Appendices to the Neural Certificates for Safe Control Policies paper

A Proofs

Before presenting proofs for the theorems of this paper, we provide some preliminaries which will be used in deriving proofs. First, we present the definition of forward invariance of a set.

Definition A.1 (Forward invariance of a set). *Given a dynamical system $\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x})$ with $\boldsymbol{f} : \mathcal{X} \to \mathcal{X} \subseteq \mathbb{R}^n$, we say a set $\mathcal{S} \subseteq \mathcal{X}$ is forward invariant if for every $\boldsymbol{x}(0) \in \mathcal{S}$, any state along the system trajectory $\boldsymbol{x}(t)$ starting from $\boldsymbol{x}(0)$ have $\boldsymbol{x}(t) \in \mathcal{S}$ for all $t \in \mathbb{R}_0^+$.*

Second, we provide the Nagumo's Theorem [1–3] which establishes a sufficient and necessary condition to verify the invariance of a sub-level set of a continuously differentiable function.

Theorem A.2 (Nagumo's Theorem [1–3]). *Consider a dynamical system $\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x})$ with $\boldsymbol{f} : \mathcal{X} \to \mathcal{X} \subseteq \mathbb{R}^n$ and a continuously differentiable scalar function $h(\boldsymbol{x}) : \mathcal{X} \to \mathbb{R}$. Define the zero sub-level set of $h(\boldsymbol{x})$ as $\mathcal{C} = \{\boldsymbol{x} \in \mathcal{X} : h(\boldsymbol{x}) \leq 0\}$. The following two conditions are equivalent:*

1) *$\mathcal{C}$ is forward invariant in a sense of Definition A.1;*
2) *$\nabla h(\boldsymbol{x})\boldsymbol{f}(\boldsymbol{x}) \leq 0$, if $\boldsymbol{x} \in \{\boldsymbol{x} \in \mathcal{X}, h(\boldsymbol{x}) = 0\}$.*

The above Nagumo's Theorem states that $\dot{h}(t) = \nabla h(\boldsymbol{x})\boldsymbol{f}(\boldsymbol{x}) \leq 0$ on the boundary of the zero sub-level set $\mathcal{C}$ is a necessary and sufficient for $\mathcal{C}$ to be forward invariant. Please refer to [1–3] for more detail and proofs of this claim.

A.1 Proof of Lemma 3.1

Assume that a barrier function $B(\boldsymbol{x})$ satisfying the three conditions in (4) can be found. Take any trajectory $\boldsymbol{x}_{\boldsymbol{u}}(t)$ in $\mathcal{X}$ that starts at some $\boldsymbol{x}(0) \in \mathcal{X}_0$ and consider the evaluation of $B(\boldsymbol{x}_{\boldsymbol{u}}(t))$ along the trajectory. The condition (4c) directly indicates the second condition in the Nagumo's Theorem A.2 holds, which is equivalent to say that $\{\boldsymbol{x} \in \mathcal{X} : B(\boldsymbol{x}) \leq 0\}$ is forward invariance according to the Nagumo's Theorem. Thus, along the trajectory $\boldsymbol{x}_{\boldsymbol{u}}(t)$, $B(\boldsymbol{x}_{\boldsymbol{u}}(t)) \leq 0$ holds for all $t \in \mathbb{R}_0^+$. Consequently, any such trajectory can never reach an unsafe state whose $B(\boldsymbol{x})$ is positive according to (4b). We conclude that the safety of the system is guaranteed. □.

A.2 Proof of Theorem 1

The proof of Theorem 1 consists of three steps.

First, we need to show that the following set

$$\mathcal{A} = \{\boldsymbol{x} \in \mathcal{X} : V(\boldsymbol{x}) \leq 0\} \tag{I.1}$$

is closed and invariant. Its closeness is straightforward, and the invariance can be proved by applying the Nagumo's Theorem A.2. Specifically, from the condition (5b), if $\boldsymbol{x} \in \{\boldsymbol{x} \in \mathcal{X} : V(\boldsymbol{x}) = 0\}$, then $\nabla V(\boldsymbol{x})\boldsymbol{f}_{\boldsymbol{u}}(\boldsymbol{x}) \leq 0$, which directly indicates the second condition in the Nagumo's Theorem holds. Thus, we can say that $\mathcal{A}$ is invariant by applying the Nagumo's Theorem.

Second, we need to define another Lyapunov function

$$V_{\mathcal{A}}(\boldsymbol{x}) = \begin{cases} 0 & \text{if } \boldsymbol{x} \in \mathcal{A} \\ V(\boldsymbol{x}) & \text{if } \boldsymbol{x} \in \mathcal{X} \backslash \mathcal{A} \end{cases}. \tag{I.2}$$

Combining the definition of the Lyapunov-like function in Definition 3.2, it is easy to show the following properties of the Lyapunov function in (I.2): (i) $V_{\mathcal{A}}(\boldsymbol{x}) = 0$ for all $\boldsymbol{x} \in \mathcal{A}$; (ii) $V_{\mathcal{A}}(\boldsymbol{x}) > 0$ for all $\boldsymbol{x} \in \mathcal{X} \backslash \mathcal{A}$ due to (I.1); and (iii) for all $\boldsymbol{x} \in \mathcal{X} \backslash \mathcal{A}$, we have

$$V_{\mathcal{A}}(\boldsymbol{x}) = V(\boldsymbol{x}) \leq -\beta(V(\boldsymbol{x})) = -\beta(V_{\mathcal{A}}(\boldsymbol{x})), \tag{I.3}$$

which is a result of directly applying the condition (5b) in Definition 3.2.

Third, based on the results obtained in the first and second steps, and also from the fact that $V_{\mathcal{A}}(\boldsymbol{x})$ is continuous on its domain and continuously differentiable at every point $\boldsymbol{x} \in \mathcal{X} \backslash \mathcal{A}$, we can directly apply Theorem 2.8 in [4] to show that $\mathcal{A}$ is asymptotically stable; that is, there exists a $\mathcal{KL}$-function γ such that for any $\boldsymbol{x}(0) \in \mathcal{X} \backslash \mathcal{A}$,

$$\|\boldsymbol{x}_{\boldsymbol{u}}(t)\|_{\mathcal{A}} \leq \gamma(\|\boldsymbol{x}(0)\|_{\mathcal{A}}, t) \tag{I.4}$$

holds for all $t \in \mathbb{R}_0^+$. Also combining $\mathcal{A} \subseteq \mathcal{X}_g$ in (5a) of Definition 3.2, it follows that

$$\|\boldsymbol{x}_{\boldsymbol{u}}(t)\|_{\mathcal{X}_g} \leq \|\boldsymbol{x}_{\boldsymbol{u}}(t)\|_{\mathcal{A}} \leq \gamma(\|\boldsymbol{x}(0)\|_{\mathcal{A}}, t) \tag{I.5}$$

holds for all $t \in \mathbb{R}_0^+$. This directly indicates that the controlled system is goal-reaching by Definition 2.2. Thus, we conclude that the existence of the Lyapunov-like function in Definition 3.2 guarantees the goal-reaching of the controlled system. $\square$.

B Experiment Details

Pendulum system. The equation of motion for the pendulum system is

$$\ddot{\alpha} = -\frac{g}{l}\sin(\alpha) - \frac{d}{ml^2}\dot{\alpha} + \frac{u}{ml^2}, \tag{I.6}$$

with the constants set as $g = 10, l = 1, m = 1, d = 0.1$. We define the state variable to be $\boldsymbol{x} = [\alpha, \dot{\alpha}]'$. Define the state domain as $\mathcal{X} = \{\boldsymbol{x} : \boldsymbol{x}_{\text{lb}} \leq \boldsymbol{x} \leq \boldsymbol{x}_{\text{ub}}\}$ with $\boldsymbol{x}_{\text{lb}} = [-\pi, -5]'$ and $\boldsymbol{x}_{\text{ub}} = [\pi, 5]'$, the unsafe state set $\mathcal{X}_u = \{\boldsymbol{x} : 2.5 \leq \|\boldsymbol{x}\|_2 \leq 3\}$, the goal state set $\mathcal{X}_g = \{\boldsymbol{x} : \|\boldsymbol{x}\|_2 = 0\}$, and the initial state set $\mathcal{X}_0 = \{\boldsymbol{x} : \|\boldsymbol{x}\|_2 \leq 2\}$.

We set the neural policy network to be one-layer linear function (without bias): $u = K\boldsymbol{x}$. For the neural barrier function $B_{\boldsymbol{\theta}}(\boldsymbol{x})$, we use a 2-16-16-1 fully connected network with $\tanh$ activation function. For the neural Lyapunov-like function $V_{\boldsymbol{\omega}}(\boldsymbol{x})$, we use a 2-16-16-1 fully connected network with $\tanh$ activation function, but the last layer is modified to be dot product operation. The learning rate is set as 10^{-3}. Note that in our experiments, we always choose 4-layer neural networks for certificate functions, and the number of nodes in layers is set as n-$8n$-$8n$-1 with n is the dimension of input layer (i.e., state dimension).

Only with Lyapunov-like certificate, the learned neural policy is

$$u = [-0.3286, -0.5950]\boldsymbol{x}. \tag{I.7}$$

With both the Lyapunov-like and barrier certificates, the learned neural policy is

$$u = [2.0120, -2.1343]\boldsymbol{x}. \tag{I.8}$$

Cartpole system. The equation of the motion for the cartpole system is

$$\ddot{x} = \frac{u + m_p \sin\theta(l\dot{\theta}^2 - g\cos\theta)}{m_c + m_p(\sin\theta)^2}, \tag{I.9a}$$

$$\ddot{\theta} = \frac{u\cos\theta + m_p l\dot{\theta}^2\cos\theta\sin\theta - (m_c + m_p)g * \sin\theta}{l(m_c + m_p(\sin\theta)^2)}, \tag{I.9b}$$

with the constants set as $m_c = 1, m_p = 1, g = 1, l = 1$. The system state variable is defined as $\boldsymbol{x} = [x, \theta, \dot{x}, \dot{\theta}]'$, where x is the position of the cart and θ is the angle between pole and upward direction. Define the state space $\mathcal{X} = \{\boldsymbol{x} : \boldsymbol{x}_{\text{lb}} \leq \boldsymbol{x} \leq \boldsymbol{x}_{\text{ub}}\}$ with $\boldsymbol{x}_{\text{lb}} = [\text{-}1.3, \text{-}1.3, \text{-}1.3, \text{-}1.3]'$ and $\boldsymbol{x}_{\text{ub}} = [1.3, 1.3, 1.3, 1.3]'$, the unsafe state set $\mathcal{X}_u = \{\boldsymbol{x} : 0.9 \leq \|\boldsymbol{x}\|_2 \leq 1.3\}$, the goal state set $\mathcal{X}_g = \{\boldsymbol{x} : \|\boldsymbol{x}\|_2 = 0\}$, and the initial state set $\mathcal{X}_0 = \{\boldsymbol{x} : \|\boldsymbol{x}\|_2 \leq 0.8\}$.

We set the neural policy network to be one-layer linear function (without bias): $u = K\boldsymbol{x}$. For the neural barrier function $B_{\boldsymbol{\theta}}(\boldsymbol{x})$, we use a 4-32-32-1 fully connected network with $\tanh$ activation function. For the neural Lyapunov-like function $V_{\boldsymbol{\omega}}(\boldsymbol{x})$, we use a 4-32-32-1 fully connected network with $\tanh$ activation function, but the last layer is modified to be dot product operation. The learning rate is set as 10^{-3}.

Only with Lyapunov-like certificate, the learned neural policy is

$$u = [-0.0652, -0.2577, -1.3080, -0.6947]\boldsymbol{x}. \tag{I.10}$$

With both the Lyapunov-like and barrier certificates, the learned neural policy is

$$u = [-1.5064, -0.7969, -3.1892, -1.5950]\boldsymbol{x}. \tag{I.11}$$

Vehicle path tracking system. The kinematic model of a wheeled vehicle tracking a reference path is given by [5]:

$$\dot{s} = \frac{v\cos\theta_e)}{1 - d_e\kappa(s)}, \tag{I.12a}$$

$$\dot{d}_e = v\sin(\theta_e), \tag{I.12b}$$

$$\dot{\theta}_e = \frac{v\tan(u)}{L} - \frac{v\kappa(s)\cos\theta_e}{1 - d_e\kappa(s)}, \tag{I.12c}$$

where $\theta_e = \theta - \theta_r$ is the angle error between the vehicle orientation θ and the reference path tangent angle θ_r; d_e is the distance error (see Figure 4. (c) in [5]); and the constants are $v = 6$ and $L = 1$. Assuming that the reference path is a unit circle. Define the state domain $\mathcal{X} = \{\boldsymbol{x} : \boldsymbol{x}_{\text{lb}} \leq \boldsymbol{x} \leq \boldsymbol{x}_{\text{ub}}\}$ with $\boldsymbol{x}_{\text{lb}} = [\text{-}0.8, \text{-}0.8]'$ and $\boldsymbol{x}_{\text{ub}} = [0.8, 0.8]'$, the unsafe state set $\mathcal{X}_u = \{\boldsymbol{x} : 0.6 \leq \|\boldsymbol{x}\|_2 \leq 0.8\}$, the initial state set $\mathcal{X}_0 = \{\boldsymbol{x} : \|\boldsymbol{x}\|_2 \leq 0.5\}$, and the goal state set $\mathcal{X}_g = \{\boldsymbol{x} : \|\boldsymbol{x} - \boldsymbol{x}_{\text{g}}\|_2 \leq 0.2\}$ with $\boldsymbol{x}_g = [\text{-}0.2, 0]'$.

We set the neural policy network to be one-layer linear function (without bias): $u = K\boldsymbol{x}$. For the neural barrier function $B_{\boldsymbol{\theta}}(\boldsymbol{x})$, we use a 2-16-16-1 fully connected network with $\tanh$ activation function. For the neural Lyapunov-like function $V_{\boldsymbol{\omega}}(\boldsymbol{x})$, we use a 2-16-16-1 fully connected network with $\tanh$ activation function, but the last layer is modified to be dot product operation. The learning rate is set as 10^{-3}. With both the Lyapunov-like and barrier certificates, the learned neural policy is

$$u = [-0.3662, -1.7802]\boldsymbol{x}. \tag{I.13}$$

UAV control. The motion of equation for a UAV flying in planar is given by:

$$\ddot{x} = \frac{-(u_1 + u_2)\sin\theta}{m}, \tag{I.14a}$$

$$\ddot{y} = \frac{(u_1 + u_2)\cos\theta - mg}{m}, \tag{I.14b}$$

$$\ddot{\theta} = \frac{r(u_1 - u_2)}{I}, \tag{I.14c}$$

where the constants are set as $m = 0.1, g = 0.1, I = 0.1, r = 0.1$. We define the state variable as $\boldsymbol{x} = [x, y, \theta, \dot{x}, \dot{y}, \dot{\theta}]'$ and the control variable $\boldsymbol{u} = [u_1, u_2]'$. Define the state domain $\mathcal{X} = \{\boldsymbol{x} : \boldsymbol{x}_{\text{lb}} \leq \boldsymbol{x} \leq \boldsymbol{x}_{\text{ub}}\}$ with $\boldsymbol{x}_{\text{lb}} = [\text{-}1, \text{-}1, \text{-}1, \text{-}1]'$ and $\boldsymbol{x}_{\text{ub}} = [1, 1, 1, 1]'$, the unsafe state set $\mathcal{X}_u = \{\boldsymbol{x} : 0.9 \leq \|\boldsymbol{x}\|_2 \leq 1\}$, the initial state set $\mathcal{X}_0 = \{\boldsymbol{x} : \|\boldsymbol{x}\|_2 \leq 0.5\}$, and the goal state set $\mathcal{X}_g = \{\boldsymbol{x} : \|\boldsymbol{x}\|_2 = 0\}$.

We set the neural policy network to be one-layer linear function (without bias): $\boldsymbol{u} = K\boldsymbol{x}$. For the neural barrier function $B_{\boldsymbol{\theta}}(\boldsymbol{x})$, we use a 6-48-48-1 fully connected network with $\tanh$ activation function. For the neural Lyapunov-like function $V_{\boldsymbol{\omega}}(\boldsymbol{x})$, we use a 6-48-48-1 fully connected network with $\tanh$ activation function, but the last layer is modified as dot product operation. The learning rate is set as 10^{-3}. With both the Lyapunov-like and barrier certificates, the learned neural policy is

$$\boldsymbol{u} = \begin{bmatrix} 0.8185, & 0.8221, & -1.9815, & 2.4234, & -0.2271, & -1.8433 \\ 0.9136, & -1.0979, & -1.8189, & -0.0967, & -5.1917, & 0.3099 \end{bmatrix}\boldsymbol{x}. \tag{I.15}$$

References

[1] Franco Blanchini. Set invariance in control. *Automatica*, 35(11):1747–1767, 1999.

[2] Franco Blanchini and Stefano Miani. *Set-theoretic methods in control*. Springer, 2008.

[3] Aaron D Ames, Xiangru Xu, Jessy W Grizzle, and Paulo Tabuada. Control barrier function based quadratic programs for safety critical systems. *IEEE Transactions on Automatic Control*, 62(8):3861–3876, 2016.

[4] Yuandan Lin, Eduardo D Sontag, and Yuan Wang. A smooth converse lyapunov theorem for robust stability. *SIAM Journal on Control and Optimization*, 34(1):124–160, 1996.

[5] Ya-Chien Chang, Nima Roohi, and Sicun Gao. Neural lyapunov control. In *Advances in Neural Information Processing Systems*, 2019.