ECE 586 BH: Interplay between Control and Machine Learning **Fall 2023**

Homework 2

Instructor: Bin Hu Due date: Oct 10, 2023

1. (60 points) Consider $x_{k+1} = Ax_k + B_k\sigma(C_k x_k + b_k)$, where σ is the ReLU activation function. We know that if there exist a positive definite diagonal matrix Λ_k such that the following matrix inequality holds

$$\begin{bmatrix} A_k^\mathsf{T} A_k - I & A_k^\mathsf{T} B_k \\ B_k^\mathsf{T} A_k & B_k^\mathsf{T} B_k \end{bmatrix} + \begin{bmatrix} C_k & 0 \\ 0 & I \end{bmatrix}^\mathsf{T} \begin{bmatrix} 0 & \Lambda_k \\ \Lambda_k & -2\Lambda_k \end{bmatrix} \begin{bmatrix} C_k & 0 \\ 0 & I \end{bmatrix} \leq 0, \tag{1}$$

then we have $\|x'_{k+1} - x_{k+1}\| \leq \|x'_k - x_k\|$. Apply the above condition to show

(a) The convex potential layer $x_{k+1} = x_k - \frac{2}{\|W_k\|^2} W_k \sigma(W_k^\mathsf{T} x + b_k)$ is 1-Lipschitz.

(b) The layer $x_{k+1} = x_k - 2W_k\sigma(W_k^\mathsf{T} x_k + b_k)$ with $W_k^\mathsf{T} W_k = I$ is 1-Lipschitz.

(c) The layer $x_{k+1} = x_k - 2W_k \operatorname{diag}(\sum_{j=1}^n |W_k^\mathsf{T} W_k|_{ij})^{-1}\sigma(W_k^\mathsf{T} x_k + b_k)$ is 1-Lipschitz (here we assume $\sum_{j=1}^n |W_k^\mathsf{T} W_k|_{ij} \neq 0$ for all i, and n is the column dimension of $W_k^\mathsf{T} W_k$).

(d) The layer $x_{k+1} = \sqrt{2} M_k^\mathsf{T} \Psi_k \sigma(\sqrt{2}\Psi_k^{-1} N_k x_k + b_k)$ with Ψ_k being any positive definite diagonal matrix and (M_k, N_k) being any matrix pair satisfying $M_k M_k^\mathsf{T} + N_k N_k^\mathsf{T} = I$ is 1-Lipschitz.

2. (10 points) Consider the deep equilibrium model $z = \sigma(Wz + Ux + b_z)$ with σ being ReLU. Suppose the model is set up in a way that there is a unique solution z for any input x (so we assume the deep equilibrium model is well-posed). Derive a matrix inequality condition to ensure $\|z' - z\| \leq L\|x' - x\|$ for any (x', x). (Notice that z is the output of x, and z' is the output of x'.)

Hint: Since σ is ReLU, we must have

$$\begin{bmatrix} y' - y \\ \sigma(y') - \sigma(y) \end{bmatrix}^\mathsf{T} \begin{bmatrix} 0 & \Lambda_k \\ \Lambda_k & -2\Lambda_k \end{bmatrix} \begin{bmatrix} y' - y \\ \sigma(y') - \sigma(y) \end{bmatrix} \geq 0$$

where Λ_k is any positive definite diagonal matrix. Use the above quadratic constraint to derive the matrix inequality condition.

3. Consider a control affine system $\dot{x} = f(x) + g(x)u$.

(a) (15 points) Suppose we want the system to avoid the area defined by $\{x : \|x - x*\| \leq 3\}$ where x^* is a prescribed point. We assume that the perfect state measurement is available. In addition, we also assume that a baseline controller is given by $u = K(x)$, where K is some neural network policy. How can we apply the control barrier function (CBF) approach to solve this problem?

(b) (15 points) Now suppose only a state estimate $\hat{x}$ is provided. A prescribed error bound is given, i.e. we always have $\|x(t) - \hat{x}(t)\| \leq r$ for our system. We still want the system to avoid the area defined by $\{x : \|x - x*\| \leq 3\}$. How can we achieve this via a CBF-based approach? You can make assumptions on Lipschitz properties as needed.