ML Track Fitting in Nuclear Physics

Thomas Britton
CHEP 2019

Special thanks to:
David Lawrence and Gagik Gavalian
- Located in Newport News
- Continuous Electron Beam accelerator Facility (CEBAF)
High Energy vs Nuclear Physics

- Nuclear Physics (NP) experiments tend to have far fewer number of tracks
  - NP tracks have more “curvature”
- NP has less uniform magnetic fields
A JLab Joint Effort

- To study the application of ML to the problems of tracking in NP both CLAS and GlueX launched a joint effort
  - Slightly different problems/approaches but methods and results are shared
GlueX Current State

- **94.4%** of total reconstruction time (CPU) spent on tracking
  - Only ~2% spent on finding tracks
  - 22.5% spent on wire based tracking
  - **69.9%** spent on time based tracking

- Event Rate 162.2 Hz /node
  - ~3.75 Hz /core

- Per track
  - Candidate: ~1.2 ms
  - Wire Based: ~3.3 ms
  - Time Based: ~10.3 ms

- Track finding is **not** the big contributor. **Track fitting is!**
GlueX Goals

• Desire to “surgically” replace existing methods
  - Looking for performance gains
    • Emphasis on shallow learning

• Desire to shift needs with changing hardware landscape
  - Exa-scale machines majority GPU based. As GlueX scales up (higher luminosity) it may not have access to a scaling supply of CPU cycles

• Focus first on the **forward drift chamber**
A Toy Problem

- Can we get a $z$ and $\theta$ from tracks in a toy detector?
  - 6 chambers of 6 layers (36 layers total)
  - 100 wires per layer
  - 50cm between chambers and 1cm between wires
  - Wire hit efficiency 90% (grey scale for distance to track)
  - One hit per layer only

This is similar to the forward drift chamber
Results from the Toy Problem

- Used Keras and Tensorflow

- Treat like a classification
  - Take a weighted average of the results

- Why is the network doing much better than estimated uncertainty? (hint: it isn’t over-training)
Crystal Lattice Analogy

- Previously “optimal” resolution was calculated based only on the wire geometries. Just like a crystal angle matters! Resolution is angle dependent with some angles having better resolutions than others.

- Know your data!

\[ \phi_{\text{truth}} - \phi_{\text{ML}} \text{ vs. } \phi_{\text{ML}} \]

*model faithfully reproduces the angular dependence on resolution*
Future GlueX Work

• Begin looking at actual data
  – First focus on the state vector
• Look for optimal models
  – Low inference times with optimal errors
• Adapt the ML methods to other portions of code
Tracks are found through an iterative process

At higher luminosity
More hits =>
More segments =>
More combinations

This forces many more iterations

e.g. Four track combinations considered even though there are none
Training Data

• VGG16 model (Convolutional)

~ 3ms inference time vs ~15ms traditionally
An Analogy

• Take each sector and define a track. This goes in the “Positive sample”

• The other combinations go in the “negative sample”
Combinatoric Reduction

- Performance gains driven by **combinatoric reduction**.
  - The more track segments the more the benefit

![Graph](1 track events)

![Graph](2 track events)
CLAS12 Future work

- Collaboration with Old Dominion University
- Predict track trajectory from 2/3 of hits (the cleaner sectors).
  - Uses Recurrent Neural Nets
  - Cleans sectors then reduce combinatorics through better track finding
- Move on to estimation of initial state vector
Conclusion

- ML techniques from HEP are finding their way into NP
  - Problems are different enough to warrant careful study in this specific application

- Both GlueX and CLAS12 are undertaking parallel R&D (with cross-talk!)
  - CLAS12 has obtained funding for this R&D

- Both tracks look promising in reducing reconstruction times
  - And increasing code portability!