

62ND INTERNATIONAL WINTER MEETING  
ON NUCLEAR PHYSICS

# HEXAQUARKS AT BELLE II

FELIX M. KEIL

---



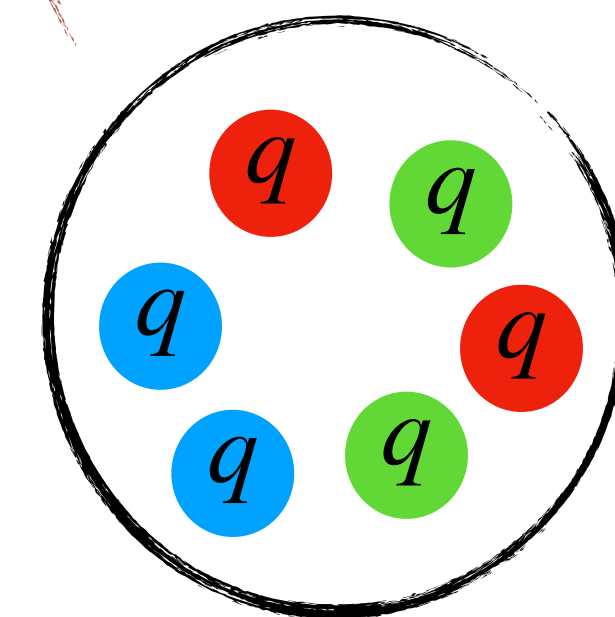
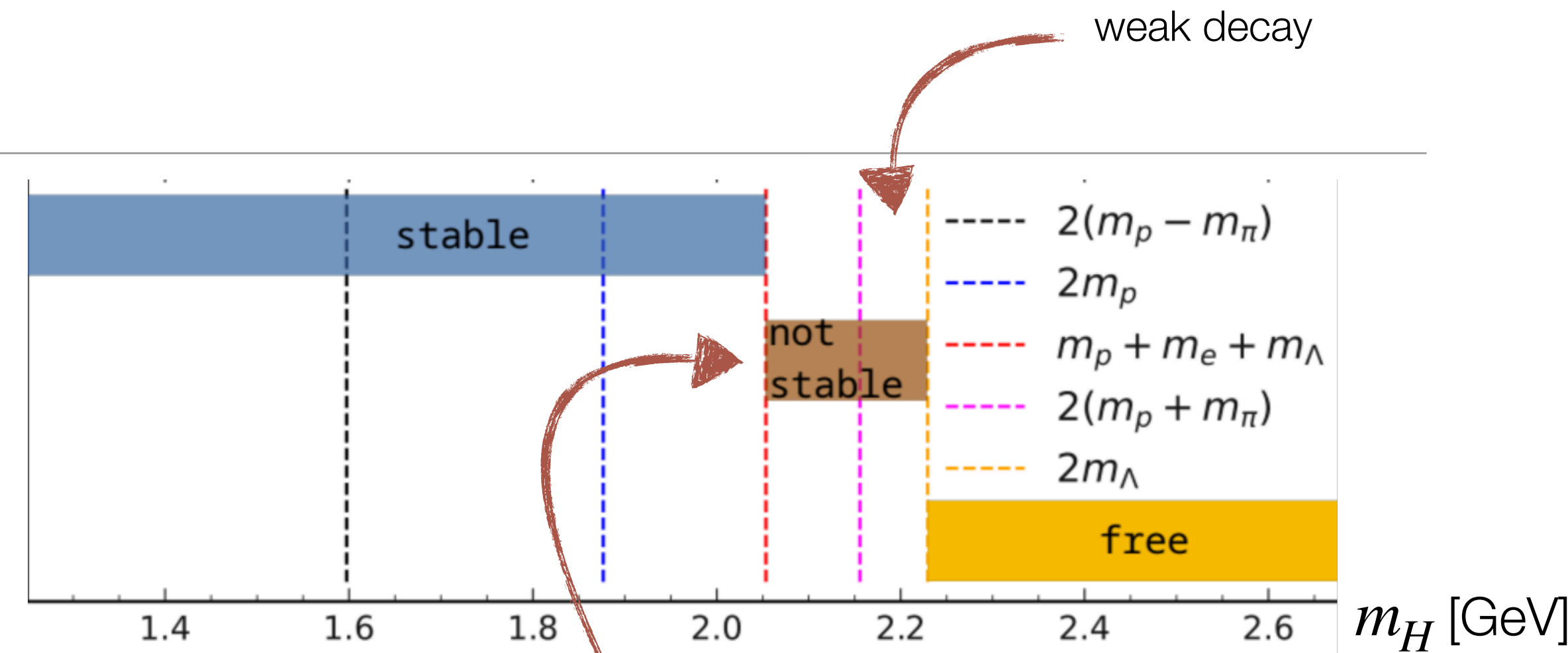
Bundesministerium  
für Forschung, Technologie  
und Raumfahrt



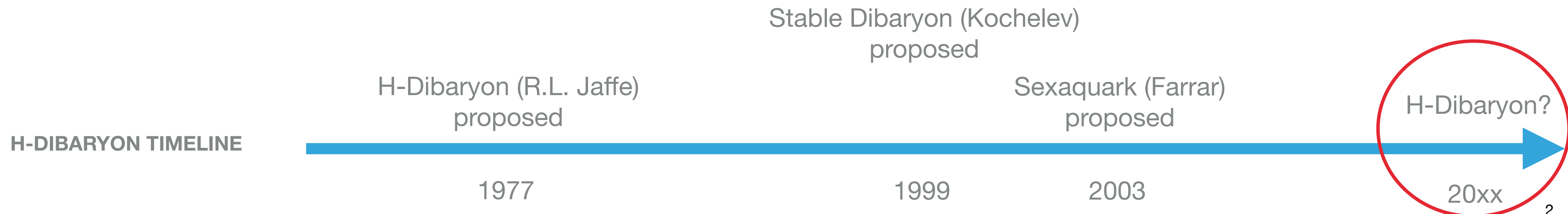
JOHANNES GUTENBERG  
UNIVERSITÄT MAINZ

# What?

- Doubly strange six-quark state
- Spinless and flavor-singlet
- udsuds content ( like  $2\Lambda^0$  )
- LQCD & ALICE & NAGARA  $\rightarrow$  Binding Energy  $\lesssim 7$  MeV



Hexaquark/H-Dibaryon?



# Why?

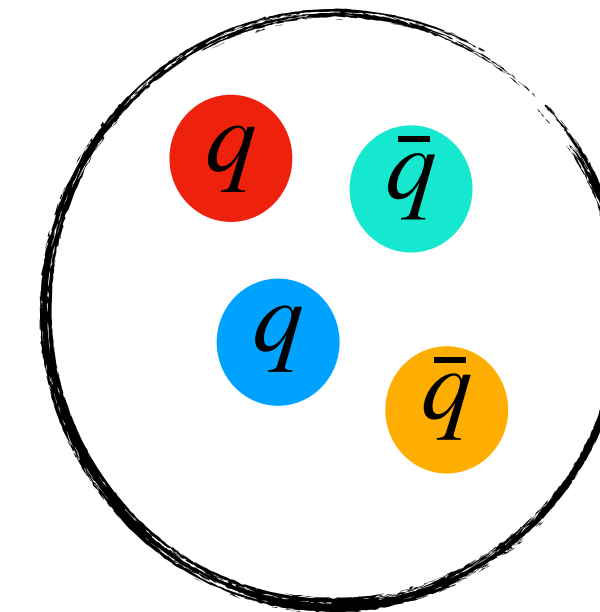
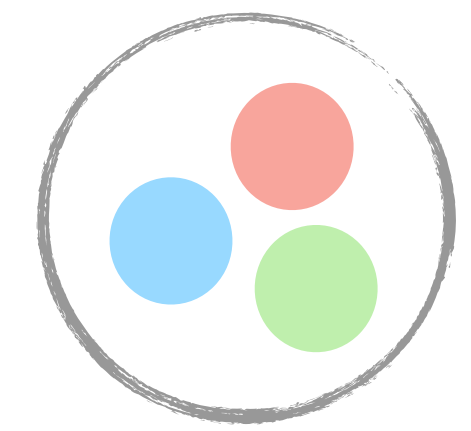
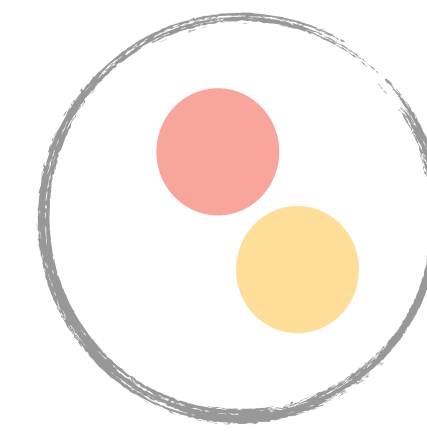
- Discoveries:

- X(3872) at Belle,
- Tetraquarks at LHCb,
- Pentaquarks at LHCb,
- Why not **six-quark** states?

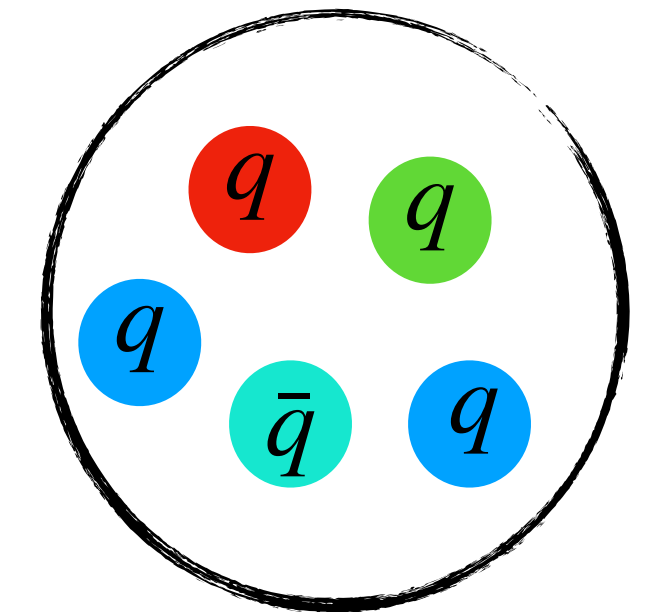
- Influence on **neutron star** EoS

- Understand  $\Lambda - \Lambda$  interaction better

- New particle = Cool 😎

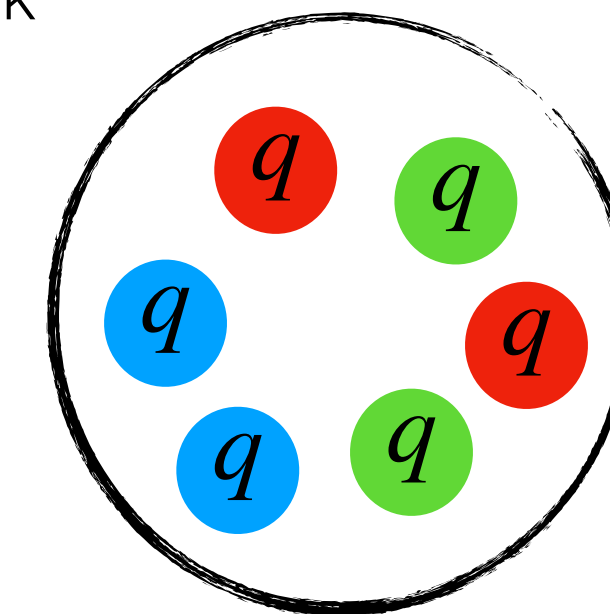


Tetraquark



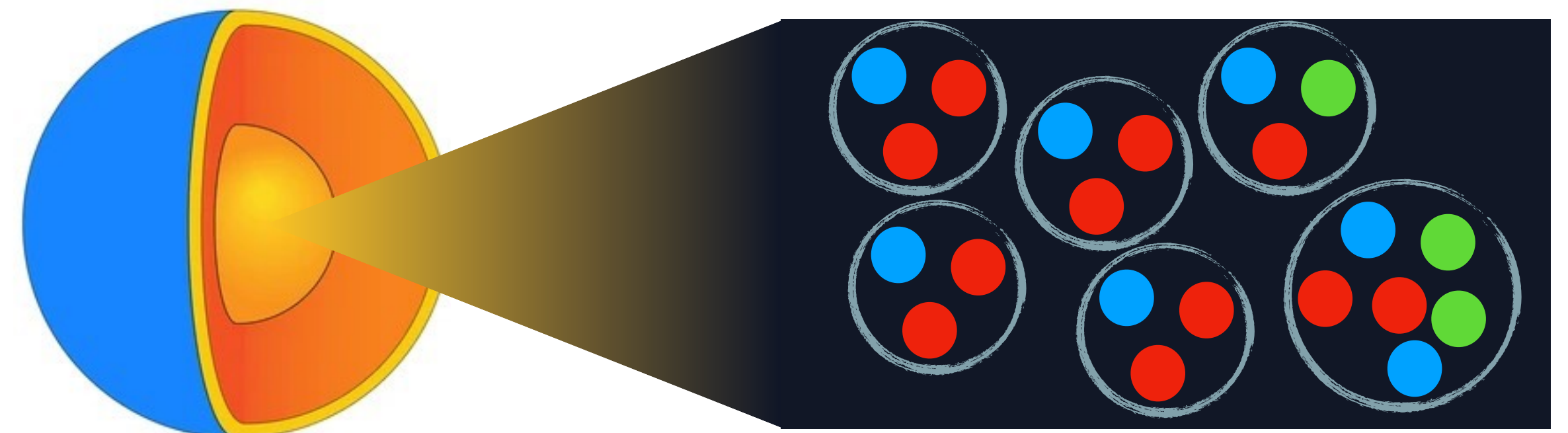
Pentaquark

QCD: No limit on number of quarks!



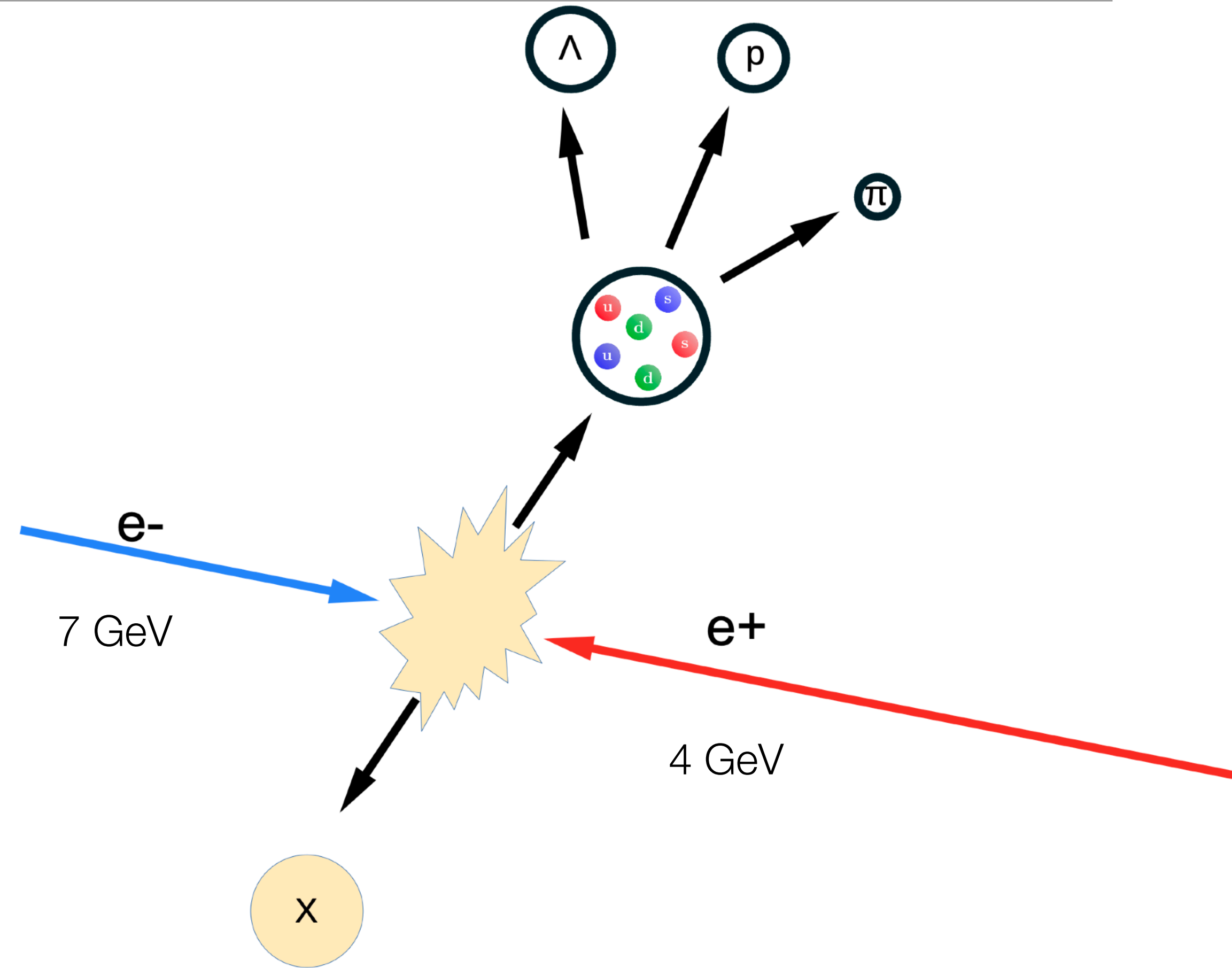
Hexaquark/H-Dibaryon?

Neutron star



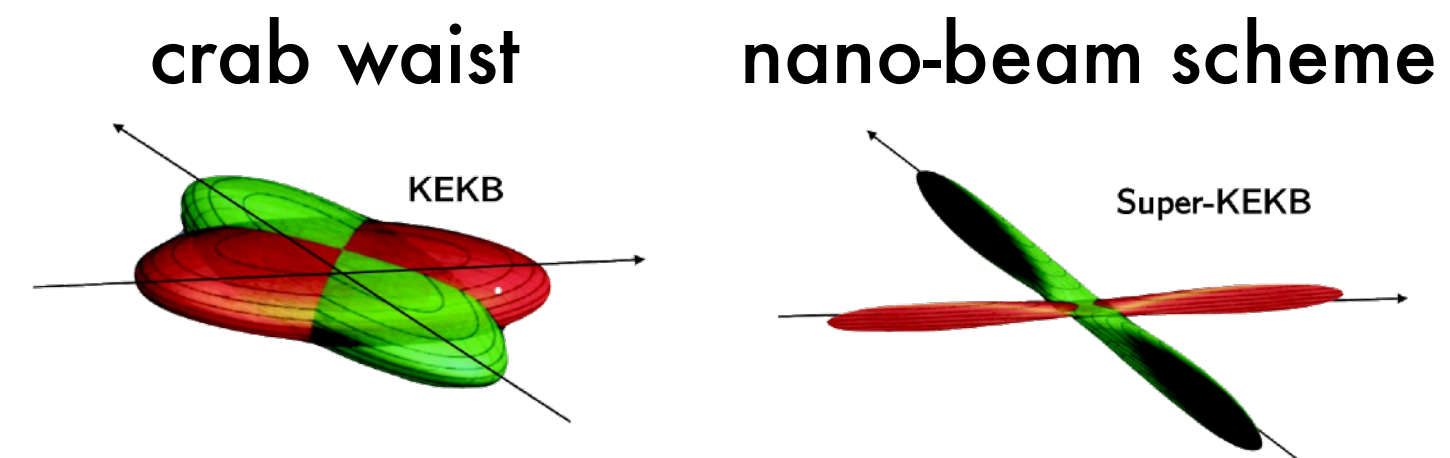
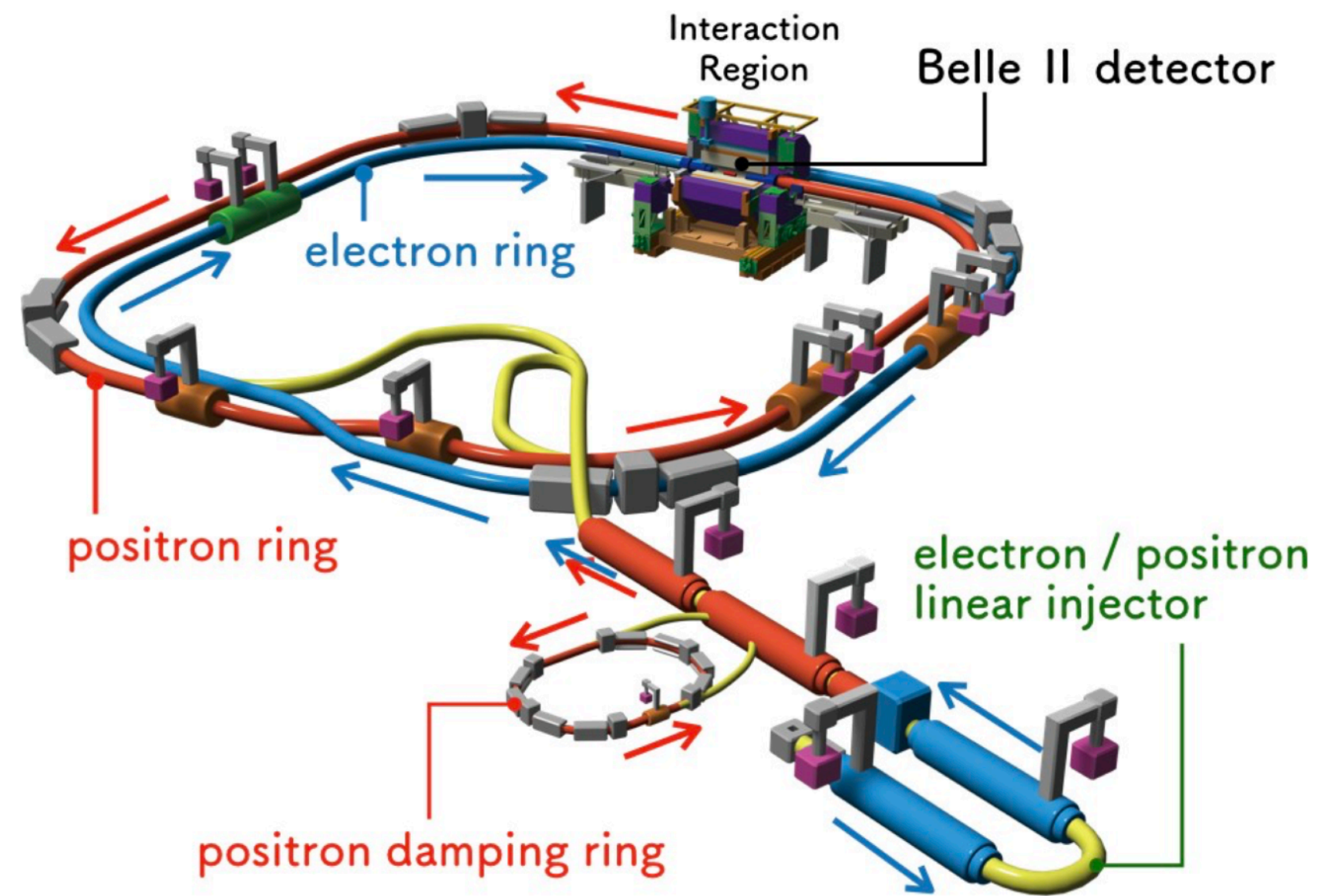
# How? - High Luminosity Frontier

- $H \rightarrow \Lambda p \pi^-$
- Requirement:
  - High density of quarks
  - Good reconstruction of  $\Lambda$  and tracking of  $p \pi^-$
  - Lots of data  $\Leftrightarrow$  high luminosity
  - Optimal having  $\Upsilon(1,2,3S)$



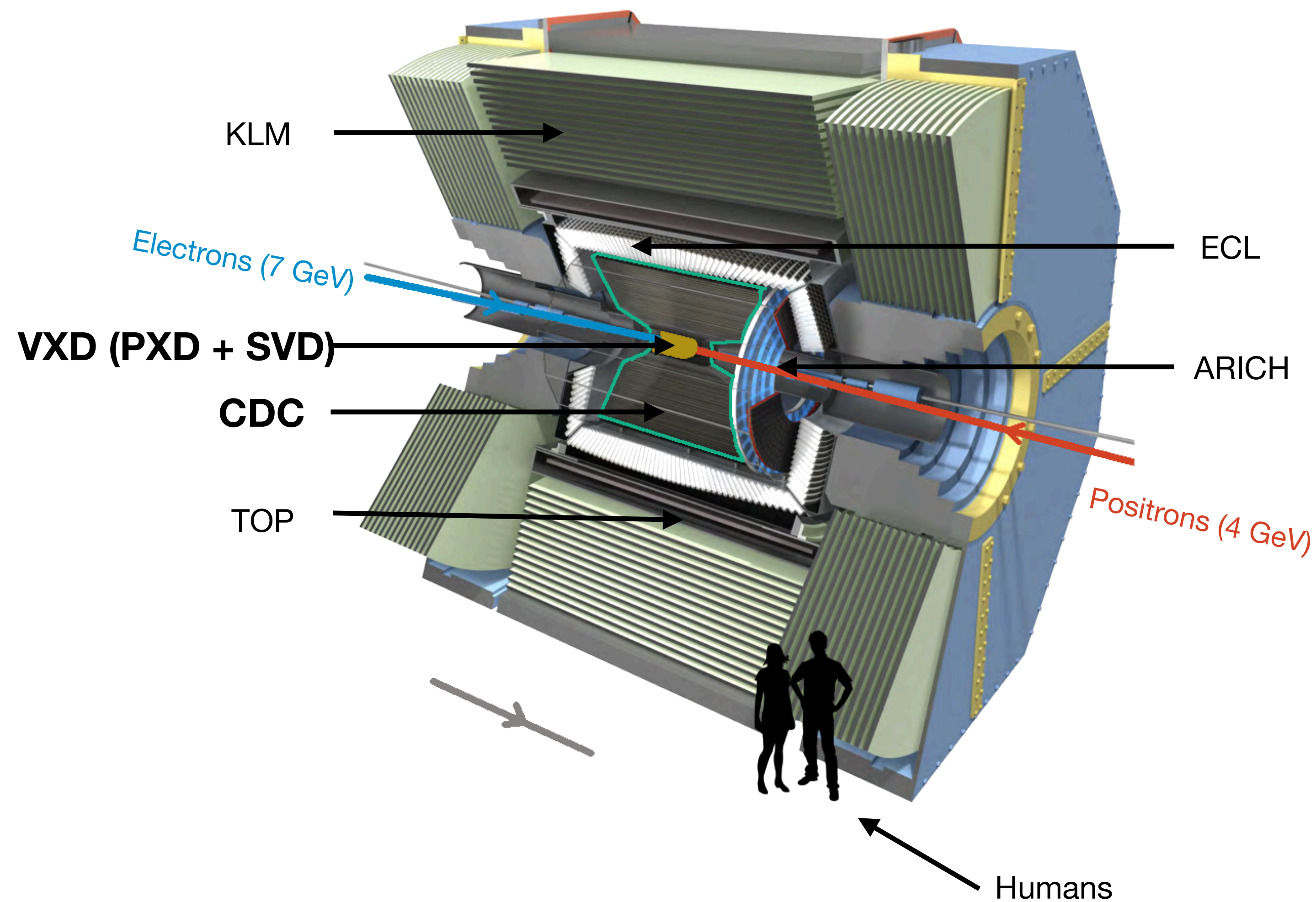
$\Rightarrow$  B-Factories (eg. **SuperKEKB + Belle II**) are optimal for such decay channels

# Briefly: SuperKEKB & Belle II



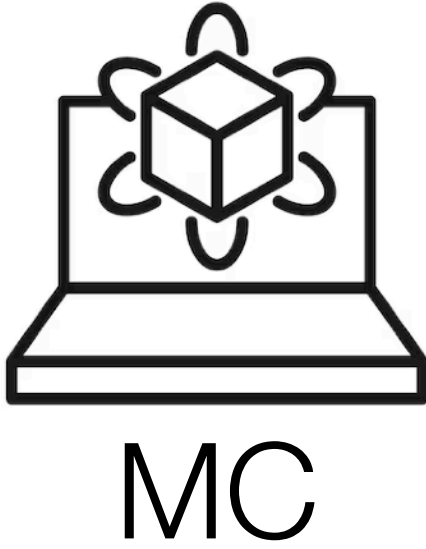
x30 instant. luminosity  
 $\sim 5.1 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$

**SuperKEKB**



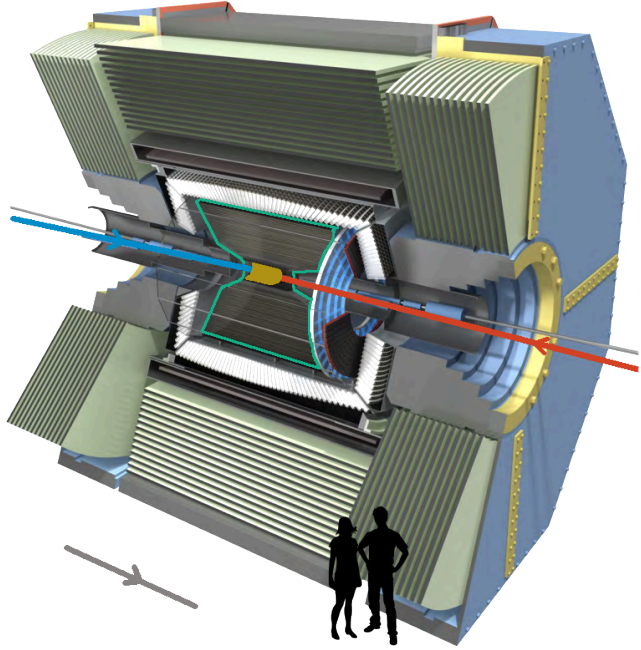
**Belle II**

# How? - Analysis Outline

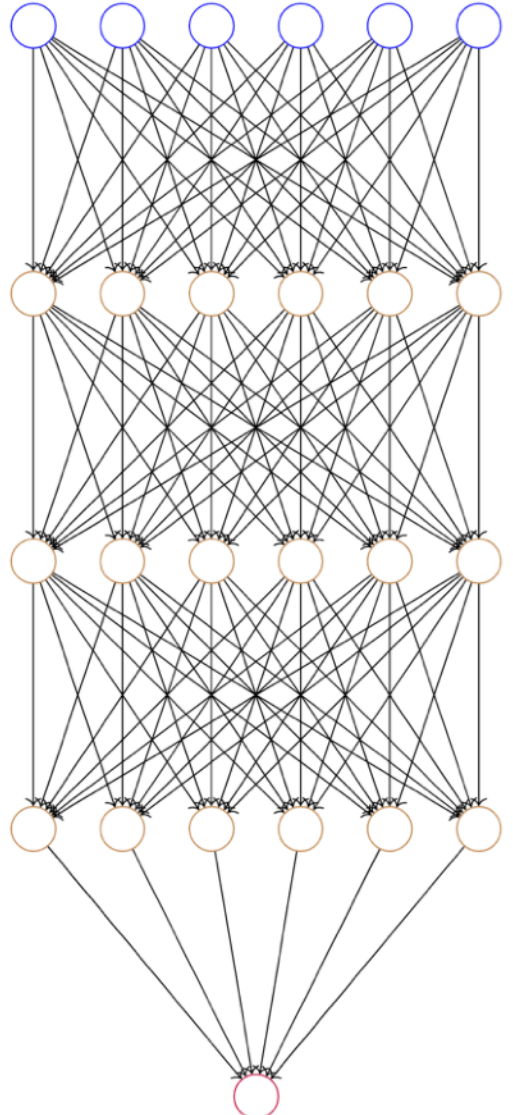


Reconstruction

Reconstruction

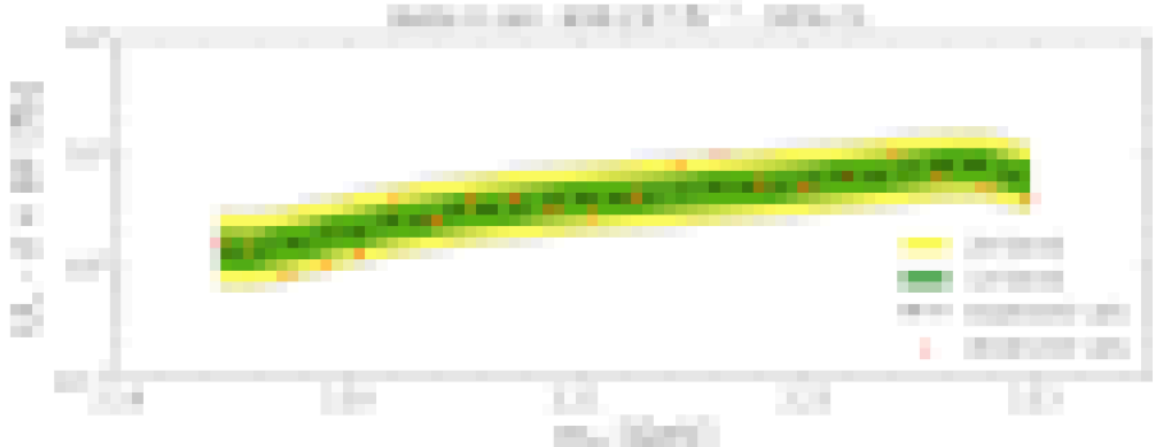
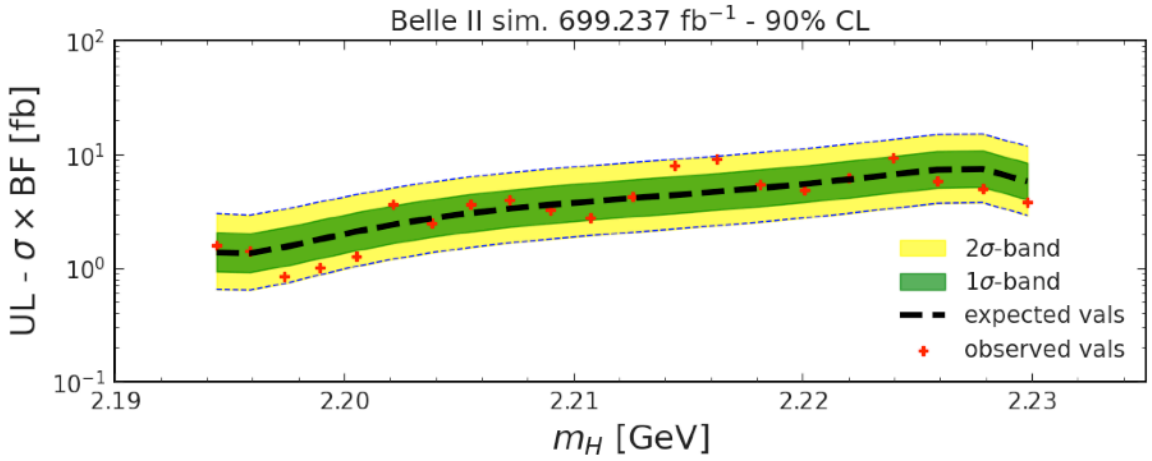


Neural Network

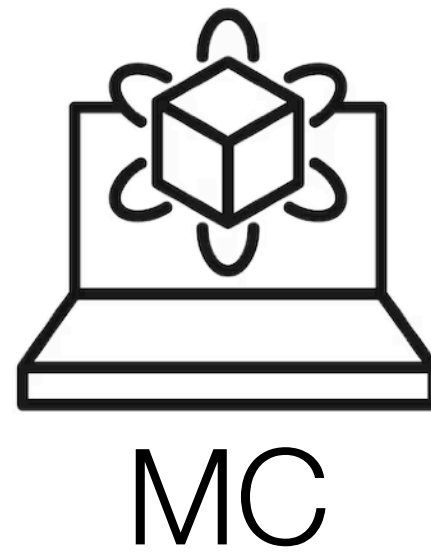


Fit + Inference

Fit + Inference

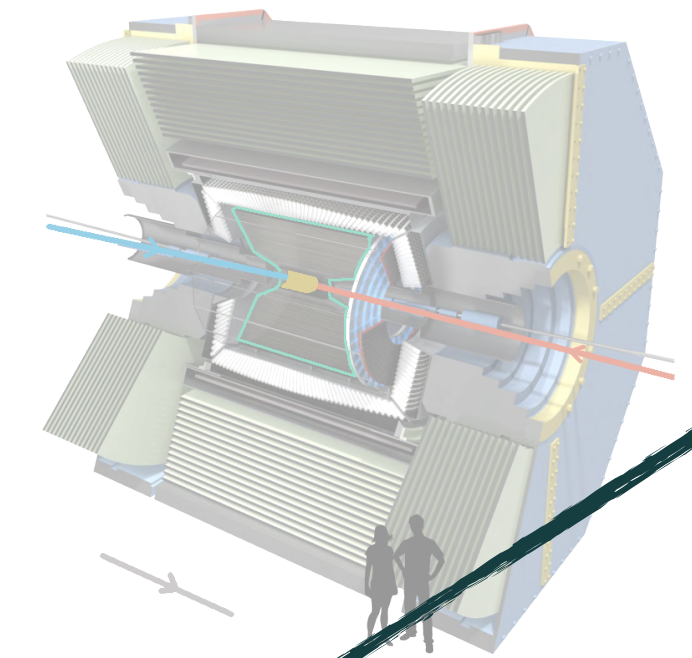


# How? - Analysis Outline

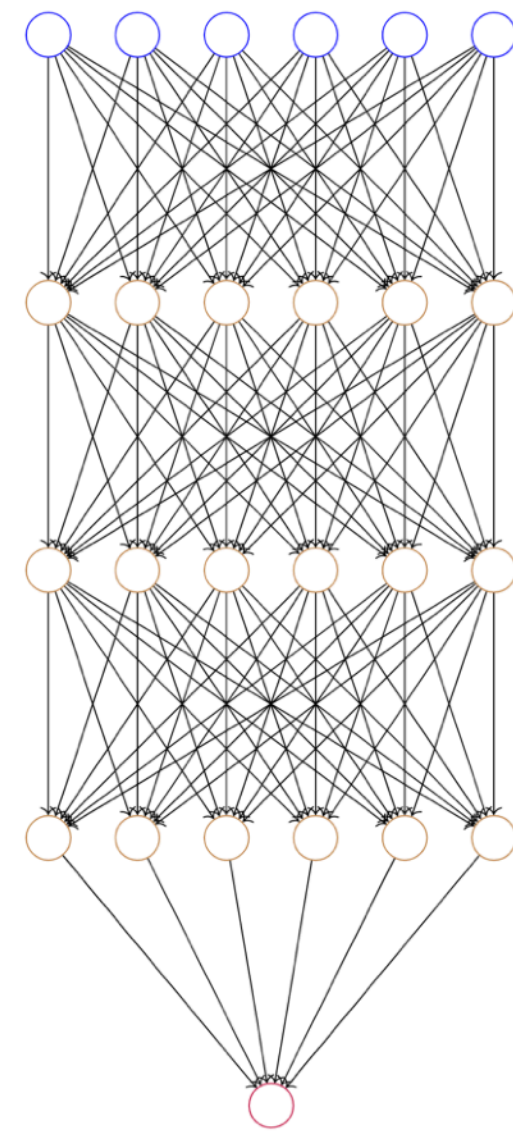


Reconstruction

Reconstruction



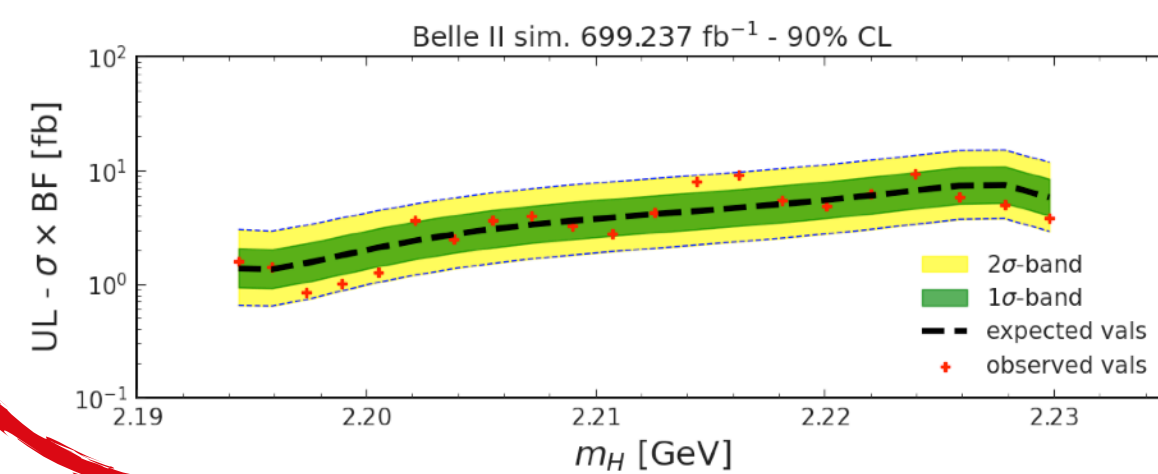
Neural Network



Blinded

Fit + Inference

Fit + Inference

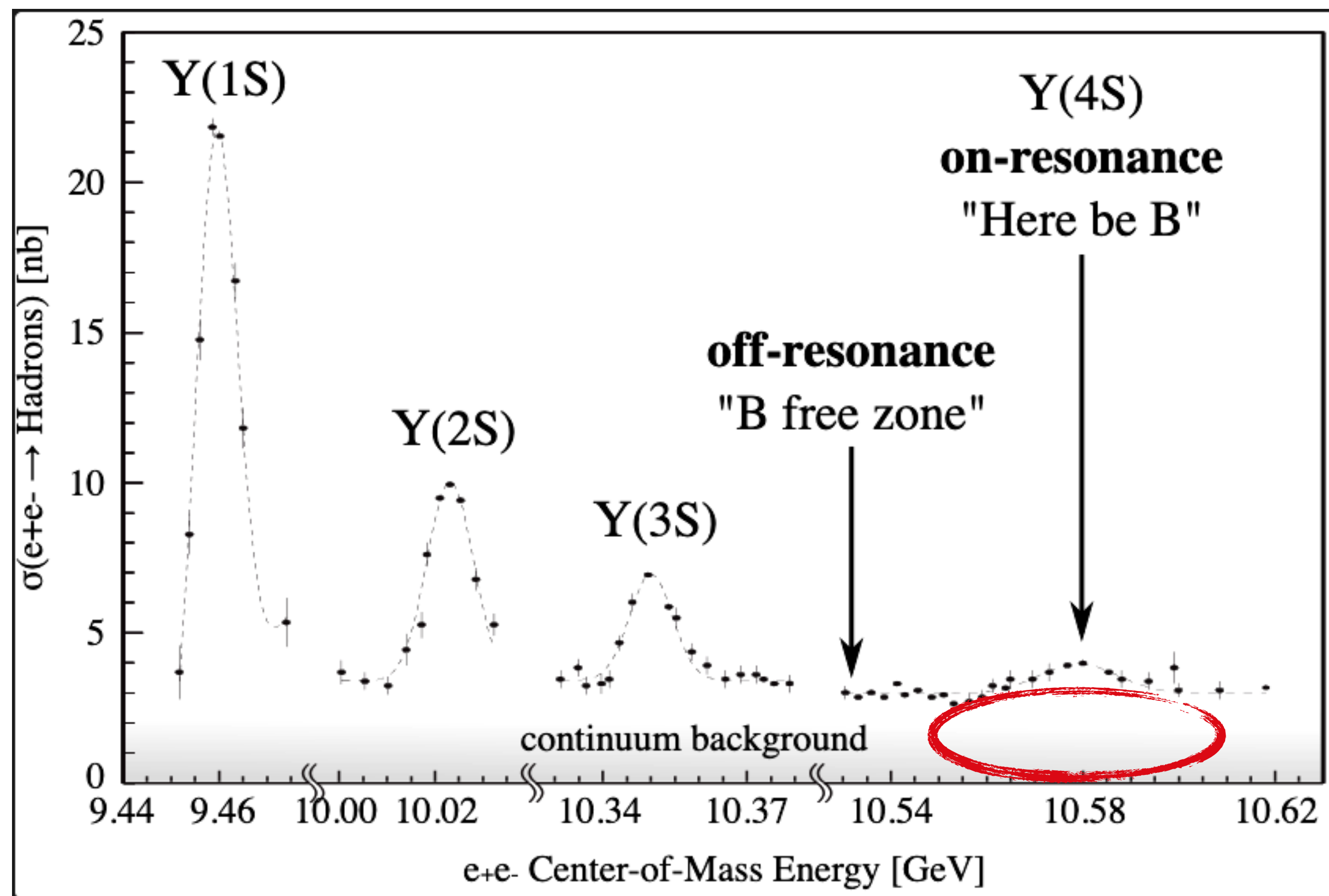


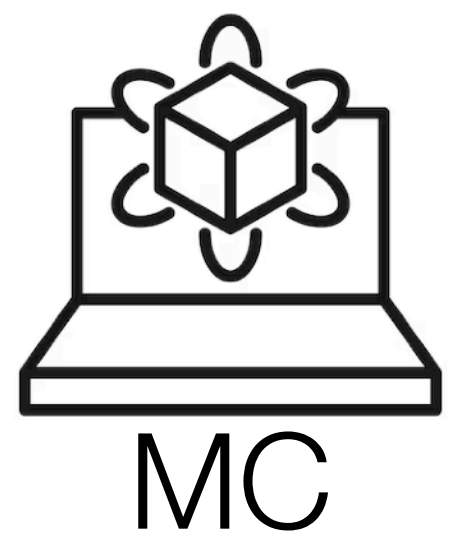
# How?

- SuperKEKB + Belle II tuned to  $\Upsilon(4S)$

•  ~~$e^+e^- \rightarrow \Upsilon(1,2,3S)$~~

→ **Continuum**





# How? - Generation

- From **continuum**

$$e^+e^- \rightarrow [H \rightarrow \Lambda p\pi^-] + X$$

- Replace  $\Xi^0(1530)$  values with hypothesised parameters for **H** in Pythia.

→ Sensitive to strangeness

→ Unlikely to be a daughter-particle

- 26 masses in region  $m_H \in [2.19, 2.23]$  GeV

KKMC

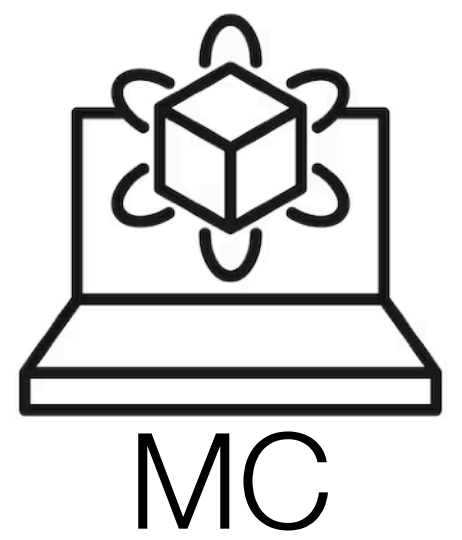


Pythia



EvtGen

# How? - Reconstruction



- Belle II → Bottom-Up

- Steering file

Build lists with FS

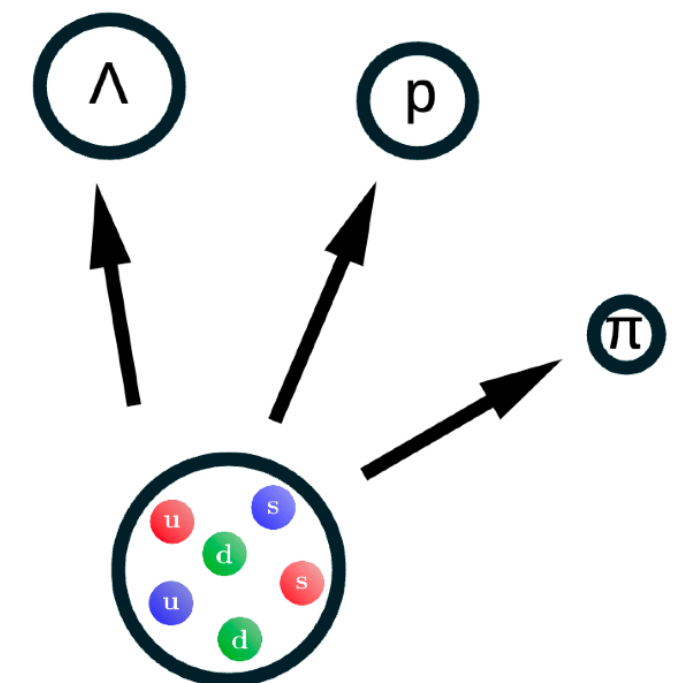
Combine lists + vertex fit

Write out new list + coarse selection

add PID weights

- Neural Network - selection

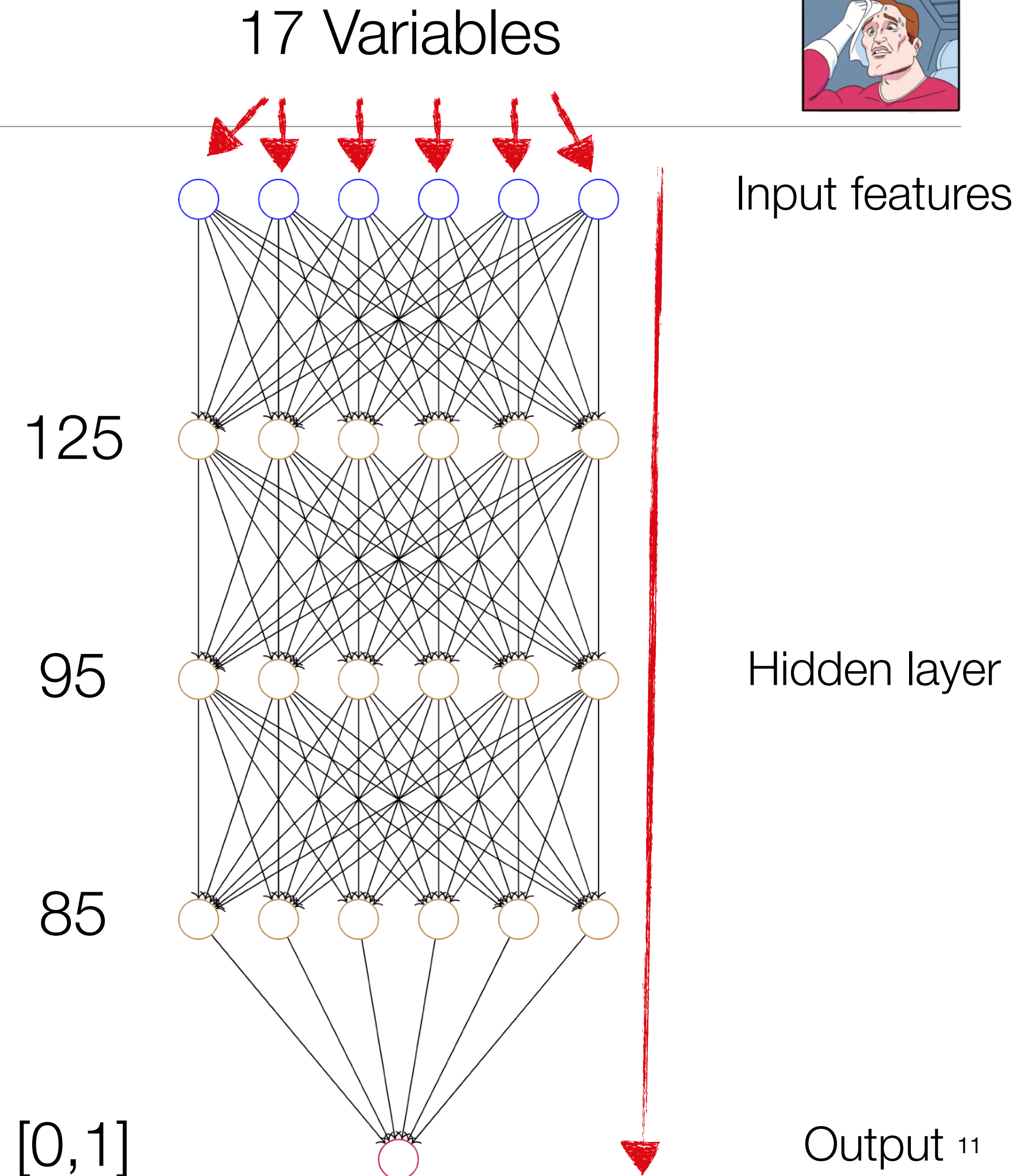
- Best Candidate - selection



# How? - Neural Network

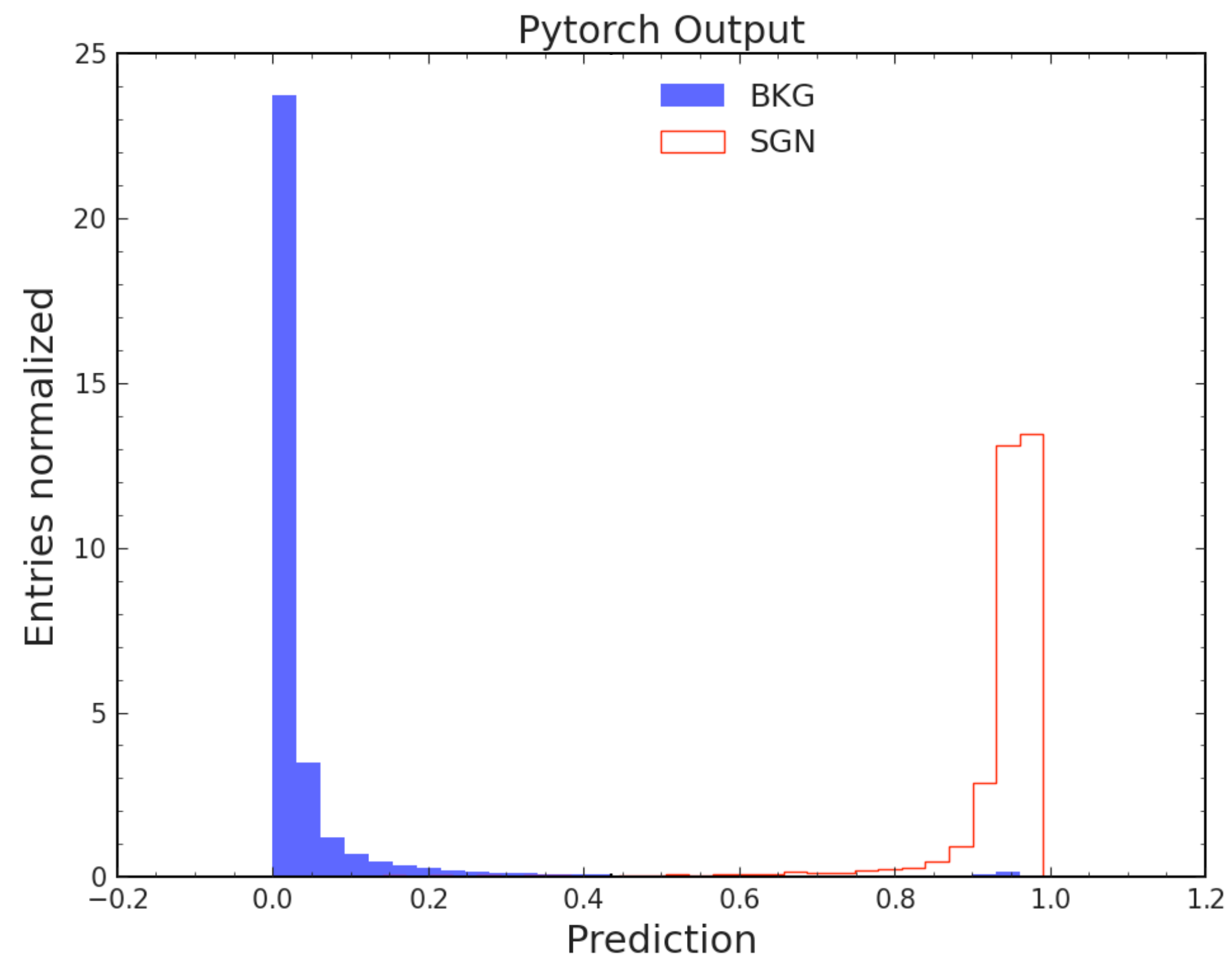


- Deep Neural Network
  - Hyperparameter tuned with Optuna
  - Variables selected via BDT importance
  - ...
- Training + validation on BKG + every other signal mass-point
- Test on data the NN has not seen
  - BKG + signal mass-point(s) not used in training/validation

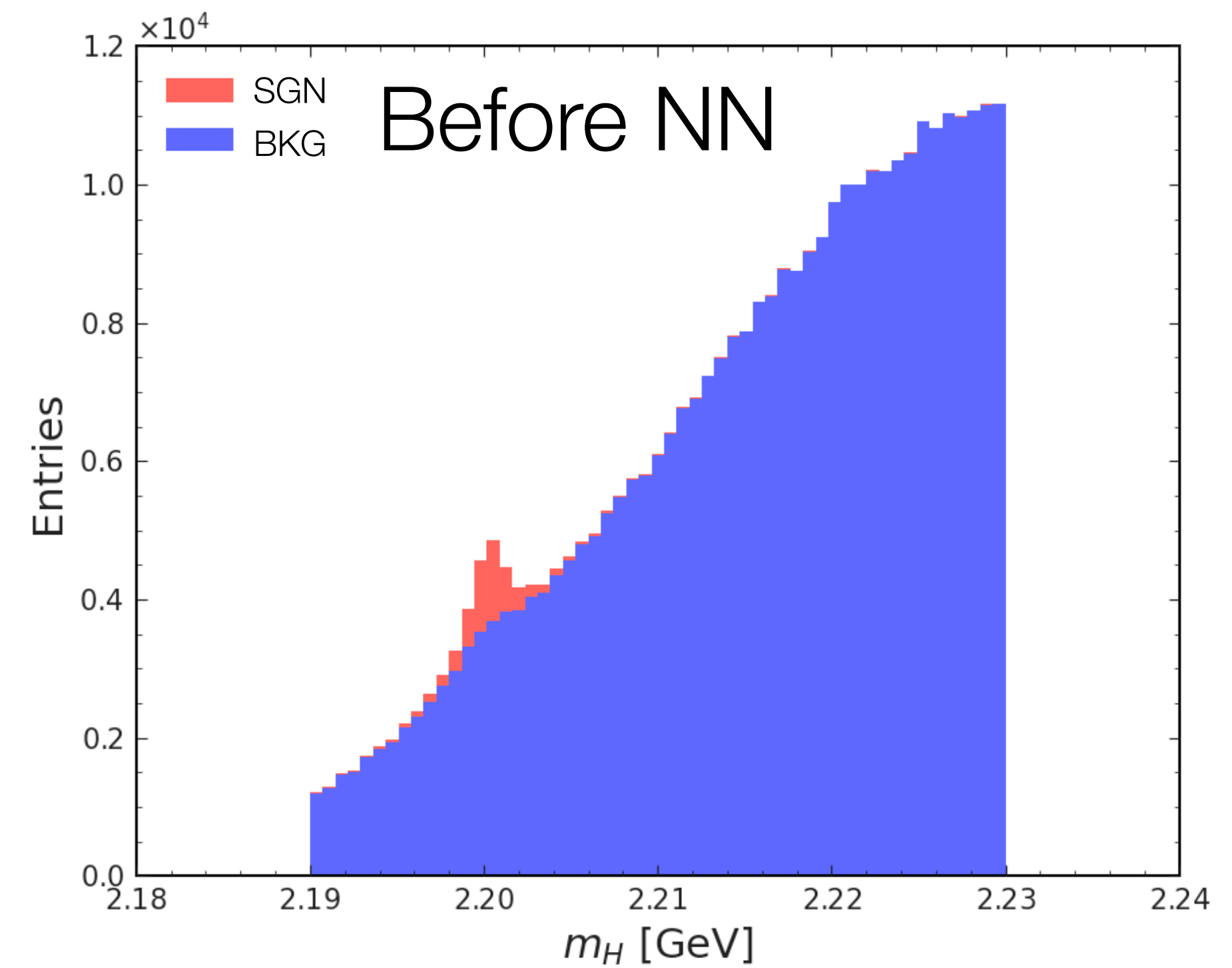




# How? - Neural Network Evaluation

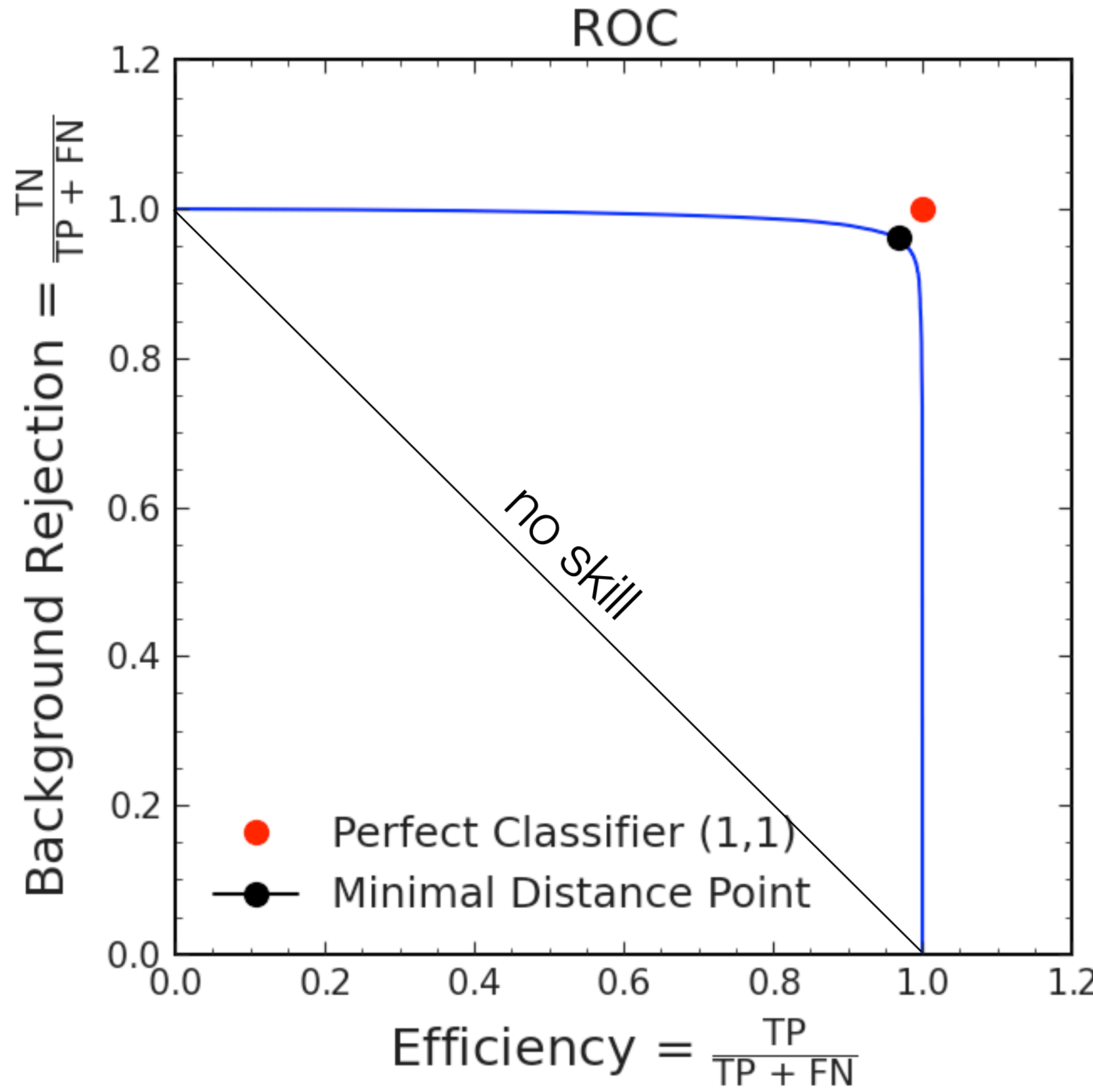
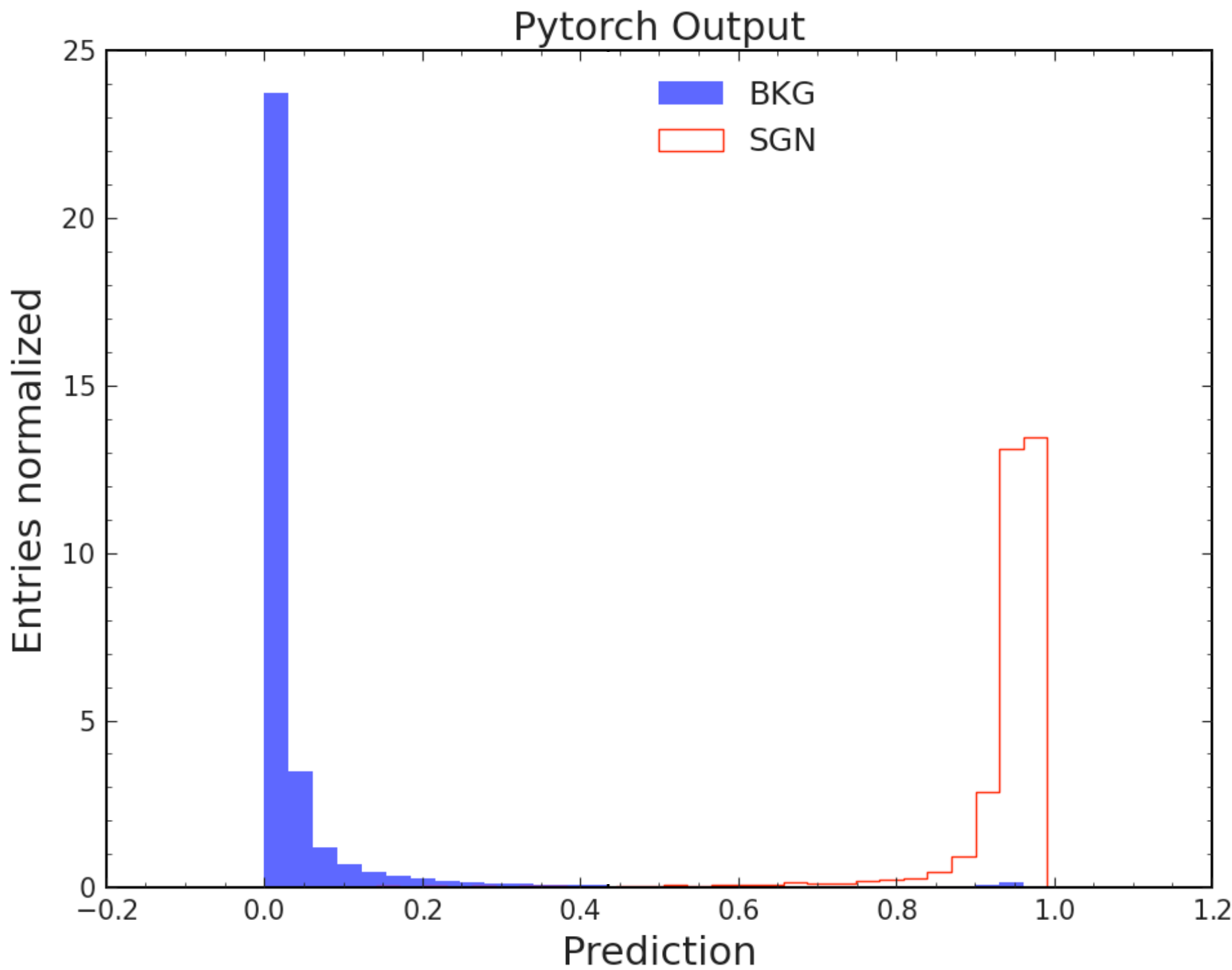


Test-set (not seen before)  
BKG + 5000 artificial signals @ 2.200 GeV injected





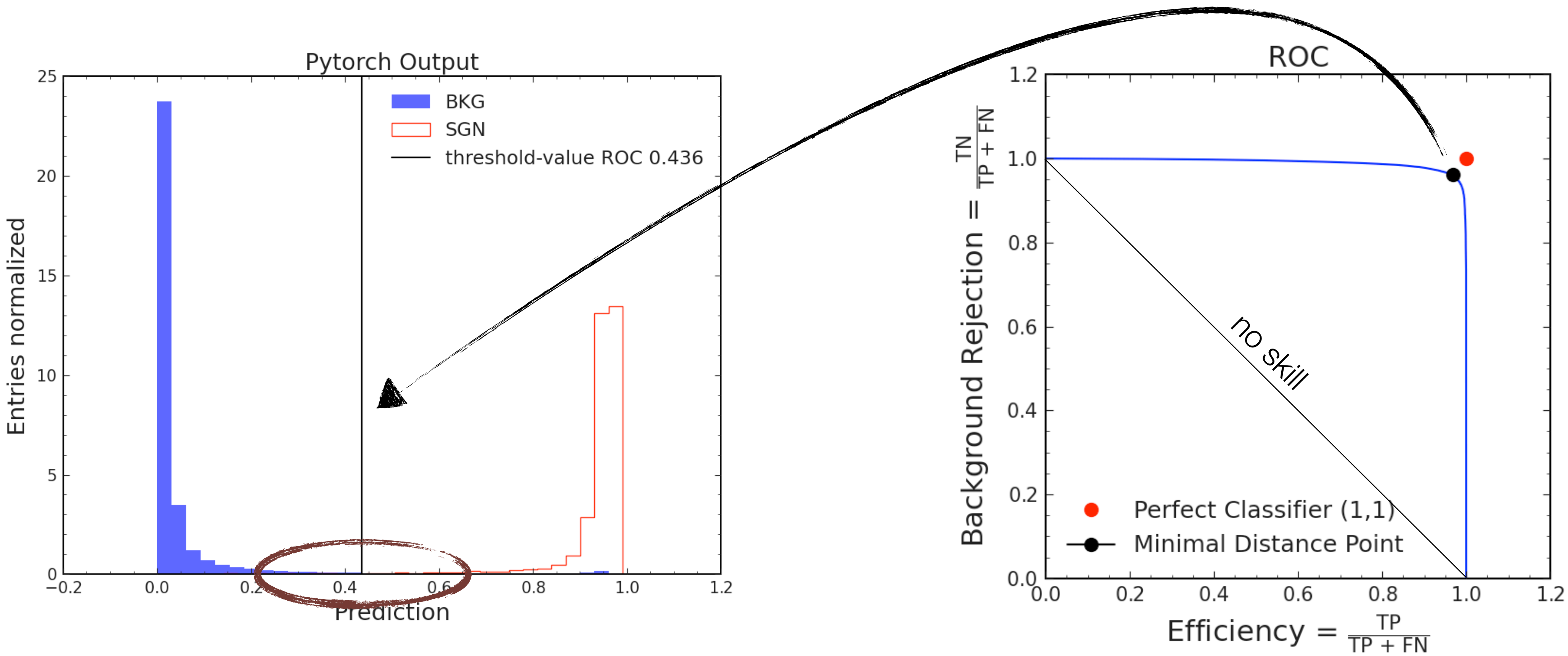
# How? - Neural Network Evaluation



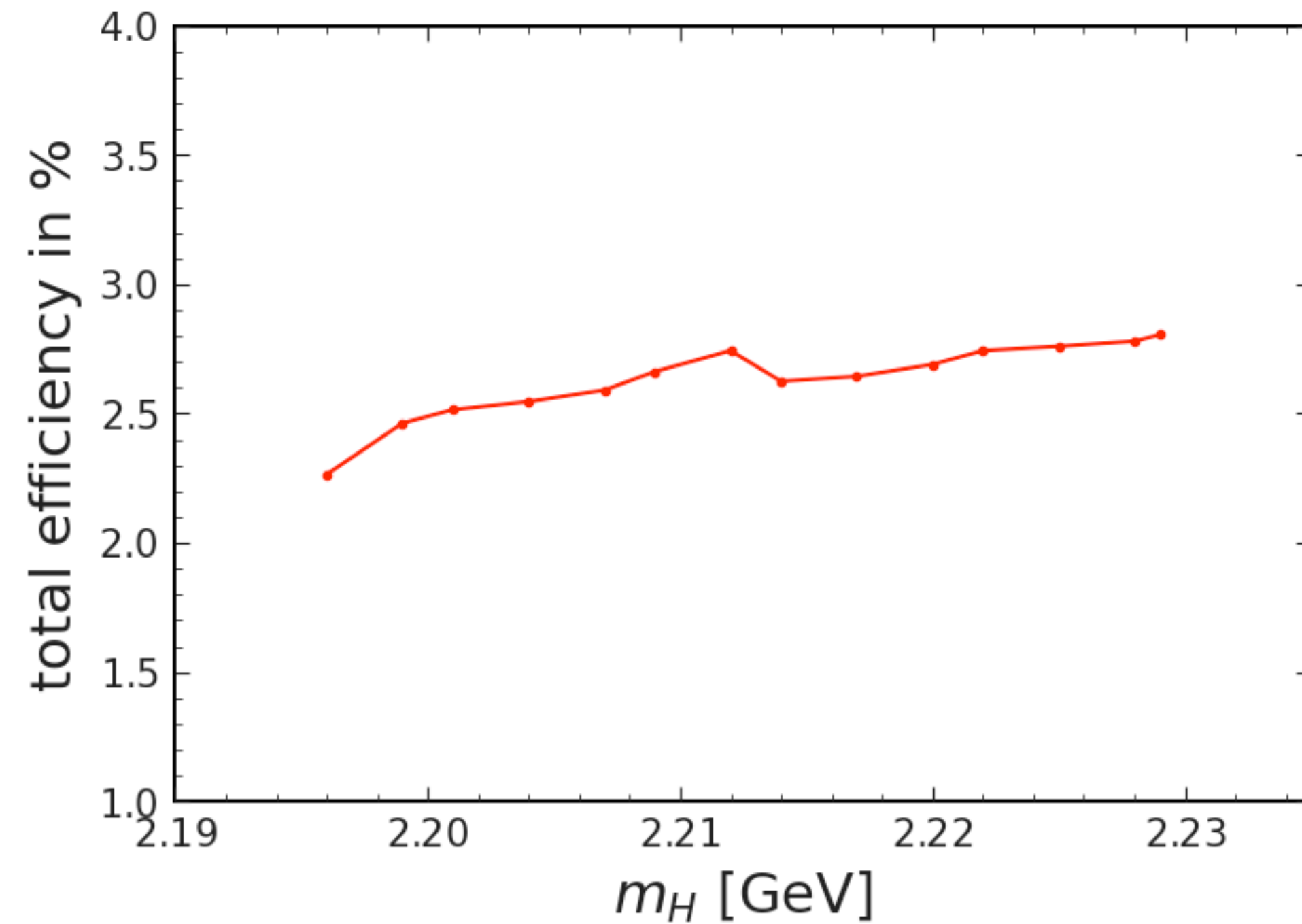


# How? - Neural Network Evaluation

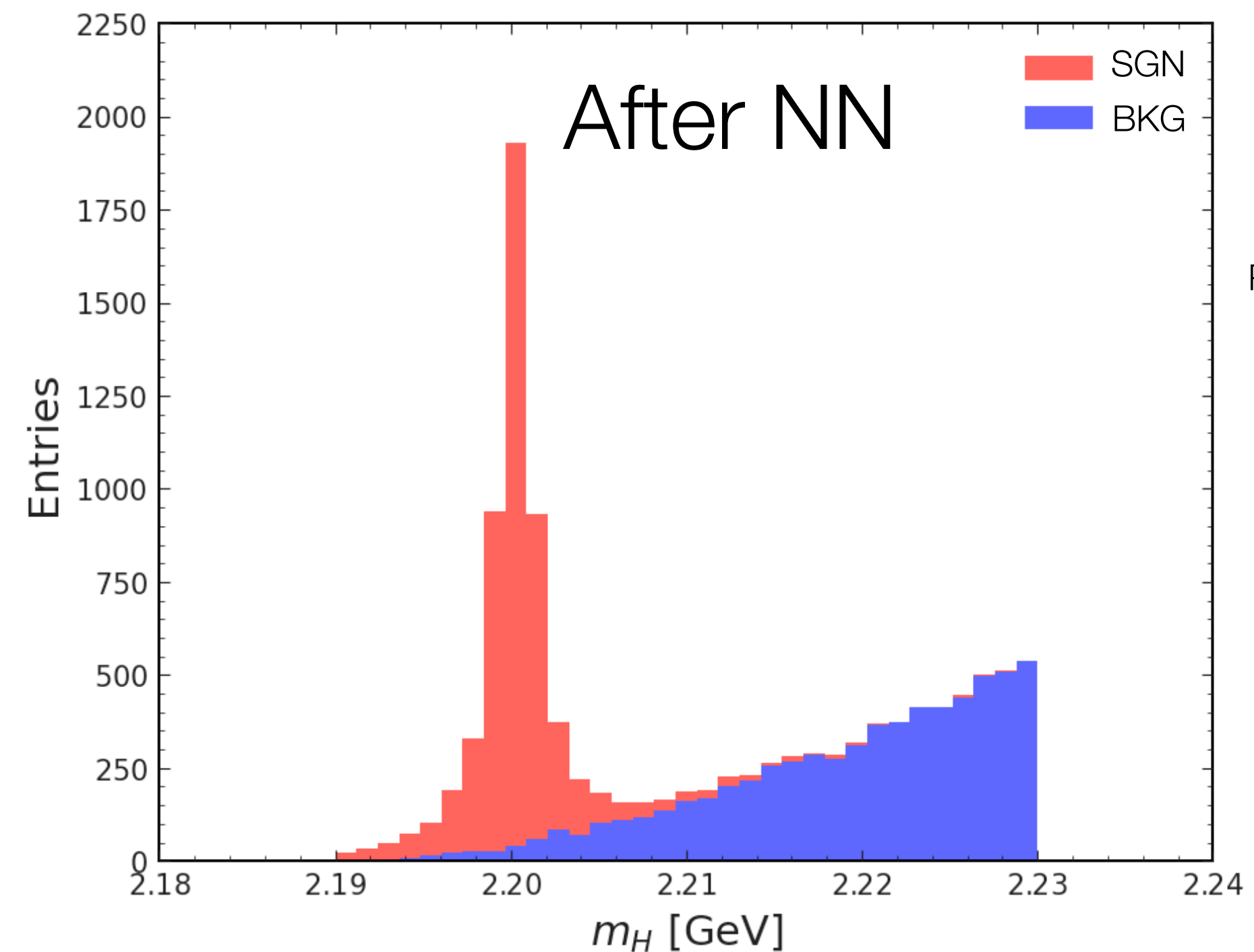
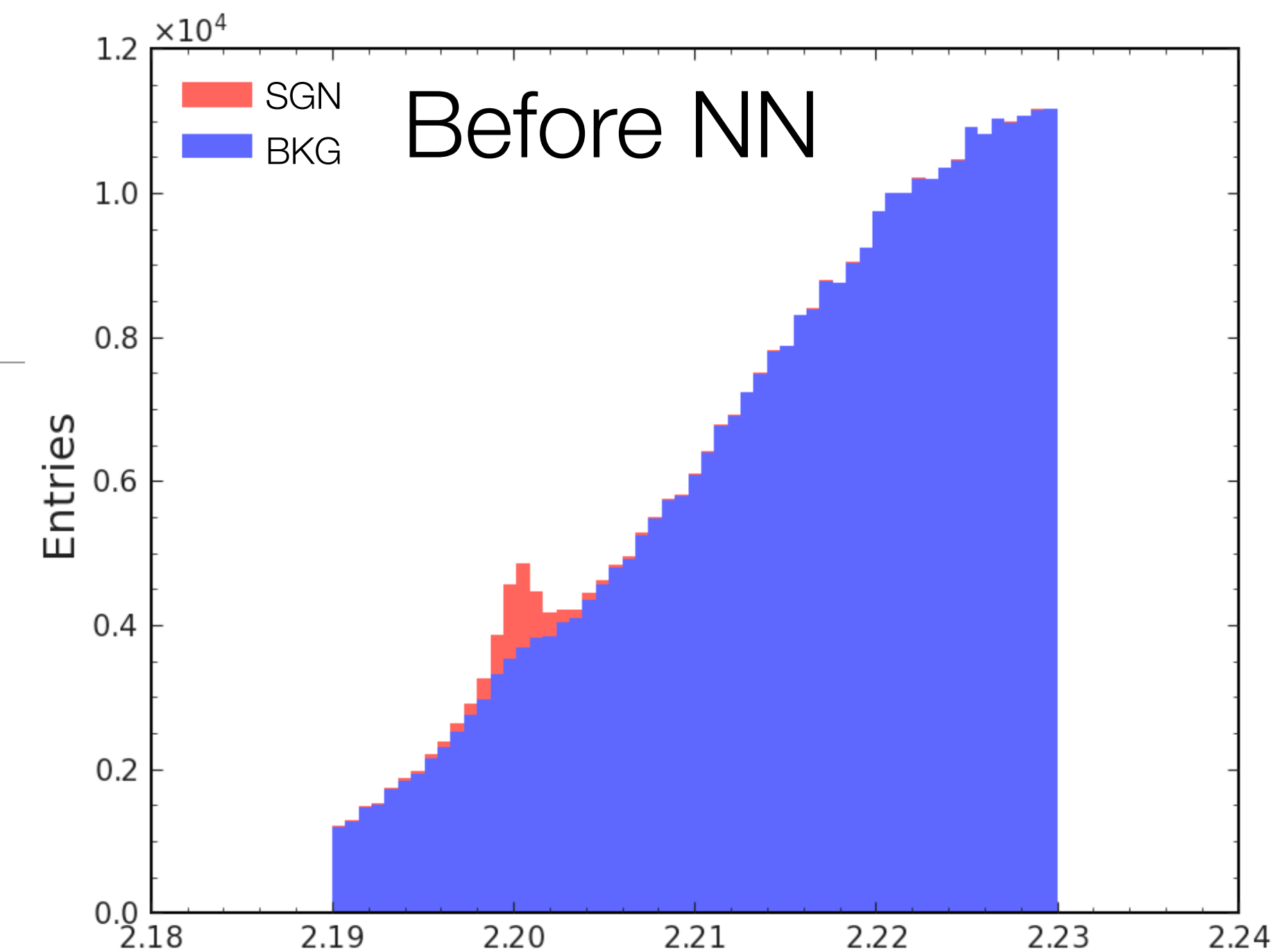
- Threshold in flat area  $\rightarrow$  Network is stable



# How? - Neural Network Evaluation



Total efficiency = Reconstruction-efficiency · NN-efficiency



$$\text{Purity} = \frac{TP}{TP + FP}$$

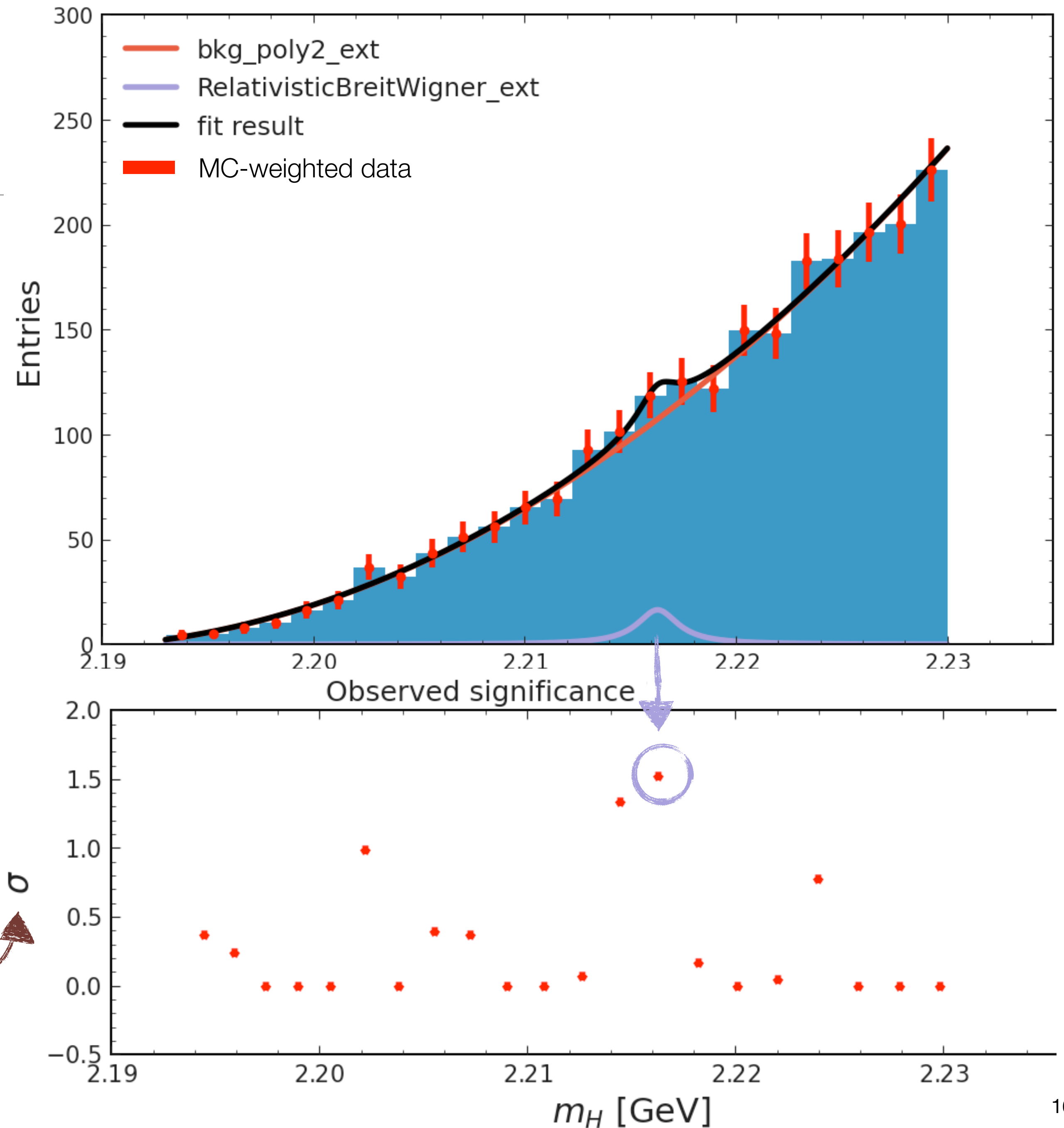
~ 90%  
( $2\sigma$  window)



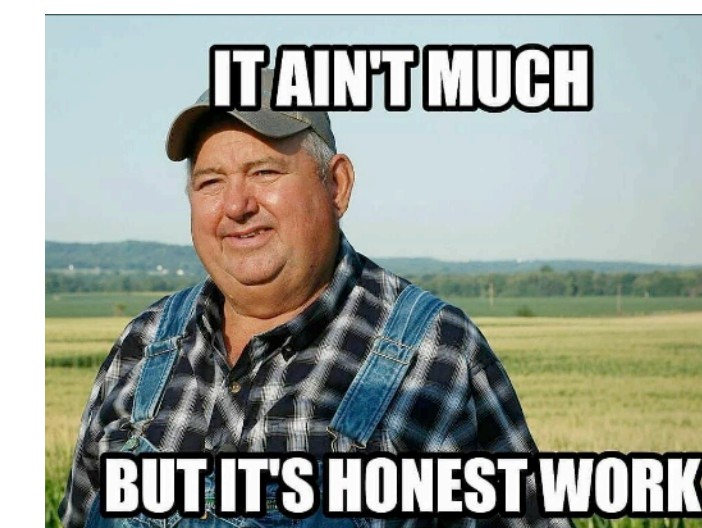
# How? - Inference

- Extended Unbinned Maximum Likelihood
- Keep it simple:
  - SGN  $\rightarrow$  Breit-Wigner profile
  - BKG  $\rightarrow$  2<sup>nd</sup> order polynomial
- Fit on scaled BKG-only—MC
  - $\rightarrow$  **No signal** injected here
  - $\rightarrow$  Peak due to statistical fluctuation
  - $\rightarrow$  Account for Look Elsewhere Effect (**LEE**)

Discovery significance  $\sigma$

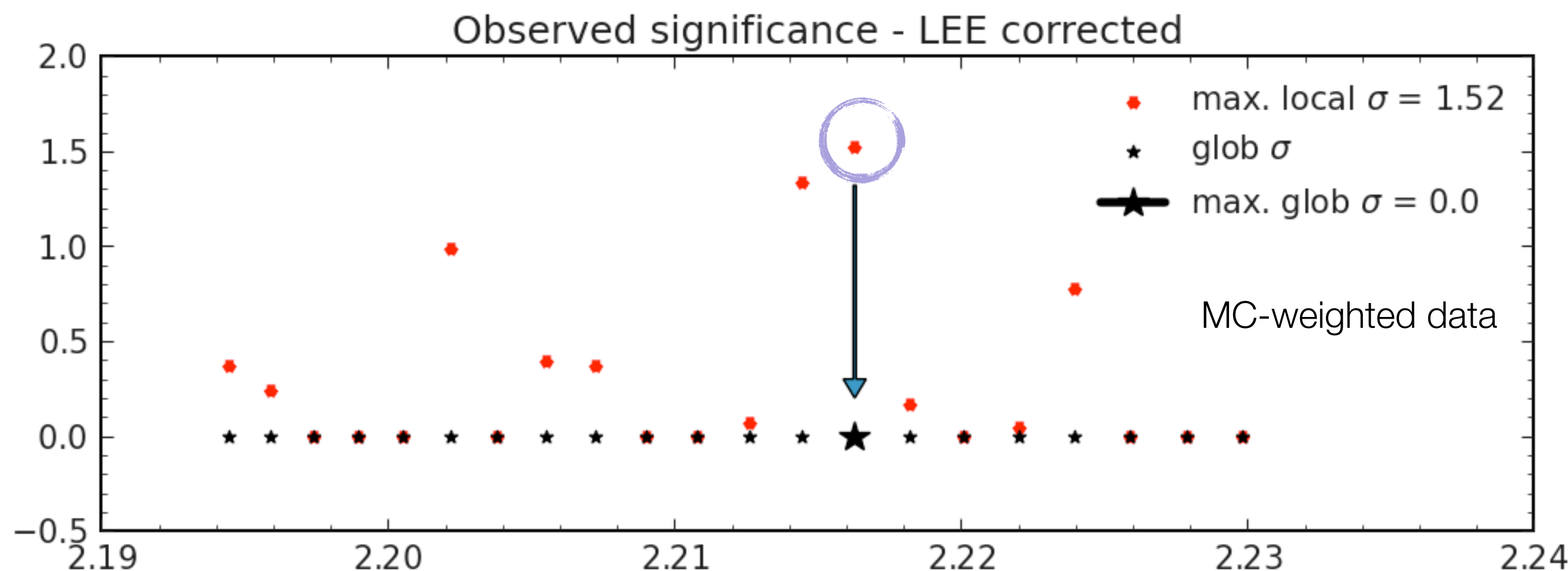


# Results on MC



- Discovery significance  $\sim 0 \sigma$
- Upper Limits  $\sim 1-10 \text{ fb}$
- Compare with on-resonance searches:

Discovery Significance  $\sigma$



→ weak decay<sup>[1]</sup>:

$$\text{BF}(\Upsilon(1,2S) \rightarrow [H \rightarrow \Lambda p \pi] + X) \sim 10^{-7} - 10^{-6}$$

$$\text{UL} \sim 0.5 - 4 \text{ fb}$$

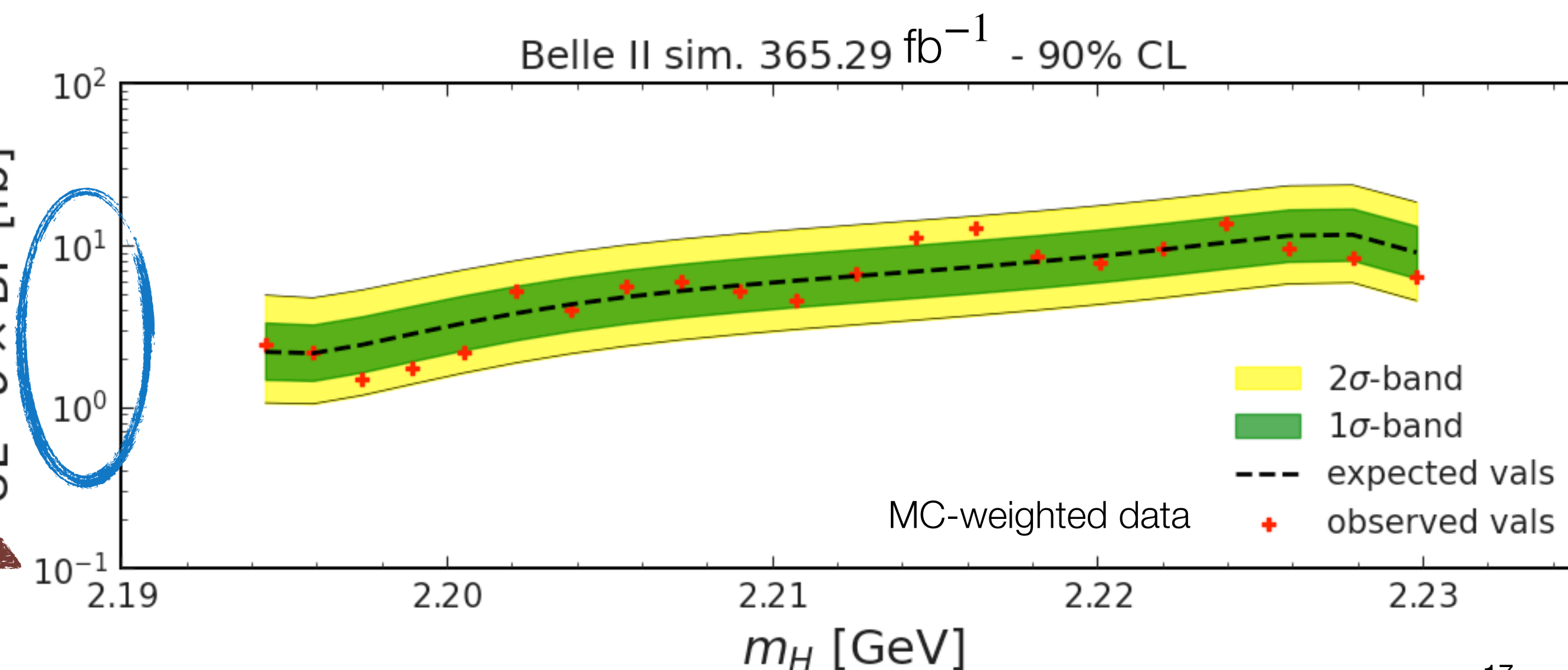
→ stable (MC-only)<sup>[2]</sup>:

$$\text{BF}(\Upsilon(3S) \rightarrow S \Lambda \Lambda) \sim 0.5 \cdot 10^{-7} - 3 \cdot 10^{-6}$$

$$\text{UL} \sim 2 - 12 \text{ fb}$$

Upper Limit

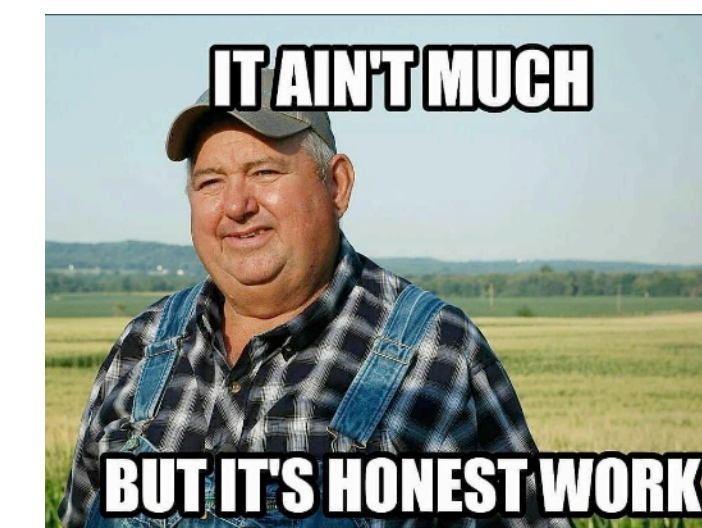
UL -  $\sigma \times \text{BF}$  [fb]



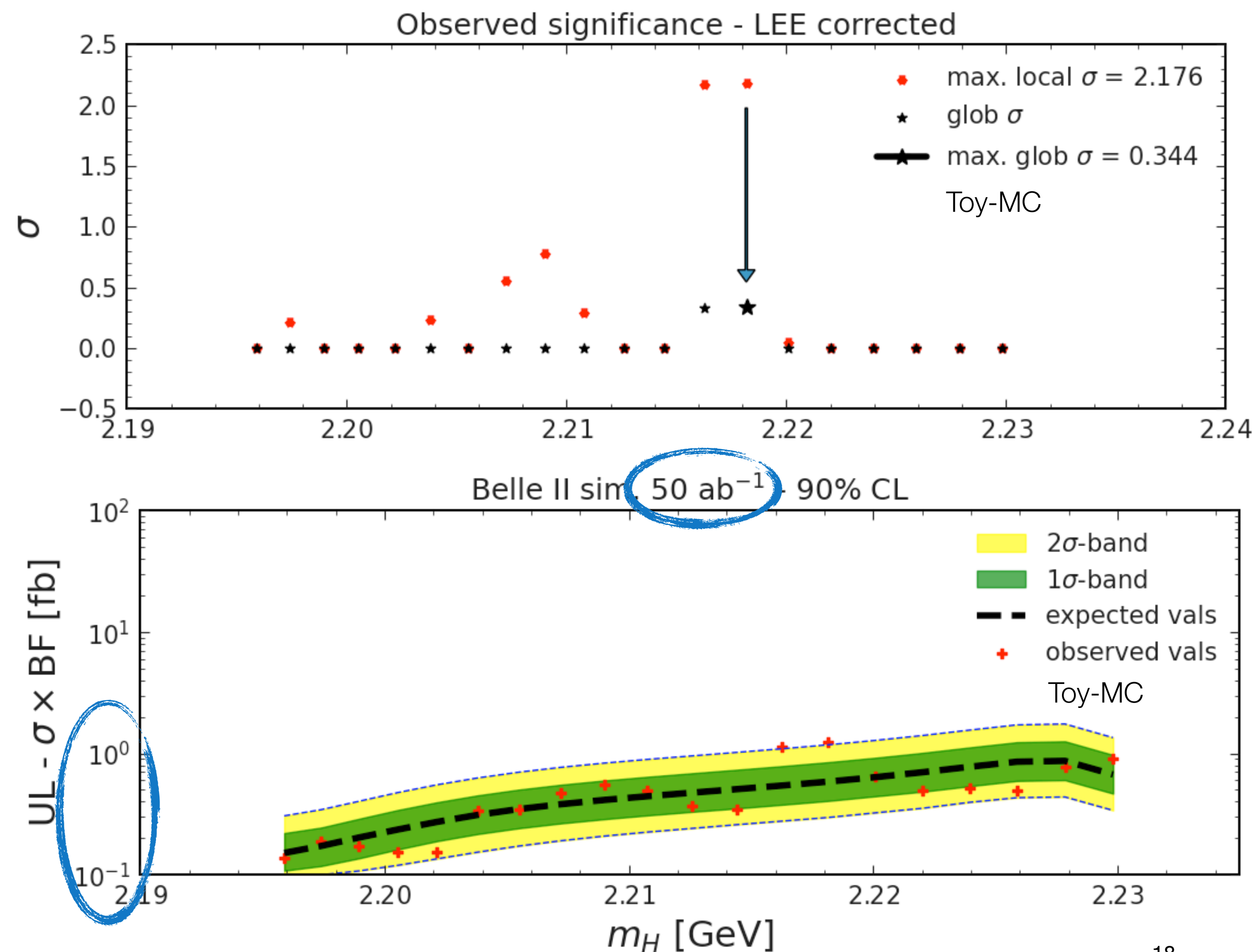
[1] <https://arxiv.org/abs/1302.4028>

[2] [https://b2-pubdb1.desy.de/pub\\_data/documents/2331/](https://b2-pubdb1.desy.de/pub_data/documents/2331/)

# Conclusion and Outlook



- **First search for a weakly decaying H-Dibaryon from continuum ready to unblind**
- Analysis-pipeline including Neural Networks **successful** for classification of this kind
- **UL of 1-10 fb** with current dataset available
- Luminosity **projection to 50 ab<sup>-1</sup>** further reduces the UL by order of magnitude



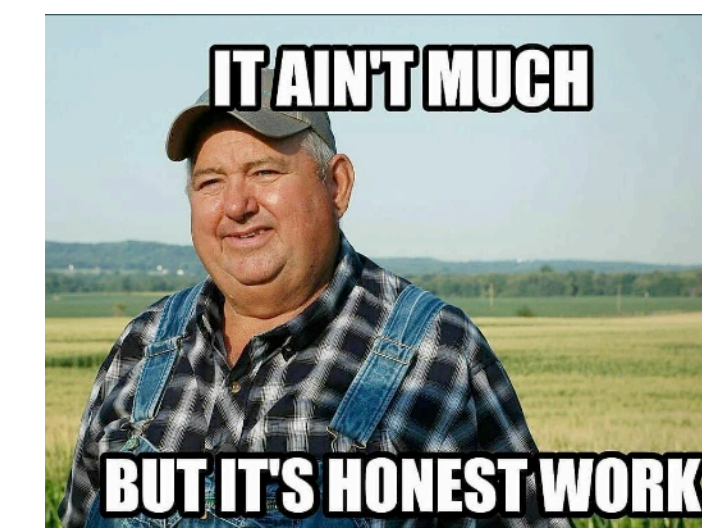
---

**Thanks!**

---

# Backup

# Results on MC

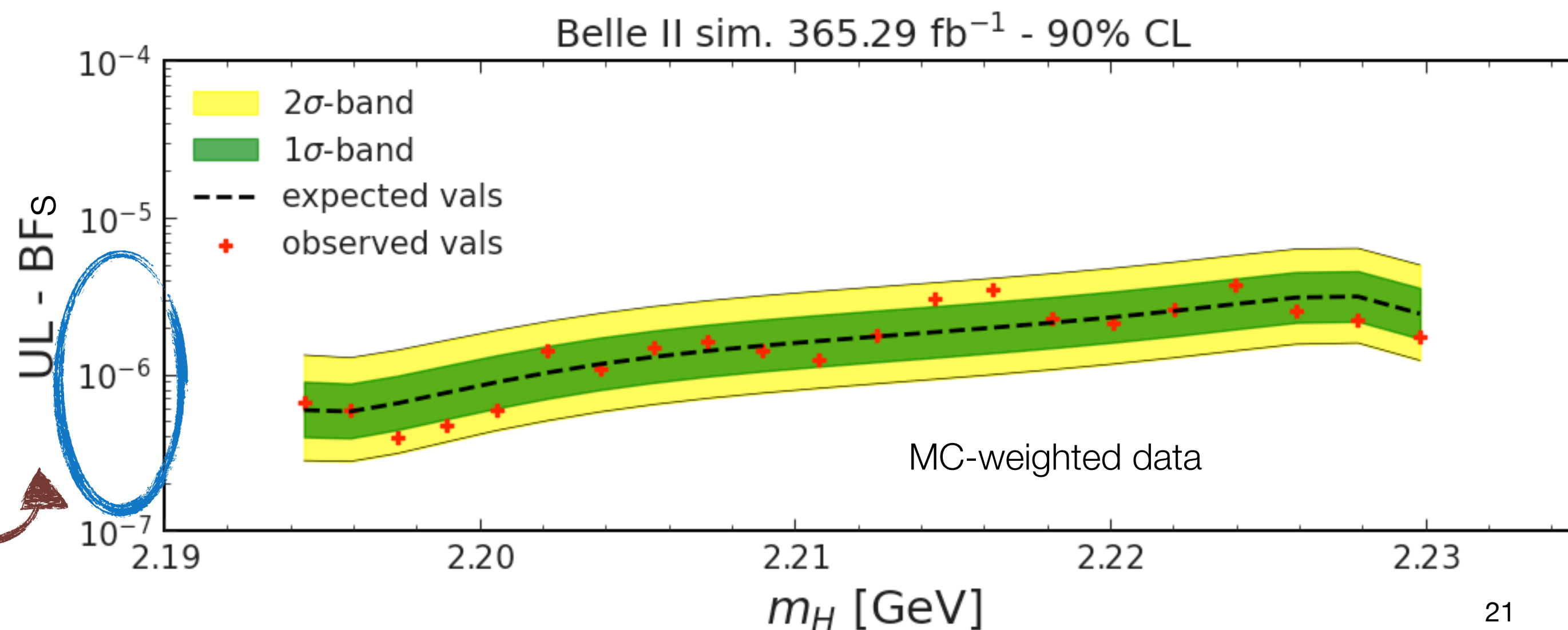
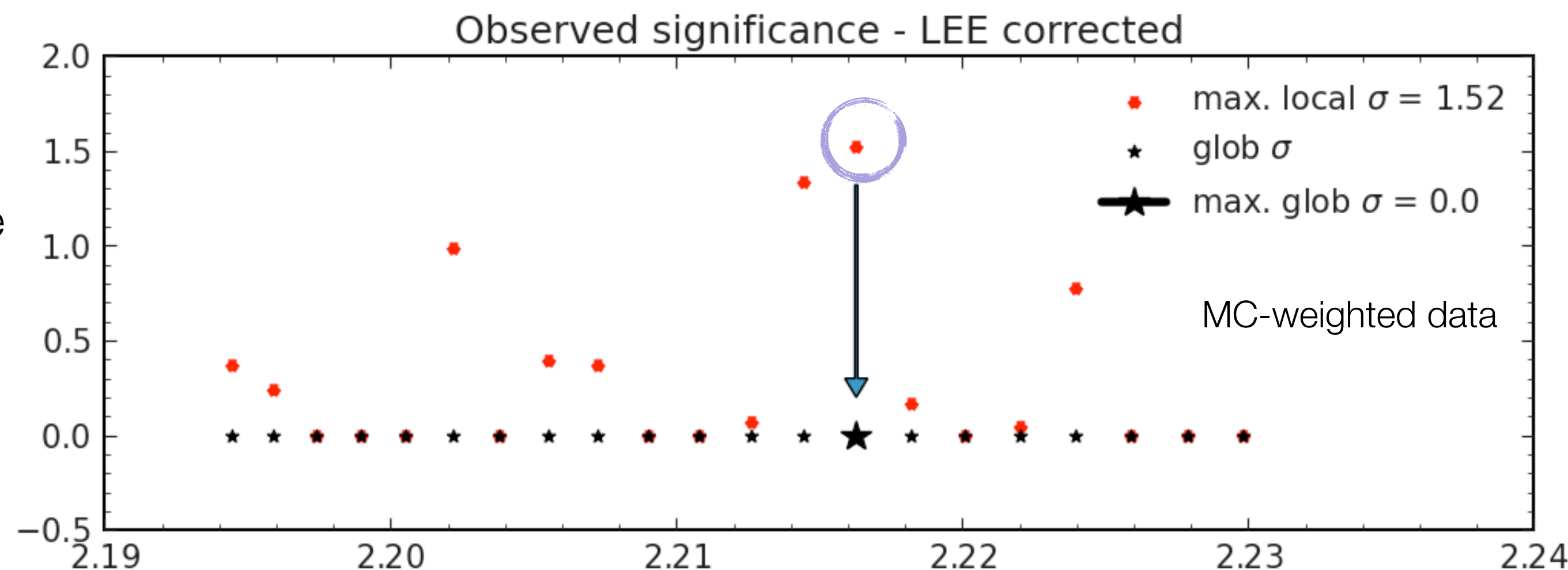


- Discovery significance  $\sim 0 \sigma$
- Upper Limits  $\sim 6 \cdot 10^{-7} - 3 \cdot 10^{-6}$
- Compare with on-resonance searches:

→ weak decay<sup>[1]</sup>:  
 $BF(\Upsilon(1,2S) \rightarrow [H \rightarrow \Lambda p \pi] + X)$   
 $UL \sim 10^{-7} - 10^{-6}$

→ stable (MC-only)<sup>[2]</sup>:  
 $BF(\Upsilon(3S) \rightarrow S \Lambda \Lambda)$   
 $UL \sim 0.5 \cdot 10^{-7} - 3 \cdot 10^{-6}$

Discovery  
Significance



Upper Limit



[1] <https://arxiv.org/abs/1302.4028>

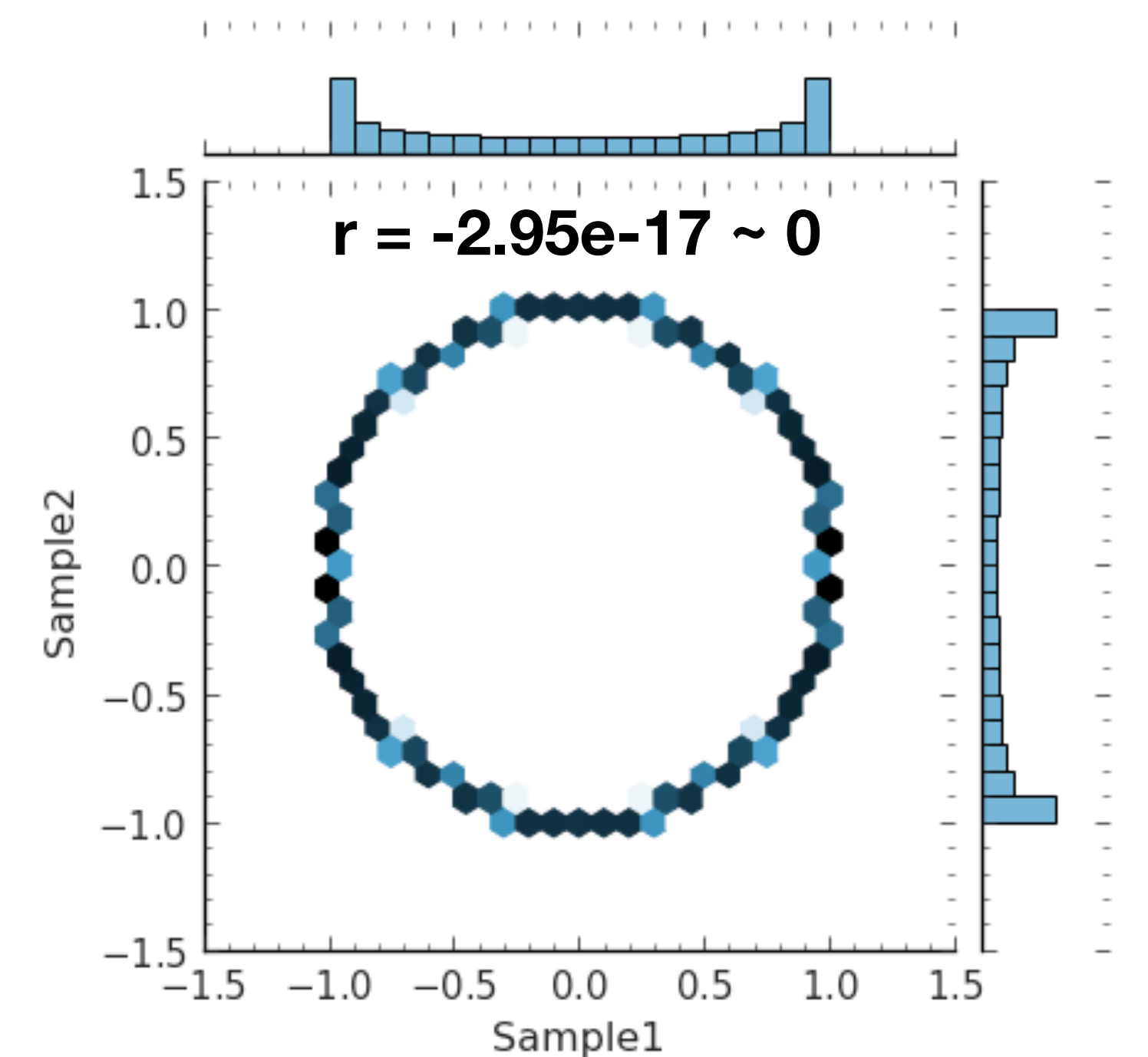
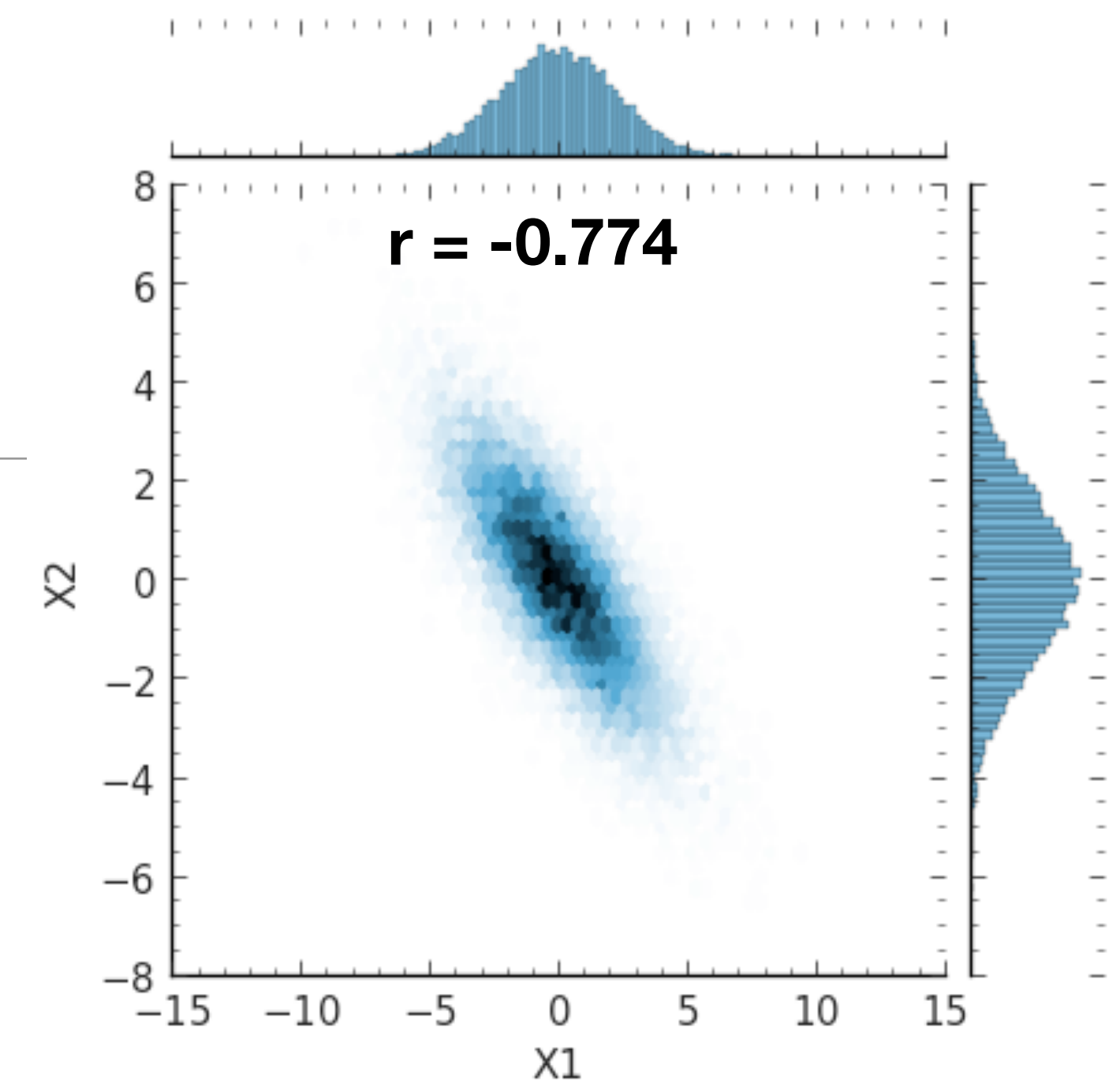
[2] [https://b2-pubdb1.desy.de/pub\\_data/documents/2331/](https://b2-pubdb1.desy.de/pub_data/documents/2331/)

$$r_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2} \sqrt{\sum (y_i - \bar{y})^2}}$$

$$r_{xy} \in [-1, 1]$$

## Backup - Correlations

- Important when using NN and or fitting PDFs
- Classic: Pearson Coefficient  
But: Linear only
- Better: Uniform Distributions (flat)
- If uncorrelated
  - **uniformly distributed** in the 2D-plane
  - $n_{expect} = N/n^2$  events

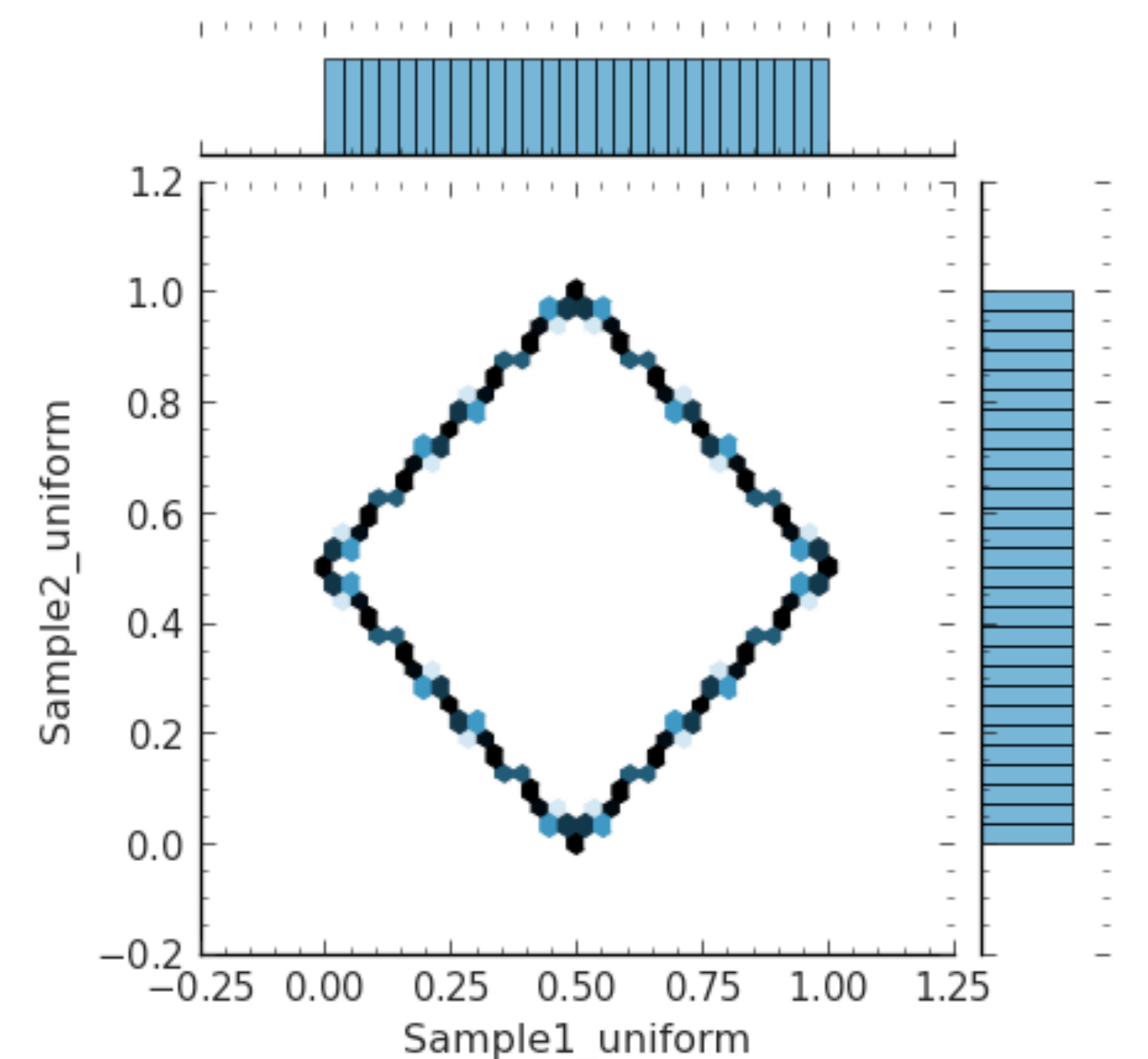
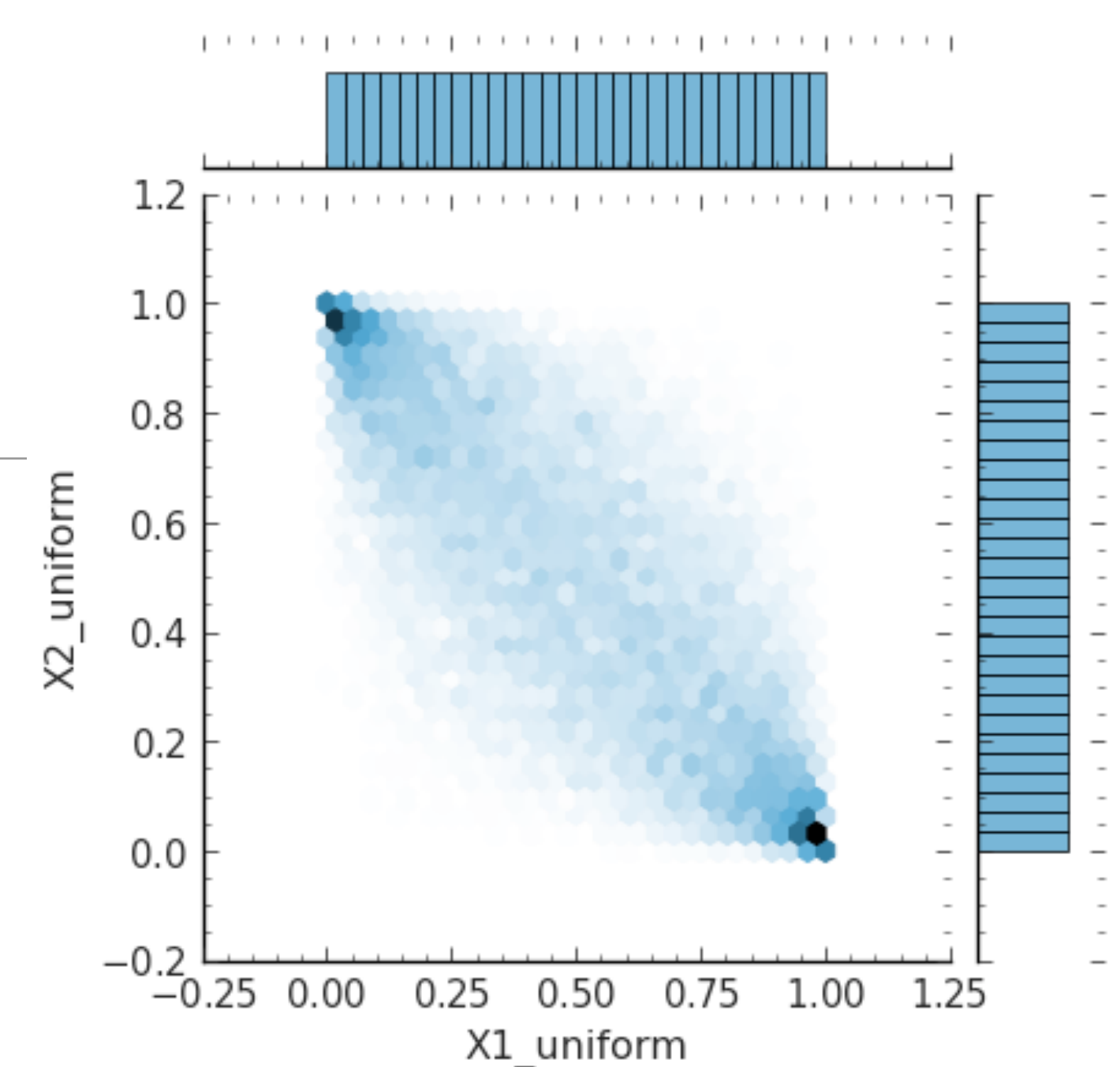


$$r_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2} \sqrt{\sum (y_i - \bar{y})^2}}$$

$$r_{xy} \in [-1, 1]$$

## Backup - Correlations

- Important when using NN and or fitting PDFs
- Classic: Pearson Coefficient  
But: Linear only
- Better: Uniform Distributions (flat)
- If uncorrelated
  - **uniformly distributed** in the 2D-plane
  - $n_{expect} = N/n^2$  events

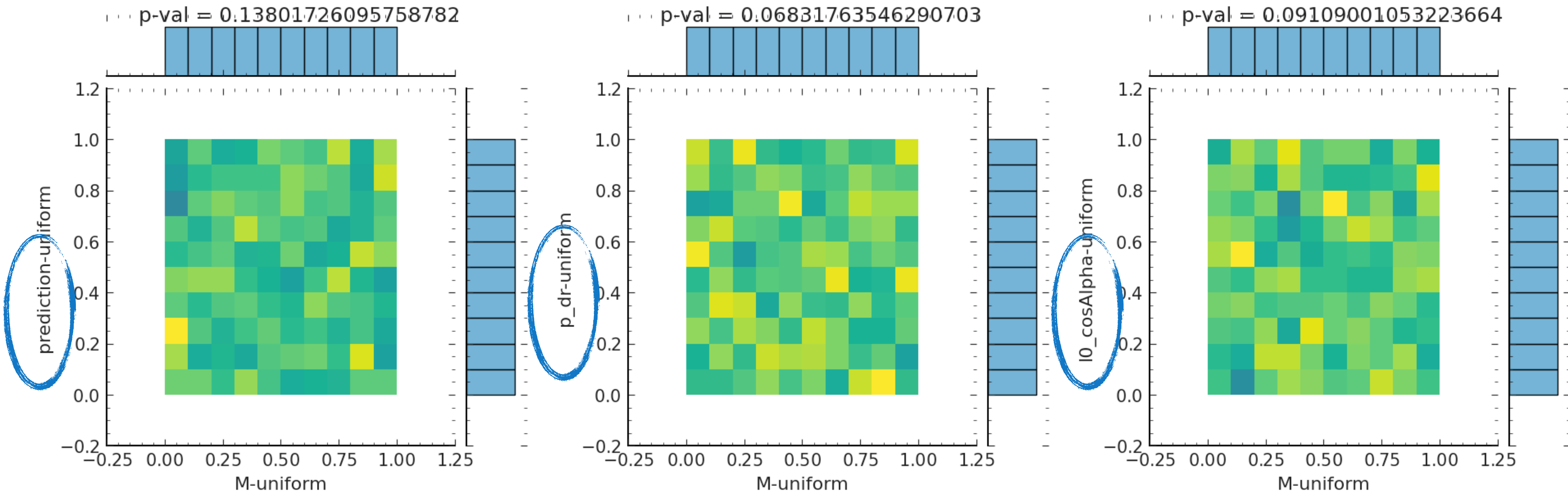


## Backup - Correlations

---

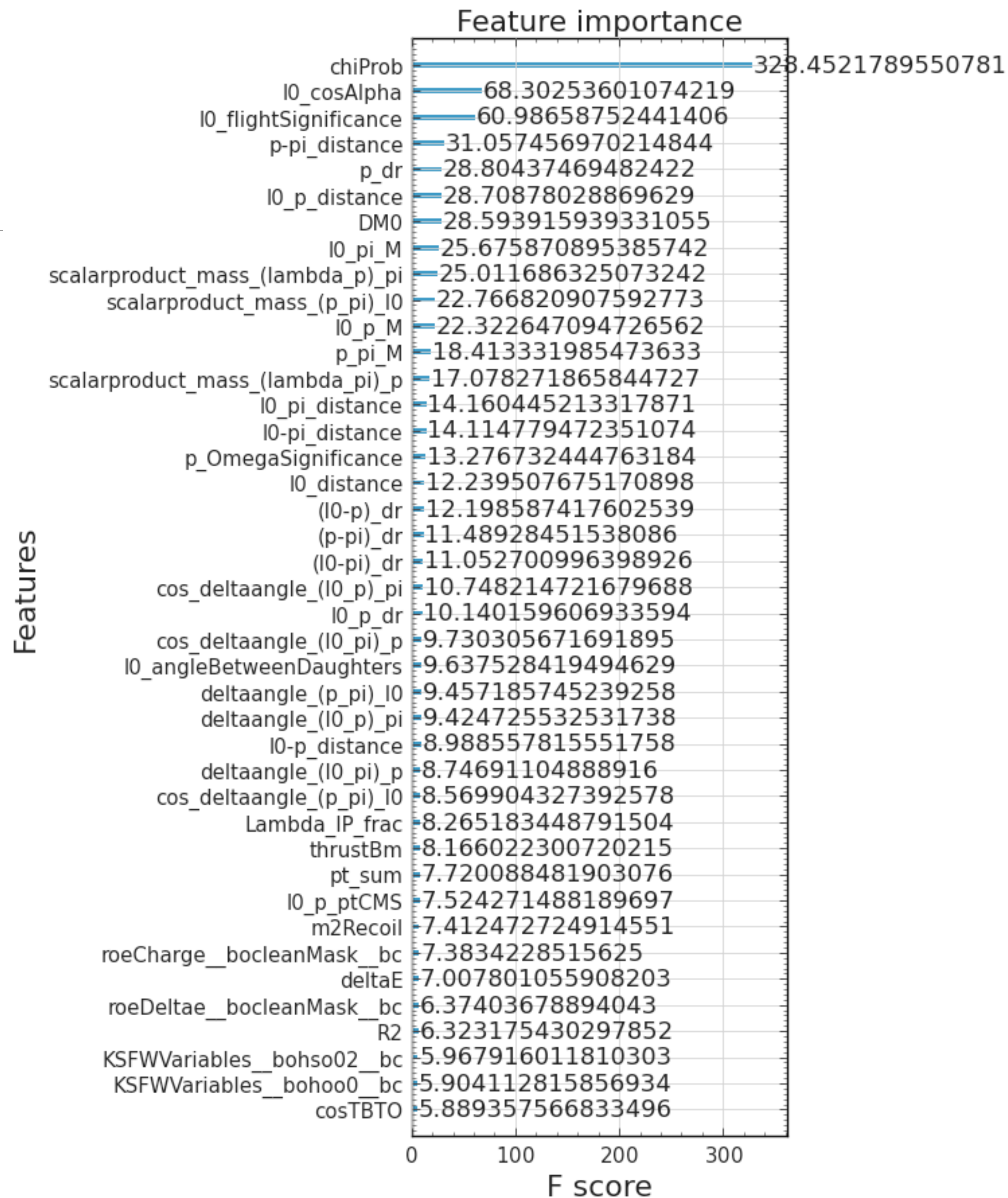
- Important when using NN and or fitting PDFs
- Better: Uniform Distributions (flat)
- If uncorrelated  $\rightarrow$  **uniformly distributed** in the 2D-plane  $\rightarrow n_{expect} = N/n^2$  events

# How? - Correlations



# Backup - Variables

```
▼ root
  ▼ variables [] 18 items
    0 "chiProb"
    1 "DM0"
    2 "l0_cosAlpha"
    3 "l0_flightSignificance"
    4 "Lambda_IP_frac"
    5 "p-pi_distance"
    6 "l0-p_distance"
    7 "l0-pi_distance"
    8 "p_OmegaSignificance"
    9 "l0_p_distance"
    10 "(p-pi)_dr"
    11 "(l0-pi)_dr"
    12 "p_dr"
    13 "(l0-p)_dr"
    14 "l0_pi_distance"
    15 "deltaangle_(p_pi)_l0"
    16 "cos_deltaangle_(p_pi)_l0"
    17 "isSignal"
  files "all"
  ► pytorch
  ► pytorch_parameter
```



# Backup - Systematics

---

Luminosity - 0.5%

$$\sigma(e^+e^- \rightarrow q\bar{q}) \cdot \text{BF}(q\bar{q} \rightarrow H + X) \cdot \text{BF}(H \rightarrow \Lambda p\pi) = \frac{N_{sig}}{\mathcal{L} \cdot \epsilon \cdot \text{BF}(\Lambda \rightarrow p\pi)}$$

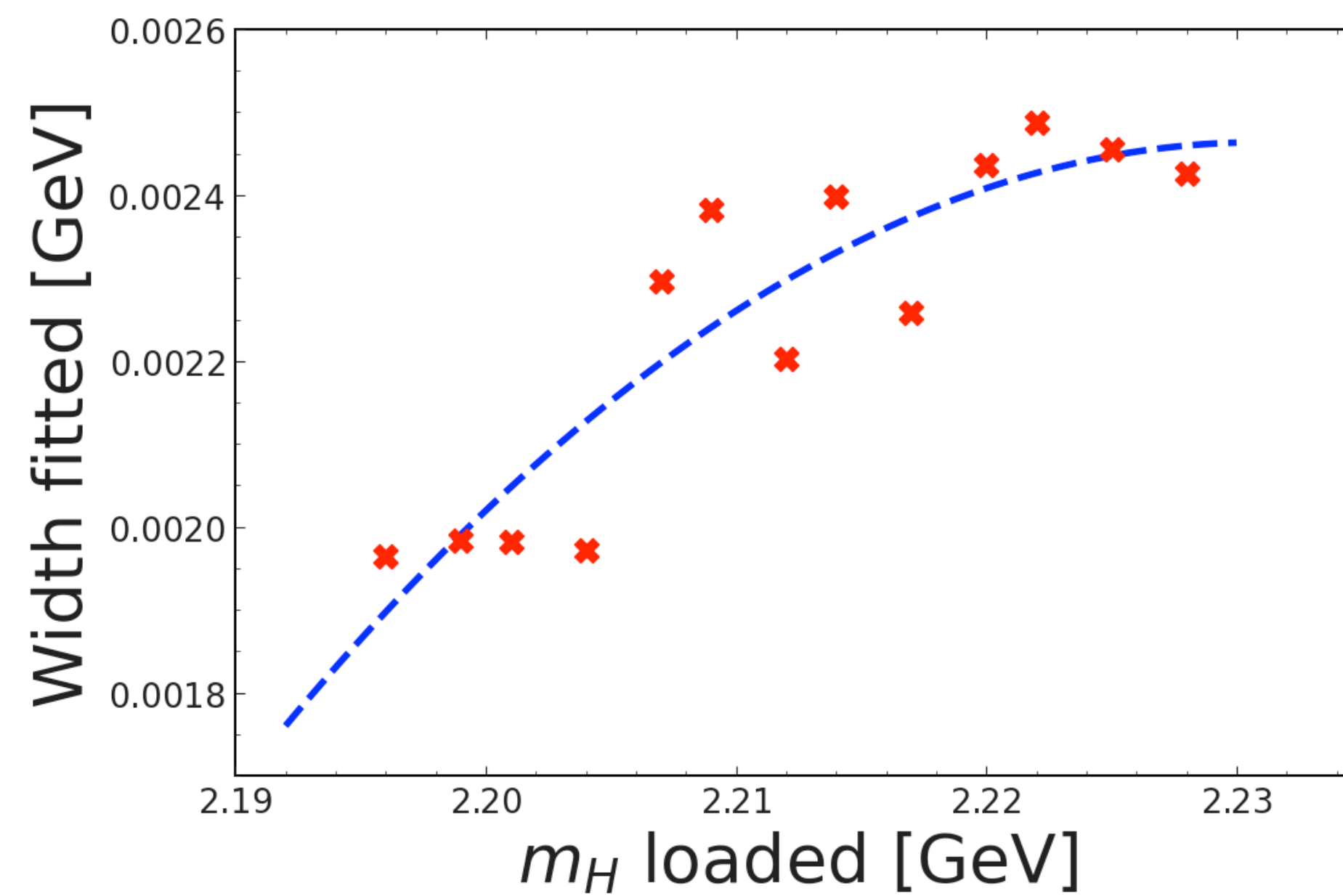
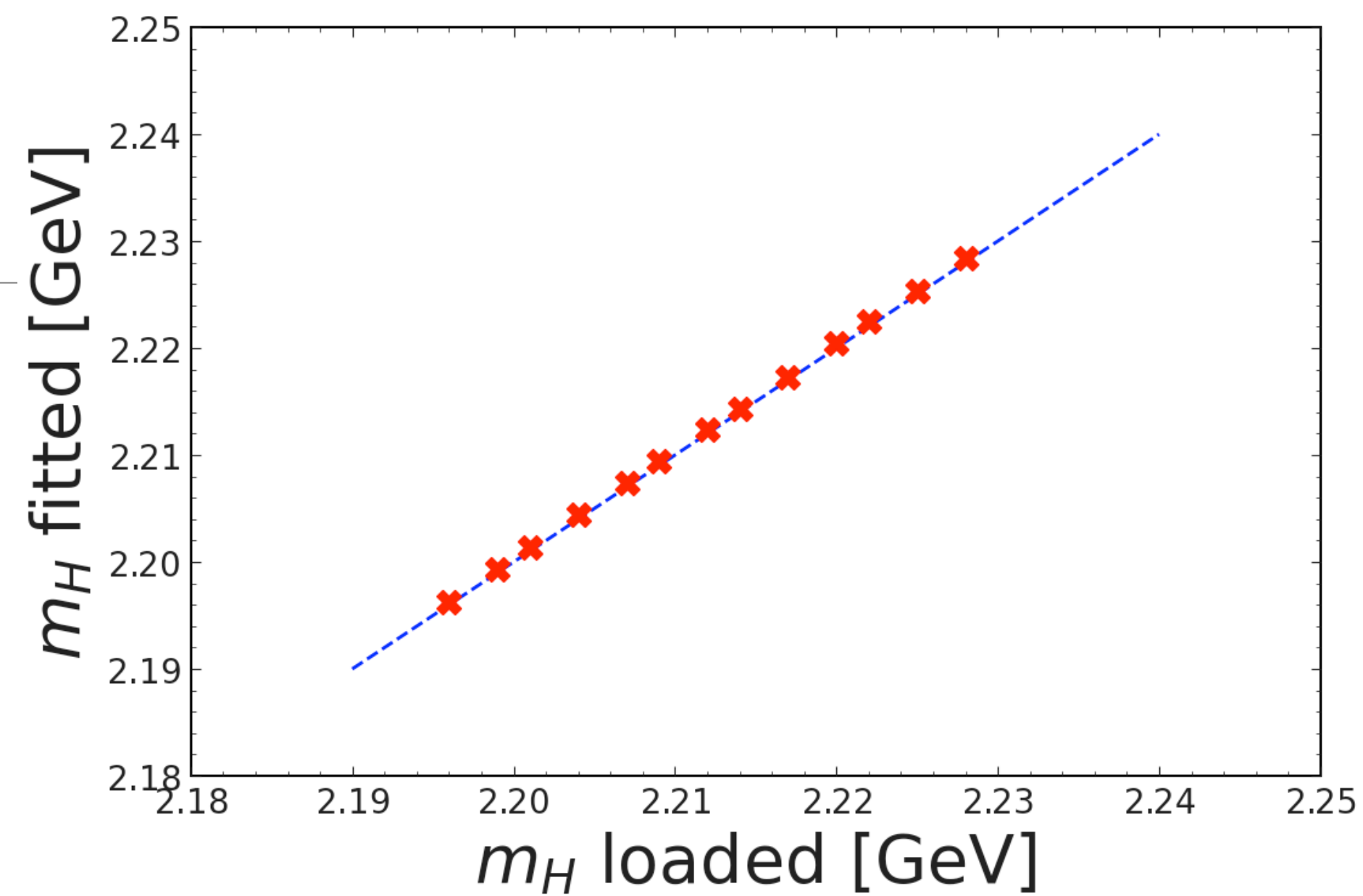
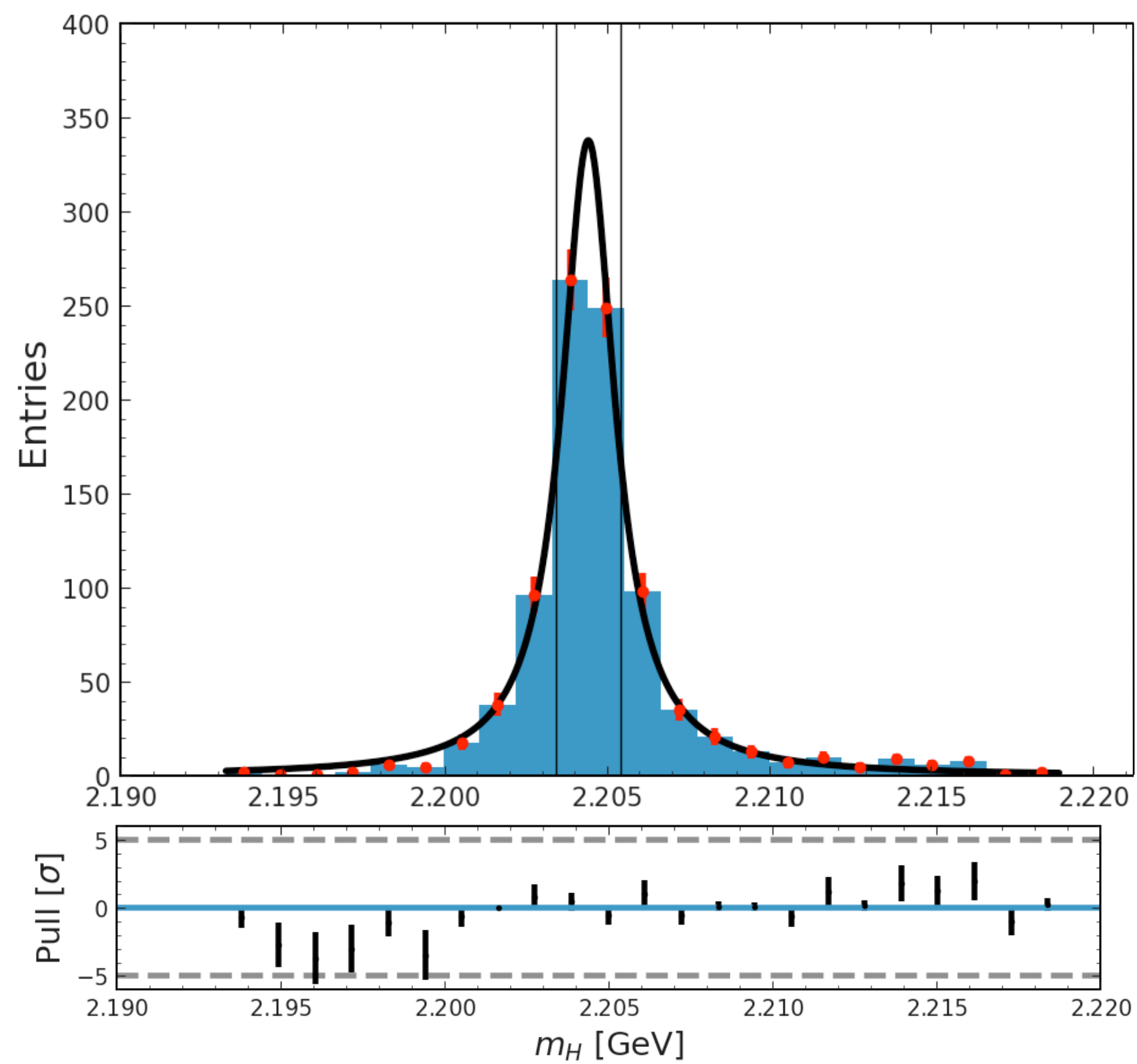
Signal Efficiency - 10.0%

BF( $\Lambda \rightarrow p\pi$ ) - 0.5%

NN efficiency - 1.0%

quadrature sum - 10.1%

# Raw-Fit



# Test statistics

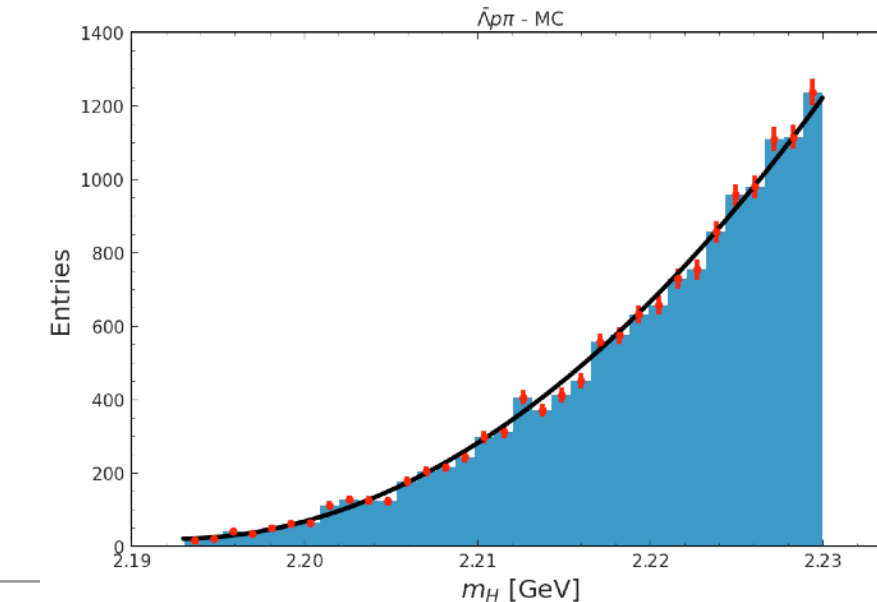
---

$$\tilde{t}_\mu = \begin{cases} -2 \ln \left( \frac{L(x; \mu, \hat{\Theta})}{L(x; 0, \hat{\Theta})} \right), & \mu < 0 \\ -2 \ln \left( \frac{L(x; \mu, \hat{\Theta})}{L(x; \mu, \hat{\Theta})} \right), & \mu \geq 0 \end{cases} \quad \leftarrow \text{Discovery}$$

$$\tilde{q}_\mu = \begin{cases} -2 \ln \left( \frac{L(x; \mu, \hat{\Theta})}{L(x; 0, \hat{\Theta})} \right), & \tilde{\mu} < 0 \\ -2 \ln \left( \frac{L(x; \mu, \hat{\Theta})}{L(x; \mu, \hat{\Theta})} \right), & 0 \leq \tilde{\mu} \leq \mu \\ 0, & \tilde{\mu} > \mu \end{cases} \quad \leftarrow \text{Upper Limit}$$

# Look elsewhere effect

Sample from this



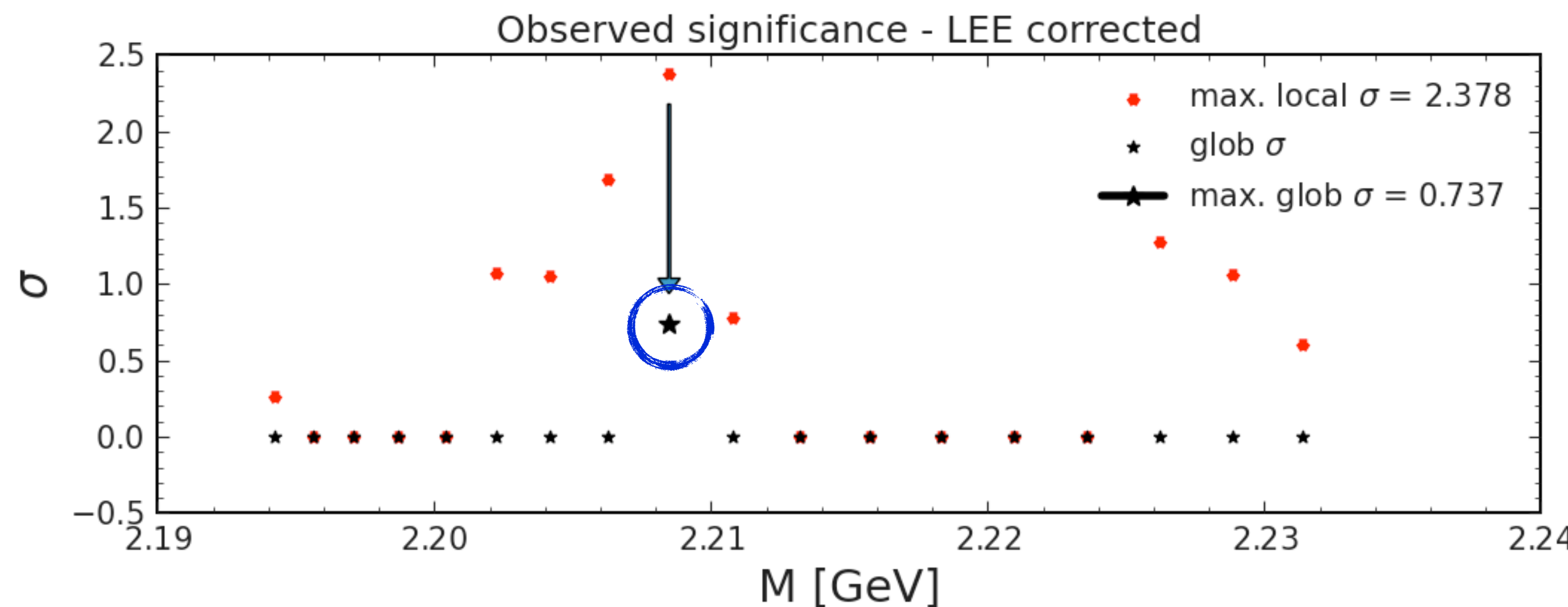
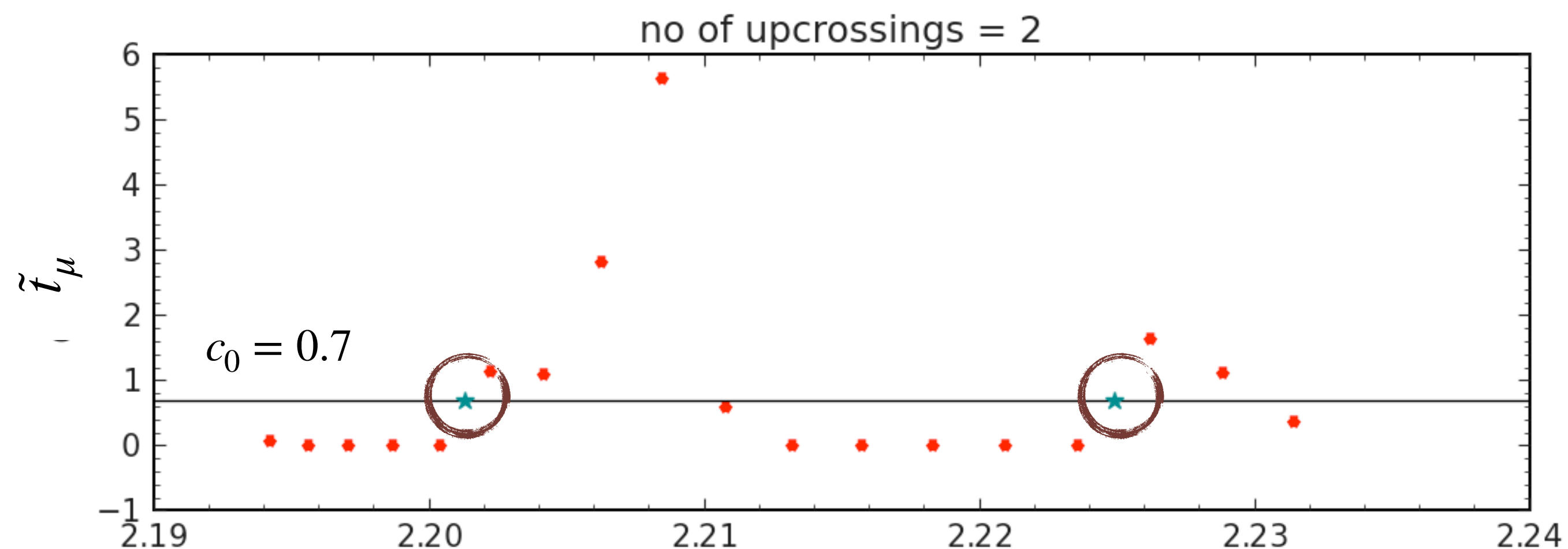
- Essentially:  
one can relate number of  
“upcrossings”  $N_{c_0}$  to  $\sigma_{global}$

- Count  $N_{c_0}$  from toy studies

- $c_0$  as low as possible

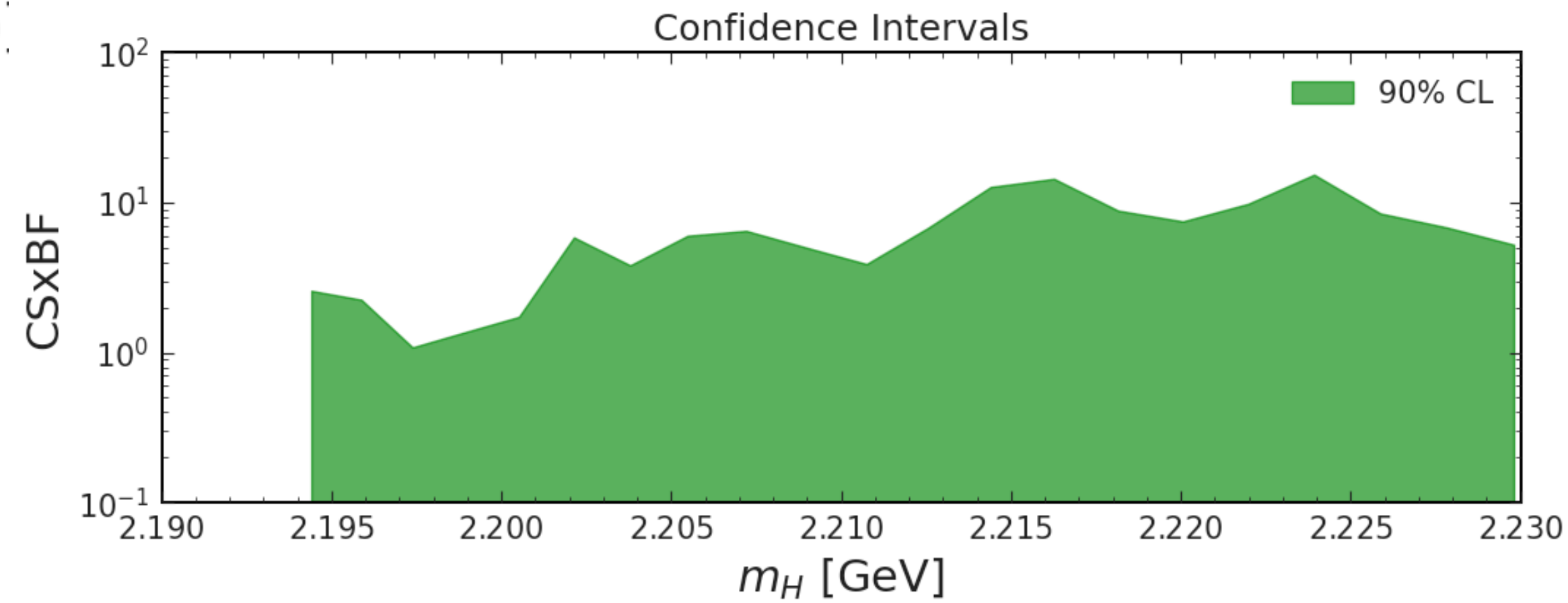
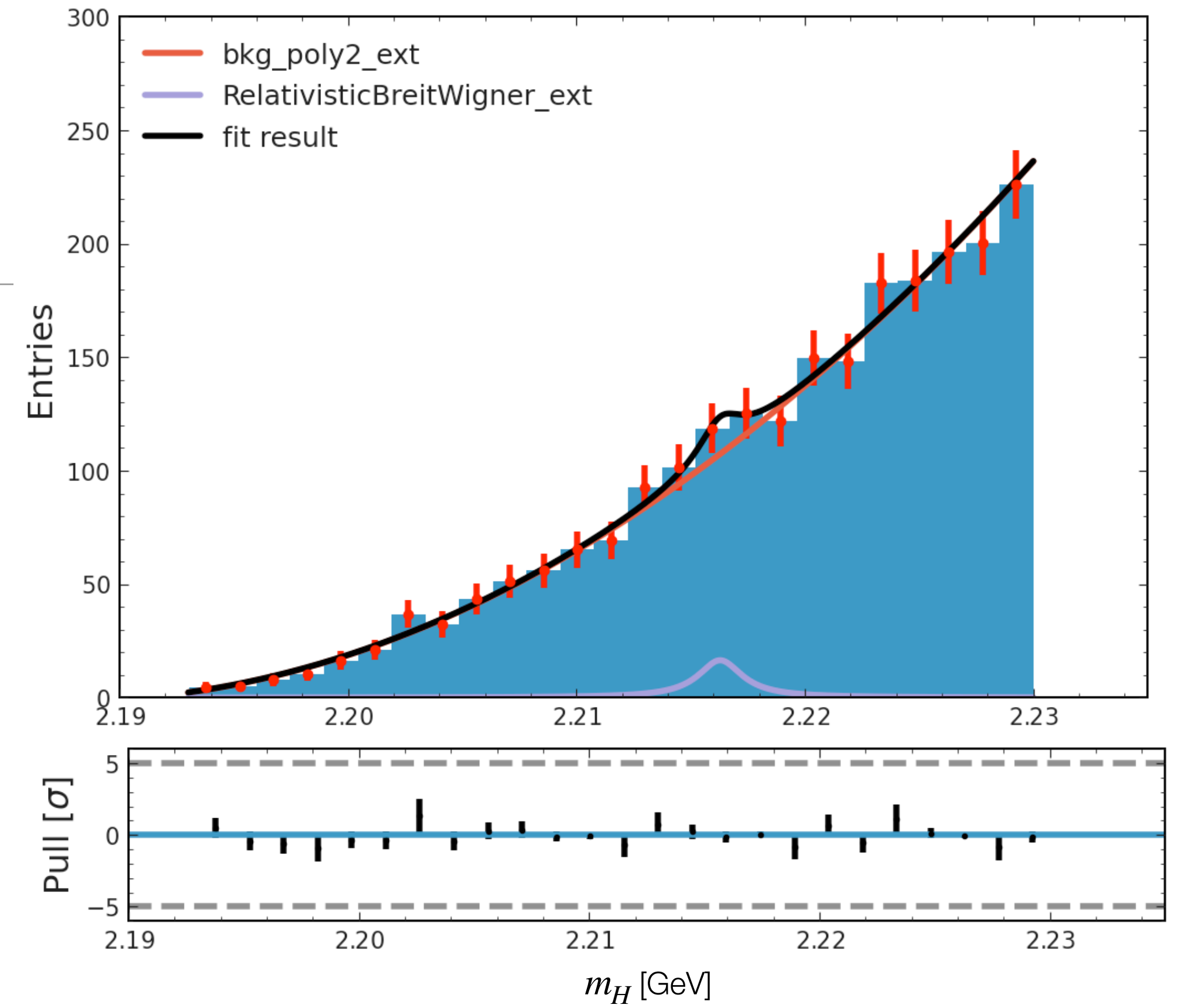
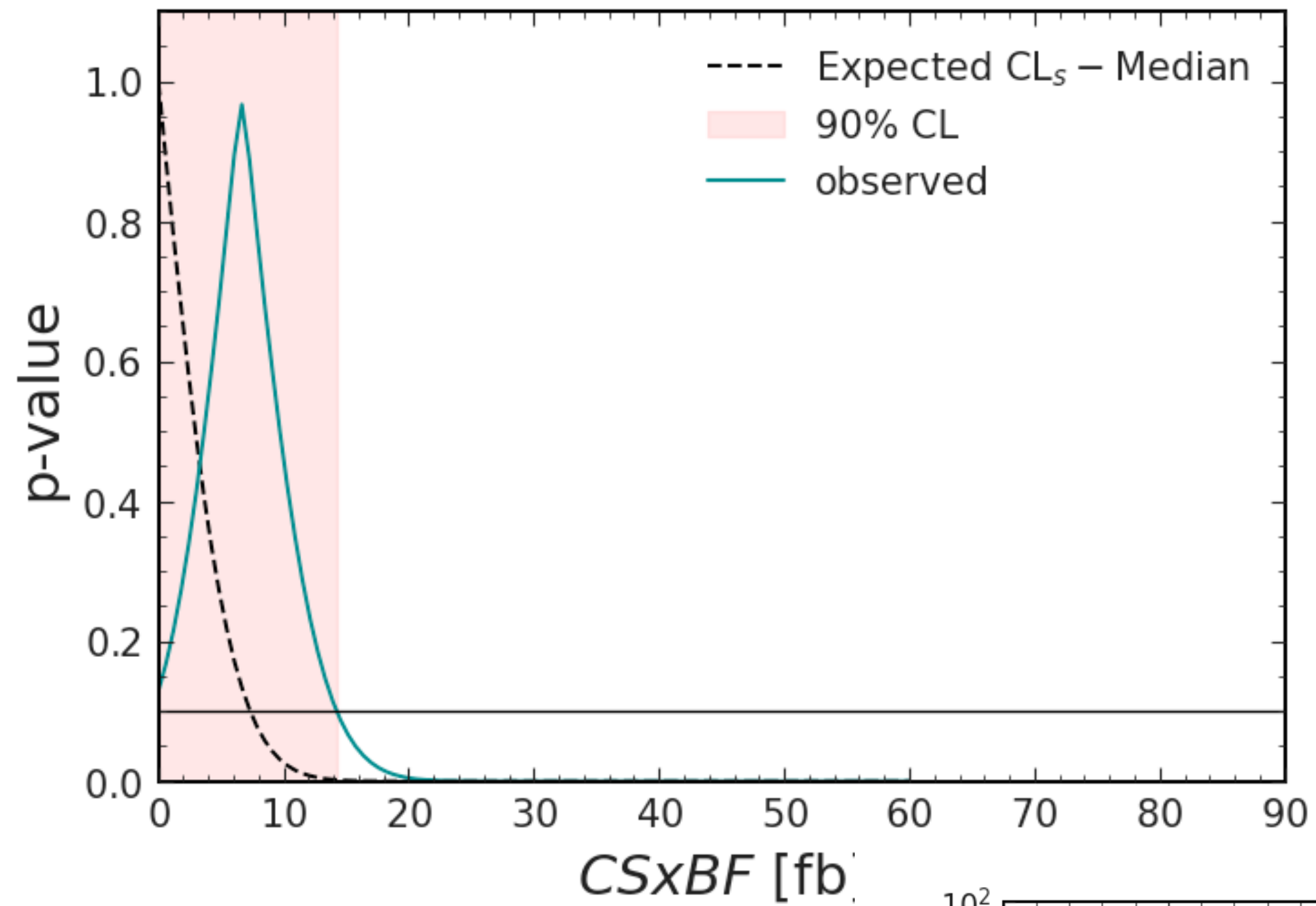
- Reasonable sample size

$$P_{global} \approx P_{local} + N_{c_0} e^{-\frac{u_{ref} - u}{2}}$$

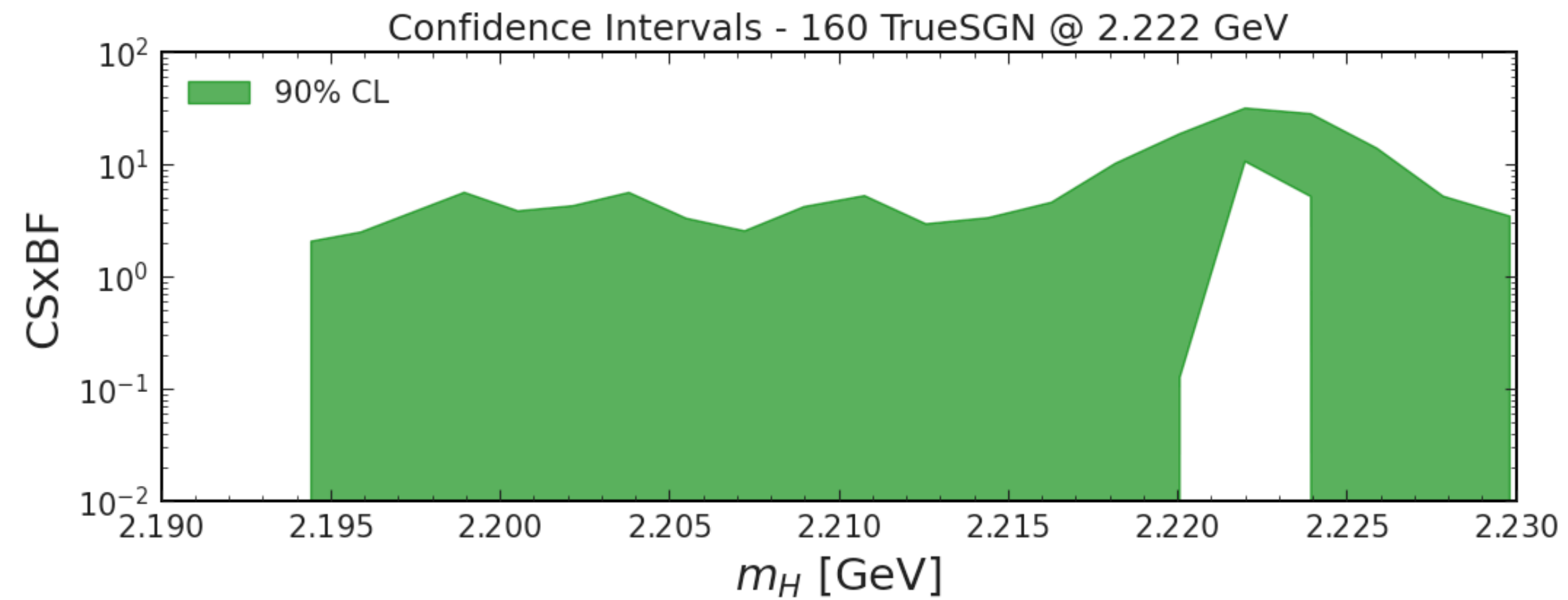
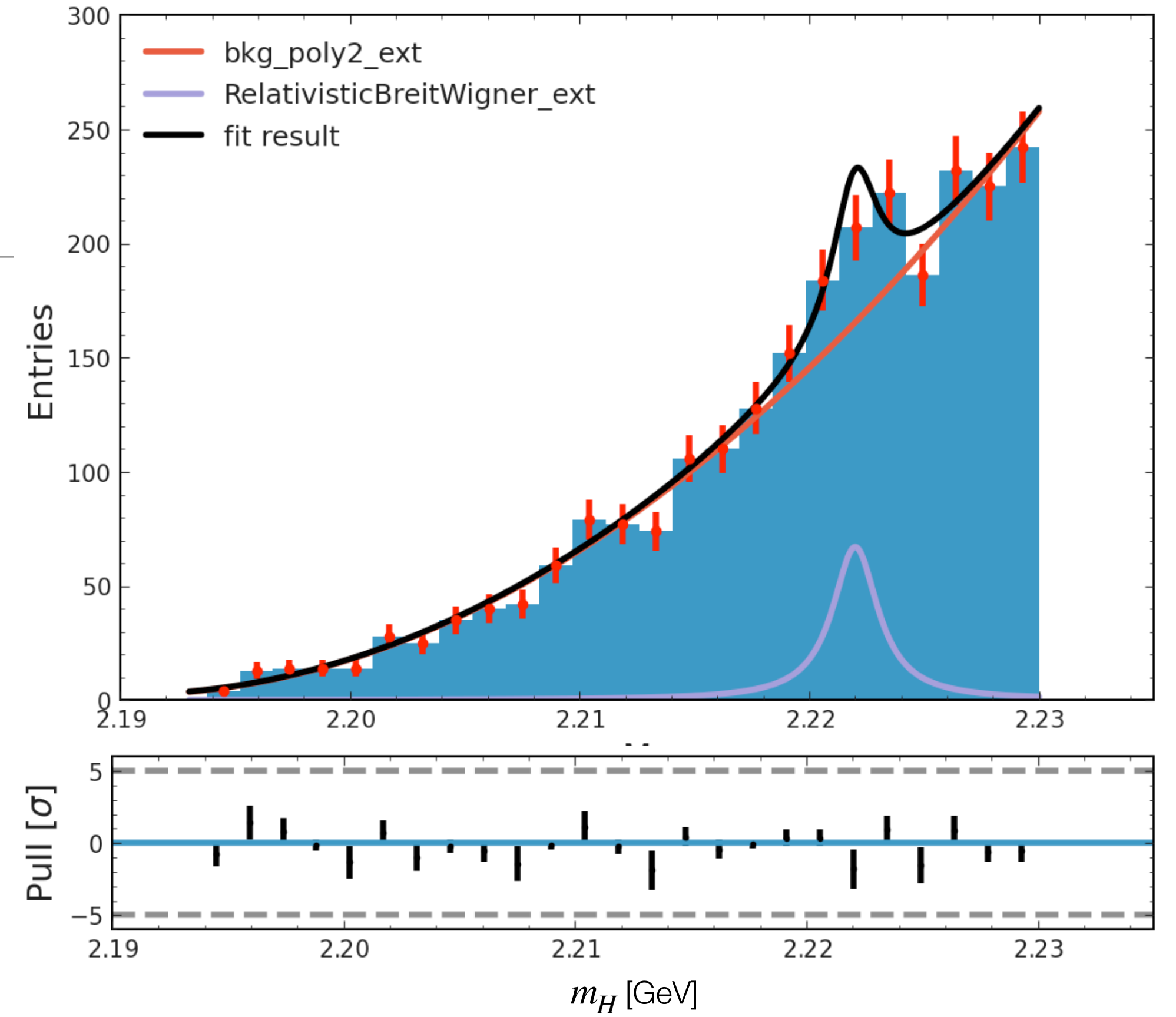
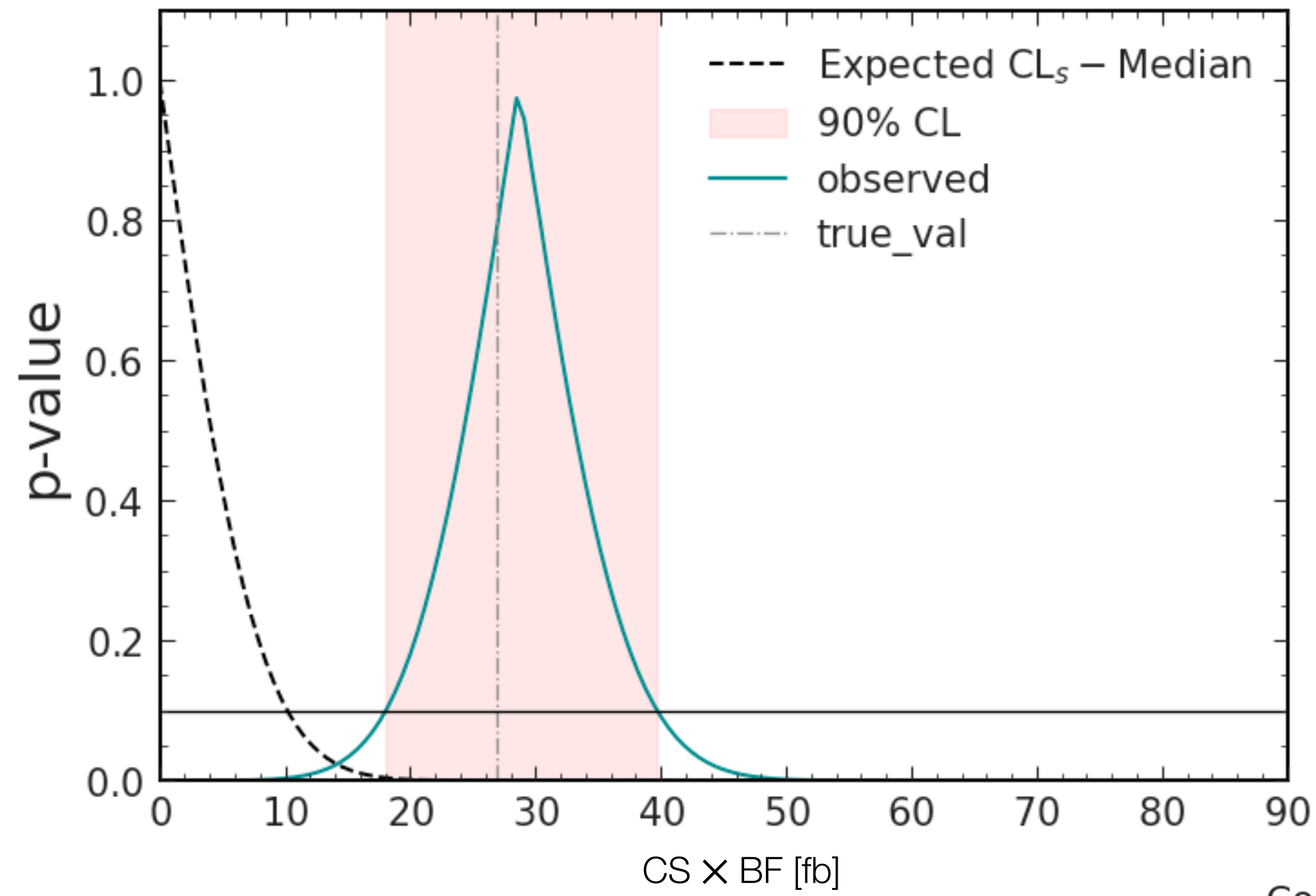


$$u = \sigma_{local,max}^2 \quad u_{ref} = 0.7^2 \quad N_{c_0} = 1.86$$

# Confidence Intervals - MC-scaled



# Confidence Intervals - SGN loaded



---

Conference End

Bormio 3000

Enjoy the view of the Alps



Ski the slopes



Get in lift to Bormio 3000 again

Have an espresso at the ski hut

