Neural Lyapunov Control

UC San Diego

Ya-Chien Chang, Nima Roohi, Sicun Gao

NEURAL INFORMATION PROCESSING SYSTEMS

OVERVIEW

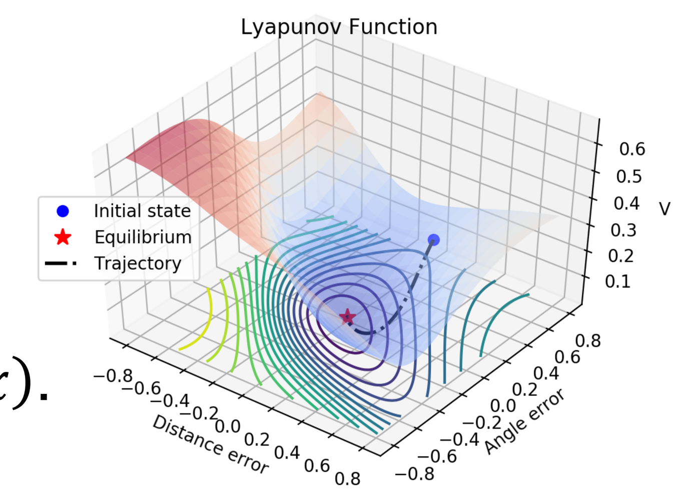
We propose new methods for learning control policies and neural network (NN) Lyapunov functions for nonlinear control problems, with provable guarantee of stability.

- The framework consists of a learner that attempts to find the control and Lyapunov functions, and a falsifier that finds counterexamples to quickly guide the learner towards solutions.
- The approach significantly simplifies the process of Lyapunov control design, provides end-to-end correctness guarantee, and can obtain much larger regions of attraction (ROA) than existing methods such as linear-quadratic regulators (LQR) and sum-of-squares (SOS) and semidefinite programming (SDP).

PRELIMINARIES

Proposition 1 (Lyapunov Functions for Asymptotic Stability)
Consider a controlled system $\frac{dx}{dt} = \mathrm{f_u}(x)$ with equilibrium at the origin, and suppose there exists a continuously differentiable function $\mathrm{V}: \mathcal{D} \to \mathbb{R}$ that satisfies the following conditions:

$$V(0) = 0,$$
$$\forall x \in \mathcal{D} \setminus \{0\}, V(x) > 0,$$
$$\forall x \in \mathcal{D} \setminus \{0\}, \nabla_{f_u} V(x) < 0,$$



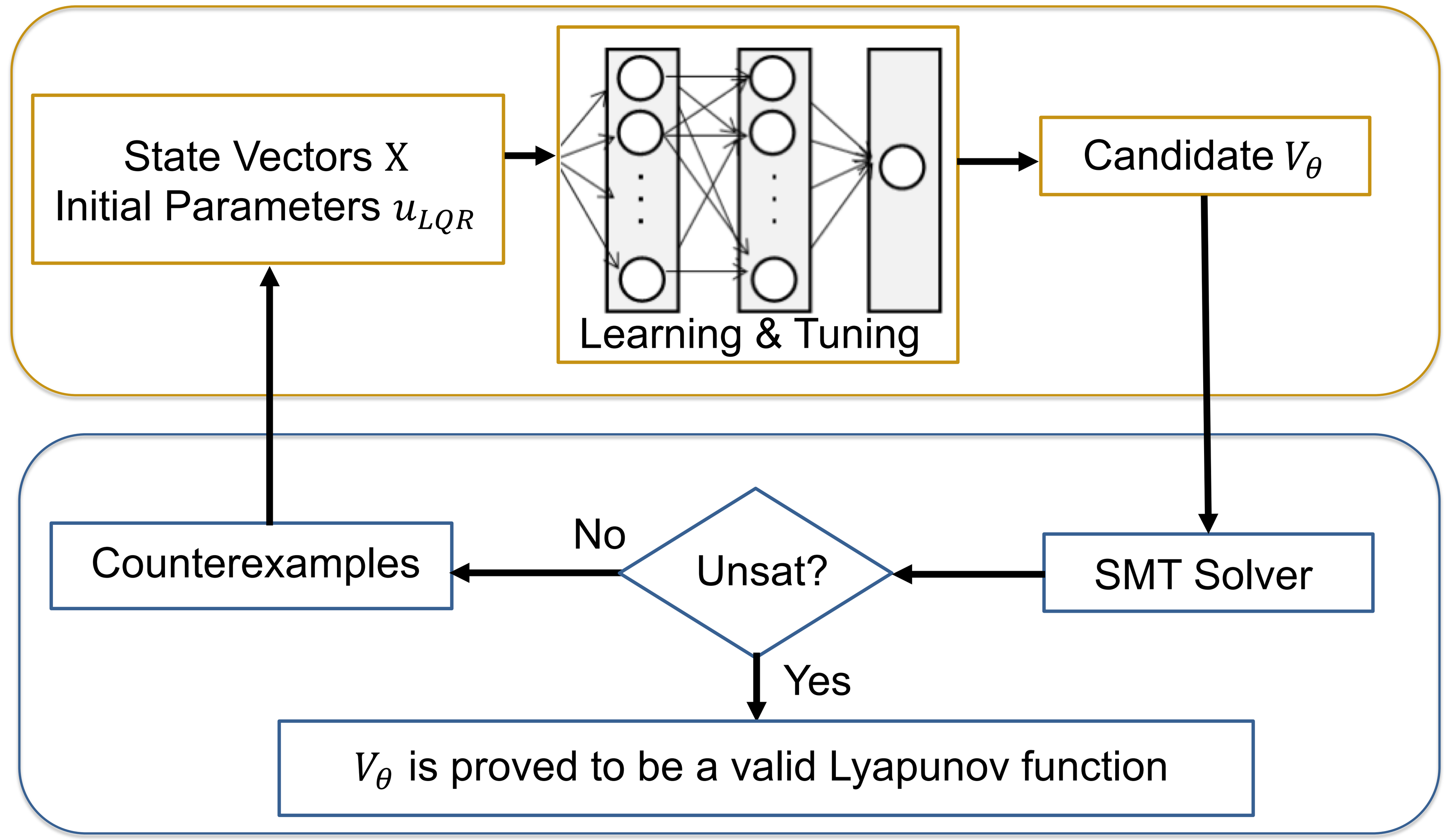
where $\nabla_{f_u} V(x) = \sum_{i=1}^{n} \frac{\partial V}{\partial x_i} \frac{dx_i}{dt} = \sum_{i=1}^{n} \frac{\partial V}{\partial x_i} [f_u]_i(x)$.

Then, the system is asymptotically stable at the origin, and V is called a Lyapunov function.

PROCEDURE

Learner :
Updates parameters in both a control function u and a neural Lyapunov function V_θ by iteratively minimizing the following Lyapunov risk:

$$L_\rho(\theta, u) =$$
$$\mathbb{E}_{x\sim\rho(\mathcal{D})}\left(\max\big(0, -V_\theta(x)\big) + \max\left(0, \nabla_{f_u} V_\theta(x)\right) + V_\theta^2(0)\right).$$



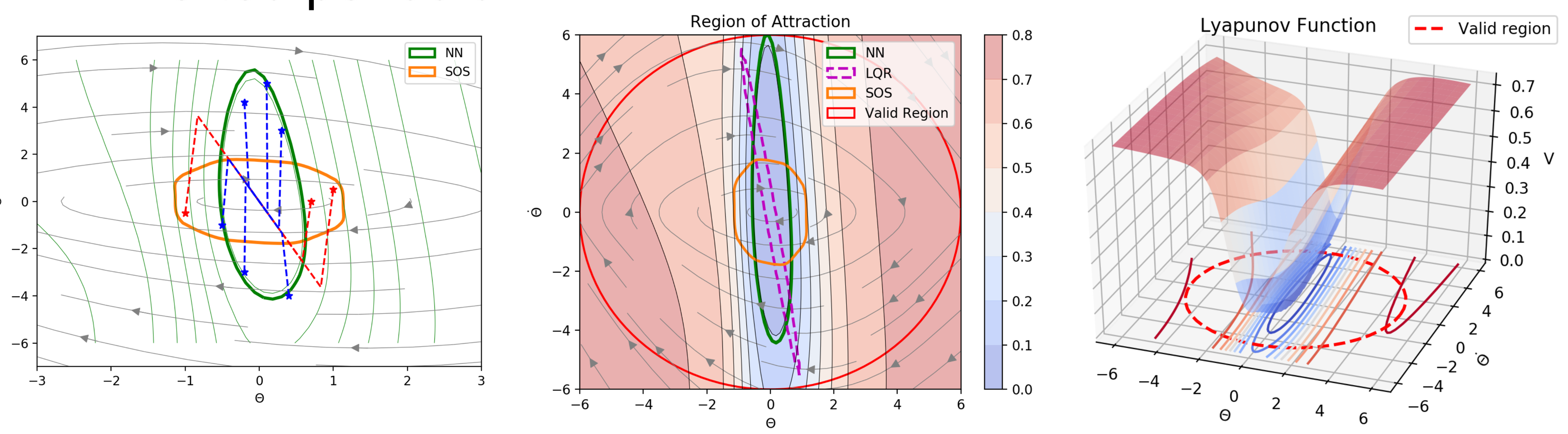
Falsifier :
Solves nonlinear constraint problems over real numbers to search for counterexample state vectors that violate the Lyapunov conditions. The falsification constraints are defined as follows:

$$\Phi_\varepsilon(x) = \left(\sum_{i=1}^{n} x_i^2 \geq \varepsilon\right) \wedge \left(V(x) \leq 0 \vee \nabla_{f_u} V(x) \geq 0\right).$$

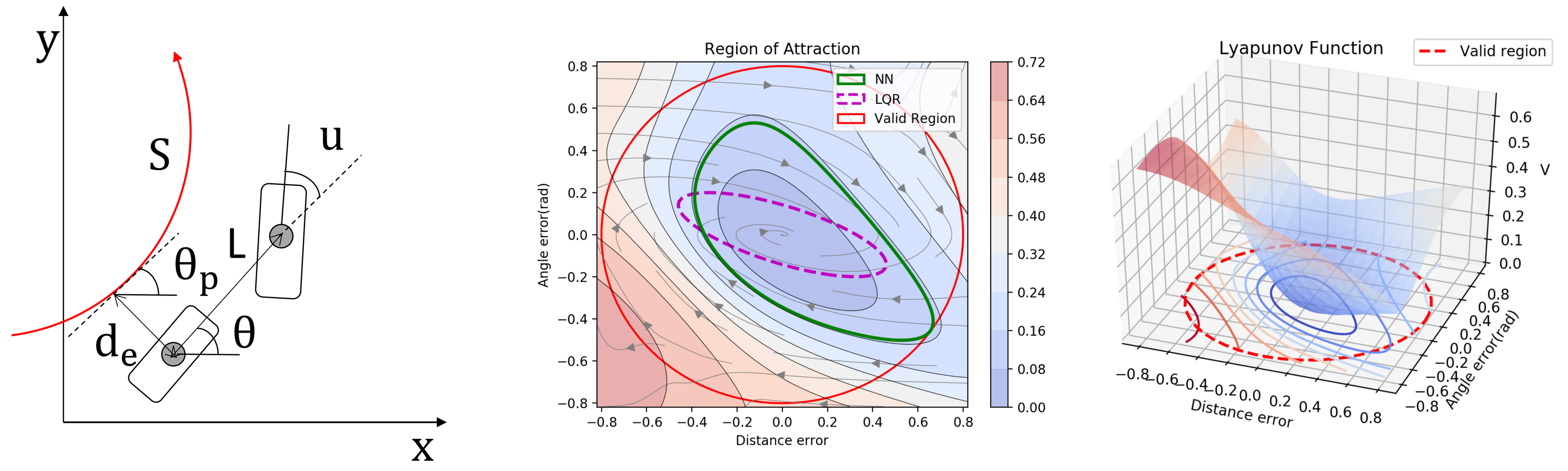
RESULTS

- In all examples, the regions of attraction are enlarged by more than three times compared with LQR results solved by defining both Q and R as identity matrices.
- In an inverted pendulum, we show trajectories (blue) are safely bounded within ROA defined by NN Lyapunov function. On the contrary, many trajectories (red) start inside the SOS region can escape.

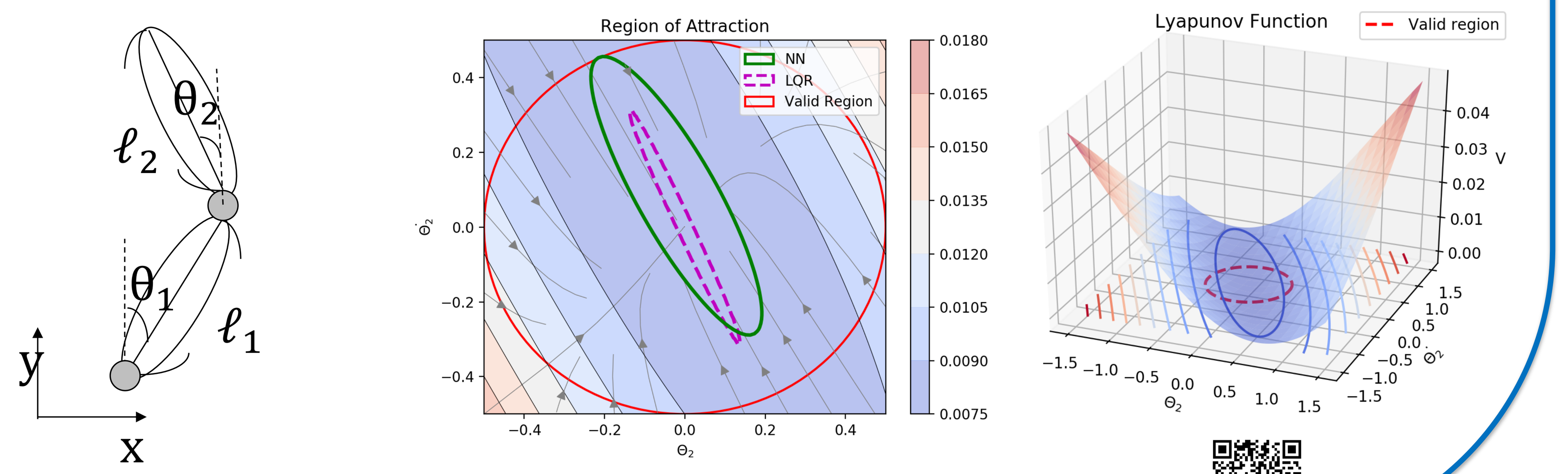
❖ Inverted pendulum



❖ Vehicle path following



❖ Humanoid balance



Project page available at