Background studies in Higgs Searches

Bruce Mellado
University of Wisconsin-Madison

Theory Experiment Interplay at the LHC
Royal Holloway, London, 08/05/10

Outline

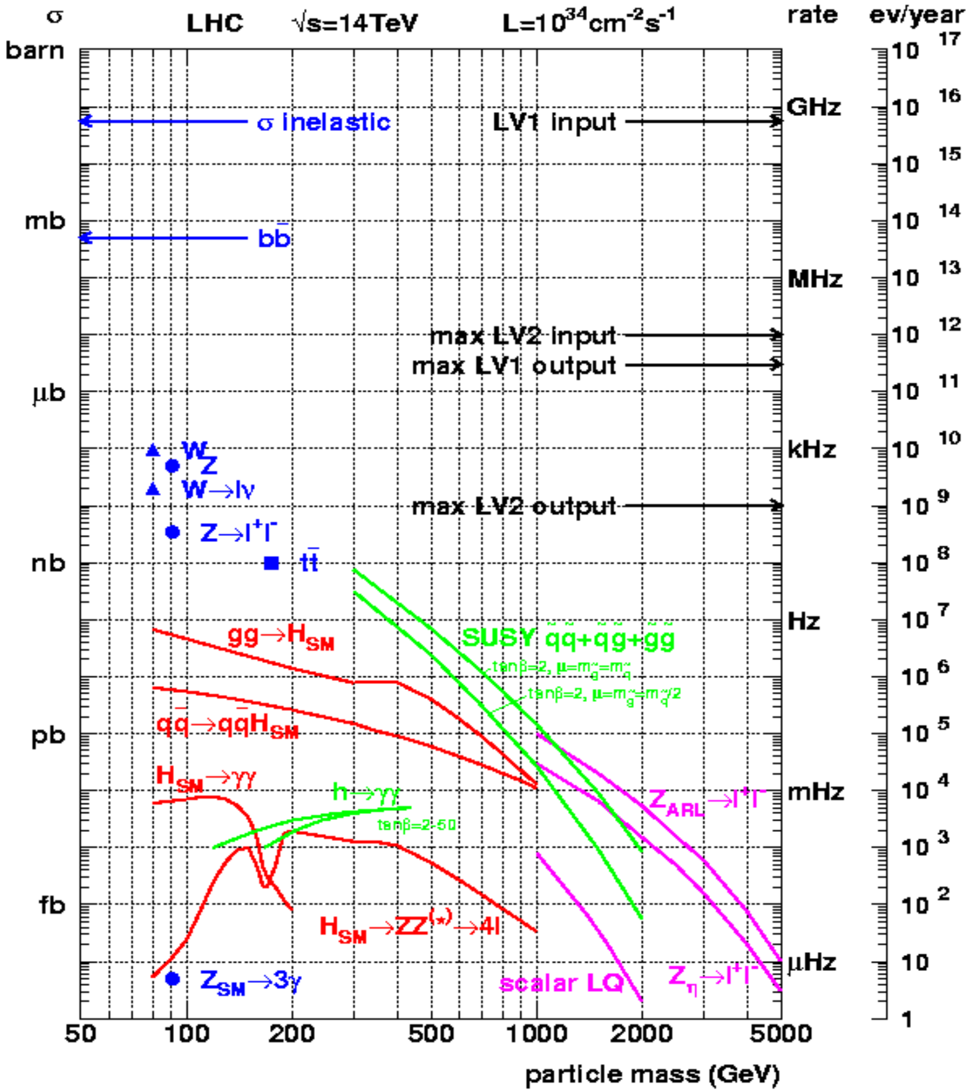
- Introduction
- Data Driven methods in prospective channels with the first data
  - $H \to \gamma\gamma$
  - $ttH \to bb$
  - $H \to \tau\tau$
  - $H(\to ZZ^{(*)} \to ll),(WW^{(*)} \to ll\nu\nu)$
  - MSSM

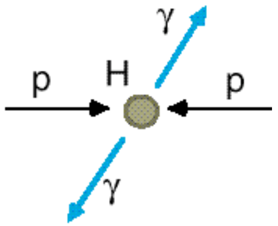
"Expected Performance of the ATLAS Experiment"
http://arxiv.org/pdf/0901.0512v1

See W. Murray's talk for a global view of discovery potential
Special thank to T.Vickey for help with slides

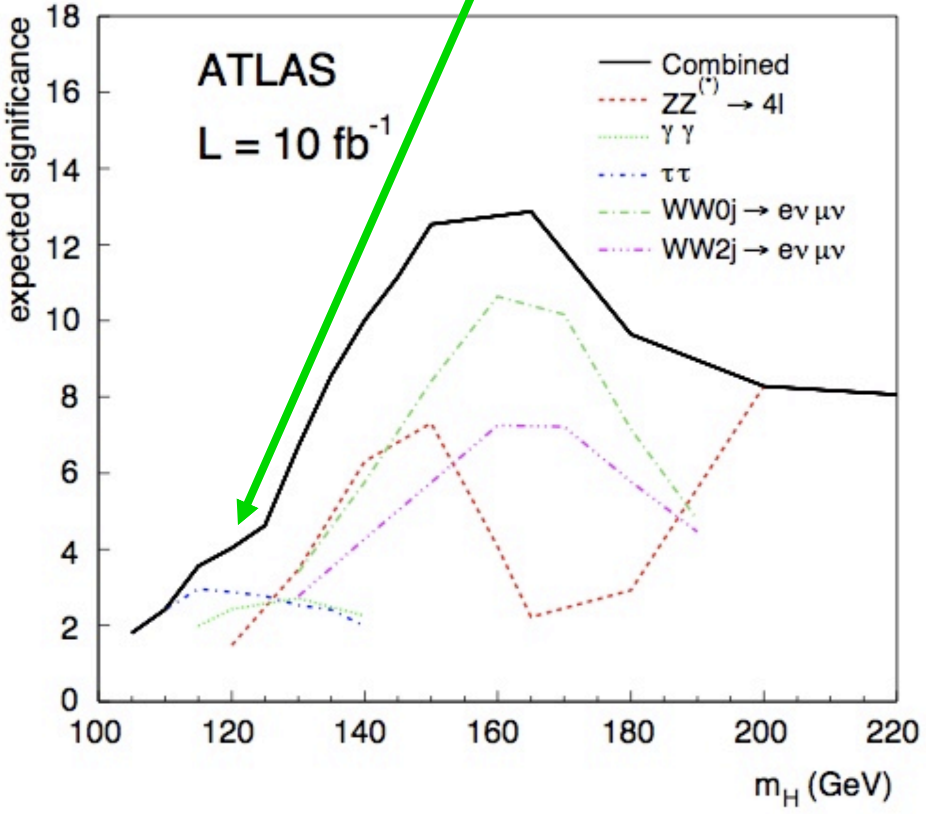
Cross-sections at LHC

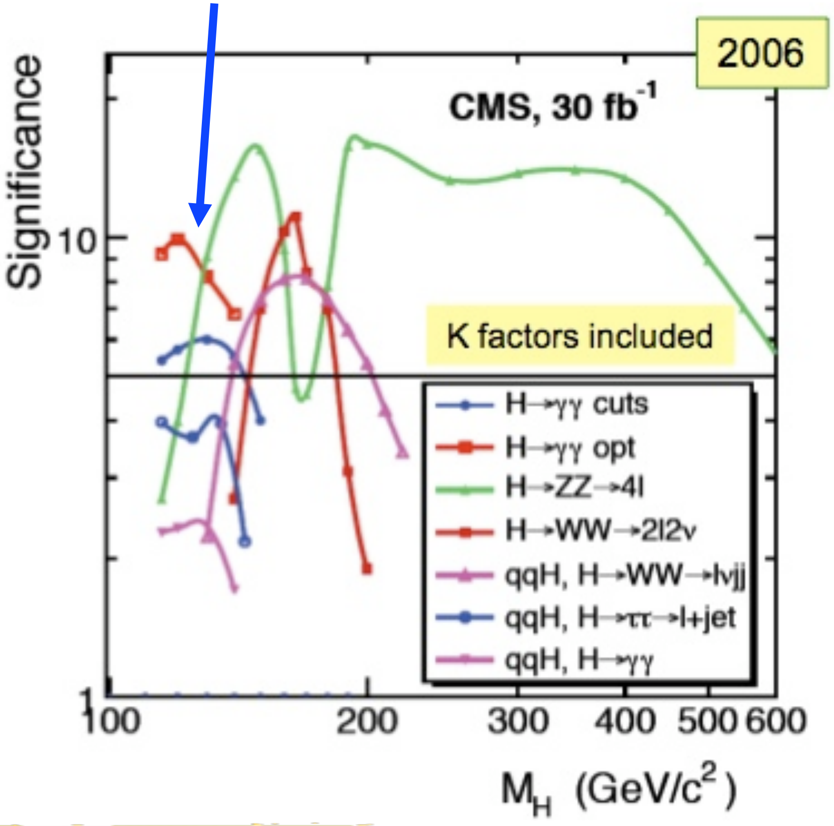
- **Search for Higgs and new physics hindered by huge background rates**
  - Known SM particles produced much more copiously
- **This makes low mass Higgs especially <u>challenging</u>**
  - Narrow resonances
  - Complex signatures
  - <u>Data Driven Methods play a very important role when subtracting backgrounds</u>
    - ❖ <u>Theoretical inputs are always needed</u>



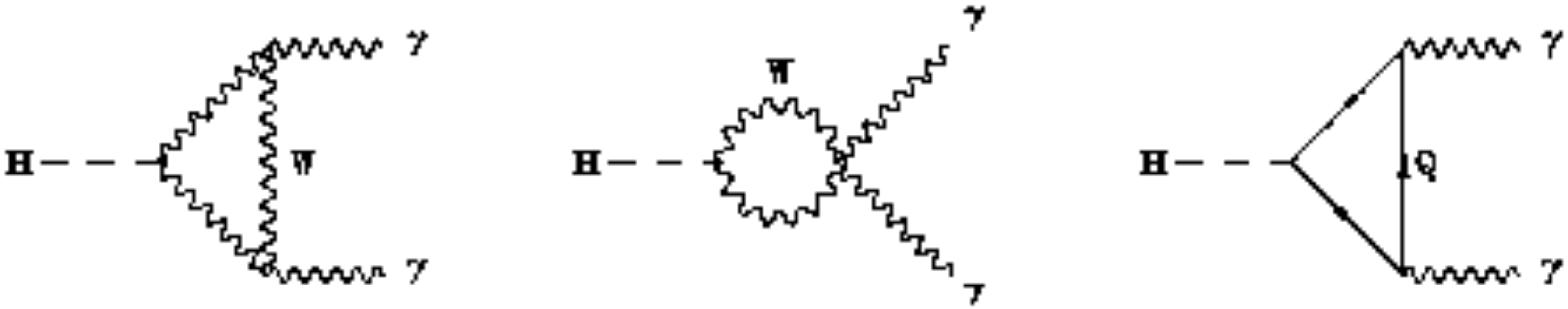


Low Mass SM Higgs: $H\rightarrow\gamma\gamma$

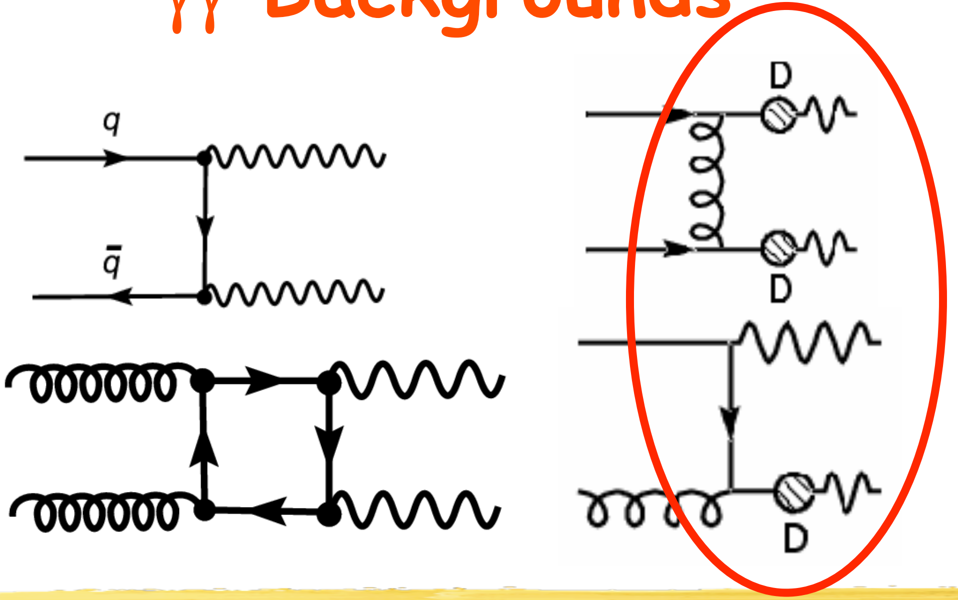




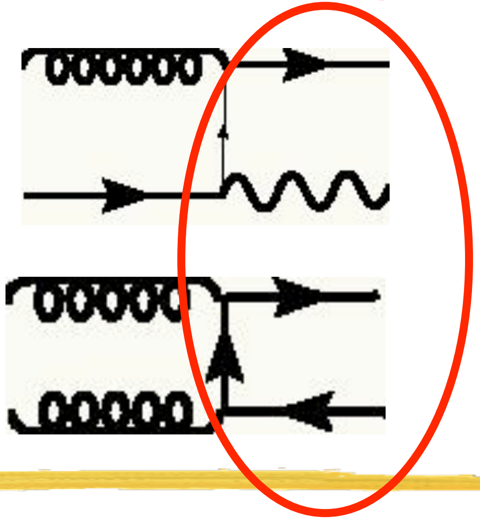
Higgs decay to γγ



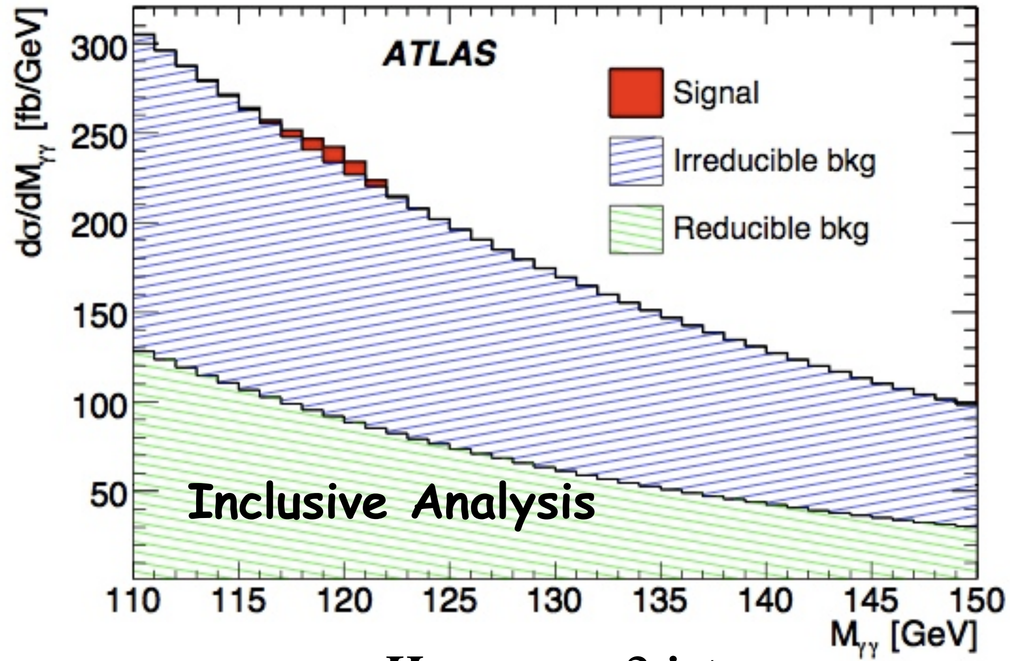
γγ Backgrounds



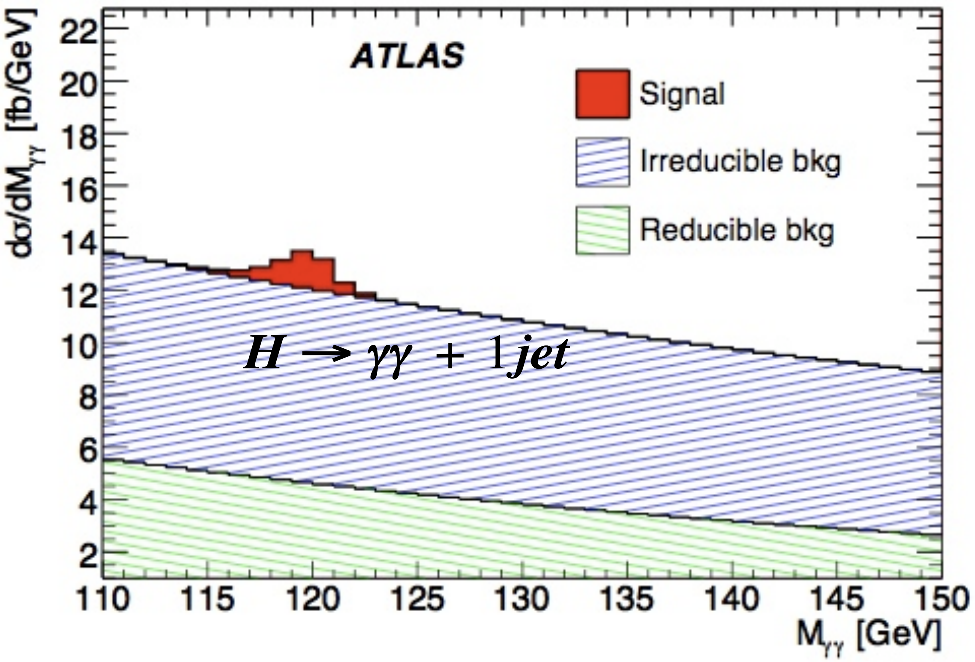
Reducible γj and jj Backgrounds



$q \rightarrow \pi^0$

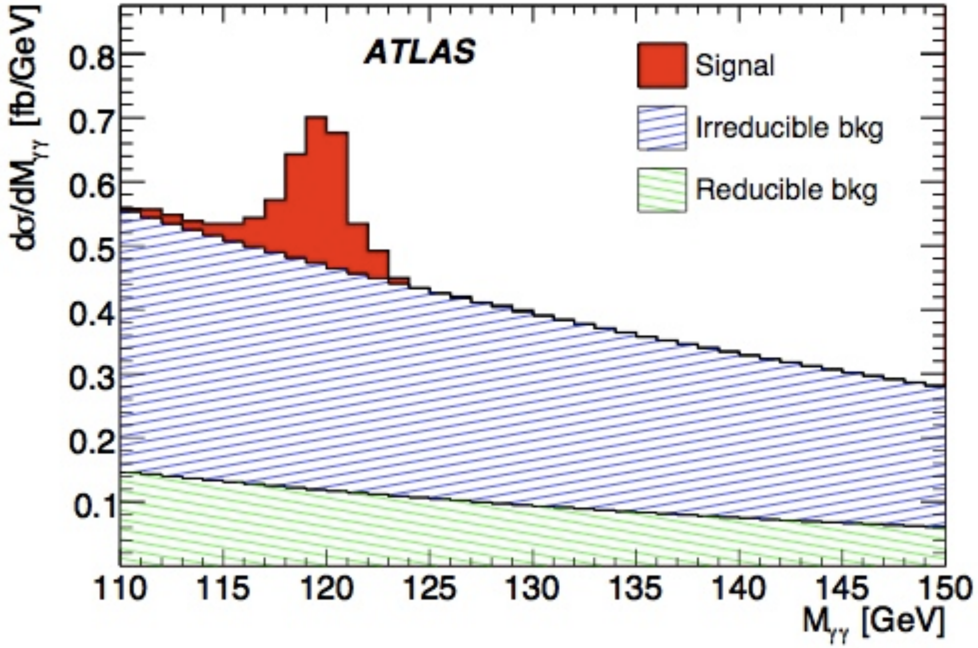


Inclusive Analysis



$H \rightarrow \gamma\gamma \, + \, 1\,jet$

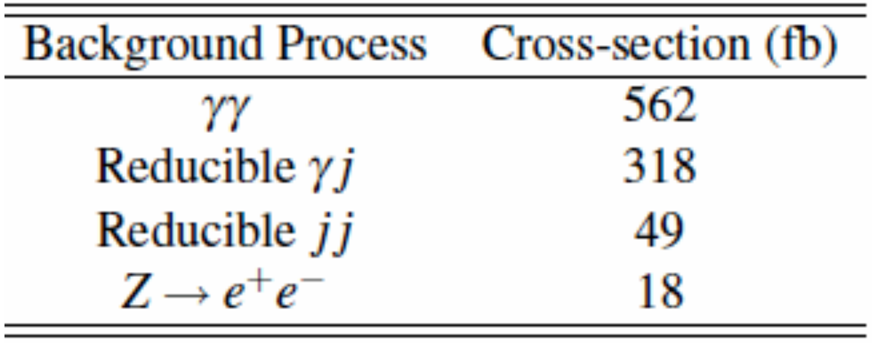
$H \rightarrow \gamma\gamma \, + \, 2\,jet$



This search is based on a side-band analysis. However, we want to understand the background composition the best we can: fraction of irreducible backgrounds and the contribution from fragmentation

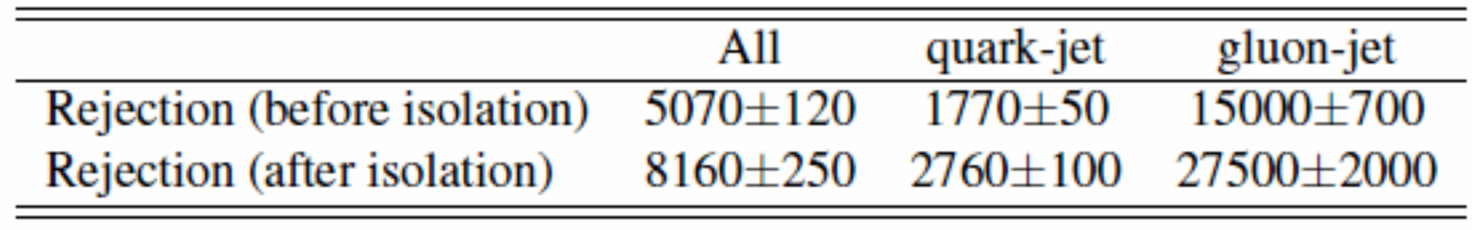
The contamination in the signal-like region from fake photons is significant

Background contributions with diphoton inv. Mass around 120 GeV. Relative contributions change little with mass in the range $110 < M_{\gamma\gamma} < 150$ GeV

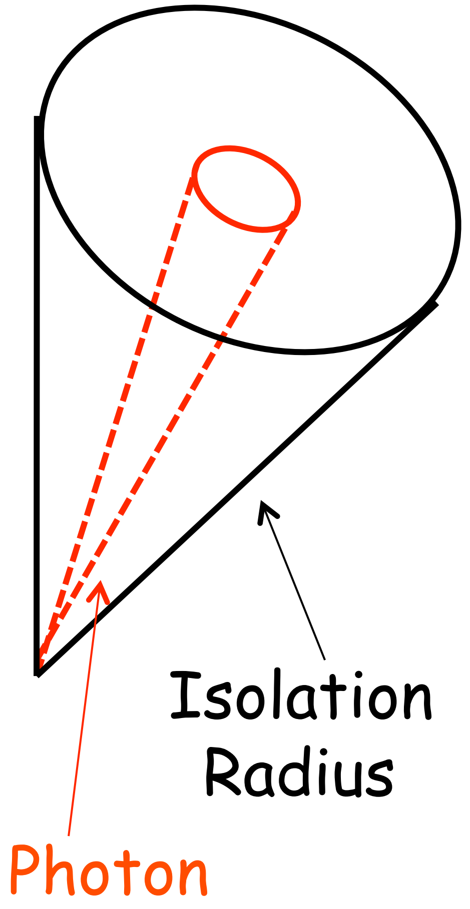


Background Process	Cross-section (fb)
$\gamma\gamma$	562
Reducible γj	318
Reducible jj	49
$Z \rightarrow e^+e^-$	18

Fake photons come mostly from fragmentation of quarks into π^0(~70% of fakes are from π^{0}s)



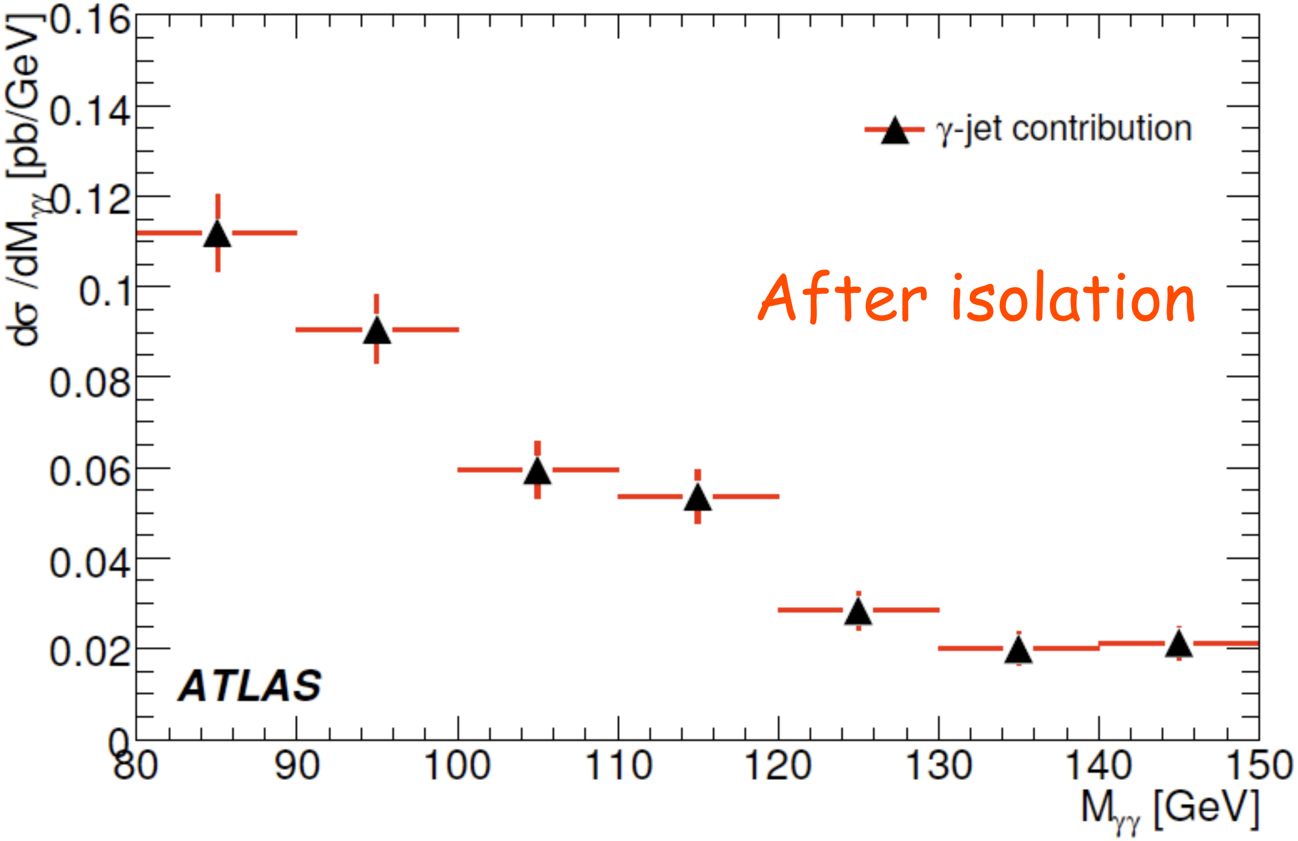
	All	quark-jet	gluon-jet
Rejection (before isolation)	5070±120	1770±50	15000±700
Rejection (after isolation)	8160±250	2760±100	27500±2000

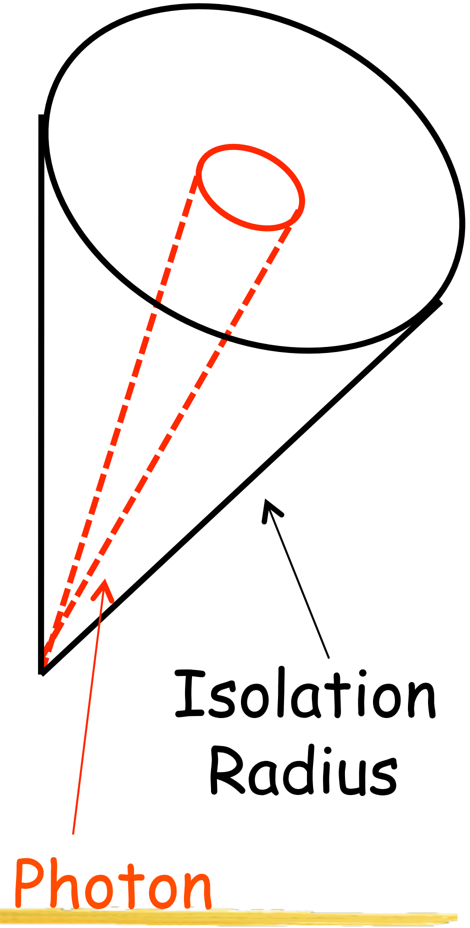


The normalization and shape from γj can be estimated with data by using a control sample with photon ID but no isolation

Observe a mixture of signal and background and use p_T^{γ}/p_T^{jet} (still need functional form from theory)

Try technique in measurement of γj, γγ cross-sections

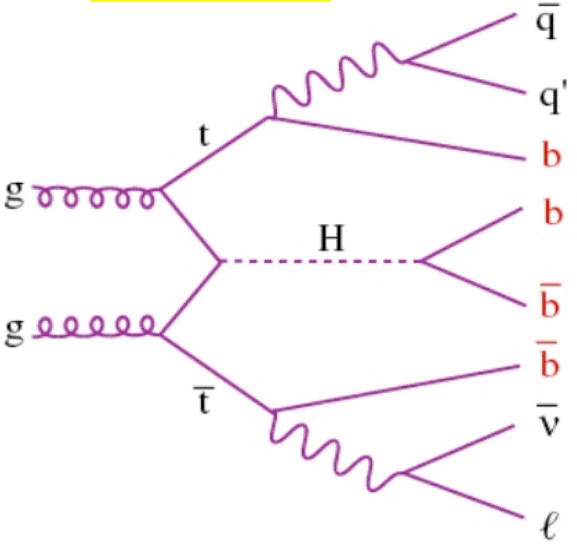




Low Mass SM Higgs: ttH→bb

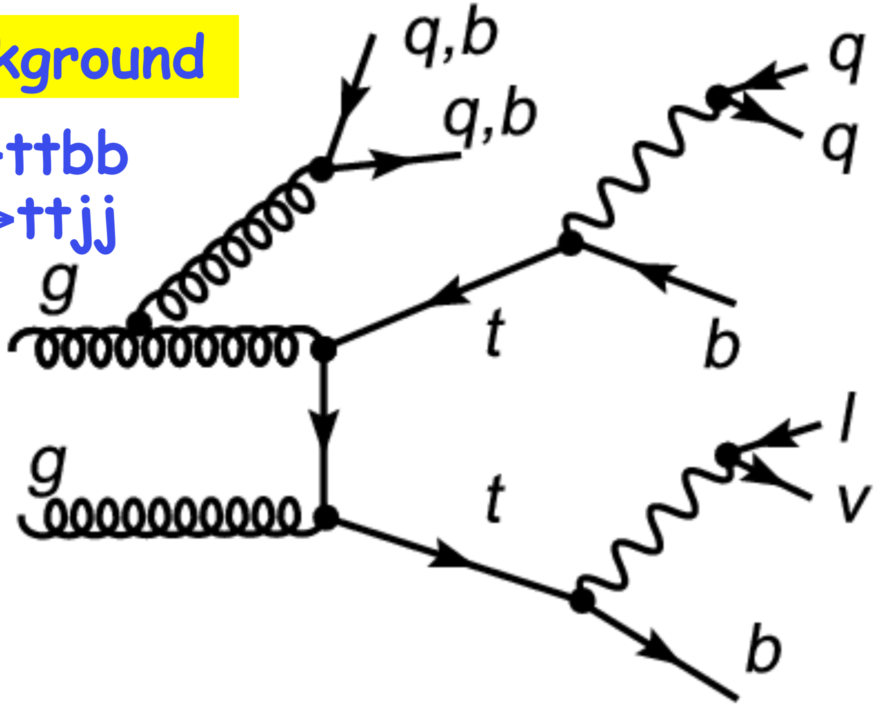
Complex final state: ttH(→bb)→lepton+ν+bbbb+jj

Signal



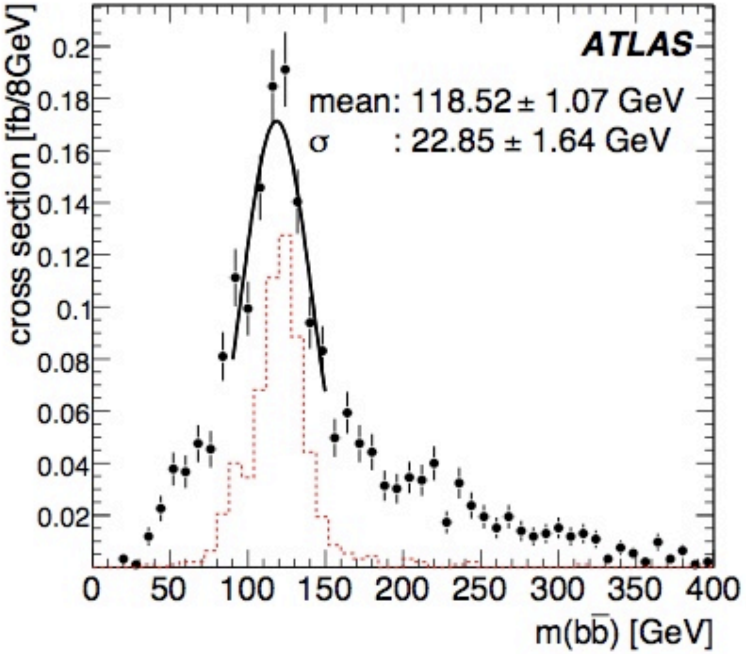
Background

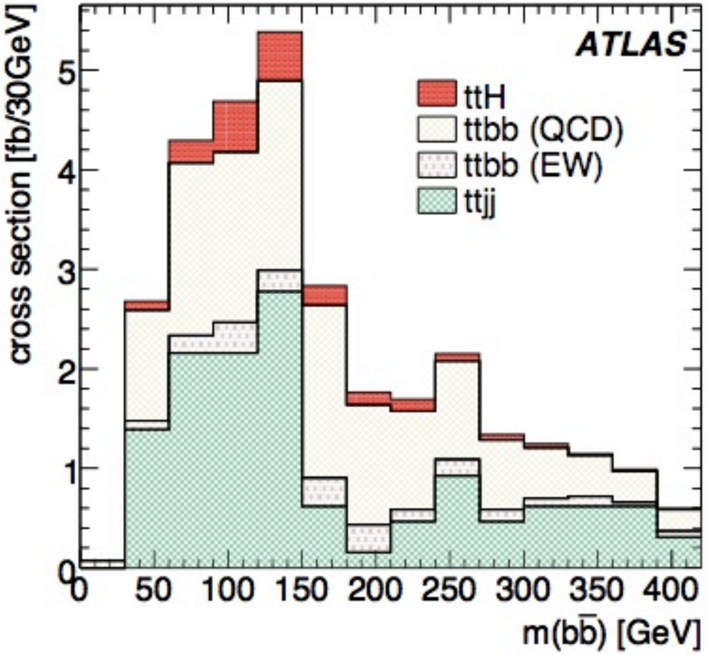
pp→ttbb
pp→ttjj



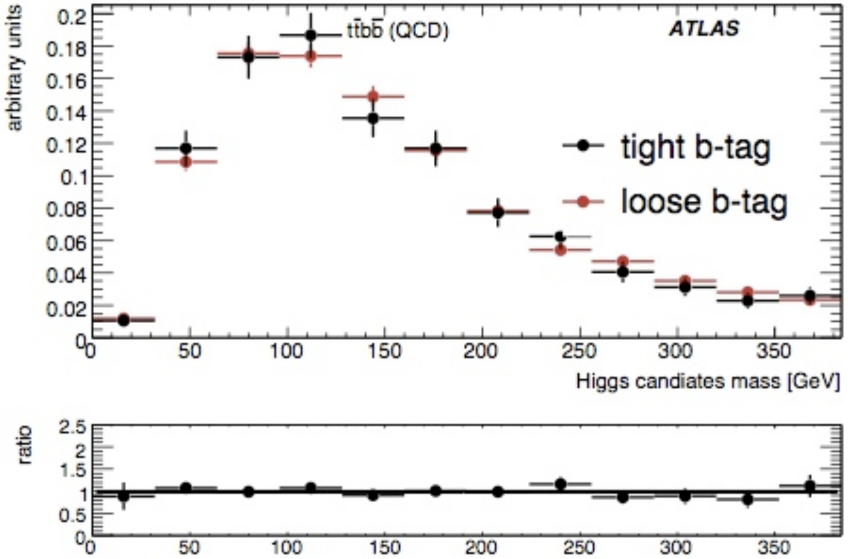
- Analysis very sensitive to b-tagging efficiency (ε_b^4)
  - Parton/Hadron level studies → $\varepsilon_b \geq 60\%$ needed
- Need ~100 times rejection against light jets and ~10 times against charm to suppress ttjj

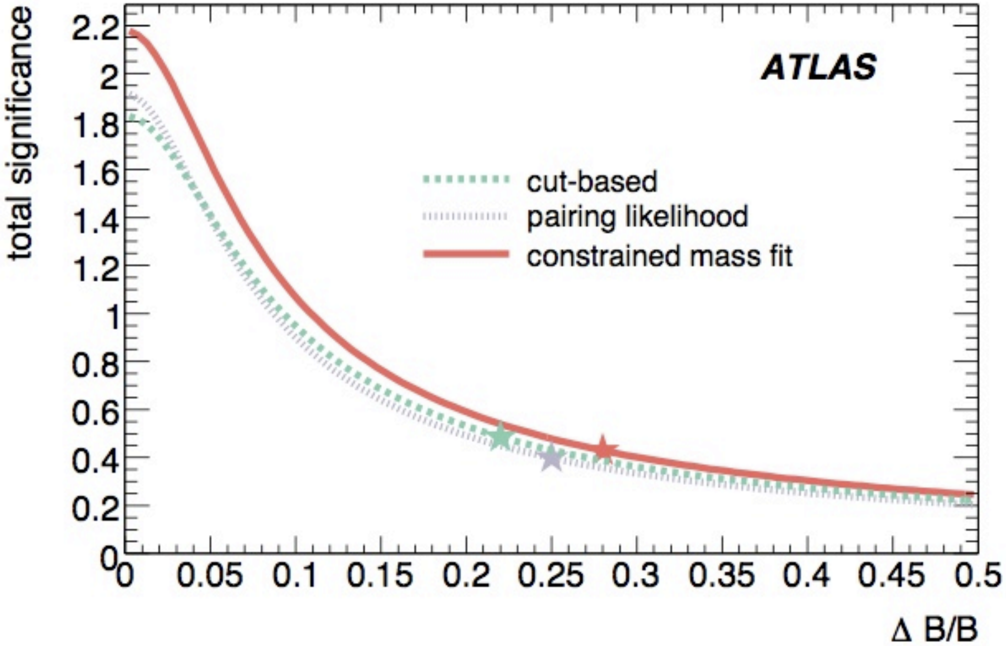
Results of cut analysis

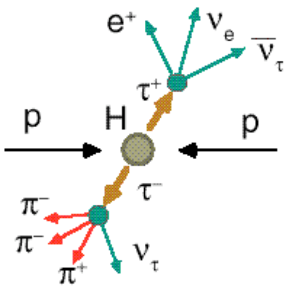




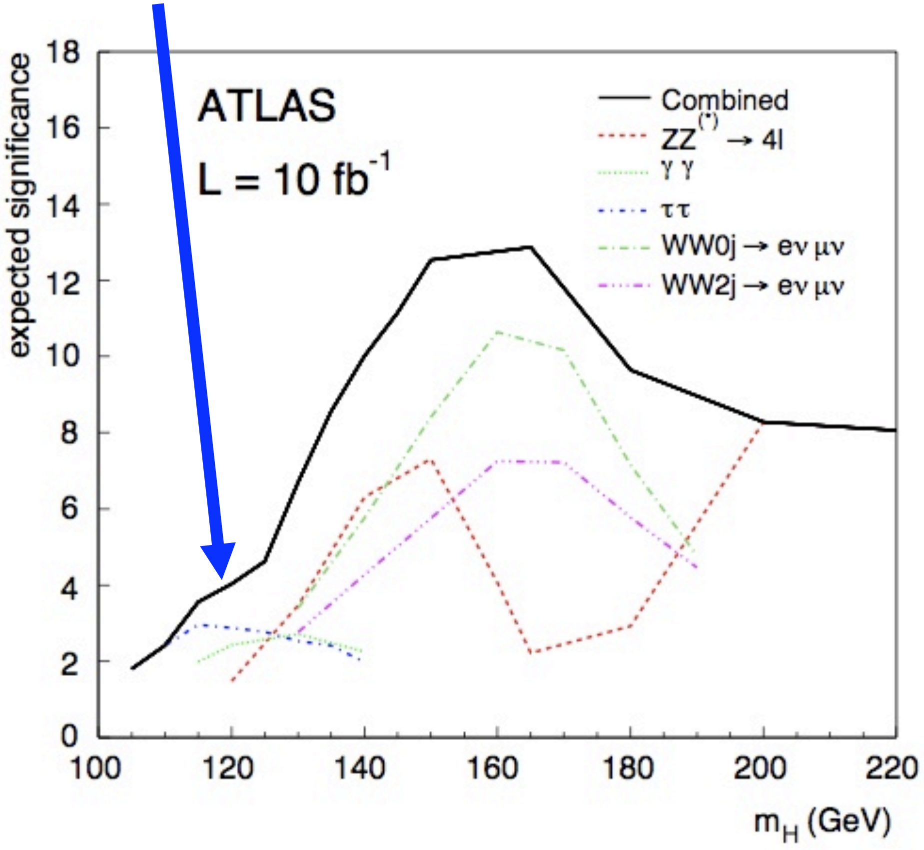
Data-driven extracion of background shape






e+
ν_e
ν̄_τ
τ+
p
H
p
τ−
π−
π−
π+
ν_τ

H→ττ


ATLAS
L = 10 fb-1
Combined
ZZ(*) → 4l
γ γ
ττ
WW0j → eν μν
WW2j → eν μν
expected significance
0
2
4
6
8
10
12
14
16
18
100
120
140
160
180
200
220
mH (GeV)

H→ττ Mass Reconstruction

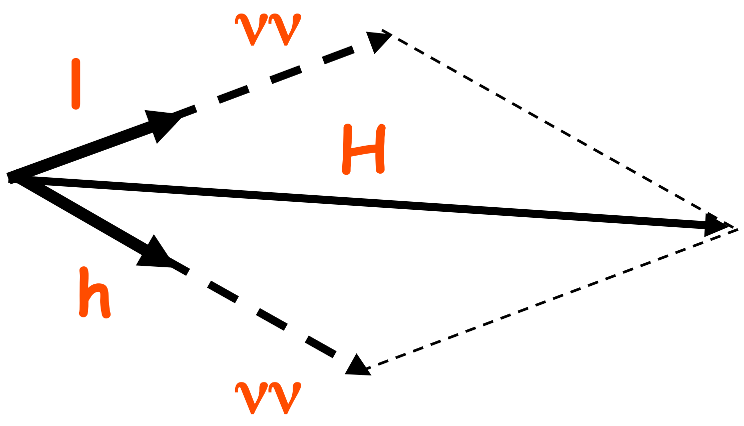
- In order to reconstruct the Z mass need to use the collinear approximation

Tau decay products are collinear to tau direction

Fraction of τ momentum carried by visible τ decay

$$\vec{P}_\tau = \frac{\vec{P}_l}{x_\tau}$$

$$M_{\tau\tau} \approx \frac{M_{ll}}{\sqrt{x_{\tau 1} x_{\tau 2}}}$$



$$\vec{P}_{T\tau 1} + \vec{P}_{T\tau 2} = \vec{P}_{Tl1} + \vec{P}_{Tl2} + \vec{P}_{Tmiss}$$

$$\Downarrow$$

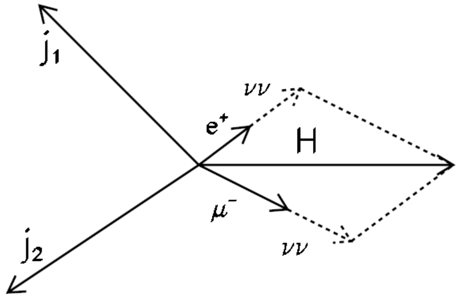
$$x_{\tau_1} = \frac{p_{T_{lep1},x} \cdot p_{T_{lep2},y} - p_{T_{lep1},y} \cdot p_{T_{lep2},x}}{p_{T_{Higgs},x} \cdot p_{T_{lep2},y} - p_{T_{Higgs},y} \cdot p_{T_{lep2},x}}$$

$$x_{\tau_2} = \frac{p_{T_{lep1},x} \cdot p_{T_{lep2},y} - p_{T_{lep1},y} \cdot p_{T_{lep2},x}}{p_{T_{Higgs},y} \cdot p_{T_{lep1},x} - p_{T_{Higgs},x} \cdot p_{T_{lep1},y}}$$

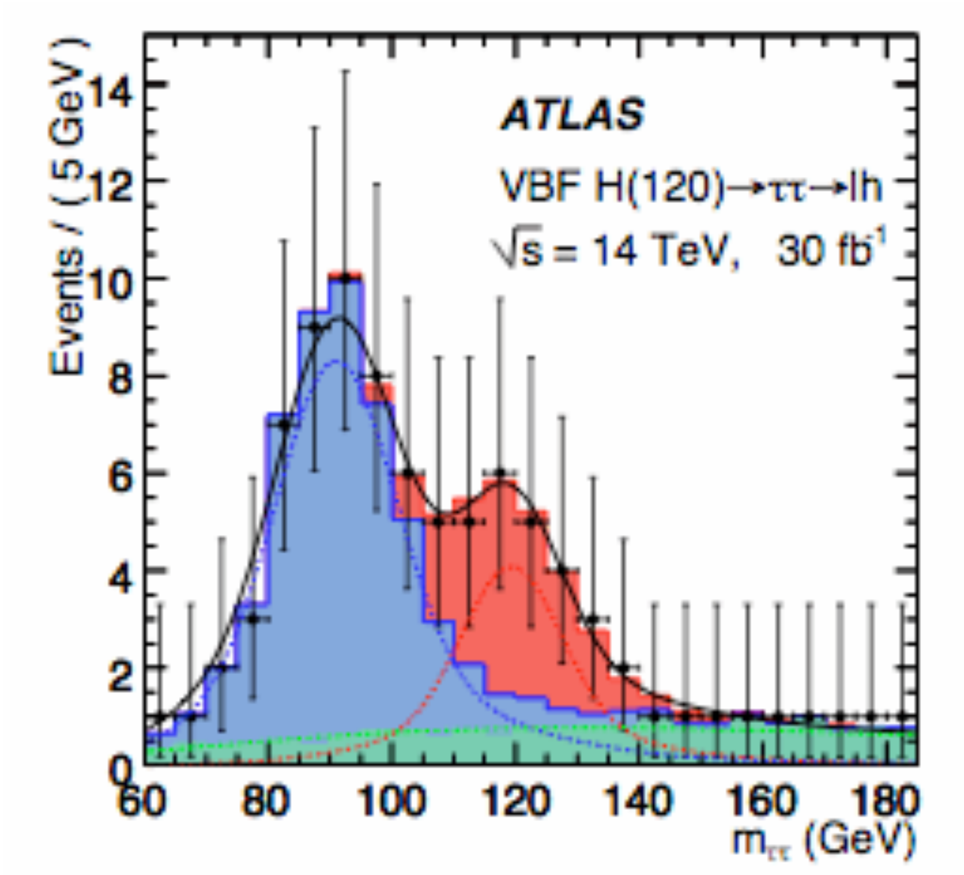
- $x_{\tau 1}$ and $x_{\tau 2}$ can be calculated if the missing E_T is known
- Good missing E_T reconstruction is essential

Low Mass SM H→ττ+jets

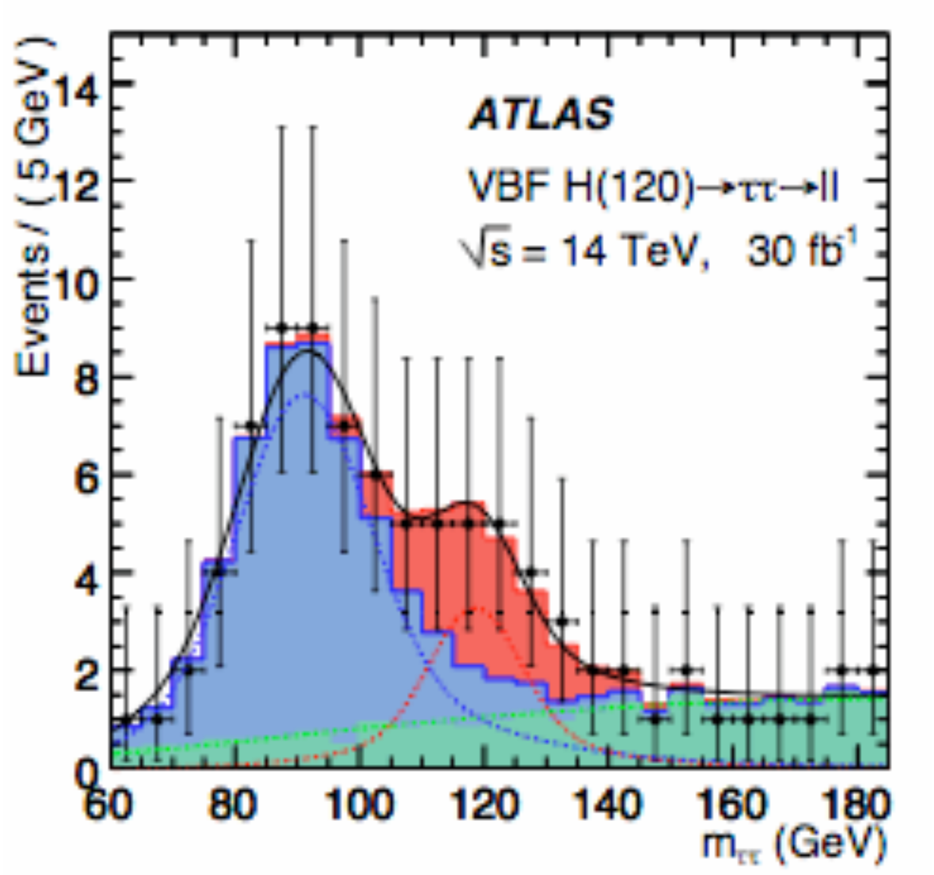
- Reconstruct Higgs mass with collinear approxim



H(→ττ→lh) +≥2jets

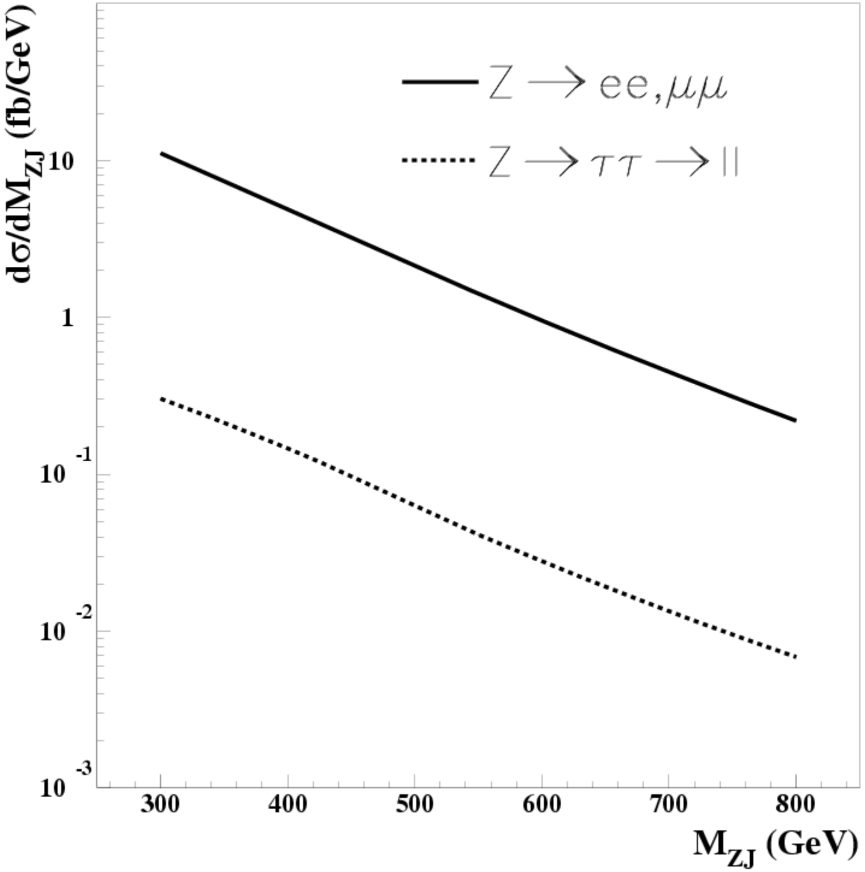


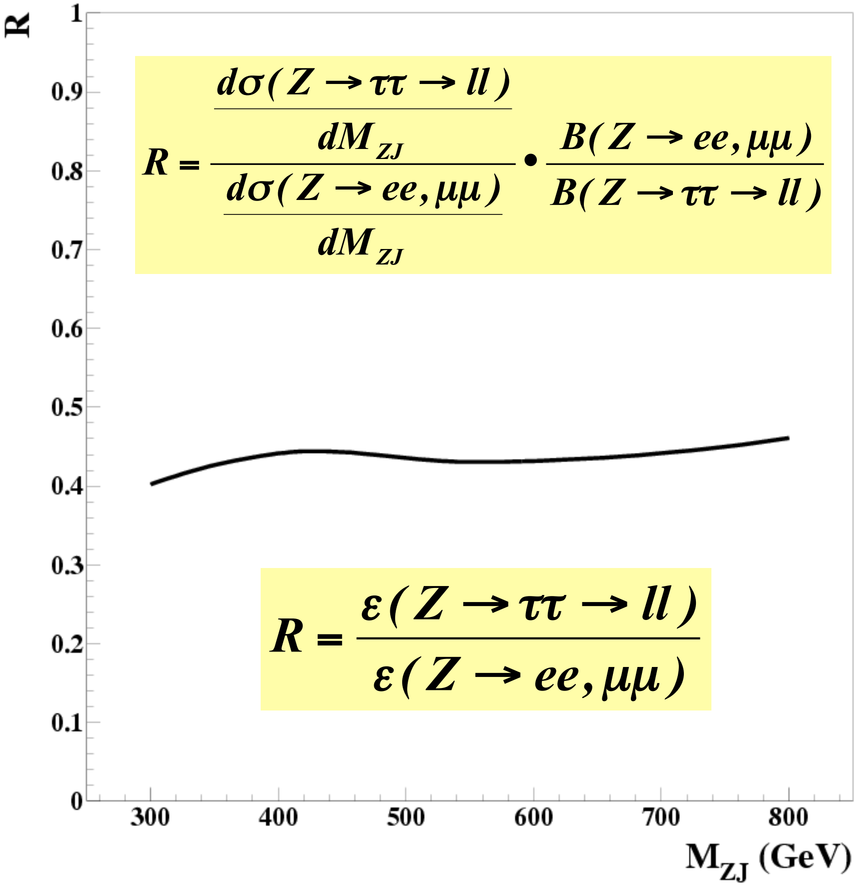
H(→ττ→2l) +≥2jets



Normalization of $Z \to \tau\tau$ using $Z \to ee,\mu\mu$

- $Z \to ee,\mu\mu$ offers about 35 times more statistics w.r.t to $Z \to \tau\tau \to ll$
  - Ratio of efficiencies depends weakly with M_{ZJ}, M_{jj} and can be easily determined with MC after validation with data



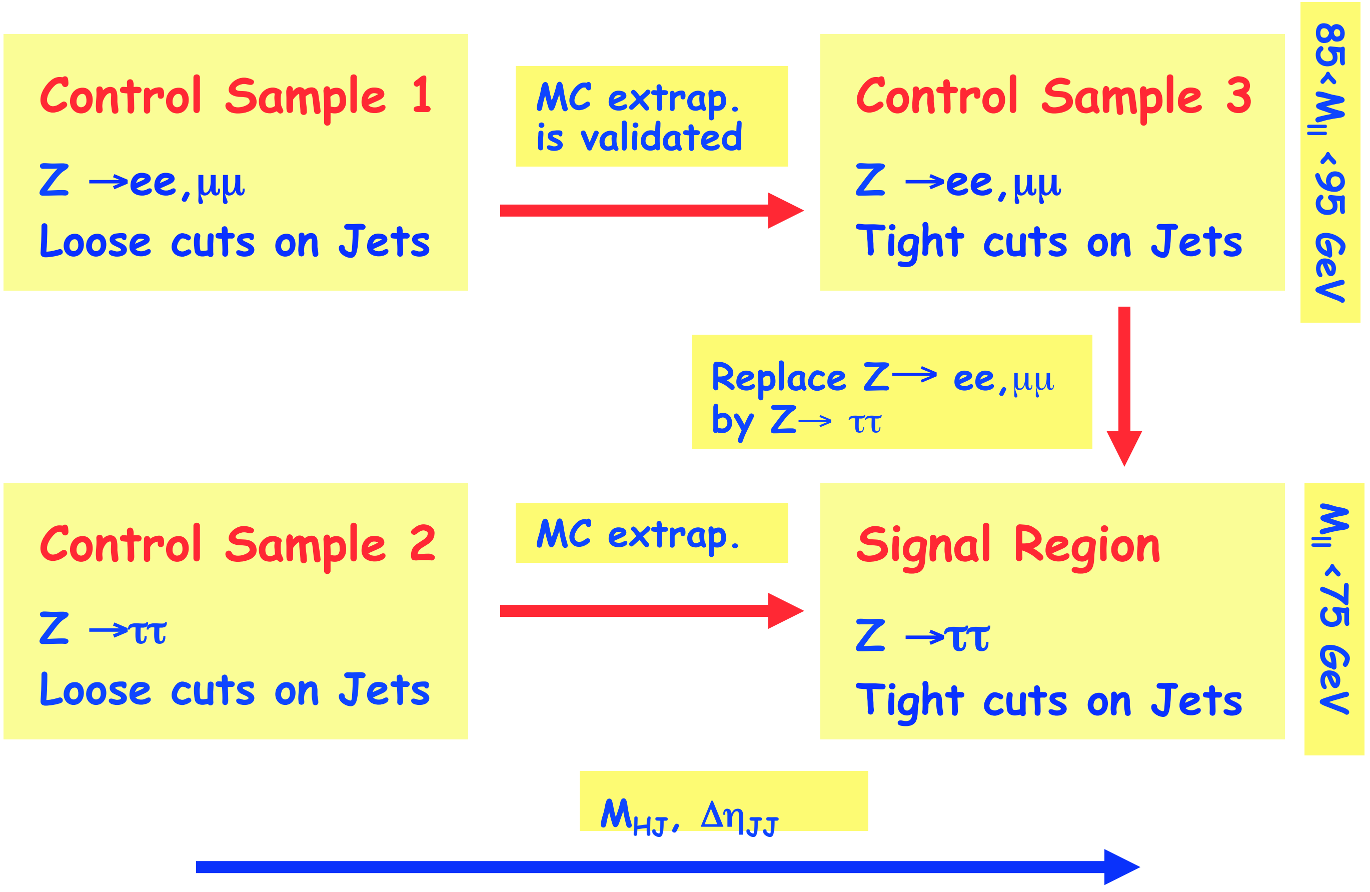


$$R = \frac{\dfrac{d\sigma(Z \to \tau\tau \to ll)}{dM_{ZJ}}}{\dfrac{d\sigma(Z \to ee,\mu\mu)}{dM_{ZJ}}} \bullet \frac{B(Z \to ee,\mu\mu)}{B(Z \to \tau\tau \to ll)}$$

$$R = \frac{\varepsilon(Z \to \tau\tau \to ll)}{\varepsilon(Z \to ee,\mu\mu)}$$

Two independent ways of extracting $Z\rightarrow\tau\tau$ shape

- Data driven and MC driven
- Similar procedure has been defined for $H\rightarrow WW^{(*)}$



Control Sample 1

$Z\rightarrow ee,\mu\mu$

Loose cuts on Jets

MC extrap. is validated

Control Sample 3

$Z\rightarrow ee,\mu\mu$

Tight cuts on Jets

$85<M_{ll}<95$ GeV

Replace $Z\rightarrow ee,\mu\mu$ by $Z\rightarrow\tau\tau$

Control Sample 2

$Z\rightarrow\tau\tau$

Loose cuts on Jets

MC extrap.

Signal Region

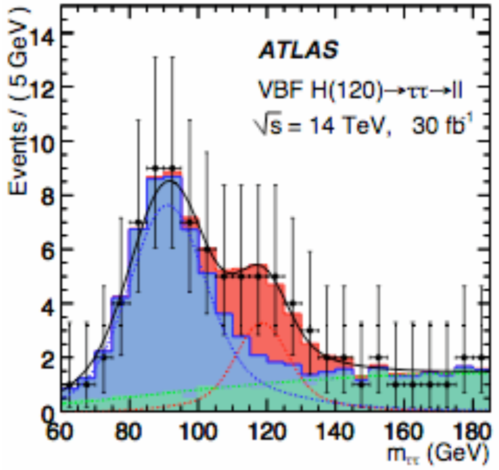
$Z\rightarrow\tau\tau$

Tight cuts on Jets

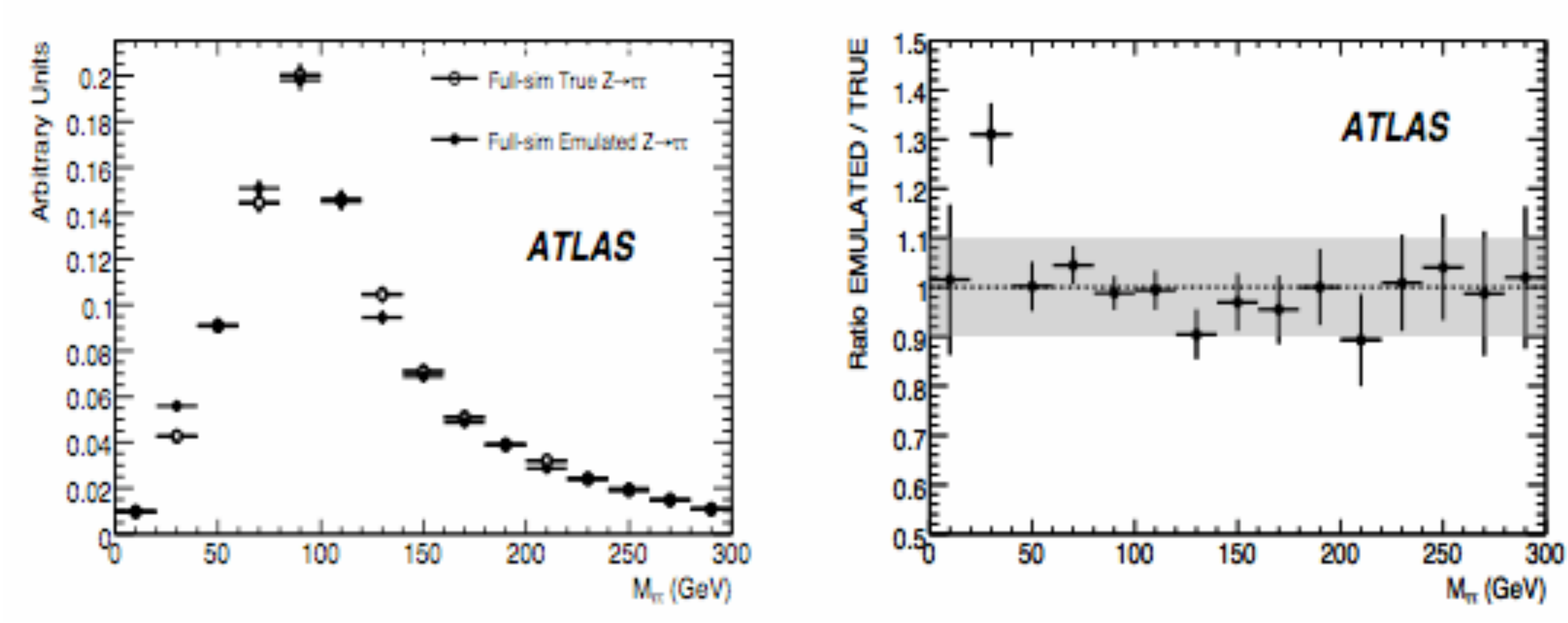
$M_{ll}<75$ GeV

M_{HJ}, $\Delta\eta_{JJ}$

Determine shape and normalization of $Z\rightarrow\tau\tau$ background

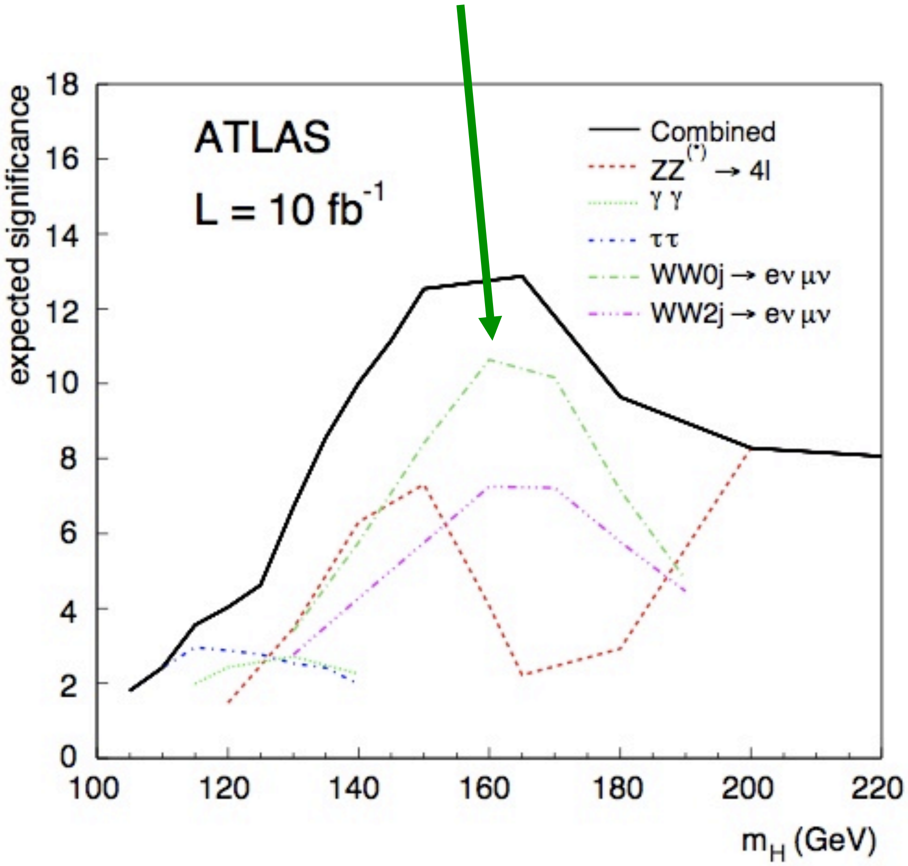


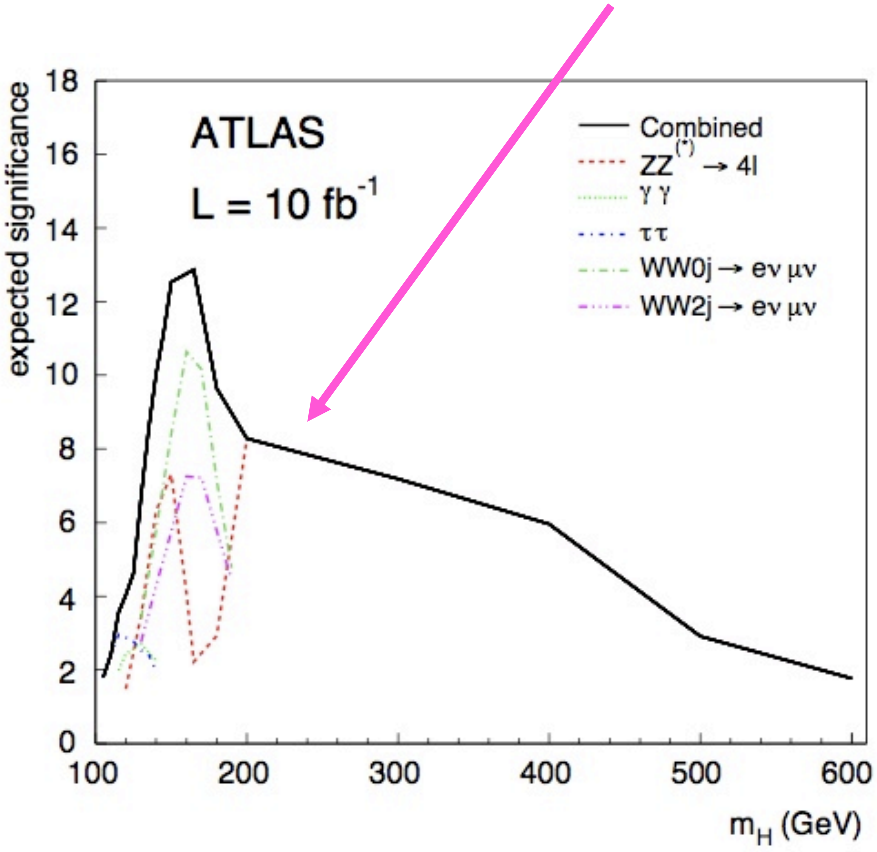
Data-Driven Extraction of Z+jets Background



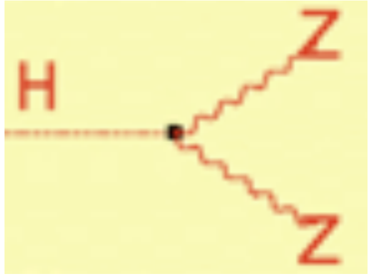
SM Higgs

$H(\to WW^{(*)}\to 2l2\nu),(\to ZZ^{(*)}\to 4l)$

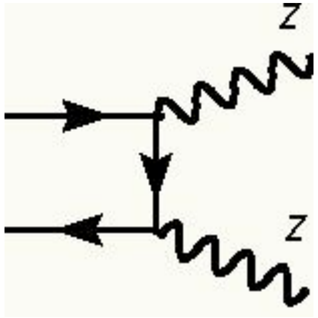


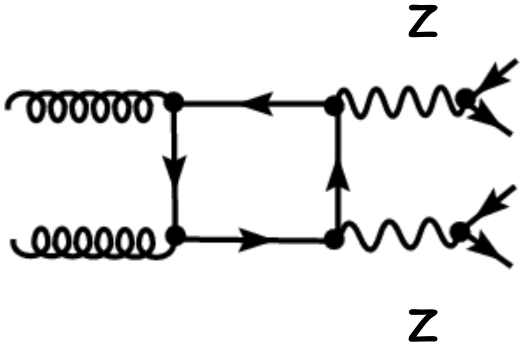


Higgs decay to Z^0Z^0

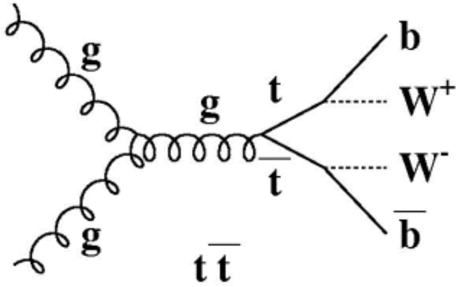


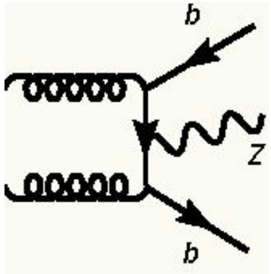
Irreducible Z^0Z^0 backgrounds



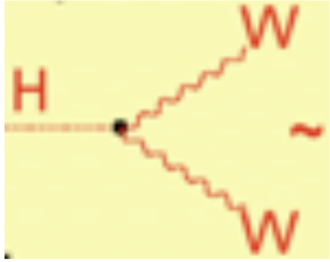


Reducible 4l backgrounds

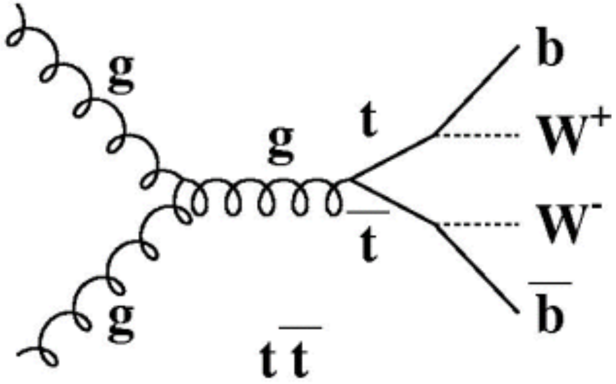


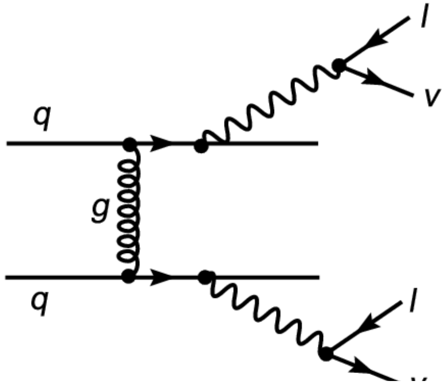


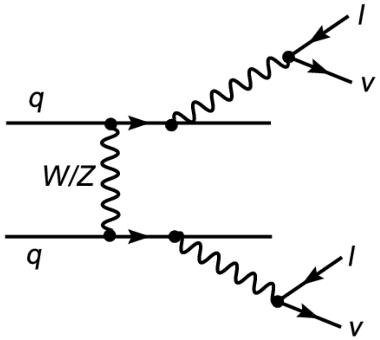
Higgs decay to W^+W^-

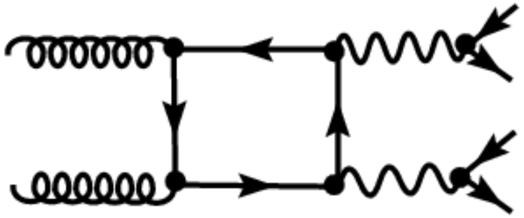


W^+W^- backgrounds







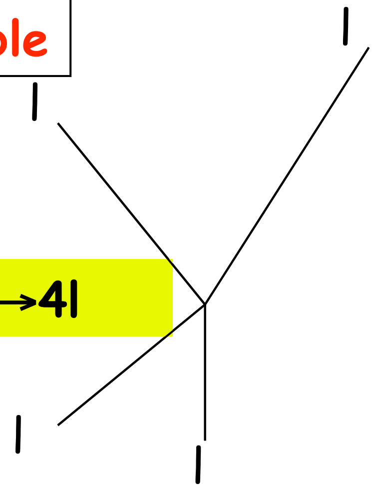


+ Single top
& non-resonant WWbb

Backgrounds Higgs→ZZ$^{(*)}$→4l (l=eμ)

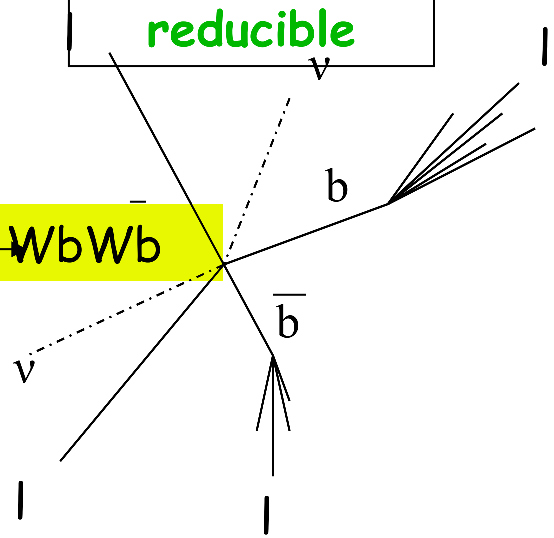
Continuum Irreducible

ZZ*/ γ*→4l



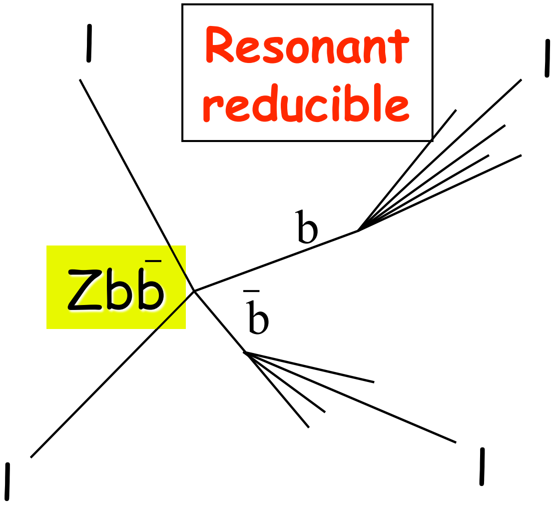
Non-Resonant reducible

$t\bar{t}$→WbW$\bar{b}$



Resonant reducible

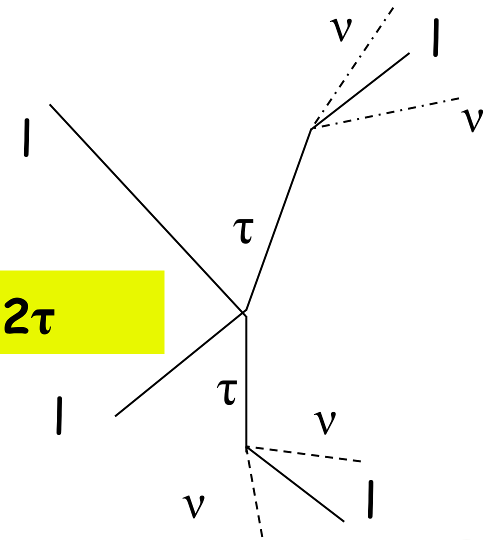
Z$b\bar{b}$



Side-bands or Z$^{(*)}$

Isolation techniques

ZZ*/ γ*→2l 2τ



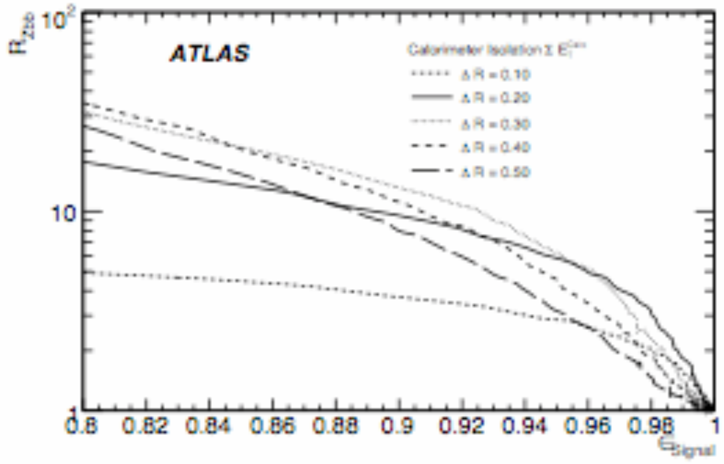


Figure 7: $Zb\bar{b}$ rejection versus $H \rightarrow 4\mu$ efficiency, for $m_H = 130GeV$, for various calorimetric isolation cone sizes.

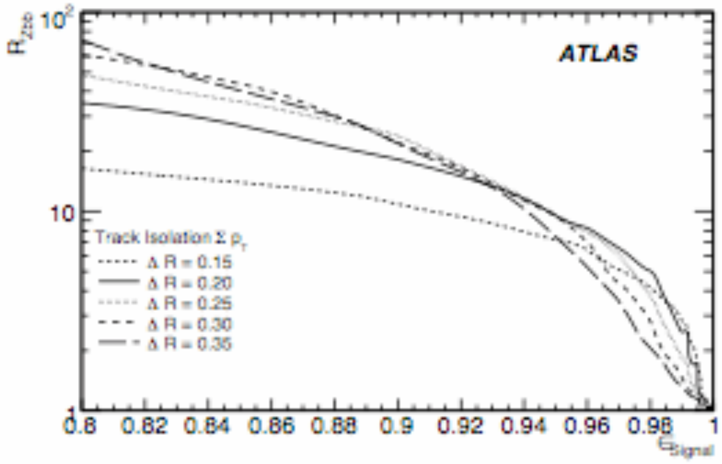


Figure 8: $Zb\bar{b}$ rejection versus $H \rightarrow 4\mu$ efficiency, for $m_H = 130GeV$, for various track isolation cone sizes.

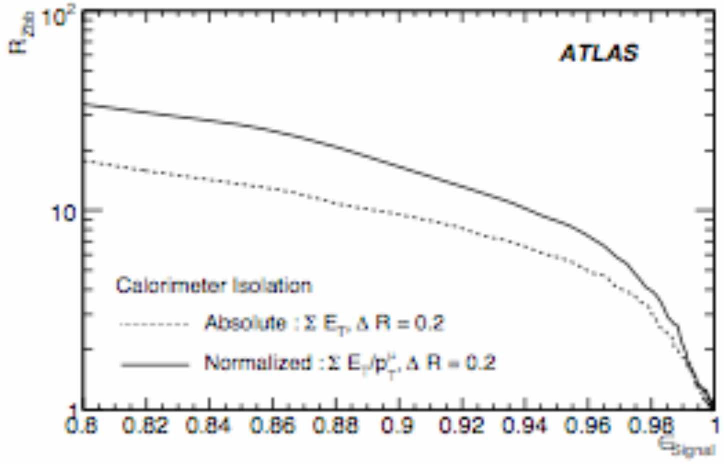


Figure 9: $Zb\bar{b}$ rejection versus $H \rightarrow 4\mu$ efficiency, for $m_H = 130GeV$, for standard and normalized calorimetric isolation calculated in a ΔR=0.2 cone around the muon track.

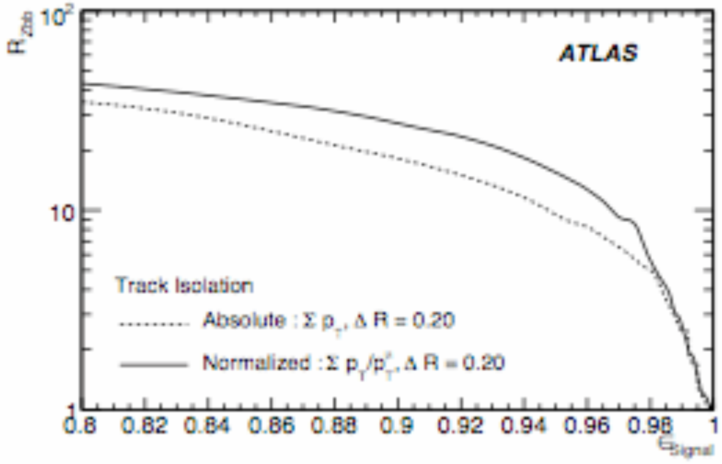
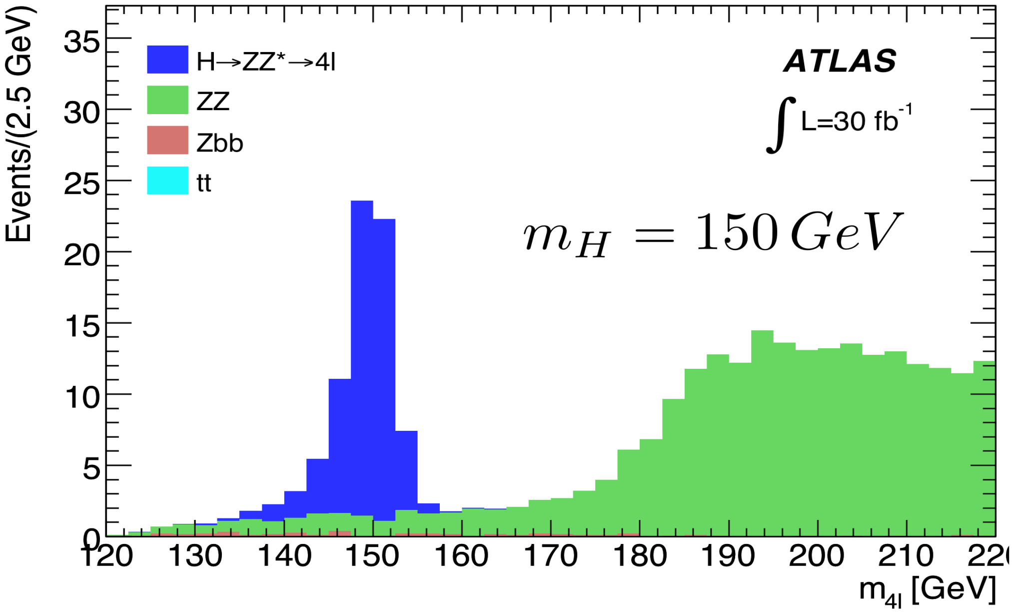
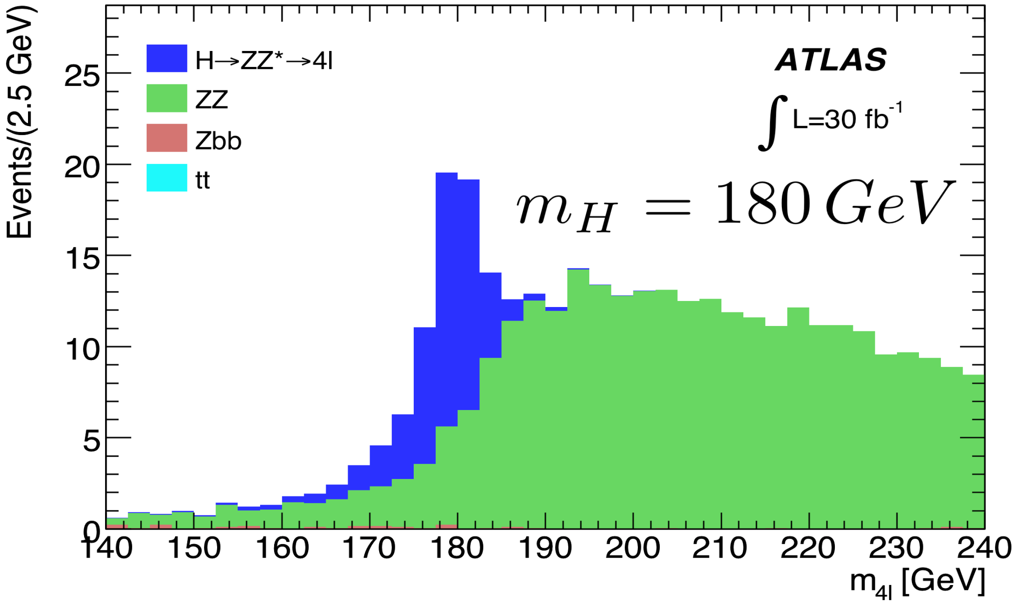
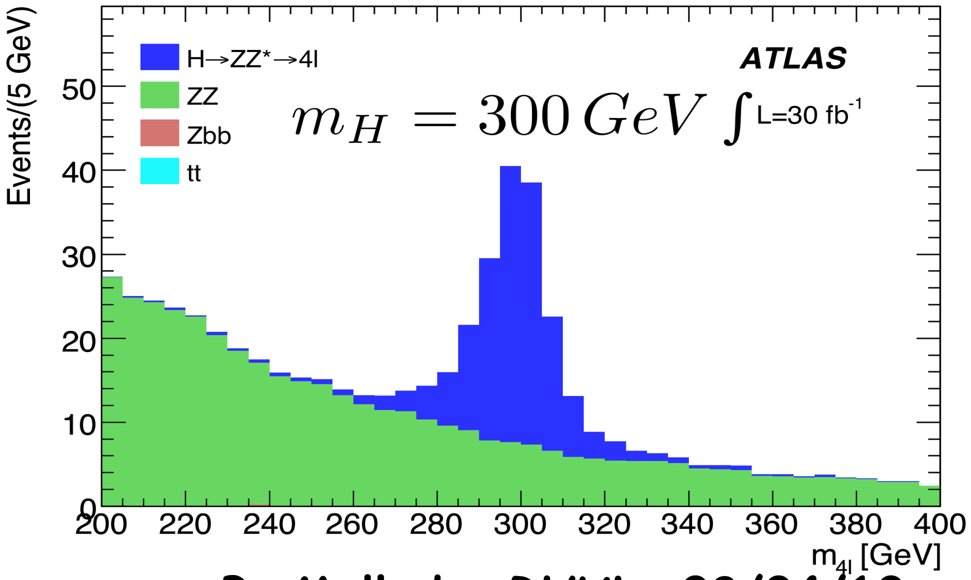


Figure 10: $Zb\bar{b}$ rejection versus $H \rightarrow 4\mu$ efficiency, for $m_H = 130GeV$, for standard and normalized track isolation calculated in a ΔR=0.2 cone around the muon track.



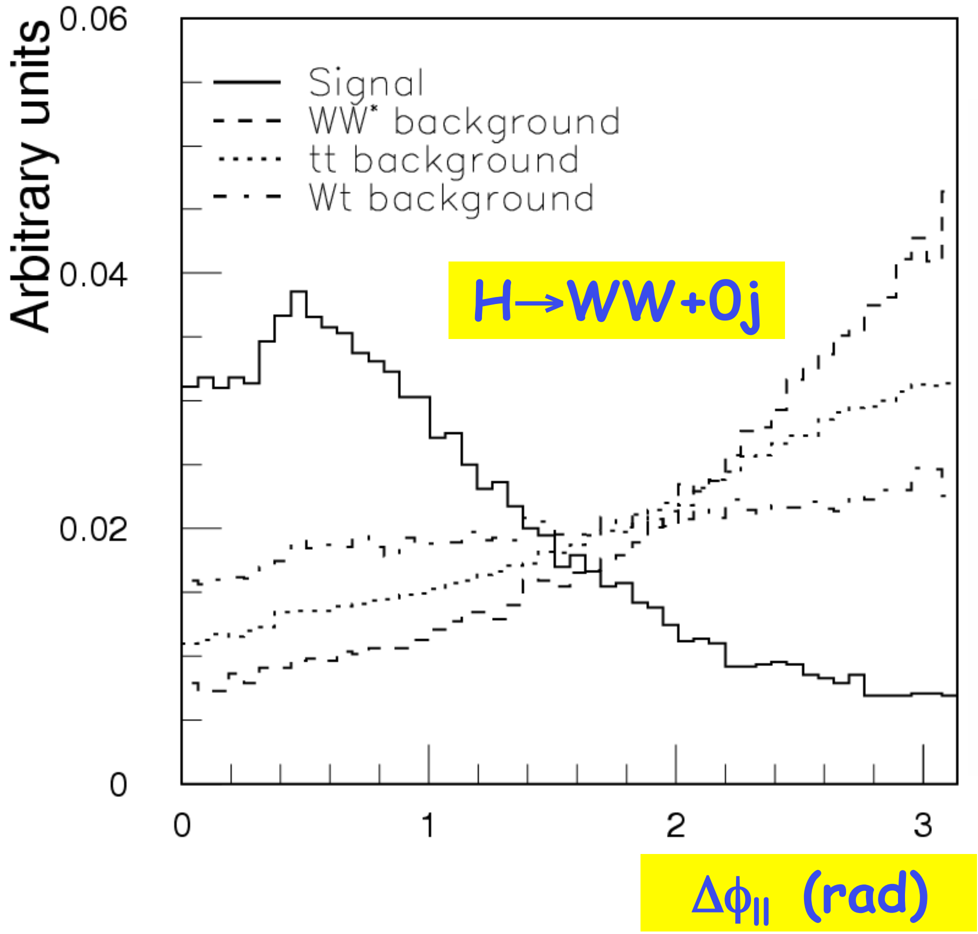


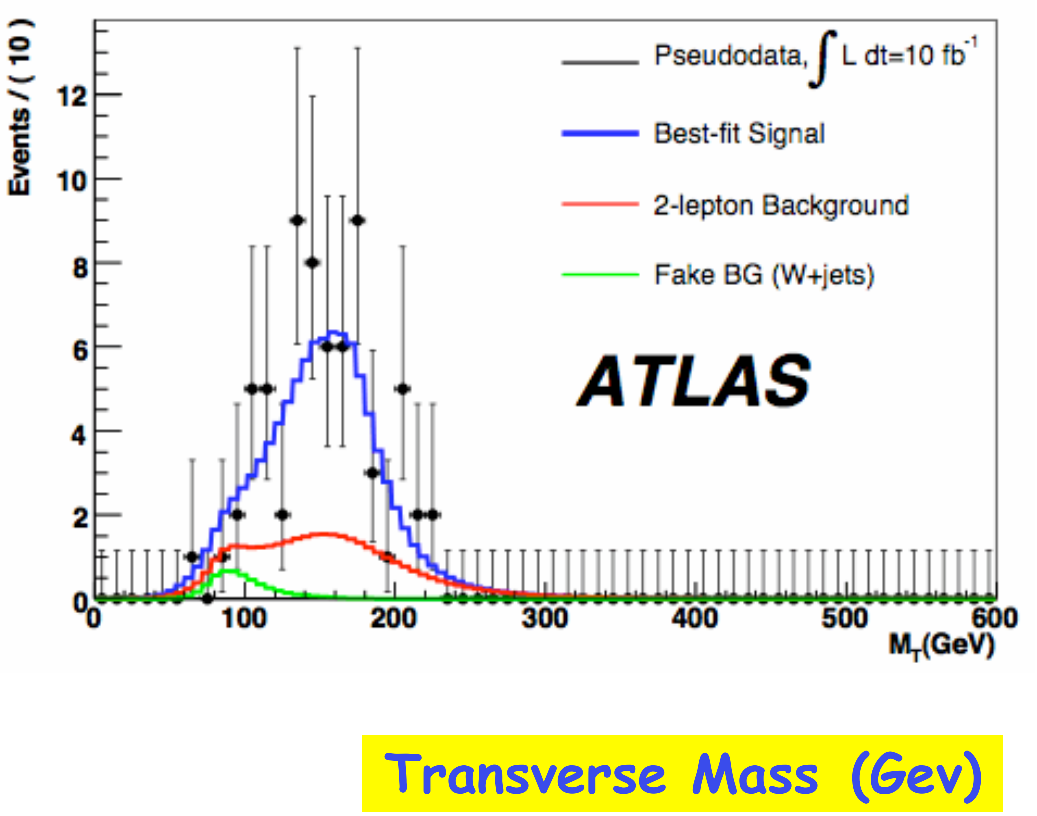


With large data samples the ZZ* background normalization can be extracted using side-bands. This is problematic at low luminosity: Can get normalization from Z$^{(*)}$

SM Higgs $H \rightarrow WW^{(*)} \rightarrow 2l2\nu$

- **Strong potential due to large signal yield, but no narrow resonance. Left basically with event counting experiment**





Normalizing VV with $Z^{(*)}$

- Strong similarities of diagrams since dominant cross-section comes from qq->V(V) via EW couplings
- Ratios VV/V expected to reduce pdf and a significant portion of the scale uncertainty
  - This is an asset especially at the very beginning of data taking when global pdf fits will not be available

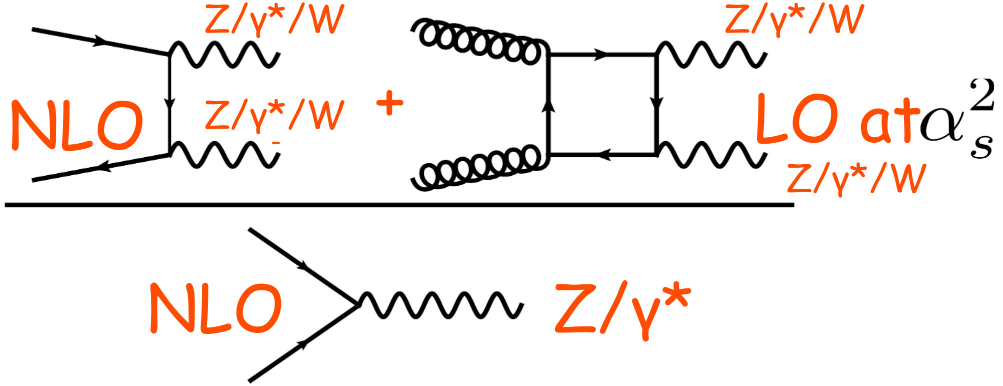
$$\overset{\text{Prediction}}{N(VV)} = \overset{\text{Theory}}{\left(\frac{\sigma(pp \to VV)}{\sigma(pp \to Z^{(*)})}\right)_{Th}} \cdot \overset{\text{Experimental efficiencies}}{\epsilon(ll \to Nl)} \cdot \overset{\text{Observed}}{N_{Obs}(Z^{(*)})}$$

M. Dittmar, F. Pauss, and D. Zurcher, Phys. Rev. **D56**, 7284 (1997)

Abdullin et al. in hep-ph/0604120 computed the ratio ZZ/Z to NLO

Ratio ZZ(WW)/Z$^{(*)}$

- **The production of ZZ and WW is enhanced by large contributions from gg->VV with gluons in the initial state**
  - **Formally a part of the NNLO contribution, but enhanced due to the large gluon flux**



Phys.Rev.D80:054023,2009

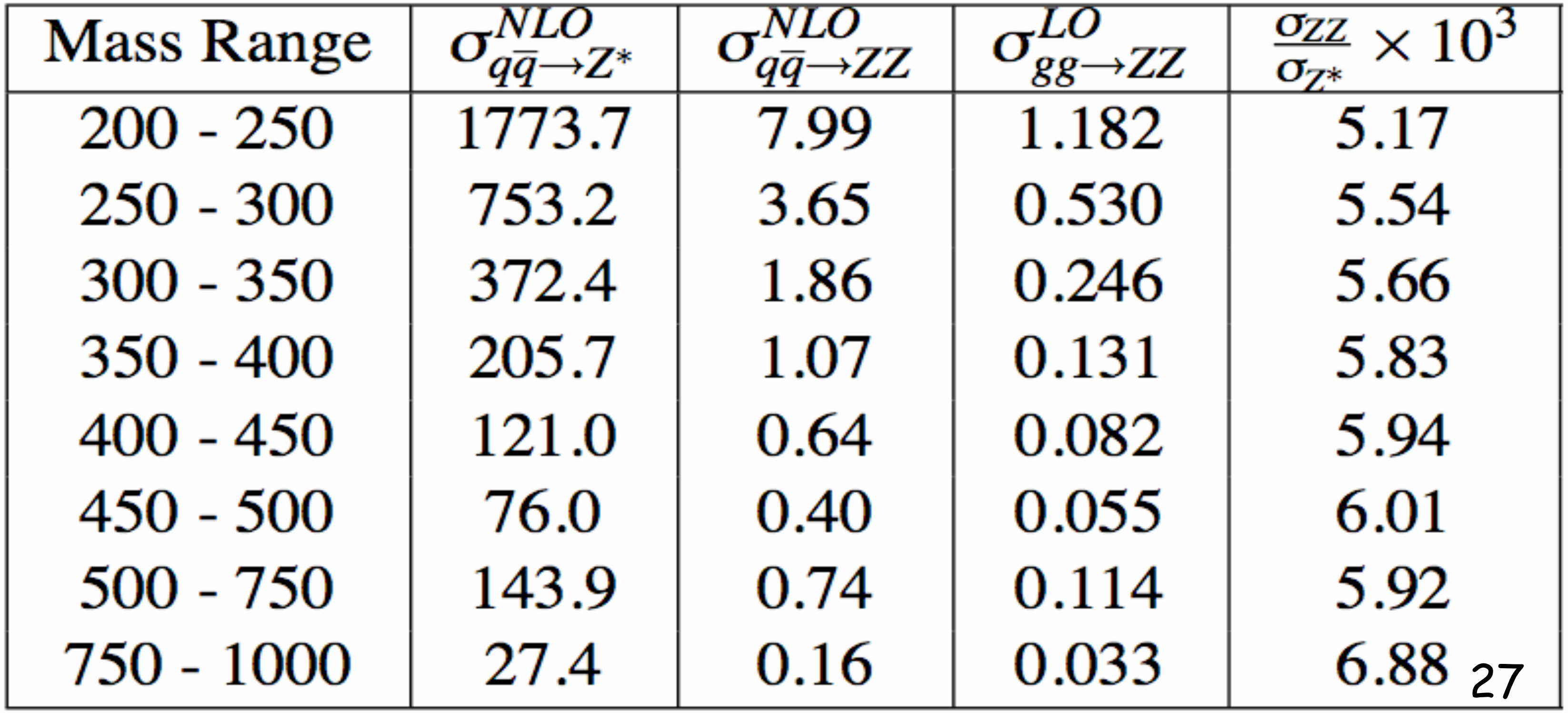
Including decays into leptons

$$R = \frac{\sigma^{NLO}_{q\bar{q}\rightarrow ZZ,WW} + \sigma^{LO}_{gg\rightarrow ZZ,WW}}{\sigma^{NLO}_{q\bar{q}\rightarrow Z^{(*)}}}$$

Nominal Values of ZZ/Z*

- **Ratios are constructed such that the invariant mass of Z* and ZZ are in the same bin**
  - **Contribution from gg->ZZ increases sigma by ~13%**
  - **Ratio depends weakly with Mass (nice surprise!)**

Cross-sections in fb

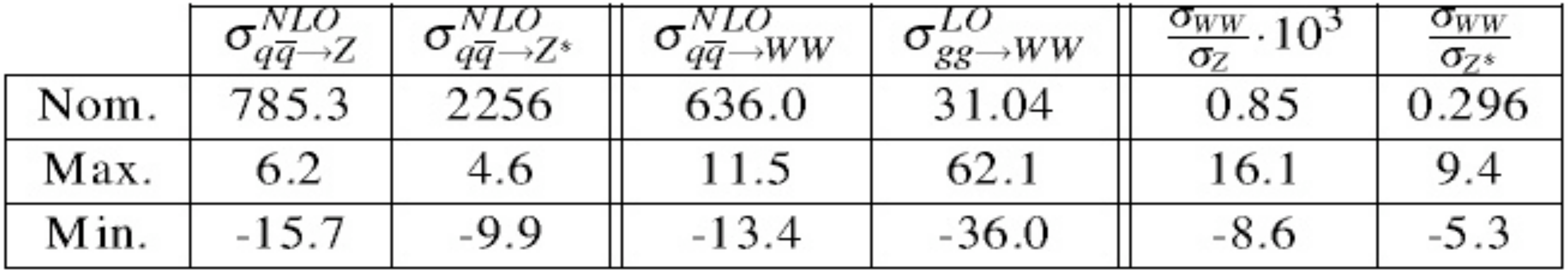


Mass Range	$\sigma^{NLO}_{q\bar{q}\to Z^*}$	$\sigma^{NLO}_{q\bar{q}\to ZZ}$	$\sigma^{LO}_{gg\to ZZ}$	$\frac{\sigma_{ZZ}}{\sigma_{Z^*}} \times 10^3$
200 - 250	1773.7	7.99	1.182	5.17
250 - 300	753.2	3.65	0.530	5.54
300 - 350	372.4	1.86	0.246	5.66
350 - 400	205.7	1.07	0.131	5.83
400 - 450	121.0	0.64	0.082	5.94
450 - 500	76.0	0.40	0.055	6.01
500 - 750	143.9	0.74	0.114	5.92
750 - 1000	27.4	0.16	0.033	6.88

Ratio WW/Z$^{(*)}$

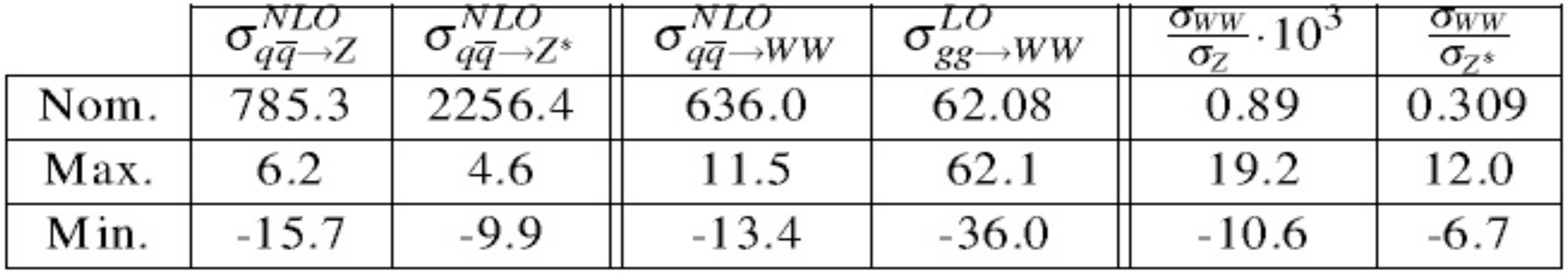
- **Scale-related uncertainties arise from changing scales by factors of 4 (*4,/4)**
  - **Pick biggest deviation of changing at the same time and in opposite directions**

M_{Z^*} >185 GeV



	$\sigma^{NLO}_{q\bar{q}\to Z}$	$\sigma^{NLO}_{q\bar{q}\to Z^*}$	$\sigma^{NLO}_{q\bar{q}\to WW}$	$\sigma^{LO}_{gg\to WW}$	$\frac{\sigma_{WW}}{\sigma_Z}\cdot 10^3$	$\frac{\sigma_{WW}}{\sigma_{Z^*}}$
Nom.	785.3	2256	636.0	31.04	0.85	0.296
Max.	6.2	4.6	11.5	62.1	16.1	9.4
Min.	-15.7	-9.9	-13.4	-36.0	-8.6	-5.3

Same as above after multiplying σ(gg->WW) by two



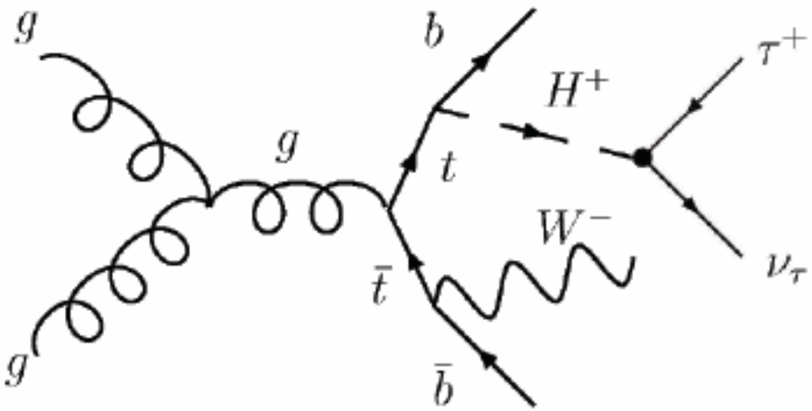
	$\sigma^{NLO}_{q\bar{q}\to Z}$	$\sigma^{NLO}_{q\bar{q}\to Z^*}$	$\sigma^{NLO}_{q\bar{q}\to WW}$	$\sigma^{LO}_{gg\to WW}$	$\frac{\sigma_{WW}}{\sigma_Z}\cdot 10^3$	$\frac{\sigma_{WW}}{\sigma_{Z^*}}$
Nom.	785.3	2256.4	636.0	62.08	0.89	0.309
Max.	6.2	4.6	11.5	62.1	19.2	12.0
Min.	-15.7	-9.9	-13.4	-36.0	-10.6	-6.7

MSSM $H^{\pm}$

Data-driven techniques also studied for Neutral Higgs ($\rightarrow \mu\mu, \tau\tau$)

Searches for Charged Higgs ($M_{H^\pm}<M_t$)

Signal Final State

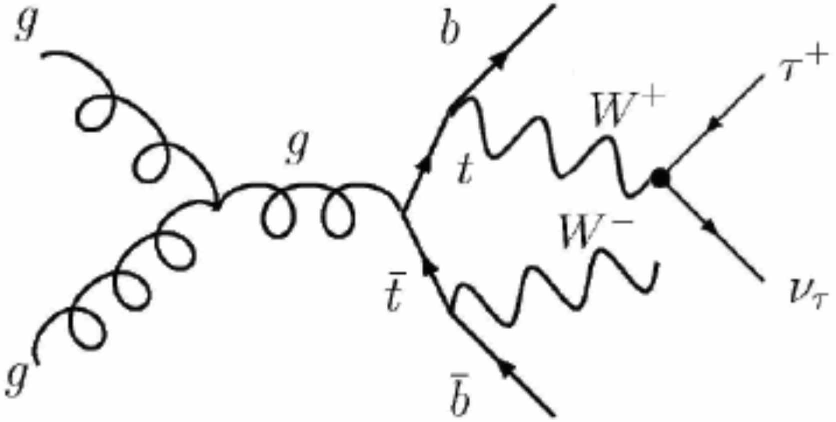


$H^+ \to \tau_H \nu;\ W \to qq$

$H^+ \to \tau_L \nu;\ W \to qq$

$H^+ \to \tau_H \nu;\ W \to l\nu$

Dominant Background



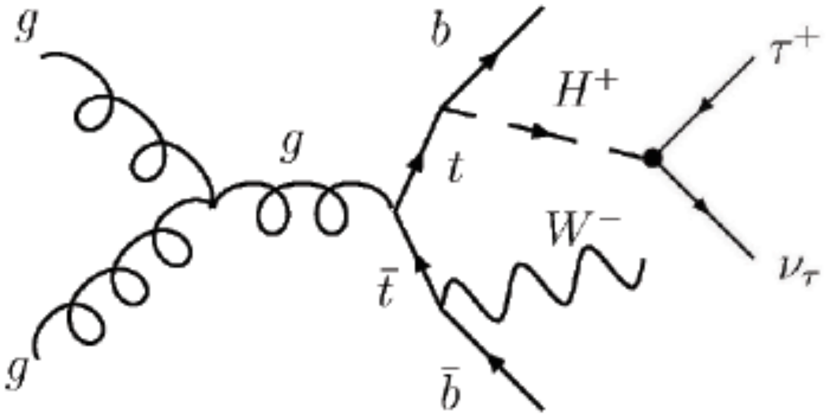
$W \to \tau_H \nu;\ W \to qq$

$W \to \tau_L \nu;\ W \to qq$

$W \to \tau_H \nu;\ W \to l\nu$

Top background (τ→l,h) extraction with Top events with μ (W→μ)

Signal Final State



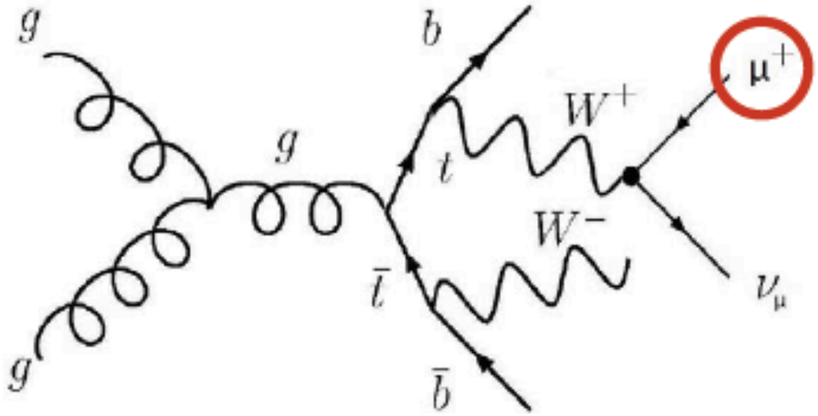
$$H^+ \to \tau_H \nu;\ W \to qq$$

$$H^+ \to \tau_L \nu;\ W \to qq$$

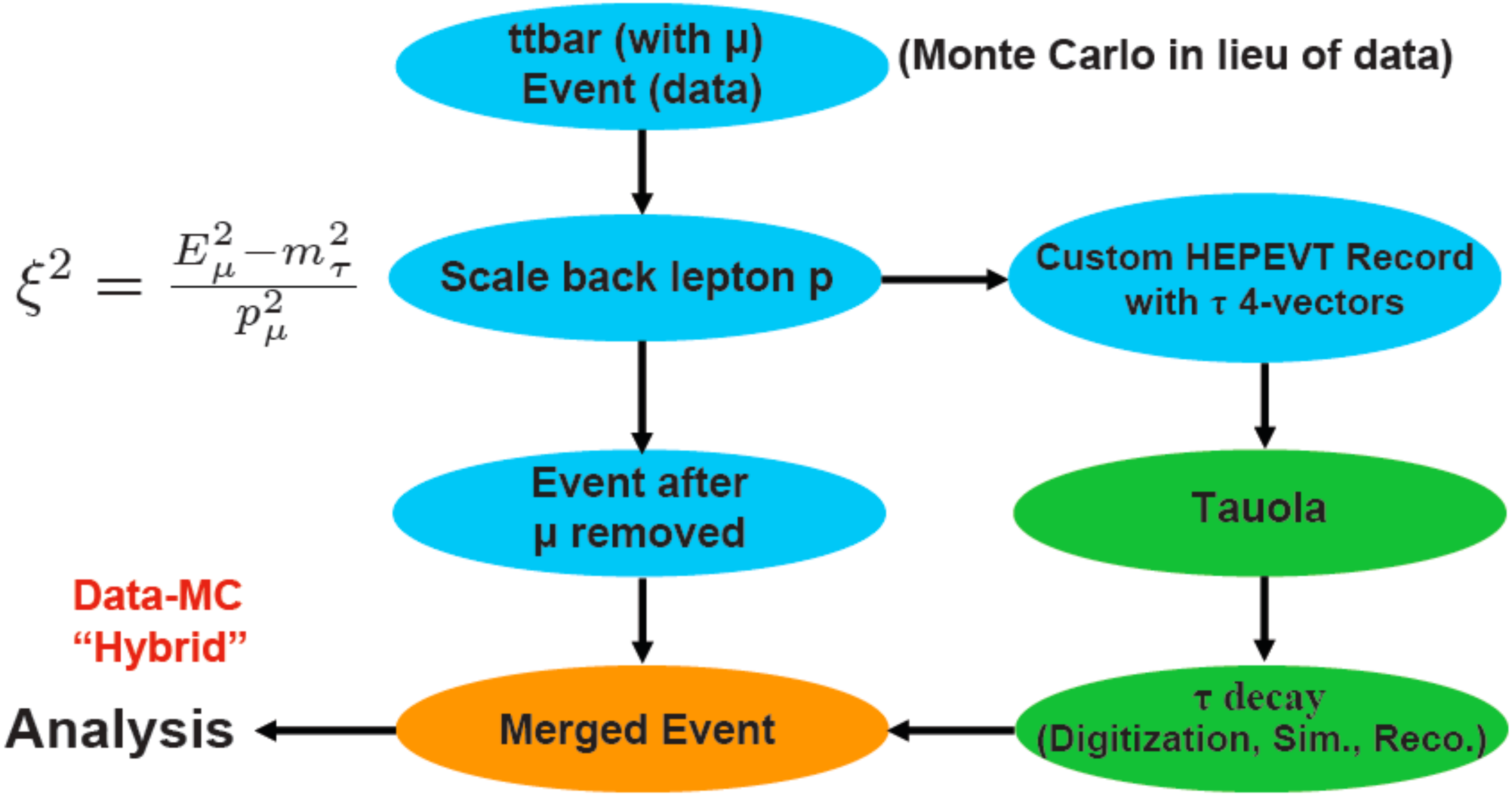
$$H^+ \to \tau_H \nu;\ W \to l\nu$$

Background Control Sample

Change muons into taus using the TAUOLA package



Leptonically- and hadronically-decaying taus can be emulated



ttbar (with μ)
Event (data)

(Monte Carlo in lieu of data)

$$\xi^2 = \frac{E_\mu^2 - m_\tau^2}{p_\mu^2}$$

Scale back lepton p

Custom HEPEVT Record
with τ 4-vectors

Event after
μ removed

Tauola

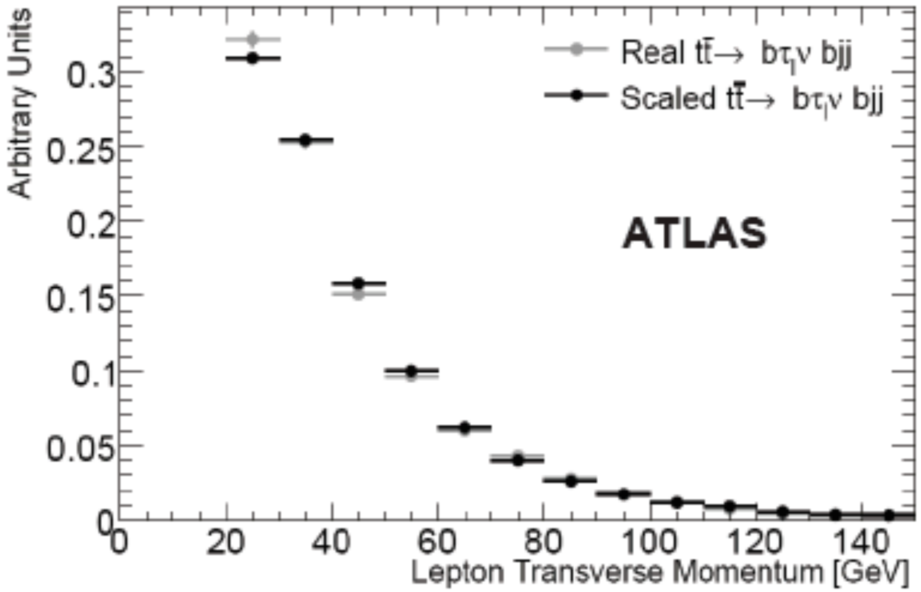
Data-MC
“Hybrid”

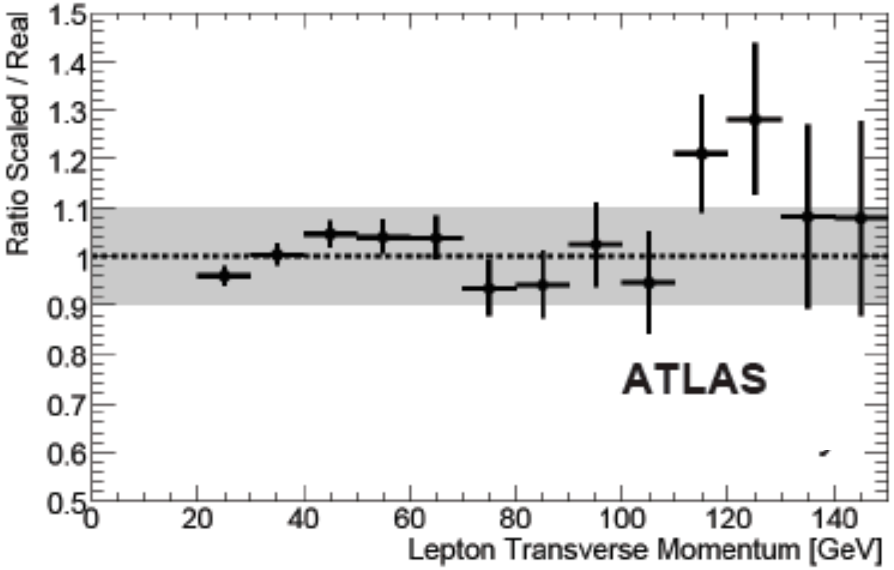
τ decay
(Digitization, Sim., Reco.)

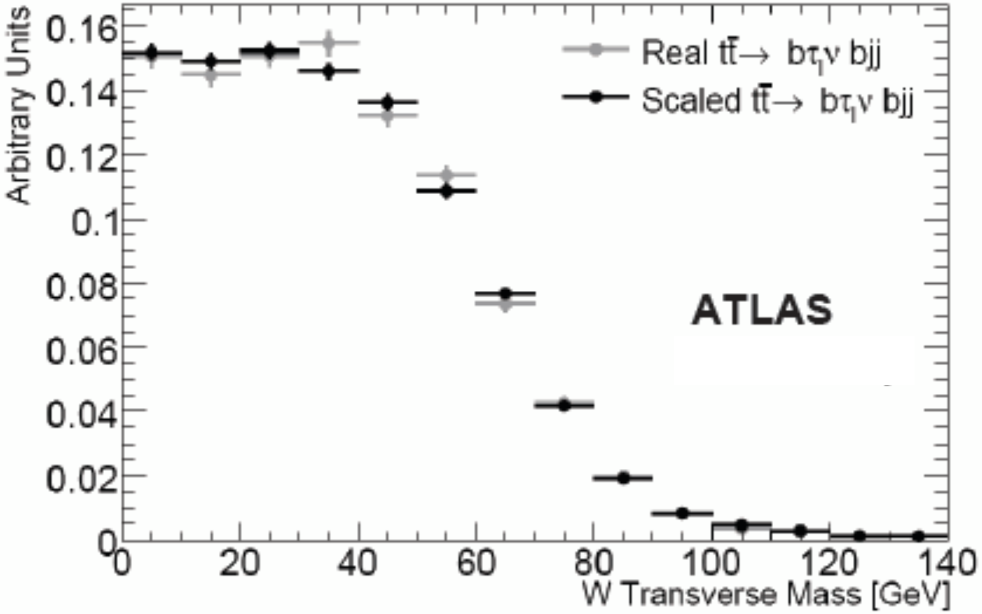
Merged Event

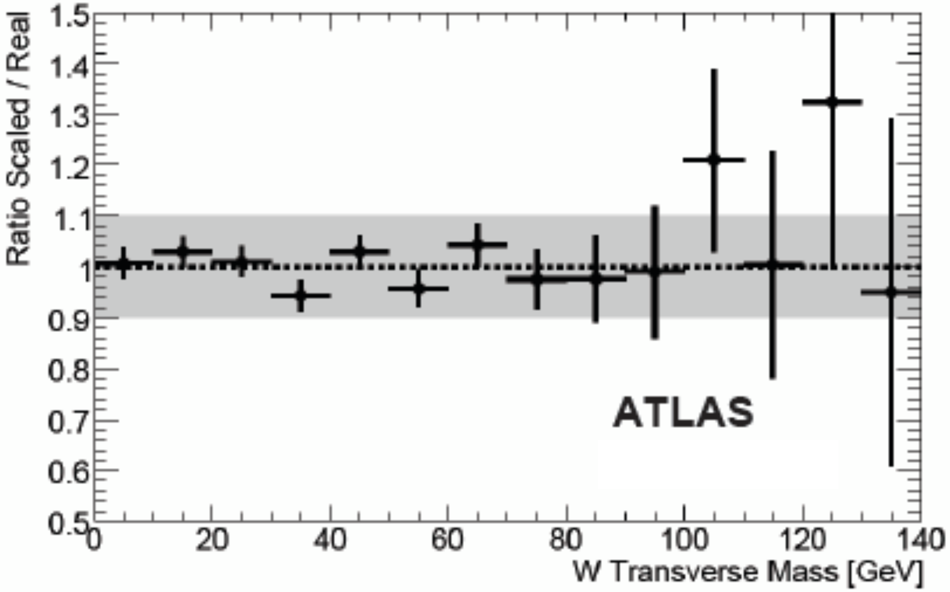
Analysis

$t\bar{t} \rightarrow b\tau_L\nu bqq$

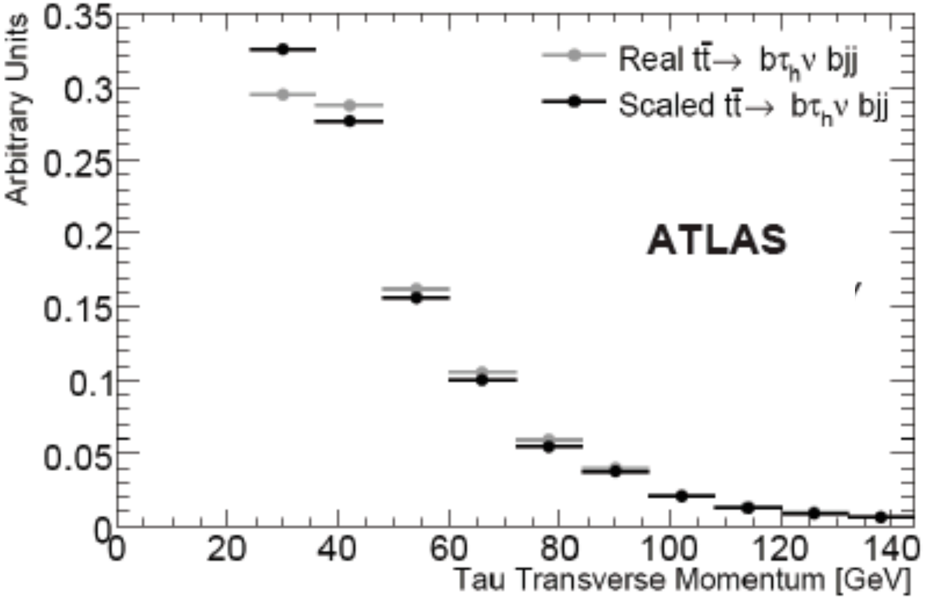


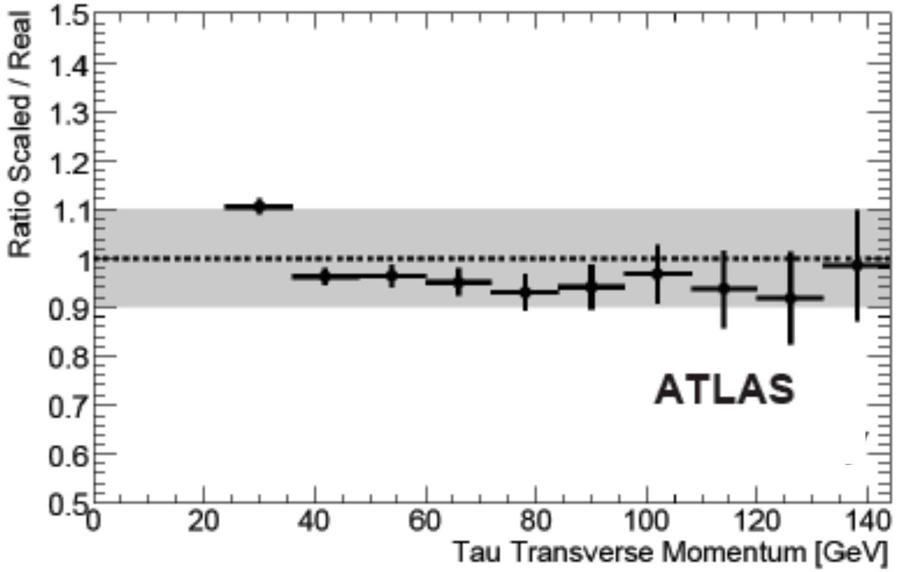


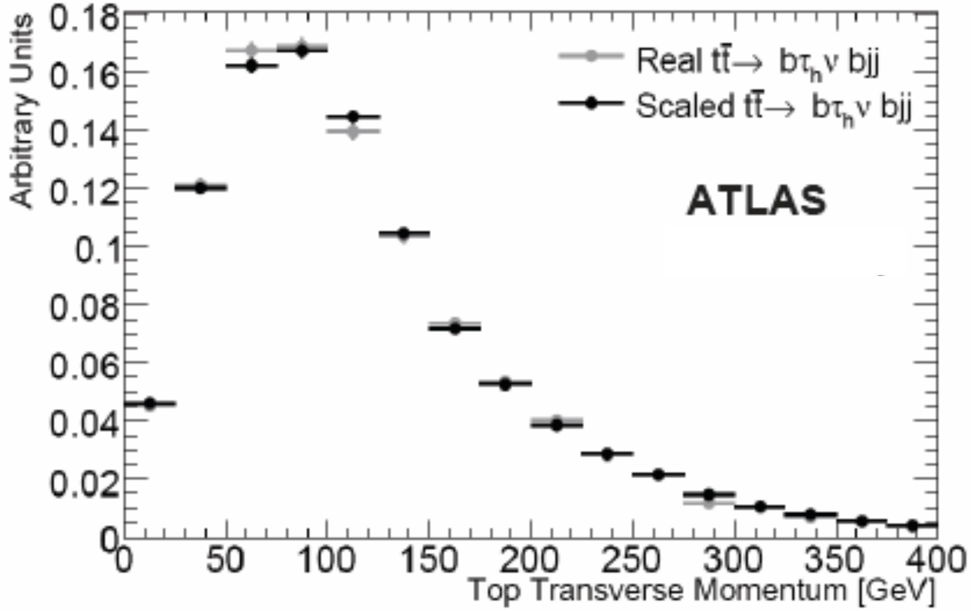


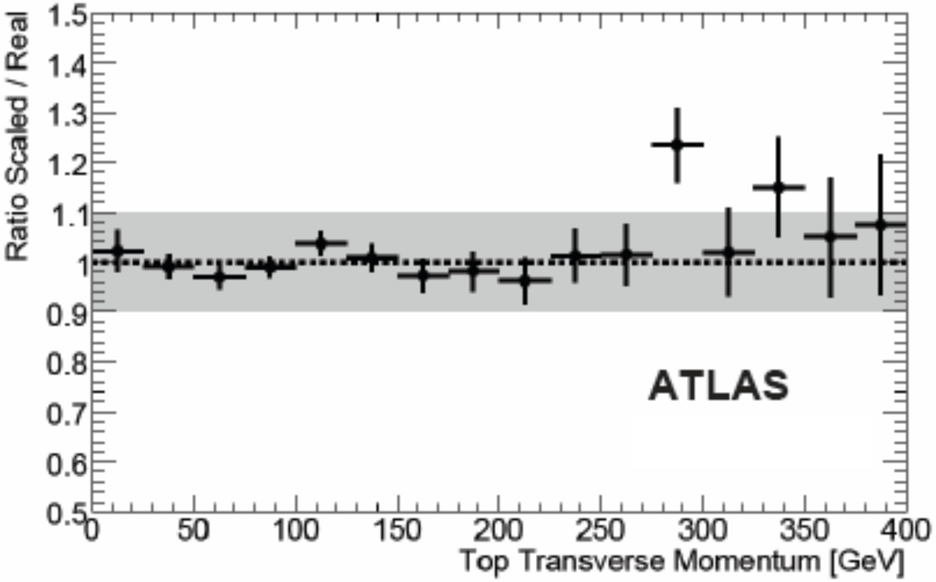


$t\bar{t} \rightarrow b\tau_H \nu bqq$









Outlook and Conclusions

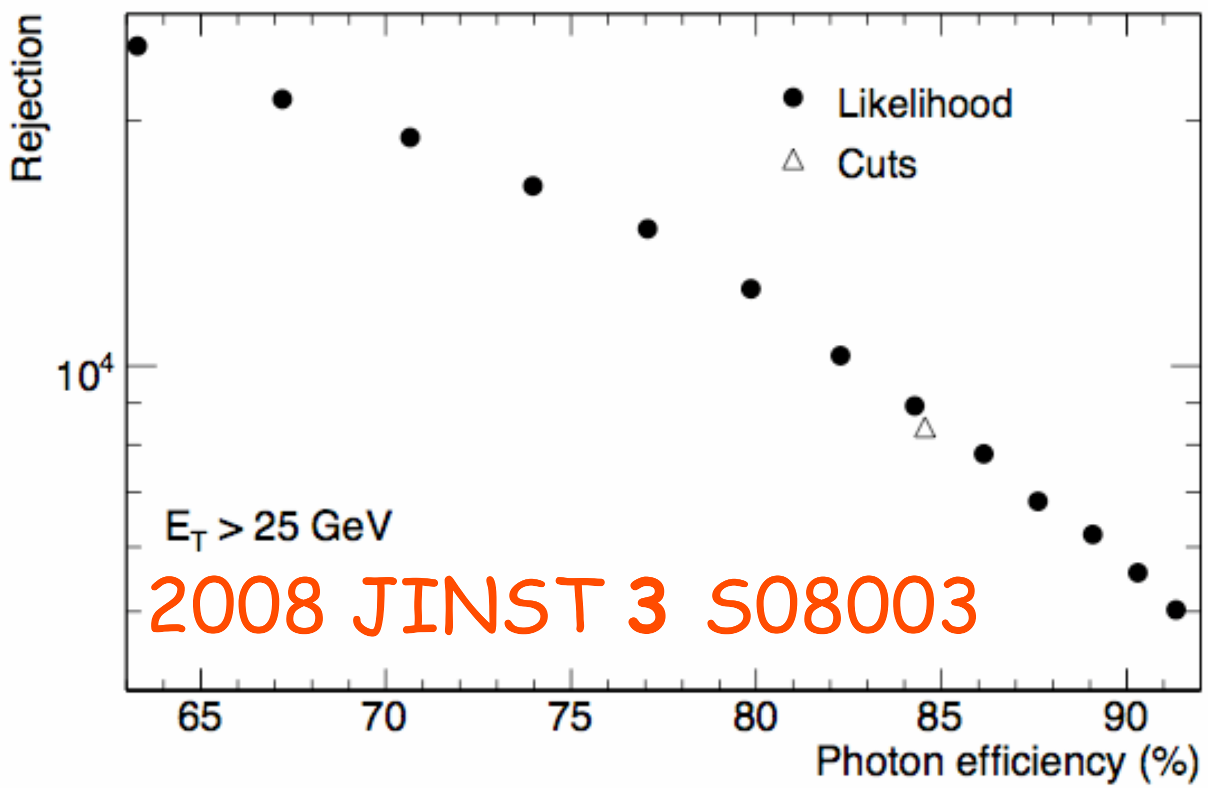
- The search for a Higgs boson is very exciting perspective for CMS and ATLAS. We should be able to make good use of the first data by exercising background extraction techniques with the first data
  - First Higgs cross-sections limits with O(100) pb^{-1}

- Higgs searches at the LHC comprise a large number of final states involving all the signatures that the CMS and ATLAS detectors can reconstruct
  - Electrons, muons, photons, τ, jets, b-jets
  - Need to understand V,VV, (V=Z,W), tt, $\gamma\gamma$, jγ and their production in association with jets

Additional Slides

Photon Identification

- To separate jets from photons is crucial for Higgs discovery
  - Need rejection of > 1000 against quark-initiated jets for ε_γ=80% to keep fake background about 20% of total background
  - Expect rejection against gluon-jets to be 4-5 times greater
- Jet rejection will be evaluated with data
  - Look into sub-leading jets in multi-jet final states with different P_T thresholds
    - Avoid trigger bias
    - Apply trigger pre-scaling if needed
    - Correct for contribution from prompt photons

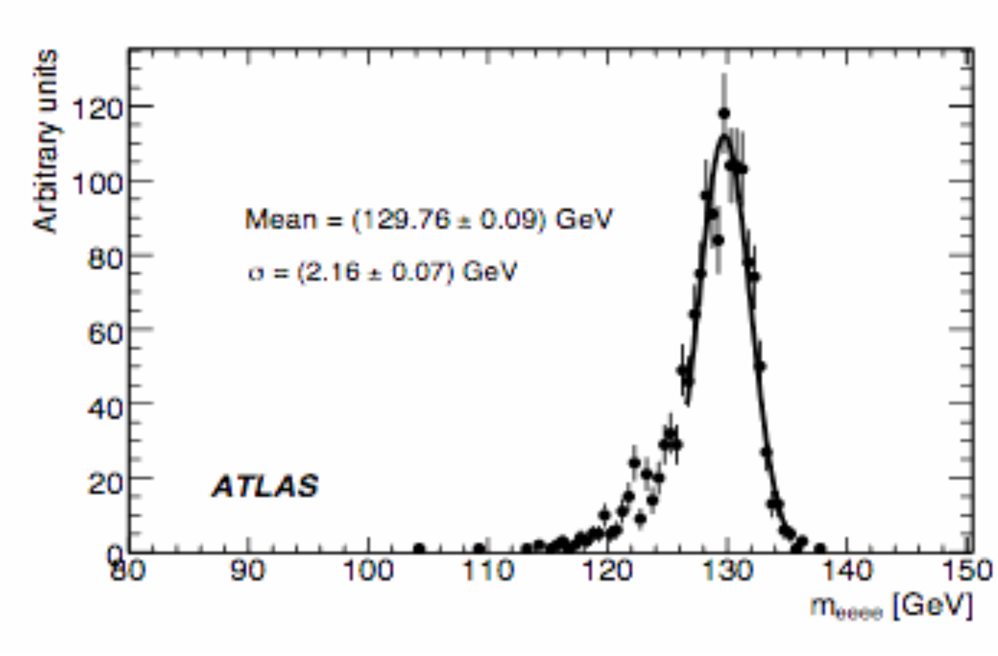


2008 JINST **3** S08003

SM Higgs→ZZ$^{(*)}$→4l

- Able to reconstruct a narrow resonance, with mass resolution close to 1%. Can achieve excellent signal-to-background > 1
  - **Major issue: Lepton ID and rejection of semi-leptonic decays of B decays. Suppress reducible background Zbb,tt→4l**

H[130 GeV]→4e



H[130 GeV]→4μ