Event building

TB2021 Analysis efforts: kick off meeting

Jonas Kunath (LLR). 30.11.2021.

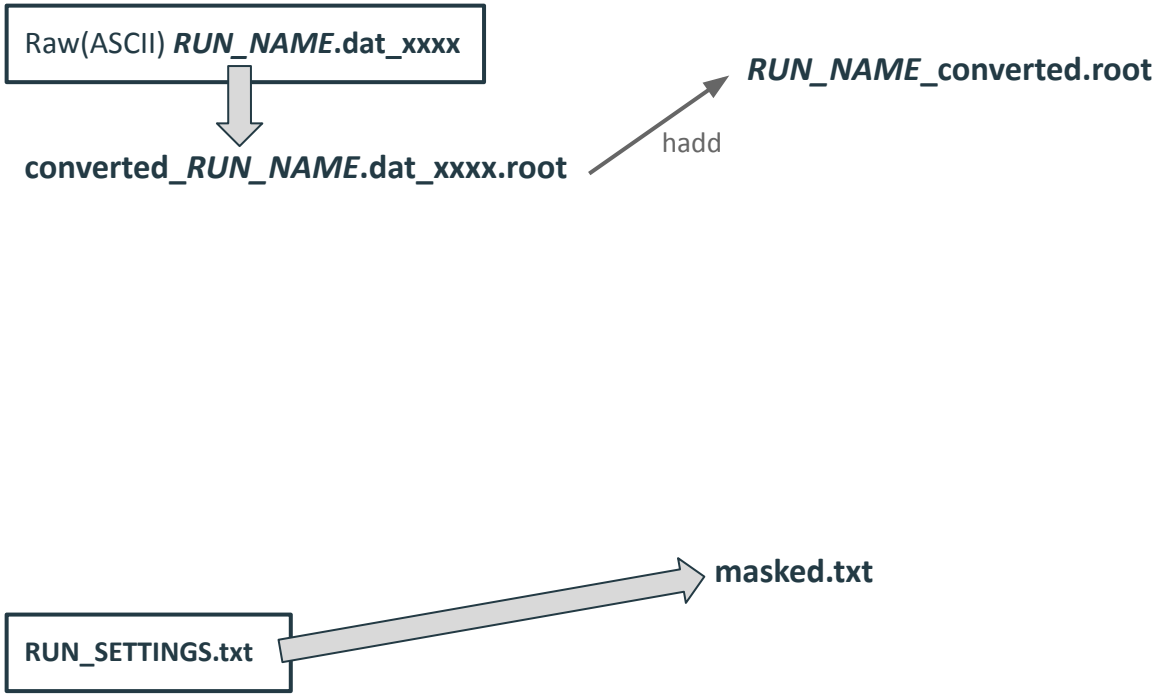
What are the buildfiles?

Event building pipeline

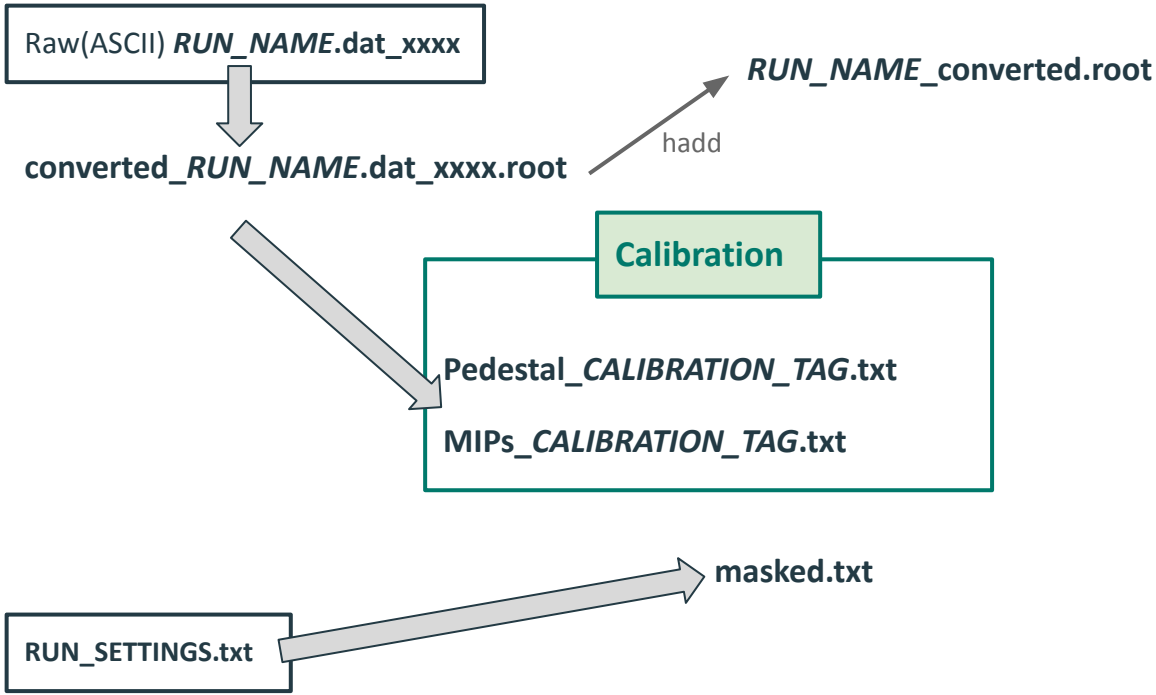
Raw(ASCII) ***RUN_NAME*.dat_xxxx**

RUN_SETTINGS.txt

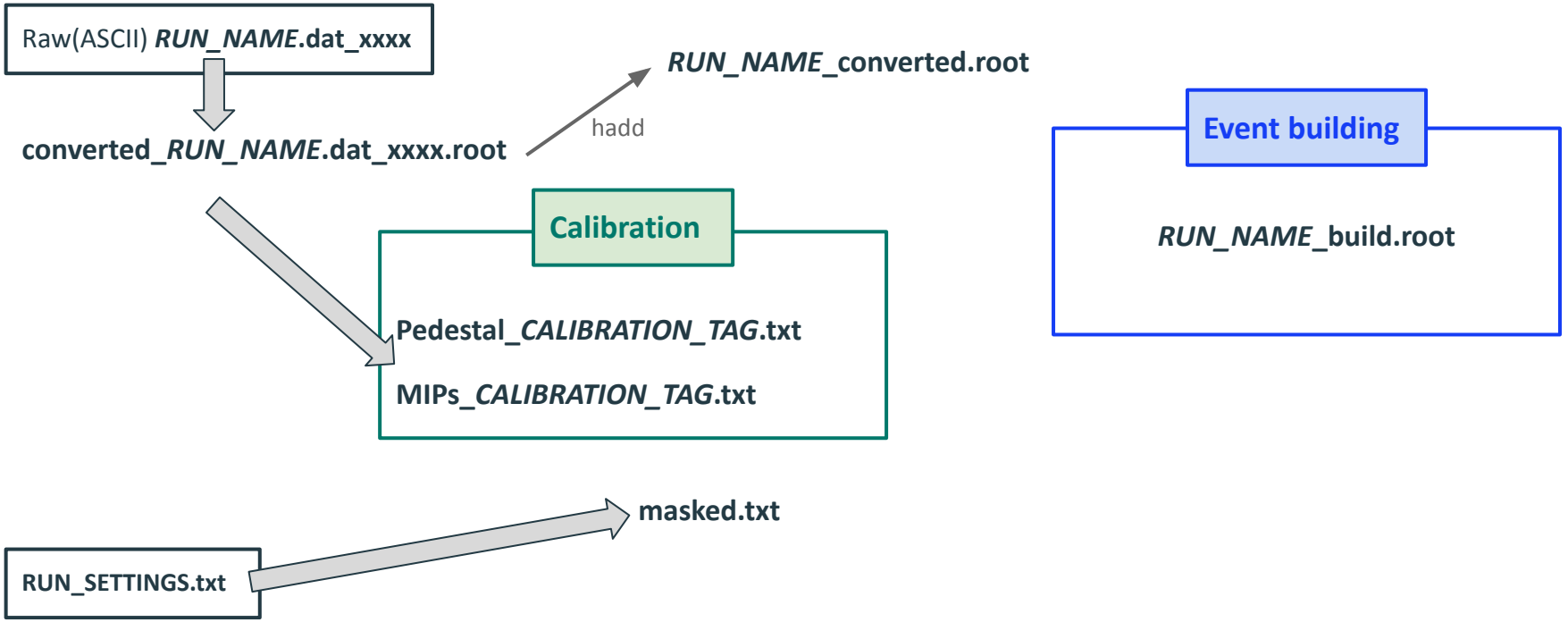
Event building pipeline



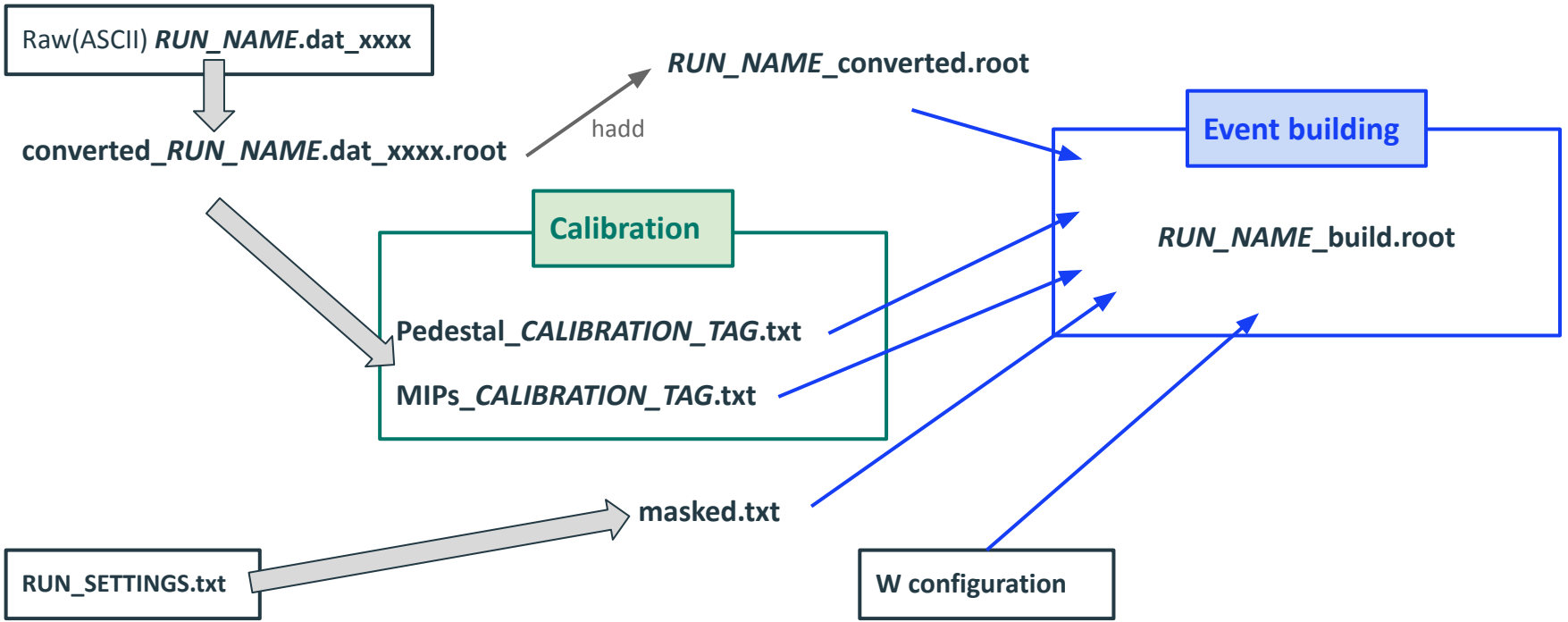
Event building pipeline



Event building pipeline

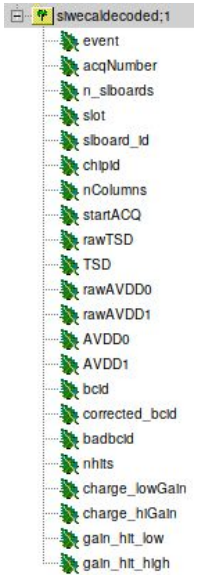


Event building pipeline



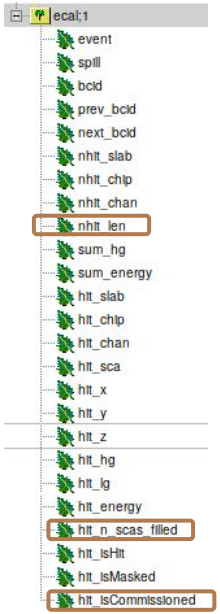
The rootfile contents

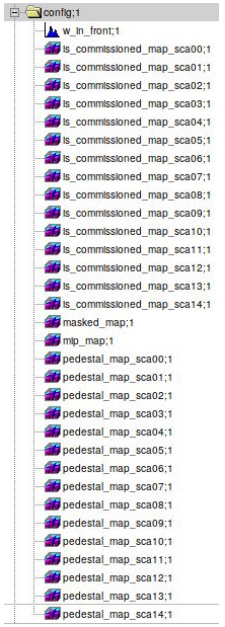
converted.root



old build.root

proposed new build.root





Why buildfiles?

Why buildfiles

- event
- spill
- bcid
- prev_bcid
- next_bcid
- nhit_slab
- nhit_chip
- nhit_chan
- nhit_len
- sum_hg
- sum_energy

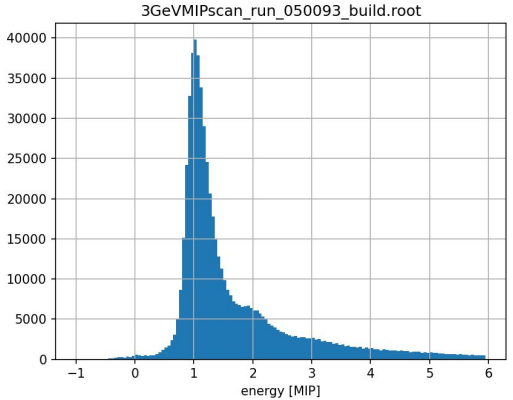
Event level information

- Useful for selection
- e.g. “nhit_slab > 10”

- hit_slab
- hit_chip
- hit_chan
- hit_sca
- hit_x
- hit_y
- hit_z
- hit_hg
- hit_lg
- hit_energy
- hit_n_scas_filled
- hit_isHit
- hit_isMasked
- hit_isCommissioned

hit arrays

- Know which hits belong together
- hit_energy: calibration applied. ==1 for average MIP deposit



Why buildfiles

- event
- spill
- bcid
- prev_bcid
- next_bcid
- nhit_slab
- nhit_chip
- nhit_chan
- nhit_len
- sum_hg
- sum_energy

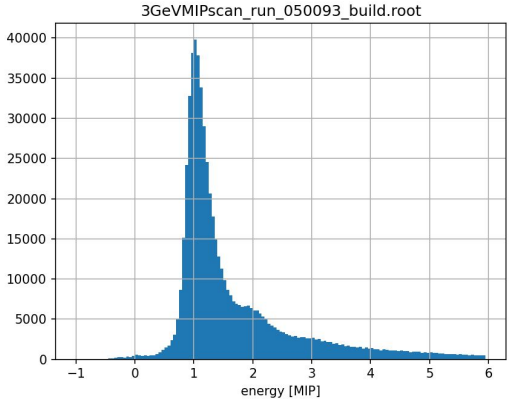
Event level information

- Useful for selection
- e.g. “nhit_slab > 10”

- hit_slab
- hit_chip
- hit_chan
- hit_sca
- hit_x
- hit_y
- hit_z
- hit_hg
- hit_lg
- hit_energy
- hit_n_scas_filled
- hit_isHit
- hit_isMasked
- hit_isCommissioned

hit arrays

- Know which hits belong together
- hit_en
  applied
  averag




```
[1]: import awkward as ak
     import matplotlib.pyplot as plt
     %matplotlib widget
     import numpy as np
     import uproot

     file = "3GeVMIPscan_run_050093_build.root"
     ar = uproot.open(file)["ecal"].arrays(filter_name="hit_*")
     n_slabs = uproot.open(file)["ecal"]["nhit_slab"].array()

[2]: hits = ar[n_slabs > 10]
     hits = hits[hits.hit_isHit == 1]

     energy = ak.flatten(hits.hit_energy).to_numpy()
     fig, ax = plt.subplots()
     ax.hist(energy, bins=np.arange(-1, 6, 0.05))
     ax.set_xlabel("energy [MIP]")
     ax.set_title(file)
     ax.grid()
```

How buildfiles?

Changes to eventbuilding.py

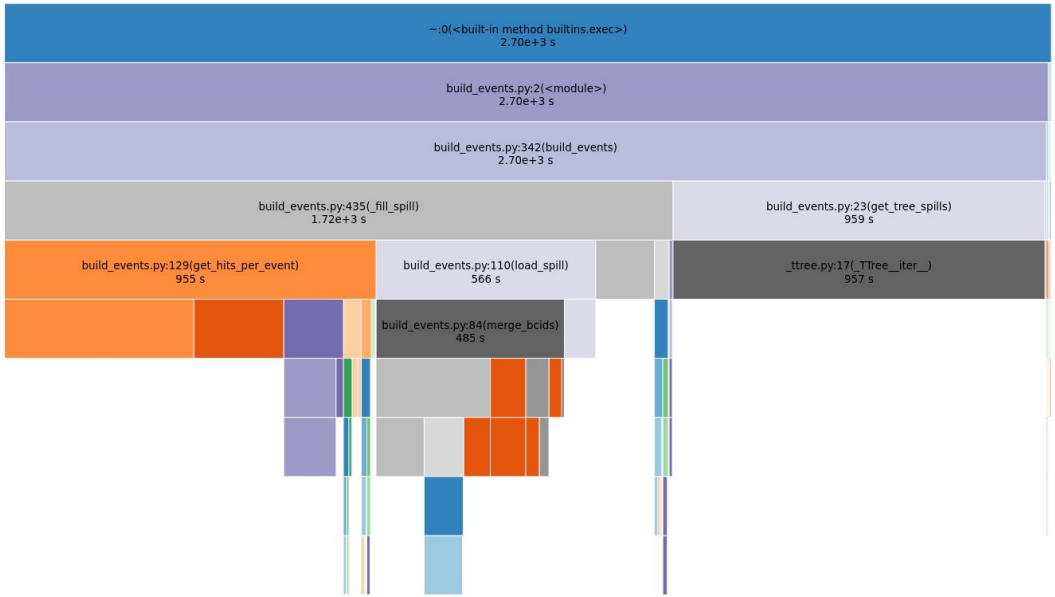
Required

- Input format of calibration .txt files changed
- Some bugs found along the way
  - But potentially replaced by fresh ones ;-)

Nice to have

- Performance improvements
  - Apply approximations for uncommissioned channels once at the start, instead of at each hit
  - Replace (critical) fo- loops by array manipulations. Especially remove channel for-loops
- Some new branches (we could still go for more)
- Save calibration histograms with the data

Is this fast enough?



Profiling shows:

- No obvious bottlenecks left
- > ⅓ of time spend in get_tree_spills (loading data to memory?)
  - I do not think this can be sped up
  - Even pure ROOT/C++ should not be faster here
- Upfront cost usually < 8s
  - Loading imports
  - Loading & improving calibration
  - Writing calibration histograms
- merge_bcids could potentially be faster
  - But why bother?

Where buildfiles?

Code and data location

The code is currently awaiting approval:
https://github.com/SiWECAL-TestBeam/SiWECAL-TB-analysis/pull/20

The data will be located at:

/eos/project-s/siw-ecal/TB2021-11/beamData/buildfiles

The data is currently only 1 file (~600MB):

- 3GeVMIPscan_run_050093_build.root

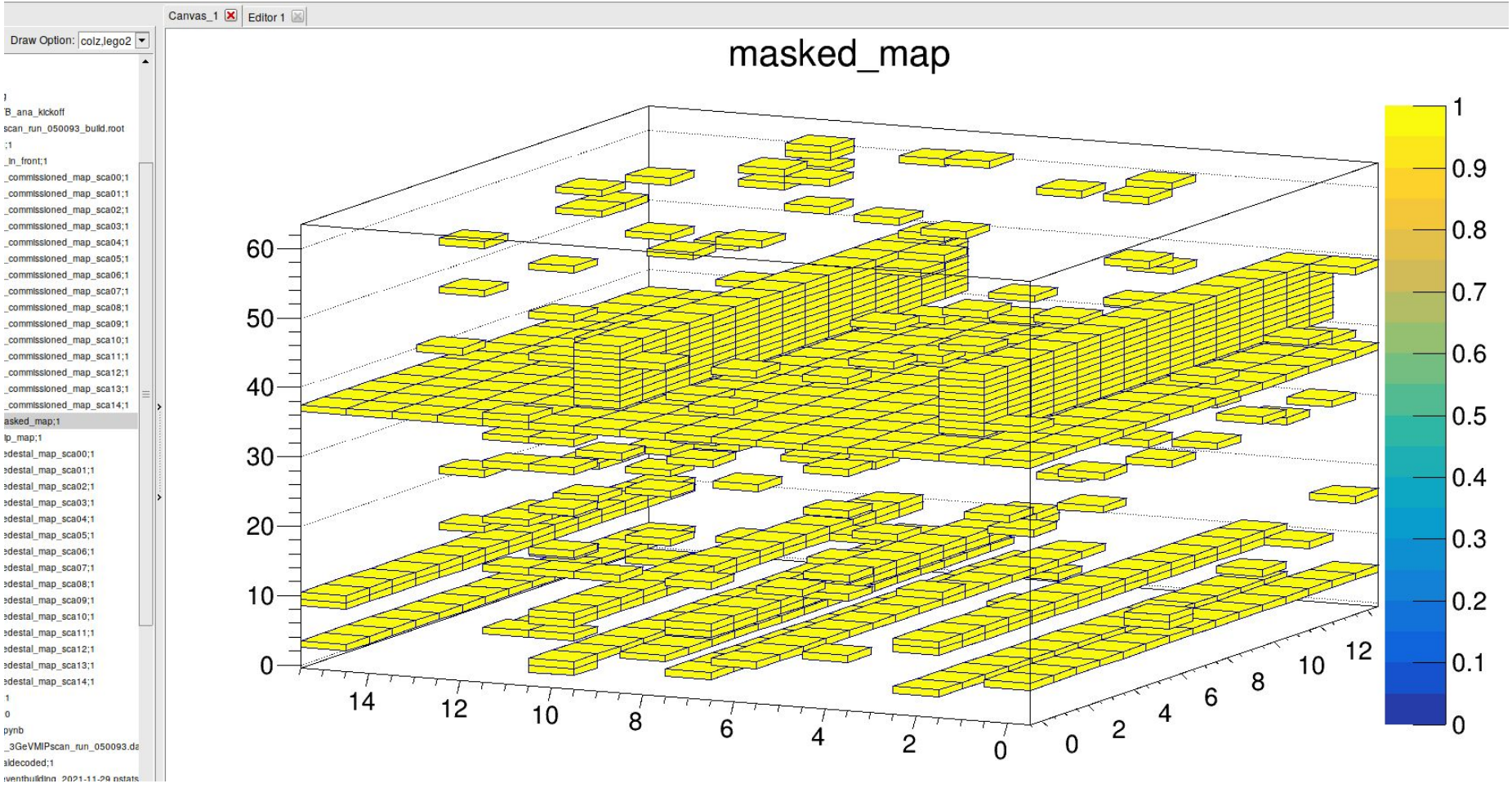
What is left to do?

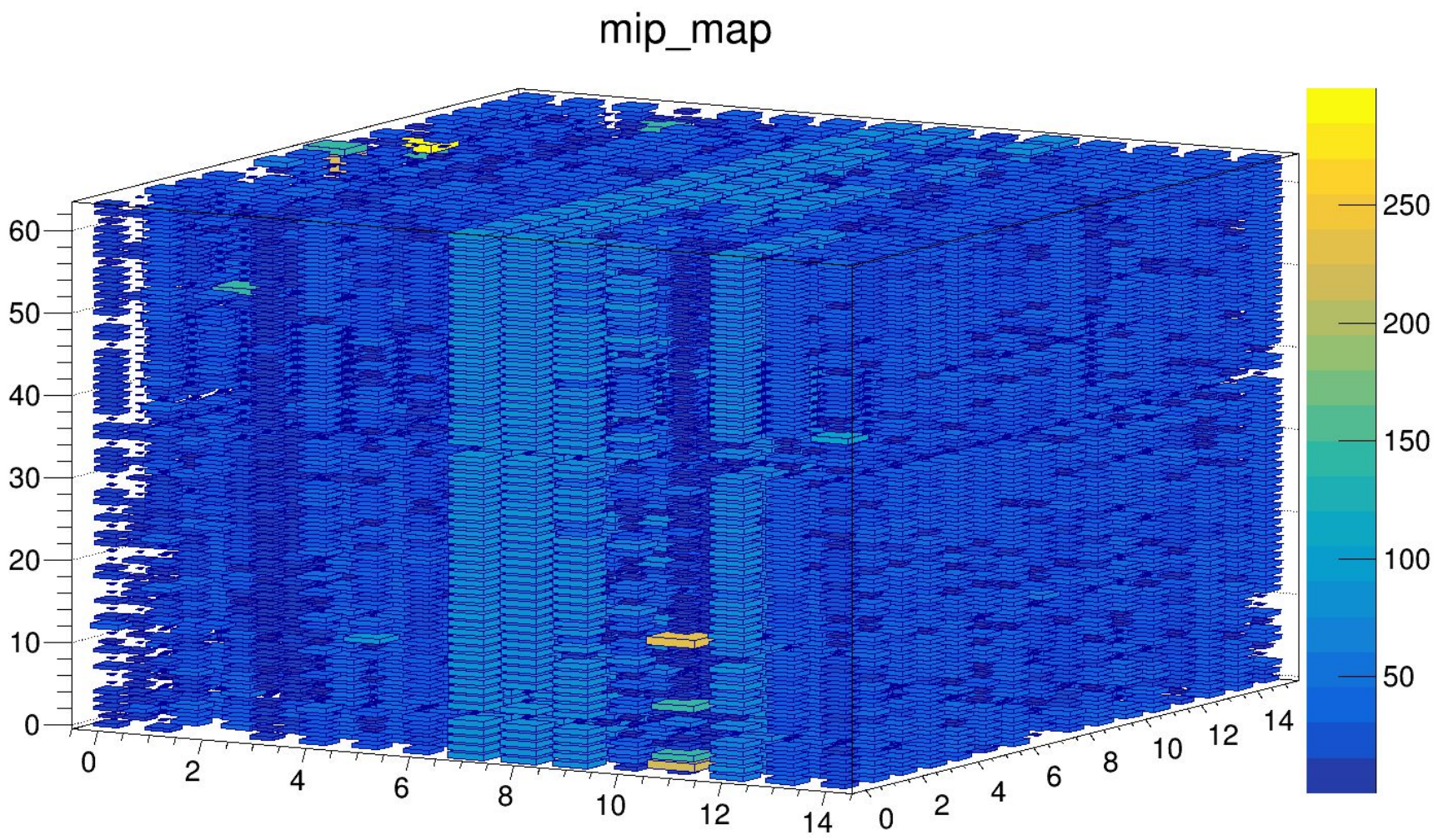
What is left to do

```
self.bcid_skip_noisy_acquisition_start = 50
self.bcid_merge_delta = 3
self.bcid_too_many_hits = 8000

self.pedestal_min_average = 200
self.pedestal_min_scas = 3
self.pedestal_min_value = 10
self.mip_cutoff = 0.5
```

- Verify the build.root file
  - There might be bugs
  - Also important: The steering defaults are not necessarily optimal
- Add/drop variables
  - hit_energy_lg: Adrián has produced calibration files - adding this should not be much coding
  - Do you like the idea of calibration histograms nect to the data?
  - Maybe only keep those channels in a hit where the gain_hit_high bit is 1? Currently: Keep all 64.
- Masking files can currently only be produced for ~½ of the data files
  - None of those are with Tungsten
  - Due to DESY_MOVING_TABLE line in DAQ. @Yuichi will fix this.


Canvas_1
Editor 1
Draw Option: colz,lego2
masked_map
60
50
40
30
20
10
0
14
12
10
8
6
4
2
0
0
2
4
6
8
10
12
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0


mip_map
60
50
40
30
20
10
0
0
2
4
6
8
10
12
14
0
2
4
6
8
10
12
14
250
200
150
100
50