A study on fiTQun performance

V. V. Gligorov, M. Guigue, G. Díaz López, LPNHE, 25.02.2025

Introduction

The idea of this study is to

1. Profile fiTQun performance
2. Understand which parts of the reconstruction take the most time
3. Begin developing some tools to allow them to be made faster while controlling the impact on physics performance

All results have been obtained running the prefit branch on a Macbook equipped with an M3 Max processor and 64 GB of ram

Profiling is done using the CLion IDE's built-in tools

Input datasets

Used simulated samples of 300 events each of

1. Electrons
2. Non-decaying muons
3. Decaying muons

Generated with WCSim v1.12.19

Particles generated with 500, 750, and 1000 MeV kinetic energy

Work ongoing to also look at proton decay and pi0 events

Flamegraph of fiTQun

This is obtained running the prefit branch "out of the box" on some decaying muon events.

Notice that the MR + Pi0 fits take > 50% of the total time (but contribute nothing on these events). Will come back to this later.

Flamegraph of fiTQun

The 6D interpolation used in the scattering table is about 35% of all the function calls in fiTQun

Impact of 6D interpolation

In order to better understand the impact of the 6D interpolation, decided to replace it with a "nearest neighbour" approach

To do this, modified the existing GetInterpVal function. The function was already computing which bin in the scattering table our event belongs to, however when going float (data point) to integer (bin index) it was rounding down instead of calculating the closest bin. I changed this to calculate and then return the closest one.

Flamegraph of fiTQun after change

Impact of removing the 6D interpolation clear

Benchmarking indicates a ~20% speedup in the time taken per event by using the closest point rather than interpolating

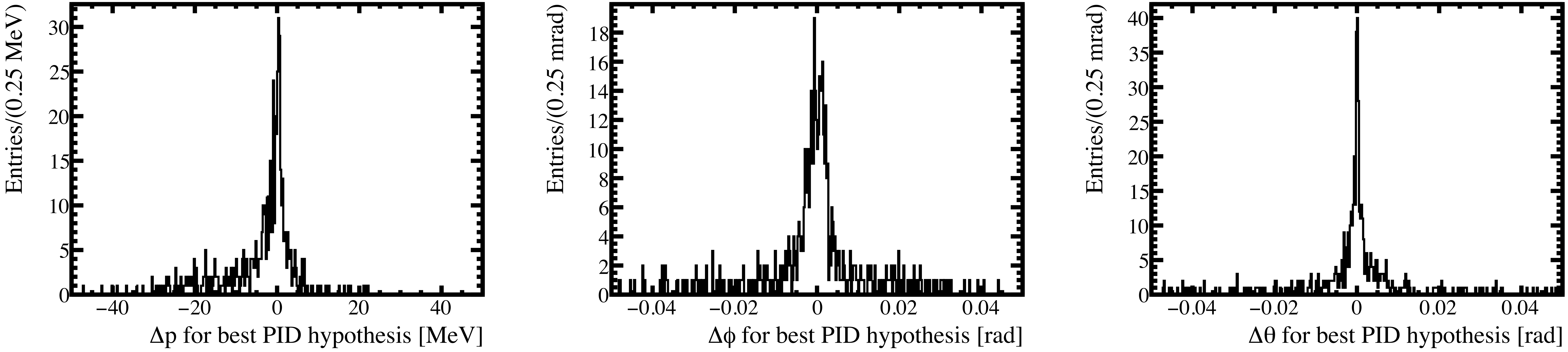
Impact on physics performance

Studied the impact on physics performance using a (very) simple configurable python script which takes the ntuples produced by fiTQun as input

Ran on each simulated sample with the 6D interpolation and the nearest neighbour approach and compared

1. PID assigned to each sub-event
2. Reconstructed particle momentum
3. Reconstructed particle direction

Physics performance differences (1)



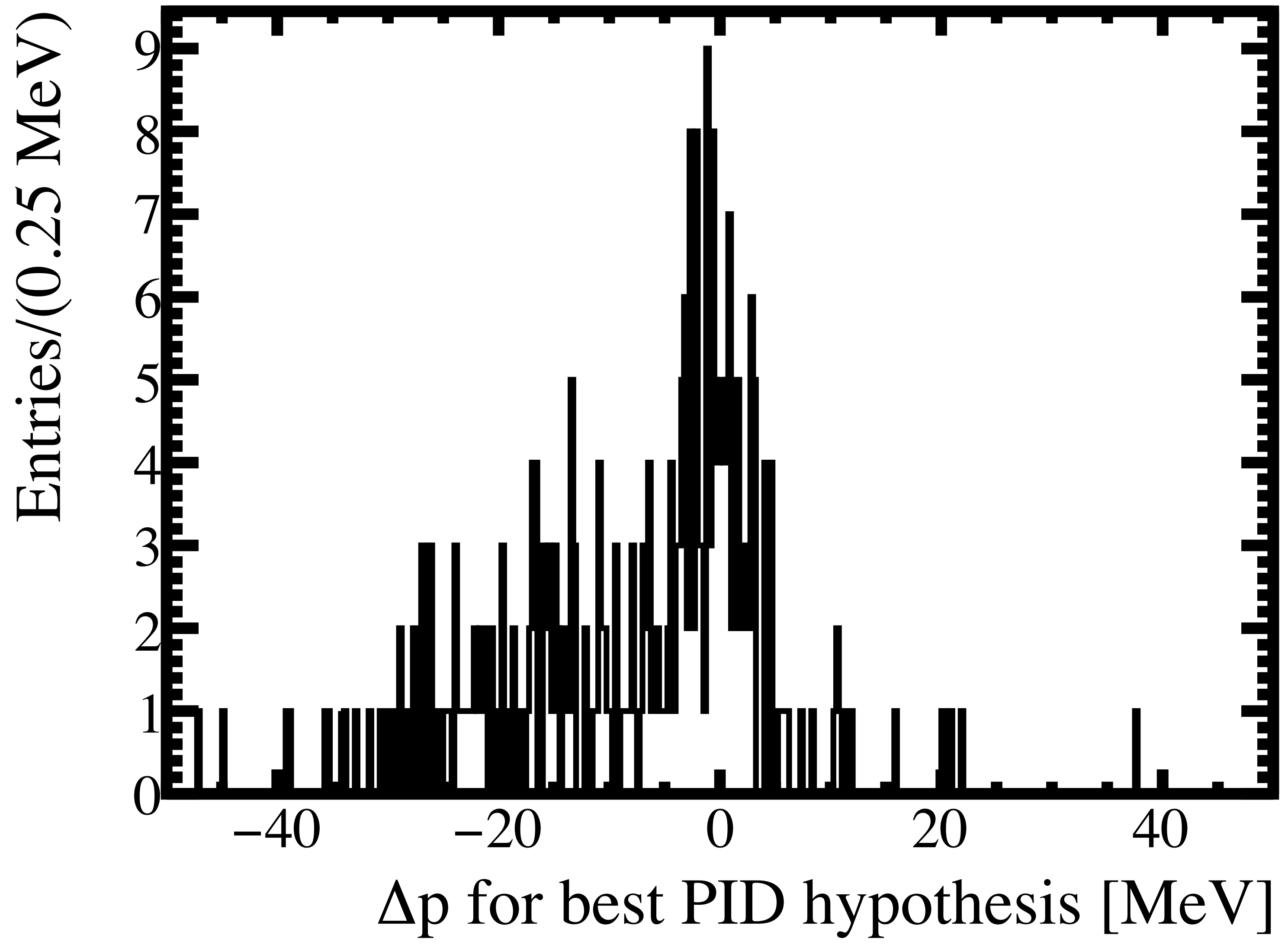
Ran over 300 muon events generated at 500 MeV in which the muon decays

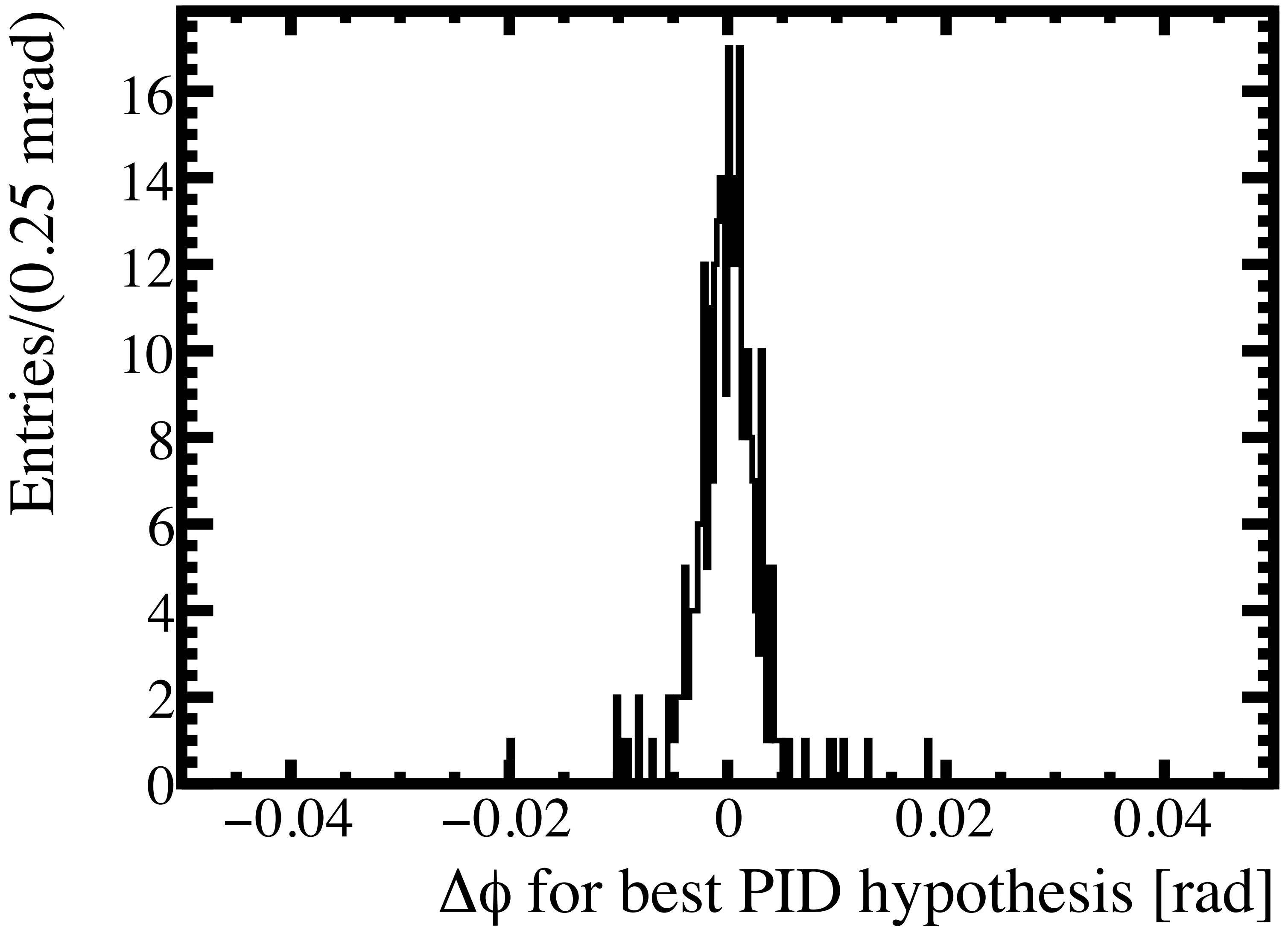
The 6D Interpolation and the closest bin method agree on the best PID for 546/547 sub-events

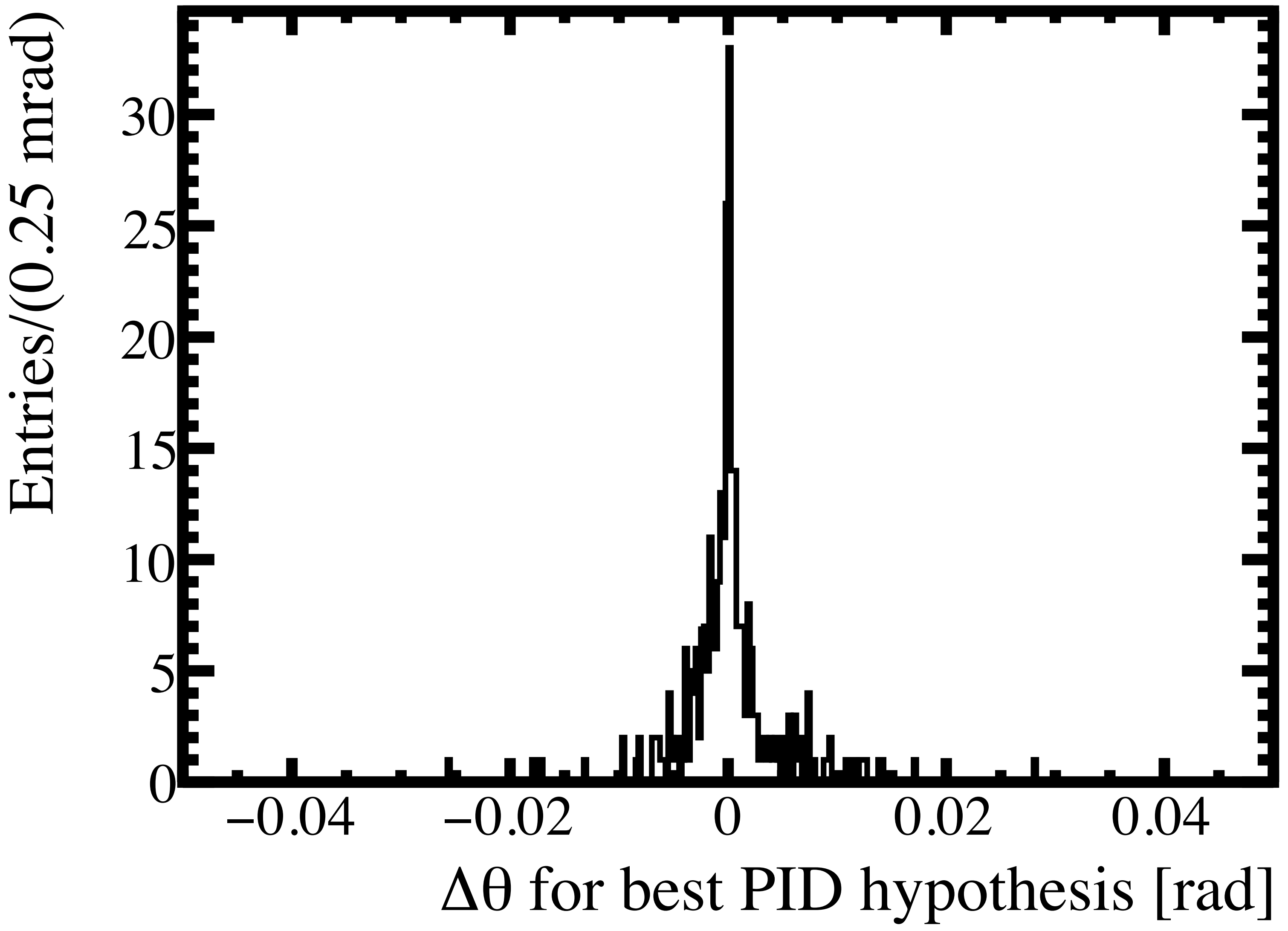
- One pi+ (6D) identified as an electron (closest)

Differences in recoed momentum and directions are fairly small, tails to be understood

Physics performance differences (2)



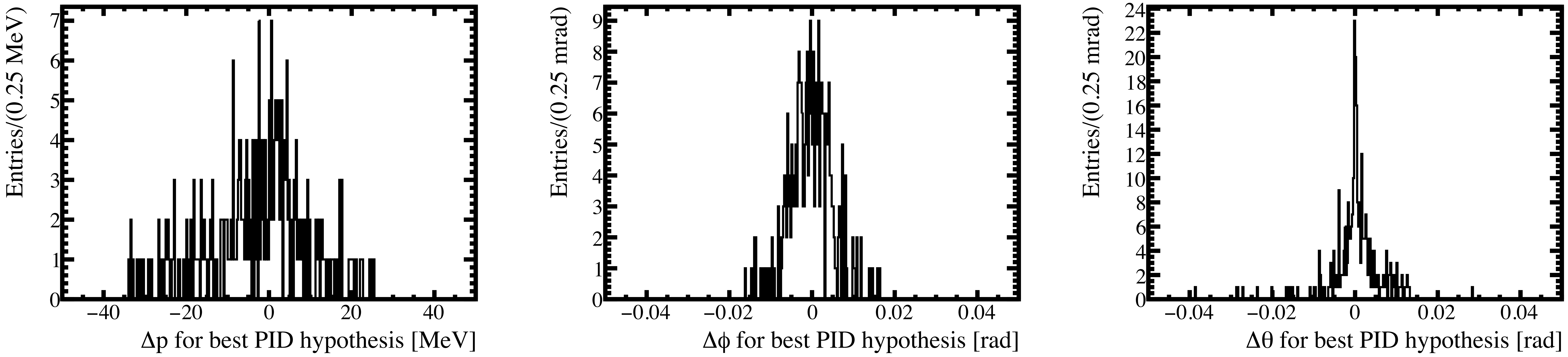




Ran over 300 muon events generated at 500 MeV in which the muon does not decay
The 6D Interpolation and the closest bin method agree on the best PID for all sub-events

Differences in recoed momentum and directions are still relatively small, but the tails become bigger and there is a bigger offset from 0 in the difference of reconstructed momenta

Physics performance differences (3)



Ran over 300 electron events generated at 500 MeV

The 6D Interpolation and the closest bin method agree on the best PID for all sub-events

Differences in recoed momentum and directions are bigger than for muons

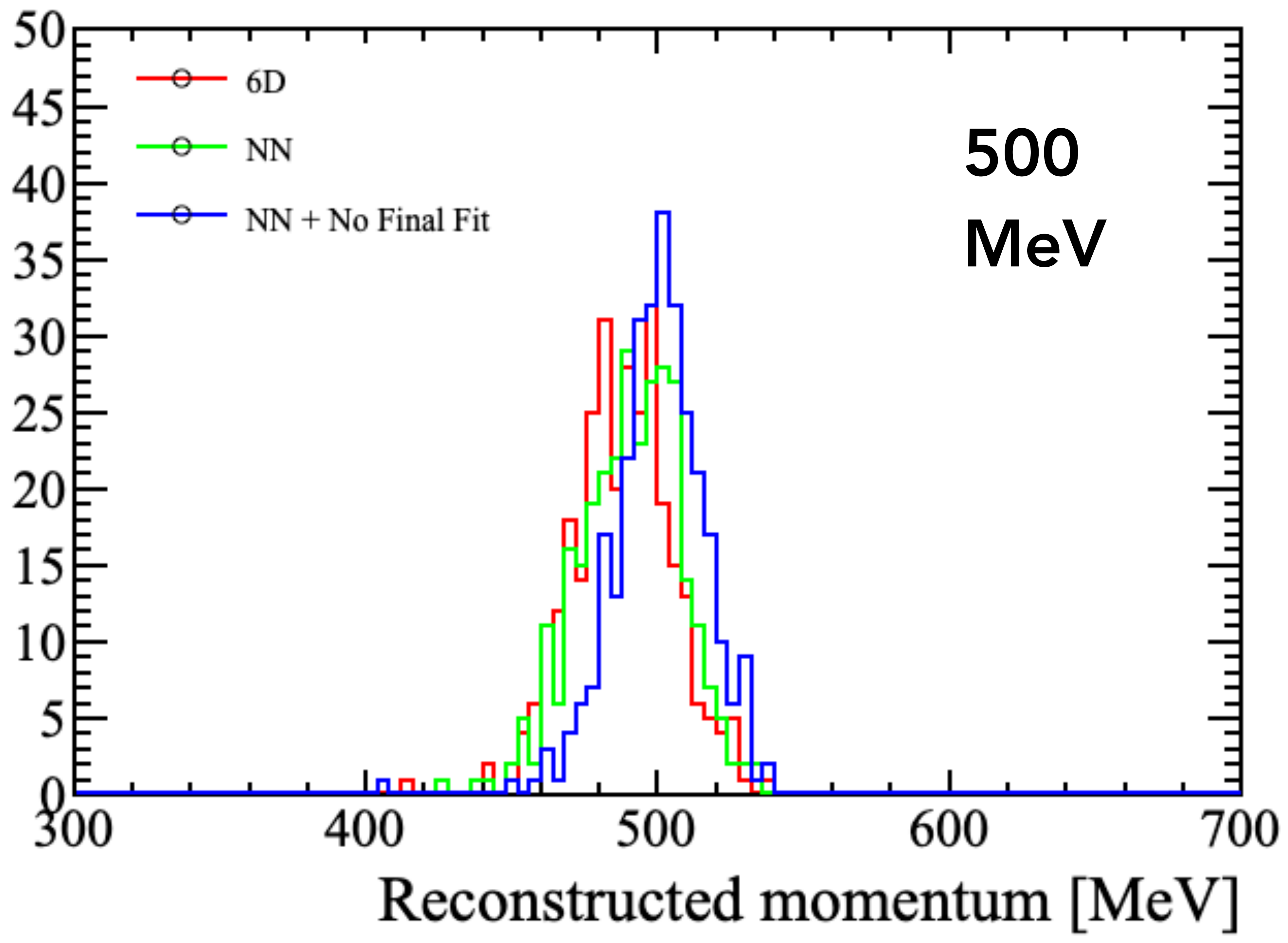
How can we make things faster?

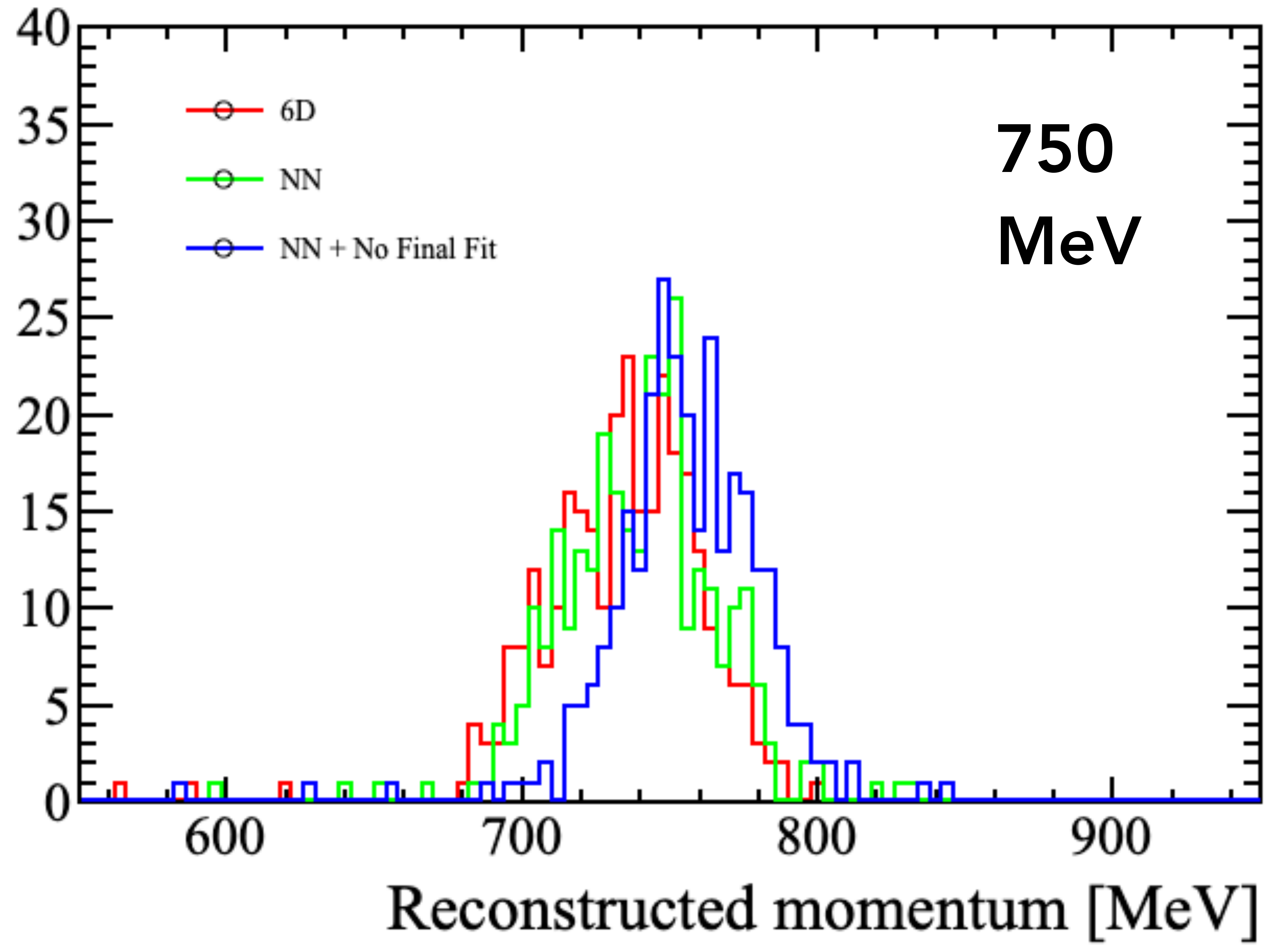
Tried the following things in addition to the nearest-neighbour approach for scattered light

1. Turn off the final fit in the One Ring fit
2. Turn off the Multi Ring fit
3. Turn off the Pi0 fit

All together these gain 4-5x in speed compared to the "out of the box" fiTQun. While (1) has an impact on the physics, (2) and (3) have no impact on the reconstructed best PID hypothesis, momentum, or direction in these specific MC samples (expected but good!).

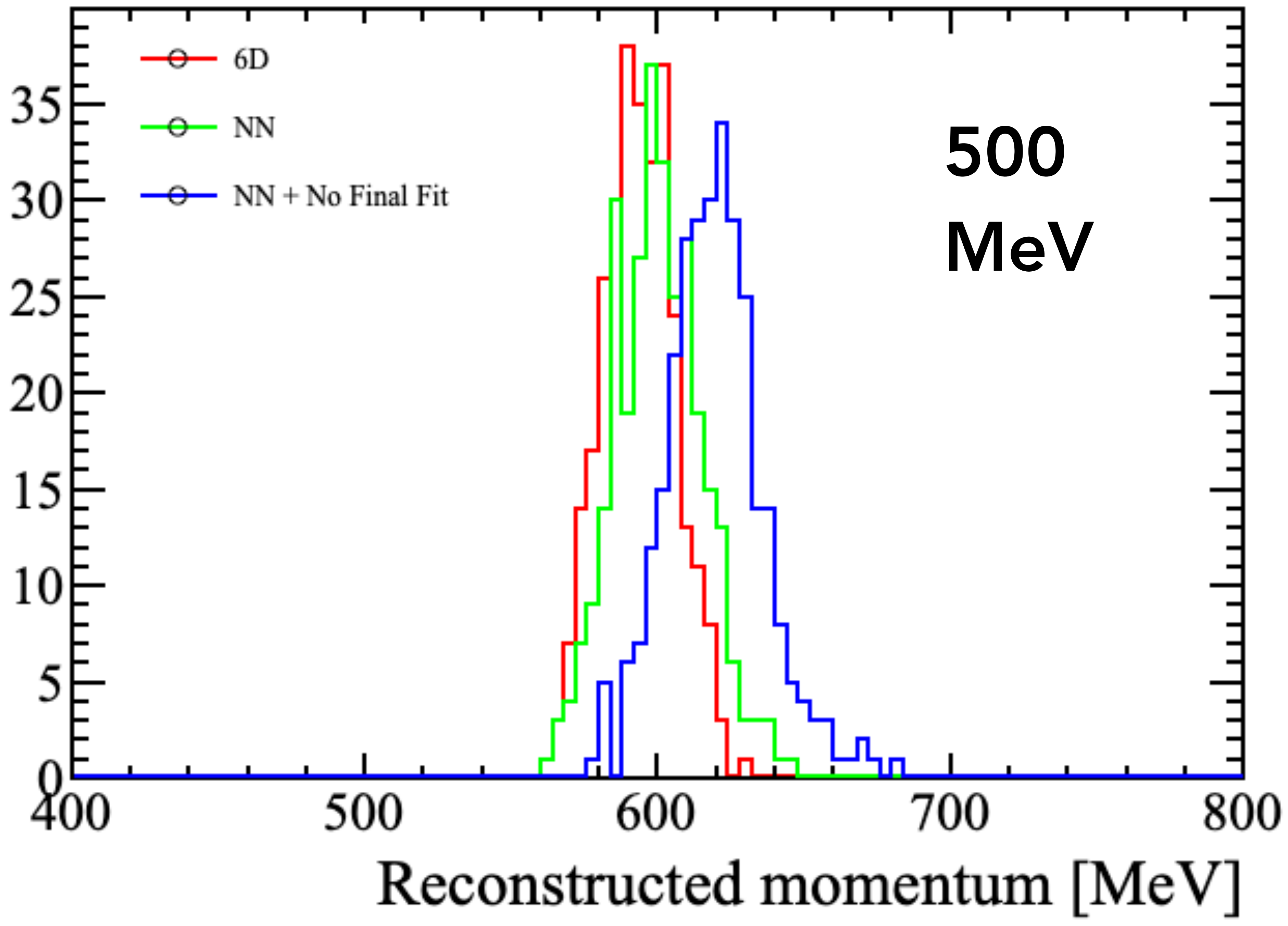
Removing final 1R fit: electrons

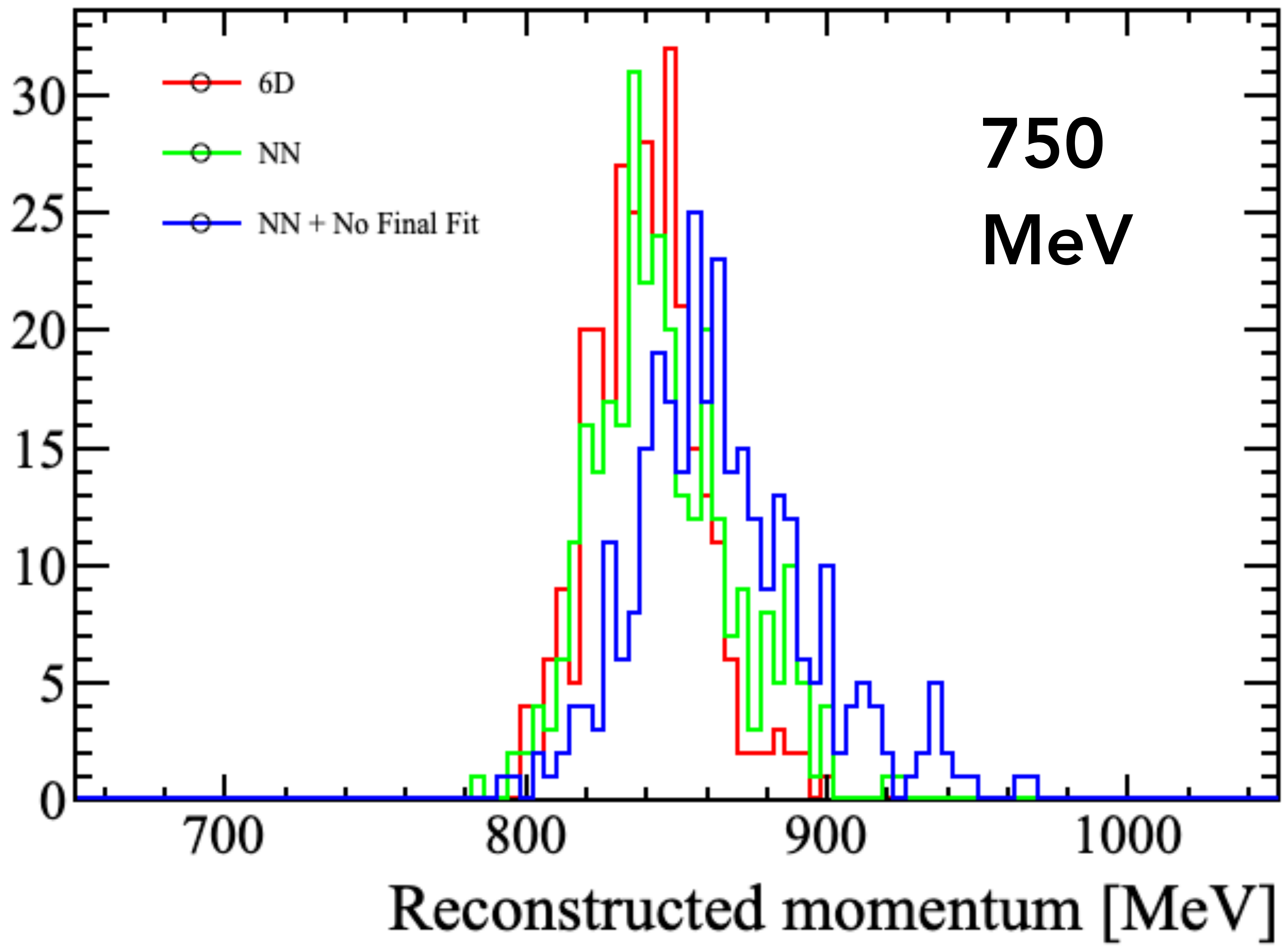




All results shown for the combination of using the nearest-neighbour approach for scattered light and removing the final One Ring fit.

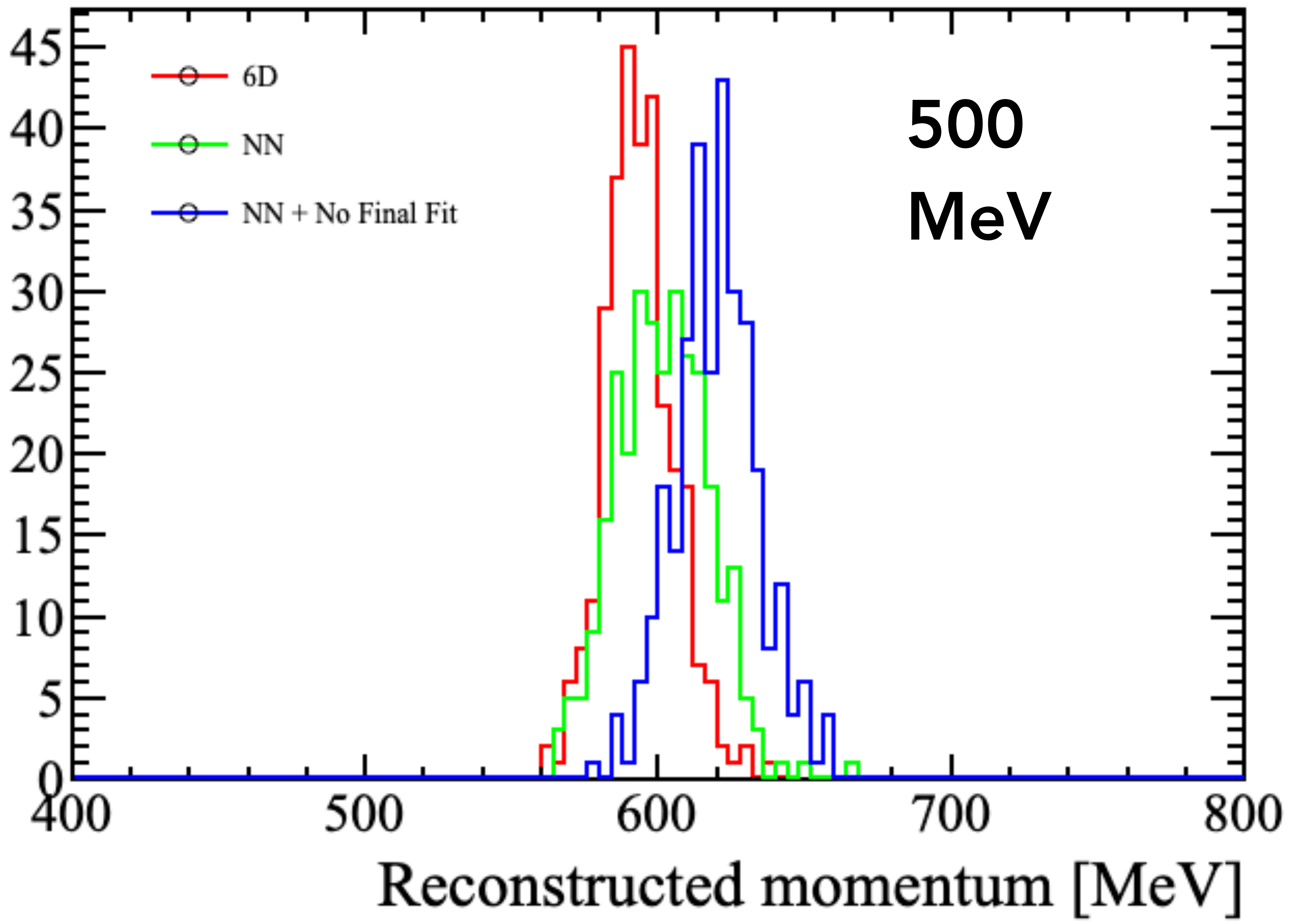
Removing final 1R fit: muons

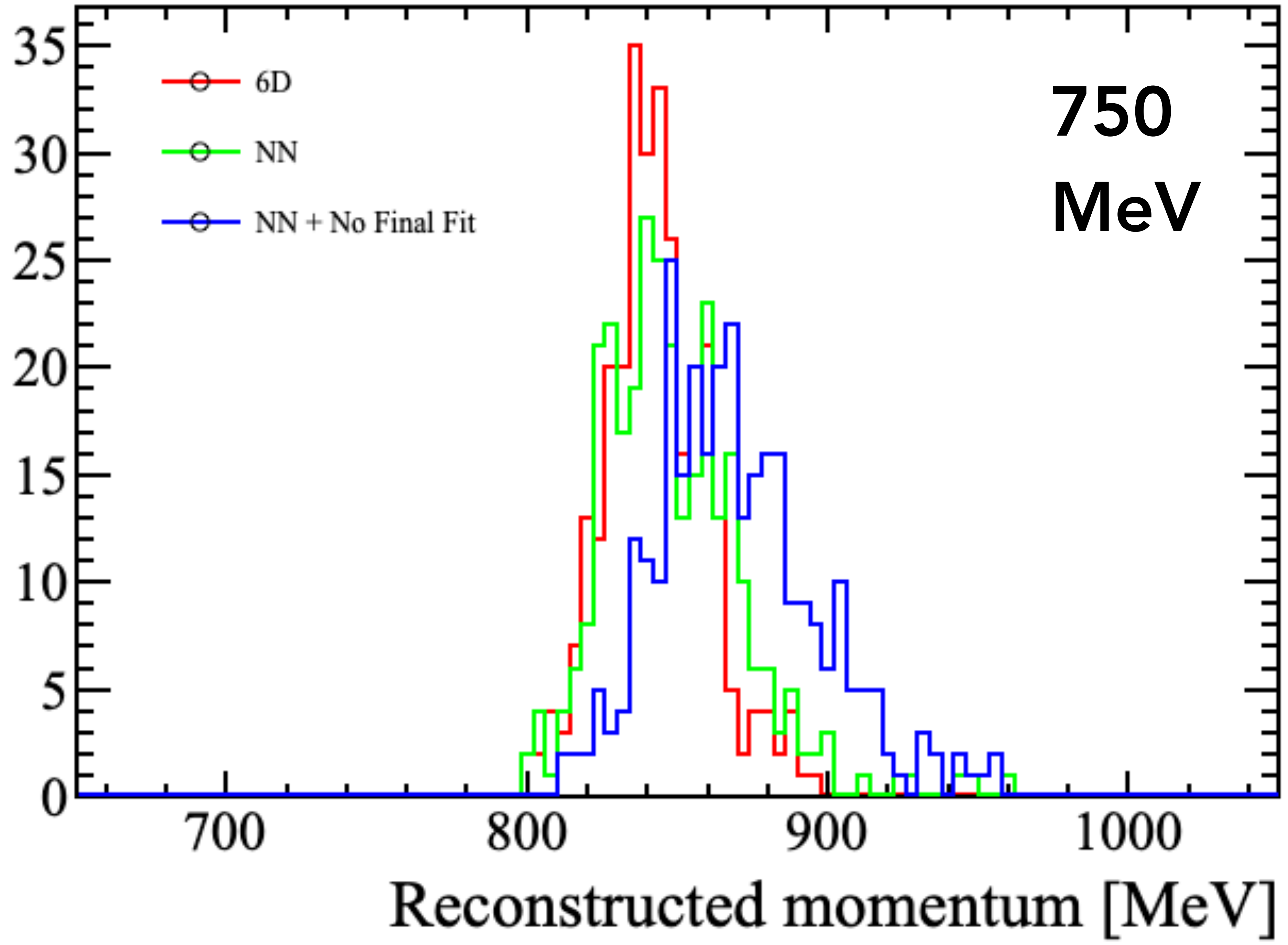




All results shown for the combination of using the nearest-neighbour approach for scattered light and removing the final One Ring fit.

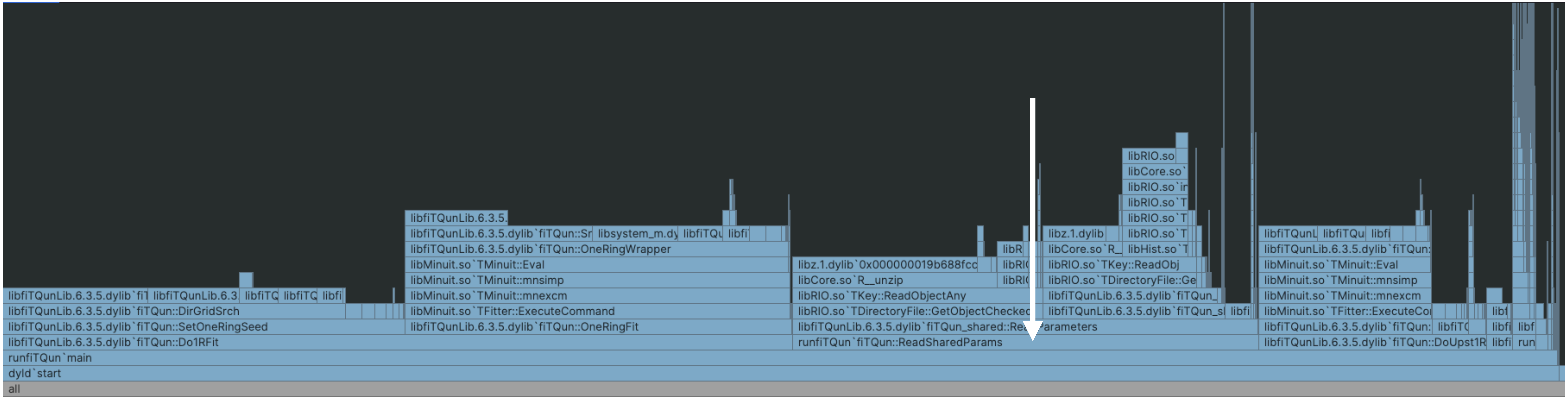
Removing final 1R fit: non-decaying muons





All results shown for the combination of using the nearest-neighbour approach for scattered light and removing the final One Ring fit.

Flamegraph of fiTQun for 1R fit



30% of time is taken by reading shared parameters?

Conclusions and next steps

- Carried an initial look into fiTQun performance.
- For processing MC already turning off the multi-ring and Pi0 fits and switching to a nearest-neighbour approach (since we do not seem to be memory-bound we could make the lookup table more granular in some regions?) could gain roughly 4x in computational time.
  - Final one ring fit seems to make things worse for electrons?
- Have written a simple script to compare two output ntuples from the reconstruction, worth committing for validation purposes?
  - How can I do more fine-grained comparisons of PID performance after cuts as people show in HK meetings?
- Very happy + motivated to keep working on this and get suggestions what is the most useful place for me to contribute.

BACKUPS

Memory usage and scaling

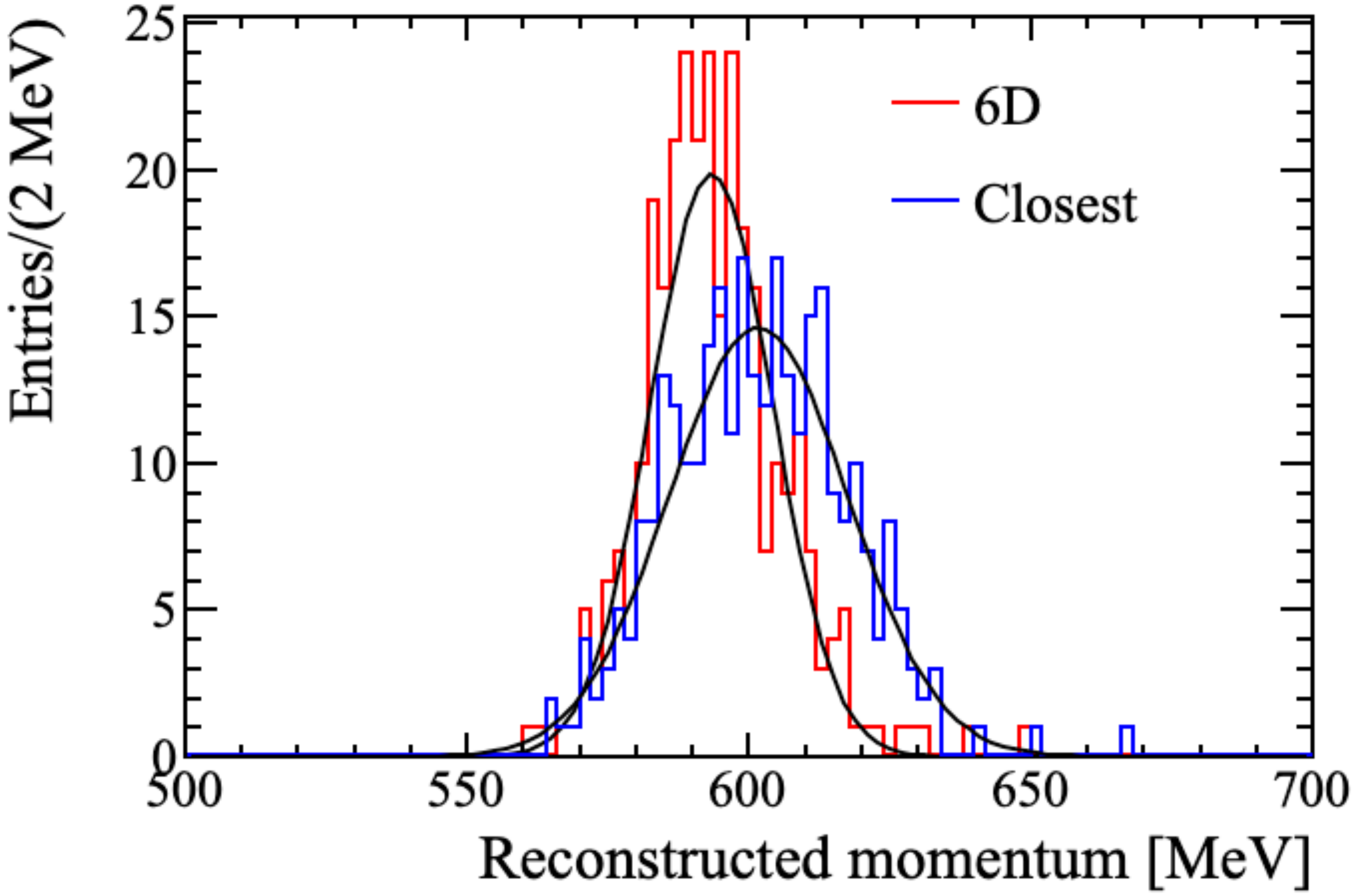
An instance of fiTQun uses about 1 GB of memory, and basically all of this is allocated at the start of the processing

This means it should be possible to run multiple jobs in parallel on more or less any modern x86 server and get a good scaling not only with the number of physical but also virtual cores

- Not the case out of the box with Apple silicon even if I use nice to set process priorities and copy input files to avoid I/O limits

To be verified on cc-in2p3 servers as a follow-up

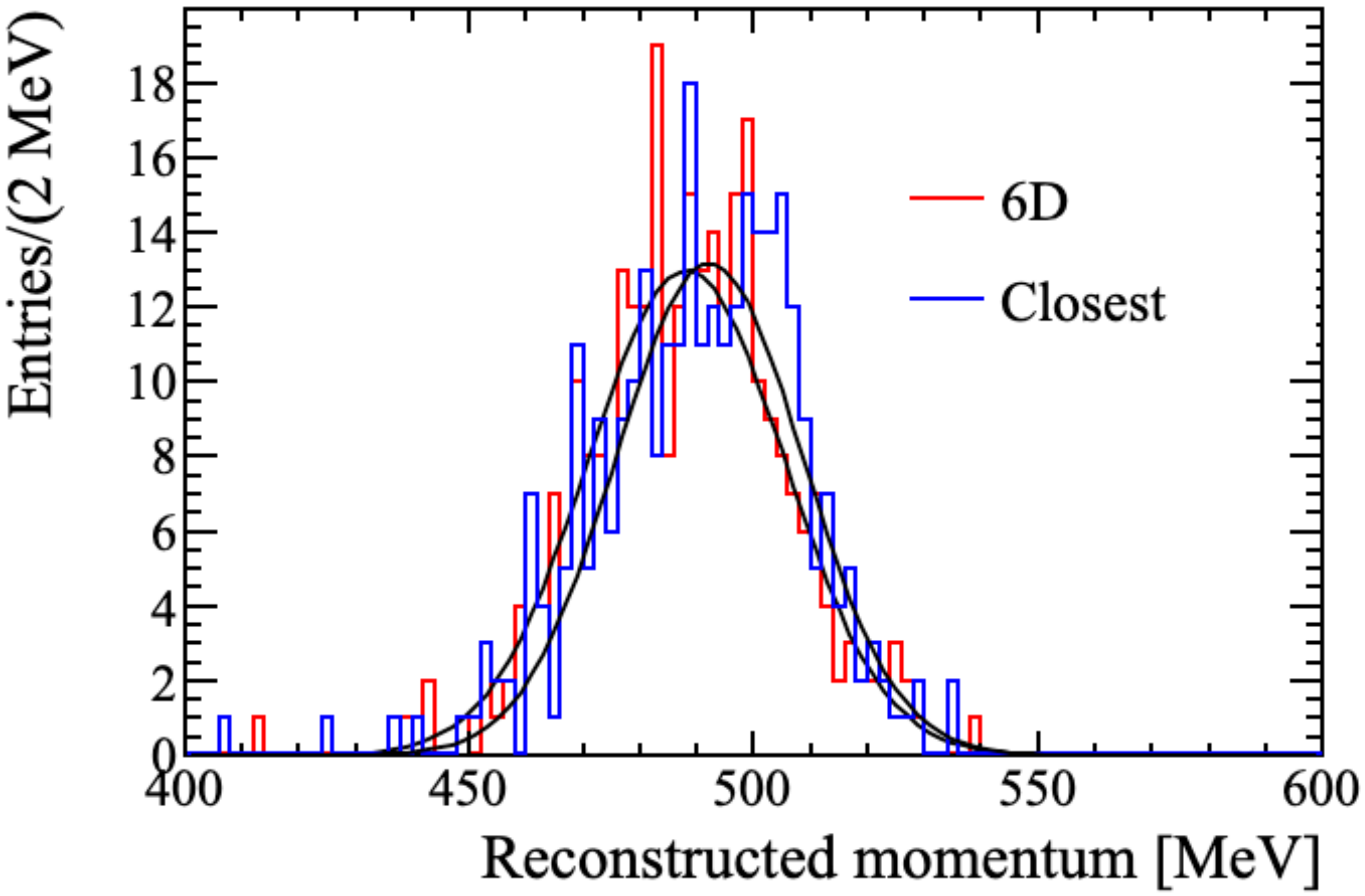
Reconstructed momenta for muons



Resolution is noticably worse at 500 MeV for the "closest" setting, 15.9 vs. 10.8 MeV
Mean is also a little bit more biased in the "closest" setting than in the 6D

Similar conclusions hold for the non-decaying muons

Reconstructed momenta for electrons



Resolution is a bit better at 500 MeV for the "closest" setting, ~16.3 vs. ~17.3 MeV
Mean is also a little bit *less* biased in the "closest" setting than in the 6D