# Probing the Shadows:
# A Search for Dark Photons (10-40 GeV) Using Scouting Data

## Analysis review

Yusuf Seday
22/07/2024

Artwork Credit: DALL-E

# Table of Contents


1. Context and Analysis Strategy: Minimal Dark Photon Model
2. Event Selection and Efficiency Calculations
3. Signal and Background Modelling
4. Systematic Uncertainties
5. Expected Exclusion Limits
6. Conclusions

# 1) Minimal Dark Photon Model

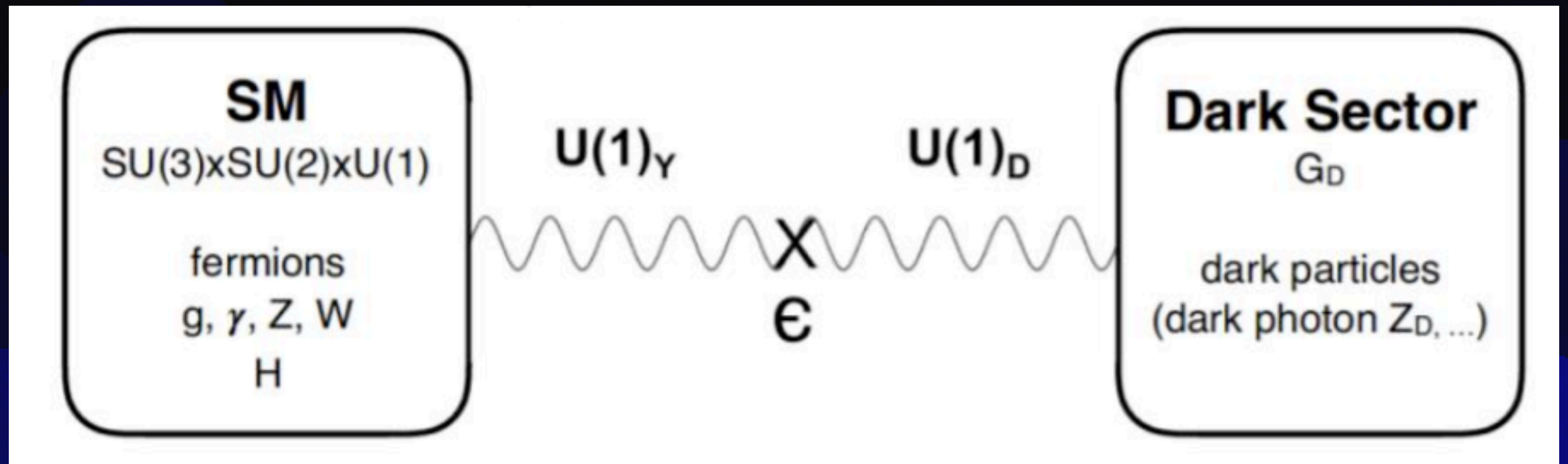
- **Minimal Dark Photon Model:**
  - Extends the Standard Model (SM) with a new U(1) gauge symmetry in the dark sector.
  - Introduces a dark photon (Z') that interacts with SM particles via kinetic mixing.

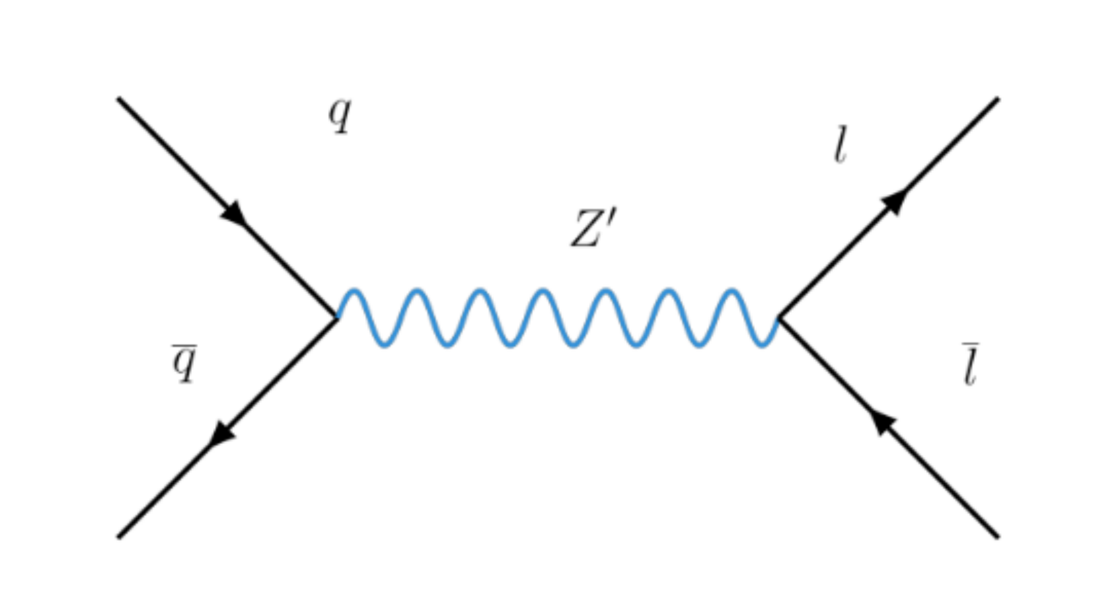
- **Kinetic Mixing:**
  - Lagrangian term: $\mathscr{L} \supset -\frac{\epsilon}{2} F_{\mu\nu} F'^{\mu\nu}$
  - $\epsilon$ is the kinetic mixing parameter allowing dark photon to couple to SM particles.

- **Free Parameters:**
  - Kinetic mixing parameter $\epsilon$.
  - Dark photon mass $m_{Z'}$
  - The decay branching fraction of the dark photon into invisible dark-sector final states, typically assumed to be either unity or zero (corresponding to whether any invisible dark-sector final states are kinematically allowed or not)†

- **Experimental Signatures:**
  - Dark photons produced in collisions, decay into SM particles (e.g., $\mu^+\mu^-$).
  - Feynman diagram: dark photon (Z') mediating between quarks and leptons.

- **Motivation:**
  - Explains astrophysical and cosmological observations suggesting dark matter.
  - Dark photons as candidates for mediating interactions between dark matter and SM.



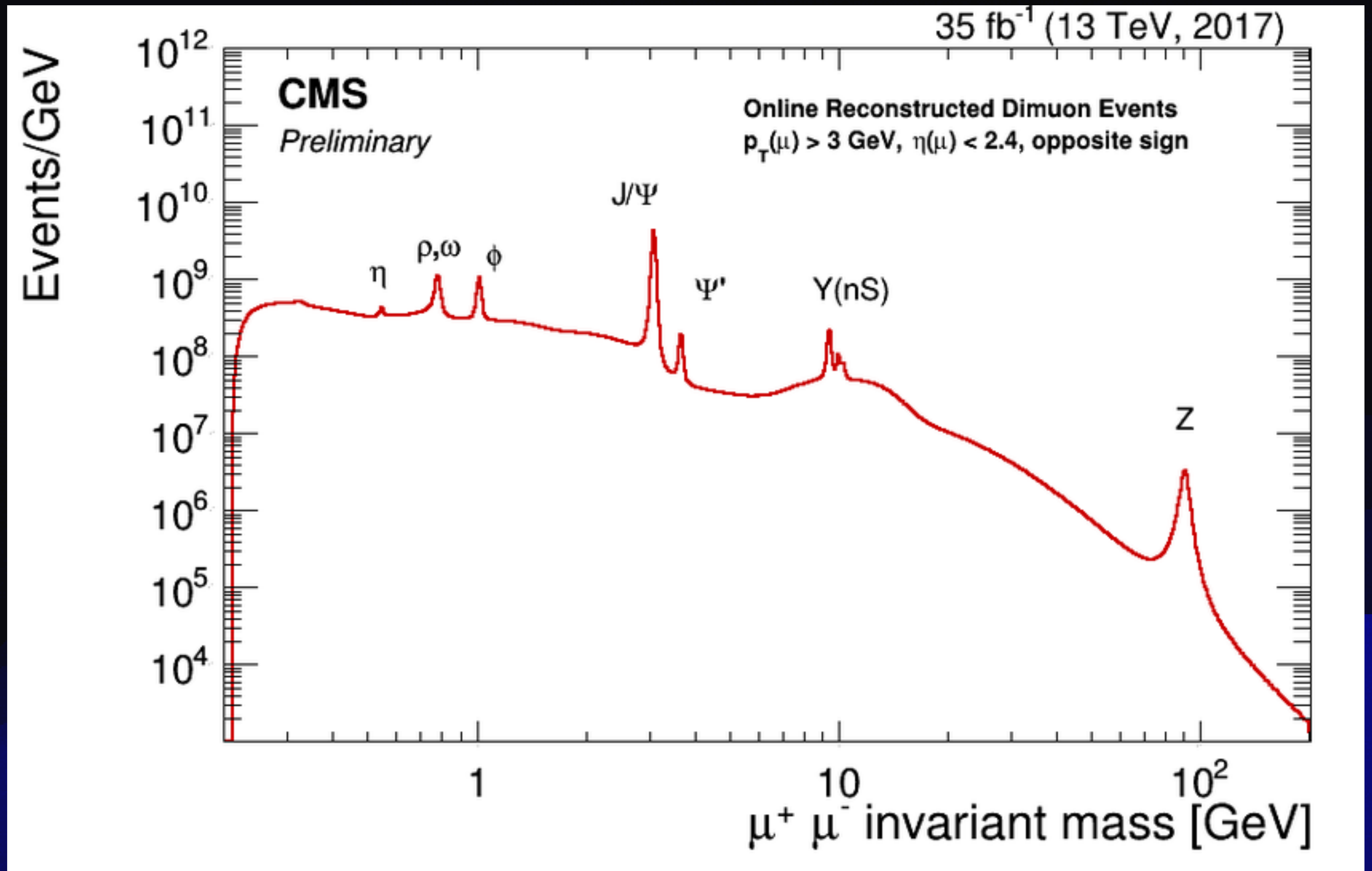



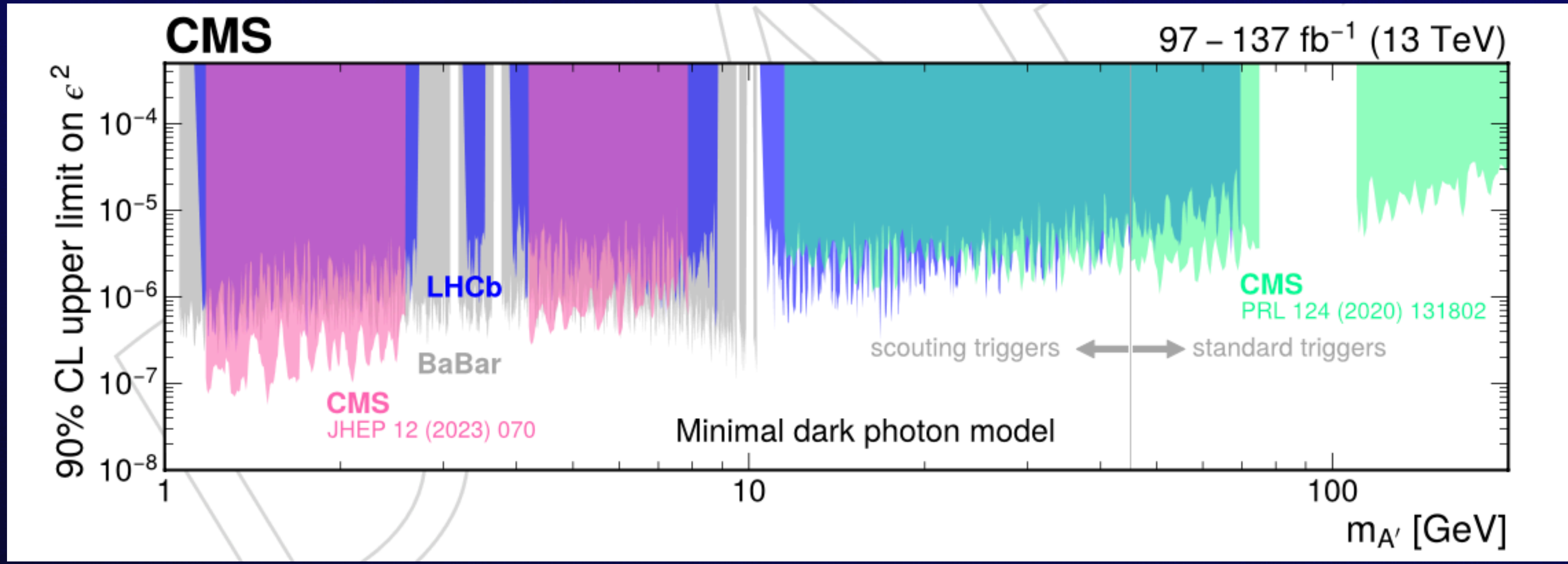


† https://arxiv.org/pdf/2104.10280

# Dimuon Resonances

- Discovery of many new particles through the resonant particle pair production in dimuon channel
- Search for a narrow dimuon resonance at low mass using scouting data recorded by the CMS
- Study the dimuon final states to test the minimal Dark Photon model
- Most recent results for the observed upper limits on the square of the kinetic mixing coefficient ε




*Observed upper limits*
**CMS PAPER EXO-23-005**

# Analysis Strategy

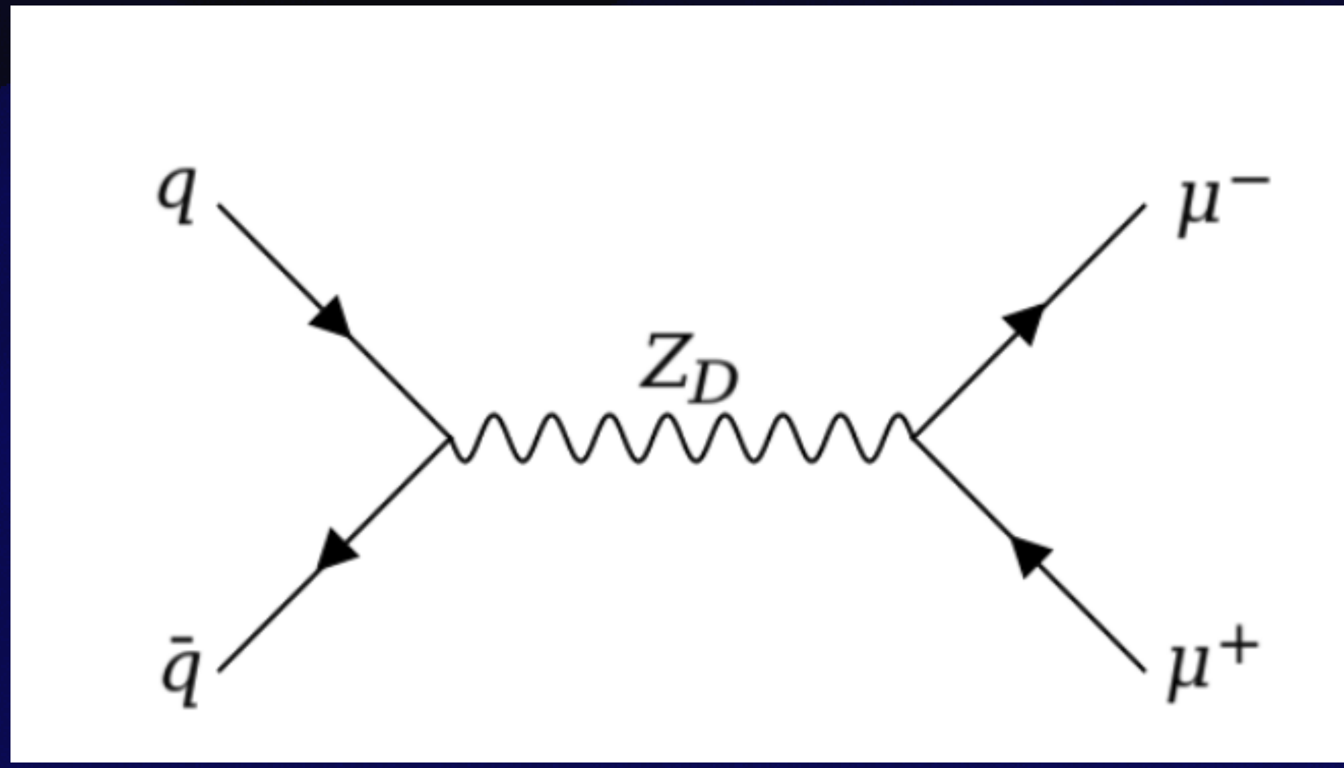
- Investigating the existence of Dark Photons, a potential BSM mediator, within the 10-40 GeV range, through the decay into oppositely charged muon pairs, utilising scouting data from the CMS experiment

- Optimising event selection for (prompt) dimuon resonance signals and efficiency calculations

- Searching for a bump in the dimuon mass spectrum using analytical signal and background Pdfs

- Study systematic uncertainties

- Establish model-independent limits for the cross-section of low-mass dimuon resonant states




An expected production channel of a Dark Potion
https://arxiv.org/pdf/2309.16003.pdf

# Employed Data

## The datasets that are used in this analysis:

- **Observed Data**
  - LHC Run 3, CMS Scouting Data:  /ScoutingPFMonitor/Run202*/RAW
- **MC: Upsilon Samples**
  - Upsilonto2Mu_UpsilonFilter_2MuFilter_TuneCP5_13p6TeV_pythia8 dataset
  - Used to compare efficiencies around the Υ region
- **MC: DY Samples**
  - Privately produced samples
  - Used for efficiency calculations

# 2) Event Selection and Efficiency Calculations

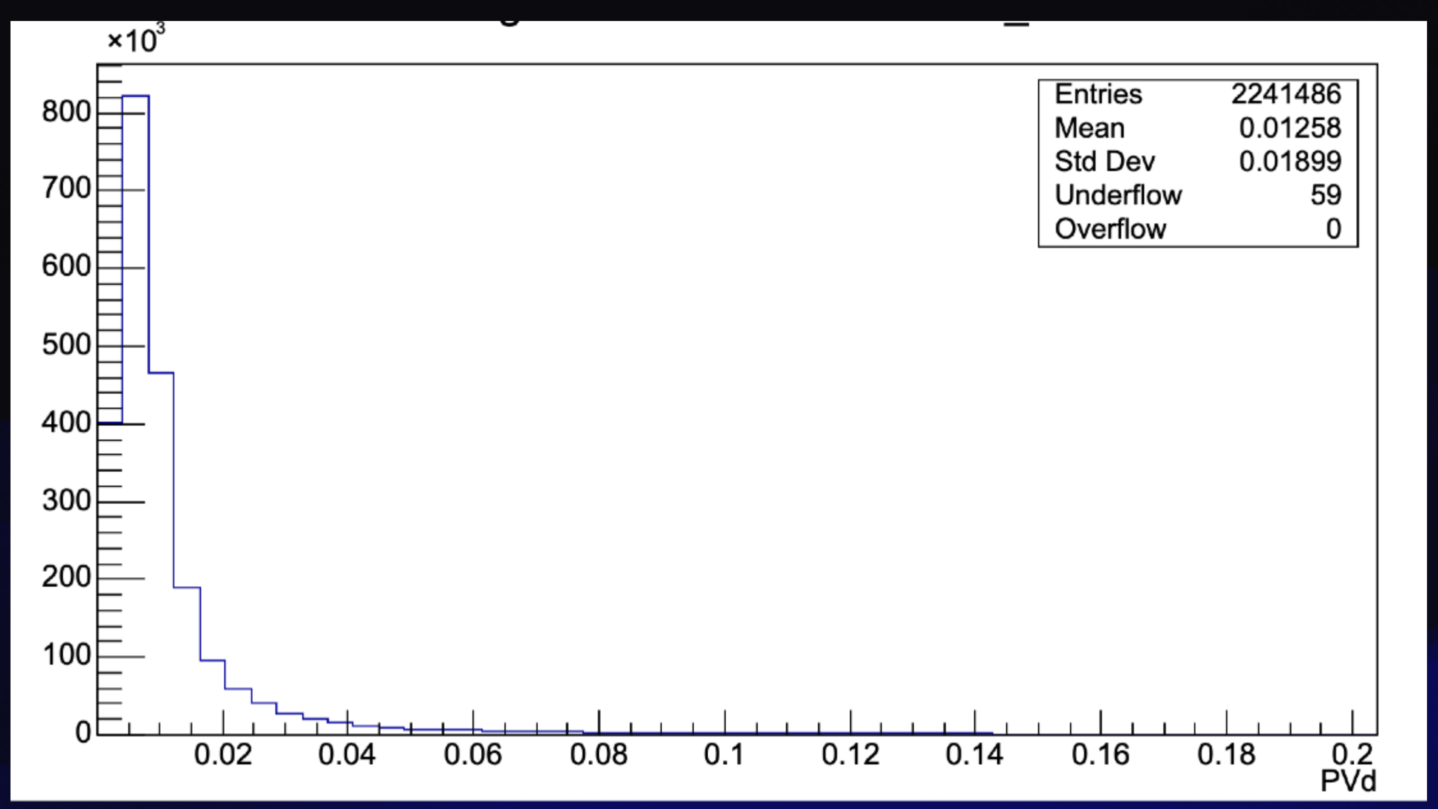
# Event Preselection

- Muon pair with opposite charges as the final state

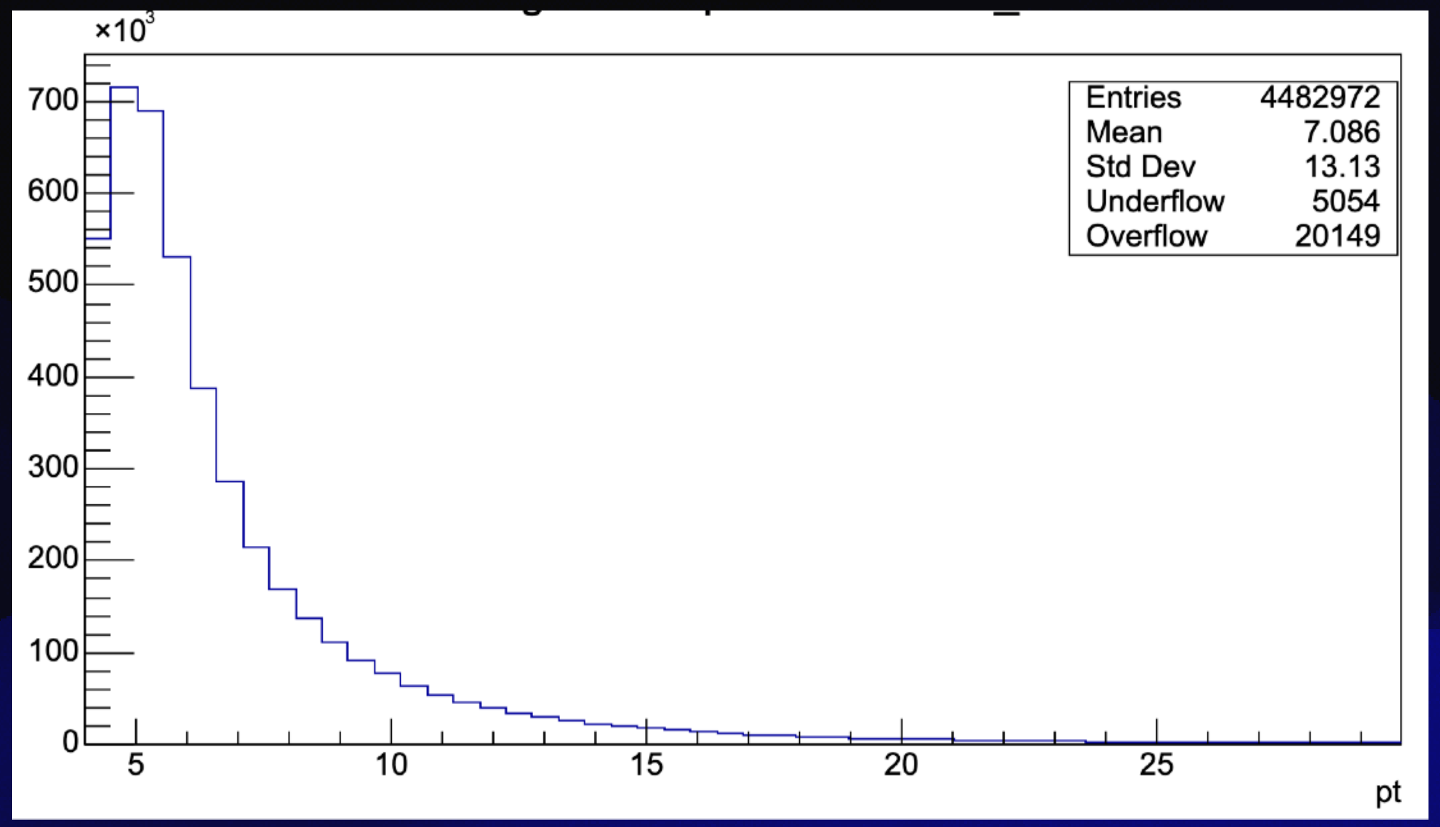
- Prompt Production
  Transverse Displacement:
  L < 0.2 cm

- Transverse Momentum:
  $p_t$ > 4 GeV

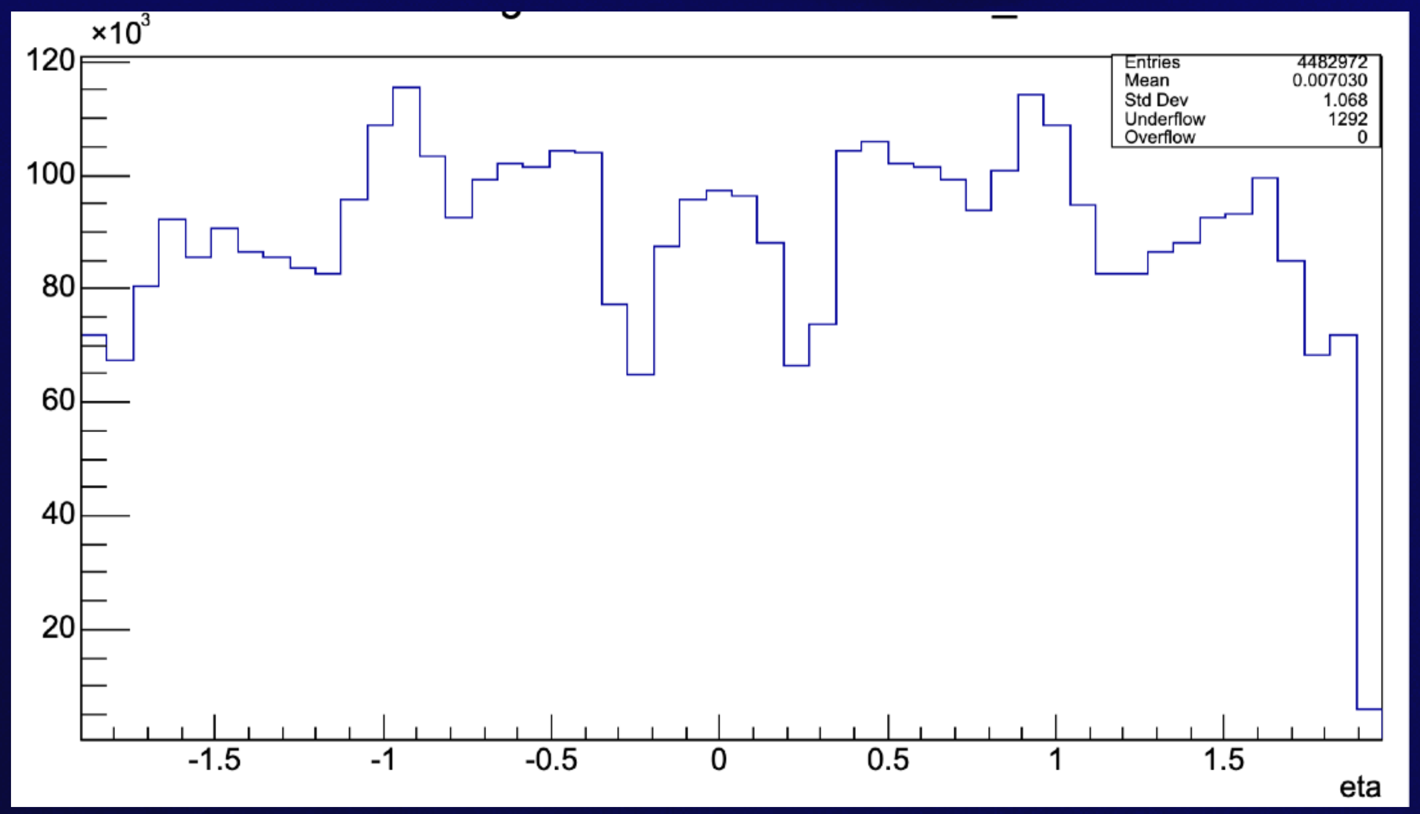
- Pseudorapidity: $|\eta| < 1.9$




*PVd Distribution*




*$p_t$ Distribution*




*$\eta$ Distribution*

# Optimising the Event Selection

- Focus: Optimisation of event selection within the 10-40 GeV mass range.
- Strategy: Employing the Upsilon (Υ) resonance as benchmark for optimising event selection through neural network training and efficiency analysis of simple cuts.
- Goal: Enhancing signal detection sensitivity for prompt dimuon events.
- Approach: Comparison of neural network performance against traditional cut-based methods and refining the parameters for each case.

# Approach 1. Employing Neural Networks

i. Optimising the signal mass window

ii. Choosing the variables for NN training

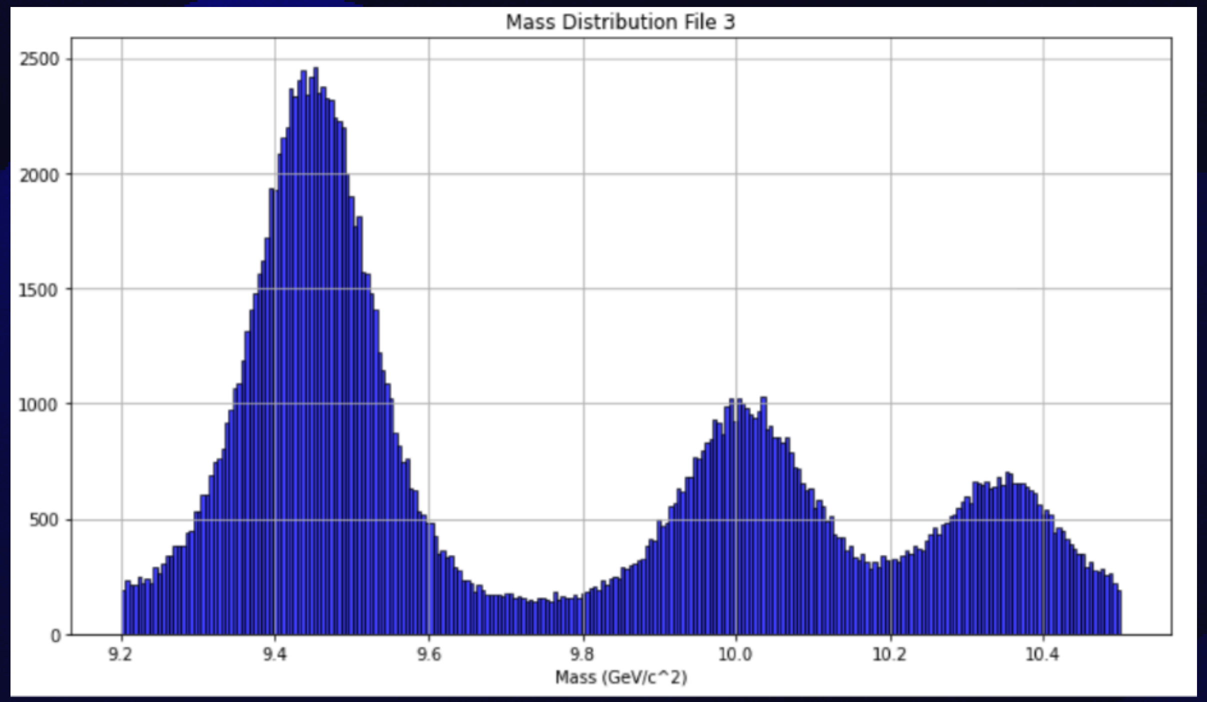
iii. Choosing the optimiser algorithm

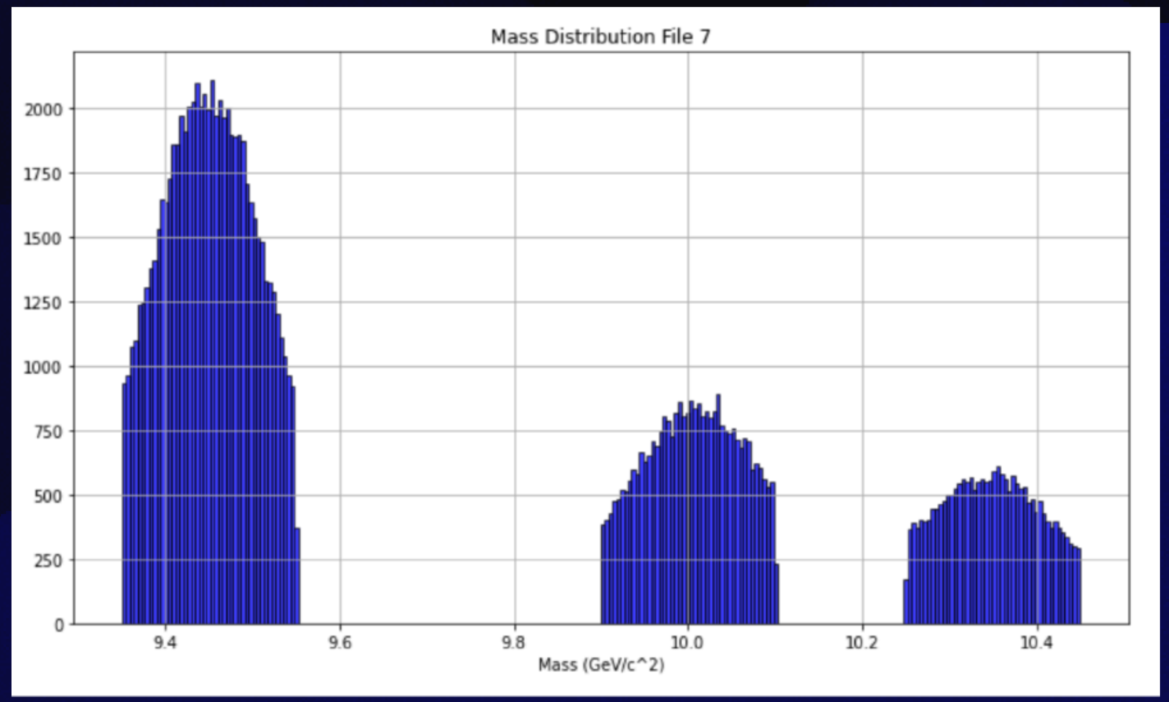
iV. MVA analysis and MVA cut

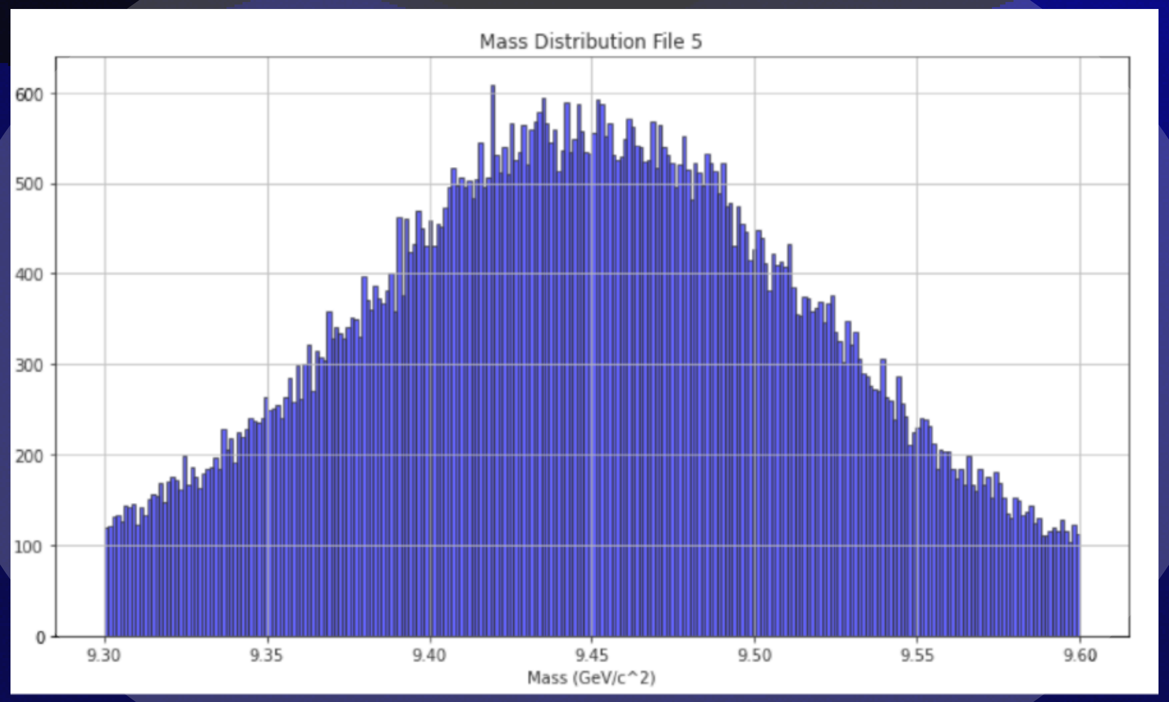
*Only for the MuonID*

# i) Optimising the Signal Mass Window

- Use all the candidate variables (will be optimised in the next slide)
- Deploy several mass windows as signal region for the training and compare the ROC curves
- Use AUC for comparison



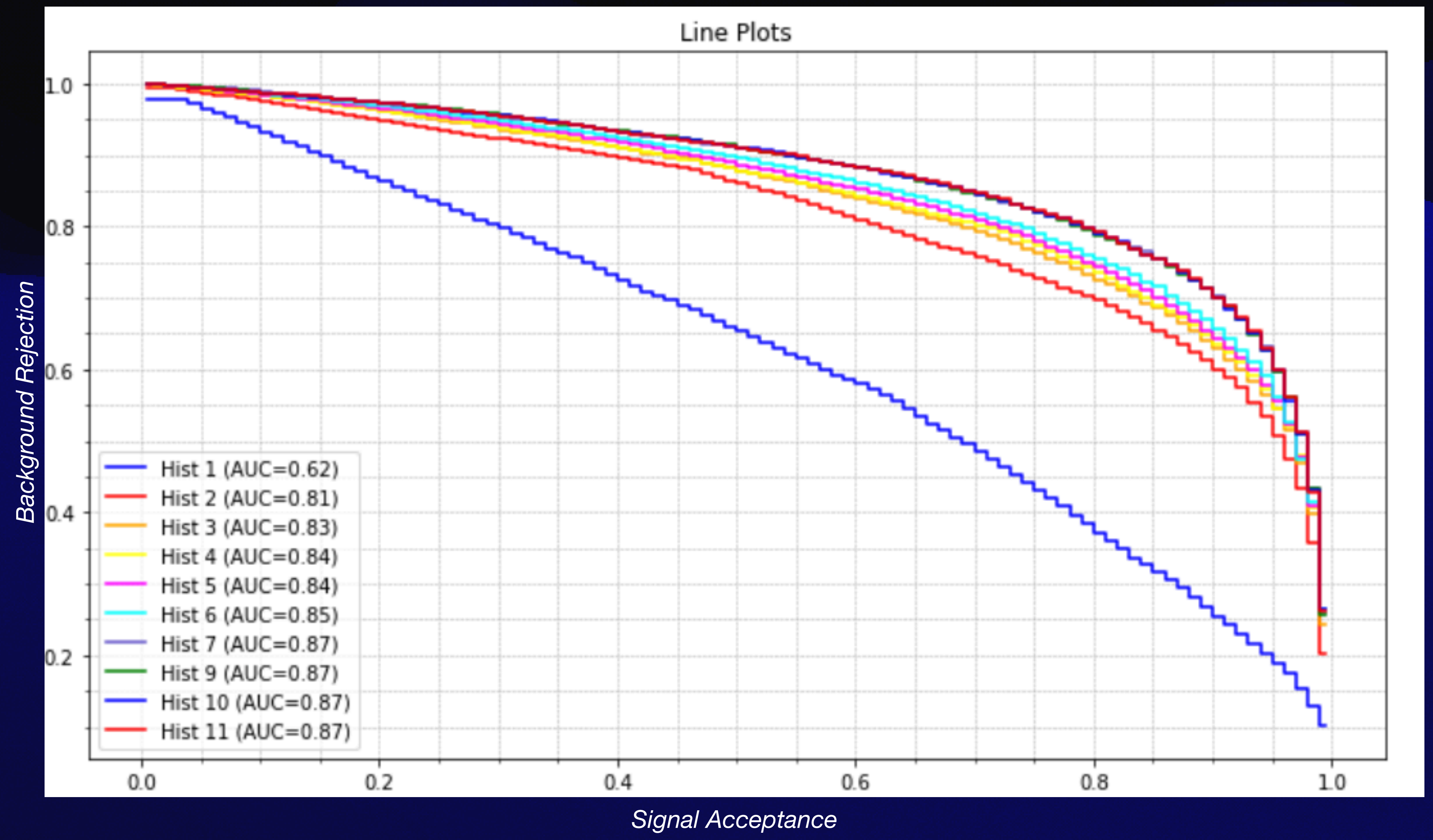

**The best response:**

| | |
|---|---|
| Mass Region | 9.3 - 9.6 GeV |
| AUC | 0.87 |
| Signal Contamination | 84% |

# ii) Choosing the Training Variables

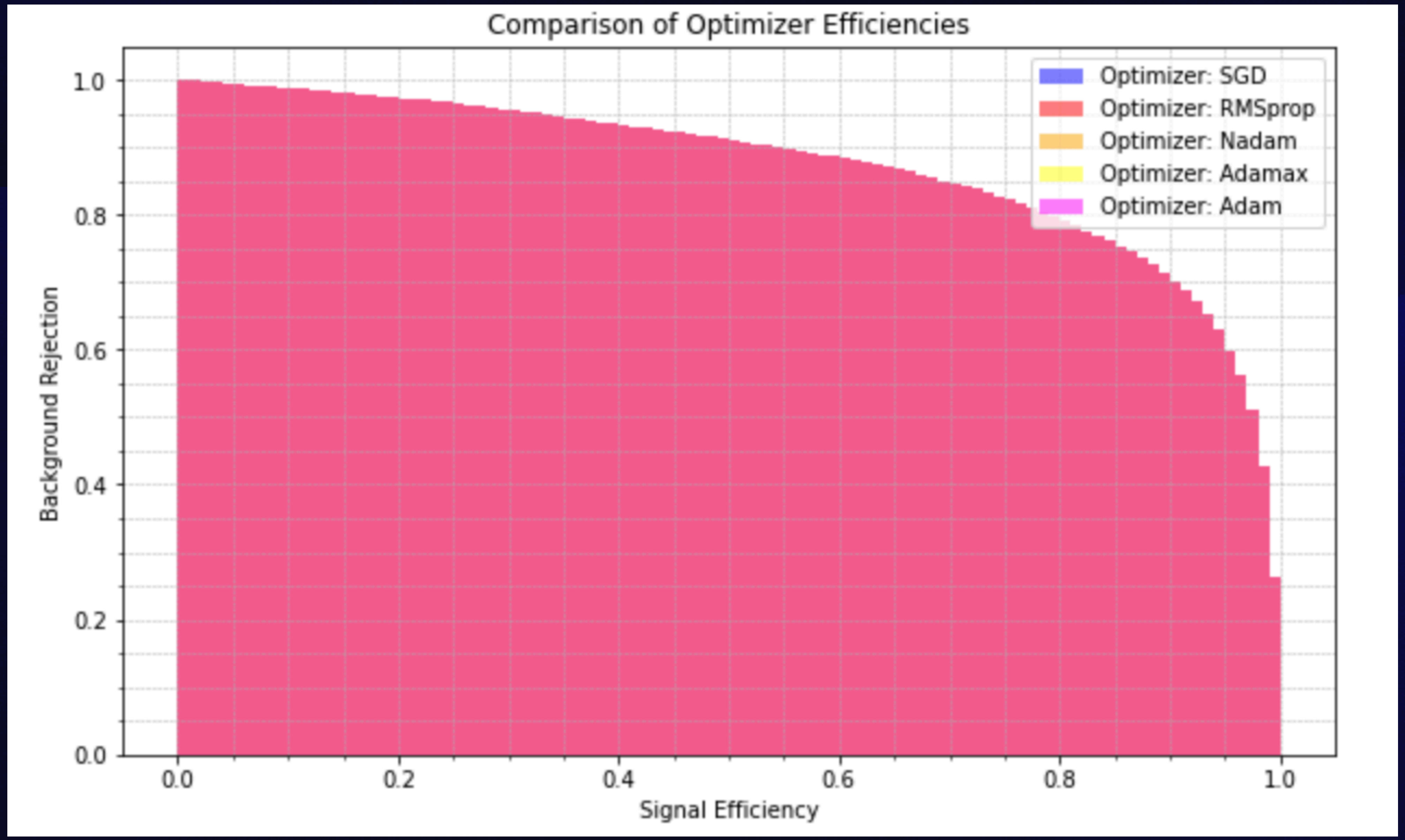
- Train with all the variables
- Modify the variable order according to importance
- Train with 1 variable and add the next one, repeat cumulatively
- Decide which variables to use

Used variables: "nmhits", "trkiso", "trkqoverp", "trklambda", "dxy", "ntklayers", "eta", "chi", "nphits"



# iii) Choosing the Optimiser Algorithm

- Deploy several optimiser algorithms (SGD, Adam, Nadam…)
- Compare the ROC curves



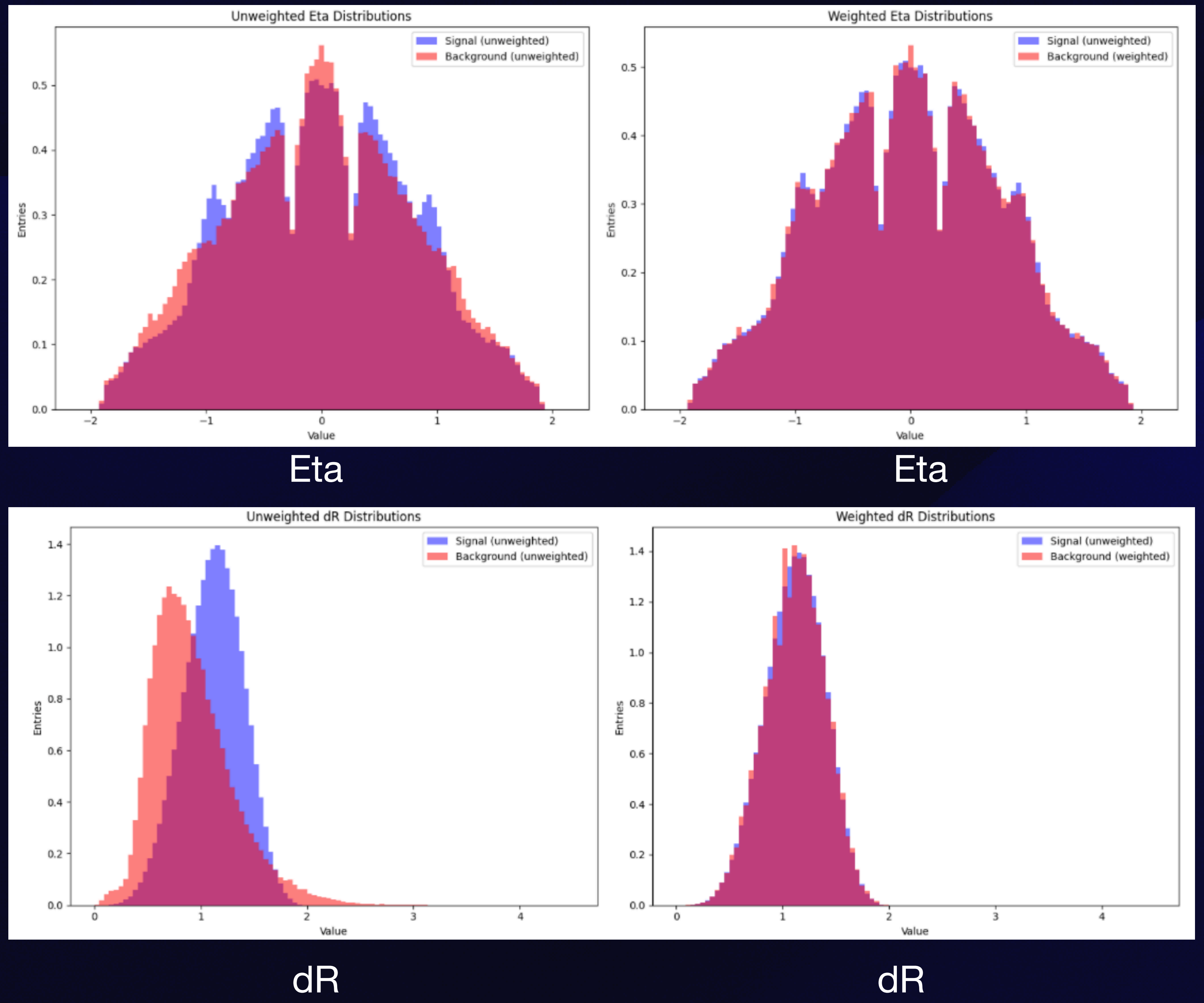

→

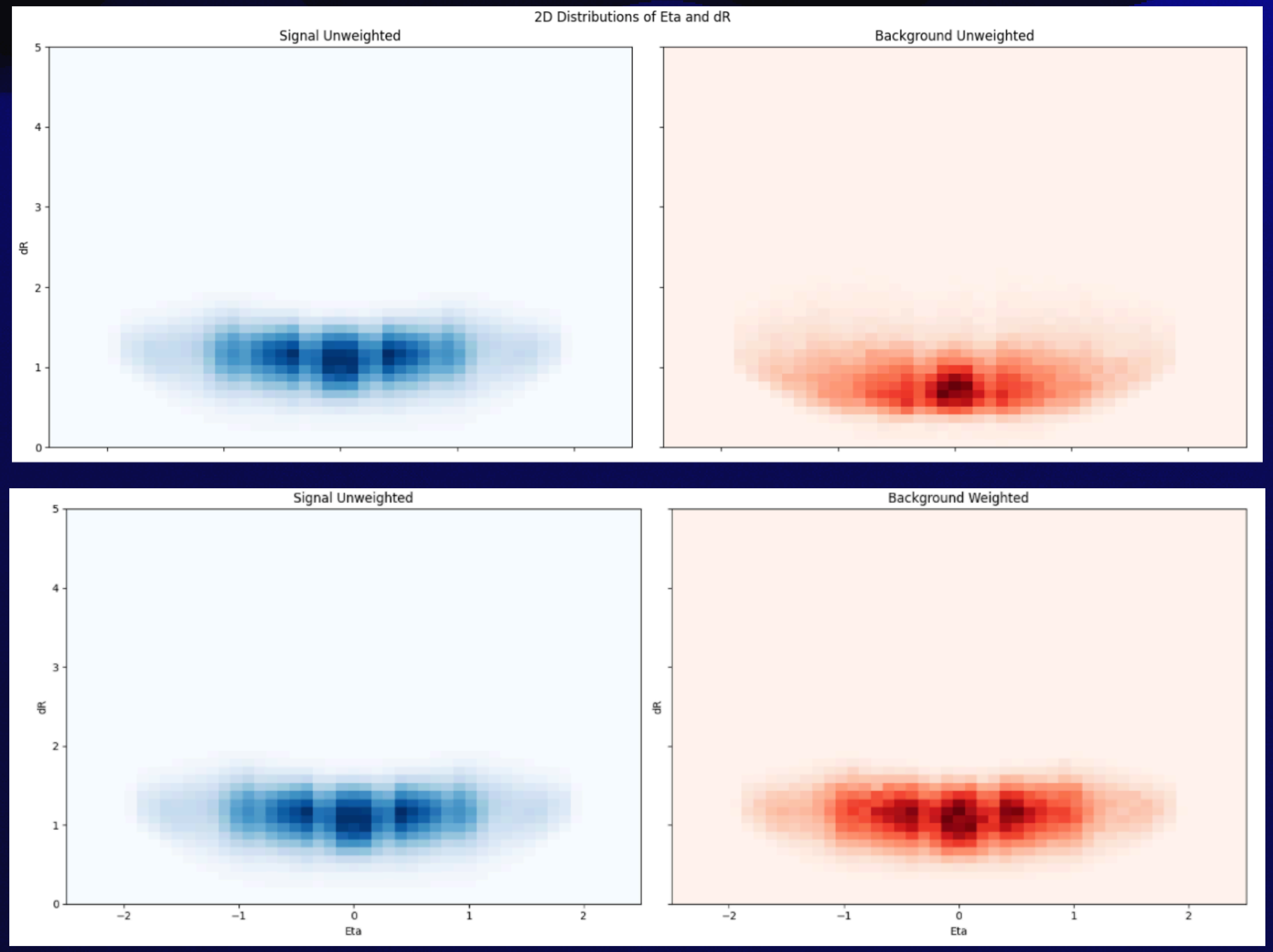
Almost same AUCs,
no visible improvement

Stick with SGD

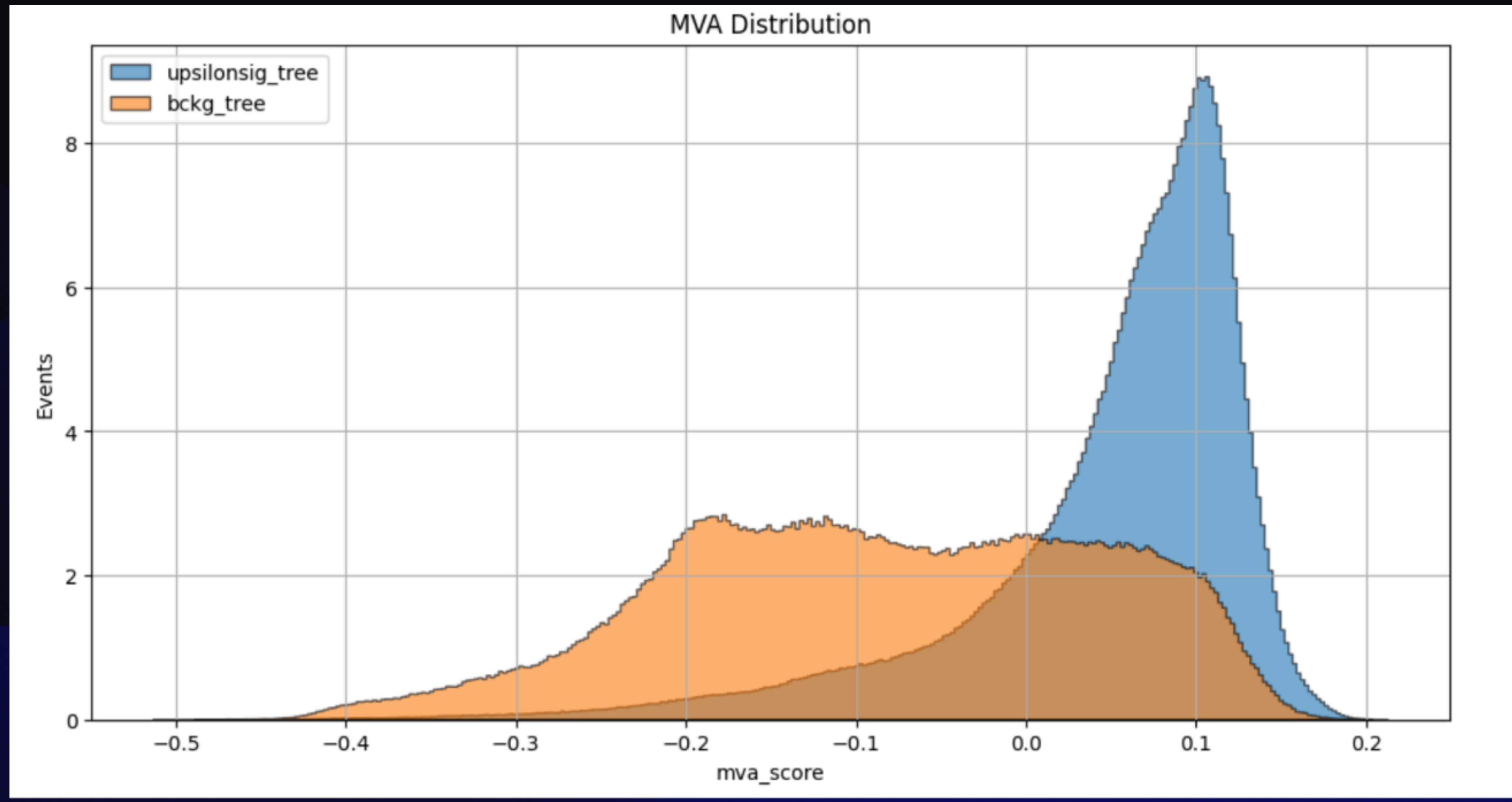
# iv) MVA Analysis and MVA Cut

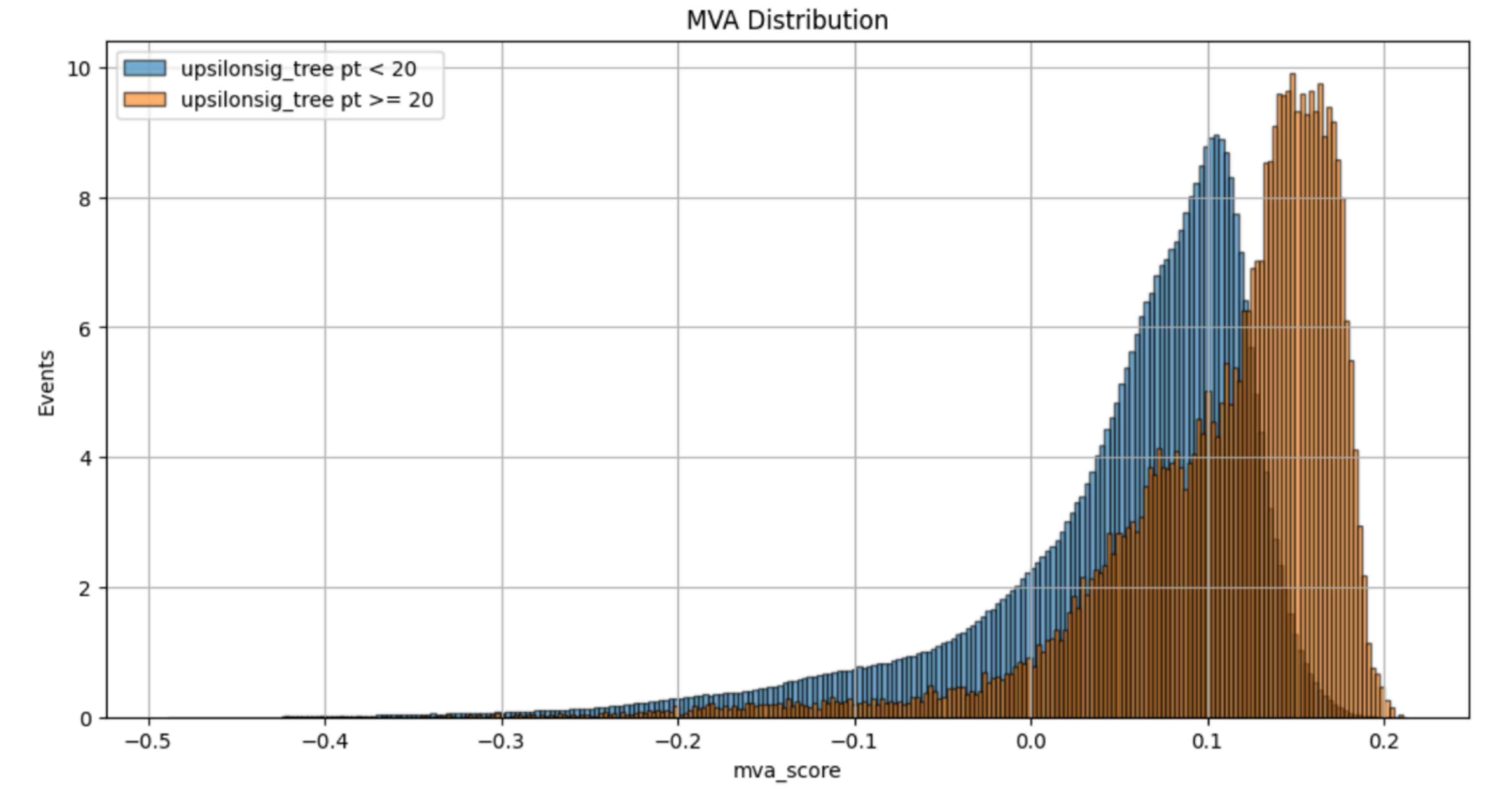
- **Re-weigh data to balance background and signal to prevent model bias in distinguishing features**



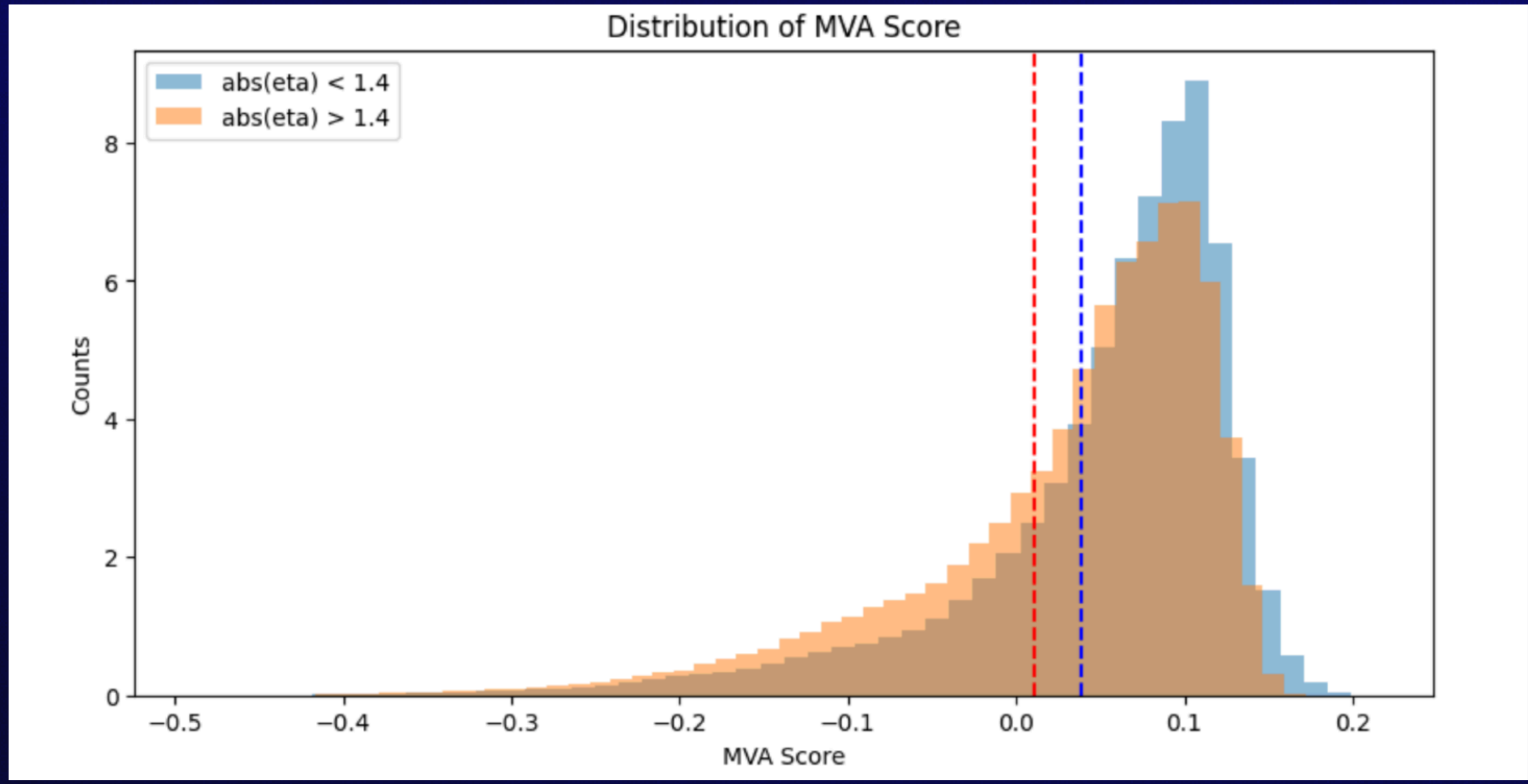



- Add MVA using the best model's weights



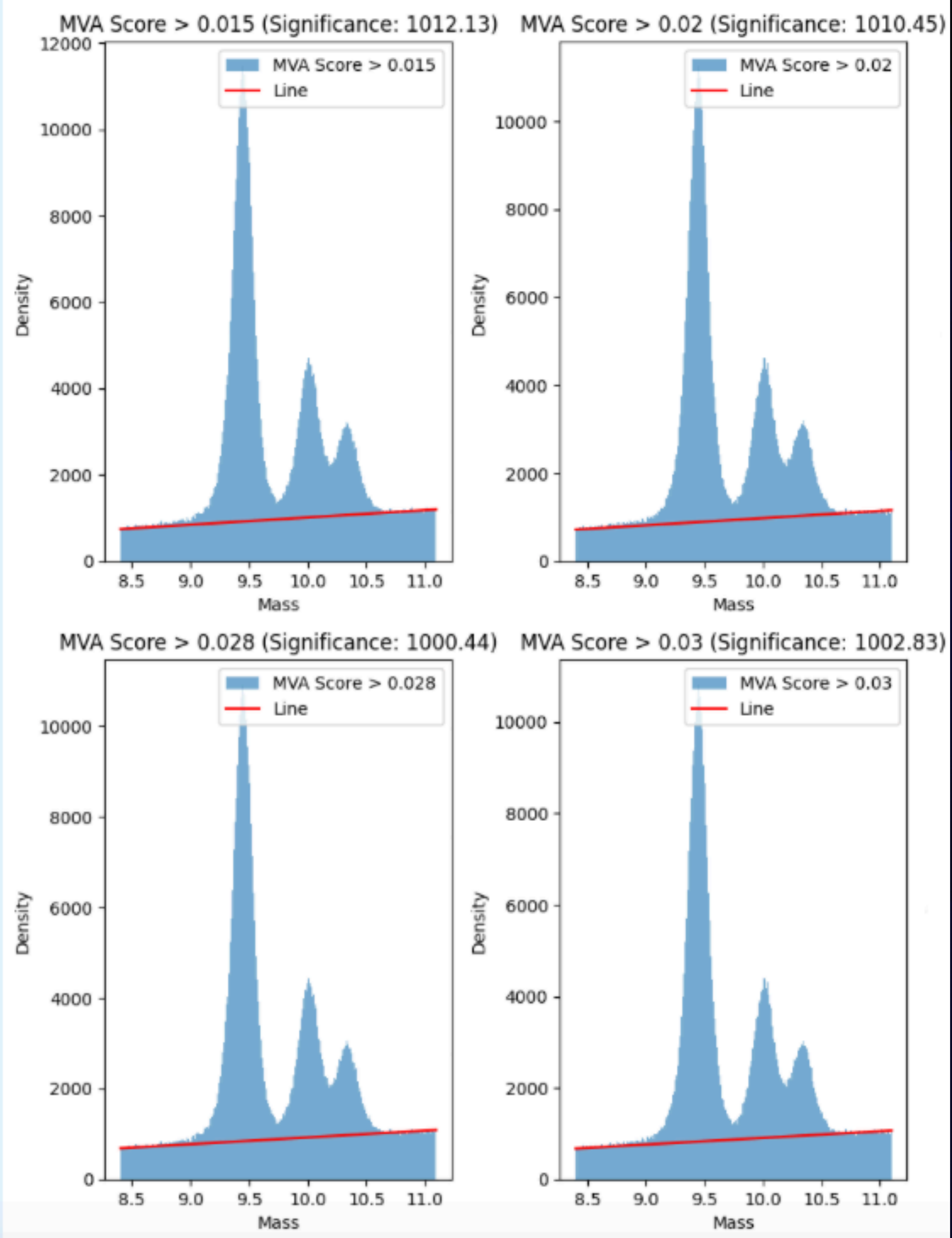



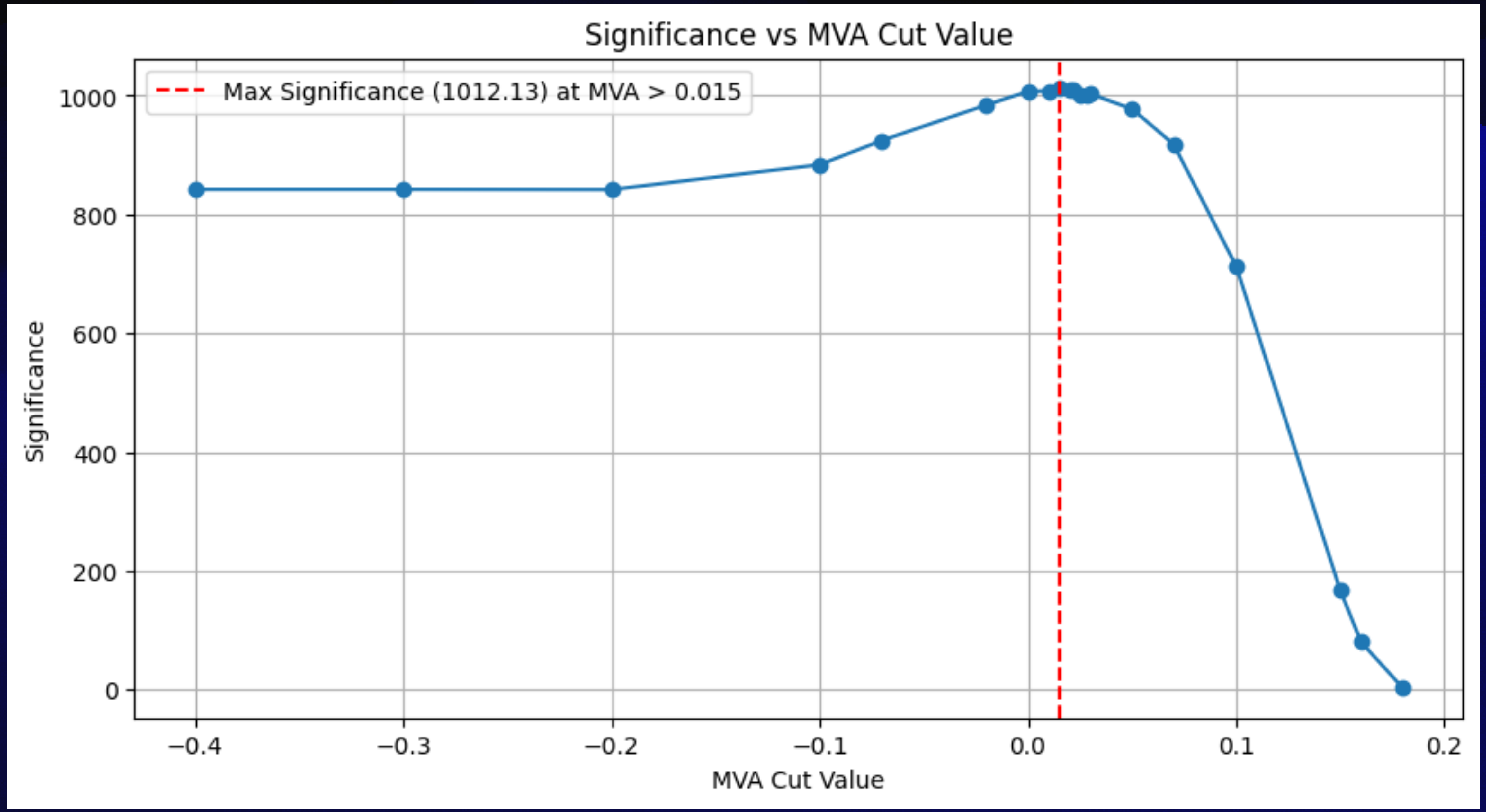

As expected




Previously this was a problem

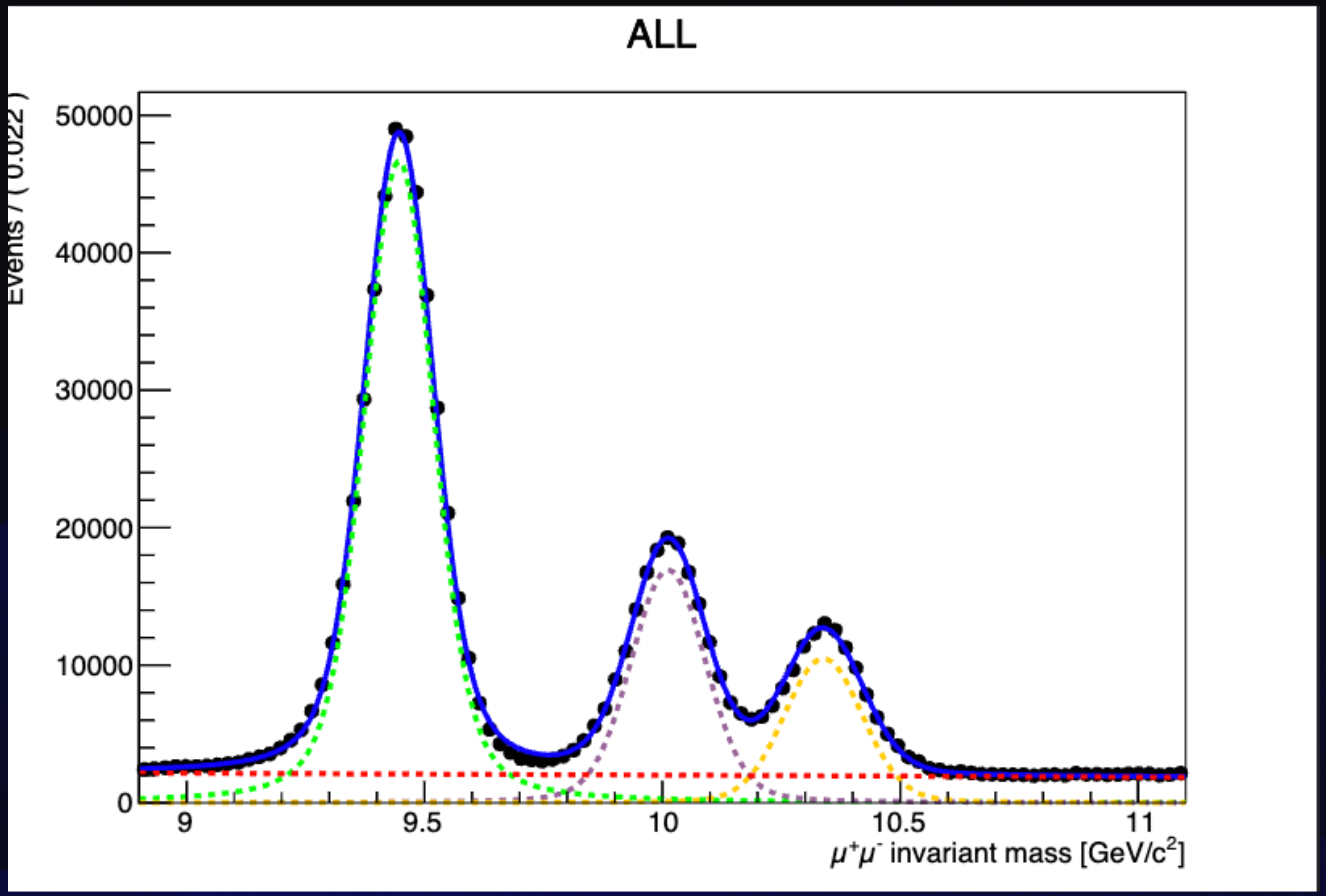
- Optimise the significance

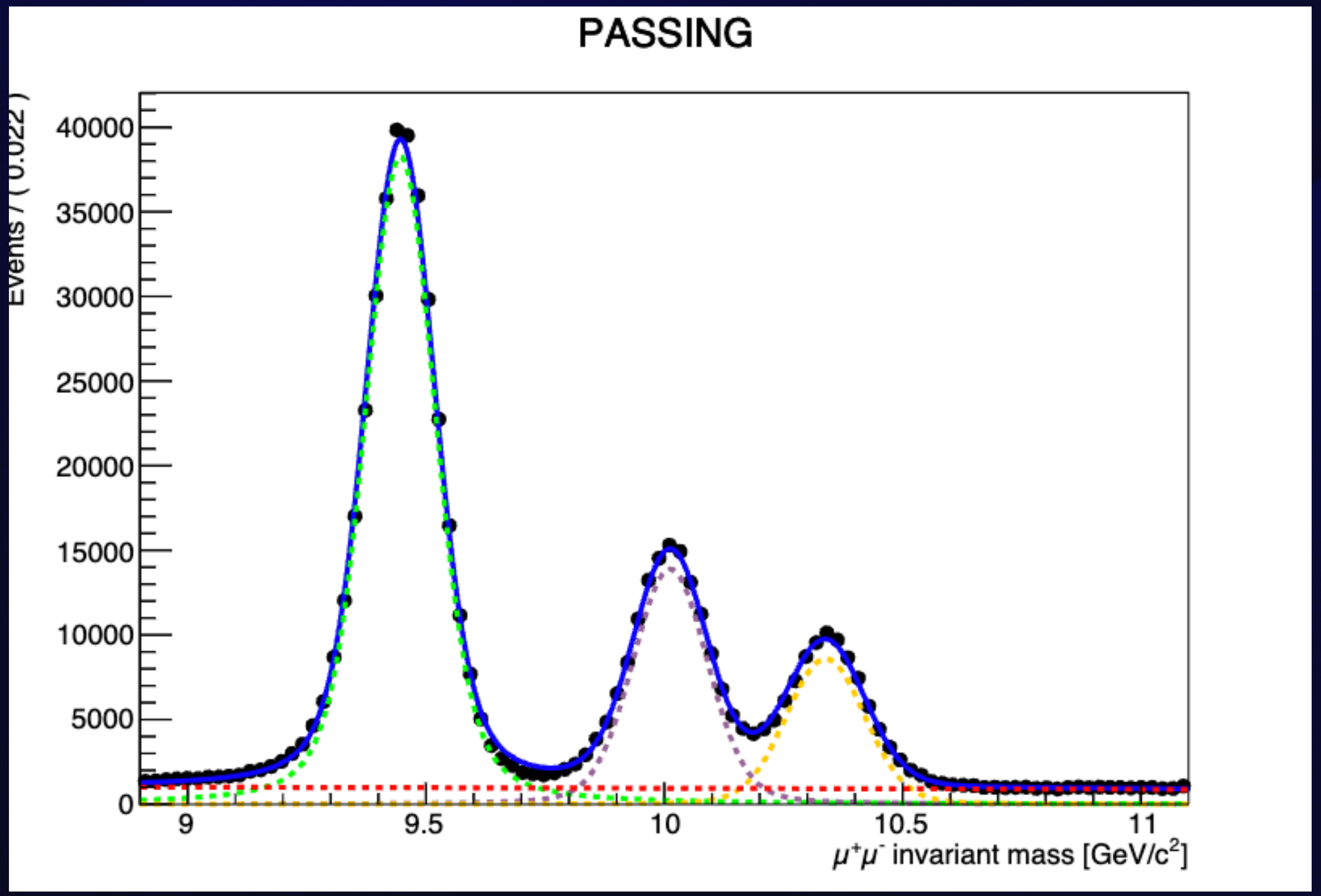


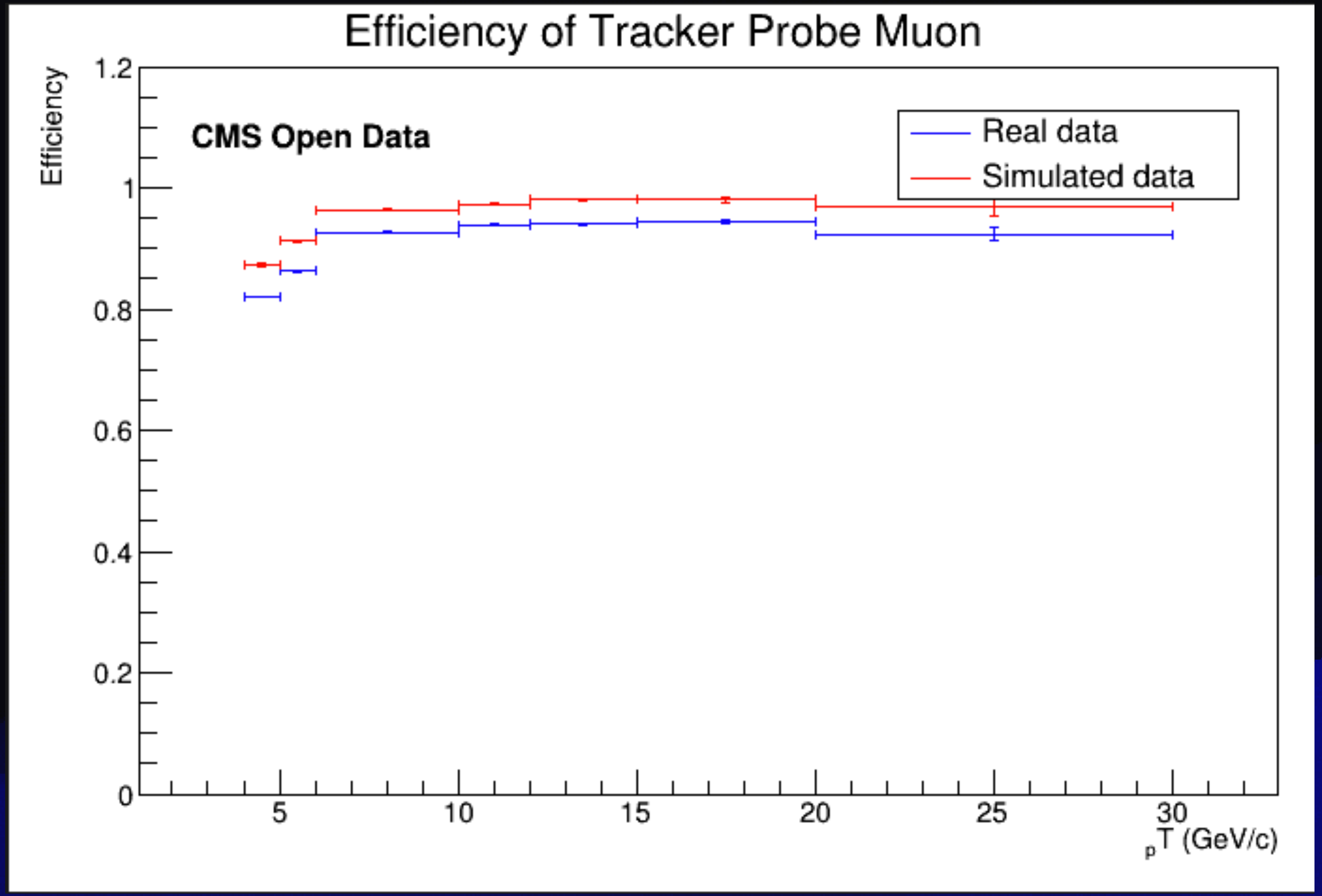



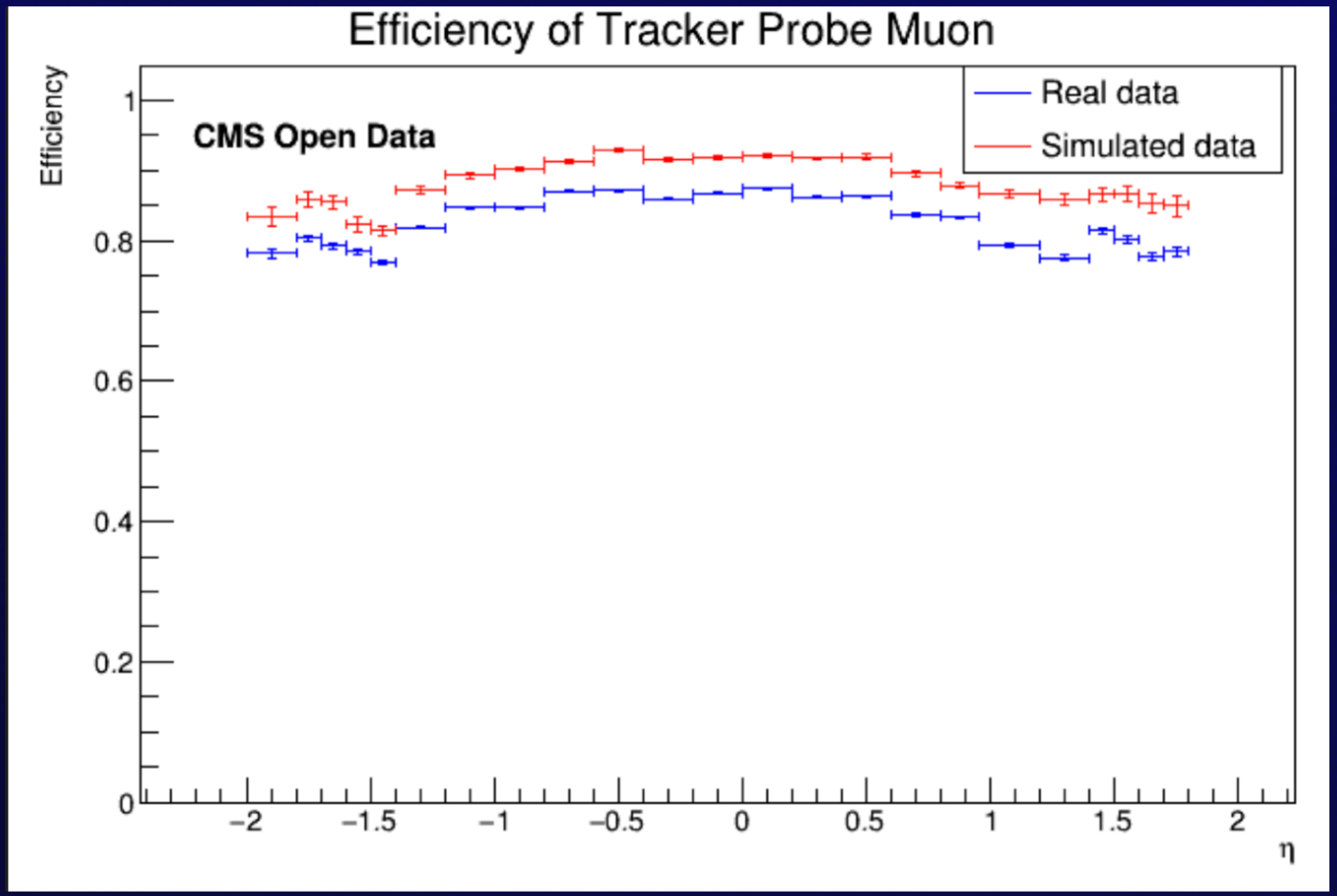

The maximum significance is 1012.13 for an MVA cut of > 0.015.

# •MuonID (1: pass, 0: fail) Efficiency with Tag'n Probe Fitting Method



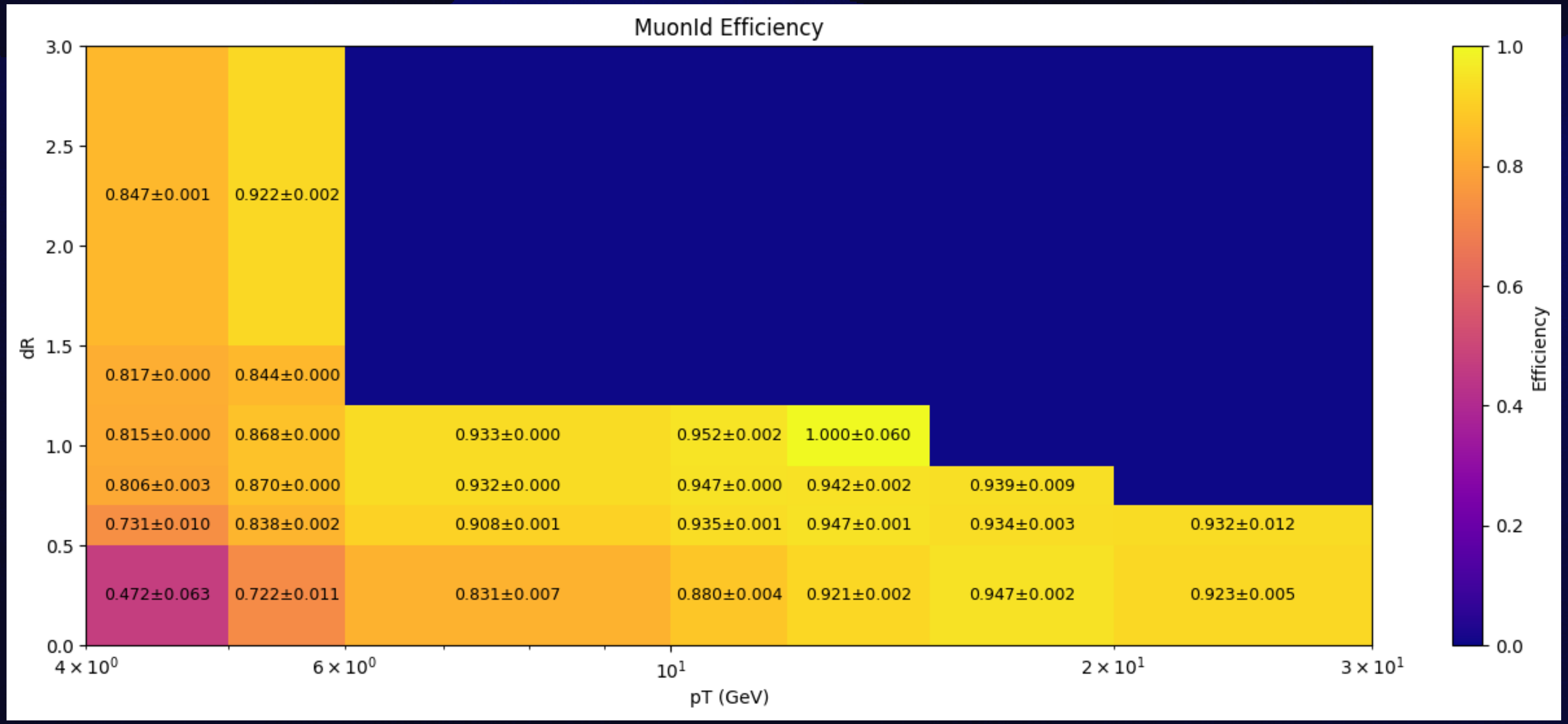



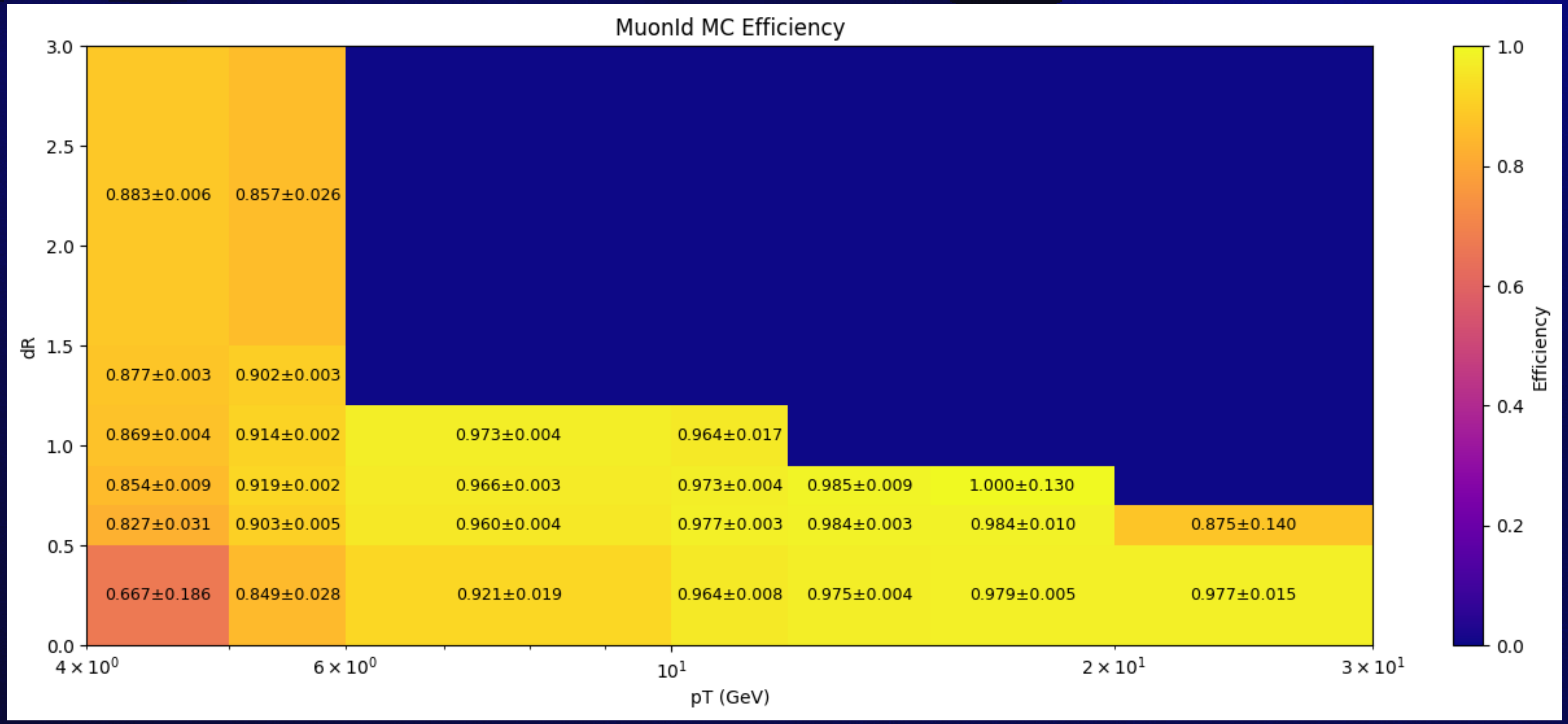




Signal Model: Voigtian Profile for each peak
Background Model: Bernstein Polynomial 1st order

# MuonId Efficiency

- Efficiencies for 2D (pT & dR) binnings
- The total efficiencies are calculated by integrating these values
- MC efficiencies for systematic uncertainties
- Efficiency does not change significantly wrt mass and taken constant as ~0.892



# Approach 2. Traditional Cuts vs Training

i. Compare the results for track isolation (trkiso)

ii. Compare the results for vertex displacement ($L_{xy}$)

*Only for the vertex variables*

# i) Simple Cut vs Training: trkiso

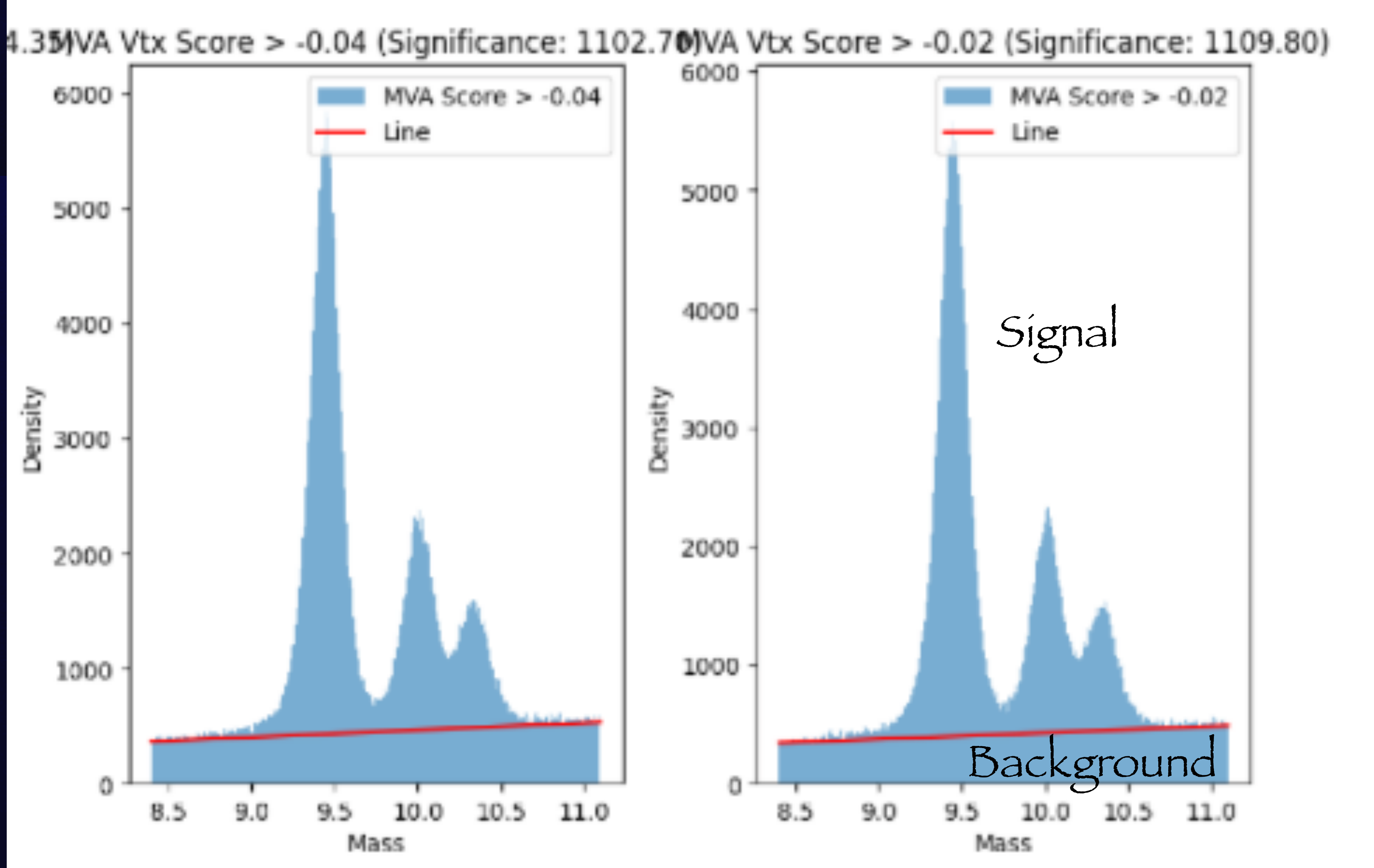
- Trkiso was used in the training for Run II data
- Use the weights of the training using track isolation, optimise the MVA cut, observe the significance
- Exclude trkiso from the training, optimise the MVA cut, optimise the trkiso cut
- Compare the results

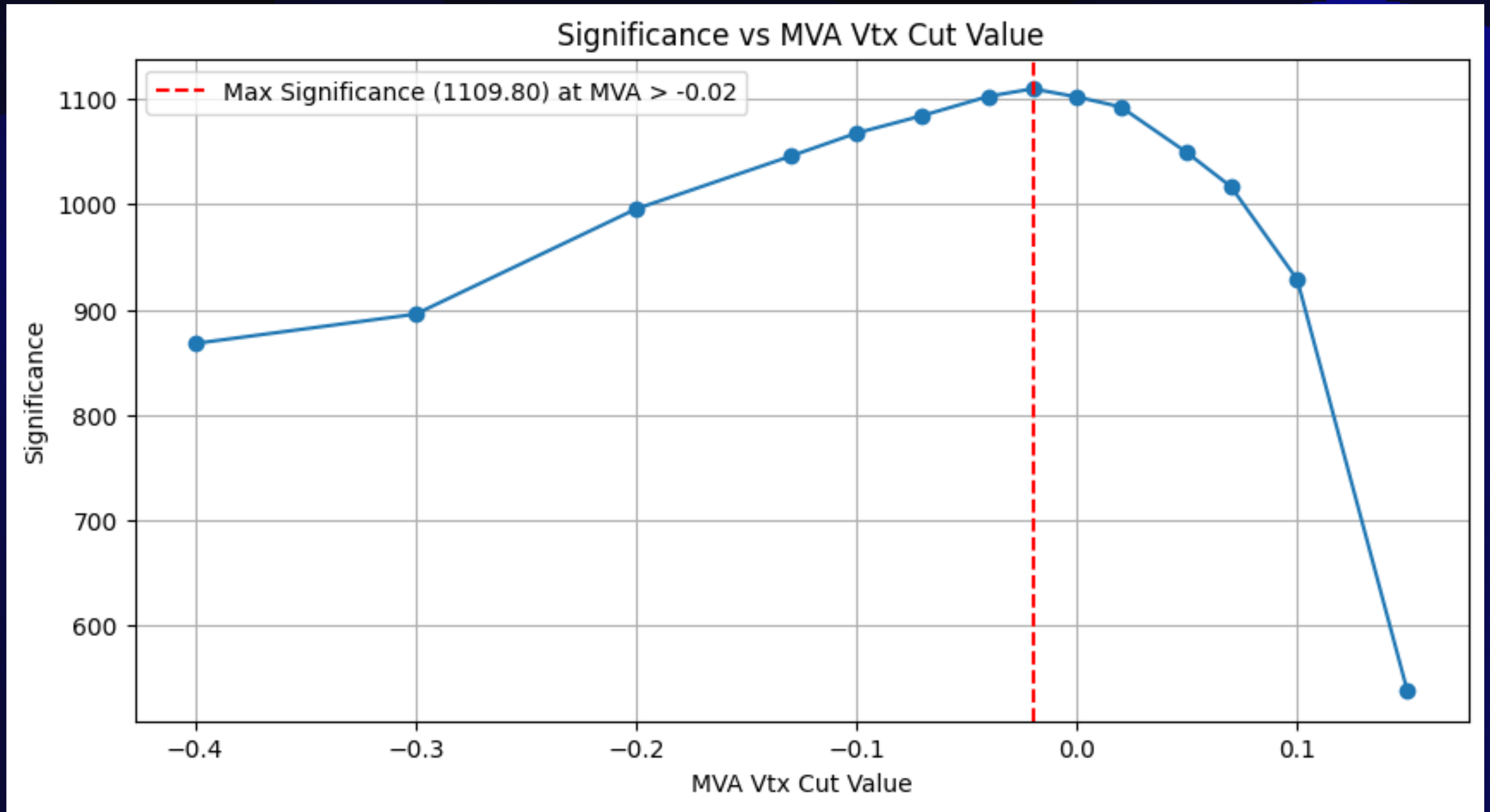
| | Trkiso included | Trkiso excluded + trkiso cut |
|---|---|---|
| Optimal significance | 6210.69 | 6009.85 |
| Signal efficiency at optimal cut: | 0.6764 | 0.8278 |
| Background rejection at optimal cut: | 0.8398 | 0.6293 |

-> Including trkiso in the training is more effective!

# ii) Simple Cut vs Training: $L_{xy}$

- Train a model using vertex parameters[†], add MVA, optimise the MVA cut, observe the significance



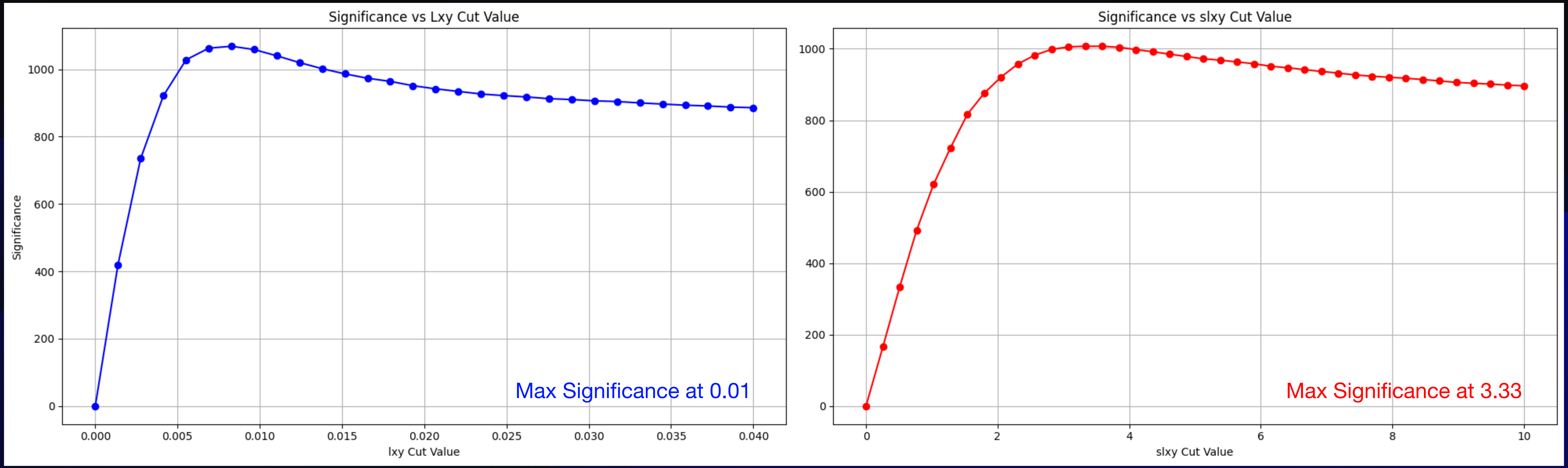




*The maximum significance is 1109.80 for an MVA cut of > -0.02.*



[†]"lxy", "vtxchi2", "slxy", "vtxErr"

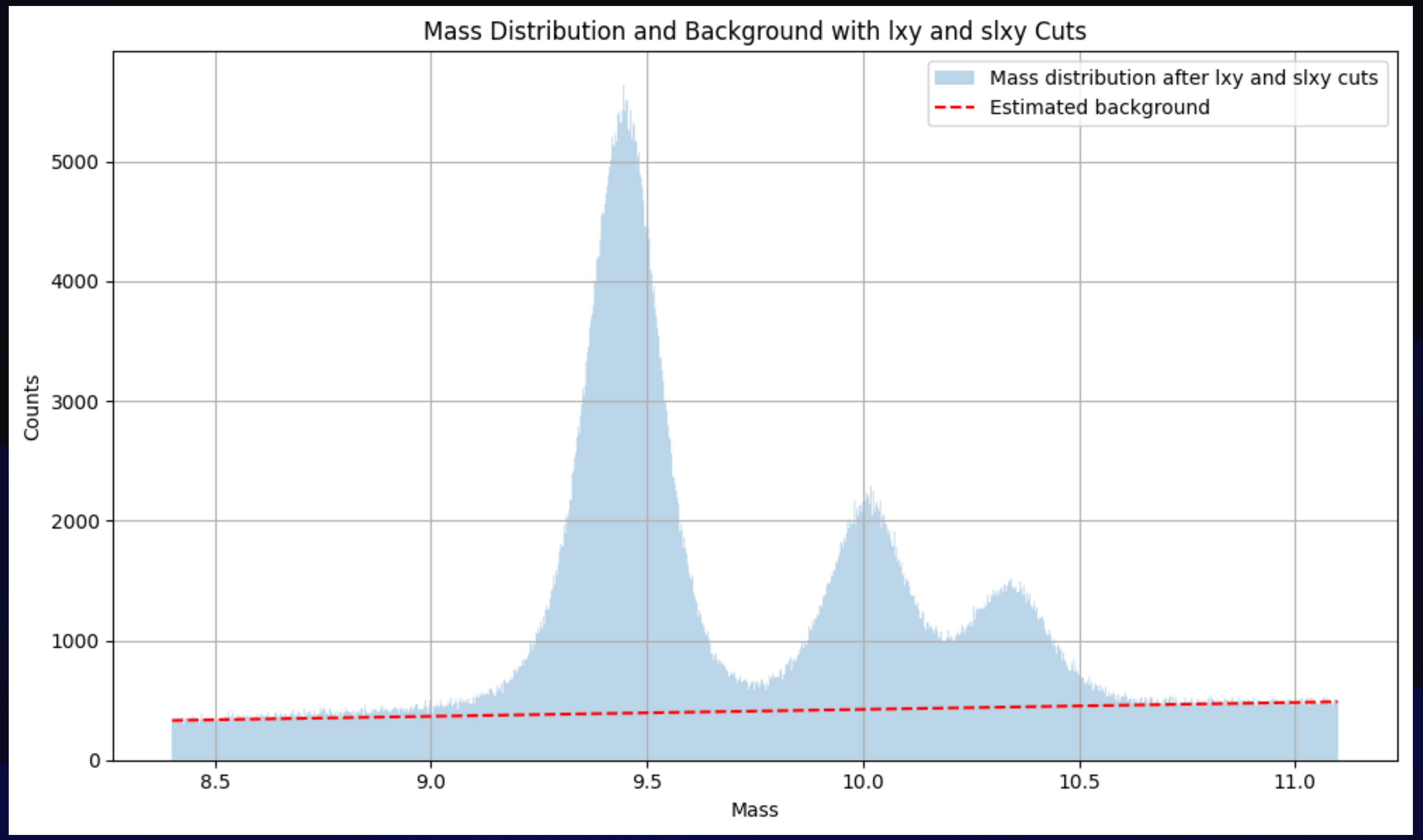
- Ignore the MVA, optimise the $L_{xy}$ and Significance ($L_{xy}/\sigma_{xy}$) cuts

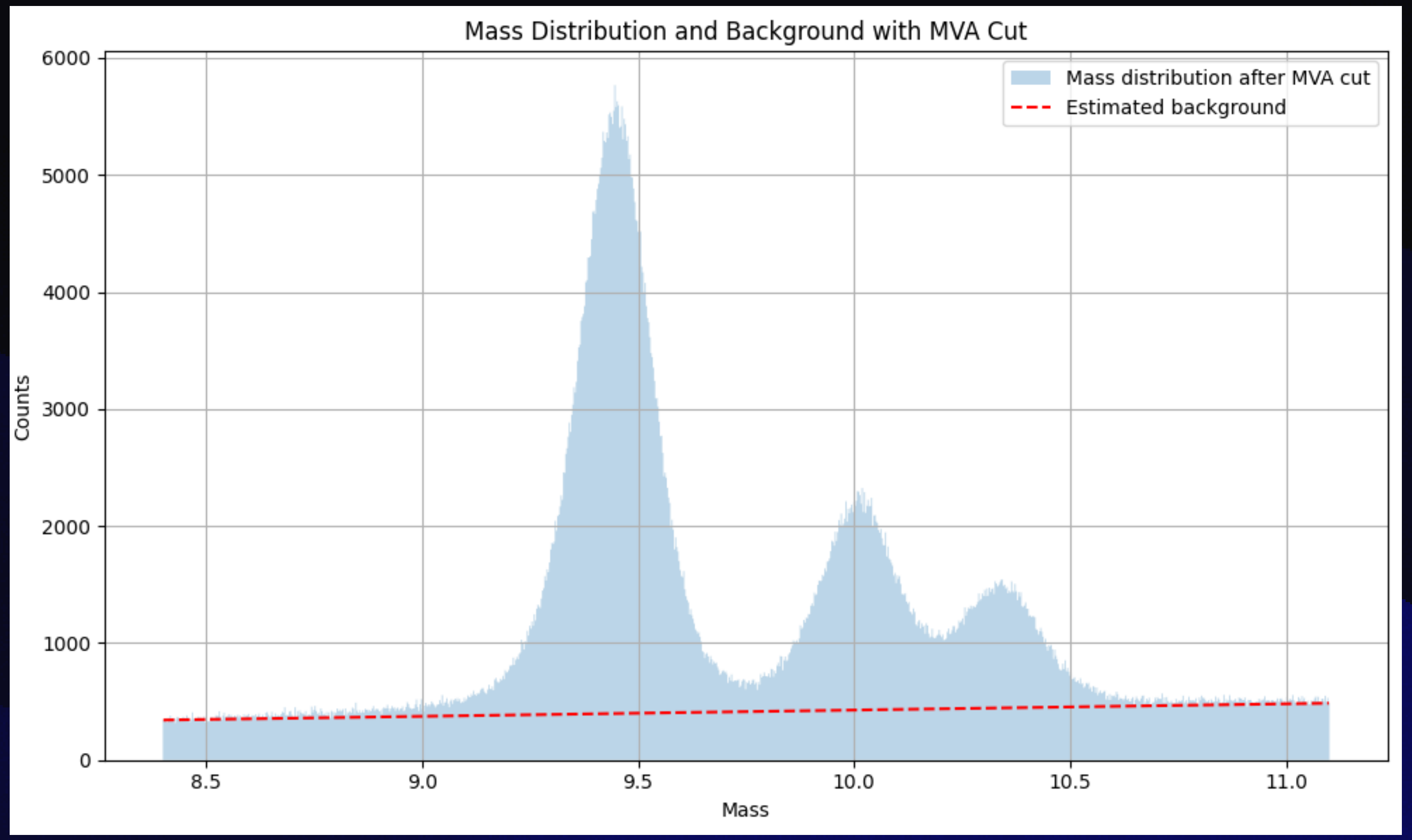



*Max Significance after both cuts: 1072*

**Significance is higher with the MVA cut !**

# Comparison of the Results



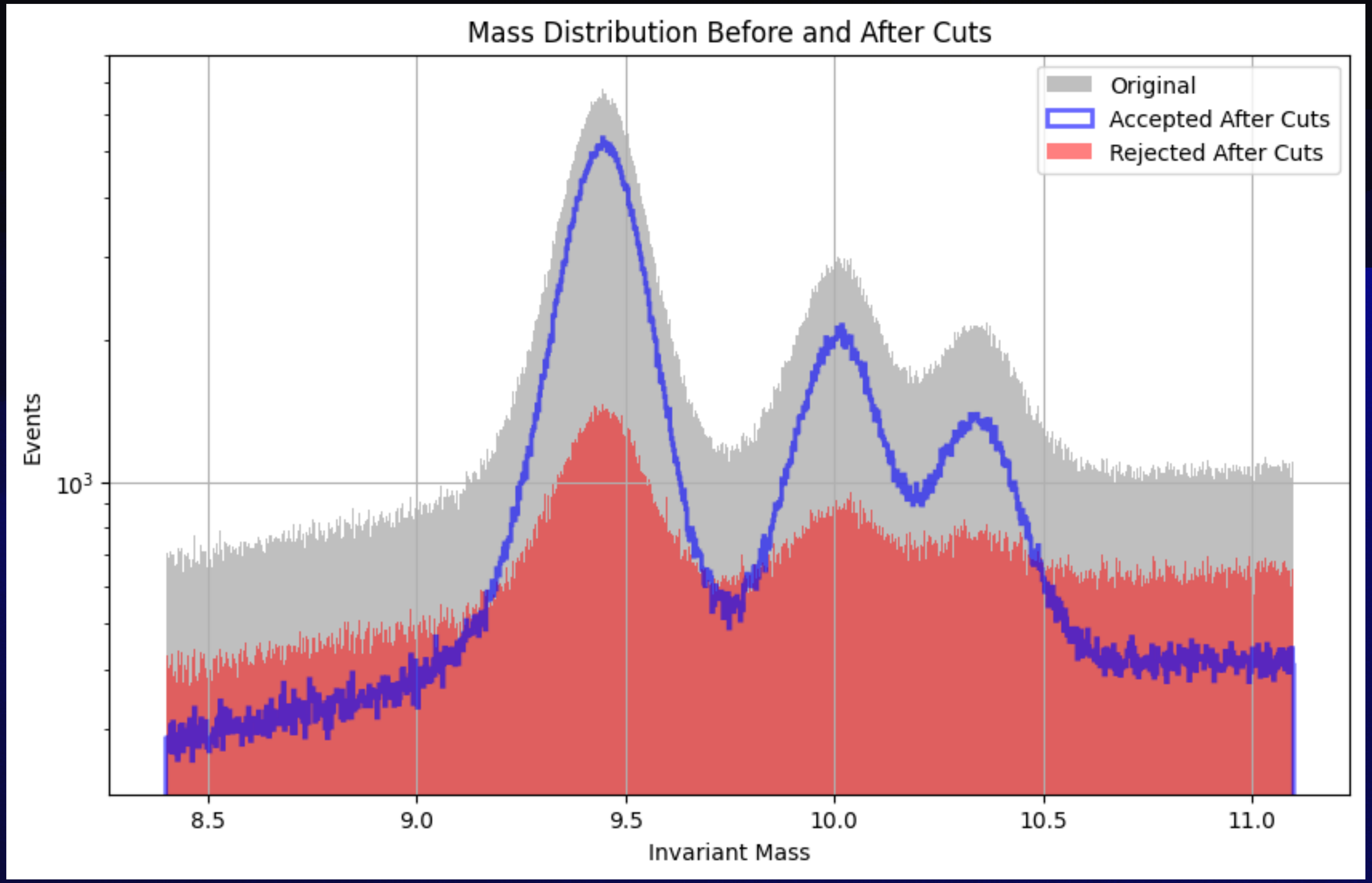




| | Lxy + Slxy Cut | MVA Cut |
|---|---|---|
| Optimal significance | 1072.85 | 1109.80 |
| Signal efficiency at optimal cut: | 0.8500 | 0.8846 |
| Background rejection at optimal cut: | 0.5320 | 0.5263 |

| Vertexing Efficiency | |
|---|---|
| Data | 0.850 |
| MC | 0.925 |
| Ratio | 0.919 |

# Summary: Event Selection

- Muon pair with opposite charges
- Transverse Momentum: $p_t > 4$ GeV
- Pseudorapidity: $|\eta| < 1.9$
- MVA Cut: MVA Score > 0.015
- MVA_VTX Cut : MVA_VTX > -0.02

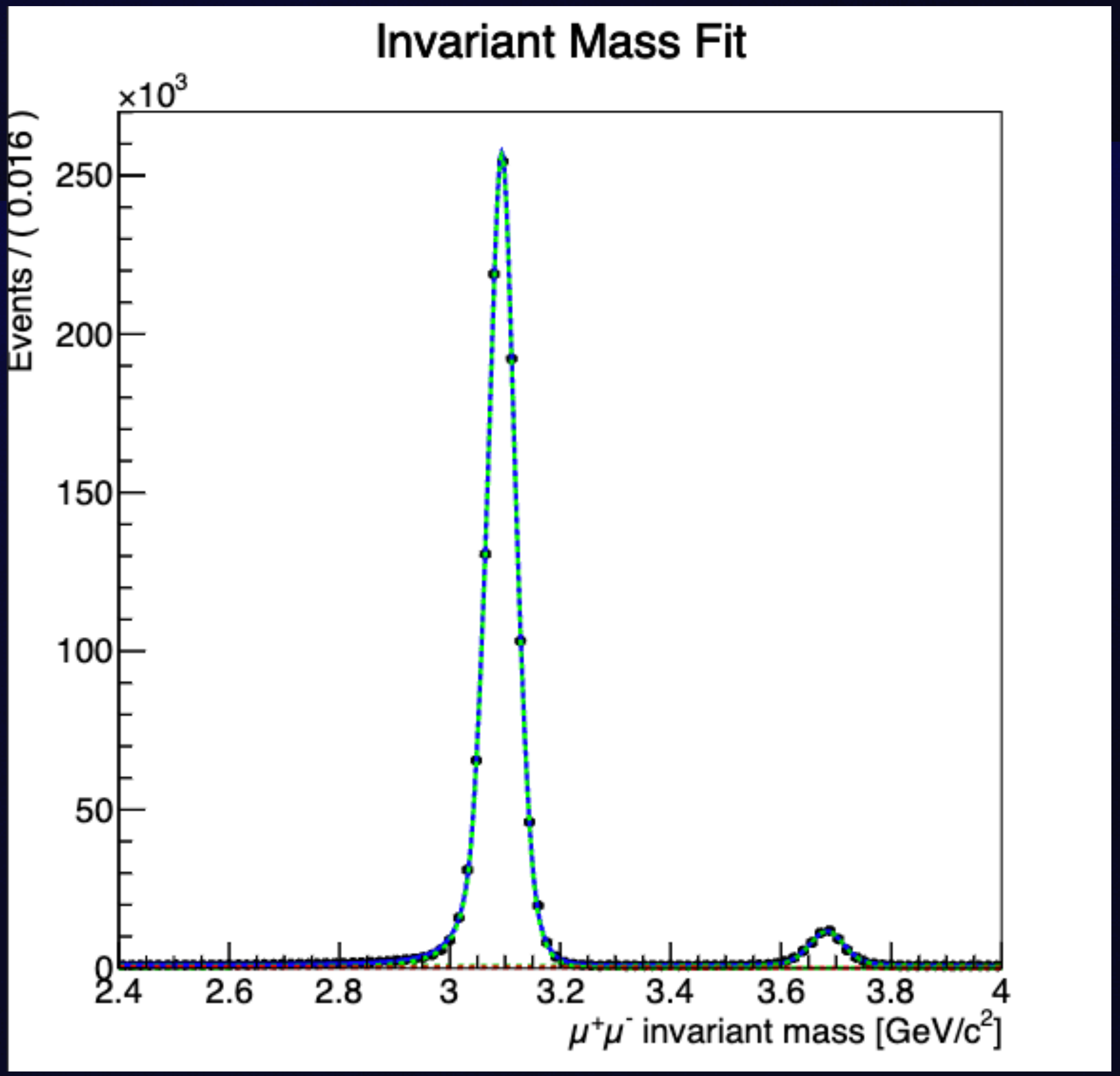


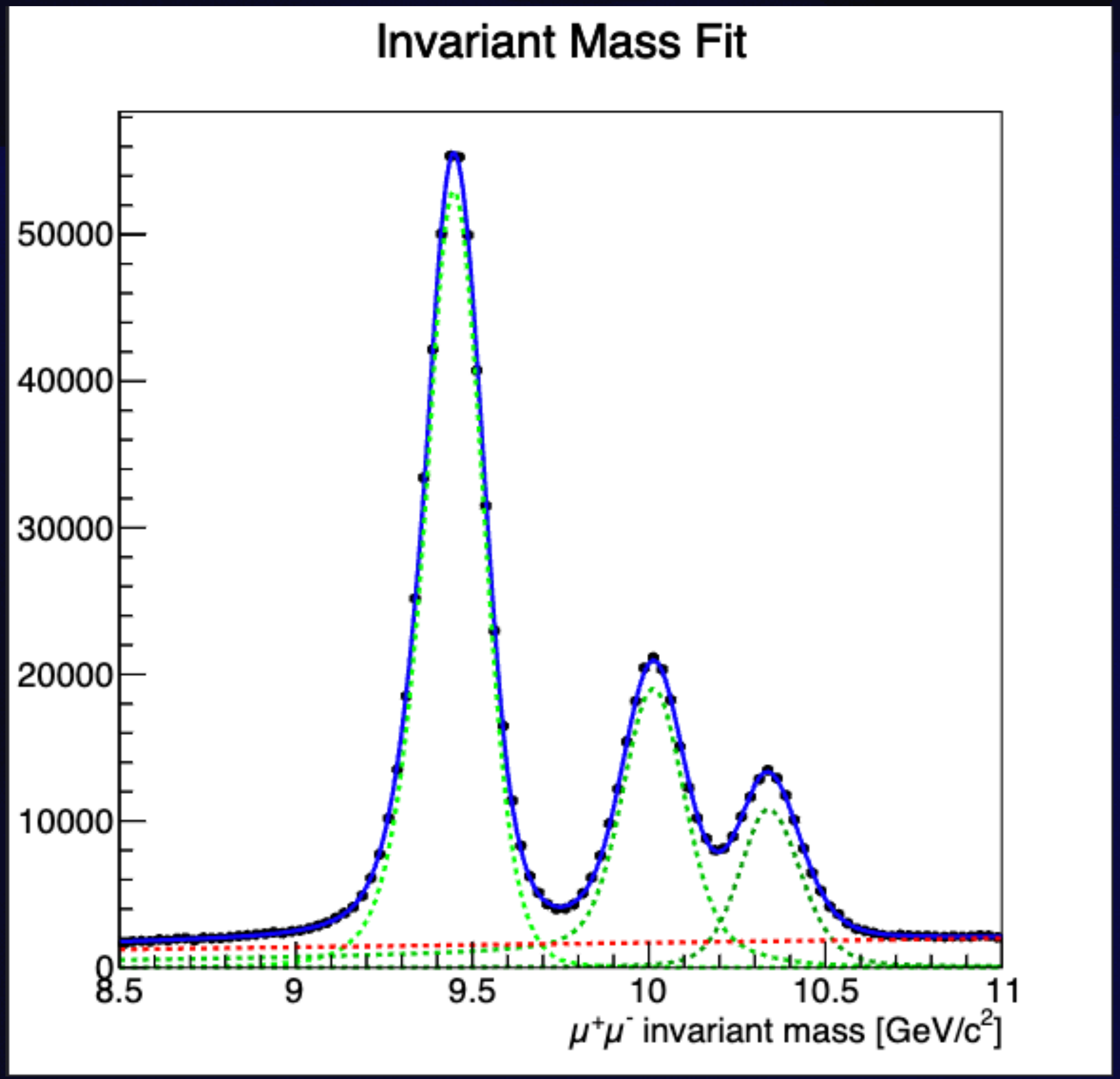
# 3) Signal and Background Modelling

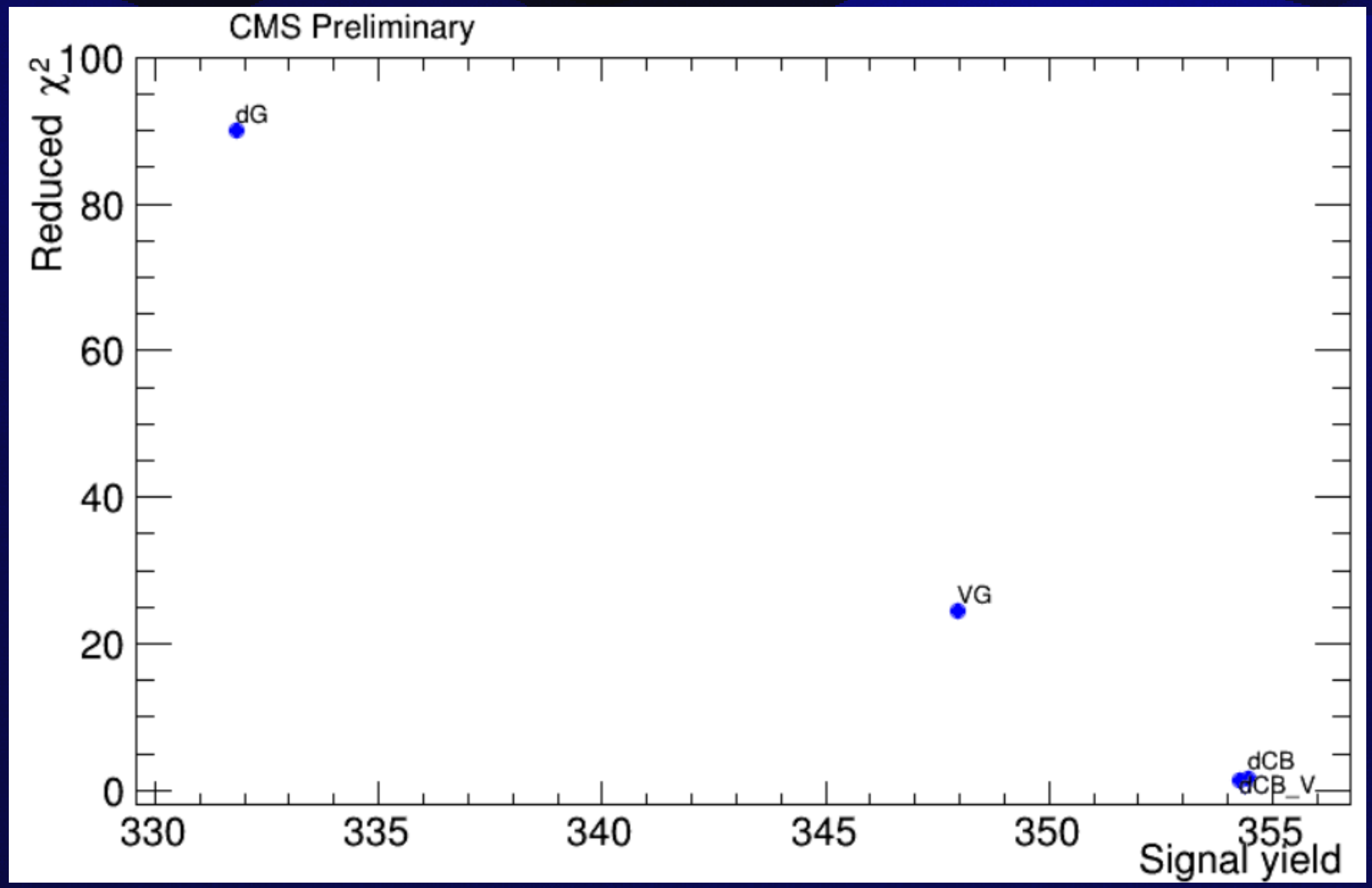
# i) Signal Modelling

**Tested Models:**

- Gaussian
- Voigtian
- dCB + Gaussian
- dCB + Voigtian (best model, also see Ludo's study)




From Ludo

# ii) Background Modelling

Discrete profiling method to vary the function choice.

The families of functions up for investigation using RooMultiPdf:

- **Bernstein Polynomial:** $B_n(x) = \sum_{v=0}^{n} \beta_v b_{v,n}(x)$, **where** $b_{v,n} = \binom{n}{v} x^v (1-x)^{n-v}$
- **Polynomial times exponential:** $P_n(x) = e^{c \cdot x} \sum_{n=1}^{n} \beta_n x^n$
- **Sum of exponentials:** $E_n(x) = \sum_{n=1}^{n} a_n e^{c_n \cdot x}$
- **Bernstein polynomial plus power law:** $B_{Pn}(x) = fB_n(x) + (1-f)x^a$

# 4) Systematic Uncertainties

# Systematic Uncertainties

(i) On efficiency of data-driven selection

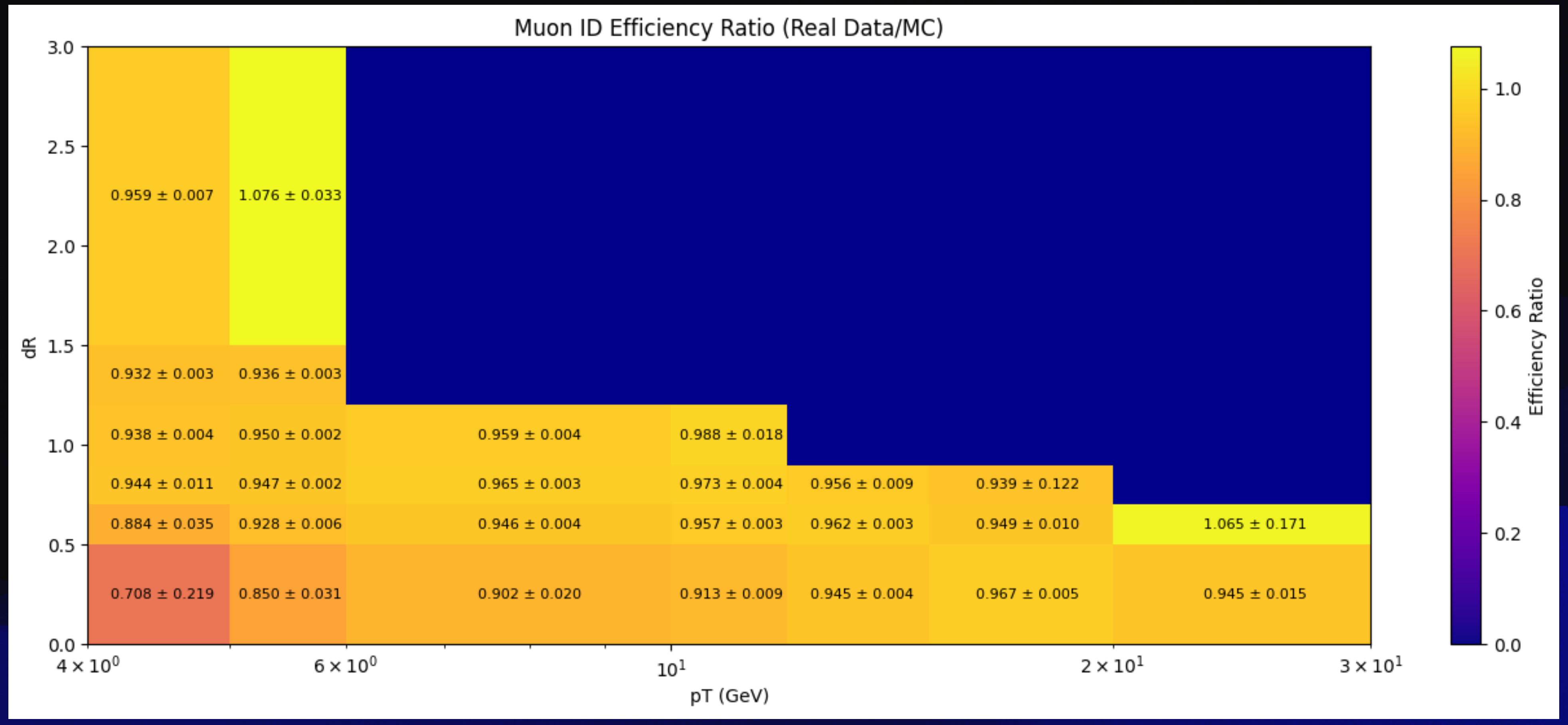
-Compare the efficiencies of MVA cut in Y region in data and MC (~6.9%)

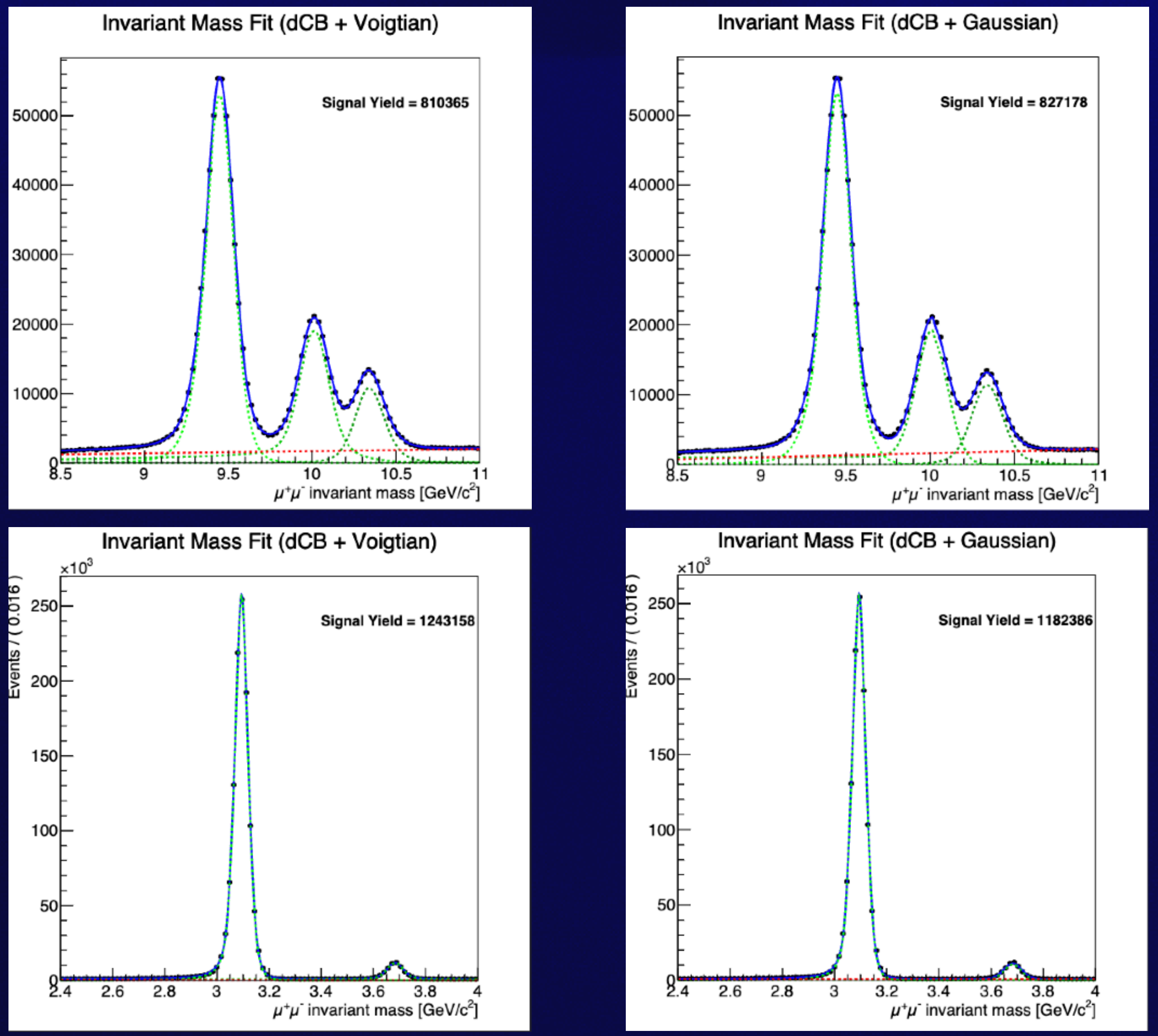
(ii) On signal modelling

-Compare yields of model candidates
-Negligible yield gap (~0.002)

(iii) On Luminosity (65.46 $fb^{-1}$)

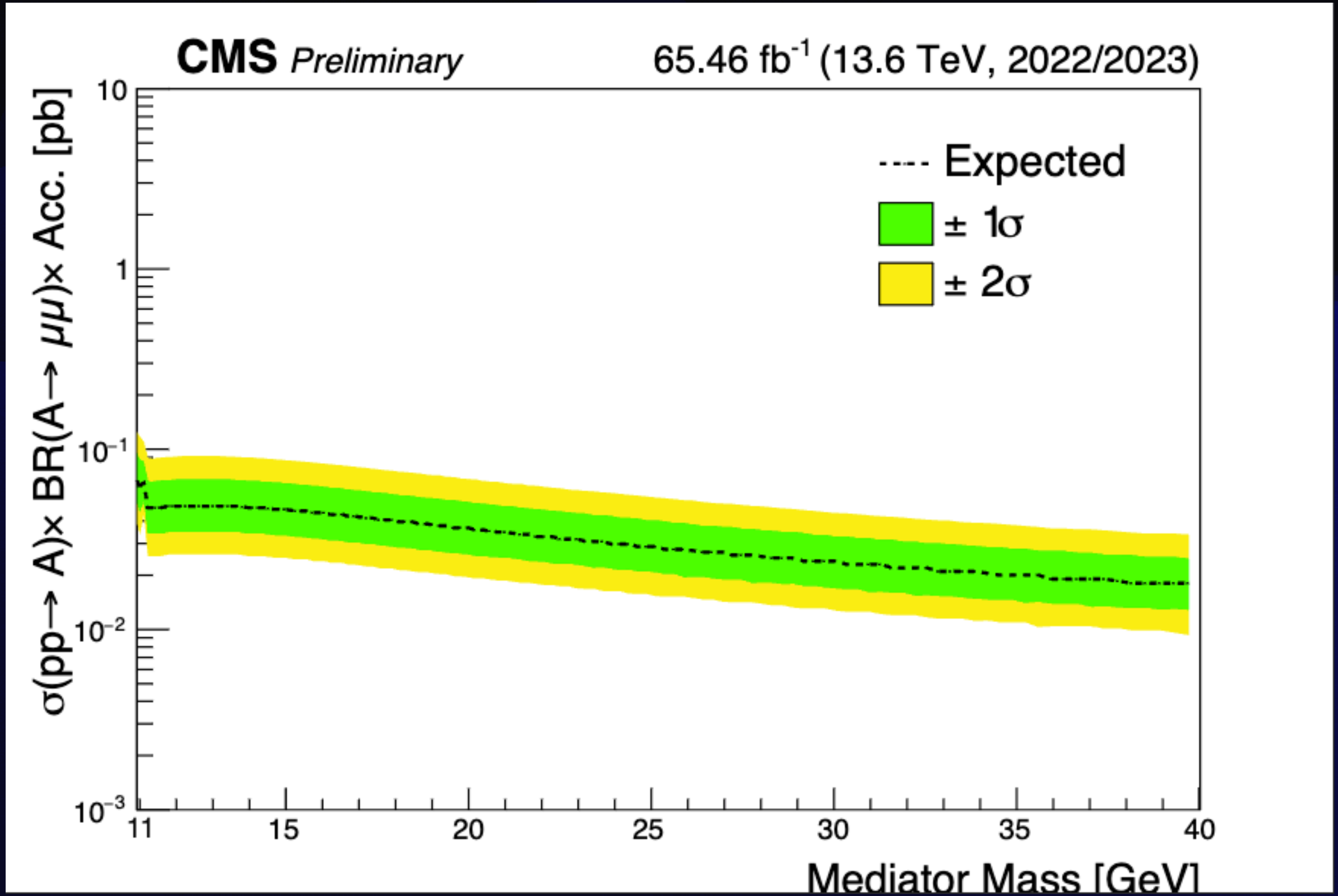
-Uncertainty ~ 2.3%



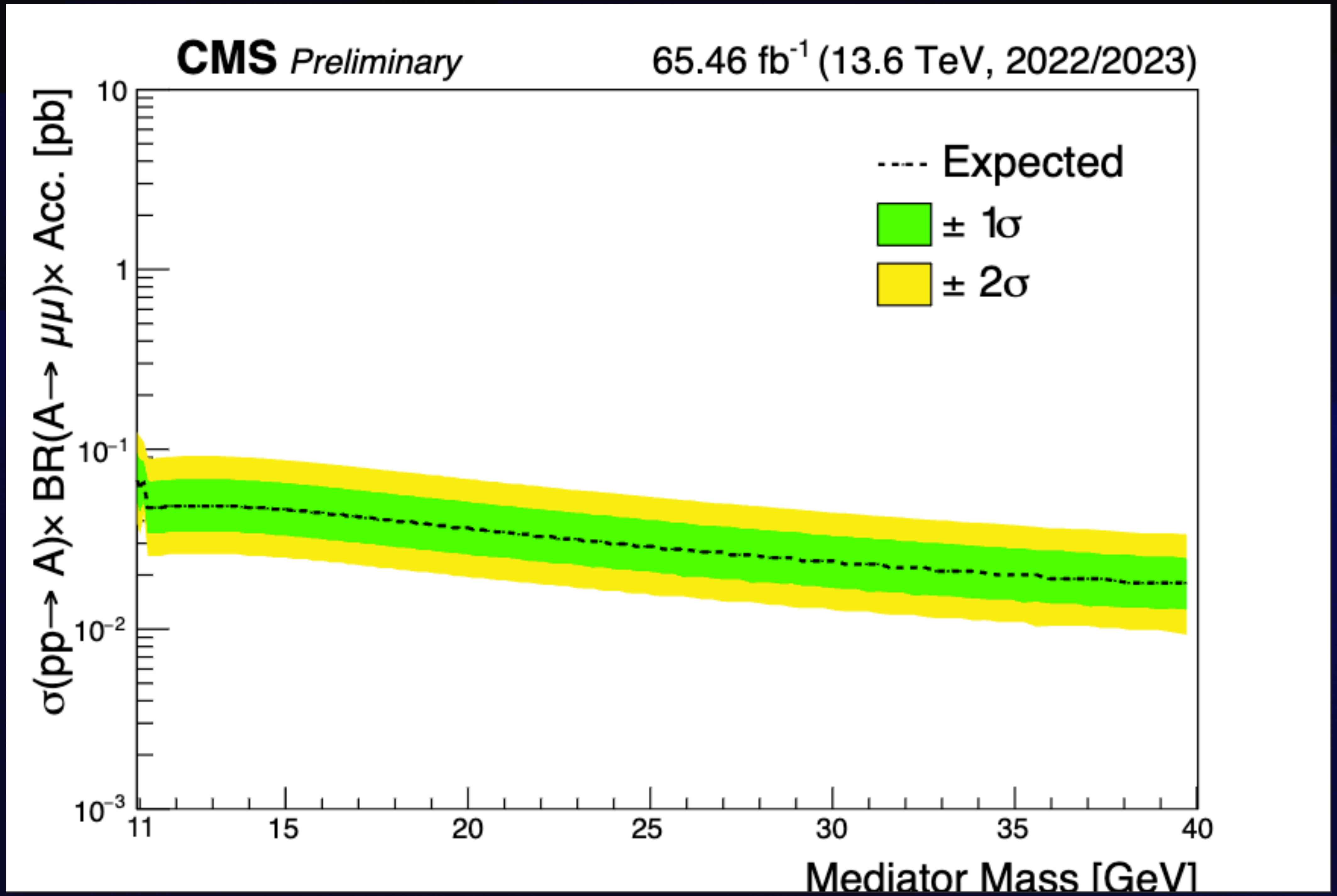
# 5) Exclusion Limits

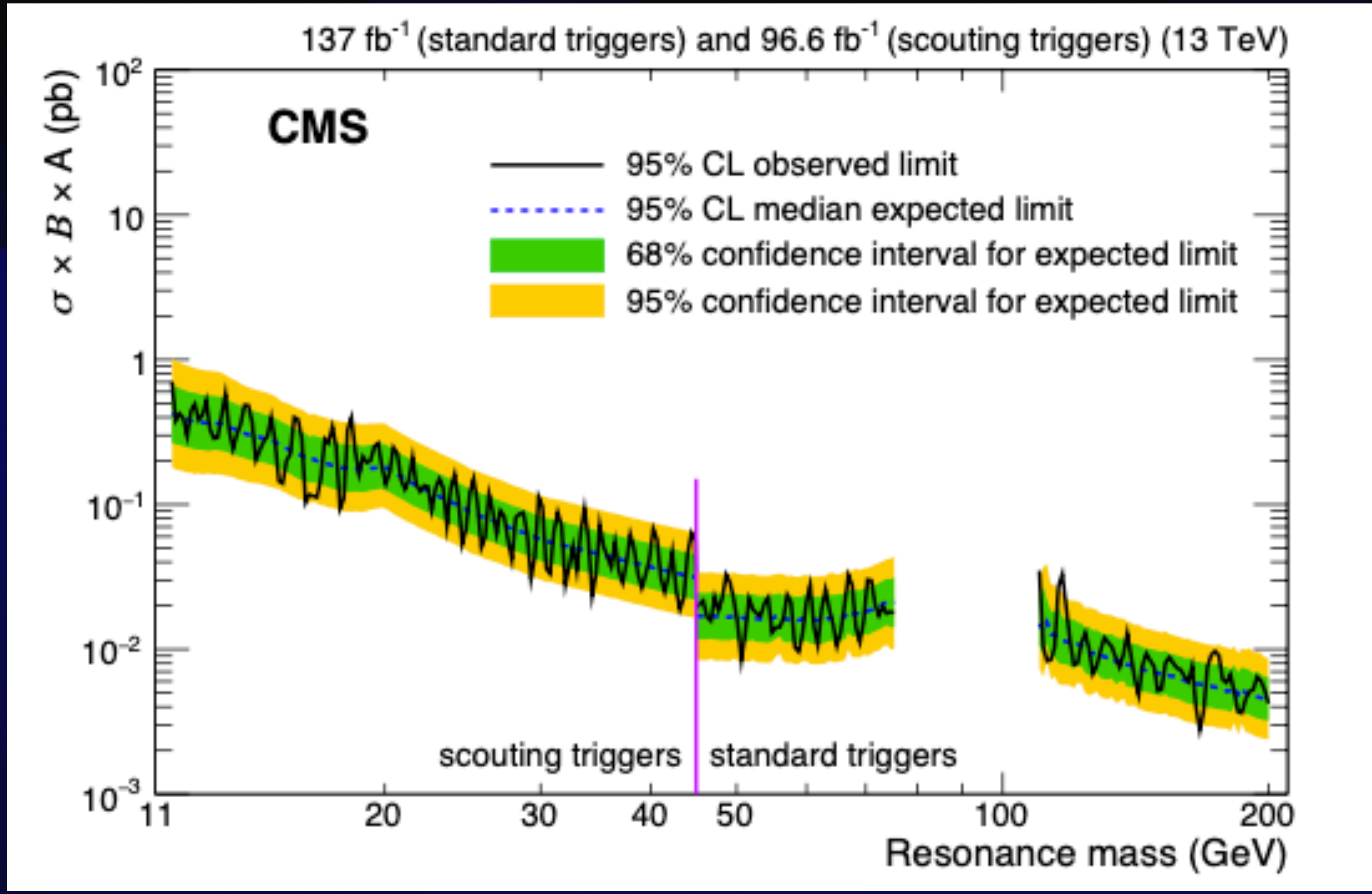
# Expected limits with discrete profiling




- Based on 65.46 $fb^{-1}$ of data collected at 13.6 TeV in 2022/2023, these limits constrain the parameter space of the minimal dark photon model.
- The trigger efficiency is taken from Run 2 and set to be constant at 0.80.

# Comparison with Run 2 results



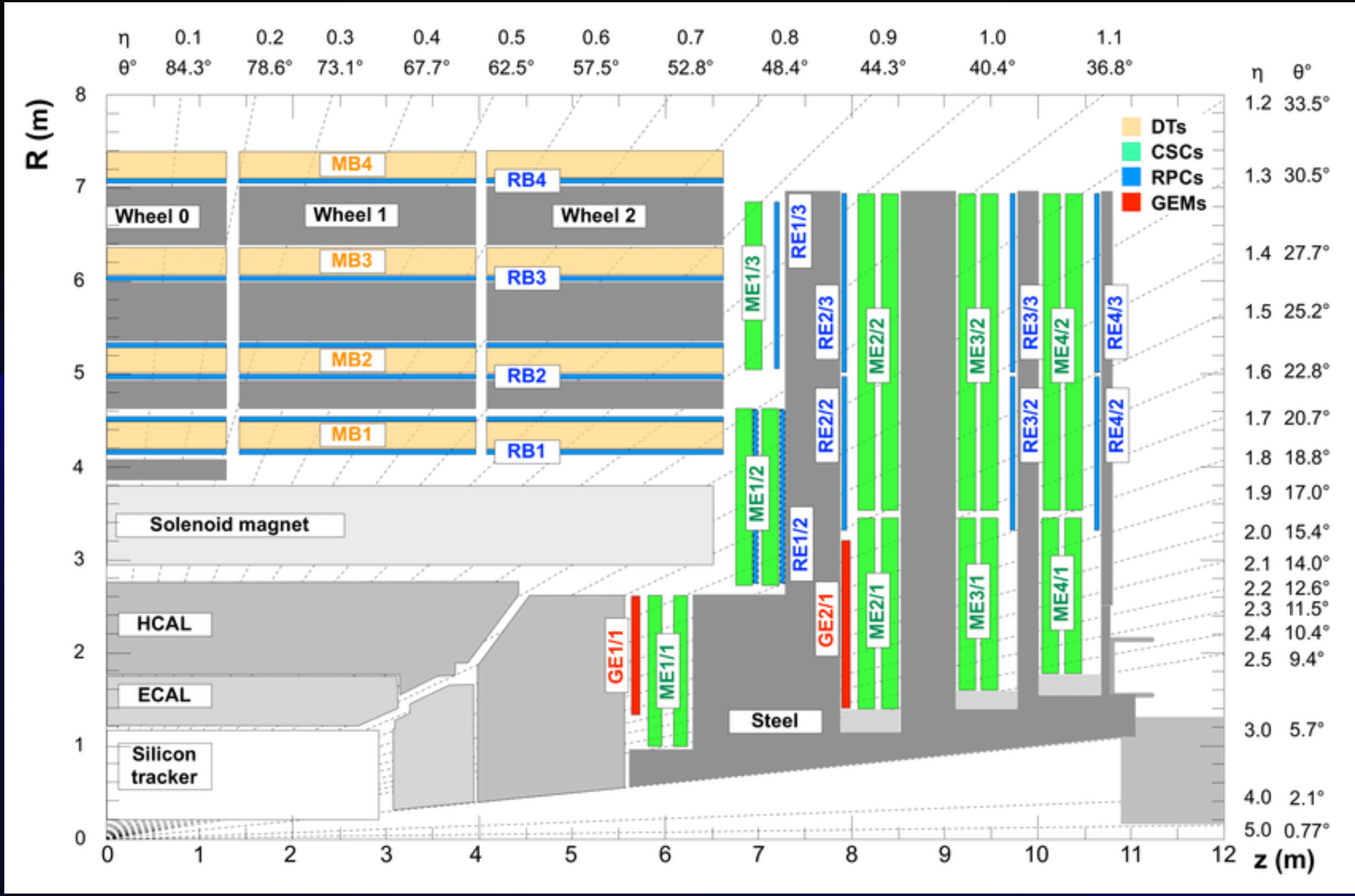

https://arxiv.org/abs/1912.04776

# Summary

- **Minimal Dark Photon Model**
  - Extends Standard Model with new U(1) gauge symmetry
  - Introduces dark photon (Z') interacting with SM particles via kinetic mixing
  - Free parameters: kinetic mixing parameter (ε), dark photon mass (mZ')
  - Experimental signatures: dark photons decay into SM particles ($\mu^+\mu^-$)
- **CMS Run 3 Data and Scouting Trigger**
  - Utilized scouting trigger for data collection
  - Analysis focused on 10-40 GeV mass range
- **Event Selection and Efficiency Calculations**
  - Muon pair with opposite charges, transverse momentum ($p_t > 4$ GeV), pseudorapidity ($|\eta| < 1.9$)
  - Neural network methods for MVA cut optimization
  - Key variables: nmhits, trkiso, trkqoverp, trklambda, dxy, ntklayers, chi, nphits
  - Efficiency calculated with Tag'n Probe fitting method
- **Signal and Background Modelling**
  - Signal models: Gaussian, Voigtian, double Crystal Ball (dCB) + Gaussian, dCB + Voigtian (best) on multiple resonances within mass range
  - Background models: Bernstein Polynomial, Polynomial times exponential, Sum of exponentials, Bernstein polynomial plus power law
  - Discrete profiling method for background
- **Systematic Uncertainties**
  - Data-driven selection efficiency: ~6.9%
  - Signal modelling: Insignificant yield gap (~0.002)
  - Luminosity uncertainty: ~2.3% (65.46 $fb^{-1}$)
- **Exclusion Limits**
  - Based on 65.46 $fb^{-1}$ of data at 13.6 TeV (2022/2023)
  - Limits constrain minimal dark photon model parameter space

# *Thank You!*

# Backup




CMS Muon Detectors