SEARCH FOR THE STANDARD MODEL HIGGS BOSON DECAYING INTO $B\bar{B}$ AND PRODUCED IN ASSOCIATION WITH A TOP QUARK PAIR IN THE ATLAS EXPERIMENT

May 9th 2016
Pheniics Days
E. Le Quilleuc, supervisors H. Bachacou, F. Déliot - CEA Saclay
Intro: LHC, ATLAS experiment overview

I. $t\bar{t}H$ analysis
   - Higgs boson production modes
   - $t\bar{t}H$ decay modes accessible by the ATLAS detector
   - Standard/boosted signatures

II. $b$-tagging in ATLAS
   - Introduction
   - Secondary vertex tagger
   - Show some studies/improvements
LHC AND ATLAS EXPERIMENTS

• LHC (Large Hadron Collider)
  • circular hadron collider located at Cern (27km circumference)
  • designed proton-proton (pp) center of mass energy \( CME = 14 \text{ TeV} \)
  • designed luminosity \( L \sim 10^{34} \text{ cm}^{-2} \text{s}^{-1} \)

• Operations
  - 2011 CME = 7 TeV, \( \int L \, dt \sim 7 \text{ fb}^{-1} \)
  - 2012 CME = 8 TeV, \( \int L \, dt \sim 20 \text{ fb}^{-1} \)
  - 2015 CME = 13 TeV, \( \int L \, dt \sim 3 \text{ fb}^{-1} \)

• ATLAS : a multi-purpose detector which probes a range of high energy processes.

• Purpose of the experiment:
  • precise SM measurements, like Higgs boson properties
  • Discover new physics with very energetic particles in the final state

ATLAS detector
I. TTH ANALYSIS IN ATLAS
In 2012, Higgs boson discovery by ATLAS and CMS (gg fusion largest cross section)

- 8 TeV → 13 TeV, factor ~2 increase of gg fusion, VBF & VH cross sections, while ~4 tH cross section (opening of the phase space)
- The tH cross section prop. to the (top Yukawa coupling)$^2$ → its precise measurement is essential in order to constrain the Higgs mechanism of the SM (Yukawa couplings prop fermion masses) and BSM theories
- $t$-quark decay $\rightarrow Wb$

- Note: lepton = electron or muon

- Higgs boson branching ratios:

- $t\bar{t} \rightarrow \text{dileptons}$: very clean signature but low statistics

- $t\bar{t} \rightarrow \text{lepton+jets}$: compromise between clean signature and fair statistics

- $H \rightarrow b\bar{b}$: highest branching ratio of the SM
TTH 8 TEV RESULTS

Signatures in the signal region (resolved)

• Lepton + jets: 1 charged lepton, >= 6 jets, >= 4 b-tags

• dilepton: 2 charged leptons, >= 4 b-tags

• Main background: $t\bar{t}$ + jets

Signal strength for individual channels and their combination
BOOSTED TTH ANALYSIS

- **Boosted ttH analysis**: top/Higgs are produced at high momenta
  - decay products are well separated
  - can be merged into large radius jets

- **Signature**: 1 charged lepton,
  1 large-R jet that contains hadronic top
  1 large radius jet that contains Higgs boson

- Suffers from very low statistics
- Currently working on the selection results will be shown with 2016 data
II. B-TAGGING IN ATLAS
B-TAGGING IN SHORT

• **b-tagging** ≡ identification of jets coming from the hadronization of *b*-quarks

• Motivation :
  • Many interesting physics processes contain *b*-quarks in the final state (t-quark and Higgs boson)
  • QCD background processes mainly produce *light*-jets (from *u*, *d*, *s*-quarks or gluons)

• **b-hadrons have** a relatively large lifetime = O(ps)

→ experimentally tracks from *b*-hadron decays
  • do not point to the hard pp collision, called primary vertex (PV)
  • cross each other into a single geometrical point called Secondary Vertex (SV)

• **Main background of b-tagging** :
  - **light** (*u*)-jets (fake *b*-hadron vertices like V0 decays, material hadronic interaction, *γ* → e⁺e⁻, detector resolution effects
  - **c-jets** (from hadronisation of *c*-quarks)
• $b$-hadron almost always decays into $c$-hadron with lifetime $O(\text{ps})$

• JetFitter (JF) : reconstruct vertices from $b$-hadron and $c$-hadron decays along the same $b$ flight axis
  
  • Tries to constrain jet tracks to intersect the $b$ flight axis ( = the jet direction for the first fit iteration)
  
  • Fit vertices and the $b$ flight axis and keep vertices high significance (1-track or multi-track vertices)

• JetFitter reco eff. increases linearly with jet $p_T$ for $u$-jets

\[ \text{"Reco" efficiency} = \frac{Z}{Y} \]

$Z$ = specific vertex reconstructed (no llr cut)

$Y$ = all $b$-jets
**JF TRACK PT SELECTION, HIGH PT**

- **Explanation**: number of fragmentation tracks (i.e., from hadronisation) increases with jet pT, more collimated tracks

  ![Fraction of Tracks vs Jet pT](image)

  **Origin of track fraction in a b-jet as function of jet pT**

- **Select tracks for vertex reconstruction with higher pT**: cut proportional to the ~ sum of tracks pT

  **Example**: track pT > 1.5 % * sum track pT → 15 % improvement of the rejection of u-jets
The vertices mass, decay length etc can be exploited to separate $b$-jets from $u$-jets.

**Flight dir. DR**: pseudo-angle between $\mathbf{p}_{\text{vtx}}$ and the $b$-flight axis.

- **$u$-jets**: peak near 0 for $u$-jets due to V0s decays (into 2 charged particles).
- **$b$-jets**: larger tail because the B-hadron decays into charged and neutral particles.

Adding flight dir. DR improves the JetFitter c-jet rejection by 20% at a 65% b-jet efficiency.
CONCLUSION

• $t\bar{t}H$ cross section measurement can be improved by the boosted analysis with 13 TeV data
• Boosted channel is under study with optimization of the selection
• $b$-tagging is an important feature for search of interesting processes at the LHC
• We improved the secondary vertex reconstruction
  • At high jet pT, with a tighter selection on tracks pT
  • By adding new information to $b$-tagging final discriminant
BACK UP
JETFITTER TRACK SELECTION @ HIGH JET PT
BDT STUDIES
MV2 DEFAULT : 2 TRAININGS

- 2 mv2 obtained with same training settings but with 2 statistically independent training samples → we evaluate their performances

Significantly large discrepancy between the evaluation of the 2 trainings

(!= overtraining here)

• Issue origin : in training, jets with SV1 and JF vertices (good jets) are treated differently than jets without these reconstructed vertices (bad jets)

• Bad jets are reweighted with factor 10e-6
**BDT WEIGHT**

- Comparison of BDT weight obtained from 2 statistically independent trainings for good/bad, signal(b) and background

- ~ good agreement of 2 BDT weights for good jets

- Large differences on the tails of BDT dist. for bad jets

- **Key point**: most u-jets in ttbar (98%) are bad jets. The discrepancy at low BDT for bad u-jets leads to huge differences in the evaluation since bad-jets are not down weighted
FAIR COMPARISON C00

- Comparison
  - **AtLeast1Tag**: slightly different new good jet definition = at least 1 of the 3 taggers IP, SV1, JF is defined (~ very similar to IP only since IP is very inclusive)
  - **mvx(c00)**: retraining with default mv2c00 settings
  - C-fraction issue forced us to consider the case c00 = “no charm in the background” → fair comparison with new training
SOME OPTIMIZATIONS

- **Minimum Node Size (MSN)**: minimum fraction of events contained in a leaf during training: \# jets in leaf / total \# jets in training

- The new training takes more jets because it is more inclusive → more events in fraction denominator → must reduce the Minimum Node Size

- Default value: 0.5%, new value: 0.1%

- For both lpxD, 3 discriminant variables
  - Llr\_IP = likelihood(IPb) - likelihood(IPu)
  - Llr\_IP\_c = likelihood(IPb) - likelihood(IPc)
  - Llr\_IP\_cu = likelihood(IPc) - likelihood(IPu)

  only 2 are linearly independent, ex:
  
  Llr\_IP\_cu = Llr\_IP - Llr\_IP\_c
JF TRACK PT SELECTION, HIGH PT

• Cut on JF track pT proportional to Scalar sum of pT over selected tracks
• With a slope $a = 0.5\%, 1.5\%, 3.5\%$
• Sample : last MC15 $Z' \rightarrow t\bar{t}$, mass $= \{0.4, 1, 3, 4\}$ TeV
• Kinematic selection (jet) : $p_T > 300$ GeV & $|\eta| < 2.5$
• Pile-up removal : JVF $> 0.5$
• Training with only JF variables