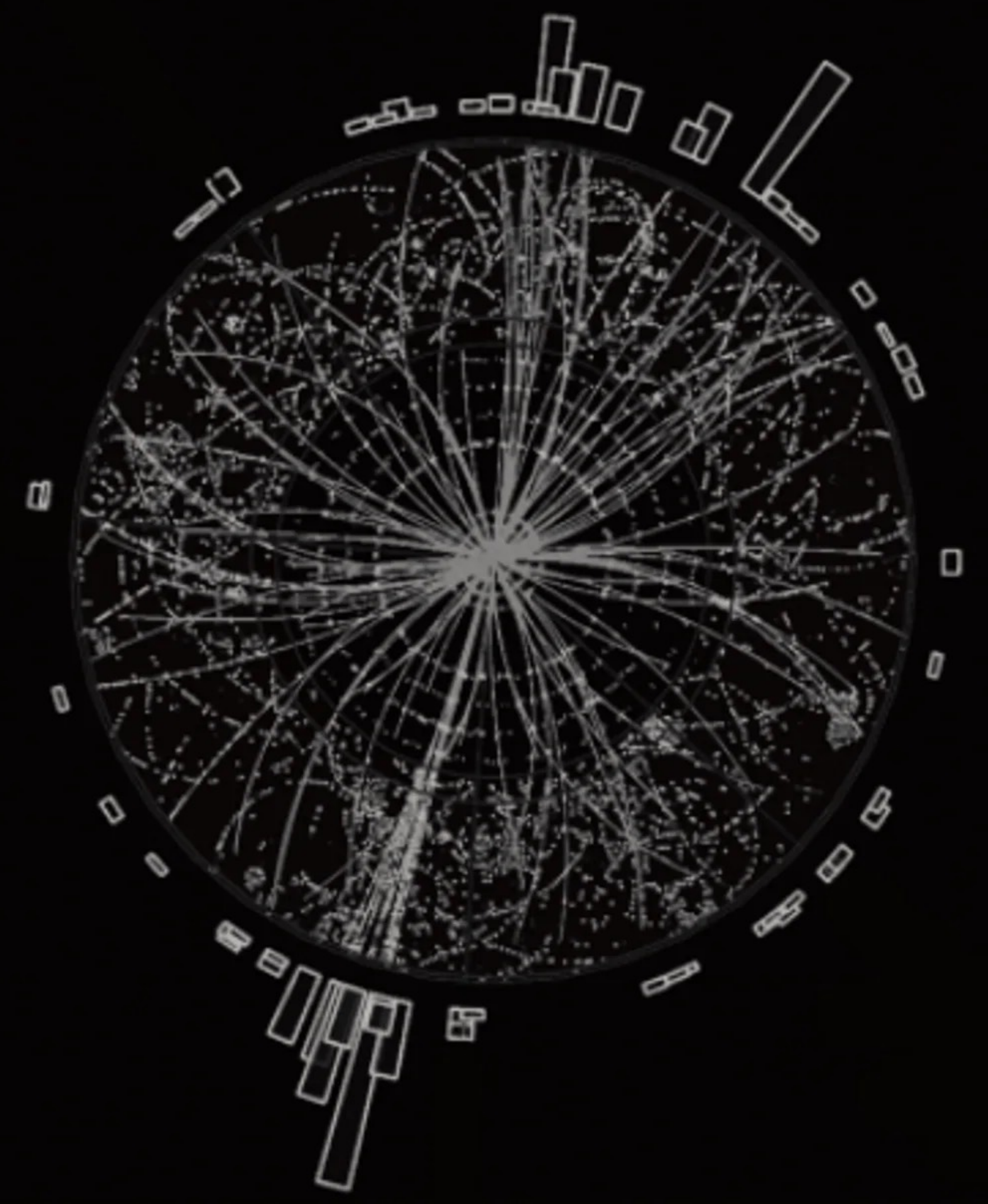
Prompt Lepton Isolation Tagger

Development and Optimisation of ATLAS State-of-the-art Isolation Algorithm



Ema Maričić

CEA Paris-Saclay & Institute of Physics Belgrade

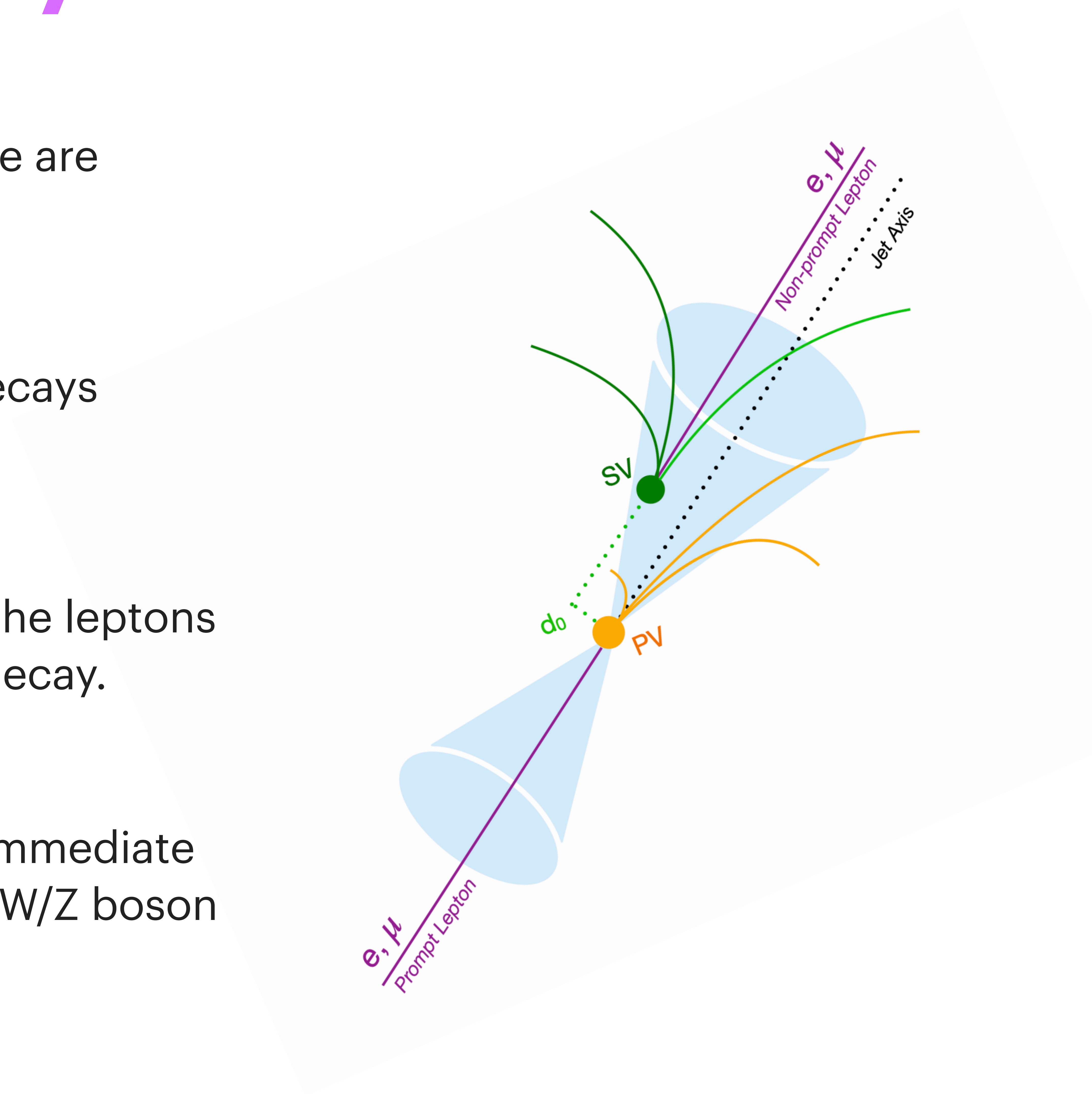
Isolation: The Very Basics

If we are doing a measurement related to top quarks at the LHC, we are usually interested in events containing Z/W/H bosons.

Since these bosons have very short lifetimes, leptons from their decays effectively come from the primary vertex.

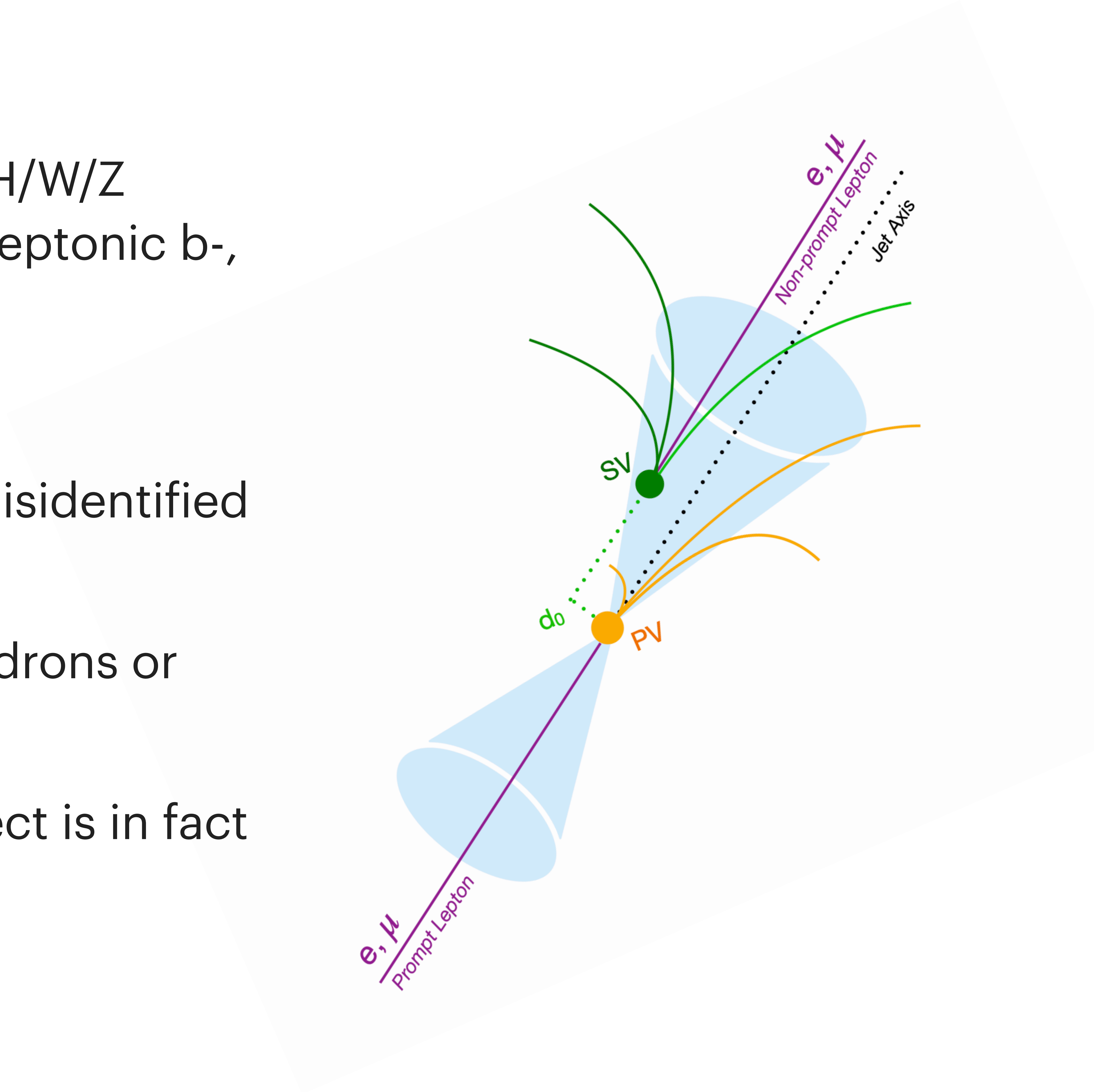
As such, they should have less activity around them compared to the leptons that come from a decay chain of a b-/c-/l-hadron or hadronic tau decay.

The idea is to combine information about the lepton itself and its immediate surroundings in order to conclude if the lepton comes from the H/W/Z boson or not.



Nomenclature

Thus, we want to differentiate **prompt leptons** coming from H/W/Z decays from **non-prompt leptons** coming mostly from semi-leptonic b-, c-, and l-decays.



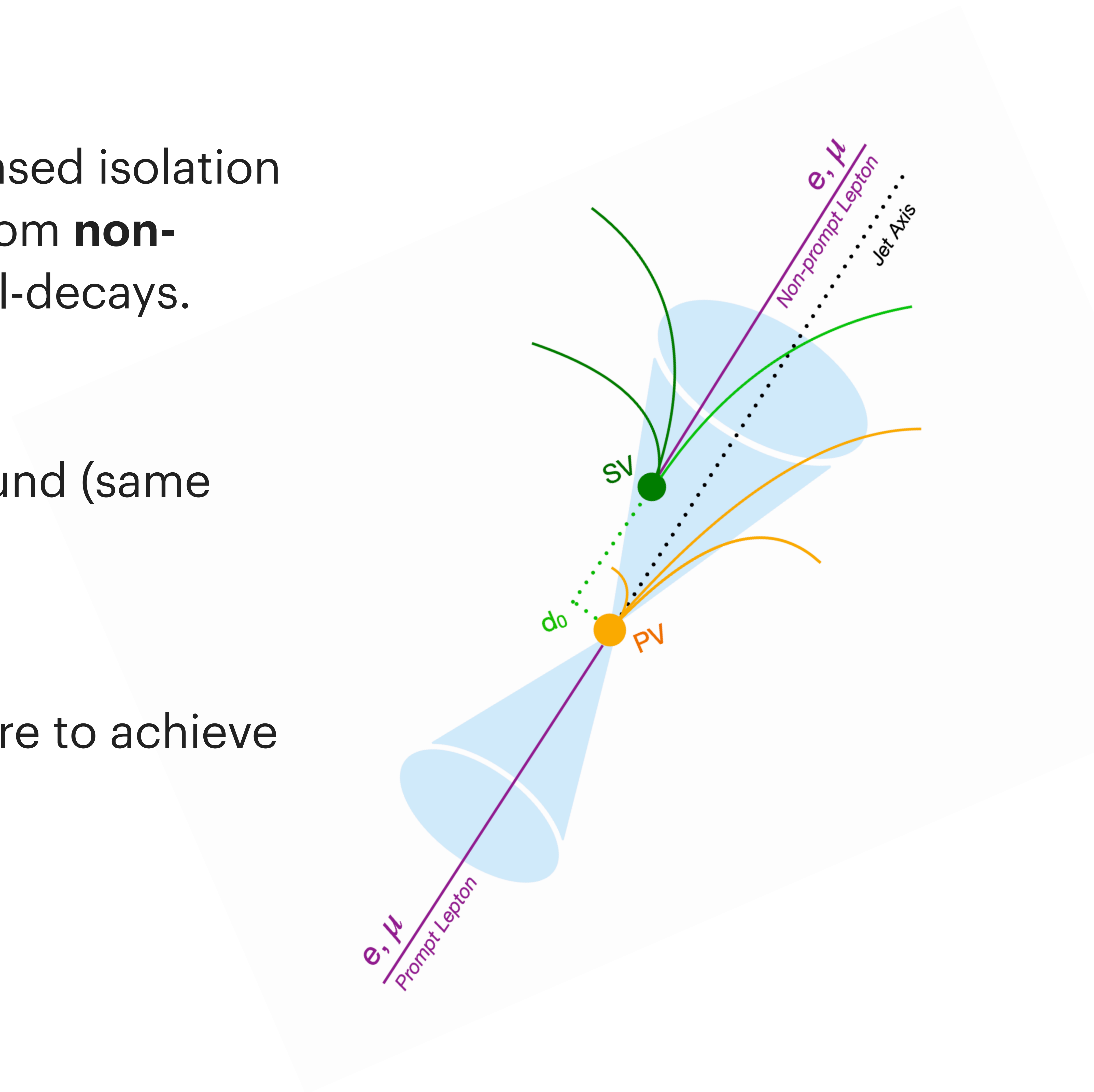
What is the difference between a non-prompt lepton and a misidentified lepton?

- Non-prompt leptons come from semi-leptonic decays of hadrons or photon conversions.

- In the case of a misidentified lepton, the reconstructed object is in fact not due to a lepton.

What is PLIT?

PLIT (Prompt Lepton Isolation Tagger) is a machine learning based isolation tool separating **prompt leptons** coming from H/W/Z decays from **non-prompt leptons** coming mostly from semi-leptonic b-, c-, and l-decays.



It's useful to analyses sensitive to non-prompt lepton background (same sign leptons and multilepton analyses such as ttH(ML)).

PLIT employs flavour-tagging tools and transformer architecture to achieve unprecedented performance.

As such, it supersedes PLIV, its predecessor.

PLIV: Predecessor of PLIT

PLIV (Prompt Lepton Improved Veto) is based on a BDT operating on input variables calculated with respect to the lepton, tracks, and track jets. It is an algorithm extensively used in Run 2 multilepton analyses.

RNN inputs - defined for tracks

RNN inputs	Description
$\Delta R(track, trackjet)$	The angular ΔR between ID track and jet^{lepton}_{track}
$p_T^{track}/p_T^{trackjet}$	p_T of ID track divided by p_T of jet^{lepton}_{track}
$z_0 \sin(\theta)$	longitudinal impact parameter scaled by $\sin(\theta)$
d_0/σ_{d_0}	transverse impact parameter significance
N^{PIX}_{hit}	number of hits of the track in the pixel detector
N^{SCT}_{hit}	number of hits of the track in the semiconductor tracker

BDT inputs - defined for leptons

Inputs	Description	BDT^{μ}_{new}	BDT^{e}_{new}
RNN	RNN using track impact parameters and other variables	✓	✓
$L^{longitudinal}_{SV\ to\ PV}/\sigma$	Secondary vertex longitudinal significance using dedicated vertexing	✓	✓
$\sum p_T^{VarCone30}/p_T$	Lepton isolation using ID tracks within a cone of $\Delta R < 0.3$ (TTVA for μ)	✓	✓
$\sum E_T^{TopoCone30}/p_T$	Lepton isolation using topological clusters within a cone of $\Delta R < 0.3$	✓	✓
$E^{\mu}_{cluster}/E_{expected}$	Relative muon calorimeter cluster energy	✓	-
$\sum^{dR<0.15}_{cluster} E_T/p_T$	Sum of cluster energy divided by lepton p_T	-	✓
N_{track} in track jet	Number of tracks clustered by the track jet	-	✓
p_T^{rel}	Lepton p_T along the track jet axis: $p \cdot \sin(<$ lepton, track jet $>)$	-	✓
$p_T^{lepton\ track}/p_T^{track\ jet}$	Lepton track p_T divided by track jet p_T	✓	✓
ΔR(lepton, track jet)	ΔR between the lepton and the track jet axis	✓	✓
p_T^{lepton} bin number	Index of the lepton p_T bin	✓	✓

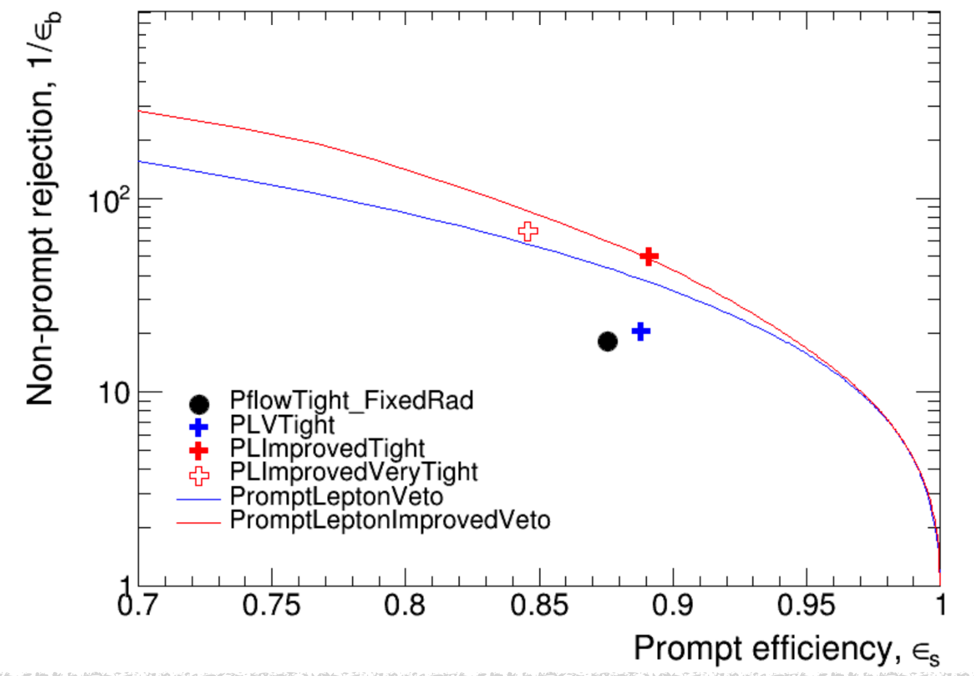
Reliance on track jets

PLIV relies on secondary vertex reconstruction as well

PLIV relies on separate b-tagging-inspired algorithms

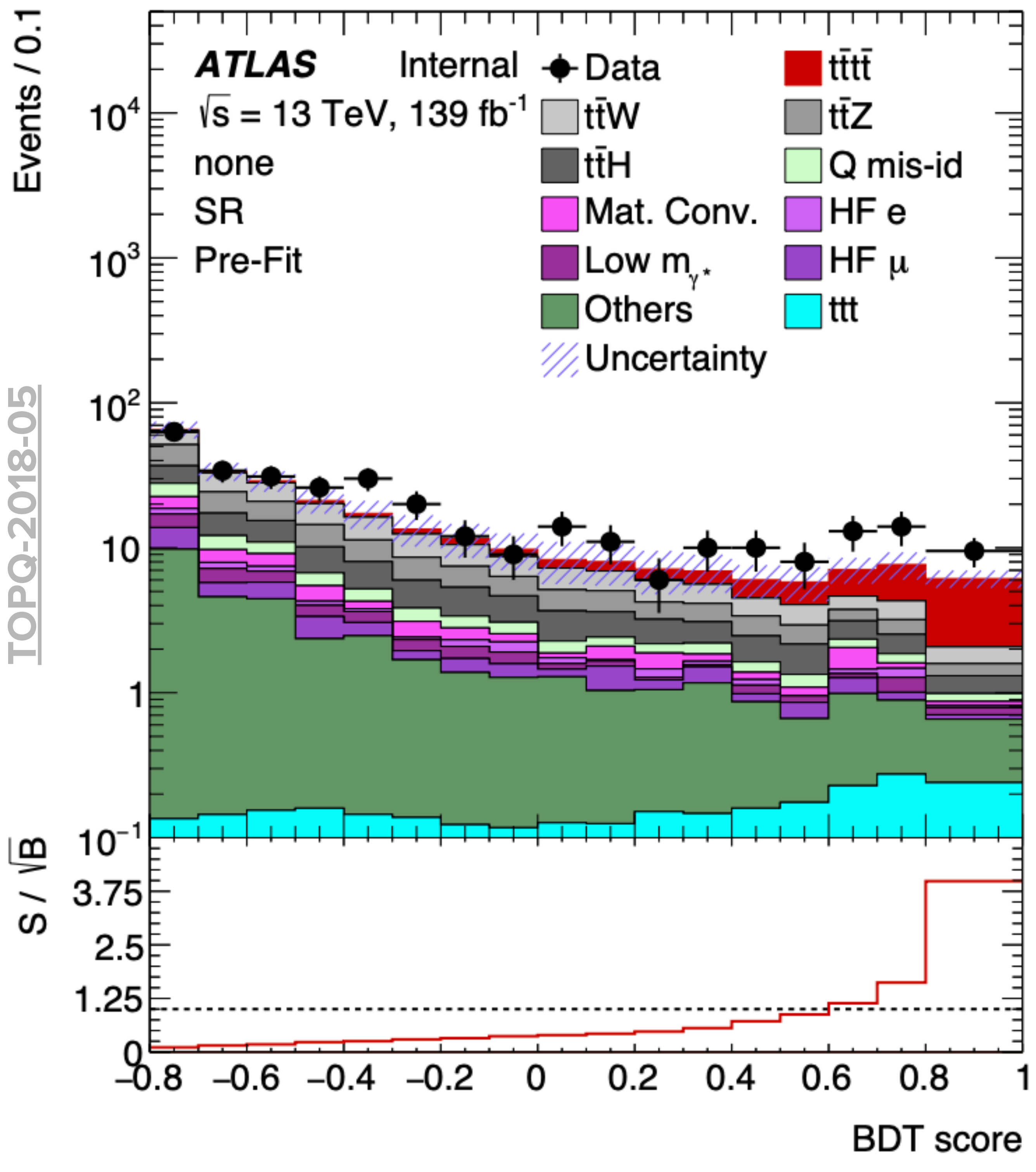
Major advantage of PLIT: No reliance on separate algorithms!

ROCs of PLIV and its predecessor



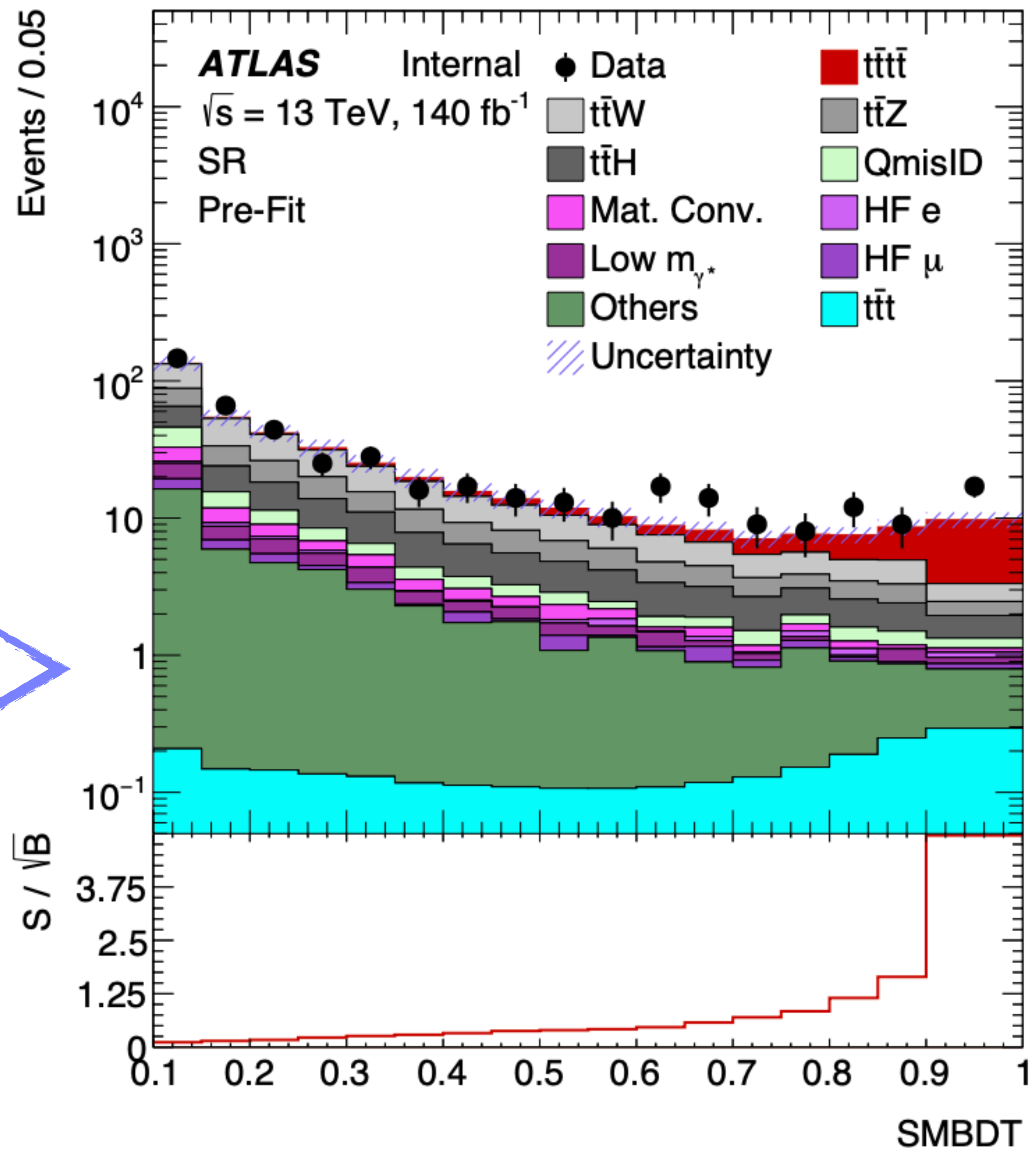
PLIV in 4-tops Run 2 analysis

Evidence Paper



Addition of PLIV
& numerous* other changes!

Discovery Paper



On the right, one sees the result of the **reanalysis effort** which uses PLIV, but also contains updates to p_T thresholds, reconstruction and b-tagging algorithms.

Input variables for PLIT

Variables computed with respect to the lepton and tracks are used as input variables:

- Lepton variables: Use of standard isolation variables
- Track variables: Up to 40 tracks with ΔR<0.4 around the lepton are associated to it. Usage of same variables as the current ATLAS b-tagging algorithm!

This large set of input variables along its advanced architecture enables PLIT not to rely on separate algorithms and efficiently separate prompt leptons from the non-prompt ones.

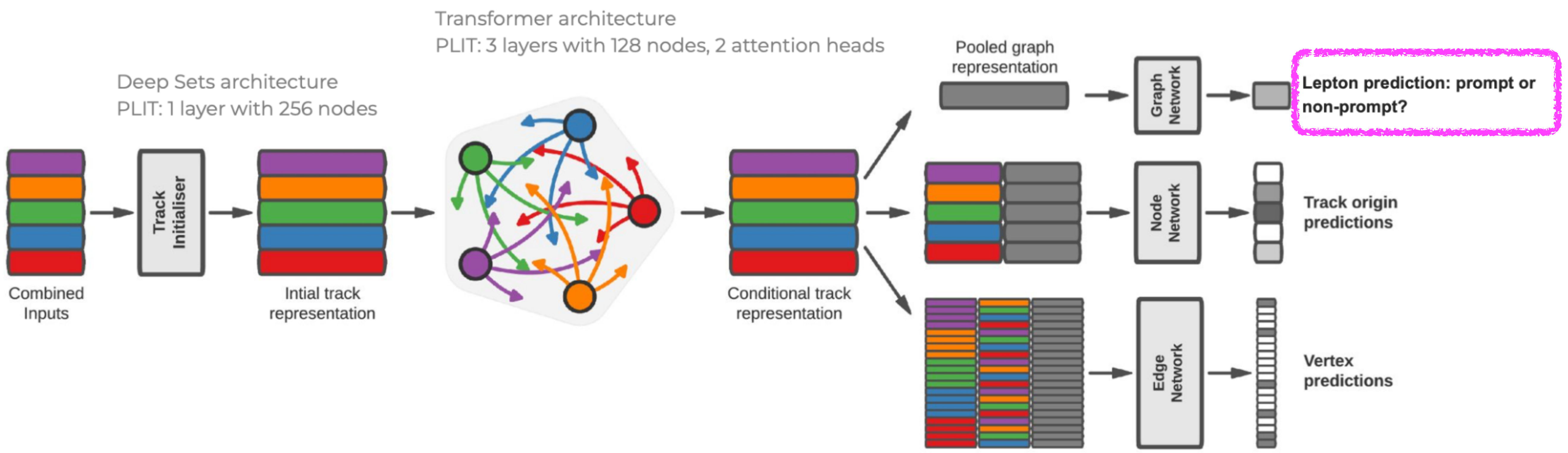
Muon Input	**Description**
p_{T}	Transverse momentum
η	Pseudorapidity
ϕ	Azimuthal angle
$p_{\mathrm{T}}^{\mathrm{varcone30}}/p_{\mathrm{T}}$	Scalar sum of the $p_{\mathrm{T}}^{\mathrm{track}}$ of the inner detector tracks associated with the primary vertex in an η–ϕ cone of $\Delta R < 0.3$ around the muon, excluding the muon track itself, divided by the muon p_{T}
$E_{\mathrm{T}}^{\mathrm{varcone30}}/p_{\mathrm{T}}$	Sum of the transverse energy of topological cell clusters in a cone of size $\Delta R = 0.3$ around the position of the muon, extrapolated to the calorimeters, after subtracting the contribution from the energy deposit of the muon itself and correcting for pile-up effects, divided by the muon p_{T}
caloClusterERel	Energy of the calorimeter cluster associated with the muon divided by the muon's expected energy loss in calorimeter
Track Input	**Description**
$\Delta\eta(\mathrm{track, lepton})$	Difference in η between track and lepton
$\Delta\phi(\mathrm{track, lepton})$	Difference in ϕ between track and lepton
$q^{\mathrm{track}}\ /\ p^{\mathrm{track}}$	Track charge divided by momentum
d_0	Transverse impact parameter
$z_0 \sin\theta$	Longitudinal impact parameter times sine of the polar angle
d_0 significance	Significance of the transverse impact parameter
$z_0 \sin\theta$ significance	Significance of the longitudinal impact parameter times sine of the polar angle
nPixHits	Number of pixel hits
nIBLHits	Number of IBL hits
nBLHits	Number of B-layer hits
nIBLShared	Number of shared IBL hits
nIBLSplit	Number of split IBL hits
nPixShared	Number of shared pixel hits
nPixSplit	Number of split pixel hits
nSCTHits	Number of SCT hits
nSCTShared	Number of shared SCT hits
muon_track	Indicates whether a track was used in reconstruction of the lepton

Model Architecture & Classification

PLIT uses flavour tagging tools and a transformer machine learning model.

Main task: Is the lepton prompt or not?

Auxiliary tasks improve the convergence of the main one:
- track origin prediction task
- vertex prediction task



Lepton truth labels are determined according to ATLAS truth classification tool:

- **prompt leptons**: coming from H/W/Z (including charge-flips)

- **non-prompt leptons**: coming from a b-, c-, l-hadron or hadronic tau decay; for electrons photon conversions and electrons from muon bremsstrahlung are taken into account

* leptons for which classification failed (“Unknown” type) not considered

Truth labels for tracks are obtained by examining the full ancestry of the truth particle linked to the track, and these are used for auxiliary tasks.

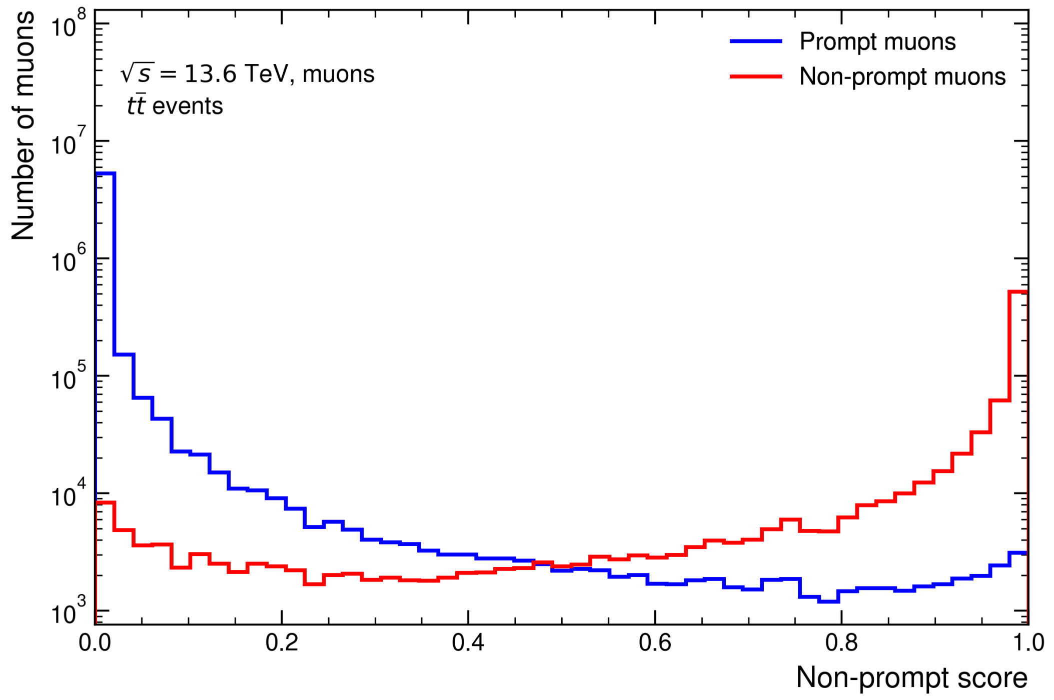
Model Training

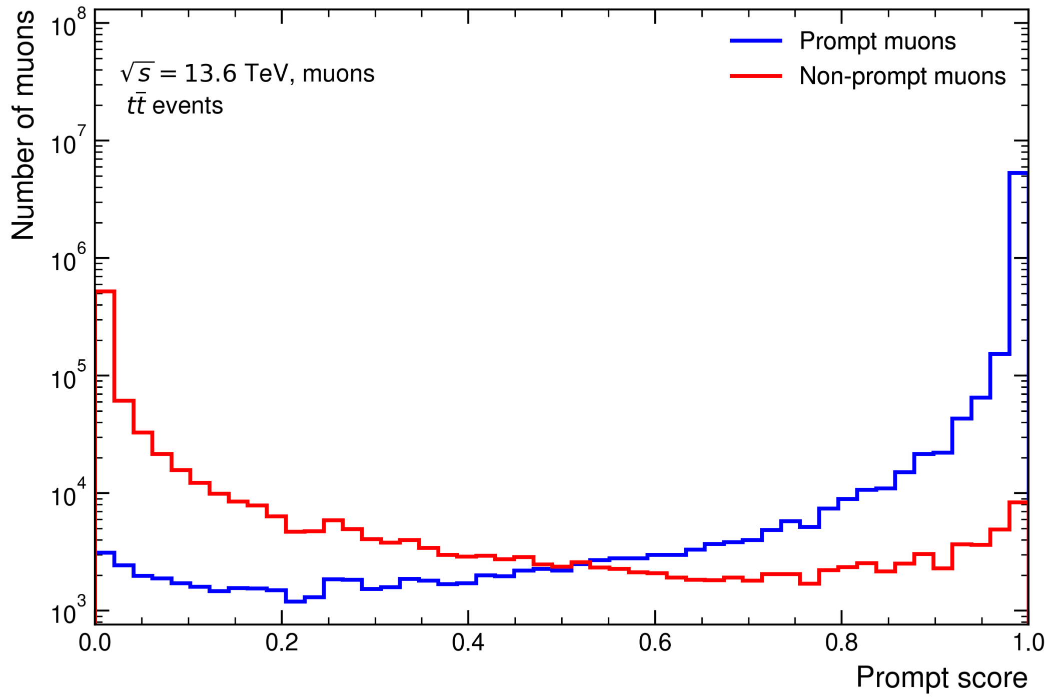
Lepton selection:
- **Medium muons and Loose electrons**
- p_T > 10 GeV
- $|\eta| < 2.5$ (μ)
- $|\eta| < 1.37$ (barrel e)
- $1.52 < |\eta| < 2.5$ (endcap e)
- $|d_0 / \sigma(d_0)| < 5$ (e) / 3 (μ)
- $|z_0 \sin\theta| < 0.5$

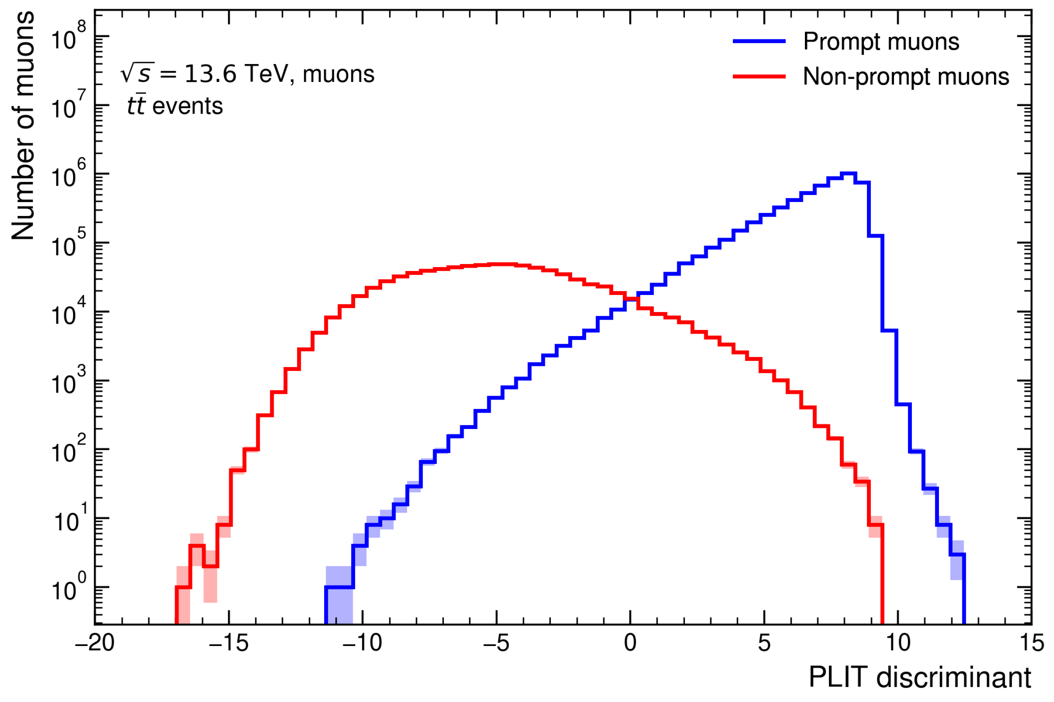
Track selection: identical to selection used for b-tagging in ATLAS (details in backup)

The trainings were done on Run 2 and Run 3 **ttbar MC** (nonallhad and allhad).

$$\text{PLIT discriminant} = \log\left(\frac{\text{prompt score}}{\text{non-prompt score}}\right)$$







PLIV Training Dataset Size

Muons: 8 million
Barrel electrons: 4.5 million
Endcap electrons: 1.4 million

PLIT Run 2 Training Dataset Size

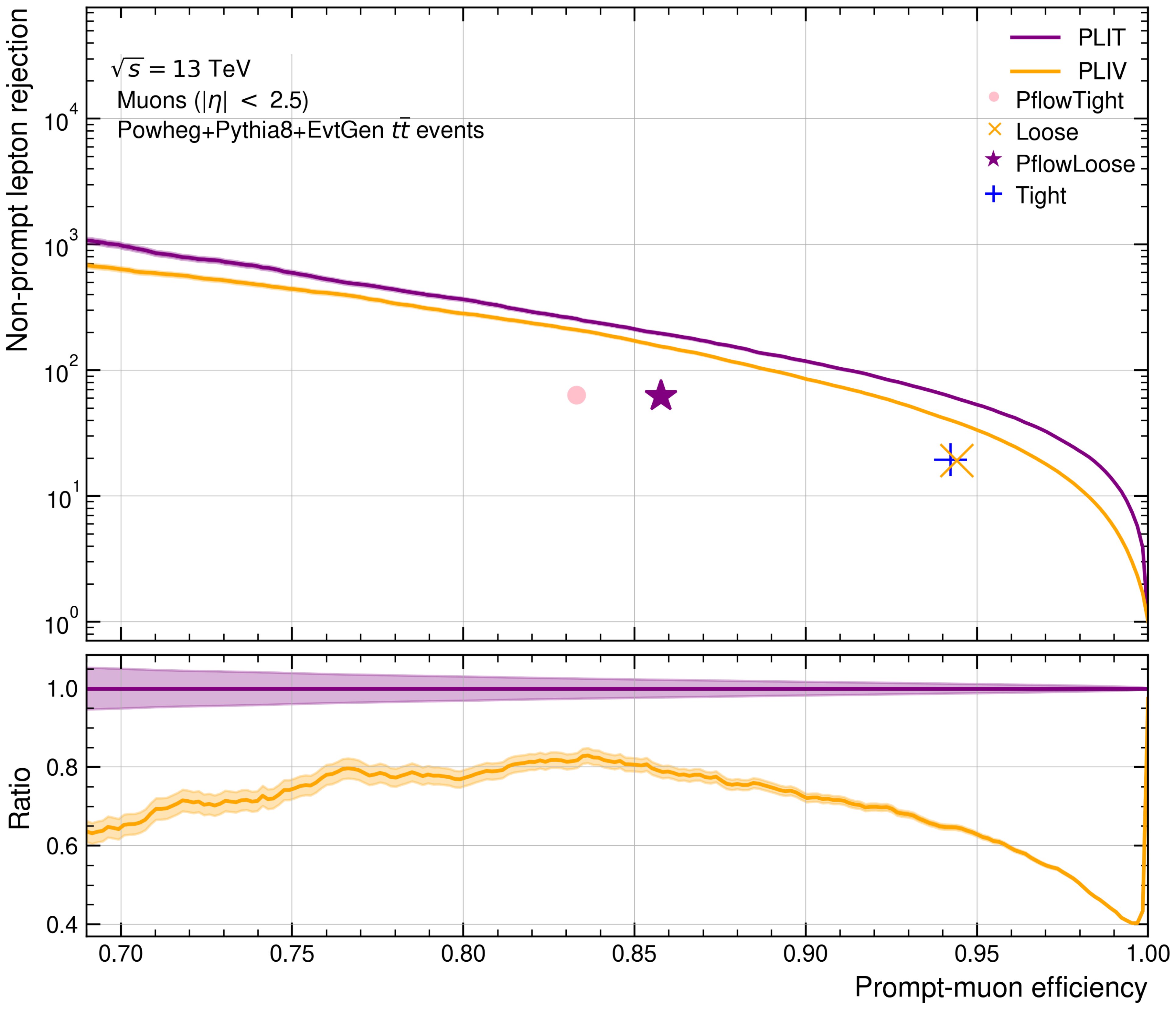
Muons: 35 million
Barrel electrons: 11.5 million
Endcap electrons: 5 million

PLIT Run 3 Training Dataset Size

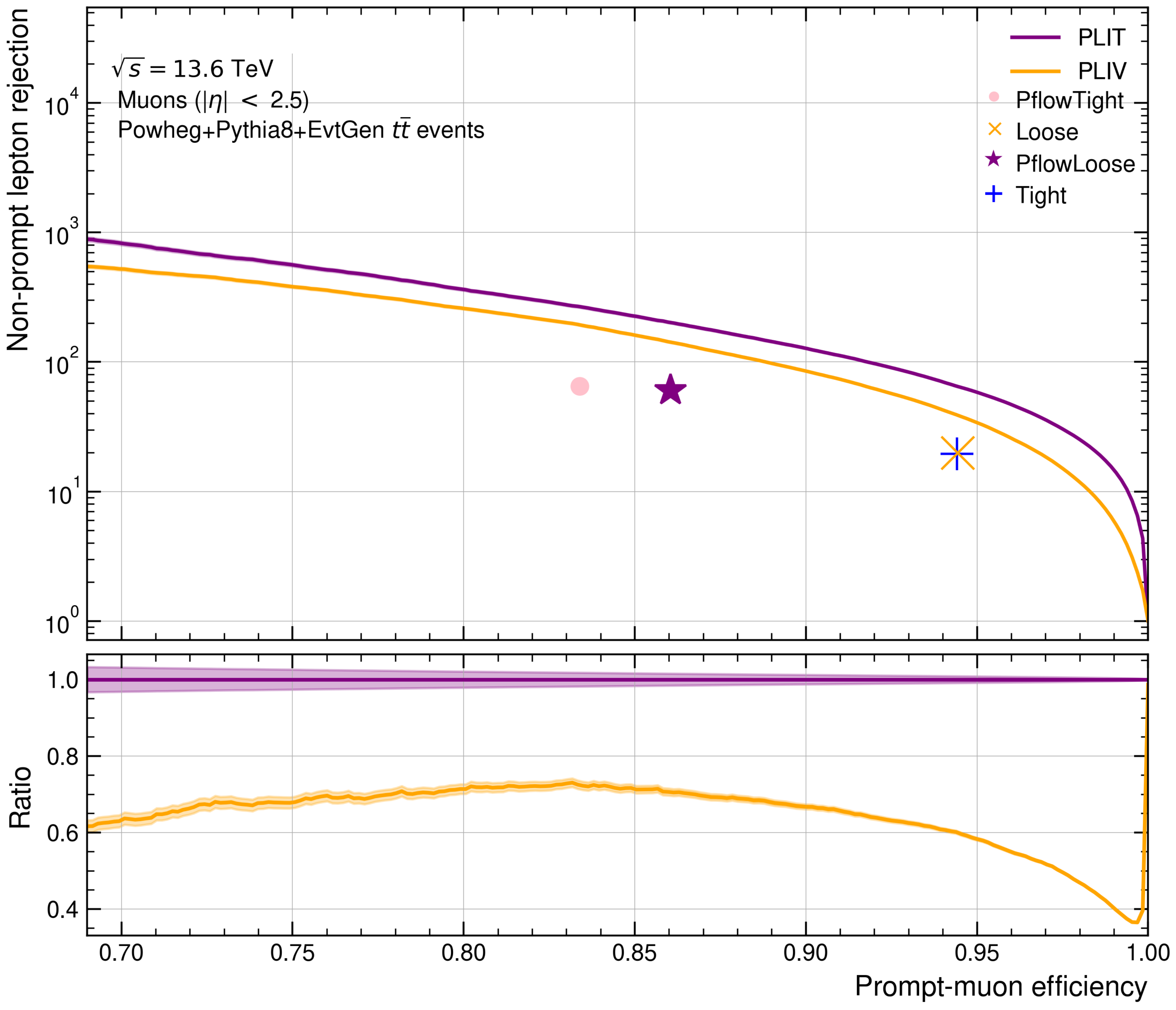
Muons: 75 million
Barrel electrons: 50 million
Endcap electrons: 18.5 million

Results for Muons

Run 2 Muons



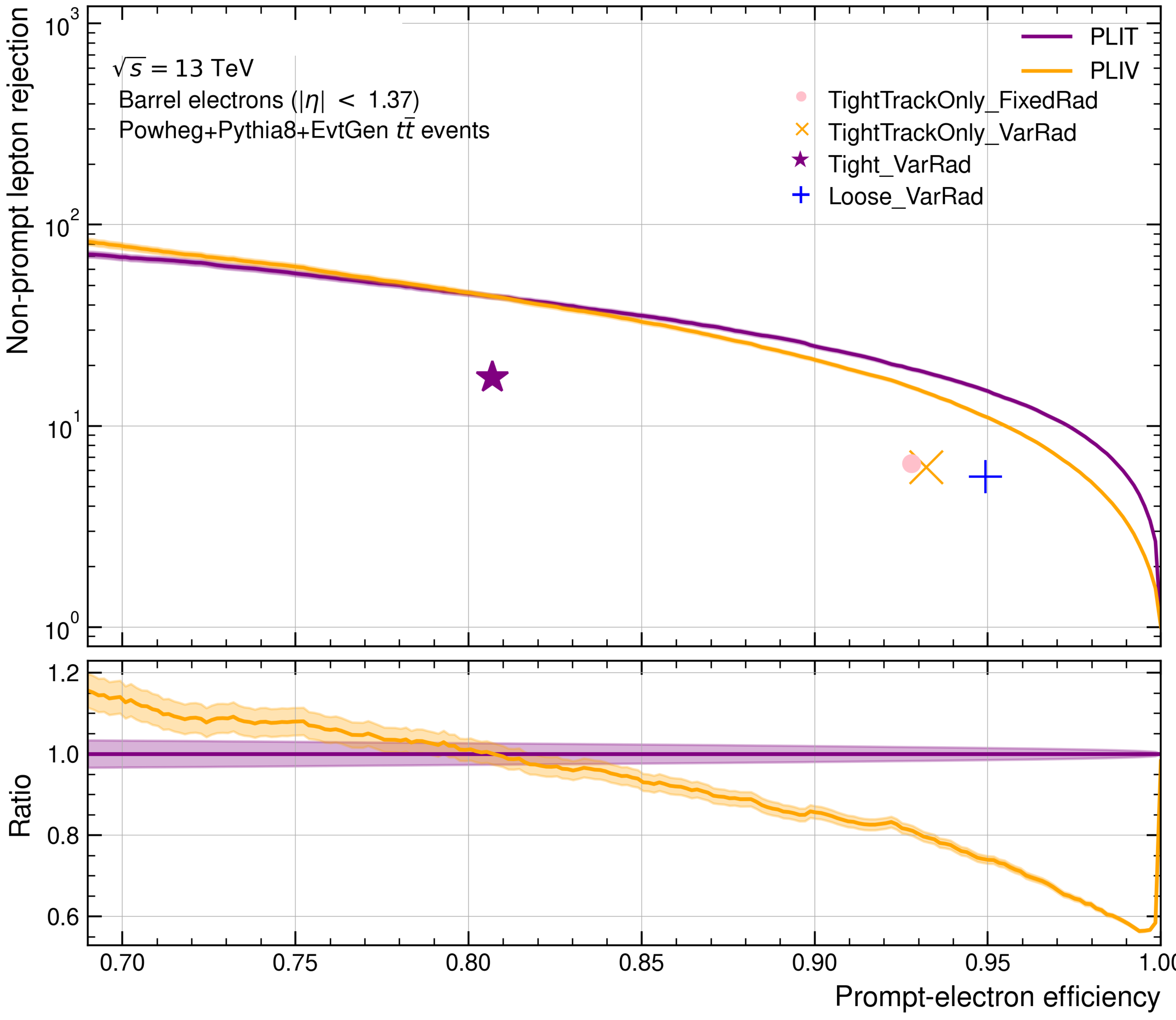
Run 3 Muons



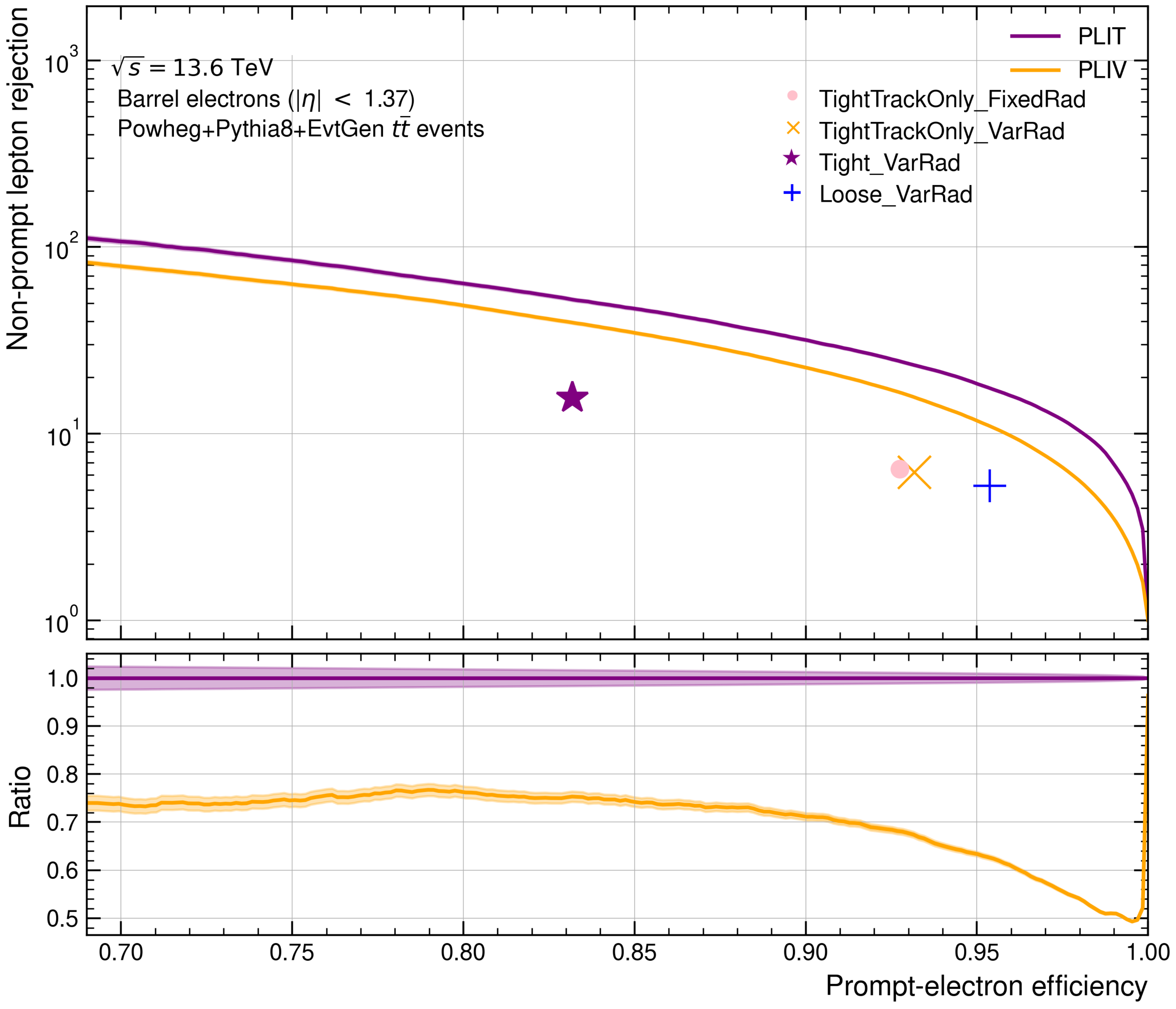
PLIT performs better than PLIV for all prompt-muon efficiencies.

Results for Barrel Electrons

Run 2 Barrel Electrons



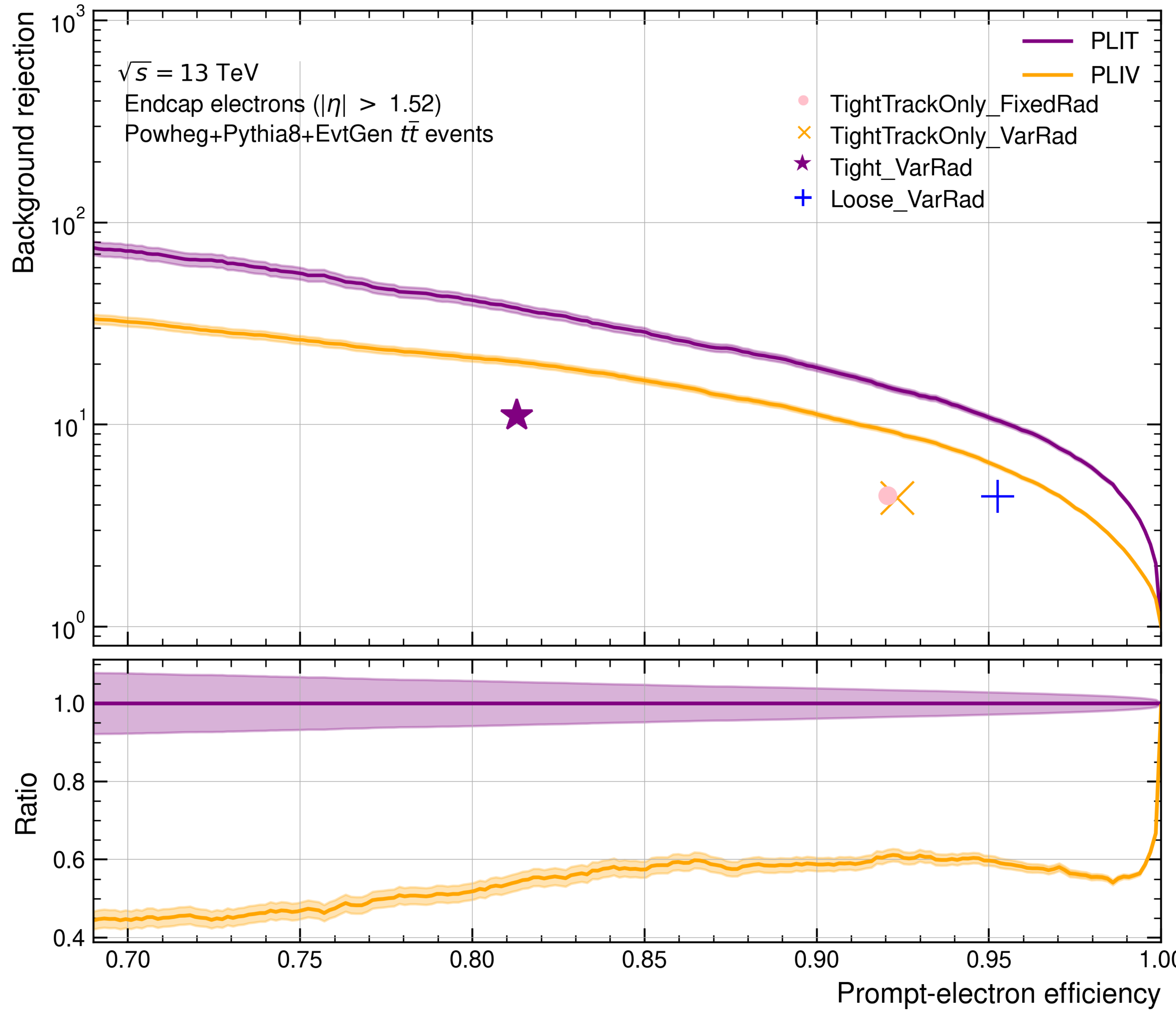
Run 3 Barrel Electrons



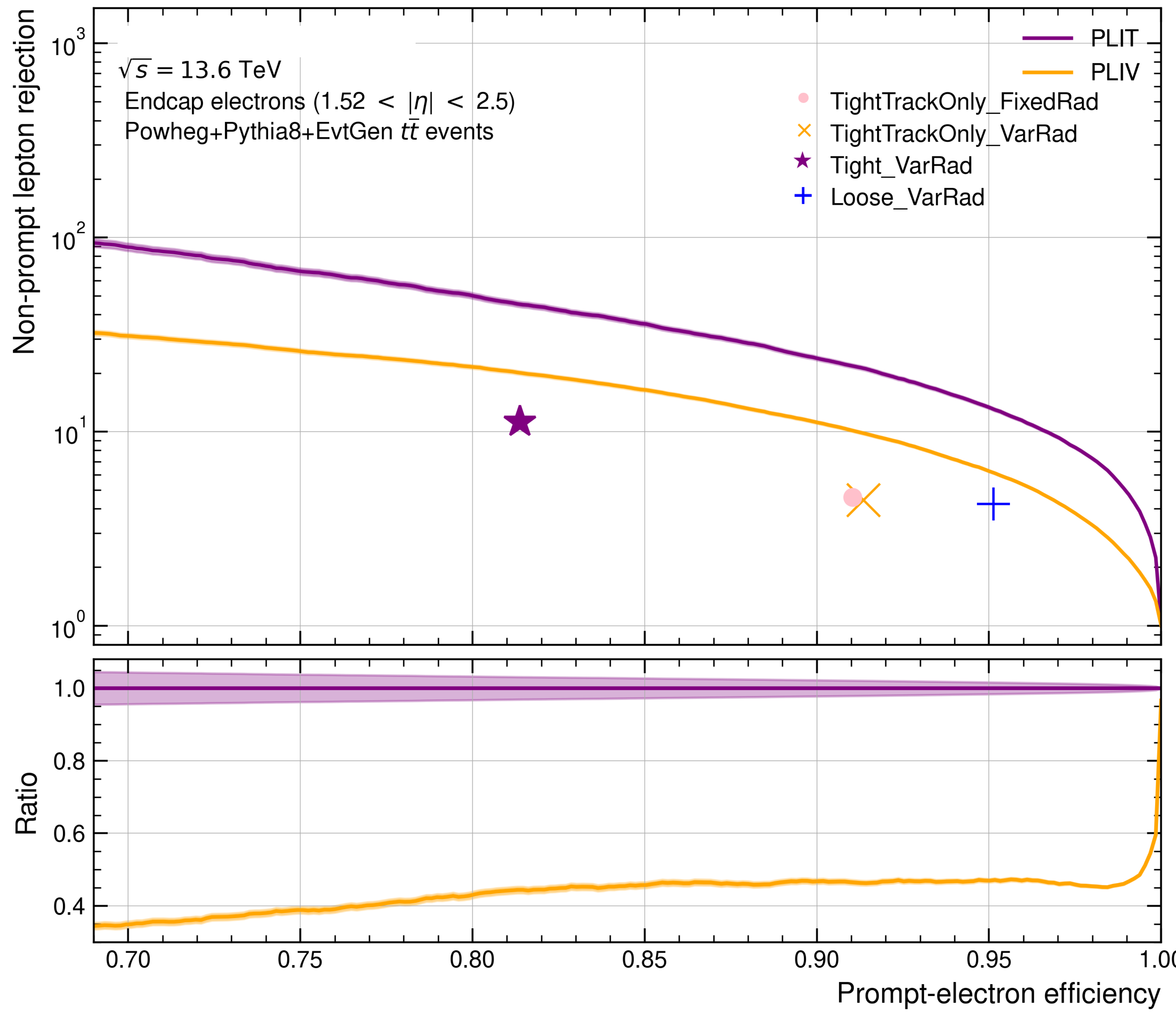
PLIT performs better than PLIV for all relevant prompt-electron efficiencies for Run 3.

Results for Endcap Electrons

Run 2 Endcap Electrons



Run 3 Endcap Electrons



PLIT performs better than PLIV for all prompt-electron efficiencies.

Ablation Study: Impact of dropping track jets

Ablation study involves comparing performance of a classifier before and after a change of its input variable list.

What are track jets?

- Track jets are reconstructed from inner detector tracks using the anti-k_t algorithm.

- They were introduced to ATLAS as a way to identify and tag b-hadron decays independently of calorimeter jet algorithms.

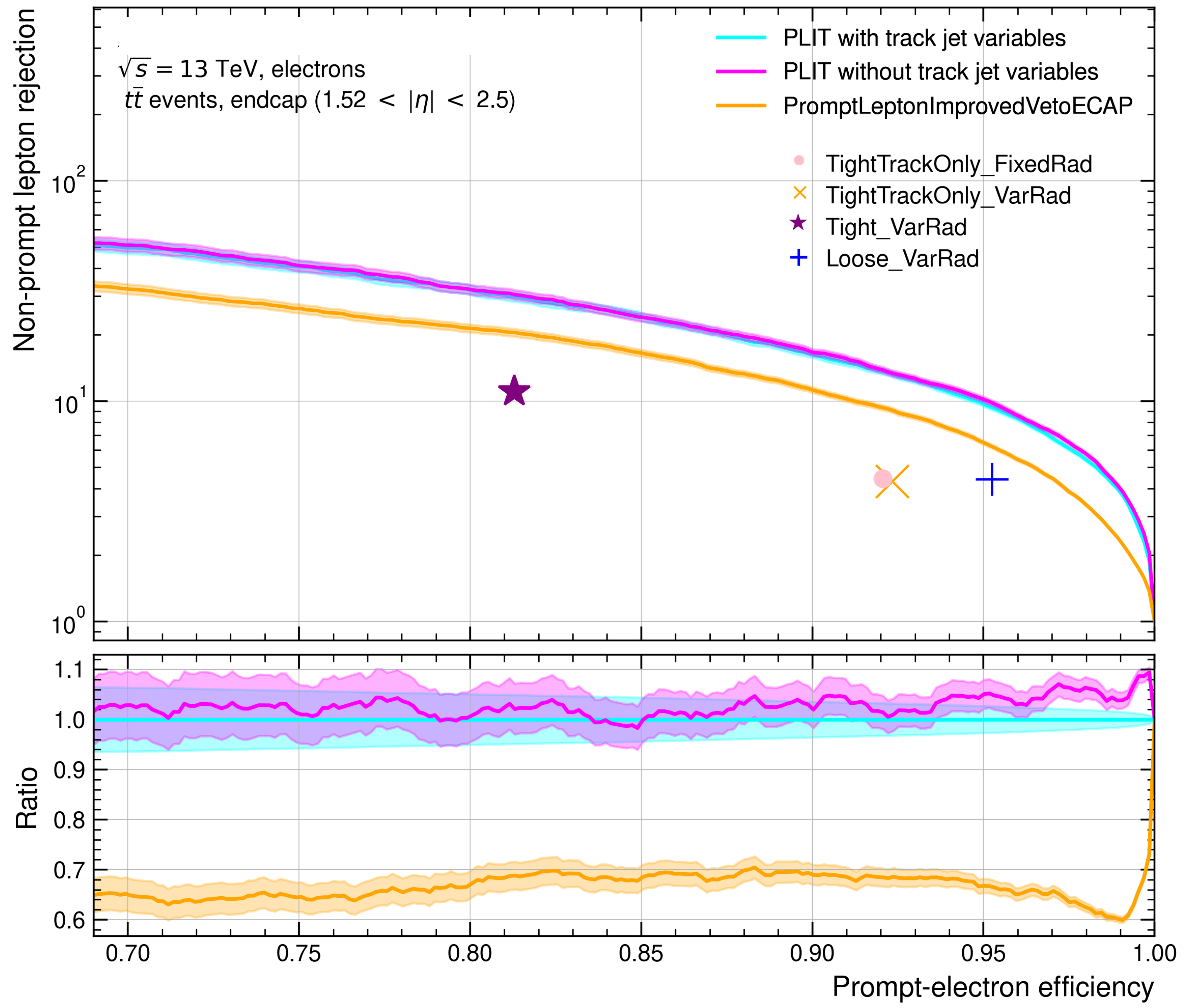
Q: Can we afford dropping them?

Markers on the left indicate the input variables that were dropped in the second training

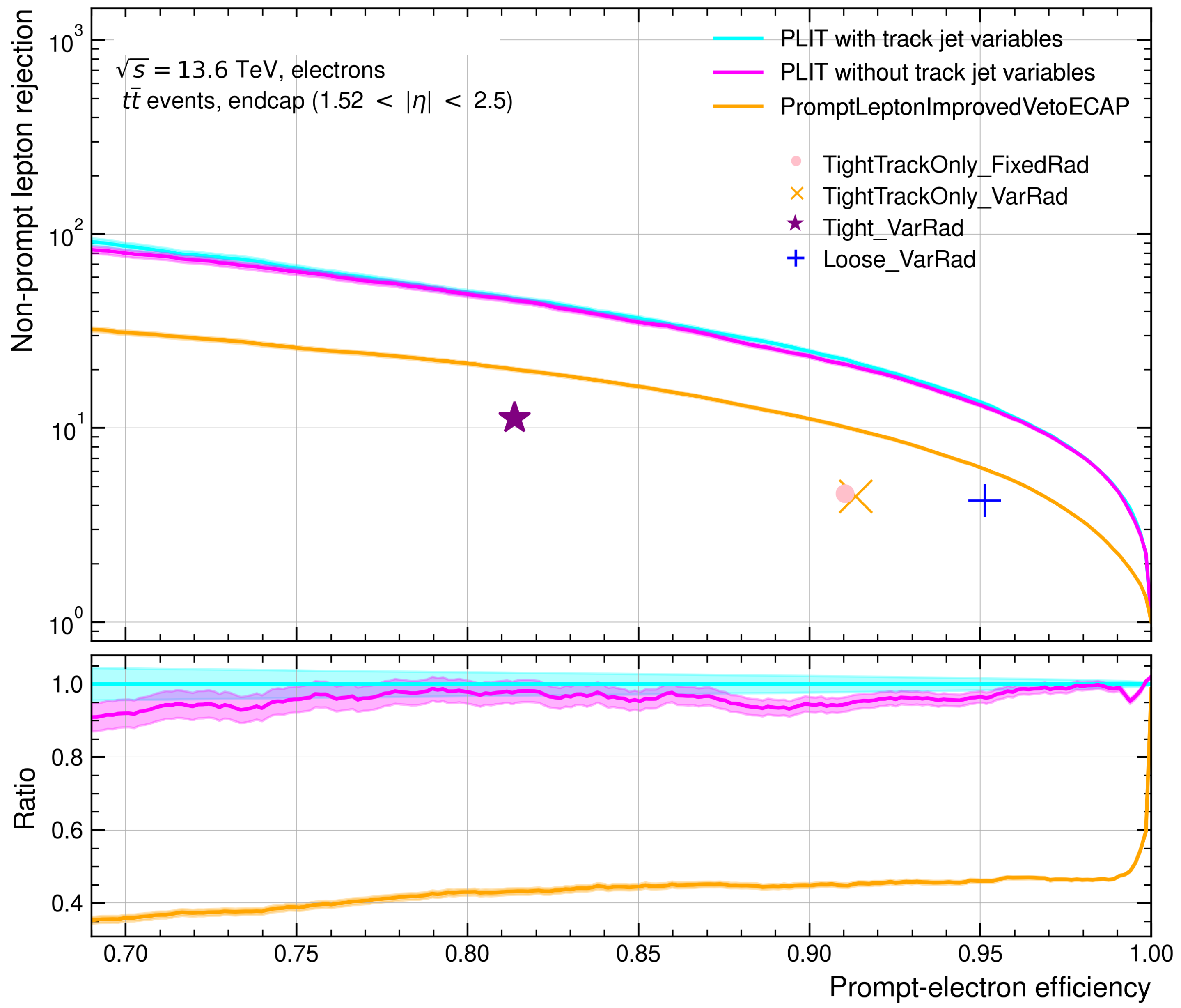
Muon Input	Description
p_{T}	Transverse momentum
η	Pseudorapidity
ϕ	Azimuthal angle
$p_{\mathrm{T}}^{\mathrm{varcone30}}/p_{\mathrm{T}}$	Scalar sum of the $p_{\mathrm{T}}^{\mathrm{track}}$ of the inner detector tracks associated with the primary vertex in an η–ϕ cone of $\Delta R < 0.3$ around the muon, excluding the muon track itself, divided by the muon p_{T}
$E_{\mathrm{T}}^{\mathrm{varcone30}}/p_{\mathrm{T}}$	Sum of the transverse energy of topological cell clusters in a cone of size $\Delta R = 0.3$ around the position of the muon, extrapolated to the calorimeters, after subtracting the contribution from the energy deposit of the muon itself and correcting for pile-up effects, divided by the muon p_{T}
caloClusterERel	Energy of the calorimeter cluster associated with the muon divided by the muon's expected energy loss in calorimeter
p_{T} / $p_{\mathrm{T}}^{\mathrm{track\ jet}}$	Lepton transverse momentum divided by transverse momentum of closest track jet
$p_{\mathrm{T}}^{\mathrm{rel}}$	Transverse momentum of the lepton with respect to the track jet axis
$\Delta R(\mathrm{muon, track\ jet})$	Distance between lepton and closest track jet
$n(\mathrm{tracks})^{\mathrm{trackjet}}$	Number of inner detector tracks associated with the closest track jet
Track Input	**Description**
$p_{\mathrm{T}}^{\mathrm{track}}$ / $p_{\mathrm{T}}^{\mathrm{track\ jet}}$	Track transverse momentum divided by the track jet transverse momentum
$\Delta R(\mathrm{track, track\ jet})$	Distance between track and closest track jet
$\Delta\eta(\mathrm{track, lepton})$	Difference in η between track and lepton
$\Delta\phi(\mathrm{track, lepton})$	Difference in ϕ between track and lepton
q^{track} / p^{track}	Track charge divided by momentum
d_0	Transverse impact parameter
$z_0 \sin\theta$	Longitudinal impact parameter times sine of the polar angle
d_0 uncertainty	Uncertainty of the transverse impact parameter
$z_0 \sin\theta$ uncertainty	Uncertainty of the longitudinal impact parameter times sine of the polar angle
d_0 significance	Significance of the transverse impact parameter
$z_0 \sin\theta$ significance	Significance of the longitudinal impact parameter times sine of the polar angle
nPixHits	Number of pixel hits
nIBLHits	Number of IBL hits
nBLHits	Number of B-layer hits
nIBLShared	Number of shared IBL hits
nIBLSplit	Number of split IBL hits
nPixShared	Number of shared pixel hits
nPixSplit	Number of split pixel hits
nSCTHits	Number of SCT hits
nSCTShared	Number of shared SCT hits
muon_track	Indicates whether a track was used in reconstruction of the lepton

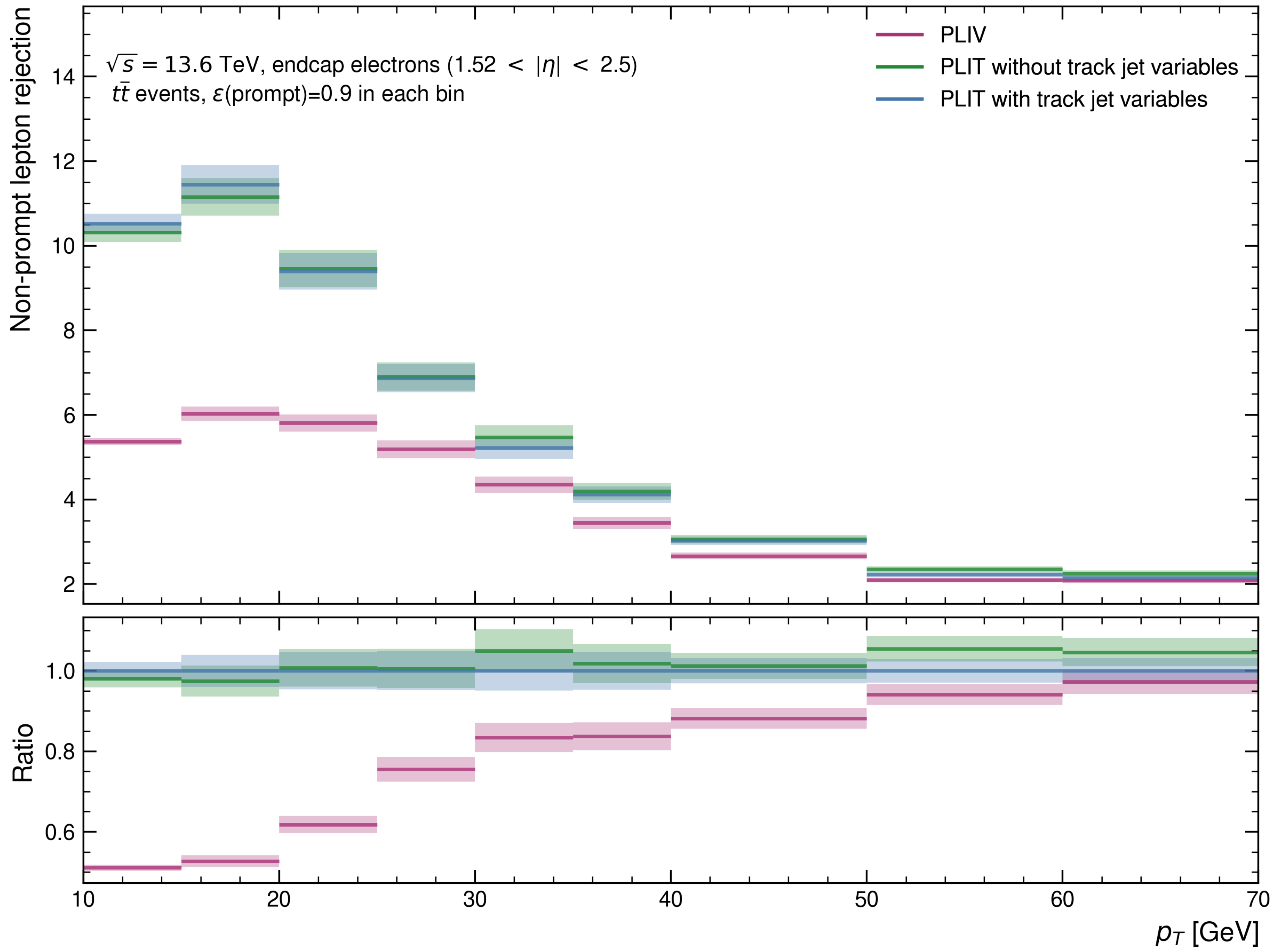
Ablation Study Results

Run 2 Endcap Electrons



Run 3 Endcap Electrons



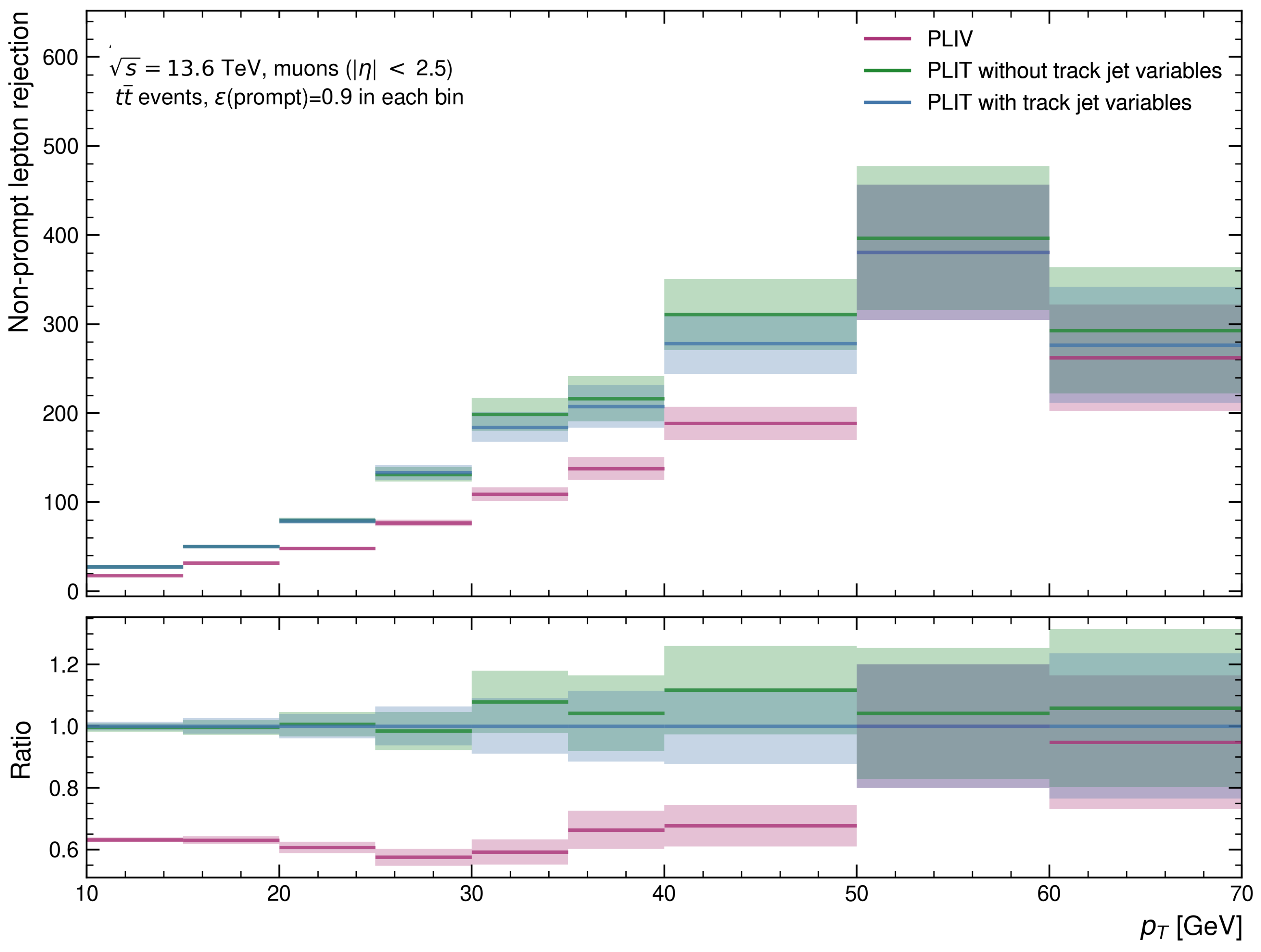


PLIT performs better than PLIV for all prompt-electron efficiencies.

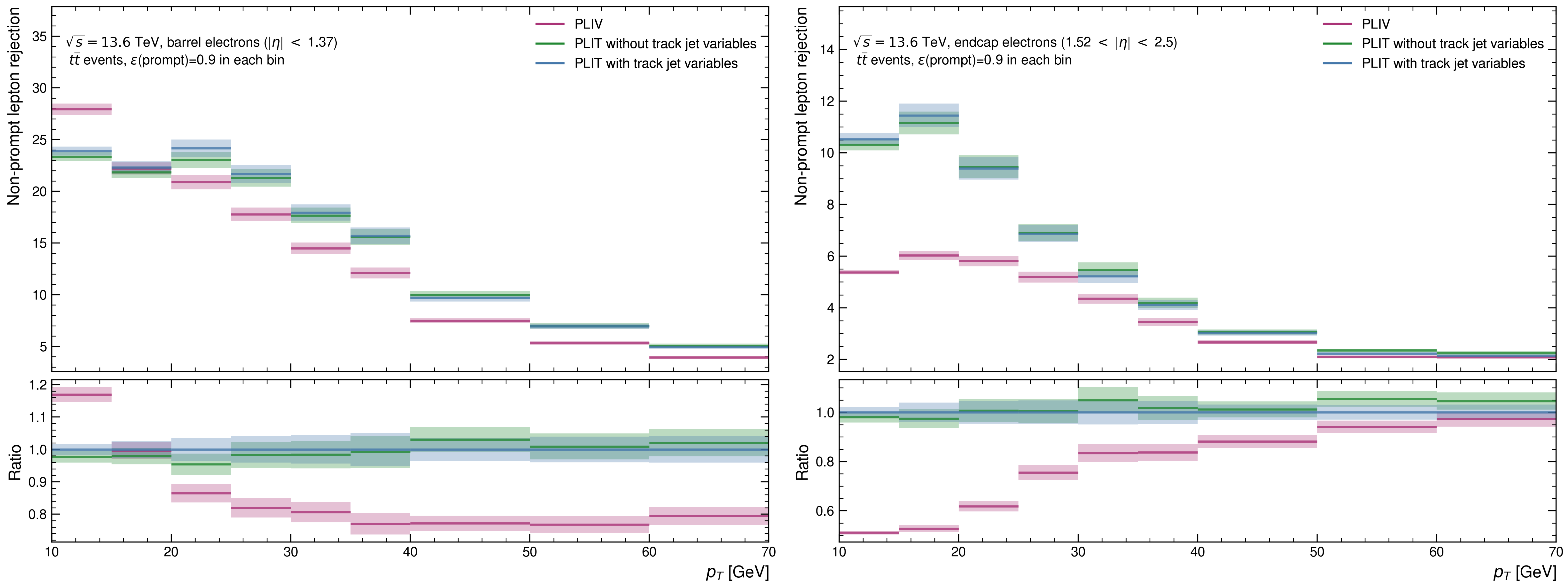
Removing track jet variables does not affect the performance for electrons!

Background rejection vs pT

Run 3 performance



Muons

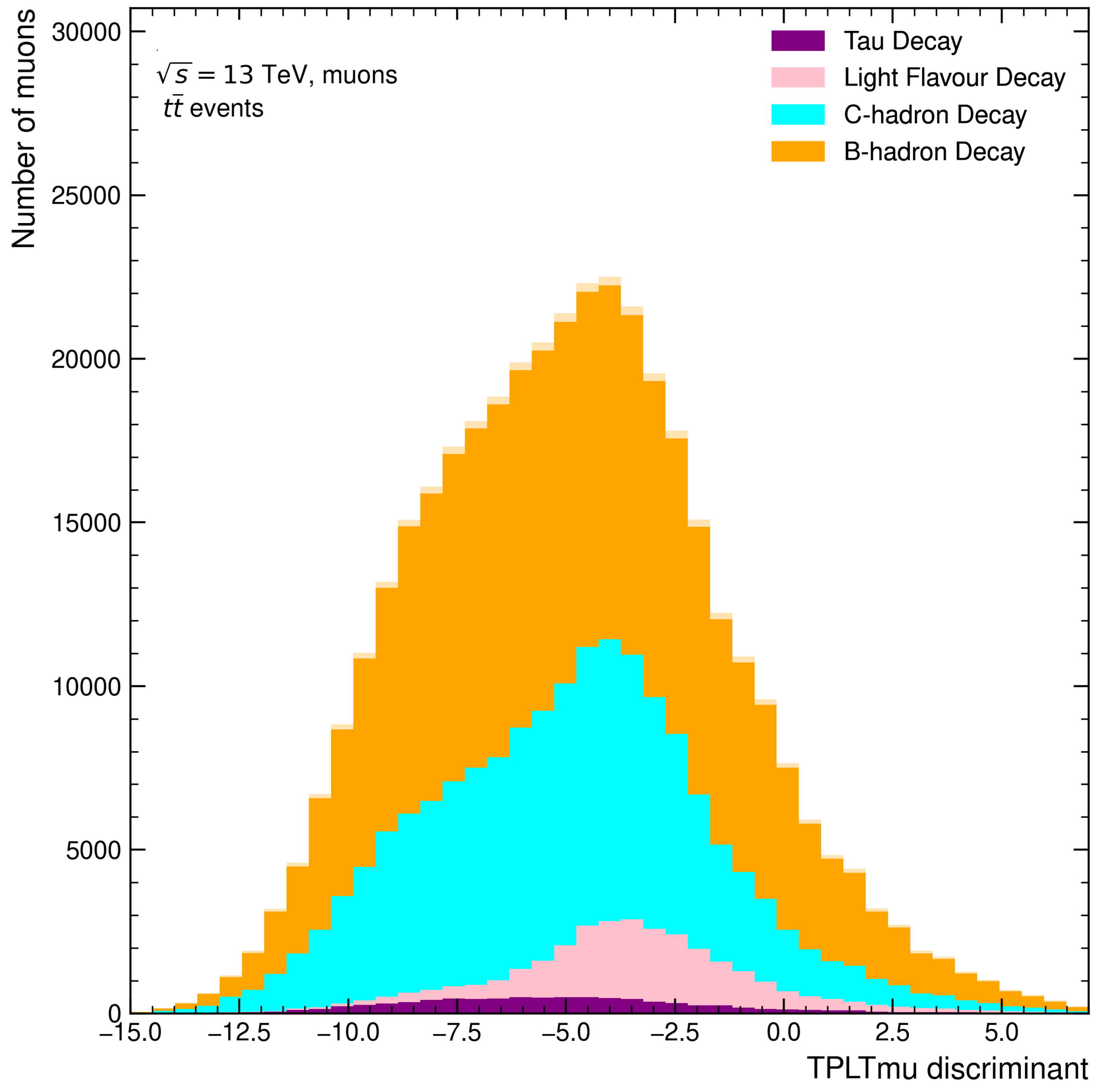


Electrons

Background Composition

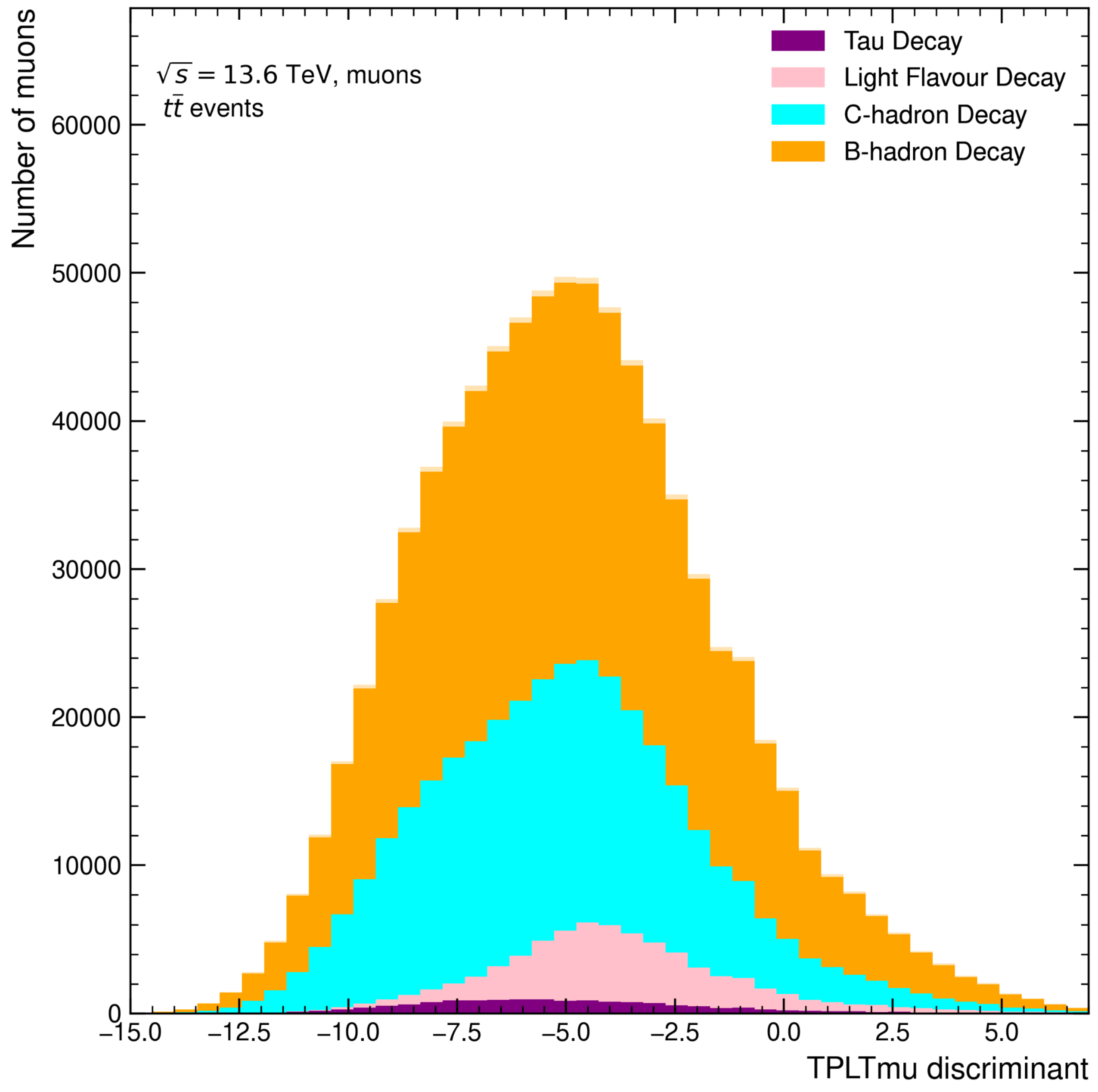
Goal: Understanding distributions of PLIT discriminants for different backgrounds

Run 2 Muons



	% of entries
Tau decays	2.2%
Light decays	6.5%
C-decays	35.2%
B-decays	56.1%

Run 3 Muons

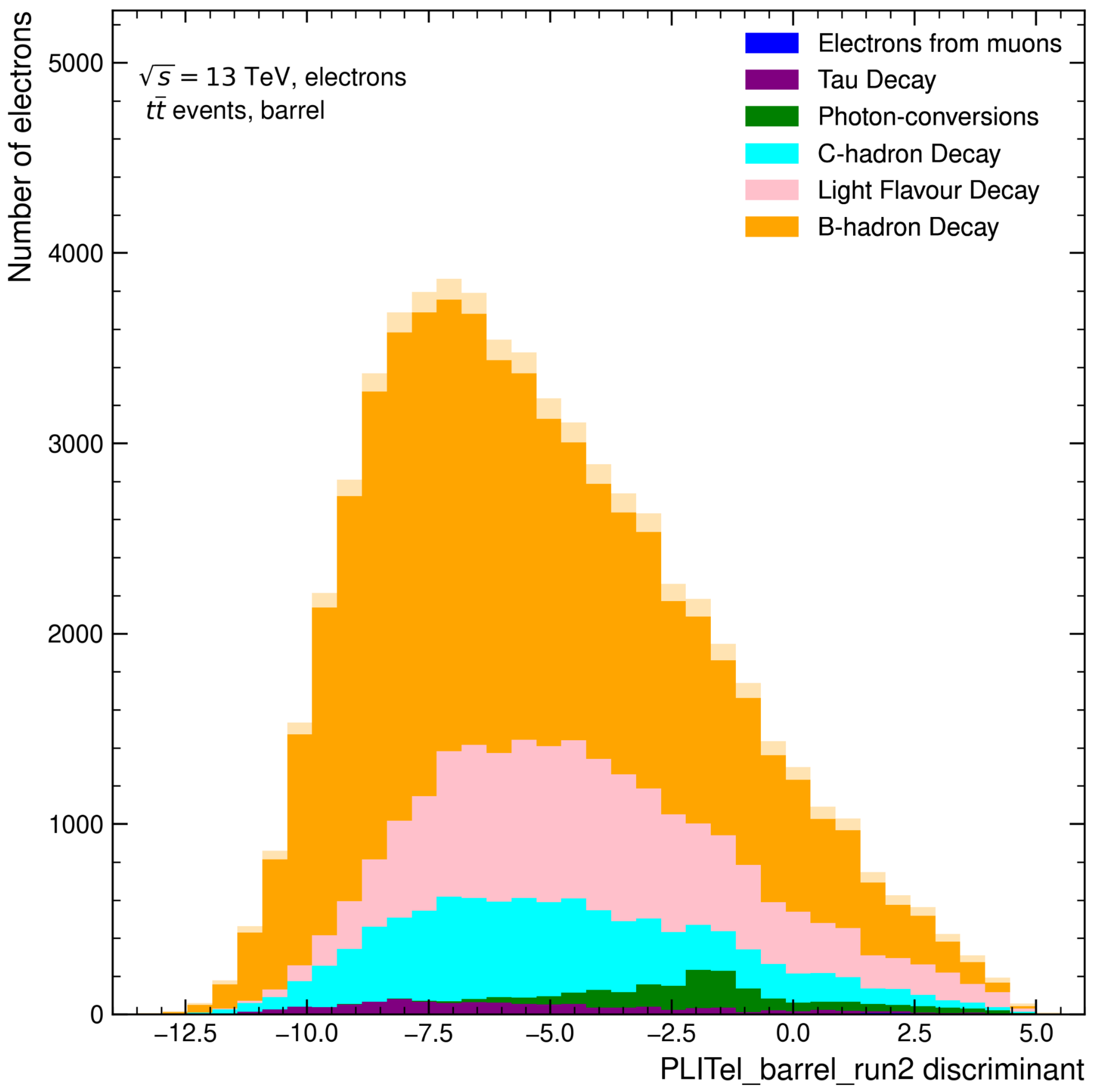


	% of entries
Tau decays	1.8%
Light decays	6.7%
C-decays	34.6%
B-decays	56.9%

Background Composition

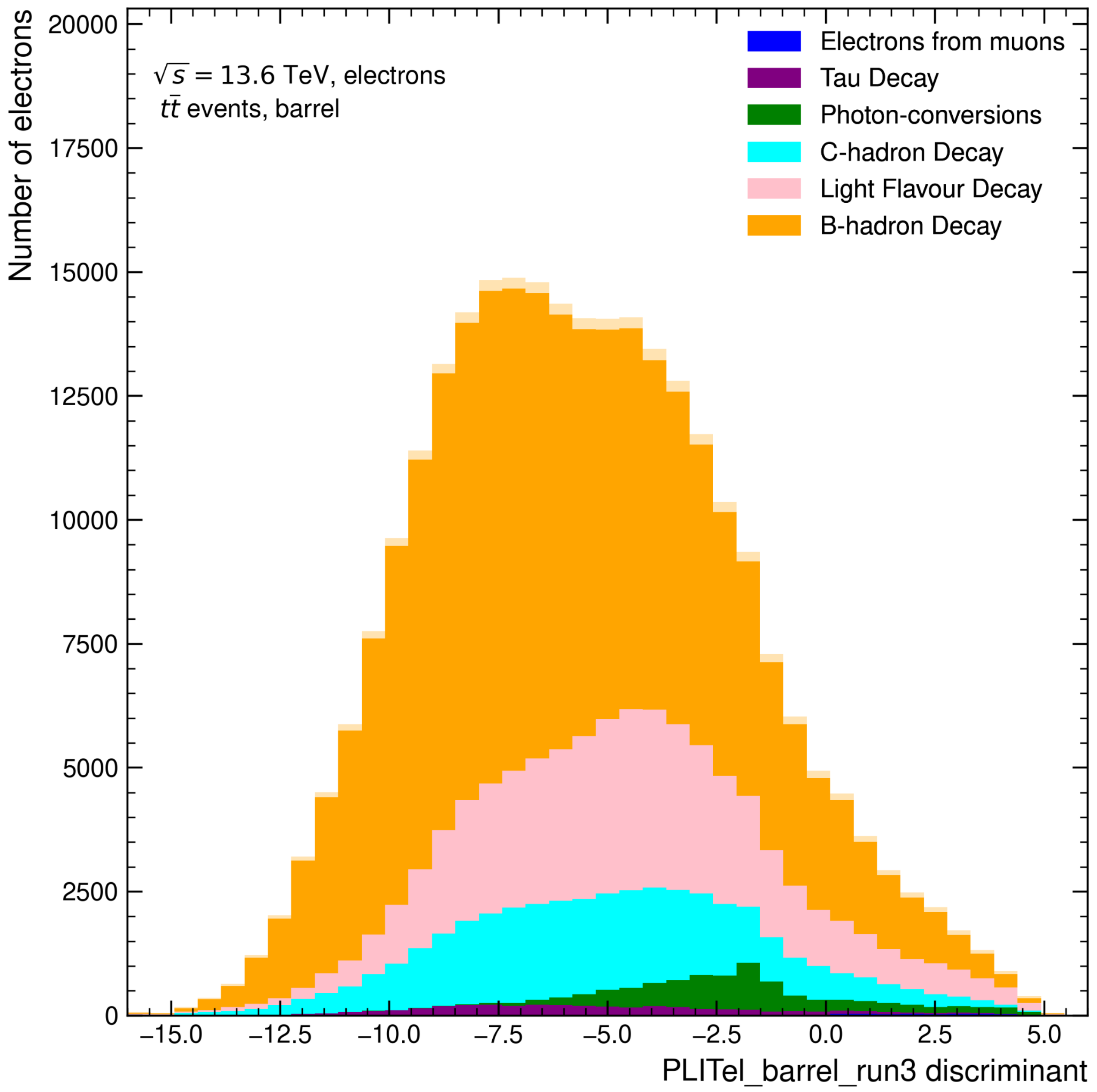
Goal: Understanding distributions of PLIT discriminants for different backgrounds

Run 2 Barrel Electrons



	% of entries
Electrons from Muons	< 0.1%
Tau decays	1.7%
Photon conversions	2.5%
C-decays	13.2%
Light decays	21.4%
B-decays	61.1%

Run 3 Barrel Electrons

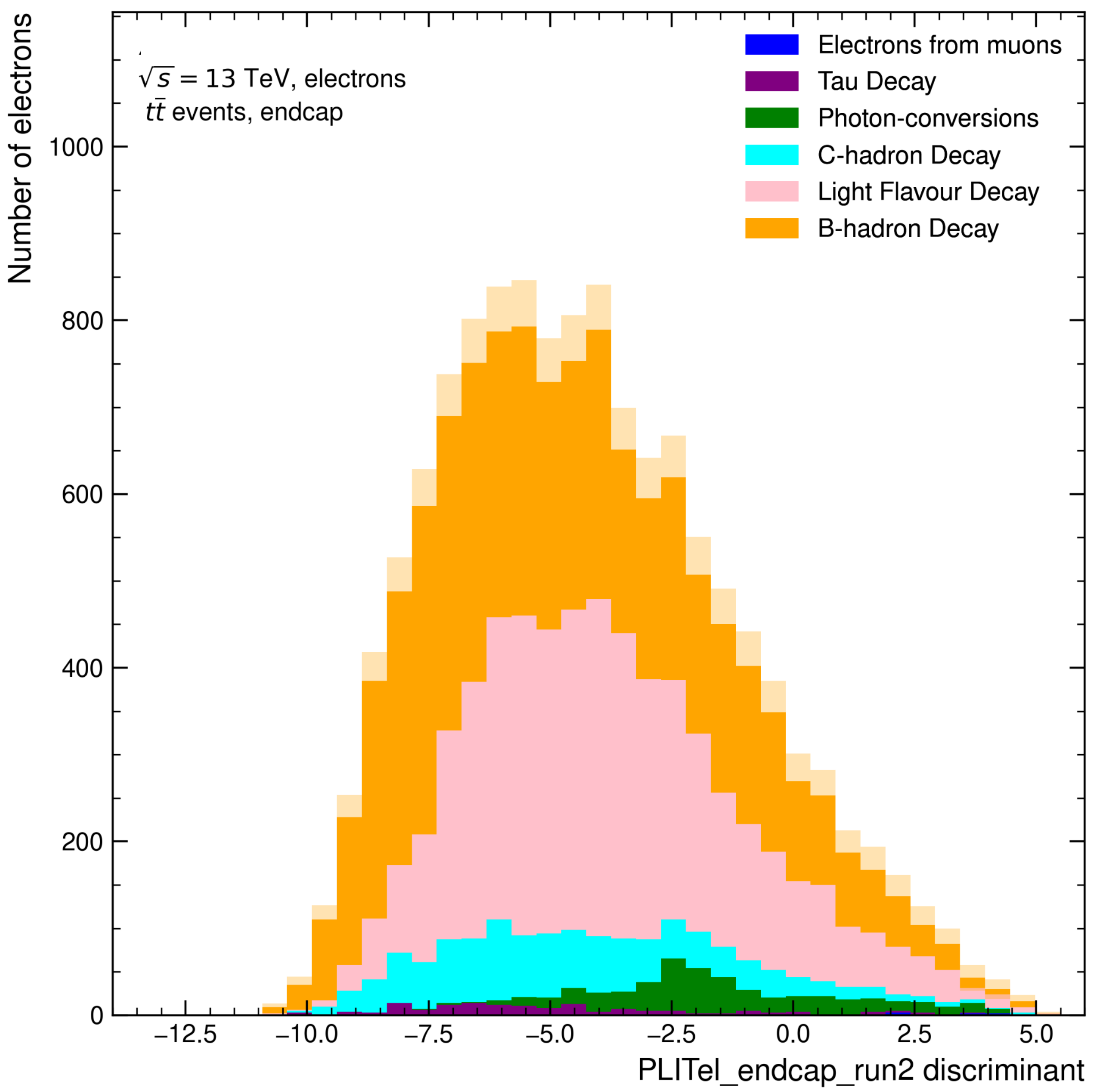


	% of entries
Electrons from Muons	< 0.1%
Tau decays	1.3%
Photon conversions	2.7%
C-decays	12.8%
Light decays	21 %
B-decays	62.1%

Background Composition

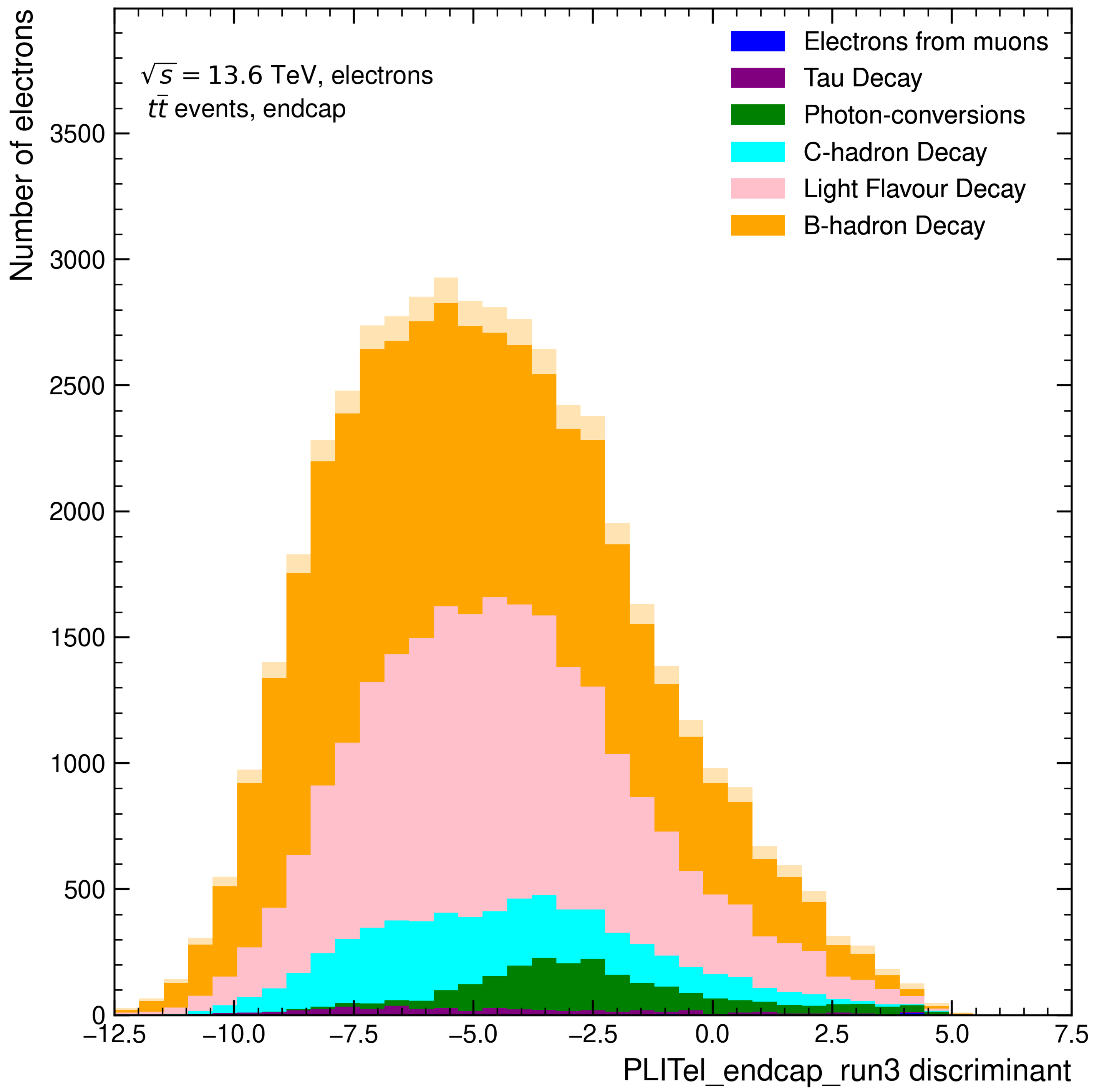
Goal: Understanding distributions of PLIT discriminants for different backgrounds

Run 2 Endcap Electrons



	% of entries
Electrons from Muons	< 0.1%
Tau decays	1.2%
Photon conversions	3.7%
C-decays	9.1%
Light decays	40.6%
B-decays	45.3%

Run 3 Endcap Electrons

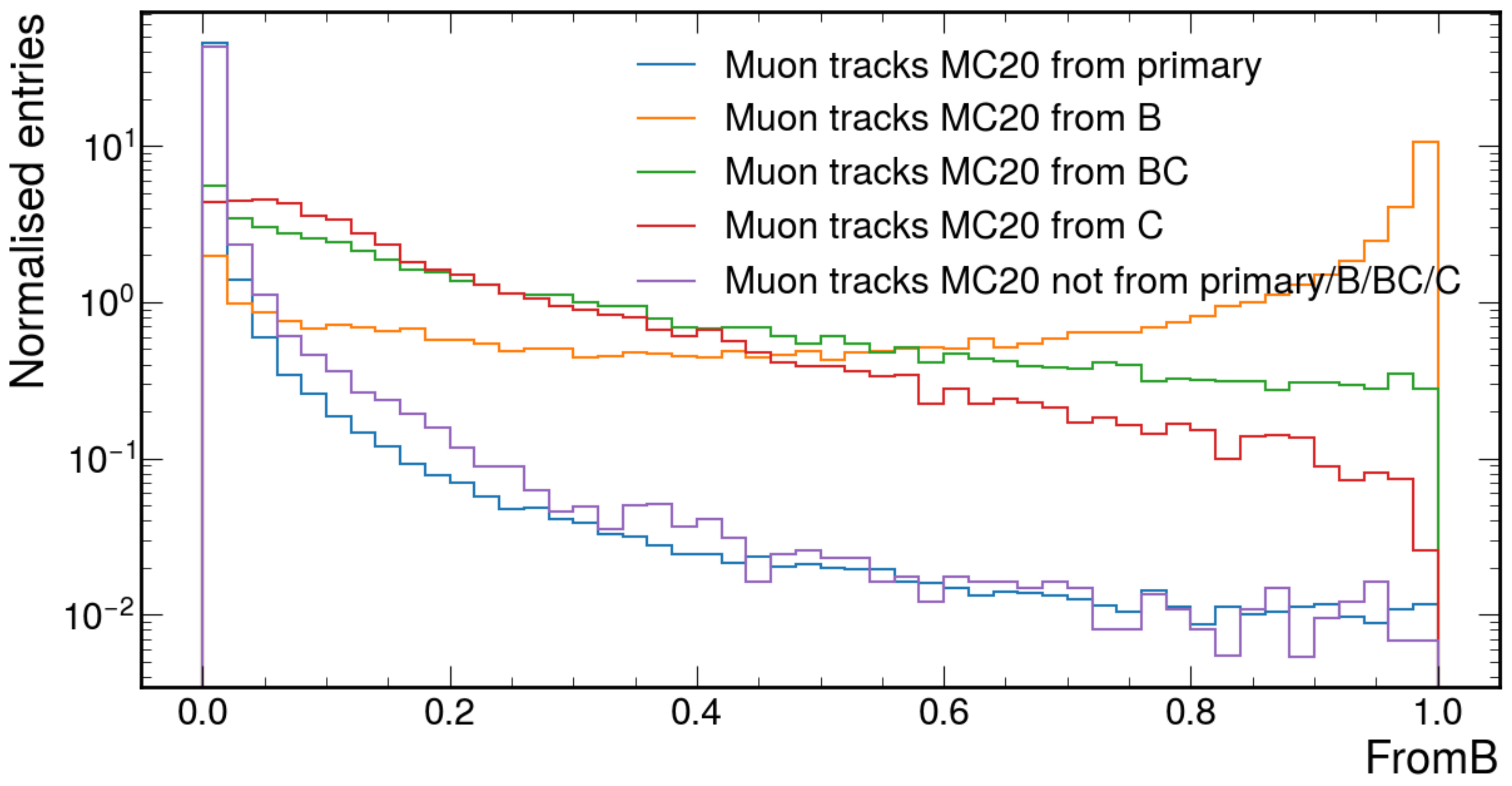


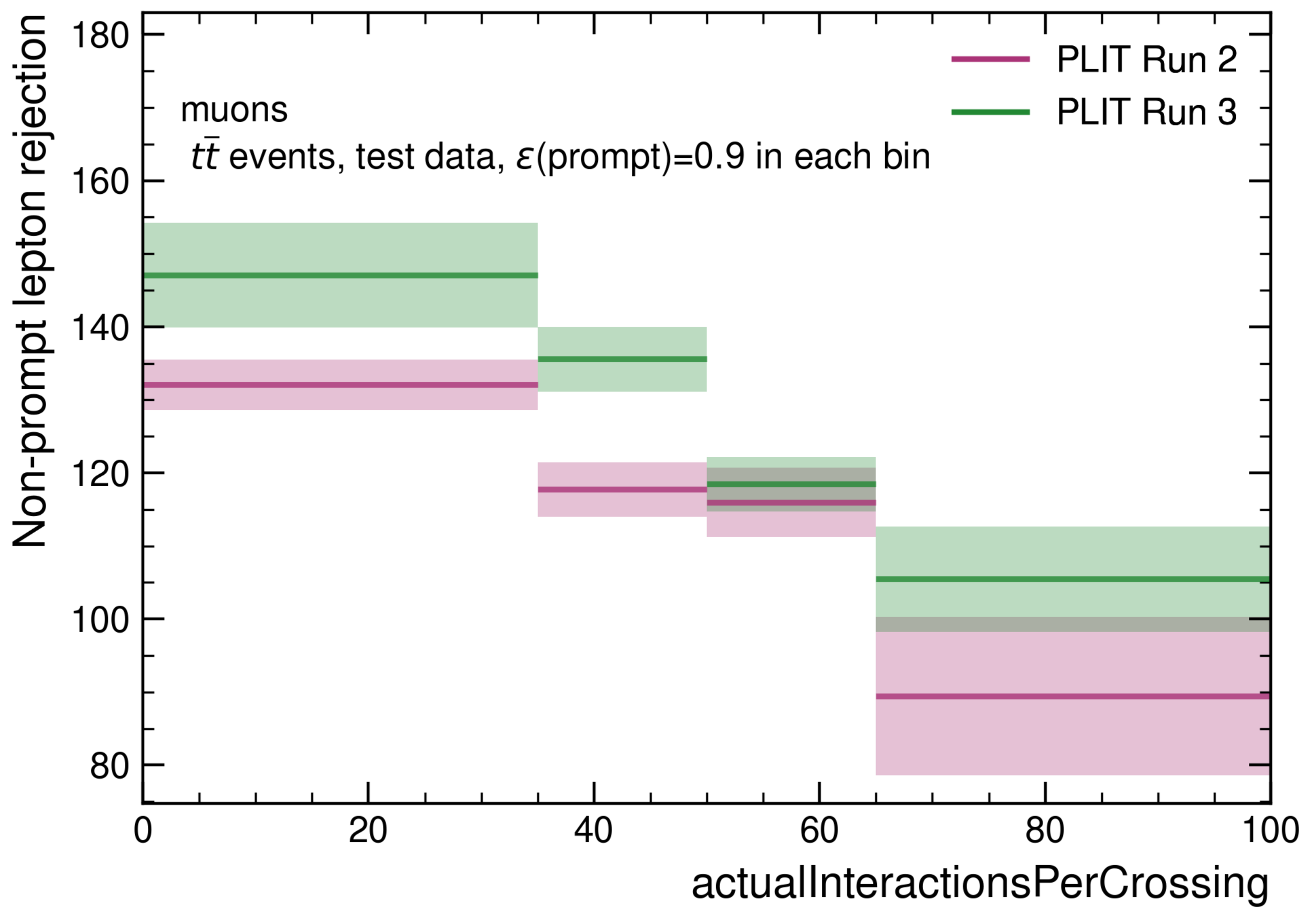
	% of entries
Electrons from Muons	< 0.1%
Tau decays	1 %
Photon conversions	4.3 %
C-decays	9.7%
Light decays	37.6%
B-decays	47.4%

Optimisation and validation of the training

Many studies were done to understand PLIT performance better:

- Examination of input distributions, studies on choice of input variables
- Background composition studies
- Impact of pile-up on the performance
- Validation of auxiliary tasks
- Training stability studies
- Impact of having a separate training for low p_T muons
- Impact of isolation pre-selection on training and on performance relatively to PLIV





Conclusions

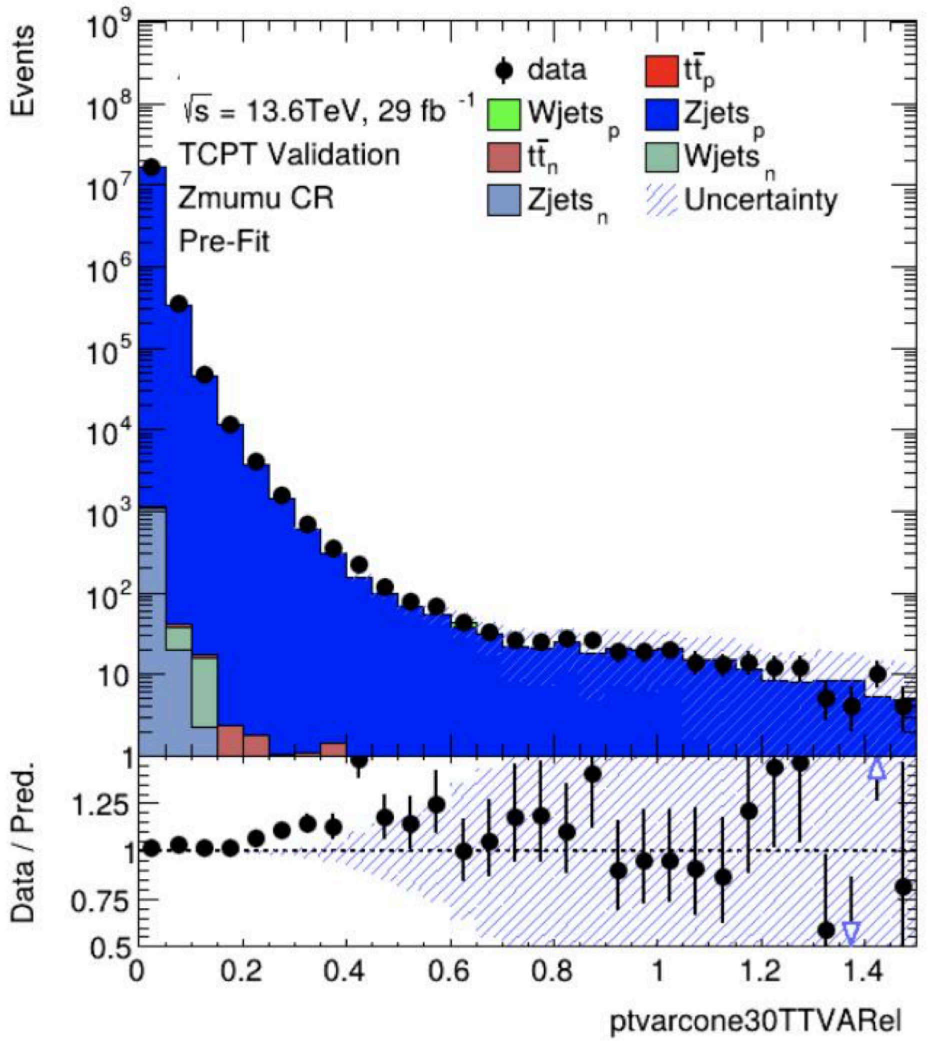
- First version of PLIT is ready.

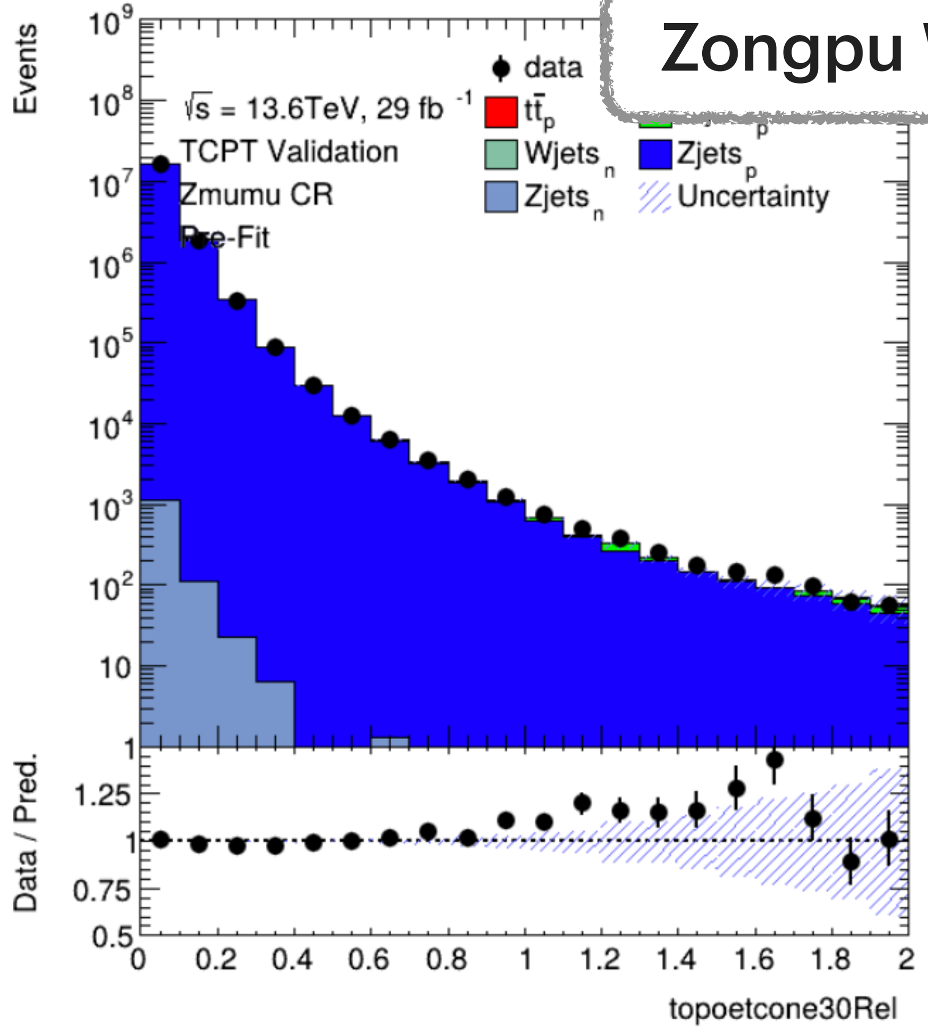
- **PLIT is performing better than PLIV in MC.**

- The current version of PLIT has been implemented in ATLAS central software.

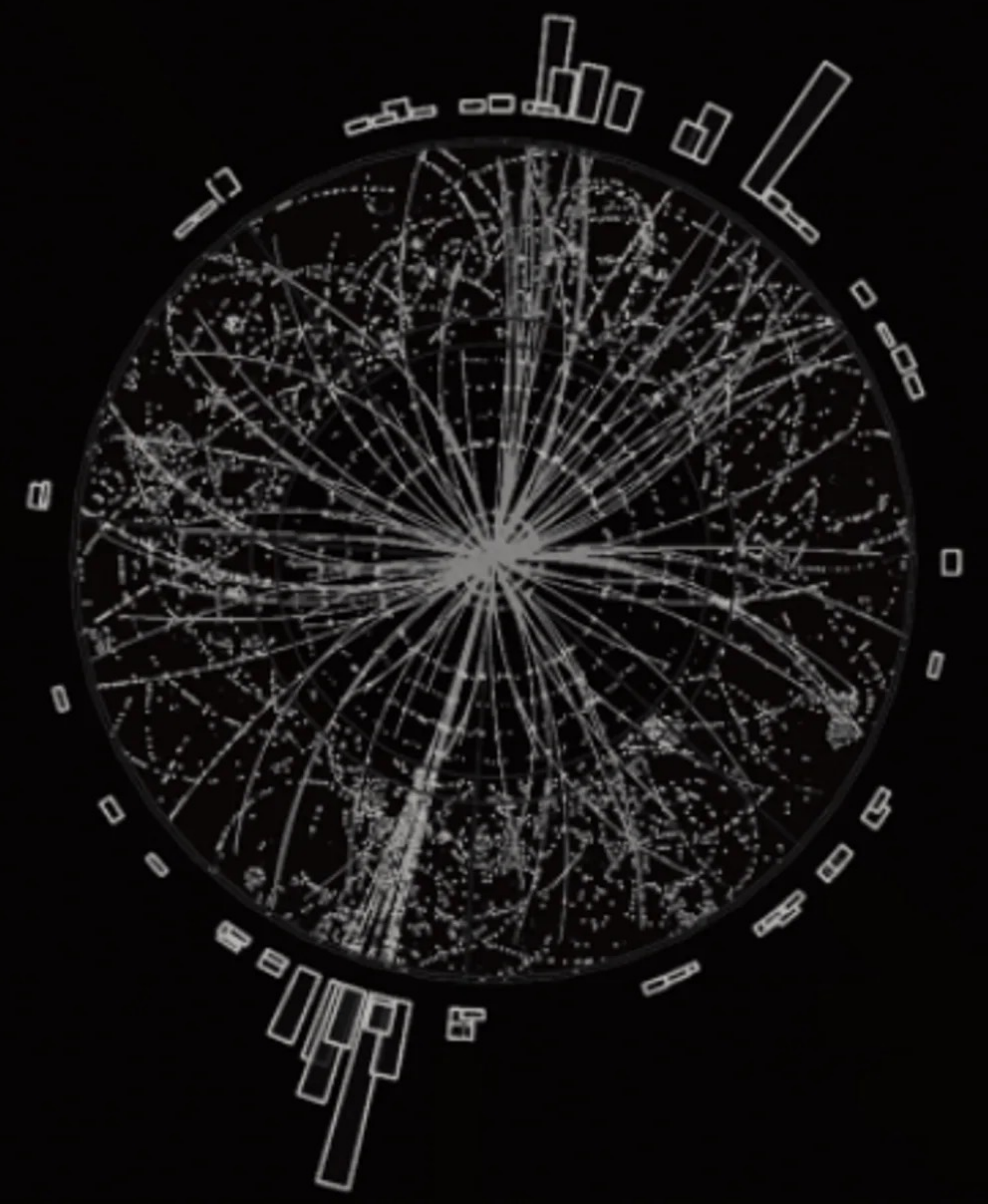
- Next steps:
  - validation of inputs in data
  - working point definition
  - calibration

Work by
Zongpu Wang

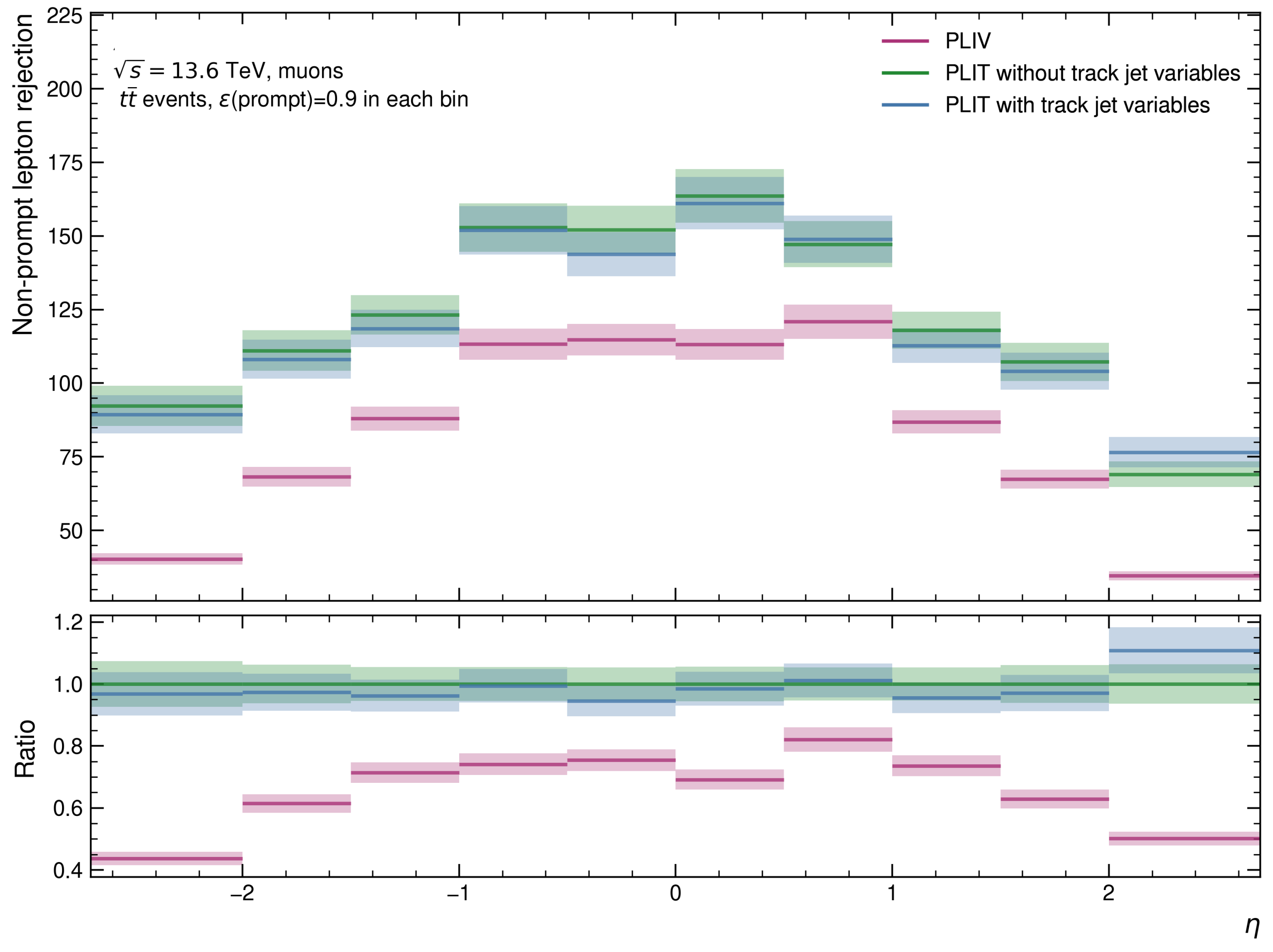


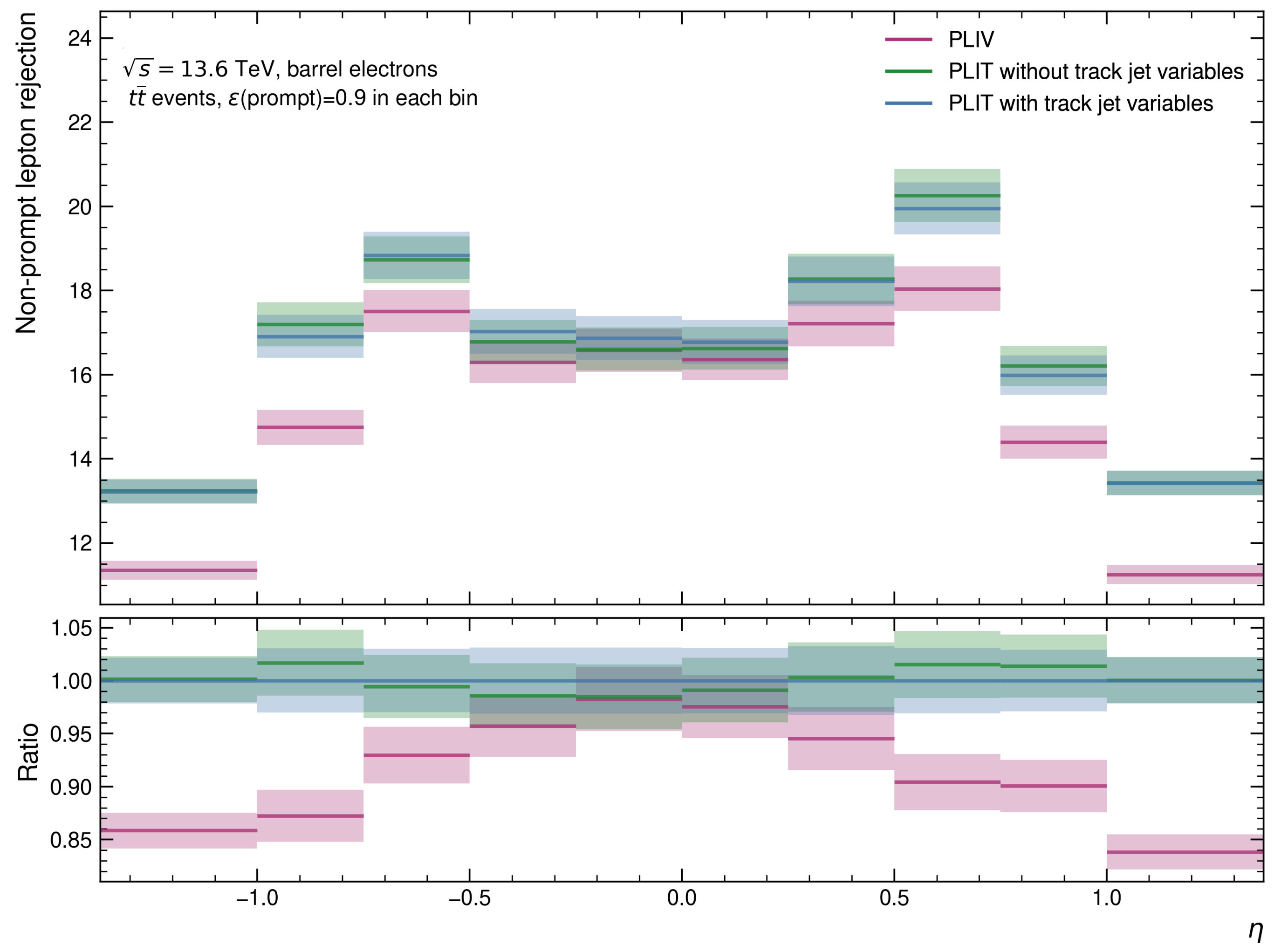


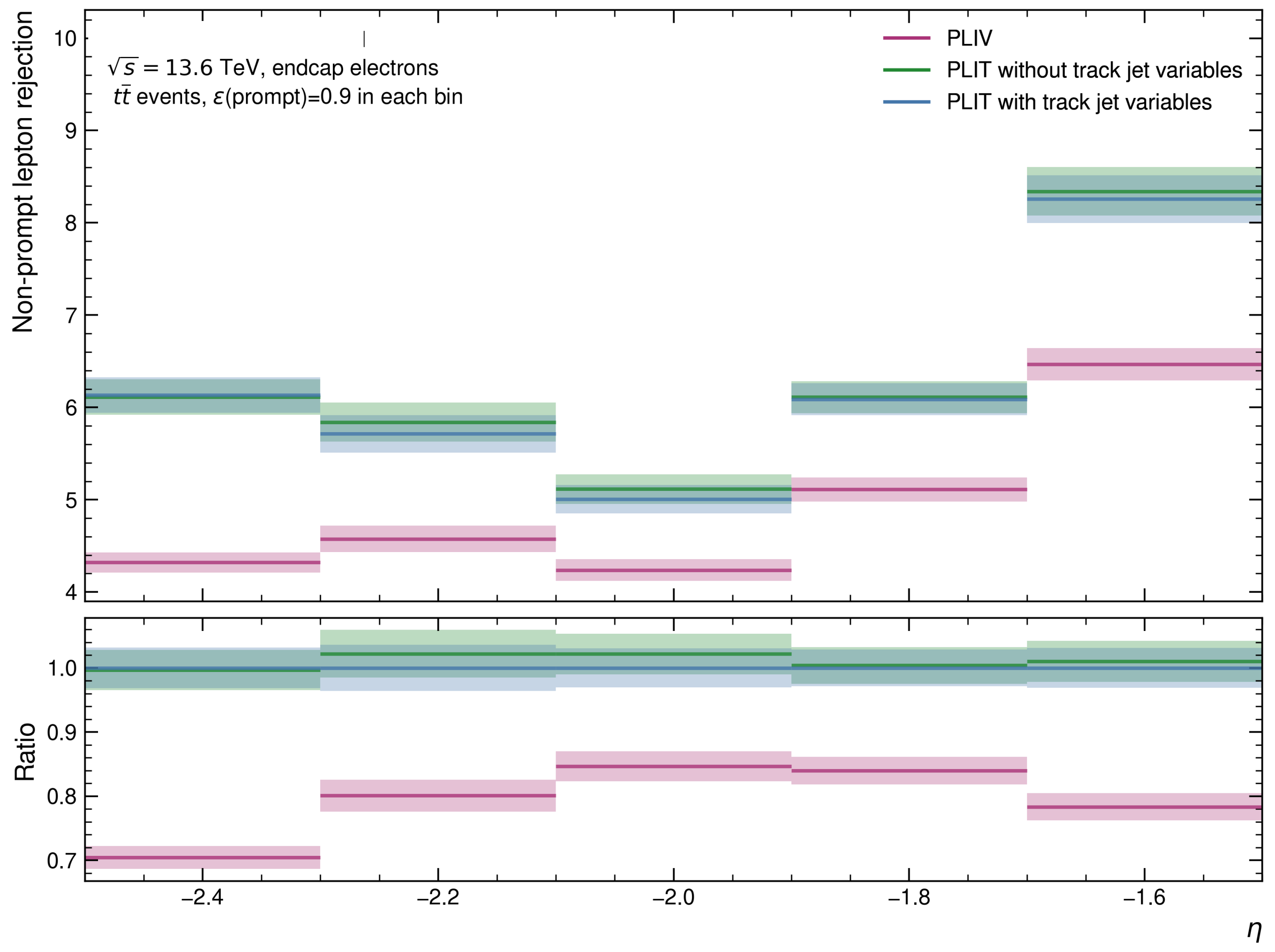
Backup

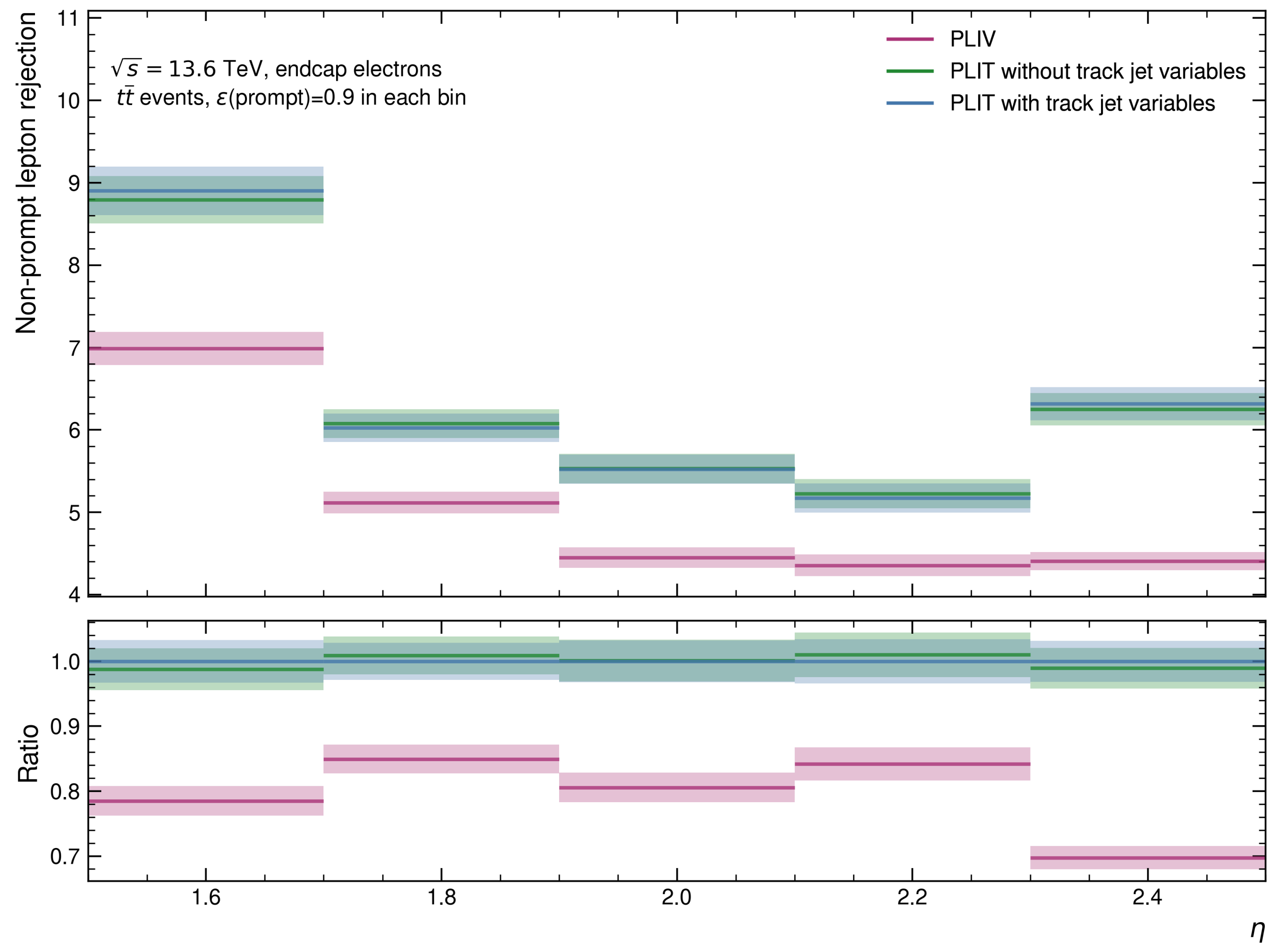


Background rejection vs η



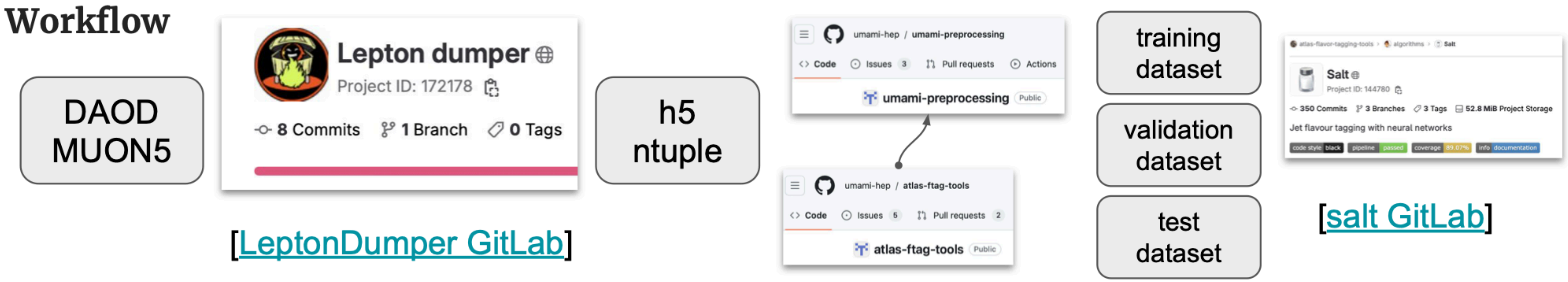






Workflow

- Inputs extracted from DAOD_MUON5 and converted in h5 files: [LeptonDumper]
- Later preprocessing with Umami and Salt: more [here]



Track selection

Object	p_T [MeV]	$\|\eta\|$	$\|d_0\|$ [mm]	$\|z_0 sin\theta\|$ [mm]	Δ R(lepton, track)	N_{Si}	N_{Si}^{shared}	N_{Si}^{hole}	N_{pix}^{hole}
track	> 500	< 2.5	< 3.5	< 5	< 0.4	≥ 8	≤ 1	≤ 2	≤ 1