Top quark mass reconstruction in the semi-leptonic channel using the Globalχ^2 algorithm

María Moreno
Regina Moles
Vicente Lacuesta
Carlos Escobar

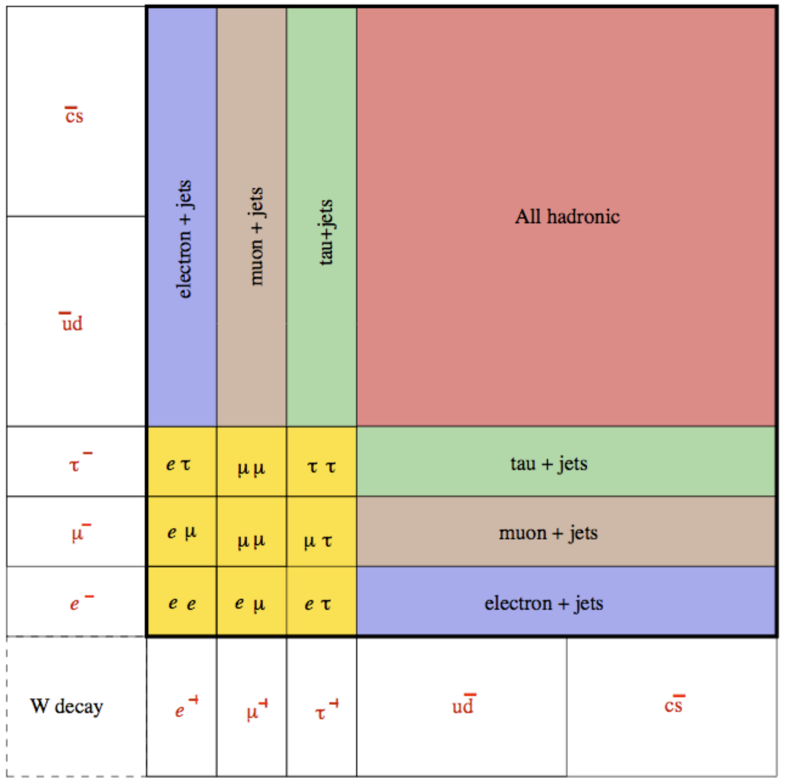
Introduction

Goals

- Contribute to the studies of the ATLAS potential to measure the top quark mass
- Introduce the GlobalChi2 method which we have used for the ATLAS Silicon Tracker alignment for years.
  - Nowadays, GlobalChi2 is the baseline alignment algorithm for the ATLAS ID alignment.
- This talk presents MonteCarlo studies for the LHC scenario of 14 TeV, for a luminosity of 10^{33} cm^{-2} s^{-1} and integrated lumi of 1 fb^{-1}.
  - The MC samples are the ones generated in the context of the comprehensive ATLAS review (aka as CSC Book):
  Expected Performance of the ATLAS Experiment : Detector, Trigger and Physics.
  ATLAS Collaboration. arXiv:0901.0512 ; CERN-OPEN-2008-020

Selecting the decay channel for *ttbar* events...

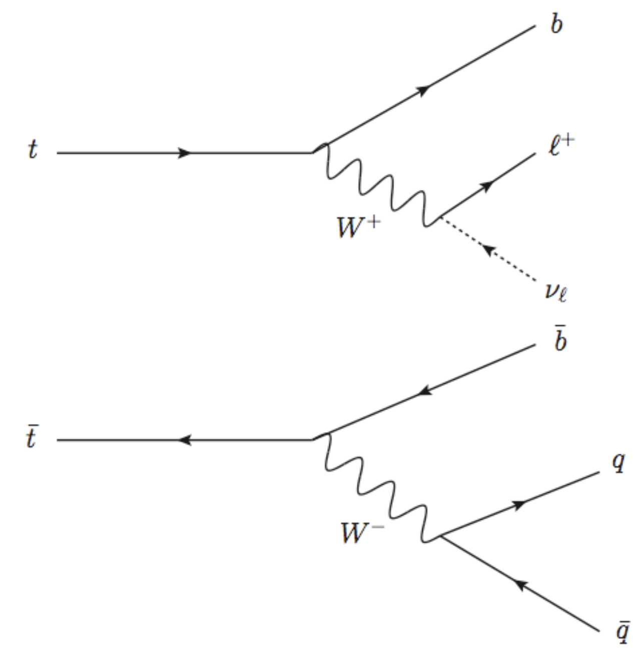
- Top quark mass reconstruction in the semi-leptonic channel.
  - Production: $\sigma_{ttbar\ (LHC)}$ = (833±100) pb @ 14 TeV
  - Top decay (99.9%) through $t \rightarrow W^+b$, where $t\bar{t} \rightarrow W^+bW^-\bar{b}$
  - Final states depending on the W decay channel:
    - Fully-Hadronic (BR=4/9): 6 jets (2x $W \rightarrow jj$)
    - Fully-Leptonic (BR=1/9): 2 leptons, 2 neutrinos and 2 jets (2x $W \rightarrow l\nu$)
    - Semi-Leptonic (BR=4/9): 1 lepton, 1 neutrino and 4 jets ($W \rightarrow l\nu + W \rightarrow jj$)
  - The Golden Channel is the semi-leptonic one (lepton = e, μ): 2.5 10^6 events/year.



Introduction

Semi-leptonic channel

- Final state objects
  - 1 lepton → Isolated and high pT lepton (e or μ): pT > 25 (20) GeV/c and |η| < 2.5
  - 1 neutrino → Missing Transverse Energy (MET): E_{MET} > 20 GeV
  - 4 jets → pT > 40 GeV/c and |η| < 2.5
    - 2 jets coming from the hadronization of the two light quarks.
    - 2 jets tagged as b-jets (i.e. coming from the hadronization of the b quarks)



Pre-calibration for jets

- A **pre-calibration** map (E and η dependent) is used for light jets and b-jets determined using MonteCarlo
- Semi-leptonic events → requirements: 2 or more light jets, 2 b-jets, 1 lepton (*e* or *μ*) and MET > 20 GeV
- Basic cuts:
  - GoodJet: pT > 40 GeV/c and |η| < 2.5
  - Good lepton: pT > 20 (25) GeV/c for *μ* (*e*) and |η| < 2.5

Traditional methods

- Actually, there are several methods based on a chi2 minimization which use an explicit reconstruction of the event topology (assuming a particle decay model) to measure the top quark mass reconstruction: ***kinematic methods*** [CSC Book]
- These methods use the kinematic information of the ttbar decay in order to build their basic chi2 and add their constraints.
  - The used chi2 in this talk is:

$$\chi^2 = \sum_{\mathrm{jets}+\ell} \left(\frac{E_i^m - E_i^f}{\sigma_{E_i}} \right)^2 + \left(\frac{M_{jj} - M_W^{PDG}}{\Gamma_W^{PDG}} \right)^2 + \left(\frac{M_{l\nu} - M_W^{PDG}}{\Gamma_W^{PDG}} \right)^2 + \left(\frac{M_{jjb_H} - M_{top}^f}{\sigma_{top_H}} \right)^2 + \left(\frac{M_{\ell\nu b_L} - M_{top}^f}{\sigma_{top_L}} \right)^2$$

Kinematic constraints

where E_i^{fit}, M_{top} are the parameters to be determined. σ_E comes from the the calibration, where $\sigma_E = \sigma_E(E,Eta)$.

- Thus, all of the existing methods minimize their chi2 with respect all the parameters in one step (“standard minimization”).
  - Then, a minimum is found for all parameters at the same time thanks to its kinematic constraints.

- An *in-situ* calibration is done: $E_i^{fit} = \alpha_i E_i^{reco}$ (*i* represents the light jets, the b-jets and lepton)
  - This allows to study the Jet Energy Scale (JES), denoted as α_i, for light jets and b jets separately.

GlobalChi2 - Kinematic fit

- The innovation of the GlobalChi2 algorithm is that the kinematic information of the ttbar decay is much more exploited as the hadronic W boson decay is used to introduce correlations, helping therefore the chi2 convergence.

  - Inside the "standard" chi2 minimization there is a nested minimization:
    - First, a chi2 minimization with respect the hadronic W boson parameters.
    - And afterwards a chi2 minimization with respect the top parameters including the information of the previous minimization (through correlations at algebra level: derivatives and covariance matrix).
      - These W boson parameters introduce correlations when determining the top quark parameters.

  - Thus, the current implementation of the GlobalChi2 method is just an improved Kinematic Fit technique. Moreover it is an analytical method (i.e. derivatives are calculated analytically), where every step can be controlled and monitored.
    - Implemented as a standalone class (TLorentzVector as input)
    - Available in CVS: http://atlas-sw.cern.ch/cgi-bin/viewcvs-atlas.cgi/groups/IFIC-SCT/tops/VKiFi/

    - *There is no time to explain the algebra but a summary can be found in the backup slides :(*

Strategy

1. Event selection:
   - Semi-Leptonic events → requirements: 2 or more light jets, 2 b-jets, 1 lepton (*e* or *μ*) and MET > 20 GeV
     - GoodJet: pT > 40 GeV/c and |eta| < 2.5
     - Good lepton: pT > 20 (25) GeV/c for *μ* (*e*) and |eta| < 2.5
2. Perform an hadronic W boson mass fit using a chi2 for the hadronic side.
   - In events with more than 2 light jets, the pair withe the smallest chi2 is kept as the pair candidate.
   - Once the light jet pair is selected, as a pre-selection, we reject W boson candidates outside the W mass peak (80 GeV/c^2) within a mass window of ±30 GeV/c^2
3. Hadronic b-jet association: choosing the closest b-jet (through ΔR) to the W boson candidate.
4. Perform the leptonic W boson mass fit determining the neutrino momentum (p_z^{ν}) using the MET.
5. Leptonic b-jet association: choosing the other b-jet.
6. Kinematic Fit using the selected candidates → GlobalChi2 → M_{top}

Common strategy in ttbar analysis with semi-leptonic channel.

GlobalChi2 specific.

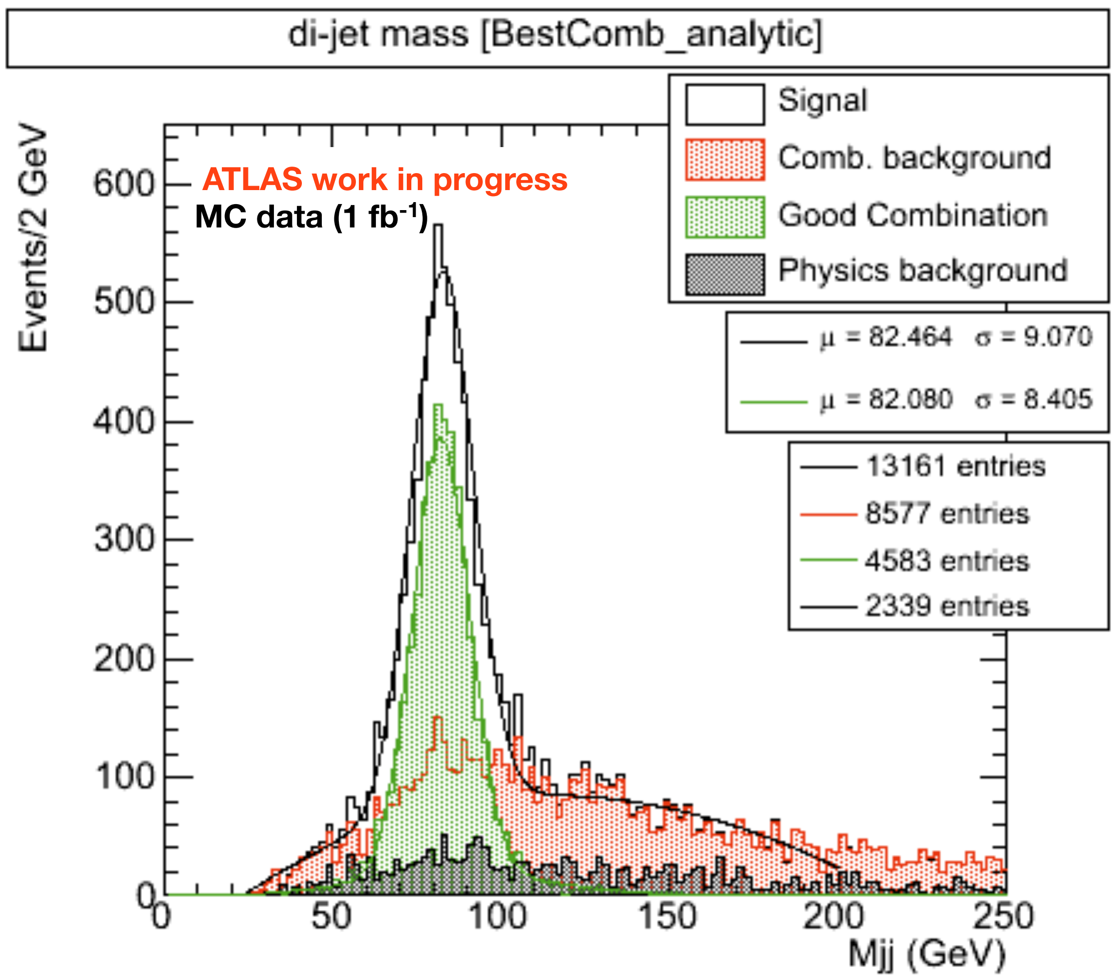
- Generally, all the plots in this talk show four curves which mean:
  - **Signal:** 1 high pT lepton (e or μ), MET> 20 GeV/c, and at least 4 high pt jets (with exactly 2 b-jets).
  - **Combinatorial Background:** From the truth, one can know if the selected candidates are the good ones or not (matching between truth and reconstruction). If not, the event is marked as “combinatorial background”.
  - **Good Combination:** If yes, the event is marked as “good combination”
  - **Physics background:** Sources of backgrounds: single top events (Wt, t and s channel), W boson production, Z +jets, etc...

Hadronic W

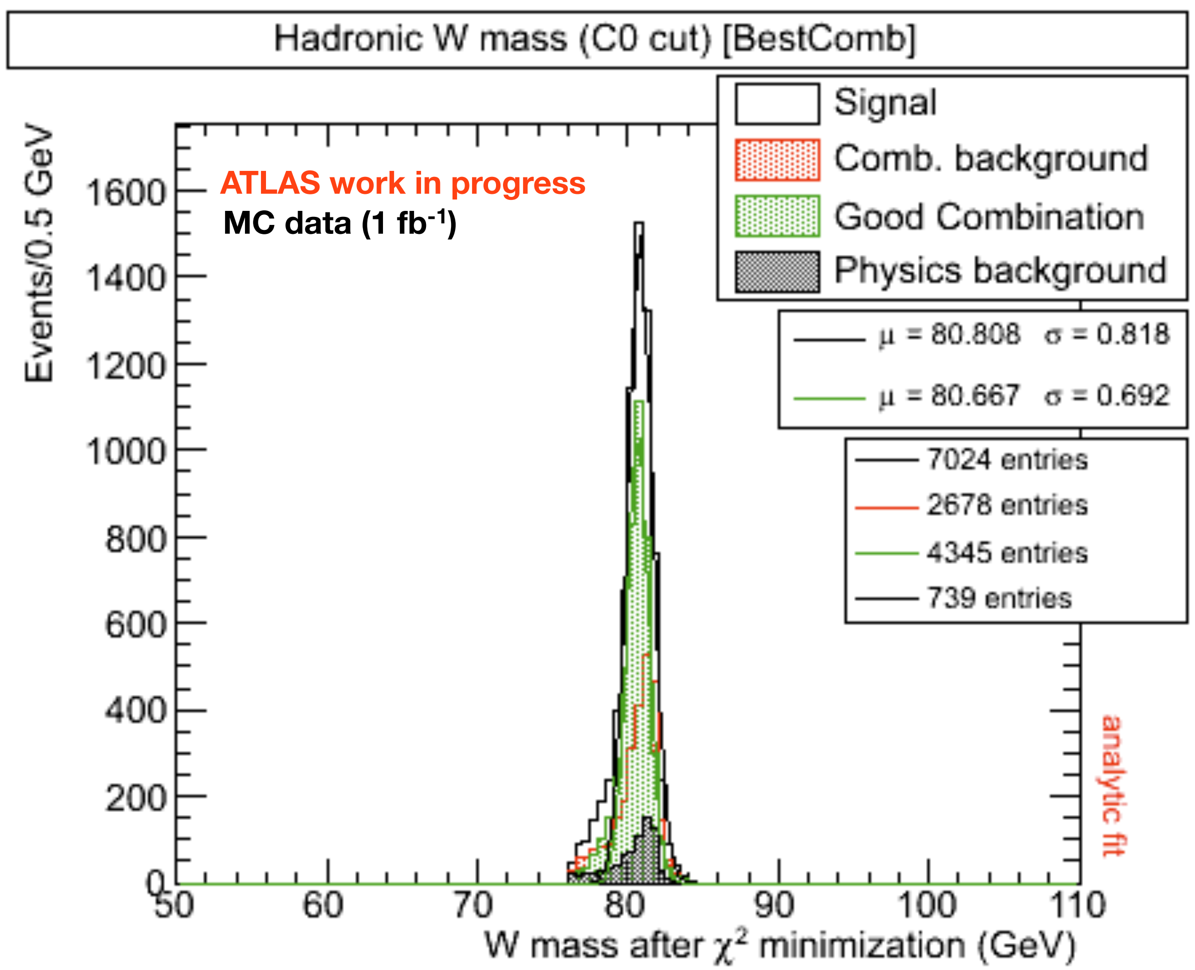
To selected good events (criterion on point 1 of previous slide), an hadronic W boson mass reconstruction is performed through a chi2 minimization, where the chi2 is:

$$\chi^2 = \left(\frac{E^m_{j0} - E^f_{j0}}{\sigma_{E_j0}}\right)^2 + \left(\frac{E^m_{j1} - E^f_{j1}}{\sigma_{E_j1}}\right)^2 + \left(\frac{M_{jj} - M_W^{PDG}}{\Gamma_W^{PDG}}\right)^2$$

Before any minimization the W boson mass is:



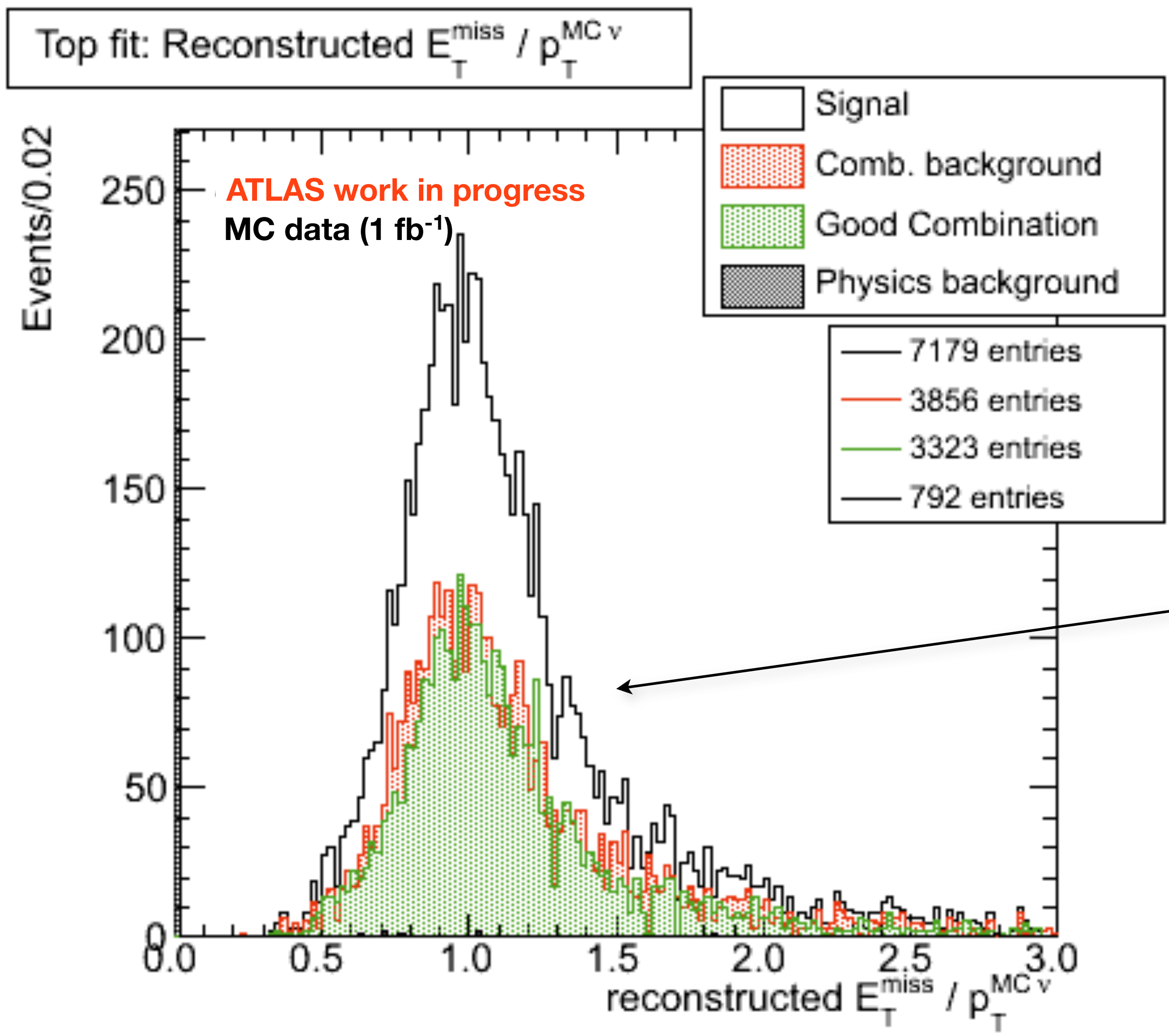
After the chi2 minimization
(*Cut0:* W boson mass window cut: $|M_{jj}-M_W| > 30$ GeV):



For events with more than 2 light jets, the pair with the smallest chi2 is kept as the hadronic W boson candidate (label BestComb)

Leptonic W

To reconstruct the leptonic W, the main difficulty here comes from the kinematics of the neutrino.
The MET is used as an estimator of the neutrino pT, but this is just an approximation as is shown in the plot:



There is more MET that the one coming from the neutrino

Four-momentum conservation for the $W \rightarrow \ell + \nu$ decay:

$$p^W = p^\ell + p_\nu \longrightarrow (p^W)^2 = (p^\ell + p^\nu)^2$$

$$M_W^2 = m_\ell^2 + 2(E_\ell, \mathbf{p}^\ell)(E_\nu, \mathbf{p}^\nu) = m_\ell^2 + 2(E_\ell E_\nu - \mathbf{p}^\ell \cdot \mathbf{p}^\nu)$$

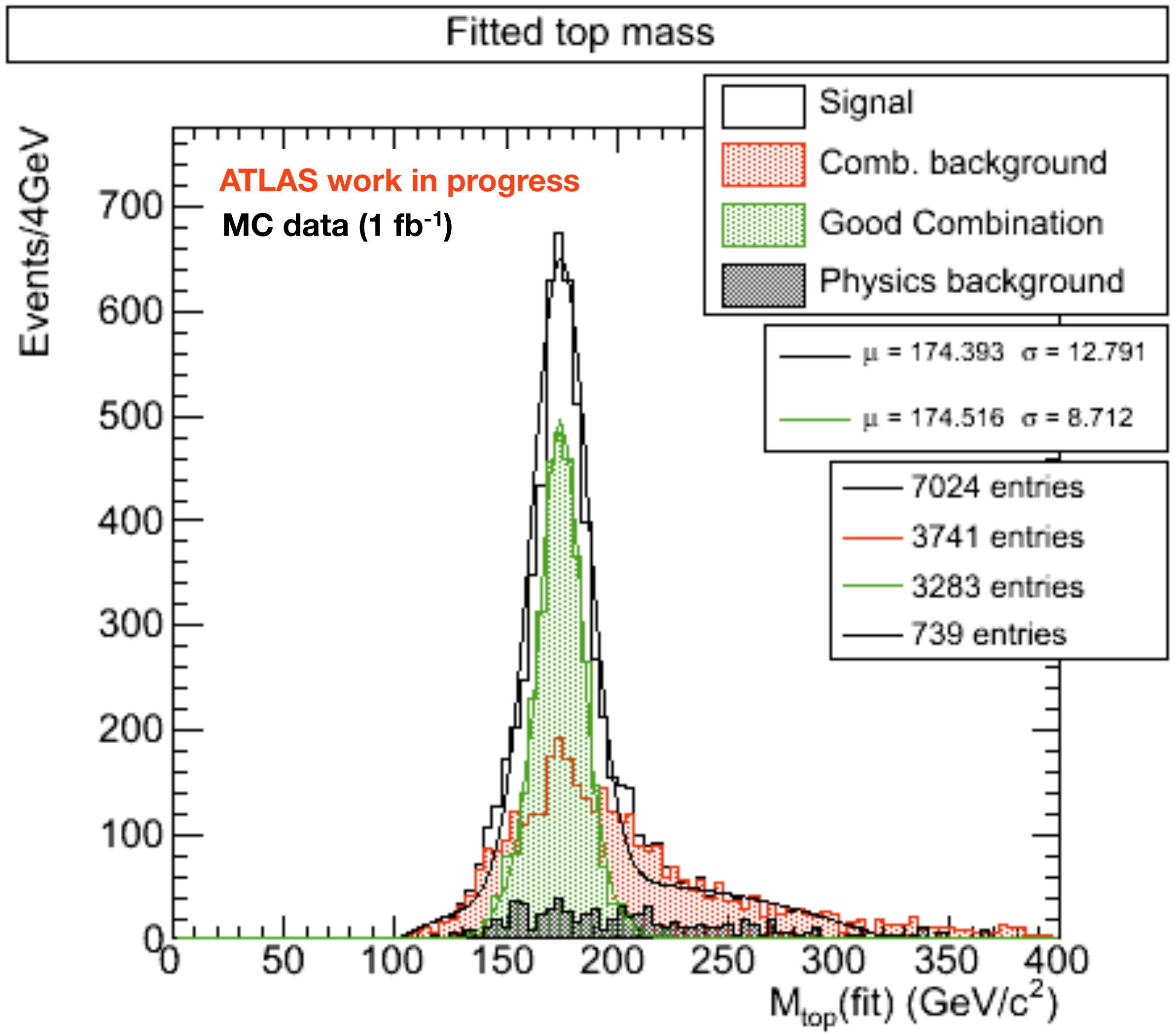
This equation is quadratic in $p_z^{\,\nu}$ and it has no solution if the measured MET fluctuates such that the neutrino-lepton invariant mass is above the W boson mass. Then, when no solution, the MET has to be re-estimated and the new $p_z^{\,\nu}$ is calculated.

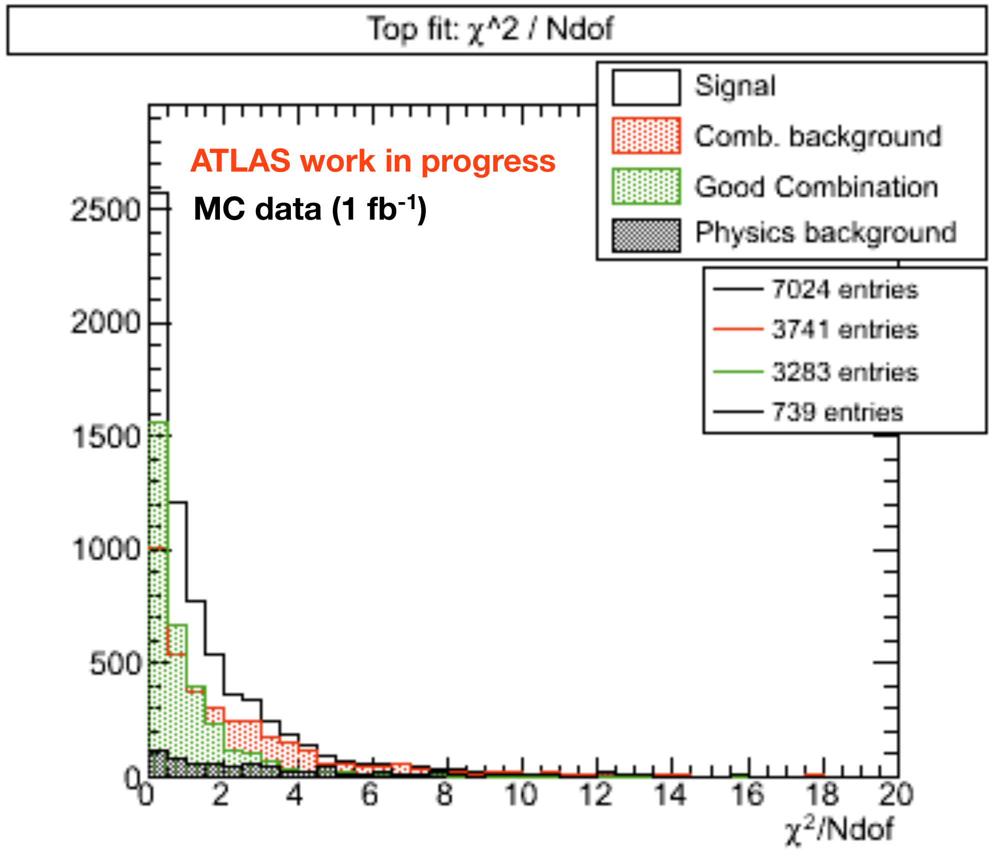
Thus, from the two solutions for $p_z^{\,\nu}$, we select the one which makes M_{jjbH} - $M_{\ell\nu bL}$ closer to zero. Afterwards, with this solution we can do the kinematic fit to extract the top quark mass....

Now, with all the candidates performed the GlobalChi2 fit: 1) $\frac{\partial \chi^2}{\partial \mathbf{W}} = 0$ and then, 2) $\frac{d\chi^2}{d\mathbf{t}} = 0$

$$\chi^2 = \sum_{\text{jets}+\ell} \left(\frac{E_i^m - E_i^f}{\sigma_{E_i}}\right)^2 + \left(\frac{M_{jj} - M_W^{PDG}}{\Gamma_W^{PDG}}\right)^2 + \left(\frac{M_{l\nu} - M_W^{PDG}}{\Gamma_W^{PDG}}\right)^2 + \left(\frac{M_{jjb_H} - M_{top}^f}{\sigma_{top_H}}\right)^2 + \left(\frac{M_{\ell\nu b_L} - M_{top}^f}{\sigma_{top_L}}\right)^2$$

The top quark mass can be extracted from the chi2 minimization wrt the top parameters where there is a nested chi2 minimization wrt the W boson parameters.



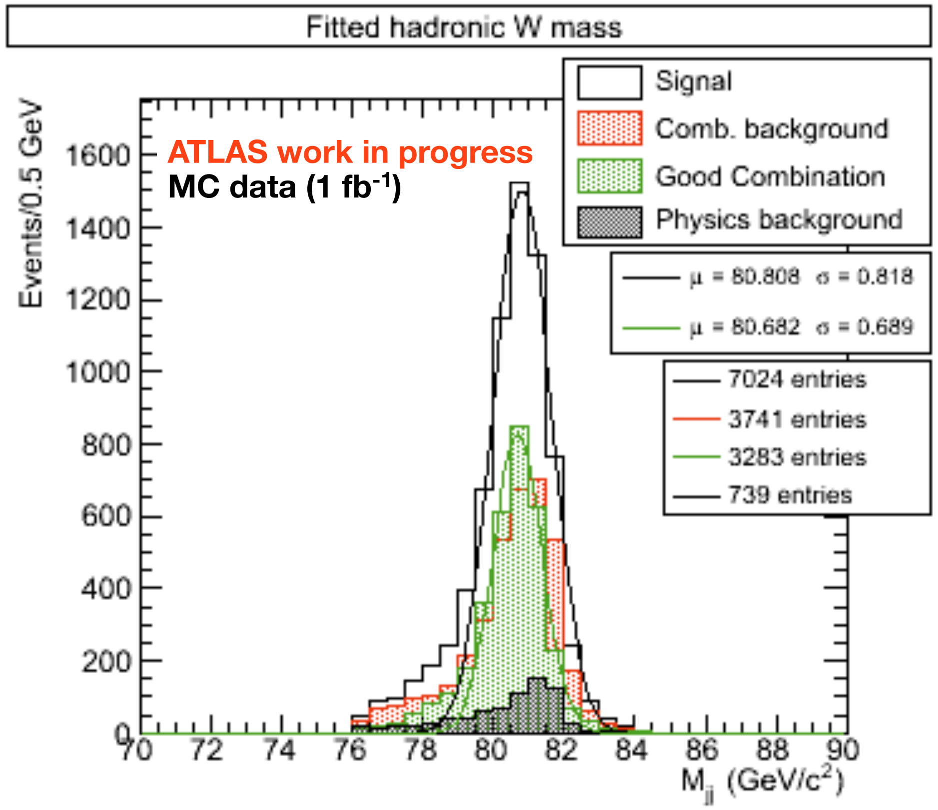


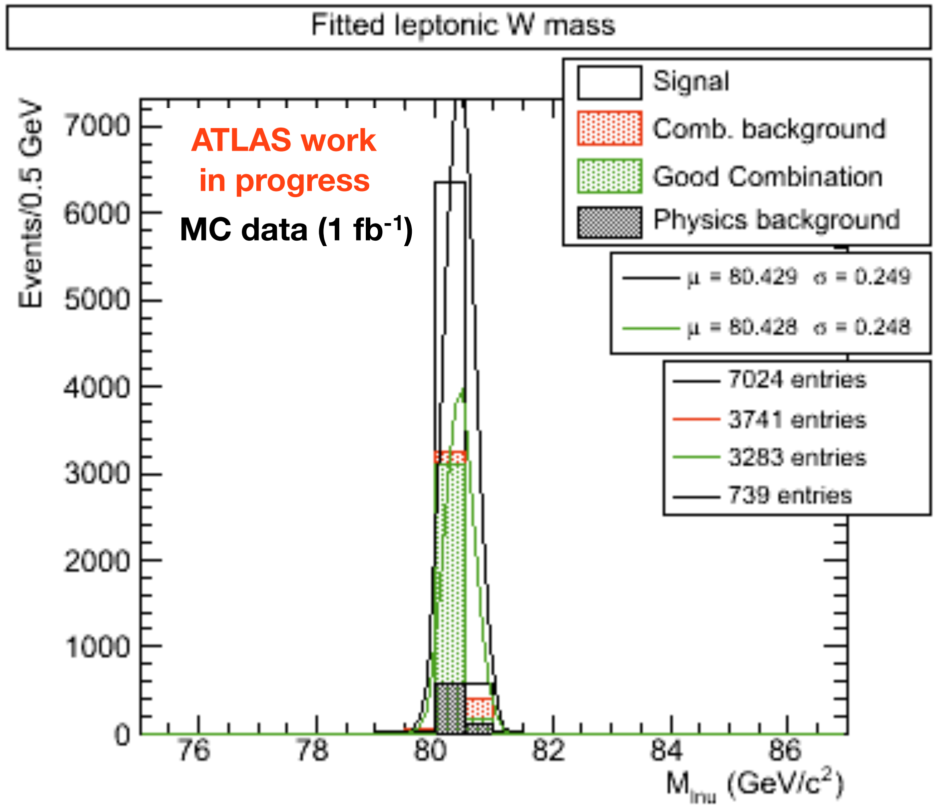
μ = (174.4 ± 0.3) GeV and σ = (12.8 ± 0.04) GeV

The simulated sample had a generated top quark mass of 175 GeV

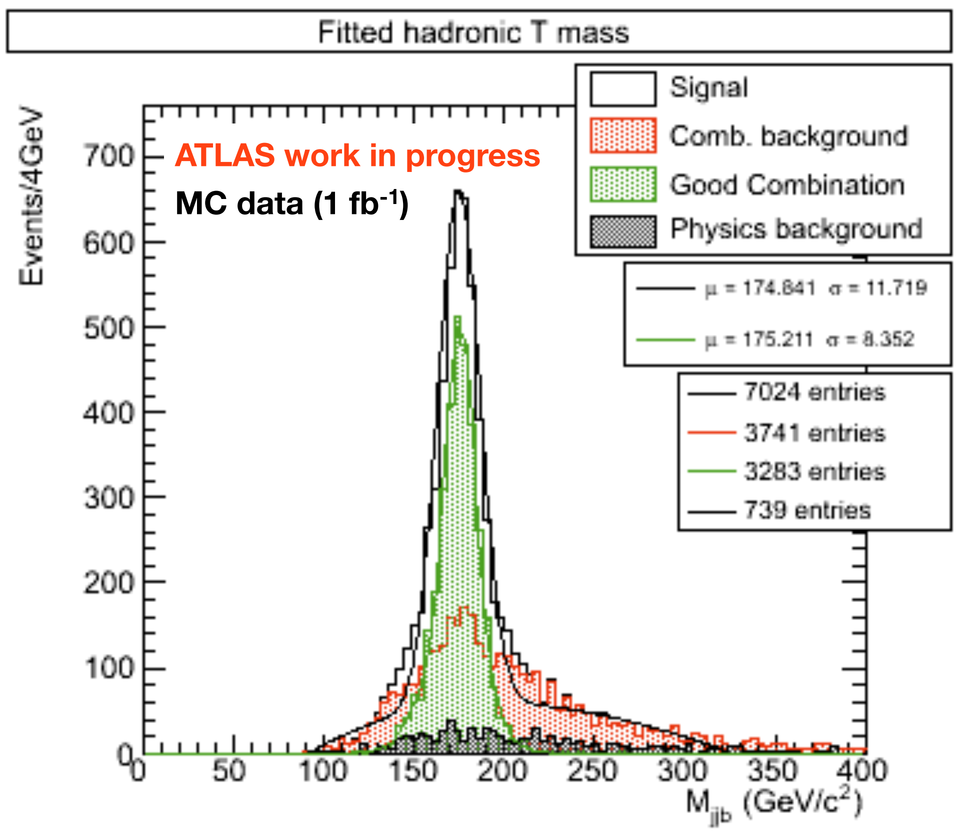
These results are compatible with the “official” ones, published in the [***CSC Book***]

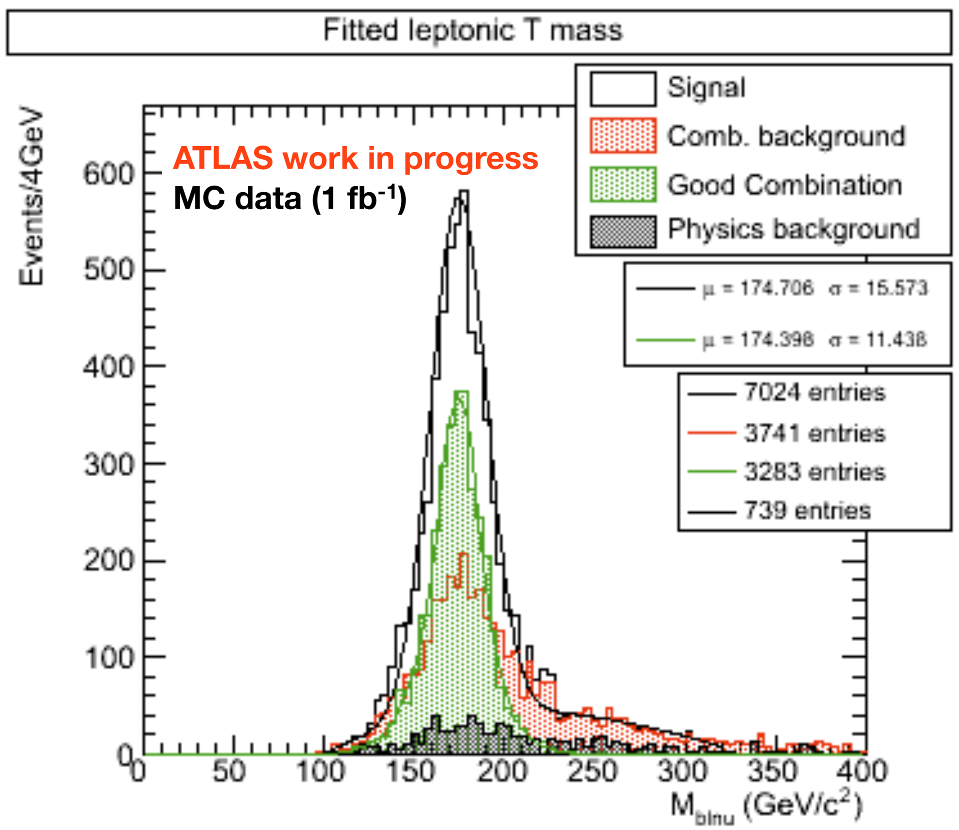
The hadronic W and the leptonic W can be reconstructed using the invariant mass of the jj and $\ell\nu$, respectively.





The hadronic top and the leptonic top can be reconstructed using the invariant mass of the jjb_H and $\ell\nu b_L$, respectively.





Summary

- A novel method has being developed based on the GlobalChi2 algorithm which is a chi2 minimization method.

- This algorithm uses the explicit reconstruction of the event topology to measure the top quark mass. i.e. it is a kinematic method. But the main difference with respect the "standard" methods is that the kinematic information of the ttbar decay is much more exploited as the hadronic W boson decay is used to introduce correlations, helping therefore the minimization convergence.

- It is expected a faster and simpler convergence when small statistics and high background, i.e. early studies.

- Results are compatible with the top quark mass peak generated for the MonteCarlo samples (175 GeV): $\mu = (174.4 \pm 0.3)$ GeV
    - The results are compatible with the official ones (see ***CSC Book***).

- No bias is introduced between the hadronic and the leptonic top reconstruction.

Contact persons:

Carlos Escobar: cescobar@ific.uv.es
Salvador Martí: martis@ific.uv.es

Backup

Nested minimization: Hadronic W

With the selected events (criterion on point 1 of previous slide), an hadronic W boson mass reconstruction is performed through a chi2 minimization, where the chi2 is:

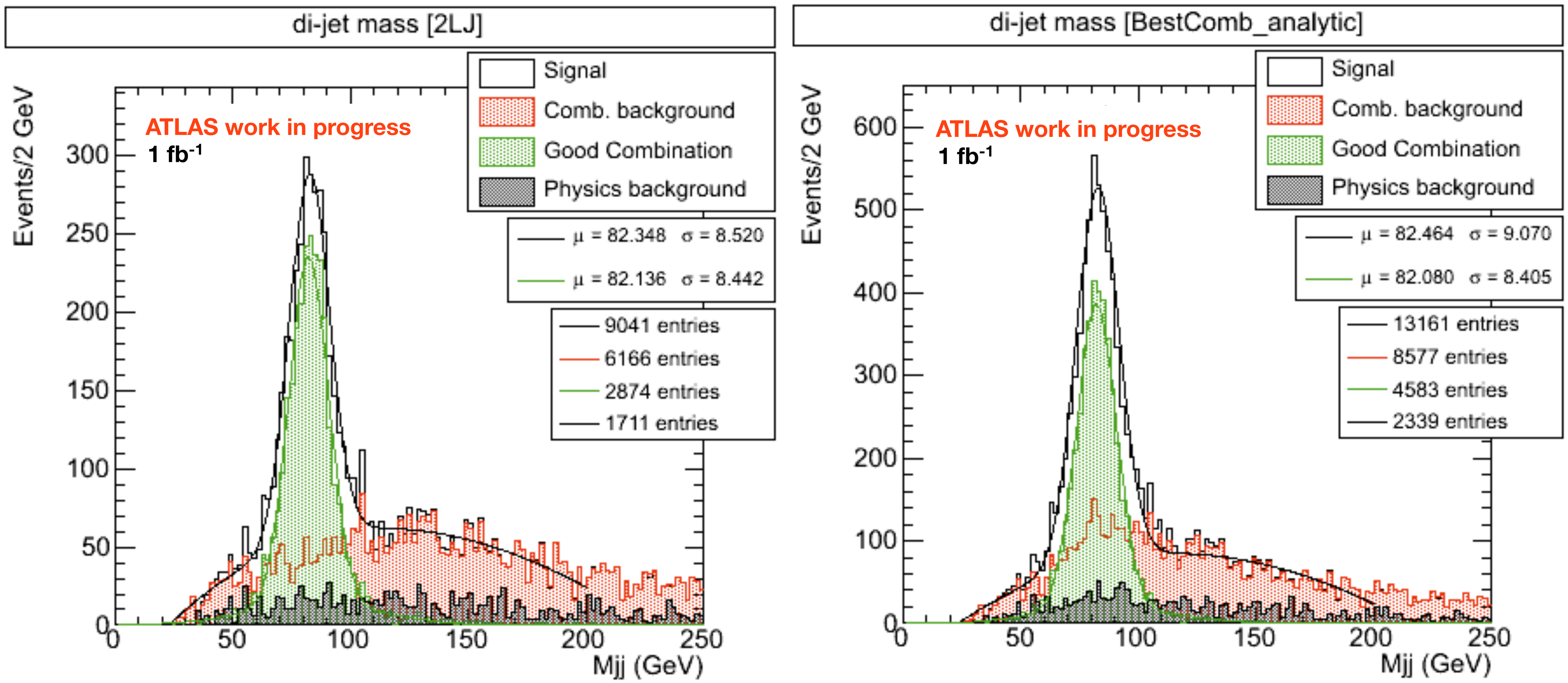
$$\frac{\partial \chi^2}{\partial \mathbf{W}} = 0$$

$$\chi^2 = \sum_{\text{jets}+\ell} \left(\frac{E_i^m - E_i^f}{\sigma_{E_i}}\right)^2 + \left(\frac{M_{jj} - M_W^{PDG}}{\Gamma_W^{PDG}}\right)^2 + \left(\frac{M_{l\nu} - M_W^{PDG}}{\Gamma_W^{PDG}}\right)^2 + \left(\frac{M_{jjb_H} - M_{top}^f}{\sigma_{top_H}}\right)^2 + \left(\frac{M_{\ell\nu b_L} - M_{top}^f}{\sigma_{top_L}}\right)^2$$

Before any minimization the W boson mass is:

Invariant mass of the light jet pair in events with only 2 light jets (label 2LJ)

Invariant mass of the light jet pair in events with more than 2 light jets (label BestComb)



For events with more than 2 light jets, the pair with the smallest chi2 is kept as the hadronic W boson candidate (label BestComb)

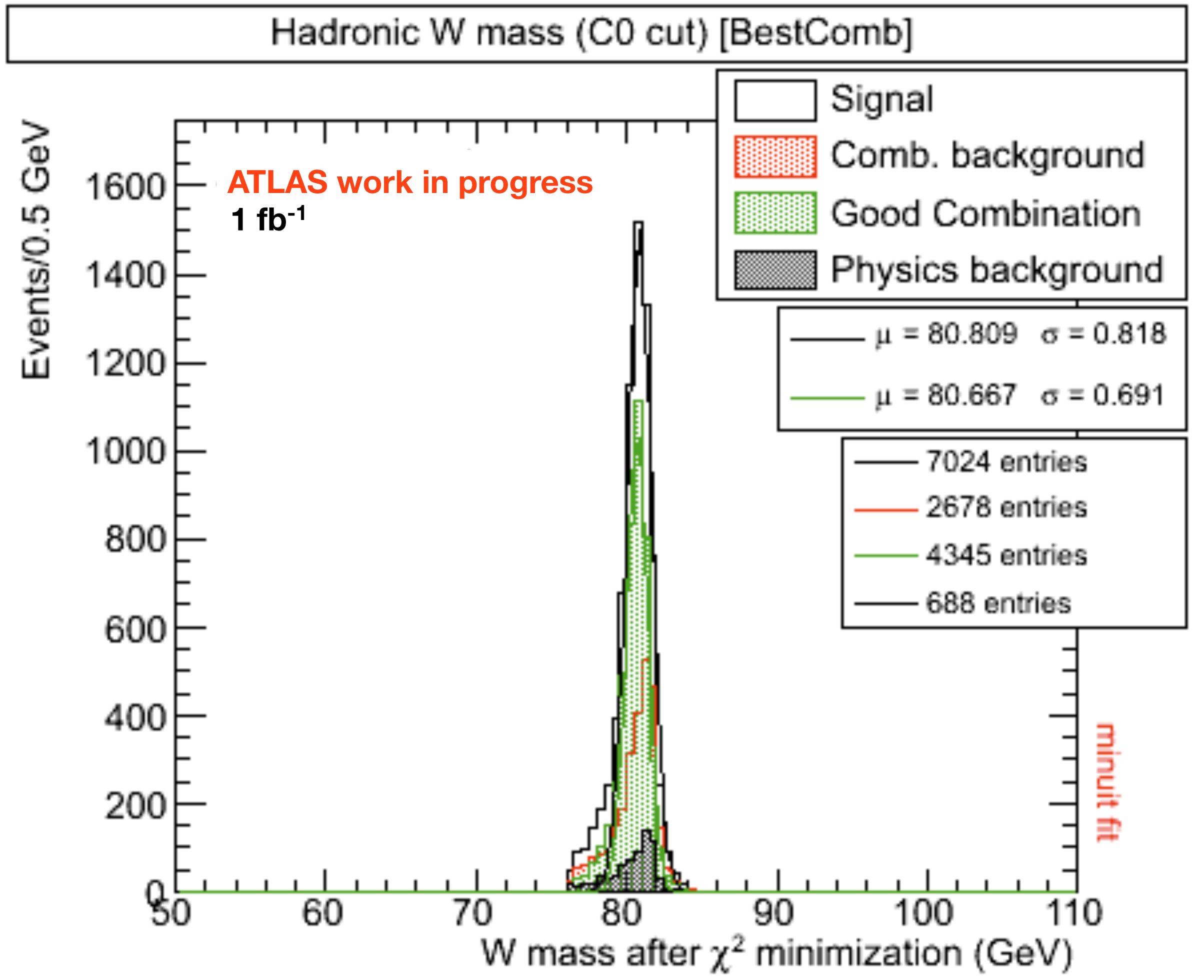
Nested minimization: Hadronic W

A "MINUIT version" (as in the others methods) and an "Analytic version" of the minimization has been implemented to perform cross-checks. Naively, both methods should give the same results. Then, the W boson mass after the chi2 minimization:

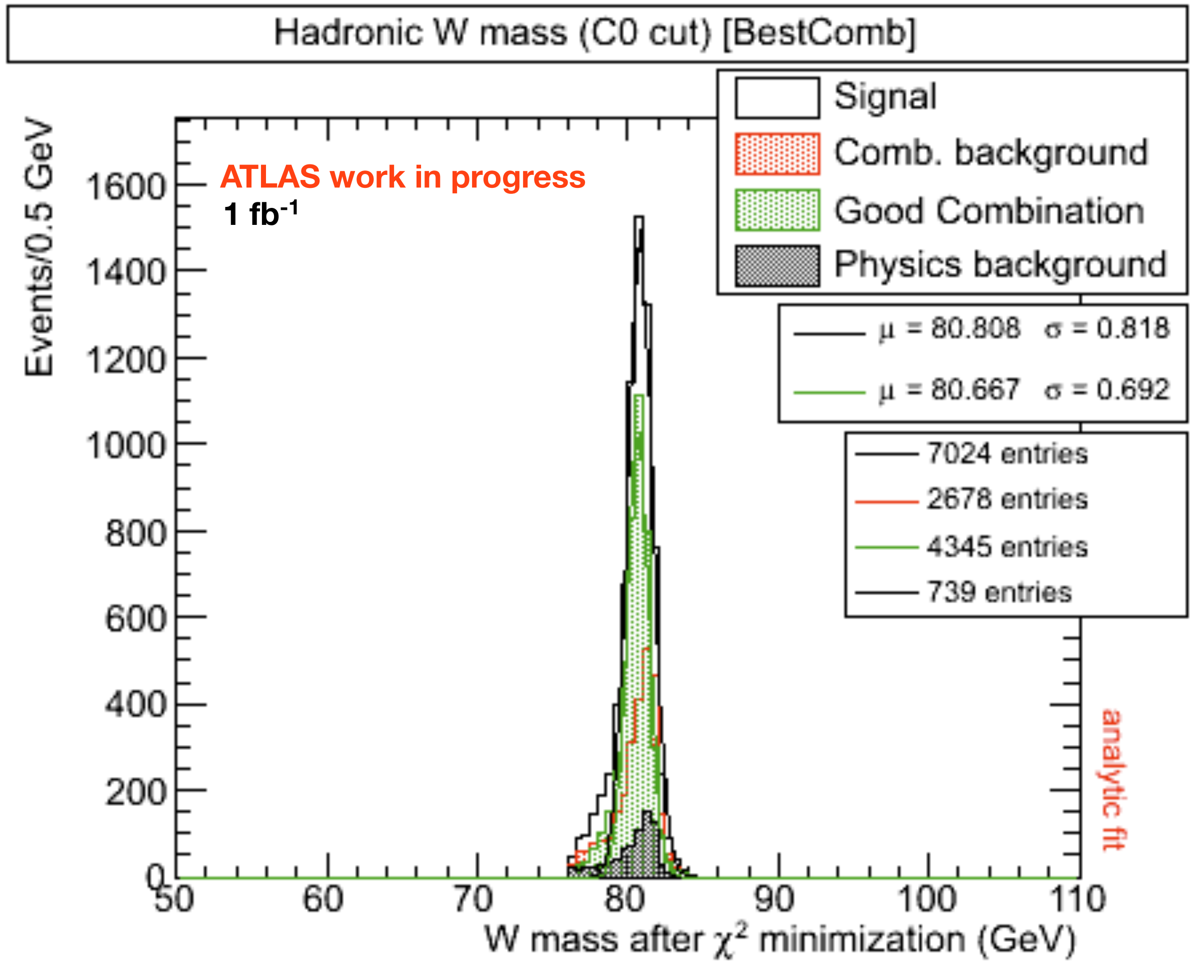
Both methods give the same results within their errors:
- Minuit: $\mu = (80.809 \pm 0.015)$ GeV/c and $\sigma = (0.818 \pm 0.014)$ GeV/c
- Analytic: $\mu = (80.808 \pm 0.015)$ GeV/c and $\sigma = (0.818 \pm 0.014)$ GeV/c

Cut0: W boson mass window cut: $|M_{jj}-M_W| > 30$ GeV

minuit (after Cut0)

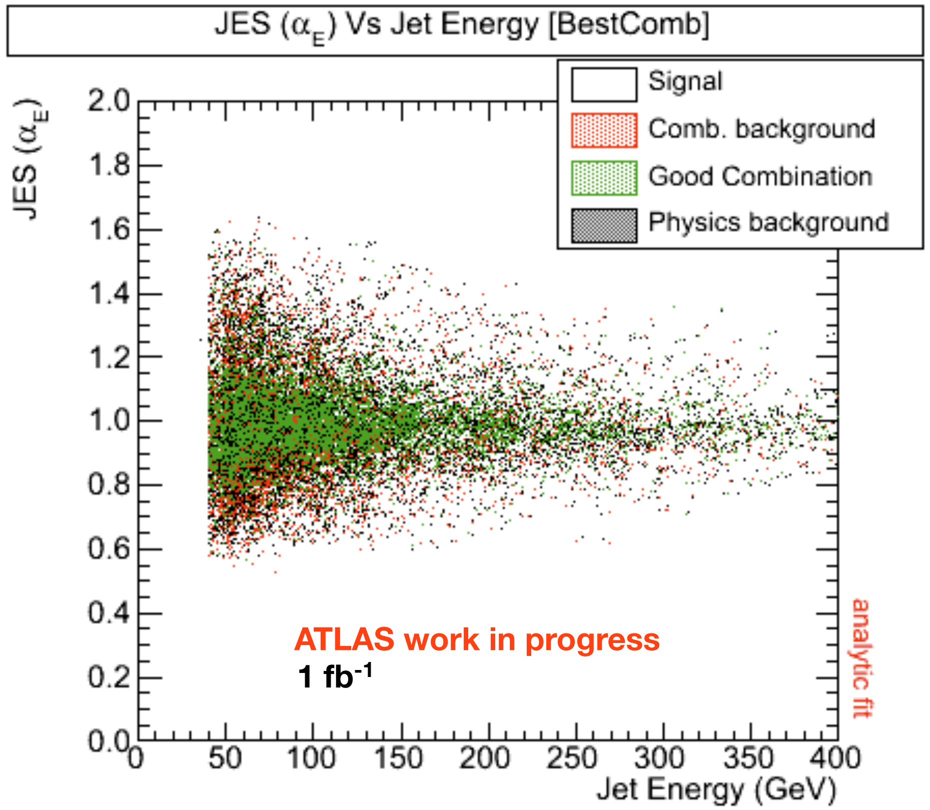


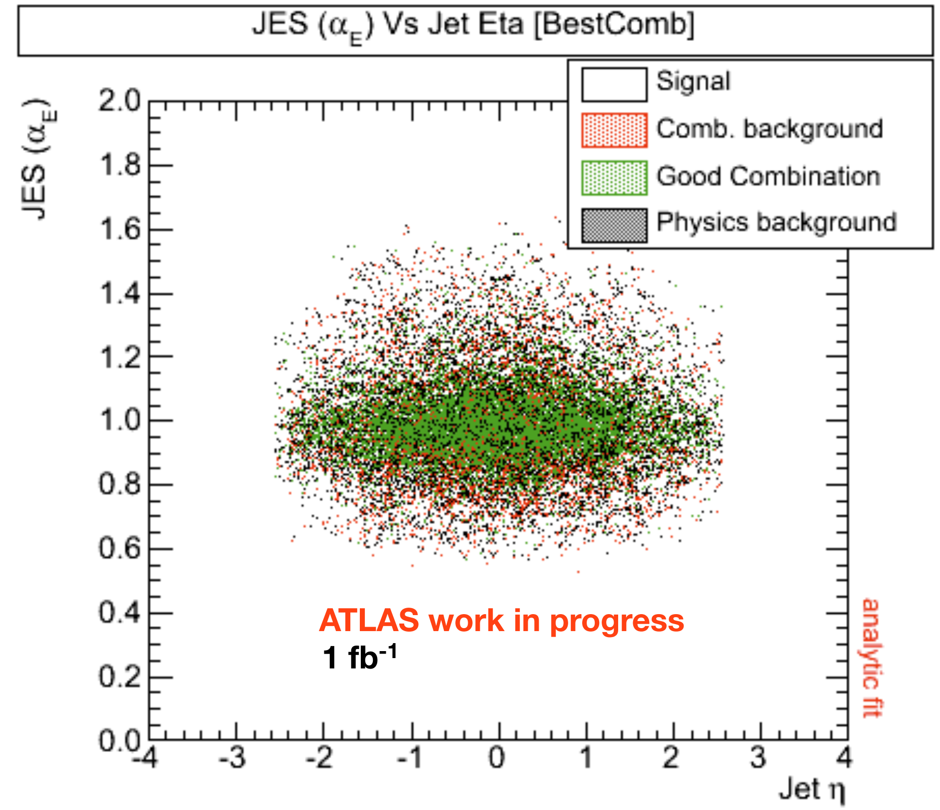
analytic (after Cut0)

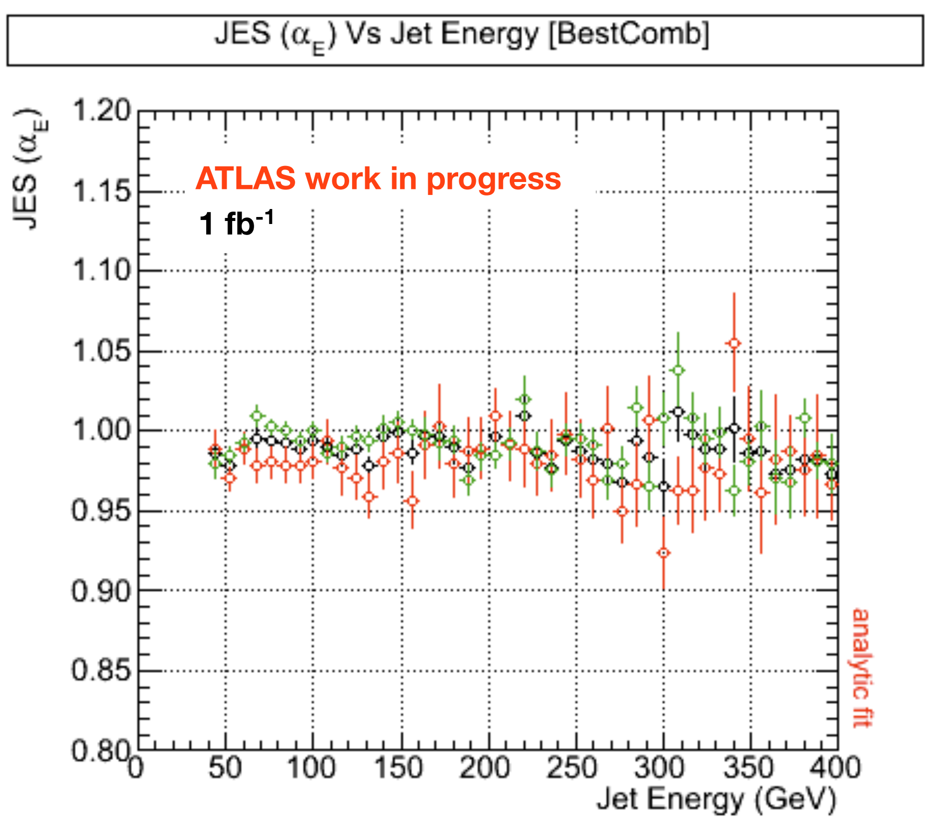


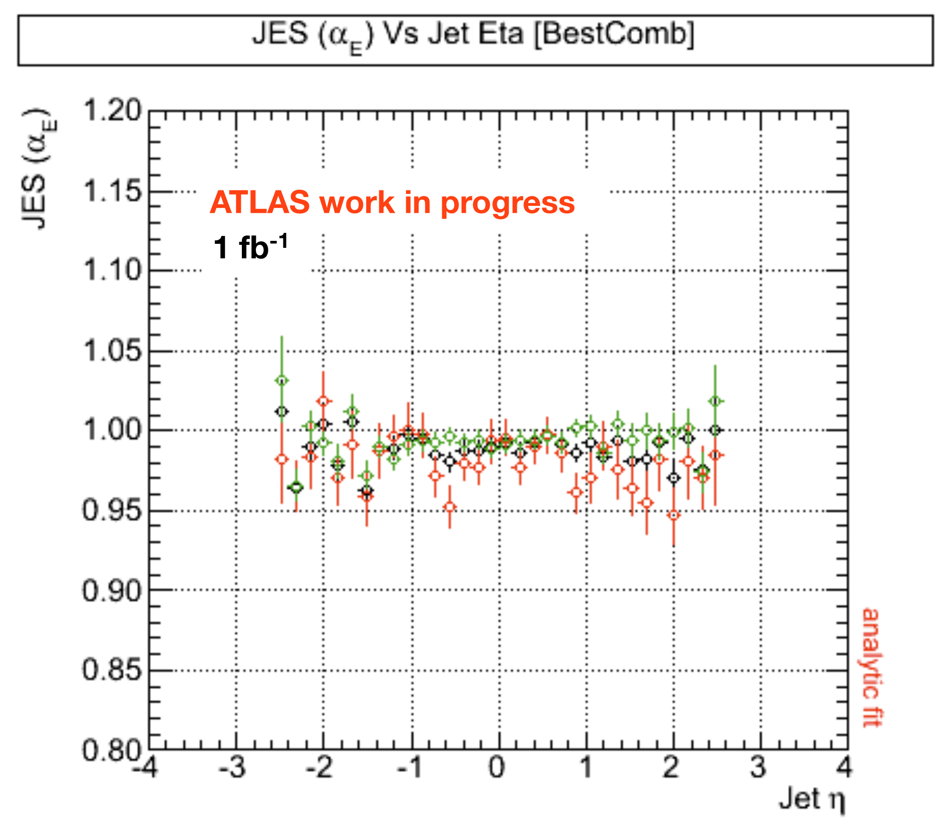
Nested minimization: Hadronic W

Jet energy scale factors behavior:
(jets have pre-calibration based on MC)



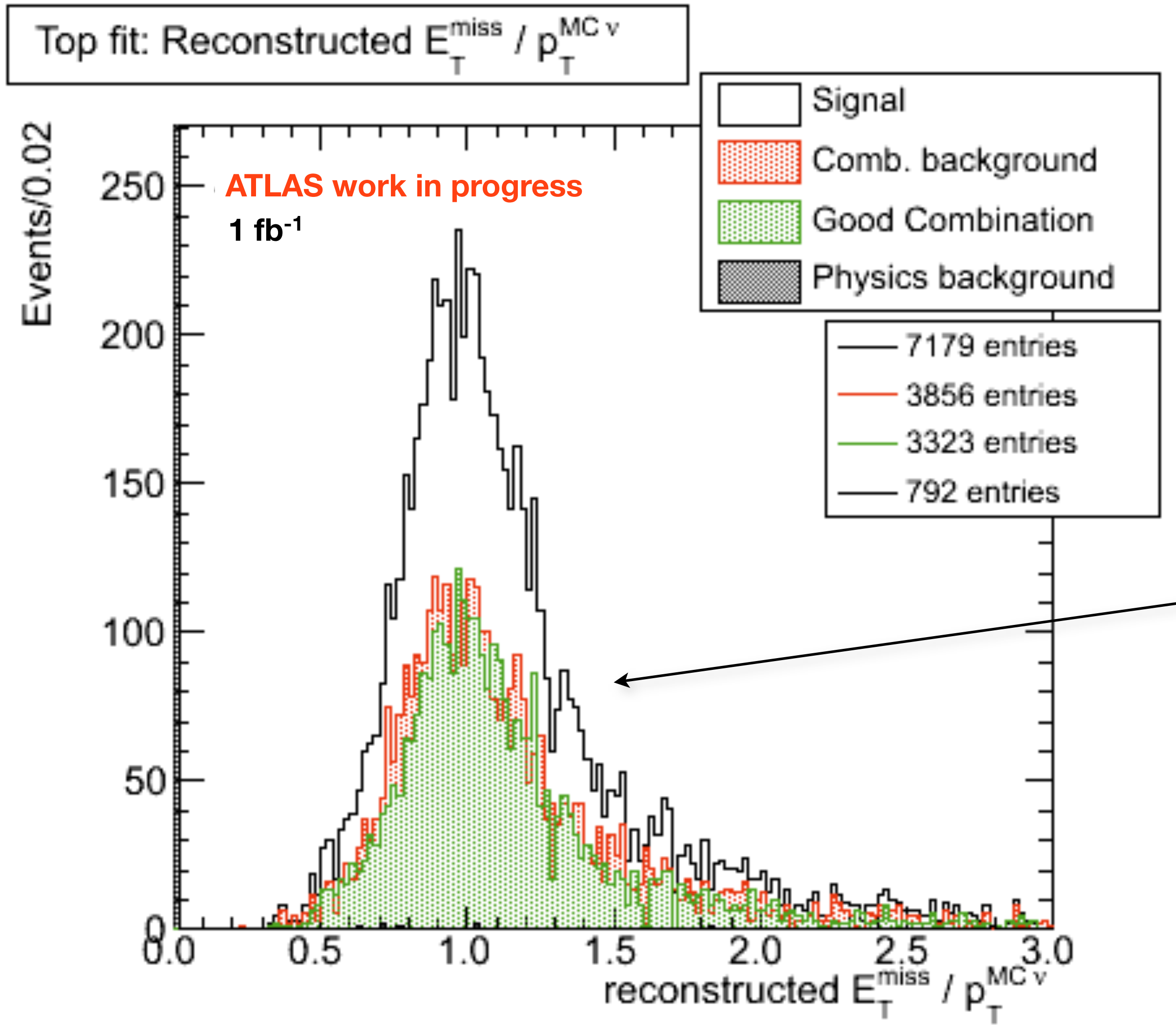






Leptonic W

To reconstruct the leptonic W, the main difficulty here comes from the kinematics of the neutrino.
The MET is used as an estimator of the neutrino pT, but this is just an approximation as is shown in the plot:



Four-momentum conservation for the $W \to \ell + \nu$ decay:

$$p^W = p^\ell + p_\nu \longrightarrow (p^W)^2 = (p^\ell + p^\nu)^2$$

$$M_W^2 = m_\ell^2 + 2(E_\ell, \mathbf{p}^\ell)(E_\nu, \mathbf{p}^\nu) = m_\ell^2 + 2(E_\ell E_\nu - \mathbf{p}^\ell \cdot \mathbf{p}^\nu)$$

This equation is quadratic in $p_z{}^\nu$ and it has no solution if the measured MET fluctuates such that the neutrino-lepton invariant mass is above the W boson mass. Then, the exact solution is:

$$p_z^\nu = \frac{-p_z^\ell\left(-M_W^2 + m_\ell^2 - 2(p_x^\ell p_x^\nu + p_y^\ell p_y^\nu)\right)}{2(E_\ell^2 - (p_z^\ell)^2)} \pm \sqrt{E_\ell^2\left[\left(M_W^2 - m_\ell^2 + 2(p_x^\ell p_x^\nu + p_y^\ell p_y^\nu)\right)^2 + 4\,\not{E}_t^2(-E_\ell^2 + (p_z^\ell)^2)\right]}$$

Leptonic W

Thus when this formula has no solution.... is because the term within the square root is negative. If that happens, one can find for which values of the MET this term becomes positive (or at least 0). By doing so, one just scales MET but preserves its direction. Therefore one has to solve a new quadratic equation, this time in terms of MET' (the new missing transverse energy):

$$E_\ell^2\left[\left(M_W^2 - m_\ell^2 + 2(p_x^\ell p_x^\nu + p_y^\ell p_y^\nu)\right)^2 + 4\,\not{E}_t^2(-E_\ell^2 + (p_z^\ell)^2)\right] = 0$$

and this equation has two solutions:

$$\not{E}_t' = \frac{-(-M_W^2 + m_\ell^2)(p_x^\ell \cos\phi_{\not{E}_t} + p_y^\ell \sin\phi_{\not{E}_t}) \pm (-M_W^2 + m_\ell^2)\sqrt{E_\ell^2 - (p_z^\ell)^2}}{2\left[E_\ell^2 - (p_z^\ell)^2 - (p_x^\ell \cos\phi_{\not{E}_t} + p_y^\ell \sin\phi_{\not{E}_t})^2\right)}$$

Thus, from the two solutions for $p_z{}^\nu$, we select the one which makes M_{jjbH} - $M_{\ell\nu bL}$ closer to zero. Afterwards, with this solution we can do the kinematic fit to extract the top quark mass....

GlobalChi algorithm

GlobalChi2 basis

Starting point

Chi2 function:

$$\chi^2 = \sum_{\text{jets}+\ell} \left(\frac{E_i^m - E_i^f}{\sigma_{E_i}} \right)^2 + \left(\frac{M_{jj} - M_W^{PDG}}{\Gamma_W^{PDG}} \right)^2 + \left(\frac{M_{l\nu} - M_W^{PDG}}{\Gamma_W^{PDG}} \right)^2 + \left(\frac{M_{jjb_H} - M_{top}^f}{\sigma_{top_H}} \right)^2 + \left(\frac{M_{\ell\nu b_L} - M_{top}^f}{\sigma_{top_L}} \right)^2$$

where $E_i^{fit} = \alpha_i E_i^m$ as has been said before.

The matrix representation is going to be used: $\chi^2 = r^T V^{-1} r$

where **r** are called residuals and represents "measurement - fit" and **V** is the covariance matrix of these measurements. In fact, the residuals **r=r(W,t)** where ***W*** represents the W boson (local) parameters and ***t*** represents the top quark mass (global) parameters.

$$\mathbf{r} = \begin{pmatrix} E^{mes}_{jet1} - E^{fit}_{jet1} \\ E^{mes}_{jet2} - E^{fit}_{jet2} \\ E^{mes}_{b_{had}} - E^{fit}_{b_{had}} \\ E^{mes}_{b_{lep}} - E^{fit}_{b_{lep}} \\ E^{mes}_{l} - E^{fit}_{l} \\ M_{jj} - M_W^{PDG} \\ M_{l\nu} - M_W^{PDG} \\ M_{jjb_{had}} - M_{top}^{PDG} \\ M_{l\nu b_{lep}} - M_{top}^{PDG} \end{pmatrix}$$

$$V^{-1} = \begin{pmatrix} 1/\sigma^2_{E_{j1}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/\sigma^2_{E_{j2}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/\sigma^2_{E_{b_{had}}} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/\sigma^2_{E_{b_{lep}}} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/\sigma^2_{E_l} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/\Gamma^2_{M_W} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1/\Gamma^2_{M_W} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/\sigma^2_{M_{jjb_{had}}} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/\sigma^2_{M_{l\nu b_{lep}}} \end{pmatrix}$$

constraints

where: $\mathbf{r} = \mathbf{r}(\mathbf{W}, \mathbf{t})$ →

- **W** represents the W fit parameters: **W** = $\{\alpha_{jet1}, \alpha_{jet2}\}$
- **t** represents the top fit parameters: **t** = $\{\alpha_{bH}, \alpha_{bL}, \alpha_{lepton}$ and $M_{top}\}$

GlobalChi2 basis

Chi2 minimization

The goal is to minimize the Chi2 with respect t, so applying the minimum condition:

$$\chi^2 = r^T V^{-1} r \quad (1) \quad \xrightarrow{\text{minimum condition...}} \quad \frac{d\chi^2}{dt} = 0 \quad (2) \quad \xrightarrow{\text{after some trivial algebra... } \chi^2 = r^T V^{-1} r} \quad \left(\frac{d\mathbf{r}}{dt}\right)^T V^{-1}\mathbf{r} = 0 \quad (3)$$

Then, as **r=r(W,t)**... one has to consider the residual derivatives...

$$d\mathbf{r} = \frac{\partial \mathbf{r}}{\partial \mathbf{W}} d\mathbf{W} + \frac{\partial \mathbf{r}}{\partial \mathbf{t}} d\mathbf{t}$$

And **dr/dt** can be calculated:

$$\frac{d\mathbf{r}}{d\mathbf{t}} = \frac{\partial \mathbf{r}}{\partial \mathbf{W}} \frac{d\mathbf{W}}{d\mathbf{t}} + \frac{\partial \mathbf{r}}{\partial \mathbf{t}} \quad (4)$$

Assumption:

- The W parameters depend on the top parameters (W's come from the top decays).
- The top parameters do not depend on the W parameters as the top mother particle and it is fixed (at simulation level the only particle with a fixed mass is the top).

Therefore, including this derivative into the previous expression: (3) + (4)

$$\left(\frac{d\mathbf{r}}{d\mathbf{t}}\right)^T V^{-1}\mathbf{r} = \left(\frac{\partial \mathbf{r}}{\partial \mathbf{W}} \frac{d\mathbf{W}}{d\mathbf{t}} + \frac{\partial \mathbf{r}}{\partial \mathbf{t}}\right)^T V^{-1}\mathbf{r} = 0$$

Now, the question is how to calculate **dW/dt**, which correlates the W boson parameters and the top quark parameters....

... the answer is simple: a chi2 minimization wrt W have to be done!

GlobalChi2 basis

Chi2 minimization wrt W boson parameters

Now, the point is to calculate: $d\mathbf{W}/d\mathbf{t}$
To do this, we need to perform a W boson fit:

$$\frac{\partial \chi^2}{\partial \mathbf{W}} = 0 \quad (6)$$

after some trivial algebra... $\chi^2 = r^T V^{-1} r$ →

$$\left(\frac{\partial \mathbf{r}}{\partial \mathbf{W}}\right)^T V^{-1}\mathbf{r} = 0 \quad (7)$$

To solve this expression one has to consider that we are close to the solution:

$$\mathbf{r} = \mathbf{r}(\mathbf{W}_0, \mathbf{t}) + \left.\frac{\partial \mathbf{r}}{\partial \mathbf{W}}\right|_{W=W_0} \delta\mathbf{W} \quad (8)$$

where: $\mathbf{W} = \mathbf{W}_0 + \delta\mathbf{W}$

7 + 8 ⇢

$$\left(\frac{\partial \mathbf{r}}{\partial \mathbf{W}}\right)^T V^{-1}\mathbf{r} = \left(\frac{\partial \mathbf{r}(\mathbf{W}_0,\mathbf{t})}{\partial \mathbf{W}}\right)^T V^{-1}\mathbf{r}(\mathbf{W}_0,\mathbf{t}) + \left(\frac{\partial \mathbf{r}(\mathbf{W}_0,\mathbf{t})}{\partial \mathbf{W}}\right)^T V^{-1} \left.\frac{\partial \mathbf{r}}{\partial \mathbf{W}}\right|_{W_0} \delta\mathbf{W} = 0$$

Now, defining E and working out dW...

defining E...

$$E \equiv \left.\frac{\partial \mathbf{r}}{\partial \mathbf{W}}\right|_{W_0}$$

E will introduce the correlations!

⇢

$$\delta\mathbf{W} = -(E^T V^{-1} E)^{-1} E^T V^{-1} \mathbf{r}(\mathbf{W}_0, \mathbf{t}) \quad (9)$$

$$\frac{d\mathbf{W}}{d\mathbf{t}} = -(E^T V^{-1} E)^{-1} E^T V^{-1} \frac{\partial \mathbf{r}(\mathbf{W}_0, \mathbf{t})}{\partial \mathbf{t}} \quad (10)$$

Now, we have dW/dt and we can continue with expression 5.

GlobalChi2 basis

Chi2 minimization wrt top quark parameters

Once one has $d\mathbf{W}/d\mathbf{t}$, we can continue: (5) + (10)

$$\left(\frac{\partial \mathbf{r}}{\partial \mathbf{W}}\frac{d\mathbf{W}}{d\mathbf{t}} + \frac{\partial \mathbf{r}}{\partial \mathbf{t}}\right)^T V^{-1}\mathbf{r} = \left(-E\left(E^T V^{-1} E\right)^{-1} E^T V^{-1}\frac{\partial \mathbf{r}}{\partial \mathbf{t}} + \frac{\partial \mathbf{r}}{\partial \mathbf{t}}\right)^T V^{-1}\mathbf{r} =$$

$$= \left(\left(I - E\left(E^T V^{-1} E\right)^{-1} E^T V^{-1}\right)\frac{\partial \mathbf{r}}{\partial \mathbf{t}}\right)^T V^{-1}\mathbf{r} = \left(\frac{\partial \mathbf{r}}{\partial \mathbf{t}}\right)^T \mathcal{W}^{-1}\mathbf{r} = 0 \quad (11)$$

Defining $\mathcal{W}$...
(remider: correlations thanks to E)
if E=0, then $\mathcal{W}=V^{-1}$, therefore "standard" Chi2 case

$$\mathcal{W} \equiv \left(I - E(E^T V^{-1} E)^{-1} E^T V^{-1}\right)^T V^{-1} \quad (12)$$

To solve this expression, once more, we have to consider that we are close to the solution:

$$\mathbf{r} = \mathbf{r}(\mathbf{W}_0, \mathbf{t}_0) + \left.\frac{\partial \mathbf{r}}{\partial \mathbf{t}}\right|_{\mathbf{t}=\mathbf{t}_0} \delta\mathbf{t}$$

(11) + (12) $\longrightarrow$

$$\left(\frac{\partial \mathbf{r}}{\partial \mathbf{t}}\right)^T \mathcal{W}\mathbf{r} + \left(\frac{\partial \mathbf{r}}{\partial \mathbf{t}}\right)^T \mathcal{W}\frac{\partial \mathbf{r}}{\partial \mathbf{t}}\delta\mathbf{t} = 0$$

Working out dt:

$$\delta\mathbf{t} = -\left[\left(\frac{\partial \mathbf{r}}{\partial \mathbf{t}}\right)^T \mathcal{W}\frac{\partial \mathbf{r}}{\partial \mathbf{t}}\right]^{-1}\left(\frac{\partial \mathbf{r}}{\partial \mathbf{t}}\right)^T \mathcal{W}\mathbf{r} = 0 \quad (13)$$

$$\delta\mathbf{t} = -\mathcal{M}^{-1}\nu \quad (14)$$

These are the top parameter corrections!

$$\mathcal{M} = \left(\frac{\partial \mathbf{r}}{\partial \mathbf{t}}\right)^T \mathcal{W}\frac{\partial \mathbf{r}}{\partial \mathbf{t}}$$

This is a Matrix (by the way, $\mathcal{M}^{-1}$ is the covariance matrix of the top fit parameters)

$$\nu = \left(\frac{\partial \mathbf{r}}{\partial \mathbf{t}}\right)^T \mathcal{W}\mathbf{r}$$

This is a Vector