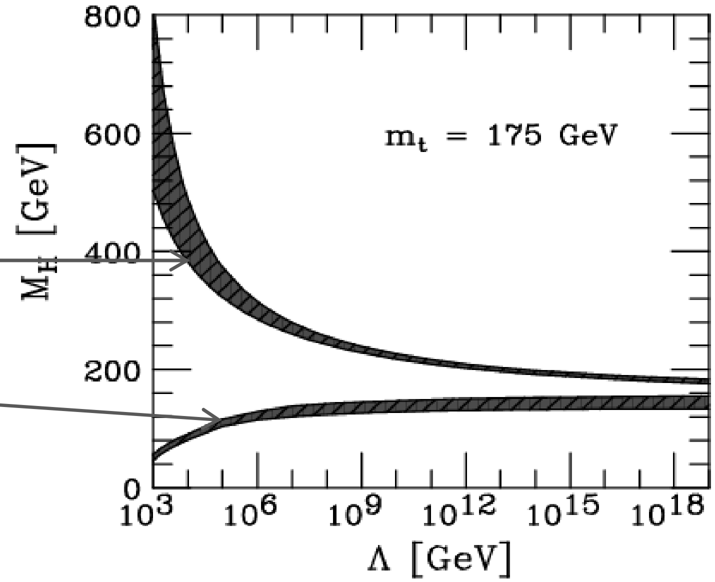
The ttH Physics Channel at the LHC

Georges Aad - CPPM Marseille
Jan Steggemann - RWTH Aachen
On behalf of the ATLAS and CMS experiments

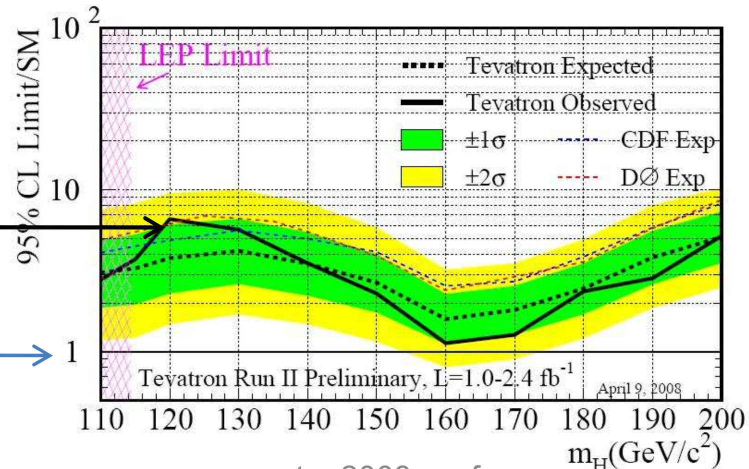
The SM Higgs Boson

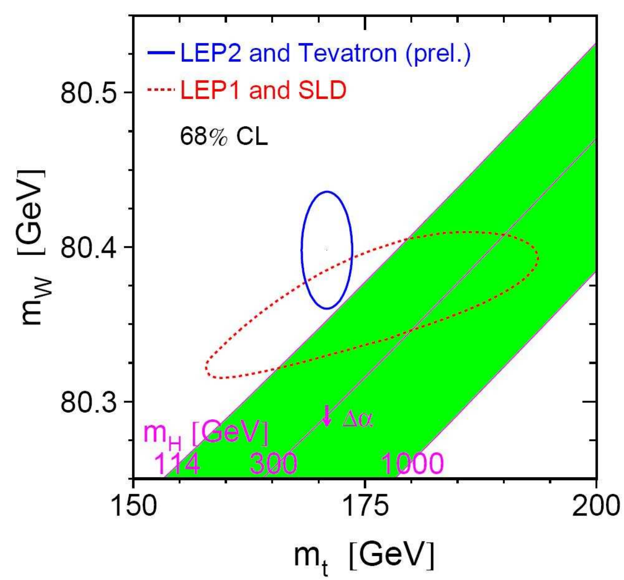
- Many theoretical arguments predict the presence of a Higgs boson with a mass domain accessible with the LHC energy
- Higgs mass theoretical constraint
  - Triviality bound:
    - $\Lambda_C \sim 10^{16(3)}$ GeV $\rightarrow m_H < 200(1000)$ GeV/c^2
  - Vacuum stability bound:
    - $\Lambda_C \sim 10^{3(16)}$ GeV $\rightarrow m_H > 70\ (130)$ GeV/c^2
- Direct limits (LEP)
  - $m_H > 114.4$ GeV/c^2 @ 95% CL
- Electroweak precision measurements fit
  - $m_H = 76^{+33}_{-24}$ GeV/c^2
  - $m_H < 144$ GeV/c^2 @ 95% CL



Tevatron limits: 5 times the SM cross section at low mass

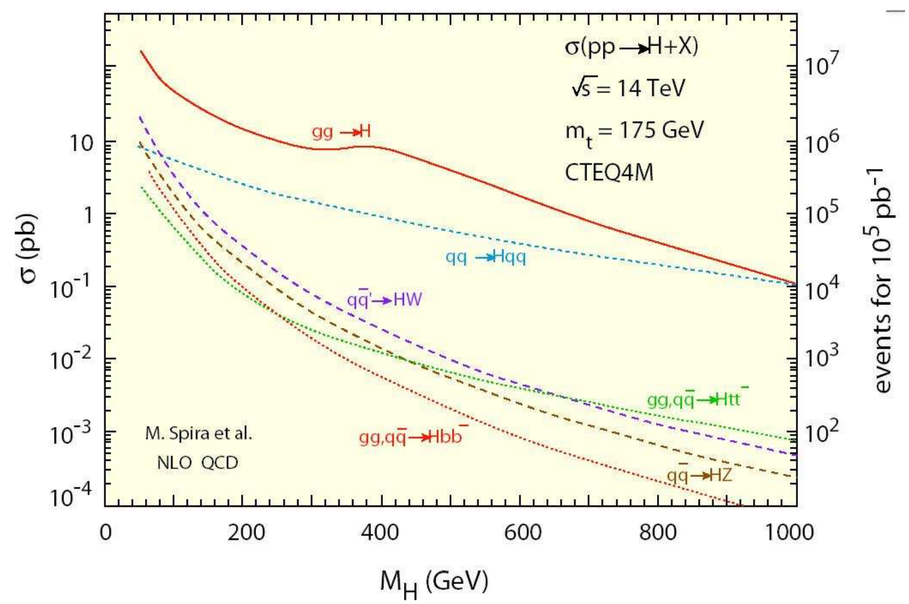
SM cross section

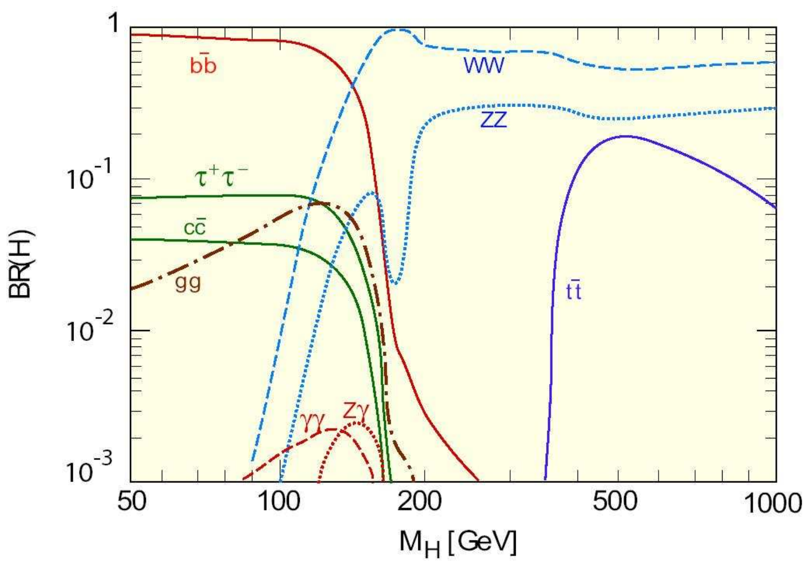




The SM Higgs at the LHC

- Production mechanism:
  - Dominated by gluon-gluon fusion
  - Vector Boson Fusion production 10 times smaller
  - Associated production with top quark pairs or gauge bosons
    - Important for low Higgs mass and Yukawa coupling measurements
    - ttH cross section : 0.519 pb (LO) @ 120 GeV/c^2, k-factor =1.2



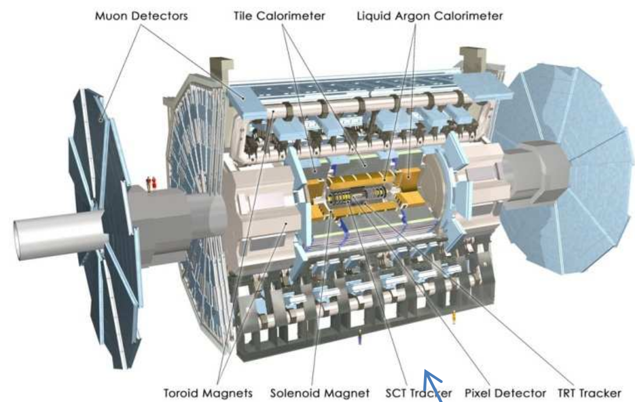


- Decay modes:
  - H→bb is the main channel for m_H < 135 GeV/c^2 (BR : 0.68 @ 120 GeV/c^2)
    - Direct production is swamped by QCD background
    - Only detectable via associated production
  - For m_H > 135 GeV/c^2 H → VV dominates

Difficult for low mass: we need to associate more than one channel

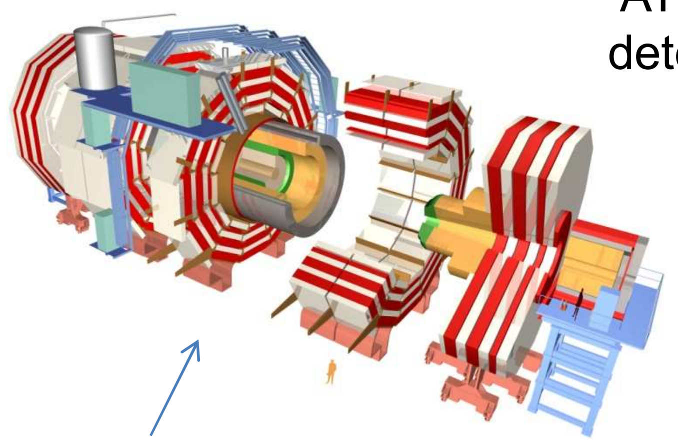
The LHC Experiments

- LHC:
  - Proton-Proton collisions @ 14 TeV
  - First run @ 10 TeV expected at the end of this summer
  - Luminosity:
    - Low luminosity regime ~$10^{33}cm^{-2}s^{-1}$
      - ~ 30 fb^{-1} between 2008 and 2011
    - High Luminosity regime ~$10^{34}cm^{-2}s^{-1}$
      - ~ 300 fb^{-1} by 2014/2015



ATLAS detector

- ATLAS and CMS:
  - Generalist experiments
  - Classic detectors composed mainly by 3 sub-systems
    - Inner tracker
    - Calorimeter system
    - Muon spectrometer
  - Good e/γ/μ/τ/b-jets identification



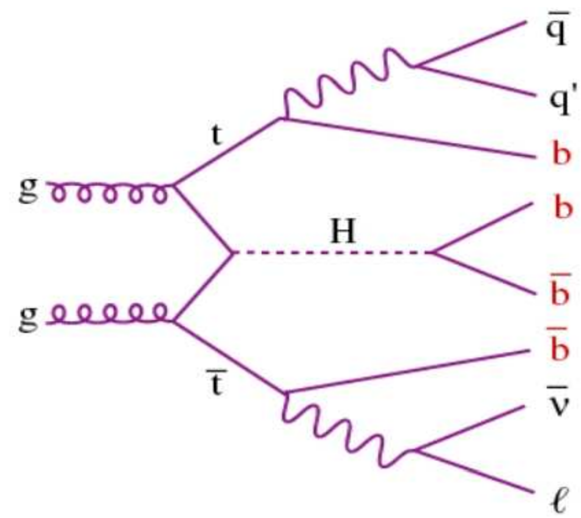
CMS detector

ttH Analysis Outline

- ttH channels phenomenology
- bb decay (detailed analysis)
  - ATLAS: lepton + jets final state
    - 3 different approaches
  - CMS: All final states
    - All hadronic, lepton+jets and di-leptonic
  - Discussion of main issues for these analysis: b-tagging, Jet Energy Scale, combinatorial background and physical background extraction from data
- WW (ATLAS) and $\gamma\gamma$ (CMS) decay (briefly reported)

ttH Channels Phenomenology

- ttH(→bb): potential discovery channel for light Higgs boson
  - All hadronic final state
    - Higher branching ratio
    - 8 jets (4 b) in the final state
  - Lepton + jet final state
    - 1 lepton, 6 jets (4 b) and one neutrino in the final state
  - Di-leptonic final state
    - Low branching ratio, high trigger efficiency
    - 2 leptons, 4 b jets and 2 neutrinos in the final state

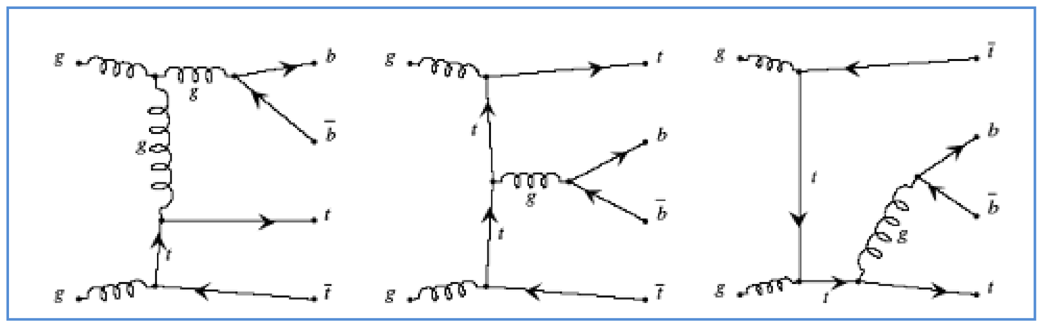


- Main backgrounds
  - tt+jets : reducible background using b-tagging
  - ttbb (EW/QCD) : irreducible background but slightly different kinematics than signal

ttH (→WW) and ttH (→γγ) decay modes are interesting for respectively higher Higgs mass and high integrated luminosity

ttH channels give a unique access to top Yukawa coupling

Some ttbb(QCD) background Feynman's diagrams



$t\bar{t}(H \to b\bar{b})$ Analysis

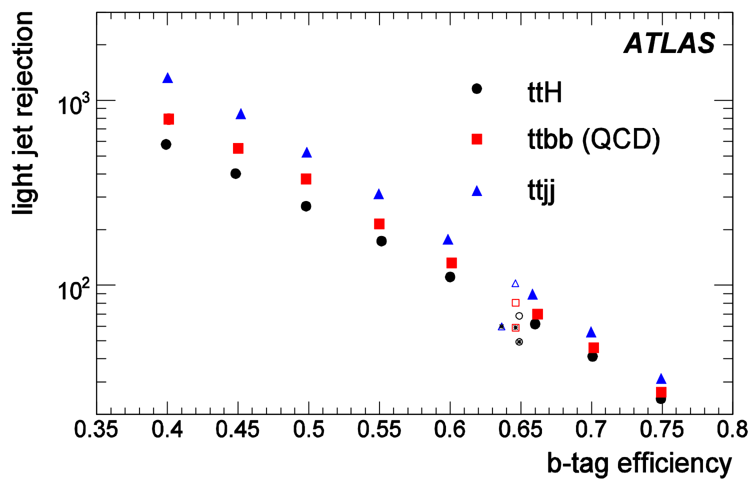
- Strategy: Reconstruct the tt system to look at the rest
- 4 b-jets at the final state
  - b-tagging is one of the most important keys for this channel
- Jet pairing and JES are also very important
- ATLAS:
  - Efforts mainly for lepton + jets channel
  - All hadronic channel analysis started, only lepton + jets channel will be reported
  - ttH samples produced with Pythia, ttbb with AcerMC+Pythia and tt+jets using MC@NLO+Herwig with an additional filter on jets
  - Higgs boson mass produced at 120 GeV/c^2
  - All sample produced with full simulation of ATLAS detector
  - All studies for 30 fb^{-1}
- CMS:
  - Combined analysis for dileptonic, lepton + jets and all-hadronic channel
  - Focus and optimization on lepton + jets channel
  - ttH & ttbb samples produced with CompHEP + Pythia, tt+jets with Alpgen + Pythia, QCD with Pythia (120-170 GeV/c, >170 GeV/c)
  - ttH with Higgs masses 115, 120, 130 GeV/c^2
  - Full simulation of CMS detector
  - All studies for 60 fb^{-1}

Pre-selection (ATLAS)

- Isolated electron or muon trigger selection applied before the pre-selection
- One isolated lepton is required
  - Electron, muon: Acceptance, identification and isolation cuts
  - Tau is not considered in this analysis
- At least 6 jets: cone 0.4 jet algorithm, $p_T > 20$ GeV/c
- At least 4 jets tagged as b-jets using life time based taggers:
  - b-tag weight combines impact parameter based tagger with a secondary vertex based tagger
    - Tight: b-tagging efficiency = 65%, light jet rejection = 60
    - Loose: b-tagging efficiency = 85%, light jet rejection = 8.6
- Scaling of the b-tagging weight to simulate 30% worsening of light jet rejection
  - More realistic description of the b-tagging performance (impact of residual misalignments coming from actual alignment exercise)

LO cross section (only tt+jets NLO+NLL)

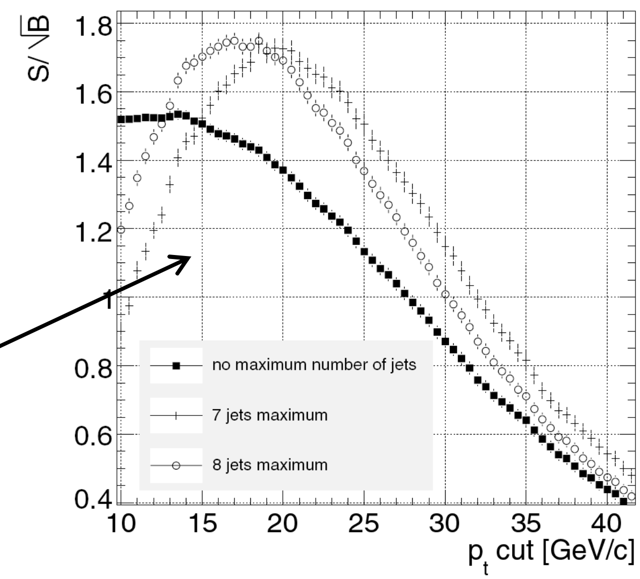
cut	ttH	ttbb(EW)	ttbb(QCD)	tt+jets
1 Lepton (fb)	56.9	141	1356	63710
+ 6 jets (fb)	36.2	76.7	665	26214
+ 4 b-jets loose (fb)	16.2	23.4	198	2589
+ 4 b-jets tight (fb)	3.76	4.2	29.6	50.7

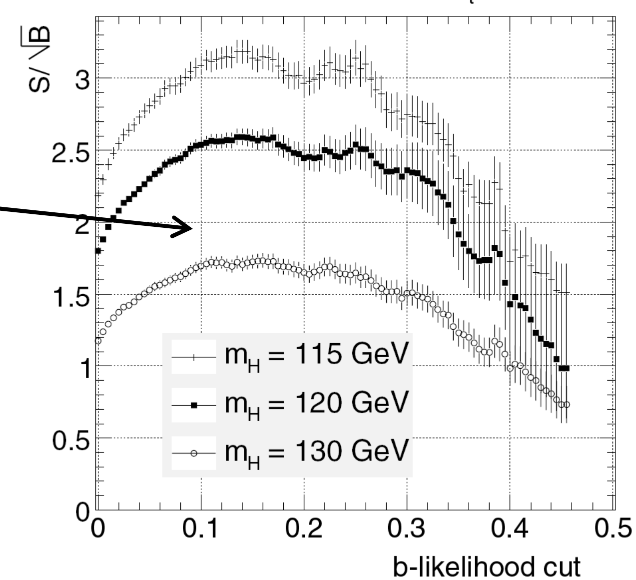


Pre-selection (CMS)

Lepton+jets channel

- Trigger: isolated muon or isolated electron
- One isolated lepton (electron or muon)
  - likelihood including p_T, isolation and identification variables
- 6 or 7 jets: cone 0.5 algorithm
  - $p_T > 20$ GeV/c, $|\eta| < 3$
  - Optimized for maximum significance
- b-jet identification
  - Combined secondary vertex algorithm
  - Likelihood ratio method using 4 jets with highest discriminator output
  - Optimized for “loose” (maximum significance) and “tight” (maximum purity) working points
- Signal efficiency (Higgs mass 120 GeV/c^2)
  - 1.55% (loose), 0.52% (tight)





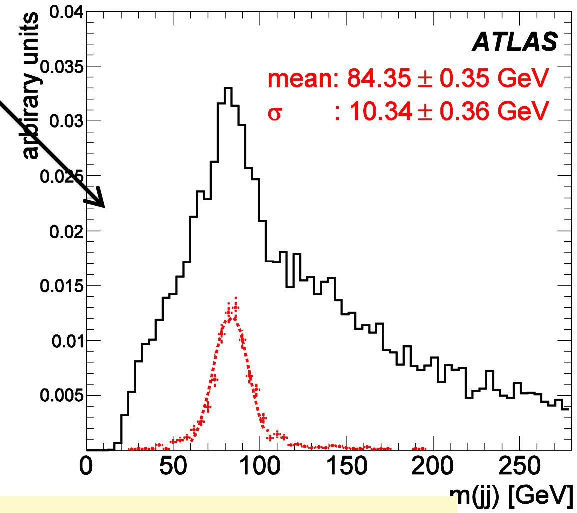
Cut-Based Reconstruction(ATLAS)

- Leptonic W reconstruction
  - Force lν mass to the W mass
  - Solve 2nd degree equation to get neutrino p_z
  - 28% no solution → neglect imaginary part
- Hadronic W reconstruction
  - After requesting 4 b-tag jets, remaining jets are considered as light jets
  - W boson candidates are formed of all combinations of light jet pairs
- top quarks candidates are formed using one W boson candidate and one b quark
  - Combination with $|m(jj) - m(W)| > 25$ GeV/c^2 or $|m(t_{reco}) - m(t_{true})| > 25$ GeV/c^2 are removed
- Combination minimizing a χ^2, based on the top quark masses, is chosen

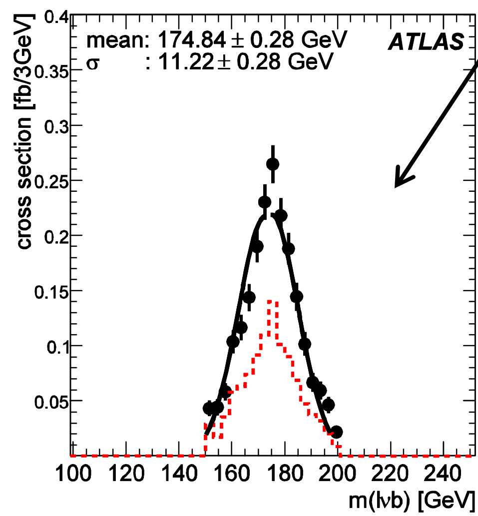
$$\chi^2 = \left(\frac{m_{\ell\nu b} - m_{top}}{19GeV}\right)^2 + \left(\frac{m_{jjb} - m_{top}}{13GeV}\right)^2$$

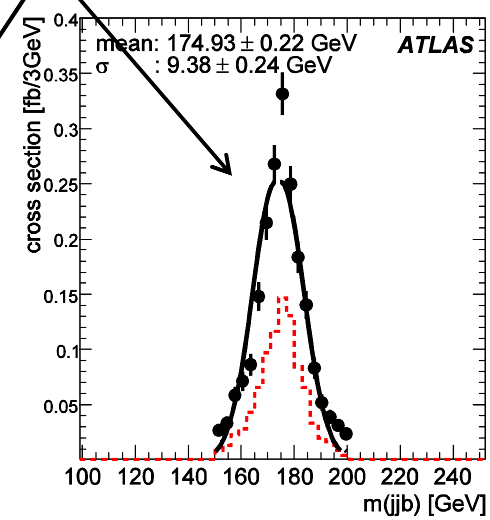
- The two remaining b jets used to reconstruct the Higgs candidates

Invariant mass of reconstructed hadronic W boson
(red for good combination)



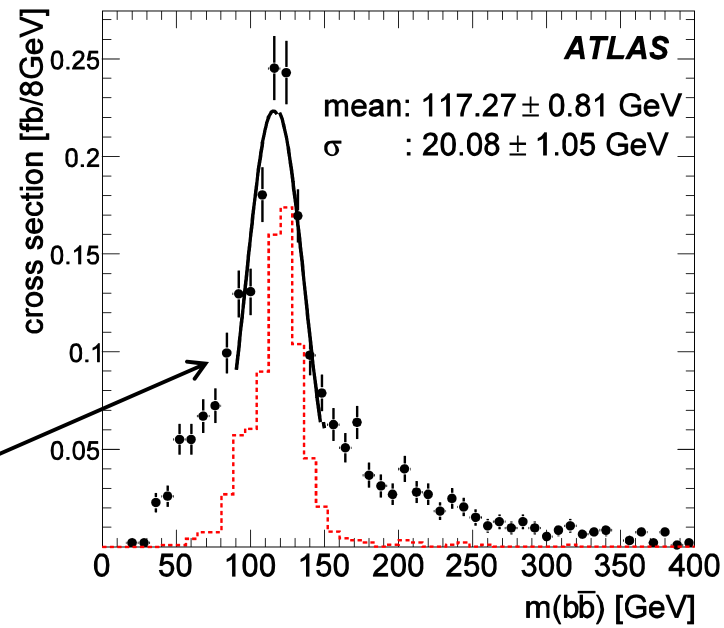
Invariant mass of reconstructed top quarks
(red for good combination)





Multivariate Based Reconstruction (ATLAS)

- Pairing likelihood:
  - Using tt system kinematic properties to build a pairing likelihood
    - 6 variables are used
    - b-jets and light jets are treated separately
    - Choose the combination that maximize the likelihood output
    - Cut: Likelihood > 0.9



Invariant mass of reconstructed Higgs boson using likelihood for signal sample
(red for good combination)

- Constrained fit:
  - Fit jet p_T and E_{Tmiss} to give the mass of the top quarks
  - Pairing likelihood is formed using the χ^2 output of the constrained fit together with b-tagging and kinematic quantities
    - 14 variables are used
    - 3D likelihood is used to take into account the correlations
  - Final selection likelihood is used to separate signal and physics background

$$\chi^2 = \sum_{i=1}^{6}\left(\frac{f^i_{jet}-1}{\sigma^i_{jet}/p^{i,initial}_{jet}}\right)^2 + \frac{(m_W^{lep}-80.425)^2}{(2.1)^2} + \frac{(m_t^{lep}-175)^2}{(1.5)^2}$$

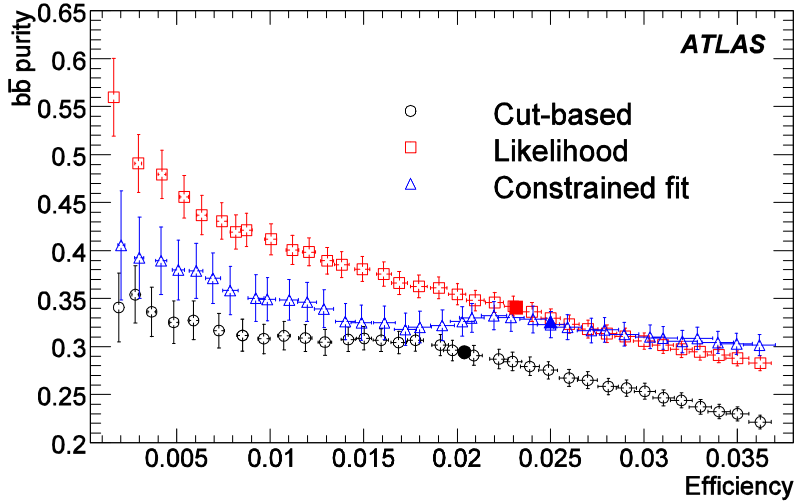
f^i_{jet} : scale factor for jet momentum

$$\sigma_{P_{light}}/P_{light} = 0.988/\sqrt{p_T} \oplus 0.035$$

$$\sigma_{P_b}/P_b = 0.888/\sqrt{p_T} \oplus 0.125$$

Comparison of Reconstruction Algorithms (ATLAS)

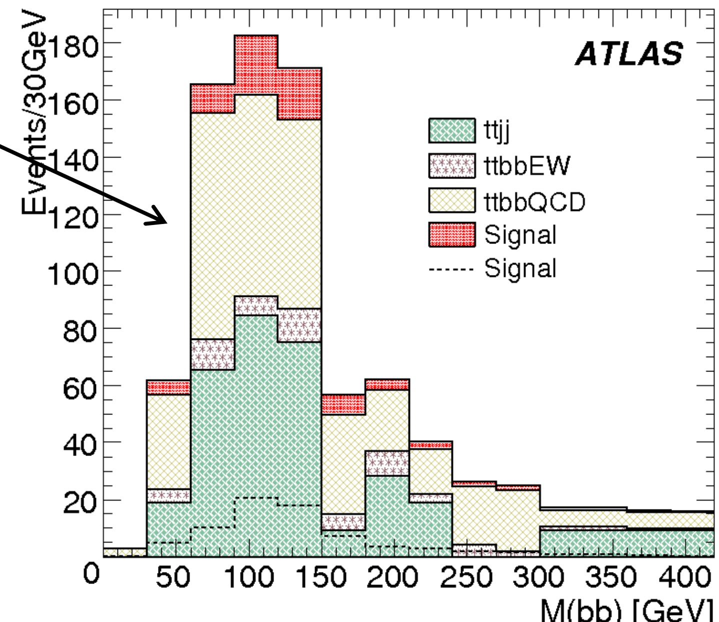
- S/√B differences relatively small among the 3 analysis
- Multivariate techniques increase Higgs bb pair purity by ~5%
- All analyses suffer from low purity mainly coming from b exchange between top and Higgs
  - Wide distribution for reconstructed Higgs mass
  - No clear signal peak on top of background distribution
  - Combinatorial background dilutes differences between ttbb and ttH



Invariant mass of reconstructed Higgs boson using constrained fit for all samples

30 fb^{-1}, LO cross section (only tt NLO+NLL)

	Cut-based	Likelihood	Constrained fit
Signal efficiency (%)	2.04	2.32	2.49
bb purity (%)	29.4	34.0	32.0
bb mass peak resolution (GeV)	22.8	20.1	22.3
S/B (*)	0.110	0.103	0.123
S/√B (*)	1.82	1.95	2.18



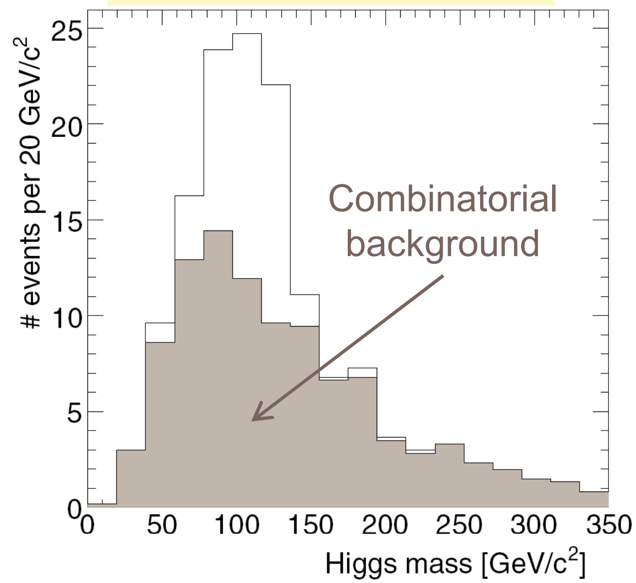
* $90 < m_H < 150$ GeV/c^2

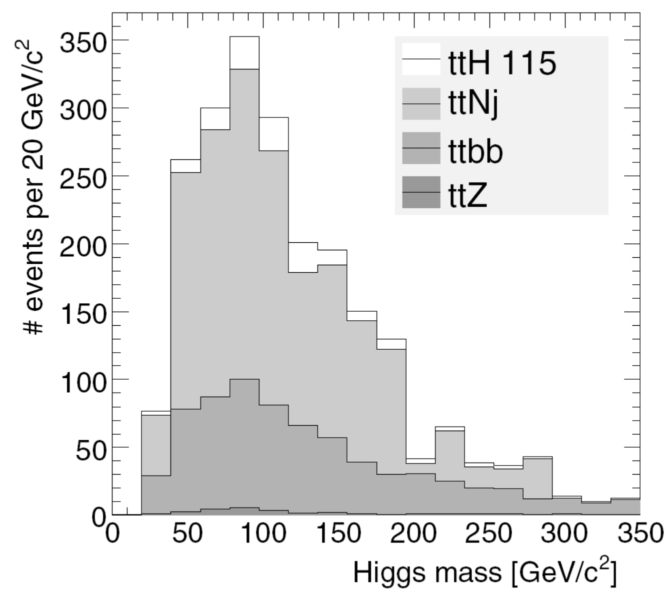
Reconstruction (CMS)

Lepton+jets channel

- Reconstruction of leptonic W:
  - Solve for neutrino p_z with W mass constraint
  - Non-real solutions: $p_z^{\nu} = p_z^{\ell}$
- Several methods for jet-parton assignment (pairing):
  - Kinematic fits, likelihood methods, mass resonances, angular distributions
  - Comparable efficiencies ~30%
- Higgs mass peak influenced by
  - Detector resolution effects
  - Inefficiencies in jet-parton assignment
- Loose/Tight (m_H = 120 GeV/c^2, 60 fb^{-1}):
  - Signal efficiency: 1.55% / 0.52%
  - Significance : 2.5 / 1.9

115GeV/c² Higgs mass

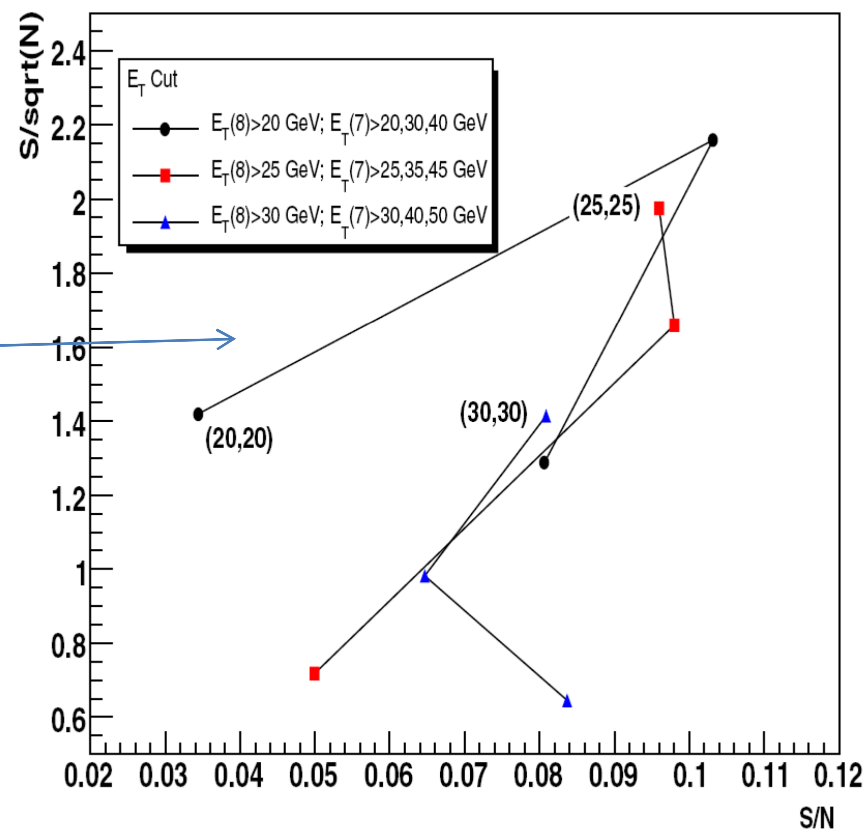




All Hadronic Channel (CMS)

- Challenge:
  - Combinatorics in events with at least 8 jets
- Jet pairing with χ^2 method
  - Use invariant masses of top quark and W
- Cut optimization for loose/tight working points:
  - Pt and |eta| of all jets
  - b-tagging discriminator
  - Centrality (whole event)
  - Higgs centrality
- Loose/Tight (m_H 120 GeV/c^2 , 60 fb^{-1}):
  - Signal efficiency : 2.55% / 0.37%
  - Significance : 2.4 / 2.0

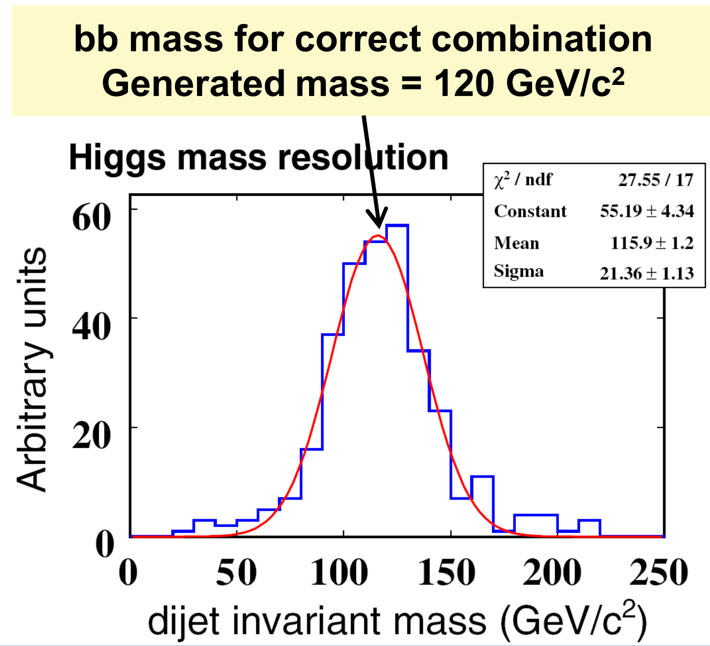
Significance vs Purity for different jet pt cuts

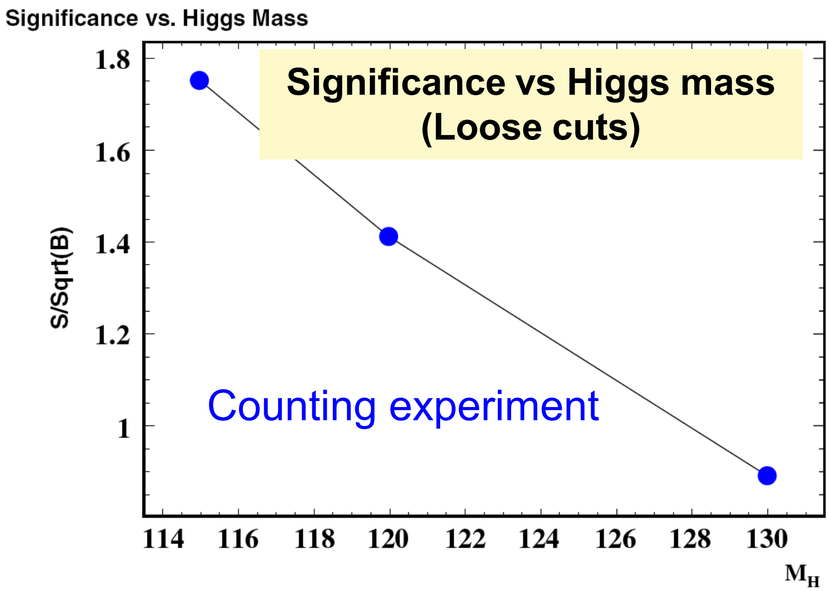


Di-leptonic Channel (CMS)

Event Selection:

- Loose
  - Two oppositely charged leptons passing likelihood cuts
  - MEt > 40 GeV
  - 4-7 jets with pt > 20 GeV/c, |eta| < 2.5
  - 3-4 jets b-tagged
- Tight (additional cuts)
  - 4-6 jets, 4 jets b-tagged





60 fb^{-1}, m_H = 120 GeV/c^2, NLO cross section for signal, LO for background(*)

	S/√B (loose)	S/√B (tight)
Lepton+jets	2.5	1.9
All hadronic	2.4	2
Di-leptonic	1.4	0.9
Combined	3.89	3.29

* tt+jets: LO cross section but NLO kinematics effect included via the Alpgen matching procedure and additional hard radiation from Pythia

Systematic Errors

- For a counting experiment the background uncertainties need to be small
  - Side band analysis is difficult with the current combinatorial background and Higgs mass resolution
- Systematic errors need to be well controlled
- Main systematic errors come from the JES, jet resolution and b-tagging efficiency

CMS main systematic errors, lepton+jet channel

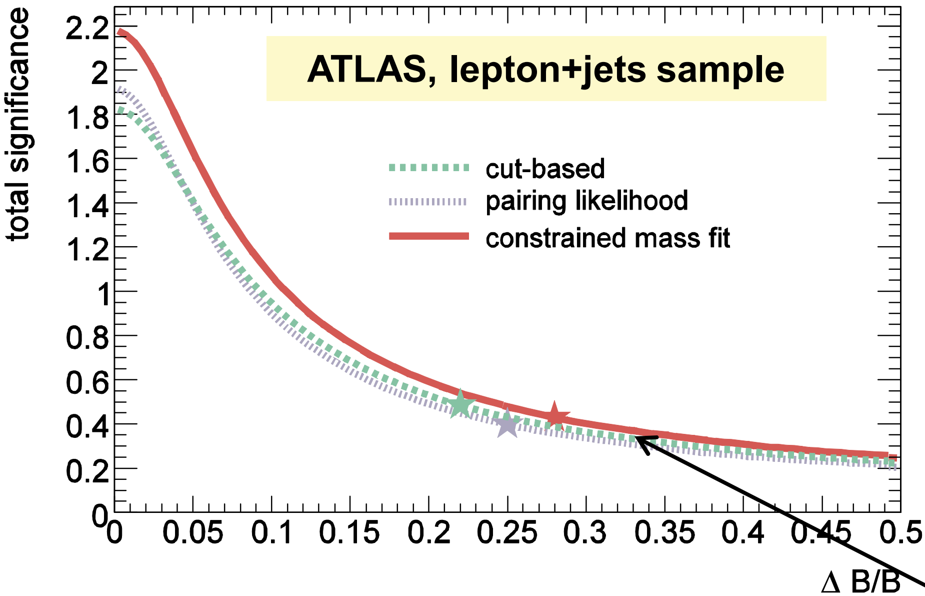
	Signal	Background
JES	3-5%	6-30%
Jet resolution	1-4%	2-8%
bc-tagging	20-25%	20-40%
uds-tag	< 2.7%	< 2.7%
All contribution	21-25%	26-40%

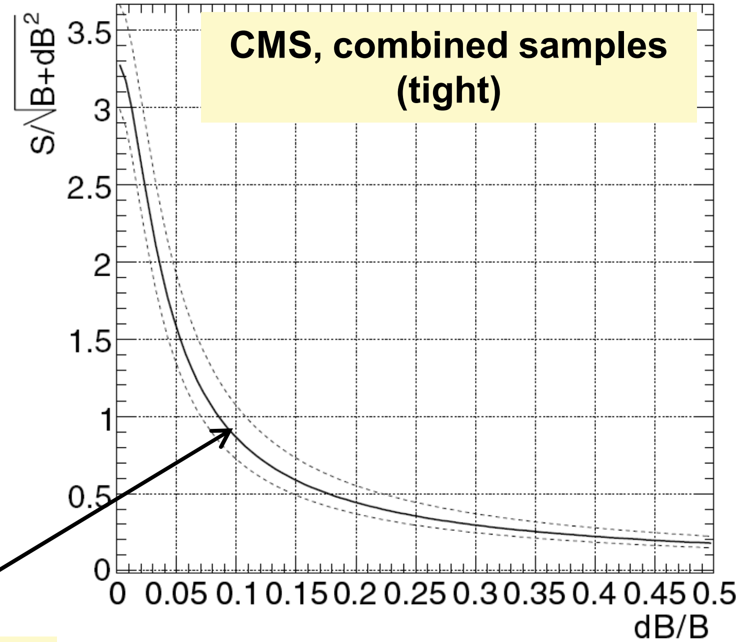
ATLAS main systematic errors, background sample

	Cut-based	Likelihood	Constrained fit
JES	5%	14%	8%
Jet resolution	7%	5.5%	14%
b-tagging efficiency	20%	20%	20%
Light jet rejection	5%	3%	10%
All contribution	22%	25%	28%

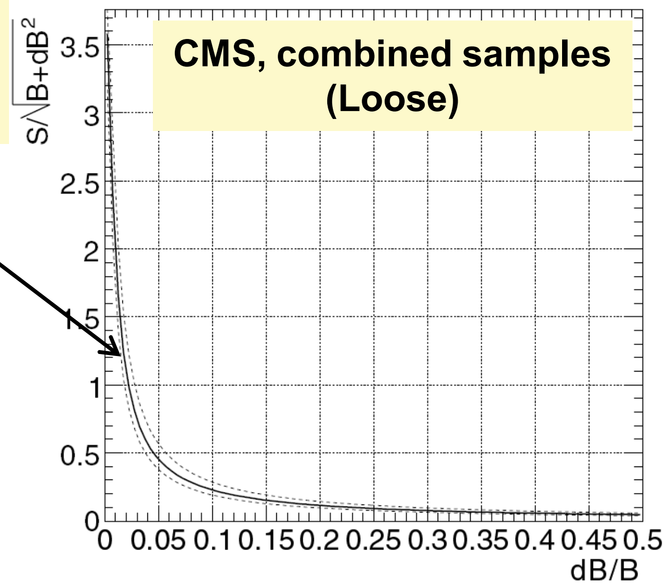
- Theoretical errors on background cross sections, especially top anti-top production, are sizable.
  - These cross sections need to be well measured in data (~5% error on top anti-top)

Significance + Systematic Errors





Significance vs background errors



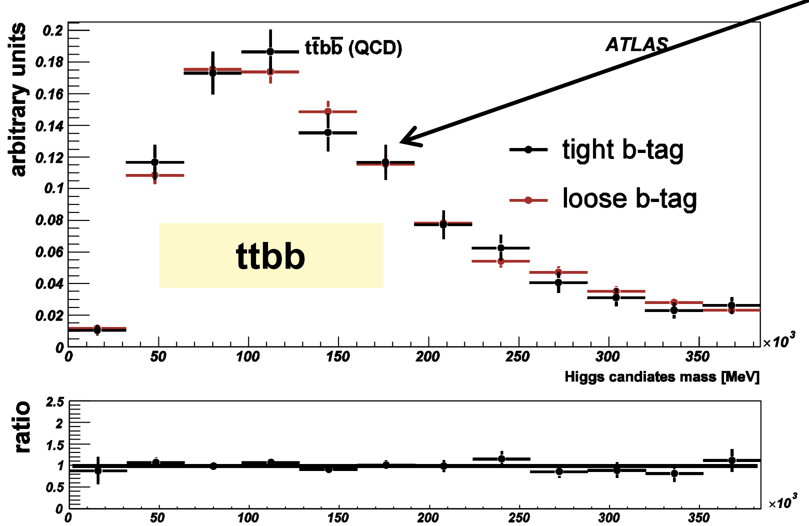
- For CMS the significance drops very fast with loose cuts
- Systematic errors on background higher than signal
- If nothing is done for the background extraction the significance will decrease to:
  - ATLAS: ~0.5 (m_H=120GeV/c^2, 30fb^{-1})
  - CMS combined: 0.13 (loose), 0.48 (tight) (m_H=120GeV/c^2, 60fb^{-1})

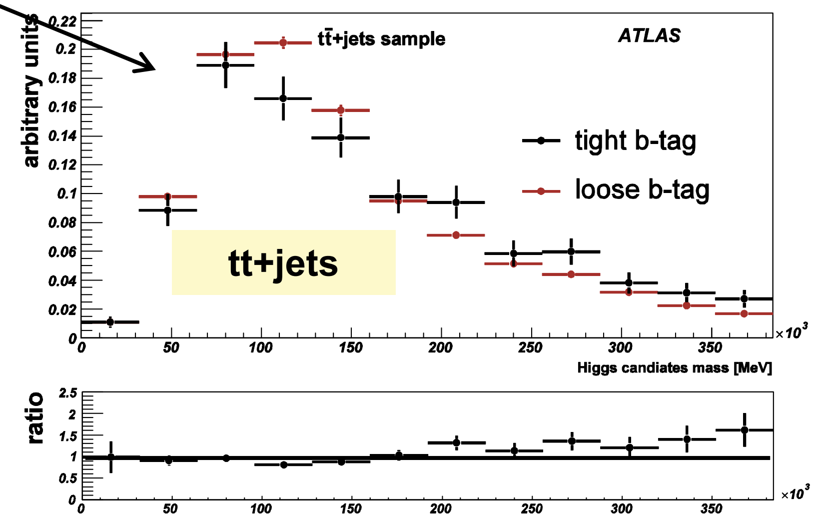
Main Issues for ttH(→bb) Channel

- After analysis Higgs purity is ~30%
  - large tails and width in the bb invariant mass
  - No visible peak on top of physical background, side band extraction very difficult
- The main problem is the b-jets exchange between top quarks and the Higgs
  - b-jet from Higgs used for the hadronic top : ~36%
  - b-jet from Higgs used for the leptonic top : ~30%
  - Higgs boson reconstructed with only one correct b-jet: ~55% (The other wrong jet is mainly coming from the top quarks)

- Big uncertainty for the tt+jets cross section
  - Extracting the background shape and normalization in data is crucial for this channel
  - Using loose and tight b-tagging cuts looks promising (ATLAS)
    - Background shape independent from b-tagging cuts

b quark pair invariant mass for loose and tight b-tagging cut

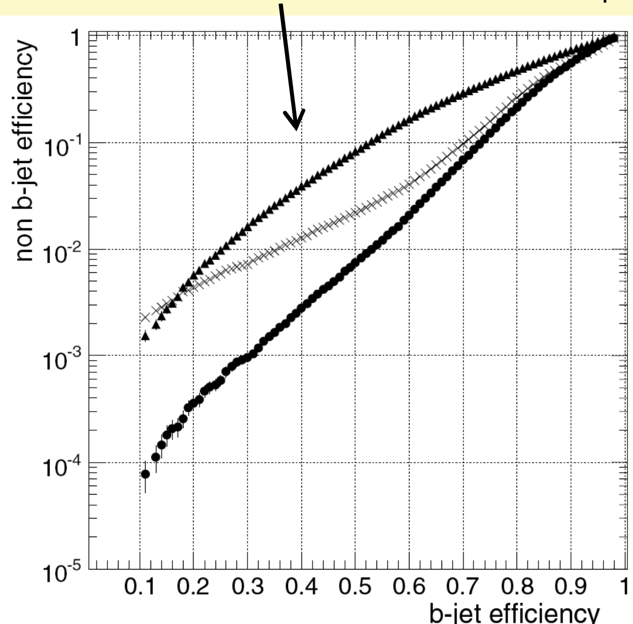




Main Issues for ttH(→bb) Channel

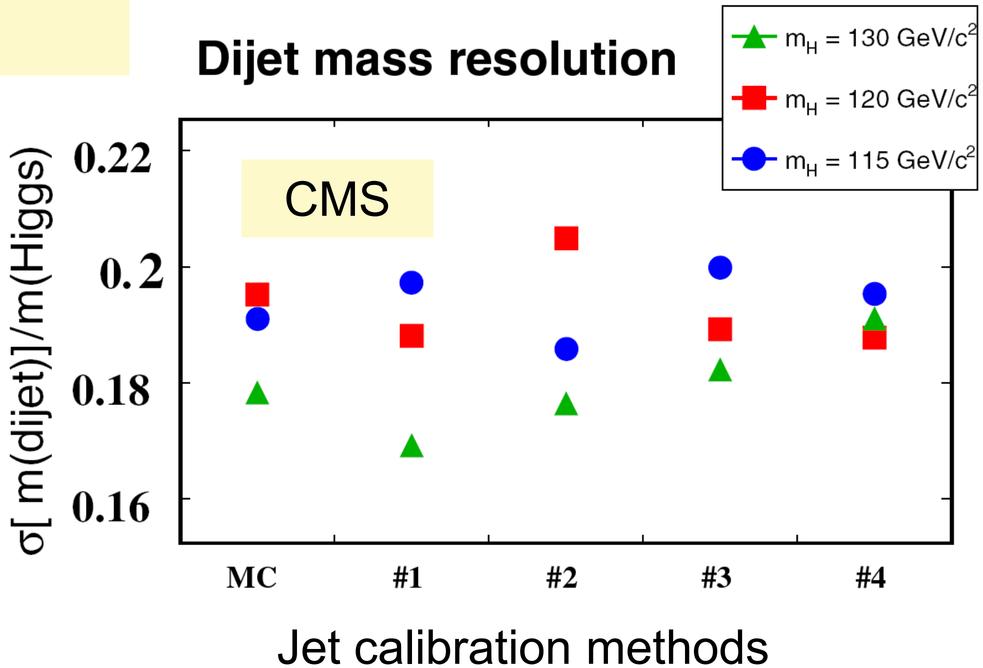
- b-tagging is crucial due to 4 b-jets in final state
- CMS, two improvements have been developed for this analysis:
  - inclusion of tertiary vertices (further charm decay vertex)
  - inclusion of soft lepton tagging
- ATLAS, latest and more powerful tagger based on reconstructing the B→D decay chain not used in this analysis

b-jet efficiency vs mis-tagging rate (CMS):
- Triangles are c-jets, crosses gluon jets, points light flavored jets
- This plot is for the ttjj sample.
- The results are about the same for the ttH sample



- bb invariant mass resolution is important to extract the signal from background
- ATLAS, resolution around 13% (*)
- CMS, resolution around 19% (*)
  - The resolution is comparable for different calibration method
  - Energy flow methods for jet calibration can improve the Higgs mass resolution
  - This is not included in the plot below

* Resolution is computed for true bb combinations



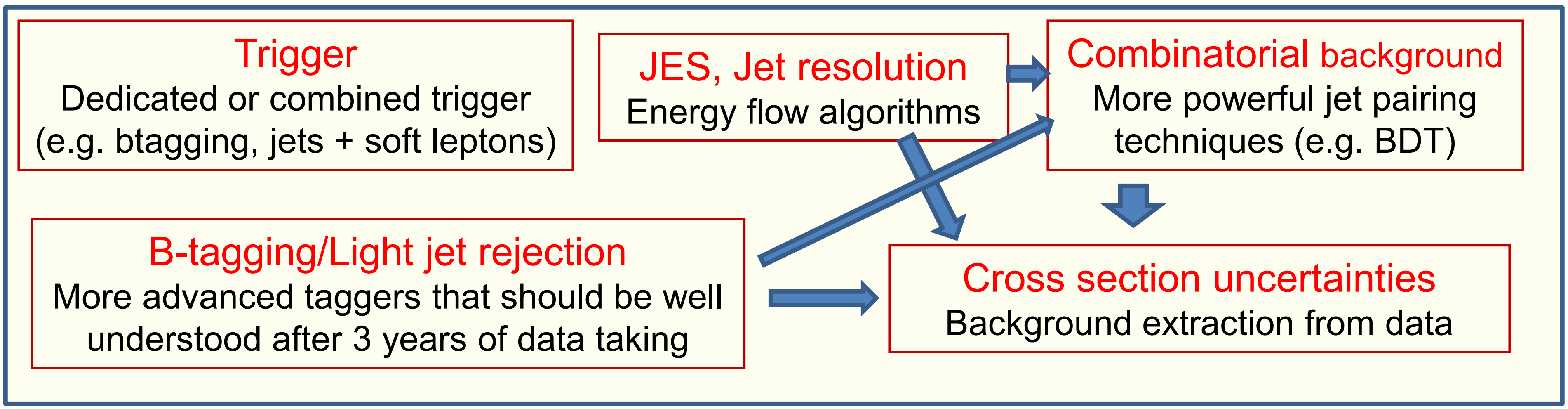
Summary of ttH(→bb) analyses

Summary table	Significance loose/tight	Luminosity
ATLAS (Lepton+jets)	2.2	30 fb^{-1}
CMS (Lepton+jets)	2.5/1.9	60 fb^{-1}
CMS(Combined)	3.9/3.3	60 fb^{-1}

Significance < 1 for both experiment if:

- Simple counting experiment
- No special treatment to extract background from data
- Systematic errors included

There is still room for improvement (⇨ ≡ Helps)



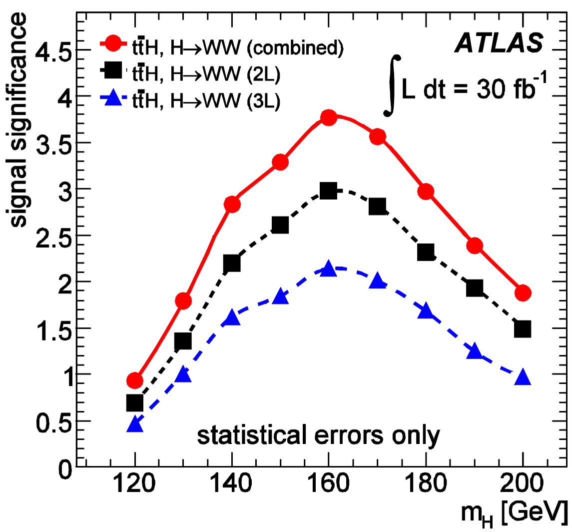
Input from theory side needed, e.g. NLO calculation for ttbb

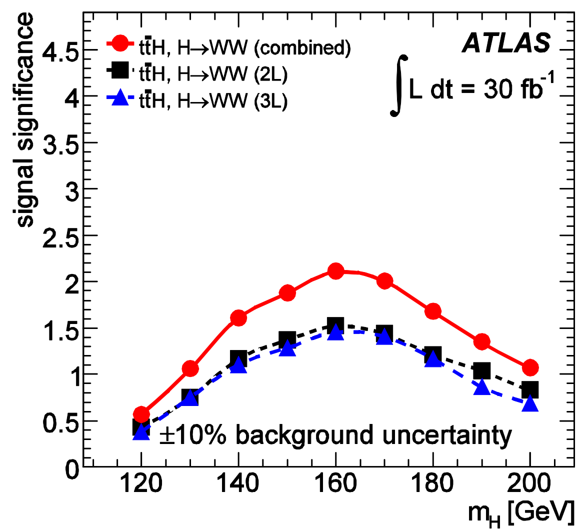
WW(*) and $\gamma\gamma$ decay modes

$WW^{(*)}$ Decay Mode (ATLAS)

- Both 2 sign-like lepton (2L) and 3 lepton (3L) decay configuration were considered
- Main backgrounds: tt, ttW, ttZ, tttt and ttbb (QCD is rather negligible)
- Event selection:
  - Trigger: isolated high p_T electron or muon
  - 2 sign-like/3 isolated electron or muon
    - Isolation combine track and calorimeter isolation
  - 4 jets (cone 0.4 algorithm)
  - Z-veto ($75 < m_{\ell\ell} < 100$ GeV/c^2)
  - For the 2L analysis the Higgs boson can be reconstructed using a constrained fit
    - 2 b-tagged jets are required to reduce the number of combinations

- Significance can reach 3.8 for the combined channels at 30 fb^{-1} total luminosity
- The systematic errors were estimated around 10% mainly coming form the JES
- The significance decreases to 2.1 after adding the systematic errors



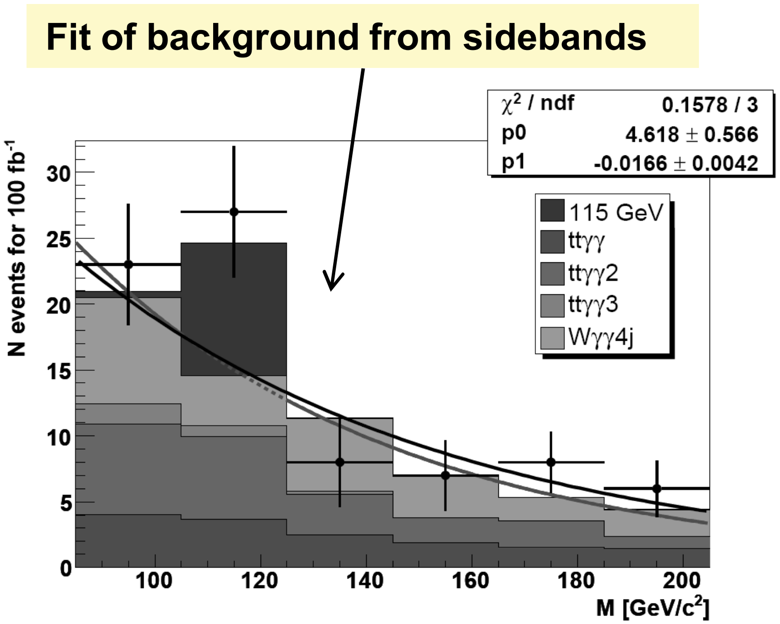


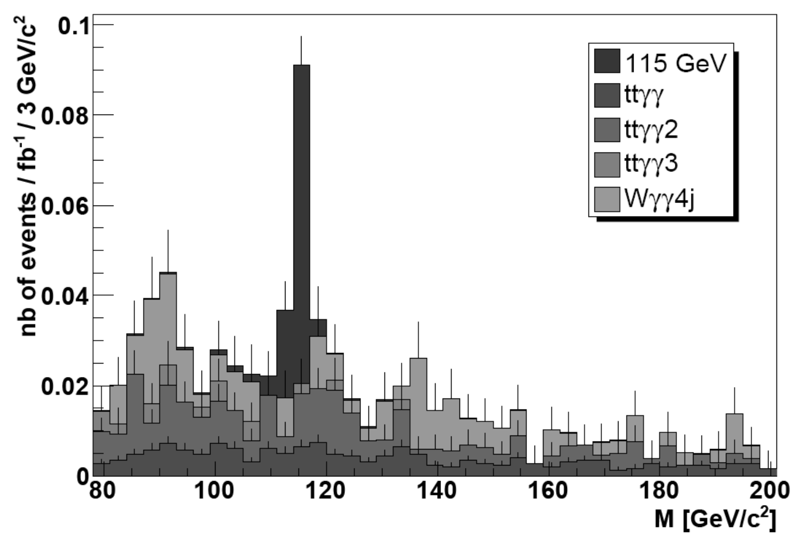
γγ Decay Mode (CMS)

Important channel for Higgs mass and Yukawa coupling measurements

- Analysis performed for 100 fb^{-1}
- Lepton+jets decay channel
- Cut-based analysis
  - Properties of photons
  - Isolated lepton
  - At least 4 jets with p_T of 60 GeV/c
  - One b-tagged jet
- S/B ratio of 4
- 11.2% selection efficiency (Higgs mass of 120 GeV/c^2)
- Significance above 3 including systematics, using sideband fit

Signal observable (3 sigma confidence) at 100 fb^{-1}





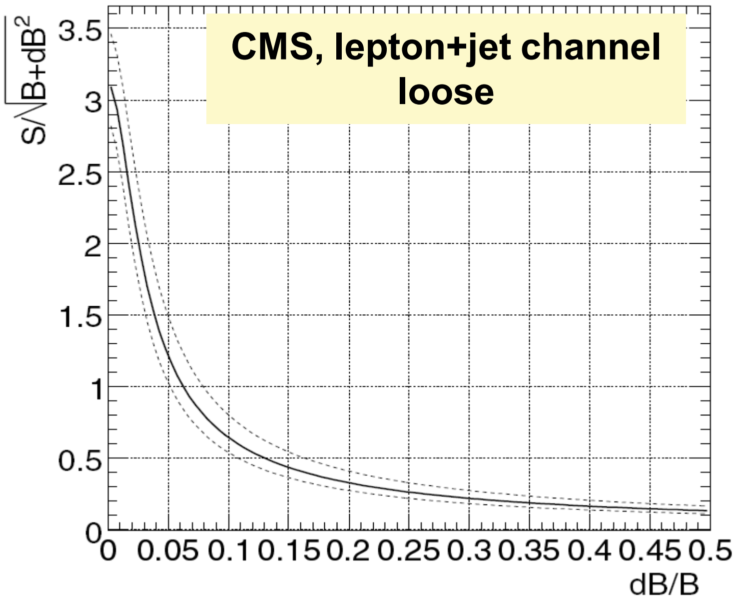
Conclusion

- ttH channel is an important channel for testing the electroweak sector of the Standard Model
- Several Higgs decay modes are considered:
  - bb decay as a discovery channel for a low mass Higgs boson
  - WW decay for medium Higgs mass around a resonant W boson pair
  - $\gamma\gamma$ decay for high luminosity
- A detailed analysis with full simulation of ATLAS and CMS detector yields a comparable conclusion for the ttH ($H \rightarrow bb$) channel
  - Without special treatment to extract background in data and improvements in b-tagging and jet pairing the signal is not observable
  - Promising work starting on this
  - Once the background shapes are understood, but maybe not the cross section/k-factor, one can stop using a counting experiment, and go for more advanced methods
- Analysis result should be update with the latest improvements, specially for b-tagging where tt($H \rightarrow bb$) is very sensible, for both experiments
- WW and $\gamma\gamma$ decay modes are also important channels for Yukawa coupling measurement

Back-up

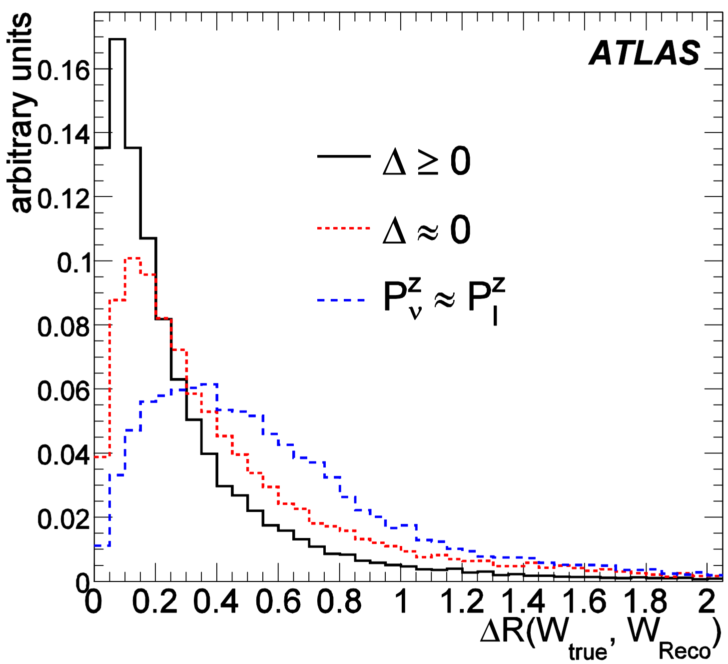
Systematic errors

- Systematic errors

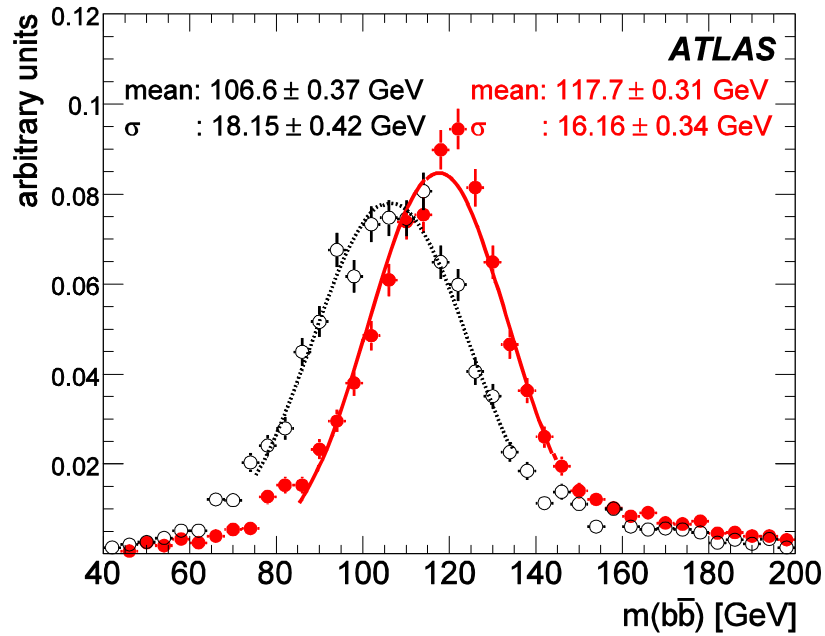


	JES	Jet resolution	Btag efficiency	Light jet rejection
ATLAS	|eta| < 3.2, 7% |eta| > 3.2, 15%	|eta| < 3.2, $\sigma=0.45*\sqrt{E}$ |eta| > 3.2, $\sigma=0.63*\sqrt{E}$	5%	10%
CMS	From 10% at 20 GeV to 3% for 50 GeV	10%	4%	10%

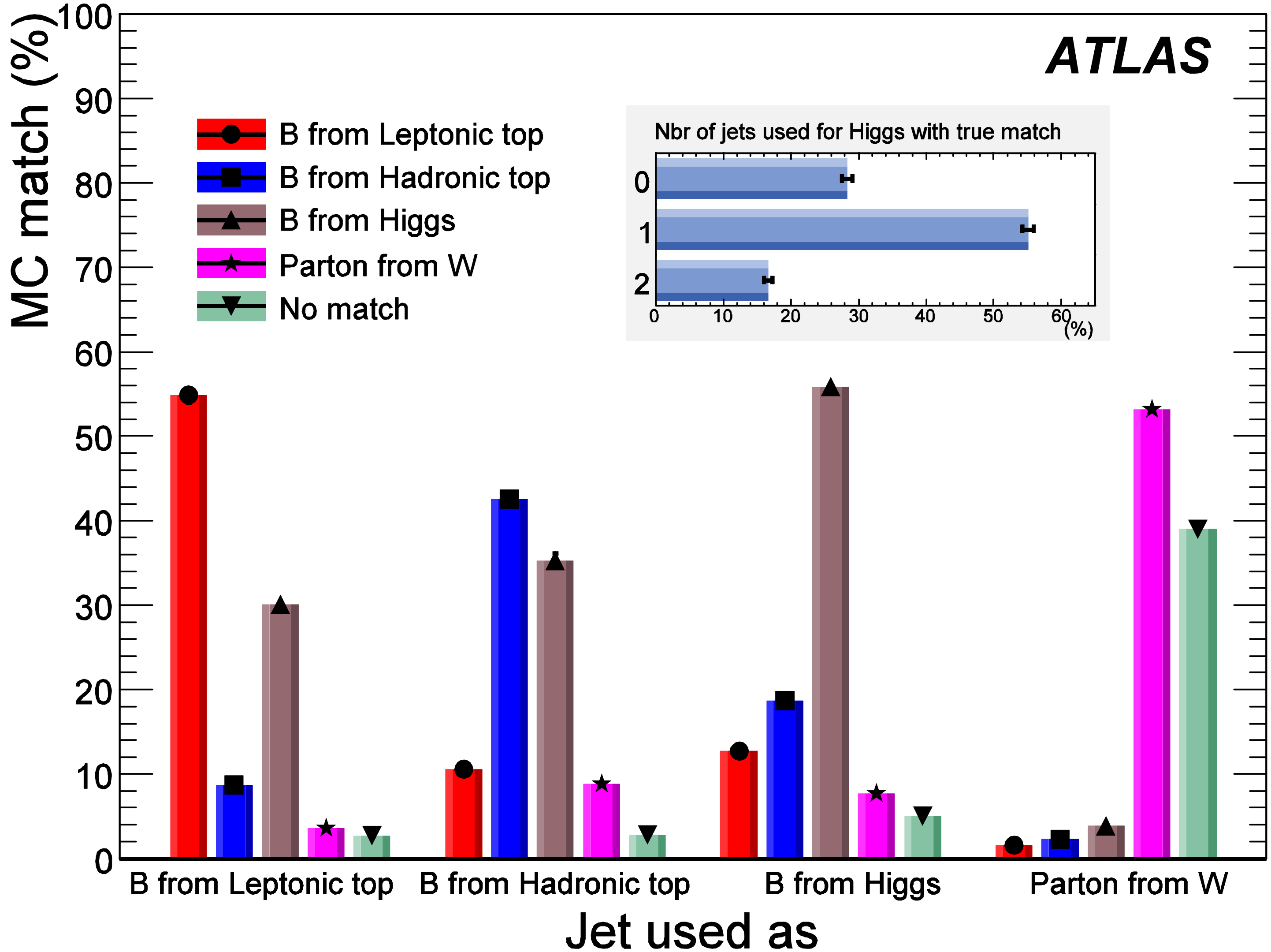
Spatial distance between leptonic W_{reco} and W_{true} (black when there is a p_z solution)



Effect of soft muons, only jets containing a muon (red after correction)



Jet pairing, No likelihood cut



Pairing likelihood templates