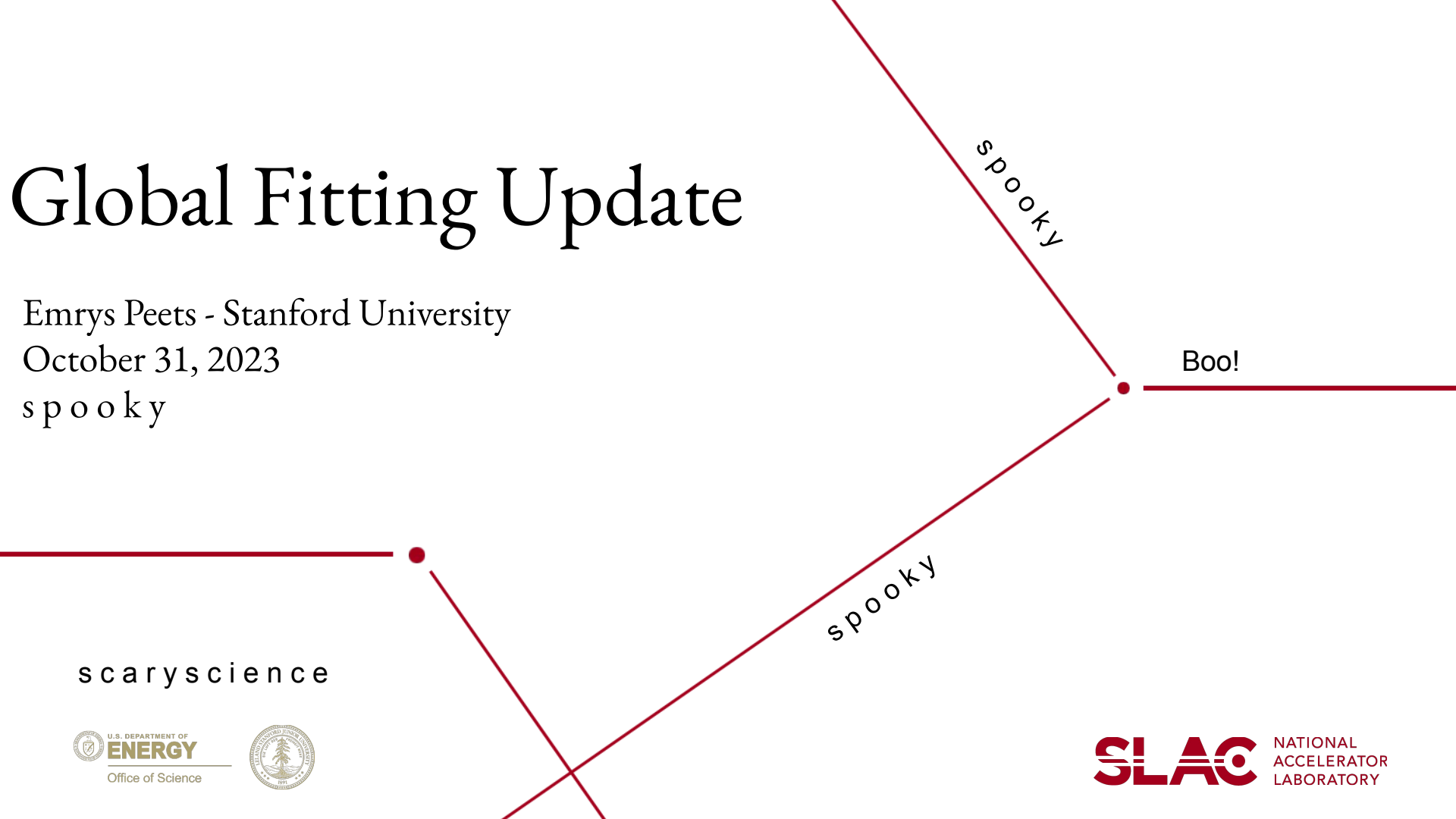
Global Fitting Update

Emrys Peets - Stanford University
October 31, 2023
s p o o k y



s c a r y s c i e n c e

U.S. DEPARTMENT OF ENERGY
Office of Science

SLAC NATIONAL ACCELERATOR LABORATORY

Quick Update

- 8 hour study over the range [50,198] MeV conducted in last update was bugged
  - parameter information was lost in the storing process for best fits
  - fixed this bug
- Conducted a 24 hour study over the range [50, 198] MeV
  - 10 functions with values > 1e-2
  - modified previous 1d histo plotter
- Conducted round two fitting procedure
  - Took parameters and functions from ^
  - added functions that had performed well in previous round two
- New Display Tool
  - that plots each set of fits for a given function

Last Update: 8 Hour Study - [50, 198] MeV

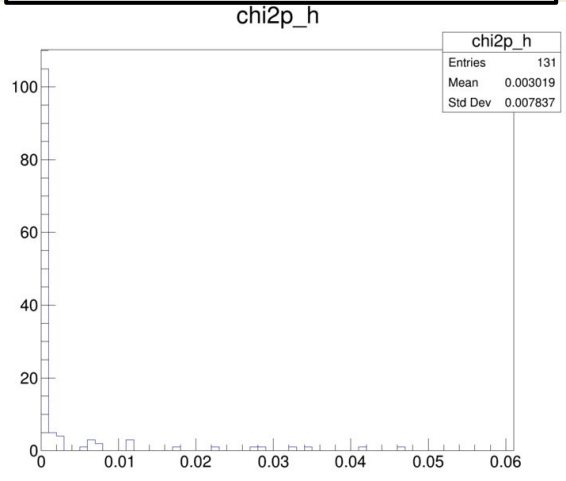
- 11 functions with pvalue >10^-2
- mean pvalue = .003

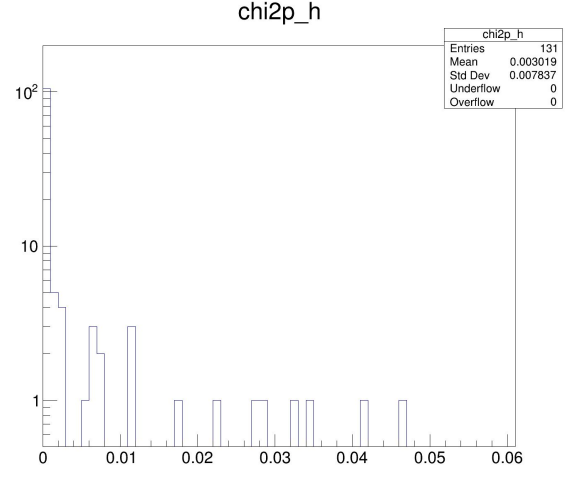
Top 20 performing functions

function_table_50MeV_198MeV

Function Name	Number of Fits	Chi2/Ndf	PValue
las3_plus_las6	1702.0	1.0443	0.0460
ua23_nolin_plus_las1	1053.0	1.0457	0.0412
las2_plus_las6	10053.0	1.0480	0.0342
ua23_nolin_plus_las3	126.0	1.0487	0.0323
dj1_plus_dj1	16601.0	1.0501	0.0286
ua23_nolin_plus_las2	8124.0	1.0506	0.0275
las1_plus_ua23	2391.0	1.0527	0.0228
dj1_plus_las2	1110.0	1.0557	0.0174
dj1_plus_ua22	2704.0	1.0599	0.0117
las1_plus_ua21	2999.0	1.0600	0.0117
las3_plus_las3	953.0	1.0605	0.0111
dj1_plus_cms1	665.0	1.0640	0.0078
las1_plus_las5	4455.0	1.0645	0.0075
dj1_plus_las6	292.0	1.0654	0.0068
las1_plus_las1	2905.0	1.0659	0.0064
las1_plus_las6	1623.0	1.0661	0.0063
ua23_er_er_10_2	8557.0	1.0680	0.0052
ua23_er_er_8_4	8728.0	1.0734	0.0029
ua23_er_er_11_2	36878.0	1.0755	0.0023
ua23_er_er_1	3562.0	1.0759	0.0022

1D Pvalue Distribution





24 Hour Study on range [50, 198] MeV

Variance Technique:

Select starting parameters from gaussian with width that is 1% of starting parameter value and change width by an additional 0.5% of mean for every 100 fits without finding better pvalue.

- Improved on pvalue for many functions found in the 8 hour study
- Stored all parameters for each of the best fits to be used in “Round 2 Fitting”
- Fixed 8 hr study bug and rewrote display scripts to reflect proper data format

24 Hour Study on range [50, 198] MeV

- 10 functions with pval > 1e-2
- Unclear if change in variance helped. 1 less function, but better top 10 fits.

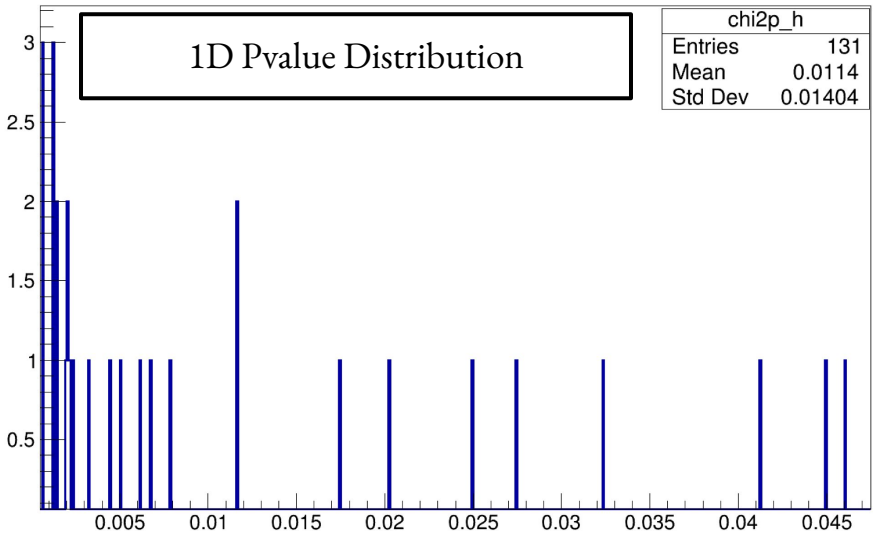
new display script modified from Cam's initial histo plotter

Top 20 performing functions

24hr_function_table_50MeV_198MeV

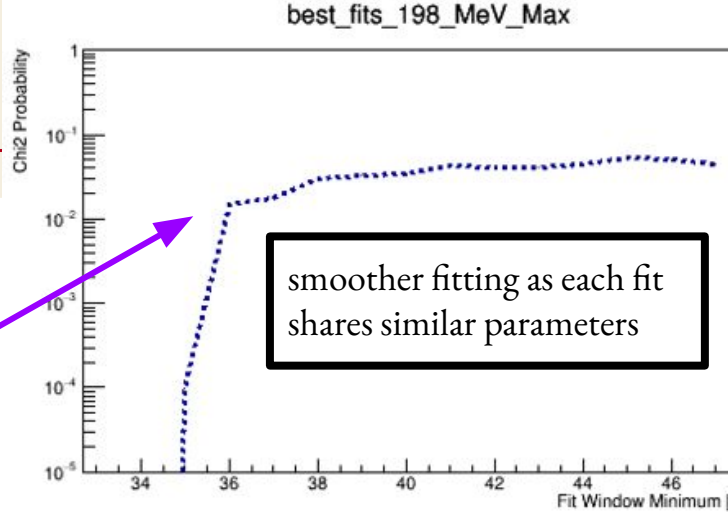
Function Name	Number of Fits	Chi2/Ndf	PValue
las3_plus_las6	1217	1.044	4.606E-02
las3_plus_las3	22667	1.045	4.498E-02
ua23_nolin_plus_las1	2531	1.046	4.128E-02
ua23_nolin_plus_las3	126	1.049	3.231E-02
ua23_nolin_plus_las2	2986	1.051	2.747E-02
las2_plus_las6	1148	1.052	2.496E-02
[illegible]	52867	1.054	2.025E-02
dj1_plus_las2	2800	1.056	1.742E-02
las1_plus_ua21	1965	1.060	1.170E-02
dj1_plus_ua22	2894	1.060	1.164E-02
dj1_plus_cms1	589	1.064	7.823E-03
dj1_plus_las6	3729	1.065	6.785E-03
las2_plus_las3	19109	1.066	6.142E-03
ua23_er_er_10_2	28827	1.068	5.089E-03
ua23_er_er_11	20320	1.070	4.450E-03
ua23_er_er_4	47251	1.072	3.247E-03
ua23_er_er_8_4	24464	1.075	2.372E-03
ua23_er_er_1	12416	1.076	2.168E-03
ua23_er_er_3	1071	1.076	2.084E-03
las2_plus_las5	12465	1.076	2.025E-03

chi2p_h

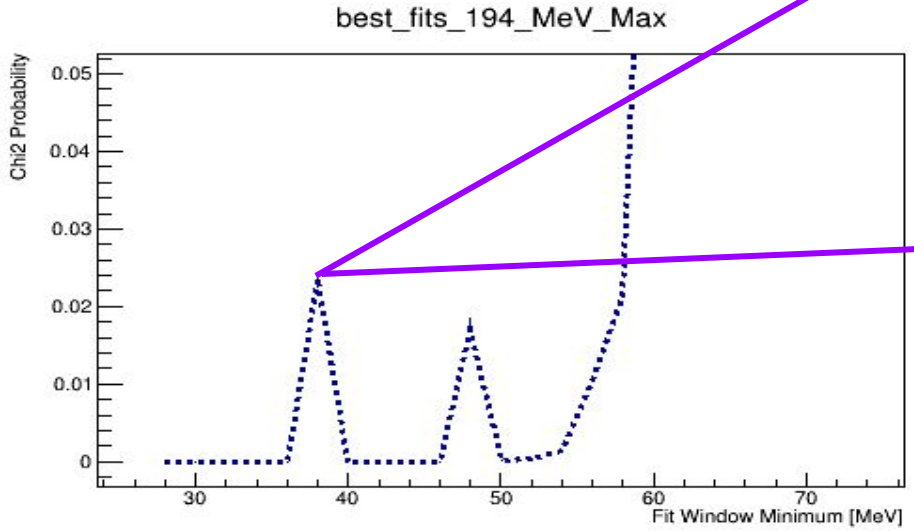


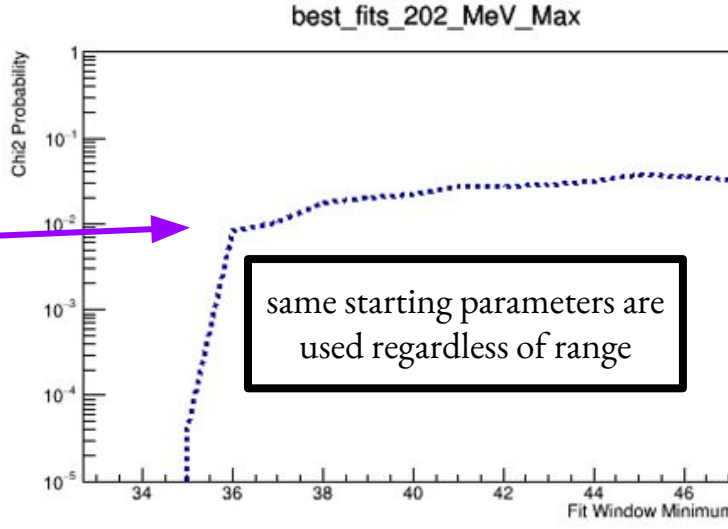
Round 2 Fitting Recap

Tested method of using parameters found for las3_plus_l as the exact seeds used in fitting each window range.



smoother fitting as each fit shares similar parameters



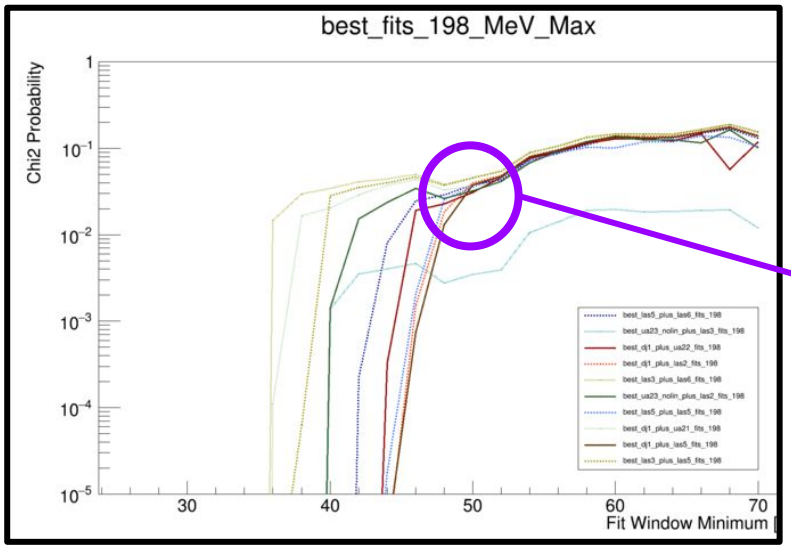


same starting parameters are used regardless of range

Round 2 Fitting Study Procedure

Fit each function to the IMD over multiple ranges of WinMin/WinMax using stored parameters as a starting place.

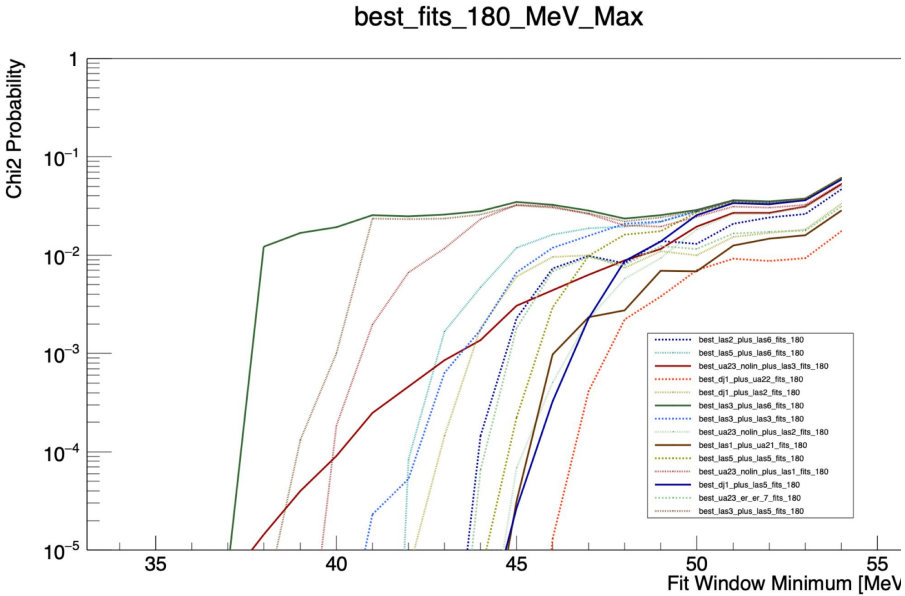
- Did not use 8 hr study results due to strings being cut off in previous storing process.
- WinMin Range: [35, 54] MeV 1 MeV step size
- WinMax Range: [180, 220] MeV 10 MeV step size

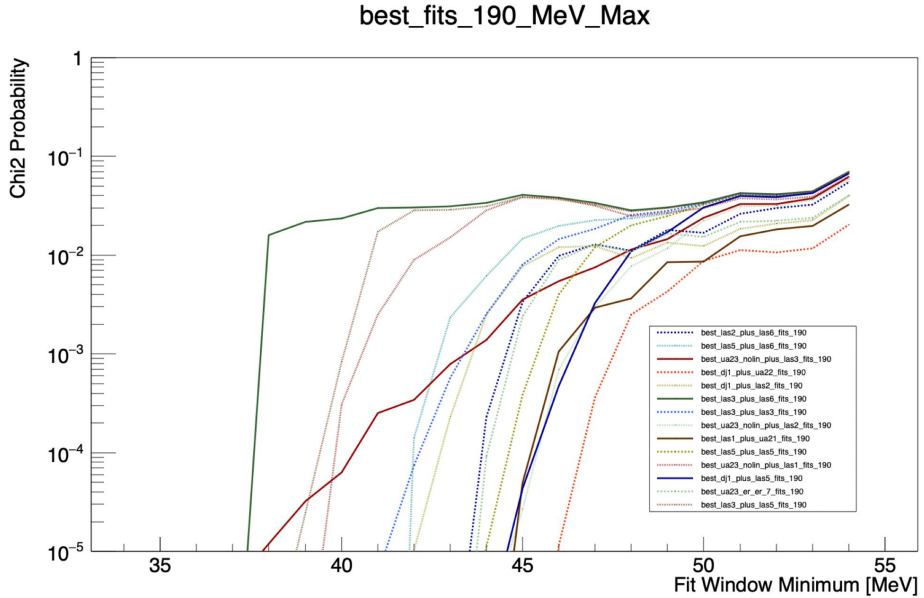


Previous R2 Study also had selection of well performing functions over the range [50, 198] MeV. Used these parameters if function's fit was better than performance during 24 hr study.

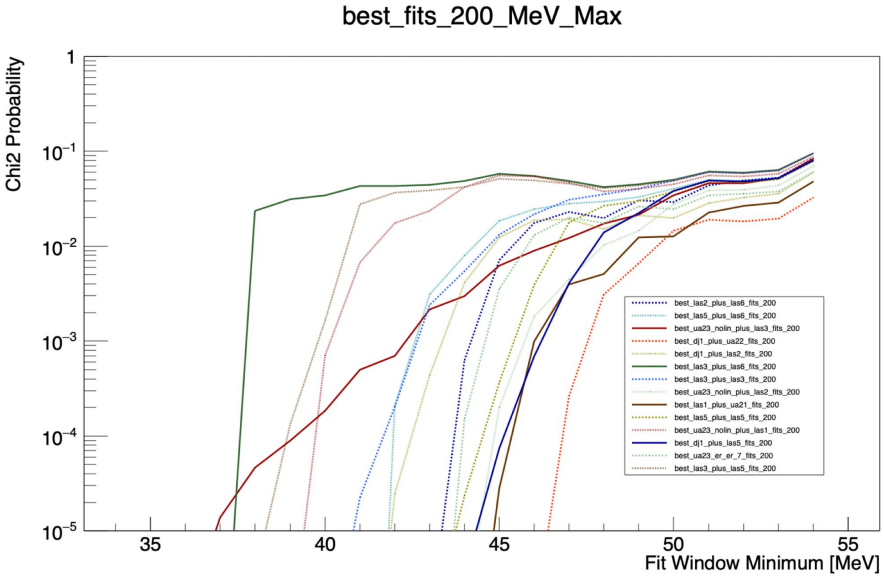
Round 2 Fitting Results

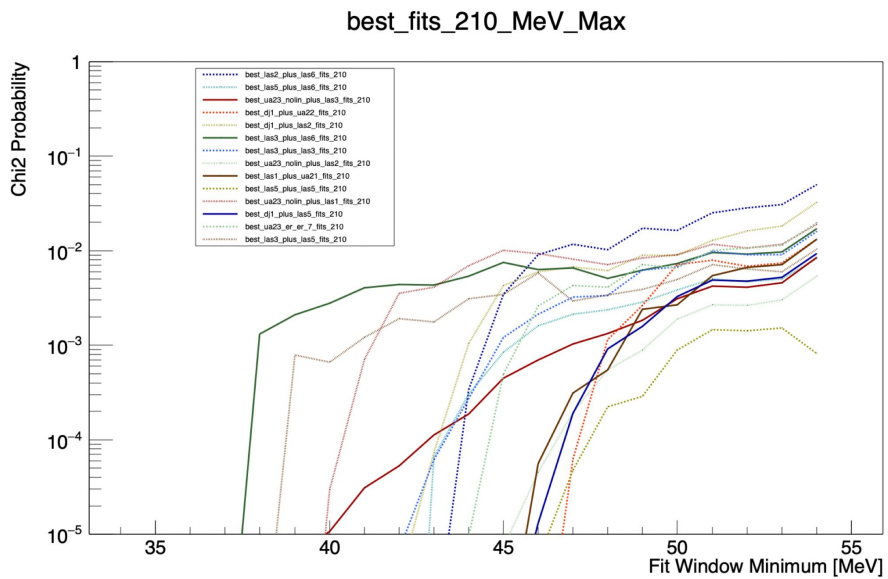
14 Total Functions Found



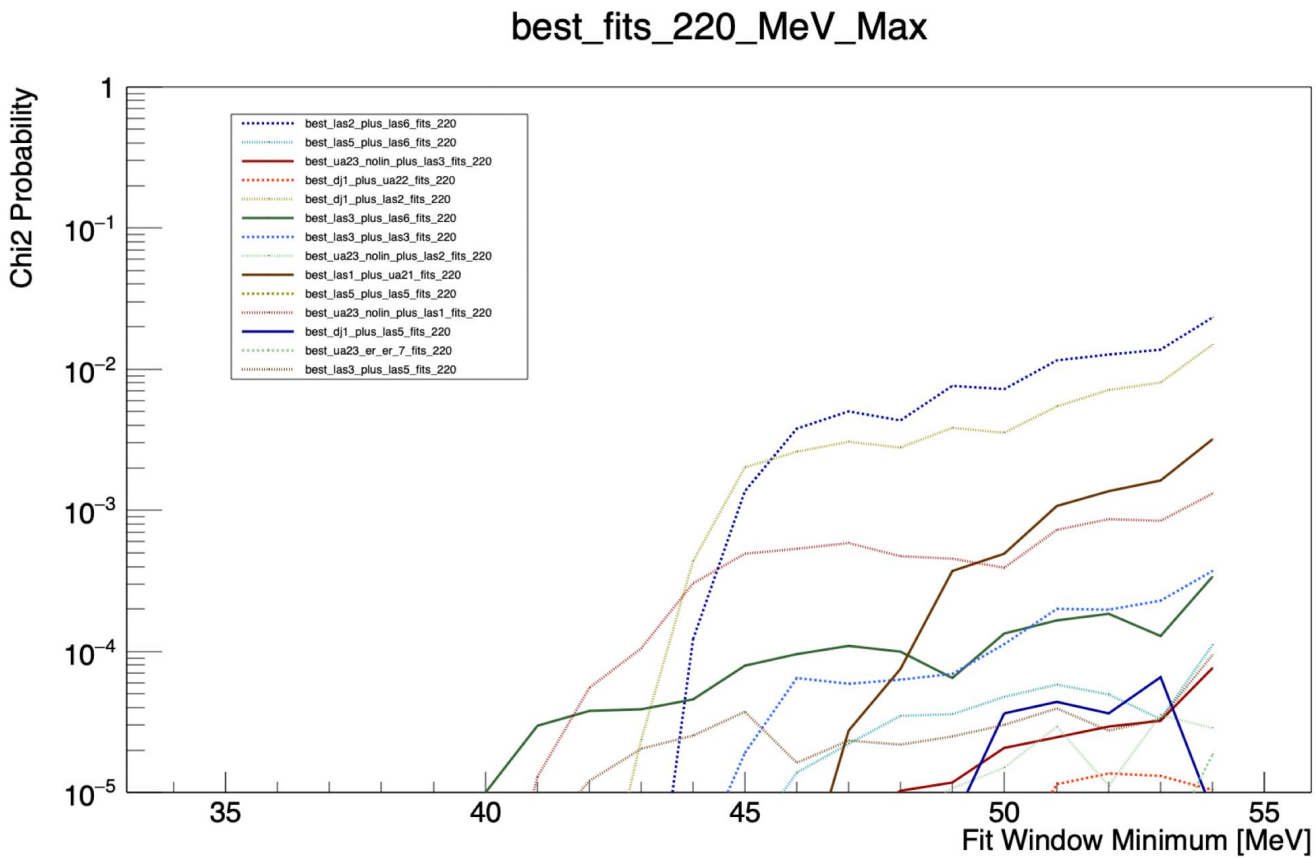


R2 Fitting Results (continued)





R2 Fitting Results (continued)

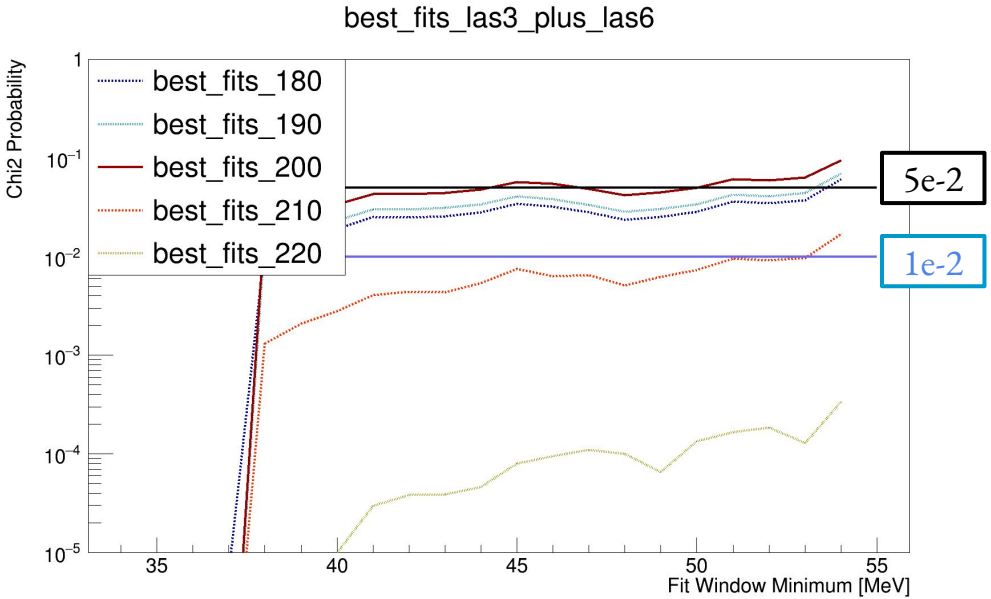


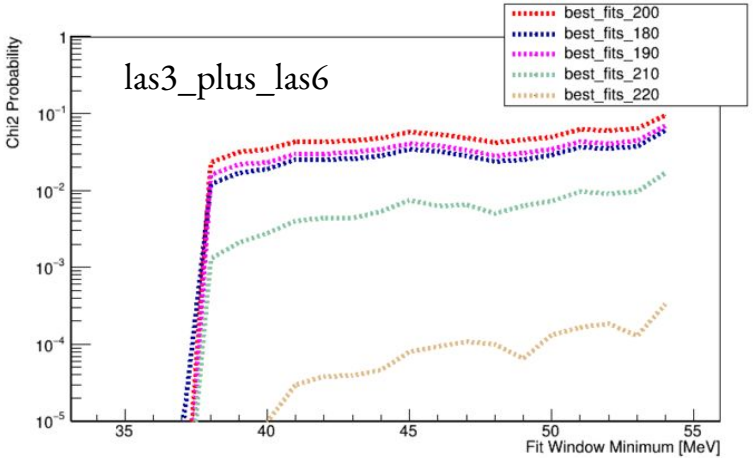
New Display Tool

Modified previous tool **merge_fits.py** to make **func_merge_fits.py**

- takes each winmax fixed plot for each function and plots them over the same range of win mins.

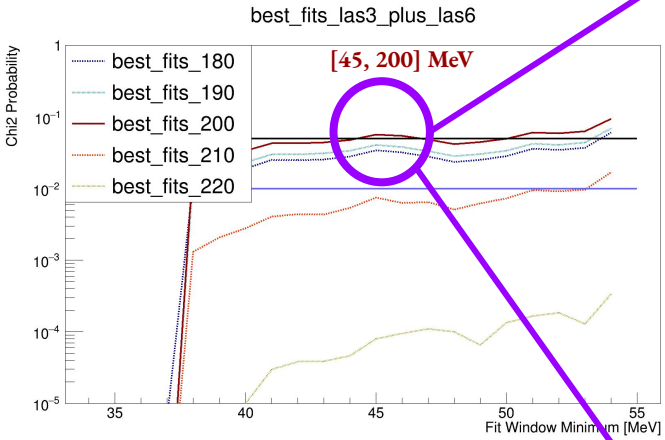
Terminal Input: python3 func_merge_fits.py -n 35 55 1 -x 180 230 10 -B 1

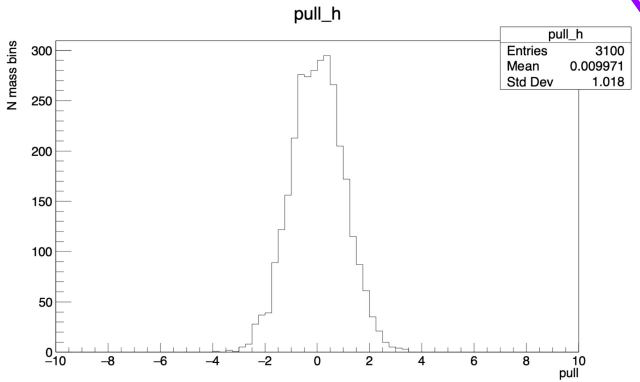


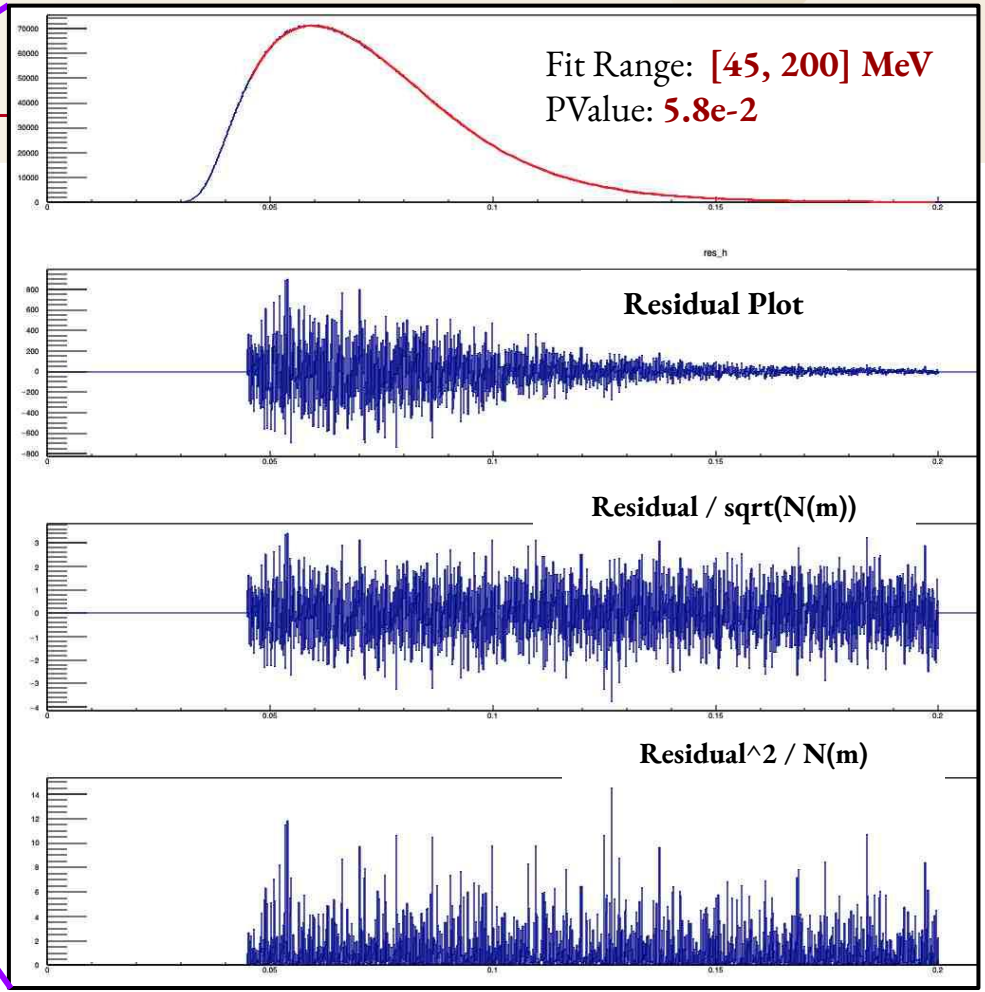


root TBrowser prototype plot (thanks Rory)

Notable Fit from Study



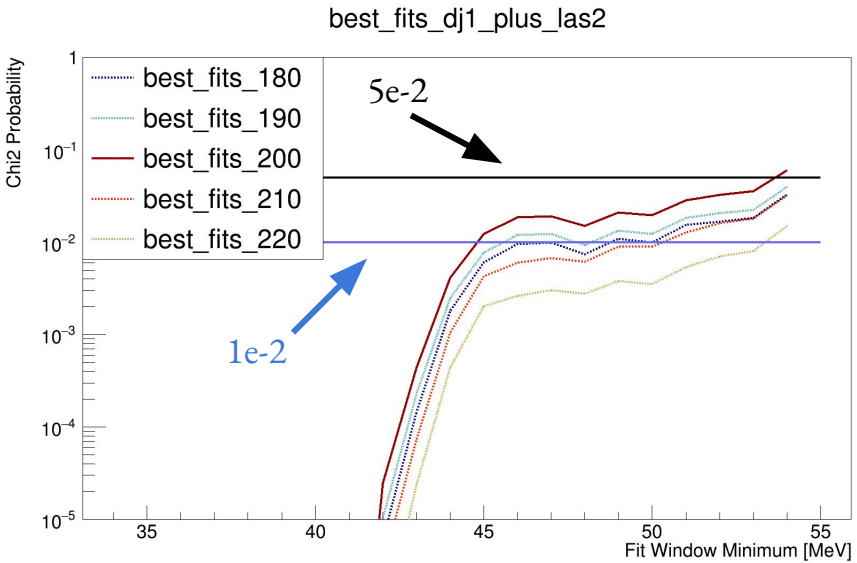


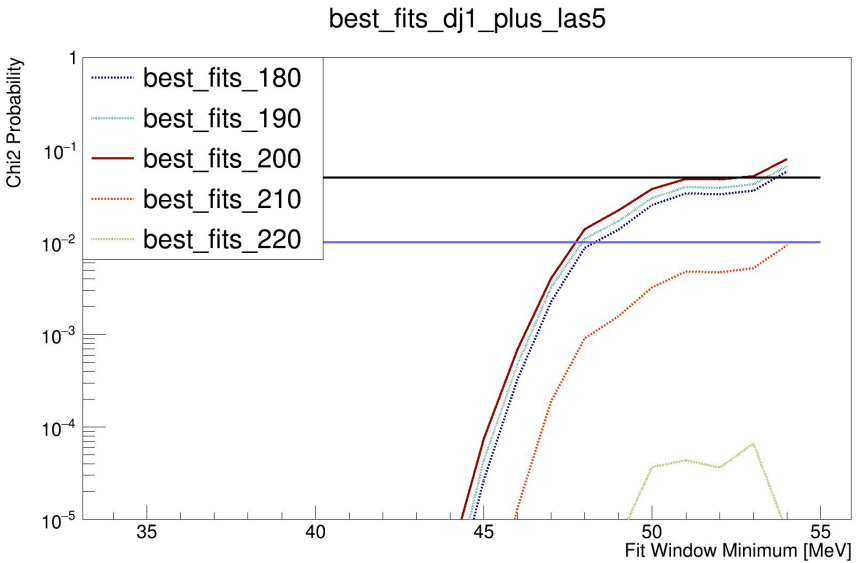


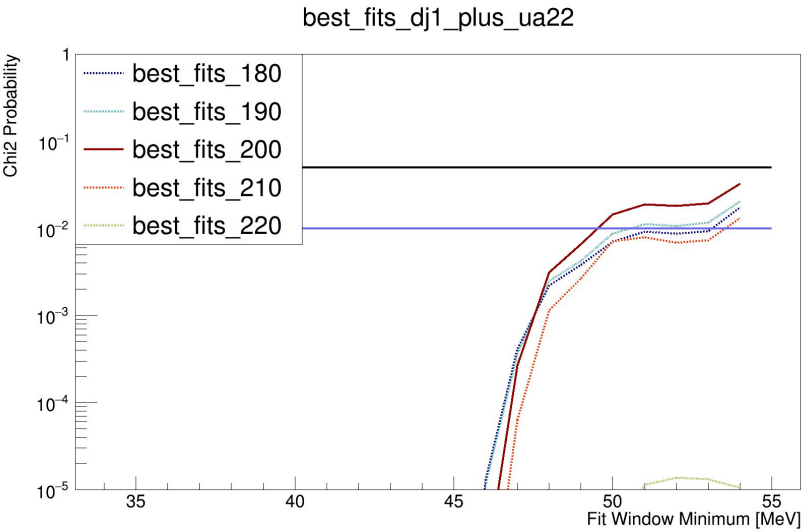
All Functions

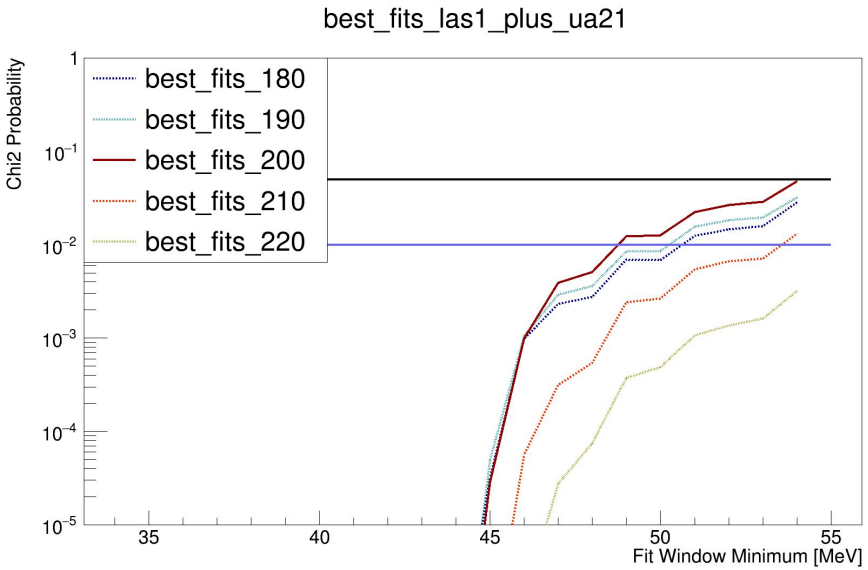
General trend for a given Window Minimum:

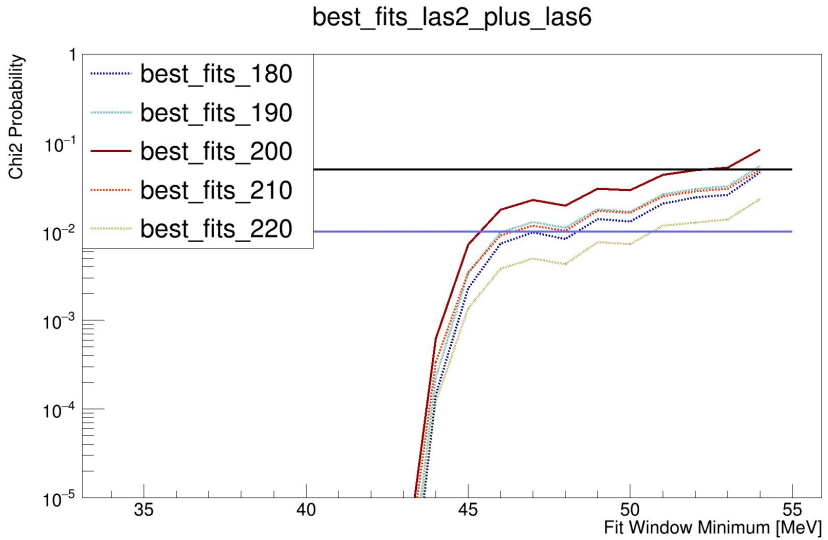
Pval(220 Max) < Pval(210 Max) < Pval(180 Max) < Pval(190 Max) < Pval(200 Max)

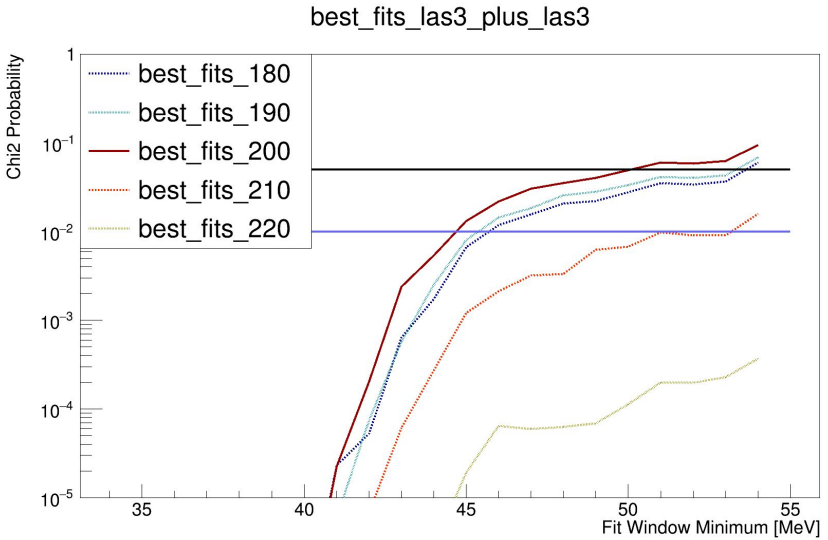


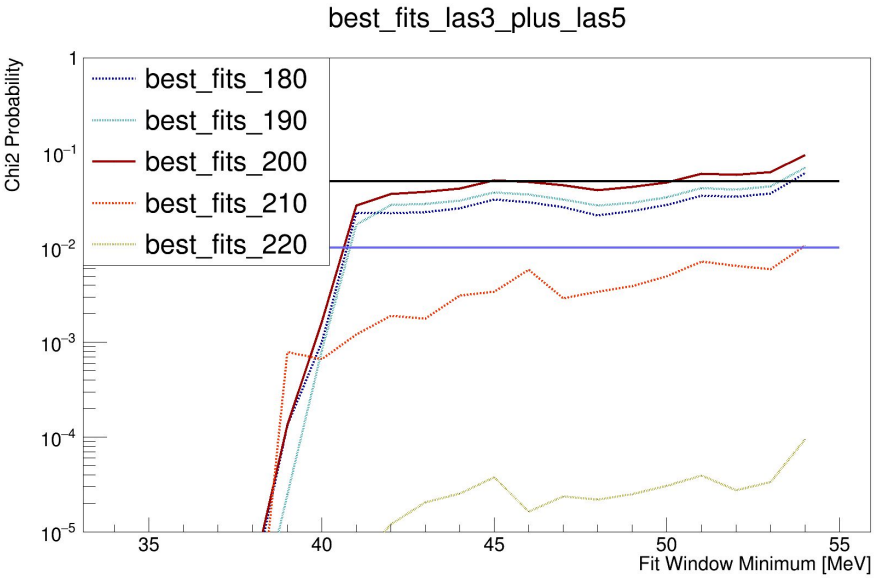


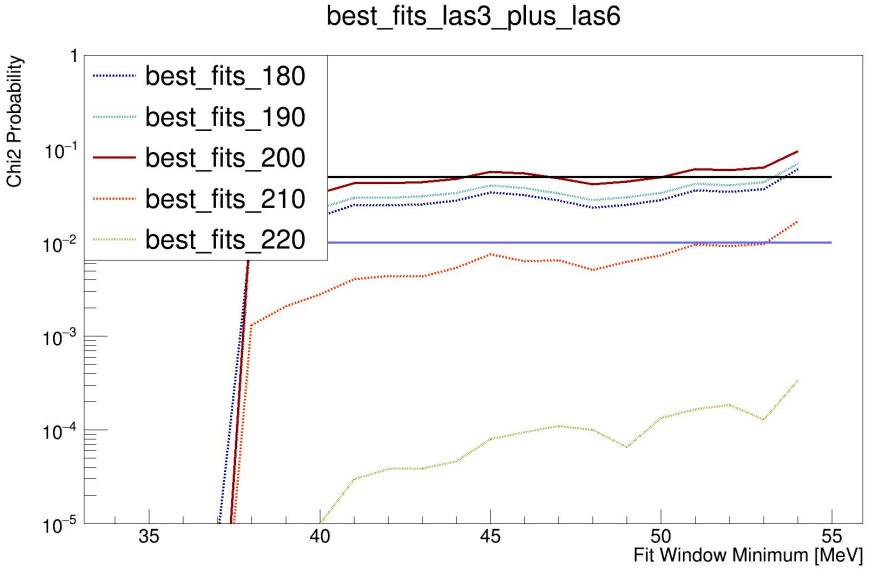

best_fits_dj1_plus_ua22
Chi2 Probability
best_fits_180
best_fits_190
best_fits_200
best_fits_210
best_fits_220
Fit Window Minimum [MeV]

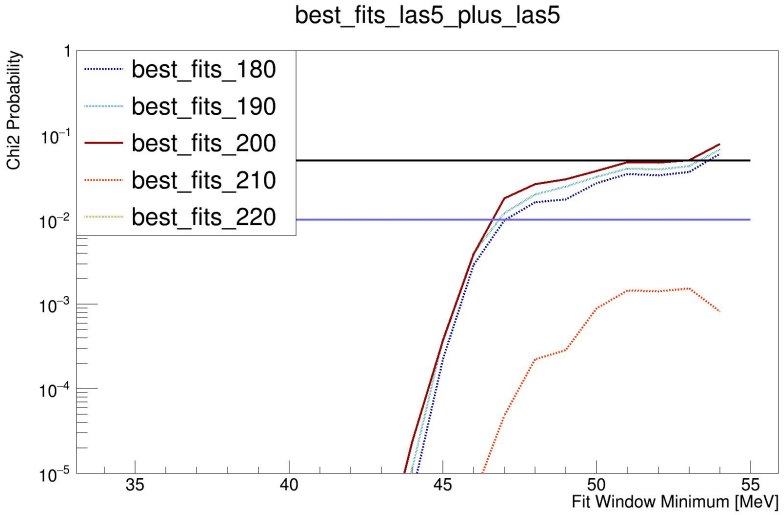

best_fits_las1_plus_ua21
Chi2 Probability
best_fits_180
best_fits_190
best_fits_200
best_fits_210
best_fits_220
Fit Window Minimum [MeV]

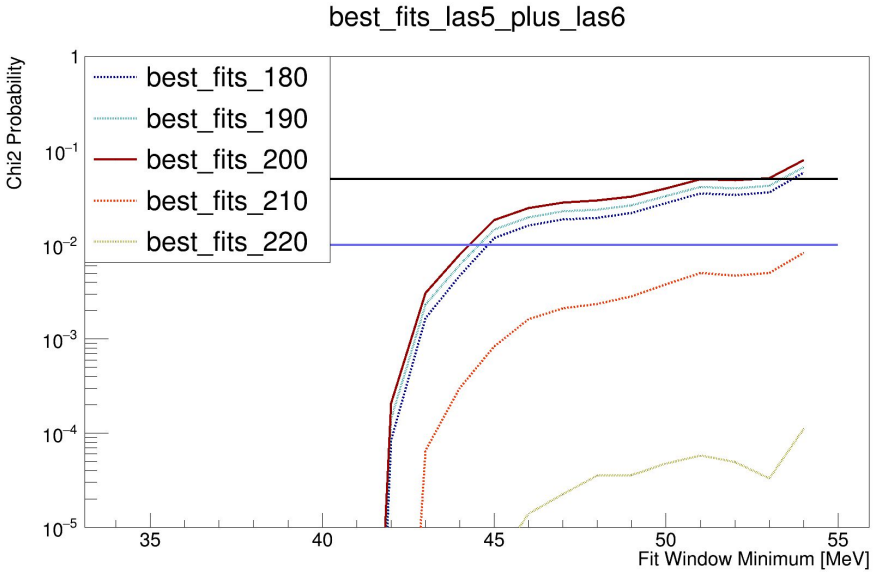

best_fits_las2_plus_las6
best_fits_180
best_fits_190
best_fits_200
best_fits_210
best_fits_220
Chi2 Probability
1
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
35
40
45
50
55
Fit Window Minimum [MeV]

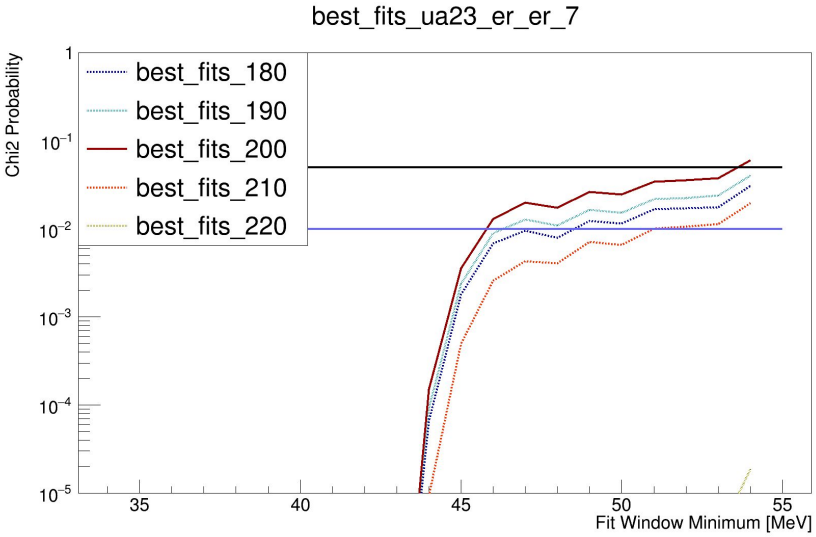

best_fits_las3_plus_las3
best_fits_180
best_fits_190
best_fits_200
best_fits_210
best_fits_220
Chi2 Probability
1
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
35
40
45
50
55
Fit Window Minimum [MeV]

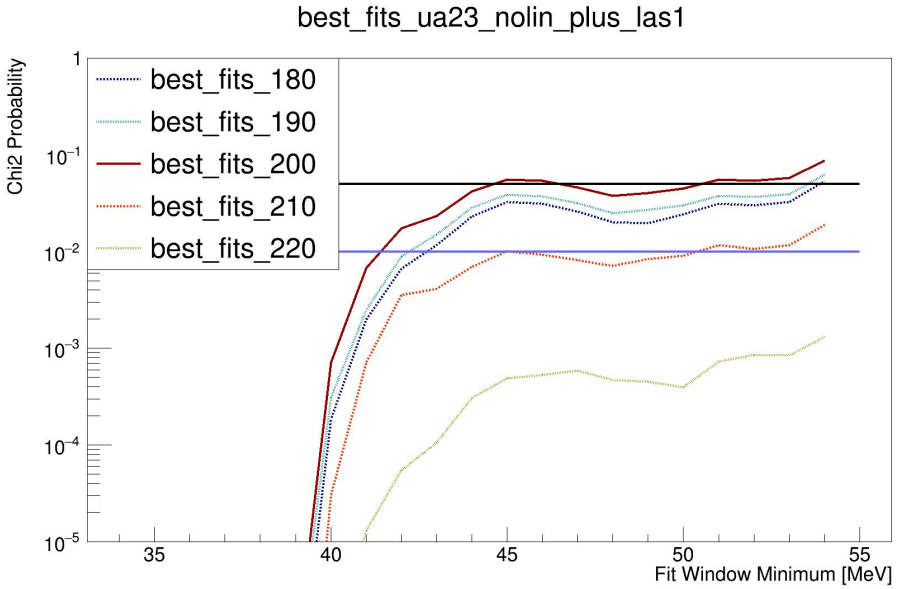

best_fits_las3_plus_las5
Chi2 Probability
1
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
best_fits_180
best_fits_190
best_fits_200
best_fits_210
best_fits_220
35
40
45
50
55
Fit Window Minimum [MeV]

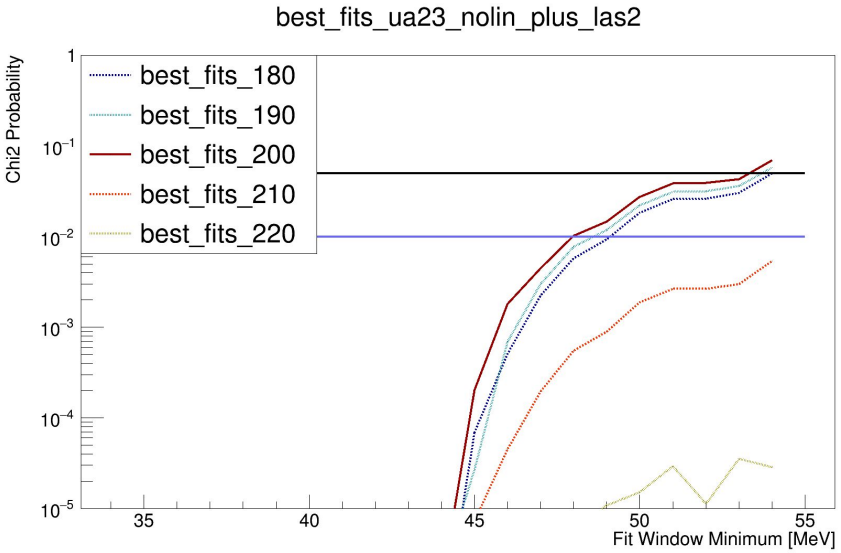

best_fits_las3_plus_las6
Chi2 Probability
1
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
best_fits_180
best_fits_190
best_fits_200
best_fits_210
best_fits_220
35
40
45
50
55
Fit Window Minimum [MeV]

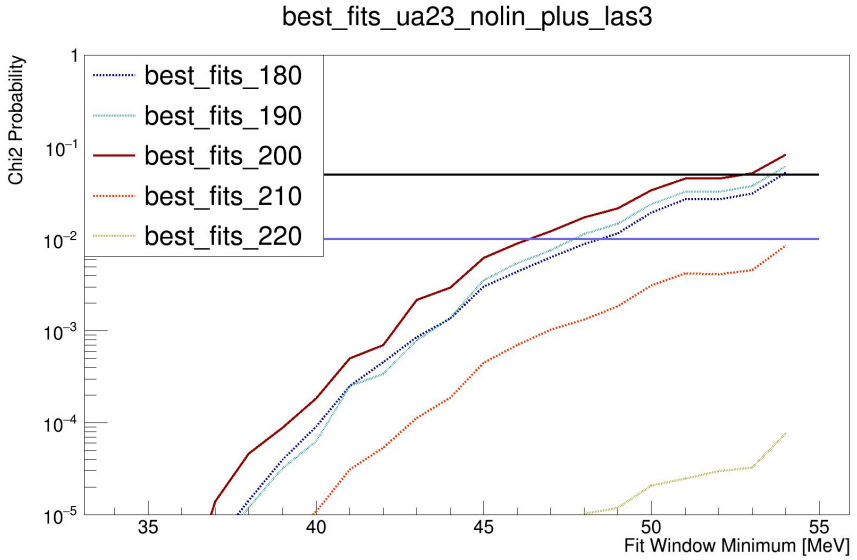

best_fits_las5_plus_las5
Chi2 Probability
1
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
best_fits_180
best_fits_190
best_fits_200
best_fits_210
best_fits_220
35
40
45
50
55
Fit Window Minimum [MeV]


best_fits_las5_plus_las6
Chi2 Probability
1
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
best_fits_180
best_fits_190
best_fits_200
best_fits_210
best_fits_220
35
40
45
50
55
Fit Window Minimum [MeV]


best_fits_ua23_er_er_7
Chi2 Probability
1
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
best_fits_180
best_fits_190
best_fits_200
best_fits_210
best_fits_220
35
40
45
50
55
Fit Window Minimum [MeV]


best_fits_ua23_nolin_plus_las1
Chi2 Probability
1
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
best_fits_180
best_fits_190
best_fits_200
best_fits_210
best_fits_220
35
40
45
50
55
Fit Window Minimum [MeV]


best_fits_ua23_nolin_plus_las2
Chi2 Probability
1
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
best_fits_180
best_fits_190
best_fits_200
best_fits_210
best_fits_220
35
40
45
50
55
Fit Window Minimum [MeV]


best_fits_ua23_nolin_plus_las3
Chi2 Probability
1
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
best_fits_180
best_fits_190
best_fits_200
best_fits_210
best_fits_220
35
40
45
50
55
Fit Window Minimum [MeV]

End of Update

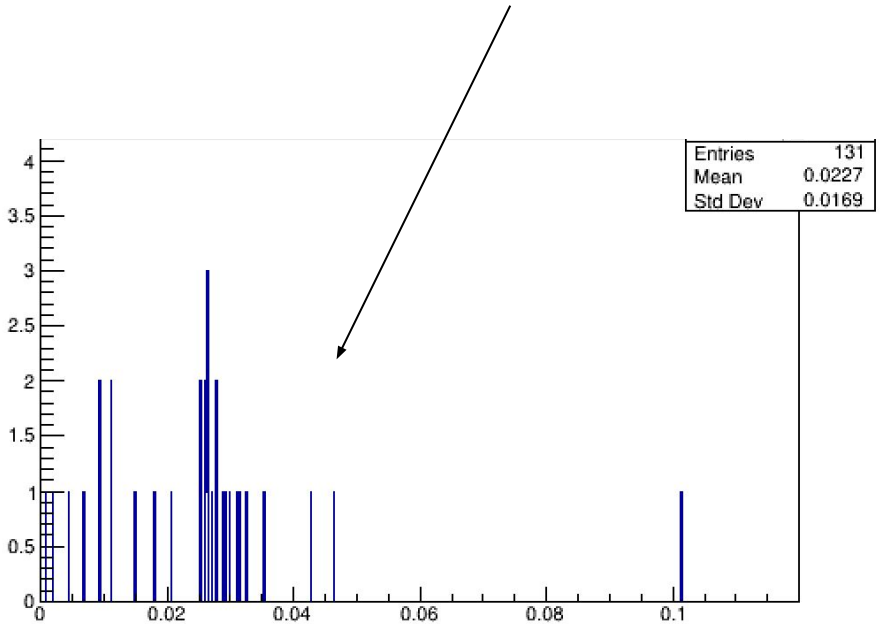
- Next step is to transition away from filtering / finding functions
  - Would like to push through single function through existing fitting infrastructure and see how a result might compare to what Matt/Cam have
  - then will generalize for a set of n functions to compare against and determine bias'

Extra Slide(s)

-Interesting parameter observation

Interesting parameter observation

shows that the error function's first parameter has a best region



accidentally ran the 1D distribution script on the first fitting parameter, this displays the distribution of this first parameter value for all functions