Developing reconstruction algorithms for Hyper-Kamiokande

N. F. CALABRIA

INFN and Politecnico di Bari

HK data accessibility

HK simulated data are natively available as ROOT files which require the compilation and installation of several analysis libraries.

I developed a simplified file format based on pandas dataframes to make HK simulated data Immediately available without any dependency from the HK codebase, for our internal uses.

I generated 100k electron gun events (vertex distribution homogeneus up to 1m from the Inner Detector Wall, direction isotropic, Kinetic Energy between 0 MeV and 1000 MeV).

Processed files in the new format are available here for Emanuele:
/lustrehome/nfcalabria/shared/100k_ranvtx_ranmom_0_1000_pandas

An example script to read them is:
/lustrehome/nfcalabria/shared/read_pandas.py

If you have a ReCaS account and you are interested in accessing those files, just let me know and I'll add your username to the ACL!

Our data format

Each file contains 6 pandas dataframes:

pmts: one row per 20-inch Photomultiplier Tube (PMT). **Row index** identifies the pmt, it is referenced by **hits** column **id**. Columns:
    x, y, z: Position of the PMT (cm)
    dir_x, dir_y, dir_z: Direction of the PMT sensitive surface as a unitary vector.
    cyl_loc: location of the PMT: 0 (top cap), 1 (barrel), 2 (bottom cap)

mpmts: same as pmts, but for 3-inch PMTs

evts: list of events. **Row index** is referenced by **hits** column **evt_key** . Columns:
    nevt, **run**: this pair identifies an event inside the file.
    ntrigger: one row per trigger. Each event is divided in triggers. In this production it's always 0 (only one subevent)

hits: list of hit 20-inch PMT information. Columns:
    id: references **pmts** dataframe **row index**
    charge: charge collected
    time: time of detection
    evt_key: references **evts** dataframe **row index**

hits2: same as hits, but for 3-inch PMTs

tracks: one row per true Monte-Carlo track. Columns:
    nevt, run: the event which this true track belongs to
    PID: particle type GEANT code (https://pdg.lbl.gov/2007/reviews/montecarlorpp.pdf)
    id, parent_id: track id and in case of decay parent track id (root track has id == 0)
    x, y, z: vertex of track (cm)
    dir_x, dir_y, dir_z: direction of initial particle momentum as a unitary vector
    mom: momentum (MeV)

Event display

The script file read_pandas contains a basic 3-D event display.

Instantiate reader with:
      reader = HKPandasReader(file_path)

Draw charge of hit 20-inch PMTs of event (nevt, run)
      reader.plotEventPmtCharge(nevt, run)

Draw time of hit 20-inch PMTs of event (nevt, run)
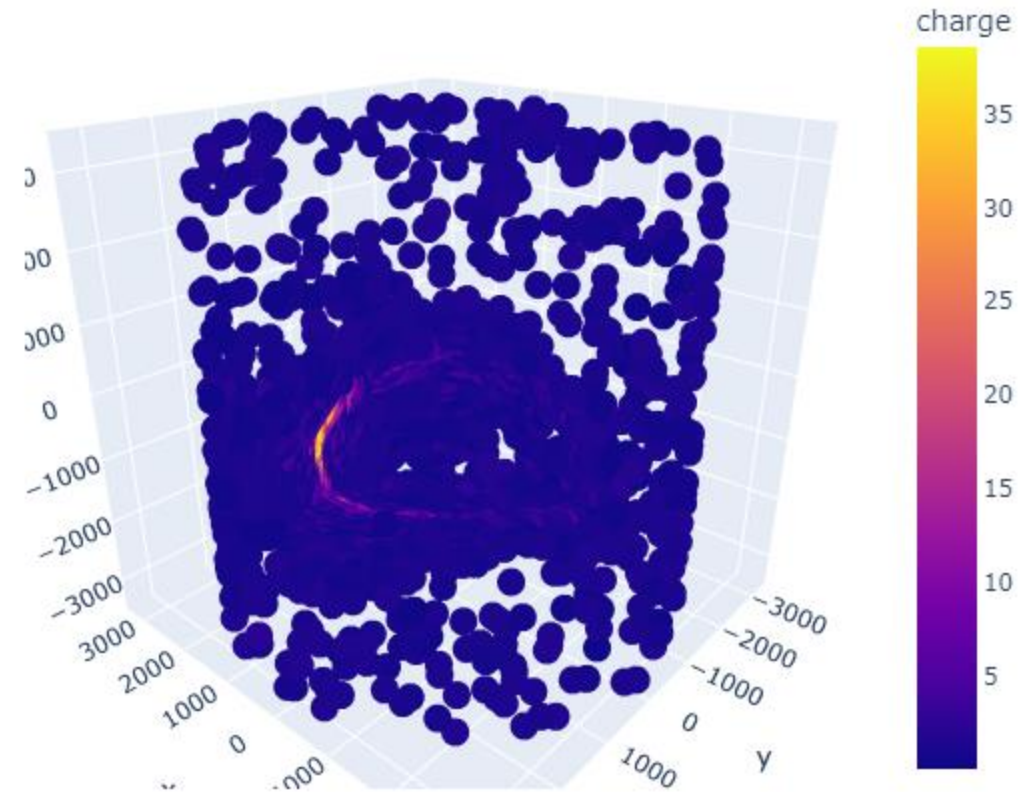   **reader.plotEventPmtTime(nevt, run)**

Draw charge of hit 3-inch PMTs of event (nevt, run)
      reader.plotEventMPmtCharge(nevt, run)

Draw time of hit 3-inch PMTs of event (nevt, run)
      reader.plotEventMPmtTime(nevt, run)

Current research topics on reconstruction in HK

- Port fiTQun, max-likelihood based algorithm, to HK

- Research on Convolutional Neural Networks (CNN) for reconstuction, mainly ResNet

- Research on Graph Neural Networks (GNN)