



Status report on
DNNROI sigproc

Hokyeong Nam
Chung-Ang University

HfFilter and Truth Smearing Parameters in Summary

Detector/Filter	Wiener_tight_U	Wiener_tight_V	Wiener_tight_W	Gaus_wide	Wire_ind
PD-SP	Sigma: 0.148788 Power: 3.76194	Sigma: 0.1596568 Power: 4.36125	Sigma: 0.13623 Power: 3.35324	Sigma: 0.12	Sigma: ~* 5.0
PD-HD	Sigma: 0.221933 Power: 6.55413	Sigma: 0.222723 Power: 8.75998	Sigma: 0.225567 Power: 3.47846	Sigma: 0.12	Sigma: ~* 0.75
PD-VD	Sigma: 0.286735 Power: 8.07074	Sigma: 0.289041 Power: 9.94129	Sigma: 0.207127 Power: 4.13956	Sigma: 0.188	Sigma: ~*0.75

☐ Parameters in sp-filters.jsonnet

Detector/Smear	smear_long_U	smear_long_V	smear_long_W	smear_tran_U	smear_tran_V	smear_tran_W
PD-HD	2.691862..	2.675020..	2.713756..	0.737721..	0.715776..	0.139806..
PD-VD	1.68404	1.68404	1.68404	0.737721..	0.715776..	0.139806..

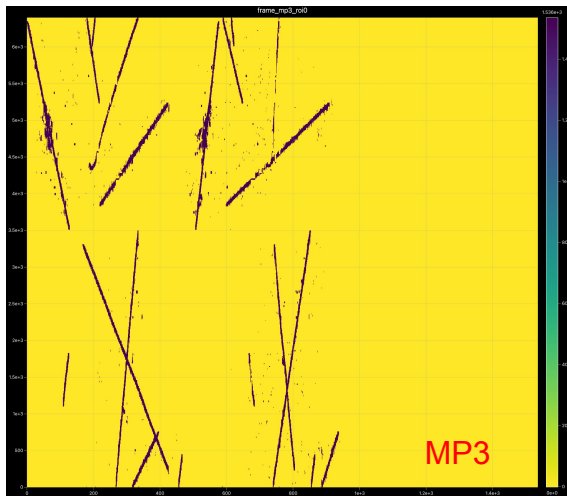
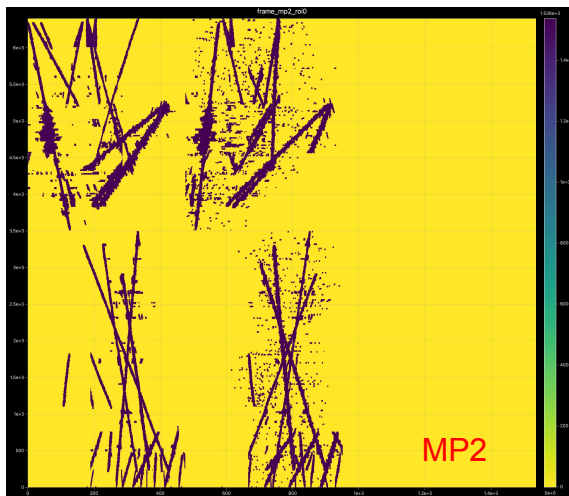
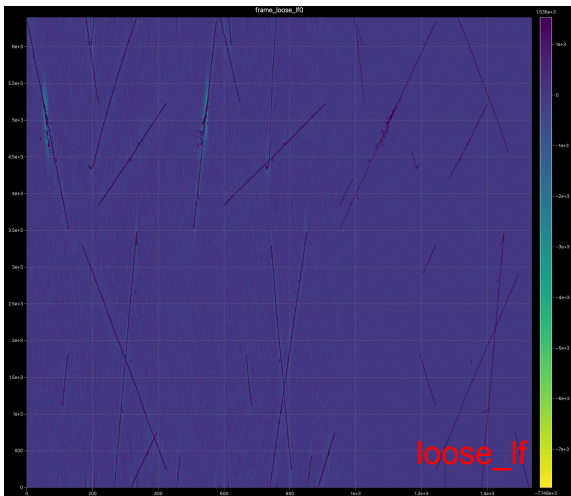
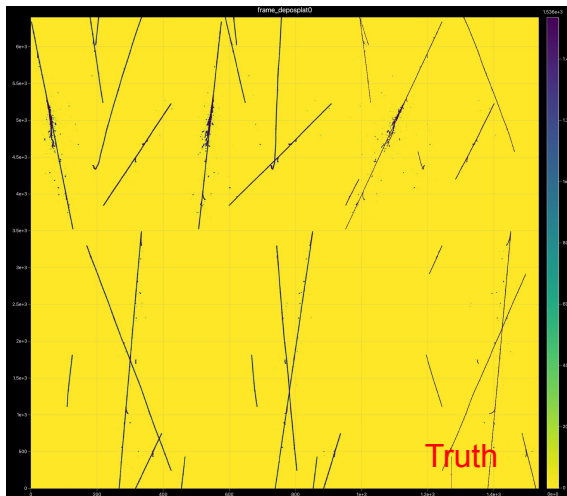
☐ Parameters in funcs.jsonnet

Training Sample Production

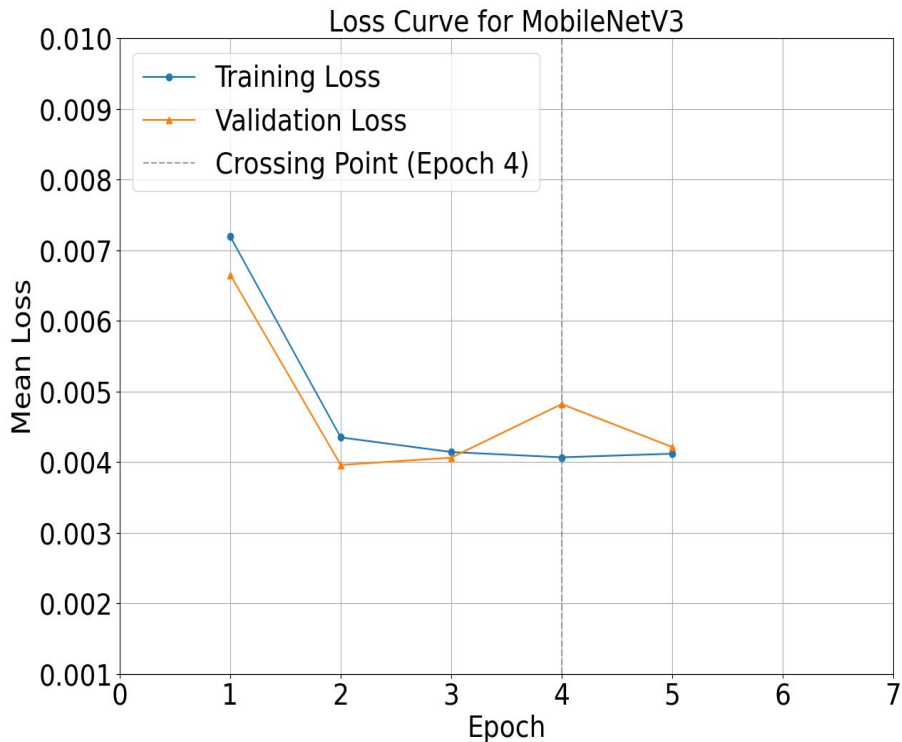
- Setup
 - Machine: dunegpvm07, dunegpvm14
 - dunesw: v10_05_00d00 with e26:prof qualifier
 - LarSoft: art 3.14.04
 - Wire-Cell: 0.30.3
- Organized by CRP
 - crp1, crp2, crp3, crp4
 - each CRP contains 250 reco files and 250 truth files
 - each file includes 2 events
- # of total images available: 2000
- Location:
 - on dunegpvms: /pnfs/dune/scratch/users/hnam/train_data_PDVD_rate2M_window3p2
 - on the WC cluster: /nfs/data/1/hnam/train_data_PDVD_rate2M_window3p2

Training Sample Production

- Additional information:
 - Sampling rate: 2 MHz
 - Time window: 3.2 ms (3.0 is default in sim)
 - Wire range: 0 - 1536
 - Time ticks: 6400
 - Naming convention: g4-rec or tru-crp<crp>-<run>-<date>.h5



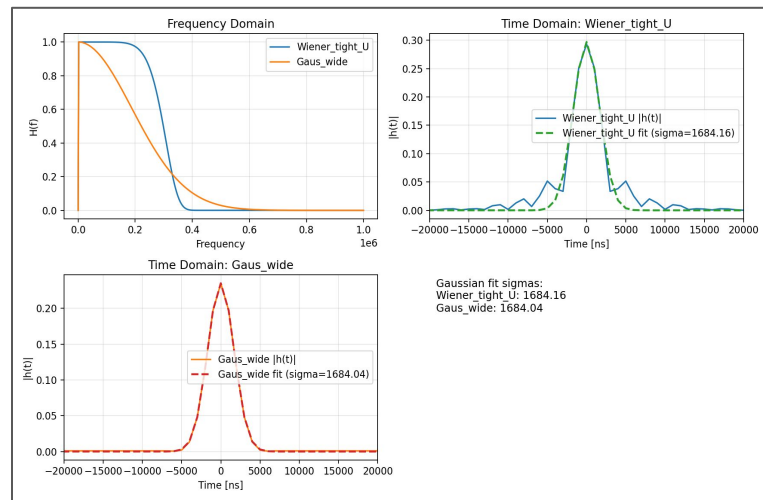
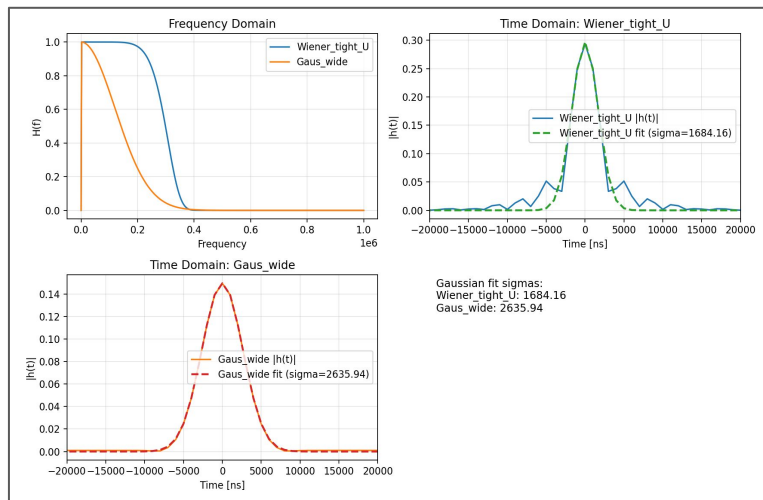
Model Training



- Training time: 46 min for 5 epoch
 - slower than PD-HD (~30 min for 50 epoch)
 - Additional pre-processing in training script caused the delay
- Trained on induction views (U & V)
- Converting the pth file to torch script file with torch.jit.script was successful
- Running test with PD-VD real data

Back Up

Gaus Filter and Truth Smearing Optimization

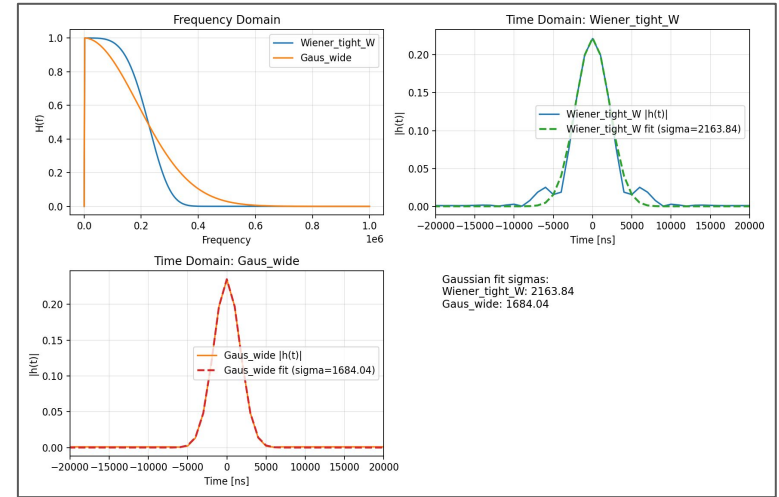
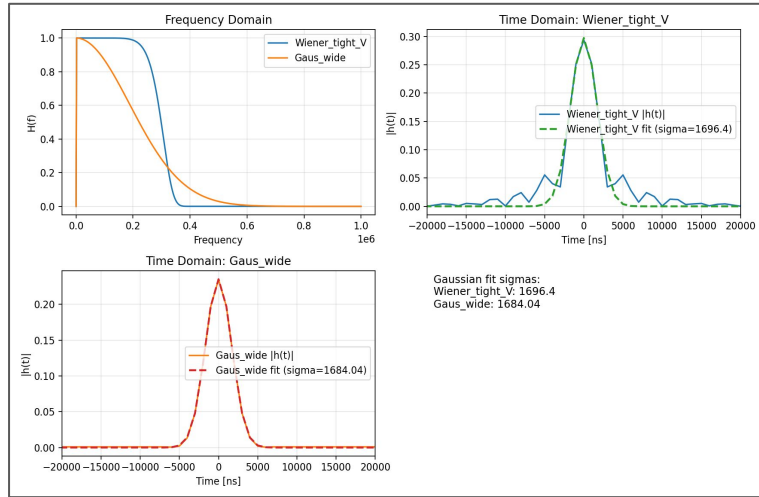


❑ PD-HD's Gaus_wide vs. PD-VD's Wiener U

❑ PD-VD's Gaus_wide vs. PD-VD's Wiener U

- Gaussian fit for PD-VD optimized Wiener filter gives a sigma of 1684.16
- The PD-HD's Gaus Filter in time domain is wider (left)
- After updating parameter for Gaus_wide, the fitted sigma is matched
- The fitted sigma for Wiener_tight is used as a parameter for truth smearing (smear_long_U)

Gaus Filter and Truth Smearing Optimization

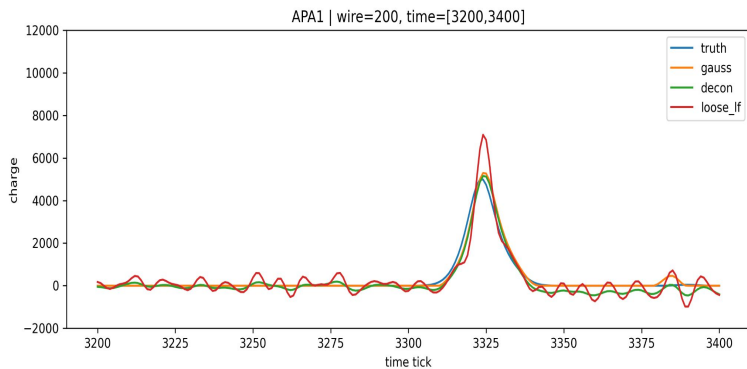


❑ PD-VD's Gaus_wide vs. PD-VD's Wiener V

❑ PD-VD's Gaus_wide vs. PD-VD's Wiener W

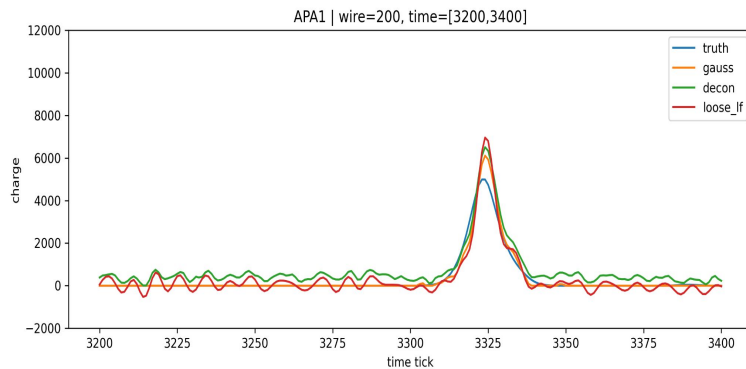
- The single Gaus_wide filter is used for all three wire planes in the current setting
- For U and V planes, the optimized Gaus_wide filter is well aligned in terms of the sigma
- For W plane, Wiener_tight_W is wider than Gauss

1D Waveform Study



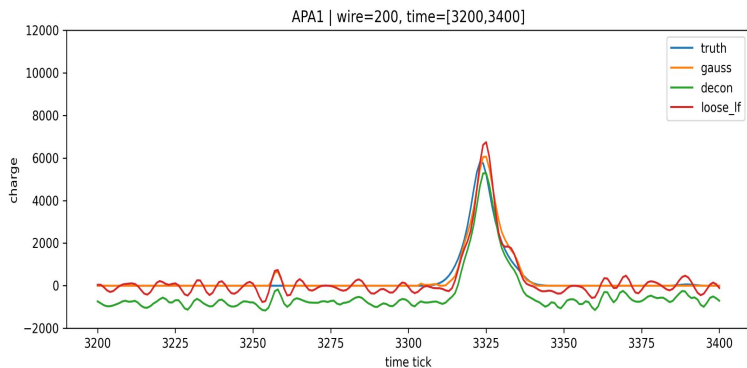
PD-HD's Gaus_wide and truth smearing

update Gaus_wide



PD-HD's truth smearing and PD-VD's Gaus_wide

update smear_long

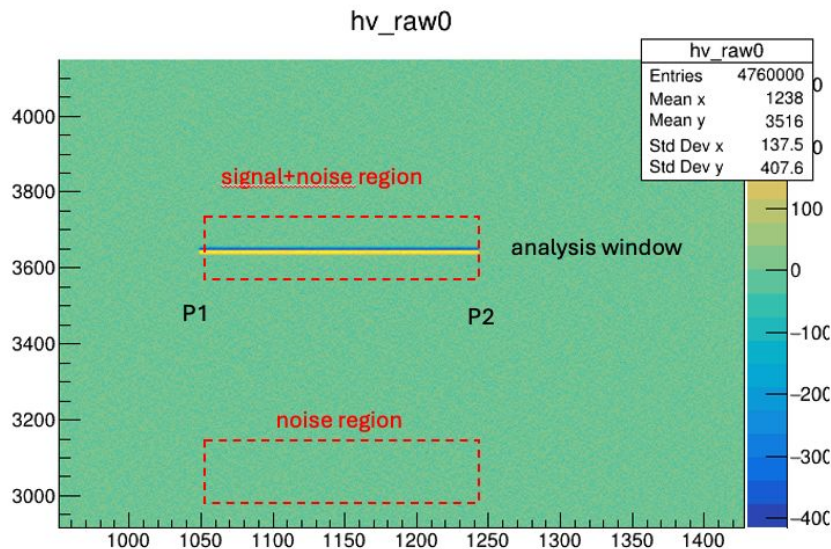


PD-VD's Gaus_wide and truth smearing

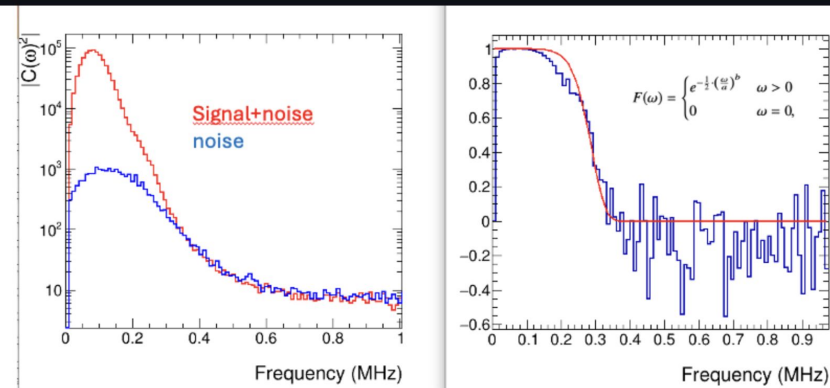
- The PD-VD's Wiener filter is default for this test

Software Filter Optimization

- There is a tool to get the optimized sigma and power for the HfFilter
 - https://github.com/WireCell/wire-cell-toolkit/blob/feature/xn_WF_fitter_script/root/test/calcFilter.C
 - https://github.com/WireCell/wire-cell-toolkit/blob/feature/xn_WF_fitter_script/root/docs/calcFilter.org
- Isochronous track sample is needed



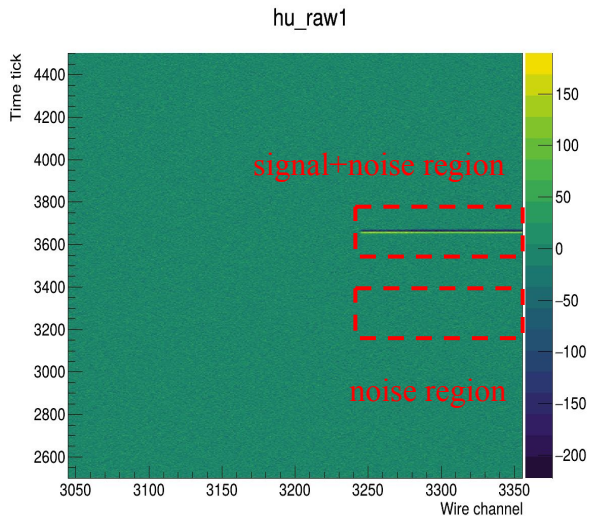
1. c1: Power spectral density comparison
 - Red: Signal + noise power density
 - Blue: Noise-only power density
2. c2: Wiener filter response
 - Normalized to maximum value of 1
 - Fitted with function: $\exp(-0.5*(x/p\theta)^{p1})$
 - Fit parameters printed to console



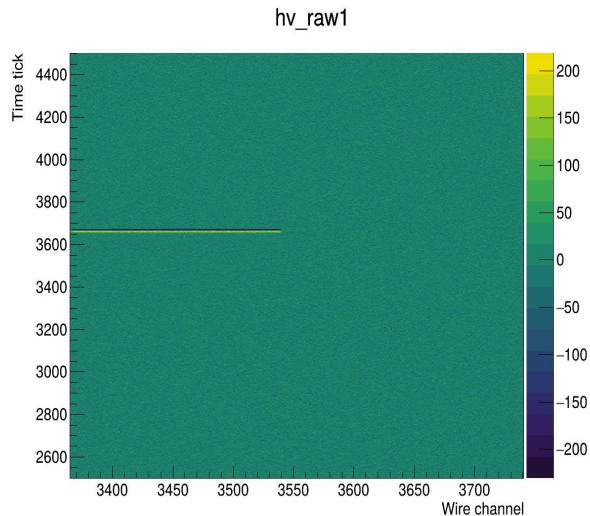
EXT NO.	PARAMETER NAME	VALUE	ERROR	STEP SIZE	FIRST DERIVATIVE
1	pθ	2.69885e-01	3.74802e-02	2.98269e-05	3.55623e-02
2	p1	8.34938e+00	6.86118e+00	5.42952e-03	-2.51530e-04

Software Filter Study

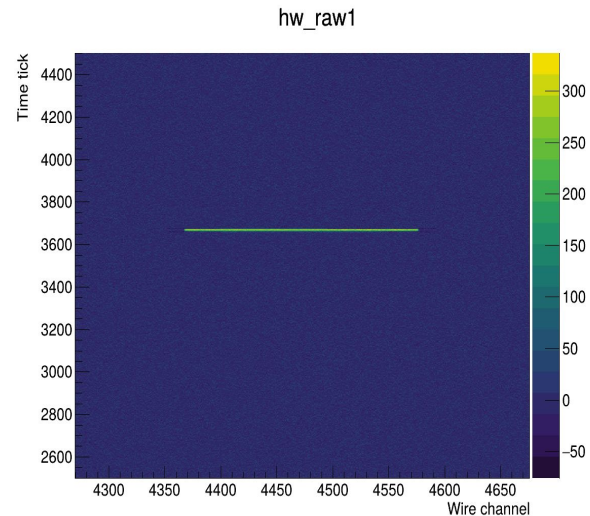
- Confirming the method to get parameters is correct



```
p0 = 0.168801
p1 = 3.11267
```



```
p0 = 0.18197
p1 = 5.35323
```



```
p0 = 0.155818
p1 = 2.73558
```

- The fitted parameters are different from PD-HD HfFilter we used

HDF5 Compression Test

```
=== Analysis by Dataset Name (Grouped) ===
```

Name	Cnt	Type	Shape	Logical	Storage	Ratio	Comp
frame_decon_charge0	1	float32	(1536, 6400)	37.50 MB	37.50 MB	1.0	None
frame_gauss0	1	float32	(1536, 6400)	37.50 MB	37.50 MB	1.0	None
frame_loose_lf0	1	float32	(1536, 6400)	37.50 MB	37.50 MB	1.0	None
frame_mp2_roi0	1	float32	(1536, 6400)	37.50 MB	37.50 MB	1.0	None
frame_mp3_roi0	1	float32	(1536, 6400)	37.50 MB	37.50 MB	1.0	None
channels_decon_charge0	1	int32	(1536,)	6.00 KB	6.00 KB	1.0	None
channels_gauss0	1	int32	(1536,)	6.00 KB	6.00 KB	1.0	None
channels_loose_lf0	1	int32	(1536,)	6.00 KB	6.00 KB	1.0	None
channels_mp2_roi0	1	int32	(1536,)	6.00 KB	6.00 KB	1.0	None
channels_mp3_roi0	1	int32	(1536,)	6.00 KB	6.00 KB	1.0	None
tickinfo_decon_charge0	1	float64	(3,)	24 B	24 B	1.0	None
tickinfo_gauss0	1	float64	(3,)	24 B	24 B	1.0	None
tickinfo_loose_lf0	1	float64	(3,)	24 B	24 B	1.0	None
tickinfo_mp2_roi0	1	float64	(3,)	24 B	24 B	1.0	None
tickinfo_mp3_roi0	1	float64	(3,)	24 B	24 B	1.0	None
TOTAL	15	-	-	187.53 MB	187.53 MB	1.0	-

❑ gzip = 0

```
=== Analysis by Dataset Name (Grouped) ===
```

Name	Cnt	Type	Shape	Logical	Storage	Ratio	Comp
frame_loose_lf0	1	float32	(1536, 6400)	37.50 MB	34.84 MB	1.1	gzip
frame_decon_charge0	1	float32	(1536, 6400)	37.50 MB	34.77 MB	1.1	gzip
frame_gauss0	1	float32	(1536, 6400)	37.50 MB	462.11 KB	83.1	gzip
frame_mp2_roi0	1	float32	(1536, 6400)	37.50 MB	192.14 KB	199.9	gzip
frame_mp3_roi0	1	float32	(1536, 6400)	37.50 MB	175.06 KB	219.4	gzip
channels_decon_charge0	1	int32	(1536,)	6.00 KB	6.00 KB	1.0	None
channels_gauss0	1	int32	(1536,)	6.00 KB	6.00 KB	1.0	None
channels_loose_lf0	1	int32	(1536,)	6.00 KB	6.00 KB	1.0	None
channels_mp2_roi0	1	int32	(1536,)	6.00 KB	6.00 KB	1.0	None
channels_mp3_roi0	1	int32	(1536,)	6.00 KB	6.00 KB	1.0	None
tickinfo_decon_charge0	1	float64	(3,)	24 B	24 B	1.0	None
tickinfo_gauss0	1	float64	(3,)	24 B	24 B	1.0	None
tickinfo_loose_lf0	1	float64	(3,)	24 B	24 B	1.0	None
tickinfo_mp2_roi0	1	float64	(3,)	24 B	24 B	1.0	None
tickinfo_mp3_roi0	1	float64	(3,)	24 B	24 B	1.0	None
TOTAL	15	-	-	187.53 MB	70.45 MB	2.7	-

❑ gzip = 2

```
=== Analysis by Dataset Name (Grouped) ===
```

Name	Cnt	Type	Shape	Logical	Storage	Ratio	Comp
frame_loose_lf0	1	float32	(1536, 6400)	37.50 MB	34.74 MB	1.1	gzip
frame_decon_charge0	1	float32	(1536, 6400)	37.50 MB	34.67 MB	1.1	gzip
frame_gauss0	1	float32	(1536, 6400)	37.50 MB	309.53 KB	124.1	gzip
frame_mp2_roi0	1	float32	(1536, 6400)	37.50 MB	50.13 KB	766.1	gzip
frame_mp3_roi0	1	float32	(1536, 6400)	37.50 MB	42.13 KB	911.4	gzip
channels_decon_charge0	1	int32	(1536,)	6.00 KB	6.00 KB	1.0	None
channels_gauss0	1	int32	(1536,)	6.00 KB	6.00 KB	1.0	None
channels_loose_lf0	1	int32	(1536,)	6.00 KB	6.00 KB	1.0	None
channels_mp2_roi0	1	int32	(1536,)	6.00 KB	6.00 KB	1.0	None
channels_mp3_roi0	1	int32	(1536,)	6.00 KB	6.00 KB	1.0	None
tickinfo_decon_charge0	1	float64	(3,)	24 B	24 B	1.0	None
tickinfo_gauss0	1	float64	(3,)	24 B	24 B	1.0	None
tickinfo_loose_lf0	1	float64	(3,)	24 B	24 B	1.0	None
tickinfo_mp2_roi0	1	float64	(3,)	24 B	24 B	1.0	None
tickinfo_mp3_roi0	1	float64	(3,)	24 B	24 B	1.0	None
TOTAL	15	-	-	187.53 MB	69.83 MB	2.7	-

❑ gzip = 6

- gzip option works now (chunk should not be zero)
- File sizes per event (1 CRU)
 - gzip = 0, reco 188 MB, truth 38 MB
 - gzip = 2, reco 70 MB, truth 744 KB
 - gzip = 6, reco 70 MB, truth 600 KB
- gzip=2 gives enough suppression rate
- For 1000 events, estimated required storage ~ 70.7 GB