Optimal Feedback Control with Neural Networks

A compositionality-based approach via Hamilton-Jacobi-Bellman PDEs

Optimal Control Problems

The project deals with **nonlinear infinite horizon optimal control problems** of the form

$$\text{minimize} \quad J(x_0, u) = \int_0^\infty e^{-\delta t} \ell(x(t), u(t))dt, \qquad (1)$$

$$\text{such that} \quad \dot{x}(t) = f(x(t), u(t)), \quad x(0) = x_0, \qquad (2)$$

where

- $x\colon \mathbb{R}_{\geq 0} \to \mathbb{R}^n$ is the state at time t with initial value $x(0) = x_0$
- $u\colon \mathbb{R}_{\geq 0} \to U \subset \mathbb{R}^m$ is the control variable
- $f\colon \mathbb{R}^n \times U \to \mathbb{R}^n$ is a controlled vector field that determines the dynamics of x in dependence of u via the differential equation (2)
- $\ell\colon \mathbb{R}^n \times U \to \mathbb{R}$ is the cost function and $\delta \in [0, \infty)$ is the discount rate

A control function $u^\star$ is called optimal for an initial value x_0 if it minimizes the costs in (1), i.e., $J(x_0, u^\star) = V(x_0)$ for the optimal value function

$$V(x_0) := \inf_u J(x_0, u).$$

It is desirable to obtain the optimal control in feedback form, i.e.,

$$u^\star(t) = F(x^\star(t))$$

for a feedback law $F\colon \mathbb{R}^n \to U$ and the corresponding optimal trajectory $x^\star$. If an (approximate) optimal value function V is known, an (approximate) optimal feedback law F can be computed from V.

Task: Compute the optimal value function V efficiently.

Curse of Dimensionality for the HJB-PDE

The optimal value function is the unique (viscosity) solution of the Hamilton-Jacobi-Bellman equation

$$\delta V(x) + \sup_{u \in U}\{-DV(x) f(x,u) - \ell(x,u)\} = 0,$$

a partial differential equation (PDE) in $\mathbb{R}^n$. Solving this PDE numerically is subject to the **curse of dimensionality**, i.e., in general the numerical **effort grows exponentially in the state dimension n**.

Deep neural networks can compute so-called compositional functions without suffering from the curse of dimensionality. One particular example for compositional functions are **separable functions** of the form

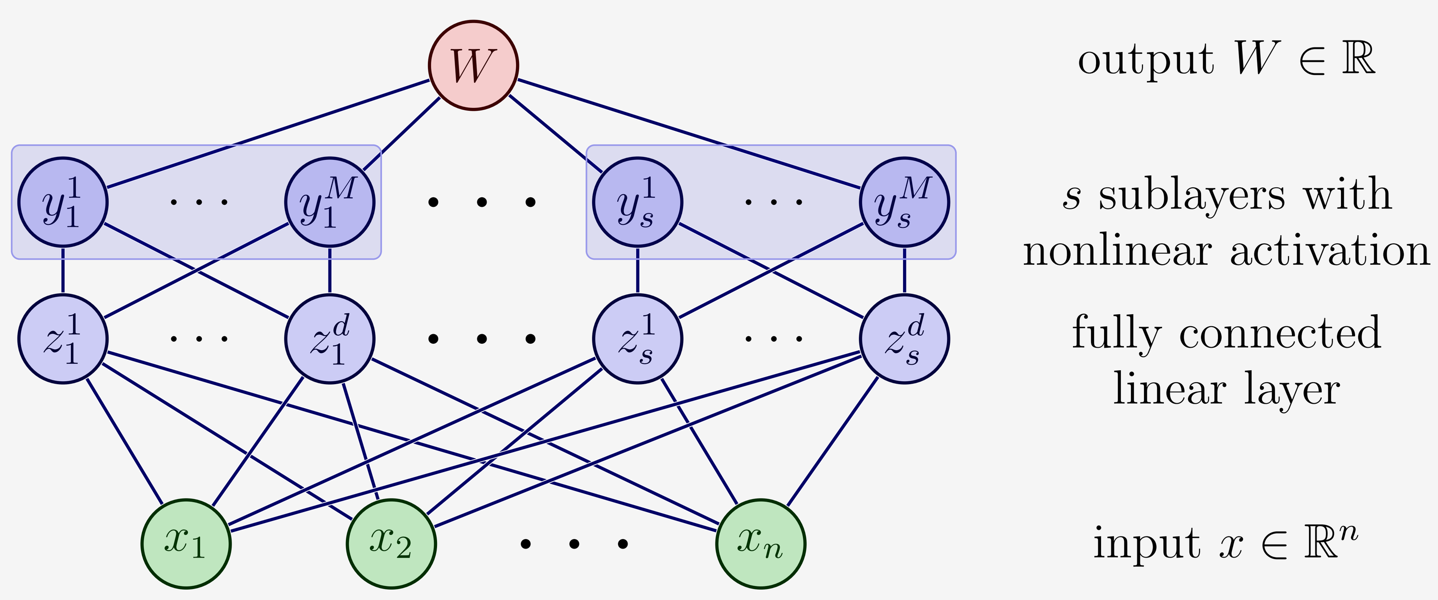
$$V(x) = \sum_{j=1}^{s} V_j(z_j),$$

where z_j are lower-dimensional components of x.

Project Goals

- Explanation of the ability of deep neural networks for the curse-of-dimensionality-free solution of high-dimensional HJB-PDEs
- Identification of structural conditions for optimal control problems allowing for compositional (approximate) optimal value functions and possibly also compositional optimal feedback laws
- Construction of neural network architectures and training algorithms for an efficient approximation

Network Architecture for Separable Functions



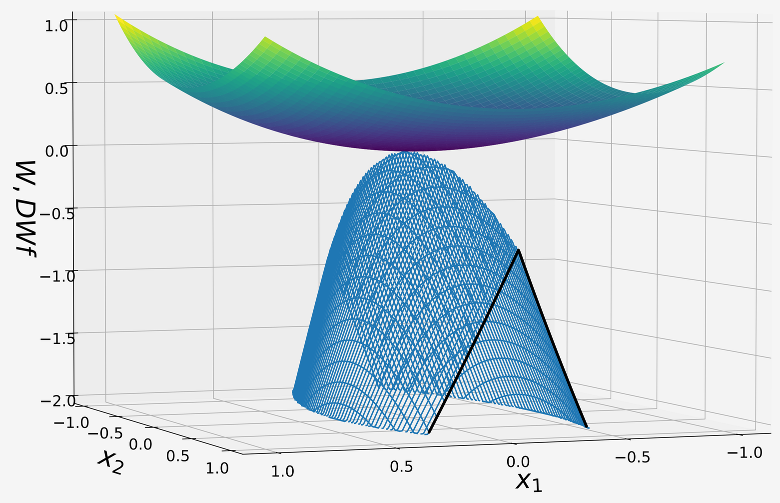
- In the training process, we iteratively evaluate the network at randomly chosen points $x_1, \ldots, x_N$ of the state space, measure the deviation from the desired function V with an appropriate loss function, and update the weights of the network accordingly

Stabilization Task: Control Lyapunov Functions

- Objective: control the trajectory to a desired set point and keep it there $\rightsquigarrow$ Compute a **control Lyapunov function** V
- Consider an interconnected control system represented as graph: one node for each subsystem with an edge between two nodes if they interact, i.e., influence their dynamics
- Assume **input-to-state stability** for each subsystem
- Assume that each cycle of the graph contains at least one node with a high level of controllability, called **active node**
- Result: existence of a **separable** control Lyapunov function $V = \sum_{j=1}^{s} V_j$

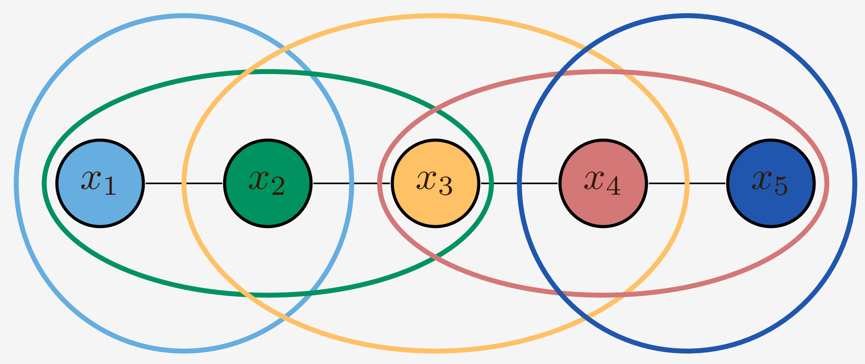
Numerical test case:

$$\dot{x}_1 = x_{10} + u,$$
$$\dot{x}_2 = x_1 - x_2 + x_1^2,$$
$$\dot{x}_j = x_{j-1} - x_j, \quad 3 \leq j \leq 10.$$



Separability via Decaying Sensitivity

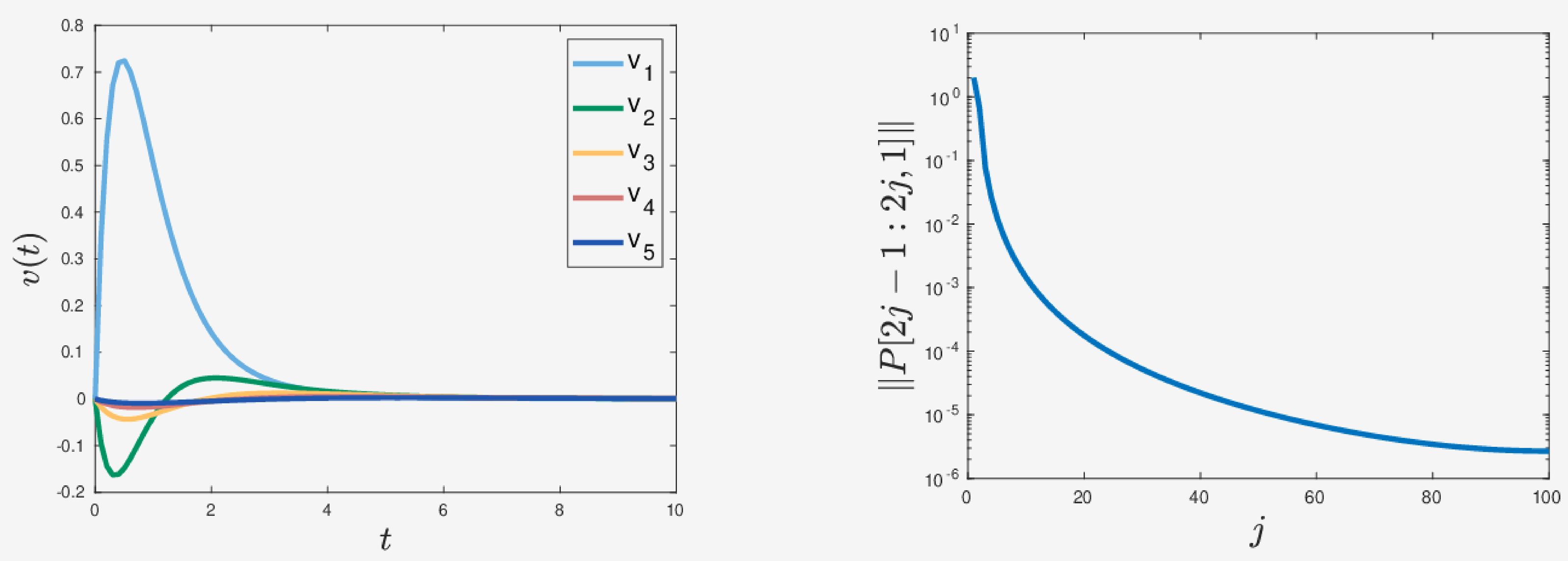
- Setting: interconnected control system represented by its graph
- Assumption: the influence of the node x_k on
$$V(x_1, \ldots, x_{j-1}, x_j, x_{j+1}, \ldots, x_s) - V(x_1, \ldots, x_{j-1}, 0, x_{j+1}, \ldots, x_s)$$
decays to 0 with an increasing graph-distance of the nodes x_k and x_j
- Construction: **overlapping decomposition** based on neighborhoods in the graph
- Result: separable approximation of $V(\cdot) \approx \sum_{j=1}^{s} \Psi_j(\cdot) + V(0)$



Example: Convoy of Vehicles

- Control objective: the first vehicle should follow some reference trajectory and all other vehicles should maintain a fixed distance L

- Decaying sensitivity yields that a perturbation in any vehicle will **decrease quickly** if the vehicles are **controlled optimally**
- Numerical simulation with 100 vehicles, dynamics $\dot{x}_i = v_i, \dot{v}_i = u_i$ and costs given as $\underbrace{(x_1 - x_{\text{ref}})^2}_{\text{reference}} + \sum_{j=1}^{99} \underbrace{(x_{j+1} - x_j - L)^2}_{\text{desired distance}} + \underbrace{r(v,u)}_{\text{regularization}}$



Velocities for the first 5 vehicles and decay in the optimal value function $V(x) = x^T P x$

References

- M. Sperl, L. Saluzzi, L. Grüne, and D. Kalise. **Separable approximations of optimal value functions under a decaying sensitivity assumption**. 62nd IEEE Conference on Decision and Control (CDC) Singapore. 259-264, 2023.
- M. Sperl, J. Mysliwitz, and L. Grüne. **Approximation of separable control Lyapunov functions with neural networks**. preprint, ePub Bayreuth, 2023.
- S. Shin, Y. Lin, G. Qu, A. Wiermann, and M. Anitescu. **Near-optimal distributed linear-quadratic regulator for networked systems**. SIAM Journal on Control and Optimization. 61(3):1113-1135, 2023.

Lars Grüne, Mario Sperl

lars.gruene@uni-bayreuth.de, mario.sperl@uni-bayreuth.de

Chair of Applied Mathematics, Institute of Mathematics

UNIVERSITÄT BAYREUTH

DFG