Shielding for Hybrid Systems

Christian Schilling

joint work with Asger Horn Brorholt,
Andreas Holck Høeg-Petersen, Peter Gjøl Jensen,
Kim Guldstrand Larsen, and Florian Lorber

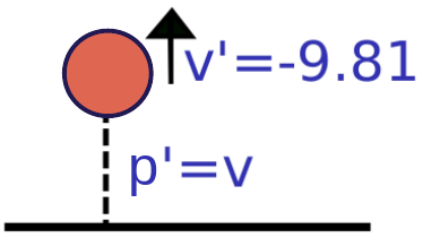
May 13, 2025

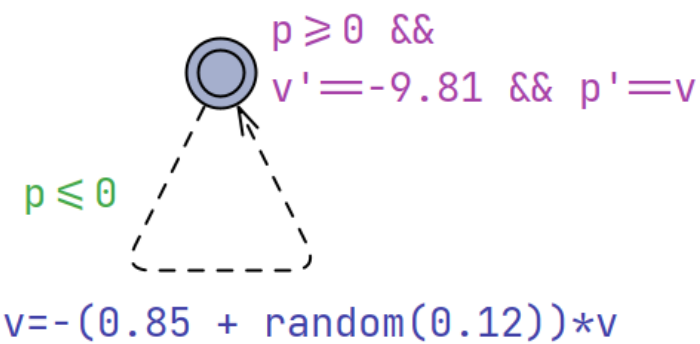
AALBORG UNIVERSITET

Overview

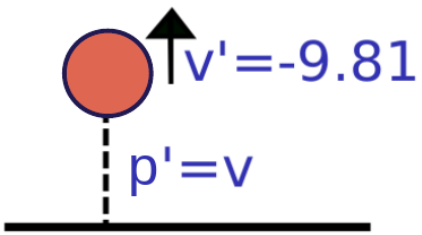
Control of physical systems

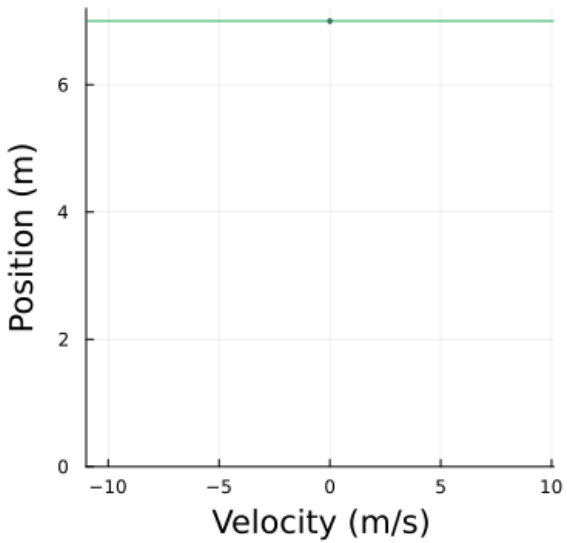
- **Physical systems have complex dynamics**
  - Continuous evolution
  - Discrete events
  - Stochastic uncertainty

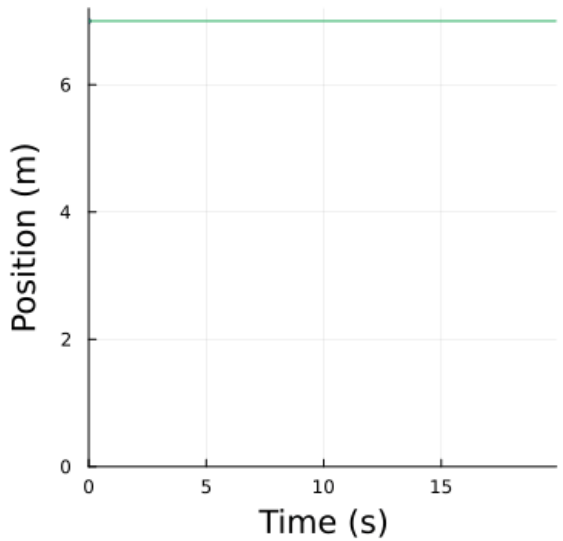


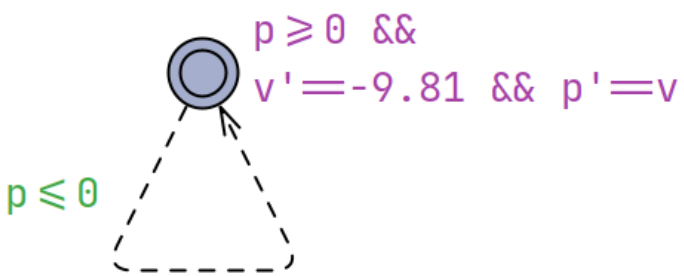


Control of physical systems



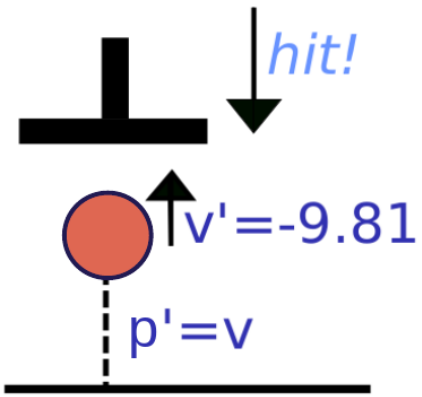




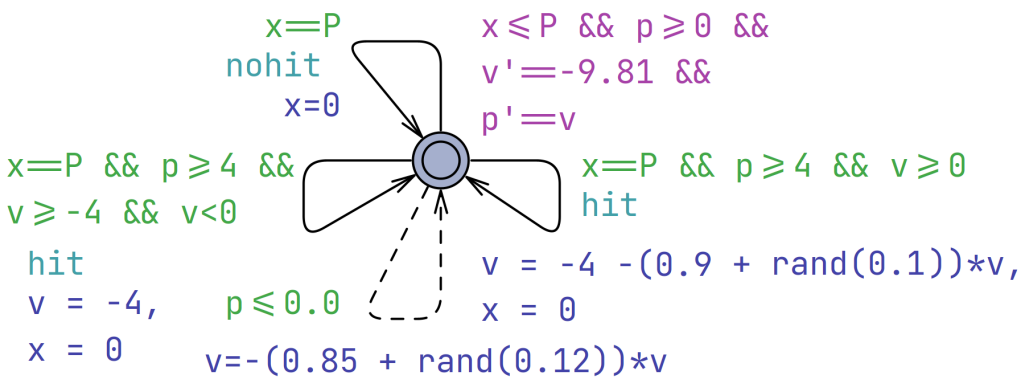



```
v=-(0.85 + random(0.12))*v
```

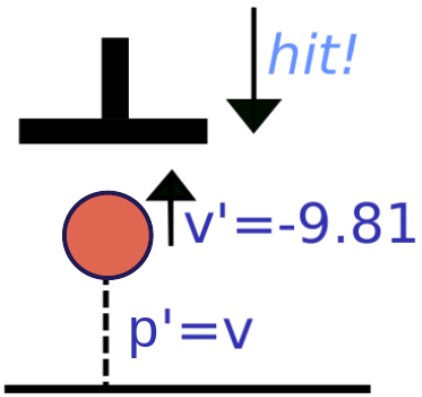
Control of physical systems

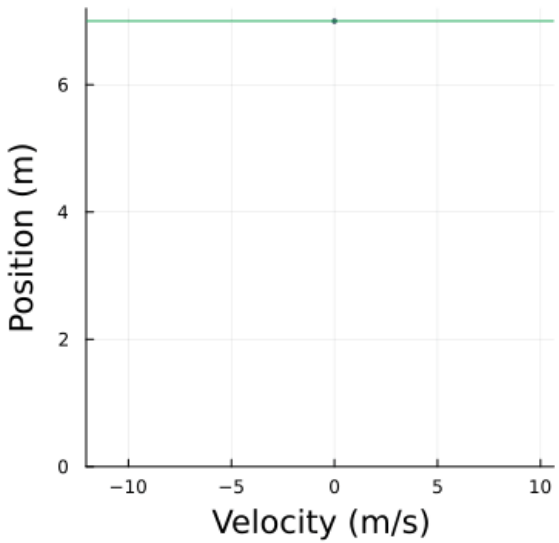


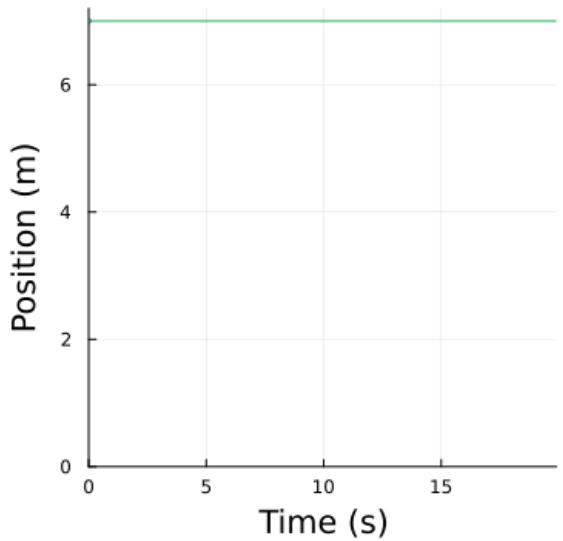
- Physical systems have complex dynamics
  - Continuous evolution
  - Discrete events
  - Stochastic uncertainty
- **Goal: control subject to some optimality criterion, e.g., minimize number of hits**

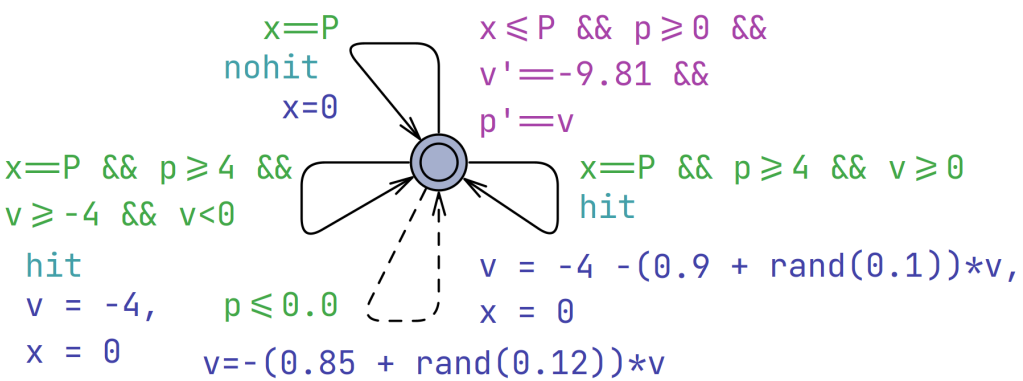


Control of physical systems

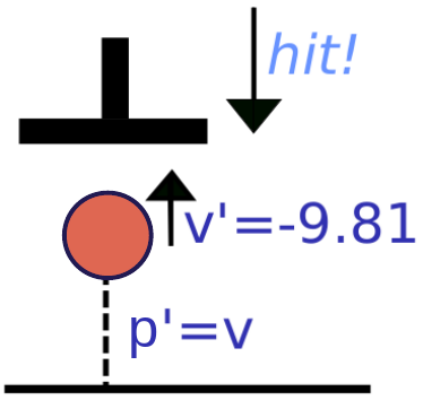




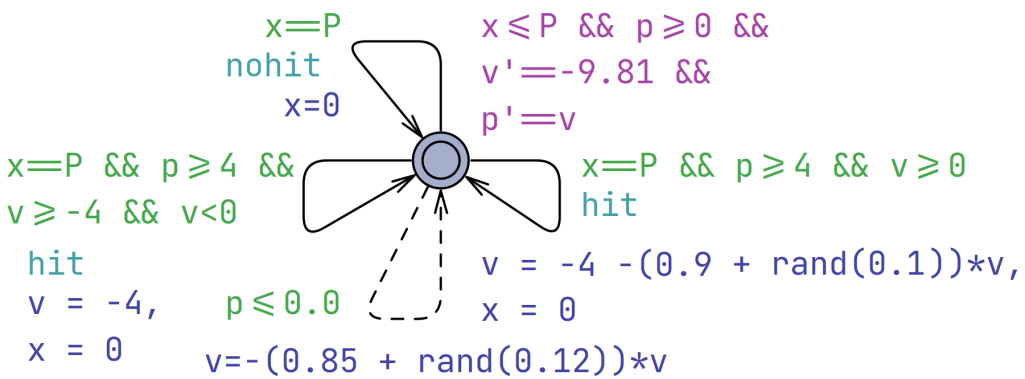




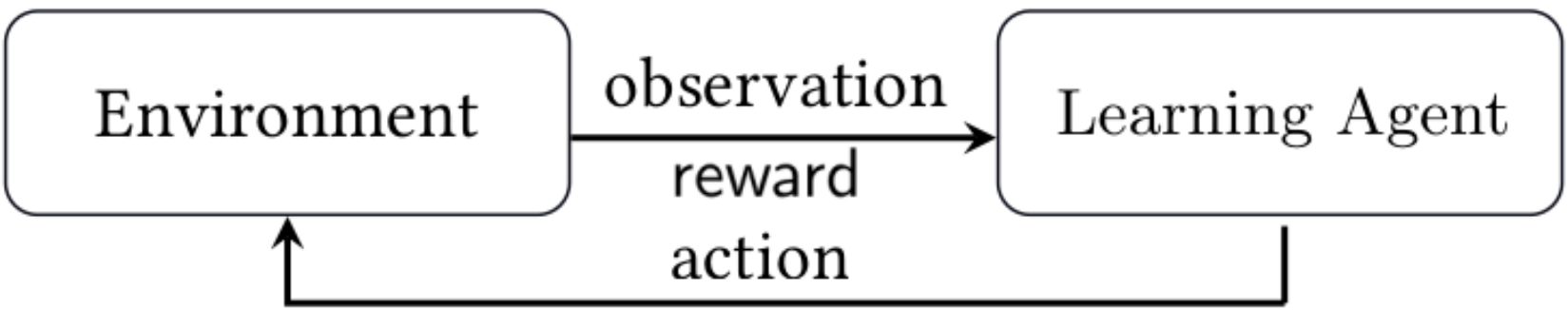
Control of physical systems



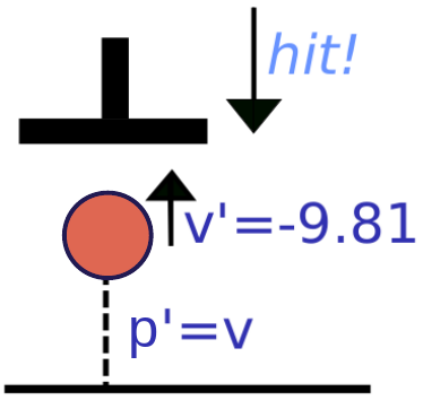
- Physical systems have complex dynamics
  - Continuous evolution
  - Discrete events
  - Stochastic uncertainty
- Goal: control subject to some optimality criterion, e.g., minimize number of hits
- **We can reinforcement-learn a controller, e.g., with Uppaal Stratego**

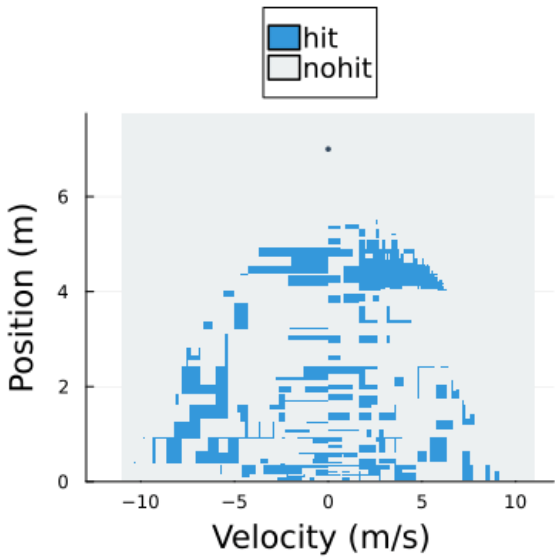


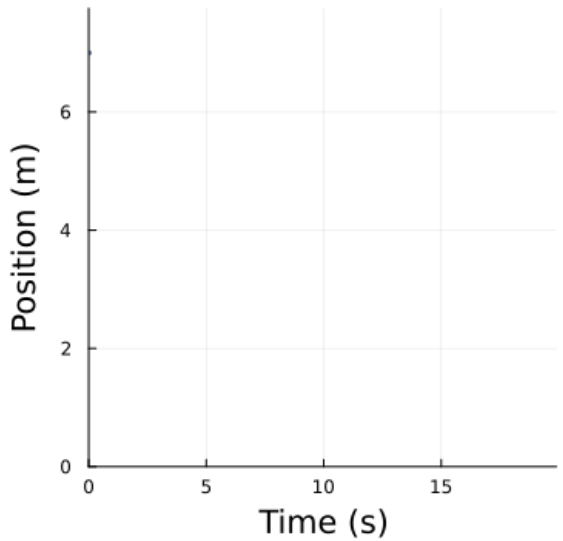
Reinforcement learning



Control of physical systems

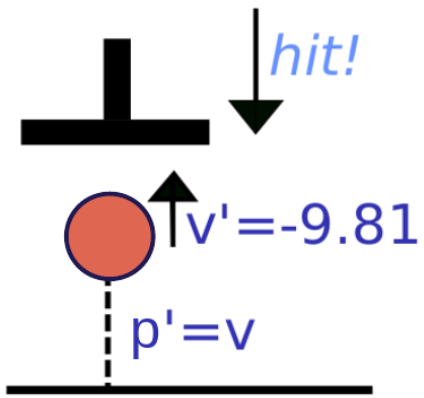






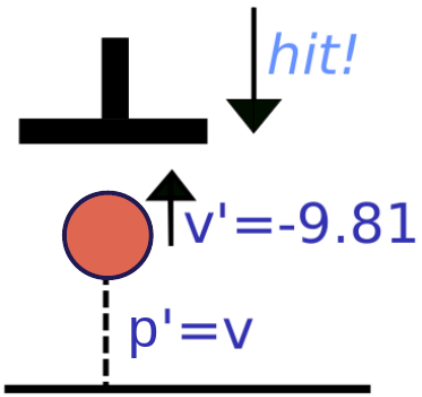
Trained for 12,000 episodes

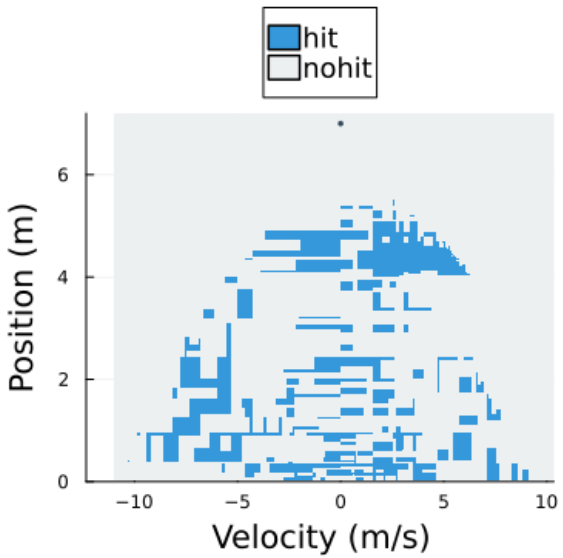
Control of physical systems

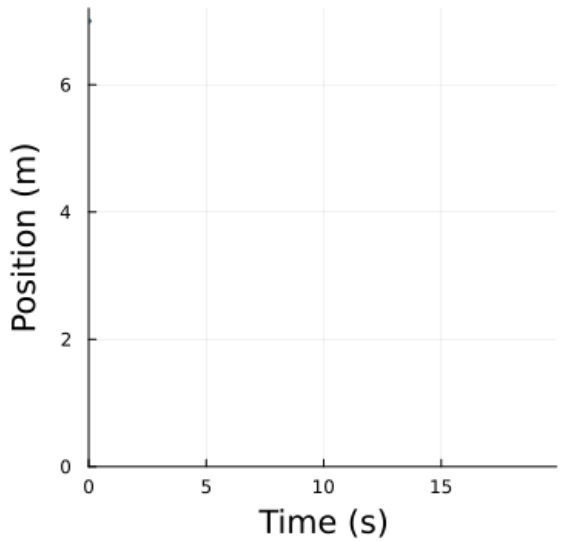


- Physical systems have complex dynamics
  - Continuous evolution
  - Discrete events
  - Stochastic uncertainty
- Goal: control subject to some optimality criterion, e.g., minimize number of hits
- We can reinforcement-learn a controller, e.g., with Uppaal Stratego
- **We also have safety constraints, e.g.,** $p = 0 \implies |v| > 1$

Control of physical systems



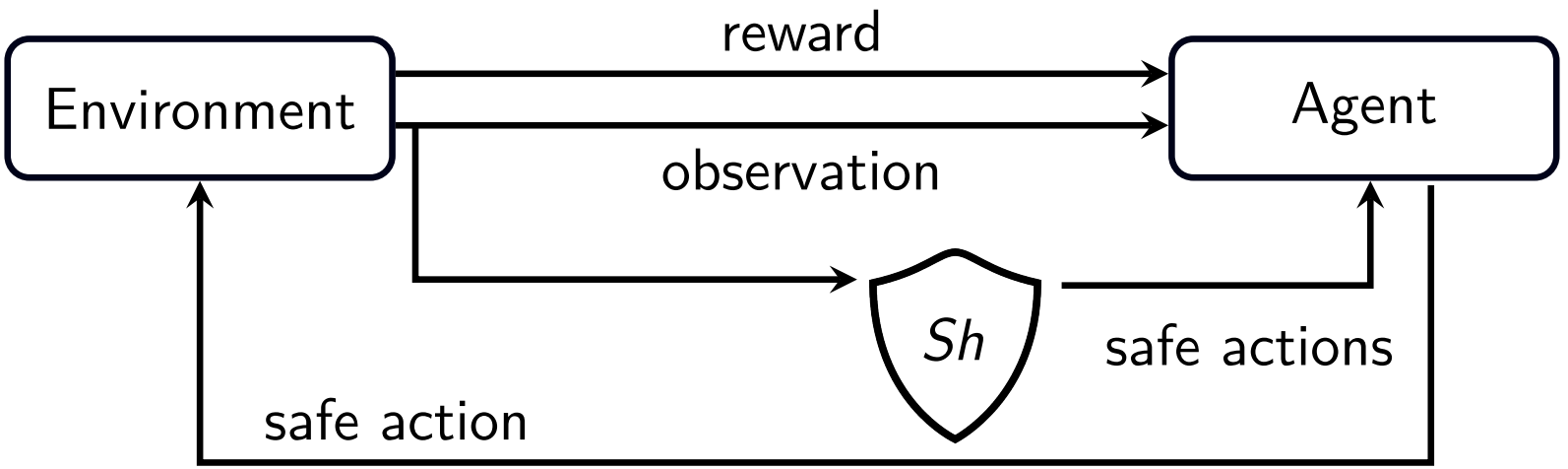




2 % of executions unsafe

Overview

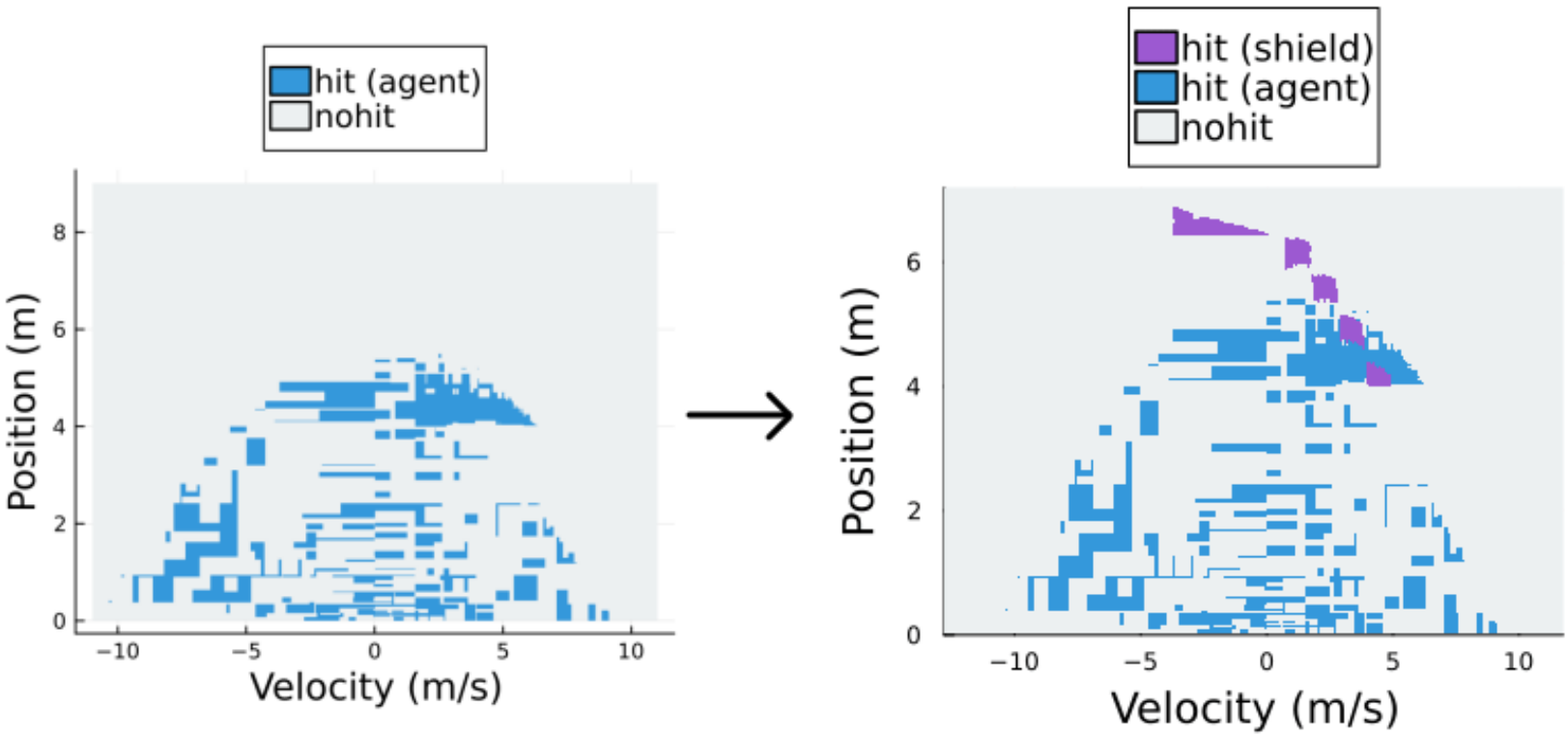
Shielding[1]



- Shield = nondeterministic safe policy
- Under a shield, *any* controller/agent is safe

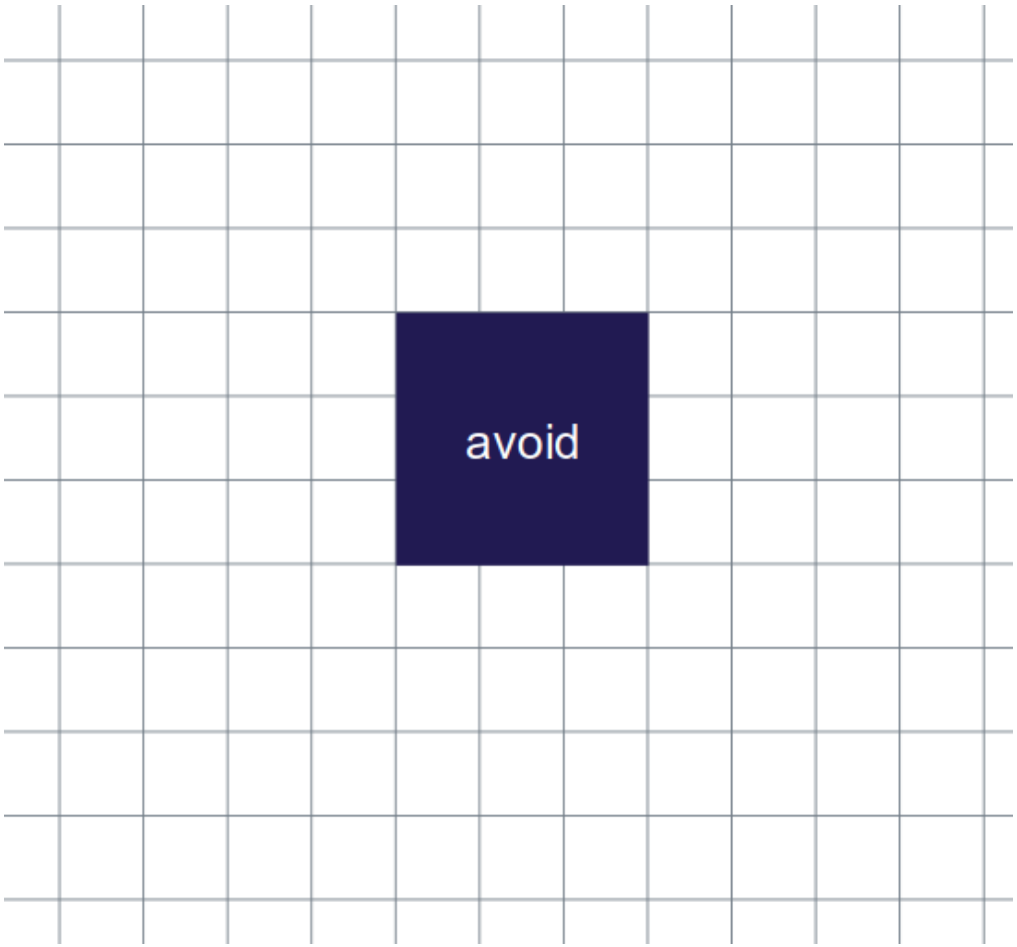
[1] Alshiekh et al. *AAAI*. 2018.

Shielding via state-space partitioning[1]



[1] Brorholt, Jensen, Larsen, Lorber, and Schilling. *AISoLA*. 2023.

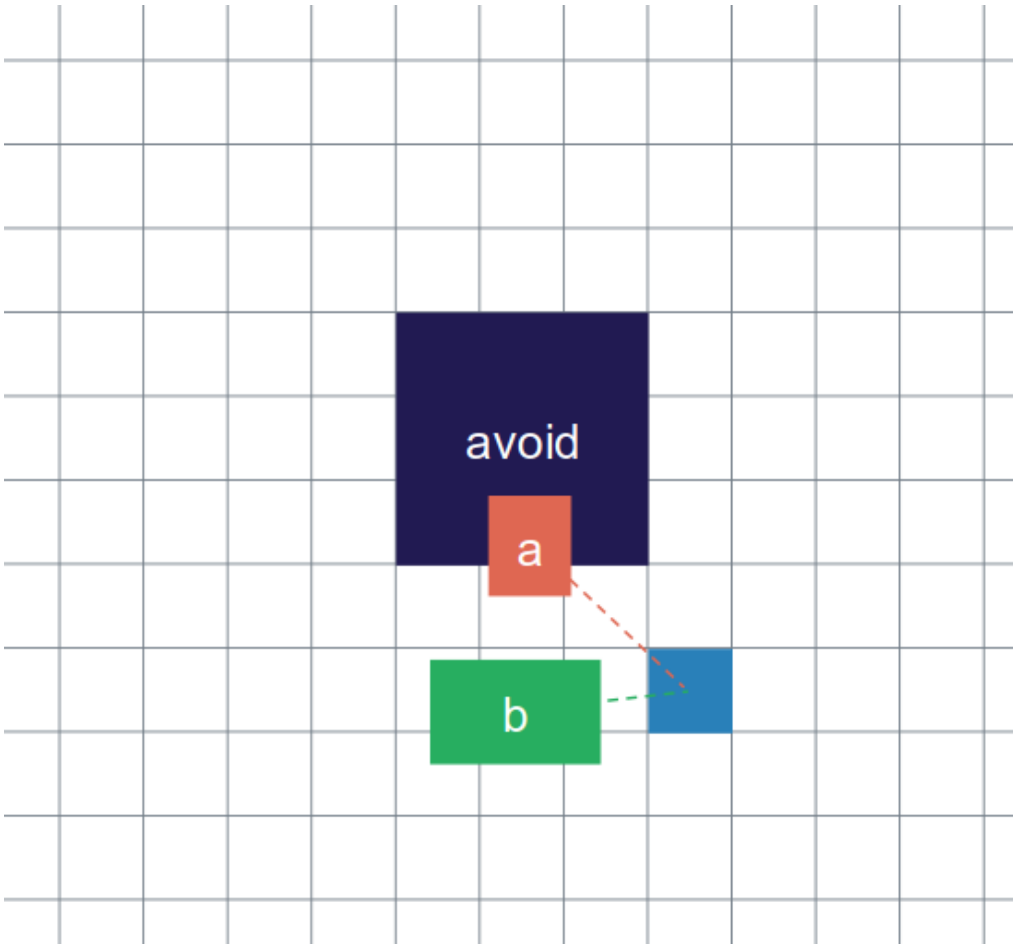
State-space partitioning and two-player game[1]



- Avoid set = blocks with unsafe states
- For each block, compute successor states under each action

[1] Brorholt, Jensen, Larsen, Lorber, and Schilling. *AISoLA*. 2023.

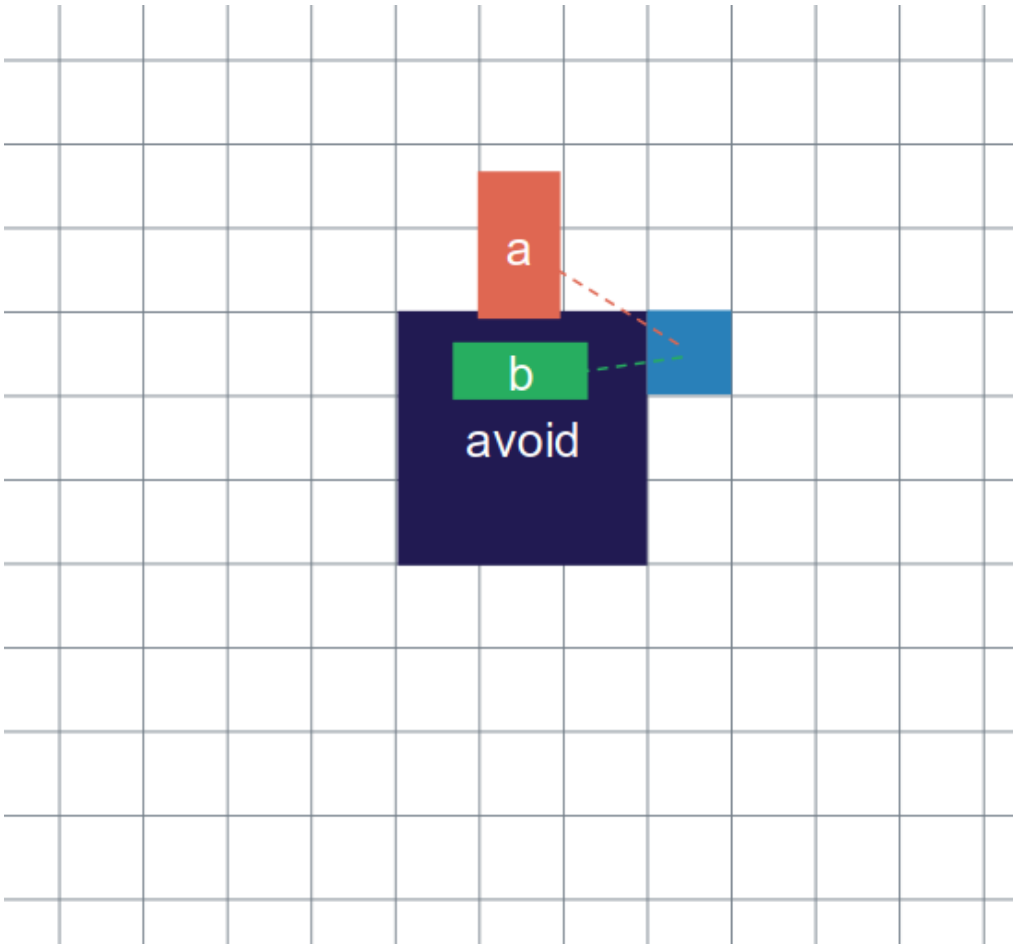
State-space partitioning and two-player game[1]



- Avoid set = blocks with unsafe states
- For each block, compute successor states under each action
- Forbid actions that reach the avoid set

[1] Brorholt, Jensen, Larsen, Lorber, and Schilling. *AISoLA*. 2023.

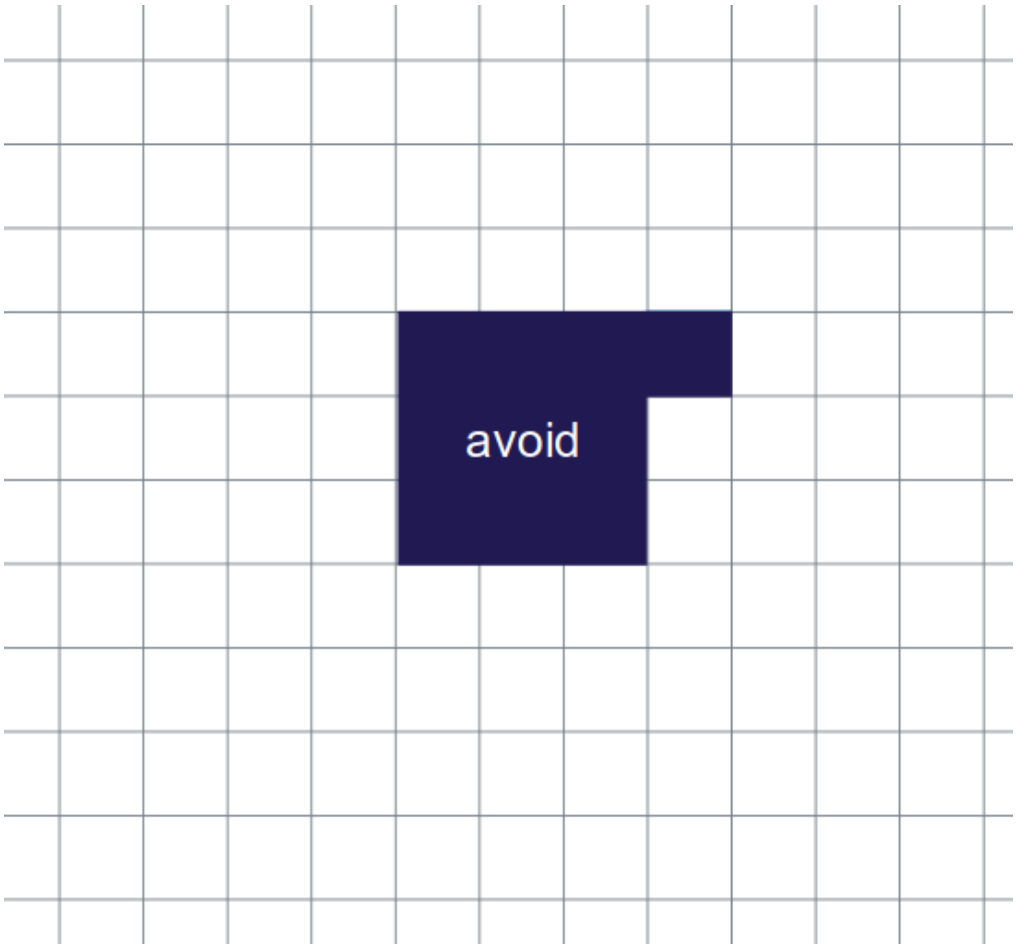
State-space partitioning and two-player game[1]



- Avoid set = blocks with unsafe states
- For each block, compute successor states under each action
- Forbid actions that reach the avoid set
- If all actions reach the avoid set,

[1] Brorholt, Jensen, Larsen, Lorber, and Schilling. *AISoLA*. 2023.

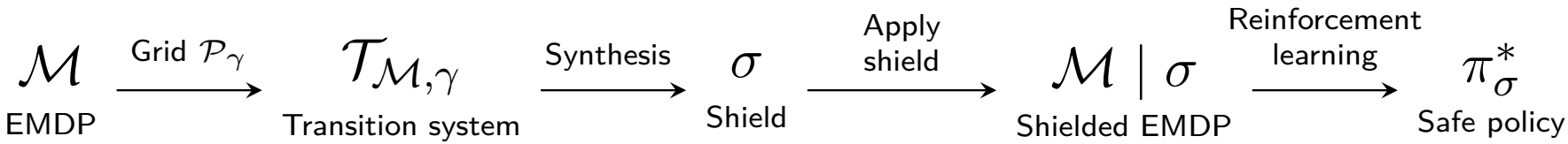
State-space partitioning and two-player game[1]



- Avoid set = blocks with unsafe states
- For each block, compute successor states under each action
- Forbid actions that reach the avoid set
- If all actions reach the avoid set, add the block to the avoid set

[1] Brorholt, Jensen, Larsen, Lorber, and Schilling. *AISoLA*. 2023.

State-space partitioning and two-player game[1]

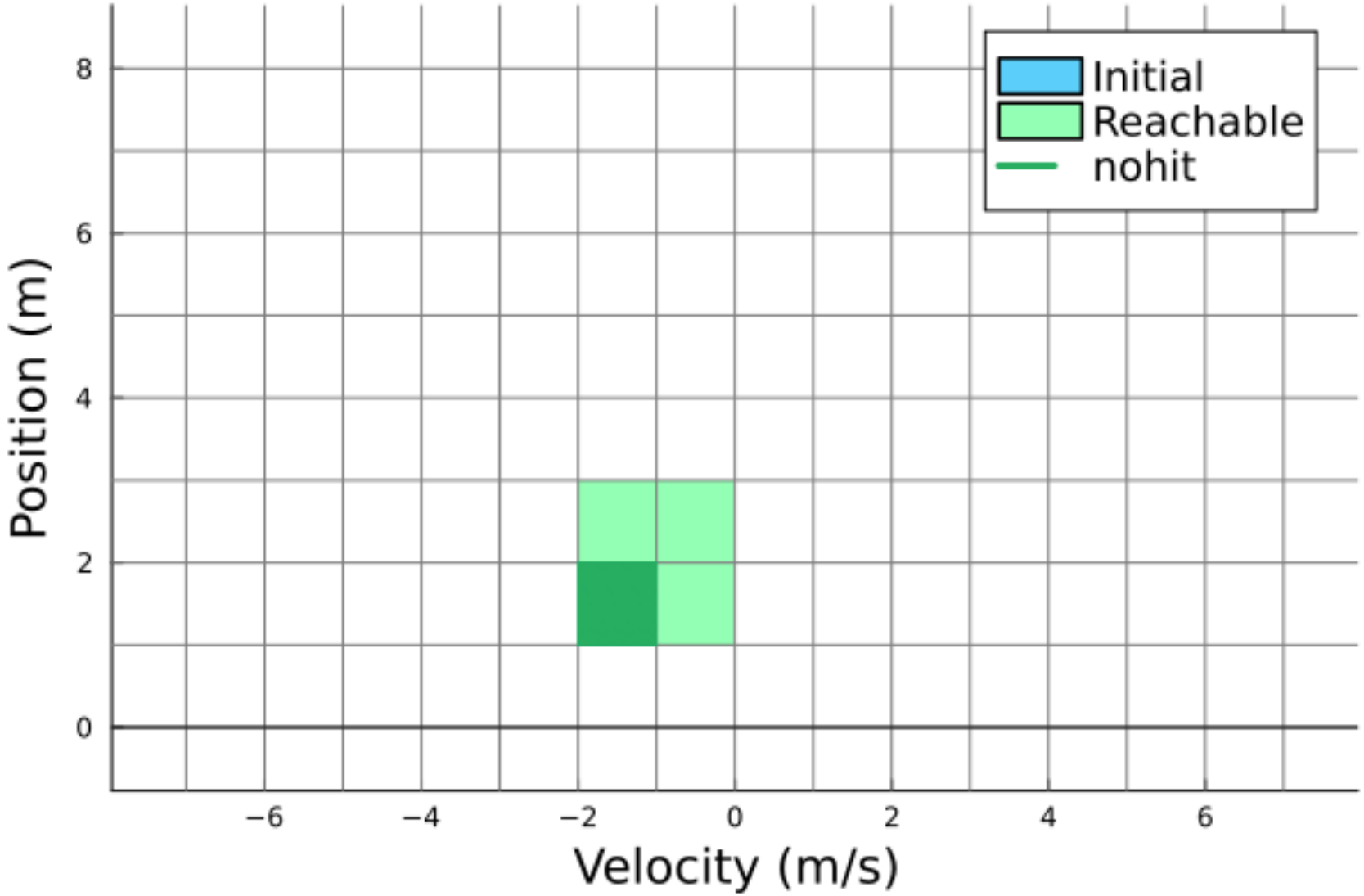


Theorem

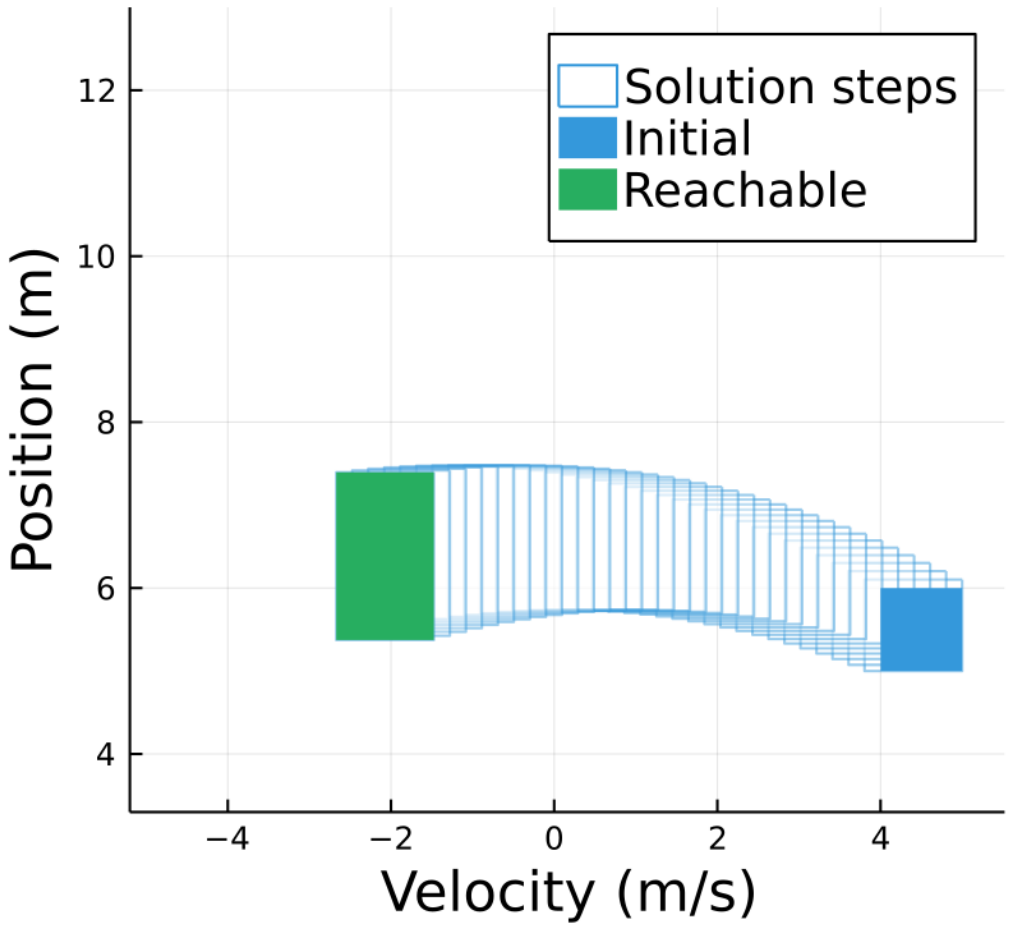
Any shield σ for $\mathcal{T}_{\mathcal{M},\gamma}$ is safe for $\mathcal{M}$

[1] Brorholt, Jensen, Larsen, Lorber, and Schilling. *AISoLA*. 2023.

How to compute reachable blocks for complex systems?

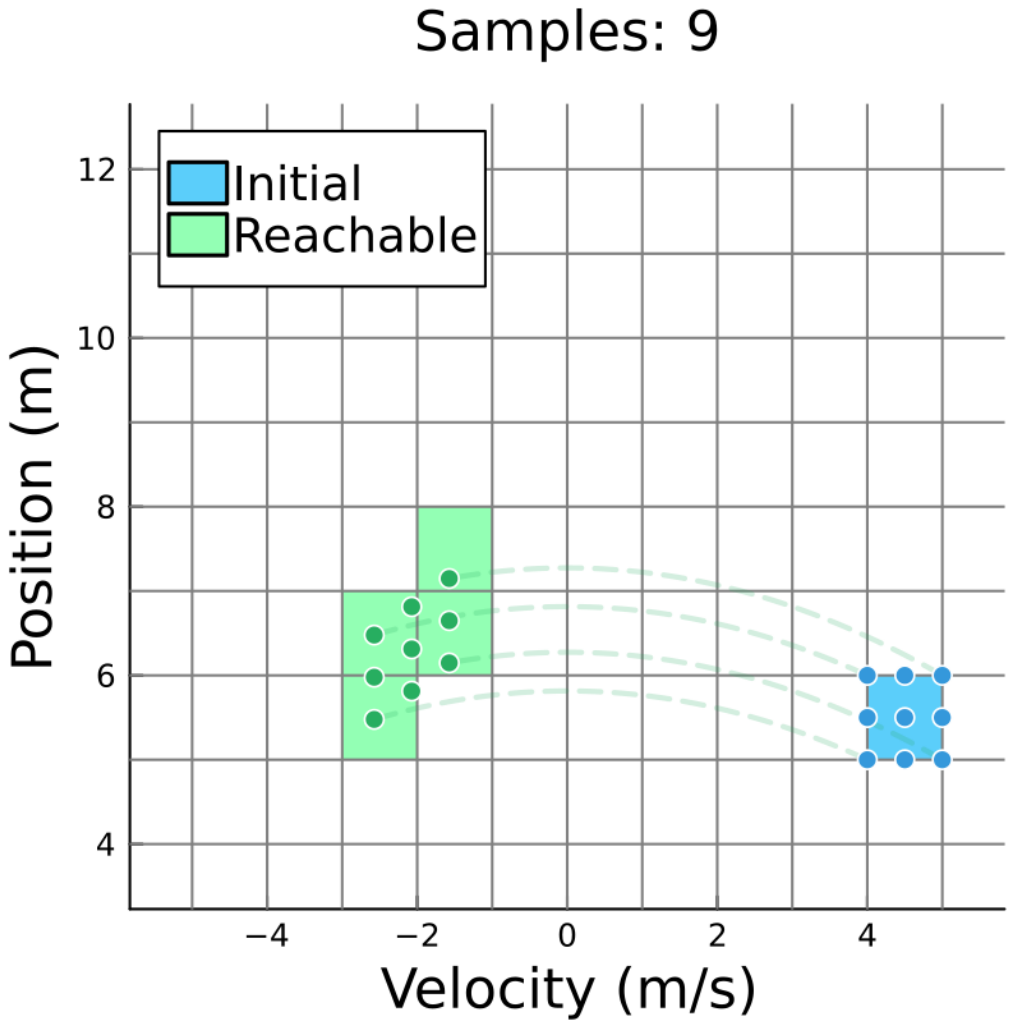


Computation via reachability tool (here: JuliaReach)



- Sound result
- Expensive

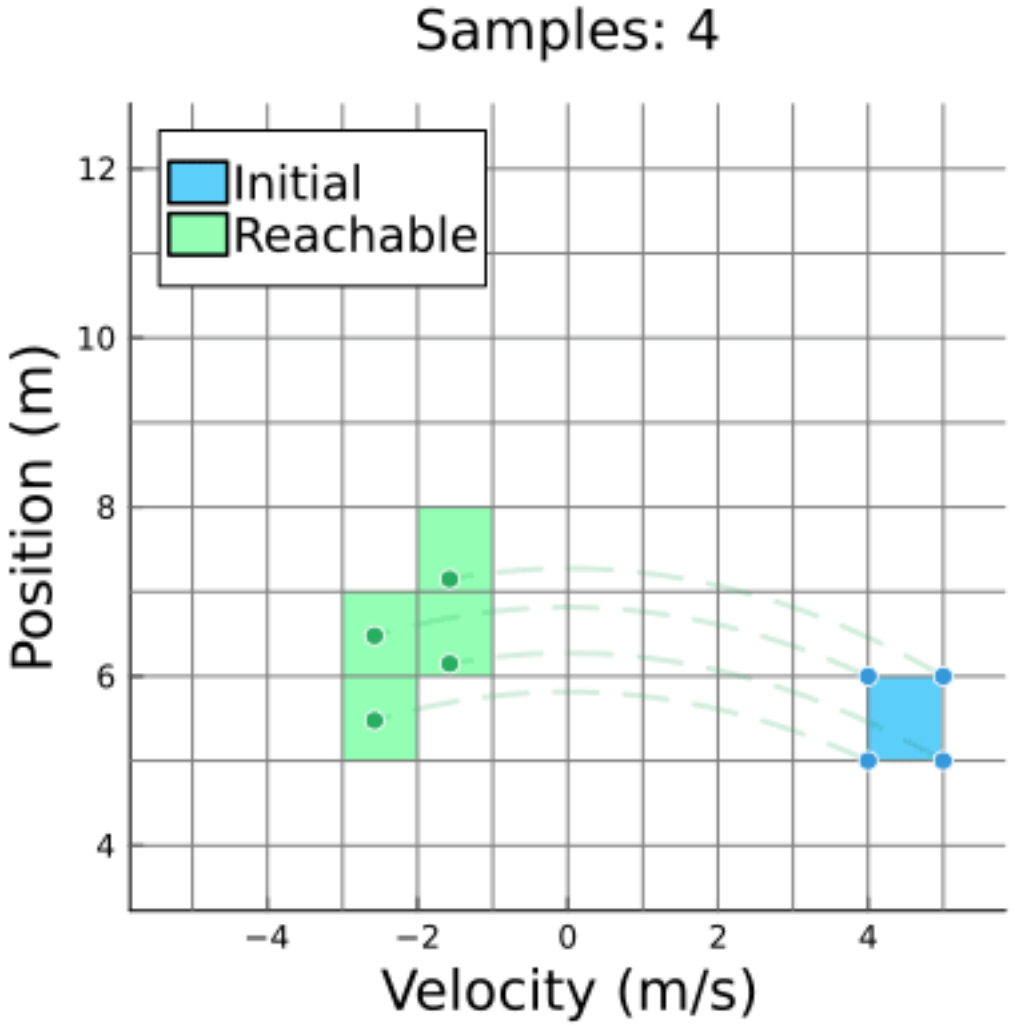
Computation via simulation-based method[1]



- Cheap
- May miss behavior
- Accuracy improves with more samples

[1] Brorholt, Jensen, Larsen, Lorber, and Schilling. *AISoLA*. 2023.

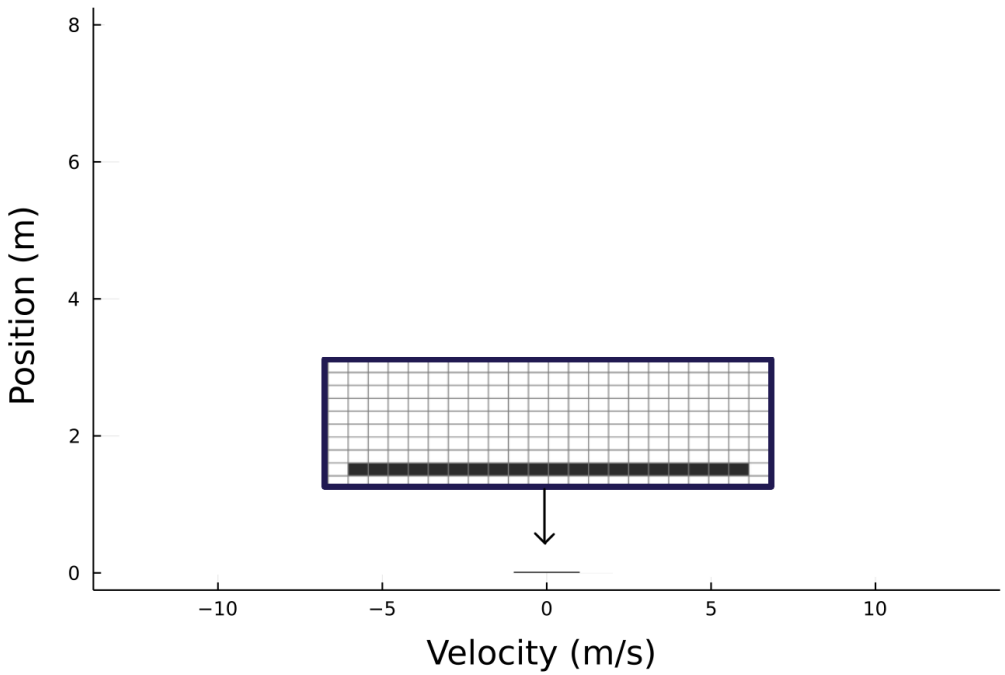
Computation via simulation-based method[1]



- Cheap
- May miss behavior
- Accuracy improves with more samples

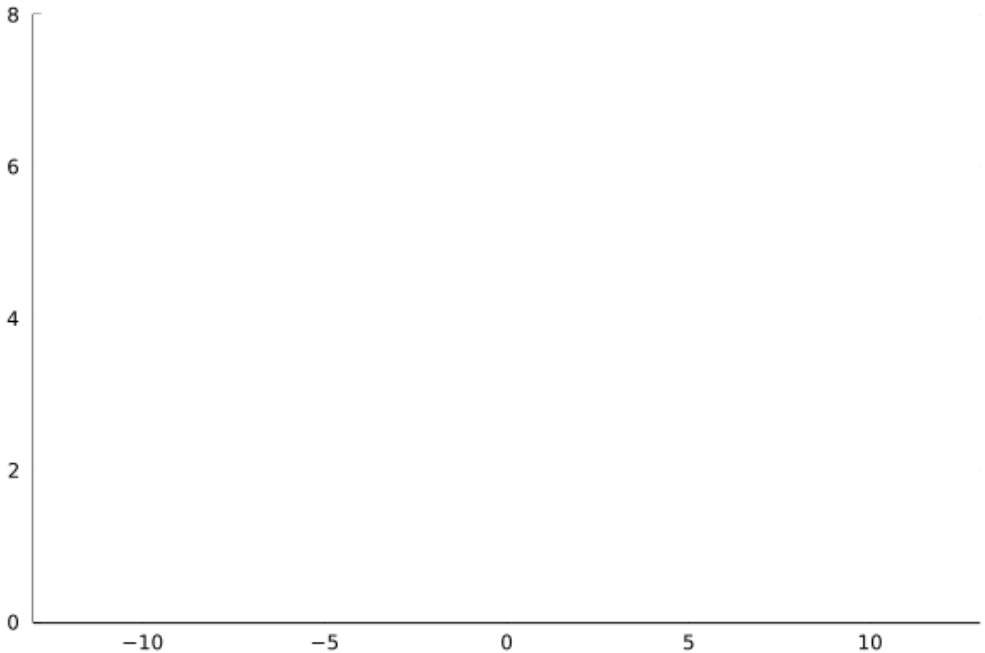
[1] Brorholt, Jensen, Larsen, Lorber, and Schilling. *AISoLA*. 2023.

Synthesis algorithm in action



- 16 samples per block
- Cell diameter 0.02
- 520,000 cells
- Time: 134 sec

Synthesis algorithm in action



- 16 samples per block
- Cell diameter 0.02
- 520,000 cells
- Time: 134 sec

Overview

Scalability

Grid size	Samples	Time
0.02	4	2m 14s
0.02	8	4m 02s
0.02	16	11m 03s
0.01	4	19m 00s
0.01	8	27m 21s
0.01	16	56m 32s

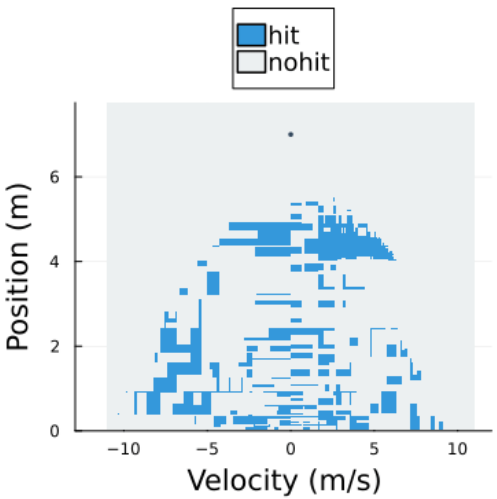
← Sampling-based

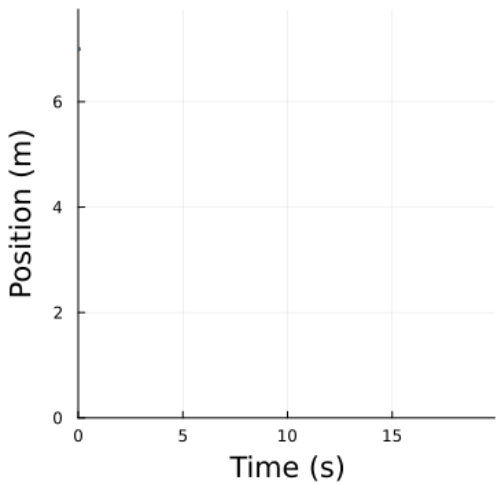
- Statistically safe ($\geq 99.99\,\%$ with confidence $99\,\%$)

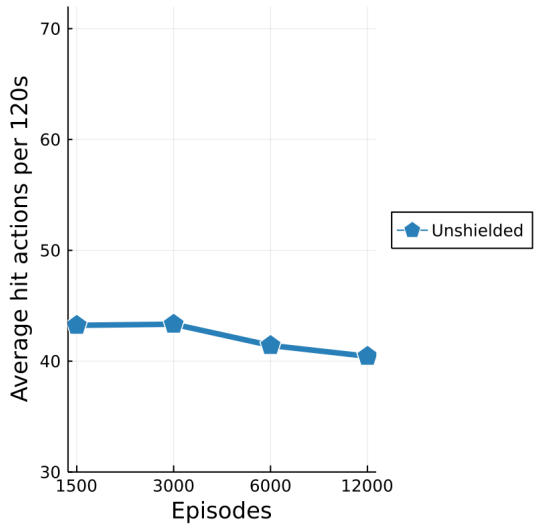
Grid size	Time
0.01	41h 05m

← Reachability-based

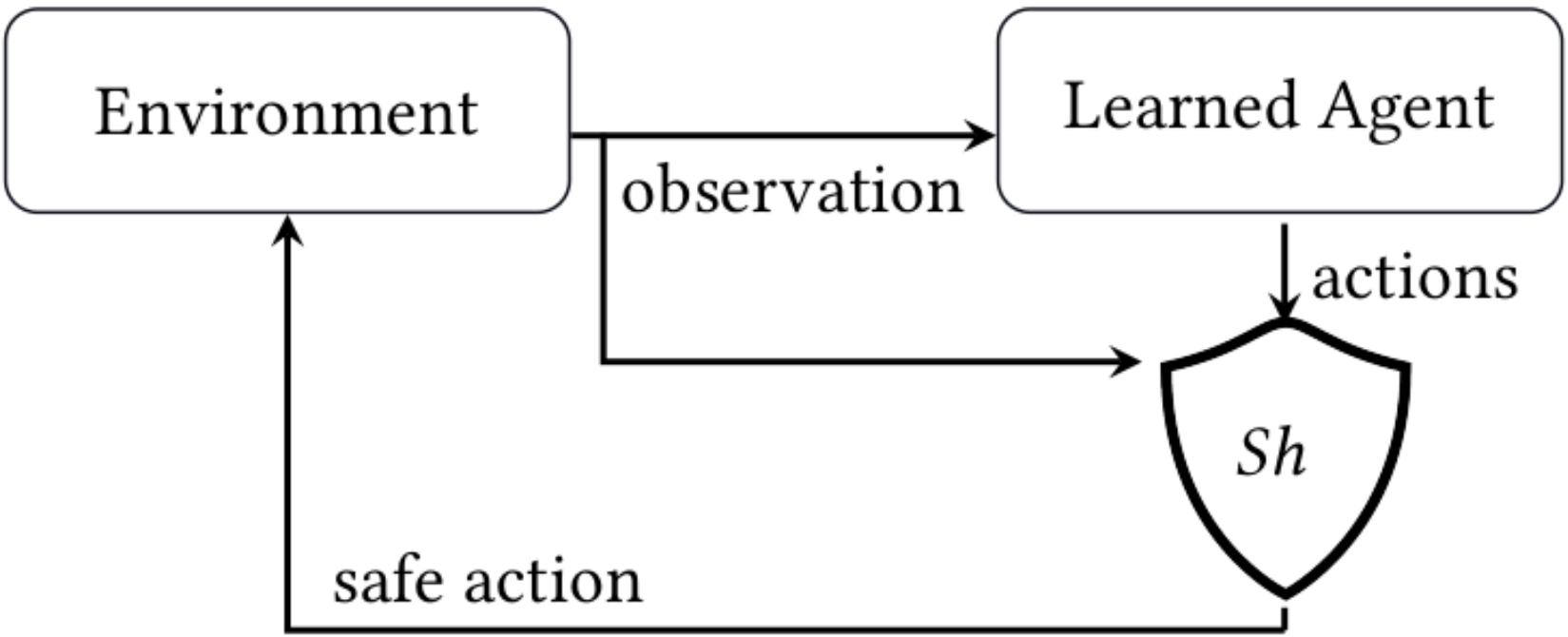
Unshielded agent



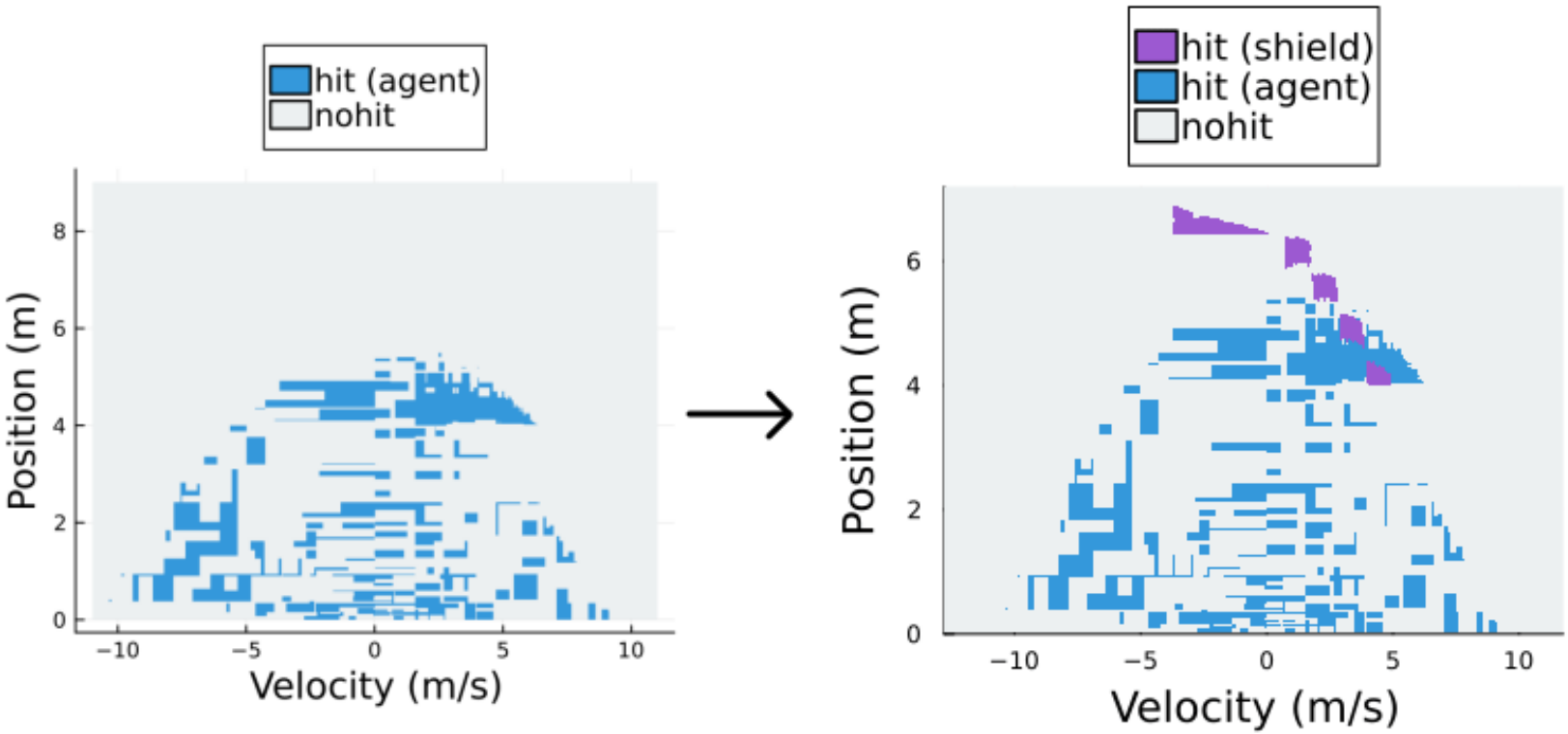




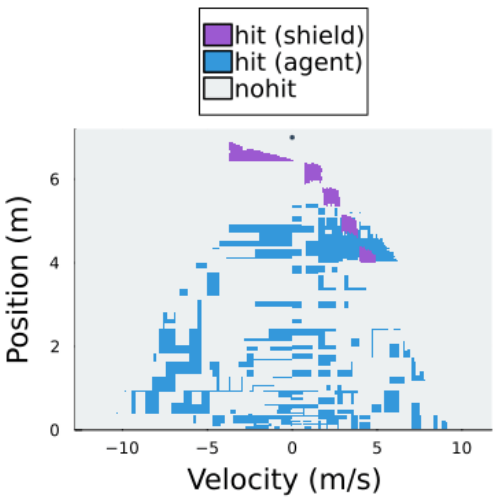
Post-shielded agent

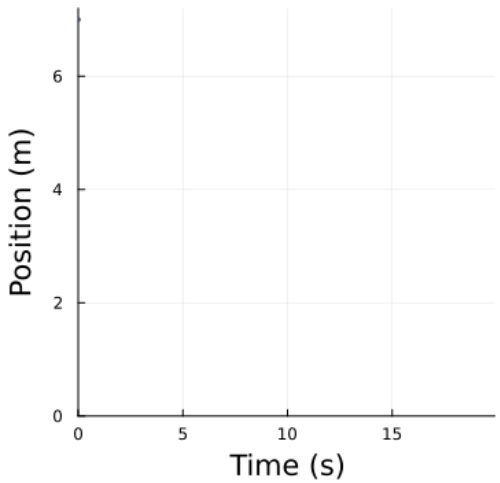


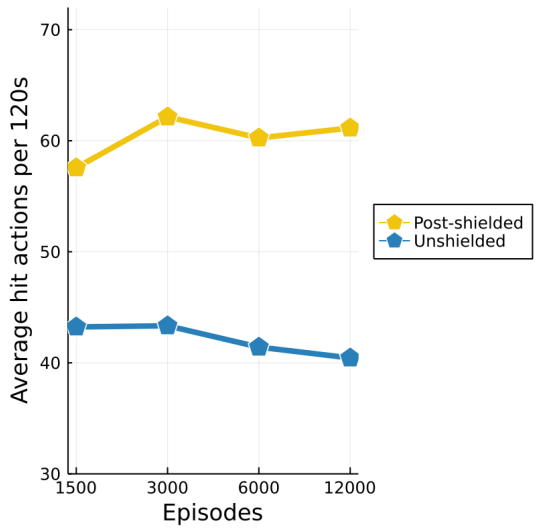
Post-shielded agent



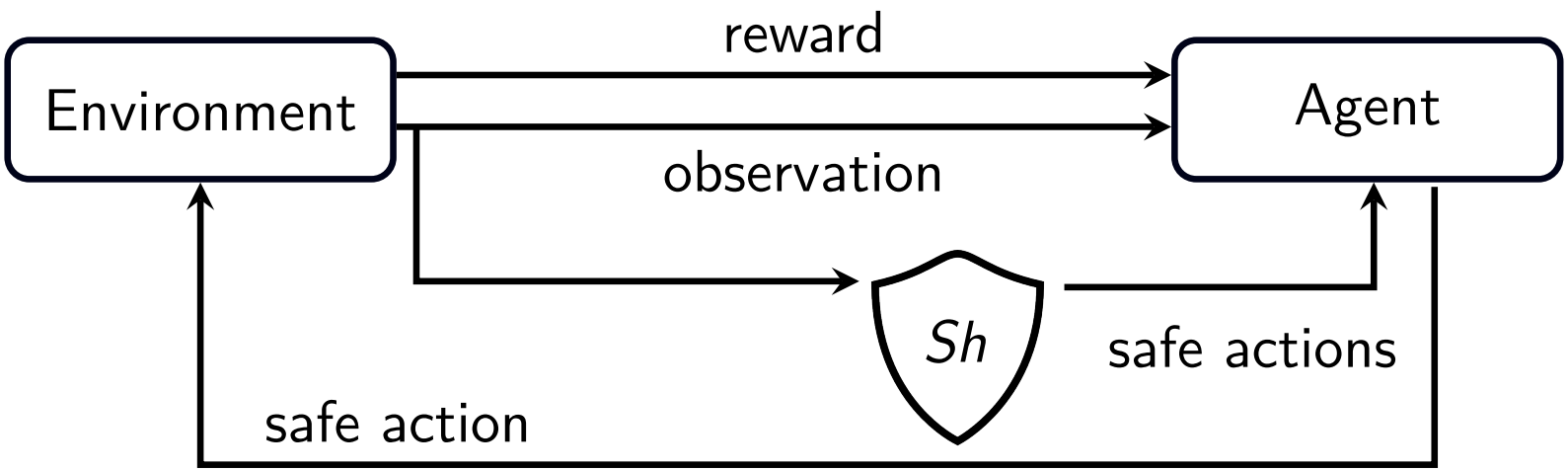
Post-shielded agent



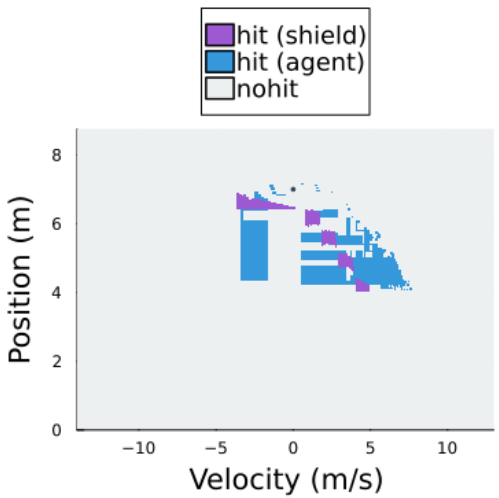


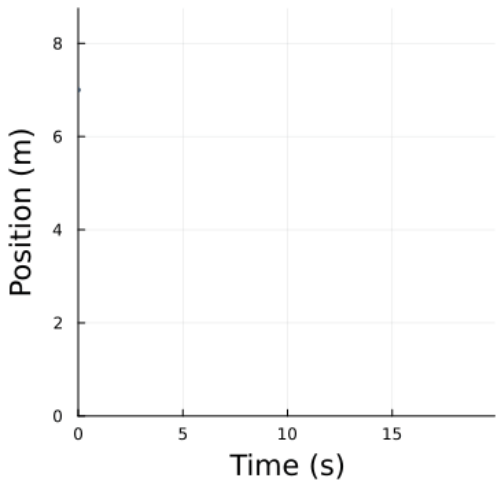


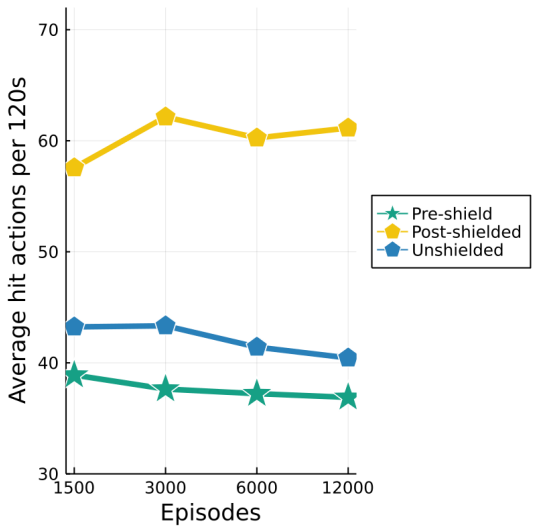
Pre-shielded agent



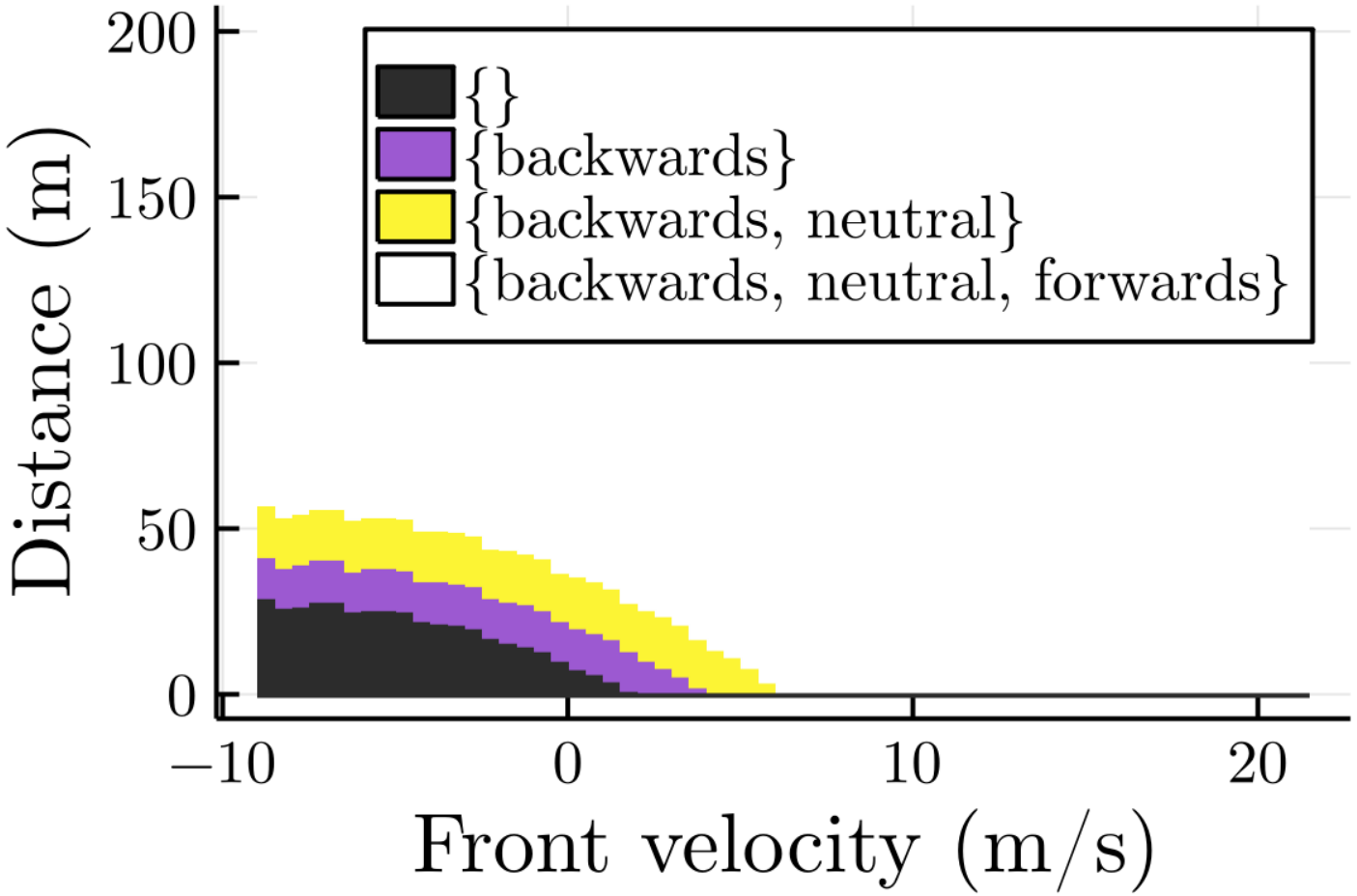
Pre-shielded agent





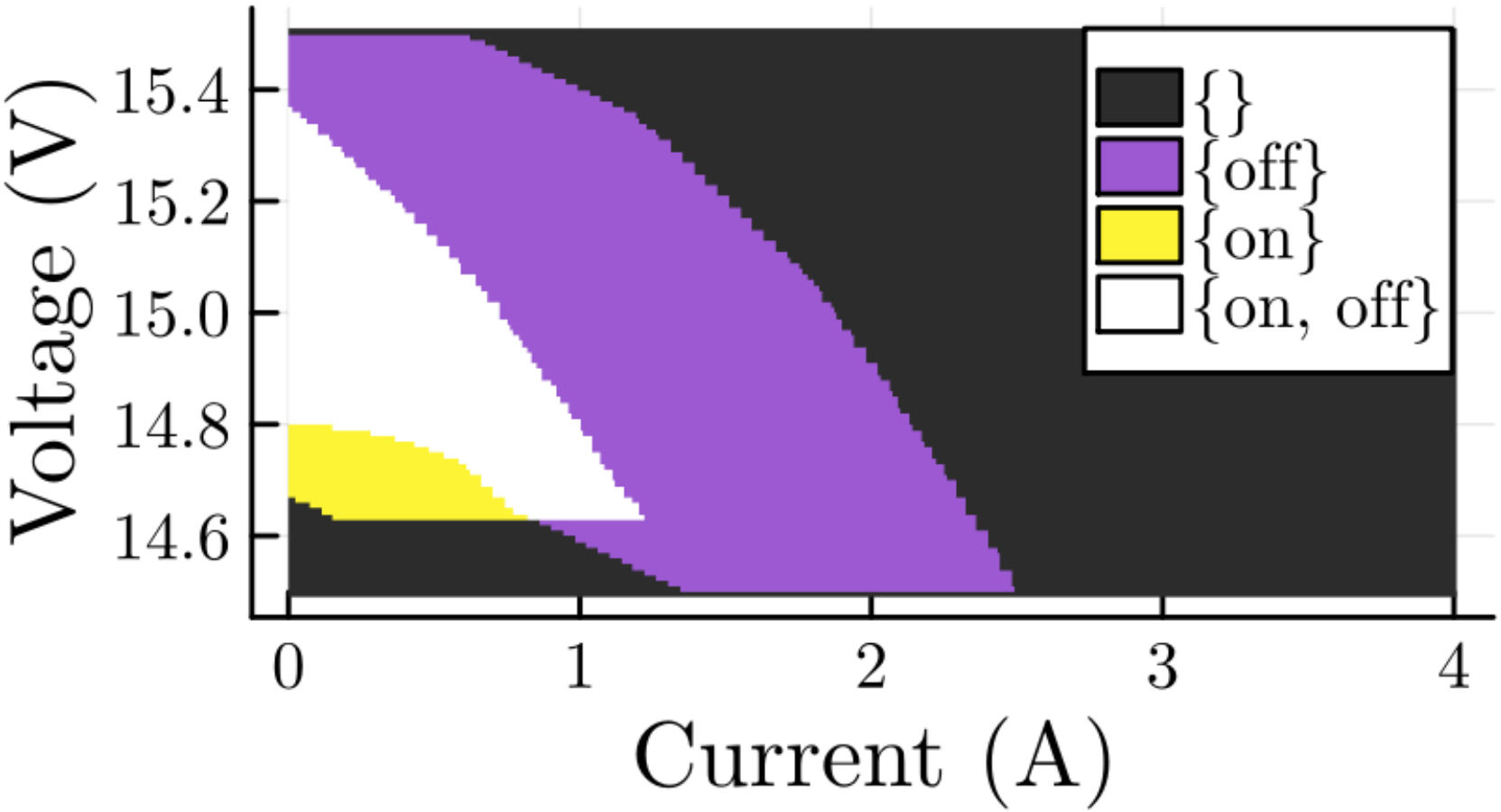


Cruise control (car behind another car)



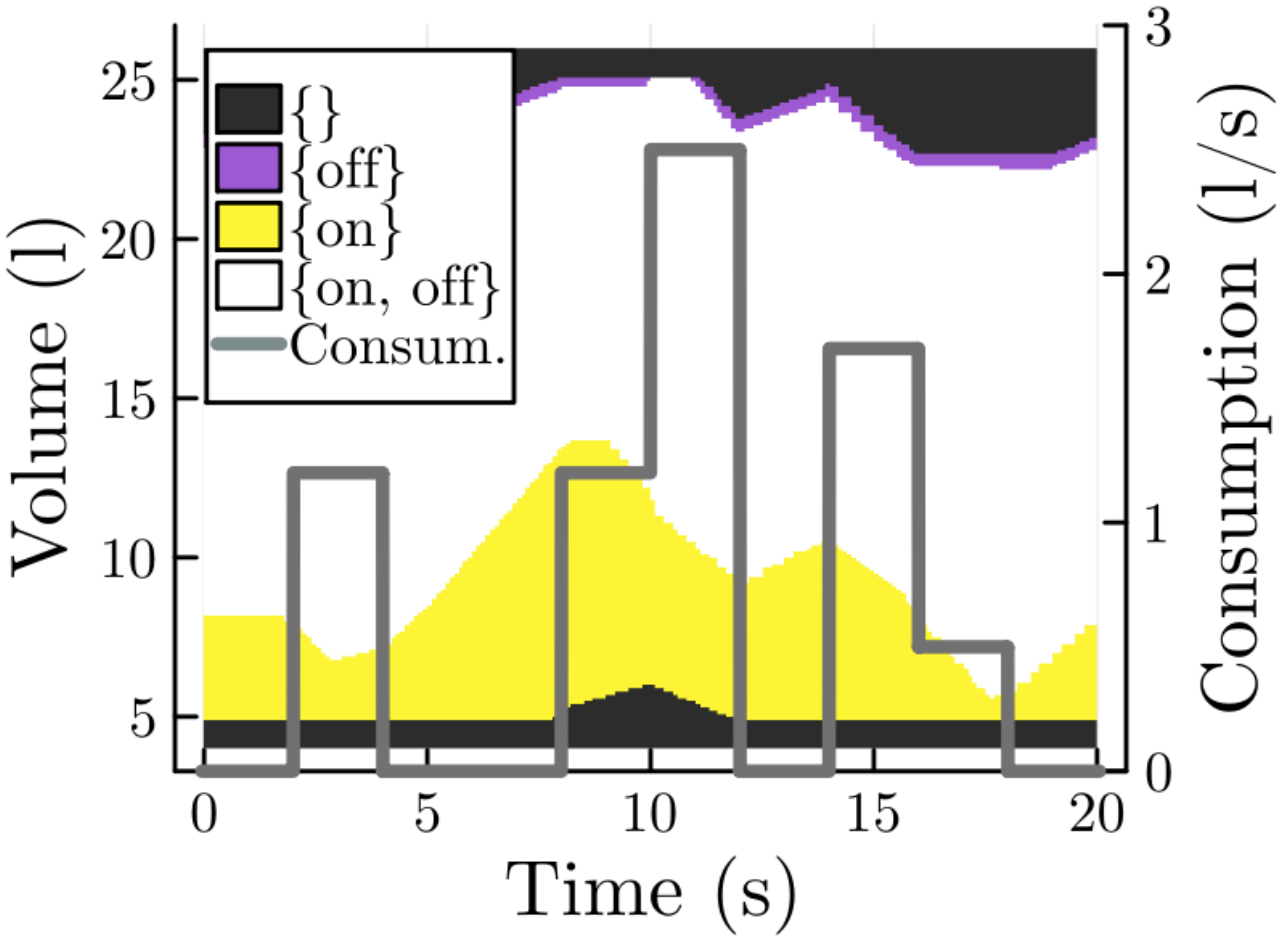
Synthesis time: 36 min

DC-to-DC boost converter



Synthesis time: 1 h 19 min

Oil pump (plot: pump is *on*)



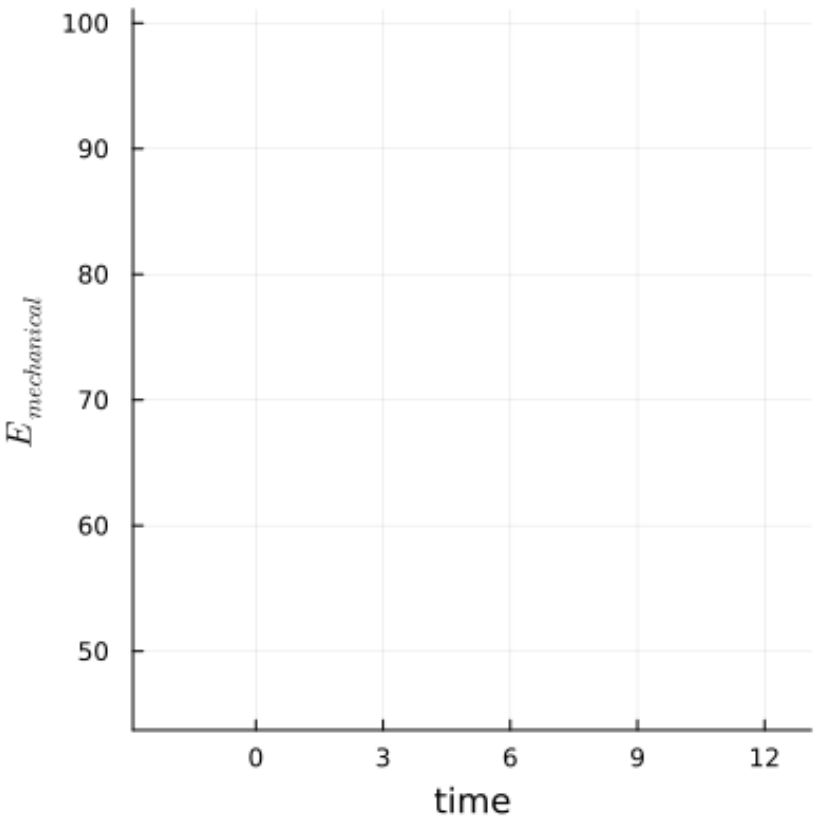
Synthesis time: 5 h 23 min

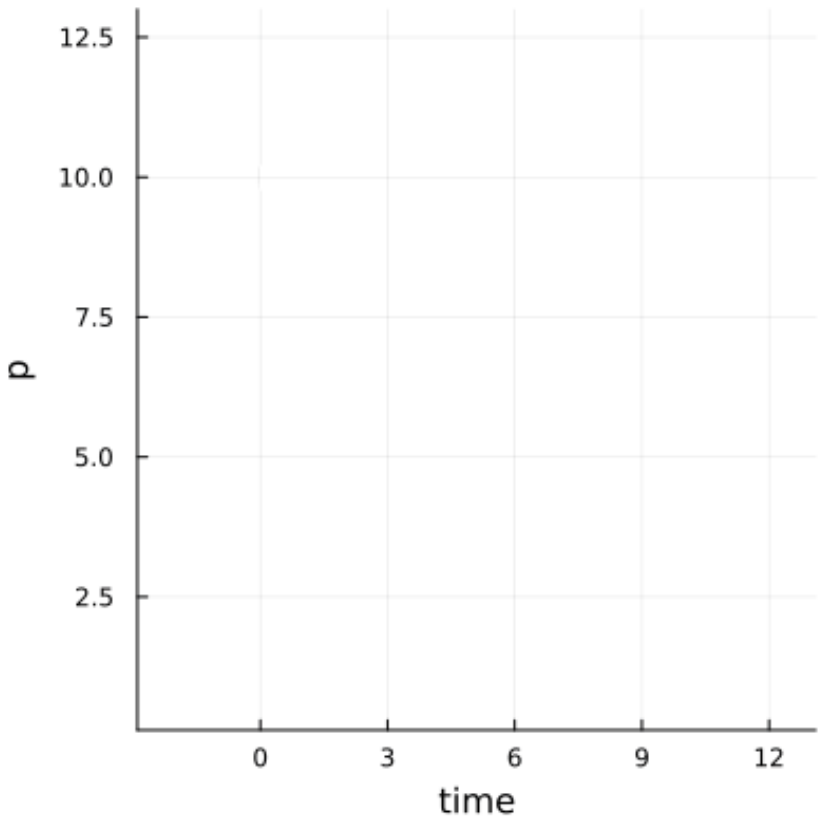
Overview

A different perspective: Mechanical energy

$$E_{mechanical} = mgp + 0.5mv^2$$

with mass m (1 kg), gravity g (9.81 m/s^2), position p, velocity v

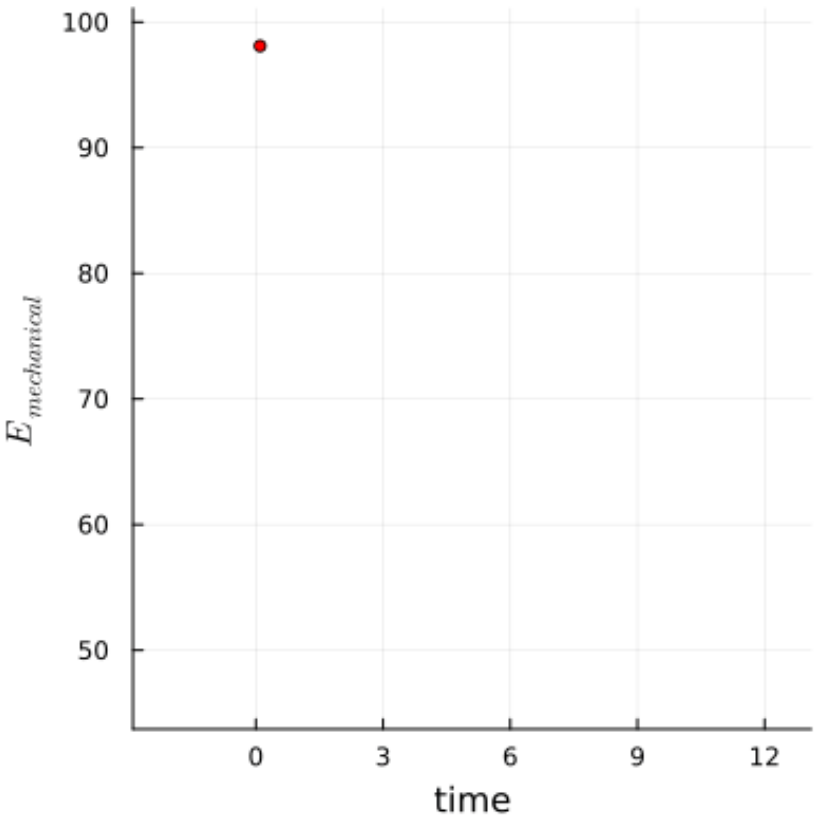


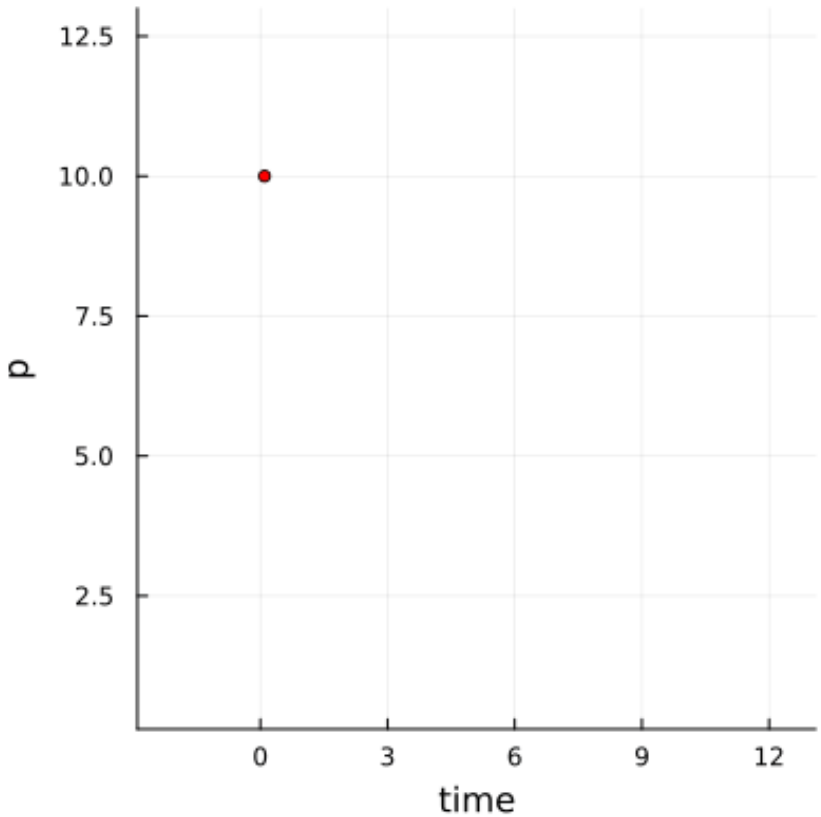


A different perspective: Mechanical energy

$$E_{mechanical} = mgp + 0.5mv^2$$

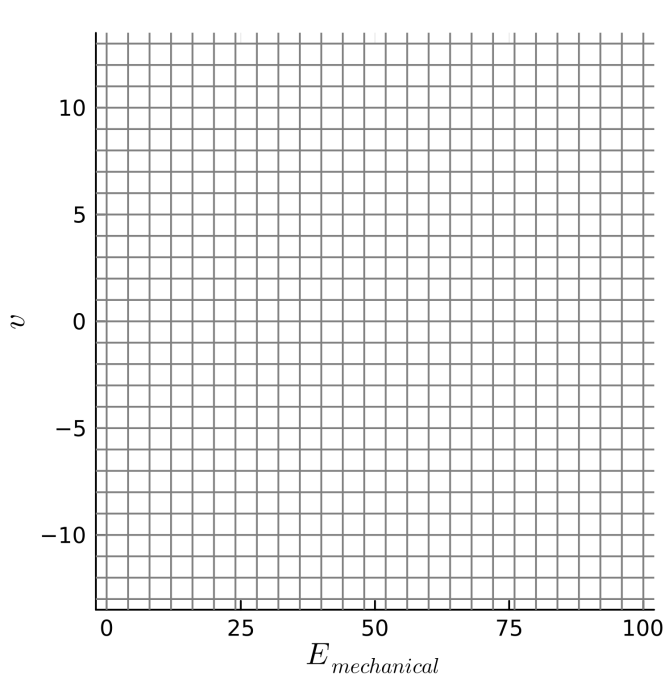
with mass m (1 kg), gravity g (9.81 m/s^2), position p, velocity v

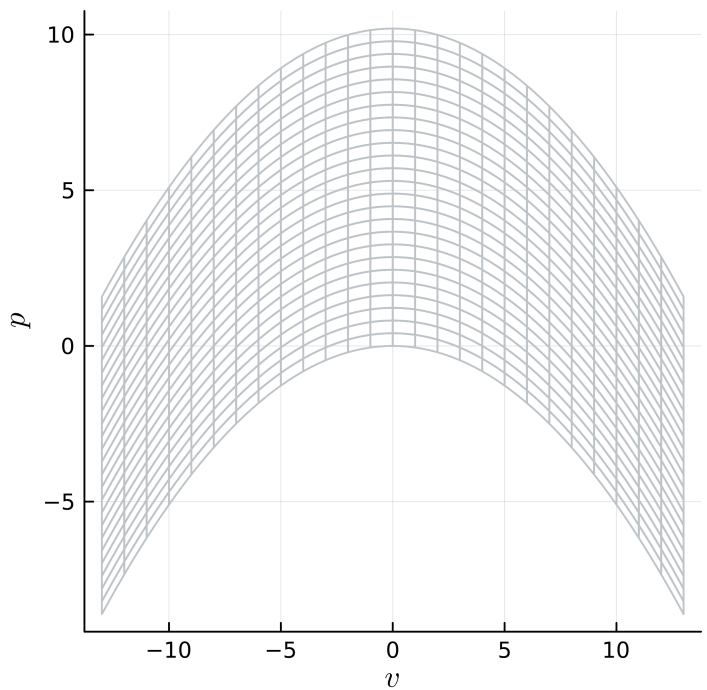




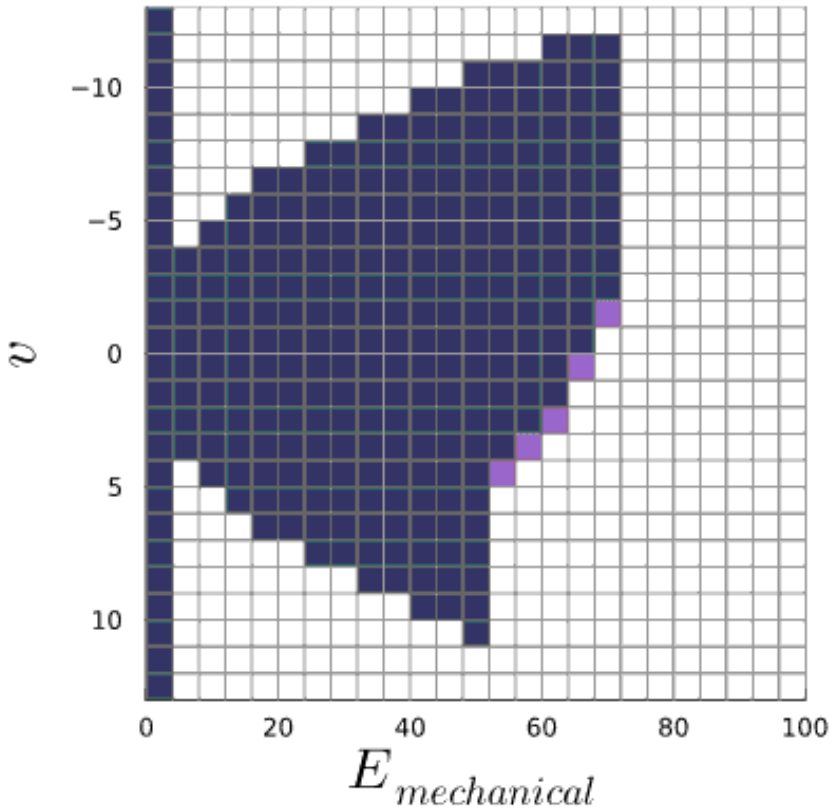
Transformed state space

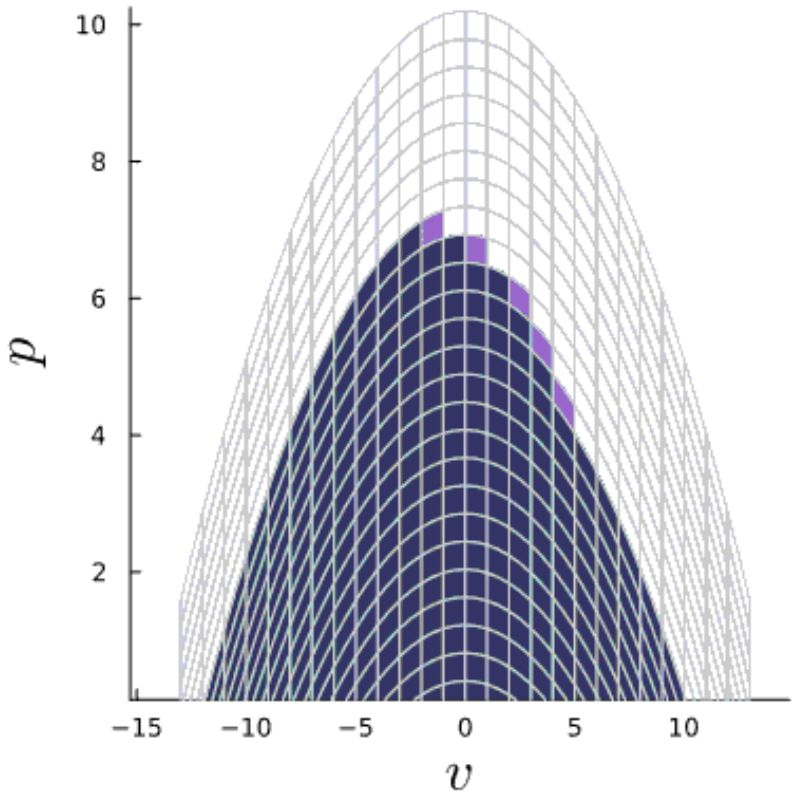
- New state dimensions ($E_{mechanical}, v$)
- We construct a regular grid again
- Much fewer cells (650) are sufficient
- Can go back and forth wrt. original state space (v, p)



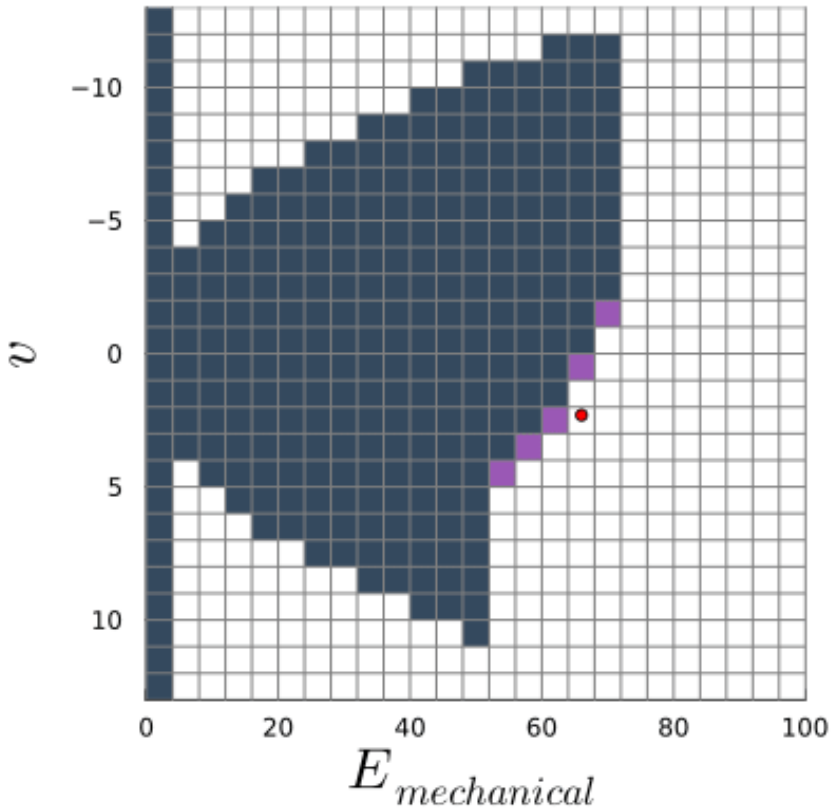


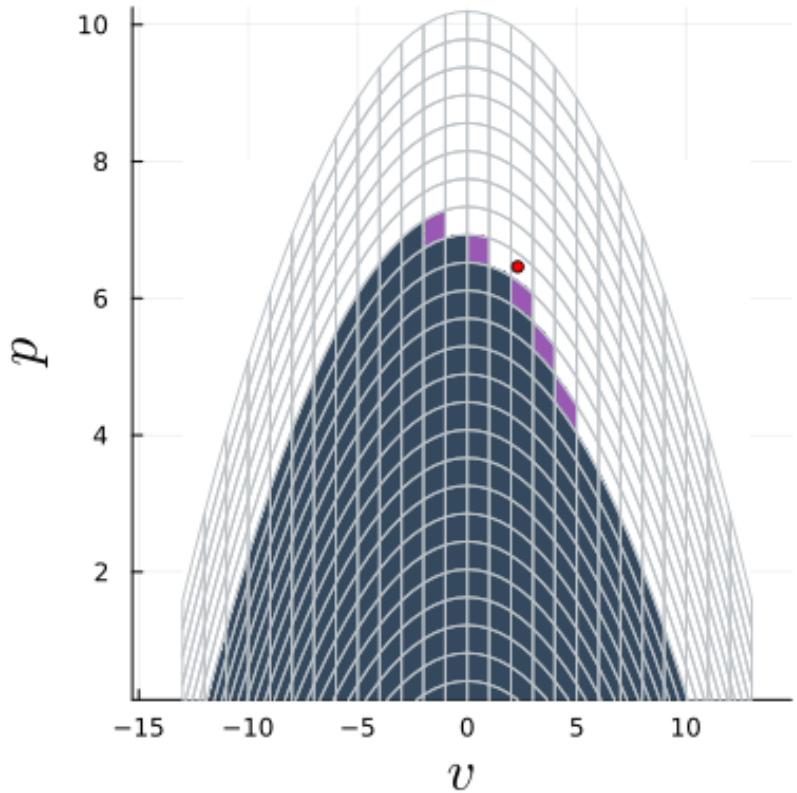
Shield in transformed state space



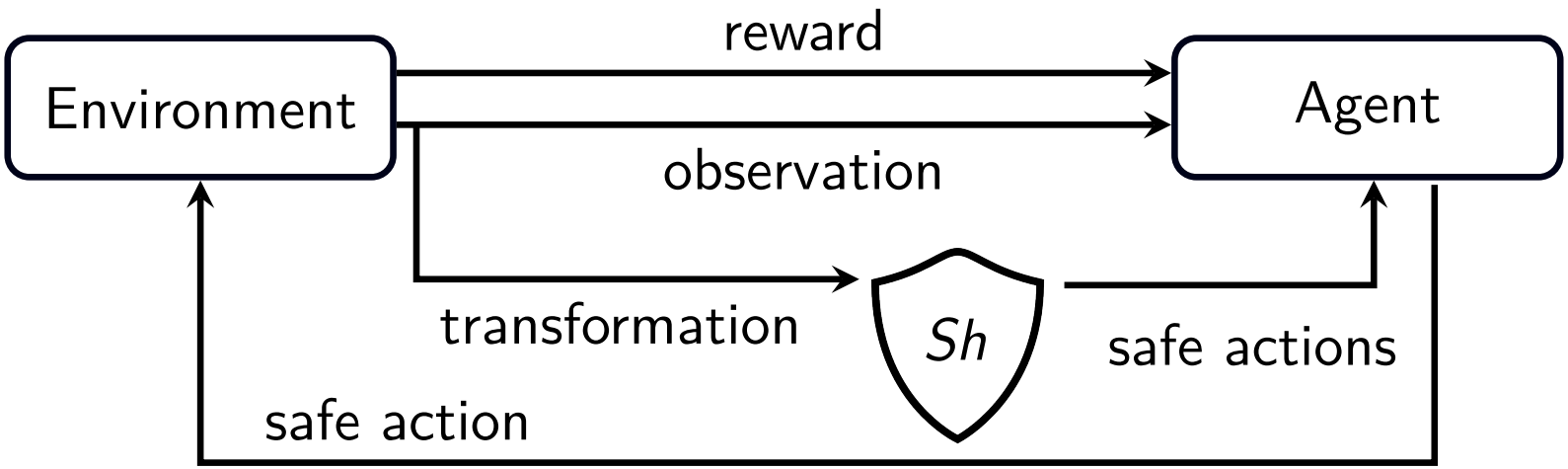


Shield in transformed state space



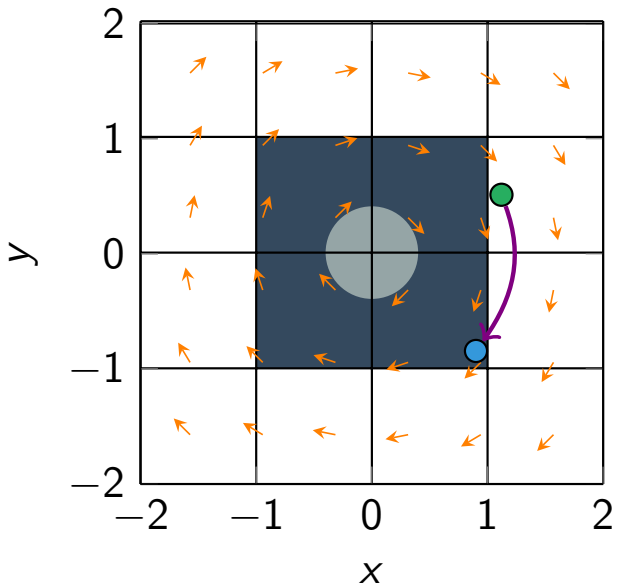


Shielding with transformation[1]

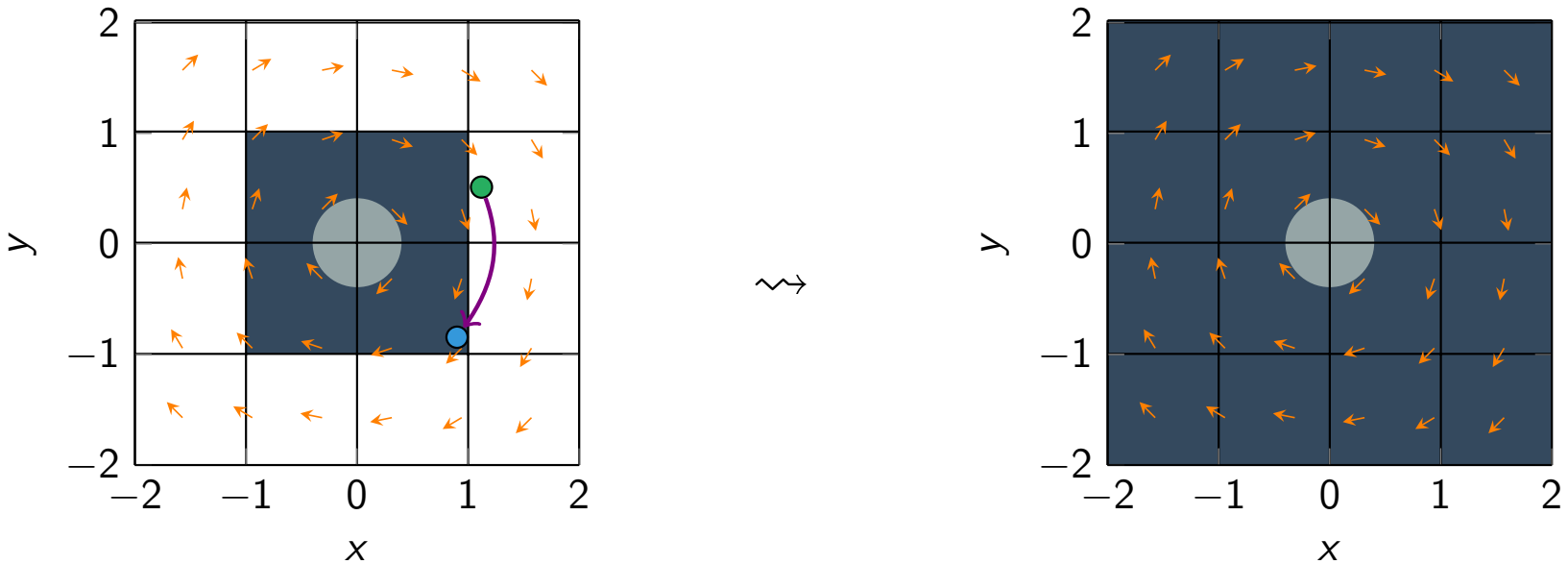


[1] Brorholt, Høeg-Petersen, Larsen, and Schilling. *AISoLA*. 2024.

Another example: Bivariate harmonic oscillator

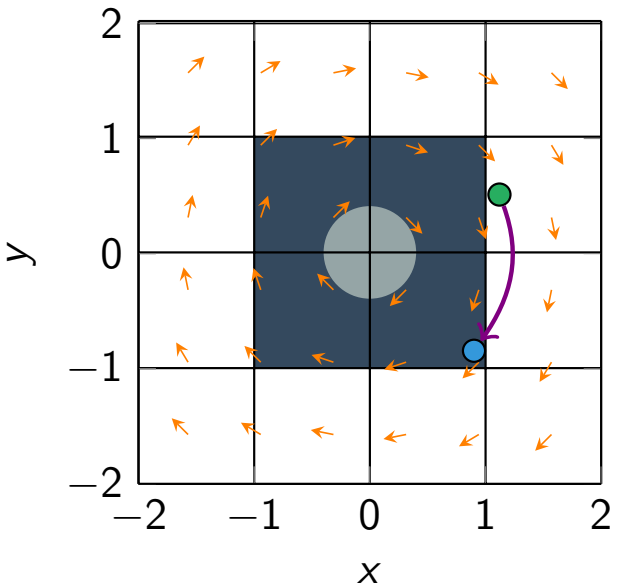


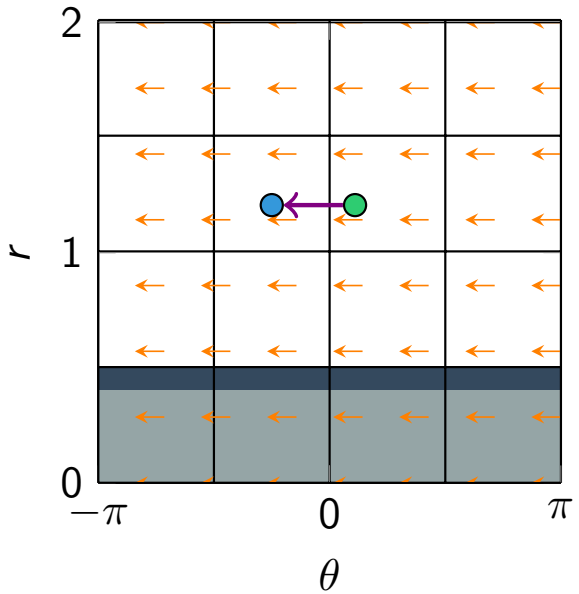
Another example: Bivariate harmonic oscillator



- All cells get into the avoid set $\rightarrow$ useless shield
- Would require a very fine-grained grid

Another example: Bivariate harmonic oscillator

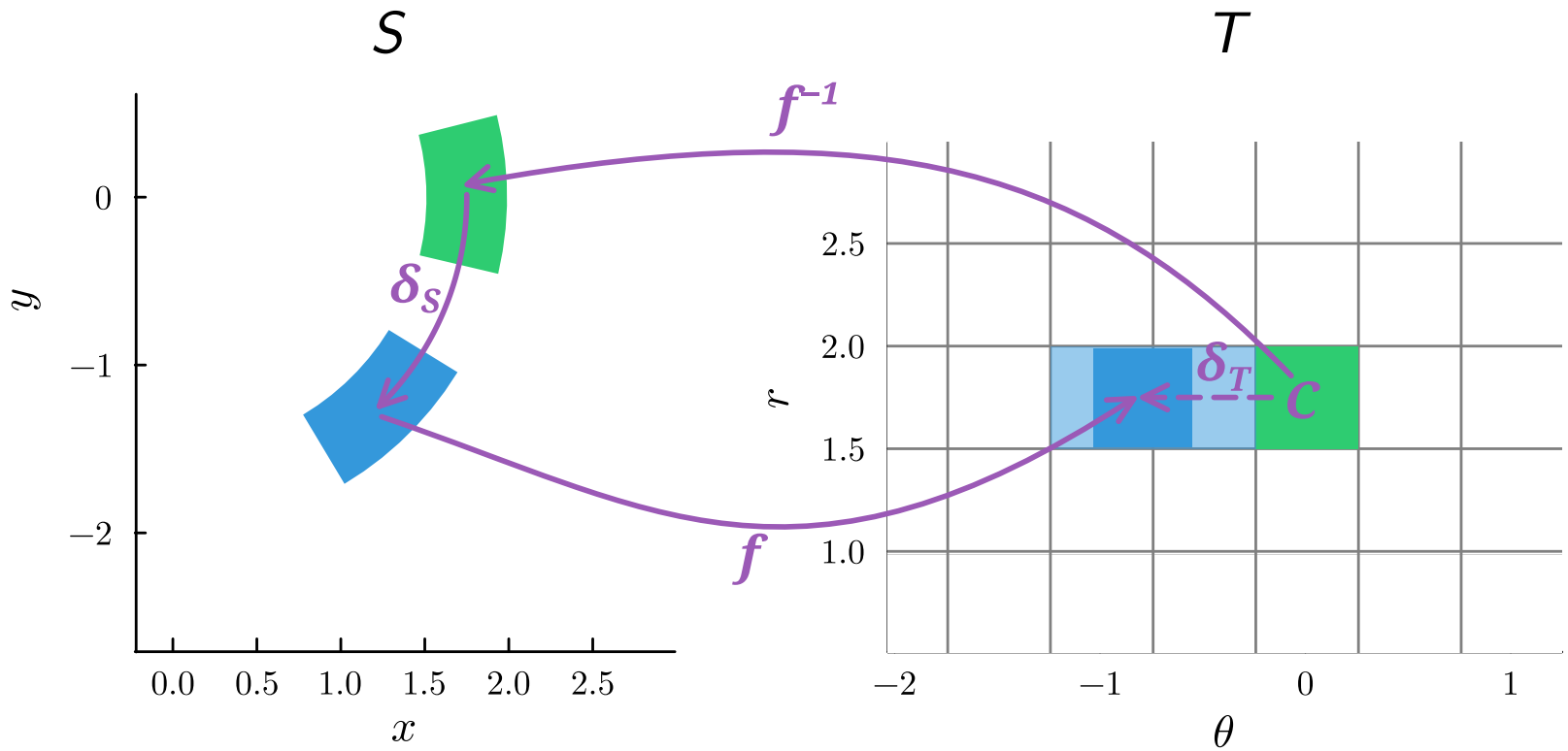




- Transformation to polar coordinates

$$f(x, y) = (\text{atan2}(y, x), \sqrt{x^2 + y^2}) =: (\theta, r)$$

Shield computation with a state-space transformation

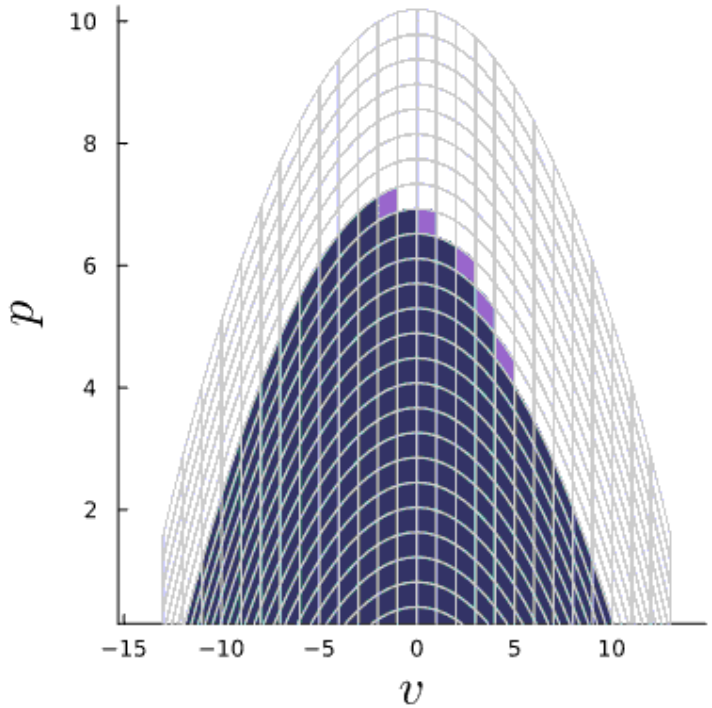


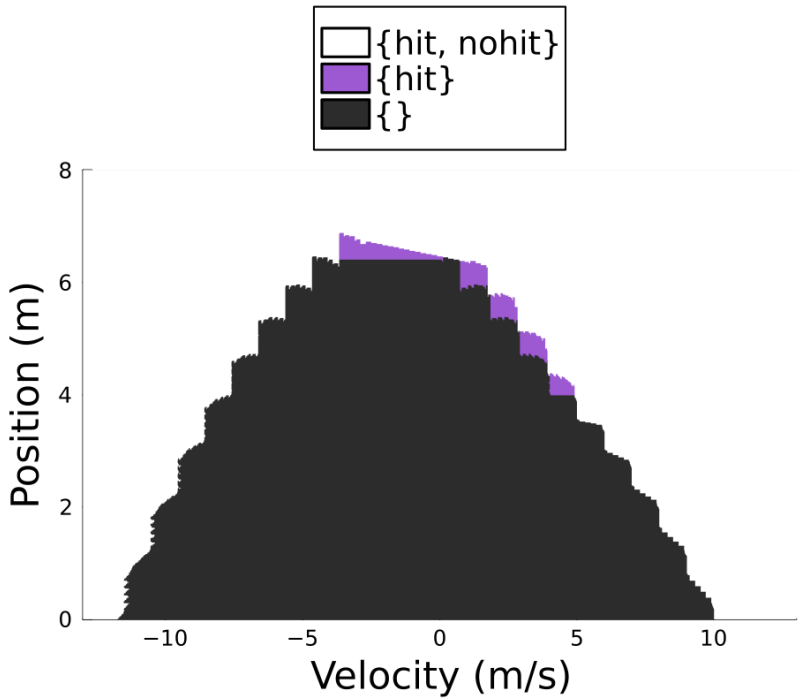
Compute successors δ_T from C (green) in transformed space T

1. Map to original state space S via f^{-1}
2. Compute successors via δ_S
3. Map back to transformed state space T via f (dark blue)

As before, identify all cells intersecting with this set (light blue)

Bouncing ball

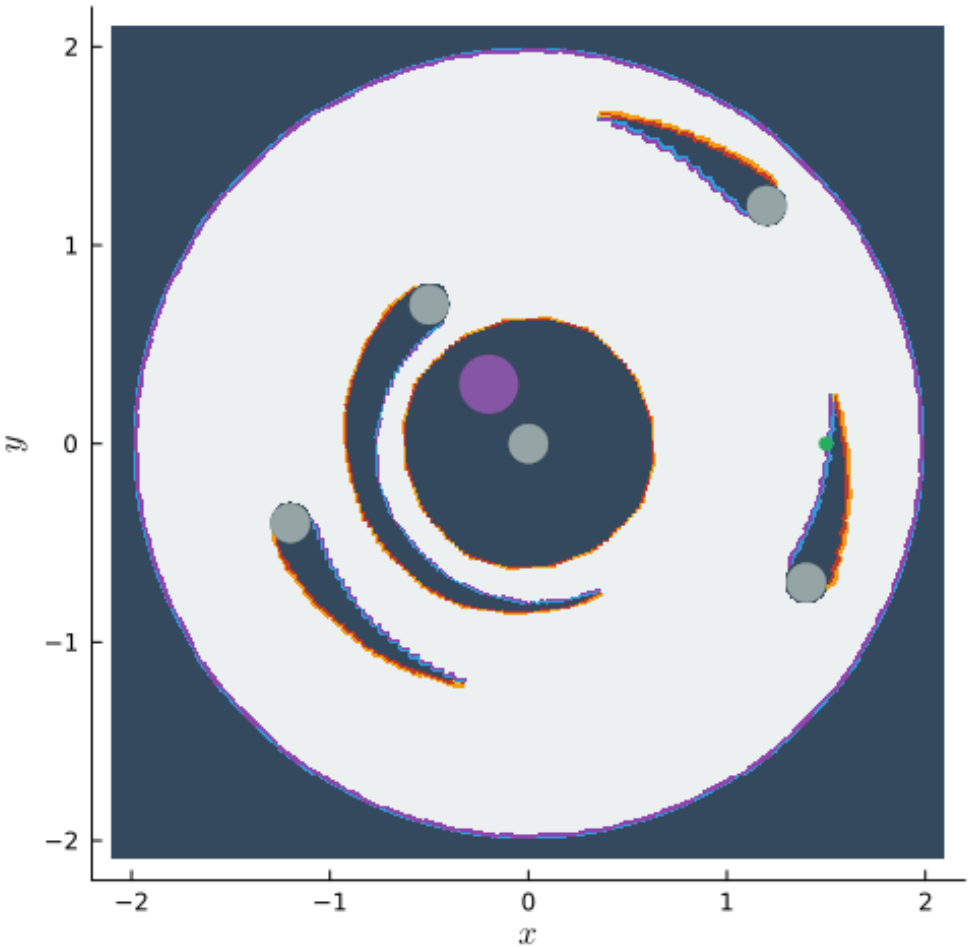


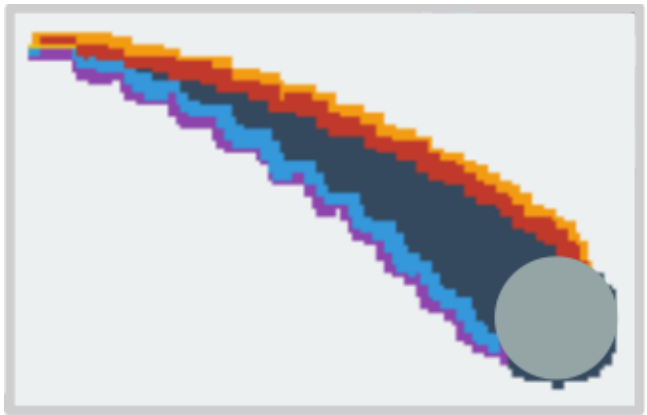


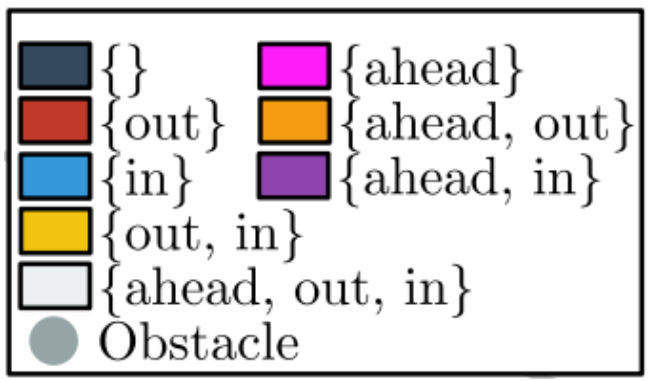
- 650 cells
- Computation: 1.3 sec
- Decision tree: 49 nodes

- 520,000 cells
- Computation: 134 sec
- Decision tree: 940 nodes

Satellite: Harmonic oscillator + {ahead, out, in}

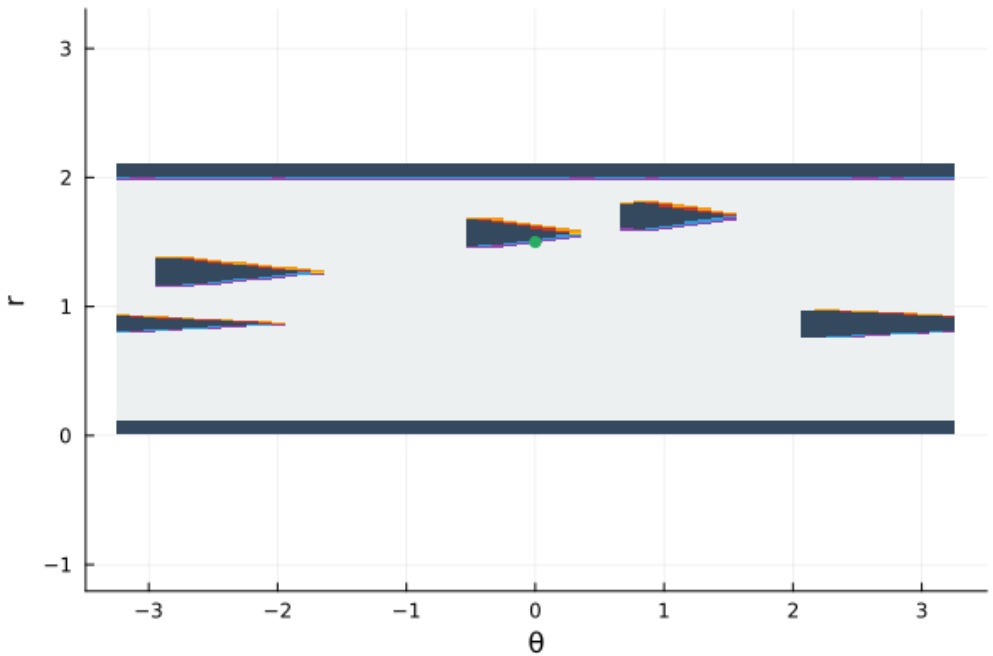


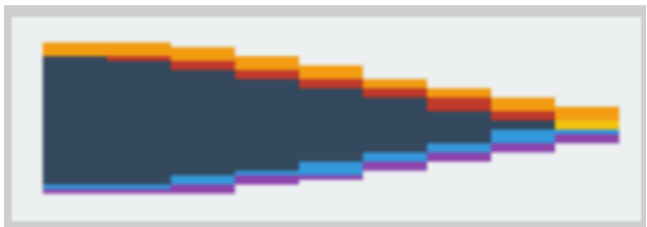


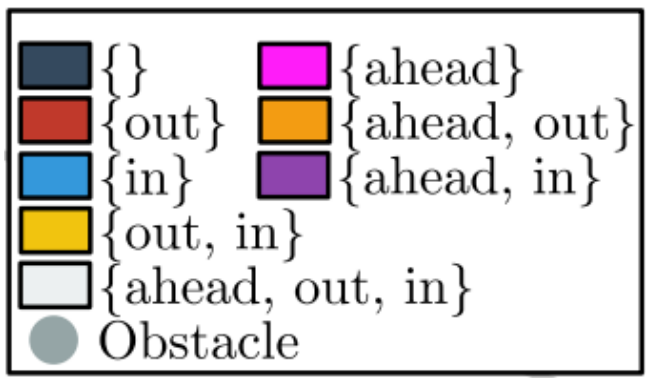


- Goal: reach a randomly spawning purple area
- Here this is impossible because the shield forbids it

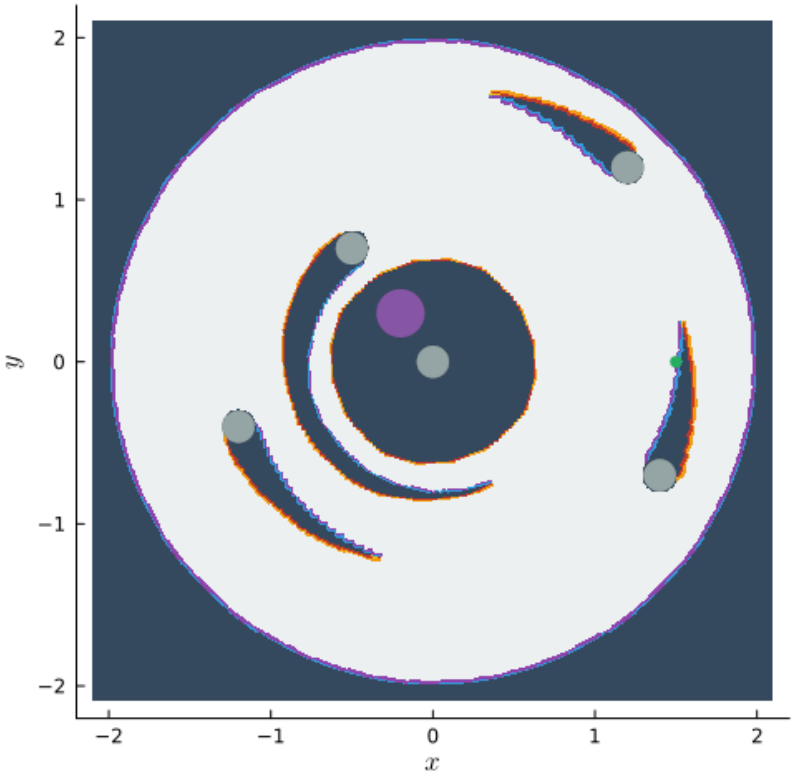
Satellite: Transformed state space

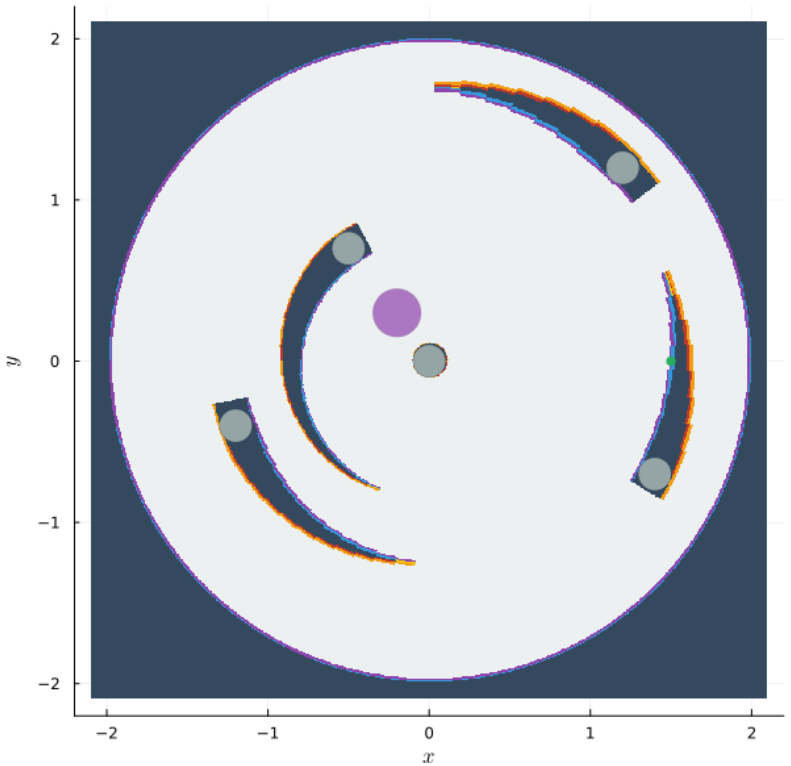






Satellite: Comparison

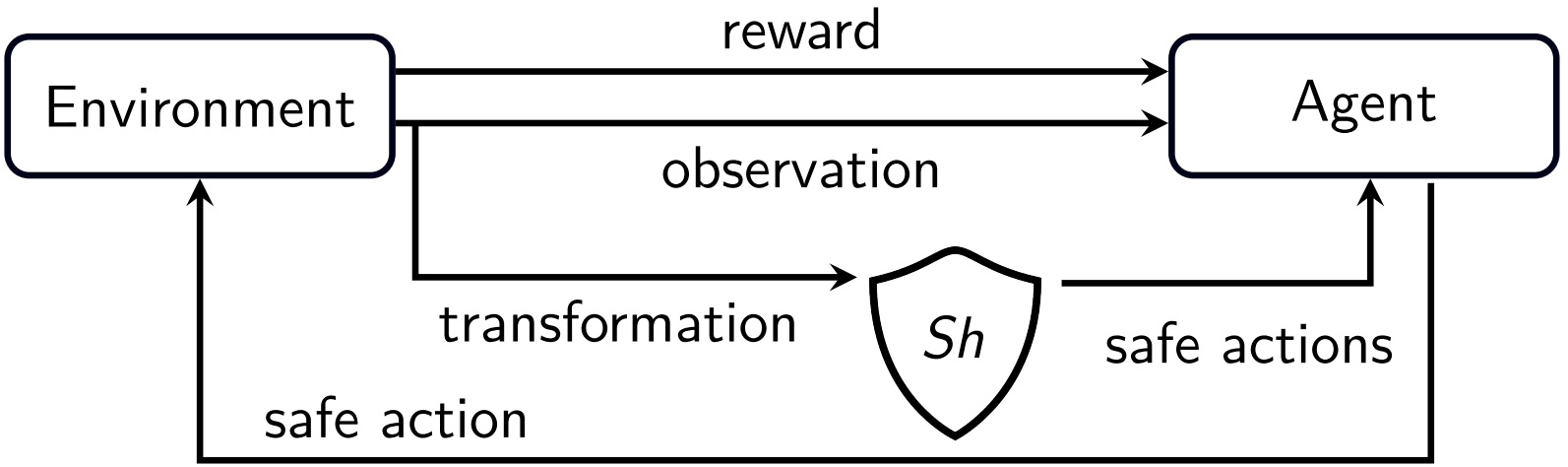




- 176,400 cells
- Computation: 161 sec
- Decision tree: 4,913 nodes

- 27,300 cells
- Computation: 10 sec
- Decision tree: 544 nodes

Impact on learning



- We evaluate the cumulative return in six scenarios
- Three shield variants: no shield, shield in S, shield in T
- Two learning variants: learn in S, learn in T
- (The diagram shows the variant *shield in T, learn in S*)

Impact on learning

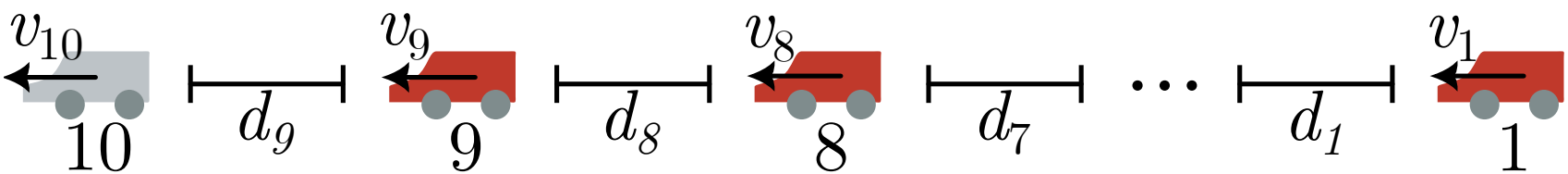
Learning	**Satellite** (↗)			**Bouncing ball** (↘)		
	None	S	T	None	S	T
S	1.123	0.786	**1.499**	39.897	37.607	**36.593**
T	0.917	0.889	**1.176**	39.128	40.024	**39.099**

Learning	**Cart-pole** (↘)		
	None	S	T
S	0.007	0.019	**0.001**
T	**0.000**	**0.000**	**0.000**

Overview

Multi-agent shields[1]

- Shield synthesis scales exponentially in number of agents
- Compositional shielding based on assume-guarantee reasoning
- Shielding even enables more efficient reinforcement learning



[1] Brorholt, Larsen, and Schilling. *AAMAS*. 2025.

Overview

Conclusion

- Shield synthesis for hybrid systems
  - Replace hard steps by simulation[1]
    - Now fully integrated in Uppaal
  - State-space transformation: more precise and efficient[2]
  - Multi-agent shielding: assume-guarantee structures[3]

Future work

- Dynamic partitioning
- Combination with symbolic approach
- Learning of suitable transformations and assumptions

[1] Brorholt, Jensen, Larsen, Lorber, and Schilling. *AISoLA*. 2023.
[2] Brorholt, Høeg-Petersen, Larsen, and Schilling. *AISoLA*. 2024.
[3] Brorholt, Larsen, and Schilling. *AAMAS*. 2025.