Charm-tagging

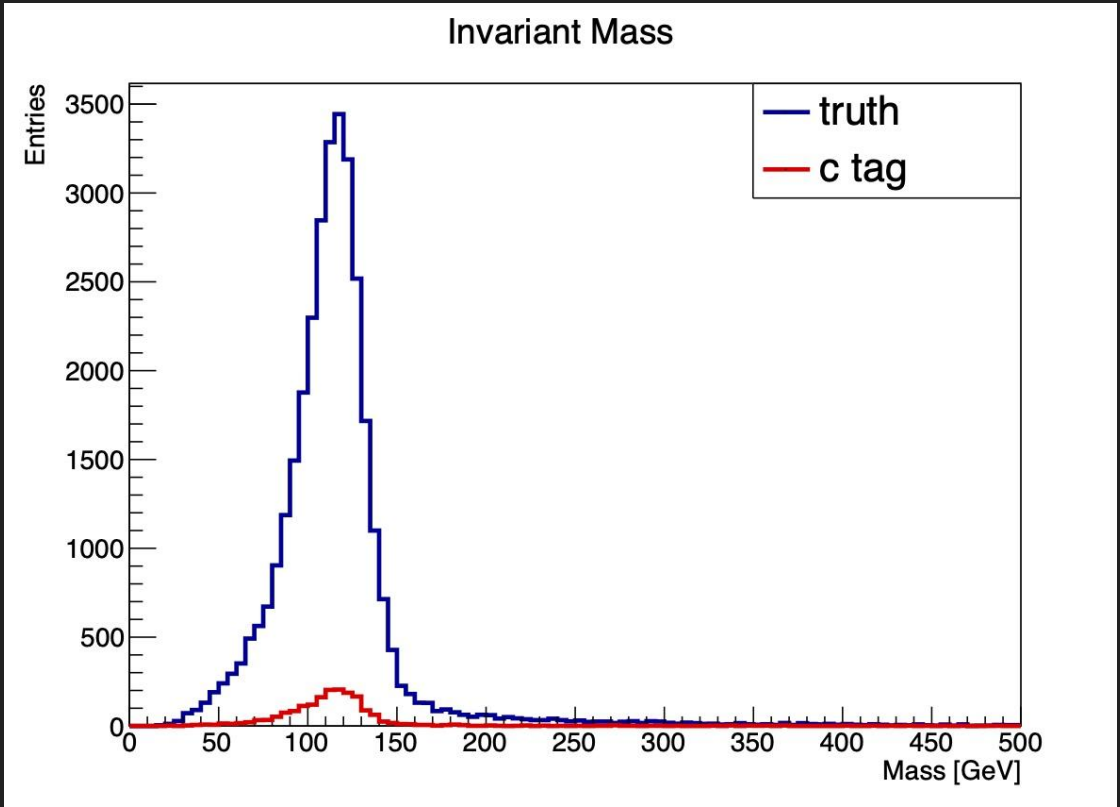
Bennett Austin

Introduction

- c-tagging variables study
  - JetFitter and SV1 variables used in DL1 algorithm
- Data samples used:
  - H->cc for charm and light jets
  - H->bb for bottom jets
- Cuts used:
  - Jet pt > 20 GeV
  - |Jet eta| < 2.5

JetFitter	$m(\mathrm{JF})$	Invariant mass of tracks from displaced vertices
	$f_E(\mathrm{JF})$	Energy fraction of the tracks associated with the displaced vertices
	$\Delta R(\vec{p}_{\mathrm{jet}}, \vec{p}_{\mathrm{vtx}})(\mathrm{JF})$	ΔR between the jet axis and the vectorial sum of momenta of all tracks attached to displaced vertices
	$S_{xyz}(\mathrm{JF})$	Significance of the average distance between PV and displaced vertices
	$N_{\mathrm{TrkAtVtx}}(\mathrm{JF})$	Number of tracks from multi-prong displaced vertices
	$N_{\mathrm{2TrkVtx}}(\mathrm{JF})$	Number of two-track vertex candidates (prior to decay chain fit)
	$N_{\mathrm{1\text{-}trk\ vertices}}(\mathrm{JF})$	Number of single-prong displaced vertices
	$N_{\geq 2\text{-}\mathrm{trk\ vertices}}(\mathrm{JF})$	Number of multi-prong displaced vertices
SV1	$m(\mathrm{SV})$	Invariant mass of tracks at the secondary vertex assuming pion mass
	$f_E(\mathrm{SV})$	Energy fraction of the tracks associated with the secondary vertex
	$N_{\mathrm{TrkAtVtx}}(\mathrm{SV})$	Number of tracks used in the secondary vertex
	$N_{\mathrm{2TrkVtx}}(\mathrm{SV})$	Number of two-track vertex candidates
	$L_{xy}(\mathrm{SV})$	Transverse distance between the primary and secondary vertex
	$L_{xyz}(\mathrm{SV})$	Distance between the primary and the secondary vertex
	$S_{xyz}(\mathrm{SV})$	Distance between the primary and the secondary vertex divided by its uncertainty
	$\Delta R(\vec{p}_{\mathrm{jet}}, \vec{p}_{\mathrm{vtx}})(\mathrm{SV})$	ΔR between the jet axis and the direction of the secondary vertex relative to the primary vertex.

Introduction

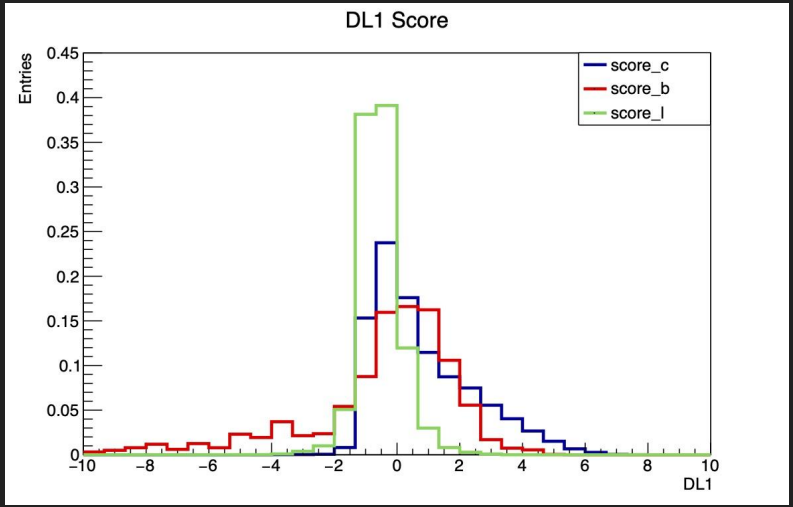


- Higgs->cc invariant mass
- Highest jet pt pairs
- Blue: truth c-jets
- Red: truth c-jets correctly tagged as c-jets by DL1

DL1 Score and efficiencies

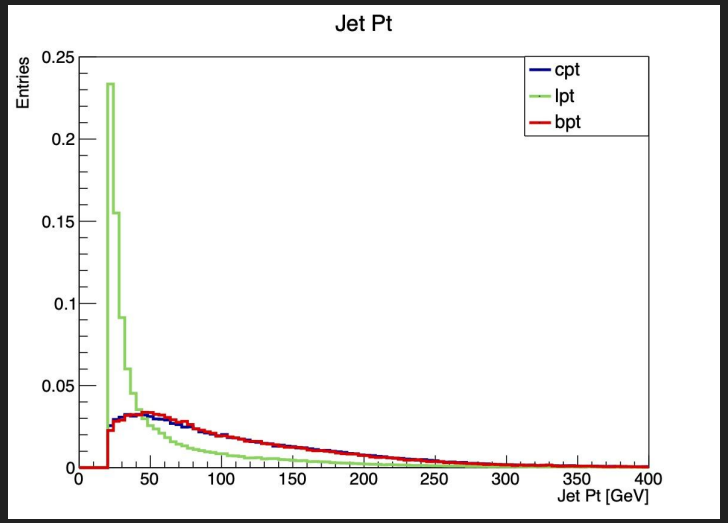
- c-jet efficiency: 19.8%
- b-jet c-tagged efficiency: 6.4%
- Light-jet c-tagged efficiency: 0.8%
- pc, pu, pb - probability of jets
- f is fraction of b-jets in background
  - f = 0.08
- Requirements for c-tag:
  - DL1 >= 1.3
  - MV2c10 <= 0.83

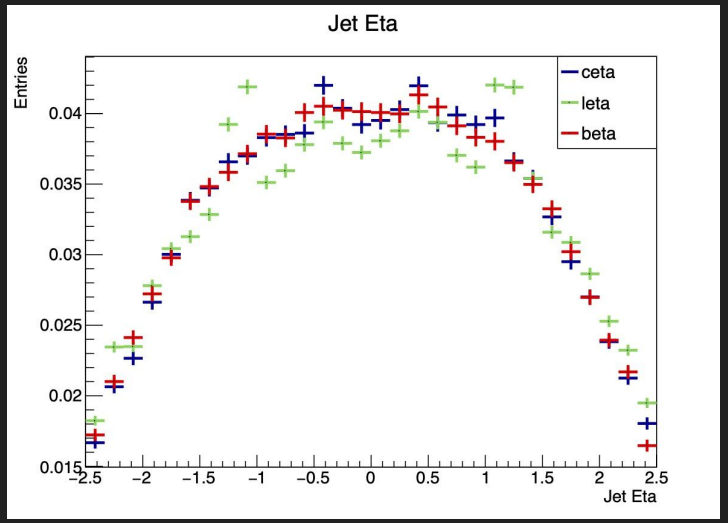
$$DL_1 = \ln \frac{p_c}{f p_b + (1-f) p_u}$$

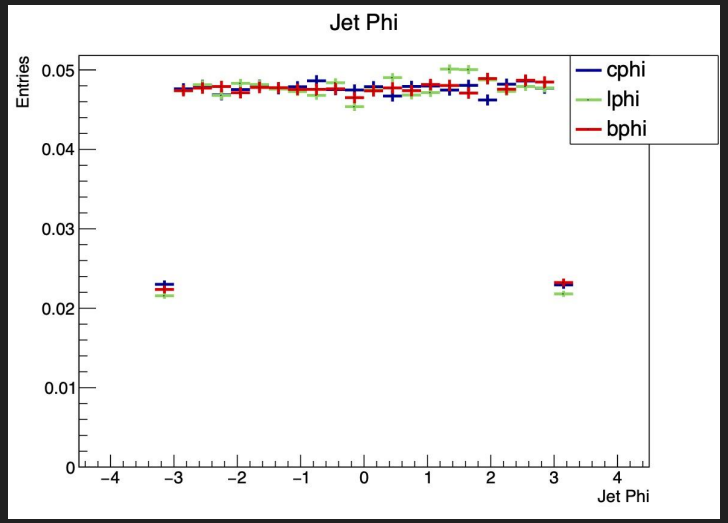


Green: light jets, red: b-jets, blue: c-jets

Jet Kinematics

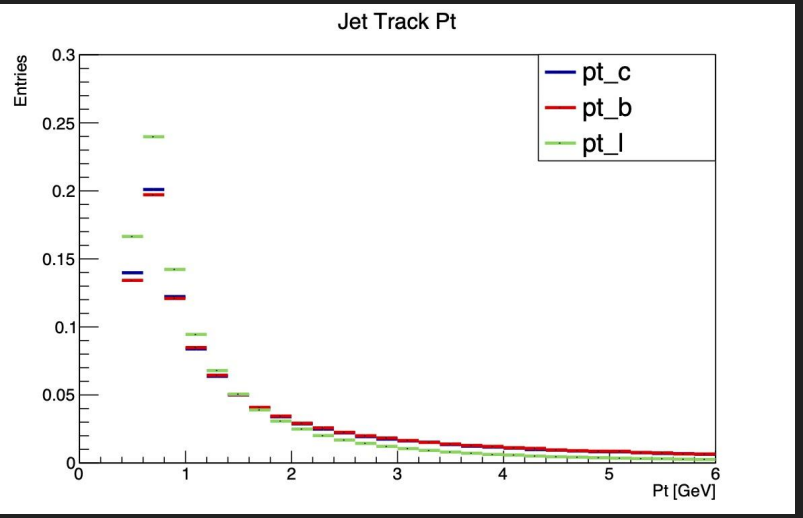


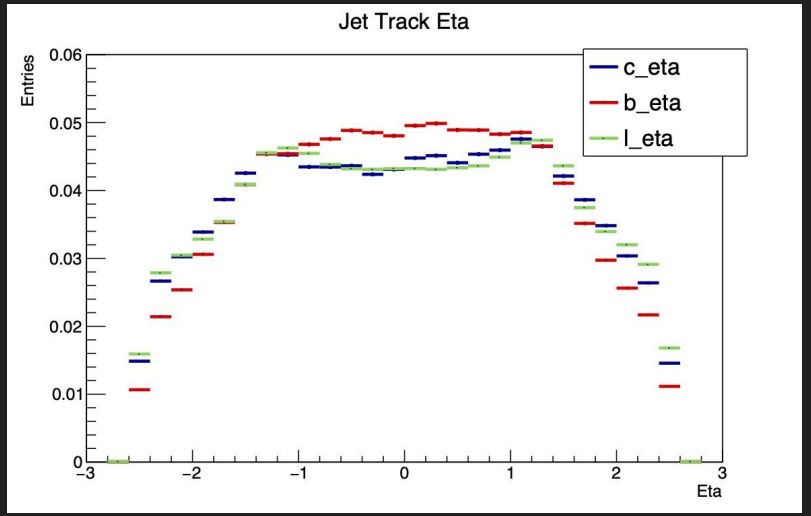


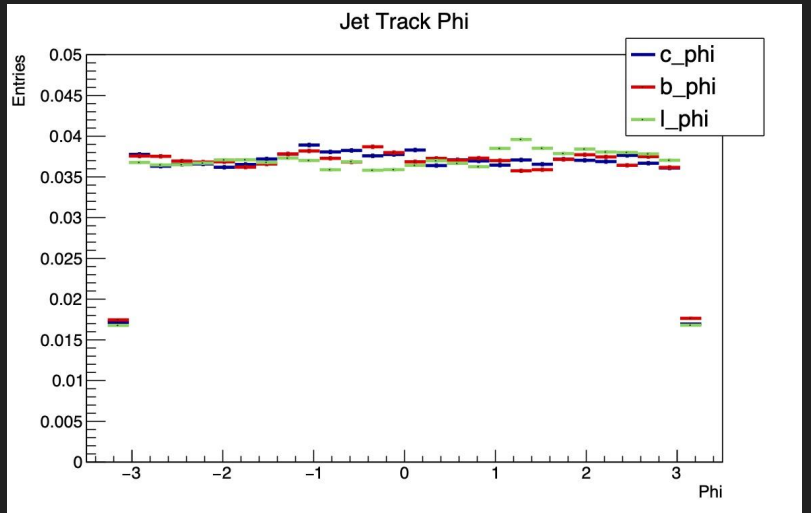


Green: light jets, red: b-jets, blue: c-jets

Track kinematics

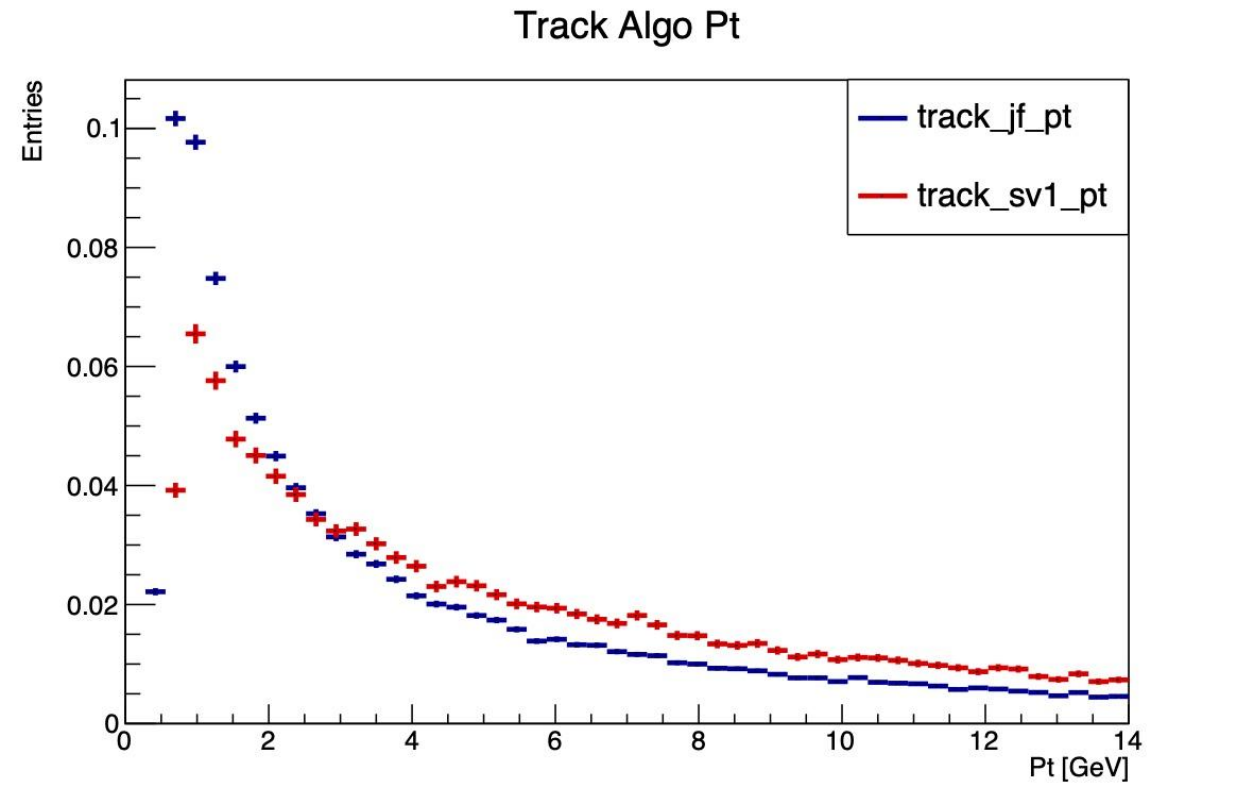






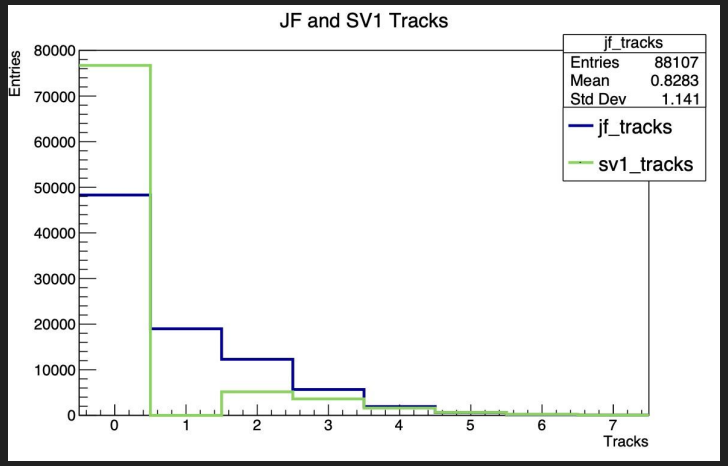
Track eta: c-jets follow a pattern similar to light jets rather than b-jets

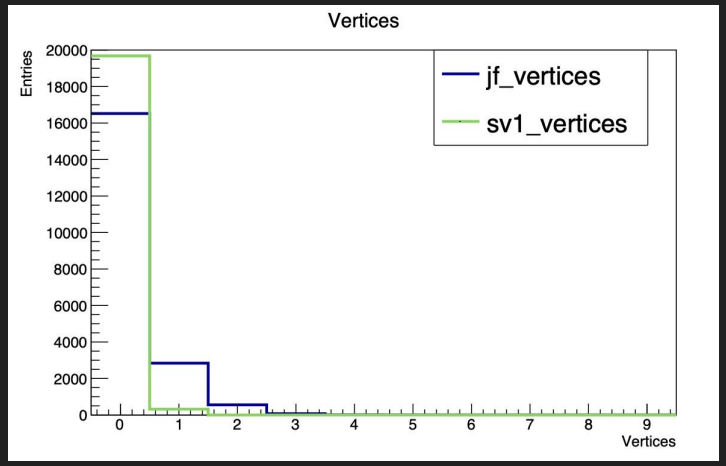
Track Pt for JF and SV1

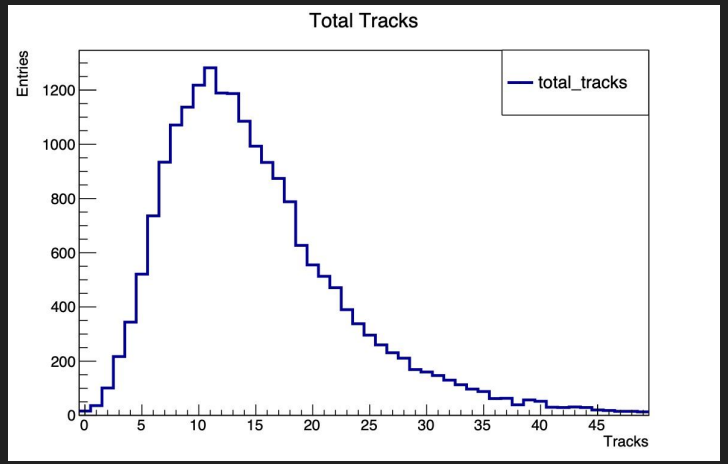


- Blue: JetFitter
- Red: SV1
- All flavours considered

Track variables

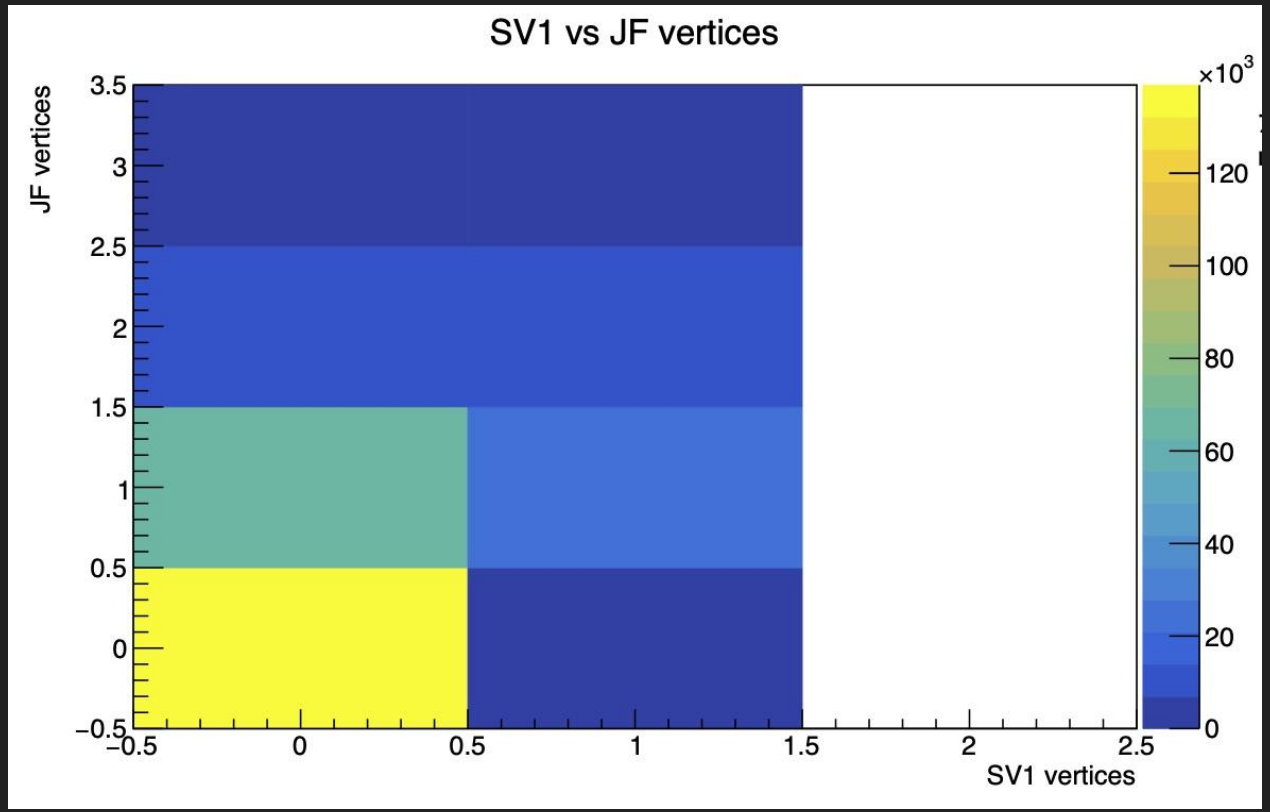






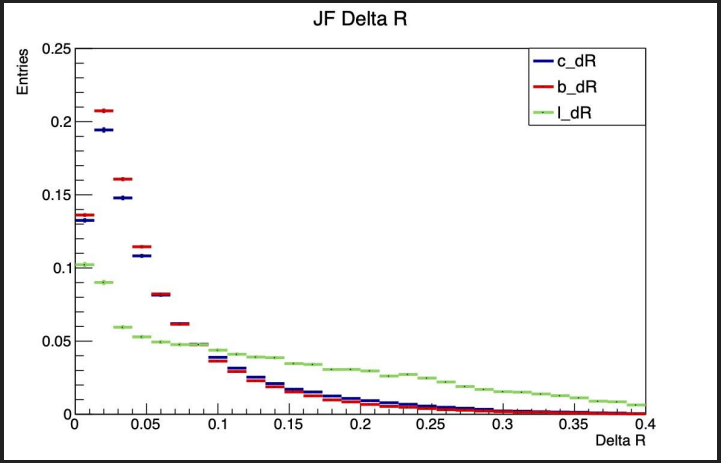
- Total tracks peak at 11
- No single track SV1 events
- No SV1 vertices > 1

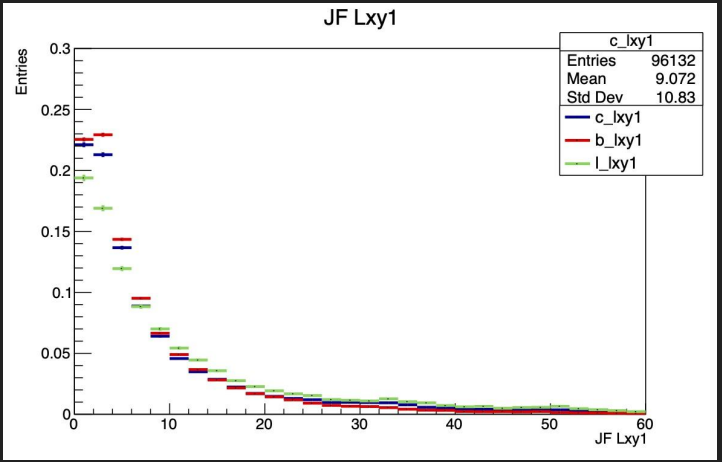
JF and SV1 vertices

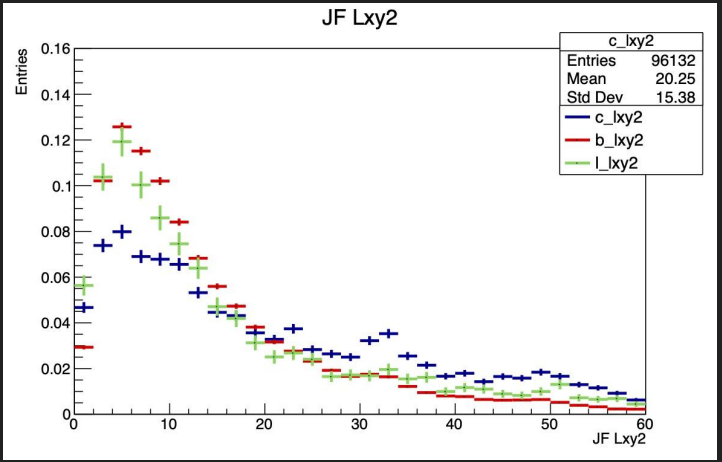


2D Plot of previous vertex graph

JF variables

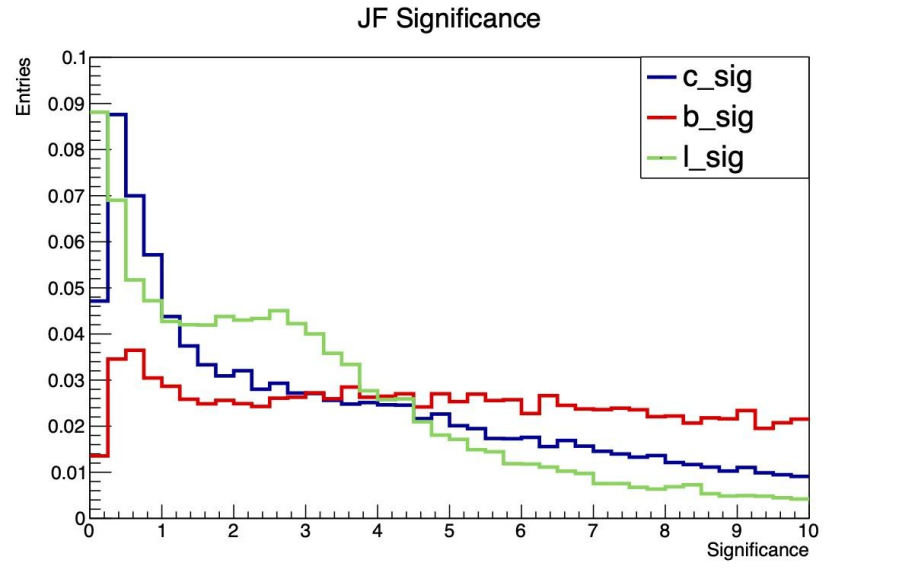


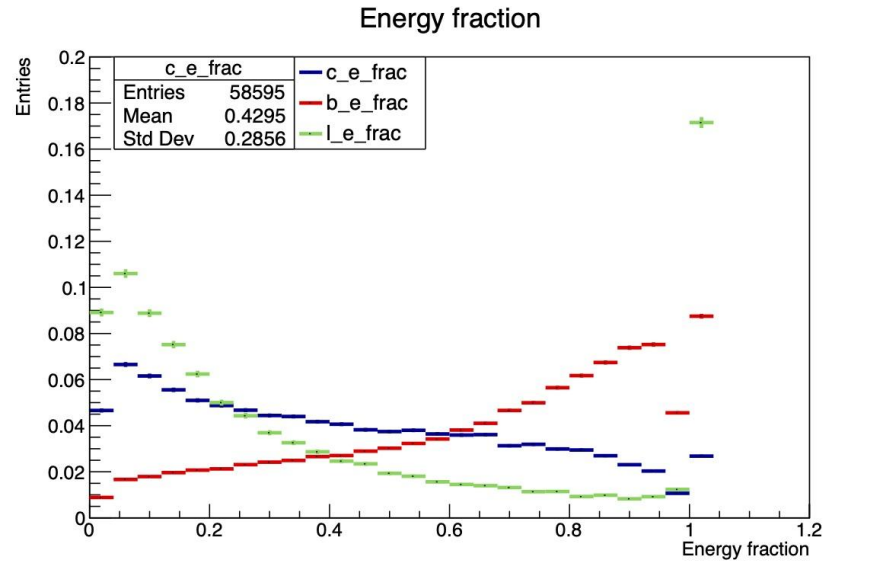




- Delta R: delta R between jet axis and direction of displaced vertices
- Lxy1: decay distance from primary to secondary vertex
- Lxy2: decay distance from secondary to tertiary vertex

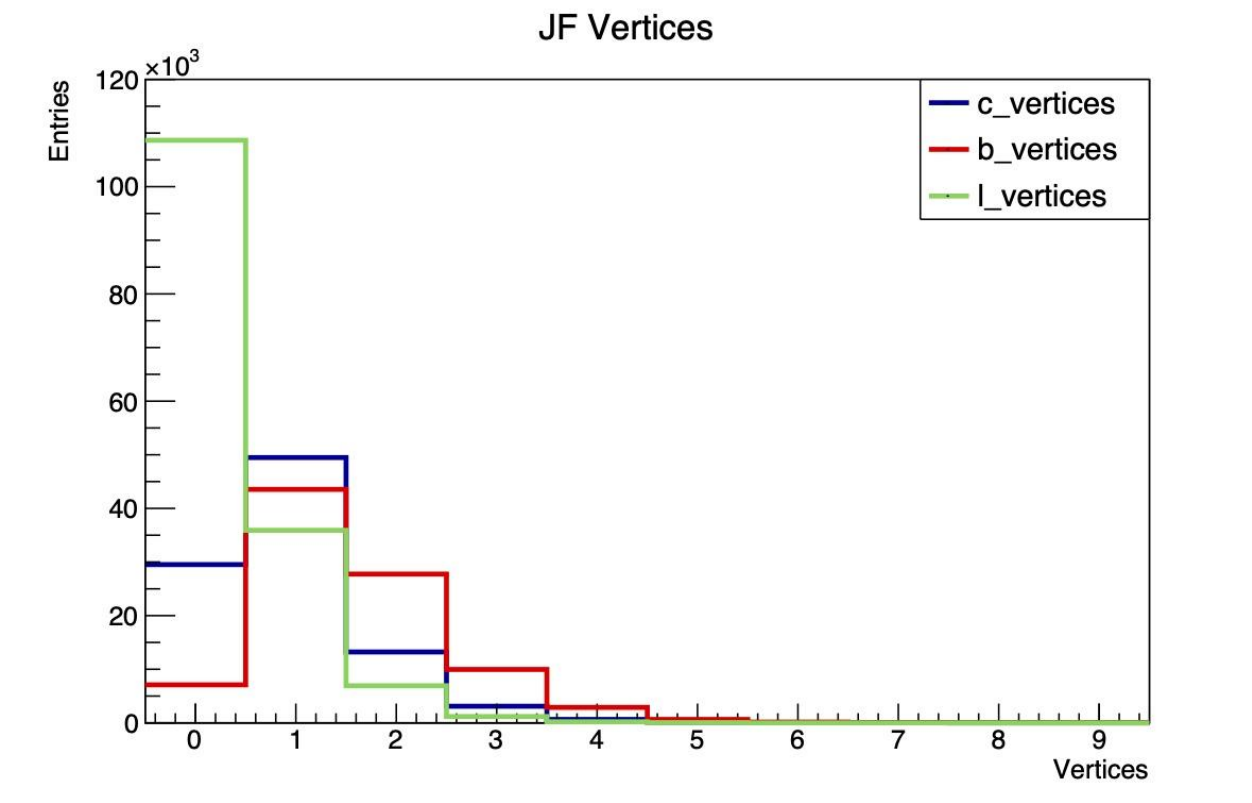
JF variables





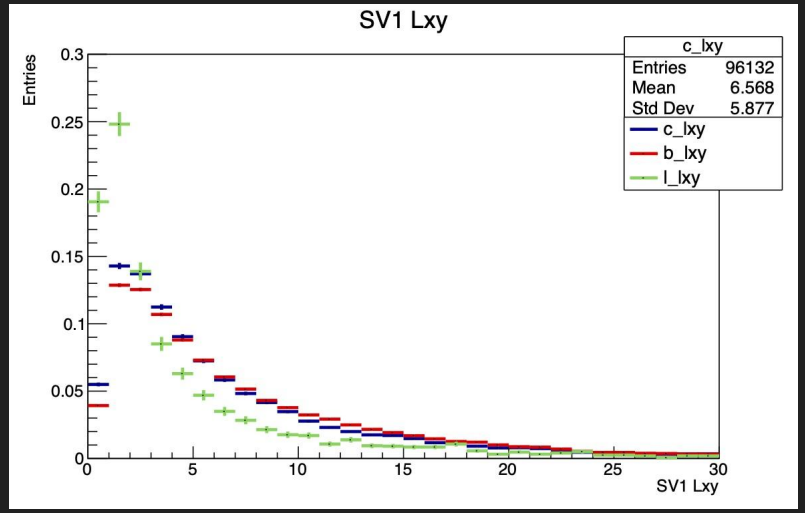
- Significance: average distance between primary vertex and displaced vertices divided by its uncertainty
- Energy fraction: ratio of energy from JF tracks to sum of all track energies

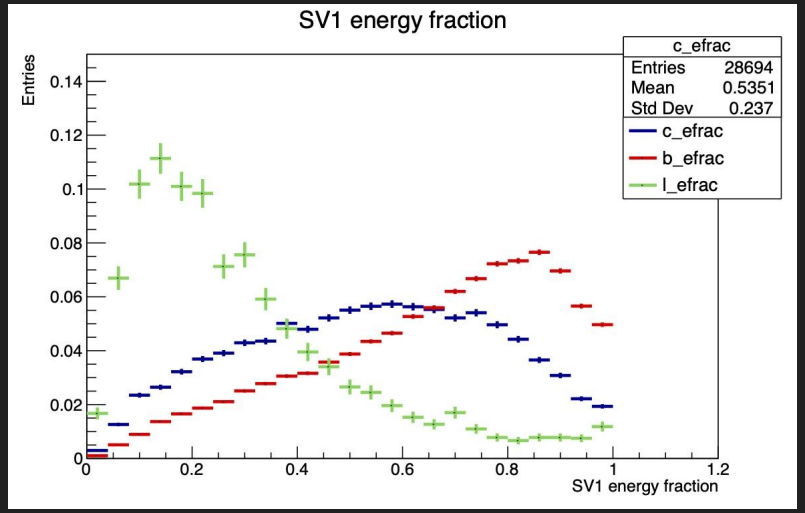
JF Variables



- b and c jets have somewhat different plots
- b-jets have much fewer 0-vertex events
- 3-vertex events are mainly b-jets

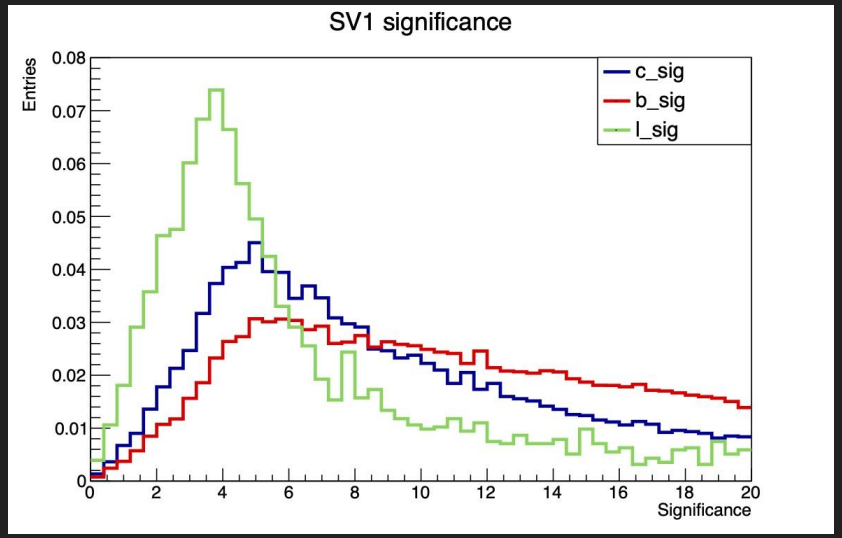
SV1 Variables



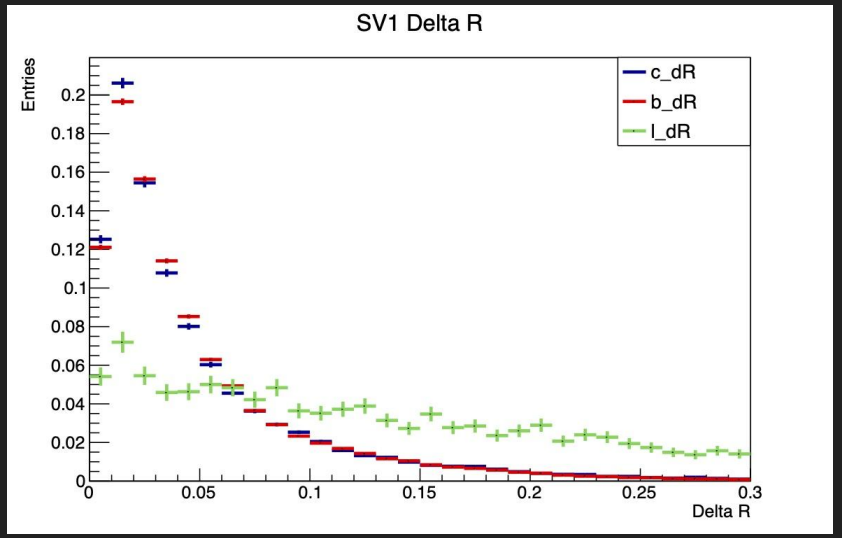


- Lxy: decay distance from primary to secondary vertex
- Energy fraction: ratio of energy from SV1 tracks to sum of all track energies

SV1 variables

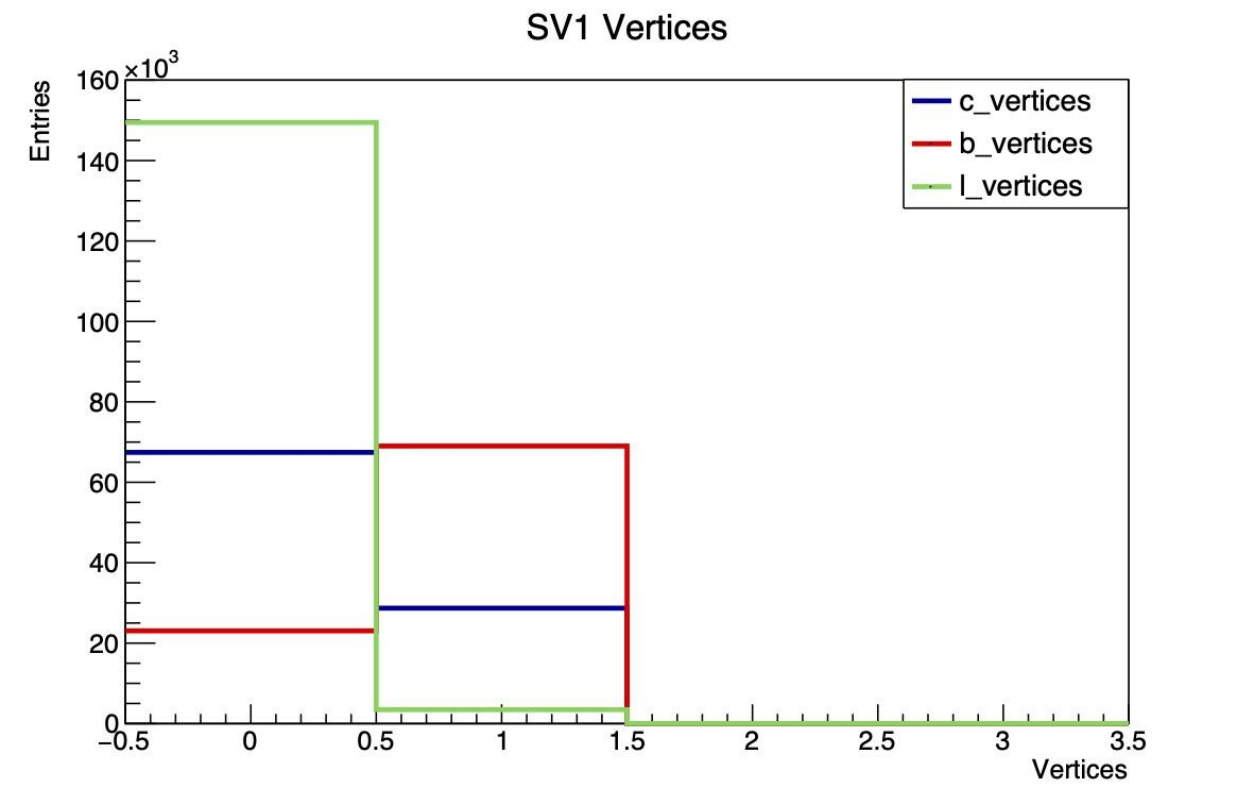


- Significance: distance from primary vertex to secondary vertex divided by its uncertainty



- Delta R: delta R between jet axis and PV->SV direction

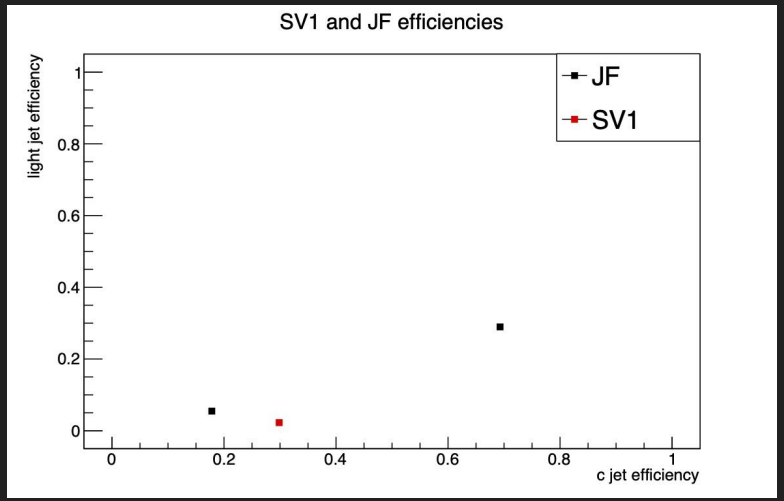
SV1 variables



- c-jets have fewer vertices than b-jets
- b-jets mainly have one SV1 vertex
-

SV1 and JF Efficiencies

- c jet SV1 efficiency: 29.8%, Light jet SV1 efficiency: 2.2%
- c jet JF efficiency: 69.29%, Light jet JF efficiency: 28.9% (one vertex requirement)
- c jet JF efficiency: 17.8%, Light jet JF efficiency: 5.47% (two vertices requirement)



Conclusion and Next Steps

- Improve c-tagging efficiency in DL1 algorithm
- b and c jets are fairly distinguishable when looking at the JF and SV1 vertices
- Machine learning: after studying the SV1 and JF variables, it will be useful to improve the DL1 algorithm
  - DL1 training
  - Improve c-tagging efficiency
  - looking into applying electron information into DL1