Optimization of Cut-Based EID Menus for Run2
&&
Selection Optimization for the ttH 3 Lepton Final State at 13TeV

Chao Wang

Supervisor: Cristinel Diaconu, Emmanuel Monnier, Lianliang Ma

Outline

- Qualification task
  - Optimization of cut-based electron identification menus for Run2 at low Pt
- ttH 3 Lepton Final State at 13TeV Analysis
  - ttH overview
  - Cut based study and selection optimization

Optimization of cut-based EID menus

Motivation:

- Lower energy background is higher due to low energy photo and pion
- Run2 has a higher luminosity than run 1 and pile-up situation us different
- Reconstruction and the identification menus have also to be retuned

  - Signal: J/ψ → ee and Z → ee
  - Background: Inelastic minimum bias and dijet filtered, ET(jet) > 17 GeV
  - The 9 variables are considered due to their separation power are:

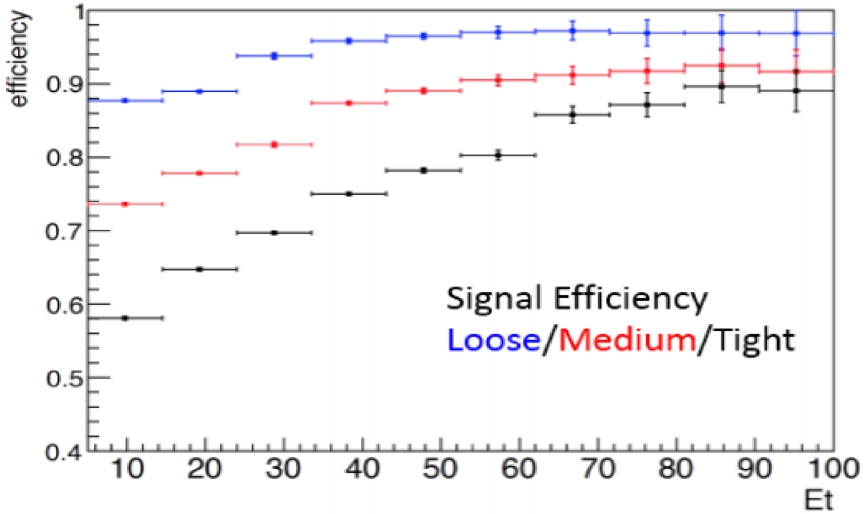
    f3, wstot, ERatio, RHad, Rη, Rϕ, Wη2, eOverP, ΔΦ

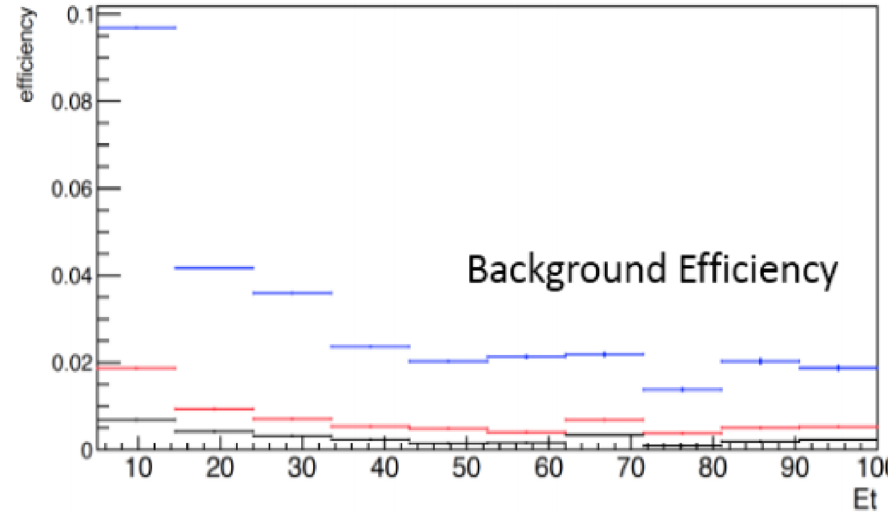
Strategy:

- Pre-selection
  – Truth match used for signal, and anti-truth match used for background – Et: [5,20] ; |eta|<2.47
- Cleaning cuts were applied to avoid cuts in tails of distributions
- Optimization carried out in groups of variables(correlation and pileup dependence)
  TMVA Cut-based method to optimize the cuts
- A target efficiency was decided for each low Et bin
  – Flatness in efficiency desired
- Post-optimization, cuts were smoothened
  – Avoid jumps in cut values from one bin to next
- Inclusiveness of menus ensured
  – Loose is looser than medium and medium is looser than tight
- Monotonicity of cut values as a function of Et
  – Cut value in the i-th eta bin always greater than or less than i-th eta bin in the next Et bin

Menus	5-10(GeV)	10-15(GeV)	15-20(GeV)
Loose	0.88	0.88	0.88
Medium	0.75	0.75	0.80
Tight	0.60	0.65	0.75

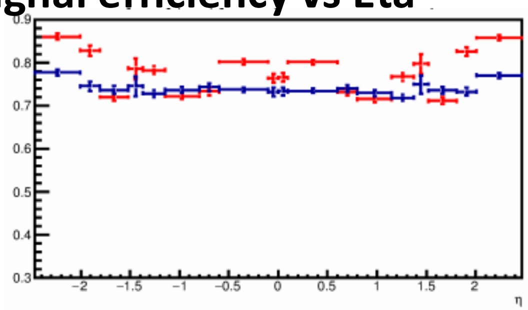
Optimization of cut-based EID menus

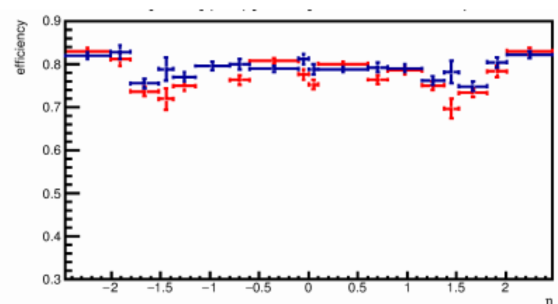


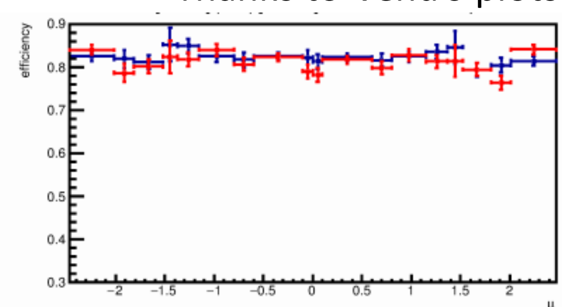


Thanks to Venu's plots

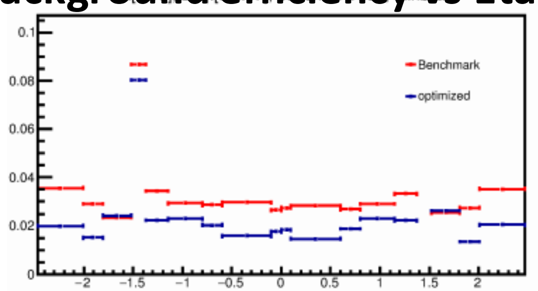
Signal efficiency vs Eta

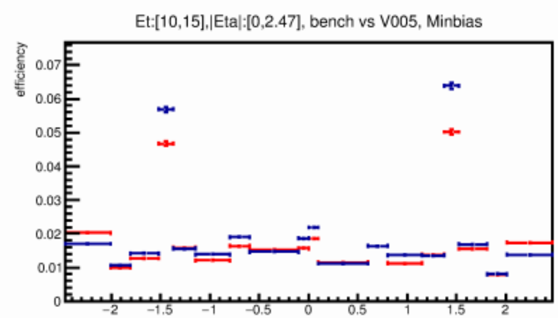


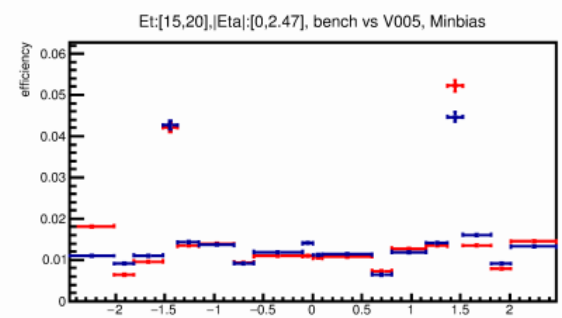




Background efficiency vs Eta

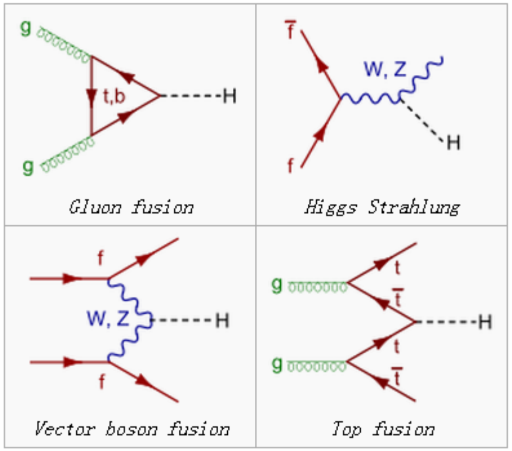






5-10 GeV **10-15 GeV** **15-20 GeV**

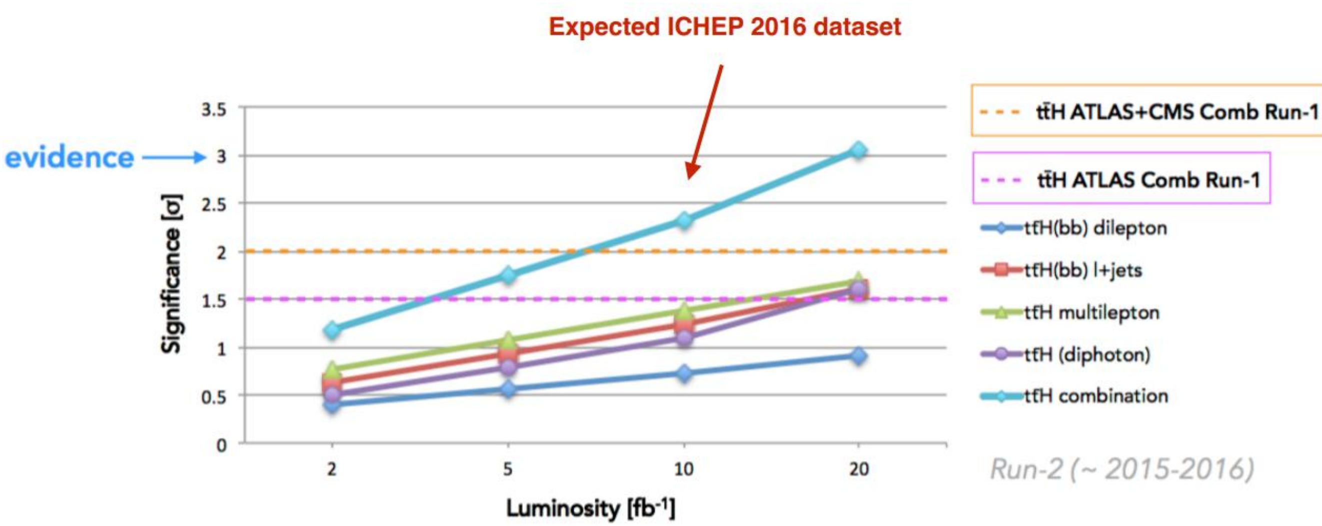
Higgs Boson (125GeV) Production at LHC



Cross section (pb) at $\sqrt{s} = 8$ (7) TeV	
ggF	19.52 (15.32)
VBF	1.58 (1.22)
WH	0.70 (0.57)
ZH	0.39 (0.31)
$t\bar{t}H$	0.13 (0.09)
Total	22.32 (17.51)

Cross Section(pb) at 13TeV	
ggF	43.92
VBF	3.748
WH	1.380
ZH	0.8696
ttH	0.5085

a factor of 4 in 13 TeV



Top and Higgs: the Special Relationship

"Higgs Boson"

"Top Quark"

Yukawa Coupling

$$y_t = \sqrt{2} m_t / v \simeq 1$$

Special role for the top quark in electroweak symmetry breaking?

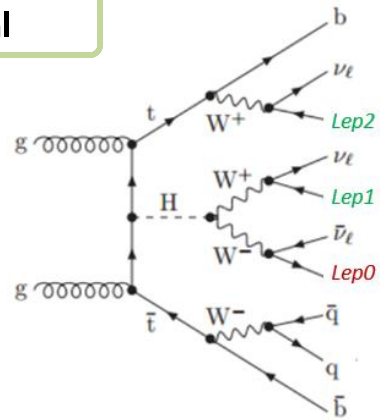
Close to unity - why?

Why is the top so heavy?

$t\bar{t}H$ production with smallest cross section.
Test properties and give direct access to top-Higgs coupling.

$t\bar{t}H$ Signature with 3 Leptons

Signal



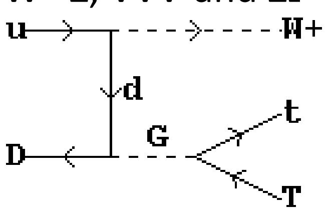
- **3 leptons, Lepton charges' sum = +/-1**
- **High jet multiplicity, b-jets**

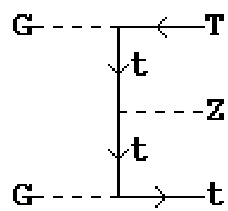
Lep0: the lepton with no same-charge partner(never fake lepton for $t\bar{t}$, Z + jets background)

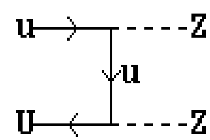
Lep1: ΔR closet to the Lep0

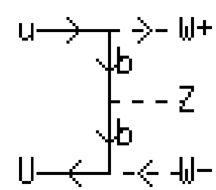
Background

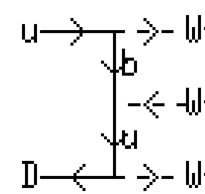
- Events same final state as the signal (irreducible backgrounds), three true charged leptons: $t\bar{t}V$, W$^{\pm}$Z, VVV and ZZ





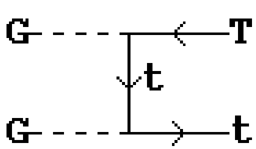


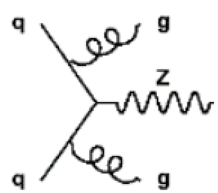


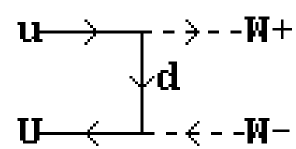


Obs: top-> bW ; W/Z can decay to leptons.

- Events with a non-prompt or a fake lepton selected as prompt lepton, events with misreconstructed charge: $t\bar{t}$, Z + jets, W$^+$W$^-$.







Signal: leptons (some fake), low jet multiplicity, in general no b-jets.

Cut-based Study and Selection Optimization

Motivation

- Toward to Run2 Condition(higher energy scale and high pile-up)
- New b-tagging and some CP tools(Lep ID and Isolation…)

Sample Usage(MC15a samples):

- Z+jets: Madgraph+Pythia;
- ttW: MadGraphPythia;
- ttbar: PowhegPythia
- VV: Sherpa

ttH(semilep + dilep): aMcAtNloHerwigpp

ttZ: MadGraphPythia

ttbar(singletop, ttww, 4Top): PowhegPythia

TTHtoLeptonsPreliminarySelection: "preliminary" selection for Moriond 2016 and Event Selection

https://twiki.cern.ch/twiki/bin/view/AtlasProtected/TTHtoLeptonsPreliminarySelection

Electrons

- pt > 10 GeV
- |eta| < 2.47, and not 1.37 < |eta| < 1.52 (use el->caloCluster()->etaBE(2))
- pass Loose Likelihood ID
- |z0 sin theta| < 2 mm
- |d0 significance| < 10
- pass isolation (Loose working point in IsolationSelectionTool)

Tau Jets (hadronically decaying tau lepton)

- abs(charge==1)
- (nTracks ==1 || nTracks ==3)
- eta : [0, 1.37], [1.52, 2.5]
- JetIDBDTMedium == 1
- pT > 25 GeV
- EleOLR in TauSelectionTool

Muons

- pt > 10 GeV
- |eta| < 2.5
- pass loose muon quality requirement: MuQuality ≤ 2 (or muon_isLoose = 1 with group ntuples)
- |z0 sin theta| < 2 mm
- |d0 significance| < 10
- pass isolation (Loose working point in IsolationSelectionTool)

Jets

- pass jet clean criteria ("LooseBad" in the JetCleaningTool)
- pt > 25 GeV
- |eta| < 2.5
- remove jets with |JVT| < 0.64 and |eta| < 2.4 and pt < 50 GeV (use jet->jetP4("JetEMScaleMomentum").eta())
- BTag: BTagMV2c20 > -0.4434 (77% eff)

Cut-based Study and Selection Optimization

- pass lepton triggers
  - in MC : HLT_mu20_iloose_L1MU15 || HLT_mu50 || HLT_e24_lhmedium_L1EM18VH || HLT_e60_lhmedium || HLT_e120_lhloose
  - in Data : HLT_mu20_iloose_L1MU15 || HLT_mu50 || HLT_e24_lhmedium_L1EM20VH || HLT_e60_lhmedium || HLT_e120_lhloose
- select leptons (electron/muon/tau) and jets following above Object Selection
- Overlap removal following run-I recommendation, which is
  - if an electron and muon candidate are within 0.1 of each other: remove the electron
  - If two electron candidates within 0.1 of each other: remove the one with lower pt
  - if an electron and a jet are within 0.3 of each other: remove the jet
  - if a muon and a jet are within 0.04+10[GeV]/pT(muon) of each other: remove the muon
  - if an electron and a tau are within 0.2 of each other: remove the tau
  - if an muon and a tau are within 0.2 of each other: remove the tau
  - if a tau and a jet are within 0.3 of each other: remove the jet

Selection in **three leptons** channel

- basic selection (see above)
- requiring exactly three light leptons : sum of lepton charge = +/- 1
- two same sign lepton pt > 20 GeV and the OS lepton pt > 10 GeV
- any of the leptons matched to any of the triggers with pt >25 GeV (see above trigger names)
- Z mass veto : [81, 101] GeV
- Invariant mass of the two opposite sign lepton great than 12 GeV

$$Z_0 \text{ significance estimator} = \sqrt{2((S+B)\ln(1+S/B)-S)}$$

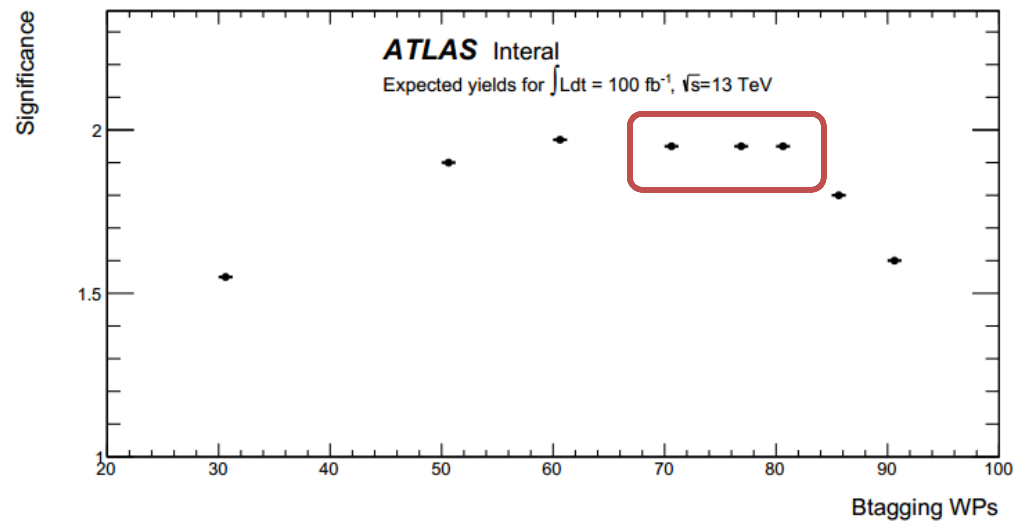
Selection Optimization -- b tagging and jet multiplicity

Task1: b tagging working point and Jets multiplicity

Goal: test and find a good b tagging working point and Jet requirement

MV2c20 tagger (AntiKt4EMTopoJets jets)

name	weight cut	b-jet efficiency [%]	purity [%]
FixedCutBEff_30	0.9540	30.17	99.87
FixedCutBEff_50	0.7535	50.26	99.11
FixedCutBEff_60	**0.4496**	**60.27**	**97.83**
FixedCutBEff_70	**-0.0436**	**70.26**	**94.86**
FixedCutBEff_77	**-0.4434**	**77.25**	**90.90**
FixedCutBEff_80	-0.5911	80.24	88.45
FixedCutBEff_85	**-0.7887**	**85.22**	**82.75**
FixedCutBEff_90	-0.9185	90.17	72.89



B Tagging Working Points	30	50	60	70	77	80	85	90
$t\bar{t}H$	52.7	80.5	92.0	102.3	107.8	110.3	114.4	118.8
Total Background	1698.0	2847.2	3412.7	4024	4735.5	5059.8	5753	7244.3

All events are normalized to 100fb^-1

Selection Optimization -- b tagging and jet multiplicity

- Best jet combinations under each b-tagged efficiency.

Here 4-0 or 3-2 means: (NJet >= 4 and NbJet >= 0) or (NJet >= 3 and NbJet >= 2)

B Tagging Working Points	30	50	60	70	77	80	85	90
Jets Combinations	4-0 or 3-2	4-1 or 3-2	4-1 or 3-2	4-2 or 4-1	4-2	4-2	4-2	4-2
Significance	1.62	1.9	1.97	1.96	2.03	2.2	2.2	2.07

- Signal events number for the best jet combinations and Run1 jets' requirement under each b-tagged efficiency

B Tagging Working Points	30	50	60	70	77	80	85	90
Jets Combinations	4-0 or 3-2	4-1 or 3-2	4-1 or 3-2	4-2 or 4-1	4-2	4-2	4-2	4-2
ttH	73	53	62	62	33	36	43	52
Run1 Jets' Requirement	4-1 or 3-2							
ttH	34	53	62	72	74	77	81	87

77% b-tagging working point and (NJet >= 4 and NbJet >= 1) or (NJet >= 3 and NbJet >= 2) are chosen as the baseline for Moriond2016

Selection Optimization -- Lepton ID and Isolation tool

Task2: Lepton ID and Isolation tool Study

Goal: Isolation working point and lepton ID

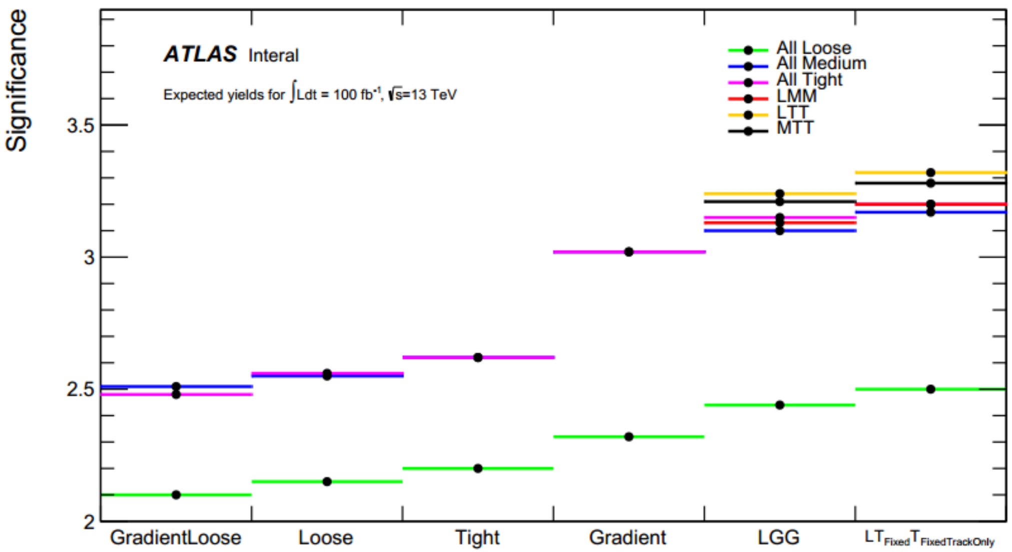
Working point	Objects	Calo isolation	Track isolation	Combined isolation
Tight	96%	99%	95%	
LooseTrackOnly	all leptons	-	99%	99%
Loose	all leptons	99%	99%	99%
Gradient	all leptons	ε=0.1143*pT+92.14 %	ε=0.1143*pT+92.14 %	ε(25 GeV) = 90%, ε(60 GeV) = 99%
GradientLoose	all leptons	ε=0.057*pT+95.57 %	ε=0.057*pT+95.57 %	ε(25 GeV) = 95%, ε(60 GeV) = 99%
FixedCutTight (previously EL0p06)	electrons	Cut: topoetcone20/pT < 0.06	Cut: ptvarcone20/pT < 0.06	-
FixedCutTightTrackOnly (previously MU0p06)	muons	-	Cut: ptvarcone30/pT < 0.06	-
FixedCutTightTrackOnly	electrons	-	Cut: ptvarcone20/pT < 0.06	-
FixedCutLoose	electrons	Cut: topoetcone20/pT < 0.2	Cut: ptvarcone20/pT < 0.15	-
FixedCutLoose	muons	Cut: topoetcone20/pT < 0.3	Cut: ptvarcone30/pT < 0.15	-

- Lepton ID: loose, medium, tight

Jets’requirement:
4 jets of which at least one must be b-tagged or exactly 3 jets of which at least 2 are b-tagged

Different Isolation working points and lepton ID requirements are applied to
lep0, lep1, lep2

Selection Optimization -- Lepton ID and Isolation tool



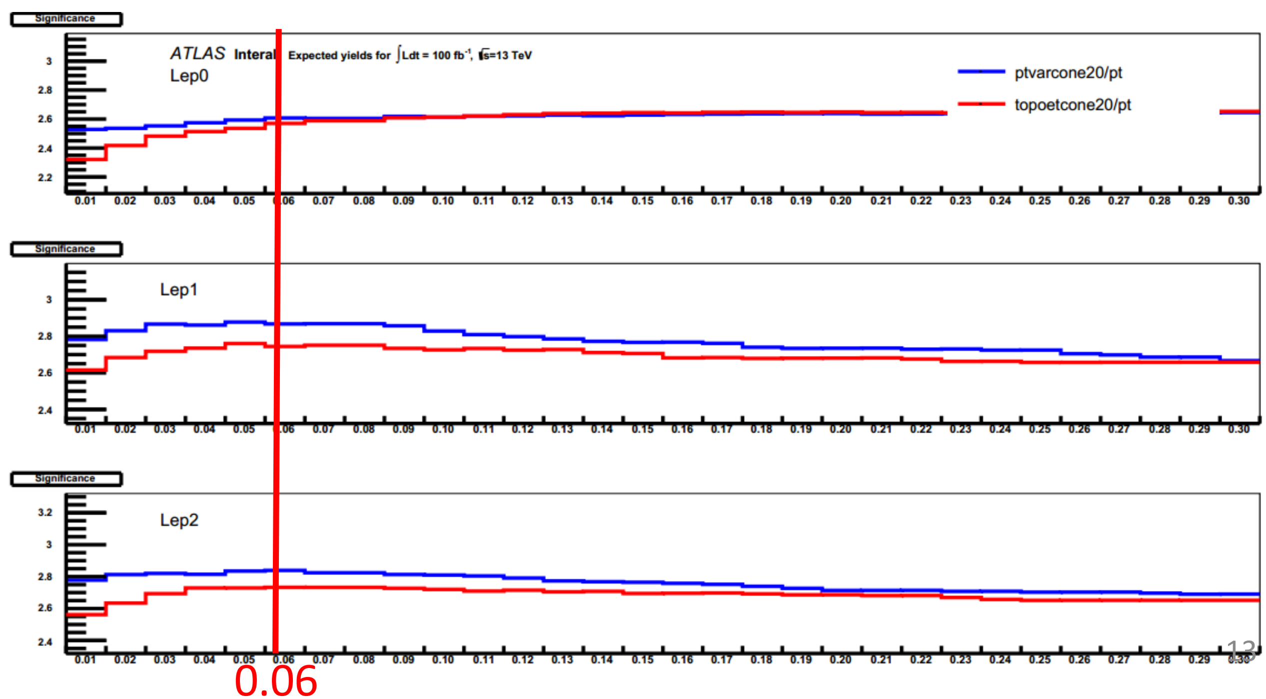
LTT and $LT_{Fixed}T_{FixedTrackOnly}$:

lep0 is loose ID and loose isolation;

lep1, lep2 are tight ID and FixedCutTight(electron) FixedCutTightTrackOnly(muon)

is the best

Value scanning under LTT LLL cuts on ptvarcone/pt and topoetcone/pt individually on 3 leptons



Summary

- **Electron Identification menu**
  - released in the official ATLAS offline and trigger software framework

- **Selection Optimization of ttH 3l final states**
  - 77% b tagging working point, jet multiplicity
  - Isolation working point and lepton ID
  - Baseline for Moriond 2016

Plan for the next

- **Estimation of fake background events (take the run1 strategy)**
- **Matrix Method study**
  - Preliminary results and need more statistics and tests


Thank you for your attention!

Backup

Variables used in the EID menu

	Variables	Loose	Medium	Tight
Shower Shape	Eratio(Demaxs1)	x	x	x
	wstot	x	x	x
	Weta2	x	x	x
	Reta	x	x	x
	Rphi		x	x
	f3		x	x
	Rhad	x	x	x
	Rhad1	x	x	x
Match	deltaEta	x	x	x
	deltaPhi			x
	EoverP			x
Track Q	nSi	x	x	x
	nPix	x	x	x
	nBlayer	x	x	x
	F HT		x	x
	nTRT			x

Discriminating variables used in the calorimeter are generically referred to as “shower-shapes”

Variables' list

Variable	Definition
nSiHits	Hits on silicon detector
nPixHits	Hits on Pixel detector
nBlayerHits	Hits on b-layer
nTRTHits	Hits on TRT detector
TRatio	TRTHighThresholdHits over TRT hits
d0	Longitudinal distance between collision point and primary vertex
deltaEta	deltaEta between cluster and track
f3	Fraction of energy deposit in the third layer of electromagnetic calorimeter
f1	Fraction of energy deposit in the first layer of electromagnetic calorimeter
rHad	Et of hadronic calorimeter over Et of electromagnetic calorimeter
rHad1	Et of first layer of hadronic calorimeter over Et of electromagnetic calorimeter
Reta	In second layer of electromagnetic calorimeter,energy deposit in the middle 3 cells over the whole 7 cells in eta direction
Weta2	Horizontal width of shower in the second layer of electromagnetic calorimeter
wstot	Shower width in the first layer of electromagnetic calorimeter
DEmaxs1	Difference between largest and second largest energy deposits in first layer of electromagnetic calorimeter over total energy deposit

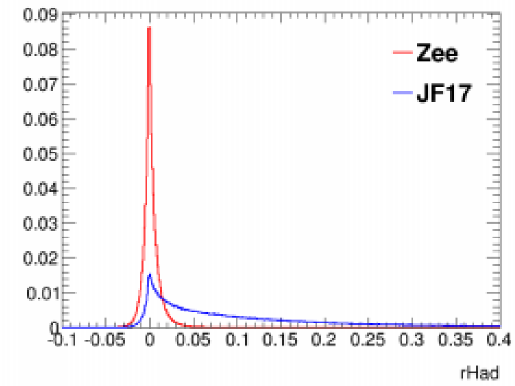


Figure 3: rHad

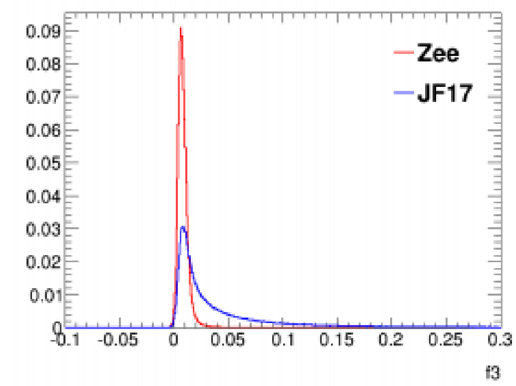


Figure 4: f3

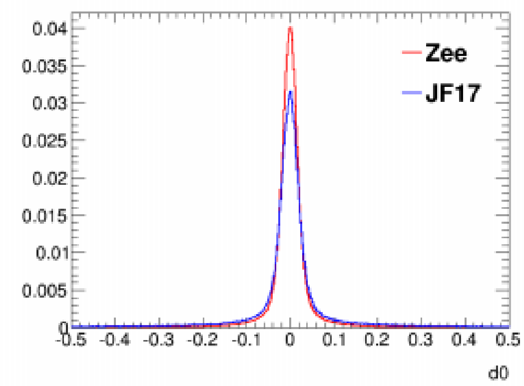


Figure 5: d0

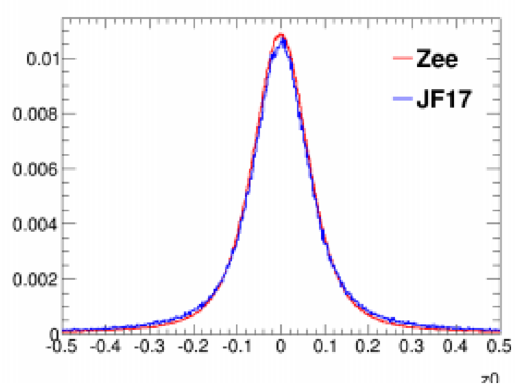


Figure 6: z0

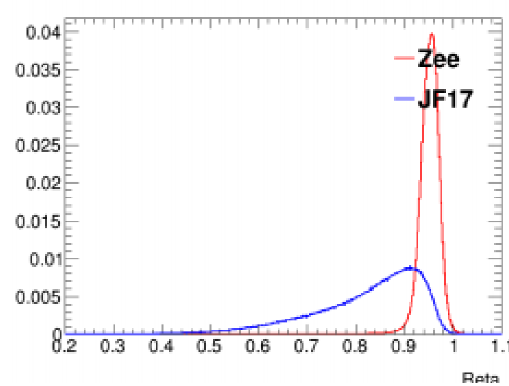


Figure 7: Reta

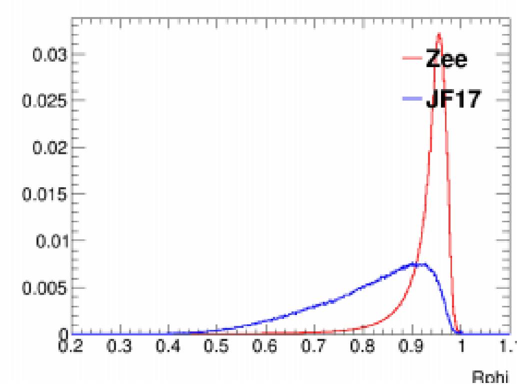


Figure 8: Rphi

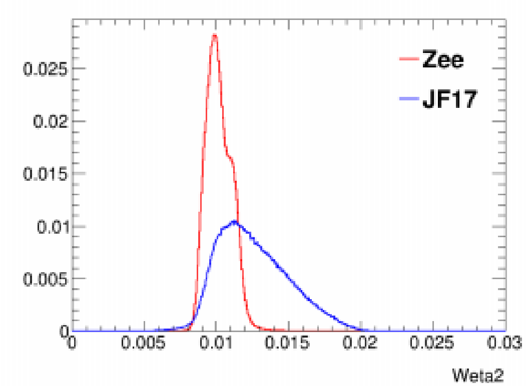


Figure 9: Weta2

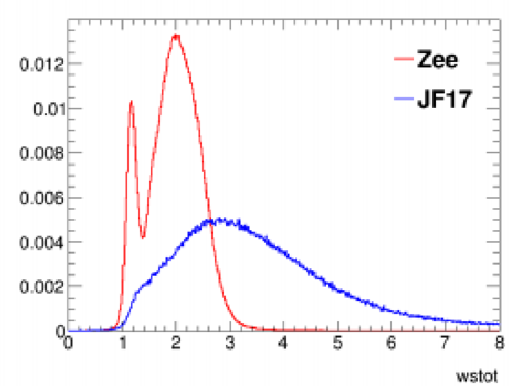


Figure 10: wstot

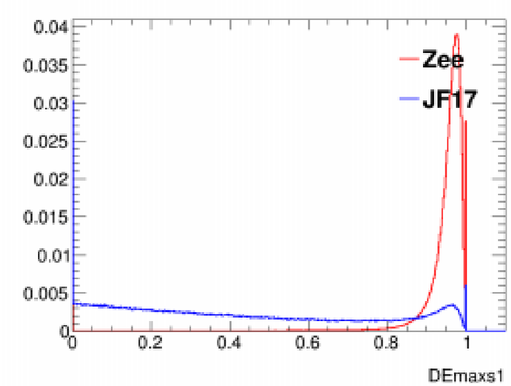


Figure 11: DEmaxs1