Searching for Dark Scalar decays with T tracks

5/5/25 QEE General Meeting

P. Hoffmann[1], G. Lanfranchi[2], C. Langenbruch[1], S. Libralon[3], M. Lucio Martinez[3], F. Martinez Vidal[3], N. Neri[4], I. Sanderswood[3], A. Usachov[5], [3]J. Velilla Serna

[1]Heidelberg University, [2]Laboratori Nazionali di Frascati dell'INFN, [3]Instituto de Física Corpuscular, University of Valencia—CSIC, [4]INFN Milano, [5]Nikhef National Institute for Subatomic Physics and VU University Amsterdam

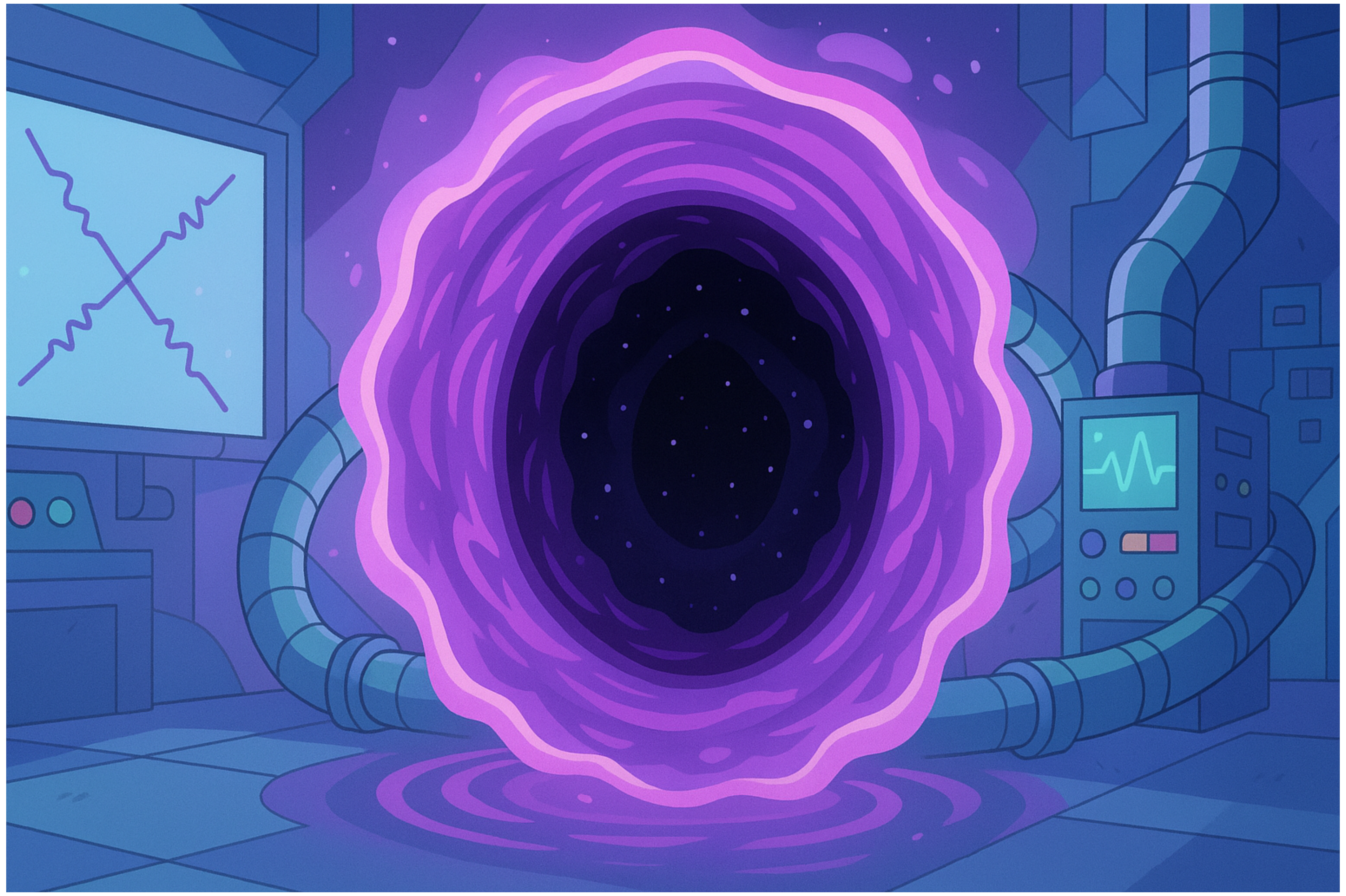
Overview

- Background and motivation
- Analysis strategy
- Data and Monte Carlo samples
- Normalisation
- Reconstruction
- Monte Carlo corrections
- Selections
- Efficiencies
- Signal & control regions
- Background studies

Background and Motivation

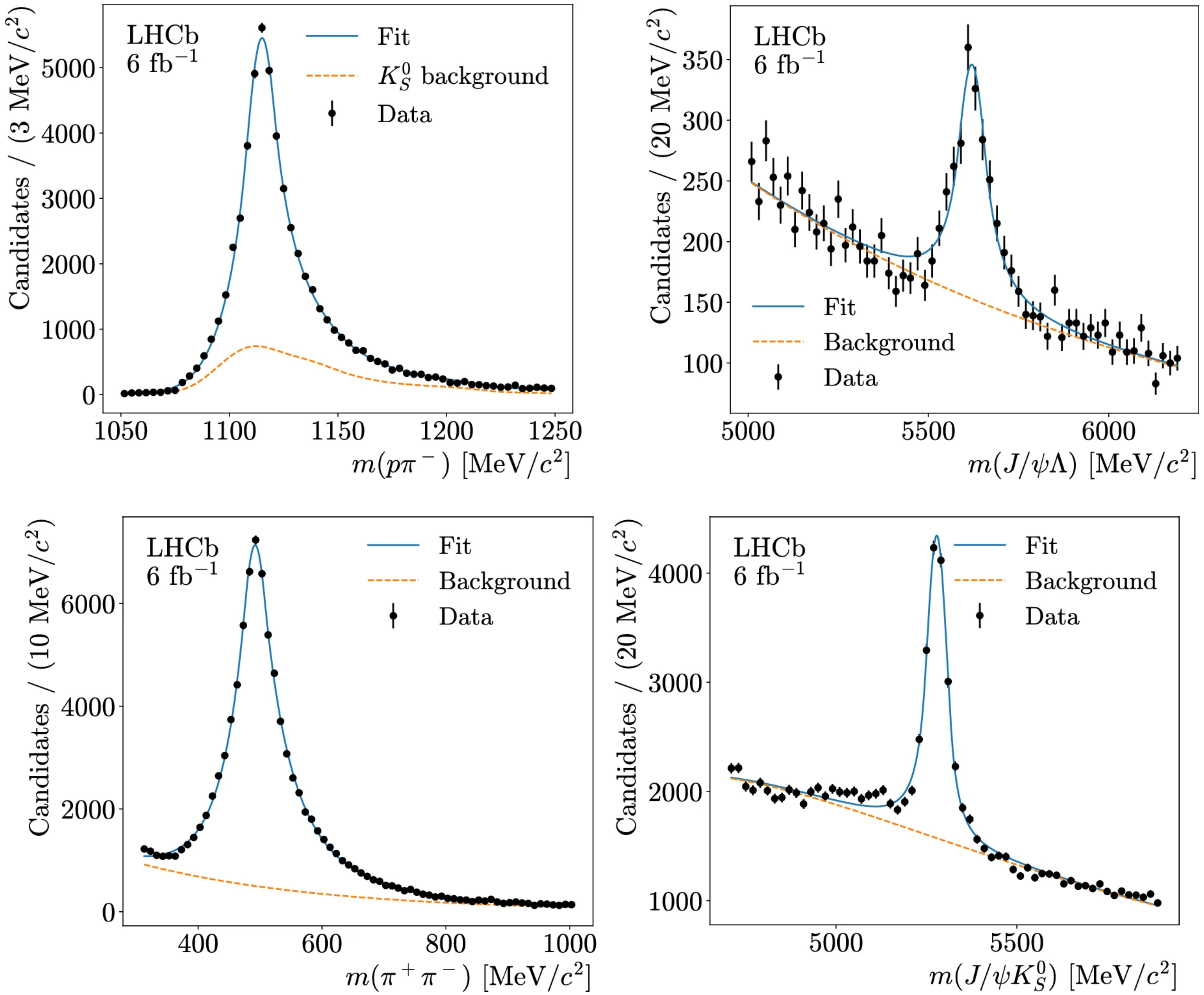
Background

- The Standard Model (SM) is incomplete: questions remain about the nature of dark matter, dark energy, matter-antimatter asymmetry, aspects of the Big Bang, etc.
- Many extensions of the SM propose a hidden sector of particles to answer these questions, which can interact with SM particles via portals, for example the Higgs portal
  - A hidden Higgs couples to the standard Higgs
  - The coupling strength determines the lifetime of the particle
  - Particles with very weak couplings are known as Feebly Interacting Particles (FIPs)



Reconstruction with T tracks

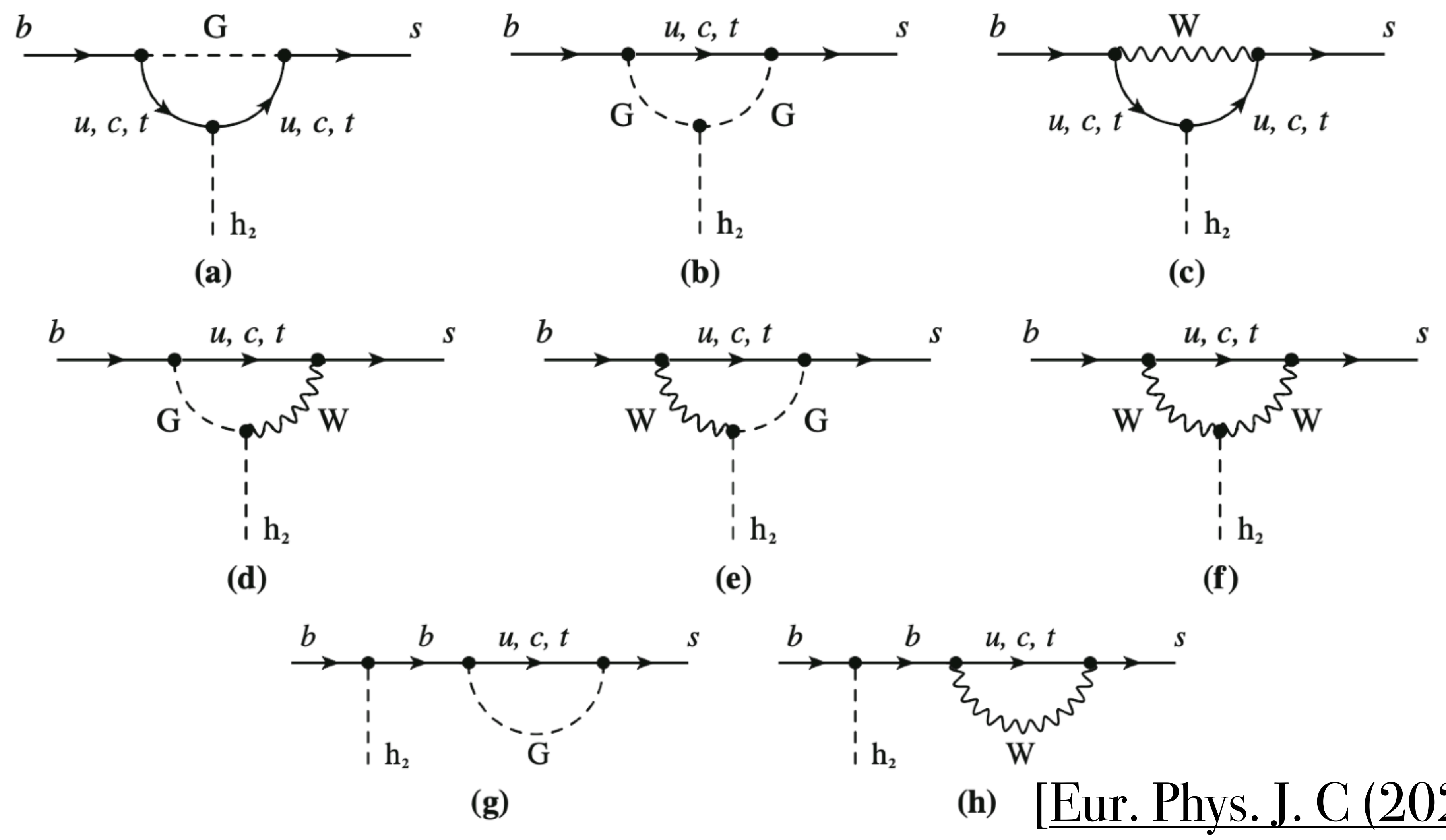
- Feasibility of using T tracks as focus of analysis demonstrated with Λ and K_S^0 decaying 6-7.6 m from PV using Run 2 data [Eur. Phys. J. C 77, 181 (2017)], [Eur. Phys. J. C 85, 7 (2025)]
  - See also talks from Giorgia and Javier at BandQ meeting
- This means FIPs decaying up to 7.6 m become plausible targets of searches
- Challenges:
  - Combinatorics
  - Momentum and vertex resolution
  - Completely new region of phase space
- For details in Run 3 see talk at general performance meeting



Analysis Strategy

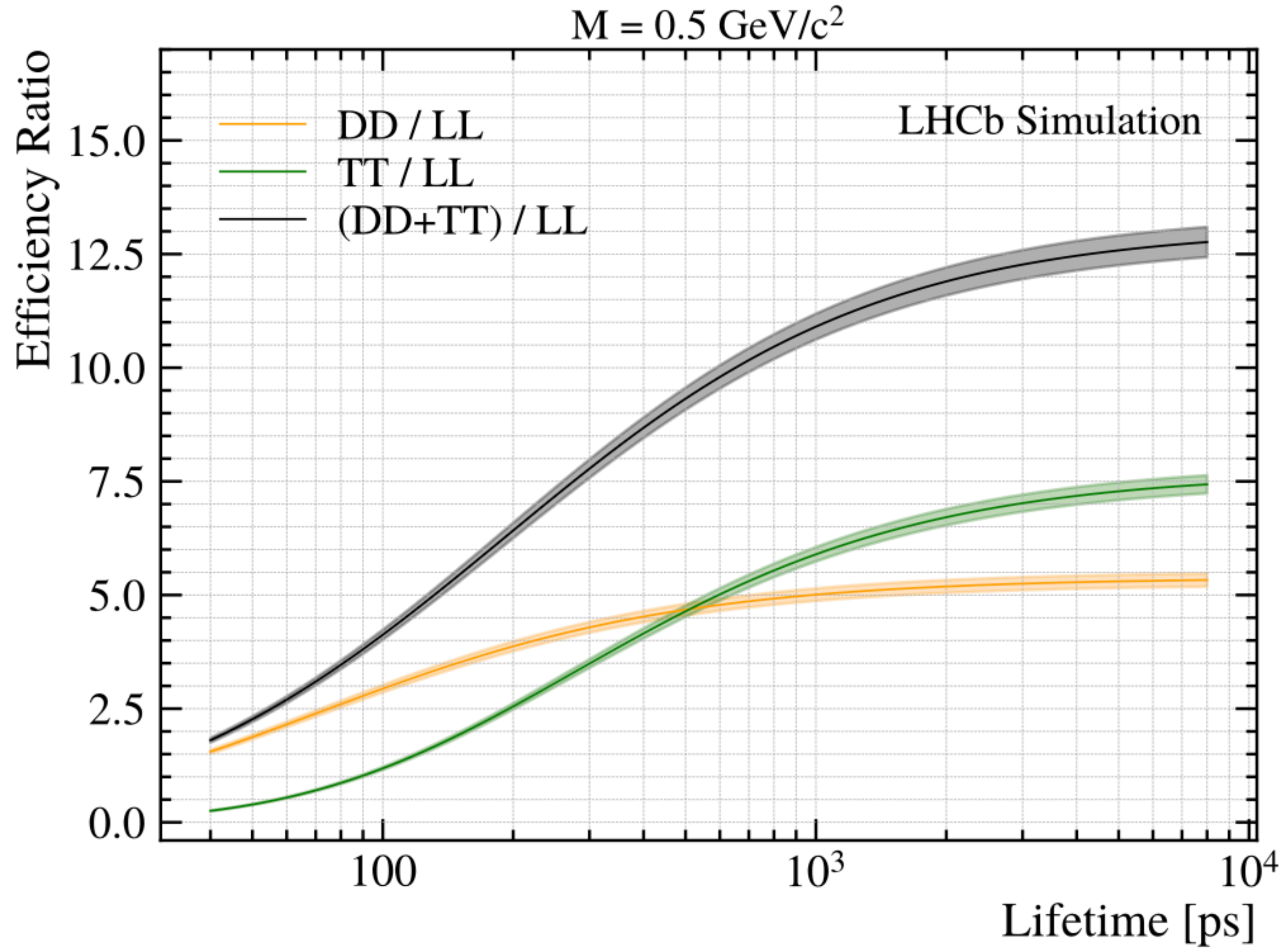
Overall Strategy

- Search for (pseudo-)scalar particles produced in $b \to \phi s$ decays, where the b hadronises to a B or B^+ and φ decays to dimuon
  - Muon ID helps to control background
  - Using T tracks unlocks higher lifetimes and therefore weaker couplings compared to Long and Downstream
- Search *inclusively* i.e. search is only for a particle decaying to two muons in the magnet region, rather than reconstructing B meson
  - No hypothesis, reconstructing or matching made with accompanying strange hadron → higher efficiencies
  - Exclusive search also being considered in parallel for improved mass resolution and to serve as cross check



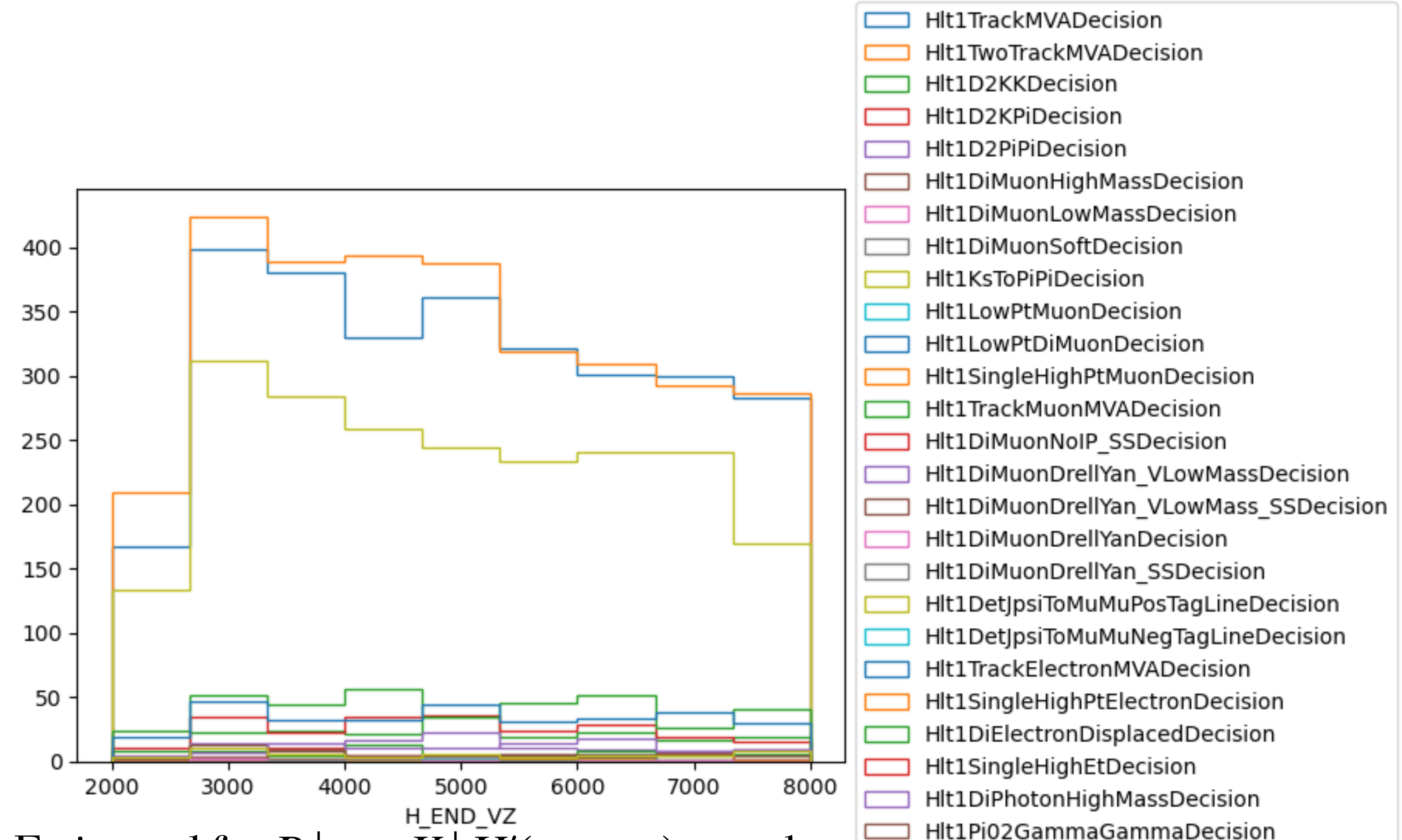
[Eur. Phys. J. C (2020) 80 :669]

Fig. 1 One-loop diagrams contributing to $b \to sh_2$ in R_ξ gauge

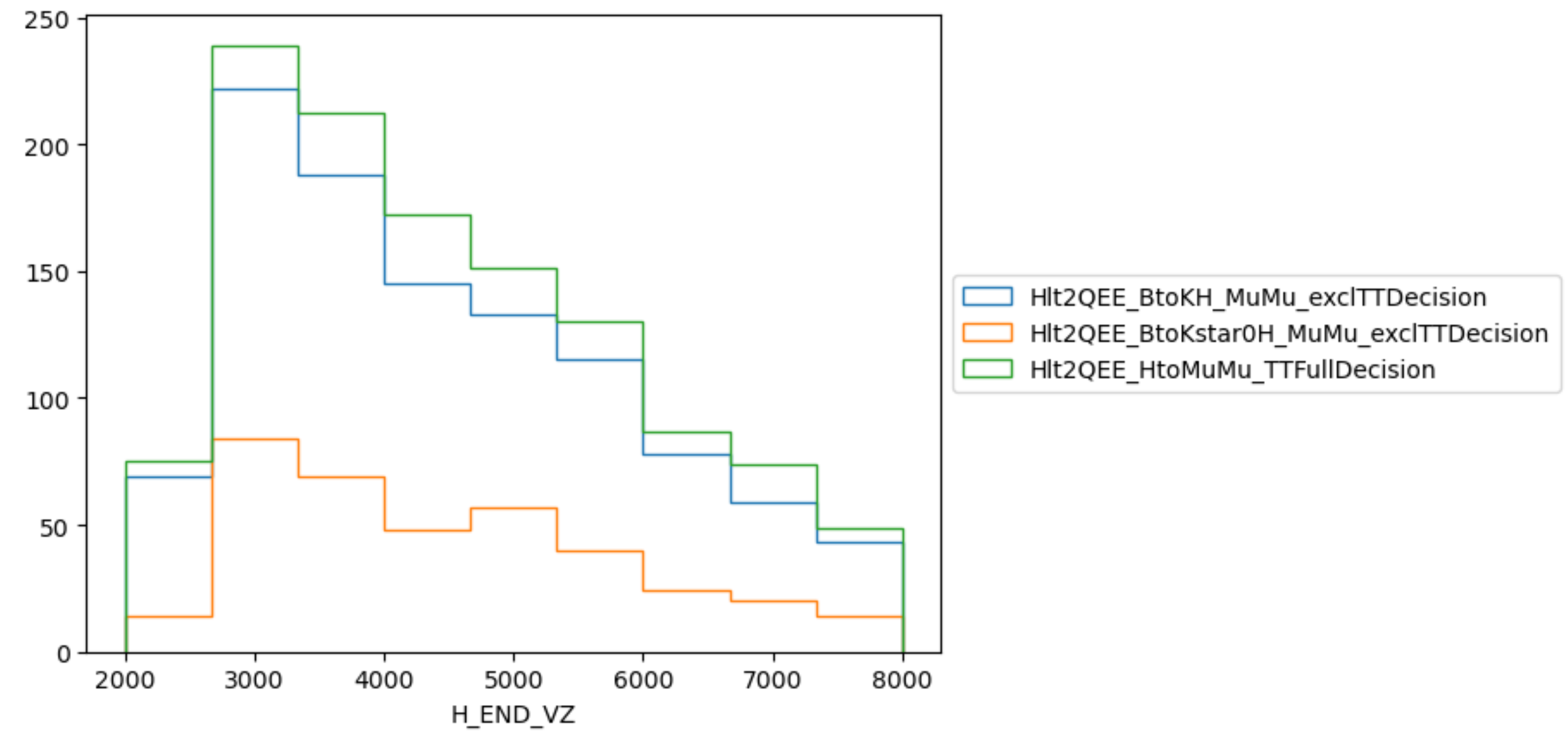


Selection Strategy

- In HLT1, trigger is independent (TIS) of dimuon candidate
  - Therefore largely depends on the accompanying strange hadron, and underlying event
  - Fully inclusive in HLT1
- In HLT2, trigger on (TOS) dimuon candidate via a dedicated trigger line (Hlt2QEE_HtoMuMu_TTFull)
- Final offline selection to be performed with an MVA approach (either NN or BDT)



Estimated for $B^+ \to K^+H'(\to \mu\mu)$ sample, $\tau_{H'}$ =1 ns, m_H3.5 GeV

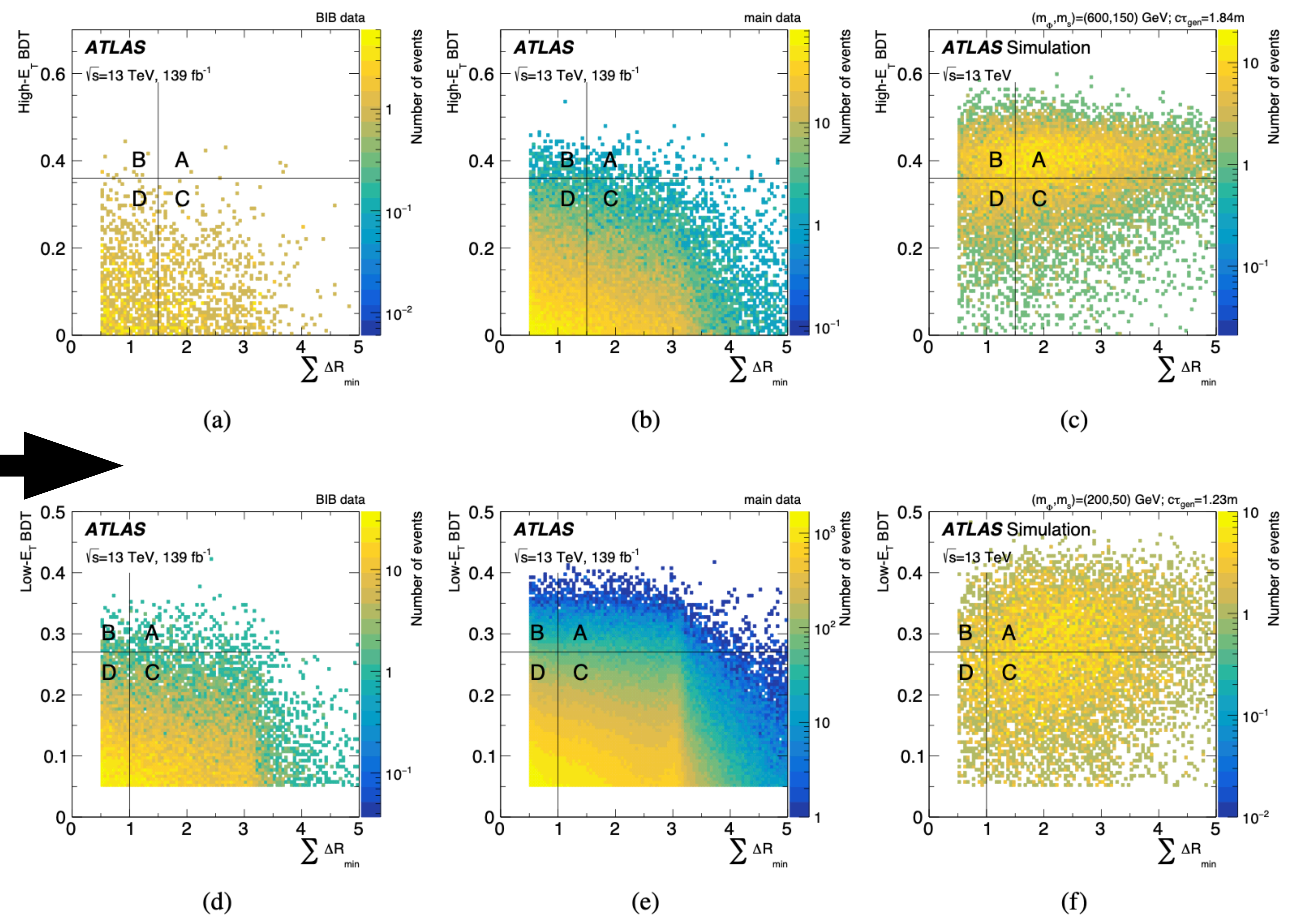


Normalisation Strategy

- Normalise to $B \to J/\psi K_S^0$ where the $K_S^0 \to \pi\pi$ is reconstructed using T tracks
- $b\overline{b}$ cross sections and fragmentation fractions in B meson drop out
- Only relative variations in T track reconstruction efficiencies are taken into account
  - Will have to be binned in p_T, η in order to account for kinematic differences
- Control sample collected and being studied in parallel for electric and magnetic dipole moment measurements

Blinding Strategy

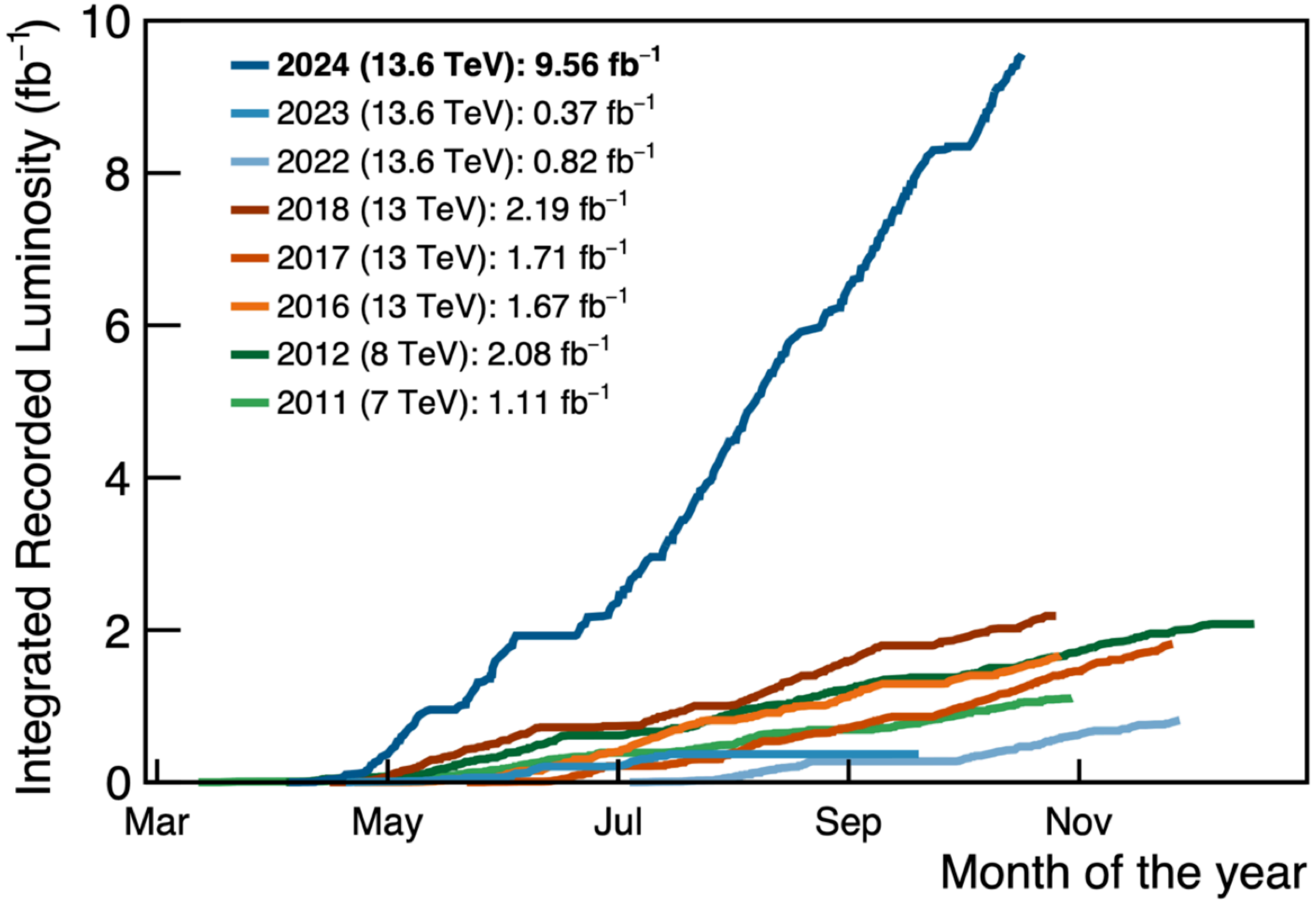
- Analysis is performed blind
- A predetermined signal region will be excluded from the analysis, and only opened at end of review
- Signal region to be determined by plane of two uncorrelated variables with a good signal-background separation (ABCD method, see e.g. https://arxiv.org/pdf/2203.01009) (WIP)
- This region is required to accurately estimate the background in the signal region using a data-driven approach



Data and MC Samples

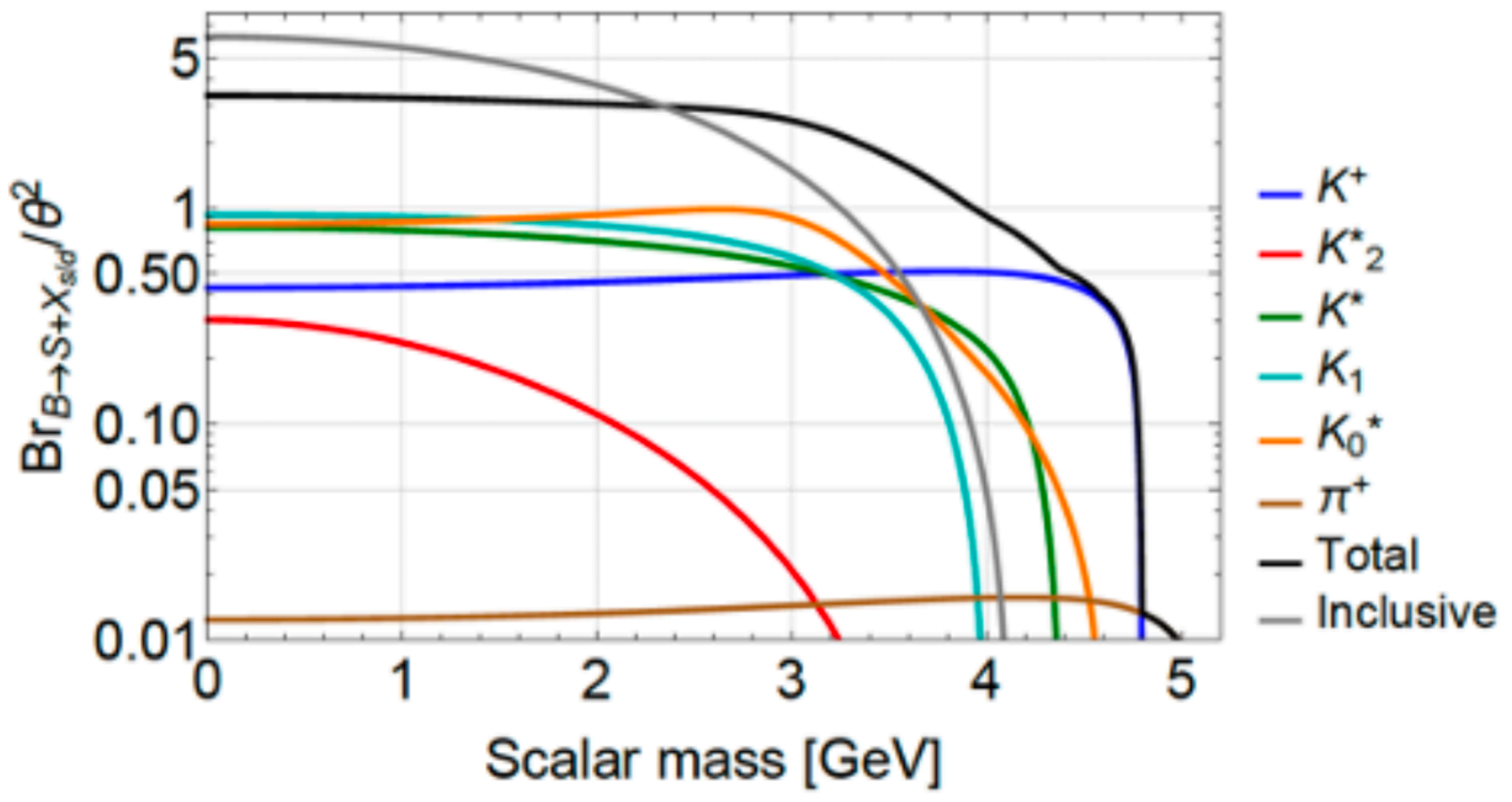
Data Samples

- Data collected in September and October of 2024 (corresponding to c3 and c4 sprucing blocks)
  - c3: μ = 4.4, *hlt1_pp_forward_then_matching*, Moore v55r12, **2.1** fb-1
  - c4: μ = 5.3, *hlt1_pp_forward_then_matching_and_downstream_200kHz*, Moore v55r13, **1.2** fb-1
- In c3 unblinded **104.5** pb^{-1} (304800:304902) and in c4 **154.7** pb^{-1} (308245:308335) for developing the analysis



Monte Carlo Samples

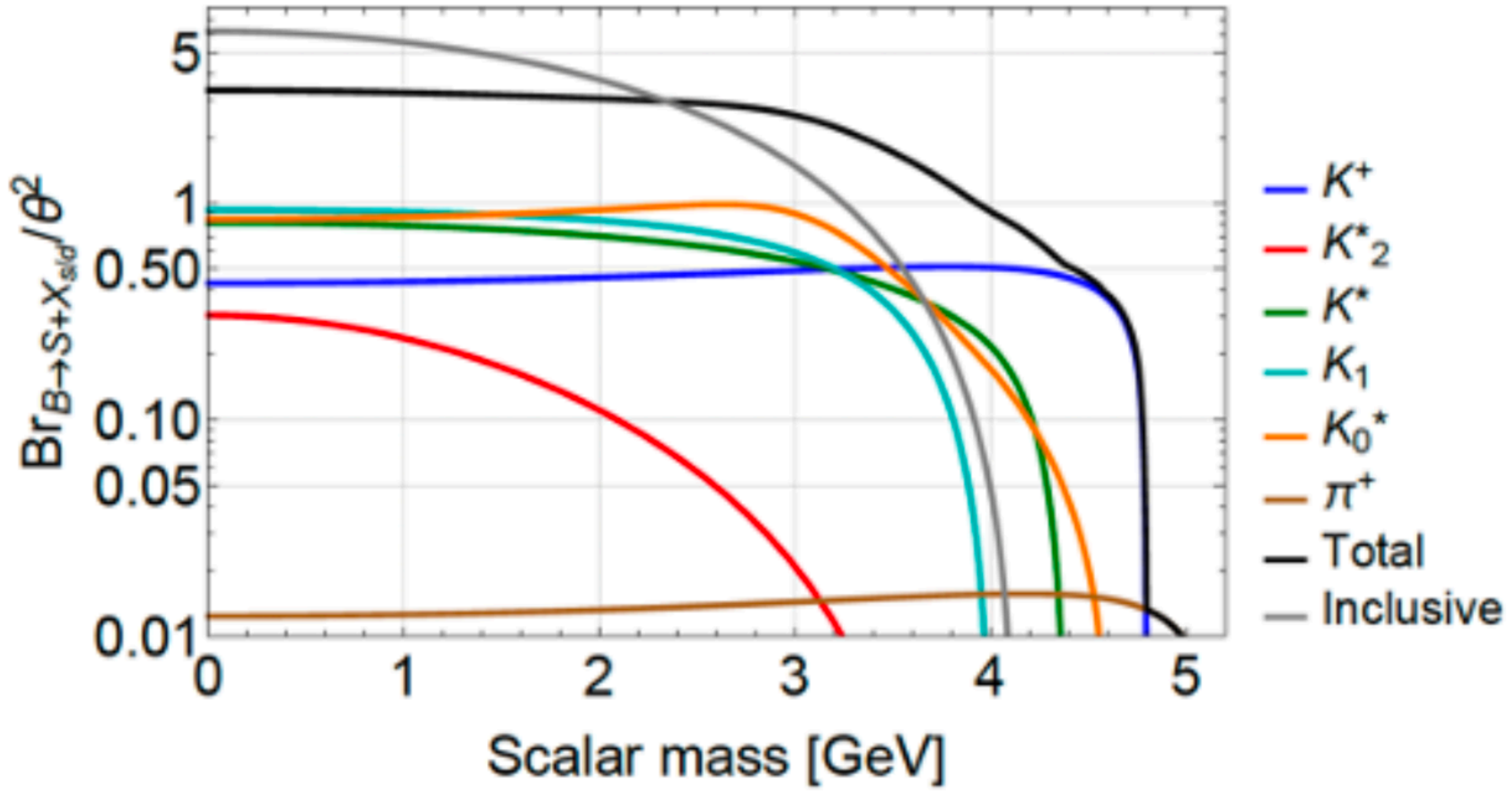
- Main exclusive decays dominating the production process to be simulated in order to study the variation of the HLT1 efficiency, as well as any variations of the kinematics of the φ particle (expected to be small)
- Cocktail decfile being prepared, relative fractions to be reweighed after production [Decfiles!2294]
- B^+ and B^0, scalars with a lifetime of 3 ns and masses of 0.25, 0.5, 0.75, 1, 1.5, 2.5, 3.5, 4, 4.5 GeV, depending on the accompanying kaon
  - Lifetime can be reweighed offline using $w = \tau_{gen}/\tau_{new} * \exp(-1/\tau_{new} + 1/\tau_{gen}) * t$



Process	$BR(m_\phi = 0)/\theta^2$	closing mass [GeV]	generated events per mass value
$B^+ \to \phi K_1^+(1270)$	$9.1^{+3.6}_{-4.0} \cdot 10^{-1}$	3.82	100k
$B^+ \to \phi K_0^{*,+}(700)$	$7.6 \cdot 10^{-1}$	4.27	–
$B^+ \to \phi K^{*,+}(892)$	$4.7^{+0.9}_{-0.8} \cdot 10^{-1}$	4.29	200k
$B^+ \to \phi K^+$	$4.3^{+1.1}_{-1.0} \cdot 10^{-1}$	4.79	100K
$B^+ \to \phi K_2^{*,+}(1430)$	$3.0 \cdot 10^{-1}$	3.85	100k
$B^+ \to \phi K^{*,+}(1410)$	$2.1^{+0.6}_{-1.1} \cdot 10^{-1}$	3.57	100k
$B^+ \to \phi K^{*,+}(1680)$	$1.3^{+0.5}_{-0.4} \cdot 10^{-1}$	3.26	100k
$B^+ \to \phi K_0^{*,+}(1430)$	$8.1 \cdot 10^{-2}$	3.82	100k
$B^+ \to \phi K_1^+(1400)$	$1.6^{+0.6}_{-1.1} \cdot 10^{-2}$	2.28	100k
$B^+ \to \phi \pi^+$	$1.3^{+0.3}_{-0.3} \cdot 10^{-2}$	5.14	100k

Monte Carlo Samples

- Some locally generated samples used for development, $B^0 \to K^*(892)(\to K\pi)H'(\to \mu\mu)$
  - Same as for plots shown in FIPs@LHCb workshop
- Corresponding to c4 conditions
  - dddb tag: dddb-20240427
  - conddb tag: sim10-2024.Q3.4-v1.2-mu100
  - geometry version: run3/2024.Q1.2-v00.00
- These samples used for early studies and efficiency estimates (shown in later plots)



Process	$BR(m_\phi = 0)/\theta^2$	closing mass [GeV]	generated events per mass value
$B^+ \to \phi K_1^+(1270)$	$9.1^{+3.6}_{-4.0} \cdot 10^{-1}$	3.82	100k
$B^+ \to \phi K_0^{*,+}(700)$	$7.6 \cdot 10^{-1}$	4.27	–
$B^+ \to \phi K^{*,+}(892)$	$4.7^{+0.9}_{-0.8} \cdot 10^{-1}$	4.29	200k
$B^+ \to \phi K^+$	$4.3^{+1.1}_{-1.0} \cdot 10^{-1}$	4.79	100K
$B^+ \to \phi K_2^{*,+}(1430)$	$3.0 \cdot 10^{-1}$	3.85	100k
$B^+ \to \phi K^{*,+}(1410)$	$2.1^{+0.6}_{-1.1} \cdot 10^{-1}$	3.57	100k
$B^+ \to \phi K^{*,+}(1680)$	$1.3^{+0.5}_{-0.4} \cdot 10^{-1}$	3.26	100k
$B^+ \to \phi K_0^{*,+}(1430)$	$8.1 \cdot 10^{-2}$	3.82	100k
$B^+ \to \phi K_1^+(1400)$	$1.6^{+0.6}_{-1.1} \cdot 10^{-2}$	2.28	100k
$B^+ \to \phi \pi^+$	$1.3^{+0.3}_{-0.3} \cdot 10^{-2}$	5.14	100k

Normalisation

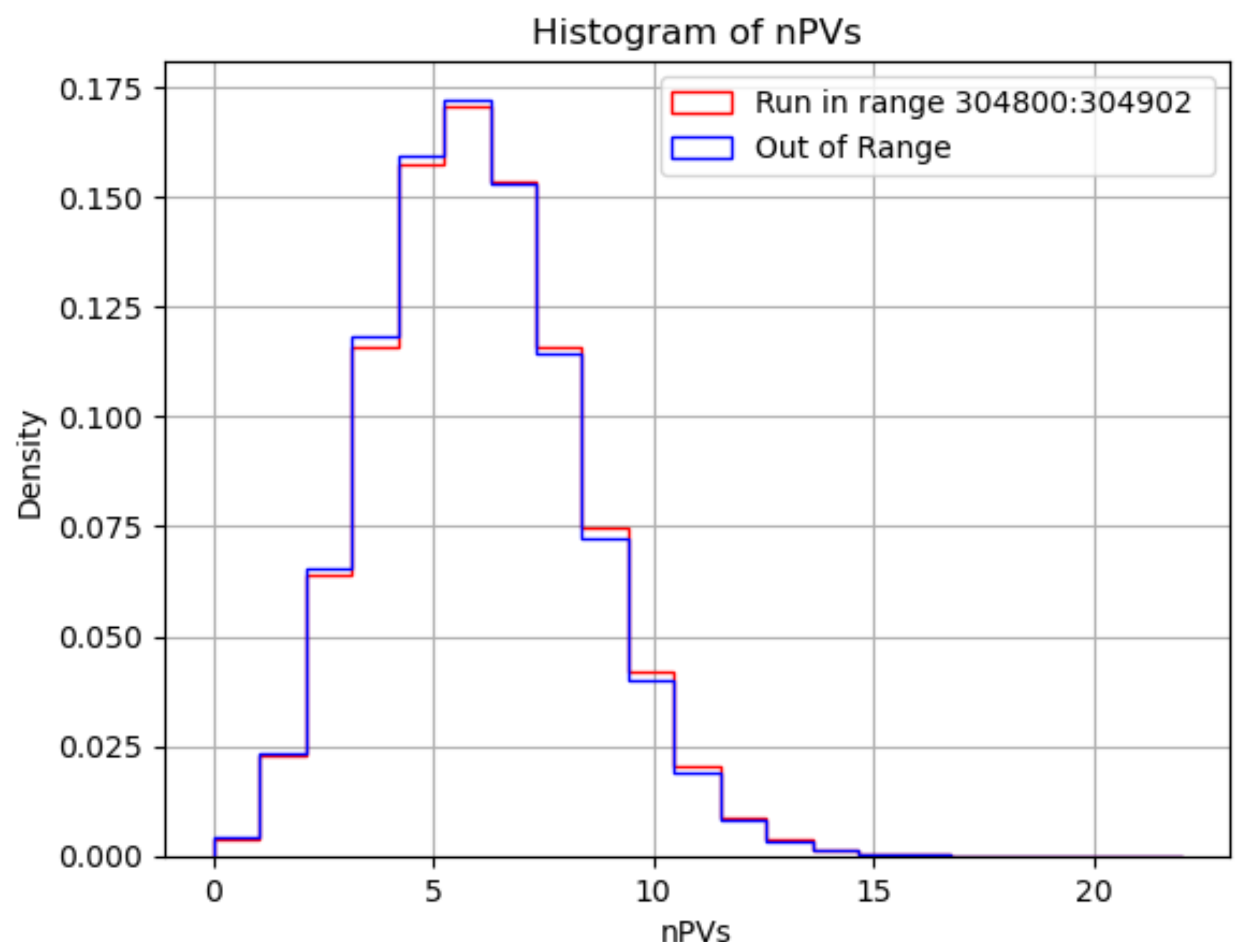
Normalisation

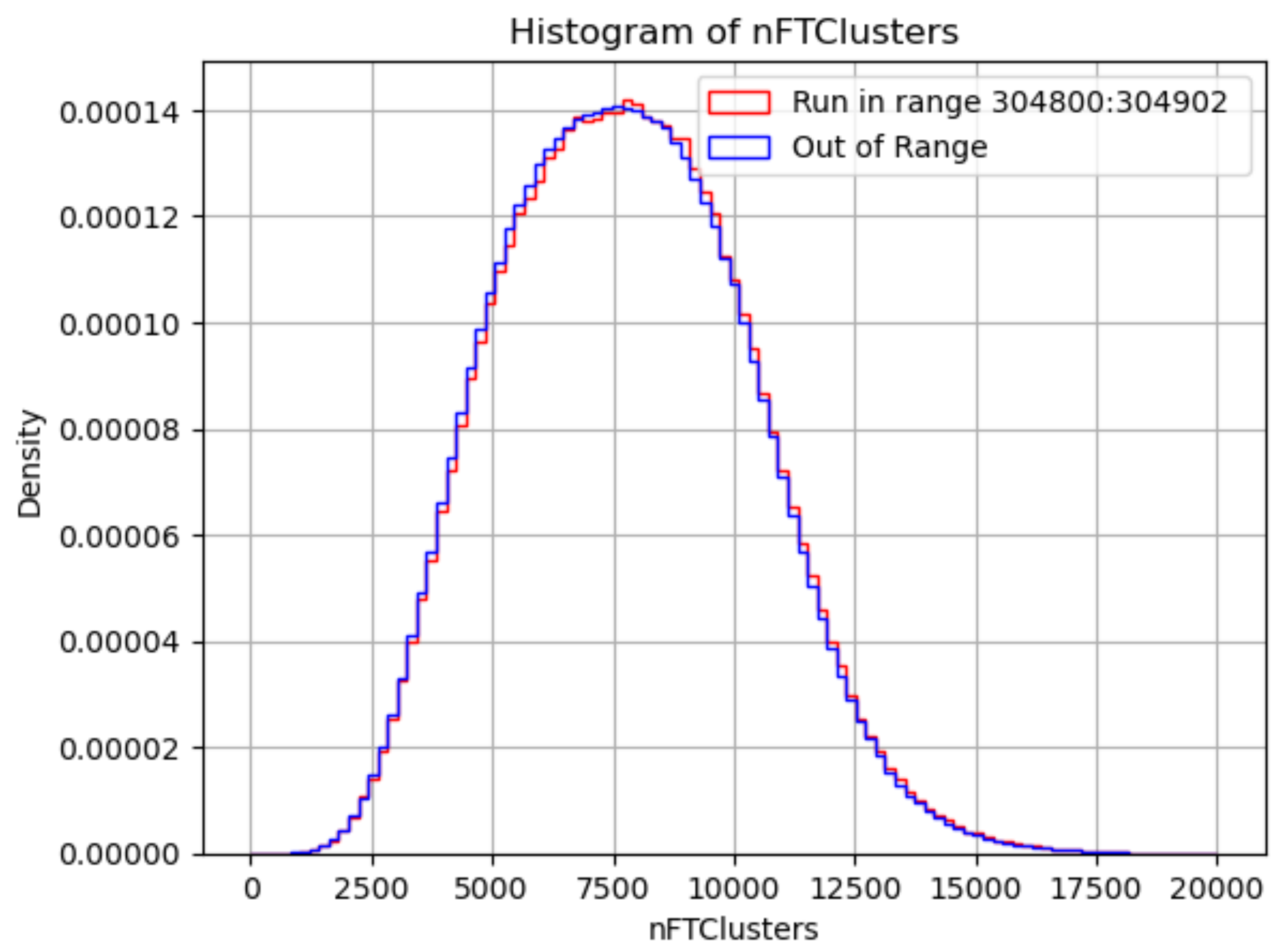
- Normalise to $B \to J/\psi(\to \mu\mu)K^0_S(\to \pi\pi)$ where the K^0_S is reconstructed using T tracks
- $b\bar{b}$ cross sections and fragmentation fractions in B meson drop out
- Number of signal events given by:
  $N_S = \mathscr{L}_{\text{int}} \times \sigma_{b\bar{b}} \times f_{B^{0,\pm}} \times BR(B^{0,\pm} \to \phi X) \times BR(\phi \to \mu^+\mu^-) \times \epsilon_{\text{sign,tot}}(\tau, m)$, where
  $\epsilon_{\text{sig,tot}}(\tau, m) = \epsilon_{\text{sig,geo}}(\tau, m) \times \epsilon_{\text{sig,HLT1-TIS}} \times \epsilon_{\text{sig,HLT2-TOS}} \times \epsilon_{\text{sig,MVA}}$, and where
  $\epsilon_{\text{sig,HLT2-TOS}} = \epsilon_{\text{sig,HLT2-recostruction}} \times \epsilon_{\text{sig,HLT2-selection}}$
- Number of control channel events given by:
  $N_{cc} = \mathscr{L}_{\text{int}} \times \sigma_{b\bar{b}} \times f_{B^0} \times BR(B^0 \to J/\Psi K_S) \times BR(J/\Psi \to \mu^+\mu^-) \times BR(K_S \to \pi^+\pi^-) \times \epsilon_{\text{cc,tot}}$, where
  $\epsilon_{\text{cc,tot}}(\tau, m) = \epsilon_{\text{cc,geo}} \times \epsilon_{\text{cc,HLT1-TIS(K}_\text{S})} \times \epsilon_{\text{cc,HLT2-TOS(K}_\text{S})} \times \epsilon_{\text{cc,offline}}$, and where
  $\epsilon_{\text{cc,HLT2-TOS(K}_\text{s})} = \epsilon_{\text{cc,HLT2-recostruction(K}_\text{s})} \times \epsilon_{\text{cc,HLT2-selection(K}_\text{s})}$

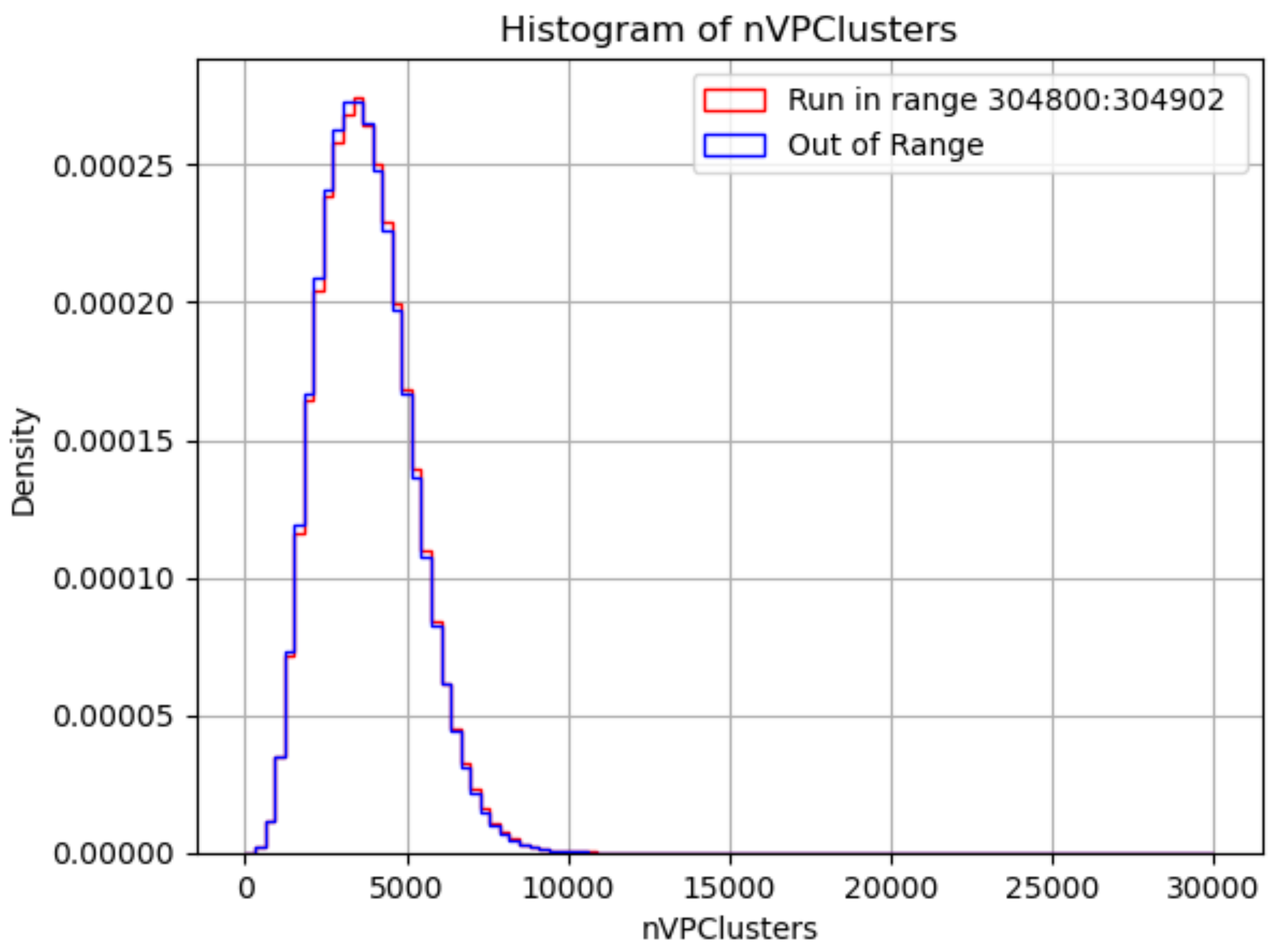
Reconstruction

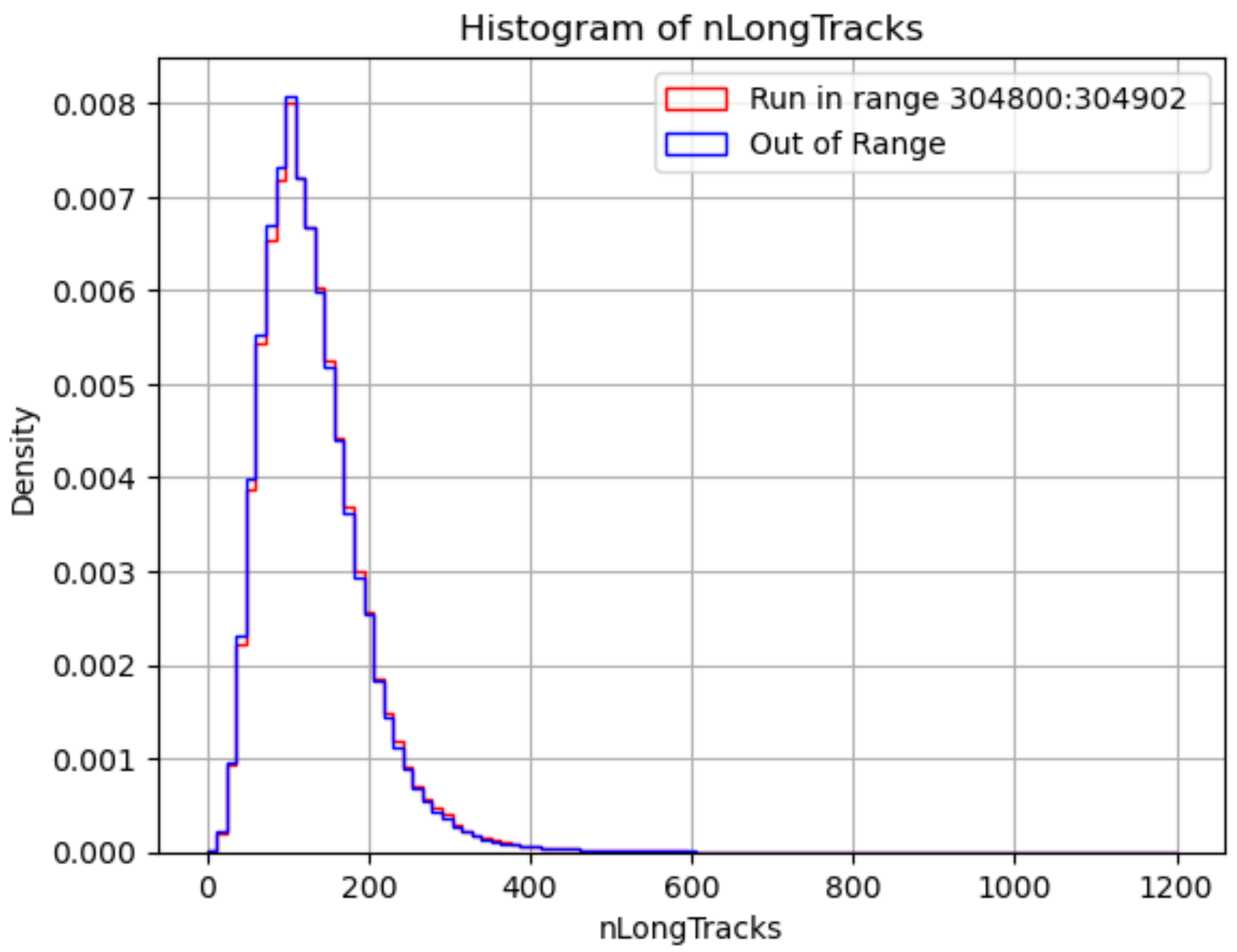
Data uniformity in un/blinded regions

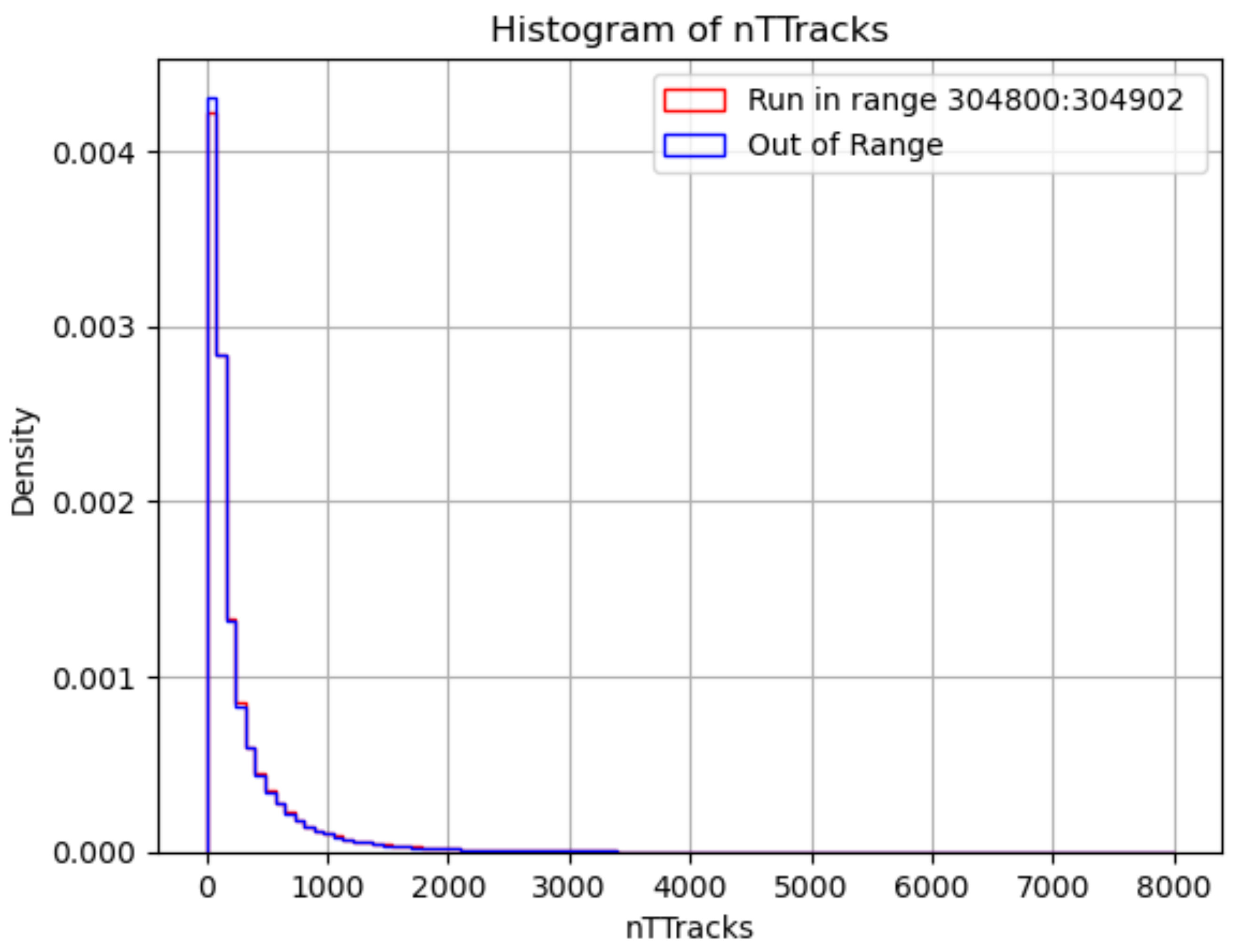
- Check unblinded region corresponds to blinded region for c3, by comparing event level variables in/out of the unblinded run range in the control sample
- Distributions agree well
- To do: check for c4, muon system variables





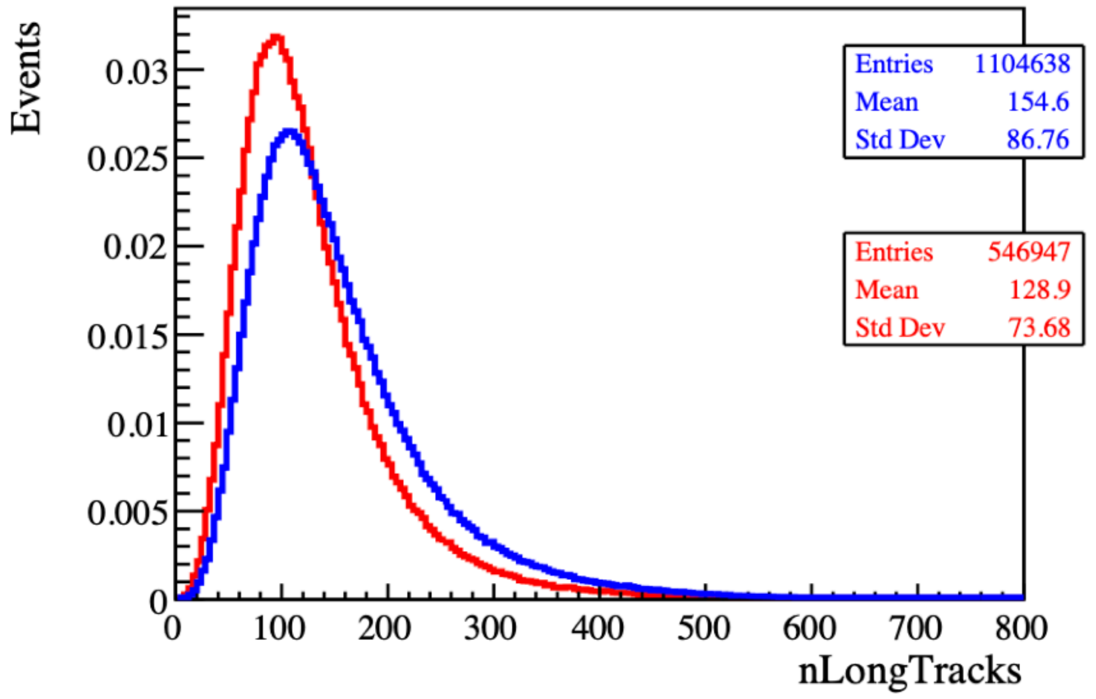


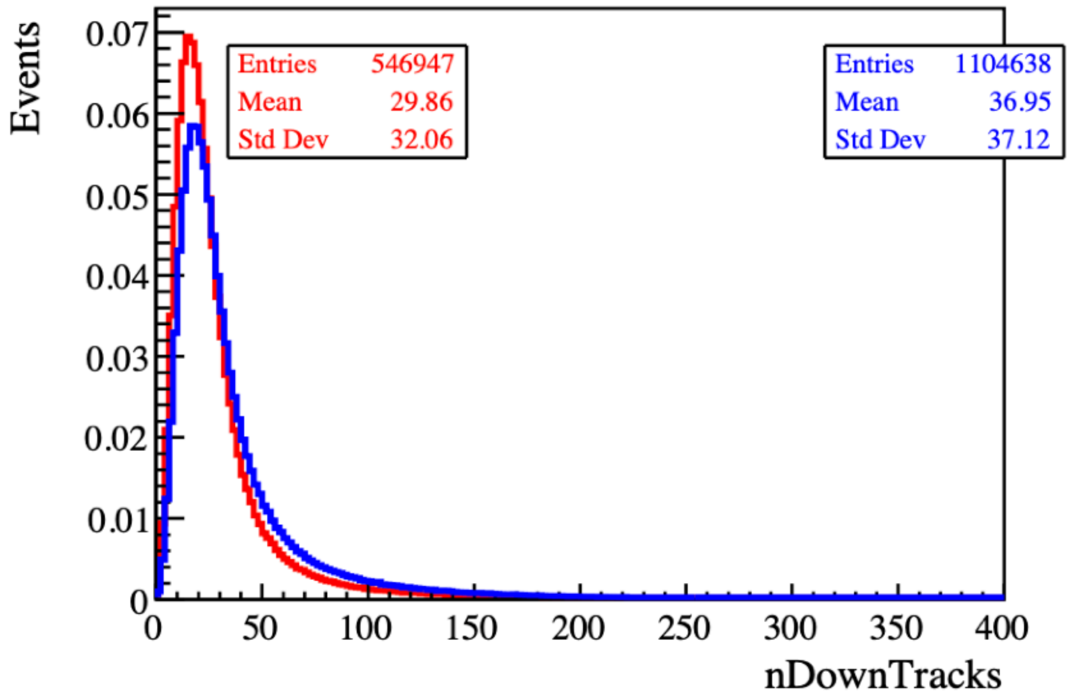


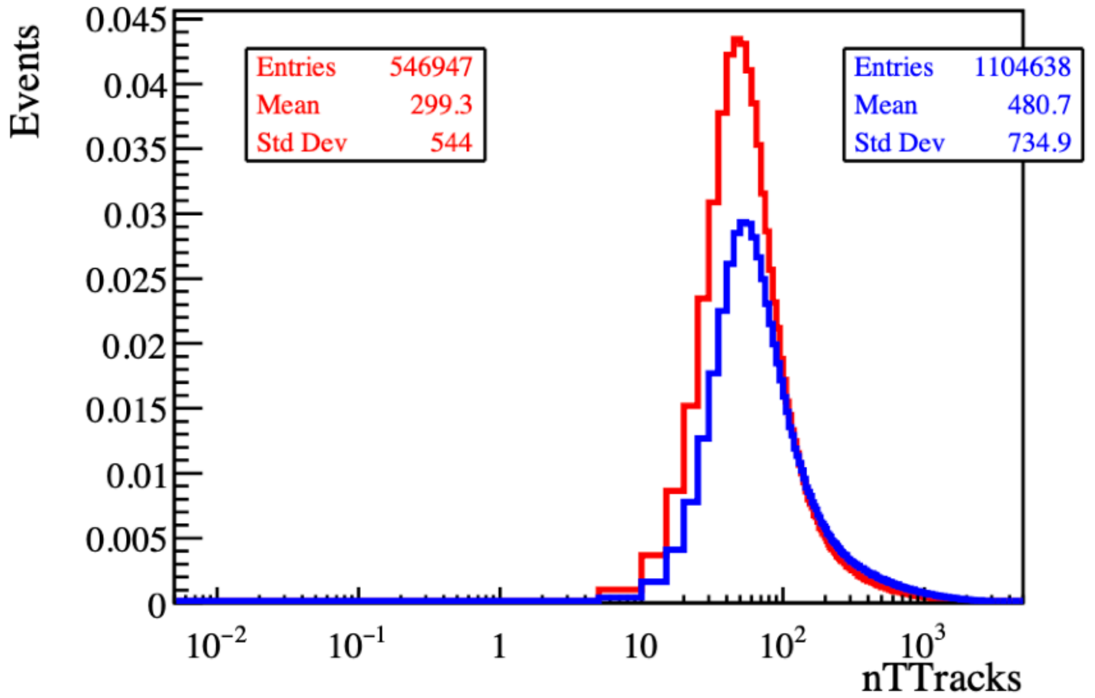


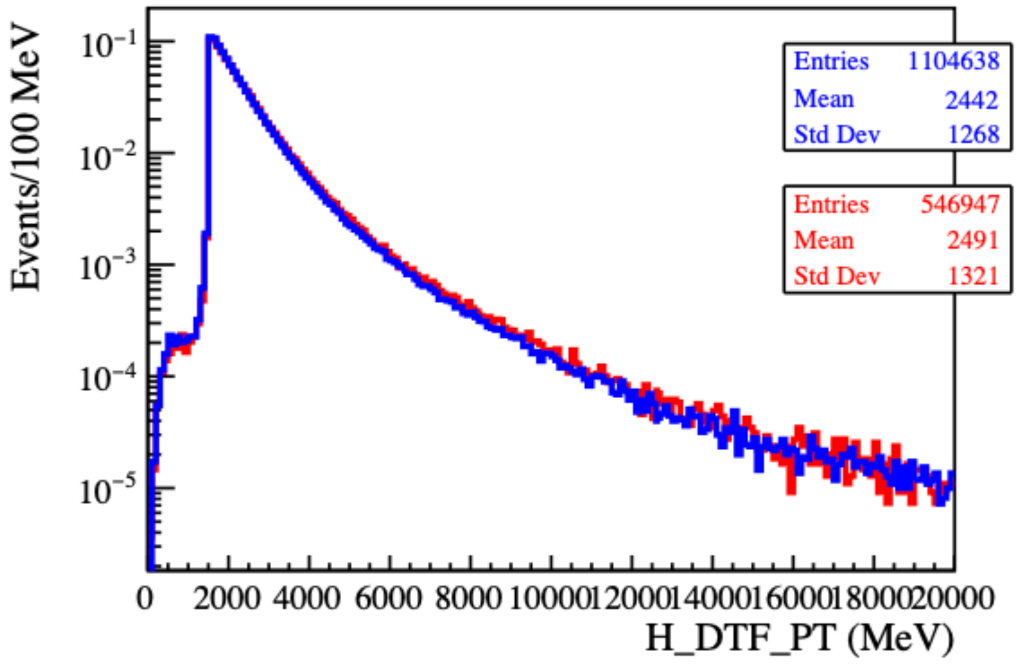
Reconstruction dependence with beam intensity

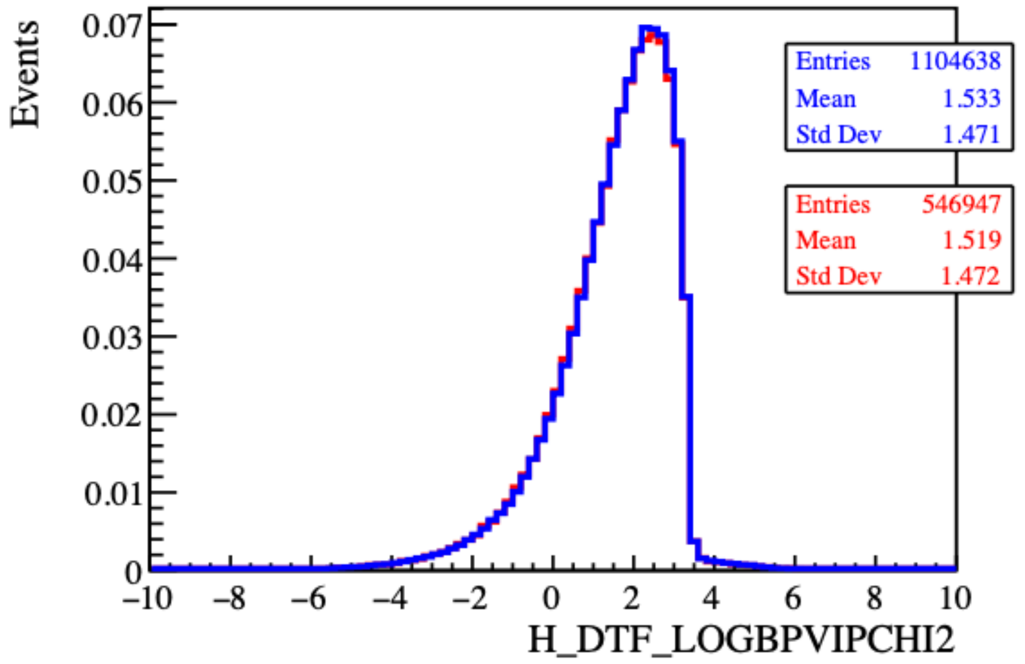
- Check event level variables in the two different blocks
  - Differ, as expected
- Check the candidate kinematics in the two different blocks
  - Show good agreement

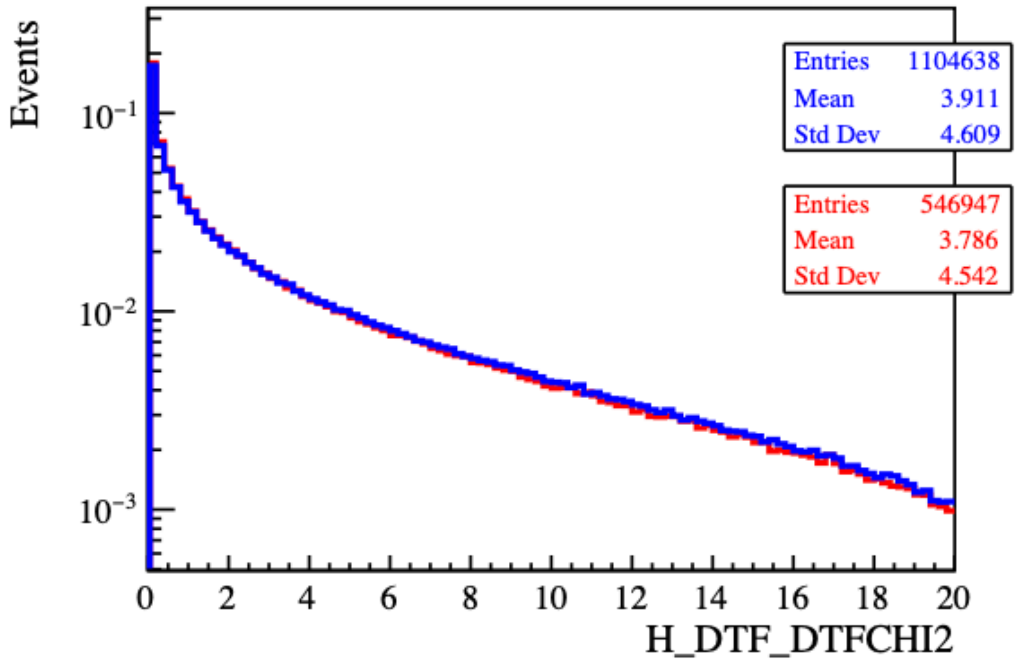


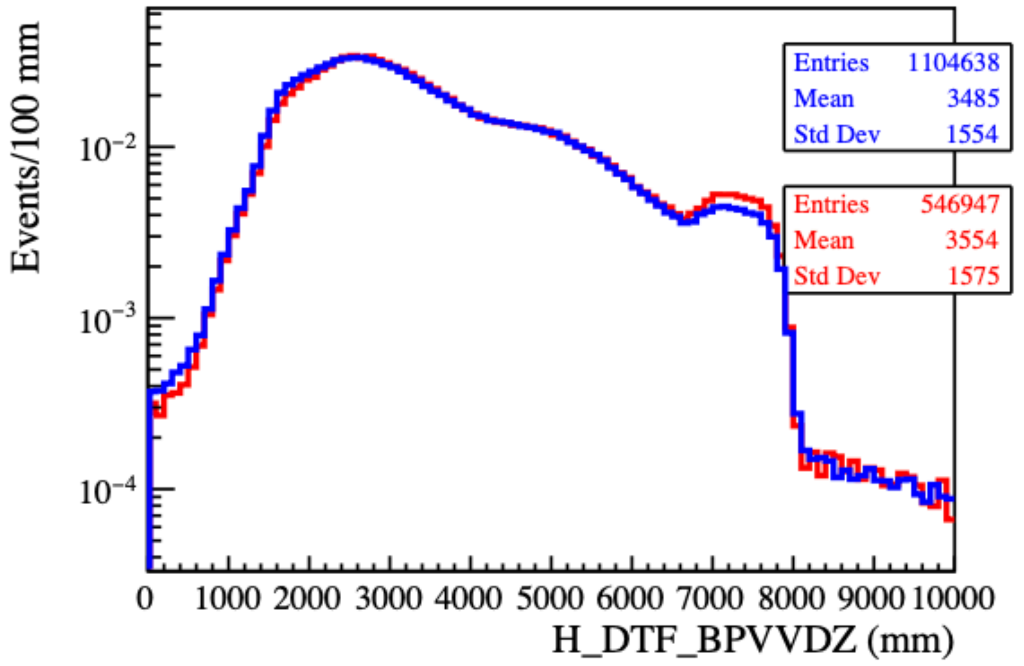


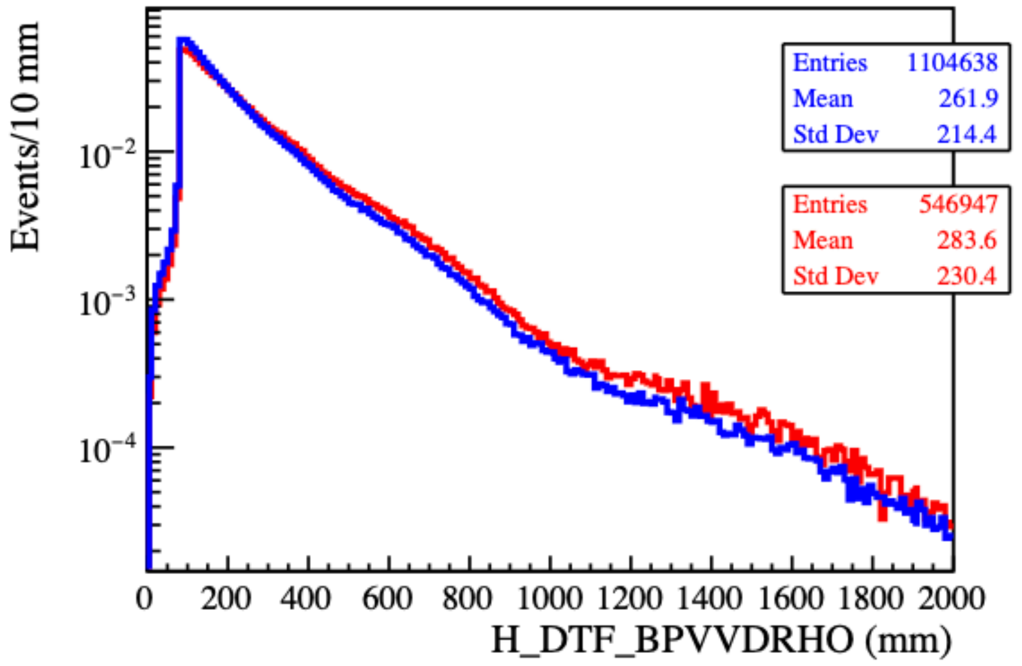


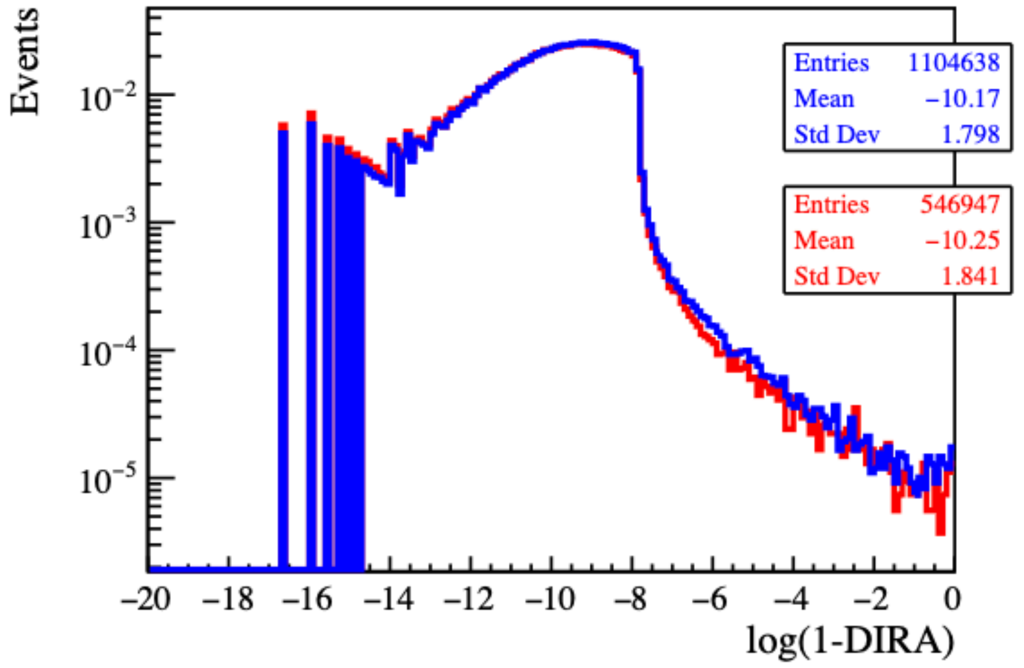






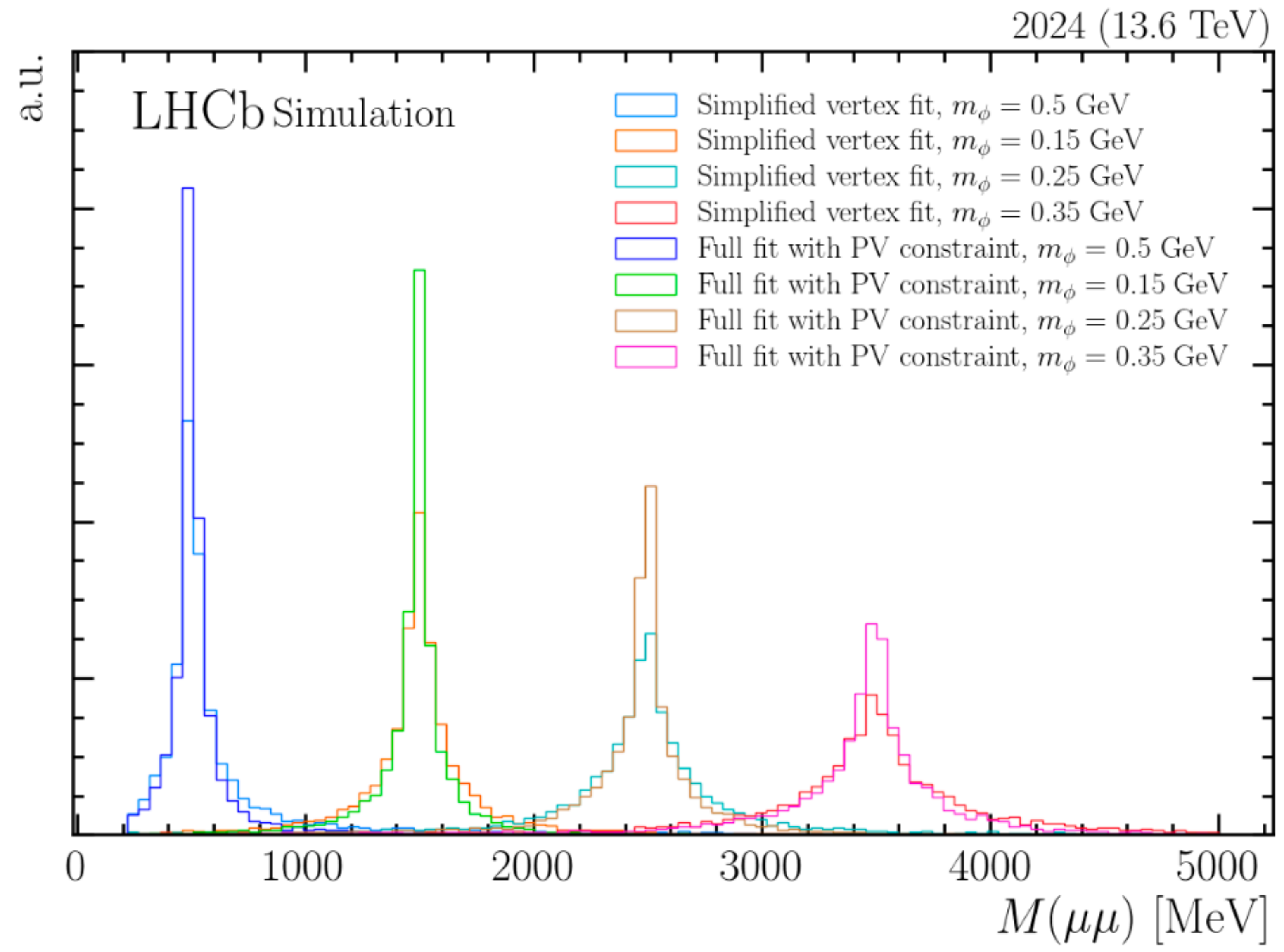


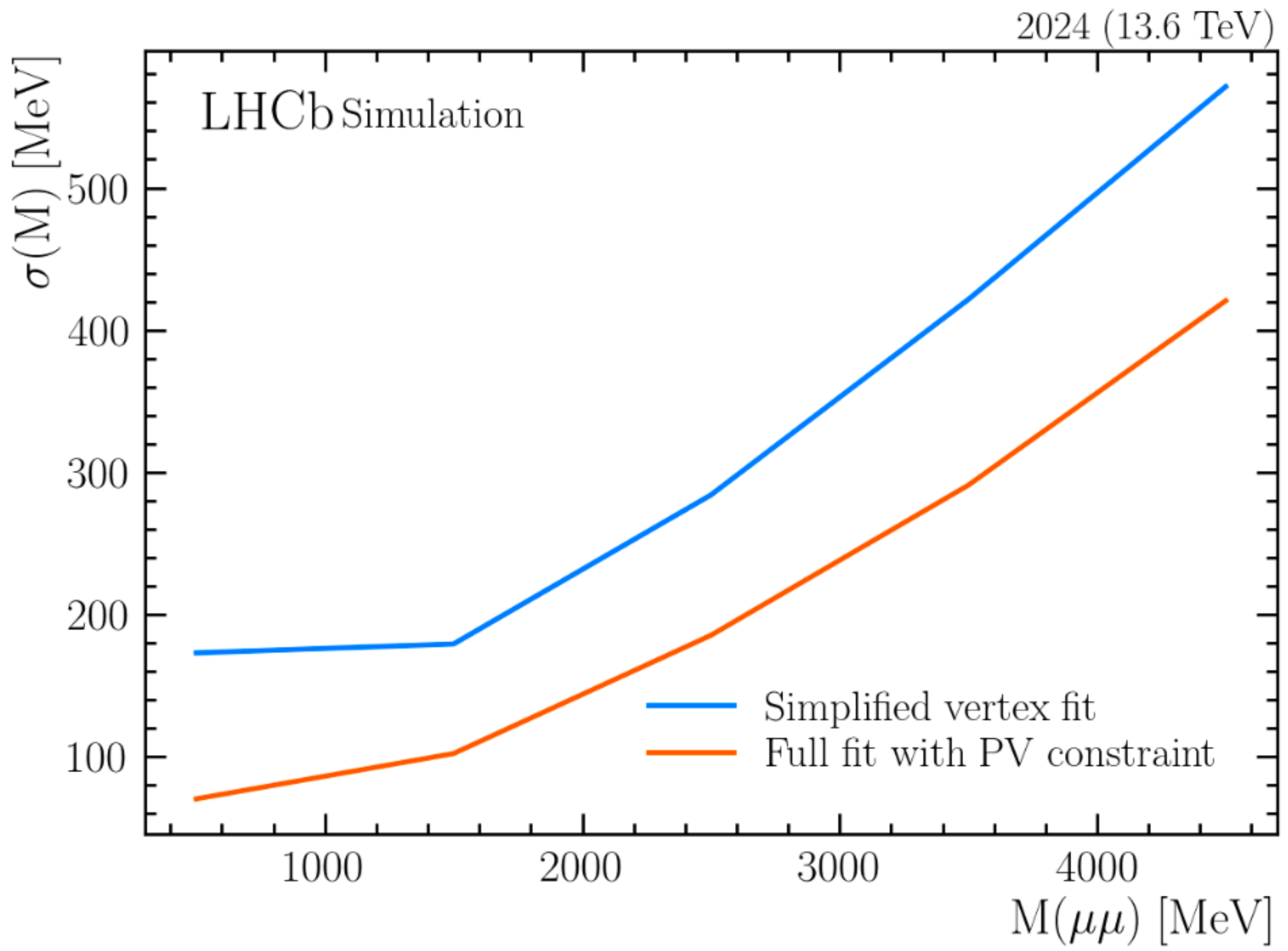




Mass resolution

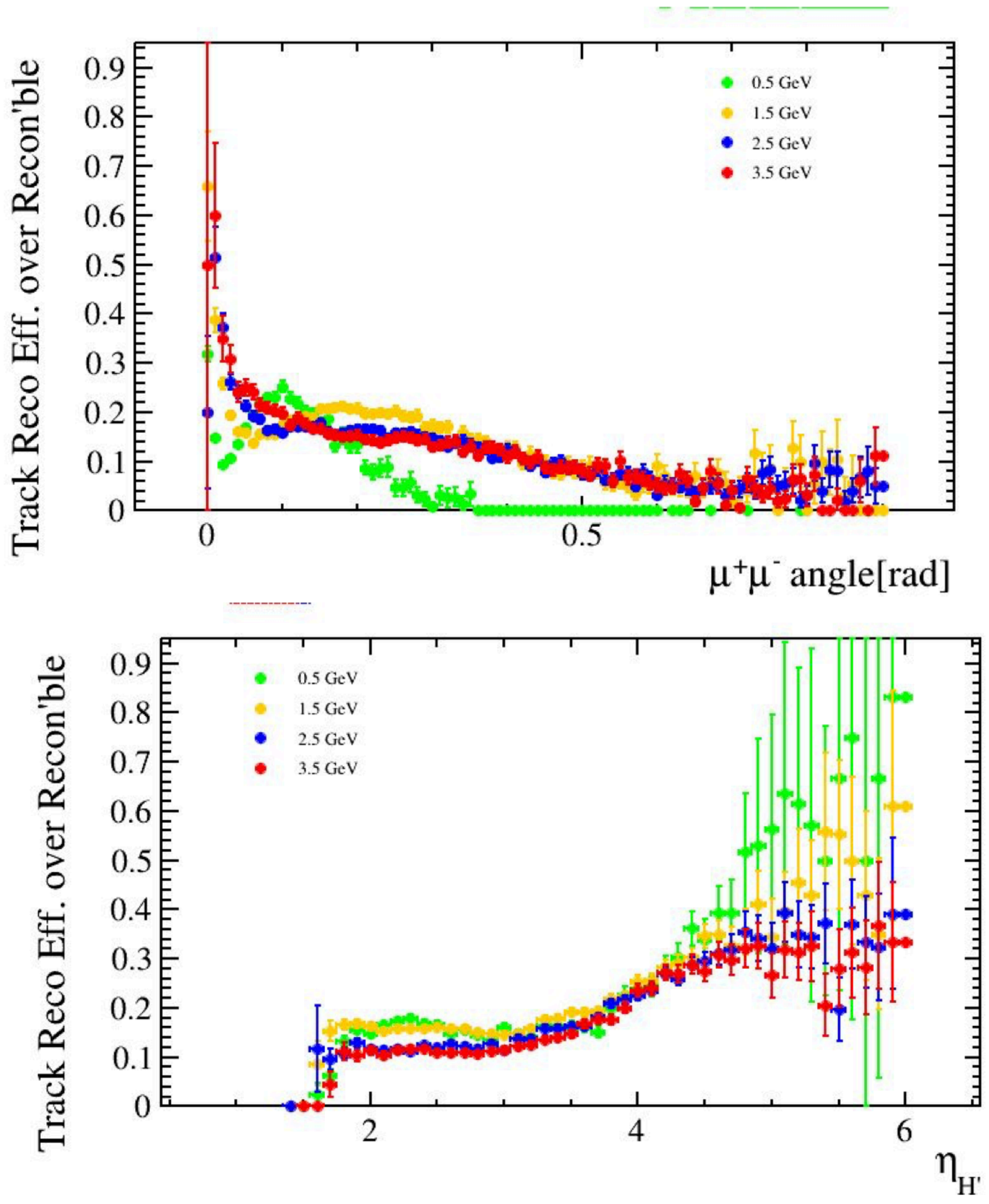
- Mass resolution degrades as the mass of the LLP increases
  - The average muon momentum increases, which leads to a lower momentum resolution
- This means the search window has to be wider as the mass increase
- The background decreases exponentially with mass, so impact of this should be reduced
- WIP: investigating ways to improve mass resolution with kinematic fitting for exclusive selection (see e.g. talk from last Friday's RTA-DPA general meeting)





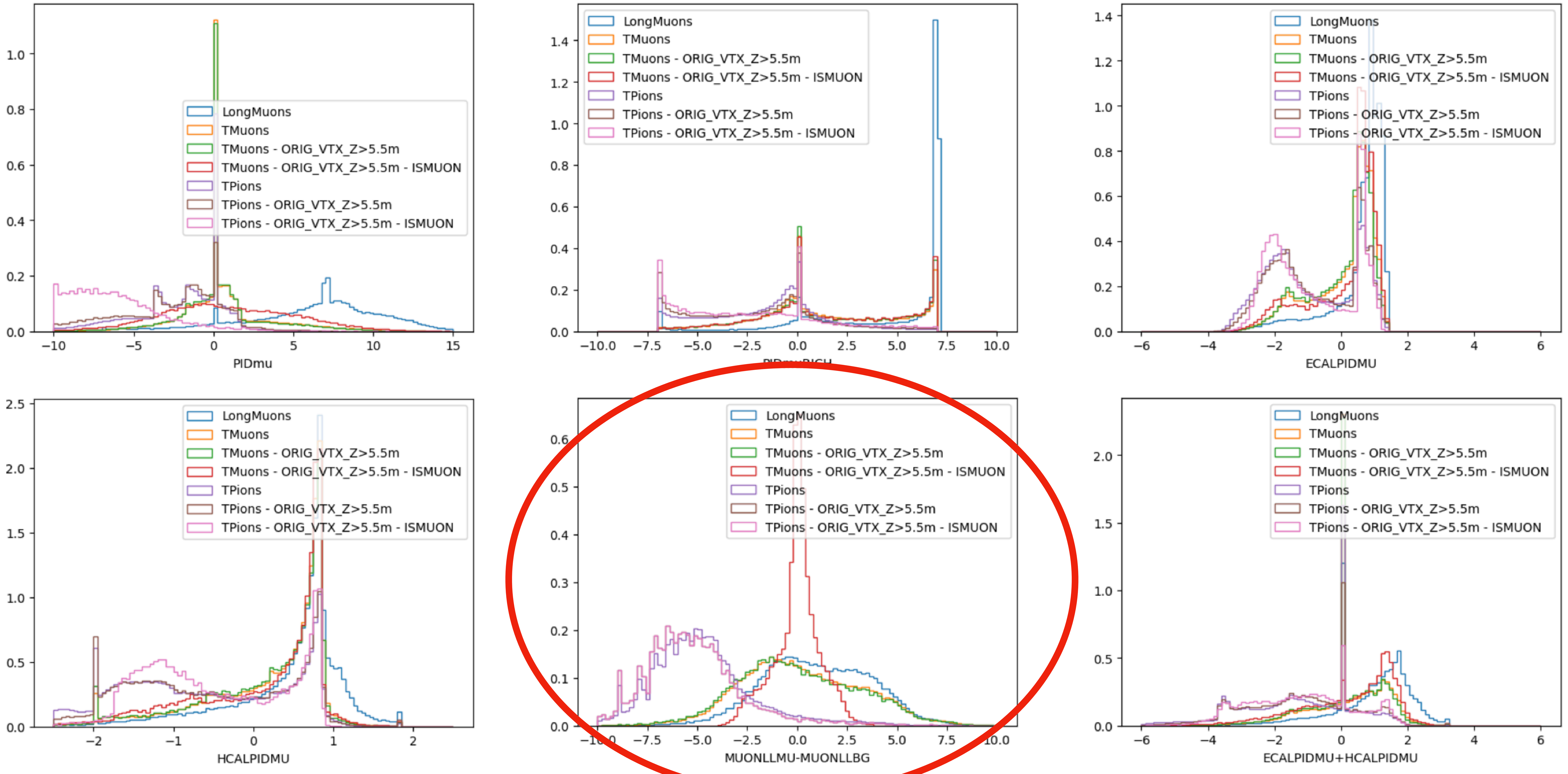
Opening angle dependence

- Assumption in the seeding that track points to origin
- Larger mass LLPs have higher Q-values, and therefore opening angles, meaning the tracks point less to the origin
- For a given opening angle, dimuon reconstruction efficiency is approximately the same



PID

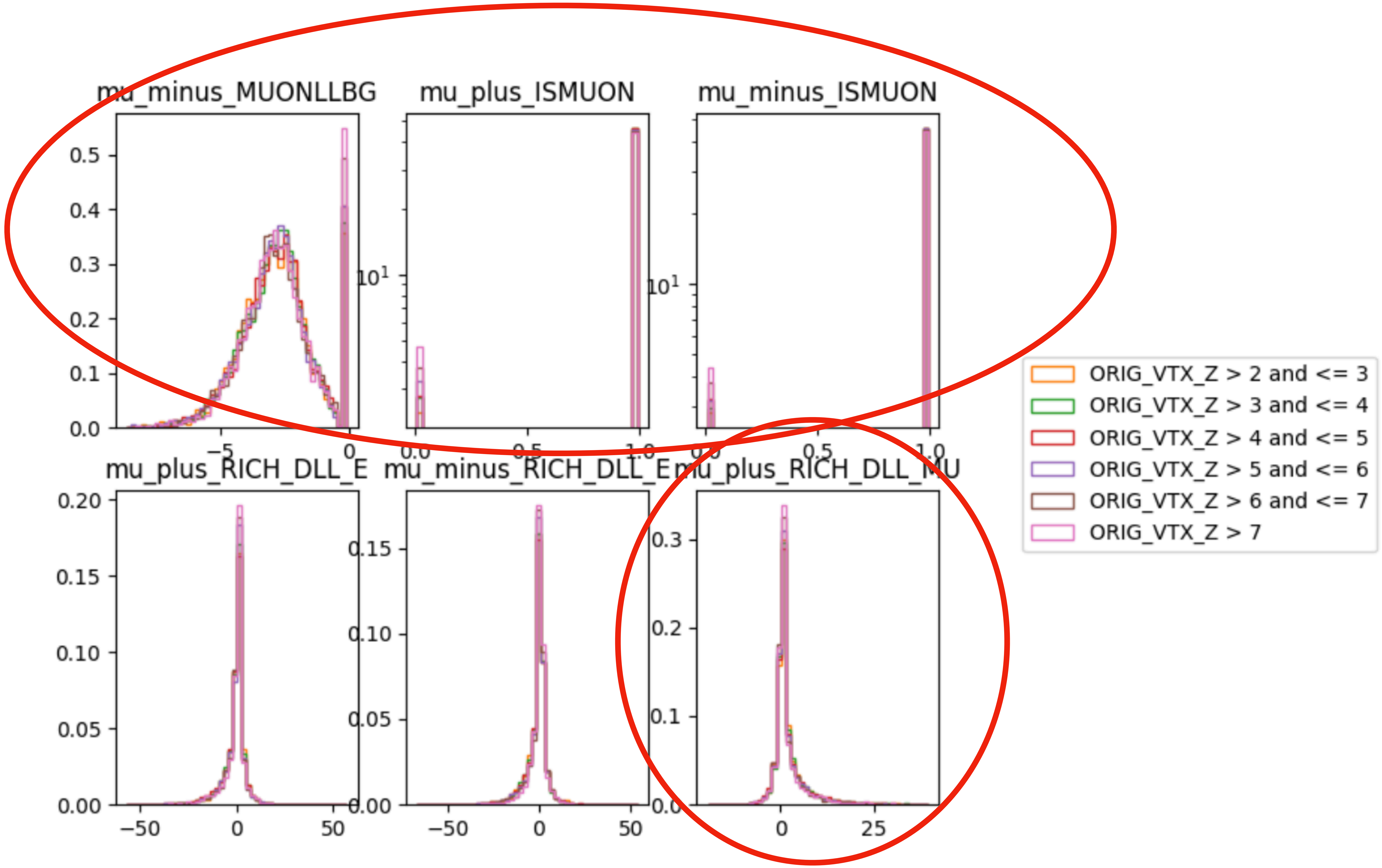
- PID less performant than for Long and Downstream, but muon ID still shows good performance
- PID calibration will be an important consideration: never done before for T tracks
- Plan is to use same calibration samples and methods for Long and/or Downstream tracks, if there is no vertex z dependence
    - e.g. identify calibration signal with Long tracks, compute PID variables for standalone T segment, then use for calibration



[Talk from Andrea 20/2/25](#) adding Calo ID to T tracks for 2025

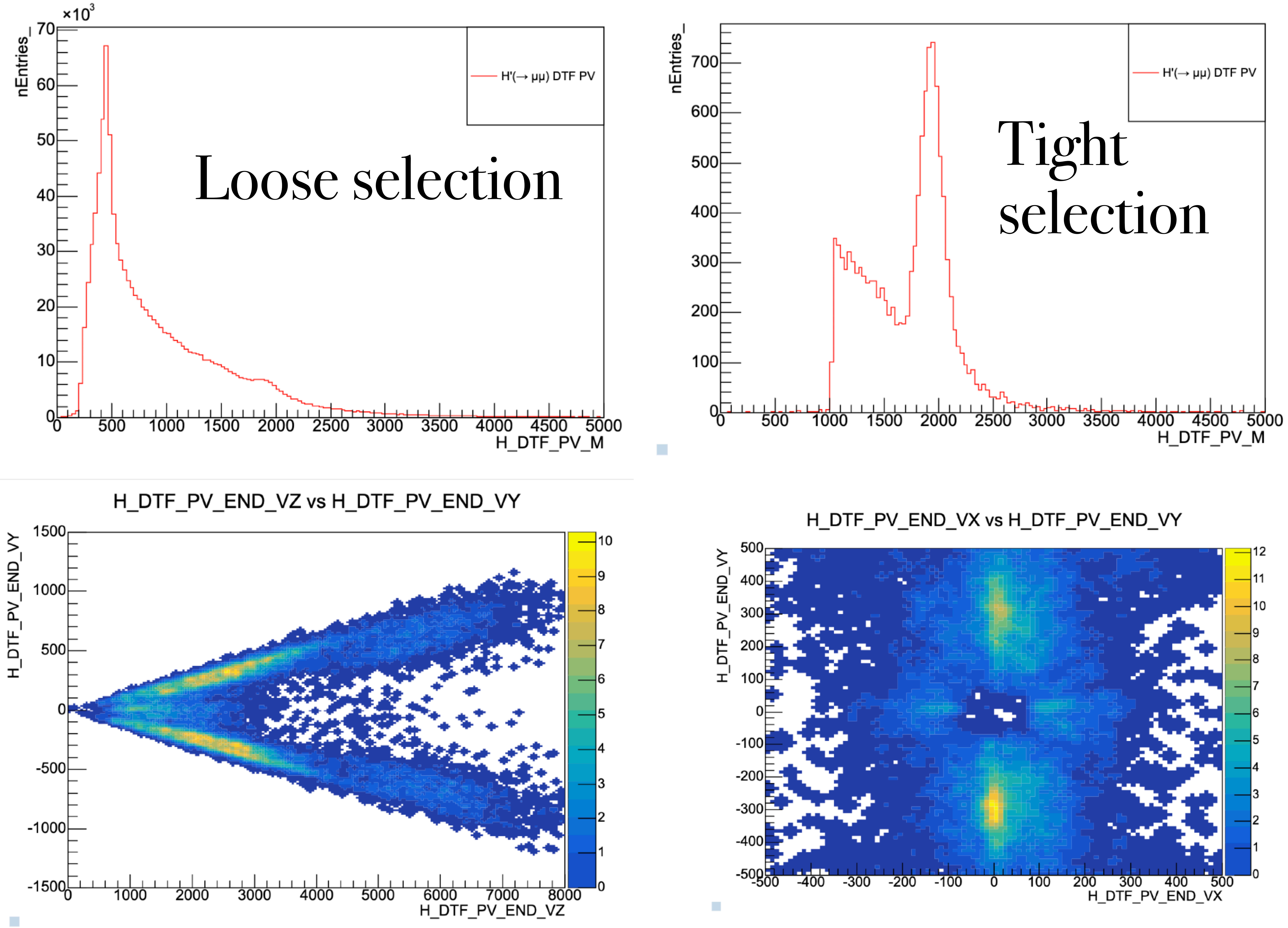
PID

- PID less performant than for Long and Downstream, but muon ID still shows good performance
- PID calibration will be an important consideration: never done before for T tracks
- Plan is to use same calibration samples and methods for Long and/or Downstream tracks, if there is no vertex z dependence
    - e.g. identify calibration signal with Long tracks, compute PID variables for standalone T segment, then use for calibration



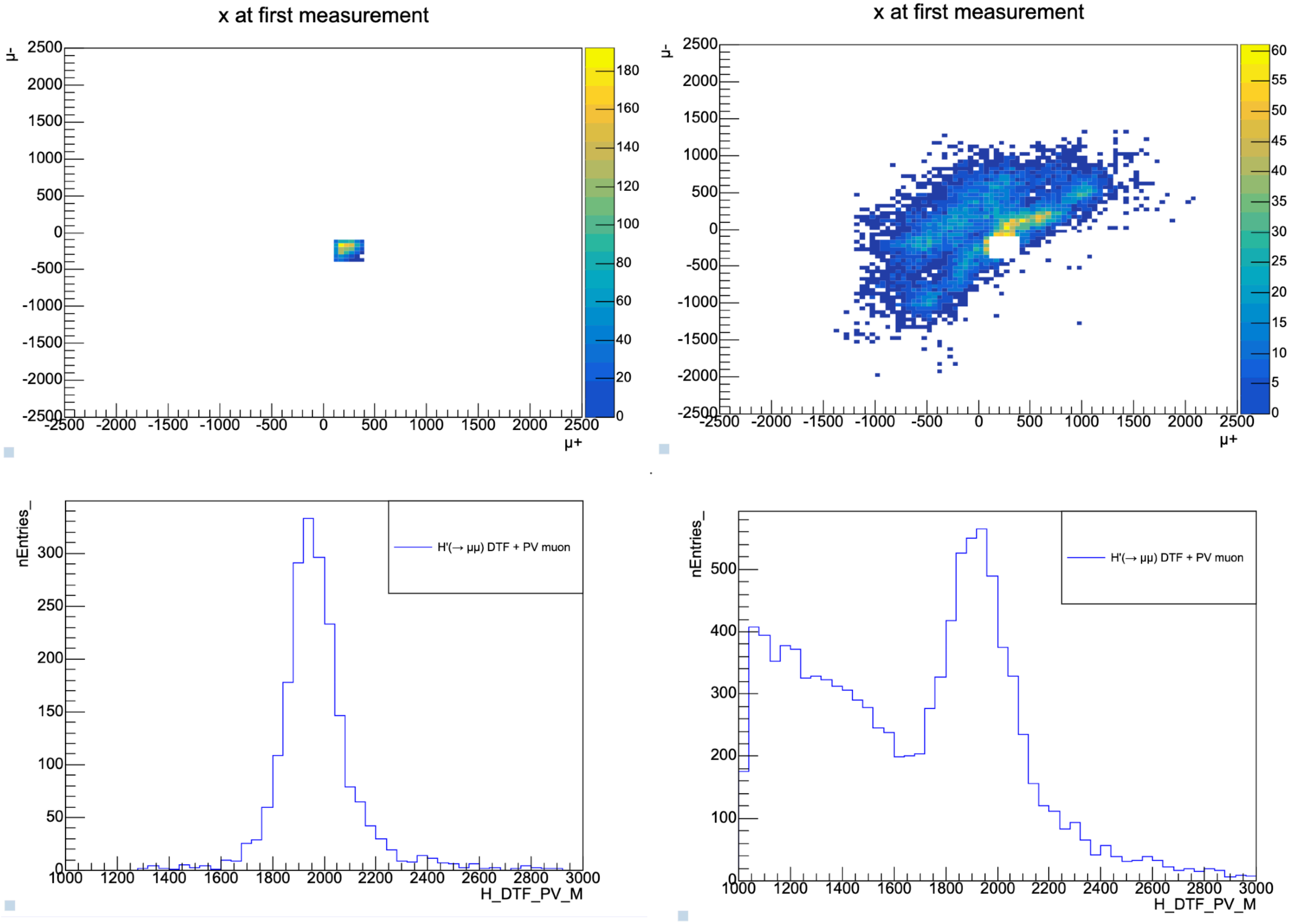
Mystery bump

- Previously reported on unexpected bump in the data
- Corresponds to hot spot in x of tracks at first measurement, but doesn't disappear when removing hotspot
- Could be a reconstruction artefact → to be further investigated



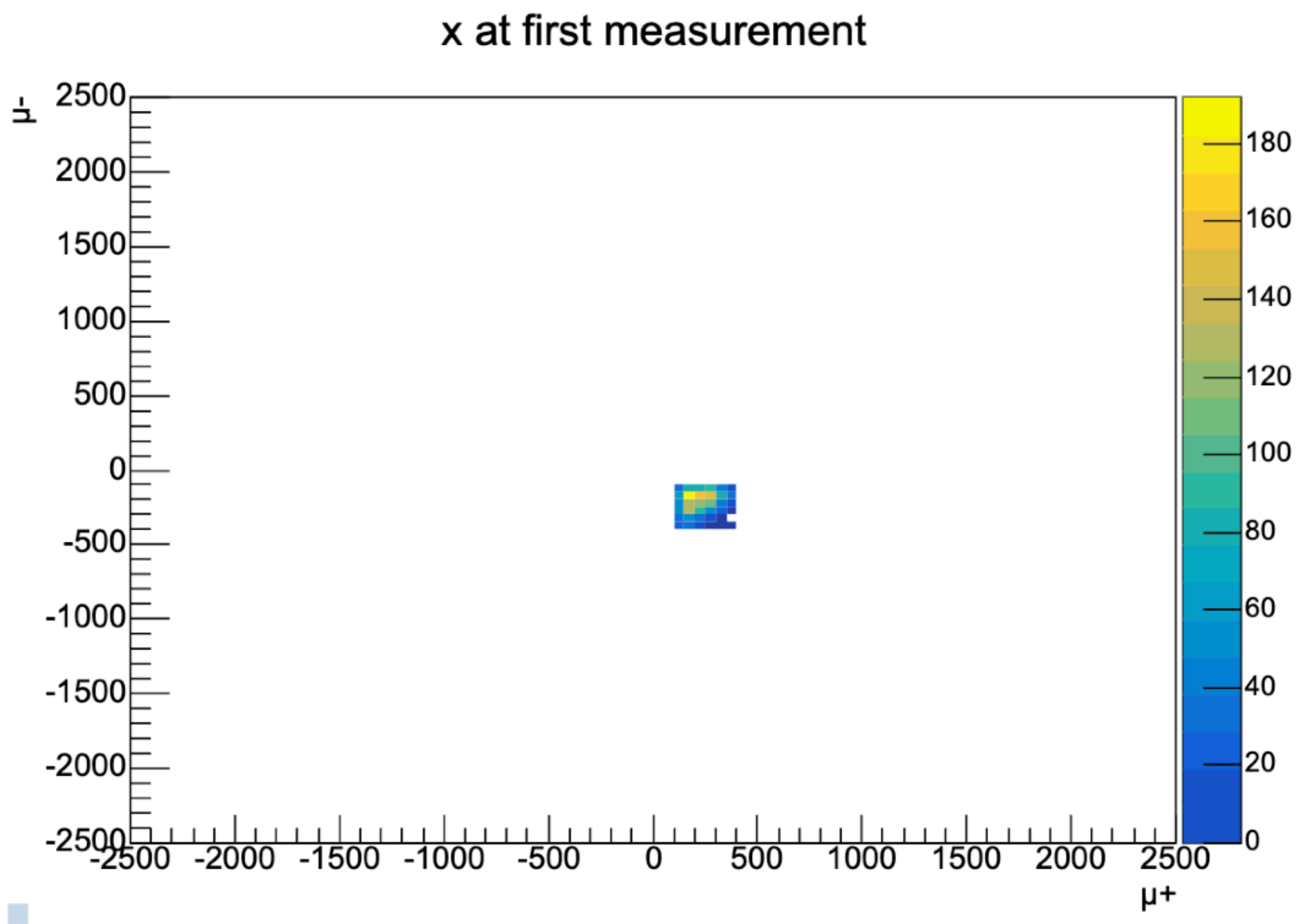
Mystery bump

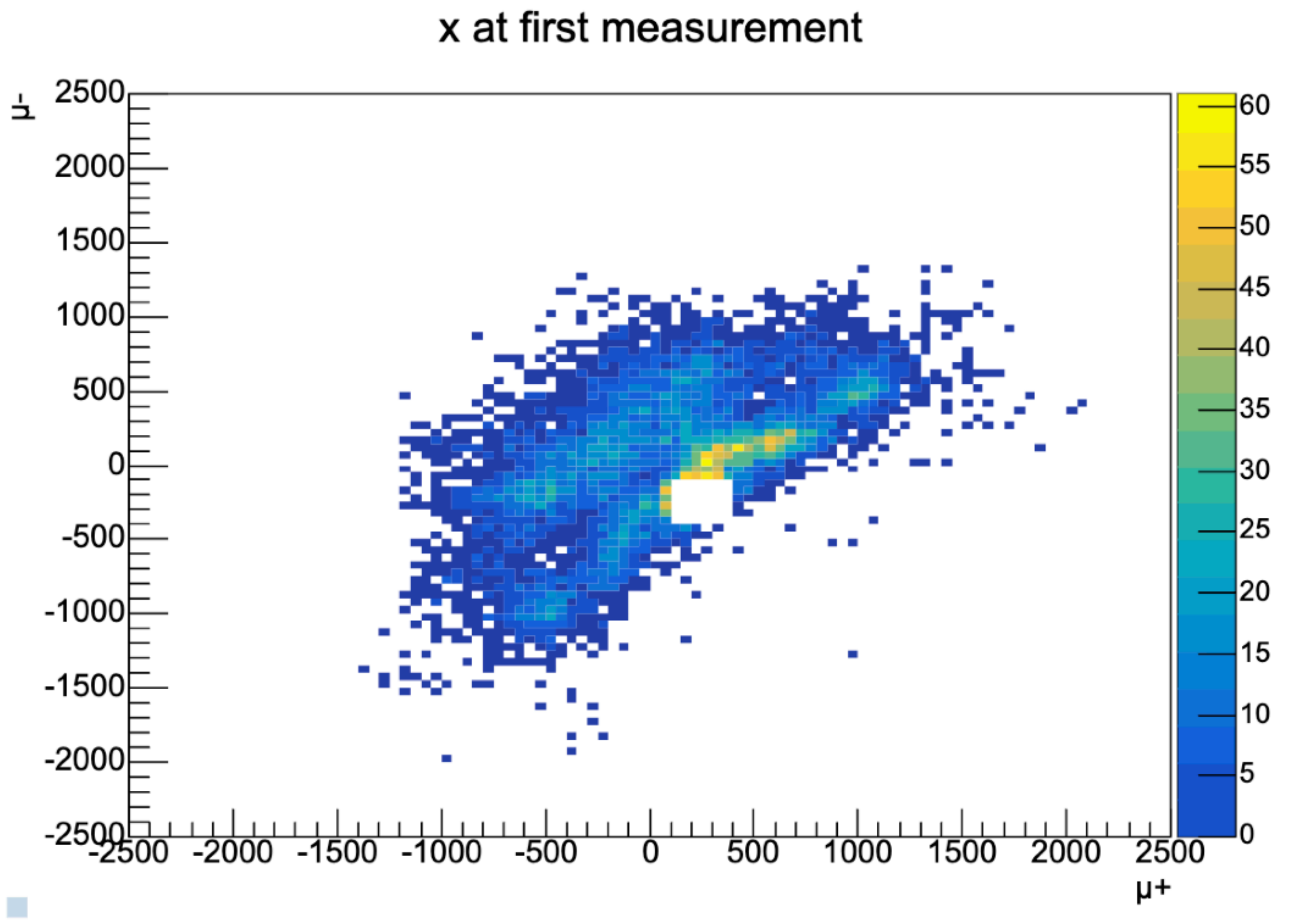
- Previously reported on unexpected bump in the data
- Corresponds to hot spot in x of tracks at first measurement, but doesn't disappear when removing hotspot
- Could be a reconstruction artefact → to be further investigated

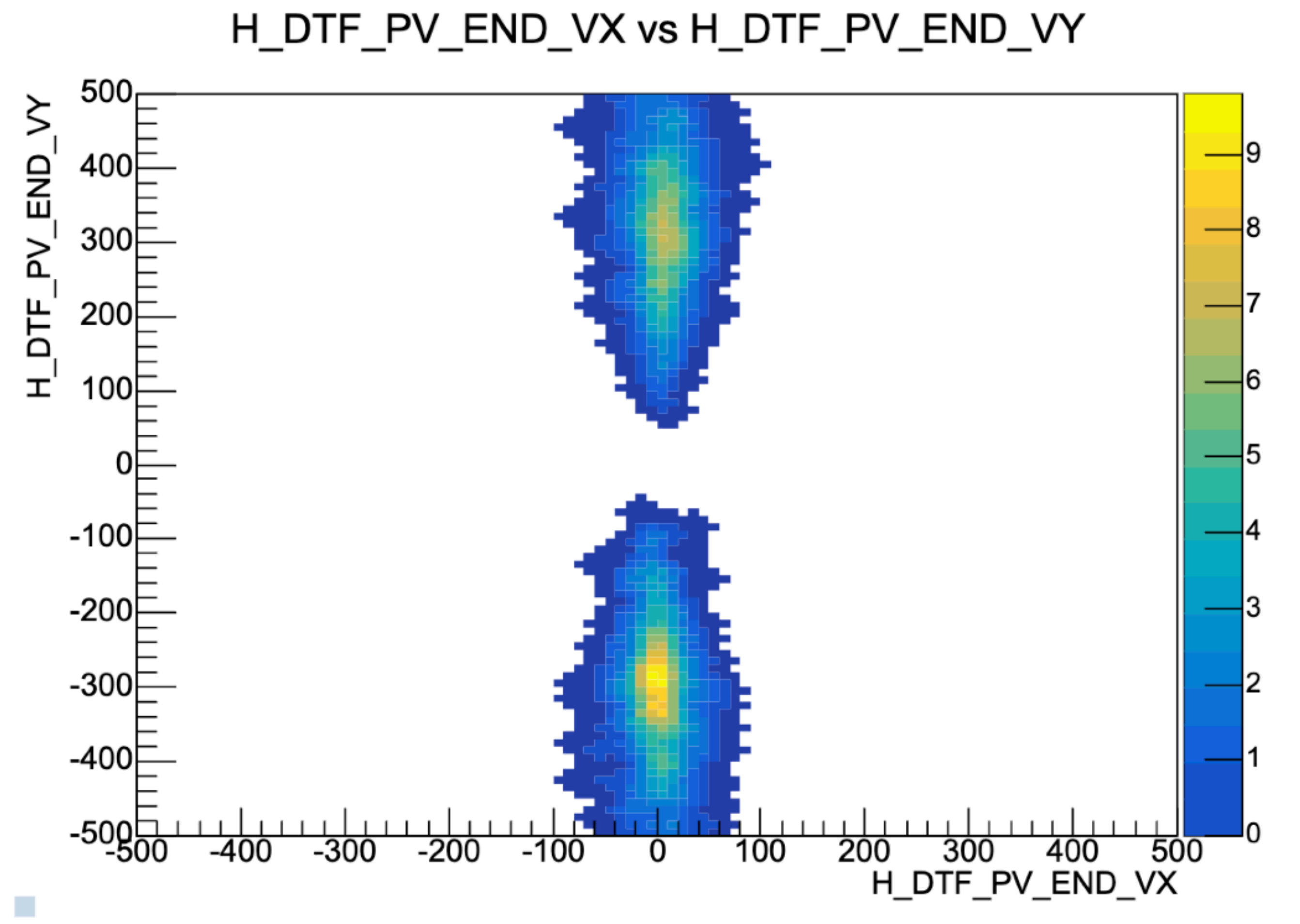


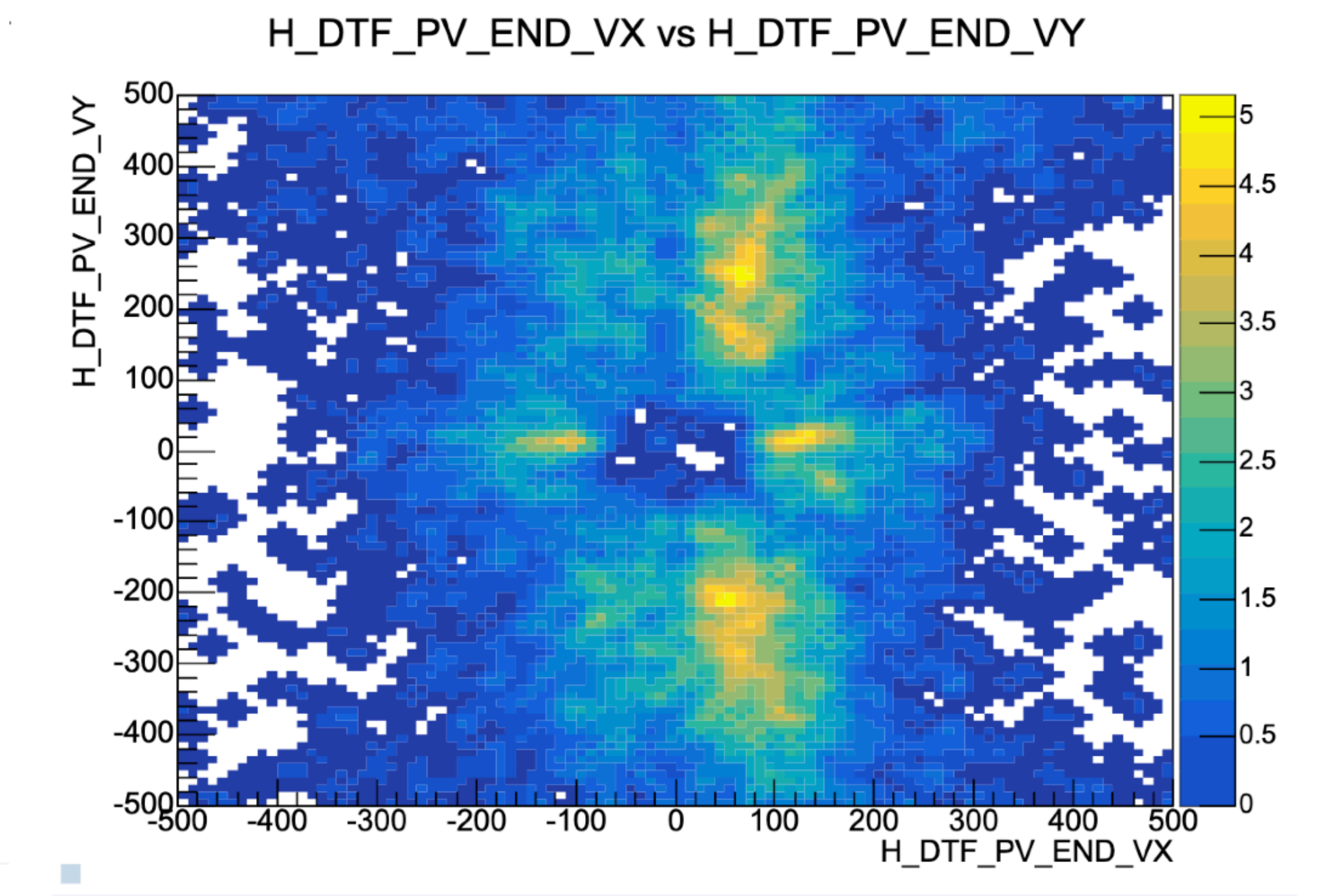
Mystery bump

- Previously reported on unexpected bump in the data
- Corresponds to hot spot in x of tracks at first measurement, but doesn't disappear when removing hotspot
- Could be a reconstruction artefact → to be further investigated









Monte Carlo Corrections

Monte Carlo Corrections

- Still under study
- Plan is to reweight distributions using the control sample control sample for taking into account data – MC discrepancies
- Following similar strategy to $B \to 4\mu$
- Reweight for track multiplicity
- Then reweight simulation with s-weighted data in regions of p and η
- Then reweight also for LLP vertex variables
- Then apply weights from control sample to the signal sample

Selections

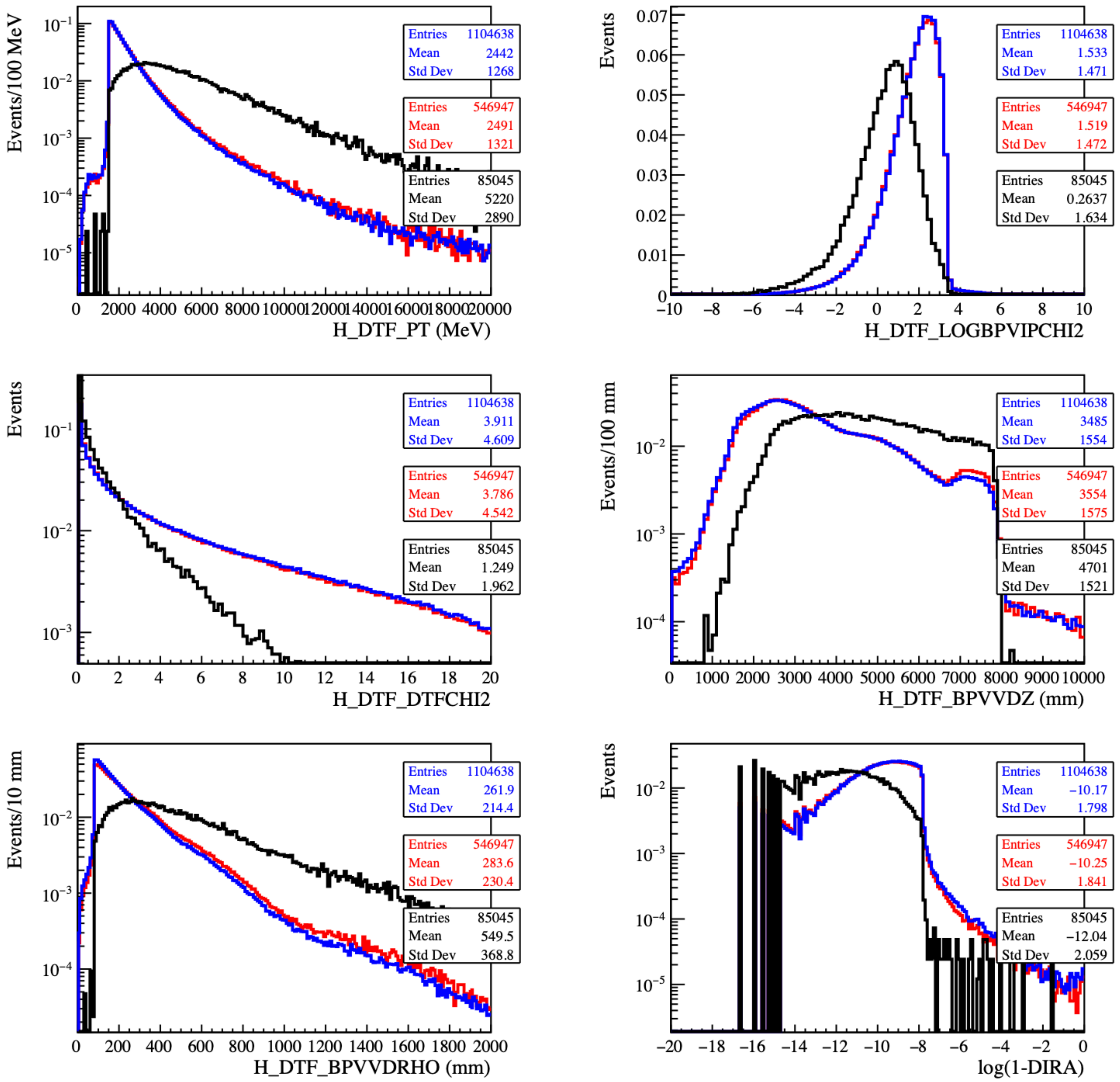
HLT2 selections

Quantity	Cut
Seed track MVA topological filter	
Single track MVA score	$P > 0.5$
Track pair MVA score	$P > 0.8$
Track cuts	
Ghost probability	$P_{\mathrm{ghost}} < 0.6$
Muon log-likelihood difference	$LL_{\mu} - LL_{\mathrm{BG}} > -5$
Muon correlated χ^2	$\chi^2_{\mu,\mathrm{corr}} < 10.0$
Muon identification	`isMuon`
Transverse momentum	$p_{\mathrm{T}} > 750\,\mathrm{MeV}$
Momentum	$p > 5000\,\mathrm{MeV}$
Number of hits	$N_{\mathrm{hits}} > 10$
Track χ^2	$\chi^2_{track} < 15.0$
Combined $\Delta\mathrm{LL}(\mu - \pi)$	$PID_{\mu} > -10.0$
Combined $\Delta\mathrm{LL}(K - \pi)$	$PID_{K} < 10.0$
Combined $\Delta\mathrm{LL}(p - \pi)$	$PID_{P} < 12.0$
Impact parameter χ^2	$\chi^2_{\mathrm{IP}} > 25.0$

Combination cuts	
z at YZ-intersection	$1500.0 < z_{\mathrm{yz}} < 8000.0$ mm
Distance of closest approach	$d_{\mathrm{CA}} < 400.0$ mm
DOCA χ^2	$\chi^2_{d_{\mathrm{CA}}} < 20000.0$
Vertex cuts	
Vertex z	$z_{\mathrm{vertex}} < 8\,\mathrm{m}$
Pointing angle (DIRA)	DIRA > 0.9996
Vertex χ^2	$\chi^2_{\mathrm{vertex}} < 20.0$
IP wrt primary vertex	$d_0^{\mathrm{BPV}} < 150.0$ mm
IP χ^2 wrt primary vertex	$\chi^2_{d_0\mathrm{BPV}} < 200.0$
Vertex $v_{d\rho}$	$v_{d\rho} > 80.0$ mm
Mass	$0.0 < m < 100000.0$ MeV

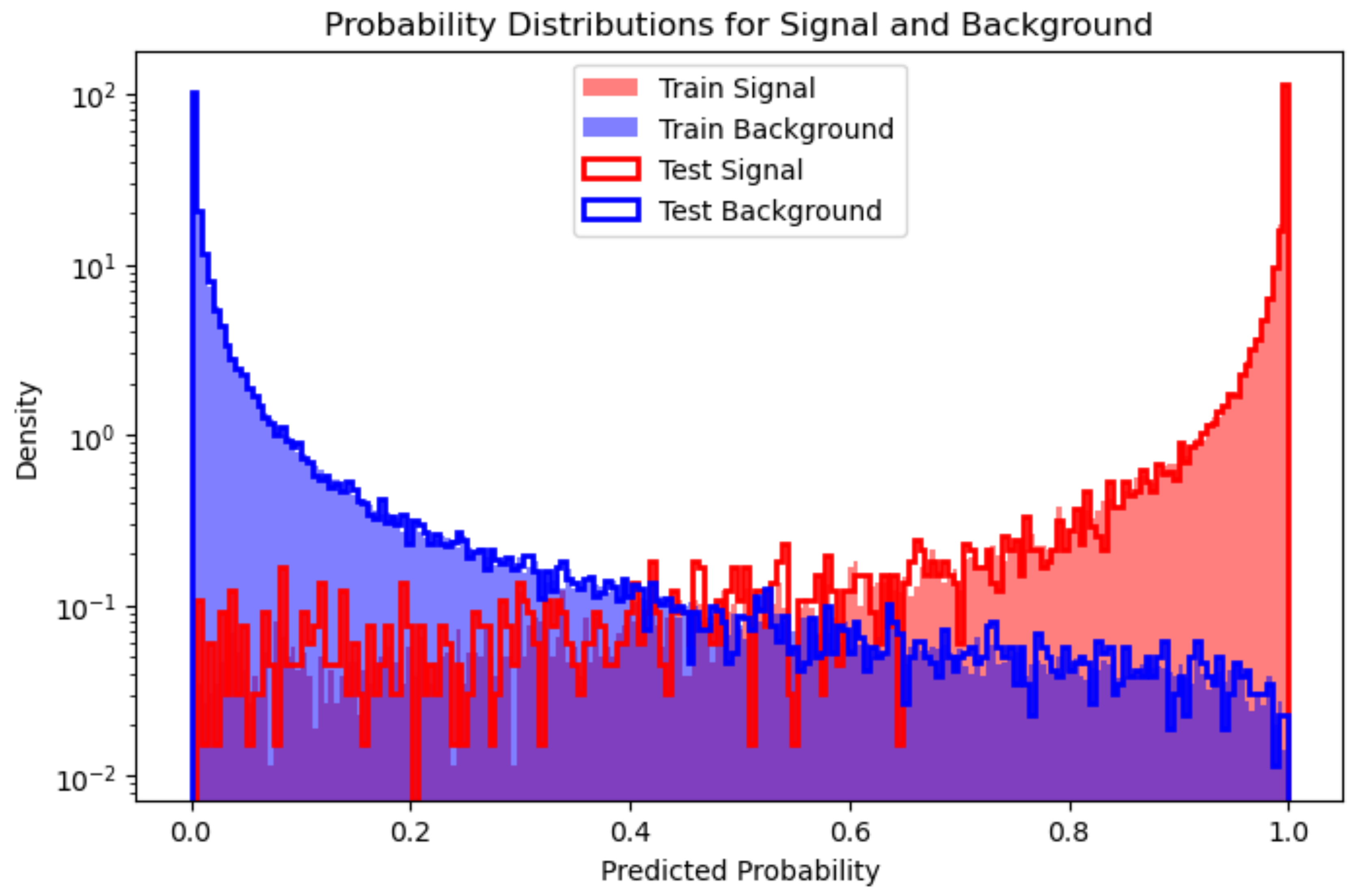
Offline selection

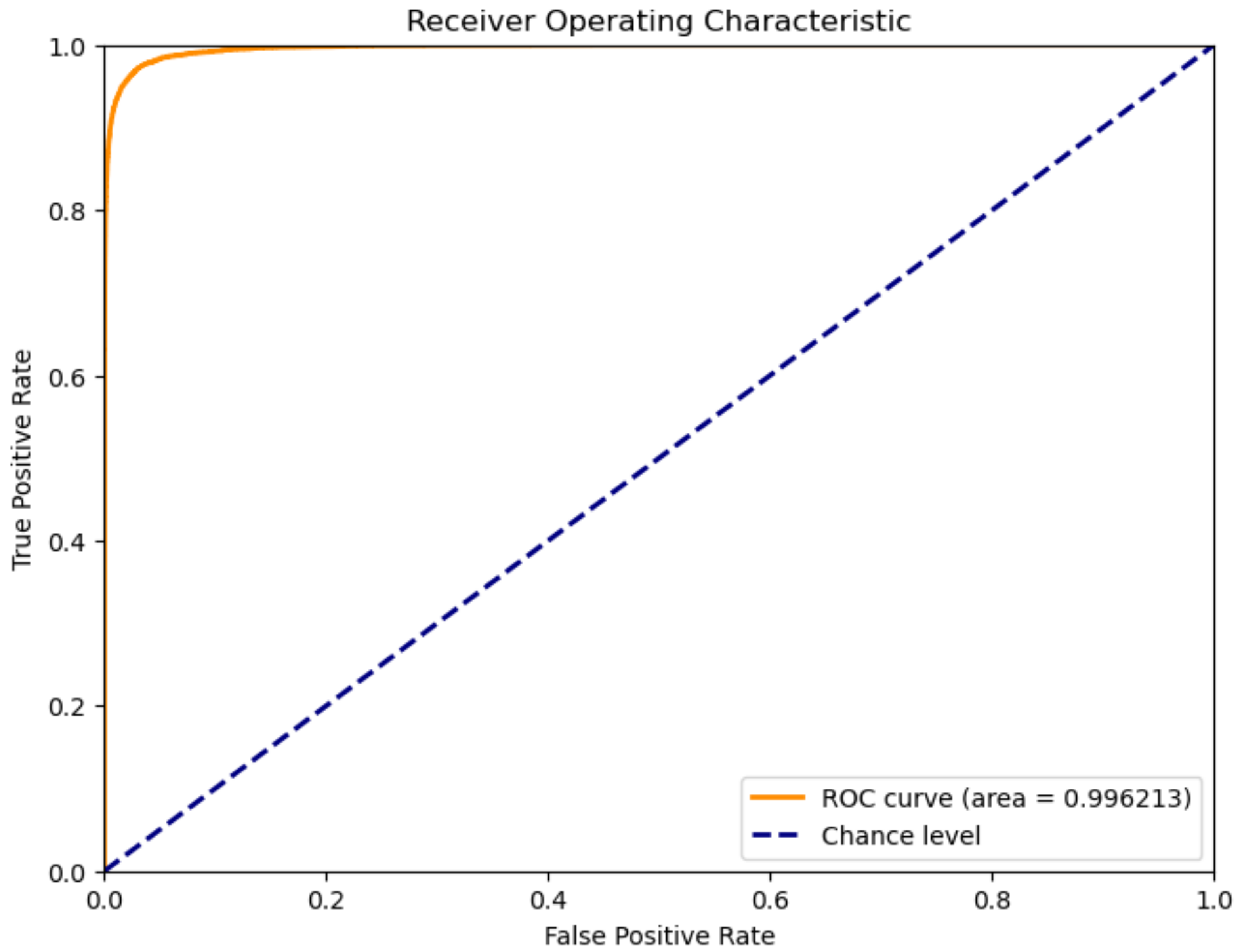
- Developing MVA to separate signal and background
  - Considering PyTorch NN and CatBoost
- Uses as features kinematic, topological and PID variables that can discriminate signal and background
  - Feature selection still to be pruned
- Use the unblinded data as background, signal taken from MC
  - Need to ensure data MC agreement (WIP)
- Estimate can reduce background to < 1 event /pb^{-1} keeping 60% of signal (O(20) events after catboost cut of > 0.995)



Offline selection

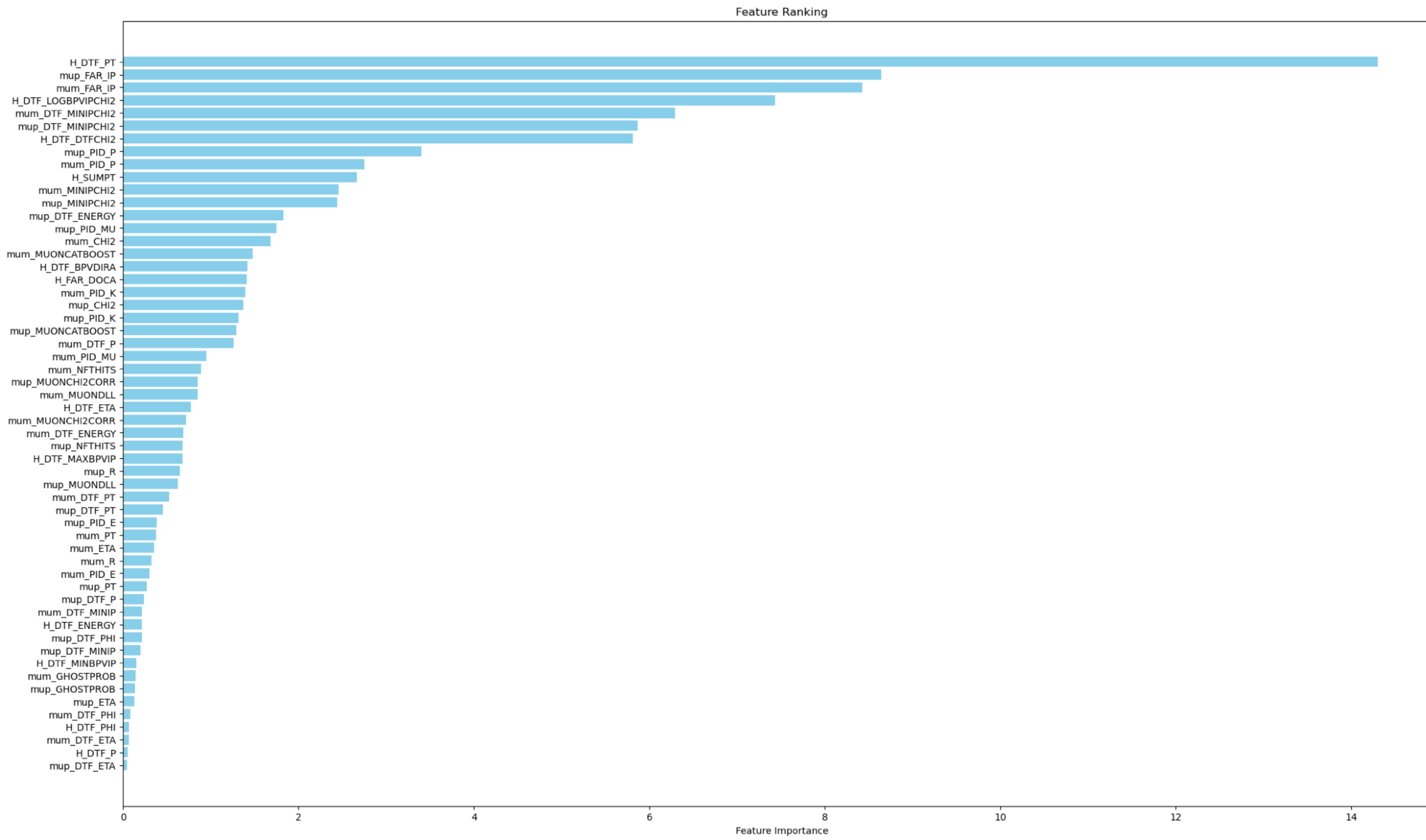
- Developing MVA to separate signal and background
  - Considering PyTorch NN and CatBoost
- Uses as features kinematic, topological and PID variables that can discriminate signal and background
  - Feature selection still to be pruned
- Use the unblinded data as background, signal taken from MC
  - Need to ensure data MC agreement (WIP)
- Estimate can reduce background to < 1 event /pb^{-1} keeping 60% of signal (O(20) events after catboost cut of > 0.995)





Offline selection

- Developing MVA to separate signal and background
  - Considering PyTorch NN and CatBoost
- Uses as features kinematic, topological and PID variables that can discriminate signal and background
  - Feature selection still to be pruned
- Use the unblinded data as background, signal taken from MC
  - Need to ensure data MC agreement (WIP)
- Estimate can reduce background to < 1 event /pb^{-1} keeping 60% of signal (O(20) events after catboost cut of > 0.995)

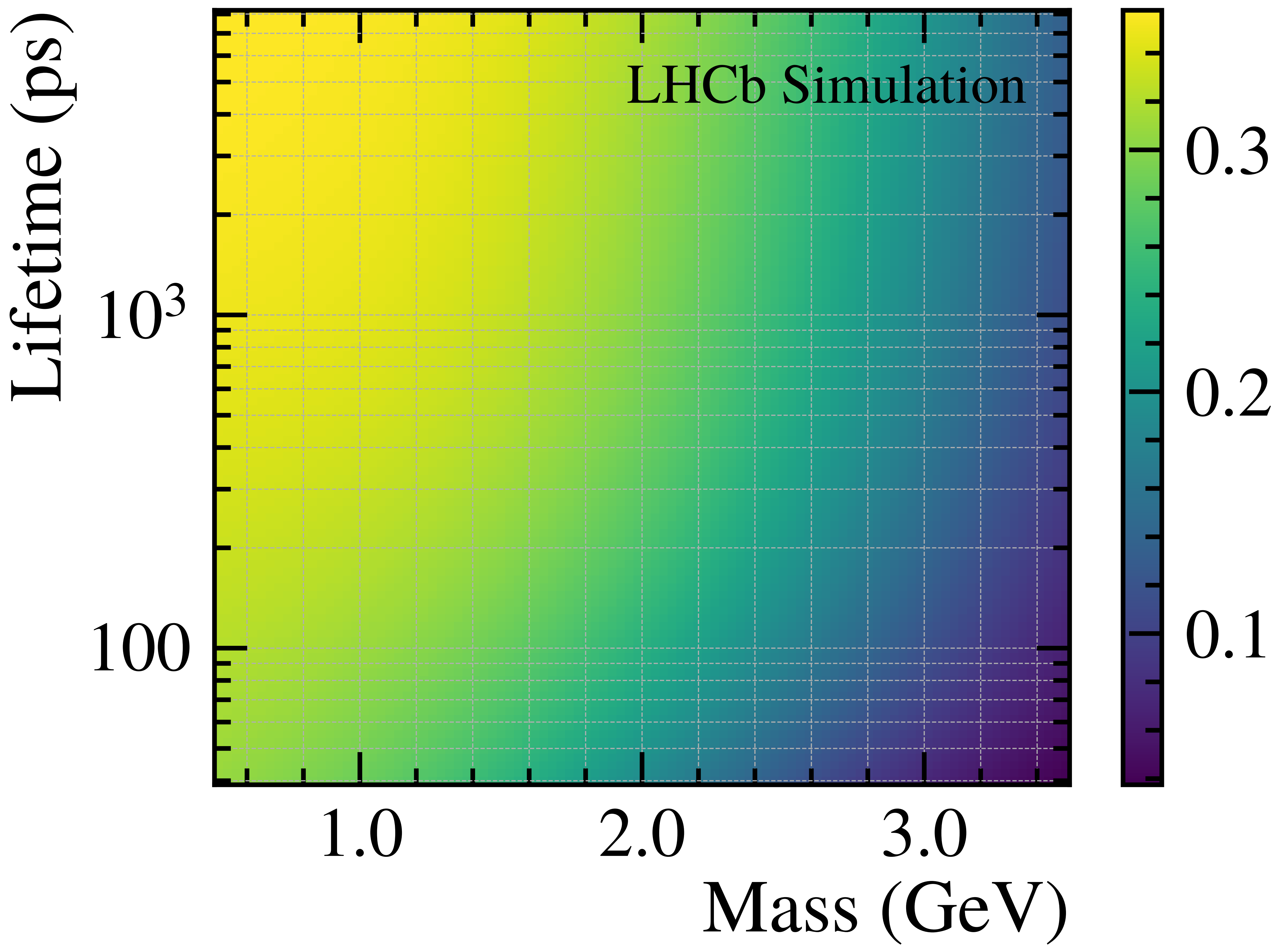


Efficiencies

Efficiencies

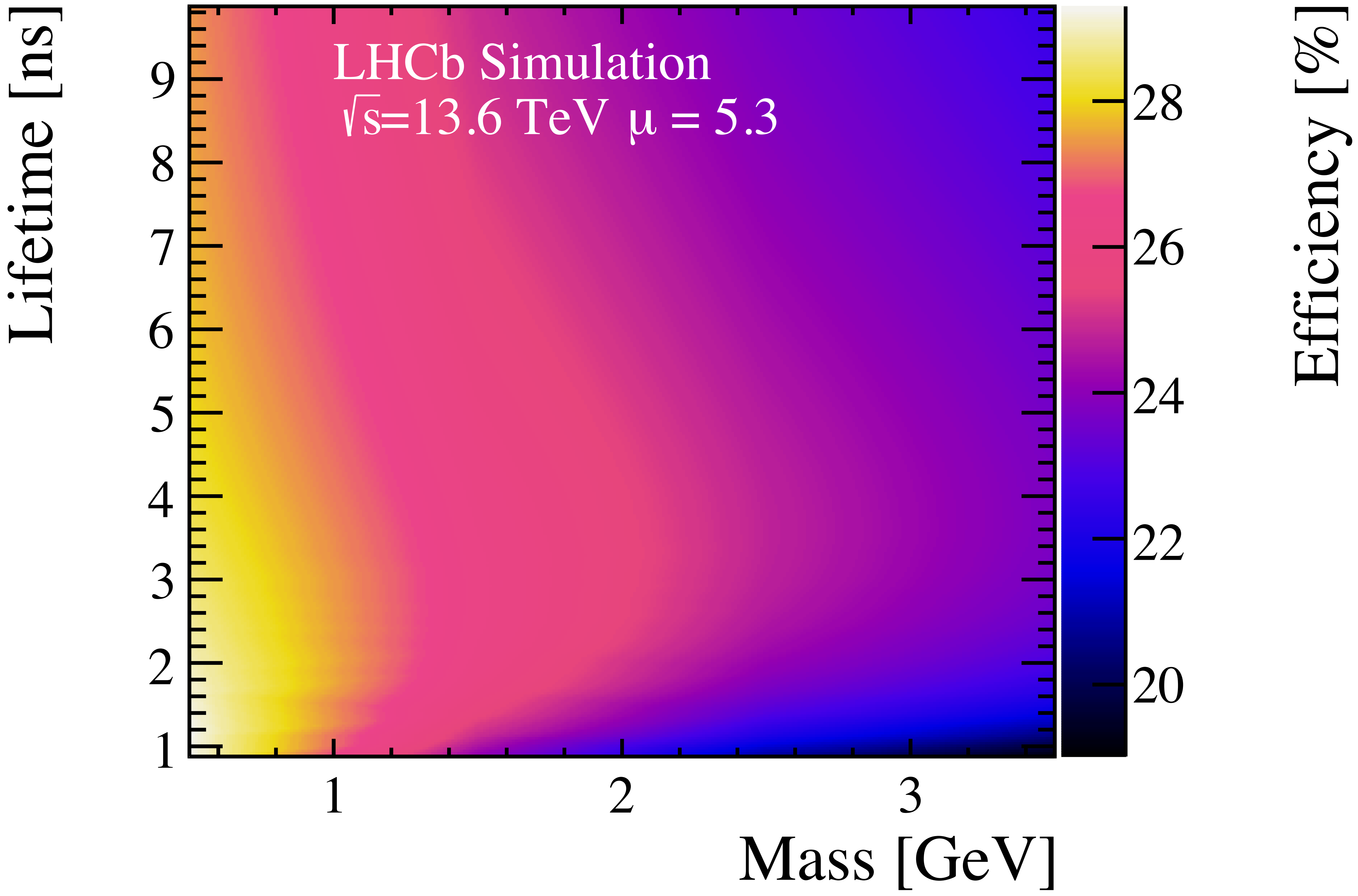
- Generator cut efficiency
- Generator cuts correspond to:

```
# tightCut.Preambulo += [
#     "from GaudiKernel.SystemOfUnits import GeV, MeV, mrad, meter, millimeter",
#     'inAcc        =  in_range ( 0.0 , GTHETA , 0.500 ) & ( GFAEVX ( GVZ, 100 * meter ) < 8000. * millimeter)',
#         ]
# tightCut.Cuts      =    {
#     '[H_10]cc'      : 'inAcc' ,
#     }
```



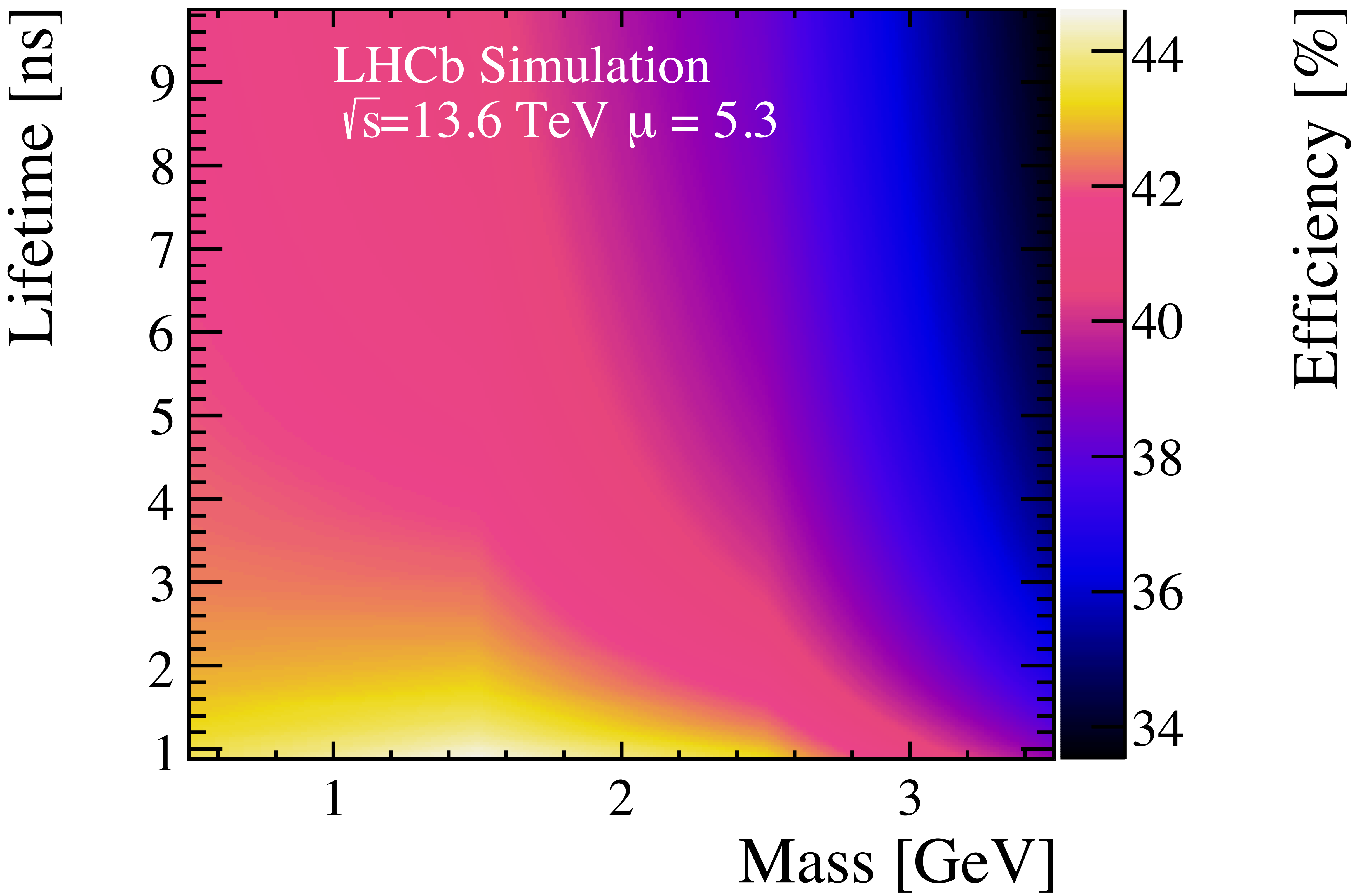
Efficiencies

- Reconstructibility efficiency
- Corresponds to fraction of tracks reconstructible as T tracks
  - Of course all tracks reconstructible as Long and Downstream are also reconstructible as T tracks but this is not considered here
    - Something to be investigated for future



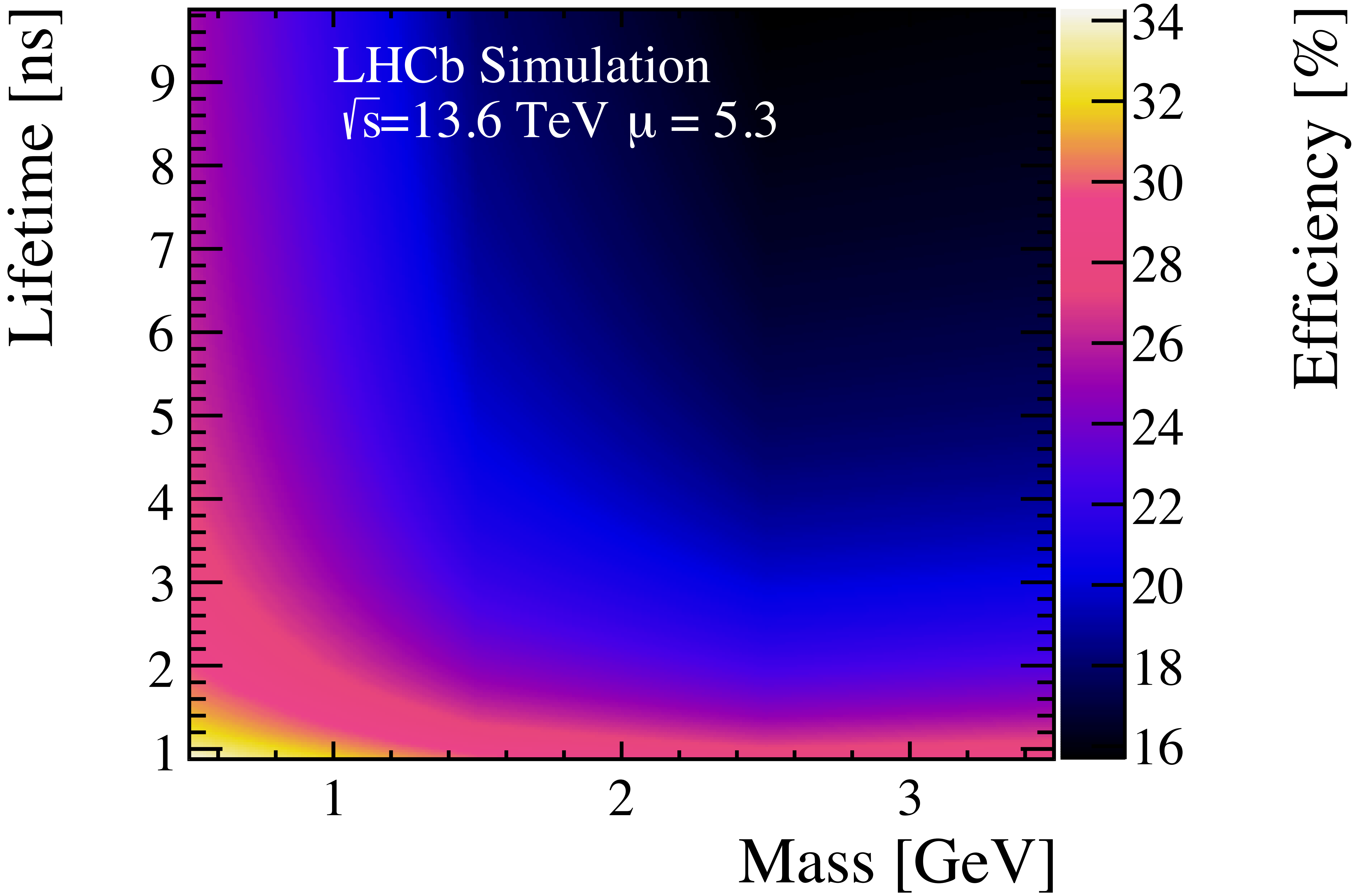
Efficiencies

- HLT1 efficiency
- Independent of the dimuon decay
- Largely depends on the associated kaon
- Shown here for $B^0 \to K^*(892)(\to K\pi)H'(\to \mu\mu)$
- Will vary for other modes, to be reevaluated with the cocktail MC when available



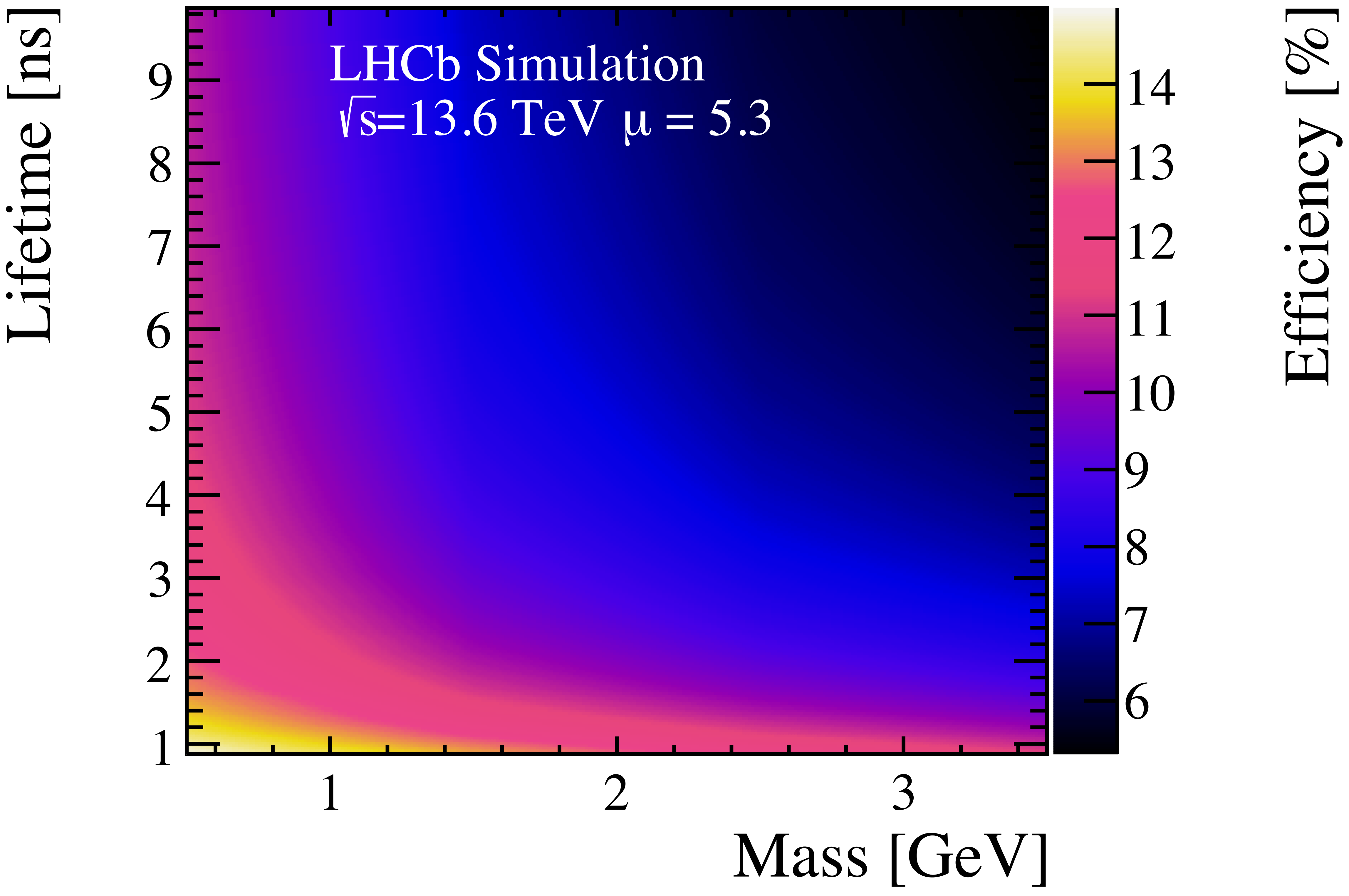
Efficiencies

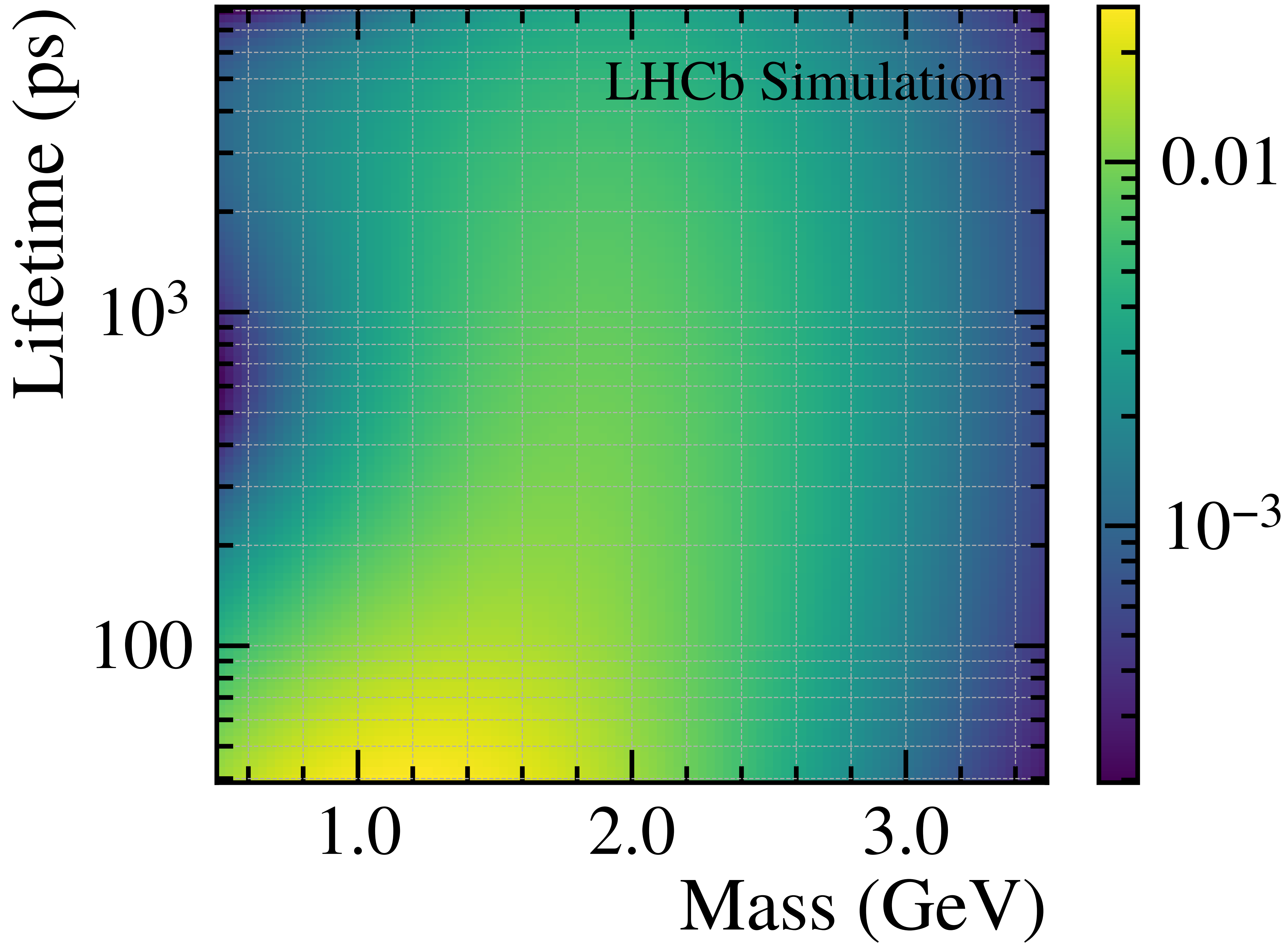
- HLT2 efficiency wrt HLT1
- Room for future improvement here for 2025-2026
- In particular with the seed track MVA topological filtering
  - (Have retuned the model to boost efficiencies by 30%, but issues with truth matching mean not implemented yet)



Efficiencies

- Total efficiencies:
  - With respect to reconstructible (left) and generated events (right)
  - Offline selection efficiencies still WIP

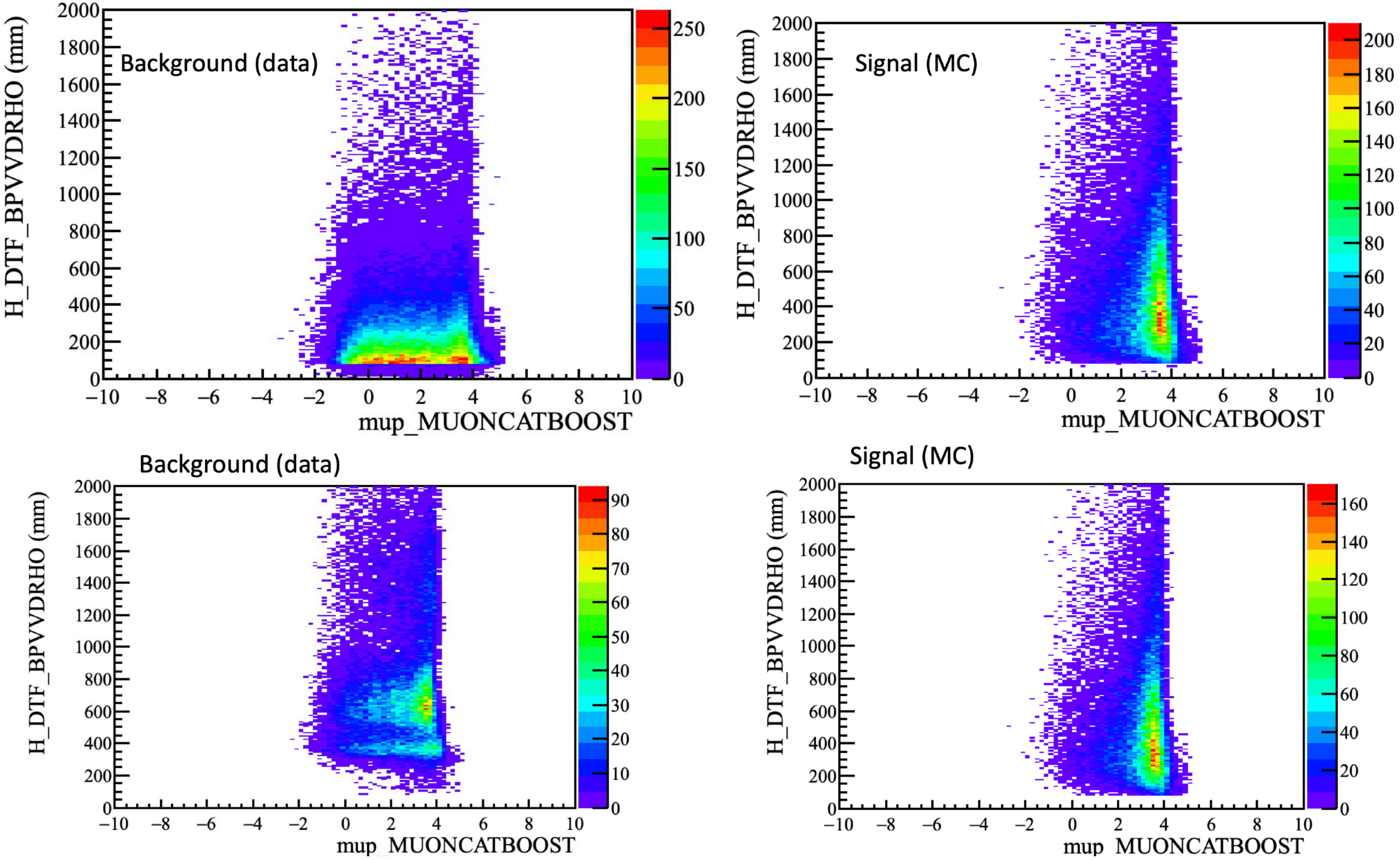




Signal and control regions

Signal and control regions

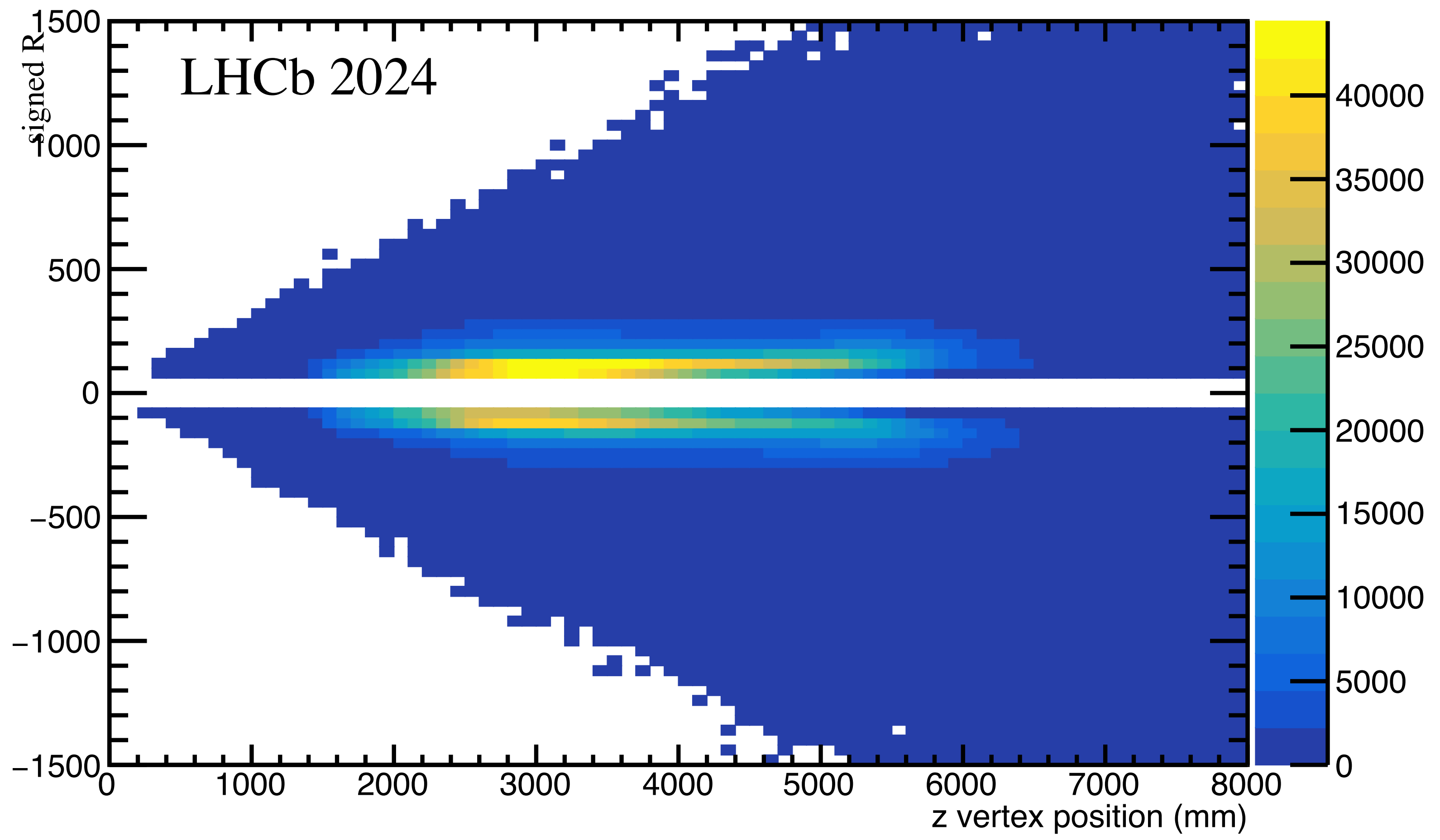
- Work still in progress to identify suitable uncorrelated variables to define signal region for ABCD method
- Need to be uncorrelated with each other and with the MVA response
  - Signal regions may be well separated before MVA cut but not after

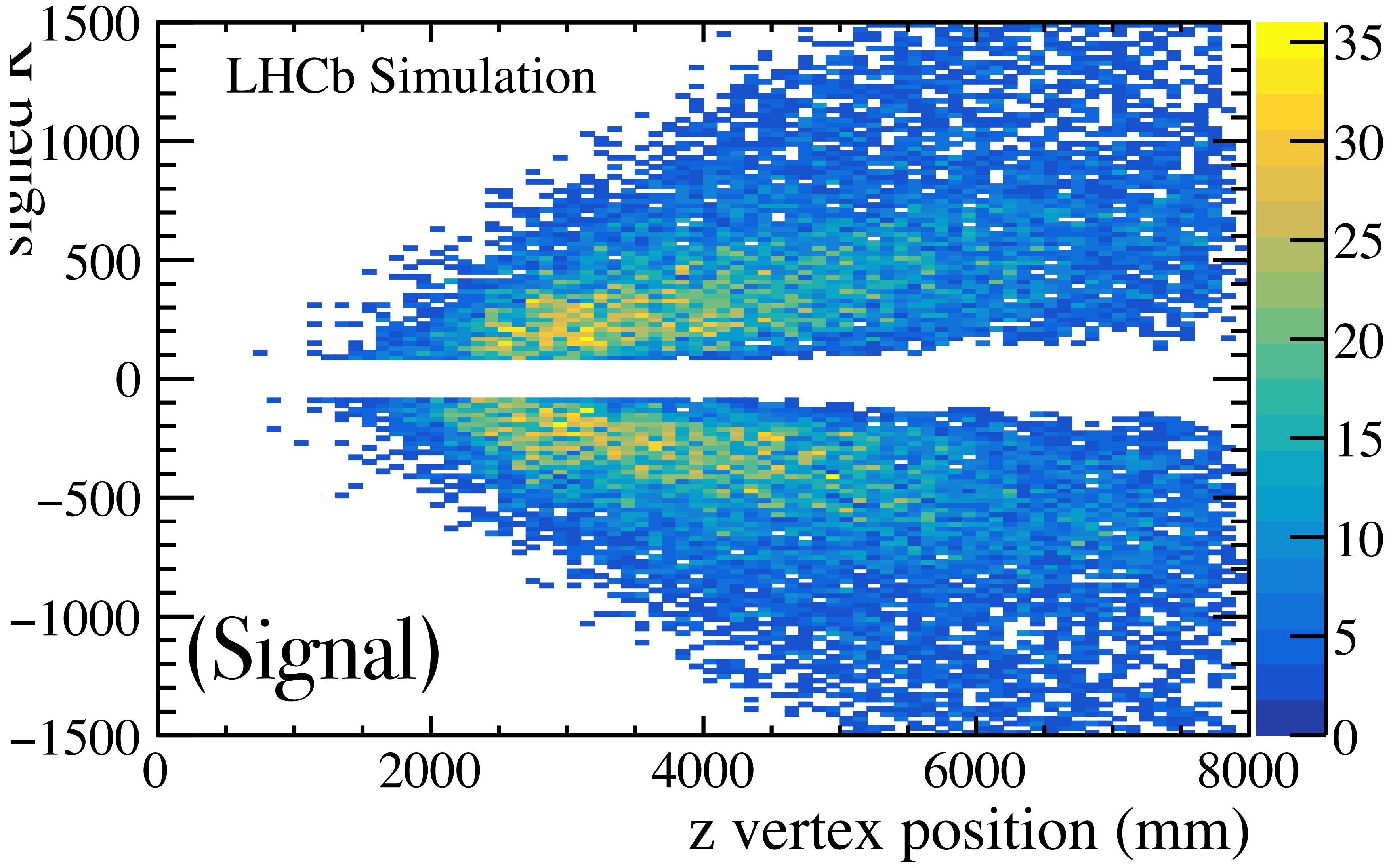


Background studies

Background studies

- Background expected to be dominated by combinatorial and material interactions
  - Mainly located around beam pipe
  - Signal is much more dispersed throughout the volume
- Work still underway to classify background in MC → requires very high statistics due to tight trigger selections
- Long term plan to do a detailed tomography of magnet region → may not be required for 2024 study but precise mapping will be required for larger statistics

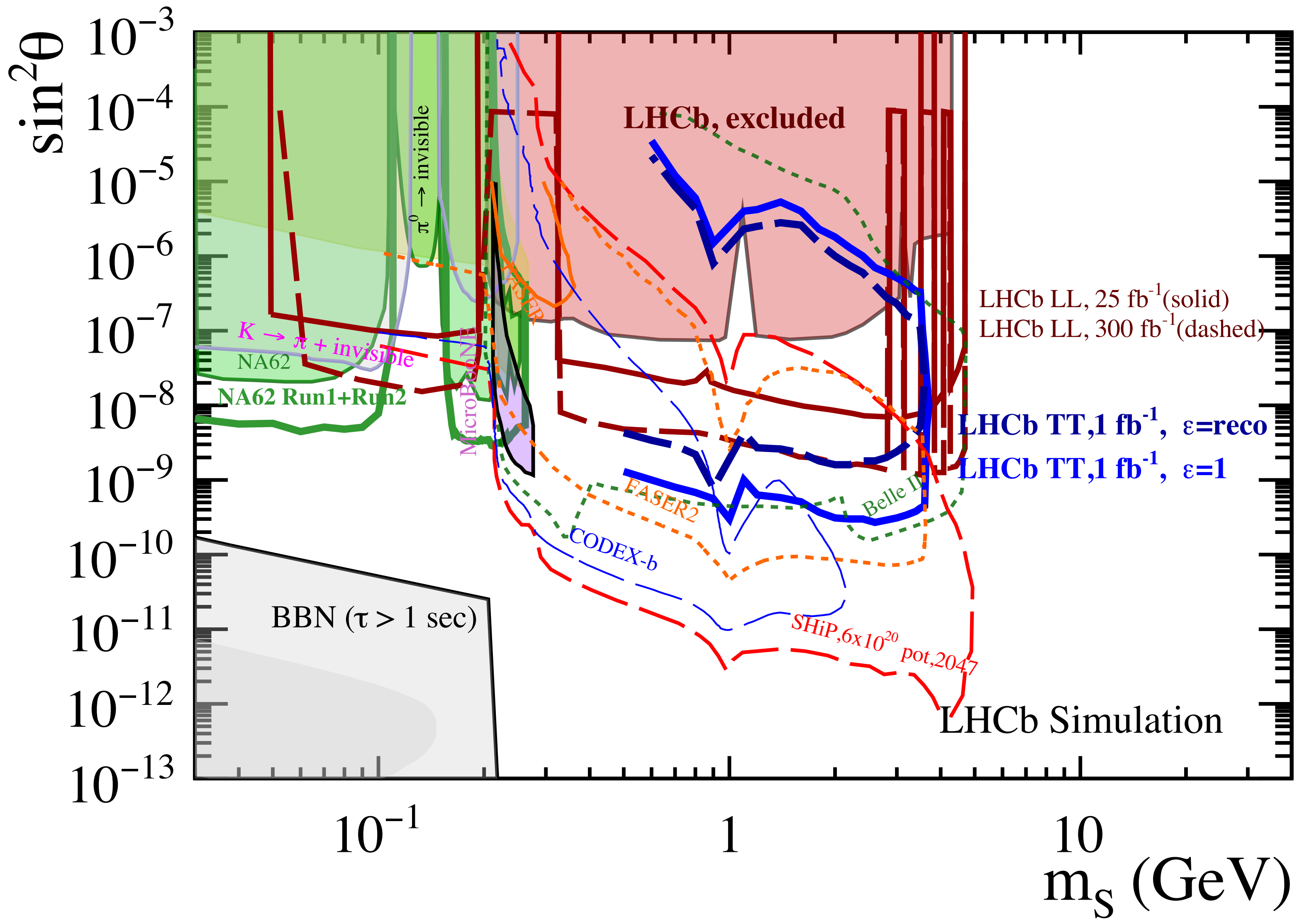


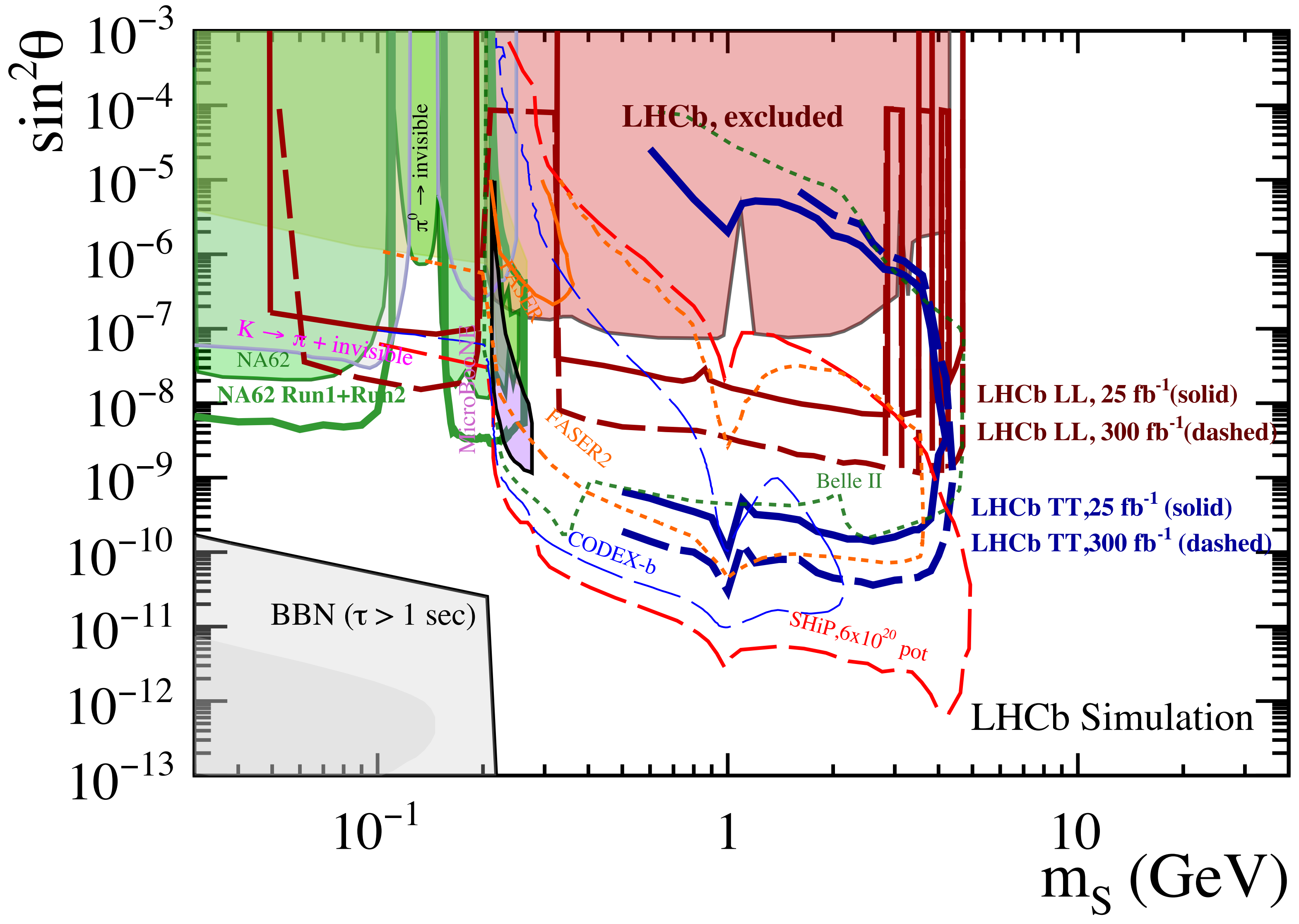


Summary and conclusions

Summary and conclusions

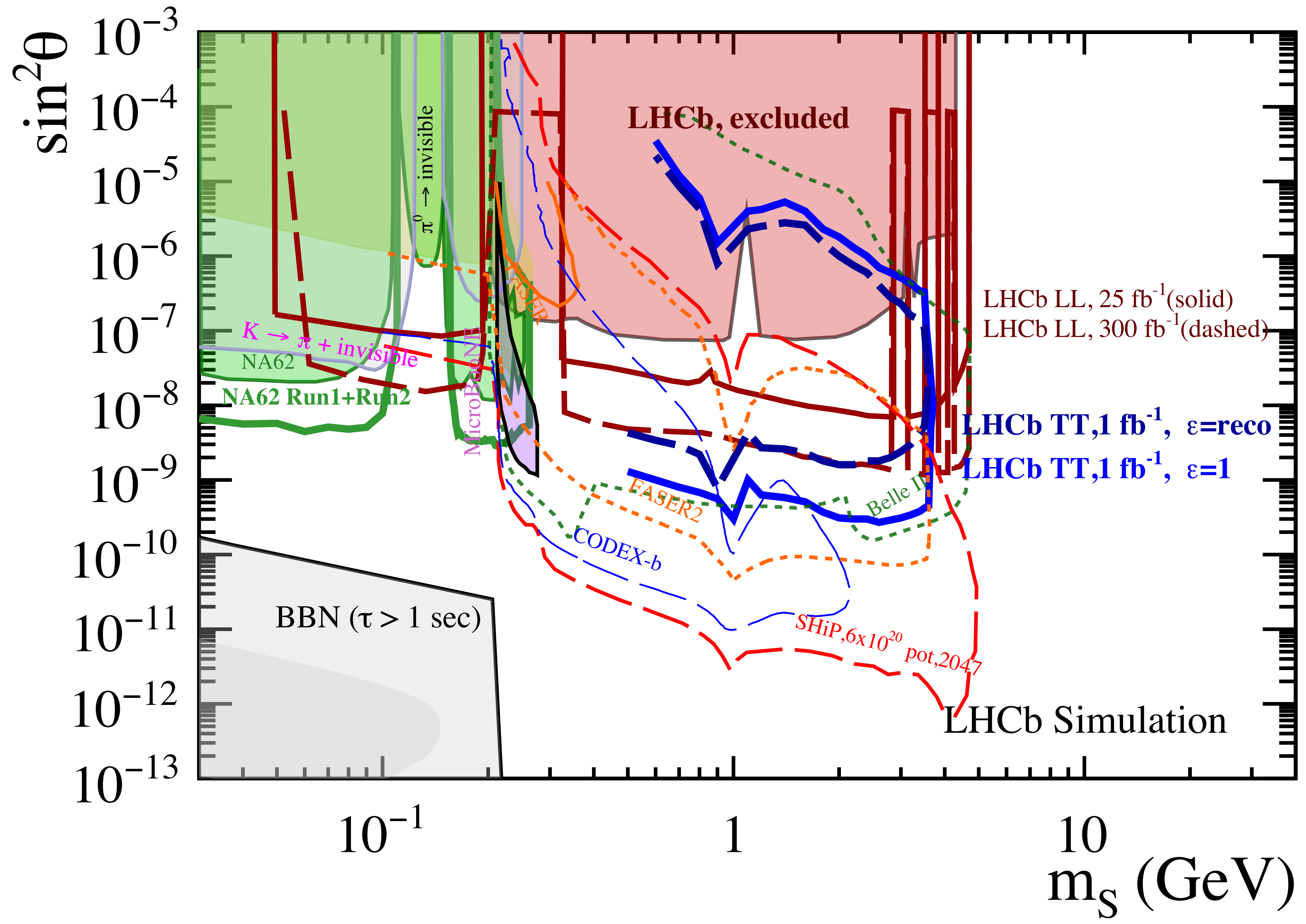
- Plots show projected sensitives for 1 fb^{-1}, 25 fb^{-1} and 300 fb^{-1} using the 2024 triggers, assuming same efficiencies for all channels as for the $B^0 \to K^*(892)(\to K\pi)H'(\to \mu\mu)$ sample, not accounting for background or offline selection efficiency
- With T tracks can make a significant impact on low coupling/high lifetime limits

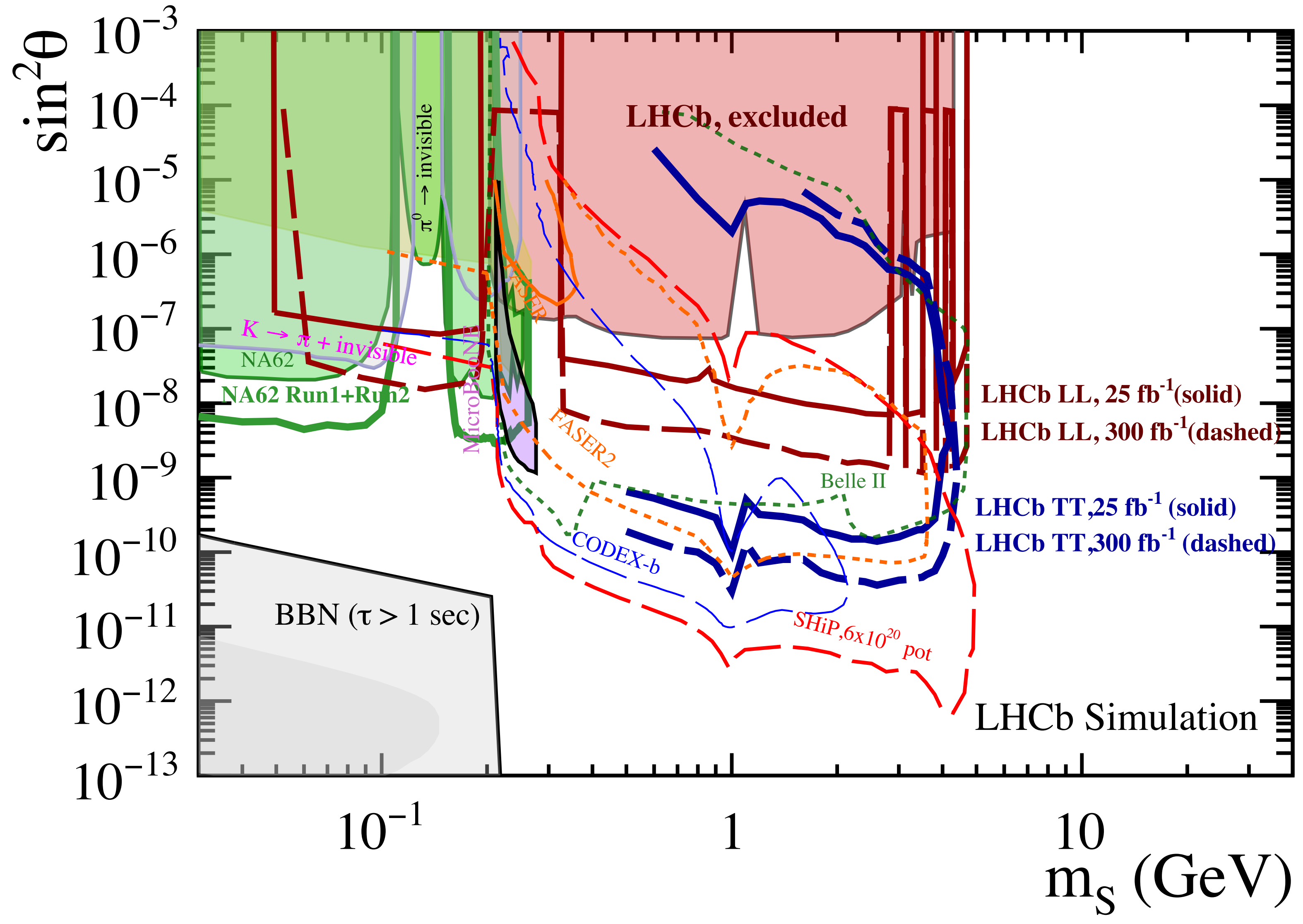




Summary and conclusions

- Good progress made on:
  - Analysis strategy
  - Understanding reconstruction
  - Selections
  - Efficiency estimates
- Still to do/in progress:
  - Full MC sample generation
  - Control channel studies
  - Detailed momentum resolution and mass resolution studies
  - Data–MC reweighing
  - MVA optimisation
  - Background studies
  - PID calibration






Thank you for listening

Backup

Data uniformity in un/blinded regions