



# Flavour Tagging Commissioning with Data

Laurie McClymont,  
Andreas Korn

Flav. Tag Alg. Meeting  
08/08/16



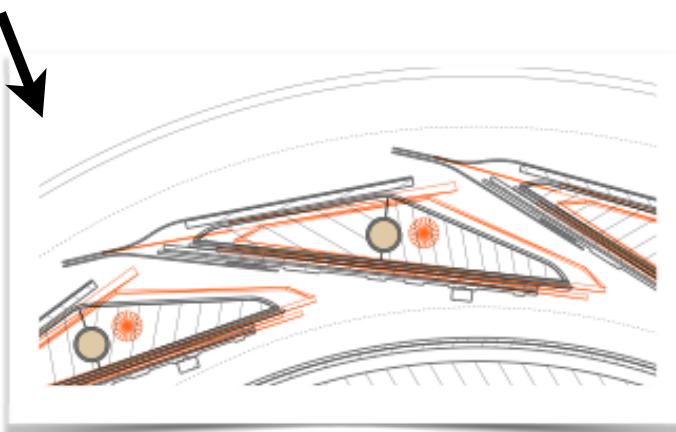
## Aims

- Comparing data to MC in dijet events to test our understanding of flavour tagging.
- This was done in 2015 data - We saw some discrepancies - [see here](#)
- Tracking group has an updated geometry (see next slide)
- Here we study the impact of using the updated geometry

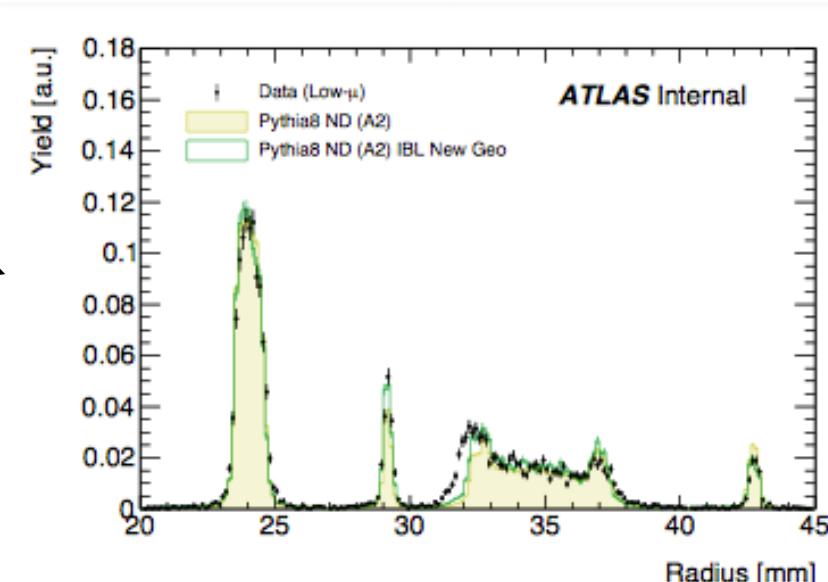
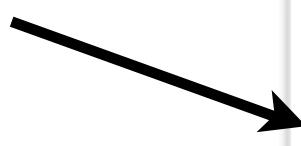


### 3 Changes in Geometry

- Geometry tag missing 23% IBL material
  - New Geo. Tag Produced
  - In validation.
- Angle of overlap issue for IBL
  - Unlikely to be a large effect.



**GEO Model**  
**IBL**



- These issues affect the error estimation and hence the  $d_0/z_0$  significance distribution
- Also the  $d_0$  and  $z_0$  resolution is also effected



## 4 Samples

- **MC Sample (New Geo):**

- Full AODs, dijet Pythia8 QCD ~4M Events. r20.7
- Geometry = “**ATLAS-R2-2015-03-15-00\_VALIDATION**”

*mc15\_13TeV.361021.Pythia8EvtGen\_A14NNPDF23LO\_jetjet\_JZ1W.merge.AOD.e3569\_s2781\_r8122\_r7676\_tid08750161\_00  
mc15\_13TeV.361022.Pythia8EvtGen\_A14NNPDF23LO\_jetjet\_JZ2W.merge.AOD.e3668\_s2781\_r8122\_r7676\_tid08750166\_00  
mc15\_13TeV.361024.Pythia8EvtGen\_A14NNPDF23LO\_jetjet\_JZ4W.merge.AOD.e3668\_s2781\_r8122\_r7676\_tid08750174\_00  
mc15\_13TeV.361023.Pythia8EvtGen\_A14NNPDF23LO\_jetjet\_JZ3W.merge.AOD.e3668\_s2781\_r8122\_r7676\_tid08750170\_00*

- **MC Sample (Old Geo):**

- Full xAOD, dijet MC sample, ~ 4M Events. , r20.7
- Geometry = “**ATLAS-R2-2015-03-01-00\_VALIDATION**”

*“mc15\_13TeV.361021.Pythia8EvtGen\_A14NNPDF23LO\_jetjet\_JZ1W.merge.AOD. e3569\_s2576\_s2132\_r7725\_r7676/”  
“mc15\_13TeV.361022.Pythia8EvtGen\_A14NNPDF23LO\_jetjet\_JZ2W.merge.AOD.e3668\_s2576\_s2132\_r7725\_r7676/”  
“mc15\_13TeV.361023.Pythia8EvtGen\_A14NNPDF23LO\_jetjet\_JZ3W.merge.AOD.e3668\_s2576\_s2132\_r7725\_r7676/”  
“mc15\_13TeV.361024.Pythia8EvtGen\_A14NNPDF23LO\_jetjet\_JZ4W.merge.AOD.e3668\_s2576\_s2132\_r7725\_r7676/”*

- **Data Sample:**

- 50ns data from stable beam collisions.
- **FTAG** derivation
- ~4M Events from 6 Runs: 270806, 270953, 271048, 271421, 271516 and 271595

*data15\_13TeV.00270806.physics\_Main.merge.DAOD\_FTAG1.r7600\_p2521\_p2667/  
data15\_13TeV.00270953.physics\_Main.merge.DAOD\_FTAG1.r7600\_p2521\_p2667/  
data15\_13TeV.00271048.physics\_Main.merge.DAOD\_FTAG1.r7600\_p2521\_p2667/  
data15\_13TeV.00271421.physics\_Main.merge.DAOD\_FTAG1.r7600\_p2521\_p2667/  
data15\_13TeV.00271516.physics\_Main.merge.DAOD\_FTAG1.r7600\_p2521\_p2667/  
data15\_13TeV.00271595.physics\_Main.merge.DAOD\_FTAG1.r7600\_p2521\_p2667/  
data15\_13TeV.00271595.physics\_Main.merge.DAOD\_FTAG1.r7600\_p2521\_p2667/  
data15\_13TeV.00271744.physics\_Main.merge.DAOD\_FTAG1.r7600\_p2521\_p2667/*

- We are using NTuples created using Run2BTagOptimisationFramework



## 5 Details and Cuts

- r20.7 with all tags recommended by CP group
- **HLT\_j60 Trigger for MC and Data with Leading Jet  $P_T > 70 \text{ GeV}$ .**
- 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:

- $n_{\text{jets}} \geq 2$  with  $P_T > 35 \text{ GeV}$ ,  $|\eta| < 2.5$   
(and  $JVT > 0.641$  if  $P_T < 50 \text{ GeV}$  and  $|\eta| < 2.4$ )
- Leading jet,  $P_T > 70 \text{ GeV}$ ,  $|\eta| < 2.5$

### Then plot subleading if subleading jet has:

- $P_T > 35 \text{ GeV}$
- $|\eta| < 2.5$
- $JVT > 0.641$  if  $P_T < 50 \text{ GeV}$  and  $|\eta| < 2.4$

### Monte Carlo Cuts

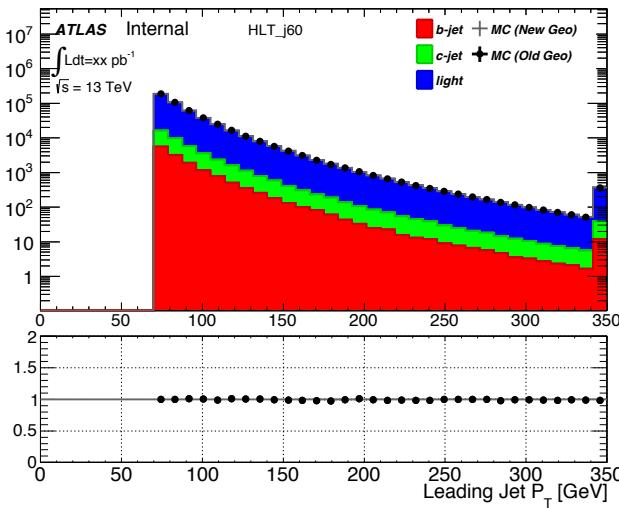
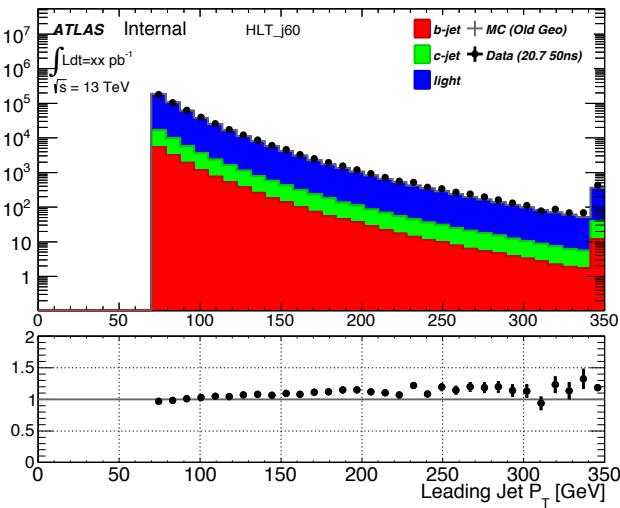
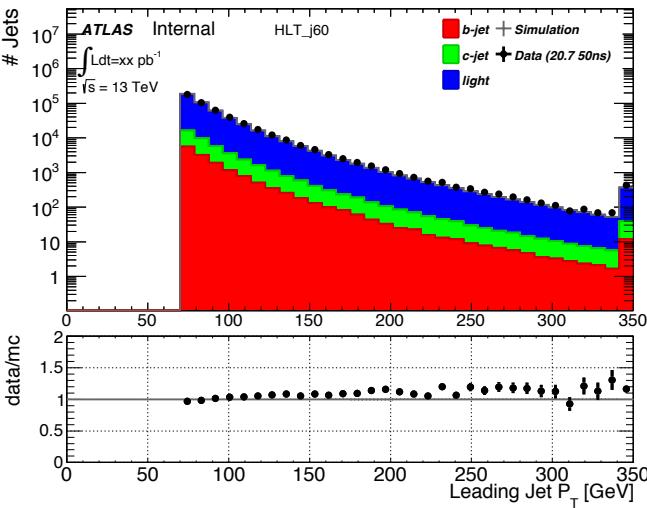
- Truth Dijet Test applied to MC to clean sample  
-  $(\text{Lead } P_T + \text{Sublead } P_T)/2 < 1.4^* \text{ Truth Lead } P_T$

- LabDr\_HadF truth matching.
- Re-weight PU in MC to match data



# 6 Kinematics

## Jet pT

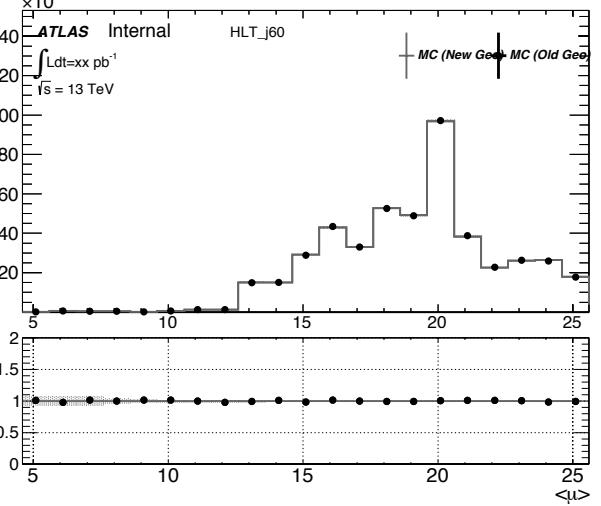
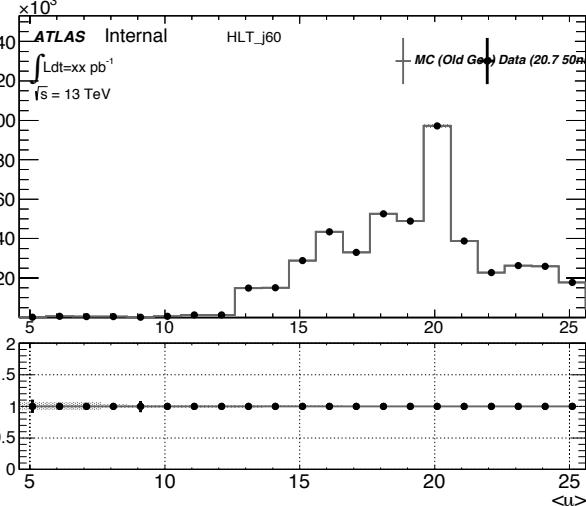
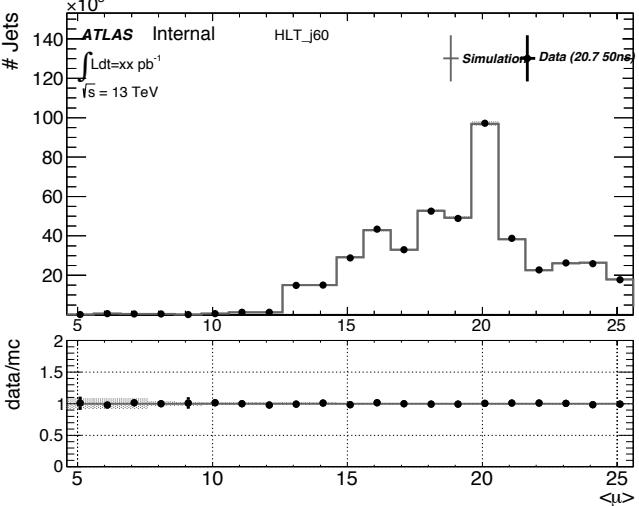


Data / New Geo MC

Data / Old Geo MC

Old Geo MC / New Geo MC

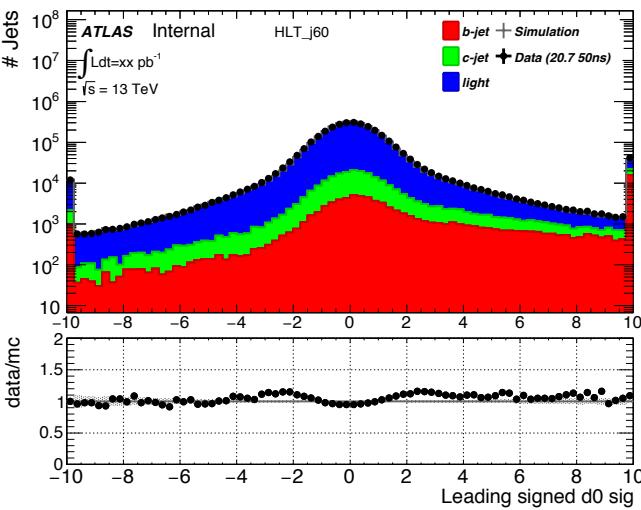
## Mu After Reweighting



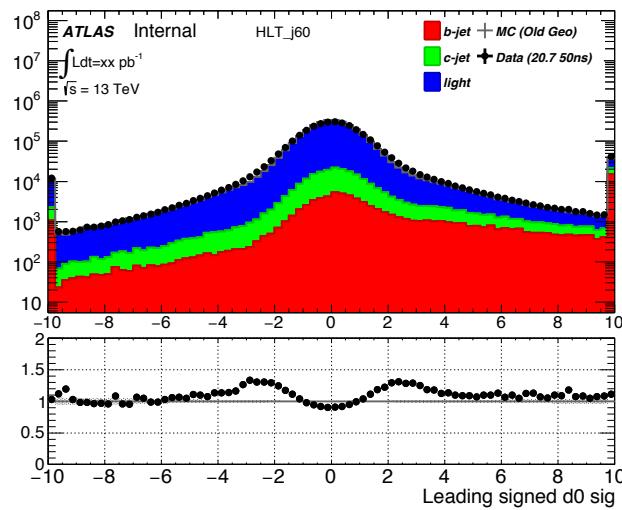


# 7 Signed d0/z0 Significance

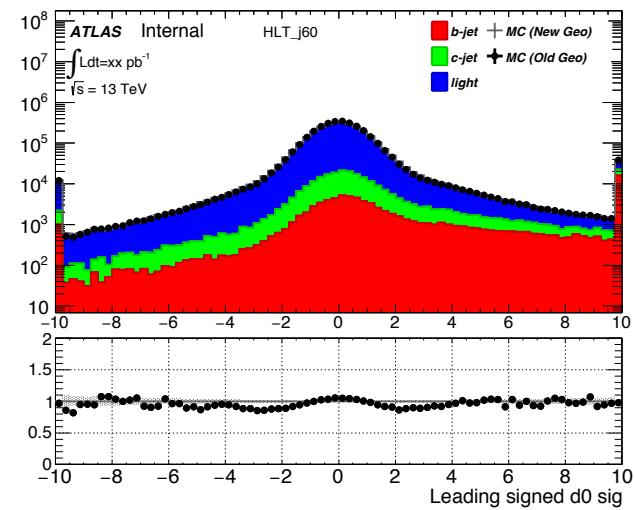
d0 Significance:



Data / New Geo MC

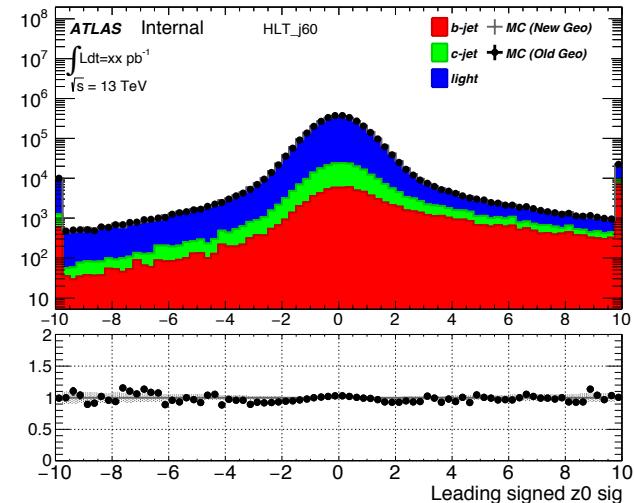
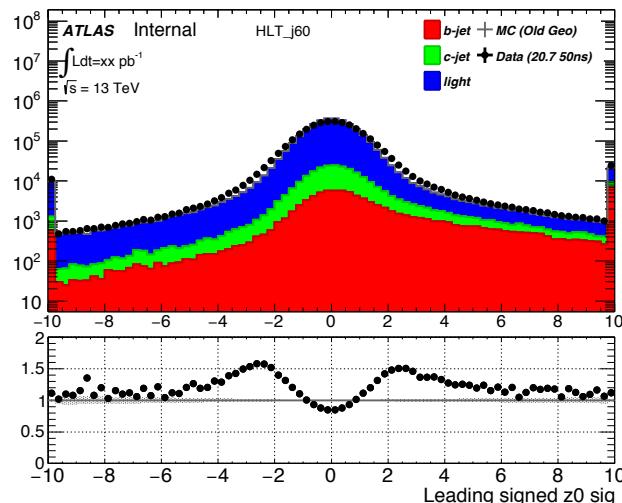
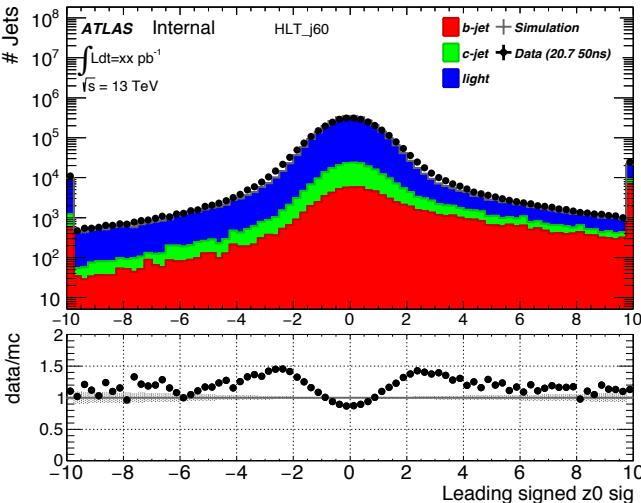


Data / Old Geo MC



Old Geo MC / New Geo MC

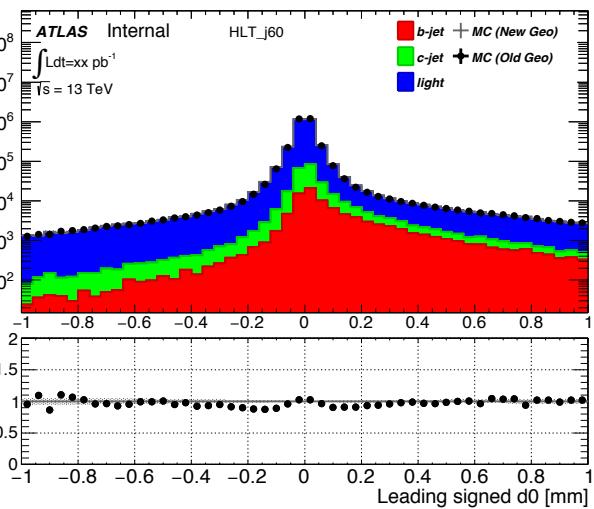
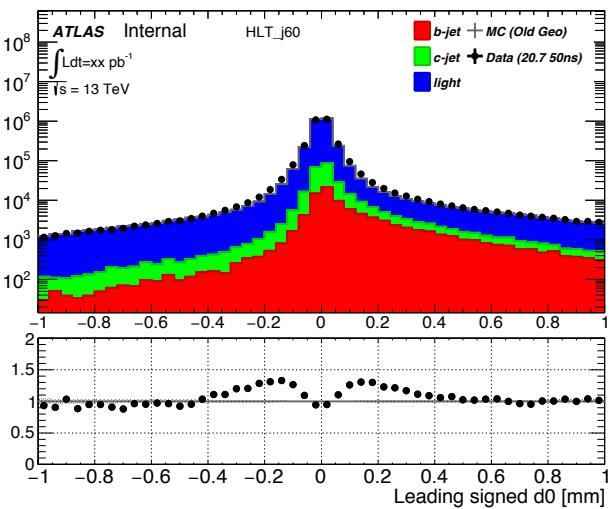
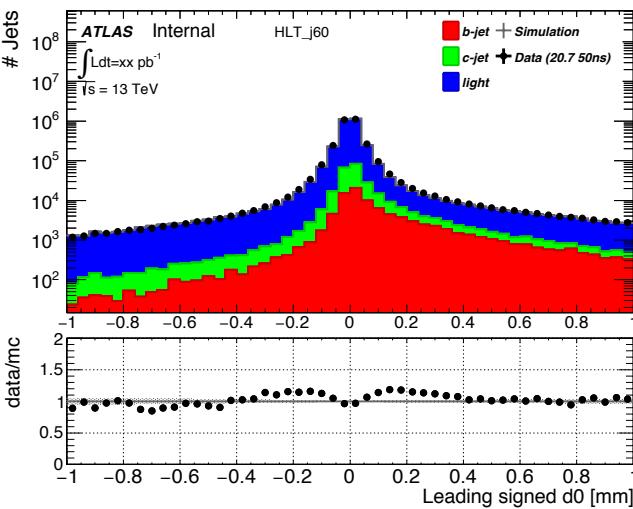
z0 Significance:





## 8 Signed d0/z0

d0:

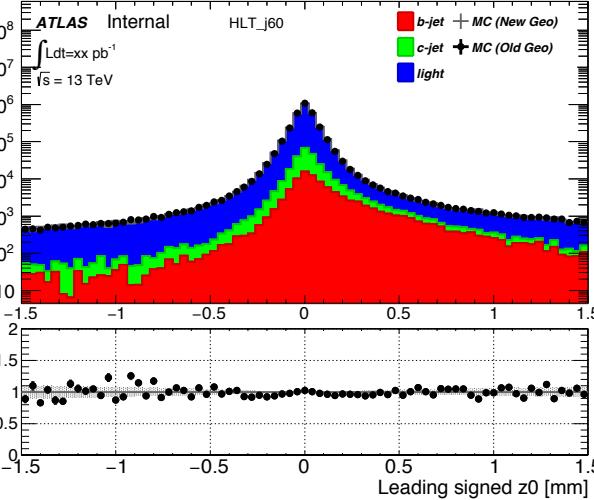
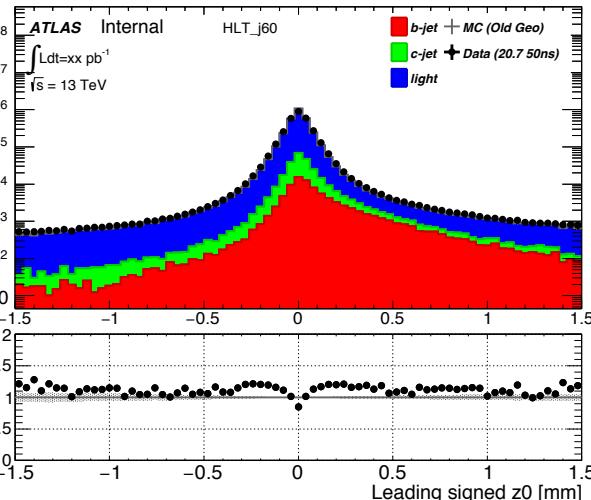
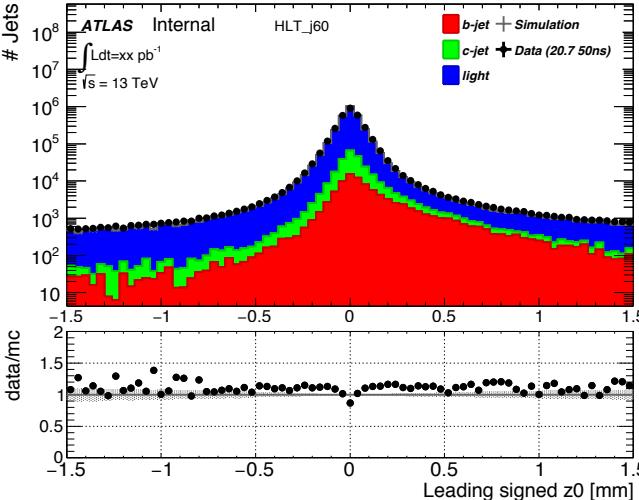


Data / New Geo MC

Data / Old Geo MC

Old Geo MC / New Geo MC

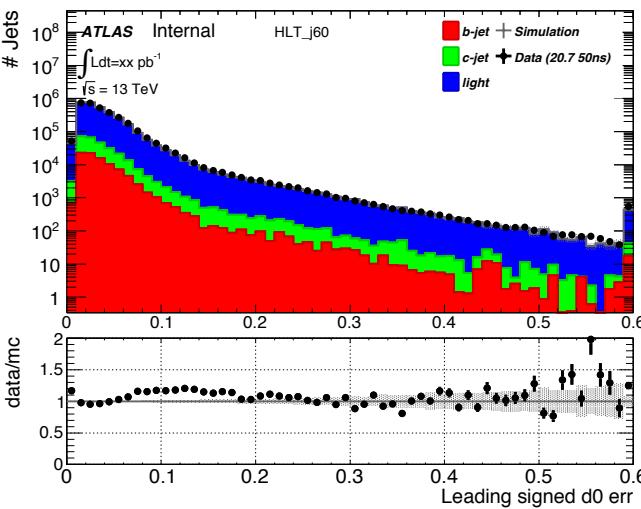
z0:



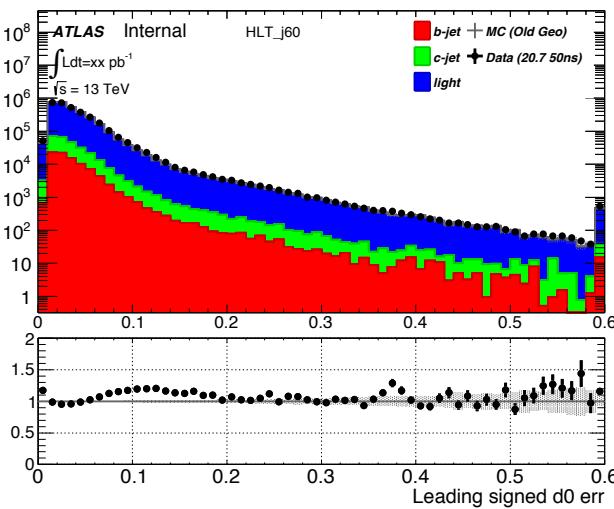


## 9 Signed d0/z0 Error

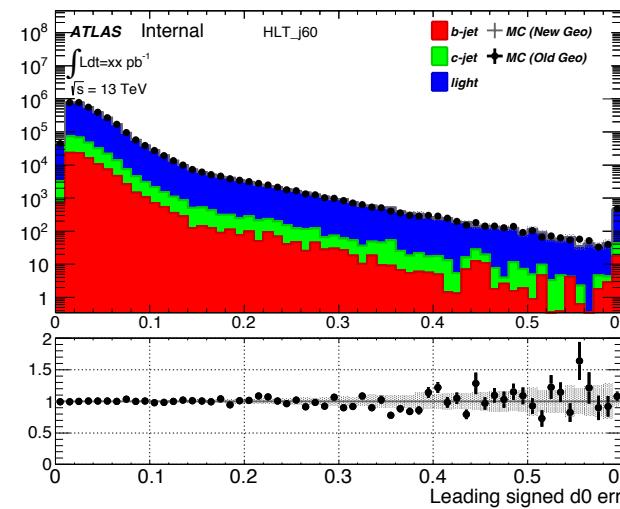
d0 Error:



Data / New Geo MC

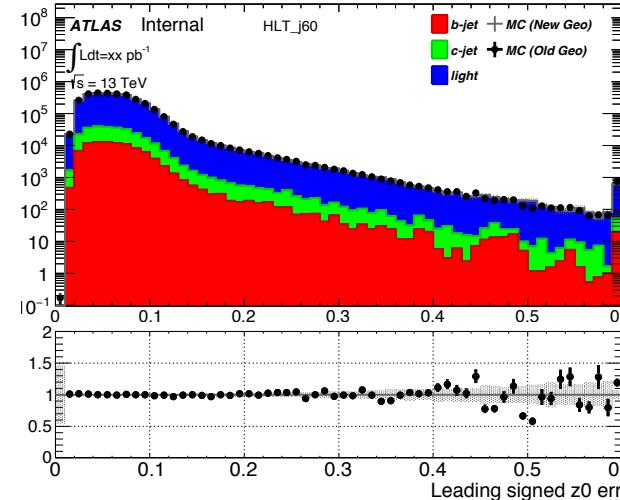
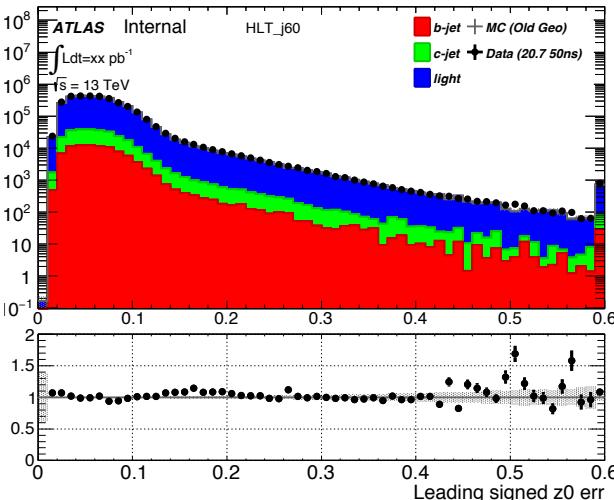
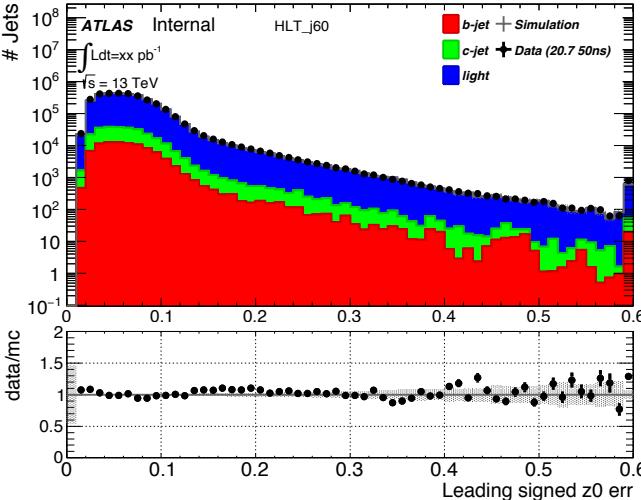


Data / Old Geo MC



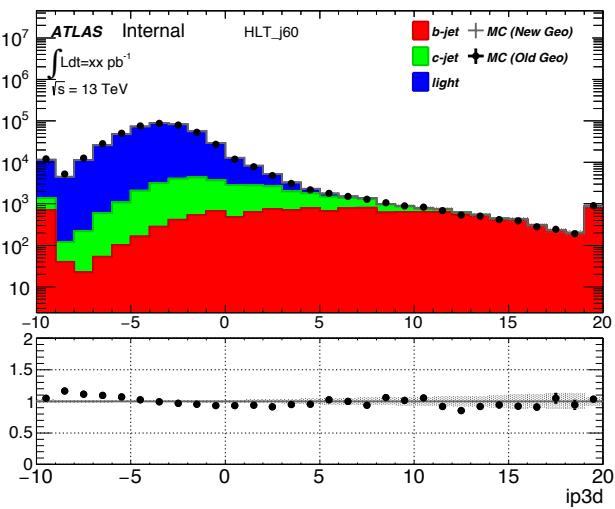
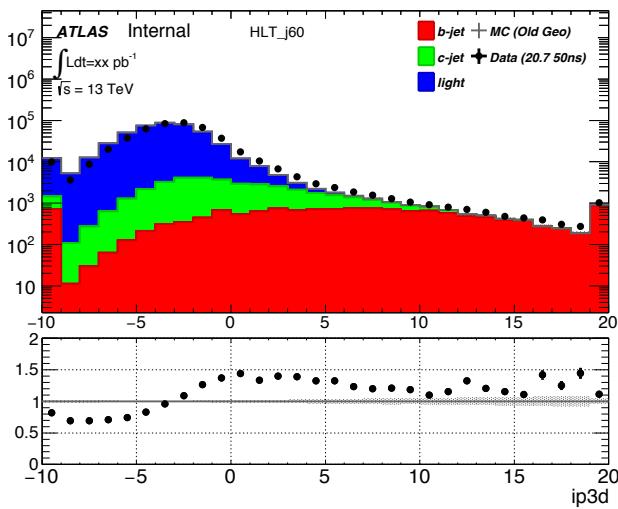
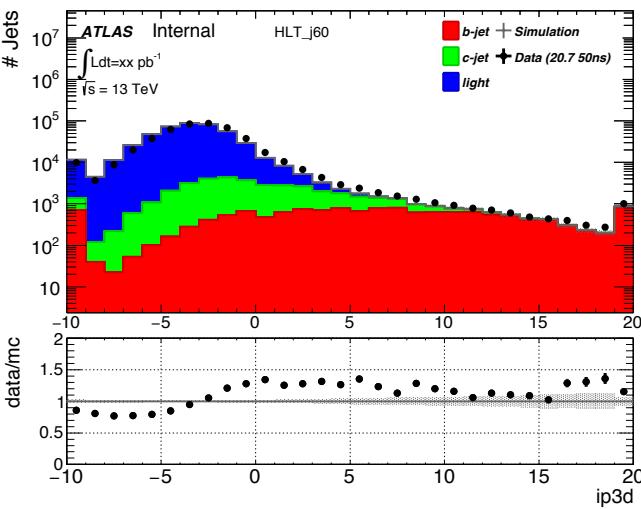
Old Geo MC / New Geo MC

z0 Error:





## IP3D

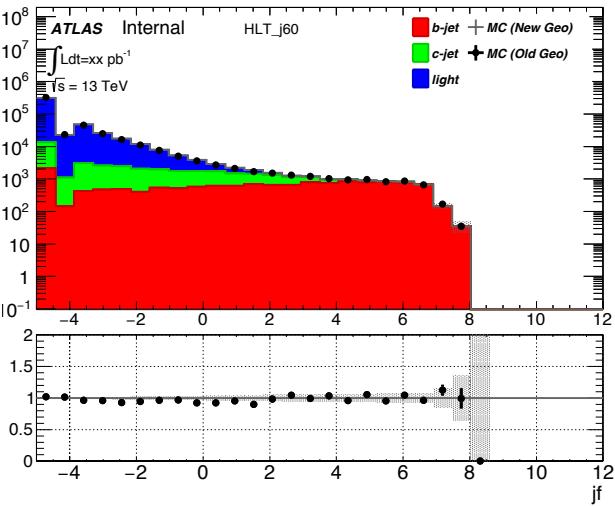
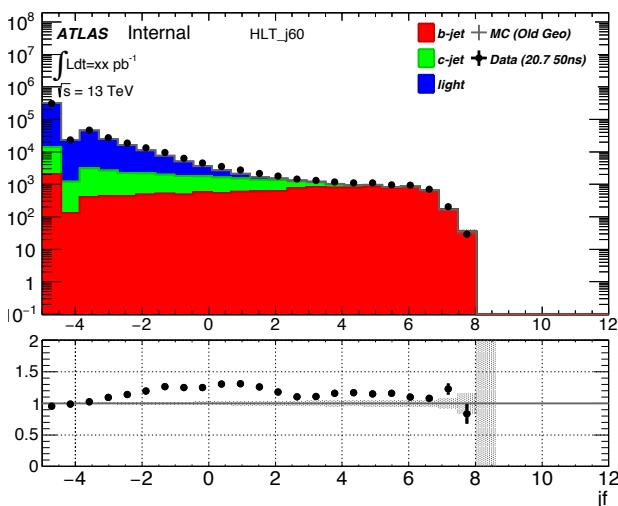
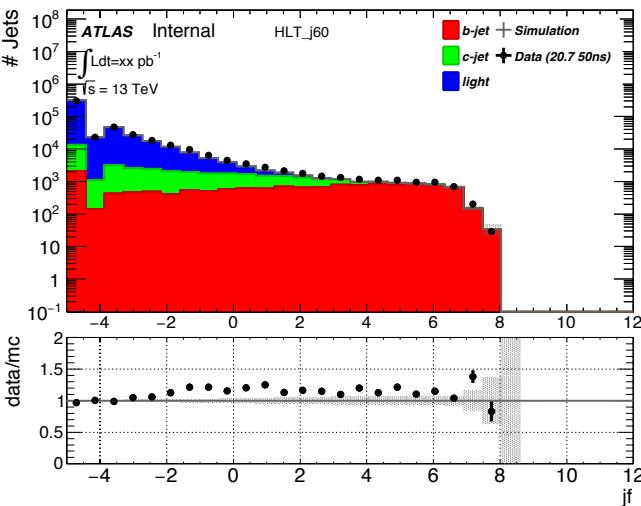


Data / New Geo MC

Data / Old Geo MC

Old Geo MC / New Geo MC

## JF

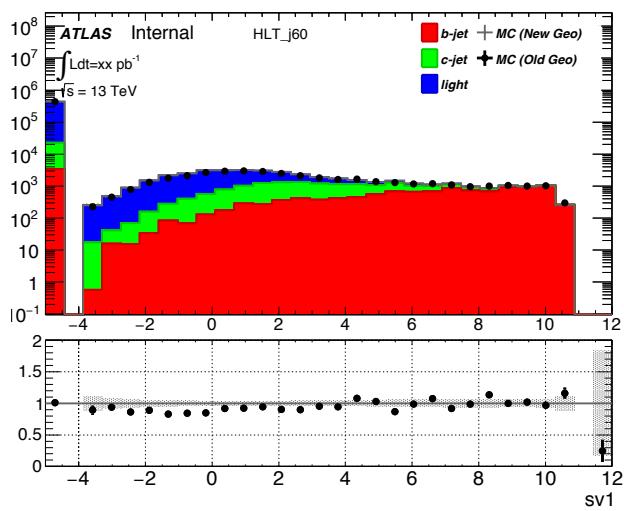
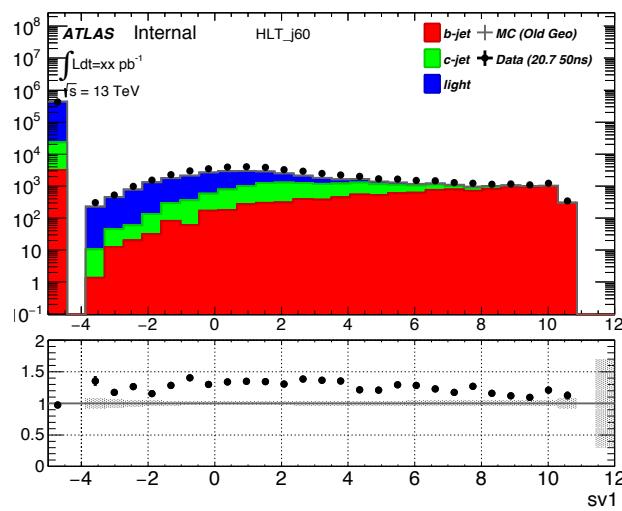
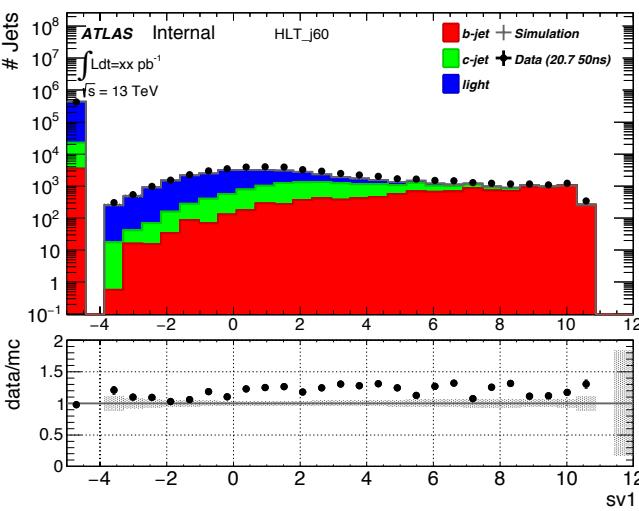




# 11 SV1 / MV2c20

UCL

## SV1

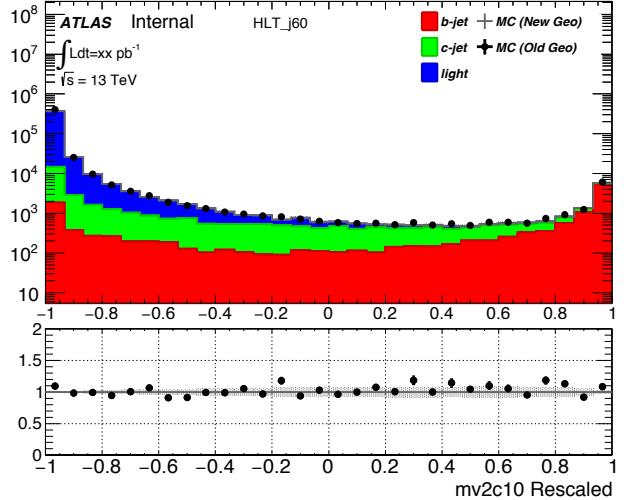
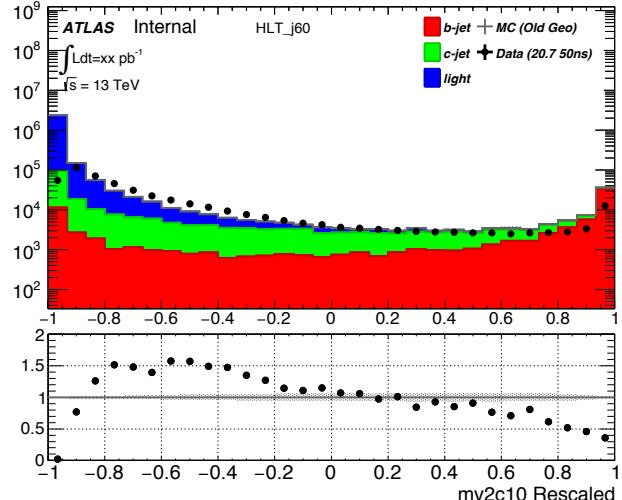
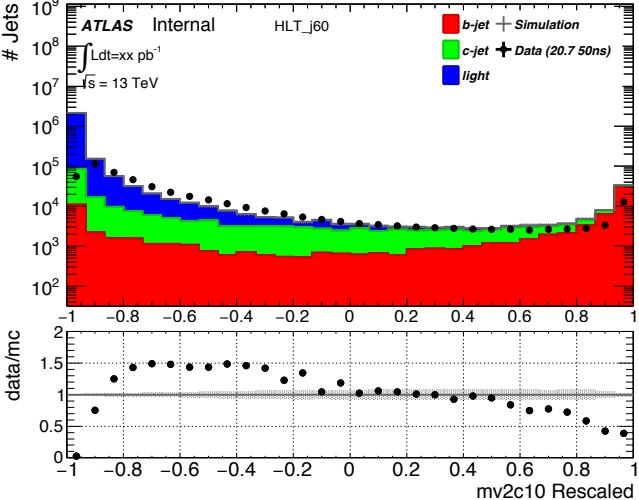


Data / New Geo MC

Data / Old Geo MC

Old Geo MC / New Geo MC

## MV2c10

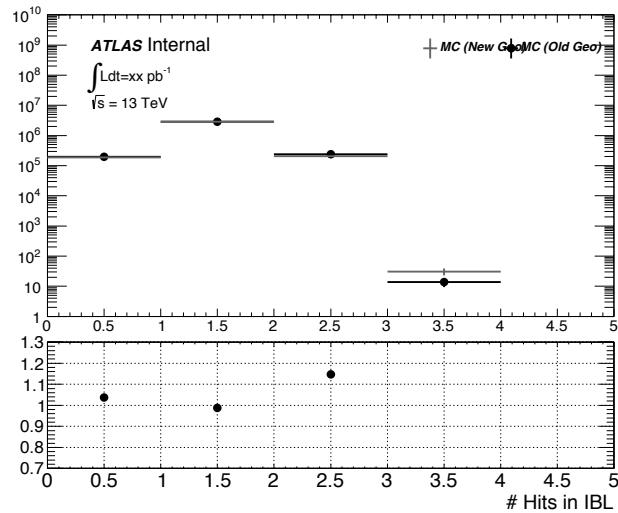
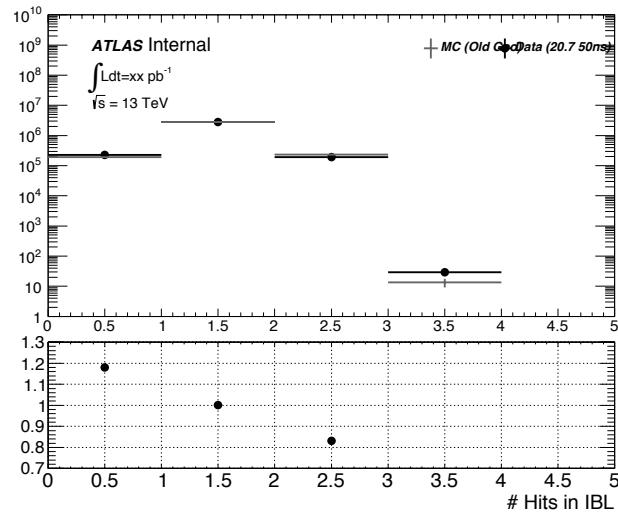
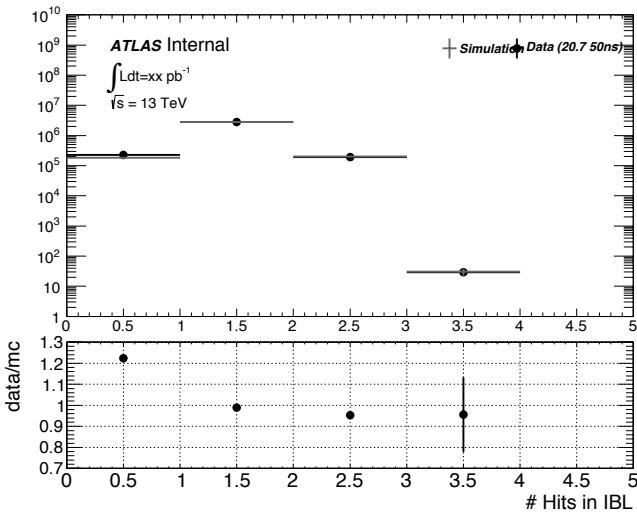




# 12 IBL Studies



## IBL Hits

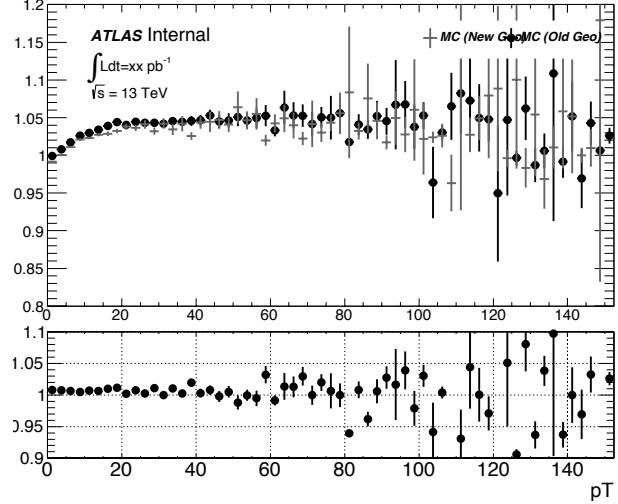
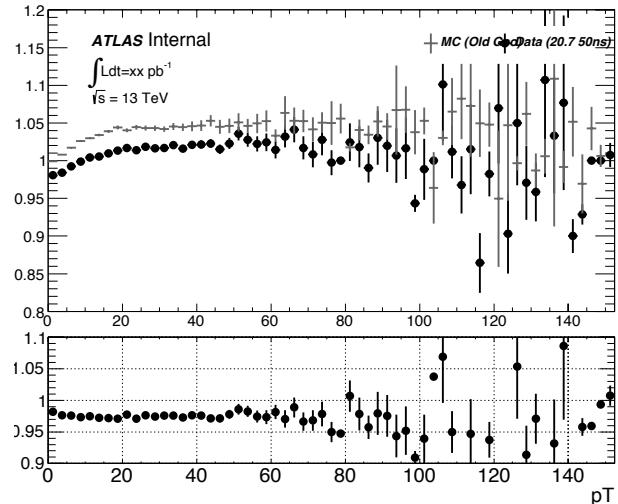
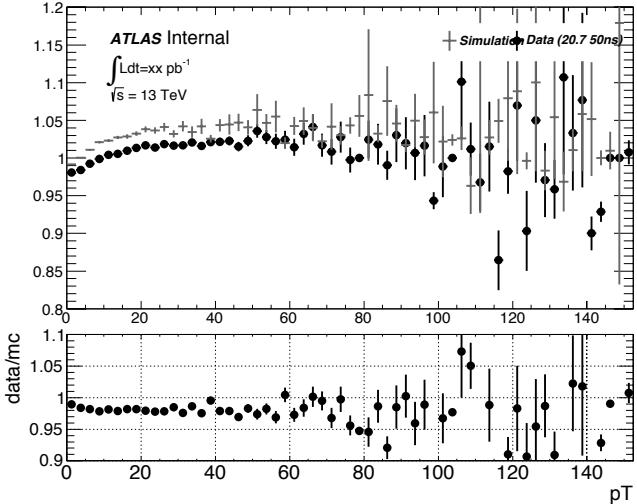


Data / New Geo MC

Data / Old Geo MC

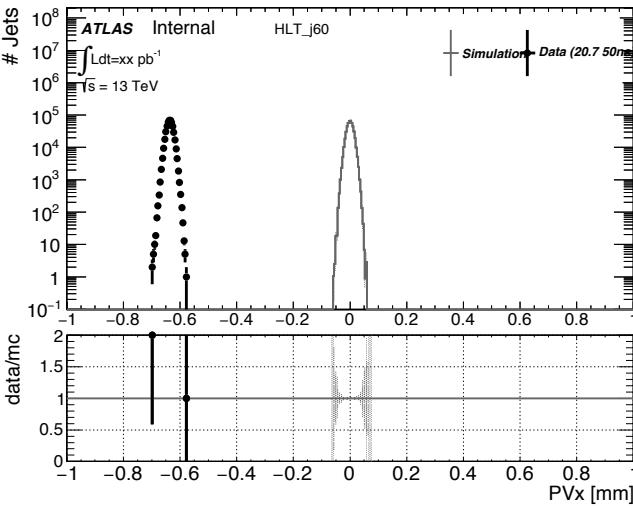
Old Geo MC / New Geo MC

## IBL hits vs track pT

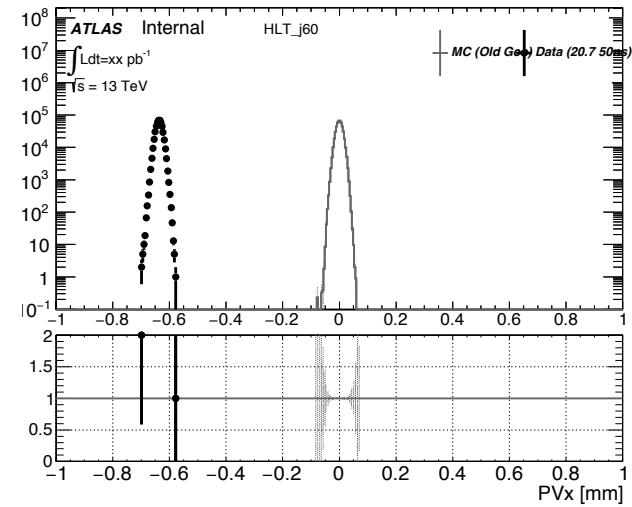




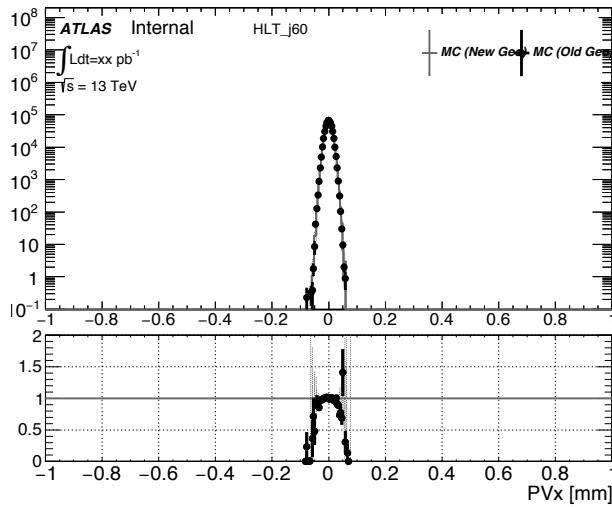
PVx



## Data / New Geo MC

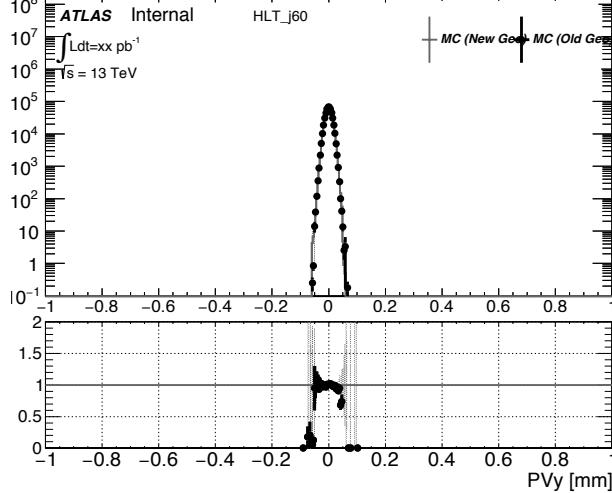
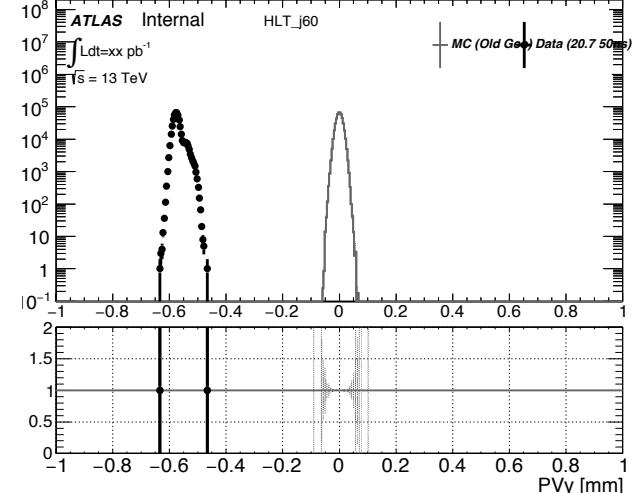
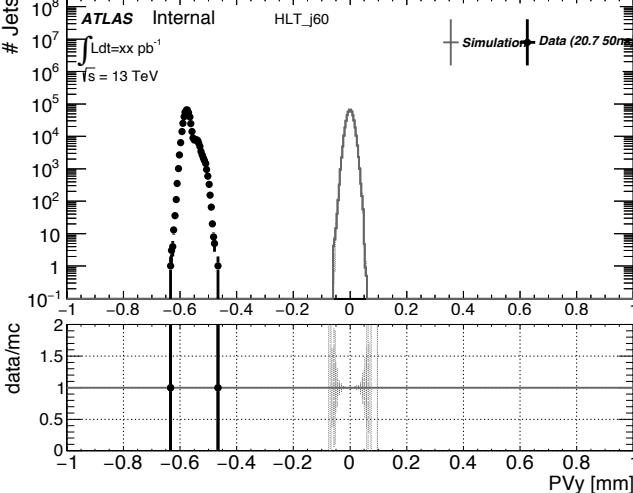


Data / Old Geo MC



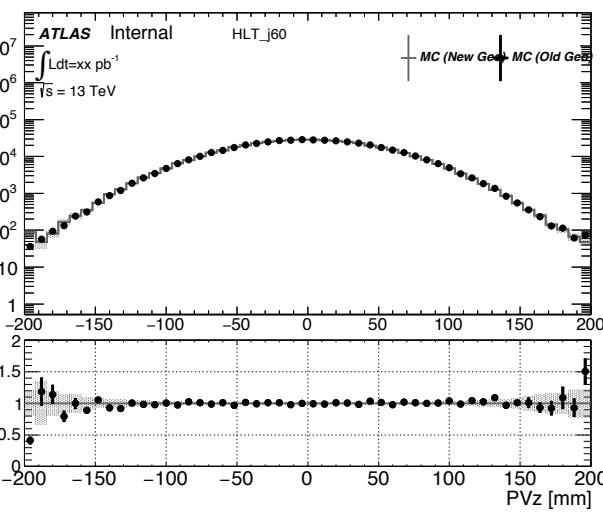
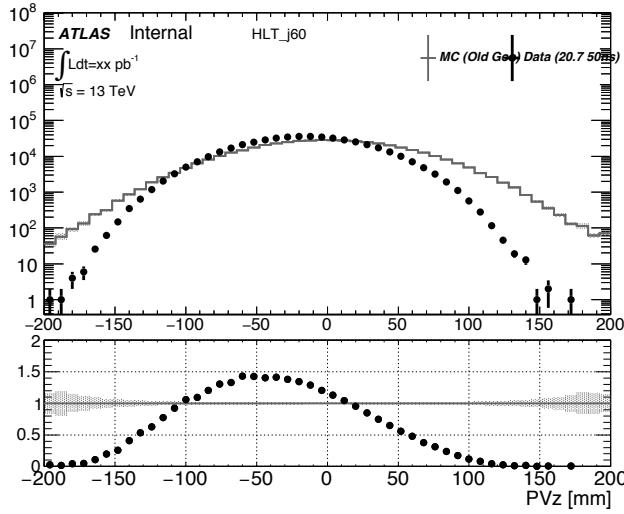
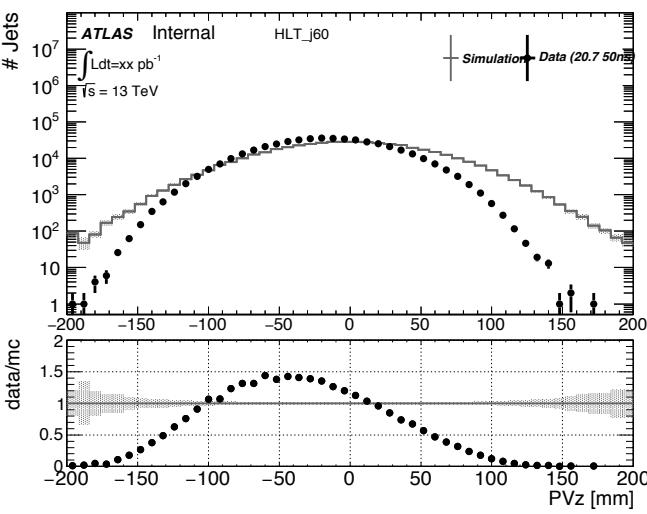
## Old Geo MC / New Geo MC

PVy





## PVz

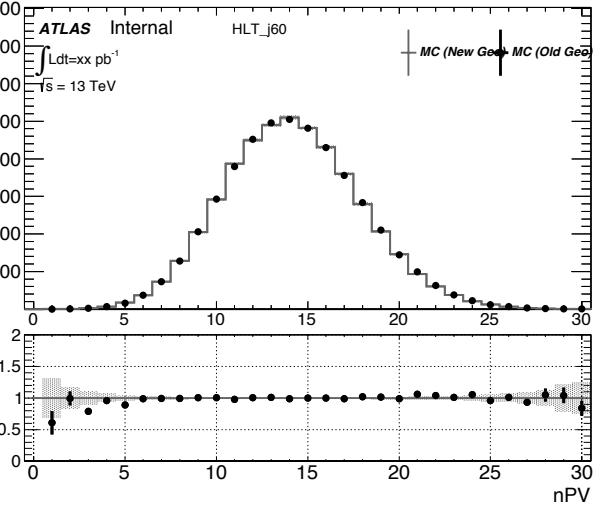
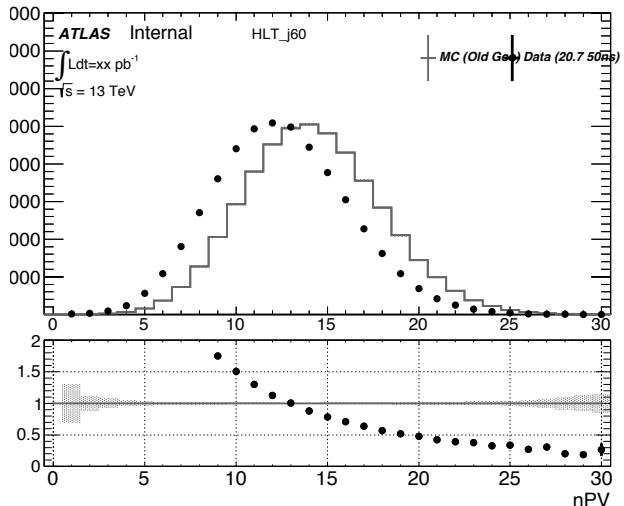
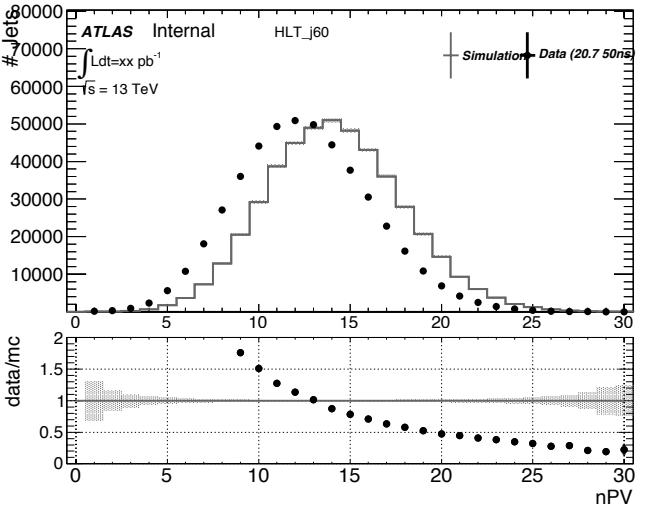


Data / New Geo MC

Data / Old Geo MC

Old Geo MC / New Geo MC

## nPV





## 15 Cutflows



Data 20p7 V23 - 50ns

All	5.53e+06		
GRL	5.252e+06	94.97%	94.97%
Trigger	9.448e+05	17.99%	17.08%
nJets	9.448e+05	100%	17.08%
jetPt	4.77e+05	50.49%	8.627%
eta	4.77e+05	100%	8.627%
badMedium	4.741e+05	99.38%	8.574%
JVT	4.741e+05	100%	8.574%
2 jets	4.741e+05	100%	8.574%
nPV	4.741e+05	100%	8.574%
Truth PV	4.741e+05	100%	8.574%
Truth pT	4.741e+05	100%	8.574%

Data 20p1 V19/V20 - 50ns

All	4.226e+06		
GRL	4.104e+06	97.12%	97.12%
Trigger	9.457e+05	23.04%	22.38%
nJets	9.457e+05	100%	22.38%
jetPt	4.947e+05	52.31%	11.71%
eta	4.947e+05	100%	11.71%
badMedium	4.919e+05	99.43%	11.64%
JVT	4.919e+05	100%	11.64%
2 jets	4.919e+05	100%	11.64%
nPV	4.919e+05	100%	11.64%
Truth PV	4.919e+05	100%	11.64%
Truth pT	4.919e+05	100%	11.64%

Data 20p7 V23 - 25ns

All	1.781e+06		
GRL	1.772e+06	99.48%	99.48%
Trigger	9.627e+04	5.433%	5.405%
nJets	9.627e+04	100%	5.405%
jetPt	4.655e+04	48.35%	2.613%
eta	4.655e+04	100%	2.613%
badMedium	4.622e+04	99.29%	2.595%
JVT	4.622e+04	100%	2.595%
2 jets	4.622e+04	100%	2.595%
nPV	4.622e+04	100%	2.595%
Truth PV	4.622e+04	100%	2.595%
Truth pT	4.622e+04	100%	2.595%



## 16 Cutflows

### MC New Geo 20p7 V23

All	1.929e+10		
GRL	1.929e+10	100%	100%
Trigger	2.687e+08	1.393%	1.393%
nJets	2.686e+08	99.96%	1.393%
jetPt	1.538e+08	57.25%	0.7973%
eta	1.374e+08	89.38%	0.7126%
badMedium	1.366e+08	99.4%	0.7083%
JVT	1.366e+08	100%	0.7083%
2 jets	1.366e+08	100%	0.7083%
nPV	1.366e+08	100%	0.7083%
Truth PV	1.366e+08	100%	0.7083%
Truth pT	8.911e+07	65.23%	0.462%

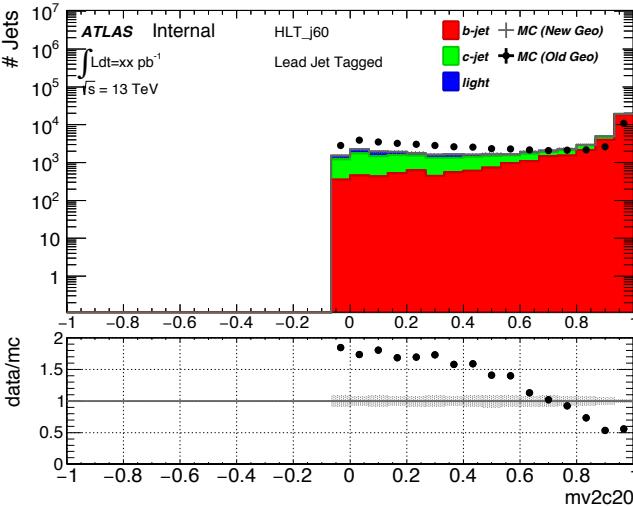
### MC Old Geo 20p7 V23

All	1.931e+10		
GRL	1.931e+10	100%	100%
Trigger	2.732e+08	1.415%	1.415%
nJets	2.732e+08	99.99%	1.415%
jetPt	1.585e+08	58.02%	0.8209%
eta	1.417e+08	89.39%	0.7338%
badMedium	1.41e+08	99.51%	0.7302%
JVT	1.41e+08	100%	0.7302%
2 jets	1.41e+08	100%	0.7302%
nPV	1.41e+08	100%	0.7302%
Truth PV	1.41e+08	100%	0.7302%
Truth pT	8.932e+07	63.36%	0.4626%



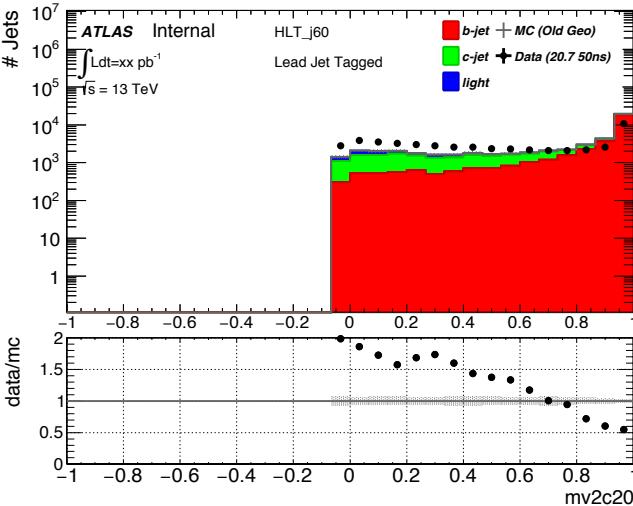
Leading jet  $\text{mv2c20} > -0.0436$ , study subleading jet

data/MC new geo

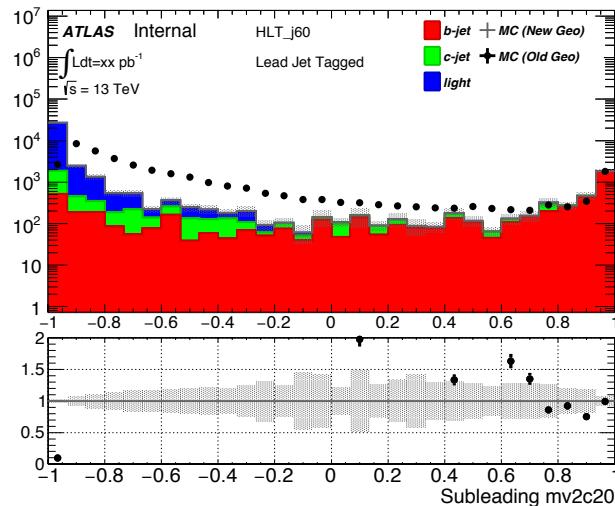


Leading Jet mv2c20

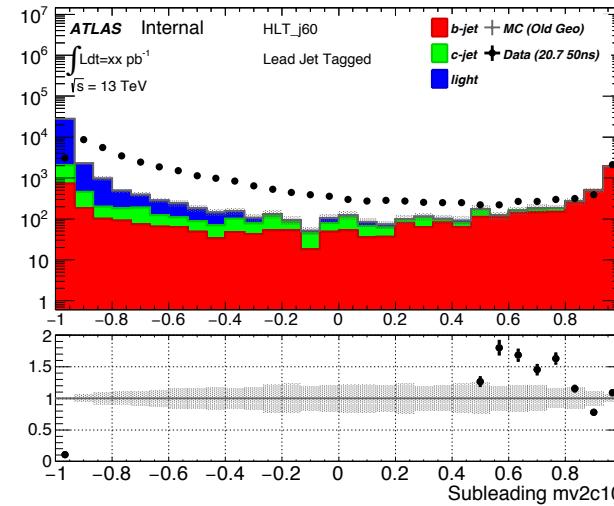
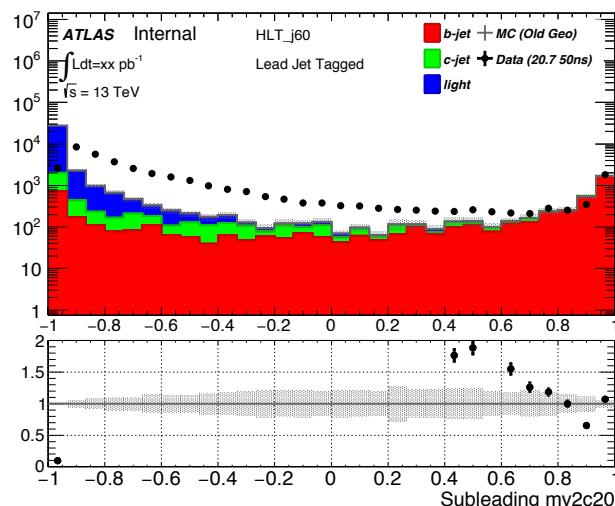
data/MC old geo



Subleading Jet mv2c20



Subleading jet mv2c10



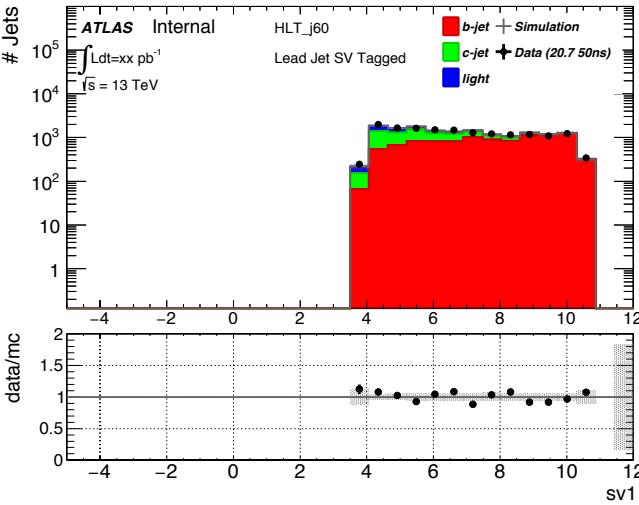


# 18 Leading SV1 Tagged

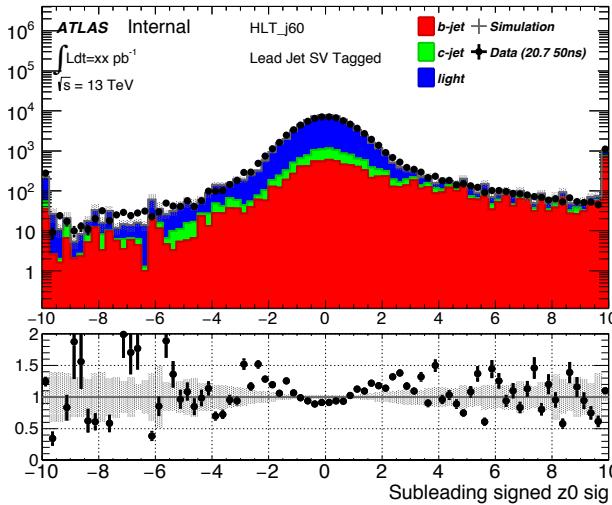


Data/New MC, Leading Jet  $SV1 > 4$ , study subleading, new Geo Only

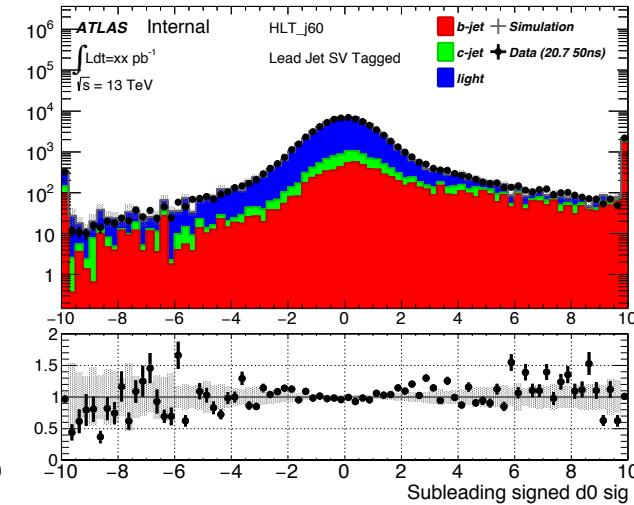
## Leading Jet SV1



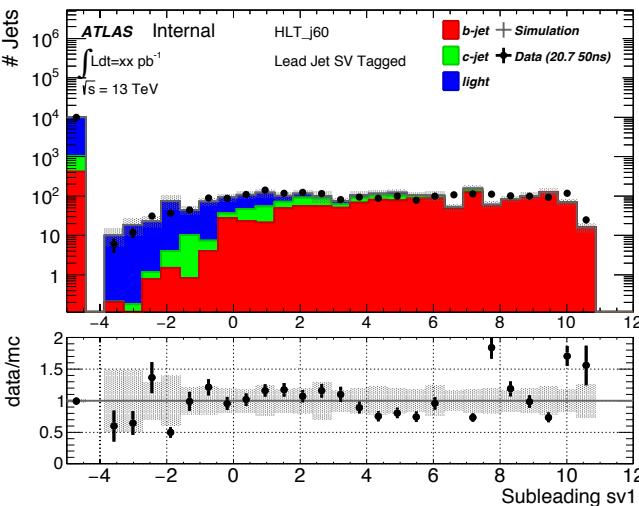
## Subleading z0 sig



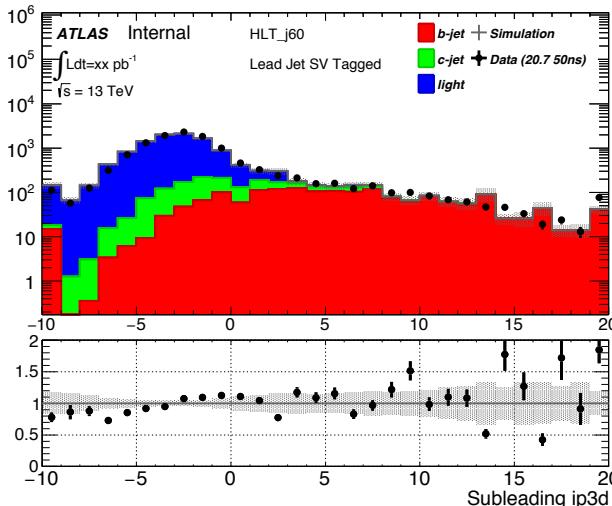
## Subleading d0 sig



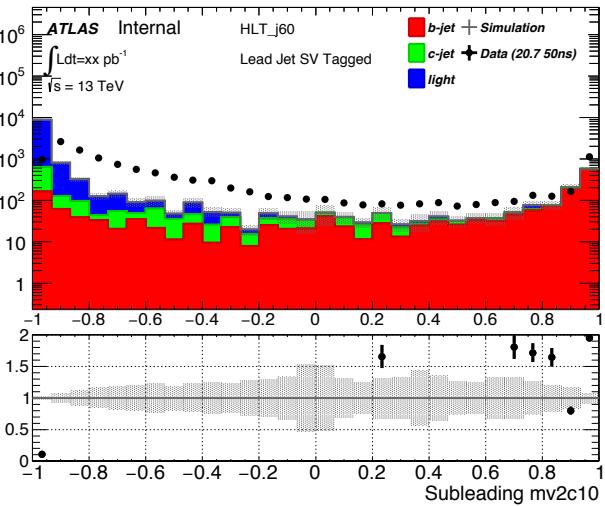
## Subleading sv1



## subleading ip3d



## subleading mv2c10

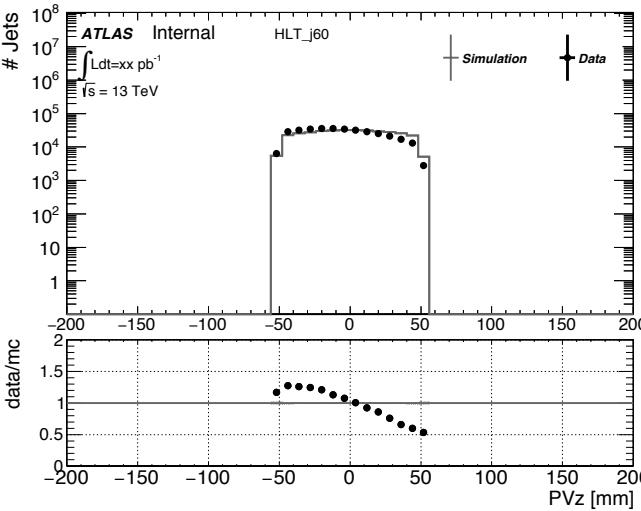




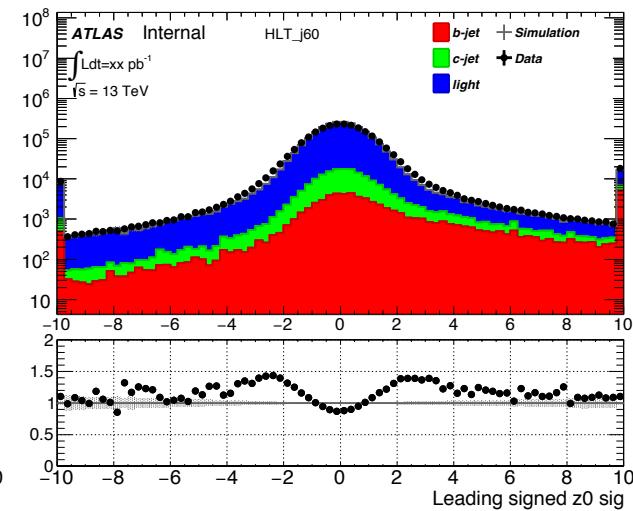
# 19 PVZ Restriction

Data/New MC for  $|PVz| < 50$  mm

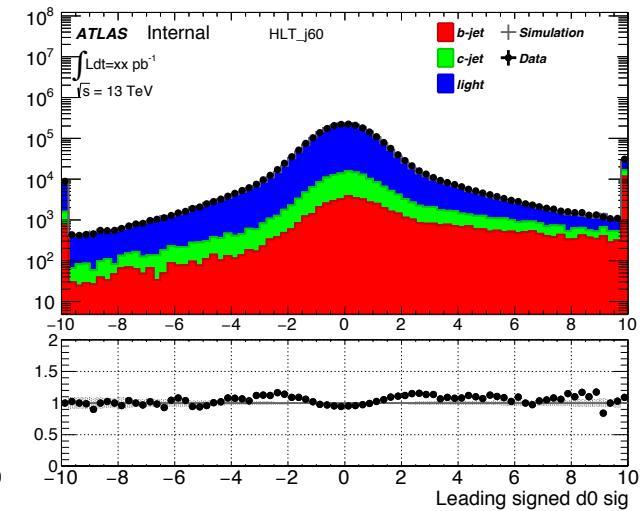
## PVz Distribution



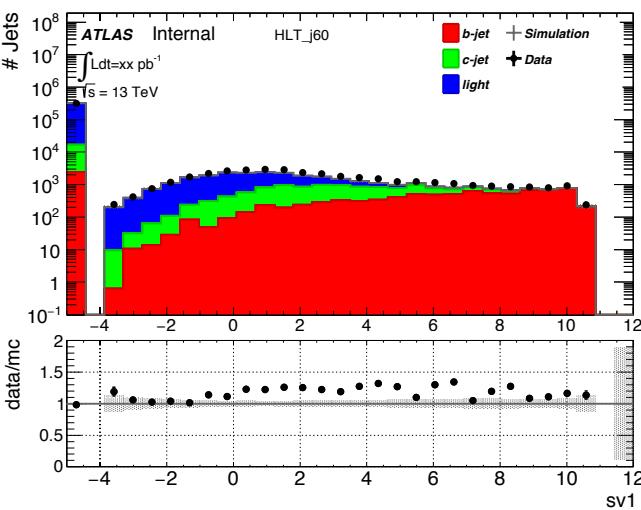
## $z_0$ sig



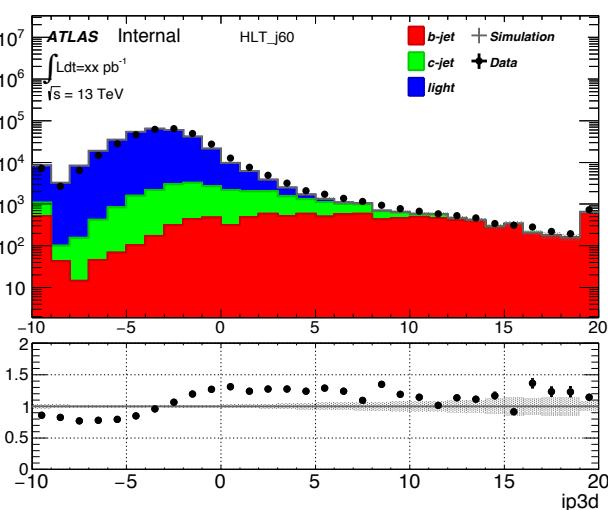
## $d_0$ sig



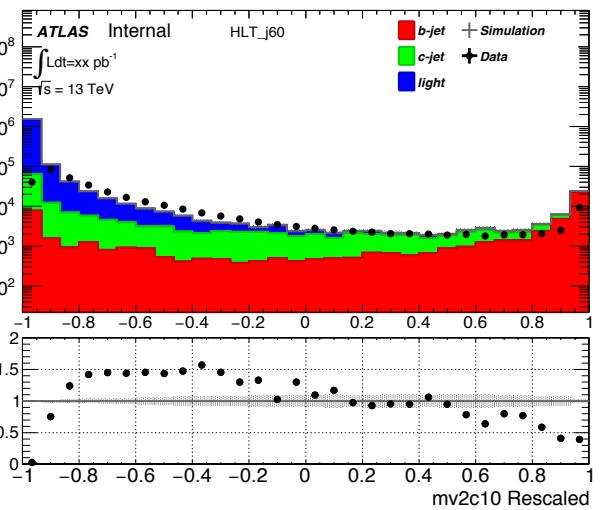
## sv1



## ip3d



## mv2c10

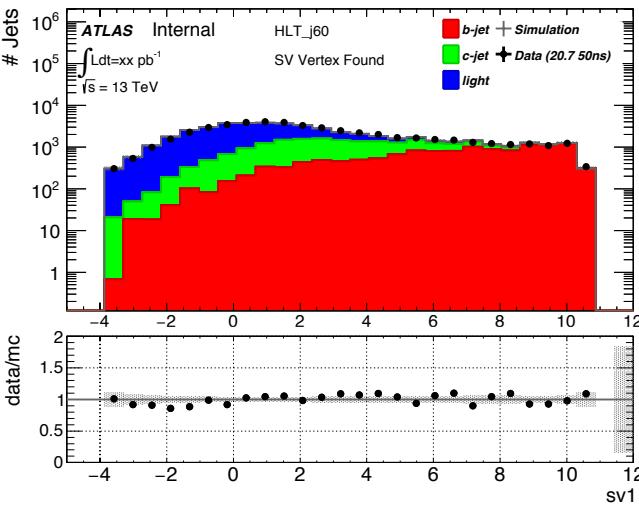




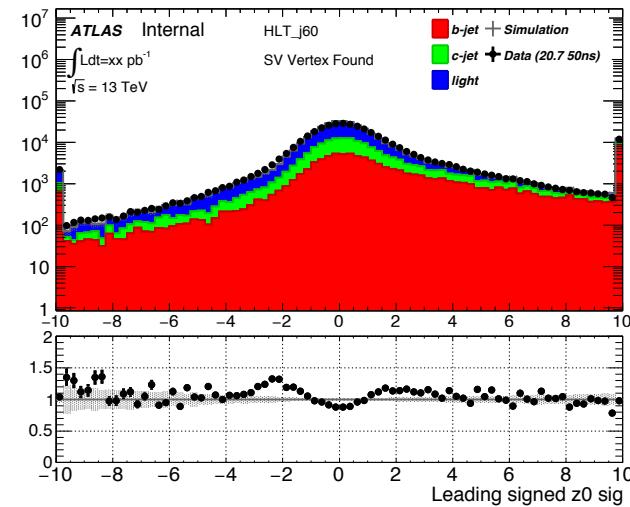
## 20 SV1 Vertex Found

### SV Vertex Found, Data/New Geo MC

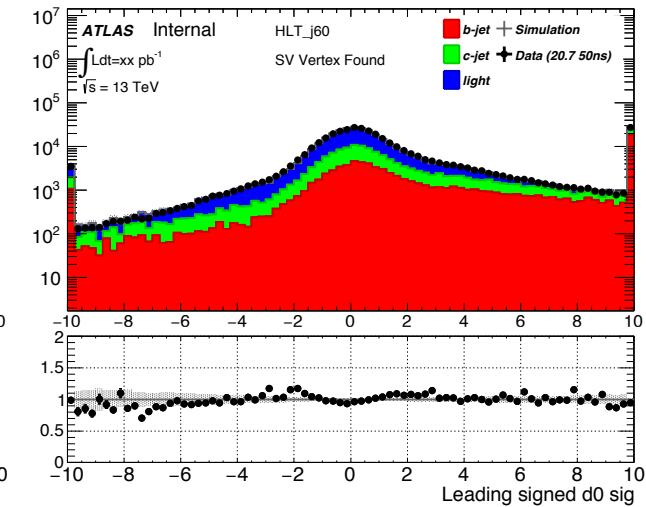
SV1



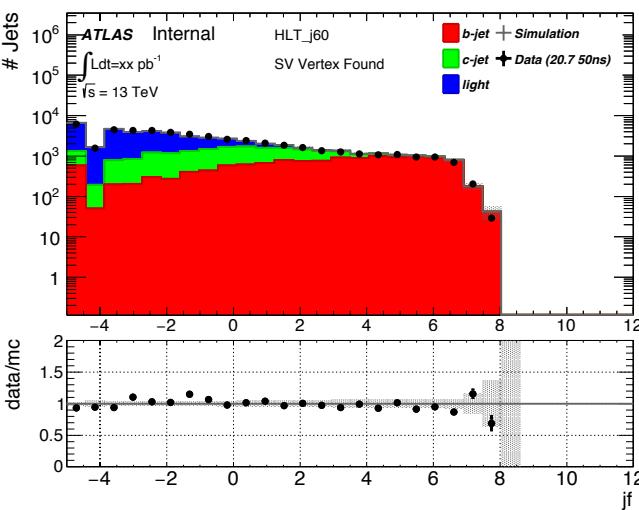
z0 sig



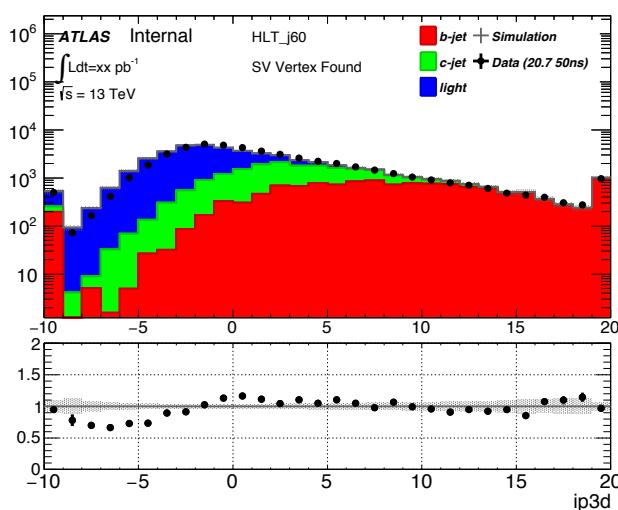
d0 sig



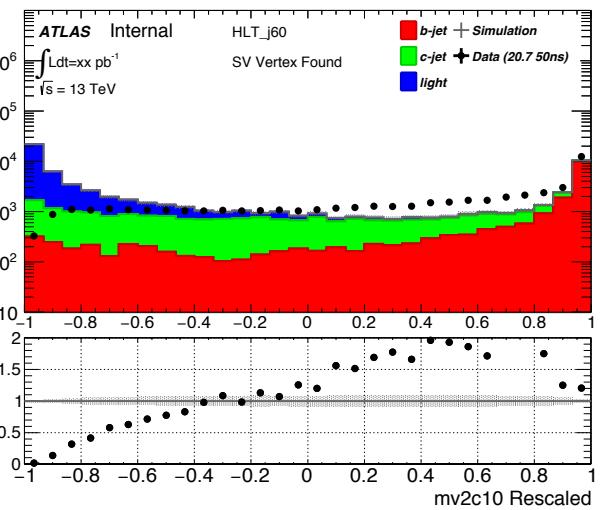
JF



IP3D



mv2c10





## 21 Flavour Fractions

### MC New Geo 20p7 V23

Jet?	# b	# c	# u		b frac	c frac	u frac
Lead	1.45e+04	2.93e+04	4.3e+05		3.06%	6.18%	90.8%
SubL	1.14e+04	2.18e+04	3.12e+05		3.31%	6.31%	90.4%

### MC Old Geo 20p7 V23

Jet?	# b	# c	# u		b frac	c frac	u frac
Lead	1.43e+04	2.99e+04	4.3e+05		3.02%	6.3%	90.7%
SubL	1.13e+04	2.14e+04	3.12e+05		3.28%	6.19%	90.5%

### MC Old Geo 20p7 V23 - mv2c20 tagged (70%) LJ

Jet?	# b	# c	# u		b frac	c frac	u frac
Lead	3.46e+04	1.27e+04	3.73e+03		67.8%	24.9%	7.31%
SubL	5.54e+03	2.89e+03	2.93e+04		14.7%	7.65%	77.7%

### MC Old Geo 20p7 V23 - SV Tagged (> 4) LJ

Jet?	# b	# c	# u		b frac	c frac	u frac
Lead	1.03e+04	3.95e+03	1.76e+03		64.4%	24.6%	11%
SubL	1.71e+03	957	9.47e+03		14.1%	7.88%	78%

### MC Old Geo 20p7 V23 - SV tagged - SV Found for LJ

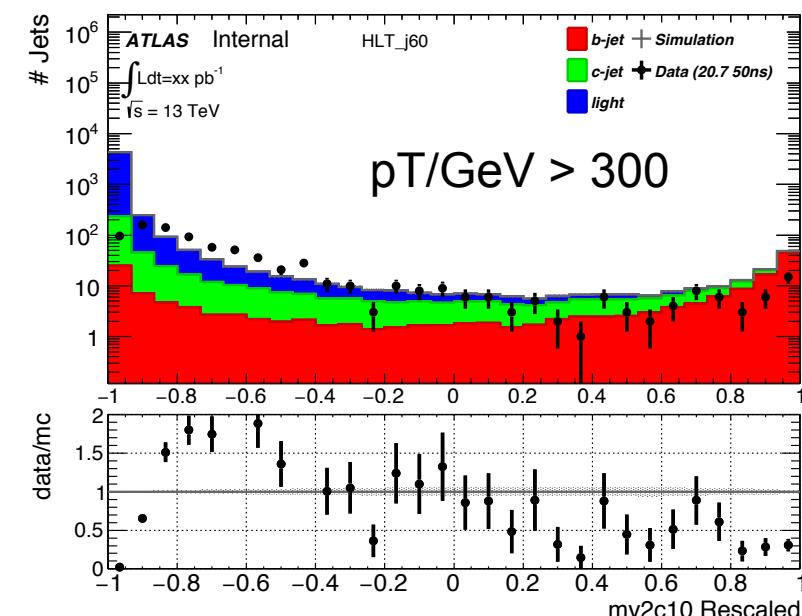
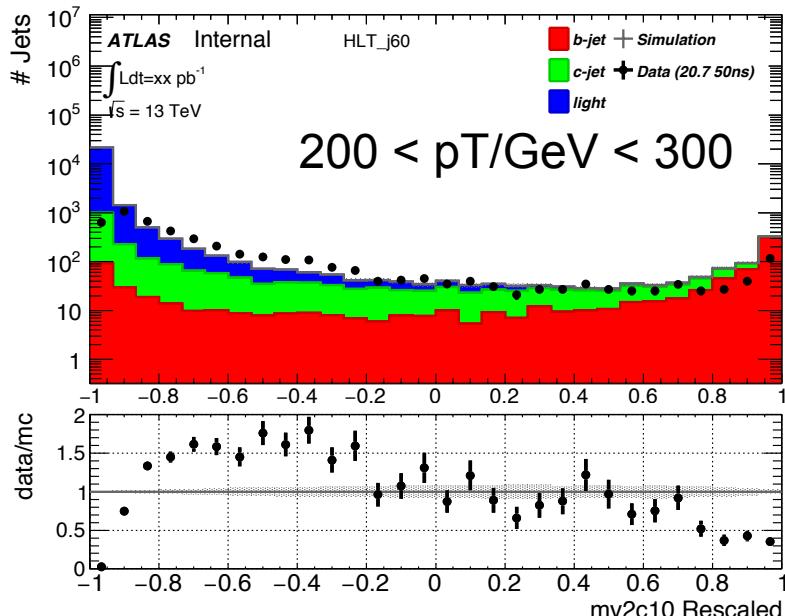
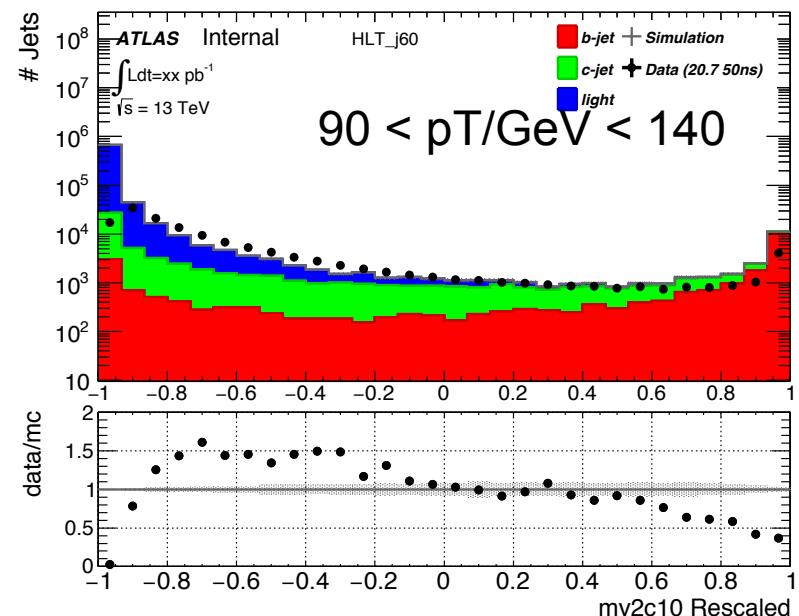
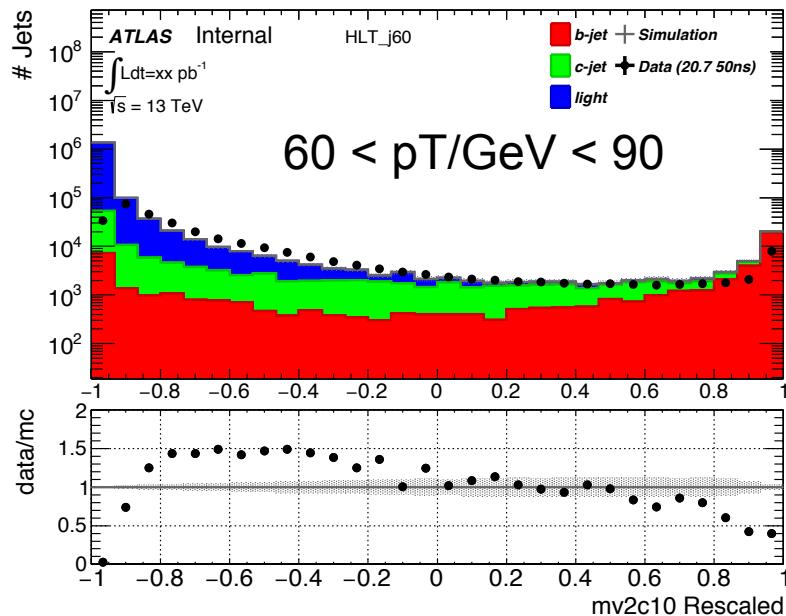
Jet?	# b	# c	# u		b frac	c frac	u frac
Lead	1.33e+04	1.21e+04	2.51e+04		26.3%	24%	49.7%
SubL	2.79e+03	2.99e+03	3.14e+04		7.49%	8.03%	84.5%

### MC Old Geo 20p7 V23 - PVz Cut

Jet?	# b	# c	# u		b frac	c frac	u frac
Lead	1.07e+04	2.19e+04	3.16e+05		3.07%	6.26%	90.7%
SubL	8.38e+03	1.59e+04	2.3e+05		3.3%	6.27%	90.4%



## 22 mv2c10 in pT bins





## Conclusions

Changes comparing new geometry to old geometry...

### Improvements!

- d0 significance modelling is much better
- # IBL hits somewhat modelled better (overlap of IBL issue)
- ip3d, sv1 and jf appear to show a small improvement

### No real change

- z0sig is still mis-modelled (although small improvements)

### Deteriorations

- mv2c10
  - This is totally wrong
  - Seen elsewhere: Slide 5 of [this talk](#)
  - Is mv2c10 mis-modelled for light jets

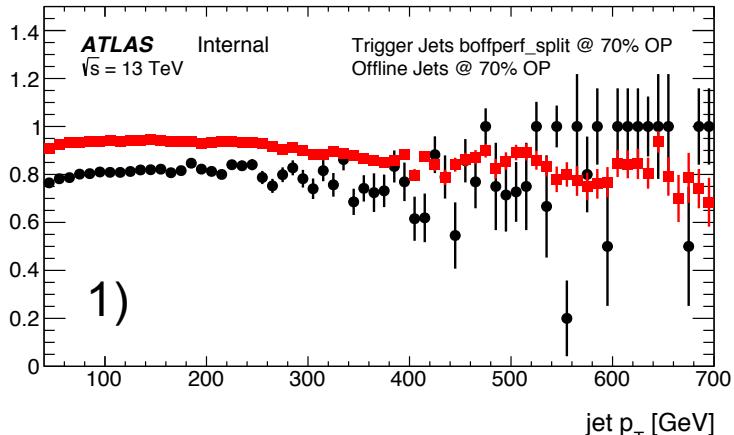


# Backup

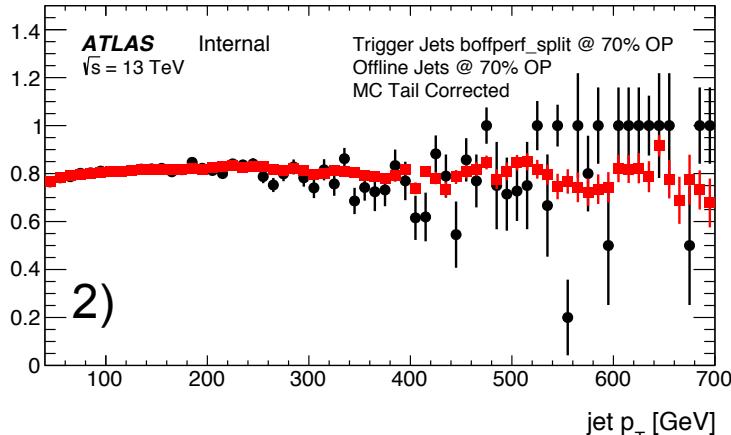


## 25 bTrigger Plots

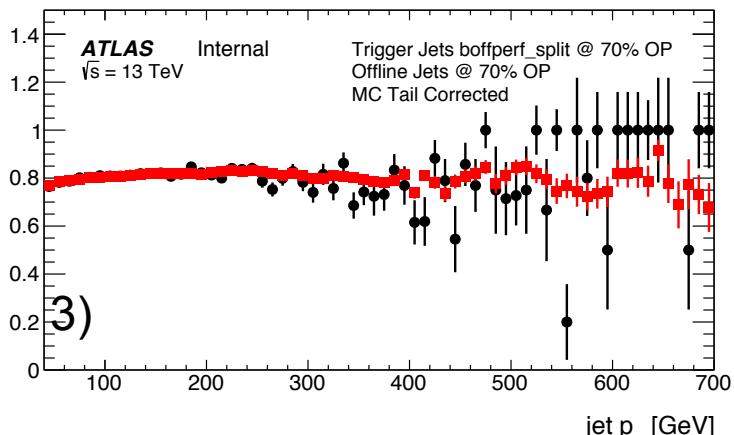
Efficiency



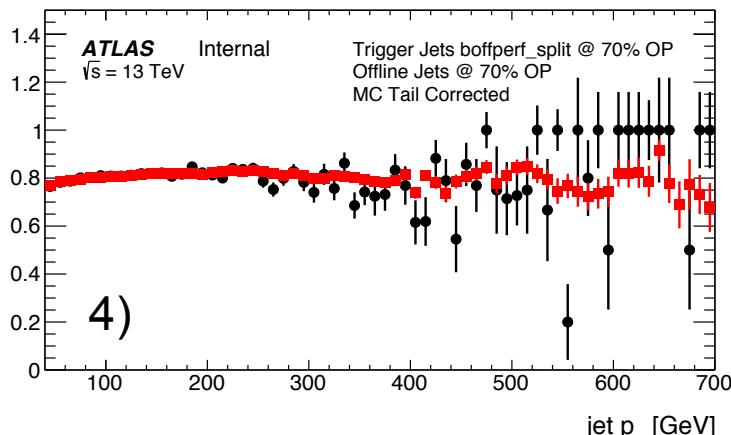
Efficiency



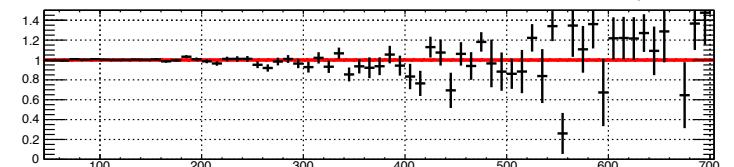
Efficiency



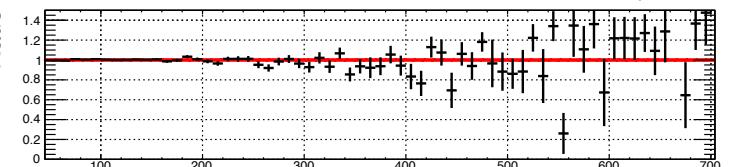
Efficiency

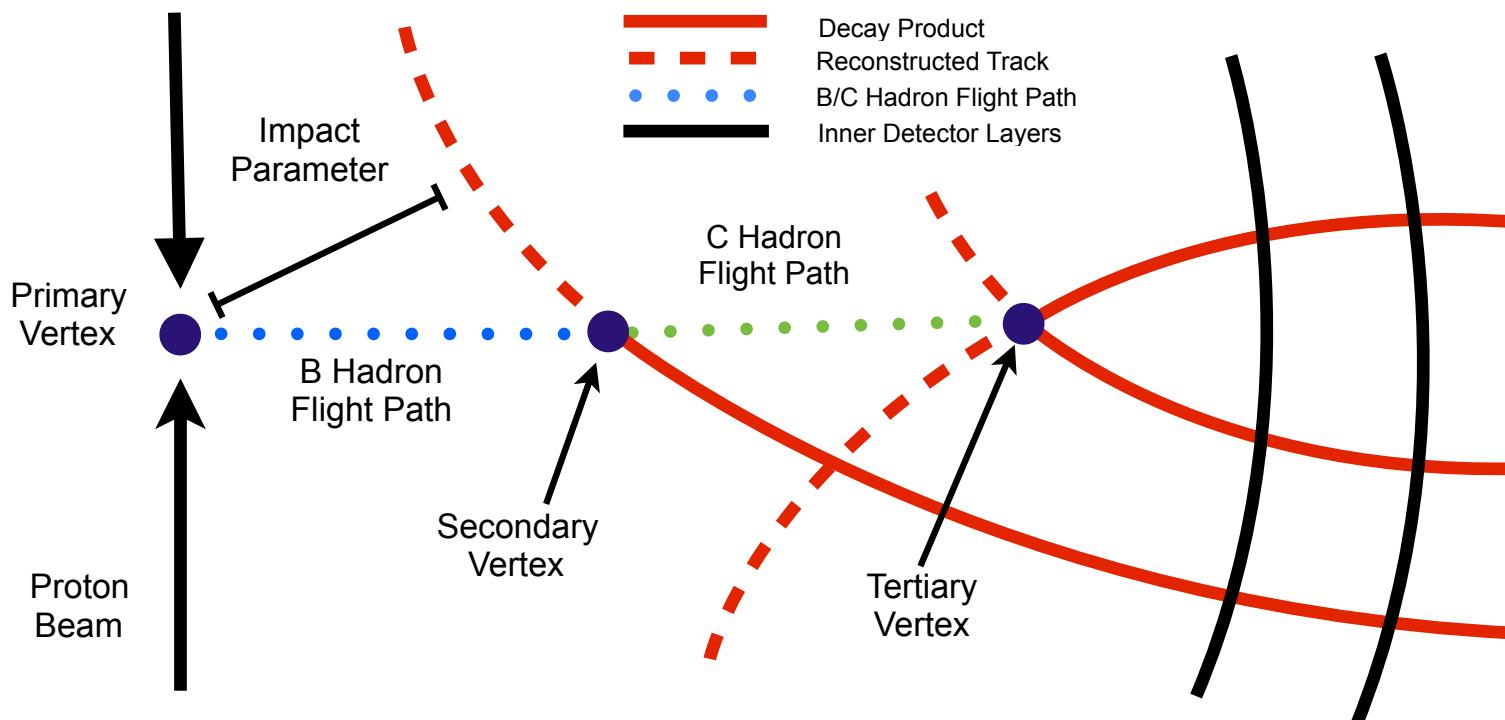


Ratio



Ratio





- IP2D and IP3D: Use impact parameter distributions to discriminate between flavours
- SV1: Search for Secondary Vertex from crossing of tracks
- Jet Fitter: Reconstructs full decay chain by searching for many vertices along a shared jet flight axis
- MV2c20: Base algorithms are combined in a BDT to give optimal performance.