First-pass clustering evaluation in SBND

Ewerton Belchior

March 20, 2025

Goal and procedure

Goal:

- First-pass hand scanning to identify potential issues like ghosts, under(over)-clustering and track gaps;
- Initial quantitative evaluation of clustering performance.

Evaluation steps:

1. Use a KDTree to query for nearest neighbors between truth 3-D points (simEnergyDepos) and reco 3-D points (output from clustering). A given match is considered successful if d < 2 cm, where d is the distance to the nearest neighbor.
2. For this evaluation, truth clusters are required to have energy > 50 MeV
3. Points of a given truth cluster can be matched to points of many different reco clusters. We want to evaluate what fraction of a given truth cluster is matched to each reco cluster (and vice-versa).
4. Define two metrics to quantify clustering performance: efficiency and purity:

Efficiency(truth cluster) = matched truth cluster charge / total truth cluster charge

Purity(reco cluster) = reco cluster points matched / reco cluster points

1. Events are said to have “bad” clusters if it has at least one cluster with either efficiency < 50% or purity < 50%. Those events are selected for inspection on next slides.

Data samples

Bee displays:

- MC: https://www.phy.bnl.gov/twister/bee/set/1f269c02-e266-414a-8c3f-831e785306a4/event/list/
- Data (run 14766): https://www.phy.bnl.gov/twister/bee/set/5caef909-f5b9-4e57-be0a-92076af514c2/event/list/
- Data (run 18005): https://www.phy.bnl.gov/twister/bee/set/67ed0429-e8c9-4979-9b57-7128acb01e74/event/list/

Data samples:

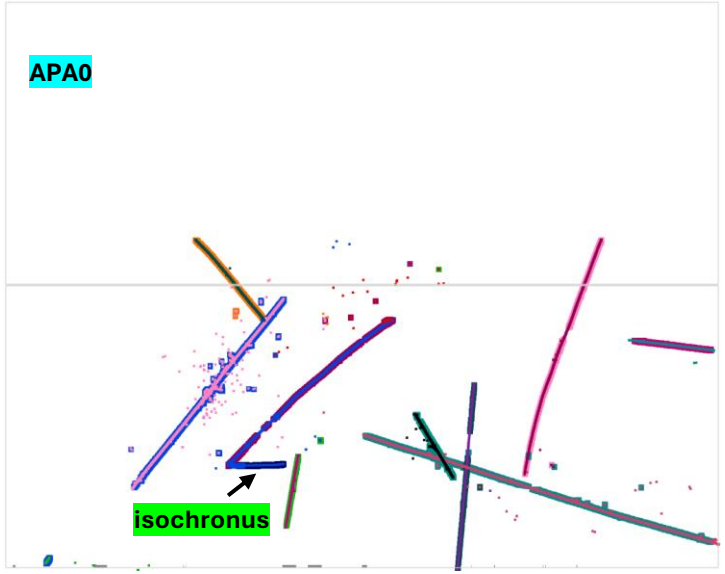
- MC: /exp/sbnd/data/users/ebatista/wirecell-clustering/eval/bnb_nu_cosmic_additional/sigproc.root
- Data (run 14766): /exp/sbnd/data/users/ebatista/wirecell-clustering/eval/run14766/sigproc.root
- Data (run 18005): /exp/sbnd/data/users/ebatista/wirecell-clustering/eval/run18005/sigproc.root

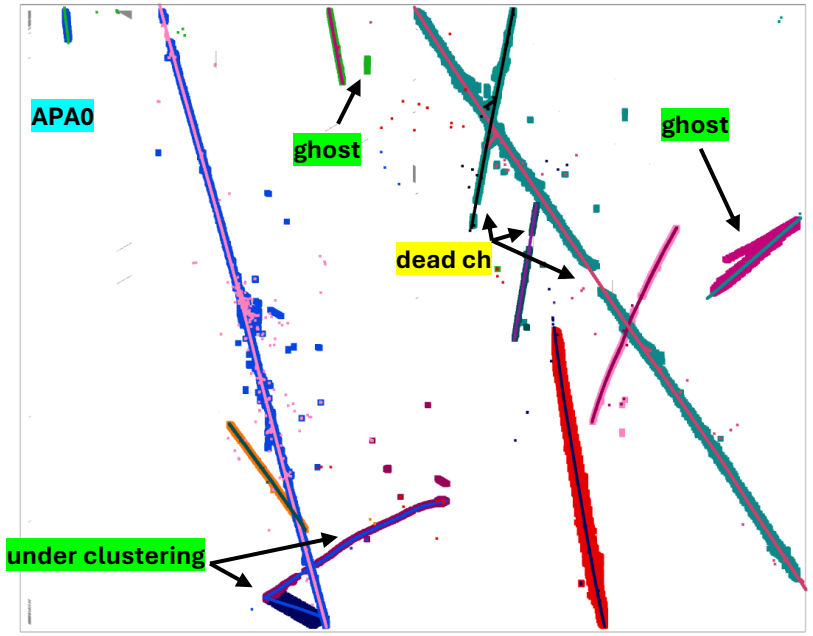
- **sbndcode version for clustering:**
- MC: v09_91_02
- Data (run 14766): v10_04_03
- Data (run 18005): v10_04_03

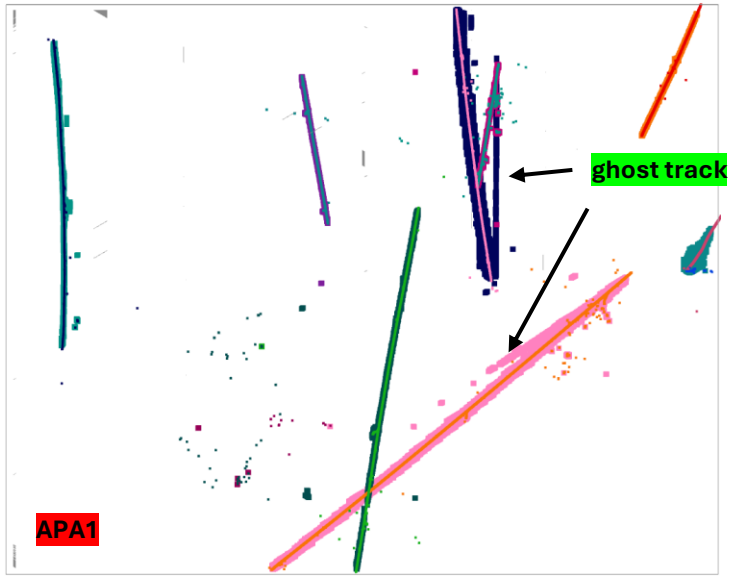
Handscan (MC)

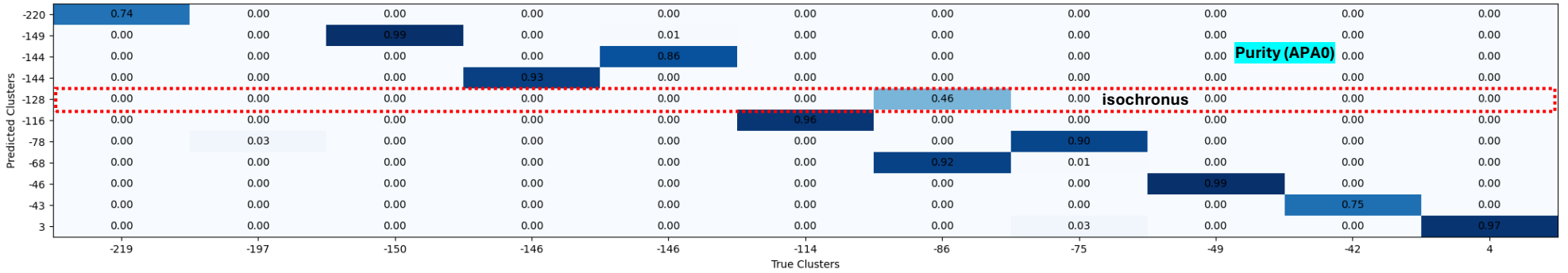
run 1, event 10

https://www.phy.bnl.gov/twister/bee/set/1f269c02-e266-414a-8c3f-831e785306a4/event/9/





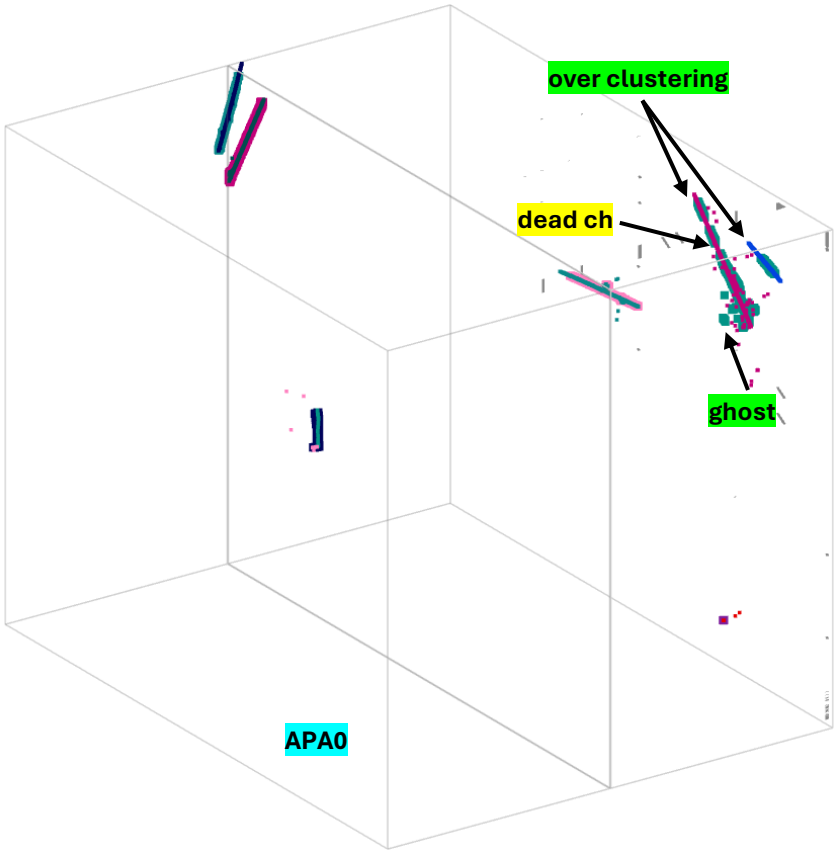


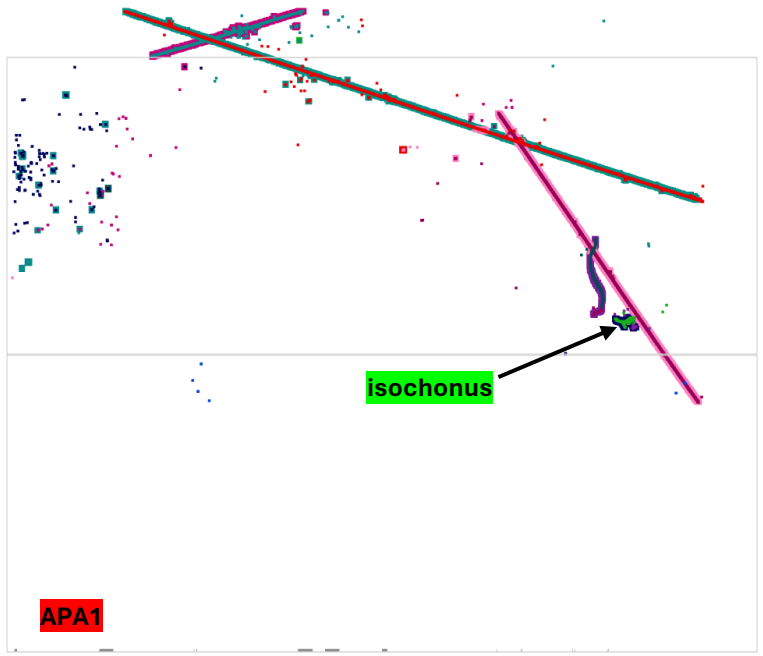


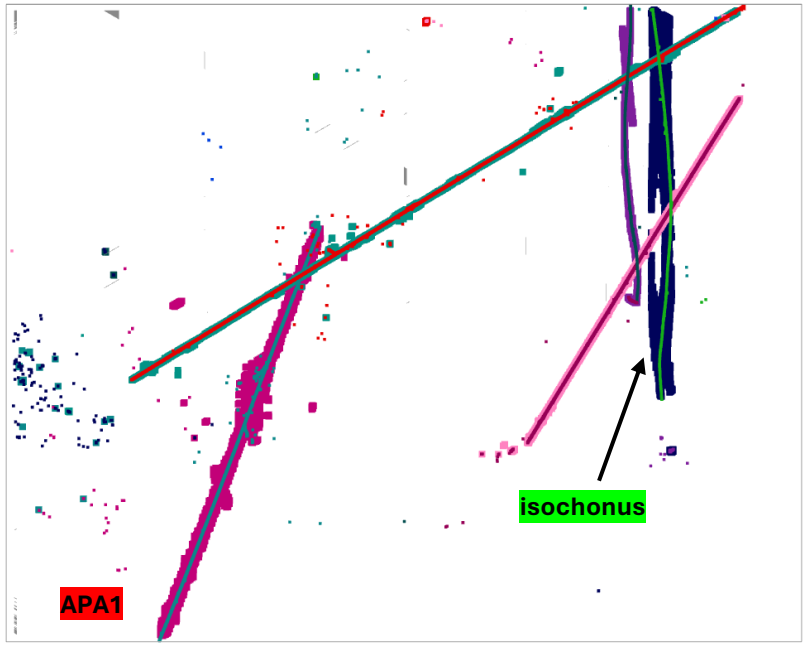
Handscan (MC)

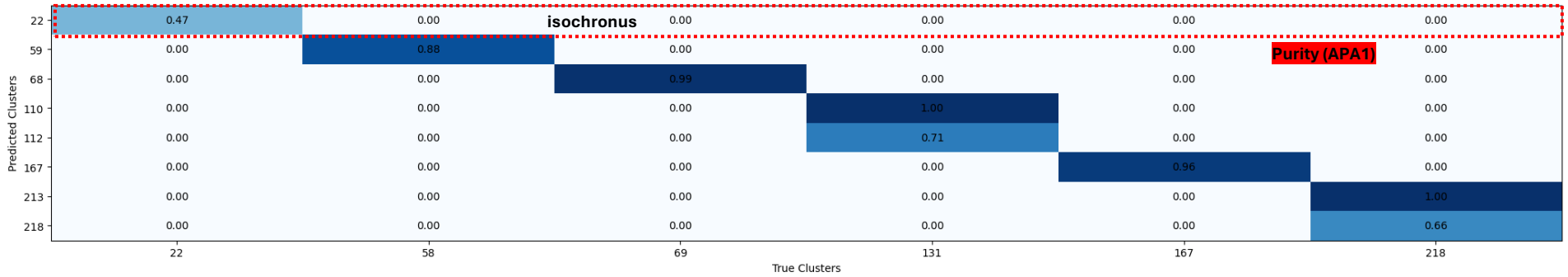
https://www.phy.bnl.gov/twister/bee/set/1f269c02-e266-414a-8c3f-831e785306a4/event/6/

run 1, event 7





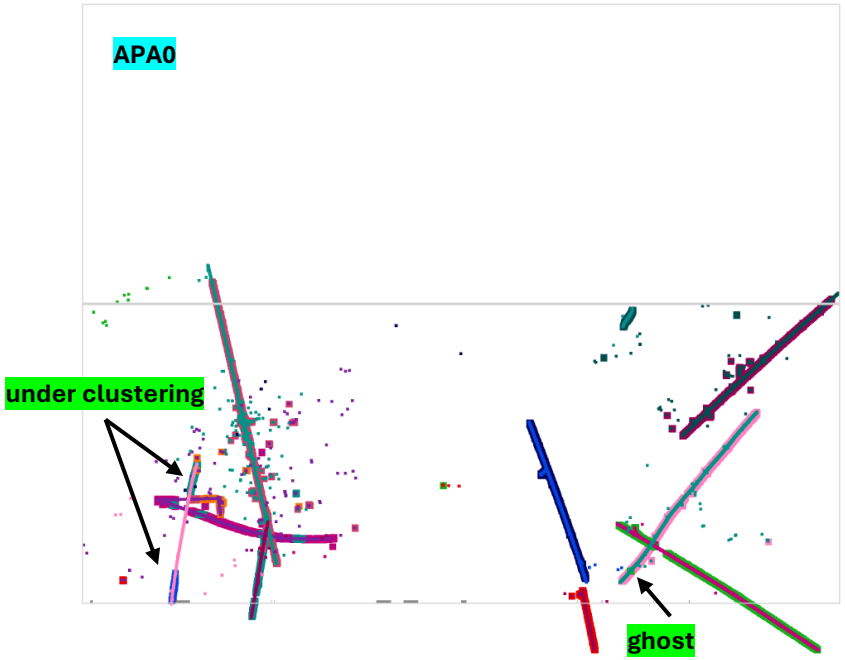


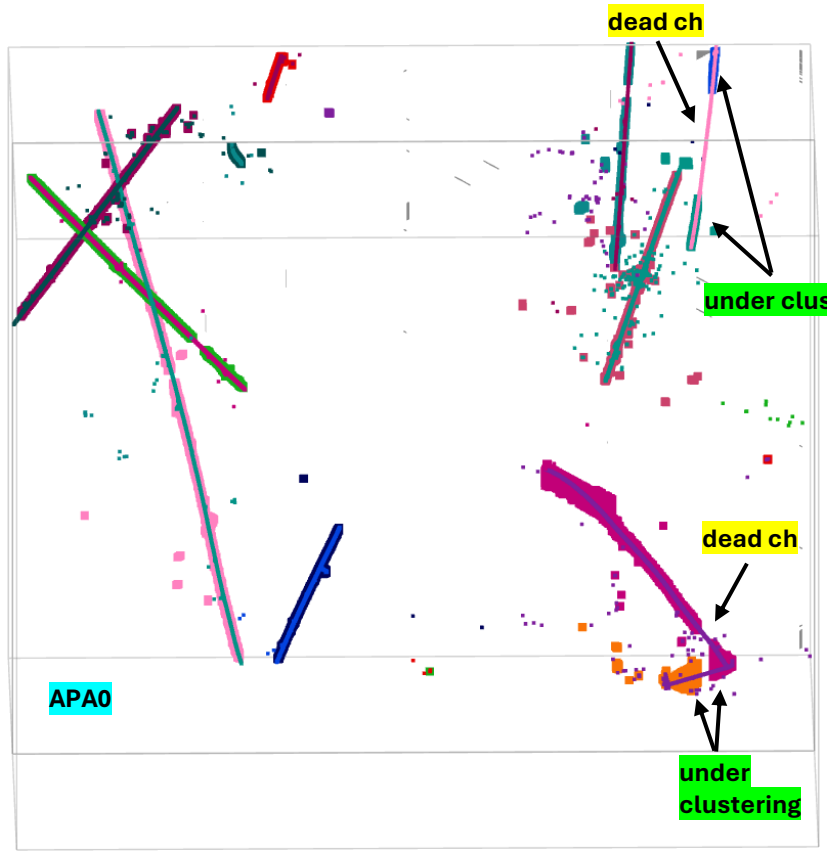


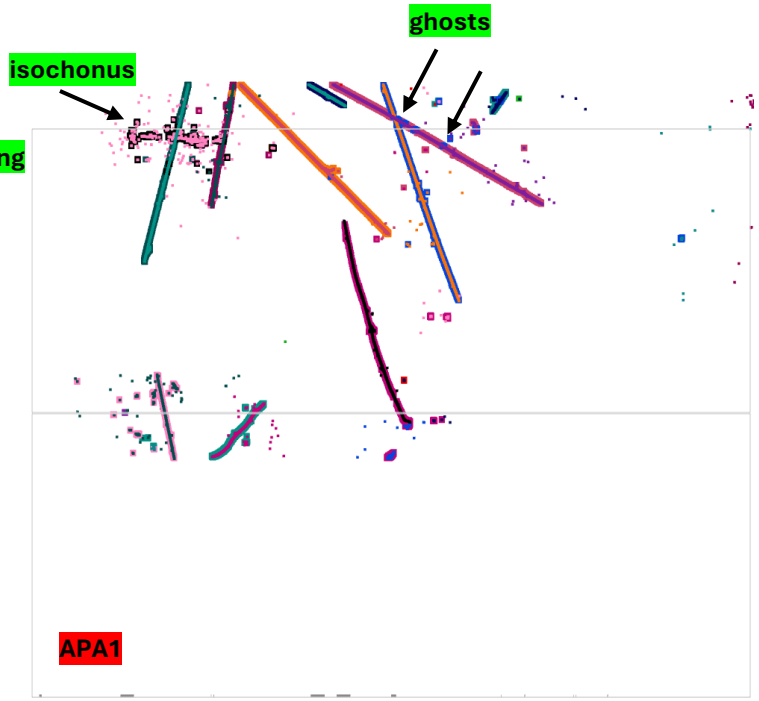
Handscan (MC)

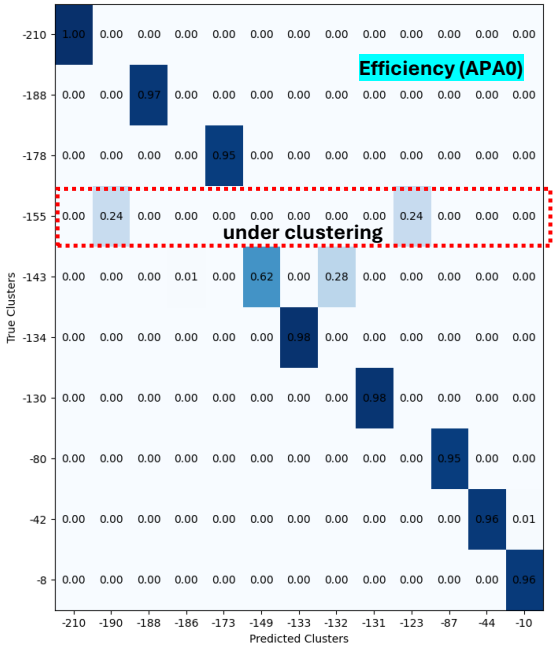
https://www.phy.bnl.gov/twister/bee/set/1f269c02-e266-414a-8c3f-831e785306a4/event/2/

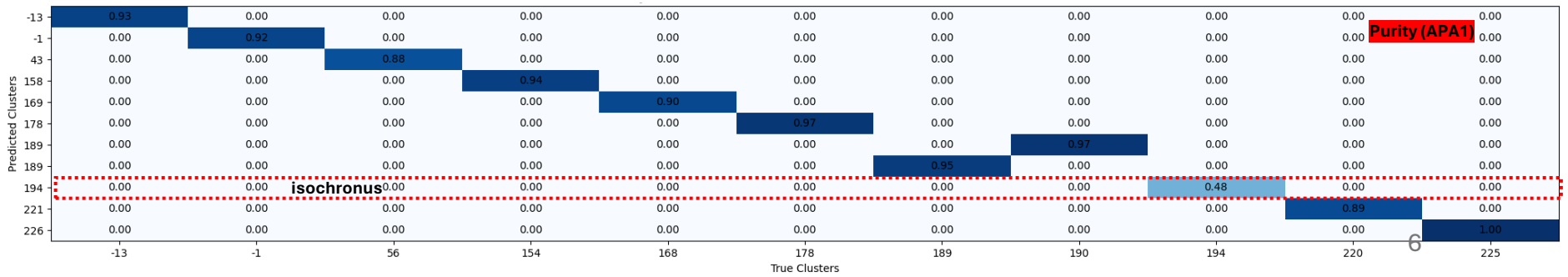
run 1, event 3







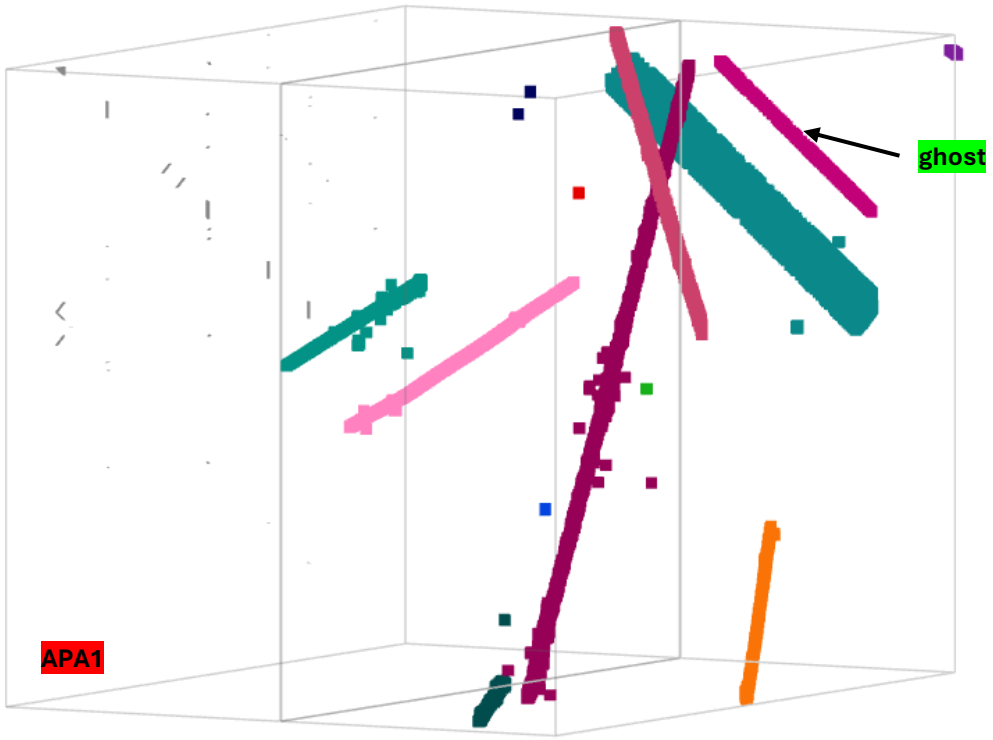




Handscan (data: run 18005)

run 18005, event 36453

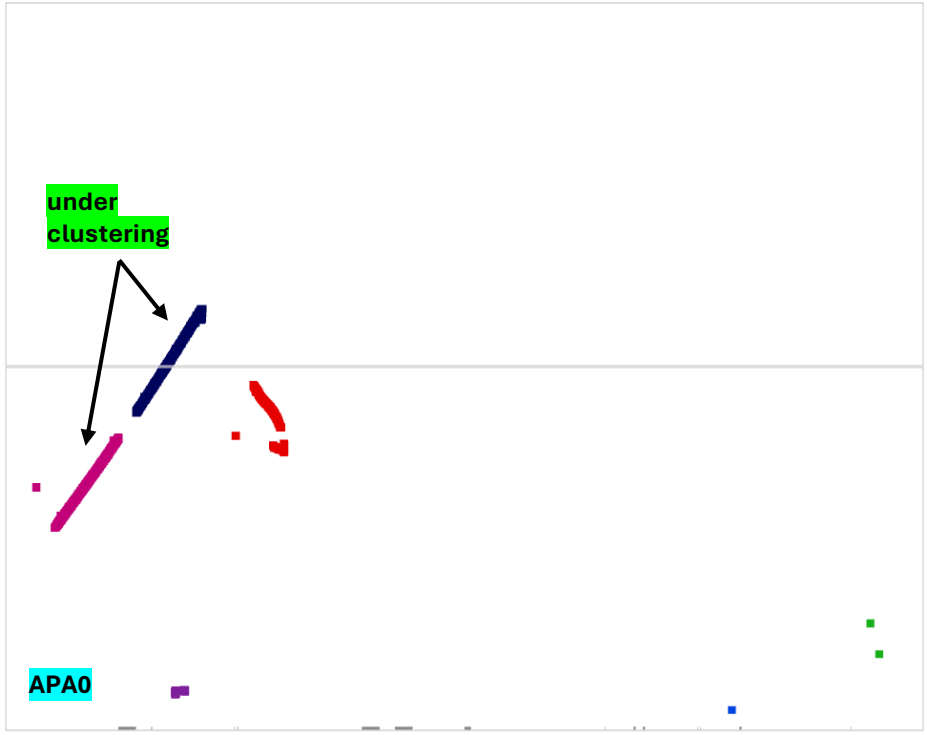
https://www.phy.bnl.gov/twister/bee/set/67ed0429-e8c9-4979-9b57-7128acb01e74/event/10/



Handscan (data: run 18005)

run 18005, event 36145

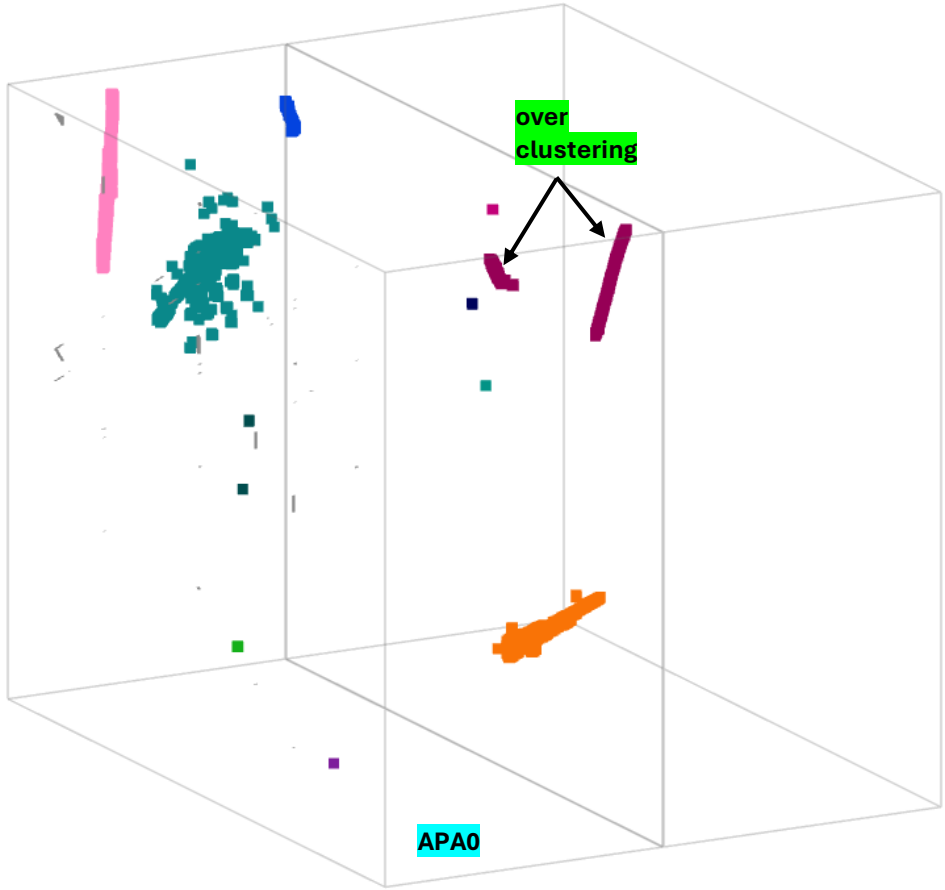
https://www.phy.bnl.gov/twister/bee/set/67ed0429-e8c9-4979-9b57-7128acb01e74/event/5/



Handscan (data: run 18005)

run 18005, event 36321

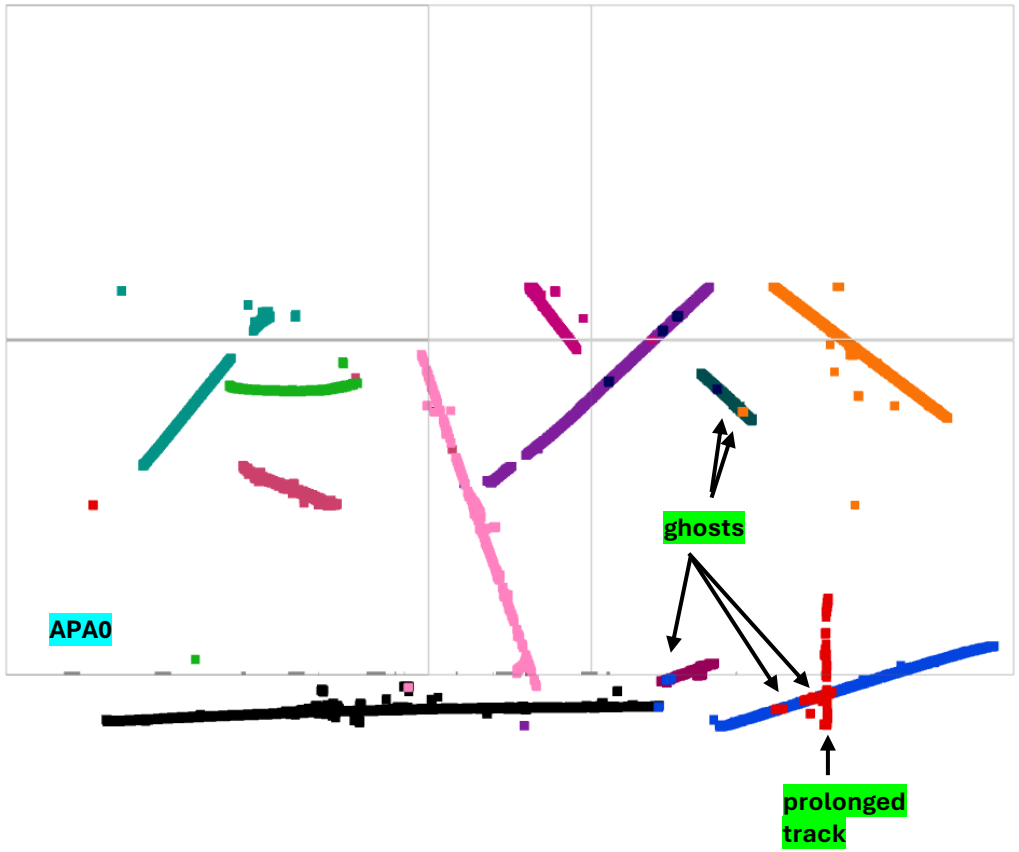
https://www.phy.bnl.gov/twister/bee/set/67ed0429-e8c9-4979-9b57-7128acb01e74/event/7/

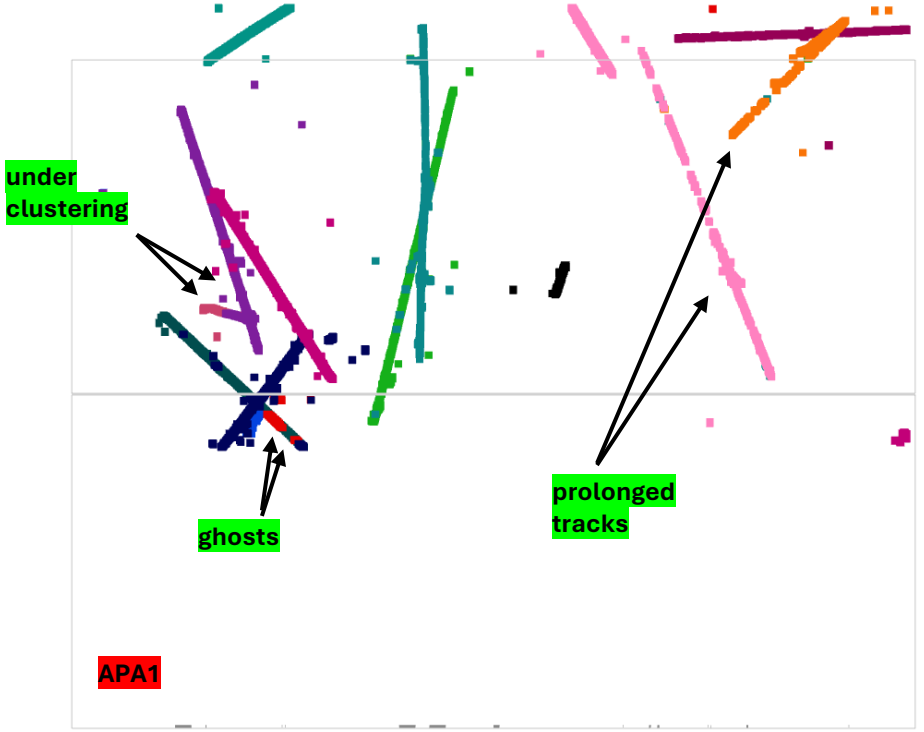


Handscan (data: run 18005)

run 18005, event 36123

https://www.phy.bnl.gov/twister/bee/set/67ed0429-e8c9-4979-9b57-7128acb01e74/event/4/





Summary

Run	Events verified*	Total APAs verified	at least one ghost in APA**	under clustering counts	over clustering counts
14766	6	12	5	1	0
18005	11	22	5	2	1
MC	21	42	12	6	3

Rates:

- At least one ghost in APA: 29% (data), 28% (MC)
- Under-clustering per APA: 9% (data), 14% (MC)
- Over-clustering per APA: 3% (data), 7% (MC)

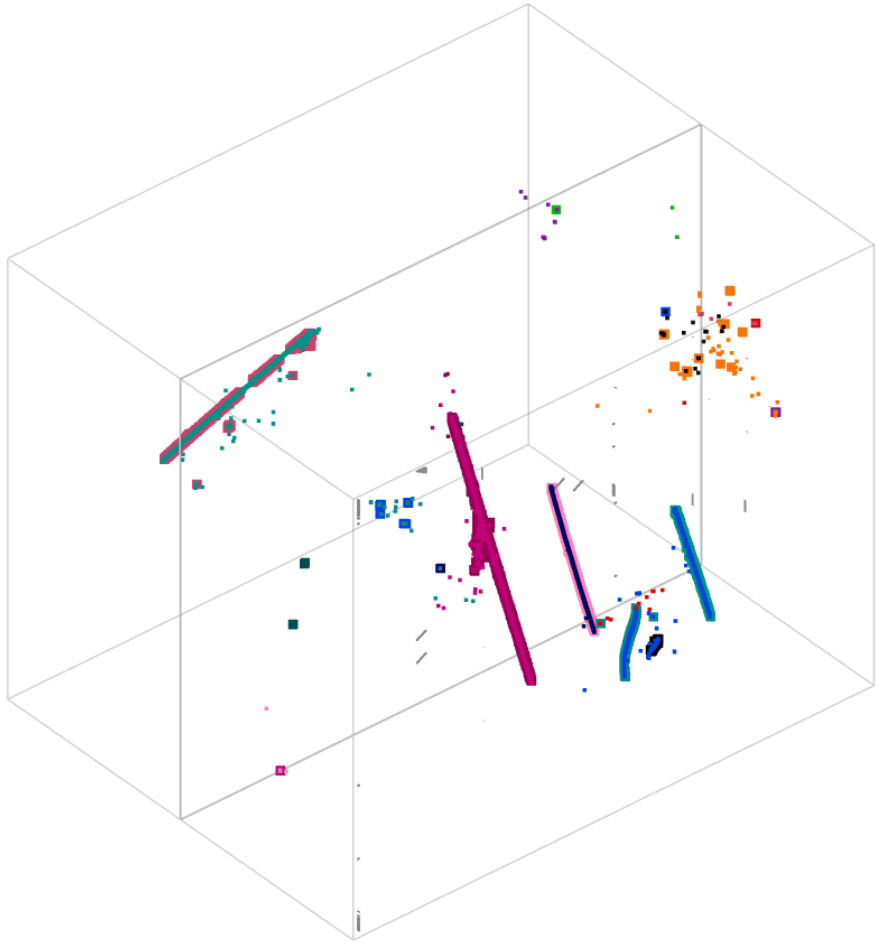
* For MC, 21 events were used to verify if efficiency > 50%, purity > 50%. From those, only 10 passed the cutoff and those are the ones used in the counting of ghosts/under-clustering/over-clustering.

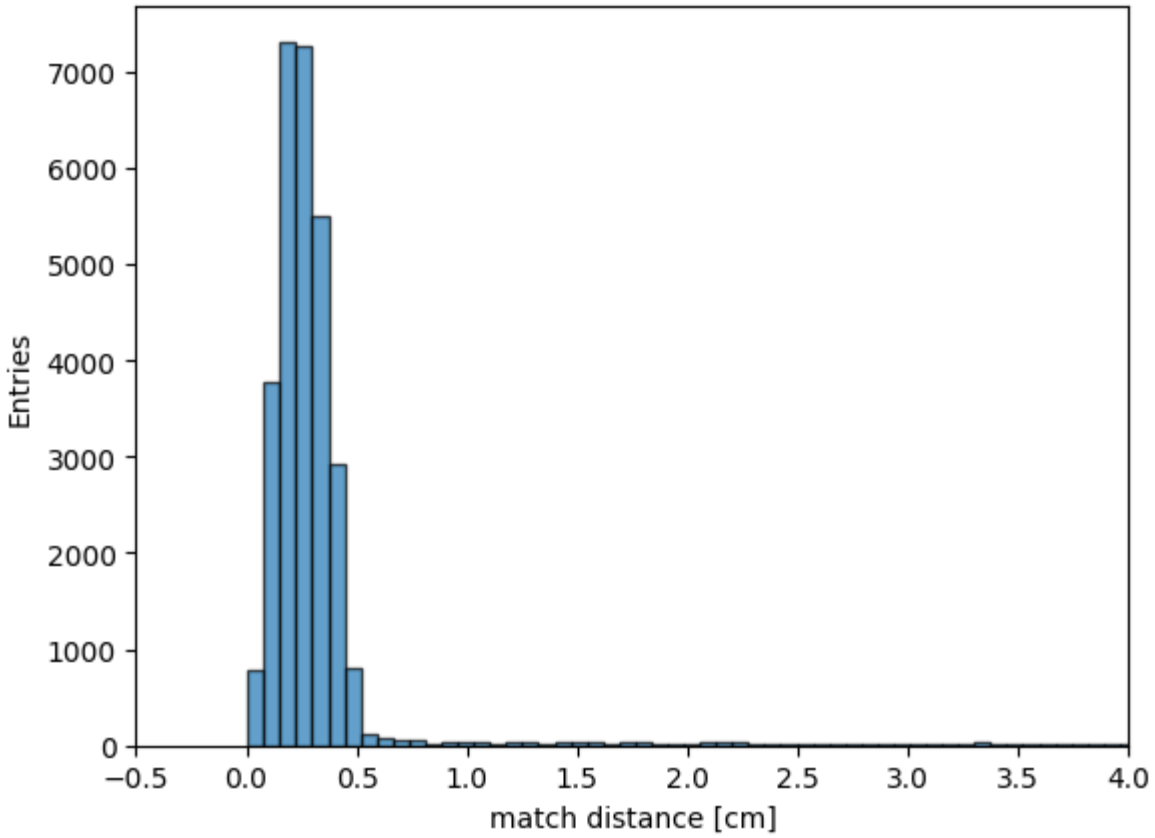
** Since MC have more dotted-like activity, it turns difficult to identify all dotted-like ghosts, so that is a rough counting for MC.

Backup

Matching distance (MC)

https://www.phy.bnl.gov/twister/bee/set/1f269c02-e266-414a-8c3f-831e785306a4/event/0/





Sample = bnb_nu_cosmic_additional
APA= apa0
event= 0
cutoff (efficiency and purity) > 0.5
min_cluster_energy= 50 MeV

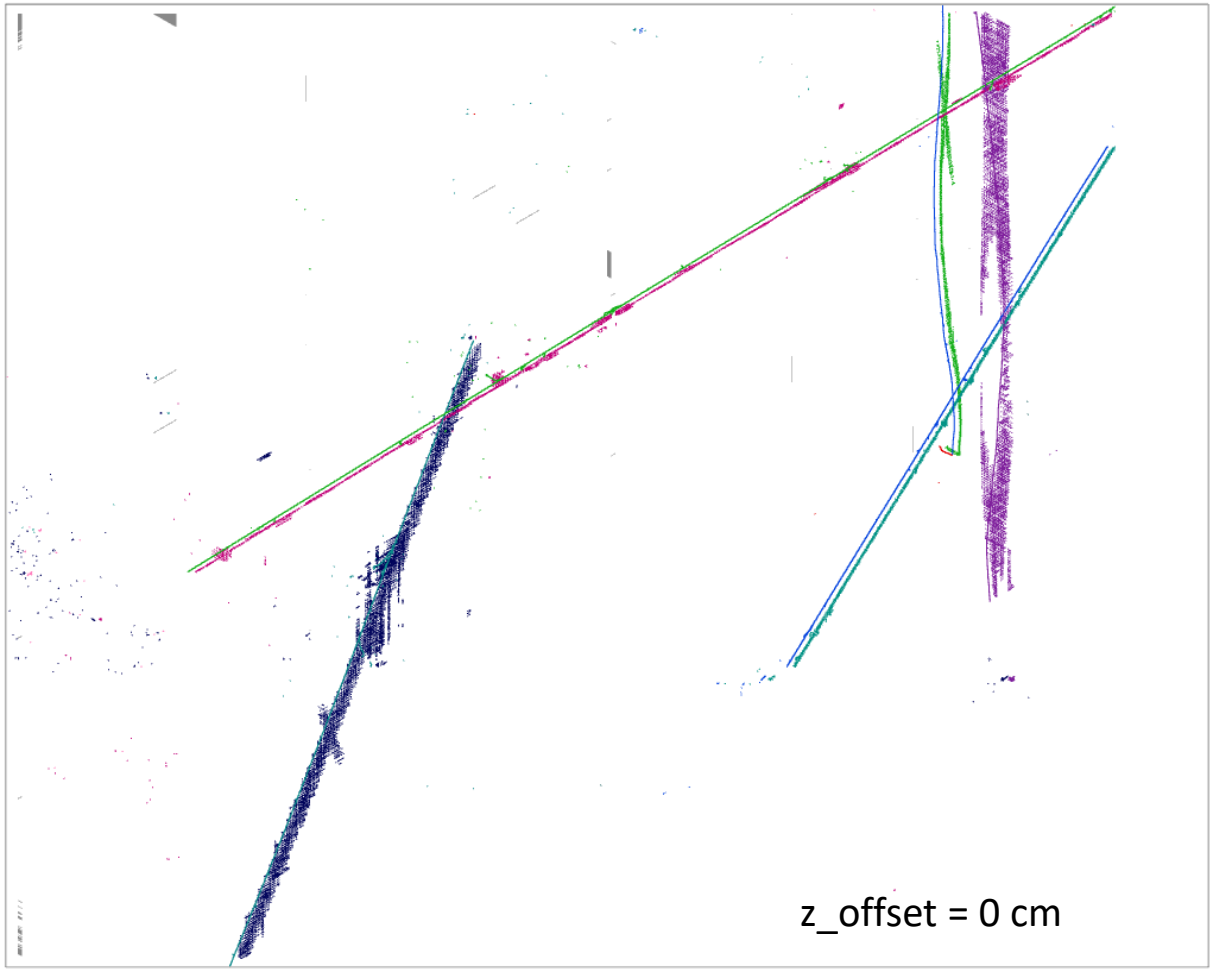
Z-offset (MC)

Sample = bnb_nu_cosmic_additional
APA= apa1
event= 6

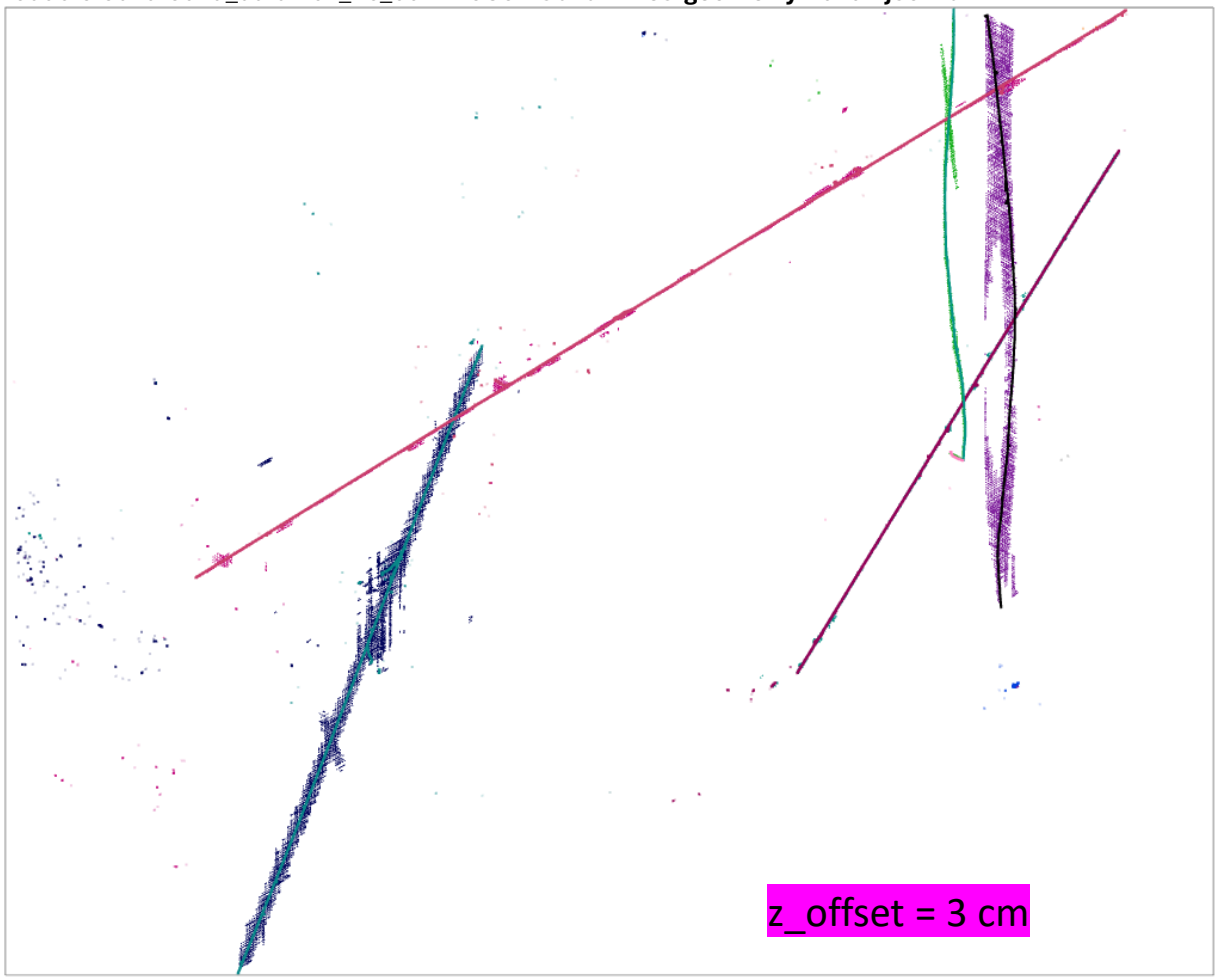
- **Clustering step**: sbndcode v09_91_02 (sbnd_data v01_24_00) => using this setup: https://github.com/HaiwangYu/wcp-porting-img/blob/5129b35992fcdc0c78e9538caf371833ed1fb431/sbnd/setup.sh#L1

Apptainer> grep -r sbnd-wires-geometry-v0200.json.bz2 /exp/sbnd/app/users/yuhw/wct-cfg/cfg
/exp/sbnd/app/users/yuhw/wct-cfg/cfg/pgrapher/experiment/sbnd/params.jsonnet: wires: "**sbnd-wires-geometry-v0200.json.bz2**",

- **SP step**: sbndcode v09_93_01_02 (sbnd_data v01_26_00)
/cvmfs/sbnd.opensciencegrid.org/products/sbnd/sbnd_data/v01_26_00/WireCell/**sbnd-wires-geometry-v0202.json.bz2**



https://www.phy.bnl.gov/twister/bee/set/b7cddce7-2b4d-4108-a5b8-768b609ba29e/event/6/



https://www.phy.bnl.gov/twister/bee/set/ad653a48-8d01-4ddb-a9a6-631c151a5cae/event/0/

Z-offset (MC)

Sample = bnb_nu_cosmic_additional
APA= apa0
event= 17

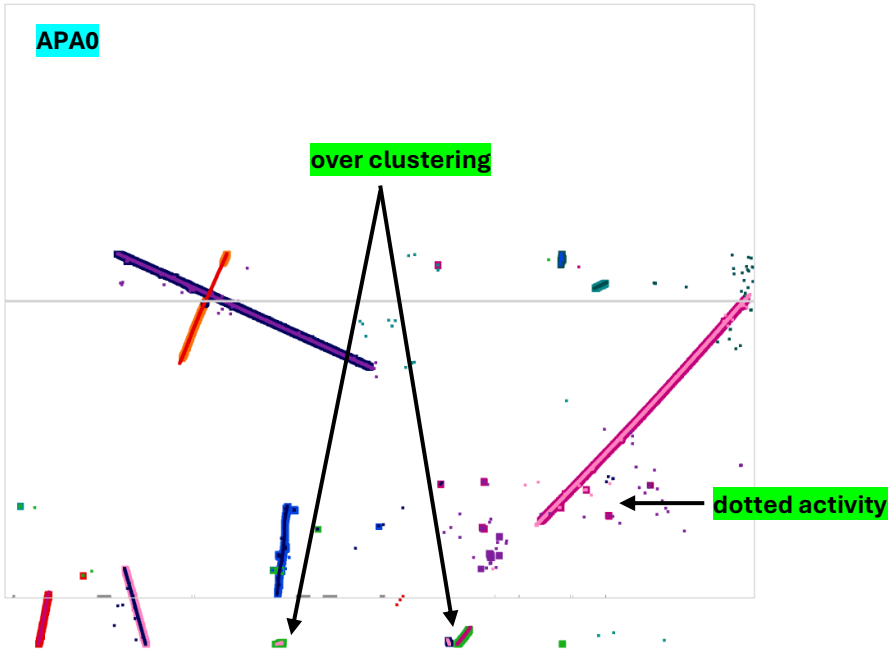
https://www.phy.bnl.gov/twister/bee/set/e47321bc-e6c0-4bff-84d9-f69127f55d0c/event/0/

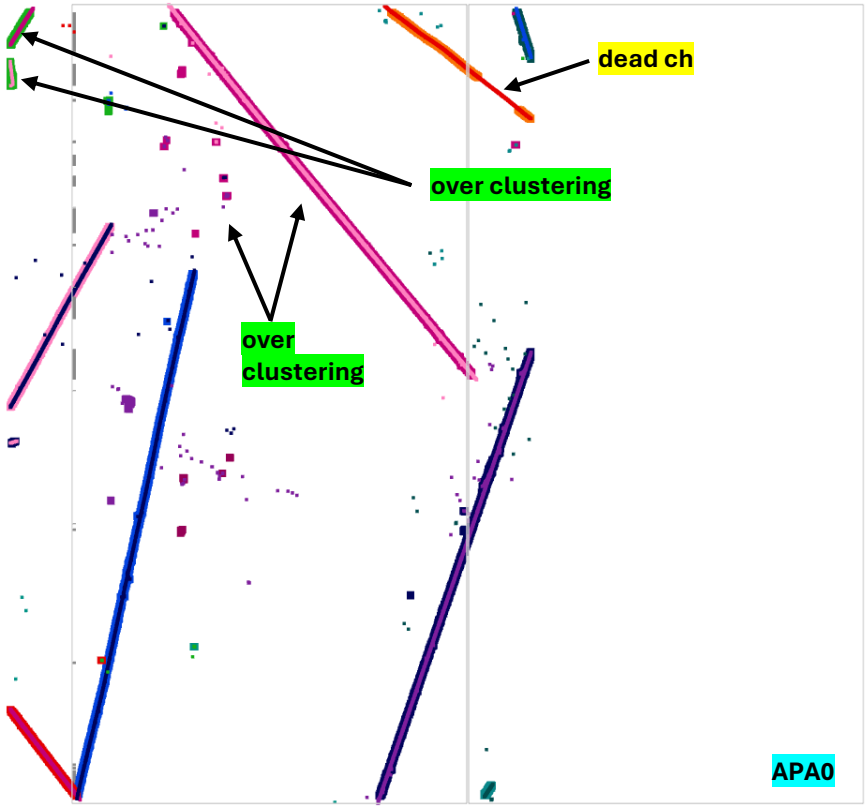
https://www.phy.bnl.gov/twister/bee/set/70e73d1f-f30b-456d-be12-b9b393e80219/event/0/

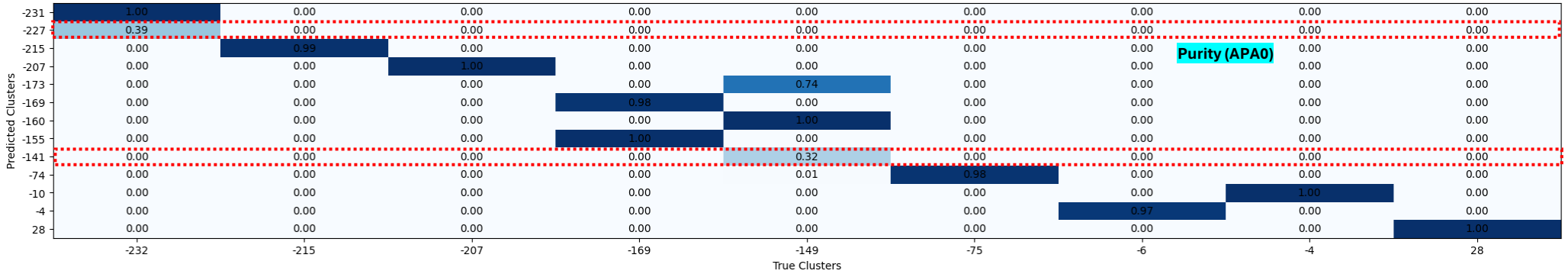
Handscan (MC)

https://www.phy.bnl.gov/twister/bee/set/1f269c02-e266-414a-8c3f-831e785306a4/event/7/

run 1, event 8



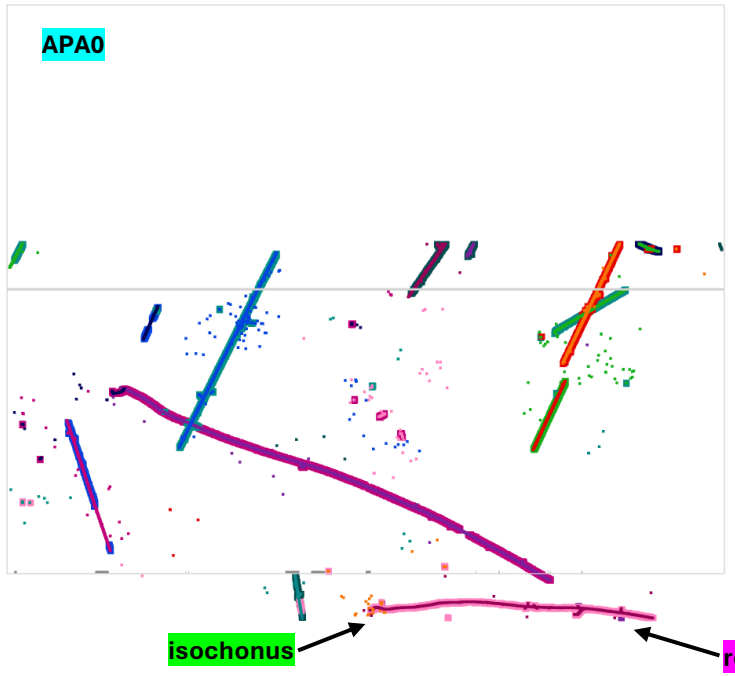


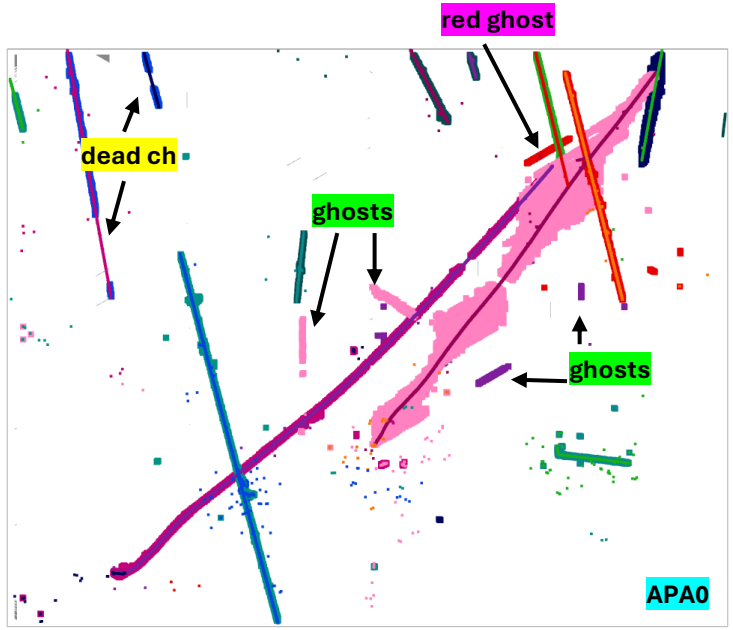


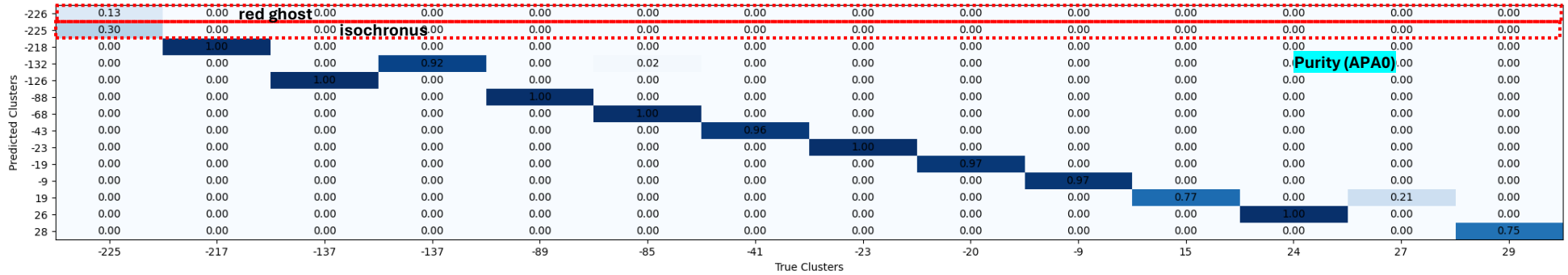
Handscan (MC)

run 1, event 11

https://www.phy.bnl.gov/twister/bee/set/1f269c02-e266-414a-8c3f-831e785306a4/event/10/



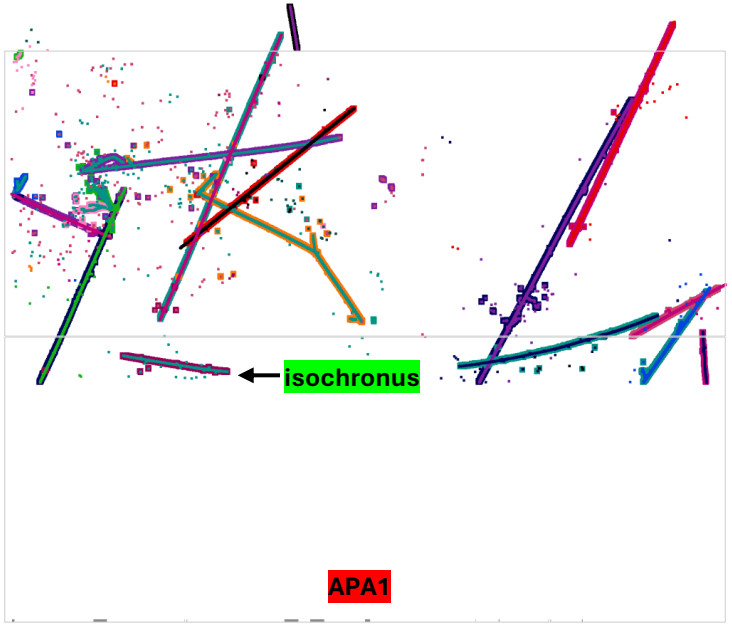


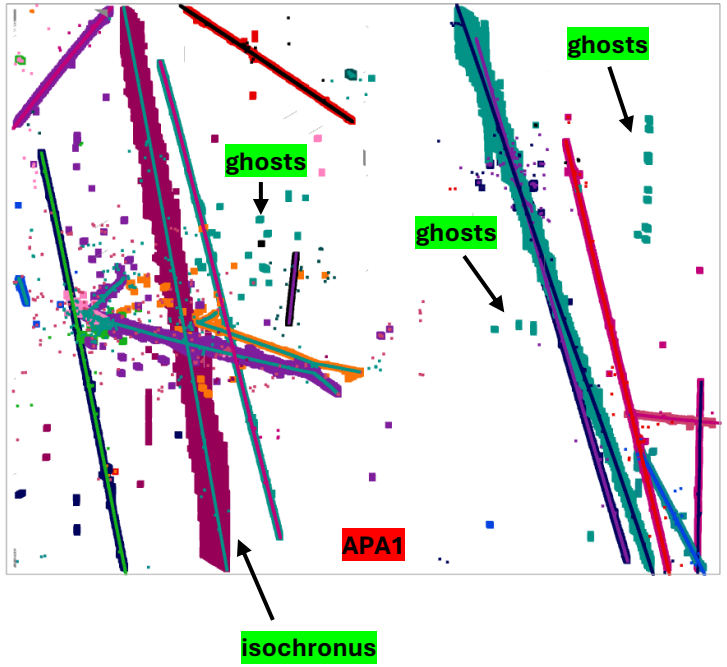


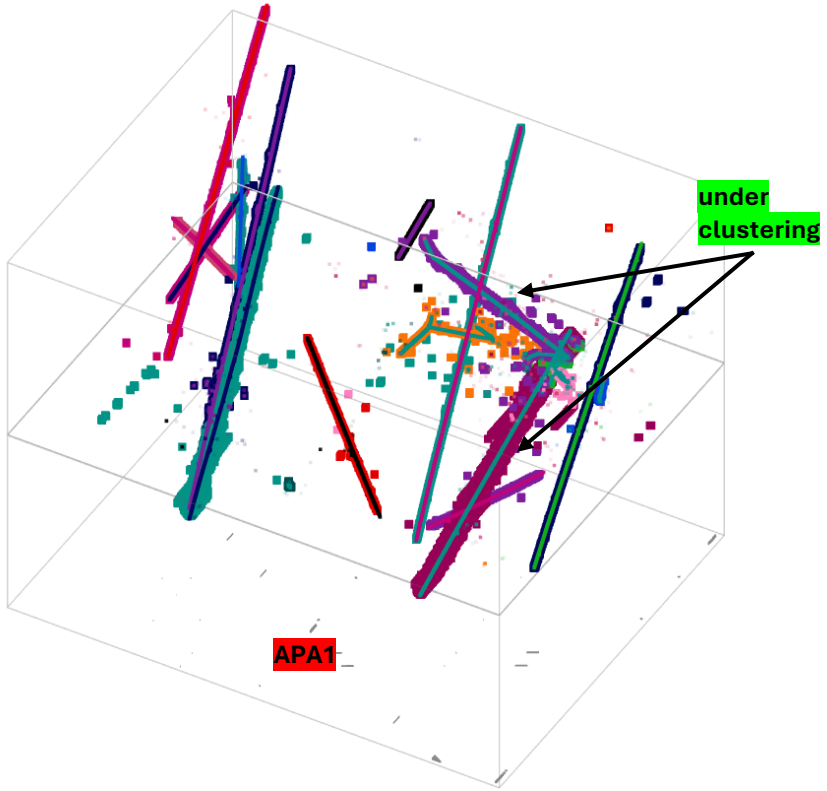
Handscan (MC)

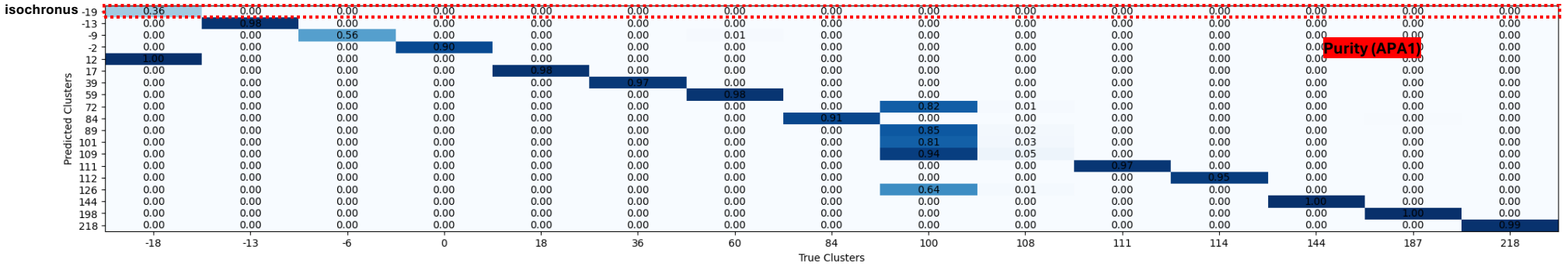
run 1, event 11

https://www.phy.bnl.gov/twister/bee/set/1f269c02-e266-414a-8c3f-831e785306a4/event/10/





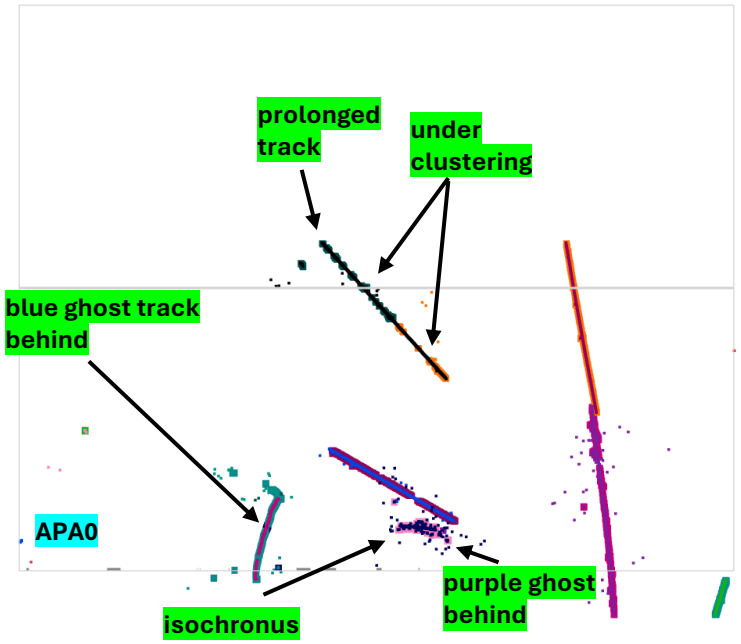


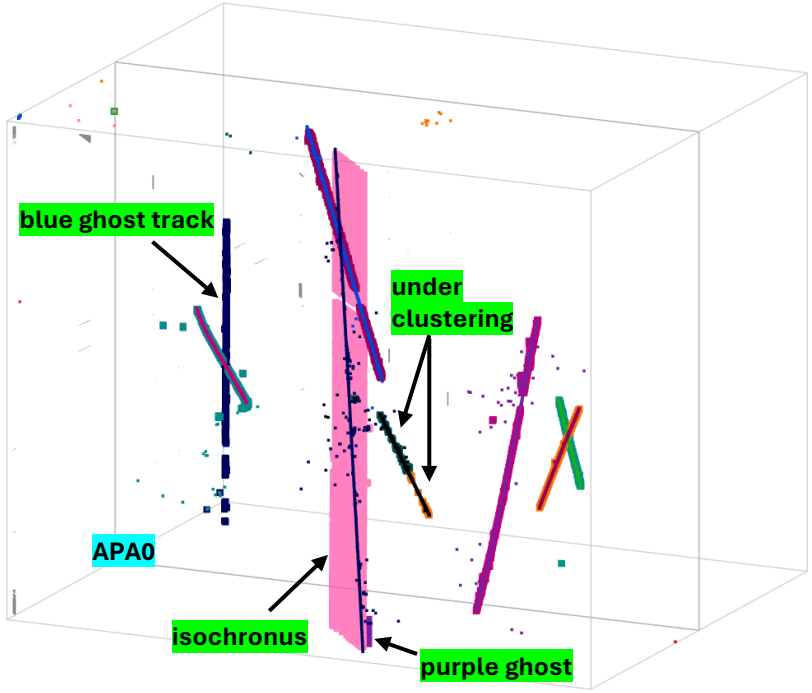


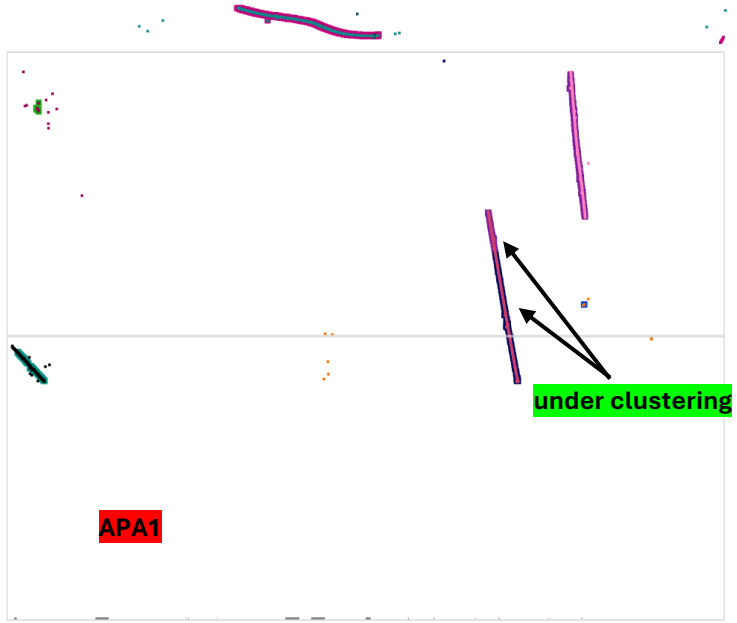
Handscan (MC)

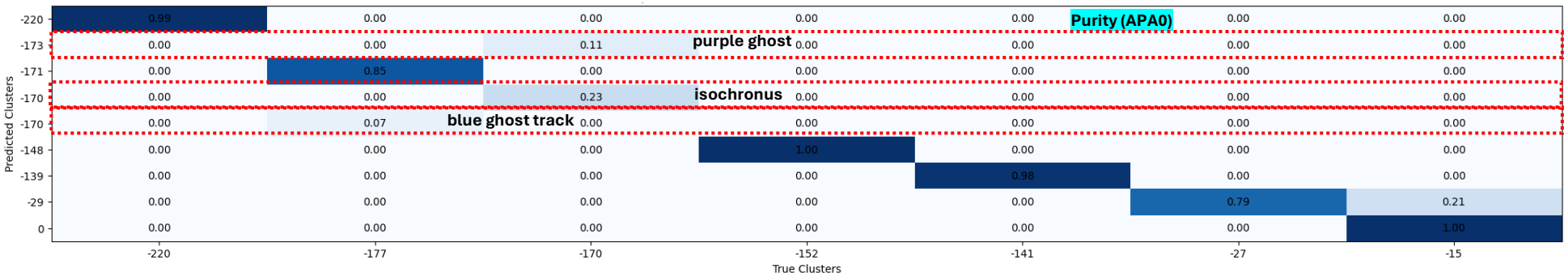
run 1, event 14

https://www.phy.bnl.gov/twister/bee/set/1f269c02-e266-414a-8c3f-831e785306a4/event/13/





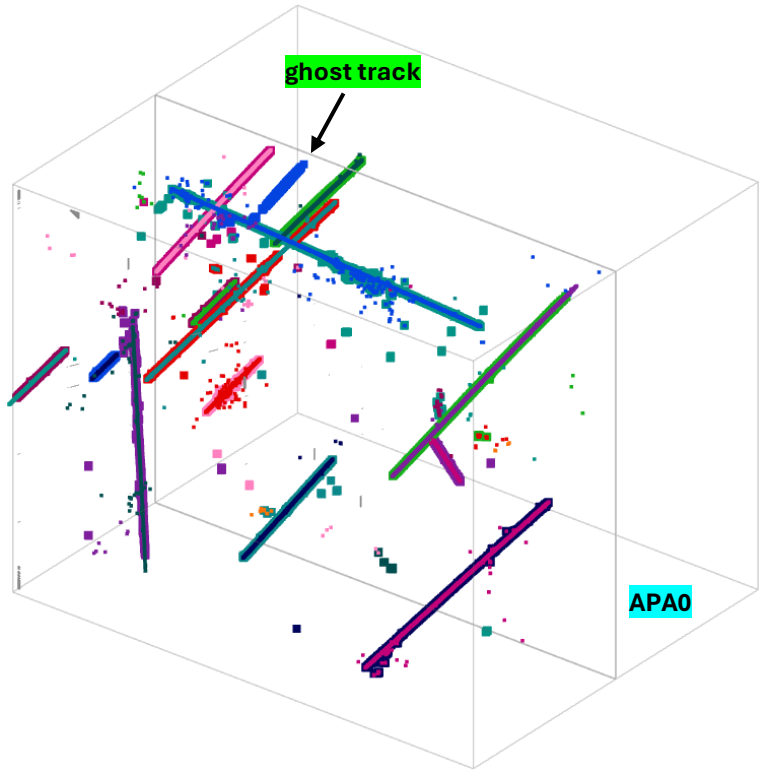


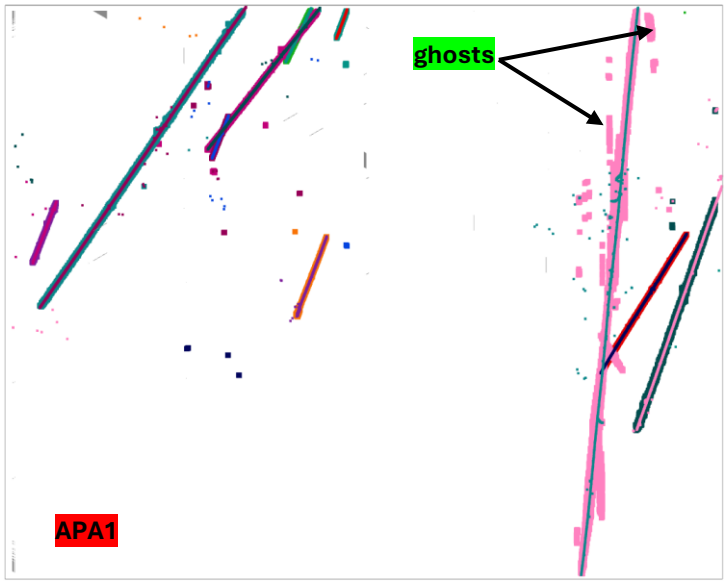


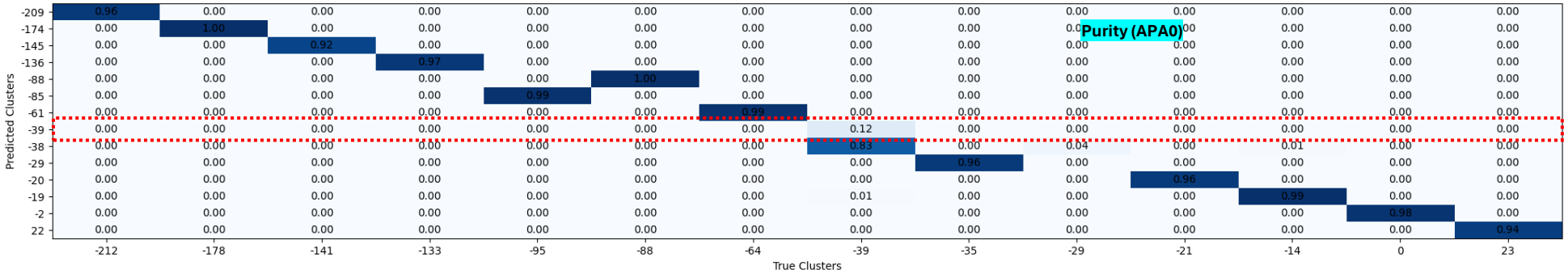
Handscan (MC)

run 1, event 18

https://www.phy.bnl.gov/twister/bee/set/1f269c02-e266-414a-8c3f-831e785306a4/event/17/



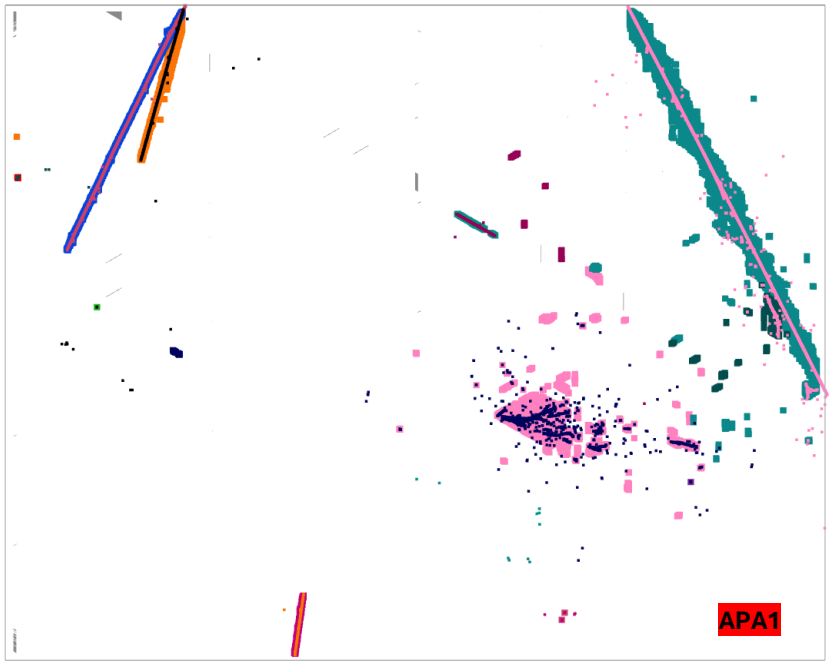


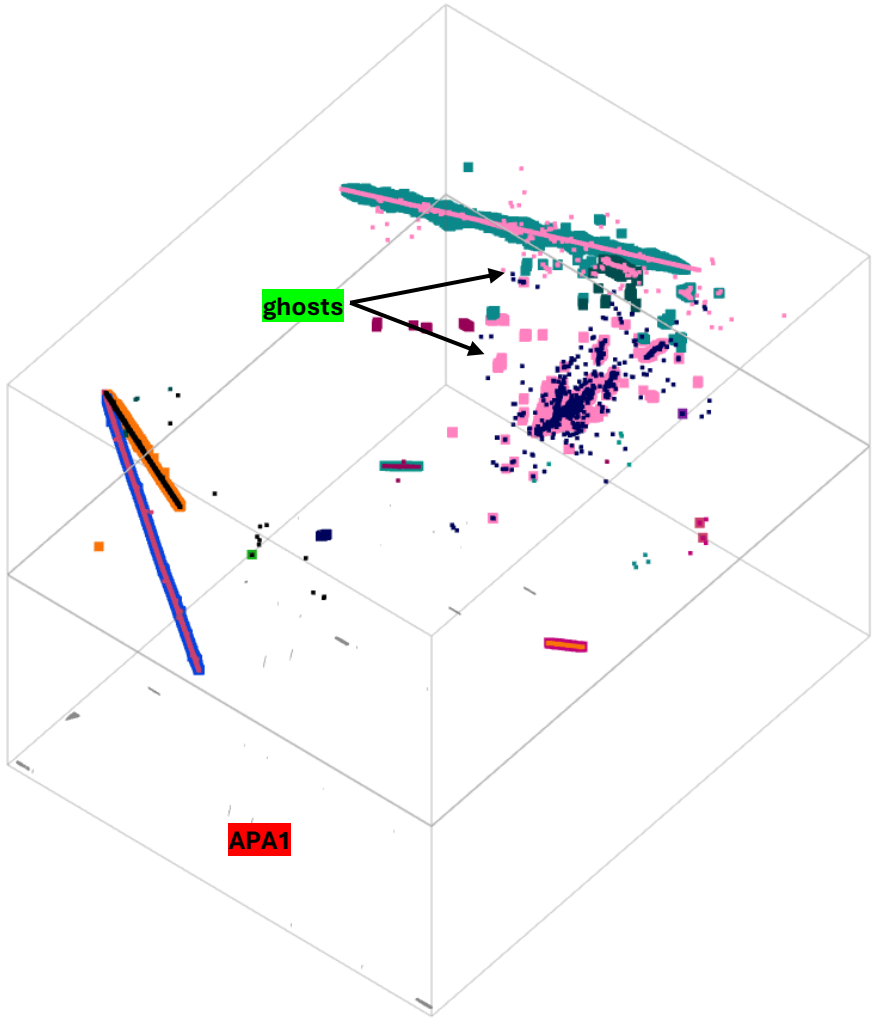


Handscan (MC)

run 1, event 20

https://www.phy.bnl.gov/twister/bee/set/1f269c02-e266-414a-8c3f-831e785306a4/event/19/

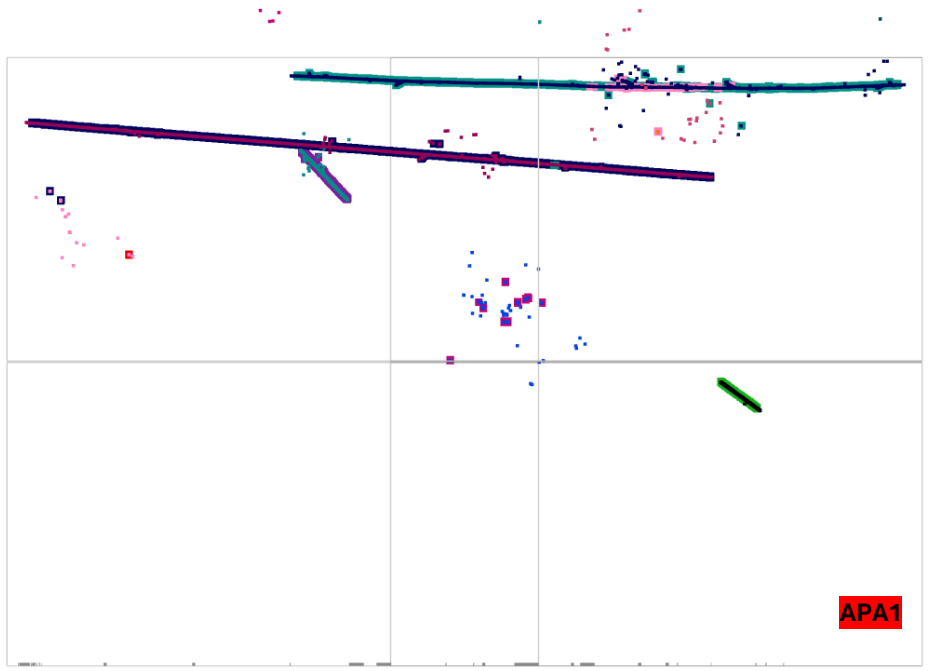


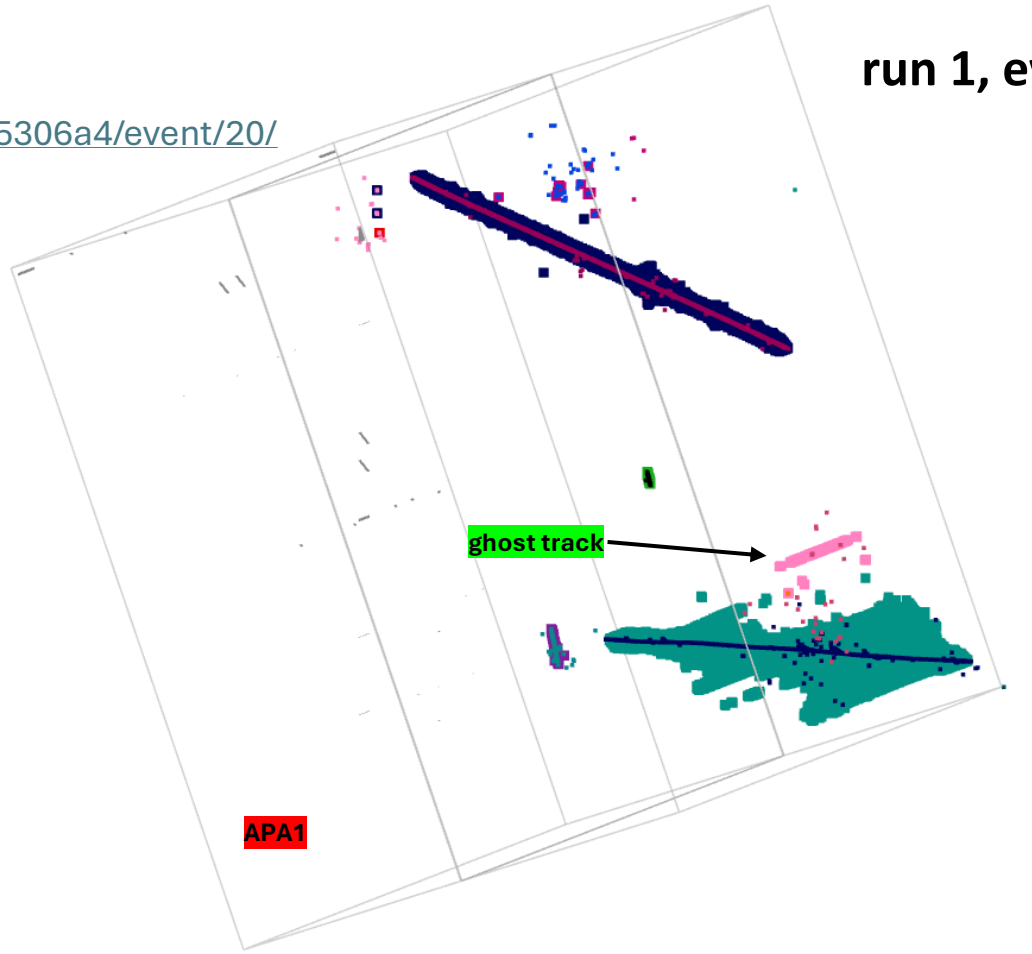


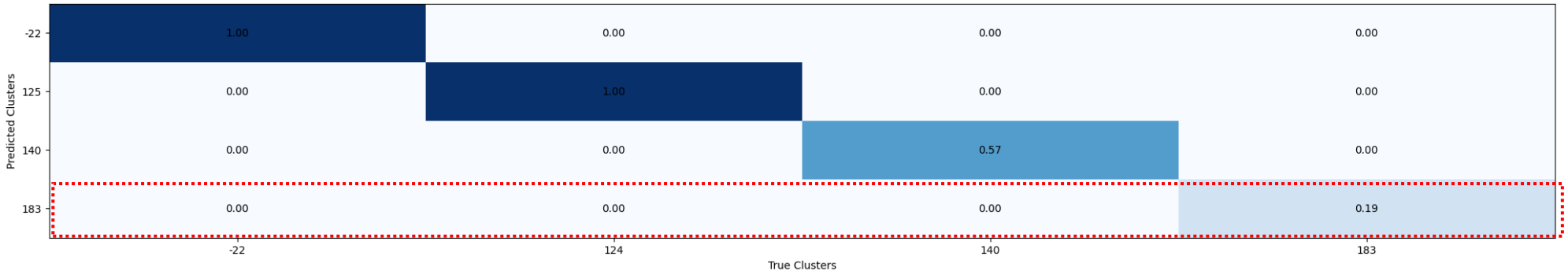
Handscan (MC)

run 1, event 21

https://www.phy.bnl.gov/twister/bee/set/1f269c02-e266-414a-8c3f-831e785306a4/event/20/



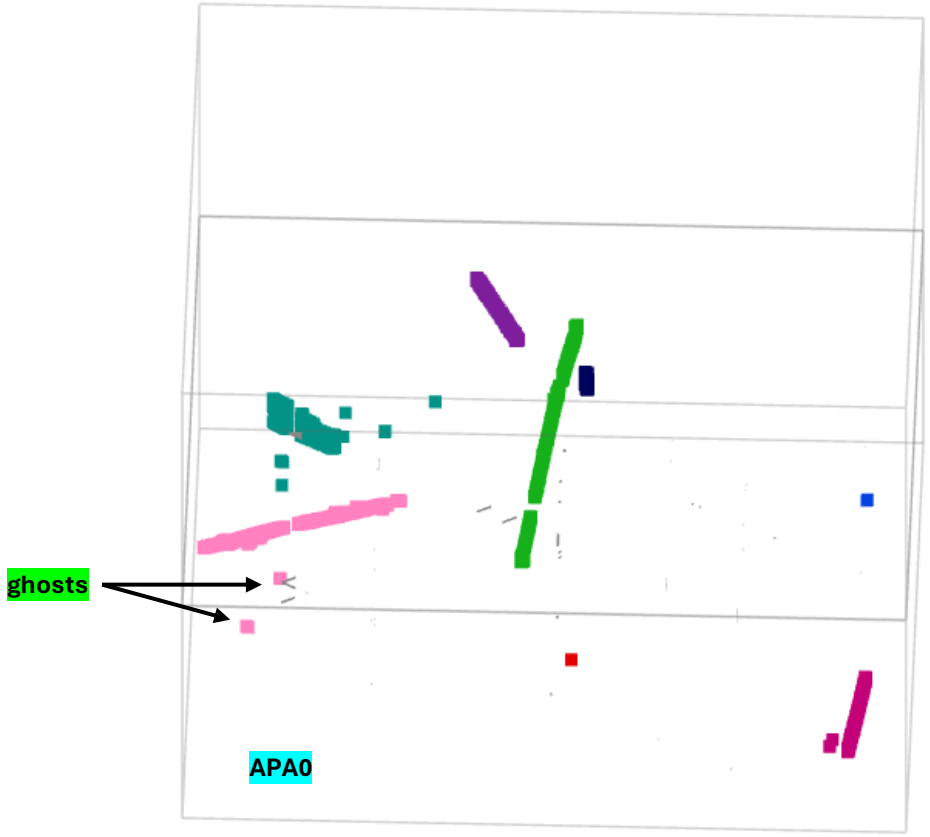




Handscan (data: run 14766)

run 14766, event 3360

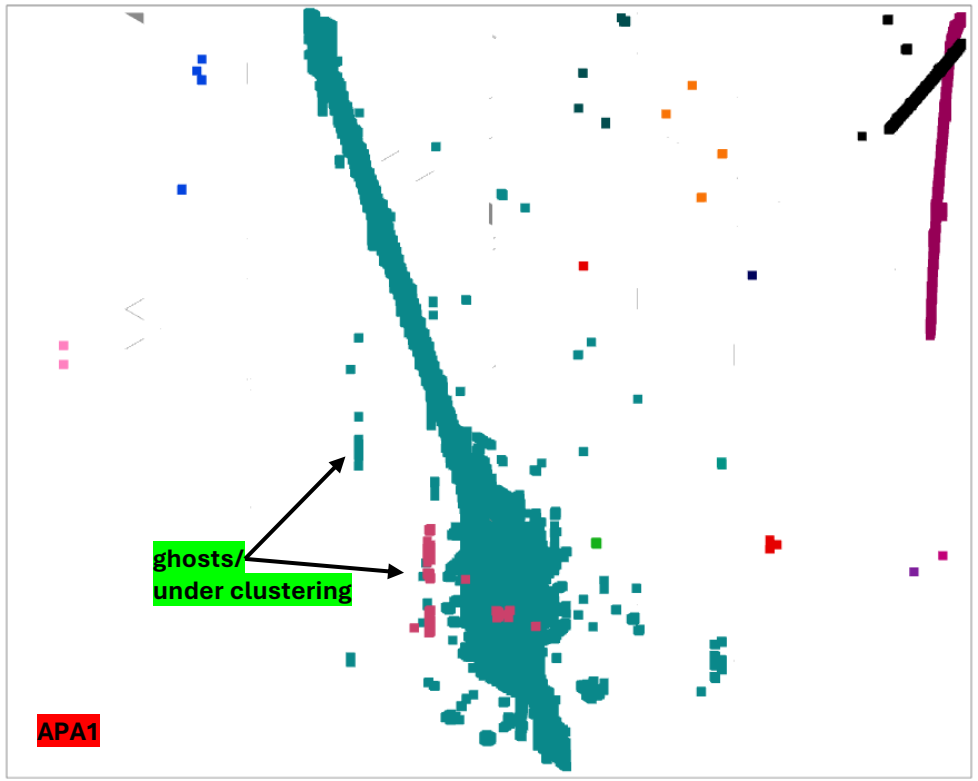
https://www.phy.bnl.gov/twister/bee/set/5caef909-f5b9-4e57-be0a-92076af514c2/event/0/



Handscan (data: run 14766)

run 14766, event 3384

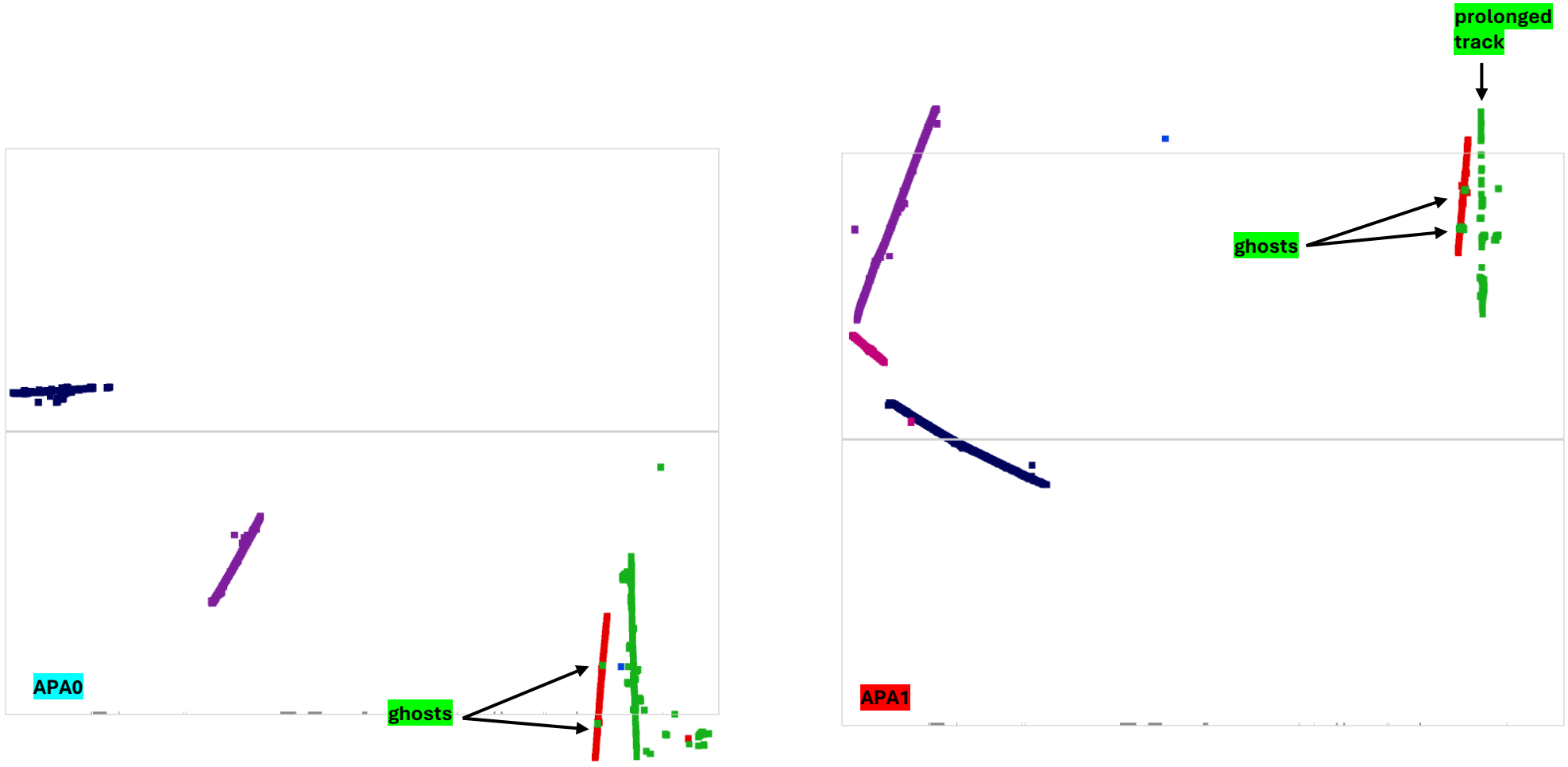
https://www.phy.bnl.gov/twister/bee/set/5caef909-f5b9-4e57-be0a-92076af514c2/event/2/



Handscan (data: run 14766)

run 14766, event 3408

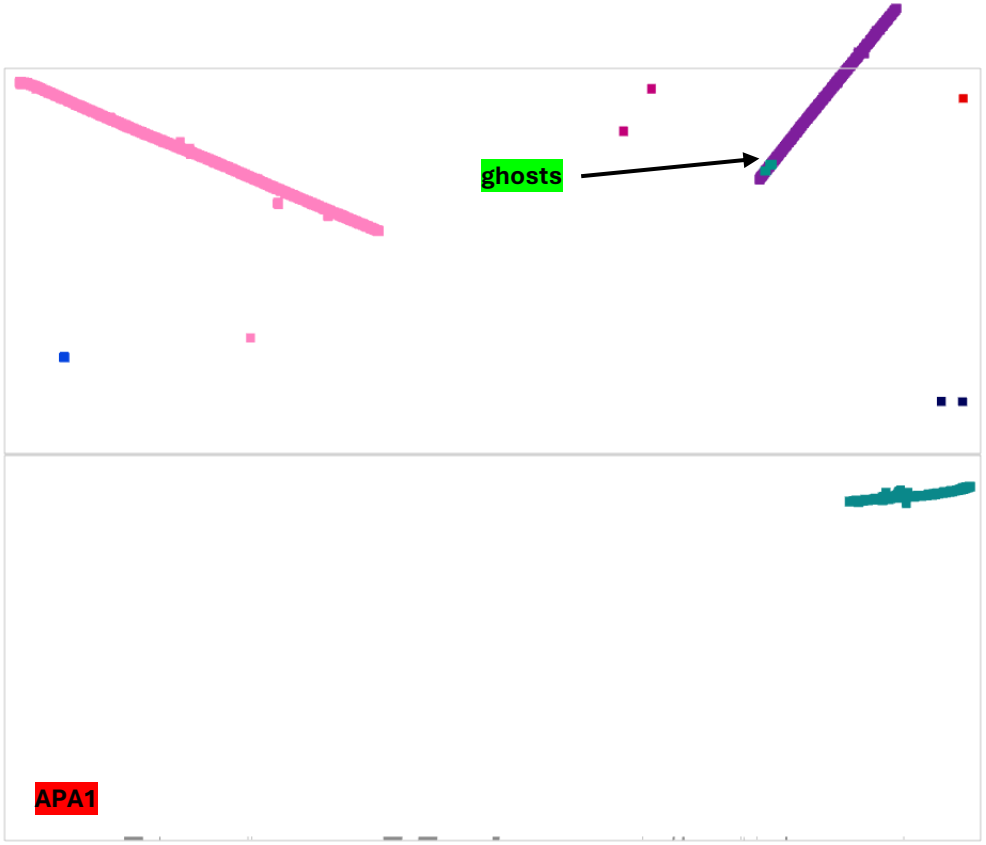
https://www.phy.bnl.gov/twister/bee/set/5caef909-f5b9-4e57-be0a-92076af514c2/event/4/



Handscan (data: run 14766)

run 14766, event 3420

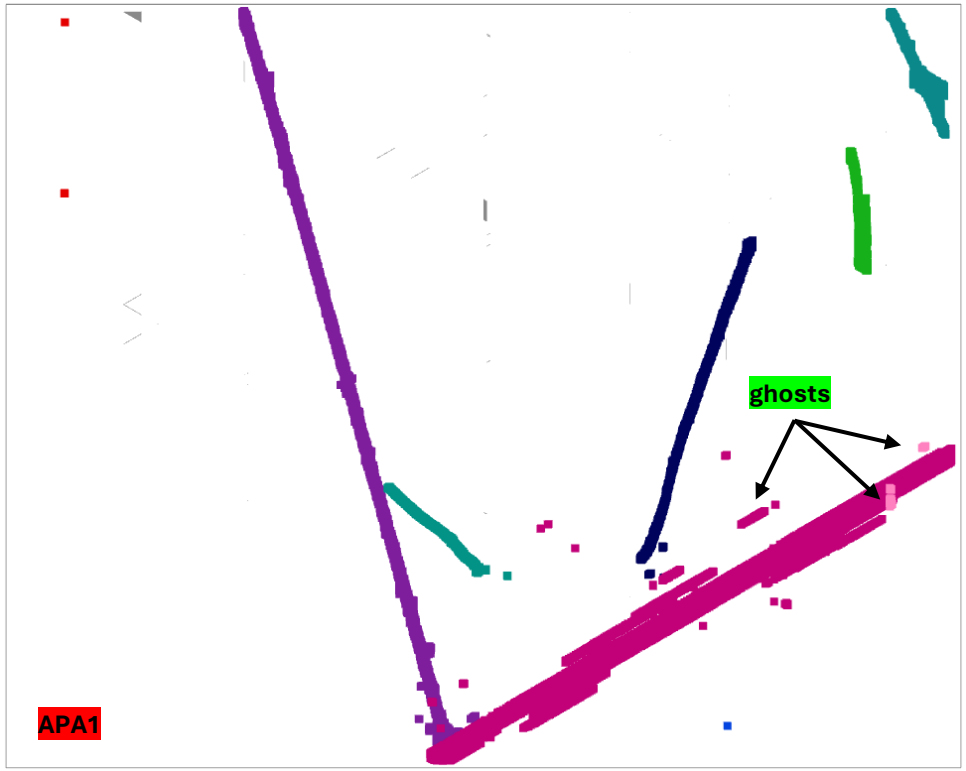
https://www.phy.bnl.gov/twister/bee/set/5caef909-f5b9-4e57-be0a-92076af514c2/event/5/



Handscan (data: run 18005)

run 18005, event 35969

https://www.phy.bnl.gov/twister/bee/set/67ed0429-e8c9-4979-9b57-7128acb01e74/event/1/



Handscan (data: run 18005)

run 18005, event 36409

https://www.phy.bnl.gov/twister/bee/set/67ed0429-e8c9-4979-9b57-7128acb01e74/event/8/