Neural Policy Synthesis and Verification with Guarantees

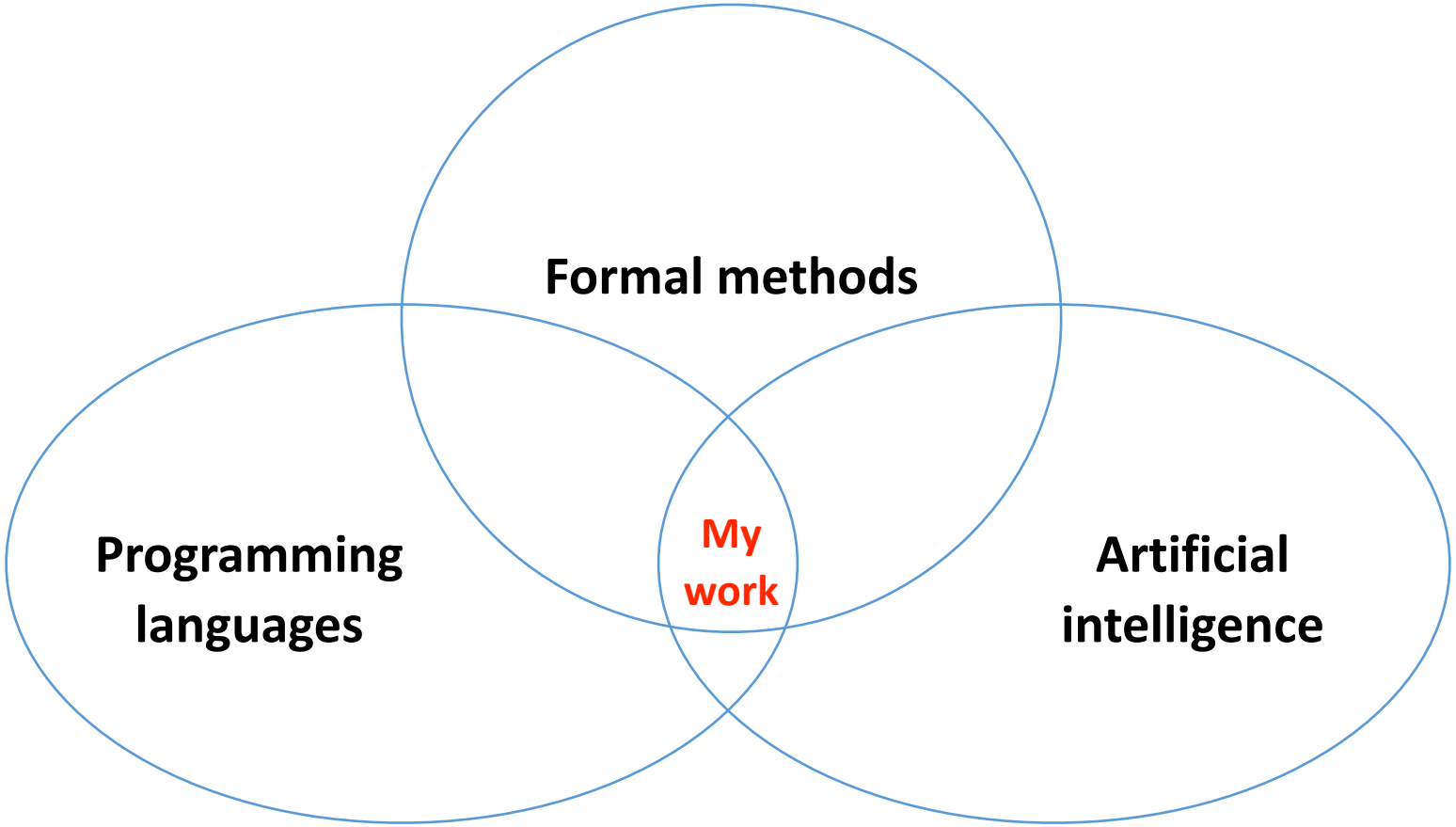
Djordje Zikelic

SCIS Research Cluster Seminar Series, August 2024

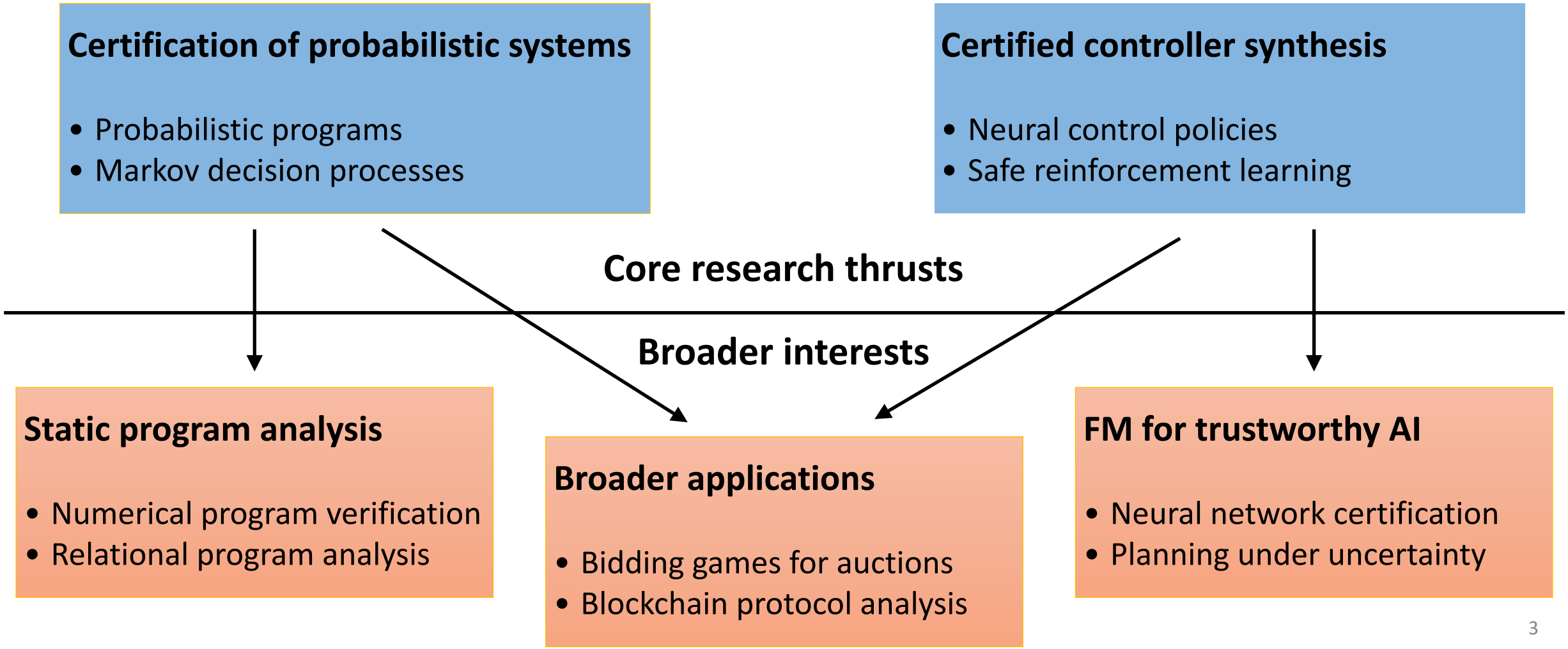
SMU
SINGAPORE MANAGEMENT UNIVERSITY

School of
Computing and
Information Systems

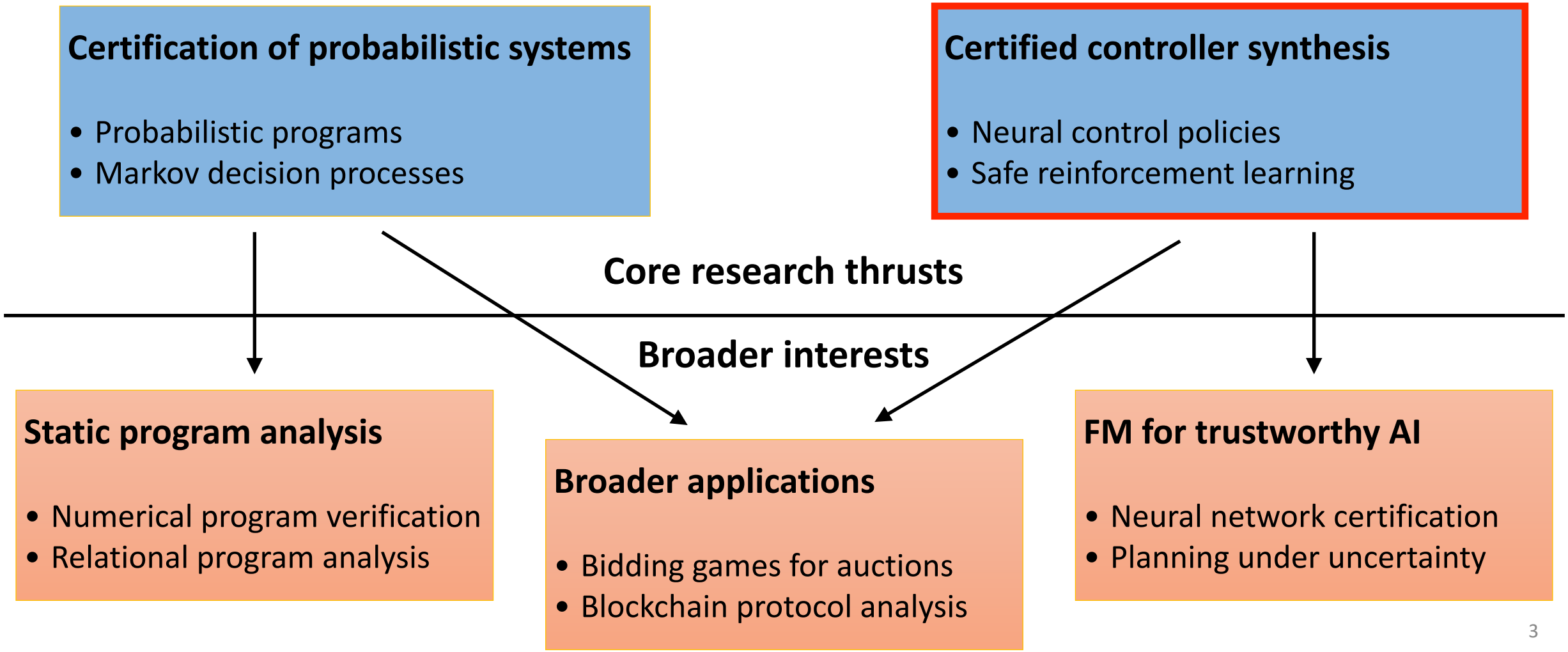
Short research introduction



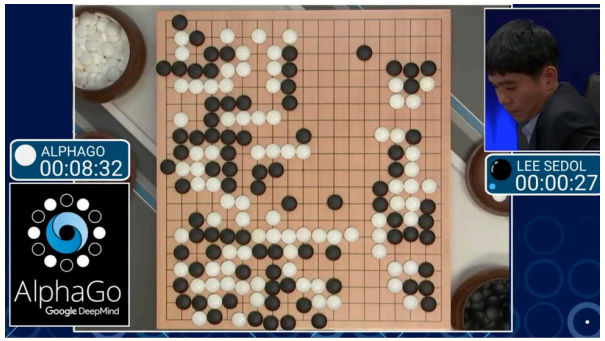
Certification of software and AI systems in the presence of **uncertainty**



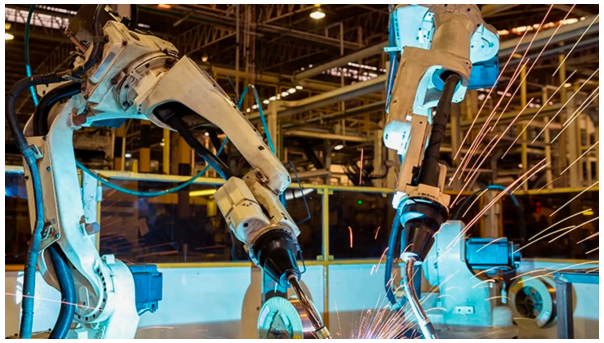
Certification of software and AI systems in the presence of **uncertainty**



Learning-enabled Policy (or Controller) Synthesis



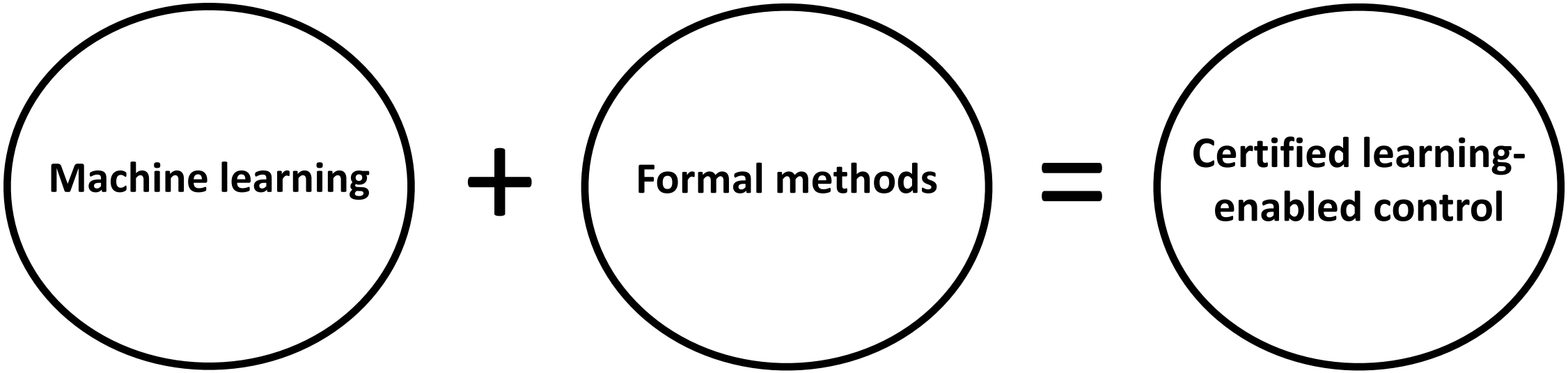
Not safety-critical



Safety-critical

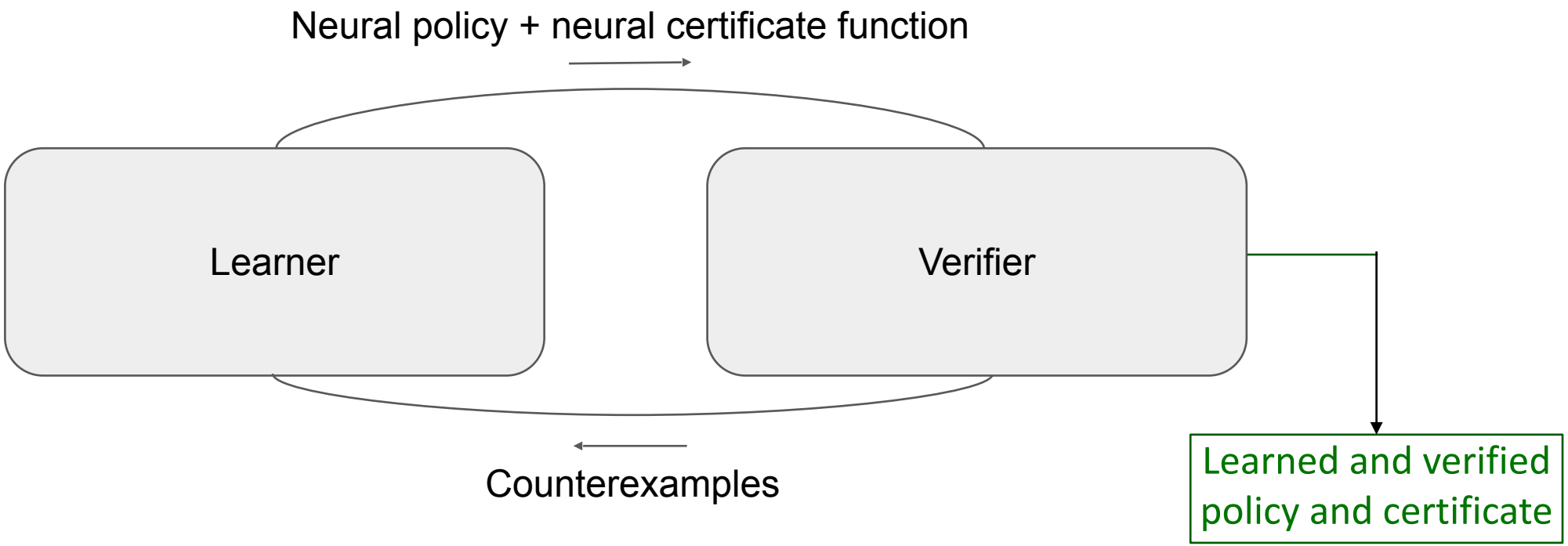
Safety-critical applications require **correctness guarantees**

New paradigm for certified neural control

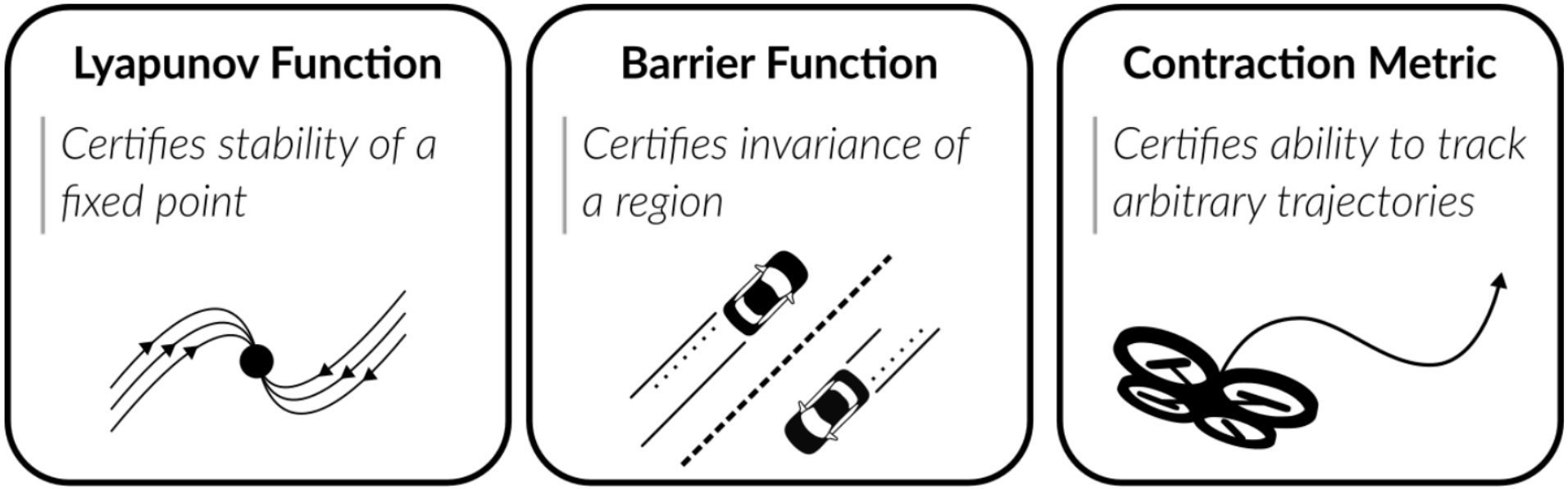


Key idea: Learn neural policy + neural certificate of correctness

Certified control through a learner-verifier loop



Prior work: Deterministic systems*



Learning neural policies with classical control theory certificates
(+ verification by reduction to SMT-solving)

*Image taken from: Dawson, Gao, Fan. *Safe Control with Learned Certificates: A Survey of Neural Lyapunov, Barrier, and Contraction Methods for Robotics and Control.* IEEE Transactions on Robotics. 2023

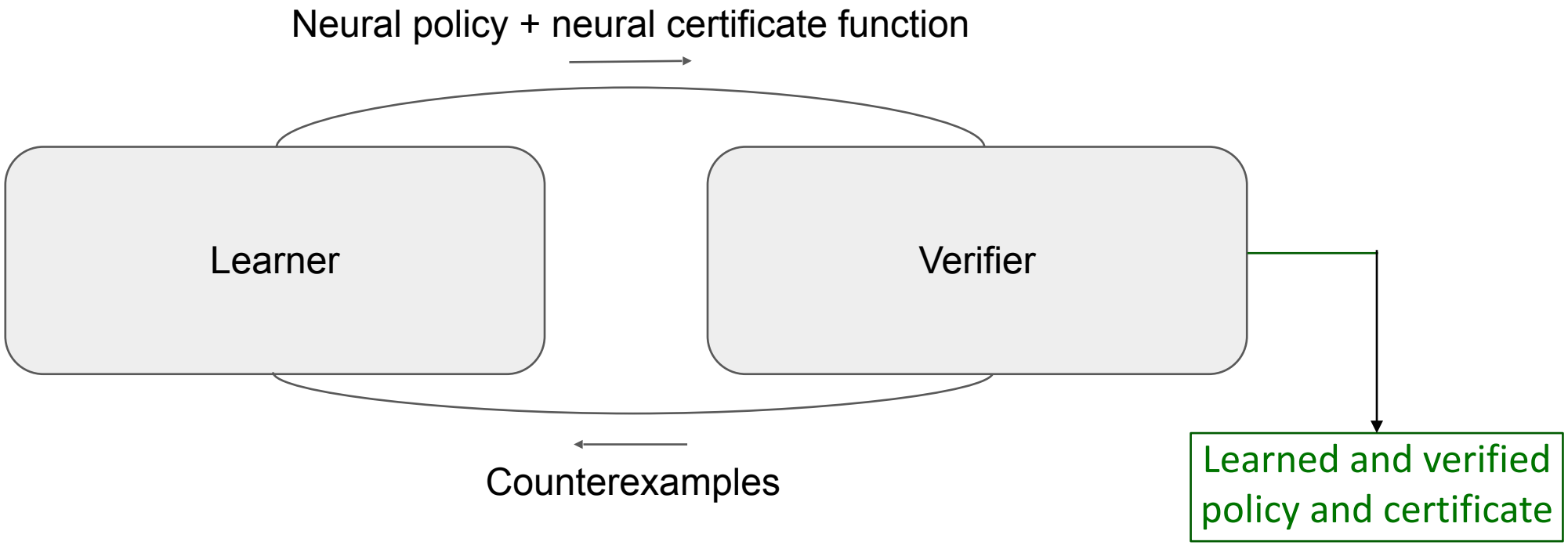
Can we guarantee neural policy correctness
in the presence of **environment uncertainty**?

A Learner-verifier Framework for Neural Stochastic Control and Verification [NeurIPS23, AAAI23, TACAS23, ATVA23, AAAI22]

Joint with Mathias Lechner (MIT, Liquid AI), Krishnendu Chatterjee (ISTA), Thomas A. Henzinger (ISTA),

Matin Ansaripour (ATVA'23, EPFL), Abhinav Verma (NeurIPS'23, Penn State)

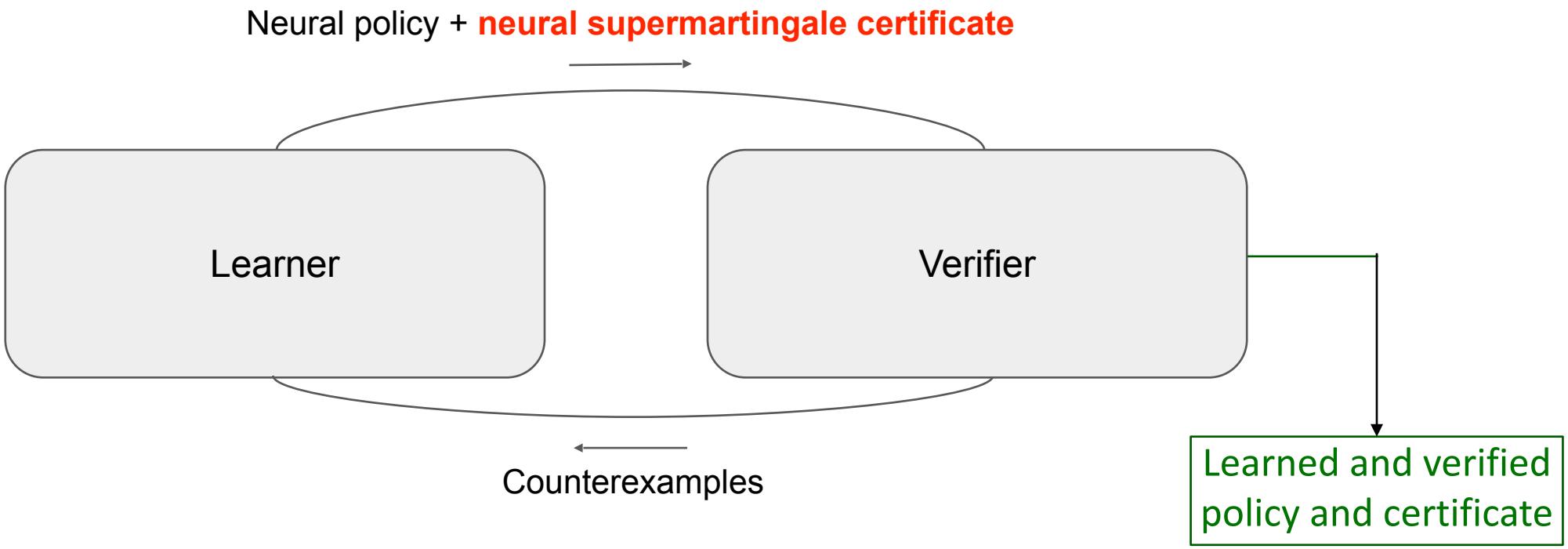
Certified control through a learner-verifier loop



Q1: What should be the certificates for stochastic dynamical systems?

Q2: How to learn and formally verify neural policies and certificates?

How to formally verify learned controllers in stochastic dynamical systems?



What are supermartingales?

Supermartingale – stochastic process <u>decreasing in expectation</u> $\mathbb{E}[X_{n+1} \mid X_n] \leq X_n$

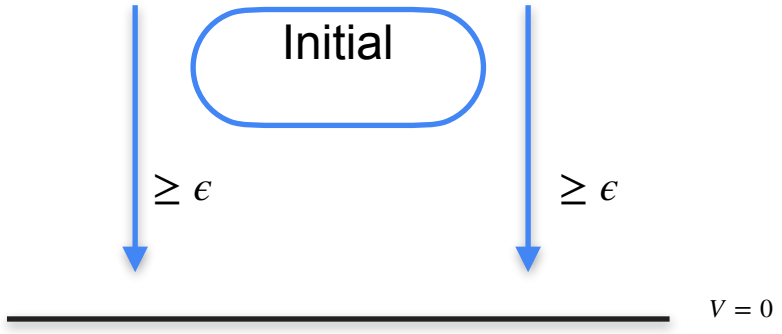
Supermartingale certificates in stochastic control

Ranking supermartingales (RSMs) for probability 1 reachability

[Kushner, Transactions on Automatic Control 1966, Chakarov, Sankaranarayanan, CAV 2013]

A measurable function $V : X \to \mathbb{R}$ for a target set X_t such that:

1. Nonnegativity. $V(x) \geq 0$ for $x \in X$
2. Strict expected decrease. $\exists \epsilon > 0$ s.t. $\mathbb{E}_{w \sim d}[V(f(x, \pi(x), w))] \leq V(x) - \epsilon$ for $x \in X \backslash X_t$



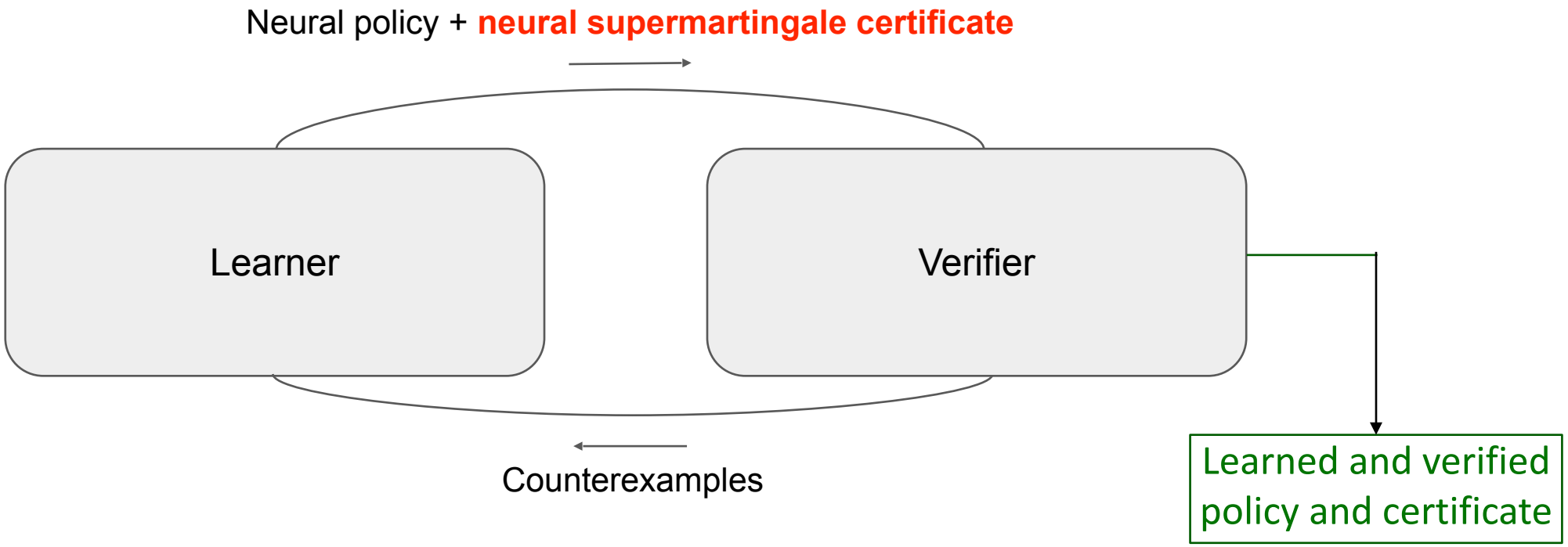
Stochastic barrier functions for probability $p \in [0,1]$ safety

[Prajna, Jadbabaie, Pappas. CDC 2004]

Contribution 1: Supermartingale certificates for more general properties

- Reach-avoidance [AAAI23]
- Sequential and boolean compositions of reach-avoidance [NeurIPS23]
- Stability (a.k.a. reach-and-stay) [ATVA23]
- Linear-time temporal logic (LTL) — work in progress!

Contribution 2: The first learner-verifier loop for stochastic control systems



$$\mathscr{L}(\theta, \nu) = \mathscr{L}_{\text{Init}}(\nu) + \mathscr{L}_{\text{Unsafe}}(\nu) + \mathscr{L}_{\text{Decrease}}(\theta, \nu)$$

Loss encodes defining properties of supermartingale certificates

Discretization
+ Lipschitz analysis
+ abstract interpretation

Key features of the framework

1. Applicable to deterministic and stochastic control systems
2. Both formal synthesis and verification of neural policies
3. Hard formal guarantees

Experiments

Environment	Reach-avoidance probability
2D system	93.3%
Inverted pendulum	92.1%
Collision avoidance	90.4%

Experiments on reach-avoid tasks

✅ Formally certified neural policies

❗ Scalability bottleneck, low dimensional environments

Some ongoing and future directions

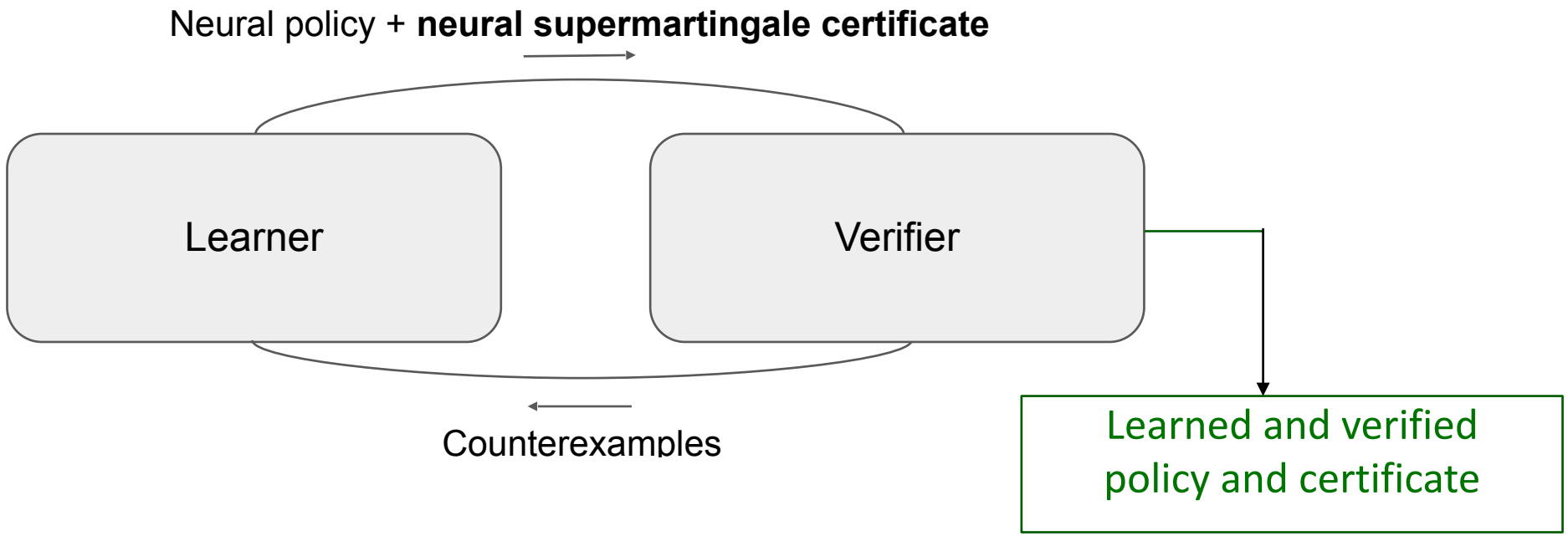
More general properties

Black-box setting & safe RL

Scalability

Multi-agent systems

Conclusion


Neural policy + **neural supermartingale certificate**
Learner
Verifier
Counterexamples
Learned and verified
policy and certificate