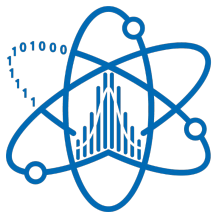




National Research
**Tomsk
State
University**



**Лаборатория
анализа данных
физики высоких энергий**

Томского
государственного
университета

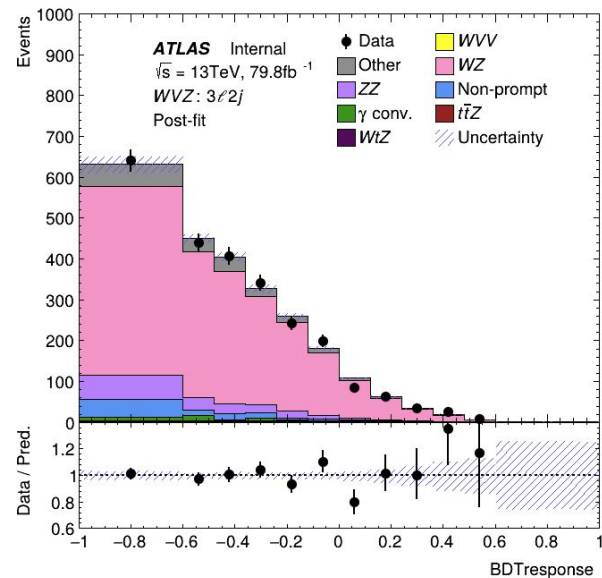
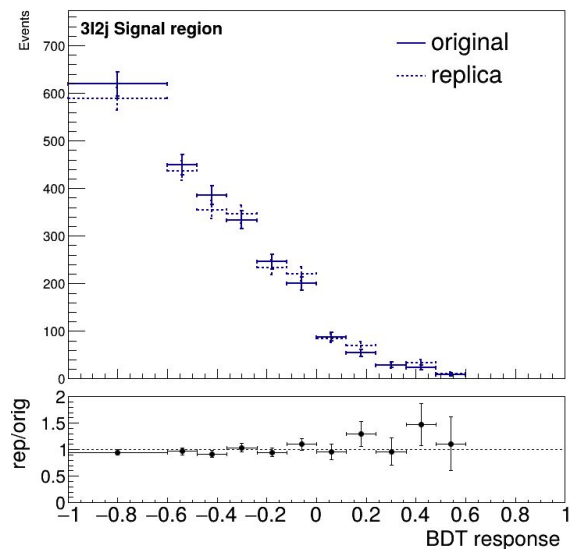
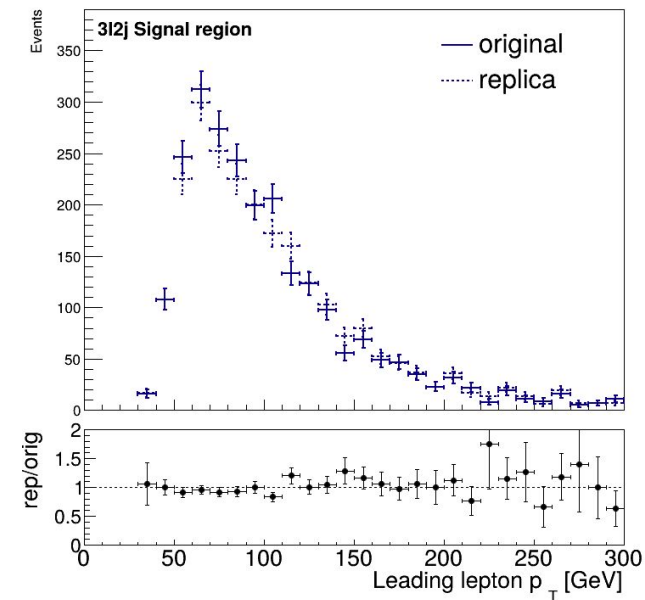
Физический анализ данных

Томский Государственный Университет

Мария Диденко

Distribution of variables: signal region

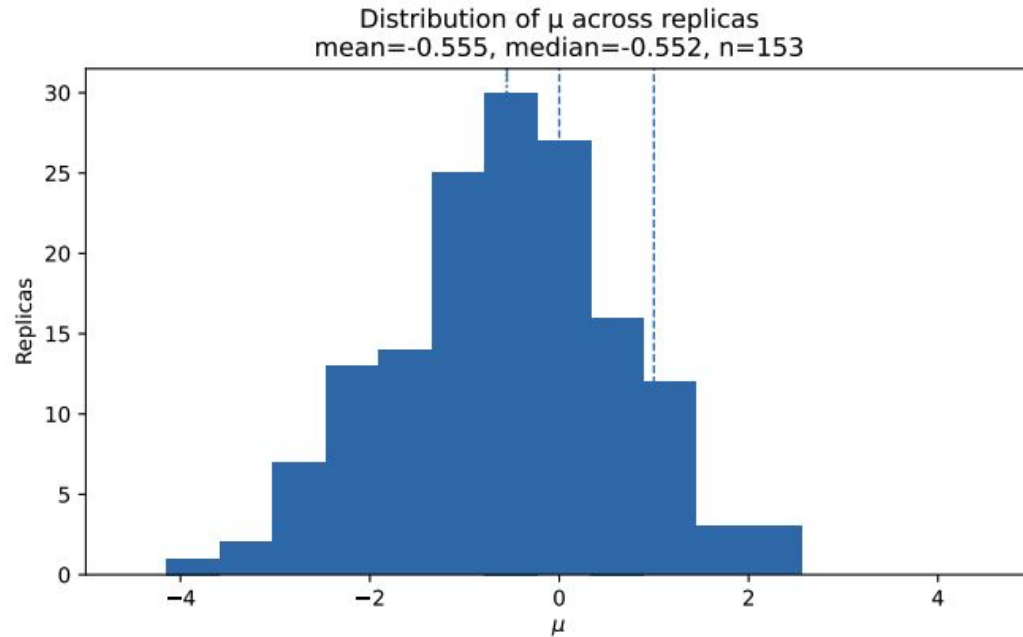
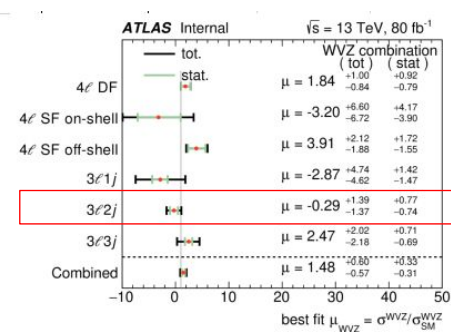
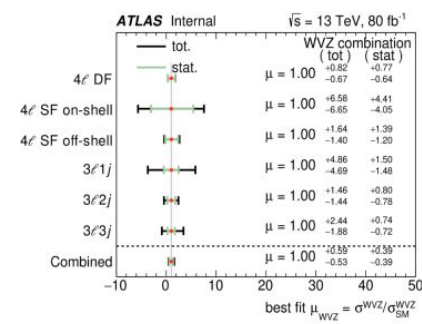
- 3l2j SR selection is applied
- Events after SR selection: orig=2438, repl=2407
- The shapes agree within statistical fluctuations (ratio plot: replica/original $\approx 1 \pm \text{stat}$)



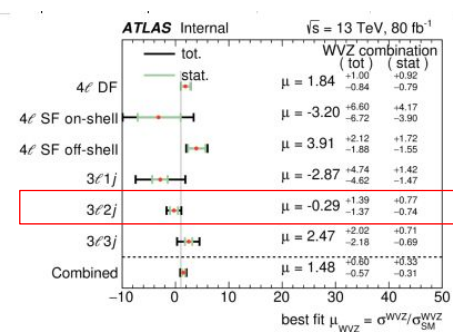
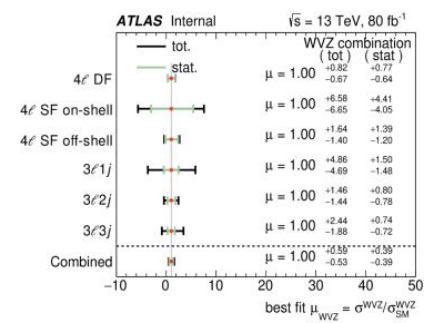
μ distribution

Most of tasks are still running

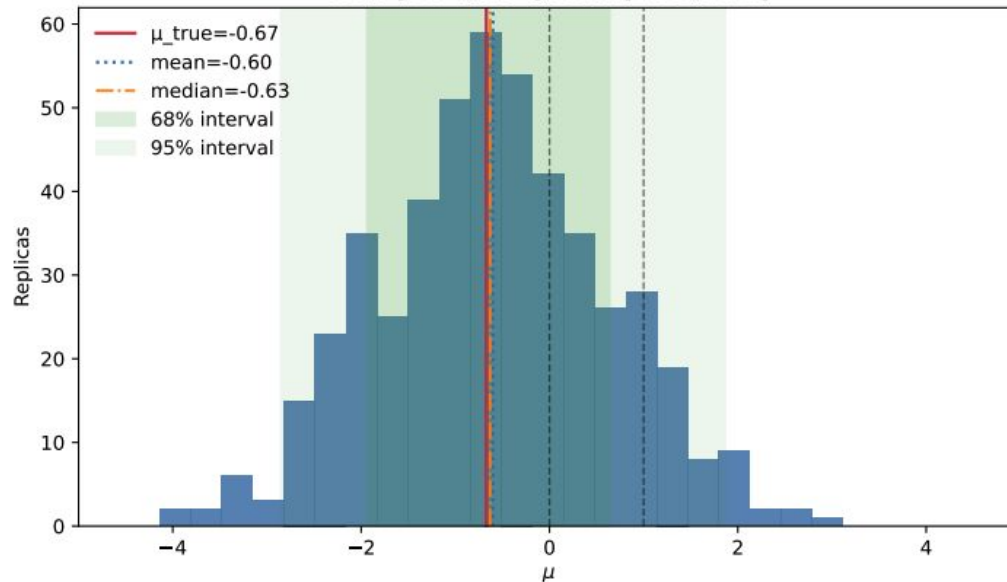
Checked 153 tasks



μ distribution



Distribution of μ across replicas
 mean=-0.60, median=-0.63, $\sigma=1.23$, var=1.51, n=486
 68%: [-1.94, 0.64], 95%: [-2.85, 1.87]



Leptons + at least 3 jets (real data)

- Checked the remaining $3\ell + \geq 3j$ region using **real data**.
- In the configuration file, the variable **newBDTG_15_3l3j** was specified, but it was **missing in the ntuples**.
- Used **newBDTG_32_3l3j_1** from the ntuples instead.
- The obtained signal strength ($\mu = 2.57$) is close to the reference value ($\mu = 2.47$)

###--- 3L preselection && at least 3 jets ---###

Region: three_lep_presel_atLeast_3jets

Type: SIGNAL

DataType: DATA

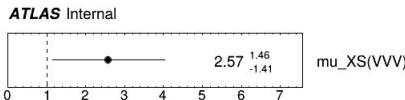
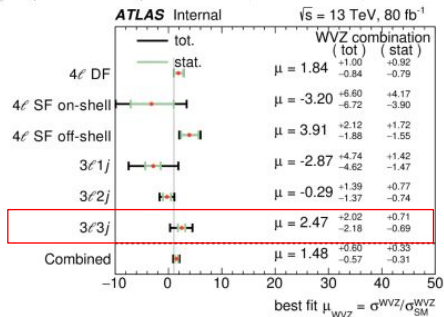
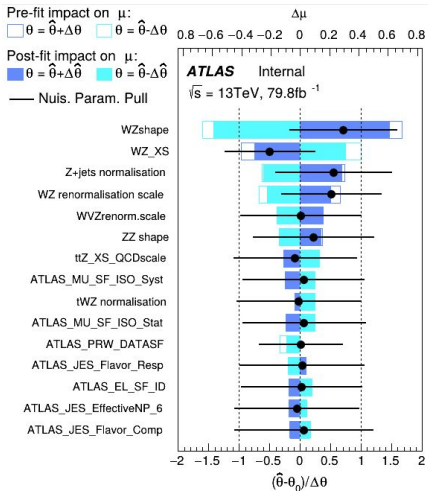
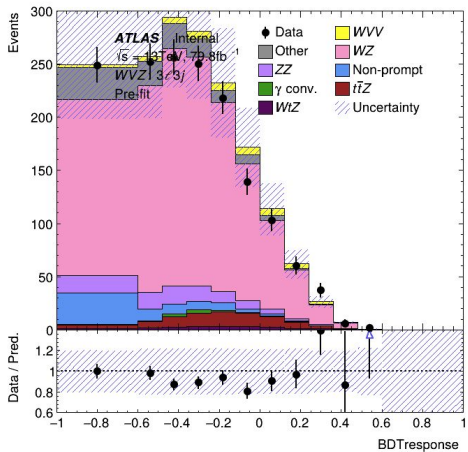
Variable: newBDTG_15_3l3j,13,-1.,1.

VariableTitle: BDT response

newBDTG_28_3l3j
newBDTG_32_3l3j
newBDTG_32_3l3j_1

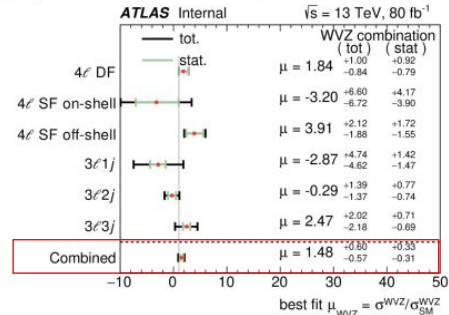
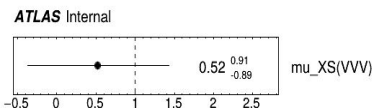
in the ntuples

in the config



Combined regions (real data)

- Combined results from **three regions**: $3\ell + 1j$, $3\ell + 2j$, and $3\ell + \geq 3j$.
- Used **real data** to verify consistency of the combined fit.
- The **post-fit BDT distributions** (bottom plots) show good agreement between data and prediction across all regions.
- The obtained **combined signal strength** is $\mu = 0.52$, while the **reference value** reported in the publication is $\mu = 1.48$:
 - The **difference mainly comes from the $3\ell + 2j$ region**:
 -0.67 vs -0.29



Pre-fit impact on μ :

$\square \theta = \hat{\theta} + \Delta\theta$ $\square \theta = \hat{\theta} - \Delta\theta$

Post-fit impact on μ :

$\blacksquare \theta = \hat{\theta} + \Delta\hat{\theta}$ $\blacksquare \theta = \hat{\theta} - \Delta\hat{\theta}$

— Nuis. Param. Pull

WZ renormalisation scale

Z+jets normalisation

ttZ_XS_QCDscale

ttWZ normalisation

γ (threelep+res1jet bin 1)

ZZ shape

WZshape

WZ_XS

ATLAS_JES_EffectiveNP_5

ATLAS_MU_SF_ISO_Syst

ATLAS_MU_SF_ISO_Stat

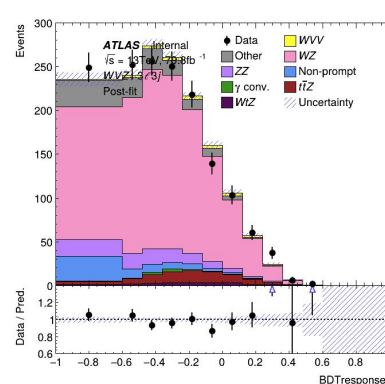
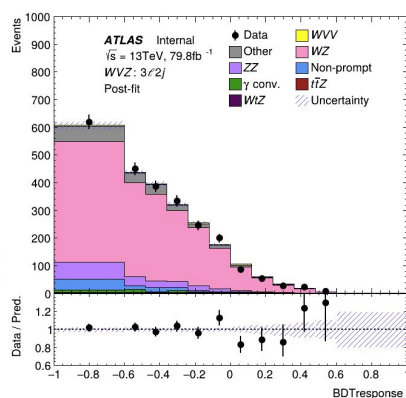
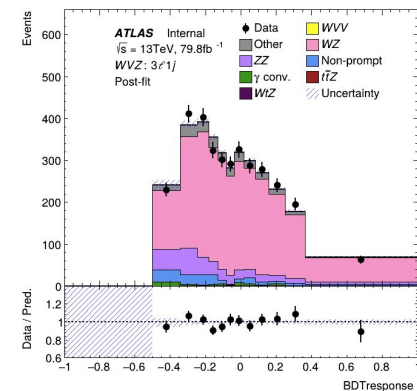
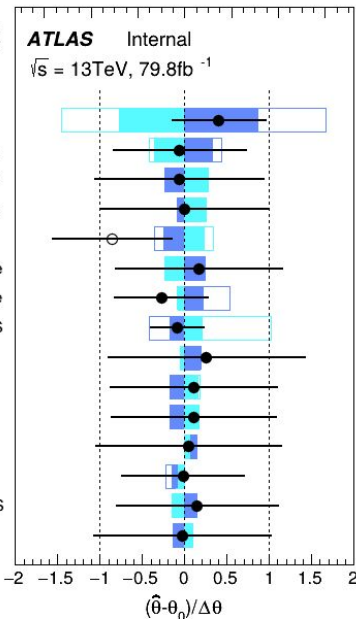
ATLAS_JES_Flavor_Comp

ATLAS_PRW_DATASF

ttZ_XS

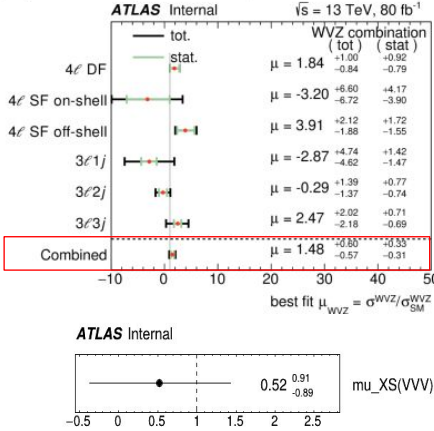
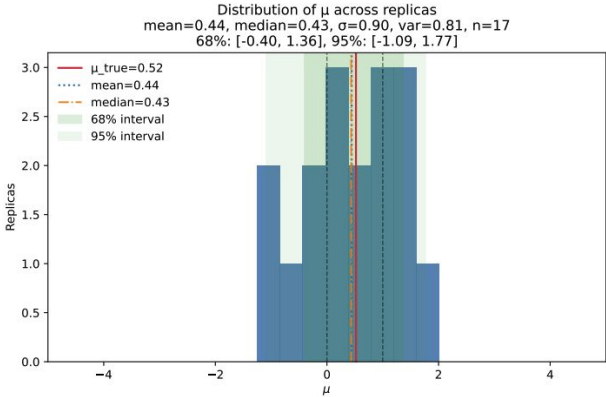
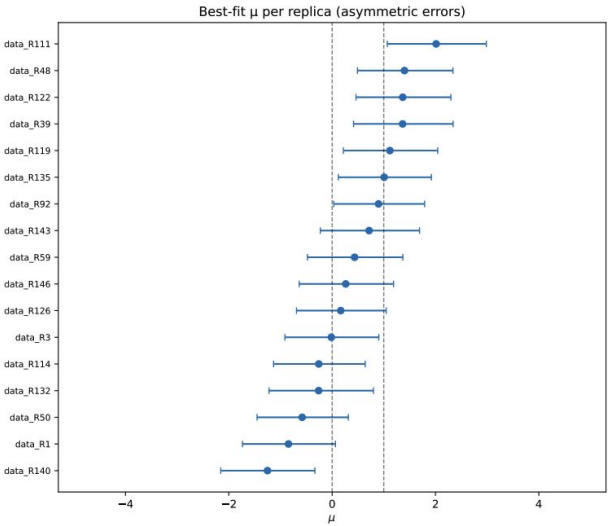
ATLAS_JES_EffectiveNP_6

$\Delta\mu$
 -0.6 -0.4 -0.2 0 0.2 0.4 0.6



Combined regions (replicas)

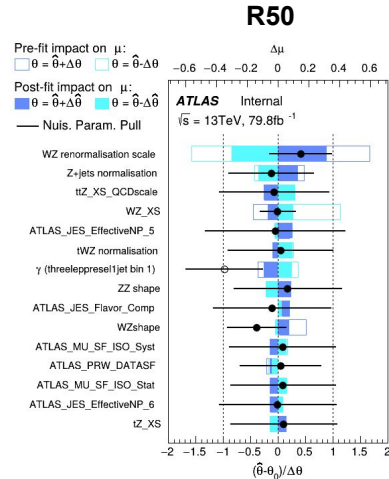
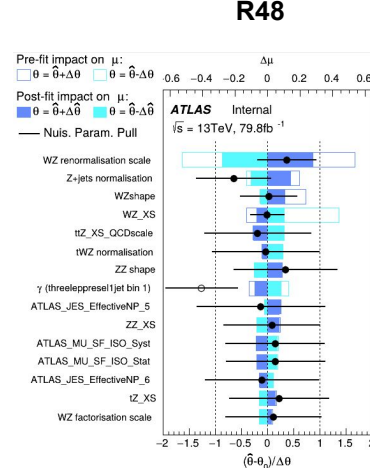
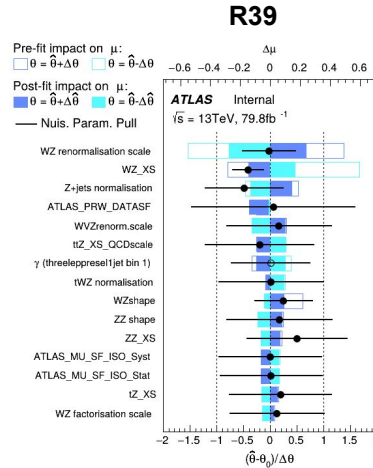
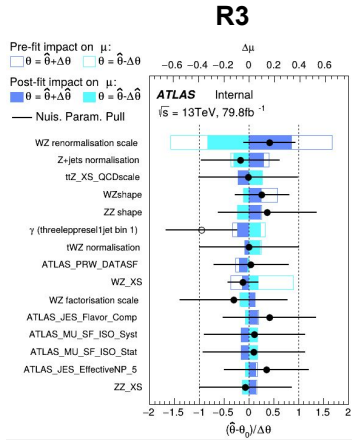
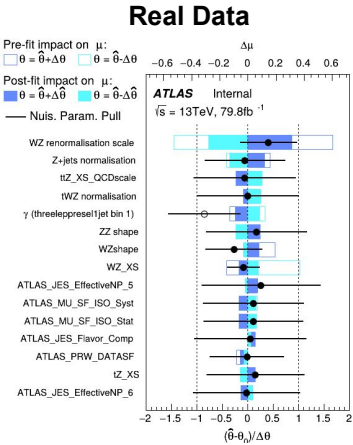
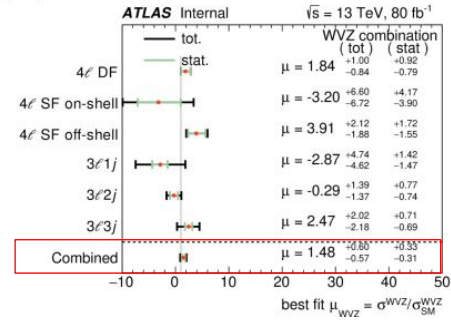
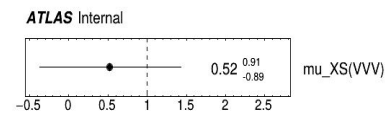
- Many replicas **fail during generation**, when TRExFitter tries to run all options simultaneously (**nwfsdpr**).
- Currently, I'm running **one option at a time (n)**, and then plan to process the remaining ones.
- So far, **17 replicas** have been successfully generated.
- The **distribution of μ** is close to the original result (**$\mu = 0.52$**), but **more statistics are needed** to confirm the stability of the result.



Replicas: 17
 μ mean: **0.4429**, median: 0.4341, std: 0.9023
Avg errUp: 0.9335, Avg [errDown]: 0.9045
-- vs $\mu_{\text{true}}=0.52$ --
Bias : -0.07706
RMSE : 0.8787
Std/Var : 0.9023 / 0.8141
Replicas coverage 68% : 76.5%
Replicas coverage 95% : 100.0%
Global 68% CI [-0.402, 1.36] => hit? True
Global 95% CI [-1.09, 1.77] => hit? True

Combined regions (replicas)

Each replica includes random statistical variations in data, which slightly change the fitted nuisance parameters and their impact on μ .



Measurement of the total and differential cross-sections of ttW production: a good starting point

- Ready bootstrap replicas already included in the workspace ✓
- Working fitting scripts available ✓
- Regularization disabled, fitting launched (XRootFit) ✓
 - evaluate the statistical variation of the data
- Running 1000 replicas: in progress ▶▶

Fit results (without regularization):

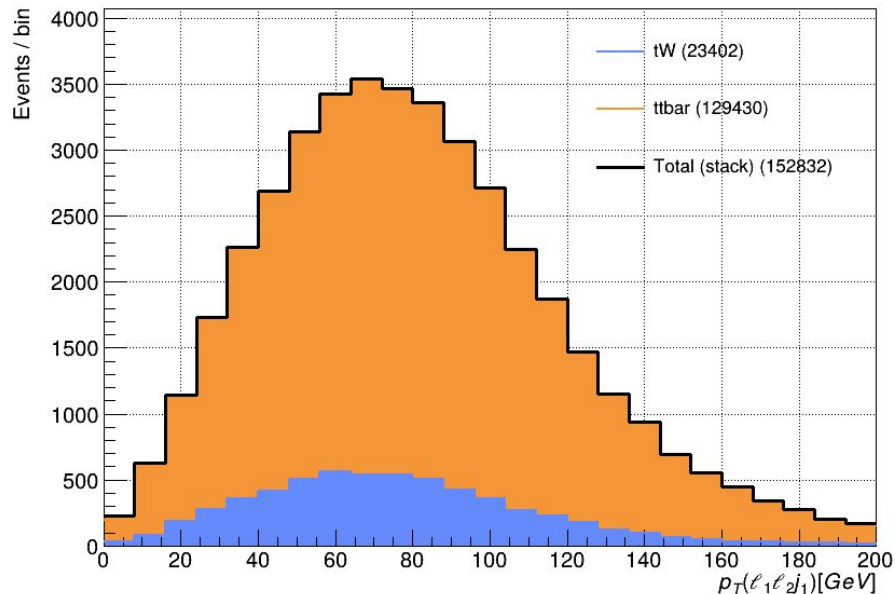
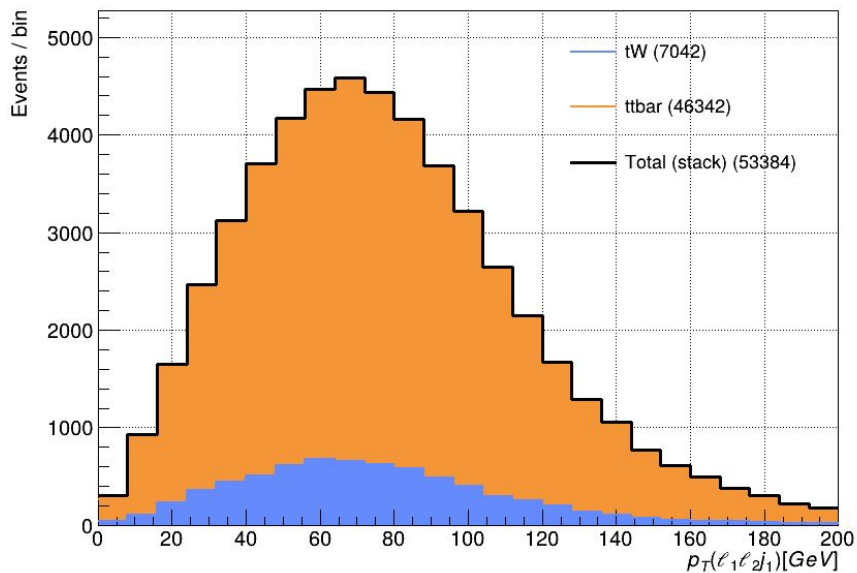
- NLL: 454.056
- Norm_ttW = 1.1393 ± 0.0932

Measurement of t-channel production of single top quarks and antiquarks

- JSON file with data: understanding of its structure is required (or conversion JSON → ROOT workspace)
- Several fitting options available: PyHF or TRExFitter
- repeating the ttW structure: conversion JSON → YAML workspace

BDT ntuples

- **Additionally:** the ntuples for the BDT are ready, and the statistics have been increased by a factor of three.

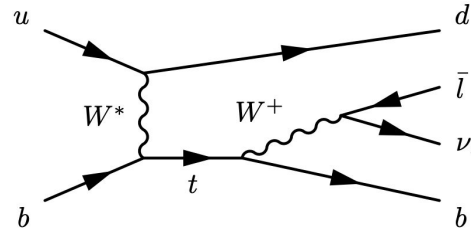
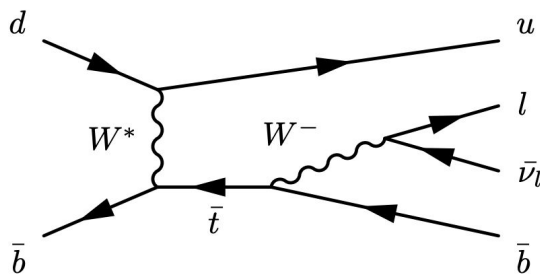


.....

16.12.2025

tW analysis overview

- This analysis uses **140 fb⁻¹ of ATLAS 13 TeV** data to measure the production cross-sections of tW processes separately **for $tq \rightarrow tW^+$ and $t\bar{q} \rightarrow tW^-$ final states**.
 - The separate measurements provide enhanced sensitivity to the u- and d-quark PDFs, since the dominant initial states differ for tW^+ ($u \rightarrow d$ transition) and tW^- ($d \rightarrow u$ transition).
- Events are selected in the single-lepton final state with one charged lepton, MET, and b-tagged jets.
- A **neural network (NN)** is trained to distinguish tW signal from background using event-level kinematic variables.
- The **NN output** distribution is then used as the discriminant **in a profile likelihood fit**.



Region strategy

Two complementary signal regions tailored to the angular correlation between the lepton and the b-jet:

- **SR-plus (SRp):** Events where the lepton and the b-jet are preferentially aligned (sensitive to **tW⁺ production**).
- **SR-minus (SRn):** Events with opposite angular correlation (sensitive to **tW⁻ production**).

This separation increases PDF sensitivity and improves constraints on the signal model.

Main Background Processes

- **t \bar{t} (top–antitop):** dominant background in single-lepton, b-jet final states.
- **Single top (tW, t-channel):** important and must be modelled accurately, especially in 1-b-tag categories.
- **W+jets:** critical for regions with one lepton and MET.
- **Z+jets / Diboson:** typically subdominant but included.
- **Fake leptons / charge mis-ID:** included where relevant.

CR name	Requirement
B-e-plus	$q_e/e = +1, \eta(e) < 1.37, E_T^{\text{miss}} < 30 \text{ GeV}$
B-e-minus	$q_e/e = -1, \eta(e) < 1.37, E_T^{\text{miss}} < 30 \text{ GeV}$
EC-e-plus	$q_e/e = +1, \eta(e) > 1.52, E_T^{\text{miss}} < 30 \text{ GeV}$
EC-e-minus	$q_e/e = -1, \eta(e) > 1.52, E_T^{\text{miss}} < 30 \text{ GeV}$
CR μ -plus	$q_\mu/e = +1, 28 \text{ GeV} < p_T(\mu) < 40 \text{ GeV} \cdot \frac{ \Delta\phi(j_1, \ell) }{\pi}$
CR μ -minus	$q_\mu/e = -1, 28 \text{ GeV} < p_T(\mu) < 40 \text{ GeV} \cdot \frac{ \Delta\phi(j_1, \ell) }{\pi}$

Table 1: Summary of the definition of the CRs.

6 Control regions

BCCI implementation

HEPData Record: [ins2764820](#)

- **workspace.json** is a JSON specification of the statistical model.
- **8 channels**: signal and control regions:
SRp, SRn, SRelép, SRelépforw, SRmuonp, SRelen, SRelenforw, SRmuonn
- **44 bins** distributed across the 8 channels
- **Observed data** provided per bin
- **Expected model**: signal + background + systematics (400+ nuisance parameters)
- **Parameter of Interest (POI)**: **negSigXsecOverSM**

Full workflow

1. Load the JSON workspace
2. Generate Poisson bootstrap replicas
3. Modify the observed data in the workspace
4. Build the pyhf statistical model
5. Perform a maximum likelihood fit (MLE)
6. Extract the POI value and uncertainties
7. Repeat the procedure 1000 times
8. Compute BC / BCa confidence intervals

```
{
  "channels": [
    {
      "name": "SRp",      // Название канала (сигнальный регион)
      "samples": [       // Физические процессы
        {
          "name": "signal", // Сигнальный процесс
          "data": [0.1, 0.2, ...], // Ожидаемые события по бинам
          "modifiers": [...] // Систематические вариации
        },
        {
          "name": "background",
          "data": [10.5, 15.3, ...],
          "modifiers": [...]
        }
      ]
    }
  ]
}
```

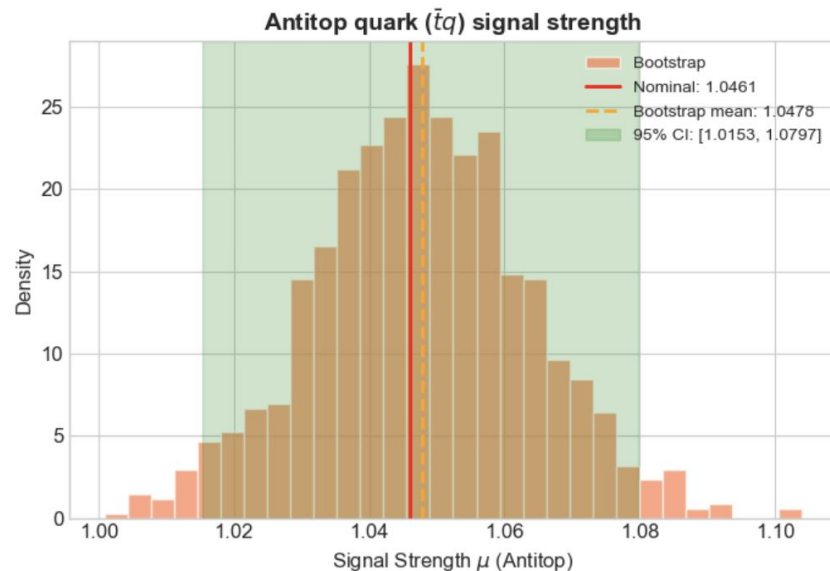
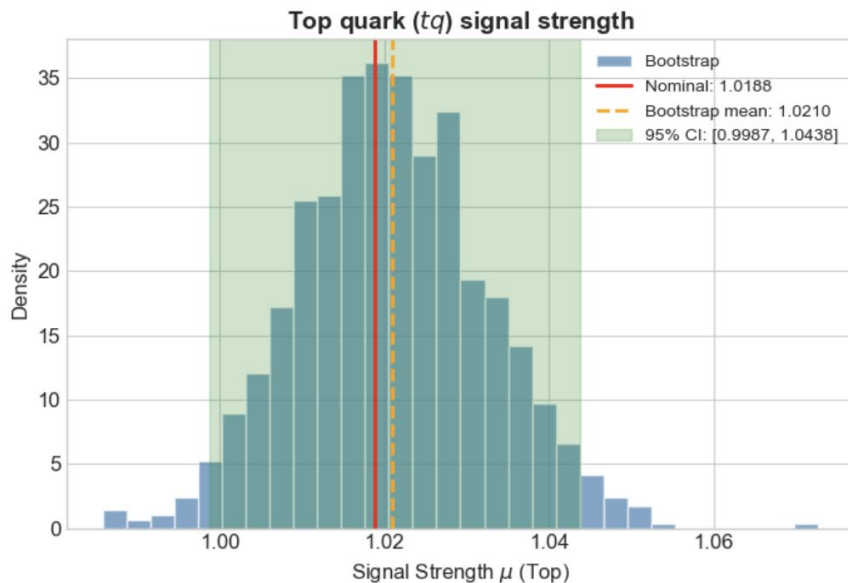
From the HEPData workspace (workspace_fixed.json):

Type	Count	TRexFitter Type	Examples
lumi	1	LUMI	lumi
staterror	8	STATERORR	staterror_SRp
normsys	208	OVERALL	sitop_mur, JET_*
histosys	207	HISTO	weight_bTagSF_*
normfactor	5	NormFactor	negSigXsecOverSM

Total: 429 systematic parameters

BCCI implementation

- **Data** $[n_1, n_2, n_3, \dots]$ per bin \rightarrow **varied**
- **Systematic uncertainties**: all normsys, histosys, luminosity, etc. \rightarrow **fixed**
- **Likelihood model**: includes **nuisance parameters**



BC/BCa implementation

$$\mu_{BC}[\alpha] = \hat{G}^{-1} \left(\Phi \left(2z_0 + z^{(\alpha)} \right) \right)$$

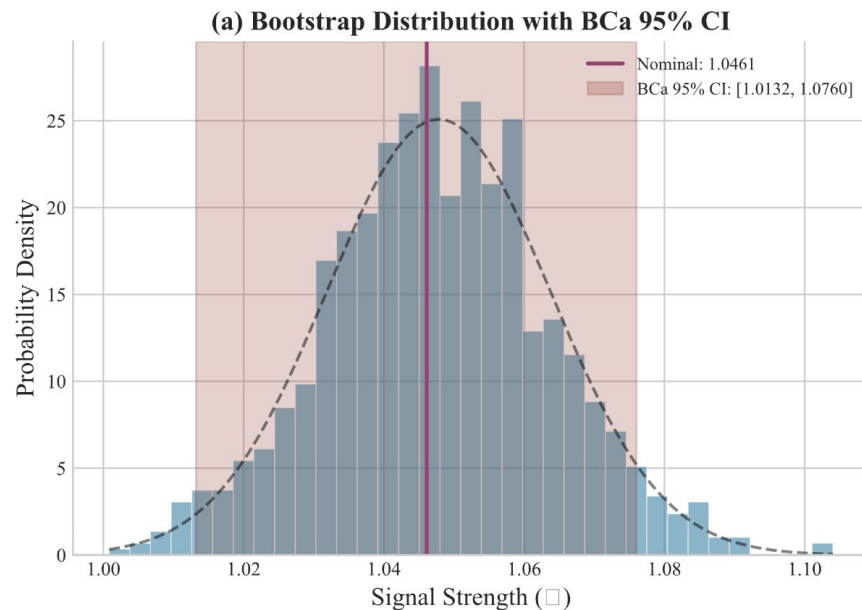
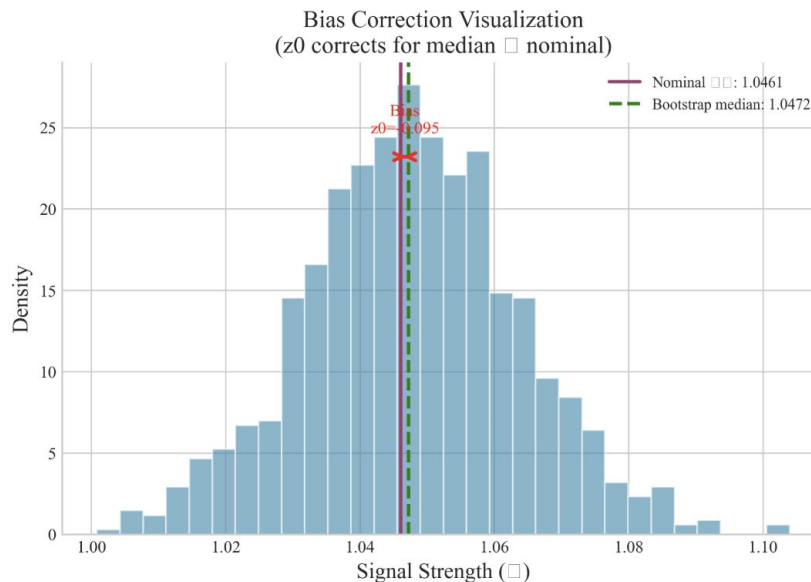
Antitop quark ($\bar{t}q$):

$z_0 = -0.0954$, $a = 0.0112$

Standard : [1.0166, 1.0790] (width=0.0624)

Percentile : [1.0153, 1.0797] (width=0.0644)

BC : [1.0132, 1.0760] (width=0.0628)

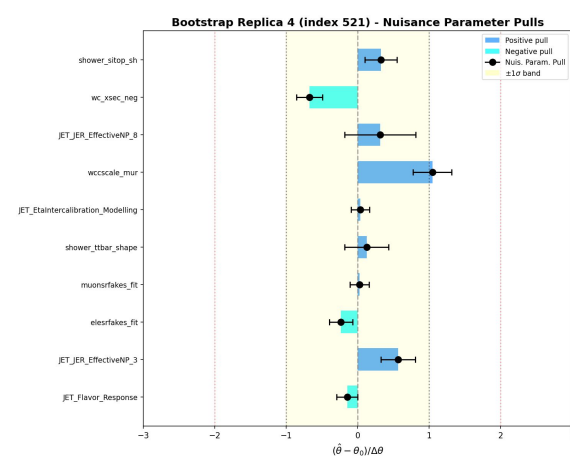
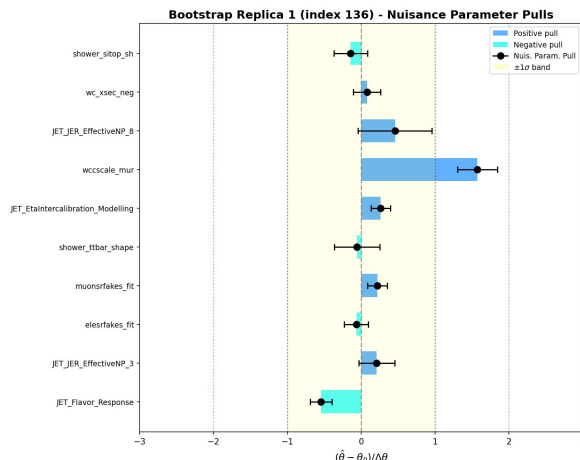
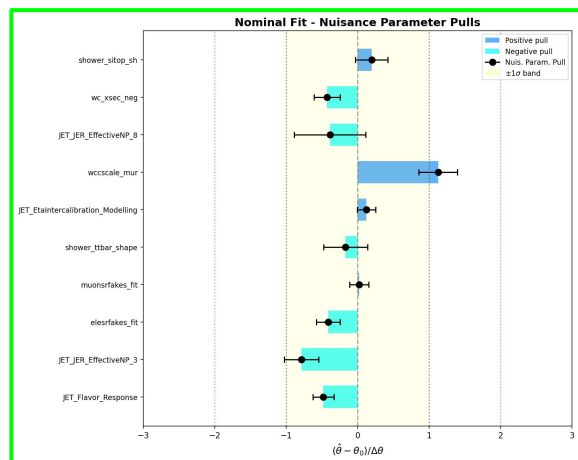


BCCI implementation

The ranking of nuisance parameters shows a **clear data-dependent behavior** of systematic uncertainties.

The **Jet_JER_Effec...** nuisance parameter is particularly **sensitive to data fluctuations** and may require a more accurate treatment or dedicated validation.

A large fraction of systematic uncertainties shows **negligible impact** on the likelihood and can potentially be **safely neglected** in simplified models.



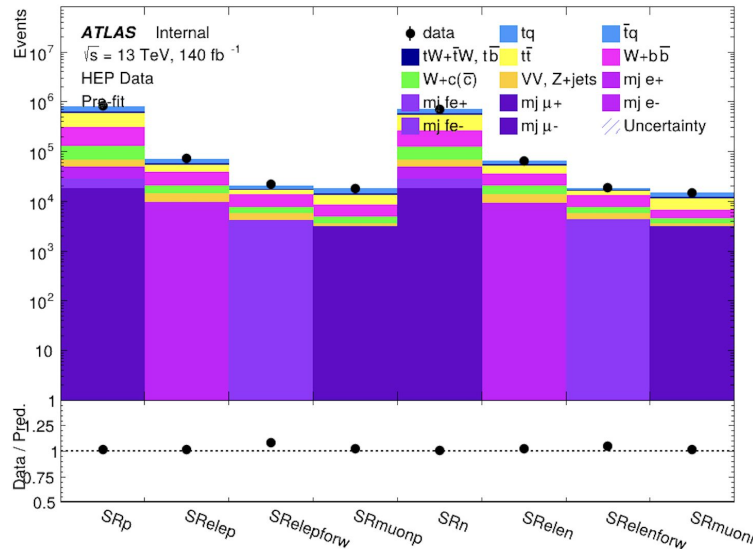
TRExfitter cross-check

- **JSON → ROOT conversion:** the converted ROOT file currently contains **only histograms**, with **variable bin widths** matching the reference.
- **TRExFitter setup:** The fit was run in **StatOnly = TRUE** mode, i.e. **systematic uncertainties were disabled** and only statistical uncertainties were considered.
- **Next step:** To reproduce the full result, we must include **all systematic uncertainties** in the workspace (400+ nuisance parameters), i.e. enable **normsys**, **histosys**, luminosity, etc., and run the **full profile-likelihood fit** in TRExFitter.

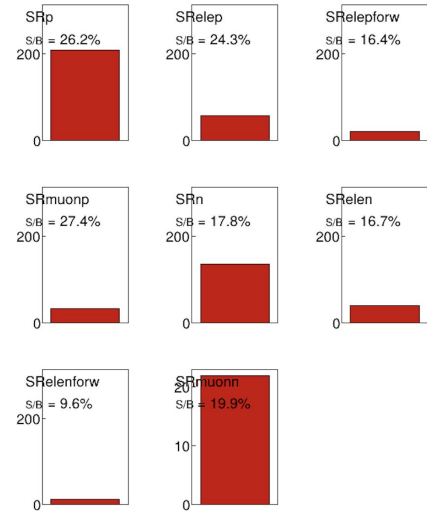
NUISANCE_PARAMETERS

$\mu_{t\bar{b}q}$: **1.0476** $+0.005$ -0.005

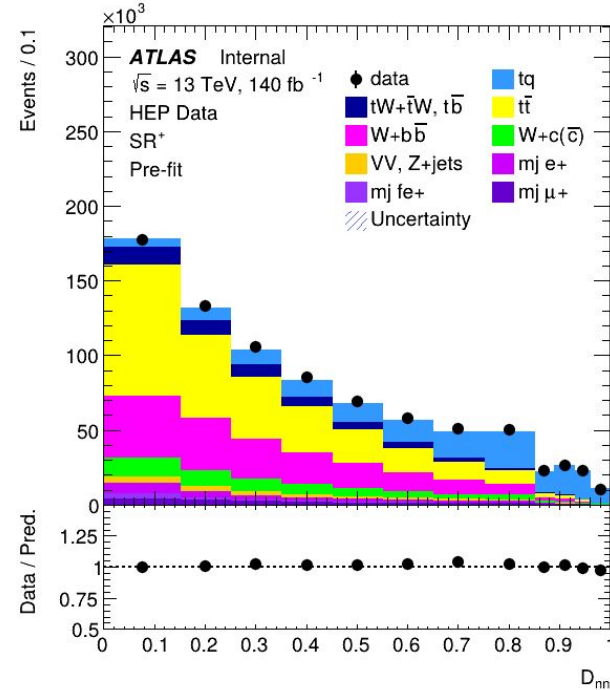
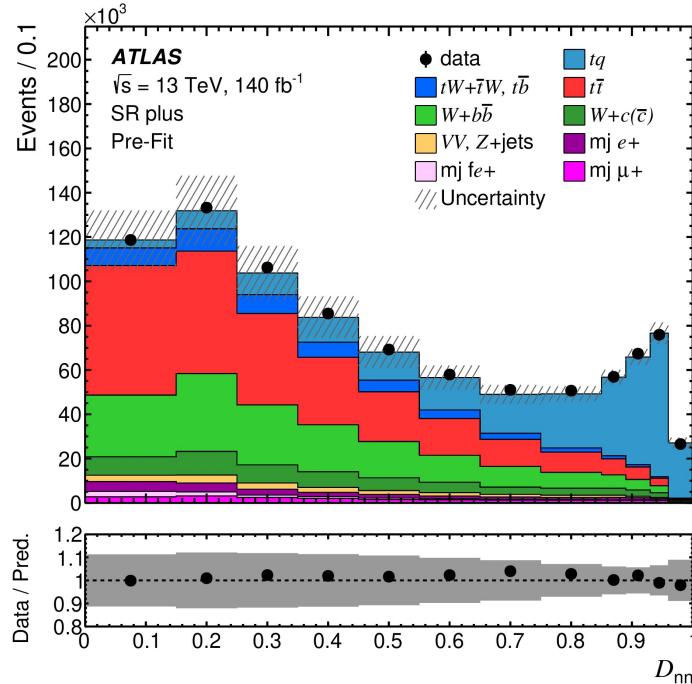
μ_{tq} : **1.03865** $+0.0035$ -0.004



ATLAS Simulation Internal
 $\sqrt{s} = 13 \text{ TeV}, 140 \text{ fb}^{-1}$
 HEP Data



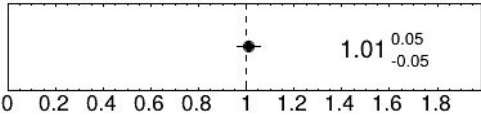
Different distribution: first and last bins don't show correct number of events.



- Issue with first and last bins was fixed by applying correct normalisation
- All sources of systematics were included: at first look everything is OK
- The distribution of DNN is the same as original

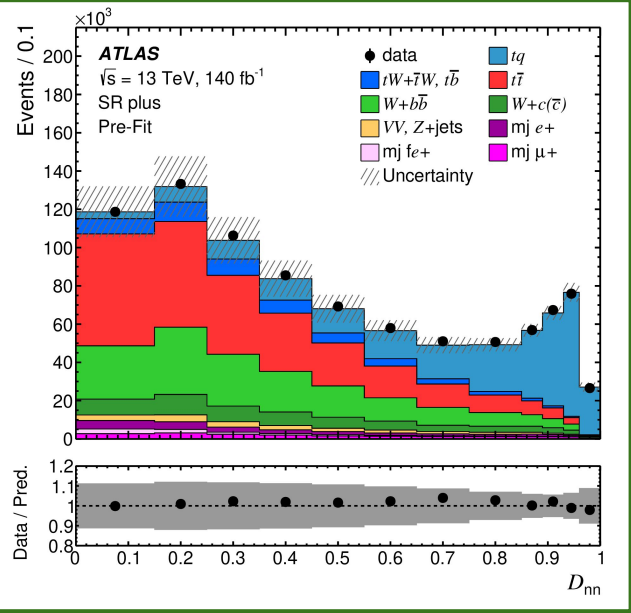
both SRs are included

ATLAS Internal

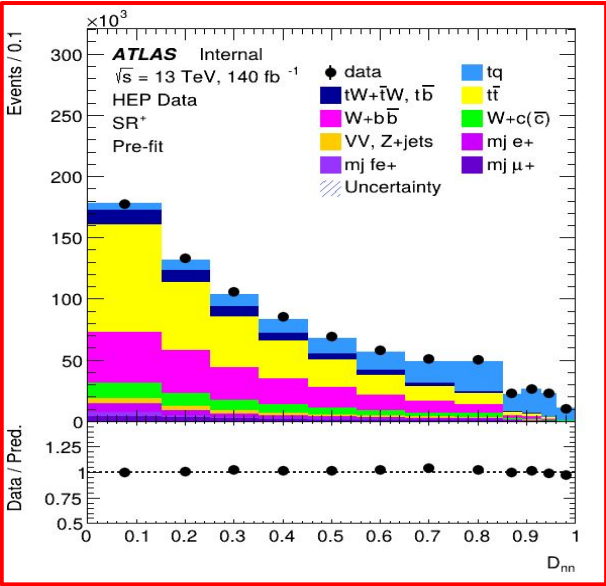


Signal strength

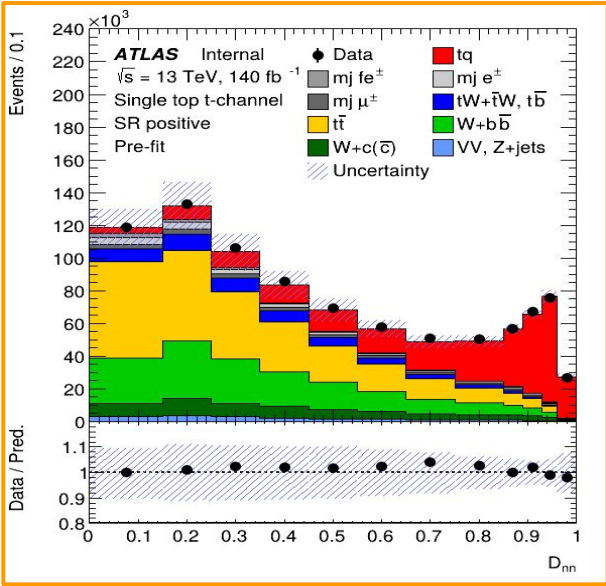
Reference Article



Issue



Fixed



Trexfitter: SR minus

Job: "HEPData_tchan"

Label: "Single top t-channel"

CmeLabel: "13 TeV"

LumiLabel: "140 fb⁻¹"

POI: "mu_sig"

ReadFrom: HIST

HistoPath: "/home/users/mdidenko/HEPData_BBCI/hepdata_bcci/workspace"

HistoFile: "histograms"

OutputDir: "results/trexfitter"

DebugLevel: 1

SystControlPlots: TRUE

SystErrorBars: TRUE

RankingMaxNP: 15

SystPruningShape: 0.005

SystPruningNorm: 0.005

SystLarge: 1

CorrelationThreshold: 0.0001

SystDataPlots: TRUE

SystControlPlots: TRUE

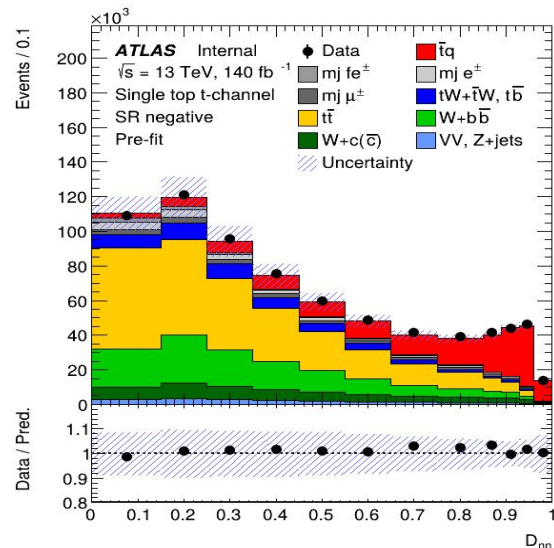
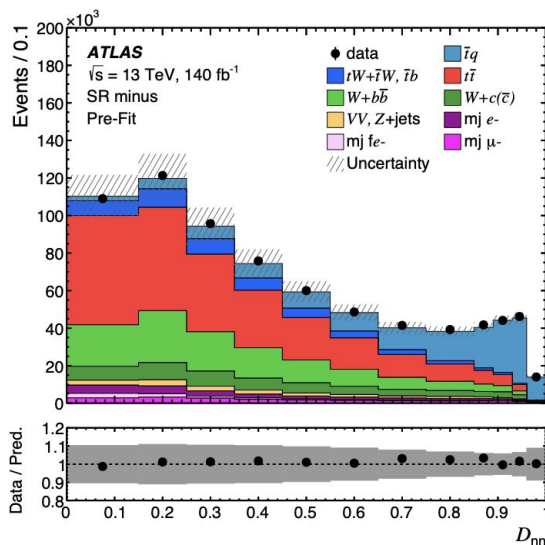
SystDataPlots: TRUE

UseGammaPulls: TRUE

PlotOptions: CHI2, NOXERR, NOENDERR, POISSONIZE

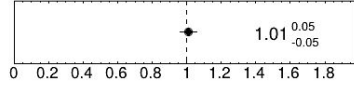
RatioYmin: 0.8

RatioYmax: 1.2



Trexfitter results

ATLAS Internal



Signal strength

Unc. component on μ :

$\sigma_{\text{POI}}^{\text{up}} \sigma_{\text{NP}}^{\text{up}}$ $\sigma_{\text{POI}}^{\text{dn}} \sigma_{\text{NP}}^{\text{dn}}$
w/ ρ = linear correlation coef.

— Nuis. Param. Pull

shower_sitop_pth1_norm

shower_sitop_norm

shower_sitop_pth1_sh

JET_JER_EffectiveNP_1

lumi_uncertainty

JET_Pileup_OffsetMu

ET_EtaIntercalibration_Modelling

wc_xsec_pos

wc_xsec_neg

γ (SRp bin 9)

weight_fjvt

γ (SRp bin 10)

JET_JER_EffectiveNP_7

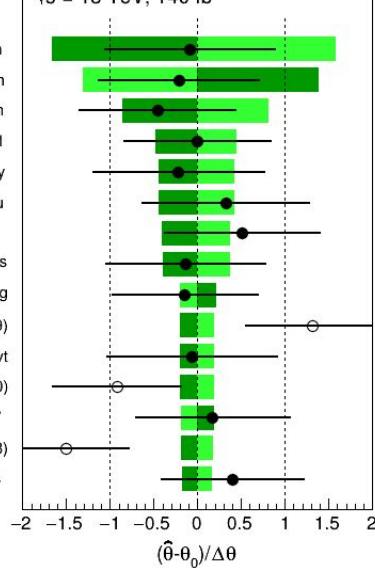
γ (SRp bin 8)

JET_JER_EffectiveNP_4

$\Delta\mu$
-0.03 -0.02 -0.01 0 0.01 0.02 0.03

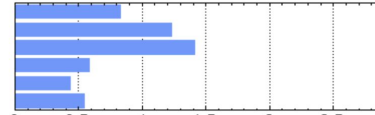
ATLAS Internal

$\sqrt{s} = 13 \text{ TeV}, 140 \text{ fb}^{-1}$



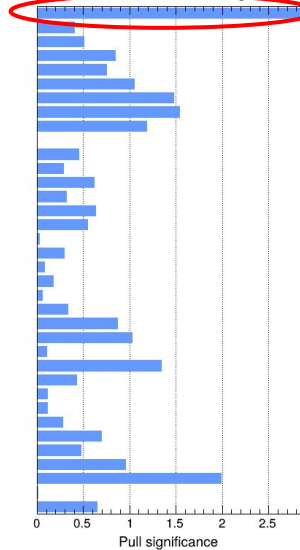
ATLAS Internal

Theory



ttbar_mur_norm
ttbar_muf_norm
shower_ttbar_norm
shower_swchan_norm
shower_sitop_pth1_norm
shower_sitop_norm

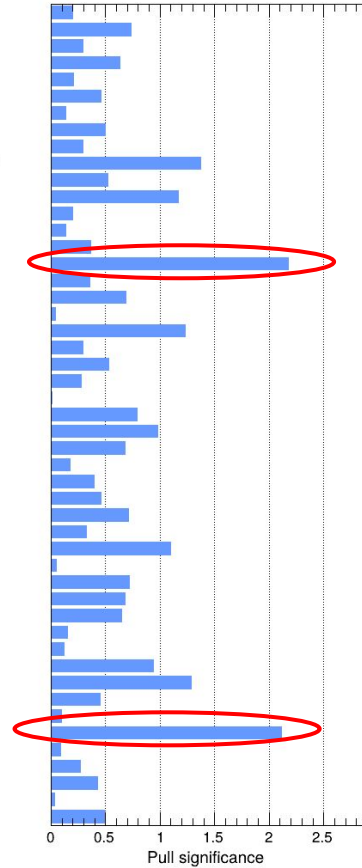
ATLAS Internal Modelling



jetsdiboson_xsec
wjets_xsec_pos
wjets_xsec_neg
ttbar_pth1
ttbar_mur_shape
ttbar_muf_shape
ttbar_hdamp
ttbar_fsr
ttbar_Var3c
swchan_mur
swchan_muf
swchan_fsr
swchan_Var3c
swchan_DS
sitop_mur
sitop_muf
sitop_fsr
sitop_Var3c
shower_ttbar_shape
shower_swchan_shape
shower_sitop_sh_CR
shower_sitop_pth1_sh_CR
shower_sitop_pth1_sh
PDF4LHC28
PDF4LHC22
PDF4LHC19
PDF4LHC18
PDF4LHC12
PDF4LHC11
PDF4LHC09
PDF4LHC08
PDF4LHC06
PDF4LHC05
PDF4LHC04
PDF4LHC02

ATLAS Internal

Instrumental

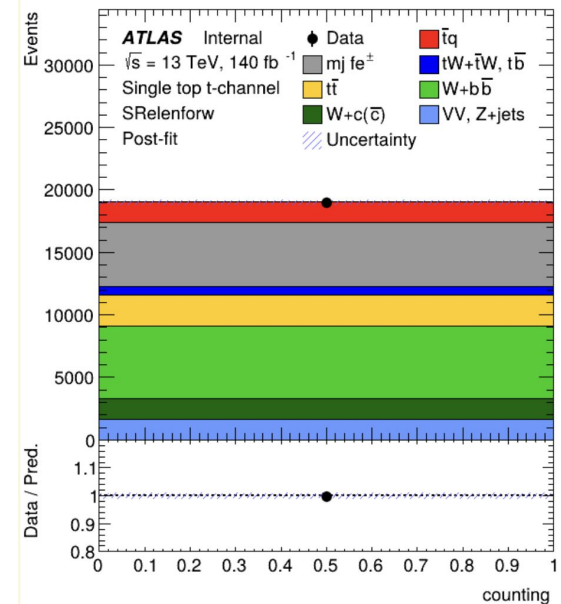
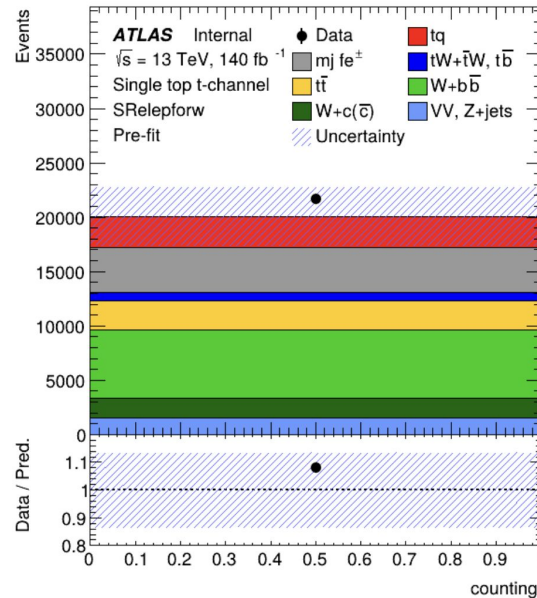
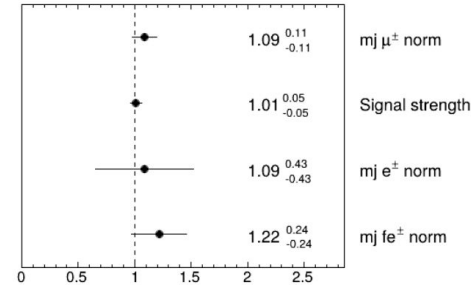


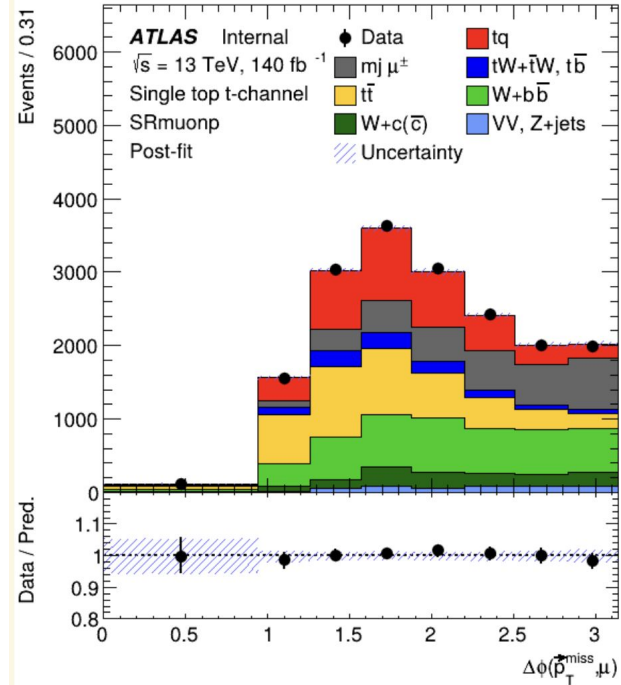
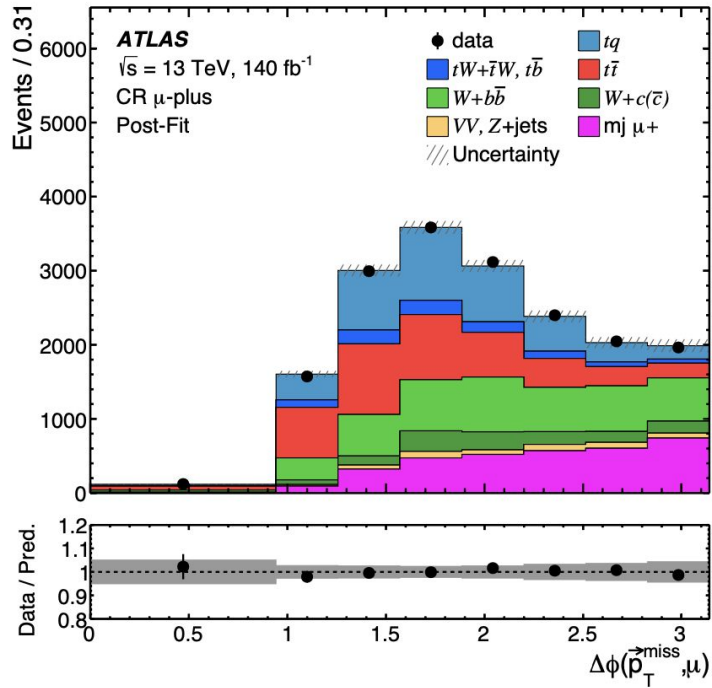
MUON_SCALE
MUON_MS
MUON_ID
MET_SoftTrk_Scale
MET_SoftTrk_ResoPerp
MET_SoftTrk_ResoPara
JET_SingleParticle_HighPt
JET_PunchThrough_MC16
JET_Pileup_RhoTopology
JET_Pileup_PtTerm
JET_Pileup_OffsetNPV
JET_Pileup_OffsetMu
JET_JER_EffectiveNP_9
JET_JER_EffectiveNP_8
JET_JER_EffectiveNP_7
JET_JER_EffectiveNP_6
JET_JER_EffectiveNP_5
JET_JER_EffectiveNP_4
JET_JER_EffectiveNP_3
JET_JER_EffectiveNP_2
JET_JER_EffectiveNP_12restTerm
JET_JER_EffectiveNP_11
JET_JER_EffectiveNP_10
JET_JER_EffectiveNP_1
JET_JER_DataVsMC
JET_Flavor_Response
JET_Flavor_Composition_ttbar
JET_Flavor_Composition_swchan
JET_Flavor_Composition_sitop
JET_EtaIntercalibration_TotalStat
JET_EtaIntercalibration_NonClosure_posEta
JET_EtaIntercalibration_NonClosure_negEta
JET_EtaIntercalibration_Modelling
JET_EffectiveNP_Statistical6
JET_EffectiveNP_Statistical5
JET_EffectiveNP_Statistical4
JET_EffectiveNP_Statistical2
JET_EffectiveNP_Statistical1
JET_EffectiveNP_Modelling4
JET_EffectiveNP_Modelling3
JET_EffectiveNP_Modelling2
JET_EffectiveNP_Modelling1
JET_EffectiveNP_Mixed3
JET_EffectiveNP_Mixed2
JET_EffectiveNP_Mixed1
JET_EffectiveNP_Detector1
JET_BJES_Response
EG_SCALE_ALL
EG_RESOLUTION_ALL

Trexfitter: normfactors

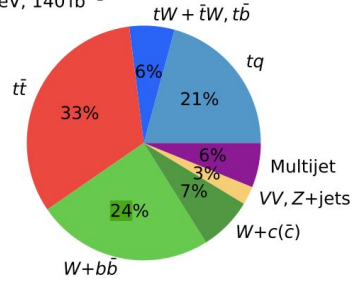
```
"measurements": [  
  {  
    "config": {  
      "parameters": [  
        {  
          "fixed": true,  
          "name": "lumi"  
        },  
        {  
          "fixed": false,  
          "name": "elesrfakes_norm"  
        },  
        {  
          "fixed": false,  
          "name": "eleforwsrfakes_norm"  
        },  
        {  
          "fixed": false,  
          "name": "muonsrfakes_norm"  
        }  
      ],  
      // ... другие параметры  
    },  
    "poi": "negSigXsecOverSM"  
  },  
  {  
    "name": "vb23"  
  }  
]
```

ATLAS Internal

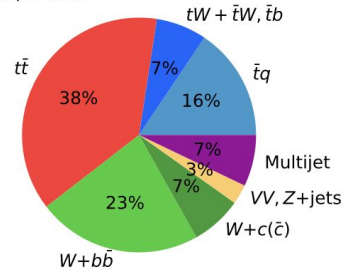




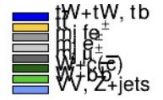
ATLAS Simulation
 $\sqrt{s} = 13 \text{ TeV}$, 140 fb^{-1}
 SR plus



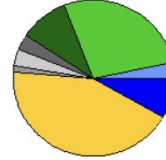
ATLAS Simulation
 $\sqrt{s} = 13 \text{ TeV}$, 140 fb^{-1}
 SR minus



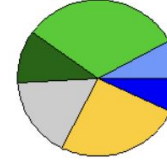
ATLAS Simulation Internal
 $= 13 \text{ TeV}$
 gle top t-channel



SRp



SRelep



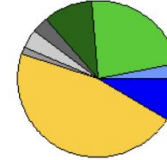
SRelepforw



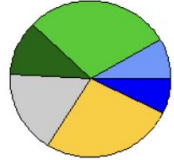
SRmuonp



SRn



SRelen



SRelenforw



SRmuonn



.....

Thank you!