On Mathematical Guarantees in Machine Learning for Safe Autonomous Driving

By Philipp Geiger, Bosch Center for Artificial Intelligence

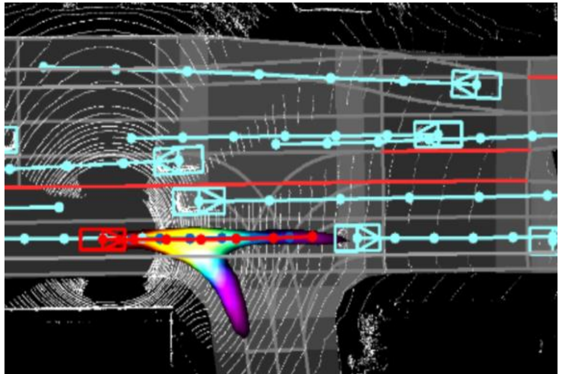
At Symposium on Applications of Mathematical Sciences, KIT, 29-09-2023

BOSCH

Introduction
Key tasks in autonomous driving (AD)

- *Control (= decision making)* of autonomous vehicles or delivery robots – needs safety

- *Modeling and simulation* of realistic human agents' multi-modal traffic behavior, e.g., to test and validate control algorithms against such models – need generality of road situations, but also robustness



Source: Zheng et al.: End-to-end Interpretable Neural Motion Planner

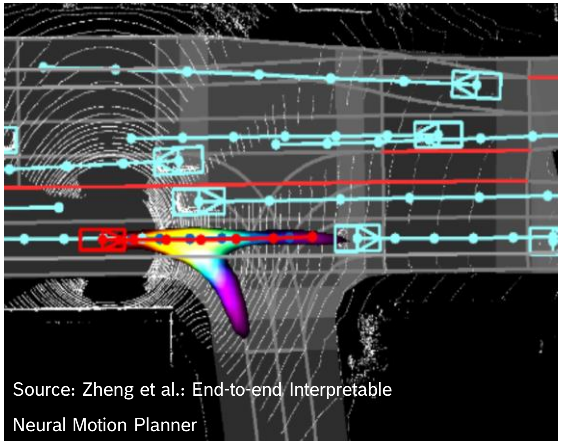
Introduction
Deep imitation learning, task formulation

Powerful approach for such control and modeling problems: machine learning (ML), and especially **deep imitation learning (IL):**

- **Given** a data set of temporal trajectories of ***states*** *s*, ***actions*** *a*,
  $(s_1, a_1), (s_2, a_2), \ldots, (s_T, a_T)$
  of ***demonstrator agent's*** *sequential decision making* -- e.g., human driver
- **Goal:** from this data, learn an ***imitator agent*** $\pi^I(a|s)$ – a probabilistic policy mapping state to action density – that *behaves similarly to demonstrator*

- More and more **cheap data available**: from drones, car sensors, etc.
- Deep IL is **flexible and scalable** - needs little human work on hand-crafting rules for each new situation
- Therefore, deep IL is booming in AD

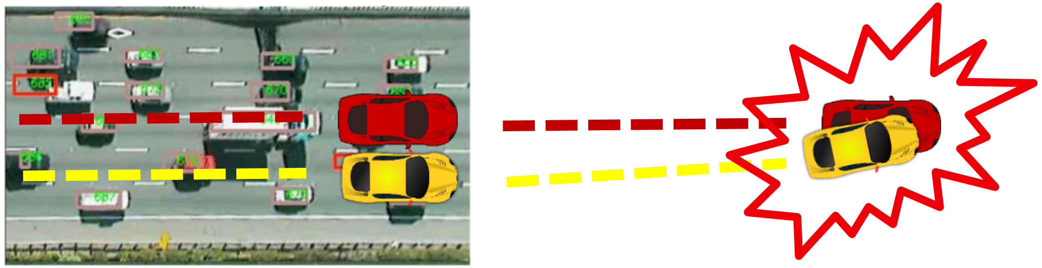
[Igl et al, '22][Bansal et al, '18] [Bhattacharyya et al, '20] [Tao et al, '21] [Deo et al, '18] [Tang et al, '19]



Source: Zheng et al.: End-to-end Interpretable Neural Motion Planner

Introduction

Problem: robustness and safety



- Various IL algorithms suffer from ***compounding error problem.*** *There are some mitigations for this.*
- But: Generally, almost no work on guaranteeable safe/robust IL
- Of course: generally in ML/IL: **fundamental problem of induction**. That's uncritical in some areas.
- But: for autonomous driving (AD) control or simulation, **we need safety/robustness arguments!**

Introduction

A broad landscape of types of mathematical guarantees in ML

(very preliminary)

Guarantee: proven statement about how a trained system will perform in deployment

Form: often relative to some benchmark – otherwise no free lunch – inherent uncertainty in ML

Prediction = offline

Control = online (key for AD)

Probabilistic statistical learning

- Often i.i.d.
- Law of large numbers, Central limit theorems, "Probably approximately correct" (PAC) bounds
- -> Often too weak/pessimistic
- Test-set based approaches (recent)
- Extreme Value Theory for AD

Adversarial robustness

- (in supervised learning)
- Take into account delibarate perturbations

Interpretability/identifiability

- Identifiability of parameters of a model (e.g., agent preferences) **-> our work (not presented today)**
- Explainability

Reinforcement learning

- RL, bandits, (Stochastic) optimization
- Convergence, in large sample limit, with enough exploration
- **Probababilistic No-Regret bounds**
- **Adversarial No-Regret bounds**
- Multi-agent -> **Convergence to (Nash-)equilibria**

A priori safety biases, e.g.,

- Obtain „safe set of actions“ via worst-case reachability games / Hamilton-Jacobi type eq. / invariant sets
- **Or via RSS from AD domain ("no-blame" if in utopia) [Shalev-Shwartz et al, `17]**
- Then, constrain a learnable policy to output into the safe set **-> our safe IL (today)**

Overall: few success stories, many limitations. **But the problem does not go away!** ML in AD is growing

Today: present one approach using a priory safety biases (constraints) for IL

BOSCH

Fail-Safe Adversarial Generative Imitation Learning

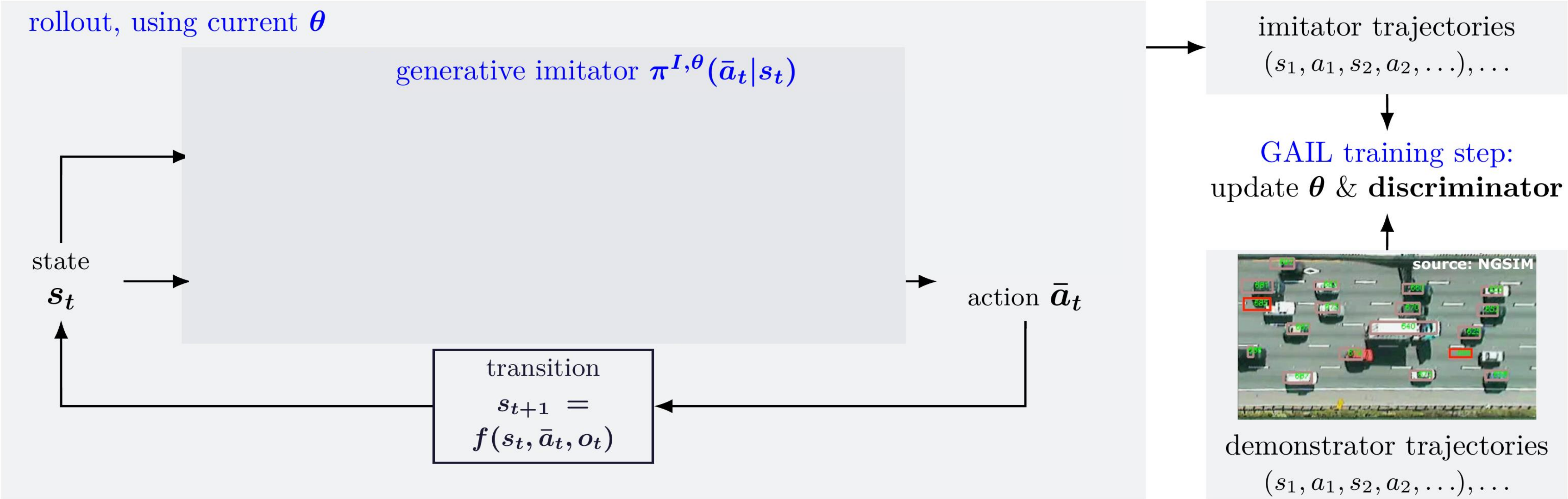
Published at TMLR

Joint work with Christoph-Nikolas Straehle

BOSCH

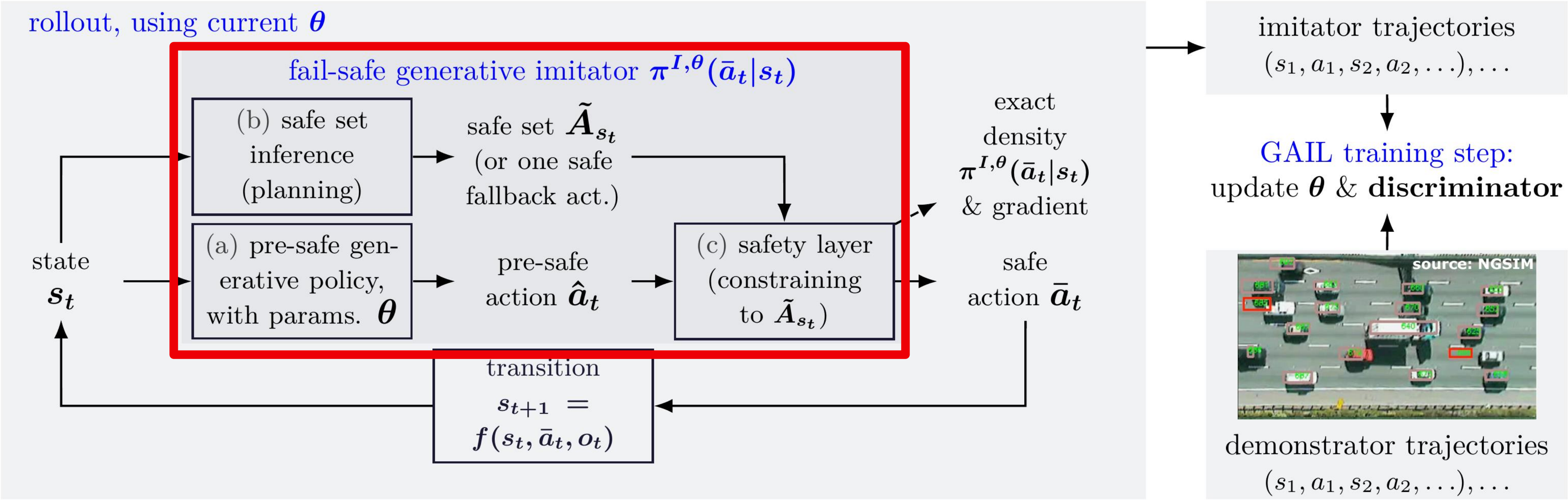
Fail-Safe Adversarial Generative Imitation Learning

Outline of our method



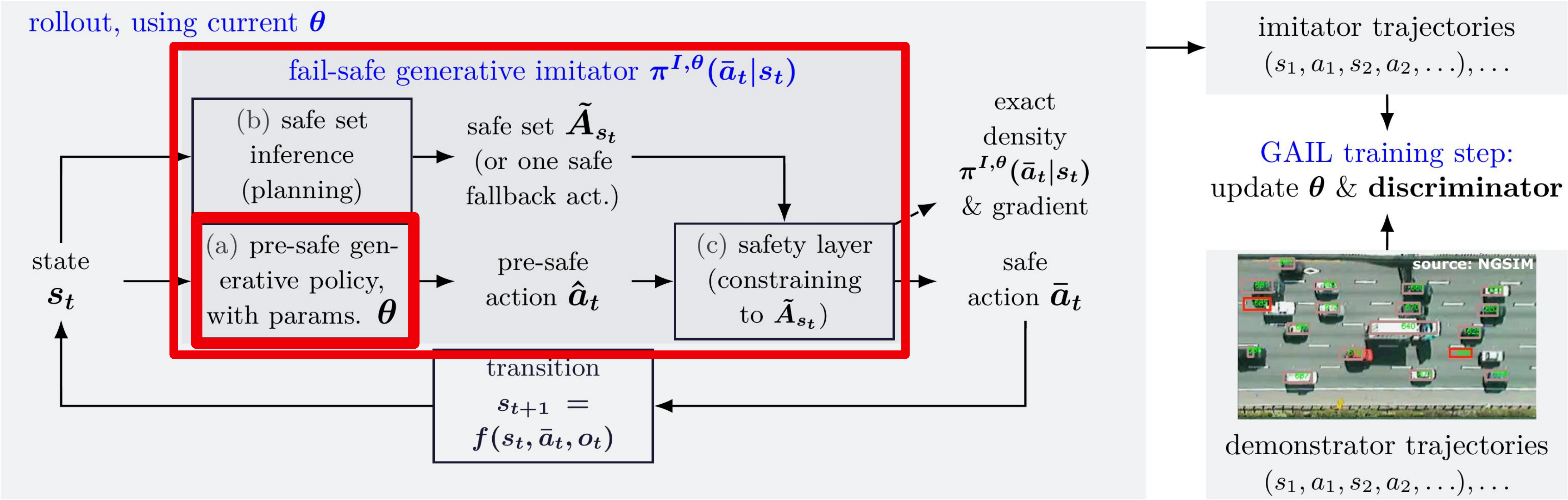
- Build on "*GAIL*": *Generative Adversarial Imitation Learning* [Ho et al, '16], based on GANs

Fail-Safe Adversarial Generative Imitation Learning
Outline of our method



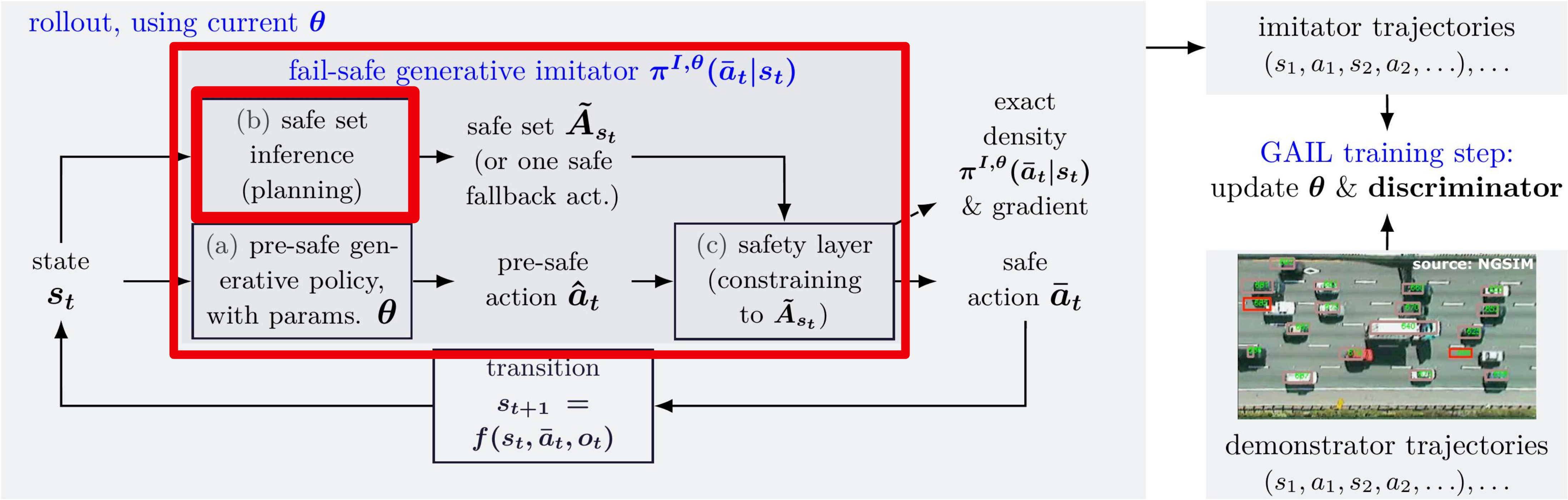
- Build on "*GAIL*": *Generative Adversarial Imitation Learning* [Ho et al, '16], based on GANs
- **Idea:** add safety, but keep closed-form policy density/gradient, for end-to-end training (no cov. shift)

Fail-Safe Adversarial Generative Imitation Learning
Outline of our method



- Build on "GAIL": *Generative Adversarial Imitation Learning* [Ho et al, '16] , based on GANs
- ``*pre-safe generative policy*" – take off-the-shelve Gaussian policy or Normalizing Flow policy with closed-form density

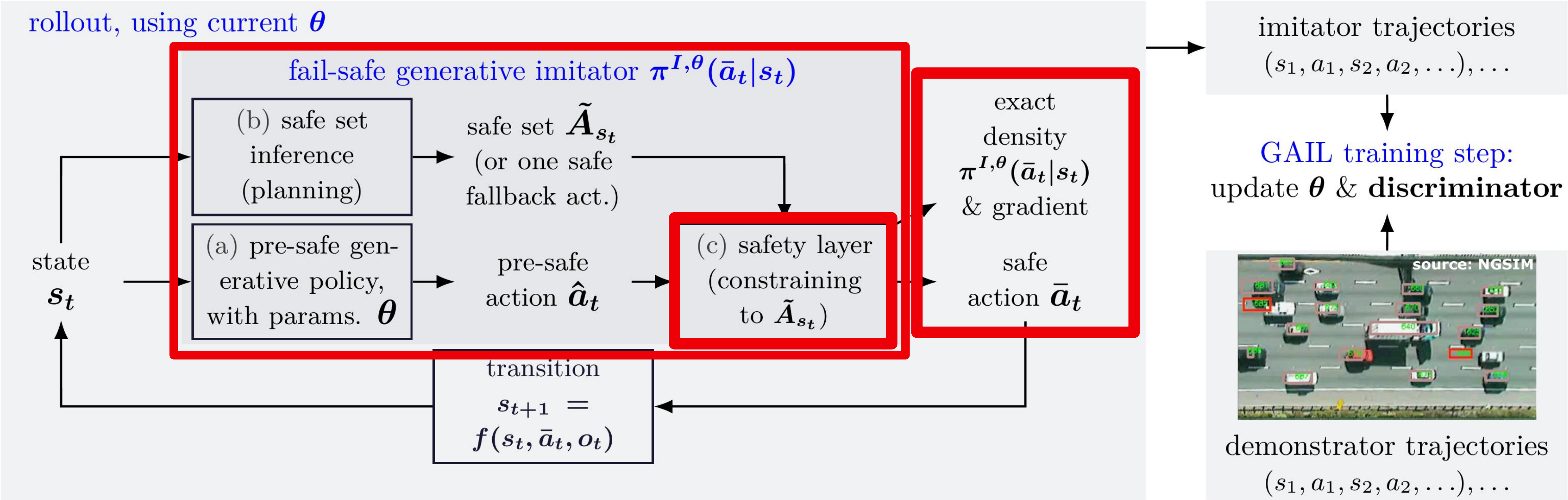
Fail-Safe Adversarial Generative Imitation Learning
Outline of our method



- Build on *"GAIL"*: *Generative Adversarial Imitation Learning* [Ho et al, '16]
- **Idea:** add safety, but keep closed-form policy density/gradient, for end-to-end training (no cov. shift)

Fail-Safe Adversarial Generative Imitation Learning

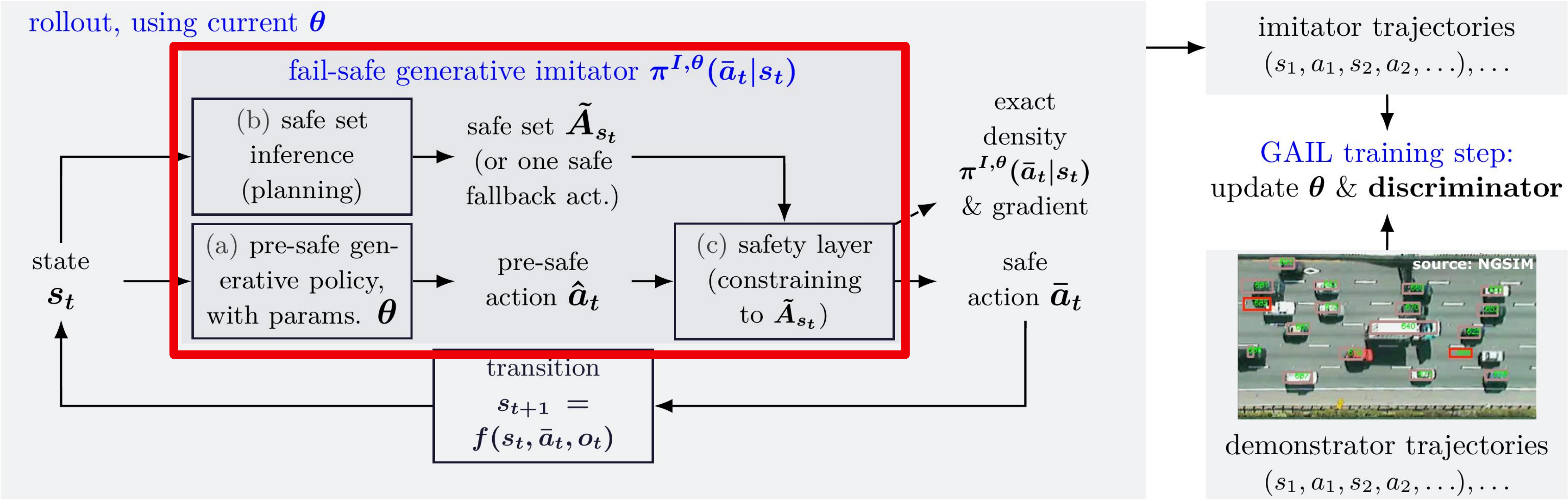
Outline of our method



- Build on *"GAIL"*: *Generative Adversarial Imitation Learning* [Ho et al, '16]
- **Idea:** add safety, but keep closed-form policy density/gradient, for end-to-end training (no cov. shift)

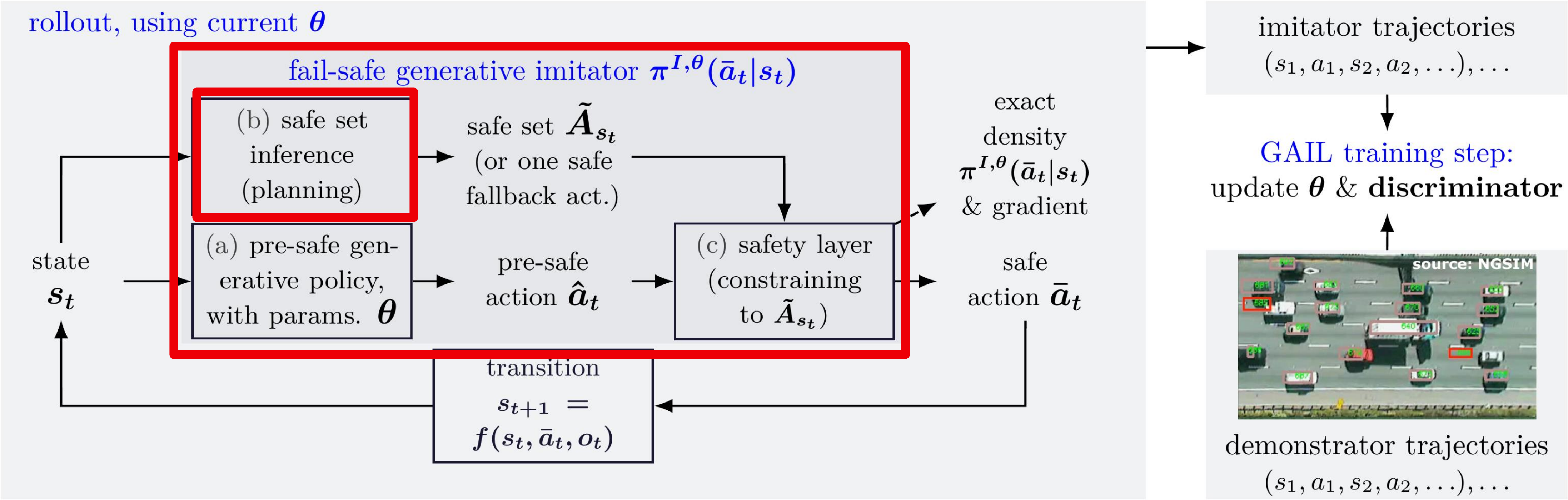
BOSCH

Fail-Safe Adversarial Generative Imitation Learning
Outline of our method



BOSCH

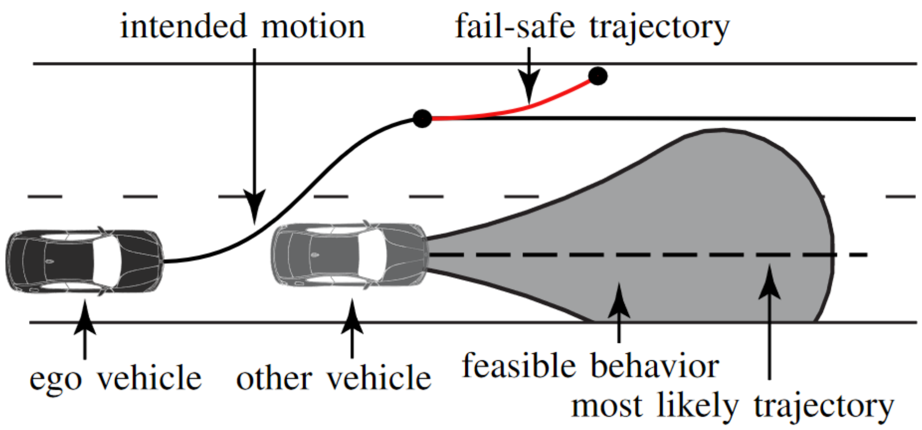
Fail-Safe Adversarial Generative Imitation Learning
Outline of our method



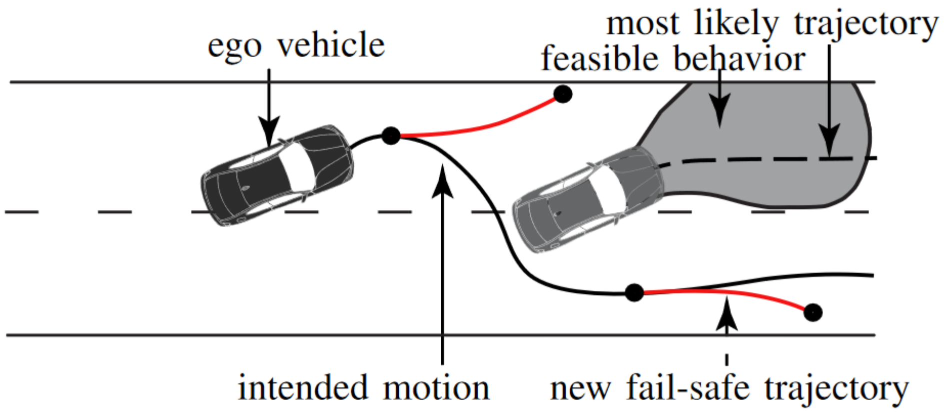
BOSCH

Fail-Safe Adversarial Generative Imitation Learning

Safe action set via sample-based reachability analysis I



(a) Initial scenario



(b) Future scenario

Image credit: "Computationally Efficient Fail-safe Trajectory Planning for Self-driving Vehicles Using Convex Optimization”

We build on the following idea from control engineering:

The **set of safe actions** is given by those potential current actions/motions, for which at least **one invariably safe future** continuation trajectory exists (no unsafe states are reached)

BOSCH

Fail-Safe Adversarial Generative Imitation Learning
Safe action set via sample-based reachability analysis II

Define ***safe action set*** $\bar{A}$ at state s and time t, via adversarial/worst-case reachability analysis

$$\bar{A}_t^s := \{a \in A : \text{ it exists } \pi_{t+1:T}, \text{ s.t. for all } \varphi_{t:T}, t < t' \leq T, \ d(s_{t'}) \leq 0 \text{ holds, given } s_t = s, a_t = a\}$$

Making this *quantitative (safety value*) instead of *qualitative (safe set yes/no)* will be helpful!

Total safety cost to go function w **:**

$$w_t(s,a) := \min_{\pi_{t+1:T}} \max_{\varphi_{t:T}} \max_{t' \in t+1:T} d(s_{t'}), \text{ for all } t$$

then

$$\bar{A}_t^s = \{a : w_t(s,a) \leq 0\}$$

Recall:

$\boldsymbol{\pi}$ ego agent policy

$\boldsymbol{\varphi}$ other agents and (adversarial) perturbations in the environment

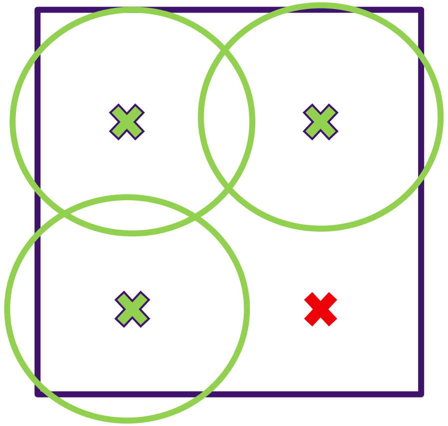
$d(s_t)$ momentary safety cost in state s_t

Fail-Safe Adversarial Generative Imitation Learning
Safe action set via sample-based reachability analysis III

1. Calculate safety of **finite** sample of actions,
2. conclude on safety of **infinite** set (inner approx. of safe set), via **Lipschitz** continuity (or convexity)!

Proposition 1 (Lipschitz constants for Lipschitz-based safety). *Assume the momentary safety cost d is α-Lipschitz continuous. Assume that for all (deterministic) ego/other policies $\pi_t \in \Pi_t, \sigma_t \in \Phi_t$, $t \in 1{:}T$, the dynamics $s \mapsto f(s, \pi_t(s), \sigma_t(s))$ as well as $a \mapsto f(s, a, \sigma_t(s))$ for fixed s are β-Lipschitz. Then $a \mapsto w_t(s, a)$ is $\alpha \max\{1, \beta^T\}$-Lipschitz.*

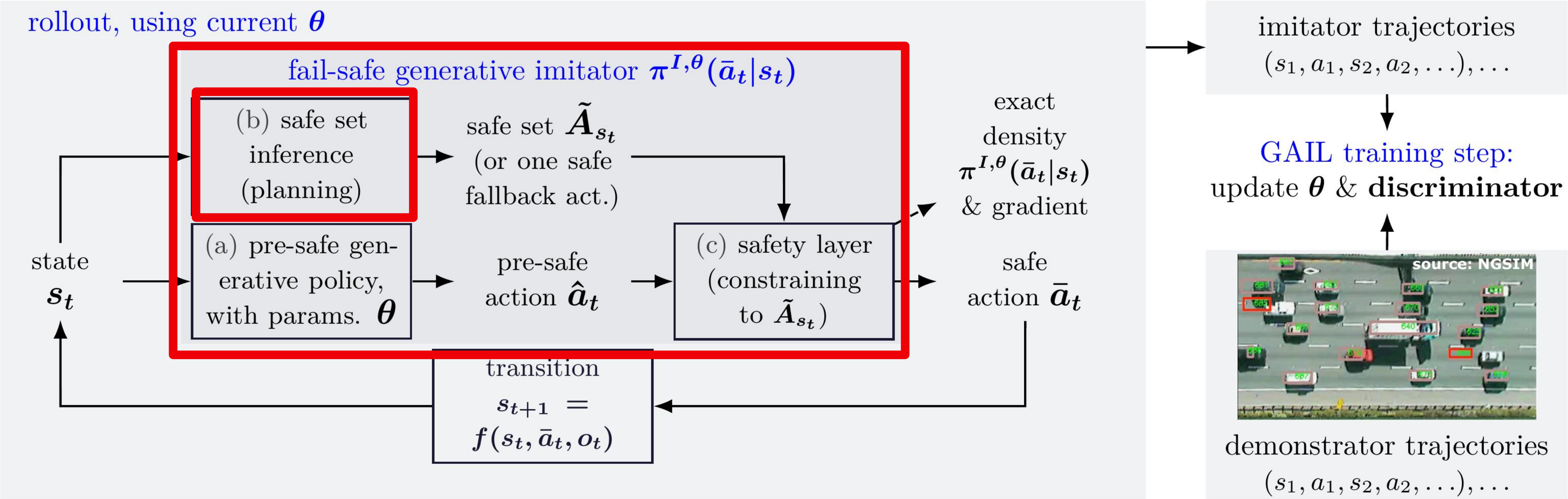
action set



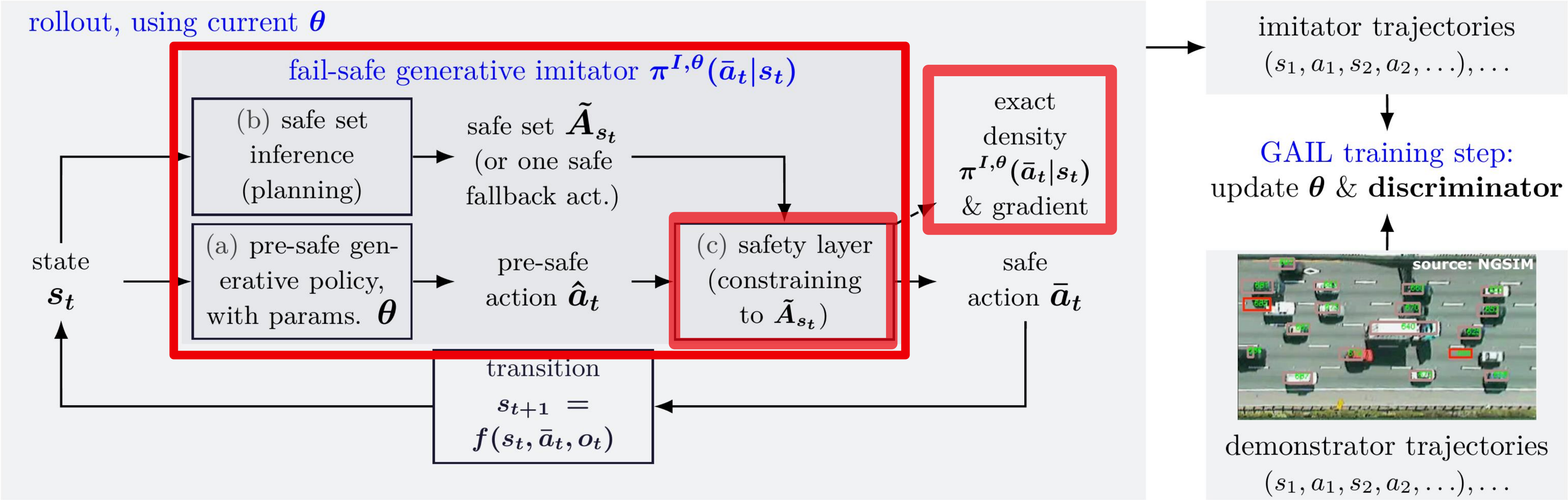
$$\text{Safety radius} = \frac{w_t(s,a)}{\alpha \max\{1, \beta^T\}}$$

Fail-Safe Adversarial Generative Imitation Learning

Outline of our method



Fail-Safe Adversarial Generative Imitation Learning
Outline of our method



BOSCH

Fail-Safe Adversarial Generative Imitation Learning

Safety layer with closed-form probability density/gradient I

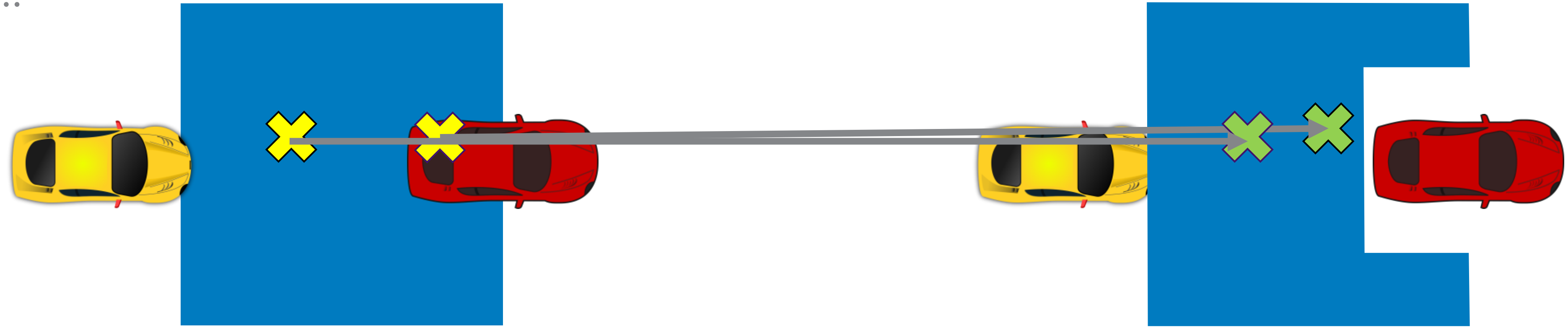
Our final neural net layer guarantees *safety* of actions:

… Pre-safe policy neural net layers …

complete action set

Safety layer
(final layer in policy neural net)

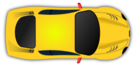
safe action set



Pre-safe action can be anywhere

Safe action always in safe action set

Legend:

 = current state

 = “action” = next state

BOSCH

Fail-Safe Adversarial Generative Imitation Learning

Safety layer with closed-form probability density/gradient II

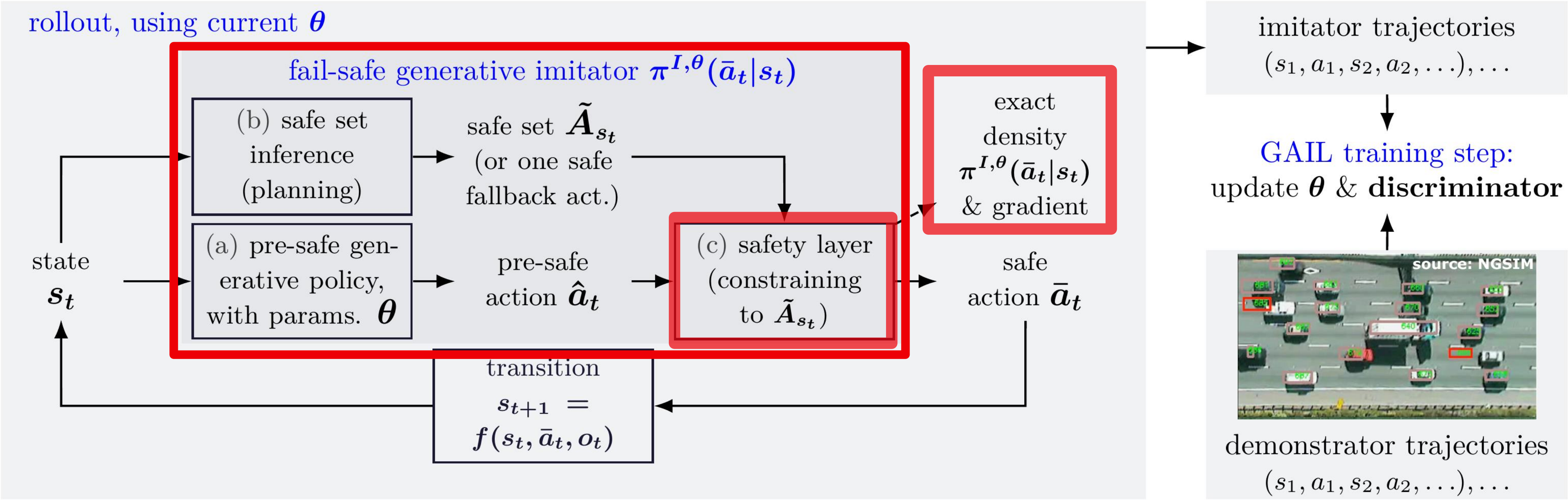
- We want to use the **change-of-variables formula**, but its **injectivity** requirements are too rigid!
- So we combine change of variables with additivity of measures to allow for countable non-injectivity
- by using ``piecewise **diffeomorphisms”** as mappings for safety layers

Proposition 3 (Closed-form density for piecewise diffeomorphism). *If g is such a piecewise diffeomorphism, $\bar{a} = g(\hat{a})$ and $\hat{a}$'s density is $p_{\hat{a}}(\hat{a})$, then $\bar{a}$'s density is*

$$p_{\bar{a}}(\bar{a}) = \sum_{k:\bar{a}\in g_k(A_k)} |\det(J_{g_k^{-1}}(\hat{a}))| p_{\hat{a}}(g_k^{-1}(\bar{a})). \tag{5}$$

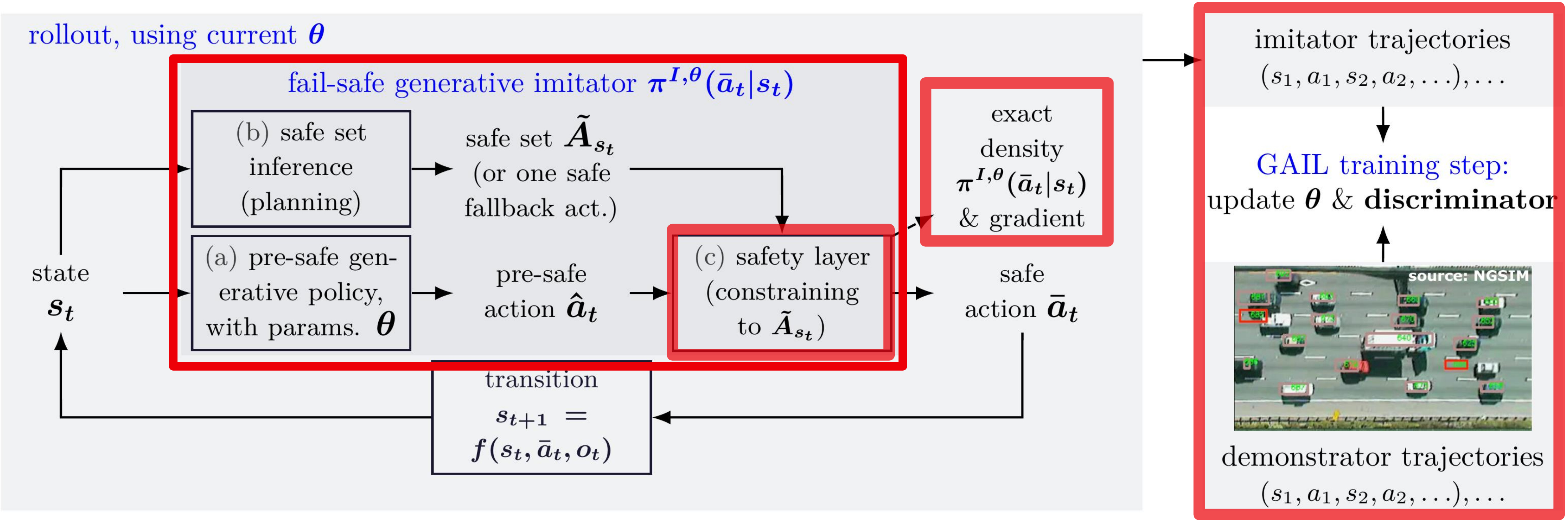
This gives us **closed-form differentiable policy density** $\pi^{I\theta}(\bar{a}|s)$ **and gradient** $\nabla_\theta \pi^{I\theta}(\bar{a}|s)$,
for **policy-gradient based training** (like GAIL, using, e.g., SAC, PG, ...)!

Fail-Safe Adversarial Generative Imitation Learning
Outline of our method



Internal | CR/AIR4.2 | 2023-09-15

Fail-Safe Adversarial Generative Imitation Learning
Outline of our method



BOSCH

Fail-Safe Adversarial Generative Imitation Learning

Imitation performance guarantees w.r.t. safety layers I

Performance difference: test-time-only safety layer versus train-and-test time safety layer (ours)?

Remark 1 (Linear error in T of end-to-end train-and-test-time safety layer). *Assume $D_{TV}(\rho^I, \rho^D) \leq \varepsilon$. Then we get*

$$|v^I - v^D| \leq 2\varepsilon T \|c^*\|_\infty.$$

BOSCH

Fail-Safe Adversarial Generative Imitation Learning
Imitation performance guarantees w.r.t. safety layers II

Performance difference: test-time-only safety layer versus train-and-test time safety layer (ours)?

Theorem 1 (Quadratic error in T of test-time-only safety layer). ***Lower bound*** *(an "existence" statement): We can construct an environment*[11] *with variable horizon T and with a demonstrator, sketched in Fig. 2 and additional details in Appendix A.3.2, a universal constant ι, and, for every $\varepsilon > 0$, an unconstrained imitator π^U with $D_{TV}(\rho^D, \rho^U) \leq \varepsilon$, such that for the induced test-time constrained imitator π^O we have, for all $T \geq 2^{12}$,*

$$|v^O - v^D| \geq \iota \min\{\boxed{\varepsilon T^2}, T\} \|c^*\|_\infty. \tag{6}$$

Upper bound *(a "for all" statement): Assume $D_{TV}(\rho^D, \rho^U) \leq \varepsilon$ and assume $\rho^U(s)$ has support wherever $\rho^D(s)$ has. Then*

$$|v^O - v^D| \leq \frac{4\varepsilon}{\nu} \boxed{T^2} \|c^*\|_\infty,$$

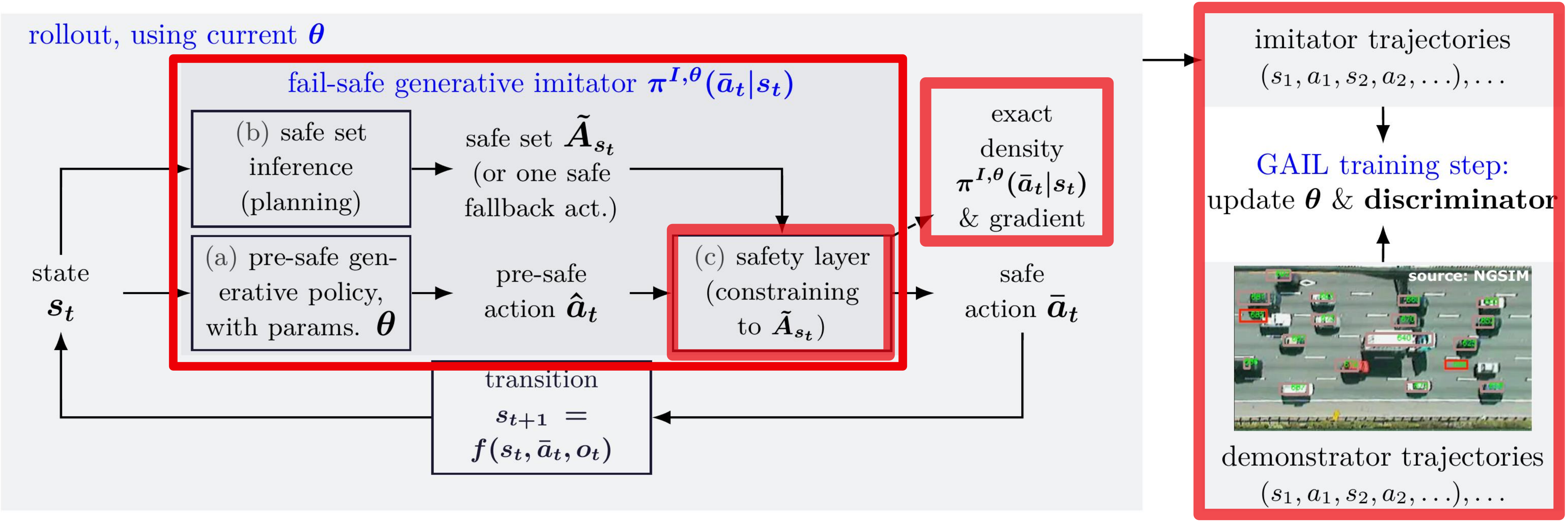
where ν is the minimum mass of $\rho^D(s)$ within the support of $\rho^D(s)$.

- T = rollout horizon
- both results are on population-level performance during test time

BOSCH

Fail-Safe Adversarial Generative Imitation Learning

Outline of our method



Internal | CR/AIR4.2 | 2023-09-15

BOSCH

Fail-Safe Adversarial Generative Imitation Learning

Experiments: driver imitation – safety and imitation performance

Method		*Imitation performance*		*Safety performance*
Pre-safe	Overall	ADE	FDE	Probability of crash/off-road
Gauss	FAGIL-E (ours)	0.59	1.70	0.00
	FAGIL-L (ours)	0.60	1.77	0.00
	GAIL Ho and Ermon (2016)	0.47	1.32	0.13
	RAIL Bhattacharyya et al. (2020)	0.48	1.35	0.22
	TTOS (Sec. 3.3)	0.60	1.78	0.00
Flow	FAGIL-E (ours)	0.58	1.69	0.00
	FAGIL-L (ours)	0.57	1.68	0.00
	GAIL Ho and Ermon (2016)	0.44	1.22	0.11
	RAIL Bhattacharyya et al. (2020)	0.53	1.50	0.11
	TTOS (Sec. 3.3)	0.59	1.72	0.00

Each method in two versions: *Gauss* vs. *Normalizing Flow* as *“pre-safe policy”*

Dataset: “highD” (highway driver trajectories)

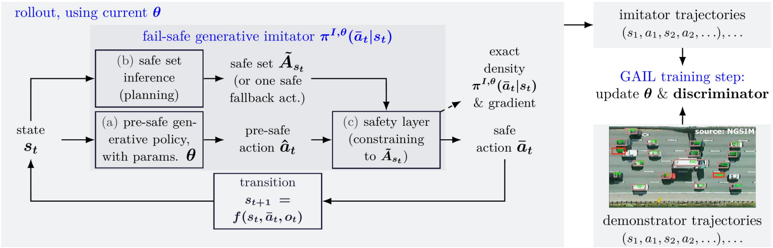
- ADE: average displacement error.
- FDE: final displacement error
- GAIL: Generative Adversarial Imitation Learning
- RAIL: Reward-augmented GAIL
- TTOS: “Test-Time-Only Safety” (train GAIL, then add safety layer at test time)

Conclusions

Fail-Safe Adversarial Generative Imitation Learning
Conclusions

- Machine learning / imitation learning on the rise for autonomous driving
- **But big open challenge to make it safe – inherent uncertainty in deployed ML/IL performance**
- Showed rough landscape of possible approaches for mathematically validated safe ML

- Our specific approach builds on generative adversarial imitation learning (GAIL) and adds
  - sample-based reachability analysis for guaranteed safe action sets,
  - safety layers with closed-form density/gradient via “piecewise” change-of-variables,
  - and the theoretical understanding of end-to-end generative training with safety layers.



- **We are always looking for students for internships and master theses with an ML background!**