Non-Convex Model Predictive Control Fast Inference using Context Sensitive Neural Networks

Lopez-Rojas Arturo D.* Dean-Leon Emmanuel* Ramirez-Amaro Karinne*

Neural networks (NNs) are increasingly used as real-time approximators of solvers in Model Predictive Control (MPC) implementations, offering significant computational advantages over traditional solvers, an approach referred to as NN-Based Optimization (NNBO). The primary advantages of this learning-based approach include reduced execution time, adaptability through fine-tuning, and portability across computational platforms [1],[2]. However, existing NNBO methods are largely limited to convex or quadratic problems. In non-convex problems (e.g., trajectory planning), the feedforward NNs used in NNBO often yield infeasible or suboptimal solutions. This limitation stems from the multi-valued nature of solution mappings in non-convex problems, where multiple local optima exist.

In this work, we are proposing a learning framework to enable NNBO in non-convex problems by formulating a set-valued mapping of the MPC's optimization problem as a multimodal learning problem. The proposed framework is organized into two stages: training and deployment. In the training stage, a network composed of two subnetworks, as shown in Figure 1, where the first subnetwork categorizes the relationship between the optimization problem parameters and their corresponding solutions, encoding this information as a one-hot vector (referred to as mode selection). The second subnetwork takes both the optimization parameters and the selected mode as inputs and learns to approximate the corresponding solution. In the second stage (deployment), the second subnetwork is used to infer solutions for a predefined number of modes. These candidate solutions are then evaluated using the original optimization problem, and the best one is selected.

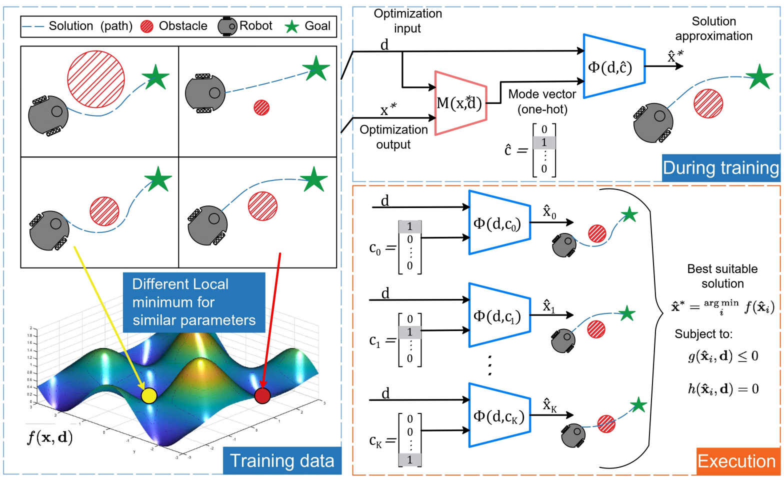

Figure 1: Proposed set-valued mapping for neural network inference in MPC problems with multiple local minima, where $\mathbf{d}$ is a vector with the MPC's optimization problem parameters and $\hat{\mathrm{x}}^*$ the predicted solution.

References

[1] Lopez-Rojas, A.D. and Cruz-Villar, C.A. (2024). Neural networks as an approximator for a family of optimization algorithm solutions for online applications. *Neural Computing and Applications.*, 36(6), 3125–3140.

[2] Castillo, A., Pryor, E., El Fathi, A., Kovatchev, B., and Breton, M. (2025). Neural networks for on-chip model predictive control: A method to build optimized training datasets and its application to type-1 diabetes. *IEEE Transactions on Cybernetics*.

*Faculty of Electrical Engineering, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden (e-mail: arturod@chalmers.se; deane@chalmers.se; karinne@chalmers.se).