Concepts for Breaking the Curse of Dimensionality for the Optimal Control HJB Equation

Karl Kunisch and Daniel Walter

University of Graz, and RICAM Linz, Austria

RICAM, October 2019

Closed loop optimal control

$$
\begin{cases}
\min\limits_{u(\cdot)\in U} J(u(\cdot),x) := \int\limits_0^\infty \ell(y(t)) + \frac{\gamma}{2}|u(t)|^2\,dt \\
\text{subject to} \quad \dot{y}(t) = f(y(t)) + g(y(t))u(t)\,, \quad y(0) = x
\end{cases}
$$

$\ell(0) = f(0) = 0,$ optimal value function

$$
V(x) := \min_{u(\cdot)\in U} J(u(\cdot),x)
$$

Hamilton-Jacobi-Bellman equation

$$
\min_{u\in U}\{DV(x)(f(x) + g(x)u) + \ell(x) + \frac{\gamma}{2}|u|^2\} = 0\,, \quad V(0) = 0\,,
$$

if $U \equiv$ linear space,

$$
u^*(x) = -\frac{1}{\gamma}g(x)^*DV(x)^*\,,
$$

then

$$
DV(x)f(x) - \frac{1}{2\gamma}DV(x)g(x)g(x)^*DV(x)^* + \ell(x) = 0\,.
$$

Close the loop

$$\dot{y}(t) = f(y(t)) - \frac{1}{\gamma} g(y(t)) g(y(t))^* DV(y(t))^* , \quad y(0) = x$$

good properties, but ...

$$\begin{cases} \min_u J(u(\cdot), x) := \frac{1}{2} \int_0^\infty |\mathcal{D} y(t)|^2 + \gamma |R^{\frac{1}{2}} u(t)|^2 \, dt \\ \text{subject to} \quad \dot{y}(t) = Ay(t) + Bu(t) , \quad y(0) = x. \end{cases}$$

under stabilizability and detectability assumption:

$$\Pi A + A^* \Pi - \Pi B R^{-1} B^* \Pi + \mathcal{D}^* \mathcal{D} = 0$$

$$u^* = -R^{-1} B^* \Pi y$$

closed loop system

$$\dot{y}(t) = A_\Pi y(t) = (A - BR^{-1} B^* \Pi) y(t), \quad y(0) = x$$

$A \sim A(t)$ as linearization of f M.Badra, T. Breiten, S.Ervedoza, J.-P. Raymond, ...

Before we start the analysis

IS HJB WORTH THE EFFORT ?

and if yes, how to get it ?

- solve it directly
- solve using tensor calculus (TT-rank)
- interpolate it from open loop data
- Taylor expansion
- Hopf formulas
- ...

Optimal HJB-based feedback stabilization of the Newell-Whitehead equation

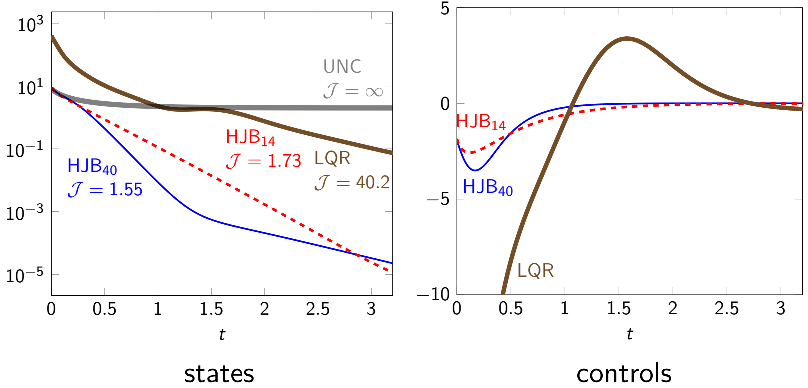
$$
\begin{aligned}
y_t &= \nu \Delta y + y(1 - y^2) + \chi_\omega(x) u(t) && \text{in } (-1,1) \times (0,\infty),\\
y_x(-1,t) &= y_x(1,t) = 0 && \text{for } t \geq 0,\\
y(x,0) &= y_0(x) && \text{in } (-1,1),
\end{aligned}
$$

Note: 0 is unstable, ± 1 are stable equilibria

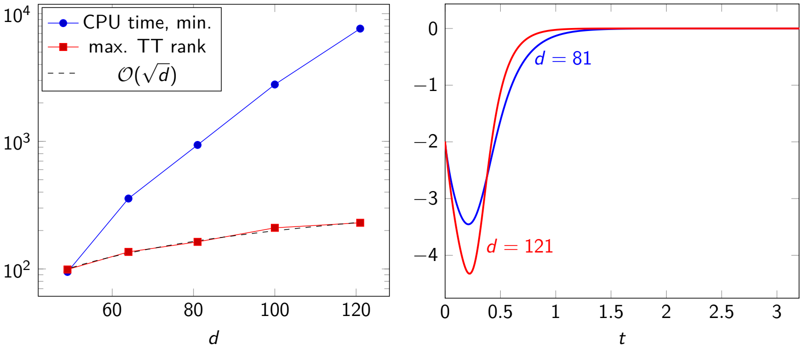
describes excitable systems such as neurons or axons, relates to Schlögel model, describing Rayleigh-Benard convection.

tensor train computations, jointly with S.Dolgov and D. Kalise, up to dimension 100

Newell-Whitehead Equation $d = 40$



2-D Newell-Whitehead Equation , 121 DoFs



Structure exploiting policy iteration

$$DV(x)f(x) - \frac{1}{2\gamma}DV(x)g(x)g(x)^*DV(x)^* + \ell(x) = 0\,.$$

$$u^*(x) = -\frac{1}{\gamma}g(x)^*DV(x)^*$$

- Solving nonlinear HJB: policy iteration (Howard's alg.), Newton method, Newton-Kleinman iteration for Riccati equations.

Successive Approximation Algorithm

Data: Initialization:*tol*, stabilizing control $u^0(x)$

while $\|V^n - V^{n+1}\| \geq tol$ **do**

1. Solve for $V^{n+1}(x)$

$$(f(x) + gu^n)^T \nabla V^{n+1}(x) + \ell(x) + \frac{1}{2\gamma}\|u^n(x)\|^2 = 0\,.$$

2. Update $u^{n+1}(x) = -\frac{1}{2} g^T \nabla V^{n+1}(x)$.
3. $n = n + 1$.

end

Result: $V^\infty(x), u^\infty(x)$

- $u^0(x)$ must be asymptotically stabilizing or
- discounting

Two 'infinities': the dynamical system

Meshfree discretization of dynamical system, e.g. pseudo-spectral collocation based on Chebysheff polynomials

- The state $x(t) = (x_1(t), \ldots, x_d(t))^t \in \mathbb{R}^d$.
- The free dynamics $f(x) : \mathbb{R}^d \to \mathbb{R}^d$ are $\mathcal{C}^1$ and separable in every coordinate $f_i(x)$

$$f_i(x) = \sum_{j=1}^{N_f} \prod_{k=1}^{d} \mathcal{F}_{(i,j,k)}(x_k)\,,$$

where $\mathcal{F}(x) : \mathbb{R}^d \to \mathbb{R}^{d \times N_f \times d}$ is a tensor-valued function.

Galerkin Approximation of the GHJB Equation

- Given $u^n(x)$, we solve the linear Generalized HJB equation

$$(f(x)+Bu^n)^T\nabla V(x)+\ell(x)+\|u^n\|^2=0\,.$$

- With $\{\phi_j(x)\}_{j=1}^{\infty}$ a complete set of d-dimensional polynomial basis functions, we approximate

$$V(x)\approx\sum_{j=1}^{N}c_j\phi_j(x)$$

- u^n $(n>0)$ is expressed in the form

$$u^n(x)=-\frac{1}{2\gamma}g^T\nabla V^n(x)=-\frac{1}{2\gamma}g^T\sum_{j=1}^{N}c_j^n\nabla\phi_j(x)\,.$$

- Every term expanded, leads to dense linear system for $V^{n+1}(x)$

$$A(c^n)c^{n+1}=b(c^n)\,.$$

The Ingredients of Policy Iteration

- Meshfree ! eg pseudo-spectral collocation based on Chebysheff polynomials
- separability of f
- Galerkin approximation of GHJB using globally supported polynomials (monomials, Legendre, ...)
- high dimensional integrals: introduce separable structure:

$$\phi_j(x) = \prod_{i=1}^{d} \phi_j^i(x_i) \qquad (\ldots M^d!)$$

- tensorize

Towards neural network based optimal feedback control

$$(P_\beta^{y_0}) \qquad \begin{cases} \inf_{y \in W_\infty,\, u \in L^2(I;\mathbb{R}^m)} \dfrac{1}{2} \displaystyle\int_I \left(|Dy(t)|^2 + \beta |u(t)|^2 \right) \, \mathrm{d}t \\ s.t. \quad \dot{y} = f(y) + Bu, \quad y(0) = y_0, \end{cases}$$

$$W_\infty = \{\, y \in L^2(I;\mathbb{R}^n) \mid \dot{y} \in L^2(I;\mathbb{R}^n) \,\},\ I = (0,\infty),\ B \in \mathbb{R}^{n\times m}.$$

our interest: optimal feedback stabilization

$$u^*(t) = F^*(y^*(t)) = -\tfrac{1}{\beta} B^\top \nabla V(y^*(t))$$

for all y_0 in a compact set $Y_0 \subset \mathbb{R}^n$ containing 0.

A.1 $Df\colon \mathbb{R}^n \to \mathbb{R}^{n\times n}$ is Lip. continuous on compacts, $f(0) = 0$.

A.2 There exists $F^*\colon \mathbb{R}^n \to \mathbb{R}^m$: the induced Nemitsky operator satisfies $\mathcal{F}^*\colon W_\infty \to L^2(I;\mathbb{R}^m)$. Further

$$\dot{y} = f(y) + B\mathcal{F}^*(y), \quad y(0) = y_0$$

admits a unique solution $y^*(y_0) \in W_\infty$, $\forall y_0 \in \mathbb{R}^n$, and

$$(y^*(y_0), \mathcal{F}^*(y^*(y_0))) \in \arg\min\,(P_\beta^{y_0}) \quad \forall y_0 \in \mathbb{R}^n.$$

A.3 $DF^*\colon \mathbb{R}^n \to \mathbb{R}^{n\times n}$ is Lip. continuous on compacts, $F^*(0) = 0$.

A.4 $\exists\, M_0$ and a bounded neighborhood $\mathcal{N}(Y_0) \subset \mathbb{R}^n$:

$$\mathbf{y}^*\colon \mathcal{N}(Y_0) \to W_\infty, \quad y_0 \mapsto \mathbf{y}^*(y_0)$$

is continuously differentiable and

$$\|\mathbf{y}^*(y_0)\|_{W_\infty} \leq M_0 \quad \forall y_0 \in \mathcal{N}(Y_0).$$

The learning problem

$$(\mathcal{P}_{y_0}) \quad \begin{cases} \min\limits_{\substack{\mathcal{F}\in\mathcal{H},\\ y\in W_\infty}} J(y,\mathcal{F}) = \frac{1}{2}\int_I \left(|Dy(t)|^2 + \beta|\mathcal{F}(y)(t)|^2\right)\,\mathrm{d}t \\ \dot{y} = f(y) + B\mathcal{F}(y), \quad y(0) = y_0, \end{cases}$$

where

$$\mathcal{H} = \left\{\mathcal{F}(y)(t) = F(y(t) : F \in \mathrm{Lip}\left(\bar{B}_{2M_0}(0);\mathbb{R}^m\right),\ F(0) = 0\right\}.$$

Yes, but ....
Bellman principle implies learn along: $S = \{y(\mathcal{F}^*(t)) : t \in I\}$.
or better

$$\begin{cases} \min\limits_{\mathcal{F}\in\mathcal{H},\,\mathbf{y}(y_0^i)\in W_\infty} \sum_{i=1}^{m} J(\mathbf{y}(y_0^i),\mathcal{F}) \\ s.t. \quad \dot{\mathbf{y}}(y_0^i) = f(\mathbf{y}(y_0^i)) + B\mathcal{F}(\mathbf{y}(y_0^i)), \quad y(0) = y_0^i, \end{cases}$$

The learning problem

$$(\mathcal{P}) \quad \begin{cases} \min\limits_{\substack{\mathcal{F}\in\mathcal{H},\\ \mathbf{y}\in L^\infty_\mu(Y_0;W_\infty)}} j(\mathbf{y},\mathcal{F}) = \int_{Y_0} J(\mathbf{y}(y_0)\mathcal{F}(\mathbf{y}(y_0)))\,\mathrm{d}\mu(y_0), \\ \dot{\mathbf{y}}(y_0) = f(\mathbf{y}(y_0)) + B\mathcal{F}(\mathbf{y}(y_0)), \quad \text{for } \mu\text{-a.e.} y_0 \in Y_0, \\ \|\mathbf{y}\|_{L^\infty_\mu(Y_0;W_\infty)} \leq 2M_0 \end{cases}$$

where $(Y_0, \mathcal{A}, \mu)$ is a complete probability space.

Proposition

$(\mathcal{P})$ admits a solution and we have equivalence to μ-a.e. solutions of $(\mathcal{P}_{y_0})$ on Y_0.

Corollary

Let $(\bar{\mathcal{F}}, \bar{\mathbf{y}})$ be an optimal solution to $(\mathcal{P})$ and assume that $\mathcal{A}$ contains the Borel σ-algebra on Y_0. If $\bar{\mathbf{y}} \in \mathcal{C}_b(\operatorname{supp}\mu; W_\infty)$, then

$$\widehat{Y}_0 := \{\, \bar{\mathbf{y}}(y_0)(t) \;\mid y_0 \in \operatorname{supp}\mu,\; t \in [0,+\infty) \,\} \subset \bar{B}_{2M_0}(0)$$

is compact and the previous proposition can be extended to $\widehat{Y}_0$.

Recap on neural networks

$$f_{i,\theta}(x) = \sigma(W_i x + b_i) \qquad \forall x \in \mathbb{R}^{N_{i-1}},\ i = 1, \dots, L-1$$
$$f_{L,\theta}(x) = W_L x + b_L \qquad \forall x \in \mathbb{R}^{N_{L-1}}$$

$\sigma \in C^1(\mathbb{R}, \mathbb{R})$ activation function

$$\theta = (W_1, b_1, \dots, W_L, b_L)$$
$$\mathcal{R} = \bigtimes_{i=1}^{L} \left(\mathbb{R}^{N_i \times N_{i-1}} \times \mathbb{R}^{N_i} \right),$$

which is uniquely determined by its *architecture*

$$\text{arch}(\mathcal{R}) = (N_0, N_1, \dots, N_L) \in \mathbb{N}^{L+1},$$

$$f_{L,\theta} \circ f_{L-1,\theta} \circ \cdots \circ f_{1,\theta}(x)$$

$$F_\theta(x) = f_{L,\theta} \circ f_{L-1,\theta} \circ \cdots \circ f_{1,\theta}(x) - f_{L,\theta} \circ f_{L-1,\theta} \circ \cdots \circ f_{1,\theta}(0) \quad \forall x \in \mathbb{R}^n$$

Recap on neural networks

$$f_{i,\theta}(x) = \sigma(W_i x + b_i) + x \qquad \forall x \in \mathbb{R}^{N_{i-1}}, \; i = 1, \dots, L-1$$
$$f_{L,\theta}(x) = W_L x + b_L \qquad \forall x \in \mathbb{R}^{N_{L-1}}$$

$\sigma \in C^1(\mathbb{R}, \mathbb{R})$ activation function

$$\theta = (W_1, b_1, \dots, W_L, b_L)$$
$$\mathcal{R} = \bigtimes_{i=1}^{L} \left(\mathbb{R}^{N_i \times N_{i-1}} \times \mathbb{R}^{N_i} \right),$$

which is uniquely determined by its *architecture*

$$\mathrm{arch}(\mathcal{R}) = (N_0, N_1, \dots, N_L) \in \mathbb{N}^{L+1},$$
$$f_{L,\theta} \circ f_{L-1,\theta} \circ \cdots \circ f_{1,\theta}(x)$$

$$F_\theta(x) = f_{L,\theta} \circ f_{L-1,\theta} \circ \cdots \circ f_{1,\theta}(x) - f_{L,\theta} \circ f_{L-1,\theta} \circ \cdots \circ f_{1,\theta}(0) \quad \forall x \in \mathbb{R}^n$$

Approximation by neural networks

Theorem

Let $\eta_1 > 0, \eta_2 > 0$, *and assume that the activation function* σ *is not a polynomial. Then for each* $\epsilon > 0$ *there exist* $L_\varepsilon \in \mathbb{N}$, $\mathrm{arch}(\mathcal{R}_\varepsilon) \in \mathbb{N}^{L_\varepsilon+1}$ *and a neural network*

$$\theta_\varepsilon = (W_1^\varepsilon, b_1^\varepsilon, \ldots, W_{L_\varepsilon}^\varepsilon, b_{L_\varepsilon}^\varepsilon) \in \mathcal{R}_\varepsilon$$

such that $\|W_1^\varepsilon\|_\infty \leq \eta_1$, $|b_i^\varepsilon|_\infty \leq \eta_2$, $i = 1, \ldots, L_\varepsilon$, *as well as*

$$|F^*(x) - F_{\theta_\varepsilon}(x)| + \|DF^*(x) - DF_{\theta_\varepsilon}(x)\| \leq \varepsilon$$

for all $|x| \leq 2M_0$.

Thus, approximate $\mathcal{F}$ by $\mathcal{F}_{\theta_\varepsilon}$!

Approximation by neural networks:cont

Assumption

$\exists\ C > 0$, such that for all $y_0 \in \mathcal{N}(Y_0)$, $\delta y_0 \in \mathbb{R}^n$ and $\delta v \in L^2(I;\mathbb{R}^n)$ there exists a unique $\delta y \in W_\infty$:

$$\dot{\delta y} = Df(\mathbf{y}^*(y_0))\delta y + BD\mathcal{F}^*(\mathbf{y}^*(y_0))\delta y + \delta v,\ \delta y(0) = \delta y_0$$

$$\|\delta y\|_{W_\infty} \le C(\|\delta v\|_{L^2(I;\mathbb{R}^n)} + |\delta y_0|).$$

Theorem

There exist $\varepsilon_1 > 0$, c such that for $\varepsilon \in (0, \varepsilon_1)$, $y_0 \in Y_0$

$$\dot{y}_\varepsilon = \mathbf{f}(y_\epsilon) + \mathcal{B}\mathcal{F}^\sigma_{\theta_\epsilon}(y_\epsilon), \quad y_\epsilon(0) = y_0$$

admits a unique solution $y_\varepsilon = \mathbf{y}_\varepsilon(y_0) \in \mathcal{Y}_{ad}$ and $\|\mathbf{y}^*(y_0) - \mathbf{y}_\varepsilon(y_0)\|_{W_\infty} \le c\varepsilon$. Moreover $\mathbf{y}_\varepsilon \in L^\infty_\mu(Y_0; W_\infty)$ with $\|\mathbf{y}_\varepsilon\|_{L^\infty_\mu(Y_0;W_\infty)} \le \frac{3}{2}M_0$.

Optimal neural network feedback law

$$(\mathcal{P}_{\mathcal{R}_\varepsilon}) \quad \begin{cases} \min\limits_{\substack{\theta\in\mathcal{R}_{ad,\varepsilon} \\ \mathbf{y}\in L^\infty_\mu(Y_0;W_\infty)}} j(\mathbf{y},\mathcal{F}_\theta)+\mathcal{G}_{\mathcal{R}_\varepsilon}(\theta), \\ \dot{\mathbf{y}}(y_0)=f(\mathbf{y}(y_0))+B\mathcal{F}(\mathbf{y}(y_0)), \quad \text{for } \mu\text{-a.e.} y_0\in Y_0, \\ \|\mathbf{y}\|_{L^\infty_\mu(Y_0;W_\infty)}\leq 2M_0 \end{cases}$$

$$\mathcal{R}_{ad,\varepsilon}=\{\theta=(W_1,b_1,..,W_{L_\varepsilon},b_{L_\varepsilon}):\|W_1\|_\infty\leq\eta_1,|b_i|_\infty\leq\eta_2,i=1,..,L_\varepsilon\}\subset\mathcal{R}_\varepsilon$$

$$\mathcal{G}_{\mathcal{R}_\varepsilon}(\theta)=I_{\mathcal{R}_{ad,\varepsilon}}(\theta)+\alpha_{\mathcal{R}_\varepsilon}\sum_{i=2}^{L_\varepsilon}\|W_i\|^2$$

Theorem

There exists $\varepsilon_1>0$ such that $(\mathcal{P}_{\mathcal{R}_\varepsilon})$ admits a global minimizer $(\theta^*_\varepsilon,\mathbf{y}^*_\varepsilon)\in\mathcal{R}_{ad,\varepsilon}\times L^\infty_\mu(Y_0;W_\infty)$ for every $0<\varepsilon\leq\varepsilon_1$.

Convergence

Theorem

$$\text{If} \qquad 0 < \alpha_{\mathcal{R}_\varepsilon} \leq \frac{\varepsilon}{2\sum_{i=2}^{L_\varepsilon} \|W_i^\varepsilon\|^2}, \quad 0 < \varepsilon < \varepsilon_2, \text{ then}$$

$$0 \leq j(\mathbf{y}_{\theta_\varepsilon^*}, \mathcal{F}_{\theta_\varepsilon^*}^\sigma) + \mathcal{G}_{\mathcal{R}_\varepsilon}(\theta_\varepsilon^*) - j(\mathbf{y}^*, \mathcal{F}^*) \leq c\varepsilon$$

for some constant $c > 0$ *independent of* ε. *In particular*

$$j(\mathbf{y}_{\theta_\varepsilon^*}, \mathcal{F}_{\theta_\varepsilon^*}) \to j(\mathbf{y}^*, \mathcal{F}^*) \text{ as } \varepsilon \to 0.$$

Theorem

Each weak accumulation point $(\widehat{\mathbf{y}}, \widehat{\mathbf{u}})$ *of* $\{\, (\mathbf{y}_{\theta_\varepsilon^*}, \mathcal{F}_{\theta_\varepsilon^*}(\mathbf{y}_{\theta_\varepsilon^*}))\,\}$ *in* $L_\mu^2(Y_0; W_\infty) \times L_\mu^2(Y_0; L^2(I; \mathbb{R}^m))$ *fulfills* $\|\widehat{\mathbf{y}}\|_{L_\mu^\infty(Y_0;W_\infty)} \leq 2M_{Y_0}$ *and*

$$\dot{\widehat{\mathbf{y}}}(y_0) = \mathbf{f}(\widehat{\mathbf{y}}(y_0)) + B\widehat{\mathbf{u}}(y_0), \widehat{\mathbf{y}}(y_0)(0) = y_0,$$
$$(\widehat{\mathbf{y}}(y_0), \widehat{\mathbf{u}}(y_0)) \in \arg\min\,(P_\beta^{y_0})$$

for μ*-a.e.* $y_0 \in Y_0$. *If* $D > 0$ *the convergence is strong.*

NN optimal feedback control for a Van der Pol oscillator

Consider

$$\begin{pmatrix} \dot{y}_1 \\ \dot{y}_2 \end{pmatrix} = \begin{pmatrix} y_2 \\ 1.5(1 - y_1^2)y_2 - y_1 + y_1^3 + u \end{pmatrix}$$

- Finite set of training initial conditions

$$\{y_0^i\}_{i=1}^5 = \{\, \pm(1,0), \pm(0,1), (6,4)\,\}\,,\ \mu = \frac{1}{5}\sum_{i=1}^{5} \delta_{y_0^i}$$

- Parameters in objective functional:

$$D = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\,,\ \beta = 10^{-3},\ \alpha_\varepsilon = 0.$$

Numerical realization

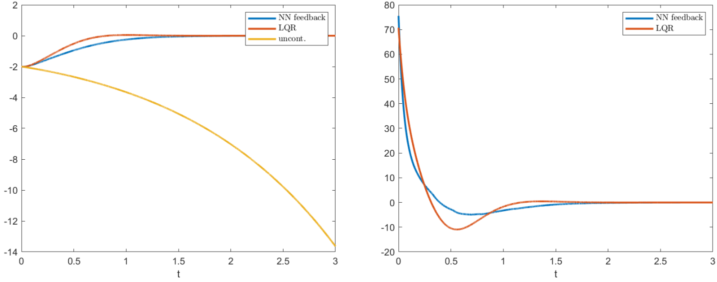
- Replace infinite time horizon by $T > 0$ sufficiently large.
- Fix $L_\varepsilon = 8$, $N_i = 2$ and $\sigma(x) = \max\{|x|^{0.1}x, 0\}$.
- Add residual connections:

$$f_{i,\theta}(x) = \sigma(W_i x + b_i) + x$$

- Assumption: Constraints are inactive.
- Solved by Barzilai-Borwein.

Validation for $y_0 = (-2, 1)$, $T = 3$

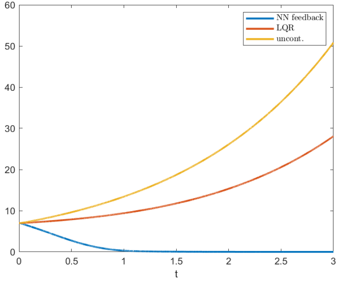
$$J(\mathbf{y}_{LQR}(y_0), \mathcal{F}_{LQR}) = 0.686, \quad J(\mathbf{y}_{\theta^*_\varepsilon}(y_0), \mathcal{F}_{\theta^*_\varepsilon}) = 0.815$$



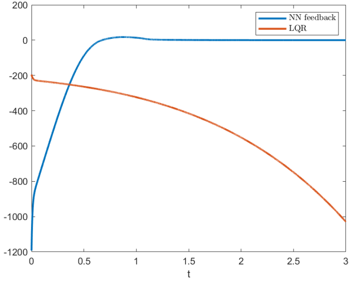
(a) States (b) Controls

Validation for $y_0 = (7, -2),\ T = 3$

$$J(\mathbf{y}_{LQR}(y_0), \mathcal{F}_{LQR}) = 759.112, \quad J(\mathbf{y}_{\theta_\varepsilon^*}(y_0), \mathcal{F}_{\theta_\varepsilon^*}) = 76.185$$



(a) States



(b) Controls