetic and the generalized mean value theorem to obtain lower bounds of the reachable set in axis-aligned directions in polynomial time [4]. This approach has been extended to deal with inputs [1], competing control inputs and disturbances [5], and neural-network controlled systems [6]; Another line of approaches [7]–[11] is based on the proof that a connected set is an inner approximation if it a) does not intersect the reachable set of the boundary of the initial set and b) contains some state of the exact reachable set [7, Thm. 4]. Hence, one can obtain an inner approximation by contracting an outer approximation, as illustrated in [10, Fig. 2]. This idea has been applied to autonomous systems, with inner approximations represented by polytopes [8], semi-algebraic sets [9], Taylor models [7], and polynomial zonotopes [10], for which the approach has been extended to systems with inputs [11, Ch. 4.3.3]. The reachable set for the boundary is commonly computed by partitioning the initial set into boxes, yielding an exponential time complexity in the state dimension [8], [9], or semi-algebraic sets, yielding polynomial time complexity [7], [10].

One can also obtain inner approximations via optimization-based techniques. Prominent approaches include Hamilton-Jacobi reachability [3], which scales exponentially in the state dimension due to gridding, and dissipativity-based approaches using sum-of-squares programming [12], [13], which scales polynomially in the state dimension, but exponentially in the degree of the polynomial representing the reachable set [14].

Figure 1 summarizes our main contribution: Using an on-the-fly linearization of the nonlinear autonomous dynamics, we obtain an affine dynamics and a higher-order error dynamics. While the Minkowski sum of the corresponding reachable sets $\mathcal{R}_{\text{aff}}$ and $\mathcal{R}_{\text{err}}$ is a common method to compute outer approximations $\widehat{\mathcal{R}}$, we are the first to obtain inner approximations $\widetilde{\mathcal{R}}$—for both time-point and time-interval reachable sets—via the Minkowski difference of the aforementioned sets. We choose a combination of polytopes and constrained zonotopes as set representations to implement a reachability algorithm with low polynomial time complexity in the state dimension.

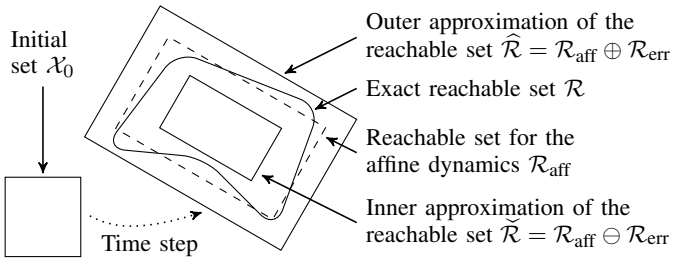

Fig. 1. Main idea: The Minkowski difference between the reachable sets for the affine dynamics $\mathcal{R}_{\text{aff}}$ and due to the error dynamics $\mathcal{R}_{\text{err}}$ returns a sound inner approximation $\widetilde{\mathcal{R}}$.

This paragraph of the first footnote will contain the date on which you submitted your paper for review. This work was supported by the European Research Council (ERC) project justITSELF under grant agreement No 817629.

Mark Wetzlinger, Adrian Kulmburg, and Matthias Althoff are with the School of Computation, Information and Technology, Technical University of Munich, 85748 Garching, Germany (e-mail: {m.wetzlinger, adrian.kulmburg, althoff}@tum.de).

This article has been accepted for publication in IEEE Control Systems Letters. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/LCSYS.2024.3421194

II. Preliminaries and Problem Statement

A. Notation and Set Operations

We denote vectors by lowercase letters, matrices by uppercase letters, and sets by calligraphic letters. For a vector $s \in \mathbb{R}^n$, $s_{(i)}$ represents its ith entry; for a matrix $M \in \mathbb{R}^{m\times n}$, $M_{(i,\cdot)}$ refers to the ith row and $M_{(\cdot,j)}$ to the jth column; for a set $\mathcal{S} \subseteq \mathbb{R}^n$, $\mathcal{S}_{(i)}$ returns the projection onto the ith dimension. The horizontal concatenation of two properly-sized matrices M_1, M_2 is denoted by $[M_1\ M_2]$. We write $0_n \in \mathbb{R}^n$ for an all-zero vector and $I_n \in \mathbb{R}^{n\times n}$ for the identity matrix. For a set $\mathcal{S} \subset \mathbb{R}^n$, we denote an outer approximation by $\widehat{\mathcal{S}} \supseteq \mathcal{S}$ and an inner approximation by $\breve{\mathcal{S}} \subseteq \mathcal{S}$. Real-valued intervals are denoted by $\mathcal{I} = [a, b] \subset \mathbb{R}^n$, where $\forall i \in \{1, \dots, n\}\colon a_{(i)} \leq b_{(i)}$ holds elementwise.

For a matrix $M \in \mathbb{R}^{n\times n}$ and a set $\mathcal{S} \subset \mathbb{R}^n$, the linear map is defined by $M\mathcal{S} \coloneqq \{Ms | s \in \mathcal{S}\}$. The operation $\texttt{cen}(\mathcal{S})$ returns the Chebyshev center of a set $\mathcal{S}$, while $\texttt{box}(\mathcal{S})$ returns the tightest enclosing axis-aligned interval. For two convex sets $\mathcal{S}_1, \mathcal{S}_2 \subset \mathbb{R}^n$, the Minkowski sum is defined as $\mathcal{S}_1 \oplus \mathcal{S}_2 \coloneqq \{s_1 + s_2 | s_1 \in \mathcal{S}_1, s_2 \in \mathcal{S}_2\}$, the Cartesian product is defined as $\mathcal{S}_1 \times \mathcal{S}_2 \coloneqq \{[s_1^\top\ s_2^\top]^\top | s_1 \in \mathcal{S}_1, s_2 \in \mathcal{S}_2\}$, the convex hull is defined as $\texttt{conv}(\mathcal{S}_1, \mathcal{S}_2) \coloneqq \{\lambda s_1 + (1-\lambda)s_2 | s_1 \in \mathcal{S}_1, s_2 \in \mathcal{S}_2, \lambda \in [0, 1]\}$, and the Minkowski difference is defined as

$$\begin{aligned}\mathcal{S}_1 \ominus \mathcal{S}_2 &\coloneqq \big\{s \,\big|\, s \oplus \mathcal{S}_2 \subseteq \mathcal{S}_1\big\} \\ &= \big\{s \,\big|\, \forall s_2 \in \mathcal{S}_2\colon s + s_2 \in \mathcal{S}_1\big\}.\end{aligned} \tag{1}$$

For three compact, convex, nonempty sets $\mathcal{S}_1, \mathcal{S}_2, \mathcal{S}_3 \subset \mathbb{R}^n$, it holds that [15, Eq. (2)]

$$\mathcal{S}_1 \ominus (\mathcal{S}_2 \oplus \mathcal{S}_3) = (\mathcal{S}_1 \ominus \mathcal{S}_2) \ominus \mathcal{S}_3 = (\mathcal{S}_1 \ominus \mathcal{S}_3) \ominus \mathcal{S}_2 \tag{2}$$

and [16, Lemma 1]

$$\texttt{conv}(\mathcal{S}_1 \ominus \mathcal{S}_3, \mathcal{S}_2 \ominus \mathcal{S}_3) \subseteq \texttt{conv}(\mathcal{S}_1, \mathcal{S}_2) \ominus \mathcal{S}_3, \tag{3}$$

which we will exploit in later derivations. We assume a time complexity of $\mathcal{O}\big((p+q)^{1.5}p^2\big)$ for the evaluation of a linear program with p variables and q constraints [17]. Finally, the Jacobian of a function $f\colon \mathbb{R}^n \to \mathbb{R}^n$ is denoted by Df with $\forall i, j \in \{1, ..., n\}\colon (Df(x))_{(i,j)} = \frac{\partial f_{(i)}(x)}{\partial x_{(j)}}$.

B. Problem Statement

We consider autonomous nonlinear continuous-time systems

$$\dot{x}(t) = f\big(x(t)\big), \tag{4}$$

where $x \in \mathbb{R}^n$ is the state vector and $f\colon \mathbb{R}^n \to \mathbb{R}^n$ is sufficiently smooth. Let the solution trajectory at time $t > 0$ for an initial state $x(0) \in \mathbb{R}^n$ be denoted by $\xi\big(t; x(0)\big)$.

Definition 1 (Exact reachable set): The exact reachable set at a time point $t > 0$ for an initial set $\mathcal{X}_0 \subset \mathbb{R}^n$ is defined as

$$\mathcal{R}(t) \coloneqq \big\{\xi\big(t; x(0)\big) \,\big|\, x(0) \in \mathcal{X}_0\big\}. \tag{5}$$

The reachable set over a time interval $\tau = [0, t_{\text{end}}], t_{\text{end}} > 0$ is the union of time-point solutions: $\mathcal{R}(\tau) = \bigcup_{t\in\tau} \mathcal{R}(t)$. □

As the exact reachable set $\mathcal{R}(t)$ cannot be computed for general nonlinear systems [2], our goal is to compute a tight inner approximation $\breve{\mathcal{R}}(t) \subseteq \mathcal{R}(t)$ instead.

III. Reachability Analysis

In this section, we first recall the set-based integration of nonlinear systems (4) in Section III-A. Thereafter in Section III-B, we present our main contribution, a set-based computation of an inner approximation $\breve{\mathcal{R}}(t)$ of the exact reachable set $\mathcal{R}(t)$.

A. Set-based Integration

For the integration of the right-hand side of (4) over a domain $\mathcal{X} \subset \mathbb{R}^n$, we use a Taylor expansion around the linearization point $x^* \in \mathbb{R}^n$,

$$\forall x \in \mathcal{X}\colon f(x) = f(x)\big|_{x=x^*} + Df(x)\big|_{x=x^*}(x - x^*) + l(x), \tag{6}$$

where $l(x) \in \mathcal{L}(\mathcal{X})$ is a vector within the Lagrange remainder $\mathcal{L}(\mathcal{X})$ defined component-wise by [18, Eq. (2)]

$$\begin{aligned}\mathcal{L}_{(i)}(\mathcal{X}) \coloneqq \big\{ &\tfrac{1}{2}(x - x^*)^\top D^2 f_{(i)}(\tilde{x})\big|_{\tilde{x}=\zeta}(x - x^*) \,\big| \\ &x \in \mathcal{X}, \zeta \in \{x^* + \alpha(x - x^*) \,|\, \alpha \in [0, 1]\}\big\}.\end{aligned} \tag{7}$$

Combining (6) and (7) yields the differential inclusion [18, Eq. (2)]

$$\forall x \in \mathcal{X}\colon f(x) \in w + A(x - x^*) \oplus \mathcal{L}(\mathcal{X}), \tag{8}$$

where $w = f(x^*)$, $A = Df(x)|_{x=x^*}$, and $\mathcal{L}_{(i)}(\mathcal{X})$ can be enclosed by $\widehat{\mathcal{L}}_{(i)} = [-\hat{l}_{(i)}, \hat{l}_{(i)}]$ with [18, Prop. 1]

$$\hat{l}_{(i)} = \tfrac{1}{2} \max_{x\in\mathcal{X}} |x - x^*|^\top \max_{\zeta\in\mathcal{X}} D^2 f_{(i)}(\tilde{x})\big|_{\tilde{x}=\zeta} \max_{x\in\mathcal{X}} |x - x^*|, \tag{9}$$

which is computed via interval arithmetic.

The integration of the linear differential inclusion (8) over a time interval $\tau_0 = [0, \Delta t]$ with step size Δt consists of two parts, computed separately due to the superposition principle:

1) Assuming $x^* = 0_n$, the reachable set of the affine dynamics $f_{\text{aff}}(x(t)) = Ax(t) + w$ is

$$\mathcal{R}_{\text{aff}}(\Delta t) = \mathcal{R}_{\text{hom}}(\Delta t) \oplus \mathcal{R}_{\text{con}}(\Delta t) \tag{10}$$

with the well-known homogeneous solution

$$\mathcal{R}_{\text{hom}}(\Delta t) = e^{A\Delta t}\mathcal{X}_0 \tag{11}$$

and the constant input solution

$$\mathcal{R}_{\text{con}}(\Delta t) = \int_0^{\Delta t} e^{A(\Delta t - \theta)} \mathrm{d}\theta\, w = A^{-1}(e^{A\Delta t} - I_n)w. \tag{12}$$

If the matrix A is not invertible, we can factor A^{-1} into the power series of the exponential matrix $e^{A\Delta t}$.

2) The abstraction error

$$\mathcal{R}_{\text{err}}\big(\mathcal{L}(\mathcal{R}(\tau_0))\big) = \int_0^{\Delta t} e^{A(\Delta t - \theta)} \mathcal{L}\big(\mathcal{R}(\tau_0)\big)\, \mathrm{d}\theta$$

due to the Lagrange remainder $\mathcal{L}\big(\mathcal{R}(\tau_0)\big)$ can be enclosed by [19, Prop. 3.7]

$$\begin{aligned}\widehat{\mathcal{R}}_{\text{err}}\big(\mathcal{L}(\mathcal{R}(\tau_0))\big) = &\bigoplus_{j=0}^{\eta} \frac{A^j \Delta t^{j+1}}{(j+1)!} \mathcal{L}\big(\mathcal{R}(\tau_0)\big) \\ &\oplus \boldsymbol{\mathcal{E}}(\Delta t)\Delta t\, \mathcal{L}\big(\mathcal{R}(\tau_0)\big),\end{aligned} \tag{13}$$

where $\boldsymbol{\mathcal{E}}(\Delta t)$ is the remainder term of the propagation matrix as in [16, Eq. (19)].

Let us now introduce the formulae for computing time-point and time-interval outer approximations, on which we base our derivations in Section III-B. The exact time-point reachable set $\mathcal{R}(\Delta t)$ at time $t = \Delta t$ can be enclosed by [18, Sec. III]

$$\mathcal{R}(\Delta t) \subseteq \widehat{\mathcal{R}}(\Delta t) = \mathcal{R}_{\text{aff}}(\Delta t) \oplus \widehat{\mathcal{R}}_{\text{err}}\big(\mathcal{L}(\mathcal{R}(\tau_0))\big). \tag{14}$$

Analogously, the time-interval reachable set $\mathcal{R}(\tau_0)$ over the first time interval τ_0 can be enclosed by [18, Sec. III]

$$\mathcal{R}(\tau_0) \subseteq \widehat{\mathcal{R}}(\tau_0) = \widehat{\mathcal{R}}_{\text{aff}}(\tau_0) \oplus \widehat{\mathcal{R}}_{\text{err}}\big(\mathcal{L}(\mathcal{R}(\tau_0))\big), \tag{15}$$

where $\widehat{\mathcal{R}}_{\text{aff}}(\tau_0)$ is an outer approximation of the time-interval affine solution

$$\widehat{\mathcal{R}}_{\text{aff}}(\tau_0) = \texttt{conv}\big(\mathcal{X}_0, \mathcal{R}_{\text{aff}}(\Delta t)\big) \oplus \mathcal{C}, \tag{16}$$

with the curvature error $\mathcal{C}$ computed as [19, Sec 3.2.2]

$$\mathcal{C} = \boldsymbol{\mathcal{F}}(\Delta t)\mathcal{X}_0 \oplus \boldsymbol{\mathcal{G}}(\Delta t)w, \tag{17}$$

over a time step size Δt and where the interval matrices $\boldsymbol{\mathcal{F}}(\Delta t)$ and $\boldsymbol{\mathcal{G}}(\Delta t)$ are computed as in [16, Eq. (18), (26)]. We will show in Section IV-B how to resolve the mutual dependency between $\widehat{\mathcal{R}}(\tau_0)$ and $\widehat{\mathcal{R}}_{\text{err}}\big(\mathcal{L}(\mathcal{R}(\tau_0))\big)$ [18, Sec. III].

B. Computation of an Inner Approximation

We now derive our novel set-based computation for inner approximations of the time-point and time-interval reachable sets. Let us first show how the Minkowski difference can be used as an inner approximation.

Lemma 1: Let $\mathcal{A} \subset \mathbb{R}^n$ be non-empty, convex, and compact, $\mathcal{B} \subseteq \mathbb{R}^n$ be centrally-symmetric around a center c (e.g., a zonotope), and $\varphi\colon \mathcal{A} \to \mathcal{B}$ be a continuous function. Moreover, let

$$\mathcal{S} \coloneqq \big\{x \in \mathbb{R}^n \,\big|\, \exists a \in \mathcal{A}\colon x = a + \varphi(a)\big\}. \tag{18}$$

Then, we have the containment relation

$$\mathcal{A} \ominus \mathcal{B} \subseteq \mathcal{S} - 2c. \tag{19}$$

Proof: We first treat the case where $\mathcal{B}$ is symmetric around the origin (i.e., $\mathcal{B} = -\mathcal{B}$, and $c = 0$). Let $x \in \mathcal{A} \ominus \mathcal{B}$, and consider the function $f_x(a) = x - \varphi(a)$ for $a \in \mathcal{A}$. Since $-\varphi(a) \in -\mathcal{B} = \mathcal{B}$, it holds that $f_x(a) \in x + \mathcal{B} \subseteq \mathcal{A}$, by the definition of the Minkowski difference. Consequently, f_x is a continuous function (since φ is continuous) that maps $\mathcal{A}$ to $\mathcal{A}$. By the Brouwer fixed-point theorem (which is the special case of the Schauder fixed-point theorem on $\mathbb{R}^n$, see [20, Theorem 11.1.]), there exists a fixed point $a^* \in \mathcal{A}$ such that $f_x(a^*) = a^*$, i.e., $x = a^* + \varphi(a^*)$. This implies $x \in \mathcal{S}$, proving that $\mathcal{A} \ominus \mathcal{B} \subseteq \mathcal{S}$ in the case $c = 0$.

If $c \neq 0$ we can define $\mathcal{A}' = \mathcal{A} - c$, $\mathcal{B}' = \mathcal{B} - c$, and $\varphi'\colon \mathcal{A}' \to \mathcal{B}'$ through $\varphi'(a') = \varphi(a' + c) - c$ for $a' \in \mathcal{A}'$. Then $\mathcal{A}' \ominus \mathcal{B}' = \mathcal{A} \ominus \mathcal{B}$, and

$$\begin{aligned}&\big\{x \in \mathbb{R}^n \,\big|\, \exists a' \in \mathcal{A}'\colon x = a' + \varphi'(a')\big\}\\ &= \big\{x \in \mathbb{R}^n \,\big|\, \exists a \in \mathcal{A}\colon x = a + \varphi(a)\big\} - 2c.\end{aligned}$$

Hence, we may use the arguments above on $\mathcal{A}'$, $\mathcal{B}'$, and φ', since $\mathcal{B}'$ is now symmetric, which proves (19). ■

Remark 1: The assumption in Lemma 1 that φ should be continuous cannot be dropped: If $\mathcal{A} = \mathcal{B} = [-1, 1]$, it holds that $\mathcal{A} \ominus \mathcal{B} = \{0\}$, yet if $\varphi(a) = \text{sign}(a)$ (with the convention $\text{sign}(0) = 1$), then $\mathcal{S} = [-2, -1) \cup [1, 2]$.

Crucially, Lemma 1 enables us to use the Minkowski difference for computing inner approximations.

Proposition 1 (Time-point reachable set): For a compact, non-empty initial set $\mathcal{X}_0 \subset \mathbb{R}^n$ and a centrally-symmetric set $\widehat{\mathcal{R}}_{\text{err}}\big(\mathcal{L}(\mathcal{R}(\tau_0))\big)$, see (13), an inner approximation of the time-point reachable set $\breve{\mathcal{R}}(\Delta t) \subseteq \mathcal{R}(\Delta t)$ at time $t = \Delta t$, can be computed by

$$\breve{\mathcal{R}}(\Delta t) = \mathcal{R}_{\text{aff}}(\Delta t) \ominus \widehat{\mathcal{R}}_{\text{err}}\big(\mathcal{L}(\mathcal{R}(\tau_0))\big) + 2c, \tag{20}$$

with $\mathcal{R}_{\text{aff}}(\Delta t)$ as in (10) and $c = \texttt{cen}\big(\widehat{\mathcal{R}}_{\text{err}}\big(\mathcal{L}(\mathcal{R}(\tau_0))\big)\big)$.

Proof: Each successor state $x(\Delta t) \in \mathcal{R}(\Delta t)$ can be expressed using an initial state $x(0) \in \mathcal{X}_0$ and an error vector $z\big(x(0)\big) \in \widehat{\mathcal{R}}_{\text{err}}\big(\mathcal{L}(\mathcal{R}(\tau_0))\big)$ that may depend on $x(0)$:

$$\begin{aligned}\mathcal{R}(\Delta t) = \big\{x(\Delta t) \,\big|\, &\exists x(0) \in \mathcal{X}_0\colon\\ &x(\Delta t) = \mathcal{R}_{\text{con}}(\Delta t) + e^{A\Delta t}x(0) + z\big(x(0)\big)\big\}.\end{aligned} \tag{21}$$

By assumption, the right-hand side f in (4) is at least Lipschitz continuous. Thus, the Picard-Lindelöf theorem entails that $x(\Delta t)$ depends continuously on $x(0)$, which in turn entails that $z\big(x(0)\big)$ must be continuous with respect to $x(0) \in \mathcal{X}_0$. Since $\mathcal{X}_0$ is non-empty and compact, and $\widehat{\mathcal{R}}_{\text{err}}\big(\mathcal{L}(\mathcal{R}(\tau_0))\big)$ is centrally-symmetric since it is a zonotope, we may use Lemma 1 to obtain (20) from (21). ■

Next, we show how to compute an inner approximation of the time-interval reachable set $\breve{\mathcal{R}}(\tau_0) \subseteq \mathcal{R}(\tau_0)$.

Proposition 2 (Time-interval reachable set): For a compact, non-empty initial set $\mathcal{X}_0 \subset \mathbb{R}^n$ and a centrally-symmetric set around c,

$$\mathcal{E} \coloneqq \mathcal{C} \oplus \widehat{\mathcal{R}}_{\text{err}}\big(\mathcal{L}(\mathcal{R}(\tau_0))\big) \tag{22}$$

computed using (17) and (13), respectively, an inner approximation of the time-interval reachable set $\breve{\mathcal{R}}(\tau_0) \subseteq \mathcal{R}(\tau_0)$ with $\tau_0 = [0, \Delta t]$ can be computed by

$$\breve{\mathcal{R}}(\tau_0) = \texttt{conv}\big((\mathcal{X}_0 \ominus \mathcal{C}) \ominus \widehat{\mathcal{R}}_{\text{err}}\big(\mathcal{L}(\mathcal{R}(\tau_0))\big), \breve{\mathcal{R}}(\Delta t) \ominus \mathcal{C}\big) + 2c. \tag{23}$$

Proof: Using (15)-(16), each reachable state in the time interval $\tau_0 \in [0, \Delta t]$ can be expressed as:

$$\begin{aligned}\mathcal{R}(\tau_0) = \big\{x(t) \,\big|\, &\exists t \in \tau_0\ \exists x(0) \in \mathcal{X}_0\colon x(t) = x(0) + \dots\\ &+ \tfrac{t}{\Delta t}\big(e^{A\Delta t}x(0) + \mathcal{R}_{\text{con}}(\Delta t) - x(0)\big) + z\big(x(0), t\big)\big\}\end{aligned}$$

with a function $z\big(x(0), t\big) \in \mathcal{E}$ that may depend on $x(0) \in \mathcal{X}_0$ and $t \in \tau_0$. Analogously to the proof of Proposition 1, $z\big(x(0), t\big)$ needs to be at least continuous in the variable $[x(0)^\top\ t]^\top \in \mathcal{X}_0 \times \tau_0$. Hence, applying Lemma 1 yields

$$\begin{aligned}\mathcal{R}(\tau_0) &\overset{(19)}{\supseteq} \texttt{conv}\big(\mathcal{X}_0, \mathcal{R}_{\text{aff}}(\Delta t)\big) \ominus \mathcal{E} + 2c\\ &\overset{(3)}{\supseteq} \texttt{conv}\big(\mathcal{X}_0 \ominus \mathcal{E}, \mathcal{R}_{\text{aff}}(\Delta t) \ominus \mathcal{E}\big) + 2c\\ &\overset{(2),\,(20),\,(22)}{=} \texttt{conv}\big((\mathcal{X}_0 \ominus \mathcal{C}) \ominus \widehat{\mathcal{R}}_{\text{err}}\big(\mathcal{L}(\mathcal{R}(\tau_0))\big), \breve{\mathcal{R}}(\Delta t) \ominus \mathcal{C}\big) + 2c,\end{aligned}$$

which yields the claim. ■

The reachable sets of later time steps $\breve{\mathcal{R}}(t_{k+1})$ and $\breve{\mathcal{R}}(\tau_k)$

can be obtained by recursively applying the above procedure, with $\widetilde{\mathcal{R}}(t_k)$ replacing $\mathcal{X}_0$. In the next section, we show how to implement a set-based reachability algorithm that computes the inner approximations derived in Propositions 1 and 2.

IV. Polynomial-Time Implementation

The choice of set representations for computing an inner approximation with the formulae (20) and (23) determines the time complexity of the resulting reachability algorithm in the state dimension. Our proposed implementation achieves a polynomial time complexity, as detailed subsequently.

A. Set Representations and Operations

Let us now introduce support functions, polytopes, and (constrained) zonotopes, as well as required operations.

Definition 2 (Support function [21, Def. 1]): The support function $\rho\colon \mathbb{R}^n \to \mathbb{R}$ of a convex, compact set $\mathcal{S} \subset \mathbb{R}^n$ in the direction $\ell \in \mathbb{R}^n$ is defined as

$$\rho(\mathcal{S}, \ell) \coloneqq \max_{s \in \mathcal{S}} \ell^\top s. \qquad \square$$

Definition 3 (Polytope [22, Sec. 1.1]): A polytope is a convex, compact set defined by the linear inequalities using a matrix $H \in \mathbb{R}^{h \times n}$ and a vector $d \in \mathbb{R}^h$:

$$\mathcal{P} \coloneqq \{ s \in \mathbb{R}^n \mid Hs \leq d \}.$$

We abbreviate $\mathcal{P} = \langle H, d \rangle_H$. $\square$

The evaluation of the linear map $M\mathcal{P}$ of a polytope with an invertible matrix $M \in \mathbb{R}^{n \times n}$ (in $\mathcal{O}(hn^2)$ operations) and the translation $\mathcal{P} + v$ by a vector $v \in \mathbb{R}^n$ (in $\mathcal{O}(hn)$ operations) follow directly from Definition 3. Additionally, we compute the Minkowski difference with another convex set $\mathcal{S} \subset \mathbb{R}^n$ (in $\mathcal{O}\big(h\mathcal{O}(\rho(\mathcal{S}, \ell))\big)$ operations) by [23, Thm. 2.2]

$$\begin{aligned} &\mathcal{P} \ominus \mathcal{S} = \langle H, \tilde{d} \rangle_H, \\ &\text{where } \forall j \in \{1, \dots, h\}\colon \tilde{d}_{(j)} = d_{(j)} - \rho\big(\mathcal{S}, H_{(j,\cdot)}^\top\big). \end{aligned} \tag{24}$$

The Chebyshev center $\texttt{cen}(\mathcal{P})$ can be computed using a single linear program (in $\mathcal{O}\big((h+n)^{1.5} n^2\big)$ operations) and a tight interval enclosure $\mathcal{I} = \texttt{box}(\mathcal{P}) \supseteq \mathcal{P}$ can be obtained via $2n$ support function evaluations (in $\mathcal{O}\big((h+n)^{1.5} n^3\big)$ operations).

Definition 4 ((Constrained) zonotope [24, Def. 1&3]): Using a vector $c \in \mathbb{R}^n$, a generator matrix $G \in \mathbb{R}^{n \times \gamma}$, a constraint matrix $K \in \mathbb{R}^{h \times \gamma}$, and a constraint offset $l \in \mathbb{R}^h$, a constrained zonotope $\mathcal{CZ} \subset \mathbb{R}^n$ is

$$\mathcal{CZ} \coloneqq \Big\{ c + \sum_{i=1}^{\gamma} G_{(\cdot,i)}\, \alpha_i \;\Big|\; \sum_{i=1}^{\gamma} K_{(\cdot,i)} \alpha_i = l,\ \alpha_i \in [-1, 1] \Big\}.$$

We abbreviate $\mathcal{CZ} = \langle c, G, K, l \rangle_{CZ}$. Omitting the equality constraints yields a zonotope $\mathcal{Z} = \langle c, G \rangle_Z$. $\square$

We require the following set operations for zonotopes and constrained zonotopes: the linear map $M\mathcal{Z}$ (in $\mathcal{O}(n^2\gamma)$ operations) [19, Eq. (2.1)], the Minkowski sum $\mathcal{Z}_1 \oplus \mathcal{Z}_2$ (in $\mathcal{O}(n)$ operations) [19, Eq. (2.1)], the support function evaluation $\rho(\mathcal{Z}, \ell)$ (in $\mathcal{O}(n\gamma)$ operations) [16, Eq. (14)], and an enclosure of the multiplication $\boldsymbol{\mathcal{M}}\mathcal{Z}$ with an interval matrix $\boldsymbol{\mathcal{M}} = \big[\underline{M}, \overline{M}\big] \in \mathbb{R}^{n \times n}$ (in $\mathcal{O}(n^2\gamma)$ operations) [19, Thm. 3.3]. The convex hull of two constrained zonotopes $\texttt{conv}(\mathcal{CZ}_1, \mathcal{CZ}_2)$ can be computed (in $\mathcal{O}(n)$ operations) according to [25, Thm. 5]. The exact conversion from a polytope to a constrained zonotope [24, Thm. 1], denoted by $\texttt{CZ}(\mathcal{P})$, is shown in [16, Alg. 1] (in $\mathcal{O}\big((h+n)^{1.5} n^3\big)$ operations).

B. Reachability Algorithm

Algorithm 1 computes a sequence of inner approximations of the time-point and time-interval reachable sets derived in Propositions 1 and 2. By Definition 1, the union of the individual inner approximations of the time-interval solutions (line 18) is an inner approximation of the reachable set over the entire time horizon $\tau = [0, t_{\text{end}}]$.

In each time step $k \in \{0, \dots, \omega - 1\}$, we first obtain a Taylor expansion of the nonlinear dynamics (4) to obtain the affine dynamics (line 5), for which we compute the time-point and time-interval solutions $\mathcal{R}_{\text{aff}}(\Delta t)$ and $\widehat{\mathcal{R}}_{\text{aff}}(\tau_0)$ (line 8). Note that we convert the start set $\widetilde{\mathcal{R}}(t_k)$ to a zonotope in order to efficiently compute the curvature error $\mathcal{C}$ (line 6). Next, we compute the Lagrange remainder (lines 10-13), for which we resolve the mutual dependency between the Lagrange remainder $\mathcal{L}$ and the time-interval reachable set $\widehat{\mathcal{R}}(\tau_0)$ as in [18, Sec. III]: We iteratively enlarge an initial guess $\overline{\mathcal{L}} = 0$ using a constant factor $\mu > 1$, and evaluate the Lagrange remainder (9) on the domain

$$\widehat{\mathcal{R}}(\tau_0) = \texttt{box}\Big(\widehat{\mathcal{R}}_{\text{aff}}(\tau_0) \oplus \widehat{\mathcal{R}}_{\text{err}}\big(\widehat{\mathcal{L}}\big)\Big) \tag{25}$$

until the containment condition (line 13) is fulfilled. Then, we represent $\widehat{\mathcal{L}}$ as a zonotope to efficiently evaluate the linear maps and Minkowski sums in (13), yielding the abstraction error set $\widehat{\mathcal{R}}_{\text{err}}\big(\widehat{\mathcal{L}}\big)$ (line 14). Finally, we apply Proposition 1 to compute an inner approximation $\widetilde{\mathcal{R}}(t_{k+1})$ of the time-point reachable set by Proposition 1 and an inner approximation $\widetilde{\mathcal{R}}(\tau_k)$ of the time-interval reachable set by Proposition 2.

Algorithm 1 only uses set operations introduced in Section IV-A, the most time-consuming of which are the $2n$ linear programs required for the enclosure of polytopes by axis-aligned intervals, e.g., for the evaluation of (25) (line 12). Hence, the overall time complexity is $\mathcal{O}\big((h+n)^{1.5} n^3\big)$, or $\mathcal{O}(n^{4.5})$ if we assume the number of constraints of the initial set $\mathcal{X}_0$ to be linear in the state dimension n. Please note that the number of steps linearly influences the time complexity.

V. Numerical Examples

We demonstrate the performance of Algorithm 1, implemented in the MATLAB toolbox CORA [26], on several benchmarks and compare it to the scaling approach [10], implemented also in CORA, and the projection approach [5], implemented in the C++ toolbox RINO [6]. All computations are performed on a 2.60GHz i7 processor with 32GB memory.

Table I shows the benchmarks with the system dimension n, the time horizon t_{end} in s, and a reference for the initial set $\mathcal{X}_0$. To measure the tightness of the results, we use [7, Sec. VI.]

$$\gamma_{\min}(t) = \min_{i \in \{1, \dots, n\}} \frac{\texttt{diam}_i\big(\texttt{box}\big(\widetilde{\mathcal{R}}(t)\big)\big)}{\texttt{diam}_i\big(\texttt{box}\big(\mathcal{R}_{\text{sim}}(t)\big)\big)}, \tag{26}$$

where $\mathcal{R}_{\text{sim}}(t)$ represents a close approximation of the exact reachable set, as it is the convex hull of simulated extremal

Algorithm 1 Inner approximation of the reachable set

Require: Nonlinear system $\dot{x}(t) = f(x(t))$, polytopic initial set $\mathcal{X}_0 = \langle H, d \rangle_H$, time horizon $\tau = [0, t_{\text{end}}]$, steps $\omega \in \mathbb{N}$
Ensure: Inner approximation of the reachable set $\breve{\mathcal{R}}(\tau)$

$\Delta t \leftarrow t_{\text{end}}/\omega, t_0 \leftarrow 0, \breve{\mathcal{R}}(t_0) \leftarrow \mathcal{X}_0, \mu \leftarrow 1.1$
for $k \leftarrow 0$ to $\omega - 1$ **do**
  $t_{k+1} \leftarrow t_k + \Delta t$, $\tau_k \leftarrow [t_k, t_{k+1}]$
  $x^*(t_k) \leftarrow \texttt{cen}(\breve{\mathcal{R}}(t_k)) + \frac{1}{2}\Delta t\, f(x^*(t_k))$
  $w, A \leftarrow$ Eq. (6) using $x^*(t_k)$
  $\mathcal{C} \leftarrow \boldsymbol{\mathcal{F}}(\Delta t)\,\texttt{box}(\breve{\mathcal{R}}(t_k)) \oplus \boldsymbol{\mathcal{G}}(\Delta t)w$
  $\mathcal{R}_{\text{hom}}(\Delta t) \leftarrow$ Eq. (11), $\mathcal{R}_{\text{con}}(\Delta t) \leftarrow$ Eq. (12)
  $\mathcal{R}_{\text{aff}}(\Delta t) \leftarrow$ Eq. (10), $\widehat{\mathcal{R}}_{\text{aff}}(\tau_0) \leftarrow$ Eq. (16)
  $\overline{\mathcal{L}} \leftarrow [0_n, 0_n]$
  repeat
    $\widehat{\mathcal{L}} \leftarrow \mu\overline{\mathcal{L}}$, $\widehat{\mathcal{R}}_{\text{err}}(\widehat{\mathcal{L}}) \leftarrow$ Eq. (13)
    $\widehat{\mathcal{R}}(\tau_0) \leftarrow$ Eq. (25), $\overline{\mathcal{L}} \leftarrow$ Eq. (9)
  until $\overline{\mathcal{L}} \subseteq \widehat{\mathcal{L}}$
  $\widehat{\mathcal{R}}_{\text{err}}(\widehat{\mathcal{L}}) \leftarrow$ Eq. (13), $c \leftarrow \texttt{cen}(\mathcal{C}) + \texttt{cen}(\widehat{\mathcal{R}}_{\text{err}}(\widehat{\mathcal{L}}))$
  $\breve{\mathcal{R}}(t_{k+1}) \leftarrow \mathcal{R}_{\text{aff}}(\Delta t) \ominus \widehat{\mathcal{R}}_{\text{err}}(\widehat{\mathcal{L}}) + 2c$ ▷ see Prop. 1
  $\breve{\mathcal{R}}(\tau_k) \leftarrow \texttt{conv}\Big(\texttt{CZ}(\breve{\mathcal{R}}(t_k) \ominus \widehat{\mathcal{R}}_{\text{err}}(\widehat{\mathcal{L}}) \ominus \mathcal{C}),$
    $\texttt{CZ}(\breve{\mathcal{R}}(t_{k+1}) \ominus \mathcal{C})\Big) + 2c$ ▷ see Prop. 2
end for
$\breve{\mathcal{R}}(\tau) \leftarrow \bigcup_{k=0}^{\omega-1} \breve{\mathcal{R}}(\tau_k)$

trajectories at time t, and $\texttt{diam}_i(\mathcal{S}) = b_{(i)} - a_{(i)} \in \mathbb{R}^n$ for an interval $\mathcal{S} = [a, b]$. For all benchmarks, we used a time step size of $\Delta t = 0.01\,\text{s}$ for both Algorithm 1 and for the projection approach [5], where we also tuned the maximum order for the Taylor models to 4. For the scaling approach [10], we used the same algorithm parameters as the authors[1].

Algorithm 1 is faster than the scaling approach [10] by orders of magnitude and even faster than the C++ implementation of the projection approach [5]. This provides empirical validation for the low polynomial time complexity derived at the end of Section IV. Please note that we have exploited the axis alignment of the initial sets $\mathcal{X}_0$—which commonly occurs in reachability analysis—to avoid solving linear programs, which has sped up our computation time on average by a factor of 3 to 4. Most computation time is spent on the range bounding in (9) using interval arithmetic, which explains the fast evaluation of the Lotka-Volterra and Biological Model benchmarks whose Hessian tensor is constant.

In all cases, the final reachable set returned by Algorithm 1 is at least 70% as tight as an estimate of the exact reachable set, see $\gamma_{\min}(t_{\text{end}})$ for Algorithm 1 in Table I. In two out of five cases, the scaling approach [10] and the projection approach [5] return a tighter final reachable set since they use higher-order abstractions that can accurately enclose the dynamics due to the constant Hessian tensor. However, the projection approach [5] only computes lower and upper bounds

[1]Repeatability package at https://codeocean.com/capsule/5233492/tree/v2.

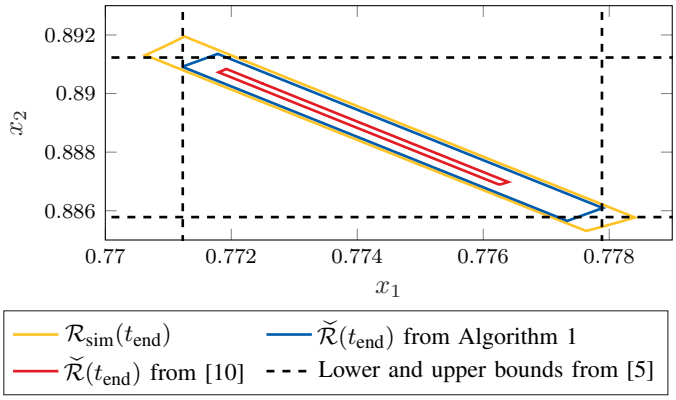

Fig. 2. Final reachable set of the Higgins Selkov benchmark.

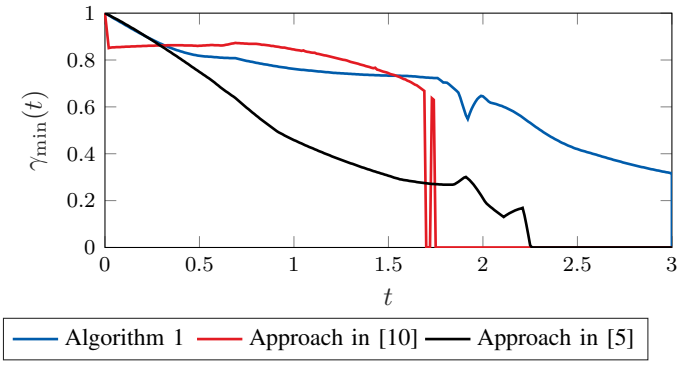

Fig. 3. Tightness metric $\boldsymbol{\gamma_{\min}(t)}$ over time for the Roessler benchmark. From $\boldsymbol{t = 3}$ onwards, all inner approximations are empty.

in axis-aligned directions, not an explicit inner approximation. This can also be seen in Figure 2, which shows the final reachable set of the Higgins Selkov benchmark computed by our approach, the approach in [10], and the bounds computed by the approach in [5]. Additionally, we plot the polygon $\mathcal{R}_{\text{sim}}(t_{\text{end}})$ constructed from the end points of the sampled extremal simulations.

Figure 3 plots $\gamma_{\min}(t)$ over time for the Roessler benchmark. Our computed inner approximation is always tighter than the one from the projection approach [5] and only marginally worse than the one from the scaling approach [10] early on. Crucially, Algorithm 1 returns non-empty inner approximations for much longer than the other approaches.

VI. Conclusion

We compute a sound inner approximation of the reachable set for nonlinear autonomous systems using the Minkowski difference between the reachable set due to the affine dynamics and the abstraction error due to the Lagrange remainder. A combination of polytopes and constrained zonotopes yields low polynomial time complexity. Future work will address non-autonomous nonlinear systems, where the main challenge is to design a reachability algorithm that can efficiently evaluate both the Minkowski difference and the Minkowski sum.

TABLE I

COMPARISON OF ALGORITHM 1 WITH THE APPROACHES [5], [10] IN TERMS OF COMPUTATION TIME (IN S) AND TIGHTNESS METRIC $\boldsymbol{\gamma_{\min}(t_{\text{end}})}$ (26).

Benchmark	n	t_{end}	$\mathcal{X}_0$	**Algorithm 1**		**Scaling approach [10]**		**Projection approach [5]**	
				Time	$\gamma_{\min}(t_{\text{end}})$	Time	$\gamma_{\min}(t_{\text{end}})$	Time	$\gamma_{\min}(t_{\text{end}})$
Jet Engine [27, Eq. (19)]	2	4	[28, Ex. 3.3.9]	**3.9**	**0.7769**	40	0.6093	16.4	0
Higgins-Selkov [29, Ex. II.1.]	2	3	[29, Ex. II.1.]	**2.5**	**0.8505**	57	0.5896	10	0.8127
Rössler [30, Eq. (2)]	3	1.5	[28, Ex. 3.4.3]	**1.1**	**0.7335**	32	0.7137	2.3	0.3089
Lotka-Volterra [31, Eq. (1)]	5	1	[28, Ex. 5.2.3]	**0.99**	0.7711	238	**0.8240**	7.8	0.7884
Biological model [32]	7	0.2	[28, Ex. 5.2.4]	**0.59**	0.7540	82	0.8760	2.2	**0.9811**

REFERENCES

[1] E. Goubault and S. Putot, "Inner and outer reachability for the verification of control systems," in *Proc. of the 22nd International Conference on Hybrid Systems: Computation and Control*, ACM, 2019, pp. 11–22.

[2] T. Gan, M. Chen, *et al.*, "Reachability analysis for solvable dynamical systems," *IEEE Transactions on Automatic Control*, vol. 63, no. 7, pp. 2003–2018, 2018.

[3] M. Chen and C. J. Tomlin, "Hamilton-Jacobi reachability: Some recent theoretical advances and applications in unmanned airspace management," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 1, pp. 333–358, 2018.

[4] E. Goubault and S. Putot, "Forward inner-approximated reachability of non-linear continuous systems," in *Proc. of the 20th International Conference on Hybrid Systems: Computation and Control*, ACM, 2017, pp. 1–10.

[5] E. Goubault and S. Putot, "Robust under-approximations and application to reachability of non-linear control systems with disturbances," *IEEE Control Systems Letters*, vol. 4, no. 4, pp. 928–933, 2020.

[6] E. Goubault and S. Putot, "RINO: Robust INner and Outer approximated reachability of neural networks controlled systems," in *Proc. of the 36th International Conference on Computer Aided Verification*, Springer, 2022, pp. 511–523.

[7] X. Chen, S. Sankaranarayanan, *et al.*, "Under-approximate flowpipes for non-linear continuous systems," in *Formal Methods in Computer-Aided Design*, IEEE, 2014, pp. 59–66.

[8] B. Xue, Z. She, *et al.*, "Under-approximating backward reachable sets by polytopes," in *International Conference on Computer Aided Verification*, Springer, 2016, pp. 457–476.

[9] B. Xue, Z. She, *et al.*, "Underapproximating backward reachable sets by semialgebraic sets," *IEEE Transactions on Automatic Control*, vol. 62, no. 10, pp. 5185–5197, 2017.

[10] N. Kochdumper and M. Althoff, "Computing non-convex inner-approximations of reachable sets for nonlinear continuous systems," in *Proc. of the 59th Conference on Decision and Control*, IEEE, 2020, pp. 2130–2137.

[11] N. Kochdumper, "Extensions of polynomial zonotopes and their application to verification of cyber-physical systems," Dissertation, Technische Universität München, 2022.

[12] H. Yin, A. Packard, *et al.*, "Reachability analysis using dissipation inequalities for uncertain nonlinear systems," *Systems and Control Letters*, vol. 142, p. 104 736, 2020.

[13] H. Yin, M. Arcak, *et al.*, "Backward reachability for polynomial systems on a finite horizon," *IEEE Transactions on Automatic Control*, vol. 66, no. 12, pp. 6025–6032, 2021.

[14] A. A. Ahmadi, G. Hall, *et al.*, "Improving efficiency and scalability of sum of squares optimization: Recent advances and limitations," in *Proc. of the 56th Conference on Decision and Control*, 2017, pp. 453–462.

[15] V. I. Danilov and G. A. Koshevoy, "Cores of cooperative games, superdifferentials of functions, and the Minkowski difference of sets," *Journal of Mathematical Analysis and Applications*, vol. 247, no. 1, pp. 1–14, 2000.

[16] M. Wetzlinger and M. Althoff, "Backward reachability analysis of perturbed continuous-time linear systems using set propagation," *arXiv preprint arXiv:2310.19083v2*, 2023.

[17] P. M. Vaidya, "An algorithm for linear programming which requires $\mathcal{O}(((M+n)n^2+(M+n)^{1.5}n)L)$ arithmetic operations," in *Proc. of the 19th Annual Symposium on Theory of Computing*, ACM, 1987, pp. 29–38.

[18] M. Althoff, O. Stursberg, *et al.*, "Reachability analysis of nonlinear systems with uncertain parameters using conservative linearization," in *Proc. of the 47th Conference on Decision and Control*, IEEE, 2008, pp. 4042–4048.

[19] M. Althoff, "Reachability analysis and its application to the safety assessment of autonomous cars," Dissertation, Technische Universität München, 2010.

[20] D. Gilbarg and N. S. Trudinger, *Elliptic Partial Differential Equations of Second Order*. Springer, 2001.

[21] C. Le Guernic and A. Girard, "Reachability analysis of linear systems using support functions," *Nonlinear Analysis: Hybrid Systems*, vol. 4, no. 2, pp. 250–262, 2010.

[22] G. M. Ziegler, *Lectures on polytopes*. Springer Science & Business Media, 2012.

[23] I. Kolmanovsky and E. G. Gilbert, "Theory and computation of disturbance invariant sets for discrete-time linear systems," *Mathematical Problems in Engineering*, vol. 4, 1998.

[24] J. K. Scott, D. M. Raimondo, *et al.*, "Constrained zonotopes: A new tool for set-based estimation and fault detection," *Automatica*, vol. 69, pp. 126–136, 2016.

[25] V. Raghuraman and J. P. Koeln, "Set operations and order reductions for constrained zonotopes," *Automatica*, vol. 139, p. 110 204, 2022.

[26] M. Althoff, "An introduction to CORA 2015," in *Proc. of the Workshop on Applied Verification for Continuous and Hybrid Systems*, 2015, pp. 120–151.

[27] E. Aylward, P. Parrilo, *et al.*, "Stability and robustness analysis of nonlinear systems via contraction metrics and SOS programming," *Automatica*, vol. 44, no. 8, pp. 2163–2170, 2008.

[28] X. Chen, "Reachability analysis of non-linear hybrid systems using Taylor models," Dissertation, RWTH Aachen University, 2015.

[29] X. Chen and S. Sankaranarayanan, "Decomposed reachability analysis for nonlinear systems," in *Proc. of the 37th Real-Time Systems Symposium*, IEEE, 2016, pp. 13–24.

[30] O. Rössler, "An equation for continuous chaos," *Physics Letters A*, vol. 57, no. 5, pp. 397–398, 1976.

[31] J. A. Vano, J. C. Wildenberg, *et al.*, "Chaos in low-dimensional Lotka–Volterra models of competition," *Nonlinearity*, vol. 19, no. 10, p. 2391, 2006.

[32] E. Klipp, R. Herwig, *et al.*, *Systems biology in practice: Concepts, implementation and application*. John Wiley & Sons, 2005.