LOWE FLASH RECONSTRUCTION & FLASH-MATCHING

HD Geometry (dune10kt_1x2x6)

Sergio Manthey Corchado

Motivation

Flash Reconstruction for LowE Neutrino Events:

- **TPC - PDS matching is required** to reconstruct the drift-coordinate of low energy interactions allowing **improved calorimetry** and **3D fiducialization.**
- **Current standard implementation of OpFlash fails** to produce pure objects for low energy events due to the presence of background.
- **Impure OpFlashes are badly reconstructed** and lead to an inefficient matching that fails frequently (> 80%) due to the presence of competitive bkg. OpFlashes (up to 4000 per event).
- Up to now there is **no flash-matching strategy implement in larsoft**. Only FlashMatchAna_module exists but does not attempt a proper TPC - PDS match, it only associates reconstructed signal OpFlashes to their truth particle interaction vertex.

This work aims to implement a flash reconstruction and a flash-matching algorithm for low energy events allowing a **full 3D reconstruction** for the region of interest in **solar neutrino studies** (4 - 30 MeV).

This Work

- Start **evaluating HD geometry** (dune10kt_1x2x6). Results for VD will follow!
- Study the **distributions of OpHits** (number, adjacency, PE & MaxPE) for **signal events**.
- OpHits on their own don't carry information to distinguish signal from bkg. Need higher order object: **OpFlash → collection of OpHits** with commonly reconstructed variable (time, charge, vertex...).
    - **LowE OpFlashes** should be sharper (in time) and **smaller (in footprint) than standard** (high energy). Flip-Flop **algorithm** clusters OpHits by time and distance to create SolarOpFlash.
- Study the **distribution of signal OpFlashes** (number & weighted distributions).
- With defined parameters evaluate generated set of MainOpFlashes (OpFlashes with highest number of PEs per event) in the context of full background.
- For a given TPC cluster, all OpFlashes with *ΔT < MaxDriftTime are candidates for association (matching). Find additional cuts to minimize bkg. candidates.

*ΔT = TPC_Time - PDS_Time

	Hit	OpFlash	Matching	Efficiency
Before this work.	Standard OpHit*.	Standard larsoft OpFlash based on DBSCAN.	No TPC - PDS matching strategy exists for lowe.	< 20%
With this work.	Standard OpHit*.	Specific SolarOpFlash for low energy point-like interactions based on (flip-flop).	Strategy based on tight vertex reconstruction and maxPE criteria.	~ 80%
After this work.	Standard OpHit*.	Updated SolarOpFlash with **centroid vertex** computation for VD.	Updated matching criteria for a better likelihood-based algo.	???

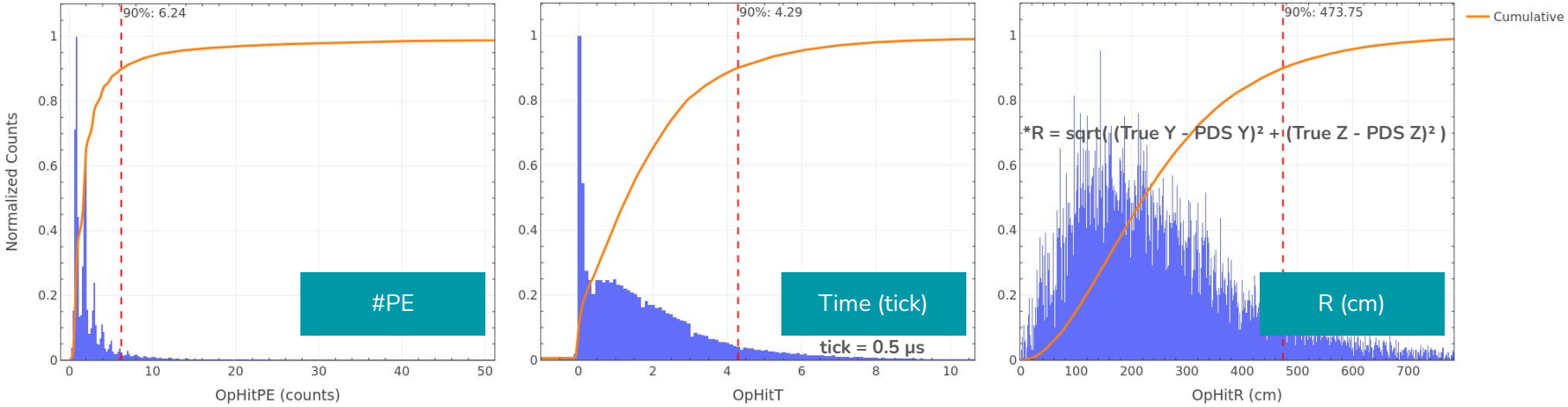
*Already tuned for low energy during deconvolution studies.

Neutrino Production

Isotropic flat CC neutrino spectrum 4-30 MeV

OpHit Distributions

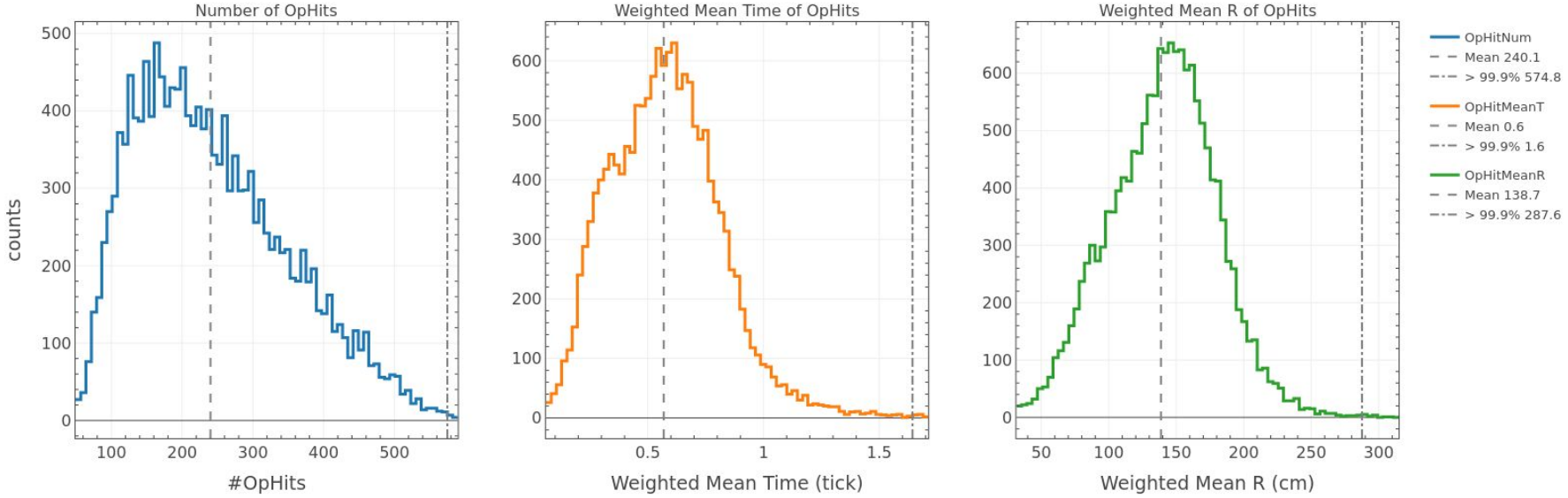
- OpHits: Reco signals of the PDS raw wvfs (> 1.5 PE). Characterised by their charge in units of Photo Electron (PE), the channel (X-Arapuca) location in 2D space (*Y, Z coordinates) and their Time (tick).
- OpHits from low energy neutrino signals arrive to the PDS plane with a **lowering #PE** count, a very **sharp scintillation peak** (< 0.25 μs) and a wide **surface coverage**.
- Better to **evaluate the weighted response** to tune clustering parameters (see next slide).



*In VD XAs surround the active volume (in HD all XAs on the same plane) → their 3D location needs to be considered .

OpHit Sample: #Number and Weighted Distributions

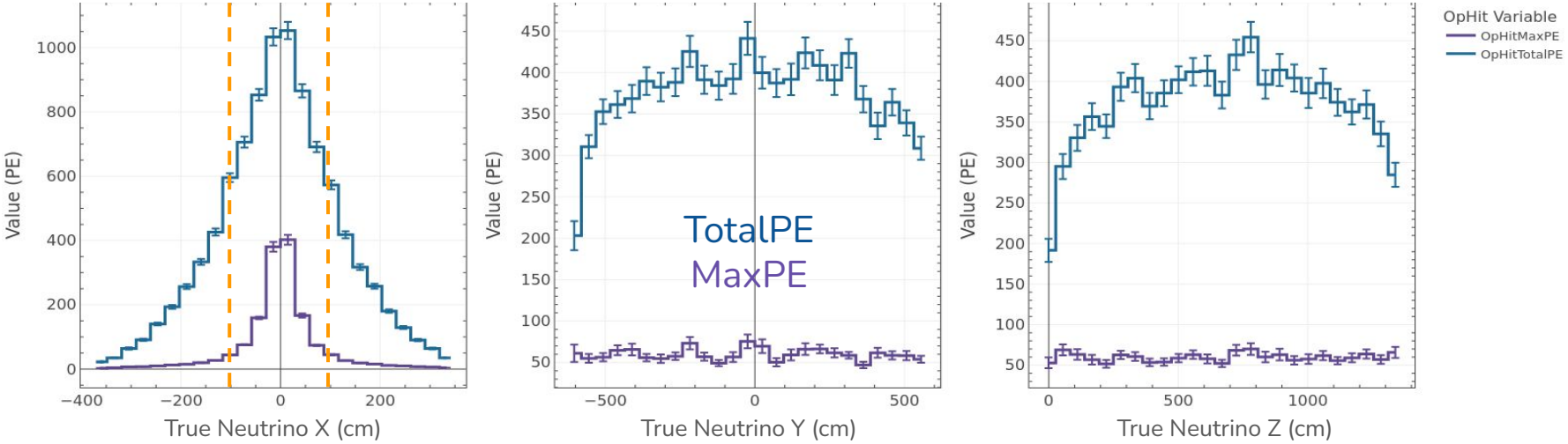
- **Found ~240 OpHits** per neutrino event for which the PE weighted Time peaks at 0.6 tick with an asymmetric distribution (**[peak - 0.6, peak + 1] tick**) and the R distribution peaks at 150 cm with OpHits reaching up **300 cm** in projected True Neutrino - PDS vertex distance.



*R = sqrt((True Y - PDS Y)² + (True Z - PDS Z)²)

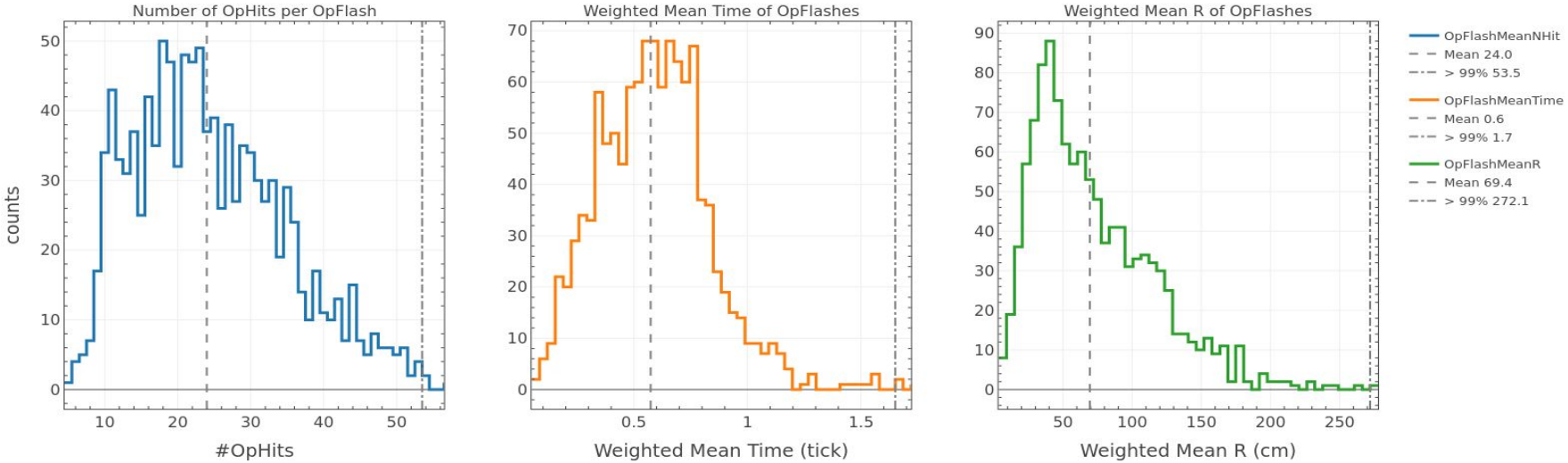
OpHit Sample: Max and TotalPE Coordinate Scan

- Evaluate total number PEs per event (**TotalPE**) & OpHit with the maximum PE count (**MaxPE**).
- **MaxPE** distribution **decreases stronger** than **TotalPE**, outside 1m in drift coordinate, value falls below 10% while **TotalPE** only reduced by 50% → Probable **reason why density based standard algorithm** (using DBScan) **fails** to produce pure samples.



OpFlash Sample: #OpHits and Weighted Distributions

- **Weighting the OpFlash** sample according to their #PE → better estimate of their properties.
- PE weighted Time peaks at 0.6 tick (same as OpHits)
- R distributions peak at 40 cm with shorter tail wrt. OpHits. → **expected improvement** due to the combined information (weighted average of Y & Z coordinates) of all OpHits in the OpFlash.



*R = sqrt((True Y - PDS Y)² + (True Z - PDS Z)²)

Result: Clustering Parameters from Signal Production

- **Flip-Flop algorithm clusters OpHits around a trigger OpHit (with #PE > Adj. OpHit #PE) up to a tunable radius *R and for a certain tick interval [t_i, t_f].**
- **OpHit clustering parameters (see slide):**
  - Cluster OpHits around trigger OpHit with relative time **[-0.6, 1] tick**.
  - Cluster OpHits around trigger OpHit up to **300 cm** in radial distance *R.
- **OpFlash signal thresholds (see silde)**:
  - Final cuts be tuned in the context of background, for now signal threshold:
  - OpFlashes should be made of at least 3 OpHist.
  - OpFlashes should be found within ~300 cm of projected distance **R.

*R = sqrt((Trigger Y - Adj. Y)² + (Trigger Z - Adj. Z)²)

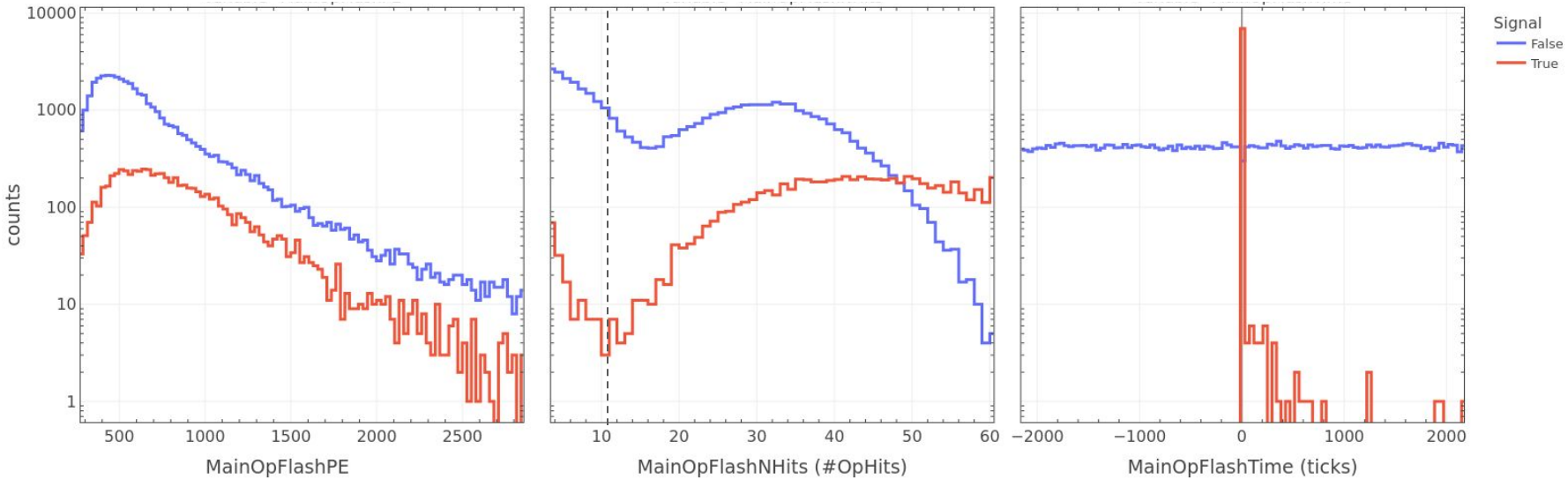
**R = sqrt((True Y - PDS Y)² + (True Z - PDS Z)²)

Wbkg Production

Isotropic flat CC neutrino spectrum 4 - 30 MeV + full standard centralAPA bkg.

MainOpFlash Sample

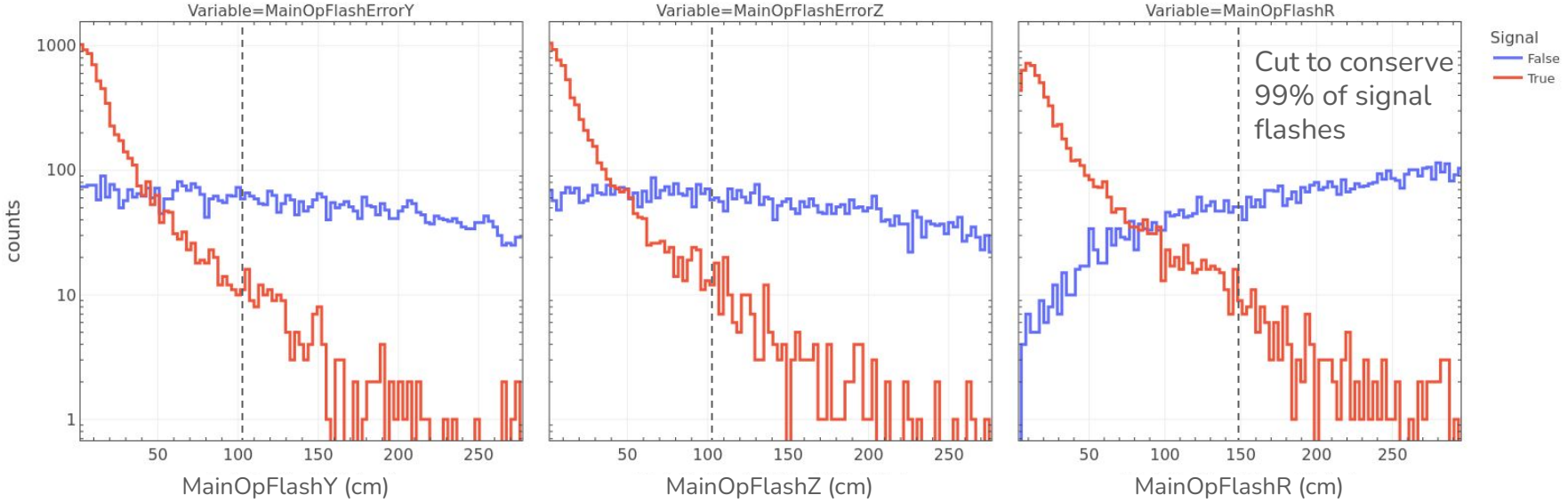
- Evaluate **MainOpFlash** (highest PE per event) by separating out the ones with *purity > 0 as **signal**.
- Found **biggest flash in the event coming from signal in only 15% of the events** (rest is **bkg.**).
- Most **signal** **OpFlash have > 10 OpHits**.



*purity evaluated in SolarNuAna. Defined as fraction of OpHits per OpFlash backtracked to signal.

MainOpFlash Sample

- Need to cut on OpFlash variable to reduce choice for any given TPC cluster → easy cut on *R.
- **ΔY & ΔZ < 100 cm** conserves 99% of signal flashes.



*R = sqrt((True Y - PDS Y)² + (True Z - PDS Z)²)

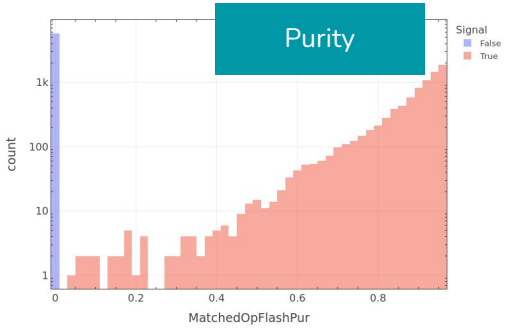
Result: Cuts from Signal+Background Production

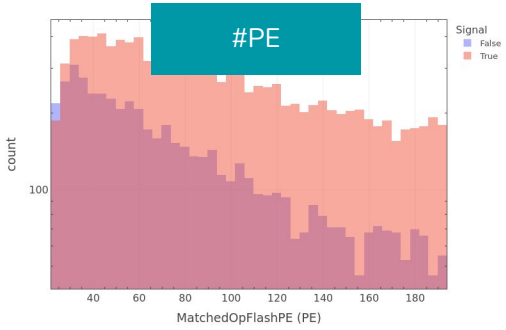
- **OpHit clustering parameters (see slide):**
  - Not updated in the context of background.
- **OpFlash cuts (see previous slides)**:
  - OpFlashes should be found within **R ~141 cm** of projected distance ($\Delta Y > 100$, $\Delta Z > 100$ cm).
  - OpFlashes should be made of at least **10 OpHits** (used as matching criteria).

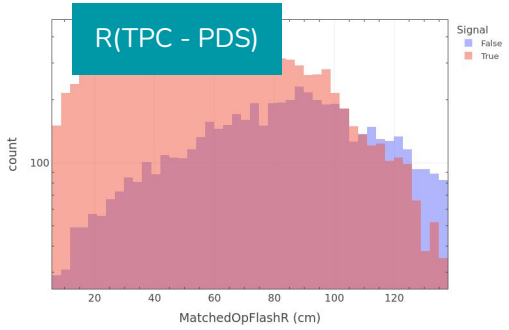
- **In the following** we attempt **proper matching: evaluating AdjOpFlashes (~40 per event) within the vicinity of a signal TPC cluster *R,** from which we select only one **MatchedOpFlash** to reconstruct the drift coordinate (X) for a given TPC cluster.
- Current strategy selects AdjOpFlash with highest #PE after adjacency and #OpHit cuts.

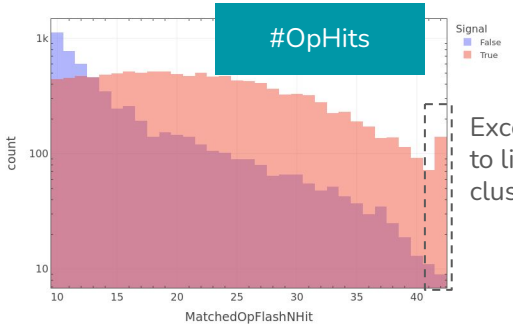
*R = sqrt((**TPC Y** - PDS Y)² + (**TPC Z** - PDS Z)²)

MatchedOpFlash Sample: Background vs Signal

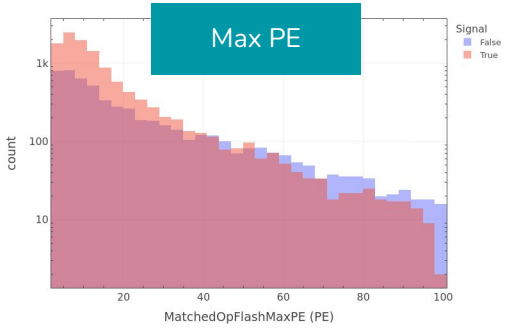


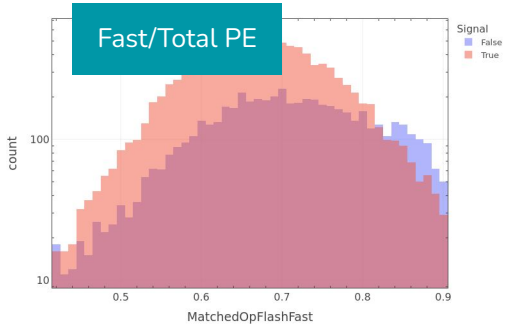






Excess due to limit in clustering R

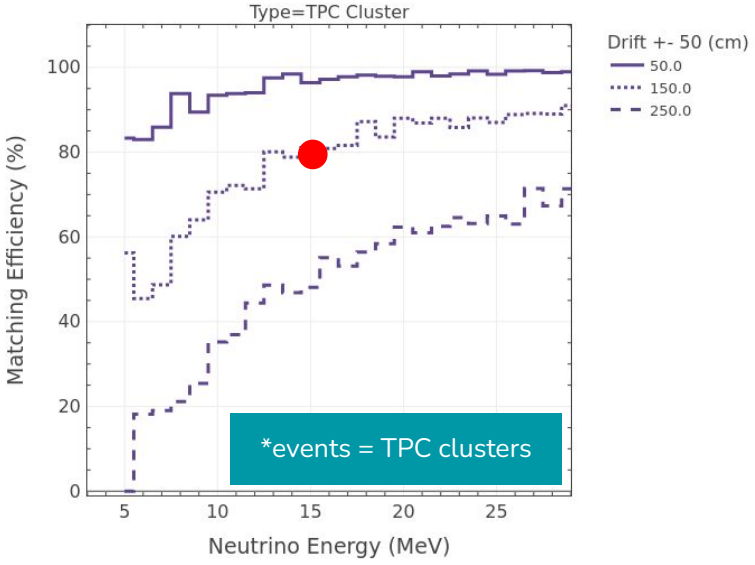




- After choosing from AdjOpFlash the **MatchedOpFlash → signal dominates!**

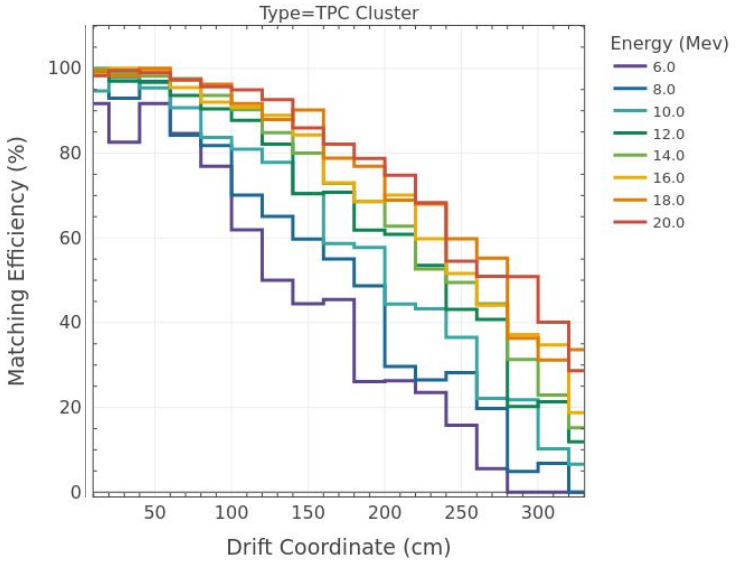
Results: Matching Efficiency (centralAPA bkg config)

- Efficiency **#correctly matched events / #events*** (correct match = match to OpFlash with purity > 0).
- **For reference**, solar event (15 MeV, +3Hits) interacting @ drift coordinate 150 cm has a **80% chance of correct match.**



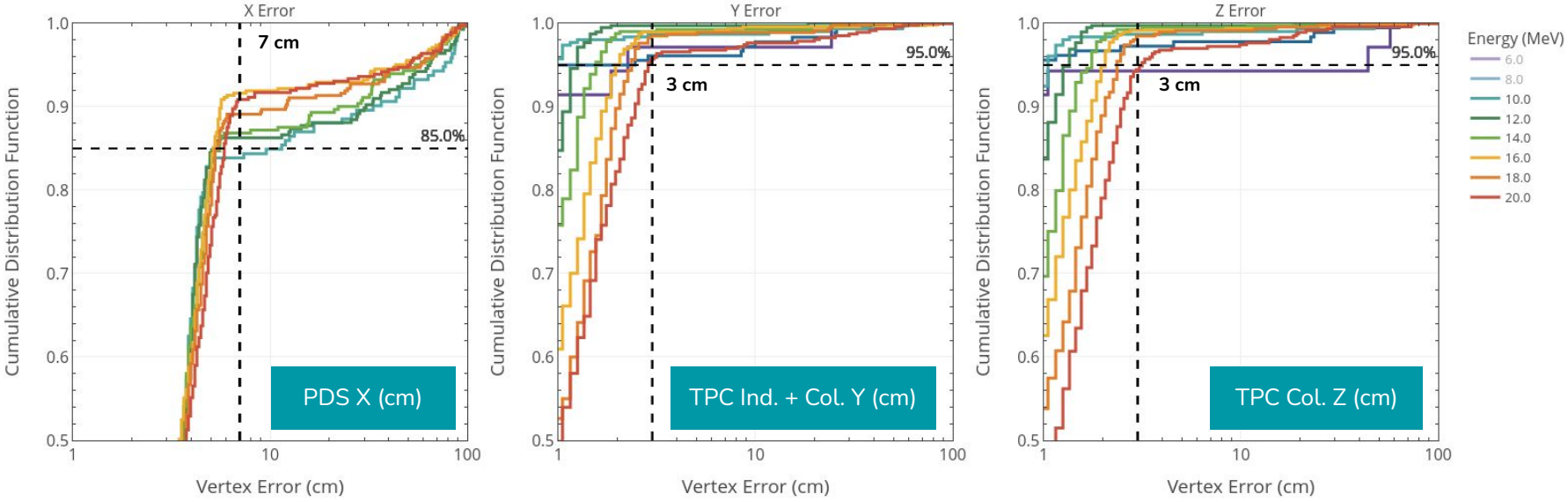
Results: Matching Efficiency (centralAPA bkg config)

- Very good overall performance compared to previous results (< 20%)!
- Efficiency for low energy region needs to be defined as a function of **neutrino energy** and **drift coordinate**.



Results: 3D Reconstruction

- Evaluate **cumulative error distributions** defined as **abs(True - Reco)**.
- Scope of full 3D reconstruction: for subsample of 10 - 20 MeV **85% of matched TPC** clusters have a drift coordinate reconstruction that agrees **within < 7 cm** wrt. **true neutrino interaction**.



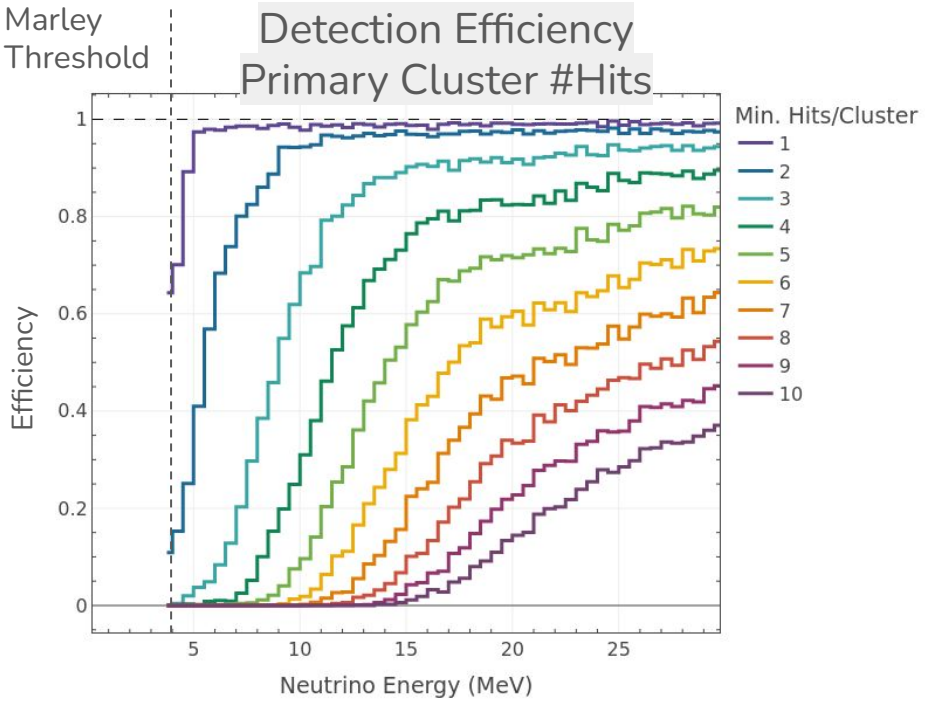
Summary & Next Steps

- OpHit clustering, OpFlash reconstructing and PDS-matching algorithms has been implemented in larsoft for low energy events in the context of background.
- **OpHit clustering parameters have been updated** to improve signal efficiency wrt previous results.
  - Asymmetric time limits wrt. trigger hit have been implemented in Flip-Flop algorithm.
  - OpHit R-adjacency parameter could be fine-tuned to reduce #PE in bkg flashes.
- Presented **OpFlash cuts** successfully **remove most bkg flashes**.
- Current settings **successfully match > 95 (80, 50) % of 15 MeV neutrinos** in the first (second, third) drift meter.
- Previously presented **Residual discriminator is not favouring** the signal OpFlash topology.
  - Need to improve residual discriminator (currently using #**highest PE approach)**.
- Additional **fine-tuning** could further improve these results.
- Apply reconstruction to solar neutrino sensitivity studies. Improve algorithms for **VD geometry.**

Backup

Low Energy Clustering

- Reconstructed **wirecell hits**: col. plane + ind. plane matching (at least 1).
- Matched col hits grouped into clusters according to time & wire proximity (25 tick & 3 channels) **ensuring cluster purity**.
- Clusters divided into **“Primary”**and **“Adjacent”** (R adj cl. < 1m & charge adj. < charge primary).



- **Detection Efficiency**: Probability of finding **at least 1 primary cluster per event**.
- Important to look for low hit multiplicity in the context of low energy interactions.

Result: Configuration (SolarOpFlash.fcl)

- Presented results lead to the following configuration for the flip-flop algorithm to generate the SolarOpFlashes.

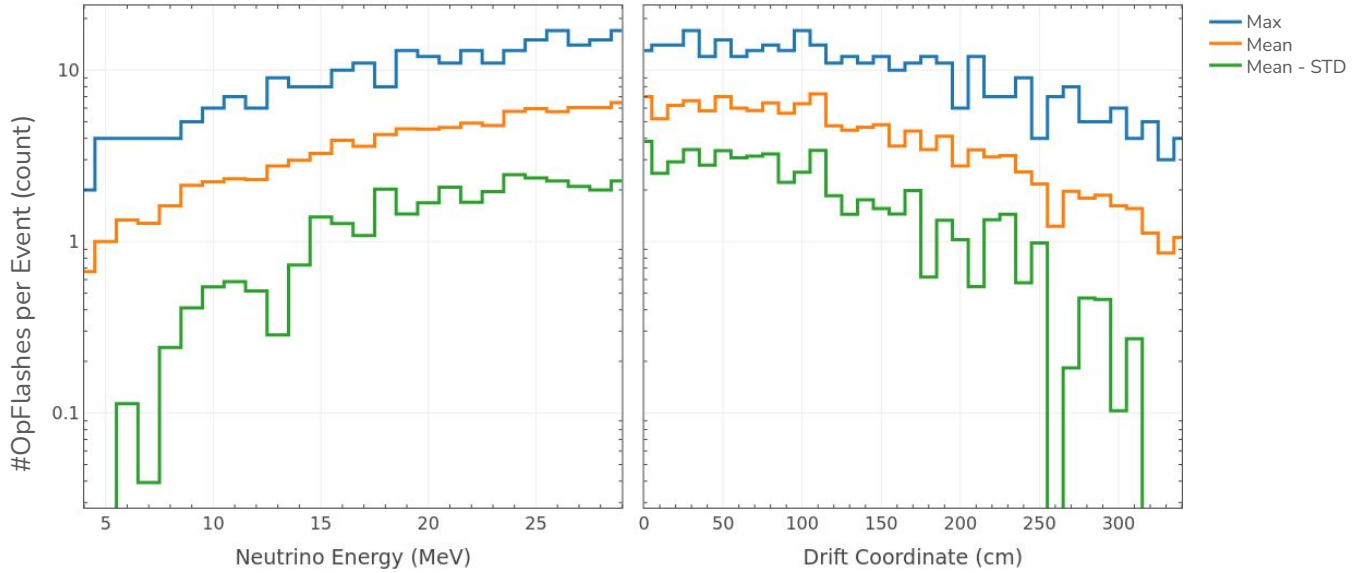
1. Avoid unnecessary large output.
2. 3. Adjust to scintillation time profile.
4. Minimize background contamination.
5. 6. Reduce flash-noise.
7. 8. Improve vertex reconstruction.

```
BEGIN_PROLOG

solar_opflash_dune10kt_1x2x6:
{
    OpHitLabel:         "ophitspe"     # The label for the process which ran the OpHit
    module_type:        "SolarOpFlash"

    Geometry:                   "HD"
    DetectorSizeX:              360          # 1 TPC drift (X) length in [cm].
    DetectorSizeY:              600          # 1 TPC Y length in [cm].
    DetectorSizeZ:              1400         # 1 TPC Z length in [cm].
    DetectorDriftTime:          4500         # Don't change this value for HD unless you know what you are doing!!!
1   OpFlashAlgoNHit:            3            # Min number of hits to consider a flash. Change to > 3 for bkg productions.
2   OpFlashAlgoMinTime:         0.60         # Negative time window to look for adj. OpHits in [tick] units.
3   OpFlashAlgoMaxTime:         1.00         # Positive time window to look for adj. OpHits in [tick] units.
4   OpFlashAlgoRad:             300          # Distance to look for adj. OpHits in [cm] units.
5   OpFlashAlgoPE:              1.5          # PE threshold to look for adj. OpHits.
6   OpFlashAlgoTriggerPE:       1.5          # PE threshold to trigger an OpFlash.
7   OpFlashAlgoHotVertexThld:   0.3          # Relative threshold to consider a hit as hot for opflash vertex determination [0-1]
8   # OpFlashAlgoCentroid:        false        # Use the centroid computation of the flash based on a likelihood estimation.
}

solar_opflash_dunevd10kt_1x8x14_3view_30deg: @local::solar_opflash_dune10kt_1x2x6
solar_opflash_dunevd10kt_1x8x14_3view_30deg.Geometry: "VD"
solar_opflash_dunevd10kt_1x8x14_3view_30deg.DetectorSizeX: 340
solar_opflash_dunevd10kt_1x8x14_3view_30deg.DetectorSizeY: 680
solar_opflash_dunevd10kt_1x8x14_3view_30deg.DetectorSizeZ: 2100
solar_opflash_dunevd10kt_1x8x14_3view_30deg.DetectorDriftTime: 8000

END_PROLOG
```

*Notice 8. OpFlashAlgoCentroid option still needs to be implemented. Probably needed in VD OpFlash reconstruction.

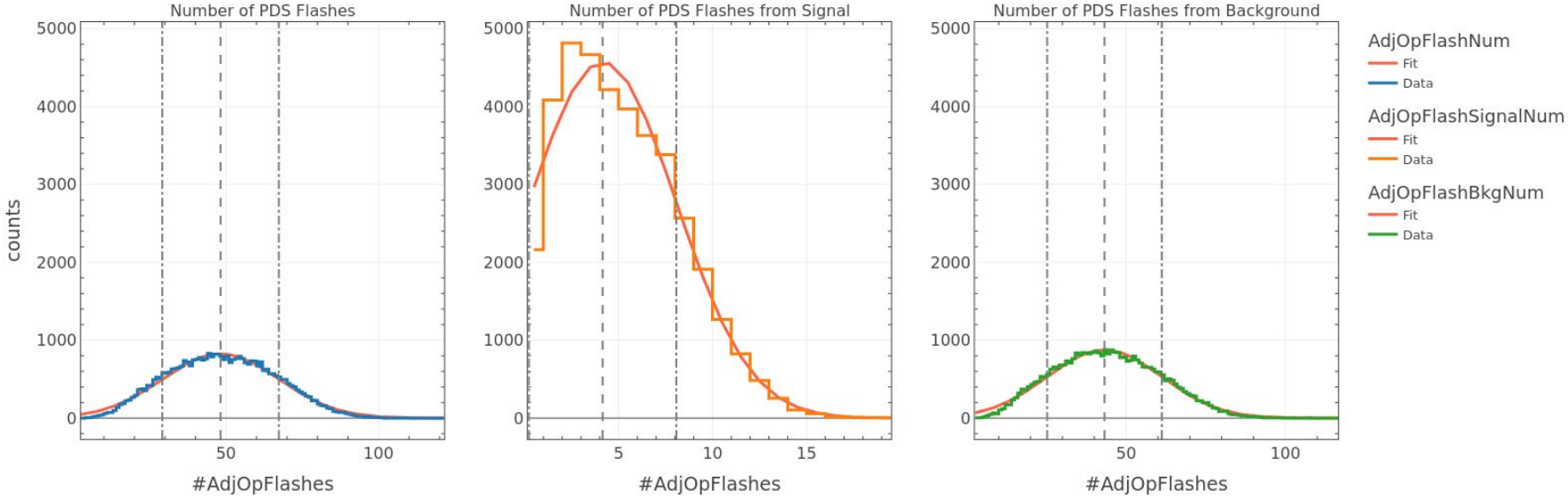
OpFlash Sample: #OpFlashes per Neutrino Energy & Drift

- Construct **OpFlash from collection of OpHits** using **clustering parameters** (summarized in slide with flip-flop algorithm), produced up to ~6 flashes on average, while ensuring **at least 1 OpFlash** for the lowest neutrino **energy (4 MeV)** at the furthest away **drift coordinate (360 cm)**.
- Perform energy and drift coordinate scan of the number of OpFlashes per event (#OpFlashes).

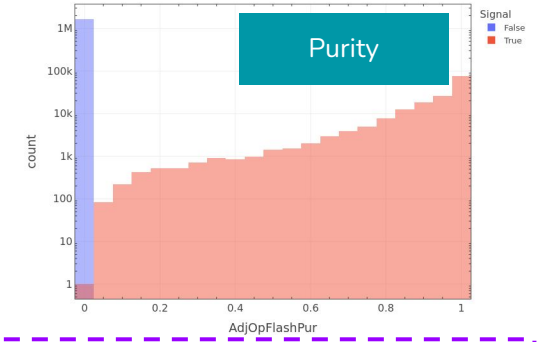


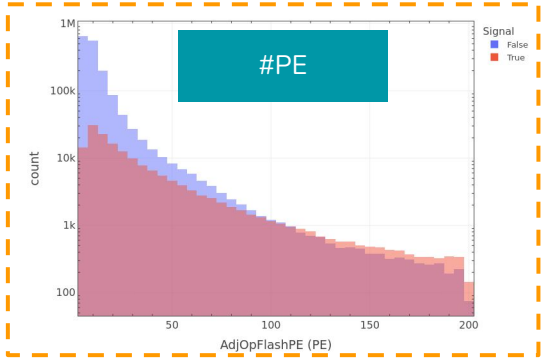
AdjOpFlash Sample: #AdjOpFlashes per TPC Cluster

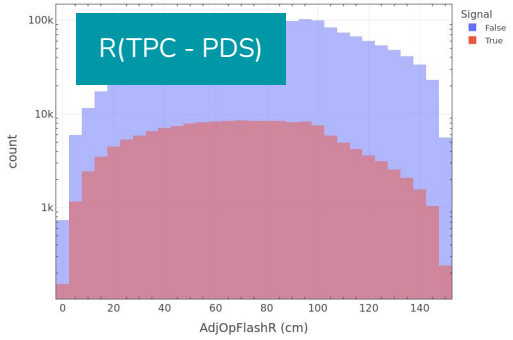
- After proposed clustering parameters and cuts, still **~40 bkg. candidate**s for any given signal TPC cluster.

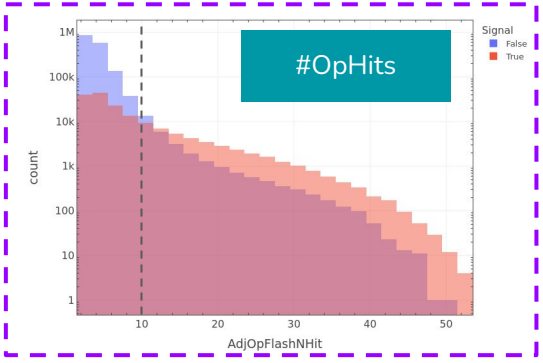


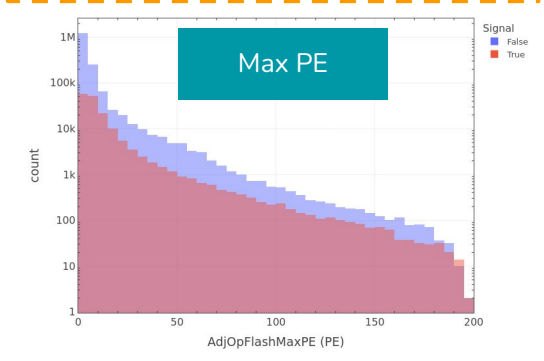
AdjOpFlash Sample: Background vs Signal

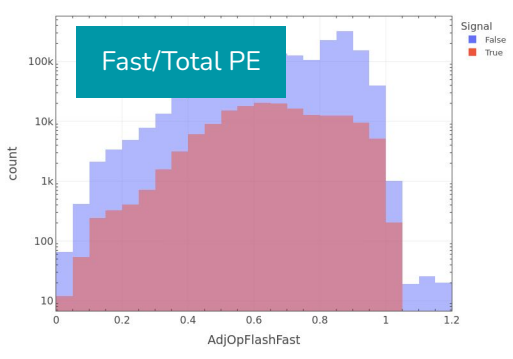












- Still AdjOpFlash sample dominated by bkg. with only **slight differences in #PE and #OpHits**.

Current Analysis Module (SolarNuAna.fcl)

- **Matching analysis** performed within the **SolarNuAna_module** (currently only module) allowing to collect both TPC and PDS info as well as all the relevant backtracked and truth info.
- Module also **allows to generate SolarOpFlash "on the fly"** if object not already existent in art/root.

1. Set to drift-window of the geometry.
2. 3. Adjust to PDS vertex deviation.
4. MaxPE/PE ratio cut for OpFlashes.
5. Min OpFlash PE cut.
6. Min #OpHit per OpFlash cut.
7. **Decide to ditch previously presented matching based on residual in favour of MaxPE discrimination.** Better strategy needs further investigation.

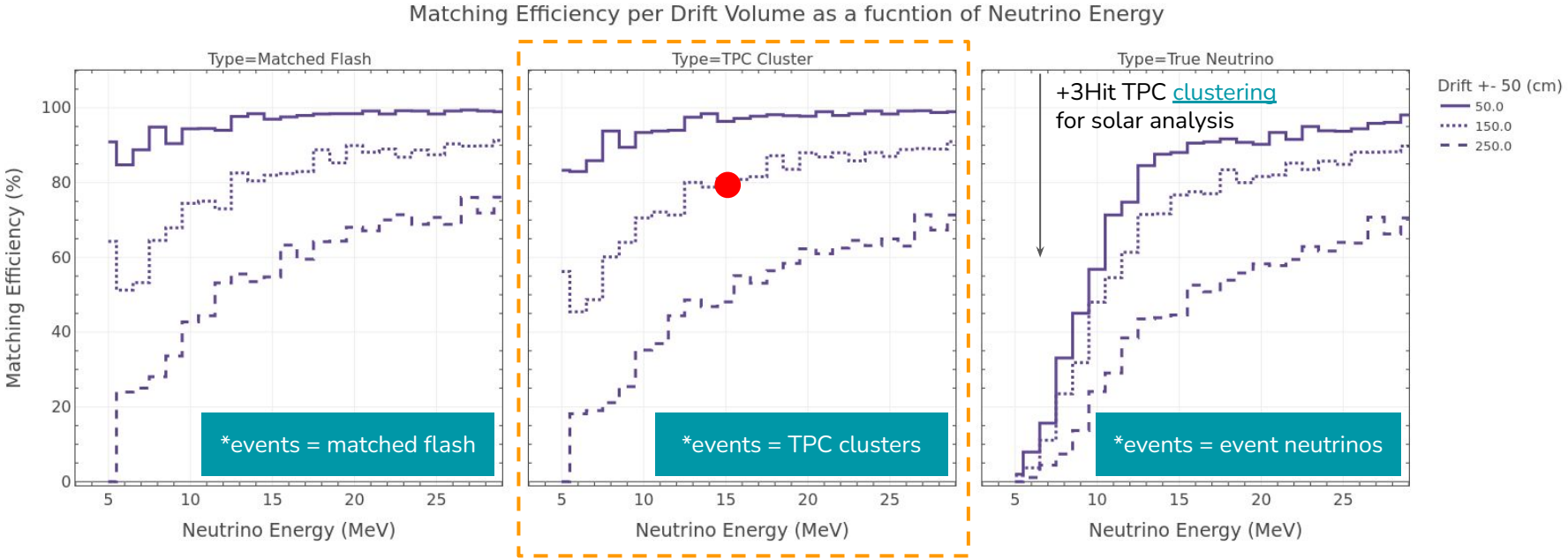
```
GenerateAdjOpFlash:         false       # Generate OpFlashes.
OpFlashAlgoNHit:            0           # Min number of hits to consider a flash. Change to 3 for bkg run to avoid huge output.
OpFlashAlgoMinTime:         0.60        # Negative time window to look for adj. OpHits in [tick] units.
OpFlashAlgoMaxTime:         1.00        # Positive time window to look for adj. OpHits in [tick] units.
OpFlashAlgoRad:             300         # Distance to look for adj. OpHits in [cm] units.
OpFlashAlgoPE:              1.5         # PE threshold to look for adj. OpHits.
OpFlashAlgoTriggerPE:       1.5         # PE threshold to trigger an OpFlash.
OpFlashAlgoHotVertexThld:   0.3         # Relative threshold to consider a hit as hot for opflash vertex determination [0-1].
# Only create SolarOpFlash from within analysis module if GenerateAdjOpFlash is true.
AdjOpFlashTime:             4500        # Time window to look for adjacent flashes in [tick] units should be one full drift length.
AdjOpFlashY:                100.        # Y distance to search for adj. OpFlashes reconstructed in [cm] units.
AdjOpFlashZ:                100.        # Z distance to search for adj. OpFlashes reconstructed in [cm] units.
AdjOpFlashMaxPERatioCut:    1.00        # Cut on the maximum OpHit PE contribution to the total OpFlash PE.
AdjOpFlashMinPECut:         20.0        # Cut on the minimum OpHit PE. From average bkg rate of ~20 PE per tick.
AdjOpFlashMinNHitCut:       10          # Cut on the minimum number of OpHits in the OpFlash.
FlashMatchByResidual:       false       # Match flashes by residual. Alternative is to match by MaxFlashPE.

SaveSignalDaughters:        false       # Save the daughters of the signal particles.
SaveSignalEDep:             false       # Save the energy deposition of the marley energy deposition in the LAr.
SaveSignalOpHits:           false       # Save the OpHits that are associated with the signal.
SaveOpFlashInfo:            false       # Save the AdjOpFlash information for each preselection cluster.
SaveTrackInfo:              false       # Save the MatchedTrack information for each preselection cluster.
```

If *GenerateAdjOpFlash* → Create SolarOpFlash from within analysis module.

*Additional cut in the context of bkg. production: OpFlashPE > 20.

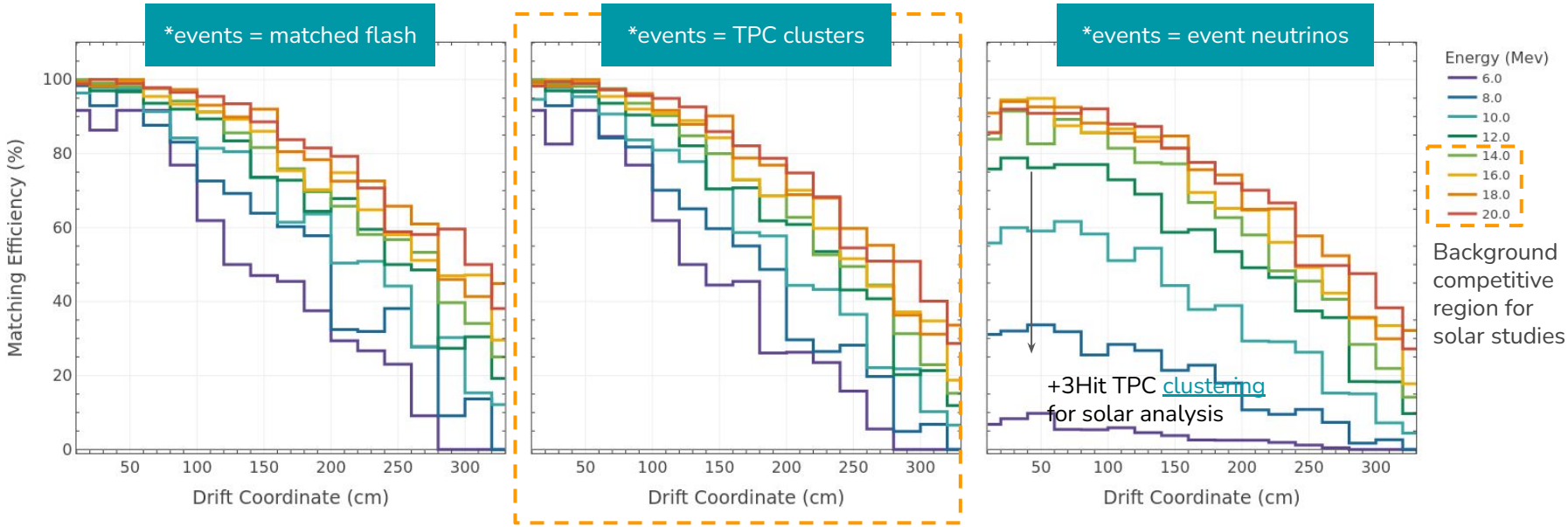
Results: centralAPA Matching Efficiency

- Efficiency **#correctly matched events / #*events** (correct match = match to OpFlash with purity > 0).
- **Relevant efficiency curves** to interpret TPC - PDS signal matching.
- **For reference**, solar event (15 MeV +3Hits) with drift 150 cm has a **80% chance of correct match.**



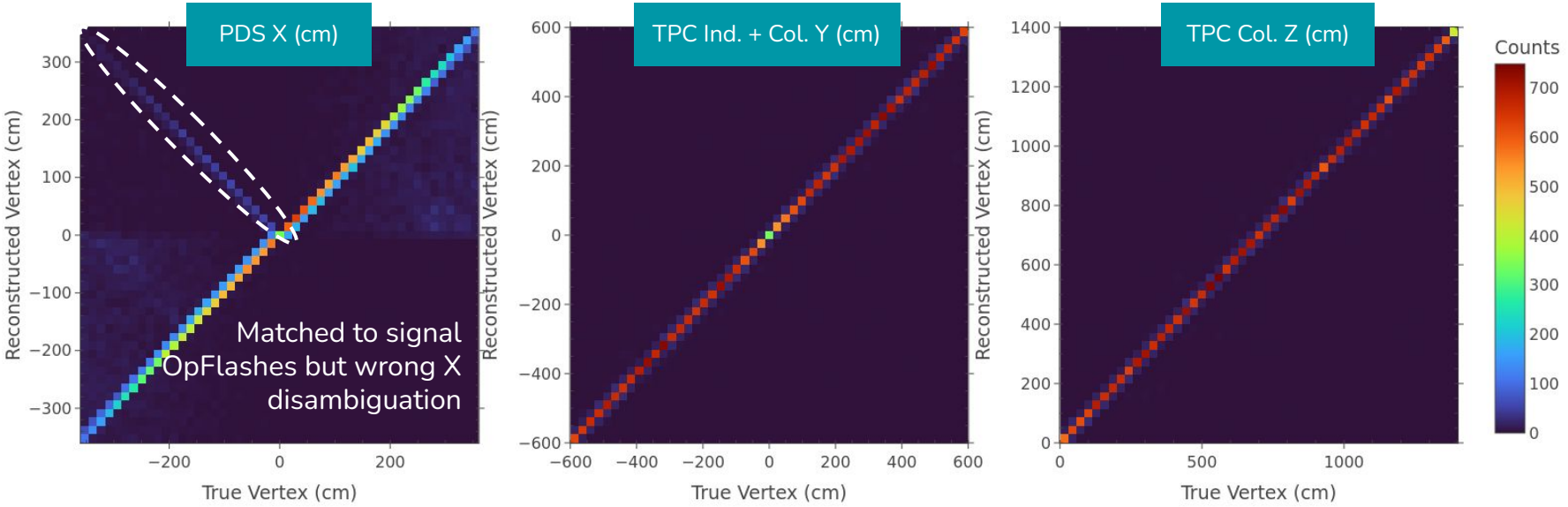
Results: centralAPA Matching Efficiency

- Very good overall performance compared to previous results!
- Efficiency best defined as a function of **neutrino energy** and **drift coordinate**.



Results: centralAPA 3D Smearing

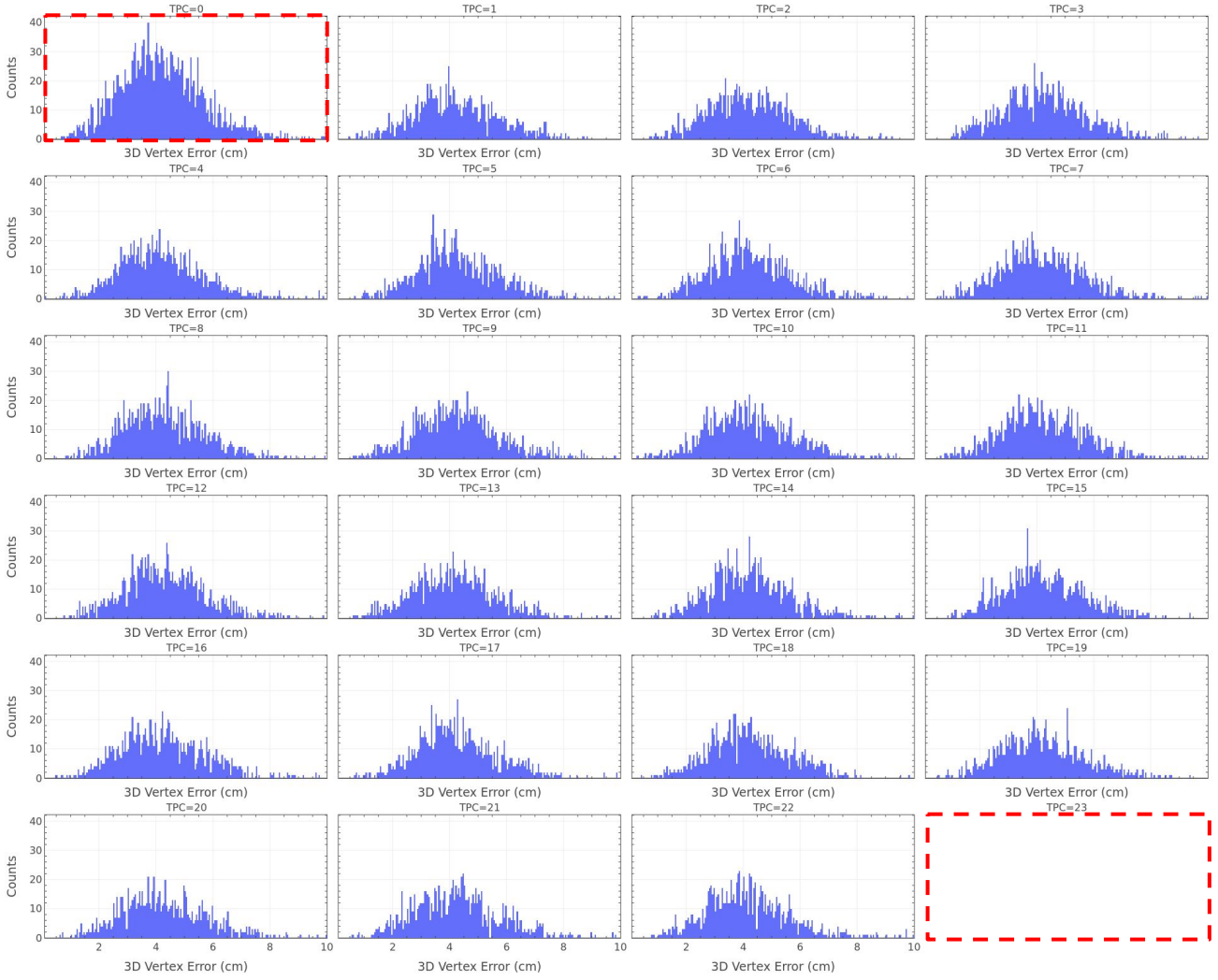
- Scope of full 3D smearing matrices for neutrino events.
- Found slight miss reconstruction of RecoX sign (see backup) due to TPC 23 events being reconstructed as TPC 0. Needs further investigation.



TPC #Events

- Showing vertex error distributions per TPC.
- 24 TPCs should be found [0-23].
- Found missing 23 contribution.
- Double the amount of clusters in TPC 0.

→ Need to investigate stage of failure (production/analysis).



GOBIERNO DE ESPAÑA
MINISTERIO DE CIENCIA E INNOVACIÓN
Ciemat
Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas

Prev. Results

Standard Flash

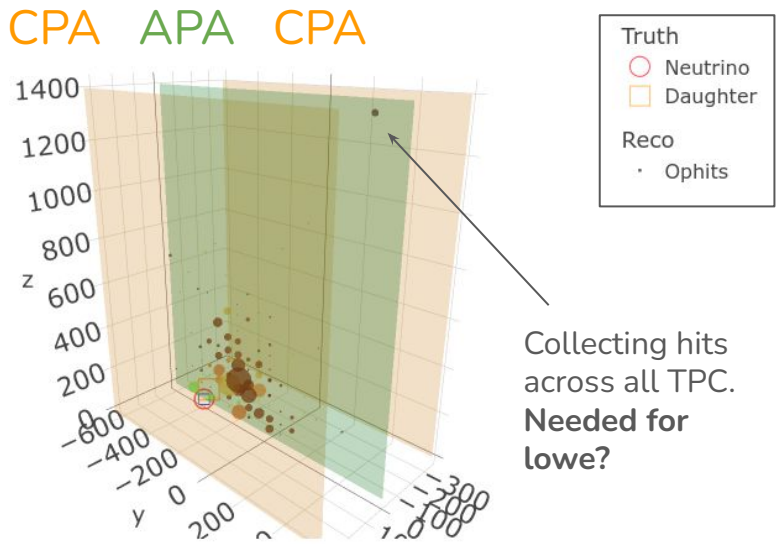
- Standard Flash algorithm build to reconstruct high energy signals using DBSCAN.

double	**Time** () const
double	**TimeWidth** () const
double	**AbsTime** () const
unsigned int	**Frame** () const
double	**PE** (unsigned int i) const
std::vector< double > const &	**PEs** () const Returns a vector with a number of photoelectrons per channel. More...
bool	**hasXCenter** () const Returns whether the estimated center on *x* direction is available. More...
double	**XCenter** () const Returns the estimated center on *x* direction (. More...
double	**XWidth** () const
double	**YCenter** () const
double	**YWidth** () const
double	**ZCenter** () const
double	**ZWidth** () const
bool	**InBeamFrame** () const
int	**OnBeamTime** () const
std::vector< double > const &	**WireCenters** () const
std::vector< double > const &	**WireWidths** () const
double	**TotalPE** () const
double	**FastToTotal** () const

* Missing important variables: Purity, Completeness, FastToTotal (present but not filled).

dune10kt_1x2x6

TNuE: **17.58MeV** Cluster: 10



Collecting hits across all TPC. **Needed for lowe?**

- OpHit size → #PE (1 - 50)
- OpHit color → **Purity (0 - 1)**

Standard Flash Distributions: Background vs Signal

- **Signal** defined according to 0 < FlashTime < 5 tick.
- **Poor signal vs background separation** in all variables.