The 11th Annual Siam Physics Congress

A new charm quark tagging algorithm at the CMS detector.

Kittikul Kovitanggoon[1], Burin Asavapibhop[1], Narumon Suwonjandee[1], Mauro Verzetti[2], Seth Moortgat[3]

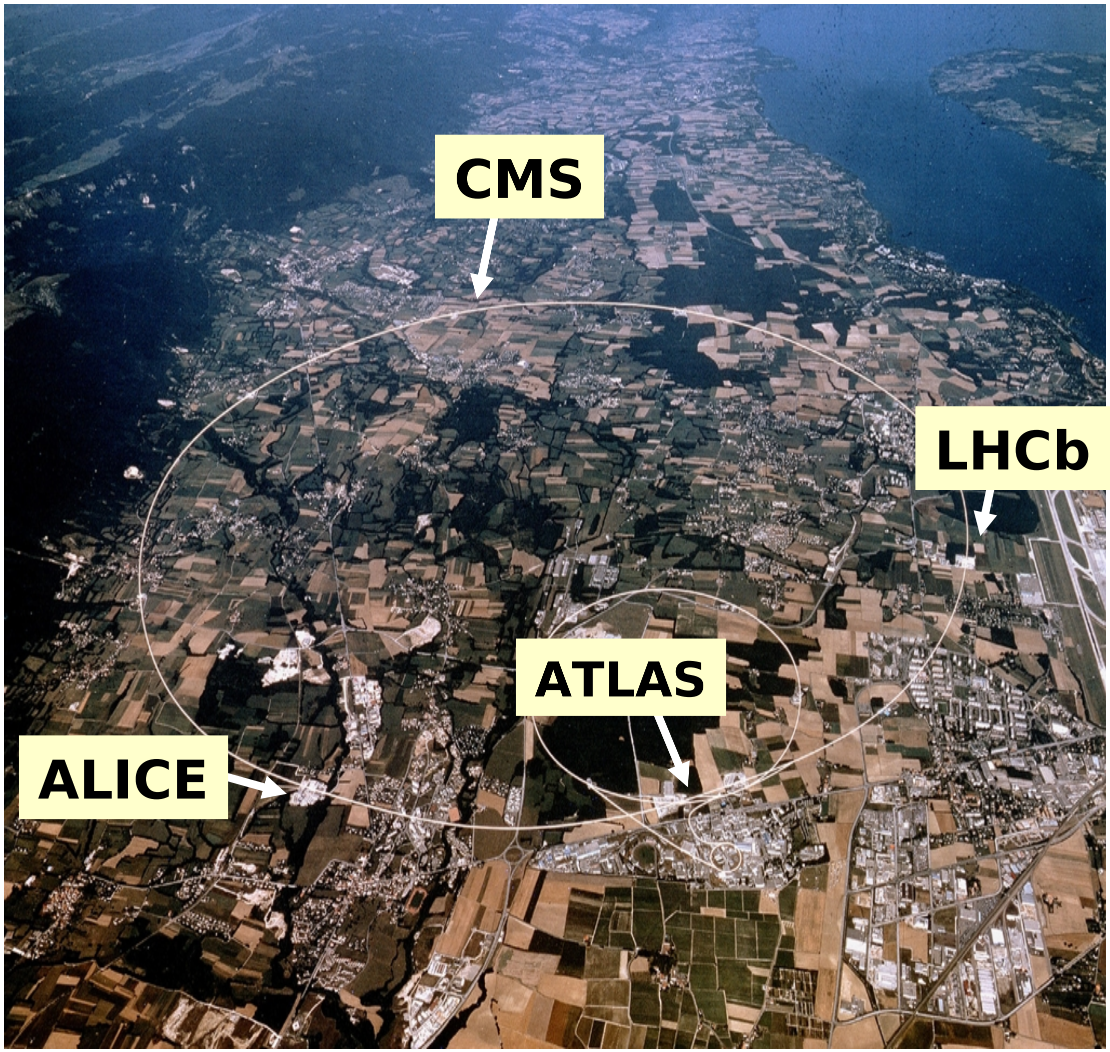
[1]Chulalongkorn University, Thailand
[2]University of Rochester, USA
[3]Vrije Universiteit Brussel, Belgium

June 8, 2016

Outline

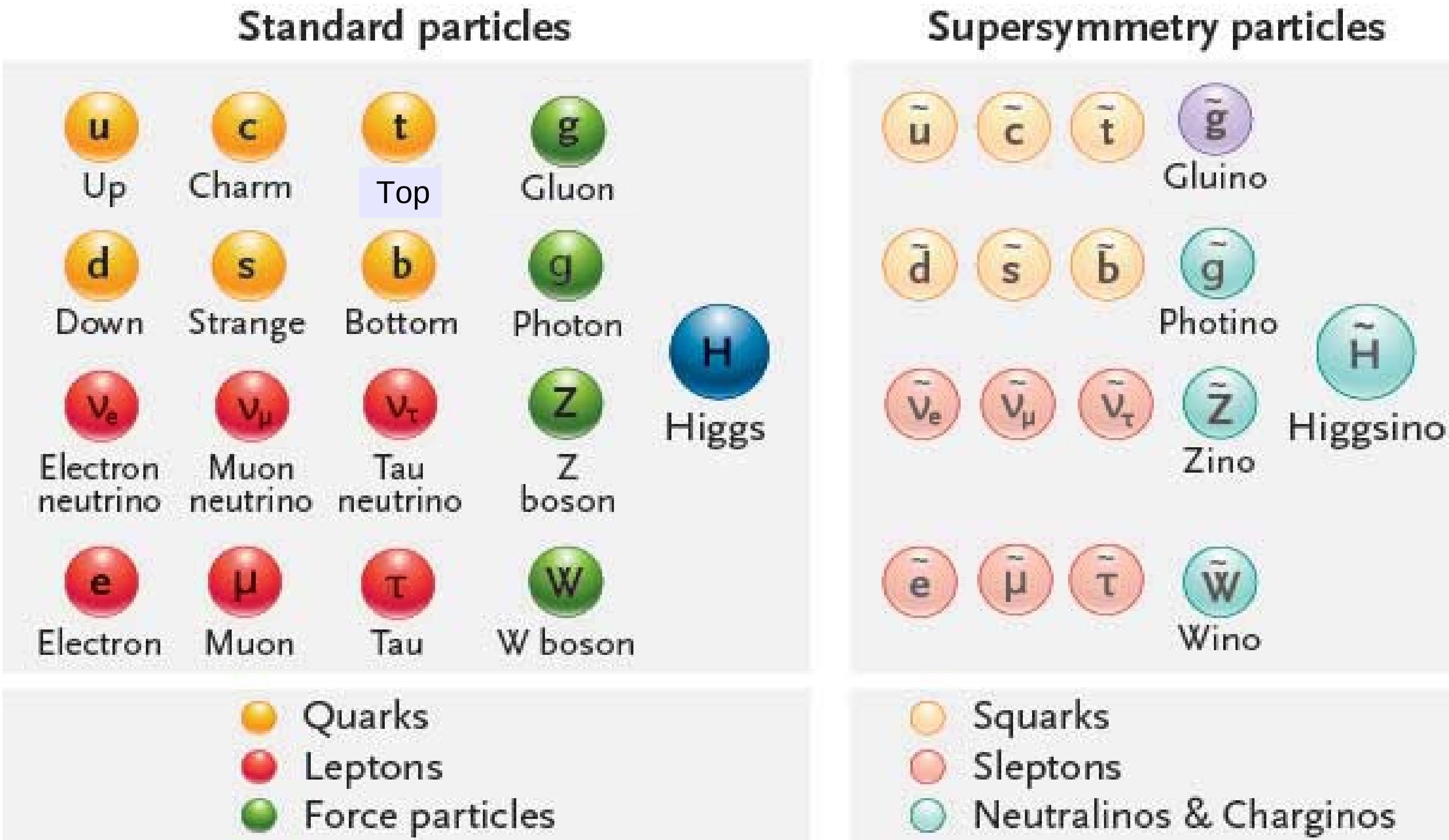
- Large Hadron Collider (LHC)
- Supersymmetry (SUSY)
- Objects and Event Reconstructions
- Compact Muon Solenoid (CMS)
- Charm quarks tagger
- Outlooks and Conclusions

Large Hadron Collider (LHC)



- 27 km in circumference
- To collide rotating beams of protons or heavy ions
- Maximum energy of proton-proton collisions at $\sqrt{s}$ = 14 TeV and 4 x 10^{34} $cm^{-2}s^{-1}$
- In 2011, collision at $\sqrt{s}$ = 7 TeV and 4 x 10^{33} $cm^{-2}s^{-1}$
- In 2012, collision at $\sqrt{s}$ = 8 TeV and 7.7 x 10^{33} $cm^{-2}s^{-1}$
- In 2015-2016, collision at $\sqrt{s}$ =13 TeV and expected 1.7x 10^{34} $cm^{-2}s^{-1}$

SUSY



- One interesting search channel is $\tilde{t} \rightarrow c + \tilde{\chi}_1^0$
- Charm quark will be detected as a jet in the detector and LSP will be invisible to our detector.
- In order to find the supersymmetric partner of the top quark in this channel, it is important to identify the charm quark jet and precisely measure the missing energy from detector.

Missing Transverse Energy

- Since the actual collisions take place at the parton level (quark, anti-quark, or gluon), thus we do not know the actual colliding energy of the partons.

- However, the plane that perpendicular to the incoming particles, **transverse plane**, has zero momentum before the collision, so after the collision the sum of outgoing momenta have to be zero.

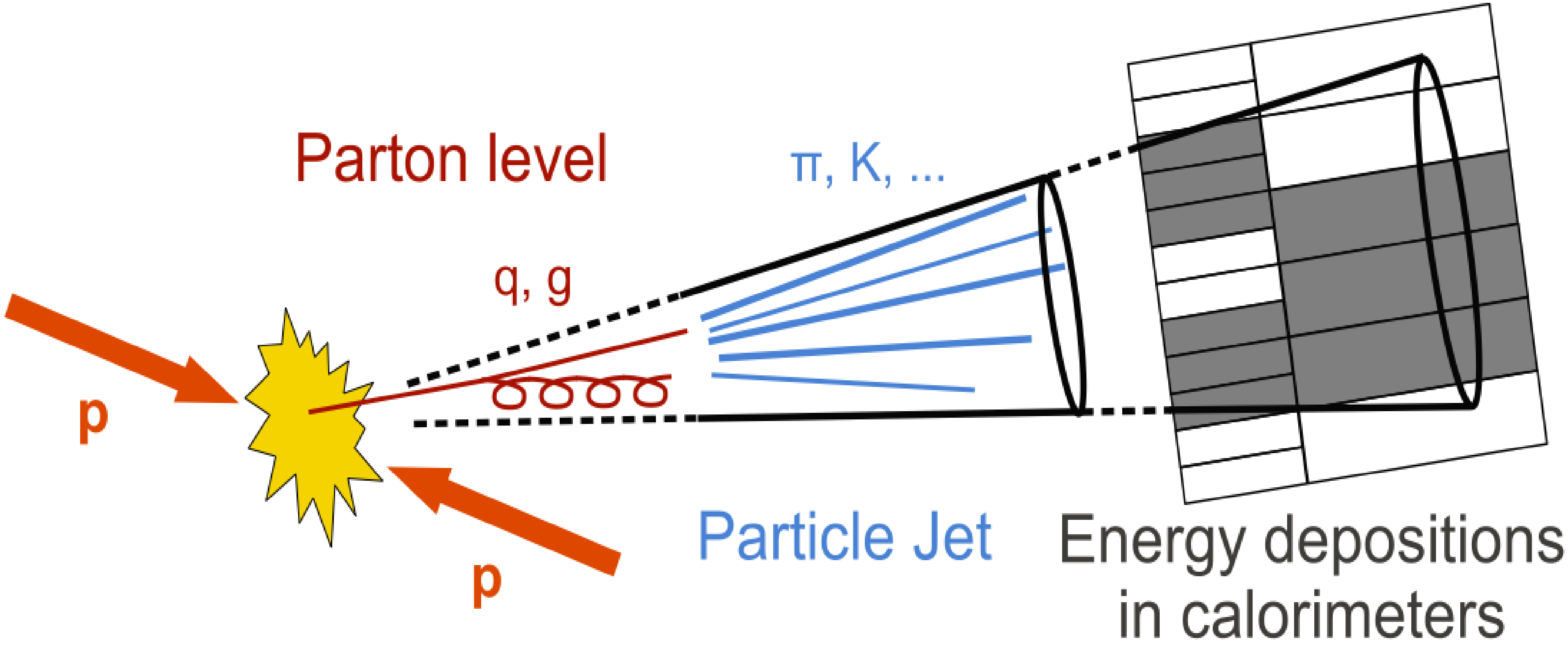
- We defined momentum measured of outgoing particles as “p_T“ or the “transverse momentum” which is the momentum in transverse plane.

- If the sum of momenta after collision is not zero, any net momentum in the transverse plane is called “**Missing Transverse Momentum** (also known as missing transverse energy, **MET**)”.

- MET is the indication of the non-detectable particles such as standard model neutrino and physics beyond standard model particle, **lightest supersymmetric particle** (**LSP**)

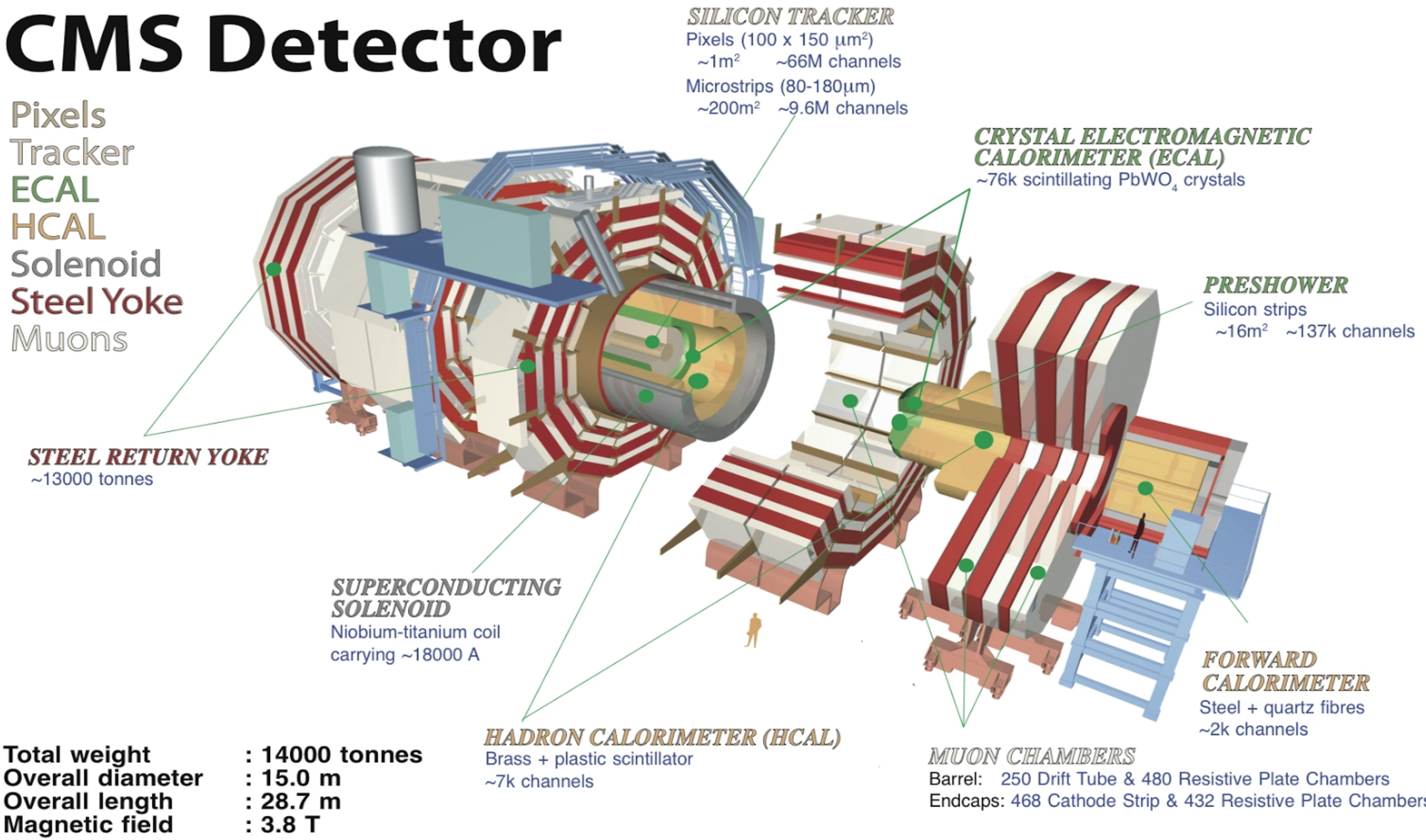
Jet

- Because of QCD confinement, particles carrying a color charge, such as quarks, cannot exist in free form.
- A high-energy quark (or gluon or anti-quark) hadronizes into a spray of hadrons (particles made from quarks, antiquarks and gluons).
- This spray is called a “jet” which can be detected.



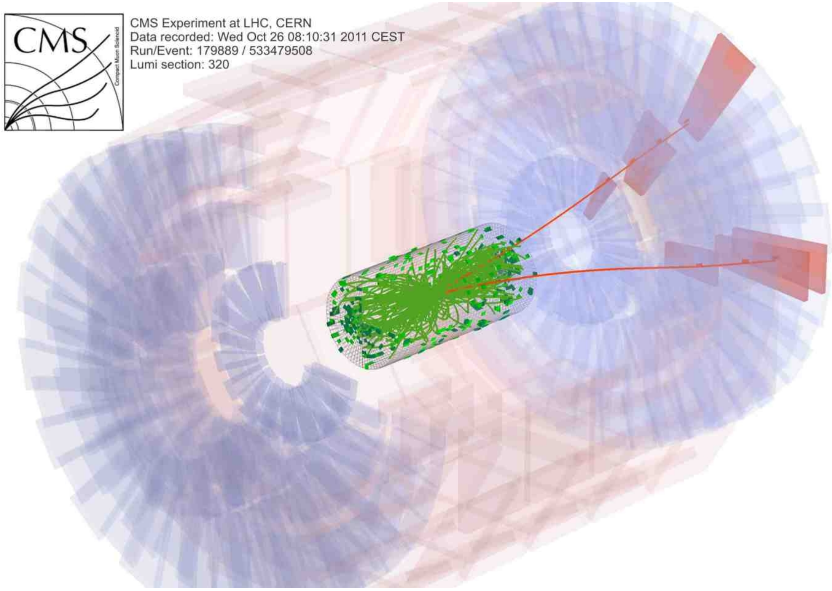
- One importmant key point of jets are the particle trajectories, we need good tracking detector.

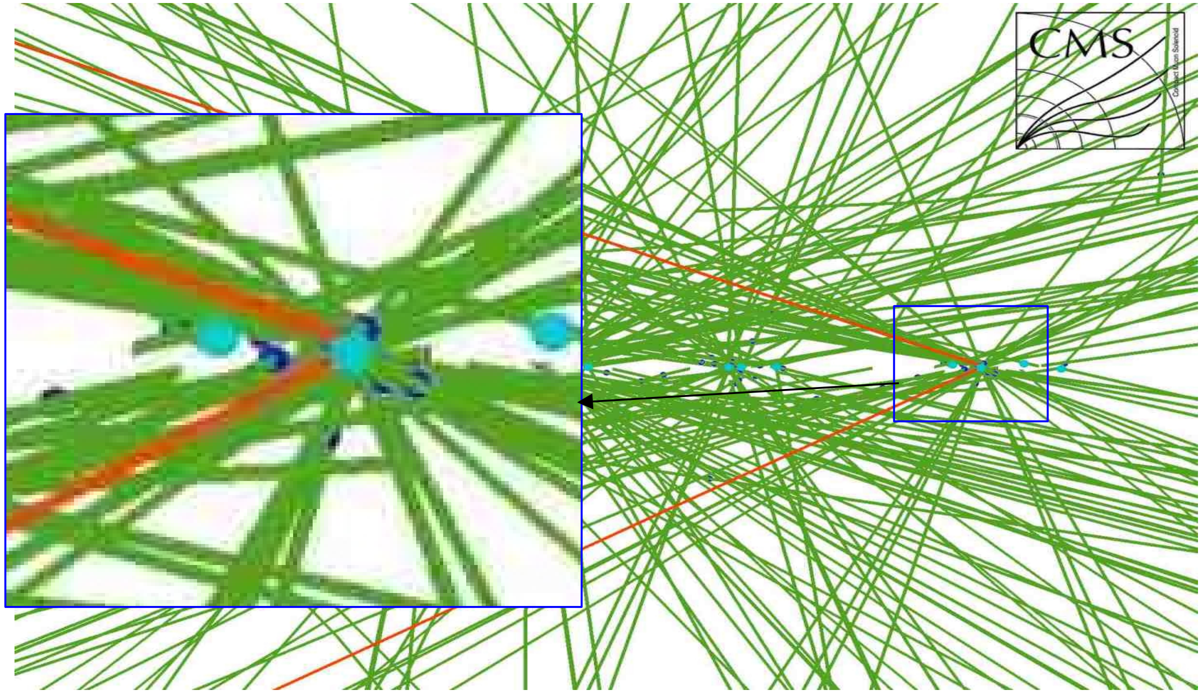
Compact Muon Solenoid (CMS)



Tracker

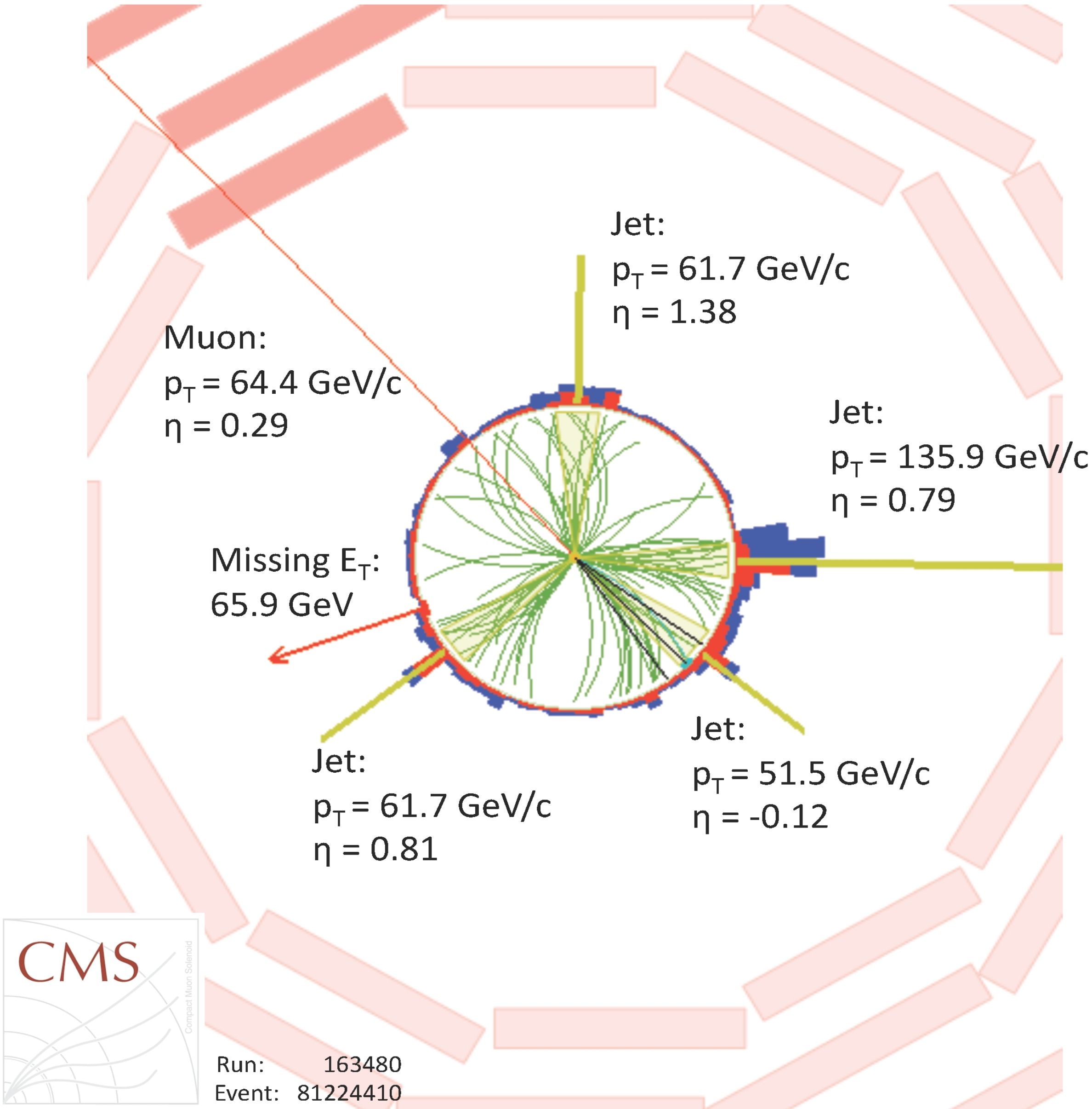
- The tracker can reconstruct the paths of high-energy muons, electrons and hadrons (particles made up of quarks) as well as see tracks coming from the decay of relatively long lived particles such as “b quarks” and “c quarks”.





- These paths can be traced back to the interaction point (or primary vertex) of the event.
- The interaction point that misplaces from beam line (or secondary vertex) is the indication of b and c quarks

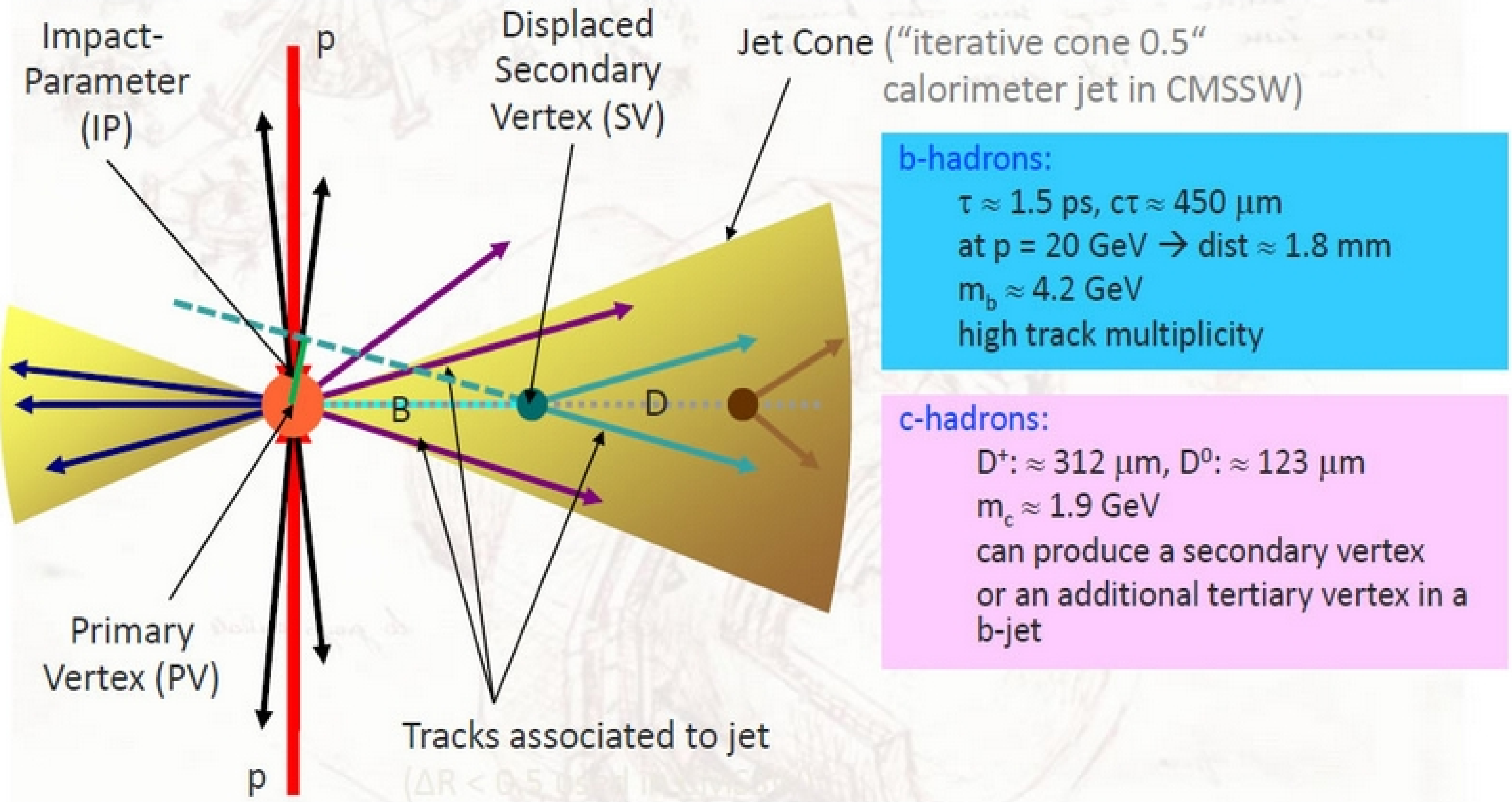
Event at CMS



Jet:
p_T = 61.7 GeV/c
η = 1.38
Muon:
p_T = 64.4 GeV/c
η = 0.29
Jet:
p_T = 135.9 GeV/c
η = 0.79
Missing E_T:
65.9 GeV
Jet:
p_T = 61.7 GeV/c
η = 0.81
Jet:
p_T = 51.5 GeV/c
η = -0.12
CMS
Run: 163480
Event: 81224410

C-quark Jet Tagger at CMS

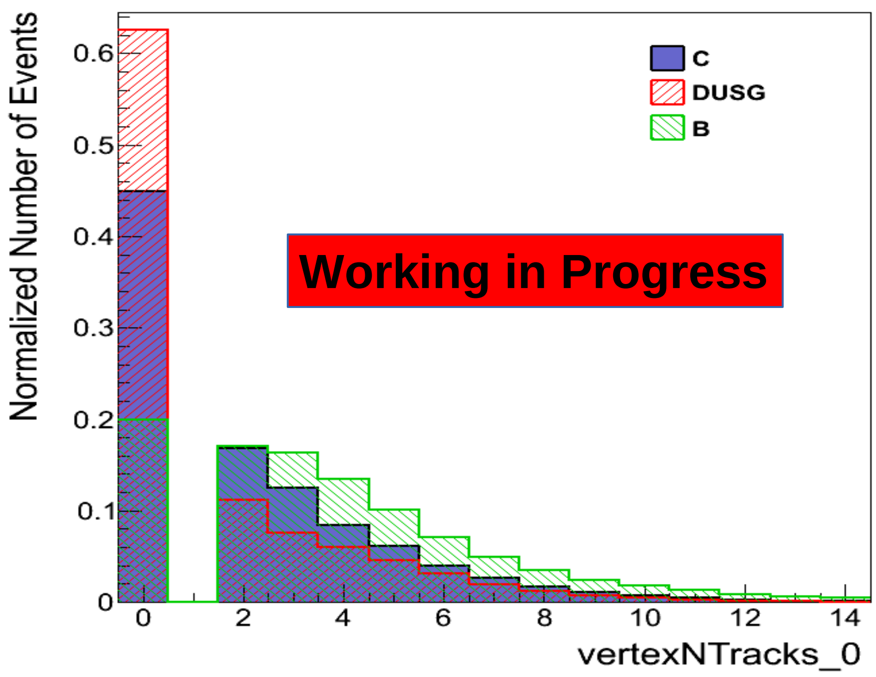
- Charm or anti-charm quark from particle collisions undergoes the hadronization process to form J/ψ meson (c and anti c) and D mesons (c or anti-c and light or anti-light quark).

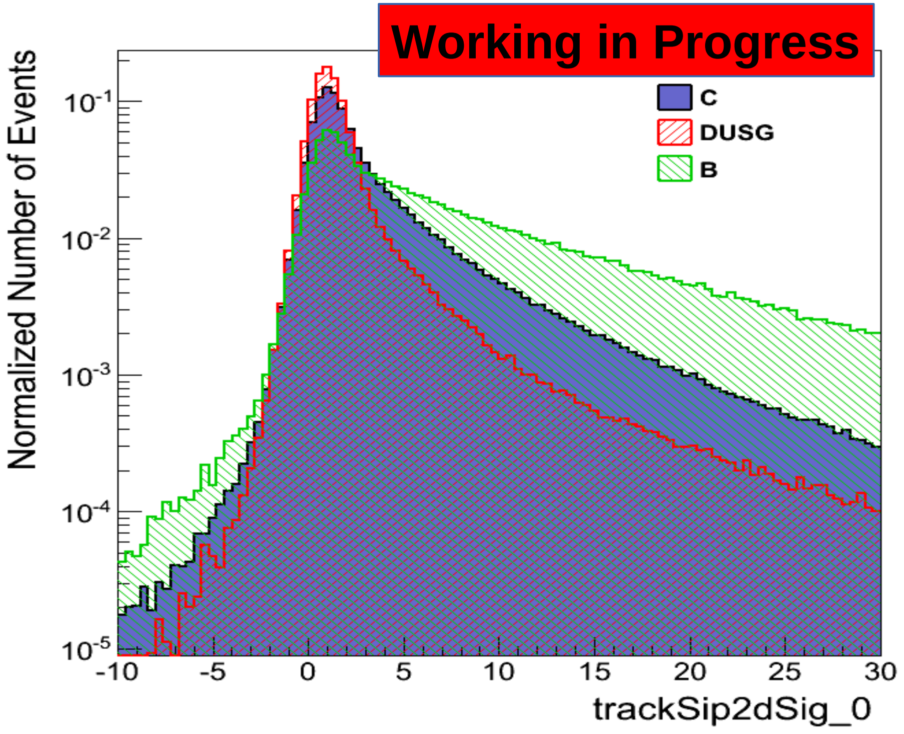


- Due to the sizable life-time, the decay of these mesons are characterized by displaced tracks with a large impact parameter (IP), a displaced secondary vertex (SV) with a large flight distance, and the soft leptons (SL) inside the jets.

C-Tagger Input Variables

- The number of tracks associated with the secondary vertex (vertexNtracks) and the signed transverse impact parameter significance (trackSip2dSig) are shown.





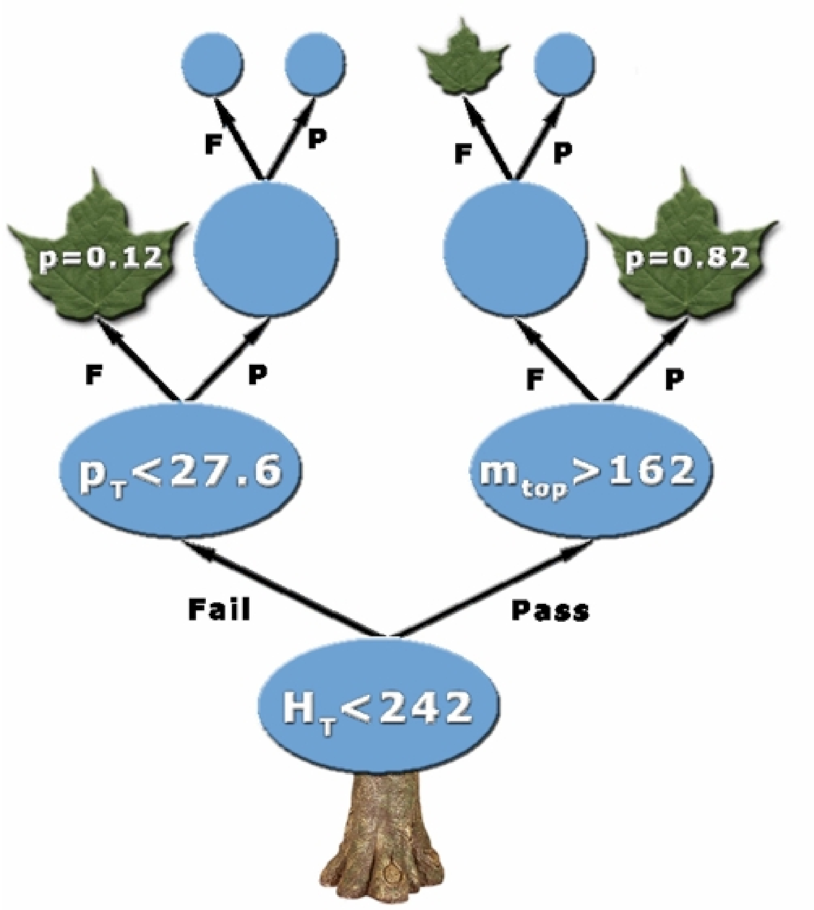
- The clear distinctions between signal (c jets) and background (light quark and gluon jets) can be seen.

- This signatures are our tools to discriminate c jets from other jets.

- 37 variables of track IP, SL, and SV are used as inputs to BDT for determine the discriminator.

Boosted Decision Tree (BDT)

• A decision tree uses successive decision nodes to categorise events as either being signal or background.



• In this analysis, decision nodes are seperated into CvsL and CvsB.

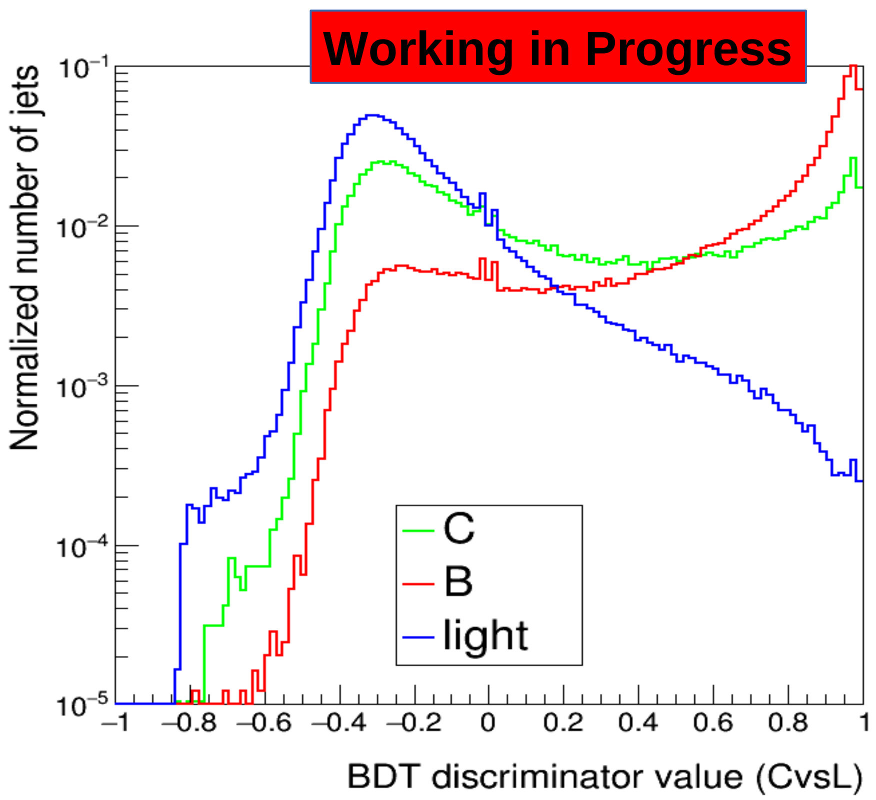
• This cutting procedure is repeated for each of the sub-nodes and stops only when one of the nodes reaches a minimum number of events or a certain signal purity.

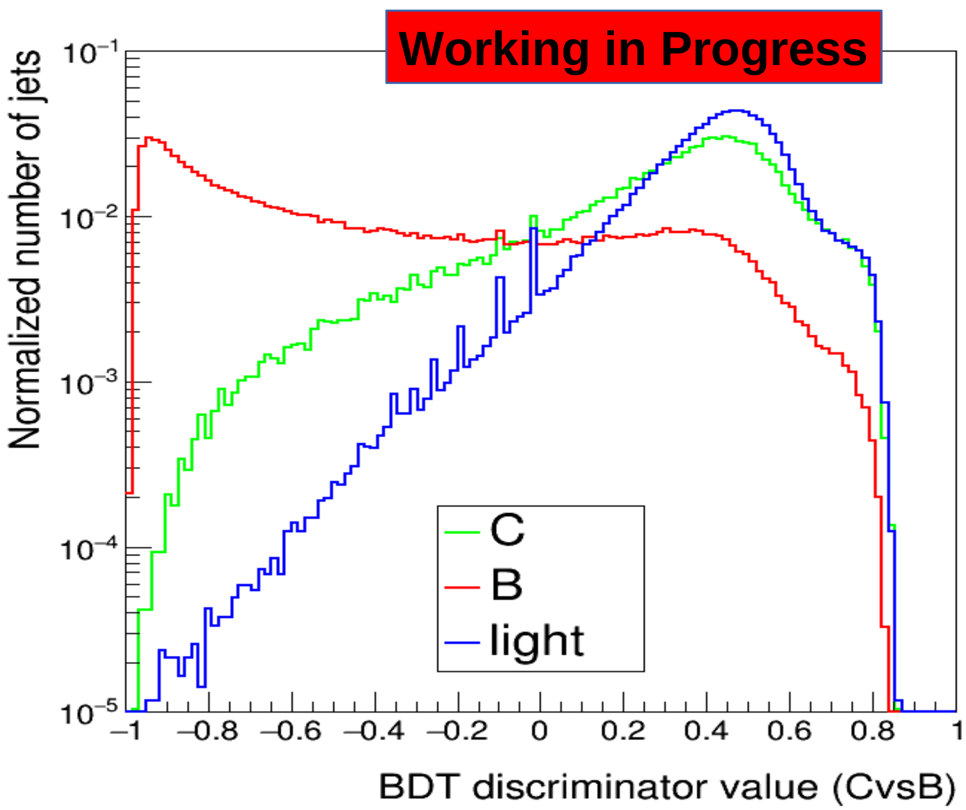
• The combination of these individual decision trees (i.e. the boosting) is done by penalizing after each individual decision tree misclassified events in the final leaves, giving them more weight in the next decision tree that is constructed.

• After the BDT is trained, events can be successively subjected to the different decision trees and are assigned an estimator (often called discriminator) value based on the number of times they end up as signal or background.

C-Tagger Discriminator

- Since Charm quark jet properties are in between the light and bottom jets.





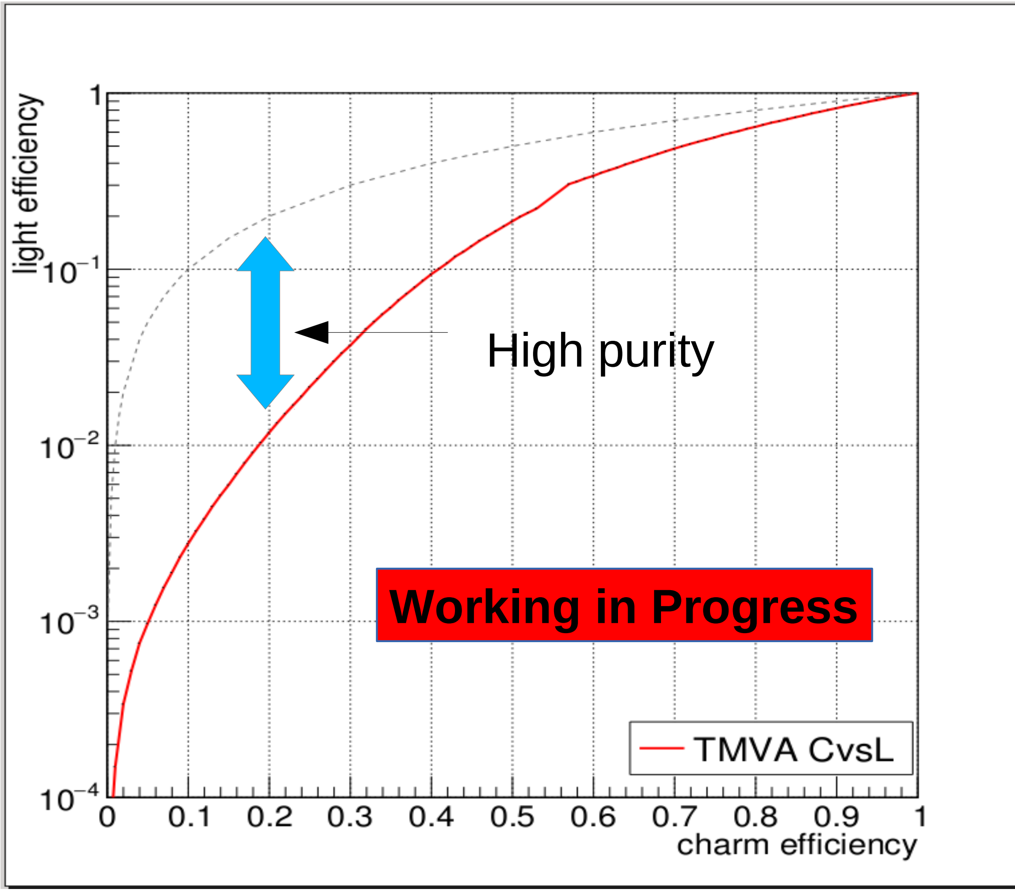
Signal (charm):	Discriminator → +1
Background (light or bottom):	Discriminator → -1

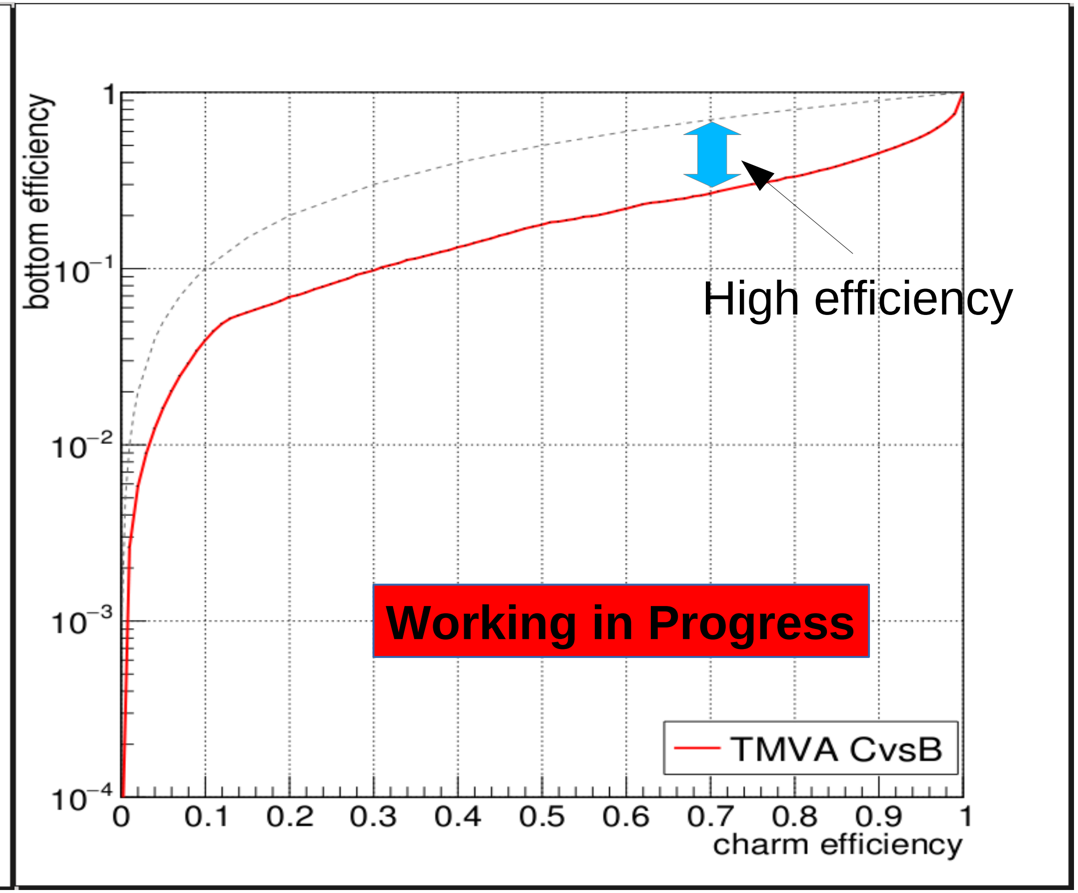
- We need to make selections on both CvsL and CvsB.

C-Tagger Performance

- The effiency, $\varepsilon_{q,}$ to tag a jet of a certain flavour q as a c-jet is defined as

$$\epsilon_q = \frac{\text{Number of jets of flavour q tagged as b}}{\text{Number of jets of flavour q}}.$$

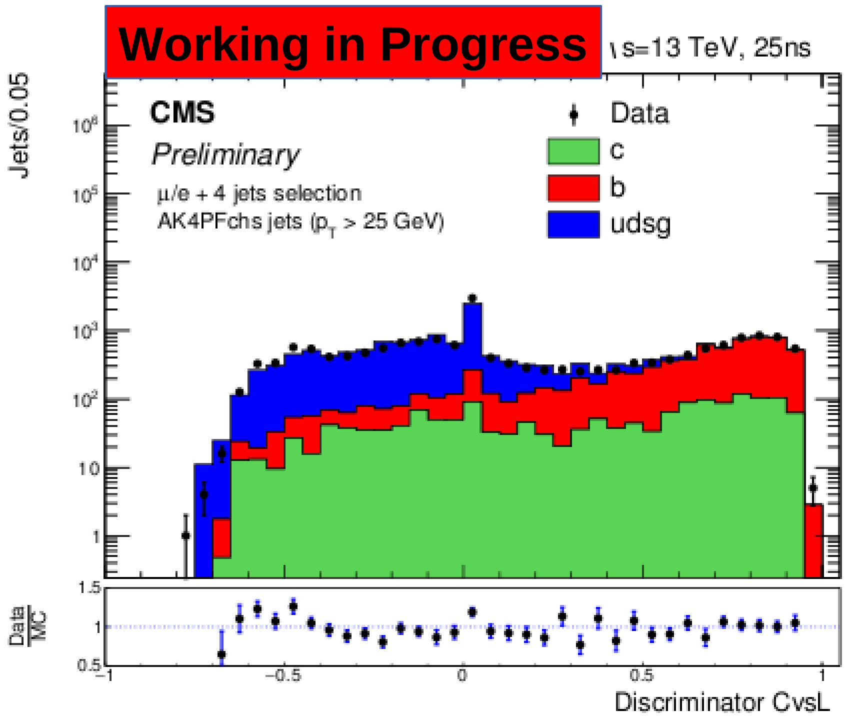


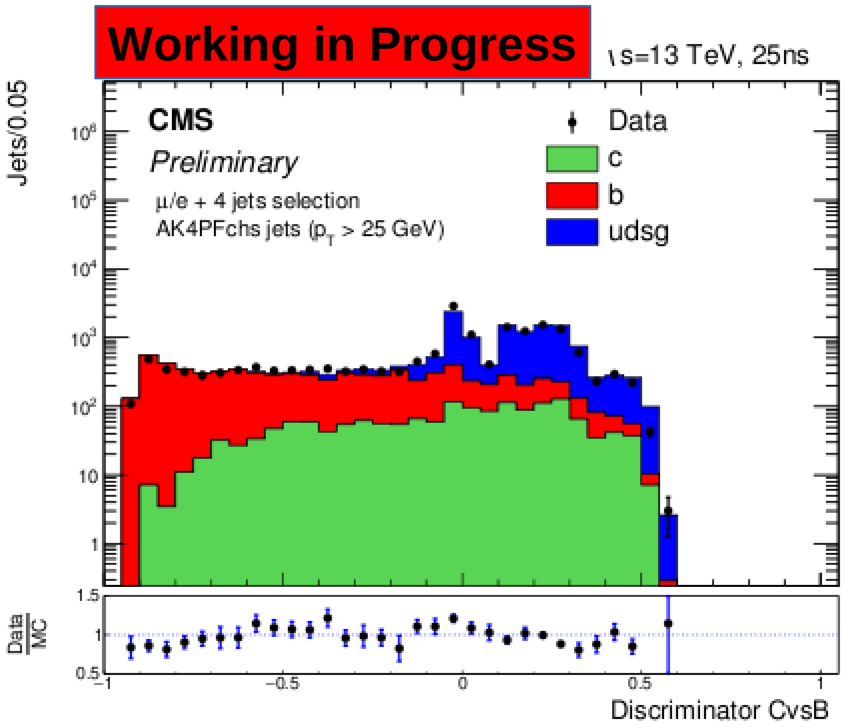


- The CvsL can effectively reject light jet background in the high purity region, whereas the CvsB can effectively reject bottom jet background in the high efficiency region.

C-Tagger Discriminator Comparison

- The comparisons between 2015 data and MC simulation are made.





- We have the good agreements on both CvsL and CvsB

Outlooks and Conclusions

Outlook:

- We are measuring the c jet efficiency in data using ttbar and W+c events.

- We are measuring the light jet efficiency in data using multijet events (negative tag method)[1].

- We are working on the public documentation: BTV-15-002.

Conclusion:

- For the first time c-tagger is avaliable at the CMS.

- C-tagger will be one of the crucial elements for new physics search such as supersymmetric top.

- Also, it can be benefit for Standard Model precision measurement.

---

1 BTV-15-001

Acknowledgments

- This research is supported by Rachadapisek Sompote Fund for Postdoctoral Fellowship, Chulalongkorn University.
- Department of Physics, Faculty of Science, Chulalongkorn University for financial support.
- CUniverse research promotion project by Chulalongkorn University.
- CMS collaboration, in particular, our b-tagger and c-tagger colleagues.
- All SPC 2016 staffs for organizing this event.

Back UP

C-quark Jets Tagger at CMS

- Identification ("tagging") of jets originating from the hadronization of c quarks (c jets).

- Due to the sizable life-time, the decay of the J/ψ meson (c and anti c) and D mesons (c or anti-c and light or anti-light quark) are characterized by displaced tracks with a large impact parameter (IP), a displaced secondary vertex (SV) with a large flight distance, and the soft leptons (SL) inside the jets.

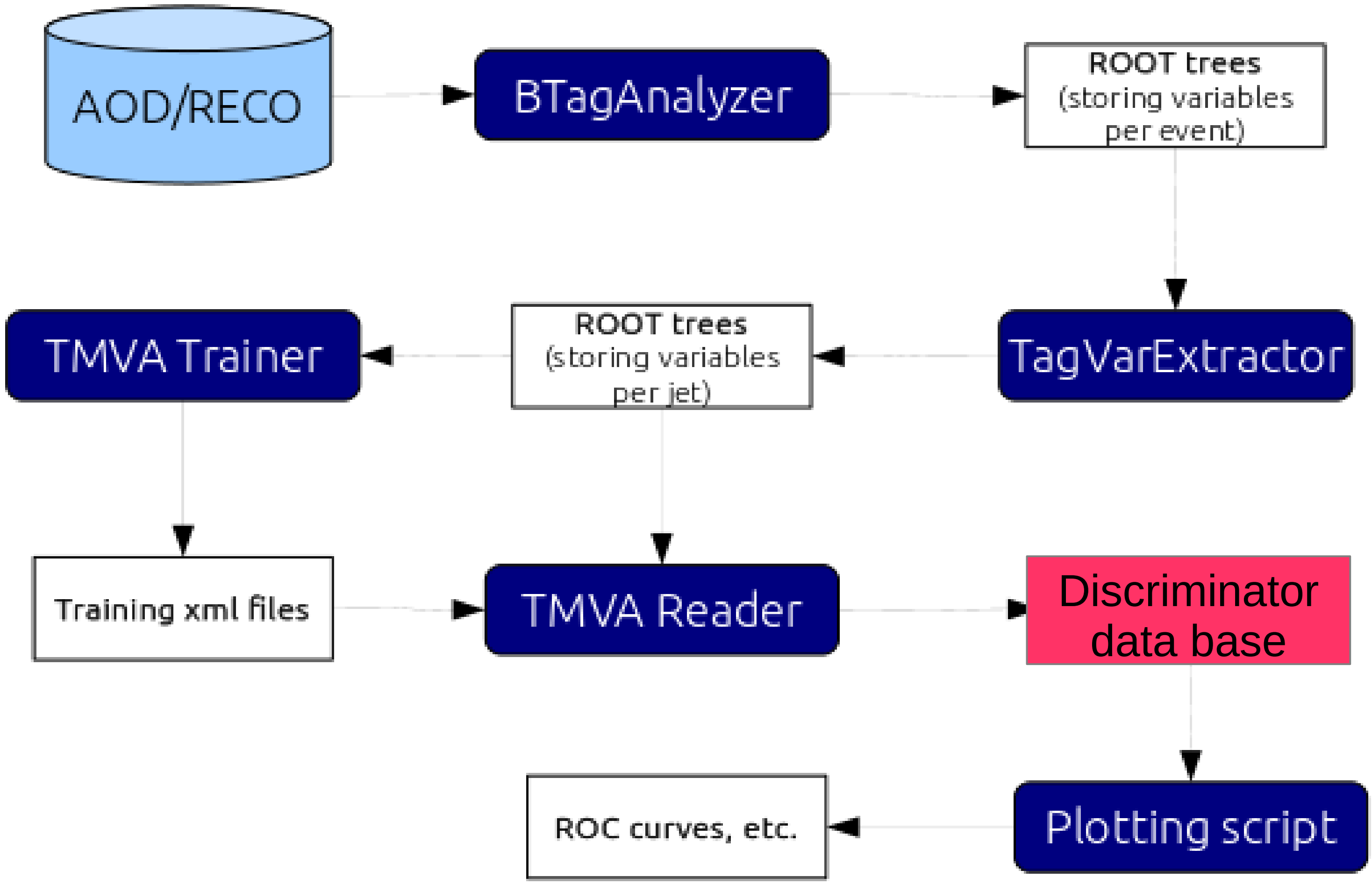
- C-tagger is a new topic studied at CMS. Started in 2013.

- C-tagger is a sub group under the well established b-tagger group.

- The method and software is based on the b-tag algorithm. The first c-tag algorithm is developed for the CMS data in 2015 and will be updated for 2016.

- As of now, there are 3 people from 3 institutes.

C-Tagger Procedures



- We are using the modified BTagAnalyzer with Toolkit for Multivariate Data Analysis (TMVA).
- Starting with well defined QCD or ttbar simulation, modified BTagAnalyzer produces a set of variables needed for c-tag purpose such as jet $p_{T,}$ jet η, track IP, SL and SV information.
- TMVA is the statistical tool providing the discriminator values to identify the possibility that such jet is more or less likely to be charm quark jet.

List of Input Variables

Table 3: Variables used (except for the p_T and η of the jet) in the development of the c-tagger with their definition.

Working in Progress

Variable	Definition
Track related variables	
jetPt	the transverse momentum of the jet
jetEta	the pseudorapidity of the jet
trackSip 3d(2d)Sig(Val)	the signed (transverse) impact parameter significance (value) of the track
trackSip3d(2d) Sig(Val) AboveCharm	the signed (transverse) impact parameter significance (value) of the track that raises the mass obtained from the summed four-momenta of the current track with the previous tracks (sorted in Sip2dSig) above the mass of the charm quark ($\simeq$ 1.5 GeV)
trackPtRel	the track transverse momentum, relative to the jet axis
trackPPar	the track momentum parallel to the jet direction, i.e. the scalar product of the jet direction and the track momentum, which basically projects the track momentum on the jet direction
trackEtaRel	$\eta_{rel} = 0.5\ln\left(\frac{E+trackPPar}{E-trackPPar}\right)$ with $E = \sqrt{\lvert\vec{p}_{track}\rvert^2 + m_\pi^2}$ and trackPPar defined just above. This is in fact the track pseudorapidity, relative to the jet axis
trackDeltaR	ΔR between the jet direction and the track momentum
trackPtRatio	the track transverse momentum, relative to the jet axis, normalized to the magnitude of its momentum
trackPParRatio	the track momentum parallel to the jet direction, normalized to the magnitude of its momentum
trackJetDist	the distance between the track and the jet axis
trackDecayLenVal	the decay length of the track calculated as the distance between the primary vertex and the point of closest approach of the track w.r.t. the jet axis
trackSumJetEtRatio	the ratio of the transverse energy of the summed four-momenta of all selected tracks and the transverse energy of the jet
trackSumJetDeltaR	ΔR between the summed four-momenta of all selected tracks and the jet direction
jetNTracks	the number of tracks associated to the jet
Secondary Vertex related variables	
vertexMass	mass of the track four momentum sum at the secondary vertex
vertexNTracks	the number of tracks associated with the secondary vertex
vertexEnergyRatio	the ratio of the energy of the summed four-momenta of all secondary vertex tracks and of all tracks associated with the jet
vertexJetDeltaR	ΔR between the summed four-momenta of all secondary-vertex tracks and the jet direction
flightDistance 3d(2d)Sig(Val)	the significance (value) of the (transverse) distance between the primary and the secondary vertex.
jetNSecondary Vertices	the number of reconstructed secondary vertices (of the type RecoVertex)
massVertex EnergyFraction	vertex mass times the fraction of the vertex energy with respect to the jet energy
vertexBoost OverSqrtJetPt	variable related to the boost of the vertex system in the flight direction
Soft Lepton related variables	
leptonSip3d(2d)	the signed (transverse) impact parameter significance of the soft lepton
leptonPtRel	transverse momentum of the soft lepton with respect to the jet axis
leptonDeltaR	ΔR between the jet direction and the soft lepton momentum
leptonRatio(Rel)	momentum of the soft lepton (parallel to jet axis) over jet energy
leptonEtaRel	pseudorapidity of the soft lepton along jet axis

Data and MC Simulations

Table 1: Simulation samples of RunIIFall15MiniAODv1/2-PU25nsData2015v1.

Working in Progress

RunIISpring15DR74-Asympt50ns_MCRUN2_74_V9A	σ[pb]
Signal	
TT_TuneCUETP8M1_13TeV-powheg-pythia8	832
TT_TuneCUETP8M1_13TeV-powheg-scaledown-pythia8	832
TT_TuneCUETP8M1_13TeV-powheg-scaleup-pythia8	832
TT_TuneCUETP8M1_mtop1695_13TeV-powheg-pythia8	832
TT_TuneCUETP8M1_mtop1755_13TeV-powheg-pythia8	832
Background	
ST_t-channel_4f_leptonDecays_13TeV-amcatnlo-pythia8	72.3
ST_tW_top_5f_NoFullyHadronicDecays_13TeV-powheg_TuneCUETP8M1	19.3
ST_tW_antitop_5f_NoFullyHadronicDecays_13TeV-powheg_TuneCUETP8M1	19.3
WJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	61524
DYJetsToLL_M-50_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	6024
QCD_Pt-30to50_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8	1652471
QCD_Pt-50to80_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8	437504
QCD_Pt-80to120_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8	106033
QCD_Pt-120to170_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8	24720
QCD_Pt-170to300_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8	8654
QCD_Pt-300to470_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8	797.3
QCD_Pt-470to600_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8	79.03
QCD_Pt-600to800_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8	25.10
QCD_Pt-800to1000_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8	4.707
QCD_Pt-1000toInf_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8	1.621
QCD_Pt-30to50_EMEnrichedPt5_TuneCUETP8M1_13TeV_pythia8	9928000
QCD_Pt-50to80_EMEnrichedPt5_TuneCUETP8M1_13TeV_pythia8	2890800
QCD_Pt-80to120_EMEnrichedPt5_TuneCUETP8M1_13TeV_pythia8	350000
QCD_Pt-120to170_EMEnrichedPt5_TuneCUETP8M1_13TeV_pythia8	62964
QCD_Pt-170to300_EMEnrichedPt5_TuneCUETP8M1_13TeV_pythia8	18810
QCD_Pt-300toInf_EMEnrichedPt5_TuneCUETP8M1_13TeV_pythia8	135