Data Commissioning of Flavour Tagging in Run 2 Data

Laurie McClymont,
Flavour Tagging Group

UCL ATLAS Meeting

07/08/15

Aims

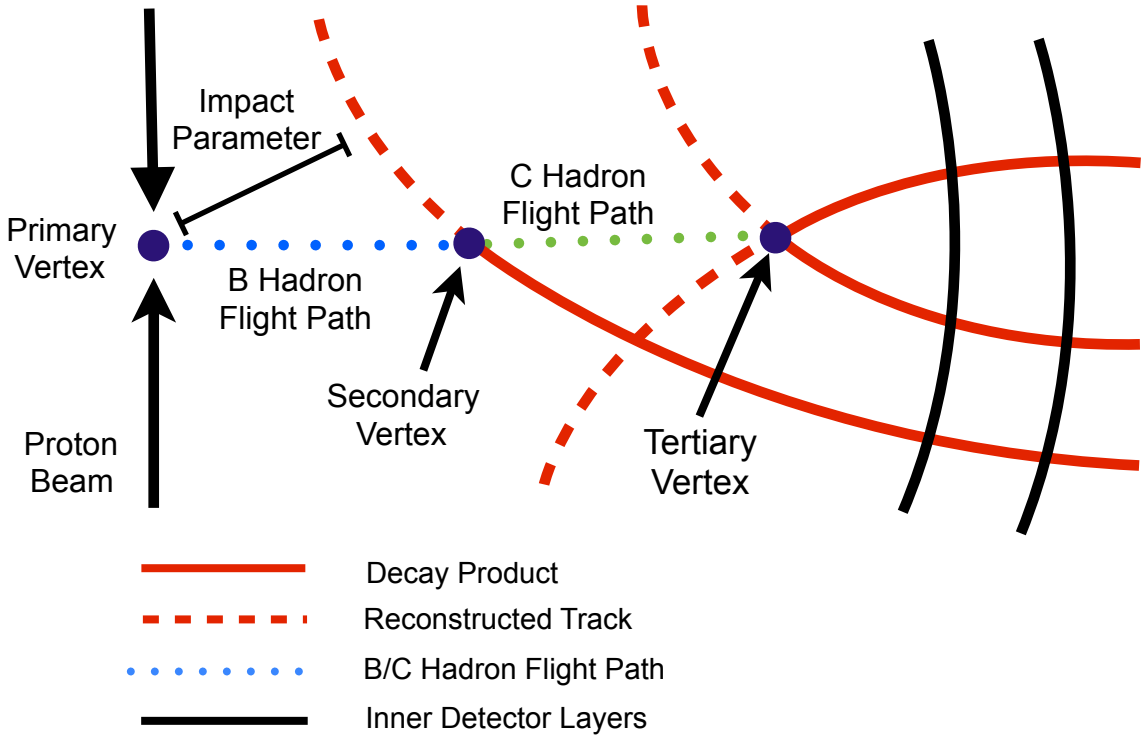
- There have been many changes that affect flavour tagging since Run 1.
  - Introduction of IBL.
  - Change from MV1 to MV2.
  - Change of collision energy.

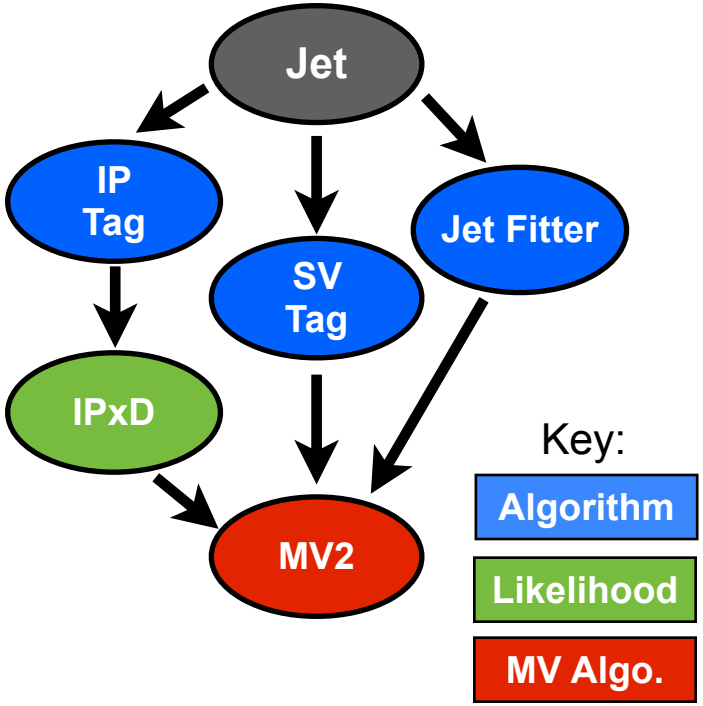
- Run 2 flavour tagging needs to be understood.
  - Flavour tagging has been tested and optimised in simulation.
  - Also need to validate flavour tagging in data.
  - This is done by comparing data to simulation.

Pub Note

- We are aiming to release a Pub Note
  - Aiming for Lepton Photon: Circulate to group by Wednesday 12th August
  - First set of plots produced, framework in place.
  - Note is written: https://cds.cern.ch/record/2032461
  - First reading has occurred and first set of comments have now been addressed.

Flavour Tagging Algorithms





- IPxD
  - Use impact parameter relative to PV for all tracks.

- SV1
  - Reconstruct a secondary vertex from the crossing of two or more tracks.
  - Look at properties of SV that discriminate

- Jet Fitter
  - Reconstruct many vertices along a flight path axis.
  - Look for properties of these vertices that can discriminate in these vertices.

MV2

- In Run 1 MV1 used combined the base taggers.
- In Run 2 MV2 will combine the inputs to the base taggers to improve performance

4 Samples

- MC Sample:

- **Full xAOD**
- 50ns dijet MC sample data
- Split into 4 slices and the re-weighted (see backup) JZ1W-JZ4W - No JZ0W used.
- ~ 8M Events.

"mc15_13TeV.361021.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ1W.merge.AOD.e3569_s2576_s2132_r6630_r6264/"
"mc15_13TeV.361022.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ2W.merge.AOD.e3668_s2576_s2132_r6630_r6264/"
"mc15_13TeV.361023.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ3W.merge.AOD.e3668_s2576_s2132_r6630_r6264/"
"mc15_13TeV.361024.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ4W.merge.AOD.e3668_s2576_s2132_r6630_r6264/"

- Data Sample:

- 50ns data from stable beam collisions.
- **FTAG** derivation
- ~6M Events from 7 Runs: 270806, 270953, 271048, 271298, 271421, 271516 and 271595
- This corresponds to 770K events passing cuts.

"data15_13TeV.00270806.physics_Main.merge.DAOD_FTAG1.f611_m1463_p2375/"
"data15_13TeV.00270953.physics_Main.merge.DAOD_FTAG1.f611_m1463_p2375/"
"data15_13TeV.00271048physics_Main.merge.DAOD_FTAG1.f611_m1463_p2375/"
"data15_13TeV.00271421.physics_Main.merge.DAOD_FTAG1.f611_m1463_p2375/"
"data15_13TeV.00271516.physics_Main.merge.DAOD_FTAG1.f611_m1463_p2375/"
"data15_13TeV.00271595.physics_Main.merge.DAOD_FTAG1.f611_m1463_p2375/"

- We are using NTuples created using Run2BTagOptimisationFramework

Details and Cuts

- 20.1.5.3 with all tags recommended by CP group
- Running xAOD fix on full xAOD

- **HLT_j60 Trigger for Data and MC with Leading Jet P_T > 70 GeV.**

- AntiKt4EMTopoJets
- Run1LooseBadCuts and "ugly" jet removal.

- Jet Calibration:

```
- calibfile  ="JES_MC15Prerecommendation_April2015.config"
- calSeg     ="JetArea_Residual_Origin_EtaJES_GSC" (_Insitu for data)
```

- GRL = *"data15_13TeV.periodAllYear_DetStatus-v63-pro18-01_DQDefects-00-01-02_PHYS_StandardGRL_All_Good.xml"*

Select event if leading jet has:
- njets ≥ 1
- $|\eta| < 2.5$
- $P_T > 70$ GeV
- JVT > 0.641 if ($P_T < 50$ GeV and $|\eta| < 2.4$)

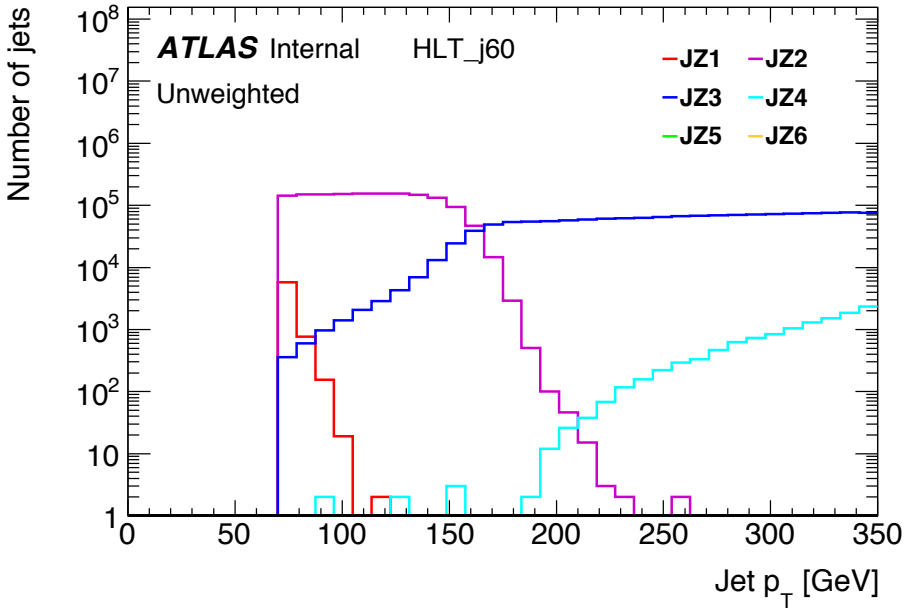
Then plot subleading if subleading jet has:
- **$P_T > 35$ GeV**
- $|\eta| < 2.5$
- JVT > 0.641 if ($P_T < 50$ GeV and $|\eta| < 2.4$)

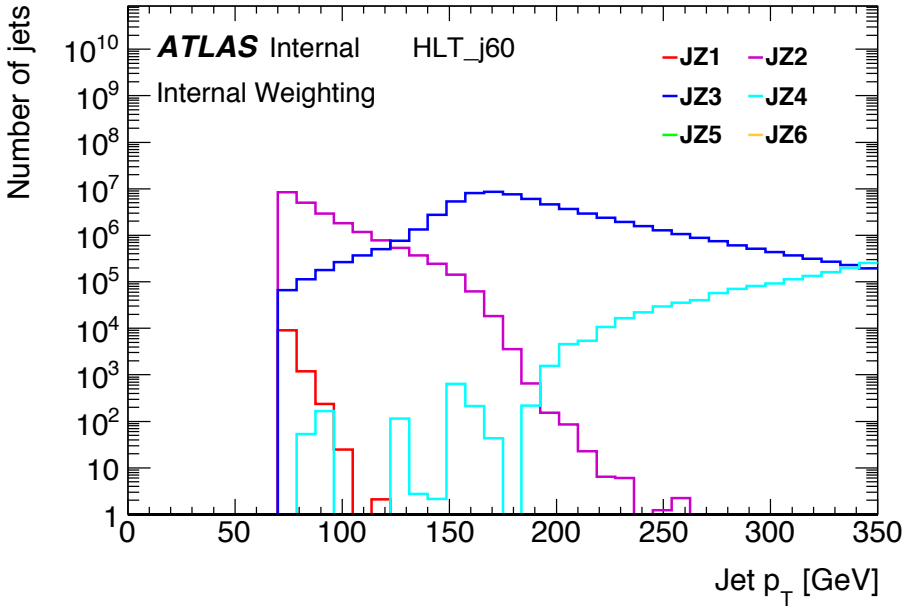
Just For MC

Truth Dijet Test applied to MC to clean sample
- (Lead P_T +Sublead P_T)/2 < 1.4* Truth Lead P_T, for njet > 1
- (Sublead P_T < 1.4 * Truth Sublead P_T), for njet =1

- LabDr_HadF truth matching.

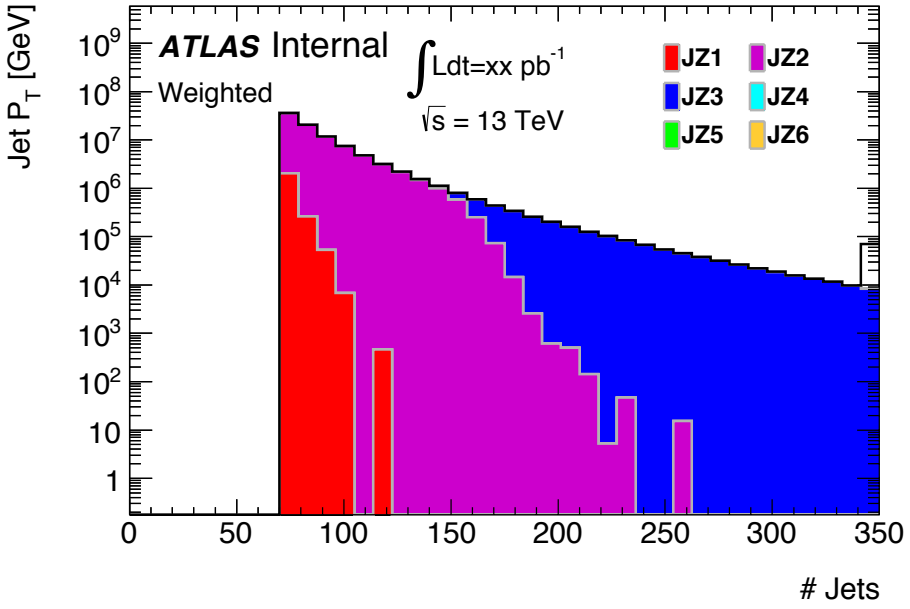
Di-jet sample re-weighting





Total Weight = ***mcwg*(Filter Eff.)*(CS[fb])*(Lumi[fb^{-1}])*** / ***(# Events)***

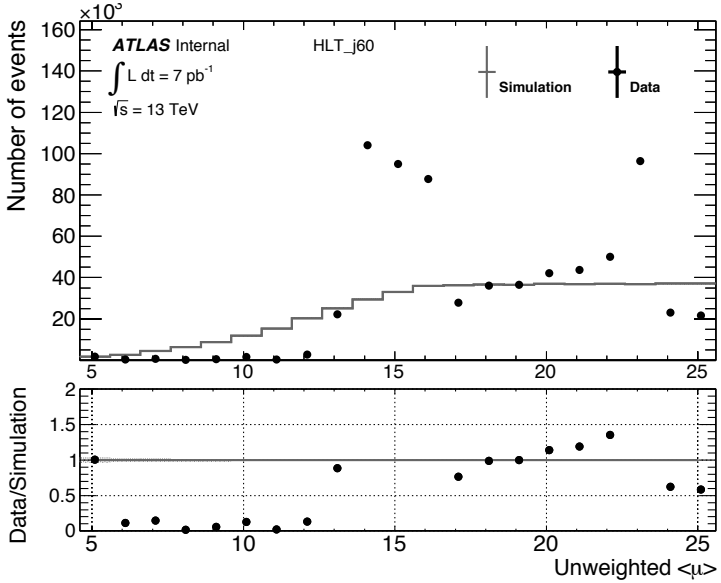
Xs(fb)	Eff.	Slice and Energy
7.8420E+13	6.7198E-04	#JZ1W 20-60 GeV
2.4334E+12	3.3264E-04	#JZ2W 60-160 GeV
2.6454E+10	3.1953E-04	#JZ3W 160-400 GeV
2.5464E+08	5.3009E-04	#JZ4W 400-800 GeV
4.5536E+06	9.2325E-04	#JZ5W 800-1300 GeV
2.5752E+05	9.4016E-04	#JZ6W 1300-1800 GeV



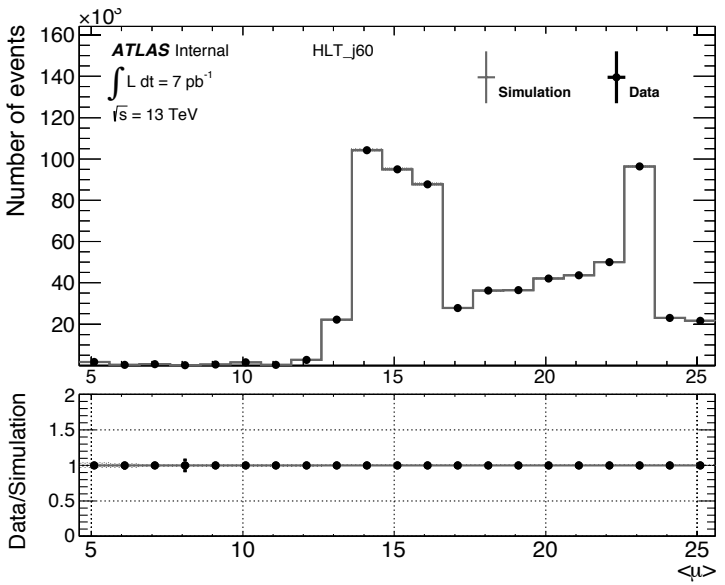
Then integral of MC is normalised to integral of data

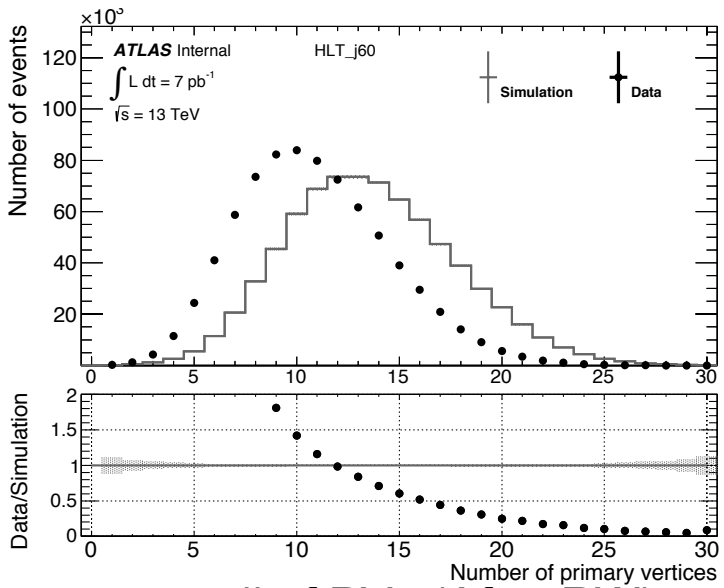
<μ> Reweighting

Before:

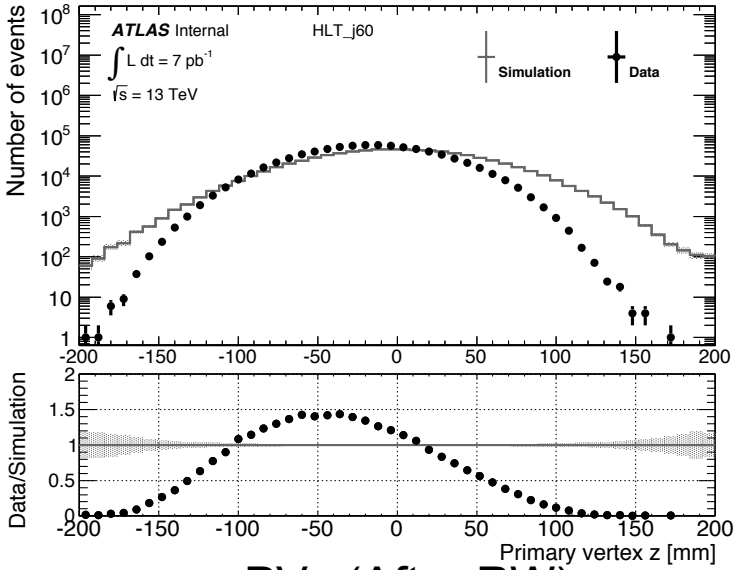


After:





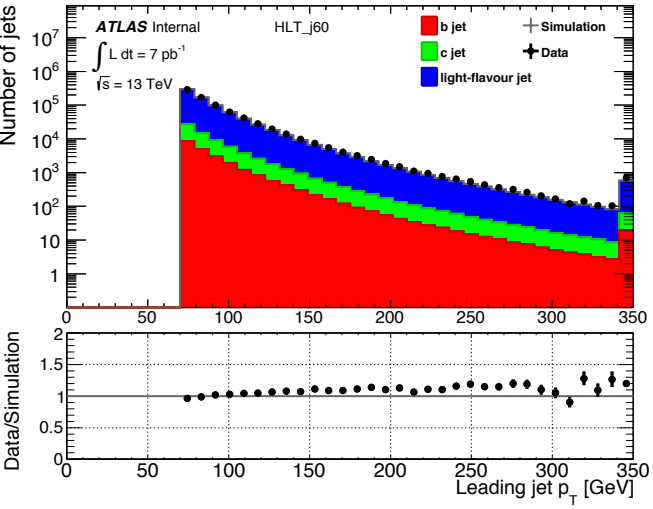
# of PVs (After RW)



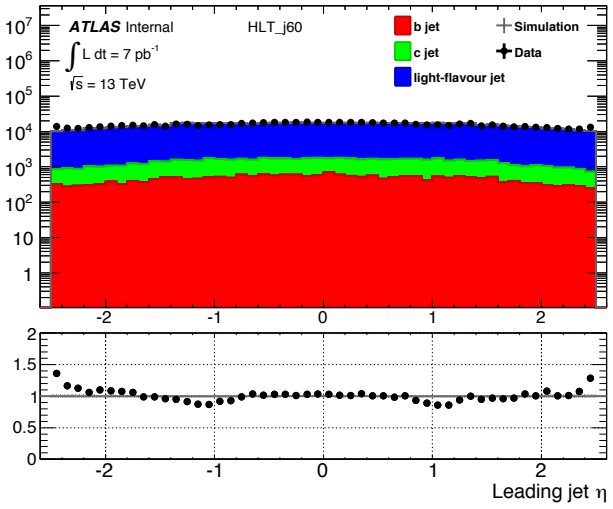
PVz (After RW)

Jet Kinematic Distributions

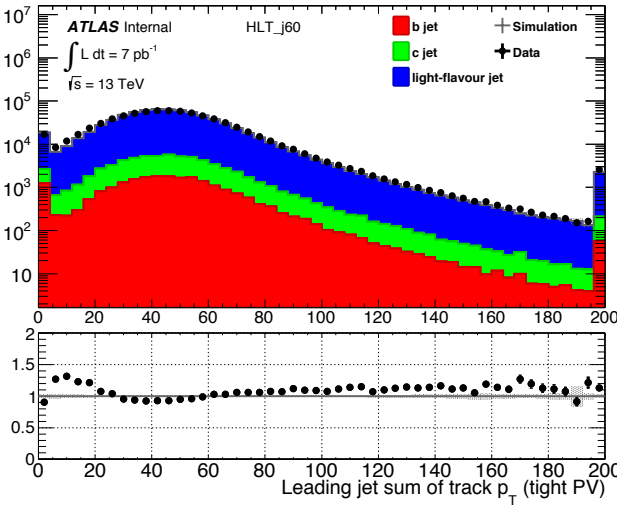
Leading Jet:



Jet P_T

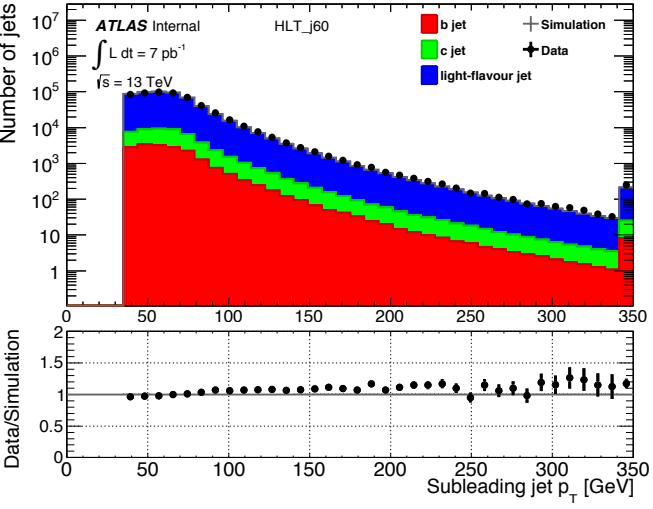


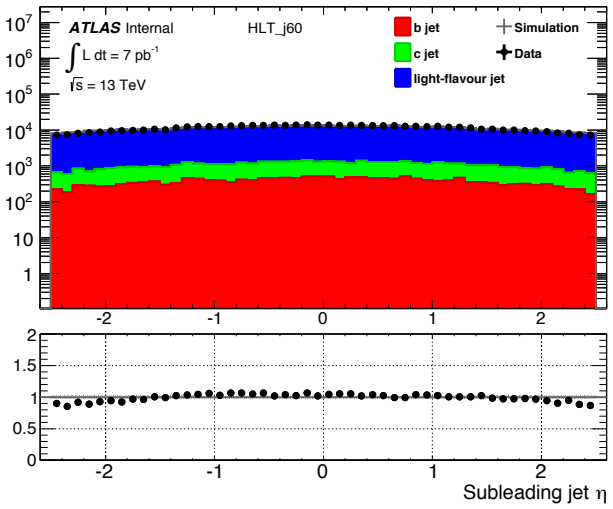
Eta

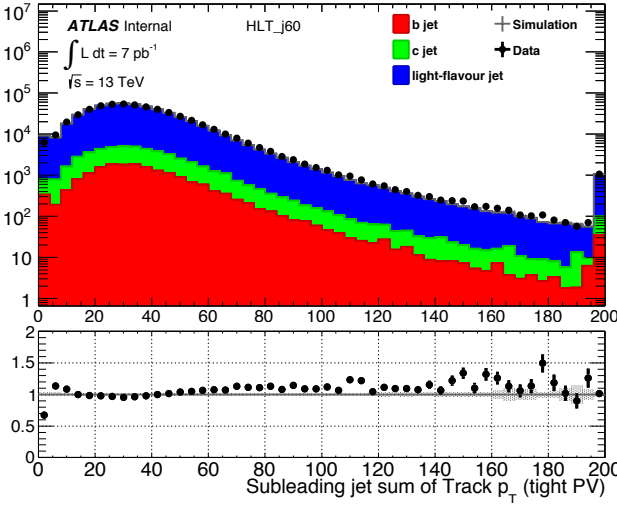


Sum of Track P_T

Sub-Leading Jet:

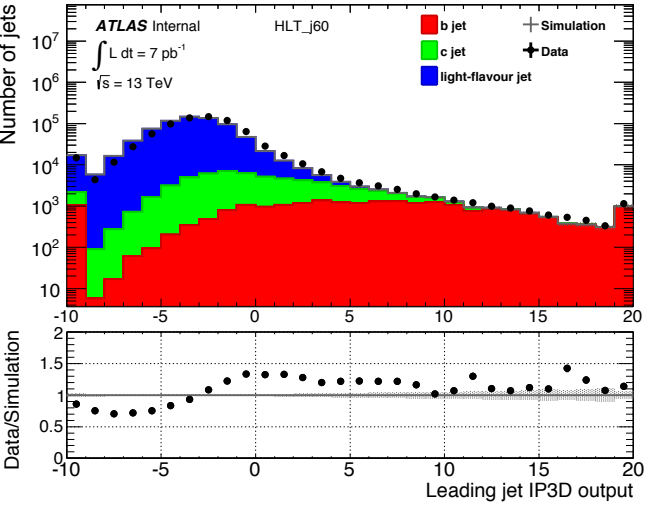


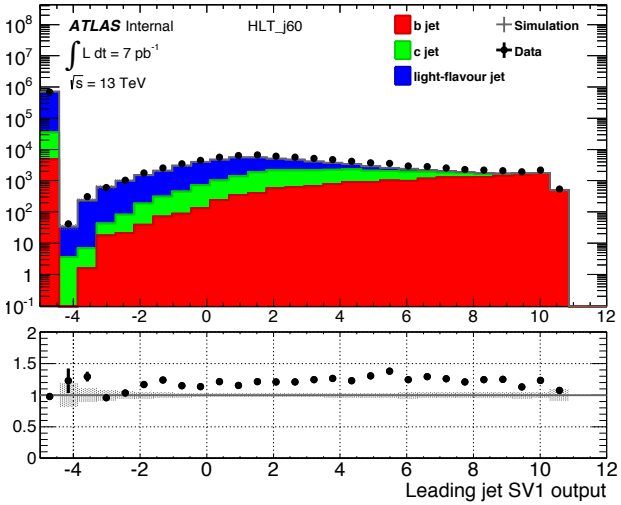


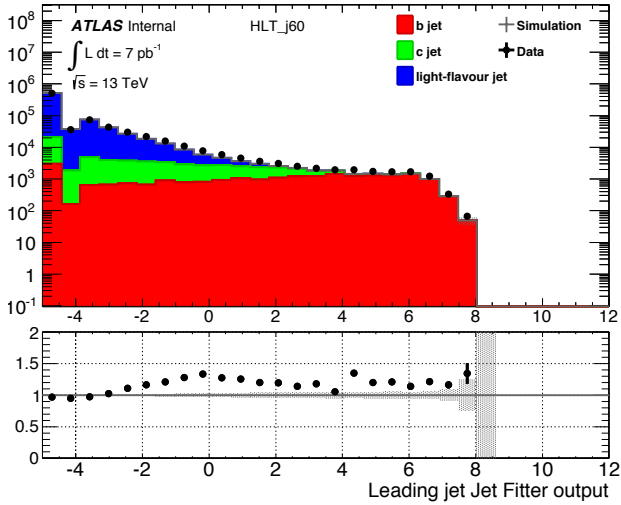


Input Taggers

Leading Jet:

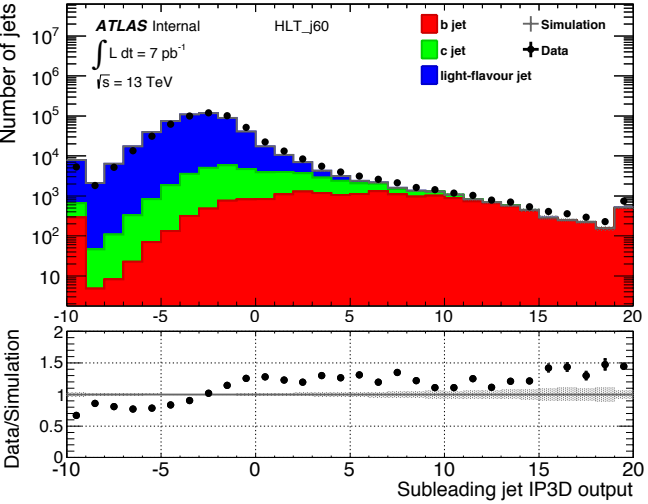


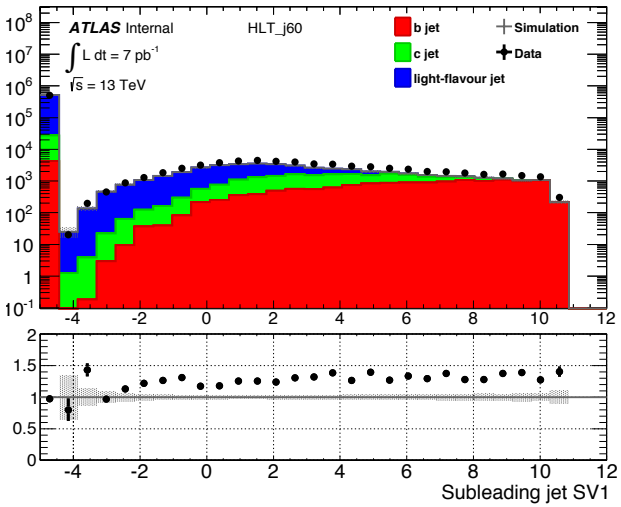


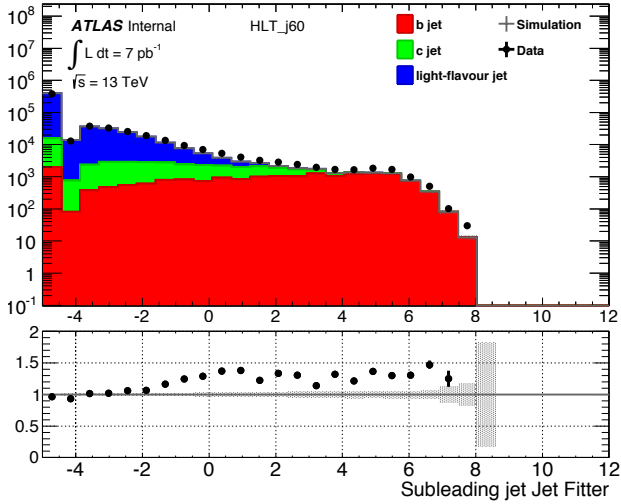


IP3D SV1 JF

Sub-Leading Jet:

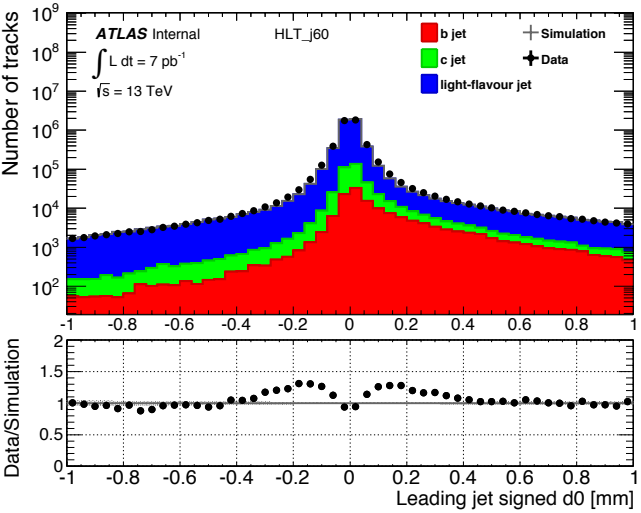




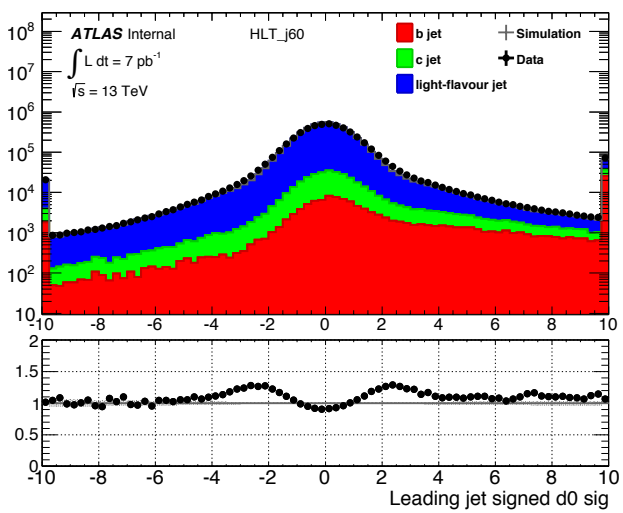


Transverse Impact Parameter

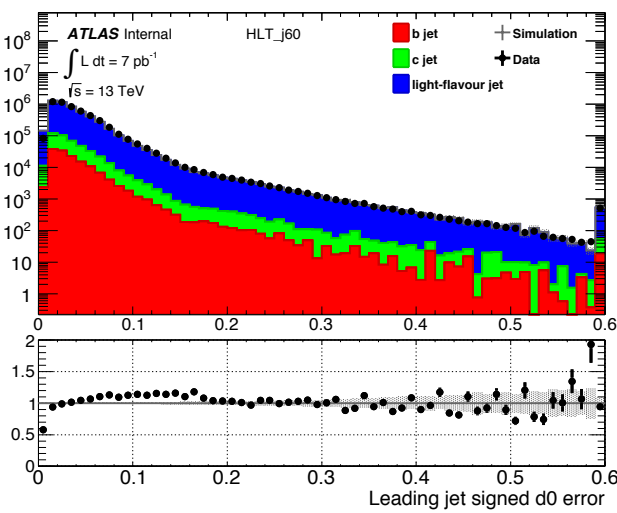
Leading Jet:



d0

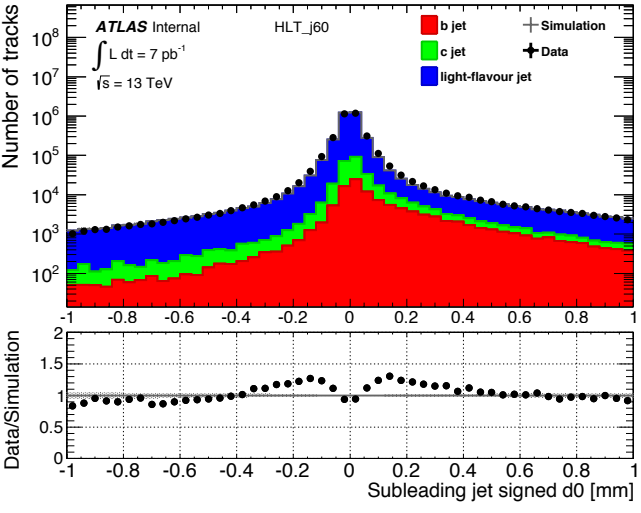


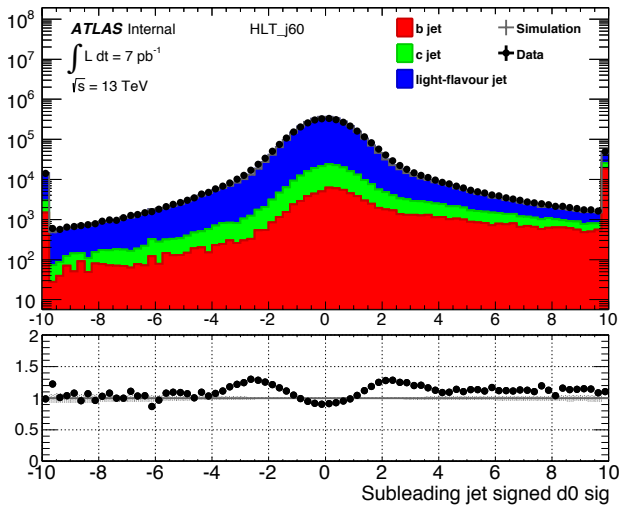
d0 significance

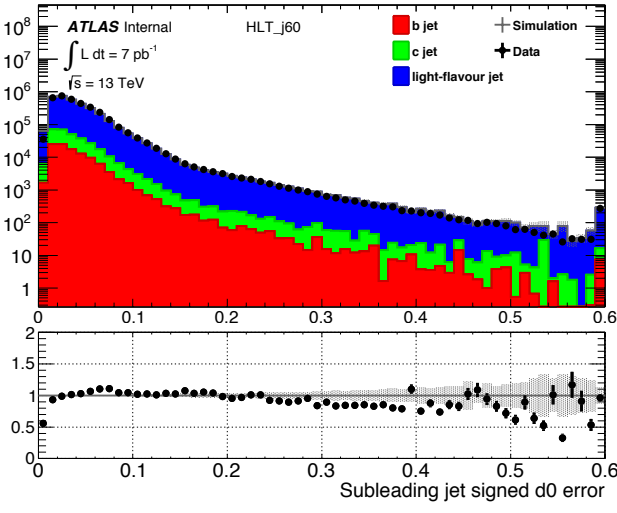


d0 error

Subleading Jet:

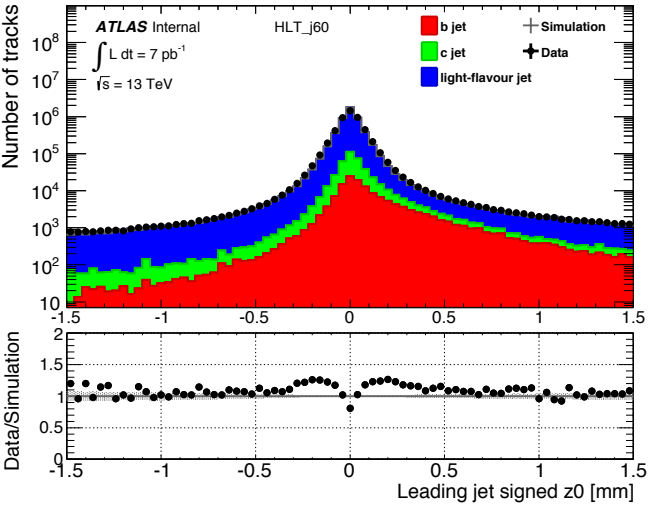




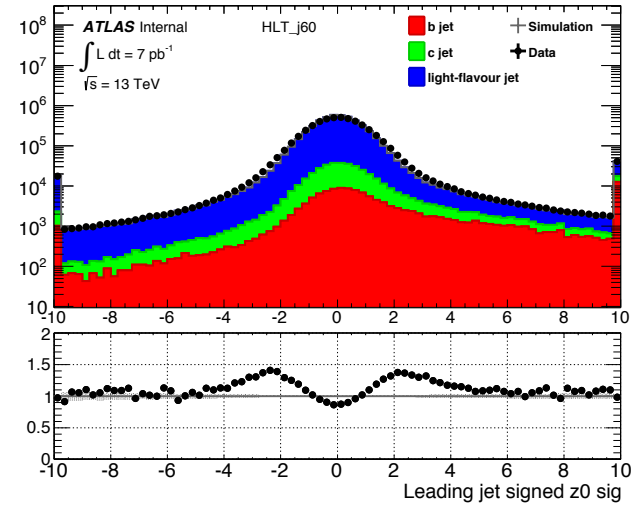


Longitudinal Impact Parameter

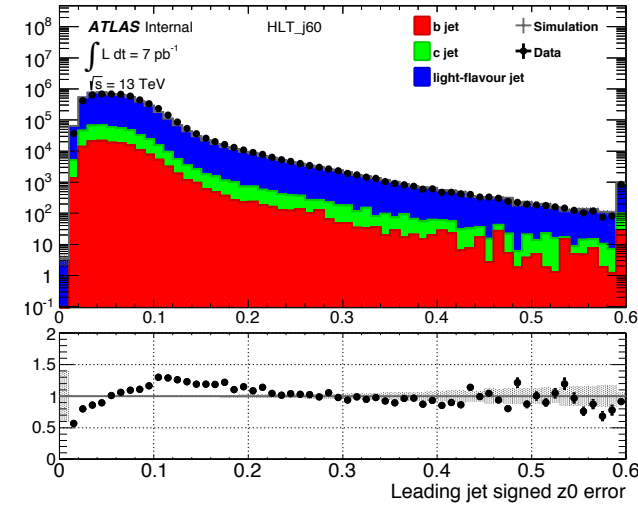
Leading jet:



z0

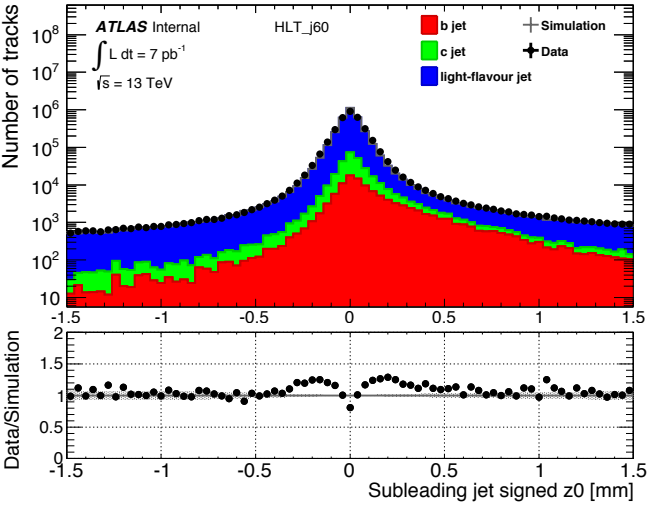


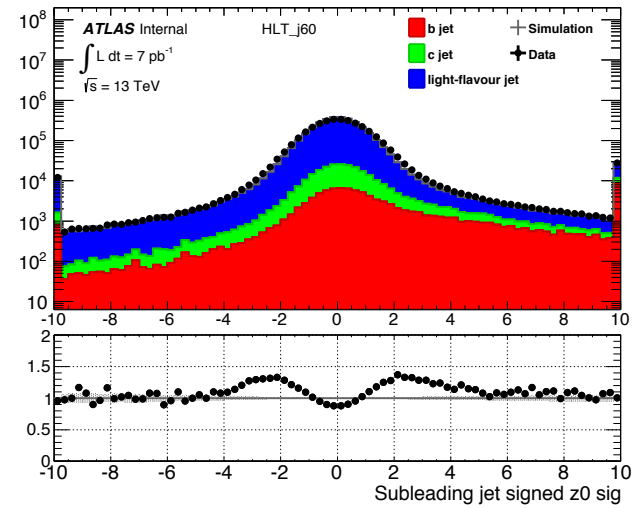
z0 significance

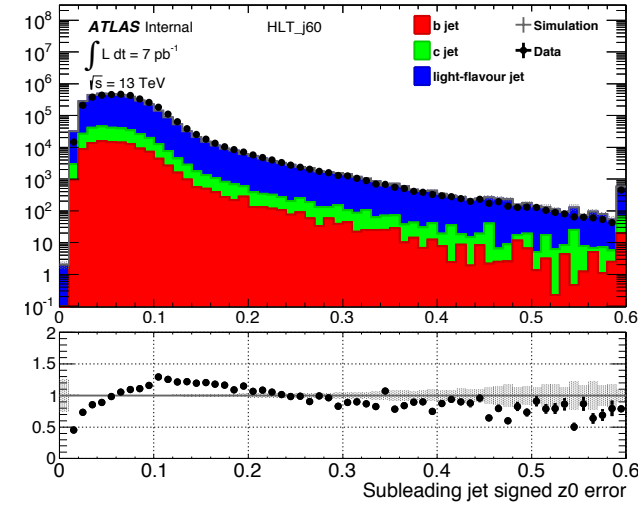


z0 error

Subleading jet:





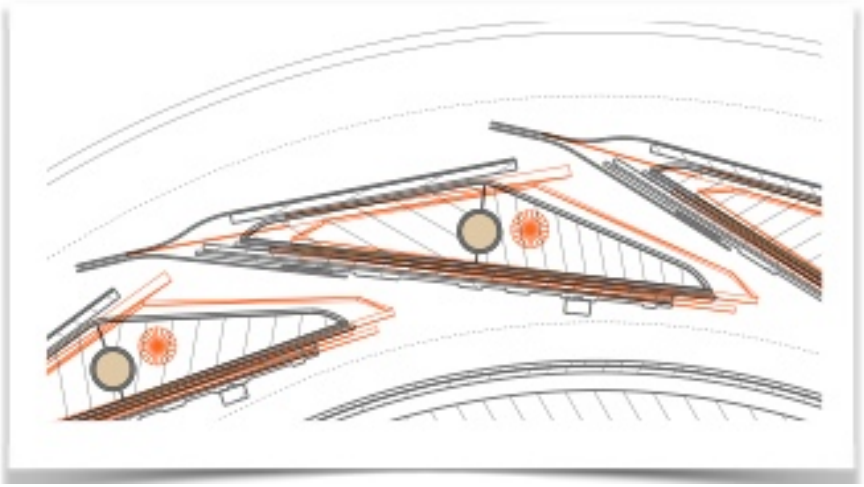


Problems

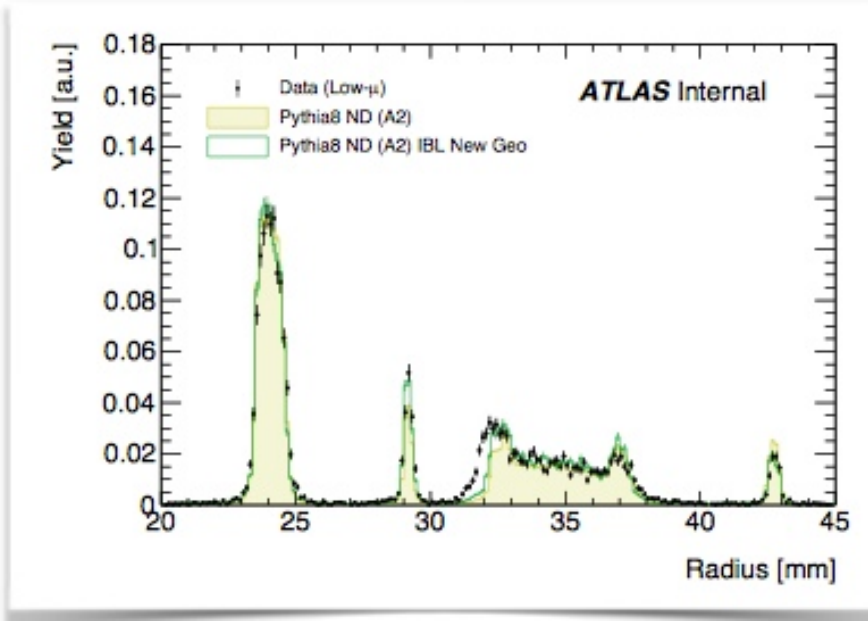
- Imperfect alignment (high pT)
- Geometry tag missing 23% IBL material (low pT)
  - New Geo. Tag Produced and validated.
- Angle of overlap issue
  - Geo. Tag avaliable for 25ns data.

Simone Pagan Griso, Heather Gray

https://indico.cern.ch/event/433839/contribution/6/attachments/1128840/1612854/PC_20jul.pdf



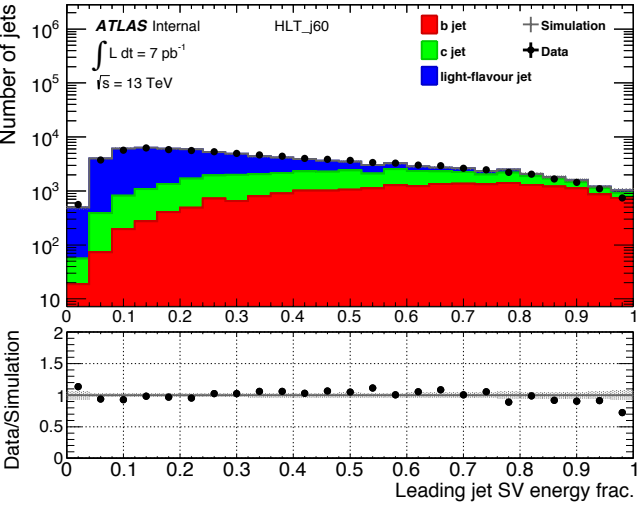
GEO Model
IBL



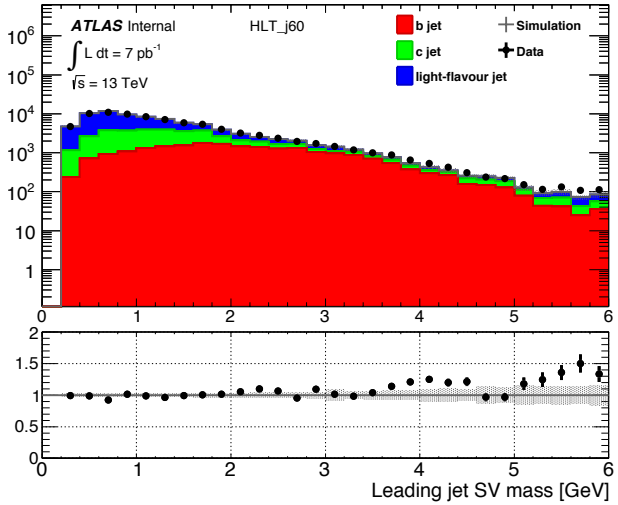
SV1 Variables

Leading Jet:

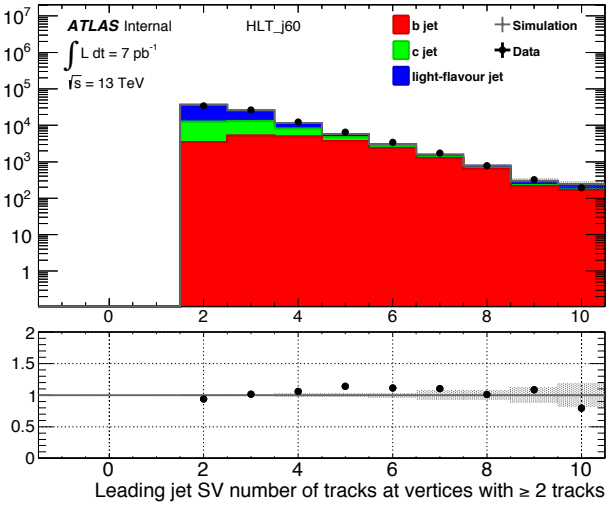
Only Filled if a Secondary Vertex is found



Energy Frac.

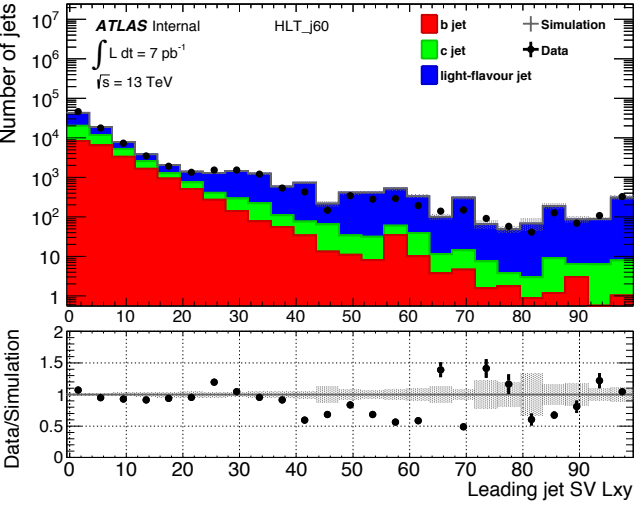


SV1 Mass

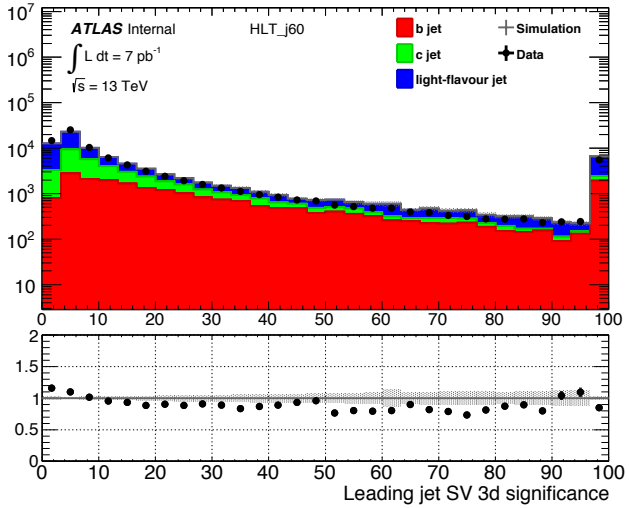


Tracks at SV1 Vertex

SV1 Lxy



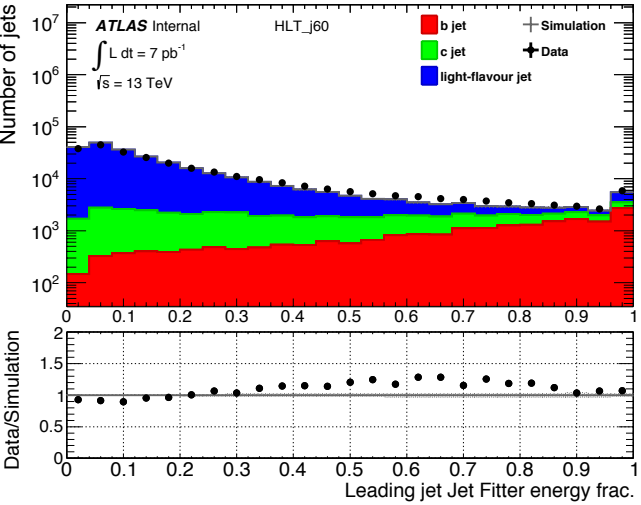
SV1 3D Sig.



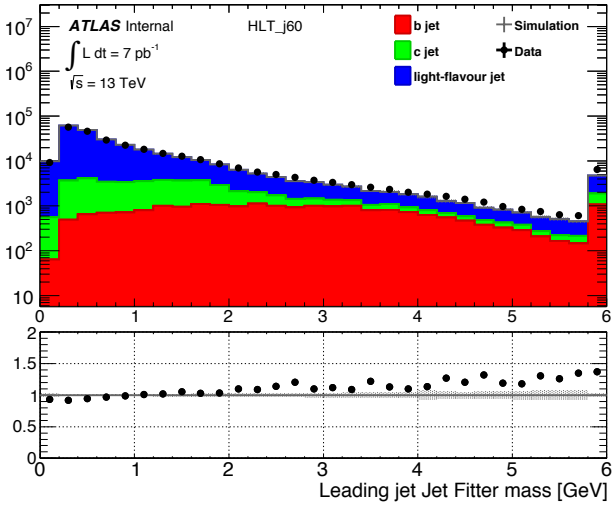
JF Variables

Leading Jet:

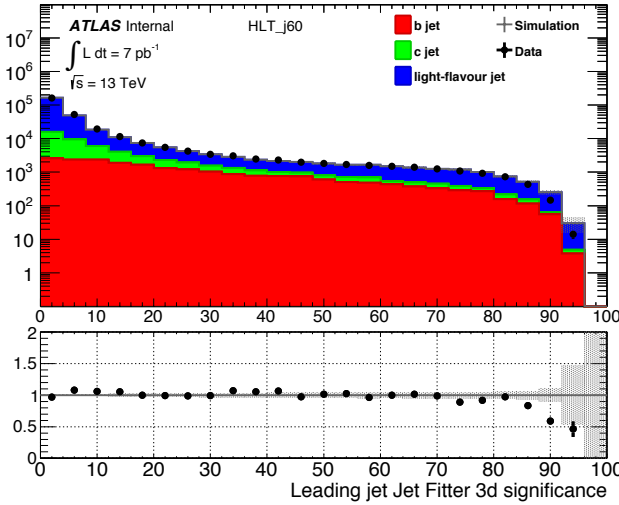
Only Filled if a Jet Fitter Vertex is found



Energy Frac.

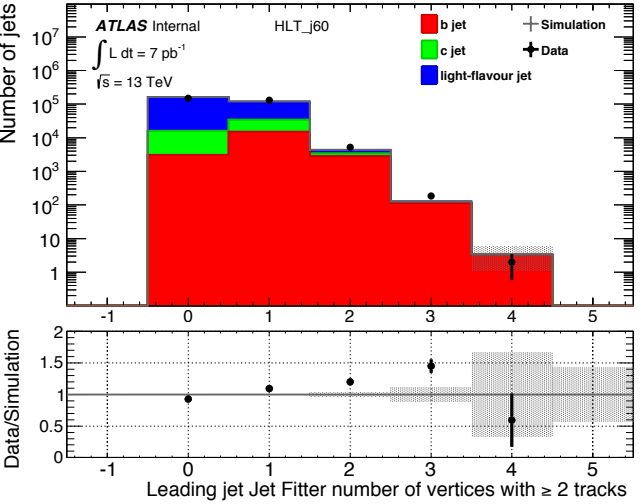


JF Mass

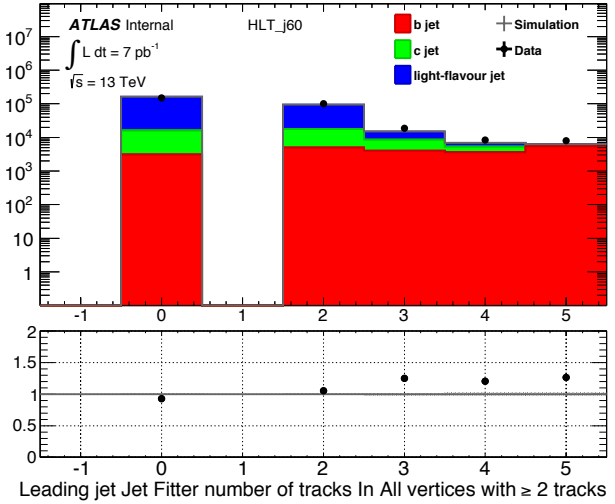


JF 3D Sig

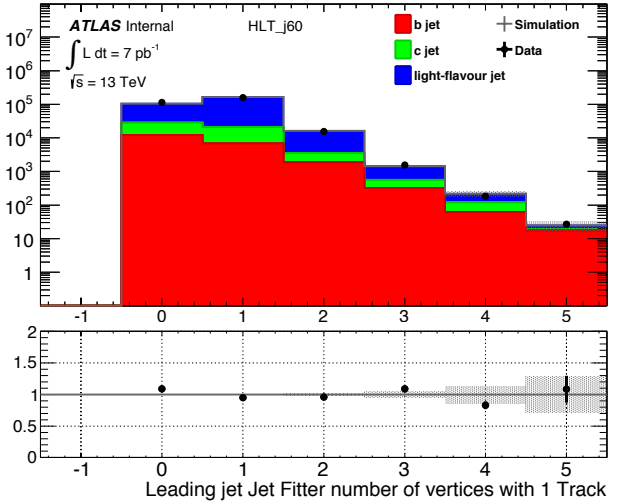
Vertices with at least 2 Tracks



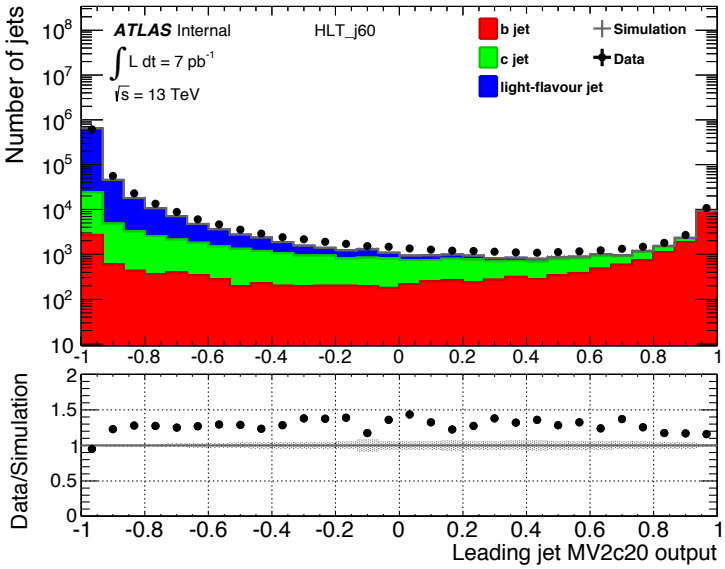
Tracks at Vertices with at least 2 Tracks

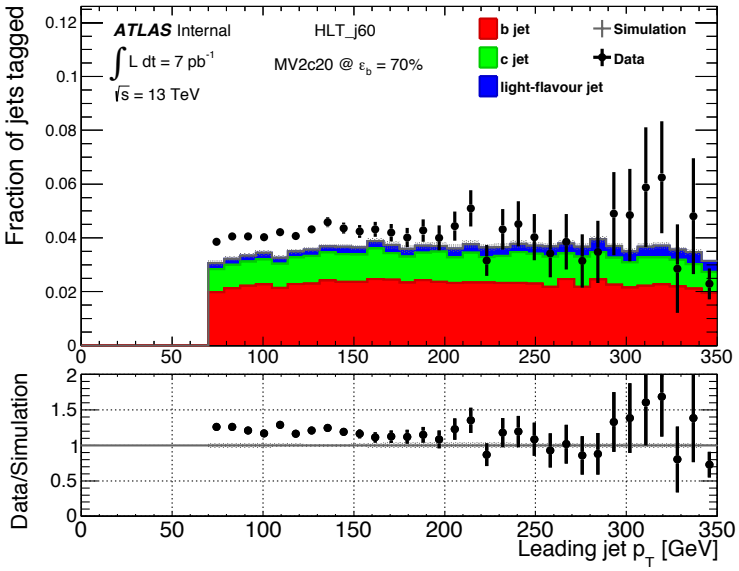


Vertices with 1 Track

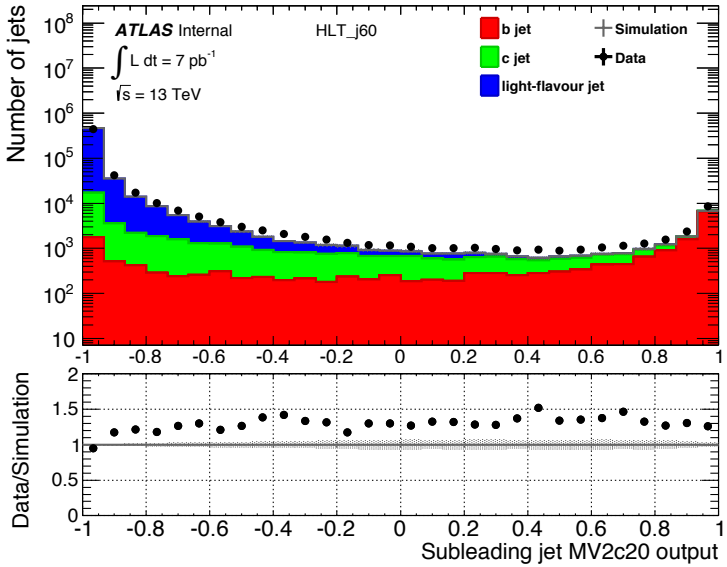


Leading Jet:





Sub-Leading Jet:

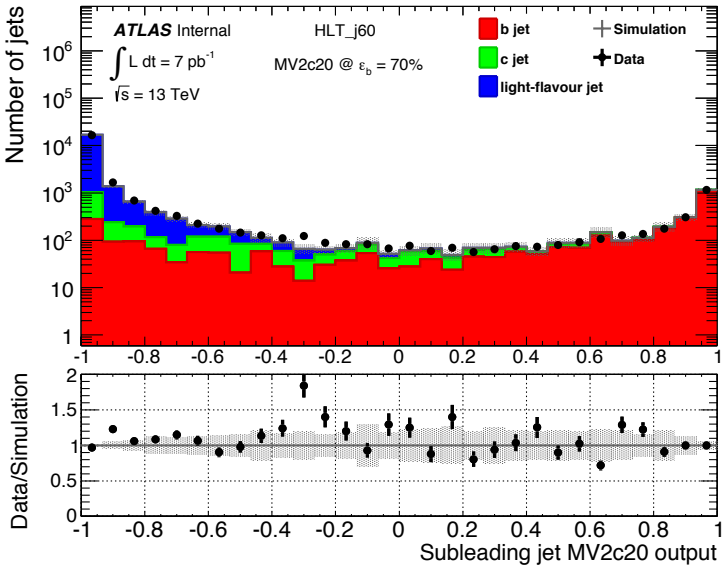


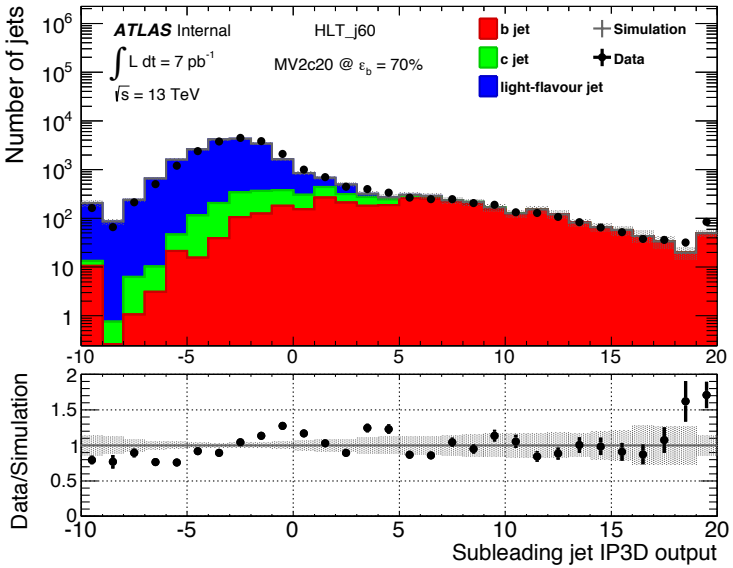
MV2c20

Tag Rate

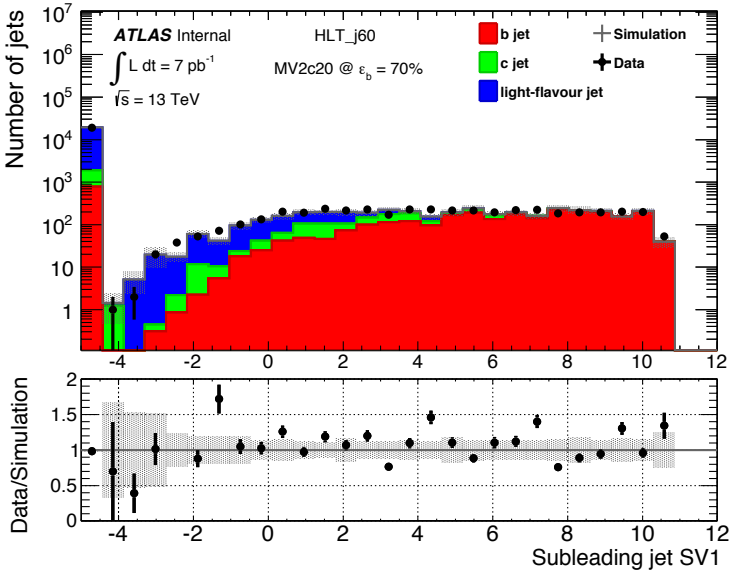
b Enhanced Sample - Subleading Jet

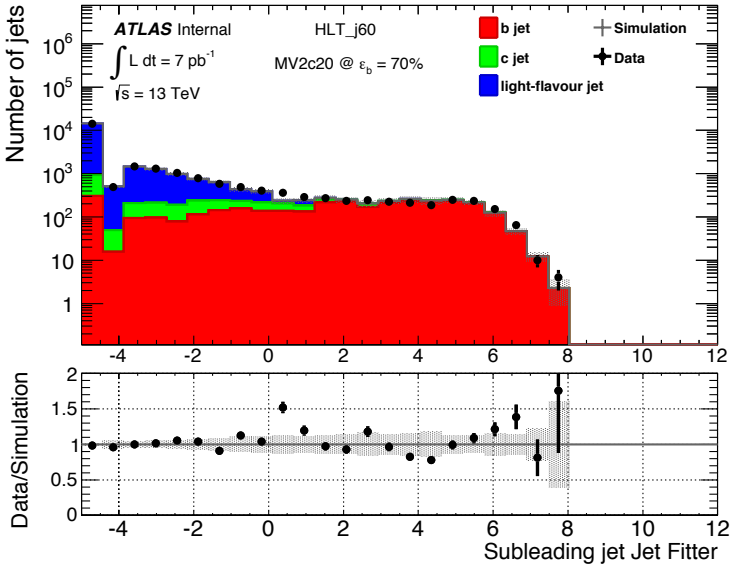
MV2c20: *Leading MV2c20 > -0.0436 which is 70% b-efficiency* IP3D:





SV1: Jet Fitter:





Conclusions

• The plot making framework is in place and first version of all plots made.
- Some good agreement
    - Kinematic Quantities.
    - SV and Jet Fitter Quantities
    - b-enhanced sample

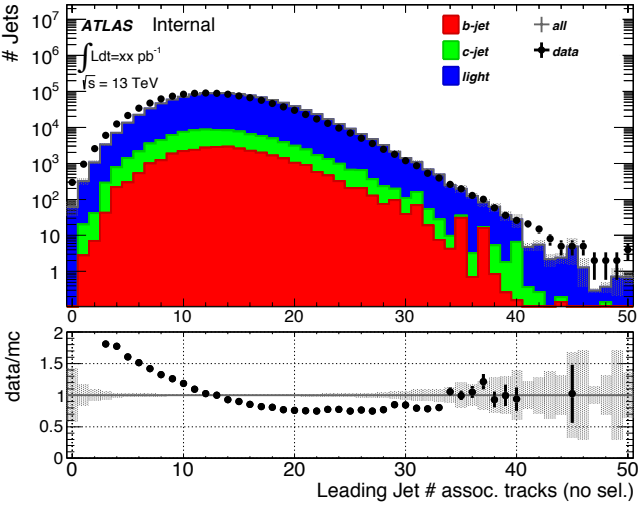
- Still some disagreements
    - Outputs of taggers.
    - Impact parameter

• Possibility to reprocess 50ns data and MC to hopefully improve I.P. resolution.

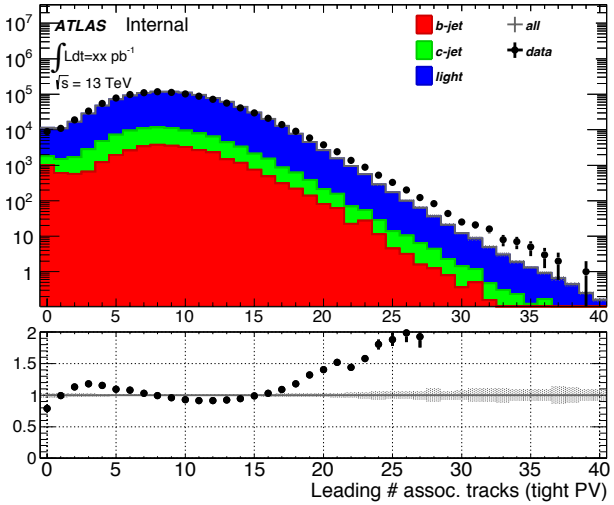
• Need to push on with documentation for Pub Note:
- Note written, plotting framework in place.
- Hopefully we will see new samples with improved I.P. resolution.

Track Distributions

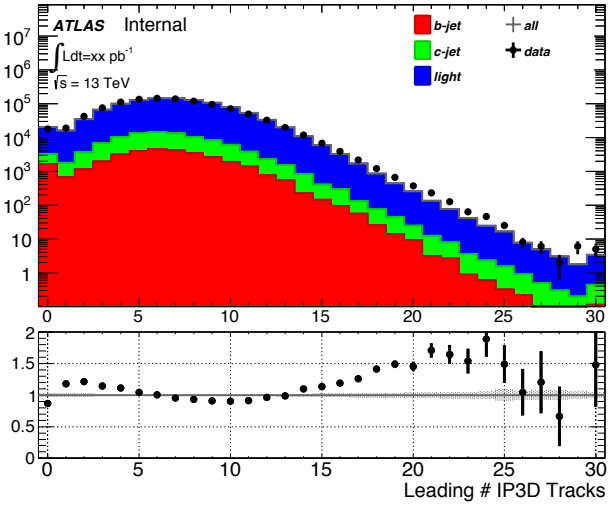
Leading Jet:



Tracks:
Directly from Track Container



Tracks:
Tight PV Selection



Tracks:
IP3D Selection
P_T > 1 GeV

Sub-Leading Jet: