

# A Safe Reinforcement Learning driven Weights-varying Model Predictive Control for Autonomous Vehicle Motion Control

Baha Zarrouki, Marios Spanakakis, Johannes Betz

**Abstract**—Determining the optimal cost function parameters of Model Predictive Control (MPC) to optimize multiple control objectives is a challenging and time-consuming task. Multiobjective Bayesian Optimization (BO) techniques solve this problem by determining a Pareto optimal parameter set for an MPC with static weights. However, a single parameter set may not deliver the most optimal closed-loop control performance when the context of the MPC operating conditions changes during its operation, urging the need to adapt the cost function weights at runtime. Deep Reinforcement Learning (RL) algorithms can automatically learn context-dependent optimal parameter sets and dynamically adapt for a Weightsvarying MPC (WMPC). However, learning cost function weights from scratch in a continuous action space may lead to unsafe operating states. To solve this, we propose a novel approach limiting the RL actions within a safe learning space representing a catalog of pre-optimized BO Pareto-optimal weight sets. We conceive a RL agent not to learn in a continuous space but to proactively anticipate upcoming control tasks and to choose the most optimal discrete actions, each corresponding to a single set of Pareto optimal weights, context-dependent. Hence, even an untrained RL agent guarantees a safe and optimal performance. Experimental results demonstrate that an untrained RL-WMPC shows Pareto-optimal closed-loop behavior and training the RL-WMPC helps exhibit a performance beyond the Pareto-front.

## 1. Contextual Challenges Introduction

In autonomous driving and complex system control, nonlinear Model Predictive Control (MPC) is a commonly used method. However, designing an effective MPC system poses several challenges, especially when setting the loss function. Here are some difficulties related to this:

### 1.1. Manual Tuning and Expert Knowledge Requirement

Setting the loss function for MPC requires extensive manual tuning and expert knowledge. This process is time-consuming and complex, often leading to suboptimal solutions, which can affect the overall performance of the system.

### 1.2. Impact of Weights on Safety Performance

The choice of weights in the loss function is crucial for safety performance. Inappropriate weights can lead to catastrophic accidents, especially in high-risk applications like autonomous driving.

### 1.3. Specificity to Operating Points

Even if suitable loss function weights are determined, they are typically only effective for specific operating points. Once operating conditions change, the reliability and effectiveness of these weights may significantly decrease. In summary, designing an MPC system requires significant effort in setting the loss function to ensure its safety and effectiveness under different operating conditions.

## 2. Related Works

In the field of Model Predictive Control (MPC), previous research has primarily focused on weight tuning to enhance system performance.

### 2.1. Traditional Approaches

In traditional approaches, Genetic Algorithm (GA) methods enhance closed-loop control performance by handling constraints, while Cooperative Particle Swarm Optimization (PSO) methods focus on improving trajectory tracking performance, widely applied in Unmanned Aerial Vehicle (UAV) domains. These methods optimize weights to enhance system stability and accuracy. However, they often require significant computational resources and are sensitive to initial parameter settings, which may lead to slow convergence or being trapped in local optima.

### 2.2. Bayesian Optimization Approaches

Bayesian Optimization (BO) aims to address the mismatch between real systems and models. Due to model simplifications, environmental changes, and parameter preset values not aligning with actual conditions, there is often a significant gap between the model and real-world conditions. BO captures uncertainties by introducing Gaussian processes and dynamically updates in real-time, helping to overcome these adverse factors and ensuring model reliability in different environments. However, BO methods are highly data-dependent and may face high computational complexity in high-dimensional spaces.

### 2.3. Reinforcement Learning Approaches

Additionally, the introduction of Reinforcement Learning (RL) has further advanced MPC development. By determining the prediction horizon of MPC through RL, system complexity can be reduced. Furthermore, RL is used to learn MPC's meta-parameters, further reducing complexity. Some studies have proposed variable-weight nonlinear MPC, which can adjust the weight matrix of the cost function online, but safety is not guaranteed. There are also approaches using RL to learn control law parameters to determine controller outputs in real-time, but safety assurance remains insufficient.

**Summary:** Although these methods have made progress in performance optimization, the issue of determining optimal cost function weights based on safety has not yet been resolved.

## 3. Proposed Work

This paper conceptualizes a weight-variable Model Predictive Control (MPC) driven by a safety reinforcement learning agent. The reinforcement learning agent autonomously adjusts the MPC cost function weights under different control conditions through foresight design, selecting the most suitable weight set from the Pareto front optimized by Multi-Objective Bayesian Optimization (MOBO). This approach allows for flexible weight adjustments in dynamic environments, ensuring optimal performance of the control system.

In experiments, this method demonstrated its adaptability in dynamically adjusting nonlinear MPC weights, particularly in simulating the control of a full-scale autonomous vehicle to follow an optimal track. Safety evidence provided by the experiments shows that even with untrained agents, this method can exhibit Pareto optimality, proving its effectiveness and safety in real-world applications.

Additionally, the paper conducts a context-related analysis of the decision-making process of the reinforcement learning agent. By testing in unseen environments, the method's generalization and robustness capabilities were evaluated, exploring the impact of continuous learning in unfamiliar settings. This research provides new insights into the application of reinforcement learning in complex control systems and demonstrates its potential to maintain performance and safety across different environments.

## 4. My Perspective on This Paper

This article cleverly combines the strengths of Bayesian optimization and reinforcement learning to propose an innovative approach. In summarizing previous work, the article accurately points out the widespread application and distinct shortcomings of Bayesian optimization and reinforcement learning in this field. By deploying reinforcement learning agents on the results of multi-objective Bayesian optimization, the authors successfully leverage the strengths and construct a more effective model. This combination strategy not

only demonstrates unique innovation but also helps readers naturally understand the advantages of the method, showcasing a writing technique worth learning.

However, the article falls short in terms of visual presentation. Compared to other articles of the same level, some charts lack detailed data support, providing only intuitive and rough data displays. This is an area for improvement, as adding more detailed data would enhance the article's persuasiveness and scientific rigor. Overall, despite areas for improvement, the article has significant highlights in methodological innovation and writing technique.