Enhancing Data-Driven Reachability Analysis using Temporal Logic Side Information

Amr Alanwar∗, Frank J. Jiang∗, Maryam Sharifi, Dimos V. Dimarogonas, and Karl H. Johansson

Abstract—This paper presents algorithms for performing data-driven reachability analysis under temporal logic side information. In certain scenarios, the data-driven reachable sets of a robot can be prohibitively conservative due to the inherent noise in the robot’s historical measurement data. In the same scenarios, we often have side information about the robot’s expected motion (e.g., limits on how much a robot can move in a one-time step) that could be useful for further specifying the reachability analysis. In this work, we show that if we can model this side information using a signal temporal logic (STL) fragment, we can constrain the data-driven reachability analysis and safely limit the conservatism of the computed reachable sets. Moreover, we provide formal guarantees that, even after incorporating side information, the computed reachable sets still properly over-approximate the robot’s future states. Lastly, we empirically validate the practicality of the over-approximation by computing constrained, data-driven reachable sets for the Small-Vehicles-for-Autonomy (SVEA) hardware platform in two driving scenarios.

I. INTRODUCTION

Reachability analysis is an essential tool that provides a principled understanding of the dynamic capabilities of a system [1], [2]. In recent years, researchers have proposed a variety of formulations in which reachability analysis provides formal guarantees on the safety of an autonomous system (i.e., for autonomous vehicles [3] and drones [4]). Traditionally, a reachable set of states is computed based on a model of the subject system using either set-propagation techniques [5]–[7] or simulation-based techniques [8]–[11]. Most techniques compute over-approximations of the robot’s reachable states to ensure that the resulting reachable set can be used for providing safety guarantees. However, these traditional approaches are sensitive to model error and do not incorporate the readily available trajectory data that robots continuously produce.

Several recent works have proposed performing reachability analysis from data [12]–[20] to overcome the limitation of prior model knowledge. By performing reachability analysis directly from data, we can form a direct link between the actual, historical performance of a robot and our prediction of its reachability, removing the dependency on the accuracy of first-principles-based modeling. Moreover, in [21], [22], authors provide formal guarantees on the over-approximation of a system’s reachability based on data that contains noise. However, in order to provide guarantees on the over-approximation of the data-driven reachable sets, the computed sets might become prohibitively conservative when the noise becomes significant. In this work, we aim to limit this conservatism whenever we have useful side information.

The main contribution of this paper is an approach for performing data-driven reachability analysis under signal temporal logic (STL) side information. We choose to use STL since it can be interpreted over continuous-time signals, supports imposing strict deadlines and robust semantics [23], and allows for the formulation of complex specifications. To the extent of the authors’ knowledge, the presented approach is novel in its use of STL formulae as side information instead of as specifications (e.g. [24], [25]) while performing reachability analysis. More specifically, the contributions of this work are as follows: (1) We provide two algorithms for performing data-driven reachability analysis under STL side information, which, in turn, reduces the conservatism of data-driven reachable sets. (2) We provide state inclusion guarantees in reachable sets by intersecting a predicate function constructed from STL side information with either reachable zonotopes or reachable constrained zonotopes. (3) We validate our approach in two driving scenarios using the Small-Vehicles-for-Autonomy (SVEA) hardware platform (e.g., in Fig. 1).

The remainder of the paper is organized as follows. In Section II, we introduce preliminary material. In Section III, we present our approach to constrain the reachable sets using STL-based side information. In Section IV, we validate the
practicality of our approach using the SVEA platform. In Section V, we conclude the paper with final remarks.

II. PRELIMINARIES AND PROBLEM STATEMENT

In this section, we start by describing our assumed model for the subject system. After establishing our assumed model, we overview some necessary preliminary material and end the section by detailing the problem that we solve in Section III.

A. Model Description

We consider a discrete-time Lipschitz nonlinear system

\[ x(k + 1) = f(x(k), u(k)) + w(k). \]  

(1)

We assume \( f \) to be an unknown twice differentiable function and \( w(k) \in \mathcal{Z}_w \) to be process noise bounded by the set \( \mathcal{Z}_w \).

B. Reachable Set and Set Representations

In the following definitions, we define the reachable sets and different set representations used in our approach.

Definition 1: (Reachable set) The reachable set \( \mathcal{R}_N \) after \( N \) steps of system (1) from a set of initial states \( X_0 \subseteq \mathbb{R}^n \) and a set of possible inputs \( U_k \subseteq \mathbb{R}^m \) is

\[ \mathcal{R}_N = \{ x(N) \in \mathbb{R}^n \mid \forall k \in \{0, \ldots, N - 1 \} : \]

\[ x(k + 1) = f(x(k), u(k)) + w(k), w(k) \in \mathcal{Z}_w, \]

\[ u(k) \in U_k, x(0) = X_0 \}. \]

Definition 2: (Zonotope) [26, 27] Given center \( c_z \in \mathbb{R}^n \) and \( \gamma_z \in \mathbb{N} \) generator vectors in a generator matrix

\[ G_z = \begin{bmatrix} g_z^1 & \cdots & g_z^\gamma_z \end{bmatrix} \in \mathbb{R}^{n \times \gamma_z}, \]

a zonotope is defined as

\[ Z = \left\{ x \in \mathbb{R}^n \mid x = c_z + G_z \beta_z, -1 \leq \beta_z^i \leq 1 \right\}. \]

We use the shorthand notation \( Z = (c_z, G_z) \) for a zonotope.

The linear map is defined and computed as follows [28]:

\[ L \mathcal{Z} = \{ Lz \mid z \in \mathcal{Z} \} = (Lc_z, LG_z). \]

(2)

Given two zonotopes \( Z_1 = (c_{z_1}, G_{z_1}) \) and \( Z_2 = (c_{z_2}, G_{z_2}) \), the Minkowski sum \( Z_1 + Z_2 = \{ z_1 + z_2 \mid z_1 \in Z_1, z_2 \in Z_2 \} \) can be computed exactly as [28]:

\[ Z_1 + Z_2 = (c_{z_1} + c_{z_2}, G_{z_1} + G_{z_2}). \]

(3)

The Cartesian product of two zonotopes \( Z_1 \) and \( Z_2 \) is defined and computed as

\[ Z_1 \times Z_2 = \left\{ \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \mid z_1 \in Z_1, z_2 \in Z_2 \right\} = \begin{bmatrix} c_{z_1} \\ c_{z_2} \end{bmatrix}, \begin{bmatrix} G_{z_1} & 0 \\ 0 & G_{z_2} \end{bmatrix} \}. \]

(4)

The noise \( w(k) \) is random but bounded by the zonotope \( w(k) \in \mathcal{Z}_w = (c_{z_w}, G_{z_w}) \). With a minor abuse of notation, we write \( Z = \text{zonotope}([l, \bar{l}] \subseteq \mathbb{R}^n) \) to represent an interval vector \( I = [l, \bar{l}] \subseteq \mathbb{R}^n \) as a zonotope where the interval vector is defined element wise. Zonotopes have been extended in [29] to represent polytopes by applying constraints on the factors multiplied with the generators.

Definition 3: (Constrained zonotope) [29] An \( n \)-dimensional constrained zonotope is defined by

\[ C = \left\{ x \in \mathbb{R}^n \mid x = c_c + G_c \beta_c, A_c \beta_c = b_c, -1 \leq \beta_c^i \leq 1 \right\}, \]

where \( c_c \in \mathbb{R}^n \) is the center, \( G_c \in \mathbb{R}^{n \times \gamma_c} \) is the generator matrix and \( A_c \in \mathbb{R}^{n \times \gamma_c} \) and \( b_c \in \mathbb{R}^{n_c} \) constrain the factors \( \beta_c \). In short, we write \( C = (c_c, G_c, A_c, b_c) \).

Definition 4: (Strip [30]) For given parameters \( y_{i,k} \in \mathbb{R}^p, H_{i,k} \in \mathbb{R}^{p \times n} \) and \( r_{i,k} \in \mathbb{R}^p \), the strip \( S_{i,k} \) of index \( i \) is the set of all possible state values satisfying

\[ S_{i,k} = \{ x \mid |H_{i,k}x - y_{i,k}| \leq r_{i,k} \}, \]

(5)

where \( | \cdot | \) and \( \leq \) are applied element wise.

Definition 5: (Nonlinear strip) For given \( h_{i,k}(x) \in \mathbb{R}^p \) and \( r_{i,k} \in \mathbb{R}^p \) the nonlinear strip \( N_{i,k} \) of index \( i \) is the set of all possible state values satisfying

\[ N_{i,k} = \{ x \mid h_{i,k}(x) \leq r_{i,k} \}. \]

(6)

We denote the Moore-Penrose pseudoinverse by \( \dagger \) and the Kronecker product by \( \otimes \). We also denote the \( j \)th column of a matrix \( A \) by \( A_j \). The Frobenius norm is denoted by \( \| \cdot \|_F \). For simplicity, we consider dimension of \( p = 1 \).

C. Signal Temporal Logic

STL is an expressive language that is able to model complex, time-varying side information. STL is based on predicates \( \nu \) which are obtained by evaluation of a predicate function \( h(x) : \mathbb{R}^n \to \mathbb{R} \), where \( \nu := \top \) (True) if \( h(x) \geq 0 \) and \( \nu := \bot \) (False) if \( h(x) < 0 \) for \( x \in \mathbb{R}^n \) [31]. In this paper, we consider side information that can be modeled with the following STL fragment:

\[ \phi := G_{[a,b]} \otimes F_{[a,b]} \otimes U_{[a,b]} \phi' \otimes \phi'' = \phi' \otimes \phi'', \]

(7)

where \( \phi, \phi', \phi'' \) are STL formulas. In addition, \( U_{[a,b]} \) is the until operator with \( a \leq b < \infty \), and \( F_{[a,b]} = U_{[a,b]} \phi = \phi = \neg F_{[a,b]} \neg \phi \) are eventually and always operators, respectively. Let \( (x, t) \models \phi \) denote the satisfaction relation. A formula \( \phi \) is satisfiable if \( \exists x : \mathbb{R}_x \to \mathbb{R}^n \) such that \( (x, t) \models \phi \). STL semantics are defined formally as follows:

Definition 6: (STL semantics) [31] The STL semantics for a signal \( x : \mathbb{R}_x \to \mathbb{R}^n \) are recursively given by:

\[ (x, t) \models \nu \Leftrightarrow h(x) \geq 0, \]

\[ (x, t) \models \neg \phi \Leftrightarrow \neg (x, t) \models \phi, \]

\[ (x, t) \models \phi' \land \phi'' \Leftrightarrow (x, t) \models \phi' \land (x, t) \models \phi'', \]

\[ (x, t) \models \exists ! \exists \{ t + a, t + b \} s.t. (x, t_1) \models \phi'' \land \forall t_2 \{ t + a, t + b \} s.t. (x, t_2) \models \phi', \]

\[ (x, t) \models F_{[a,b]} \phi \Leftrightarrow \exists t_1 \{ t + a, t + b \} s.t. (x, t_1) \models \phi, \]

\[ (x, t) \models G_{[a,b]} \phi \Leftrightarrow \forall t_1 \{ t + a, t + b \} s.t. (x, t_1) \models \phi. \]

We omit the time to simplify the notation and write \( x \models \phi \).

D. Data-Driven Reachability Analysis

In this section, we show how we compute data-driven reachable sets from recorded trajectories. Consider \( K \) input-state data trajectories of length \( T_j, j = 1, \ldots, K \), from system (1), given by \( \{ u^{(j)}(k) \}_{k=0}^{T_j} \}, \{ x^{(j)}(k) \}_{k=0}^{T_j} \}. \) Denote the following matrices containing the set of all data sequence:

\[ X = [x^{(1)}(0) \cdots x^{(1)}(T_1) \cdots x^{(K)}(0) \cdots x^{(K)}(T_K)], \]

\[ U = [u^{(1)}(0) \cdots u^{(1)}(T_1-1) \cdots u^{(K)}(0) \cdots u^{(K)}(T_K-1)], \]

\[ X_+ = [x^{(1)}(1) \cdots x^{(1)}(T_1) \cdots x^{(K)}(1) \cdots x^{(K)}(T_K)], \]

\[ X_- = [x^{(1)}(0) \cdots x^{(1)}(T_1-1) \cdots x^{(K)}(0) \cdots x^{(K)}(T_K-1)]. \]

The total number of data points is denoted by \( T = \sum_{j=1}^{K} T_j \), and the set of all data by \( D = \{ U_-, X \}. \)
Algorithm 1 Reachability analysis for Lipschitz nonlinear system

**Input:** input-state trajectories \( D = (U_-, X) \), initial set \( \mathcal{X}_0 \), process noise zonotope \( Z_w \) and matrix zonotope \( M_w = (C_{M_w}, G_{M_w}) \), Lipschitz constant \( L^* \), covering radius \( \delta \), input zonotope \( U_k \), data-driven zonotope \( \tilde{Z}_k-1 \)

**Output:** data-driven zonotope \( \tilde{Z}_k \)

1. \( \tilde{M} = (X_+ - C_{M_w}) \left[ \begin{array}{c} X_+ - 1 \otimes x_k^* \\ U_+ - 1 \otimes u_k^* \end{array} \right] \)
2. \( \tilde{l} = \max_j \left( (X_+)_j - \tilde{M} \left[ \begin{array}{c} x_k^* \\ u_k^* \end{array} \right] \right) \)
3. \( \tilde{l} = \min_j \left( (X_+)_j - \tilde{M} \left[ \begin{array}{c} x_k^* \\ u_k^* \end{array} \right] \right) \)
4. \( \tilde{Z}_L = \text{zonotope}(\tilde{l}, \tilde{I}) - Z_w \)
5. \( \tilde{Z}_c = \{ \text{diag}(L^*, \ldots, L^*) \} \)
6. \( \tilde{Z}_k = \tilde{M}(1 \times \tilde{Z}_k-1 \times U_k) + Z_L + Z_c + Z_w \)

After collecting the data offline, we calculate an over-approximation of the reachable sets online using Algorithm 1 [22]. We compute a least-squares model \( \tilde{M} \) at a linearization point \((u_k^*, x_k^*)\) in line 1 where \( M_w = (C_{M_w}, G_{M_w}) \) is a the noise matrix zonotope [22] with center matrix \( C_{M_w} \) and a list of generator matrices \( G_{M_w} \). Then, we compute a zonotope that over-approximates the model mismatch and the nonlinearity terms in lines 2 to 4. Given that the data have a limited covering radius, we compute a Lipschitz zonotope in line 5 to provide guarantees. Next, we perform the reachability recursion in line 6 given the previously computed zonotopes. Note that the Lipschitz constant \( L^* \) and covering radius \( \delta \) can be computed as proposed in [22], [32], [33].

**E. Problem Statement**

Now that we have introduced the necessary preliminaries, we can detail the problem that we aim to solve.

**Problem 2.1:** Given the STL side information \( \phi_k = \phi_{1,k} \land \ldots \land \phi_{n_{\phi,k}} \) with \( \phi_{i,k} \) of the form (7), \( i = 1, \ldots, n_{\phi,k} \), a historical data set \( D = \{ U_-, X \} \) collected from an unknown system model, noise zonotope \( Z_w \), and input zonotope \( U_k \), compute the STL reachable set \( R_N \) at time step \( k = N \) starting from initial zonotope \( \mathcal{X}_0 \) that properly over-approximates the set of states \( \mathcal{R}_{\phi,N} \) where

\[
\mathcal{R}_{\phi,N} = \{ x(N) \in \mathbb{R}^n \mid \forall k \in \{0, \ldots, N-1\} : x(k+1) = \phi_{k+1}, x(k+1) = f(x(k), u(k)) + w(k), \mathcal{X}_0, u(k) \in U_k, x(0) = \phi_0 \}.
\]

The reachable set \( \mathcal{R}_N \) can be represented by a zonotope \( \tilde{Z}_N \supseteq \mathcal{R}_{\phi,N} \) or a constrained zonotope \( C_N \supseteq \mathcal{R}_{\phi,N} \).

**III. REACHABILITY ANALYSIS GIVEN STL SIDE INFORMATION**

In the previous section, we showed how to generate a data-driven reachable set from input-state data. In this section, we show how to incorporate STL formulas in data-driven reachability analysis. Algorithm 2 summarizes our proposed approach using zonotopes. The input to the algorithm is the data-driven zonotope \( \tilde{Z}_k \) from Algorithm 1 and STL side information \( \phi_{i,k}, i = 1, \ldots, n_{\phi,k} \) of the form (7).

**Algorithm 2. Reachability analysis under STL side information using zonotopes**

**Input:** data-driven zonotope \( \tilde{Z}_k = (\tilde{c}_k, \tilde{G}_k) \), STL side information \( \phi_{i,k}, \forall i = 1, \ldots, n_{\phi,k} \)

**Output:** STL zonotope \( \tilde{Z}_k = (\tilde{c}_k, \tilde{G}_k) \)

1. \( \tilde{c}_k = \tilde{c}_k, \tilde{G}_k = \tilde{G}_k \)
2. for \( i = 1, \ldots, n_{\phi,k} \) do
3. Construct \( h_{i,k}(x) \) from \( \phi_{i,k} \)
4. \( \lambda_{i,k}^* = \min_{\lambda_{i,k}} \| G_k \|_F^2 \)
5. if \( h_{i,k}(x) \) is linear then
6. if \( h_{i,k}(x) \) is nonlinear then
7. \( \tilde{c}_k = \tilde{c}_k + \lambda_{i,k}^* (y_{i,k} - H_{i,k} \tilde{c}_k) \)
8. \( \tilde{G}_k = (I - \lambda_{i,k}^* H_{i,k}) \tilde{G}_k \lambda_{i,k}^* \)
9. else if \( h_{i,k}(x) \) is nonlinear then
10. \( \tilde{G}_k = (I - \lambda_{i,k}^* H_{i,k}) \tilde{G}_k \lambda_{i,k}^* \)
11. \( \tilde{c}_k = \tilde{c}_k - \lambda_{i,k}^* (h_{i,k}(x) - c_{L,i,k} \tilde{c}_k + x_{i,k} - c_{L,i,k}) \)
12. end if
13. end for

In line 4, we use Algorithm 3 to compute reachable sets under STL side information using constrained zonotopes. Similar to Algorithm 2, we construct \( h_{i,k}(x) \) from \( \phi_{i,k} \) in line 3. Then, we provide intersection between constrained zonotope \( \tilde{c}_k \) and the \( h_{i,k}(x) \) in line 7. In case of nonlinear \( h_{i,k}(x) \), we provide the intersection in lines 10 to 13. In both Algorithms 2 and 3, we guarantee state inclusion by providing an over-approximated intersection between the data-driven reachable set \( \tilde{R}_k \) and the \( h_{i,k}(x) \) in line 7. To guarantee state inclusion in the STL generated set in case of nonlinear \( h_{i,k}(x) \), we use the minima of the STL side information for the linear case. In the nonlinear case, we use the linear approximation of the STL side information to compute a conservative zonotope that over-approximates the reachable set.
Algorithm 3 Reachability analysis under STL side information using constrained zonotopes

**Input:** data-driven constrained zonotope $\hat{C}_k=(\hat{c}_k, \hat{G}_k, \hat{A}_k, \hat{b}_k)$, STL side information $\phi_i,k, i=1, \ldots, n_{\phi,k}$

**Output:** STL constrained zonotope $\hat{C}_k=(\bar{c}_k, \bar{G}_k, \bar{A}_k, \bar{b}_k)$

1. $\bar{c}_k = \hat{c}_k$, $\bar{G}_k = \hat{G}_k$, $\bar{A}_k = \hat{A}_k$, $\bar{b}_k = \hat{b}_k$
2. for $i = 1, \ldots, n_{\phi,k}$ do
   3. Construct $h_{i,k}(x)$ from $\phi_i,k$
   4. if $h_{i,k}(x)$ is linear then
      5. if $h_{i,k}(x) = r_{i,k} - |H_{i,k}x - y_{i,k}|$
         6. $\bar{c}_k = \hat{c}_k$, $\bar{G}_k = \hat{G}_k$
      7. $\bar{A}_k = \begin{bmatrix} 0 & H_{i,k} \bar{G}_k \end{bmatrix}$, $\bar{b}_k = \begin{bmatrix} \bar{b}_k \end{bmatrix}^\top$
   8. else if $h_{i,k}(x)$ is nonlinear then
      9. $\bar{c}_k = \hat{c}_k$, $\bar{G}_k = \hat{G}_k$
   10. Compute $Z_{L,i,k} = \{c_{L,i,k}, G_{L,i,k}\}$ [3, p.65]
   11. $\bar{A}_k = \begin{bmatrix} \partial_{h_{i,k}}|x^*_{i,k} \hat{G}_k & -r_{i,k} & G_{L,i,k} \end{bmatrix}$
   12. $\bar{b}_k = \begin{bmatrix} -h_{i,k}(x^*_{i,k}) - \partial_{h_{i,k}}|x^*_{i,k} (\hat{c}_k - x^*_{i,k}) - c_{L,i,k} \end{bmatrix}$
13. end if
14. end for

Consider the Lagrange remainder $Z_{L,i,k} = (c_{L,i,k}, G_{L,i,k})$ [3, p.65] results in
$$\frac{\partial h_{i,k}}{\partial x}|x^*_{i,k} \hat{G}_k z_k \in -h_{i,k}(x^*_{i,k}) \frac{\partial h_{i,k}}{\partial x}|x^*_{i,k} (\hat{c}_k - x^*_{i,k}) - Z_{L,i,k} + r_{i,k} d.$$ (12)

Inserting (12) in (10) results in
$$x \in \hat{c}_k + \lambda_i,k \left[ \frac{\partial h_{i,k}}{\partial x}|x^*_{i,k} \hat{G}_k \right] z_k,$$
$$\bar{b}_k = \begin{bmatrix} -h_{i,k}(x^*_{i,k}) - \frac{\partial h_{i,k}}{\partial x}|x^*_{i,k} (\hat{c}_k - x^*_{i,k}) - c_{L,i,k} \end{bmatrix}$$

Note that $z_k \in [-1_{\gamma_2 \times 1}, 1_{\gamma_2 \times 1}]$ as $d \in [-1_{p \times 1}, 1_{p \times 1}]$, $z \in [-1_{\gamma_2 \times 1}, 1_{\gamma_2 \times 1}]$, and $z \in [-1_{\gamma_2 \times 1}, 1_{\gamma_2 \times 1}]$. Thus, the center and the generator of the over-approximating zonotope are $\bar{c}_k$ and $\bar{G}_k$, respectively.

**Theorem 2:** Algorithm 3 provides reachability analysis with state inclusion guarantees under STL side information, i.e., $\hat{Z}_k \supseteq R_{\phi,k}$.

**Proof:** Similar to the proof of Theorem 1, we omit the proof for the linear case as it follows immediately from [29, Prop.1] and prove the guaranteed intersection in the nonlinear case as follows: Let $x \in (\hat{C}_k \cap N_i,k)$, then there is a $z_k \in [-1_{\gamma_2 \times 1}, 1_{\gamma_2 \times 1}]$ such that
$$x = \hat{c}_k + \hat{G}_k z_k,$$ (13)
$$\bar{A}_k z_k = \bar{b}_k.$$ (14)

Given that $x$ is inside the intersection of the constrained zonotope $\hat{C}_k$ and $N_{i,k}$, there exists a $d \in [-1_{p \times 1}, 1_{p \times 1}]$ such that
$$h_{i,k}(x^*_{i,k}) + \frac{\partial h_{i,k}}{\partial x}|x^*_{i,k} (x - x^*_{i,k}) + \cdots = r_{i,k} d.$$ (15)

Inserting (13) into (15) results in
$$h_{i,k}(x^*_{i,k}) + \frac{\partial h_{i,k}}{\partial x}|x^*_{i,k} (\hat{c}_k + \hat{G}_k z_k - x^*_{i,k}) + \cdots = r_{i,k} d.$$ (16)

We combine (16) and (14) while considering the Lagrange remainder yielding
$$\begin{bmatrix} \bar{A}_k & 0 & 0 \\ \frac{\partial h_{i,k}}{\partial x}|x^*_{i,k} \hat{G}_k & -r_{i,k} & G_{L,i,k} \end{bmatrix} z_k = \begin{bmatrix} z_k \\ \bar{b}_k \end{bmatrix}.$$ (17)

Note that we consider the superset consisting the equality (16) by solving it for all $d \in [-1_{p \times 1}, 1_{p \times 1}]$. Then, we can assure that (16) is also satisfied.
ment compared to original data-driven reachable sets.

IV. Evaluation

In this section, we detail the application of our method to two examples. Readers can find an overview video of our experiments conducted at the Smart Mobility Lab at [https://bit.ly/DataReachSTL].

For our experimental platform, we represent a vehicle V with a SVEA vehicle [25]. We use historical data sets of length 1000 points gathered from the same car from other driving scenarios than the presented ones. We perform a single-step reachability analysis for each example, and we manually operate the car such that its behavior satisfies the known side information. Measurements for both the historical data sets and our two examples are made using a motion capture system. The assumed process noise zonotope is \( Z_{\theta} = \{ 0, [0.9 \ 0.9]^{T} \} \) and measurement noise zonotope of value \( [0.01 \ 0.01]^{T} \). For both examples, let V and its environment be defined over \( \mathbb{R}^2 \). In other words, V’s state \( x \in \mathbb{R}^2 \) is written as \( x = [x_1 \ x_2]^{T} \), where \( x_1 \) and \( x_2 \) are the x and y positions of V. Now, in the following sections, we will introduce our two scenarios for V and present the results for each case.

A. Parking Lot Example

In this example, we consider side information that contains only linear spatial constraints. Suppose V is parked in the parking lot and is scheduled to depart the parking lot soon. As denoted in Fig. 2, let the set of states corresponding to the parking region be \( P \subset \mathbb{R}^2 \) and the set of states corresponding to the outside of the parking region (the street) be \( O \subset \mathbb{R}^2 \). Note, the entrance and exit of the parking lot is considered both part of the parking region and the street. We know that V is scheduled to leave the parking region within 25 seconds of the start of our scenario. Thus, we can write the following STL formula as the known side information about V:

\[
\phi_p ::= G_{[0,25]}(P) \land F_{[0,25]}(P \land O) \land G_{[25,40]}(O). \tag{17}
\]

We can find the functions \( h_1 \) to \( h_5 \), which encode (17):

\[
\begin{align*}
  h_1(x_1, x_2) &= 1.7175 - |x_1 - 0.2805|, t \in [0, 25], \\
  h_2(x_1, x_2) &= 2.429 - |x_2 - 0.839|, t \in [0, 25], \\
  h_3(x_1, x_2) &= 1.3045 - |x_1 + 0.3225|, t \in [24, 25], \\
  h_4(x_1, x_2) &= 0.453 - |x_2 + 1.137|, t \in [24, 25], \\
  h_5(x_1, x_2) &= 1 - |x_2 + 1.665|, t \in [25, 40],
\end{align*}
\]

where \( h_1 \) and \( h_2 \) models our knowledge of V’s time within the region \( P \), \( h_3 \) and \( h_4 \) encodes V eventually reaching the exit region \( P \land O \) before \( t = 25 \), and \( h_5 \) corresponds to our knowledge of when V departs to \( O \). Fig. 3 shows a snapshot of the data-driven reachable sets before and after being constrained by \( \phi_p \) at \( t = 1 \). We show the unconstrained, data-driven reachable sets in Fig. 4a and the STL reachable sets constrained by \( \phi_p \) in Fig. 4b.

Then, suppose we know the upper limit of V’s capability to move forward and change heading between each sampling time. Let this set be denoted by \( T(x) \). Then, we can expand (17) into the following STL formula as the known side information about V: \( \phi_\theta ::= G_{[0,40]}(T(x)) \land G_{[0,25]}(P) \land F_{[0,25]}(P \land O) \land G_{[25,40]}(O) \). Now, we find the additional functions \( h_6, h_7 \), which encode the constraints corresponding to \( G_{[0,40]}(T(x)) \). Let \( \theta \) be the heading angle and \( \theta_c \) be the known, maximum heading angle change between each sampling time. We derive the constrained rectangular region \( T(x) \), shown in Fig. 6, with the following equations using the edges coordinates \( x_i^\theta, y_i^\theta, i = 1, \ldots, 4 \):

\[
\begin{align*}
  h_6(x_1, x_2) &= 0.5|c_2 - c_3| - | - m_2 x_1 + x_2 - 0.5(c_1 + c_4)|, \\
  h_7(x_1, x_2) &= 0.5|c_1 - c_4| - | - m_1 x_1 + x_2 - 0.5(c_2 + c_3)|,
\end{align*}
\]

where \( m_1 = \frac{y_2^\theta - y_1^\theta}{x_2^\theta - x_1^\theta} \) and \( c_i = -m_i x_i^\theta + y_i^\theta \) for \( i = 1, 2, c_3 = -m_2 x_2^\theta + y_2^\theta \), and \( c_4 = -m_1 x_3^\theta + y_3^\theta \). Both \( h_4 \) and \( h_5 \) are

---

TABLE I: Average volumes in the parking example.

| Zonotope       | Constrained zonotope |
|----------------|----------------------|
| No constraints | 9.722                |
| \( \phi_p \) constraints | 9.311 7.042       |
| \( \phi_\theta \) constraints | 0.124 0.076     |

---

TABLE II: Average volumes in the roundabout example.

| Zonotope       | Constrained zonotope |
|----------------|----------------------|
| No constraints | 9.722                |
| \( \phi_r \) constraints | 9.109 5.956       |
V. Conclusion

We have provided an approach to achieve less conservative, data-driven reachable sets. We have shown that known, STL-based side information can be used to constrain reachable zonotopes post-analysis, while still maintaining safety guarantees on the resulting constrained zonotopes. In future work, we will evaluate our approach on more complex scenarios and potentially apply the work to multi-agent tasks.
Reference:

[1] A. B. Kurzhanski and P. Varaiya, “Ellipsoidal techniques for reachability analysis,” in International Workshop on Hybrid Systems: Computation and Control, pp. 202–214, 2000.

[2] T. Pecsvaradi and K. S. Narendra, “Reachable sets for linear dynamical systems,” Information and control, vol. 19, no. 4, pp. 319–344, 1971.

[3] M. Althoff, Reachability analysis and its application to the safety assessment of autonomous cars. PhD thesis, Technische Universität München, 2010.

[4] J. H. Gillula, G. M. Hoffmann, H. Huang, M. P. Vitus, and C. J. Tomlin, “Applications of hybrid reachability analysis to robotic aerial vehicles,” The International Journal of Robotics Research, vol. 30, no. 3, pp. 335–354, 2011.

[5] M. Berz and K. Makino, “Rigorous reachability analysis and domain decomposition of Taylor models,” in International Workshop on Numerical Software Verification, pp. 90–187, 2017.

[6] S. V. Rakovic, E. C. Kerrigan, D. Q. Mayne, and J. Lygeros, “Reachability analysis of discrete-time systems with disturbances,” IEEE Transactions on Automatic Control, vol. 51, no. 4, pp. 546–561, 2006.

[7] N. Kochdumper and M. Althoff, “Sparse polynomial zonotopes: A novel set representation for reachability analysis,” arXiv preprint arXiv:1901.01780, 2019.

[8] A. Donzé and O. Maler, “Systematic simulation using sensitivity analysis,” in International Workshop on Hybrid Systems: Computation and Control, pp. 174–189, 2007.

[9] T. Lew and M. Pavone, “Sampling-based Reachability Analysis: A Random Set Theory Approach with Adversarial Sampling,” arXiv preprint arXiv:2008.10180, 2020.

[10] A. A. Julius, G. E. Fainekos, M. Anand, I. Lee, and G. J. Pappas, “Robust test generation and coverage for hybrid systems,” in International Workshop on Hybrid Systems: Computation and Control, pp. 329–342, 2007.

[11] J. Maidens and M. Arcak, “Reachability analysis of nonlinear systems using matrix measures,” IEEE Transactions on Automatic Control, vol. 60, no. 1, pp. 265–270, 2014.

[12] A. Devonport and M. Arcak, “Data-driven reachable set computation using adaptive gaussian process time series and Monte Carlo methods,” in American Control Conference, pp. 2629–2634, IEEE, 2020.

[13] F. Djeumou, A. P. Vinod, E. Goubault, S. Putot, and U. Topcu, “On-The-Fly Control of Unknown Smooth Systems from Limited Data,” arXiv preprint arXiv:2009.12733, 2020.

[14] A. Devonport, F. Yang, L. E. Ghaoui, and M. Arcak, “Data-Driven Reachability Analysis with Christoffel Functions,” arXiv preprint arXiv:2104.13902, 2021.

[15] A. Chakrabarty, A. Raghunathan, S. Di Cairano, and C. Danielson, “Data-driven estimation of backward reachable and invariant sets for unmodeled systems via active learning,” in 2018 IEEE Conference on Decision and Control, pp. 372–377, IEEE, 2018.

[16] A. Chakrabarty, C. Danielson, S. Di Cairano, and A. Raghunathan, “Active Learning for Estimating Reachable Sets for Systems With Unknown Dynamics,” IEEE Transactions on Cybernetics, pp. 1–12, 2020.

[17] R. E. Allen, A. A. Clark, J. A. Staruk, and M. Pavone, “A machine learning approach for real-time reachability analysis,” in IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 2202–2208, 2014.

[18] A. R. R. Matavalam, U. Vaidya, and V. Ajjarapu, “Data-driven approach for uncertainty propagation and reachability analysis in dynamical systems,” in 2020 American Control Conference (ACC), pp. 3393–3398, IEEE, 2020.

[19] A. J. Thorpe, K. R. Ortiz, and M. M. Oishi, “Data-driven stochastic reachability using hilbert space embeddings,” arXiv preprint arXiv:2010.08036, 2020.

[20] S. Bak, S. Bogomolov, P. S. Duggirala, A. R. Gerlach, and K. Potomkin, “Reachability of black-box nonlinear systems after koopman operator linearization,” arXiv preprint arXiv:2105.00886, 2021.

[21] A. Alanwar, A. Koch, F. Allgöwer, and K. H. Johansson, “Data-Driven Reachability Analysis Using Matrix Zonotopes,” in Proceedings of the 3rd Conference on Learning for Dynamics and Control, vol. 144, pp. 163–175, 2021.

[22] A. Alanwar, A. Koch, F. Allgöwer, and K. H. Johansson, “Data-Driven Reachability Analysis from Noisy Data,” arXiv preprint arXiv:2105.07229, 2021.

[23] G. E. Fainekos and G. J. Pappas, “Robustness of temporal logic specifications for continuous-time signals,” Theoretical Computer Science, vol. 410, no. 42, pp. 4262–4291, 2009.

[24] L. Lindemann and D. V. Dimarogonas, “Barrier Function-based Collaborative Control of Multiple Robots under Signal Temporal Logic Tasks,” IEEE Trans. Control Network Syst., vol. 7, no. 4, pp. 1916–1928, 2020.

[25] F. J. Jiang, Y. Gao, L. Xie, and K. H. Johansson, “Ensuring safety for vehicle parking tasks using Hamilton-Jacobi reachability analysis,” in 2020 59th IEEE Conference on Decision and Control (CDC), pp. 1416–1421, 2020.

[26] W. Kühn, “Rigorously computed orbits of dynamical systems without the wrapping effect,” Computing, vol. 61, no. 1, pp. 47–67, 1998.

[27] A. Girard, “Reachability of uncertain linear systems using zonotopes,” in Hybrid Systems: Computation and Control, LNCS 3414, pp. 291–305, 2005.

[28] A. Alanwar, J. J. Rath, H. Said, K. H. Johansson, and M. Althoff, “Distributed Set-Based Observers Using Diffusion Strategy,” arXiv preprint arXiv:2003.10347, 2020.

[29] J. K. Scott, D. M. Raimondo, G. R. Marseglia, and R. D. Braatz, “Constrained zonotopes: A new tool for set-based estimation and fault detection,” in Automatica, vol. 69, pp. 126–136, 2016.

[30] V. T. H. Le, C. Stoica, T. Alamo, E. F. Camacho, and D. Dunur, “Zonotope-based set-membership estimation for multi-output uncertain systems,” in IEEE International Symposium on Intelligent Control, pp. 212–217, 2013.

[31] O. Maler and D. Nickovic, “Monitoring temporal properties of continuous signals,” in Formal Techniques, Modelling and Analysis of Timed and Fault-Tolerant Systems, pp. 152–166, Springer, 2004.

[32] J. M. Montenbruck and F. Allgöwer, “Some problems arising in controller design from big data via input-output methods,” in 55th Conference on Decision and Control, pp. 6525–6530, IEEE, 2016.

[33] C. Novara, L. Fagiano, and M. Milanese, “Direct feedback control design for nonlinear systems,” Automatica, vol. 49, no. 4, pp. 849–860, 2013.