Neural Stochastic Control

JingDong Zhang
School of Mathematical Sciences
Fudan University

Background

- **Lyapunov Method in Machine Learning:** The recent work (Chang et al. ,2019) proposed an neural framework of learning the Lyapunov function and the linear control function simultaneously for stabilizing ODEs. The noise is ubiquitous in the real world systems, which calls for control methods in stochastic settings.
- **Stochastic Stability Theory of SDEs:** The positive effects of stochasticity have also been cultivated in control fields for SDEs. It inspires us to cultivate control method for SDEs with the help of noise, instead of focusing on deterministic control and regard noise as negative part.
- **Classic control methods:** Existing control methods for SDEs just use the hard constrained optimization to find online deterministic control and lack the exponential stability, we focus on learning offline neural stochastic control policy with stability guarantee.

Problem Statement

$$\mathrm{d}\boldsymbol{x}(t) = F(\boldsymbol{x}(t))\mathrm{d}t + G(\boldsymbol{x}(t))\mathrm{d}B_t,\ \ t \geq 0,\ \boldsymbol{x}(0) = \boldsymbol{x}_0 \in \mathbb{R}^d$$

we set $F(0) = 0$ and $G(0) = 0$ so that $x = 0$ is a zero solution

Assumption 2.1 *(Locally Lipschitzian Continuity) For every integer $n \geq 1$, there is a number $K_n > 0$ such that*

$$\|F(\boldsymbol{x}) - F(\boldsymbol{y})\| \leq K_n\|\boldsymbol{x} - \boldsymbol{y}\|,\ \|G(\boldsymbol{x}) - G(\boldsymbol{y})\|_{\mathrm{F}} \leq K_n\|\boldsymbol{x} - \boldsymbol{y}\|,$$

for any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d$ with $\|\boldsymbol{x}\| \vee \|\boldsymbol{y}\| \leq n$. $\longrightarrow$ Existence and uniqueness

How to stabilize the zero solution with the only diffusion term?

(⇔Benefit from noise)

Fundamental Theory

Exponential Stability

Theorem 2.2 *Mao (2007) Suppose that Assumptions 2.1 holds. Suppose further that there exist a function* $V \in C^2(\mathbb{R}^d; \mathbb{R}_+)$ *with* $V(\mathbf{0}) = 0$*, constants* $p > 0$*,* $c_1 > 0$*,* $c_2 \in \mathbb{R}$ *and* $c_3 \geq 0$*, such that* (i) $c_1\|\boldsymbol{x}\|^p \leq V(\boldsymbol{x})$, (ii) $\mathcal{L}V(\boldsymbol{x}) \leq c_2 V(\boldsymbol{x})$, *and* (iii) $|\nabla V^\top(\boldsymbol{x})G(\boldsymbol{x})|^2 \geq c_3 V^2(\boldsymbol{x})$ *for all* $\boldsymbol{x} \neq 0$ *and* $t \geq 0$*. Then,*

$$\limsup_{t\to\infty} \frac{1}{t} \log \|\boldsymbol{x}(t; t_0, \boldsymbol{x}_0)\| \leq -\frac{c_3 - 2c_2}{2p} \quad a.s.. \tag{2}$$

In particular, if $c_3 - 2c_2 > 0$*, the zero solution of Eq. (1) is exponentially stable almost surely.*

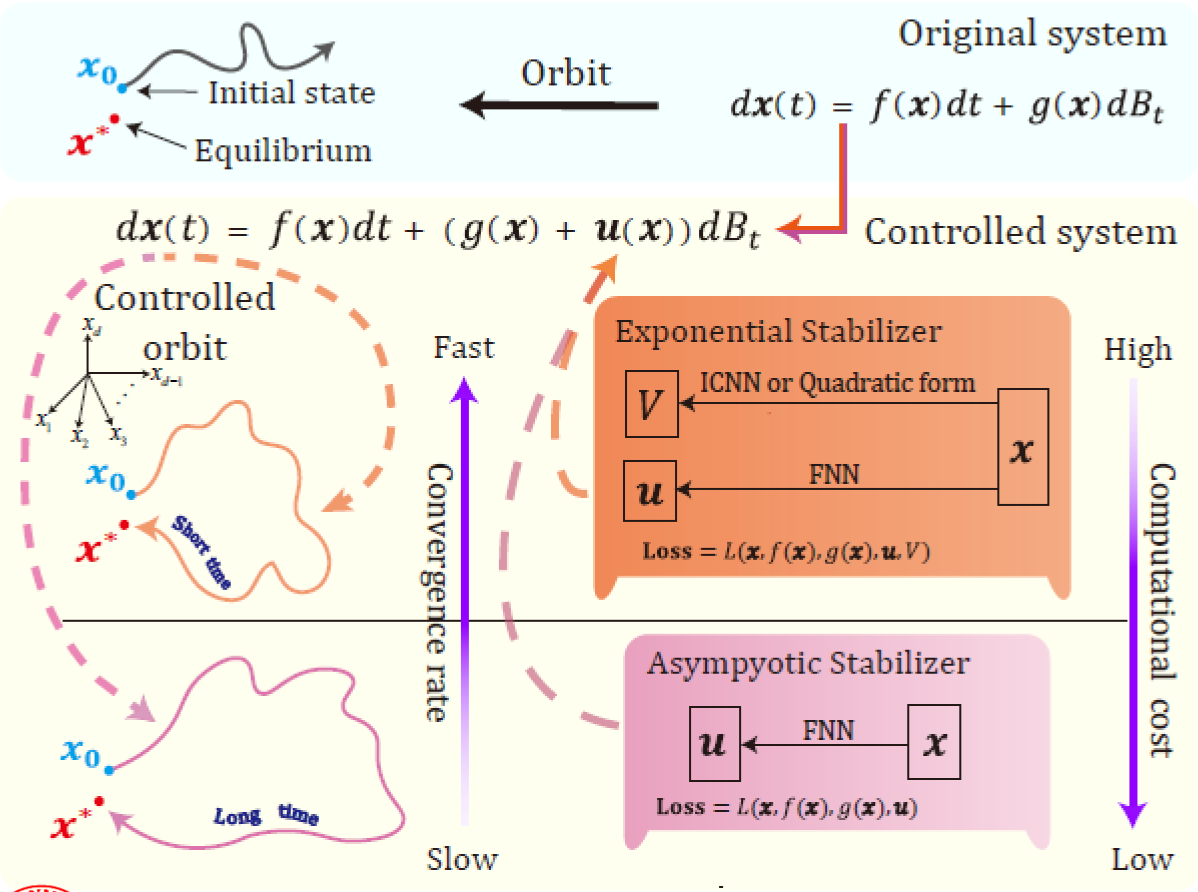
Asymptotic Stability

Theorem 2.3 *Appleby et al. (2008) Suppose that Assumptions 2.1 holds. Suppose further* $\min_{\|\boldsymbol{x}\|=M} \|\boldsymbol{x}^\top G(\boldsymbol{x})\| > 0$ *for any* $M > 0$ *and there exists a number* $\alpha \in (0,1)$ *such that*

$$\|\boldsymbol{x}\|^2(2\langle \boldsymbol{x}, F(\boldsymbol{x})\rangle + \|G(\boldsymbol{x})\|_{\mathrm{F}}^2) - (2-\alpha)\|\boldsymbol{x}^\top G(\boldsymbol{x})\|^2 \leq 0, \quad \forall \boldsymbol{x} \in \mathbb{R}^d. \tag{3}$$

Then, the unique and global solution of Eq. (1) satisfies $\lim_{t\to\infty} \boldsymbol{x}(t, \boldsymbol{x}_0) = \mathbf{0}$ *a.s., and we call this property as asymptotic attractiveness.*

Overall Workflow



- Add control u to diffusion
- (ES) Parameterize the auxiliary function V and control u
- (AS) Parameterize control u
- Physics-informed construction of V: ICNN & Quadratic form

ES:Construction of V

Input convex neural network (ICNN)

$$z_1 = \sigma_0(W_0 \boldsymbol{x} + b_0), \quad z_{i+1} = \sigma_i(U_i z_i + W_i \boldsymbol{x} + b_i),$$

$$g(\boldsymbol{x}) \equiv z_k, \quad i = 1, \cdots, k-1,$$

$$V(\boldsymbol{x}) = \sigma_{k+1}(g(\mathcal{F}(\boldsymbol{x})) - g(\mathcal{F}(\boldsymbol{0}))) + \varepsilon \|\boldsymbol{x}\|^2,$$

$$\sigma(x) = \begin{cases} 0, & \text{if } x \le 0, \\ (2dx^3 - x^4)/2d^3, & \text{if } 0 < x \le d, \\ x - d/2, & \text{otherwise} \end{cases}$$

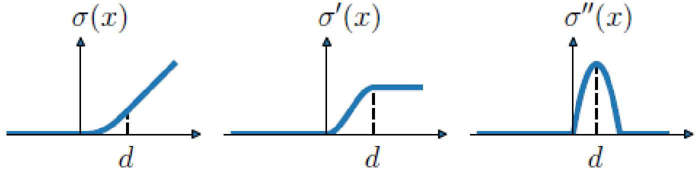


Figure 2: The smoothed **ReLU** $\sigma(\cdot)$.

Amos, B., Xu, L., and Kolter, J. Z. Input convex neural networks. In International Conference on Machine Learning, pp. 146–155. PMLR, 2017.

ES:Construction of V, u

Quadratic Form

$$V(\boldsymbol{x}) = \boldsymbol{x}^\top \left[\varepsilon I + V_{\boldsymbol{\theta}}(\boldsymbol{x})^\top V_{\boldsymbol{\theta}}(\boldsymbol{x})\right] \boldsymbol{x},$$

where V_θ is parameterized by some multilayered neural network (NN), $\varepsilon > 0$ is a hyperparameter.

Control function s.t. $\boldsymbol{u(0) = 0}$

$$u(x) = \mathrm{NN}(x) - \mathrm{NN}(0) \text{ or } u(x) = \mathrm{diag}(x)\mathrm{NN}(x)$$

Gallieri, M., Salehian, S. S. M., Toklu, N. E., Quaglino, A., Masci, J., Koutnik, J., and Gomez, F. Safe interactive model-based learning. arXiv preprint arXiv:1911.06556, 2019.

ES: Loss function

We can extract the following sufficient condition for exponential stability

$$\frac{(\nabla V(\boldsymbol{x})^\top g_{\boldsymbol{u}}(\boldsymbol{x}))^2}{V(\boldsymbol{x})^2} - b \cdot \frac{\mathcal{L}V(\boldsymbol{x})}{V(\boldsymbol{x})} \geq 0, \ \ b > 2, \ \boldsymbol{x} \neq 0.$$

The loss can be defined as

$$L_{N,b,\varepsilon}(\boldsymbol{\theta}, \boldsymbol{u}) \ = \frac{1}{N}\sum_{i=1}^{N} \max\left(0, \frac{b \cdot \mathcal{L}V(\boldsymbol{x}_i)}{V(\boldsymbol{x}_i)} - \frac{(\nabla V(\boldsymbol{x}_i)^\top g_{\boldsymbol{u}}(\boldsymbol{x}_i))^2}{V(\boldsymbol{x}_i)^2}\right)$$

AS: Loss function

Under the similar structure for controller , we define the loss function from the sufficient condition of asymptotic stability as

$$L_{N,\alpha}(\boldsymbol{u}) = \frac{1}{N}\sum_{i=1}^{N}\left[\max\left(0, (\alpha-2)\|\boldsymbol{x}_i^\top g_{\boldsymbol{u}}(\boldsymbol{x}_i)\|^2 + \|\boldsymbol{x}_i\|^2(\langle \boldsymbol{x}_i, f(\boldsymbol{x}_i)\rangle + \|g_{\boldsymbol{u}}(\boldsymbol{x}_i)\|_{\mathrm{F}}^2)\right)\right]$$

α is an adjustable parameter, which is related to the convergence time and the energy cost using the controller u

Convergence time and energy cost

We define the convergence time of the system under neural stochastic controller as the following stopping time

$$\tau_\epsilon \triangleq \inf\{t > 0 : \|\boldsymbol{x}(t)\| = \epsilon\}$$

ε is some predefined threshold value

The energy consumed in the control process until the stopping time is defined as

$$\mathcal{E}(\tau_\epsilon, T_\epsilon) \triangleq \mathbb{E}\left[\int_0^{\tau_\epsilon \wedge T_\epsilon} \|\boldsymbol{u}\|^2 \mathrm{d}t\right] = \mathbb{E}\left[\int_0^{T_\epsilon} \|\boldsymbol{u}\|^2 \mathbb{1}_{\{t<\tau_\epsilon\}} \mathrm{d}t\right]$$

Theoretical upper bounds

Theorem 4.2 *(Estimation for ES)* For ES stabilizer $\boldsymbol{u}(\boldsymbol{x})$ in (14) with $\langle \boldsymbol{x}, f(\boldsymbol{x})\rangle \leq L\|\boldsymbol{x}\|^2$, $\varepsilon < \|\boldsymbol{x}_0\|$, under the same notations and conditions in Theorem 2.2 with $c_3 - 2c_2 > 0$, we have

$$\begin{cases} \mathbb{E}[\tau_\epsilon] \leq T_\epsilon = \dfrac{2\log\left(V(\boldsymbol{x}_0)/c_1\varepsilon^p\right)}{c_3 - 2c_2}, \\ \mathcal{E}(\tau_\epsilon, T_\epsilon) \leq \dfrac{{k_{\boldsymbol{u}}}^2\|\boldsymbol{x}_0\|^2}{{k_{\boldsymbol{u}}}^2 + 2L}\left[\exp\left(\dfrac{2({k_{\boldsymbol{u}}}^2 + 2L)\log\left(V(\boldsymbol{x}_0)/c_1\varepsilon^p\right)}{c_3 - 2c_2}\right) - 1\right]. \end{cases}$$

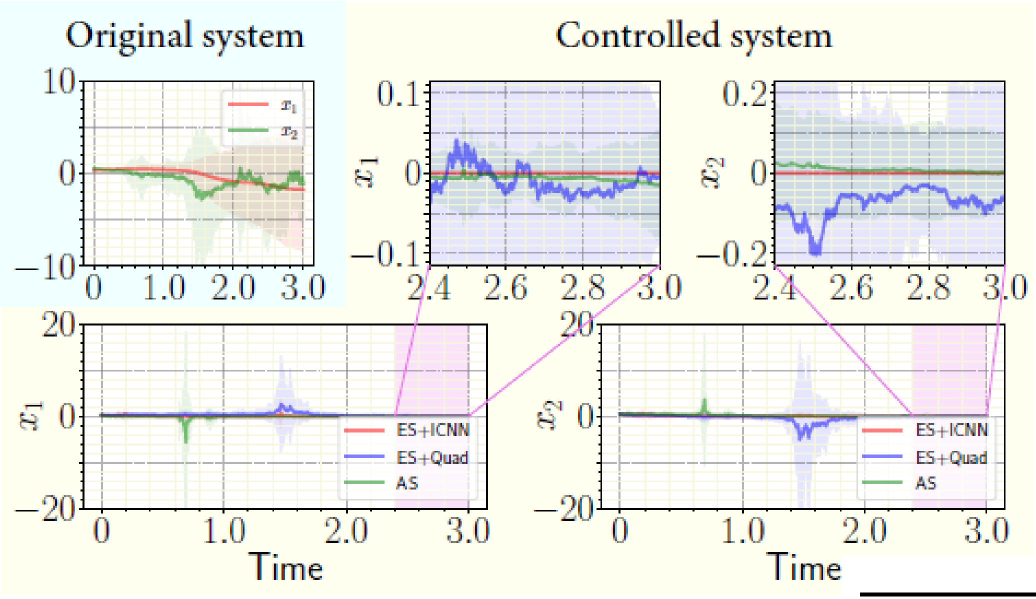
Theorem 4.3 *(Estimation for AS)* For (14) with $\langle \boldsymbol{x}, f(\boldsymbol{x})\rangle \leq L\|\boldsymbol{x}\|^2$, $\varepsilon < \|\boldsymbol{x}_0\|$, under the same notations and conditions in Theorem 2.3, if the left term in (3) further satisfies $\max_{\|\boldsymbol{x}\|\geq\varepsilon} \|\boldsymbol{x}\|^{\alpha-4}(\|\boldsymbol{x}\|^2(2\langle \boldsymbol{x}, f(\boldsymbol{x})\rangle + \|\boldsymbol{u}(\boldsymbol{x})\|_{\mathrm{F}}^2) - (2-\alpha)\|\boldsymbol{x}^\top \boldsymbol{u}(\boldsymbol{x})\|^2) = -\delta_\varepsilon < 0$, then for NN controller $\boldsymbol{u}(\boldsymbol{x})$ with Lipschitz constant $k_{\boldsymbol{u}}$, we have

$$\begin{cases} \mathbb{E}[\tau_\epsilon] \leq T_\epsilon = \dfrac{2\left(\|\boldsymbol{x}_0\|^\alpha - \varepsilon^\alpha\right)}{\delta_\varepsilon \cdot \alpha}, \\ \mathcal{E}(\tau_\epsilon, T_\epsilon) \leq \dfrac{{k_{\boldsymbol{u}}}^2\|\boldsymbol{x}_0\|^2}{{k_{\boldsymbol{u}}}^2 + 2L}\left[\exp\left(\dfrac{2({k_{\boldsymbol{u}}}^2 + 2L)\left(\|\boldsymbol{x}_0\|^\alpha - \varepsilon^\alpha\right)}{\delta_\varepsilon \cdot \alpha}\right) - 1\right]. \end{cases}$$

Hyperparameters $b = \frac{c_3}{c_2}$, α can be adjusted according to the estimation !

Results: Harmonic Linear Oscillator

Stabilize unstable zero solution in the original SDE with ES&AS



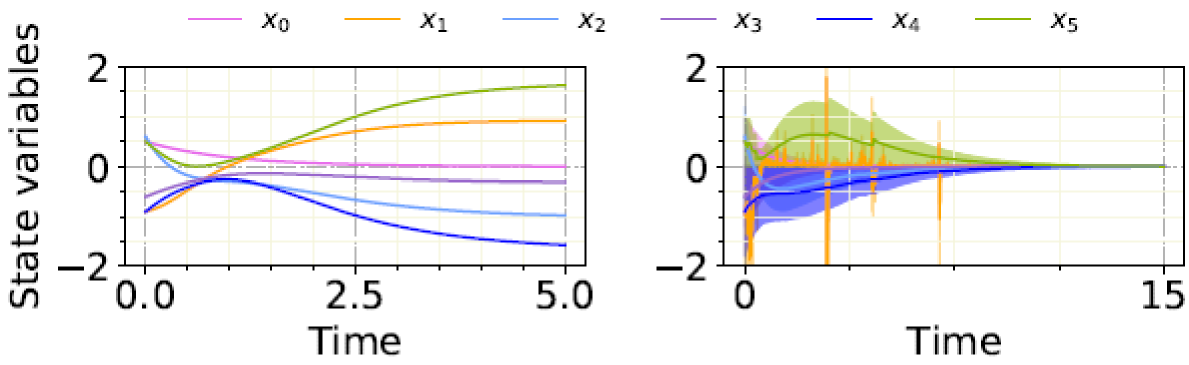
Training time

Convergence time

	Tt	Ni	Di	Ct
ES(+ICNN)	276.385s	121	1e-9	**0.459**
ES(+Quadratic)	78.071s	**107**	0.049	3.683
AS	**4.839s**	184	0.027	2.027

Results: Model free and pinning control

For 6-D Cell Fate Dynamics $\boldsymbol{x} = (x_1, \cdots, x_6)$, we combine the NODE method to learn the original dynamic from data and then find pinning control, i.e. we only control x_2 to stabilize the unstable state.



Thank you !