Study of the coupling of the Higgs boson to the top quark in the ATLAS experiment

A.Chomont

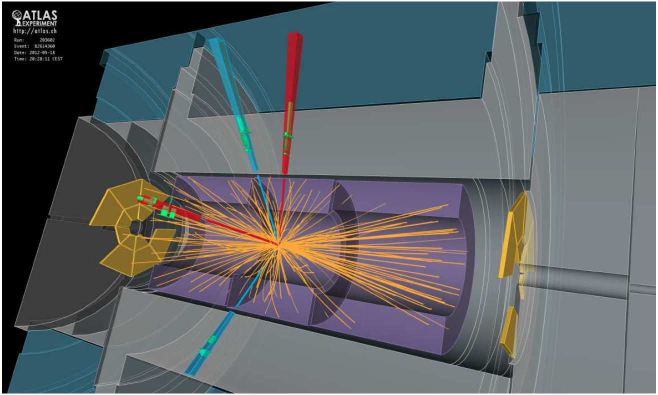
JRJC

LPC Clermont-Ferrand

November 19th 2015

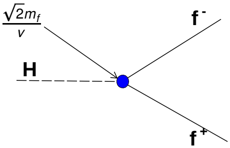
Motivation

- Higgs boson, discovered in July 2012 by the ATLAS and CMS experiments at the LHC
- Properties (spin, parity, **couplings**) ?



Standard Model and Higgs coupling to fermions

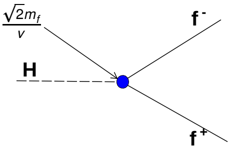
- Yukawa coupling of Higgs to fermions proportional to fermions mass
- For now, only observation of coupling of Higgs to fermion: $H \rightarrow \tau\tau$



with $\nu \simeq 246 GeV$ the vacuum expectation value

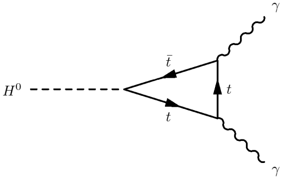
Standard Model and Higgs coupling to fermions

- Yukawa coupling of Higgs to fermions proportional to fermions mass
- For now, only observation of coupling of Higgs to fermion: $H \rightarrow \tau\tau$



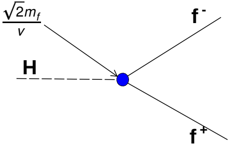
with $\nu \simeq 246 GeV$ the vacuum expectation value

- Top quark, heaviest fermion $\rightarrow$ should couple strongly to Higgs boson (coupling at around 1)
- Coupling already indirectly observed in the case of SM Higgs decaying in two photons (we assume that there is no new physics in the loop)



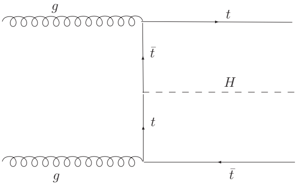
Standard Model and Higgs coupling to fermions

- Yukawa coupling of Higgs to fermions proportional to fermions mass
- For now, only observation of coupling of Higgs to fermion: $H \rightarrow \tau\tau$



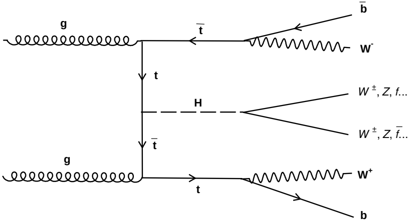
with $\nu \simeq 246 GeV$ the vacuum expectation value

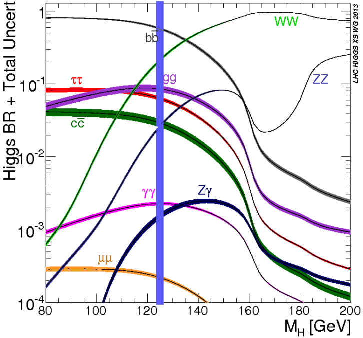
- Top quark, heaviest fermion $\rightarrow$ should couple strongly to Higgs boson (coupling at around 1)
- The goal is to do a direct measurement of this coupling at the tree level



Choice of the multileptonic signature

- Higgs boson and top quark not stable → we observe the product of their decay in the detector
- Top quark decays almost exclusively in a W boson and a b quark
  - W boson will decay to a pair $q\bar{q}'$ or to $l\bar{\nu}$

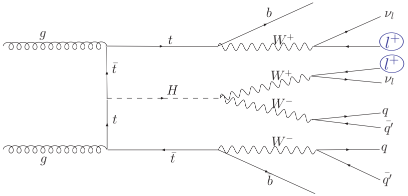




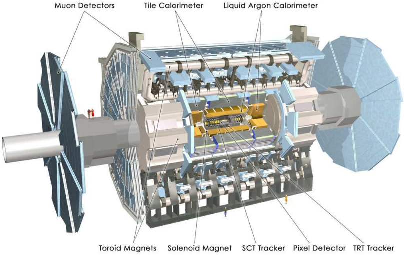
- Multileptonic signature separated in charge and flavours (e, μ, τ) can be used for search of $t\bar{t}H$

Multileptons signature

- Decays of Higgs to WW, ZZ and $\tau\tau$ targeted
- 8 different channels can be considered for Run 2
  - Light leptons channels: 2lSS, 3l, 4l
  - Light+tau channels: 2τ+1l, 2lSS+τ, 2lOS+τ, (l+τ)SS, 2τ+jets
- Focus on 2lSS channel
  - Description of estimation of data-driven background



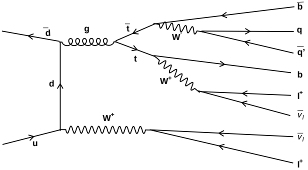
Experimental environment and personal contribution

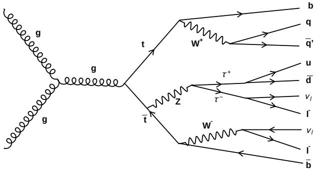


- ATLAS detector at the LHC
  - Work on the calibration of the TileCal using Laser system

- Run 1 at 8TeV in ATLAS: $20fb^{-1}$of data, discovery of Higgs boson
  - Personal work on Run 1: Test of sensibility on the $t\bar{t}H$ signal
- Run 2 at 13TeV in ATLAS: $4fb^{-1}$this year, more luminosity expected in the future, Higgs production
  - Personal contribution on Run 2 ongoing $t\bar{t}H$ analysis: data-driven backgroung estimation, fitting tools

Backgrounds

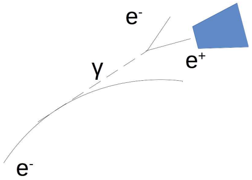




- $t\bar{t}W$ process
- $t\bar{t}Z$ process
  - Few SM processes with similar signatures
    - True physical same-sign background: $t\bar{t}W$, $t\bar{t}Z$, VV estimated from MC simulation
    - Instrumental backgrounds estimated from the data (mainly $t\bar{t}$ events)
      - Fake leptons (jets or secondary lepton from B-decay reconstructed as primary electron)
      - Electrons with a mis-identification of the charge (Charge MisId)

Charge MisId estimation

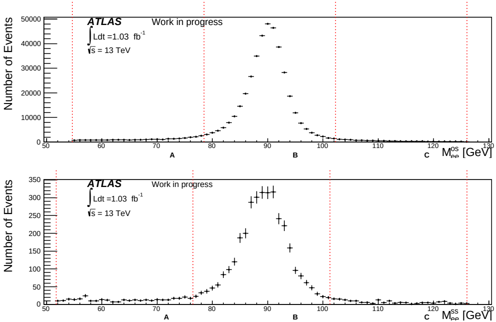
- Mis-identification of the charge of a lepton is an important background originating from two processes
  - High p_T electron with straight track
  - Trident process with an electron radiating a photon converting to a pair of electrons
- Negligible effect on muon



- Results shown afterward using $1fb^{-1}$of data, from the Run 2 of the LHC

QMisid rates estimation

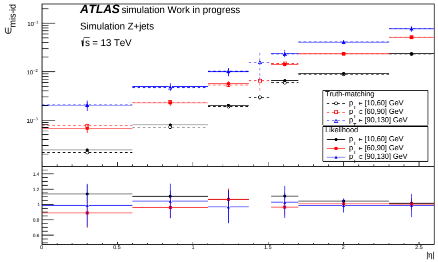
- Rate of QMisid computed from $Z \to e^+e^-$ mass peak region and used to reweight OS data
- Background substraction done using a side-band method: $N_Z = n_B - \frac{n_A+n_C}{2}$



- The QMisid rate is defined as $\frac{N_{SS}}{N_{OS}+N_{SS}}$
    - Supposition that rates are independent of the physical characteristics (energy, momentum . . .) of the electron

QMisid: method for rates estimation

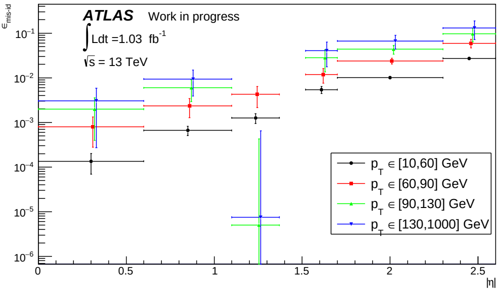
- ϵ_i rate of charge Misid for a single electron in region i (regions defined in η, p_T,E ...) and we obtain for N_{tot} true opposite-sign events:
$$N_{ss} = N_{tot}[(1-\epsilon_i)\epsilon_j + (1-\epsilon_j)\epsilon_i] \simeq N_{tot}(\epsilon_i + \epsilon_j)$$
- The rates, ϵ_i and ϵ_j, are obtained by likelihood minimization and are highly dependent on the choice of the binning



- Closure test: good agreement between rates from LH method and truth matching

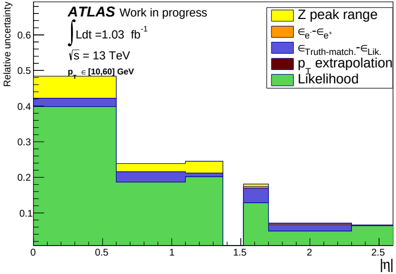
QMisid: First Rates estimation for Run 2

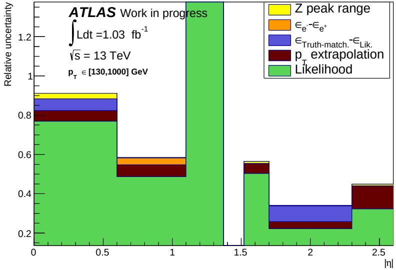
- Rates obtained using Likelihood method from $1fb^{-1}$of data
- Rates for last bin in p_T obtained by extrapolation of rates in the next to last bin in p_T (bin [90,130]GeV)



- Large uncertainties particularly in p_T bin [90,130] and $|\eta|$ bin [1.1,1.37] $\rightarrow$ to be improved with full 2015 statistics

Systematics

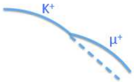




- Uncertainties include:
  - Statistical uncertainty from the likelihood method
  - Statistical uncertainty on the p_T dependent correction factor (last p_T bin, p_T >130GeV)
  - Difference between rates from truth matching and likelihood method on Z samples
  - Stability of rates due to definition of Z-peak region definition

Fakes rate estimation

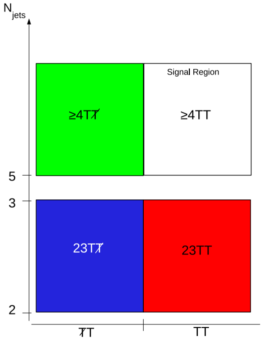
- Leptons fakes are objects reconstructed as prompt leptons, leptons coming from a W boson, a Z boson or a τ (decay results of top or Higgs)
  - jets
  - Non prompts leptons due to decays of b-hadrons for example
  - Trident process with an electron radiating a photon converting to a pair of electrons
- Fakes impact both muons and electrons



- Estimation as for the QMisid done directly from the data

Fake factor method

- Four regions defined based on jet multiplicity and 2 leptons categories (Tight T, anti-Tight $\not{T}$)



- Tight lepton: object used in the analysis which can be fakes
- Anti-Tight lepton: almost all fakes

- Ratio of $TT/T\not{T}$ estimated in region without signal and supposed to be independent w.r.t the number of jets
- Then the ratio is applied in the high multiplicity region

Fake factor method

- Fake factor θ is defined as (for electrons):

$$\theta_e = \frac{TT}{T\not{T}}(2-3jets) = \frac{TT(N_{ee}^{data} - N_{ee}^{PromptSS} - N_{ee}^{QMisId})}{T\not{T}(N_{e\not{e}}^{data} - N_{e\not{e}}^{allPrompt})}$$

- PromptSS: $t\bar{t}V$, VV
- QMisId: prompt opposite-sign events with a charge mis-identification (data-driven in TT region)

- In the case of $\mu^{\pm}\mu^{\pm}$ channel, same definition of θ_{μ} as for θ_e (without the QMisId terms)

Fake factor method

- Number of fakes in signal region obtained from θ_e, θ_μ
  - for $e^\pm e^\pm$ region: $N_{ee}(njets) = N_{e\not{e}}(njets) \times \theta_e$
  - for $\mu^\pm \mu^\pm$ region: $N_{\mu\mu}(njets) = N_{\mu\not{\mu}}(njets) \times \theta_\mu$
  - for $e^\pm \mu^\pm$ region:
    $N_{e\mu}(njets) = N_{e\not{\mu}}(njets) \times \theta_\mu + N_{\mu\not{e}}(njets) \times \theta_e$

Systematic uncertainties

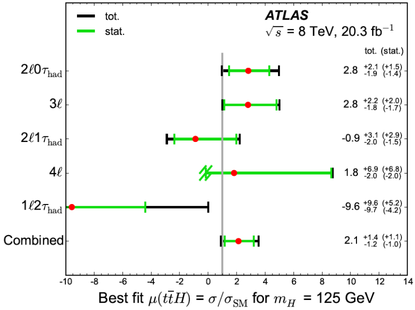
- Validity of the extrapolation flow 2-3 jets region to $\geq$ 4jets region
  - Closure test performed on simulated $t\bar{t}$ events
  - Comparison of real ss fakes in signal region to number predicted by $N_{l\not{l}} \times \theta$
- Uncertainty on substracted backgrounds (QMisId, PromptSS)
- Composition of 2-3 jets region
  - Presence of additional non-$t\bar{t}$ fake sources, prompt processes w.r.t signal region $\rightarrow$ bias on the θ estimation
  - Estimated by changing definition of low multiplicity region adding supplementary selection for example

	4 jets	$\geq$5 jets
$e^{\pm}e^{\pm}$	37.4 (35.4)	38.1 (36.2)
$\mu^{\pm}\mu^{\pm}$	37.8	37.9
$e^{\pm}\mu^{\pm}$	27.2 (26.1)	28.0 (27.1)

Statistical uncertainties [%] on fake estimate in Run 2 (5 fb^{-1})

Statistical treatment

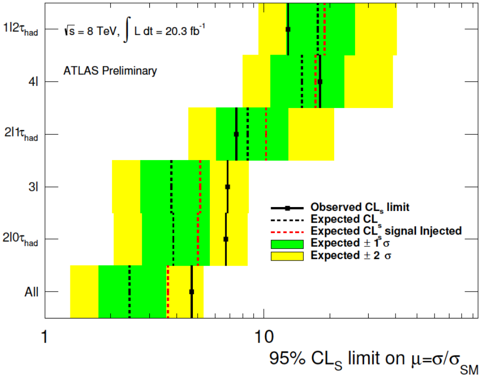
- After choice of a signal region and estimation of background, a fit on data is performed
- Signal strength defined as $\mu = \frac{\sigma_{meas}}{\sigma_{SM}}$
- Maximum likelihood fit of μ with floating systematic uncertainties used to obtain the observed value
- The final result combines the sensitivity obtained in all $t\bar{t}H$ multilepton channels



Phys. Lett. B 749 (2015) 519-541

Limit setting and result stability

- In the case of Run 1, where the $t\bar{t}H$ was not seen a limit is set on μ
- In the case of a limit on μ below 1 it means the hypothesis can be rejected



ATLAS-CONF-2015-006

- Cross-check of the result stability versus background cross-section performed

$$\mu(t\bar{t}H) = 2.1 - 1.4(\frac{\sigma(t\bar{t}W)}{232fb} - 1) - 1.3(\frac{\sigma(t\bar{t}Z)}{206fb} - 1)$$

Conclusion

- Run 1 $t\bar{t}H$ multileptons results
  - Search for $t\bar{t}H$ in multileptonic final states performed: $\mu = 2.1^{+1.4}_{-1.2}$
  - 2lSS channels one of the most sensitive one: $\mu = 2.8^{+2.1}_{-1.9}$
  - First personal participation on an analysis (cross-check of the resuts stability, test of signal sensibility)
- Run 2 ongoing $t\bar{t}H$ multileptons analysis
  - Estimation of data-driven background
  - Development of fitting tools and framework
- Run 2 data analysis on-going with the observation of $t\bar{t}H$ process expected before the end of the Run 2 of the LHC

Backup

QMisid: Likelihood method

- ϵ_i rate of charge Misid for a single electron in region i (regions defined in η, p_T,E ...) and we obtain for N_{tot} true opposite-sign events:
  $$N_{ss} = N_{tot}[(1-\epsilon_i)\epsilon_j + (1-\epsilon_j)\epsilon_i] \simeq N_{tot}(\epsilon_i + \epsilon_j)$$
- Then we suppose that all same-sign events in Z peak are produced by QMisid $\rightarrow$ N^{ij}_{SS} described by Poisson distribution
- From this the probablity for both electrons to produce a charge flip is
  $$P(\epsilon_i, \epsilon_j | N^{ij}_{SS}, N^{ij}) = \frac{[N^{ij}(\epsilon_i+\epsilon_j)]^{N^{ij}_{SS}} e^{-N^{ij}(\epsilon_i+\epsilon_j)}}{N^{ij}_{SS}!} (= L_{i,j})$$
- The likelihood is then $L(\epsilon | N_{SS}, N) = \prod_{i,j} L_{i,j}$
- The rates, ϵ_i and ϵ_j, are obtained by minimizing the likelihood and are highly dependent on the choice of the binning

QMisid: Extrapolation in p_T

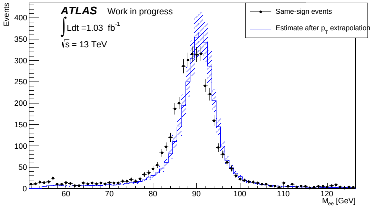
- Rates for last bin in p_T obtained by extrapolation of rates in the next to last bin in p_T (bin [90,130]GeV)
- p_T dependent correction factor extracted from $t\bar{t}$ events
- So rates in last bin obtained by:

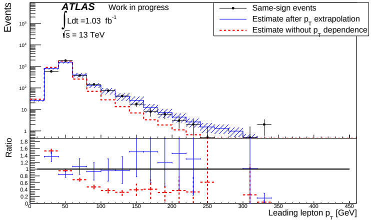
$$\epsilon(|\eta|,p_T > 130GeV) = \epsilon(|\eta|,p_T \in [90,130]GeV) \times \alpha_{t\bar{t}}(|\eta|,p_T > 130GeV)$$

- with $\alpha_{t\bar{t}}$ being defined only in the highest p_T bin as:

$$\alpha_{t\bar{t}}(|\eta|, p_T) = \frac{\epsilon(|\eta|,p_T)_{t\bar{t}}}{\epsilon(|\eta|,p_T \in [90,130]GeV)_{t\bar{t}}}$$

- The statistical uncertainty on α is taken as a systematic uncertainty for the final result





Materials in ATLAS detector