HPS Collaboration Meeting
10/19/2026

Towards unblinding, part 3

Emrys Peets

Stanford University
SLAC National Accelerator Laboratory

U.S. DEPARTMENT OF ENERGY
Office of Science

HPS

Jefferson Lab

Stanford University

SLAC NATIONAL ACCELERATOR LABORATORY

Outline

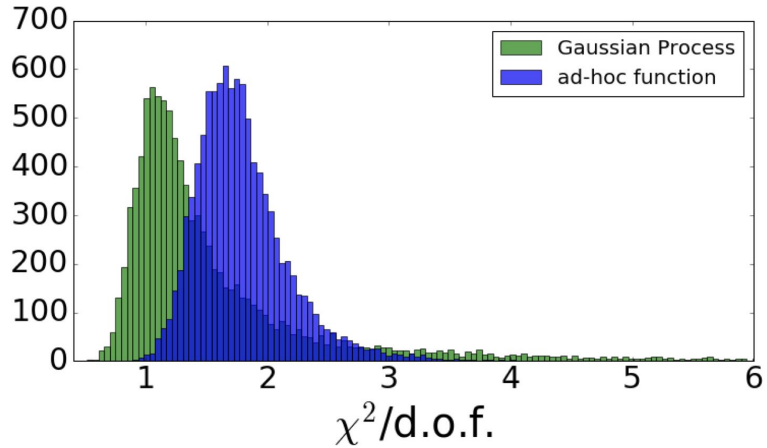
I. Literature review
II. Signal Modeling
III. Background Modeling
    A. Polynomial-mean GP background model
    B. Calibration, independent response checks and full luminosity statistics
IV. 2021 10% results and extraction examples
V. APEX and X17 Application
VI. Unblinding Review and next steps

Gaussian processes in resonance searches

SLAC

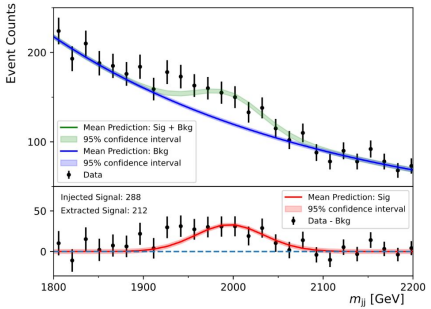
Modeling Smooth Backgrounds & Generic Localized Signals with Gaussian Processes

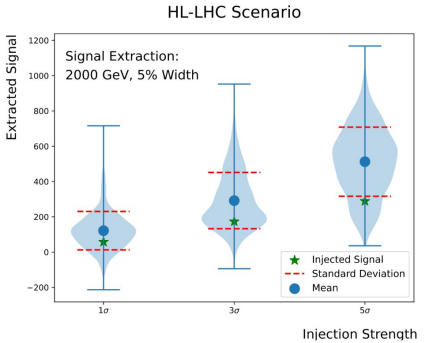
- Compared GPR with ad-hoc background forms



Gaussian Process Regression as a Sustainable Data-driven Background Estimate Method at the (HL)-LHC

- regularization used to combat fit instability and improve background only modelling





CMS Signal Extraction Study

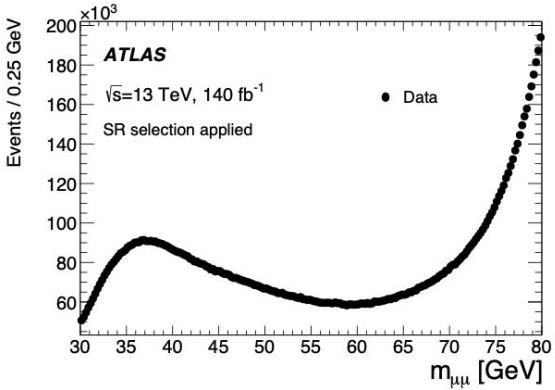
ATLAS dimuon search with a polynomial GP mean

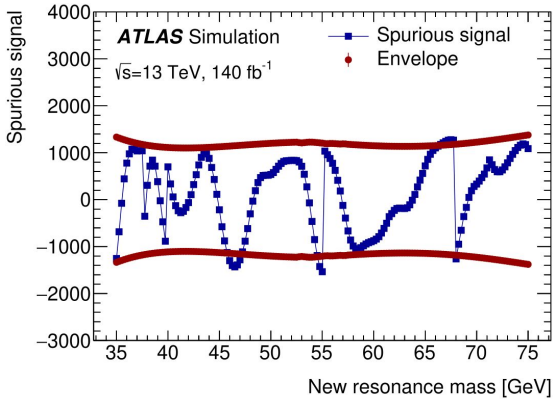
Ninth-degree polynomial prior mean from simulation

RBF covariance models the residual structure

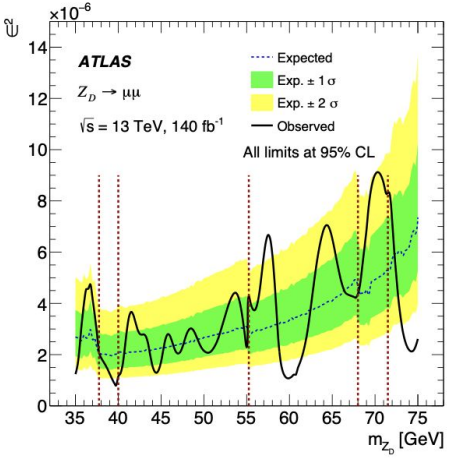
Hyperparameters selected on reweighted simulation

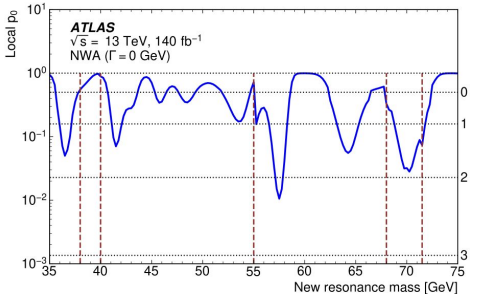
Spurious-signal and injected-yield tests





m_X [GeV] range	k	ℓ [GeV]
$35 \le m_X < 38$	4×10^9	4.4
$38 \le m_X < 40$	2×10^8	8.0
$40 \le m_X < 55$	2×10^8	7.1
$55 \le m_X < 68$	2×10^8	8.7
$68 \le m_X < 71.5$	5×10^8	12.1
$71.5 \le m_X < 75$	5×10^8	11.4





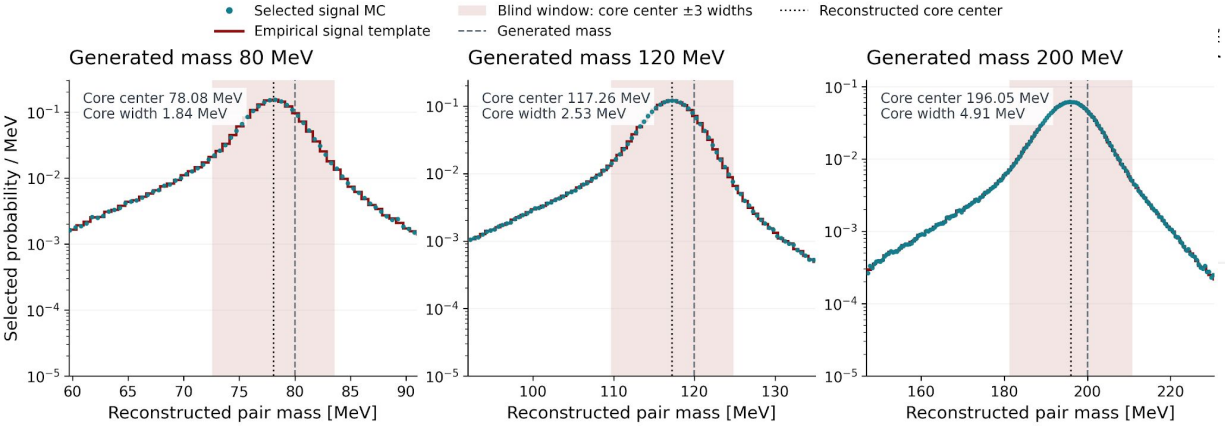
ATLAS Collaboration, JHEP 06 (2026) 165

Search for dimuon resonance in the 35 to 75 GeV mass range using 140 fb^{-1} of 13 TeV pp collisions with the ATLAS detector

Signal Modeling in HPS

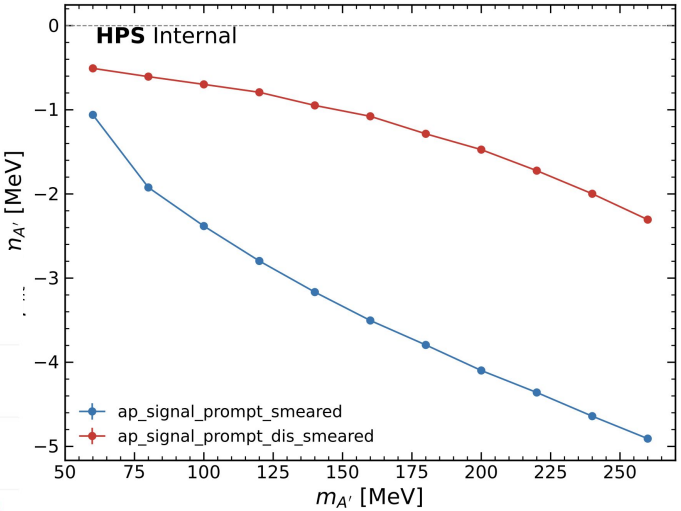
Matthew showed that signal MC shapes have improved, though offsets remain.

- we thus blind an offset mass center, and use MC shape normalized
- signal morphing in extra slides



MC peak displacement

$$c(m) = m - 3.2243308693 - 2.2139928114\ln(m/150)$$



Signal Modeling in HPS

Calibrated signal in the Poisson mean

$$\lambda_i = b_{0,\,i} + (L\theta)_i + a\,R\,S_i(m)$$

$$S_i(m) = \int_{\text{bin } i} p_{\text{MC}}(x \mid m)\,dx$$

MC shape normalization

Coupling conversion

$$\epsilon^2_{\text{up}} = \frac{2\alpha N_f N^{\text{up}}_{\text{sig}}}{3\pi m_{A'} f_{\text{rad}} (dN_{\text{bkg}}/dm)}$$

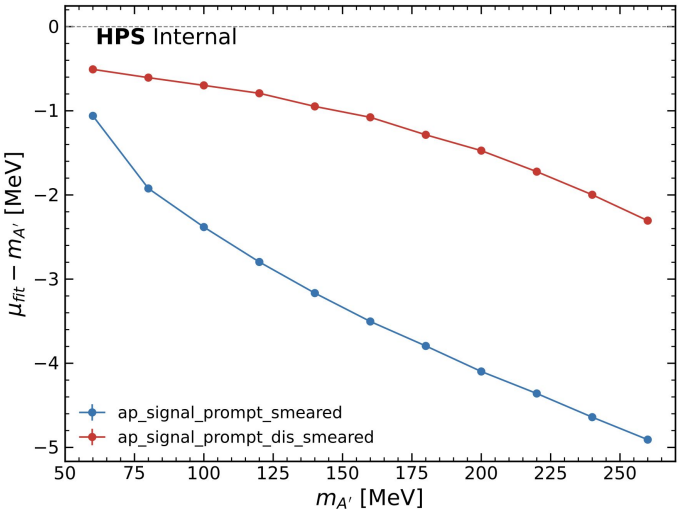
$N_f = 1 \quad (m < 2m_\mu)$

$N_f = 1 + \sqrt{1-4r}\,(1+2r) \quad (m \geq 2m_\mu)$

$r = (m_\mu/m)^2$

MC peak displacement

$$c(m) = m - 3.2243308693 - 2.2139928114\ln(m/150)$$



Background modeling in new hybrid analysis

Fit the sidebands
Learn β; predict b_{SB}
$C_{eff} = LL^T$

→

Test a signal yield
$\lambda = b_0 + L\theta + aT$
Hold b_0, L and T fixed

→

Profile θ
Minimize Poisson NLL
$+ \frac{1}{2}\theta^T\theta$

→

Build the profile
Repeat at each a
Form q(a) and CLs

Local Poisson mean

$$\lambda(a, \theta) = b_0 + L\theta + aT$$

$b_0 = b_{SB} + \delta S$
sideband reference + offset

correlated background
Gaussian constraint

$T = R\,S$
signal response × template

Constrained local likelihood

$$\mathcal{L}(a, \theta) = [\Pi_i \, \mathrm{Pois}(n_i \mid \lambda_i)] \exp(-\tfrac{1}{2}\theta^T\theta)$$

HPS-GPR Hybrid Analysis

To calibrate meaningfully using a functional form, it should be imbedded into the gaussian process. Have already been using this in X17 and phi→K+K- analyses.

1 Fit a polynomial reference as GP mean

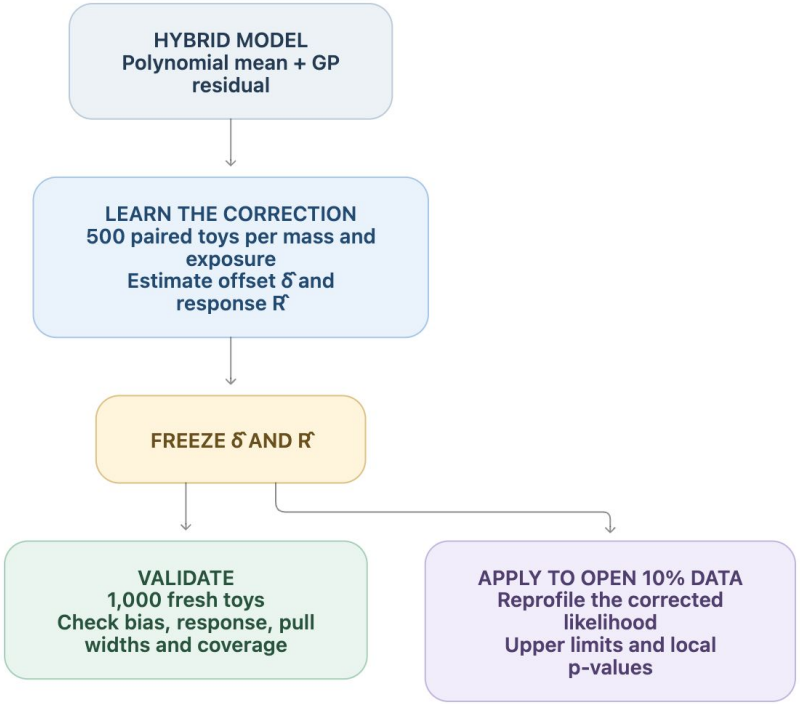
Fit the broad spectrum; freeze the source and model settings.

2 Calibrate and confirm

For each mass, injected toys determine δ and R; fresh toys check fitted errors.

3 Fit the sidebands, profile bkg, extract

Refit the polynomial mean at each mass; condition the GP. Predict across the blind window; extract corrected yields and limits.



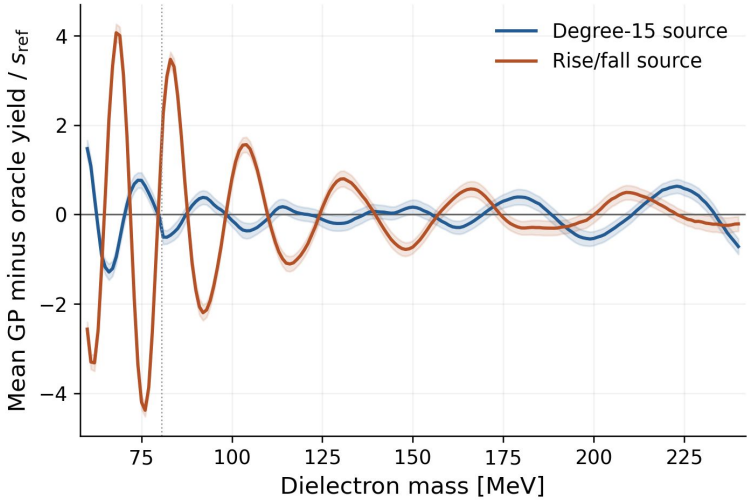
$$\text{Corrected yield} = \frac{\text{Fitted yield} - \hat{\delta}}{\hat{R}} \qquad \text{Corrected error} = \frac{\text{Fitted error}}{\hat{R}}$$

HPS-GPR Hybrid Analysis

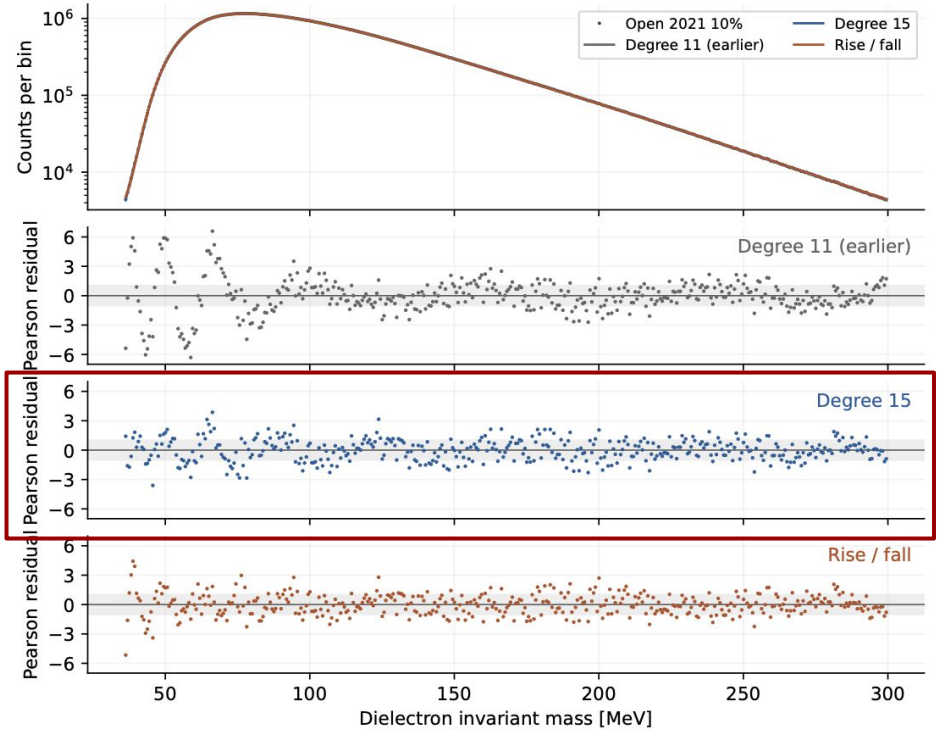
To calibrate meaningfully using a functional form, it should be imbedded into the gaussian process.

1 Fit a polynomial reference as GP mean

Fit the broad spectrum; freeze the source and model settings.



visually good fit not necessarily the best

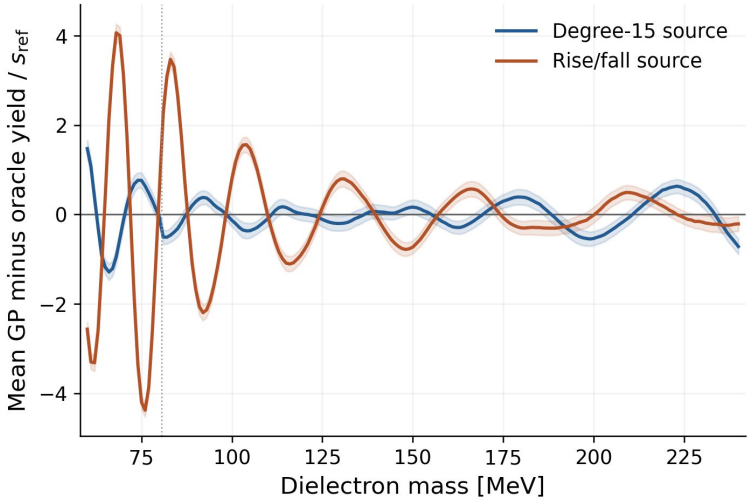


HPS-GPR Hybrid Analysis

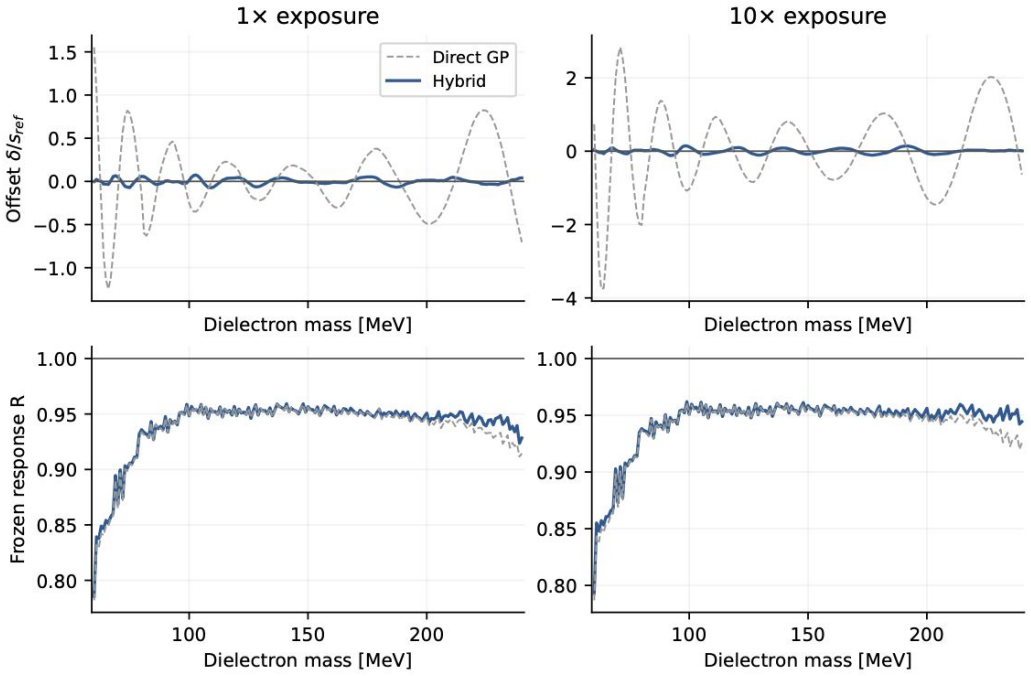
To calibrate meaningfully using a functional form, it should be imbedded into the gaussian process.

1 Fit a polynomial reference as GP mean

Fit the broad spectrum; freeze the source and model settings.



hybrid drastically improves offset, not signal absorption.
direct GP bias degrades significantly with more statistics



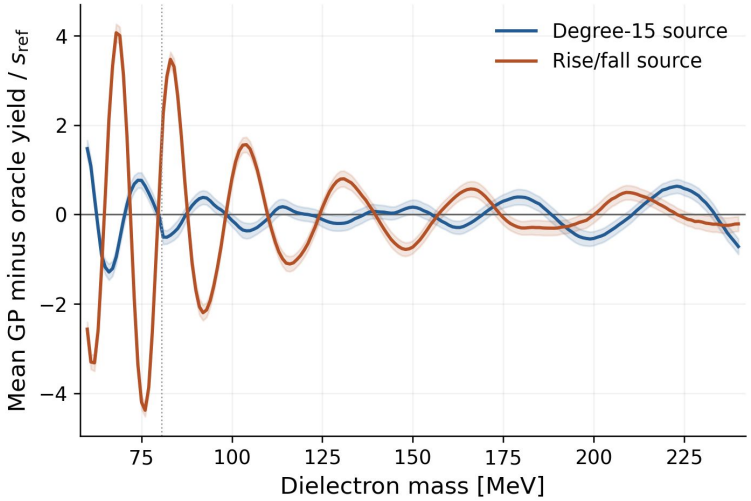
HPS-GPR Hybrid Analysis

To calibrate meaningfully using a functional form, it should be imbedded into the gaussian process.

1 Fit a polynomial reference as GP mean

Fit the broad spectrum; freeze the source and model settings.

$$f(x) = p_{15}(x;\beta) + g(x)$$
$$g \sim \mathcal{GP}(0,k), \quad x = \log(m/\mathrm{GeV})$$
$$k(x,x') = C\exp\left[-\frac{(x-x')^2}{2\ell^2}\right]$$



Test-mass region	C	ℓ
60–80 MeV (rising)	105.877	0.280784
81–240 MeV (falling)	100.994	0.273196

Frozen hyperparameters using separate calibration toys from each statistical sample.

HPS-GPR Hybrid Analysis

To calibrate meaningfully using a functional form, it should be imbedded into the gaussian process.

2 Calibrate and confirm

For each mass, injected toys determine δ and R; fresh toys check fitted errors.

Calibration Procedure

use 500 paired calibration toys (0 and $3s_{ref}$) to determine offset and signal response

verify that applying correction to 1000 independently drawn toys extracts appropriate signal yields is not significantly offset, and has appropriate pull widths

offset

$$\delta = \langle \hat{A}_0 \rangle$$

signal response

$$R = \frac{\langle \hat{A}_{\text{inj}} - \hat{A}_0 \rangle}{A_{\text{inj}}}$$

corrections applied

$$\hat{a} = \frac{\hat{A} - \delta}{R} \qquad \sigma_a = \frac{\sigma_A}{R}$$

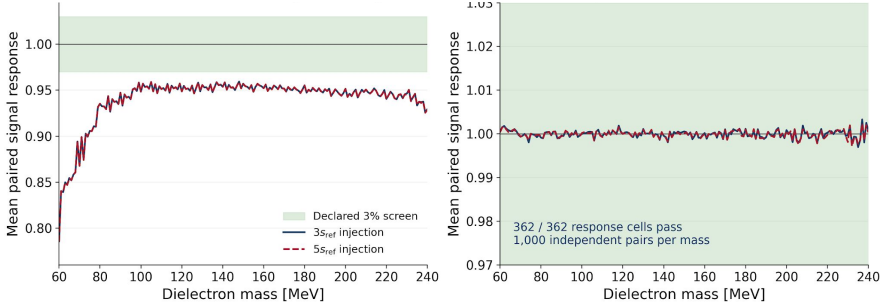
HPS-GPR Hybrid Analysis

To calibrate meaningfully using a functional form, it should be imbedded into the gaussian process.

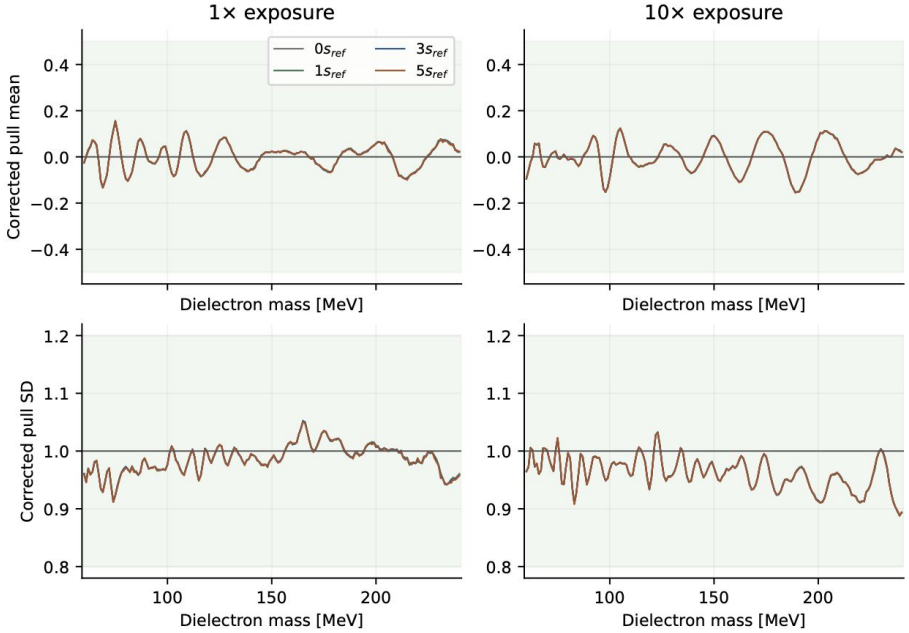
2 Calibrate and confirm

For each mass, injected toys determine δ and R;
fresh toys check fitted errors.

Signal Response Improvement (10% Statistics)



Calibration Results (1000 "Fresh Toys")



HPS-GPR Hybrid Analysis

To calibrate meaningfully using a functional form, it should be imbedded into the gaussian process.

3 Fit the sidebands, profile bkg, extract

Refit the polynomial mean at each mass; condition the GP. Predict across the blind window; extract corrected yields and limits.

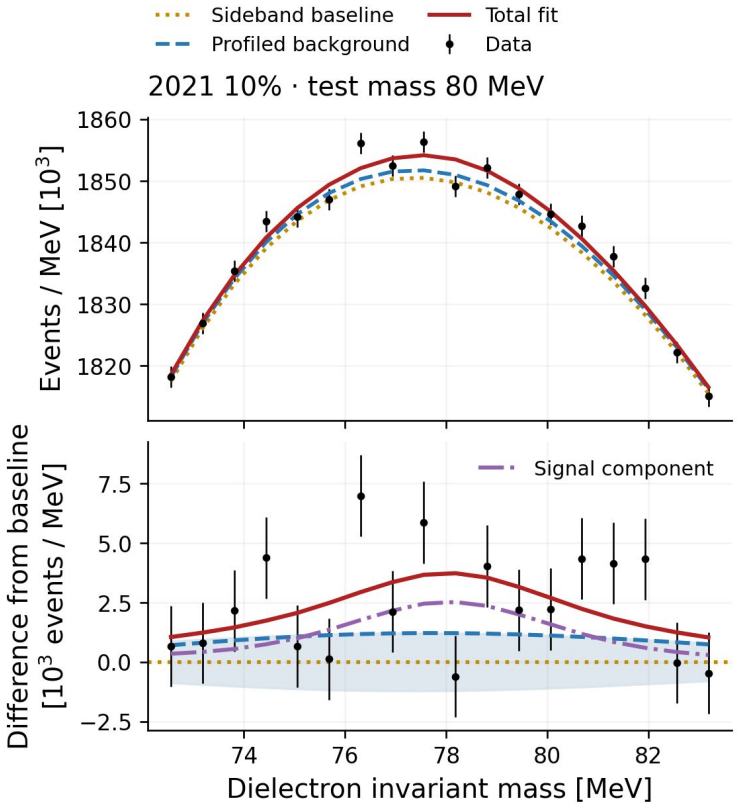
Fixed sideband reference

$b_0 = b_{\mathrm{SB}} + \delta S$

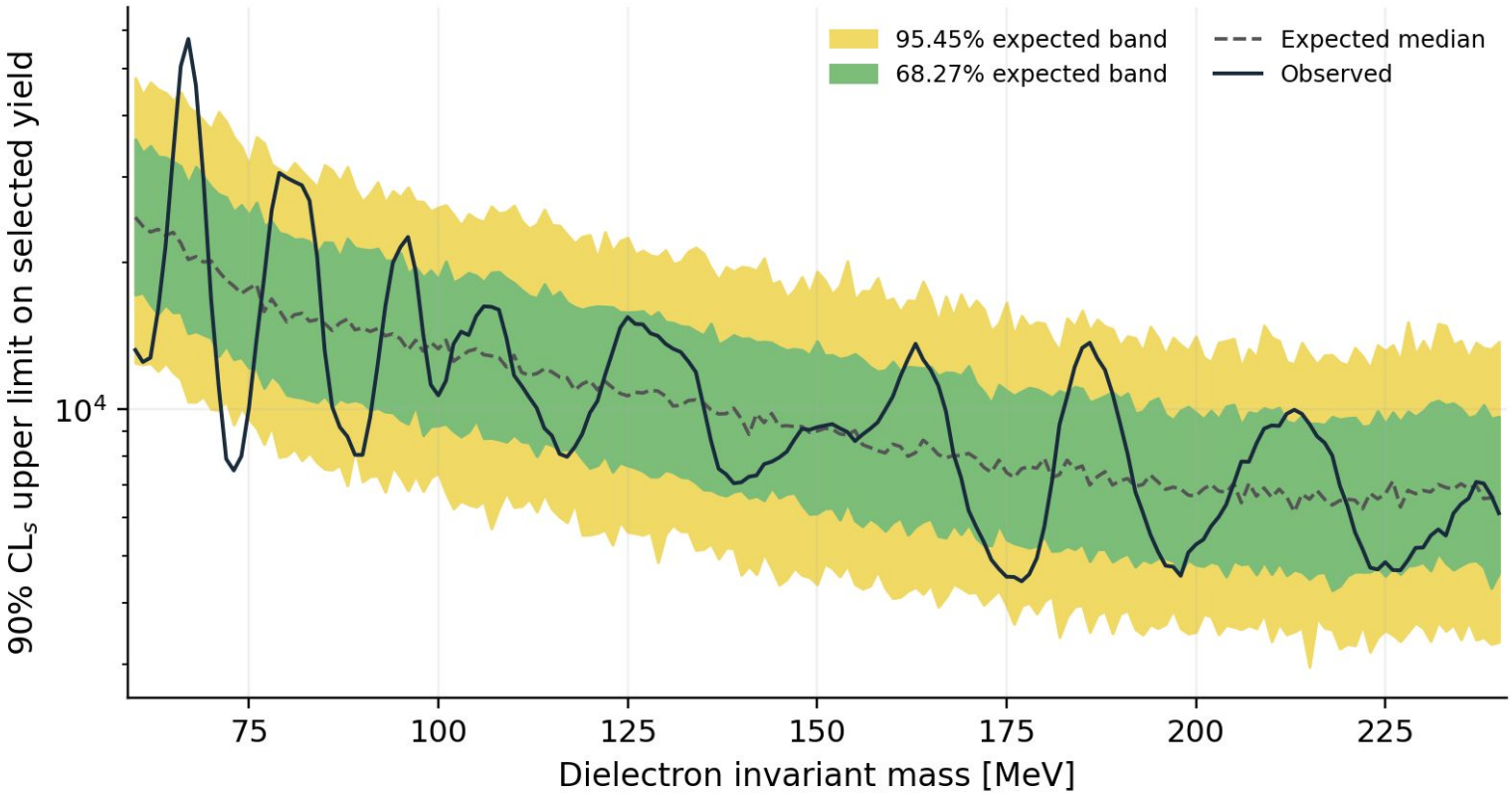
Coherent background change
θ profiled for each signal yield $L\theta$

Calibrated signal component

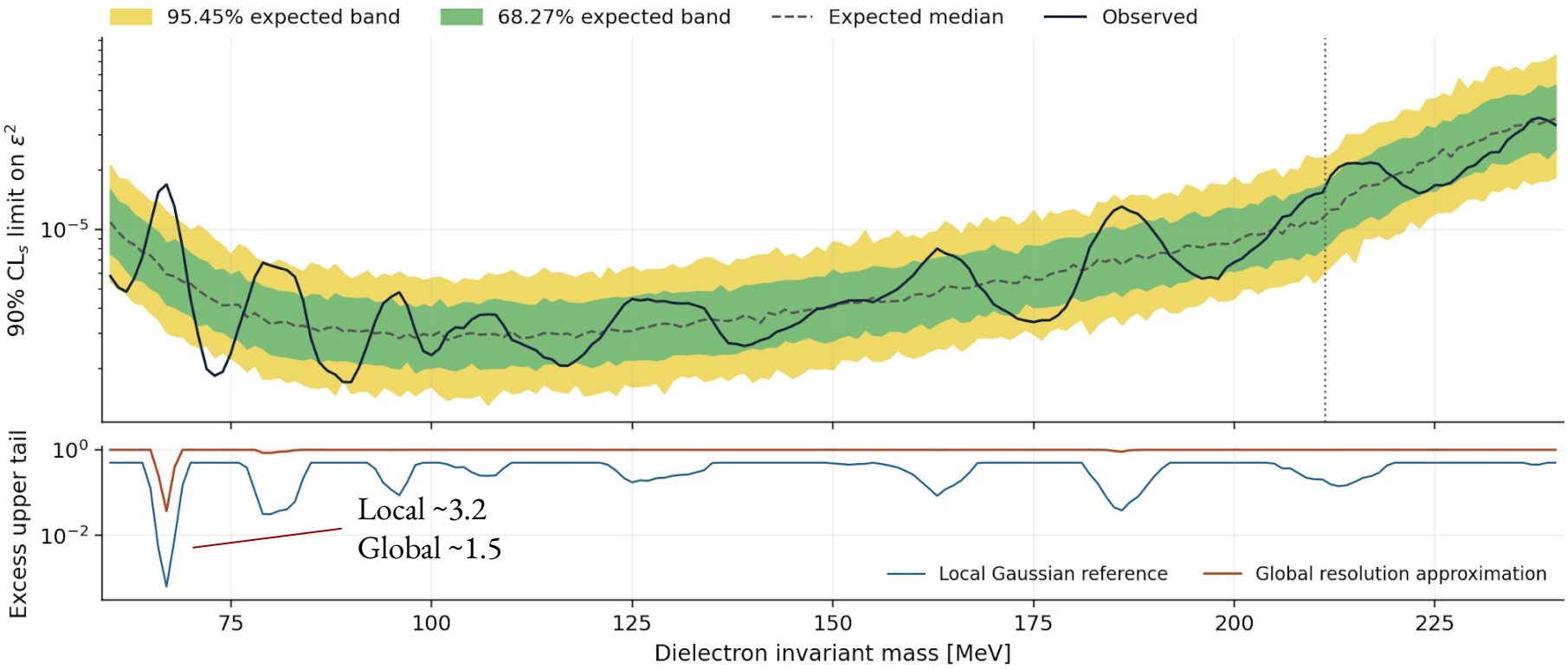
$a\,R\,S$



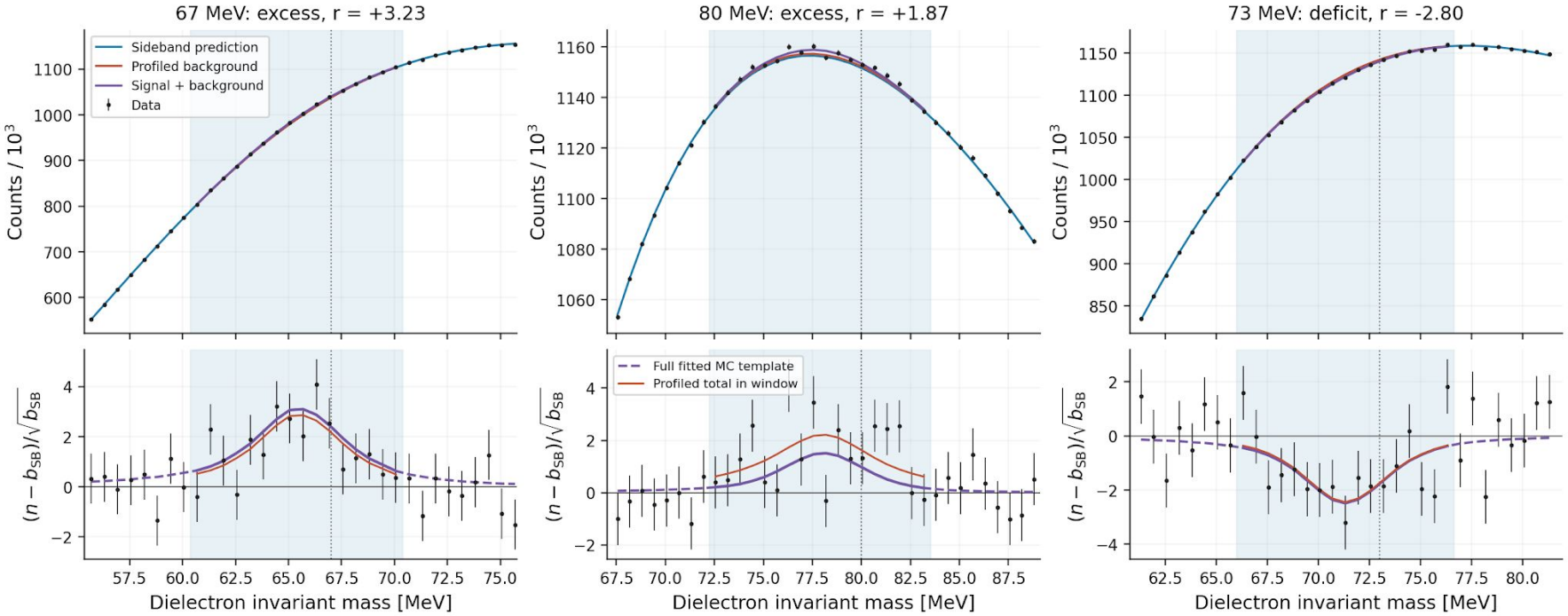
2021 10% signal yield upper limit result



2021 10% hybrid coupling result

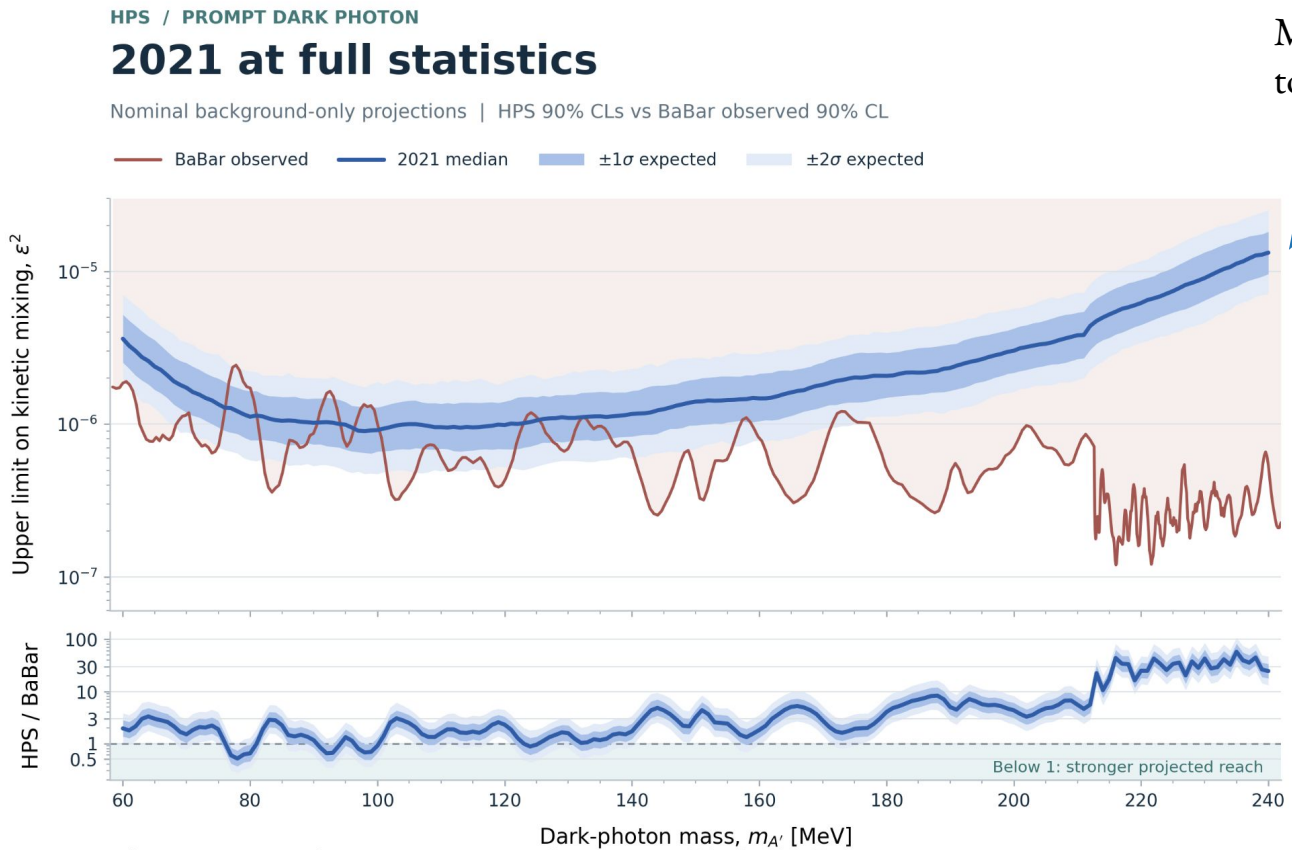


2021 10% extraction examples



Shaded region: ±3-width blind window. Dashed tails show the full MC template beyond the fit window.

Full statistics expectation



Median based on 10% *10 toy ensemble

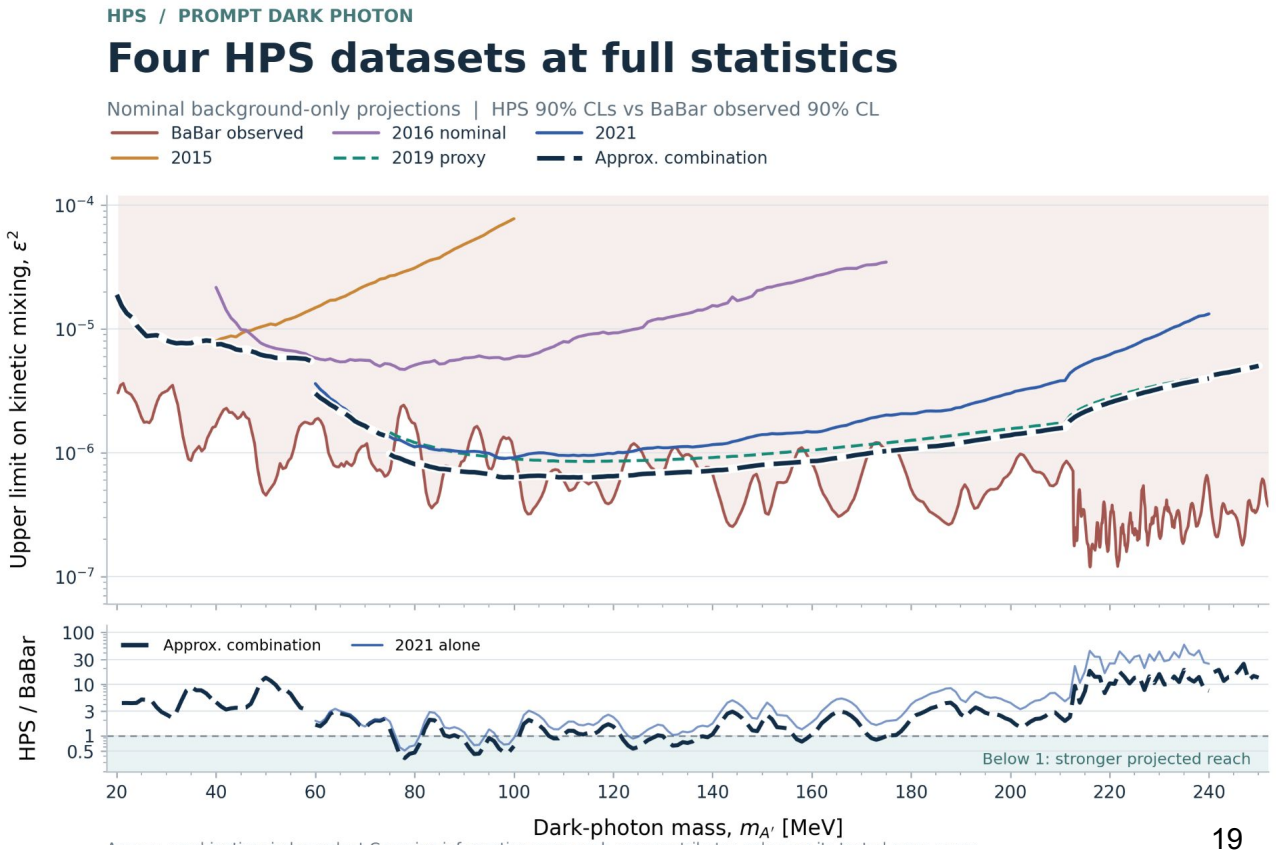
Babar (mid)

All HPS* Reach

Simultaneous reach improves.

Compared to Babar, at 78 MeV, projected ε^2 limits are **1.9 × lower for 2021** and **2.7× lower combined** .

→ Sensitivity to true signal and simultaneous fits is powerful with individual datasets.

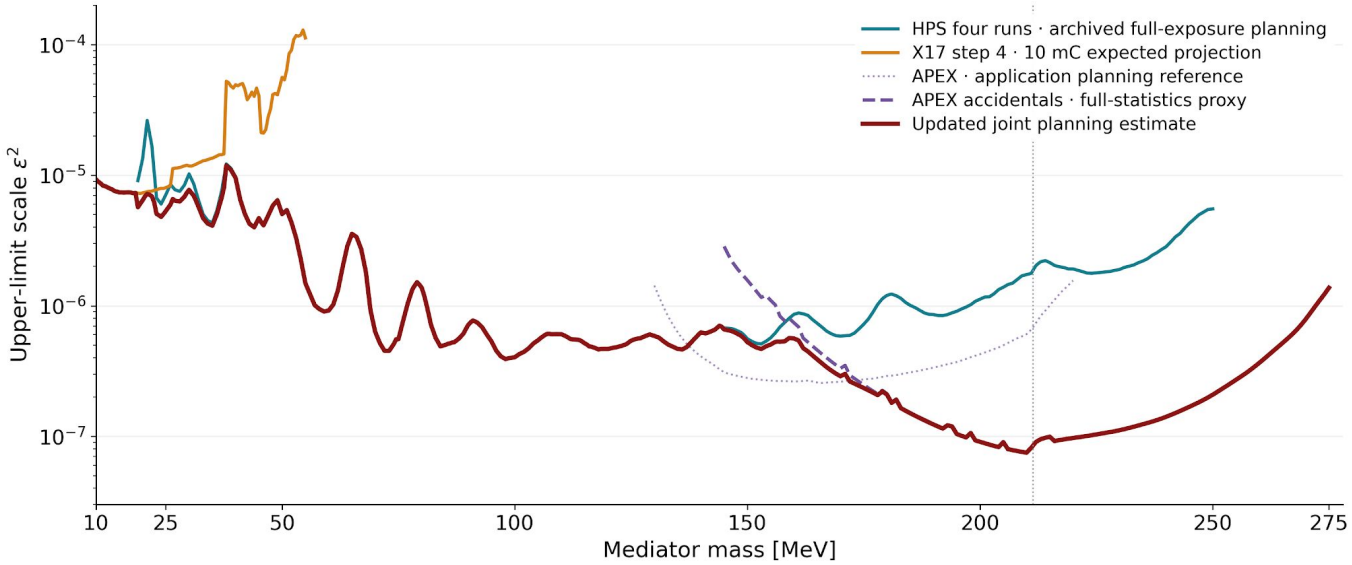


HPS, APEX and X17 projected reach (2-3ish year plan)

Have joined X17 and APEX

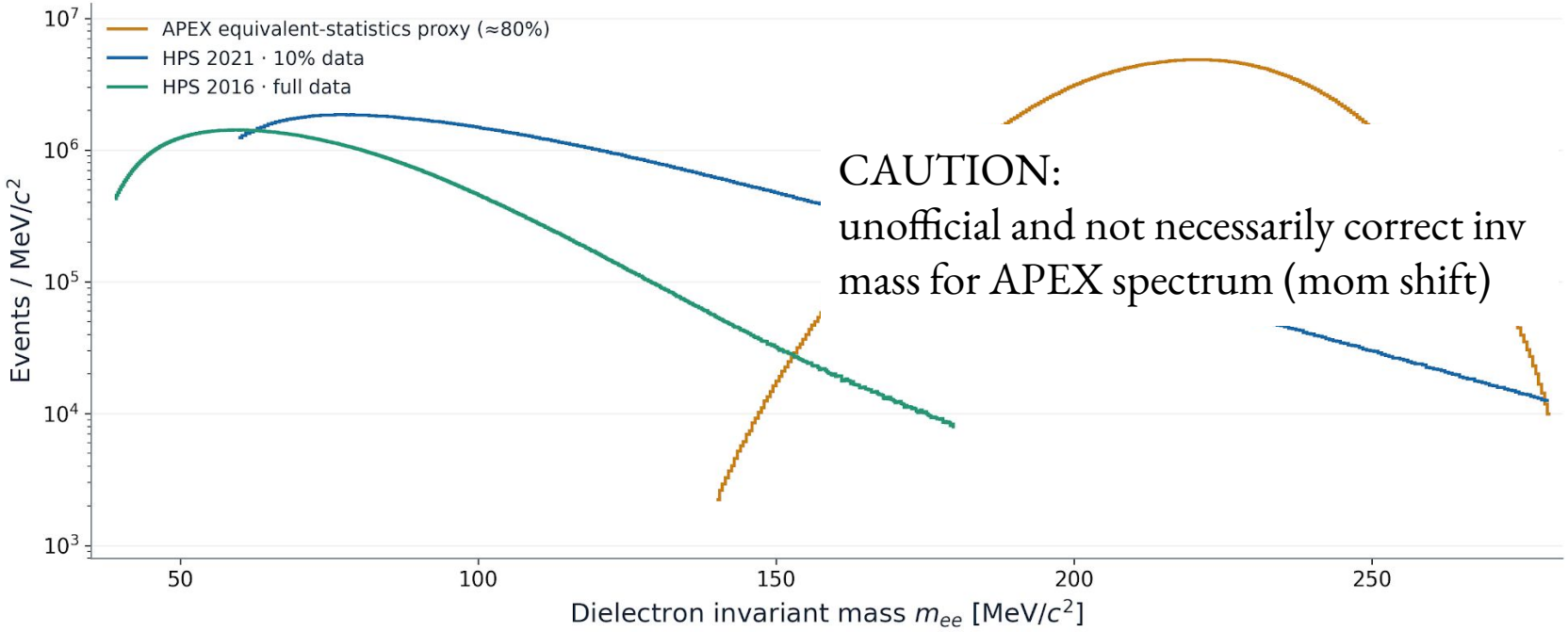
Working with these analysts will help push GP to a standard and accepted form.

Will perform joint searches with HPS and APEX at high mass, and X17 at low mass.



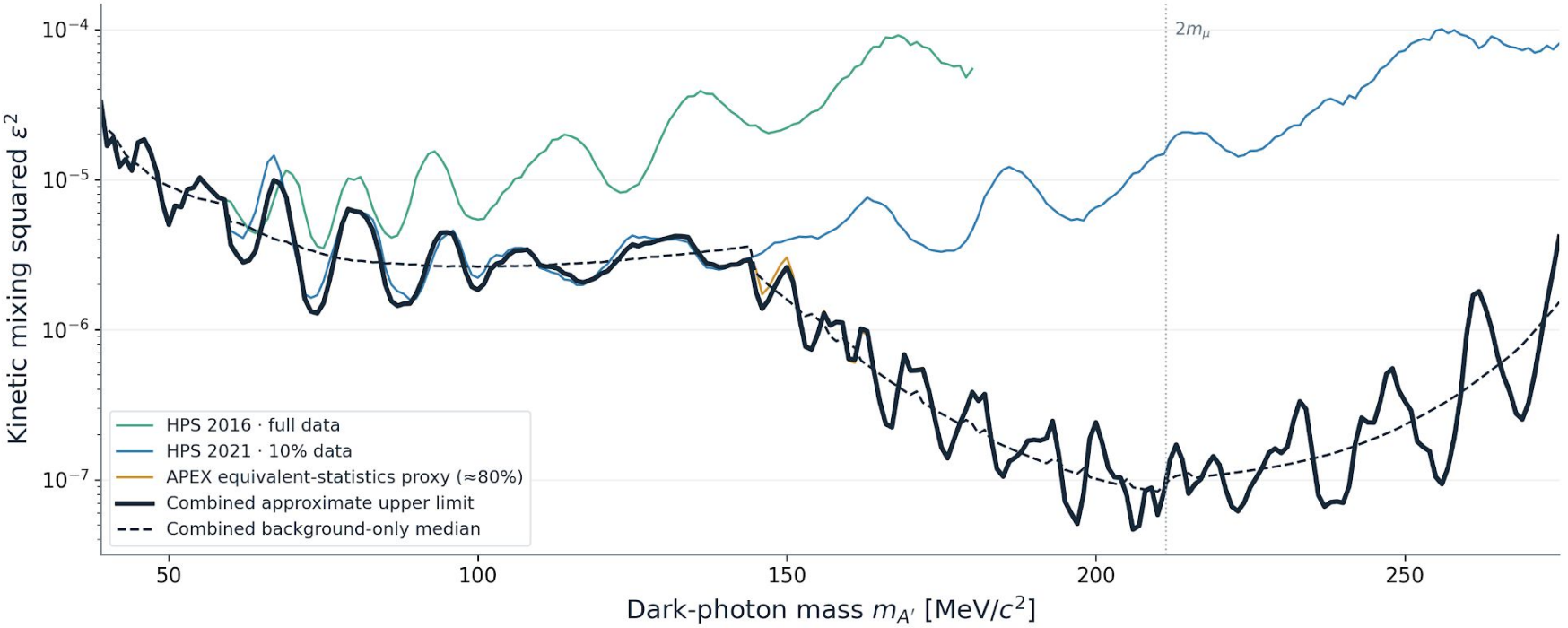
APEX accidental pairs as a background proof of principle

Accidentals: unrelated electron and positron tracks reconstructed as a coincidence.



~About 80% of full APEX statistics.

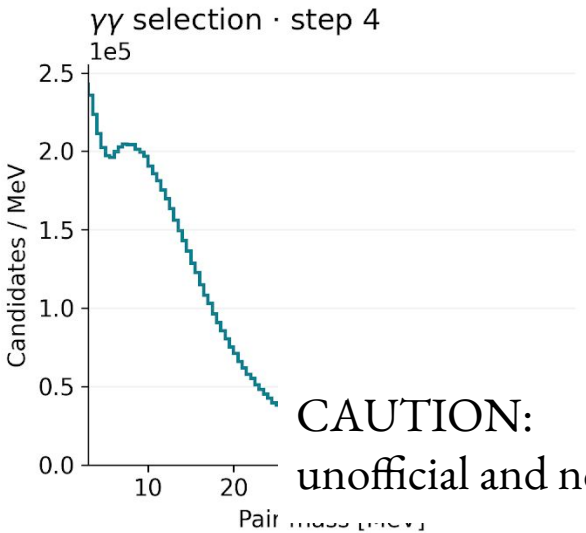
APEX + HPS available-statistics limits (high mass complement)

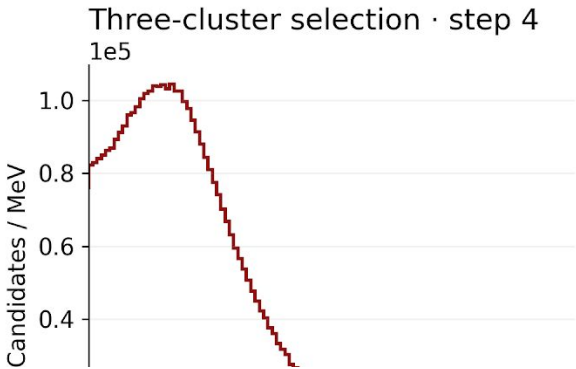


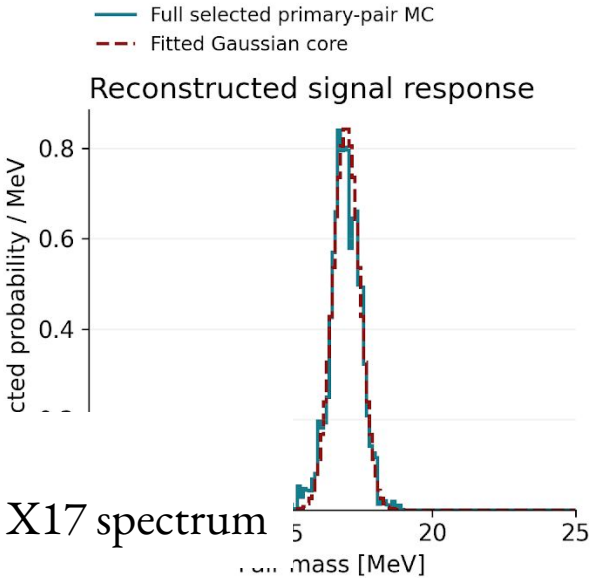
APEX projection scales event counts by 1.25.

X17 bump-hunt studies

Alternative candidate selections and reconstructed MC signal response



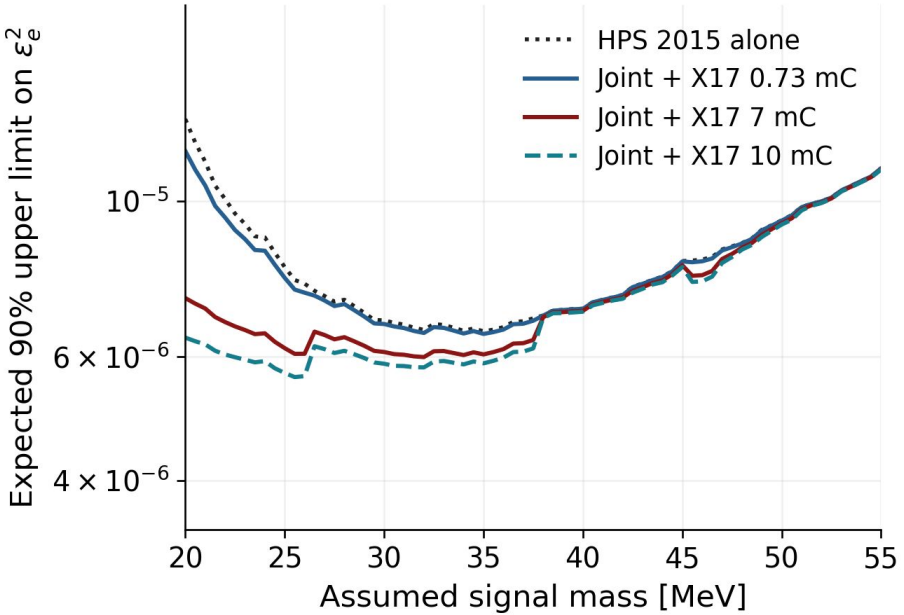




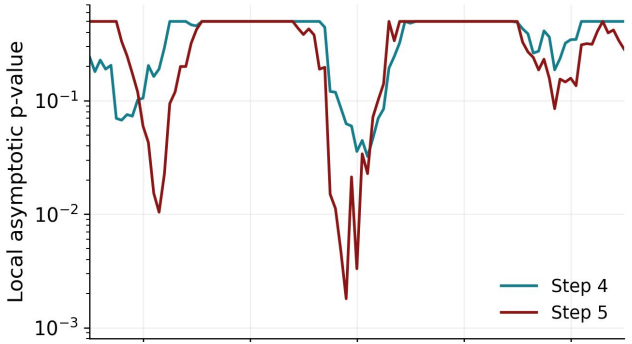
CAUTION:
unofficial and not necessarily correct inv mass for X17 spectrum

X17 + HPS 2015, expected joint reach (low mass complement)

Common-coupling projection



Observed X17 only: local p-values



ARBITRARY MASS

Protophobic: $g_p = 2g_u + g_d \simeq 0$

Dark photon: charge coupling

Plotted ε^2 assumes this case.

Conclusions

Atlas may have published first, but this approach remains novel

- different formulation of alpha
- length scale in log Mass
- alternative regularization technique
- tested on many more unique datasets with different shapes

For 2021 Analysis:

We have successfully calibrated our modeling technique. Pulls mean and widths are just about optimal. Signal extraction approaches unity across full search region.

Will have draft PRD by Nov 10th if unblinding next week. Rapid publication of 2021 is possible, at least 1 more HPS publication when combined with other experiments (with or without 2019).

Spare slides

I. GPR Supplemental Information
II. Signal Morphing
III. Using same parameters for 2016 10% worked to fit 100%
IV. Direct GP bias
V. Slightly more info on fixing hyperparameters
VI. Last two presentations in full

Hybrid background: polynomial mean + GP residuals

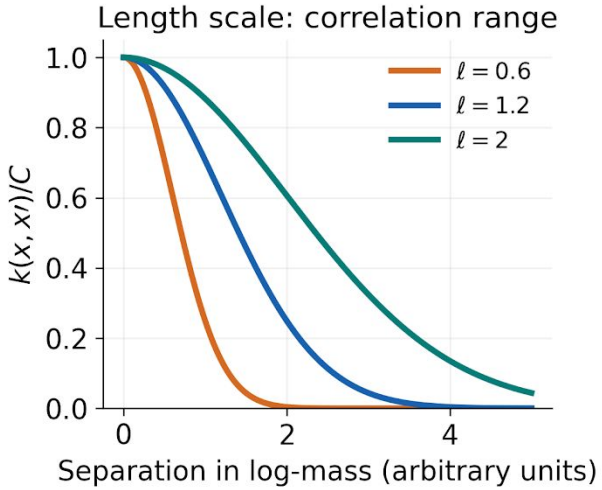
$$f(x) = p_{15}(x;\beta) + g(x)$$
$$g \sim \mathcal{GP}(0, k), \quad x = \log(m/\mathrm{GeV})$$
$$k(x,x') = C\exp\left[-\frac{(x-x')^2}{2\ell^2}\right]$$

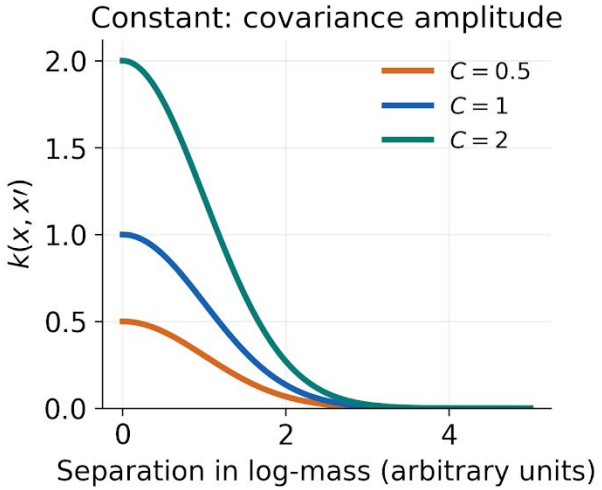
ℓ Residual smoothness in log mass

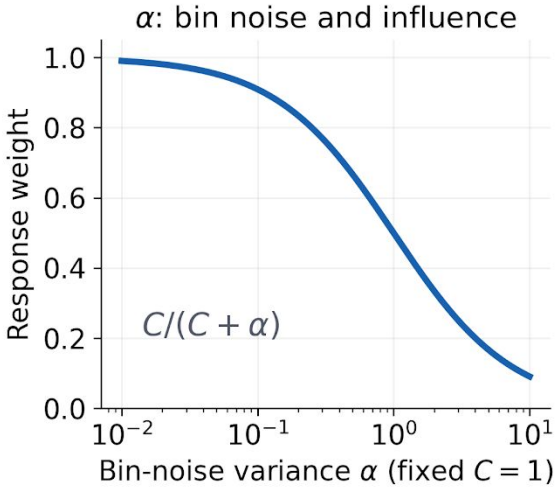
C Residual covariance amplitude

$\alpha_i = 1/n_i \quad (n_i > 0)$ Counting noise

Refit β; keep C and ℓ fixed by region.







Illustrative kernel response. The production C and ℓ follow a frozen regional rule.

Signal Morphing (1/2)

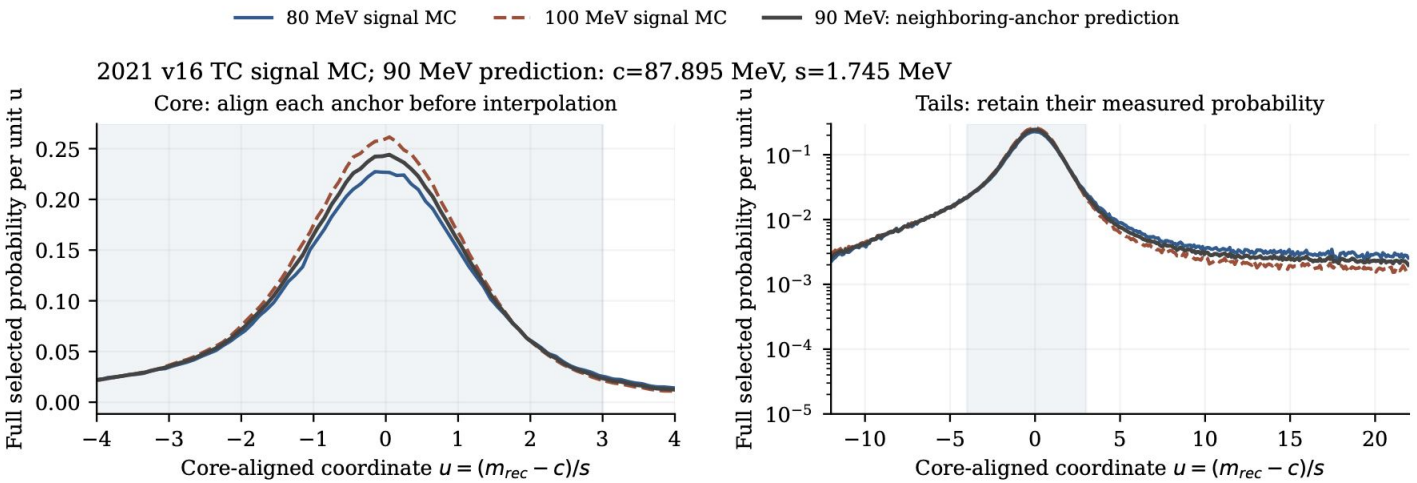
For a tested mass m between simulated masses m_L and m_R, set $t = (m - m_L)/(m_R - m_L)$. Interpolate the core center and width,

$$c_m = (1-t)c_L + tc_R, \qquad s_m = (1-t)s_L + ts_R.$$

For each anchor, define its complete aligned cumulative distribution $C_j(u) = P_j[(m_{\mathrm{rec}} - c_j)/s_j \le u]$. The prediction is

$$F_m(x) = (1-t)C_L\left(\frac{x - c_m}{s_m}\right) + tC_R\left(\frac{x - c_m}{s_m}\right).$$

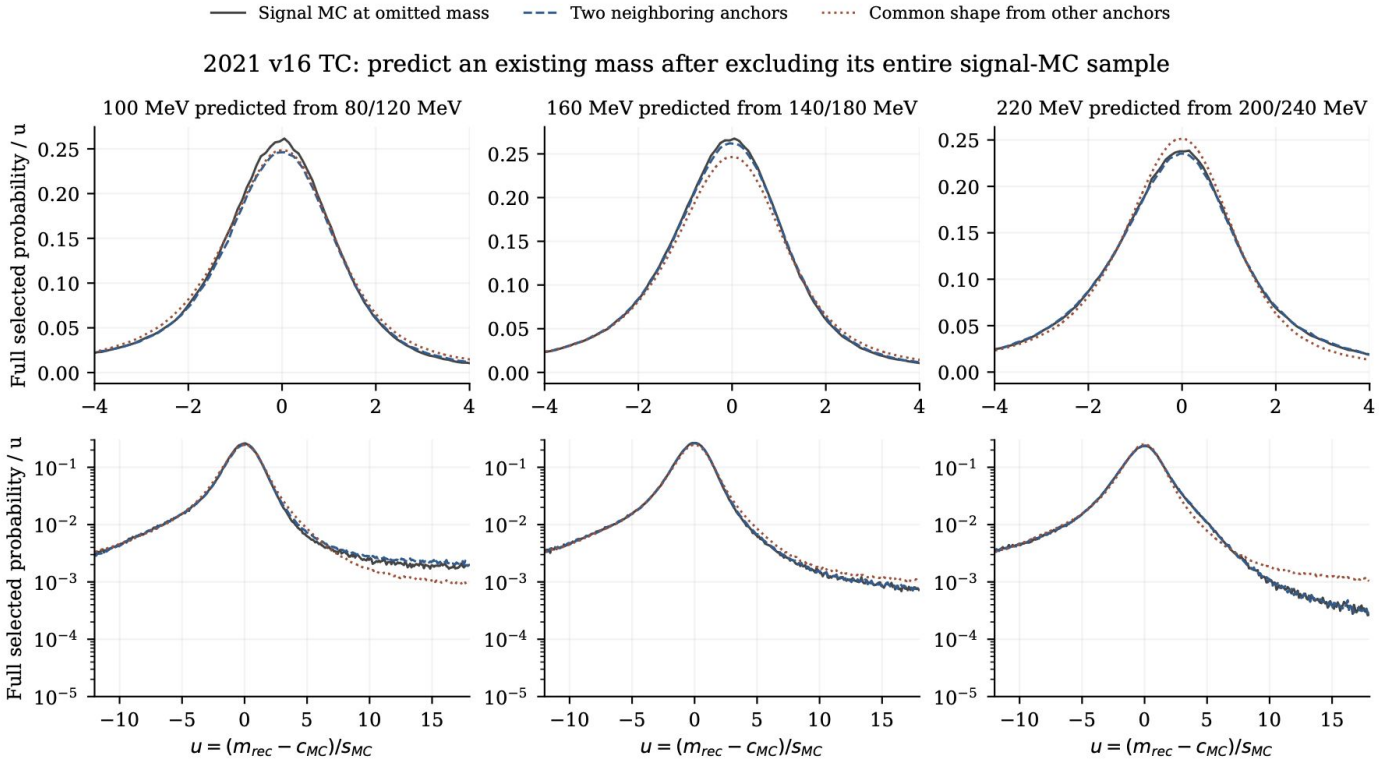
Bin probabilities are differences of this cumulative distribution at bin edges. The weights are nonnegative and sum to one, so the construction preserves probability and moves both the core and the tails. The



Signal Morphing (2/2)

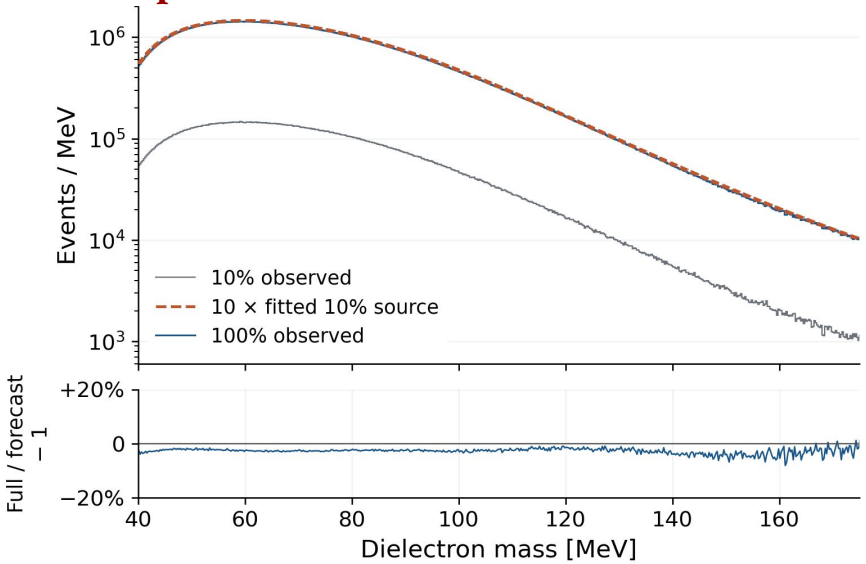
Gaussian approximation inadequate for 2021

- Although a common template could be made, it isn't the best option.
- using MC "neighbors best"

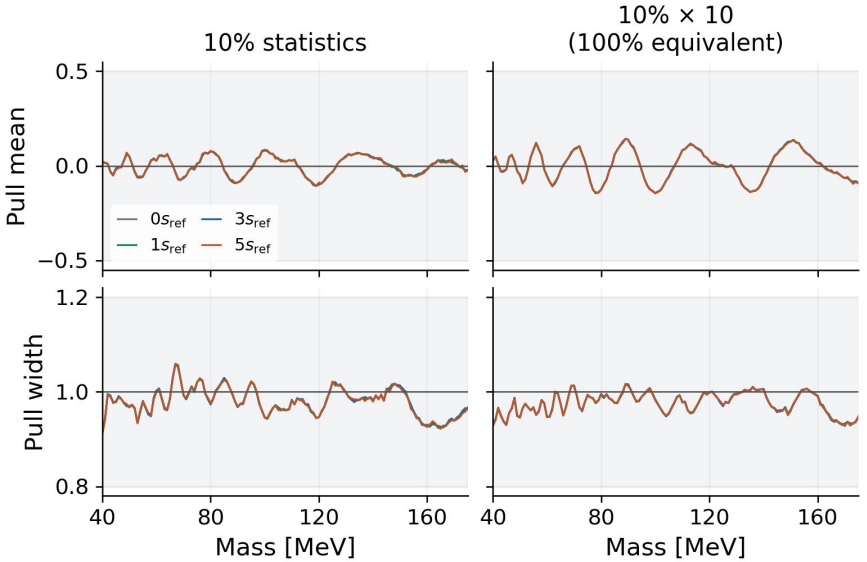


Calibration can extend to full statistics

2016 spectrum: scaled vs. observed



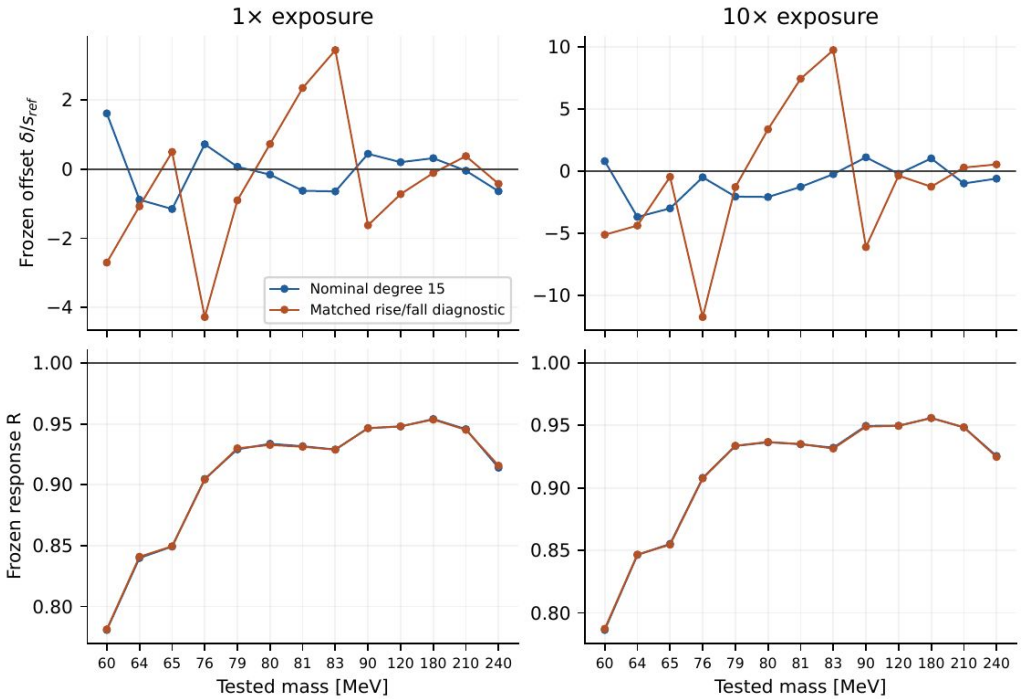
2016 nominal-source confirmation



- 2016 used fixed, campaign-specific residual-GP hyperparameters.
- Nominal-source pull checks passed at 10% and 10% × 10.
- The 10% × 10 calibration values (δ, R) were also used for the actual 100% analysis.

Bias and closure at tenfold synthetic exposure

Archived direct-GP comparator



Current hybrid, ×10

Pull means

−0.155 to +0.123

Pull widths

0.888–1.033

Paired response

0.998–1.002

No Holm rejection in the joint retention family. Six profile-CLs pointwise deficits remain.

Fixing hyperparameters

Residual GP settings are frozen by mass region.

The same values enter calibration, confirmation and data.

Test-mass region	C	ℓ
60–80 MeV (rising)	105.877	0.280784
81–240 MeV (falling)	100.994	0.273196

Selected on regional training toys; held fixed for independent confirmation.

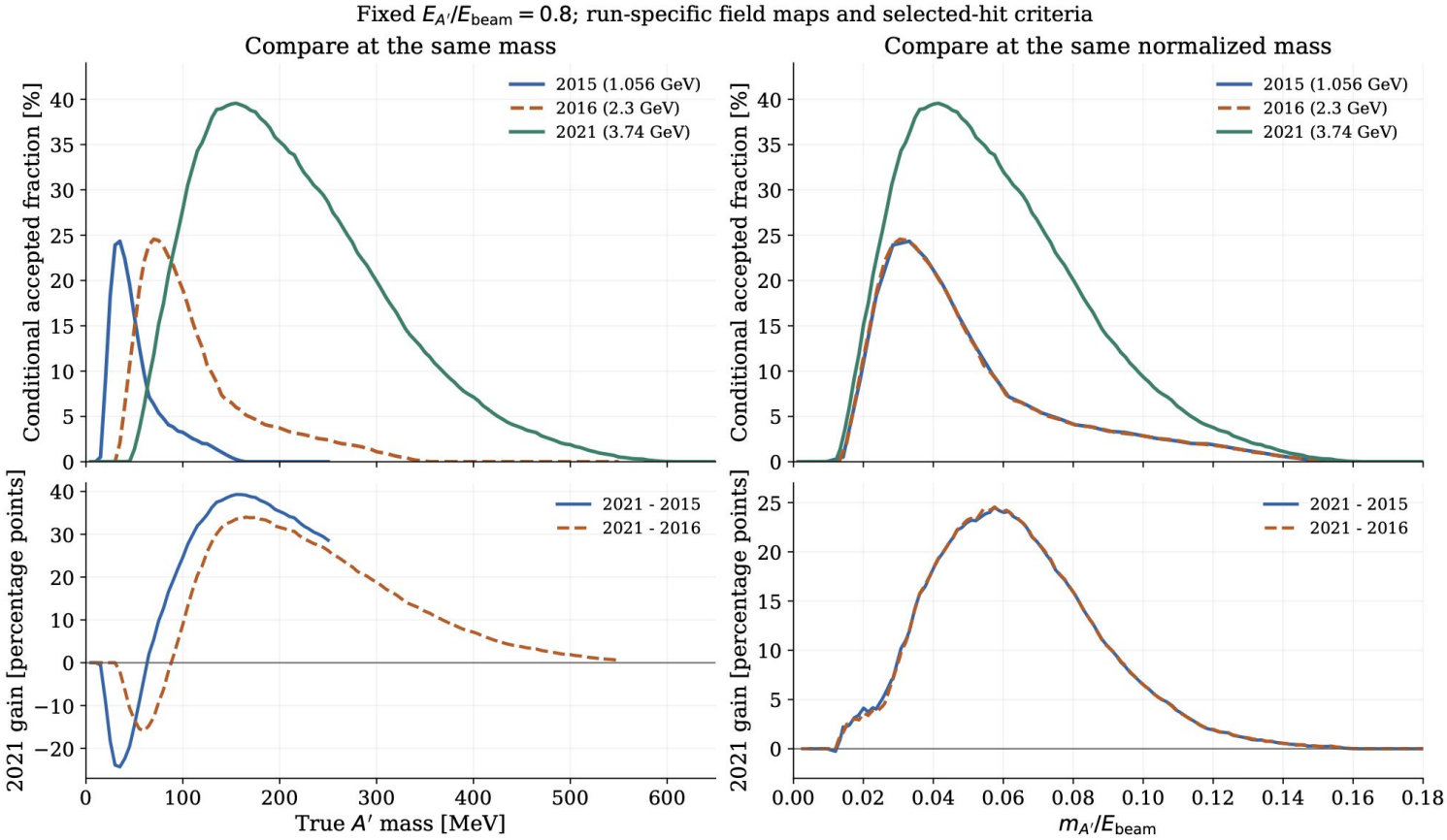
The polynomial coefficients are refit from each sample's sidebands.

List of updates to address feedback

I. better geometric acceptance plotting
II. Change ref matching to have separate “background truth” on which to calibrate
    A. Used 100 toys to create “truth” in GP Mean, and functional form
    B. Used separate 100 toys to test extraction – looks fine
III. Target Constrained MC for new signal template
    A. 2021 TC and 2016 TC differ substantively
IV. Extraction displays using new templates
V. Finalized Global Pvalue Procedure
VI. Made website

Acceptance Improvements: 2015/2016 → 2021

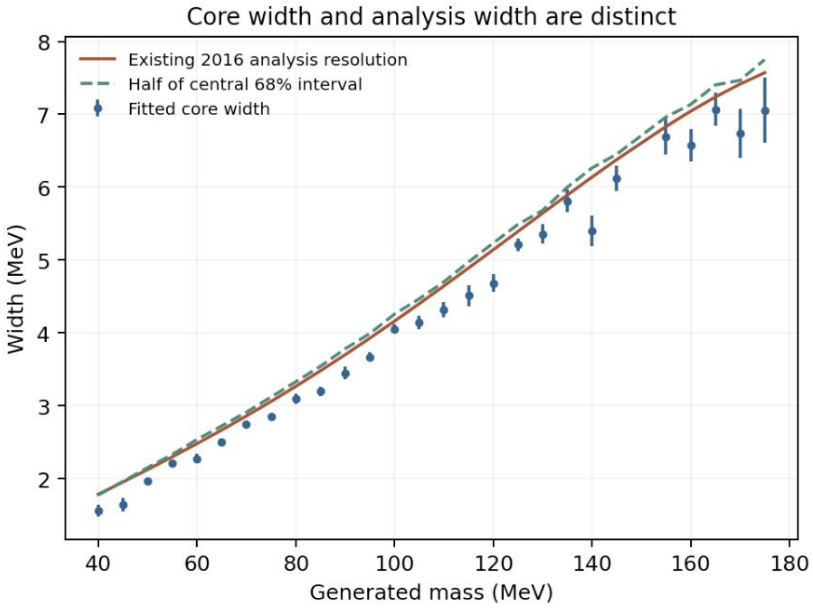
Reassuring!

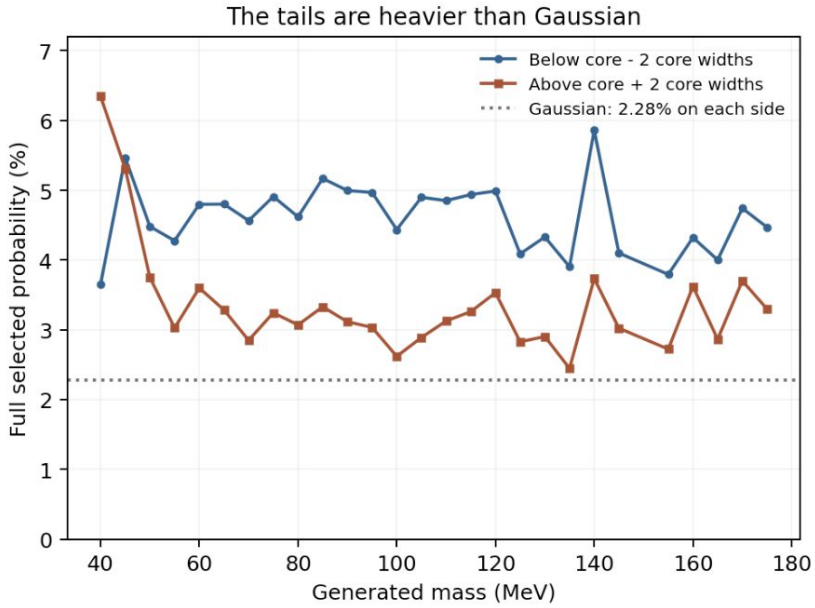


Mass Resolution / Signal Shape in 2016 TC MC

Slight discrepancy noticed with resolution curve, need to double check but could be due to not fitting with strict gaussian (better fit comes from affine + gaussian)

Tails differing from gaussing indicate need to use signal MC in fitting



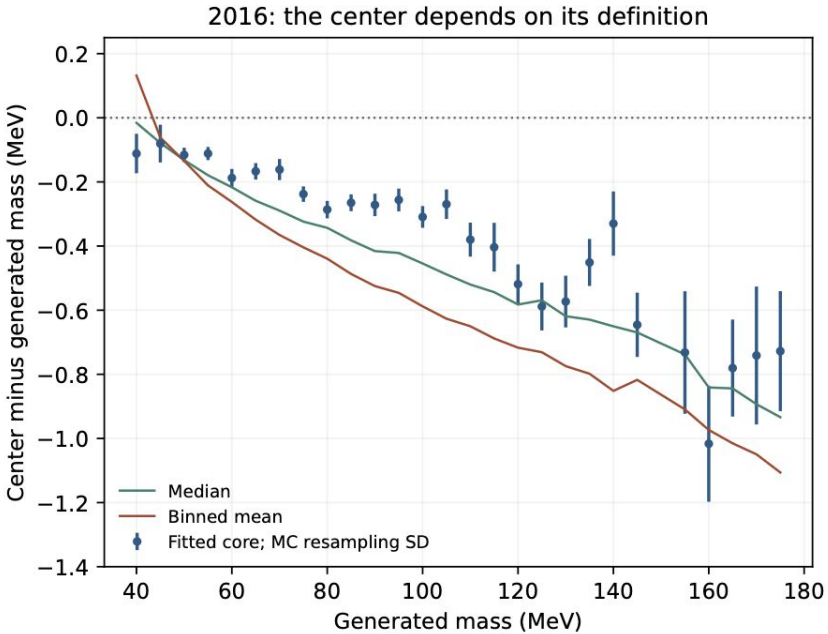


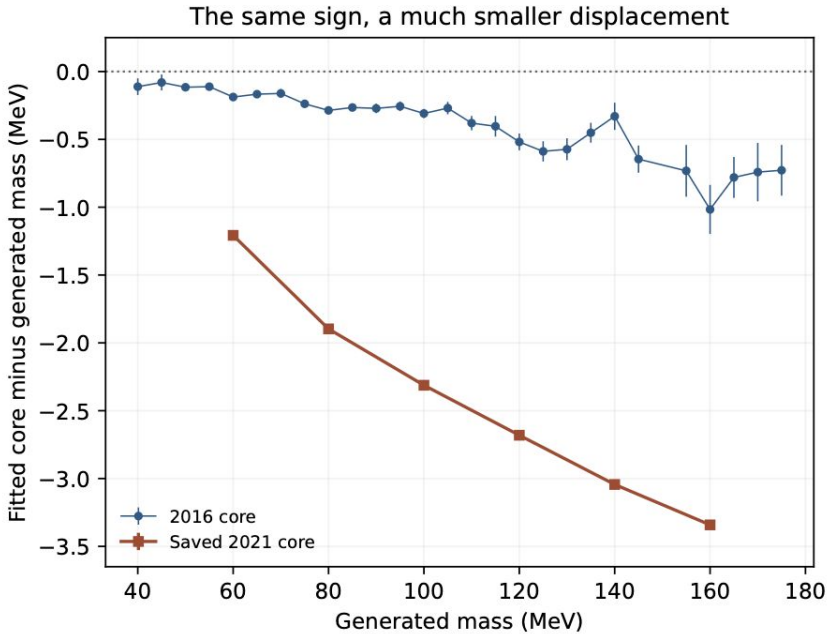
2016 TC MC Signal Location

Looks like we need to shift search position in determining upper limits

- much less of an effect in 2016 vs 2021

very different displacement compared to 2021

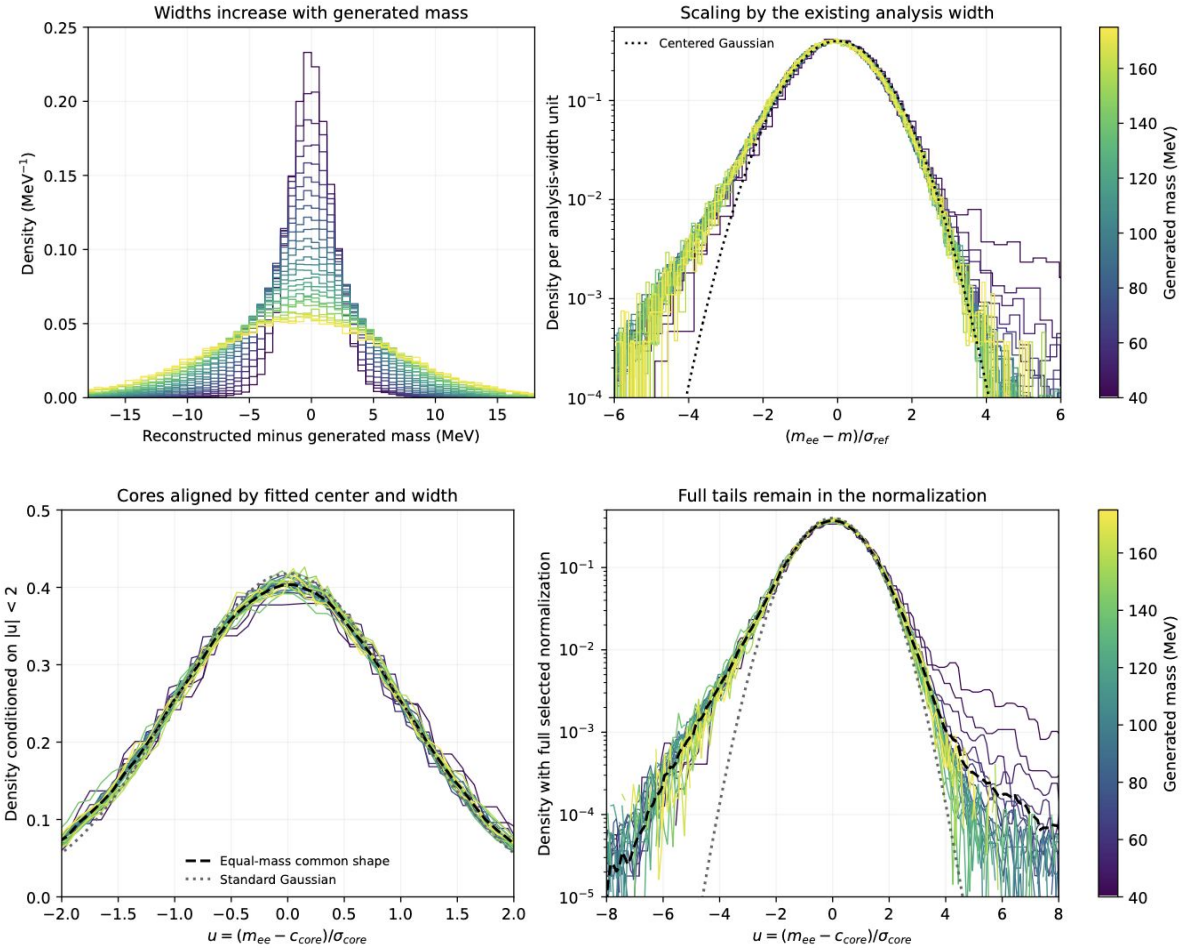




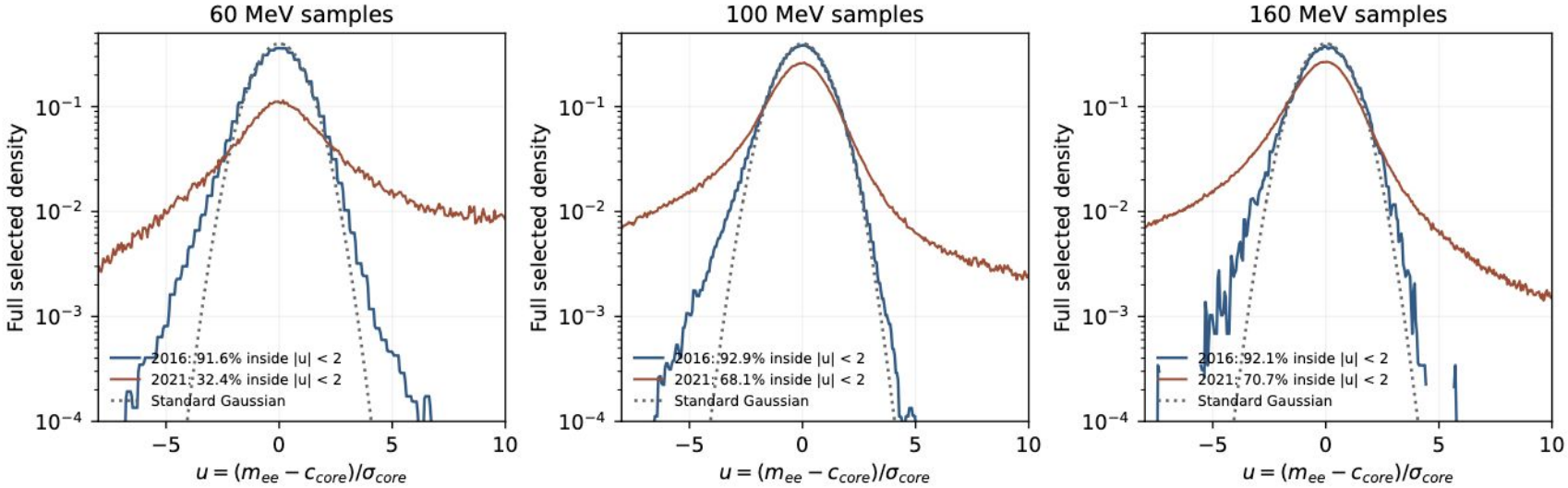
2016 MC Shape

Gaussian approximates much better than 2021, but common shape better fit

- thus, search now uses signal MC shape

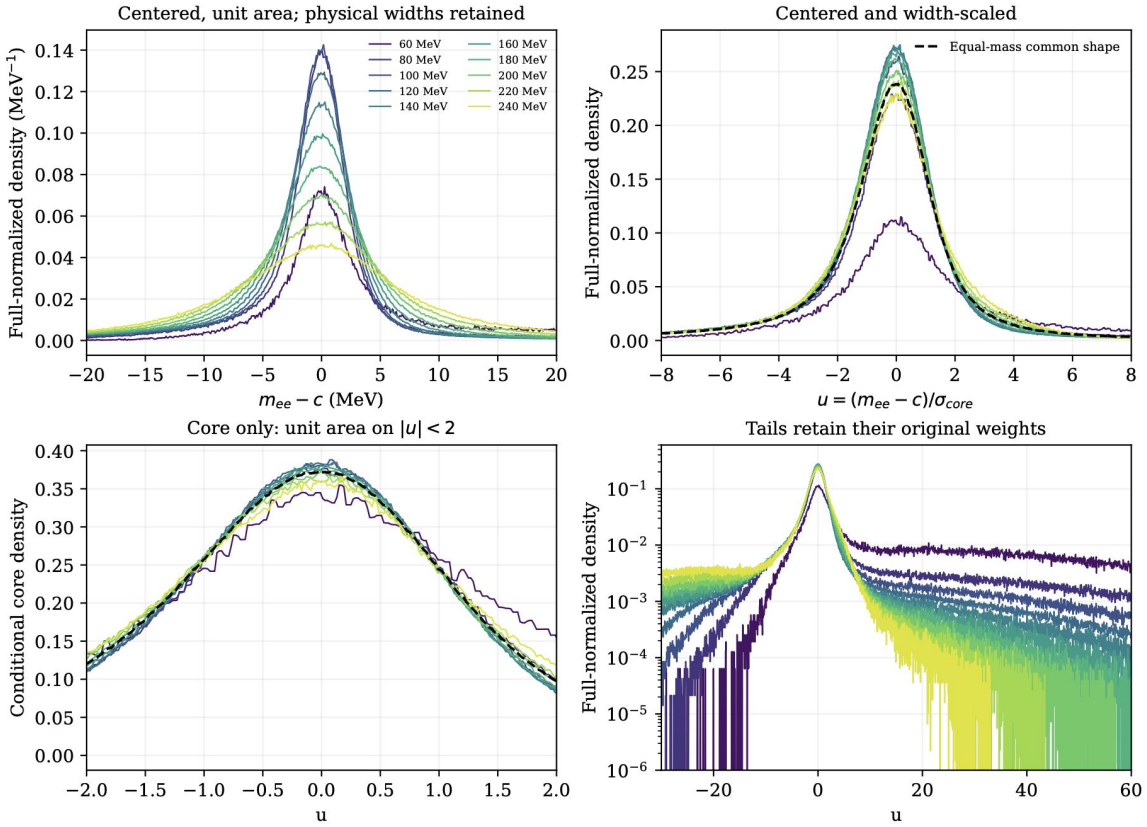


Visual Comparison – 2016 vs 2021 Signal MC

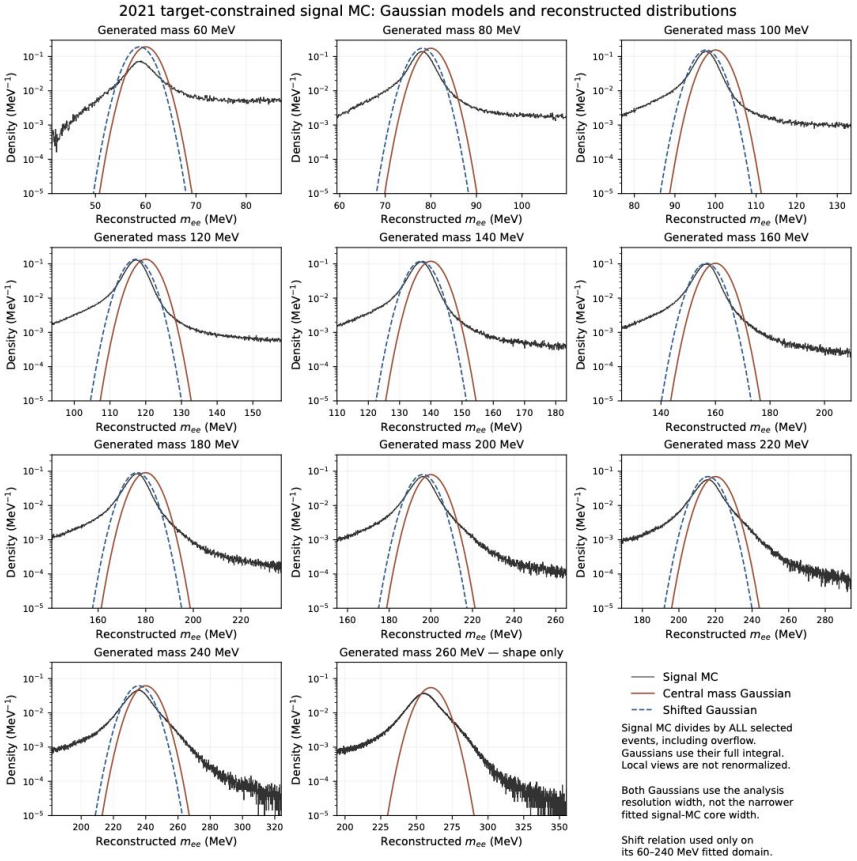


2021 MC Shape

large amounts of signal outside the core
problematic if true to signal in data
- lots of leakage

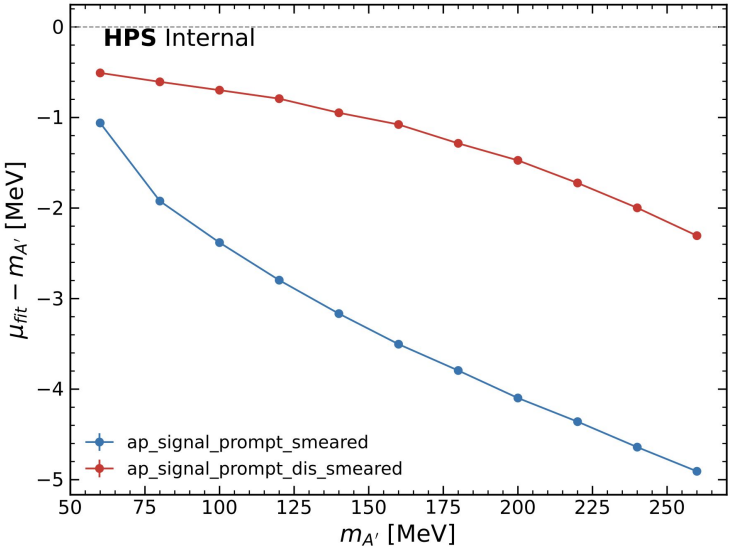


2021 MC Location



Moving signal shape

$$c(m) = m - 3.2243308693 - 2.2139928114\ln(m/150)$$



From M. Gignac, mostly agrees with my determined displacement

Calibrating away bias'

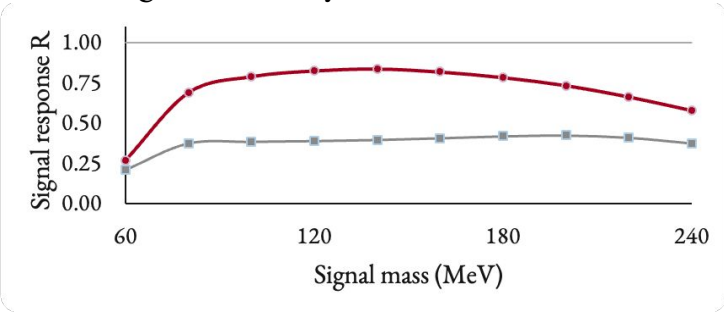
Procedure:

- training ensemble 100 toys
  - (learn background-only fitted-yield bias → δ)
  - measure response at each mass
- Use response and bias at each mass to calibrate different set of background toys (verify it works)
- apply response and bias to observations to correct for them

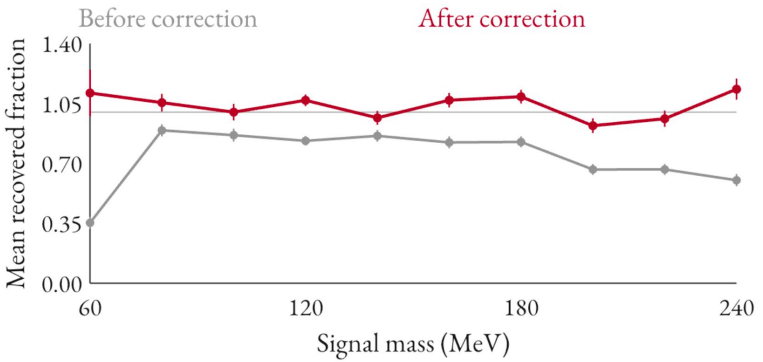
simplification for purpose of presentation

$$A_c = \frac{\hat{A} - \delta}{R}$$

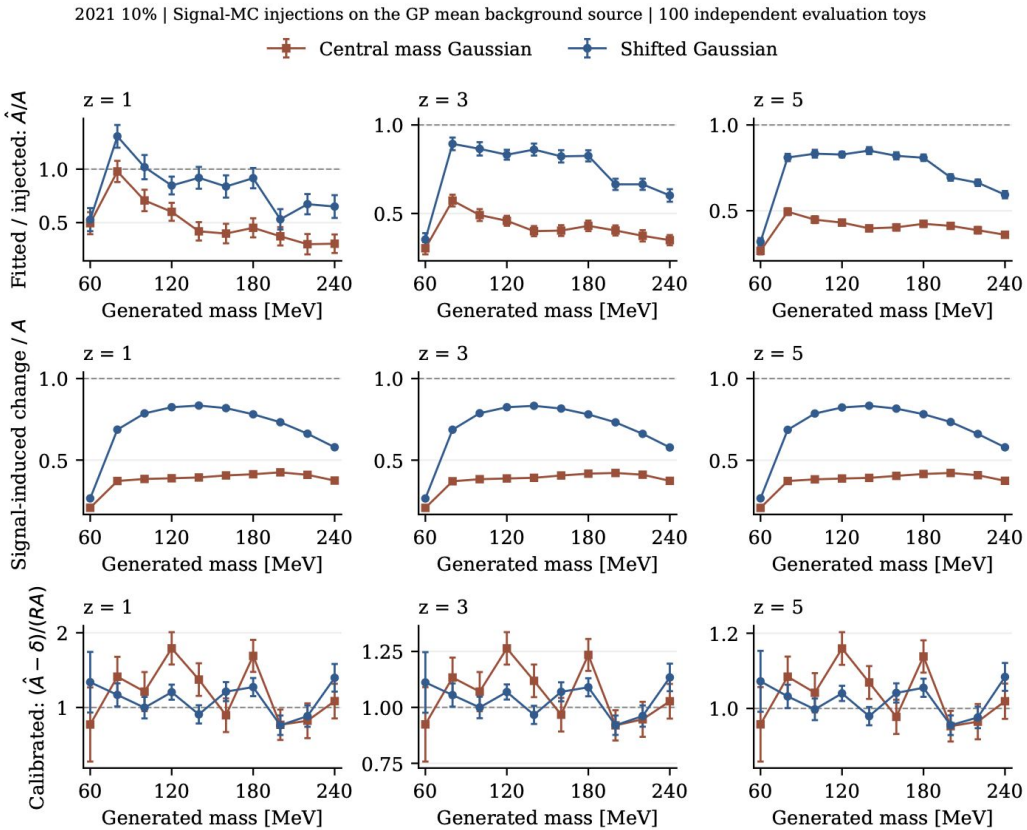
Background-only fluctuation scale



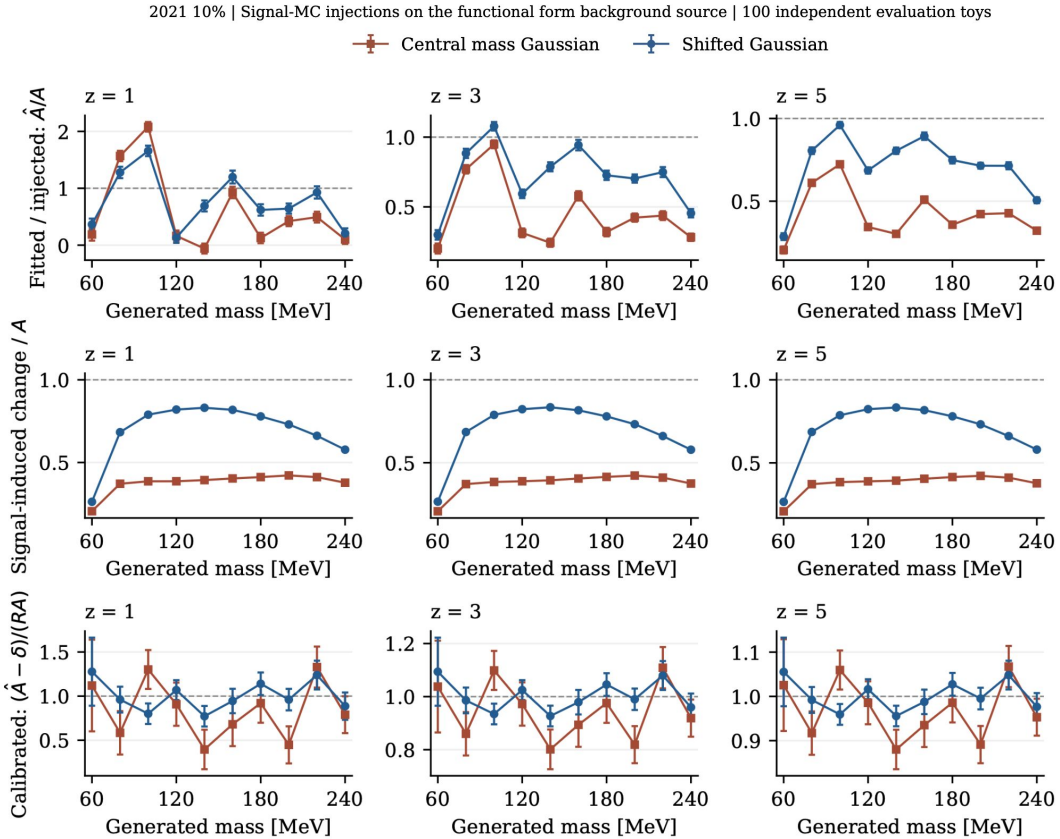
2021 10% · shifted Gaussian extraction



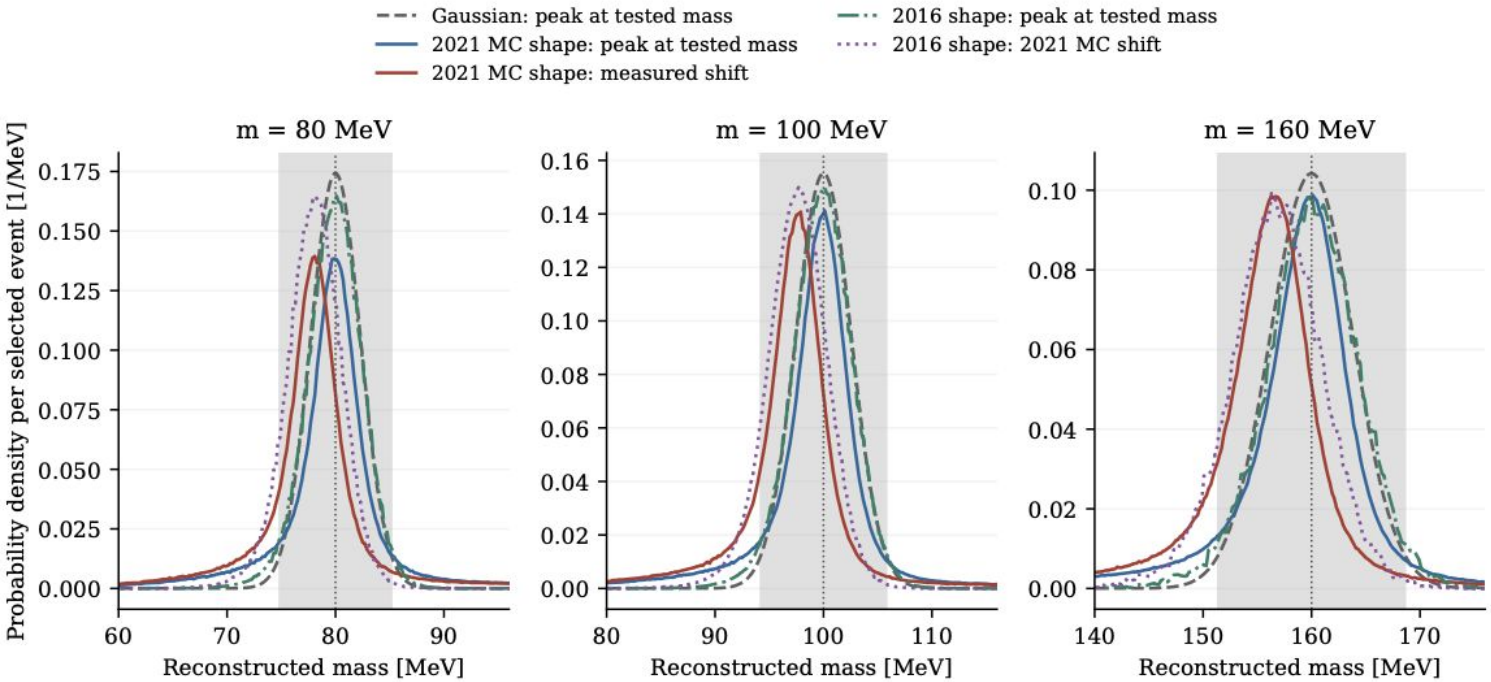
GP Mean Bias Removal



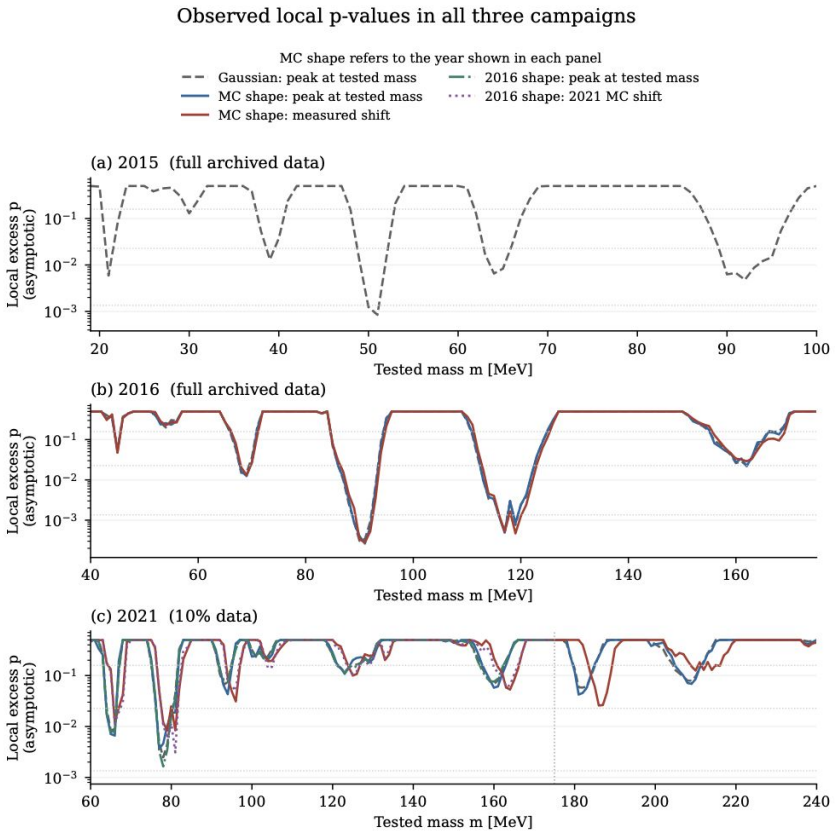
Functional Form Bias Removal

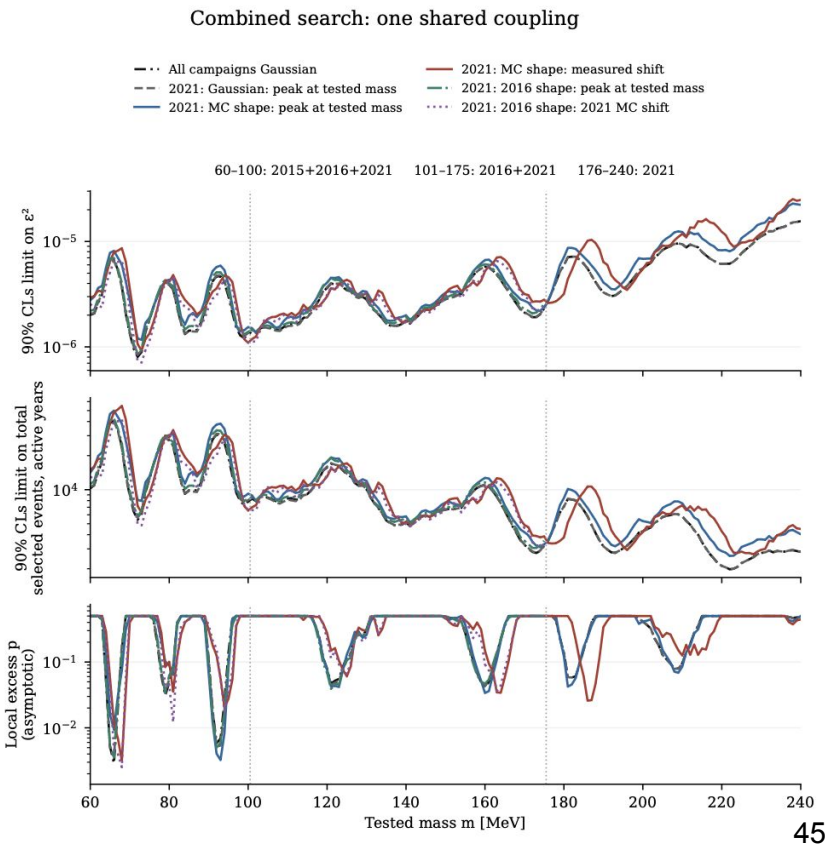


Different Signal Hypotheses

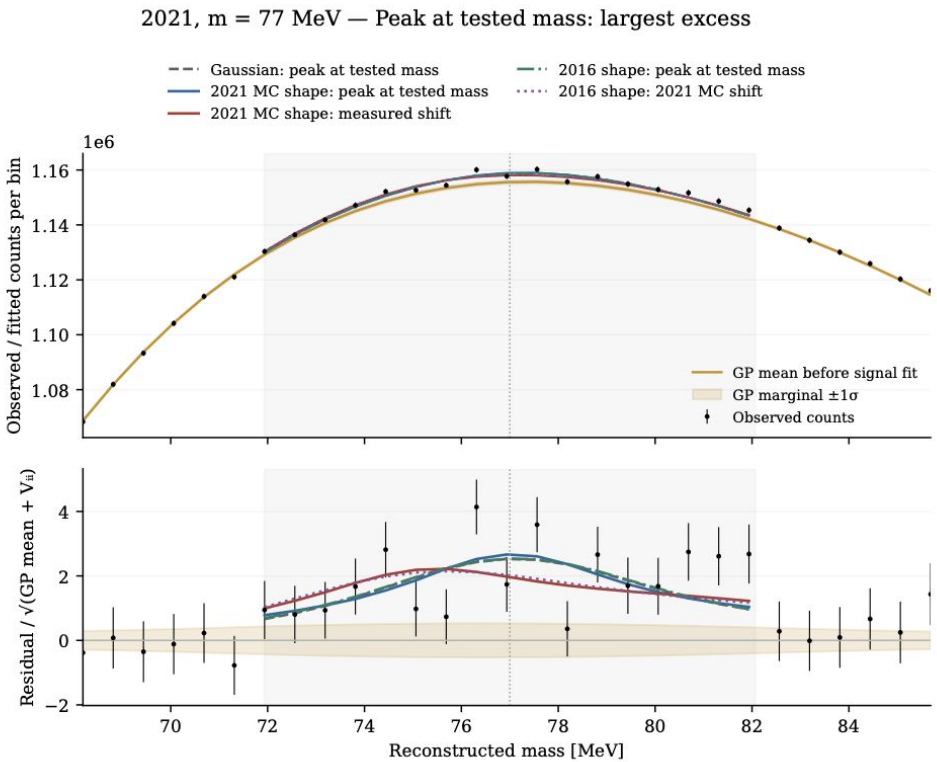


Local Significances across the different hypotheses





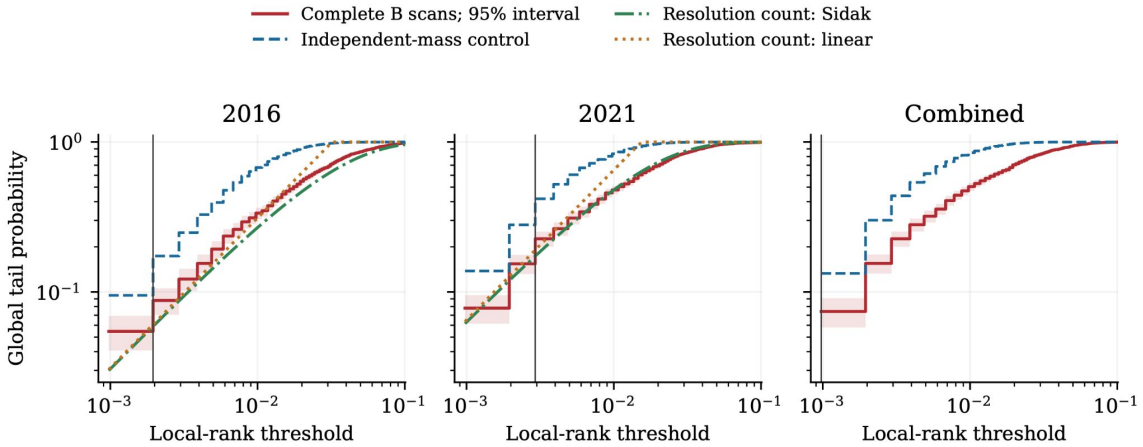
Different extractions



Using toys to compute Global Signifcances

Approaches bonferelli/sidak approximation!

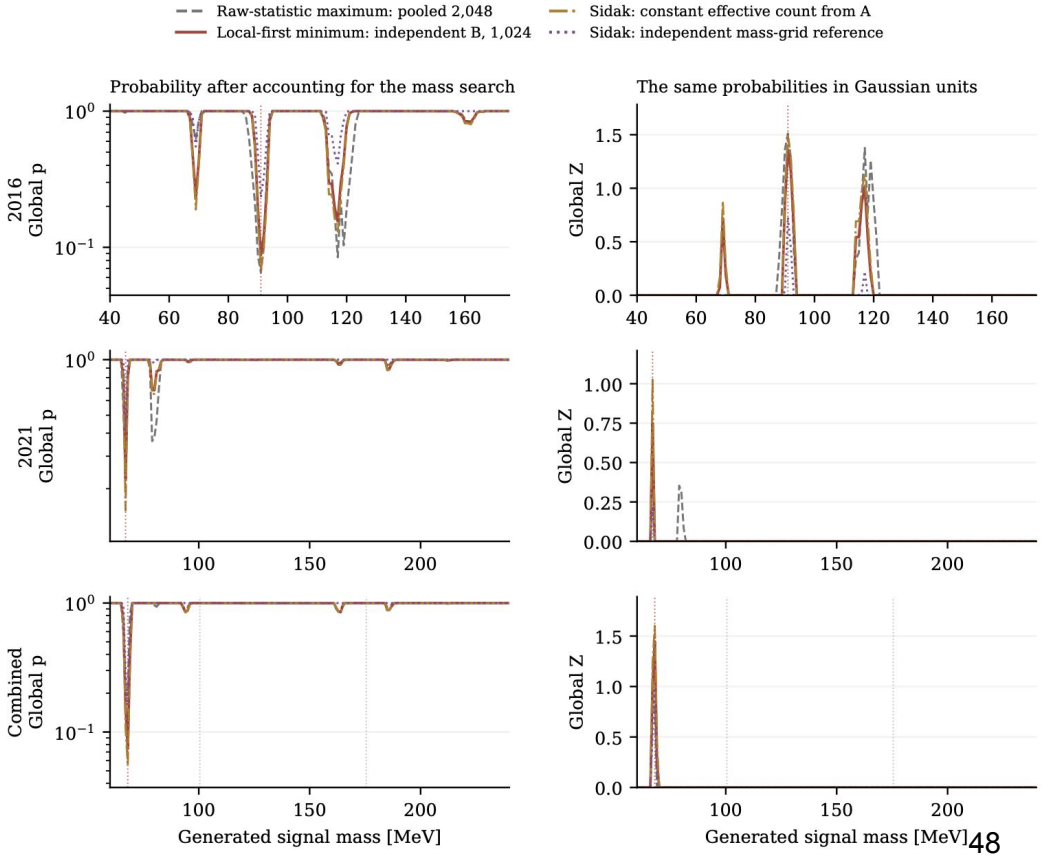
- great, so we can use it in our final result



Global Significances

Dataset Considerations:

- 2016: vast majority of signal in shifted MC core
- 2021 asymmetric blind window used, v16 template used
- 2015 Gaussian Used



Signal Morphing

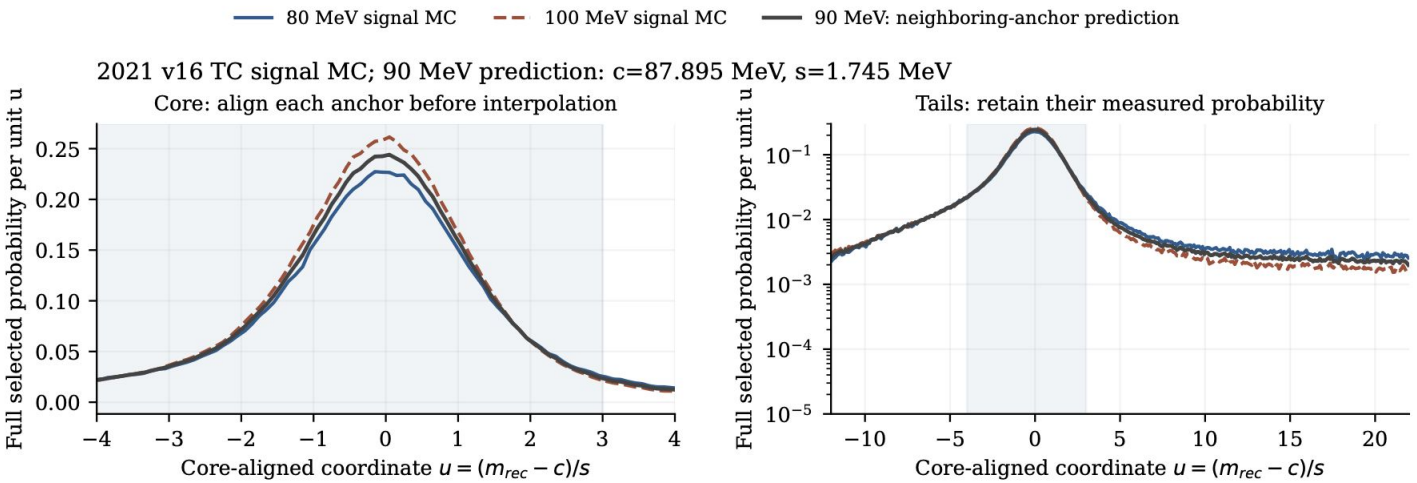
For a tested mass m between simulated masses m_L and m_R, set $t = (m - m_L)/(m_R - m_L)$. Interpolate the core center and width,

$$c_m = (1-t)c_L + tc_R, \qquad s_m = (1-t)s_L + ts_R.$$

For each anchor, define its complete aligned cumulative distribution $C_j(u) = P_j[(m_{\mathrm{rec}} - c_j)/s_j \le u]$. The prediction is

$$F_m(x) = (1-t)C_L\left(\frac{x - c_m}{s_m}\right) + tC_R\left(\frac{x - c_m}{s_m}\right).$$

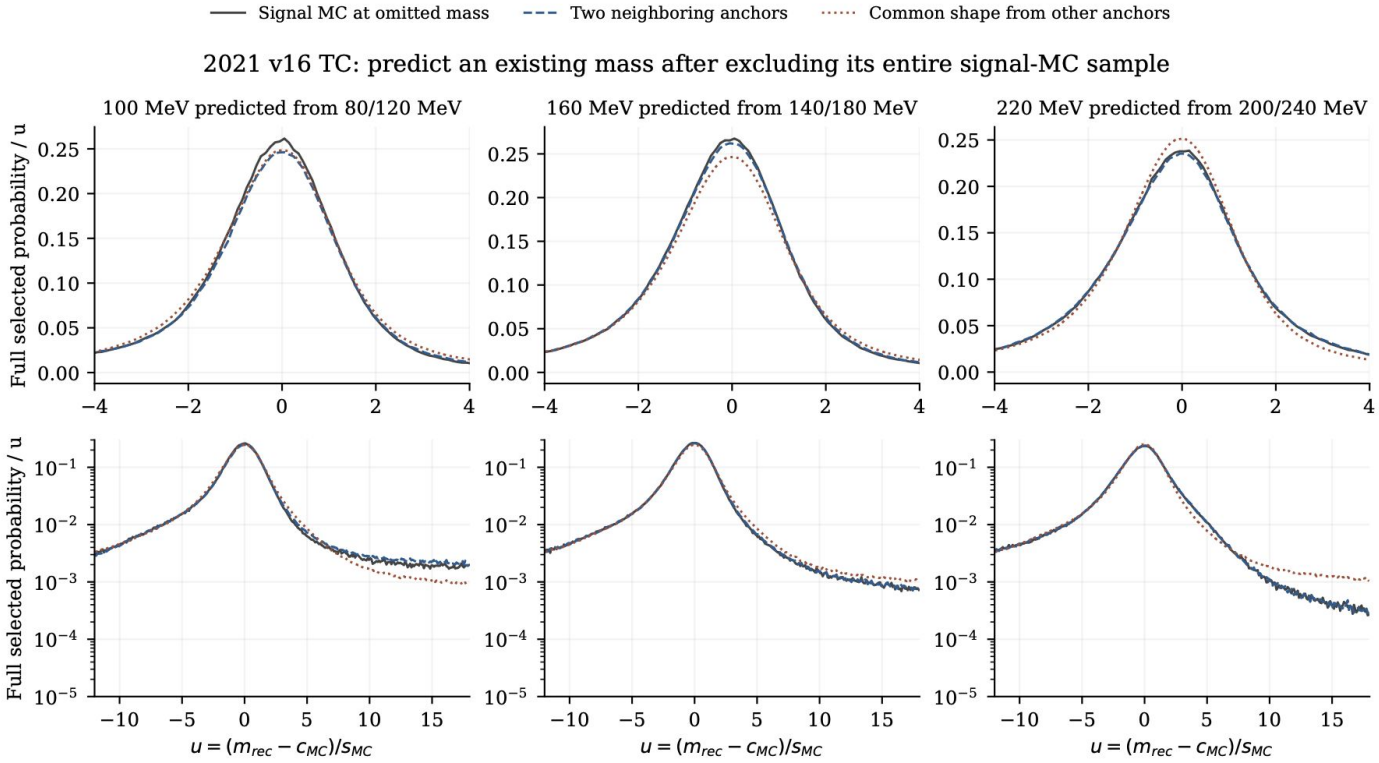
Bin probabilities are differences of this cumulative distribution at bin edges. The weights are nonnegative and sum to one, so the construction preserves probability and moves both the core and the tails. The



Signal Morphing

Gaussian approximation inadequate for 2021

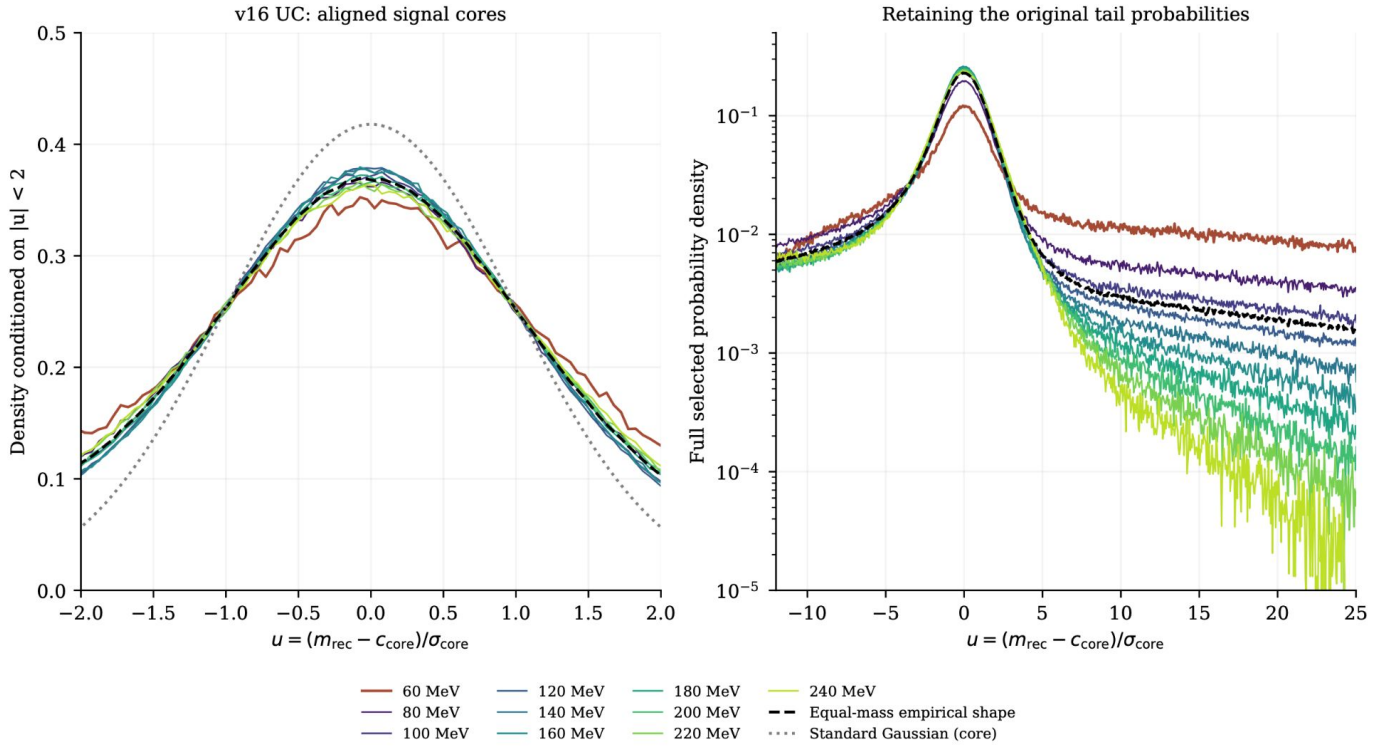
- Although a common template could be made, it isn't the best option.
- using MC "neighbors best"



UC Signal MC

Gaussian inadequate everywhere

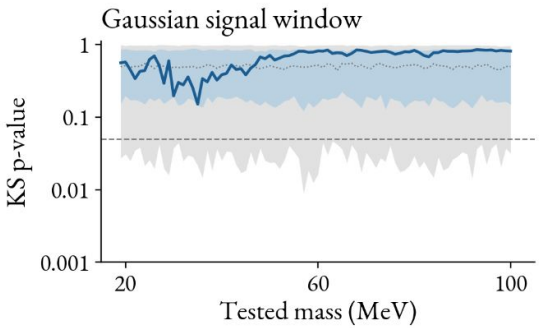
Common Shape inadequate at tails



Sideband fit quality: KS Statistics

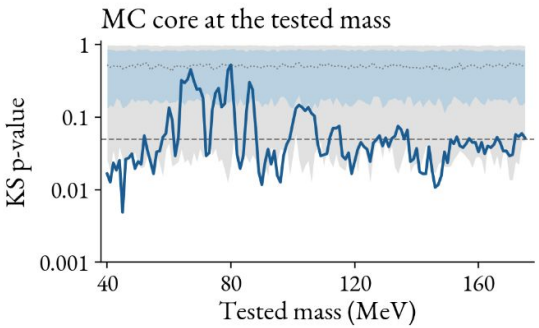
2015 · full sample

±2.25 reference widths



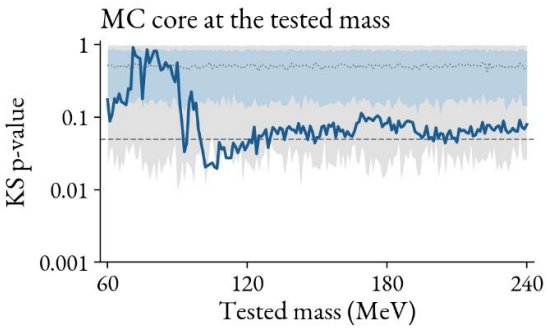
2016 · full sample

±2.5 core widths



2021 · 10% sample

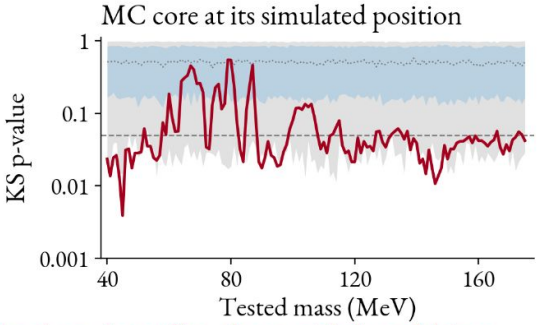
-4 to +3 core widths

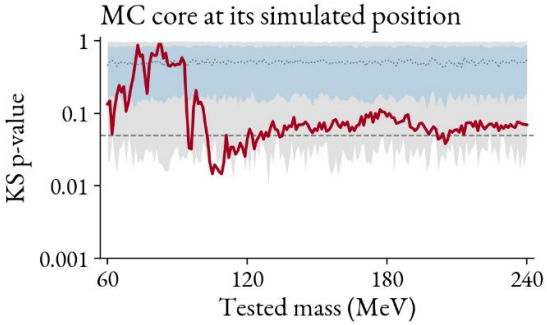


Toys made from fitted GP mean

Colored line: observed data
Light grey: middle 95% of toys
Blue band: middle 68% of toys
Dotted: toy median; dashed: 0.05

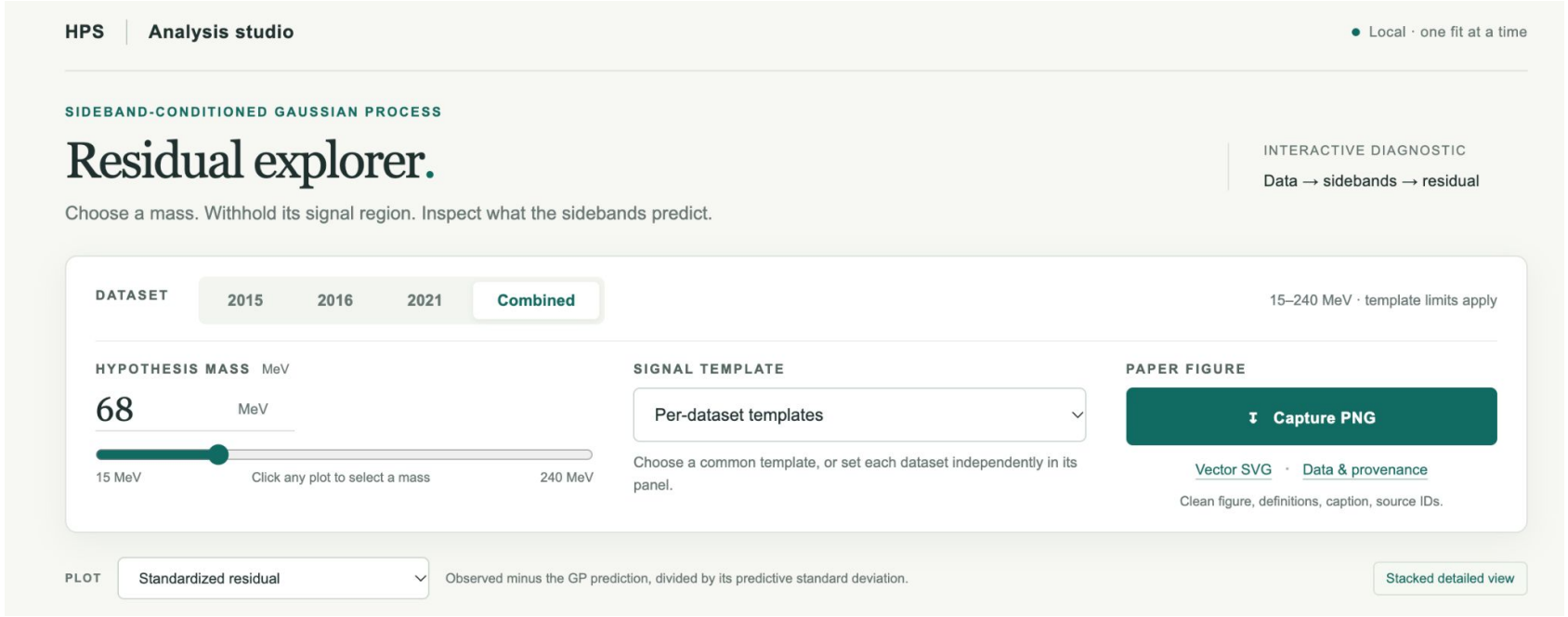
512 new toys per mass, ranked against 1,024 saved reference toys.





Toy ranges are pointwise; neighboring data fits share sideband bins.

Resonance Search Website



The joint fit uses a common signal strength with dataset-specific yields. Its availability and statistical interpretation are reported below.

Shared-coupling likelihood fit

3 participating datasets · 68.0 MeV

SHARED COUPLING Ψ = E² / 10⁴	JOINT LIKELIHOOD SCORE	JOINT LOCAL ASYMPTOTIC P	CONDITIONAL CALIBRATED JOINT P
635.525 ± 226.156	+2.81	0.0025	9.99e-4
1σ Hessian · ψ = e² / 10⁴	Signed likelihood root · shared parameter	Excess-only · p = 0.5 for deficits	Conditional local rank p; 0/1,000 exceedances; 95% binomial interval [0, 0.0037]. Zero exceedances do not imply zero probability.

A single coupling ψ = e² / 10⁴ sets the expected yield in each dataset. This joint likelihood is different from adding individual scores or count sums.

▸ Joint fit components and provenance

2015 · 2015 · full data
m = 68.0 MeV · Gaussian · nominal mass
Gaussian · nominal mass
Analysis binning 0.250 MeV
Native-count analysis bins · refit on change
Withheld: 60.00–76.00 MeV
Center 68.00 · width 3.53 · bin 0.250 MeV
Standardized residual
Invariant mass [MeV]
Data residual + √n bars · GP prediction ±2σ · Withheld region · Fitted signal

2016 · 2016 · full data
m = 68.0 MeV · Signal MC · 2016 · ±3.5 core widths
Signal MC · 2016 · ±3.5 σ
Analysis binning 0.250 MeV
Native-count analysis bins · refit on change
Withheld: 58.50–77.00 MeV
Center 67.84 · width 2.65 · bin 0.250 MeV
Standardized residual
Invariant mass [MeV]
Data residual + √n bars · GP prediction ±2σ · Withheld region · Fitted signal

2021 · 2021 · 10% data
m = 68.0 MeV · Signal MC · v16 TC · −4/+3 core widths
Signal MC · v16 TC · −4/+
Analysis binning 0.625 MeV
Native-count analysis bins · refit on change
Withheld: 60.38–71.00 MeV
Center 66.52 · width 1.58 · bin 0.625 MeV
Standardized residual
Invariant mass [MeV]
Data residual + √n bars · GP prediction ±2σ · Withheld region · Fitted signal

PLOT Standardized residual
Observed minus the GP prediction, divided by its predictive standard deviation.

01 / Standardized residual
91.0 MeV · cached inputs · 1 dataset

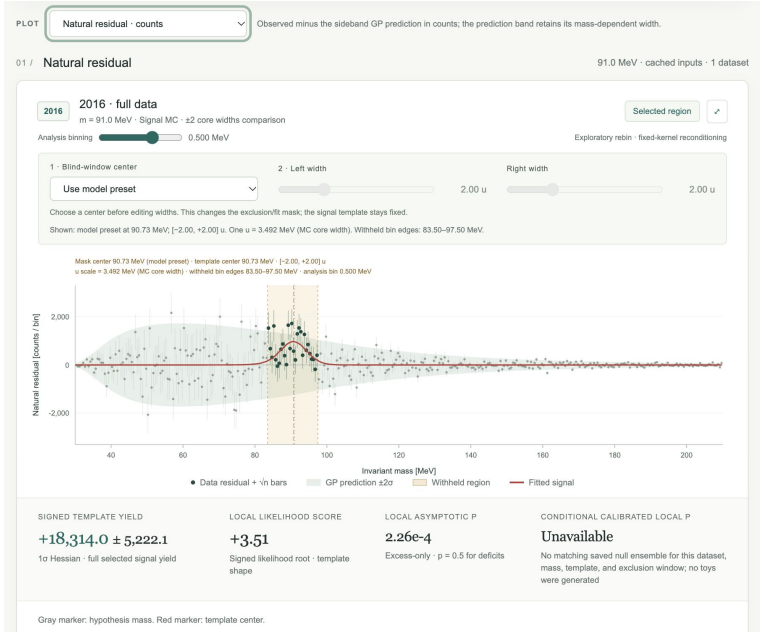
2016 · 2016 · full data
m = 91.0 MeV · Signal MC · ±2 core widths comparison
Selected region
Analysis binning 0.500 MeV
Exploratory rebin · fixed-kernel reconditioning

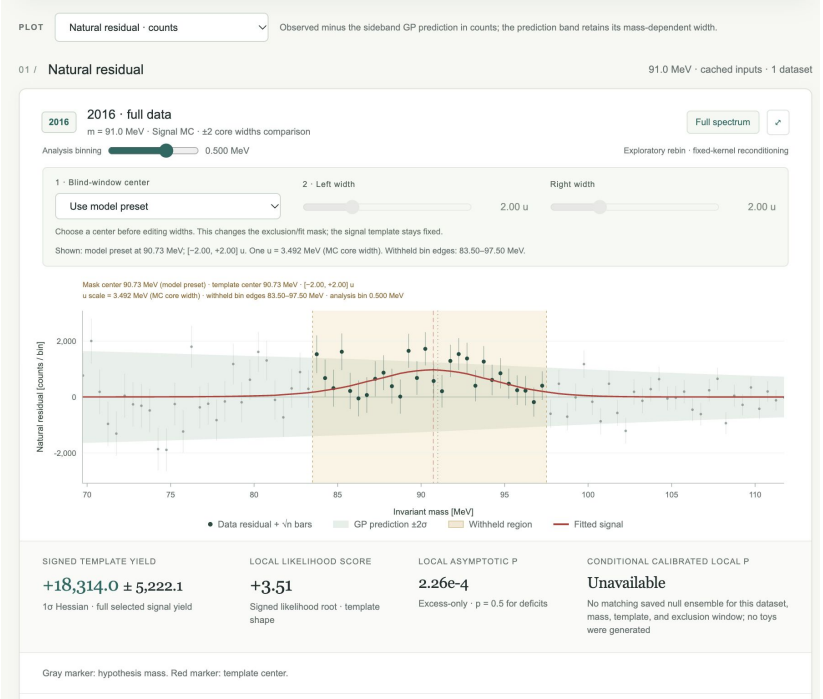
1 · Blind-window center
Use model preset
2 · Left width 2.00 u
Right width 2.00 u
Choose a center before editing widths. This changes the exclusion/fit mask; the signal template stays fixed.
Shown: model preset at 90.73 MeV; [−2.00, +2.00] u. One u = 3.492 MeV (MC core width). Withheld bin edges: 83.50–97.50 MeV.

Mask center 90.73 MeV (model preset) · template center 90.73 MeV · [−2.00, +2.00] u
u scale = 3.492 MeV (MC core width) · withheld bin edges 83.50–97.50 MeV · analysis bin 0.500 MeV
Standardized residual
Invariant mass [MeV]
Data residual + √n bars · GP prediction ±2σ · Withheld region · Fitted signal

SIGNED TEMPLATE YIELD	LOCAL LIKELIHOOD SCORE	LOCAL ASYMPTOTIC P	CONDITIONAL CALIBRATED LOCAL P
+18,314.0 ± 5,222.1	+3.51	2.26e-4	Unavailable
1σ Hessian · full selected signal yield	Signed likelihood root · template shape	Excess-only · p = 0.5 for deficits	No matching saved null ensemble for this dataset, mass, template, and exclusion window; no toys were generated

Residual denominator: √(GP background + GP variance). Gray marker: hypothesis mass. Red marker: template center.


PLOT
Natural residual · counts
Observed minus the sideband GP prediction in counts; the prediction band retains its mass-dependent width.
01 / Natural residual
91.0 MeV · cached inputs · 1 dataset
2016
2016 · full data
m = 91.0 MeV · Signal MC · ±2 core widths comparison
Selected region
Analysis binning
0.500 MeV
Exploratory rebin · fixed-kernel reconditioning
1 · Blind-window center
Use model preset
2 · Left width
2.00 u
Right width
2.00 u
Choose a center before editing widths. This changes the exclusion/fit mask; the signal template stays fixed.
Shown: model preset at 90.73 MeV; [−2.00, +2.00] u. One u = 3.492 MeV (MC core width). Withheld bin edges: 83.50–97.50 MeV.
Mask center 90.73 MeV (model preset) · template center 90.73 MeV · [−2.00, +2.00] u
u scale = 3.492 MeV (MC core width) · withheld bin edges 83.50–97.50 MeV · analysis bin 0.500 MeV
Natural residual [counts / bin]
2,000
0
-2,000
40
60
80
100
120
140
160
180
200
Invariant mass [MeV]
Data residual + √n bars
GP prediction ±2σ
Withheld region
Fitted signal
SIGNED TEMPLATE YIELD
+18,314.0 ± 5,222.1
1σ Hessian · full selected signal yield
LOCAL LIKELIHOOD SCORE
+3.51
Signed likelihood root · template shape
LOCAL ASYMPTOTIC P
2.26e-4
Excess-only · p = 0.5 for deficits
CONDITIONAL CALIBRATED LOCAL P
Unavailable
No matching saved null ensemble for this dataset, mass, template, and exclusion window; no toys were generated
Gray marker: hypothesis mass. Red marker: template center.


PLOT
Natural residual · counts
Observed minus the sideband GP prediction in counts; the prediction band retains its mass-dependent width.
01 / Natural residual
91.0 MeV · cached inputs · 1 dataset
2016
2016 · full data
m = 91.0 MeV · Signal MC · ±2 core widths comparison
Full spectrum
Analysis binning
0.500 MeV
Exploratory rebin · fixed-kernel reconditioning
1 · Blind-window center
Use model preset
2 · Left width
2.00 u
Right width
2.00 u
Choose a center before editing widths. This changes the exclusion/fit mask; the signal template stays fixed.
Shown: model preset at 90.73 MeV; [−2.00, +2.00] u. One u = 3.492 MeV (MC core width). Withheld bin edges: 83.50–97.50 MeV.
Mask center 90.73 MeV (model preset) · template center 90.73 MeV · [−2.00, +2.00] u
u scale = 3.492 MeV (MC core width) · withheld bin edges 83.50–97.50 MeV · analysis bin 0.500 MeV
Natural residual [counts / bin]
2,000
0
-2,000
70
75
80
85
90
95
100
105
110
Invariant mass [MeV]
Data residual + √n bars
GP prediction ±2σ
Withheld region
Fitted signal
SIGNED TEMPLATE YIELD
+18,314.0 ± 5,222.1
1σ Hessian · full selected signal yield
LOCAL LIKELIHOOD SCORE
+3.51
Signed likelihood root · template shape
LOCAL ASYMPTOTIC P
2.26e-4
Excess-only · p = 0.5 for deficits
CONDITIONAL CALIBRATED LOCAL P
Unavailable
No matching saved null ensemble for this dataset, mass, template, and exclusion window; no toys were generated
Gray marker: hypothesis mass. Red marker: template center.

2015 · full data

m = 68.0 MeV · Gaussian · nominal mass

Gaussian · nominal mass

Analysis binning 0.250 MeV

Native-count analysis bins · refit on change

1 · Blind-window center

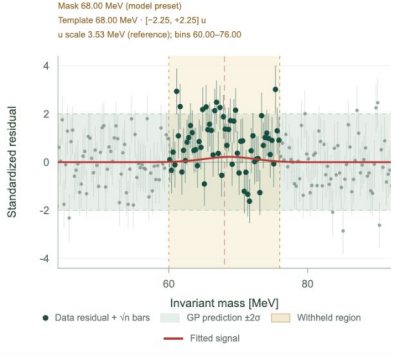
Use model preset

2 · Left width 2.25 u

Right width 2.25 u

Choose a center before editing widths. This changes the exclusion/fit mask; the signal template stays fixed.

Shown: model preset at 68.00 MeV; [−2.25, +2.25] u. One u = 3.527 MeV (reference resolution). Withheld bin edges: 60.00–76.00 MeV.



YIELD FROM THE JOINT FIT

+792.5 ± 280.2

1σ Hessian · signal yield in full analysis support

2016 · full data

m = 68.0 MeV · Template preset: Signal MC · ±2 core widths comparison

Signal MC · ±2 core width

Analysis binning 0.500 MeV

Exploratory rebin · fixed-kernel reconditioning

1 · Blind-window center

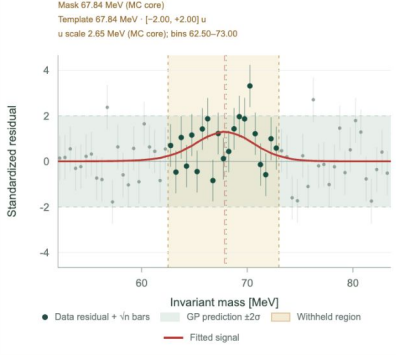
MC core

2 · Left width 2.00 u

Right width 2.00 u

Choose a center before editing widths. This changes the exclusion/fit mask; the signal template stays fixed.

Shown: MC core at 67.84 MeV; [−2.00, +2.00] u. One u = 2.651 MeV (MC core width). Withheld bin edges: 62.50–73.00 MeV.



YIELD FROM THE JOINT FIT

+16,278.5 ± 5,754.8

1σ Hessian · full selected signal yield

2021 · 10% data

m = 68.0 MeV · Signal MC · v16 TC · −4/+3 core widths

Signal MC · v16 TC · −4/+

Analysis binning 1.250 MeV

Exploratory rebin · fixed-kernel reconditioning

1 · Blind-window center

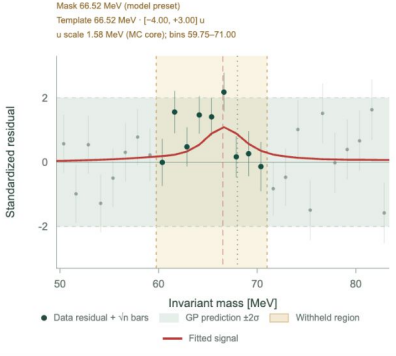
Use model preset

2 · Left width 4.00 u

Right width 3.00 u

Choose a center before editing widths. This changes the exclusion/fit mask; the signal template stays fixed.

Shown: model preset at 66.52 MeV; [−4.00, +3.00] u. One u = 1.585 MeV (MC core width). Withheld bin edges: 59.75–71.00 MeV.



YIELD FROM THE JOINT FIT

+21,350.8 ± 7,547.9

1σ Hessian · full selected signal yield

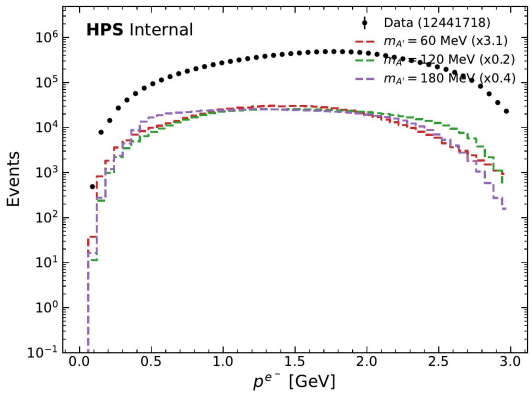
Unblinding 2021 Checklist

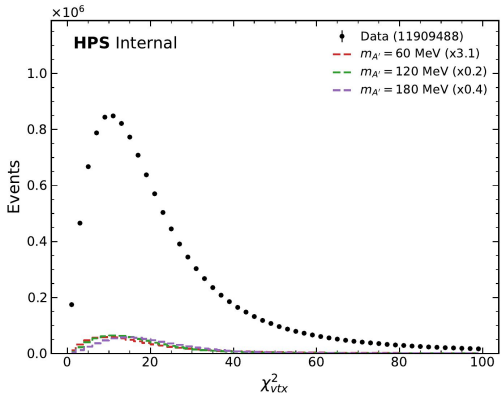
I. Finalize Event Selection
II. Background Model Validation
    A. Quality Fits/Signal Extraction
    B. Background Profiling
    C. Viable Search Window for 2021
III. Signal Model Choices
    A. Window Range, signal template
IV. Discovery Framework
    A. Asymptotic Likelihood
V. No siggie? No Biggie! – upper limits
    A. CLs to determine signal yield in individual datasets
    B. CLs to determine shared coupling when combining datasets

Note: Results may be independent of coupling strength → We are looking at peaking structure over a QED background for different incident beam energies. This can be input to a variety of frameworks. Asymptotic probabilities and independent experiment hypotheses should be published and considered meaningful.

Event Selection (2021) (Pre-Selection+)

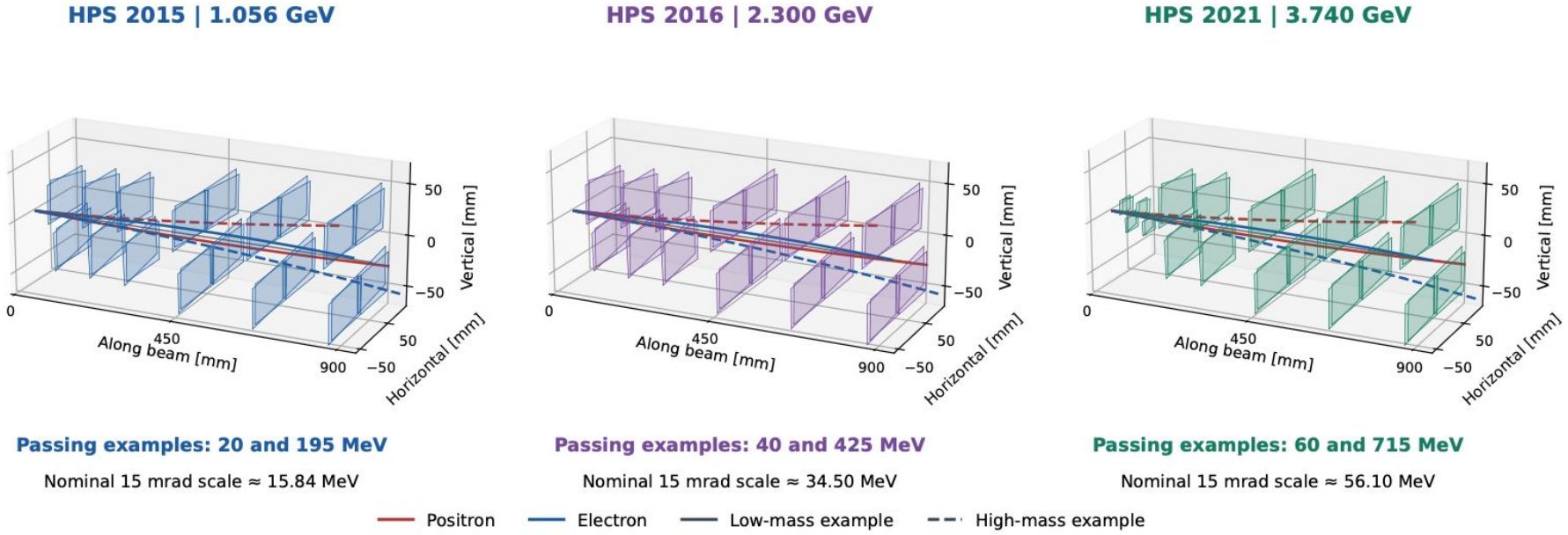
Selection	**2021 preselection + prompt sample** — Thanks Sarah!!
Trigger / topology	Single-2 or Single-3; one good vertex; E(positron cluster) > 0.2 GeV
Momenta [GeV]	p(track) > 0.4; p(e−) < 2.9; p(sum) > 2.8; p(vertex) < 4.0
Quality / timing	χ^2(track)/ndf < 20; electron: 8 hits (update); no vertex χ^2 cut; inherited timing





v5.0.5 preselection examples; these plots were not reprocessed with the updated eight-hit cut.

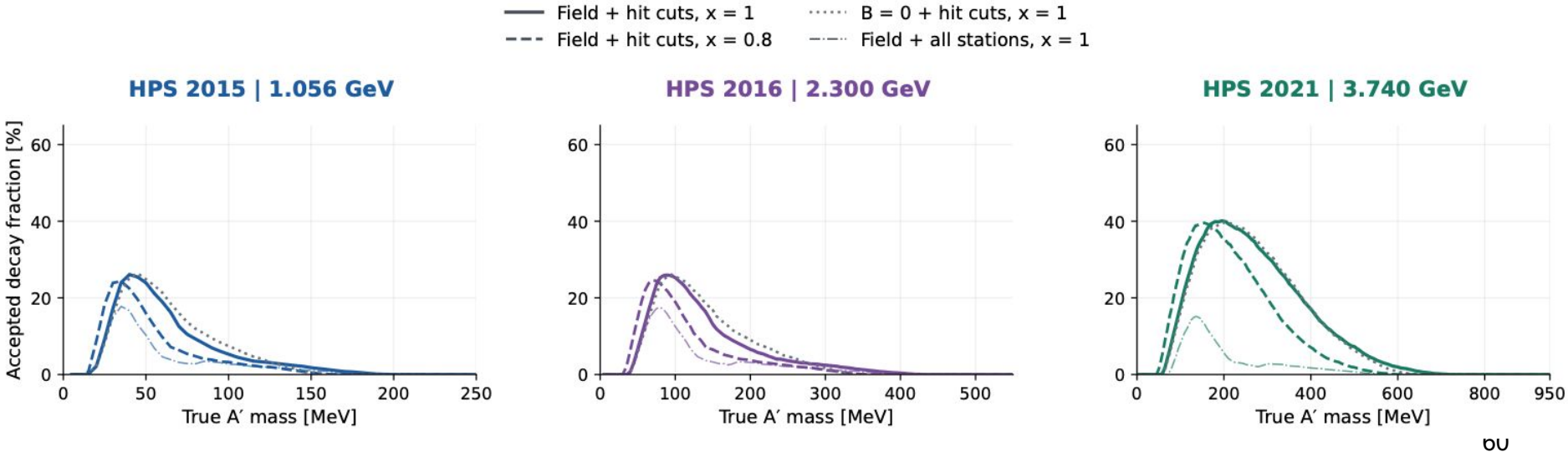
Acceptances by Geometry



Note goes into acceptances more thoroughly, will start search in 2021 100% at 56 MeV

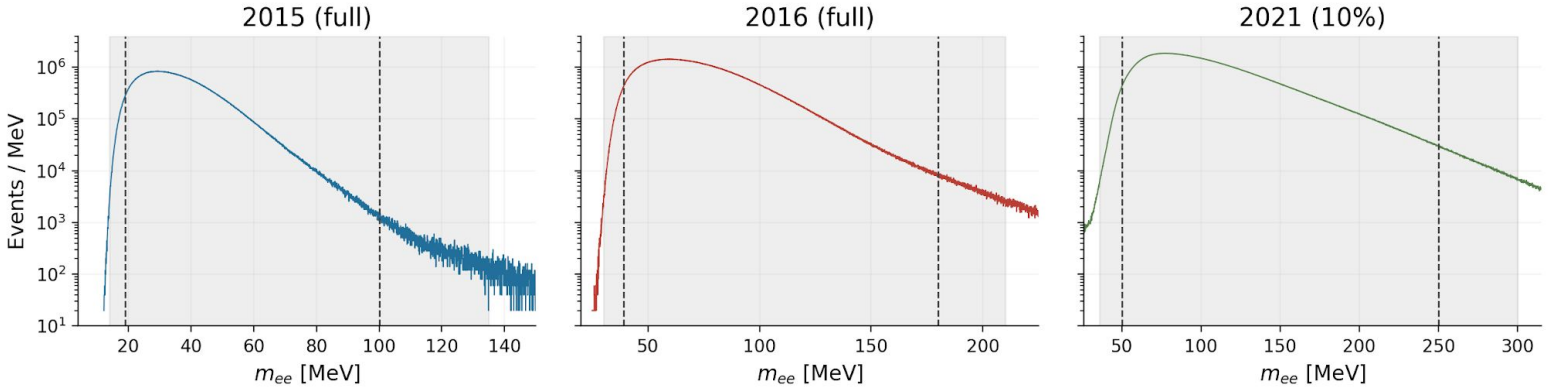
Acceptances by Campaign

Run	$B_y(\boldsymbol{r}_{\mathrm{map},0})$ [T]	Sampled masses [MeV]	Peak [MeV]	Peak fraction
2015	-0.2402	20–195	40	26.1%
2016	-0.5234	40–425	90	25.9%
2021	-0.8596	60–715	195	40.1%



HPS datasets for this phase

Run	Beam [GeV]	Sample	Search [MeV]	GP support [MeV]
2015	1.056	1170 nb^{-1}	19–100	14–135
2016	2.3	10608 nb^{-1}	39–180	30–210
2021	3.74	10%	50–250	36–300



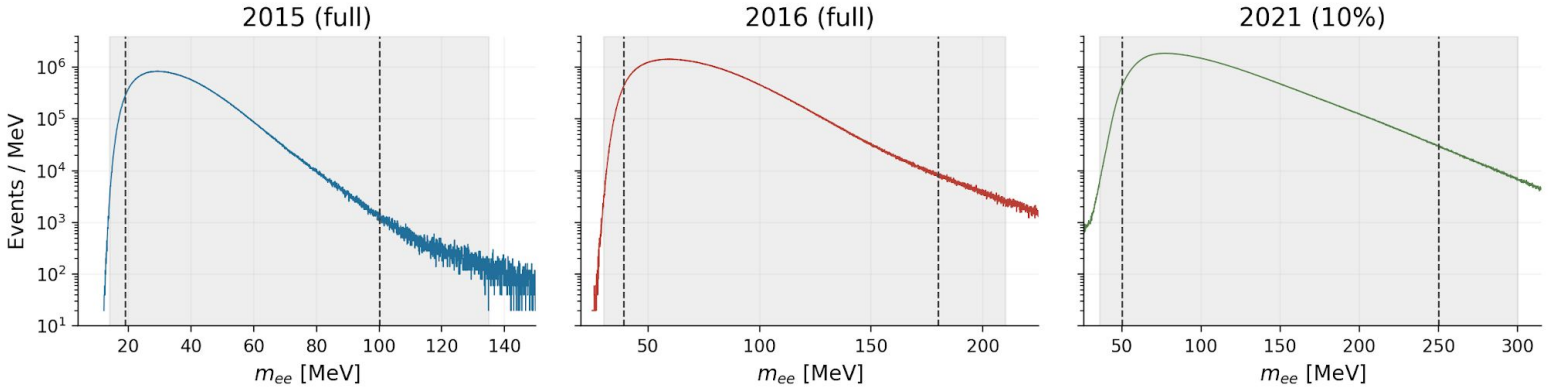
Gray: GP support. Dashed lines: search boundaries.

HPS datasets for this phase

Run	Beam [GeV]	Sample	Search [MeV]	GP support [MeV]
2015	1.056	1170 nb^{-1}	19–100	14–135
2016	2.3	10608 nb^{-1}	39–180	30–210
2021	3.74	10%	50–250	36–300

Updated ranges
2021 support: 36–300 MeV
2015 search: 19–100 MeV

Fit support supplies sidebands; the search range sets tested masses.



Gray: GP support. Dashed lines: search boundaries.

Unblinding 2021 Checklist

I. ~~Finalize Event Selection~~
II. Background Model Validation
    A. Quality Fits/Signal Extraction
    B. Background Profiling
    C. Viable Search Window for 2021
III. Signal Model Choices
    A. Window Range, signal template
IV. Discovery Framework
    A. Asymptotic Likelihood
V. No siggie? No Biggie! – upper limits
    A. CLs to determine signal yield in individual datasets
    B. CLs to determine shared coupling when combining datasets

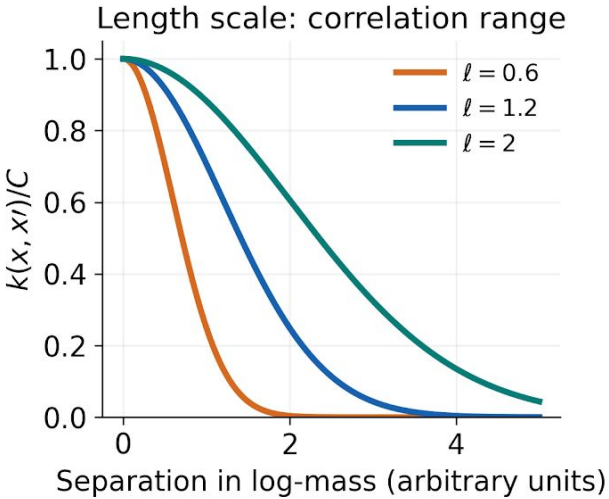
Background Modeling: GPR

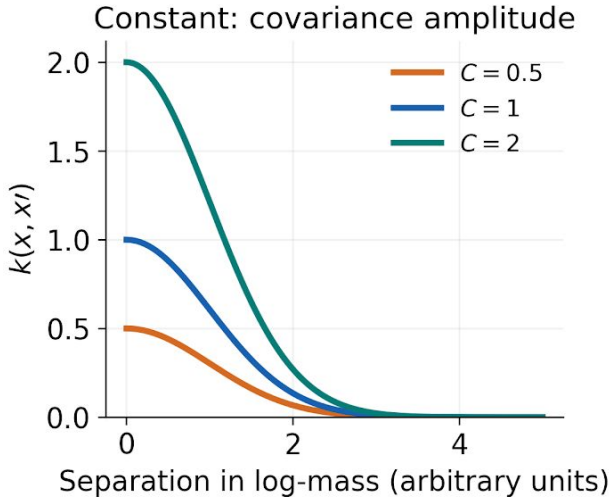
$$k(x, x') = C \exp\left[-\frac{(x - x')^2}{2\ell^2}\right]$$

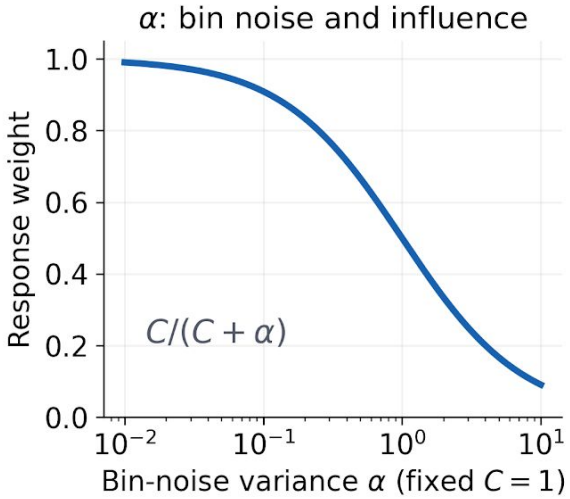
ℓ: smoothness along log mass

C: variance of the log background

$\alpha \approx 1/y$: noise variance of each bin







Background Profiling

Train on sidebands
Predict b and C_{eff}
Factor $C_{\text{eff}} = LL^T$

→

Test one yield
$\lambda = b + L\theta + Aw$
Hold b, L, w fixed

→

Profile the background
Minimize Poisson NLL
plus $\frac{1}{2}\theta^T\theta$

→

Build the profile
Record $\ell_p(A)$
Fit $\widehat{A}$ or test A_{90}

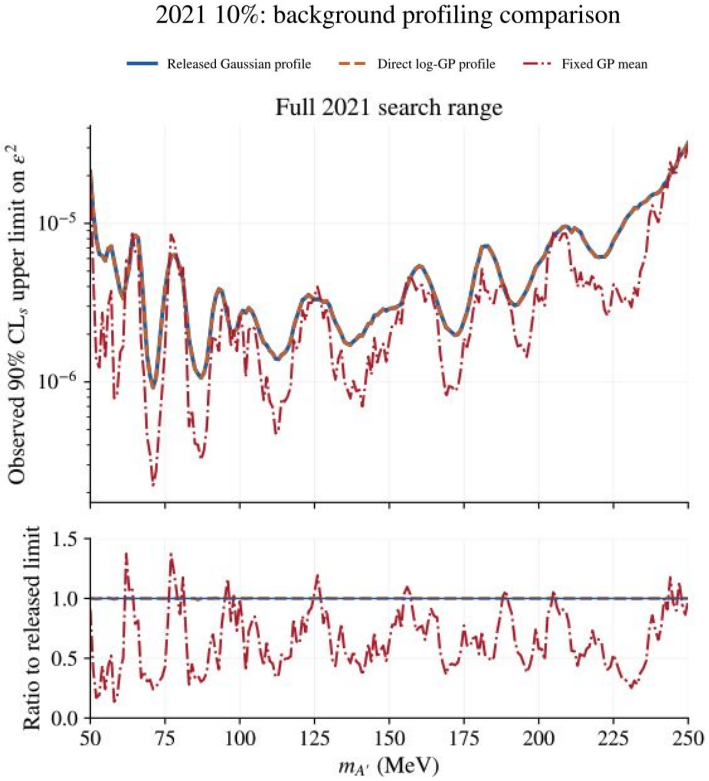
local Poisson mean

$$\boldsymbol{\lambda}_d(A_d, \boldsymbol{\theta}_d) = \boldsymbol{b}_d + L_d\boldsymbol{\theta}_d + A_d\boldsymbol{w}_d$$

projected bkg profiled bkg signal yield * template

profiled local likelihood

$$\mathcal{L}_d(A_d, \boldsymbol{\theta}_d) = \left[\prod_{i=1}^{N_d(m)} \text{Pois}(n_{d,i} \mid b_{d,i} + (L_d\boldsymbol{\theta}_d)_i + A_d w_{d,i}(m))\right] \exp\left(-\frac{1}{2}\boldsymbol{\theta}_d^\top\boldsymbol{\theta}_d\right)$$



Background Profiling

Train on sidebands
Predict b and C_{eff}
Factor $C_{\text{eff}} = LL^T$

→

Test one yield
$\lambda = b + L\theta + Aw$
Hold b, L, w fixed

→

Profile the background
Minimize Poisson NLL
plus $\frac{1}{2}\theta^T\theta$

→

Build the profile
Record $\ell_p(A)$
Fit $\widehat{A}$ or test A_{90}

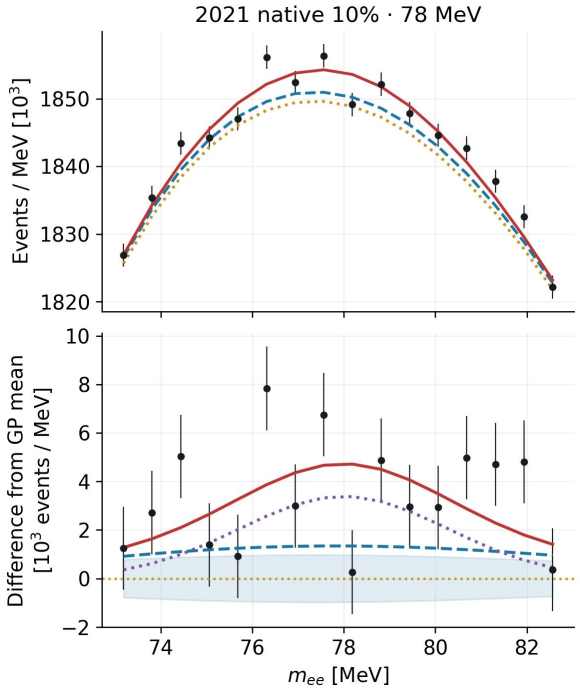
local Poisson mean

$$\boldsymbol{\lambda}_d(A_d, \boldsymbol{\theta}_d) = \boldsymbol{b}_d + L_d\boldsymbol{\theta}_d + A_d\boldsymbol{w}_d$$

projected bkg profiled bkg signal yield * template

profiled local likelihood

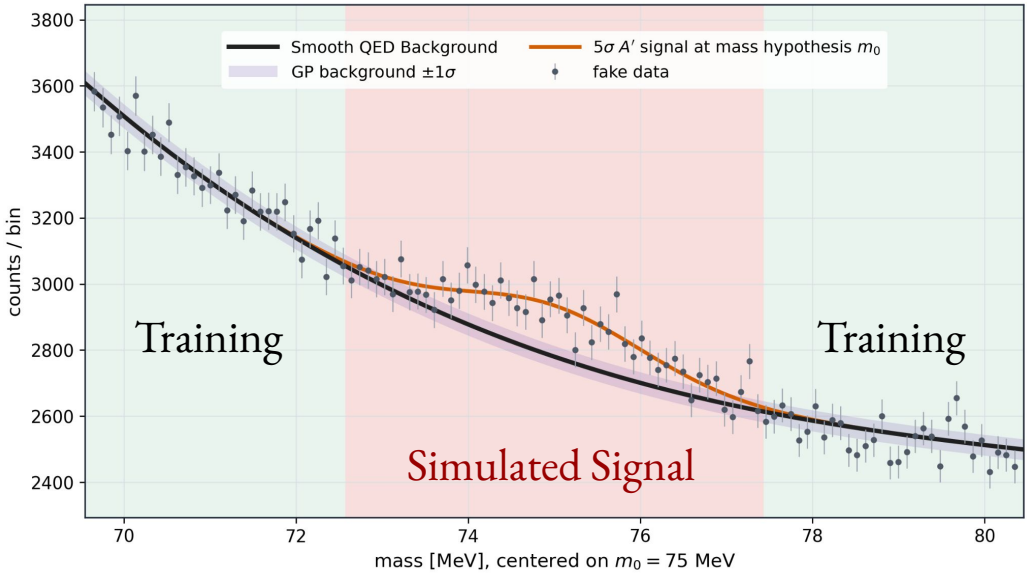
$$\mathcal{L}_d(A_d, \boldsymbol{\theta}_d) = \left[\prod_{i=1}^{N_d(m)} \text{Pois}(n_{d,i} \,|\, b_{d,i} + (L_d\boldsymbol{\theta}_d)_i + A_d w_{d,i}(m))\right] \exp\left(-\frac{1}{2}\boldsymbol{\theta}_d^\top\boldsymbol{\theta}_d\right)$$



Data: $n - b_{GP}$ GP baseline: 0
Background: $b_{prof} - b_{GP}$
Total fit: $b_{prof} + Aw - b_{GP}$ Signal: Aw
Shading: GP constraint width

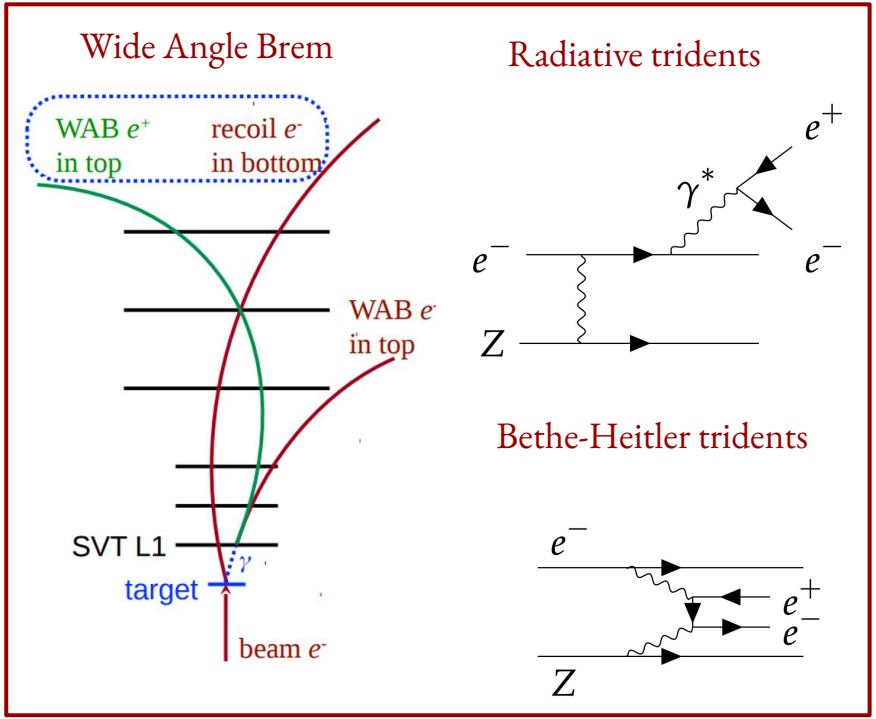
Prompt A′ Signal Model and Backgrounds

If A′ exists within the acceptance of HPS, it will present itself as a gaussian* excess above background in the IMD.



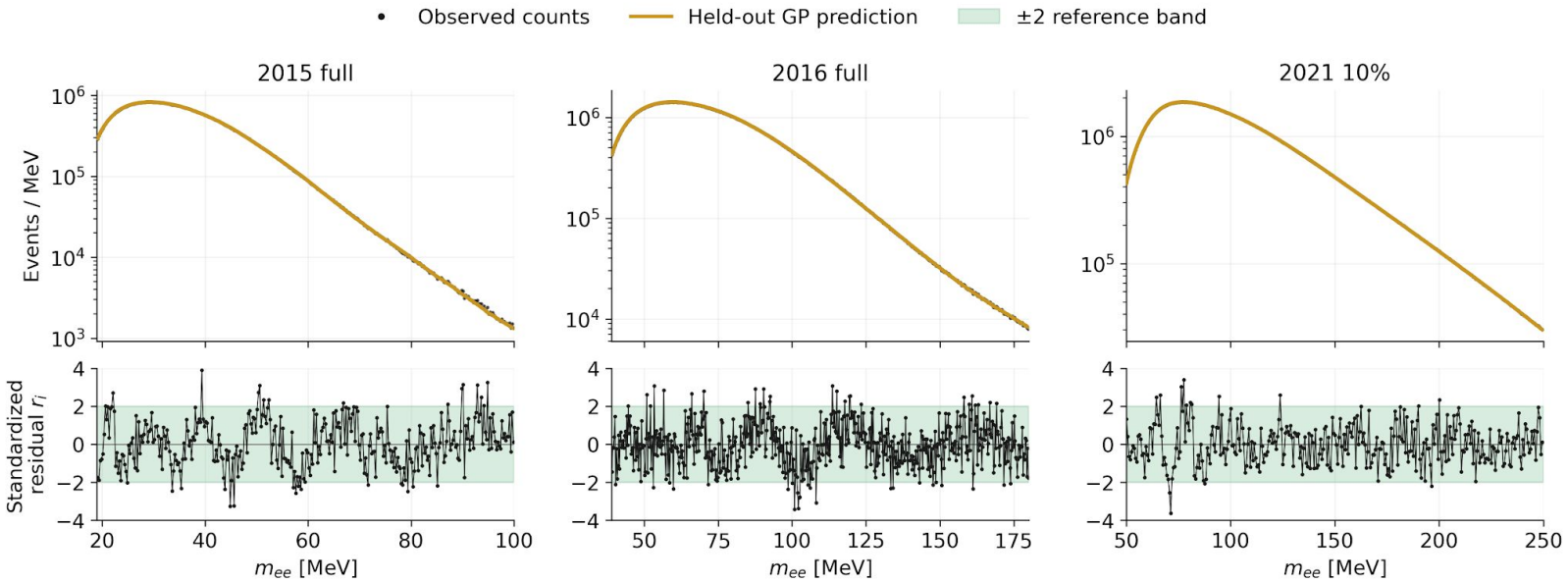
natural width of A′ << detector resolution
observed signal width = experimental mass resolution

Primary Backgrounds



GP predictions across the search regions

Each bin is predicted with its local ±2.25σ neighborhood excluded.



$$r_i = \frac{n_i - \hat{b}_i}{\sqrt{\hat{b}_i + C_{\mathrm{GP},ii}}}$$

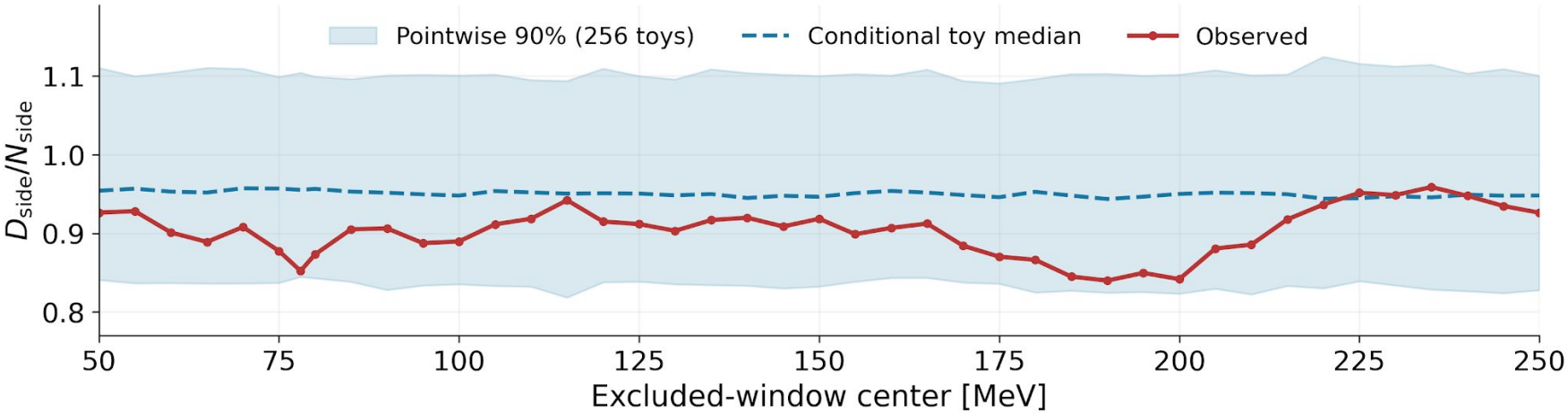
r_i: standardized data–GP difference in mass bin i.
Denominator: counting and GP uncertainty combined.
0 = agreement. Green: ±2. Bins are correlated.

Sideband fit quality

2021 10%: exclude ±2.25σ at 42 centers across 50–250 MeV

$$D_{\text{side}} = 2 \sum_{i \in \text{side}} \left[n_i \log\left(\frac{n_i}{\hat{b}_i}\right) - n_i + \hat{b}_i \right]$$

D_{side}: total Poisson discrepancy from the GP
N_{side}: number of scored sideband bins



256 poisson generated spectra from GP – full study included in next note version
Typical: near the toy median (~0.95), within the blue reference band.
Observed: 0.84–0.96.

Background validation strategy

Generate toys from analytic source fits

Fit datasets with functional forms; fluctuate background counts and add signal template.

Repeat the complete extraction

Fit the GP sidebands, then extract a signed signal yield from background-only and injected toys.

Check bias, uncertainty and recovery

Use mean pulls, pull widths and paired injected-yield response; inspect threshold and support dependence.

Pseudoexperiments information

2021 functional form **fShiftSigPowTail** (excluding 65 MeV region)

$$B_d(m) = N_d S(m; t_d, w_d)(m - m_{0,d})^{a_d} \exp\left[-\frac{m - m_{0,d}}{\theta_d} + c_{1,d}u_d + c_{2,d}u_d^2\right]$$

$$S(m; t, w) = [1 + \exp(-(m - t)/w)]^{-1} \qquad u_d = (m - m_{\mathrm{mid},d})/m_{\mathrm{scale},d}$$

2015/2016 functional form **fGenGammaThresh**

$$G_d(m) = N_d S(m; t_d, w_d)(m - m_{0,d})^{a_d} \exp\left[-\left(\frac{m - m_{0,d}}{\lambda_d}\right)^{p_d}\right]$$

Pseudoexperiments information

2021 functional form **fShiftSigPowTail** (excluding 65 MeV region)

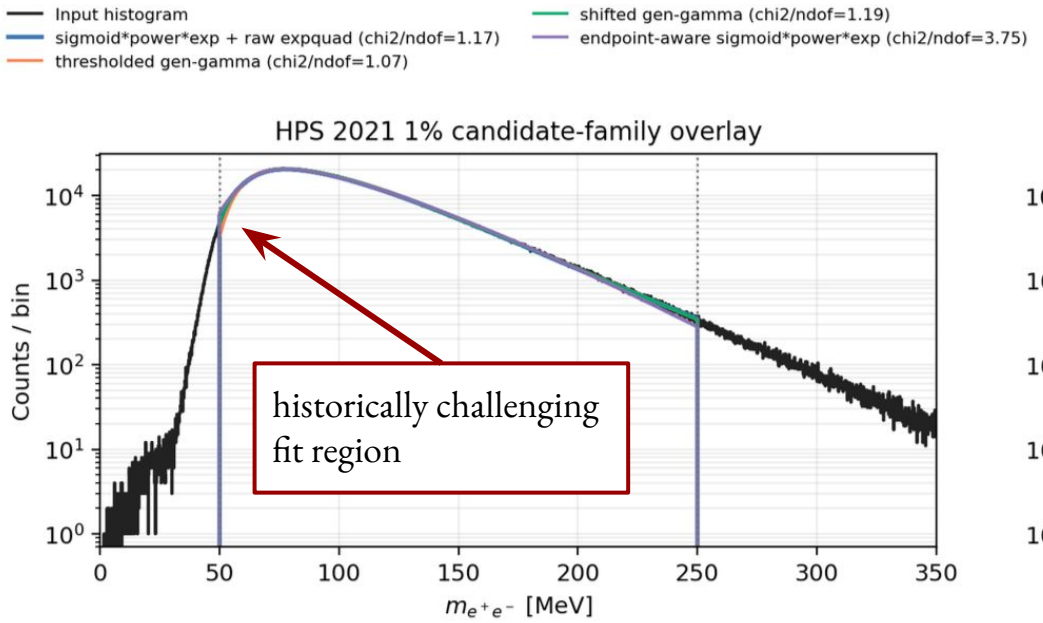
$$B_d(m) = N_d S(m; t_d, w_d)(m - m_{0,d})^{a_d} \exp\left[-\frac{m - m_{0,d}}{\theta_d} + c_{1,d}u_d + c_{2,d}u_d^2\right]$$

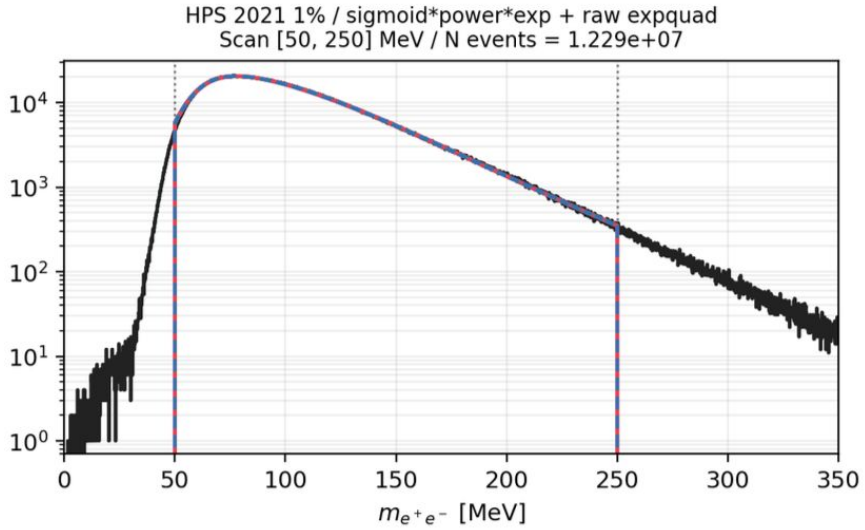
$$S(m; t, w) = [1 + \exp(-(m - t)/w)]^{-1} \qquad u_d = (m - m_{\mathrm{mid},d})/m_{\mathrm{scale},d}$$

Four different statistical validation tests conducted

I. 1% Fit * 10 (to double check 10% decisions)
II. 1% Fit * 100 (to predict 100%)
III. 10% Fit (to validate modeling on 10% data compare with scaled 1%)
IV. 10% Fit * 10 (to predict 100% and compare with scaled 1%)

Example Fits to 1% Dataset





Pseudoexperiments

Validation metrics

$$p_t(z) = \frac{\hat{A}_t(z) - A_{\mathrm{inj},\,t}(z)}{\sigma_{A,\,t}(z)}$$

Â: fitted yield; Ainj: injected yield.

σA: this toy's post-injection fit error.

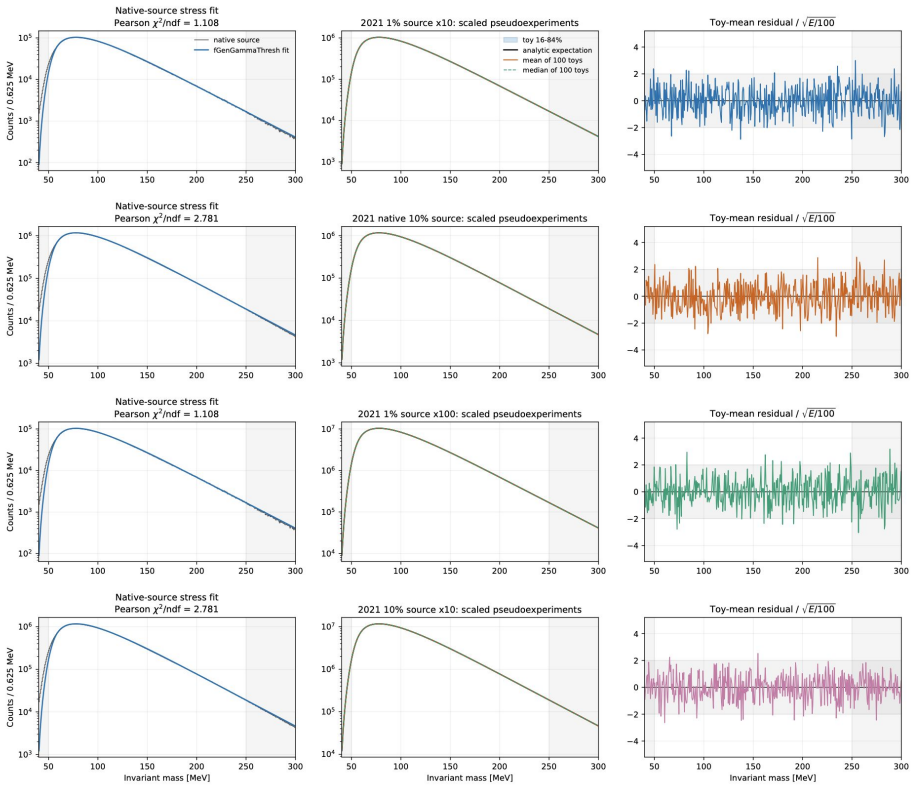
Pull mean → yield bias

- Target: mean near 0.

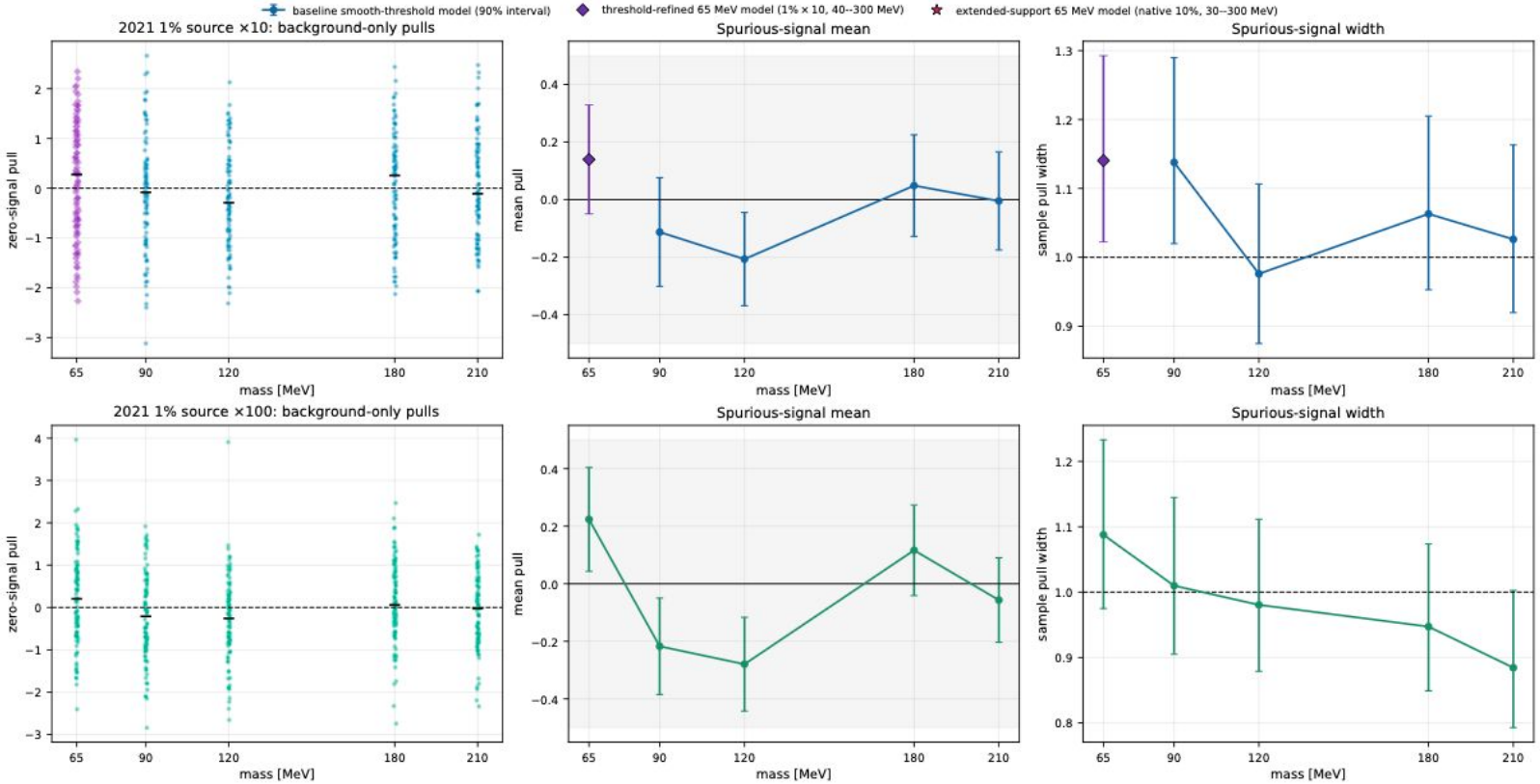
Pull width → uncertainty scale

- Target: width near 1.

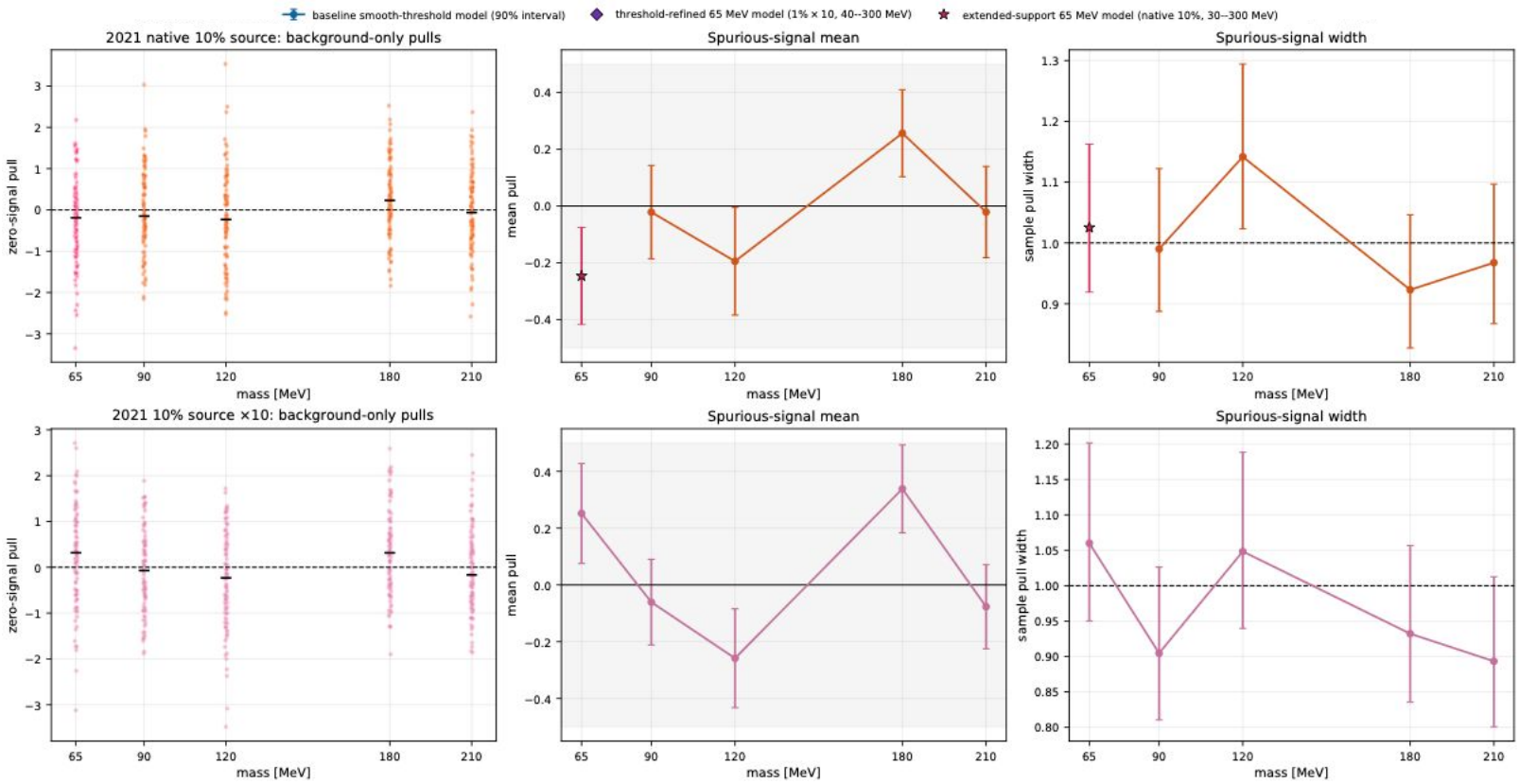
Test zero-signal toys and 1, 3, 5 σ injections.



BKG Only Pseudoexperiments (1% Source)

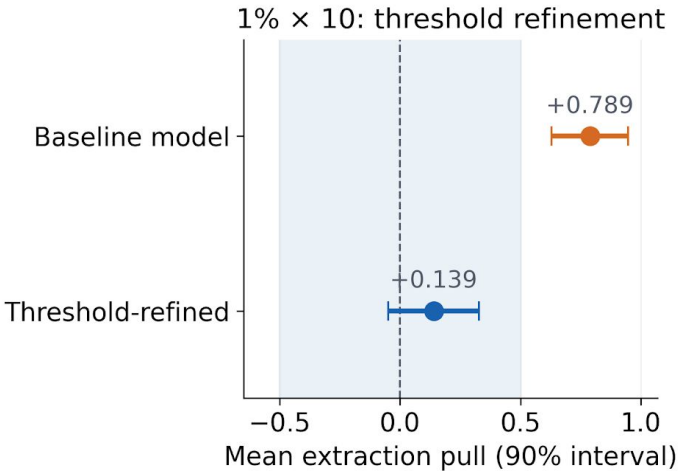


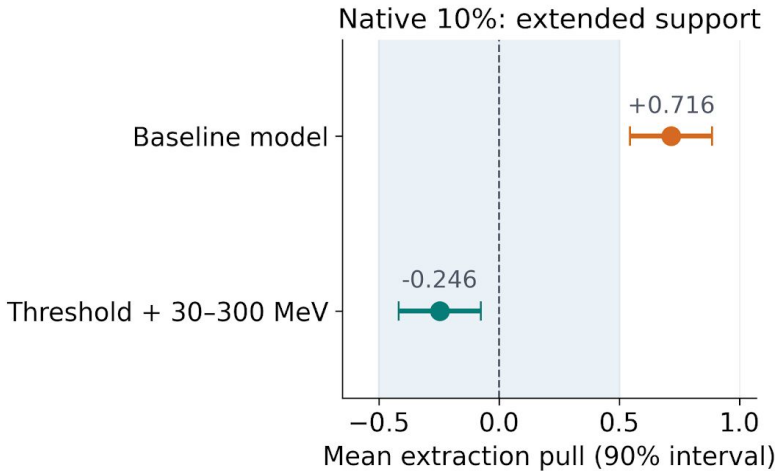
BKG Only Pseudoexperiments (10% Source)



Explanation of threshold modeling and alternate form

- 1% × 10: scale the fitted 1% mean; degree-five threshold continuation, with 40–300 MeV GP support.
- Native 10%: fit the saved 10% source; degree-six threshold continuation and 30–300 MeV support.





Signal Injection

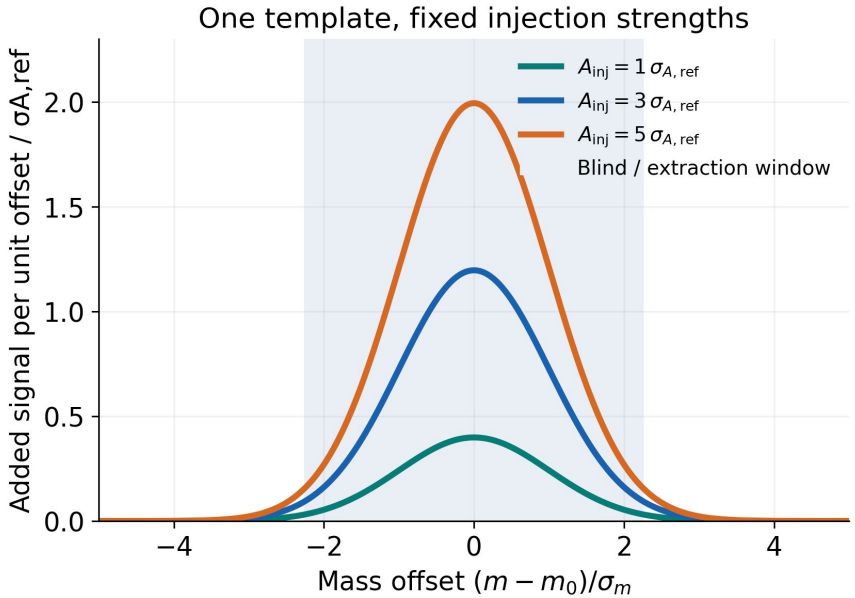
Choose the yield before adding signal

$$A_{\text{inj}}(z) = z\,\sigma_{A,\text{ref}}, \quad z \in \{0, 1, 3, 5\}$$

$\sigma_{A,ref}$ is the yield error from this same background toy and mass, before injection.

Add z × $\sigma_{A,ref}$ events with the full Gaussian template. Reuse the same background toy at each strength.

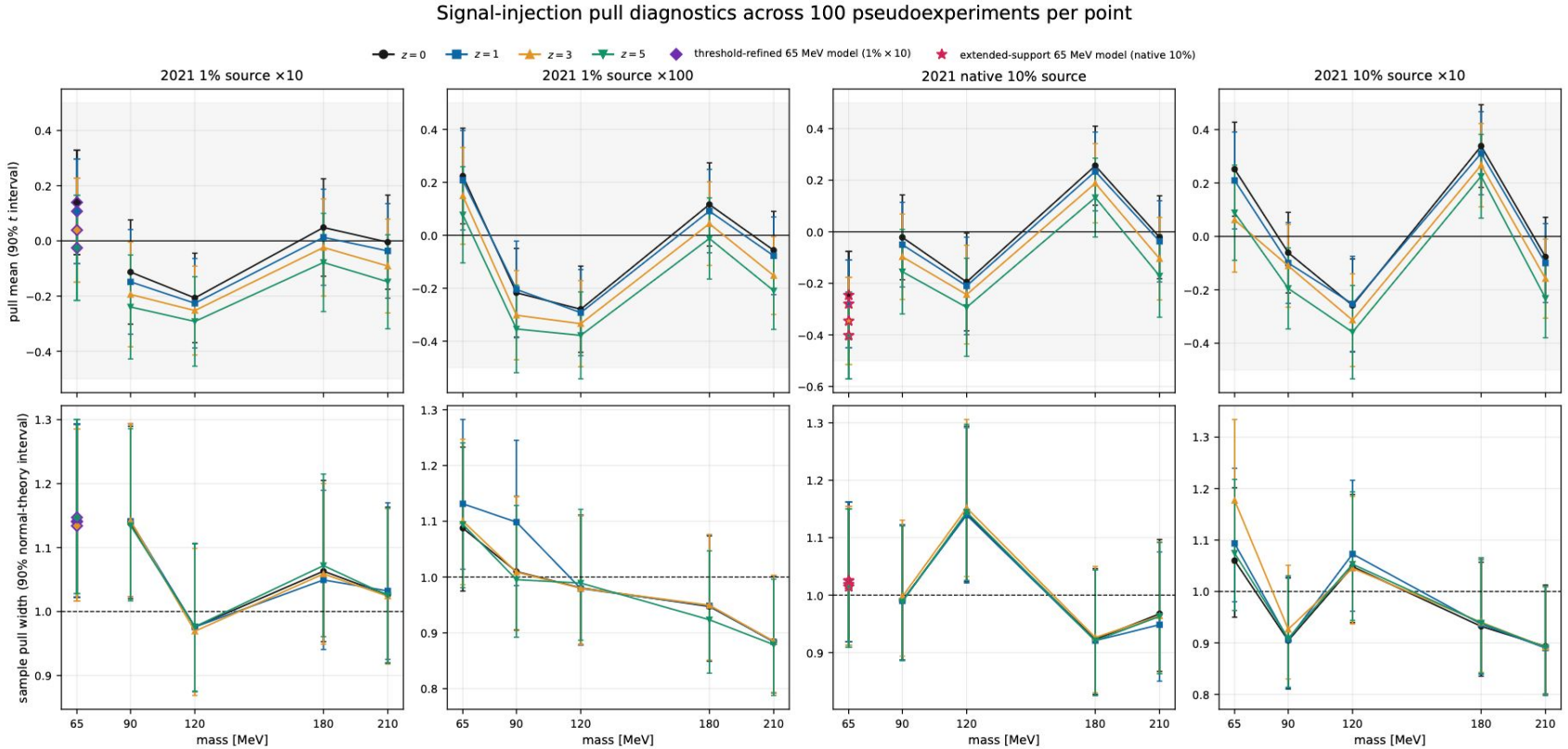
Refit the sidebands. Allow the fitted signal yield to be positive or negative.



Illustration; area under each curve gives its injected yield.

- z sets the injected yield. The measured significance comes from the new fit.

Injected Signal Studies



Validation: evidence and remaining limits

Window and recovery

±2.25σ training/extraction; refined 65 MeV tests recover 96.6–96.9% of the injected amplitude.

Spurious fitted yield

20/20 background-only mean pulls lie within ±0.5

Returned uncertainty

pull-widths ~+-10% which is about what you want with 100 toys

Unblinding 2021 Checklist

I. ~~Finalize Event Selection~~
II. ~~Background Model Validation~~
   A. ~~Quality Fits/Signal Extraction~~
   B. ~~Background Profiling~~
   C. ~~Viable Search Window for 2021~~
III. ~~Signal Model Choices~~
   A. ~~Window Range, signal template~~
IV. Discovery Framework
   A. Asymptotic Likelihood
V. No siggie? No Biggie! – upper limits
   A. CLs to determine signal yield in individual datasets
   B. CLs to determine shared coupling when combining datasets

Discovery Statistic Formulation

Bounded profile-likelihood ratio (non-negative signal)

$$q_{0,d}(m) = -2\ln\frac{\mathcal{L}_d\left(A_d = 0, \hat{\hat{\boldsymbol{\theta}}}_d(0)\right)}{\mathcal{L}_d\left(\hat{A}_d, \hat{\boldsymbol{\theta}}_d\right)}$$

Profile the background in both hypotheses.

An excess gives $q_0 > 0$; a deficit is assigned $q_0 = 0$.

Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic formulae for likelihood-based tests of new physics. Eur. Phys. J. C, 71:1554, 2011

Observed Upper Limits

1. Test a yield A at one mass. Refit the background nuisance parameters at each A.
2. The likelihood ratio measures disagreement with A. Set $\tilde{q}_A = 0$ when the data prefer a larger yield.
3. CL_{s+b} and CL_b are tails of the same statistic under signal + background and background alone.
4. Form their ratio, CL_s. Increase A until $CL_s(A_{90}) = 0.10$ for the 90% upper limit.

$$\tilde{q}_{A_d}(m) = \begin{cases} 0, & \hat{A}_d^{\text{unc}} > A_d, \\ -2\ln \dfrac{\mathcal{L}_d\left(A_d, \hat{\hat{\boldsymbol{\theta}}}_d(A_d)\right)}{\mathcal{L}_d\left(\hat{A}_d^{\text{unc}}, \hat{\boldsymbol{\theta}}_d\right)}, & 0 \le \hat{A}_d^{\text{unc}} \le A_d, \\ -2\ln \dfrac{\mathcal{L}_d\left(A_d, \hat{\hat{\boldsymbol{\theta}}}_d(A_d)\right)}{\mathcal{L}_d\left(0, \hat{\hat{\boldsymbol{\theta}}}_d(0)\right)}, & \hat{A}_d^{\text{unc}} < 0, \end{cases}$$

$$\mathrm{CL}_b(A_d; m) = P\left(\tilde{q}_{A_d} \ge \tilde{q}_{A_d}^{\text{obs}} \,\middle|\, \text{bkg}\right)$$

$$\mathrm{CL}_{s+b}(A_d; m) = P\left(\tilde{q}_{A_d} \ge \tilde{q}_{A_d}^{\text{obs}} \,\middle|\, A_d + \text{bkg}\right)$$

$$\mathrm{CL}_s(A_d; m) = \frac{\mathrm{CL}_{s+b}(A_d; m)}{\mathrm{CL}_b(A_d; m)}$$

If the unrestricted best yield is negative, the denominator uses the physical A = 0 fit.

Signal yield and coupling limits

BEST yield relation

$$\epsilon^2_{\text{up}} = \frac{2\alpha N_f N^{\text{up}}_{\text{sig}}}{3\pi m_{A'} f_{\text{rad}} (dN_{\text{bkg}}/dm)}$$

N_{sig}^{up}: 90% signal-yield limit

f_{rad}: radiative background fraction

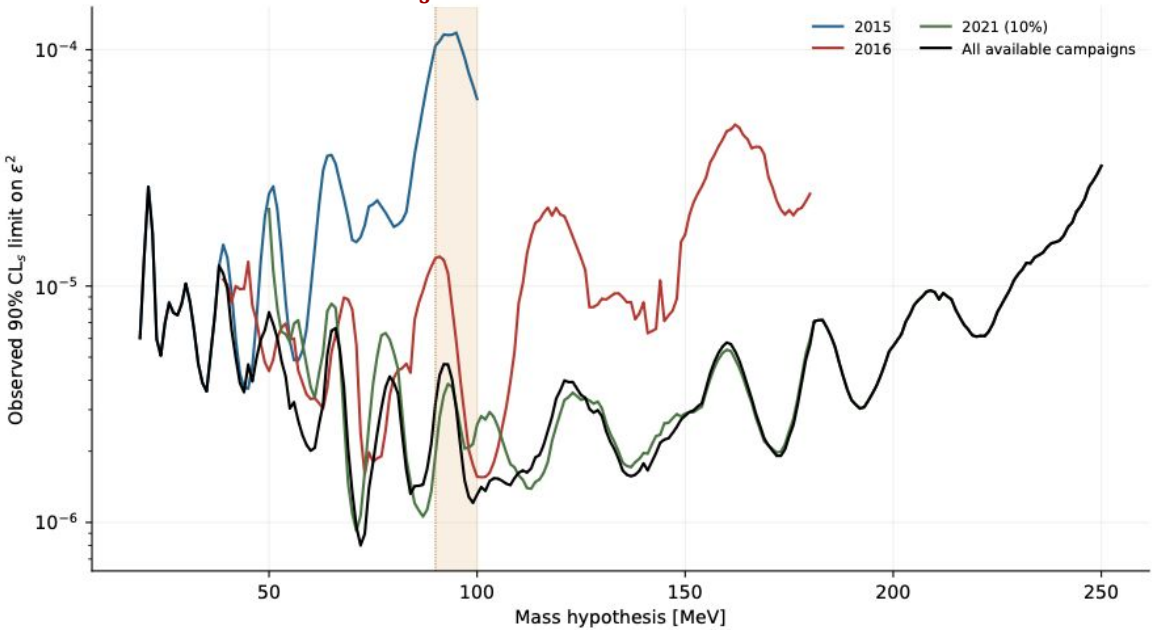
dN_{bkg}/dm: local background density

$N_f = 1 / BR(A' \rightarrow e^+e^-)$

α: fine-structure constant

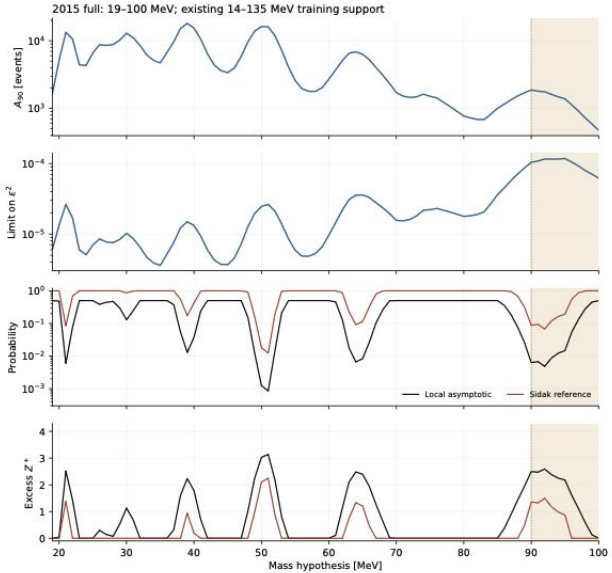
$m_{A'}$: tested mass

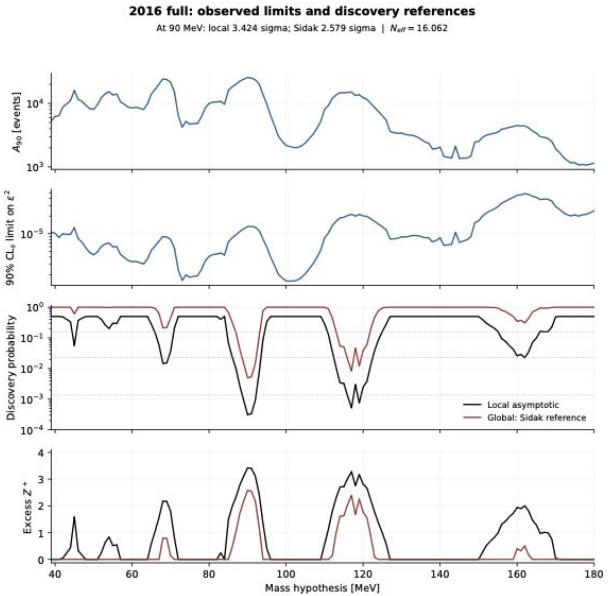
Observed 90% CL_s limits

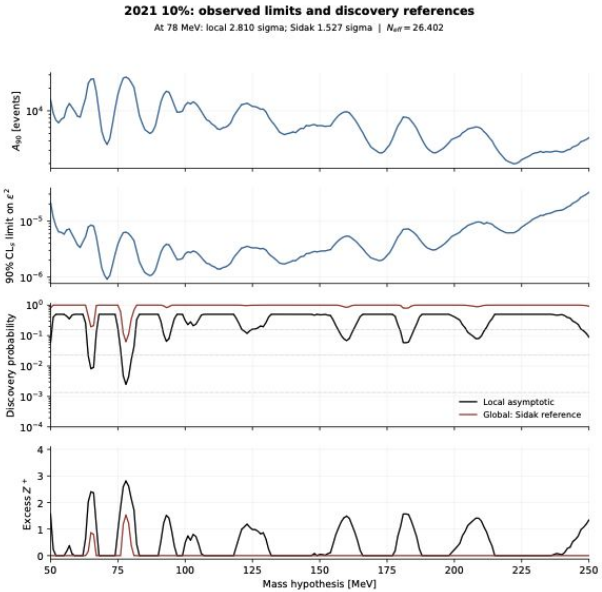


BEST: Bjorken et al., Phys. Rev. D 80, 075018 (2009).

Observed Limits and Discovery Pvalues for individual datasets







Combined discovery significance

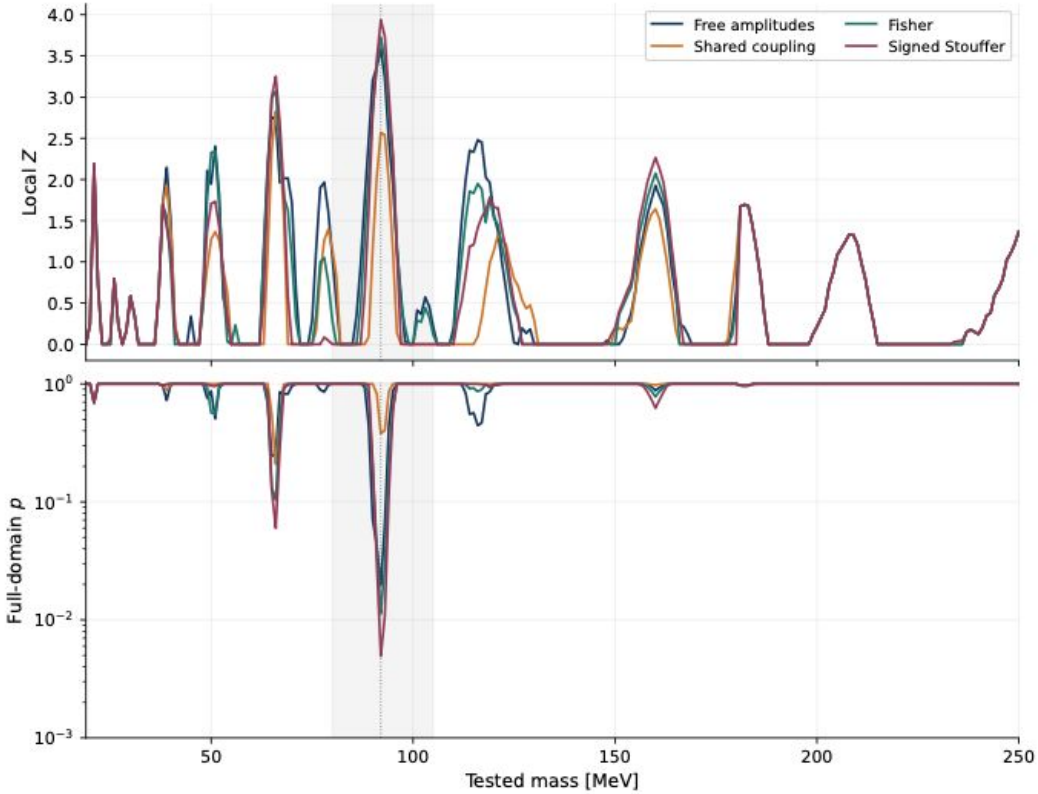
Combined Significance dependent on applied phenomenological framework

Shared Coupling:

- Local: 2.8 σ
- Global: 0.8-1.4* σ

Individual Experiments*:

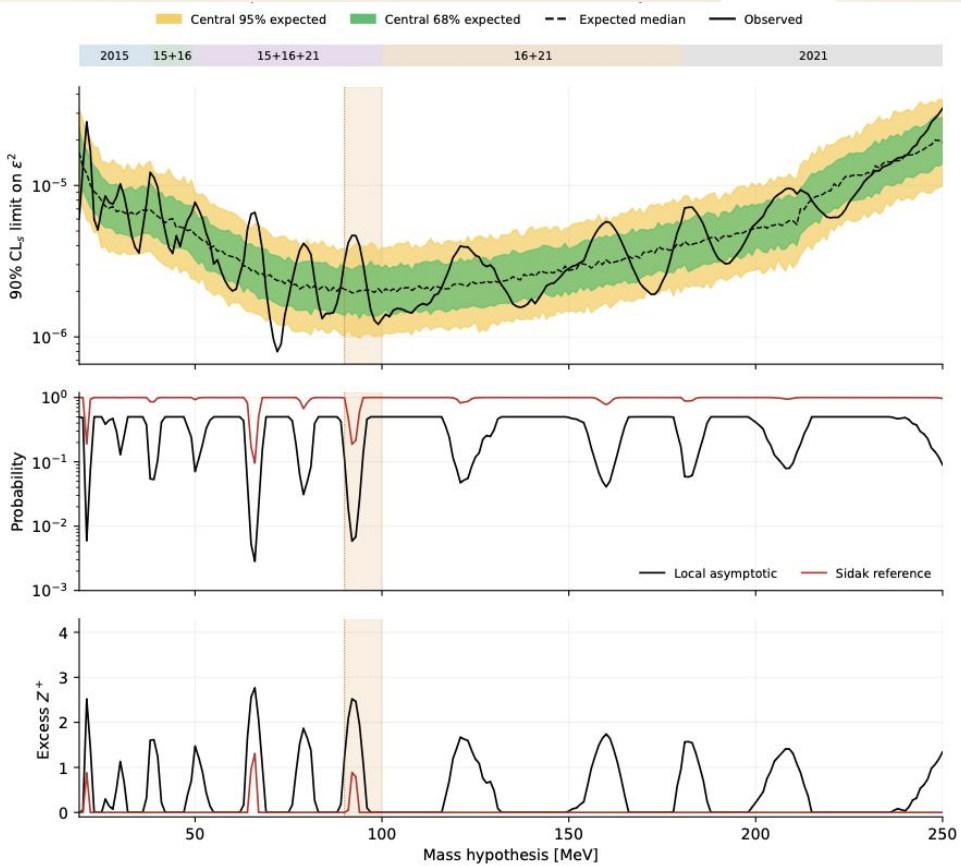
- Local: 3.6 - 3.9 σ
- Global: 2.1 - 2.6 σ



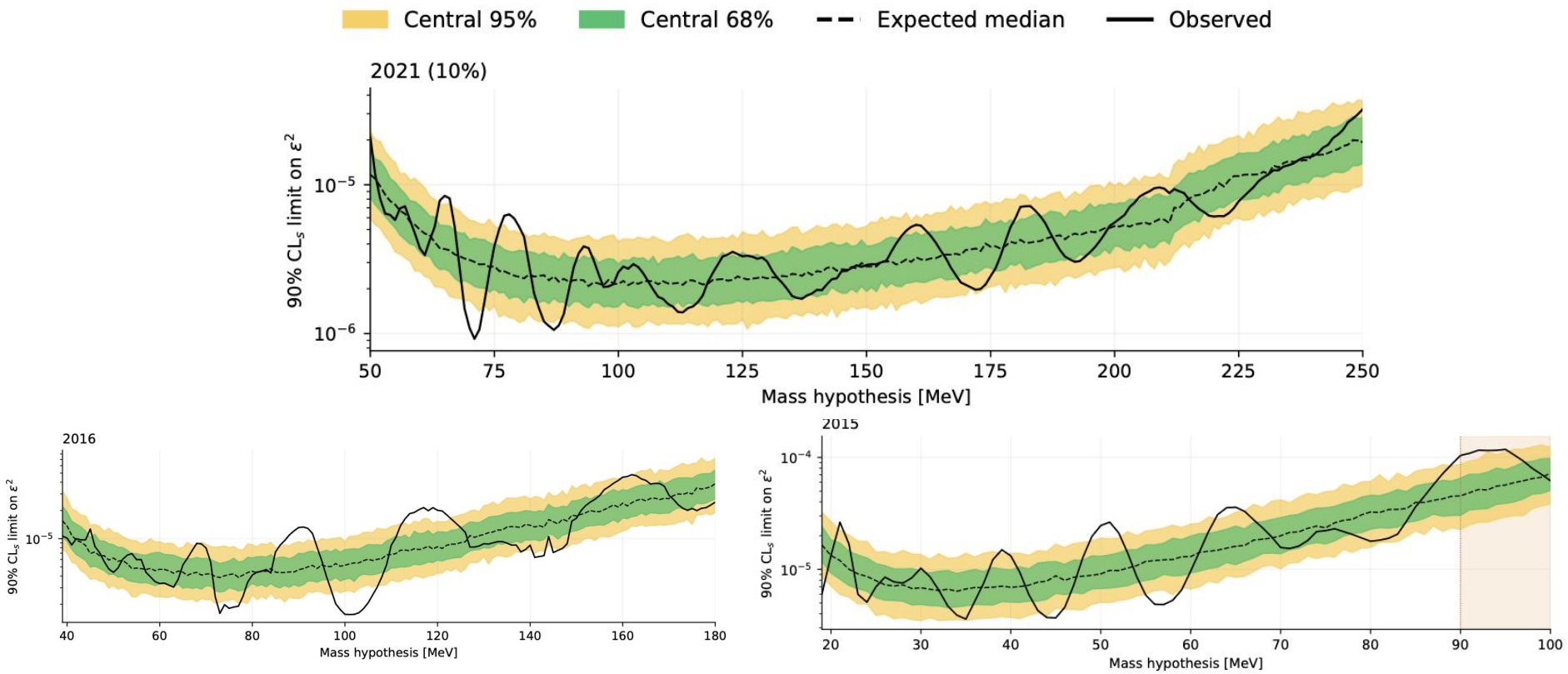
Combined Search Results with Uncertainty Bands

300 Toys for bands

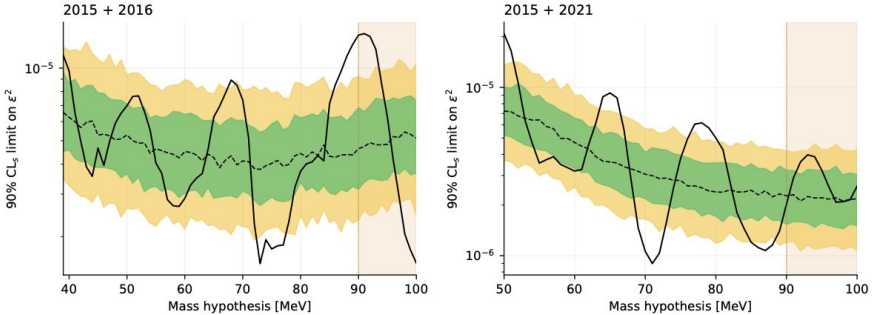
LEE penalty factor of ~35

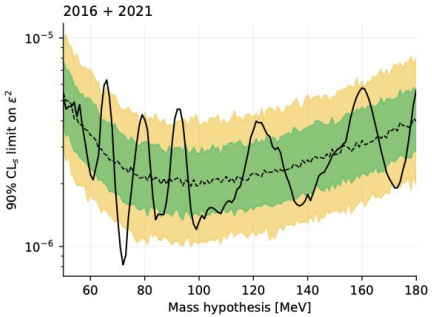


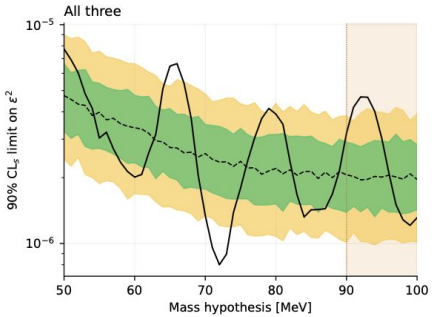
Individual Limits – gaussian signal

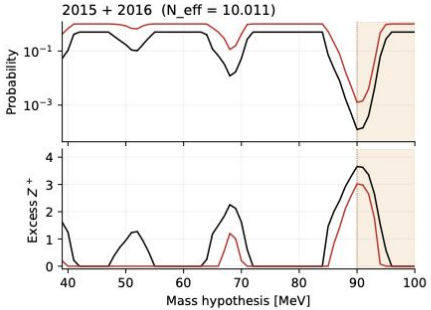


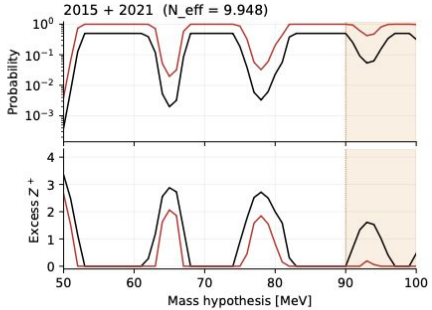
All Combined Regions

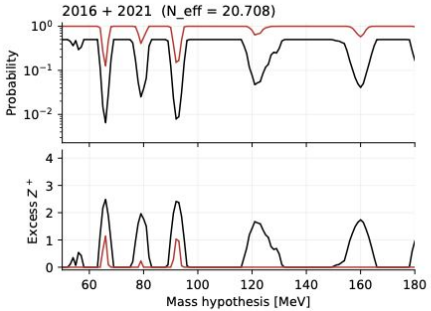


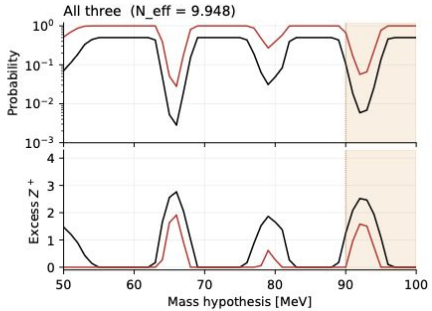












Unblinding 2021 Checklist

~~I. Finalize Event Selection~~
~~II. Background Model Validation~~
- ~~A. Quality Fits/Signal Extraction~~
- ~~B. Background Profiling~~
- ~~C. Viable Search Window for 2021~~

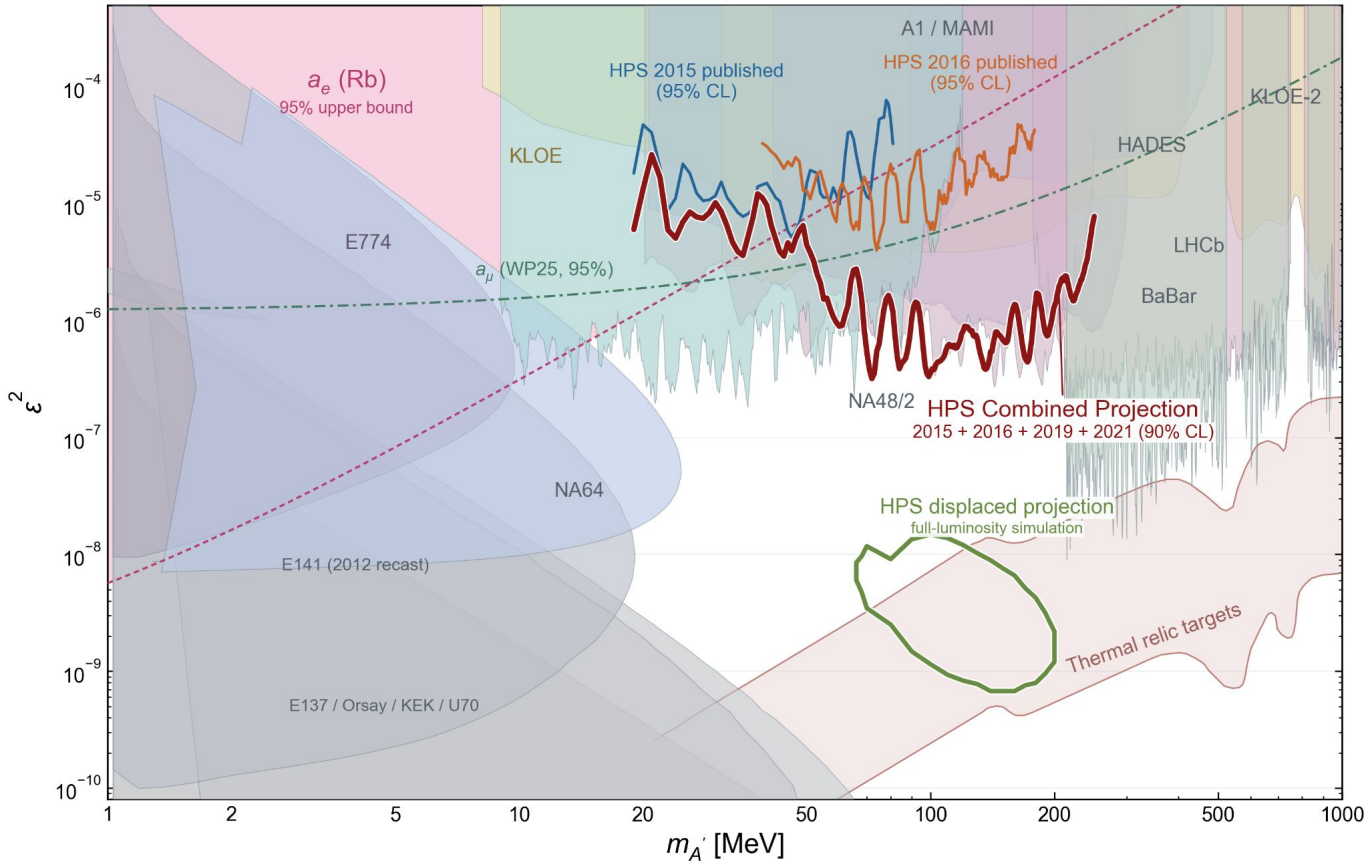
~~III. Signal Model Choices~~
- ~~A. Window Range, signal template~~

~~IV. Discovery Framework~~
- ~~A. Asymptotic Likelihood~~

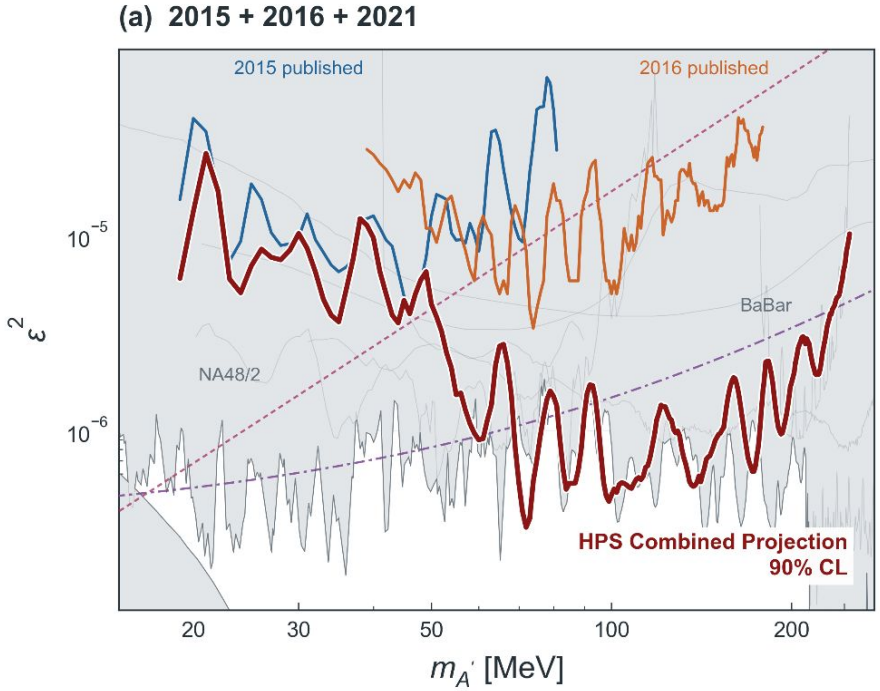
~~V. No siggie? No Biggie! – upper limits~~
- ~~A. CLs to determine signal yield in individual datasets~~
- ~~B. CLs to determine shared coupling when combining datasets~~

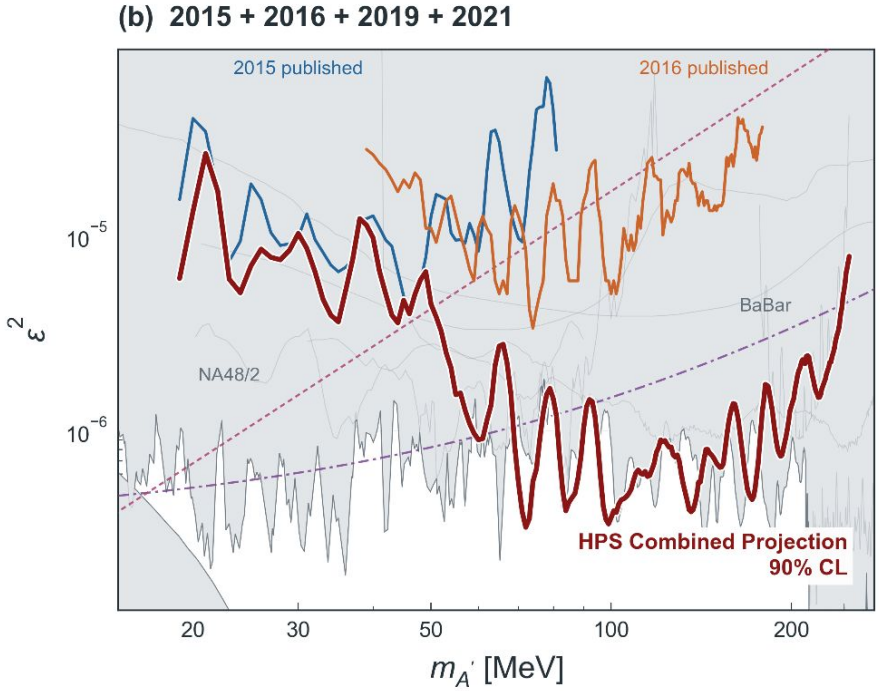
now i show projections + addl studies

Projected Contour

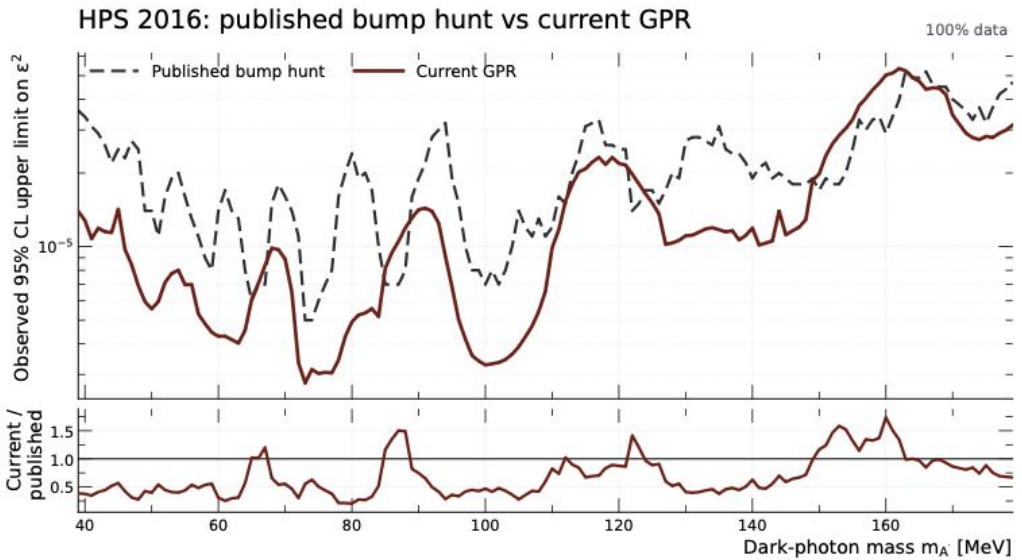


Projected Contours



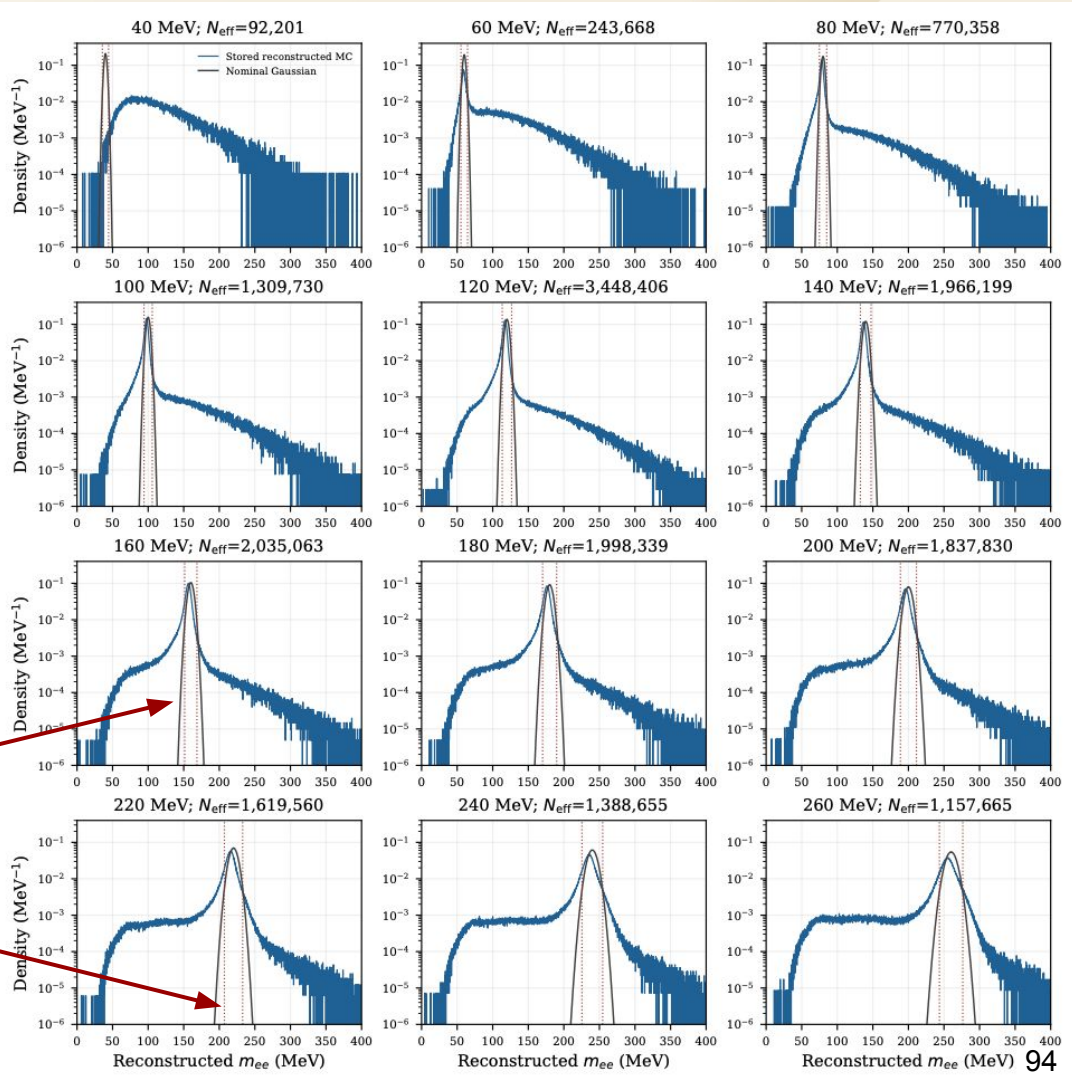


Observed Limit Comparison with 2016 Publication



Using Smeared MC Signal

Definitely deviates
from gaussian shape!

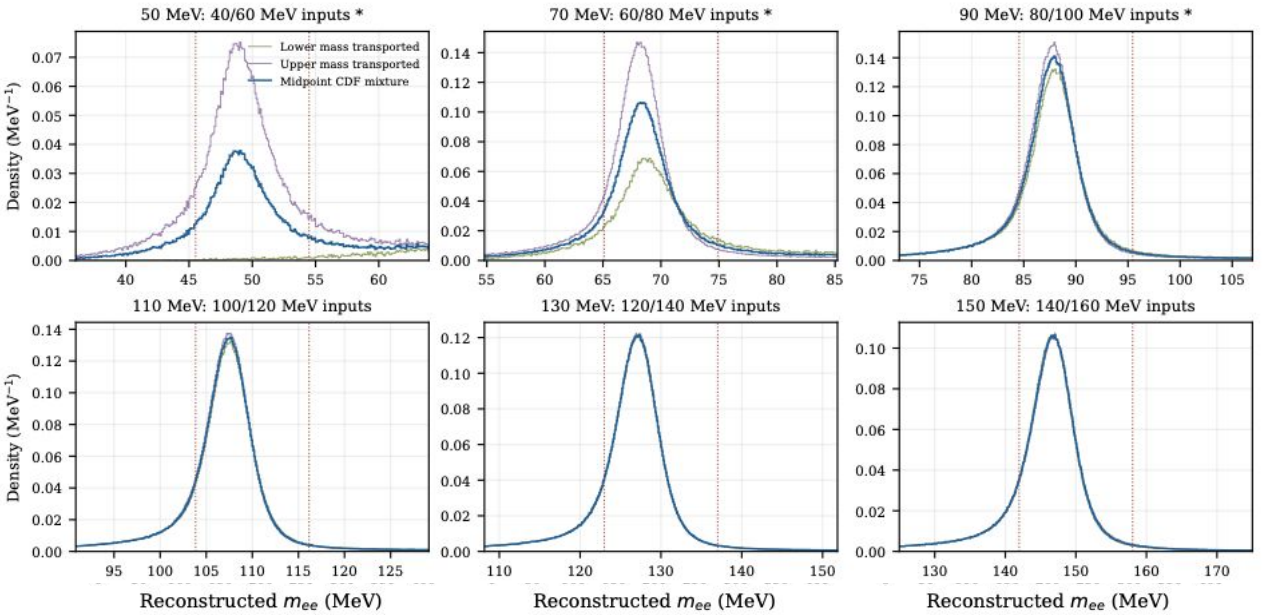


gaussian signal template

blind window

“Morphing” Signal Templates

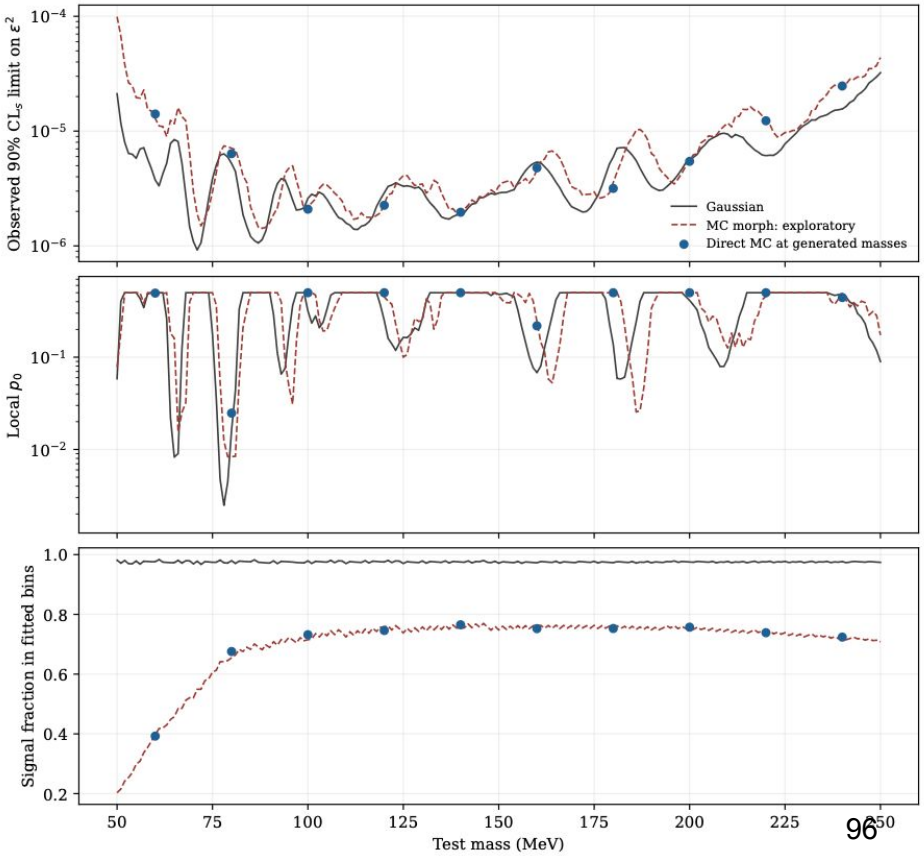
To have a signal template for every search hypothesis we use a CDF mixture between signal MC samples that exist



Upper limits using MC

Use same analysis approach with normalized MC template and +- 2.25 sigma blind window

- more leakage
  - may need to blind asymmetrically to account for different shapes
- need to truth match
- shows rough agreement with gaussian approach
  - shouldnt hold back from unblinding as we can do both MC signal and gaussian
  - other datasets would also need MC generated to do comparable search



Conclusions

Should be good to go regarding unblinding?

Gaussian signal shape for old datasets probably shouldnt be combined with signal shape from MC

- combined analysis can be done with gaussian for all three
- 2021 Result can show both signal shapes as they are not too different

Extra Slides

I. GP Correlations and expected upper limits
II. Additional Local Likelihood using GP on projected significance
III. 2021 Threshold fit optimization
IV. Signal Extraction Display
V. Mass Resolution Studies
VI. 2015 95% CLs Comparison
VII. Projection to Babar Plot
VIII. Two Signal Hypothesis
IX. Potential fit catalogues
X. Interpretation slides

Before Unblinding Sanity Checks – driving questions

SLAC

What would a signal look like in full 2021 dataset?

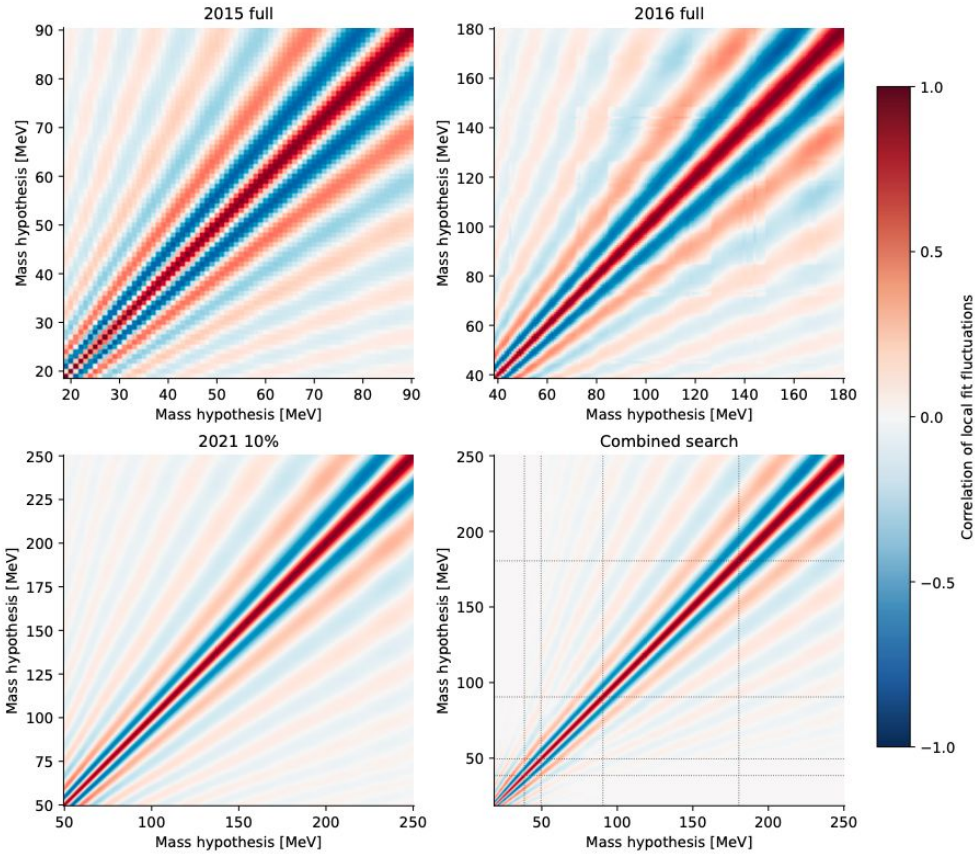
- must be able to map correlations if they exist in invariant mass
- must be able to characterize shape of observed upper limit with signal injected

What is the max significance a signal could be in 2021?

GPR Correlations in mass (1/n)

This mapping shows that for a local 1 sigma fluctuation there is a strong anti-correlation nearby that oscillates as you move away

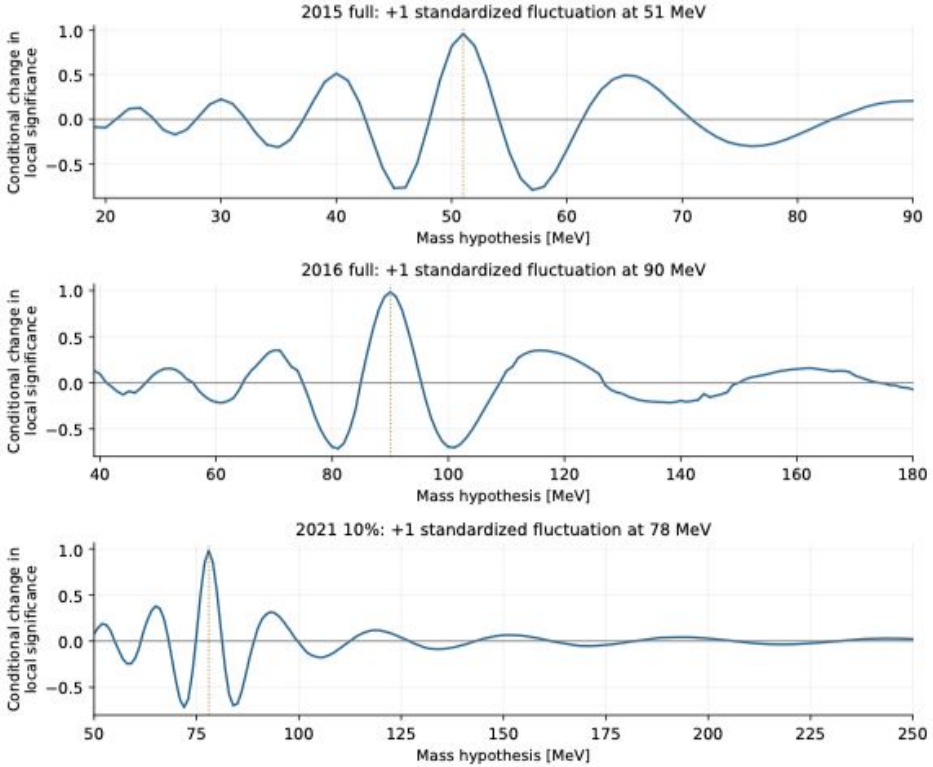
Makes sense!
Signal in a training sideband will of course spoil the training nearby



GPR Correlations in mass (1/n)

Size/shape of correlated "echoes" dependent on

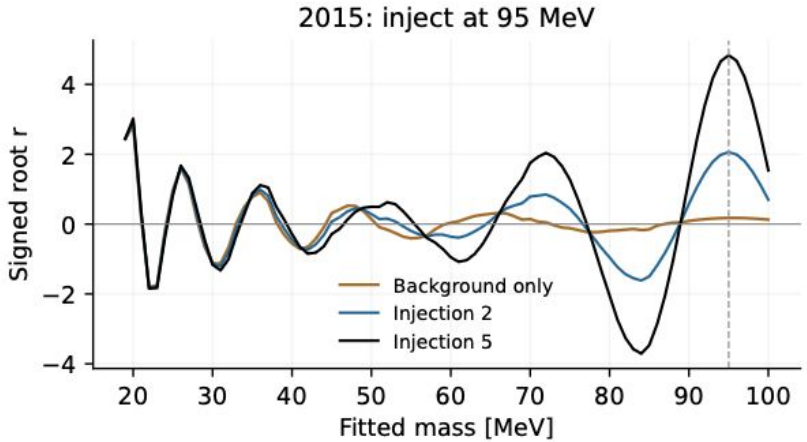
- mass resolution
- location of peak
- underlying background of nearby regions

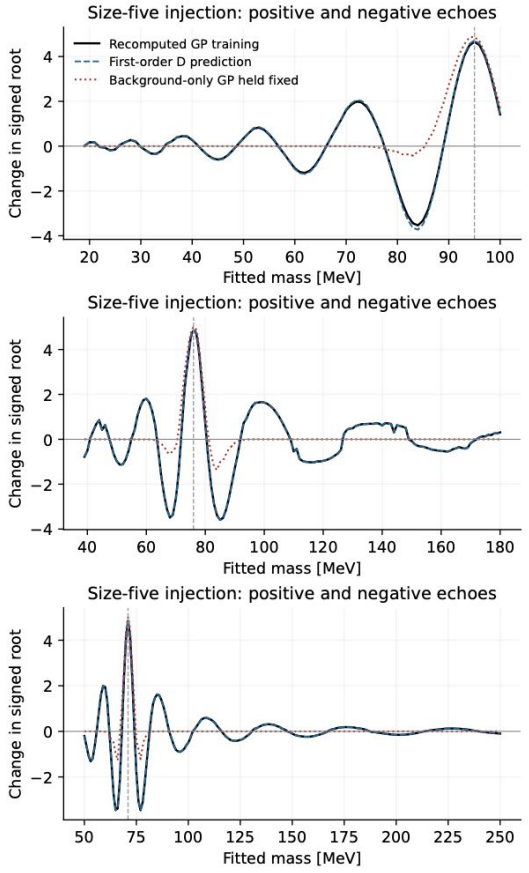


More on Echoes

First order prediction for location of features

$$\Delta r_m^{\text{lin}} = \sum_{i:B_i>0} D_{im} \frac{\delta n_i}{\sqrt{B_i}}$$





Projection of max signal if regions of interest are true signal

$$S_{100\%} = kS_f, \qquad B_{100\%} = kB_f, \qquad Z_{100\%} \simeq \frac{kS_f}{\sqrt{kB_f}} = \sqrt{k}\, Z_f$$

$$\boxed{Z^{\text{target}}_{1\%\to 100\%} = 10 Z_{1\%}} \qquad \boxed{Z^{\text{target}}_{10\%\to 100\%} = \sqrt{10}\, Z_{10\%}}.$$

Predicting Echoes for Unblinded 2021

SLAC

Conditional persistence test: one region at a time

- Scale the fitted 10% continuum to full exposure (k = 10).
- Choose a signal yield whose Asimov fit reaches $\sqrt{10}\ Z_{10}\%$.
- Draw 20 Poisson spectra; rerun the moving-window GP and extraction.

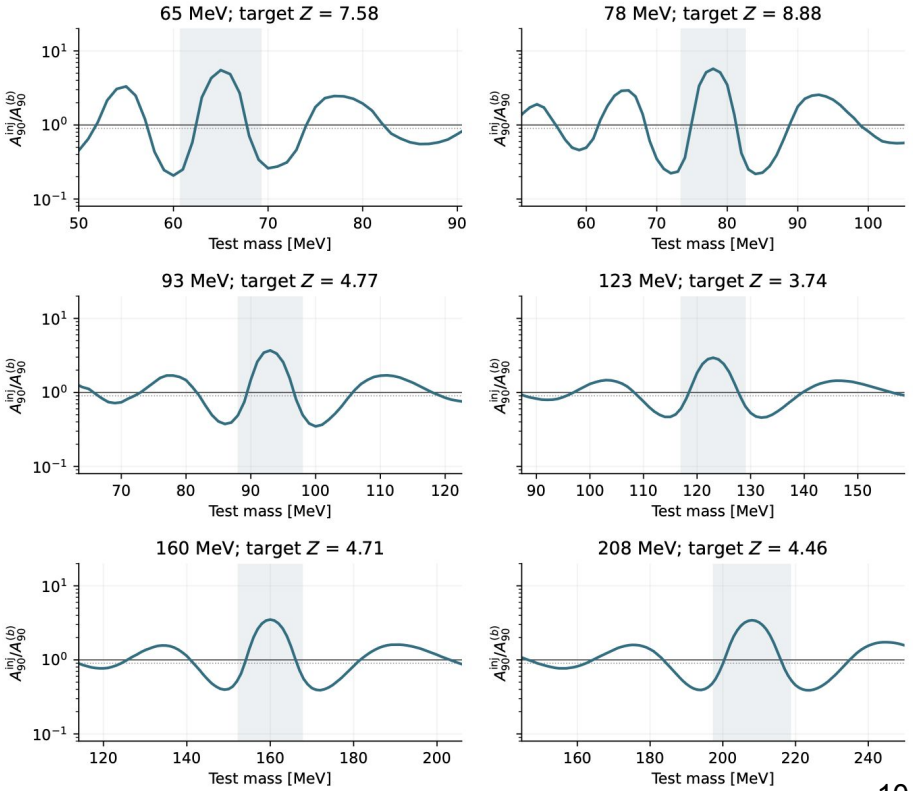
R compares injected and continuum limits. R < 1 means tighter neighboring limits; D measures the echo depth.

$$R(m) = \frac{A_{90}^{\text{injected}}(m)}{A_{90}^{(b)}(m)}, \qquad D = 1 - R(m_{\text{echo}}), \qquad m_{\text{echo}} \in m_0 \pm [2, 8]\, \sigma_m(m_0)$$

Predictions of upper limit structure

Have proven ability to predict observed upper limit structure to first order

projected shape of upper limits for different regions of interest in full 2021 dataset



Local Likelihood using GP on projected significance: pt1

Model the signed-root scan response

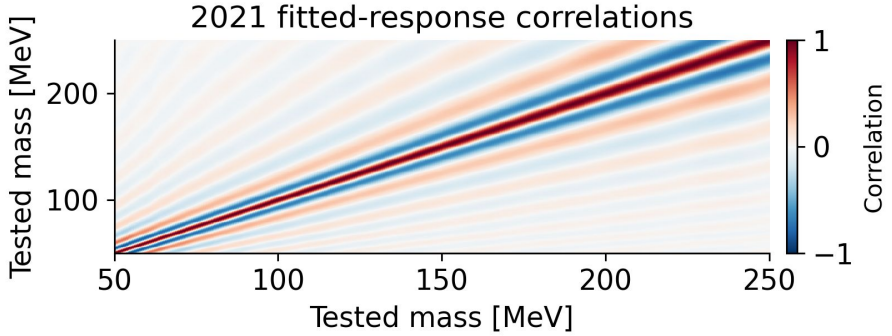
$$r^* \simeq a + D^T \xi, \quad \xi \sim N(0, I)$$

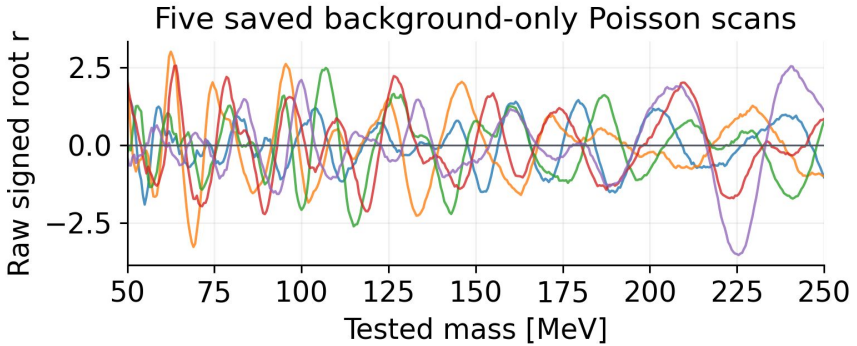
$$\Sigma_r = D^T D, \quad R_{ij} = \frac{\Sigma_{r,ij}}{s_i s_j}$$

D maps bin fluctuations into the signed-root mass scan.

Retain the response mean a, scale s and mass correlation R.

Draw correlated fields and record their largest excess.





Ananiev & Read, JINST 18 (2023) P05041; HPS v5.8.5.3.
2021 native 10%; ±2.25σ; 50–250 MeV; frozen-source conditional calibration.

Local Likelihood using GP on projected significance: pt2

One mass versus the full search

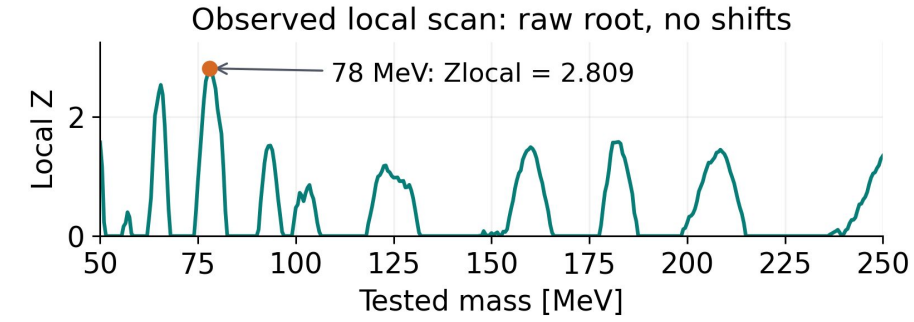
$$Z_{\text{local}} = \max(r_{\text{obs}}, 0)$$

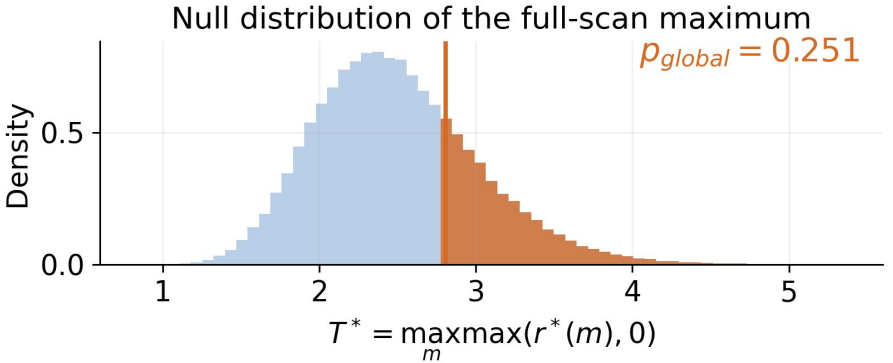
$$T = \max_{m} \max(r(m), 0)$$

$$p_{\text{global}} = P_B(T^* \geq T_{\text{obs}})$$

Local: the excess at one mass, with the raw asymptotic mapping.
Global: the largest excess anywhere in 50–250 MeV.

Poisson check: 57/256 exceedances.
95% interval: 0.173–0.279.





Ananiev & Read, JINST 18 (2023) P05041; HPS v5.8.5.3.
2021 native 10%; ±2.25σ; 50–250 MeV; frozen-source conditional calibration.

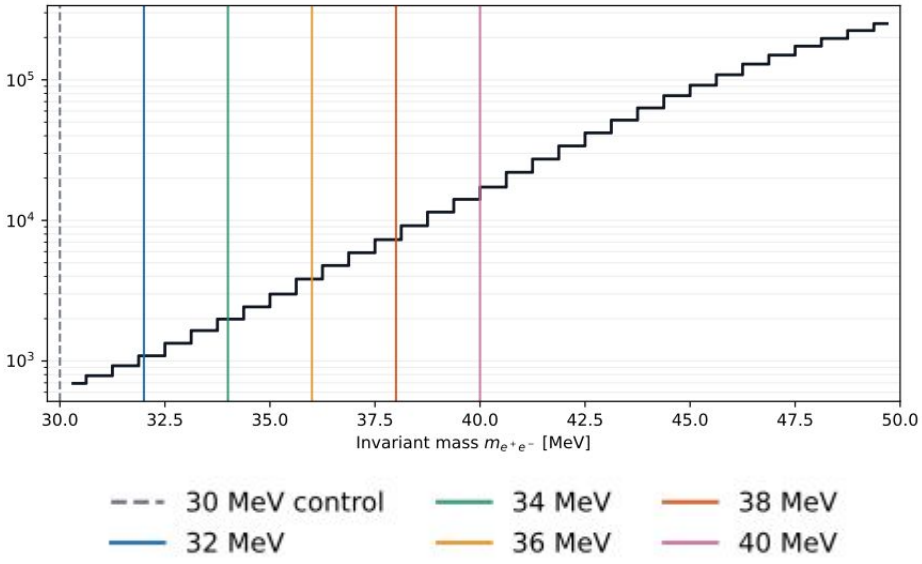
2021 Threshold Fit optimization

Study done after seeing apparent biases with GP fit support starting at 40 MeV in pseudoexperiment tests.

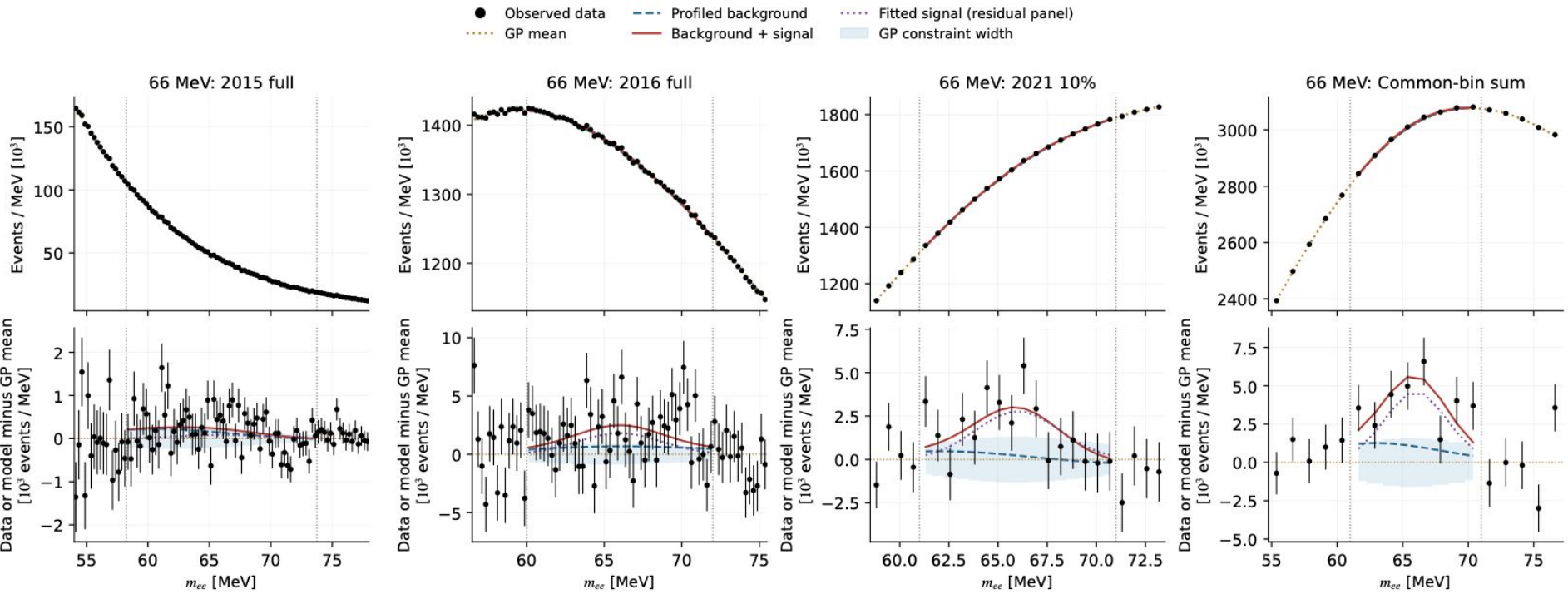
- selected GP fit support range and procedure for selection for fully unblinded 2021 dataset

Rising edge support selected: 36 MeV
Note goes into more detail

10% 2021 Threshold Region



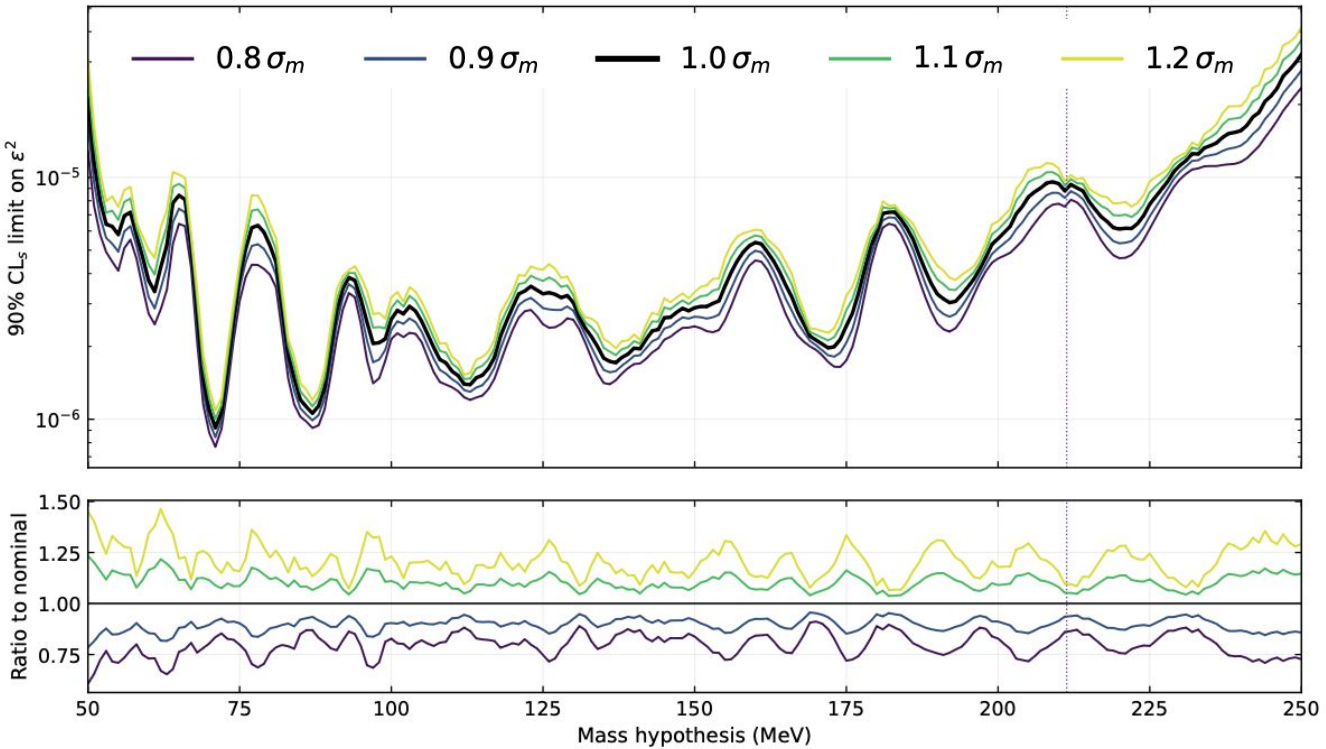
Extraction Display



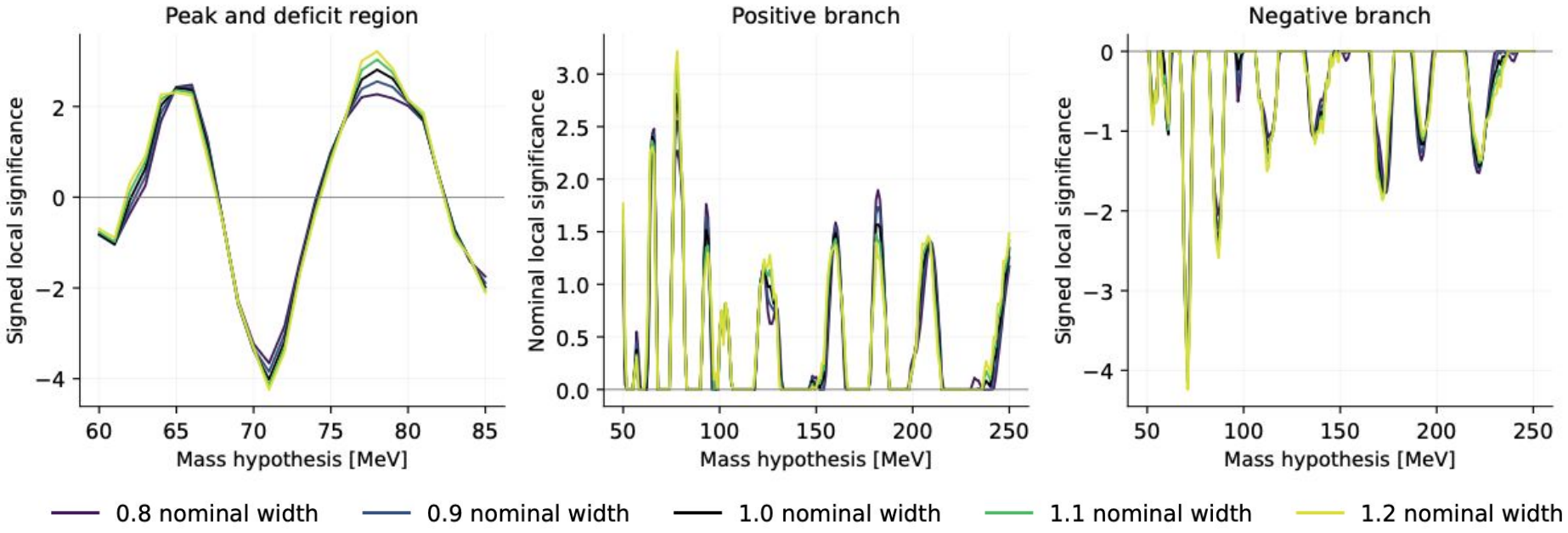
Wiggling the mass resolution pt 1

Impacts

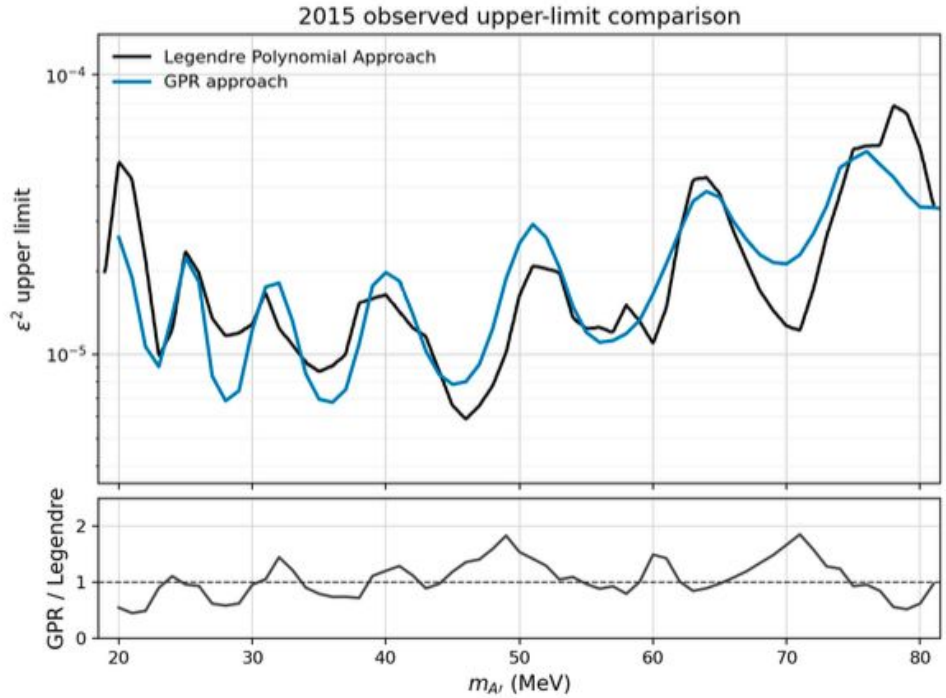
- blind window size
  - (fit uncertainty)
- signal width
- bkg density calculation (marginally)



Wiggling the mass resolution pt 2



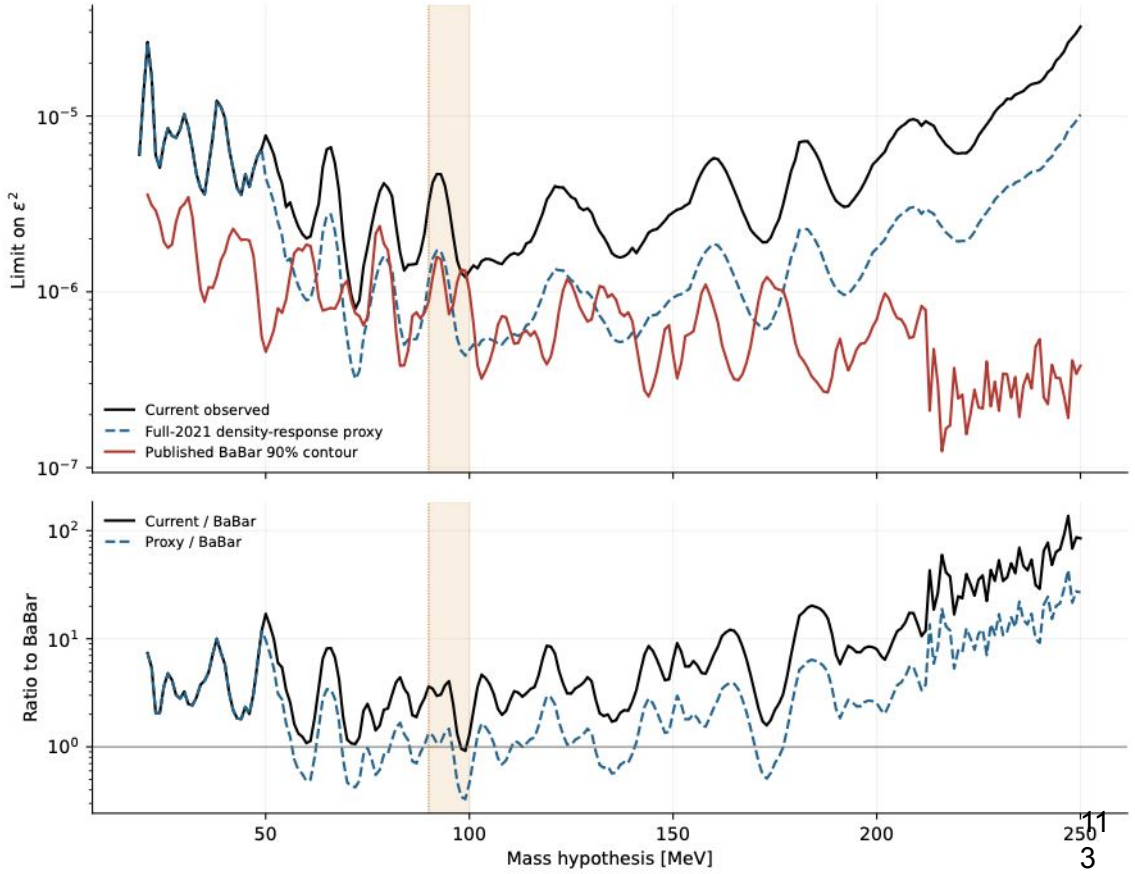
2015 Comparison



Babar Comparison

Currently have small region of new exclusion

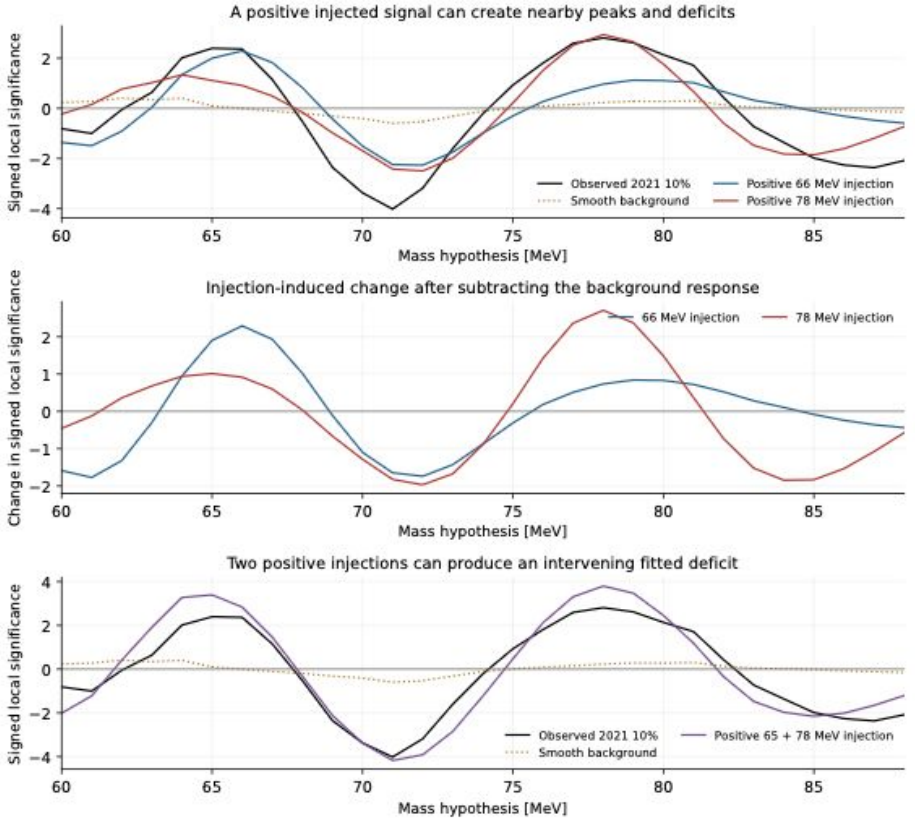
- Exciting!



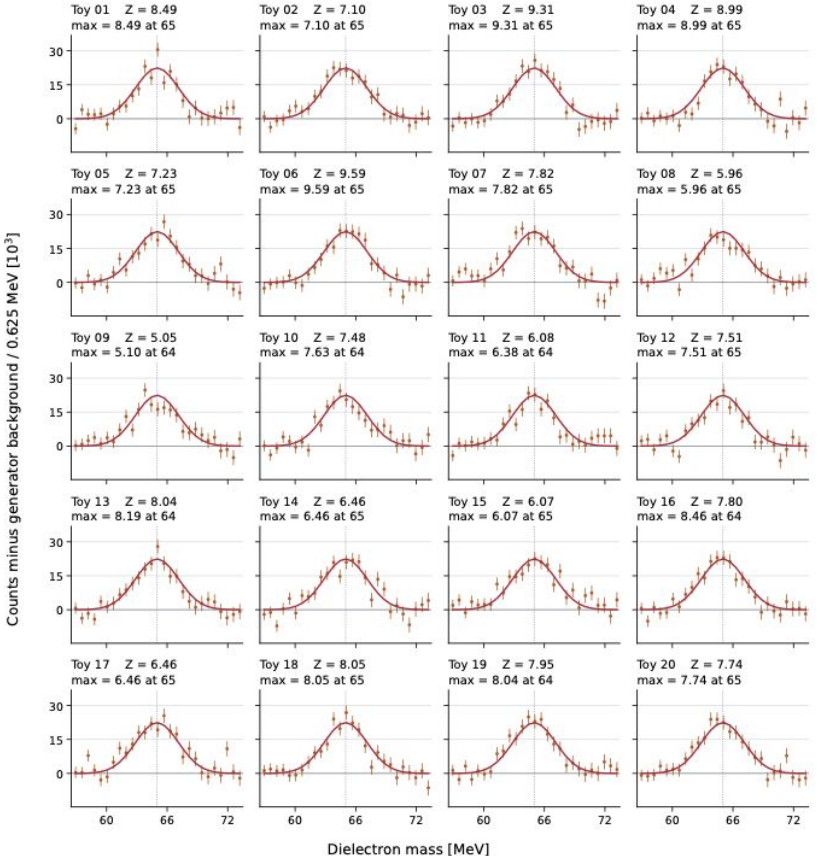
Two Signal Modeling

Two signals can create constructive/destructive effects in observed upper limit

- 2 mild excesses could create deeper deficit than a single nearby signal

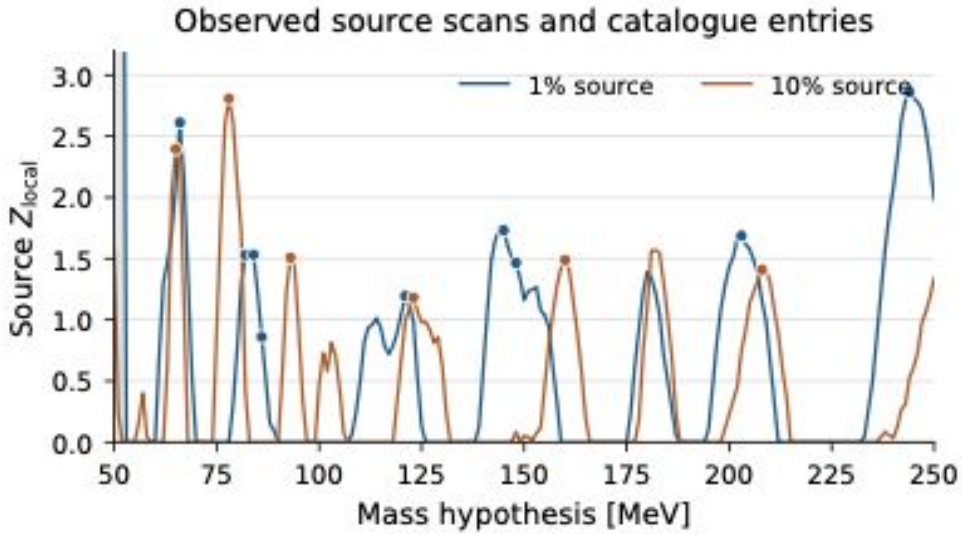


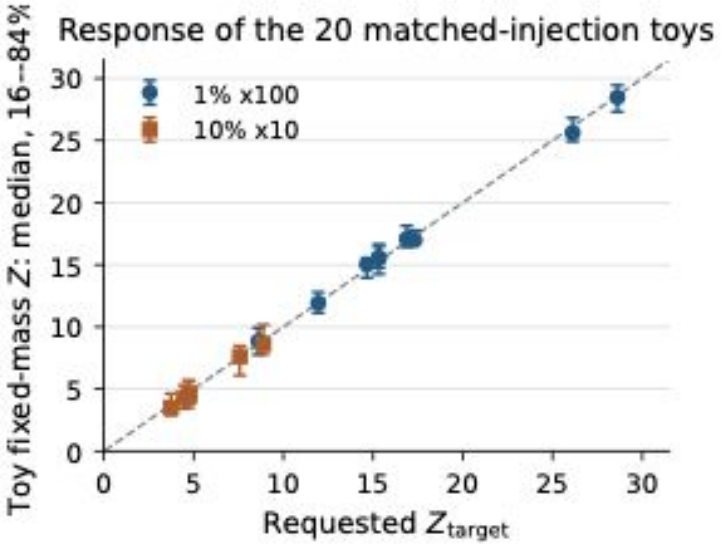
hypothetical fit catalogue (65 MeV toys)



2021 1% vs 10%

Catalogue overview





All "Candidates"

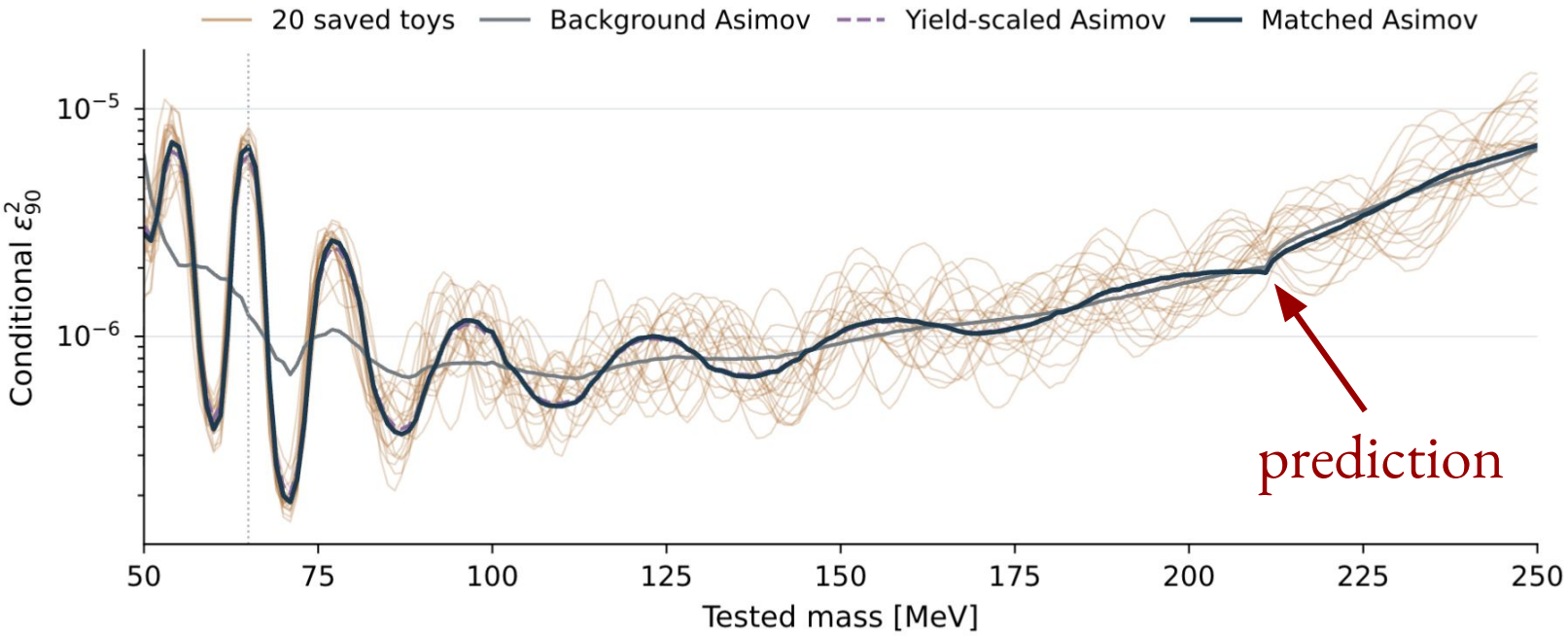
Source	Region [MeV]	m_0 [MeV]	Z_f	Target	Z_A^{yield}	R_a	Toy median [16,84%] fixed mass	Max_{20} fixed mass
1%	65	66	2.61	26.12	23.43	1.12	25.62 [24.93, 26.82]	27.79
10%	65	65	2.40	7.58	6.76	1.12	7.63 [6.10, 8.47]	9.59
1%	75	82	1.53	15.32	13.99	1.10	15.54 [14.25, 16.47]	17.92
10%	75	78	2.81	8.88	8.39	1.06	8.64 [7.80, 10.18]	11.62
1%	92	$86^{\dagger}$	0.86	8.61	7.99	1.08	8.86 [7.77, 9.92]	10.45
10%	92	93	1.51	4.77	4.45	1.07	4.63 [3.92, 5.58]	6.33
1%	120	121	1.19	11.94	11.13	1.07	11.91 [11.13, 12.79]	13.04
10%	120	123	1.18	3.74	3.64	1.03	3.51 [2.89, 4.63]	4.92
1%	160	$148^{\dagger}$	1.47	14.66	13.59	1.08	15.03 [13.96, 15.53]	16.53
10%	160	160	1.49	4.71	4.54	1.04	4.39 [3.50, 5.71]	7.35
1%	210	203	1.69	16.87	15.50	1.09	17.09 [16.43, 18.13]	19.02
10%	210	208	1.41	4.46	4.16	1.07	4.23 [3.41, 5.30]	6.90
1%	+84	84	1.53	15.33	14.05	1.09	15.57 [14.75, 16.65]	17.05
1%	+145	145	1.73	17.32	16.07	1.08	16.98 [16.78, 17.76]	18.77
1%	+244	244	2.86	28.61	25.31	1.13	28.48 [27.31, 29.45]	30.30

Predictions of observed upper limit

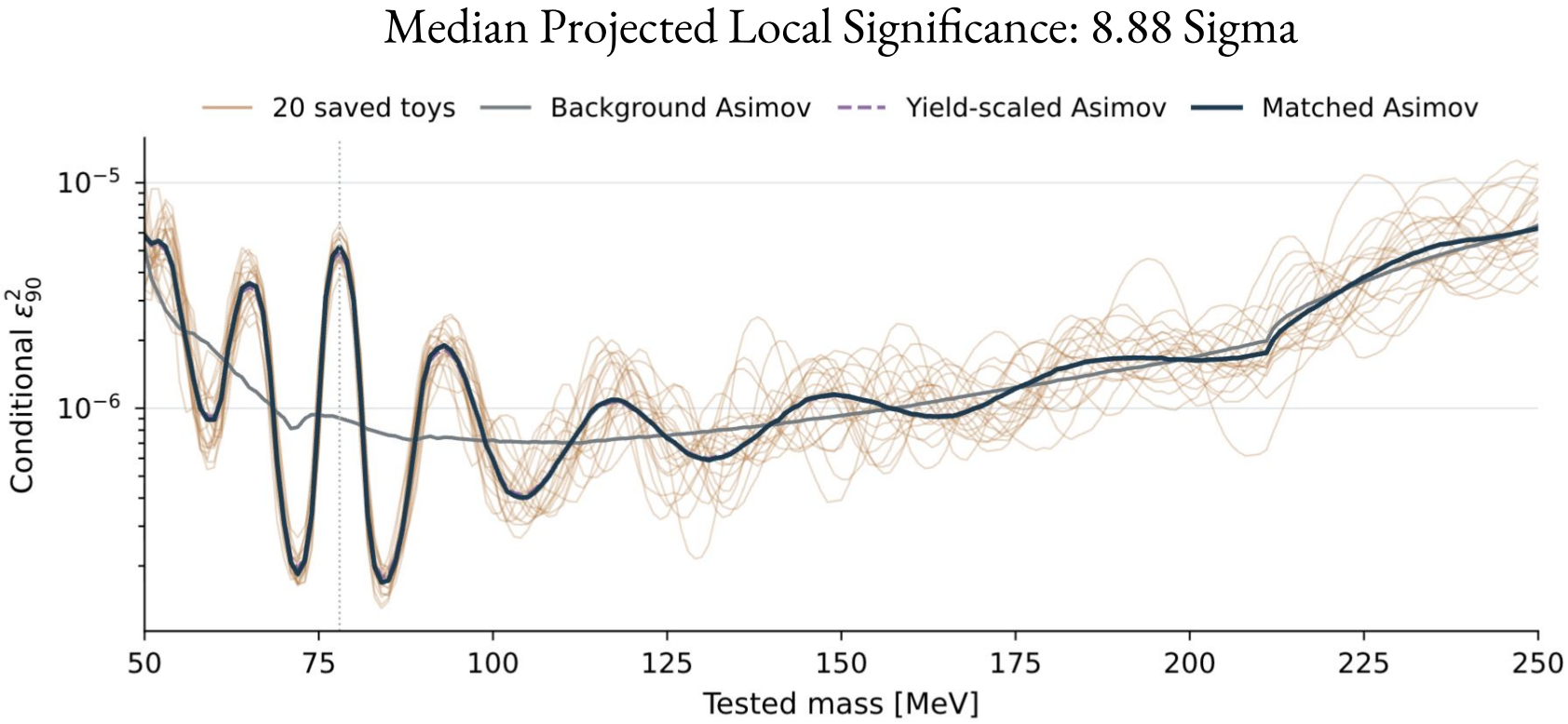
following catalogue shows that we are able to predict shapes of upper limits to first order!

100% Prediction: 65 MeV

Median Projected Local Significance: 7.6 Sigma

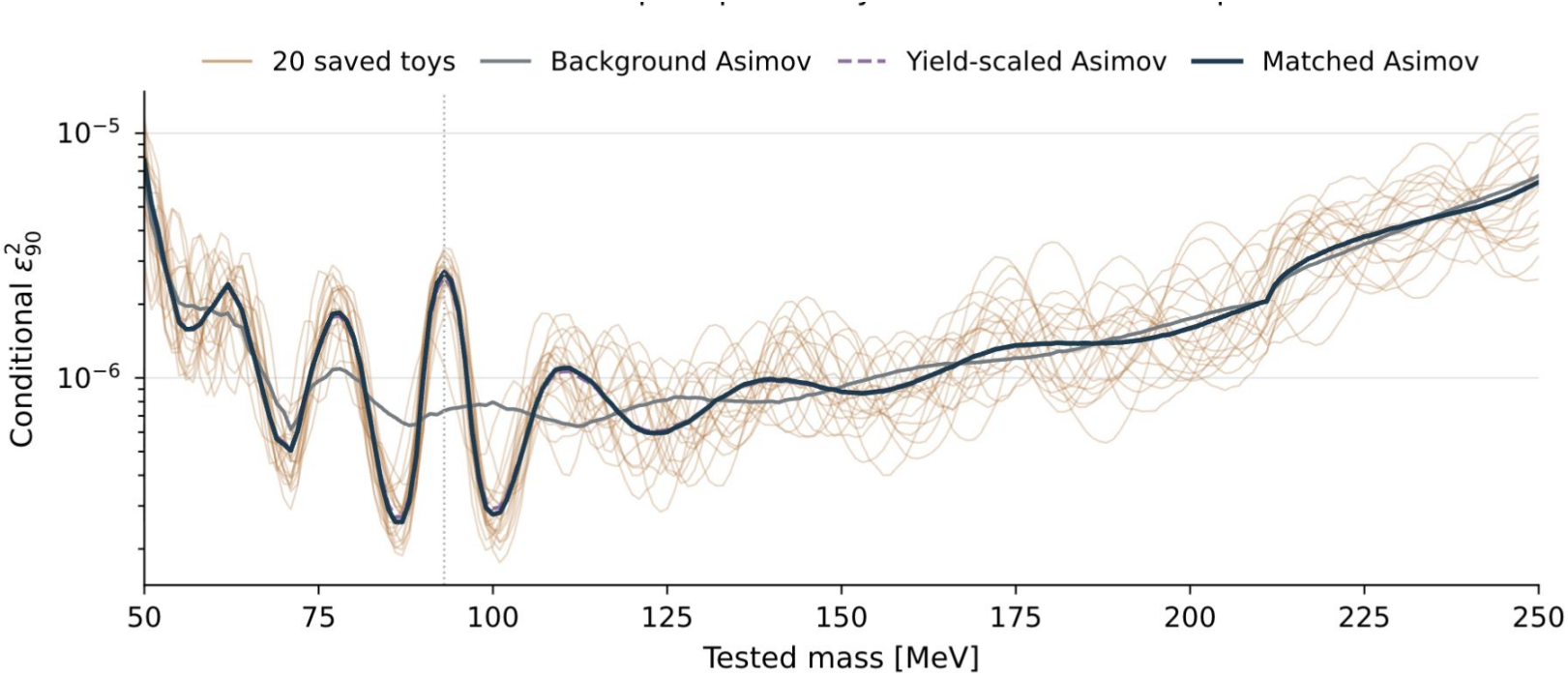


100% Prediction: 78 MeV



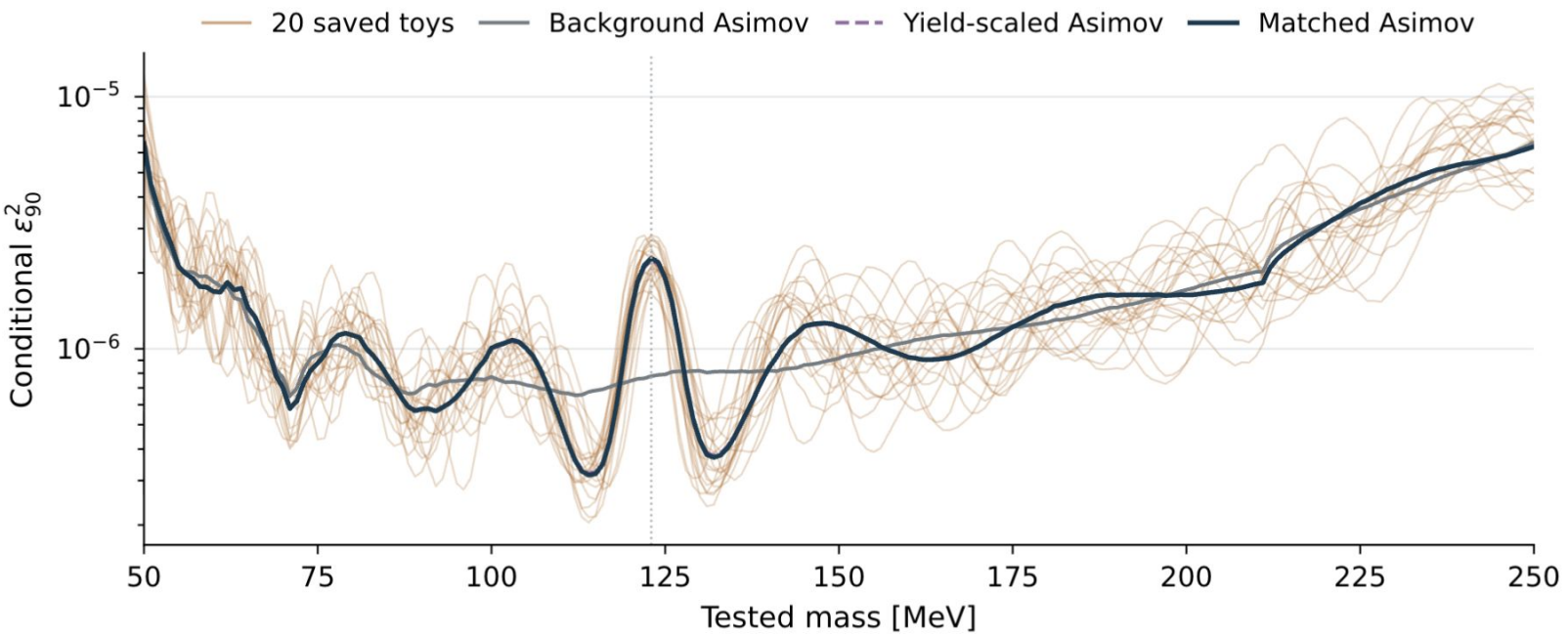
100% Prediction: 93 MeV

Median Projected Local Significance: 4.77 Sigma



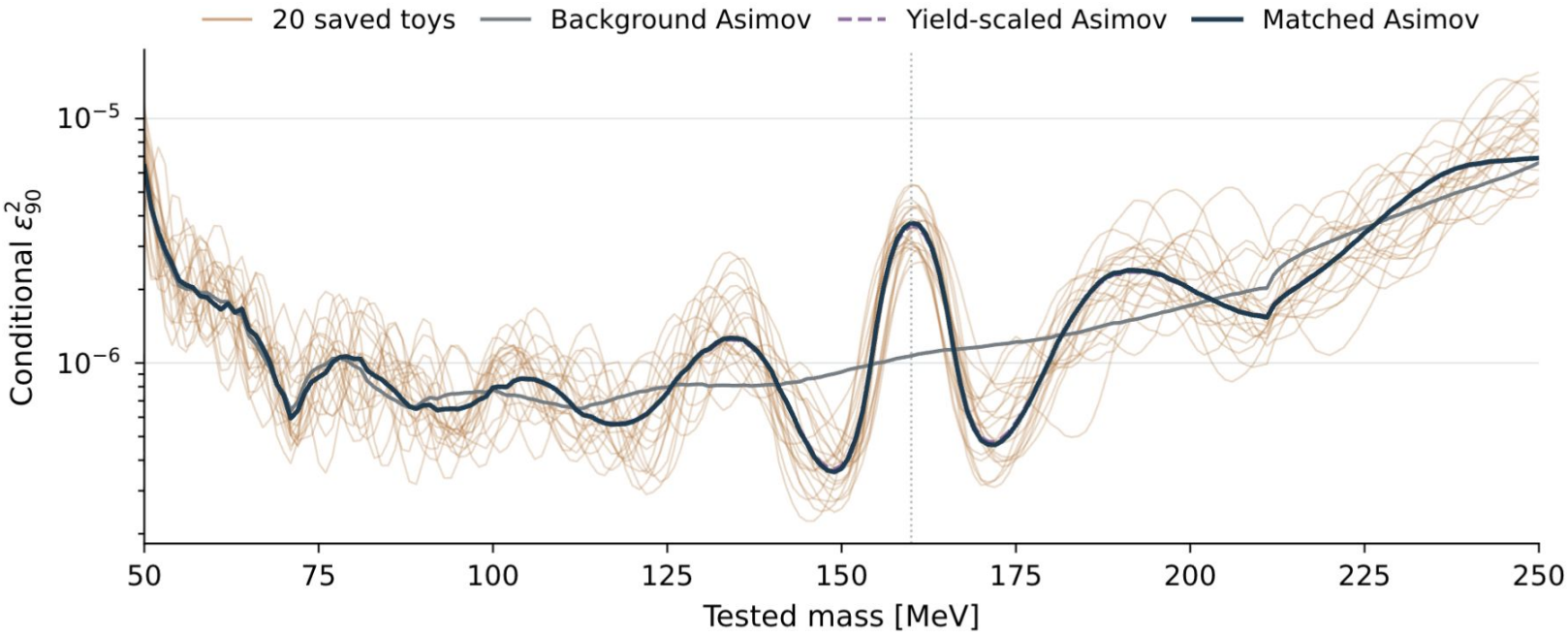
100% Prediction: 120 MeV

Median Projected Local Significance: 3.74 Sigma



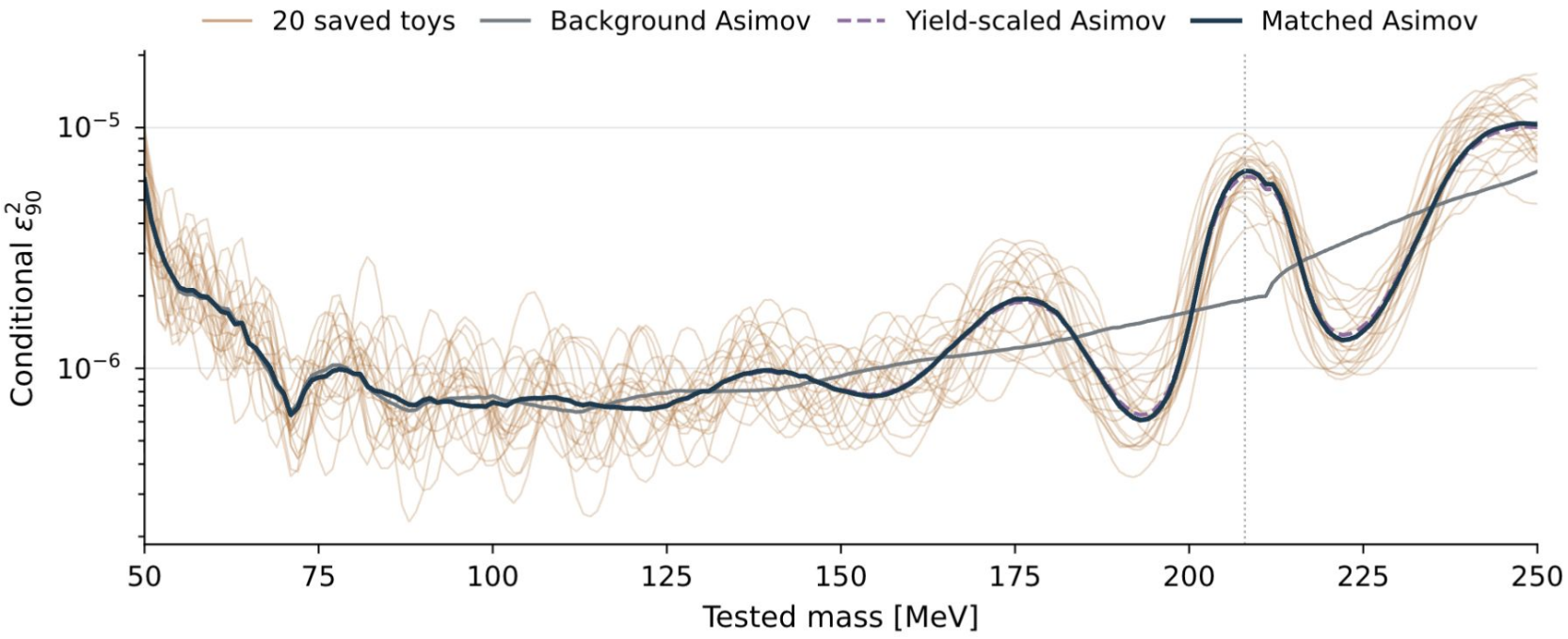
100% Prediction: 160 MeV

Median Projected Local Significance: 4.71 Sigma



100% Prediction: 208 MeV

Median Projected Local Significance: 4.46 Sigma



Interpretations

Study of Agreement in observation between datasets

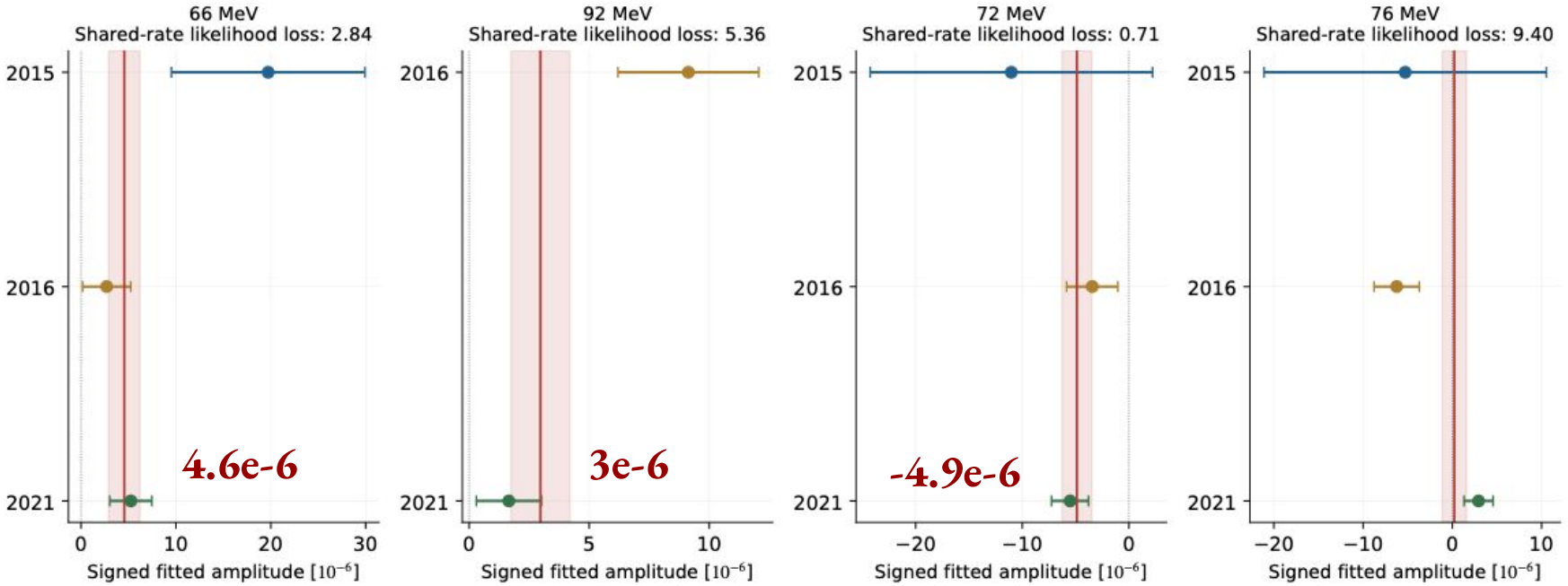
SLAC

Loss metric

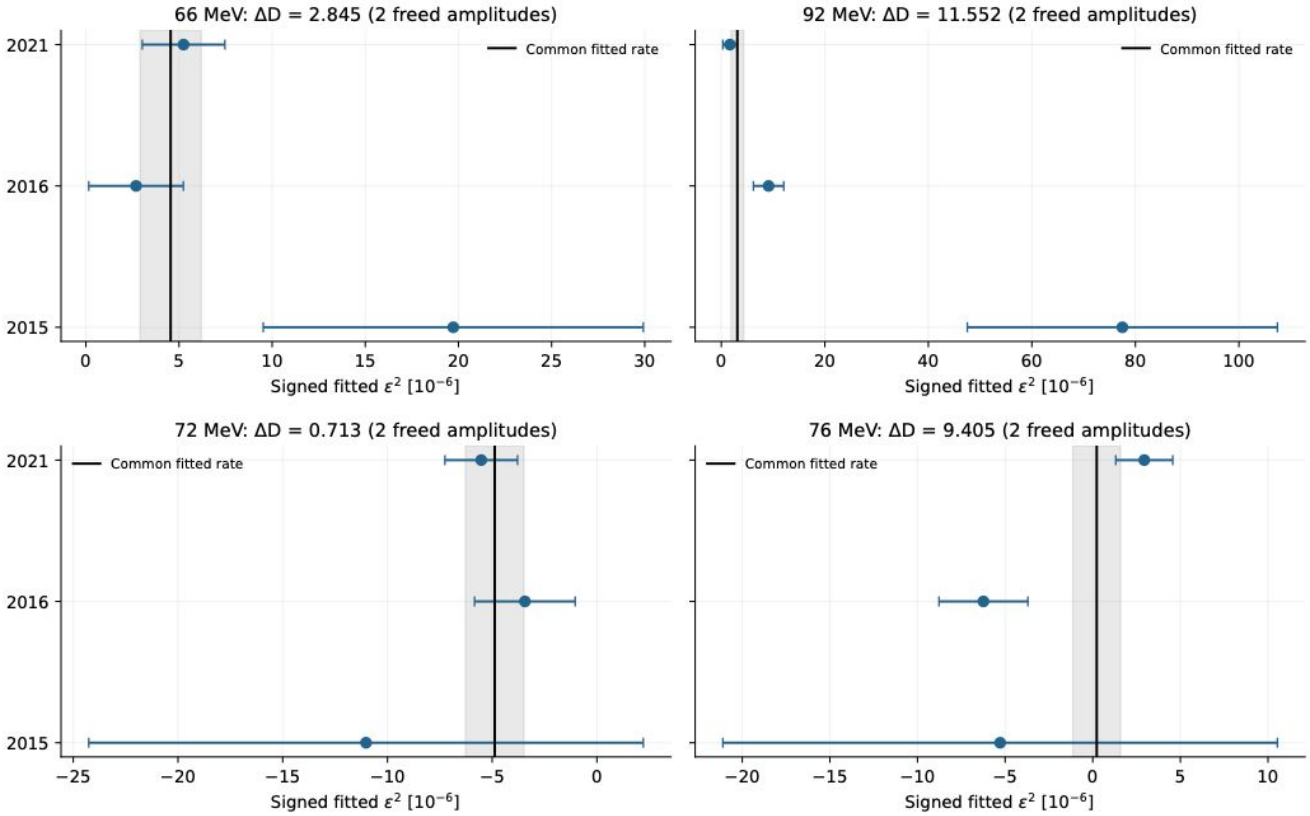
- large loss implies individual datasets describe observation better than shared coupling

$$\Delta D = 2\left[\mathrm{NLL}_{\mathrm{common}} - \sum_{d} \mathrm{NLL}_{d,\mathrm{independent}}\right]$$

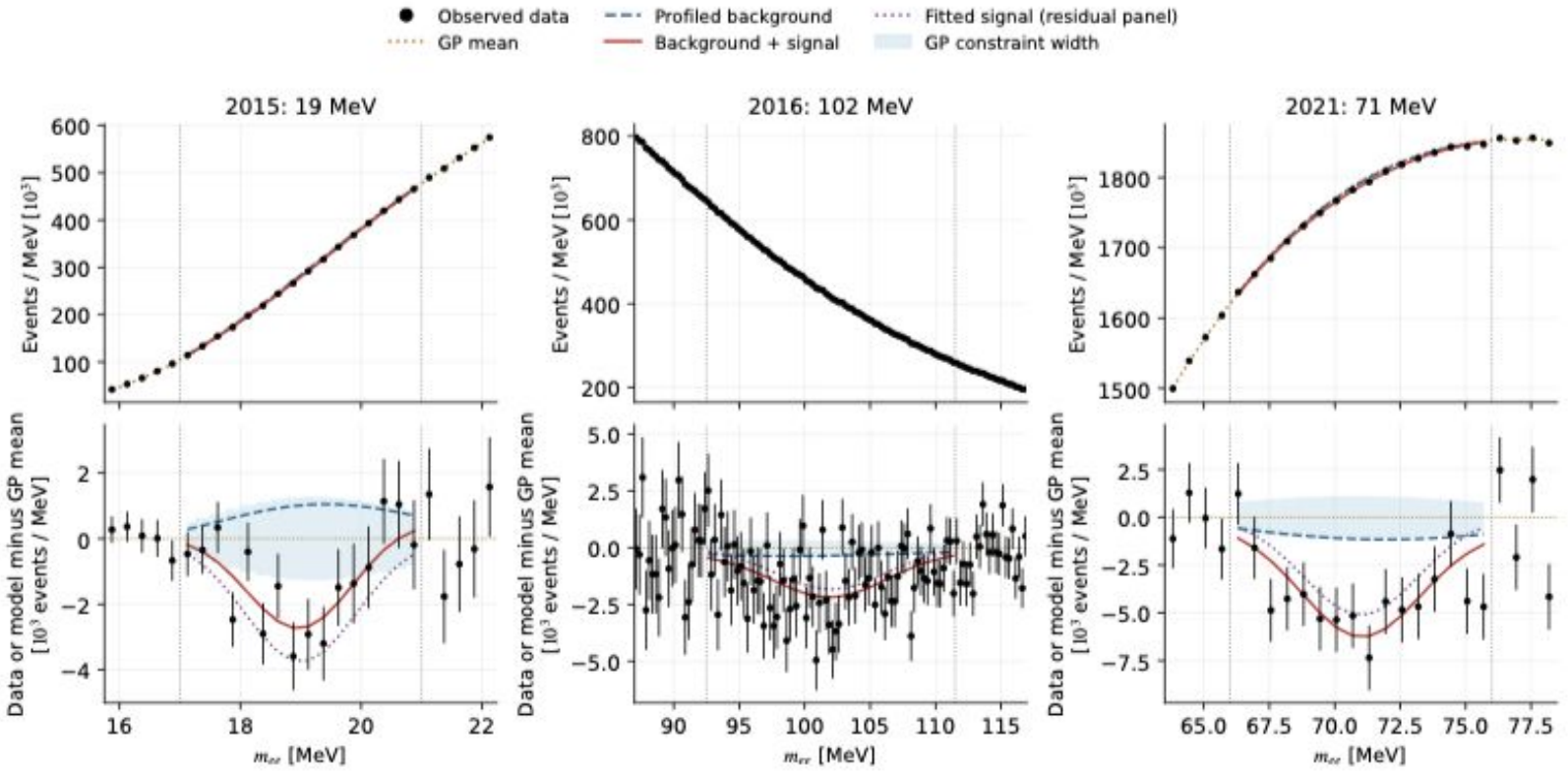
Common rate before 2015 Extension



Common rate after 2015 extension



Leading Deficits



Signal Model for alternative frameworks

Interpolated bkg

Beam Energy

$$\lambda_i = b_i + L_i\theta_i + \mu K_i(m_i, \sigma_i)(E_i/2.30\,\mathrm{GeV})^{\beta}$$

Profiled Background

Radiative Fraction + dataset dependent info

power law

Things to consider post-unblinding

SLAC

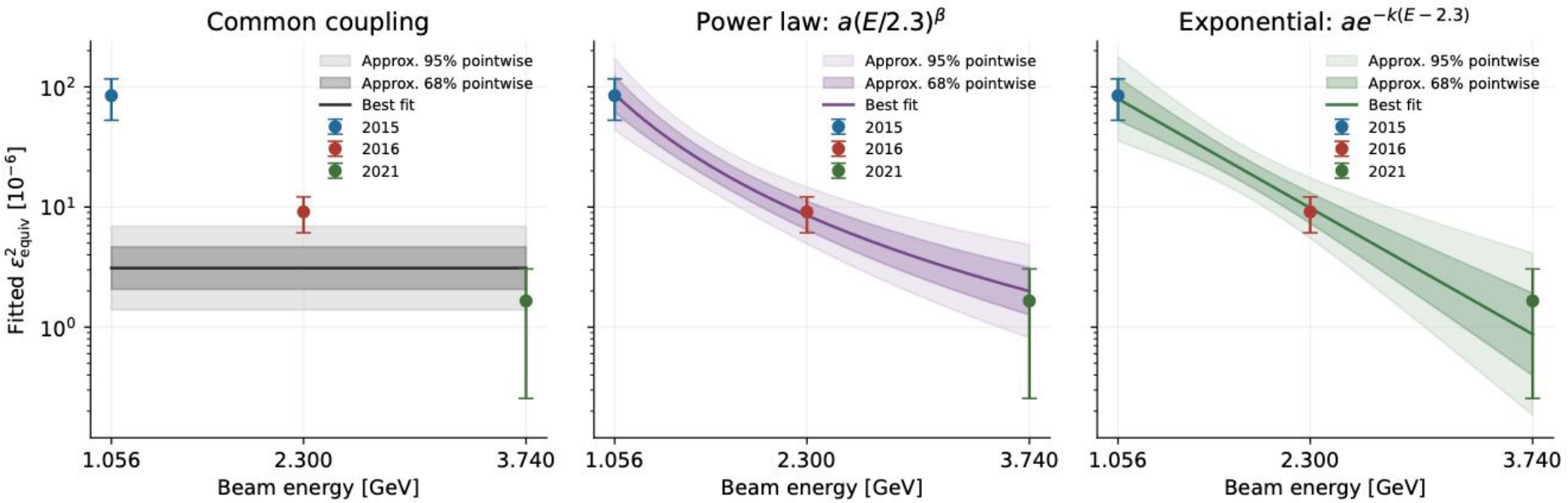
For a fixed common region of interest at 92 MeV

$$Q_G = \sum_i \max(z_i, 0)^2$$

Model or rule	$\sqrt{Q_G}$	Local p_G	Z_G
Common coupling	2.447	7.21×10^{-3}	2.447
Independent positive amplitudes	4.195	1.33×10^{-4}	3.646
Fitted power law	4.182	6.90×10^{-5}	3.812
Fitted exponential	4.148	7.77×10^{-5}	3.782
Equal-weight signed Stouffer	–	3.71×10^{-5}	3.962
Fisher of signed one-sided probabilities	–	6.91×10^{-5}	3.811

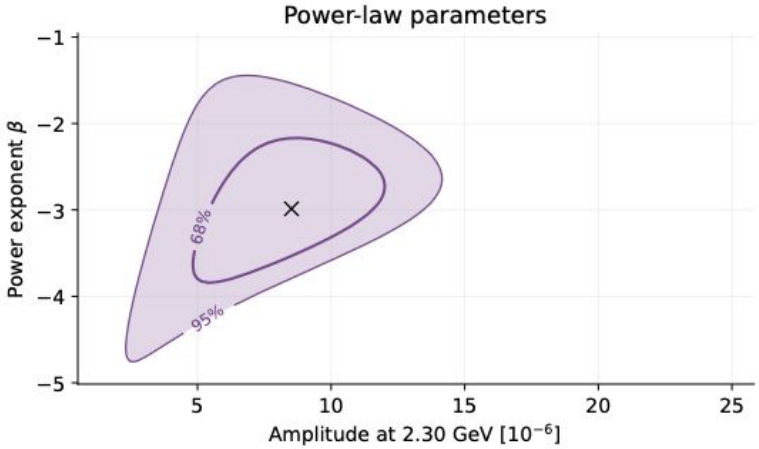
Alternative Combination Methods

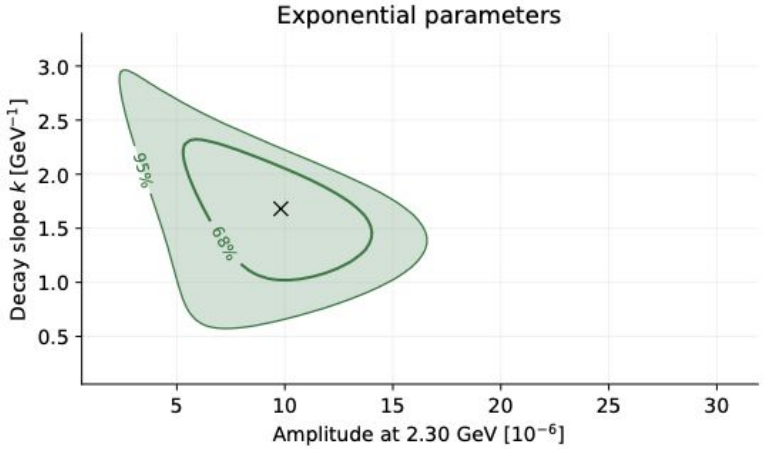
Beam energy and Rate normalization assumptions



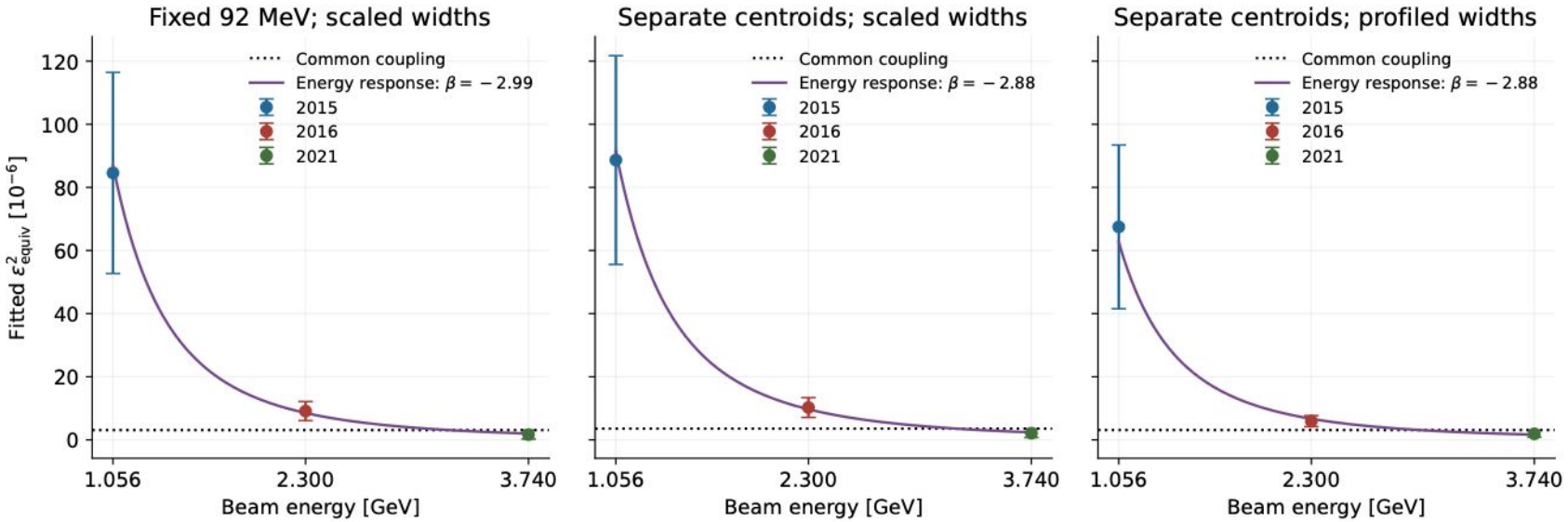
Two Parameterizations

$R(E)$	$\widehat{a}$	conditional slope interval, 68% profile
1	3.102×10^{-6}	none
$(E/2.30)^{\beta}$	8.516×10^{-6}	$\beta = -2.987\ [-3.517, -2.470]$
$\exp[-k(E-2.30)]$	9.817×10^{-6}	$k = 1.681\ [1.257, 2.087]\ \mathrm{GeV}^{-1}$





Profiling centroid and widths

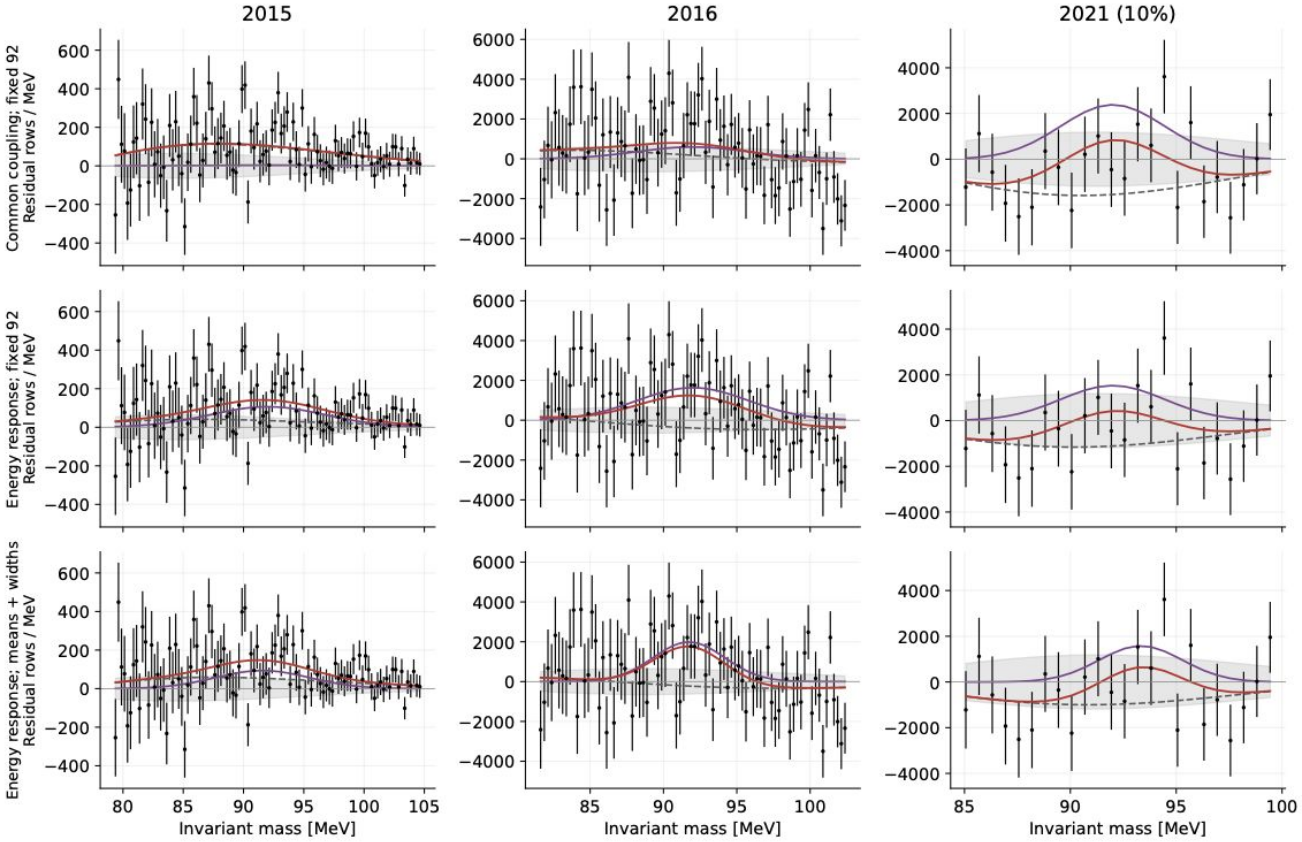


The Different Scenarios

Shared Coupling

Rate Floating within Energy Normalization

Rate Floating within Energy Normalization+ Mass Resolution Profiled



Updates and Studies Since Initial Unblinding (08/04/2026)

I. Geometric Acceptance Study for 2015/2016/2021
II. Pseudoexperiments of various Statistics
    A. 1% fit data scaled to 10% and 100%
    B. 10% fit data fit as 10% and scaled to 100%
III. 2021 GP fit support optimization
    A. rising edge
IV. New observed limit results and all combinations of limits
V. 2016 100% Comparison
VI. Contour revitalization (15/16/21 and 15/16/19/21 Projections)
VII. GPR Echo characterization across all three datasets
    A. individual/combined dataset search correlations
    B. predictions for 100% unblinding
VIII. Scaling Mass Resolution Study
IX. Profiling Mass resolution as a nuisance
X. Beam Energy relationship to extracted regions of interest
XI. “Two-bump” hypothesis