Third International Workshop on
Prospects for Charged Higgs Discovery at Colliders
Uppsala, Sweden, 27-30 September 2010

W+jets background in charged Higgs searches

Fabien Tarrade
Brookhaven National Laboratory, New York

BROOKHAVEN NATIONAL LABORATORY

on behalf of the ATLAS collaboration

Introduction

Charged Higgs :

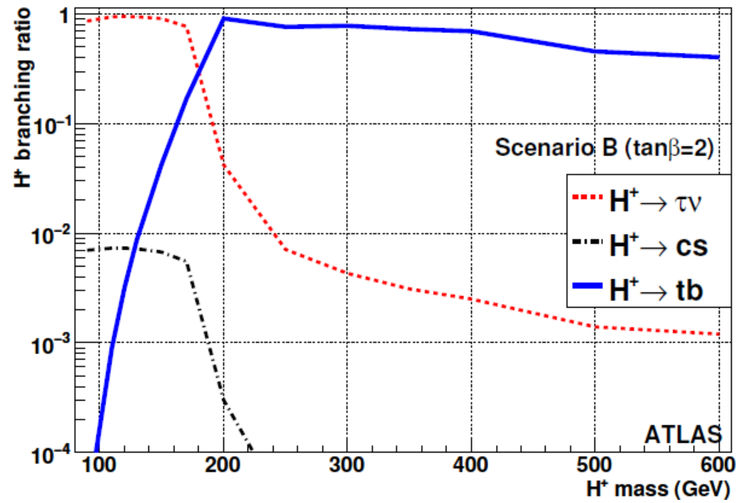
- Importance of charged Higgs searches
[see talk by Arnaud Ferrari and Simonetta Gentile]

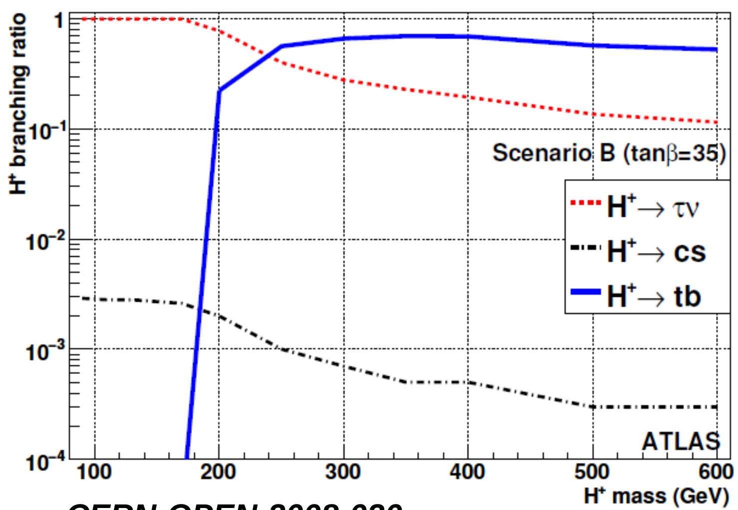
- $m_{H+} < m_{top}$ (light charged Higgs) :
production at LHC via:
$$gg \to t\bar{t} \to bW\ bH^{+}$$

- Importance of the **hadronic τ-jet**:
with the exception of small tan β values:
$$BR\ (H^{+} \to \tau^{+} \nu_{\tau}) \sim 1$$
[see talk by Miika Klemetti and Yann Coadou]

- For tan β < 1, **BR ($H^{+} \to c\bar{s}$)**
may reach 40 % for m_{H+} ~130 GeV
[see talk by Un-Ki Yang]

- **W+jets** events could be an important background





CERN-OPEN-2008-020

Motivations 1/2

$H^+ \rightarrow \tau^+ \nu_\tau$:

- **$m_{H+} < m_{top}$** (light charged Higgs), production at LHC via: **$gg \rightarrow t\bar{t} \rightarrow bW\, bH^+$**
- Look at **$H^+ \rightarrow \tau^+ \nu_\tau$**:
  [see talk by Miika Klemetti for more details]
  - Hadronic τ+ lepton: $H^+ \rightarrow \tau\text{-jet}\ \nu_\tau$ $W \rightarrow l\ \nu_l$
  - Leptonic τ: $H^+ \rightarrow \tau\ \nu_\tau \rightarrow l\ \nu_\tau\ \nu_l$ $W \rightarrow qq'$
  - Di Leptonic: $H^+ \rightarrow \tau\ \nu_\tau \rightarrow l\ \nu_\tau\ \nu_l$ $W \rightarrow l\ \nu_l$
- **Signature**:
  - 2 b-jets + 2 jets + [1 τ-jet or 1 l] + MET (missing E_T)
  - 2 b-jets + 1 τ-jet + 1 l + MET
  - 2 b-jets + 2 l + MET
- **W+jets** events could be an important **background**:
  - W (→τ-jet ν_τ) + jets (a jet → fake l)
  - W (→l ν_l) + jets (a jet → fake τ-jet)
  - W (→τ ν_τ →l ν_l ν_τ) + jets (a jet → fake τ-jet)

electron (e)
muon (μ)
lepton (l)=e, μ
τ-jets: τ lepton decaying to hadron(s)

Simulation with 1fb^{-1} at $\sqrt{s}$=7TeV

Process	Number of events after no cut	all cuts
Signal $m_{H^+} = 90$ GeV	2.5×10^3	282
Signal $m_{H^+} = 110$ GeV	2.5×10^3	330
Signal $m_{H^+} = 130$ GeV	2.5×10^3	326
Signal $m_{H^+} = 150$ GeV	2.5×10^3	284
SM $t\bar{t}$ not hadronic	87.3×10^3	1194
Single top Wt-channel	5.7×10^3	55
Single top t-channel	20.4×10^3	43
Single top s-channel	0.9×10^3	3
$Z \rightarrow ll$ + jets	3.1×10^6	4
$W \rightarrow l\nu$ + jets	3.2×10^7	42
Wbb + jets	8.7×10^3	12
Zbb + jets	2.8×10^4	11

ATL-PHYS-PUB-2010-09

Motivations 2/2

$H^+ \to c\bar{s}$:

- **$m_{H+} < m_{top}$** (light charged Higgs), production at LHC via: **$gg \to t\bar{t} \to bW\ bH^+$**
- Look at **$H^+ \to c\bar{s}$** :
  [see talk by Un-Ki Yang]
  - Hadronic H: $H^+ \to c\bar{s}$ $W \to l\ \nu_l$
- **Signature**:
  - 2 b-jets + 2 jets + l + MET
- **W+jets** is a background:
  - W ($\to \tau\ \nu_\tau \to l\ \nu_l\ \nu_\tau$) + jets
  - W ($\to l\ \nu_l\ \nu_\tau$) + jets
- As for **$H^+ \to \tau^+ \nu_\tau$**, for **$H^+ \to c\bar{s}$** the main background is **$t\bar{t}$** [see talk by Martin Flechl for more details]

Simulation with 1fb^{-1} at $\sqrt{s}$=7TeV

Process	Expected number	
	no cut	all cuts
$H^+ \to c\bar{s}$, 90 GeV	9.5×10^3	148
$H^+ \to c\bar{s}$, 110 GeV	9.5×10^3	144
$H^+ \to c\bar{s}$, 130 GeV	9.5×10^3	98
$H^+ \to c\bar{s}$, 150 GeV	9.5×10^3	56
SM $t\bar{t}$, not all hadronic	87.4×10^3	1370
Single top, Wt-channel	5.7×10^3	18
Single top, t-channel	20.4×10^3	33
Single top, s-channel ($e\nu$)	448	1
Single top, s-channel ($\mu\nu$)	448	1
Wbb + jets	5.6×10^3	9
$W \to e\nu$ + jets	39.2×10^3	2
$W \to \mu\nu$ + jets	38.7×10^3	3
$W \to \tau\nu$ + jets	39.0×10^3	0

ATL-PHYS-PUB-2010-09

Data Sample

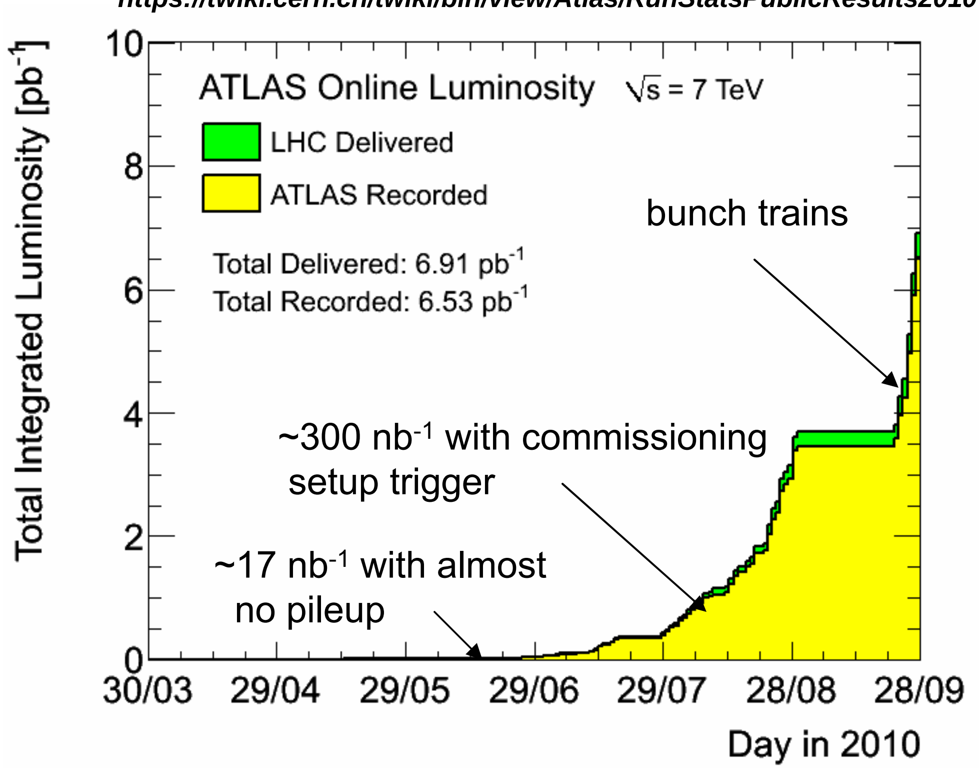
LHC :

- LHC luminosity still increasing **exponentially**

ATLAS :

- Detector performing very **well** [see talk by Domizia Orestano]
- Just 6 months of data taking at **$\sqrt{s}$=7 TeV**
- Rediscovery of the **Standard Model** : W, Z, top
- Results shown today use a **subset** of full data sample

https://twiki.cern.ch/twiki/bin/view/Atlas/RunStatsPublicResults2010

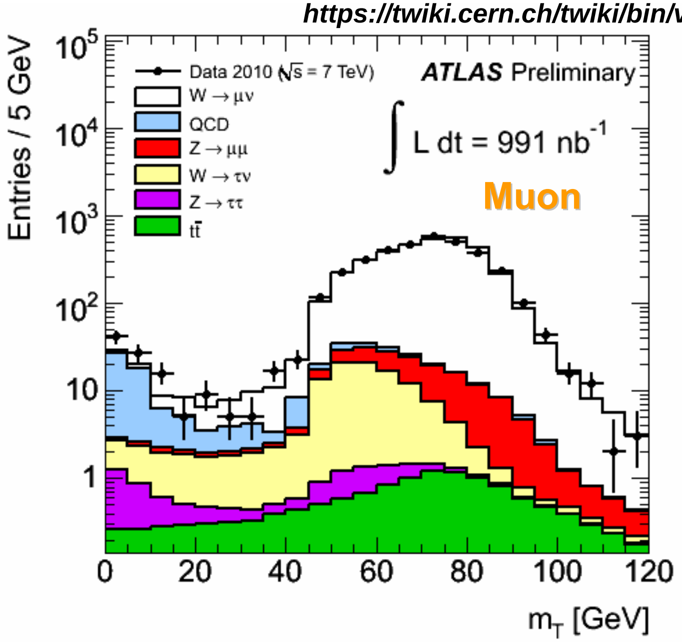


Rediscovering the SM : W±

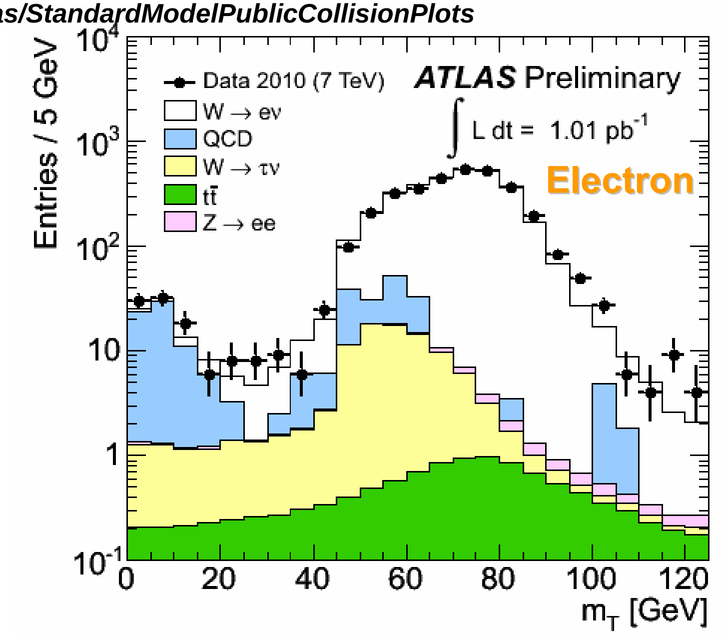
Electron and Muon channel :

- Lepton p_T>20 GeV and MET>25 GeV
- Transverse mass: $m_T = \sqrt{2(p_T^{\mu})(E_T^{\text{miss}})(1-\cos(\varphi^{\mu}-\varphi^{E_T^{\text{miss}}}))}$

https://twiki.cern.ch/twiki/bin/view/Atlas/StandardModelPublicCollisionPlots



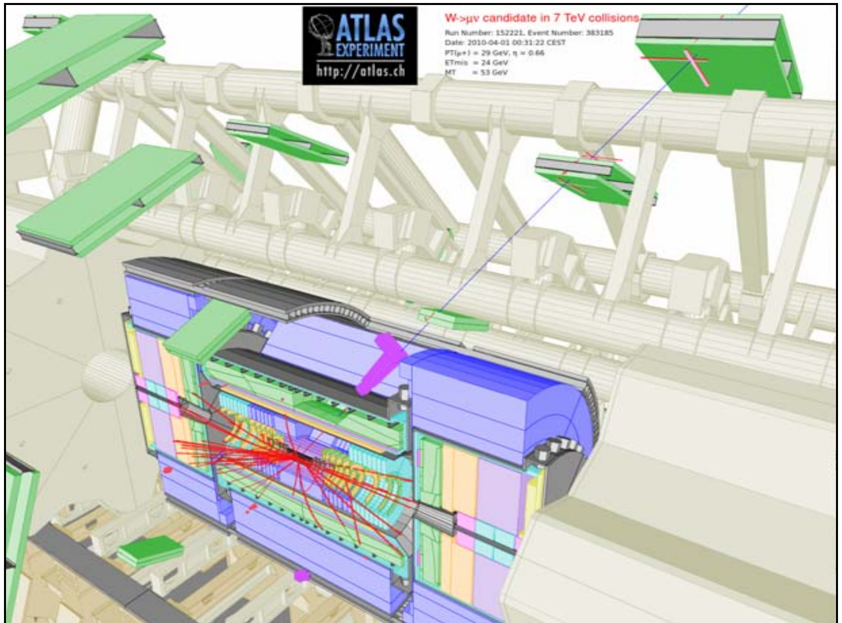
3425 events with m_T>40 GeV



3233 events with m_T>40 GeV

- Good agreement between **Data** and **Monte Carlo** (MC)

$W^{\pm}$ Event Display



W→μ ν candidate in 7 TeV collisions

https://twiki.cern.ch/twiki/bin/view/Atlas/EventDisplayPublicResults

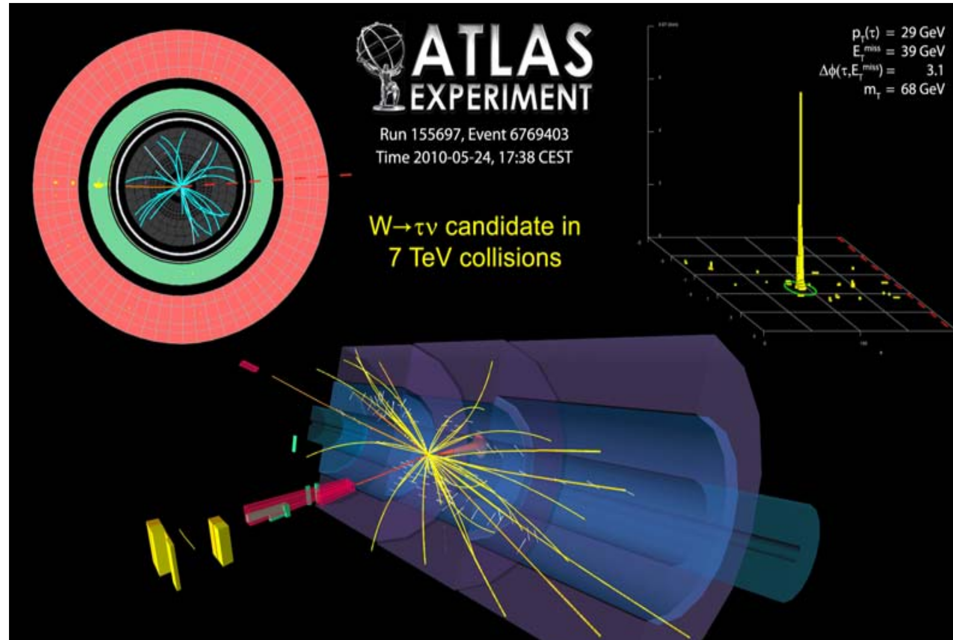
Candidate for W→μν decay, collected on 1 April 2010.

Event properties: $p_T(\mu^+)$ = 29 GeV

$\eta(\mu^+)$ = 0.66

E_T^{miss}= 24 GeV

M_T = 53 GeV



W→τ-jet ν candidate in 7 TeV collisions

https://twiki.cern.ch/twiki/bin/view/Atlas/EventDisplayPublicResults

A candidate for a W->τν decay, with a hadronically decaying tau

Event properties:

$p_T(\tau)$ = 29 GeV

E_T^{miss} = 39 GeV

$\Delta\varphi(\tau, E_T^{miss})$ = 3.1

m_T = 68 GeV

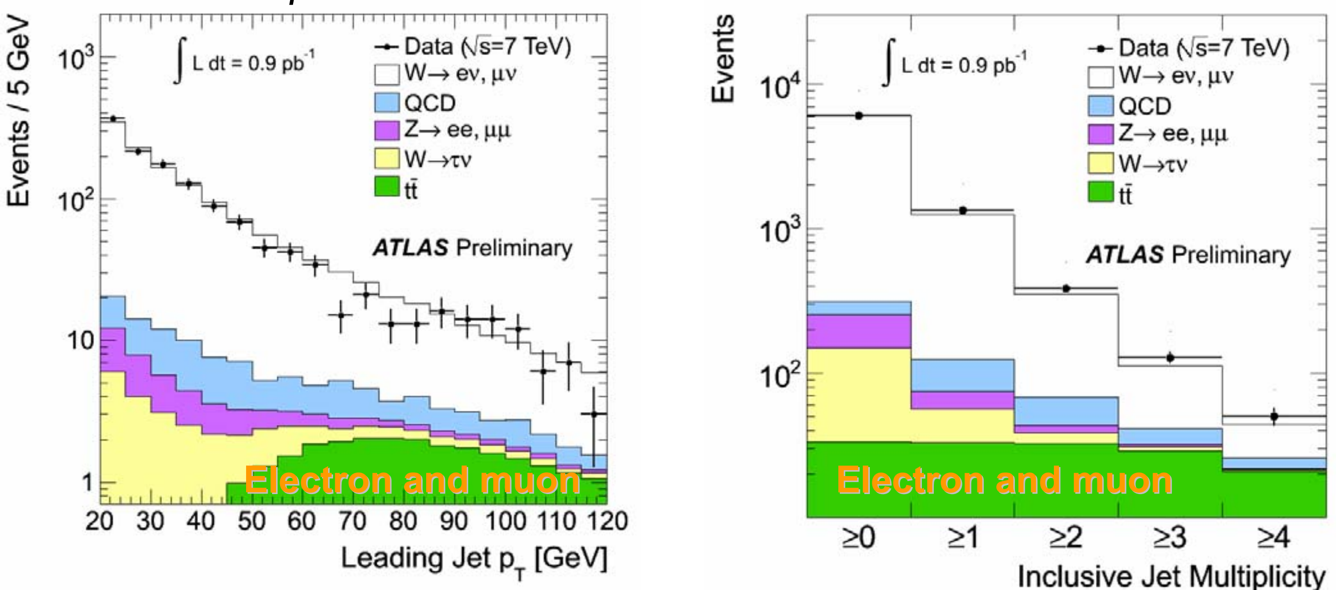
collected on 24 May 2010

Rediscovering the SM : $W^{\pm}$ + jets

Electron and Muon channel :

- Jets produced in association with **$W \rightarrow l\nu_l$**, where **l=e,μ**
- Jet algorithm : **Anti-k_T** with **R=0.4** and **$|\eta|<2.8$** and **$p_T>20$** GeV

https://twiki.cern.ch/twiki/bin/view/Atlas/StandardModelPublicCollisionPlots



The MC is normalized to the inclusive data sample

- Good agreement between **Data** and **MC**

W+jets background

Charged Higgs Searches :

- **W+jets** events can be an important background
- **Large uncertainty** on Monte Carlo prediction of **W + jets** events
- Difficulties in **accurately simulating** events with **jets misidentified** as **leptons** (e,μ)
- Difficulties in **accurately simulating** events with **jets misidentified** as **τ-jets**
- Need to estimate the **W+jets background** with a **data driven estimation**

Methods :

- Try to **review results** on the different methods based on **Data** and **MC**
- Look at the **methods** developed and used in **different analyses** :
  top, MSSM Higgs, SM Higgs

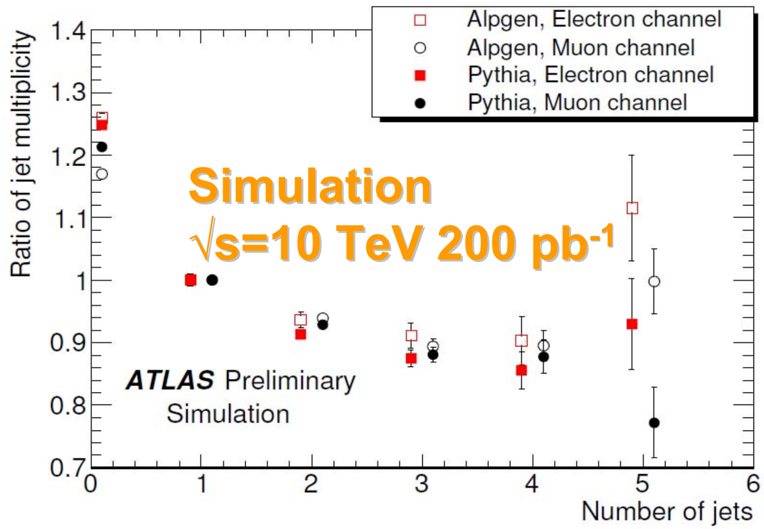
Monte Carlo

W+jets background

Data driven estimate for W+jets background (applied on MC) :

- W to Z **ratio** predicted with a small **uncertainty** and **jet multiplicity** distribution for Z events can be **measured** to reduce the MC uncertainty on the fraction of **W+jets**

ATL-PHYS-PUB-2009-087



Ratio of reconstructed jet multiplicity for $W \rightarrow l\nu$+jets over $Z \rightarrow ll$+jets events

- Used in the **$t\bar{t}$** cross section **analysis**

- Extrapolation from a **control region (CR)** with **0 or 1 jet** into the **t** (top quark) **signal region (SR)** with **≥ 4 jets**

- Estimation of the number of **W+jets** background events using the **formula**:

$$(W^{\rm SR}/W^{\rm CR})_{\rm data} = (Z^{\rm SR}/Z^{\rm CR})_{\rm data} \cdot C_{\rm MC}$$

where: $$C_{\rm MC} = \frac{(W^{\rm SR}/W^{\rm CR})_{\rm MC}}{(Z^{\rm SR}/Z^{\rm CR})_{\rm MC}}$$ ***[data=MC for this study]***

- **Expected** total **uncertainties** on the **W+jets** background estimation: **23.9% (19.6%)** for the $W \rightarrow e\,\nu_e$ + jets ($W \rightarrow \mu\,\nu_\mu$ + jets) and with 200 pb^{-1} (early data scenario)

- This **data driven** technique can be used for **charged Higgs searches** to estimate **W+jets** background in the **e,μ channels**

W+jets background

Data driven estimation of the lepton misidentification rate (applied on MC):

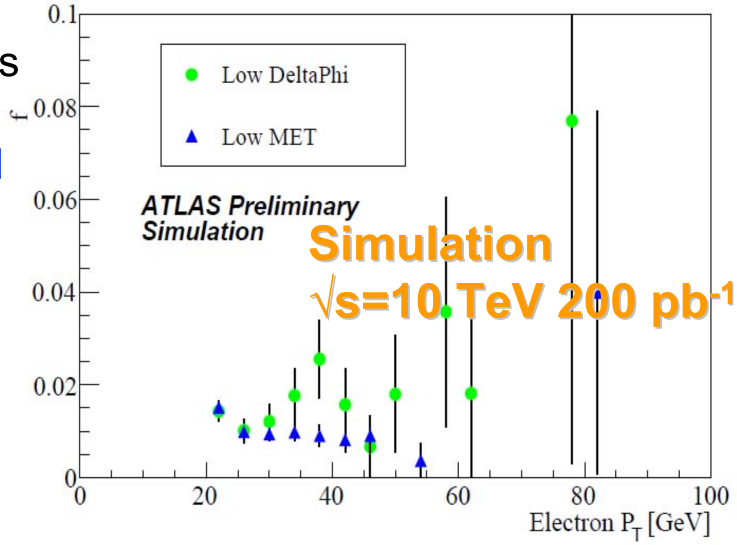
- **Probability** of a jet to **fake a lepton** can be determined from the data
- Used in the context of the $t\bar{t}$ cross section **analysis**
- **Tight(loose)** selection dominated by **real(fake)** leptons
- Probability of a **real (fake)** lepton to be **reconstructed** as a **tight lepton ε (f)**:

$$\epsilon = \frac{N_{\text{tight,real}}}{N_{\text{tight,real}} + N_{\text{loose,real}}}$$

$$f = \frac{N_{\text{tight,fake}}}{N_{\text{tight,fake}} + N_{\text{loose,fake}}}$$

[data=MC for this study]

ATL-PHYS-PUB-2009-086



Tight fake probabilities f measured as a function of p_T

- Measure **efficiency ε, f** in independent samples **dominated** by **real** and **fake leptons** respectively
- Measure **ε** using a **tag** and **probe** method in events with **two leptons**:
  - **M_{ll}** within **M_Z ± 5 GeV**
  - **MET < 15 GeV**

W+jets background

Data driven estimation of the lepton misidentification rate (applied on MC) :

- Measure **f** in two samples dominated by **fakes**:
  low ΔΦ: **MET> 15 GeV** and **ΔΦ < 1 rad** and a tight or loose lepton
  low MET: **MET< 15 GeV** and a tight or loose lepton

- Consider **different types** of events:
  - at the **reconstruction level** : 2 tight lepton [**TT**], loose and tight lepton [**TL/LT**]
  - at the **truth level** : 2 real leptons [**RR**], real and fake [**RF/FR**]
  which are related to each other in an **efficiency matrix** (index 1 (2) for the 1st (2nd) lepton):

$$\begin{bmatrix} N_{\mathrm{TT}} \\ N_{\mathrm{TL}} \\ N_{\mathrm{LT}} \end{bmatrix} = \begin{bmatrix} \epsilon_1\epsilon_2 & \epsilon_1 f_2 & f_1\epsilon_2 \\ \epsilon_1(1-\epsilon_2) & \epsilon_1(1-f_2) & f_1(1-\epsilon_2) \\ (1-\epsilon_1)\epsilon_2 & (1-\epsilon_1)f_2 & (1-f_1)\epsilon_2 \end{bmatrix} \begin{bmatrix} N_{\mathrm{RR}} \\ N_{\mathrm{RF}} \\ N_{\mathrm{FR}} \end{bmatrix}$$

[data=MC for this study]

- The **number of fakes** is obtained by the formula:

$$N_{\mathrm{Fake}} = \left[\frac{f_2(\epsilon_2-1)}{\epsilon_2-f_2} + \frac{f_1(\epsilon_1-1)}{\epsilon_1-f_1}\right] N_{\mathrm{TT}} + \frac{f_2\epsilon_2}{\epsilon_2-f_2} N_{\mathrm{TL}} + \frac{f_1\epsilon_1}{\epsilon_1-f_1} N_{\mathrm{LT}}$$

- **Expected** total **uncertainties** of **50(100)%** for the μ(e) analysis with 200 pb^{-1} (early data scenario)

- This **data driven** estimation of the **lepton misidentification rate** is used for **charged Higgs searches** for example for charged Higgs studies at √s=10 TeV (ATL-PHYS-PUB-2010-006)

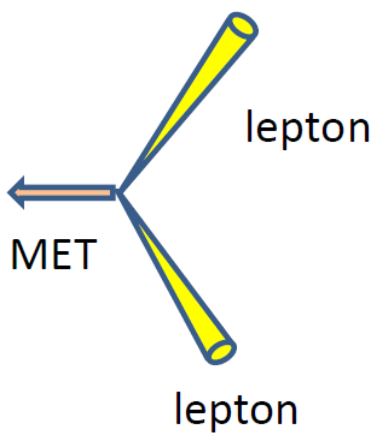
Data at $\sqrt{s}$=7 TeV

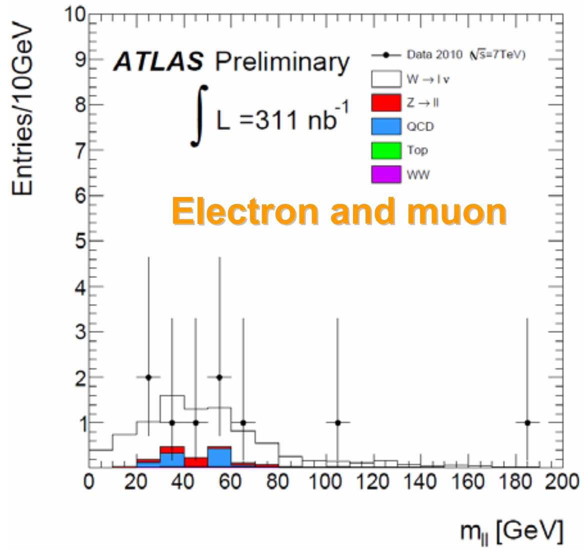
W+jets background

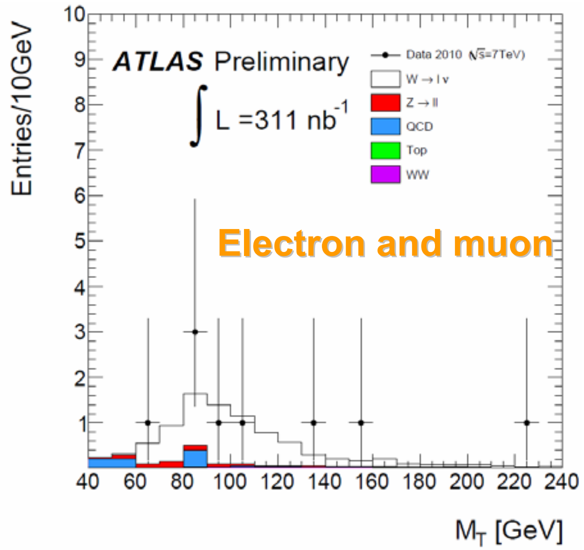
Observation of the Background from W+jets to the $H^{SM} \rightarrow WW^{*} \rightarrow l\nu_l\, l\nu_l$ at $\sqrt{s}$=7 TeV :

- **Measurement** of W+jets->l(good)+l(fake) background
- Search for candidate events with a **tight** and a **loose** lepton and look into the **same sign combinations** (as an additional cross check)
- **MET>25 GeV** and **veto** on events with a least **two jets** (p_T>20 GeV and $|\eta|$<2.8)
- **9 events pass** these requirements and are **consistent** with **MC** expectation

ATLAS-COM-CONF-2010-092







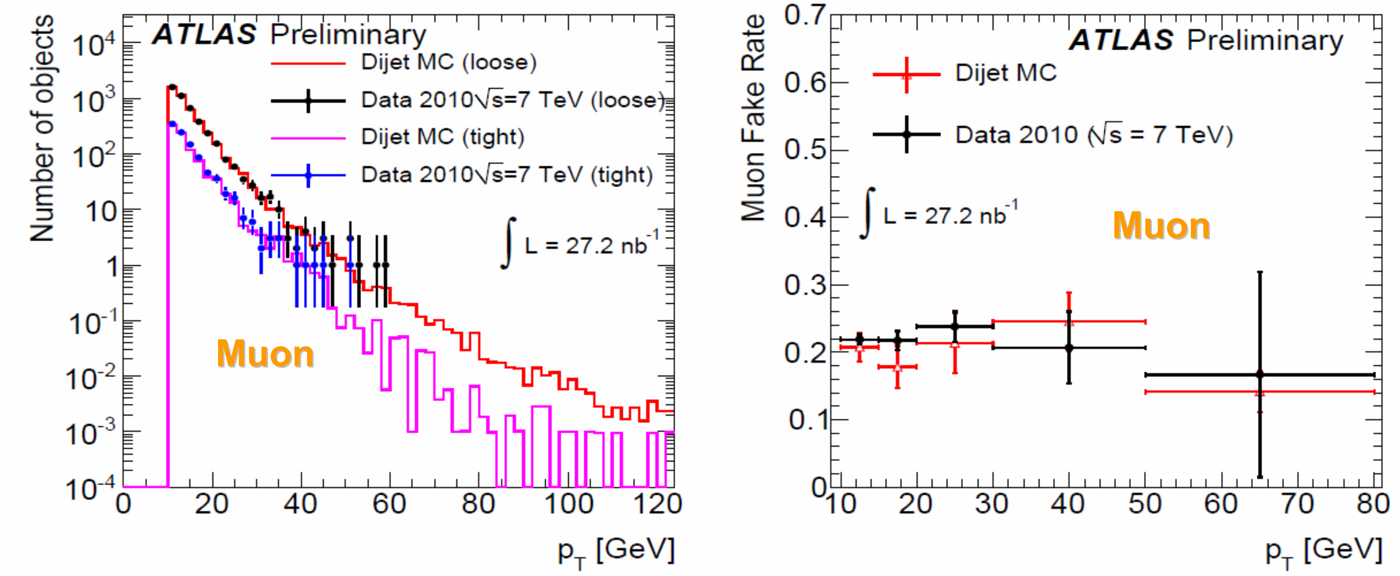
W+jets background

Observation of the Background from W+jets to the $H^{SM} \rightarrow WW^{*} \rightarrow l\nu_l\, l\nu_l$ at $\sqrt{s}$=7 TeV :

- **Extract** the **fake lepton** rates from data:
  - use real data sample dominated by **fake leptons** (e.g dijet data)
  - fake rate is **calculated** with very **loose selection** (e.g fakeable object)

$$f_\ell \equiv \frac{N_{\text{tight object}}}{N_{\text{loose object}}}$$

ATLAS-COM-CONF-2010-092



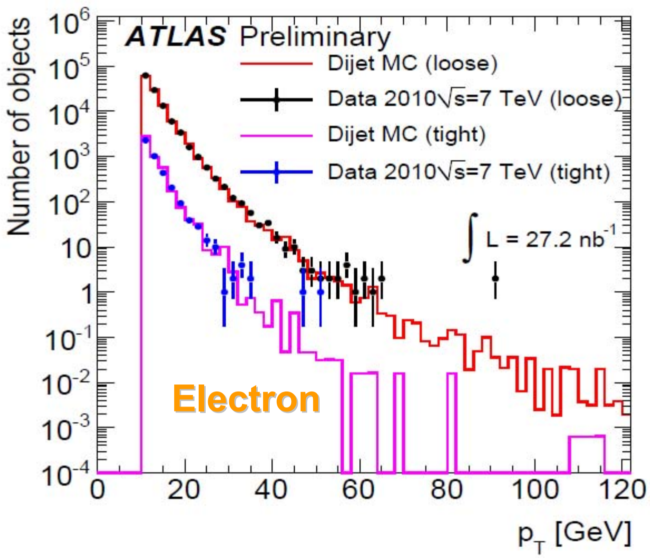
W+jets background

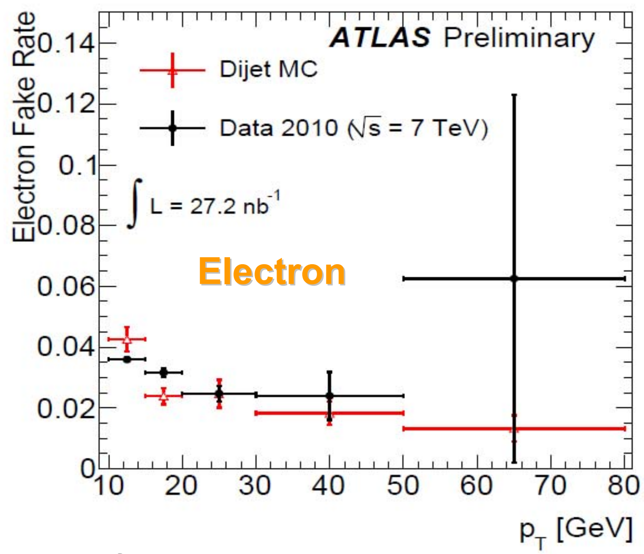
Observation of the Background from W+jets to the H→WW*→lν_l lν_l at √s=7 TeV :

- **Lepton rates** are **extracted** from dijet with the following **uncertainties** (electron/muon)

Object	Electron	Muon
Real lepton contamination	$< 1\%$	$< 1\%$
EW veto selection bias	5%	1%
Sample dependence	39%	39%
Data Statistics	2%	3%
Total	39%	39%

ATLAS-COM-CONF-2010-092





- This **data driven** technique can be used for **charged Higgs searches** to estimate **W+jets** background where W(→l ν_l) + jets (a jet fake a l)

Conclusion and Outlook

Conclusion :

- **Charged Higgs** are clear evidence of **beyond** Standard Model physics with early sensitivity
- **W+ jets** can be an important **background** for **charged Higgs** searches
- Several methods to estimate the **W+jets background** with a **data driven estimation**
- Already some estimate using the first **~1 pb^{-1}** data at **$\sqrt{s}$=7 TeV**:
  - **W+jets** background estimation
  - **lepton fake rate** estimation
- Already some **nice results**

Outlook :

- More data to come before the **end of 2010**: 10-100 pb^{-1} (try to reach the luminosity of 10^{32})
- More **methods and results** to come
- **Apply and develop** methods for **charged Higgs** searches

Back-up Slides

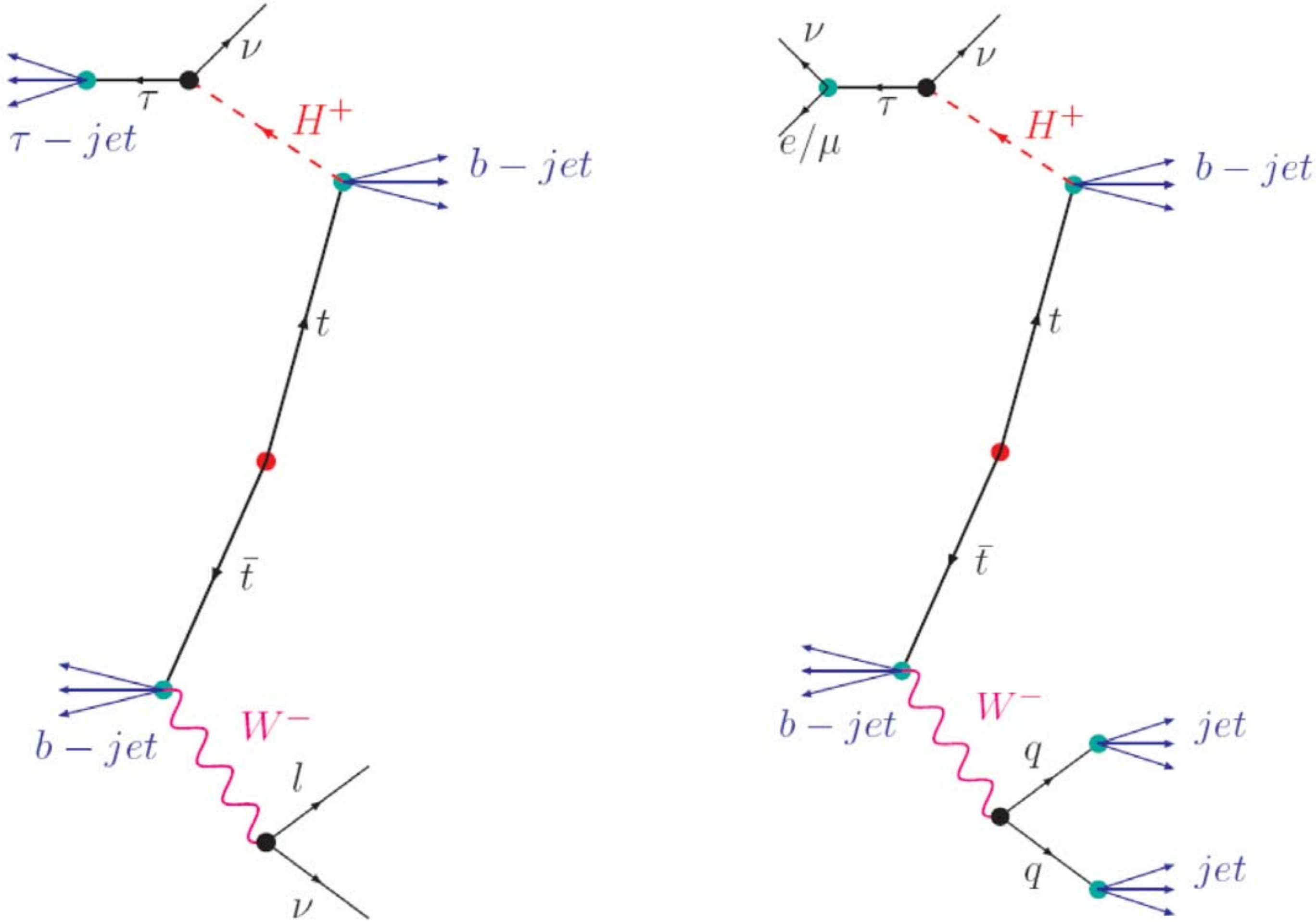
Abstract

- Abstract :
  W+jets background in charged Higgs searches

"W + jets events with one lepton in the final state could be an important background for charged Higgs boson searches. In the $H^+ \rightarrow \tau\nu$ channel for instance, W + jets will contribute to the backgrounds where a light jet fakes a hadronic τ-jet. Below the top quark mass, where the $H^+ \rightarrow \tau\nu$ will appear as an excess of τ leptons compared to electrons or muons from SM model expectation, the understanding of fake τ from background sources including W + jets will be essential to establish the existence of a viable signal. Data driven method for the measurement of W + jets background are needed. In this talk, we will describe the available methods to estimate this background from data and its relevance to charged Higgs searches."

Charged Higgs

- **Feynman** diagrams :



W+jets background

Data driven estimate for W+jets background (applied on MC) :

The expected relative uncertainties on the W+jets background estimation.

	Electron analysis	**Muon analysis**
Statistical for 200 pb^{-1}	11.3%	8.3%
Purity of control samples	17.0%	12.7%
Monte Carlo correction factor	12.1%	12.1%
JES ($\pm$10%)	3.6%	2.3%
JES ($\pm$5%)	3.0%	0.7%
Lepton energy scale	0.4%	0.7%
total uncertainty	23.9%	19.6%

W+jets background

Data driven estimation of the lepton misidentification rate (applied on MC) :

Jets Misidentified as Leptons

The largest systematic associated with jets faking leptons arises from measuring the fake rate in control samples and extrapolating to the signal region. Fortunately, we have two different control samples, and we can look at the variation between the two control samples to estimate how much the fake rates varies in different types of events. The difference between the control samples may be smaller than the difference between the control samples and the signal region. Out of caution we take twice the difference between the two predictions as our systematic.

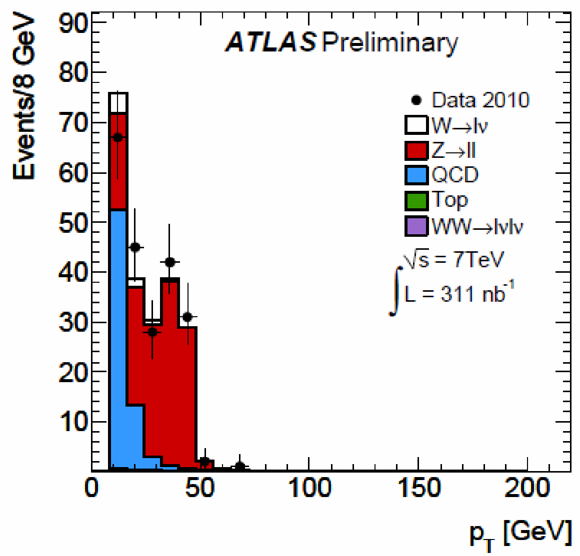
Based on the current studies with QCD Monte-Carlo samples with limited statistics we estimate an uncertainty in the fake rates of 100% for the early data ($50\,\mathrm{pb}^{-1}$) for both electrons and muons; and of 50% for the muons and 100% for the electrons for $100\,\mathrm{pb}^{-1}$ and $200\,\mathrm{pb}^{-1}$.

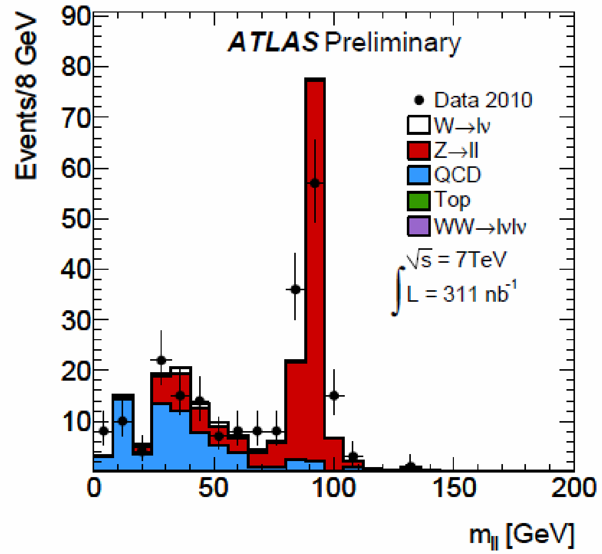
W+jets background

Observation of the Background from W+jets to the $H \rightarrow WW^* \rightarrow l\nu_l\, l\nu_l$ at $\sqrt{s}$=7 TeV :

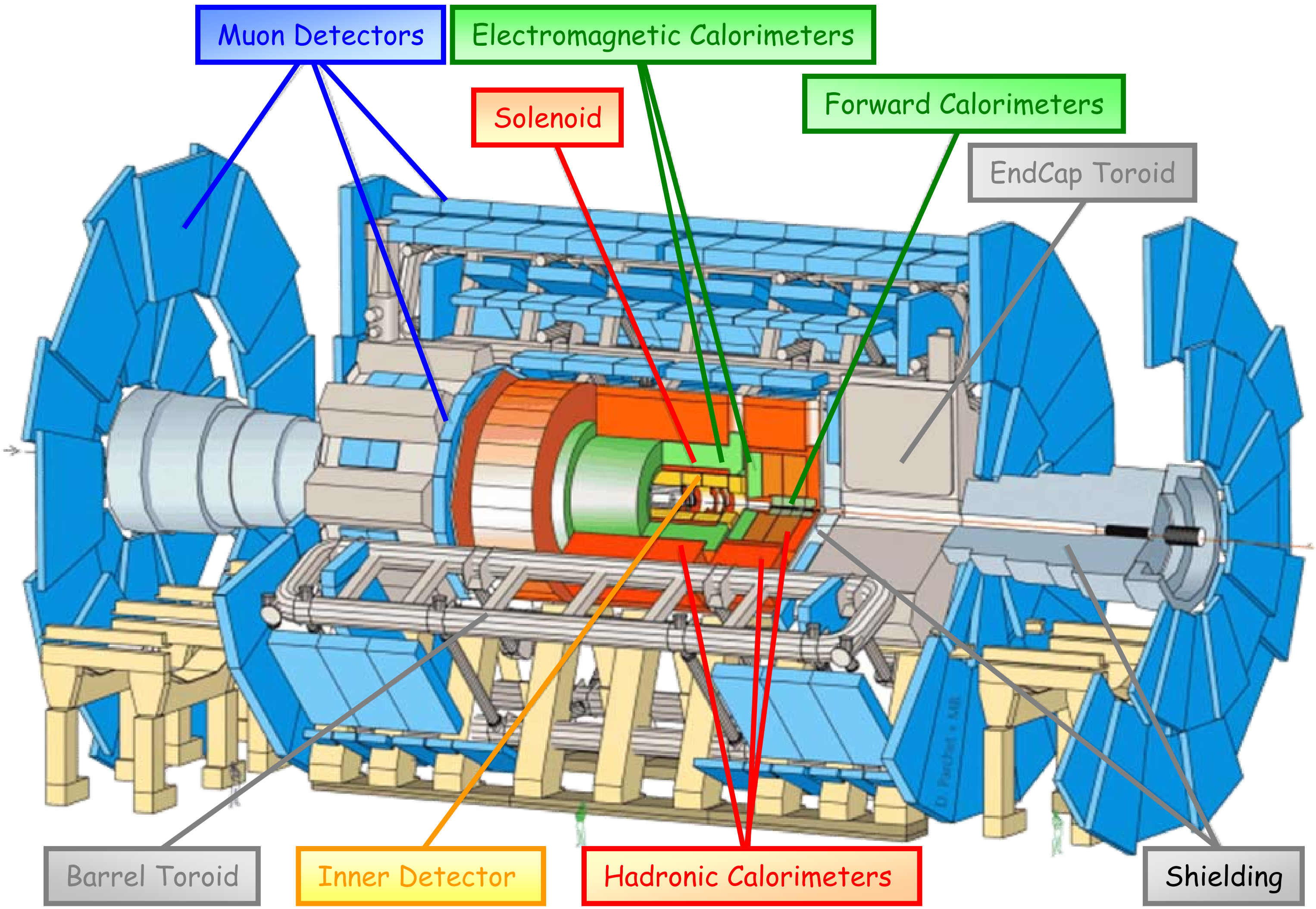
	Loose leptons	Tight leptons
electrons	118972	4171
muons	4414	971

Observed numbers of loose and tight lepton candidates after applying a W/Z veto on events with a jet trigger of 5 GeV p_T threshold.





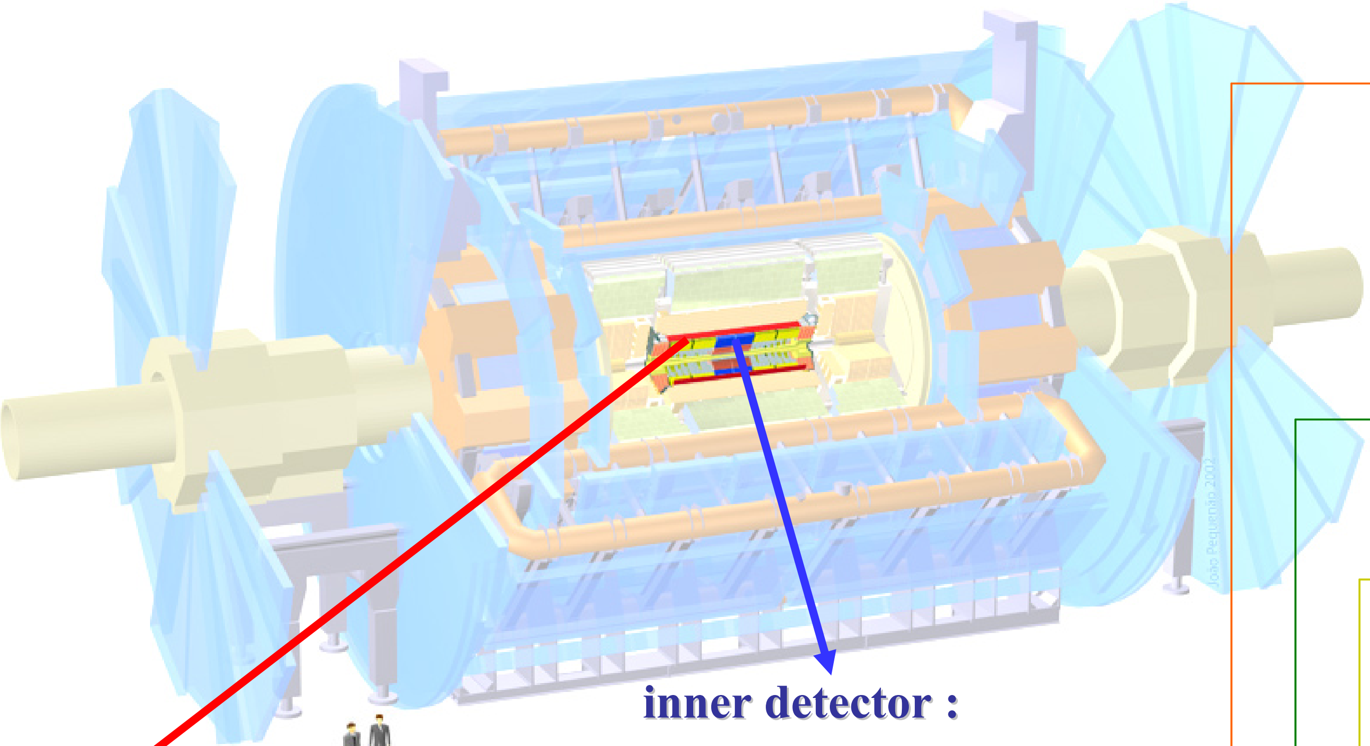
ATLAS Layout

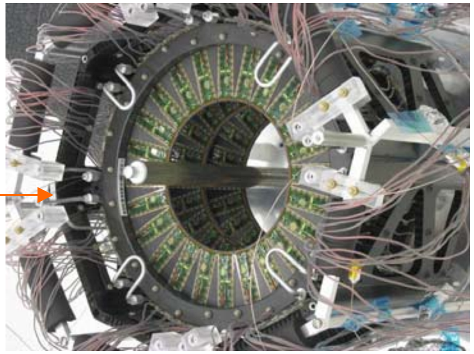


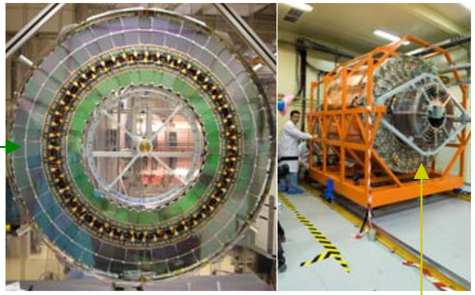
Inner Detector

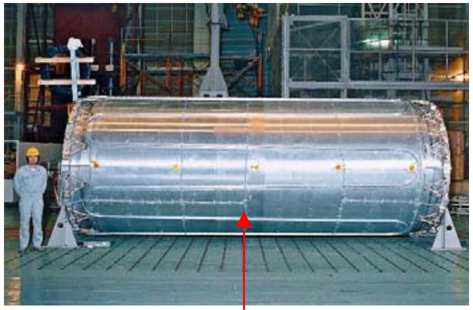
- **Impulsion resolution :**

  $\sigma(p)/p = 0.05\ \%\ p\ (\mathrm{GeV}) \oplus\ 1\%$ for $|\eta|<2.5$









Central Solenoid 2T

inner detector :

pixels detector

SCT (Semi-Conductor Tracker)

TRT (Transition Radiation Tracker)

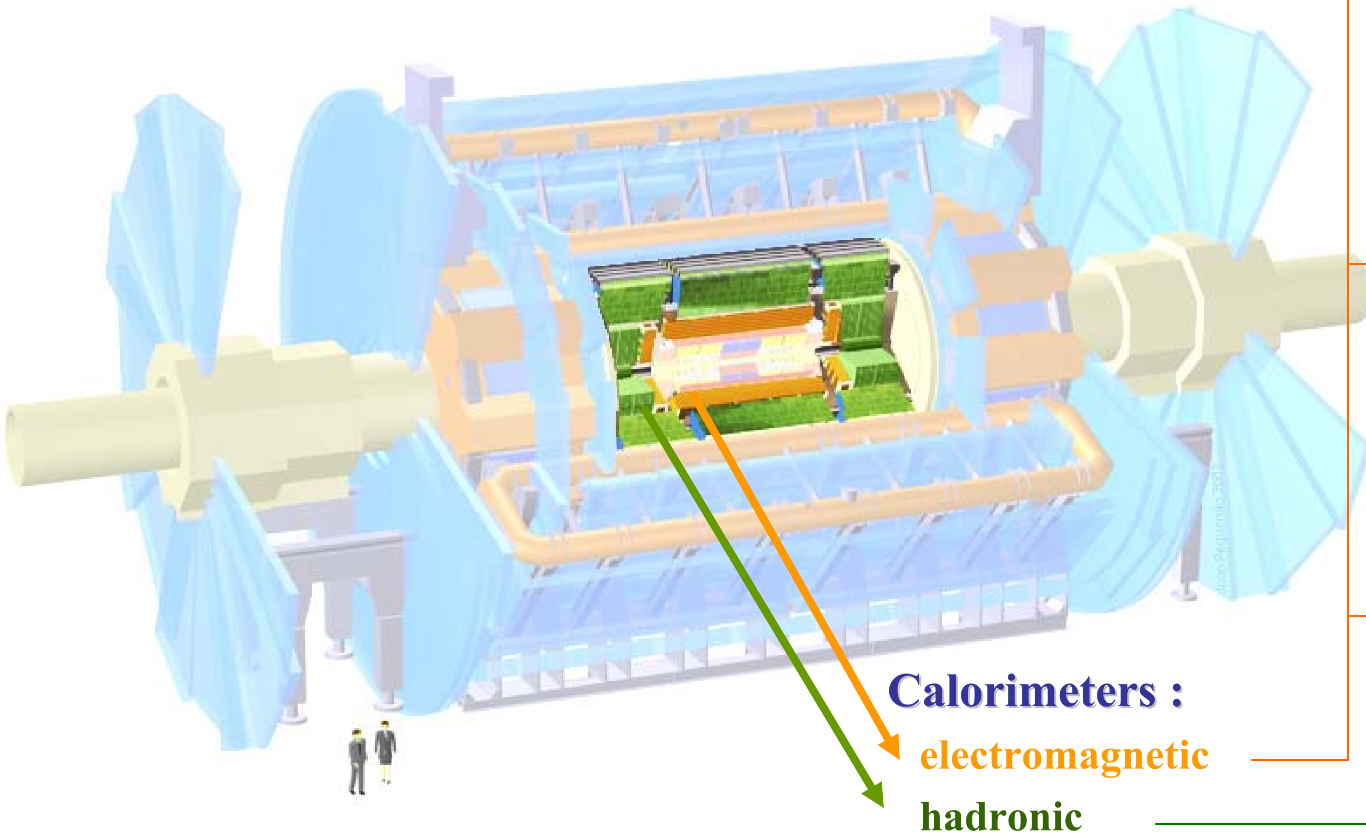
Calorimeters

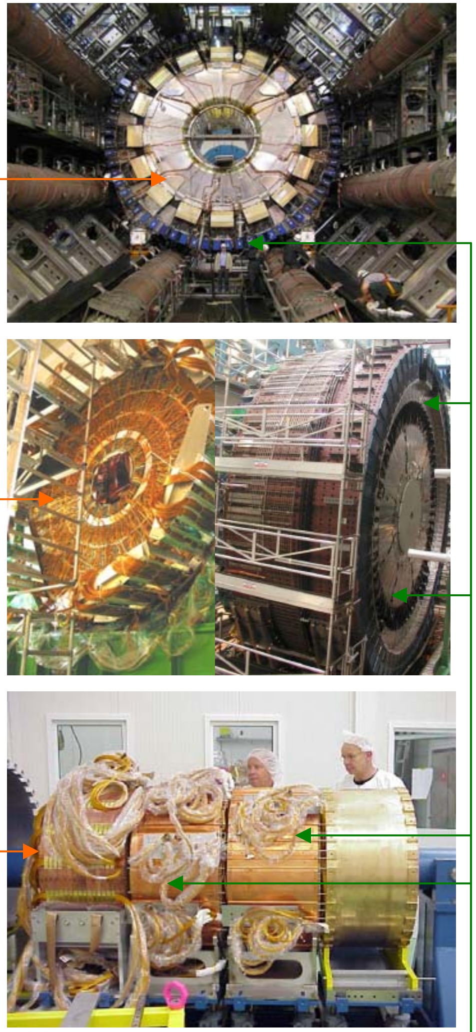
- **Energy resolution (GeV) :**

  electromagnetic : $\sigma(E)/E = 10\%/\sqrt{E} \oplus 0.3/E \oplus 0.7\%$ for $|\eta|<3.2$

  hadronic : $\sigma(E)/E = 50\%/\sqrt{E} \oplus 3\%$ for $|\eta|<3$

  : $\sigma(E)/E = 100\%/\sqrt{E} \oplus 5\%$ for $3<|\eta|<5$



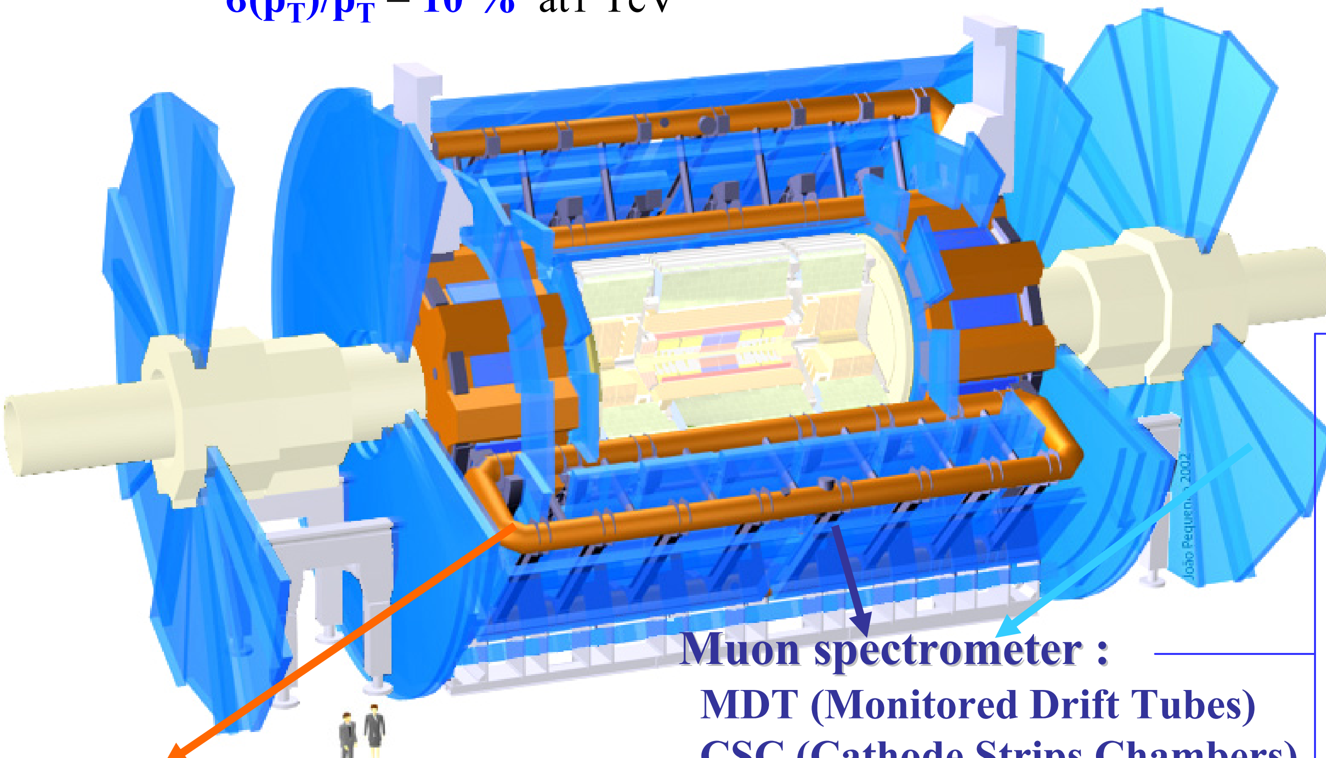


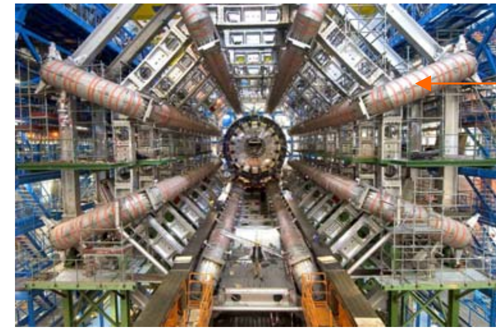
Muon Spectrometer

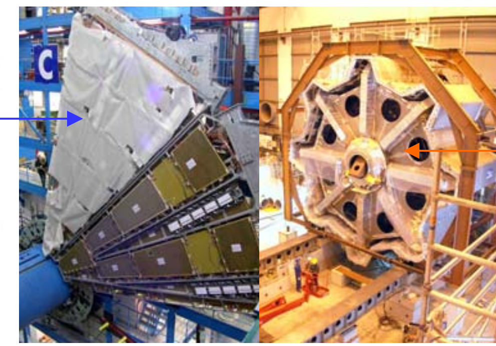
- **Impulsion resolution :**

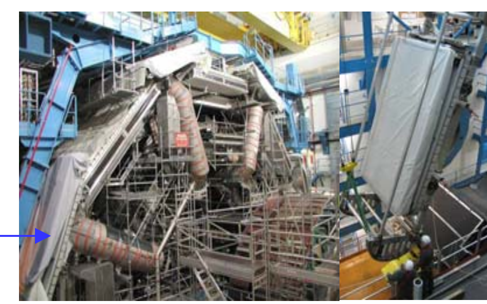
  $\sigma(p_T)/p_T < 3\%$ for $10<p_T<250$ GeV and for $|\eta|<2.7$

  $\sigma(p_T)/p_T = 10\ \%$ at1 TeV









Muon spectrometer :

MDT (Monitored Drift Tubes)
CSC (Cathode Strips Chambers)
RPC (Resistive Plate Chambers)
TGC (Thin Gap Chambers)

barrel toroid: 8 separate coils and 2 end-cap toroids