Search for the SM Higgs boson in the $ttH \to \ell\nu b\ jjb\ bb$ channel

(m_H=120 GeV/c^2)

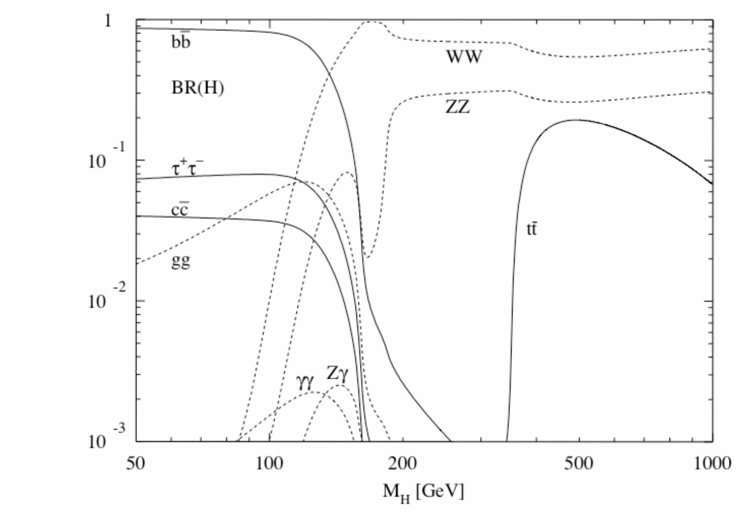
Georges Aad, Lorenzo Feligioni, Alexandre Rozanov,
Laurent Vacavant, Jean-Baptiste de Vivie
Physique ATLAS France
Seignosse 12-09-07

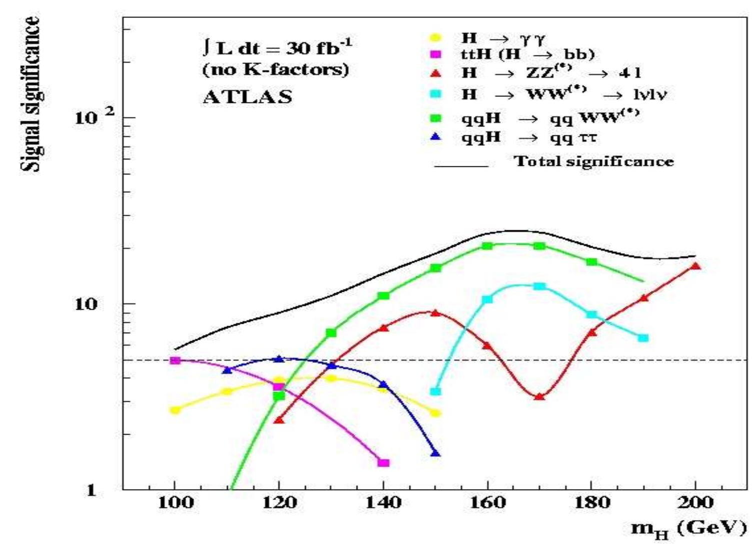
- Introduction
- ttH analysis in the lepton + jets channel
  - Analysis steps
  - Key points
    - combinatorial background
    - physics background shape
  - ttjj backgrounds with different Monte Carlo generator
- Charged Higgs analysis in tbH channel (Mossadek Talby, Remi Zaidan)
- ttH(H→WW) analysis (Feng Lu*, Emmanuel Monnier, Lian You Shan*, Huaquio Zhang*)

*IHEP Beijing/CPPM

Light Higgs boson in ATLAS

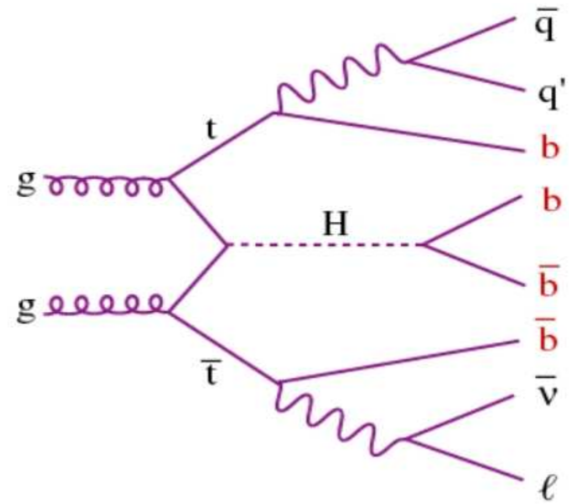
- Preferred scenario by the electroweak fit
  - $m_H = 76^{+33}_{-24}$ GeV/c^2
- Leading disintegration:
  - $m_H \leq 135$ GeV/c^2
    - $H \rightarrow bb$
  - $m_H \geq 135$ GeV/c^2
    - $H \rightarrow ZZ^{(*)}$
    - $H \rightarrow WW^{(*)}$
- "light" Higgs discovery:
  - Difficult situation
  - Need to combine more than one channel
    - $ttH \rightarrow ttbb$
    - $gg \rightarrow H \rightarrow \gamma\gamma$
    - VBF



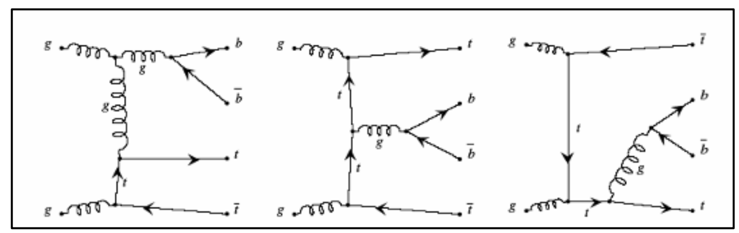


The ttH(→bb) analysis

- Challenging channel for the discovery of a light Higgs boson at ATLAS
- Yukawa coupling
- Lepton + jets final state (BR=0.29)
  - Cleanest signature
  - 1 isolated lepton (electron or muon)
  - At least 6 jets, at least 4 b jets
    - need for large b-tagging efficiency (ε^4_b)
  - Missing energy from neutrino



ttbb QCD production →



Data	Cross section (pb)	K-factor
ttH	0.519 (LO) × 0,68 (H→ bb)*	1.2
ttbb(QCD)	8.1 (LO)	1.2 ?!?
ttbb(EW)	0,89 (LO)	1 ?!?
tt (+jets)	760 (NLO)	-

* For m_H = 120 GeV/c^2

ttH(H→bb) analysis status

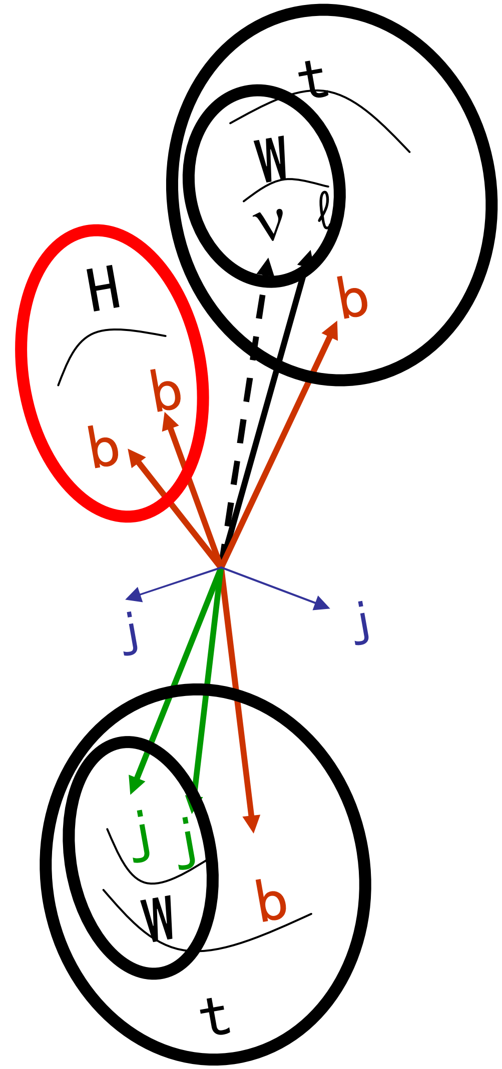
- First Analysis Results based on pre-CSC (rel 11) samples
  - combinatorial likelihood
    - reduce combinatorial background, minimal impact on Higgs mass spectrum
  - selection likelihood
    - exploit kinematical distributions to discriminate against physics background
- Studies carried out on several different aspects
  - Impact of combinatorial background on final results
  - Signal extraction techniques
  - Validation of mixed simulations for production of high statistics samples
  - Comparison of generators for background simulations
  - *On-going:*
    - *understanding differences in performances with CSC samples.*
    - *collaboration with Genova Group for Trigger studies*

Semi-leptonic Analysis: pre-CSC samples

- Reconstruction 11.0.41, ttH(bb) rootuple version 3.03
- Signal (tth): Samples 5340-5341
  - Pythia 6.3 (New showering)
  - 85100 events ~652 fb^{-1}
- Background (ttbb – QCD): Sample 5855
  - AcerMC + Pythia 6.3
  - 45500 events ~17 fb^{-1}
- Background (ttbb – EW): Sample 5214
  - AcerMC + Pythia 6.3
  - 18848 events ~77 fb^{-1}
- Background (ttjj): Sample 5212
  - MC@NLO + Herwig, filter to get 2 additional jets (15%)
  - 1085855 (14% negative weight) events ~7 fb^{-1}

Summary of pre-selection cuts
(1 lepton, > 6 jets, 4b)

- **Muon: Staco Collection (HightPt algorithm)**
  - p_T > 20GeV, $|\eta|$ < 2.5
  - χ^2/ndof < 20
  - Isolation cone 0.20 < 15 GeV
  - D0 vertex < 0.05 mm
- **Electron: Electron collection**
  - p_T > 25GeV, $|\eta|$ < 2.5
  - IsEM & 0x4ff
  - Pix hits > 1
  - Isolation cone 0.20 < 10 GeV
- **Soft muon: Staco Collection (HightPt & LowPt algorithm)**
  - Not passing muons cuts
  - 4 GeV < p_T < 100 GeV, $|\eta|$ < 2.5
  - χ^2/ndof < 4
  - Isolation cone 0.30 > 3 GeV
  - Removing overlap in LowPt algorithm
- **Jets: Cone 0.4**
  - Correction of the 4-vector by adding soft muons ($\Delta R < 0.4$)
  - AtlFast calibration (for out-of-cone)
  - Removing overlap with electrons ($\Delta R < 0.2$ & $p_{Telectron}/p_{Tjet} > 0.75$)
- **b-tagging:**
  - SV1+IP3D > 4
- **Neutrino p_z**
  - Solve second degree equation
  - Take $\Delta = 0$ approximation when $\Delta < 0$



Reconstruction

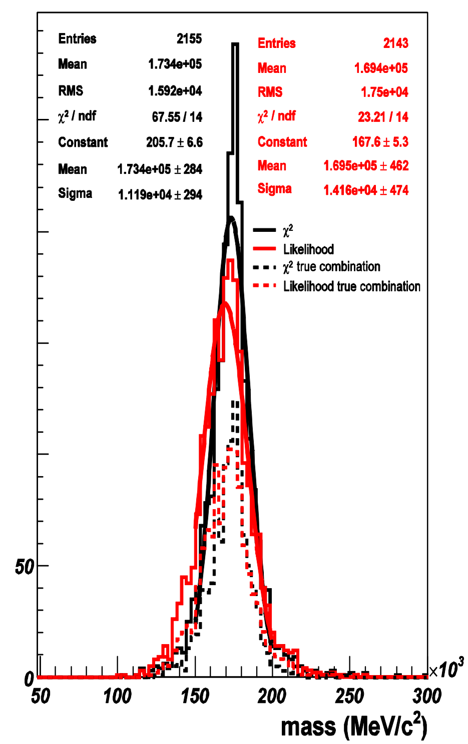
- Cut-based analysis steps
  1. Reconstruct hadronic W using all combinations of light jets pairs
     - Mass window cut: $|m_{jj} - m_W| < 25$ GeV/c^2
  2. Associate b tagged jets and W candidates to reconstruct leptonic and hadronic top quarks
     - Choose best combination by minimizing χ^2

$$\chi^2 = \frac{(m_{l\nu b} - m_t)^2}{(\sigma_{m_{lb\nu}})^2} + \frac{(m_{jjb} - m_t)^2}{(\sigma_{m_{jjb}})^2}$$

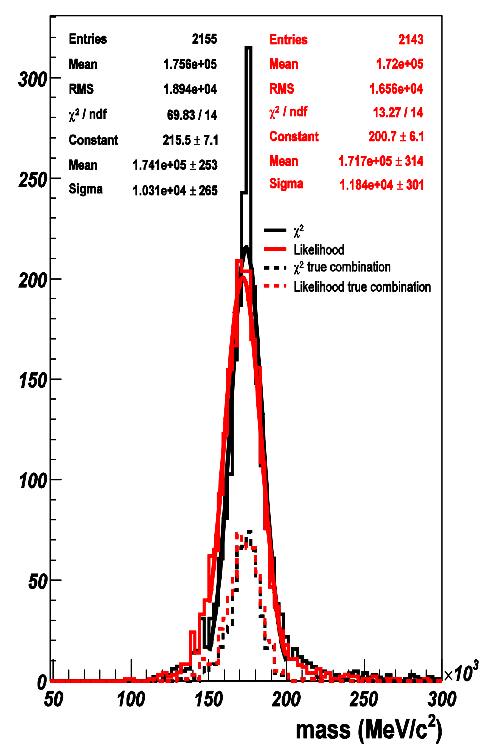
     - Cut: $|m_{jjb}\ (m_{l\nu b}) - m_t| < 25$ GeV/c^2

- Likelihood steps
  1. PDF's based on mass and angular distributions
  2. Choose combination with highest likelihood output
  3. Reject events with no combination passing a likelihood cut of 0.9

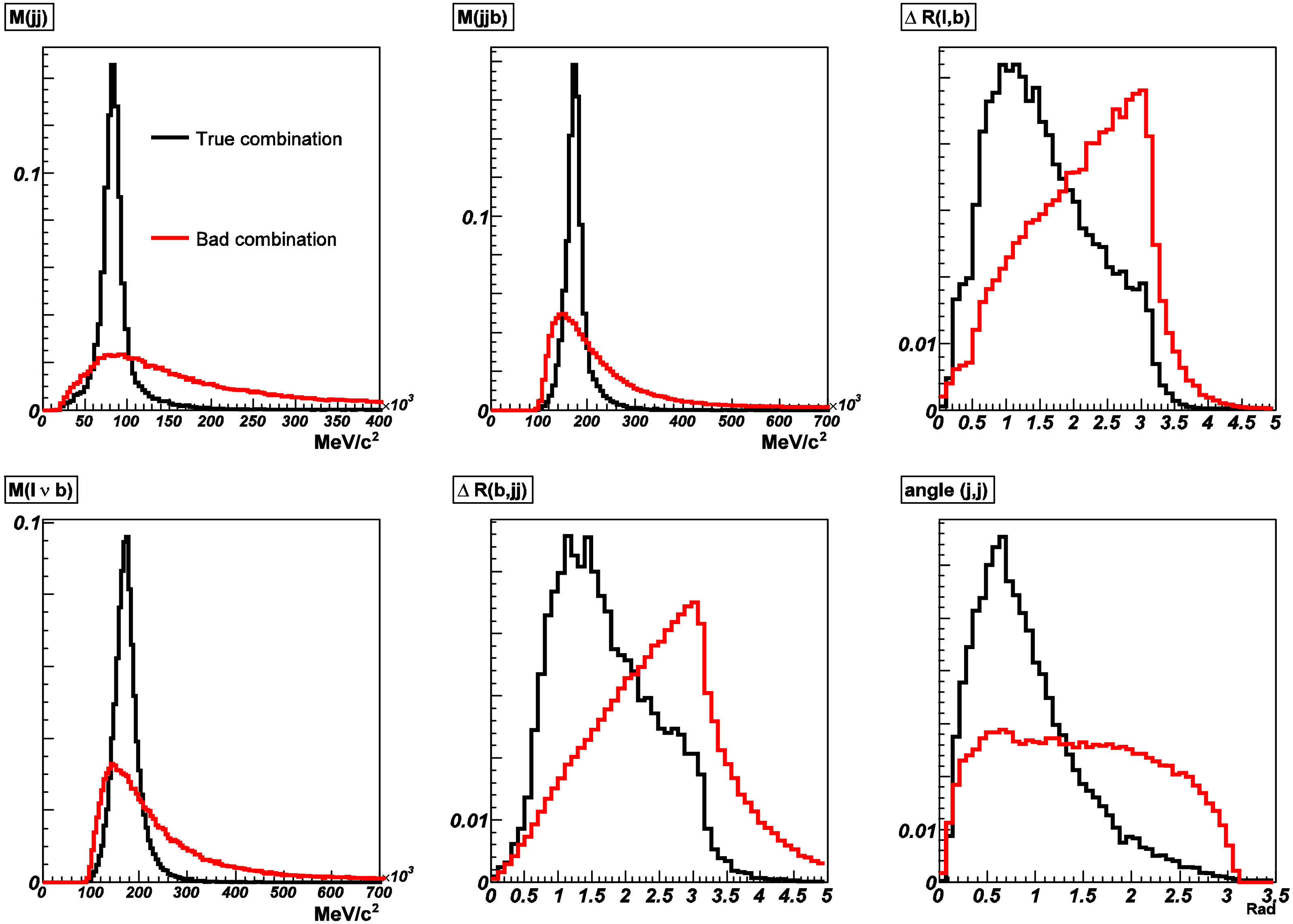
Leptonic top



Hadronic top

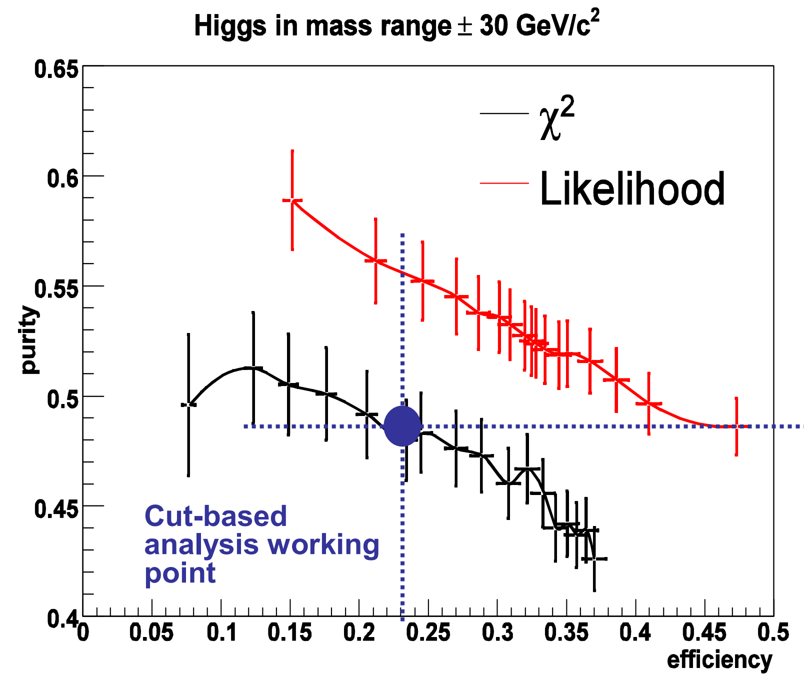


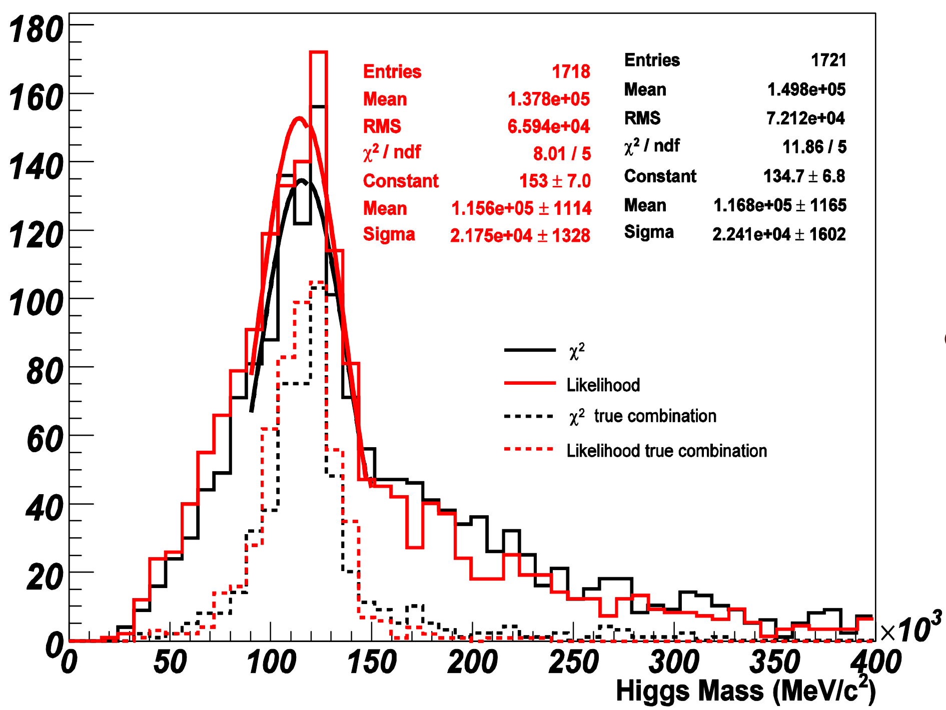
Combinatorial Likelihood templates



Likelihood vs cut-based

- Gain of efficiency vs purity
- Same efficiency
  - Gain 8% purity
- Same purity
  - Gain 22% efficiency

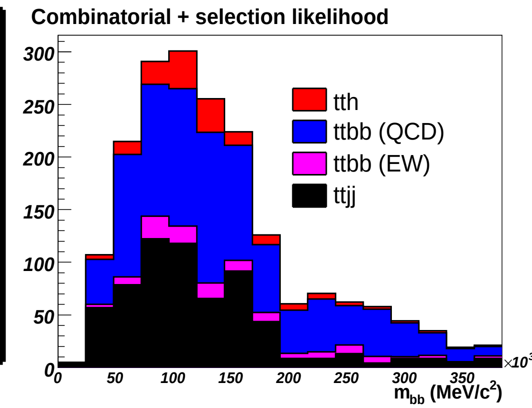


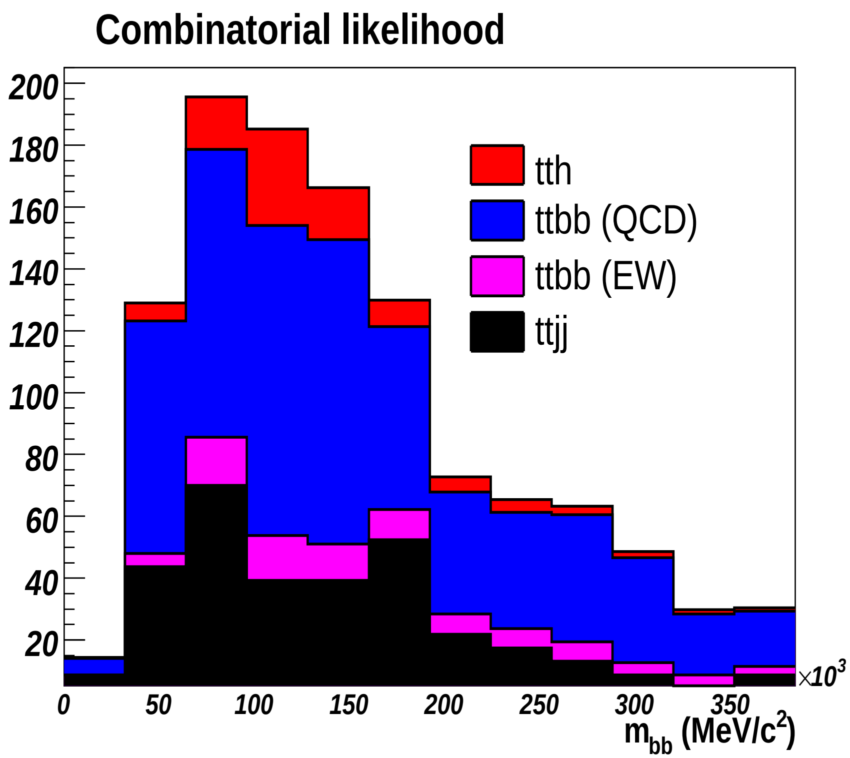


- Likelihood analysis slightly better for m_{bb}
  - smaller tails
  - more events in mass range

Results for 30 fb^{-1}

	tth normalized	ttbb normalized	ttbb (EW) normalized	ttjj (ttbb*) normalized (ttbb)	(S/√B)
Cut-based analysis	783 33	74 132	50 20	14 (35) 61 (153)	2.3
Combinatorial likelihood	1048 44	92 165	59 23	12 (45) 52 (184)	2.8





- Selection Likelihood studies:
  - Kinematical and b-tag variables used to build a selection likelihood
    - Small increase in significance
    - Bias in the reconstructed Higgs mass
    - No statistics to build pdf's

	tth	ttbb	ttbb (EW)	ttjj	(S/√B)
Selection Likelihood	66	252	32	140	3.2

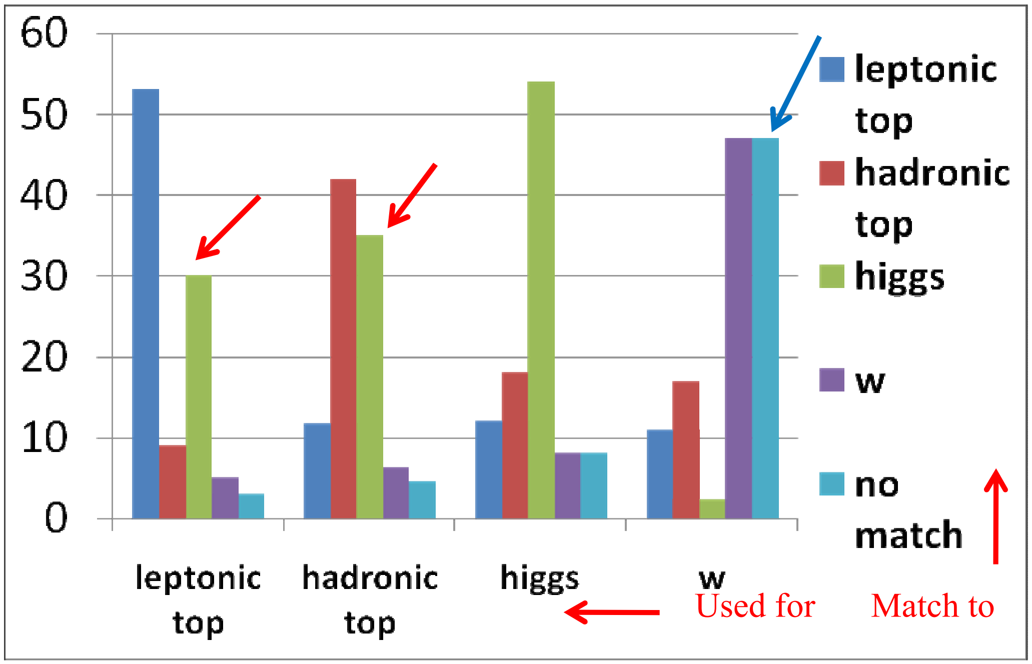
* ttbb double counting removed

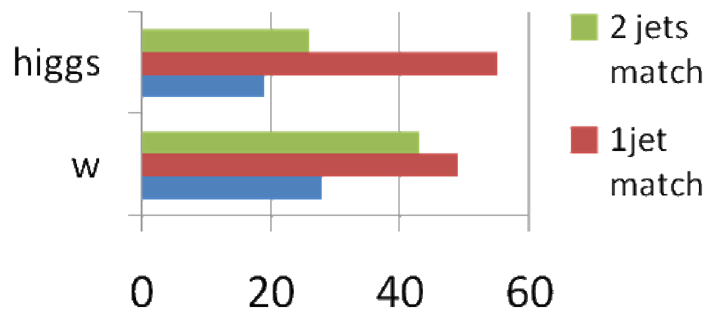
Studies for combinatorial background

It's clear that if the reconstructed objects do not match to truth we can not have the right combination

Matching after jets pre-selection and b-tagging

6 partons match = 34%	right btag = 25%
4 b partons match = 76%	right btag =58%
higgs b partons match = 85%	right btag =75%
W partons match = 47%	right btag = 31%





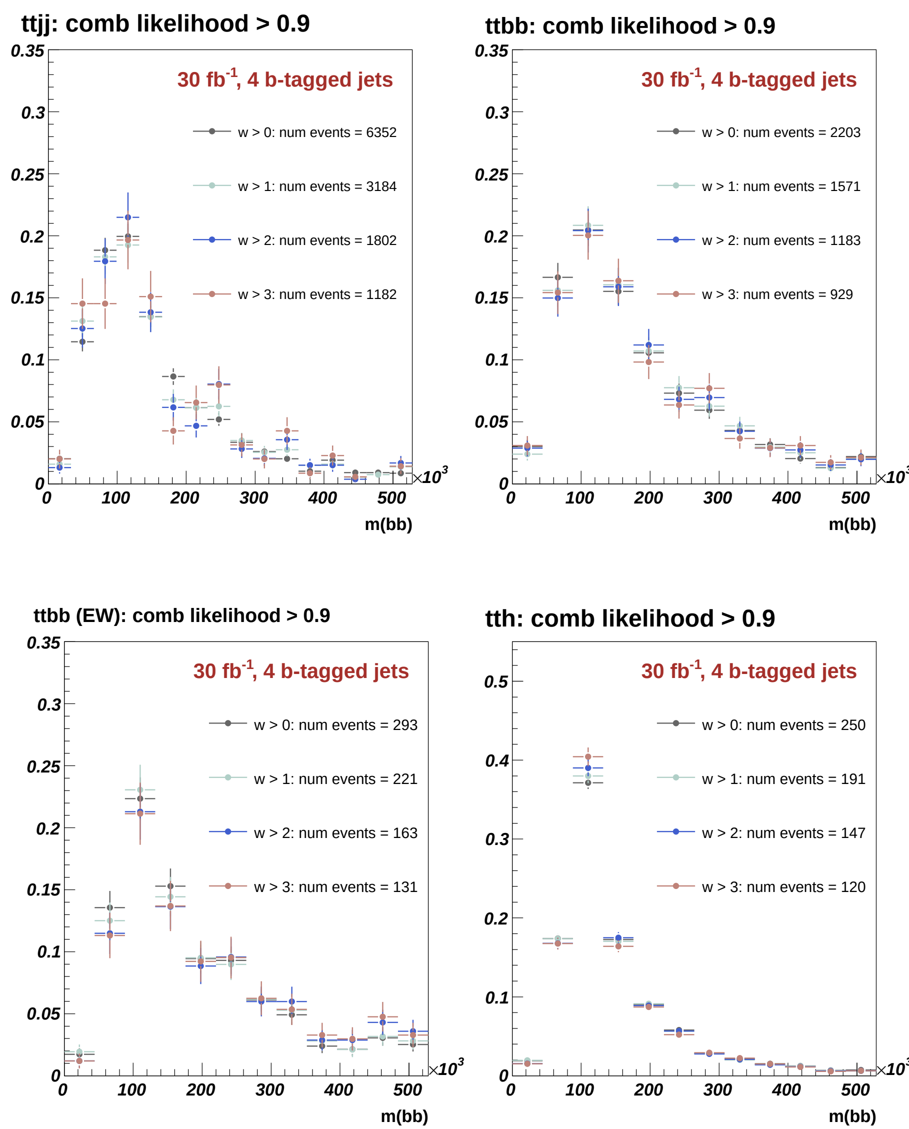
Big exchange of b jets between Higgs and top quarks

Lots of unmatched jets used to reconstructed hadronic W

For more details:
http://indico.cern.ch/getFile.py/access?contribId=3&resId=0&materialId=slides&confId=11888

Strategies for Understanding Background Shapes with DATA: Control Samples

- Reference analysis: combinatorial likelihood based preselection
  - Likelihood > 0.9
- Final working point of the analysis involve a tight cut on b-tag to suppress tt+jets background
  - b-tag weight > 4.
- Relaxing this cut means diluting the signal
- Shapes don't depend strongly on weight cut
  - Can go at loose b-tag to check background shapes
  - Loose dependence of b-tag shape modulation from weight value



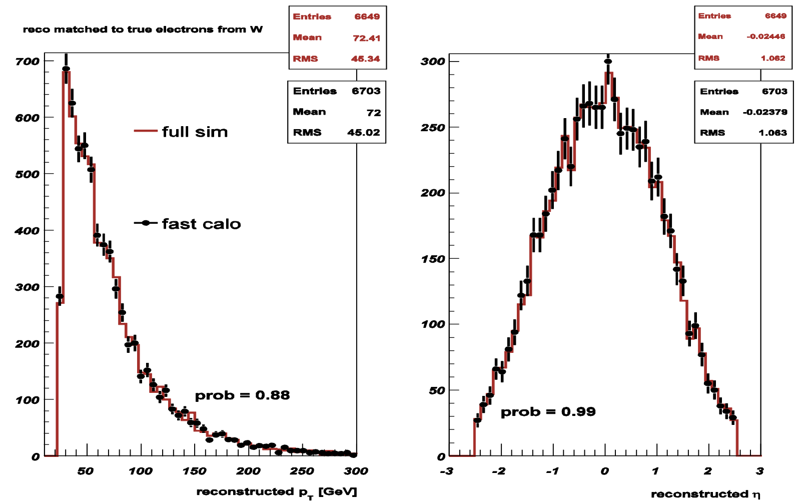
Fast Calo Simulation:
Validation for ttH events

- Kinematical variables are used to identify b-jets coming from the Higgs boson decay.
  - High combinatorial background content
    - ⇒ Similar shapes for signal and background
- High statistics needed to correctly evaluate shapes
  - Fast Calorimeter could be used to reduce simulation time
  - Fully simulated inner detector: correlation between b-tagging and kinematical distributions
- For this study:
  - used 18250 events from dataset:

trig1_misal1_mc12_V1.005870.ttH_poslepnu_jj_bb.recon.ESD.v12000601

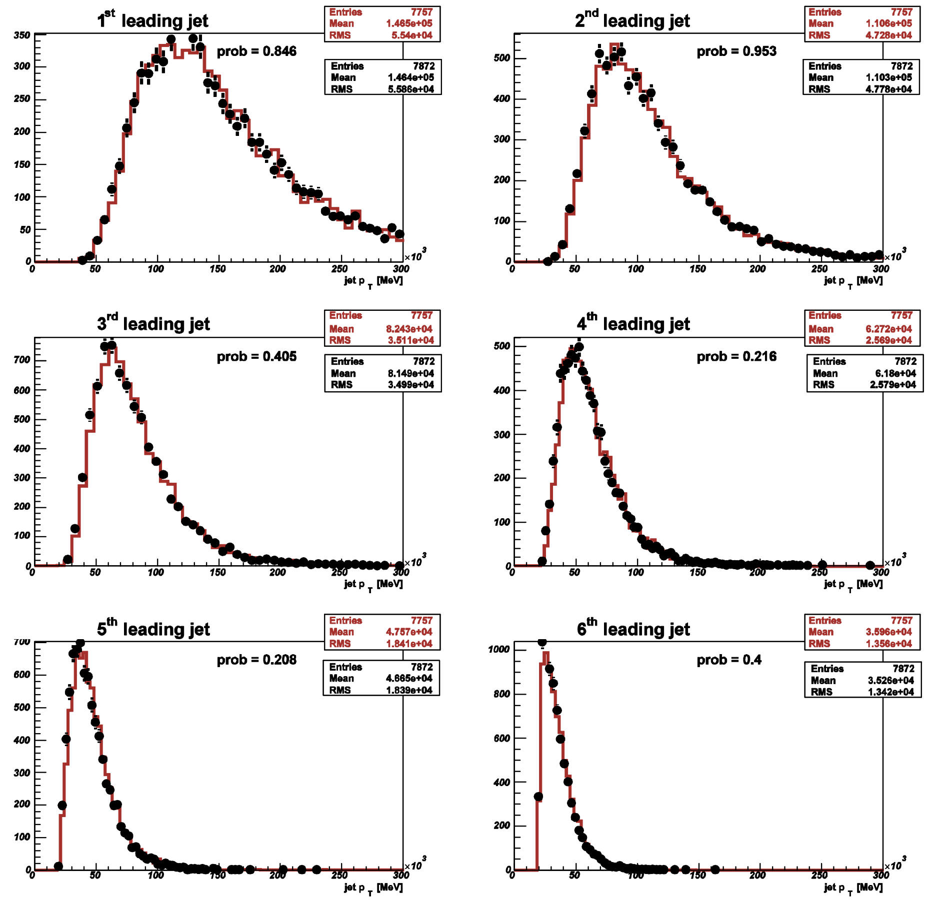
- Calorimeter re-simulated with [FastCaloSim-00-03-12]

Reconstructed electron →

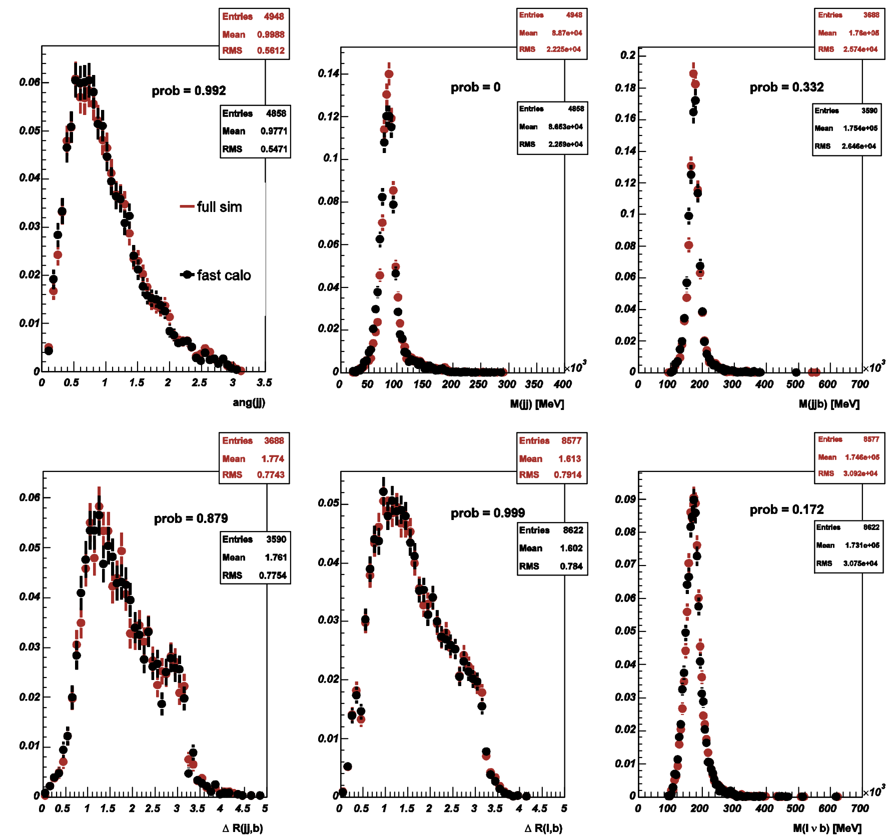


Fast Calorimeter Simulation Twiki
https://twiki.cern.ch/twiki/bin/view/Atlas/FastCaloSimttHHbb
Michael Duehrssen's talk at Glasgow
http://indico.cern.ch/getFile.py/access?contribId=33&sessionId=26&resId=2&materialId=slides&confId=12732

Fast Calo: final objects



Comb Likelihood Templates



Very good agreement, between the two simulations. Starting production of large ttjj sample next.

Systematic study of MC production
ALPGEN vs. MC@NLO

⇒ Try to use Alpgen to get a second estimate of the ttjj background

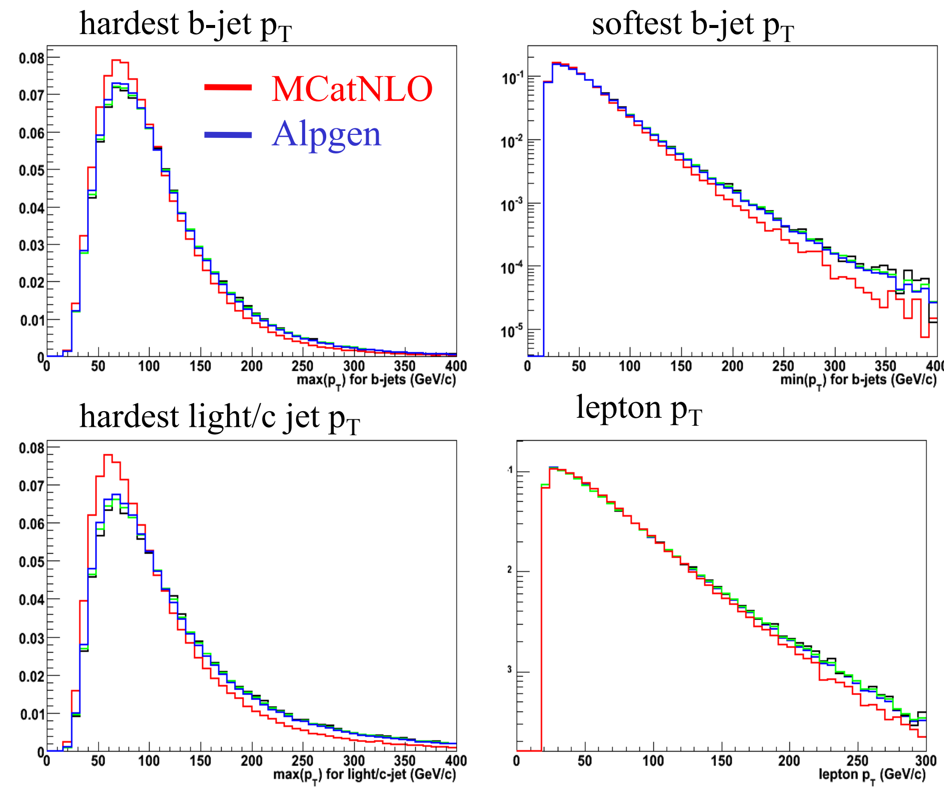
- ➢ very loose cuts (VL) : $p_T \geq 15$ GeV/c and $\Delta R_{jj} \geq 0.4$
- ➢ soft Higgs cuts (SO) : $p_T \geq 20$ GeV/c and $\Delta R_{jj} \geq 0.7$
- ➢ hard SUSY cuts (HO) : $p_T \geq 40$ GeV/c and $\Delta R_{jj} \geq 0.7$

✓ The fraction of events with 4 b-labelled jets is similar in MCatNLO and Alpgen ~ 40%

✓ VL and SO compatible

- 90% and 80% of the selected events are from the tt + ≥ 3 jets sample

New ALPGEN Version has bug fix for this sample cross section

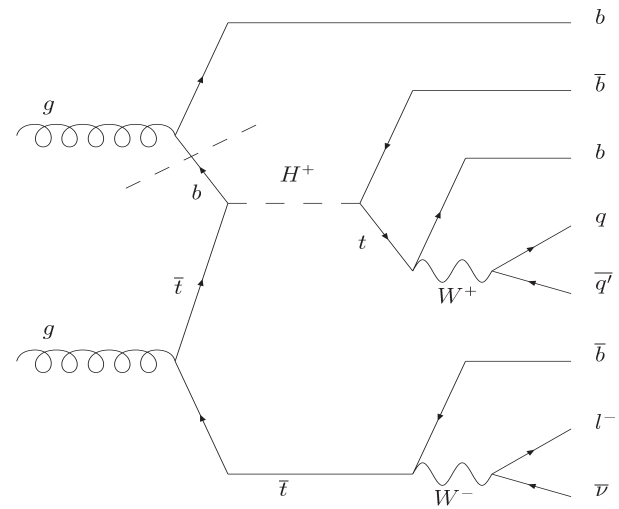


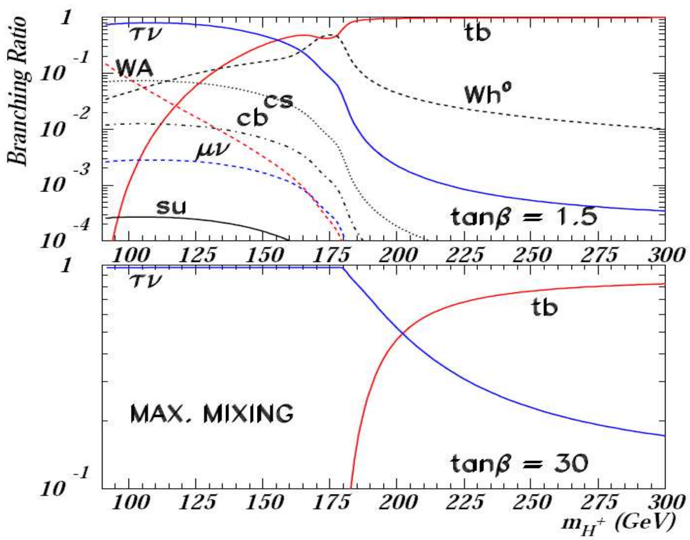
	MCatNLO* 5212	Alpgen** VL	Alpgen** SO	Alpgen** HO
events in mass window (30 fb^{-1})	~ 173	~ 310	~ 293	~ 157

Charged Higgs Boson

The gg/gb→t(b)H± ; H± → tb channel:

- Cross section: σ = 1.3 pb for m_H=250 GeV/c² tanβ=40
- Similar final state than the SM ttH channel:
    - Same Backgrounds, similar analysis...
- High BR for heavy Charged Higgs boson.
- Suffers from high SM background.





Concurrent channels:

Heavy Charged Higgs boson:

- gg/gb → t(b)H± ; H± →τν
- Lower branching ratio
- Cleaner channel with less background

Light Charged Higgs boson:

- t → H± b (H± → lν, H± → qqbar)

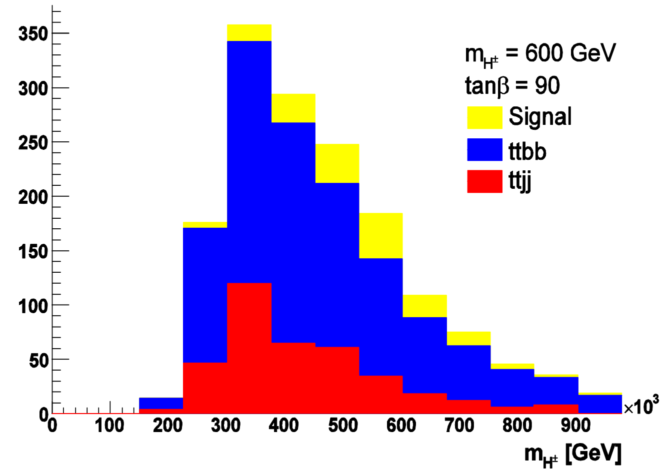
Charged Higgs Boson analysis

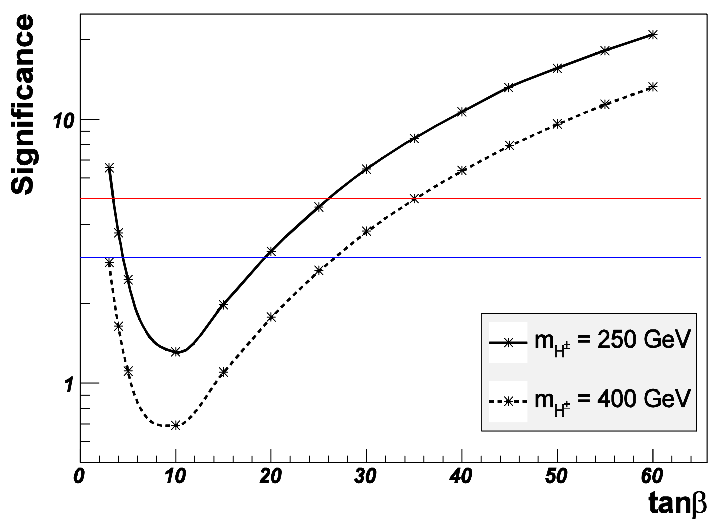
Pre-selection:

- ➢1 isolated lepton
- ➢at least 6 jet with 4 b tagged jets:

Reconstruction:

- ➢Combinatorial Likelihood function based on 8 kinematic variables



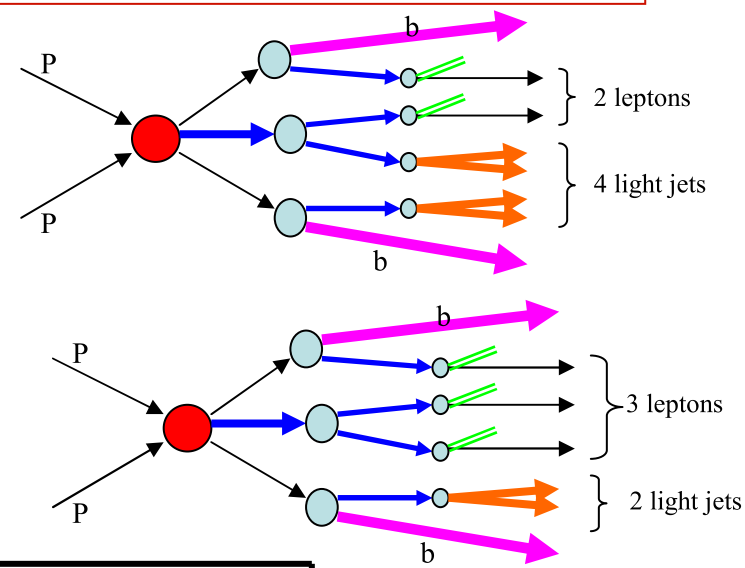


Challenging analysis:

- ➢Huge combinatorial background:
  - ➢500 possible combination per event
  - ➢25% purity for the Higgs boson!
- ➢The signal shows no peak in the invariant mass distribution:
  - ➢Shape analysis is very difficult
  - ➢Perfect understanding of the background is crucial

ttH(H→WW) analysis

- ttH→$2b4j2l2V_l$
  →6jets + 2samecharge lep. + MissEt
- Possible BGs:
  tt(1l), tt(2l), ttZ, ttW,
  tttt, ttWW, ttbb...
- ttH→$2b2j3l3V_l$
  →4jets + 3leptons + MissEt
- Possible BGs:
  tt(2l),ttZ,ttW,tttt,ttWW,ttbb



Normalized to 30 fb^{-1}	σ_{NLO} (fb)	2L	3L
ttH160 Pythia	10.9	52.6±1.0	0.1±0.04
ttH3l(160) Pythia	6.9	6.8±0.3	23.4±0.5
ttbar	461000	0	63.3(2 events)
ttW+nq	235.7	88.9±3.4	22.8±1.8
ttW	246	44.0±3.0	13.1±1.6
ttZ	144	44.2±2.8	33.4±2.4
ttbb	3091	13.2(8events)	0
tttt	3.2	2.1±0.1	0.65±0.04

S/sqrt(S+B):

2L: 3.3

3L: 1.9

Big Uncertainty from ttbar

(Lack of MC)

Conclusions

- tth analysis:
  - preliminary results with pre-CSC samples
  - likelihood based analysis: S/√B ~3 for 30 fb-1
  - finalizing the analysis using CSC data is in progress
  - start addressing more subtle, even though not less critical, issues (true as well for tbH analysis):
    - limits on the understanding of our combinatorial background
    - strategies for validating background simulations using early data
    - signal extraction
- Other analysis related to ttH in progress
  - tbH for the search of heavy charged Higgs decaying to top and b quarks
  - ttH(H→WW) for higgs search, also interesting for Yukawa coupling

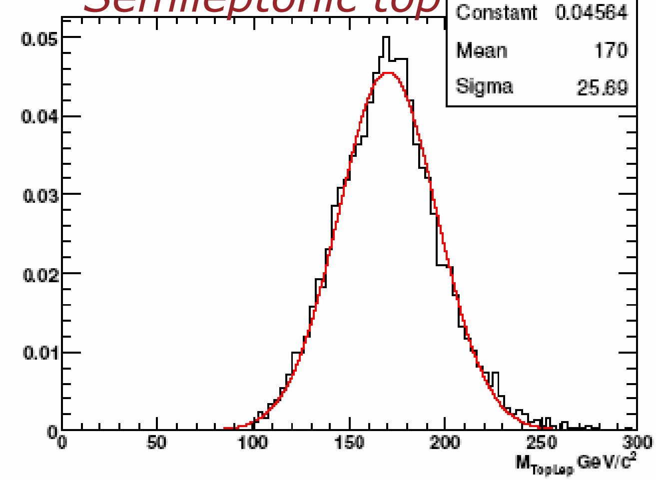
Backup

ttH(H→bb): CMS

- 3 final states considered
- ttH semileptonic:
  - Low energetic jets accepted if associa with tracks from PV
    - 3 Likelihoods
    - Constrained fit on W masses
    - b-tagging used to associate b top and higgs
    - Kinematic function discriminate hig energetic Higgs products

Cms note 2006 119

Semileptonic top



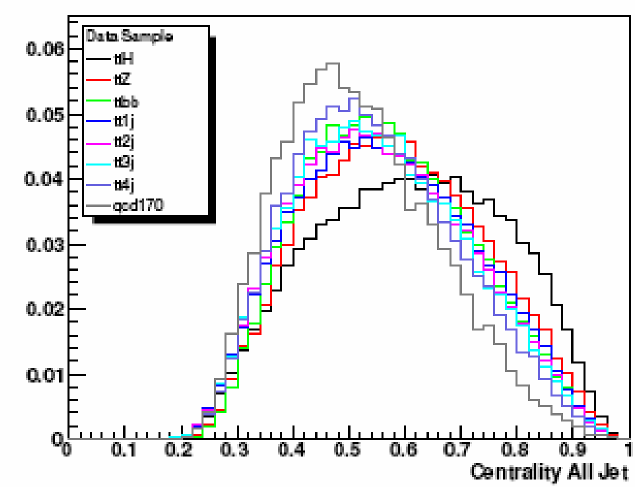
b-tagging

$\varepsilon_b = 54\%$
$R_c = 6.2$
$R_l = 62.5$

- Dilepton analysis
  - counting experiment
    - 2 opposite charged leptons (e,μ)
    - substantial missing energy (2 ν's)
- All hadronic
  - minimize χ^2 mass

$$\chi^2_{mass} = \left(\frac{m_{W^+} - m_{jj}}{\sigma(m_W)}\right)^2 + \left(\frac{m_{W^-} - m_{jj}}{\sigma(m_W)}\right)^2 + \left(\frac{m_t - m_{jjj}}{\sigma(m_t)}\right)^2 + \left(\frac{m_{\bar{t}} - m_{jjj}}{\sigma(m_t)}\right)^2$$

  - Using centrality variable



CMS summary

	s-lep μ	s-lep e	dilep	All had
S/√B	1.8	1.6	1.4	2.4