Max Zhao

IRIS-HEP Project Proposal - Efficient implementation of algorithms to reconstruct charged particles trajectories

The efficacy of machine learning algorithms in the track reconstruction process is especially salient considering the dramatic developments of GPU hardware capabilities in recent years. Now, machine learning algorithms can be implemented efficiently which is important for expected increases in data flow for experiments like ATLAS. The long term goal of this project is to achieve higher performance in track reconstruction with a machine learning algorithm that embeds the Kalman filter. Ideally, this would achieve higher performance through learning with the Kalman filter as prior knowledge, something that other machine learning approaches to tracking like graph neural networks do not have.

Track reconstruction necessarily involves parsing through data with noise. To that extent, Kalman filtering is applied to various tasks for its uncertainty reducing properties. Some examples include the combinatorial Kalman filter used for track finding in ATLAS IBL [1] and the Kalman filter responsible for track fitting in ACTS [2]. There are however some inherent limitations to the traditional deterministic Kalman filters. For one, the Kalman filter only has single step memory, limiting the predictive value of prior measurements. Additionally, while the traditional Kalman filter is provably optimal for linear systems, it does not support nonlinear transformation [3]. There are variations of the Kalman filter that effectively handle nonlinear systems, but they often either make assumptions of local behavior like the Extended Kalman filter or incur higher computational costs like the Non-linear Kalman filter [4].

There are several machine learning architectures that could possibly meet these challenges. Certain architectures have already shown promise in matching and improving the performance of traditional Kalman filters in other applications. For example, recurrent neural networks (RNN) and even basic multilayer perceptrons have shown that they can learn behaviors of Kalman filters when applied to time-dependent medical data [5]. RNNs especially are especially promising candidates that we will be investigating in this project due to being optimized for sequential data. For example, long short-term memory networks, a type of RNN, can retain memory of past nodes for arbitrary amounts of cycles which are managed by an internal network. There are also many other interesting approaches that may involve convolutional neural networks like the Backpropagation Kalman filter [6].

Given the magnitude of the goals, the specific project for this summer term will be to lay the groundwork and conduct early stage investigation of the design. We will identify the machine learning architecture that suits the problem from the aforementioned architectures or variations thereof. From there, we will need to develop methods to bias the neural network such that it emulates the desired properties of the Kalman filter. This will allow us to train the network on simulated data to hopefully improve performance over the normal Kalman filter. The objective is to demonstrate the efficacy of this approach to fill the niche of Kalman filtering in track reconstruction. As such, the primary software deliverable will be a small-scale network that can become relative in performance to the existing algorithms. If this can be achieved, it will serve as a starting point for the development of a working algorithm that can contribute directly to the track reconstruction process. I will be working on this project under the supervision and mentorship of Prof. Heather Gray (UC Berkeley, LBNL) and Johannes Wagner (UC Berkeley, LBNL).

Timeline

- Weeks 1-2:
    - Review relevant literature concerning applicable neural network architectures
    - Review code of current implementations of Kalman filtering and machine learning tracking algorithms.
- Week 3
    - Gather a set of simulated data to generate an appropriate training set for the network.
- Weeks 4-5
    - Create a list of candidate possible neural network architectures to employ with preliminary testing for stability.
- Weeks 6-9
    - Create a framework for assessing performance based on existing algorithms.
    - Test different network designs with varying parameters.
- Weeks 10-11
    - Finalize and assess results. Prepare the final presentation. Determine appropriate future avenues of research.

References

[1] Gjersdal, H., et al. “Straight line track reconstruction for the ATLAS IBL testbeam with the EUDET telescope.” *ATLAS Note*, 2014.
[2] Ai, Xiaocong, et al. “A GPU-Based Kalman Filter for Track Fitting” *Computing and Software for Big Science,* vol. 5, 2021.
[3] Pei, Yan, et al. “An Elementary Introduction to Kalman Filtering” *arXiv preprint arXiv:1710.04055*, 2019.
[4] Ai, Xiaocong, et al. “A Non-Linear Kalman Filter for track parameters estimation in High Energy Physics” *arXiv preprint arXiv:2112.09470v1*, 2021.
[5] Krishnan, R., et al. “Deep Kalman Filters” *arXiv preprint arXiv:1511.05121*, 2015.
[6] Haarnoja, Tuomas, et al. “Backprop KF: Learning Discriminative Deterministic State Estimators” *arXiv preprint arXiv:1605.07148v4*, 2017.