2x2 simulation production: Review and current status

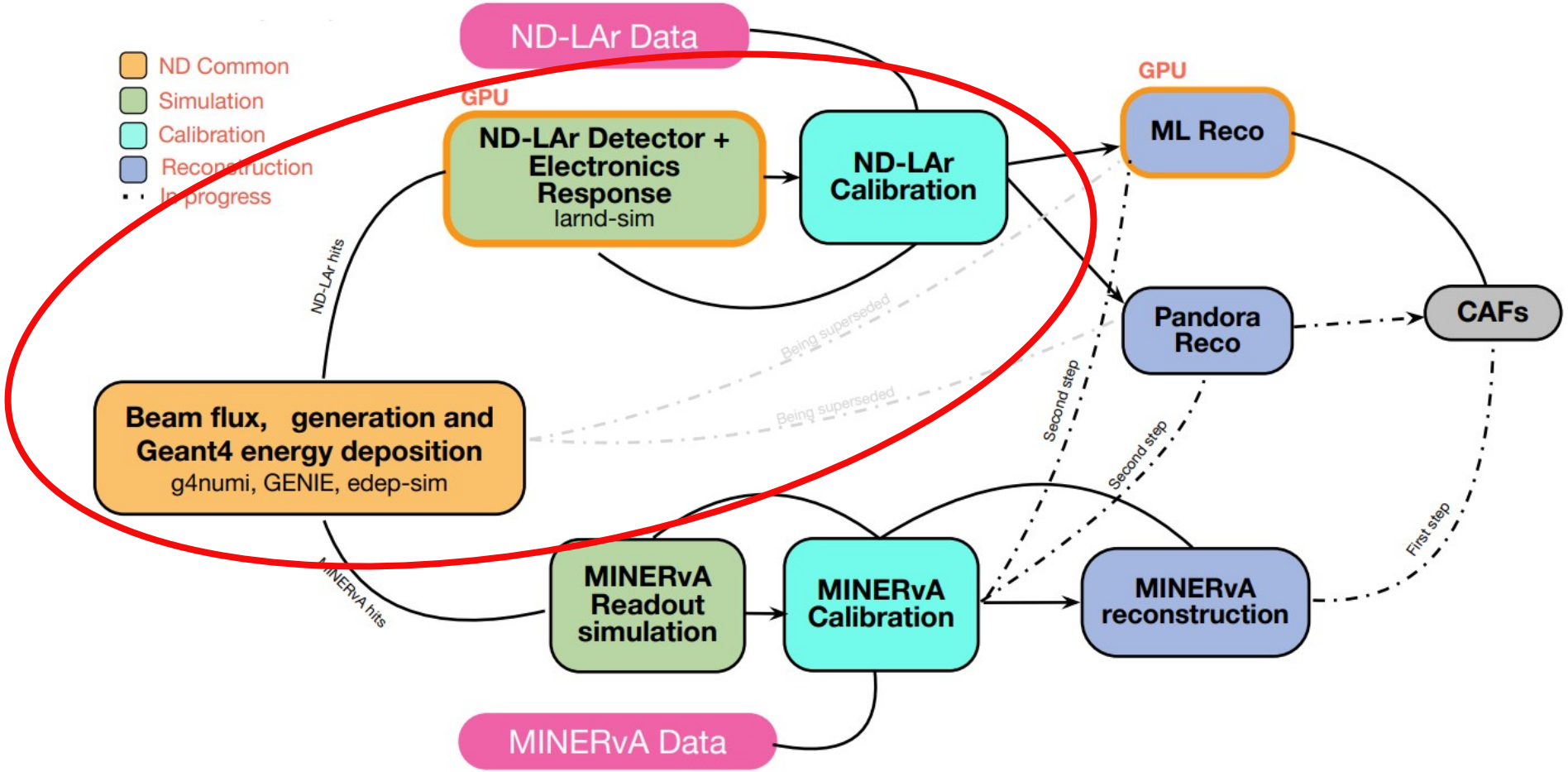
Matt Kramer (LBNL)

ND prototypes analysis workshop
May 19 2023

Outline

- Brief review of the 2x2 simulation chain
  - Provide newcomers some context for rest of workshop
  - If you aren’t new, sorry for the repetition
- Current status
- Plans for the near future
- Ways to get involved!

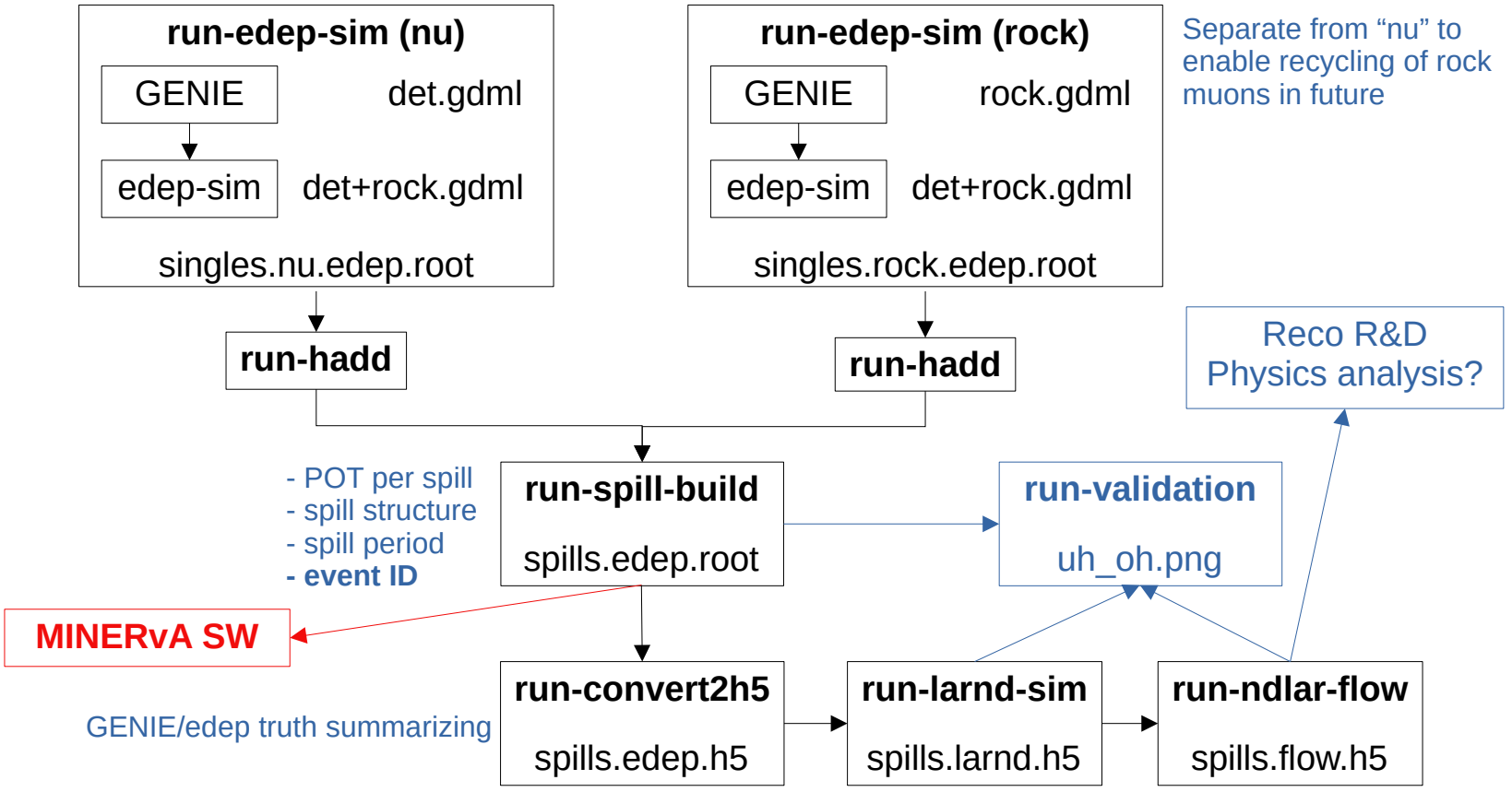
This again


ND Common
Simulation
Calibration
Reconstruction
In progress
ND-LAr Data
GPU
ND-LAr Detector + Electronics Response
larnd-sim
ND-LAr Calibration
GPU
ML Reco
Pandora Reco
CAFs
ND-LAr hits
Being superseded
Being superseded
Second step
Second step
First step
Beam flux, generation and Geant4 energy deposition
g4numi, GENIE, edep-sim
MINERvA hits
MINERvA Readout simulation
MINERvA Calibration
MINERvA reconstruction
MINERvA Data

Our corner: The pieces

- **GENIE:** Event generator. Produces neutrino interactions, given NuMI beam flux description and hall/detector geometry.
- **edep-sim:** Geant4 wrapper. Propagates particles through the geometry, stores energy deposits in “active” volumes
- **hadd:** Merges edep-sim files to make later steps less awkward
- **Spill builder:** Overlays fiducial and rock events, applies spill exposure (events per spill), spill structure, spill separation
- **larnd-sim:** Detector simulation (GPU-based): Charge drift, scintillation light, charge/light readout, etc.
- **ndlar_flow**: Calibration, low-level reconstruction
- **validation:** Plots!

Our corner, visualized



Where's the code?

- https://github.com/DUNE/2x2_sim
  - Bash scripts to drive the chain: GENIE+edep-sim, larnd-sim, ndlar_flow
  - Decoupled from any workflow / production system
  - All options set via env vars; simple to map to any production DB
  - Now there's an evolving wiki
  - Tutorial this afternoon!
- https://github.com/lbl-neutrino/fireworks4dune
  - Simple integration with FireWorks workflow manager
    - Python, MongoDB ("pile of JSON"), YAML — familiar-enough tech
  - "Runner" script
  - Scripts that populate Mongo DB with job specs
  - YAML files with job specs (env vars etc.)
  - NERSC Slurm jobs

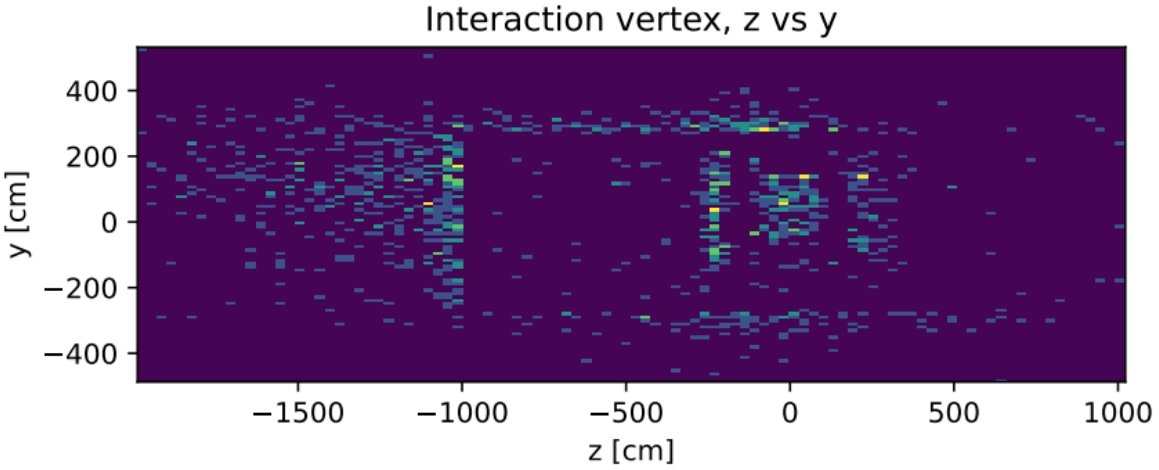
Where are we now?

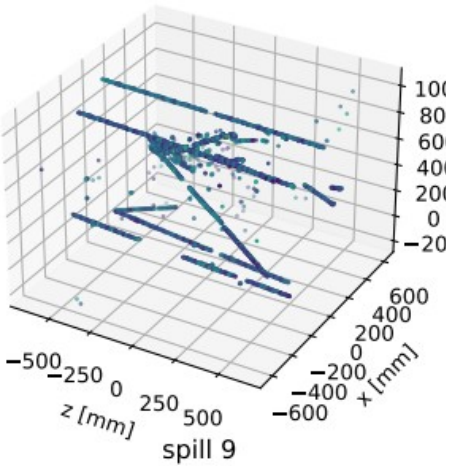
- Eventual goal: **~1E22 POT of simulated 2x2 + MINERvA data,** corresponding to ~10x 1 year in NuMI (RHC expected)
- Goal declared at Bern analysis workshop: **10 weeks of simulated 2x2 + MINERvA data** (up to larnd-sim)
  - 5E13 POT/spill * 3000 spills/hr * 24 * 7 * 10 = **2.5E20 POT**
- Now iterating on “mini” productions of **1E19 POT** (RHC so far)
  - Reasonable statistics for reco R&D (200k spills)
  - Validation results, analyzer feedback → updates for next iteration
  - Stress test of production infrastructure
  - Scale to 2.5E20 POT after validation feedback cycle converges

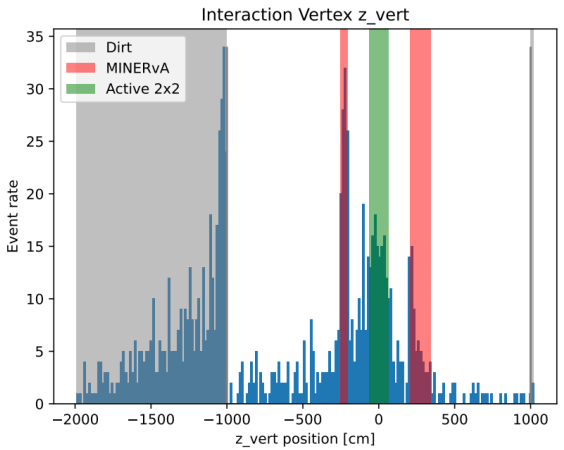
Current production: MiniRun3

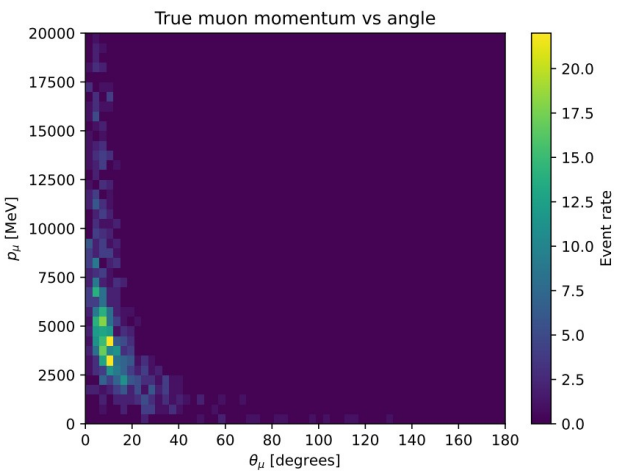
- ~1E19 POT RHC (~2.7 days)
  - 5E13 POT/spill, 200k spills, ~1 neutrino / ~5 rock muons per spill
  - 1,024 files (minus 2 failures)
  - Includes edep-sim (ROOT; for MINERvA), larnd-sim (HDF5; for mlreco), ndlar_flow (HDF5; for Pandora, later for mlreco)
- Available from NERSC, Fermilab, web
  - https://github.com/DUNE/2x2_sim/wiki/MiniRun3-file-locations
  - SAM cataloging in progress
- Intended purpose:
  - Process by MINERvA, Pandora, mlreco
  - Identify issues with the production itself as well as above packages

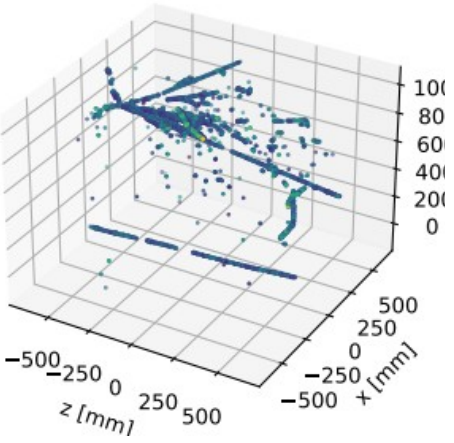
Some validation plots (one 1E16 POT output)









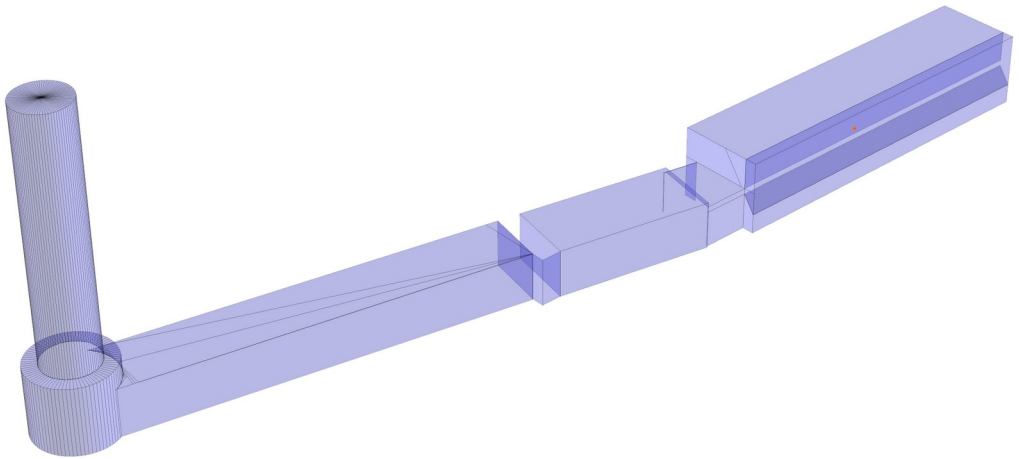


Extensions of MiniRun3

- Repeat for FHC (starting from GENIE)?
- Repeat for different pixel thresholds (starting from larnd-sim)?
  - Suppress/enhance inductive hits, for reco benchmarking and devel
  - Inform the decision on actual thresholds for 2x2 operation
- **Rerun ndlar_flow as improvements are implemented**
  - We're already doing this
  - flow_v5 released; flow_v6 on the way
  - Changes tracked on our wiki
  - Future productions ("MiniRun4", …) will also be "reflowed" regularly

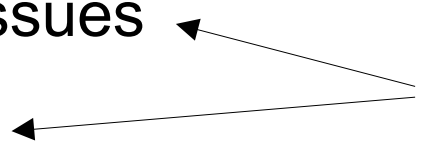
Planned MiniRun4 improvements

- Optimized beam window for rock muons (thanks Alex!)
  - ~80% reduction in CPU hours (for current hall geometry)
- Geometry updates:
  - **Cavern:** Largely done (thanks Zach and Alex!)
    - Expect significant change in rock muons
  - 2x2/cryostat/plumbing/etc: Maybe?
- larnd-sim
  - Beam trigger?
  - Improvements to light simulation
    - Trigger, backtracking, etc.
- ndlar_flow
  - Light reconstruction
  - Charge-light association
  - Improved merging of inductive (“fuzzy˝) hits



Ways to get involved

- Geometry improvements
- Validation plots
  - Add new ones
  - Inspect existing ones
  - Define criteria for “good or bad”
- Play with the data, look for issues
- Learn how to run the chain
- Port the chain to other GPU clusters
- Improve larnd-sim, ndlar_flow, ...



Summary

- MiniRun3 being processed by MINERvA, Pandora, mlreco
- Extensions to MiniRun3 planned (?)
  - FHC, alternate pixel thresholds
- MiniRun4 planned
  - Geometry, larnd-sim improvements
  - Extend to larger statistics?
- ndlar_flow under rapid development
  - Regular “reflowing” of MiniRun3
- Plenty of opportunities to get involved!