Charge asymmetry at the LHC and application in top physics within the CMS experiment.

Camille BELUFFI
Supervised by Jérémy Andrea and Marco Cardaci (IPHC Strasbourg)

Introduction

- Master thesis → Internship on behalf of the IPHC Strasbourg, based at CERN.
- **CMS** experiment → **quark top** analysis using the data from **proton-proton collisions in 2010 at the LHC**.
- First : selection of top-antitop events ($t\bar{t}$).
- Then : determination of the main background of the "muon+jets" channel, W+jets → **charge asymmetry method**.

Plan

The LHC, CMS and top quark

Events selection

The charge asymmetry method to estimate the W+jets background

Checking of the method

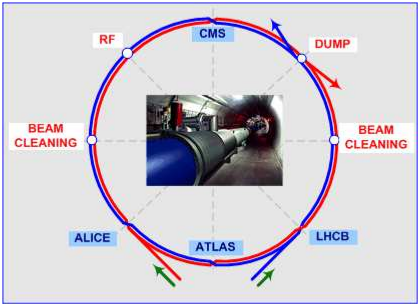
Measure of the charge asymmetry in the data.

W+jets background estimation with the 2010 LHC data.

Perspectives

Conclusion

LHC, "Large Hadron Collider"



- At the France-Switzerland border, 27 km of circumference.
- **Proton-proton** collider $\longrightarrow$ two protons beams collides at a center of mass energy of $\sqrt{s} = 7$ TeV (14 TeV after the next technical stop).
- **Production of particules** not included in the proton : E=mc^2 $\longrightarrow$ conversion of the energy to matter.
- Four big experiments : ALICE, ATLAS, CMS et LHC-b.
- The luminosity on the 30th of October was $\sim$5.22 fb^{-1} (36 pb^{-1} en 2010).

CMS, "Compact Muon Solenoïd"

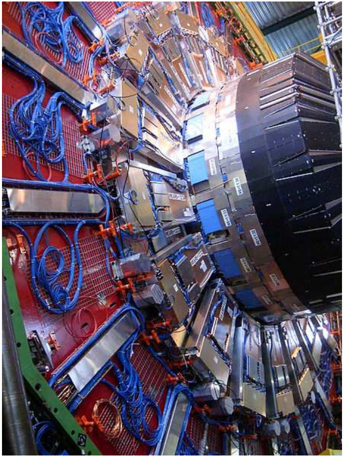
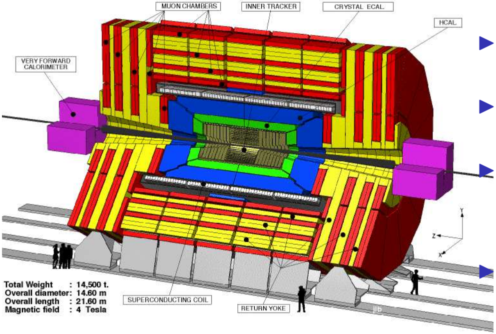


FIG.: CMS endcap.

- Particles detector,
- 14,000-tonnes (Eiffel Tower weight),
- 21 meters of length, 16 meters of diameter,
- In 100 meters under the french village Cessy.

CMS, "Compact Muon Solenoïd"



- **tracker**, closest part to the beam (silicium detectors)
  $\longrightarrow$ detection of charged particles.
- **ECAL** lead tungstene cristals
  $\longrightarrow$ detection of electrons and photons.
- **HCAL** copper layers and plastic scintillators
  $\longrightarrow$ detection of hadrons.
- **Supraconductor solénoïd magnet**
  homogenious magnetic field of 3.8 Tesla
  $\longrightarrow$ bend the charged particles trajectory to measure their transverse impulsion.
- **muons chambers** gaseous detectors
  $\longrightarrow$ detection of muons.

Top physics

Quark top physics :

- **Heaviest** quark of the SM : 173.3±1.1 GeV/c^2
- Decays in $W+b$ **before hadronizing**.
- **New physics research** : many new physics events show the same signature than $t\bar{t}$ events.

Top quark

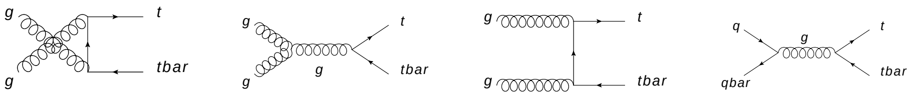


FIG.: Production of $t\bar{t}$ for p-p collisions (strong interaction) : gluons fusion (70%) or $q\bar{q}$ annihilation (30%)

Top decays in 100% of cases in $W+b$, and the W decays either in lepton-neutrino, either in quarks pair. There are 3 final states for $t\bar{t}$ events :

- **"dileptonic"** channel(10%)
- **"lepton+jets"** channel(45%)
- **"all hadronic"** channel(45%)

Signal and background

- channel **"lepton+jets", "muon+jets".**

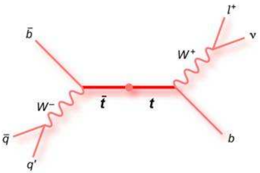


FIG.: Production of a $t\bar{t}$ pair in the "lepton+jets" channel.

- Events with the same particles in the final state or because of bad reconstruction $\rightarrow$ same signature than a "muon+jets $t\bar{t}$" event signature $\Rightarrow$ **background**.
- backgrounds : mostely **W+jets**, $t\bar{t}$ (else than "muon+jets"), QCD et single top.

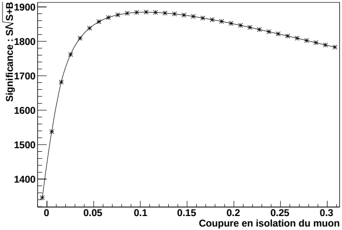
Signal events selection

Selection with a "muon+jets" topology :

- One good "trigger" muon with $p_T > 9$, 11 or 15 GeV/c,
- **Only one muon** isolated with $p_T > 20$ GeV/c and $|\eta| < 2.1$.
- no second muon with $p_T > 20$ GeV/c and $|\eta| < 2.1$,
- **at least 3 jets** with $p_T > 30$ GeV/c and $|\eta| < 2.4$. Cleaning jet applied to reject the jets reconstructed as leptons.
- **No electron**.

Isolation variable : optimisation

Cut on the isolation variable applied in 2010 CMS analysis : RelIso < 0.05.
For this analysis $\rightarrow$ otpimisation.
Significance $= S/\sqrt{S+B}$ (B the number of QCD events and everything else considered as signal).
Maximum of the plots $\rightarrow$ which cut is the better compromise between background rejection and honest statistical error $\rightarrow$ here : RelIso < 0.1.



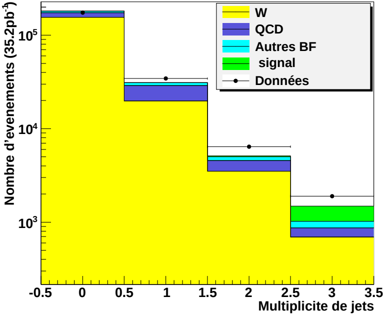


FIG.: Jet multiplicity distribution for Monte Carlo predictions (histograms normalised at 35.2 pb−1), compared to the data (black dots).

$\Rightarrow$ We can see that the **Monte Carlo** underestimates the data $\Rightarrow$ **MC not reliable** for background determination.

The charge asymmetry method

p-p collisions at the LHC $\rightarrow$ more W^+ producted than W^- (PDF).
$\rightarrow$ **Charge asymmetry** = number of selected events with a muon $\neq$ number of selected events with an antimuon.

- $$A = \frac{N_W^+ - N_W^-}{N_W^+ + N_W^-} = \frac{1}{R}, \quad (1)$$
  with N_W^+ and N_W^- the number of events that contain W^+ and W^-.
- muon+jets channel $\rightarrow N_{tot}^{\mu\pm} = N_{t\bar{t}}^{\mu\pm} + N_W^{\mu\pm} + N_{otherBF}^{\mu\pm}$
- Hypothesis : no charge asymmetry in $t\bar{t}$ and in the other backgrounds but same charge asymmetry for W+jets and single top :
  $$N_{W,ST}^{tot} = R \times (N_{tot}^{\mu+} - N_{tot}^{\mu-}). \quad (2)$$
  with $N_{tot}^{\mu+} - N_{tot}^{\mu-}$ measured in the data.
- Study of the correlations between R and the events selection, R **will then be estimated from the data.**

R in W+jets

R depends on the transverse impulsion and pseudo-rapidity of the muon.

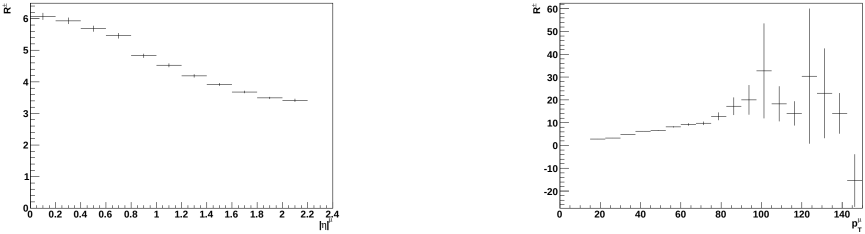


FIG.: R as a function of η (left) and p_T (right), for MC W+jets events.

R in W+jets

R depends also on the jets multiplicity.

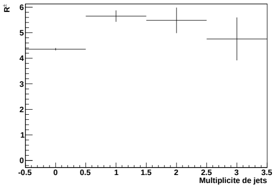
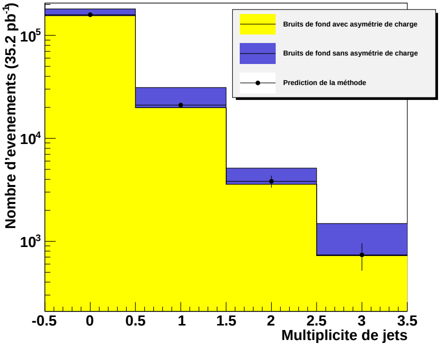


FIG.: R dependency as a function of the jets multiplicity for MC W+jets events.

- The R dependency as a function of the partons is discussed in Eur.Phys.J.C (2010) 69 :67-73.
- These dependencies must be taken in account $\rightarrow$ **same cuts** applied in the determination of R and for the events selection.

Checking of the method $\rightarrow$ R **determined from MC samples** with charge asymmetry $\rightarrow$ **applied on the MC**$\rightarrow$ matching between the yellow histograms and the dots.



Jets multiplicity	0	1	2	$\geq$3
MC truth (histograms)	154871	19849	3579	722
Method prediction (black dots)	158679	21030	3820	736

Measure of R from the data.

1. Different events selection **in the data** to **cut** most of the QCD (cut RelIso $<$ 0.005) and the more $t\bar{t}$ as possible (Njets=0) $\rightarrow$ **sample rich in W+jets.**
2. **Correction for the rest of the contaminations with the MC.** QCD contamination estimated from the data $\rightarrow$ matrix method.
3. R measured at 0 jets $\rightarrow$ propagation to higher jet multiplicities using **MC**, by doing the hypothesis that the **ratio** between R at 0 jet and R at i jets is **the same in the MC and in the data**. We then obtain :

$$\rightarrow R_{corr}^{data}(njet = i) = \underbrace{\frac{R_W^{MC}(njet = i)}{R_W^{MC}(njet = 0)}}_{\alpha} \times R_{corr}^{data}(njet = 0) \quad (3)$$

Matrix method

$\Rightarrow$ Estimation of QCD events for RelIso<0.005.

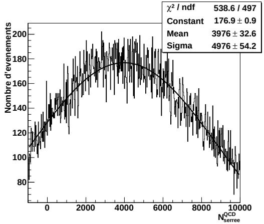
- Selection with a loose isolation (RelIso<0.5) and with a tight isolation (RelIso<0.005).
-

$$N_{loose} = N_{loose}^{QCD} + N_{loose}^{signal} \tag{4}$$
$$N_{tight} = N_{loose}^{QCD} \times \epsilon_f + N_{loose}^{signal} \times \epsilon_s \tag{5}$$
$$\tag{6}$$

- ϵ_f : probabilty that a fake lepton passes the tight cut
  ϵ_s : probability that a real lepton (isolated) passes the tight cut
- $\Rightarrow N_{tight}^{QCD} = N_{loose}^{QCD} \times \epsilon_f$
- $\Rightarrow \epsilon_f$ et ϵ_s estimated from the MC$\rightarrow$ large uncertainty on ϵ_f.

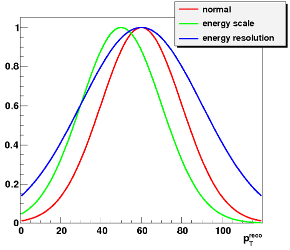
"Toy MC" method to estimate the uncertainties on $N_{tight}^{QCD/signal}$:

- Statistical uncertainty $\rightarrow$ decorrelate $N_{tight/loose}$ and follow a Poissonian law
- Systematic uncertainty due to $\epsilon \rightarrow$ variation of ϵ_f and ϵ_s following Gaussian laws centrée sur $\epsilon_{f/s}$ et de largeur $\Delta\epsilon_{f/s}$ (erreur sur ϵ_s de 1% et erreur sur ϵ_f de 100%
- Equation system (**??**) resolved 100 000 times $\rightarrow$ fill in histograms with N_{tight}^{QCD} et N_{tight}^{signal}
- Adjusting the histograms with a Gaussian $\rightarrow$ $\mu=$ number of events, σ = uncertainty on that number



Uncertainties on R

2 types of uncertainties : **statistical** et **systematic**.
Sources of the systematic uncertainties are due to the use of MC to predict the variation of R :



- Uncertainty on the jet resolution (JER).
- Uncertainty on the jet scale (JES).
- Uncertainty on the proportion of gluons and quark in the initial proton (PDF).
- Uncertainty related to the R propagation as a function of Njets.
- Uncertainty on the QCD contamination in R^{data}_{corr}.
- Fraction of events with charge asymmetry but W+jets.

Uncertainties on R

Uncertainty source (%)	0 jet	1 jet	2 jets	≤3 jets
QCD ($R^{data}_{corr}(njet = 0)$)	3.4	3.1	3.2	3.3
MC ($R^{data}_{corr}(njet = 0)$)	4.2	3.8	4.0	4.1
JES	0	0.8	2.2	7.4
JER	0	< 0.1	2.0	1.0
ECA	0	0.6	2.4	4.1
PDF	10.6	9.6	9.9	10.4
Statistical uncertainty on α[4]	0	7.9	18.5	36.5
Statistical uncertainty on $N^{tot}_{W,TS}$[5]	1.8	6.1	15.6	37.1
Total uncertainty on $N^{tot}_{W,TS}$	**12.0**	**14.7**	**26.9**	**54.0**

TAB.: Differents uncertainty sources (%) that have been propagated to the value of $N^{tot}_{W,TS}$. The last ligne was obtained by summing quadratically all the uncertainties.

[4] Uncertainty due to the lack of MC W+jets to determine R with precision.

[5] Uncertainty due to the lack of data.

Method tested on the **LHC 2010 data** : collisions at $\sqrt{s} = 7$ TeV with an integrated luminosity of 35.2 pb^{-1}.

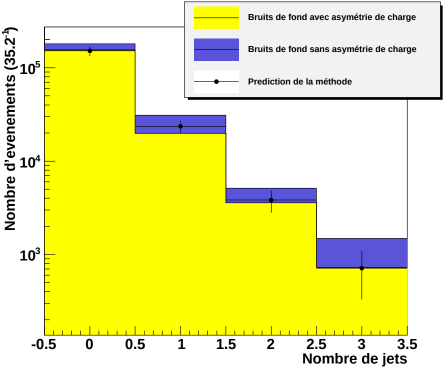


FIG.: Jet multiplicity distribution for MC prediction (histograms) and for LHC data on which the method was applied (black dots).

The dots should match the yellow histogram.

For Njets ≥ 3, the method predicts 713.7 ± 385.3 and the MC truth worths 722.4 ± 7.2.

Improvements

- Main uncertainty sources : 1)not enough W+jets MC $\rightarrow$ large uncertainty on the propagation factor α, 2)not enough data $\rightarrow$ large statistical error.
- projection : if the luminosity would be higher and with more MC :

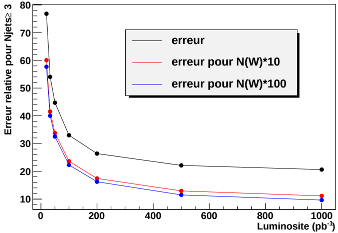


FIG.: $\Rightarrow$ With 100 times more and an integrated luminosity of 1 fb^{-1} $\rightarrow$ Final error is about 10%.

Conclusion

- $t\bar{t}$ "muon+jets" channel $\rightarrow$ selection.
- Background : how many events after selection ? $\rightarrow$ method based on the charge asymmetry of the W, uses the data.
- Method never tested before.
- Measure of the charge asymmetry for 0 jets than propagation to higher jets multiplicities using MC (onlu MC input).
- Estimation of the number of W+jets events : promising results.
- Main uncertainty sources : lack of MC W+jets and data,
- but the uncertainty drops when the integrated luminosity increases and with more MC $\rightarrow$ we can reach 10% of uncertainty.
- Uncertainties about the PDF have to be done properly.

Thank you for your attention ! Questions ?

BACK UP

PDF

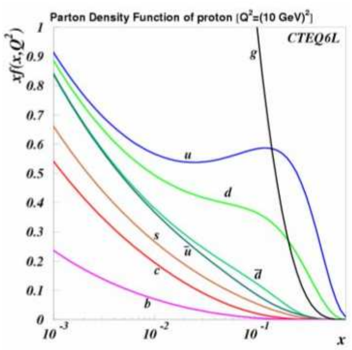


FIG.: Proton Density Function : probability of presence for a parton as a function of the total energy of the proton.

Backgrounds

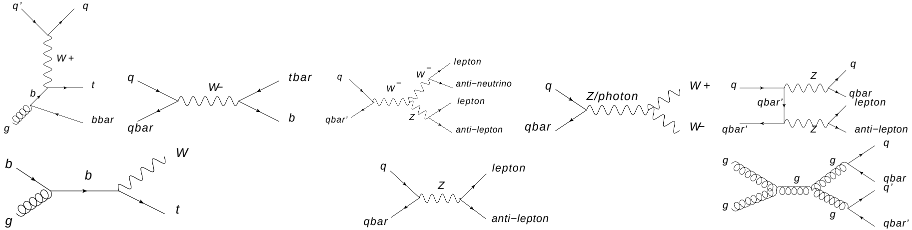


FIG.: Production channel for single top t-like, s-like, WZ, WW, ZZ, single top tW, Z+jets and QCD.

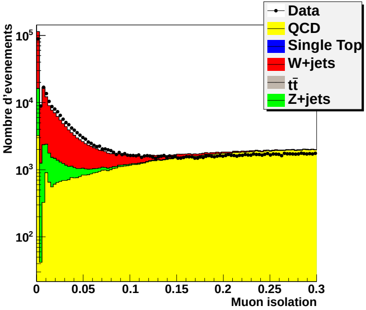
Channel	Cross sections (pb) for a collision energy of 7 TeV
QCD	84679
W+jets	31314
Z+jets	3048
$t\bar{t}$	157.5
Top solitaire voie t	20.9
Top solitaire voie tW	10.6
Top solitaire voie s	4.6
WW	4.5

Isolation variable

Defined as following :

$$RelIso = \frac{TrkIso + EcalIso + HcalIso}{p_T^{lepton}} \quad (7)$$

where TrkIso is the sum of the tracks p_T in a cone with a radius $\Delta R = 0.3$ around the lepton, EcalIso+HcalIso is the sum of the transverse energies from the calorimetric towers, excluding the track of the lepton considered. The sum is divided by the lepton' p_T to make the variable more independant.



Equation for the computing of the total uncertainty on $N_{W,TS}^{tot}$

$$|\Delta N_{W,TS}^{tot}| = \sqrt{(N_{tot}^{\mu+} - N_{tot}^{\mu-})^2 \times \Delta R^2 + R^2 \times (N_{tot}^{\mu+} + N_{tot}^{\mu-})} \quad (8)$$