Prompt lepton tagger with the ATLAS detector

Fudong He

University of **S**cience and **T**echnology of **C**hina

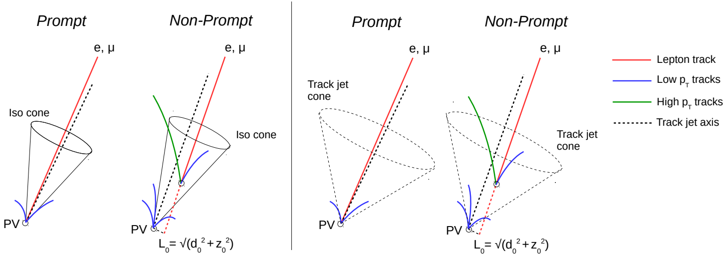
October 24, 2019

OUTLINE

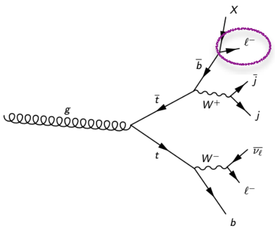
- **Introduction of the prompt lepton tagger algorithm.**
- **Newly developed calorimeter isolation variable**
- **Develop dedicated lifetime observables**
- **Training procedures and inputs**
- **Current results for electrons and muons**
- **Conclusion**

Introduction

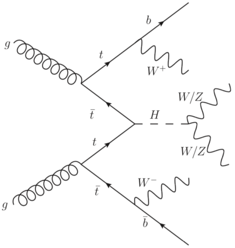
- *Prompt Lepton Tagger aims to identify prompt electrons/muons from W/Z decays:*
  - Non-prompt leptons produced in B/C decays are significant background for many analyses
  - Use isolation and lifetime observables to veto non-prompt leptons
- Current status:
  - The *PromptLeptonVeto* BDT is the previous version of the *Prompt Lepton Tagger that we developed last year.*
    - *Gives* $\sim 70\%$ *higher rejection of non-prompt leptons compared to ATLAS official isolation working point (FixedCutTight).*
    - *Used and increase the sensitivity in the* ttH, VVV *and* ttW *analysis.*



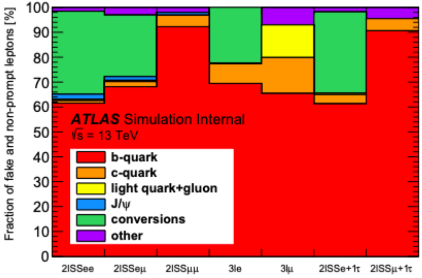
Introduction



$t\bar{t}$ contribute to same sign 2 lepton channel



ttH Feynman diagram

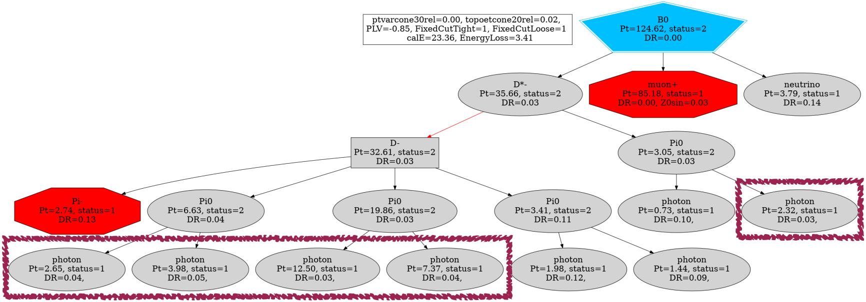


- The non-prompt background is the limiting factor in an early observation of ttH.
- The isolation cut significantly reduces the background from $t\bar{t}$, but the cross section of signal process like ttH is small enough.

Newly developed calorimeter isolation variable

Non-prompt muon event display

- **Non-prompt muon below passes standard isolation requirement**
  - B^0 decays to muon, ν_μ and D^- within the small angle to the muon ($\Delta R = 0.03$)
  - D^- further decays to a π^- and several π^0s which decay to photons
  - *Results in large calorimeter energy deposit close to the reconstructed muon*

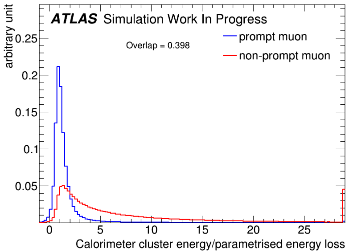


- Standard isolation tools do not include the core energy around the lepton.

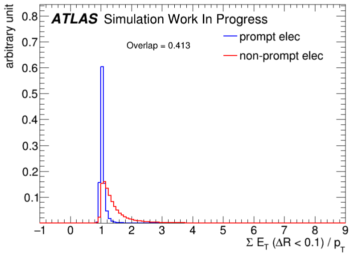
Newly developed isolation variables

- For muon, $E^{\mu}_{\text{cluster}}/E_{expected}$ = muon cluster energy/ Parameterised energy loss.
  - E^{μ}_{cluster}: collect all the energy of the clusters associated to the muon trajectory.
  - $E_{expected}$: refers to the estimated energy loss of the muon by reconstruction.
  - **The boosted b-jet will contribute to the non-prompt muon cluster energy.**
- For electron, we use $\Sigma E^{e}_{T,cluster}(\Delta R < 0.15)/p^{e}_{T}$
  - Electron will deposit all its energy does not have the "energy loss".

$E^{\mu}_{\text{cluster}}/E_{expected}$ – muon



$\Sigma E^{e}_{T,cluster}(\Delta R < 0.15)/p^{e}_{T}$ – electron

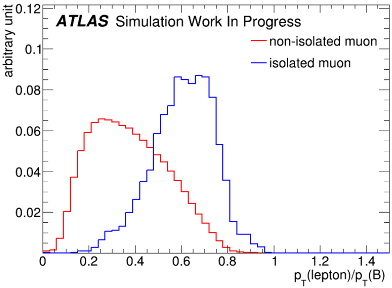


Develop dedicated lifetime observables

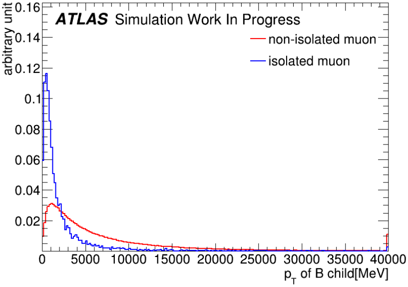
Properties of B semi-leptonic decays on MC

- Inclusive reconstructed non-prompt muons with $p_{\mathrm{T}} > 10$ GeV
- Isolated non-prompt muons: in addition passed isolation requirement.

p_T^{μ}/p_T^{B}



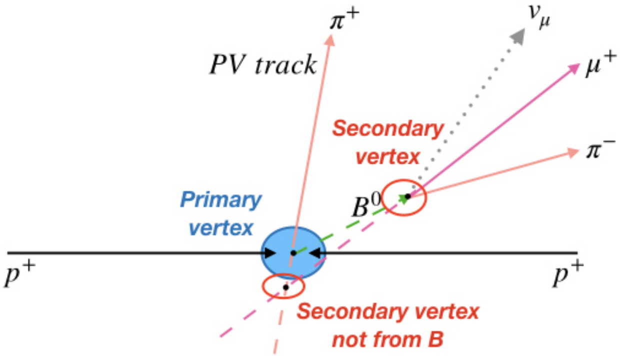
p_T of additional charged hadrons produced in B decay



- Properties of **isolated non-prompt** muons:
  - Muon carries the majority of B hadron p_T
  - Charged hadrons (that are produced in B decay) have low p_T
- Collect looser tracks with p_T > 500 MeV while BTag use 1 GeV typically.
  - Will include in average $\sim$ 8.43% more tracks from B decay.

Vertex fit using muon and track pair

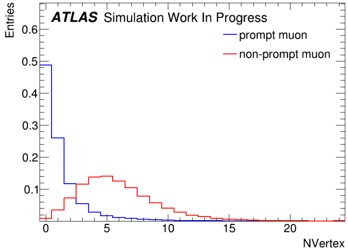
- Fit each lepton + ID track combination for 2-track vertices.
- Merge 2-track vertices that close to each other.



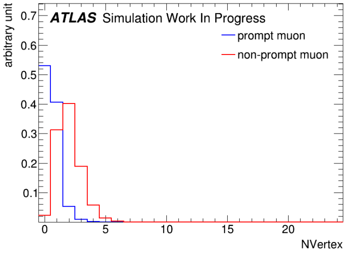
Secondary vertex multiplicity - muon

- $\sim 47\%$ of prompt muons will have reconstructed vertex. $\sim 98\%$ of non-prompt muons will have reconstructed vertex
  - Merge 2-track vertices will reduce the number of vertex for each leptons.
  - One vertex (primary vertex) reconstructed for **prompt** muon is expected.
  - For **non-prompt** muons, we will more often find both primary vertex and secondary vertex.
- **Similar observation for electrons.**
  - Vertices from material interaction are also considered.

N 2-track vertices for inclusive muons



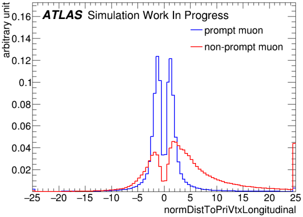
N vertices for inclusive muons – after merging



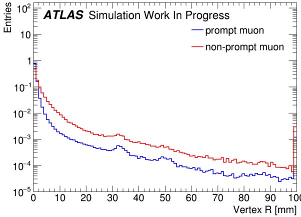
Merged vertex variables

- $l^{\text{longitudinal}}_{\text{SV to PV}}/\sigma$: the longitudinal significance distance between SV and the PV.
- Performed additional vertex selection for electron vertices: clean the SV alone the lepton track direction.

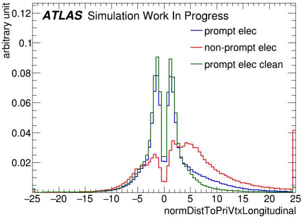
$l^{\text{longitudinal}}_{\text{SV to PV}}/\sigma$ – muon



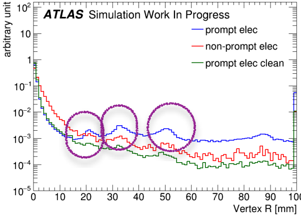
$\sqrt{x^2+y^2}$ – muon



$l^{\text{longitudinal}}_{\text{SV to PV}}/\sigma$ – electron



$\sqrt{x^2+y^2}$ – electron



Training procedures and inputs

Training procedures and inputs

- BDT^{μ}_{new}: 8 input variables; p_T bin normalization is performed; do loose isolation pre-selection on training samples.
- BDT^{e}_{new}: 10 input variables; train forward electrons and central electrons separately; p_T bin normalization is performed; do loose isolation pre-selection on training samples.
- p_T bin normalization - suppress the strong p_{T} dependence
  - p_T bins (muons): 10, 15, 20, 25, 31, 40, 10000 GeV
  - p_T bins (electrons): 10, 20, 40, 10000 GeV
  - Normalize non-prompt muon yield to prompt muon yield in each p_T bin, to remove p_{T} dependence – check_ptnorm.pdf

Inputs	Description	BDT^{μ}_{new}	BDT^{e}_{new}
RNNIP	RNN using track impact parameters and other variables[1]	✓	✓
$l^{\text{longitudinal}}_{\text{SV to PV}}/\sigma$	Secondary vertex longitudinal significance using tracks with $p_{\mathrm{T}} > 500$ MeV	✓	✓
$p_T\text{VarCone30}/p_T$	Lepton isolation using ID tracks within a cone of $\Delta R < 0.3$	✓	✓
$E_T\text{TopoCone30}/p_T$	Lepton isolation using topological clusters within a cone of $\Delta R < 0.3$	✓	✓
$E^{\mu}_{\text{cluster}}/E_{expected}$	Relative muon calorimeter cluster energy	✓	-
$\sum^{dR<0.15}_{\text{cluster}} E_T/p_T$	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: $p \cdot \sin(<\text{lepton, track jet}>)$	-	✓
$p_T^{\text{lepton track}}/p_T^{\text{track jet}}$	Lepton track p_T divided by track jet p_T	✓	✓
$\Delta R(\text{lepton, track jet})$	ΔR between the lepton and the track jet axis	✓	✓
p_T^{lepton} bin number	Index of the bin of lepton p_T	✓	✓

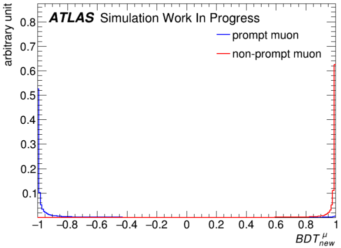
[1] will re-train it dedicated to prompt lepton tagging

Performance

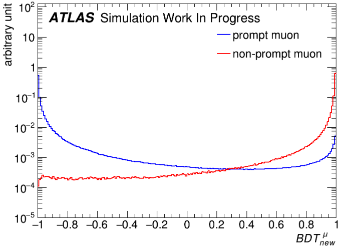
BDT score of muon

- Perform on medium reconstructed muon with p_T > 10 GeV

BDT^{μ}_{new}



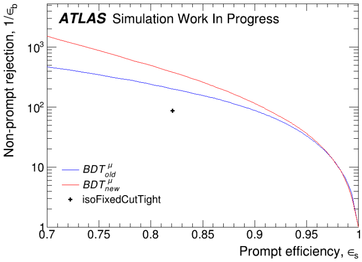
BDT^{μ}_{new} –logy



- BDT type: Gradient BDT.
- BDT configuration: NTrees=1000; NCuts=100; MaxDepth=7; MinNodeSize=0.05%;

Current performance for the muons

- Check non-prompt muon rejection for medium reconstructed muons with $p_{\mathrm{T}}^{lepton} > 10$ GeV
- *Prompt Lepton Tagger*(s) reject more non-prompt muons compared to isolation cuts
- BDT^{μ}_{new} is the recommend one now.
- Substantial improvement with updated algorithms

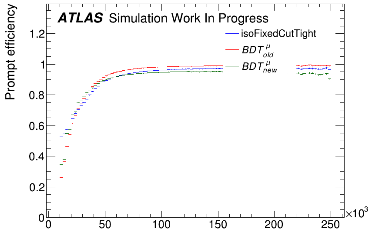


Algorithm	prompt eff. cut	prompt efficiency(%)	Bkg. rejection (1/eff.)
BDT^{μ}_{new}	-0.875	$82.14^{+0.01}_{-0.01}$	$381.6^{+5.8}_{-5.9}$
BDT^{μ}_{old}	-0.805	$82.14^{+0.01}_{-0.01}$	$201.1^{+2.2}_{-2.2}$
isoFixedCutTight	1.	$82.14^{+0.01}_{-0.01}$	$85.3^{+0.6}_{-0.6}$

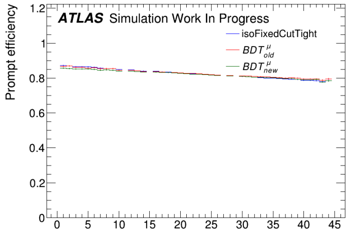
Muon efficiency

- Prompt muon efficiency at same overall efficiency of p_T > 10 GeV muons as the FixedCutTight

eff. vs. p_T – prompt

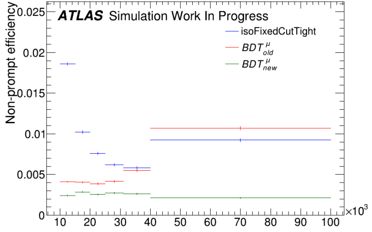


eff. vs. Average Interactions – prompt

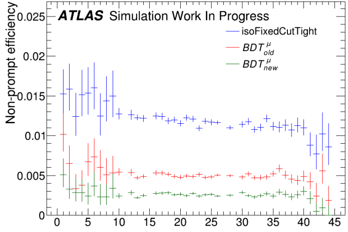


- BDT^{μ}_{new} has better non-prompt muon rejection in high p_T region.

eff. vs. p_T – non-prompt



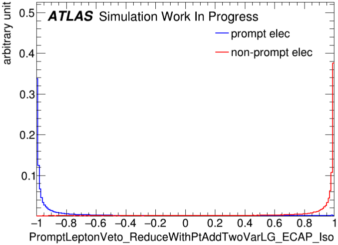
eff. vs. Average Interactions – non-prompt



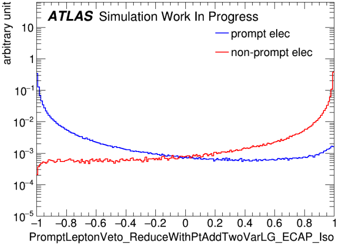
BDT score of electron

- Use the same BDT type and configuration as the muons.
- Perform on p_T > 10 GeV **forward** electron (above) and **central** electron (below).

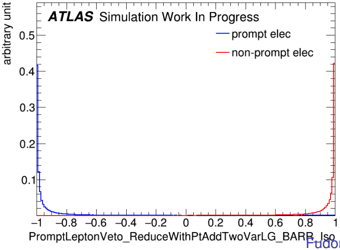
$BDT^{e}_{new,forward}$



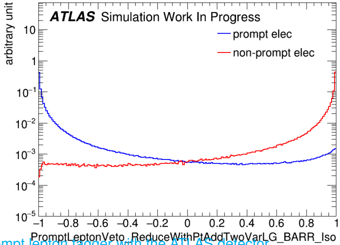
$BDT^{e}_{new,forward}$ – logy



$BDT^{e}_{new,central}$



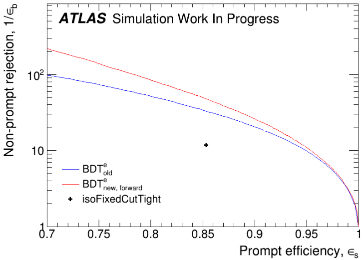
$BDT^{e}_{new,central}$ – logy



Current performance for the electrons

- Substantial improvement with updated algorithms

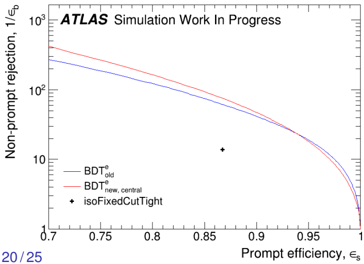
$BDT^{e}_{new,forward}$



- Perform on p_T > 10 GeV **forward** electron.

Algorithm	prompt eff. cut	prompt efficiency(%)	Bkg. rejection (1/eff.)
$BDT^{e}_{new,forward}$	-0.615	$85.37^{+0.04}_{-0.04}$	$47.8^{+0.7}_{-0.7}$
BDT^{e}_{old}	0.045	$85.37^{+0.04}_{-0.04}$	$32.9^{+0.4}_{-0.4}$
isoFixedCutTight	1.	$85.37^{+0.04}_{-0.04}$	$11.7^{+0.1}_{-0.1}$

$BDT^{e}_{new,central}$



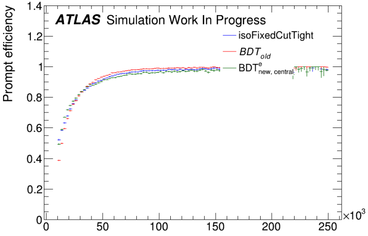
- Perform on p_T > 10 GeV **central** electron.

Algorithm	prompt eff. cut	prompt efficiency(%)	Bkg. rejection (1/eff.)
$BDT^{e}_{new,central}$	-0.705	$86.76^{+0.02}_{-0.02}$	$75.4^{+0.8}_{-0.8}$
BDT^{e}_{old}	-0.535	$86.76^{+0.02}_{-0.02}$	$61.2^{+0.6}_{-0.6}$
isoFixedCutTight	1.	$86.76^{+0.02}_{-0.02}$	$13.7^{+0.1}_{-0.1}$

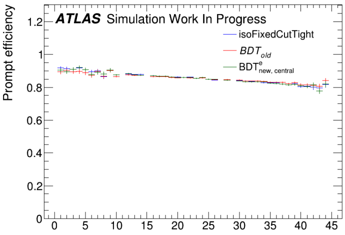
Forward electron efficiency

- Prompt forward electron efficiency at same overall efficiency of p_T > 10 GeV forward electrons as the FixedCutTight

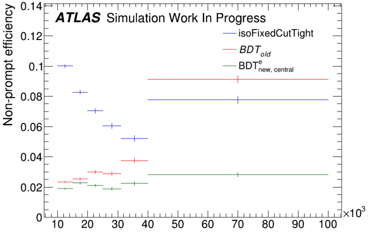
eff. vs. p_T – prompt



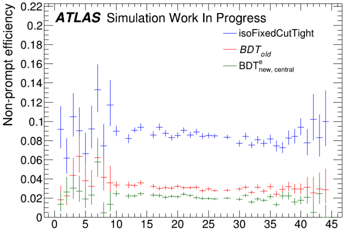
eff. vs. Average Interactions – prompt



eff. vs. p_T – non-prompt



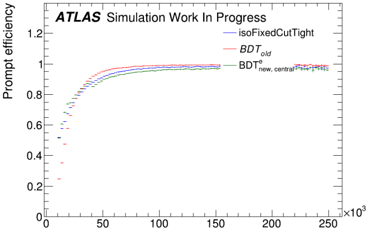
eff. vs. Average Interactions – non-prompt



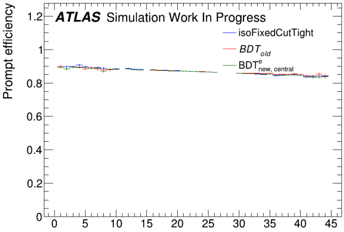
Central electron efficiency

- Prompt central electron efficiency at same overall efficiency of p_T > 10 GeV central electrons as the FixedCutTight

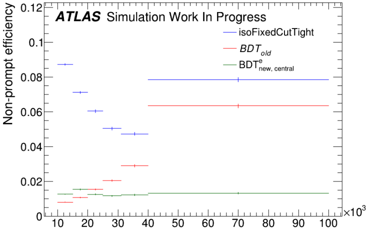
eff. vs. p_T – prompt



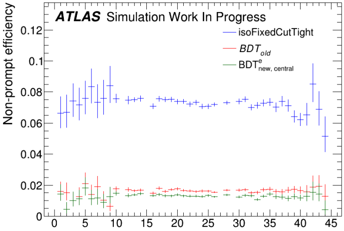
eff. vs. Average Interactions – prompt



eff. vs. p_T – non-prompt



eff. vs. Average Interactions – non-prompt



Conclusion

Conclusion

- BDTs combined isolation information and lifetime information outperform standard isolation working points in selecting prompt leptons.
- Dedicated isolation variable and lifetime variables are developed for the prompt lepton tagging.
- New BDTs with optimized training procedure have shown better performance than the current BDTs (PromptLeptonVeto).
- This study will be finalize sooner after adding new RNN variable.

Thank you for your time !

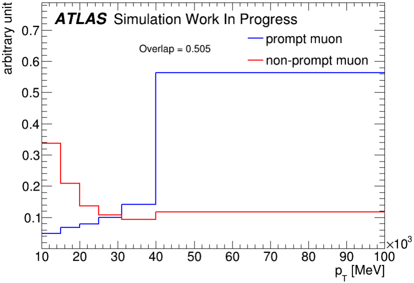
BACKUP

Soft vertex finding

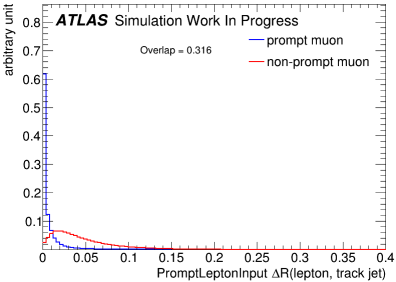
- Collect loose ID tracks within a cone of $\Delta R = 0.4$ centered at the muon
- Fit 2-track vertexes using muon track + ID track pairs
  - Vertex fitter candidates: Trk::TrkVKalVrtFitter, Trk::FastVertexFitter and Trk::SequentialVertexFitter.
  - Three fitters above show similar performance in our case.
  - Choose Trk::FastVertexFitter for our study since it has higher efficiency and straightforward theory.
- Select vertexes with fit probability at $< 3\%$
- Sort selected 2-track vertexes by track pT , then start with highest pT track
- Iteratively add each 2-track vertex, starting with closest vertex, then fit merged vertex
- Include ID tracks that fail 2-track vertex fit with lepton track - aim to reconstruct charm decay vertex.

Muon inputs

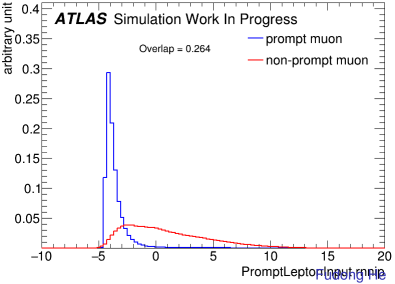
XPtBin



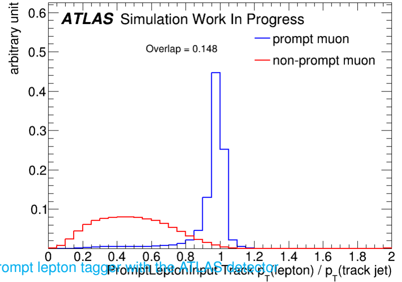
$\Delta R(\text{lepton, track jet})$



RNNIP

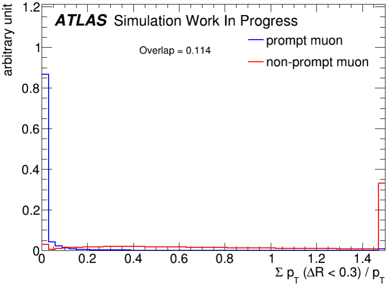


$p_T^{\text{lepton track}} / p_T^{\text{track jet}}$

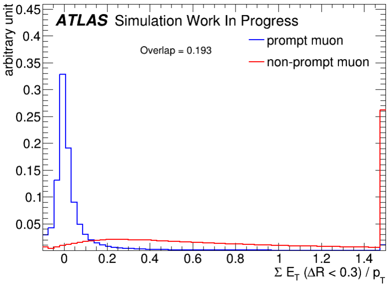


Muon inputs

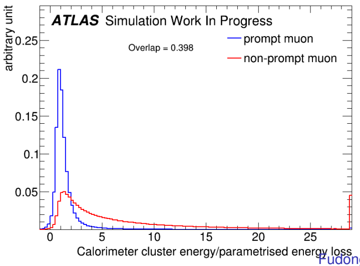
p_TVarCone30$/p_T$



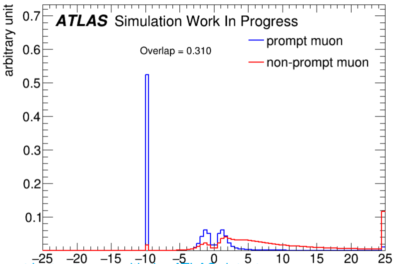
E_TTopoCone30$/p_T$



$E^{\mu}_{\text{cluster}}/E_{expected}$ – muon



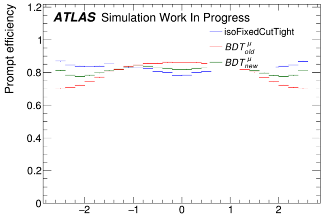
$l^{\text{longitudinal}}_{\text{SV to PV}}/\sigma$



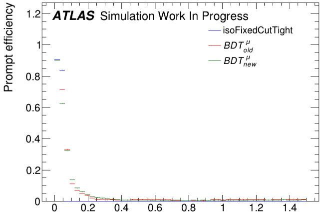
Muon efficiency

- Prompt muon efficiency at same overall efficiency of p_T > 10 GeV muons as the FixedCutTight

eff. *eta*

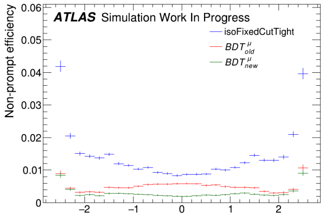


eff. P_TVarCone30$/p_T^{lepton}$



- ReduceWithPt_Iso outperforms ReduceWithPt in topoetcone30rel < 0.3 region.

eff. *eta*



eff. P_TVarCone30$/p_T^{lepton}$