Inelastic Dark Matter Search at Belle 2

T. LAM, S. LONGO, L. PIILONEN

2025 BELLE II SUMMER WORKSHOP

VIRGINIA TECH

Outline

- Next Steps
- Upper Limit Setting
- Signal Extraction
- Background Suppression
- iDM Explanation

Inelastic Dark Matter

- Motivation: During early Universe, we might want to have a production channel for dark matter
- Features such a channel would need:
  - coupling to standard model for thermalization with SM
  - having enough annihilations to be consistent with the observed relic abundance

SOURCE: M. Duerr et al. “Invisible and displaced dark matter signatures at Belle II”. https://arxiv.org/pdf/1911.03176.pdf

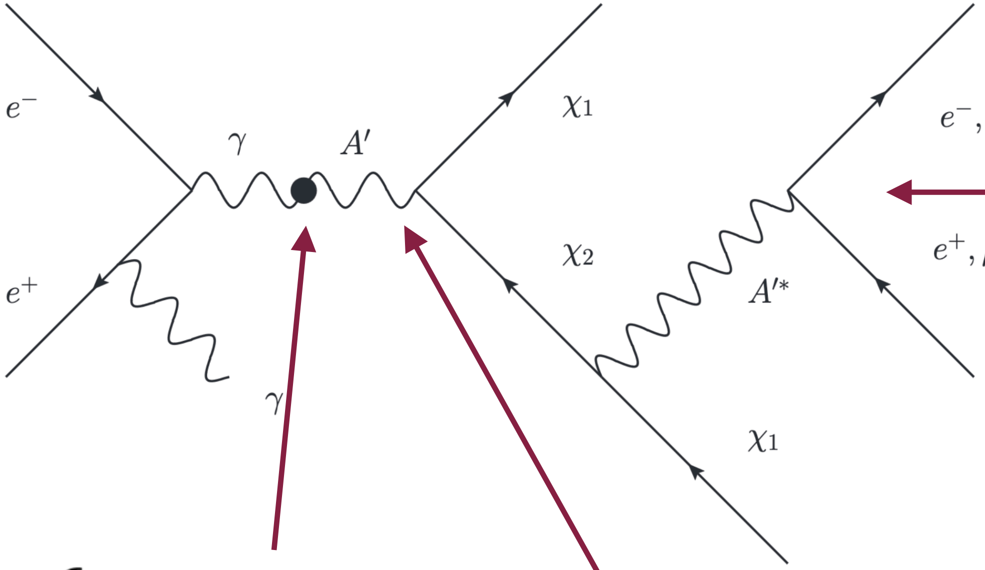
Inelastic Dark Matter

- Model Constituents
  - Dark Fermions (χ_1, χ_2), dark photon (A')
- Parameters:
  - kinetic mixing term (ϵ),
  - dark photon to dark fermions coupling ($\alpha_D = g_X^2/4\pi$)
  - stable dark fermion mass (m_{χ_1}), mass difference (Δm), and dark photon mass ($m_{A'}$ or m_X)

$$\mathcal{L} = \mathcal{L}_{\rm SM} - \frac{1}{4}\hat{X}_{\mu\nu}\hat{X}^{\mu\nu} + \frac{1}{2}m_{\hat{X}}^2\hat{X}_\mu\hat{X}^\mu - \frac{\epsilon}{2c_{\rm W}}\hat{X}_{\mu\nu}\hat{B}^{\mu\nu},$$

$$\mathcal{L}_\psi = i\overline{\chi_1}\not{\partial}\chi_1 + i\overline{\chi_2}\not{\partial}\chi_2 + \frac{i}{2}g_X\hat{X}_\mu\overline{\chi_2}\gamma^\mu\chi_1 - \frac{i}{2}g_X\hat{X}_\mu\overline{\chi_1}\gamma^\mu\chi_2 - \frac{1}{2}m_{\chi_1}\overline{\chi_1}\chi_1 - \frac{1}{2}m_{\chi_2}\overline{\chi_2}\chi_2.$$

SOURCE: M. Duerr et al. "Invisible and displaced dark matter signatures at Belle II". https://arxiv.org/pdf/1911.03176.pdf

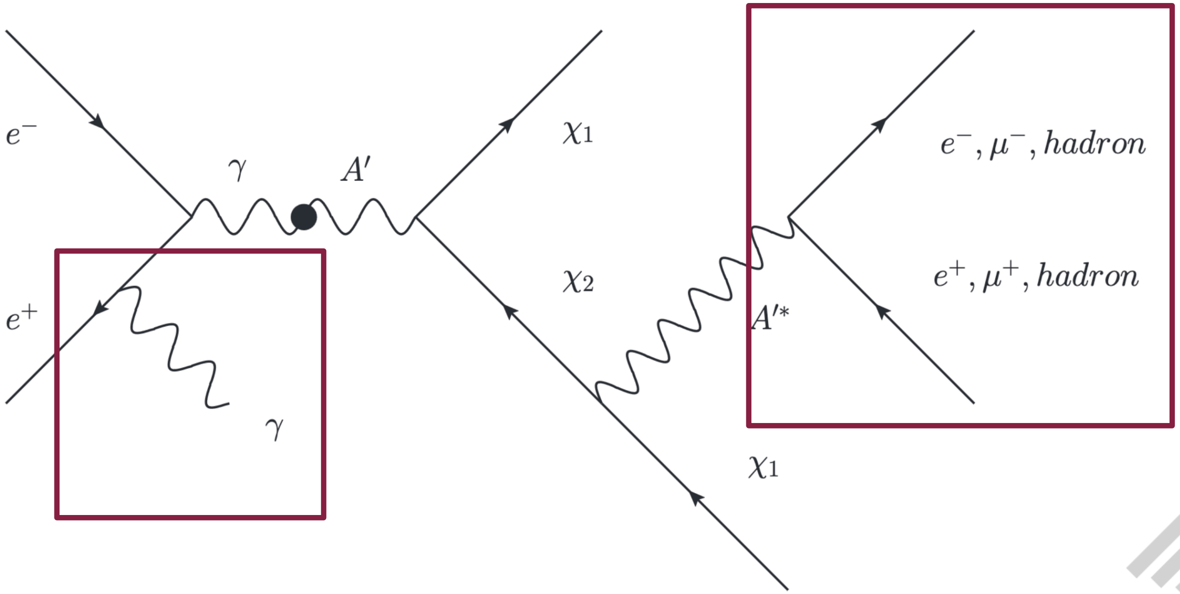


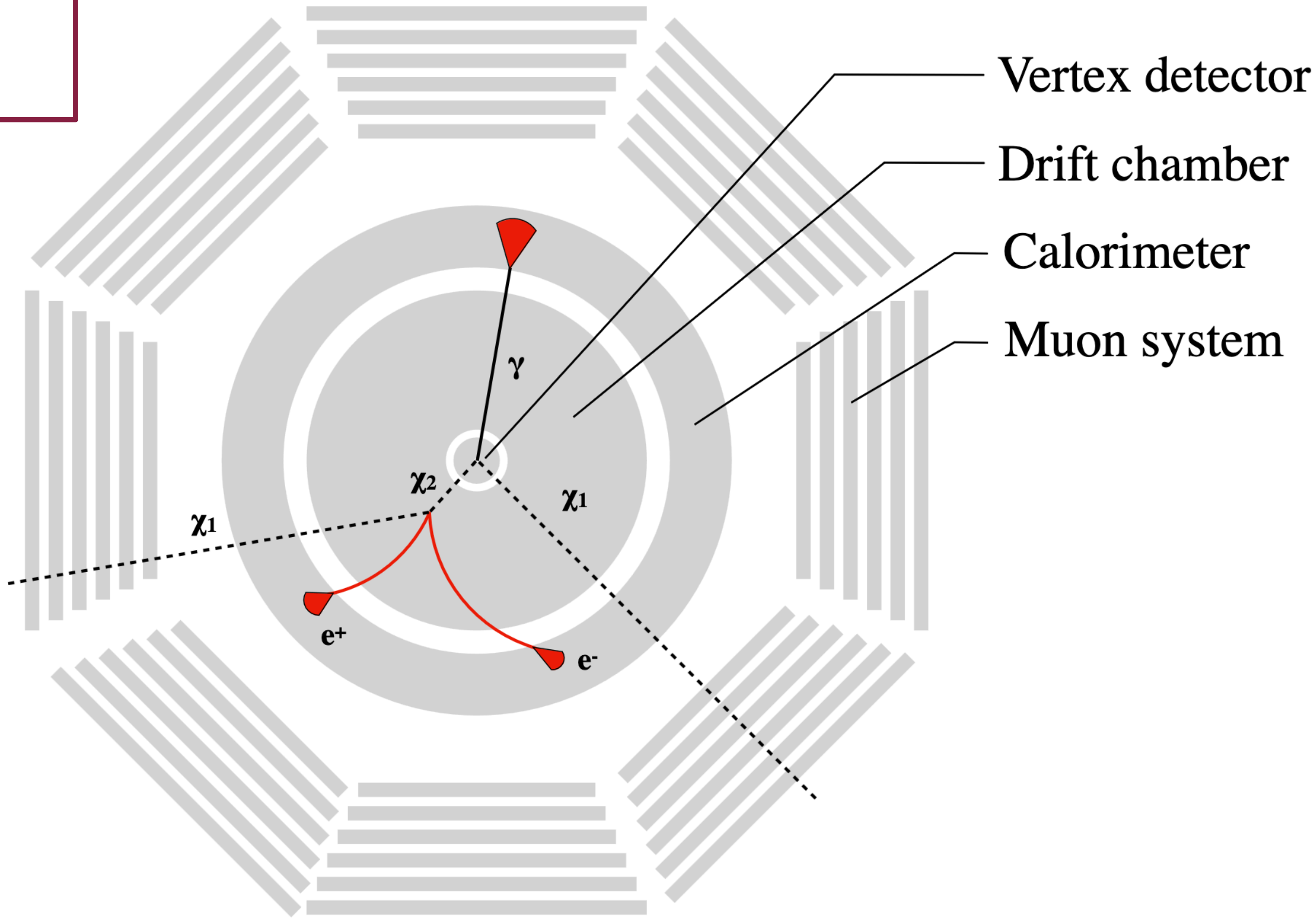
$$\frac{\epsilon}{2c_W}\hat{X}_{\mu\nu}\hat{B}^{\mu\nu},$$

$$\frac{\epsilon}{2c_W}\hat{X}_{\mu\nu}\hat{B}^{\mu\nu},$$

$$+\frac{i}{2}g_X\hat{X}_\mu\overline{\chi_2}\gamma^\mu\chi_1 .$$

SOURCE: M. Duerr et al. "Invisible and displaced dark matter signatures at Belle II". https://arxiv.org/pdf/1911.03176.pdf



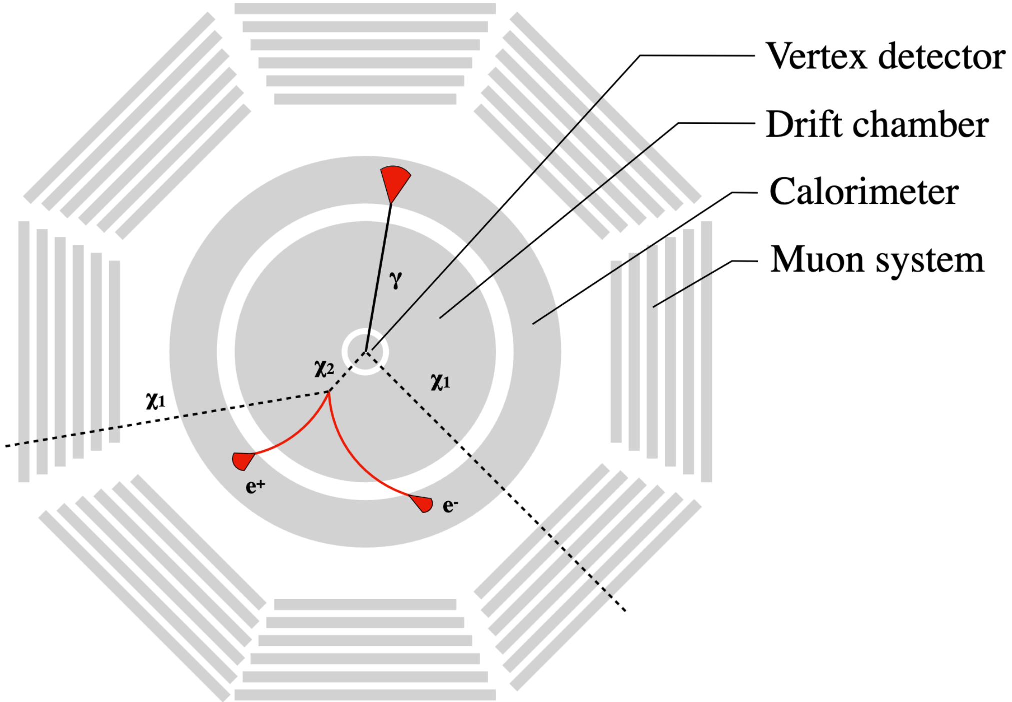


SOURCE: M. Duerr et al. "Invisible and displaced dark matter signatures at Belle II". https://arxiv.org/pdf/1911.03176.pdf

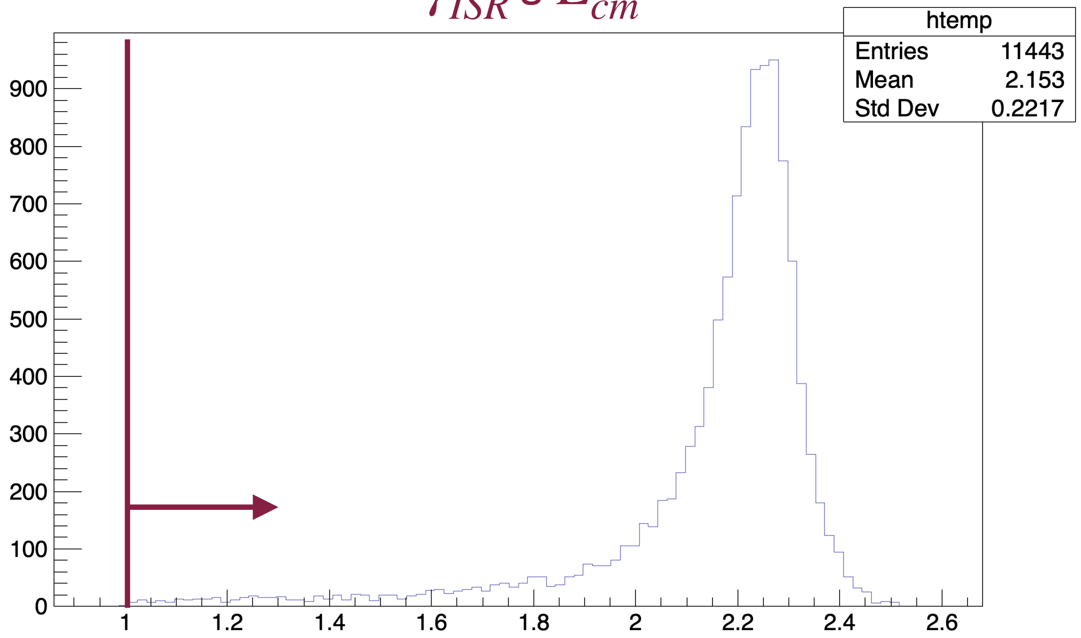
Signal: ISR Photon

ISR = initial state radiation

- Since this is our initial trigger, need to separate from background: an energetic photon (>1GeV), comes from the barrel, etc.



γ_{ISR}'s E_{cm}

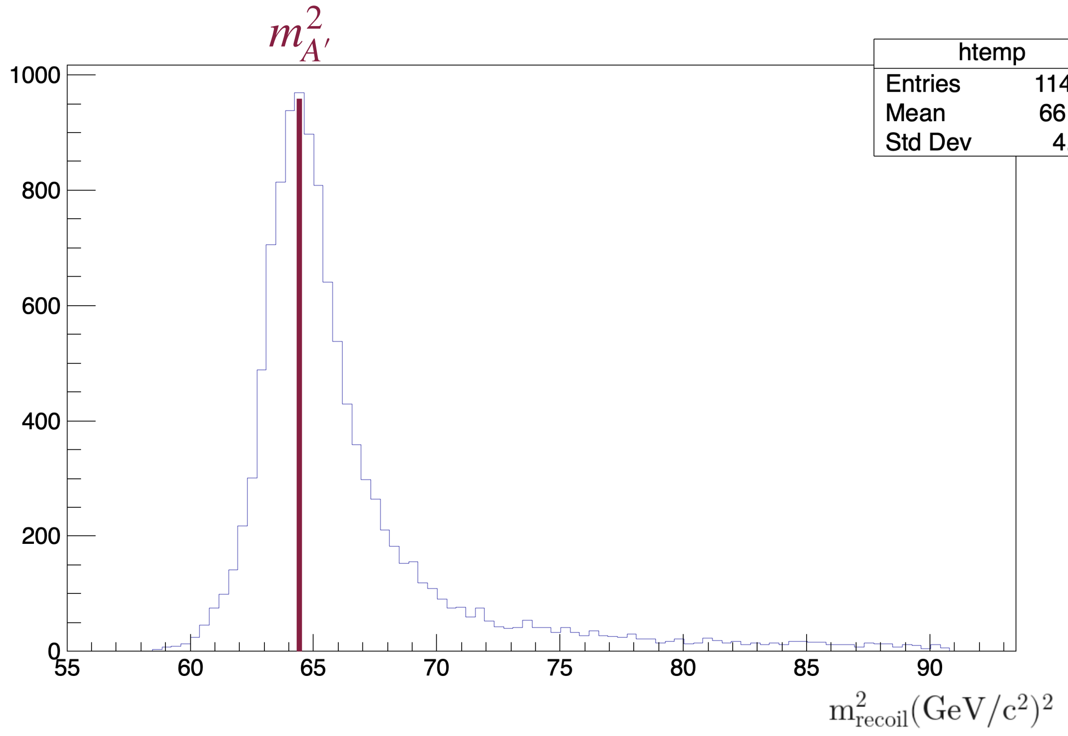


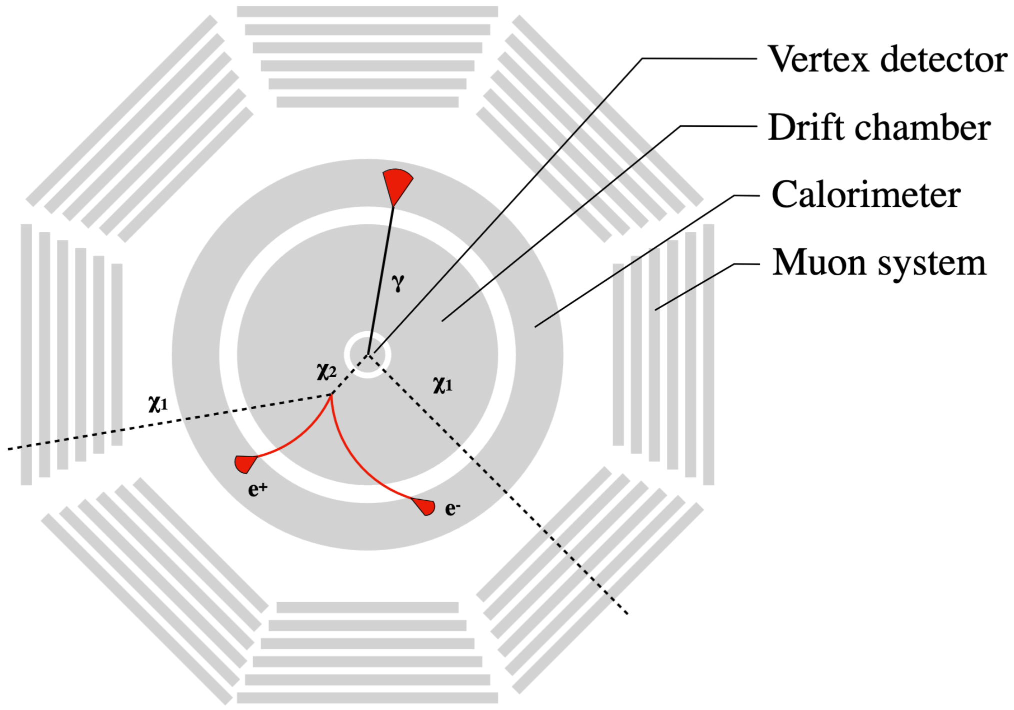
$m_{A'} = 8.0\text{GeV}, \Delta m = 1.28\text{GeV},$
$m_{\chi_1} = 3.2\text{GeV}, \epsilon = 10^{-4}, \alpha_D = 0.1$

SOURCE: M. Duerr et al. "Invisible and displaced dark matter signatures at Belle II". https://arxiv.org/pdf/1911.03176.pdf

Signal: A'

- Since we know the momentum and direction of the ISR photon + initial conditions, then $p_{ini} - p_{\gamma} = p_{A'}$



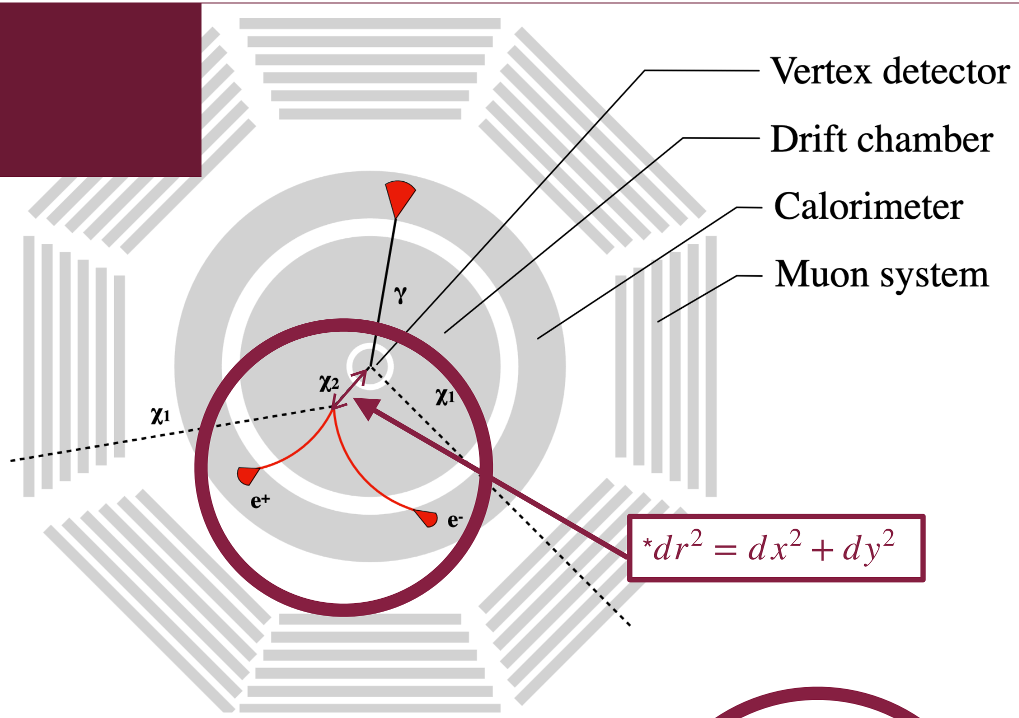


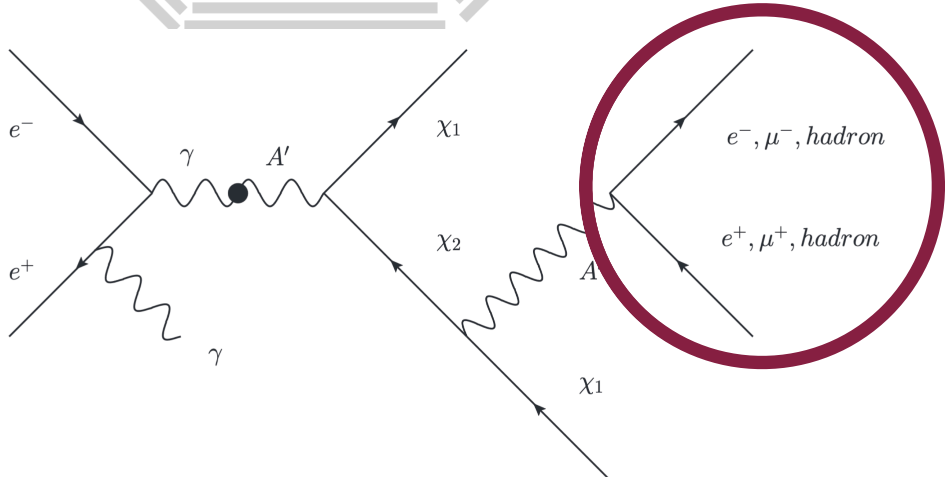
$m_{A'} = 8.0\text{GeV}, \Delta m = 1.28\text{GeV},$
$m_{\chi_1} = 3.2\text{GeV}, \epsilon = 10^{-4}, \alpha_D = 0.1$

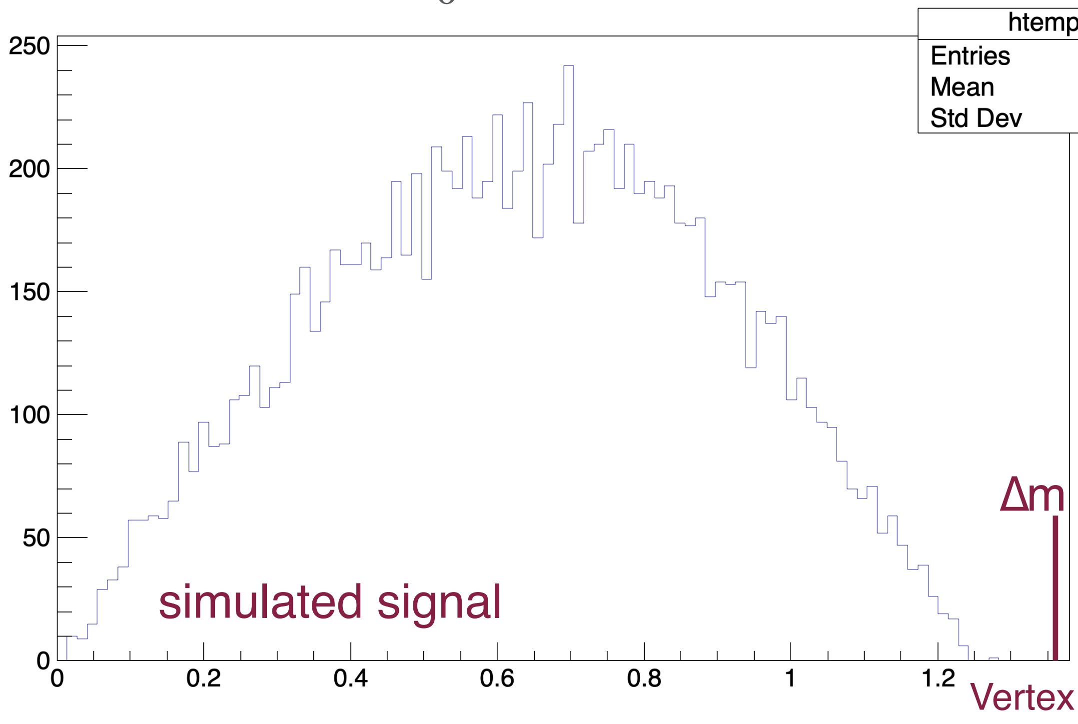
SOURCE: M. Duerr et al. "Invisible and displaced dark matter signatures at Belle II". https://arxiv.org/pdf/1911.03176.pdf

Displaced Vertex

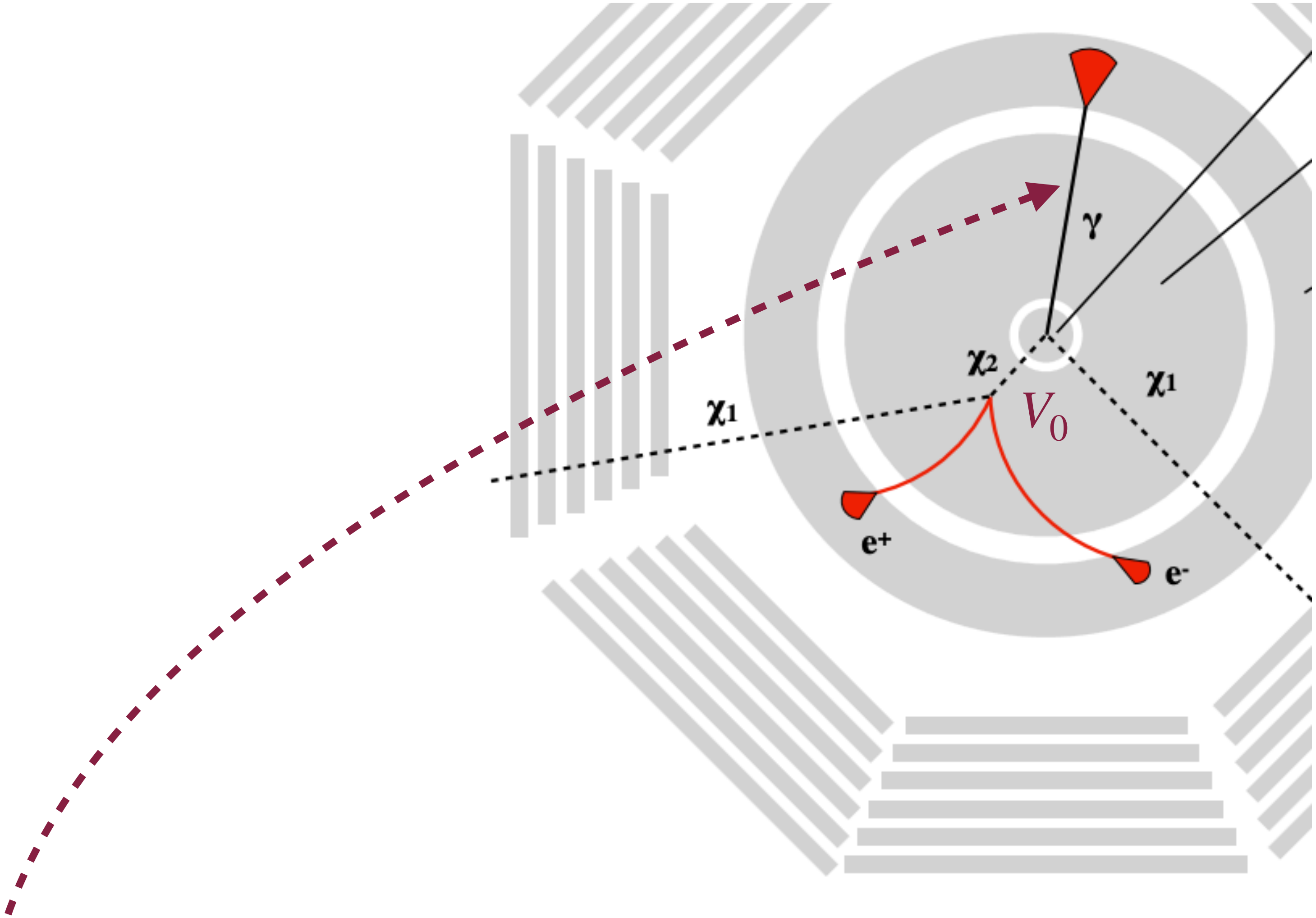
- Two oppositely charged tracks emerge from a displaced vertex, comprising of l^+l^-. Powerful in rejecting backgrounds by using m_{V_0}, dr



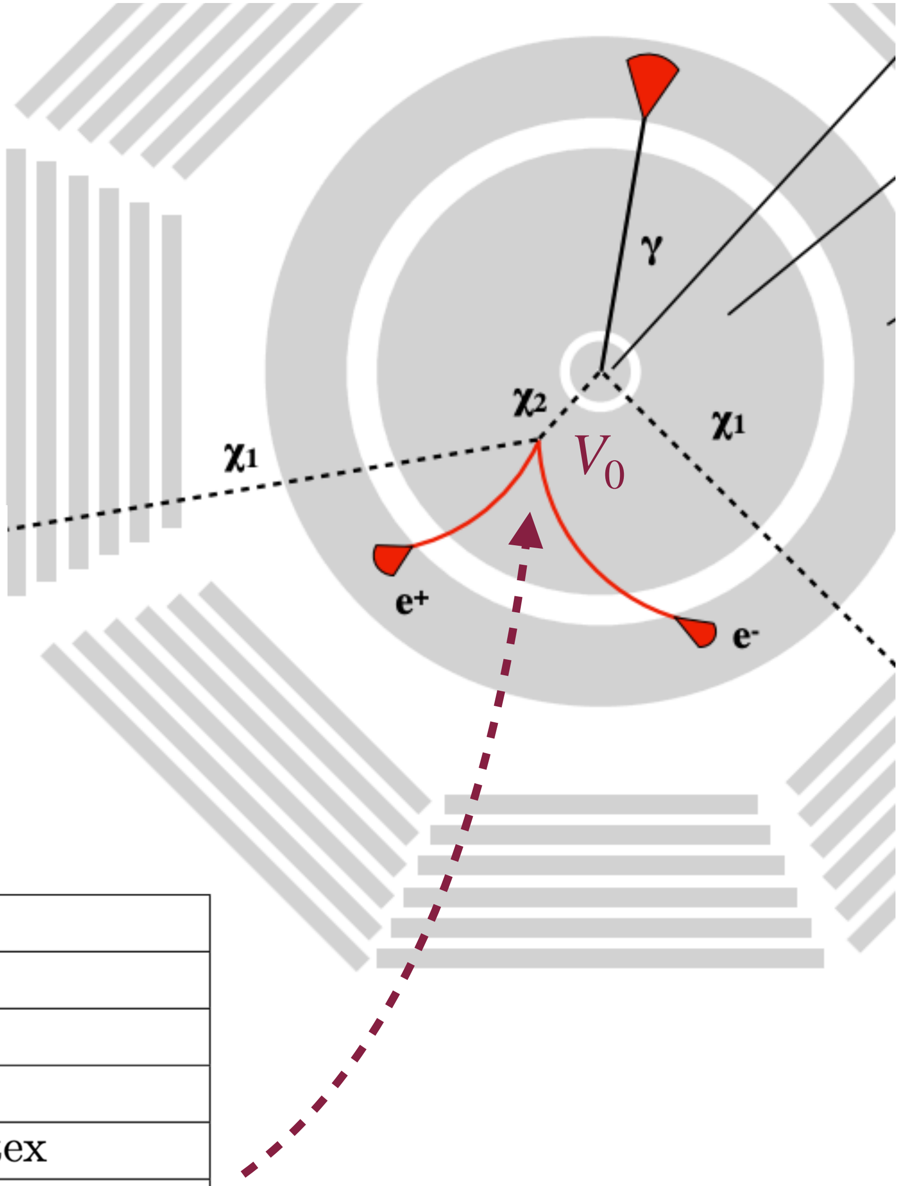




$m_{A'} = 8.0\text{GeV}$, $\mathbf{\Delta m_\chi = 1.28}$ **GeV**,
$m_{\chi_1} = 3.2\text{GeV}$, $\epsilon = 10^{-4}$, $\alpha_D = 0.1$

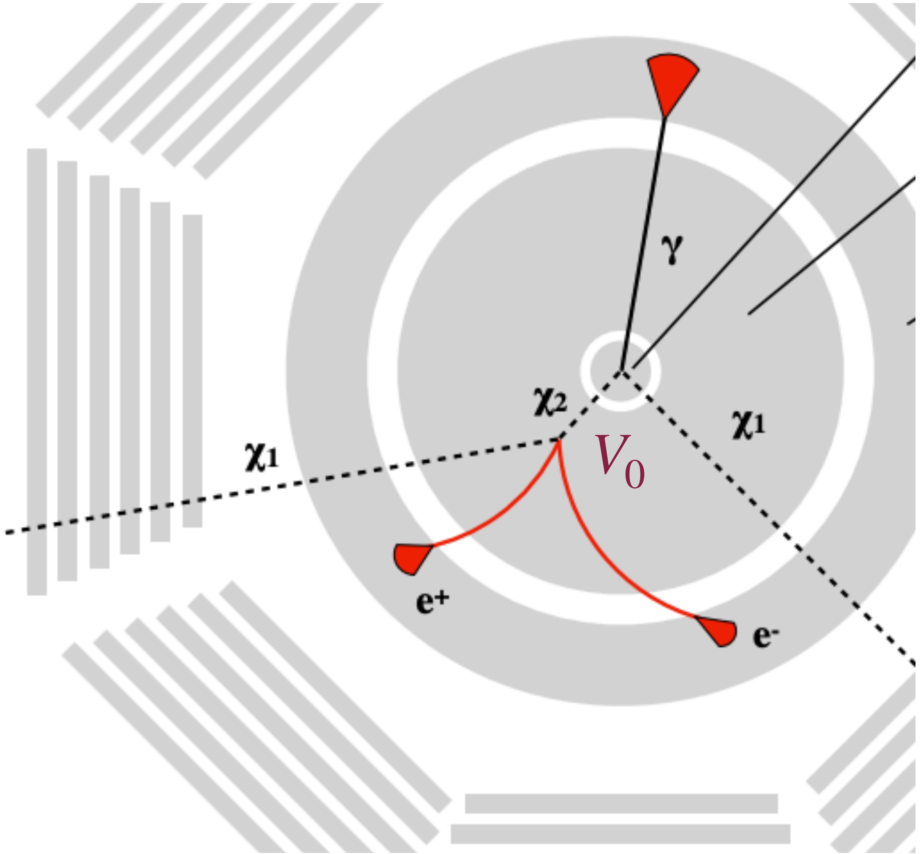


Cut on ISR Photon	Variable Description
$E_{cms} > 1$ GeV	ISR CM frame energy
clusterReg = 2	Selecting ECL clusters in the barrel
ZernikeMVA > 0.7	MVA output trained on Zernike Moments



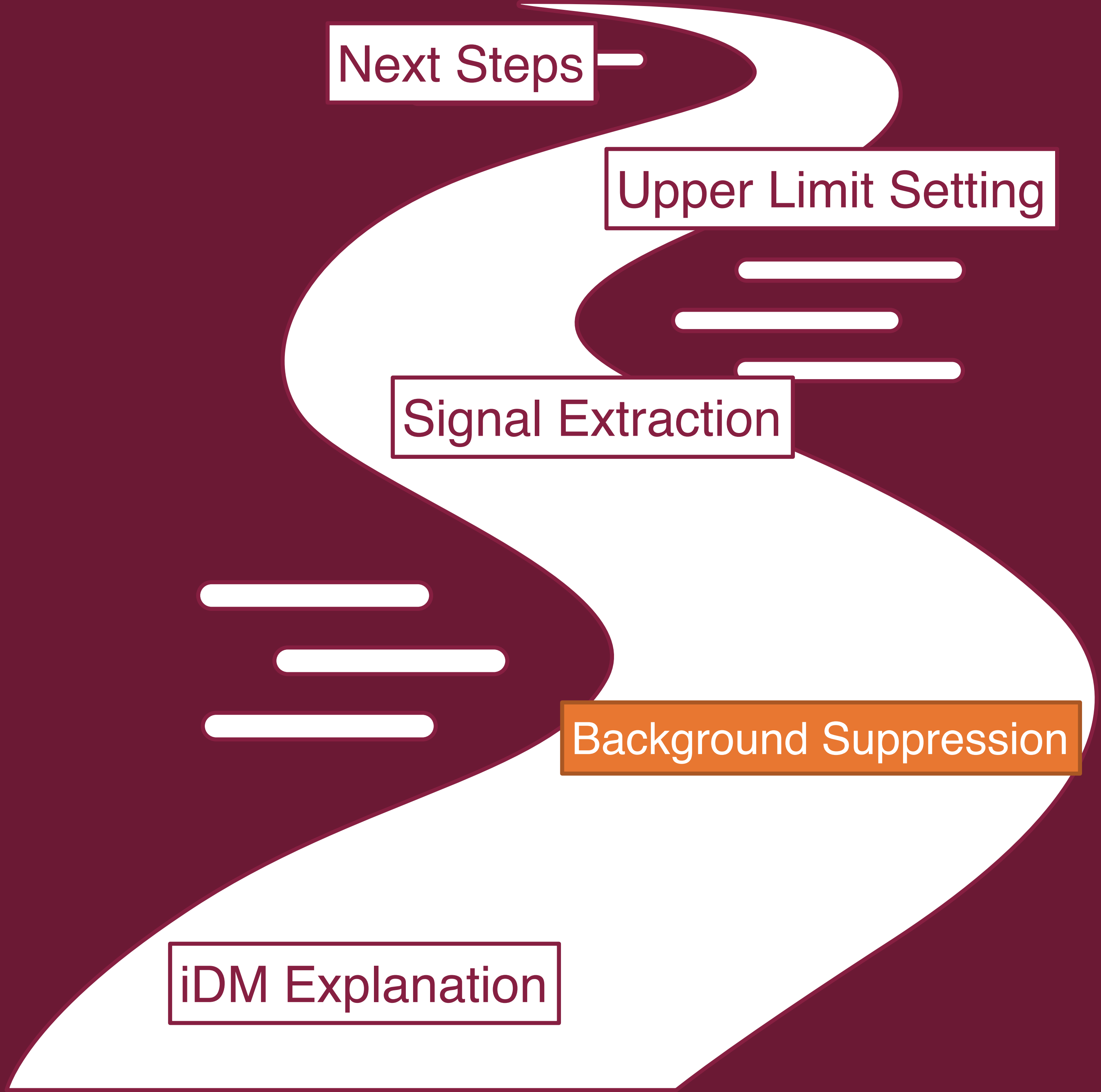
Cut on V_0 daughters	Variable Description
$dr > 0.1$ cm	distance ($\|d_0\|$ for tracks) w.r.t IP
nCDCHits > 4	Number of CDC hits associated to the tracks
$E_{ECL} < 2$ GeV or NaN	ECL Cluster Energy

Cut on V_0	Variable Description
$p > 0.1$ GeV/c	V_0 Momentum
$p_{cm} < 2.0$ GeV/c	V_0's momentum in CM frame
dcos $\theta \in (-0.6, 0.95)$	Polar angle of vertex w.r.t. IP
$dr > 0.5$	Transverse distance with respect to IP for a vertex
$d_0 < 0.5$ cm	d_0, distance from IP in $r - \phi$ plane, of each daughter
$z_0 < 1$ cm	d_0, or z distance from IP, of each daughter
$\vec{p} \cdot d\vec{r}/p\, dr < 0.999$	Displaced vertex's Pointing Vector
$m_{\gamma \to e^+e^-} > 0.01$ GeV/c^2	Invariant vertex mass, assuming it comes from photo-conversions

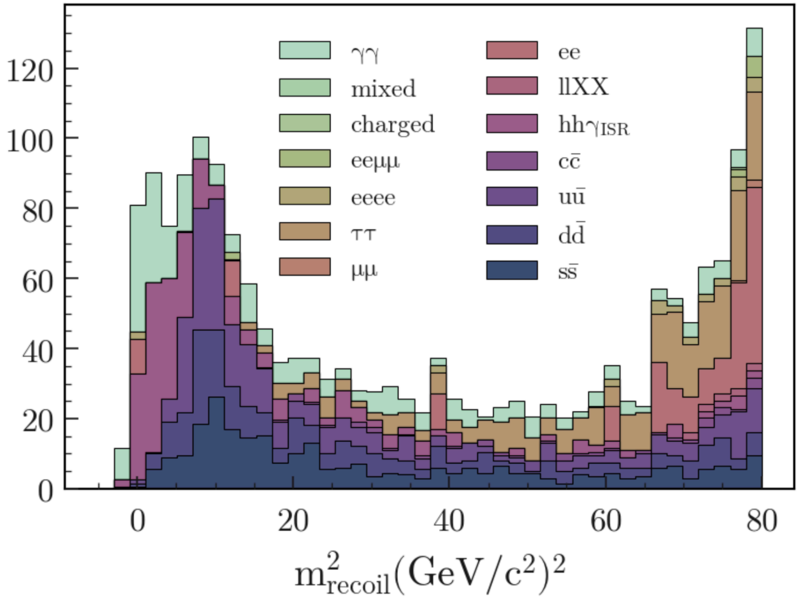


Cut on Event	Variable Description
$\frac{\vec{p}_{ISR}\cdot\vec{dr}_{V_0}}{\lvert p_{ISR}\rvert\lvert dr_{V_0}\rvert} < 0.88$	Angle between the ISR momentum and vertex position
$\theta_{ISR,V_0} > 0.5$	Angle between ISR and V_0's daughters' momenta
$\theta_{p_{recoil}} < 2.1$ rad and $\notin (0.45, 0.62)$ rad	Polar angle of recoil momentum
$n_{\text{ISR Photon}} = 1$	Number of ISR Photons in the event
$n_{\text{Tracks}} = 2$	Number of tracks in the event
$n_{\text{Signature}} \geq 1$	Number of events with displaced vertex + ISR signature

Outline


Next Steps
Upper Limit Setting
Signal Extraction
Background Suppression
iDM Explanation

Background: Composition

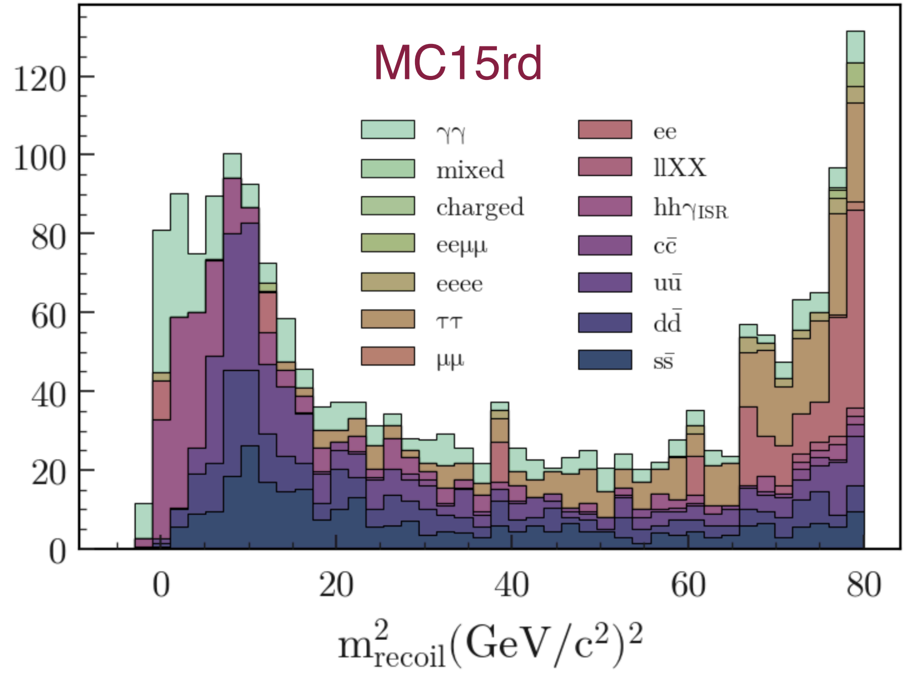


- While this iDM signature is clean, some standard model processes can fake it
- Example:
  - Direct radiative pair production with an undetected photon (i.e. $e^+e^- \rightarrow t^+t^-\gamma(\gamma)$ with $t = e, \mu, \pi$)
  - Photo-conversions ($\gamma \rightarrow e^+e^-$) from direct radiative pair production or radiative photon pair production ($e^+e^- \rightarrow \gamma\gamma(\gamma)$)
  - Meson Decays (e.g. $e^+e^- \rightarrow K^0_S K^0_L \gamma$ where $K^0_S \rightarrow \pi^+\pi^-$)

SOURCE: M. Duerr et al. "Invisible and displaced dark matter signatures at Belle II". https://arxiv.org/pdf/1911.03176.pdf

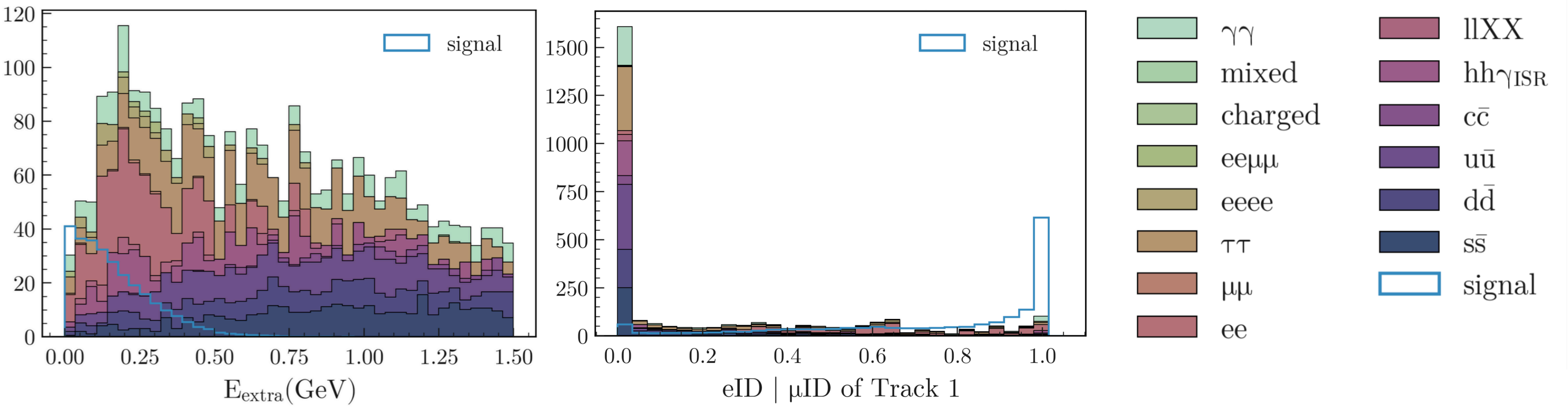
Background vs. Signal Separation

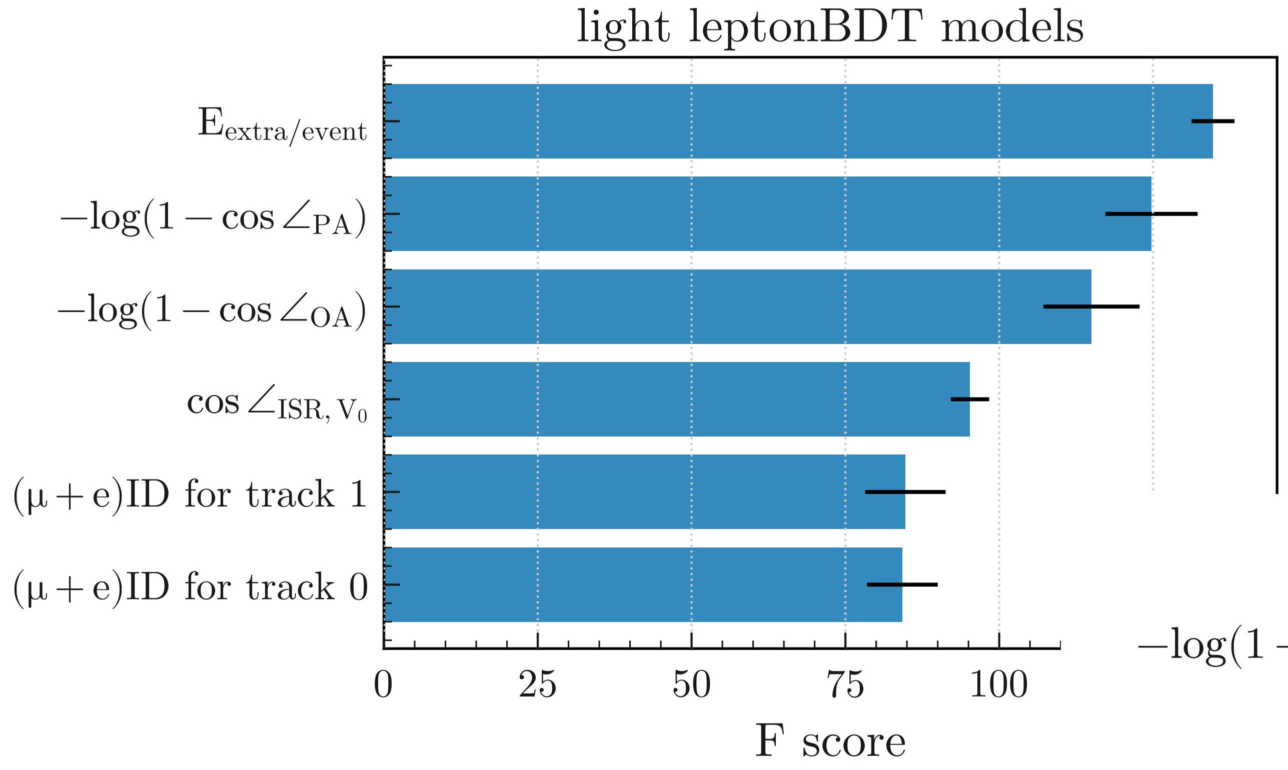
- Applying all aforementioned pre-selections, we go from the entire Belle 2 dataset down to something manageable
- For more fine-tuned cuts to further suppress backgrounds:
  - Introduce a BDT using physically meaningful variables
  - Make signal model dependent cuts

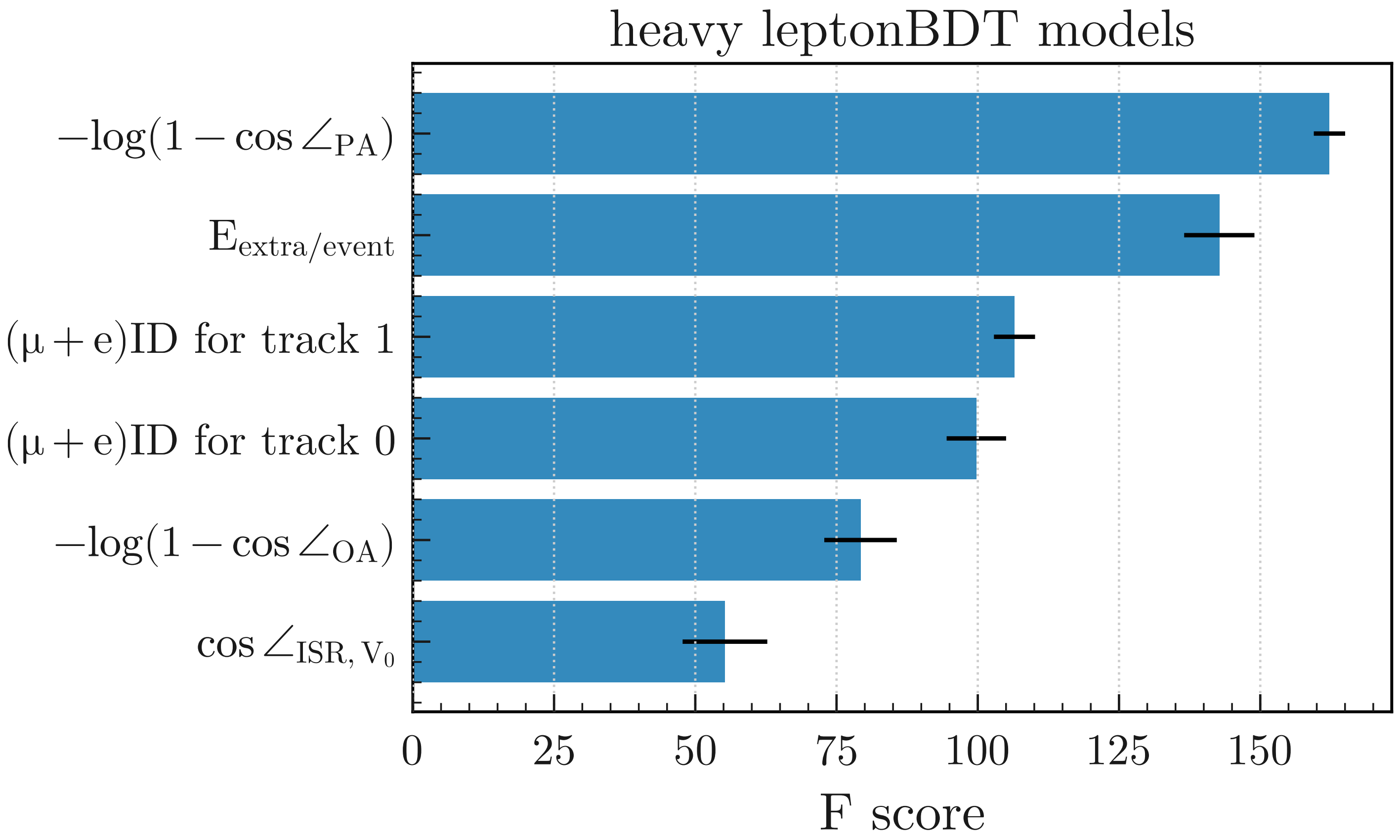


Features for BDT

- Total Extra Energy (GeV): For any additional tracks and photons (besides those used by signal), sum over their energies
- LeptonID: Checking the particle ID of each track forming the vertex

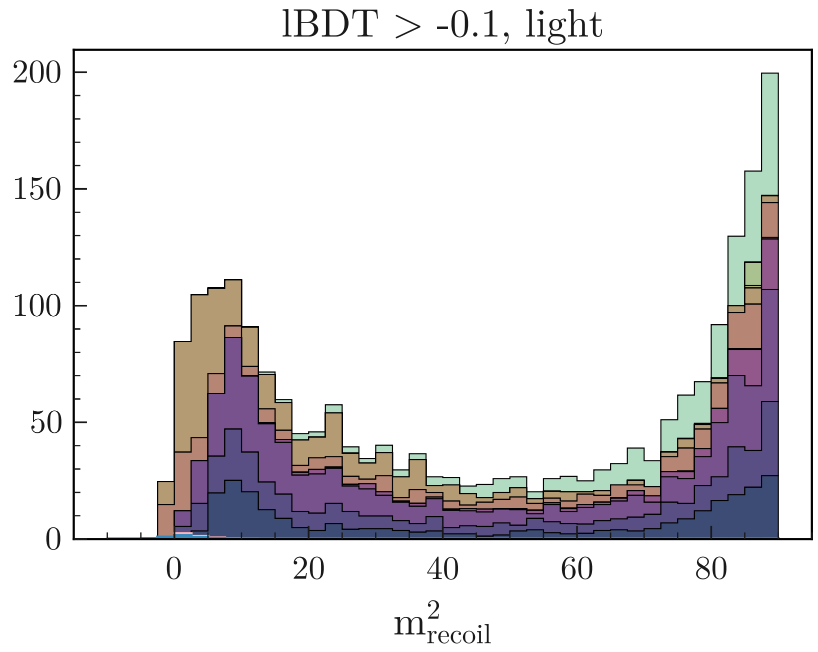


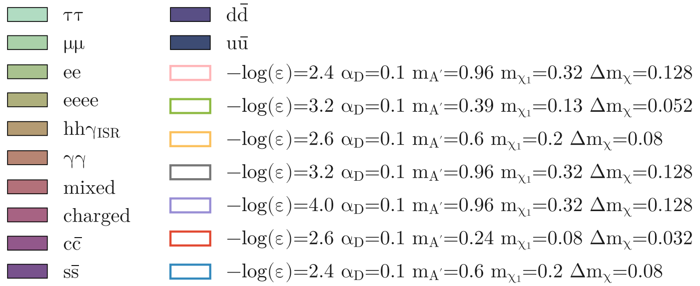


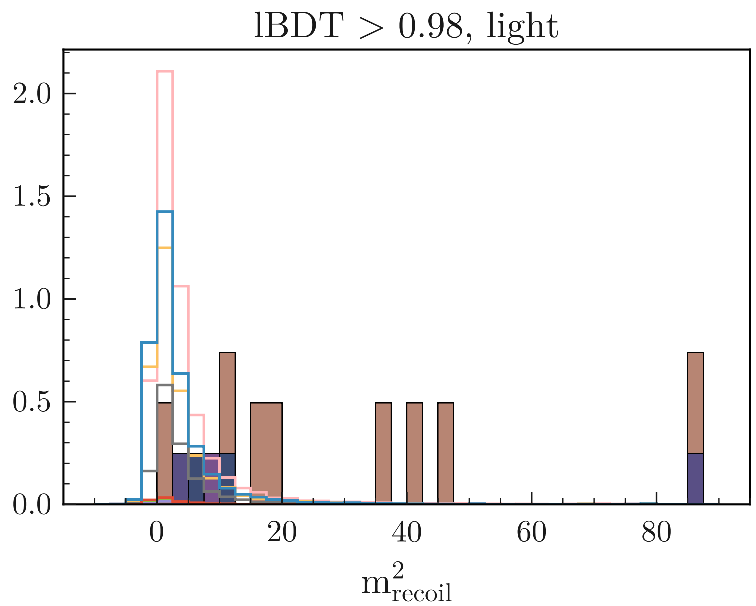


- light: $m_{A'} \leq 0.96$
- medium: $0.96 \leq m_{A'} \leq 3.9$
- heavy: $m_{A'} \leq 3.9$
- (units are GeV/c^2)

link

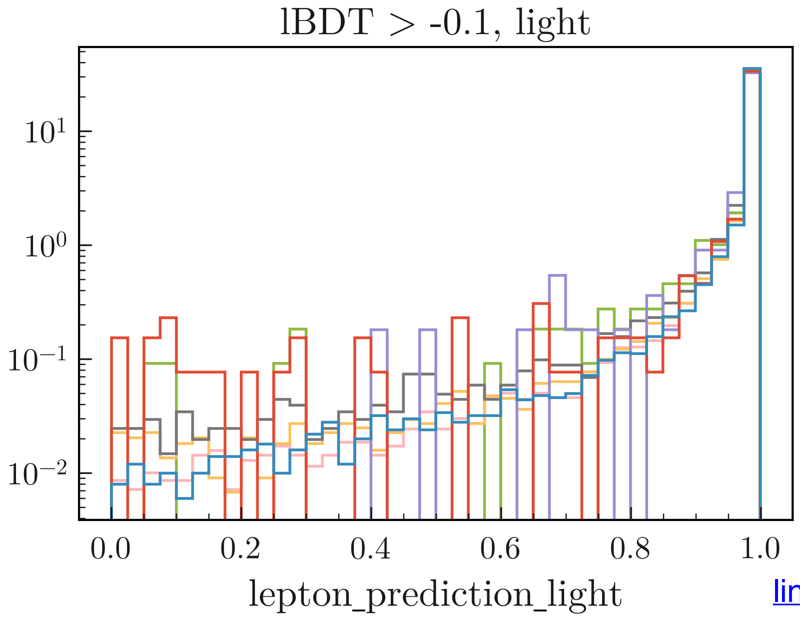




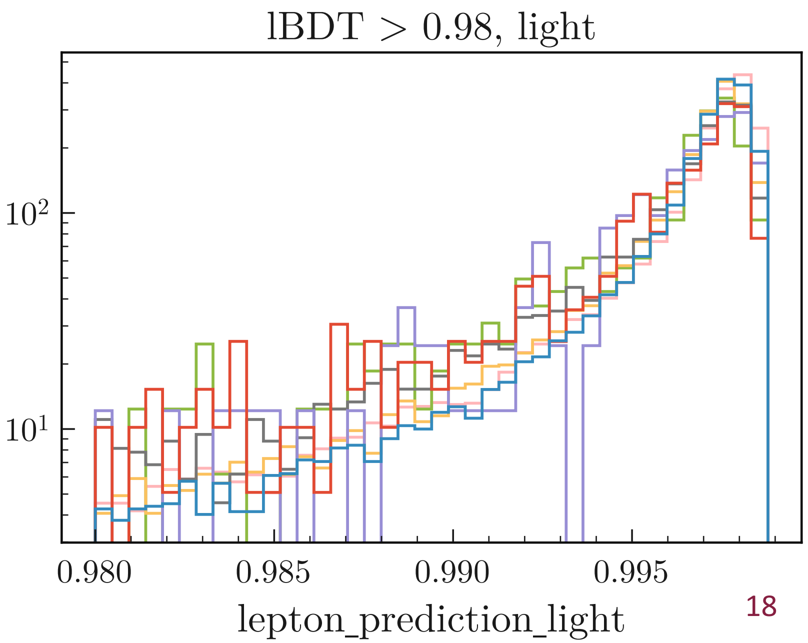


BDT performance
(used Punzi FOM to optimize)

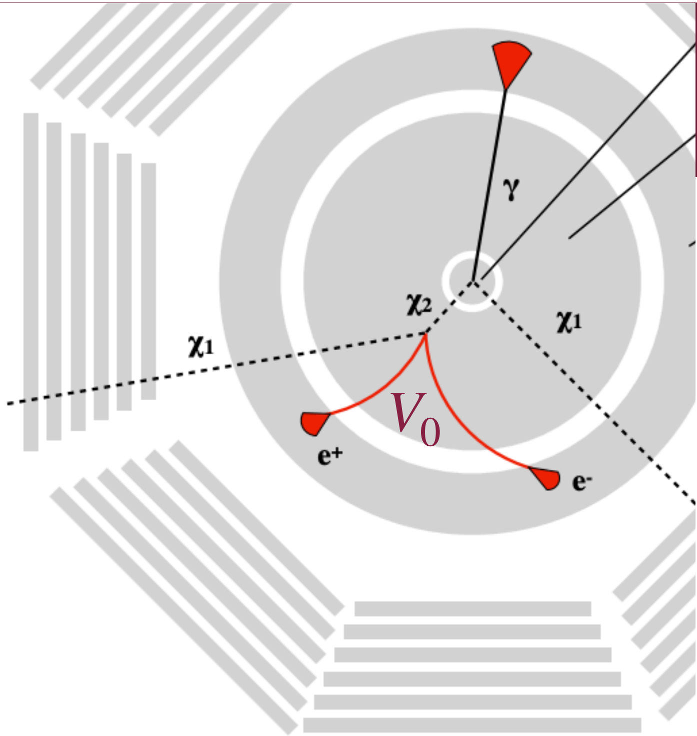
link



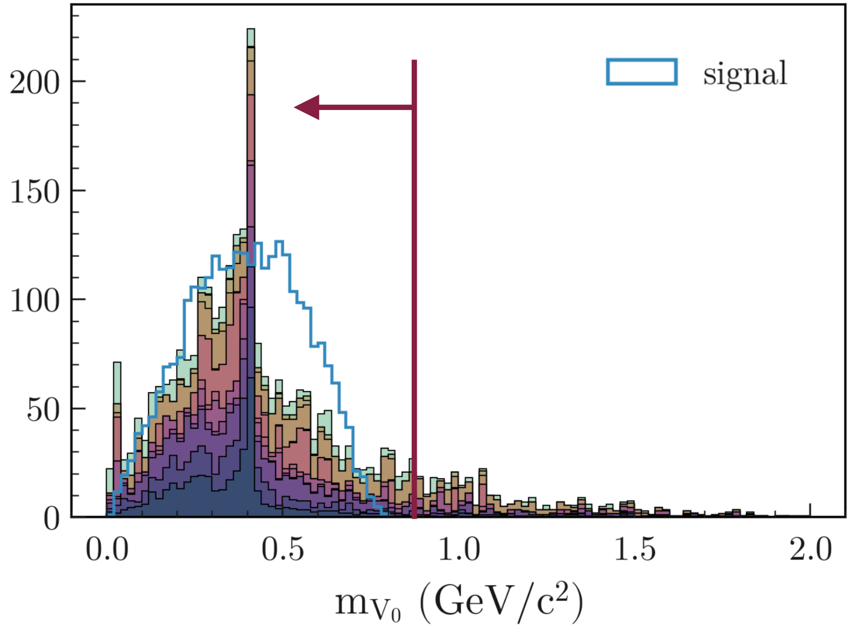
link

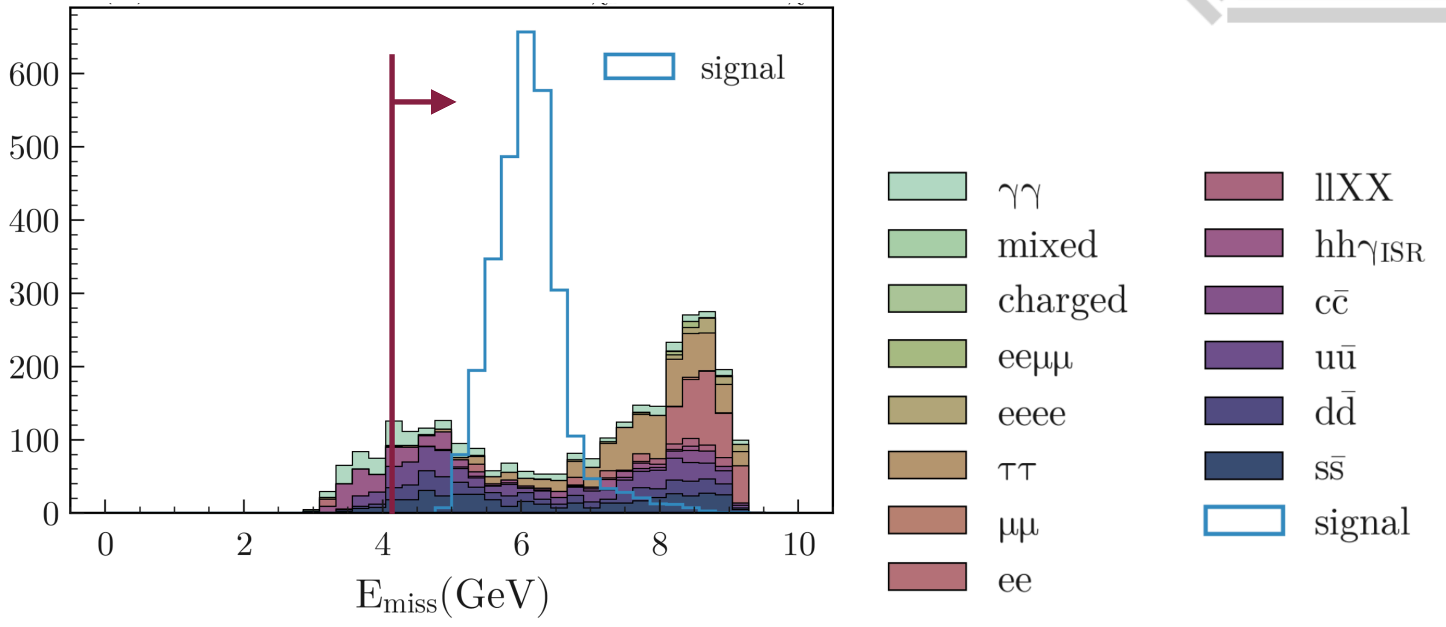


Signal Dependent Cuts



- In addition, we can also do signal dependent cuts
  - Vertex Mass $(m_{V_0}) \leq \Delta m_\chi$,
  - $E_{miss} = E_{\chi_{1,I}} + E_{\chi_{1,II}} \leq 2 m_{\chi_1} c^2$



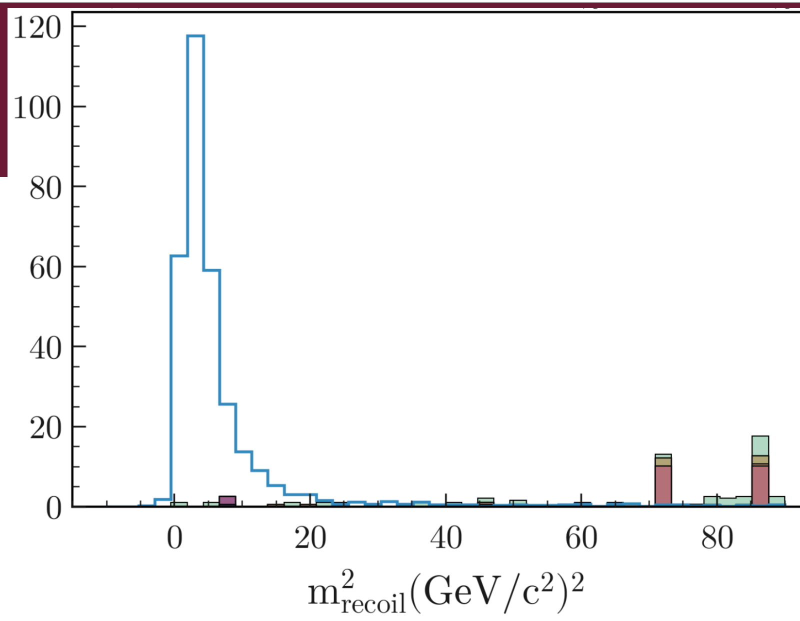


$-\log(\varepsilon)=3.6\ \alpha_D=0.1\ m_{Z'}=6.0\ m_{\chi_1}=2.0\ \Delta m_\chi=0.8$

Outline

- iDM Explanation
- Background Suppression
- Signal Extraction
- Upper Limit Setting
- Next Steps

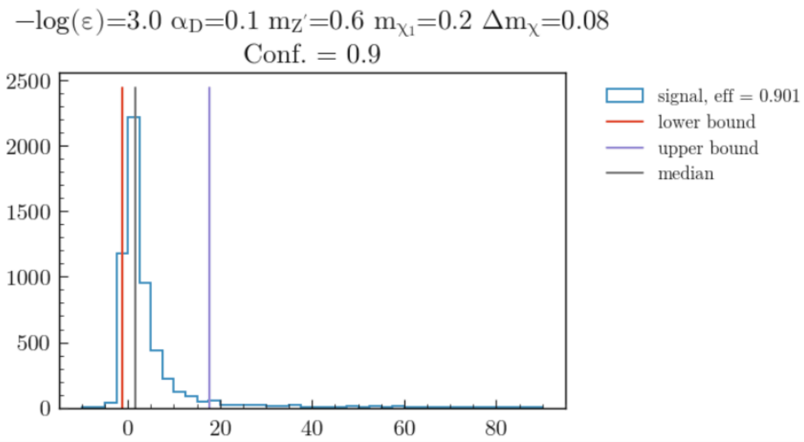
Signal Extraction

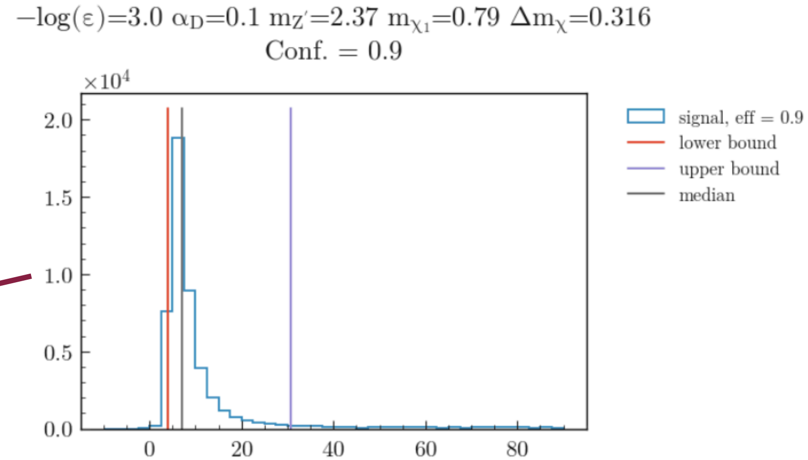


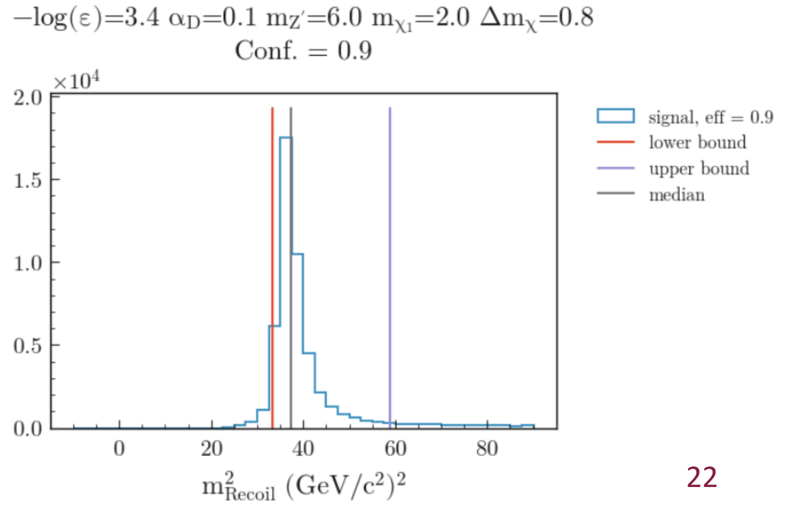
- Given how few events we have remaining, let's do a single bin counting experiment
- For a given signal window and its dependent cuts, estimate signal vs. background
- $\mathscr{L} = P_{\propto}(\mu + b \,|\, n) \times P_{\propto}(\tau b \,|\, m)$
  - 1st term: data used to extract μ signal events with b backgrounds
  - 2nd term: information from MC or a control sample (where τ is a pre-determined scale factor)
- pyhf is used

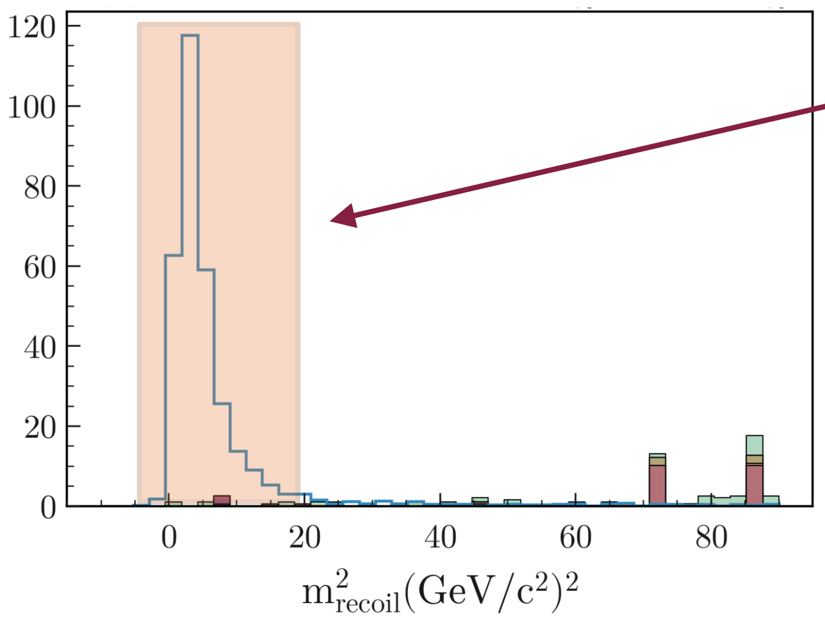
Defining Signal Window

- Since we’re not fitting, we just need to define the region where we count events
- Do this by defining 90% CL (90% of events w/ equal areas from the median)

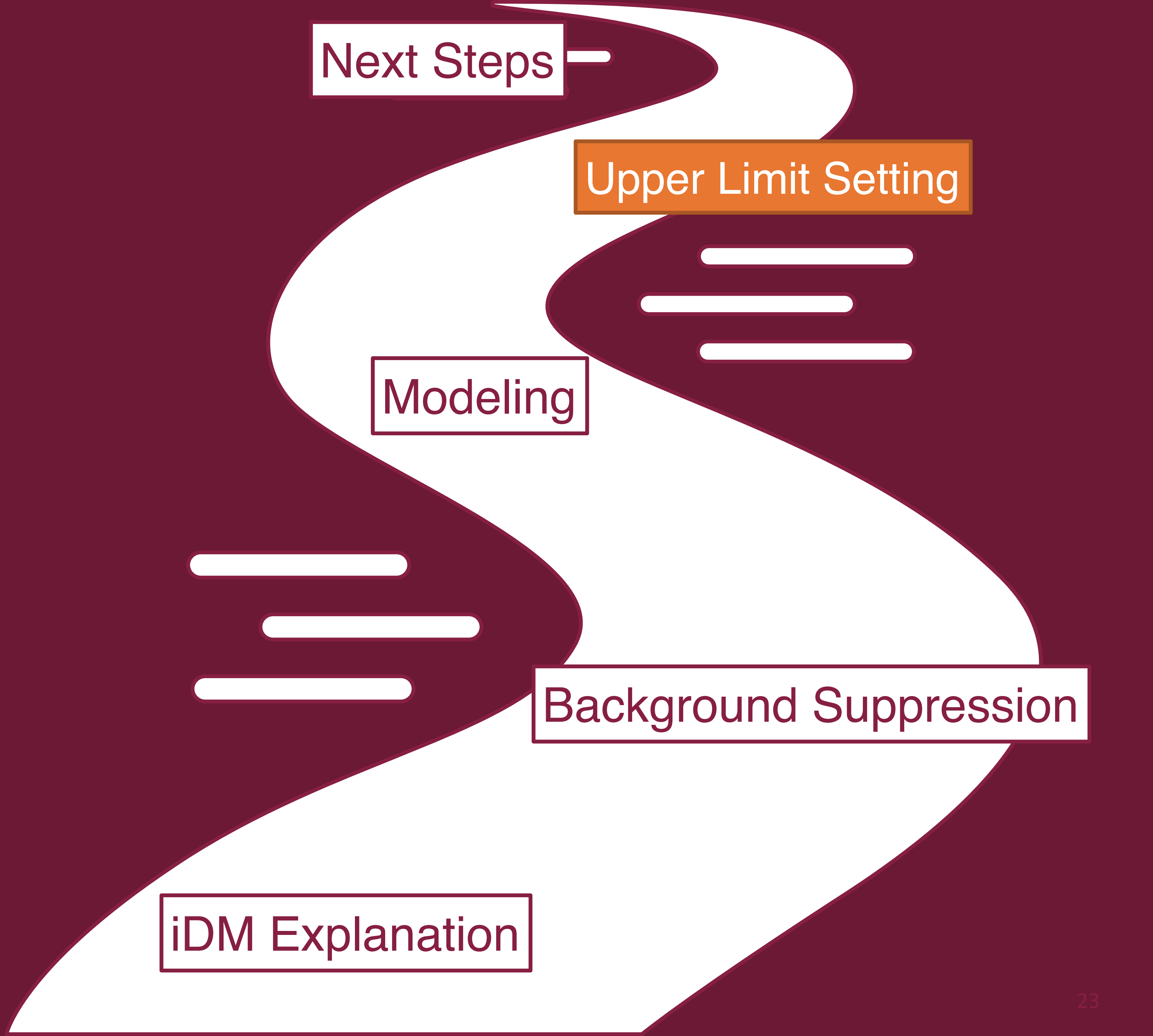








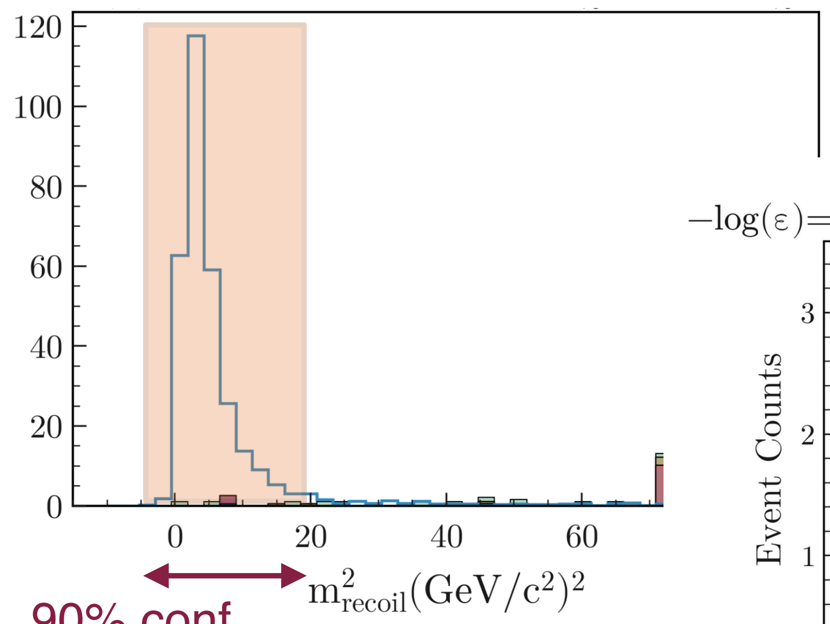
Outline


Next Steps
Upper Limit Setting
Modeling
Background Suppression
iDM Explanation

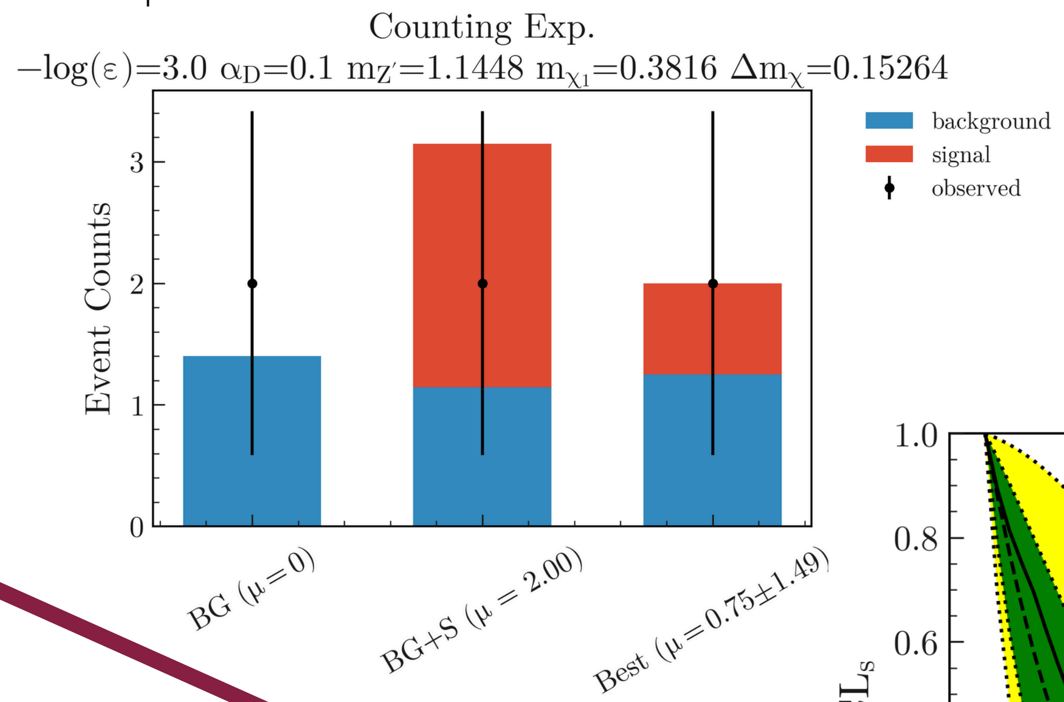
Upper Limit

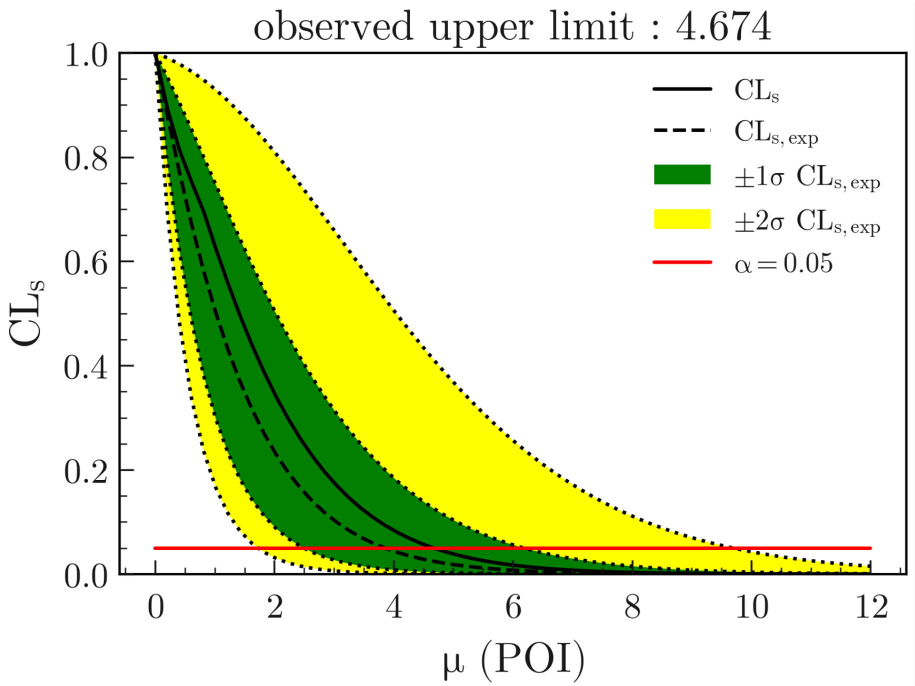
- With our selection, our signal extraction method, we should have some idea of how will perform.
- Strategy:
  - Bootstrap our background MC samples and create a toy dataset
  - Apply selection and obtain a signal yield -> convert to upper limit
  - Do this N times and obtain some default set of values
  - Use the median value from these toy trials for sensitivity plots

What a given trial looks like

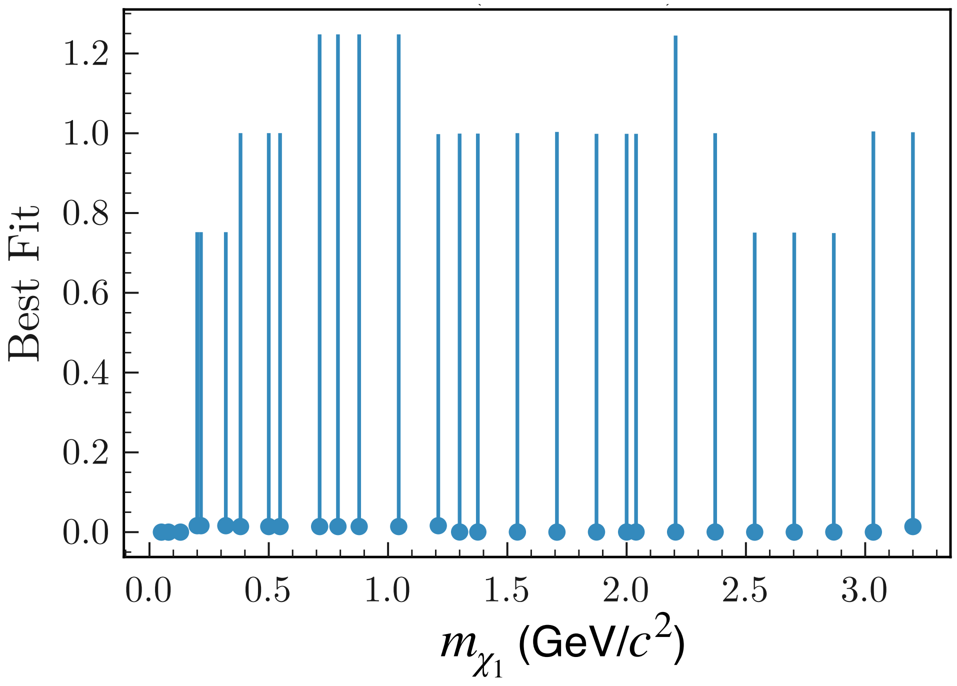


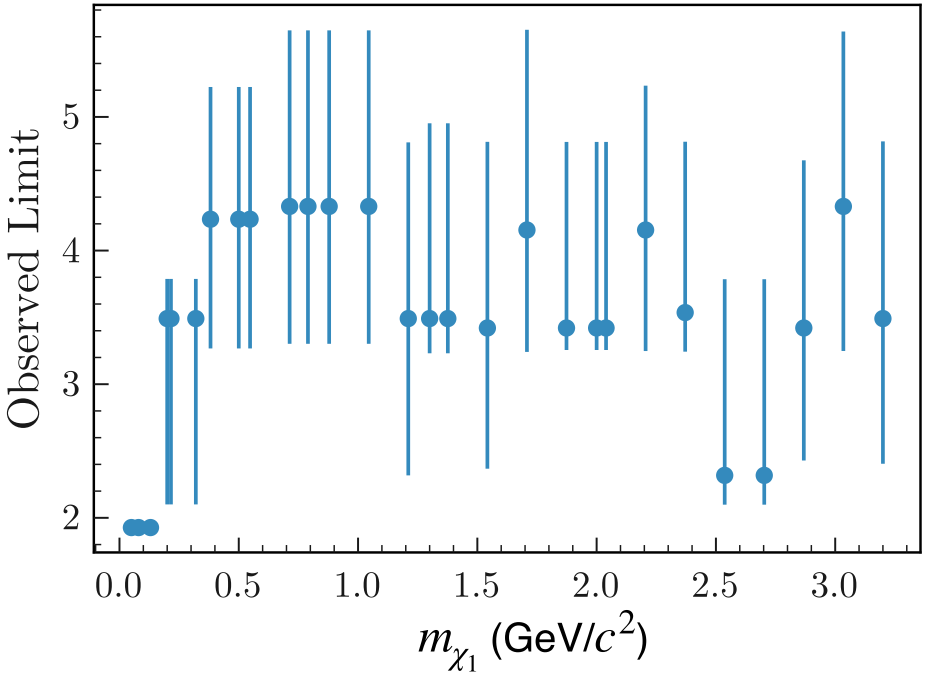
90% conf. interval





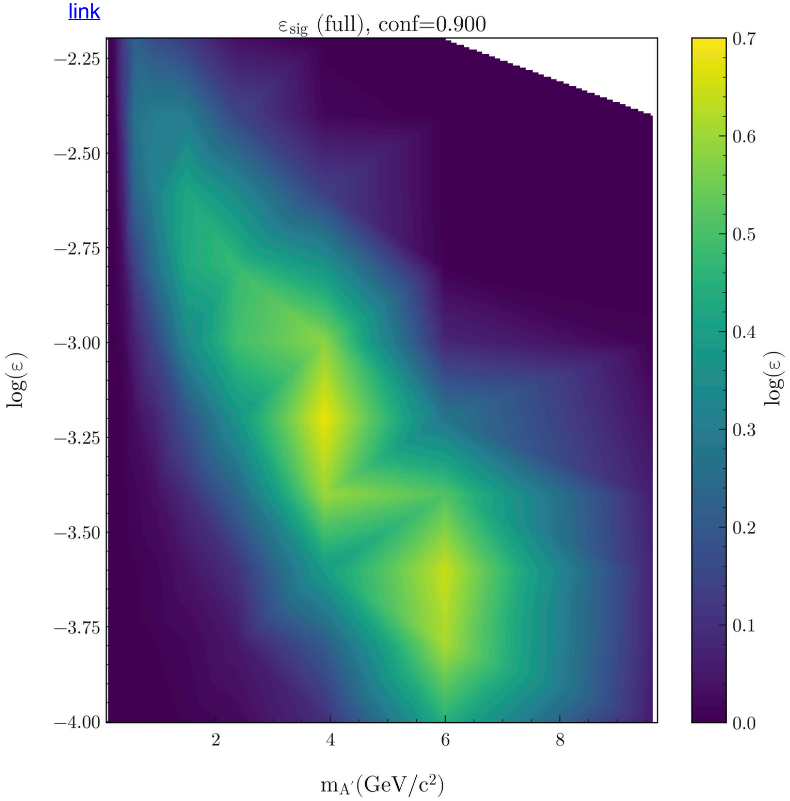
link

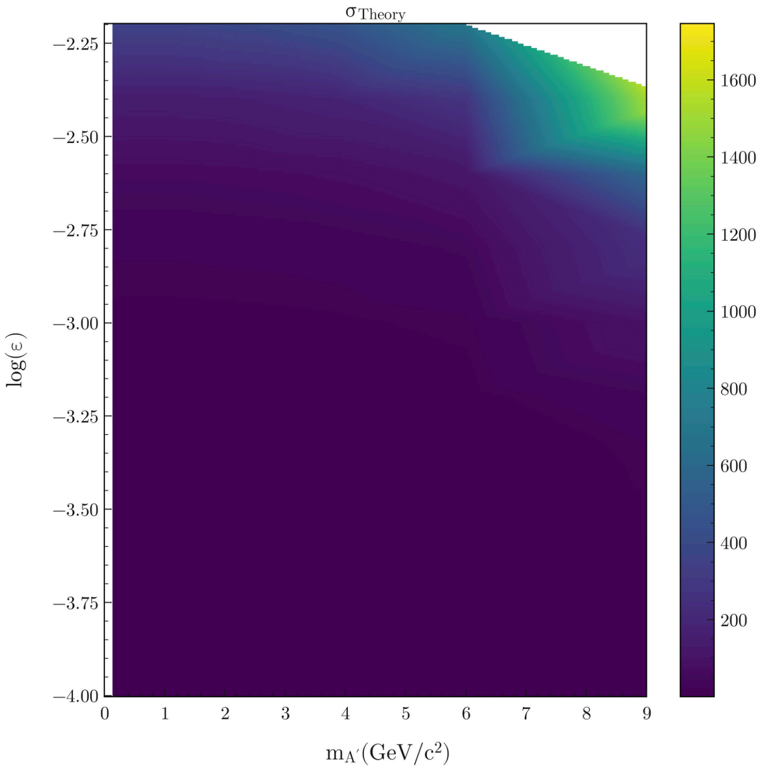


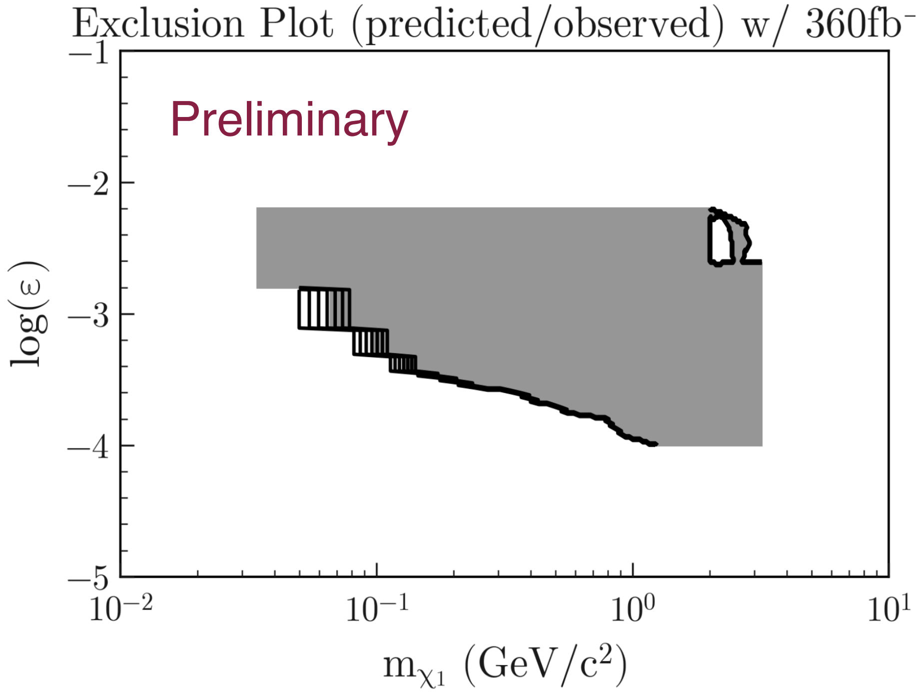


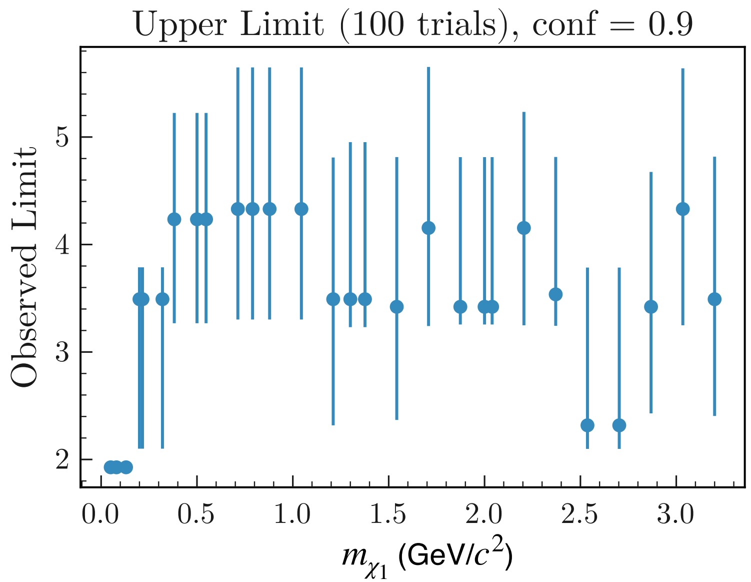
- After performing **100 samplings per model**, we get some 95% observed UL
  - (from best fit signal value and knwoeldge of background)
- From here, we want to convert the median upper limit value to a cross section that relates # of events to something meaningful
  - i.e. cross-section

link

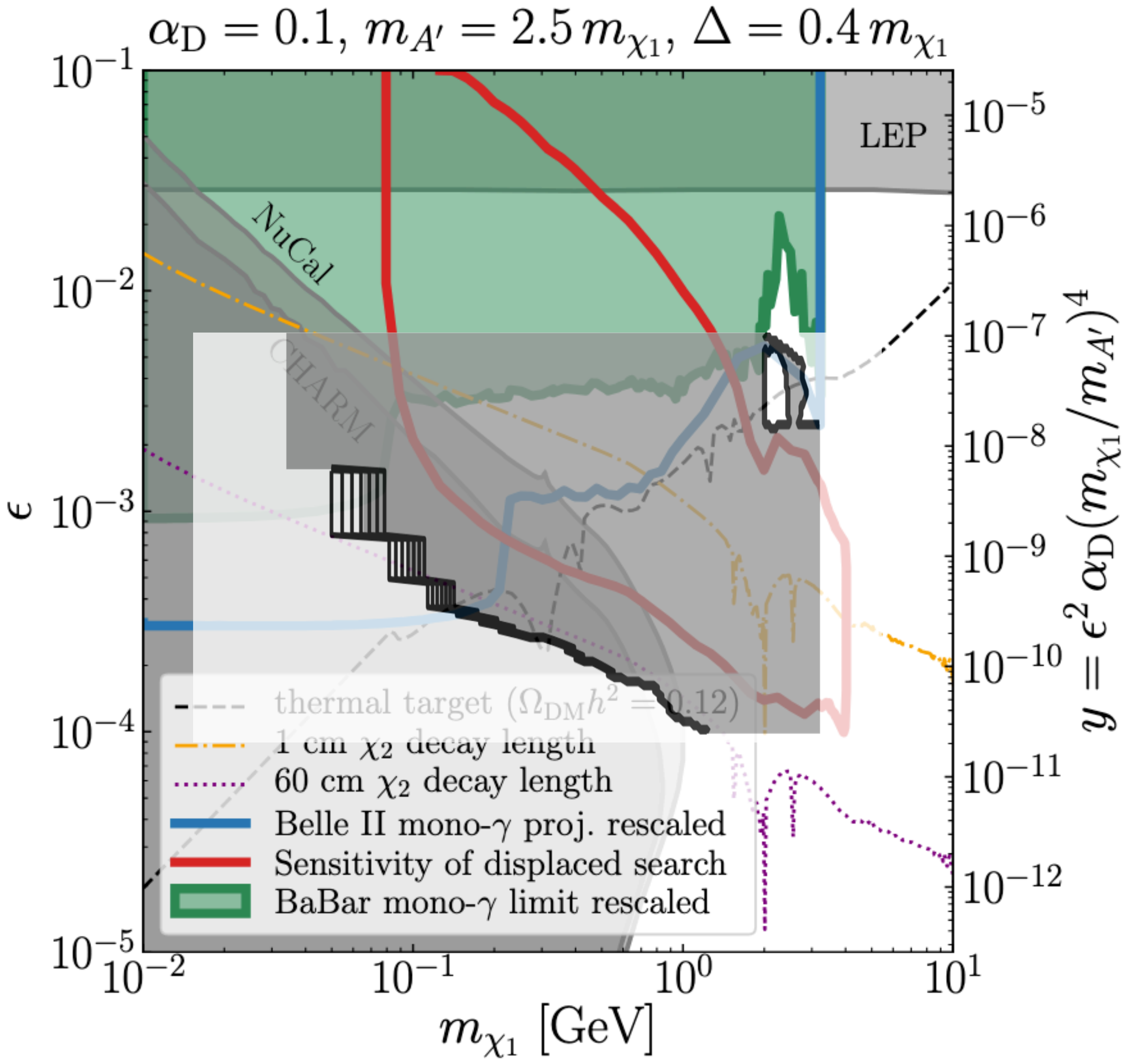








$$\sigma_{observed} = \frac{U(m_{A'}, \Delta m_{\chi}, m_{\chi 1})}{\eta(m_{\chi 1}, m_{A'}, \epsilon, \Delta m) \times L}$$



$\alpha_D = 0.1,\ m_{A'} = 2.5\,m_{\chi_1},\ \Delta = 0.4\,m_{\chi_1}$
LEP
NuCal
CHARM
ϵ
$y = \epsilon^2 \alpha_D (m_{\chi_1}/m_{A'})^4$
m_{χ_1} [GeV]
thermal target ($\Omega_{DM}h^2 = 0.12$)
1 cm χ_2 decay length
60 cm χ_2 decay length
Belle II mono-γ proj. rescaled
Sensitivity of displaced search
BaBar mono-γ limit rescaled

Outline

- iDM Explanation
- Background Suppression
- Modeling
- Upper Limit Setting
- Next Steps

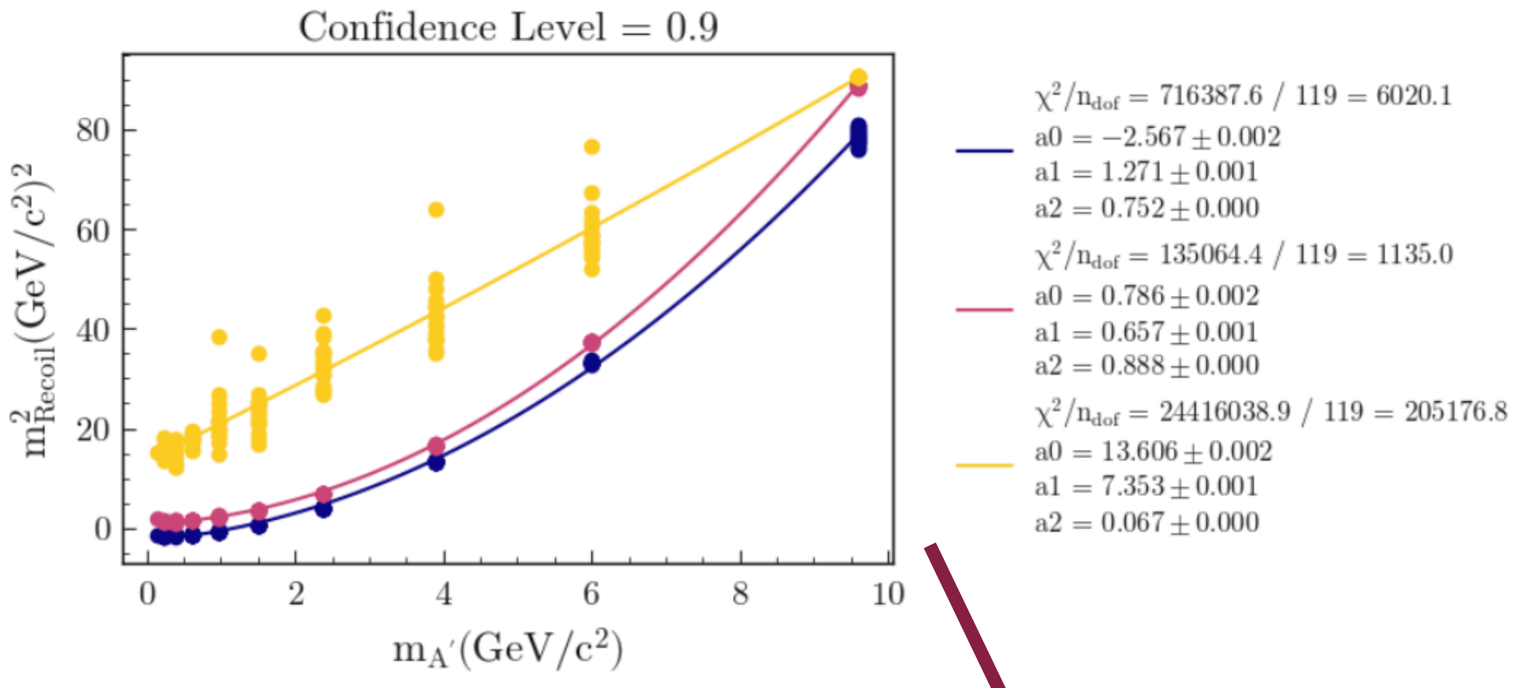
Next Steps

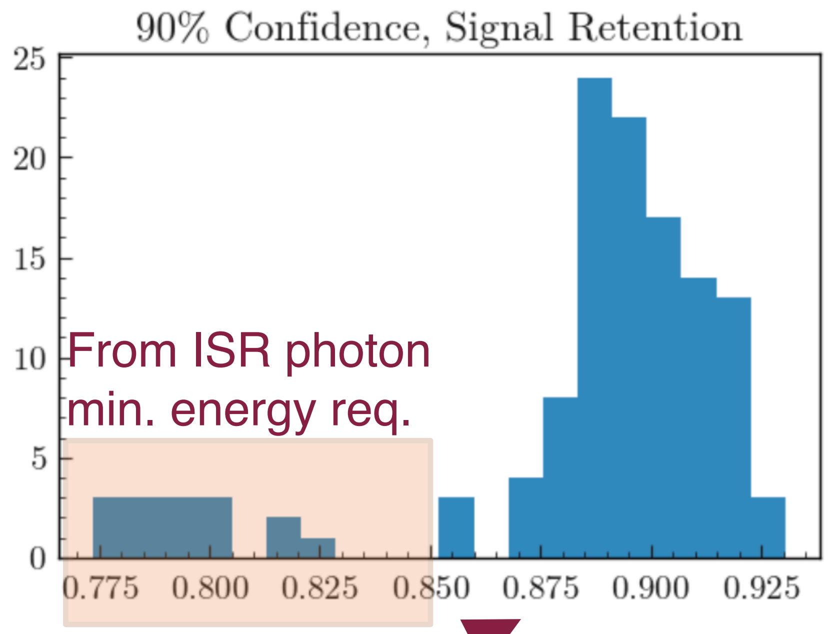
- Control Sample (K-Short Veto, γγ sample, ττ)
- Understanding BDT systematics
- Systematics?
  - Additional sanity checks on produced signal
- Better contour plotting code…
- Write…

That's all folks!

Extra

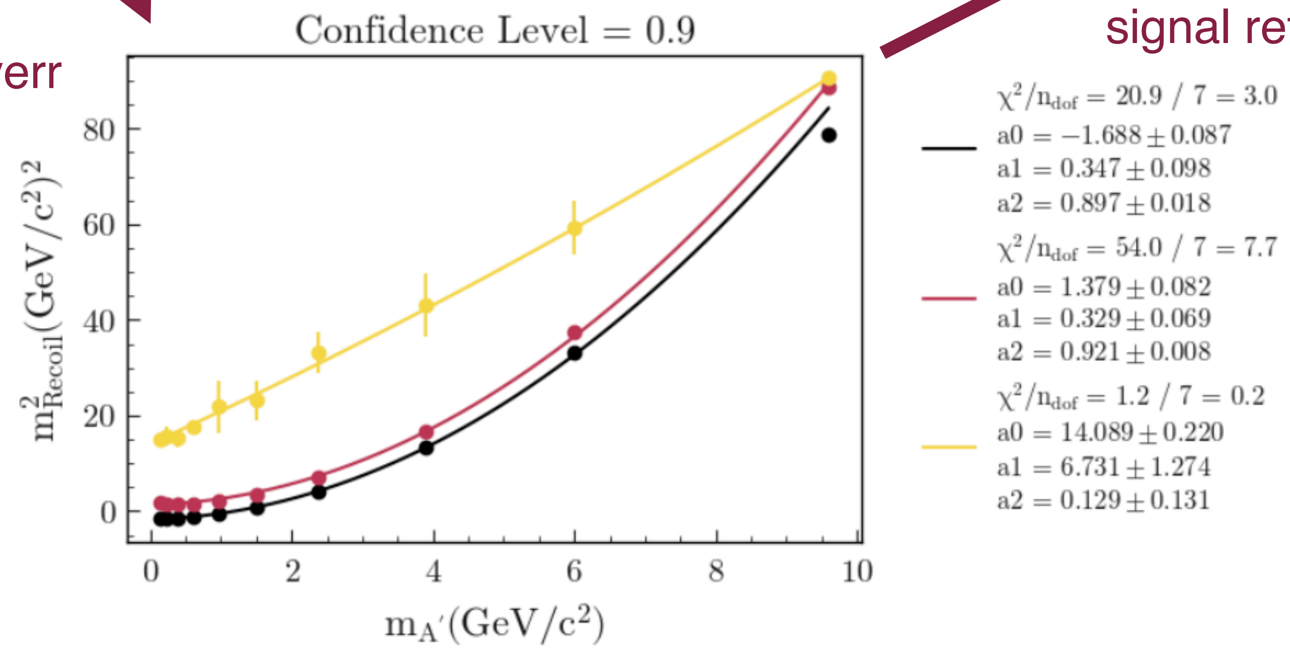
Analysis Details





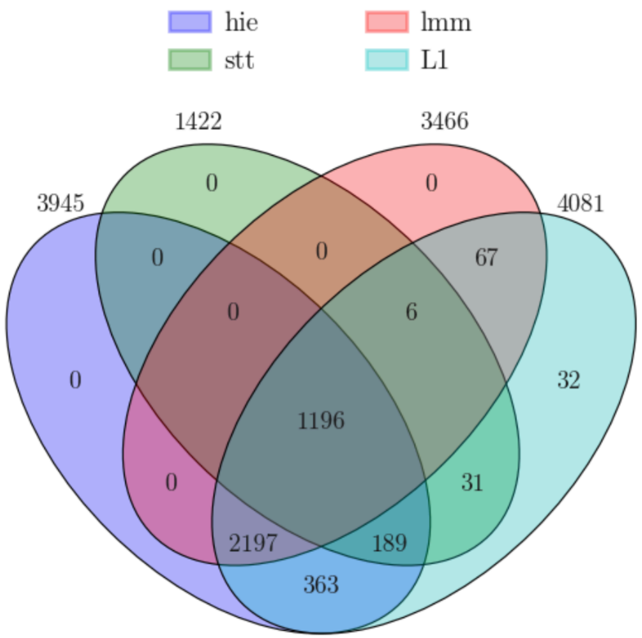
grouping by mA'
averaging and std for y, yerr

reapply and check
signal retention

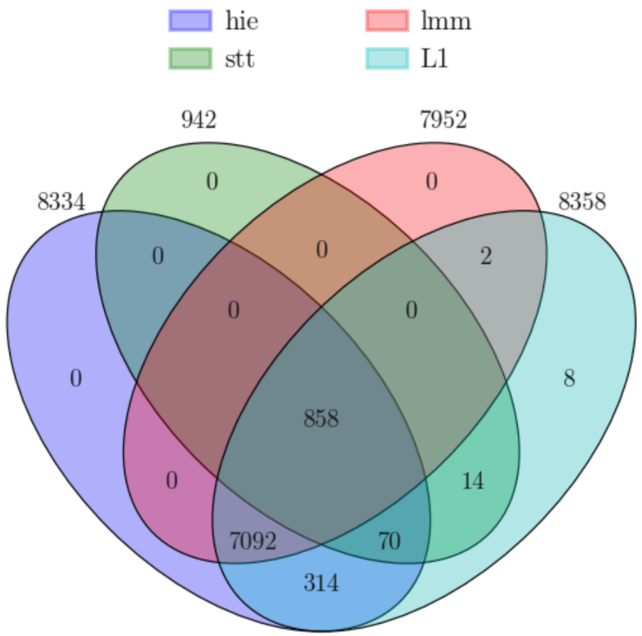


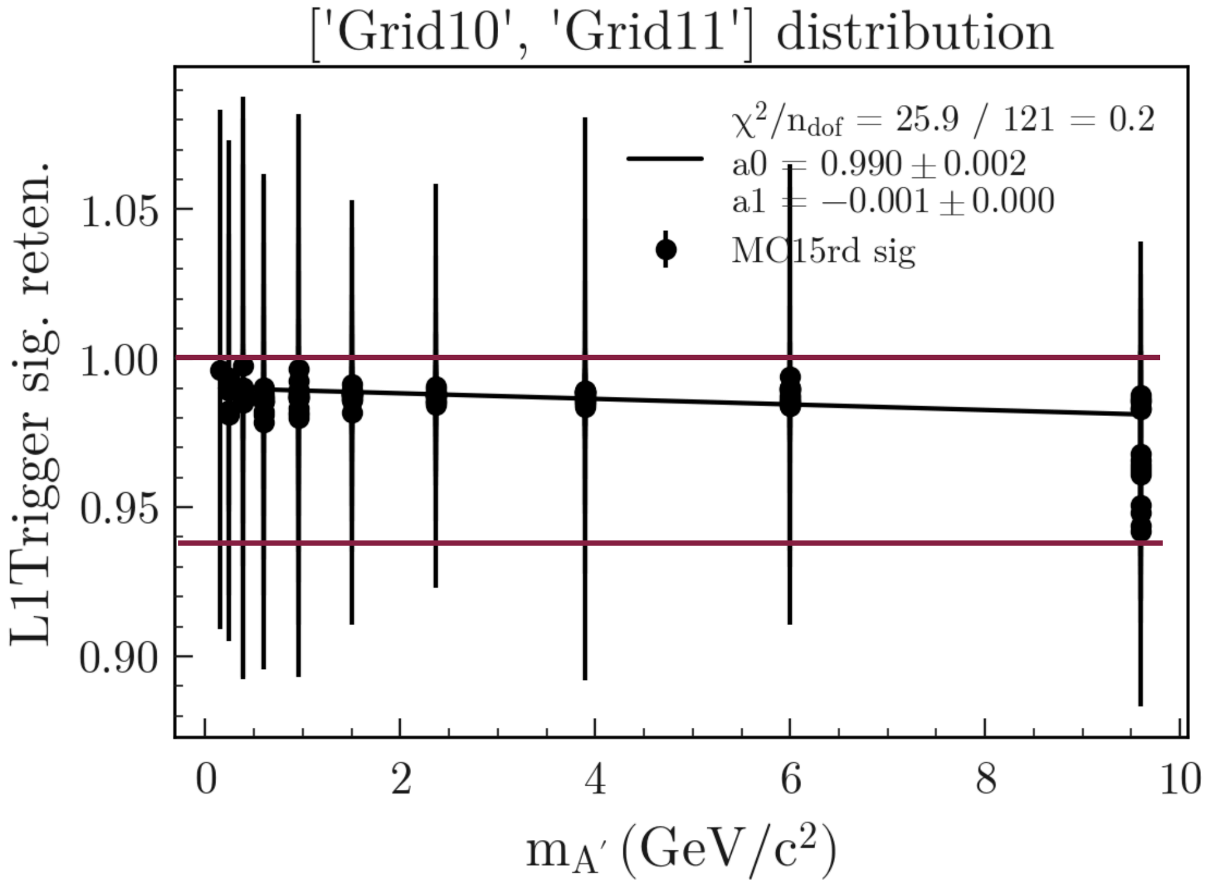
slide 22

Venn Diagram for $-\log(\varepsilon)=3.2\ \alpha_D=0.1\ m_{A'}=9.6\ m_{\chi_1}=3.2\ \Delta m_\chi=1.28$

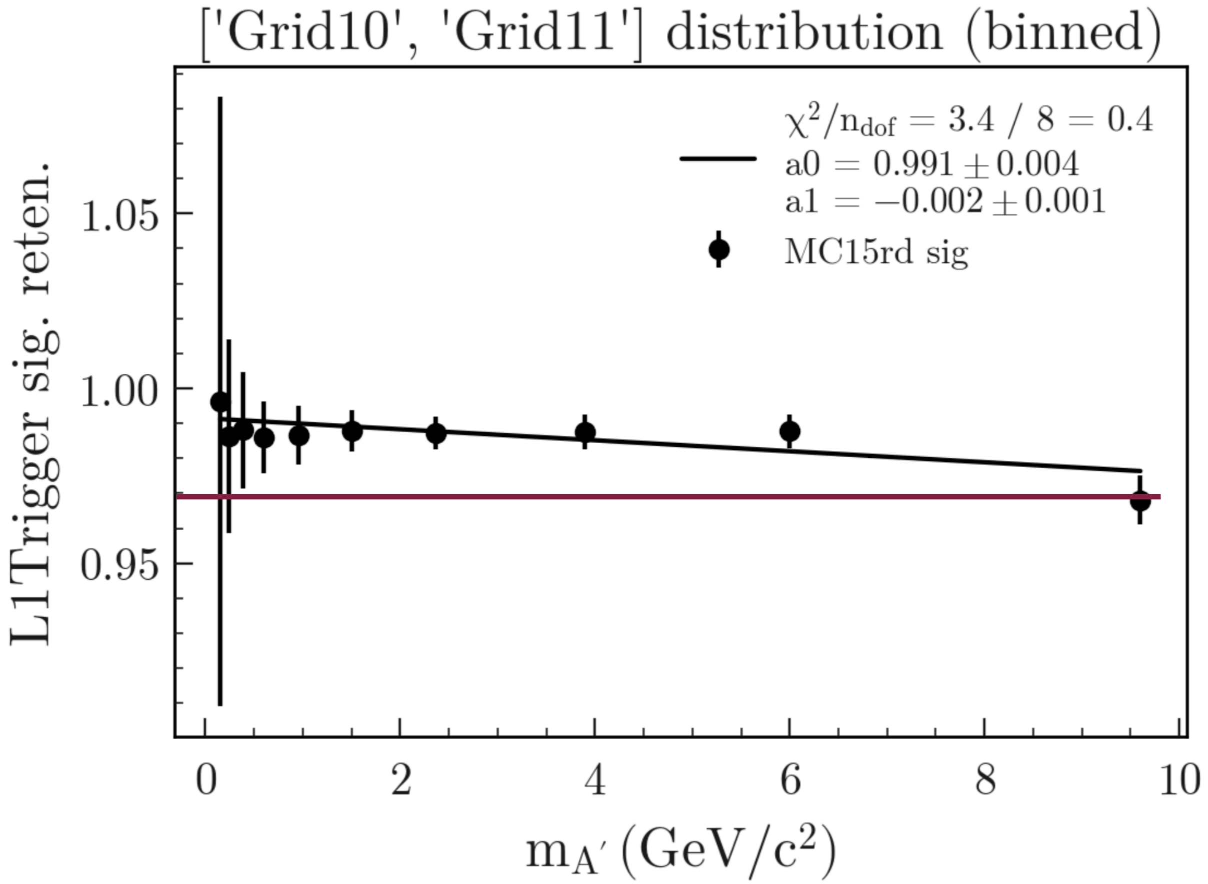


Venn Diagram for $-\log(\varepsilon)=4.0\ \alpha_D=0.1\ m_{A'}=3.9\ m_{\chi_1}=1.3\ \Delta m_\chi=0.52$

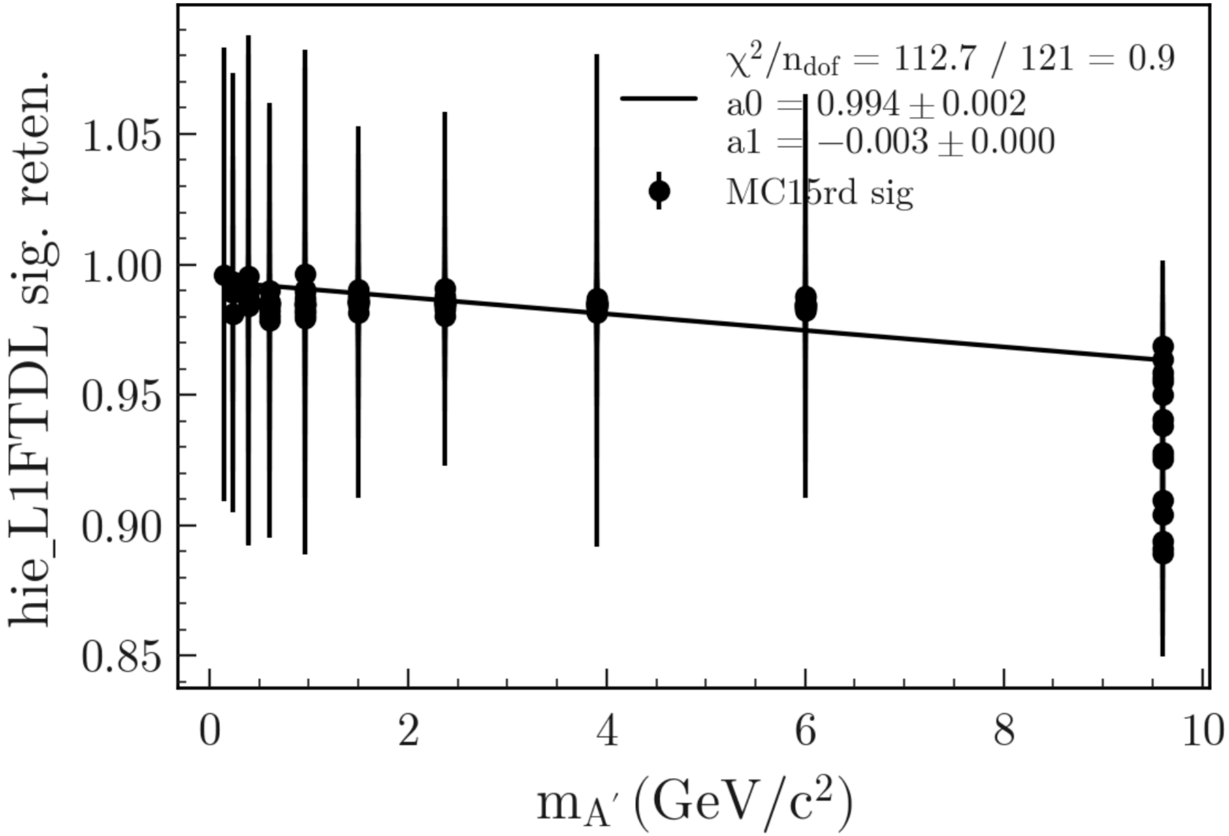




- data point = $\langle eff \rangle \pm \sigma_{eff}$ $\forall$ models
- $\sigma_{eff} = (n/d)\sqrt{1/n + 1/d}$
- Showed all ε, m_χ

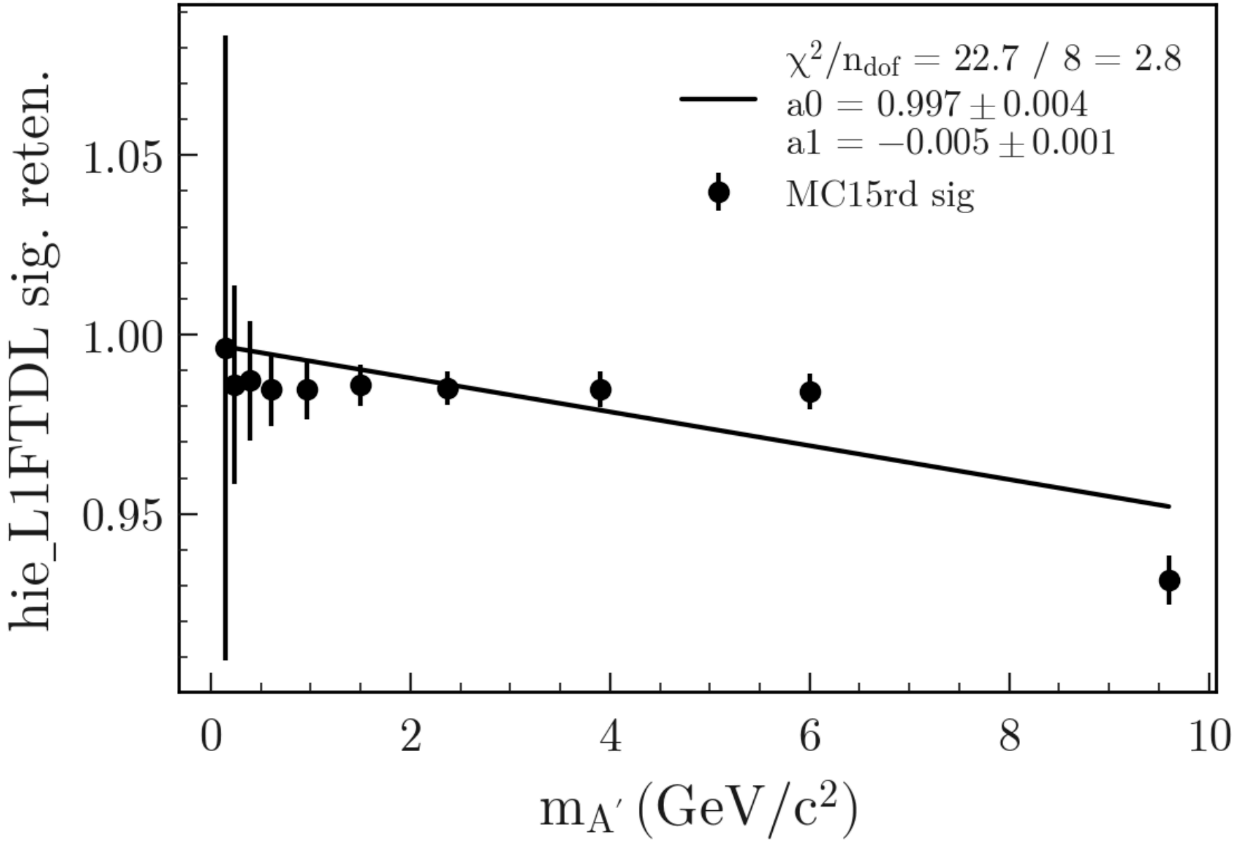


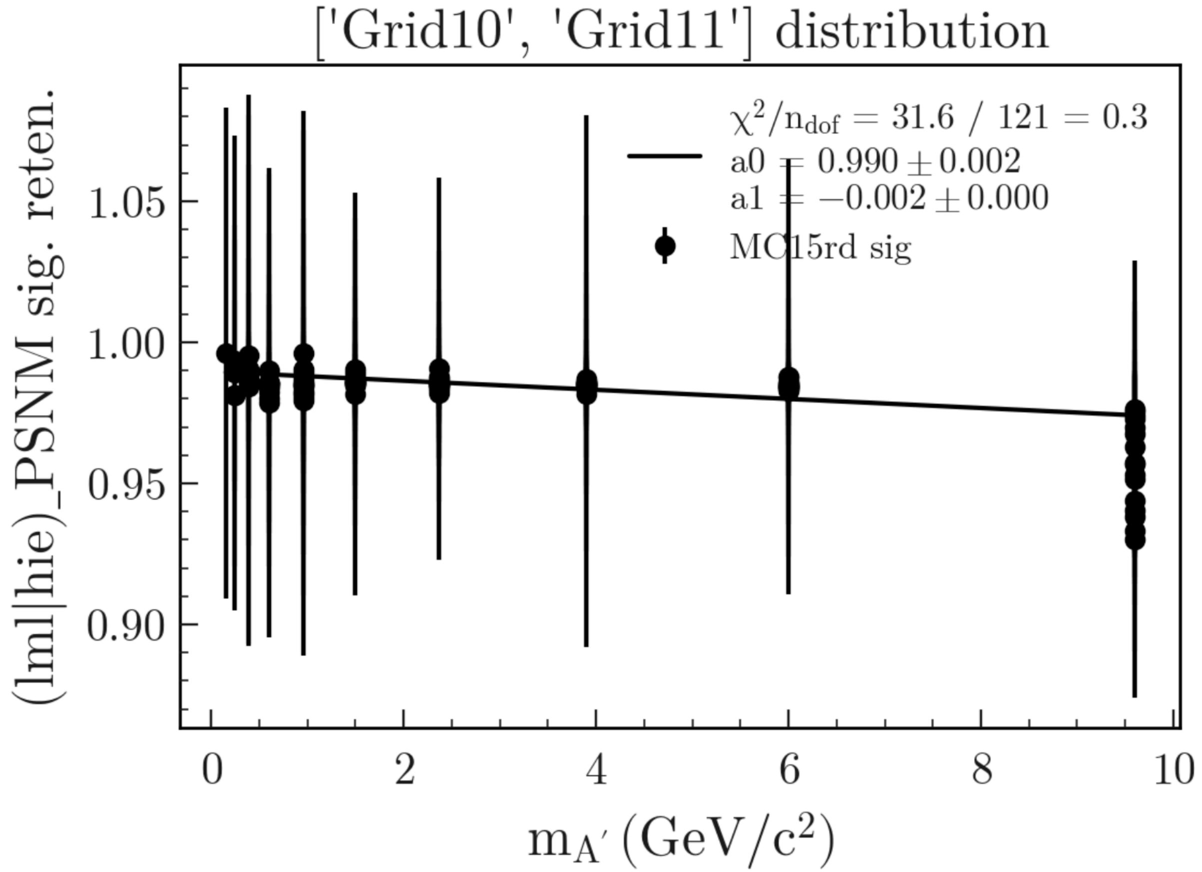
- $\langle eff \rangle_{w=1/\sigma_{eff}} \pm \langle \sigma_{eff} \rangle / \sqrt{n}$
- Averaged over ε, m_χ



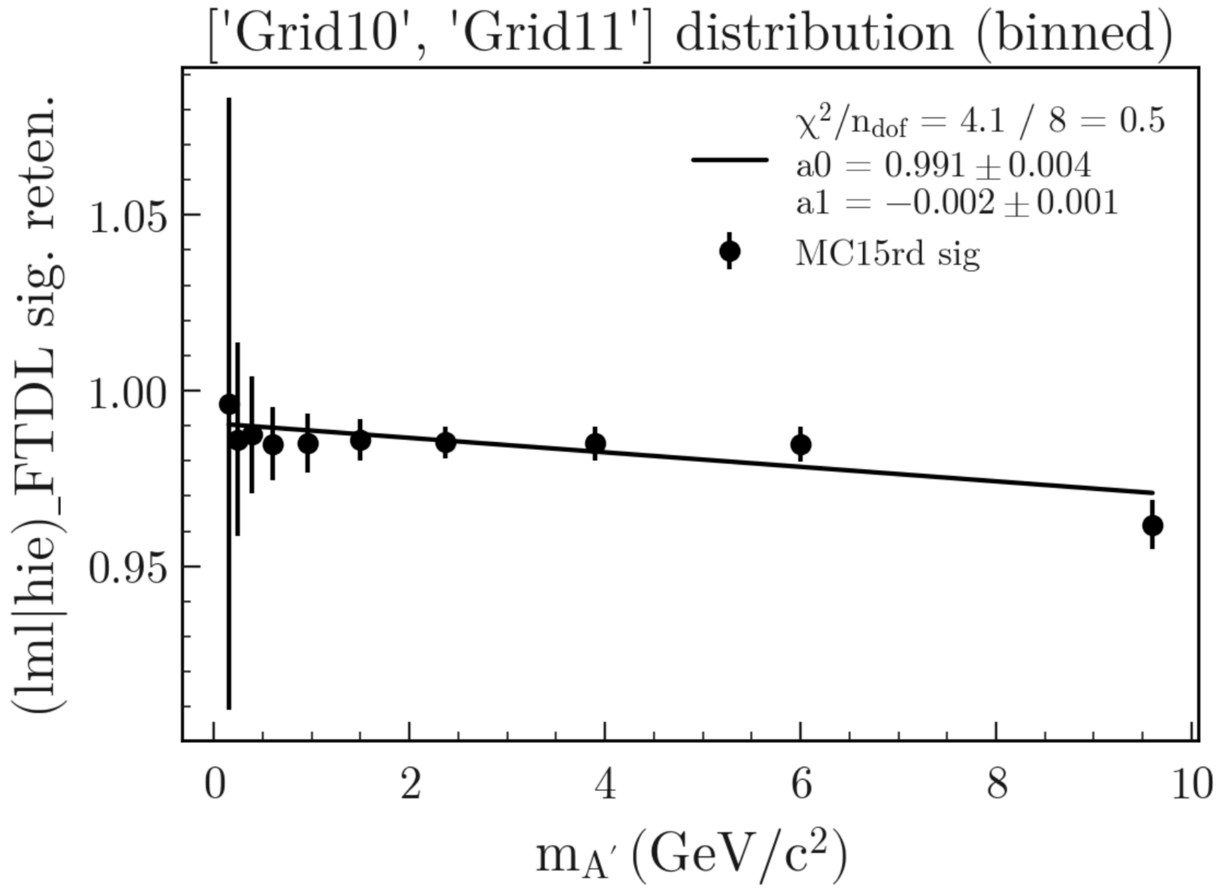
- data point = $n/d \pm \sigma_{eff}\ \forall$ models
- $\sigma_{eff} = (n/d)\sqrt{1/n + 1/d}$
- Showed all ε, m_χ

- $\langle n/d \rangle_{w=1/\sigma_{eff}} \pm \langle \sigma_{eff} \rangle / \sqrt{n}$
- Averaged over ε, m_χ
- Let me check LOWE'S to check





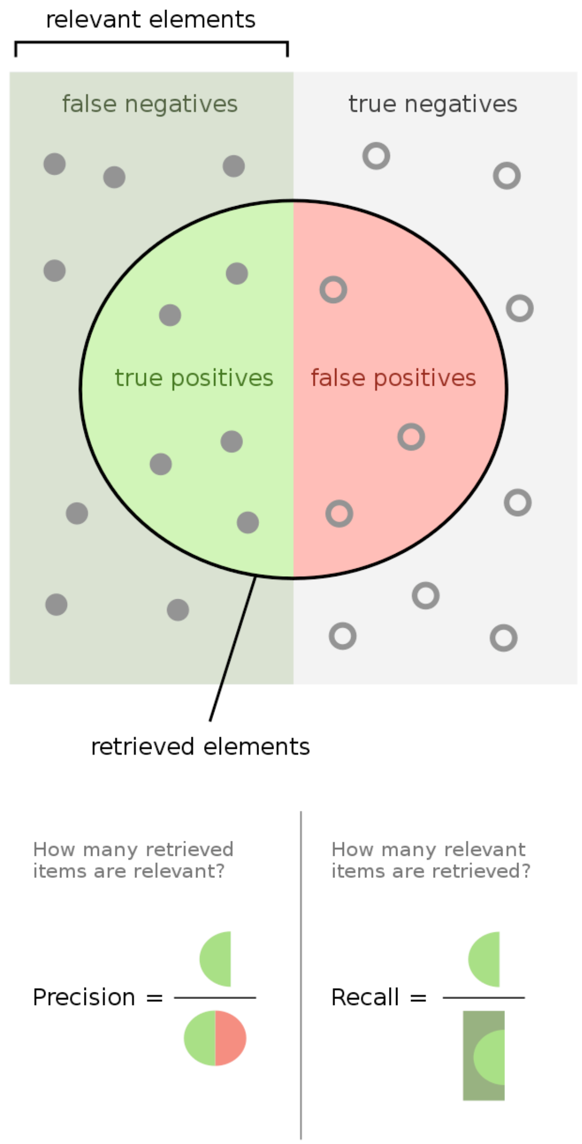
- data point = $n/d \pm \sigma_{eff}$ $\forall$ models
- $\sigma_{eff} = (n/d)\sqrt{1/n + 1/d}$
- Showed all ε, m_χ



- $\langle n/d \rangle_{w=1/\sigma_{eff}} \pm \langle \sigma_{eff} \rangle / \sqrt{n}$
- Averaged over ε, m_χ
- Let me check LOWE'S to check

What is an F-score?

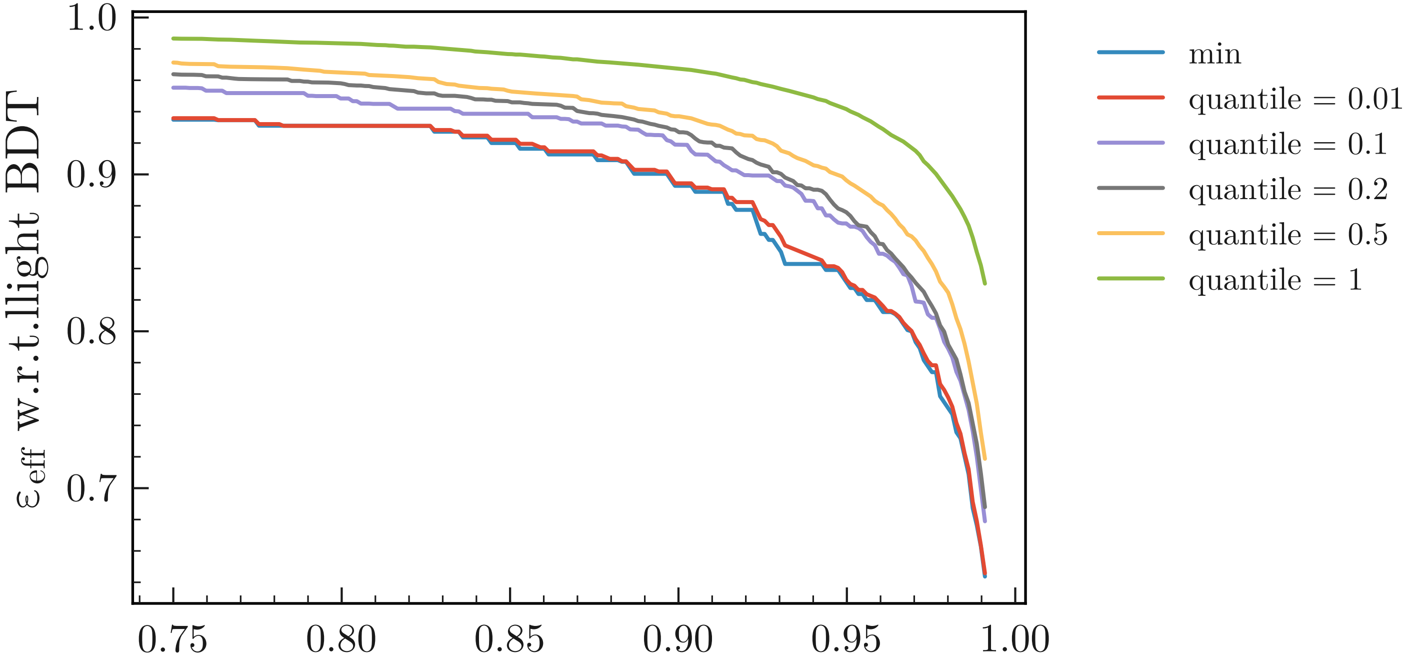
The F-score (also known as the F1 score or F-measure) is a **metric used to evaluate the performance of a Machine Learning model**. It combines precision and recall into a single score.



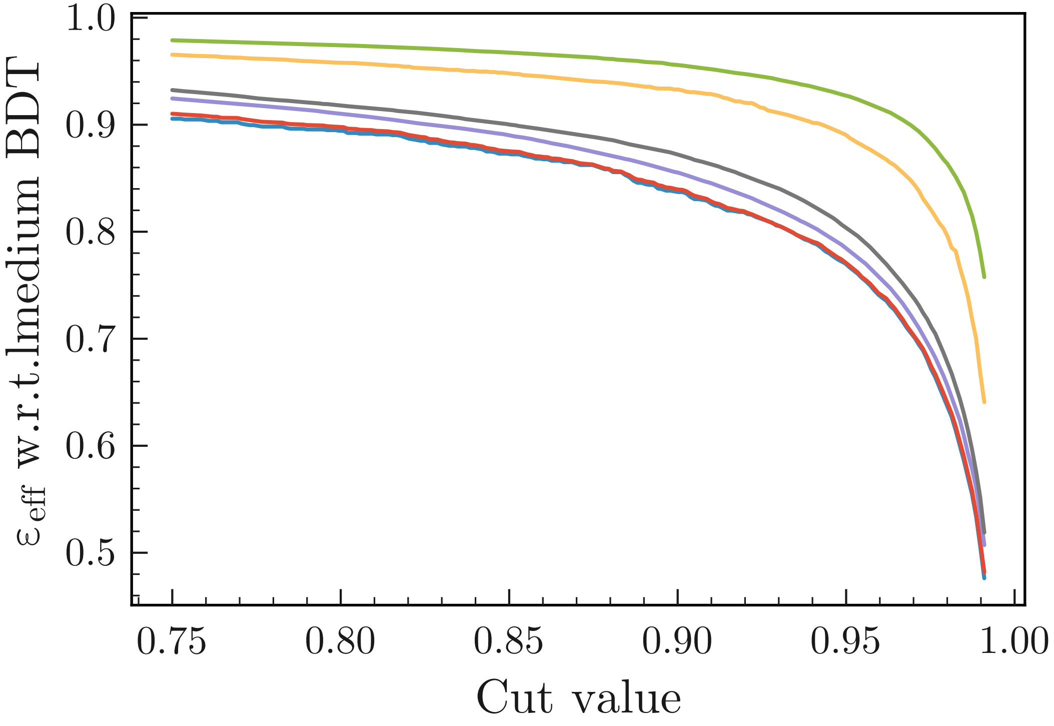
F-measure formula:

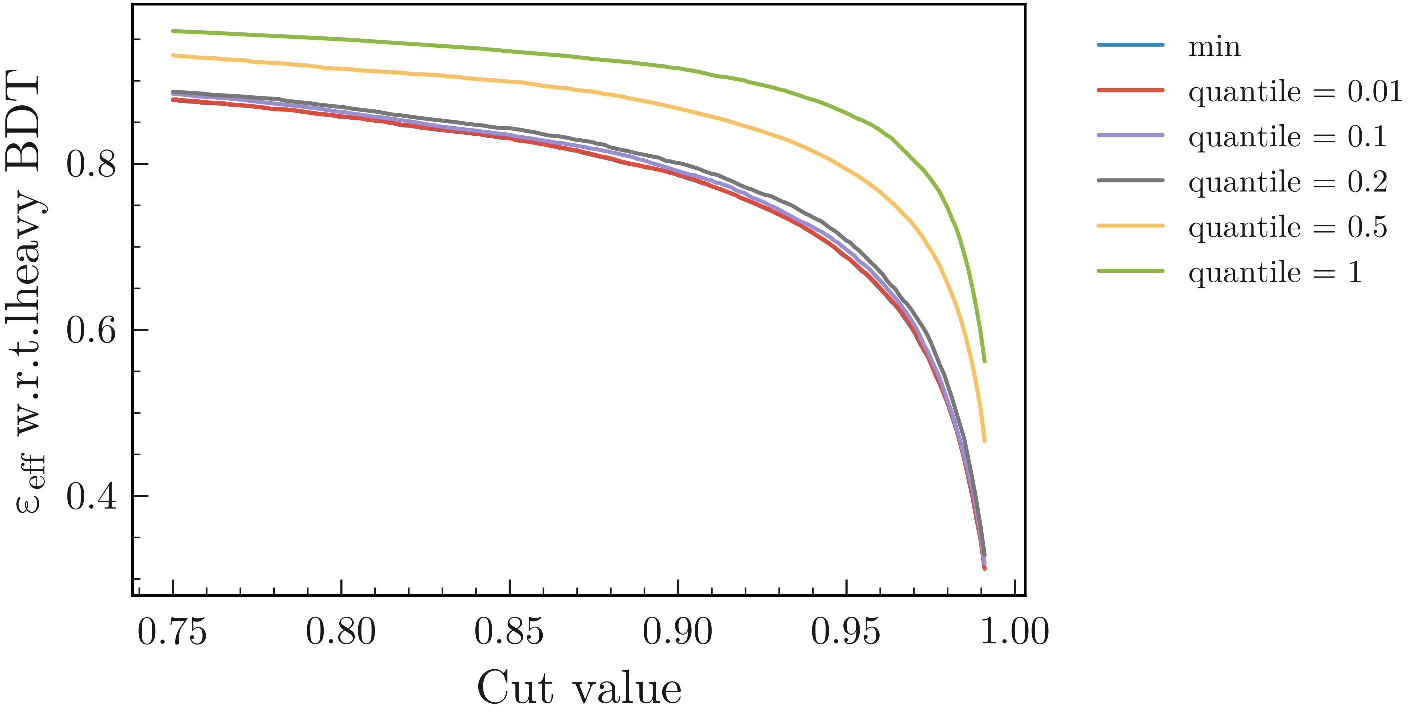
- **F-score = 2 * (precision * recall) / (precision + recall)**

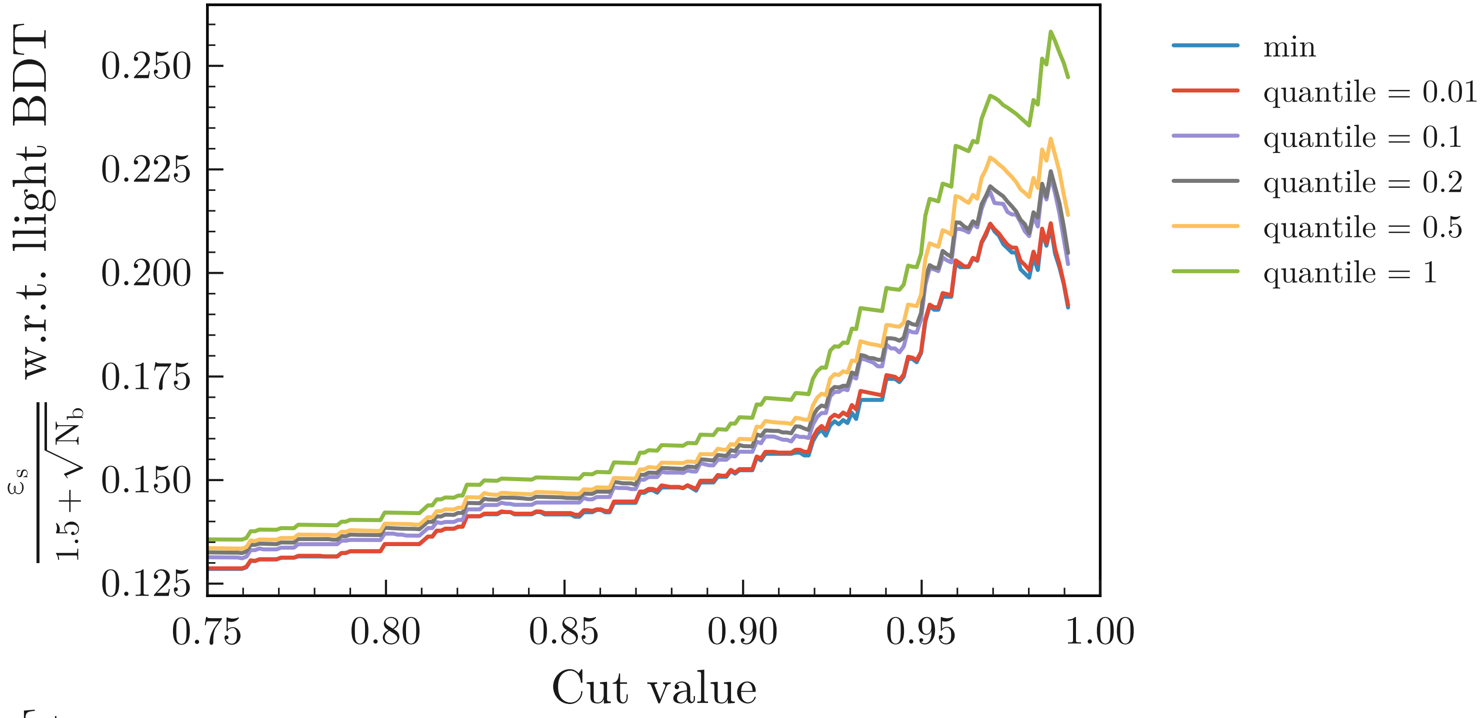
slide 17



- $\varepsilon_{eff} = N_{\text{post cut \& reco}}/N_{\text{post reco}}$
- Could do $\varepsilon_{eff} = N_{\text{post cut \& reco}}/N_{gen}$ but this would now depend on ϵ (kinetic mixing) so hesitant



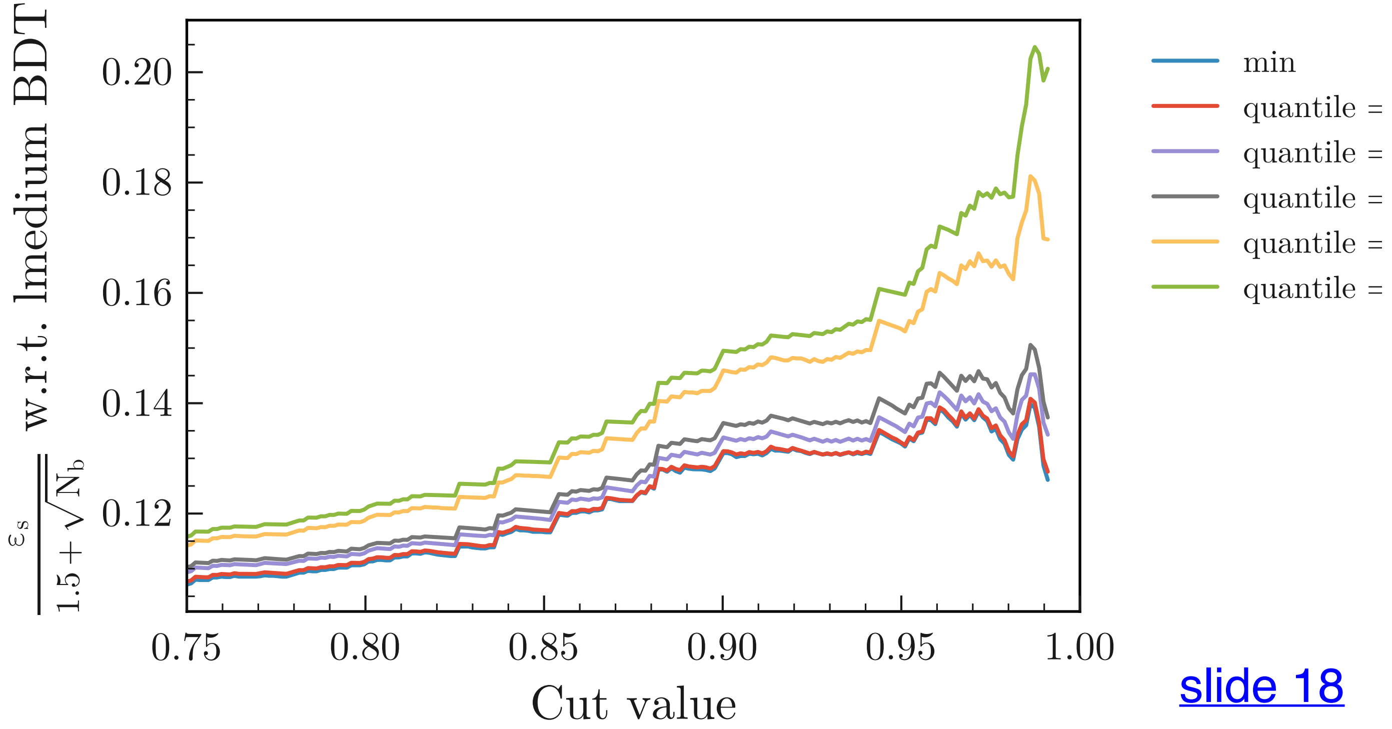


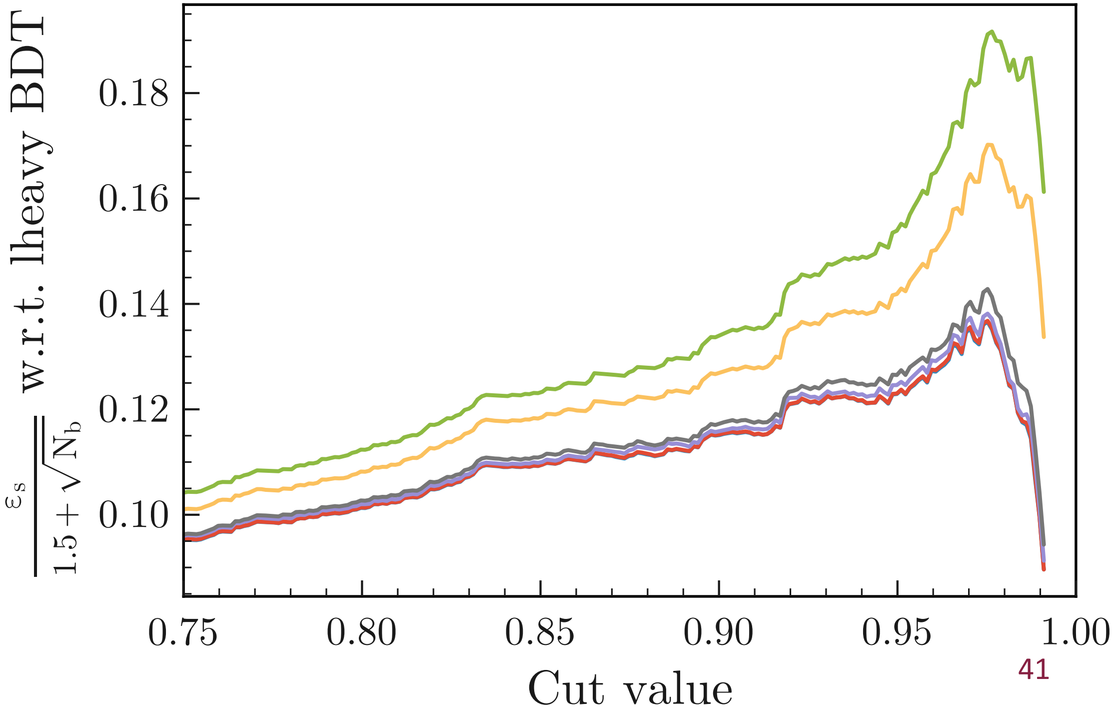



```
all_BDT_cuts = {
    "light": 0.9716231155778894,
    "medium": 0.9546683417085426,
    "heavy": 0.9813115577889447
}
```

```
separated_BDT_cuts = {
    "light": 0.991,
    "medium": 0.9885778894472361,
    "heavy": 0.9897889447236181
}
```

```
lepton_BDT_cuts = {
    "light": 0.9861557788944724,
    "medium": 0.9861557788944724,
    "heavy": 0.9752562814070351
}
```



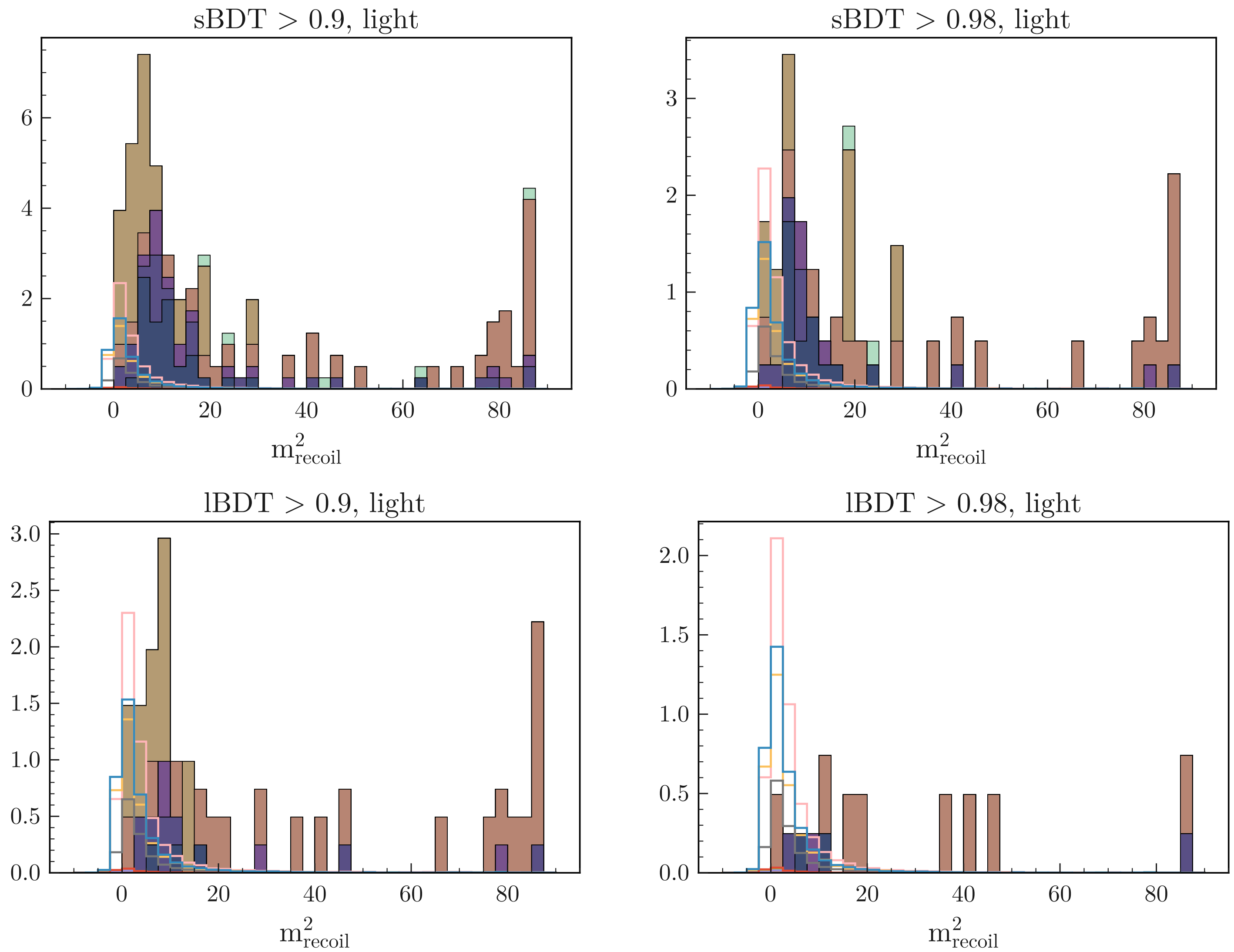


slide 18

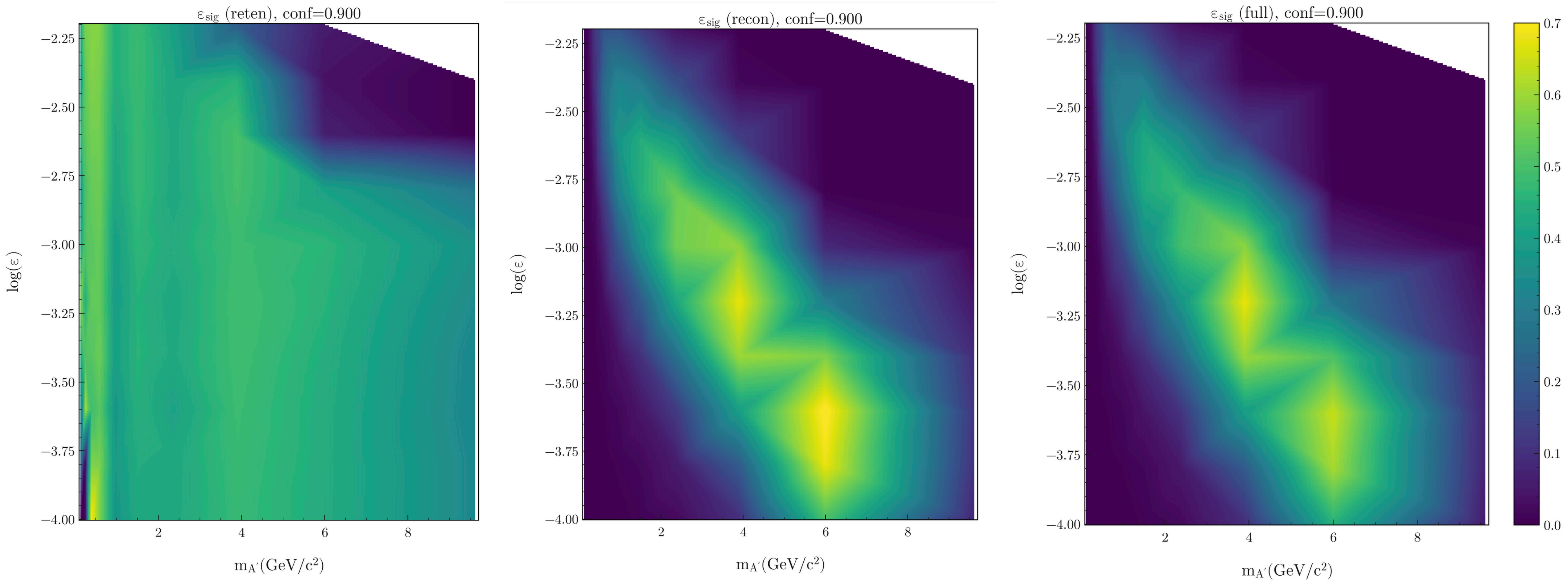
Playing around with BDTs

slide 18

- One subtle thing we've been playing around with is the influence of leptonID is utilized



- Model 1 (top): muon and electron ID are done separately, and choose highest output
- Model 2 (bottom): single leptonID variable
- Model 3 (in-progress): muon and electron ID variables at the same time

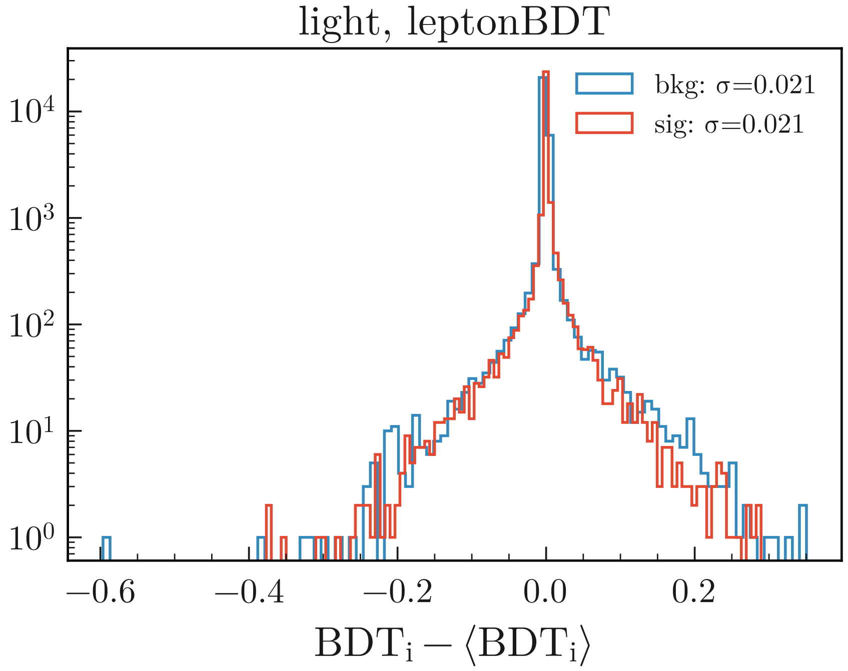


- Still preliminary but these are efficiency tables used for upper limits
  - Retention: # of events post cut & reconstruction / # of events post reconstruction
  - Full: # of events post cut & reconstruction / # of generated events

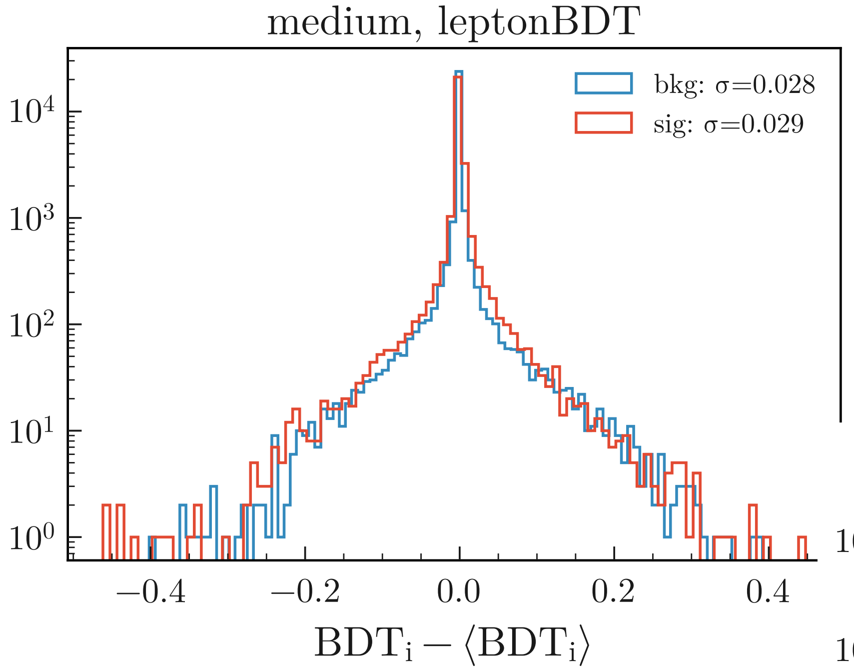
slide 27

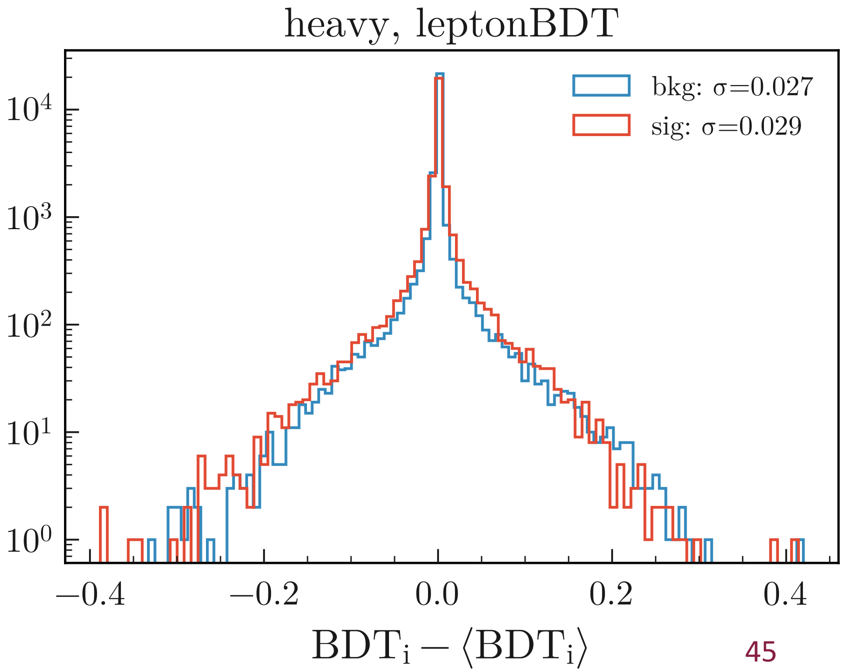
Minor details on BDT usage

- Actually doing what Tia Crane called a “Boosted Decision Forest”
- Train N = 4 (can increase later) BDTs and, using the ideas from random forests, average them (also helps define uncertainty)



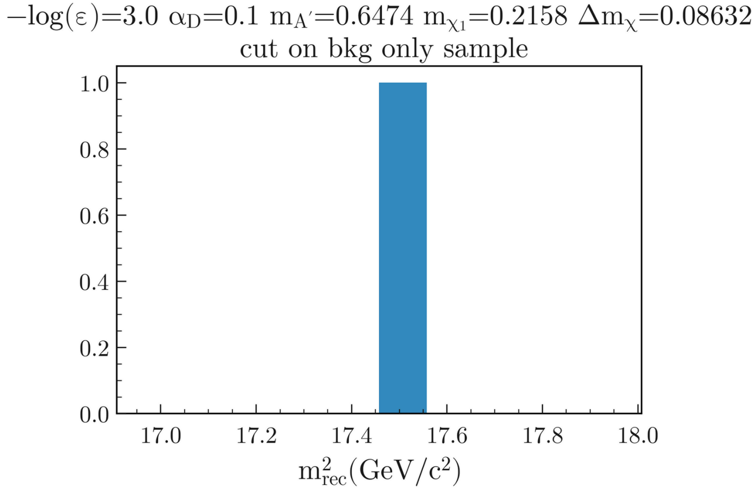
- Usage: Measure for systematics by fluctuating BDT cuts post PFOM cut by ~0.02 - 0.03



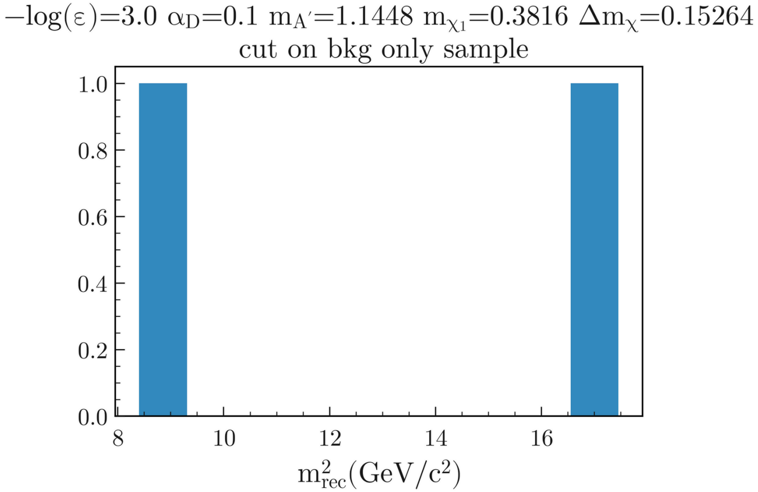


- BDT_i = prediction for an event with the i'th trial/tree
- $\langle BDT_i \rangle$ = average over all the N=4 trials
- (Maybe there are other measures for BDT variance but… This was first thought

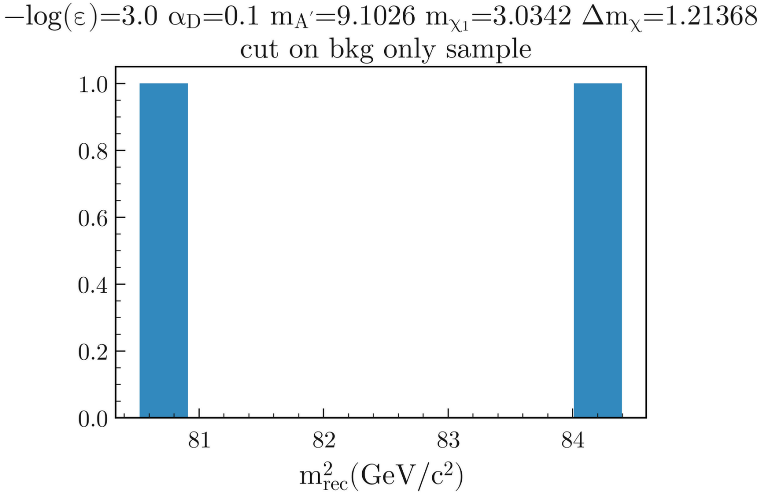
```
0.2158 [17.45806506] []
```



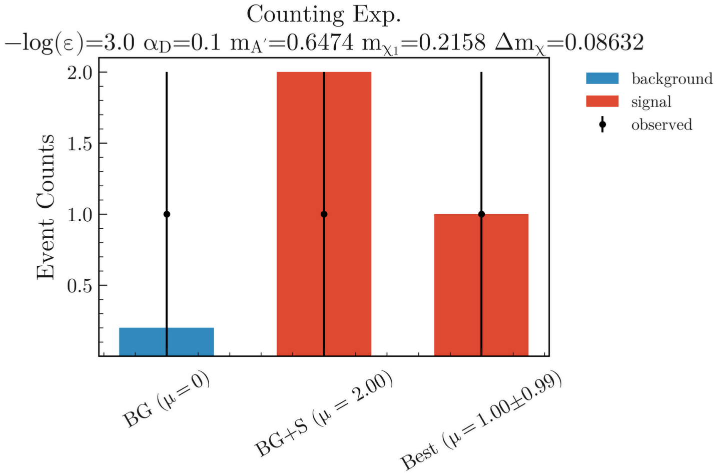

```
0.3816 [17.45806506  8.40740307] [8.8219062  8.40740307 8.40740307 8.40740307]
```



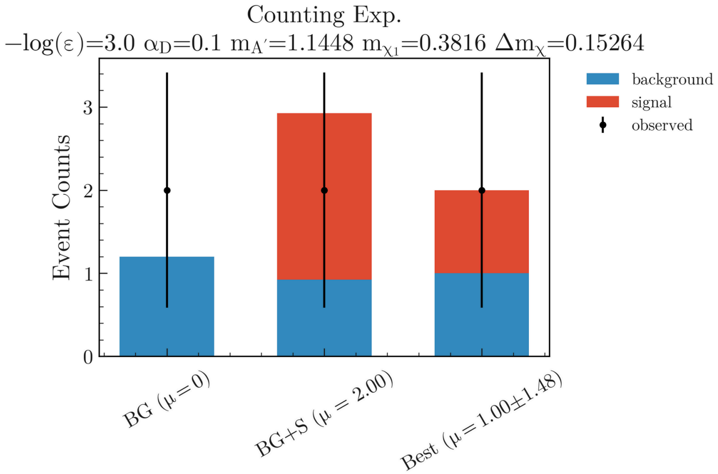

```
3.0342 [84.39782496 80.5251381 ] [76.95944912 84.39782496 84.39782496 80.5251381  80.5251381  80.5251381
 80.5251381  79.53473904 79.07492489 81.44058545]
```



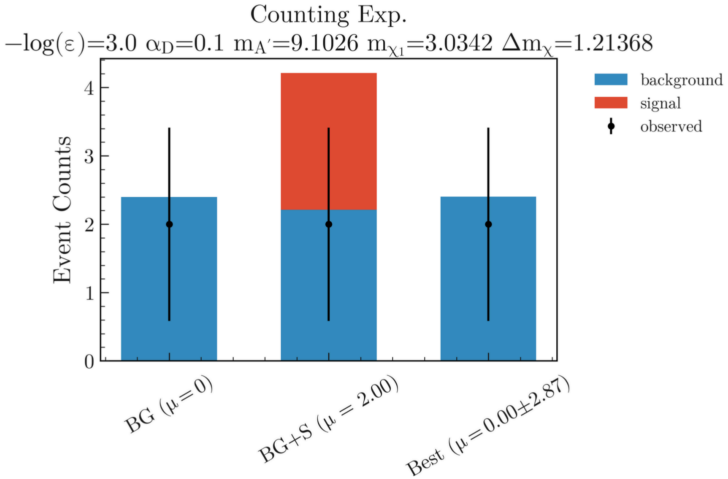

```
Initial signal guess = 2
  bkg_norm = 0.000 +/- 0.125
  mu = 1.001 +/- 0.994
```




```
Initial signal guess = 2
  bkg_norm = 1.000 +/- 0.499
  mu = 1.000 +/- 1.477
```




```
Initial signal guess = 2
  bkg_norm = 2.402 +/- 0.692
  mu = 0.000 +/- 2.868
```



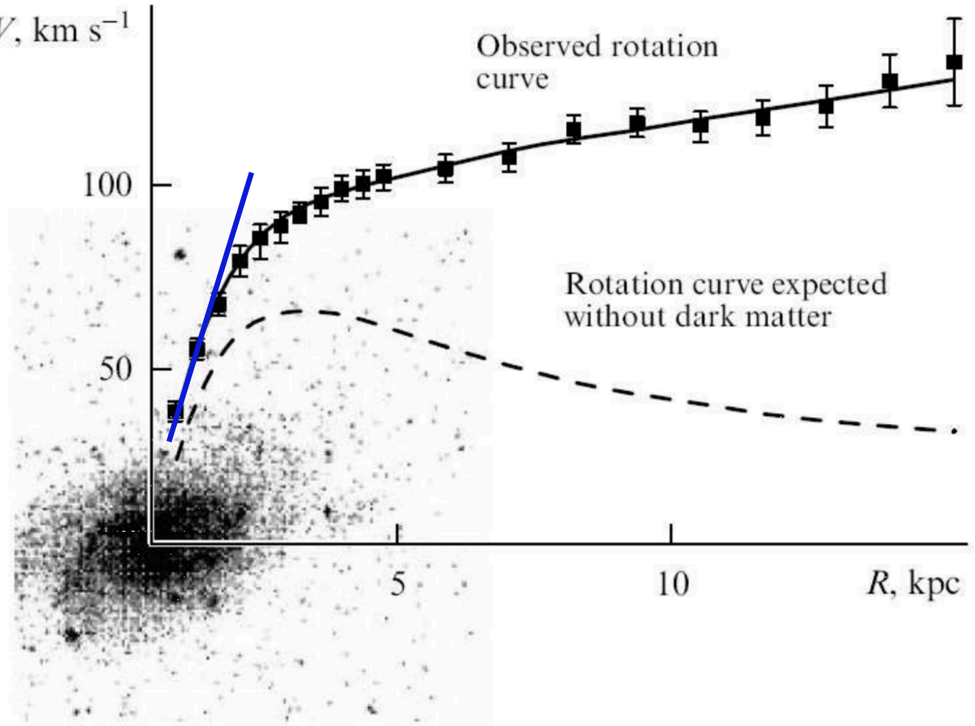
slide 25

DM Evidence

Evidence for Dark Matter

- From Astronomy/Cosmology [1]
  - Galaxy rotation curves [2]
  - Bullet Cluster
  - CMB Anisotropies
  - Lyman α-Forests
  - Nucleosynthesis

$$v^2 = \frac{GM(r)}{r}$$



M33 Galaxy's rotation curves [3]

constant density

$$v^2 = \frac{4\pi G\rho_0}{3} r^2$$

outside central density

w/o DM

$$v^2 = \frac{GM_{tot}}{r}$$

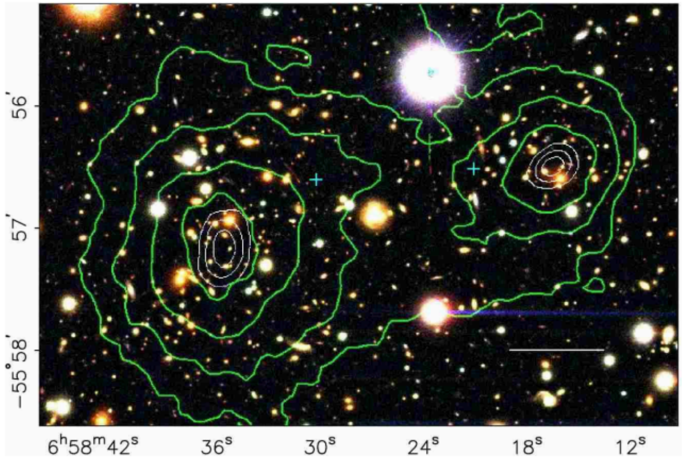
[1]: D. Gruen. "Cosmology Overview." SSI 2020:The Almost Invisibles: Exploring the Weakly Coupled Universe. (2020) (Slide 25)
[2]: V.C Rubin, W.K. Ford Jr. "Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions" (1970)
[3]: A.V. Zasov et al. "Dark Matter in galaxies." (2017) arxiv:1710.10630
[4]: D. Clowe et al. "A Direct Empirical Proof of the Existence of Dark matter." (2004) arxiv:0608407

Evidence for Dark Matter

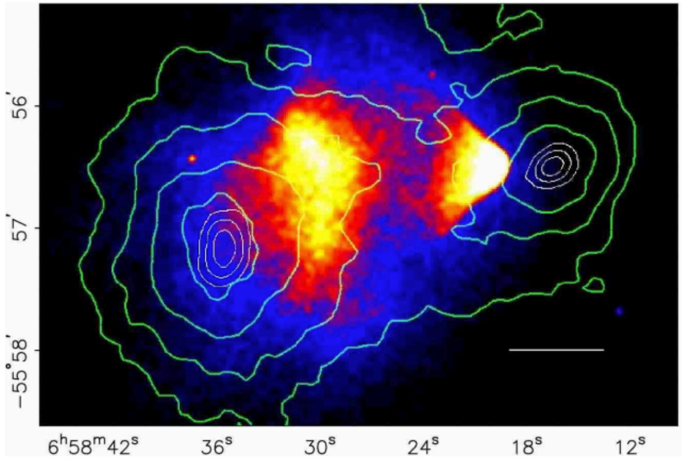
- From Astronomy/Cosmology [1]
  - Galaxy rotation curves
  - Bullet Cluster [2]
  - CMB Anisotropies
  - Lyman α-Forests
  - Nucleosynthesis

PC: [2]

Gravitational Lensing



Plasma Clouds

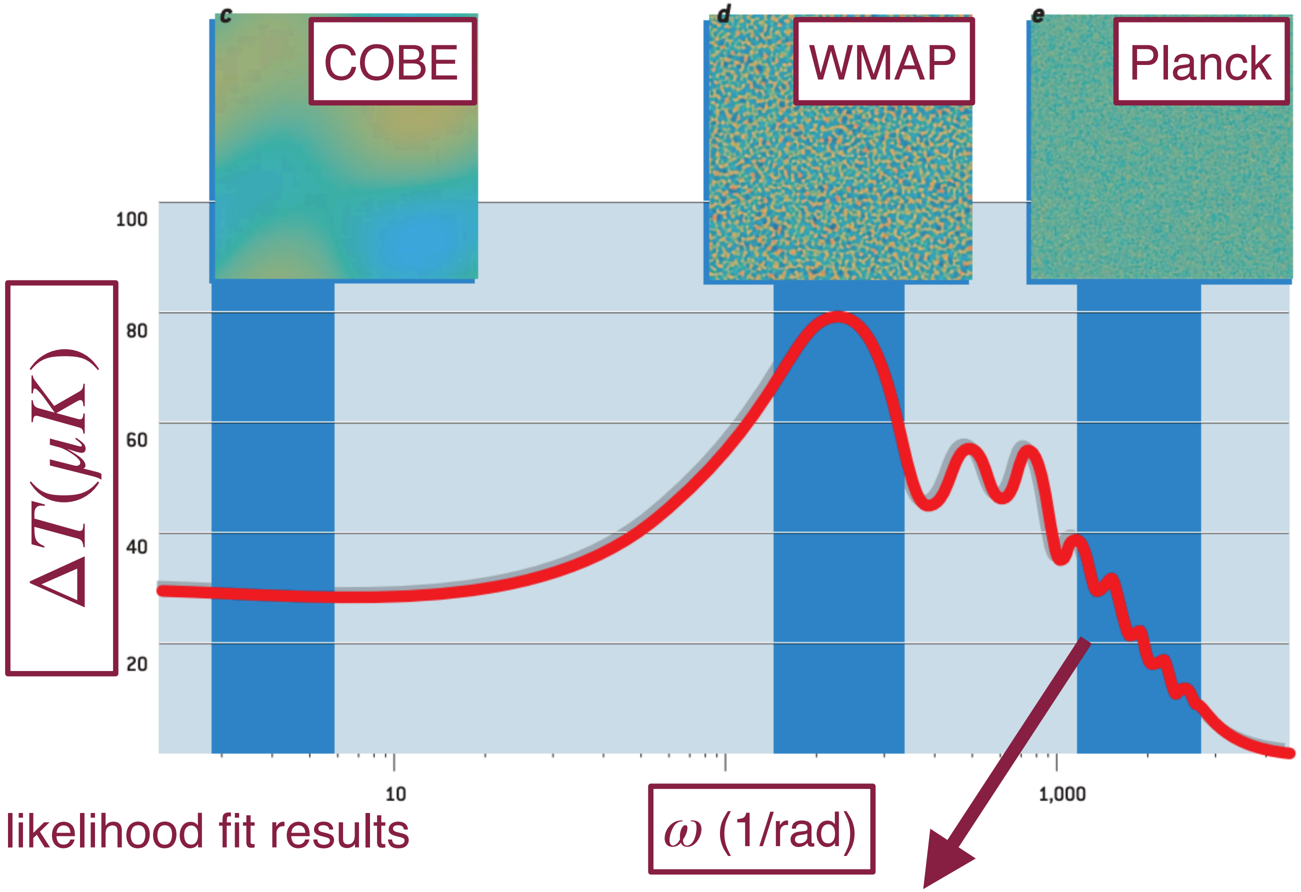


[1]: D. Gruen. “Cosmology Overview.” SSI 2020:The Almost Invisibles: Exploring the Weakly Coupled Universe. (2020) (Slide 25)
[2]: D. Clowe et al. “A Direct Empirical Proof of the Existence of Dark matter.” (2004) arxiv:0608407

Evidence for Dark Matter

PC: [3]

- From Astronomy/Cosmology [1]
  - Galaxy rotation curves
  - Bullet Cluster
  - CMB Anisotropies [4,5,6]
  - Lyman α-Forests
  - Nucleosynthesis



baryon $\Omega_b h^2 = 0.02237 \pm 0.00015$

cold DM $\Omega_c h^2 = 0.1200 \pm 0.0012$

total matter $\Omega_m h^2 = 0.1430 \pm 0.0011$

[1]: D. Gruen. “Cosmology Overview.” SSI 2020:The Almost Invisibles: Exploring the Weakly Coupled Universe. (2020) (Slide 25)

[2]: https://www.nasa.gov/mission_pages/planck/multimedia/pia16874.html

[3]: W. Hu and M. White. “The Cosmic Symphony.” http://background.uchicago.edu/~whu/Papers/HuWhi04.pdf

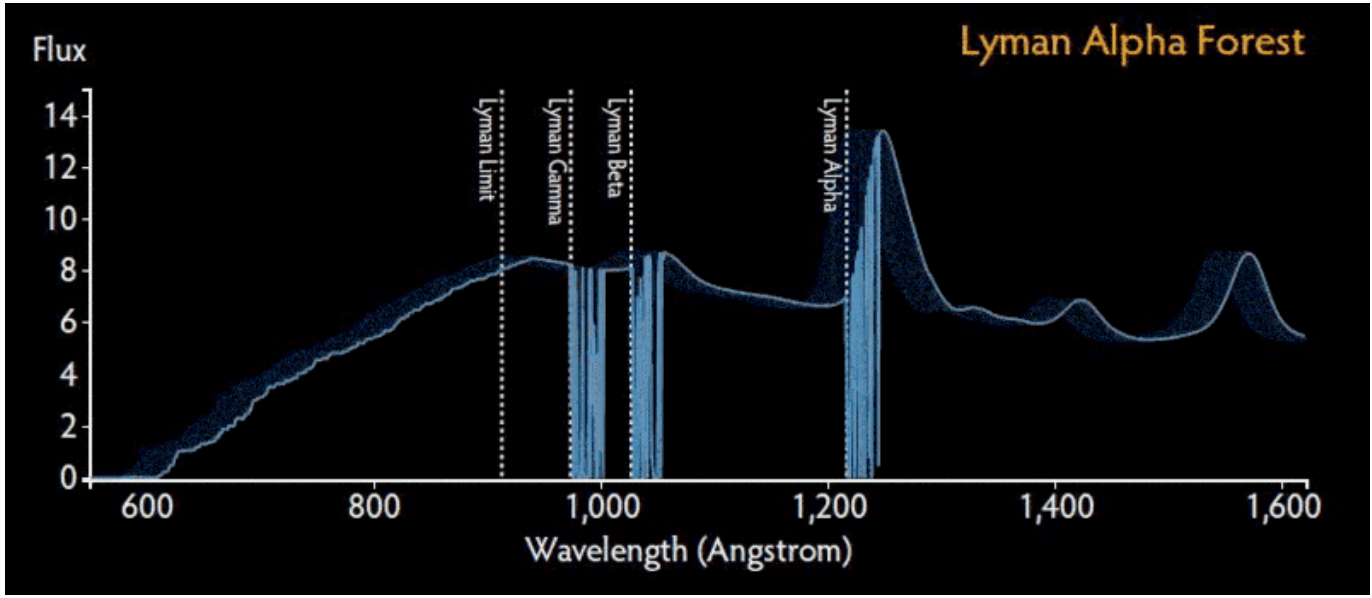
[4]: S. Das et al. PRL 107, 021301 (2011).

[5]: E. Komatsu et al. ApJS **192** 18 (2011)

[6]: Planck Collaboration. arxiv:1807.06209

Evidence for Dark Matter

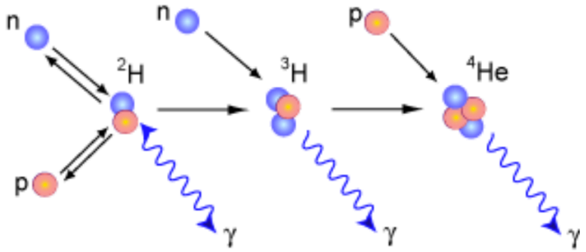
- From Astronomy/Cosmology [1]
  - Galaxy rotation curves
  - Bullet Cluster
  - CMB Anisotropies
  - Lyman α-Forests [2]
  - Nucleosynthesis

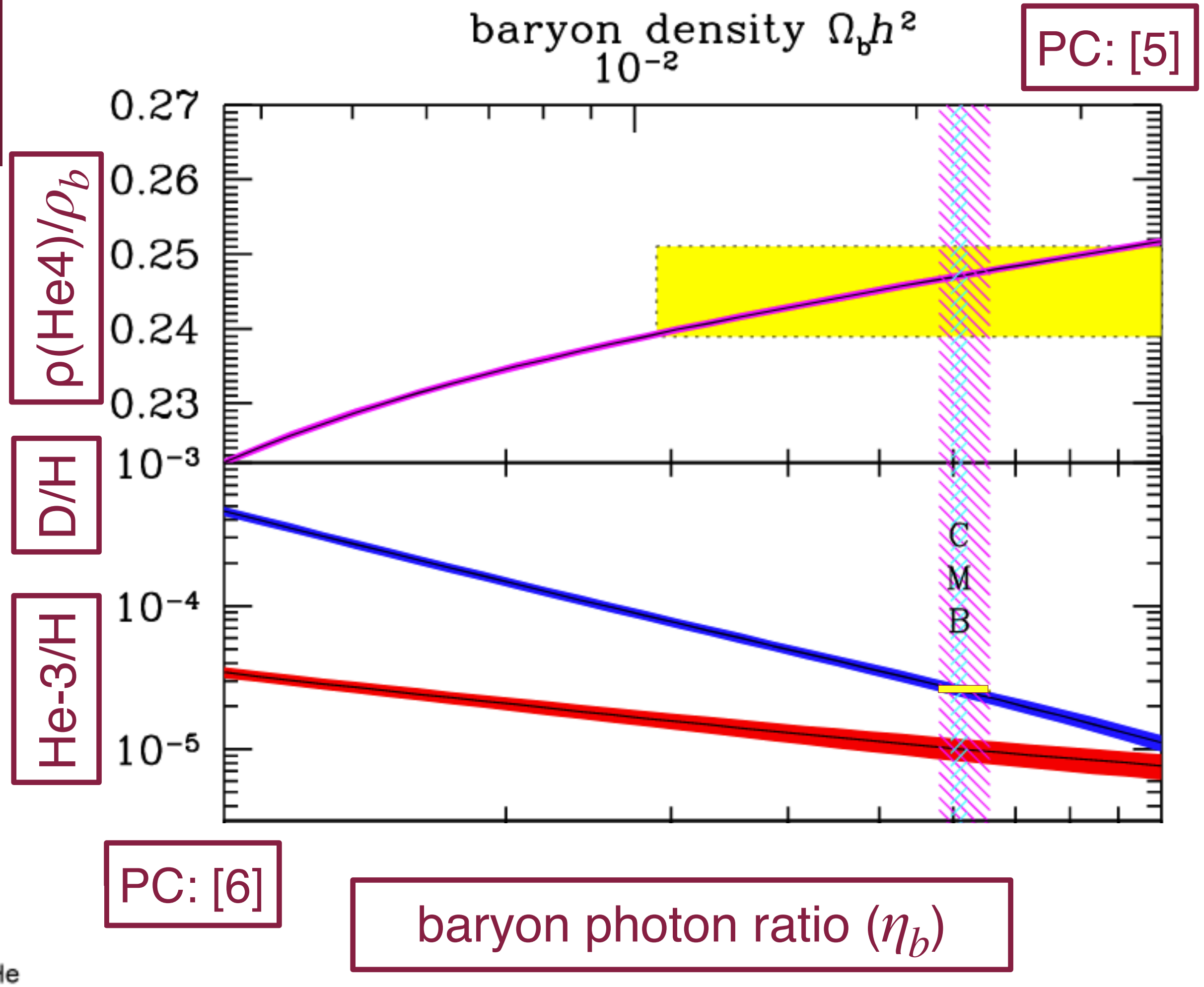


[1]: D. Gruen. "Cosmology Overview." SSI 2020:The Almost Invisibles: Exploring the Weakly Coupled Universe. (2020) (Slide 25)
[2]: https://phys.org/news/2024-07-dark-forest-unknown-particle.html

Evidence for DM

- From Astronomy/Cosmology [1]
  - Galaxy rotation curves
  - Bullet Cluster
  - CMB Anisotropies
  - Lyman α-Forests
  - Nucleosynthesis





[1]: D. Gruen. "Cosmology Overview." SSI 2020:The Almost Invisibles: Exploring the Weakly Coupled Universe. (2020) (Slide 25)
[2]: A. Dar et al. "Dark Matter and the Big Bang Nucleosynthesis." arxiv:9405010.
[3]: K. Jedamzik and M. Pospelov. "Big Bang nucleosynthesis and particle dark matter." arXiv:0906.2087
[4]: G. Steigman. "Primordial Nucleosynthesis: The Predicted and Observed Abundances and Their Consequences." arxiv:1008.4765.
[5]: B.D. Fields et al. "Big Bang Nucleosynthesis." arxiv:1412.1408. (updated 2017).
[6]: https://www.atnf.csiro.au/outreach/education/senior/cosmicengine/bigbang.html