

Soft Lepton Tagger

Bennett Austin

LBL+Nikhef Charm Tagging

4/27/21

Introduction

- Soft-lepton variables study
 - Soft electron and muon variables
 - Input into the DL1 algorithm
 - Improve b-jet and c-jet identification
- Data samples used:
 - WpH->cc for charm and light jets
 - WpH->bb for bottom jets
- Cuts used:
 - Jet pt > 20 GeV
 - |Jet eta| < 2.5

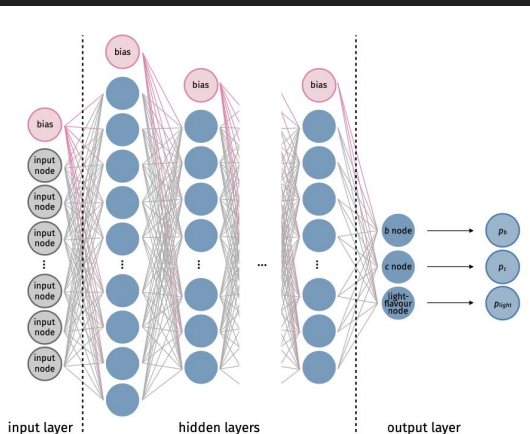
c-quark fragmentation [%]

D^+	26.39 ± 1.39
D^0	57.72 ± 2.41
D_s^+	6.91 ± 0.45
Λ_c^+	5.26 ± 0.31

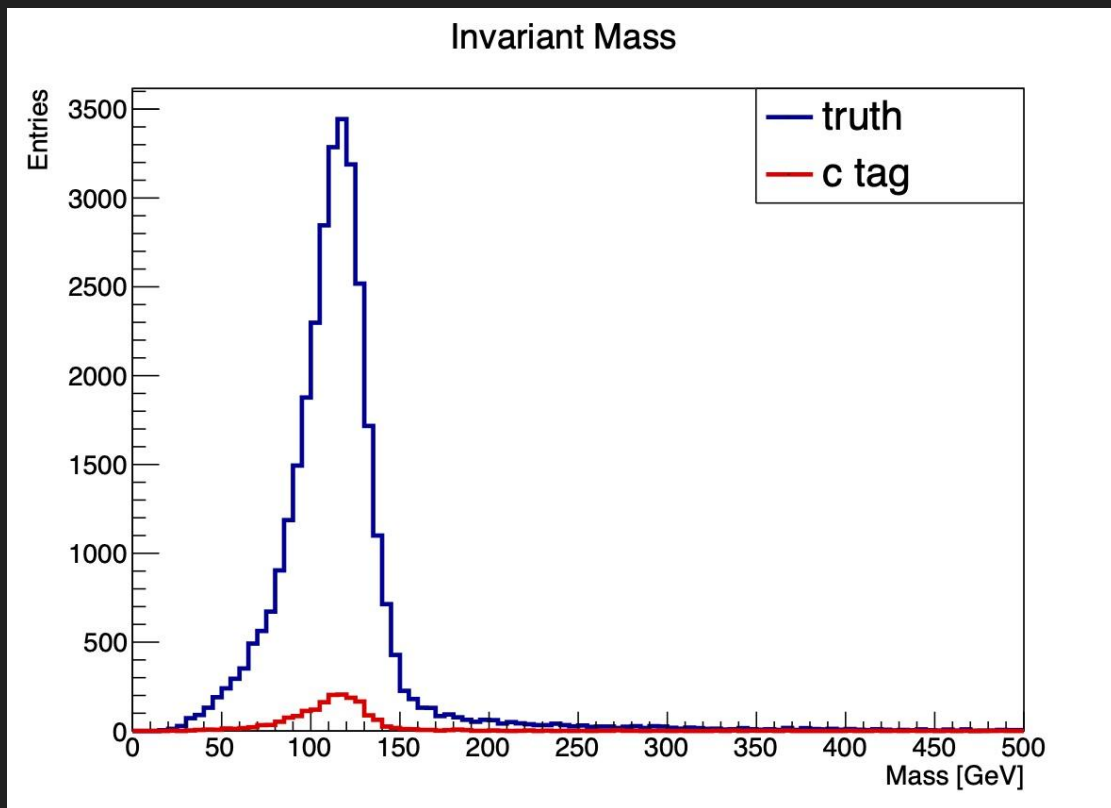
D-meson inclusive decay modes

Species	Inclusive decay mode	Fraction [%]
D^+	$\rightarrow e^+ \text{ semileptonic}$	(16.07 ± 0.30)
	$\rightarrow \mu^+ \text{ anything}$	(17.6 ± 3.2)
	$\rightarrow K^- \text{ anything}$	(25.7 ± 1.4)
	$\rightarrow \bar{K}^0 \text{ or } K^0 \text{ anything}$	(61 ± 5)
D^0	$\rightarrow \bar{K}^*(892)^0 \text{ anything}$	(23 ± 5)
	$\rightarrow e^+ \text{ semileptonic}$	(6.49 ± 0.11)
	$\rightarrow \mu^+ \text{ anything}$	(6.7 ± 0.6)
	$\rightarrow K^- \text{ anything}$	(54.7 ± 2.8)
	$\rightarrow \bar{K}^0 \text{ or } K^0 \text{ anything}$	(47 ± 4)
	$\rightarrow K^*(892)^- \text{ anything}$	(15 ± 9)

[Marko Stamenkovic](#)



Introduction

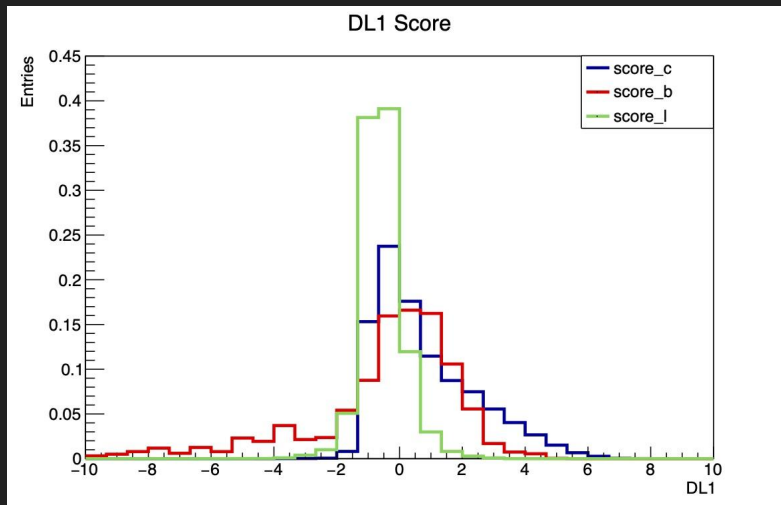


- Higgs→cc invariant mass
- Highest jet pt pairs
- Blue: truth c-jets
- Red: truth c-jets correctly tagged as c-jets by DL1
- c-tag efficiency is low

DL1 Score and efficiencies

- c-jet efficiency: 19.8%
- b-jet c-tagged efficiency: 6.4%
- Light-jet c-tagged efficiency: 0.8%
- p_c , p_u , p_b - probability of jets from training
- f is fraction of b-jets in background
 - $f = 0.08$
- Requirements for c-tag:
 - $DL1 \geq 1.3$
 - $MV2c10 \leq 0.83$

$$DL_1 = \ln \frac{p_c}{f p_b + (1 - f) p_u}$$

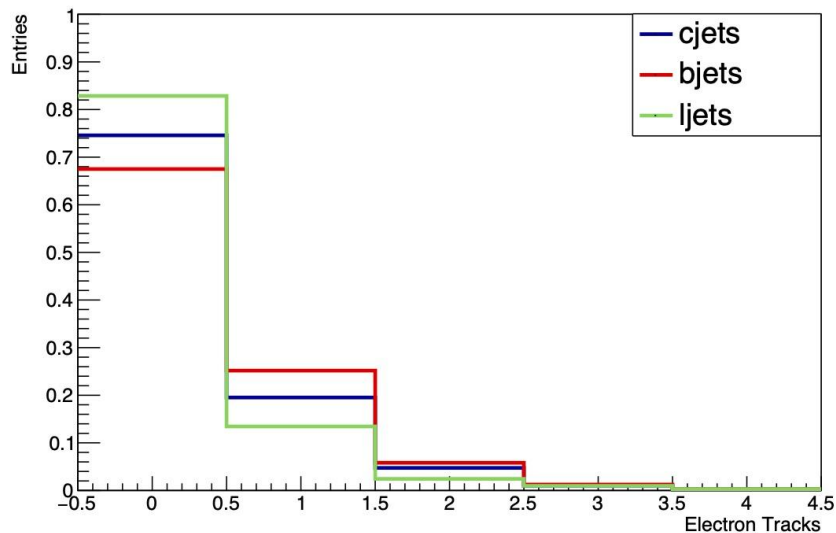


Green: l-jets, red: b-jets, blue: c-jets.
Normalized with sum of weights

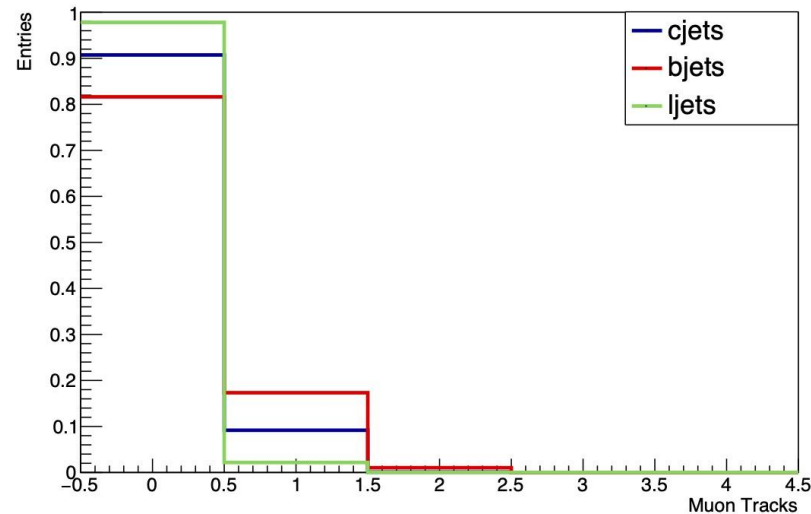
Lepton Variables

Lepton tracks per jet

Electron Tracks Per Jet



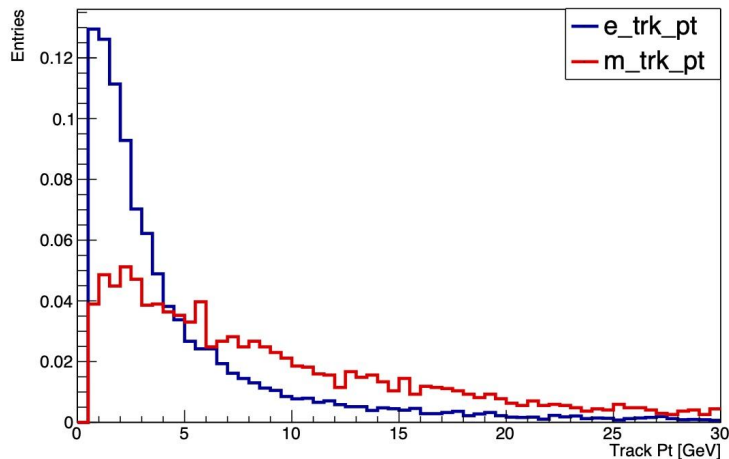
Muon Tracks Per Jet



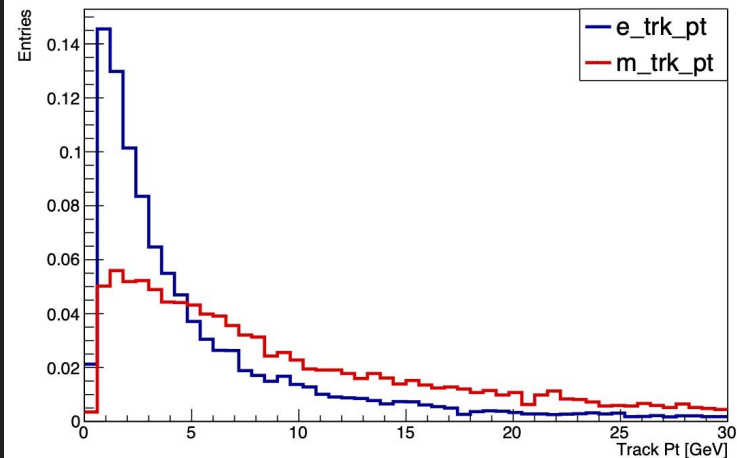
- Truth leptons matched to a track
- More electrons than muons

Lepton Track Pt

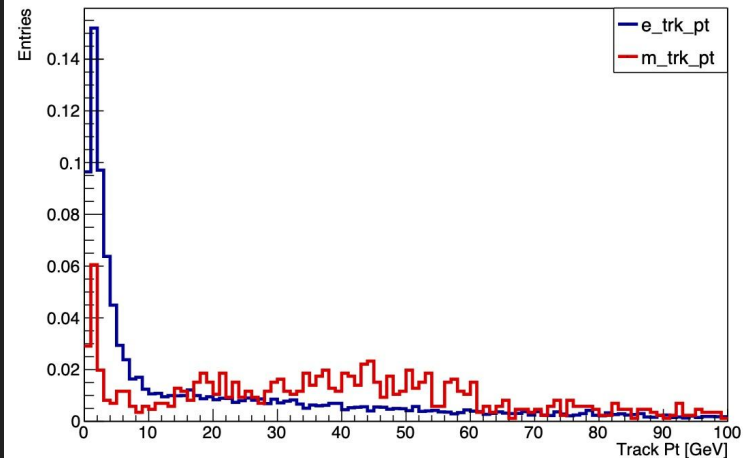
c-jet Electron and Muon Track Pt



b-jet Electron and Muon Track Pt

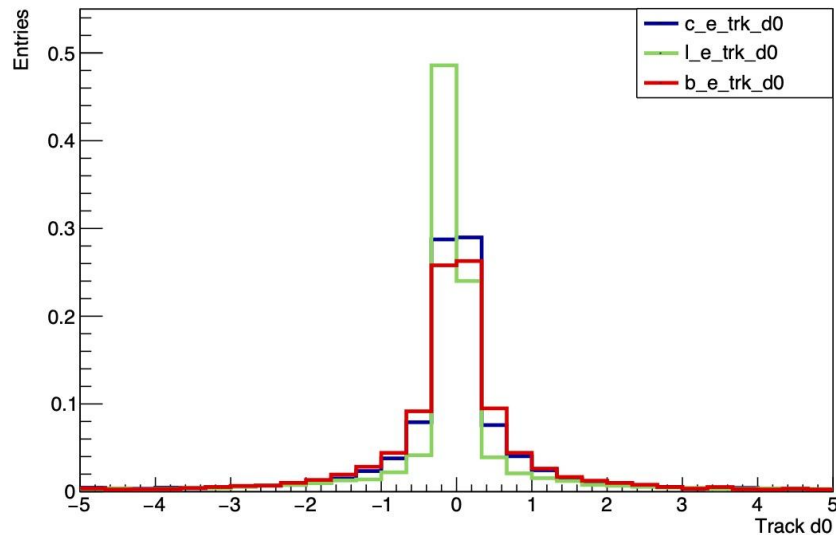


light-jet Electron and Muon Track Pt

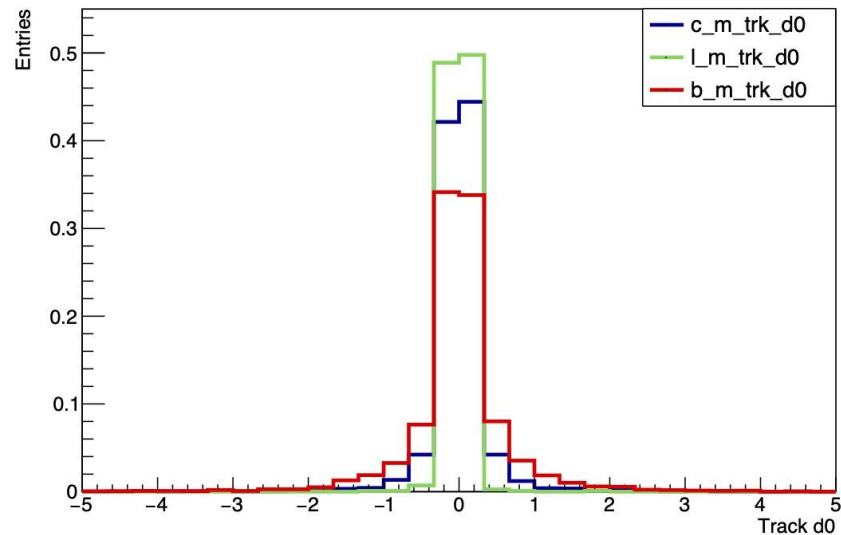


Lepton track d0

Electron Track d0

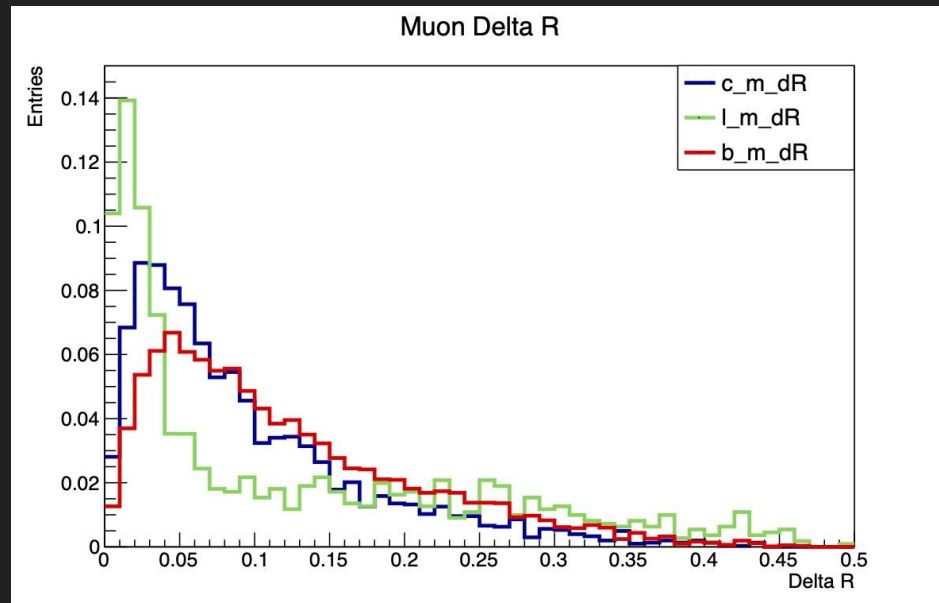
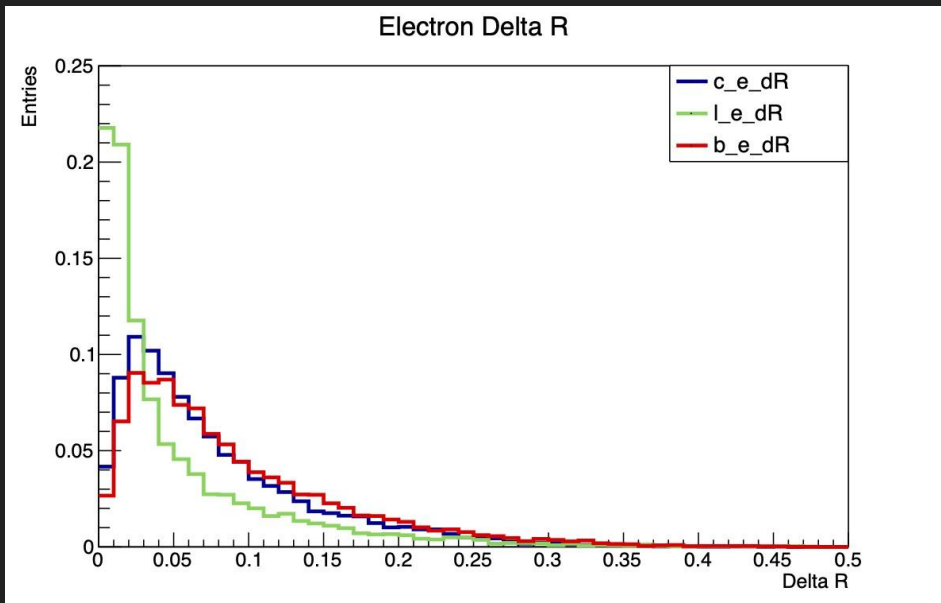


Muon Track d0



- Track d0: transverse impact parameter;
distance of closest approach of the track to
the primary vertex in the transverse direction

Lepton dR

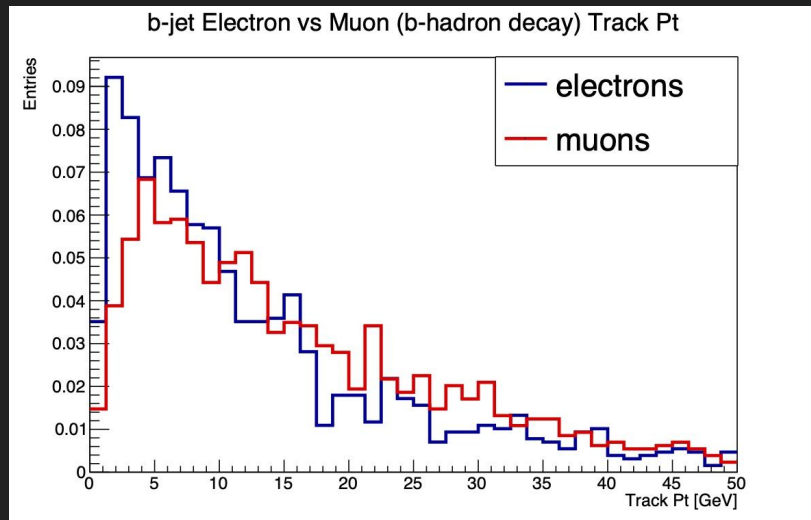


- Delta R: delta R between the lepton track and the jet axis

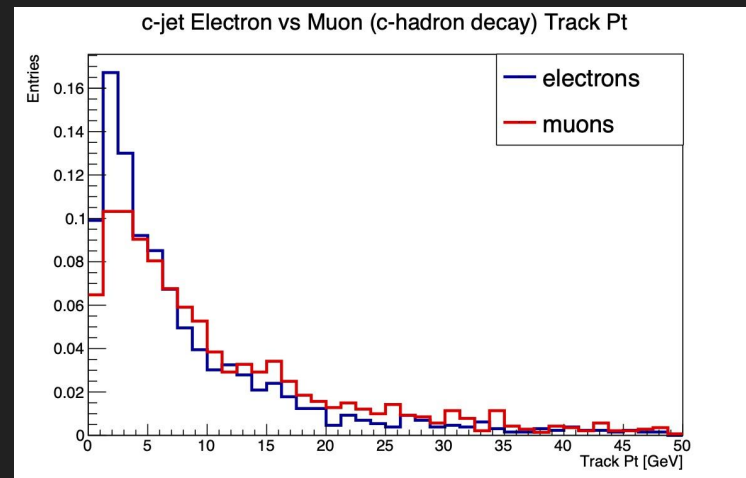
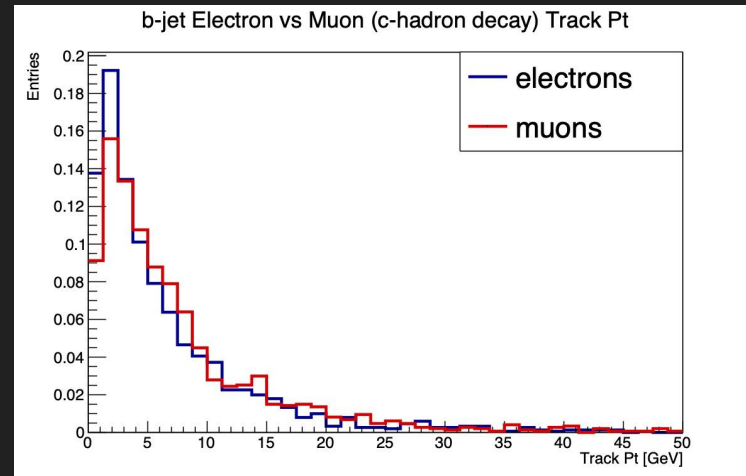
Heavy Flavor Decays

The following variables were plotted by selecting lepton tracks coming from charm and bottom mesons and baryons

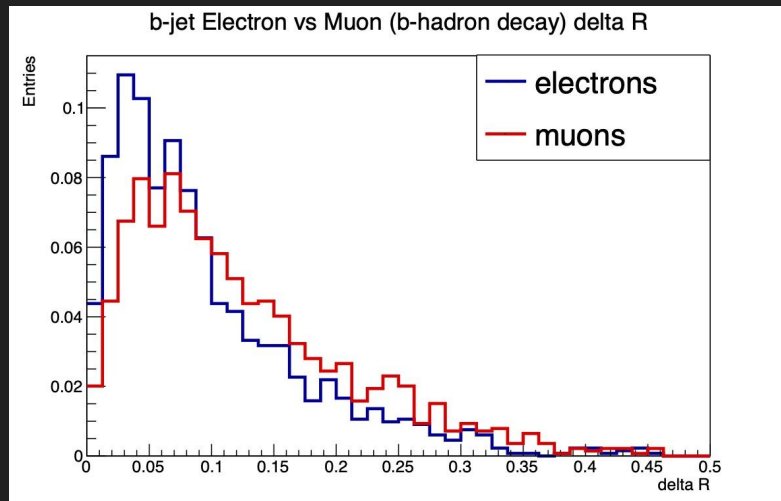
Lepton Track Pt



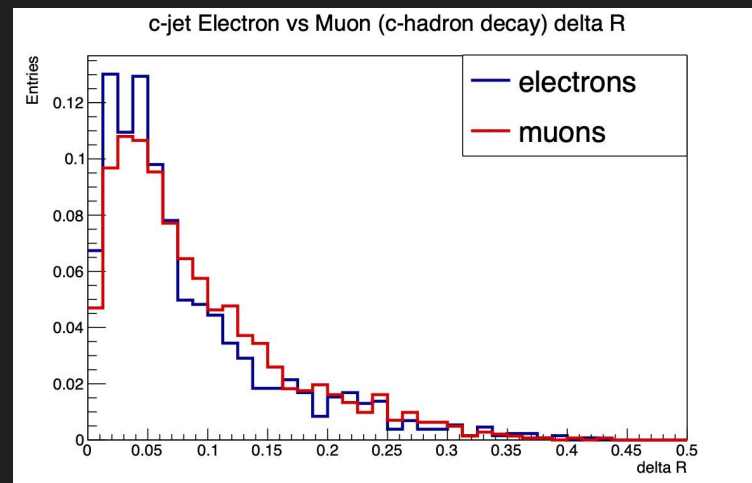
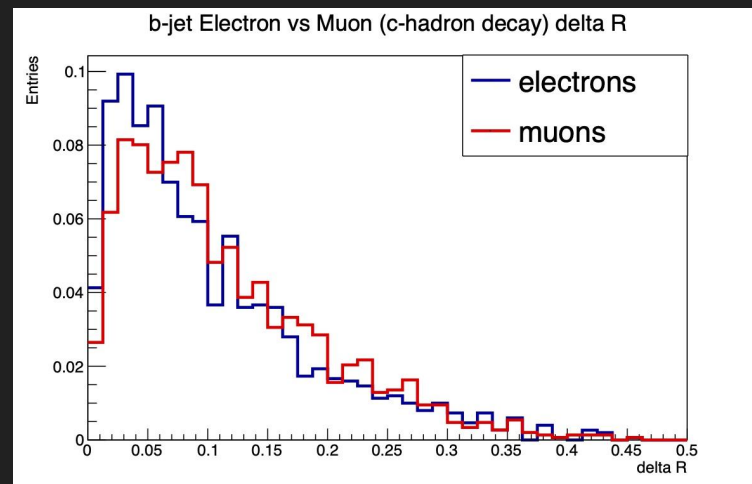
- The b-hadron decay seems to have larger track pt for both types of leptons.
- c-hadron decay is roughly the same for both types of jets.



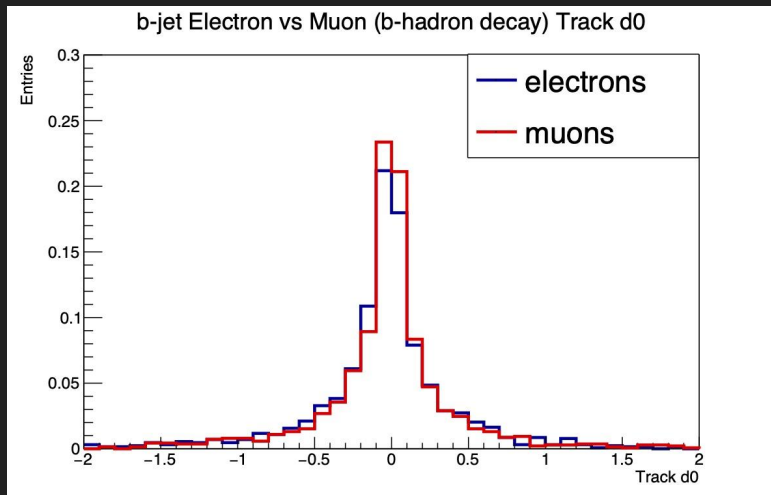
Lepton dR



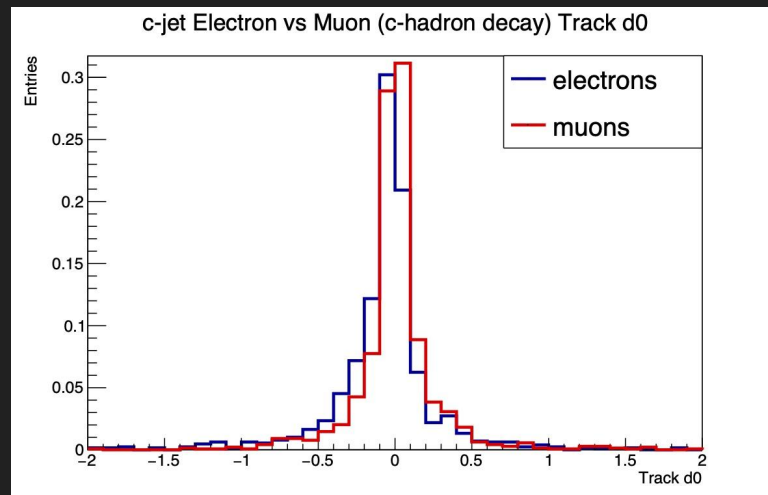
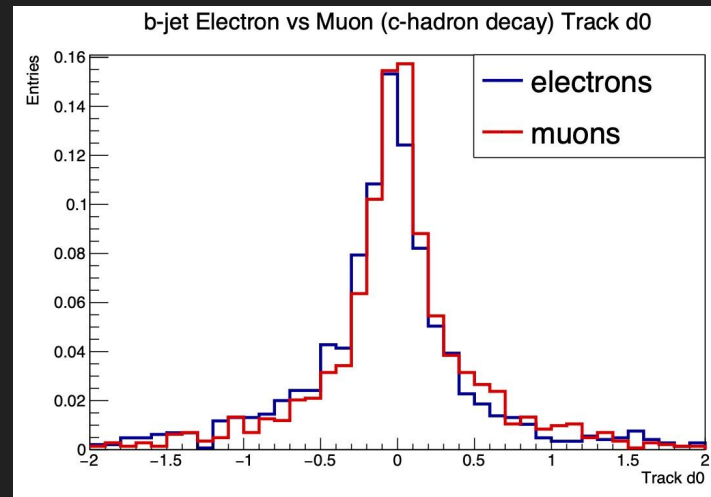
- Delta R: delta R between the lepton track and the jet axis



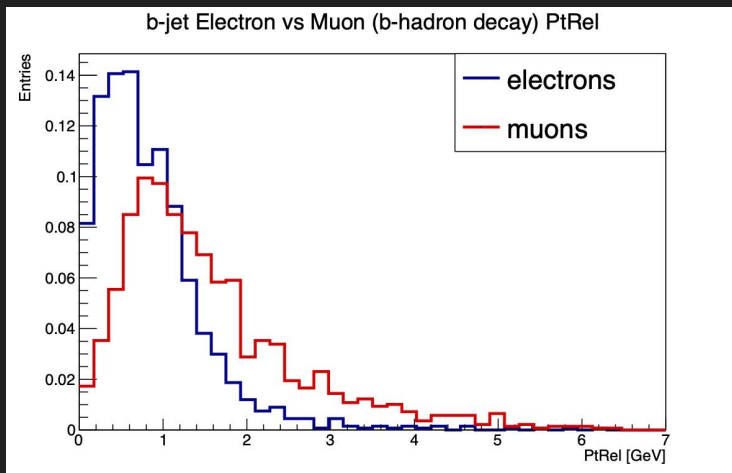
Lepton Track d0



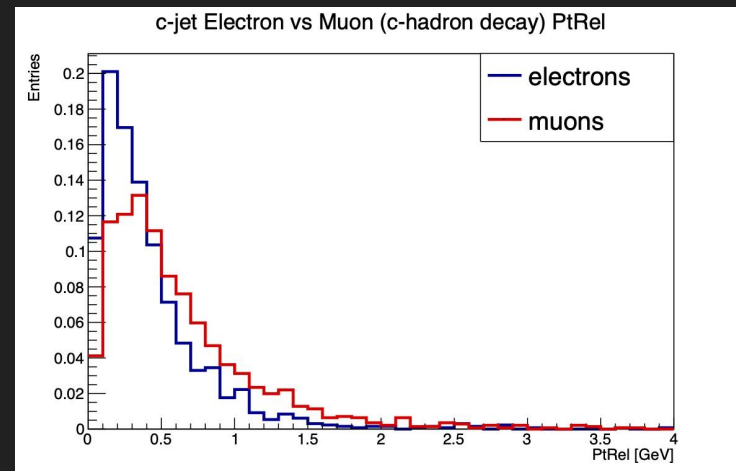
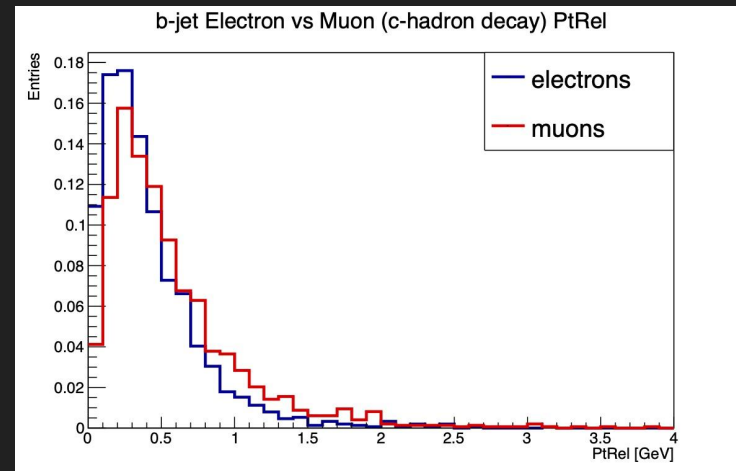
- b-jets have a slightly larger spread in track d0, especially when comparing the c-hadron decays



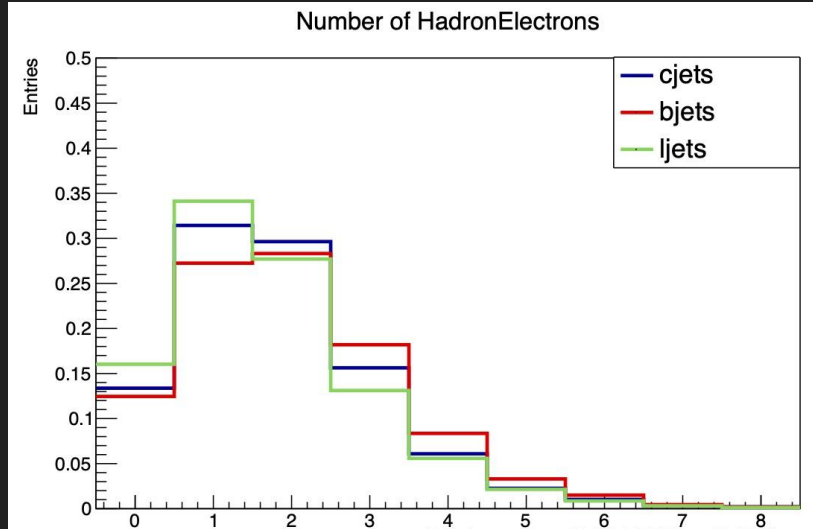
Lepton Track PtRel



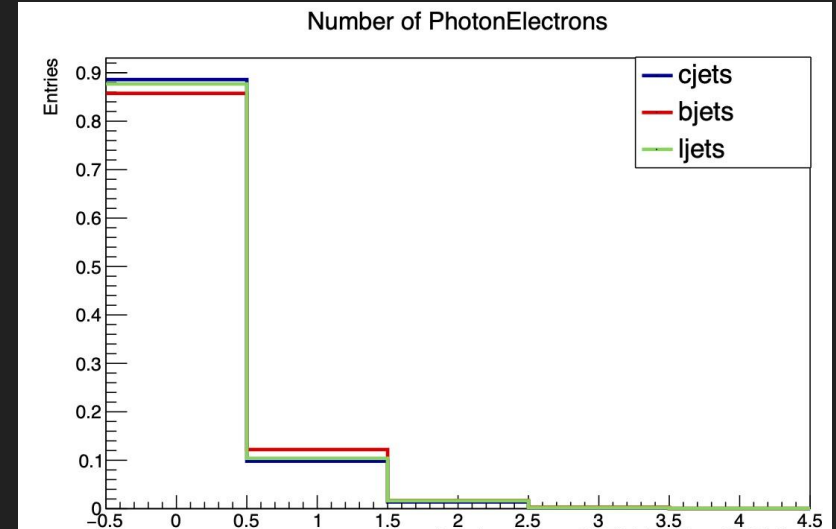
- PtRel: track pt relative to jet axis;
track momentum perpendicular to jet
axis



Electron identification



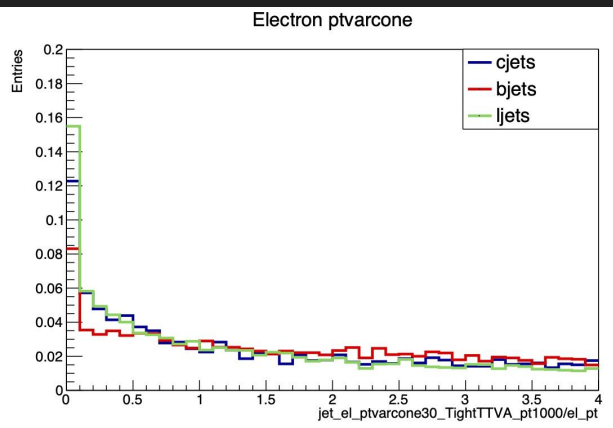
- Hadrons identified as electrons
- Lots of fake hadrons



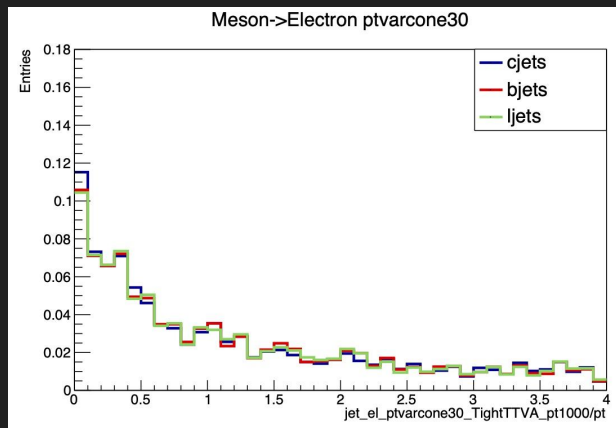
- Photons identified as electrons

Ptvarcone30

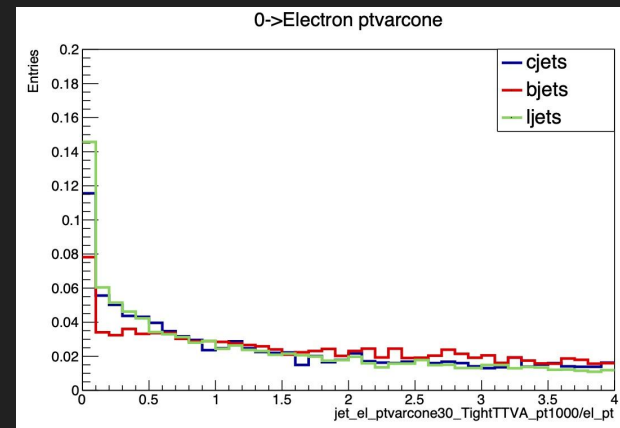
- PtvarconeXX: computed by summing the transverse momentum of the selected tracks within a cone centered around the electron track
 - XX is the size of the cone: $dR = XX/100$



All electrons (including fakes)



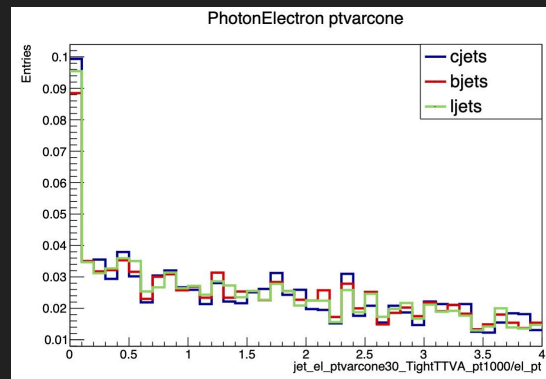
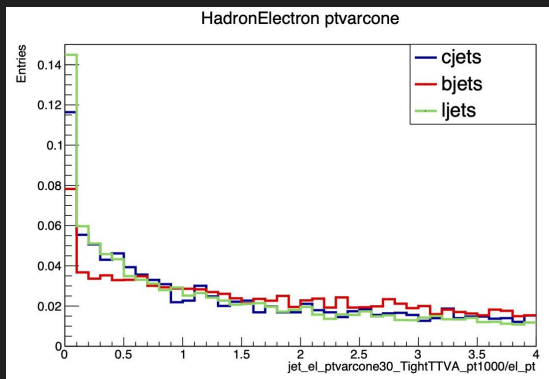
Electrons with meson origin



Electrons with fake electron origin

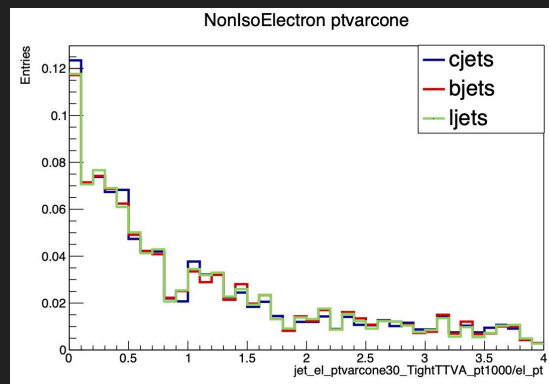
Ptvarcone30

Hadrons
identified as
electrons



Photons
identified as
electrons

Truth matched
electrons

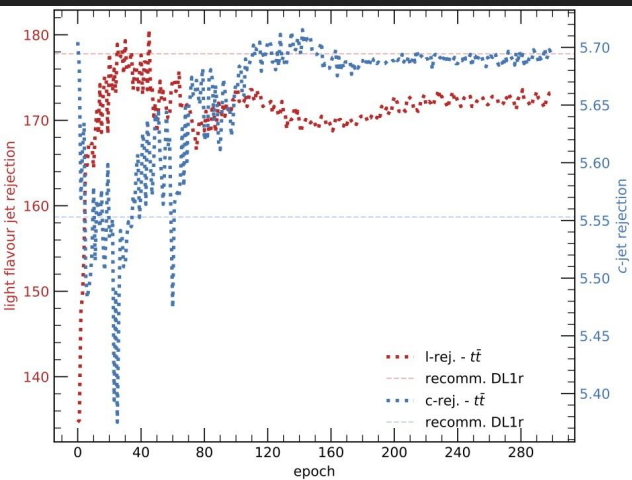


DL1 Trainings

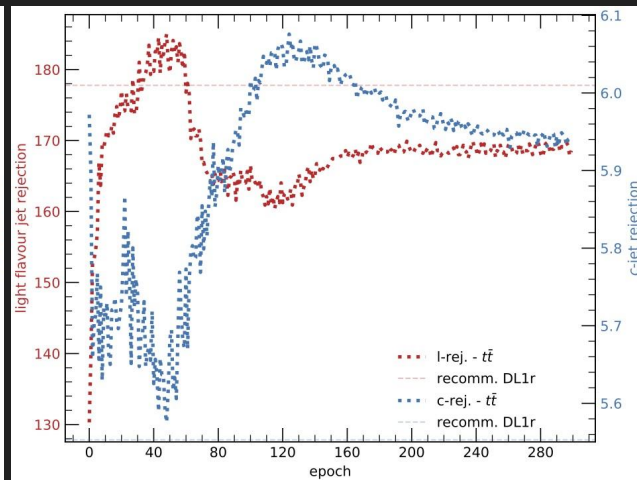
Preliminary DL1 Trainings

- Default configuration: standard training variables (JF, SV1, etc.)
- Default+Electron configuration: added soft electron pt, pTrel, dR, eta, phi, d0, and z0
- Default+Muon configuration: added soft muon pt, pTrel, dR, eta, phi, IP3D d0, and IP3D z0.
- Default and default+electron configurations:
 - Ran training for 300 epochs
 - Learning rate: 0.01
 - ROC Model evaluation epoch: 280
- Default+muon configuration
 - Ran training for 270 epochs
 - Learning rate: 0.005
 - ROC Model evaluation epoch: 250
- DNN Parameters used in all configurations:
 - Batch size: 15,000
 - Activation functions: relu
 - Layers and nodes: [256, 128, 60, 48, 36, 24, 12, 6]
 - 5,000,000 b-jets

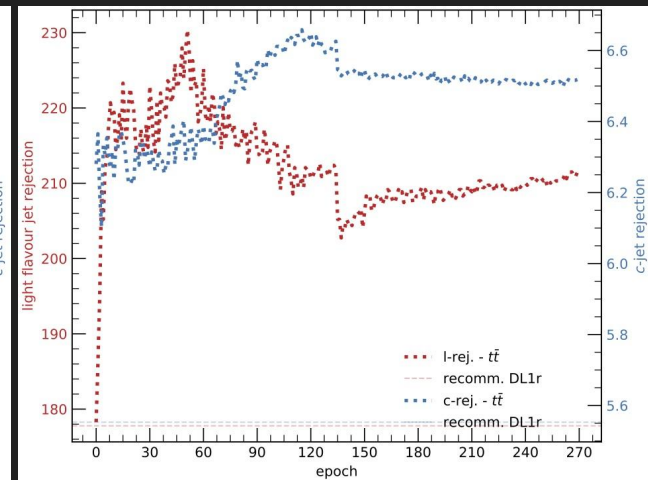
Rejection (ttbar)



Default

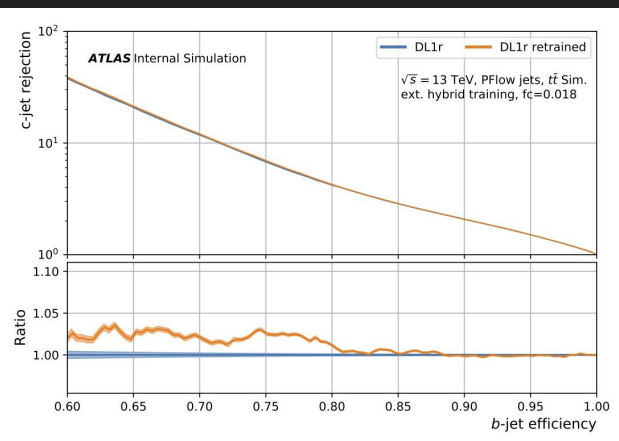


Default+electron

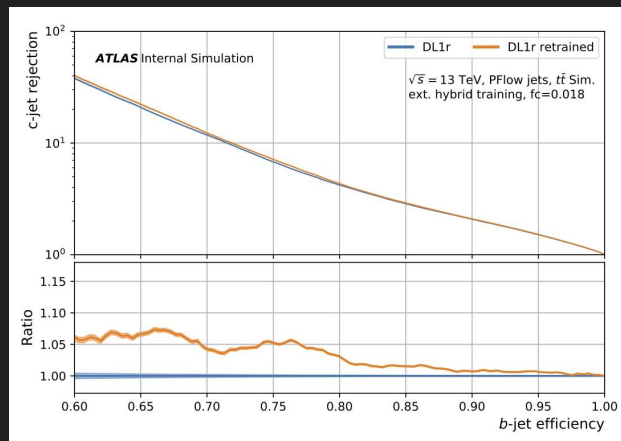


Default+muon

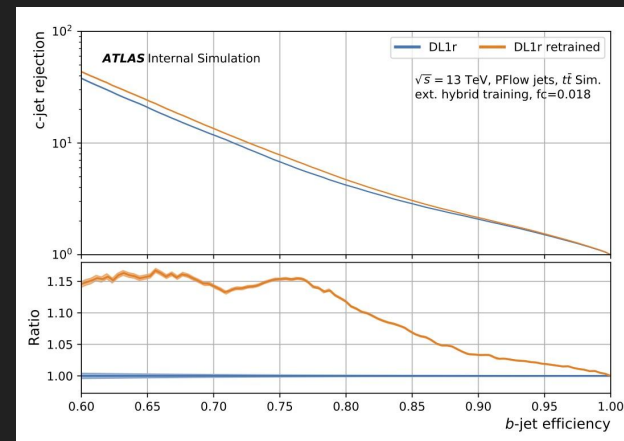
c-rej vs b-eff



Default

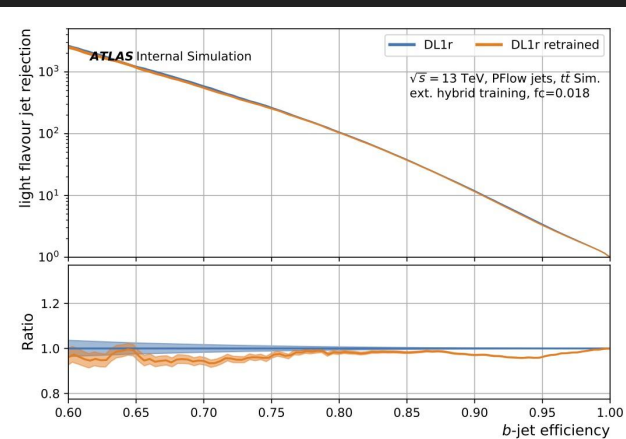


Default+electron

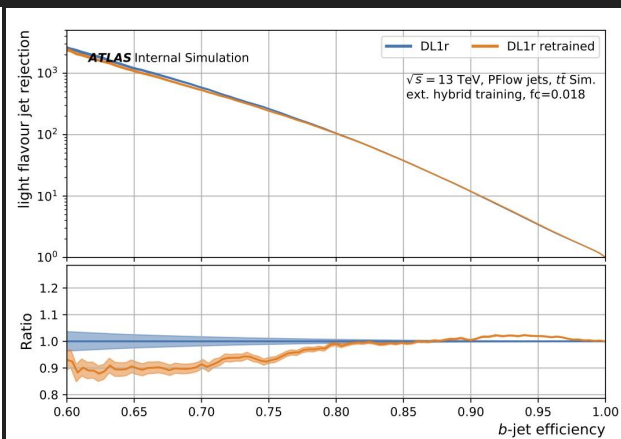


Default+muon

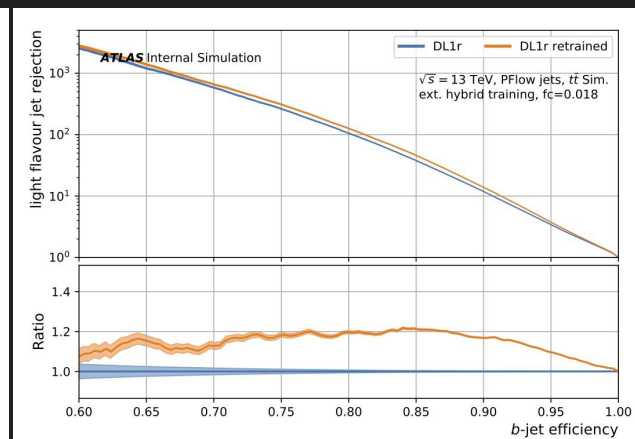
u-rej vs b-eff



Default



Default+electron



Default+muon

Further Trainings for the Muon Configuration

- Default+Muon configuration: added soft muon pt, pTrel, dR, eta, phi, IP3D d0, and IP3D z0.
- The next set of slides show the training results for the muon setting with various hyperparameter changes.
- Parameters that stayed constant:
 - Batch size = 15,000
 - Activation functions: relu
 - 5,000,000 b-jets

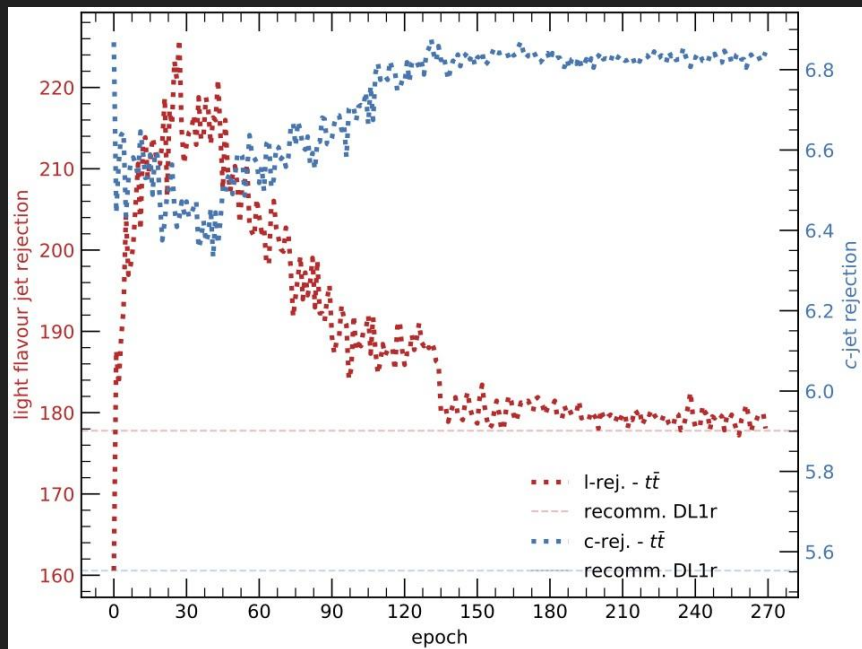
Hyperparameter Investigation

Learning rate	Layers and nodes	c-rej	u-rej	accuracy	loss
0.01	[512, 256, 128, 60, 48, 36, 24, 12, 6]	6.8	180	0.705	0.64
0.005	[256, 128, 60, 48, 36, 24, 12, 6]	6.5	210	0.70	0.65
0.005	[512, 256, 128, 60, 48, 36, 24, 12, 6]	6.7	185	0.705	0.64
0.1	[512, 256, 128, 60, 48, 36, 24, 12, 6]	6.75	205	0.695	0.65
0.0025	[256, 128, 60, 48, 36, 24, 12, 6]	6.6	210	0.695	0.64

Hyperparameter Investigation (cont.)

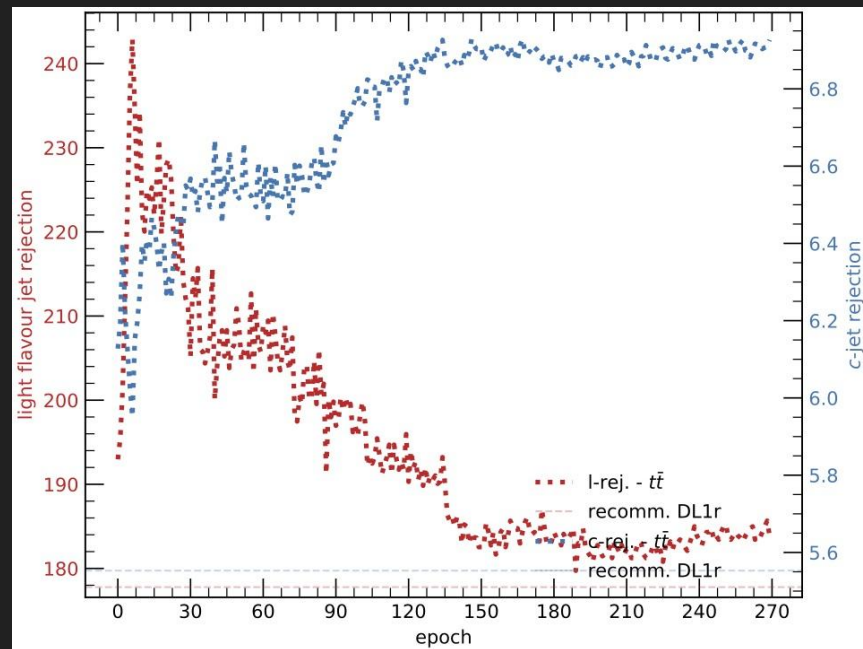
Learning rate	Layers and nodes	c-rej	u-rej	accuracy	loss
0.01	[256, 60, 48, 36, 24, 12, 6]	6.55	225	0.70	0.65
0.0025	[256, 60, 48, 36, 24, 12, 6]	6.65	210	0.70	0.65
0.0025	[512, 256, 128, 60, 48, 36, 24, 12, 6]	6.9	185	0.70	0.64
0.15	[256, 128, 60, 48, 36, 24, 12, 6]	6.4	200	0.7	0.65
0.01	[256, 128, 60, 48, 36, 24, 12, 6]	6.5	190	0.7	0.65

Rejection plots (ttbar)



$lr = 0.01$

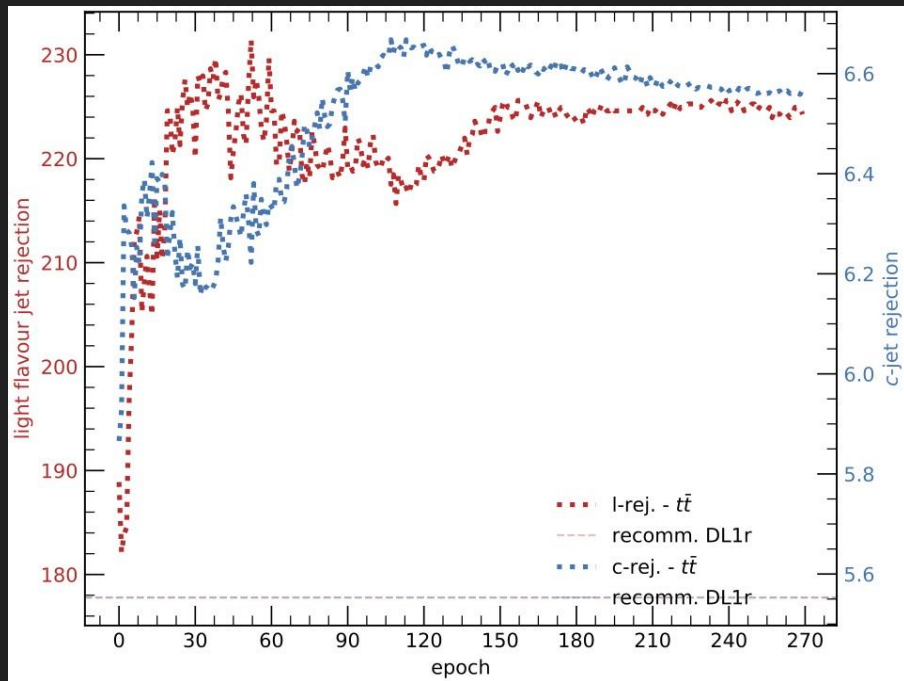
Added a layer with 512 nodes



$lr = 0.0025$

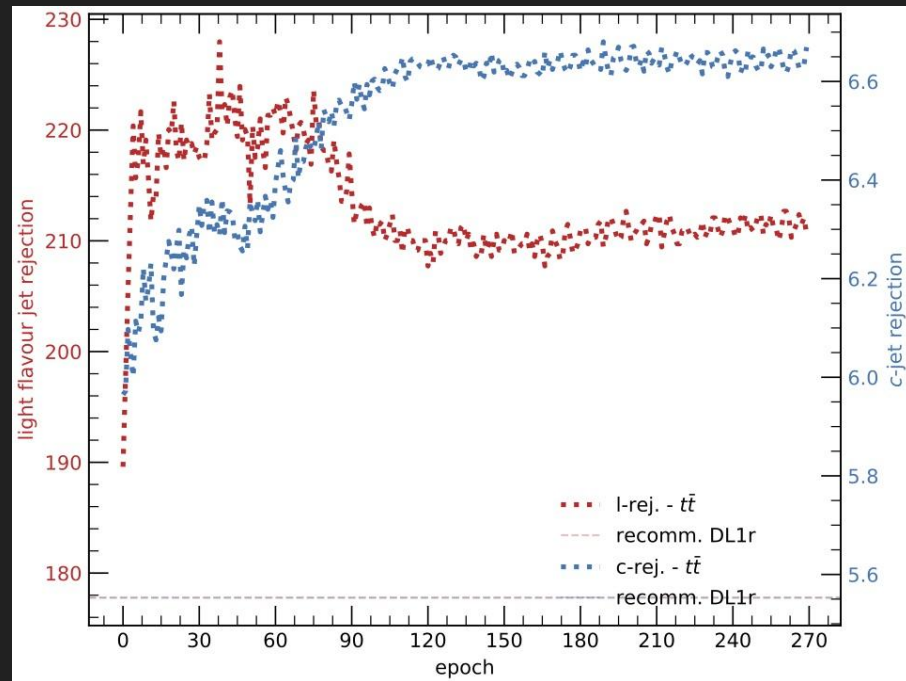
Added a layer with 512 nodes

Rejection plots (ttbar)



$lr = 0.01$

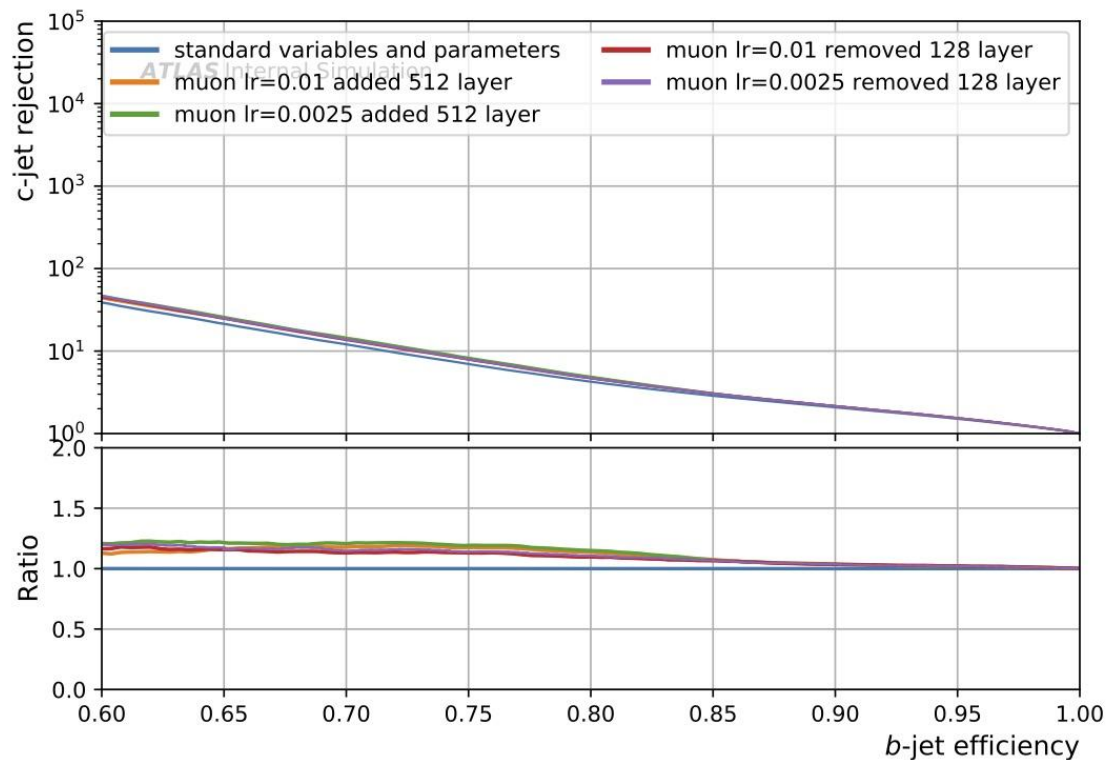
Removed a layer with 128 nodes



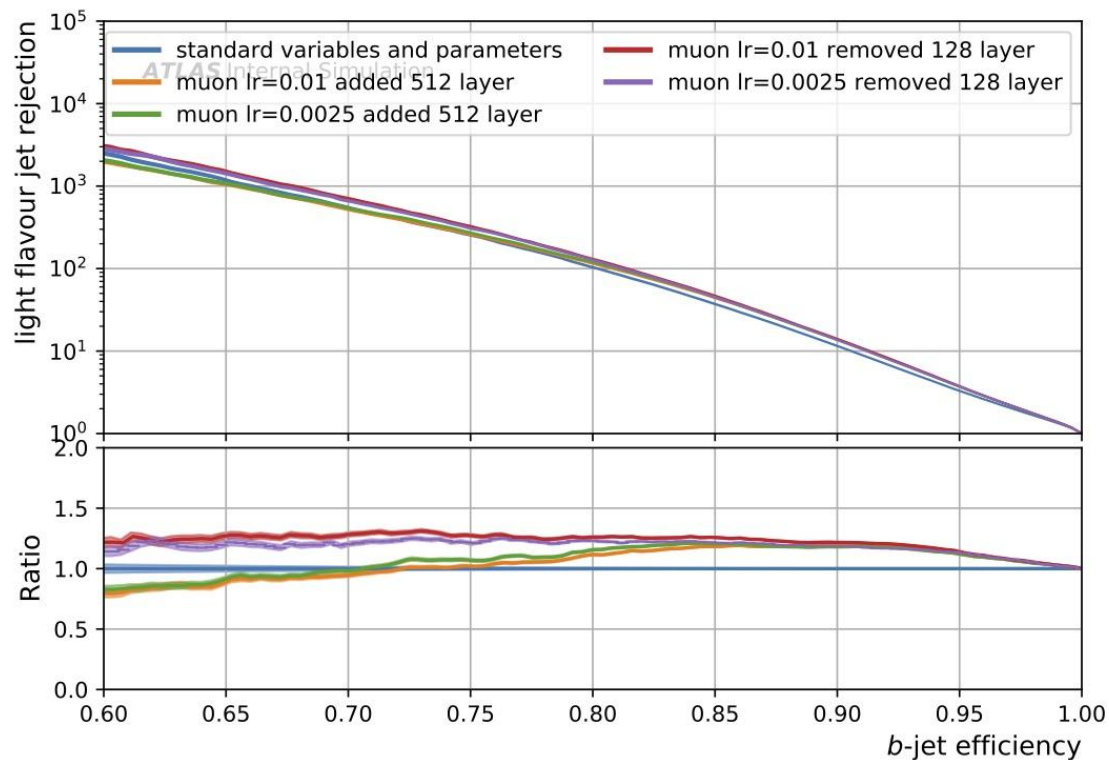
$lr = 0.0025$

Removed layer a with 128 nodes

c-rej vs. b-eff (at epoch 250)



u-rej vs. b-eff (at epoch 250)



Conclusion

- For c-rejection, muons outperformed electrons, which outperformed the default setting
- For u-rejection, muons outperformed both electrons and the default setting; electrons produced lower u-rejection than the default configuration
- Trade-off between c-rej and u-rej
 - Largest c-rej. from muons: 6.9
 - Largest u-rej. from muons: 225
 - These were achieved with different hyperparameter settings
- Electron identification
 - Lots of hadrons (pions, kaons, etc.) and some photons identified as electrons
 - Need more work for electron identification in jets
 - Affects performance of the soft electron tagger in DL1
 - Still need further training for electron variables with hyperparameter optimization