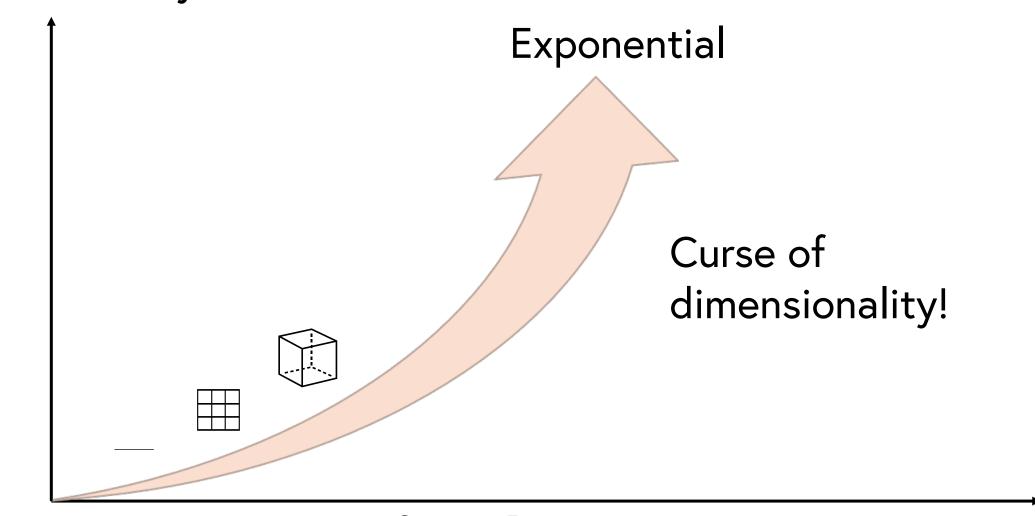


Motivation

Reachability analysis is a powerful tool for synthesizing safety controllers for autonomous systems.

Grid-based reachability methods are intractable. Time/Memory



State Dimension

Learning-based reachability methods are approximate. Trained DeepReach Solution



No formal guarantees.

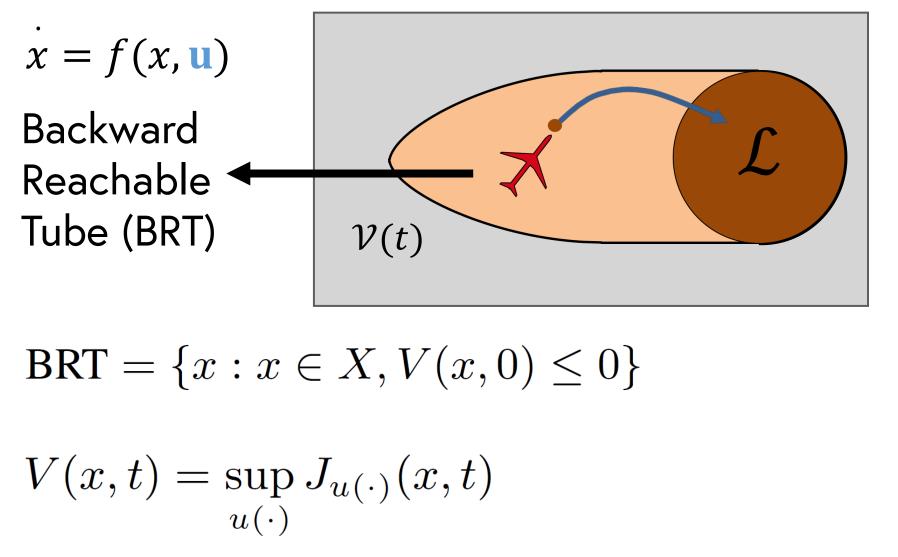
Main Goal

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.

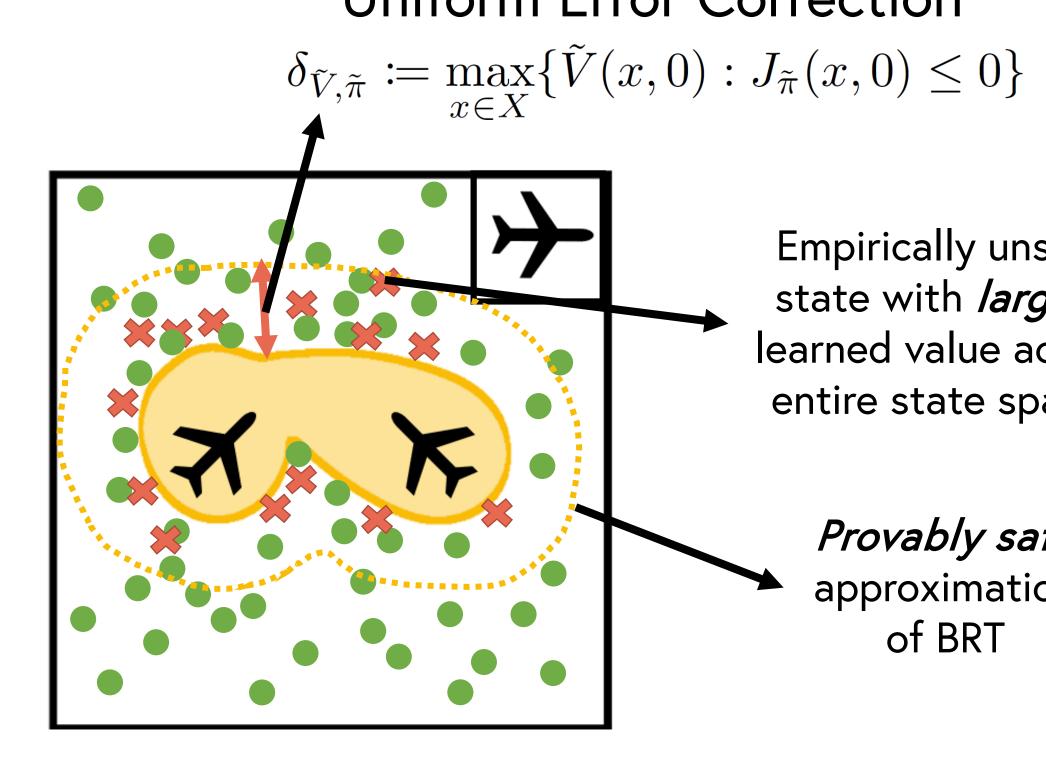


```
u^*(x,t) = \arg\max\langle \nabla V(x,t), f(x,u) \rangle
```

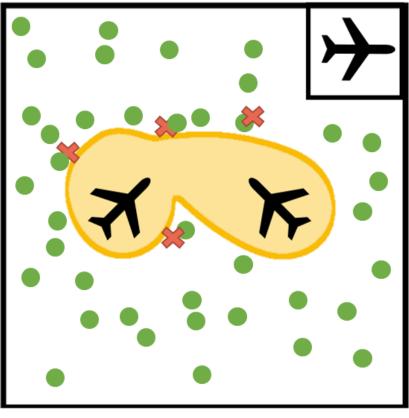
Generating Formal Safety Assurances for High-Dimensional Reachability

Albert Lin and Somil Bansal

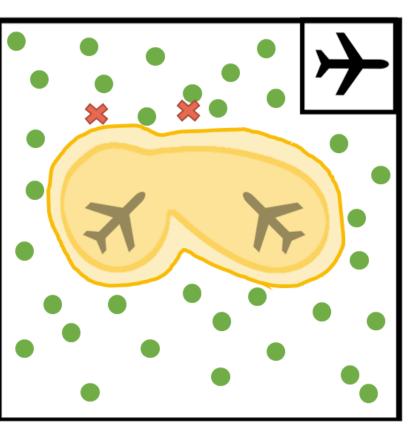
Main Contributions



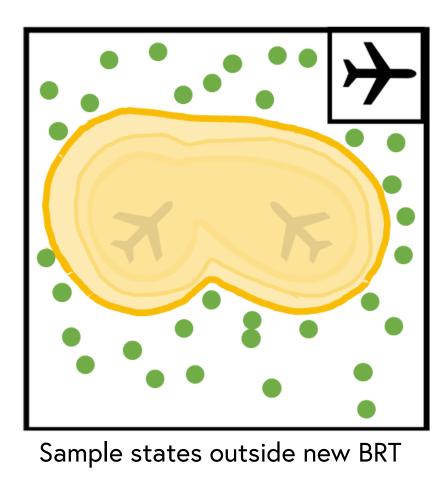
Computing a Probabilistic Error Bound



Sample states outside trained BRT



Sample states outside new BRT





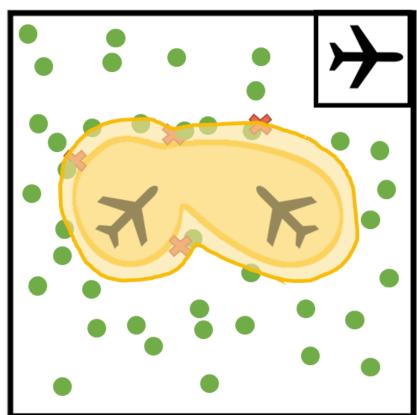
PRINCETON UNIVERSITY



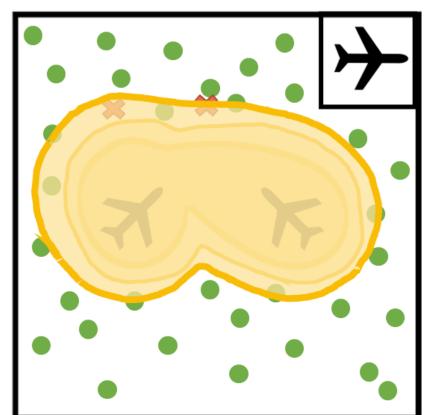
Uniform Error Correction

Empirically unsafe state with *largest* learned value across entire state space.

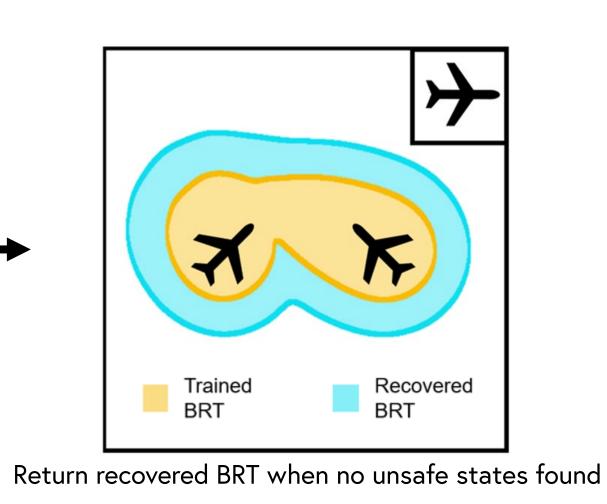
Provably safe approximation of BRT



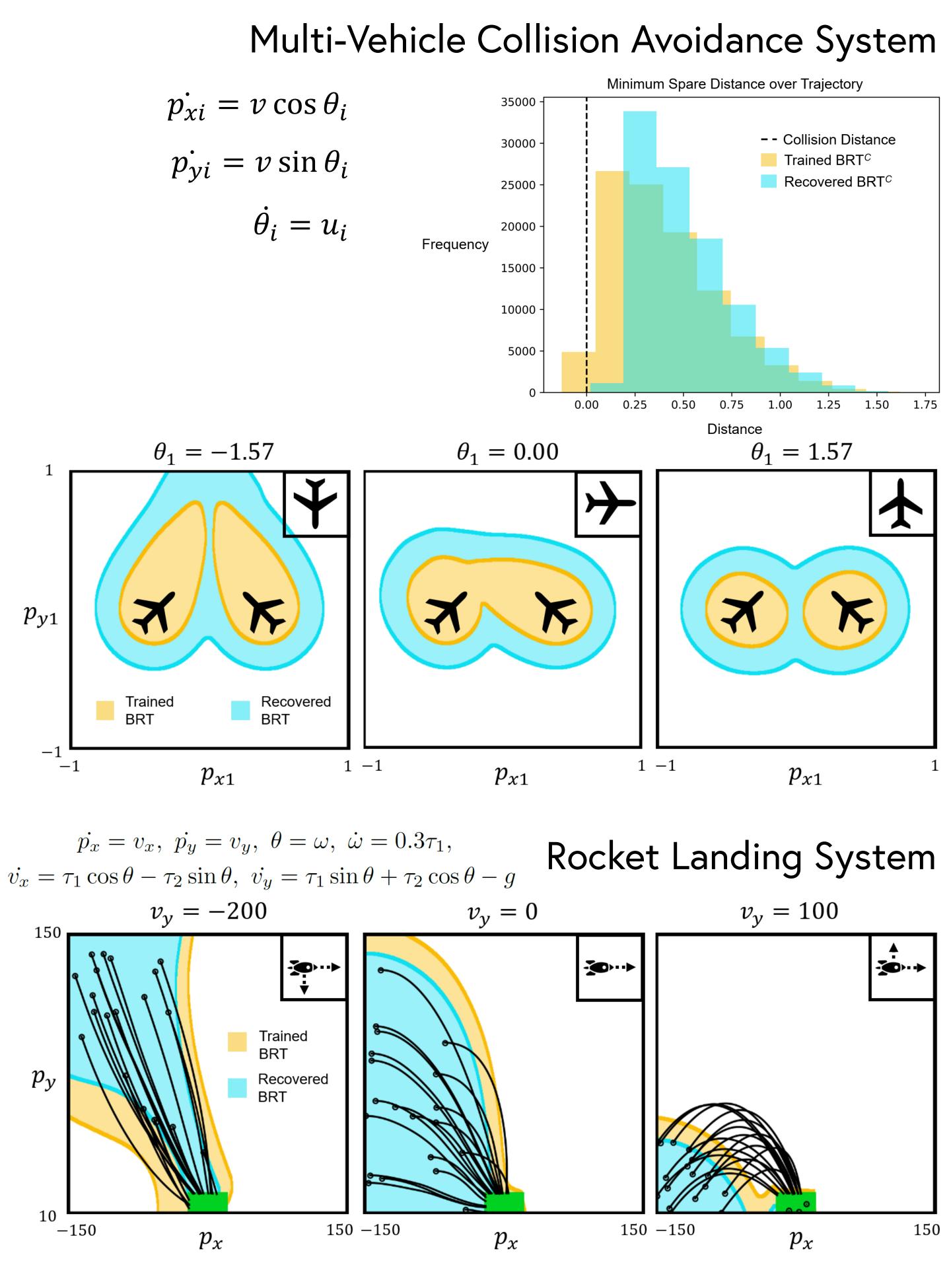
Expand BRT to capture unsafe states

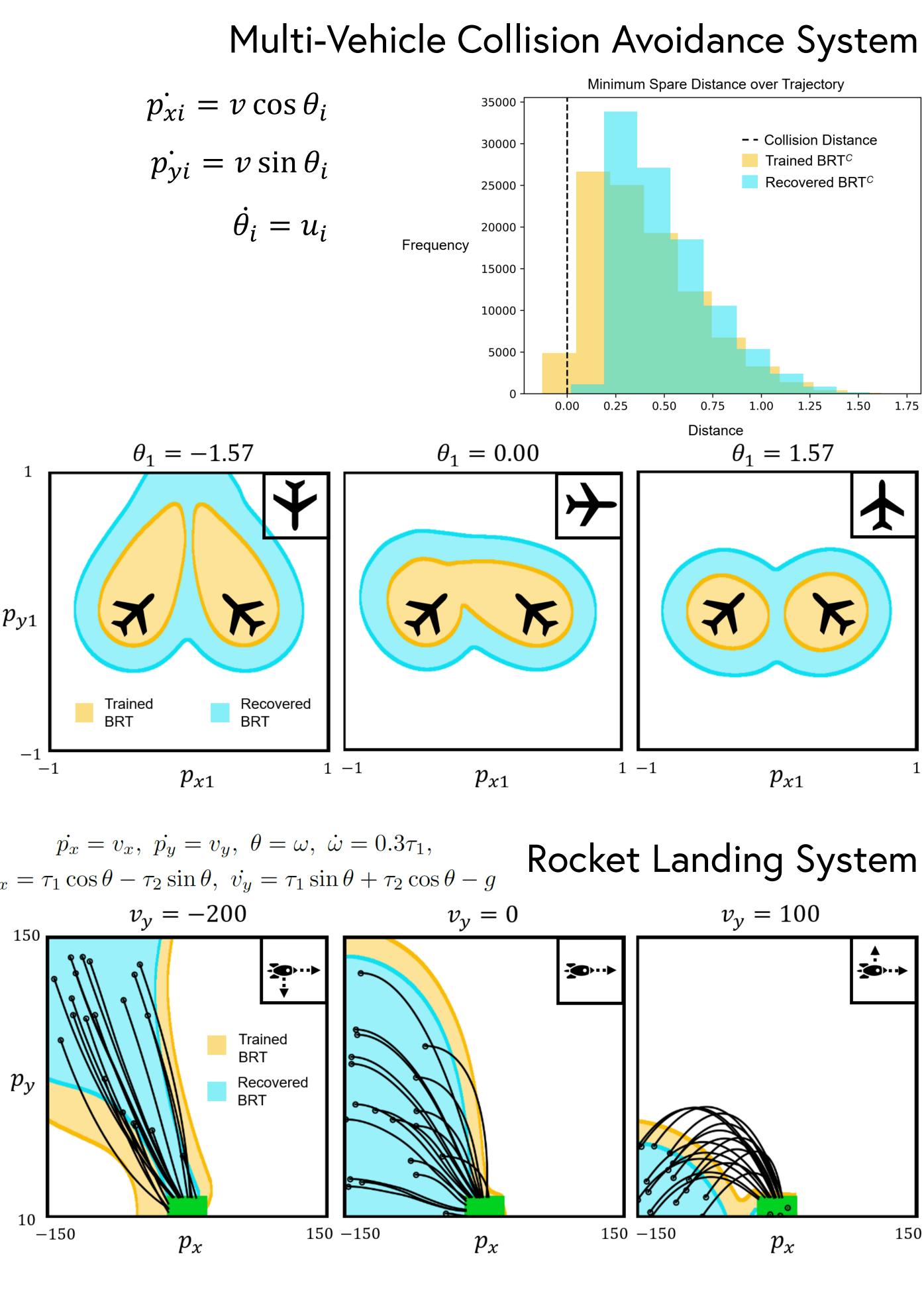


Expand BRT to capture unsafe states

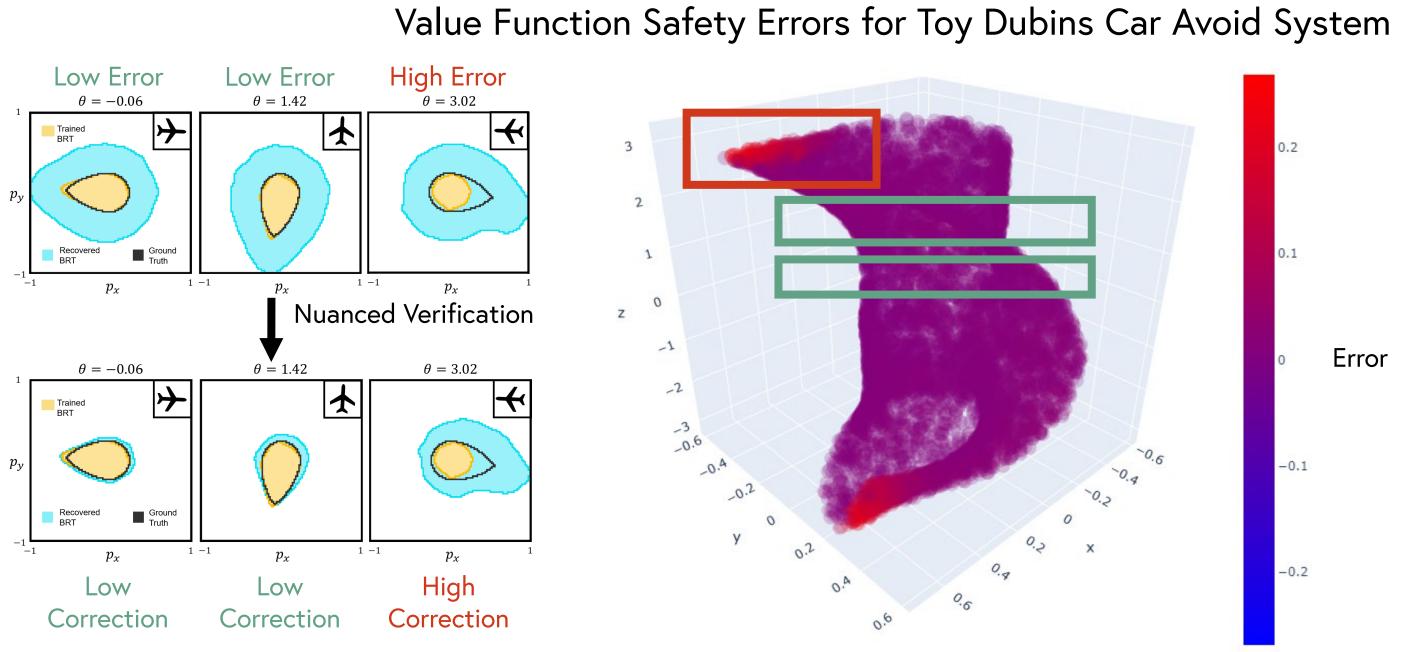








Exploring a More Nuanced Approach





Results