DUNE
“Data Selection” and FD2-VD
(Trigger and Data Filter)

Josh Klein

CDR Review

June 18, 2021

Penn
University of Pennsylvania

DUNE

Outline

- Definitions, Requirements, and Constraints
- DUNE Trigger Basics
- Performance to date
- Changes for VD

Data Selection

Data rate from each module ~10 Tb/s
Storage <=30 PB/year for all modules (SP-FD-22)

"Data Selection" is not *just* triggering---
May also include region-of-interest selection, high-level filtering

Like all experimentalists, we are both paranoid and greedy; we would like to keep as much as possible

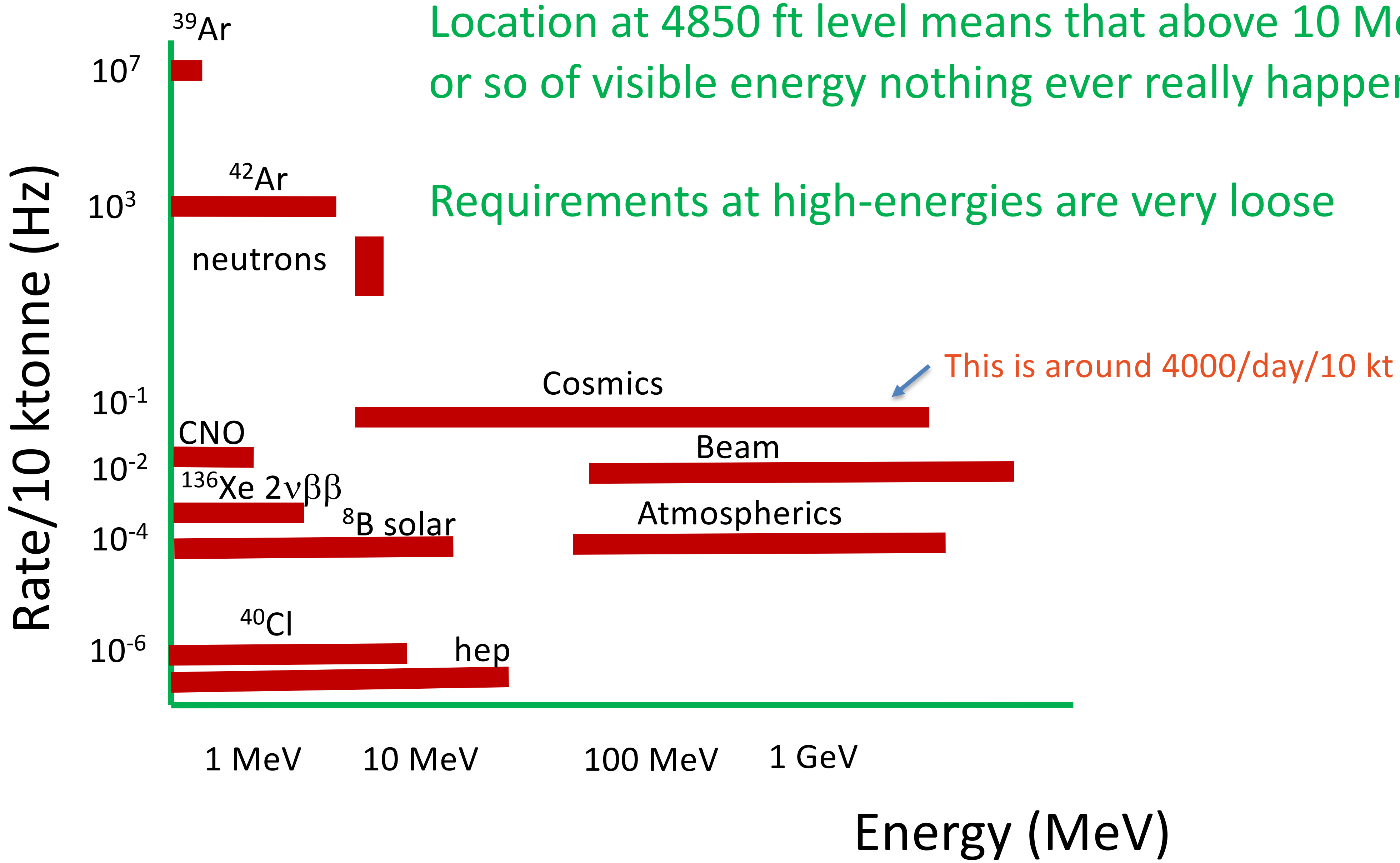
Data Selection

- Studies to date have been for HD APAs
- We expect algorithms depend only weakly on HD/VD differences
- Will highlight where we expect these to matter

Requirements

Location at 4850 ft level means that above 10 MeV or so of visible energy nothing ever really happens.

Requirements at high-energies are very loose



Requirements

Three Regimes

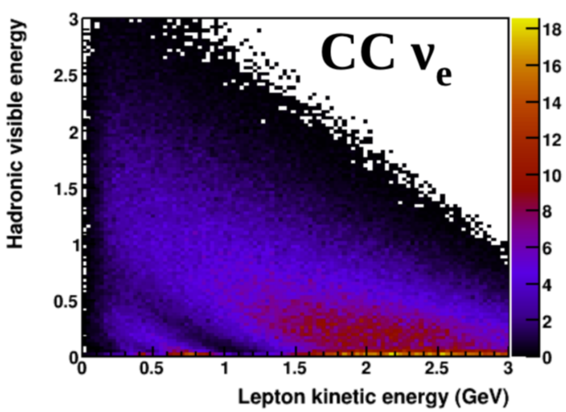
- “High” Energy – interactions producing more than 100 MeV of visible energy
- “Low” Energy – Below 100 MeV but practically speaking in 5-20 MeV regime
- Time-correlated – non-beam events that come in time-correlated ways, like SN bursts

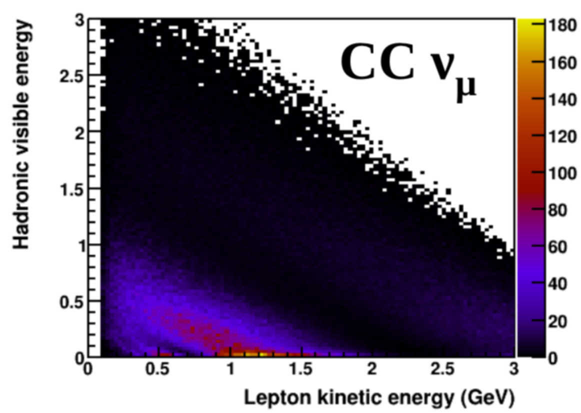
Requirements

"High" Energy Triggering

(Requirement SP-DAQ-8) The DAQ shall trigger and acquire data on visible energy deposition > 100 MeV. Data acquisition may be limited to the area in which activity was detected.

(Requirement FD-SciEng 882) The Far Detector shall be > 90% efficient for any interaction that leaves > 100 MeV of visible ionization energy inside the fiducial volume.
→The efficiency covers hardware and software performance of the detector while operating. Visible ionization is in electron equivalent ionization.





- Long baseline analyses assume 100% efficiency at >200 MeV
- Very little acceptance below 200 MeV because of NC background
- K+ from PDK can have KE < 50 MeV but total visible energy ~100 MeV

Requirements

"Low" Energy Triggering

620 (Requirement) The far detector shall have high efficiency for any interaction leaving < 100 MeV of visible ionization energy inside the fiducial volume.
→This requirement separates the performance at lower energies where supernova, solar events may occur. The lowest threshold is expected to be 5 MeV. The efficiency is expected to be a function of energy in this energy range.
→→The efficiency defined here is for TPC or Photon Detector combined.

2321 (Specification) The DAQ shall trigger and acquire data on visible energy deposition > 10 MeV of single neutrino interactions. Those triggers will normally be fired using a pre-scaling factor, in order to limit the data volume.
→Energy here is "visible energy", defined as deposited energy in the active volume as ionization and/or scintillation.

(We do not anticipate needing a pre-scale at E_{vis}~10 MeV, but ROI selection is likely)

Requirements

SN Burst Triggering

315 (Requirement) Far detector shall be capable of collecting low energy (<100 MeV) charged current electron neutrino interactions on Ar40 nucleus that arrive in a short period of time (<100 sec) . The final state electron shall be detected and its energy measured.
→ Most events from a supernova neutrino burst are neutrino absorption on Ar40 producing electrons from 5 to 100 MeV, arriving < one minute time frame. An electron and a K40 nucleus is in the final state. The total event count from a galactic supernova (10 kpc) is expected to be approximately 3,000 events for 40 kt of LAr.

2263 SP-FD-23 (Specification) The DAQ architecture shall provide a mechanism for triggering on galactic supernova bursts and recording neutrino interactions associated with those bursts over a 30 second period, with a goal of 100 seconds. During this period, the full raw data information must be stored.
→95% efficiency for a SNB producing at least 60 interactions with a neutrino energy >10 MeV in 12 kt of active detector mass during the first 10 seconds of the burst.

188 (Requirement) The FD data acquisition shall have the capability of notifying the world-wide science community of an astronomical event in a timely way.
→The SNEWs collaboration will require notification within minutes with only 1 false trigger allowed per year. To fulfill this requirement DAQ must have online capability of identifying a real SN burst signal.
→ →<~ 1 minute processing and response is desirable

Requirements

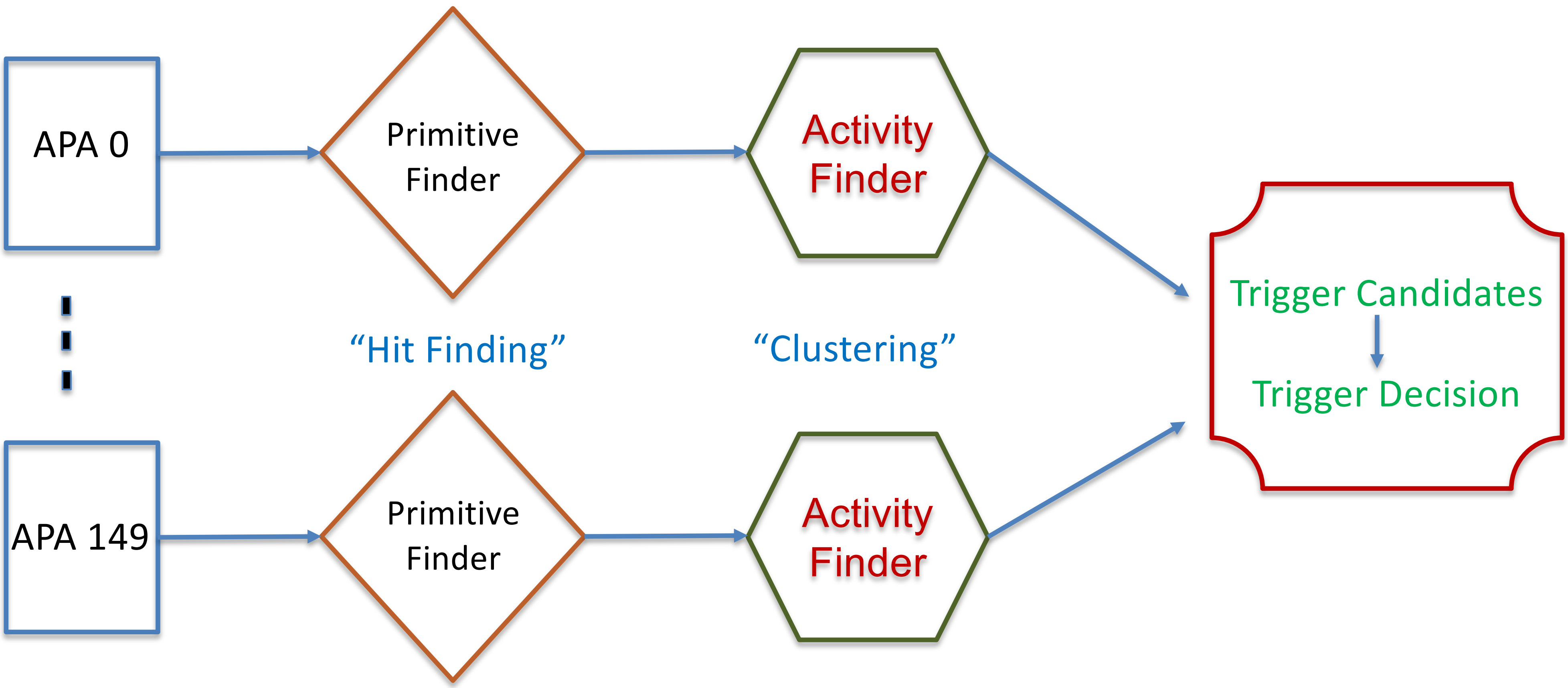
General DUNE Approach

We satisfy these requirements with the following general strategies:

- High-energy triggering is as inclusive as possible
- Low-energy triggering exploits TPC topological information to be semi-inclusive, assumes ROI selection to reduce data volume, and provides a summary data stream for ~zero threshold event analyses.
- SN Burst triggers exploit large DAQ buffers and write out everything for 10 s before and 90 s after detected burst

TPC Triggering

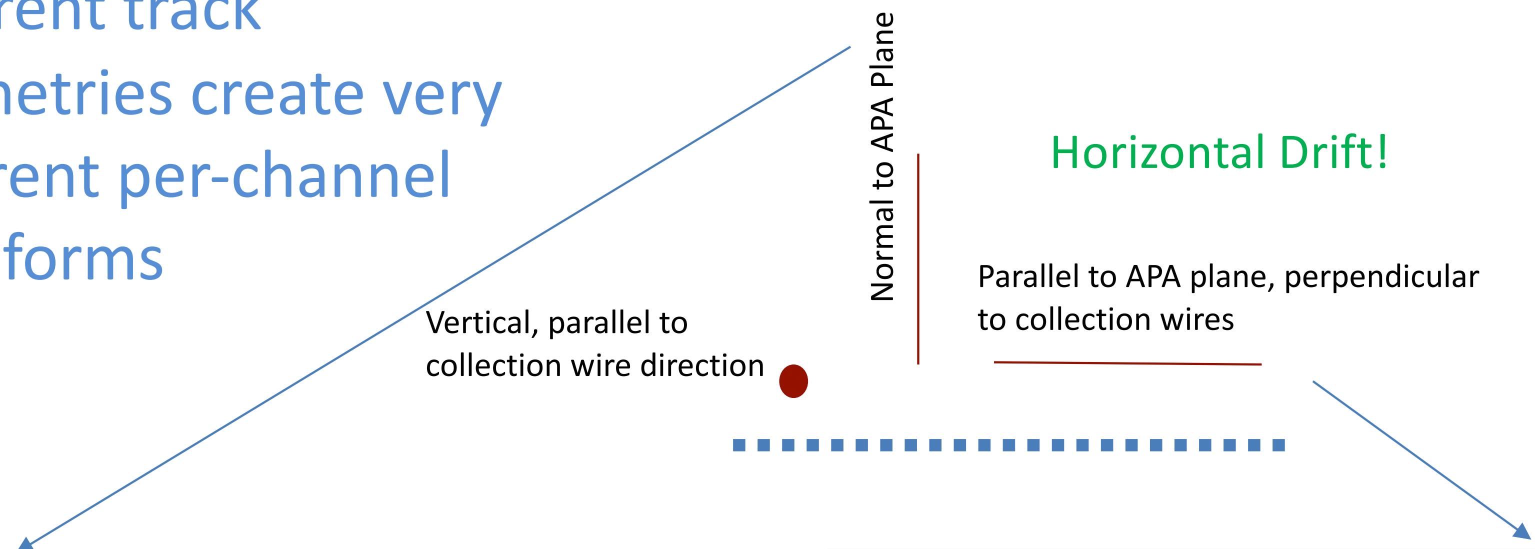
Basics

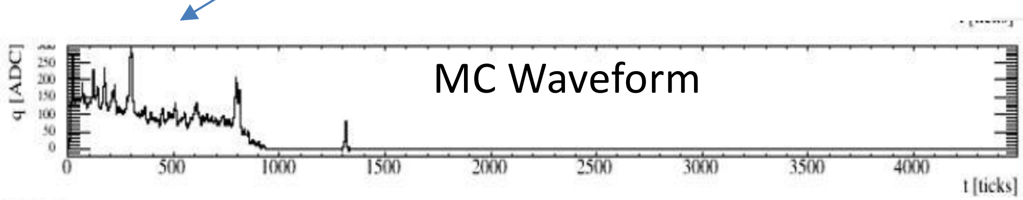


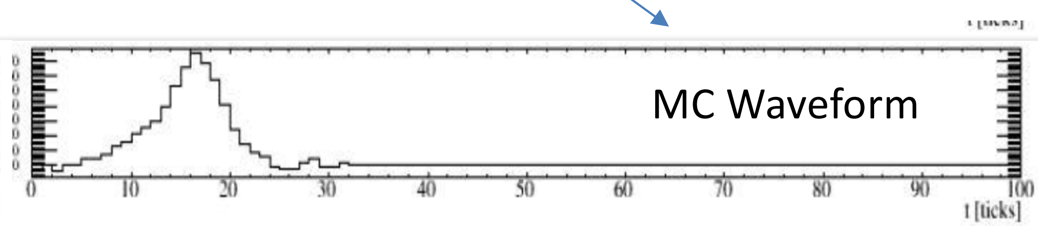
- Activity Finders = Low Energy, High Energy, Exclusive channels…
- Candidate Finder = SN bursts, High E event, beam event, etc.

Trigger Primitives (TPs)

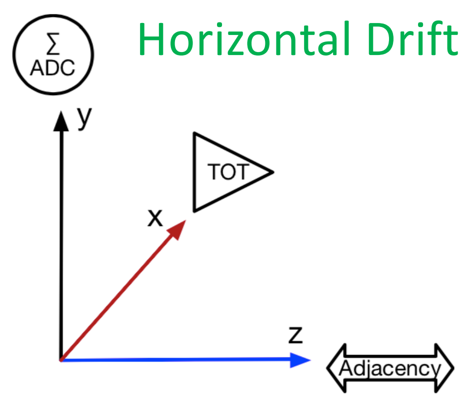
Different track geometries create very different per-channel waveforms







D. Last



Horizontal Drift!

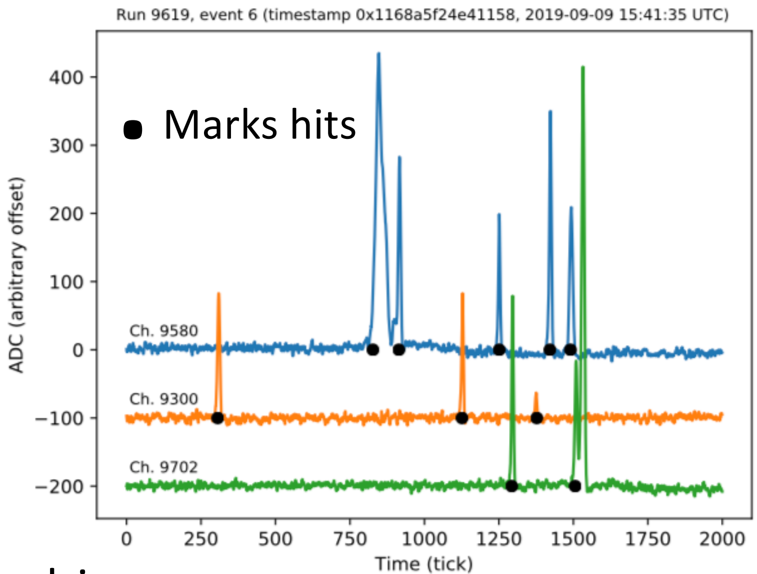
TPs include "orthogonal" basis of information
<u>(collection wires only so far—very little difference for VD)</u>

D. Rivera

Trigger Primitives (TPs)

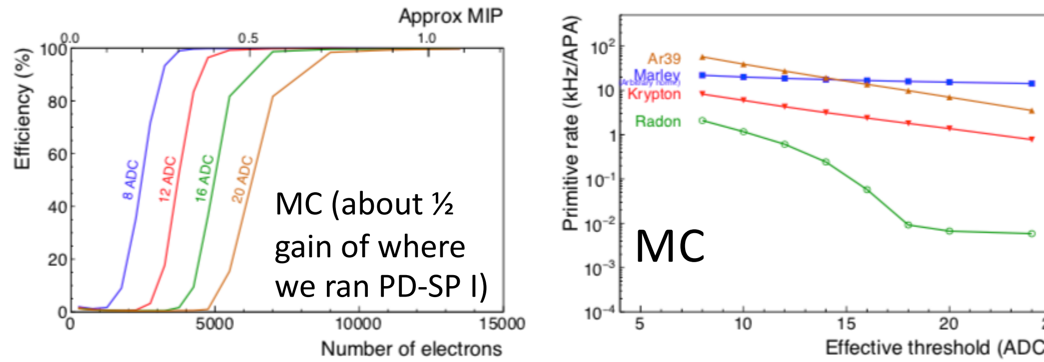
Hit finding

Example from ProtoDUNE-SP



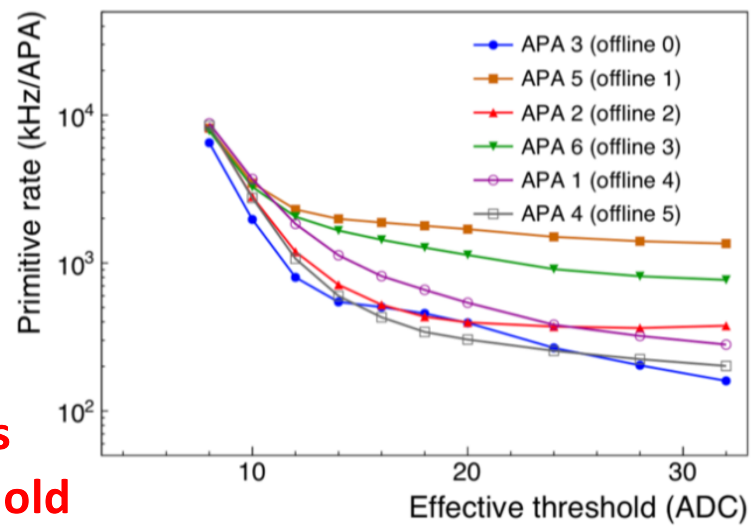
P. Rodrigues

Raw noise in ENC at PD I was ~600 e (collection)



TP threshold was around 1/4 MIP-equivalent, or around 250 keV_{ME} (per wire)

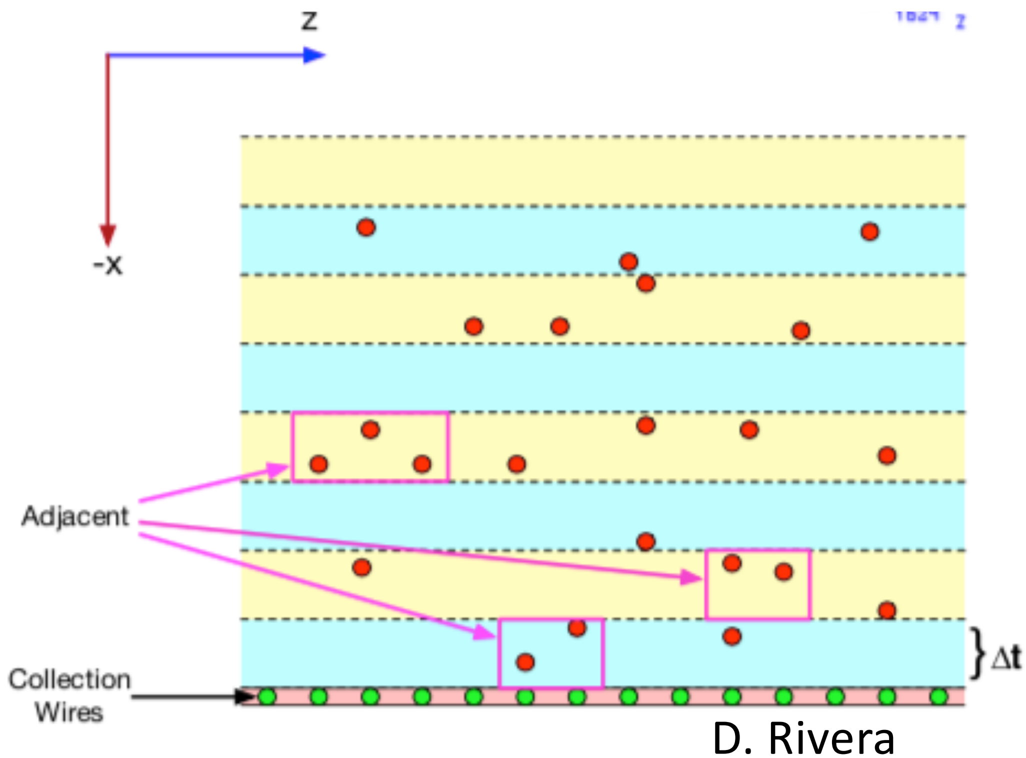
Raw noise RMS in PD I was 3-4 ADC above pedestal



Need to study this for VD strips instead of HD wires
Signal/Noise may be different and influence threshold

Trigger Activity
Clustering and Cutting

Cannot simply sum up all charge---in 10 ktonnes and a full drift, this is about a GeV of charge.



Need some kind of clustering as first stage of triggering

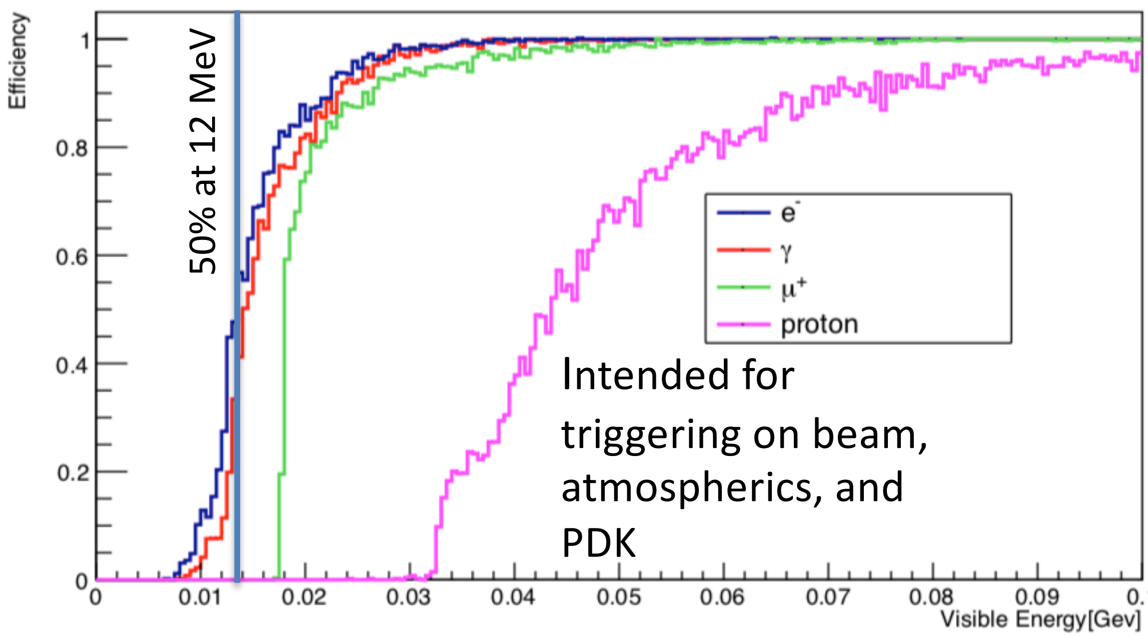
Radiological rate acceptable* with:

- N_{adj} >= 8 wires
- Cluster charge sum > 7000 ADC counts
- Max integrated wire charge > 6500 counts
- Max time-over-threshold >= 45 ticks

***acceptable rate was so that 5.4 ms readout of all channels had data rate < 25% of cosmic data rate**

High-Energy Trigger Efficiency

Differential



D. Rivera

Threshold is set by background rate!

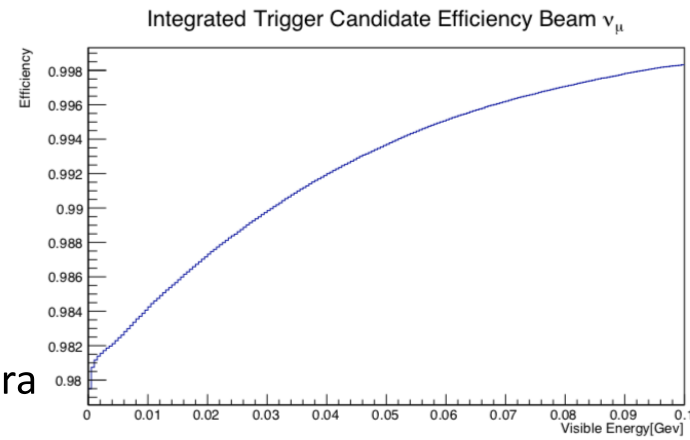
Moving to ROI readout allows a lower threshold

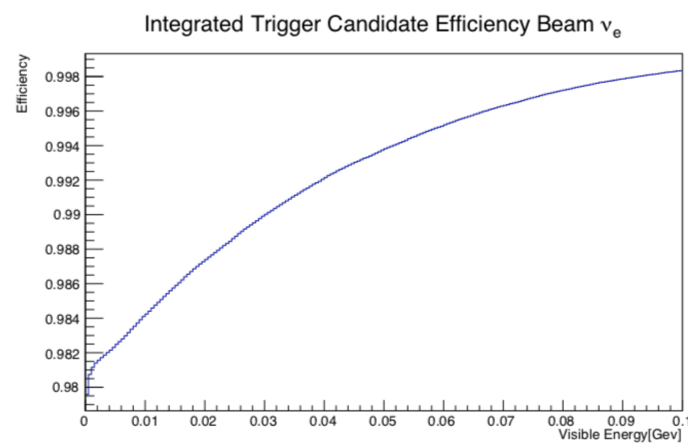
The trigger is intentionally *inclusive*

but different species have different topologies and thus different efficiencies

- Integrated efficiency ϵ_I is given by :

$$\epsilon_I(E_{vis}) = \frac{\int_{E_{vis}}^{\infty} n_{trig}(E)dE}{\int_{E_{vis}}^{\infty} n_{evt}(E)dE}$$





D. Rivera

Moving Lower in E

Region-of-Interest (ROI) Readout

Can have a higher trigger rate if data/trigger is smaller size:

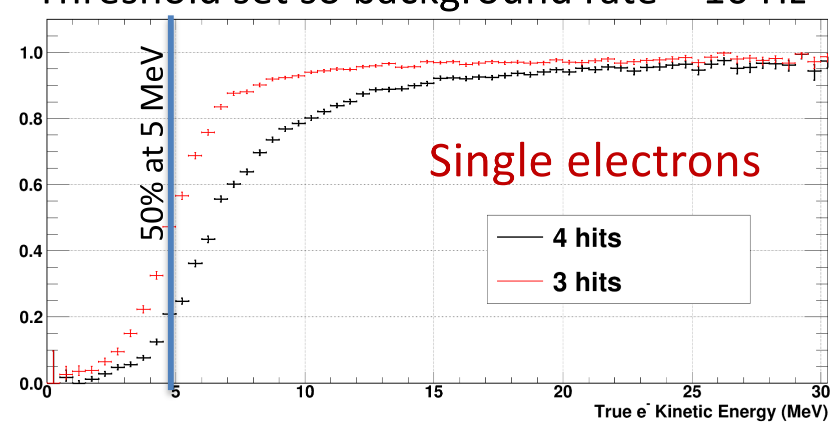
1. Halve readout window
2. Write out only APAs with trigger activity (TA)
3. Use a much narrower readout (100 μs) window around hits ("zero suppression")
4. Fully localize TA and use 100 μs window for readout

Table 2:

Data Reduction Approach	Event Size (Uncompressed)	Max Trigger Rate	Enabled Physics
Nominal	6.075 GB	0.078 Hz	Beam, NDK, Atm.
2.7 ms Readout Snapshot	3.3075 GB	0.156 Hz	Unknown
APA-Localization (Cosmics)	0.243 GB	1.95 Hz	*hep* solar ν
APA-Localization (Low-E)	0.041 GB	11.7 Hz	^{8}B solar ν, neutrons,Rn
Zero Suppression	0.040 GB	12.0 Hz	^{8}B solar ν, neutrons,Rn
TA Localization +Zero Suppression	14.6 kB	32.5 kHz	^{42}Ar, ^{40}Cl, *pep* solar ν?

Of course, ROI can depend on type of Trigger Candidate

Threshold set so background rate ~ 10 Hz

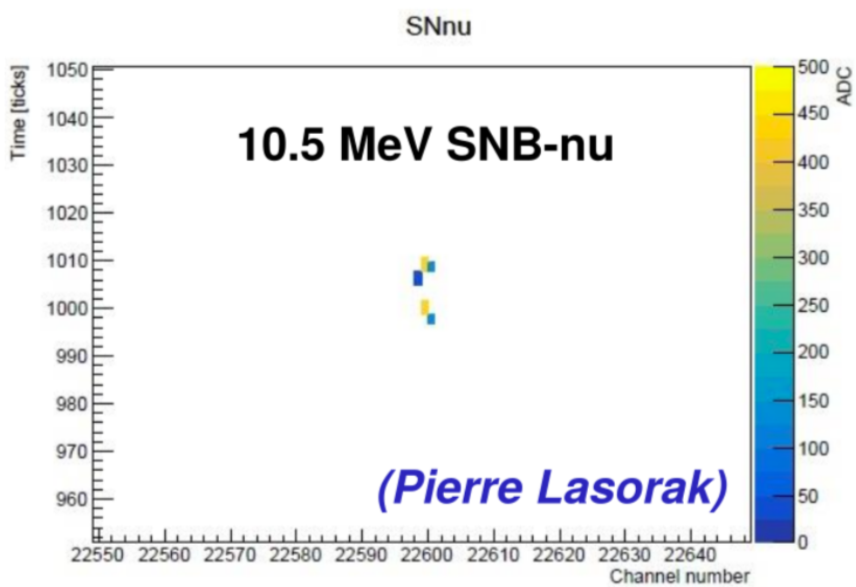


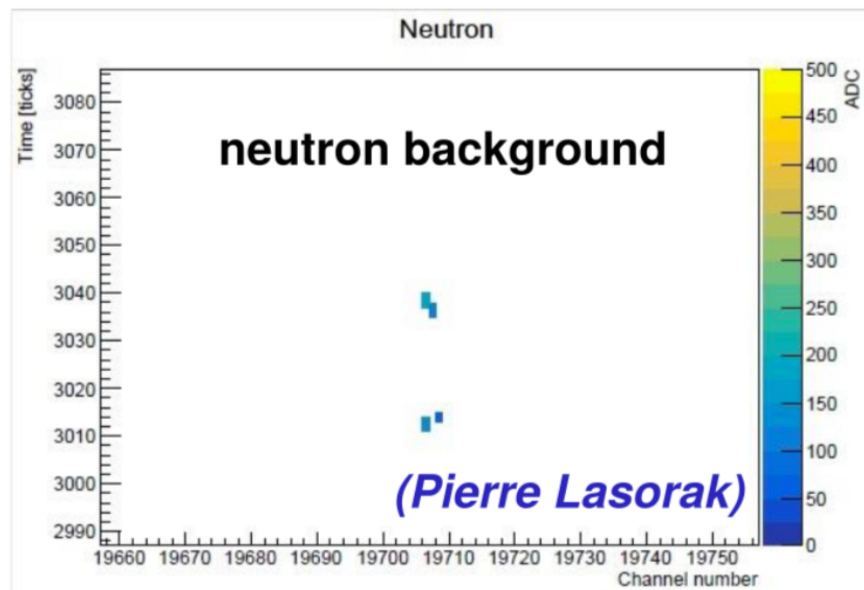
Thiago Bezerra

Moving Lower in E

What if we want to reduce neutrons?

Neutrons produce a γ cascade when they capture, electrons are more track-like
(But solar and SN neutrinos include de-excitation γs):



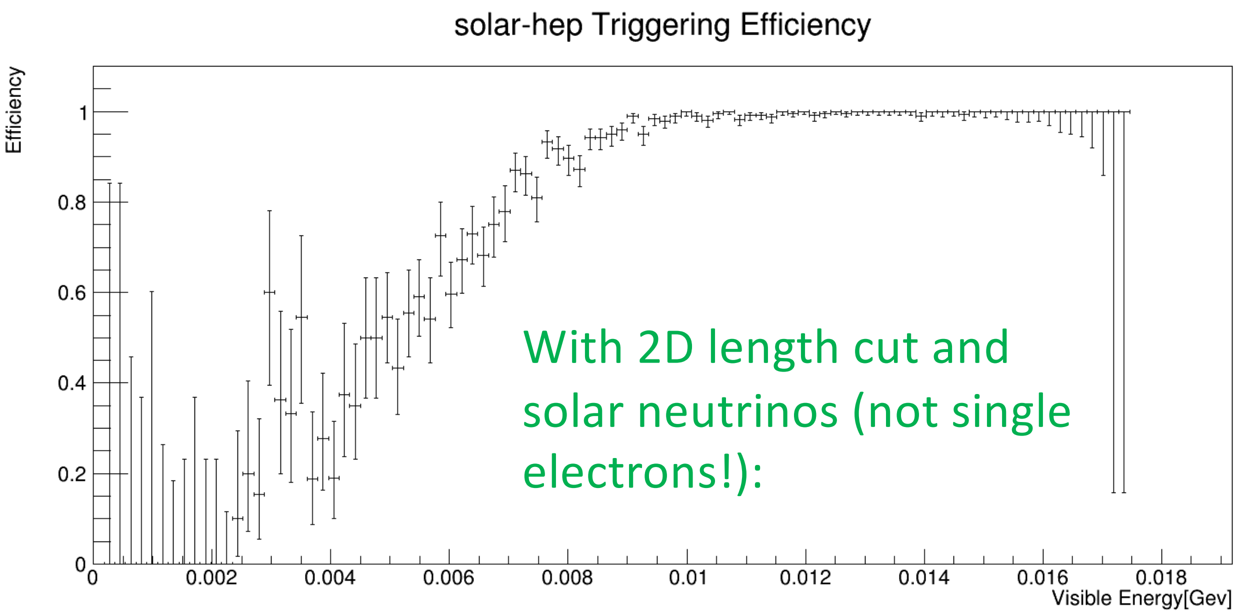


Moving Lower in E

Can exploit topology of TPC:

Use (collection-wire) primitives to create a "2D track length cut"

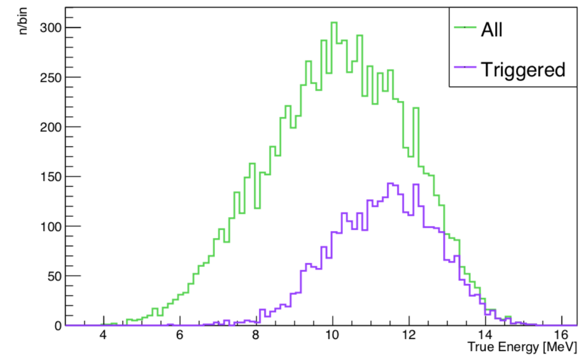
$$\mathcal{L} = \sqrt{(adj_{max} * 5mm)^2 + (TOT_{max} \times v_{drift})^2}$$



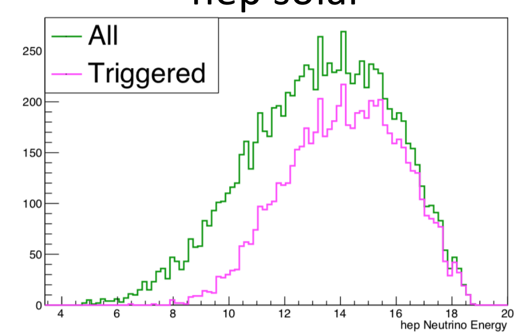
With 2D length cut and solar neutrinos (not single electrons!):

Trigger rate is ~ 100 Hz from backgrounds

^{8}B solar



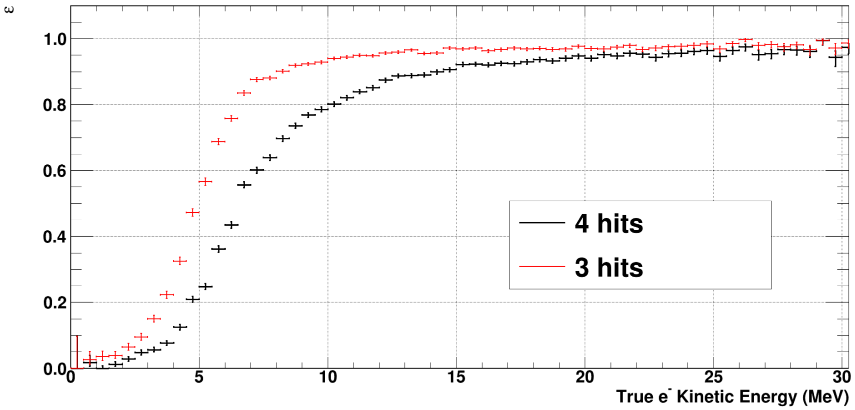
hep solar



D. Rivera

Supernova Bursts

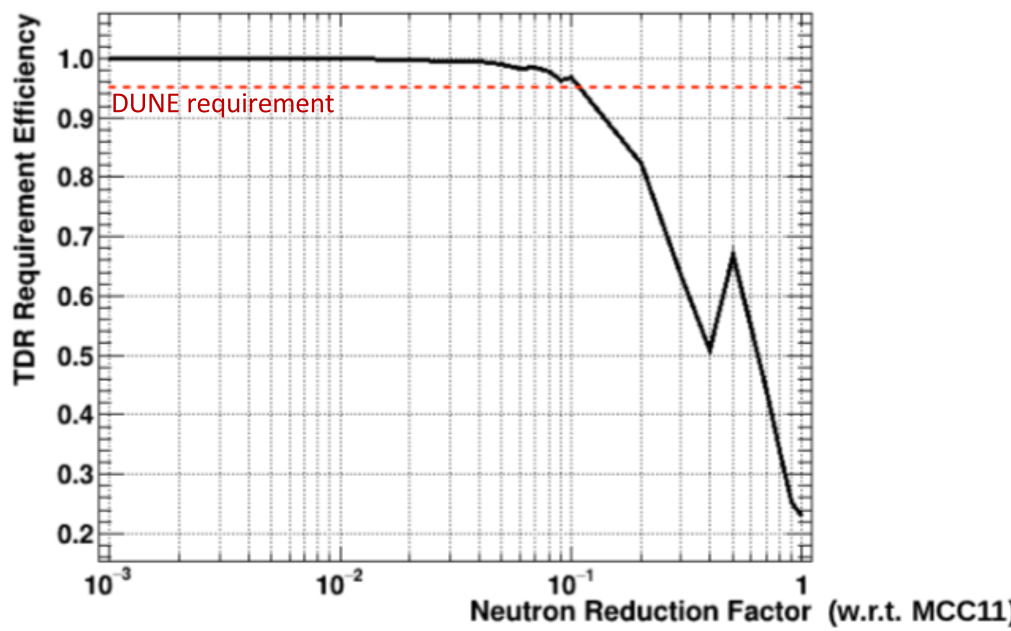
Additional handle: Time and energy profile



T. Bezerra

Can accommodate more aggressive "trigger activity" threshold but lowering single-interaction threshold hurts without energy-weighting

Not including energy profile



Note: we read out ***everything*** for 100 s if we detect a burst---event efficiency does not matter (much) except outside of that window

Changes for Vertical Drift

- Orientation of collection strips not critically different from APA collection wires
  - Induction strips/wires not currently used in triggering
  - Impact of a smaller collection pitch will have to be examined
- Signal/noise may be different and require/allow different TP thresholds
  - Top-side electronics may also have different response
- ^{39}Ar rate/channel higher because of higher imaged volume
- PDS System will be different, may allow greater opportunities for photon-based triggering

Summary

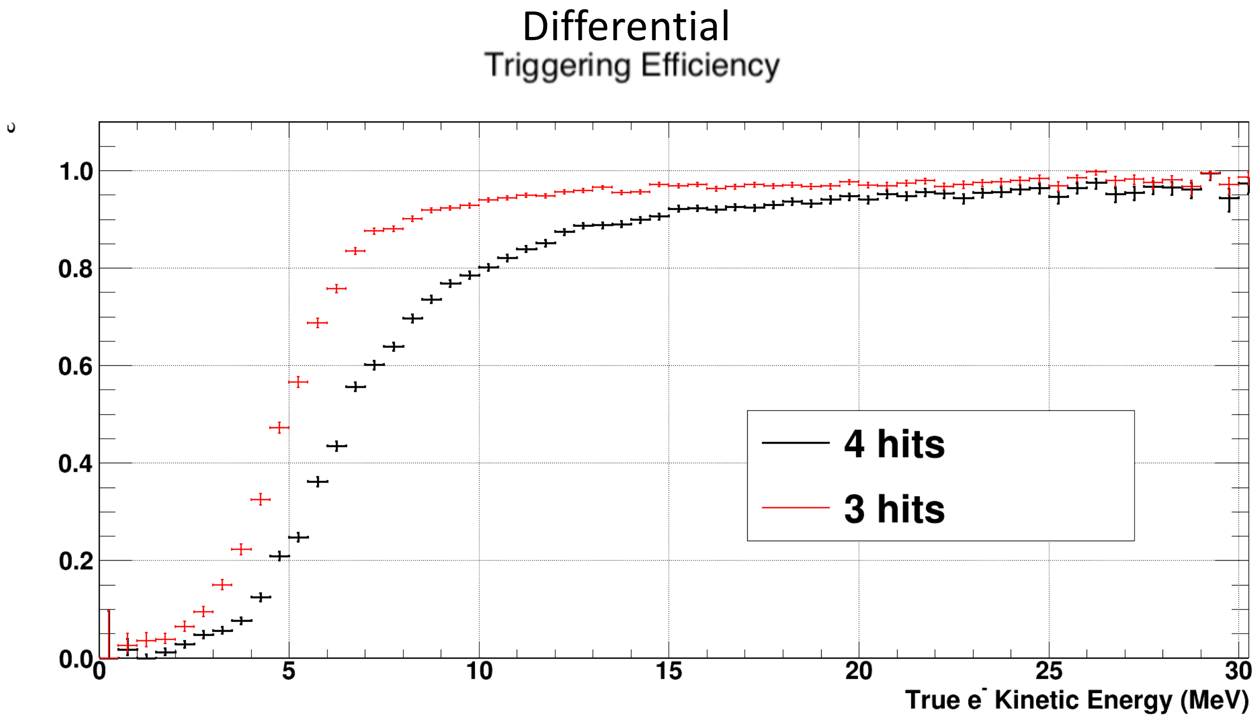
- Requirements on data selection exist but are fairly loose
- For HD, requirements satisfied by using TPC collection information
- Low-energy program can also be accommodated with tighter ROI for readout
  - Radiologicals will limit threshold more than system will
- Still need to test high-energy algorithms on VD collection strips
- Future work (PDS inclusion; induction wire Trigger Primitives; more aggressive SN burst finding) moving forward (see A. Thea talk)

Backups

Moving Even Lower in E

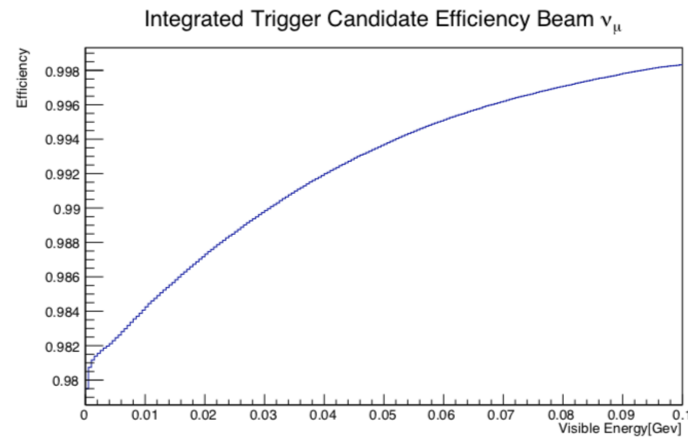
- Include u/v wire trigger primitives
  - This will be affected by HD/VD differences
- Write all of them continuously
- Threshold is ~250 keV depending on signal/noise
- Can run "offline" algorithms as sophisticated as desired
- Data will be overwhelmingly ^{39}Ar

High-Energy Trigger Efficiency



- Integrated efficiency ϵ_I is given by :

$$\epsilon_I(E_{vis}) = \frac{\int_{E_{vis}}^{\infty} n_{trig}(E)dE}{\int_{E_{vis}}^{\infty} n_{evt}(E)dE}$$



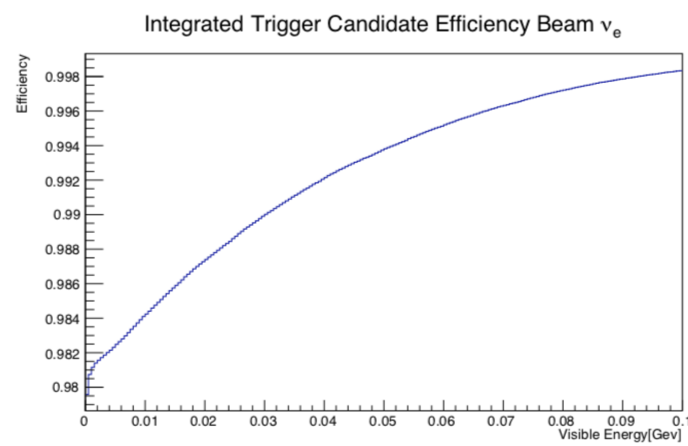
Thiago Bezerra

Threshold is set by background rate!

Moving to ROI readout allows a lower threshold

The trigger is intentionally *inclusive*

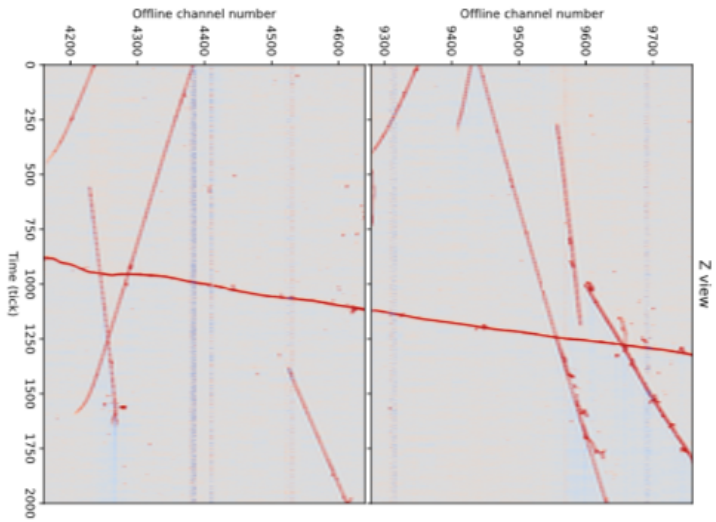
but different species have different topologies and thus different efficiencies

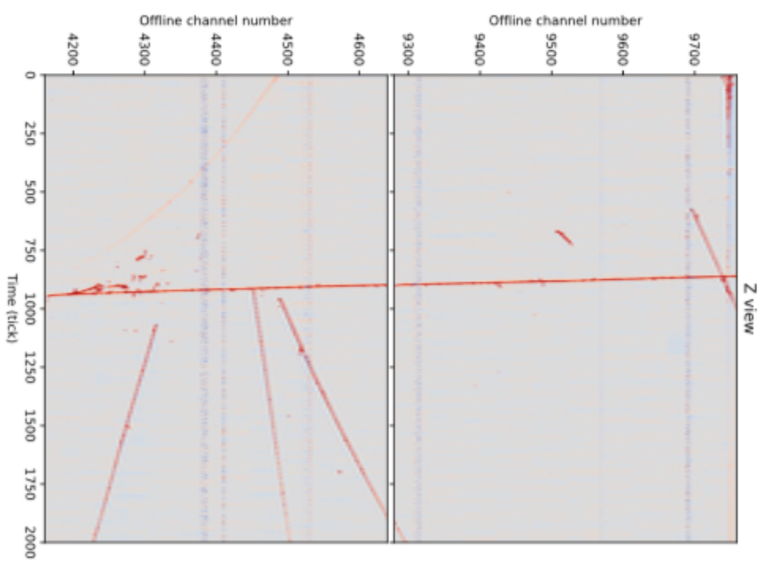


D. Rivera

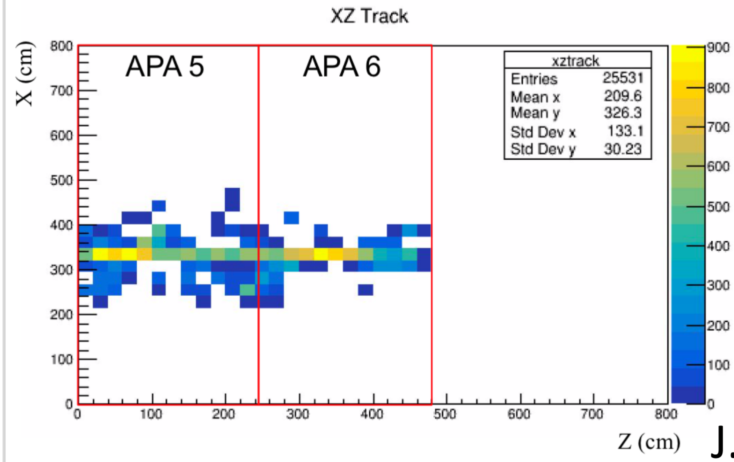
Performance at PD-SP 1

Horiztonal Muon Trigger (exclusive)

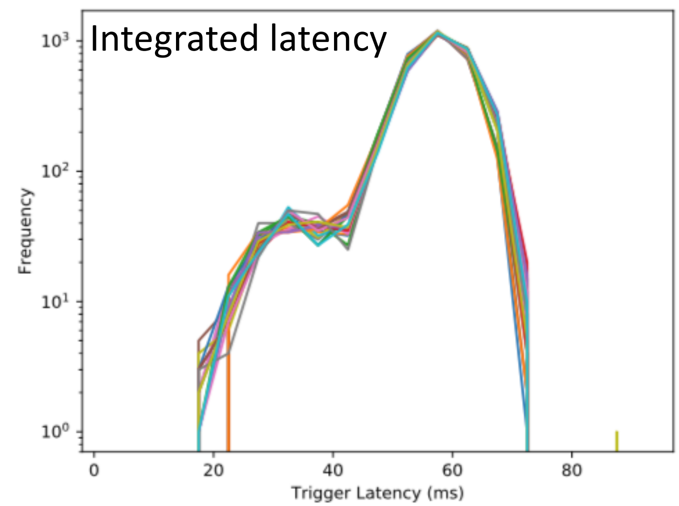




Average over several events



J. Sensenig



Time from data arriving at FELIX to Trigger Decision
Buffer depth was 1 second

DUNE Trigger Basics

- Primary detector system (TPC) is very slow, data buffered for 10 s
- Leaves plenty of time to exploit TPC topology in trigger
- PDS system can also be used to trigger independently or w/ TPC
- (Calibrations and other auxiliary systems can also trigger)
- System intended to be as *inclusive* as possible
  - e.g., everything with E > threshold
  - Constraint here will be from low-energy backgrounds (later)
- Event readout can be:
  - All channels for some time window ~ drift
  - A narrow region of interest (ROI) in channel and time space
  - A very long time window (for supernovas) up to 100 s for all channels

TPC Triggering

Basics

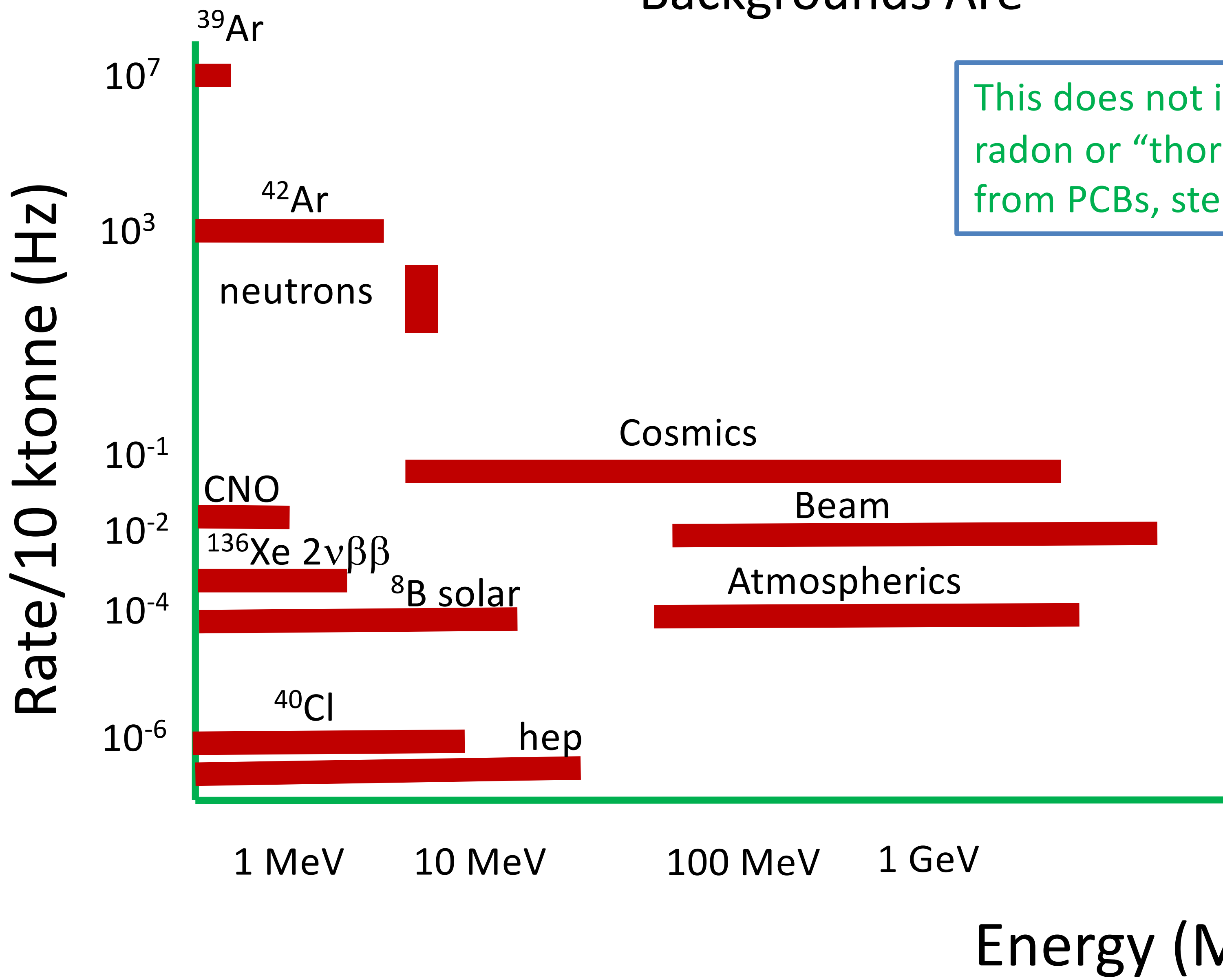
A quick note about the word “threshold”….

- Trigger Primitives have a “hit” threshold
- Trigger Decision based on various “event” thresholds
- Supernova Bursts have a “burst threshold”
  - Once a burst is triggered, the data is acquired with zero threshold for 100 s

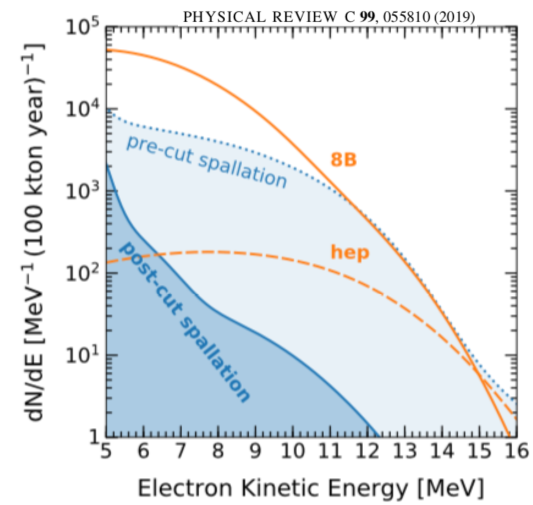
Triggering is Not the Challenge

Backgrounds Are

This does not include unsupported radon or "thoron," or U/Th gammas from PCBs, steel, APAs, or $^{40}Ar(\alpha,\gamma)$



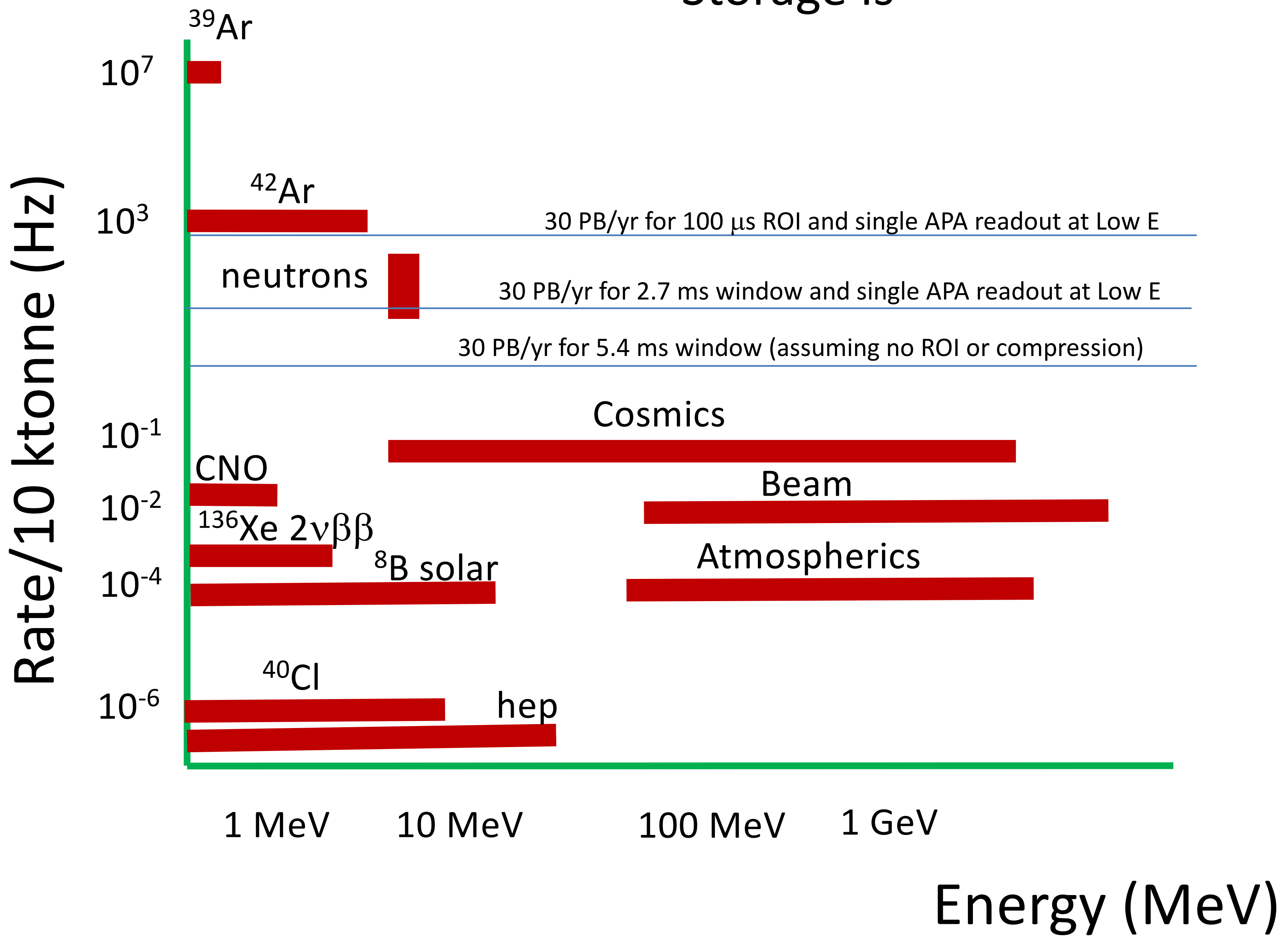
Zhu, Li, Beacom



Or spallation

Triggering is Not the Challenge

Storage Is

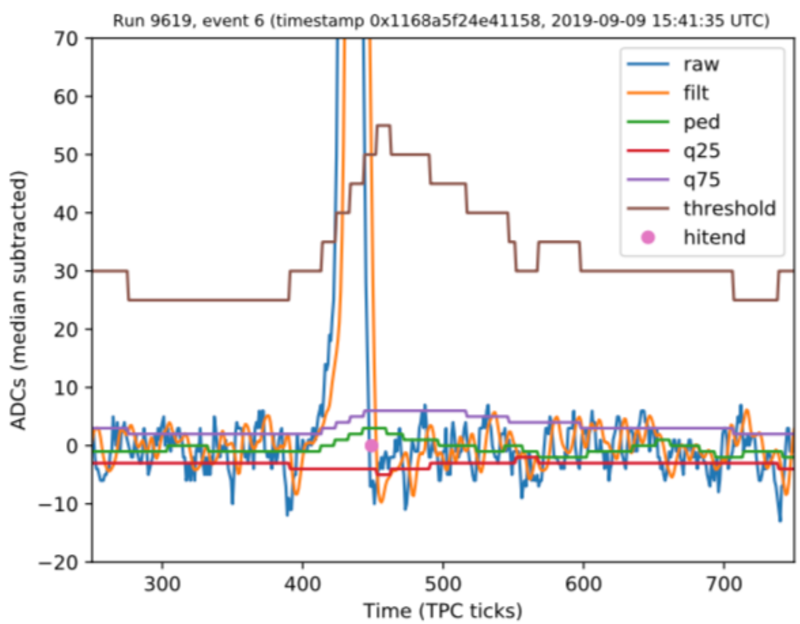


Trigger Primitives (TPs)

Hit finding

At ProtoDUNE-SP I tried two approaches:

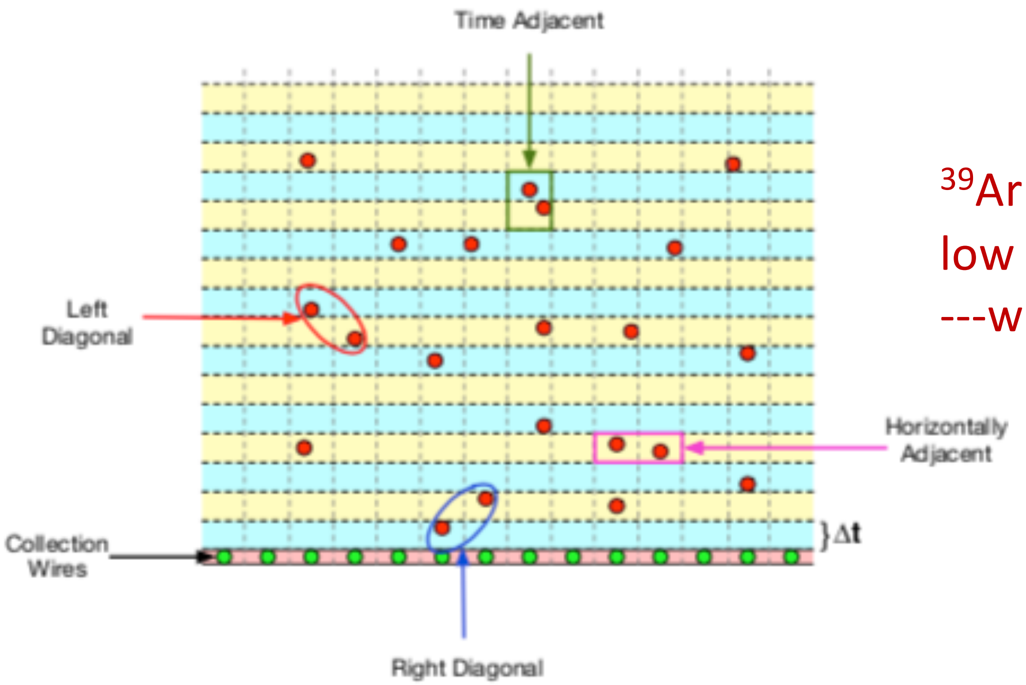
2. Noise RMS-dependent threshold
    - Find baseline and rms via "frugal streaming" (arXiv:1407.1121)
    - Filter (7-tap FTIR)
    - Apply threshold in noise sigma (5σ for PD-SP I)



P. Rodrigues

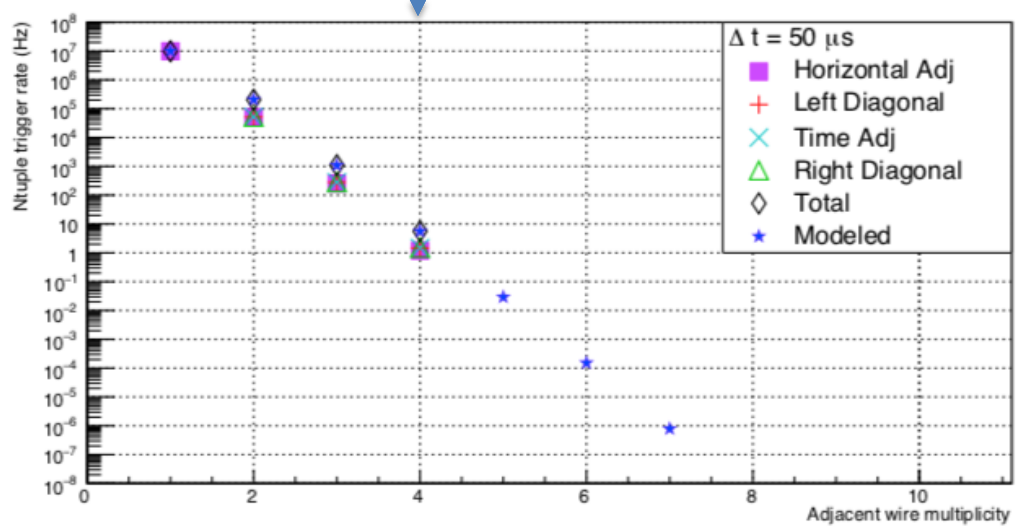
Triggering is Not the Challenge

Timing Is



^{39}Ar is such a high rate that pileup can create low E track-like events in slow drift medium ---worse for longer drift volumes.

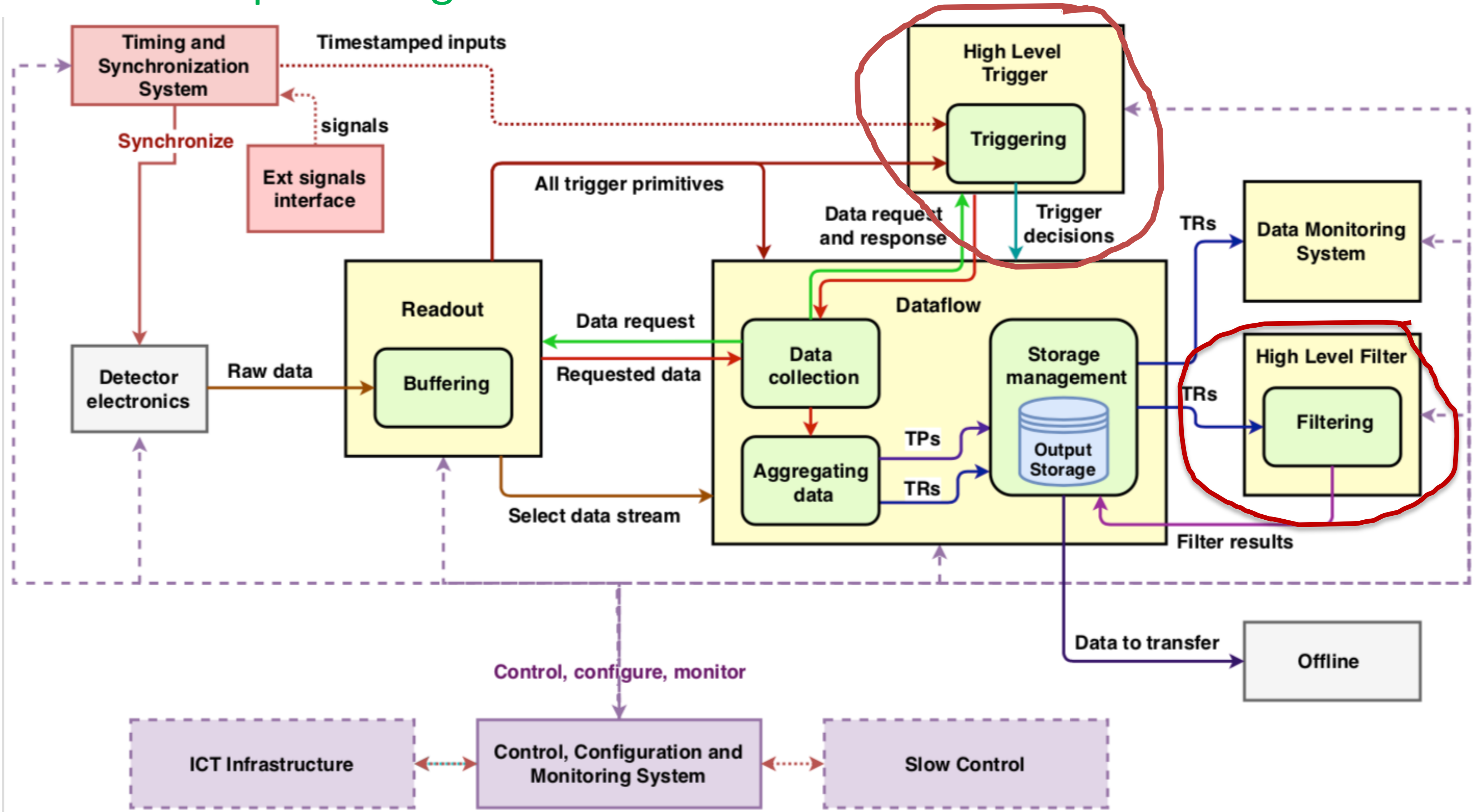
About 4 MeV MIP-equiv



D. Rivera

TPC Triggering

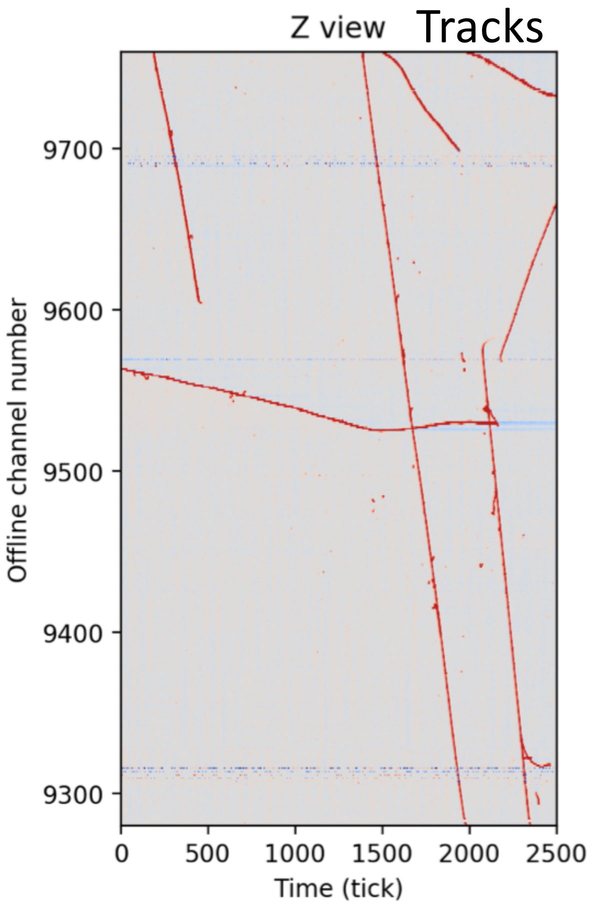
More Complete Diagram

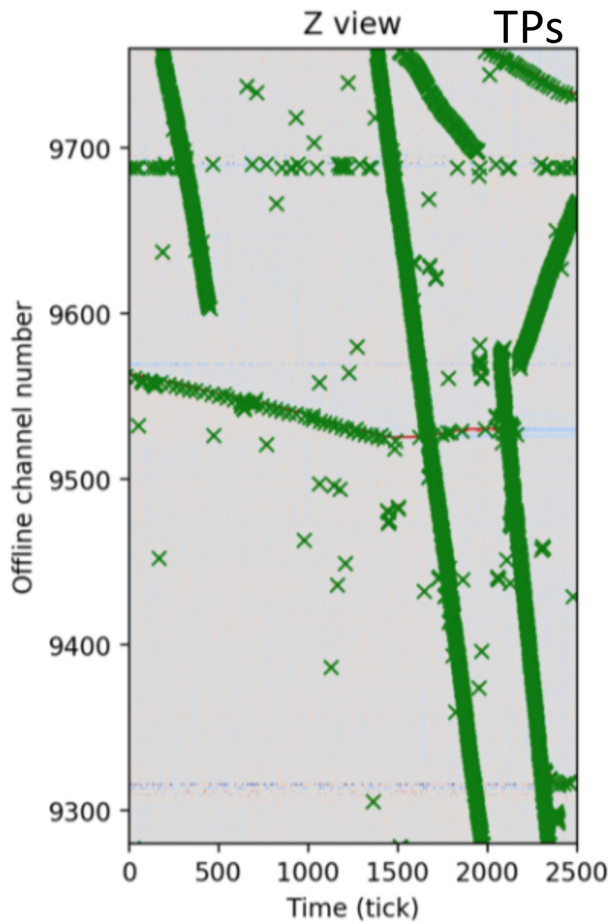


Trigger Primitives (TPs)

Hit finding

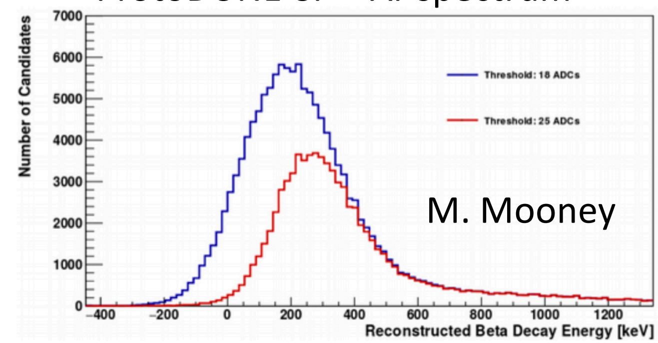
Run 11044, event 5 (timestamp 0x11955baa4c000a0, 2020-03-09 17:22:51 UTC)



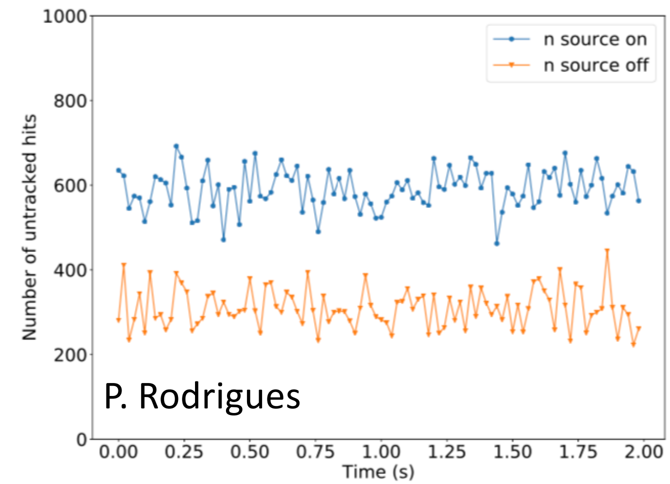


TP threshold was around 1/4 MIP-equivalent, or around 250 keV_{ME} (per wire)

ProtoDUNE-SP ^{39}Ar spectrum

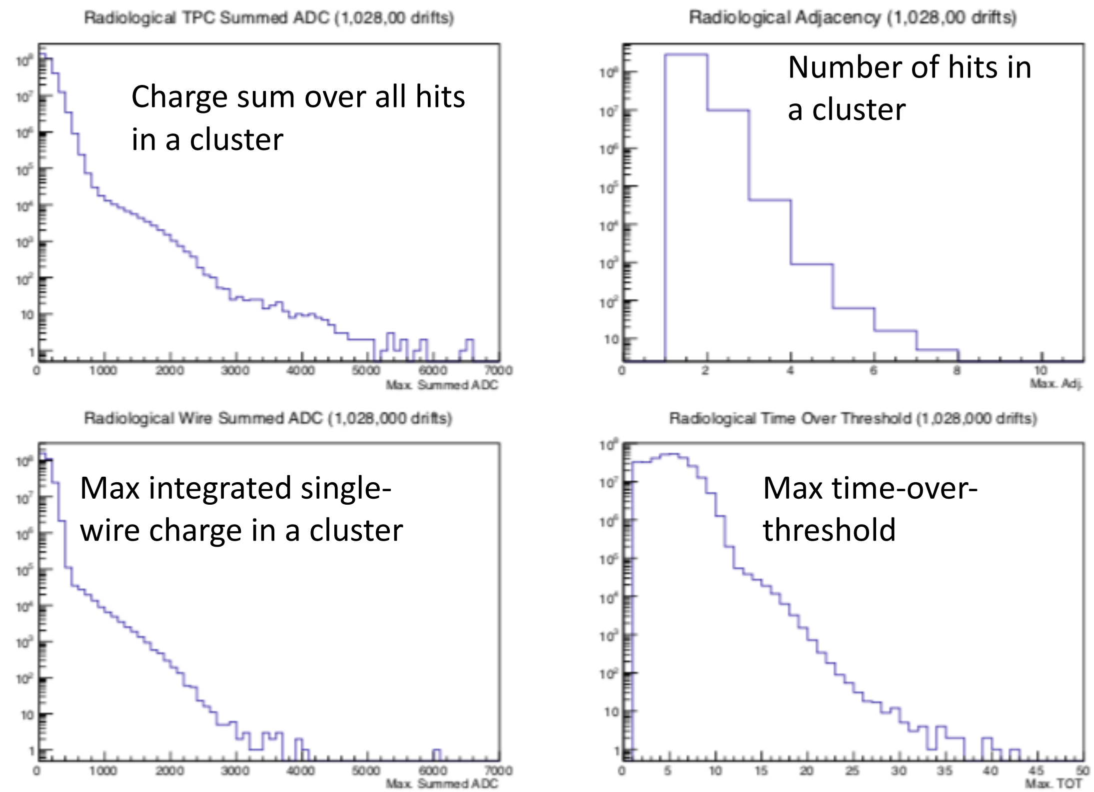


Some efficiency for neutrons



P. Rodrigues

Trigger Candidates
Cutting



Rivera and Last

Radiological rate acceptable* with:

- N_{adj} >= 8 wires
- Cluster charge sum > 7000 ADC counts
- Max integrated wire charge > 6500 counts
- Max time-over-threshold >= 45 ticks

***acceptable rate was so that 5.4 ms readout of all channels had data rate < 25% of cosmic data rate**

Moving Lower in E

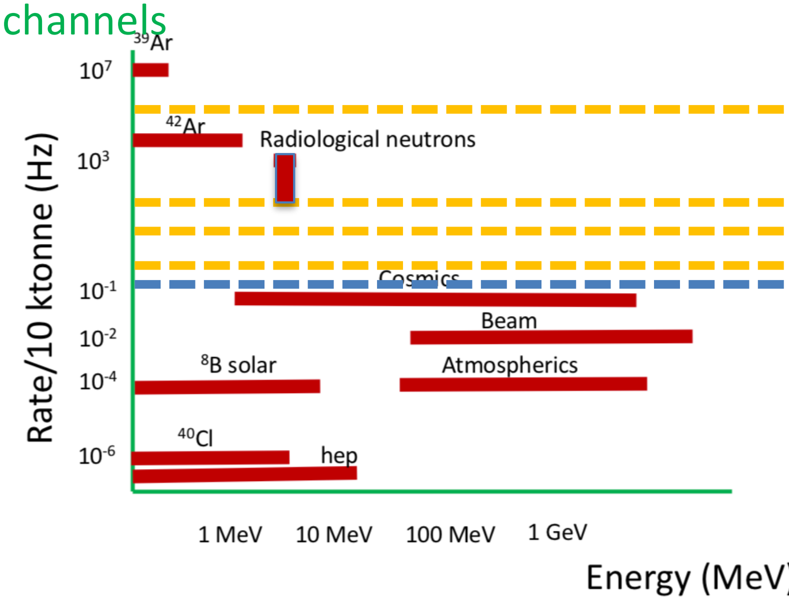
Region-of-Interest (ROI) Readout

Can have a higher trigger rate if data/trigger is smaller size:

1. Halve readout window
   - "Free" because for low E events there is no trigger ambiguity
2. Write out only APAs with trigger activity (TA)
   - At low E, pretty safe and big reduction
   - Cosmics require ~6 APAs on average
3. Use a much narrower readout (100 μs) window around hits ("zero suppression")
   - Big enough for any deconvolution in processing
   - Will definitely do this for laser calibrations
4. Fully localize TA and use 100 μs window for readout
   - Most aggressive; need to read out "box" around hit channels

Table 2:

Data Reduction Approach	Event Size (Uncompressed)	Max Trigger Rate	Enabled Physics
Nominal	6.075 GB	0.078 Hz	Beam, NDK, Atm.
2.7 ms Readout Snapshot	3.3075 GB	0.156 Hz	Unknown
APA-Localization (Cosmics)	0.243 GB	1.95 Hz	*hep* solar ν
APA-Localization (Low-E)	0.041 GB	11.7 Hz	^{8}B solar ν, neutrons,Rn
Zero Suppression	0.040 GB	12.0 Hz	^{8}B solar ν, neutrons,Rn
TA Localization +Zero Suppression	14.6 kB	32.5 kHz	^{42}Ar, ^{40}Cl, *pep* solar ν?



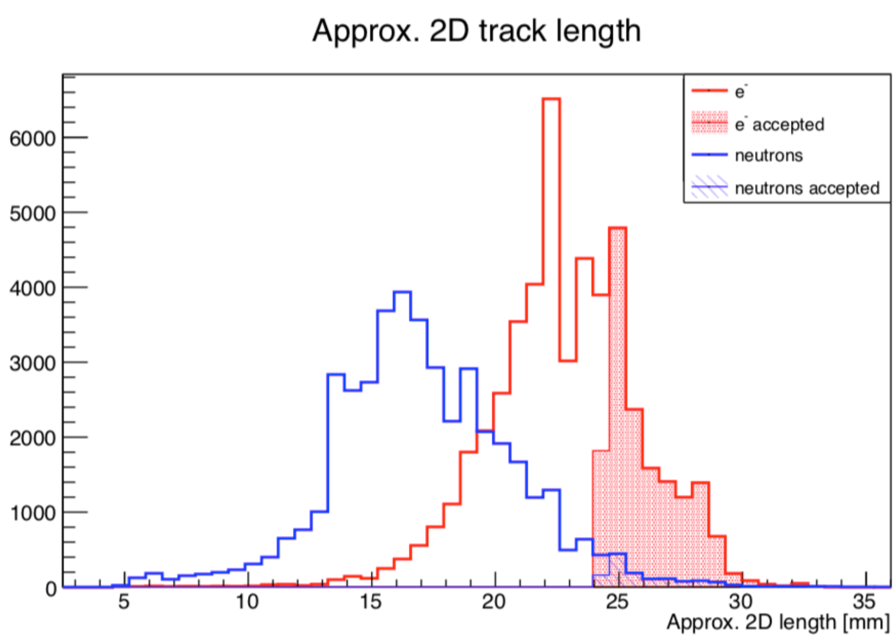
Of course, ROI can depend on type of Trigger Candidate

Moving Lower in E

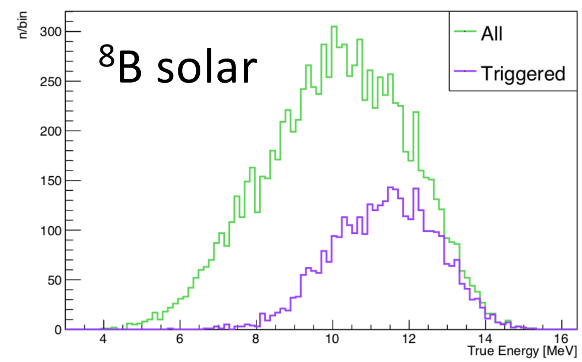
Can exploit topology of TPC:

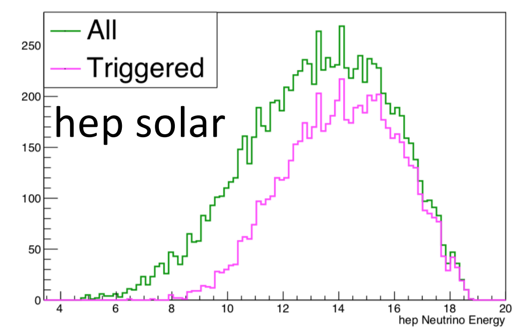
Use (colletion-wire) primitives to create a “2D track length cut”

$$\mathcal{L} = \sqrt{(adj_{max} * 5mm)^2 + (TOT_{max} \times v_{drift})^2}$$



D. Rivera



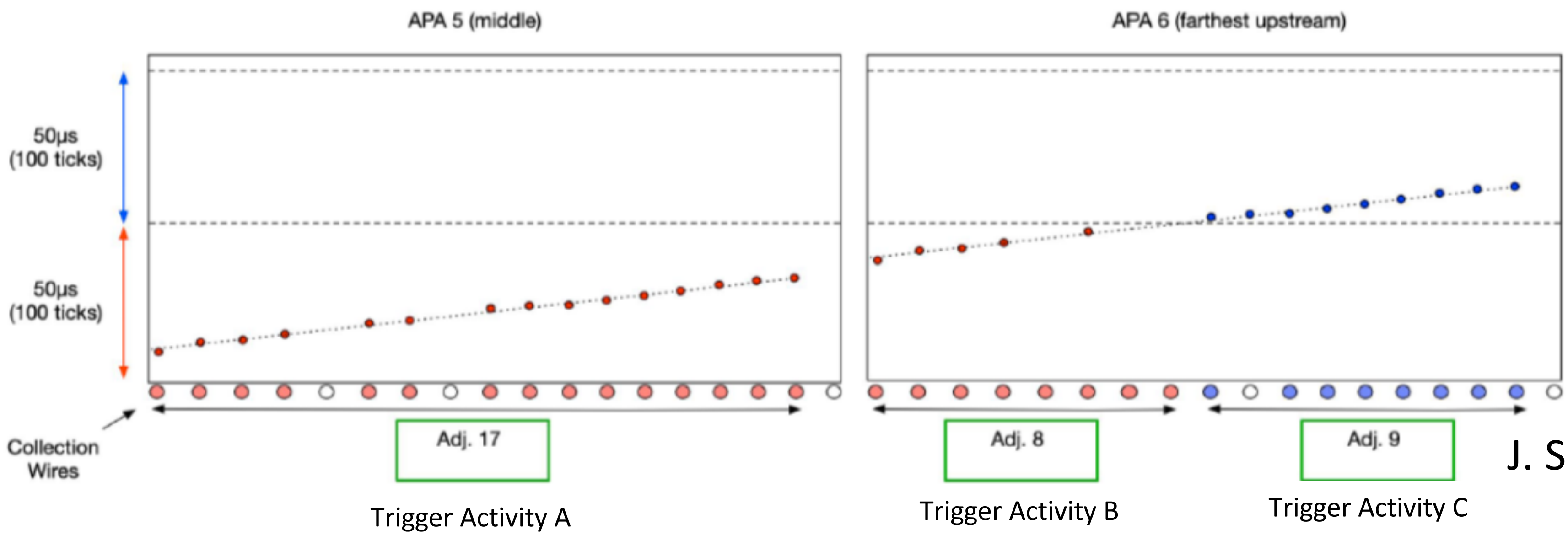


Further Plans

- Including PDS
  - Naturally inclusive trigger (e.g., Npe>threshold)
    - Trigger threshold maps to energy pretty cleanly and simply
    - But that also means higher background rates
    - Also will depend on channel-level thresholds
    - And depends strongly on light yield and uniformity
  - Efficiency likely easy(-ier) to model
    - Can be calibrated and measured relatively easily
  - Can be fast
    - Helps reduce spallation-induced fake supernova bursts
    - But readout buffers are so big (10 s) that speed not critical for trigger decision
  - Can reduce background rates via fiducialization---even better for VD!
    - External neutrons and γs will capture/convert near edges of volume
    - Can get lower trigger rates for low-energy physics
    - Might do this in High-Level Filter
  - Noise uncorrelated with TPC
    - Can help reject unexpected triggers from noisy wires
  - In principle singlet/triplet PID can reject αs for very low-E program with u/g Ar
    - Xenon loading is an interesting opportunity/complication
    - Will need to require nitrogen (and other) contamination to be very low

Implementation at PD-SP 1

- Cosmic rate too high for an inclusive trigger
- Picked reasonably low-rate exclusive channel---horizontal muons
- Allowed us to exercise entire TP→TA→TC chain



J. Sensenig

Trigger activity from individual APAs stitched together to form Trigger Candidates (Gaps of up to 4 wires allowed in Trigger Activity cluster)