Nonlinear Network Structures for Optimal Control

Frank L. Lewis and Murad Abu-Khalaf
Advanced Controls, Sensors, and MEMS (ACSM) group

System

$$\dot{x} = f(x) + g(x)u(x)$$

Cost

$$V = \int_0^\infty \left[Q(x) + W(u) \right] dt$$

The Usual Suspects

$$Q(x) = x^T Q x \qquad W(u) = u^T R u$$

Definition 2.1: *Admissible Controls*

Let $\Psi(\Omega)$ denote the set of admissible controls. A control $u:\Re^n \to \Re^m$ is defined to be admissible with respect to the state penalty function $Q(x)$ on Ω, denoted $u \in \Psi(\Omega)$, if:

- u is continuous on Ω,
- $u(0) = 0$,
- u stabilizes (1) on Ω,
- $\int_0^\infty [Q(x) + W(u)]dt < \infty$, $\forall x \in \Omega$

A stabilizing control may not be admissible!

NONLINEAR QUADRATIC REGULATOR

Generalized HJB Equation

$$GHJB(V,u) \overset{\Delta}{=} \frac{\partial V}{\partial x}^T \left(f + gu \right) + Q + u^T R u = 0,$$
$$V(0) = 0.$$

Optimal Control (SVFB)

$$u^*(x) = -\frac{1}{2} R^{-1} g^T(x) \frac{\partial V^*(x)}{\partial x}$$

Hamilton-Jacobi-Bellman (HJB) Equation

$$HJB(V^*) \overset{\Delta}{=} \frac{\partial V^*}{\partial x}^T f + Q - \frac{1}{4} \frac{\partial V^*}{\partial x}^T g R^{-1} g^T \frac{\partial V^*}{\partial x} = 0,$$
$$V^*(0) = 0.$$

PROBLEM- HJB usually has no analytic solution
SOLUTION- Successive Approximation

$u^{(0)}$ a stabilizing control

$$\frac{\partial V^{(i)T}}{\partial x}\left(f + gu^{(i)}\right) + Q + W(u^{(i)}) = 0,\ V^{(i)}(0) = 0$$

$$u^{(i+1)} = -\frac{1}{2}R^{-1}g^T \frac{\partial V^{(i)T}}{\partial x}$$

A contraction map
(Saridis)

Saridis and Beard used Galerkin
Approx to allow for GHJB solution

Converges to optimal solution

Gives u(x) in SVFB form

For Constrained Controls

NONLINEAR NONQUADRATIC REGULATOR

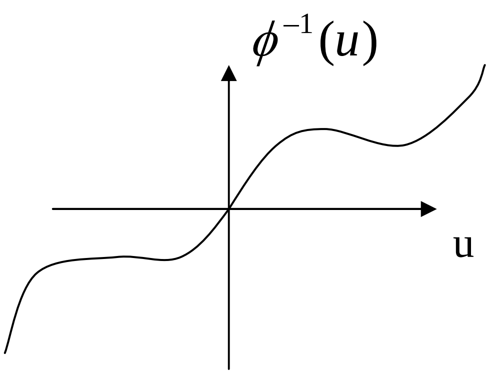
$$V = \int_0^\infty \left[Q(x) + W(u) \right] dt$$

with

$$W(u) = 2\int_0^u \left(\phi^{-1}(\mu) \right)^T R d\mu,$$

Nonquadratic form- Lyshevsky

PD if $u\phi^{-1}(u) > 0 \; when \; u \neq 0$



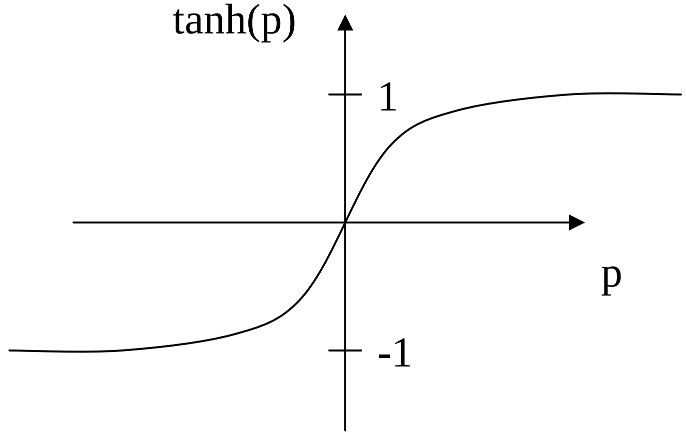
New GHJB is

$$\frac{\partial V^T}{\partial x}\left(f+gu\right)+Q(x)+2\int\left(\phi^{-1}(u)\right)^T R\,du=0,\ \ V(0)=0$$

$$u=-\phi\left(\frac{1}{2}R^{-1}g^T\frac{\partial V^T}{\partial x}\right)$$

Natural, exact,
no approximation

u(t) constrained if $\phi(.)$ is a saturation function!



Problem- cannot solve HJB

Solution- Use Successive Approximation on GHJB

Iterate:

$u^{(0)}$ a stabilizing control

$$\frac{\partial V^{(i)T}}{\partial x}\left(f + gu^{(i)}\right) + Q(x) + 2\int_0^{u^{(i)}} \left(\phi^{-1}(\mu)\right)^T R\, d\mu = 0,\ V^{(i)}(0) = 0$$

$$u^{(i+1)} = -\phi\left(\frac{1}{2} R^{-1} g^T \frac{\partial V^{(i)T}}{\partial x}\right)$$

Lemma 3.1: *Improved Saturated Control Law*

If $u^{(i)} \in \Psi(\Omega)$, and $V^{(i)}$ satisfies the equation $GHJB(V^{(i)}, u^{(i)}) = 0$ with the boundary condition $V^{(i)}(0) = 0$, then the new control derived as

$$u^{(i+1)}(x) = -\phi\left(\frac{1}{2} R^{-1} g^T(x) \frac{\partial V^{(i)}(x)}{\partial x}\right)$$

is an admissible control for the system on Ω. Moreover, if the control bound $\phi(\cdot)$ is monotonically non-decreasing and $V^{(i+1)}$ is the unique positive definite function satisfying the equation $GHJB(V^{(i+1)}, u^{(i+1)}) = 0$, with the boundary condition $V^{(i+1)}(0) = 0$, then $V^{(i+1)}(x) \le V^{(i)}(x) \quad \forall x \in \Omega$.

Theorem 3.2: *Convergence of Successive Approximations*

If $u^{(0)} \in \Psi(\Omega)$, then

1. $V^{(i)} \to V^*$ uniformly on Ω
2. $u^{(i)} \in \Psi(\Omega), \quad \forall i \ge 0$
3. $u^{(i)} \to u^*$

Lemma 3.4: *Optimal Saturated Control has the Largest Stability Region*

The saturated control u^* has a stability region that is the largest of any other saturated control $u^{(i)}$ that is admissible with respect to $Q(x)$ and the system (f, g).

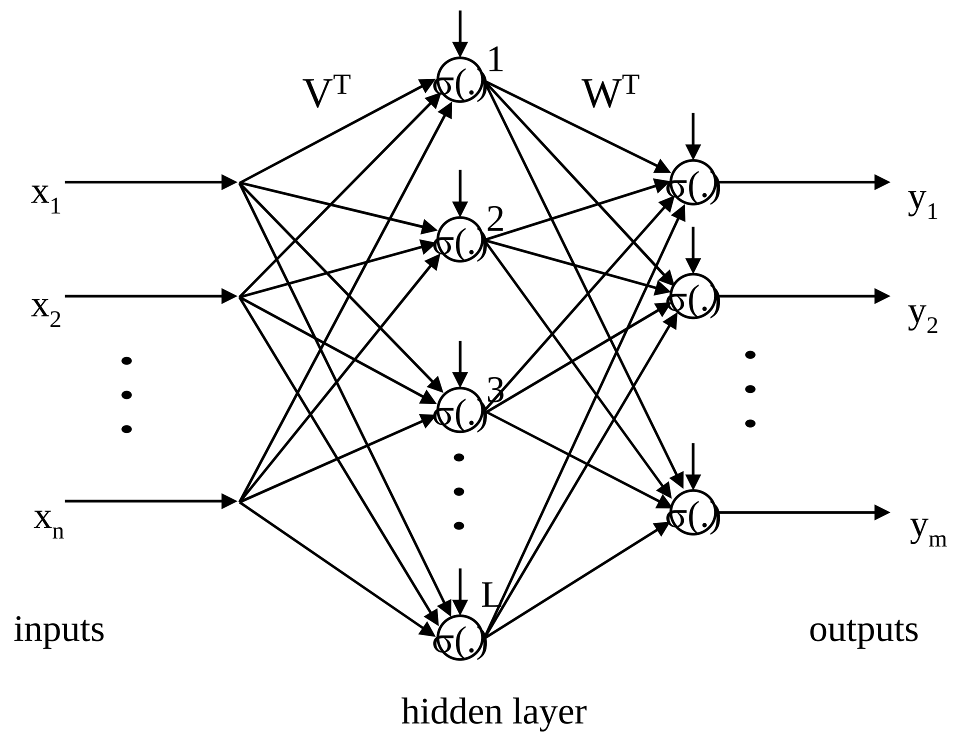
Note that there may be stabilizing saturated controls that have larger stability regions than u^*, but are not admissible with respect to $Q(x)$ and the system (f, g).

Problem- Cannot solve GHJB!

Solution- Neural Network to approximate $V^{(i)}(x)$

$$V_L^{(i)}(x) = \sum_{j=1}^{L} w_j^{(i)} \sigma_j(x) = W_L^{T(i)} \overline{\sigma}_L(x),$$

Select basis set $\overline{\sigma}_L(x)$



Two-Layer Neural Network with adjustable output weights

Cost gradient approximation

$$\frac{\partial V_L{}^{(i)}}{\partial x} = \frac{\partial \overline{\sigma}_L(L)}{\partial x}^T W_L{}^{(i)} = \nabla \overline{\sigma}_L{}^T(x) W_L{}^{(i)}$$

Let

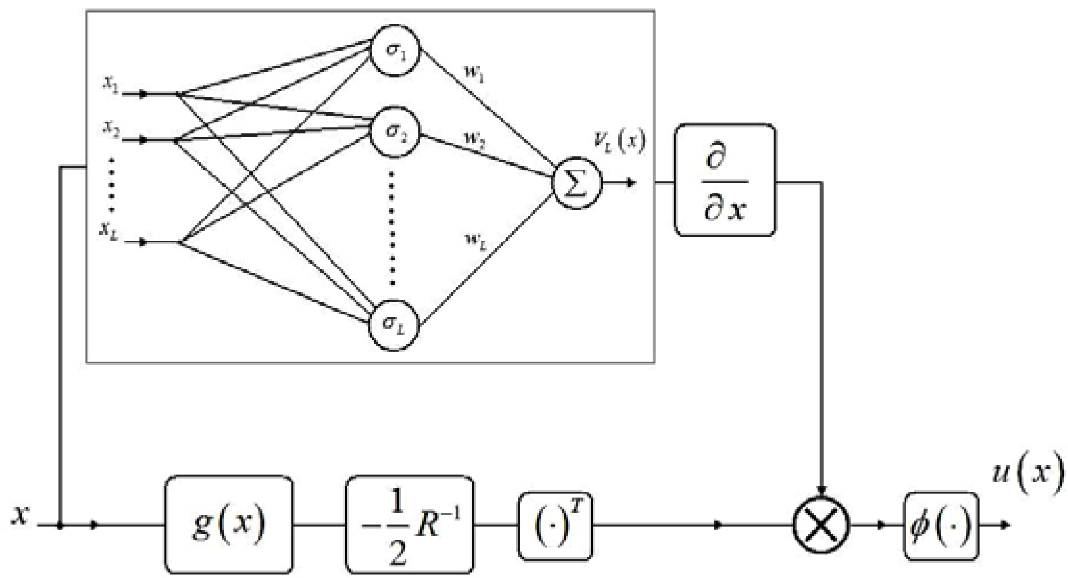
$$\phi(p) = A \cdot \tanh(\frac{p}{A} \cdot)$$

Nonzero residual!

Then GHJB is

$$W_L{}^{(i)T} \nabla \overline{\sigma}_L (f + g u^{(i)}) + Q + 2u^{(i)} R A \tanh^{-1}\left(\frac{u^{(i)}}{A}\right) + A^2 R \ln\left(1 - \left(\frac{u^{(i)}}{A}\right)^2\right) = \varepsilon(x)$$

$$u^{(i)}(x) = -A \tanh\left(\frac{1}{2A} R^{-1} g^T(x) \nabla \overline{\sigma}_L^T W_L{}^{(i-1)}\right)$$



Neural-network-based nearly optimal saturated control law.

To minimize the residual error in a LS sense

Evaluate the GHJB at a number of points $x_1, x_2, \ldots, x_N$ on Ω

Note, if

$$u^{(i)}(x) = -A\tanh\left(\frac{1}{2A}R^{-1}g^T(x)\nabla\bar{\sigma}_L^T W_L^{(i-1)}\right)$$

$$A(x, W_L^{(i)}) \equiv \nabla\bar{\sigma}_L(f + gu^{(i)})$$

$$-b(x, W_L^{(i)}) \equiv Q + 2u^{(i)}RA\tanh^{-1}\left(\frac{u^{(i)}}{A}\right) + A^2R\ln\left(1 - \left(\frac{u^{(i)}}{A}\right)^2\right)$$

Then, GHJB is

$$W_L^{(i)T}\left\lfloor A(x, W_L^{(i-1)T})\right\rfloor = b(x, W_L^{(i)T}) + \varepsilon(x)$$

Evaluating this at N points gives

L x N coefficient matrix

$$W_L^{(i)T}\left[A(x_1,W_L^{(i-1)T})\;A(x_2,W_L^{(i-1)T})\cdots A(x_N,W_L^{(i-1)T})\right]$$

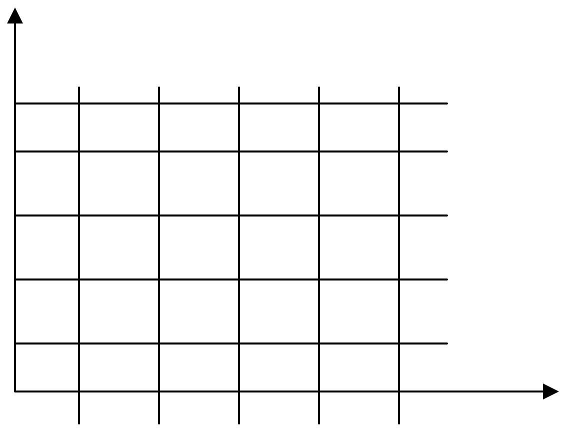
$$=\left[b(x_1,W_L^{(i)T})\;b(x_{2,}W_L^{(i)T})\cdots b(x_N,W_L^{(i)T})\right]$$

Solve by LS

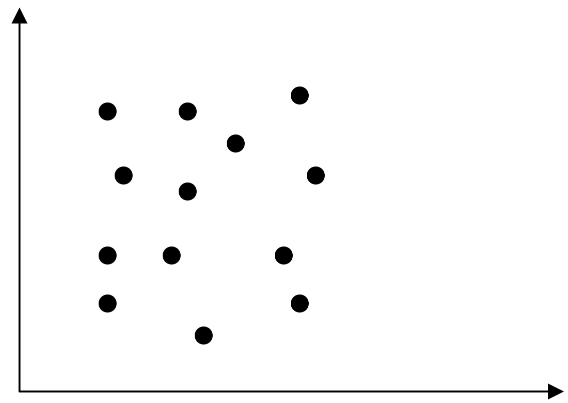
NN Training Set!

Select the N sample points x_k

Uniform Mesh Grid in R^n



Random selection- Montecarlo



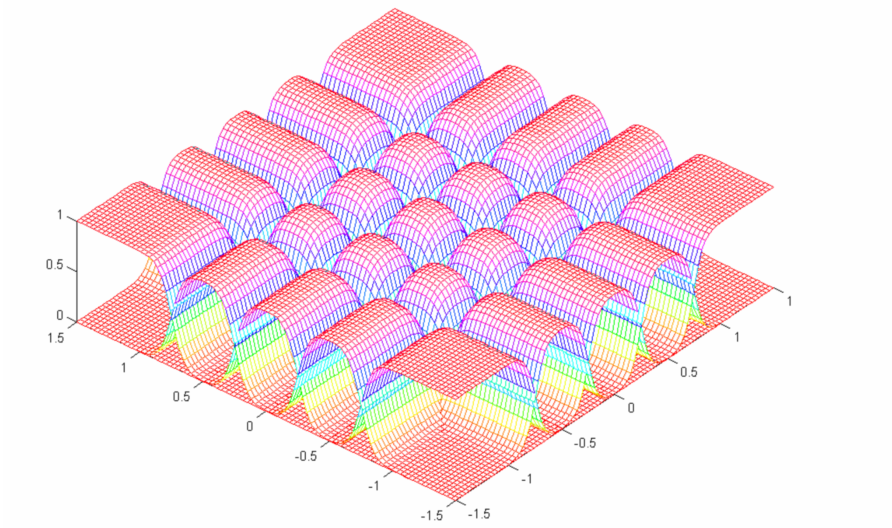
Approximation error is $\frac{1}{N^{2/n}}$
(Barron)

Approximation error is $\frac{1}{N}$

Montecarlo overcomes NP-complexity problems!

ASIDE-

Useful for reducing complexity of fuzzy logic systems?



Uniform grid of
Separable Gaussian activation functions for RBF NN

Lemma 3.1: Equation (28) will have a unique solution when

$$\sum_{k=1}^{N} A(x_k, W_L^{(i-1)}) A^T (x_k, W_L^{(i-1)}) \geq \rho I$$

where ρ is a positive constant, and I is the identity matrix. This is a persistency of excitation (PE) condition on $A(x_k, W_L^{(i-1)})$.

This guides the choice of the N sample vectors x_k

NN Training Set must be PE

Algorithm and Proofs work for any Q(x) in

$$V = \int_0^{\infty} \left[Q(x) + W(u) \right] dt$$

Constrained input given by

$$W(u) = 2\int_0^{u} \left(\phi^{-1}(\mu)\right)^T R d\mu,$$

CONSTRAINED STATE CONTROL

$$Q(x,k) = x^T Q x + \sum_{l=1}^{n_c} \left(\frac{x_l}{B_l - \alpha_l} \right)^k$$

k large and even

MINIMUM-TIME CONTROL

$$V = \int_0^{\infty} \left[\tanh(x^T Q\, x) + 2\int_0^{u} \left(\phi^{-1}(\mu)\right)^T R d\mu \right] dt$$

For small R and $x^T Q x >> 0$ this is approx. $V = \int_0^{t_s} 1 dt,$

Example: Linear system

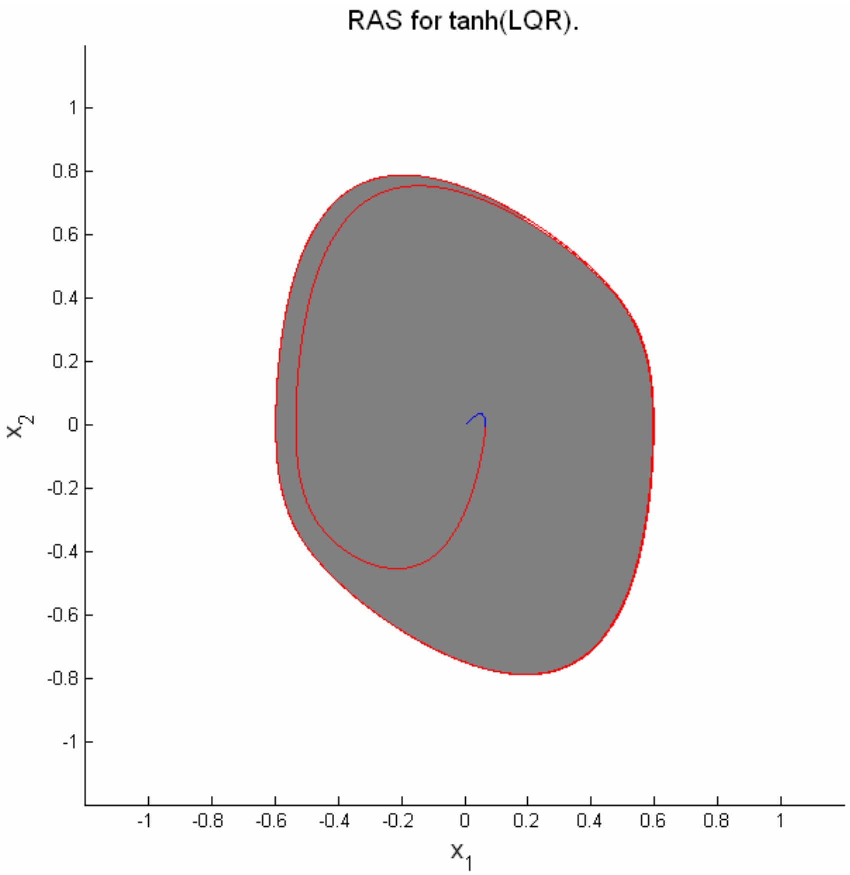
$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & -0.5 \\ 1 & 1.5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ -1 \end{bmatrix} u, \quad |u| \le 1$$

$$V_{15}(x_1, x_2) = w_1 x_1^2 + w_2 x_2^2 + w_3 x_1 x_2 + w_4 x_1^4 +$$

$$w_5 x_2^4 + w_6 x_1^3 x_2 + w_7 x_1^2 x_2^2 + w_8 x_1 x_2^3 + w_9 x_1^6 + w_{10} x_2^6$$

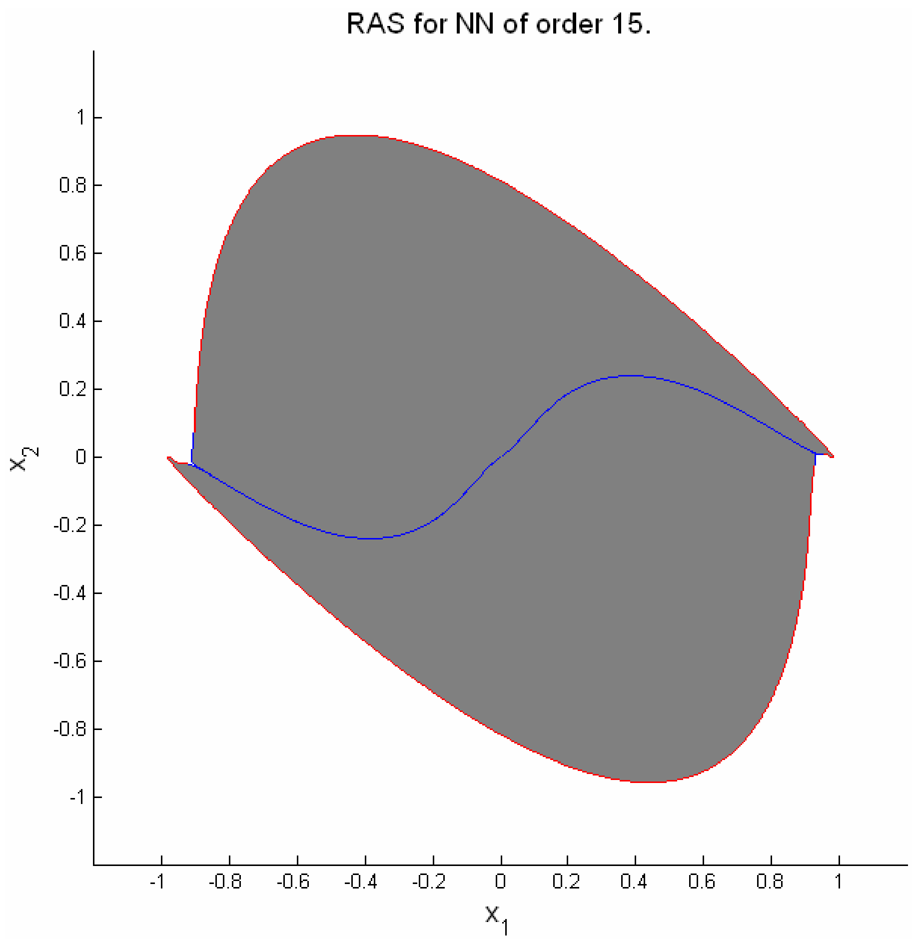
$$w_{11} x_1^5 x_2 + w_{12} x_1^4 x_2^2 + w_{13} x_1^3 x_2^3 + w_{14} x_1^2 x_2^4 + w_{15} x_1 x_2^5$$

$$-0.4 \le x_1 \le 0.4, \quad -0.4 \le x_2 \le 0.4,$$



Region of asymptotic stability for the initial controller,

$$u_0 = \tanh\left(LQR = -0.4142x_1 + 3.4142x_2\right)$$



Region of asymptotic stability
for the nearly optimal controller,

$$u = \tanh\left(\frac{1}{2}\frac{\partial V_{15}}{\partial x_2}\right),$$

$$W_{15\times1} = \begin{bmatrix} 8.85 & -0.76 & 3.51 & -2.52 & -1.64\ldots \\ -2.86 & -2.24 & 1.69 & 11.09 & 7.51\ldots \\ 20.61 & 21.57 & 24.35 & 20.84 & 10.53 \end{bmatrix}^T.$$

Example: Nonlinear oscillator

$$\dot{x}_1 = x_1 + x_2 - x_1(x_1^2 + x_2^2),$$
$$\dot{x}_2 = -x_1 + x_2 - x_2(x_1^2 + x_2^2) + u.$$

$$|u| \le 1$$

$$\begin{aligned} V_{24}(x_1, x_2) = {} & w_1 x_1^2 + w_2 x_2^2 + w_3 x_1 x_2 + w_4 x_1^4 + w_5 x_2^4 + \\ & w_6 x_1^3 x_2 + w_7 x_1^2 x_2^2 + w_8 x_1 x_2^3 + w_9 x_1^6 + w_{10} x_2^6 \\ & w_{11} x_1^5 x_2 + w_{12} x_1^4 x_2^2 + w_{13} x_1^3 x_2^3 + w_{14} x_1^2 x_2^4 + w_{15} x_1 x_2^5 + \\ & w_{16} x_1^8 + w_{17} x_2^8 + w_{18} x_1^7 x_2 + w_{19} x_1^6 x_2^2 + w_{20} x_1^5 x_2^3 \\ & + w_{21} x_1^4 x_2^4 + w_{22} x_1^3 x_2^5 + w_{23} x_1^2 x_2^6 + w_{24} x_1 x_2^7 \end{aligned}$$

$$u_0 = \tanh\left(-5x_1 - 3x_2,\right)$$

$$u = -\tanh\begin{pmatrix} 2.62x_1 + 4.23x_2 + 0.39x_2^3 - 4.0x_1^3 - 8.65x_1^2x_2 \\ -8.94x_1x_2^2 - 5.53x_2^5 + 2.26x_1^5 + 5.78x_1^4x_2 + \\ 11.00x_1^3x_2^2 + 2.57x_1^2x_2^3 + 2.00x_1x_2^4 + 2.08x_2^7 \\ -0.49x_1^7 - 1.65x_1^6x_2 - 2.71x_1^5x_2^2 - 2.19x_1^4x_2^3 \\ -0.76x_1^3x_2^4 + 1.77x_1^2x_2^5 + 0.87x_1x_2^6 \end{pmatrix}$$