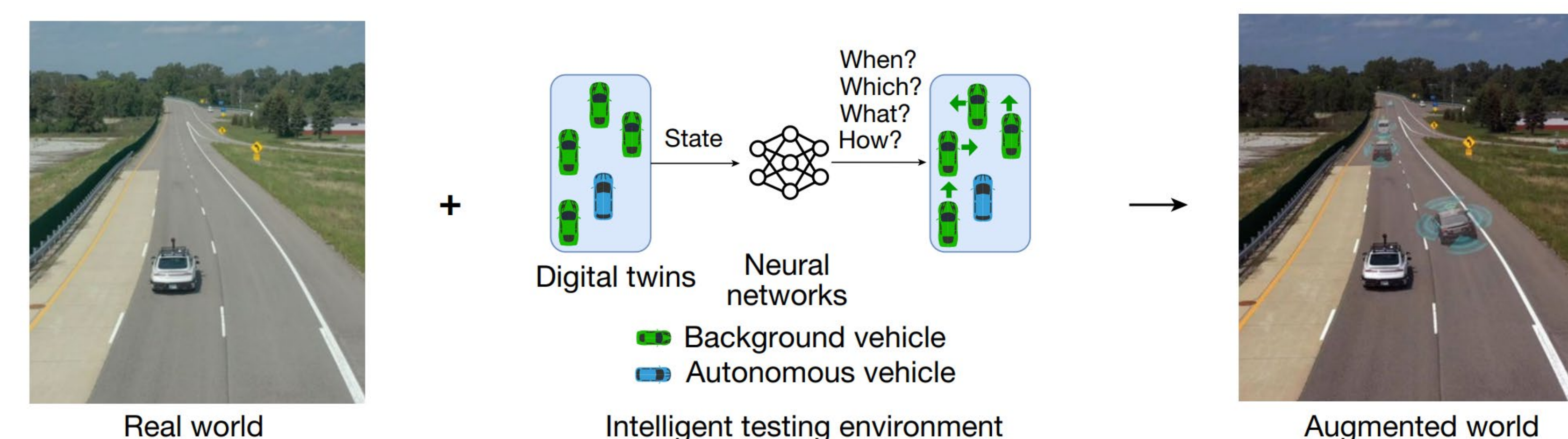# Dense Reinforcement Learning for Safety Validation of Autonomous Vehicles

Shuo Feng, Haowei Sun, Xintao Yan, Haojie Zhu, Zhengxia Zou, Shengyin Shen & Henry X. Liu
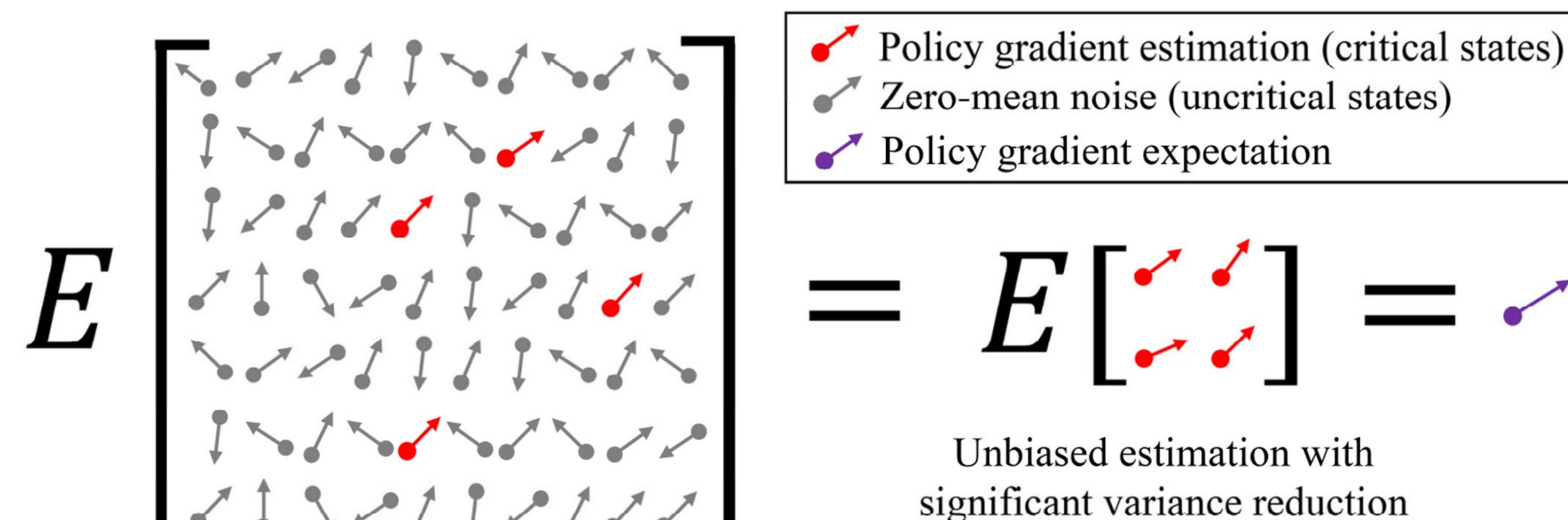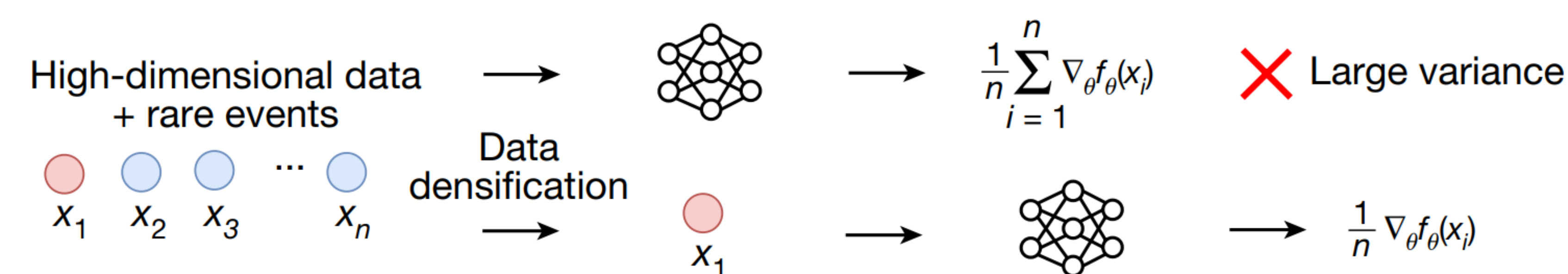
Introduction

- **AV testing under naturalistic driving environment** requires high economic and time costs: rarity of safety critical events

- **Dense deep-reinforcement-learning (D2RL):** remove non-safety-critical-states and densify the information

- **Naturalistic and Adversarial Driving Environment (NADE):** Improve the evaluation efficiency by 10^3 to 10^5 times

- **Field test** for highly automated vehicle (Autoware) in both highway and urban test tracks using **augmented reality (AR)**
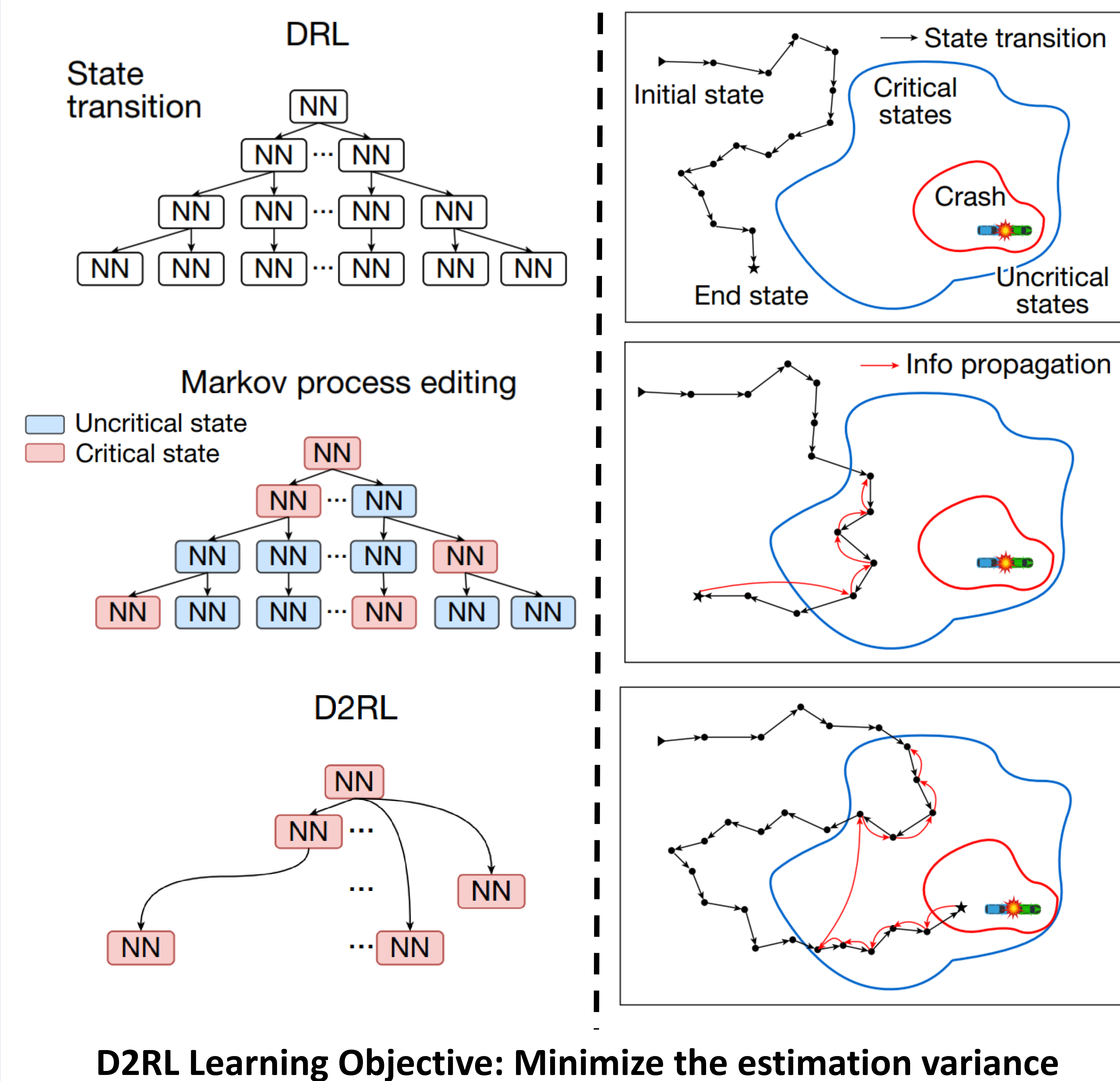
Highlights

Real world   Intelligent testing environment   Augmented world

Digital twins   Neural networks
- Background vehicle
- Autonomous vehicle

Dense Learning

High-dimensional data + rare events $\rightarrow$ $\frac{1}{n}\sum_{i=1}^{n}\nabla_\theta f_\theta(x_i)$ ✗ Large variance

Data densification $\rightarrow$ $\frac{1}{n}\nabla_\theta f_\theta(x_i)$

x_1 x_2 x_3 ... x_n

- Policy gradient estimation (critical states)
- Zero-mean noise (uncritical states)
- Policy gradient expectation

$$E\begin{bmatrix} \end{bmatrix} = E\begin{bmatrix} \end{bmatrix} = $$

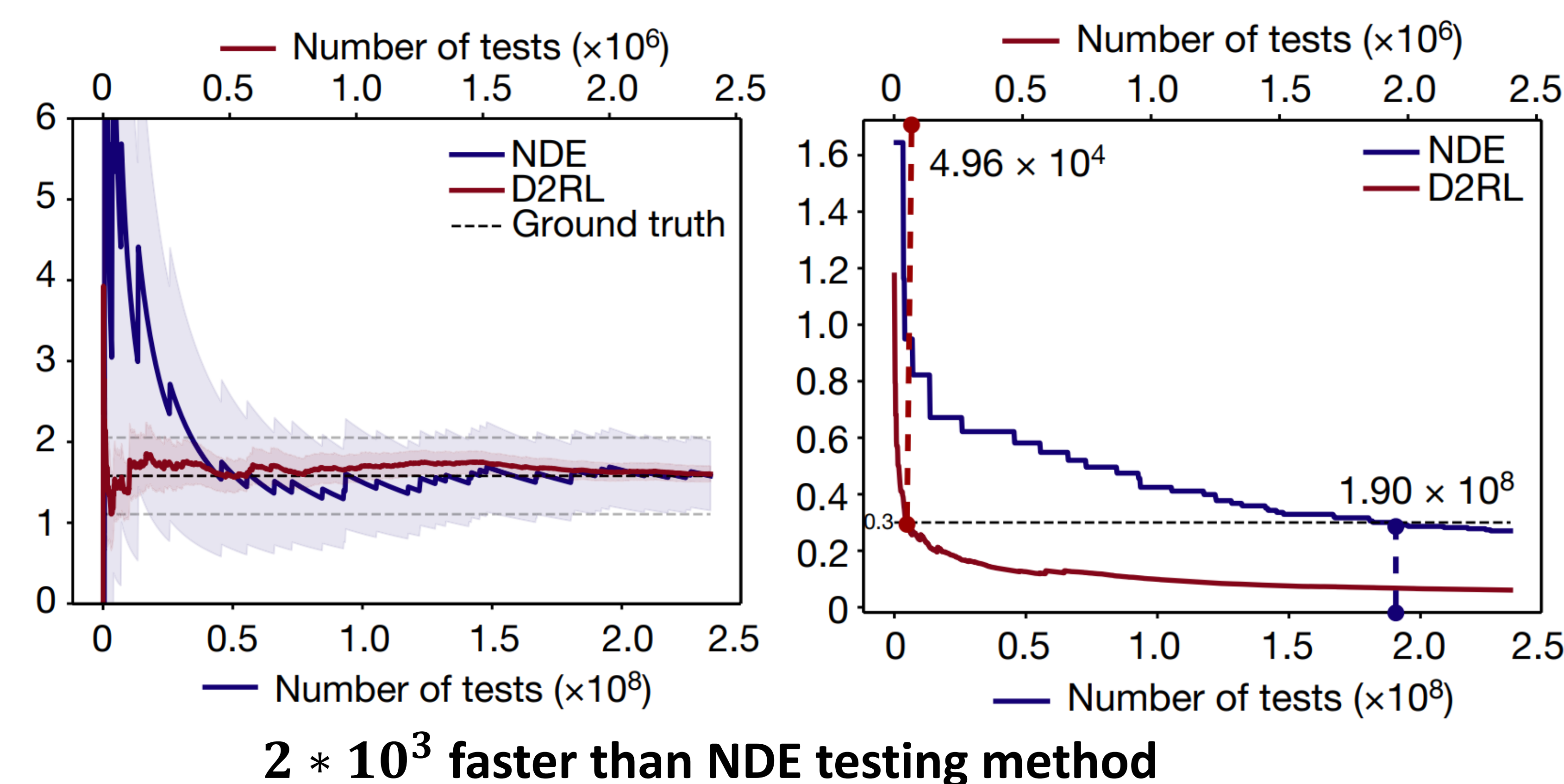Unbiased estimation with significant variance reduction

- **Neural Network is trained with informative data only**
- **Dense-learning approach reduces gradient estimation variance**
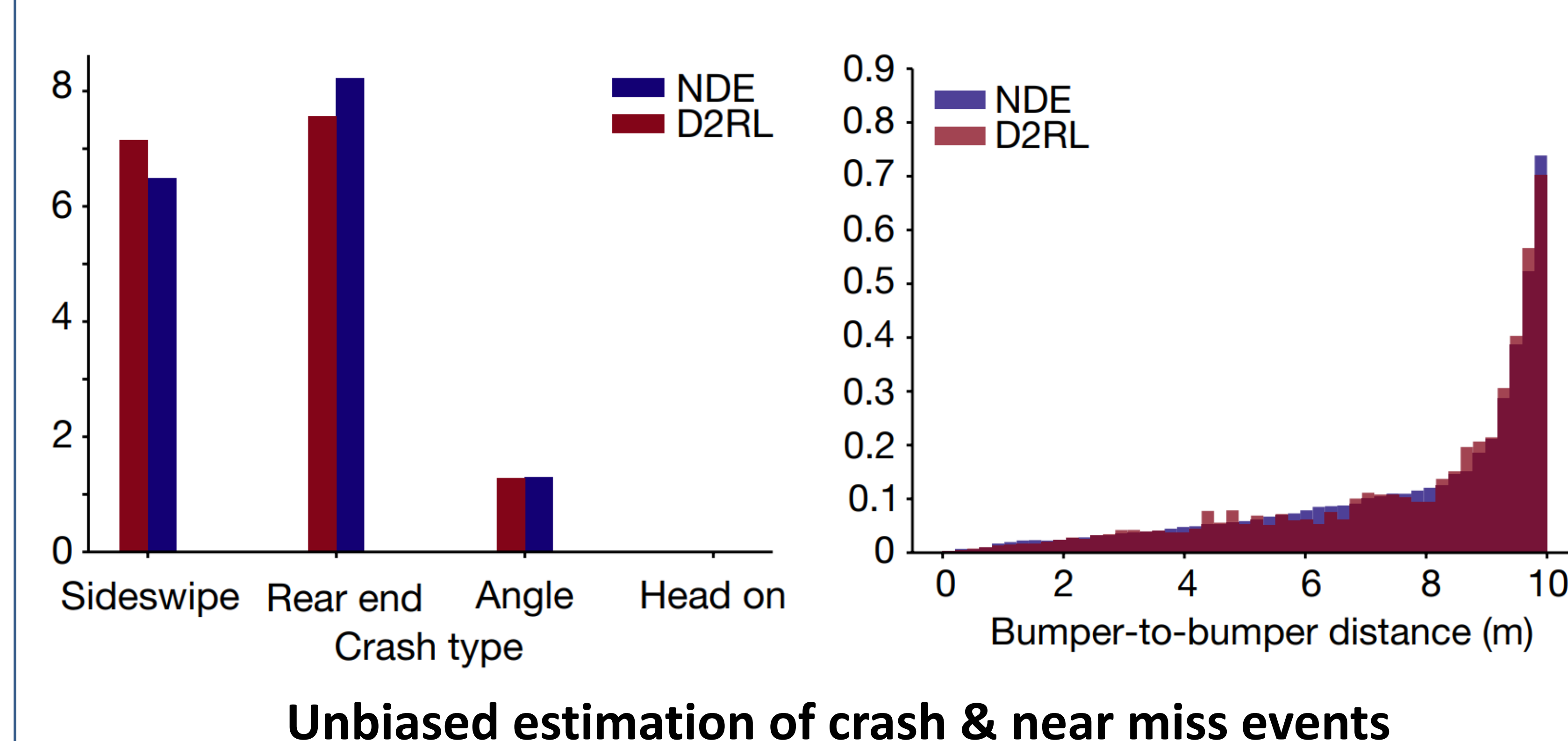- **Enable deep-learning applications in safety-critical systems**

Dense Deep Reinforcement Learning

DRL

State transition

→ State transition
Initial state
Critical states
Crash
End state
Uncritical states

Markov process editing

- Uncritical state
- Critical state

→ Info propagation

D2RL

D2RL Learning Objective: Minimize the estimation variance

Simulation Test

Number of tests (×10⁶)
NDE
D2RL
Ground truth
Number of tests (×10⁸)

4.96×10^4
1.90×10^8
0.3

$2 * 10^3$ faster than NDE testing method

Simulation Test (Cont'd)

NDE
D2RL

Crash type: Sideswipe, Rear end, Angle, Head on

Bumper-to-bumper distance (m)

Unbiased estimation of crash & near miss events

Field Test at ACM and Mcity

a — By-wire control, On-board cameras, LiDAR, DSRC OBU, Radar sensors, Image rendering, IMU + RTK, On-board computer

b — Two-lane, Highway loop 50 mph, Overpass, Highway entrance, Highway exit, Highway loop 65 mph, Three-lane, Two-lane, Overpass

c — Roundabout, Four-way stop intersection, Signalized intersections, Unprotected left-turn, Roundabout, Roundabout, 1,000-foot highway

d — Simulation view, (15.67 mph) Foward >>, Ego vehicle view, Front camera view, Augmented camera view

e — Crash rate (×10⁻⁶), Number of tests at ACM (×10⁵), D2RL

g — Crash rate (×10⁻⁵), Number of tests at Mcity (×10⁵), D2RL

i — Crash rate (×10⁻⁷), Crash types at Mcity: Sideswipe, Rear End, Angle, Head On, D2RL

10^5 faster than NDE testing method

Feng, S., Sun, H., Yan, X., Zhu, H., Zou, Z., Shen, S., and Liu H.X. (2023). Dense reinforcement learning for safety validation of autonomous vehicles. Nature 615, 620–627. https://doi.org/10.1038/s41586-023-05732-2

UNIVERSITY OF MICHIGAN