Development of ATLAS B-taggers based on BDT

Hai-Jun Yang, Xuefei Li & Bing Zhou
University of Michigan, Ann Arbor

ATLAS B-tagging Meeting
May 6, 2009

Outline

- Introduction
- B-jet decays with ‘long’ life-time
- MC samples (ttbar, WH120, WH400)
- B-tagging discriminating variables
- BDT B-taggers (for light-jets, C-jets, τ-jets)
- Performance comparison of ATLAS B-taggers
- Cross checks of ATLAS B-taggers
- Further improvement by combining B-taggers
- Summary

Introduction

- Goal: To develop ATLAS B-taggers using BDT and to evaluate B-tagging performance using various B-tagger classifications (Likelihood, ANN, BDT).

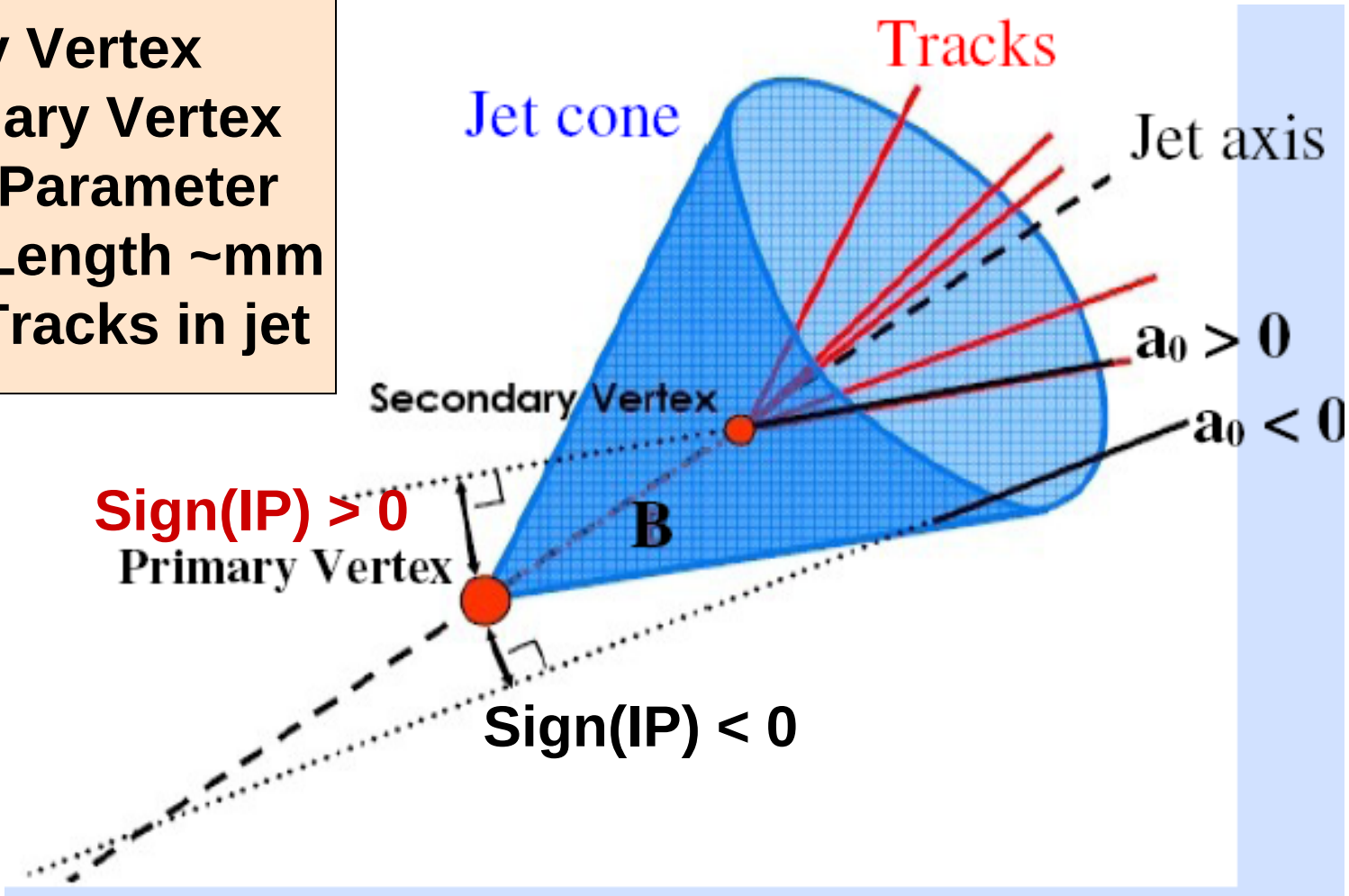
- Our starting point: using the same set of variables as an initial try and to confirm results documented in a previous ATLAS note
  - Ref: J. Bastos, ATL-PHYS-PUB-2007-019
  - H. Yang, talk at ATLAS B-tagging meeting on 02/09/09 based on ATLAS MC V13 samples

- Building a set of new discriminating variables to improve the b-tagging performance.

- Test the b-tagging performance using V14 MC samples

B-jet decays with ‘long’ life-time

- Primary Vertex
- Secondary Vertex
- Impact Parameter
- Decay Length ~mm
- No. of Tracks in jet



MC Data Samples

- Ttbar (DS105200) based on release V14.
- Using WH(120), WH(400 GeV) with H→bb, cc, uu for test
- Preselection cuts: Et (jet) > 15 GeV, $|\eta|<2.5$

Jets	**Training**	**For Performance Test**		
Et>15GeV,$\|\eta\|$<2.5	**Ttbar**	**Ttbar**	**WH120**	**WH400**
B jets	150001	327222	74235	95854
C jets	31549	68610	160120	203140
τ jets	19120	42507	N/A	N/A
Light jets	223332	486742	545190	738951

Input variables for BDT b-tagging algorithm

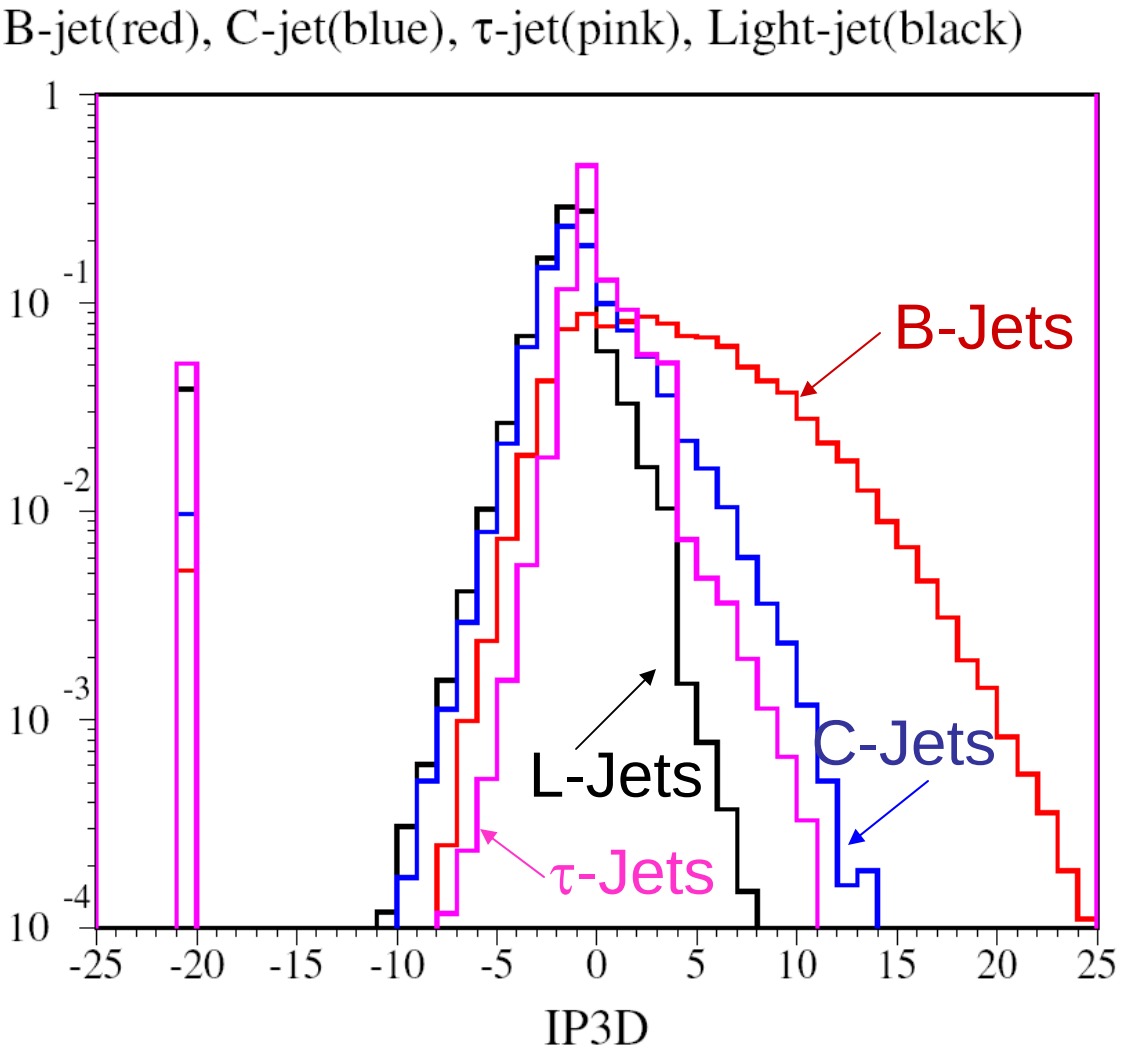
- **Performance of BDT algorithm depends on**
  - **selecting a set of discriminating variables;**
  - **using advanced training process.**

- **Existing variables for an initial try, 11/17 are selected:**
  - IP2D, IP3D, SV1: jet weights from IP and secondary vertices
  - Softe: jet weight from soft electron based tagger
  - jet-mass: mass of particles which participate in the vertex fit
  - Efrac: ratio between the total energy of charged particles in the vertex and the total energy of all particles in the jet
  - d0sig_max, z0sig_max: the largest transverse and longitudinal impact parameter significance of tracks in the jet
  - ptTrk_max: the largest transverse momentum of tracks in the jet
  - Nvertex_2track: Number of two-track vertices
  - Ntrack: Number of tracks in the jet

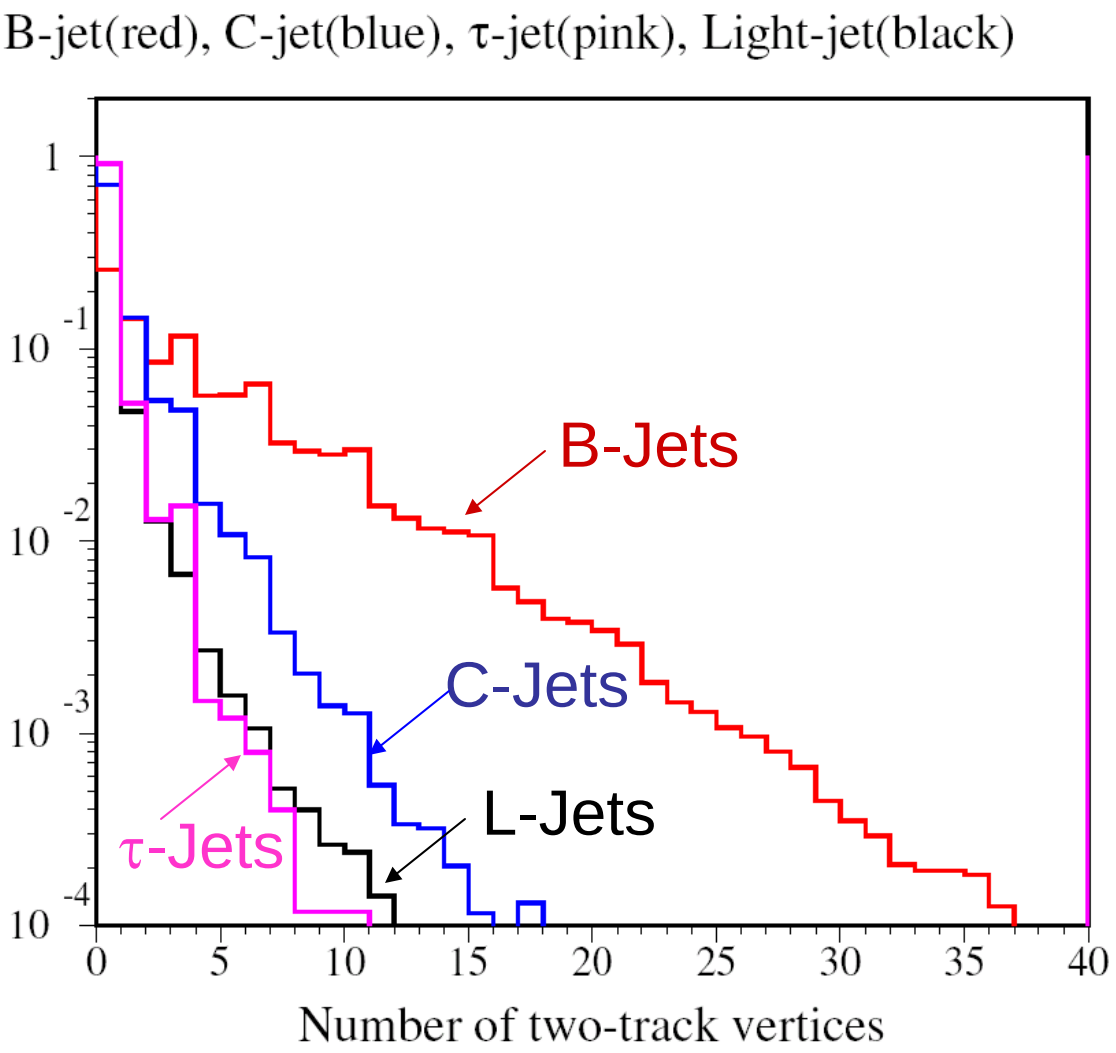
NEW variables for BDT b-tagging algorithm

- **Building additional 27 new variables, 9 / 27 are selected:**
  - Ntrack_distance_150: number of tracks in the jet with distance between PV and track-jet cross point greater than 150 microns
  - Ntrack_z0_100: number of tracks in the jet with longitudinal IP greater than 100 microns
  - Ntrack_z0_05: number of tracks in the jet with longitudinal IP significance greater than 0.5
  - 2d_dl: 2D decay length of the jet (mm)
  - d0sig_avg: average of transverse IP significance from two leading tracks which have the largest d0sig
  - d0_avg: average of transverse IP from two leading tracks
  - d0_avg: average of longitudinal IP from two leading tracks
  - Sumtrkpt_jetE: ratio between sum of track Pt w.r.t jet axis direction and jet energy
  - SumEpt_jetE: ratio between sum of Electron Pt w.r.t jet axis direction and jet energy

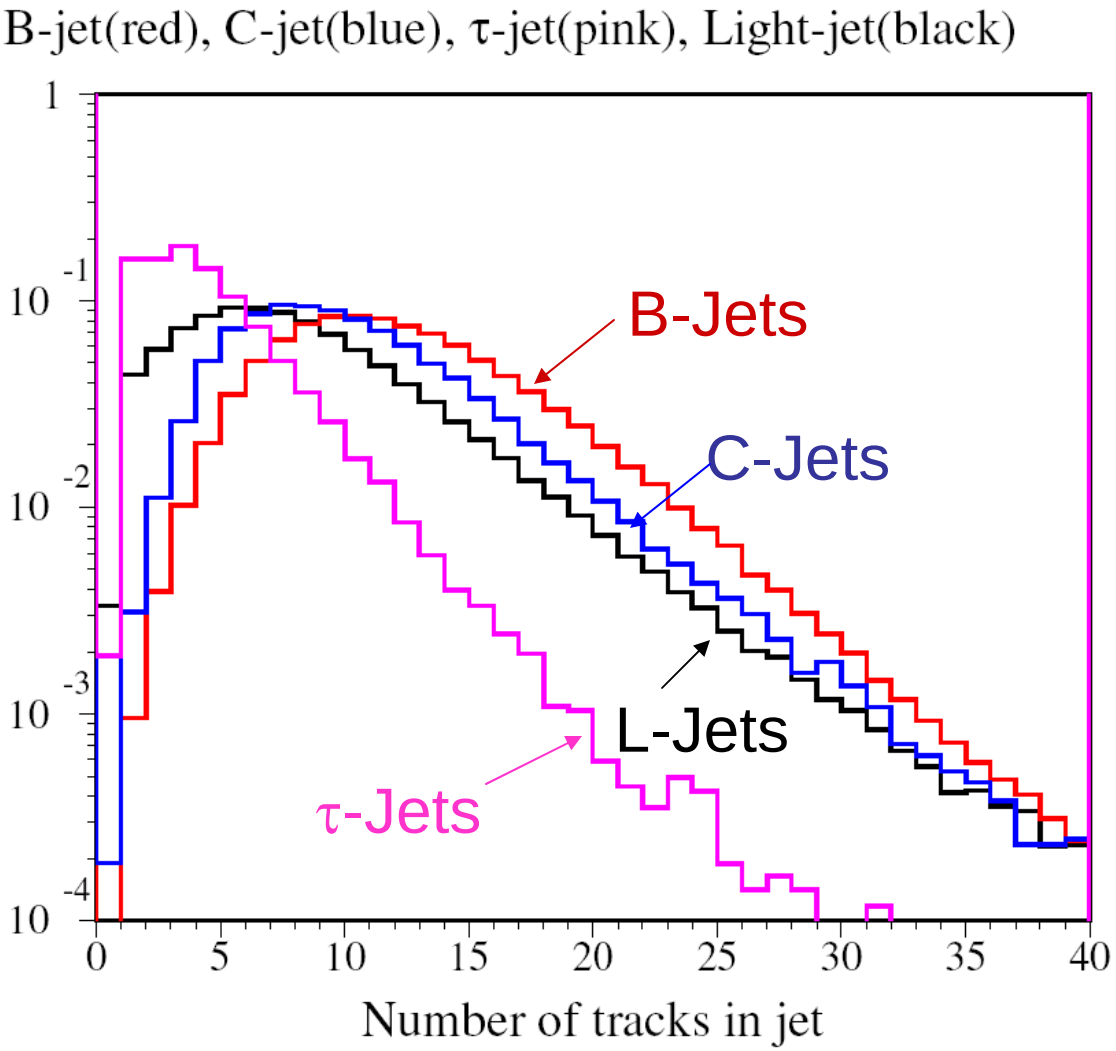
Jet Weight from 3D Impact Parameters



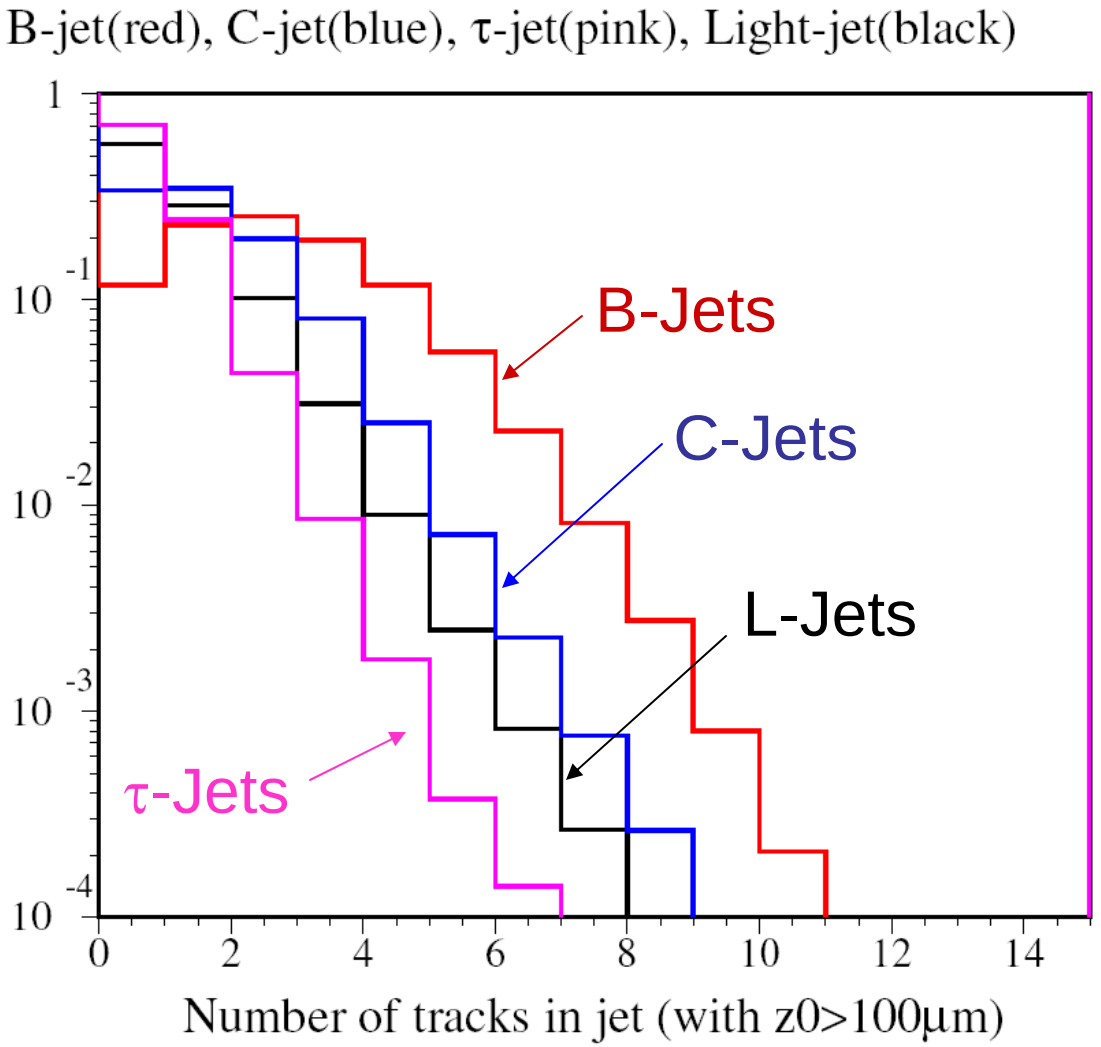
Number of 2-track vertices



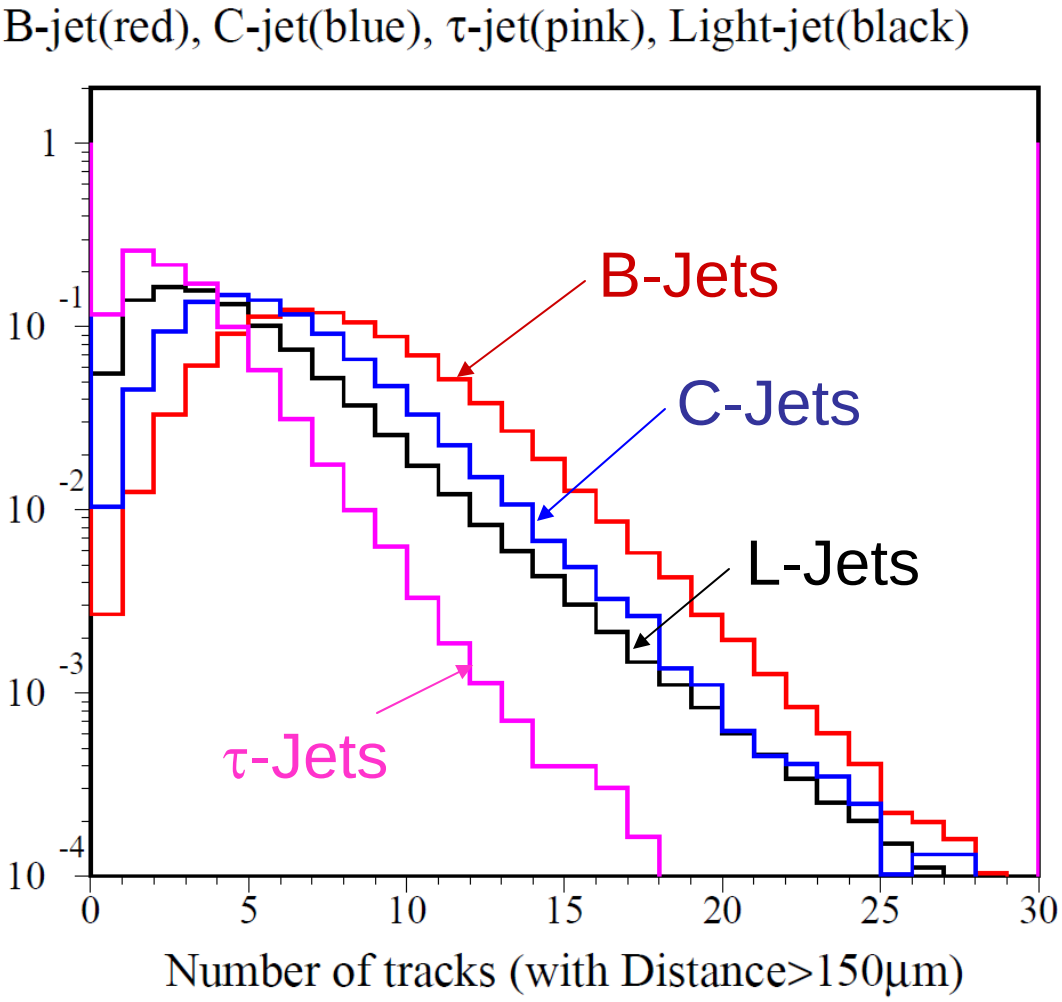
Number of tracks in jet



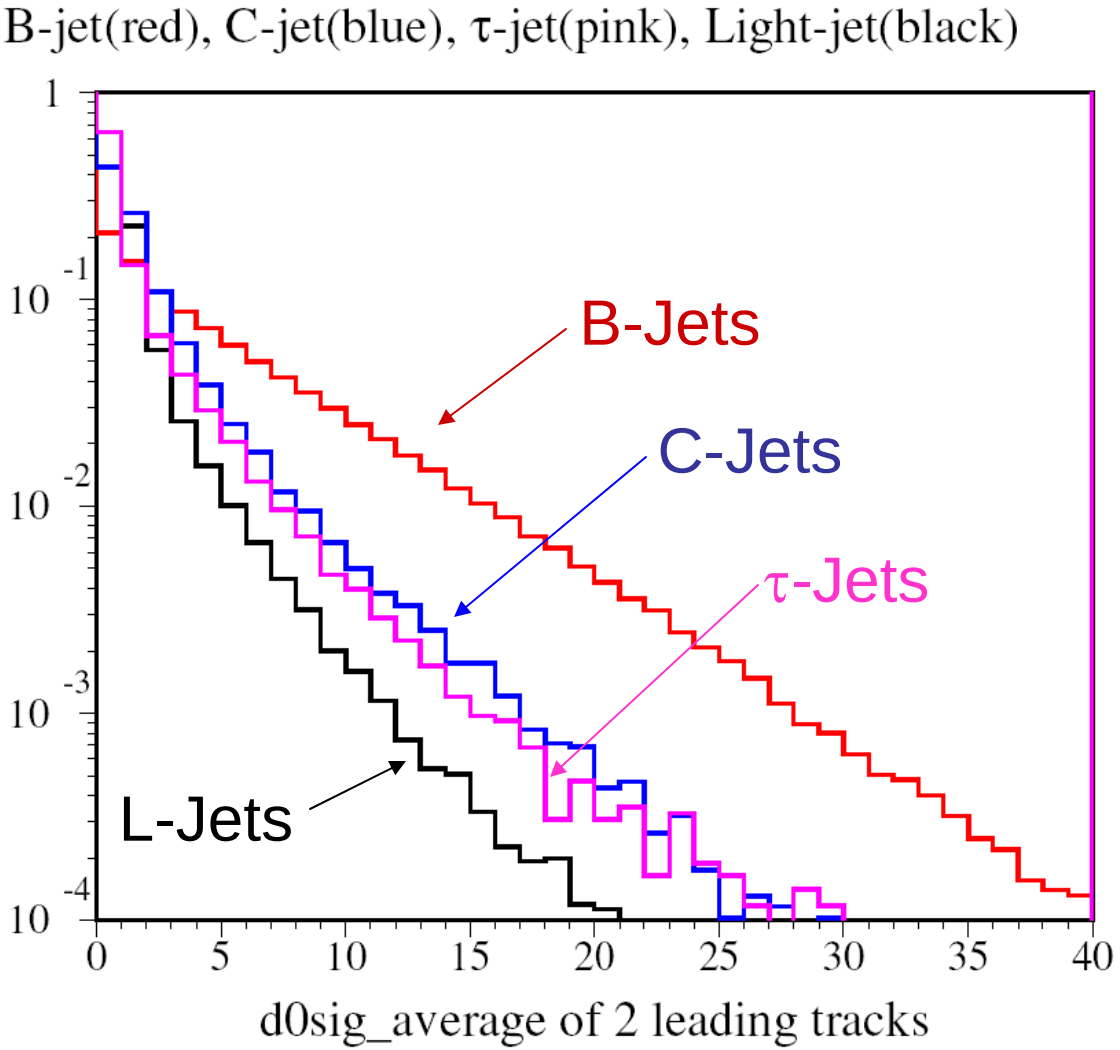
Number of tracks in jet with longitudinal IP > 100 microns (new)



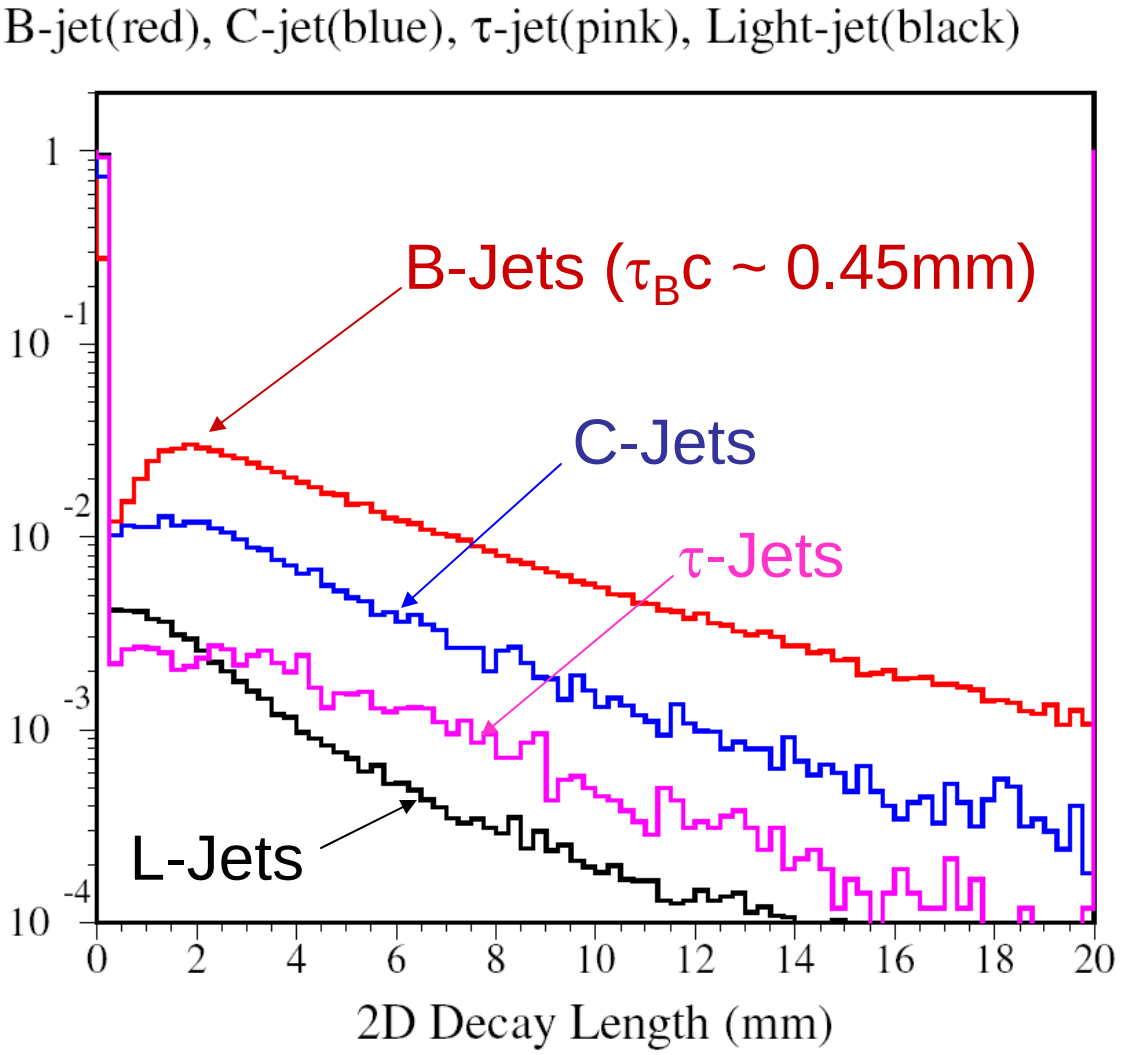
Number of tracks with distance from PV to track-jet cross point > 150microns (new)



Average of impact parameter significance from two leading tracks in jet (new)



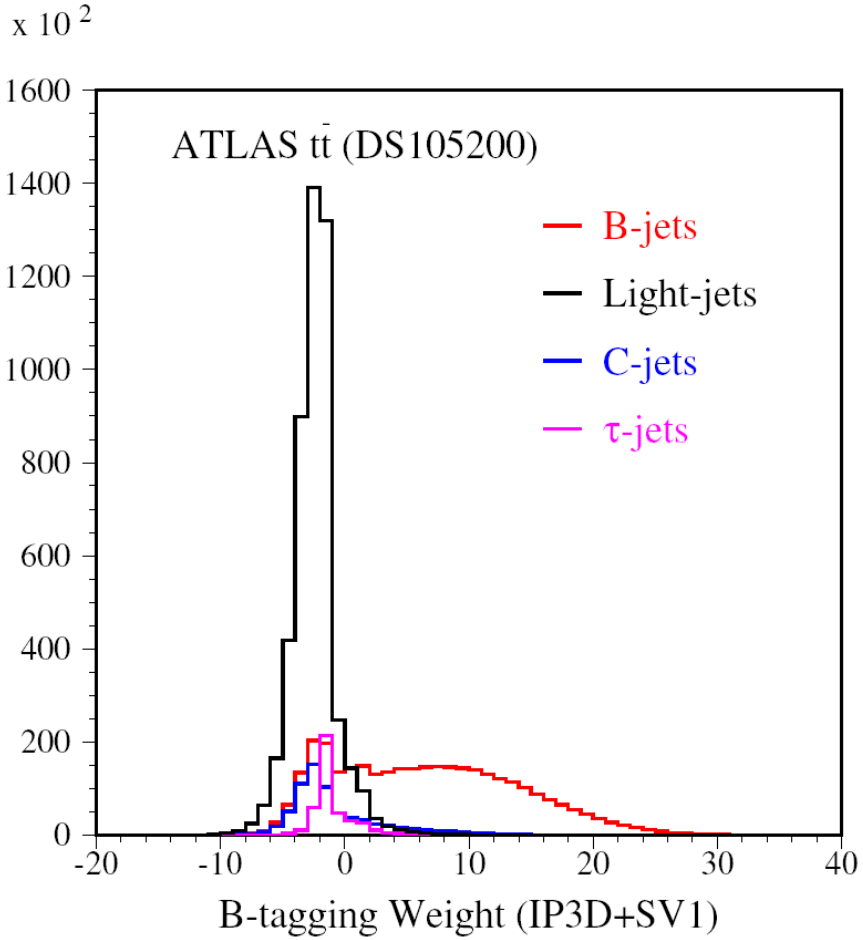
2D Decay Length (new)

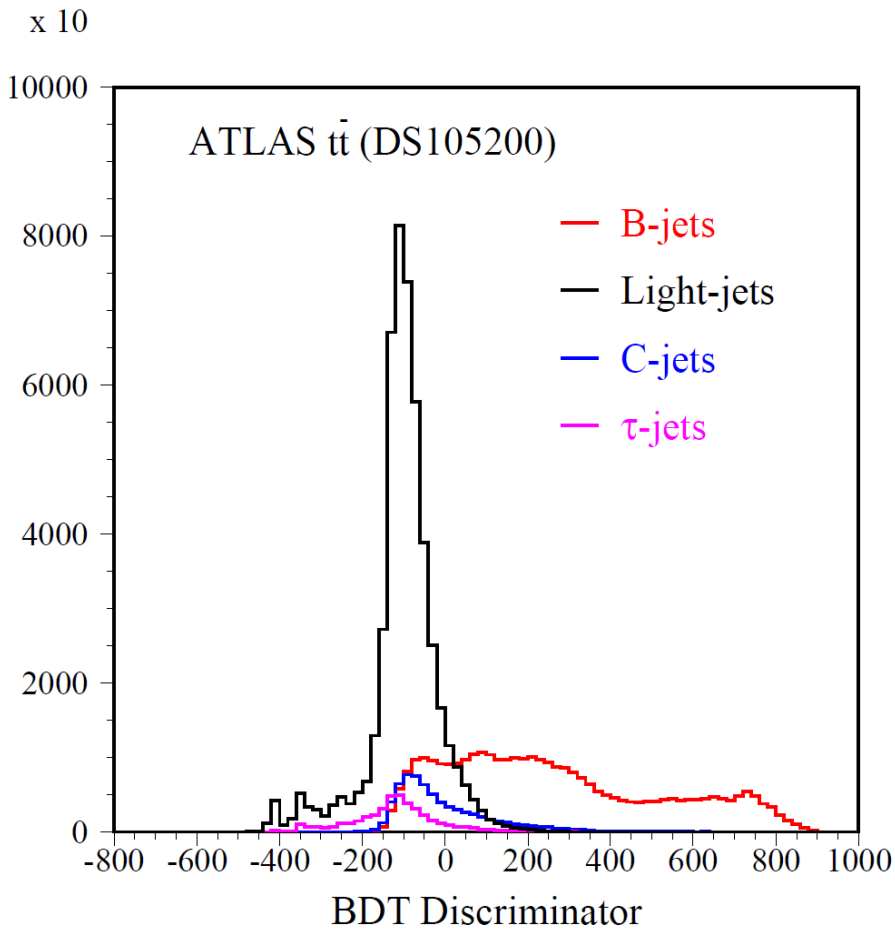


BDT B-taggers

1. **BDT_bl: B Jets vs Light Jets**
2. BDT_bc: B Jets vs C Jets
3. BDT_bt: B Jets vs τ Jets

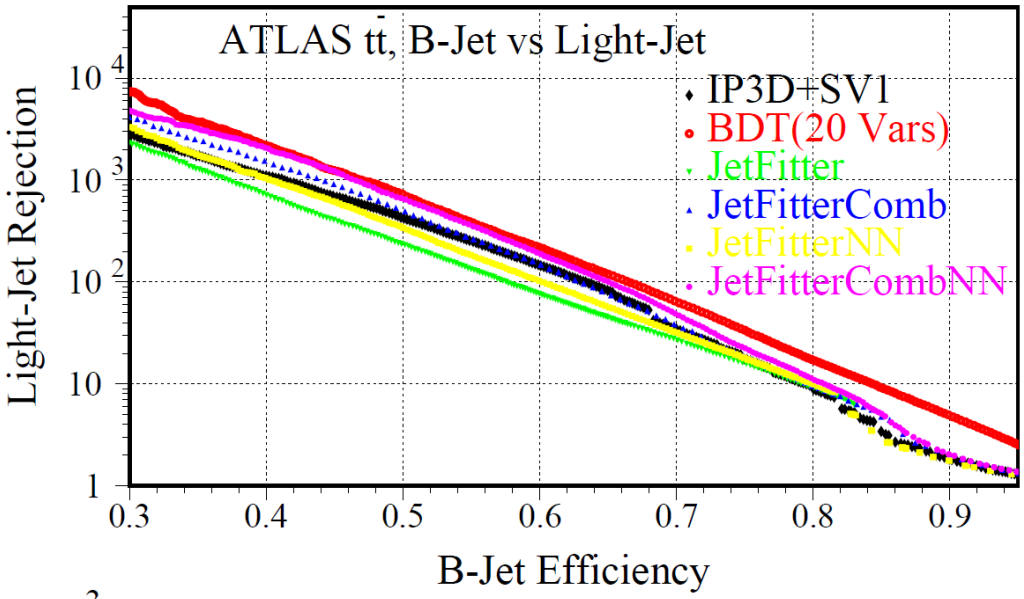
Btaggers: IP3D+SV1 and BDT_bl



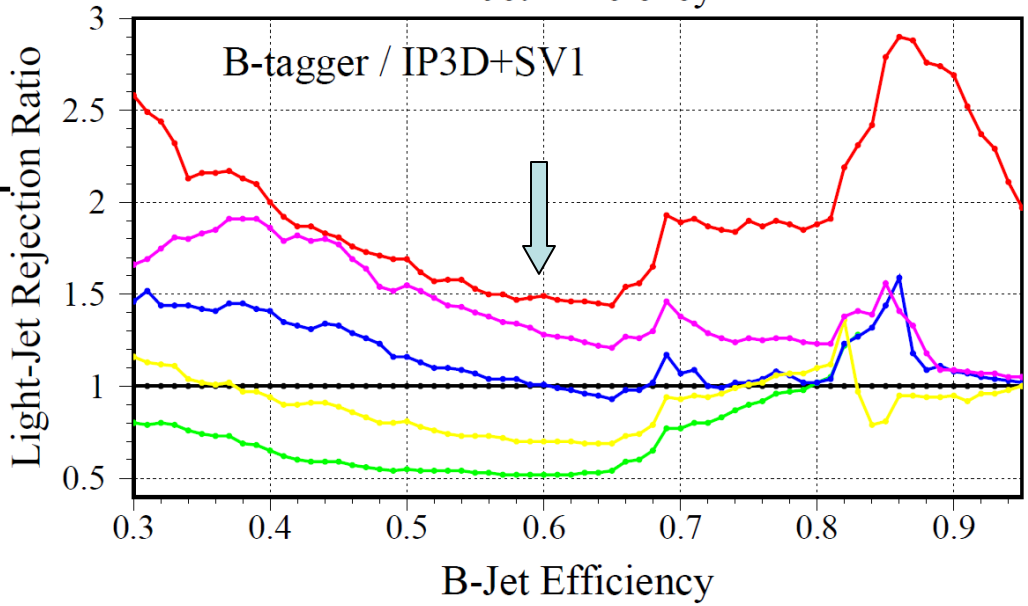


ttbar Samples

Light jet rejection vs b-tagging efficiency

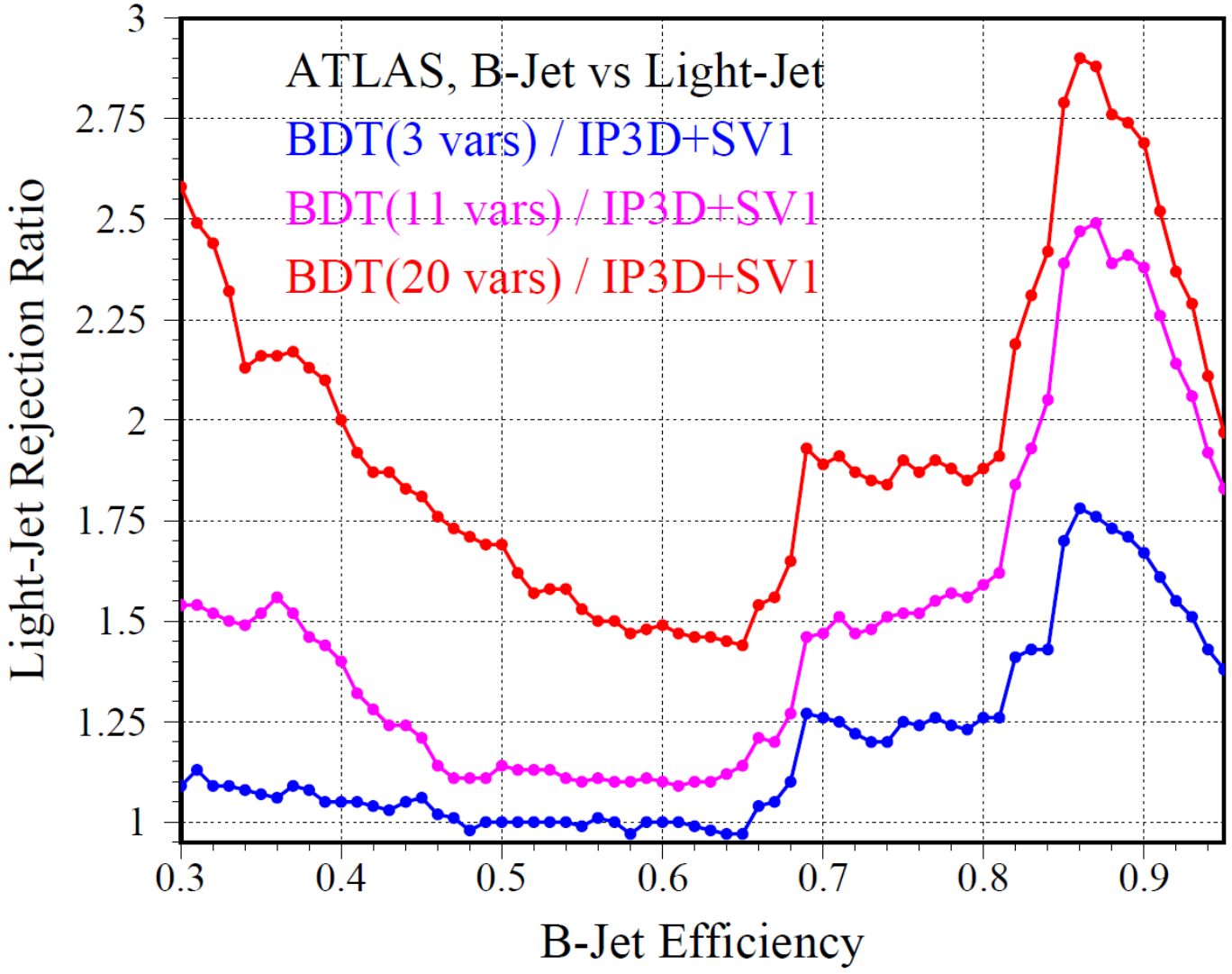


➔For 60% B-Jet Efficiency, BDT B-tagger has ~ 50% improvement over IP3D+SV1



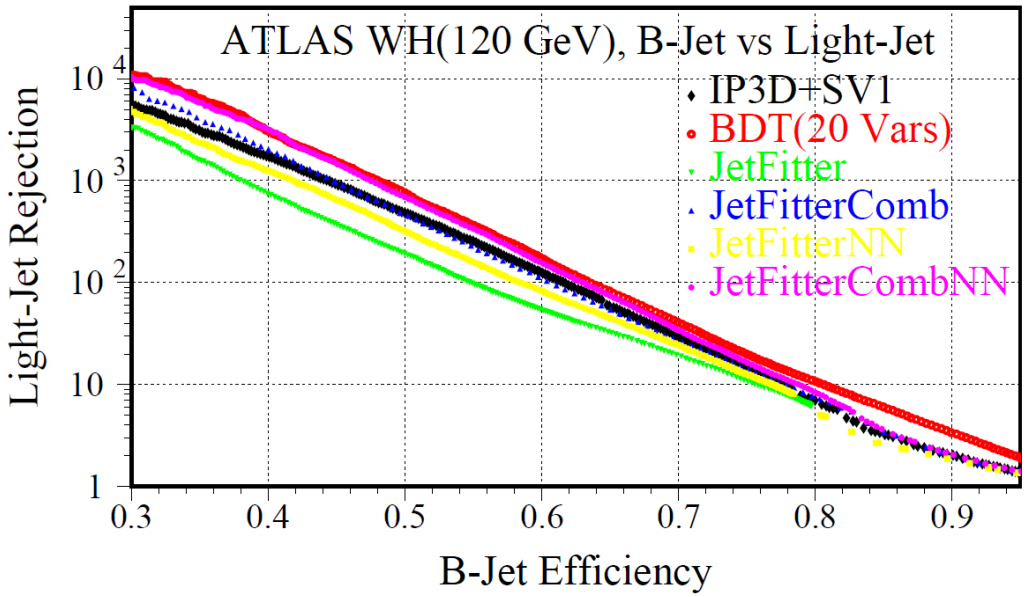
$$\frac{\text{Light-Jet Rejection (B-tagger)}}{\text{Light-Jet Rejection (IP3D+SV1)}}$$

BDT Performance vs Variables Used in BDT



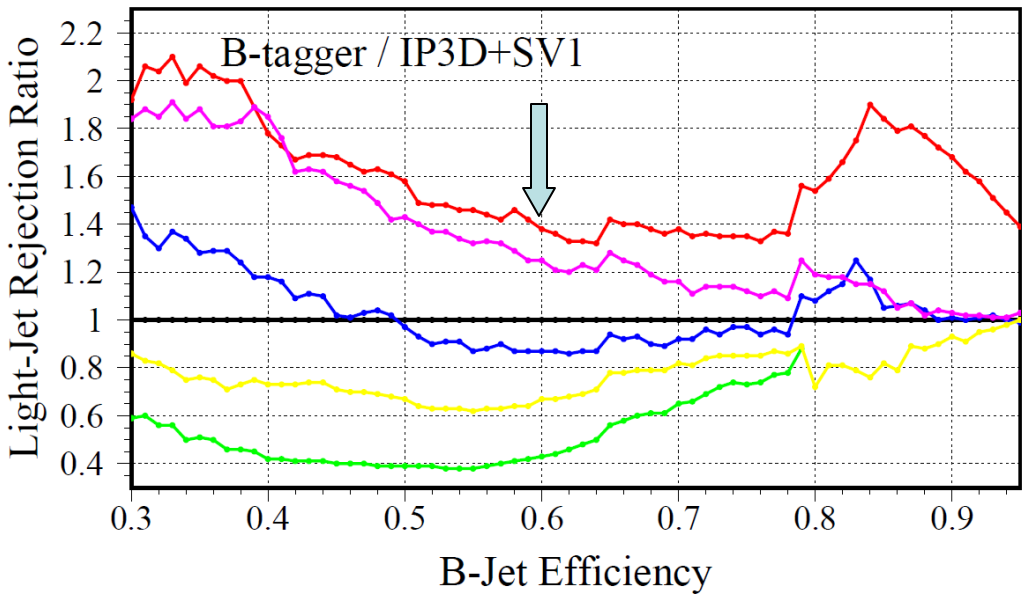
WH120 Samples

Light jet rejection vs b-tagging efficiency



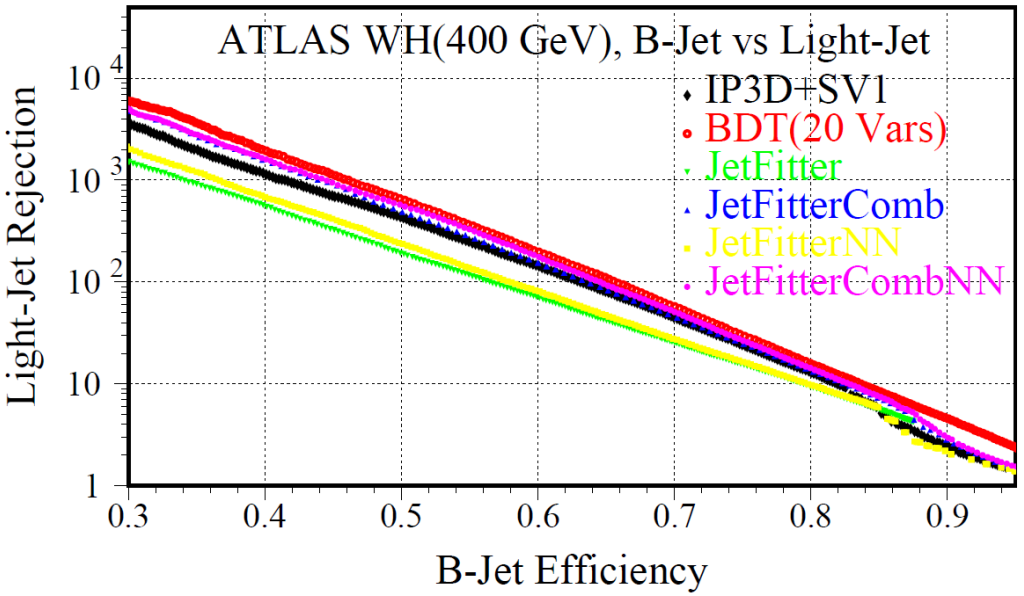
➔For 60% B-Jet Efficiency, BDT B-tagger has ~ 40% improvement over IP3D+SV1

$$\frac{\text{Light-Jet Rejection (B-tagger)}}{\text{Light-Jet Rejection (IP3D+SV1)}}$$



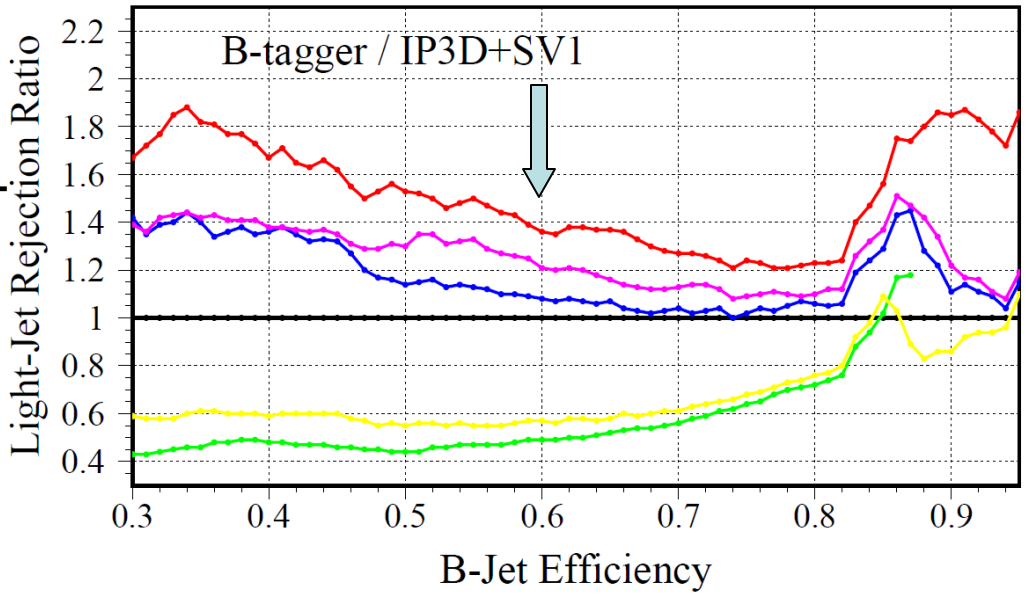
WH400 Samples

Light jet rejection vs b-tagging efficiency



➔For 60% B-Jet Efficiency, BDT B-tagger has ~ 35% improvement over IP3D+SV1.

$$\frac{\text{Light-Jet Rejection (BDT)}}{\text{Light-Jet Rejection (IP3D+SV1)}}$$



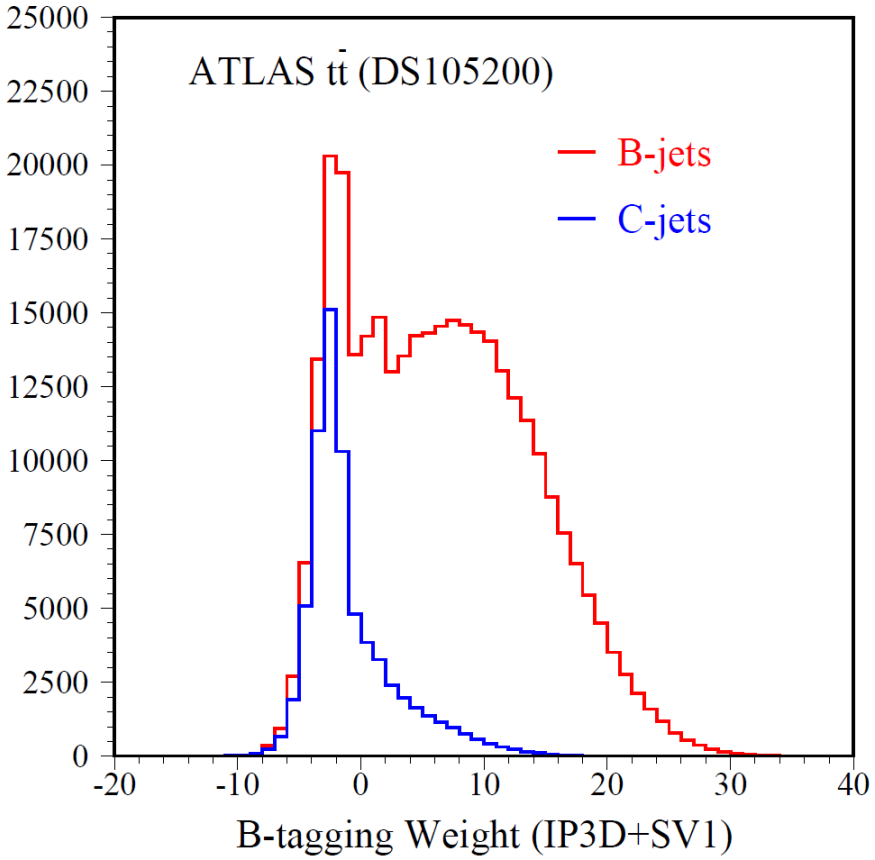
Comparison of B-taggers
(only show the best three)

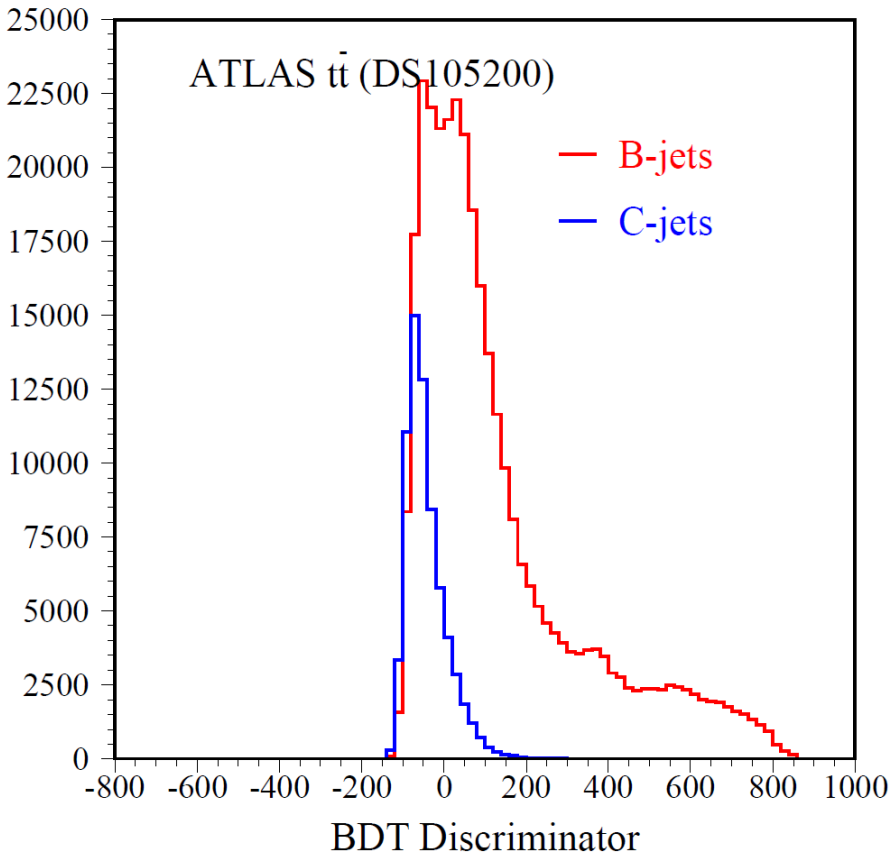
L- jet Rejection / B-jet Efficiency		IP3D+SV1 Rejection	JetFitterCombNN Rejection	BDT_bl Rejection
Ttbar:	Eff_b = 70%	34.5 ±0.3	47.8 ±0.5	65.3 ±0.8
Ttbar:	Eff_b = 60%	146 ± 3	188 ± 4	219 ± 5
Ttbar:	Eff_b = 50%	429 ± 13	663 ± 24	725 ± 28
WH120:	Eff_b = 70%	29.4 ±0.2	34.1 ±0.3	40.7 ±0.4
WH120:	Eff_b = 60%	125 ± 2	156 ± 3	173 ± 3
WH120:	Eff_b = 50%	485 ± 15	691 ± 25	764 ± 29
WH400:	Eff_b = 70%	45.0 ±0.4	50.7 ±0.4	57.3 ±0.5
WH400:	Eff_b = 60%	144 ± 2	175 ± 3	196 ± 3
WH400:	Eff_b = 50%	432 ± 10	560 ± 16	660 ± 20

BDT B-taggers

1. BDT_bl: B Jets vs Light Jets
2. **BDT_bc: B Jets vs C Jets**
3. BDT_bt: B Jets vs τ Jets

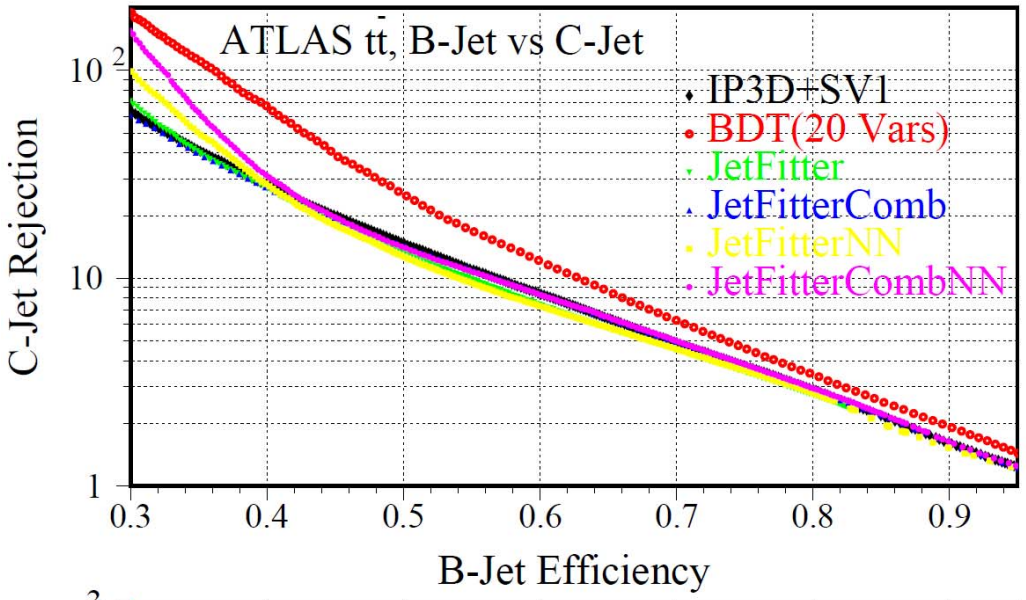
Btaggers: IP3D+SV1 and BDT_bc





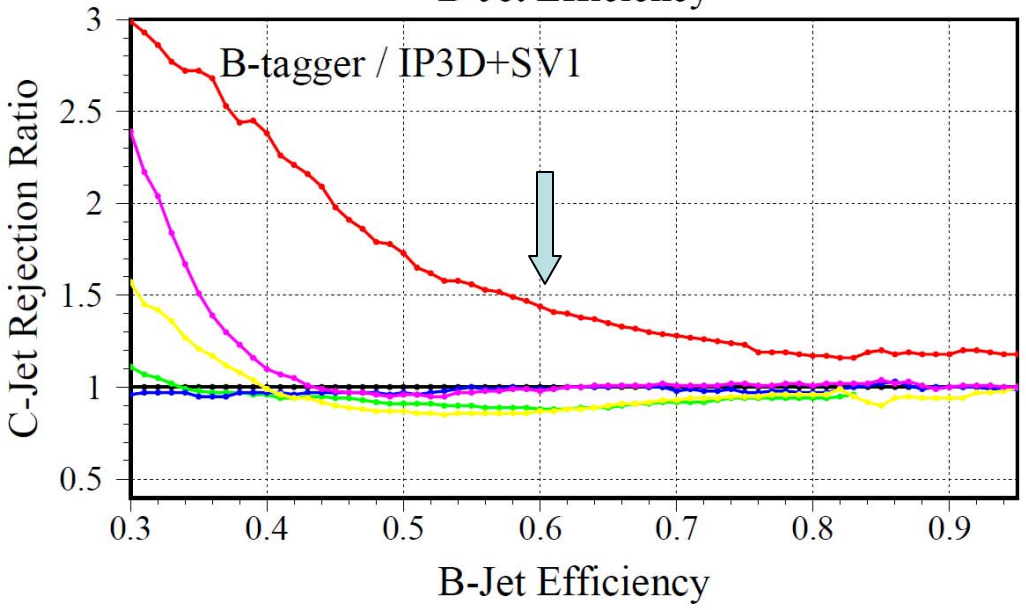
ttbar Samples

C jet rejection vs b-tagging efficiency



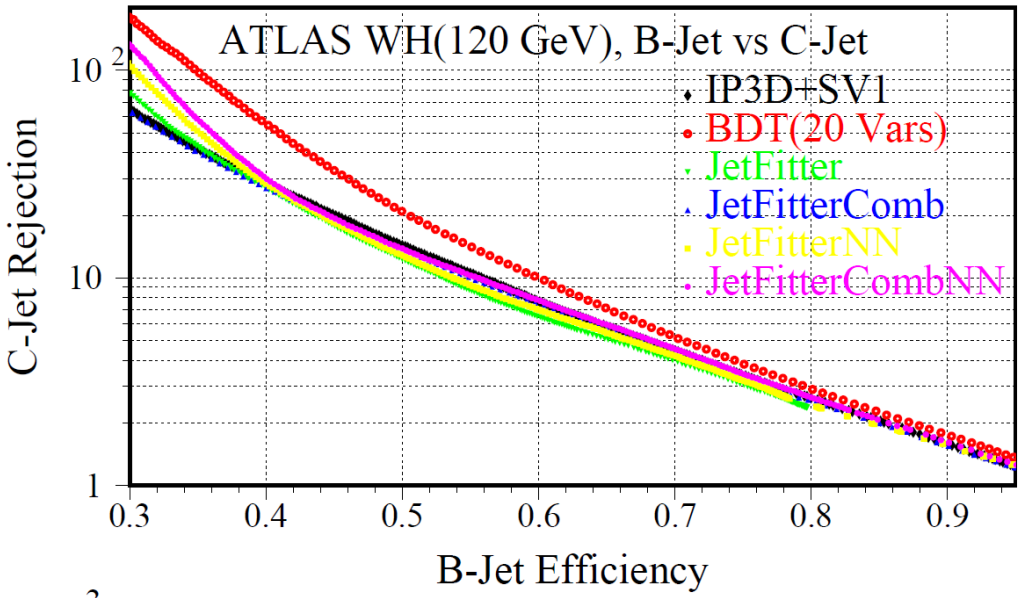
➔For 60% B-Jet Efficiency, BDT B-tagger has ~ 40% improvement over IP3D+SV1

$$\frac{\text{C-Jet Rejection (B-tagger)}}{\text{C-Jet Rejection (IP3D+SV1)}}$$



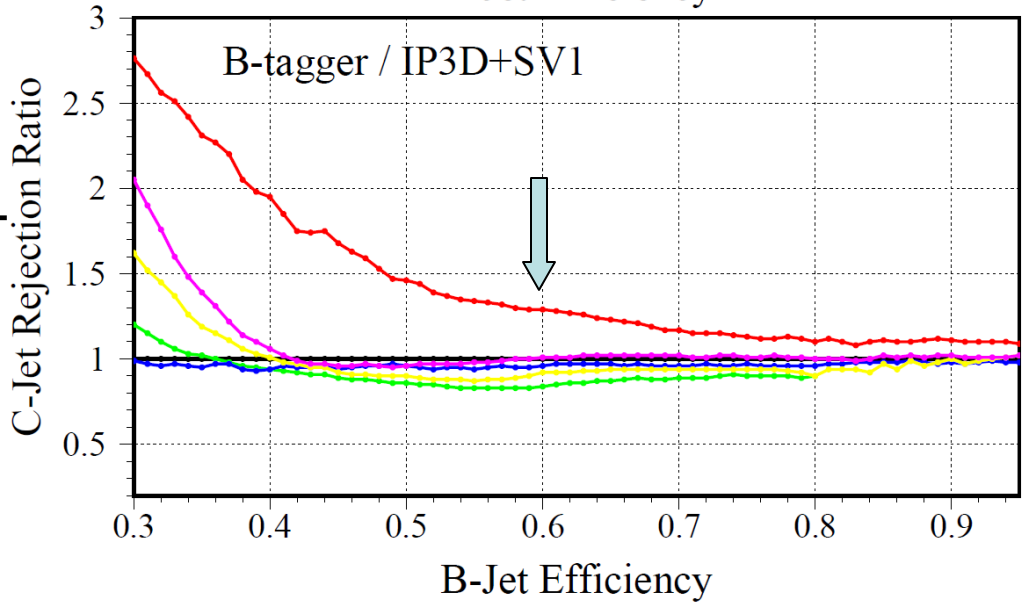
WH120 Samples

C jet rejection vs b-tagging efficiency



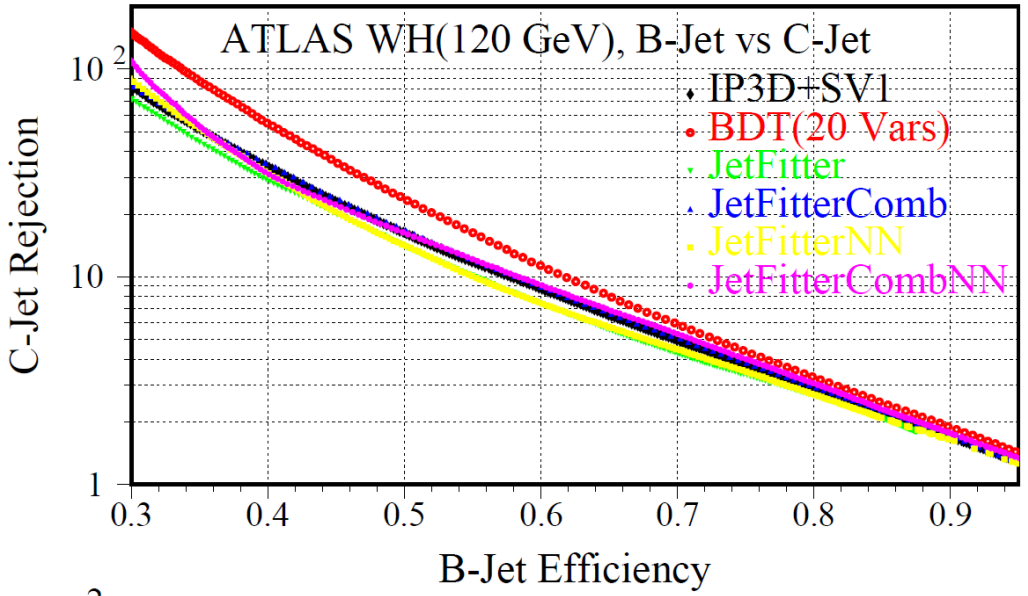
➔For 60% B-Jet Efficiency, BDT B-tagger has ~ 30% improvement over IP3D+SV1

$$\frac{\text{C-Jet Rejection (B-tagger)}}{\text{C-Jet Rejection (IP3D+SV1)}}$$



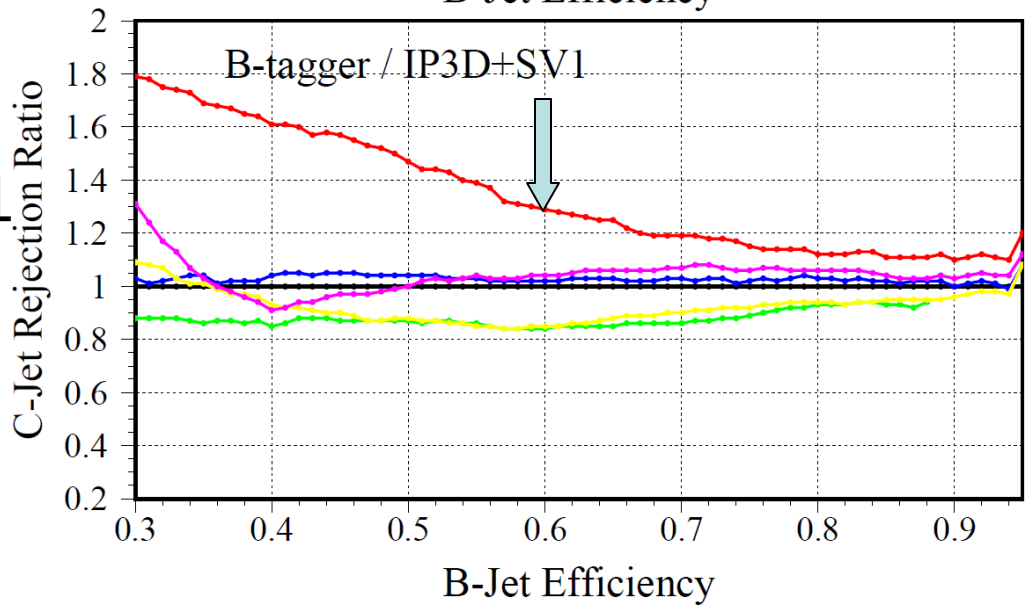
WH400 Samples

C jet rejection vs b-tagging efficiency



➔For 60% B-Jet Efficiency, BDT B-tagger has ~ 30% improvement over IP3D+SV1

$$\frac{\text{C-Jet Rejection (B-tagger)}}{\text{C-Jet Rejection (IP3D+SV1)}}$$



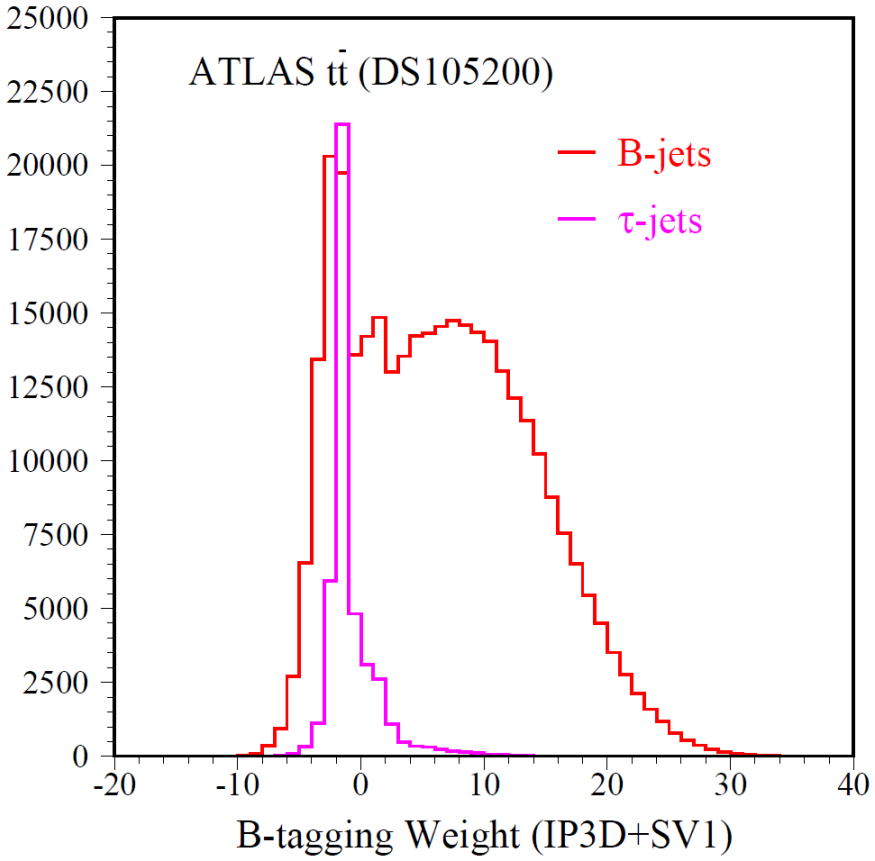
Comparison of B-taggers
(only show the best three)

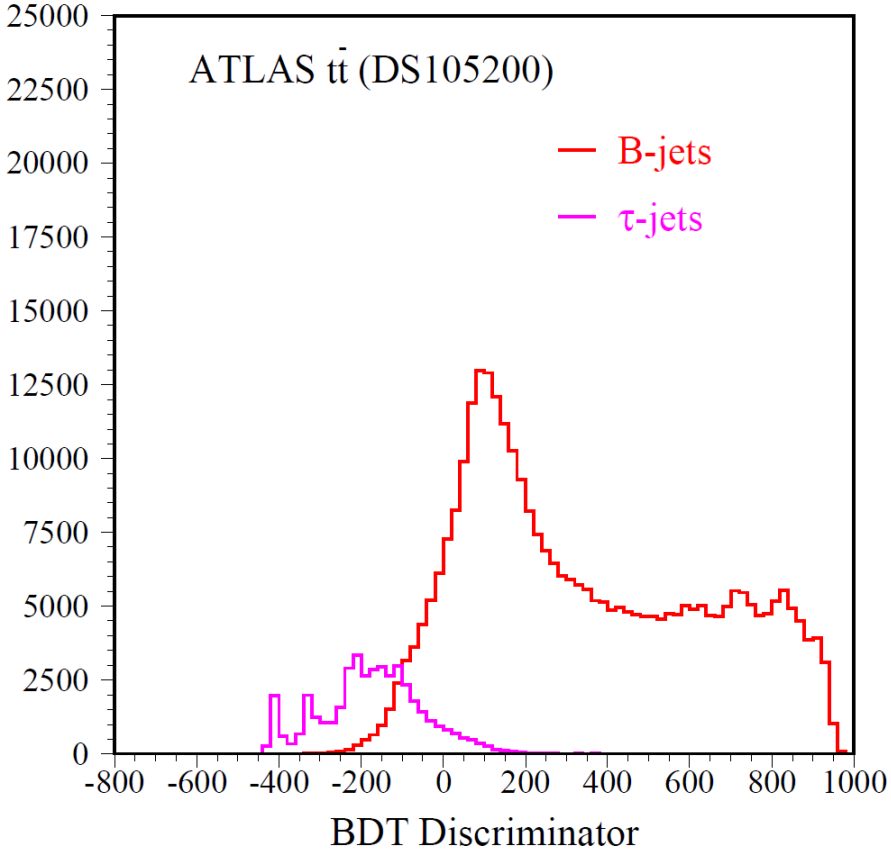
C- jet Rejection / B-jet Efficiency		IP3D+SV1 Rejection	JetFitterCombNN Rejection	BDT_bc Rejection
Ttbar:	Eff_b = 70%	4.9 ± 0.04	5.0 ± 0.04	6.3 ± 0.06
Ttbar:	Eff_b = 60%	8.4 ± 0.09	8.3 ± 0.09	12.1 ± 0.16
Ttbar:	Eff_b = 50%	14.7 ± 0.22	14.0 ± 0.20	25.5 ± 0.49
WH120:	Eff_b = 70%	4.5 ± 0.02	4.5 ± 0.02	5.2 ± 0.03
WH120:	Eff_b = 60%	7.7 ± 0.05	7.8 ± 0.05	9.9 ± 0.08
WH120:	Eff_b = 50%	14.3 ± 0.14	13.8 ± 0.13	20.9 ± 0.24
WH400:	Eff_b = 70%	4.9 ± 0.02	5.3 ± 0.03	5.9 ± 0.03
WH400:	Eff_b = 60%	8.8 ± 0.06	9.1 ± 0.06	11.3 ± 0.08
WH400:	Eff_b = 50%	16.2 ± 0.14	16.1 ± 0.14	23.8 ± 0.26

BDT B-taggers

1. BDT_bl: B Jets vs Light Jets
2. BDT_bc: B Jets vs C Jets
3. **BDT_bt: B Jets vs τ Jets**

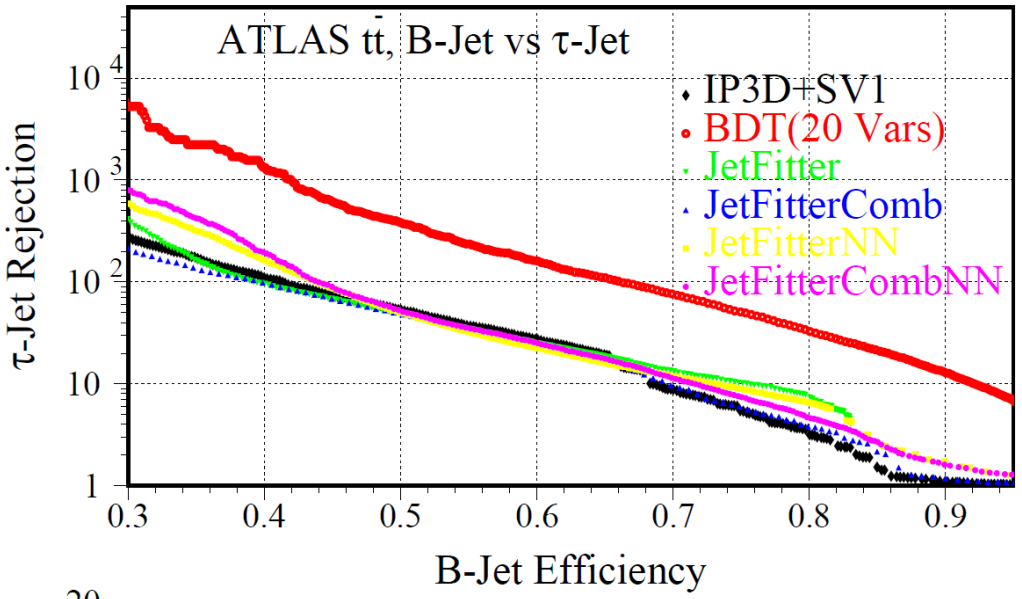
Btaggers: IP3D+SV1 and BDT_bt





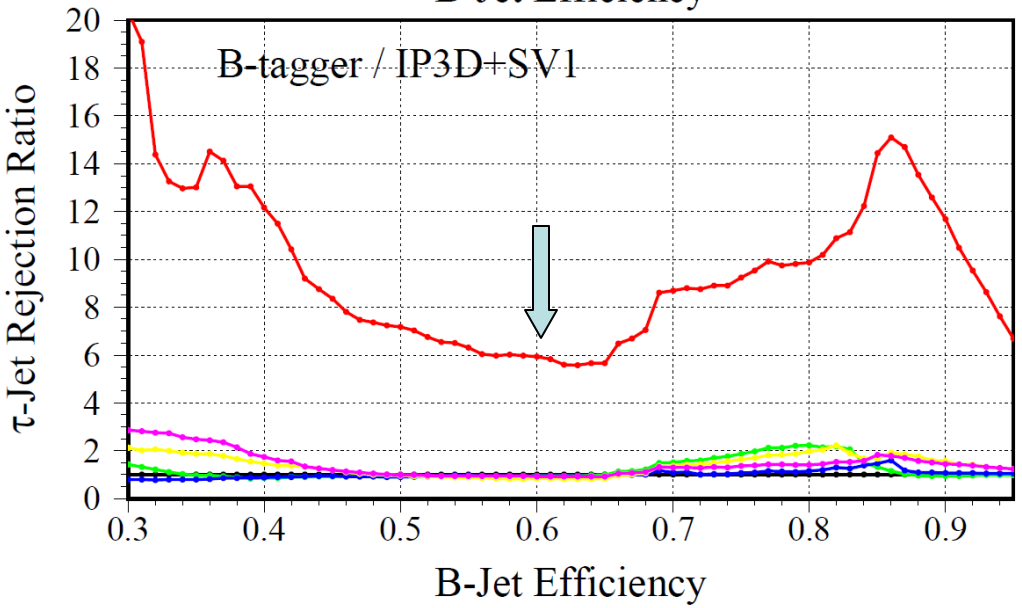
ttbar Samples

τ-jet rejection vs b-tagging efficiency



➔For 60% B-Jet Efficiency, BDT B-tagger has >~500% improvement over IP3D+SV1

$$\frac{\tau\text{-Jet Rejection (B-tagger)}}{\tau\text{-Jet Rejection (IP3D+SV1)}}$$

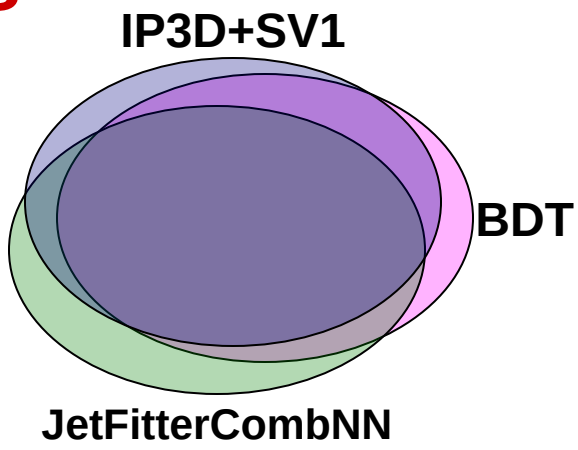


Comparison of B-taggers
(only show the best three)

τ- jet Rejection / B-jet Efficiency		IP3D+SV1 Rejection	JetFitterCombNN Rejection	BDT_bt Rejection
Ttbar:	Eff_b = 70%	8.8 ± 0.1	11.3 ± 0.2	76 ± 3.2
Ttbar:	Eff_b = 60%	27.1 ± 0.7	25.1 ± 0.6	161 ± 10
Ttbar:	Eff_b = 50%	53.2 ± 2.0	52.6 ± 2.0	382 ± 36

Cross checks of ATLAS Btaggers: Overlapped B-jets

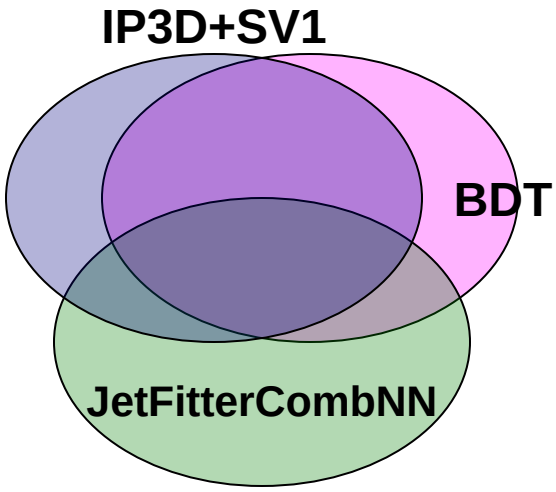
- Apply cuts for IP3D+SV1, BDT and JetFitterCombNN with 60% of B jet efficiency, respectively.
- Then calculate the overlapped B-jets passed these cuts. ***Overlapped efficiency = (A.and.B) / (A.or.B)***



No. of B-jets	IP3D+SV1	JetFitterCombNN	BDT_bl
IP3D+SV1	198196	182993/213399 = 85.8%	190827/205564 = 92.8%
JetFitterCombNN		198196	182386/214006 = 85.2%
BDT_bl			198196

Btaggers: Overlapped Light-jets

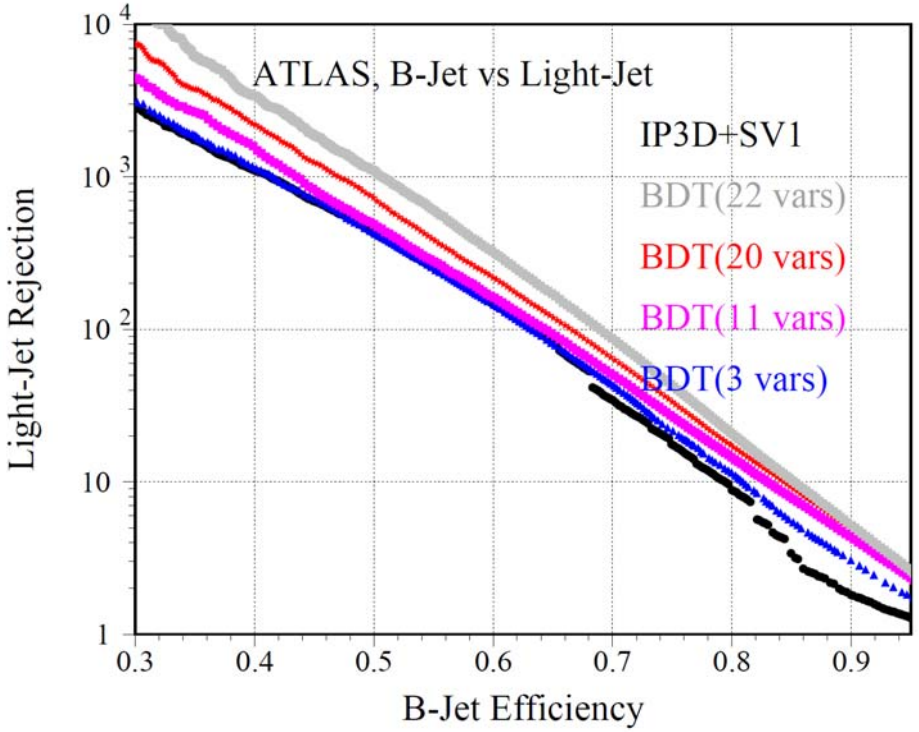
- Apply cuts for IP3D+SV1, BDT and JetFitterCombNN with 60% of B jet efficiency, respectively.
- Then calculate the overlapped light-jets passed these cuts. ***Overlapped efficiency = (A.and.B) / (A.or.B)***

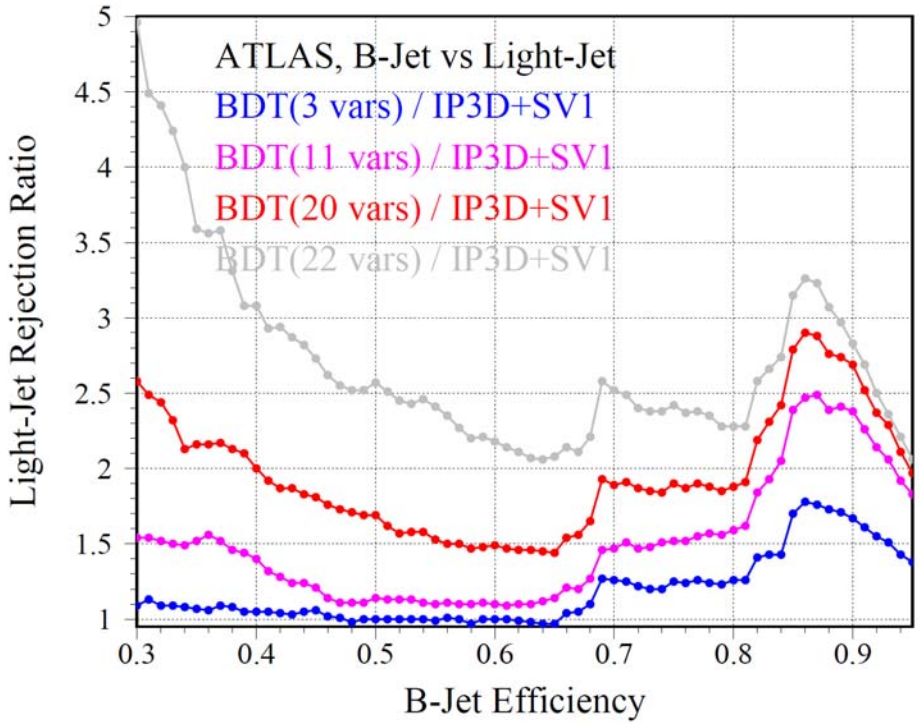


No. of L-jets	IP3D+SV1	JetFitterCombNN	BDT_bl
IP3D+SV1	3536	1446/4875=30%	2027/3886=52%
JetFitterCombNN		2785	1178/3984=30%
BDT_bl			2377

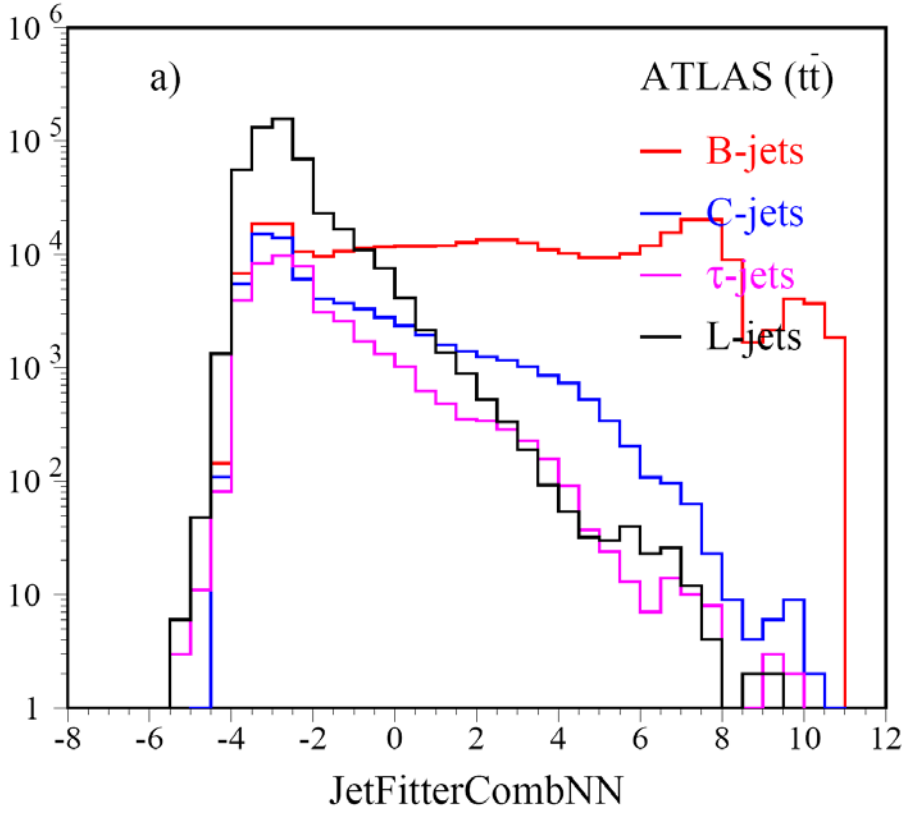
BDT B-tagger: Further Improvement

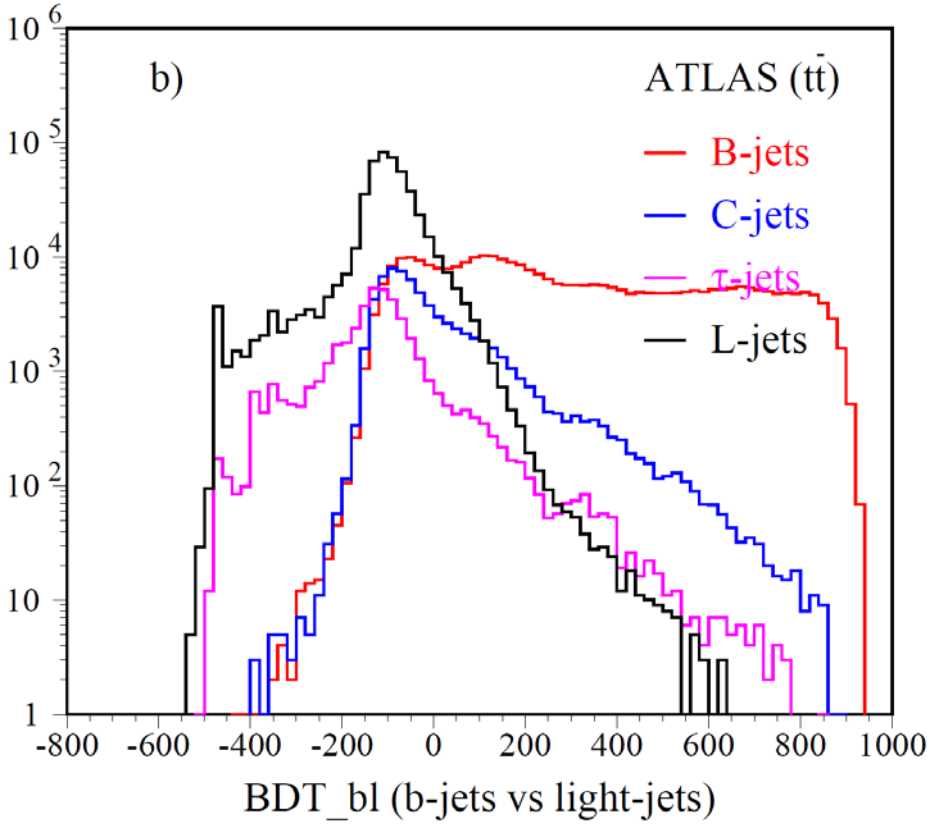
→ 3 vars: IP2D, IP3D and SV1
→ 11 vars: 11 existing variables
→ 20 vars: 11 exisiting variables plus 9 new variables
→ 22 vars: 20 vars plus IP3D+SV1 and JetFitterCombNN





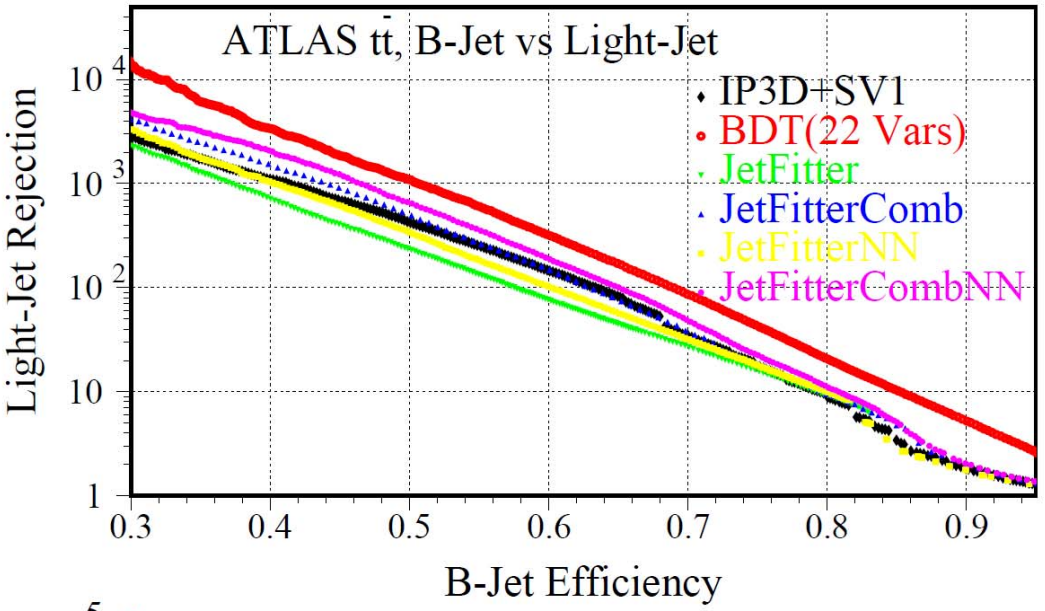
JetFitterCombNN and BDT_bl



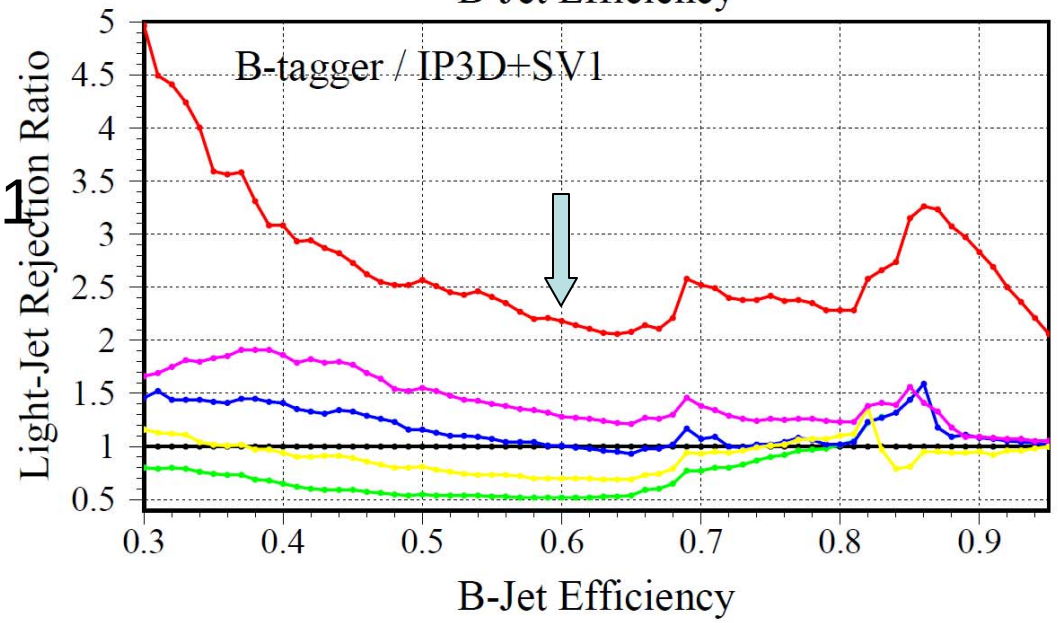


ttbar Samples

Light jet rejection vs b-tagging efficiency



➔For 60% B-Jet Efficiency, BDT B-tagger has ~ 120% improvement over IP3D+SV1

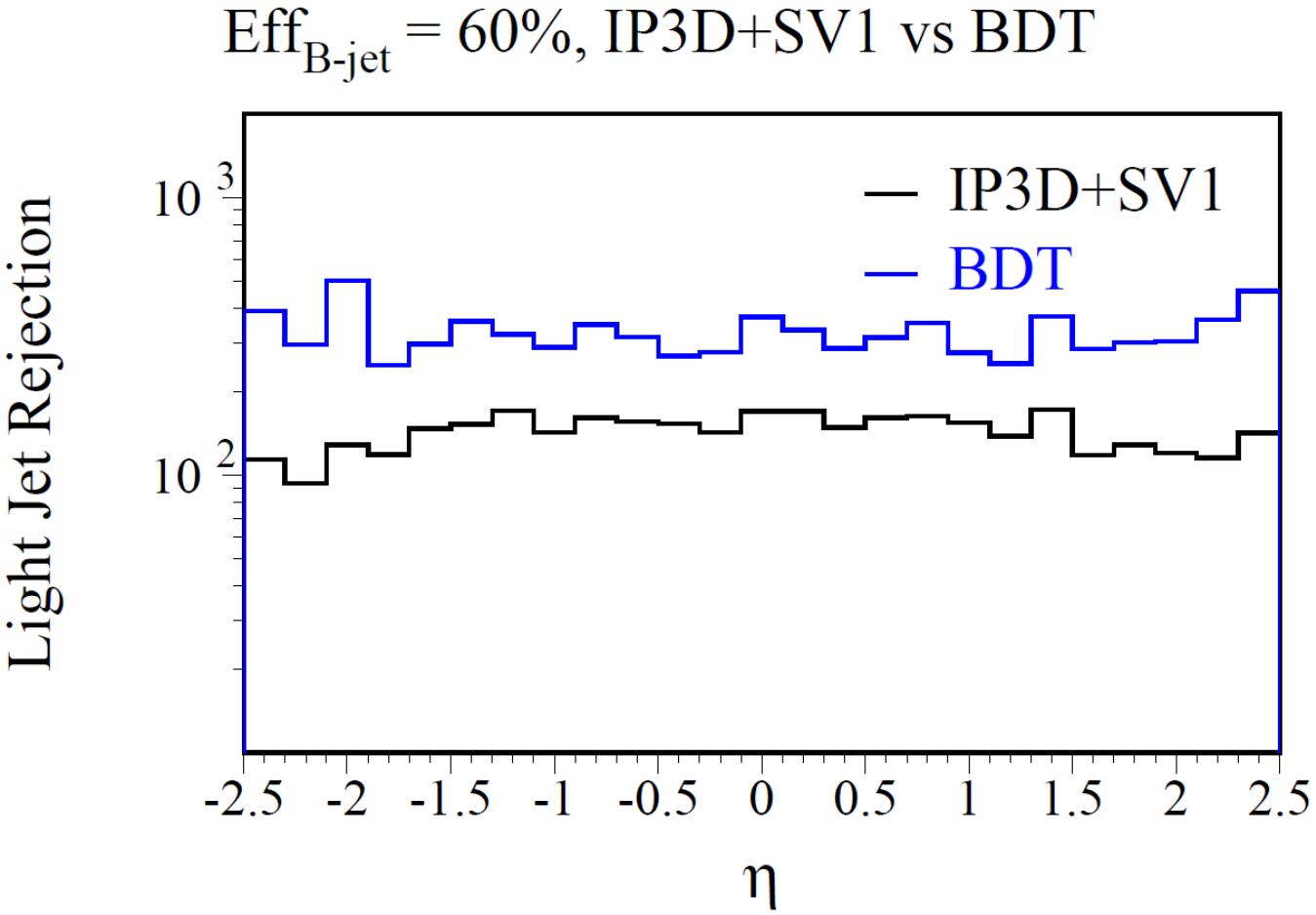


$$\frac{\text{Light-Jet Rejection (B-tagger)}}{\text{Light-Jet Rejection (IP3D+SV1)}}$$

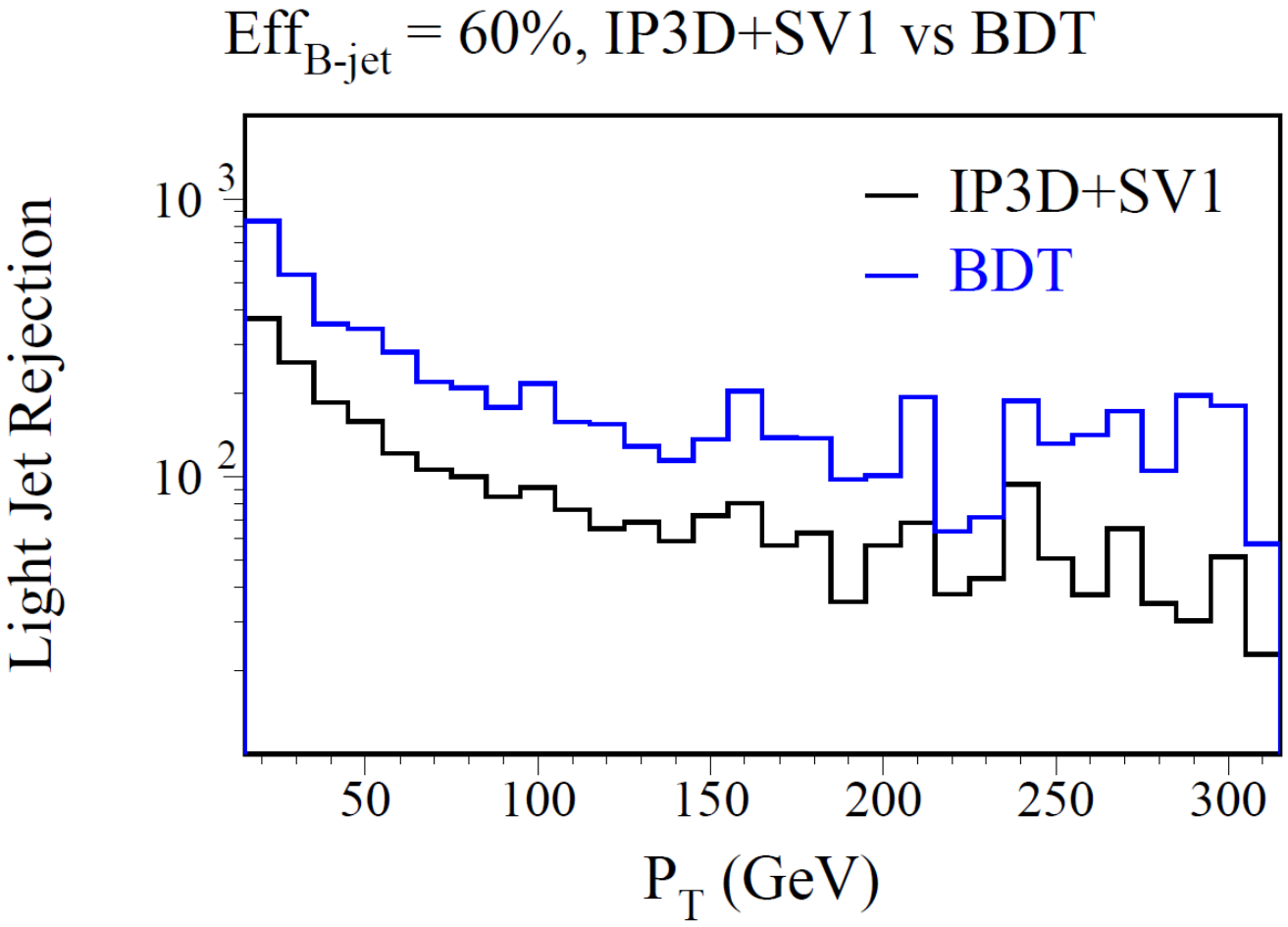
Comparison of B-taggers
(only show the best three)

L- jet Rejection / B-jet Efficiency		IP3D+SV1 Rejection	JetFitterCombNN Rejection	BDT_bl Rejection
Ttbar:	Eff_b = 70%	34.5 ±0.3	47.8 ±0.5	87.1 ±1.2
Ttbar:	Eff_b = 60%	146 ± 3	188 ± 4	319 ±8.2
Ttbar:	Eff_b = 50%	429 ± 13	663 ± 24	1100 ± 52
WH120:	Eff_b = 70%	29.4 ±0.2	34.1 ±0.3	53.9 ±0.5
WH120:	Eff_b = 60%	125 ± 2	156 ± 3	253 ±5.5
WH120:	Eff_b = 50%	485 ± 15	691 ± 25	1133 ± 52
WH400:	Eff_b = 70%	45.0 ±0.4	50.7 ±0.4	82.4 ±0.9
WH400:	Eff_b = 60%	144 ± 2	175 ± 3	281 ±5.5
WH400:	Eff_b = 50%	432 ± 10	560 ± 16	945 ± 34

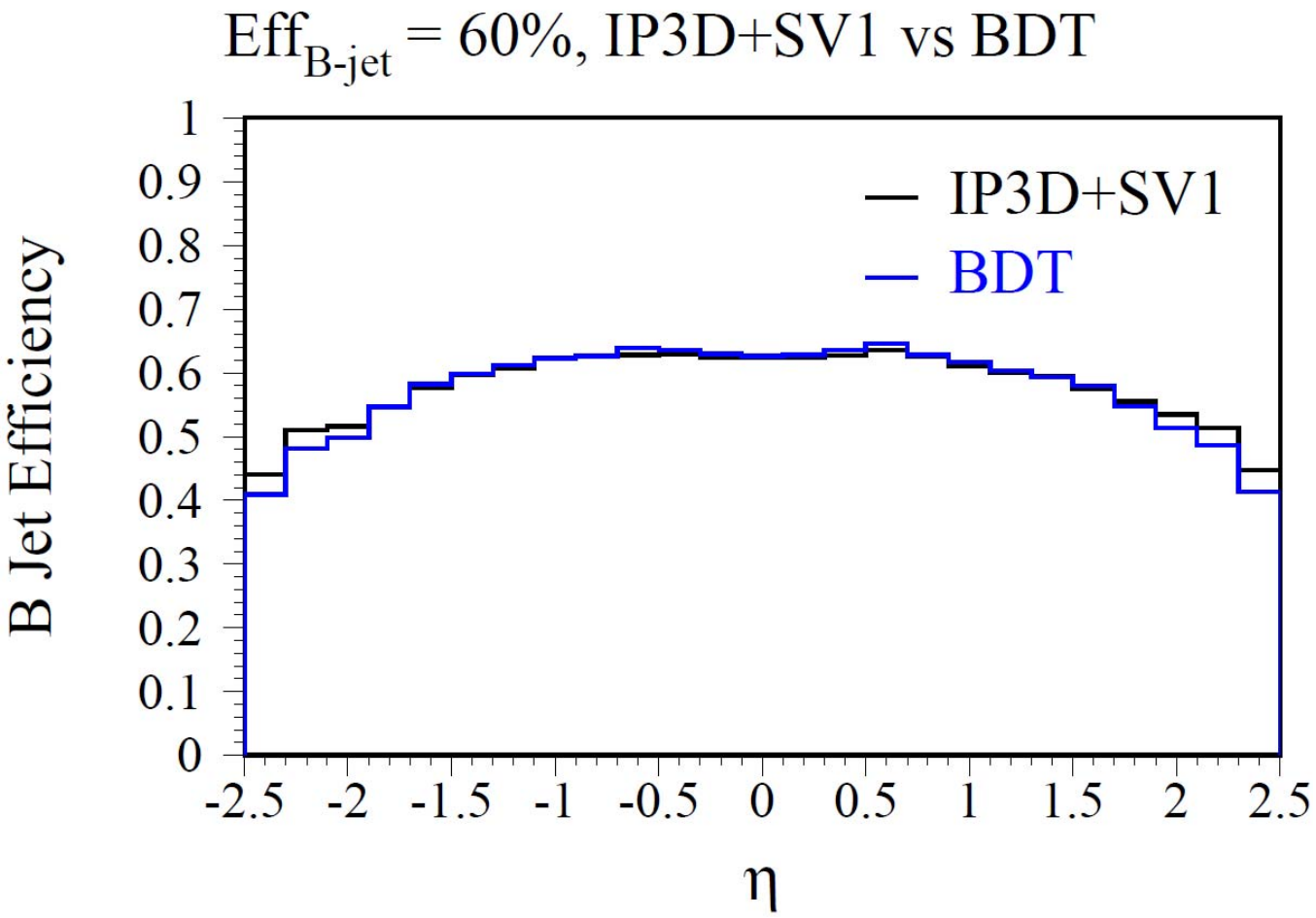
BDT_bl vs IP3D+SV1
(light jet rejection vs η)



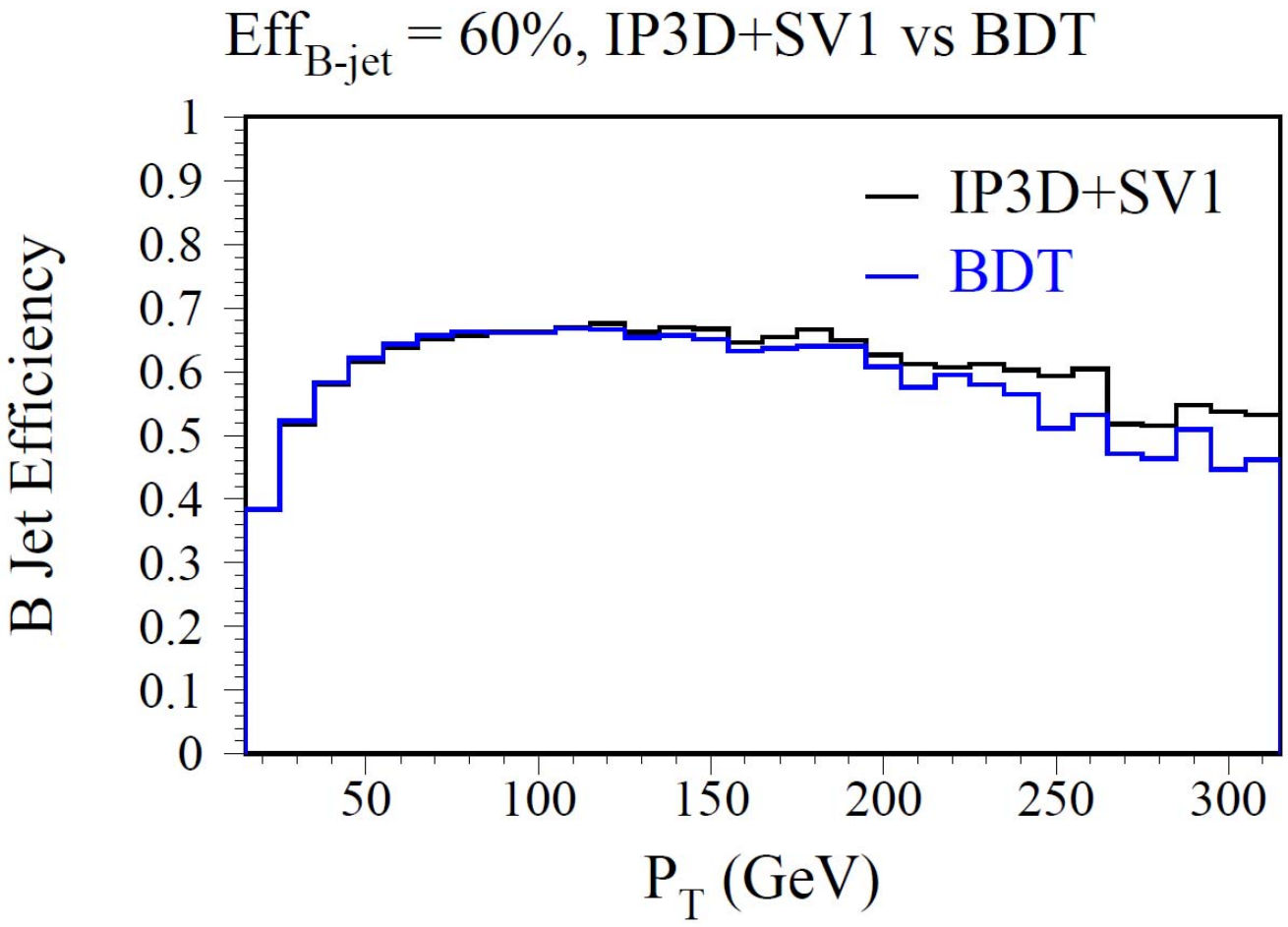
BDT_bl vs IP3D+SV1
(light jet rejection vs P_T)



BDT_bl vs IP3D+SV1
(B-Jet Efficiency vs η)



BDT_bl vs IP3D+SV1
(B-Jet Efficiency vs P_T)



Summary and Future Plan

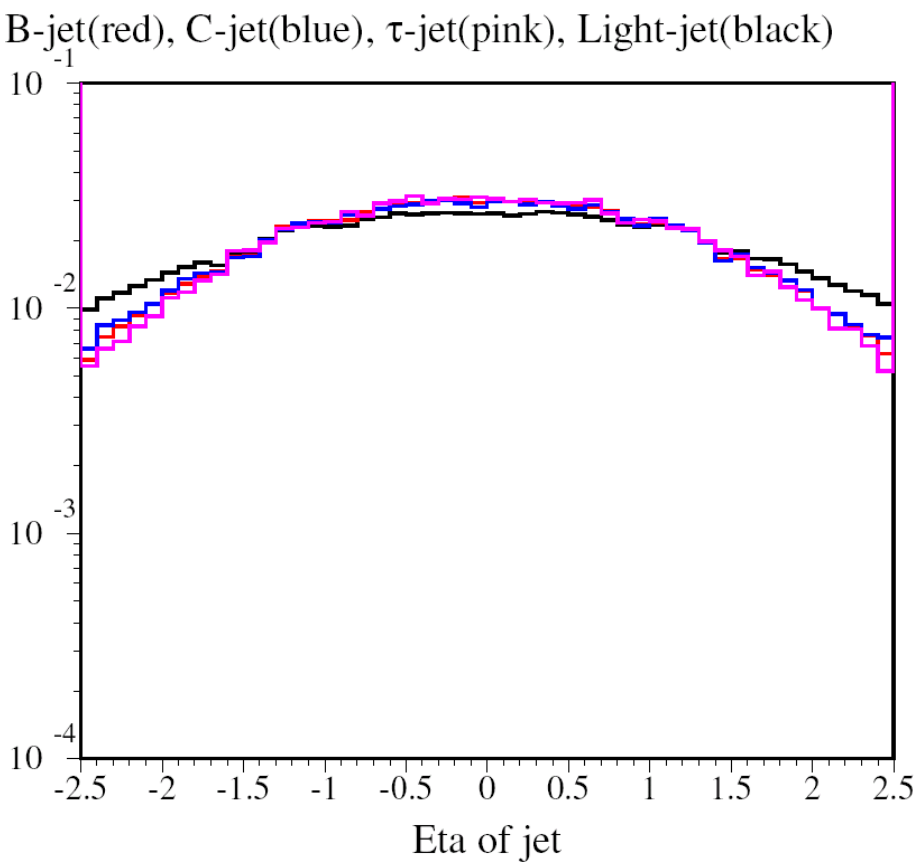
- B-tagger based on BDT technique has been developed and successfully implemented in offline Athena AnalysisExamples package.
- By building 9 new discriminating variables, BDT B-tagger improves light jet rejection by a factor of ~1.5 compared to ATLAS default B-tagger (IP3D+SV1).
- By combining JetFitterCombNN, the BDT B-tagger improves light jet rejection by a factor of ~ 2.2 compared to ATLAS default B-tagger (IP3D+SV1).

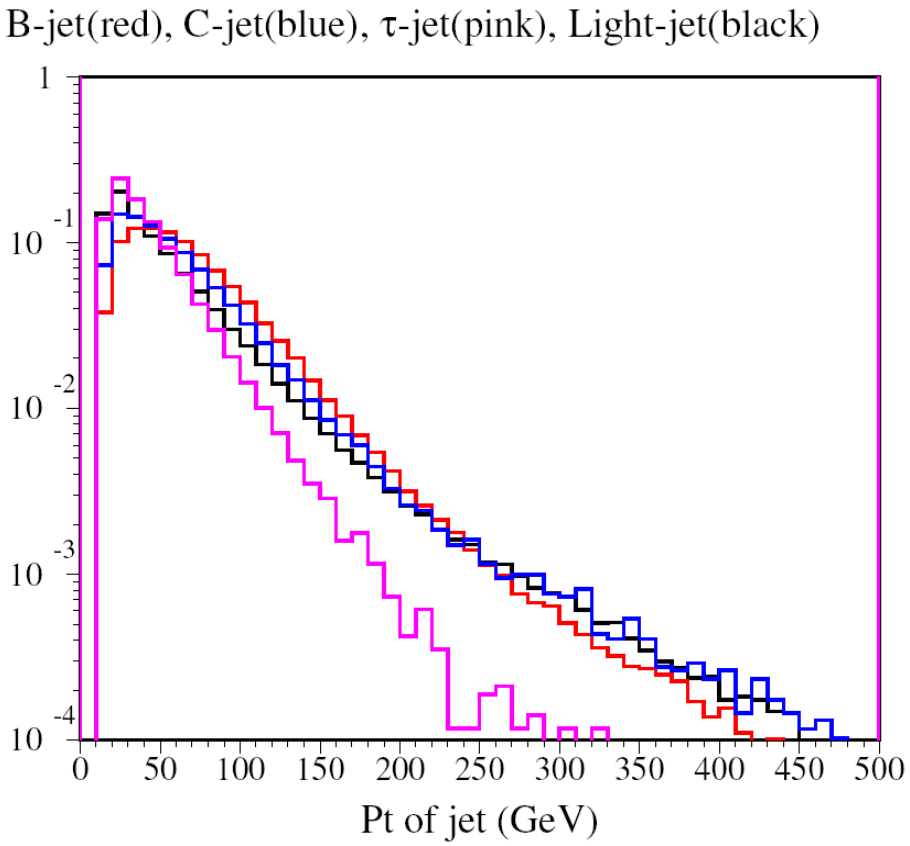
- We are preparing a note and planning to implement BDT B-taggers in the ATLAS official Jet-Tag package.

Backup Slides

BDT Input Variables	Relative Gini Index Contribution (%)		
	BDT_bl(for light-jets)	BDT_bc(for C-jets)	BDT_bt(for τ-jets)
Eleven Existing Discriminating Variables			
IP2D	0.4	0.73	0.56
IP3D	**26.3**	**12.75**	**1.97**
SV1	**42.24**	**55.18**	**20.06**
softe	**1.11**	0.83	0.37
d0sig_max	0.4	0.65	0.55
z0sig_max	**1.63**	**1.14**	0.77
mass	0.18	**8.75**	**8.14**
efrac	**13.39**	**1.16**	1.54
pt_max	**1.31**	0.61	**4.80**
nvertex_2track	0.08	**1.31**	**3.26**
ntrack	**2.84**	0.60	**45.43**
Nine New Discriminating Variables			
ntrack_distance_150	**5.76**	**2.56**	**2.46**
ntrack_z0_100	**1.44**	0.26	**1.93**
ntrack_z0sig_05	0.12	0.38	1.62
2d_dl	**0.76**	**4.76**	**1.98**
d0sig_avg	0.36	0.63	0.36
d0_avg	0.20	1.04	1.14
z0_avg	0.41	**4.77**	**2.09**
sumtrkpt_jetE	0.39	0.29	0.30
sumept_jetE	0.67	**1.60**	0.69

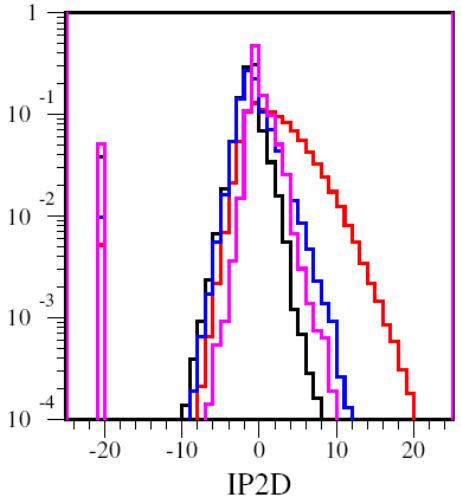
Eta and Pt of Jets

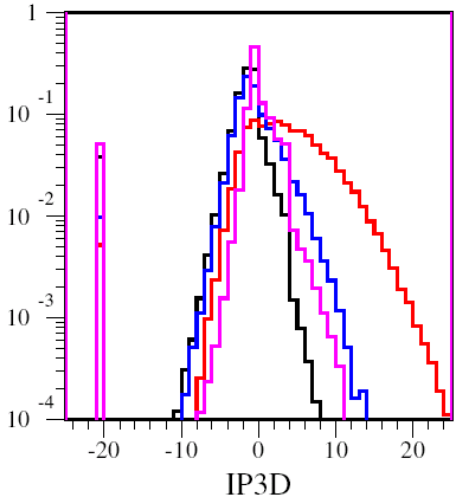


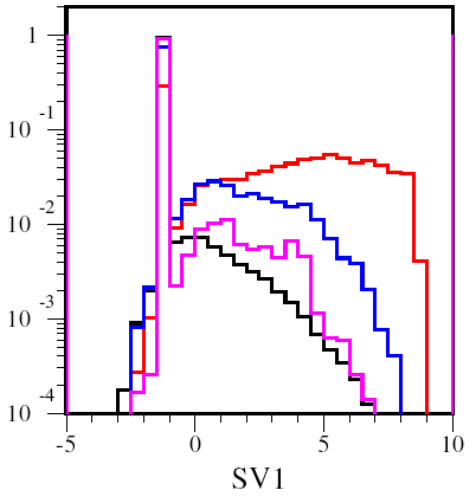


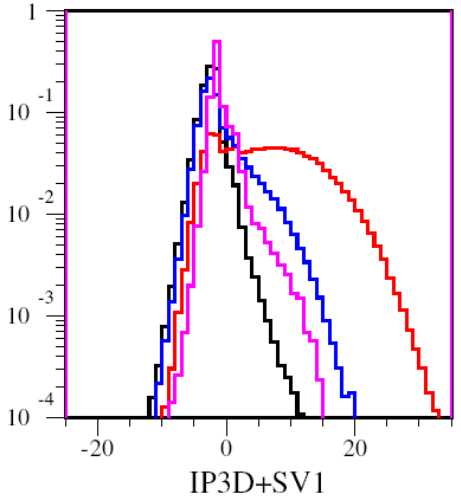
List of Existing Variables

B-jet(red), C-jet(blue), τ-jet(pink), Light-jet(black)





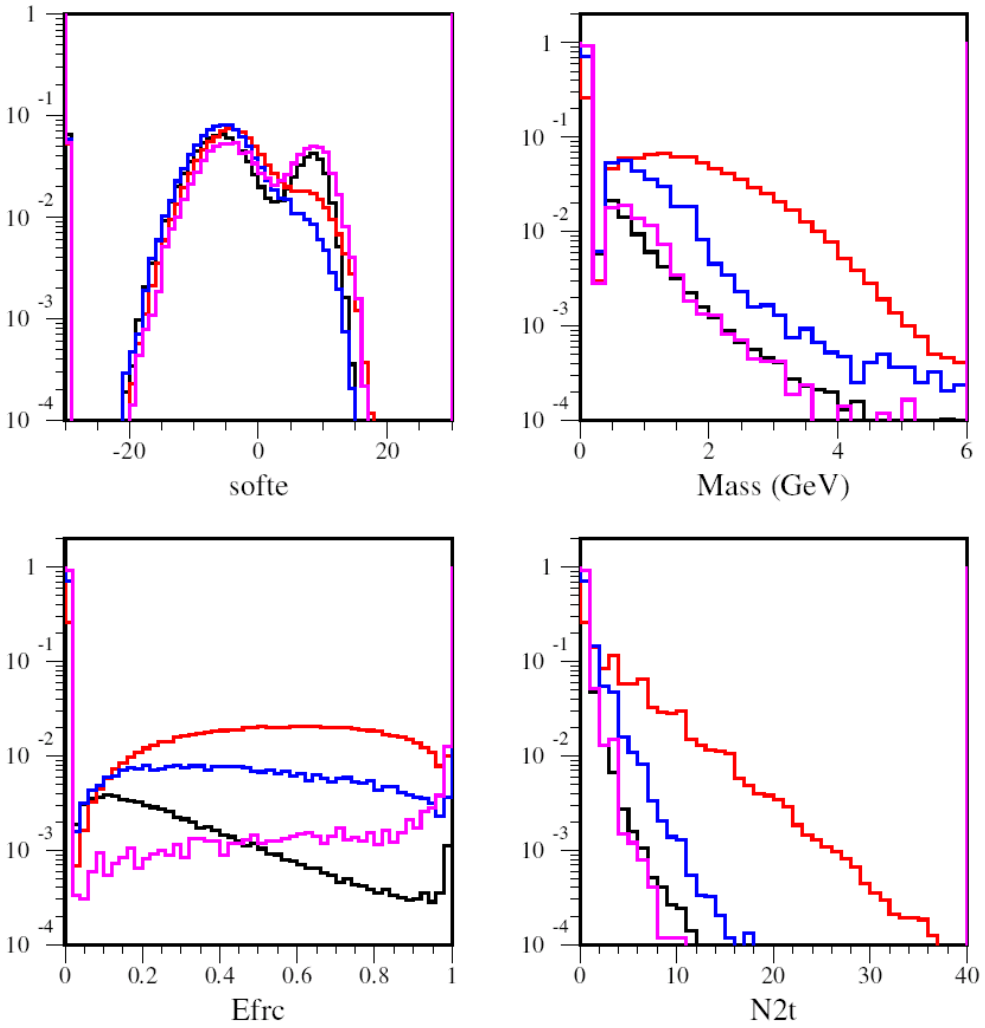




- IP2D – jet weight from transverse impact parameters
- IP3D – jet weight from 3D impact parameters
- SV1 – jet weigh from secondary vertices
- IP3D+SV1

More Existing Variables

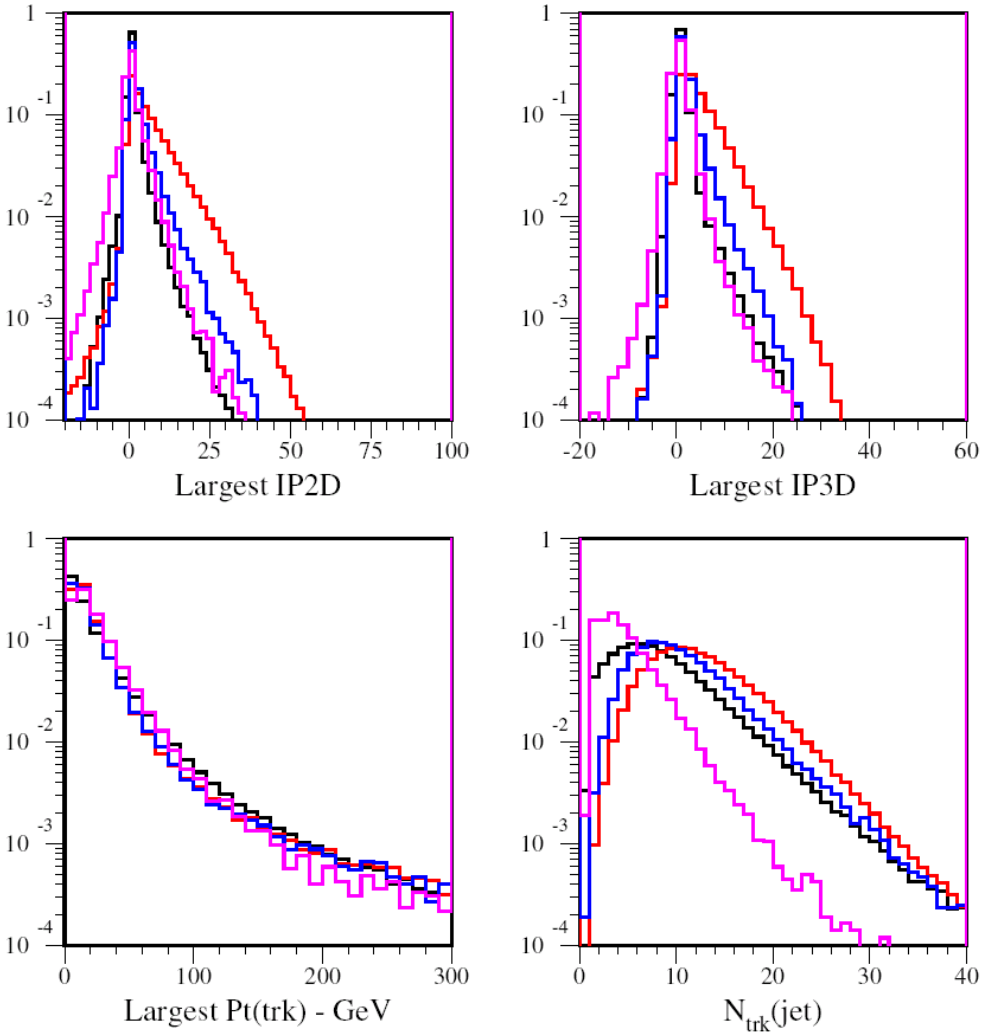
B-jet(red), C-jet(blue), τ-jet(pink), Light-jet(black)



- Softe – soft electron based tagger
- Mass – mass of particles which participate in vertex fit
- Efrc – energy ratio between particles in vertex and in jet
- N2t – No. of 2-track vertices

More Existing Variables

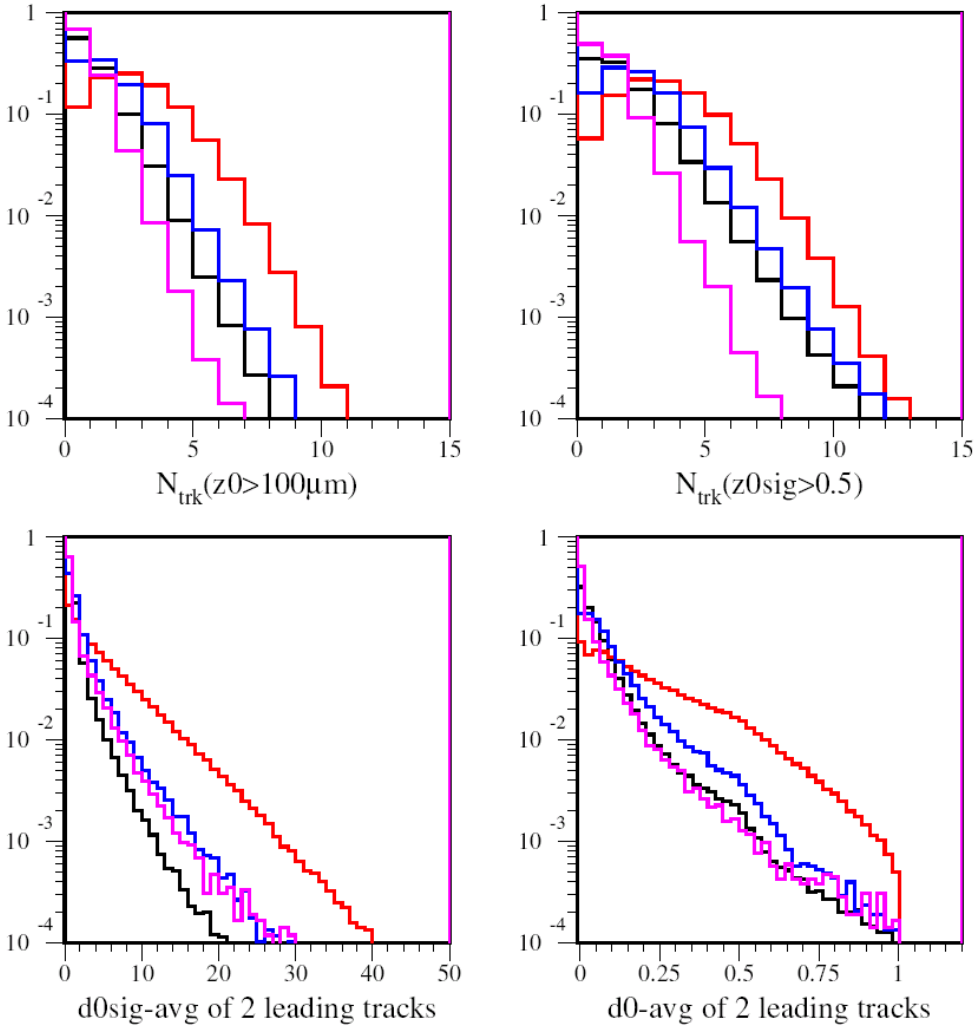
B-jet(red), C-jet(blue), τ-jet(pink), Light-jet(black)



- Largest IP2D – largest transverse IP significance of tracks in the jet
- Largest IP3D – largest longitudinal IP significance of tracks in the jet
- Largest Pt – largest transverse momentum of tracks in the jet
- Ntrk(jet) – track multiplicity in jet

New Variables

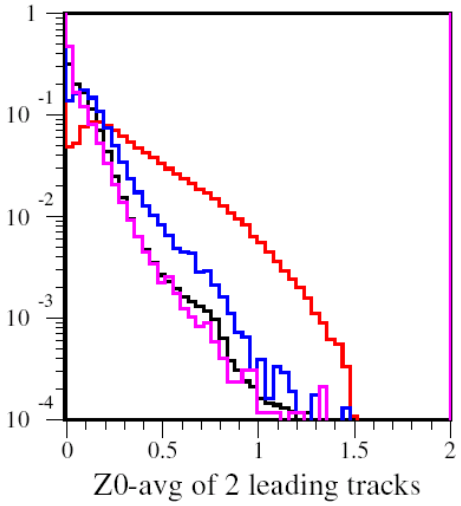
B-jet(red), C-jet(blue), τ-jet(pink), Light-jet(black)

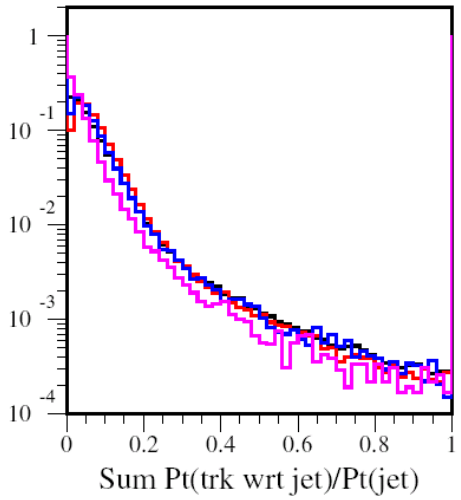


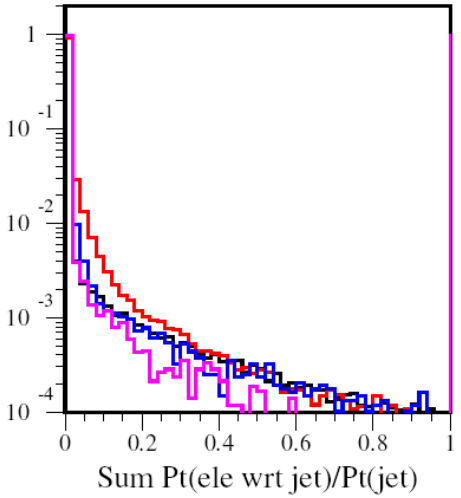
- Ntrk(z0>100microns)
- Ntrk(z0sig>0.5)
- d0sig average of 2 leading tracks with maximum d0sig
- d0 value average of 2 leading tracks with maximum d0sig

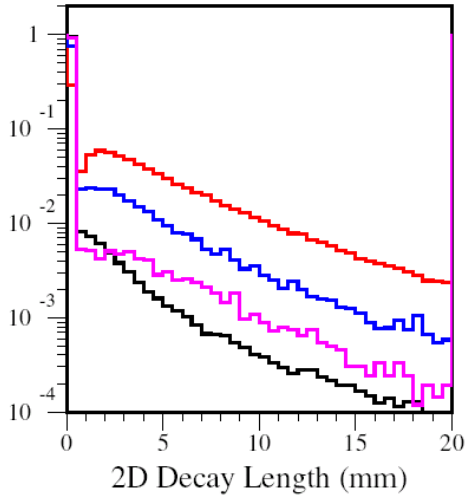
New Variables

B-jet(red), C-jet(blue), τ-jet(pink), Light-jet(black)









- Z0 value average of 2 leading tracks with maximum z0sig
- Sum Pt of tracks which are related to jet direction / E(jet)
- Sum Pt of electron which is related to jet direction / E(jet)
- 2D decay length (mm)

ATLAS B-taggers (Existing)

- IP3D+SV1 (combination of IP3D and SV1, default **weight**)

- IP2D (jet weight from 2D-transverse Impact Parameters)
- IP3D (jet weight from 3D Impact Parameters)
- JetFitter (only SV based, likelihood based)
- JetFitterComb (JetFitter combined with IP3D-SV+IP based, likelihood based)
- JetFitterNN (JetFitter standalone, NN based)
- JetFitterCombNN (JetFitter combined, NN based)

Weight = log (Prob(b-jet) / Prob(light-jet))