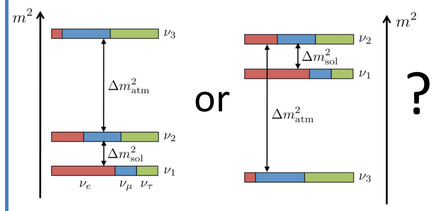
High Performance Computing Workflow for Liquid Argon Time Projection Chamber Neutrino Experiments

S. Berkman, G. Cerati, P. Gartung, W. Ketchum, K. Knoepfel, J. Kowalkowski, B. Norris, M. Paterno,
T. Peterka, A. Reinsvold Hall, R. Ross, S. Sehrish, S. Syed, M. Wospakrik, O. Yildiz

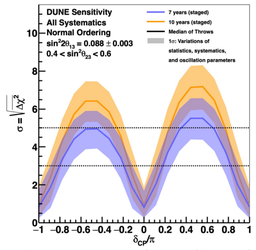
Why neutrinos?

Big open questions to address with current and future neutrino experiments[1]:

Which is the heaviest?

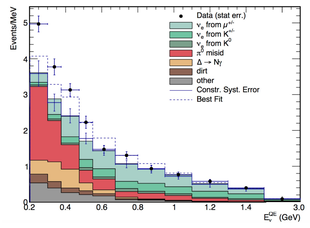


Is there CP violation[2]?



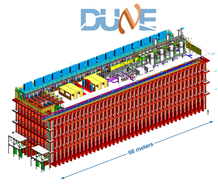
Assumes normal mass hierarchy

Is there other new physics[3]?



Neutrino Physics in LArTPCs

- Series of operating, commissioning, and planned liquid argon time projection chambers (LArTPCs) in the US
- Detectors are growing, and more neutrinos are expected
- Challenges ahead to process the data efficiently



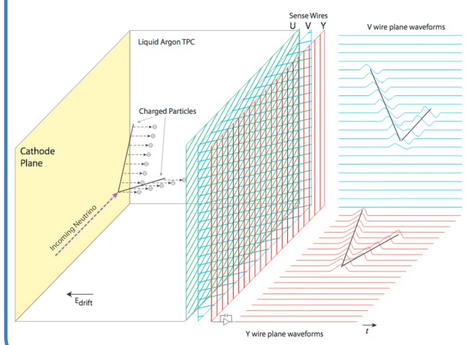
1x wires → 5x wires → 100x wires

1x mass → 5x mass → 500x mass

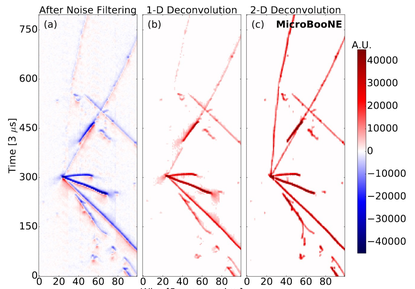
1x POT/year → 10x POT/year

Event Reconstruction in LArTPCs

1. Detector Output [4]

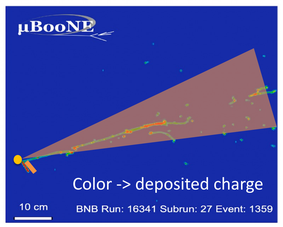


2. Signal Processing [5]



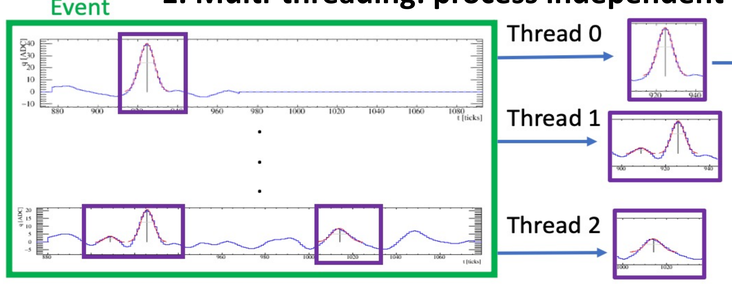
3. Hit Finding [6]

4. Pandora Pattern Recognition [7]



- Given computing trends, parallelism is necessary to gain speed increases: typically parallelize at event level, can also take advantage of parallelism within events

1. Multi-threading: process independent data regions on separate threads



2. Vectorization: perform same operation in parallel on different data bins

- In LArSoft[6] framework used by all Fermilab-based LArTPC experiments:
  - Up 2x faster from vectorization with avx-512+icc and pragmas over most time-consuming loops
  - Up to 17x faster from multi-threading within events using TBB with 95% parallel fraction
- Used by Icarus and ProtoDUNE experiments

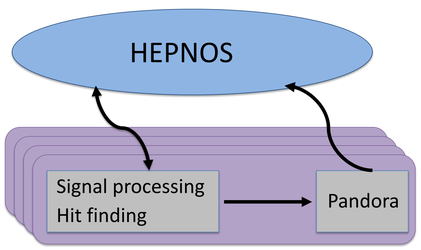
HPC Workflow

- Developing workflow for Theta HPC at ALCF:
  - Enable more parallelism than available in typical grid production environment in signal processing and hit finding algorithms
  - Develop and incorporate new tools
  - Future: Efficiently include AI reconstruction algorithms

Local build of LArSoft with spack[8,9] on Theta:

1. Enable custom compiler options for specific pieces of code: ie. avx512+icc for optimal vectorization speed increases in hit finding
2. Build experiment software releases with spack

Use HEPNOS[10,11] to manage the data during processing



Full workflow running on Fermilab local computers, close to complete on Theta

Conclusions

- LArTPC reconstruction algorithms can be parallelized to efficiently use computing resources, including HPCs, and enable physics of interest
- Working towards making these improvements available to experiments through development of a HPC workflow for data processing

References:

[1] M. Sajjad Athar et al., Prog. Part. Nucl. Phys. 124 (2022), 103947
[2] DUNE Collaboration, arXiv:2002.03005[hep-ex]
[3] MiniBooNE Collaboration, PRL 121 (2018), 221801
[4] MicroBooNE Collaboration, JINST 12 (2017) , P02017
[5] MicroBooNE Collaboration, JINST **13 (2018),** P07006
[6] S. Berkman *et al*, JINST 17 (2022) 01, P01026
[7] MicroBooNE Collaboration, Eur.Phys.J.C 78 (2018), 82
[8] C. Green, J. Amundson, L. Garren, P. Gartung, and M. Paterno, Marc, https://doi.org/10.1051/epjconf/201921405013
[9]https://indico.fnal.gov/event/47281/contributions/206026/attachments/143674/181920/Spack%20-%20HSF%20Presentation%20%281%29.pdf
[10] https://hepnos.readthedocs.io/en/latest/
[11] https://bytebucket.org/jkowalkowski/heponhpc-data/wiki/highlights/accomplishment-hepnos-2018.pdf