# Neural Stochastic Control

**Anonymous Author(s)**
Affiliation
Address
email

## Abstract

Control problems are always challenging since they arise from the real-world systems where stochasticity and randomness are of ubiquitous presence. This naturally and urgently calls for developing efficient neural control policies for stabilizing not only the deterministic equations but the stochastic systems as well. Here, in order to meet this paramount call, we propose two types of controllers, viz., the exponential stabilizer (ES) based on the stochastic Lyapunov theory and the asymptotic stabilizer (AS) based on the stochastic asymptotic stability theory. The ES can render the controlled systems exponentially convergent but it requires a long computational time; conversely, the AS makes the training much faster but it can only assure the asymptotic (not the exponential) attractiveness of the control targets. These two stochastic controllers thus are complementary in applications. We also investigate rigorously the linear control in both convergence time and energy cost and numerically compare it with the proposed controllers in these terms. More significantly, we use several representative physical systems to illustrate the usefulness of the proposed controllers in stabilization of dynamical systems.

## 1 Introduction

In the field of controlling dynamical systems, one of the major missions is to find efficient control policies for stabilizing ordinary differential equations (ODEs) to targeted equilibriums. The policies for stabilizing linear or polynomial dynamical systems have been fully developed using the standard Lyapunov stability theory, e.g., the linear quadratic regulator (LQR) Khalil (2002) and the sum-of-squares (SOS) polynomials through the semi-definite planning (SDP) Parrilo (2000). As for stabilizing more general and nonlinear dynamical systems, linearization technique around the targeted states is often utilized and thus the existing control policies are effective in the vicinity of the targeted states (Sastry & Isidori, 1989) but likely lose efficacy in the region far away from those states. Moreover, in real applications, the explicit forms of the controlled nonlinear systems are often partially or completely unknown, so it is very difficult to directly design controllers only using the Lyapunov stability theory. To overcome these difficulties, designing the controllers

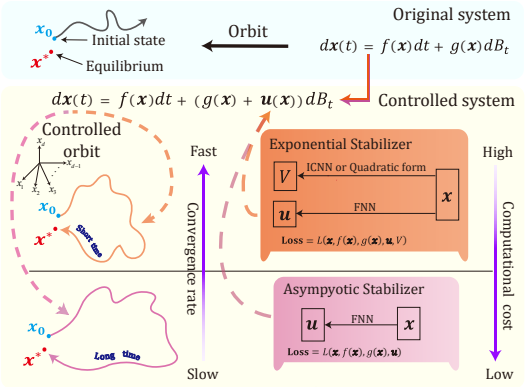


Figure 1: Sketches of the two frameworks of neural stochastic controller. Both the ES and AS find control function $u$ with fully connected feedforward NN (FNN).

Submitted to 36th Conference on Neural Information Processing Systems (NeurIPS 2022). Do not distribute.

via training neural networks (NNs) become one of the mainstream approaches in the community of cybernetics (Polycarpou, 1996). Recent outstanding developments using NNs include enlarging the safe region (Richards et al., 2018), learning the stable dynamics (Takeishi & Kawahara, 2021), and constructing the Lyapunov function and the control function simultaneously (Chang et al., 2019). In Kolter & Manek (2019) a projected NN has been constructed to directely learn a stable dynamical system and fits the observed time series data well, but it did not focus on learning a control policy to stabilize the original dynamics. All these existing developments are formulated only for deterministic systems but inapplicable directly to the dynamical systems described by stochastic differential equations (SDEs), requiring us to include the stochasticity appropriately into the use of neural controls to different types of dynamical systems.

The stability theory for stochastic systems has been systematically developed in the past several decades. Representative contributions in the literature include the Lyapunov-like stability theory for SDEs Mao (2007), stabilization of unstable states in ODEs only using noise perturbations Mao (1994b), and the stability induced by randomly switching structures Guo et al. (2018). Generally, for any SDEs governed by $\mathrm{d}\boldsymbol{x} = f(\boldsymbol{x})\mathrm{d}t + g(\boldsymbol{x})\mathrm{d}B_t$, control policies as $\boldsymbol{u} = (\boldsymbol{u}_f, \boldsymbol{u}_g)$ are introduced, which transforms the original equations into the controlled system $\mathrm{d}\boldsymbol{x} = [f(\boldsymbol{x})+\boldsymbol{u}_f(\boldsymbol{x})]\mathrm{d}t+[g(\boldsymbol{x})+\boldsymbol{u}_g(\boldsymbol{x})]\mathrm{d}B_t$. Appropriate forms of control policies are able to steer the controlled system to the equilibriums that are unstable in the original SDEs. Traditional control methods focus on designing deterministic control $\boldsymbol{u}_f$ and regard noise as negative part. Innovatively, we treat noise as a beneficial part and design stochastic control $\boldsymbol{u}_g$ to achieve the stabilization.

In this article, we articulate two frameworks of neural stochastic control which can complement each other in terms of convergence rate and the computational time of training NNs. Additionally, we analytically investigate the convergence time and the energy cost for the classic linear control and numerically compare them to the proposed neural stochastic control. The major contributions of this article are multi-folded, including:

- designing two frameworks of neural stochastic control, viz., the ES and the AS, and presenting their advantages in the stochastic control,
- computing the convergence time and the energy cost of particular stochastic neural control,
- demonstrating the efficacy of the proposed stochastic neural control in important control problems arising from representative physical systems, and we make our code available at `https://anonymous.4open.science/r/Neural-Stochastic-Control-DE3F/`.

### 1.1 Related Works

● **Lyapunov Method in Machine Learning** The recent work Chang et al. (2019) proposed an NN framework of learning the Lyapunov function and the linear control function simultaneously for stabilizing ODEs. In comparison, we select several specific types of NNs which have typical properties of the Lyapunov function. For instance, we use the input convex neural network (ICNN) Amos et al. (2017), constructing a positively definite convex function as a neural Lyapunov function (Kolter & Manek, 2019; Takeishi & Kawahara, 2021), and we construct the NN in a quadratic form (Richards et al., 2018; Gallieri et al., 2019) for linear or sublinear systems where the SDP method is often used to find the SOS-type Lyapunov function (Henrion & Garulli, 2005; Jarvis-Wloszek et al., 2003; Parrilo, 2000).

● **Stochastic Stability Theory of SDEs** Stochastic stability theory for SDEs have been systematically and fruitfully achieved in the past several decades (Kushner, 1967; Arnold, 2007; Mao, 1991, 1994a). The positive effects of stochasticity have also been cultivated in control fields (Mao et al., 2002; Deng et al., 2008; Caraballo et al., 2003; Appleby et al., 2008; Mao et al., 2007). These, therefore, motivate us to develop *only* neural stochastic control to stabilize different sorts of dynamical systems in this article. More stochastic stability theory for different kinds of systems are included in Appleby et al. (2006); Appleby (2003); Caraballo & Robinson (2004); Wang & Zhu (2017).

## 2 Preliminaries

To begin with, we consider the SDE which is written in a general form as:

$$\mathrm{d}\boldsymbol{x}(t) = F(\boldsymbol{x}(t))\mathrm{d}t + G(\boldsymbol{x}(t))\mathrm{d}B_t, \;\; t \geq 0, \; \boldsymbol{x}(0) = \boldsymbol{x}_0 \in \mathbb{R}^d, \tag{1}$$

where $F : \mathbb{R}^d \to \mathbb{R}^d$ is the drift function, and $G : \mathbb{R}^d \to \mathbb{R}^{d\times r}$ is the diffusion function with $\mathbb{R}^{d\times r}$, a space of $d \times r$ matrices with real entries, and $B_t \in \mathbb{R}^r$, a $r$-dimensional ($r$-D) Brownian motion. Without loss of generality, we set $F(\mathbf{0}) = \mathbf{0}$ and $G(\mathbf{0}) = \mathbf{0}$ so that $\boldsymbol{x}_0 = \mathbf{0}$ is a zero solution Eq. (1).

**Notations.** Denote by $\|\cdot\|$ the $L^2$-norm for any given vector in $\mathbb{R}^d$. Denote by $|\cdot|$ the absolute value of a scalar number or the modulus length of a complex number number. For $A = (a_{ij})$, a matrix of dimension $d \times r$, denote by $\|A\|_{\mathrm{F}}^2 = \sum_{i=1}^d \sum_{j=1}^r a_{ij}^2$ the Frobenius norm.

**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 $\|\boldsymbol{x}\| \vee \|\boldsymbol{y}\| \leq n$.*

**Definition 2.1** *(Derivative Operator) Define the differential operator $\mathcal{L}$ associated with Eq.* (1) *by*

$$\mathcal{L} \triangleq \sum_{i=1}^d F_i(\boldsymbol{x})\frac{\partial}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^d [G(\boldsymbol{x})G^\top(\boldsymbol{x})]_{ij}\frac{\partial^2}{\partial x_i \partial x_j}.$$

**Definition 2.2** *(Exponential Stability) The zero solution of Eq.* (1) *is said to be almost surely exponentially stable, if $\limsup_{t\to\infty} \frac{1}{t}\log\|\boldsymbol{x}(t;\boldsymbol{x}_0)\| < 0$ $a.s.$ for all $\boldsymbol{x}_0 \in \mathbb{R}^d$. Here and throughout, $a.s.$ stands for the abbreviation of almost surely.*

Then, the following Lyapunov stability theorem will be used in the establishment of our main results.

**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,*

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

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

The following asymptotic theorem also will be used in the establishment of our main results.

**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*

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

*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.*

## 3 Designing Stable Stochastic Controller

Here, we assume that the zero solution of the following SDE:

$$\mathrm{d}\boldsymbol{x} = f(\boldsymbol{x})\mathrm{d}t + g(\boldsymbol{x})\mathrm{d}B_t \tag{4}$$

is unstable. Note that, for any nontrivial targeted equilibrium $\boldsymbol{x}^*$, a direct transformation $\boldsymbol{y} = \boldsymbol{x} - \boldsymbol{x}^*$ can make the zero solution as the equilibrium of the transformed system. Thus, our mission is to stabilize the zero solution only. As such, we are to use the NNs to design the control $\boldsymbol{u} : \mathbb{R}^d \to \mathbb{R}^{d\times r}$ with $\boldsymbol{u}(\mathbf{0}) = \mathbf{0}$ and apply it to Eq. (4) as

$$\mathrm{d}\boldsymbol{x} = f(\boldsymbol{x})\mathrm{d}t + [g(\boldsymbol{x}) + \boldsymbol{u}(\boldsymbol{x})]\mathrm{d}B_t. \tag{5}$$

Since $\boldsymbol{u}$ is integrated with $\mathrm{d}B_t$ in the controlled system (5), we regard it as a stochastic controller. In what follows, two frameworks of neural stochastic control, the exponential stabilizer (ES) and the asymptotic stabilizer (AS), are articulated, respectively, in Sections 3.1 and 3.2. All these control policies are intuitively depicted in Figure 1.

### 3.1 Exponential Stabilizer

Once we find the Lyapunov function $V$ and the neural controller $\boldsymbol{u}$, making the controlled system (5) meet all the conditions assumed in Theorem 2.2, the equilibrium $\mathbf{0}$ can be exponentially stabilized. To this end, we first provide two different types of functions for constructing $V$, which actually could be complementary in applications. Then, we design the explicit forms of control function and loss function.

**ICNN $V$ Function.** We use the ICNN (Amos et al., 2017) to represent the candidate Lyapunov function $V$. This guarantees $V$ as a convex function with respect to the input $\boldsymbol{x}$. In order to further make $V$ as a true Lyapunov function, we use the following form:

$$\begin{aligned} \boldsymbol{z}_1 &= \sigma_0(W_0\boldsymbol{x} + b_0), \;\; \boldsymbol{z}_{i+1} = \sigma_i(U_i\boldsymbol{z}_i + W_i\boldsymbol{x} + b_i), \\ g(\boldsymbol{x}) &\equiv \boldsymbol{z}_k, \;\; i = 1, \cdots, k-1, \\ V(\boldsymbol{x}) &= \sigma_{k+1}(g(\mathcal{F}(\boldsymbol{x})) - g(\mathcal{F}(\mathbf{0}))) + \varepsilon\|\boldsymbol{x}\|^2, \end{aligned} \tag{6}$$

as introduced in Manek & Kolter (2020). Here, $W_i, b_i$ are real-valued weights, $U_i$ are positive weights, $\sigma_i$ are convex, monotonically non-decreasing activation functions in the $i$-th layer, $\varepsilon$ is a small positive constant, and $\mathcal{F}$ is a continuously differentiable and invertible function. In our framework, we require $V \in C^2(\mathbb{R}^d; \mathbb{R}_+)$ according to Definition 2.1; however, each activation function $\sigma_i \equiv \sigma$ in Manek & Kolter (2020) is $C^1$ only. Thus, we modify the original function as:

$$\sigma(x) = \begin{cases} 0, & \text{if } x \leq 0, \\ (2dx^3 - x^4)/2d^3, & \text{if } 0 < x \leq d, \\ x - d/2, & \text{otherwise} \end{cases} \tag{7}$$

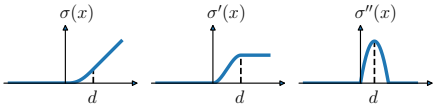


Figure 2: The smoothed **ReLU** $\sigma(\cdot)$.

which not only approximates the typical **ReLU** activation but also becomes continuously differentiable to the second order (see Figure 2).

**Quadratic $V$ Function.** For any $\boldsymbol{x} \in \mathbb{R}^d$, let $V_{\boldsymbol{\theta}} \in \mathbb{R}^d$ be a multilayered feedforward NN of the input $\boldsymbol{x}$ and with $\tanh(\cdot)$ as the activation functions, where $\boldsymbol{\theta}$ is the parameter vector. To meet the condition used in Definition 2.1, we cannot use the **ReLU**, a non-smooth function, as the activation function. Hence, we use the candidate Lyapunov function as:

$$V(\boldsymbol{x}) = \boldsymbol{x}^\top \left[\varepsilon I + V_{\boldsymbol{\theta}}(\boldsymbol{x})^\top V_{\boldsymbol{\theta}}(\boldsymbol{x})\right] \boldsymbol{x}, \tag{8}$$

which was introduced in Gallieri et al. (2019). Here, $\varepsilon$ is a small positive constant.

**Control Function.** We introduce a multi-layer feedforward NN (FNN), denoted by $\mathbf{NN}(\boldsymbol{x}) \in \mathbb{R}^r$, to design the controller $\boldsymbol{u}$. Since we require $\boldsymbol{u}(\mathbf{0}) = \mathbf{0}$, we set $\boldsymbol{u}(\boldsymbol{x}) \triangleq \mathbf{NN}(\boldsymbol{x}) - \mathbf{NN}(\mathbf{0})$ or $\boldsymbol{u}(\boldsymbol{x}) \triangleq \mathrm{diag}(\boldsymbol{x})\mathbf{NN}(\boldsymbol{x})$ with $r = d$. Here, $\mathrm{diag}(\boldsymbol{x})$ is a diagonal matrix with its $i$-th diagonal element as $x_i$.

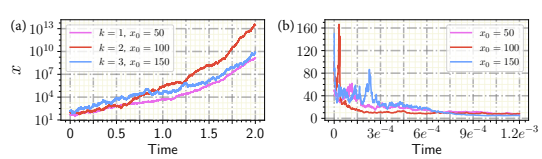


Figure 3: Sampled trajectories produced by Eq. (9) using, respectively, the stochastic linear control $u(x) = kx$ (a) and the nonlinear control $u(x) = 2x^2$ (b).

**Remark 3.1** *As reported in Chang et al. (2019), a single-layer NN without the bias constants in its arguments, which degenerates as linear control, could sufficiently take effect in the stabilization of many deterministic systems. However, this is NOT always the case for achieving the stabilization of highly nonlinear systems or even SDEs. The following proposition with Figure 3 provides an example, where neither the classic linear controller nor the stochastic linear controller can stabilize the unstable equilibrium in a particular SDE. The proof of this proposition is included in Appendix A.1.*

**Proposition 3.2** *Consider the following* 1*-D SDE:*

$$\mathrm{d}x(t) = x(t)\log|x(t)|\mathrm{d}t + u(x(t))\mathrm{d}B_t, \tag{9}$$

**Algorithm 1:** Exponential Stabilizer

**Input:** Data $\{\boldsymbol{x}_i\}_{i=1}^n$ sampled from $\mu(\Omega)$, iteration step $m$, learning rate $\gamma$, training error $\delta$, coefficient functions $f$ and $g$, initial parameters $\boldsymbol{\theta}_0$, and parameters $\varepsilon$ used in (6) and $b$ used in (11).
**Output:** Controller $\boldsymbol{u}(\boldsymbol{x}_i)$ and Lyapunov function $V(\boldsymbol{x}_i)$ in the form of (6) or (8).
**for** $r = 0$ **to** $m - 1$ **do**
  Compute $\nabla V(\boldsymbol{x}_i), HV(\boldsymbol{x}_i),\ i = 1, \cdots, n$
  Compute ES loss: $L(\boldsymbol{\theta}_r, \boldsymbol{u}_r)$ from (12)
  $\boldsymbol{u}_{r+1} = \boldsymbol{u}_r - \gamma \cdot \nabla_{\boldsymbol{u}} L(\boldsymbol{\theta}_r,\ \boldsymbol{u}_r),\ \boldsymbol{\theta}_{r+1} = \boldsymbol{\theta}_r - \gamma \cdot \nabla_{\boldsymbol{\theta}} L(\boldsymbol{\theta}_r,\ \boldsymbol{u}_r)$ ▷ Update parameters
  **if** $L(\boldsymbol{\theta}_{r+1}, \boldsymbol{u}_{r+1}) \leq \delta$ **then**
    break

*with a zero solution $x^* = 0$. Then, for $u(x) = kx$ with any $k$ and $x_0 \neq 0$, $x^* = 0$ is neither exponentially stable nor of globally asymptotic attractiveness almost surely. For $u(x) = 2x^2$, $x^* = 0$ is of globally asymptotic attractiveness. For $u(x) \equiv 0$, the deterministic system cannot be stabilized by any classic linear controller.*

**Loss Function.** When the learning procedure updates the parameters in the NNs such that the constructed $V$ and $\boldsymbol{u}$ with the coefficient functions, $f$ and $g_{\boldsymbol{u}} \triangleq g + \boldsymbol{u}$, in the controlled system (5) meet all the conditions assumed in Theorem 2.2, the exponential stability of the controlled system is assured. Thus, we demand a suitable loss function to evaluate the likelihood that those conditions are satisfied. First, from Theorem 2.2, it follows that $V(\boldsymbol{x}) \geq \varepsilon \|\boldsymbol{x}\|^2$ for all $\boldsymbol{x} \in \mathbb{R}^d$. Thus, Conditions (ii)-(iii) together with $c_3 - 2c_2 > 0$ in Theorem 2.2 equivalently become

$$\inf_{\boldsymbol{x} \neq 0} \frac{(\nabla V(\boldsymbol{x})^\top g_{\boldsymbol{u}}(\boldsymbol{x}))^2}{V(\boldsymbol{x})^2} \geq b \cdot \sup_{\boldsymbol{x} \neq 0} \frac{\mathcal{L}V(\boldsymbol{x})}{V(\boldsymbol{x})}, \quad b > 2. \tag{10}$$

These conditions further imply that

$$\frac{(\nabla V(\boldsymbol{x})^\top g_{\boldsymbol{u}}(\boldsymbol{x}))^2}{V(\boldsymbol{x})^2} - b \cdot \frac{\mathcal{L}V(\boldsymbol{x})}{V(\boldsymbol{x})} \geq 0, \quad b > 2,\ \boldsymbol{x} \neq 0. \tag{11}$$

With these reduced conditions, we design the ES loss function for the controlled system (5) as follows.

**Definition 3.1** *(*ES loss) Consider a candidate Lyapunov function $V$ and a controller $\boldsymbol{u}$ for the controlled system (5). Then, the ES loss is defined as

$$L_{\mu,b,\varepsilon}(\boldsymbol{\theta}, \boldsymbol{u}) = \mathbb{E}_{\boldsymbol{x} \sim \mu}\big[\max\big(0, \frac{b \cdot \mathcal{L}V(\boldsymbol{x})}{V(\boldsymbol{x})} - \frac{(\nabla V(\boldsymbol{x})^\top g_{\boldsymbol{u}}(\boldsymbol{x}))^2}{V(\boldsymbol{x})^2}\big)\big],$$

where the state variable $\boldsymbol{x}$ obeys the distribution $\mu$. In practice, we consider the following empirical loss function:

$$L_{N,b,\varepsilon}(\boldsymbol{\theta}, \boldsymbol{u}) \ = \frac{1}{N} \sum_{i=1}^{N} \max\big(0, \frac{b \cdot \mathcal{L}V(\boldsymbol{x}_i)}{V(\boldsymbol{x}_i)} - \frac{(\nabla V(\boldsymbol{x}_i)^\top g_{\boldsymbol{u}}(\boldsymbol{x}_i))^2}{V(\boldsymbol{x}_i)^2}\big), \tag{12}$$

where $\{\boldsymbol{x}_i\}_{i=1}^N$ are sampled from the distribution $\mu = \mu(\Omega)$ and $\Omega$ is some closed domain in $\mathbb{R}^d$.

For convenience, we summarize the developed framework in Algorithm 1. Here, $b$ is a hyperparameter that can be adjusted as required by solving a specific problem.

**Remark 3.3** *In Section 5, we show numerically that the conditions reduced in* (11) *are sufficiently effective for designing the ES loss. Actually, it is not necessary to design the loss function using the conditions in* (10)*.*

Now, for controlling any nonlinear ODEs or SDEs, we design the ES according to Algorithm 1. As such, using the ES framework can not only stabilize those unstable equilibriums (constant states)

**Algorithm 2:** Asymptotic Stabilizer

**Input:** Data $\{\boldsymbol{x}_i\}_{i=1}^n$ sampled from $\mu(\Omega)$, parameter $\alpha \in (0,1)$ used in (13), and all other parameters, $m$, $\gamma$, $\delta$, $f$, and $g$, defined in the same manner as those in Algorithm 1.
**Output:** Controller $\boldsymbol{u}(\boldsymbol{x}_i)$.
**for** $r = 0$ **to** $m-1$ **do**
  Compute loss function: $L(\boldsymbol{u}_r)$ from (13)
  $\boldsymbol{u}_{r+1} = \boldsymbol{u}_r - \gamma \cdot \nabla_{\boldsymbol{u}} L(\boldsymbol{u}_r)$ ▷ Update parameters
  **if** $L(\boldsymbol{u}_{r+1}) \leq \delta$ **then**
    break

of the given systems, but also can stabilize those unstable oscillators, e.g., the limit cycle. This is because the solution corresponding to the oscillator can be regarded as a zero solution of the controlled system after appropriate transformations are implemented.

Another point needs attention. During the construction of $V$ in (6), $\varepsilon\|\boldsymbol{x}\|^2$, the $L^2$-regularization, is used to guarantee the positive definiteness of $V$. However, often in the application of the Lyapunov stability theory, the form of $\|\boldsymbol{x}\|^2$ is not always a suitable candidate for the Lyapunov function. It may restrict the generalizability of using our framework, so it needs necessary adjustments. The following example illustrates this point.

**Example 3.4** *Consider a 2-D SDE as follows:*

$$\begin{cases} \mathrm{d}x_1(t) = x_2(t)\mathrm{d}t, \\ \mathrm{d}x_2(t) = [-2x_1(t) - x_2(t)]\mathrm{d}t + x_1(t)\mathrm{d}B_t. \end{cases}$$

*In Appendix A.2, the zero solution of this system is validated to be exponentially stable almost surely; however, $k\|\boldsymbol{x}\|^2$ for any $k \in \mathbb{R}$ cannot be a useful auxiliary function to identify the exact stability of the zero solution.*

To be candid, using the current framework takes a longer time for training and constructing the neural Lyapunov function. In the next subsection, we thus establish an alternative control framework that can reduce the training time.

### 3.2 Asymptotic Stabilizer

Here, in light of Theorem 2.3, we are to establish the second framework, the AS, for stabilizing the unstable equilibrium of system (5). This framework only makes the equilibrium asymptotically attractive almost surely. Its control function is designed in the same way as the one used in the ES framework, whereas the loss function is differently designed.

**Definition 3.2** *(*AS loss) Utilization of the notations used in Definition 3.1, the loss function for the controlled system (5) with the controller $\boldsymbol{u}$ is defined as:

$$L_{\mu,\alpha}(\boldsymbol{u}) = \mathbb{E}_{\boldsymbol{x}\sim\mu}\big[\max\big(0, (\alpha-2)\|\boldsymbol{x}^\top g_{\boldsymbol{u}}(\boldsymbol{x})\|^2 + \|\boldsymbol{x}\|^2(\langle \boldsymbol{x}, f(\boldsymbol{x})\rangle + \|g_{\boldsymbol{u}}(\boldsymbol{x})\|_{\mathrm{F}}^2)\big)\big].$$

Akin to Definition 3.1, we set the empirical loss function as:

$$L_{N,\alpha}(\boldsymbol{u}) = \frac{1}{N}\sum_{i=1}^{N}\big[\max\big(0, (\alpha-2)\|\boldsymbol{x}_i^\top g_{\boldsymbol{u}}(\boldsymbol{x}_i)\|^2 + \|\boldsymbol{x}_i\|^2(\langle \boldsymbol{x}_i, f(\boldsymbol{x}_i)\rangle + \|g_{\boldsymbol{u}}(\boldsymbol{x}_i)\|_{\mathrm{F}}^2)\big)\big]. \quad (13)$$

Here, $\alpha$ is an adjustable parameter, which is related to the convergence time and the energy cost using the controller $\boldsymbol{u}$. We show in Appendix A.6 the influence of selecting different $\alpha$. For convenience, we summarize the AS framework in Algorithm 2.

## 4 Convergence Time and Energy Cost

The convergence time and the energy cost are the crucial factors to measure the quality of a controller (Yan et al., 2012; Li et al., 2017; Sun et al., 2017). In this section, we provide a comparative study

between the traditional stochastic linear control and the AS, the above-articulated neural stochastic control.

To this end, we first present a theorem on the estimations of the convergence time and the energy cost for the stochastic linear control on a general SDE.

**Theorem 4.1** *Consider the SDE with a stochastic linear controller as:*

$$\mathrm{d}\boldsymbol{x} = f(\boldsymbol{x})\mathrm{d}t + u(\boldsymbol{x})\mathrm{d}B_t, \quad \boldsymbol{x}(0) = \boldsymbol{x}_0 \in \mathbb{R}^d, \tag{14}$$

*where* $\langle \boldsymbol{x}, f(\boldsymbol{x})\rangle \leq L\|\boldsymbol{x}\|^2$ *and* $u(\boldsymbol{x}) = k\boldsymbol{x}$ *with* $|k| > \sqrt{2L}$*. Then, for* $\epsilon < \|\boldsymbol{x}_0\|$*, we have*

$$\begin{cases} \mathbb{E}[\tau_\epsilon] \leq T_\epsilon = \dfrac{2\log\left(\|\boldsymbol{x}_0\|/\epsilon\right)}{k^2 - 2L}, \\ \mathcal{E}(\tau_\epsilon, T_\epsilon) \leq \dfrac{k^2\|\boldsymbol{x}_0\|^2}{k^2 + 2L}\left[\exp\left(\dfrac{2(k^2+2L)\log\left(\|\boldsymbol{x}_0\|/\epsilon\right)}{k^2 - 2L}\right) - 1\right], \end{cases}$$

*where, for a sufficiently small* $\epsilon > 0$*, we denote the stopping time by* $\tau_\epsilon \triangleq \inf\{t > 0 : \|\boldsymbol{x}(t)\| = \epsilon\}$ *and denote the energy cost by*

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

The proof of this theorem is provided in Appendix A.3.

Now, we numerically compare the performances of the linear controller $u(x) = kx$ and the AS on the convergence time and the energy cost of the controls applied to system (14) with specific configurations (see Figure 4). We numerically find that $u(x) = kx$ can efficiently stabilize the equilibrium for $k > k_\mathrm{c} = 5.6$. Without loss of generality, we fix $k = 6.0$, and compare the corresponding performances. As clearly shown in Figure 4, the AS outperforms $u(x) = kx$ from both perspectives, the convergence speed and the energy cost. In the simulations, the energy cost $\mathcal{E}(\tau_\epsilon, T_\epsilon)$ defined above is computed in a finite-time duration as $\mathcal{E}(\tau_\epsilon, T \wedge T_\epsilon)$, where $T < \infty$ is selected to be appropriately large. We leave more results of the comparison study in Appendix A.4.5.

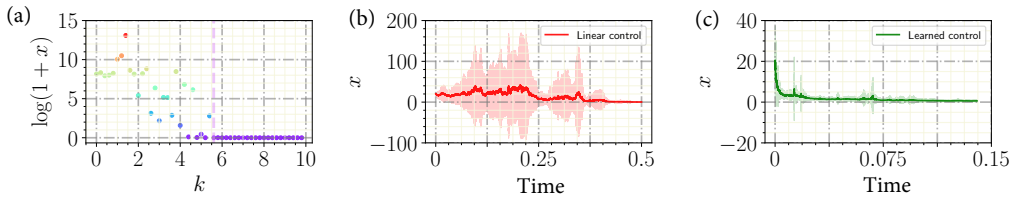


Figure 4: The performances of system (14) with specific configurations: $f(x) = x\log(1+x)$ and $u(x) = kx$ with $d = 1$, $k = 0.2 \cdot j$, and $j = 1, \cdots, 50$. (a) The convergence value of $\log(1+x)$ changes with $k$. The control becomes successful for $k > k_c$ ($k_\mathrm{c} = 5.6$ is indicated by the vertical dashed line). (b) The sampled trajectories using the linear controller with $k = 6.0$, where the energy cost, as $\epsilon = 0.1$, is computed as $38,388$. (c) The sampled trajectories using the AS, where the energy cost, as $\epsilon = 0.1$, is computed as $1,438$ only. All the simulating results, including the energy costs, are obtained through averaging the quantities produced by 20 randomly-sampled trajectories on the time duration $[0,\ T]$ with $T = 1$.

## 5 Experiments

In this section, we demonstrate the efficacy of the above-articulated frameworks of stochastic neural control, the ES and the AS, on several representative physical systems. We also compare these two frameworks, highlighting their advantages and weaknesses. The detailed configurations for these experiments are included in Appendix A.4. Additional illustrative experiments are included in Appendix A.5.

### 5.1 Harmonic Linear Oscillator

First, consider the harmonic linear oscillator $\ddot{y} + 2\beta\dot{y} + w^2 y = 0$, where $w$ is the natural frequency and $\beta > 0$ is the damping coefficient representing the strength of the external force on the vibrator (Dekker, 1981). Although this system is exponentially stable, the system with stochastic perturbations $\ddot{y} + (2\beta + \xi_2)\dot{y} + (w^2 + \xi_1)y = 0$ becomes unstable even if $\mathbb{E}[\xi_1(t)] = \mathbb{E}[\xi_2(t)] = 0$ (Arnold et al., 1983). Now, we apply the nonlinear ES(+ICNN), the linear ES(+Quadratic), and the nonlinear AS, respectively, to stabilizing this unstable dynamics (A.4.1) with $w^2 = 1$, $\beta = 0.5$, $\zeta_1 = -3$, and $\zeta_2 = 2.15$, the results are shown in Figure 5. Indeed, we find that the two nonlinear stochastic neural controls are more robust than the linear control, and that the ES(+ICNN), rather than the AS, makes the controlled system more stable.

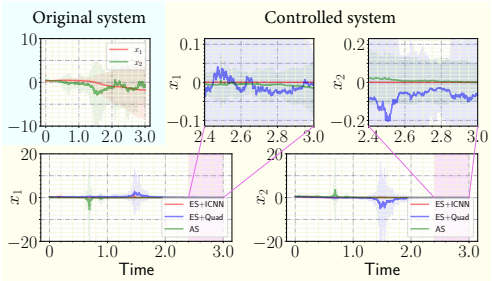


Figure 5: The solid lines are obtained through averaging the 20 sampled trajectories, while the shaded areas stand for the variance regions.

In Table 1 we compute the training time (Tt) used for the loss function converging to 0, the number of iterations (Ni), the distance (Di) between the trajectory and the targeted equilibrium at time $T = 4$, and the convergence time (Ct) when the distance between the trajectory and the targeted equilibrium is less than 0.05. The results are obtained through averaging the corresponding quantities produced by 20 randomly-sampled trajectories and the detailed training configurations are shown in Appendix A.4.1.

Table 1: Performance on Harmonic Linear Oscillator

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

### 5.2 Inverted Pendulum

Controlling the inverted pendulum is one of the classic nonlinear control problems (Anderson, 1989; Huang & Huang, 2000), which is governed by $ml^2\ddot{\theta} = mgl\sin\theta - \beta\dot{\theta}$. Here, $m$ is the mass of the pendulum, $g$ is the gravitational acceleration, $l$ is the pole length, and $\beta$ is the friction coefficient. Mathematically, the pendulum can be written as a system with two state variables: $\theta$, the angle deviating from the vertical position $x = 0$, and $\dot{\theta}$, the angular velocity. Denote the 2-D state variable by $\boldsymbol{x} = (\theta, \dot{\theta})$. Then, we apply the AS articulated in Algorithm 2 to steer the system to the equilibrium $\boldsymbol{x} = \mathbf{0}$ (sustaining the pendulum inverted). The schematic diagram of the inverted pendulum system and successful control results are shown in Figure 6. The sampled trajectories are initiated from different angles and angular velocities [i.e., $(\theta(0), \dot{\theta}(0)) = (-5,\ 5),\ (-1,\ 3),\ (3,\ -4)$].

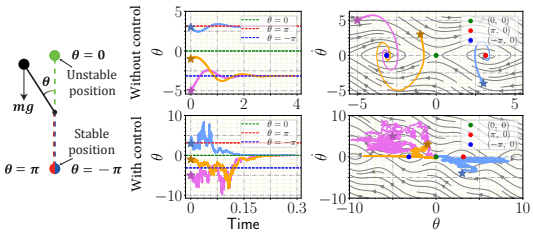


Figure 6: A schematic diagram of the inverted pendulum system (left), the sampled trajectories of the angle $\theta$ (middle), and the corresponding orbits in the phase space (right).

### 5.3 Stuart-Landau Equations

Finally, we show that our frameworks are beneficial to realizing the control and the synchronization of complex networks. To this end, we consider the single StuartLandau oscillator which is governed by the following complex-valued ODE:

$$\dot{Z} = (\beta + \mathrm{i}\gamma + \mu|Z|^2)Z, \quad Z \in \mathbb{C}. \tag{15}$$

This equation is a paradigmatic model undergoing the so-called Andronov bifurcation (Kuznetsov, 2013): Stability of the equilibrium changes and the limit cycle emerges as the parameter passes some critical value. In what follows, we consider two cases based on system (15).

● **Case 1** We set $\beta = -25$, $\gamma = 1$, and $\mu = 1$, so that system (15) has a stable equilibrium $\rho = 0$ and an unstable limit cycle $\rho = 5$. Here, $Z = x + \mathrm{i}y = \rho e^{\mathrm{i}\theta}$. Now, the AS steers the dynamics to the unstable limit cycle, as successfully shown in Figure 7. The trajectories (the left column) and the phase orbits (the right column) of system (15), initiated from 30 randomly-selected initial states, without control (the upper panels) and with control of the AS (the lower panels). The initial values inside (resp., outside) the limit cycle $\rho = 5$ are indicated by the blue (resp., purple) pentagrams.

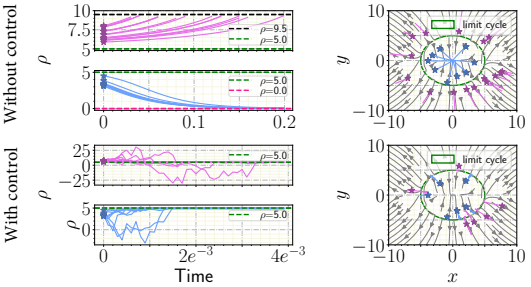


Figure 7: The trajectories (the left column) and the phase orbits (the right column) of system (15).

● **Case 2** Next, we consider the synchronization problem of the coupled Stuart-Landau equations. Successful deterministic methods have been systematically developed for realizing synchronization, including the adaptive control with time delay Selivanov et al. (2012) and the open-loop temporal network controller Zhang & Strogatz (2021). These methods majorly depend on the technique of linearization in the vicinity of the synchronization manifold. Here, we show how to apply our framework to achieving the synchronization in the coupled system. We set the corresponding Laplace matrix $L = (L_{jk})_{n\times n}$ as $\sum_{k=1}^{n} L_{jk} = 0$, which guarantees the synchronization manifold is an invariant manifold of the coupled system (Pecora & Carroll, 1998). Specifically, we select as $n = 20$, $\sigma = 0.01$, $c_1 = -1.8$, $c_2 = 4$, and $L_{jk} = \delta_{jk} - \frac{1}{n}$, where $\delta_{jk}$ is the Kronecker function. Then, we apply the AS to this system and realize the stabilization of the synchronous manifold, as shown in Figure 8.

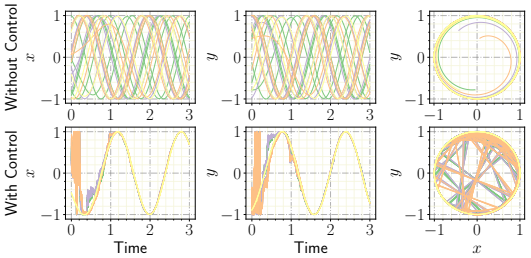


Figure 8: The dynamics of the first (resp., second) component of the coupled oscillators are shown in the panels in the left (resp., middle) column. The dynamics in the phase space (the right column).

In addition to the above controlled systems, we include the other illustrative examples, the controlled reservoir computing and the controlled Lorenz system, in Appendix A.5.

## 6 Conclusion and Future Works

In this article, we have proposed two frameworks of neural stochastic control for stabilizing different types of dynamical systems, including the SDEs. We have shown that the neural stochastic control outperforms the classic stochastic linear control in both the convergence time and the energy cost for typical systems. More importantly, using several representative physical systems, we have demonstrated the advantages of our frameworks and showed part of their weaknesses possibly emergent in real applications. Also, we present some limitations of the proposed frameworks in Appendix A.7.

Moreover, we suggest several directions for further investigations: (i) acceleration of the training process of the ES, (ii) the basin stability of the neural stochastic control (Menck et al., 2013), (iii) the trade-off between the deterministic controller $\boldsymbol{u}_g$ using the NNs and the stochastic controller $\boldsymbol{u}_f$ using the NNs, (iv) the safe learning in Robotic control with small disturbances (Berkenkamp et al., 2017), and (v) the design of the purely data-driven stochastic neural control.

## References


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

Anderson, C. W. Learning to control an inverted pendulum using neural networks. *IEEE Control Systems Magazine*, 9(3):31–37, 1989.

Appleby, J. A. Stabilisation of functional differential equations by noise. *Systems & Control Letters*, 2003.

Appleby, J. A., Mao, X., and Rodkina, A. On stochastic stabilization of difference equations. *Discrete & Continuous Dynamical Systems*, 15(3):843, 2006.

Appleby, J. A., Mao, X., and Rodkina, A. Stabilization and destabilization of nonlinear differential equations by noise. *IEEE Transactions on Automatic Control*, 53(3):683–691, 2008.

Arnold, L. *Stochastic Differential Equations: Theory and Applications*. Stochastic Differential Equations: Theory and Applications, 2007.

Arnold, L., Crauel, H., and Wihstutz, V. Stabilization of linear systems by noise. *SIAM Journal on Control and Optimization*, 21(3):451–461, 1983.

Berkenkamp, F., Turchetta, M., Schoellig, A. P., and Krause, A. Safe model-based reinforcement learning with stability guarantees. *arXiv preprint arXiv:1705.08551*, 2017.

Caraballo, T. and Robinson, J. C. Stabilisation of linear pdes by stratonovich noise. *Systems & Control Letters*, 53(1):41–50, 2004.

Caraballo, T., Garrido-Atienza, M. J., and Real, J. Stochastic stabilization of differential systems with general decay rate. *Systems & Control Letters*, 48(5):397–406, 2003.

Chang, Y.-C., Roohi, N., and Gao, S. Neural lyapunov control. In *Proceedings of the 33rd International Conference on Neural Information Processing Systems*, pp. 3245–3254, 2019.

Dekker, H. Classical and quantum mechanics of the damped harmonic oscillator. *Physics Reports*, 80(1):1–110, 1981.

Deng, F., Luo, Q., Mao, X., and Pang, S. Noise suppresses or expresses exponential growth. *Systems & Control Letters*, 57(3):262–270, 2008.

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

Guo, Y., Lin, W., and Chen, G. Stability of switched systems on randomly switching durations with random interaction matrices. *IEEE Transactions on Automatic Control*, 63(1):21–36, 2018.

Henrion, D. and Garulli, A. *Positive Polynomials in Control*, volume 312. Springer Science & Business Media, 2005.

Huang, S.-J. and Huang, C.-L. Control of an inverted pendulum using grey prediction model. *IEEE Transactions on Industry Applications*, 36(2):452–458, 2000.

Jaeger, H. The echo state approach to analysing and training recurrent neural networks-with an erratum note. *Bonn, Germany: German National Research Center for Information Technology GMD Technical Report*, 148(34):13, 2001.

Jaeger, H. Echo state network. *Scholarpedia*, 2(9):2330, 2007.

Jaeger, H. and Haas, H. Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication. *Science*, 304(5667):78–80, 2004.

Jarvis-Wloszek, Z., Feeley, R., Tan, W., Sun, K., and Packard, A. Some controls applications of sum of squares programming. In *42nd IEEE International Conference on Decision and Control (IEEE Cat. No. 03CH37475)*, volume 5, pp. 4676–4681. IEEE, 2003.

Khalil, H. K. Nonlinear systems third edition. *Patience Hall*, 115, 2002.

Kolter, J. Z. and Manek, G. Learning stable deep dynamics models. *Advances in Neural Information Processing Systems*, 32:11128–11136, 2019.

Kushner, H. J. Stochastic stability and control. Technical report, Brown Univ Providence RI, 1967.

Kuznetsov, Y. A. *Elements of applied bifurcation theory*, volume 112. Springer Science & Business Media, 2013.

Li, A., Cornelius, S. P., Liu, Y.-Y., Wang, L., and Barabási, A.-L. The fundamental advantages of temporal networks. *Science*, 358(6366):1042–1046, 2017.

Lukoševičius, M. and Jaeger, H. Reservoir computing approaches to recurrent neural network training. *Computer Science Review*, 3(3):127–149, 2009.

Ma, H., Ho, D. W., Lai, Y.-C., and Lin, W. Detection meeting control: Unstable steady states in high-dimensional nonlinear dynamical systems. *Physical Review E*, 92(4):042902, 2015.

Manek, G. and Kolter, J. Z. Learning stable deep dynamics models. *arXiv preprint arXiv:2001.06116*, 2020.

Mao, X. *Stability of Stochastic Differential Equations with Respect to Semimartingales*. Longman, 1991.

Mao, X. *Exponential Stability of Stochastic Differential Equations*. Marcel Dekker, 1994a.

Mao, X. Stochastic stabilization and destabilization. *Systems & Control Letters*, 23(4):279–290, 1994b.

Mao, X. *Stochastic differential equations and applications*. Elsevier, 2007.

Mao, X., Marion, G., and Renshaw, E. Environmental brownian noise suppresses explosions in population dynamics. *Stochastic Processes and Their Applications*, 97(1):95–110, 2002.

Mao, X., Yin, G. G., and Yuan, C. Stabilization and destabilization of hybrid systems of stochastic differential equations. *Automatica*, 43(2):264–273, 2007.

Menck, P. J., Heitzig, J., Marwan, N., and Kurths, J. How basin stability complements the linear-stability paradigm. *Nature Physics*, 9(2):89–92, 2013.

Parrilo, P. A. *Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization*. California Institute of Technology, 2000.

Pecora, L. M. and Carroll, T. L. Master stability functions for synchronized coupled systems. *Physical Review Letters*, 80(10):2109, 1998.

Polycarpou, M. Stable adaptive neural control scheme for nonlinear systems. *IEEE Transactions on Automatic Control*, 41(3):447–451, 1996. doi: 10.1109/9.486648.

Richards, S. M., Berkenkamp, F., and Krause, A. The lyapunov neural network: Adaptive stability certification for safe learning of dynamical systems. In *Conference on Robot Learning*, pp. 466–476. PMLR, 2018.

Sastry, S. and Isidori, A. Adaptive control of linearizable systems. *IEEE Transactions on Automatic Control*, 34(11):1123–1131, 1989. doi: 10.1109/9.40741.

Selivanov, A. A., Lehnert, J., Dahms, T., Hövel, P., Fradkov, A. L., and Schöll, E. Adaptive synchronization in delay-coupled networks of stuart-landau oscillators. *Physical Review E*, 85(1):016201, 2012.

Sparrow, C. *The Lorenz Equations: Bifurcations, Chaos, and Strange Attractors*, volume 41. Springer Science & Business Media, 2012.

Sun, Y.-Z., Leng, S.-Y., Lai, Y.-C., Grebogi, C., and Lin, W. Closed-loop control of complex networks: A trade-off between time and energy. *Physical Review Letters*, 119(19):198301, 2017.

Takeishi, N. and Kawahara, Y. Learning dynamics models with stable invariant sets. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pp. 9782–9790, 2021.

Wang, B. and Zhu, Q. Stability analysis of markov switched stochastic differential equations with both stable and unstable subsystems. *Systems & Control Letters*, 105:55–61, 2017.

Yan, G., Ren, J., Lai, Y.-C., Lai, C.-H., and Li, B. Controlling complex networks: How much energy is needed? *Physical Review Letters*, 108(21):218703, 2012.

Yang, S.-K., Chen, C.-L., and Yau, H.-T. Control of chaos in lorenz system. *Chaos, Solitons & Fractals*, 13(4):767–780, 2002.

Yang, T., Yang, L.-B., and Yang, C.-M. Impulsive control of lorenz system. *Physica D: Nonlinear Phenomena*, 110(1-2):18–24, 1997.

Zhang, Y. and Strogatz, S. H. Designing temporal networks that synchronize under resource constraints. *Nature Communications*, 12(1):1–8, 2021.

Zhu, Q., Ma, H., and Lin, W. Detecting unstable periodic orbits based only on time series: When adaptive delayed feedback control meets reservoir computing. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 29(9):093125, 2019.

## Checklist

1. For all authors...
   (a) Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope? [Yes]
   (b) Did you describe the limitations of your work? [Yes] See the Section 6
   (c) Did you discuss any potential negative societal impacts of your work? [N/A]
   (d) Have you read the ethics review guidelines and ensured that your paper conforms to them? [Yes]
2. If you are including theoretical results...
   (a) Did you state the full set of assumptions of all theoretical results? [Yes]
   (b) Did you include complete proofs of all theoretical results? [Yes] See the supplementary material.
3. If you ran experiments...
   (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? [Yes] See the supplementary material.
   (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? [Yes] See the supplementary material.
   (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? [Yes] See the Section 5

(d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? [Yes] See the supplementary material.

4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets...
   (a) If your work uses existing assets, did you cite the creators? [Yes]
   (b) Did you mention the license of the assets? [N/A] The repos we used do not mention the license.
   (c) Did you include any new assets either in the supplemental material or as a URL? [Yes]
   (d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? [No]
   (e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? [N/A]
5. If you used crowdsourcing or conducted research with human subjects...
   (a) Did you include the full text of instructions given to participants and screenshots, if applicable? [N/A]
   (b) Did you describe any potential participant risks, with links to Institutional Review Board (IRB) approvals, if applicable? [N/A]
   (c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? [N/A]