Corrections of Typos and Citation in My PhD Thesis
Learning Robot Control Via Bilevel Optimization

Dear Readers,

There is a note on DiffMPC Citation and KKT differentiation.

In Section 2.3.1 of my thesis, I cited DiffMPC [Amos et al., NeurIPS 2018] as requiring inversion of a large KKT matrix and incurring $O(N^2)$ computational complexity. This statement was influenced by a similar claim made in the PDP framework [Jin et al., NeurIPS 2020]. However, after revisiting the original DiffMPC work, I would like to clarify that this characterization is technically inaccurate.

While generic implicit differentiation of optimization problems through KKT conditions may involve large matrix inversion, DiffMPC avoids this by leveraging the time-recursive structure of LQR. It applies Riccati-based fixed-point differentiation, which yields linear-time complexity in the time horizon N when computing gradients of scalar loss functions (i.e., vector-Jacobian products). No explicit inversion of a full KKT matrix is required.

Therefore, it would be more accurate to classify DiffMPC and its variants as vector-Jacobian product-based methods, a possible amendment that could be reflected in Table 2.1.

This clarification does not affect the design or contributions of NeuroMHE, which relies on a different recursive estimation structure, but I appreciate the feedback from the research community and will ensure this is correctly addressed in future work.

In addition, there are three minor typos related to notation and labelling in Chapter 5:

1. On the top of Page 107, in the sentence "…between each quadrotor and the load, represented by the gradients $\partial u_{0|t}^{*,i}/\partial u_{0|t}^{*,l}$, $\partial u_{0|t}^{*,l}/\partial x_t^i$, $\partial u_{0|t}^{*,l}/\partial x_t^i$, $\mathbf{X}_{l,t}^i$, $\mathbf{X}_{i,t}^l$", the term $\partial u_{0|t}^{*,l}/\partial x_t^i$ was carelessly repeated twice. The first instance should be $\partial u_{0|t}^{*,i}/\partial x_t^l$, as indicated in Eq. (5.18).

2. On Page 124, the same typo (mentioned in point 1) appears in the note below Table 5.2.

3. In Line 18 of Algorithm 7 on Page 118, the sensitivities were carelessly labelled as $\overline{\mathbf{Y}}_t$; it should be $\overline{\mathbf{X}}_t$, as indicated in Algorithm 6.

I apologize for any confusion caused and thank you for your understanding and patience regarding these minor corrections.

王秉亨

WANG BINGHENG

07 May 2025