

Title: The Standard Model Experiment: Experimental Tools

Date: Jul 16, 2015 11:00 AM

URL: <http://pirsa.org/15070053>

Abstract:



The Standard Model: Experiment (Experimental Techniques)

Pierre Savard
University of Toronto and TRIUMF

TRISEP, Perimeter Institute
15 July 2015

INTRODUCTION

The goal of this lecture is to give you a quick tour of some of the analysis techniques that are used at Hadron Colliders

I've picked a recent ATLAS analysis as an example (Higgs decay to two W bosons). This is a challenging analysis that makes use of many techniques and strategies for background estimations. I encourage you to look at the paper: <http://arxiv.org/pdf/1412.2641.pdf>

The internal documentation for this analysis has over 1000 pages

My thanks to my ATLAS colleague Dr. Corrinne Mills of Edinburgh for contributing many slides in this lecture

INTRODUCTION

In principle, we want the reconstruction of our events to be as efficient as possible. However, we are forced to make selections that reduce the reconstruction efficiency in order to reject backgrounds

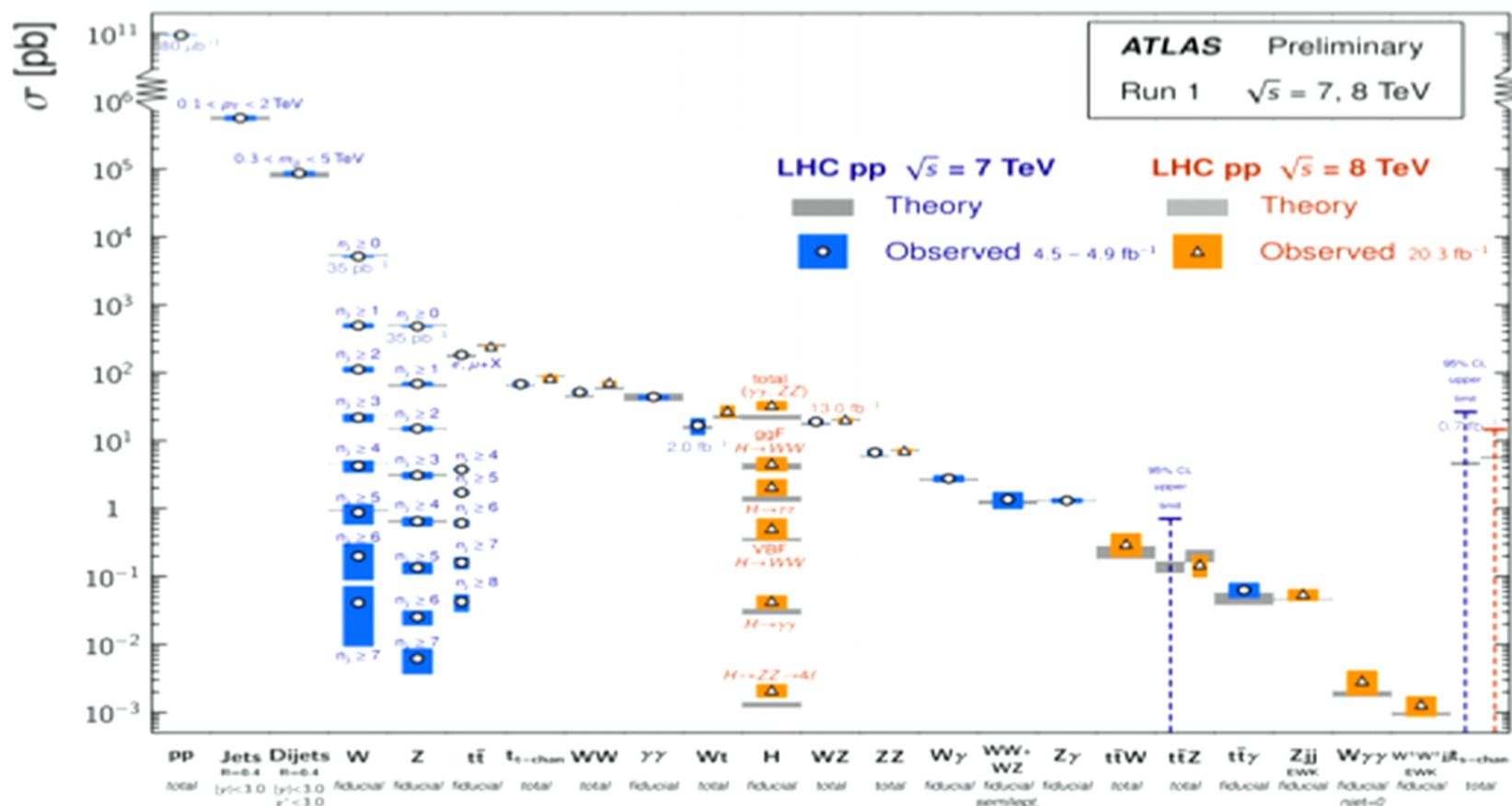
The Higgs \rightarrow WW analysis involves:

- Fake electron rejection
- Fake muon rejection
- Jet reconstruction
- Pileup mitigation for jets
- B-tagging (and vetoing)
- Jet veto
- Missing Et reconstruction in a high pileup environment
- Event categorization
- Topological categorization
- Use of control and validation regions to normalize backgrounds and check background modeling
- Use of multivariate techniques
- Etc.

ATLAS CROSS SECTION MEASUREMENTS

Standard Model Production Cross Section Measurements

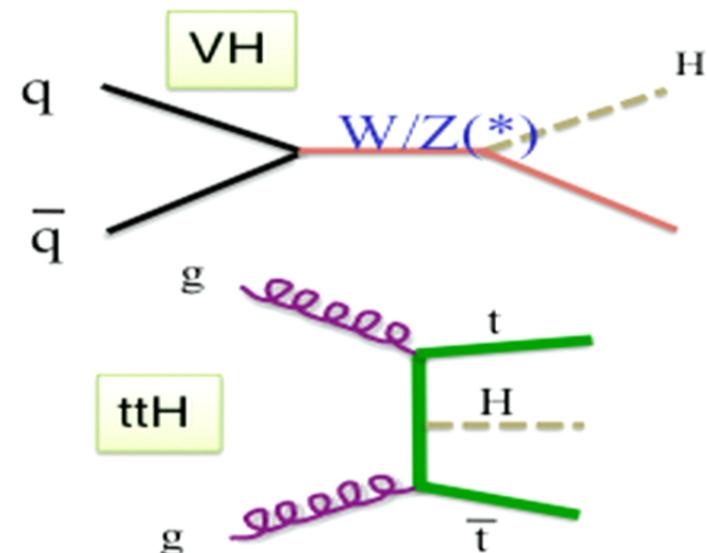
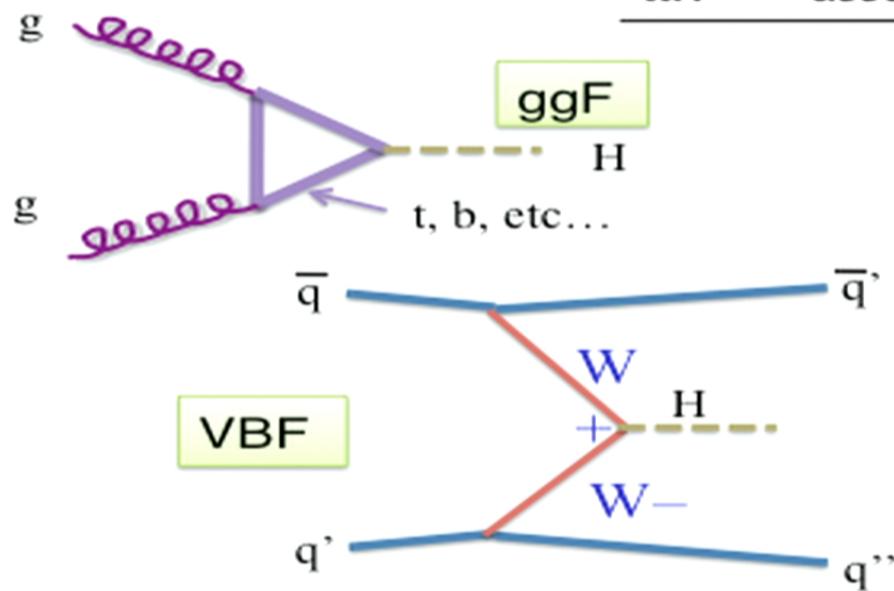
Status: March 2015



Higgs Production

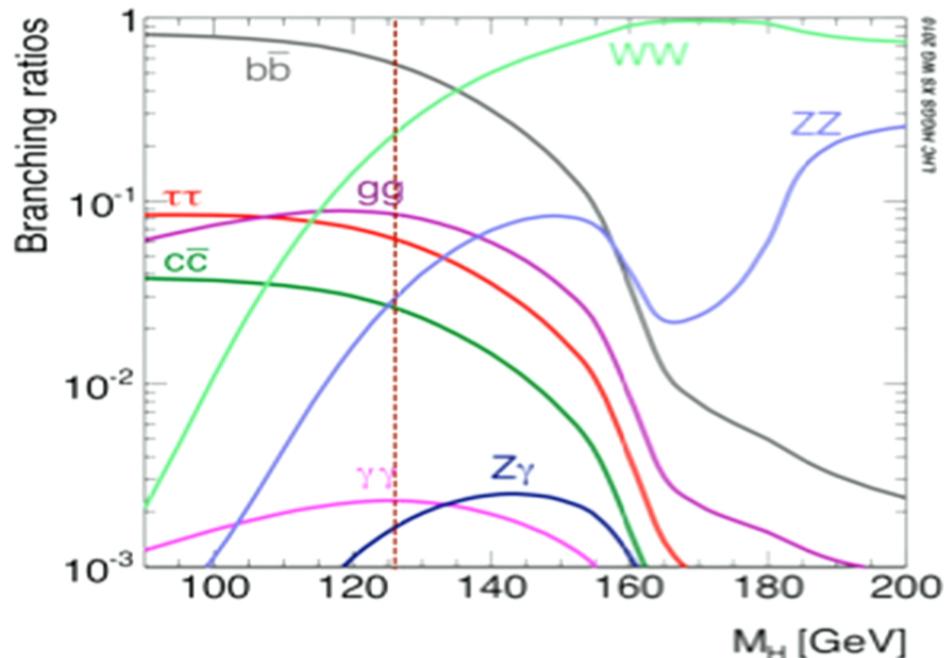
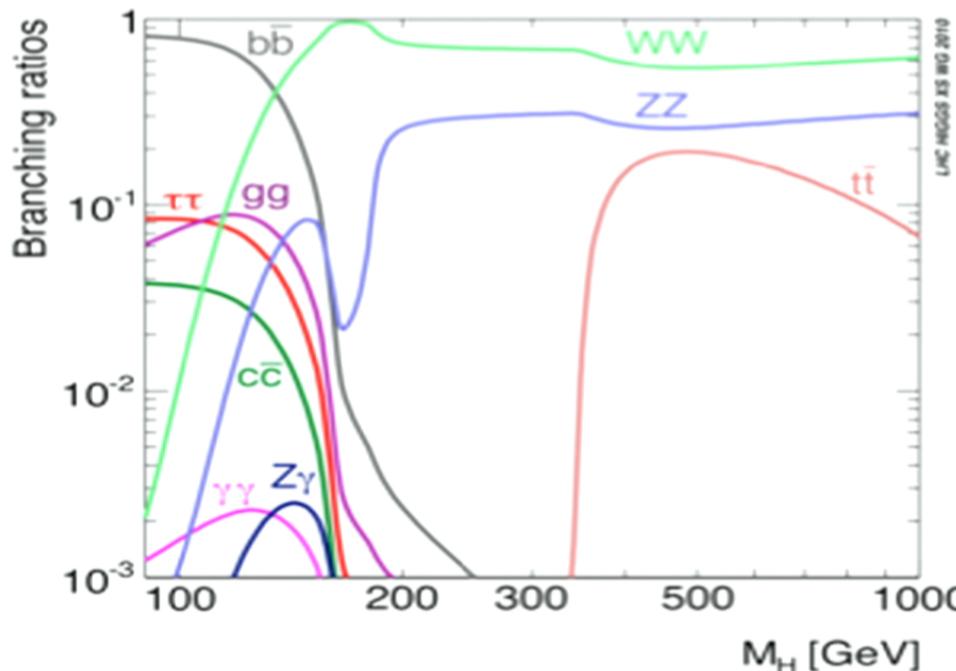
Cross sections for
 $m_H = 125 \text{ GeV}$:

	process	8 TeV	13 TeV
ggF	gluon-gluon fusion	19 pb	44 pb
VBF	vector-boson fusion	1.6 pb	3.7 pb
VH	associated production	1.1 pb	2.2 pb
ttH	associated production	0.13 pb	0.51 pb

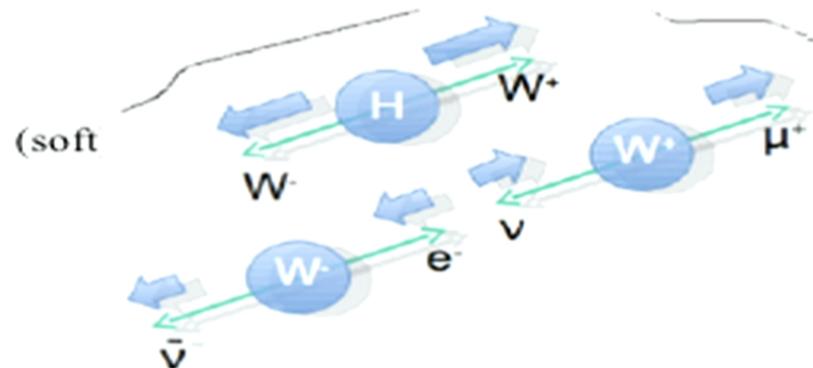
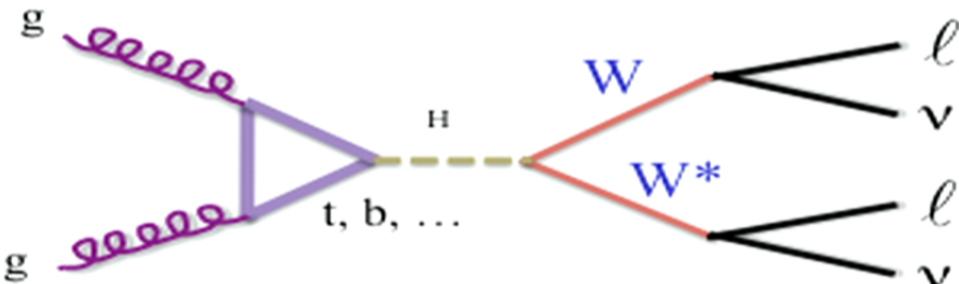


Higgs Decays

- At a mass of 125 GeV, we can observe many decay modes: Higgs measurements require that we fully exploit the detector capabilities ($e, \mu, \tau, \gamma, etmiss, jets, tag HF-jets, trigger$)



$$H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$$



- Large Br to WW :
 - many signal events
 - But final state features low p_T lepton and neutrinos
- Can't fully reconstruct final state because of neutrinos
 - Missing E_T reconstruction is important (and challenging in presence of pileup)

- Exploit spin 0 kinematics
- Use transverse mass as main discriminating variable

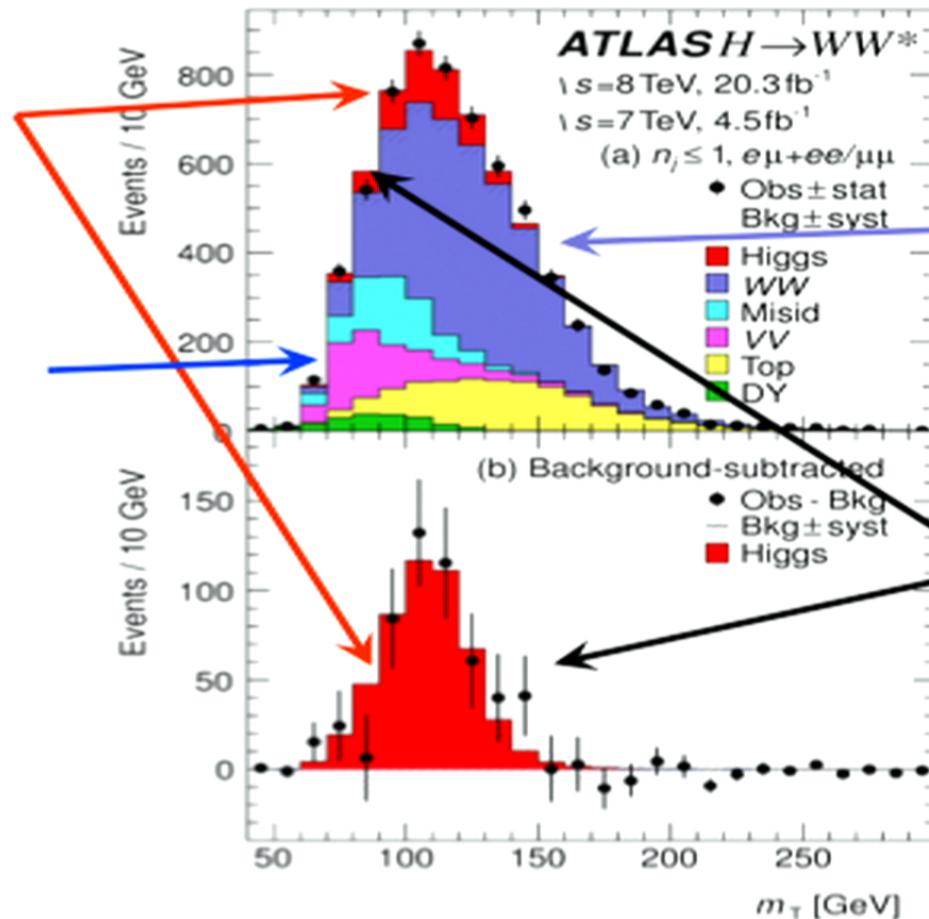
$$M_T^2 = (E_T^{\ell\ell} + E_T^{\text{miss}})^2 - (\vec{p}_T^{\ell\ell} + \vec{E}_T^{\text{miss}})^2$$

$$(E_T^{\ell\ell})^2 = (\vec{p}_T^{\ell\ell})^2 + (m_{\ell\ell})^2$$

The $H \rightarrow WW$ dataset

peaked signal

other backgrounds



dominant non-resonant WW diboson background

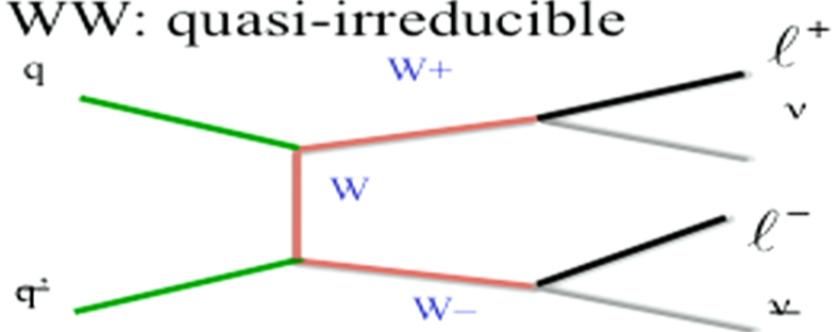
2011+2012 data

Fit m_T distribution to extract signal yield
How do we get to this point?

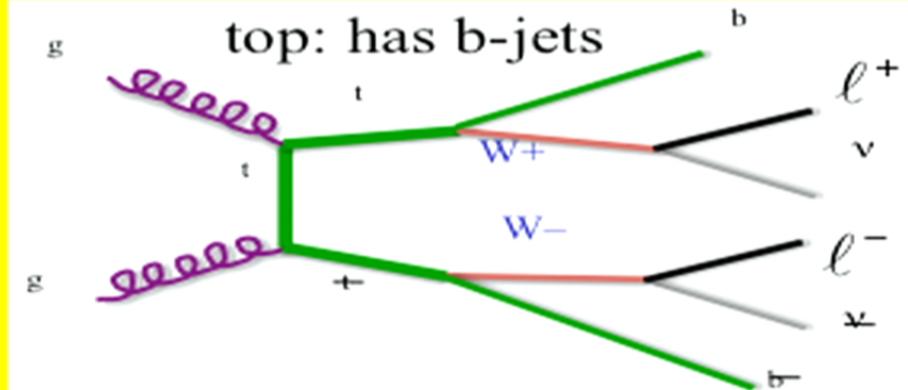
$$H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$$

Backgrounds:

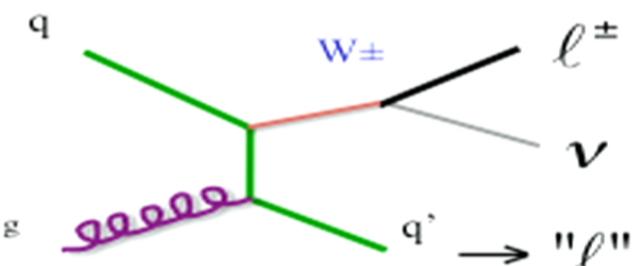
WW: quasi-irreducible



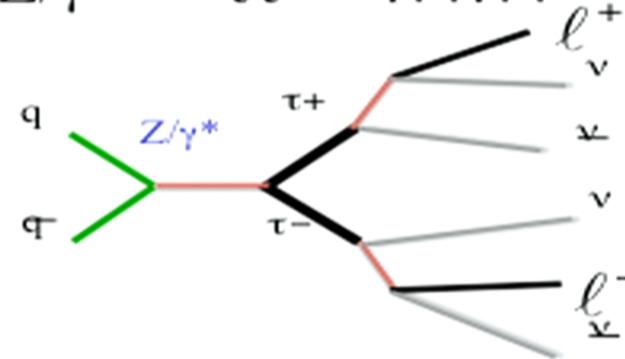
top: has b-jets



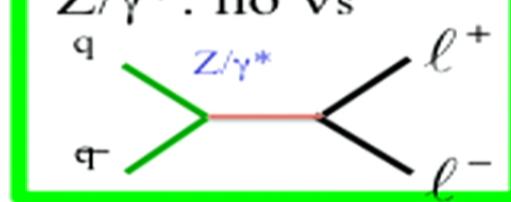
W+jets: fake leptons



$Z/\gamma^* \rightarrow \tau\tau \rightarrow |\nu\nu|\nu\nu$



Z/γ^* : no ν s

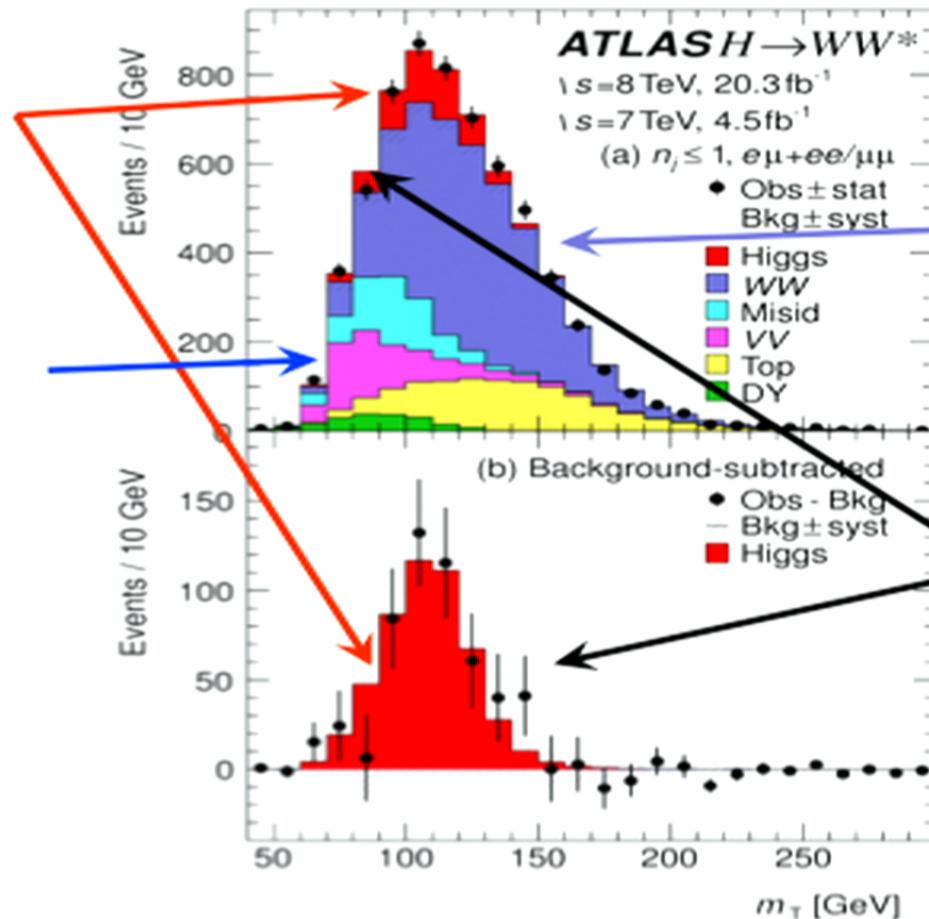


+ WZ , $W\gamma^*$, $W\gamma$

The $H \rightarrow WW$ dataset

peaked signal

other backgrounds



dominant non-resonant WW diboson background

2011+2012 data

Fit m_T distribution to extract signal yield
How do we get to this point?

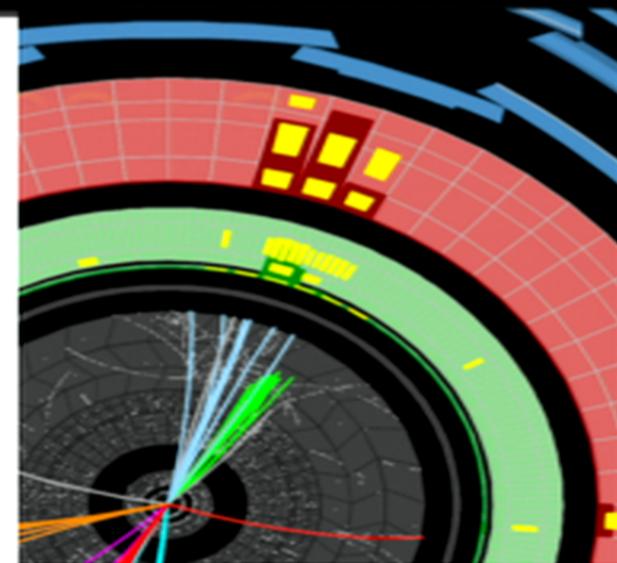
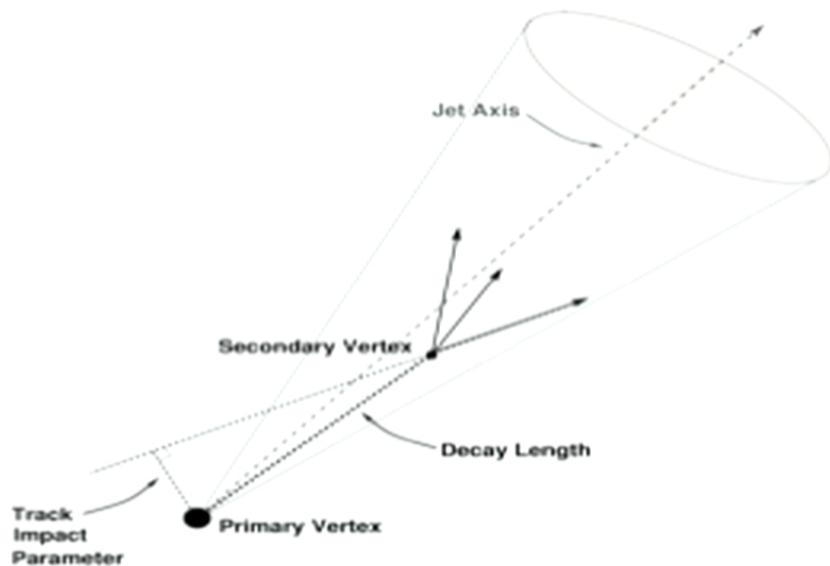
Muons and Electrons

- **Muons:** combined track from inner detector and muon spectrometer
 - *Track segments need to match*
 - *Muons need to be isolated (in calorimeter and tracker)*
 - *Small energy deposited in the calorimeter along the track*
 - *Impact parameter should be small*
- **Electrons:** cluster in EM calorimeter combined with inner detector track
 - *Momentum and energy should match*
 - *Isolation in calorimeter and tracker*
 - *Shower shape consistent with that of an electron*
 - *Impact parameter should be small*
- **Fake backgrounds drive the selections:**
 - *Tension between efficiency and suppression of background from fake and non-prompt leptons*
 - *W+jets vs H(WW):*
 - $\sigma(pp \rightarrow H \rightarrow WW \rightarrow l\nu l\nu) = 0.23 \text{ pb}$ vs $\sigma(pp \rightarrow W \rightarrow l\nu) = 12,000 \text{ pb}$

Jets and B-Tagging

Final state quarks and gluons **fragment** to produce additional quarks and gluons and **hadronize** into colorless particles

- Identified as a collimated “jet” of particles in tracker and calorimeter



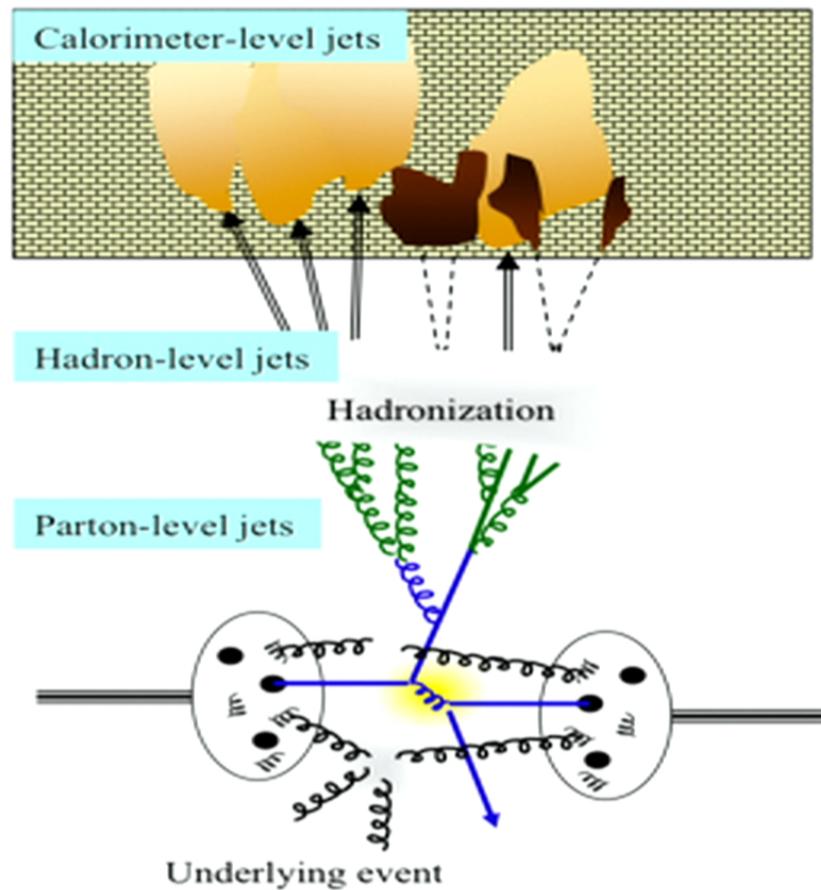
Hadrons containing heavy quarks (b, c) travel some distance in the detector and can be identified by the resulting displaced tracks and decay vertices

- “tagging” of b-jets: we use 85% efficient operating point

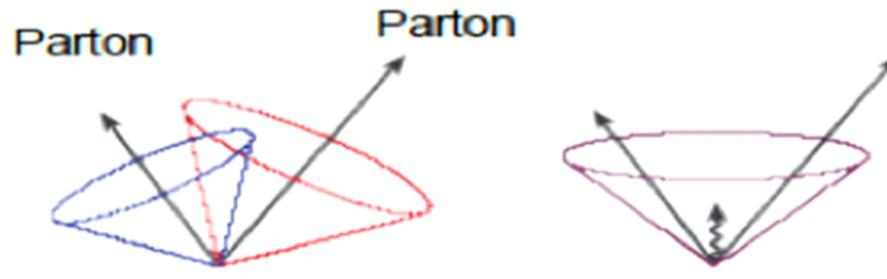
Jets

Some jet algorithm requirements:

- Theoretical considerations
 - *Infrared safe*
 - *Collinear safe*
 - *Invariant under boosts*
 - *Easy to compare with theory*
- Experimental considerations
 - *Detector independence*
 - *Easy to calibrate*
 - *Robust under pileup/underlying event*
 - *High reconstruction efficiency*

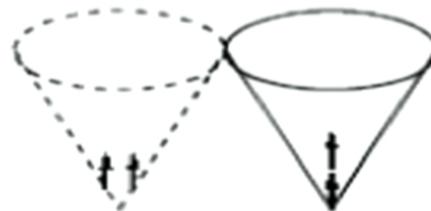


Jets



infrared sensitivity

energy split



collinear sensitivity

K_T Algorithm

Algorithm description :

- Define a distance

$$D_{ij} = \min(P_{Ti}^2, P_{Tj}^2) \frac{\Delta R_{ij}^2}{R^2} \quad D_i = P_{Ti}^2$$

- Compute all $\{D_{ij}, D_i\}$ and $d = \min(\{D_{ij}, D_i\})$
 - if $d = D_{ij}$: combine jet i with jet j
 - if $d = D_i$: define jet i as a final jet
- Exhaust all proto-jets

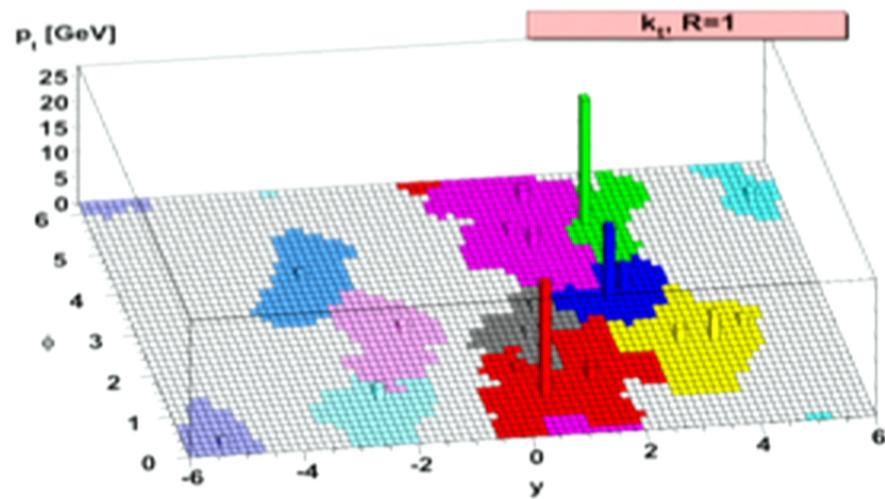
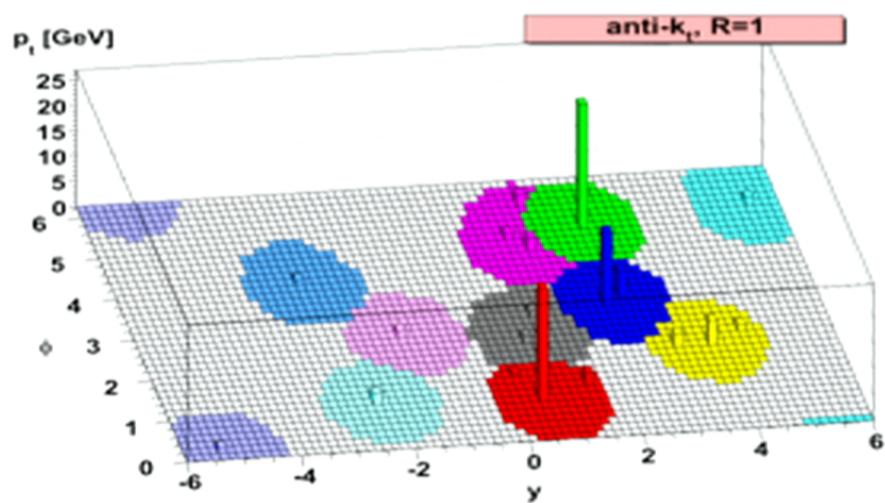
Variants : Anti-kt and Cambridge. In distance formula replace P_T^2 \longrightarrow P_T^{2p}

p = 1 : K_T

p = 0 : Cambridge

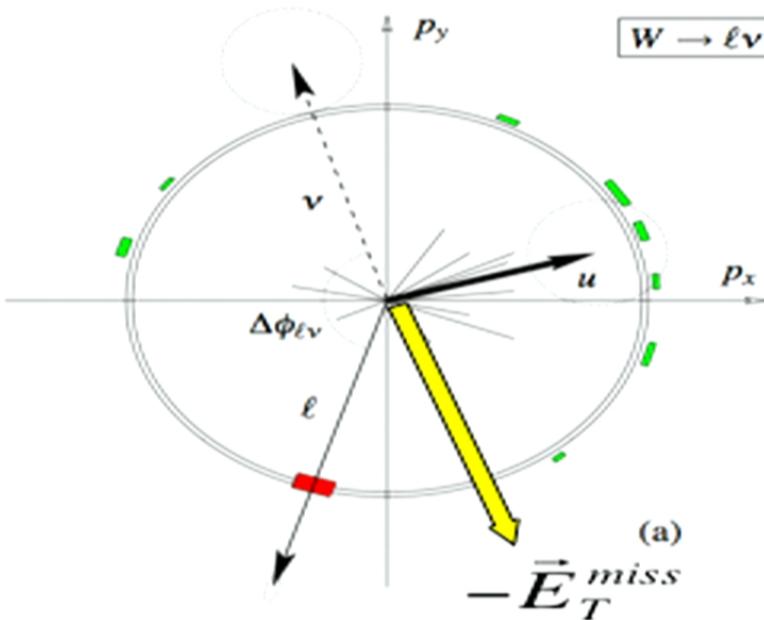
p = -1: Anti-K_T

Jet Algorithms



Missing Transverse Energy

$$\vec{E}_T^{miss} = - \sum_{\text{objects } i} \vec{p}_T^{\text{object}}$$



Detect neutrinos by summing up all of the other objects in the event

fully reconstructed objects (leptons, jets, photons) and “soft” particles

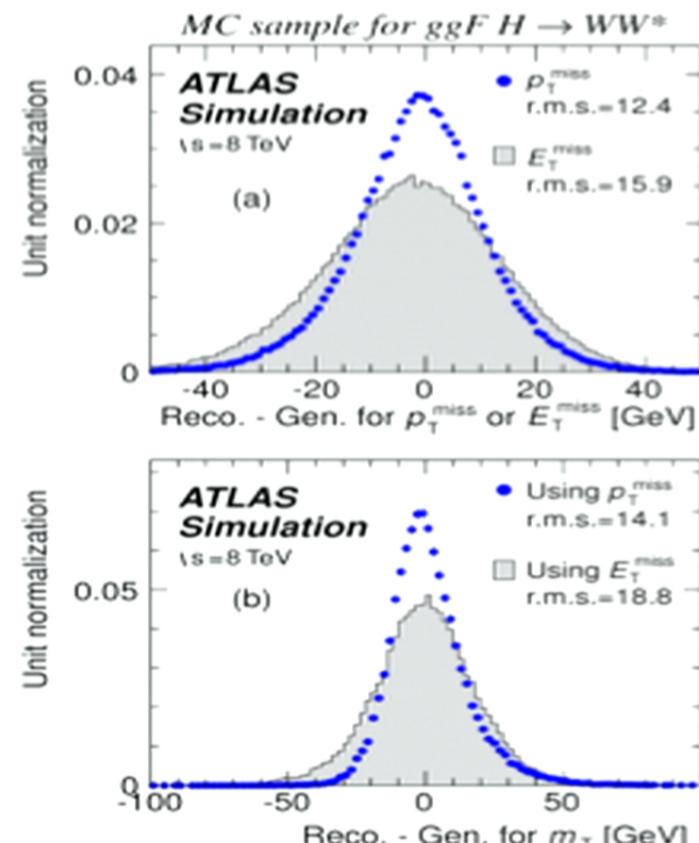
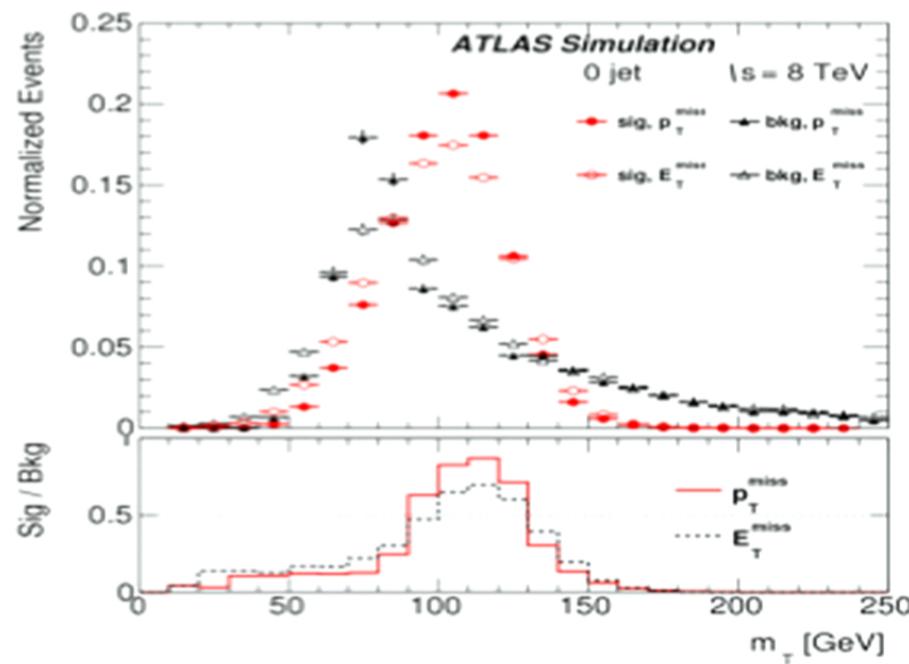
New, track-based measurement of “soft” particles improves transverse mass resolution

Invoke conservation of momentum in the plane transverse to the beam:

“Missing Transverse Energy” corresponds to the transverse energy of the neutrinos in the event

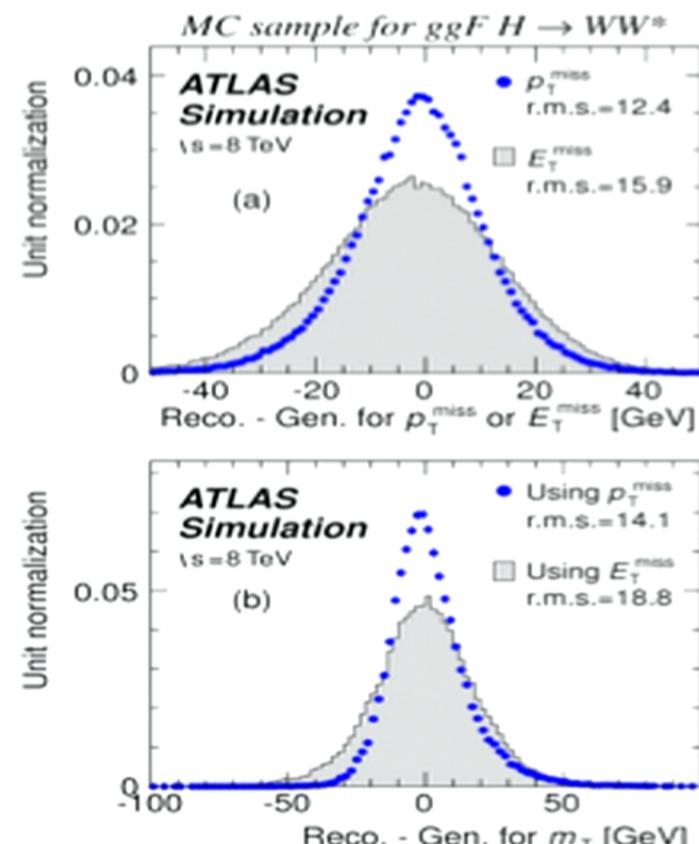
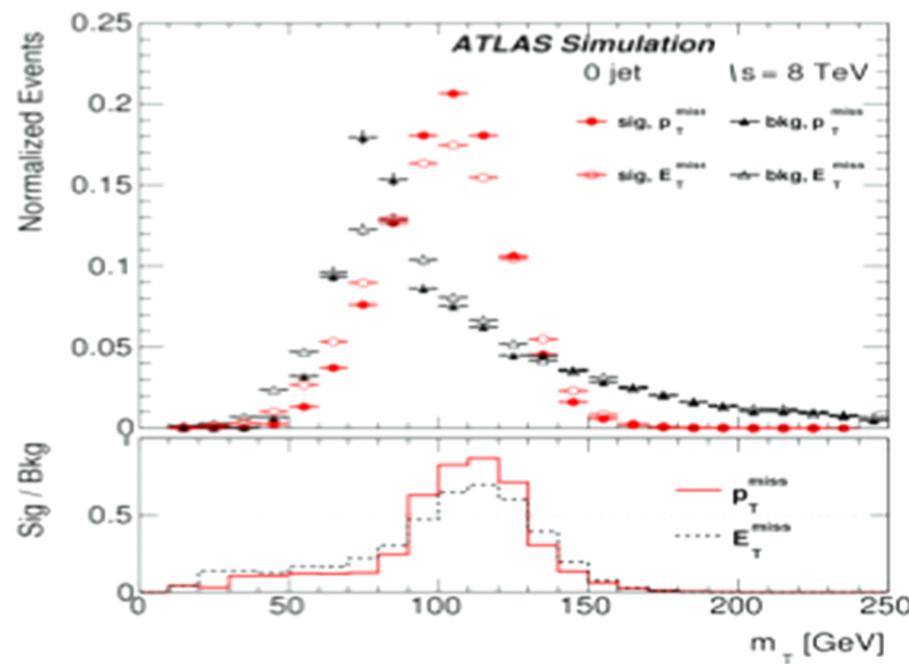
Track p_T^{miss}

- ETmiss from resolved objects (leptons, jets, photons) and “soft term” of not-otherwise-identified objects, usually hadrons
- Soft term from tracks gives better resolution than hadron clusters for high pileup



Track p_T^{miss}

- ETmiss from resolved objects (leptons, jets, photons) and “soft term” of not-otherwise-identified objects, usually hadrons
- Soft term from tracks gives better resolution than hadron clusters for high pileup

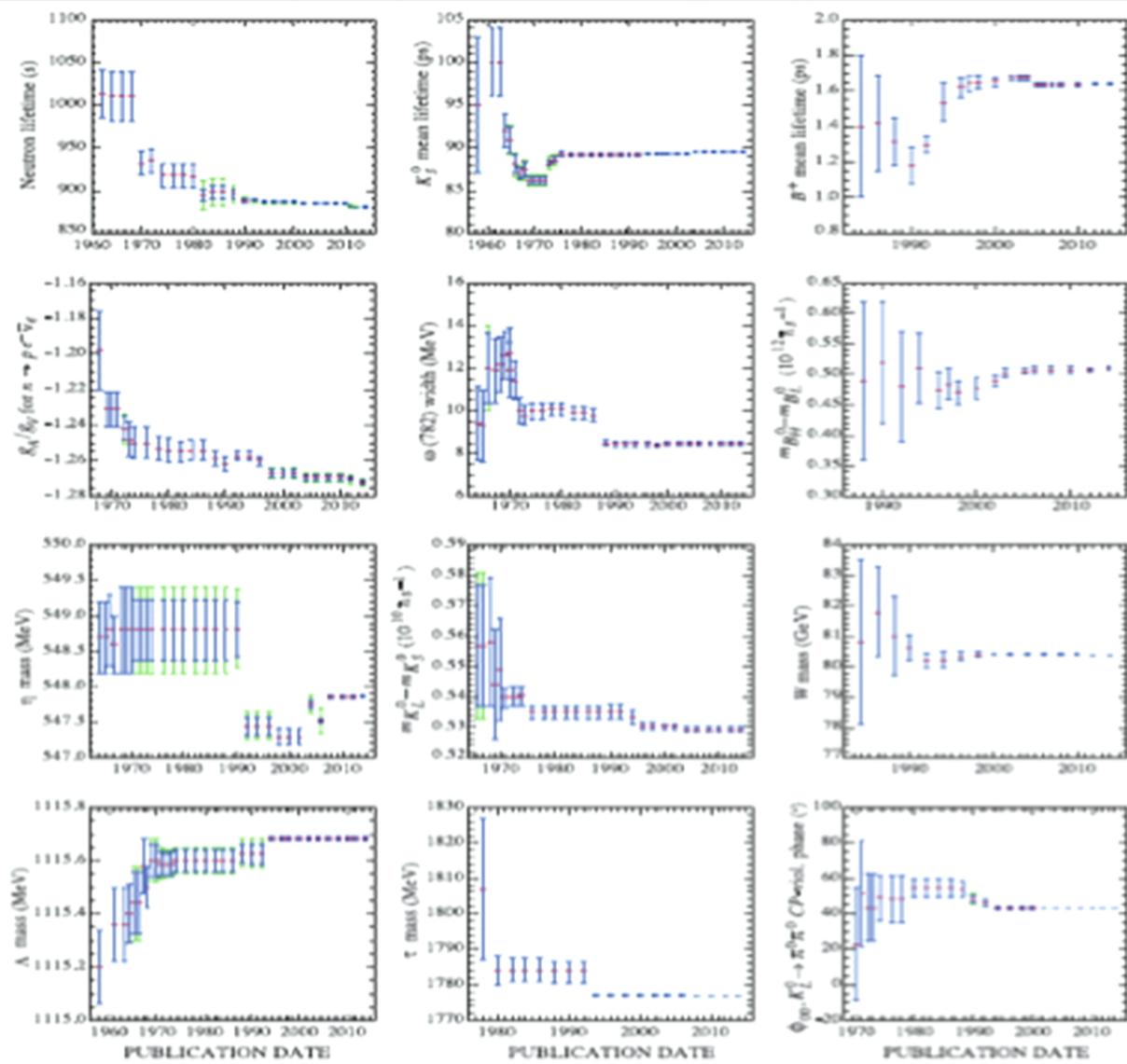


Most of our analyses are “blinded”

→ We do not look at the signal rich region until we demonstrate that we understand our backgrounds around the signal region

Why do we blind?

→ From the PDG:



Analysis Blinding

- Blinding requirements for WW analysis:
 - Look at regions where $S/B < 2\%$
- Not possible to blind WW analysis for all mH
 - Judgment call during search phase for the Higgs: what we really cared about was the low mH signal region
- Definition of the signal region:
 - Use $\Delta\varphi(ll)$ and $m(ll)$ cuts
 - Transverse mass bound corresponding to lower bound for 110 and upper bound for 140 → veto $(0.75)(110) < mT < (1.0)(140)$

Blinded Region

$82.5 < M_T < 140$

and

$\Delta\varphi(ll) < 1.8$

and

$m_{ll} < 50$

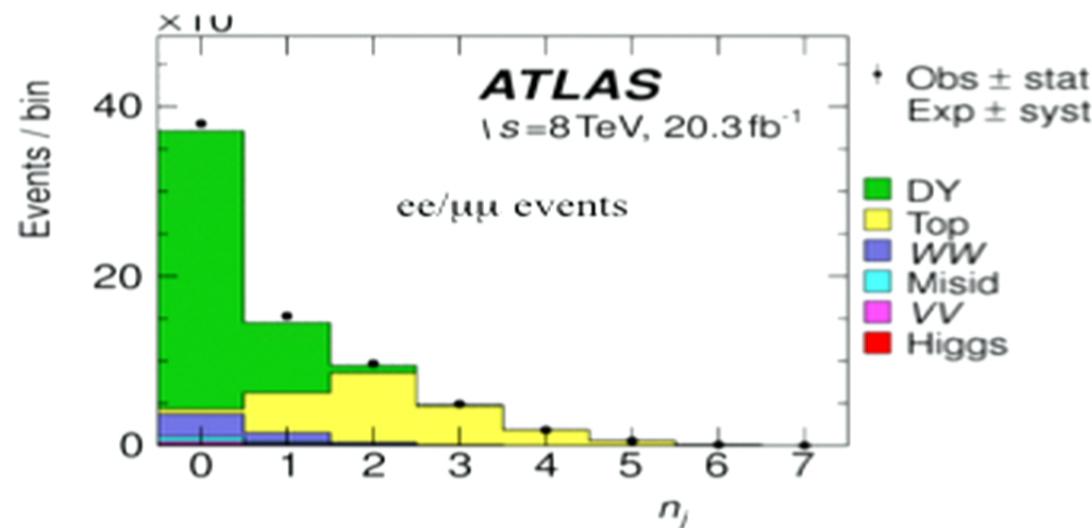
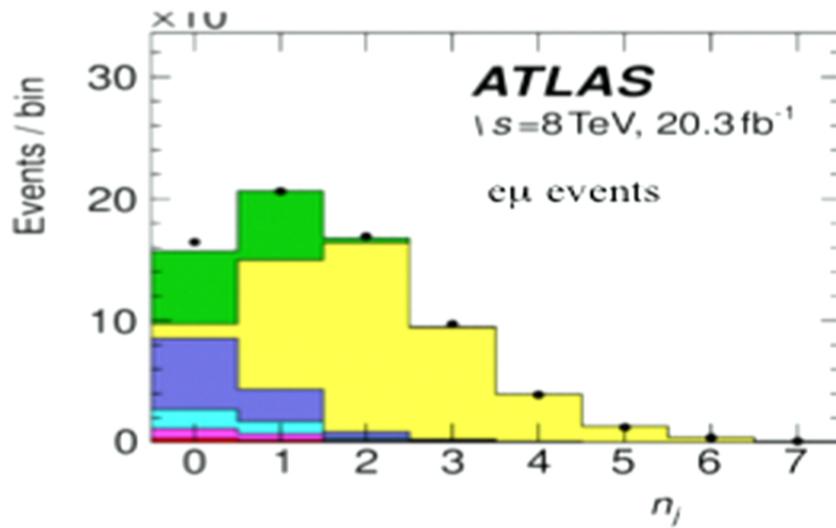
and

0 jets or 0 b-tags

Event Categorization

Basic selection: two charged leptons (e or μ) and missing transverse energy*:

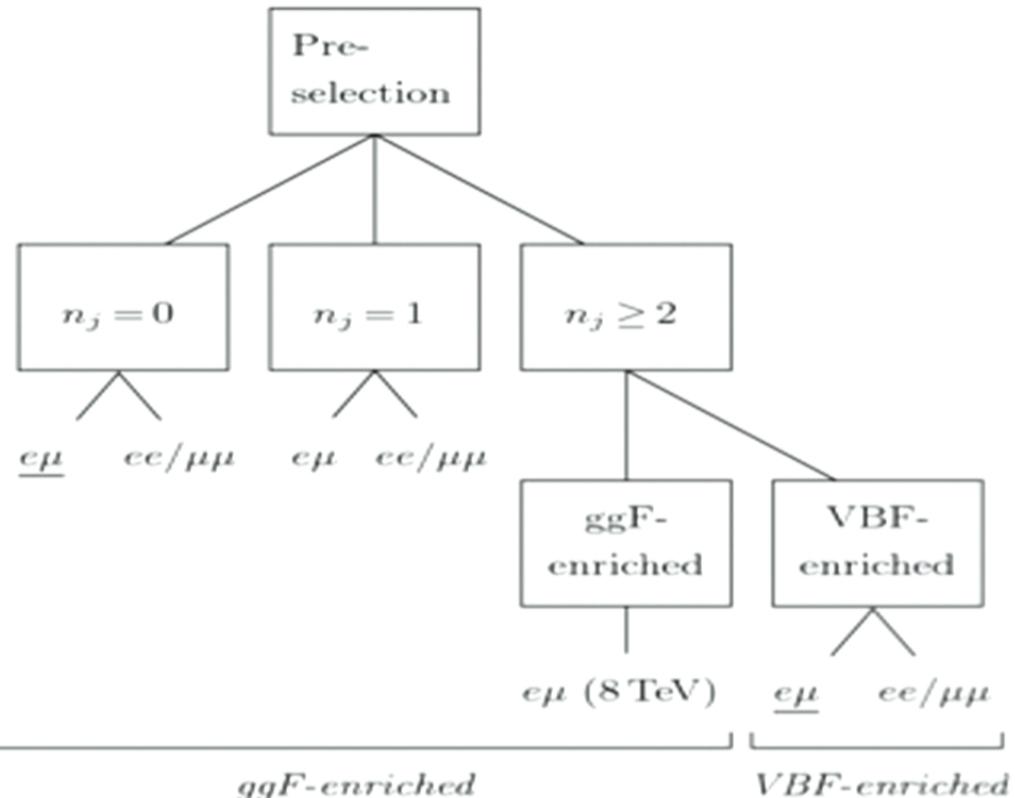
- Analysis strategy: categorize events by dominant signal, background
- Number of jets: separate top from ggF signal, VBF from ggF
- Lepton pair flavor: $ee/\mu\mu$ has challenging Z/γ^* background



$$H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$$

Selections (snapshot):

- Leading lepton $p_T > 22$
- Subleading lepton $p_T > 10$
- Sample divided in number of jets above 25 GeV
- Missing E_T uses tracking and calorimeter, cut depends on final state
- Topological selections include cuts on m_{\parallel} , $p_{T\parallel}$, $\Delta\phi_{\parallel}$
- Signal extracted from fit to transverse mass in bins of subleading lepton p_T and m_{\parallel}

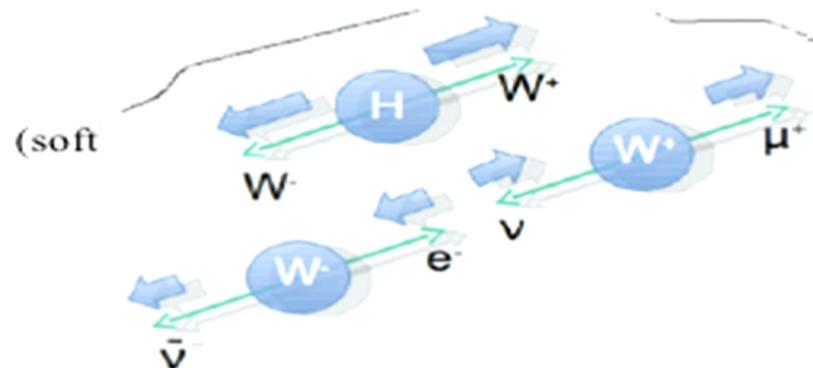
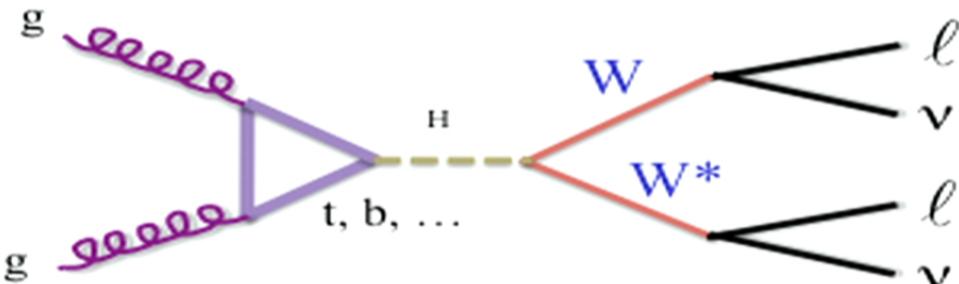


14 January 2015

c. Mills (Edinburgh)

22

$$H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$$



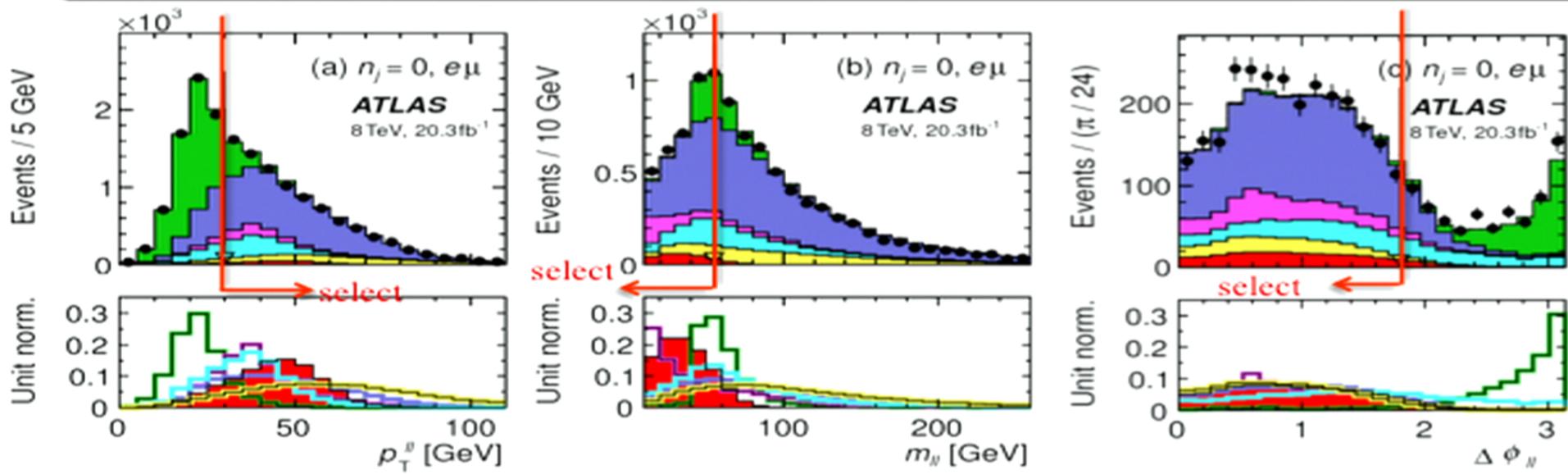
- Large Br to WW :
 - many signal events
 - But final state features low p_T lepton and neutrinos
- Can't fully reconstruct final state because of neutrinos
 - Missing E_T reconstruction is important (and challenging in presence of pileup)

- Exploit spin 0 kinematics
- Use transverse mass as main discriminating variable

$$M_T^2 = (E_T^{\ell\ell} + E_T^{\text{miss}})^2 - (\vec{p}_T^{\ell\ell} + \vec{E}_T^{\text{miss}})^2$$

$$(E_T^{\ell\ell})^2 = (\vec{p}_T^{\ell\ell})^2 + (m_{\ell\ell})^2$$

Topological Selections

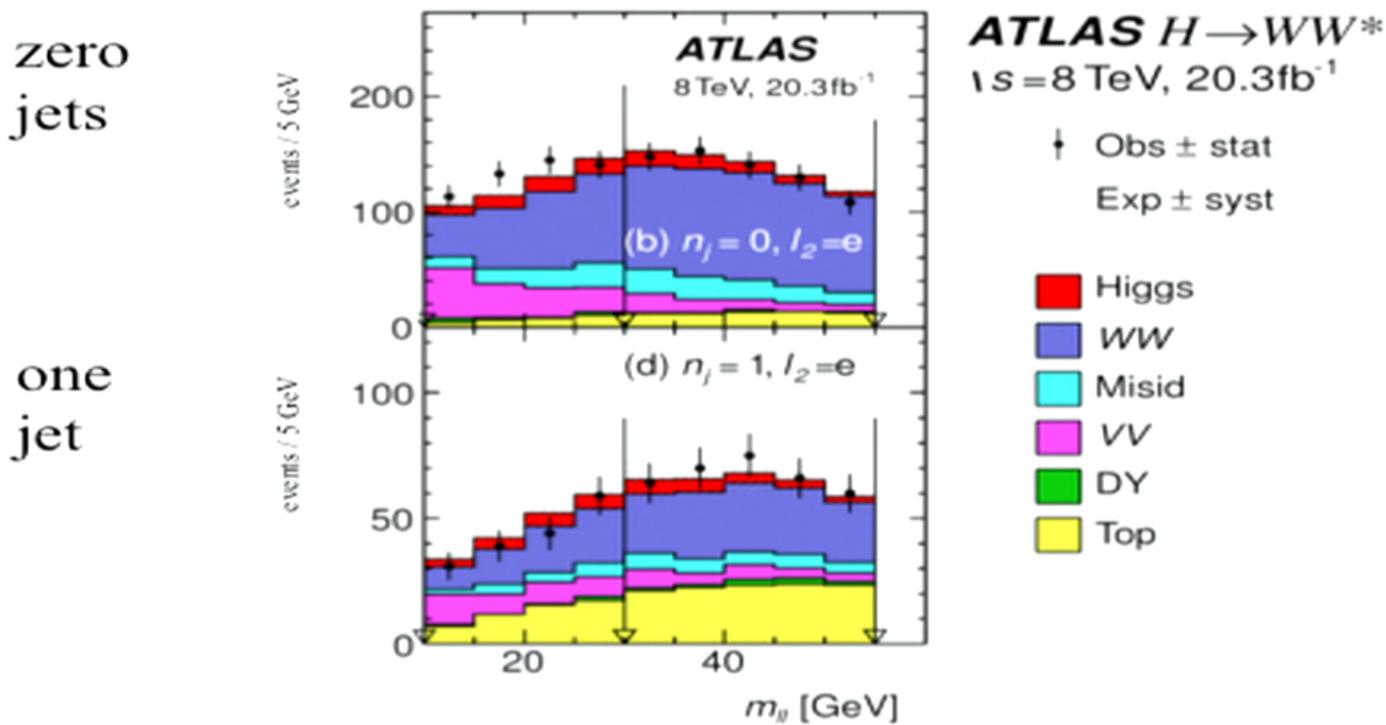


- Reject dominant BG like WW and $Z \rightarrow \tau\tau$ by using the spin correlations and lepton kinematics
 - Leptons tend to not be back-to-back, and one lepton is soft because of the off-shell W ($m_H < 2m_W$)
- Importance of $WZ/W\gamma/W\gamma^*$ and $W+jets$ grows



Gluon Fusion Categorization

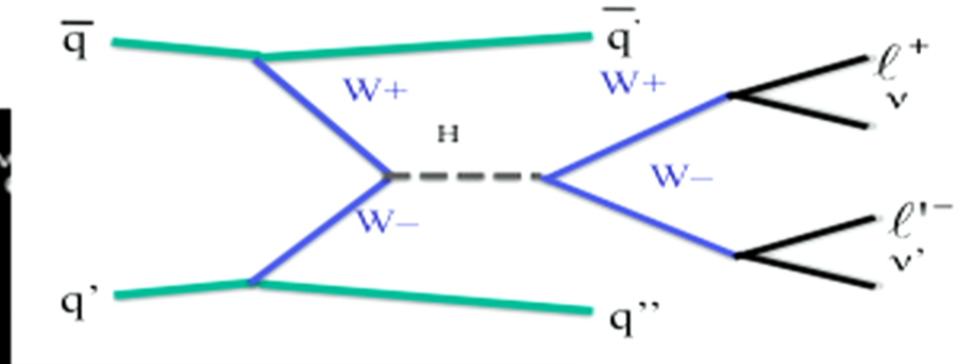
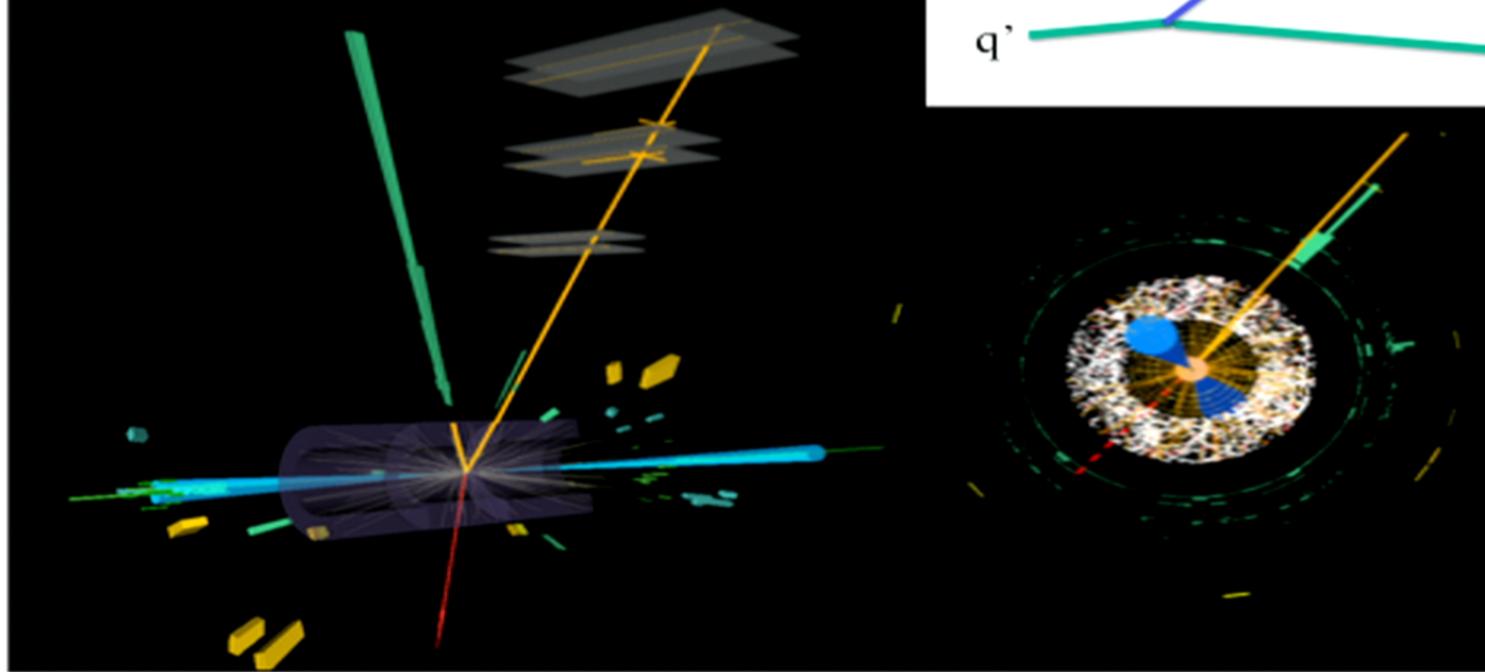
- Dilepton invariant mass m_{\parallel} separates sources of background
- Zero jet events have minimal top quark background



VBF Topology



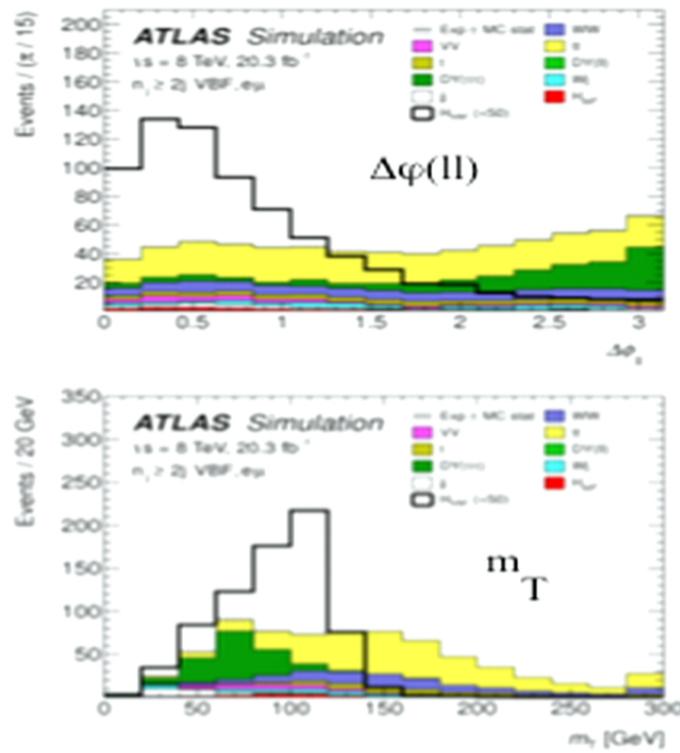
Run 214680, Event
17 Nov 2012



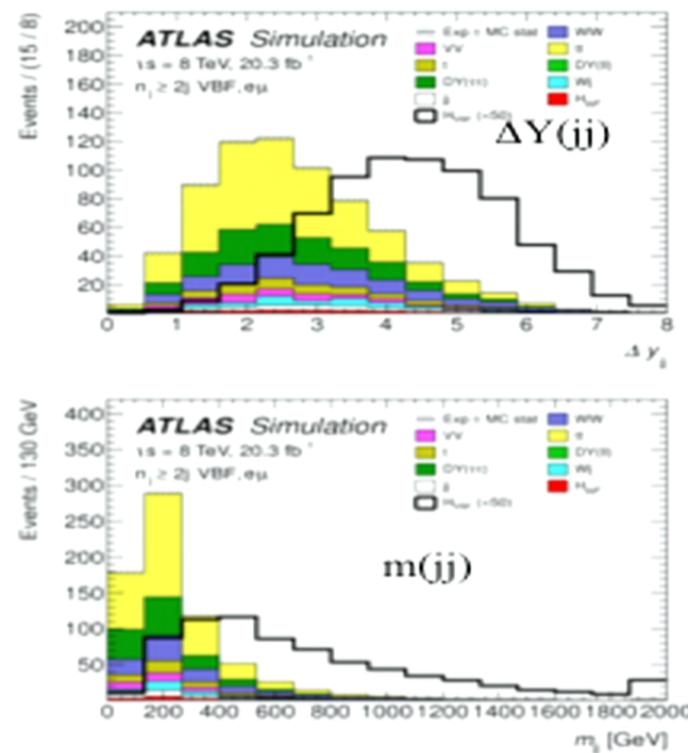
VBF Topology

- Aim to reject dominant top—anti-top background and ggF Higgs

Higgs boson kinematics

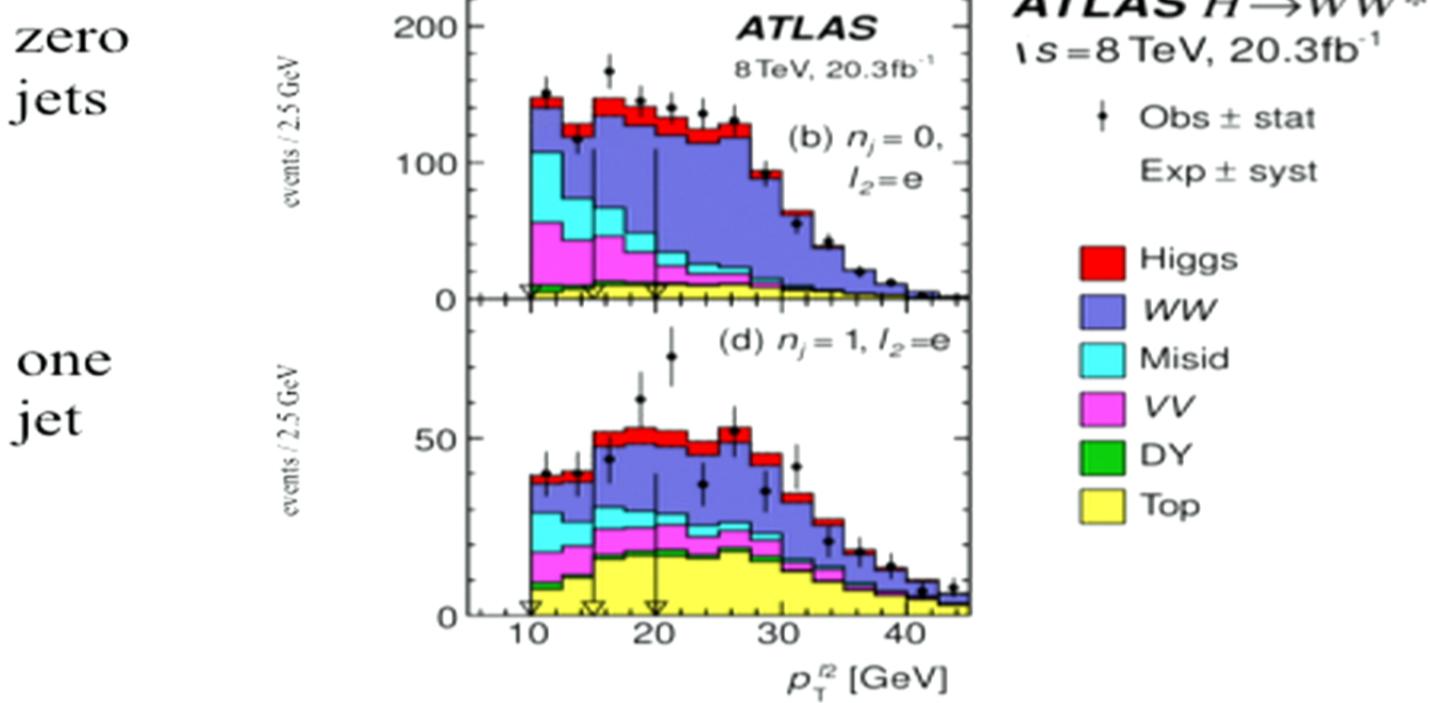


VBF jet kinematics



Gluon Fusion Categorization

- **Subleading lepton p_T** also separates sources of background
- Steeply falling distribution for $W+jets$, $VV = WZ, W\gamma, W\gamma^*$



Boosted Decision Tree

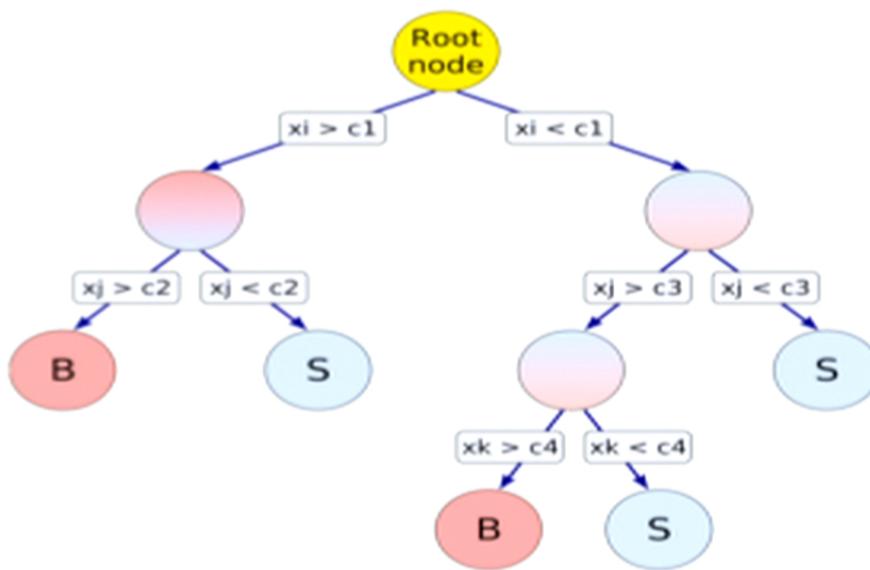


Figure 18: Schematic view of a decision tree. Starting from the root node, a sequence of binary splits using the discriminating variables x_i is applied to the data. Each split uses the variable that at this node gives the best separation between signal and background when being cut on. The same variable may thus be used at several nodes, while others might not be used at all. The leaf nodes at the bottom end of the tree are labeled "S" for signal and "B" for background depending on the majority of events that end up in the respective nodes. For regression trees, the node splitting is performed on the variable that gives the maximum decrease in the average squared error when attributing a constant value of the target variable as output of the node, given by the average of the training events in the corresponding (leaf) node (see Sec. 8.12.3).

<http://tmva.sourceforge.net/docu/TMVAUsersGuide.pdf>

Boosted Decision Tree

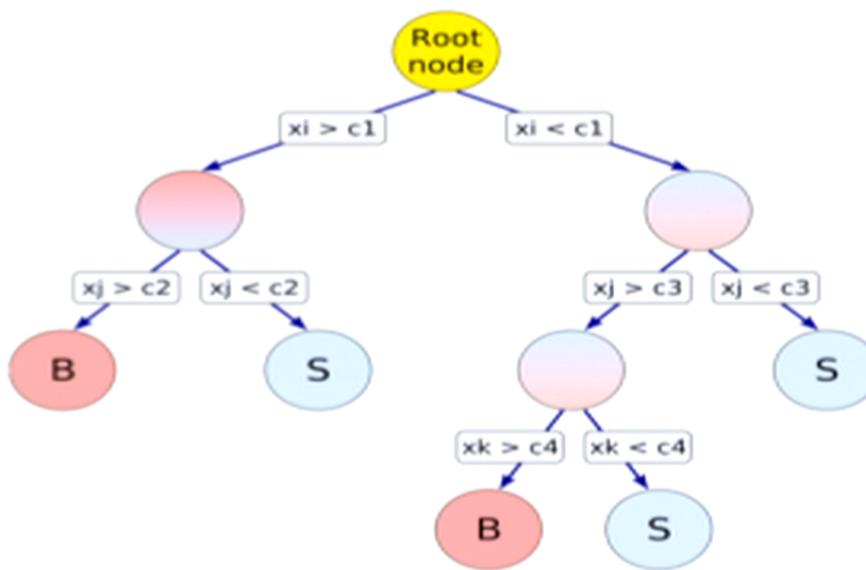


Figure 18: Schematic view of a decision tree. Starting from the root node, a sequence of binary splits using the discriminating variables x_i is applied to the data. Each split uses the variable that at this node gives the best separation between signal and background when being cut on. The same variable may thus be used at several nodes, while others might not be used at all. The leaf nodes at the bottom end of the tree are labeled "S" for signal and "B" for background depending on the majority of events that end up in the respective nodes. For regression trees, the node splitting is performed on the variable that gives the maximum decrease in the average squared error when attributing a constant value of the target variable as output of the node, given by the average of the training events in the corresponding (leaf) node (see Sec. 8.12.3).

<http://tmva.sourceforge.net/docu/TMVAUsersGuide.pdf>

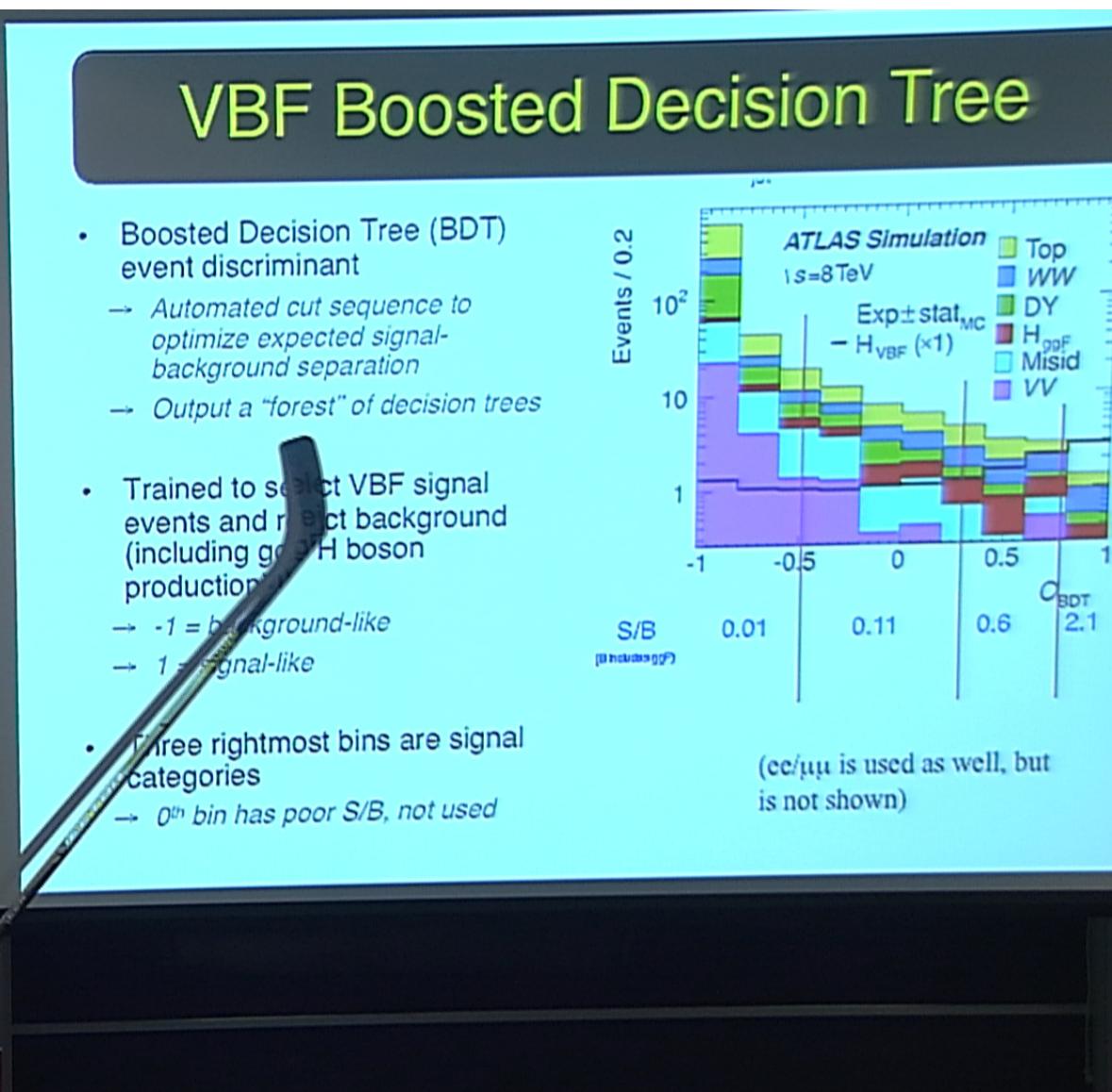
$$\left. \frac{\partial \mathcal{P}_n}{\partial z^*} \right|_{z^* = 0} = \sum_{\substack{\{z^*\} \in \text{Solutions} \\ (n-1)!}} \frac{\mathcal{I}_n(z^*, \kappa, \dots)}{\mathcal{J}(\{z^*\})}$$

$M_n \quad C_n \quad \bar{C}_n$

$\frac{\partial \mathcal{P}_n}{\partial z^*} \Big|_{z^* = 0} = \sum_{\substack{\{z^*\} \in \text{Solutions} \\ (n-1)!}} \frac{\mathcal{I}_n(z^*, \kappa, \dots)}{\mathcal{J}(\{z^*\})}$

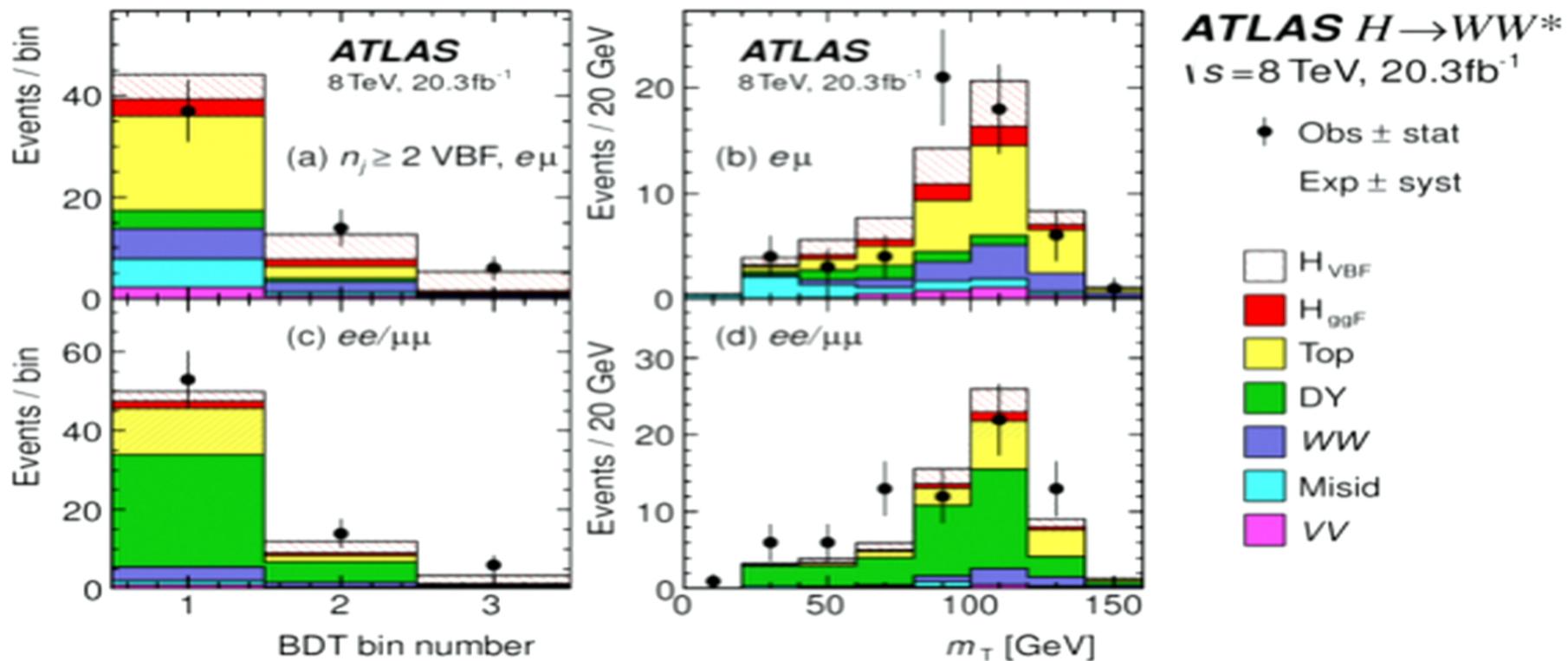
$\mathcal{P}_f' \Psi (\{z \in \kappa\})$

$A_{ab} = \begin{cases} \frac{\kappa_a \kappa_b}{z_a - z_b} & a \neq b \\ 0 & a = b \end{cases}$
 $B_{ab} = \begin{cases} \frac{\epsilon_a \epsilon_b}{z_a - z_b} & a \neq b \\ 0 & a = b \end{cases}$
 $C_{ab} = \begin{cases} \frac{\epsilon_a \epsilon_b}{z_a - z_b} & a \neq b \\ -\sum_{c \neq a} C_{ac} & a = b \end{cases}$
 $\det' \Psi = (\mathcal{P}_f' \Psi)^2 =$



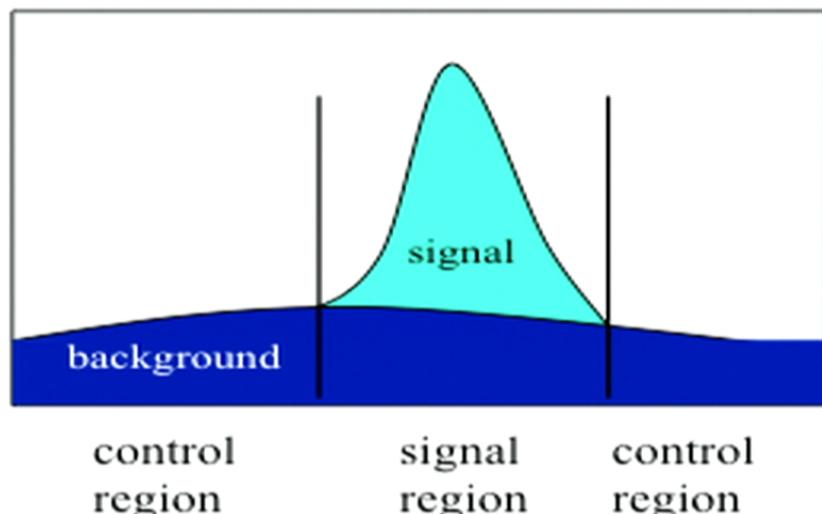
BDT Output

- Results (BDT discriminant and dilepton transverse mass m_T)



Background Estimation

- After Njet and lepton flavor classification, additional criteria to improve the signal-to-background ratio
- Most backgrounds have one or two key distinguishing features
- The events that *fail* that requirement are typically enriched in the target background. **We use these “control regions” to normalize the backgrounds**



- Classic sideband subtraction
- Systematic uncertainty on model used to extrapolate to signal region

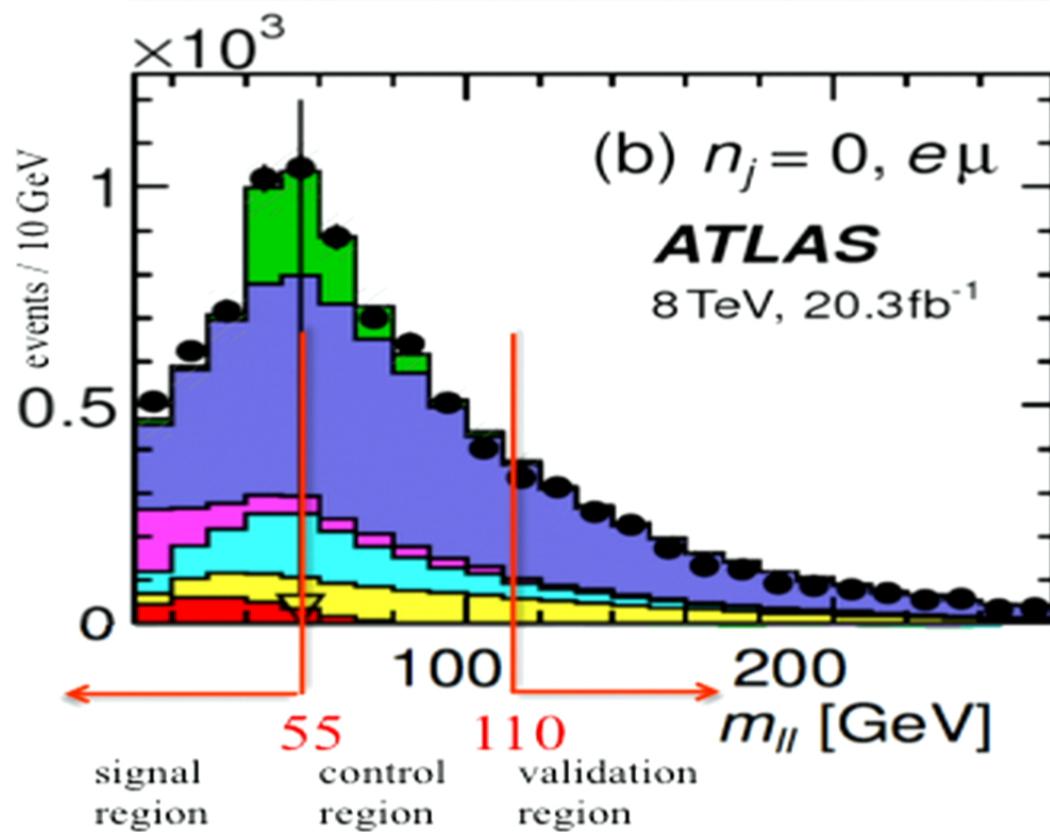
Background Overview

- Some detail on selected topics in next slides
- Rightmost column: yield (and systematic uncertainty) in zero-jet categories (most sensitive region)

process	mechanism	discriminants	0-jet yield
WW	two Ws	m_{\parallel} , m_T	2250 ± 95
top quarks	two Ws (+b-jets)	N(jet), b-tags	307 ± 24
$Z/\gamma^* \rightarrow \tau\tau$	leptonic τ decay	$\Delta\varphi(\text{ll})$, $m_{\tau\tau}$	78 ± 21
$Z/\gamma^* \rightarrow ee/\mu\mu$	E_T^{miss} resolution	E_T^{miss} , recoil	
WZ / $W\gamma^*$	lost/merged lepton	opposite charge lepton pair	
$W\gamma$	photon conversion	conversion veto	420 ± 40
W+jets	second lepton from jet	lepton ID, isolation	360 ± 60
QCD	both leptons from jets	lepton ID, isolation	16 ± 5
$H \rightarrow WW$	the signal...		300 ± 50

WW Background

WW dominates zero- and one-jet ggF categories



Estimated background
 $N_{SR} = \left(\frac{N_{SR}^{MC}}{N_{CR}^{MC}} \right) (N_{CR}^{\text{data}} - N_{CR}^{\text{other}})$

Normalize to data
subtract other processes

extrapolate using MC simulation

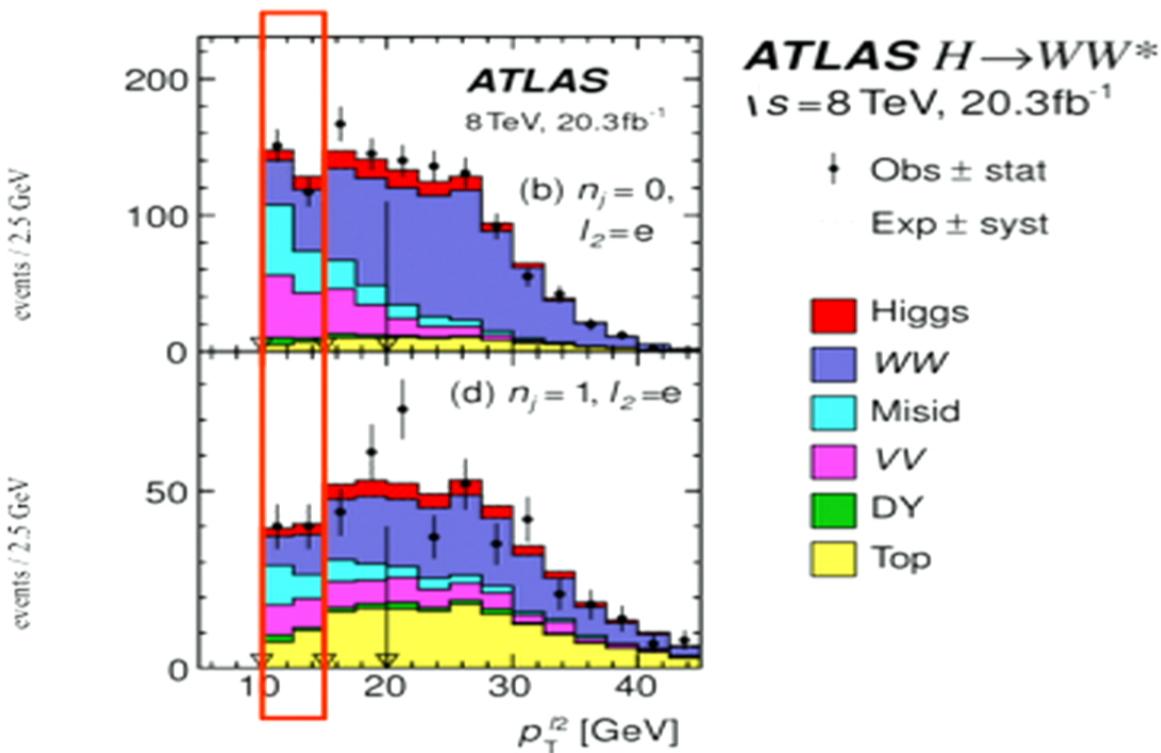
Validation region to test extrapolation
Total uncertainty 4.2% in 0-jet categories; equal parts statistical, theoretical, experimental

Adding more signal: low- p_T bin

the “low- p_T ” data ($10 < p_T^{\text{sublead}} < 15 \text{ GeV}$) add $\sim 20\%$ signal acceptance but challenging backgrounds

zero
jets

one
jet



Top background in 1-jet bin

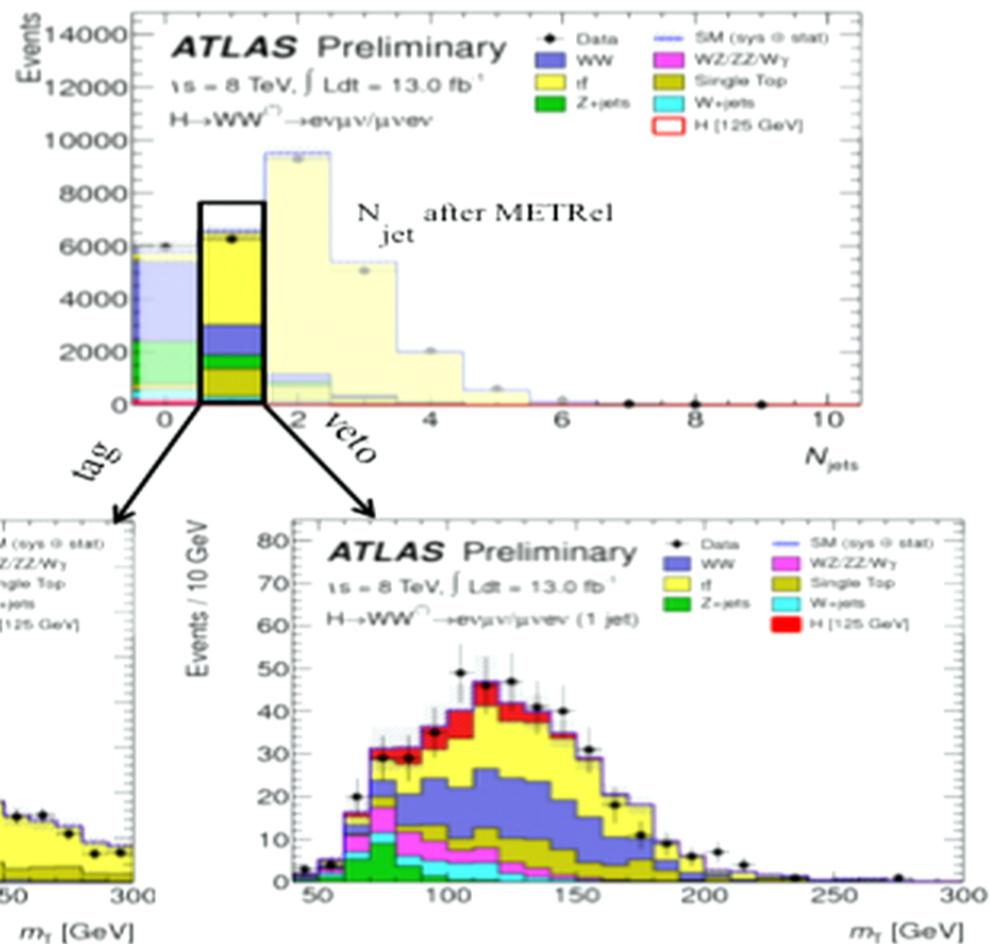
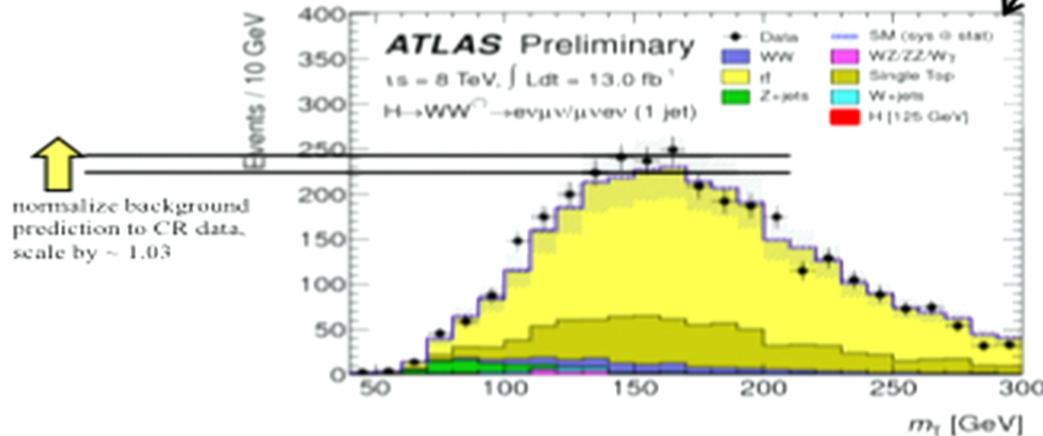
Veto b-tagged jets for signal region

- 85% *b*-jet efficiency operating point
⇒ aggressive veto, only 85% efficient for signal
- Uncertainty on efficiency 5-18%
- Largest BG in 1-jet SR, 44% of total

Tagged events form control region

Total uncertainty on 1-jet top = 37%

b-tagging is leading systematic on 1-jet background yield, at 11%



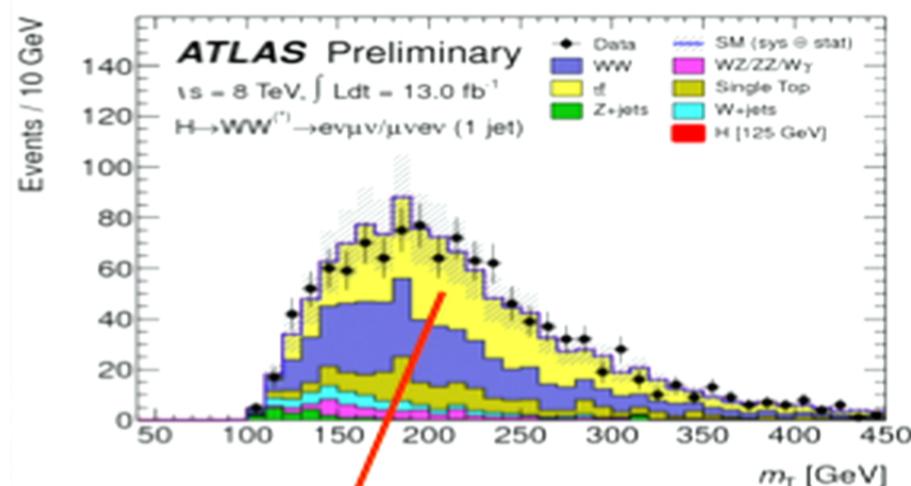
Extrapolation Systematics

Theoretical uncertainties:

- QCD **scale** and **PDFs**, usual prescription
- **Parton shower and underlying event**: Pythia8 / Pythia6 / Herwig
- **Modelling and shape**: MC@NLO vs. MCFM
- shape syst. applied in fit

WW background extrapolation uncertainties

	Scale	PDFs	PS/UE	Modelling
α_{WW}^{0j}	2.5%	3.7%	4.5%	3.5%
α_{WW}^{1j}	4%	2.9%	4.5%	3.5%



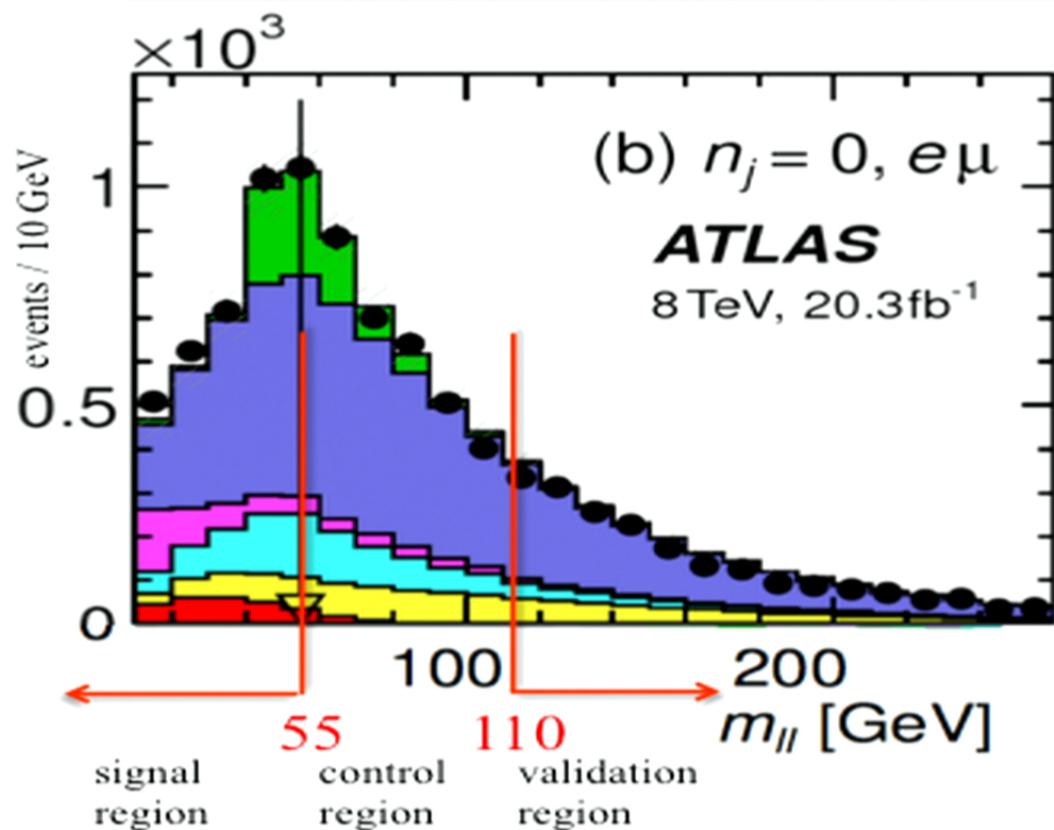
summary of uncertainties:

Background	Stat. (%)	Theory (%)	Expt. (%)	Crosstalk (%)	Total (%)
WW, H+ 0-jet	3.3	7.2	1.5	6.2	13
WW, H+ 1-jet	9	8	12	34	54
top, H+ 1-jet	2	8	29	1	37

corelations
 correctly treated
 in fit: total 1-jet
 BG uncertainty is
 16%

WW Background

WW dominates zero- and one-jet ggF categories



Estimated background
 $N_{SR} = \left(\frac{N_{SR}^{MC}}{N_{CR}^{MC}} \right) (N_{CR}^{\text{data}} - N_{CR}^{\text{other}})$

Normalize to data
subtract other processes

extrapolate using MC simulation

Validation region to test extrapolation
Total uncertainty 4.2% in 0-jet categories; equal parts statistical, theoretical, experimental

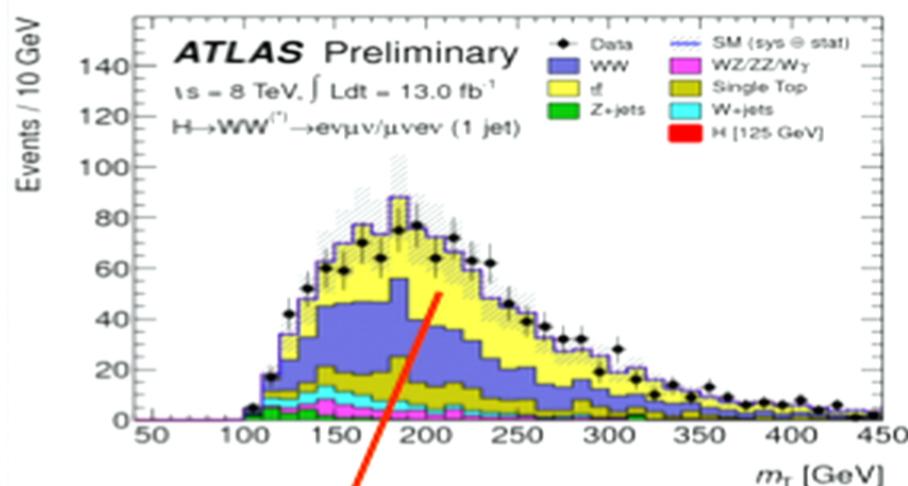
Extrapolation Systematics

Theoretical uncertainties:

- QCD **scale** and **PDFs**, usual prescription
- **Parton shower and underlying event**: Pythia8 / Pythia6 / Herwig
- **Modelling and shape**: MC@NLO vs. MCFM
- shape syst. applied in fit

WW background extrapolation uncertainties

	Scale	PDFs	PS/UE	Modelling
α_{WW}^{0j}	2.5%	3.7%	4.5%	3.5%
α_{WW}^{1j}	4%	2.9%	4.5%	3.5%

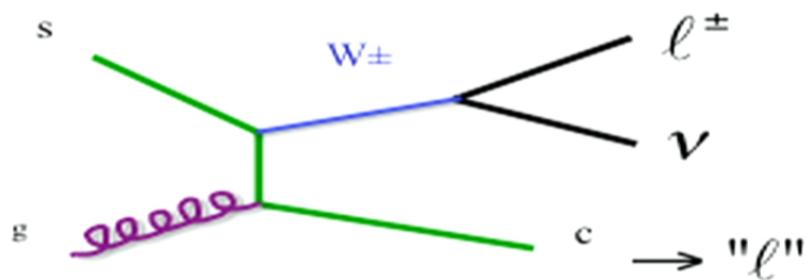


summary of uncertainties:

Background	Stat. (%)	Theory (%)	Expt. (%)	Crosstalk (%)	Total (%)
WW, H+ 0-jet	3.3	7.2	1.5	6.2	13
WW, H+ 1-jet	9	8	12	34	54
top, H+ 1-jet	2	8	29	1	37

corelations
 correctly treated
 in fit: total 1-jet
 BG uncertainty is
 16%

W+jets and Fake leptons



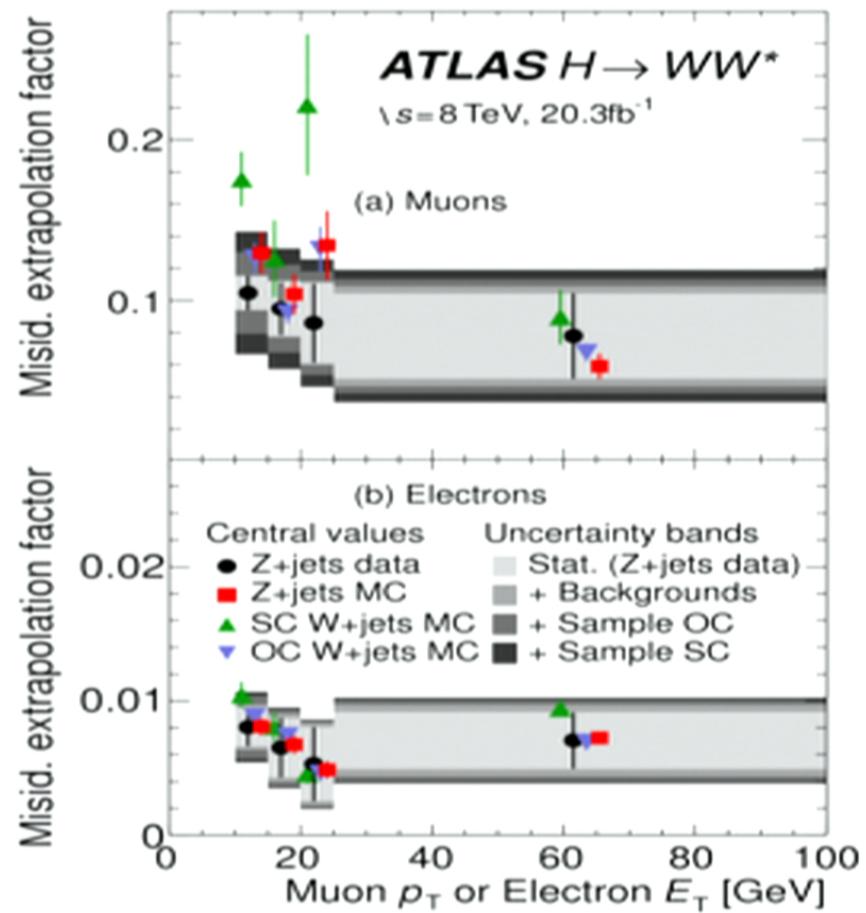
- W+jets events pass signal region selection when a jet produces a lepton
 - Heavy flavor hadron decays
 - Or unusual fragmentation and detector signature

$$\text{ID ID} = \text{ID anti-ID} \times \frac{\text{ID}}{\text{anti-ID}}$$

signal region prediction control region (data) sets normalization transfer factor from QCD-enriched data

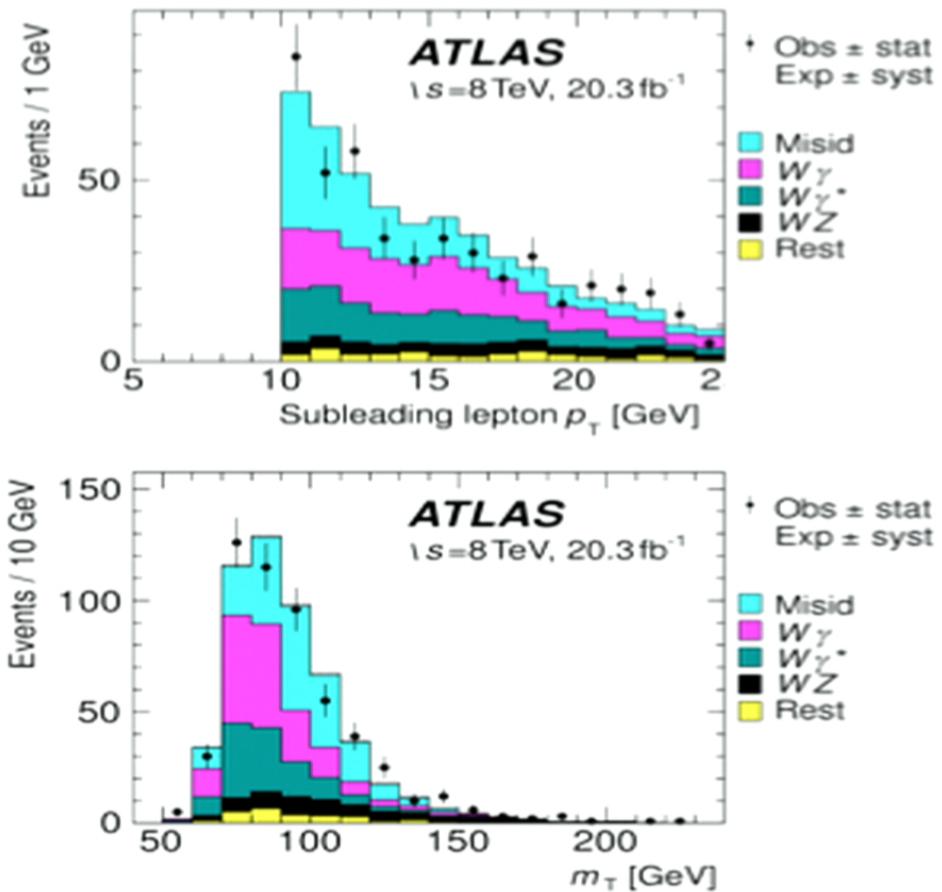
W+jets Control Region

- Flavor of jet (light quark, gluon, bottom or charm quark) influences transfer factor
- Multijet data for transfer factors incurred large systematics (40-50%) in previous analysis iterations
→ *difference of extrapolation factors between multijet vs. W+jets*
- Now: Use Z+jets data, a better analog (20-35% systematic, + stat. uncertainty → will improve)

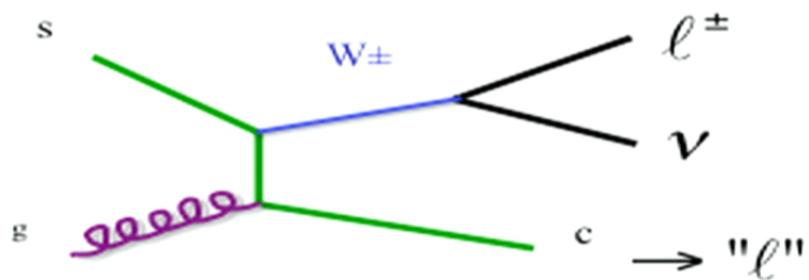


Other Diboson Backgrounds

- “Other dibosons” = “Not WW”
- WZ , $W\gamma^*$, $W\gamma$: Z , γ , γ^*
- Produce lepton with same or opposite charge as W with equal probability
- **Normalize “non-WW” using same-sign data**
→ *Need robust $W+jets$ prediction*
- In this control region, extrapolate from same-sign to opposite-sign



W+jets and Fake leptons

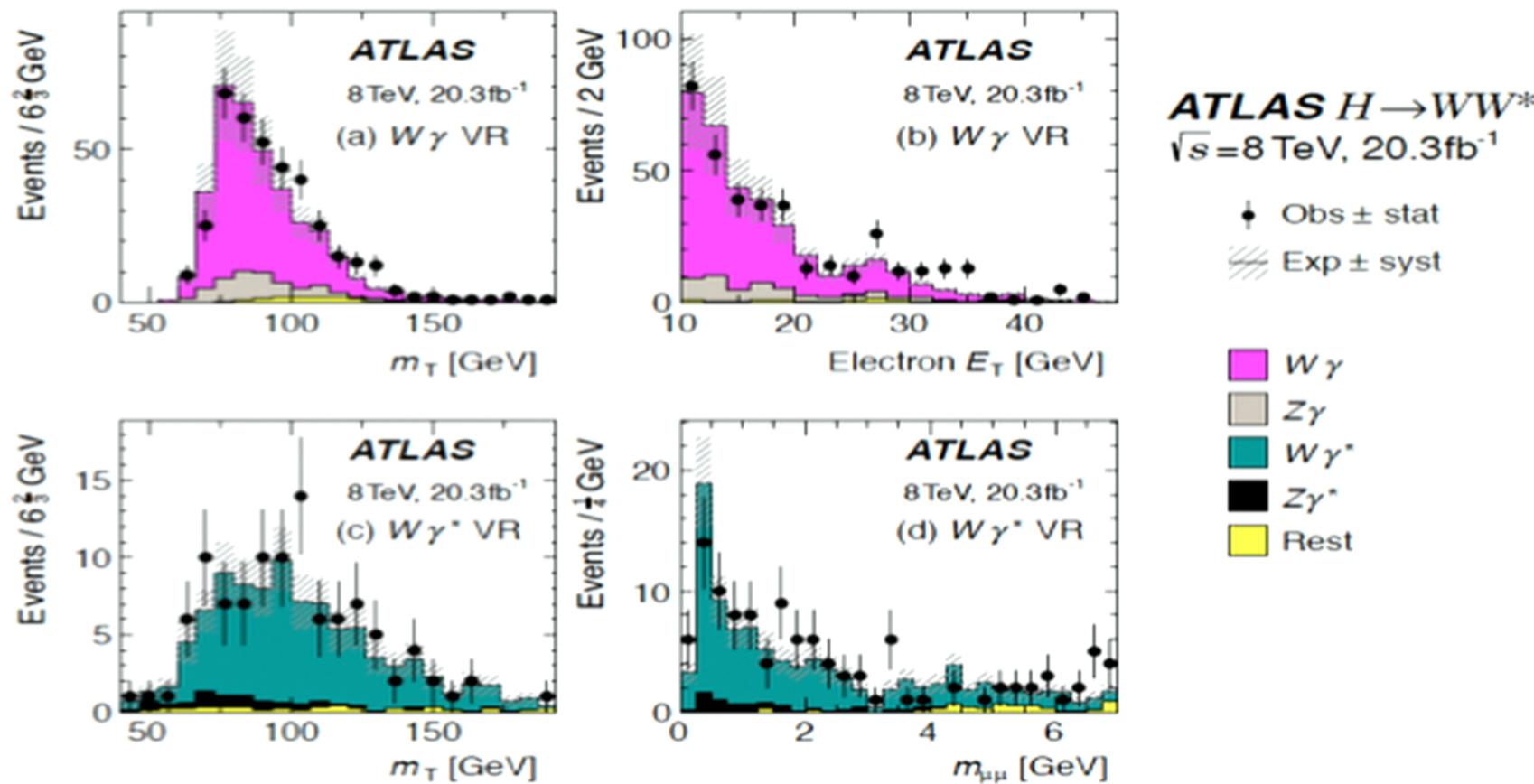


- W+jets events pass signal region selection when a jet produces a lepton
 - Heavy flavor hadron decays
 - Or unusual fragmentation and detector signature

$$\text{ID ID} = \text{ID anti-ID} \times \frac{\text{ID}}{\text{anti-ID}}$$

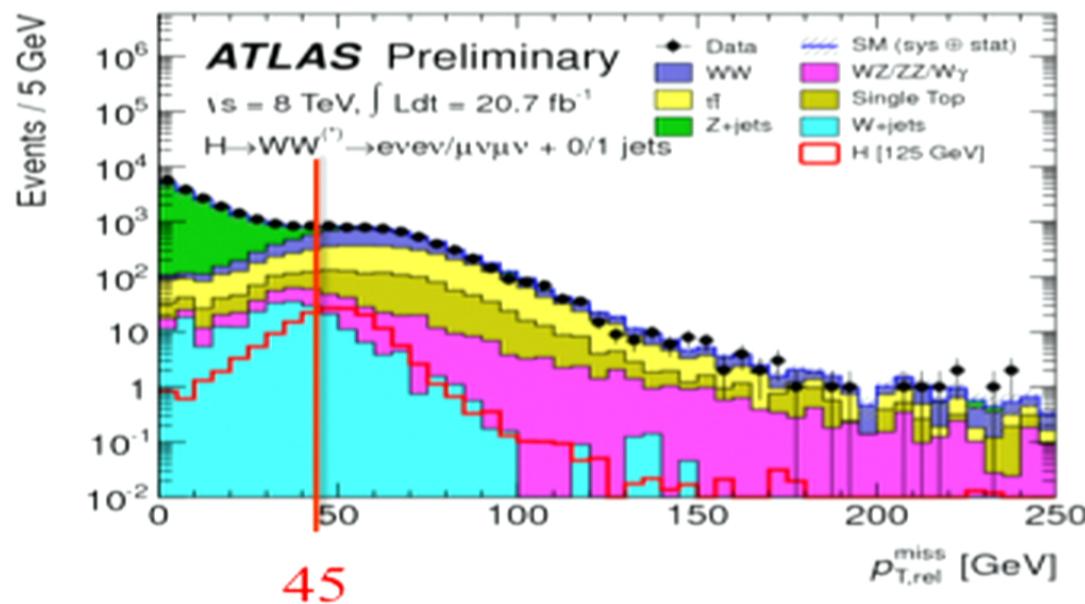
signal region prediction control region (data) sets normalization transfer factor from QCD-enriched data

Validations Regions



Z/ γ^* Background

Measure missing transverse energy only with tracks: more robust against extra interactions (pileup)



$$\vec{E}_T^{\text{miss}} = - \sum_{\text{objects}i} \vec{p}_T^{\text{object}}$$

- Left: additional rejection after initial ETmiss requirement
→ *First inclusion of ee+ $\mu\mu$ channels in 2012 WW analysis*

Control Region Yields

TABLE XIX. Control region event yields for 8 TeV data. All of the background processes are normalized with the corresponding β given in Table XX or with the data-derived methods as described in the text; each row shows the composition of one CR. The N_{sig} column includes the contributions from all signal production processes. For the VBF-enriched $n_j \geq 2$, the values for the bins in O_{BDT} are given. The entries that correspond to the target process for the CR are given in bold; this quantity corresponds to N_{bold} considered in the last column for the purity of the sample (in %). The uncertainties on N_{bkg} are due to sample size.

Control regions	Summary			Composition of N_{bkg}					Purity		
	N_{obs}	N_{bkg}	N_{sig}	N_{WW}	N_{top}	N_{misid}	N_{VV}	N_{DY}	$N_{ee/\mu\mu}$	$N_{\tau\tau}$	$N_{\text{bold}}/N_{\text{bkg}}$ (%)
$n_j = 0$											
CR for WW	2713	2680 ± 9	28	1950	335	184	97	8.7	106		73
CR for top quarks	76013	75730 ± 50	618	8120	56210	2730	1330	138	7200		74
CR for VV	533	531 ± 8	2.2	2.5	1.1	180	327	19	2.7		62
CR for $Z/\gamma^* \rightarrow \tau\tau$	4557	4530 ± 30	23	117	16.5	239	33	28	4100		91
$n_j = 1$											
CR for WW	2647	2640 ± 12	4.3	1148	1114	165	127	17	81		43
CR for top quarks	6722	6680 ± 12	17	244	6070	102	50	6	204		91
CR for VV	194	192 ± 4	1.9	1	3.1	65	117	4.7	0.8		61
CR for $Z/\gamma^* \rightarrow \tau\tau$	1540	1520 ± 14	18	100	75	84	27	7	1220		80
$n_j \geq 2$ ggF											
CR for top quarks	2664	2660 ± 10	4.9	561	1821	129	101	10	44		68
CR for $Z/\gamma^* \rightarrow \tau\tau$	266	263 ± 6	2.6	13	34	18	4.1	0.1	194		74
$n_j \geq 2$ VBF											
CR for top quarks, bin 1	143	142 ± 2	2.1	1.9	130	2.1	0.8	6.3	1.1		92
CR for top quarks, bin 2–3	14	14.3 ± 0.5	1.8	0.6	11.6	0.2	0.2	0.9	0.2		81
CR for $Z/\gamma^* \rightarrow \tau\tau$	24	20.7 ± 0.9	2.4	0.9	1.2	0.6	0.2	0.8	17		82

$$H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$$

Results:

- Observed (expected) significance: 6.1σ (5.8σ)
- Observed (expected) significance for VBF: 3.2σ (2.7σ)

Combined $WW \rightarrow \ell\nu\ell\nu$ signal strength
 $\mu = 1.08^{+0.16}_{-0.15}$ (stat.) $^{+0.16}_{-0.13}$ (syst.)

