Studies of Systematic Uncertainties with Profile Likelihood Fits in the Context of a W-Boson Polarisation Measurement at ATLAS

Ph. Stolte
Supervisors: K. Kröninger, A. Quadt
Universität Göttingen

6th Annual Workshop of the Helmholtz Alliance: December 4, 2012

GEFÖRDERT VOM
Bundesministerium für Bildung und Forschung
PHYSICS AT THE TERA SCALE
Helmholtz Alliance
GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN
FSP 101 ATLAS

Motivation

Why are profile likelihood fits so interesting?

- Meas. of most top quark properties limited by systematic effects
- Reduce systematic uncertainties by using **profile likelihood fits**

What are characteristics of profile likelihood fits used here?

- Fit parameters are used for certain sources of systematic variation
  $\Rightarrow$ Nuisance parameters: Adjust size of systematics using data

Aim of analysis:

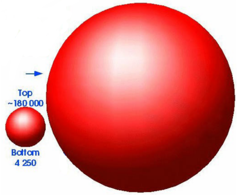
$\Rightarrow$ Explore *potential* and study important *features* of profile likelihood fits

$\Rightarrow$ Illustrate *setup* of fitting procedure

Application:

$\Rightarrow$ Test *Wtb* vertex structure

$\Rightarrow$ Measure the W helicity fractions F_0, F_L and F_R



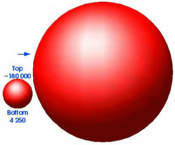
The Top Quark and Its Decay

The top quark:

- Enormous mass of about: $m_t = 173.5 \pm 0.6 \pm 0.8 \frac{\text{GeV}}{c^2}$
- Assumed to carry charge $Q = \frac{2}{3}e$ and spin $s = \frac{1}{2}$, short lifetime, no hadronisation

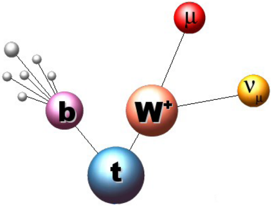
Top quark decays:

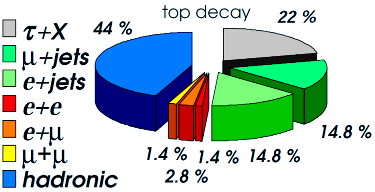
- Decay almost exclusively into b quark and W boson
- b quarks hadronise to b jets $\Rightarrow$ *b-tagging*
- W boson can decay hadronically or leptonically



Measure W helicity fractions in muon+jets channel:

- Signature: 1 isolated charged μ, E_T^{miss} due to ν, ≥ 4 high-p_T jets, ≥ 1 b-tag
- Background processes: mostly W+jets, single top and QCD multijet production



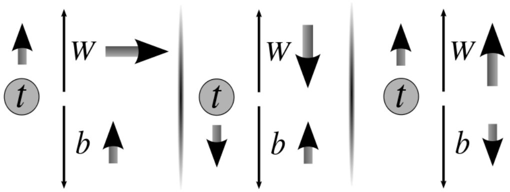


W Boson Polarisation

Possible helicity states:

- W boson: spin $s = 1$ leads to 3 possibilities for spins in t rest frame

longitudinal left-handed right-handed



Calculation of the helicity fractions: (W rest frame)

- Defined as: $F_0 = \frac{\Gamma(t \to W_0 b)}{\Gamma(t \to Wb)}$ $F_L = \frac{\Gamma(t \to W_L b)}{\Gamma(t \to Wb)}$ $F_R = \frac{\Gamma(t \to W_R b)}{\Gamma(t \to Wb)}$
- Angular distribution of particles originating from top decay:

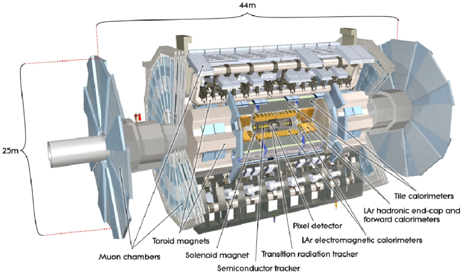
$$\frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}\cos\theta^*} = \frac{3}{4}\sin^2\theta^* F_0 + \frac{3}{8}(1-\cos\theta^*)^2 F_L + \frac{3}{8}(1+\cos\theta^*)^2 F_R$$

- NNLO QCD calculations yield:

$$F_0 = 0.687(5) \qquad F_L = 0.311(5) \qquad F_R = 0.0017(1)$$

The ATLAS Experiment

- ATLAS one of four detector experiments located at the LHC at CERN
- 44 m long, 25 m in diameter, weight of ≈ 7000 t, *four main parts*:



Inner detector (ID):
- Immersed in 2 T magn. field
- Pixel detector
- Semi-conductor tracker
- Trans. rad. tracker

Calorimeter system:
- Two sampling calorimeters:
- EM and hadronic cal. system

Muon spectrometer:
- 4 different types of tubes/chambers
- μ detection based on magnetic deflection of tracks in toroid magnets

Magnet System:
- ID surrounded by solenoid, outside calorimeters outer toroidal magn. field

3 level **trigger system** used to select signal and reject background events

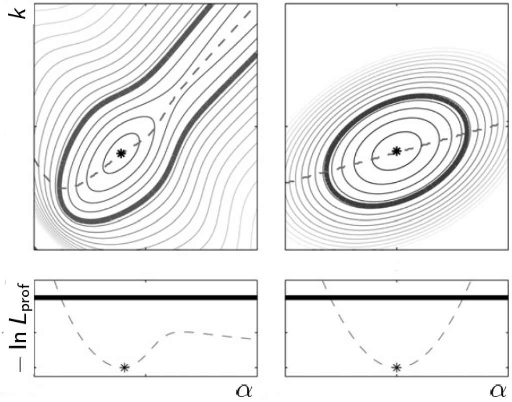
Profile Likelihood (PL)

Definition of PL:

- $L_{\text{prof}}(\vec{x}|\vec{\alpha}) = \sup_{\vec{k}} \mathcal{L}(\vec{x}|\vec{\alpha}, \vec{k})$
- PL maximised with respect to $\vec{k}$
- with:
  - $\vec{\alpha}$ vec. par. of interest
  - $\vec{k}$ vec. of nuisance parameters
  - $\vec{x}$ contains n independent observables

Graphical representation:[1]

- Usually minimise negative logar. likelihood
- Contour plot: 2D problem
- Bottom plot: Neg. log. PL
- Asterisk: Optimal param. choice



[1] A. Raue et al., *IET Syst. Biol., 5:120-130*, 2011

Template Fits

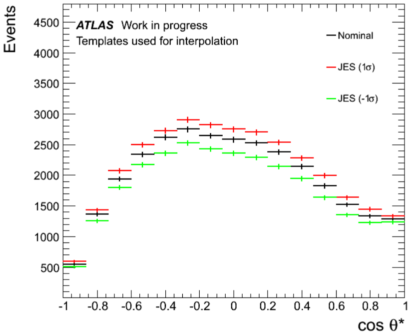
Template fits:

- ...based on idea of adding nominal templates for sig./bkg. contrib. $\Rightarrow H_S$ and H_B
- ...fitted by normalisation parameters $\vec{N}$
- Templates obtained from MC simulations and auxiliary measurements

$\Rightarrow$ Gain number of sig./bkg. events

- Write in condensed form:

$$H_{\text{sum}}(\vec{N}_S, \vec{N}_B) = H_S(\vec{N}_S) + H_B(\vec{N}_B)$$



Idea: Fit distribution H_{sum} to data/pseudo-data H_{data} (Poisson model)

Likelihood function: Minimise negative logarithmic likelihood

$$-\ln \mathcal{L}_{\vec{N}_S,\vec{N}_B} = \sum_{m=1}^{N_{\text{bins}}} \left(H_{\text{sum},m} - H_{\text{data},m} \cdot \ln H_{\text{sum},m}\right) + \text{const.}$$

Profile Likelihood Analyses with Template Fits

Fit parameters and templates:

- PL fits based on template fits:

$\Rightarrow$ Nominal templates for sig./bkg. contributions
$\Rightarrow$ ...fitted by normalisation parameters $\vec{N}$

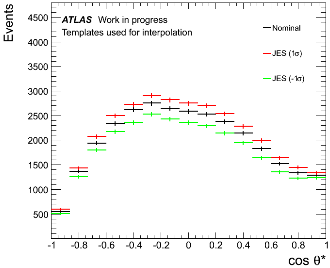
- To adjust size of systematic effects:

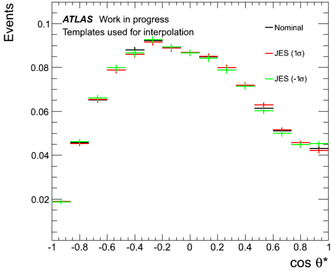
$\Rightarrow$ Templates for syst. uncertainties added to fit
$\Rightarrow$ $\pm 1\sigma$-variations for all systematics
$\Rightarrow$ Use nuisance parameters $\vec{k}$
to estimate corresponding effects

- Nuisance parameters:

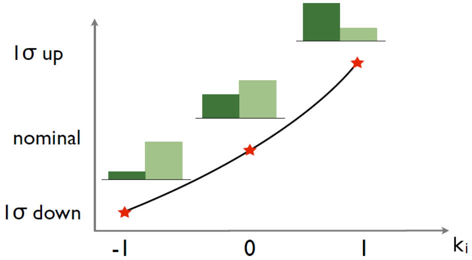
$\Rightarrow$ Gaussian-distributed
$\Rightarrow$ Nominal value: 0, $\pm 1\sigma$-variation: ± 1

- $\pm 1\sigma$-templates:
Absolute numbers and shape change





Interpolation/Morphing



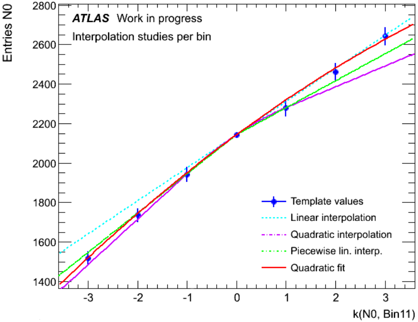
Idea of morphing:

- During fit:
  Perform interpolation/morphing step
- Shape of templates morphed from:
  $H_{-1\sigma}$ to H_{nom} to $H_{+1\sigma}$

- Morphing controlled by underlying nuisance parameters
- Morphing step performed for all nuisance parameters
  using all signal/bkg. contributions

Interpolation Methods

Idea: Produce also $\pm 2\sigma$, $\pm 3\sigma$ variations



Interpolation methods:

- Linear interpolation:

$$H_{\text{int},ij} = H_{\text{nom},ij} + k_j \cdot \frac{H_{+1\sigma,ij} - H_{-1\sigma,ij}}{2}$$

- Piecewise linear interpolation:

If $k_j > 0$: $H_{\text{int},ij} = H_{\text{nom},ij} + k_j \cdot \frac{H_{+1\sigma,ij} - H_{\text{nom},ij}}{2}$

If $k_j < 0$: $H_{\text{int},ij} = H_{\text{nom},ij} + k_j \cdot \frac{H_{\text{nom},ij} - H_{-1\sigma,ij}}{2}$

- Quadratic interpolation in $k_j \in [-1, 1]$:

$$H_{\text{int},ij} = H_{\text{nom},ij} + k_j \cdot \frac{H_{+1\sigma,ij} - H_{-1\sigma,ij}}{2} + k_j^2 \cdot \left(\frac{H_{+1\sigma,ij} + H_{-1\sigma,ij}}{2} - H_{\text{nom},ij} \right)$$

if $|k_j| > 1$ use linear interpolation so that resulting interp. function is continuous

- Quadratic fit using all data points per bin: nominal, $\pm 1\sigma$, $\pm 2\sigma$, $\pm 3\sigma$
  ...resulting fitted function serves as interpolation function

- Note: j refers to the corresponding source while i denotes bin of the templates

PLFs to Measure the W Boson Polarisation

Idea: Fit MC distribution H_{sum} to data/pseudo-data H_{data}

Fit parameters and templates:

- $\cos\theta^*$-distributions as input templates with 15 bins
- Fit parameters:
  - 3 signal parameters for helicity states: N_0, N_L, N_R
  - 3 bkg. parameters N_B for Wjets, QCD, RemBkg

  $\Big\}\ \hat{=}$ # of events

  - Due to morphing step: histos H depend on $\vec{k}$
- Signal samples: Consider selection efficiencies $\vec{\varepsilon}_{N_S}$
- Summarise signal and bkg templates:

$$H_S(\vec{k}, \vec{N}_S, \vec{\varepsilon}_{N_S}) = H_{N_0}(\vec{k}) \cdot N_0 \varepsilon_{N_0} + H_{N_L}(\vec{k}) \cdot N_L \varepsilon_{N_L} + H_{N_R}(\vec{k}) \cdot N_R \varepsilon_{N_R}$$
$$H_B(\vec{k}, \vec{N}_B) = H_{\text{Wjets}}(\vec{k}) \cdot N_{\text{Wjets}} + H_{\text{QCD}}(\vec{k}) \cdot N_{\text{QCD}} + H_{\text{RemBkg}}(\vec{k}) \cdot N_{\text{RemBkg}}$$

- Write in condensed form:

$$H_{\text{sum}}(\vec{k}, \vec{N}_S, \vec{N}_B, \vec{\varepsilon}_{N_S}) = H_S(\vec{k}, \vec{N}_S, \vec{\varepsilon}_{N_S}) + H_B(\vec{k}, \vec{N}_B)$$

Likelihood function:

$$-2 \ln L^{\text{prof}}_{\vec{N}_S, \vec{N}_B, \vec{k}} = 2 \sum_{m=1}^{N_{\text{bins}}} \left(H_{\text{sum},m} - H_{\text{data},m} \cdot \ln H_{\text{sum},m}\right) + \sum_{i=1}^{N_{\text{bkg}}} \frac{(N_{B,i} - N_{B,\text{exp},i})^2}{\sigma^2_{N_{B,\text{exp},i}}} + \sum_{j=1}^{N_{\text{prof}}} k_j^2 + \text{const.}$$

- Last sum *penalty term*: Gaussian constraint on k_j with $\mu = 0$, $\sigma = 1$
- Second sum: Gaussian constraint on $N_{B,i}$ with $\mu = N_{B,\text{exp},i}$, $\sigma_i = \sigma_{N_{B,\text{exp},i}}$

PLFs to Measure the W Boson Polarisation

Idea: Fit MC distribution H_{sum} to data/pseudo-data H_{data}

Fit parameters and templates:

- $\cos\theta^*$-distributions as input templates with 15 bins
- Fit parameters:
  - 3 signal parameters for helicity states: N_0, N_L, N_R
  - 3 bkg. parameters N_B for Wjets, QCD, RemBkg

  $\Big\} \hat{=}$ # of events
  - Due to morphing step: histos H depend on $\vec{k}$
- Signal samples: Consider selection efficiencies $\vec{\varepsilon}_{N_S}$
- Summarise signal and bkg templates:

$$H_S(\vec{k}, \vec{N}_S, \vec{\varepsilon}_{N_S}) = H_{N_0}(\vec{k}) \cdot N_0 \varepsilon_{N_0} + H_{N_L}(\vec{k}) \cdot N_L \varepsilon_{N_L} + H_{N_R}(\vec{k}) \cdot N_R \varepsilon_{N_R}$$
$$H_B(\vec{k}, \vec{N}_B) = H_{\text{Wjets}}(\vec{k}) \cdot N_{\text{Wjets}} + H_{\text{QCD}}(\vec{k}) \cdot N_{\text{QCD}} + H_{\text{RemBkg}}(\vec{k}) \cdot N_{\text{RemBkg}}$$

- Write in condensed form:

$$H_{\text{sum}}(\vec{k}, \vec{N}_S, \vec{N}_B, \vec{\varepsilon}_{N_S}) = H_S(\vec{k}, \vec{N}_S, \vec{\varepsilon}_{N_S}) + H_B(\vec{k}, \vec{N}_B)$$

Likelihood function:

$$-2\ln L^{\text{prof}}_{\vec{N}_S, \vec{N}_B, \vec{k}} = 2\sum_{m=1}^{N_{\text{bins}}} \left(H_{\text{sum},m} - H_{\text{data},m} \cdot \ln H_{\text{sum},m}\right) + \sum_{i=1}^{N_{\text{bkg}}} \frac{(N_{B,i} - N_{B,\text{exp},i})^2}{\sigma^2_{N_{B,\text{exp},i}}} + \sum_{j=1}^{N_{\text{prof}}} k_j^2 + \text{const.}$$

- Last sum *penalty term*: Gaussian constraint on k_j with $\mu = 0$, $\sigma = 1$
- Second sum: Gaussian constraint on $N_{B,i}$ with $\mu = N_{B,\text{exp},i}$, $\sigma_i = \sigma_{N_{B,\text{exp},i}}$

PLFs to Measure the W Boson Polarisation

Idea: Fit MC distribution H_{sum} to data/pseudo-data H_{data}

Fit parameters and templates:

- cos θ^*-distributions as input templates with 15 bins
- Fit parameters:
  - 3 signal parameters for helicity states: N_0, N_L, N_R
  - 3 bkg. parameters N_B for Wjets, QCD, RemBkg

  $\Big\} \hat{=}$ # of events
  - Due to morphing step: histos H depend on $\vec{k}$
- Signal samples: Consider selection efficiencies $\vec{\varepsilon}_{N_S}$
- Summarise signal and bkg templates:

$$H_S(\vec{k}, \vec{N}_S, \vec{\varepsilon}_{N_S}) = H_{N_0}(\vec{k}) \cdot N_0 \varepsilon_{N_0} + H_{N_L}(\vec{k}) \cdot N_L \varepsilon_{N_L} + H_{N_R}(\vec{k}) \cdot N_R \varepsilon_{N_R}$$
$$H_B(\vec{k}, \vec{N}_B) = H_{\text{Wjets}}(\vec{k}) \cdot N_{\text{Wjets}} + H_{\text{QCD}}(\vec{k}) \cdot N_{\text{QCD}} + H_{\text{RemBkg}}(\vec{k}) \cdot N_{\text{RemBkg}}$$

- Write in condensed form:

$$H_{\text{sum}}(\vec{k}, \vec{N}_S, \vec{N}_B, \vec{\varepsilon}_{N_S}) = H_S(\vec{k}, \vec{N}_S, \vec{\varepsilon}_{N_S}) + H_B(\vec{k}, \vec{N}_B)$$

Likelihood function:

$$-2 \ln L^{\text{prof}}_{\vec{N}_S, \vec{N}_B, \vec{k}} = 2 \sum_{m=1}^{N_{\text{bins}}} \left(H_{\text{sum},m} - H_{\text{data},m} \cdot \ln H_{\text{sum},m} \right) + \sum_{i=1}^{N_{\text{bkg}}} \frac{(N_{B,i} - N_{B,\text{exp},i})^2}{\sigma^2_{N_{B,\text{exp},i}}} + \sum_{j=1}^{N_{\text{prof}}} k_j^2 + \text{const.}$$

- Last sum *penalty term*: Gaussian constraint on k_j with $\mu = 0$, $\sigma = 1$
- Second sum: Gaussian constraint on $N_{B,i}$ with $\mu = N_{B,\text{exp},i}$, $\sigma_i = \sigma_{N_{B,\text{exp},i}}$

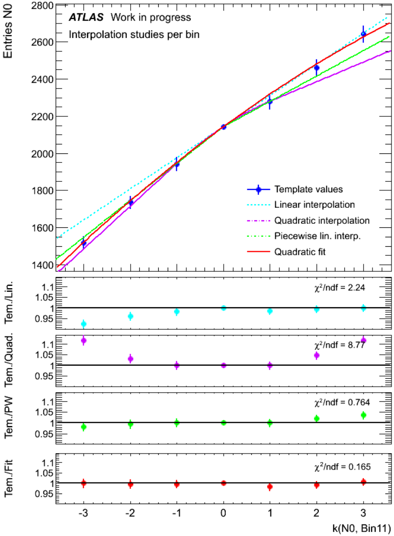
Studying Different Interpolation Methods

Approach:

- Use test systematic uncertainty: JES
- JES not split into components, as first test: toy example
- Use variations up to $\pm 3\sigma$

Evaluating interpolation methods:

- Linear: $k_j = \pm 1$ does not yield $H_{\pm 1, ij}$
- Quadratic & piecewise linear: if $|k_j| > 1$ discrepancies between histogram values and interp. curves
- Quadratic fit: Deviations small!
- χ^2/ndf-values emphasise observation



Comparison between Fitting/Interpolation Methods I

Approach:

- Estimation of systematic effects: Compare PL fit with simple delta method
- Use different interpolation methods for PL fit
- Pseudo-data based on SM expectation: $F_0 = 0.7$, $F_L = 0.3$, $F_R = 0.0$; assume $k = 0$

- Template fit:
  - Estimate systematic uncertainties using ensemble tests!
  - Perform 1,000 pseudo-experiments (PE) based on template fit w/o NPs
  - Perform 2,000 PEs for nominal and $\pm 1\sigma$-variation
  - Results of ensemble tests translated into distr. for hel. fractions
  - Take largest difference between mean of nominal and var. as uncertainty
  - Add stat. and syst. uncertainty quadratically

- Profile Likelihood Fit:
  - Perform 1,000 PEs and obtain stat.+syst. unc. directly from PE distributions

Comparison between Fitting/Interpolation Methods II

Results:

Method	Hel. Fraction	Mean	Uncertainty	Mean Pull	Pull Sigma
PL linear	F_0	0.701	0.051	0.006 ± 0.025	0.782
	F_L	0.299	0.031	0.008 ± 0.026	0.805
	F_R	0.000	0.025	-0.002 ± 0.023	0.736
PL quad. int.	F_0	0.702	0.051	0.020 ± 0.026	0.803
	F_L	0.299	0.031	-0.009 ± 0.026	0.824
	F_R	-0.001	0.025	-0.032 ± 0.024	0.759
PL quad. fit	F_0	0.701	0.054	0.010 ± 0.025	0.794
	F_L	0.299	0.033	0.003 ± 0.025	0.778
	F_R	-0.000	0.026	0.017 ± 0.025	0.806
Template fit	F_0	-	0.060	-	-
	F_L	-	0.033	-	-
	F_R	-	0.036	-	-

- Distributions do not reveal any bias ($\rightarrow$ Backup: pull distributions)
- Fits with piecewise linear interpolation do not converge
- Given uncertainty with small error: $\lesssim 0.001$

$\Rightarrow$ Profile likelihood fits lead to smaller uncertainties of the helicity fraction!

Conclusion and Outlook

- **Conclusion**
  - Presentation of a PL fit and its features; illustration of setup
  - Application: Measurement of the W boson polarisation at ATLAS
  - Studies performed with different interp. methods used during the fit
  - Quadratic fit describes variations best
  - Total unc. on helicity fractions can be reduced by applying PL fits compared to template fit using ensemble tests to estimate syst. unc.
- **Outlook**
  - Add more different systematics via nuisance parameters to the PL fit
  - Check whether a systematic effect can be profiled:
    $\Rightarrow$ Solely continuous sources of systematic effect are appropriate (e.g. effects related to detector modelling)
    $\Rightarrow$ Discrete sources need to be evaluated differently (e.g. MC generator, parton showering)
- **In sum:** PLFs provide adequate possibility to reduce uncertainties, constitute promising tool for future analyses

BACKUP

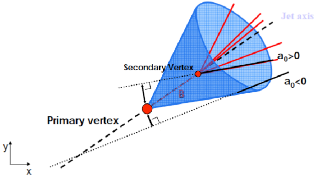
Backup Slides: Further Plots and Tables

Fundamentals of the W Polarisation Measurement

Data: collected with ATLAS at $\sqrt{s} = 7$ TeV in 2011 based on $\int \mathcal{L}\mathrm{d}t = 4.7\ \mathrm{fb}^{-1}$

Monte Carlo (MC) samples:
- MC samples (MC11c) used to calculate acceptances and bkg. contributions
- Top reference cross-sections: MSTW 2008 NLO PDF set at $m_t = 172.5\ \frac{\mathrm{GeV}}{c^2}$
- $t\bar{t}$ samples for all three signal templates: PROTOS

Definition of physics objects:
- Muons: exploit 3 level trigger, $p_\mathrm{T} > 25$ GeV, $|\eta| < 2.5$, hit and isolation req.
- Jets: use anti-k_t algorithm, $\Delta R = 0.4$,
  $p_\mathrm{T} > 25$ GeV, $|\eta_\mathrm{det}| < 2.5$, $|\mathrm{JVF}| > 0.75$
- $E_\mathrm{T}^\mathrm{miss}$: due to ν, energy scale dependent
- b-tagging: *MV1* tagger, $\varepsilon_b = 70\%$

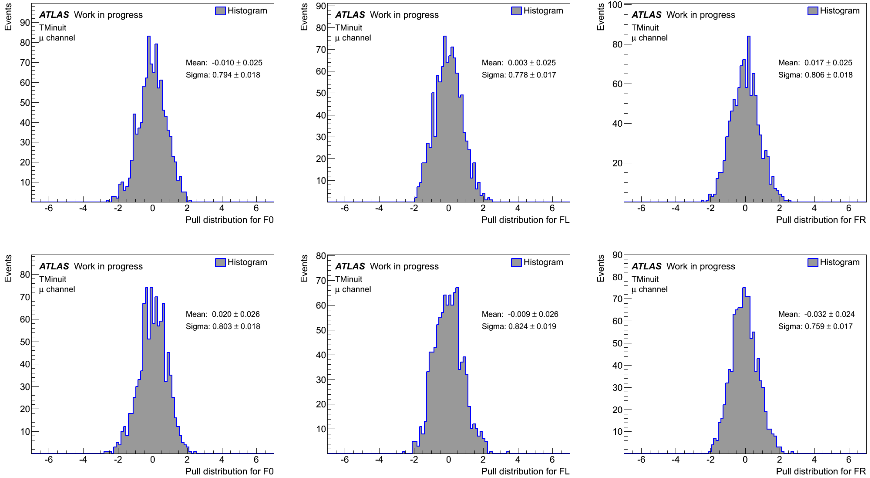


Event selection:
Trigger requirements, $p_\mathrm{T}(\mu) > 20$ GeV,
overlap removal, $p_\mathrm{T}(\mathrm{jet}) > 25$ GeV, $E_\mathrm{T}^\mathrm{miss} > 25$ GeV,
QCD multijet veto $m_\mathrm{T}(W) + E_\mathrm{T}^\mathrm{miss} > 60$ GeV, at least 1 b-tagged jet, LAr noise bursts

Event Reconstruction: Use kinematic likelihood provided by KLFitter tool

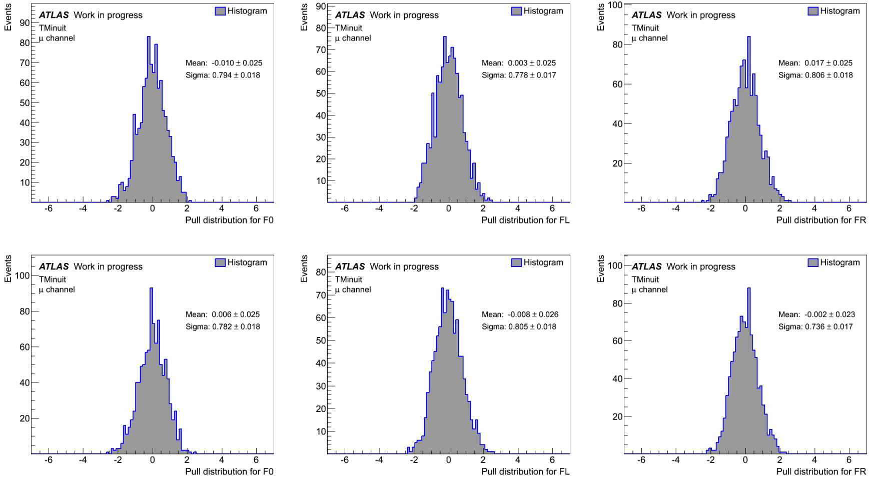
Pull Distributions: Quad. Fit and Quad. Interp.

Compare pull distributions based on quadratic fit (top) and quadratic interp. (bottom)



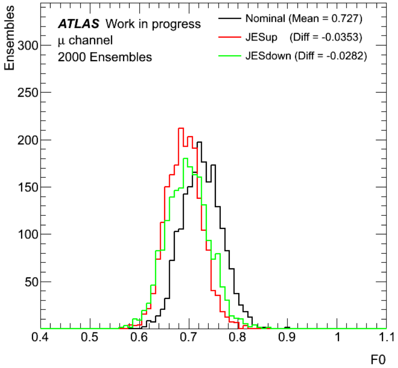
Pull Distributions: Quad. Fit and Linear Interp.

Compare pull distributions based on quadratic fit (top) and linear interp. (bottom)

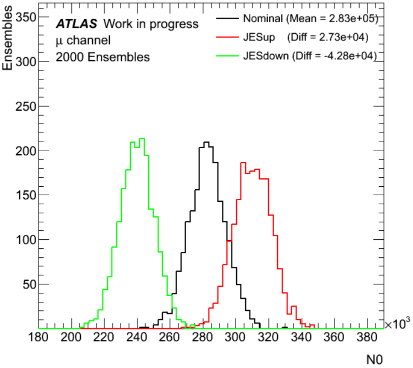


Ensemble Tests to Estimate JES Uncertainty on F_0 and N_0

JES uncertainty on F_0

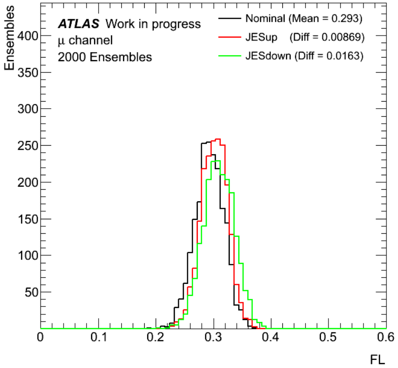


JES uncertainty on N_0

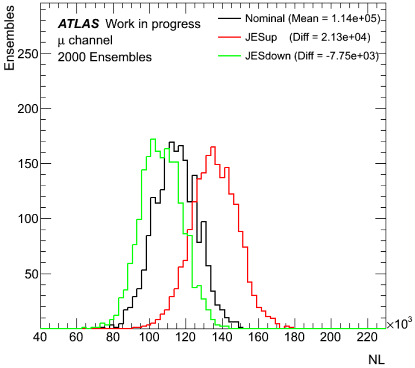


Ensemble Tests to Estimate JES Uncertainty on F_L and N_L

JES uncertainty on F_L

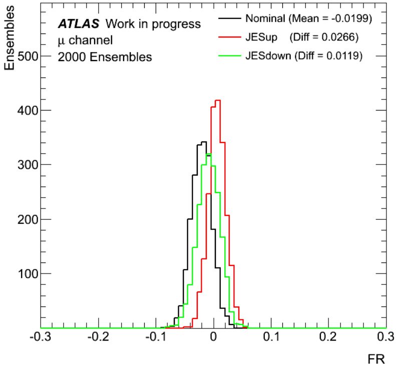


JES uncertainty on N_L



Ensemble Tests to Estimate JES Uncertainty on F_R and N_R

JES uncertainty on F_R



JES uncertainty on N_R