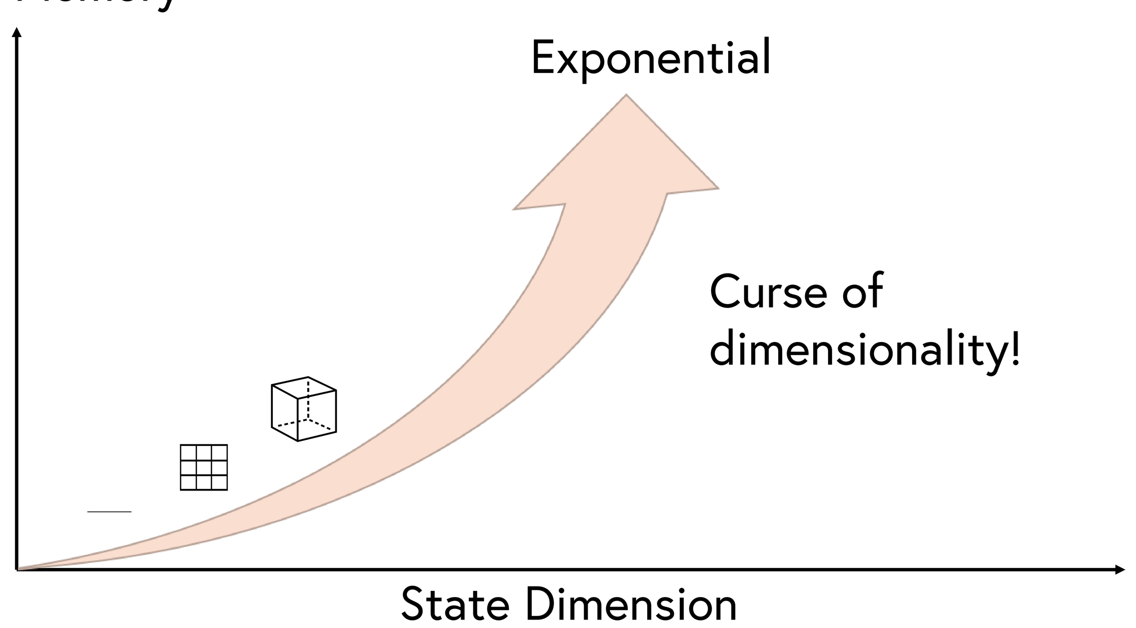
Generating Formal Safety Assurances for High-Dimensional Reachability

Albert Lin and Somil Bansal

Project Website

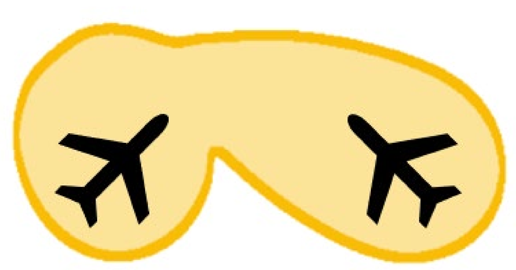
Motivation

Grid-based reachability methods are intractable.



Learning-based reachability methods are approximate.

Trained DeepReach Solution



No formal guarantees.

Main Goal

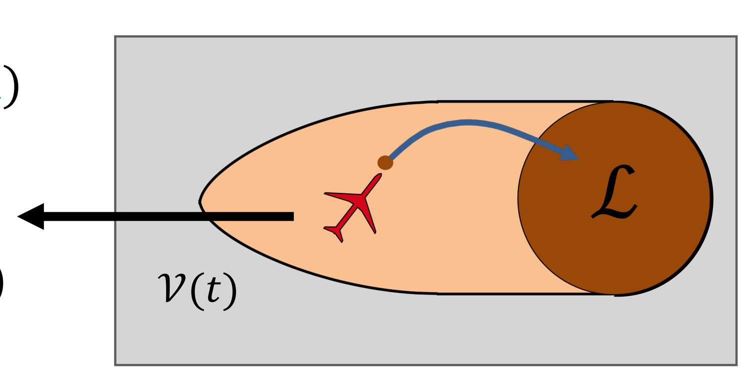
Compute formal safety guarantees for general nonlinear high-dimensional dynamical systems.

Background

Backward Reachable Tube

All states for which, for all possible control actions, the system state will reach a target set $\mathcal{L}$ at some time t within a time horizon T.

$\dot{x} = f(x, u)$



$$\mathrm{BRT} = \{x : x \in X, V(x,0) \leq 0\}$$

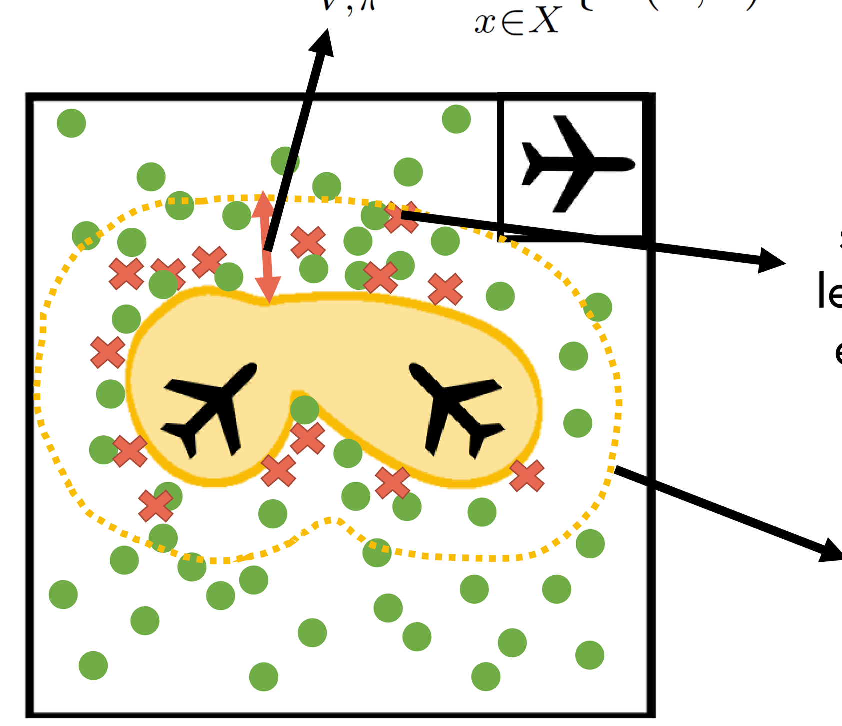
$$V(x,t) = \sup_{u(\cdot)} J_{u(\cdot)}(x,t)$$

$$u^*(x,t) = \arg\max_u \langle \nabla V(x,t), f(x,u) \rangle$$

Main Contributions

Uniform Error Correction

$$\delta_{\tilde{V},\tilde{\pi}} := \max_{x \in X}\{\tilde{V}(x,0) : J_{\tilde{\pi}}(x,0) \leq 0\}$$



Computing a Probabilistic Error Bound

Violation rate $1 - \epsilon \in (0,1)$

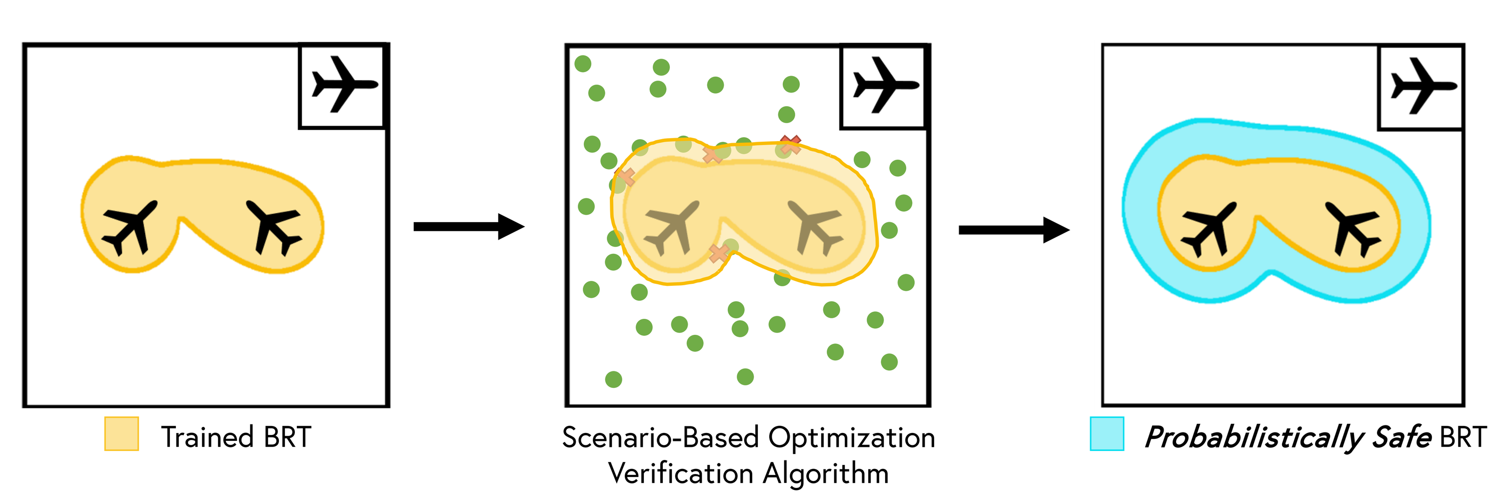
Confidence $1 - \beta \in (0,1)$

$$N \geq \frac{2}{\epsilon}\left(\ln \frac{1}{\beta} + 1\right)$$

Algorithm 1: Scenario Optimization Verification

Require: $X,\ N,\ M,\ \tilde{V}(x,0),\ J_{\tilde{\pi}}(x,0)$

1: $\delta_0 \leftarrow -\infty$
2: **for** $i = 0, 1, \ldots, M-1$ **do**
3: $\quad \mathcal{D}_i \leftarrow$ Sample N states IID from $\{x : x \in X, \tilde{V}(x,0) > \delta_i\}$
4: $\quad$ **if** $\exists x \in \mathcal{D}_i : J_{\tilde{\pi}}(x,0) \leq 0$ **then**
5: $\quad\quad \delta_i \leftarrow \max_{x \in \mathcal{D}_i}\{\tilde{V}(x,0) : J_{\tilde{\pi}}(x,0) \leq 0\}$
6: $\quad$ **else**
7: $\quad\quad$ **break**
8: $\quad$ **end if**
9: **end for**
10: **return** $\hat{\delta} := \delta_i$



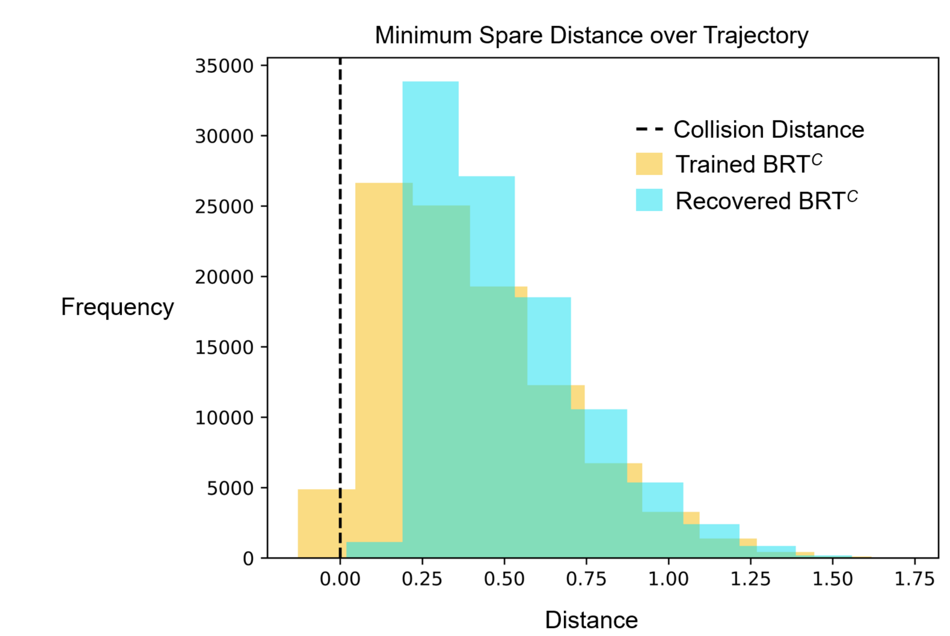
PRINCETON UNIVERSITY

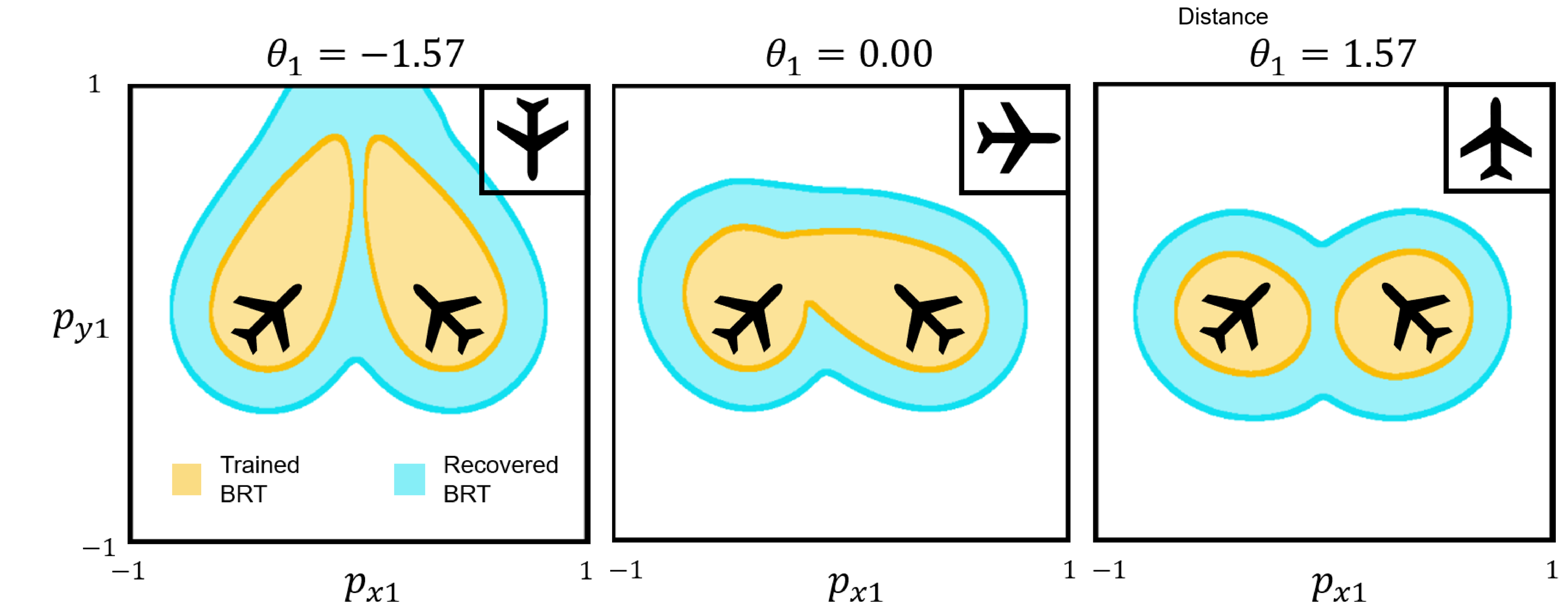
USC University of Southern California

Results

Multi-Vehicle Collision Avoidance System

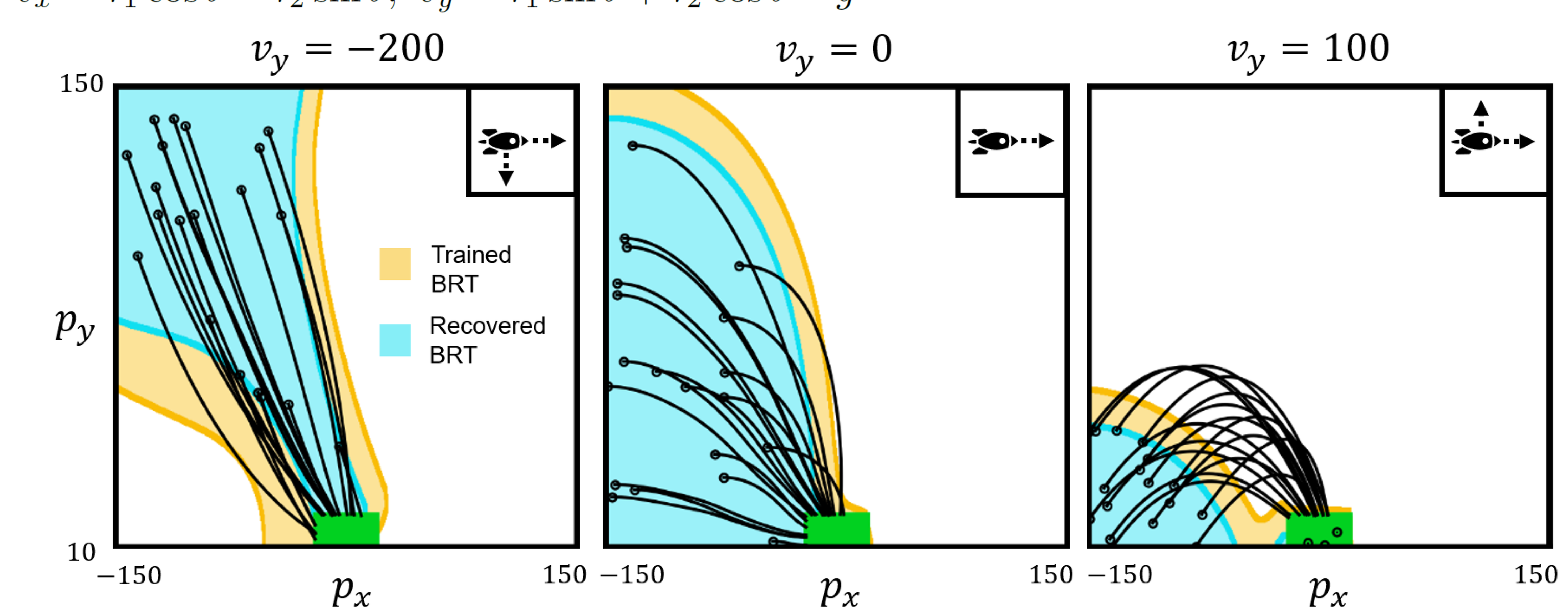
$$\dot{p_{xi}} = v\cos\theta_i$$
$$\dot{p_{yi}} = v\sin\theta_i$$
$$\dot{\theta}_i = u_i$$





Rocket Landing System

$\dot{p_x} = v_x,\ \dot{p_y} = v_y,\ \theta = \omega,\ \dot{\omega} = 0.3\tau_1,$
$\dot{v_x} = \tau_1\cos\theta - \tau_2\sin\theta,\ \dot{v_y} = \tau_1\sin\theta + \tau_2\cos\theta - g$



Exploring a More Nuanced Approach

Value Function Safety Errors for Toy Dubins Car Avoid System