Re-training of the b-tagging DL1 tagger

Dominik Derendarz & Adam Trzupek
17/12/2019
HI jet weekly

Idea

- Retrain one of the official ALTAS taggers (DL1) adding FCal ΣE_T information on heavy ion data
  - ➡ Not sure how feasible would be to have different networks trained for different centrality classes
  - ➡ Possible solution would be to add variable describing the event (not a jet) centrality
  - ➡ Use the example of the DL1 training show during ATLAS tracking workshop: *https://indico.cern.ch/event/795039/sessions/303160/#20190607*
  - ➡ Btag information extracted with *FlavourTagPerformanceFramework*

Samples

- Validation sample of Pythia dijet events (JZ1-4) with b filter overlaid with 2018 data

  ➡ JIRA ATLHI-240

`mc16_5TeV.42027{1|2|3|4}.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ1_bbfilter.recon.AOD.e7383_d1521_r11472`

  ➡ Signal only sample JZ1-4 (1k per slice)

`mc16_valid.42027{1|2|3|4}.Pythia8EvtGen_A14NNPDF23LO_jetjet_JZ4_bbfilter.recon.AOD.e7383_s3428_r11320`

- Sample of Pythia+Powheg $t\bar{t}$ events (40k events)

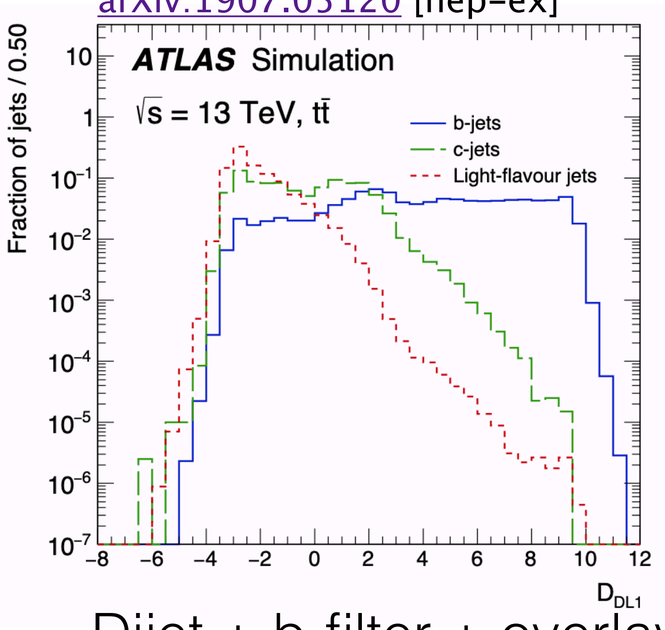
`mc16_13TeV.410470.PhPy8EG_A14_ttbar_hdamp258p75_nonallhad.deriv.DAOD_FTAG1.e6337_s3126_r10201_p3703`

DL1 tagger

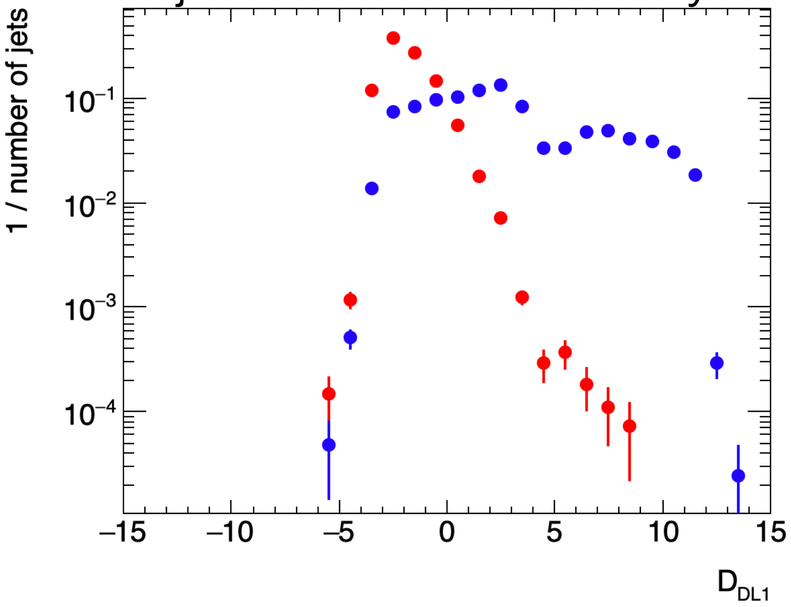
Low level taggers

Input	Variable	Description
Kinematics	p_T	Jet p_T
	η	Jet $\lvert\eta\rvert$
IP2D/IP3D	$\log(P_b/P_{light})$	Likelihood ratio between the b-jet and light-flavour jet hypotheses
	$\log(P_b/P_c)$	Likelihood ratio between the b- and c-jet hypotheses
	$\log(P_c/P_{light})$	Likelihood ratio between the c-jet and light-flavour jet hypotheses
SV1	$m(SV)$	Invariant mass of tracks at the secondary vertex assuming pion mass
	$f_E(SV)$	Energy fraction of the tracks associated with the secondary vertex
	$N_{TrkAtVtx}(SV)$	Number of tracks used in the secondary vertex
	$N_{2TrkVtx}(SV)$	Number of two-track vertex candidates
	$L_{xy}(SV)$	Transverse distance between the primary and secondary vertex
	$L_{xyz}(SV)$	Distance between the primary and the secondary vertex
	$S_{xyz}(SV)$	Distance between the primary and the secondary vertex divided by its uncertainty
	$\Delta R(\vec{p}_{jet}, \vec{p}_{vtx})(SV)$	ΔR between the jet axis and the direction of the secondary vertex relative to the primary vertex.
JetFitter	$m(JF)$	Invariant mass of tracks from displaced vertices
	$f_E(JF)$	Energy fraction of the tracks associated with the displaced vertices
	$\Delta R(\vec{p}_{jet}, \vec{p}_{vtx})(JF)$	ΔR between jet axis and vectorial sum of momenta of all tracks attached to displaced vertices
	$S_{xyz}(JF)$	Significance of average distance between PV and displaced vertices
	$N_{TrkAtVtx}(JF)$	Number of tracks from multi-prong displaced vertices
	$N_{2TrkVtx}(JF)$	Number of two-track vertex candidates (prior to decay chain fit)
	$N_{1\text{-trk vertices}}(JF)$	Number of single-prong displaced vertices
	$N_{\geq 2\text{-trk vertices}}(JF)$	Number of multi-prong displaced vertices
JetFitter c-tagging	$L_{xyz}(2^{nd}/3^{rd}vtx)(JF)$	Distance of 2^{nd} or 3^{rd} vertex from PV
	$L_{xy}(2^{nd}/3^{rd}vtx)(JF)$	Transverse displacement of the 2^{nd} or 3^{rd} vertex
	$m_{Trk}(2^{nd}/3^{rd}vtx)(JF)$	Invariant mass of tracks associated with 2^{nd} or 3^{rd} vertex
	$E_{Trk}(2^{nd}/3^{rd}vtx)(JF)$	Energy fraction of the tracks associated with 2^{nd} or 3^{rd} vertex
	$f_E(2^{nd}/3^{rd}vtx)(JF)$	Fraction of charged jet energy in 2^{nd} or 3^{rd} vertex
	$N_{TrkAtVtx}(2^{nd}/3^{rd}vtx)(JF)$	Number of tracks associated with 2^{nd} or 3^{rd} vertex
	$Y_{trk}^{min}, Y_{trk}^{max}, Y_{trk}^{avg}(2^{nd}/3^{rd}vtx)(JF)$	Min., max. and avg. track rapidity of tracks at 2^{nd} or 3^{rd} vertex

arXiv:1907.05120 [hep-ex]

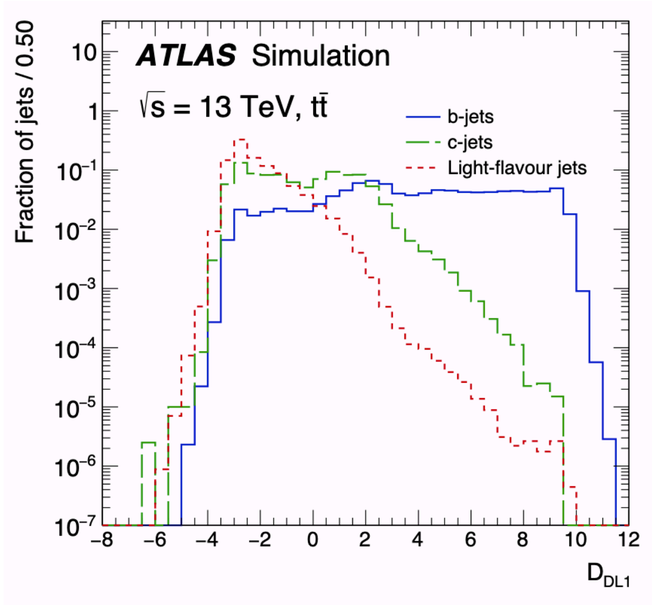


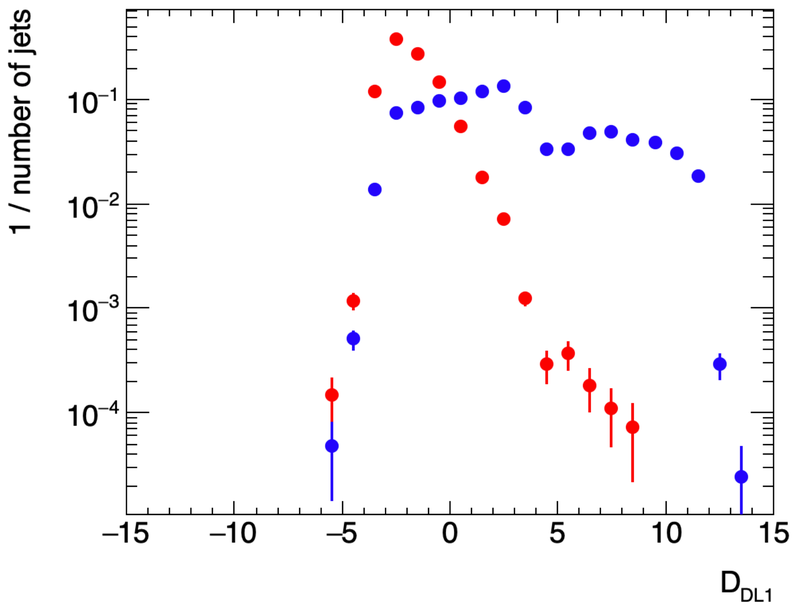
Dijet + b-filter + overlay



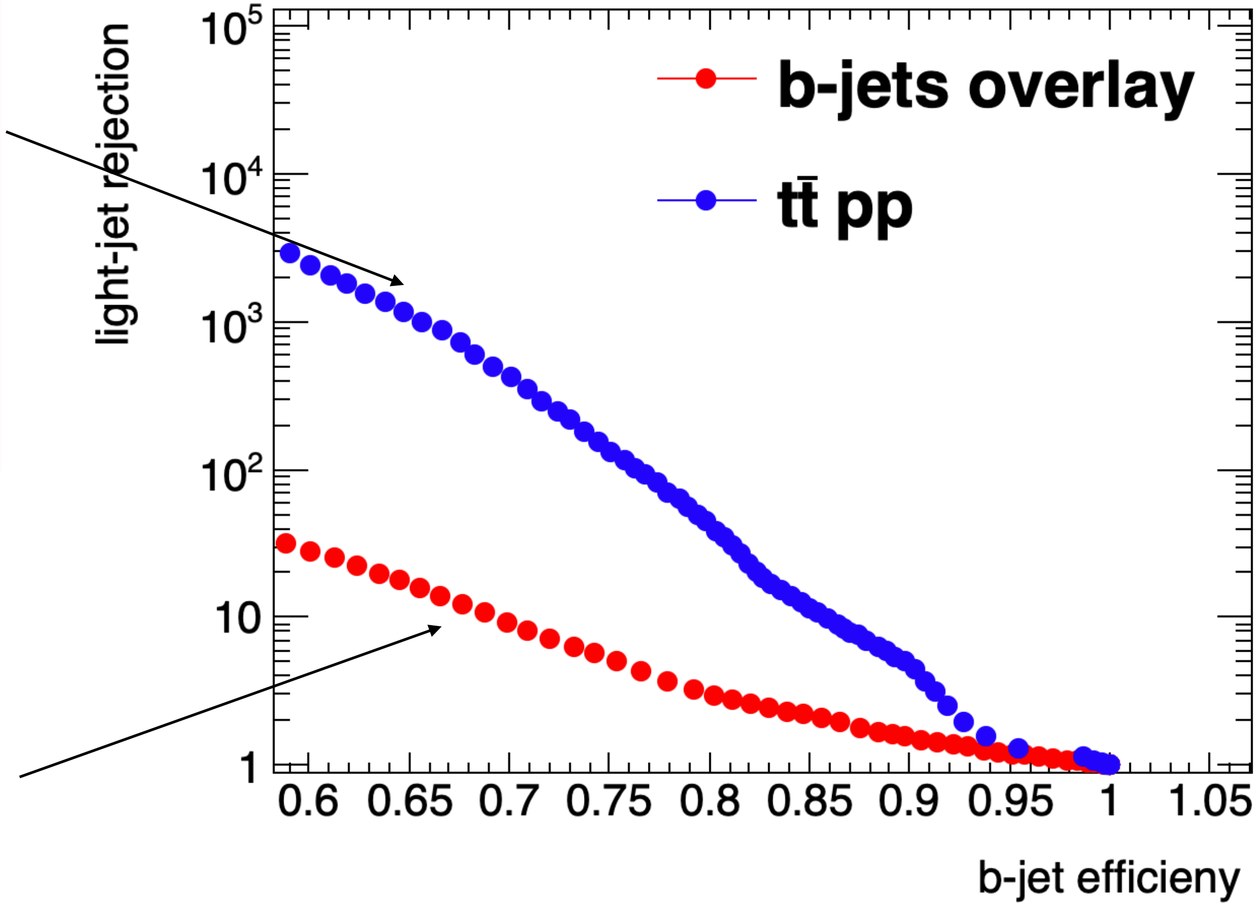
$$D_{\mathrm{DL1}} = \ln\left(\frac{p_b}{f_c \cdot p_c + (1 - f_c) \cdot p_{\mathrm{light}}}\right)$$

DL1 tagger





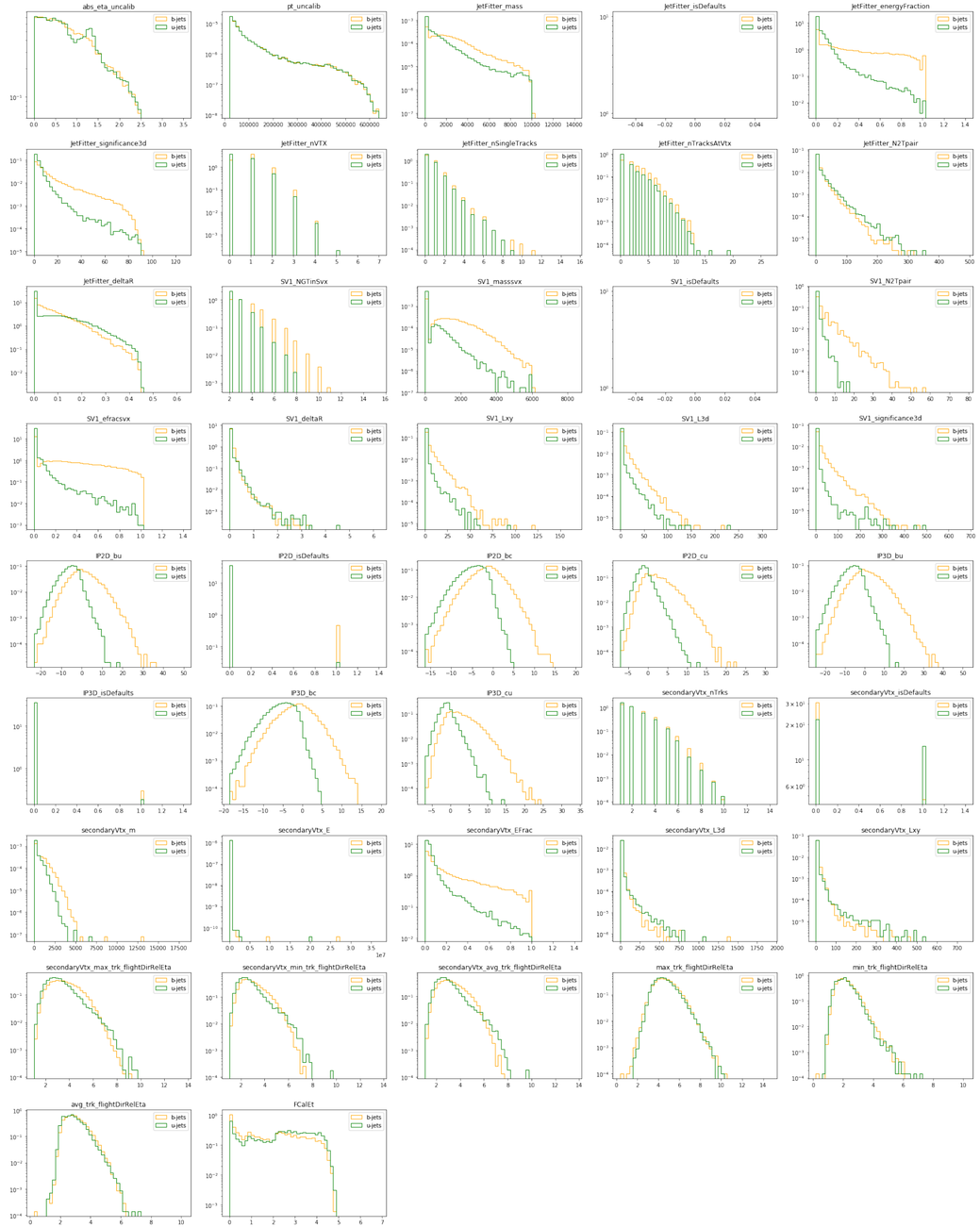
Dijet + b-filter + overlay



(different jet collections)

DL1 training

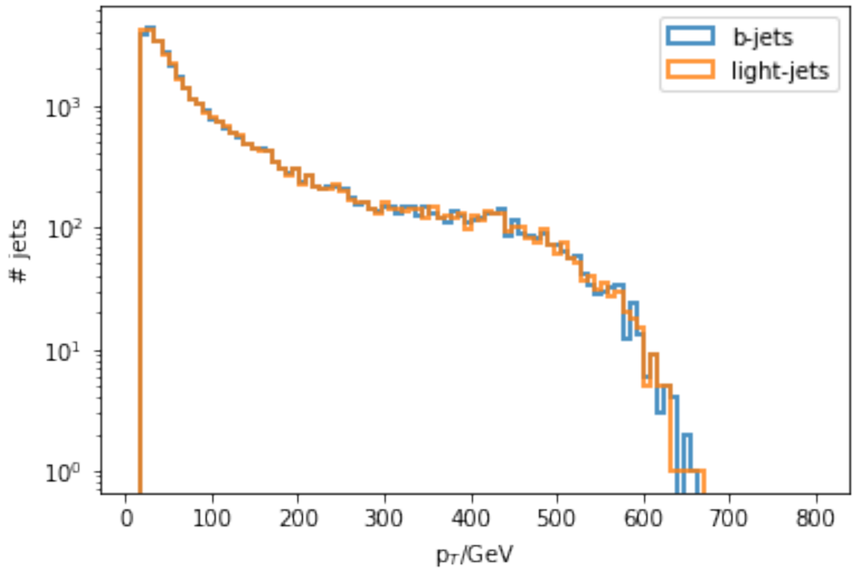
Input	Variable	Description
Kinematics	p_T	Jet p_T
	η	Jet $\lvert\eta\rvert$
IP2D/IP3D	$\log(P_b/P_{\text{light}})$	Likelihood ratio between the b-jet and light-flavour jet hypotheses
	$\log(P_b/P_c)$	Likelihood ratio between the b- and c-jet hypotheses
	$\log(P_c/P_{\text{light}})$	Likelihood ratio between the c-jet and light-flavour jet hypotheses
SV1	$m(\text{SV})$	Invariant mass of tracks at the secondary vertex assuming pion mass
	$f_E(\text{SV})$	Energy fraction of the tracks associated with the secondary vertex
	$N_{\text{TrkAtVtx}}(\text{SV})$	Number of tracks used in the secondary vertex
	$N_{\text{2TrkVtx}}(\text{SV})$	Number of two-track vertex candidates
	$L_{xy}(\text{SV})$	Transverse distance between the primary and secondary vertex
	$L_{xyz}(\text{SV})$	Distance between the primary and the secondary vertex
	$S_{xyz}(\text{SV})$	Distance between the primary and the secondary vertex divided by its uncertainty
	$\Delta R(\vec{p}_{\text{jet}}, \vec{p}_{\text{vtx}})(\text{SV})$	ΔR between the jet axis and the direction of the secondary vertex relative to the primary vertex.
JetFitter	$m(\text{JF})$	Invariant mass of tracks from displaced vertices
	$f_E(\text{JF})$	Energy fraction of the tracks associated with the displaced vertices
	$\Delta R(\vec{p}_{\text{jet}}, \vec{p}_{\text{vtx}})(\text{JF})$	ΔR between jet axis and vectorial sum of momenta of all tracks attached to displaced vertices
	$S_{xyz}(\text{JF})$	Significance of average distance between PV and displaced vertices
	$N_{\text{TrkAtVtx}}(\text{JF})$	Number of tracks from multi-prong displaced vertices
	$N_{\text{2TrkVtx}}(\text{JF})$	Number of two-track vertex candidates (prior to decay chain fit)
	$N_{\text{1-trk vertices}}(\text{JF})$	Number of single-prong displaced vertices
	$N_{\geq\text{2-trk vertices}}(\text{JF})$	Number of multi-prong displaced vertices
JetFitter c-tagging	$L_{xyz}(2^{\text{nd}}/3^{\text{rd}}\text{vtx})(\text{JF})$	Distance of 2^{nd} or 3^{rd} vertex from PV
	$L_{xy}(2^{\text{nd}}/3^{\text{rd}}\text{vtx})(\text{JF})$	Transverse displacement of the 2^{nd} or 3^{rd} vertex
	$m_{\text{Trk}}(2^{\text{nd}}/3^{\text{rd}}\text{vtx})(\text{JF})$	Invariant mass of tracks associated with 2^{nd} or 3^{rd} vertex
	$E_{\text{Trk}}(2^{\text{nd}}/3^{\text{rd}}\text{vtx})(\text{JF})$	Energy fraction of the tracks associated with 2^{nd} or 3^{rd} vertex
	$f_E(2^{\text{nd}}/3^{\text{rd}}\text{vtx})(\text{JF})$	Fraction of charged jet energy in 2^{nd} or 3^{rd} vertex
	$N_{\text{TrkAtVtx}}(2^{\text{nd}}/3^{\text{rd}}\text{vtx})(\text{JF})$	Number of tracks associated with 2^{nd} or 3^{rd} vertex
	$Y^{\text{min}}_{\text{trk}}, Y^{\text{max}}_{\text{trk}}, Y^{\text{avg}}_{\text{trk}}(2^{\text{nd}}/3^{\text{rd}}\text{vtx})(\text{JF})$	Min., max. and avg. track rapidity of tracks at 2^{nd} or 3^{rd} vertex

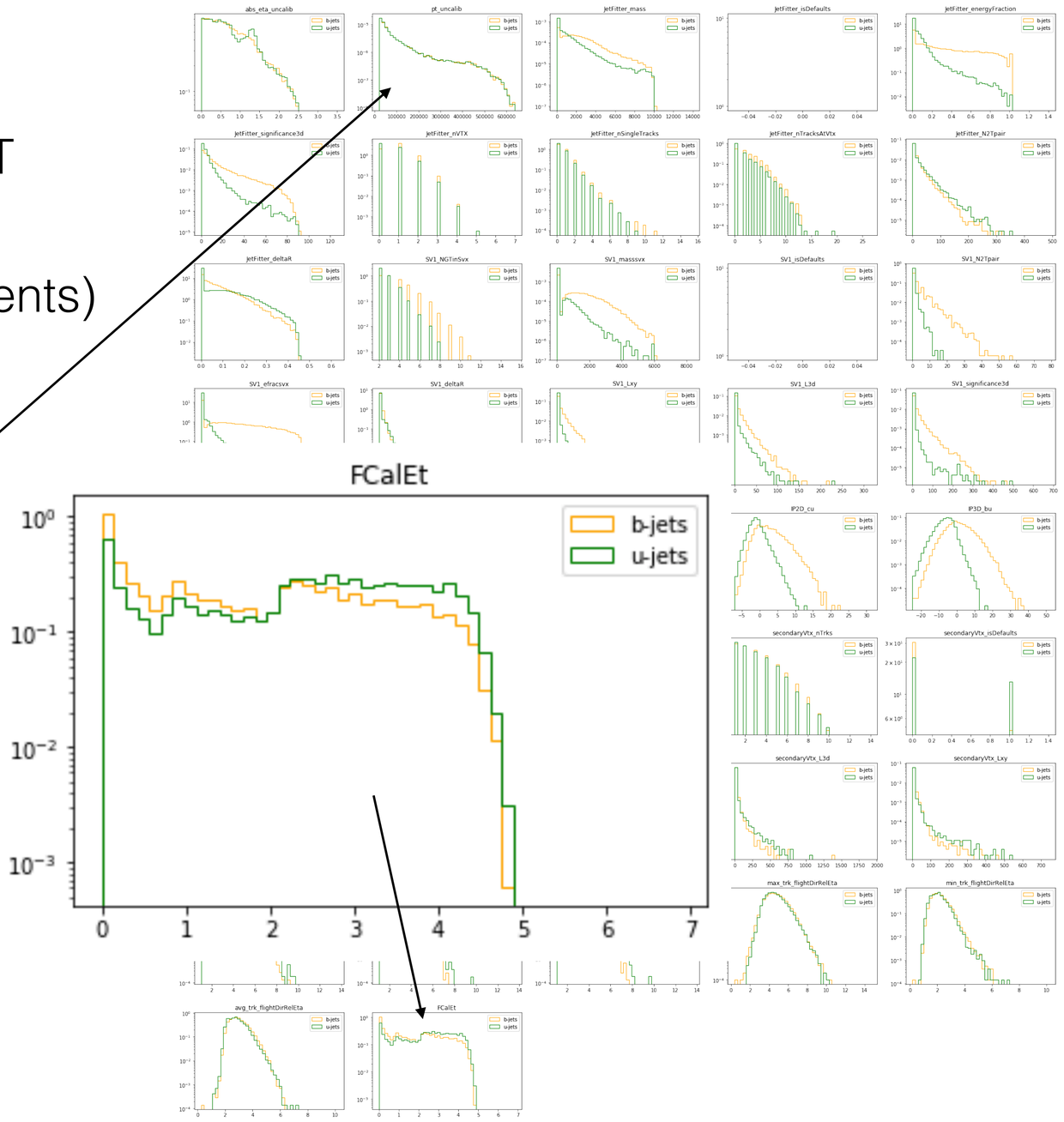


DL1 training

Light and b-jets pT (downsample light jets pT spectra to match b-jets)

➡ ~35k jets (both b & light) (even number events)



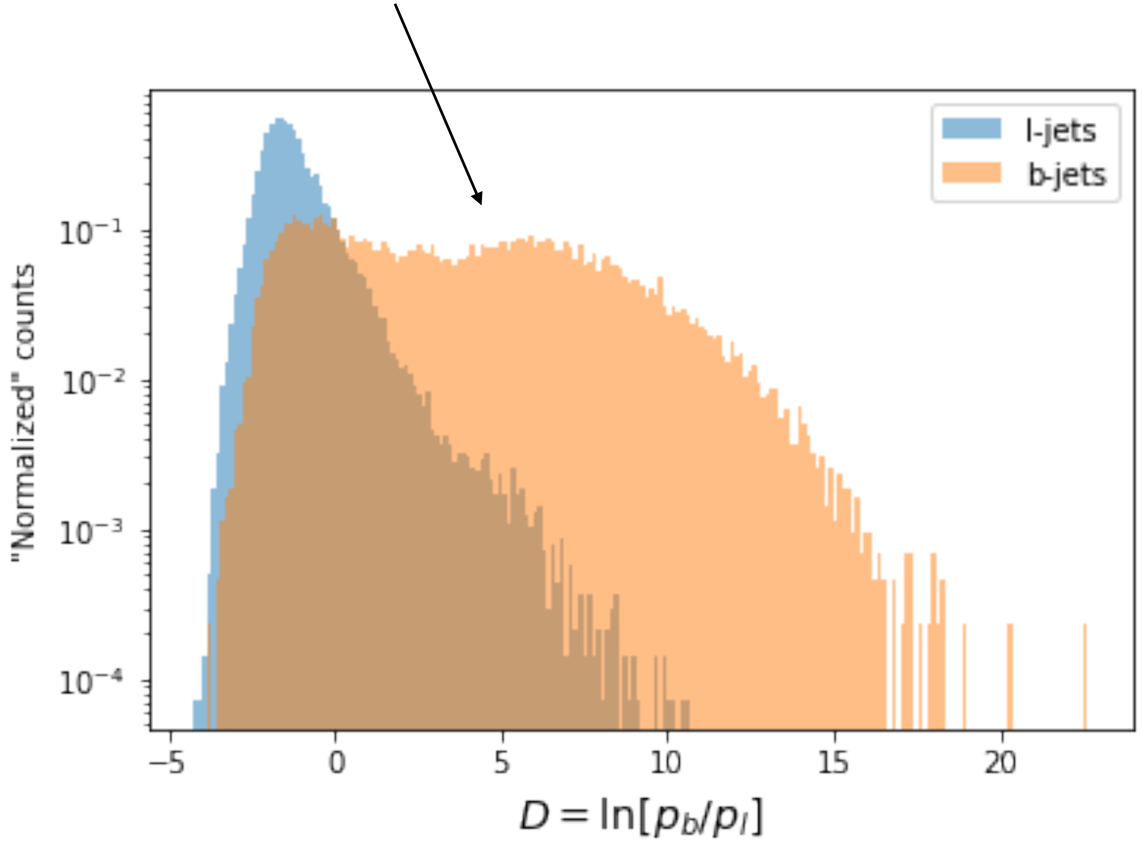


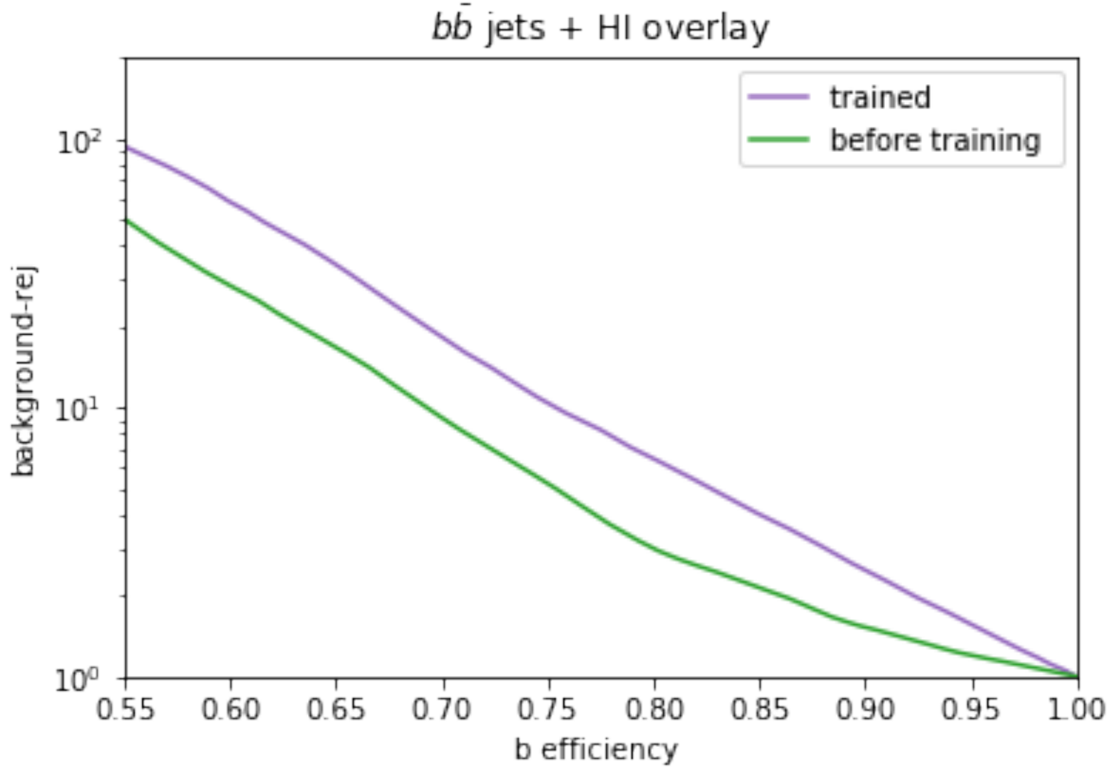
DL1 training

Light and b-jets pT (downsample light jets pT spectra to match b-jets)

➡ ~35k jets (both b & light) (even number events)

➡ tested on ~130k jets (odd number events)



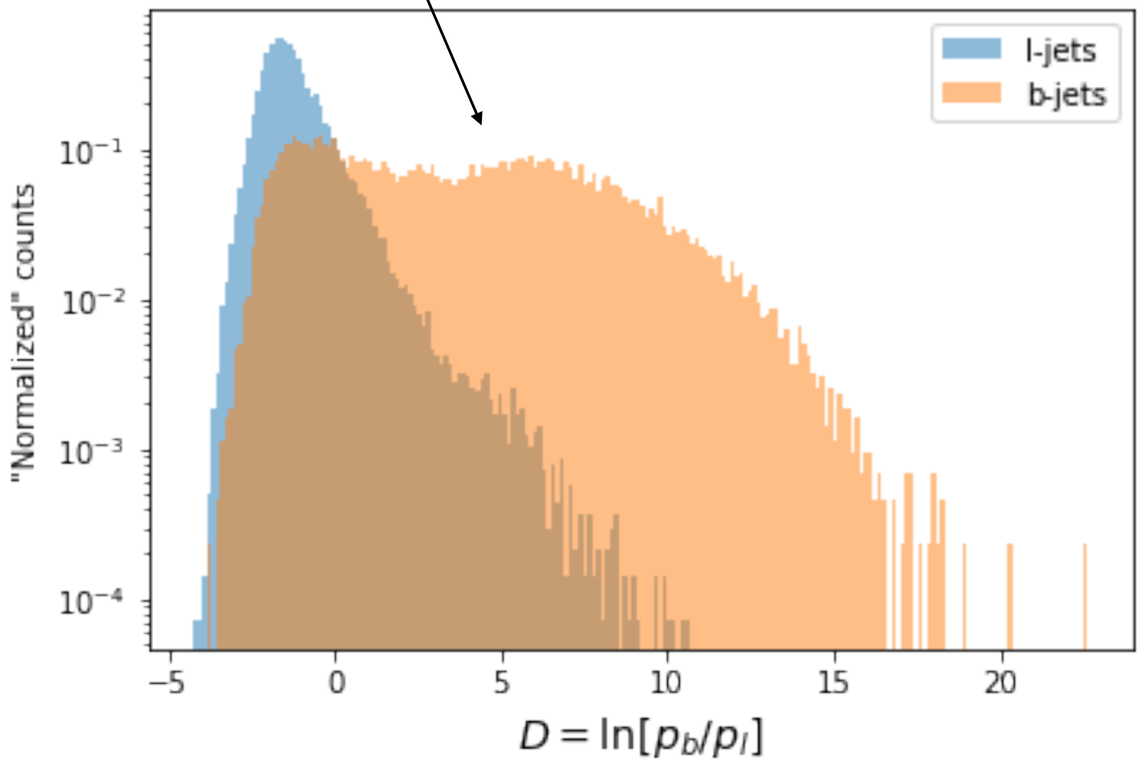


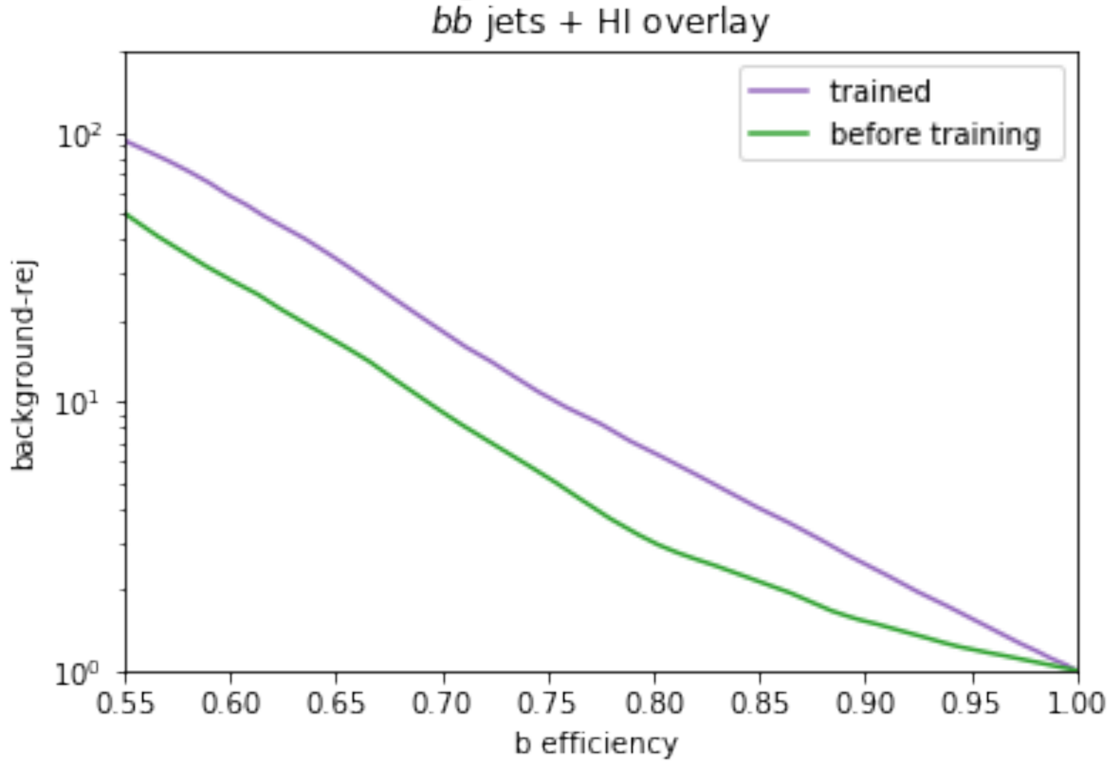
DL1 training

Light and b-jets pT (downsample light jets pT spectra to match b-jets)

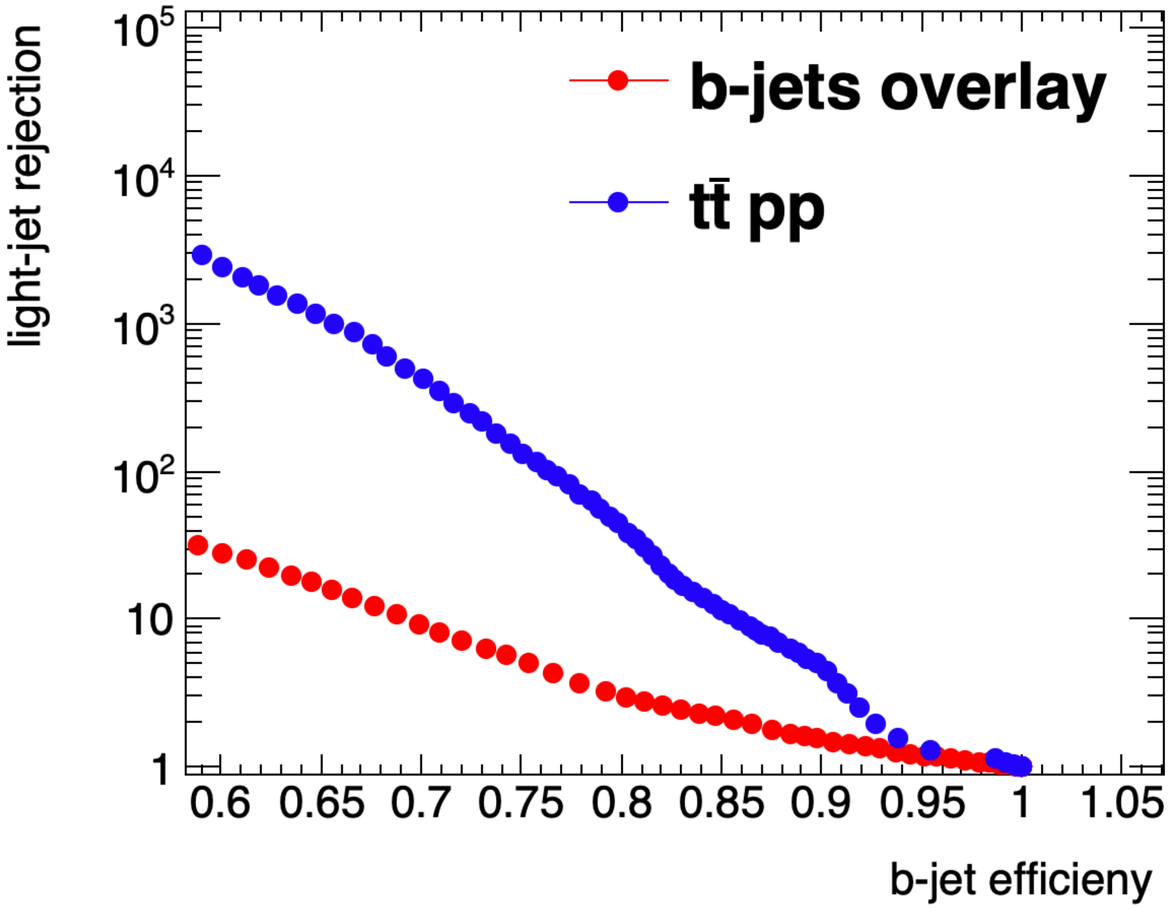
➡ ~35k jets (both b & light) (even number events)

➡ tested on ~130k jets (odd number events)

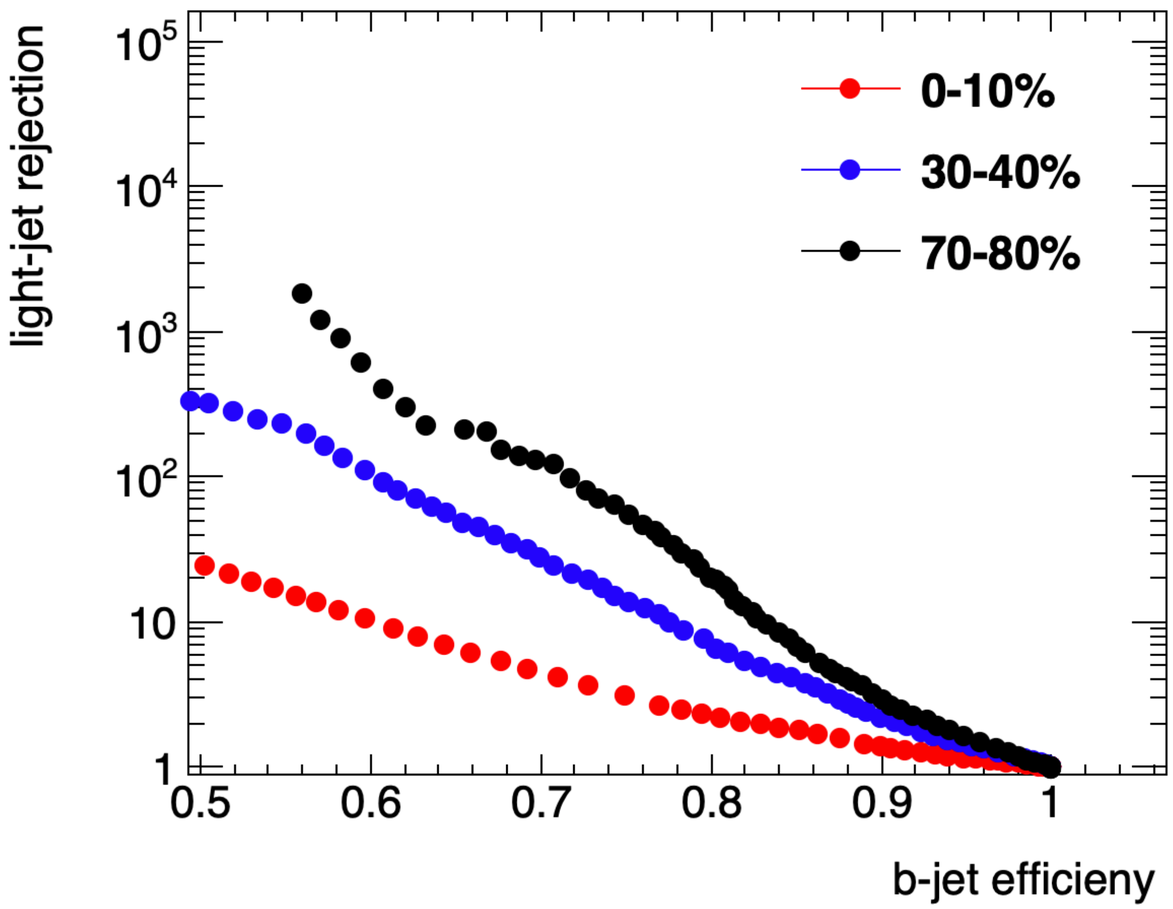




DL1 tagger - centrality



Dijet + b-filter + overlay (default)

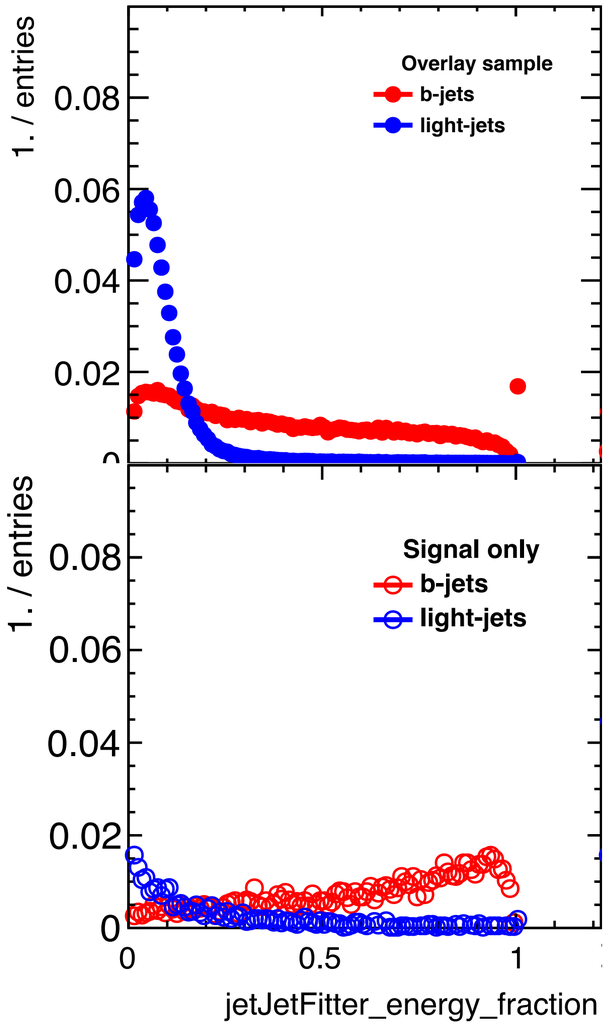


Improvement comes form proper treatment of the centrality dependence of the tagger inputs.

DL1 tagger - inputs

JetFitter energy fraction

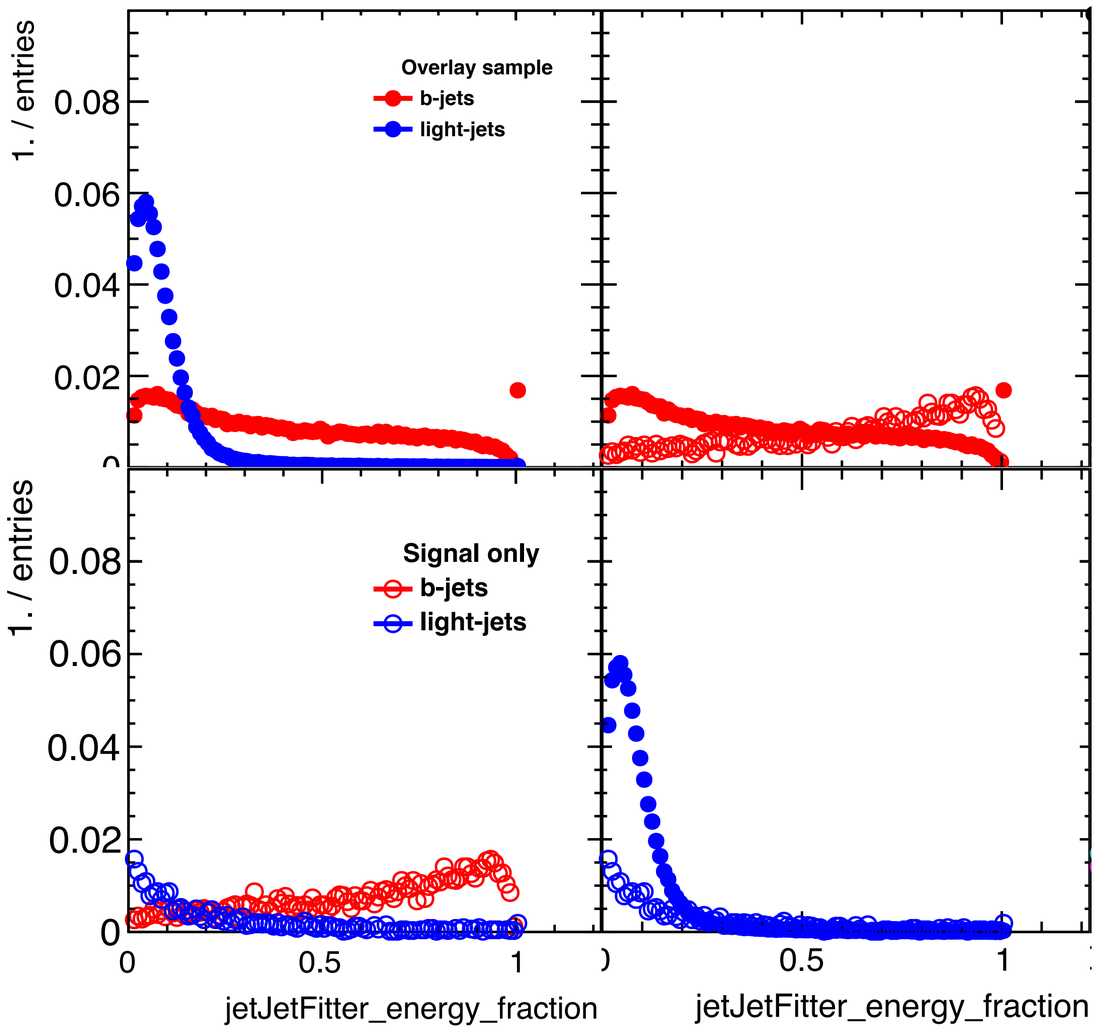
Distributions in overlay and signal only samples for b-jets significantly different



DL1 tagger - inputs

JetFitter energy fraction

Distributions in overlay and signal only samples for b-jets significantly different

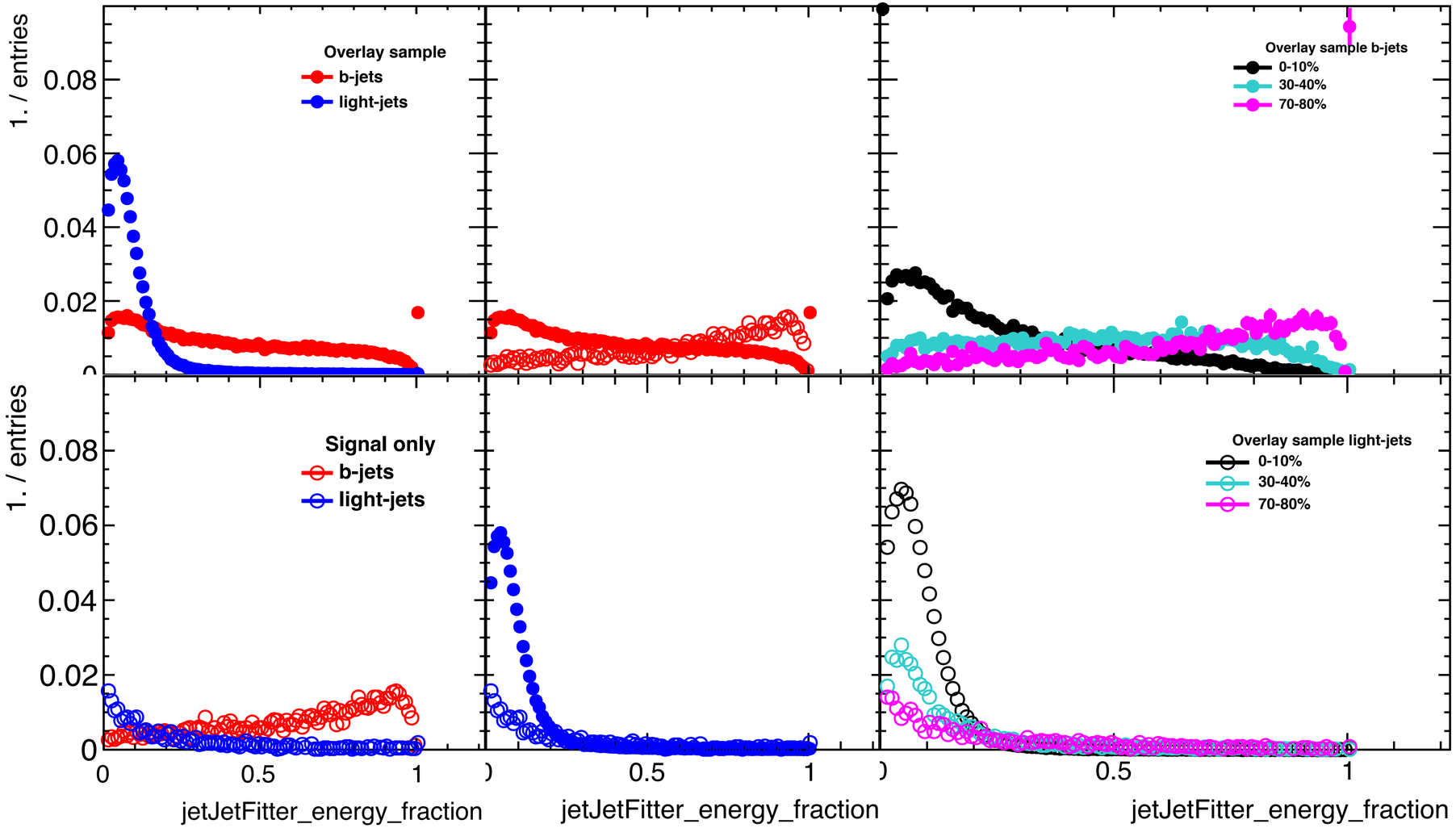


DL1 tagger - inputs

JetFitter energy fraction

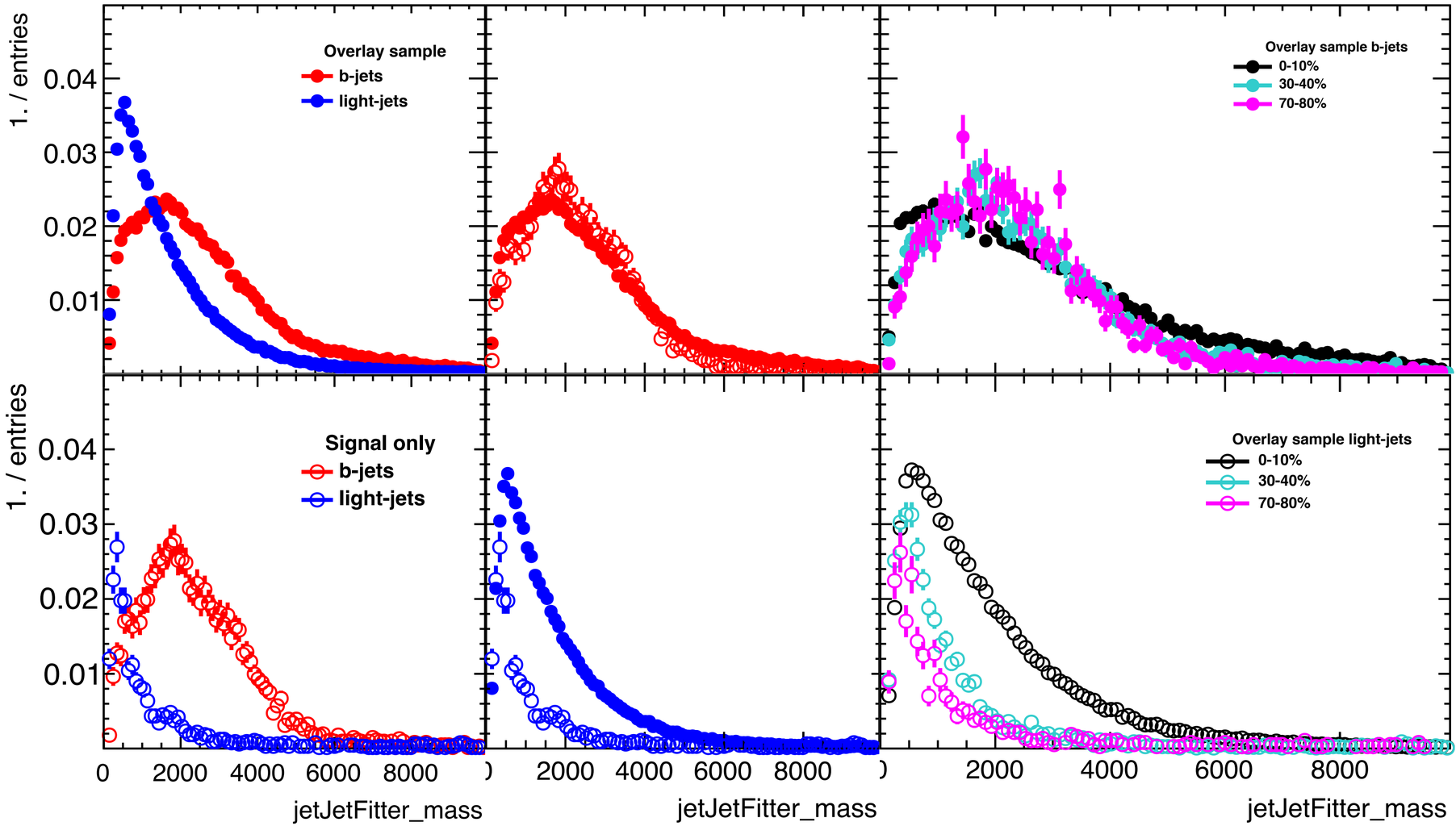
Distributions in overlay and signal only samples for b-jets significantly different

70-80% from the overlay very similar to signal only distribution for b-jets



DL1 tagger - inputs

JetFitter jet mass

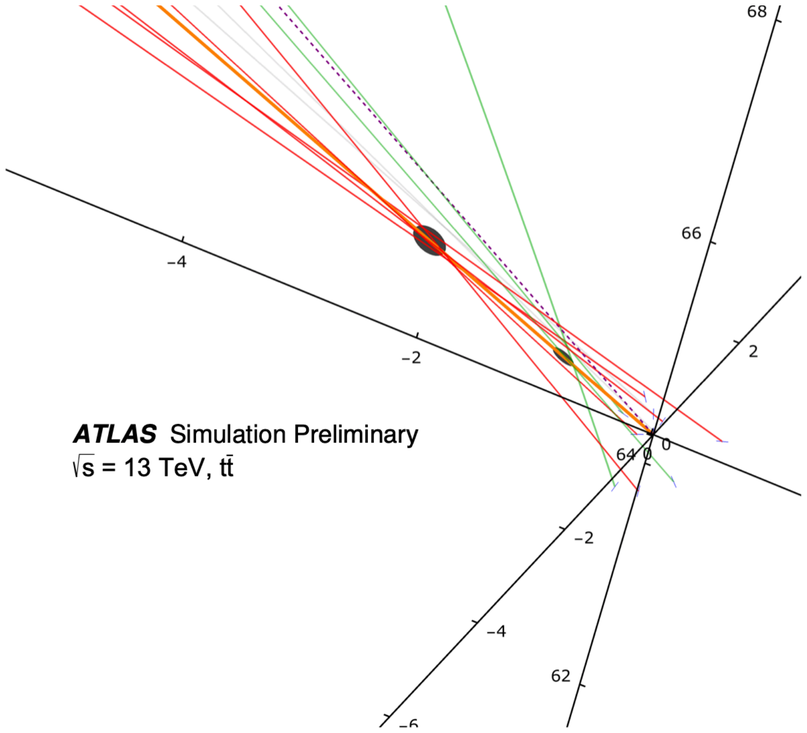


Jet fitter

ATL-PHYS-PUB-2018-025

Table 2: Fractions of selected b-jets with JetFitter vertices reconstructed in different topologies, their purity and the average generated, $< N_{\text{Bdec}} >$, and reconstructed, $< N^{\text{JF}}_{\text{Bdec}} >$, B-hadron charged decay multiplicity for JetFitter decay chains in simulated $t\bar{t}$ events.

	JF Vert. All	≥1 Single Trk 0 Multi Trk	0 Single Trk 1 Multi Trk	≥1 Single Trk 1 Multi Trk	2 Multi Trk	≥3 Multi Trk
Fraction	0.893	0.147	0.414	0.227	0.102	0.004
Purity	0.846	0.684	0.894	0.825	0.839	0.769
$< N_{\text{Bdec}} >$	4.9	3.8	4.8	5.1	6.3	7.3
$< N^{\text{JF}}_{\text{Bdec}} >$	3.0	1.1	2.9	3.7	4.9	6.0



Tracks from secondary vertex

Tracks from tertiary vertex

Fight axis

Jet fitter

ATL-PHYS-PUB-2018-025

Table 2: Fractions of selected b-jets with JetFitter vertices reconstructed in different topologies, their purity and the average generated, $< N_{\mathrm{Bdec}} >$, and reconstructed, $< N^{\mathrm{JF}}_{\mathrm{Bdec}} >$, B-hadron charged decay multiplicity for JetFitter decay chains in simulated $t\bar{t}$ events.

	JF Vert. All	≥1 Single Trk 0 Multi Trk	0 Single Trk 1 Multi Trk	≥1 Single Trk 1 Multi Trk	2 Multi Trk	≥3 Multi Trk
Fraction	0.893	0.147	0.414	0.227	0.102	0.004
Purity	0.846	0.684	0.894	0.825	0.839	0.769
$< N_{\mathrm{Bdec}} >$	4.9	3.8	4.8	5.1	6.3	7.3
$< N^{\mathrm{JF}}_{\mathrm{Bdec}} >$	3.0	1.1	2.9	3.7	4.9	6.0

Fractions of selected b-jets with JetFitter vertices reconstructed in different topologies

MC sample	Any topo.	≥ 1 Single Trk 0 Multi Trk	0 Single Trk 1 Multi Trk	≥ 1 Single Trk 1 Multi Trk	2 Multi Trk	≥ 3 Multi Trk	0 Single Trk 0 Multi Trk
pp(JZ1,2,3,4; 1K evt)	0.81	0.17	0.41	0.16	0.064	0.003	0.19
pp (overlay, JZ1,2,3,4)	0.85	0.16	0.33	0.22	0.13	0.014	0.15
pp (overlay, JZ1,2,3,4, FCalEt > 3.76 TeV)	0.93	0.11	0.19	0.30	0.28	0.046	0.07
pp (overlay, JZ1,2,3,4 FCalEt <0.056 TeV)	0.80	0.18	0.41	0.14	0.06	0.002	0.20

Signal only

Dijet + b-filter + overlay

Jet fitter

ATL-PHYS-PUB-2018-025

Table 2: Fractions of selected *b*-jets with JetFitter vertices reconstructed in different topologies, their purity and the average generated, $< N_{\mathrm{Bdec}} >$, and reconstructed, $< N^{\mathrm{JF}}_{\mathrm{Bdec}} >$, *B*-hadron charged decay multiplicity for JetFitter decay chains in simulated $t\bar{t}$ events.

	JF Vert. All	≥1 Single Trk 0 Multi Trk	0 Single Trk 1 Multi Trk	≥1 Single Trk 1 Multi Trk	2 Multi Trk	≥3 Multi Trk
Fraction	0.893	0.147	0.414	0.227	0.102	0.004
Purity	0.846	0.684	0.894	0.825	0.839	0.769
$< N_{\mathrm{Bdec}} >$	4.9	3.8	4.8	5.1	6.3	7.3
$< N^{\mathrm{JF}}_{\mathrm{Bdec}} >$	3.0	1.1	2.9	3.7	4.9	6.0

Purity of tracks for selected *b*-jets with JetFitter vertices reconstructed in different topologies

	MC sample	Any topo.	≥ 1 Single Trk 0 Multi Trk	0 Single Trk 1 Multi Trk	≥ 1 Single Trk 1 Multi Trk	2 Multi Trk	≥ 3 Multi Trk	0 Single Trk 0 Multi Trk
Signal only	pp(JZ1,2,3,4; 1K evt)	0.89	0.81	0.92	0.87	0.86	0.72	0.52
Dijet + b-filter + overlay	pp (overlay, JZ1,2,3,4)	0.51	0.43	0.65	0.49	0.43	0.35	0.25
	pp (overlay, JZ1,2,3,4, FCalEt > 3.76 TeV)	0.22	0.19	0.26	0.22	0.22	0.22	0.12
	pp (overlay, JZ1,2,3,4 FCalEt <0.056 TeV)	0.89	0.79	0.92	0.87	0.89	0.84	0.46

Jet fitter

MC sample	Any topo.	≥ 1 Single Trk 0 Multi Trk	0 Single Trk 1 Multi Trk	≥ 1 Single Trk 1 Multi Trk	2 Multi Trk	≥ 3 Multi Trk	0 Single Trk 0 Multi Trk
Average generated number of tracks for selected *b*-jets with JetFitter vertices in different topologies							
pp(JZ1,2,3,4; 1K evt)	4.9	3.9	4.9	5.2	6.3	7.2	5.1
pp (overlay, JZ1,2,3,4)	4.9	4.1	4.8	5.0	5.5	5.9	5.0
pp (overlay, JZ1,2,3,4, **FCalEt > 3.76 TeV**)	4.9	4.4	4.6	4.8	5.2	5.5	4.9
pp (overlay, JZ1,2,3,4 **FCalEt <0.056 TeV**)	4.9	4.0	5.0	5.3	6.2	7.4	5.2
Average reconstructed number of tracks for selected *b*-jets with JetFitter vertices in different topologies							
pp(JZ1,2,3,4; 1K evt)	3.1	1.3	3.0	4.2	5.7	8.0	0.2
pp (overlay, JZ1,2,3,4)	4.4	2.0	3.4	5.4	7.3	10.4	0.6
pp (overlay, JZ1,2,3,4, **FCalEt > 3.76 TeV**)	7.2	3.7	4.6	7.3	9.5	12.2	1.5
pp (overlay, JZ1,2,3,4 **FCalEt <0.056 TeV**)	3.1	1.3	3.0	4.3	5.4	7.5	0.2

Signal only

Dijet + b-filter + overlay

Signal only

Dijet + b-filter + overlay

Summary and next steps

Btagging algorithms running in the heavy ion reconstruction

➡ Adding FCal ΣE_T to the DL1 tagger will make it more robust

➡ Performance in the peripheral events quite similar to performance in pp collisions

➡ Performance in the central events degrading most probably due to tracks from UE background

Next steps

➡ Try to reduce contribution from UE tracks

- Higher p_T threshold was shown in the past to be useful
- Rerunning of the low level algorithms should be possible in derivation step

➡ Use larger sample for training

- Inclusive dijet sample

➡ Validate on data

- Performance of algorithm on the light jets should be straight forward
- Use muon tagged jets as a source of b jets