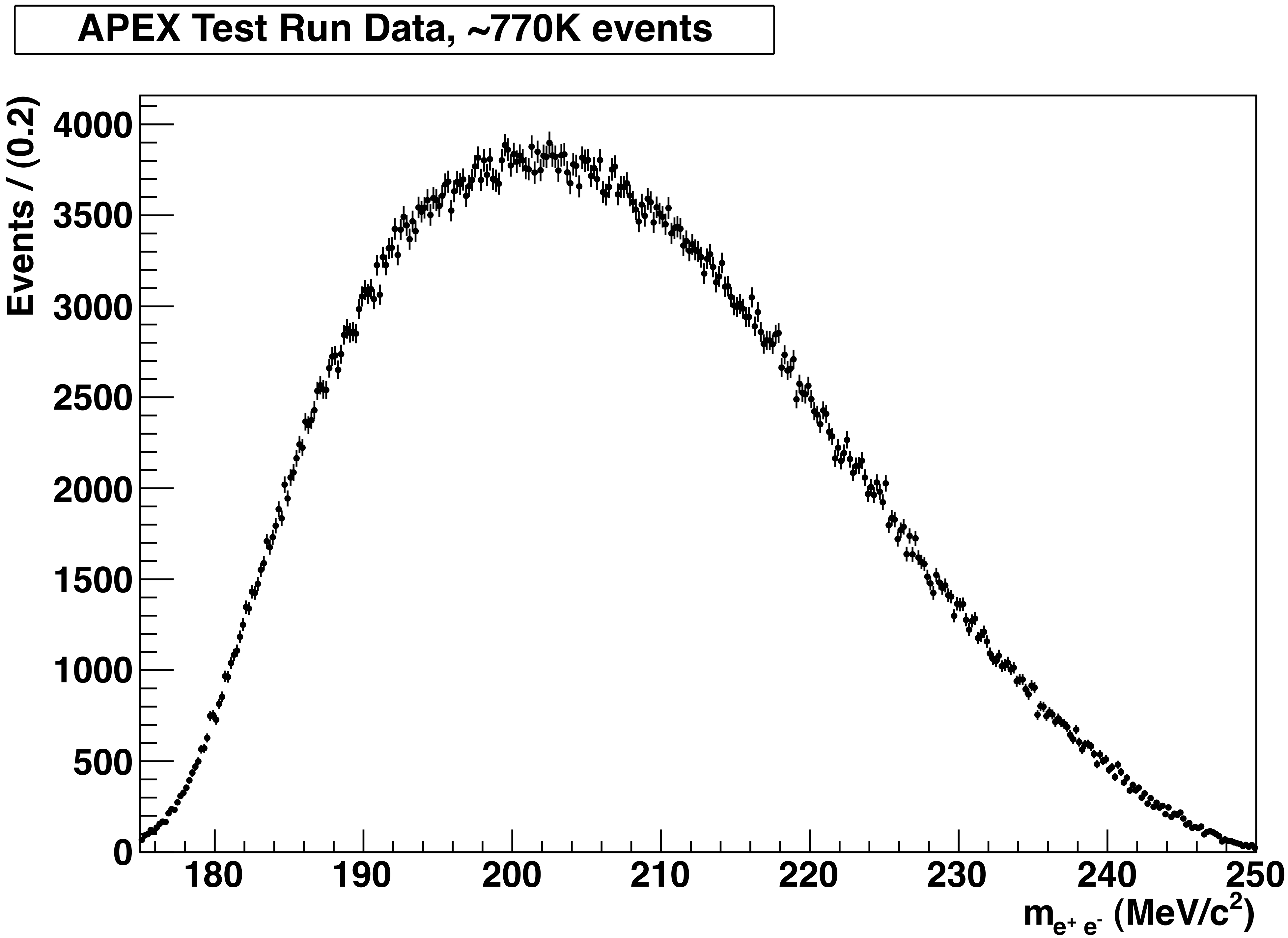
APEX Resonance Search



James Beacham
New York University

Outline

1) APEX resonance search

- Objectives of resonance search
- Procedure
- Optimization for binned analysis
- Characteristics of limit plots
- Test run unblinding and final results
- Plans for full run analysis

2) In progress: Update of reach calculation for full run using new acceptance

The Objectives

- Identify the small, narrow A' resonance on a smooth, high-statistics background spectrum
- Determine an upper limit on the number of signal events, S, consistent with background
- Upper limit on S ⟶ upper limit on coupling

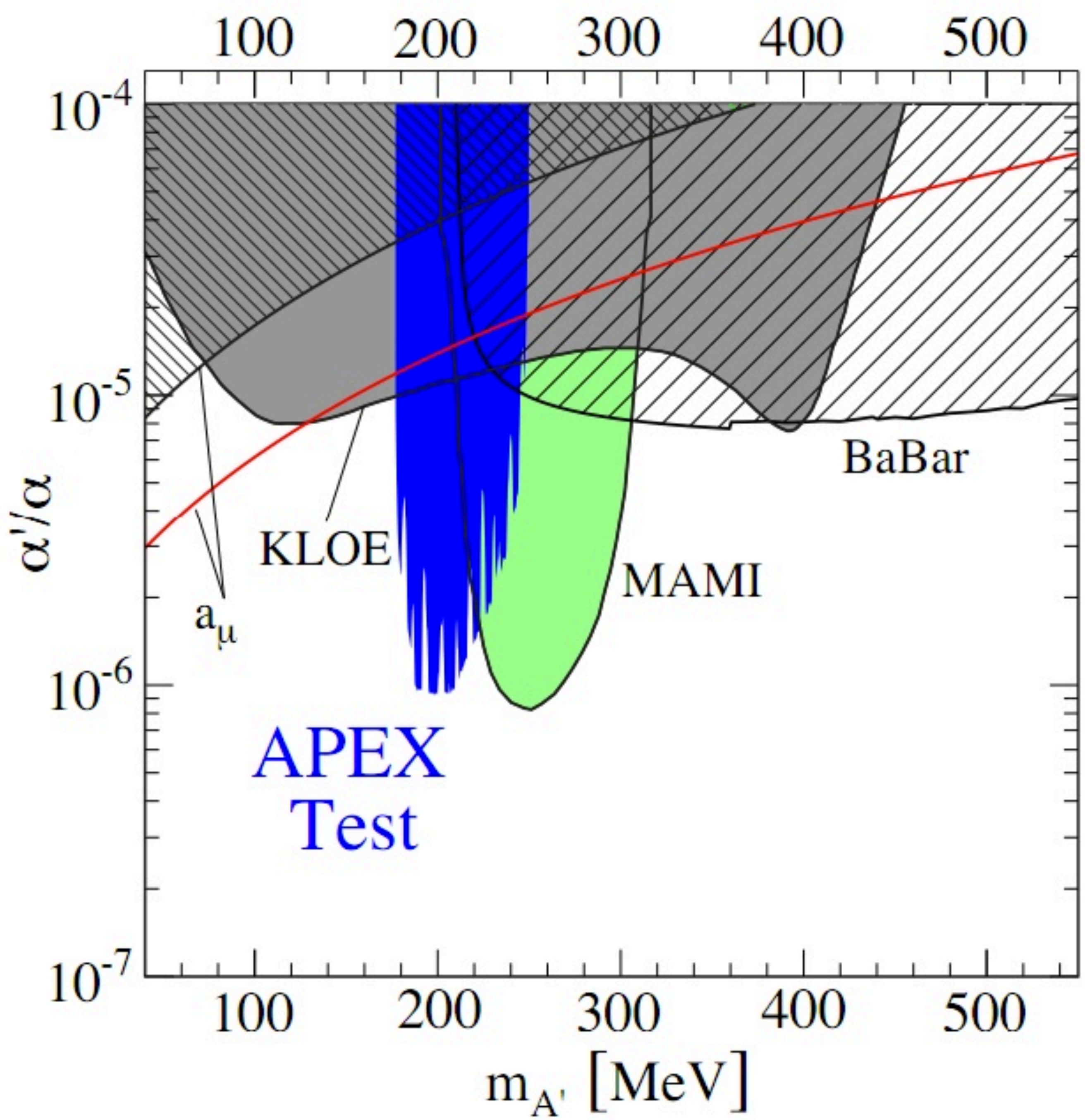
$$\frac{d\sigma(A')}{d\sigma(\gamma^*)} = \left(\frac{3\pi\epsilon^2}{2N_{\text{eff}}\alpha}\right)\frac{m_{A'}}{\delta m} = \frac{S_{\delta m}}{B^{\gamma^*}_{\delta m}}$$

(See APEX proposal)

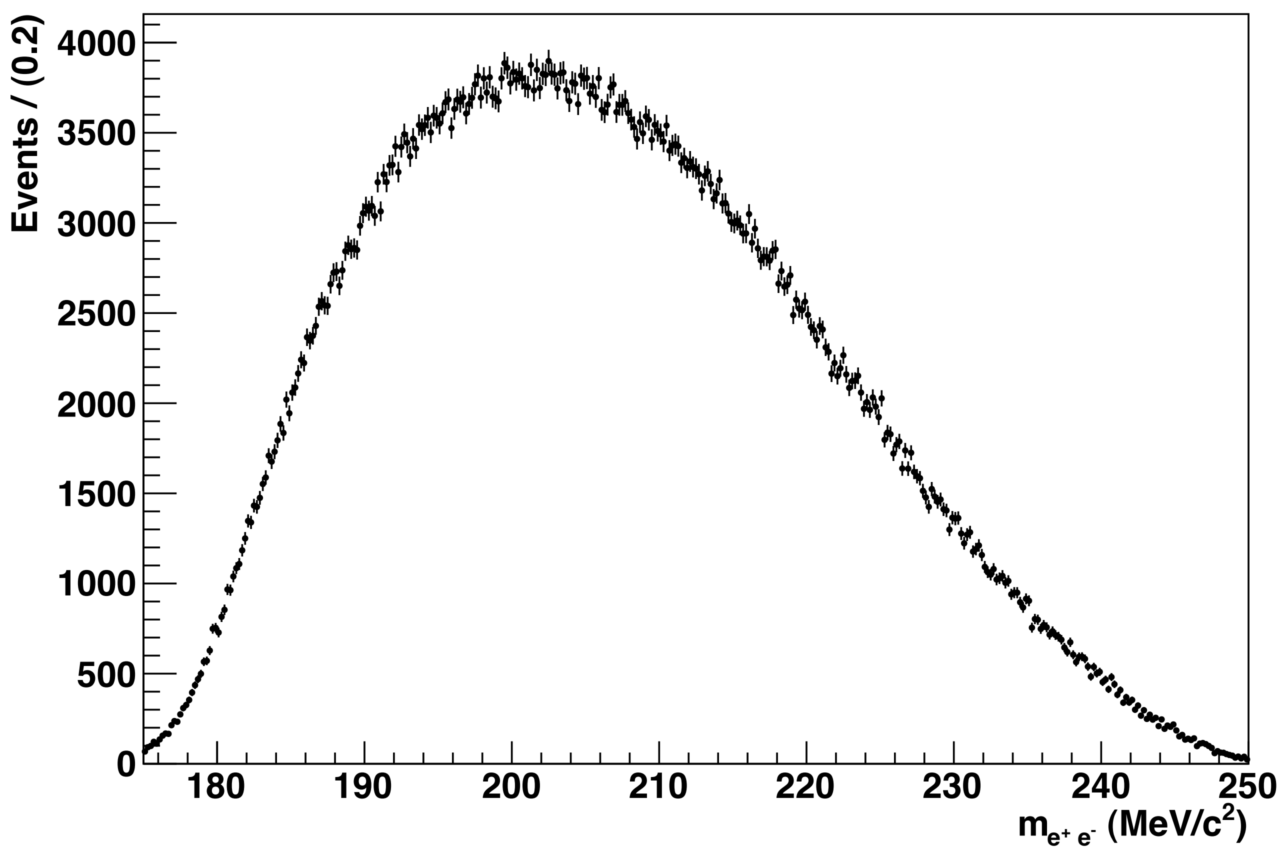
Normalize all backgrounds to γ^* background

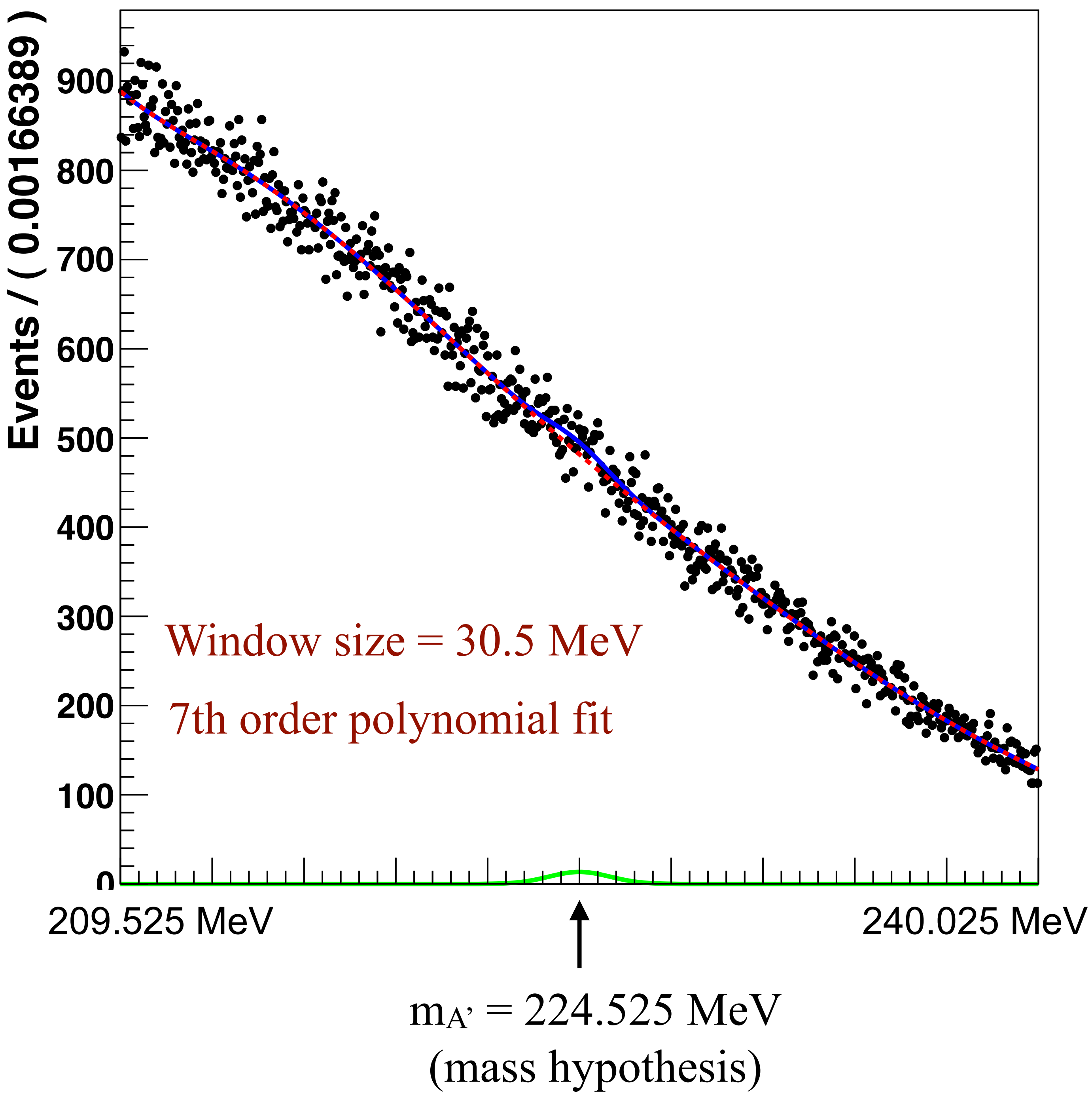
- Ratio *f* of radiative-only cross section to full trident cross section determined via Monte Carlo to vary linearly from 0.21 to 0.25 across APEX mass range

$$\left(\frac{\alpha'}{\alpha}\right)_{max} = \left(\frac{S_{max}/m_{A'}}{f\cdot\Delta B/\Delta m}\right)\times\left(\frac{2N_{\text{eff}}\alpha}{3\pi}\right)$$



The Procedure





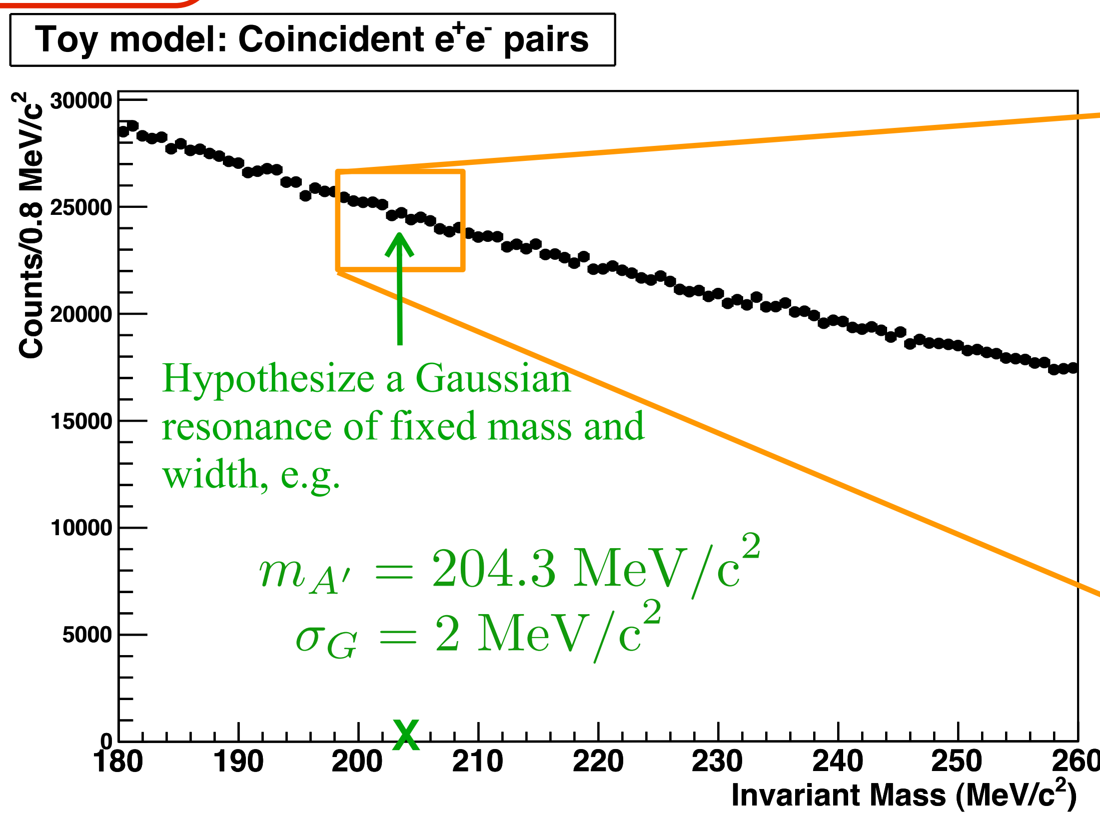
Procedure: Scan spectrum; look for significant excesses; determine upper limit on number of signal events, S, at each mass hypothesis.

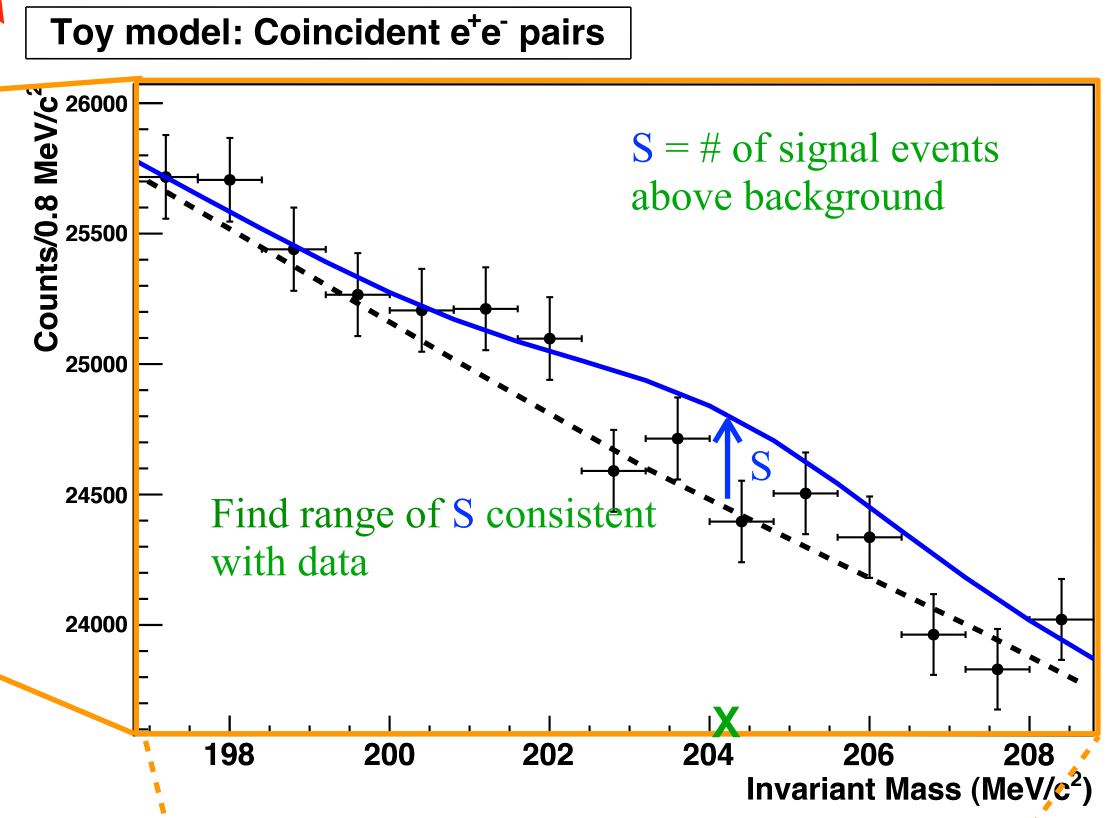
Challenge: Very high statistics (~770,000 events); regular raster scan methods (linear or quadratic background models) may be too crude.

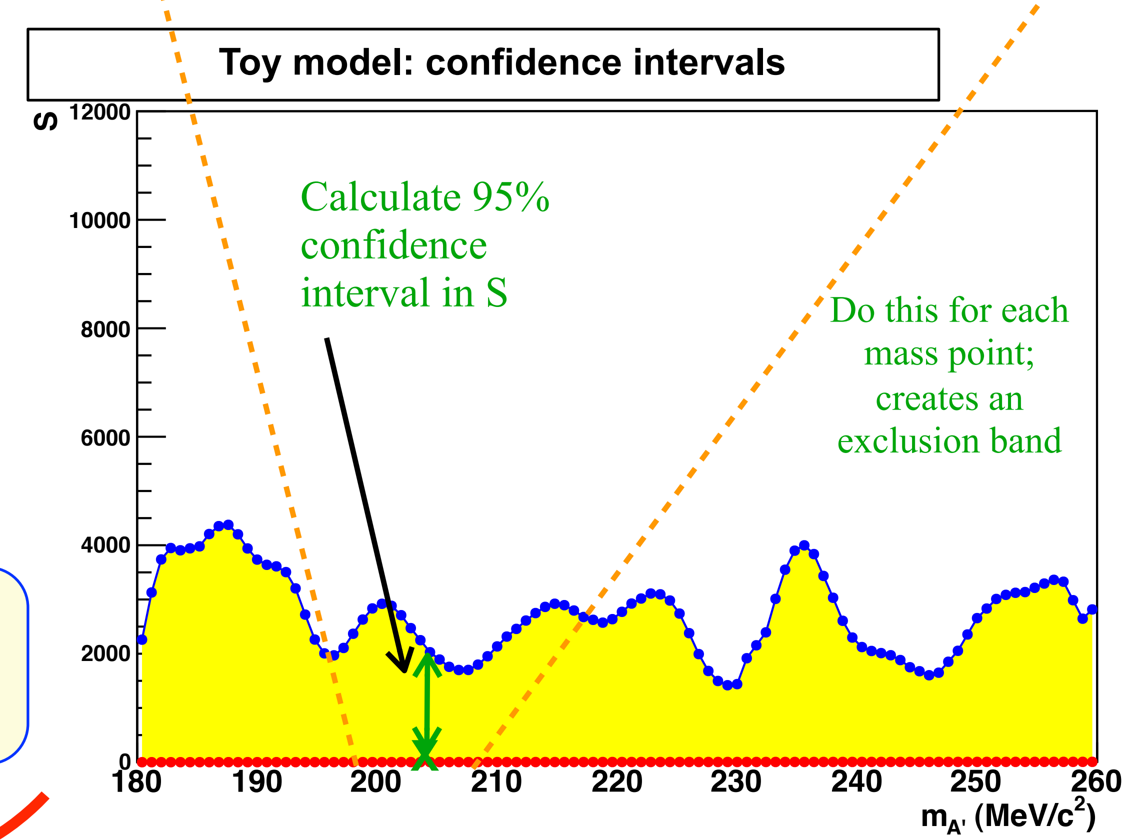
Final choice for test run: 30.5 MeV wide window and 7th order polynomial background fit.

Center for Cosmology and Particle Physics

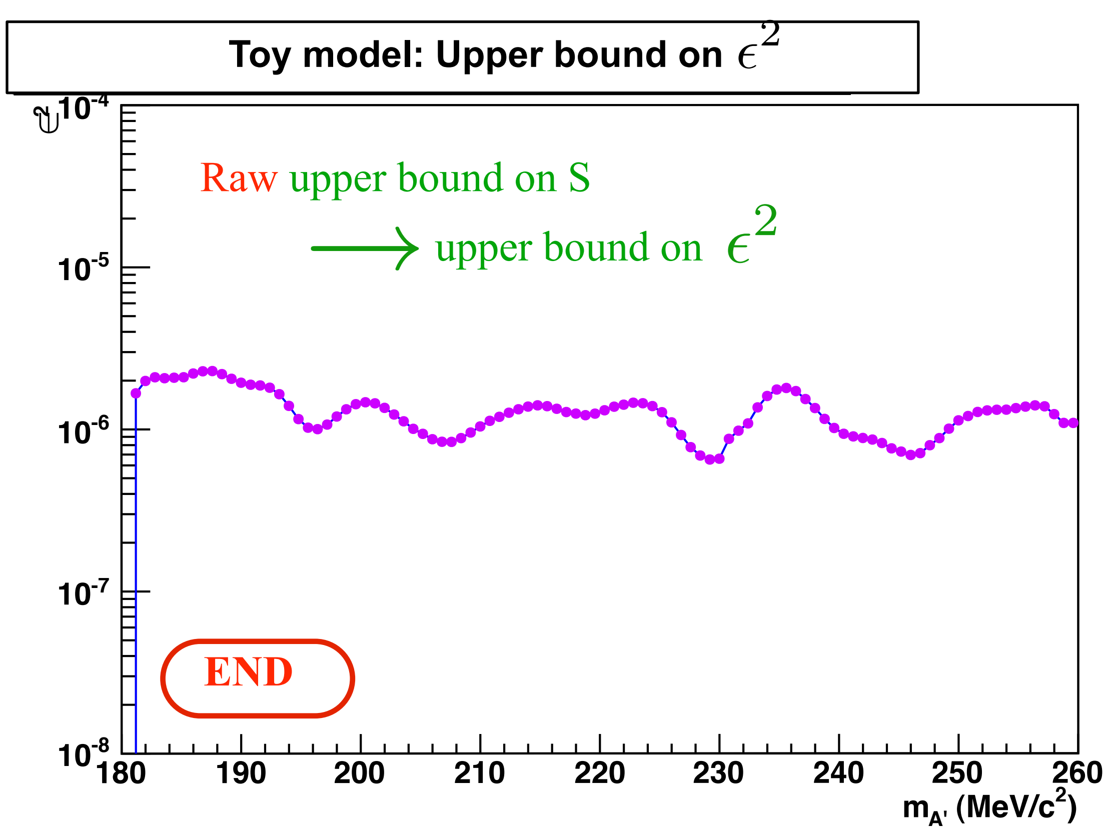
START: S = 0







One more step...



Test statistic based on **Profile Likelihood Ratio**

Probability model:

$$P(m_{e^+e^-} \mid m_{A'}, \sigma, S, B, a_i) = \frac{S \cdot \mathrm{N}(m_{e^+e^-} \mid m_{A'}, \sigma) \; + \; B \cdot \mathrm{Polynomial}(m_{e^+e^-}, a_i)}{S + B}$$

This becomes a likelihood function, L, as a function of the model parameters.

- Separate parameters into
  - Parameters of interest (physics parameters: mass, # signal events, etc.)
  - Nuisance parameters (systematics, i.e., everything else: background shape, etc.)

A test statistic that incorporates nuisance parameters is the **profile likelihood ratio**:

The PLR is the maximum likelihood under the null hypothesis as a fraction of its largest possible value; large values of λ indicate that the null is reasonable

$$\lambda(S) = \frac{L(S, \hat{\hat{B}}, \hat{\hat{a}}_i)}{L(\hat{S}, \hat{B}, \hat{a}_i)}$$

Parameters of interest (POI), set to values for null hypothesis (S = 0)

Conditional MLE of nuisance parameters (i.e., for S = 0)

MLE of POI

MLE of nuisance parameters

Test statistic is then $-2 \log \lambda$

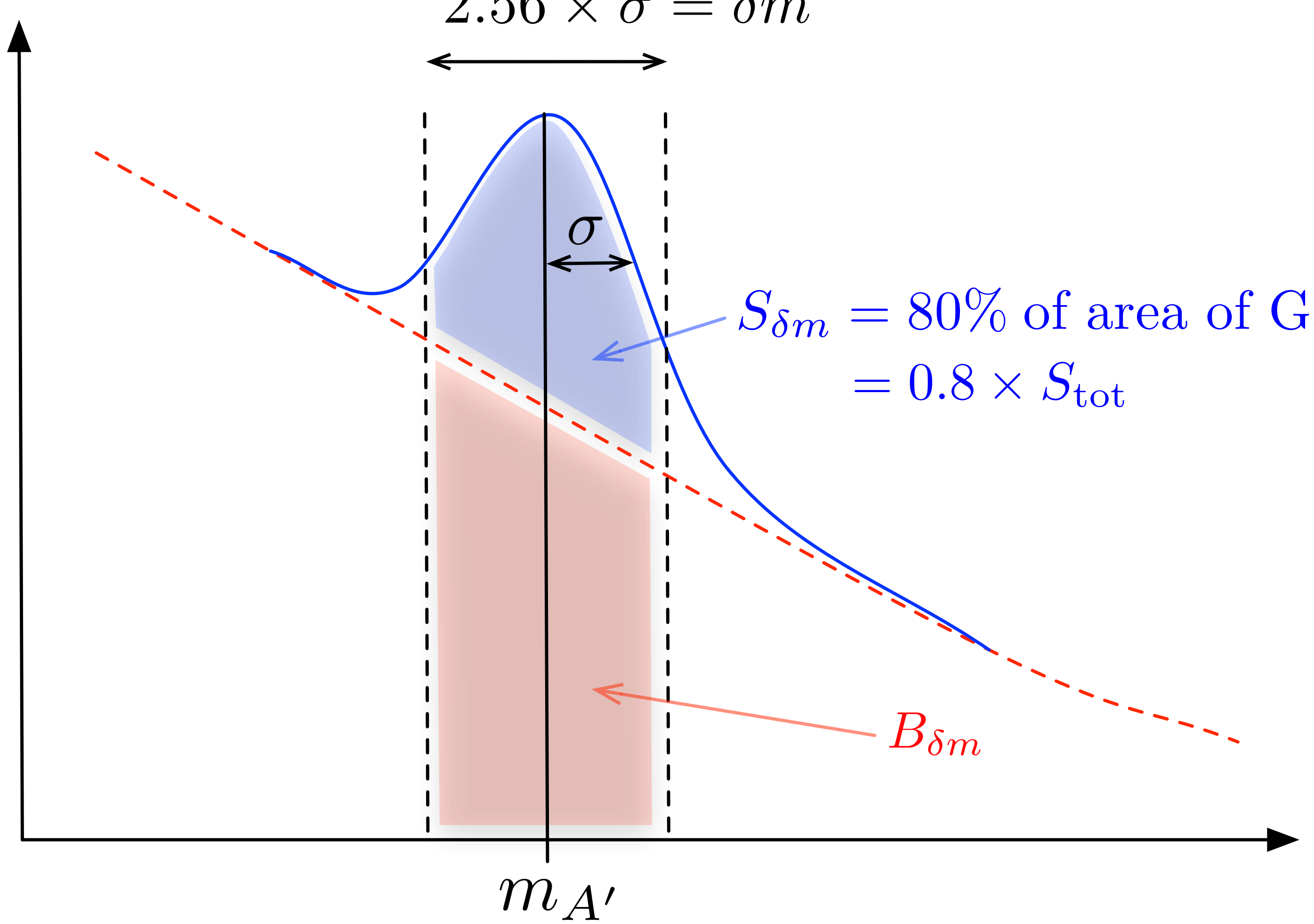
Wilks's Theorem: $-2 \log \lambda$ is distributed as a χ^2 with # d.o.f. = # POI

Upper limit on S yields an upper limit on ϵ^2

$$\epsilon^2(m_{A'}) = \frac{S_{\delta m}}{B_{\delta m}} \frac{F(m_{A'})\ \delta m}{m_{A'}} \left(\frac{2N_{\text{eff}}\alpha}{3\pi}\right)$$

$\delta m = \text{mass window} = 2.56 \times \sigma$ σ = width of Gaussian ~ 1 MeV (mass-dependent)

Background factor: $F \sim 5$ (mass-dependent) $N_{\text{eff}} = 1$

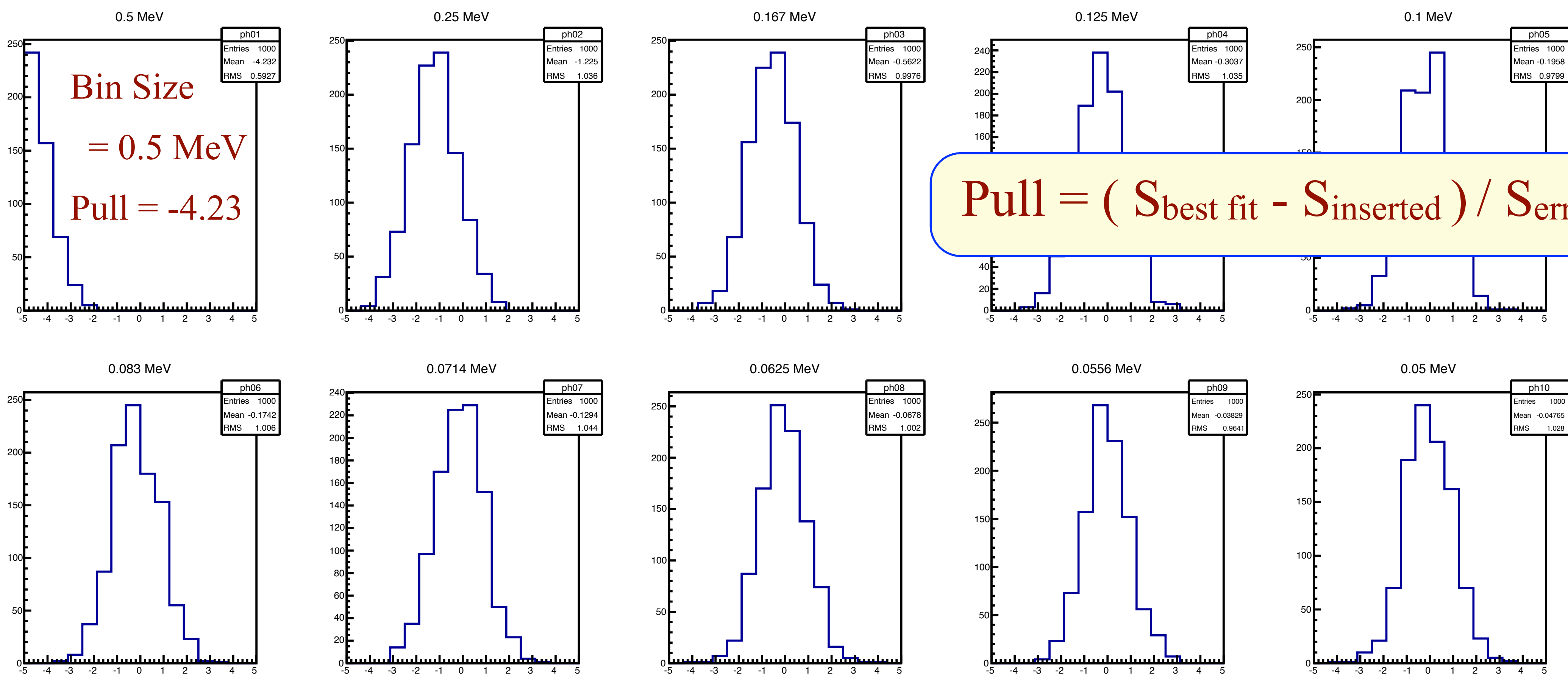


Setting limit on S: Binned analysis

We would ideally do an unbinned analysis; such a large number of events makes this intractably time consuming. A binned analysis is manageable and any pulls due to choice of statistical tool (ROOT, Mathematica, etc.) can be made negligible by choosing small enough bins.

Example: With ROOT, do ensemble studies at a representative mass point; pseudodata sets based on spectrum of 10% of test run data, scaled up

Here, $S_{inserted} = 0$; large pulls for 0.5 MeV bins; pulls negligible at 0.06/0.05 MeV bins



$$\text{Pull} = (S_{best\ fit} - S_{inserted}) / S_{error,\ fit}$$

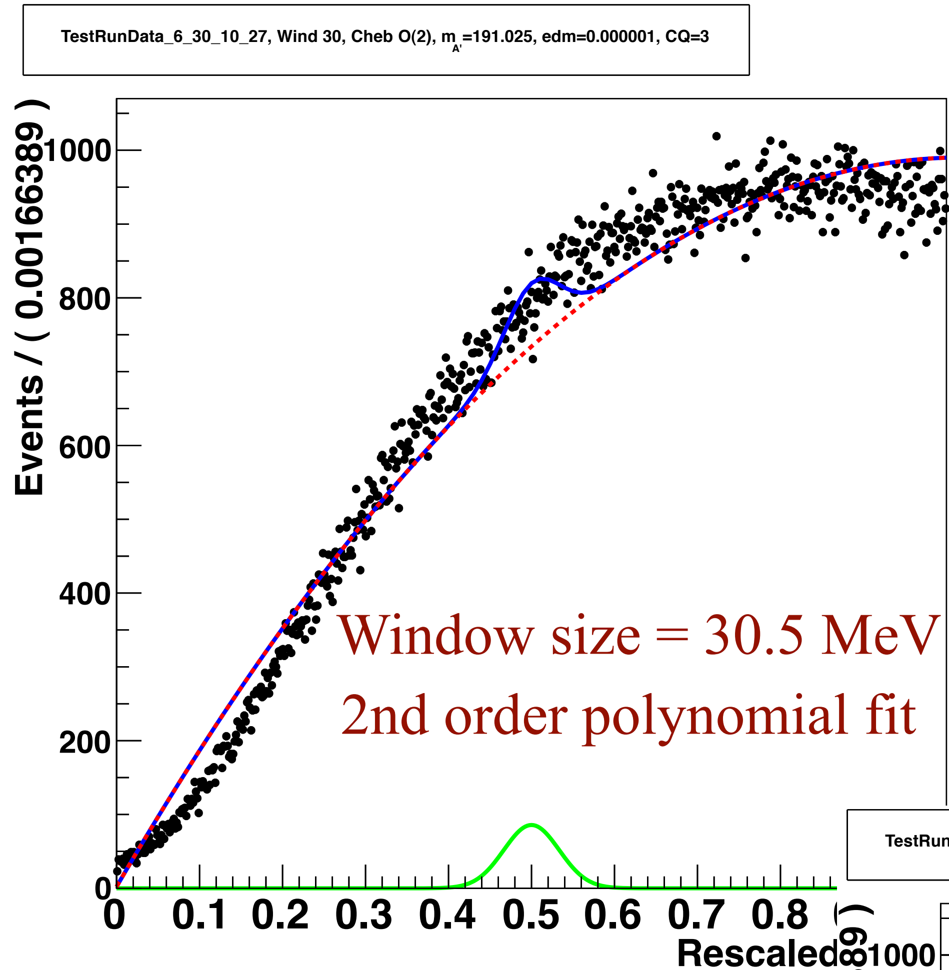
Bin Size = 0.0625 MeV

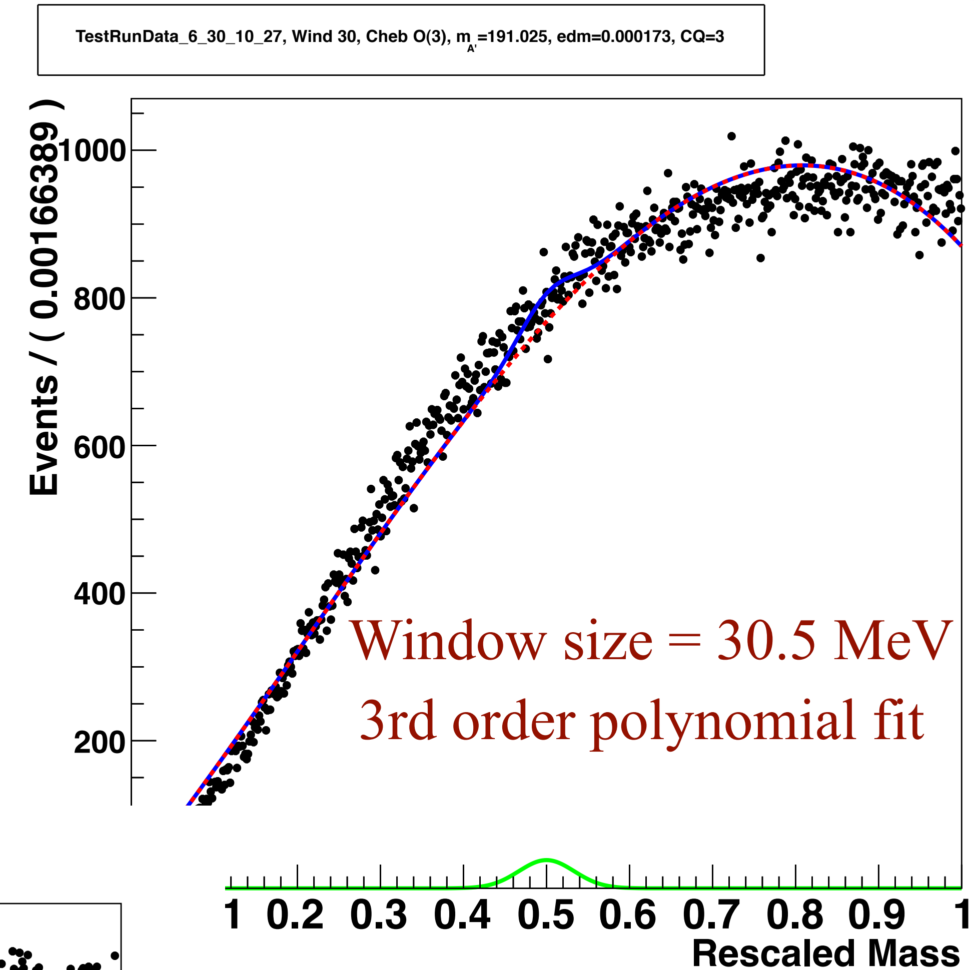
Pull = -0.067

Bin Size = 0.05 MeV

Pull = -0.048

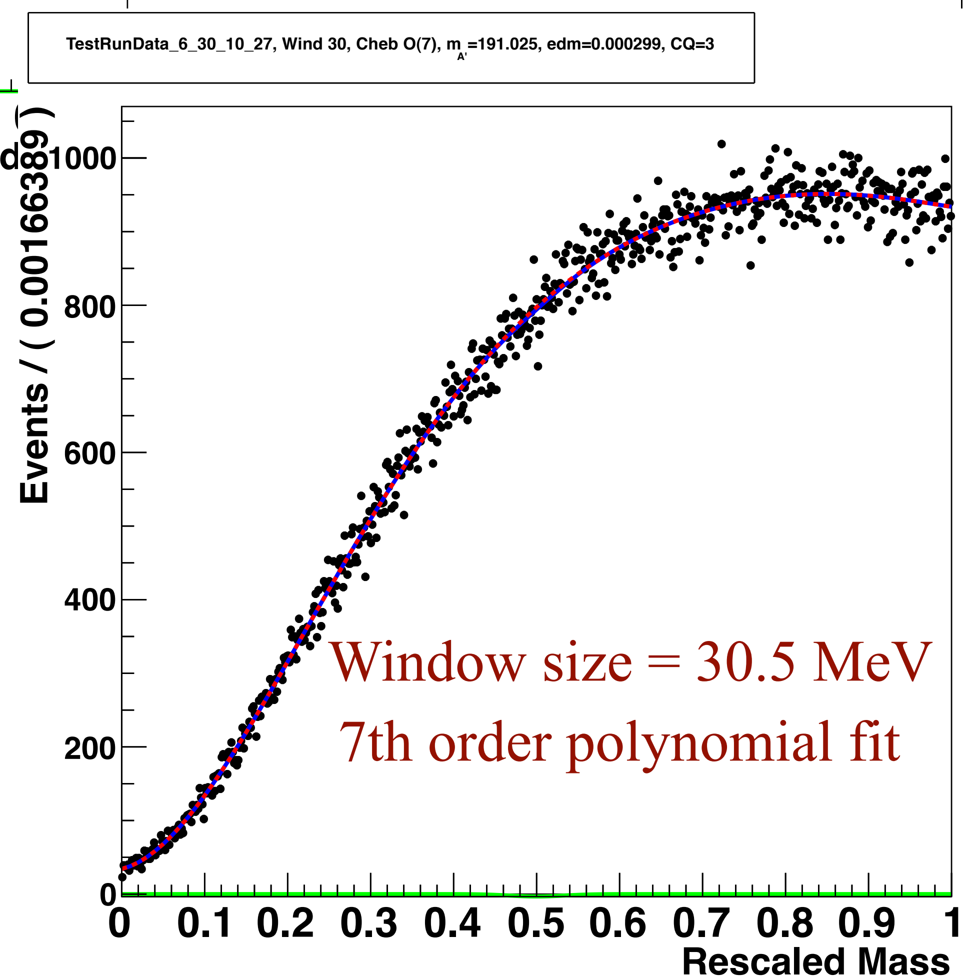
Center for Cosmology and Particle Physics





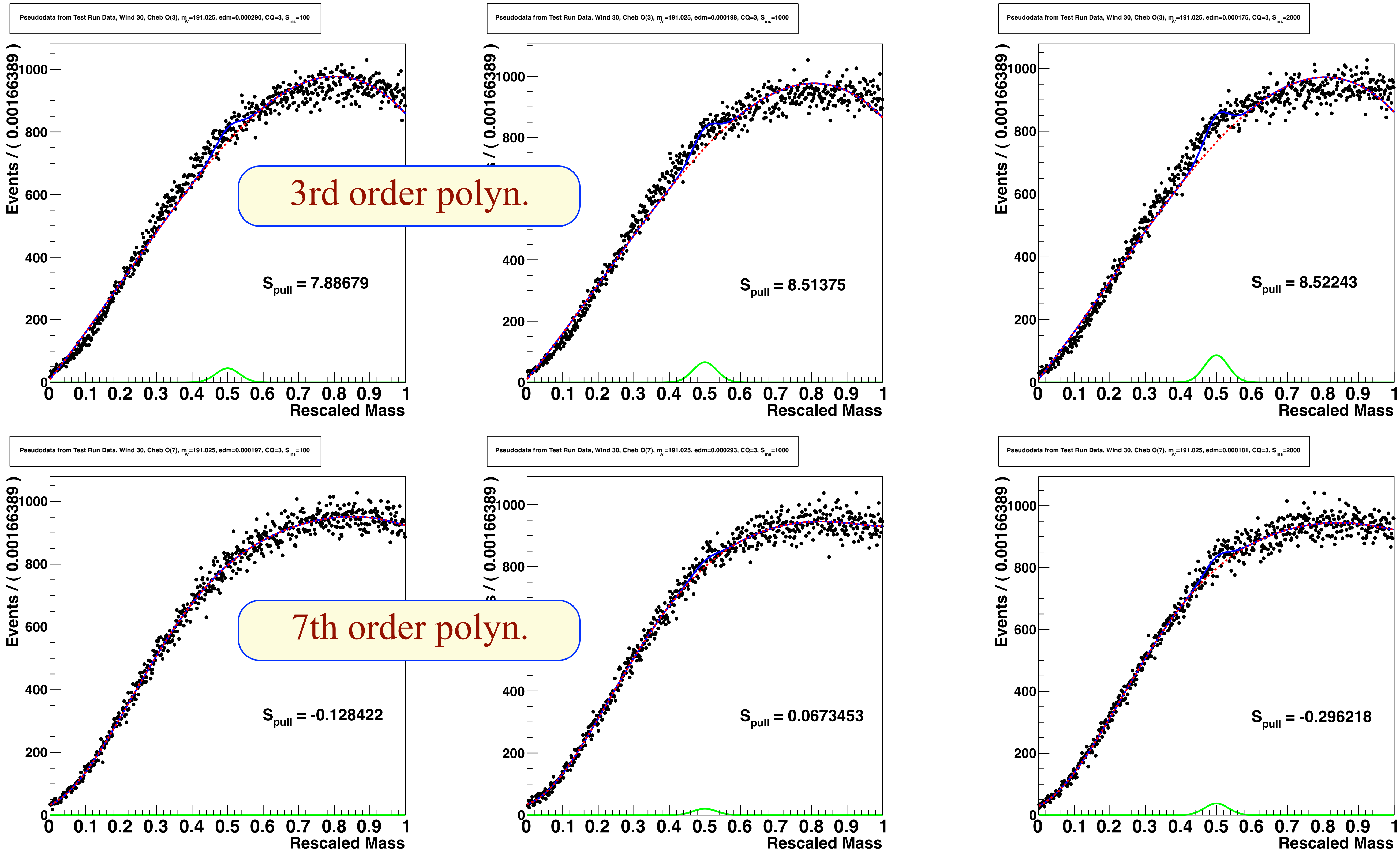
Lower order polynomials fail to capture the background shape and lead to spurious peaks

$S_{inserted} = 0$



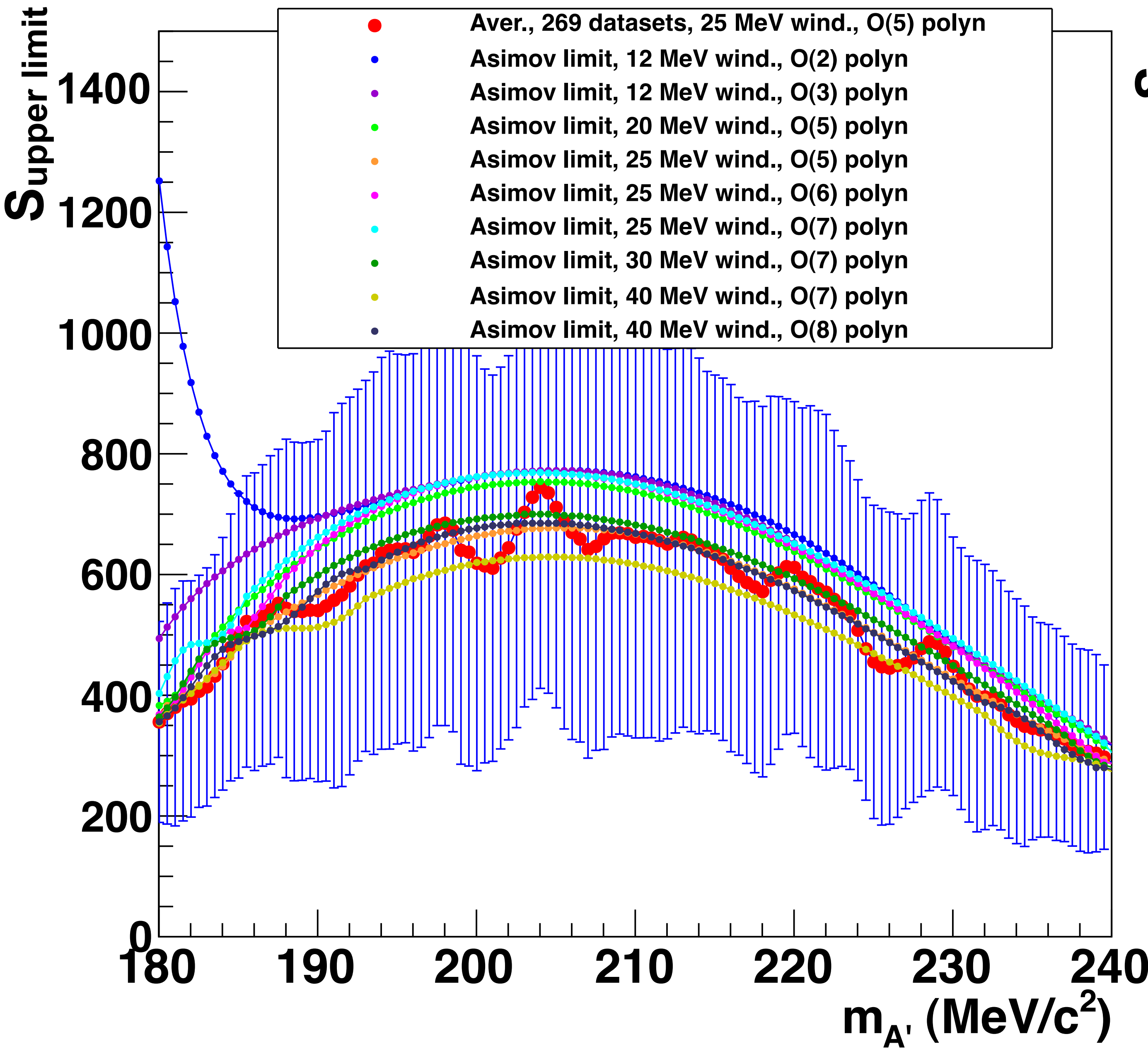
Center for Cosmology and Particle Physics

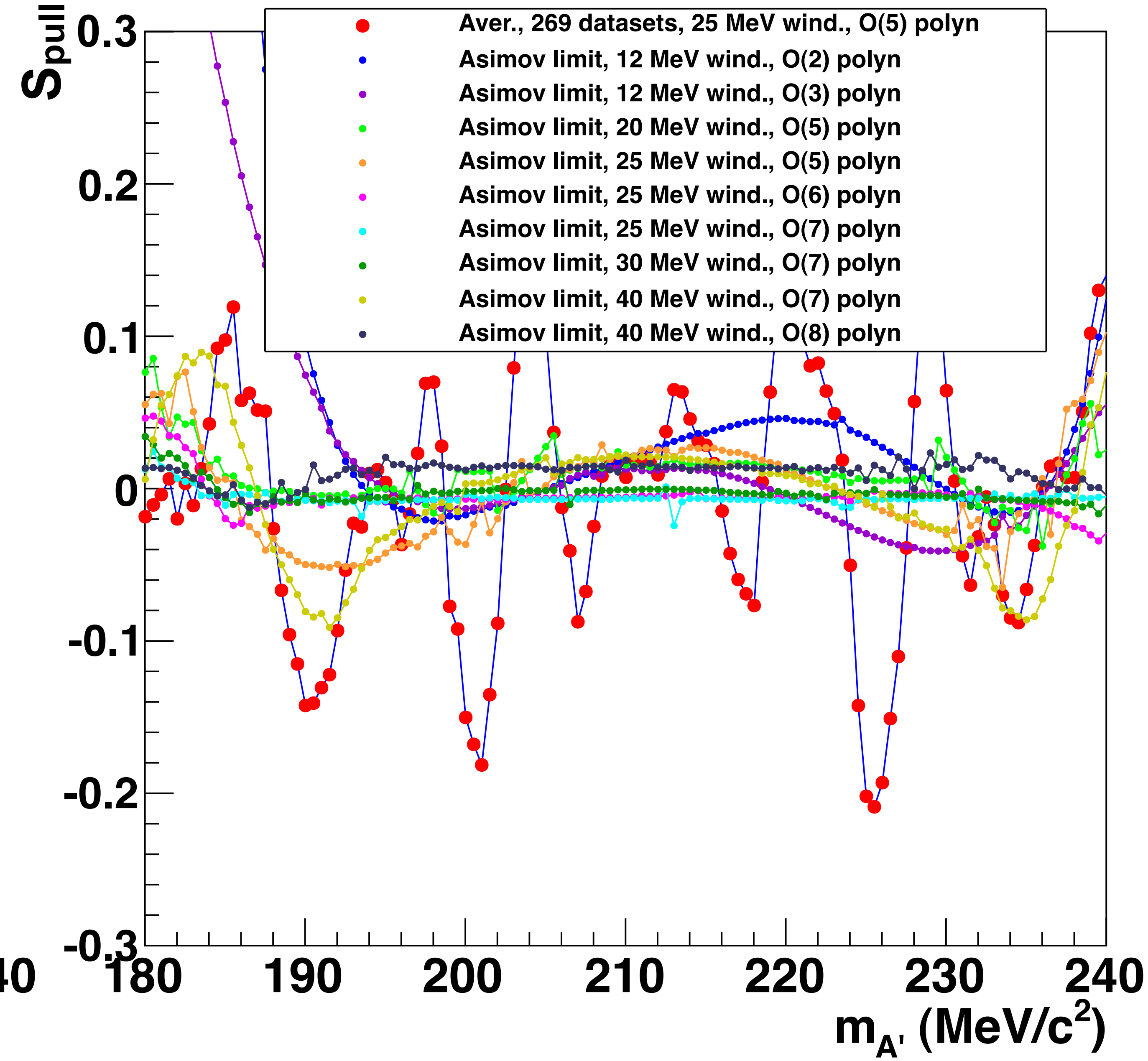
$S_{inserted} = 100$ | $S_{inserted} = 1000$ | $S_{inserted} = 2000$



Representative pseudodata sets; not an ensemble

Setting limit on S: Window size and polynomial order



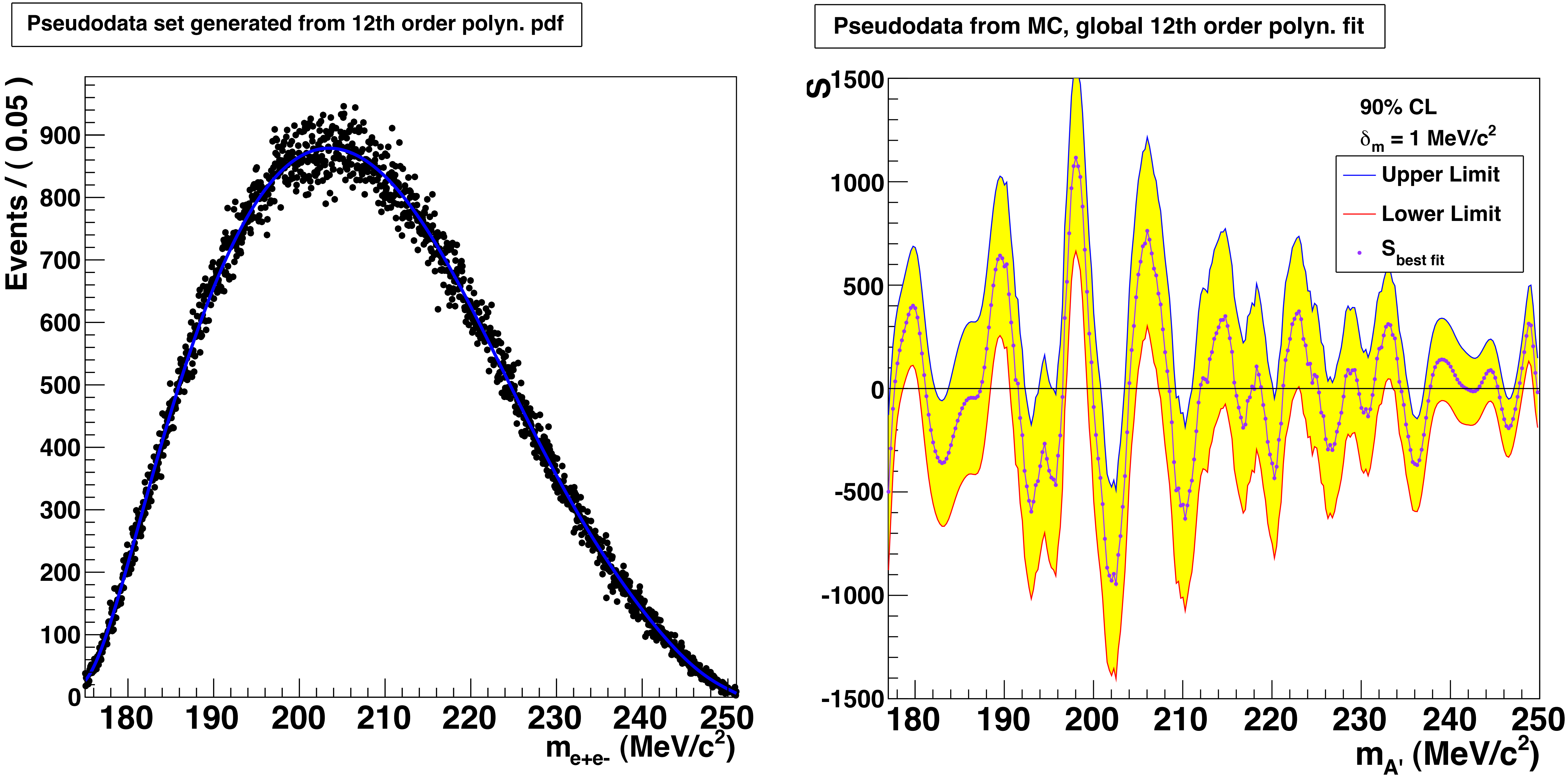


30 MeV window with 7th order polynomial (dark green line) optimizes between sensitivity in S and minimal pull

Center for Cosmology and Particle Physics

Characteristics of limit plots

Particular choice of polynomial basis and order captures the shape of the actual data or MC to a greater or lesser extent

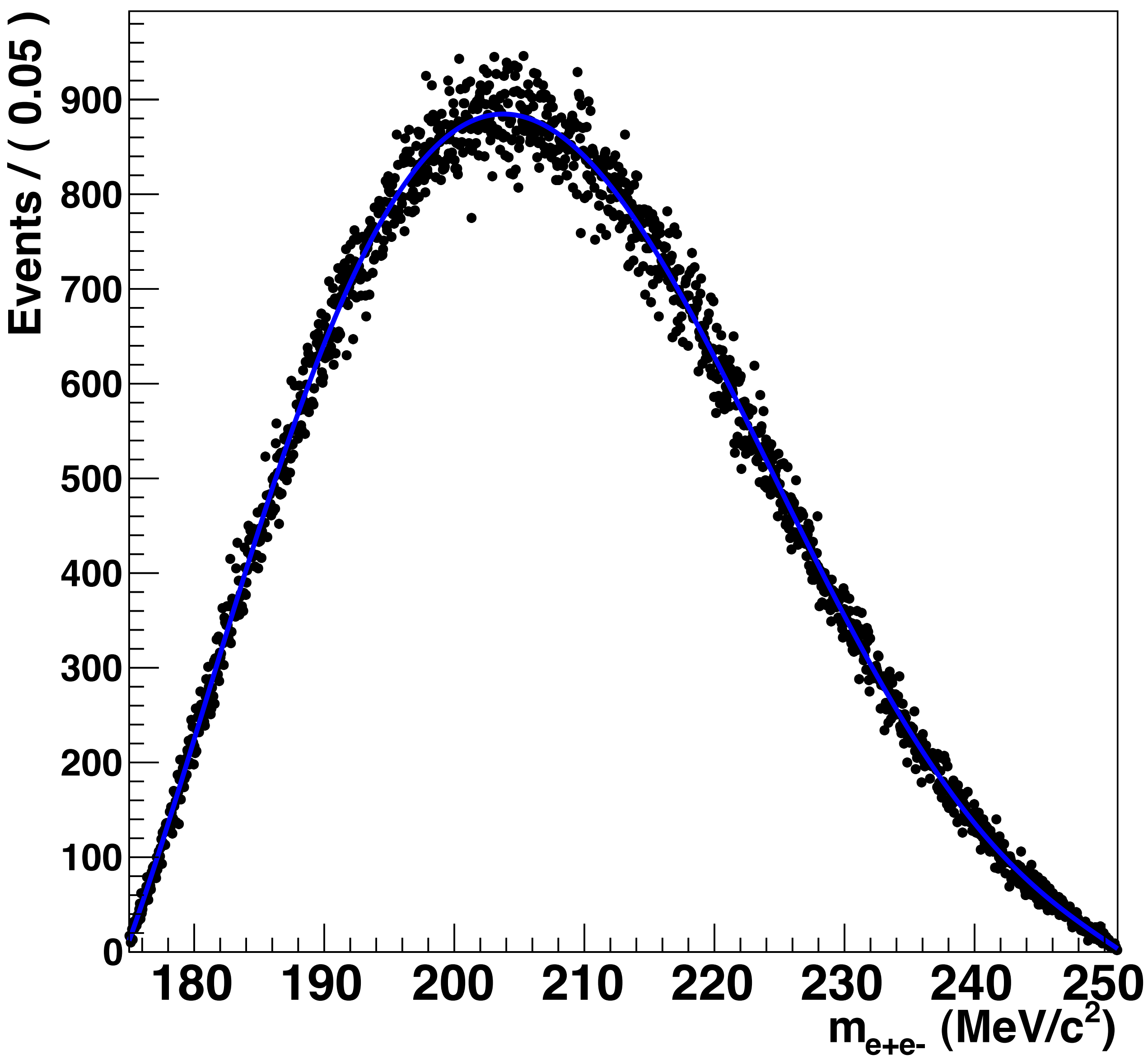


12th order global polyn. fit, 7th order polyn. background in window

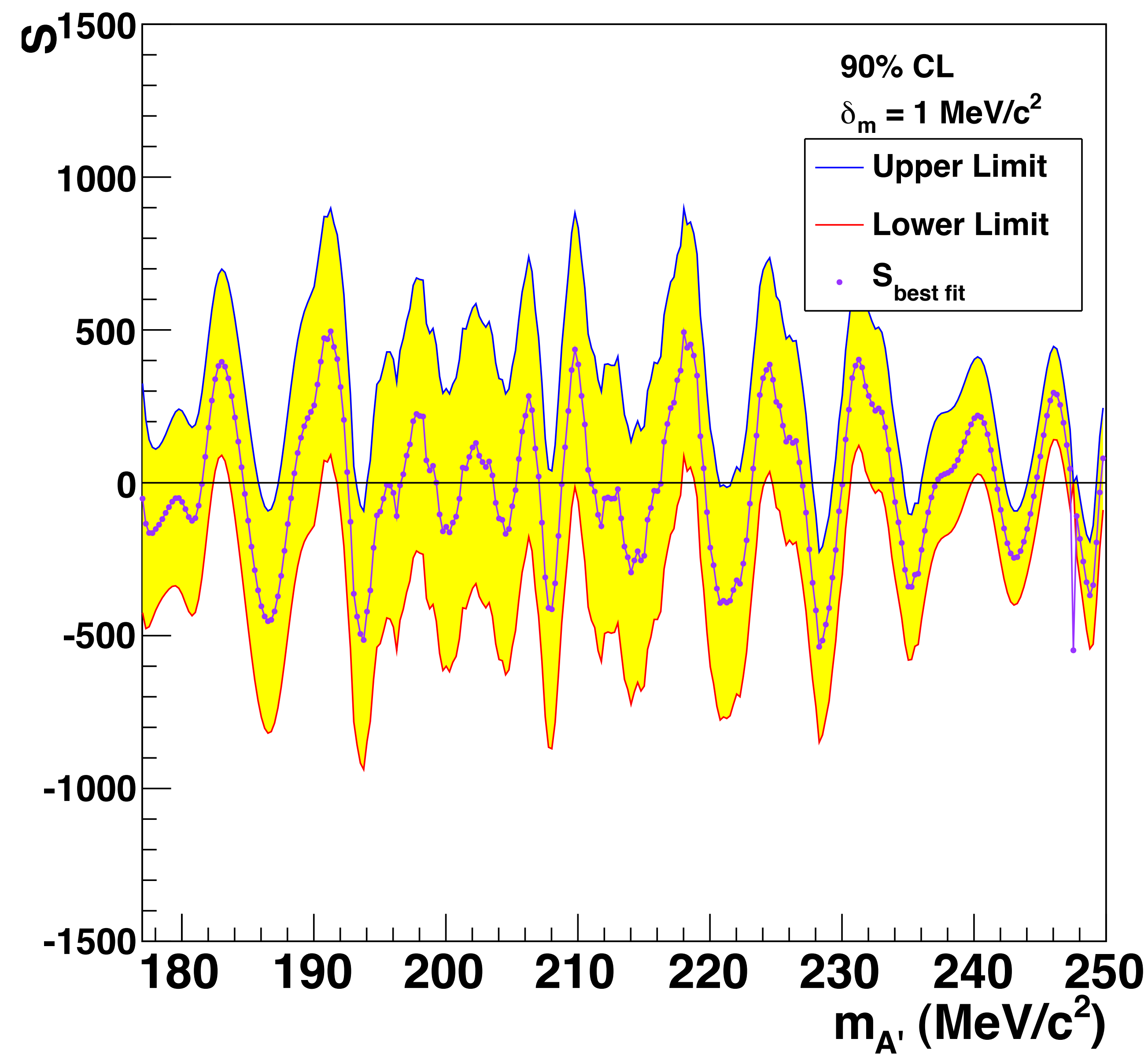
Characteristics of limit plots

Particular choice of polynomial basis and order captures the shape of the actual data or MC to a greater or lesser extent

Pseudodata set generated from 6th order polyn. pdf



Pseudodata from MC, global 6th order polyn. fit

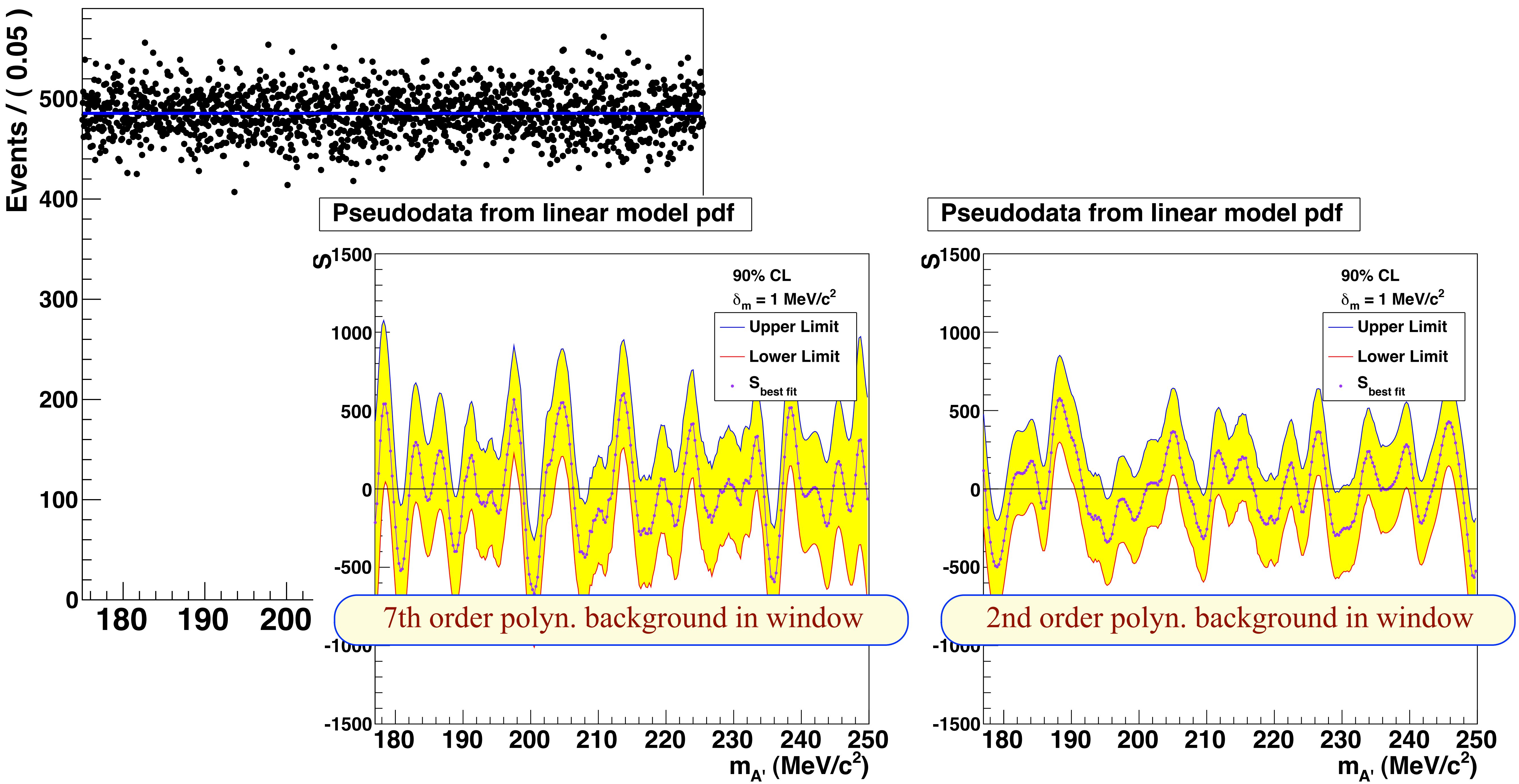


6th order global polyn. fit, 7th order polyn. background in window

Center for Cosmology and Particle Physics

Characteristics of limit plots

Linear model pdf

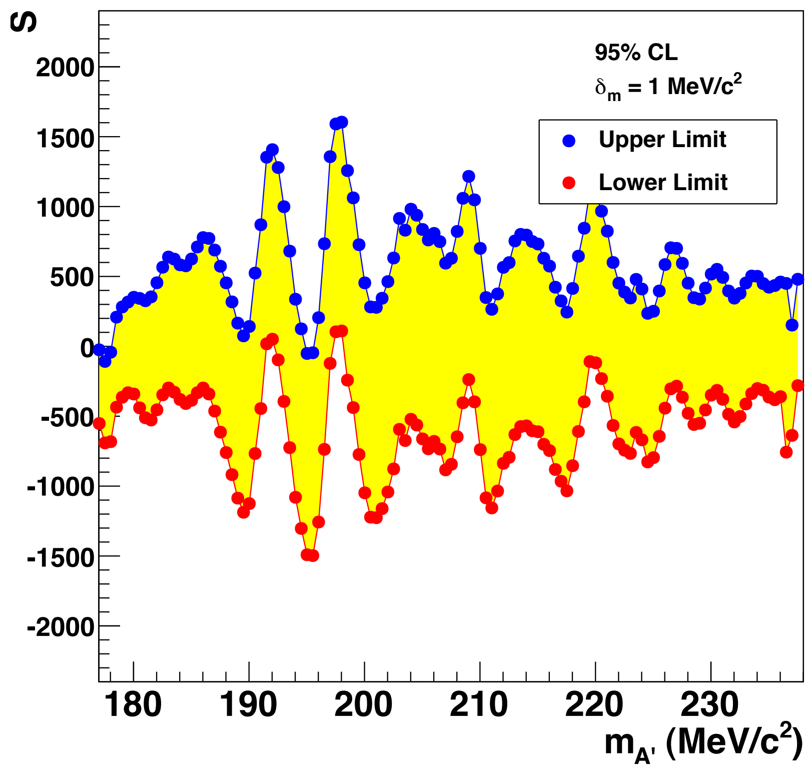


Other background parameterizations yield wiggles, too, though location and size varies; a systematic effect subdominant to statistical error

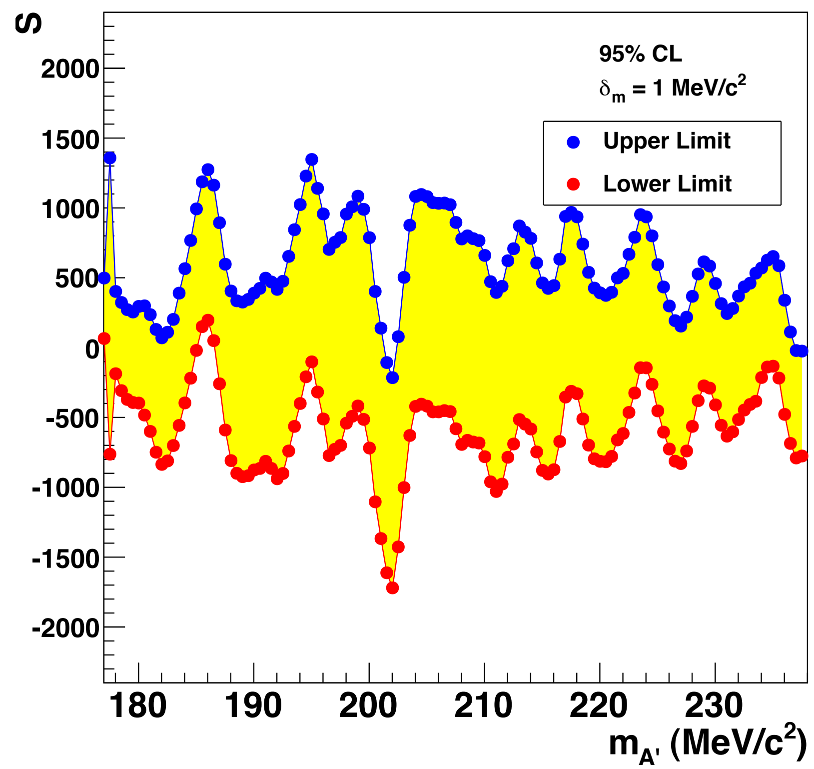
Center for Cosmology and Particle Physics

Blind analysis: Nine histograms; eight pseudodata, one real

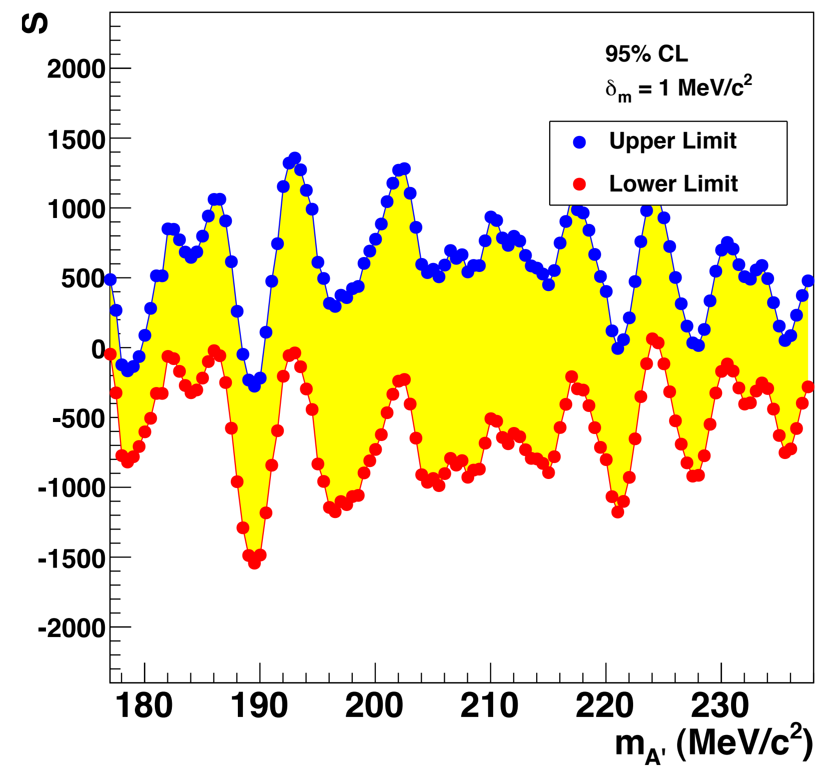
Hist 1



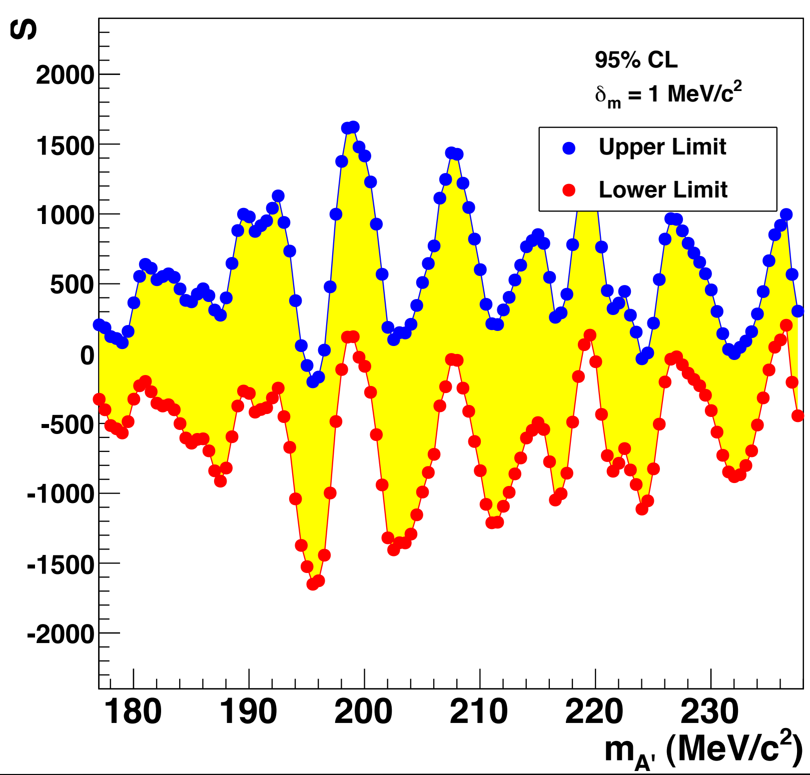
Hist 2



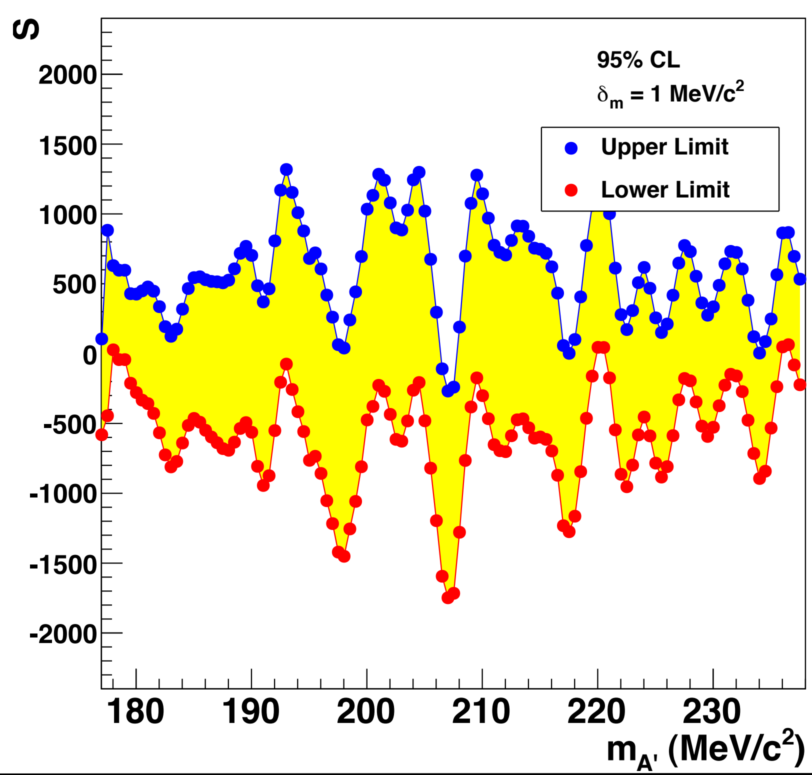
Hist 3



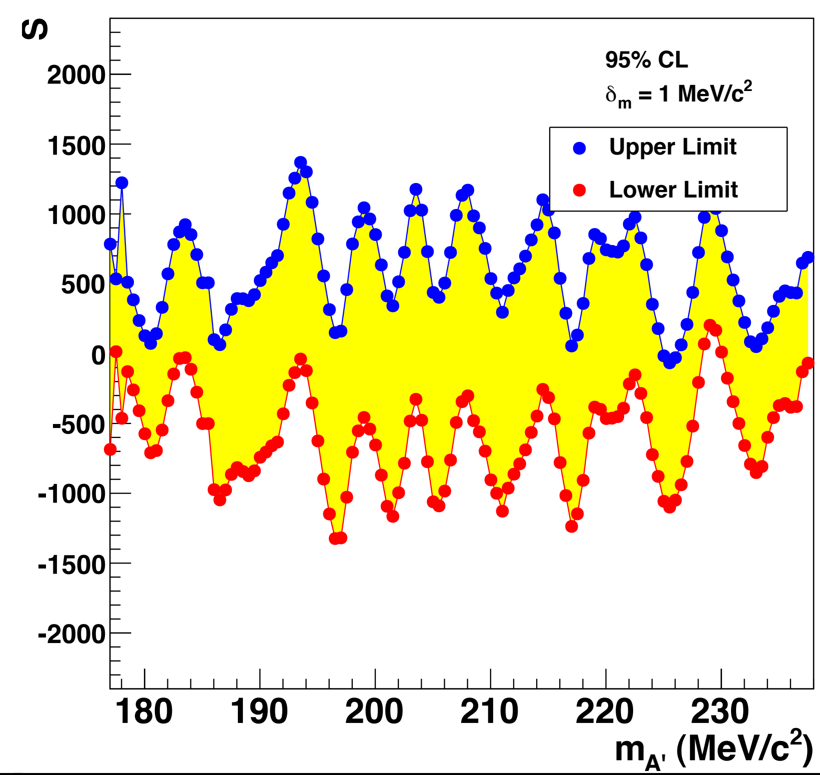
Hist 4



Hist 5

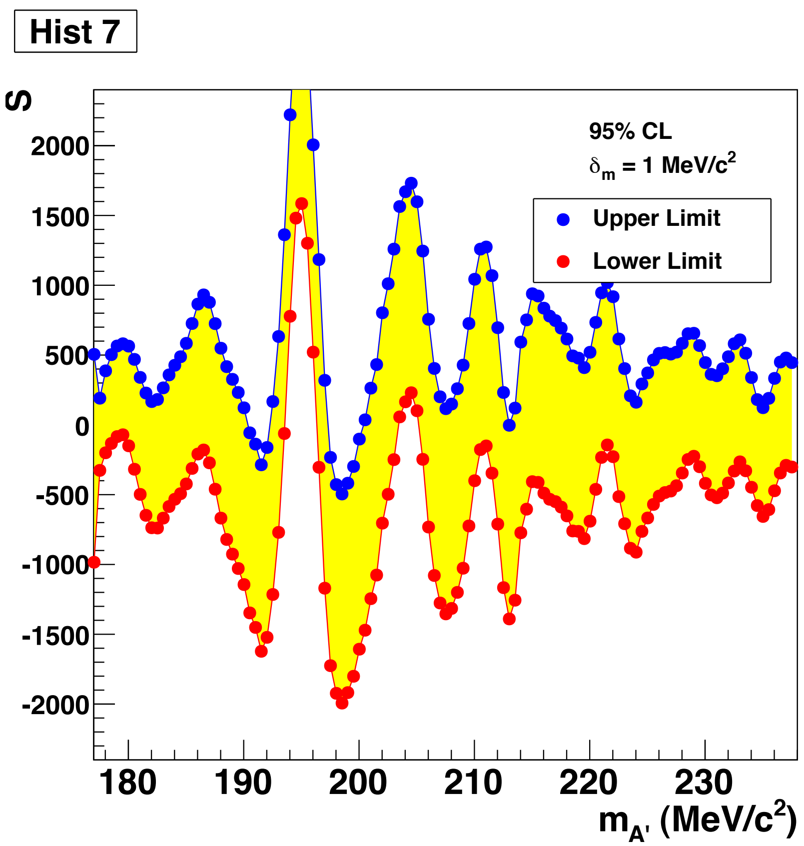


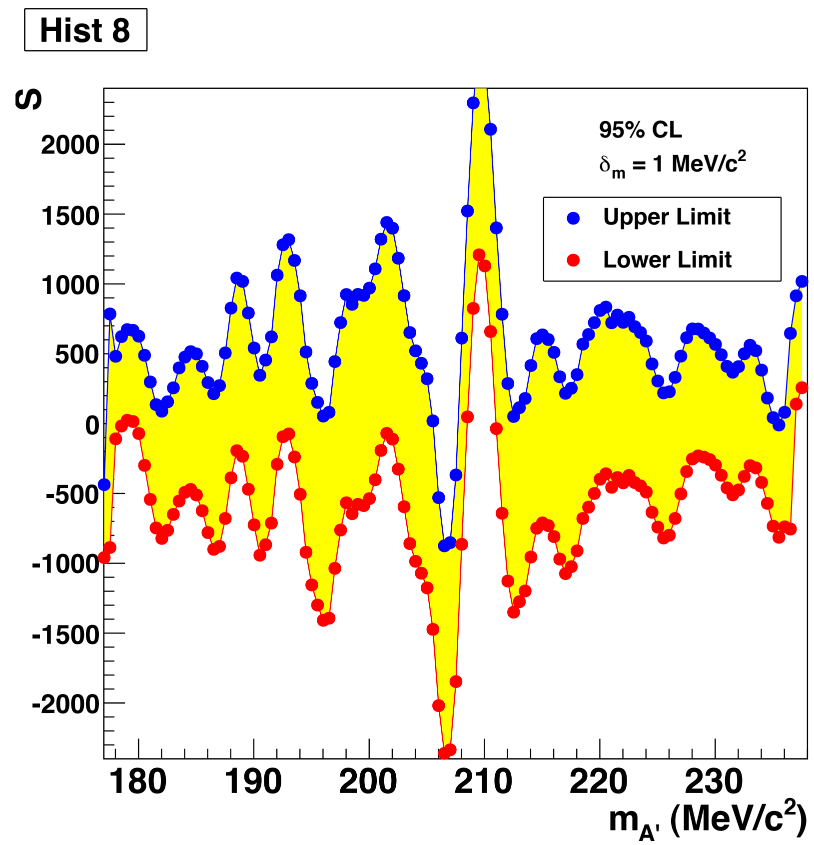
Hist 6

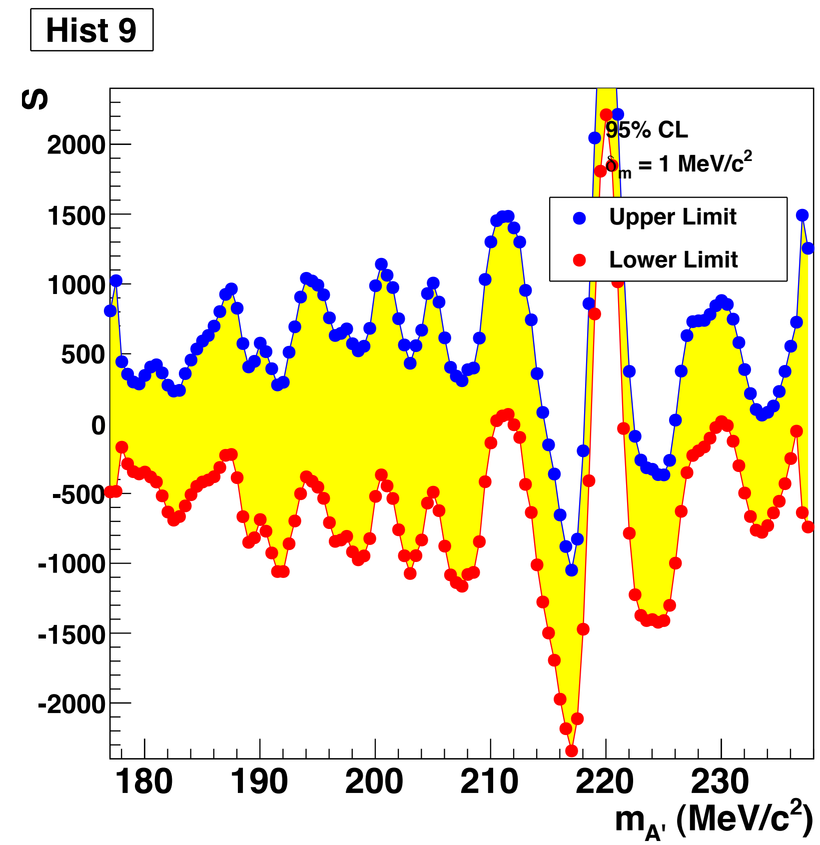


CENTER FOR COSMOLOGY AND PARTICLE PHYSICS

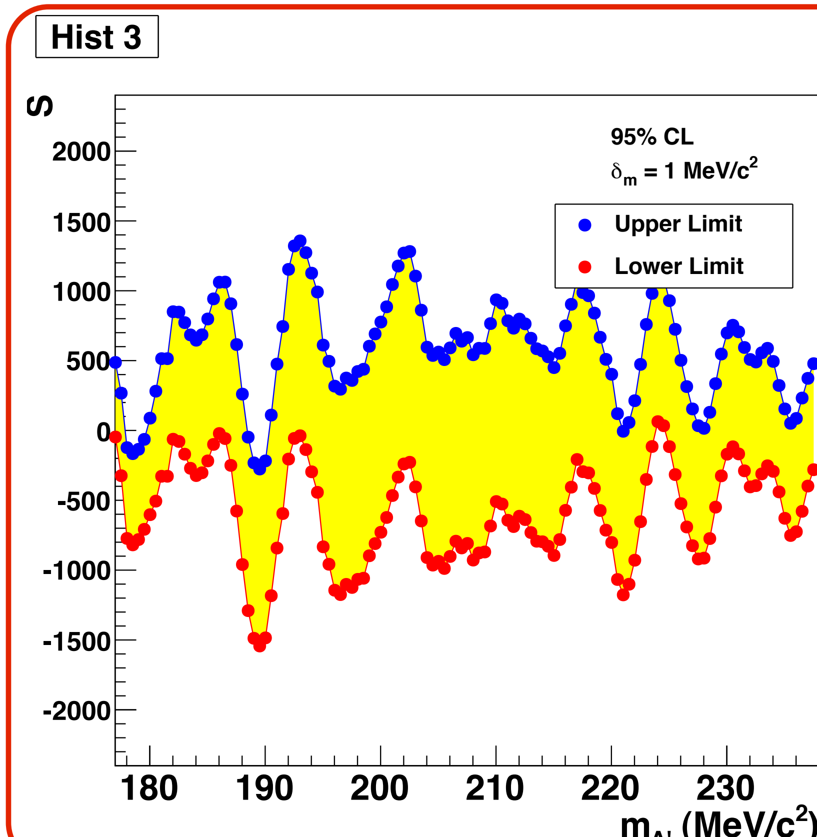
Blind analysis: Nine histograms; eight pseudodata, one real





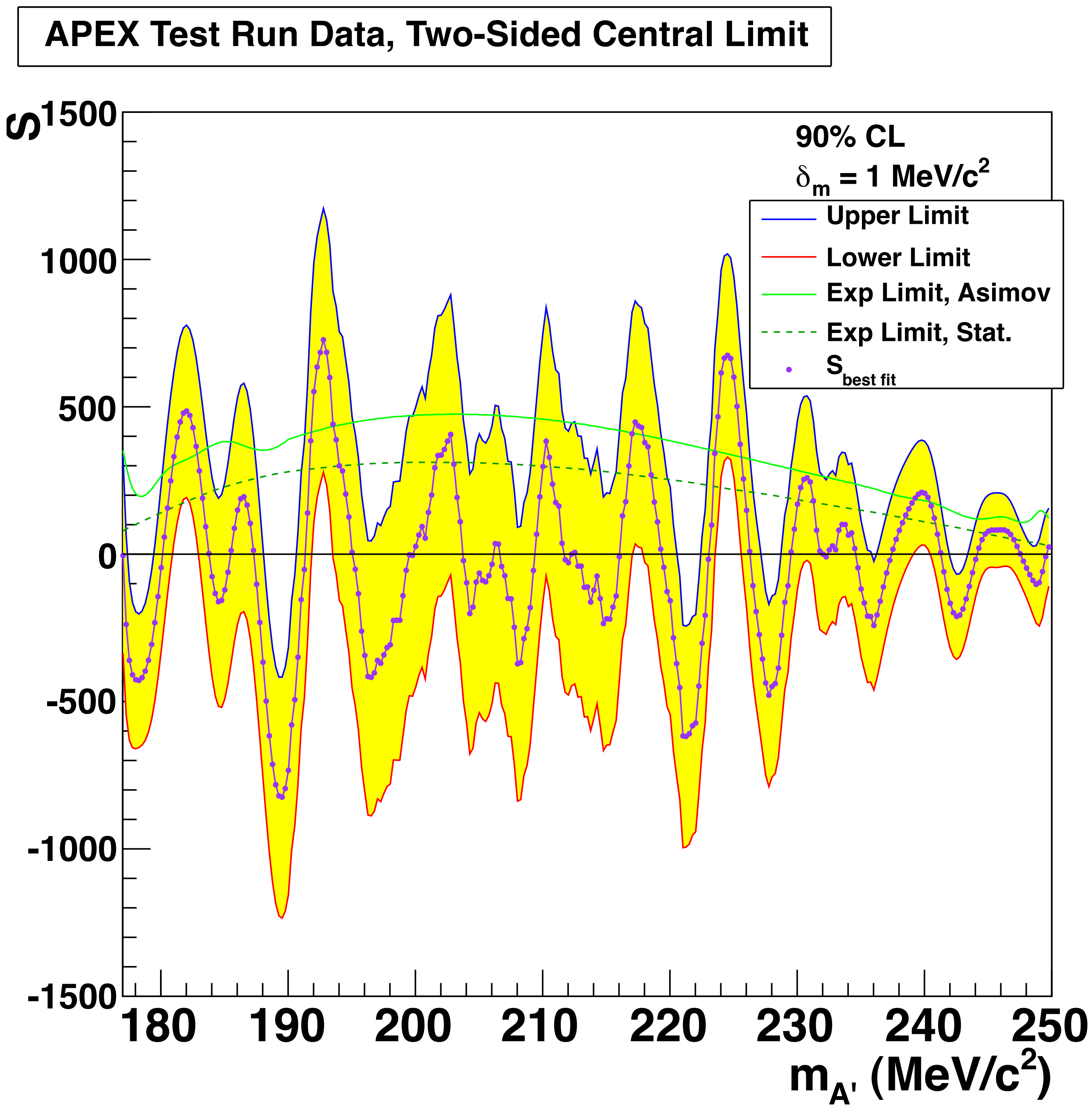


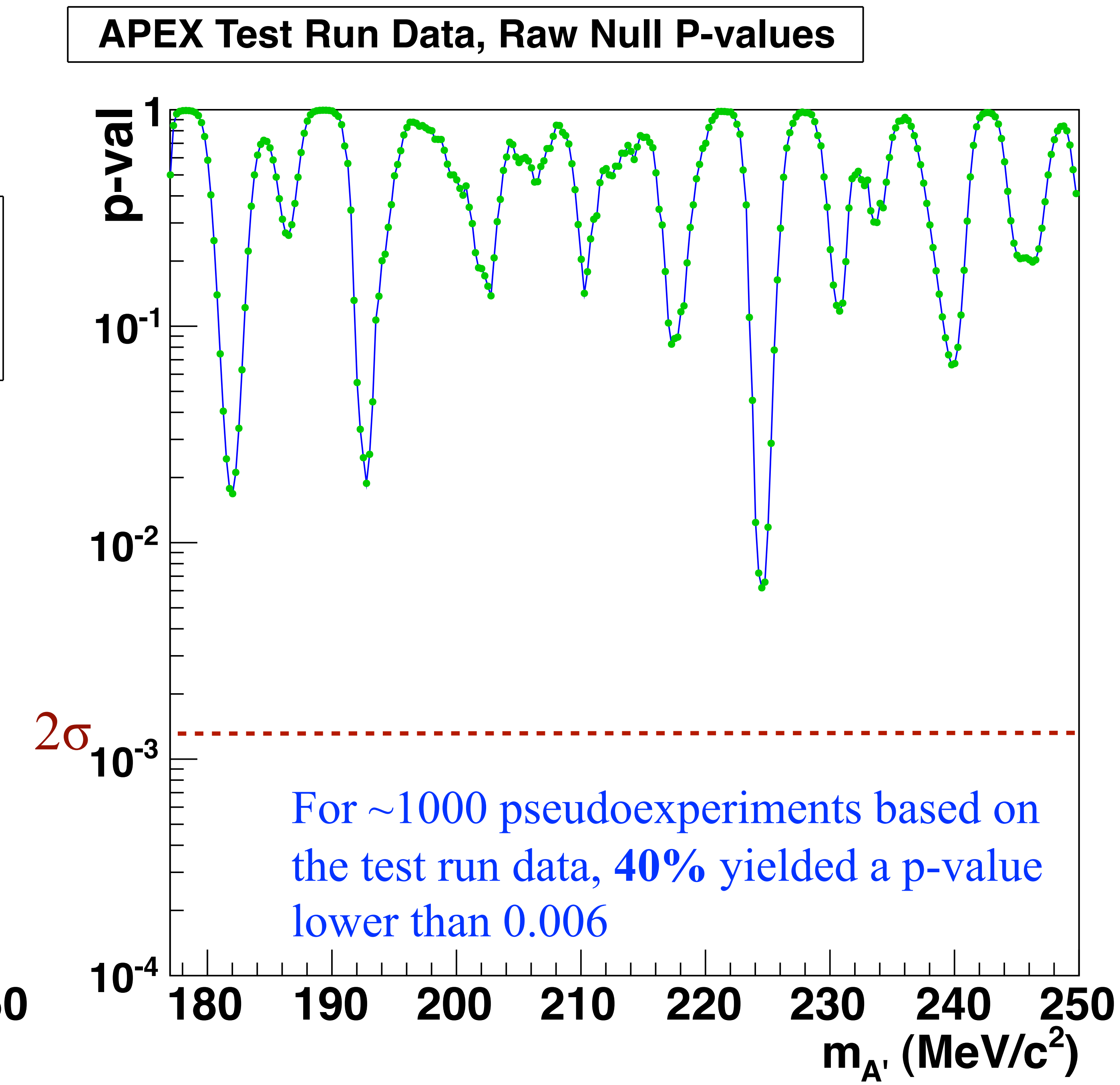
Unblinding: Histogram #3



Additionally, 40 similar histograms were prepared, with peaks of various significances inserted; method found all inserted excesses of nominal significance > 2 sigma

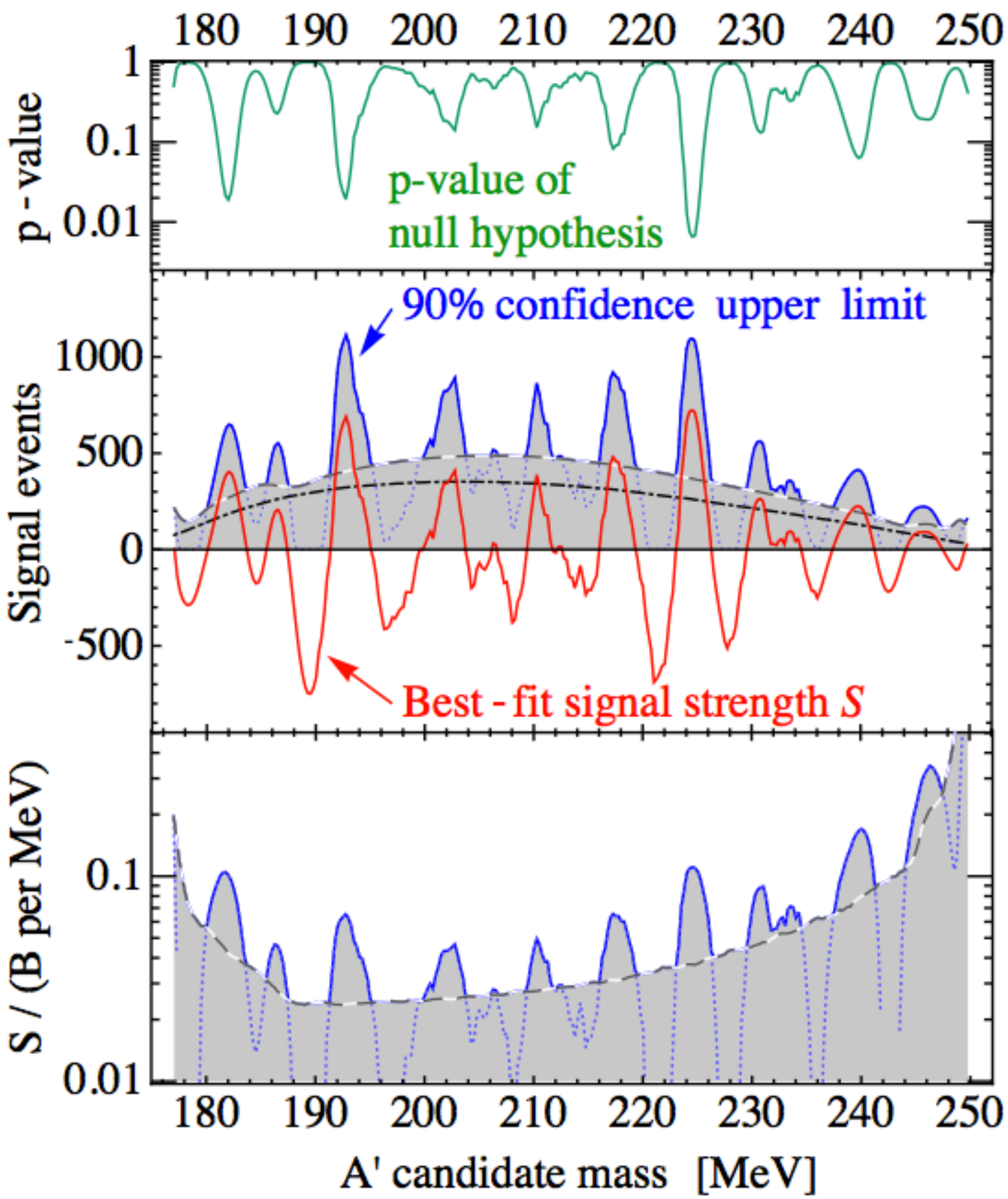
Results from scan of test run data: S, P-values

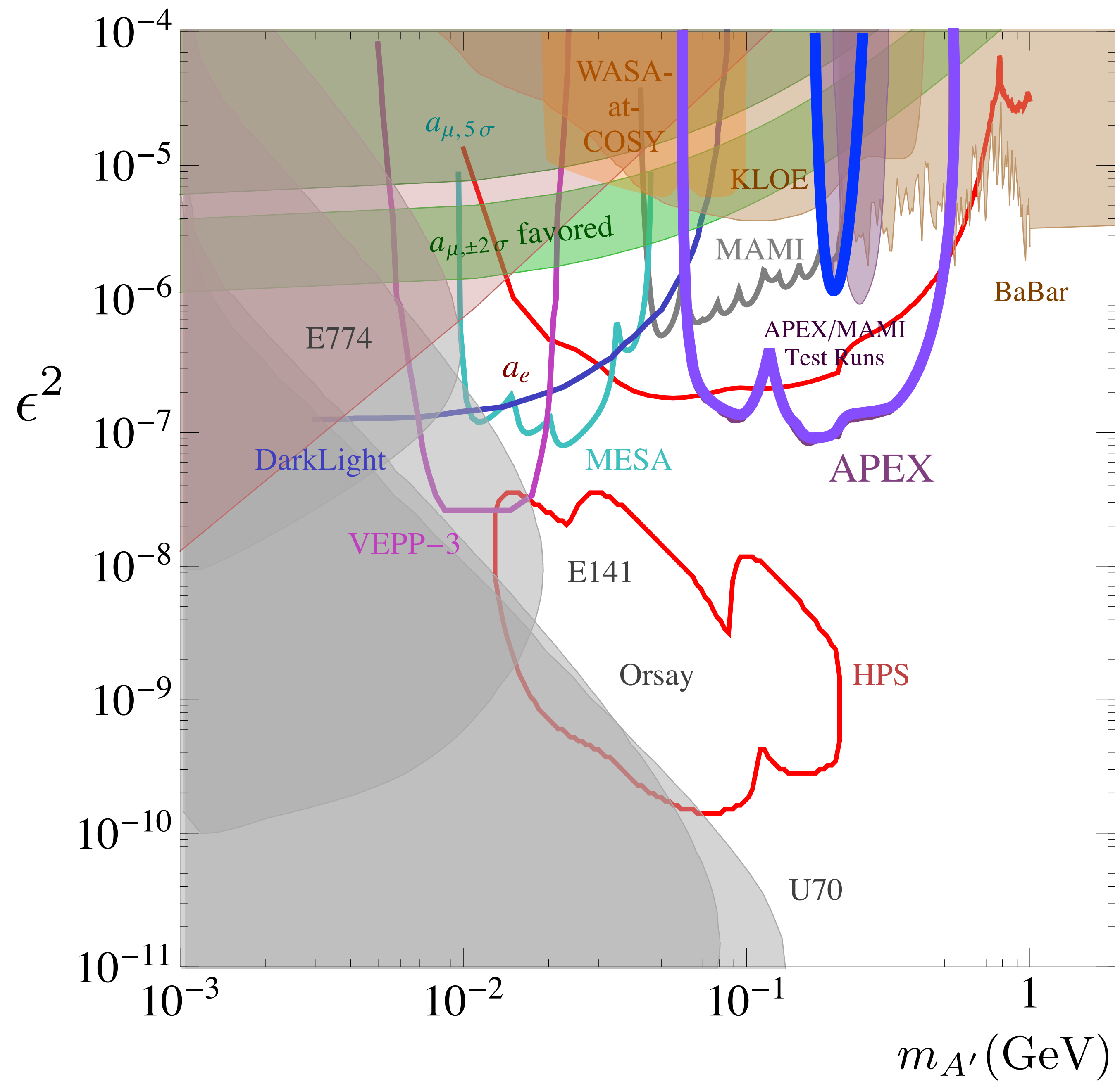




Two-sided central limit; plot in paper is constrained to physical values of S

Upper limit on coupling





Plans for full run

Look closer at unbinned analysis

- Should be possible to address computational issues

Compare multiple methods for binned analysis

Investigate mass-dependent window sizes

Full MC needed to complement resonance search

End Part 1

In progress: Update of reach estimate for full run

Proposal-style reach calculation, with the inclusion of new LeRose acceptance

- Monte Carlo events are simulated for the Bethe-Heitler, radiative tridents, and the continuum trident background including the full interference effects between the diagrams. The latter background is computationally intensive, and only a small statistics sample is generated, sufficient to obtain the cross-section from MadEvent.
- The cross-section ratio of the full continuum background (with interference effects) to the sum of the Bethe-Heitler and radiative tridents is calculated, and represents a multiplicative factor by which the latter must be multiplied to get the background cross-section.
- The rates of all reactions impinging the spectrometer acceptance were calculated by integrating over a chosen target profile, which usually extended from 4.5 to 5.5 degrees. For Bethe-Heitler, radiative tridents, and the continuum trident background, the calculation of the rate was performed as a function of invariant mass.
- Using the expressions in Appendix B, we calculated the mass resolution δ_m. We then tiled the acceptance region with bins of size $2.5 \times \delta_m$ in invariant mass.
- As a function of α'/α, the total number of signal (S) and background (B) events was calculated with the help of the formulas in §3 for each bin.
- We then set $S/\sqrt{B} = 2$, and solved for α'/α.

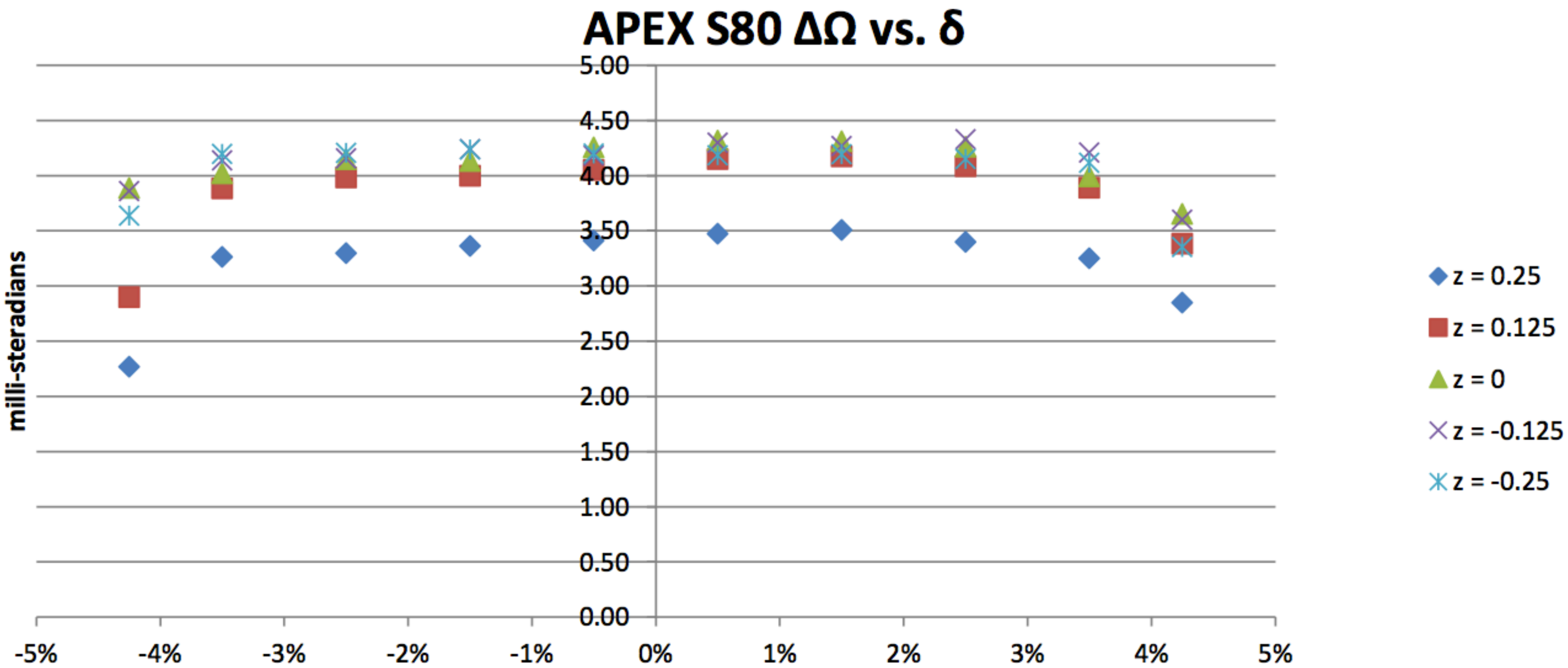
APEX PAC37

Changes from this prescription:

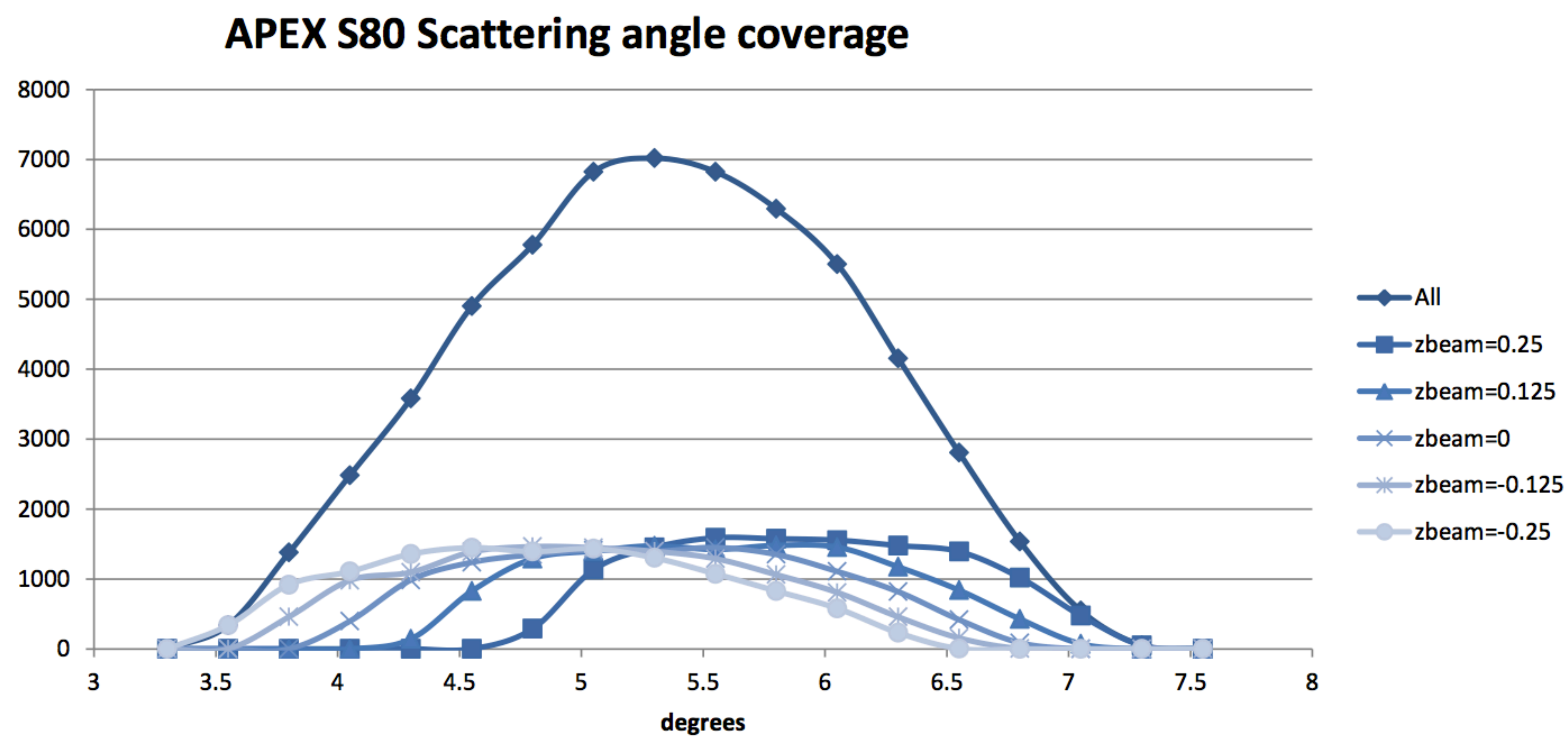
- Monte Carlo for radiative trident and Bethe-Heitler backgrounds, generated at four beam energies (not necessarily those being used for full run), including correction factors for interference effects
- Ten foils/angular settings between 4.5 and 5.5 degrees ⟶ Utilize MC multiple times
- Each "run" gives an unbinned dataset of invariant masses; tile the spectrum in bins of 2.5*(mass resolution) and calculate 2σ reach in ϵ^2
- New here is the imposition of the LeRose S80 acceptance numbers
- The following is pre-preliminary

Center for Cosmology and Particle Physics

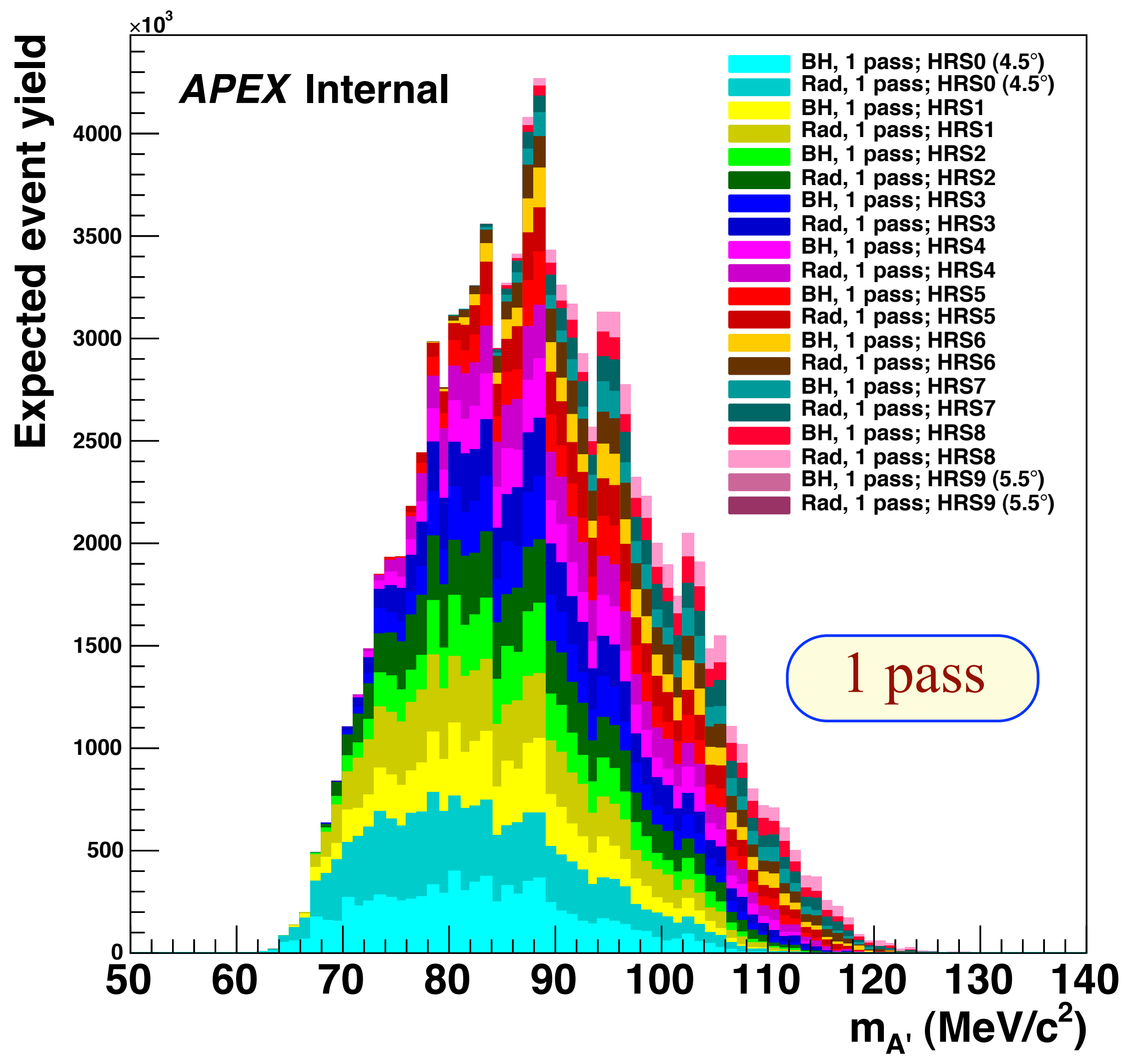
In progress: Update of reach estimate for full run

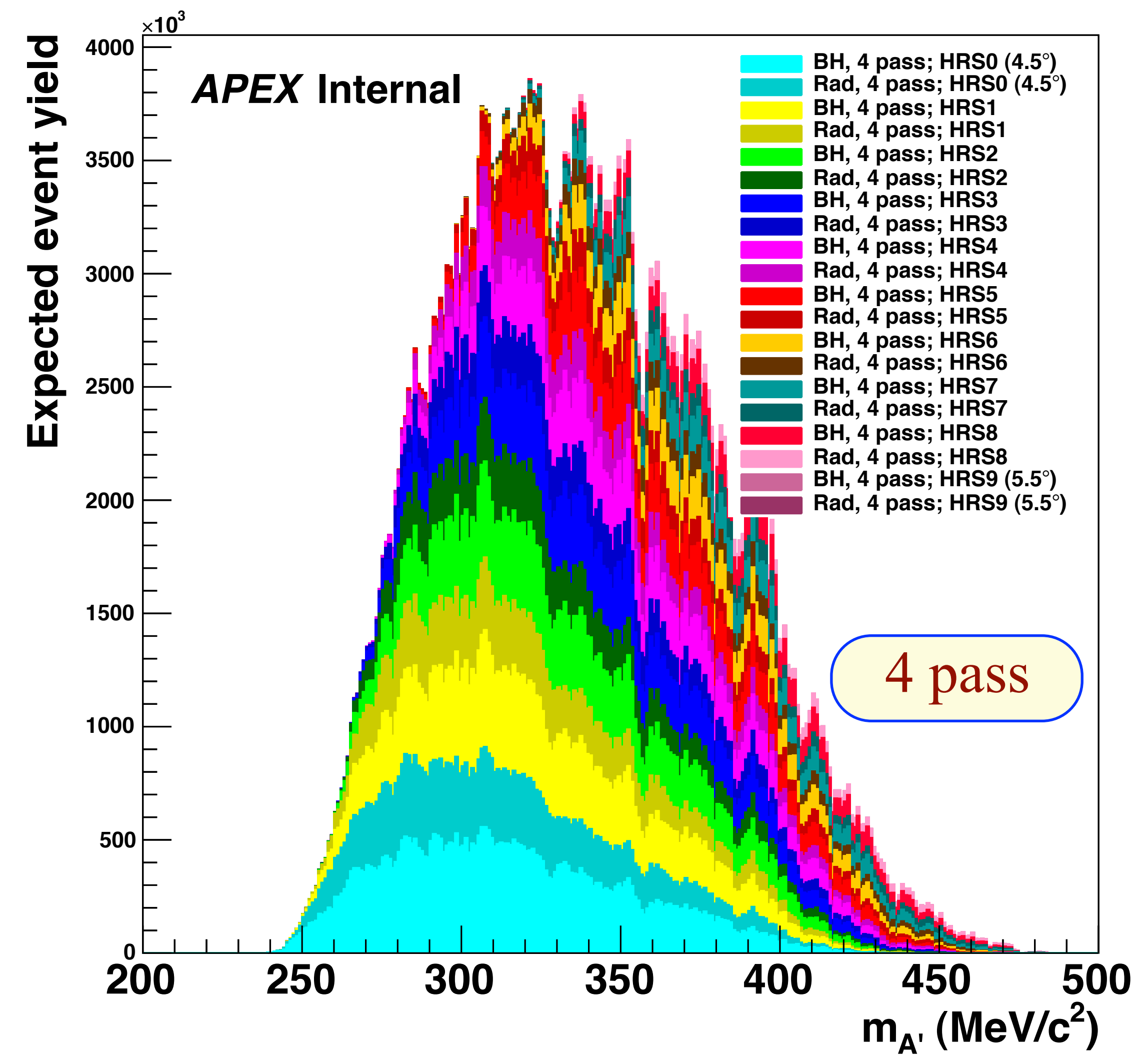


New acceptance for septum magnets calculated by John LeRose



In progress: Update of reach estimate for full run



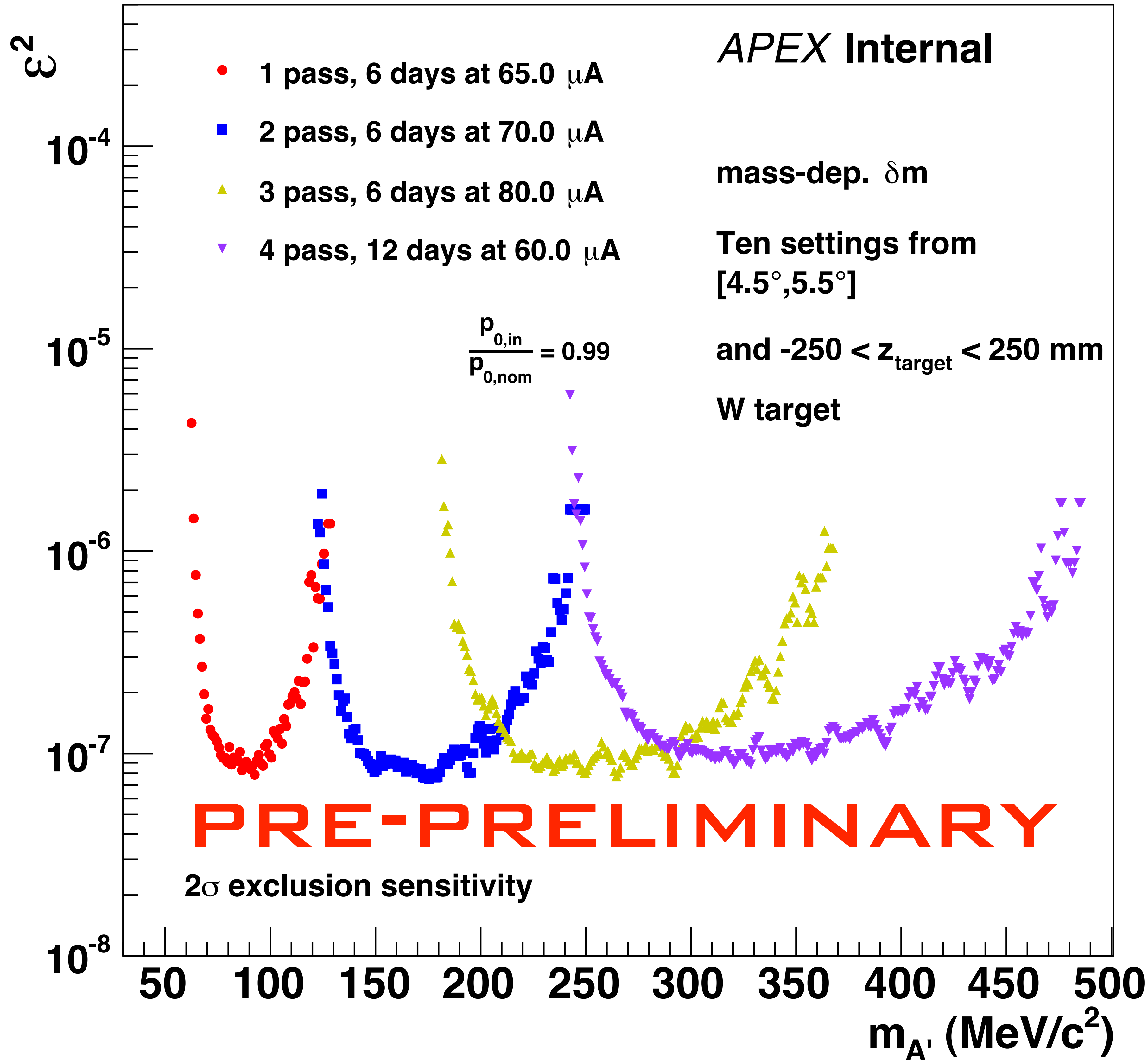


Monte Carlo for radiative trident and Bethe-Heitler backgrounds

E_{beam}	$p_0 = 0.99*E_{beam}/2$	Time [days]	Current [μA]	T [rad. lengths]
1.056	0.523	6	65	0.007
2.056	1.018	6	70	0.04
3.056	1.513	6	80	0.08
4.056	2.008	12	60	0.08

CENTER FOR COSMOLOGY AND PARTICLE PHYSICS

In progress: Update of reach estimate for full run



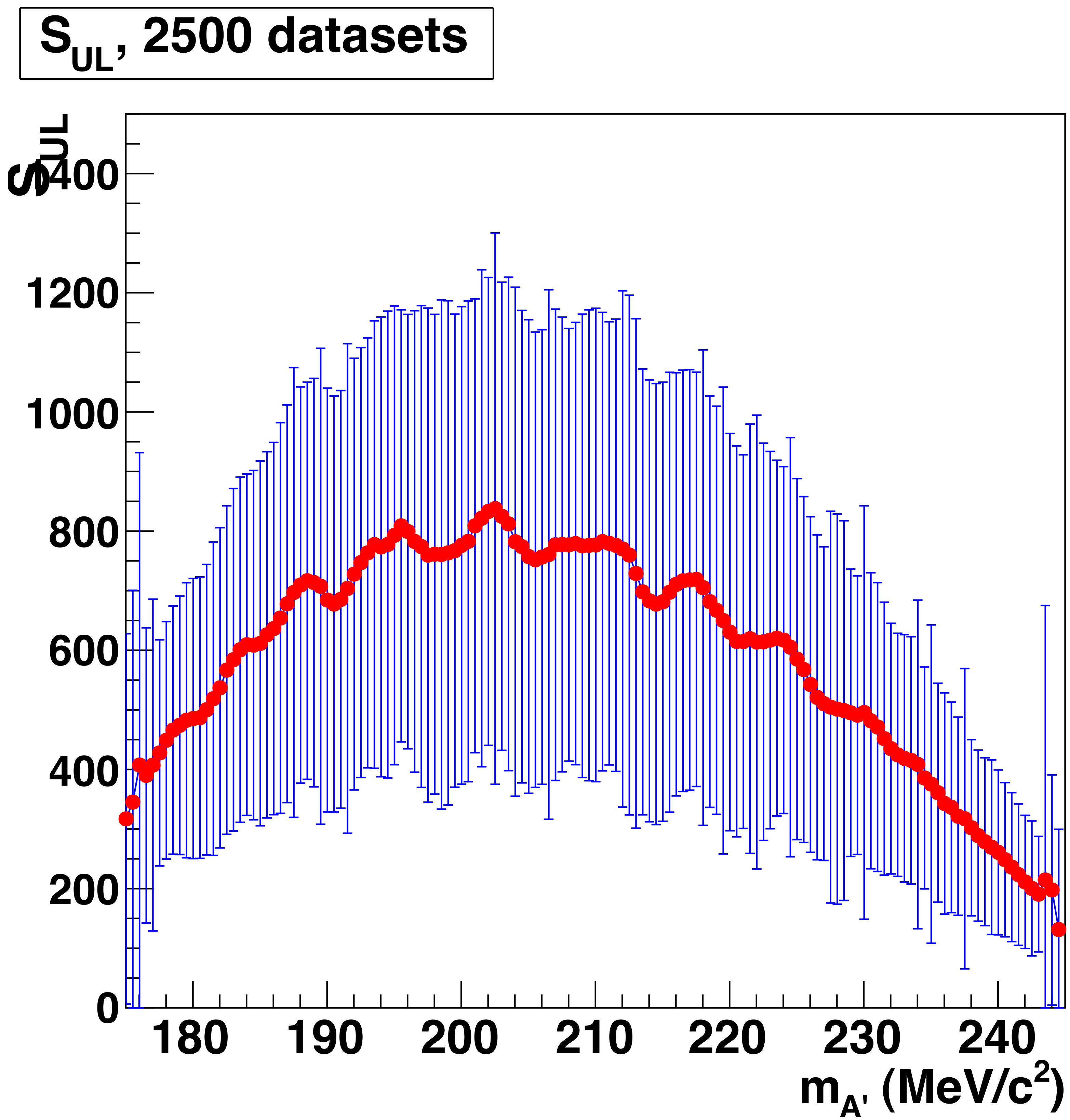
Proposal-style reach calculation with LeRose S80 acceptance numbers

STILL VALIDATING; WILL CHANGE; MACHINERY IN PLACE

End

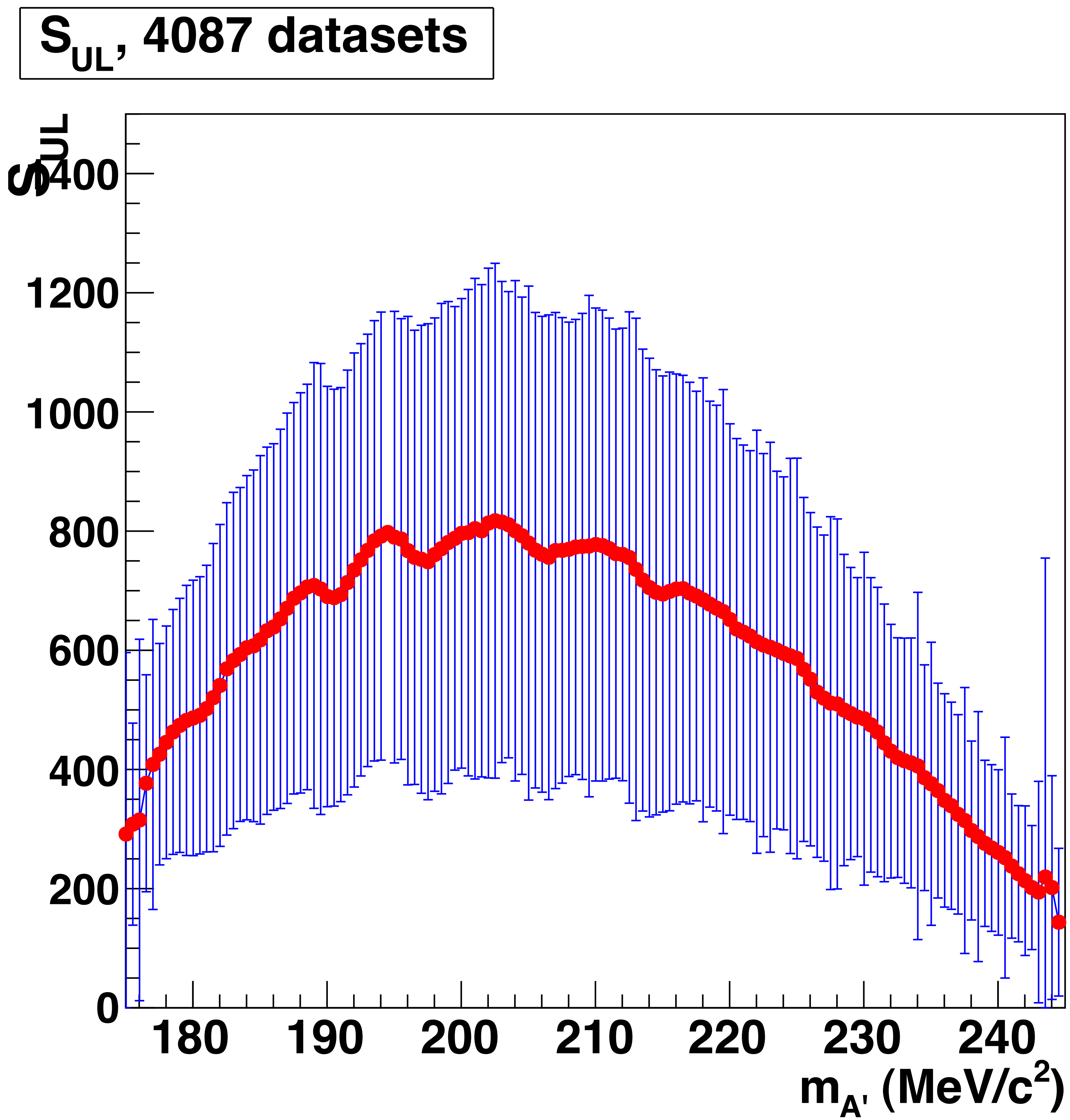
Backups

Asimov data is the limit of an ensemble

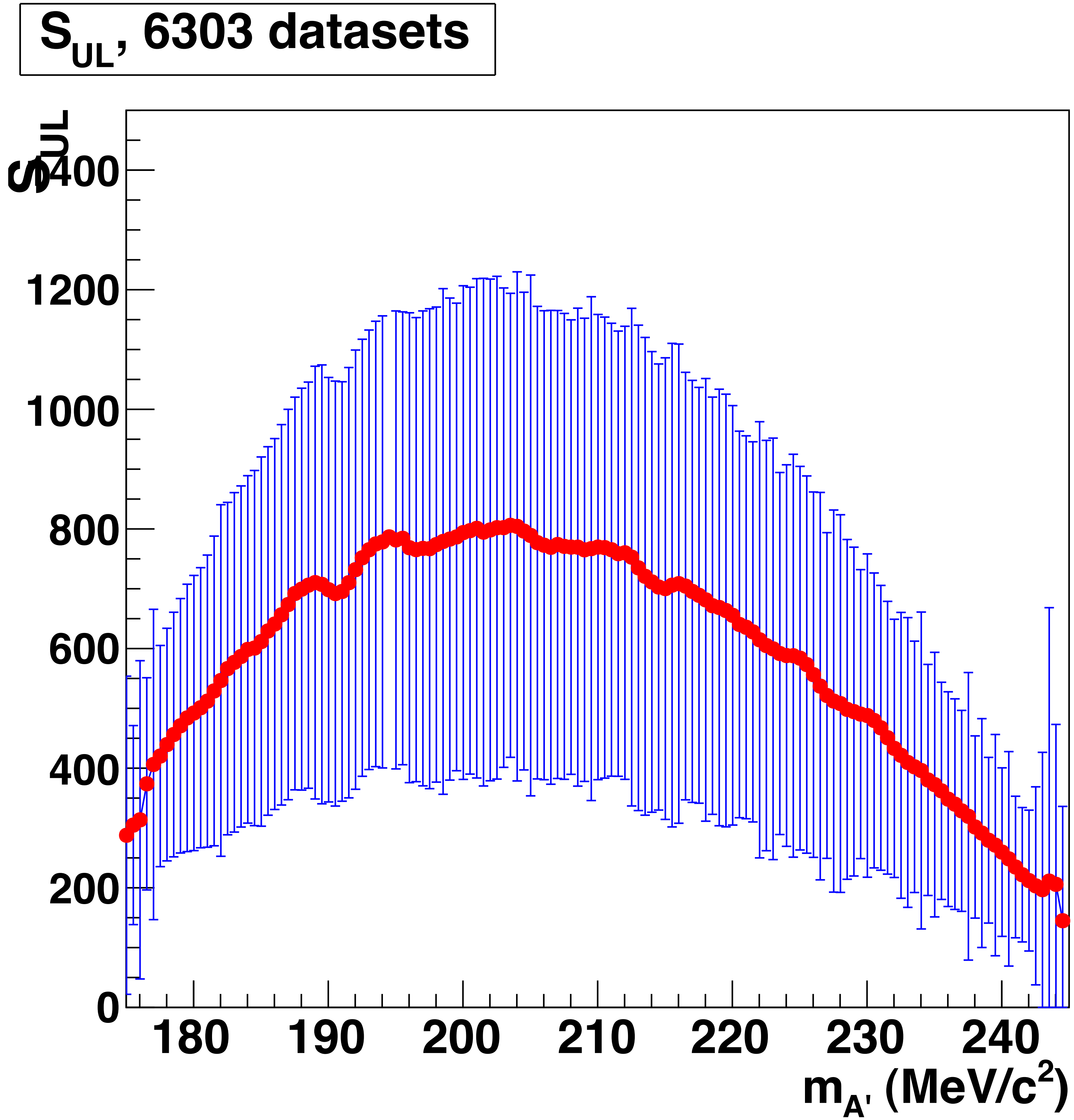


Center for Cosmology and Particle Physics

Asimov data is the limit of an ensemble



Asimov data is the limit of an ensemble



Center for Cosmology and Particle Physics

Asimov data is the limit of an ensemble