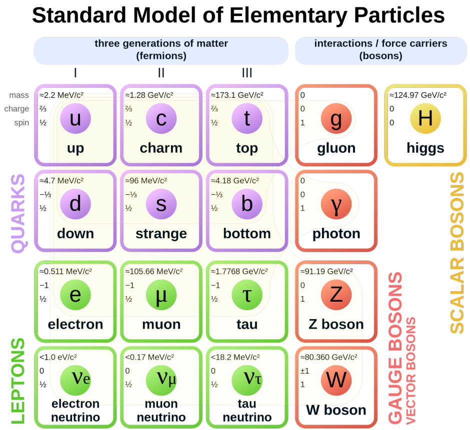
Measurement of the production cross section for the single *top* quark in association with a Z boson at the ATLAS detector

Tesi di Laurea Magistrale di:
Alberto Plebani

Università degli Studi di Milano

Relatrice: **Professoressa Lidia Dell'Asta**
Correlatore: **Professore Attilio Andreazza**

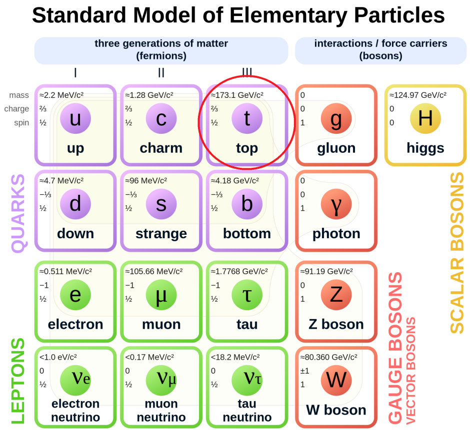
The Standard Model and the *top* quark



Standard Model (SM)

- Gauge theory describing 12 elementary particles, their interactions mediated by 4 vector bosons, and a scalar (Higgs boson)

The Standard Model and the *top* quark



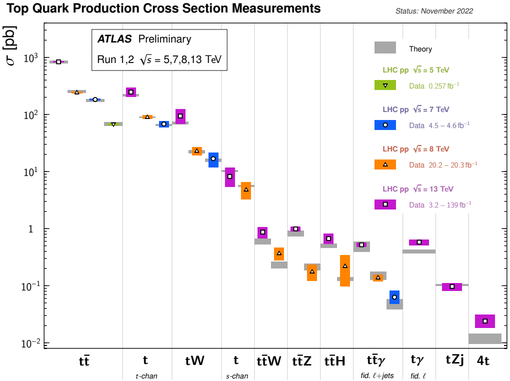
Standard Model (SM)

- Gauge theory describing 12 elementary particles, their interactions mediated by 4 vector bosons, and a scalar (Higgs boson)

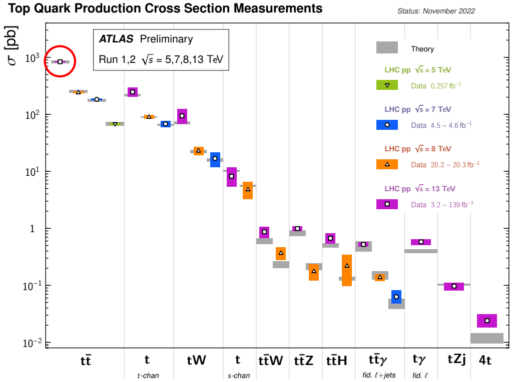
Top quark

- Heaviest SM particle (172.67 ± 0.30 GeV) $\Rightarrow$ Almost unitary coupling with Higgs boson
- Decay: $t \to Wb$ (BR=100%) with $\tau_t = 5 \times 10^{-25}$s $\Rightarrow$ No bound states

Top quark production at the LHC



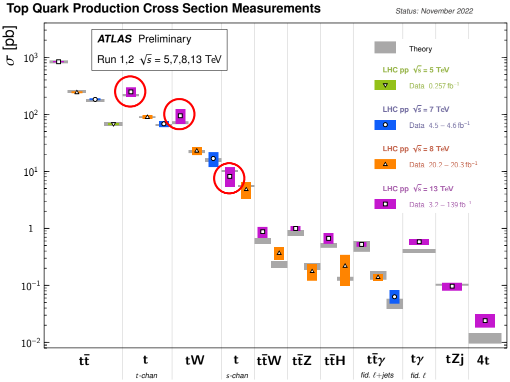
Top quark production at the LHC



Production

- Strong $t\bar{t}$ production: $\sigma_{t\bar{t}} \simeq 830\text{pb}$ at $\sqrt{s} = 13$ TeV

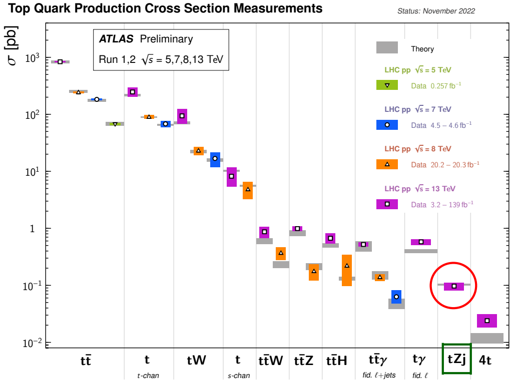
Top quark production at the LHC



Production

- Strong $t\bar{t}$ production: $\sigma_{t\bar{t}} \simeq 830\text{pb}$ at $\sqrt{s} = 13$ TeV
- Electroweak single *top* production:
  - t-channel, tW and s-channel

Top quark production at the LHC



Production

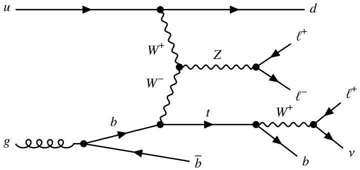
- Strong $t\bar{t}$ production: $\sigma_{t\bar{t}} \simeq 830$pb at $\sqrt{s} = 13$ TeV
- Electroweak single *top* production:
  - t-channel, tW and s-channel
- Associated productions: rarer processes, first observed at LHC thanks to unprecedent amount of data collected
  - $t\bar{t}W$, $t\bar{t}Z$ and $t\bar{t}H$
  - tZq

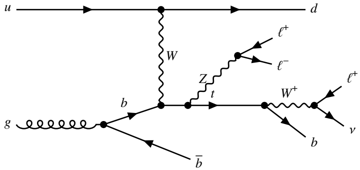
State of the art

- tZq: t-channel single *top* production in association with a Z boson

State of the art

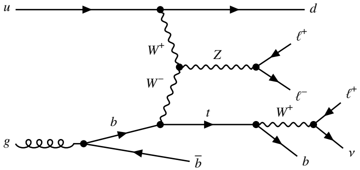
- tZq: t-channel single *top* production in association with a Z boson

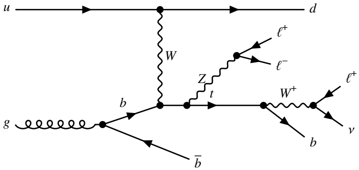




State of the art

- tZq: t-channel single *top* production in association with a Z boson

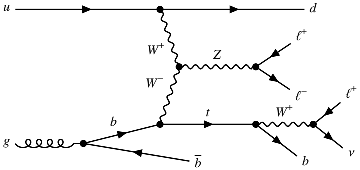


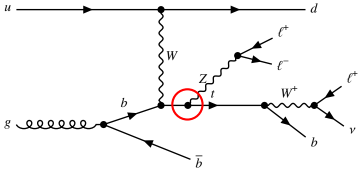


- Predicted cross section: $\sigma_{tZq} = 101.7^{+5.3}_{-1.3}(\text{scale}) \pm 1.1(\text{PDF})$ fb at $\sqrt{s} = 13$ TeV

State of the art

- tZq: t-channel single *top* production in association with a Z boson

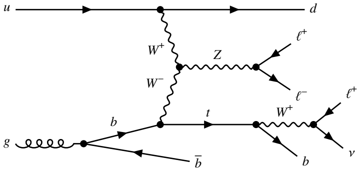


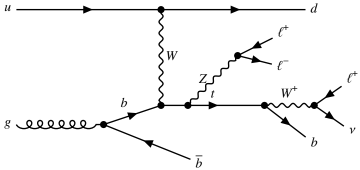


- Predicted cross section: $\sigma_{tZq} = 101.7^{+5.3}_{-1.3}(\text{scale}) \pm 1.1(\text{PDF})$ fb at $\sqrt{s} = 13$ TeV
- Sensitive to t-Z coupling, background to tHq (t-H Yukawa coupling)

State of the art

- tZq: t-channel single *top* production in association with a Z boson

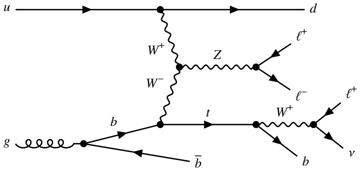


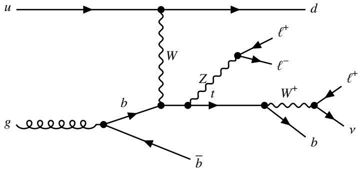


- Predicted cross section: $\sigma_{tZq} = 101.7^{+5.3}_{-1.3}(\text{scale}) \pm 1.1(\text{PDF})$ fb at $\sqrt{s} = 13$ TeV
- Sensitive to t-Z coupling, background to tHq (t-H Yukawa coupling)
- Observed by both ATLAS and CMS in 2019 for the first time

State of the art

- tZq: t-channel single *top* production in association with a Z boson





- Predicted cross section: $\sigma_{tZq} = 101.7^{+5.3}_{-1.3}(\text{scale}) \pm 1.1(\text{PDF})$ fb at $\sqrt{s} = 13$ TeV
- Sensitive to t-Z coupling, background to tHq (t-H Yukawa coupling)
- Observed by both ATLAS and CMS in 2019 for the first time
- ATLAS measurement (139 fb^{-1}, $\sqrt{s} = 13$ TeV):
  $\sigma_{tZq} = 97^{+14}_{-13}$ fb $= 97^{+13}_{-12}(\text{stat}) \pm 7(\text{syst})$ fb

Object of my analysis

Analysis strategy and goals

- Previous analysis measured the cross section with a 15% uncertainty

Object of my analysis

Analysis strategy and goals

- Previous analysis measured the cross section with a 15% uncertainty
- The aim of my work is to improve the precision of this measurement, using the same dataset, by:

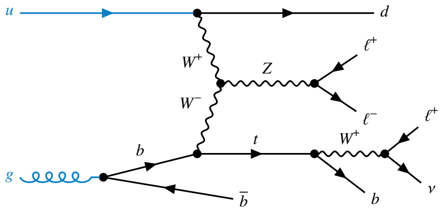
Object of my analysis

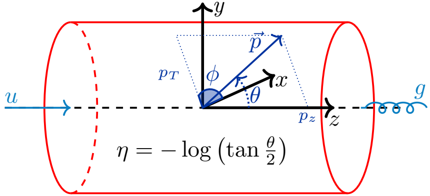
Analysis strategy and goals

- Previous analysis measured the cross section with a 15% uncertainty
- The aim of my work is to improve the precision of this measurement, using the same dataset, by:
  - Increasing signal acceptance by improving event selection
  - Better controlling of the backgrounds:
    - Increase separation between signal and backgrounds
    - Better modelling of specific sources

Trilepton final state (BR=1.76%)

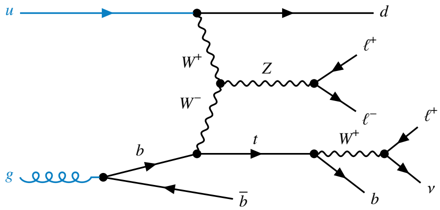
- Exactly 3 leptons (e or μ):
  - At least one $\ell^+\ell^-$ pair with $|m_{\ell\ell} - m_Z| < 10$ GeV
  - Hadronic τ leptons not included

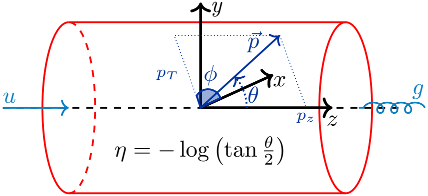




Trilepton final state (BR=1.76%)

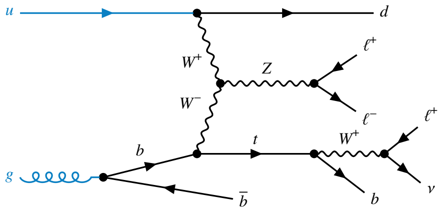
- Exactly 3 leptons (e or μ):
  - At least one $\ell^+\ell^-$ pair with $|m_{\ell\ell} - m_Z| < 10$ GeV
  - Hadronic τ leptons not included
- 2 to 5 hadronic jets:
  - Exactly 1 light forward jet
  - 1 to 2 b-jets

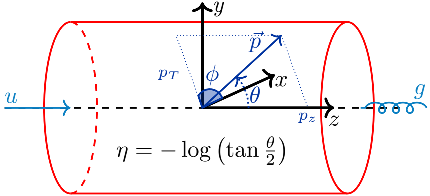




Trilepton final state (BR=1.76%)

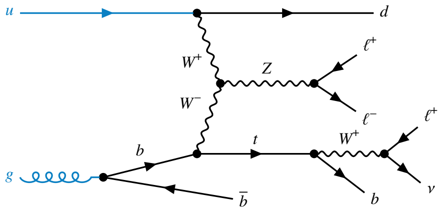
- Exactly 3 leptons (e or μ):
  - At least one $\ell^+\ell^-$ pair with $|m_{\ell\ell} - m_Z| < 10$ GeV
  - Hadronic τ leptons not included
- 2 to 5 hadronic jets:
  - Exactly 1 light forward jet
  - 1 to 2 b-jets
- $p_T(\ell_i) > 27,\ 15,\ 10$ GeV
- $p_T(\text{jets}) > 25$ GeV
- b-jets efficiency: 85%

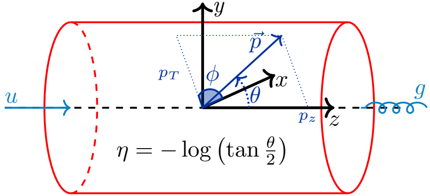




Trilepton final state (BR=1.76%)

- Exactly 3 leptons (e or μ):
  - At least one $\ell^+\ell^-$ pair with $|m_{\ell\ell} - m_Z| < 10$ GeV
  - Hadronic τ leptons not included
- 2 to 5 hadronic jets:
  - Exactly 1 light forward jet
  - 1 to 2 b-jets
- $p_T(\ell_i) > 27,\ 15,\ 10$ GeV
- $p_T(\text{jets}) > 25$ GeV
- b-jets efficiency: 85%





Not presented

- Multiple event selections tested, changing lepton p_T thresholds, b-jets efficiency and jet multiplicity

Main background sources

$t\bar{t}Z$

Diboson plus jets

Fakes

Main background sources

$t\bar{t}Z$

- Irreducible background (same final state)
- Trilepton final state has 4 jets, of which 2 b-jets
- Higher jet multiplicity $\Rightarrow$ higher contribution

Diboson plus jets

Fakes

Main background sources

$t\bar{t}Z$

- Irreducible background (same final state)
- Trilepton final state has 4 jets, of which 2 b-jets
- Higher jet multiplicity $\Rightarrow$ higher contribution

Diboson plus jets

- Mainly WZ (irr.) but also WW, ZZ (red.)
- Dominant background in 1 b-jet events
- Higher b-tagging eff. $\Rightarrow$ higher contribution

Fakes

Main background sources

$t\bar{t}Z$

- Irreducible background (same final state)
- Trilepton final state has 4 jets, of which 2 b-jets
- Higher jet multiplicity $\Rightarrow$ higher contribution

Diboson plus jets

- Mainly WZ (irr.) but also WW, ZZ (red.)
- Dominant background in 1 b-jet events
- Higher b-tagging eff. $\Rightarrow$ higher contribution

Fakes

- $t\bar{t}$ and Z+jets (red.)
- Two prompt leptons,the third is a non-prompt (fake) lepton
- Lower $p_T(\ell_3)$ thresholds $\Rightarrow$ higher contribution

Main background sources

$t\bar{t}Z$

- Irreducible background (same final state)
- Trilepton final state has 4 jets, of which 2 b-jets
- Higher jet multiplicity $\Rightarrow$ higher contribution

Diboson plus jets

- Mainly WZ (irr.) but also WW, ZZ (red.)
- Dominant background in 1 b-jet events
- Higher b-tagging eff. $\Rightarrow$ higher contribution

Fakes

- $t\bar{t}$ and Z+jets (red.)
- Two prompt leptons,the third is a non-prompt (fake) lepton
- Lower $p_T(\ell_3)$ thresholds $\Rightarrow$ higher contribution

- Other small backgrounds: tWZ, $t\bar{t}H$, $t\bar{t}W$, 3 and 4 tops production

Main background sources

$t\bar{t}Z$

- Irreducible background (same final state)
- Trilepton final state has 4 jets, of which 2 b-jets
- Higher jet multiplicity $\Rightarrow$ higher contribution

Diboson plus jets

- Mainly WZ (irr.) but also WW, ZZ (red.)
- Dominant background in 1 b-jet events
- Higher b-tagging eff. $\Rightarrow$ higher contribution

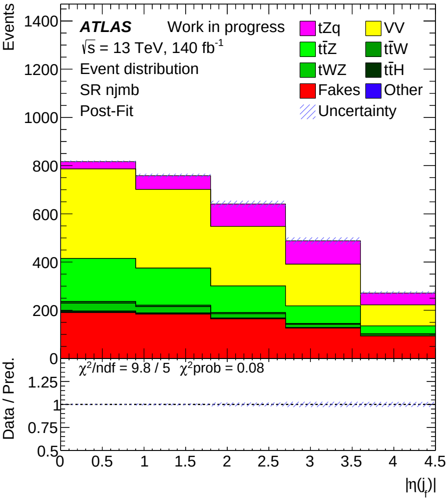
Fakes

- $t\bar{t}$ and Z+jets (red.)
- Two prompt leptons,the third is a non-prompt (fake) lepton
- Lower $p_T(\ell_3)$ thresholds $\Rightarrow$ higher contribution

- Other small backgrounds: tWZ, $t\bar{t}H$, $t\bar{t}W$, 3 and 4 tops production
- Our selection is loose on lepton p_T thresholds and b-jet identification efficiency, and also includes high jet multiplicity

Event yields

	Event yields
tZq	326.3 ± 11.2
$t\bar{t}Z$	549.5 ± 26.3
Diboson	1205.0 ± 39.6
Fakes	757.4 ± 27.0
$t\bar{t}W$	21.9 ± 0.8
tWZ	97.1 ± 15.3
$t\bar{t}H$	17.5 ± 0.6
Other	0.3 ± 0.0
Total	2974.9 ± 100.5



Multi Variate Analysis (MVA)

Neural Network training

- Need for a variable which best separates signal and backgrounds
- Training: feed input variables to a NN, which train itself at classifying separately signal and backgrounds

Multi Variate Analysis (MVA)

Neural Network training

- Need for a variable which best separates signal and backgrounds
- Training: feed input variables to a NN, which train itself at classifying separately signal and backgrounds

Binary NN

- Signal trained versus all backgrounds together
- 1 output, with $O \in [0, 1]$

Multi Variate Analysis (MVA)

Neural Network training

- Need for a variable which best separates signal and backgrounds
- Training: feed input variables to a NN, which train itself at classifying separately signal and backgrounds

Binary NN

- Signal trained versus all backgrounds together
- 1 output, with $O \in [0, 1]$

Not presented

- Tested also a comparison with a Multi Class NN, which proved to be worse

Input variables

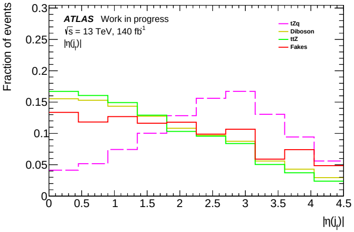
Variable selection:

- 16 input variables, chosen such that different sources had different shapes

Input variables

Variable selection:

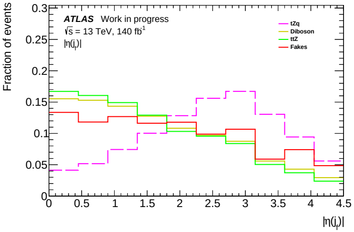
- 16 input variables, chosen such that different sources had different shapes
- Some examples in the plots:
  - $|\eta(j_f)|$: only signal has a forward jet ($\eta = -\log(\tan\theta/2)$)

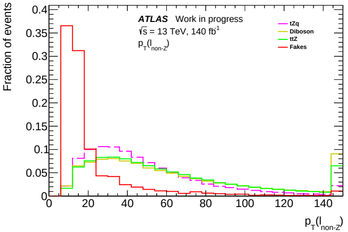


Input variables

Variable selection:

- 16 input variables, chosen such that different sources had different shapes
- Some examples in the plots:
  - $|\eta(j_f)|$: only signal has a forward jet ($\eta = -\log(\tan\theta/2)$)
  - $p_T(\ell_{\text{non-Z}})$: Fakes' third lepton has low p_T

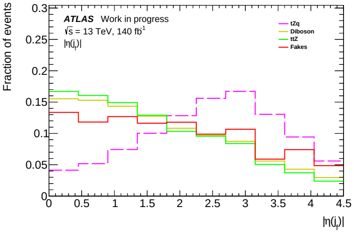


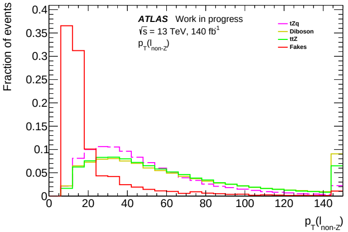


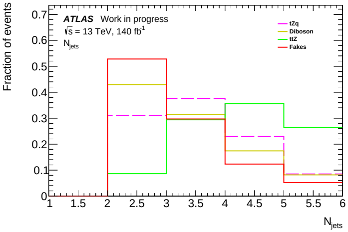
Input variables

Variable selection:

- 16 input variables, chosen such that different sources had different shapes
- Some examples in the plots:
  - $|\eta(j_f)|$: only signal has a forward jet ($\eta = -\log(\tan\theta/2)$)
  - $p_T(\ell_{\text{non-Z}})$: Fakes' third lepton has low p_T
  - N_{jets}: $t\bar{t}Z$ displays higher jet multiplicity





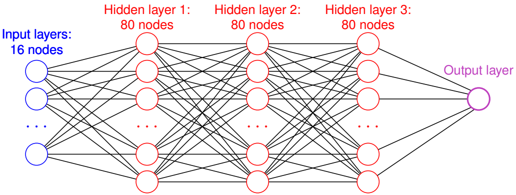


Binary NN: setup and hyperparameters

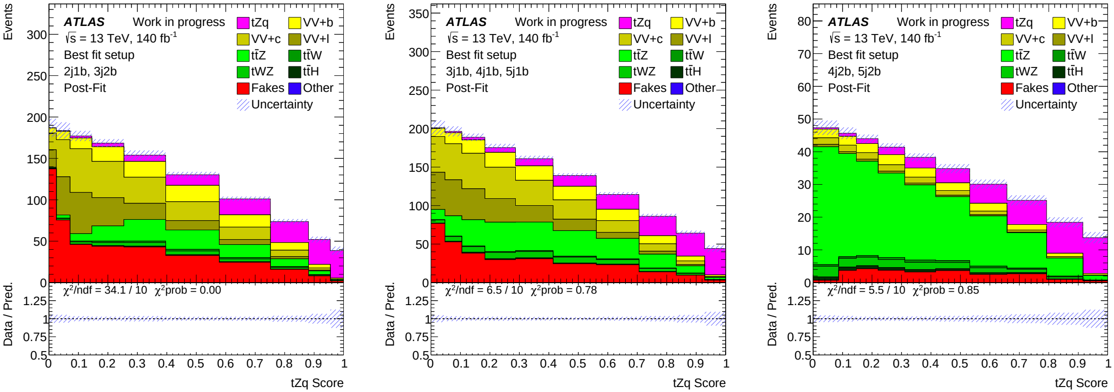
- Hyperparameters optimised using `Grid search` algorithm, tested 3600 configurations
  - Choice based on separation, AUC and loss function
- Used k-fold cross validation with $k = 4$

Hyperparameters

- 16 input variables
- 3 Hidden layers
- Nodes in each layer: 80,80,80
- Activation functions: *eLu*, *eLu*, *eLu*
- Learning rate: 0.001
- Dropout rate: 0.2



Binary training: splitting similar jet multiplicity events



NN output distributions in similar jet multiplicity events

Binned profile Likelihood fit

- Build a Likelihood which is the product of Poisson and Gaussian distributions

Binned profile Likelihood fit

- Build a Likelihood which is the product of Poisson and Gaussian distributions
  - Poisson: signal plus background events in each bin
  - Gaussian: Nuisance Parameters (NP, θ_j) distribution, referring to systematics uncertainties

Binned profile Likelihood fit

- Build a Likelihood which is the product of Poisson and Gaussian distributions
  - Poisson: signal plus background events in each bin
  - Gaussian: Nuisance Parameters (NP, θ_j) distribution, referring to systematics uncertainties

$$\mathcal{L}(\vec{n}|\vec{\theta},\vec{\mu}) = \prod_{i=1}^{N^{bins}} P\left(n_i|S_i(\vec{\theta},\vec{\mu}) + B_i(\vec{\theta},\vec{\mu})\right) \times \prod_{j=1}^{N^{NP}} G\left(\theta_j\right)$$

Binned profile Likelihood fit

- Build a Likelihood which is the product of Poisson and Gaussian distributions
  - Poisson: signal plus background events in each bin
  - Gaussian: Nuisance Parameters (NP, θ_j) distribution, referring to systematics uncertainties

$$\mathcal{L}(\vec{n}|\vec{\theta},\vec{\mu}) = \prod_{i=1}^{N^{bins}} P\left(n_i|S_i(\vec{\theta},\vec{\mu}) + B_i(\vec{\theta},\vec{\mu})\right) \times \prod_{j=1}^{N^{NP}} G\left(\theta_j\right)$$

- Cross section extracted by a *binned profile Likelihood* fit, in which the Parameter of Interest (POI) is the signal strength $\mu(tZq) = \frac{\sigma^{\text{measured}}}{\sigma^{\text{theory}}}$

Binned profile Likelihood fit

- Build a Likelihood which is the product of Poisson and Gaussian distributions
  - Poisson: signal plus background events in each bin
  - Gaussian: Nuisance Parameters (NP, θ_j) distribution, referring to systematics uncertainties

$$\mathcal{L}(\vec{n}|\vec{\theta},\vec{\mu}) = \prod_{i=1}^{N^{bins}} P\left(n_i|S_i(\vec{\theta},\vec{\mu}) + B_i(\vec{\theta},\vec{\mu})\right) \times \prod_{j=1}^{N^{NP}} G\left(\theta_j\right)$$

- Cross section extracted by a *binned profile Likelihood* fit, in which the Parameter of Interest (POI) is the signal strength $\mu(tZq) = \frac{\sigma^{\text{measured}}}{\sigma^{\text{theory}}}$
- Best values of the POI is the Maximum Likelihood Estimator (MLE), which is the parameter for which the Negative-Log-Likelihood reaches a minima

Binned profile Likelihood fit

- Build a Likelihood which is the product of Poisson and Gaussian distributions
  - Poisson: signal plus background events in each bin
  - Gaussian: Nuisance Parameters (NP, θ_j) distribution, referring to systematics uncertainties

$$\mathcal{L}(\vec{n}|\vec{\theta},\vec{\mu}) = \prod_{i=1}^{N^{bins}} P\left(n_i|S_i(\vec{\theta},\vec{\mu}) + B_i(\vec{\theta},\vec{\mu})\right) \times \prod_{j=1}^{N^{NP}} G\left(\theta_j\right)$$

- Cross section extracted by a *binned profile Likelihood* fit, in which the Parameter of Interest (POI) is the signal strength $\mu(tZq) = \frac{\sigma^{\text{measured}}}{\sigma^{\text{theory}}}$
- Best values of the POI is the Maximum Likelihood Estimator (MLE), which is the parameter for which the Negative-Log-Likelihood reaches a minima
- Asimov fit $\Rightarrow$ data substituted by Monte Carlo $\Rightarrow$ POI MLE is 1

Binned profile Likelihood fit

- Build a Likelihood which is the product of Poisson and Gaussian distributions
  - Poisson: signal plus background events in each bin
  - Gaussian: Nuisance Parameters (NP, θ_j) distribution, referring to systematics uncertainties

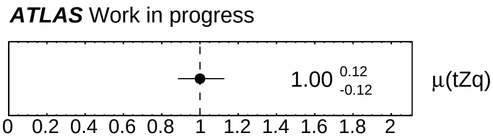
$$\mathcal{L}(\vec{n}|\vec{\theta},\vec{\mu}) = \prod_{i=1}^{N^{bins}} P\left(n_i|S_i(\vec{\theta},\vec{\mu}) + B_i(\vec{\theta},\vec{\mu})\right) \times \prod_{j=1}^{N^{NP}} G\left(\theta_j\right)$$

- Cross section extracted by a *binned profile Likelihood* fit, in which the Parameter of Interest (POI) is the signal strength $\mu(tZq) = \frac{\sigma^{\text{measured}}}{\sigma^{\text{theory}}}$
- Best values of the POI is the Maximum Likelihood Estimator (MLE), which is the parameter for which the Negative-Log-Likelihood reaches a minima
- Asimov fit $\Rightarrow$ data substituted by Monte Carlo $\Rightarrow$ POI MLE is 1

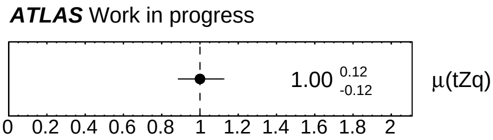
Not presented

- Data-driven estimation of Diboson and Fakes to account for data/MC discrepancies

Results


ATLAS Work in progress
1.00 0.12 -0.12
μ(tZq)
0 0.2 0.4 0.6 0.8 1 1.2 1.4 1.6 1.8 2

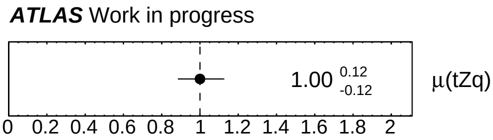
Results



Results

- MLE for the POI: $\mu(tZq) = 1.00 \ \pm 0.12 = 1.00 \ ^{+0.10}_{-0.09}(\text{stat}) \ ^{+0.08}_{-0.07}(\text{syst})$

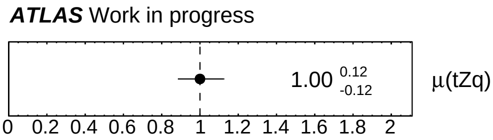
Results



Results

- MLE for the POI: $\mu(tZq) = 1.00\ \pm 0.12 = 1.00\ ^{+0.10}_{-0.09}$(stat) $^{+0.08}_{-0.07}$(syst)
  - Measurement is stat-dominated, just like previous analysis
  - Theoretical uncertainties dominate the systematics (6%, instrumental only 4%)

Results



Results

- MLE for the POI: $\mu(tZq) = 1.00 \pm 0.12 = 1.00\ ^{+0.10}_{-0.09}(\text{stat})\ ^{+0.08}_{-0.07}(\text{syst})$
  - Measurement is stat-dominated, just like previous analysis
  - Theoretical uncertainties dominate the systematics (6%, instrumental only 4%)
- 20% expected improvement with respect to the previous analysis (15%), using same dataset

Ratio of *top* and *antitop* cross sections in tZq production

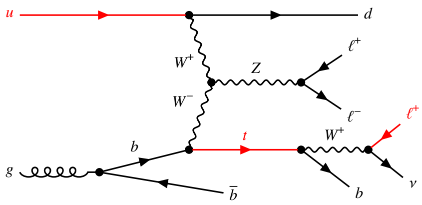
Motivation

- tZq process can occur with two different initial state quarks:

Ratio of *top* and *antitop* cross sections in tZq production

Motivation

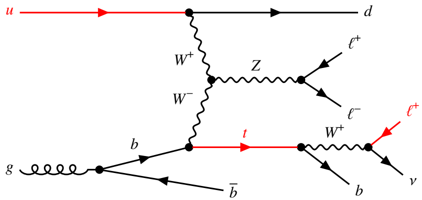
- tZq process can occur with two different initial state quarks:
  - "up" type (u, c) $\Rightarrow$ *top* (t) $\Rightarrow$ ℓ^+

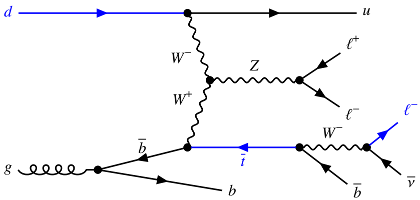


Ratio of *top* and *antitop* cross sections in tZq production

Motivation

- tZq process can occur with two different initial state quarks:
  - "up" type (u, c) $\Rightarrow$ *top* (t) $\Rightarrow$ ℓ^+
  - "down" type (d, s) $\Rightarrow$ *antitop* ($\bar{t}$) $\Rightarrow$ ℓ^-

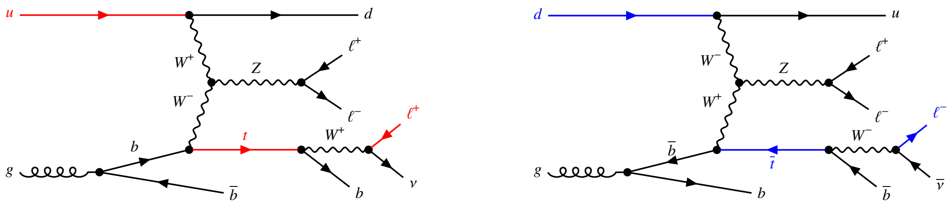




Ratio of *top* and *antitop* cross sections in tZq production

Motivation

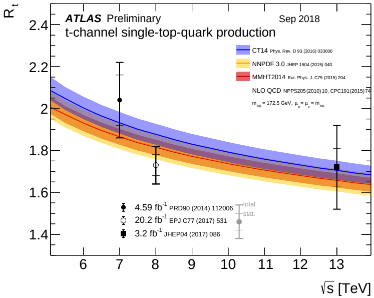
- tZq process can occur with two different initial state quarks:
  - "up" type (u, c) ⇒ *top* (t) ⇒ ℓ^+
  - "down" type (d, s) ⇒ *antitop* ($\bar{t}$) ⇒ ℓ^-
- In proton, *up* more present than *down* ⇒ expected asymmetry, with the size of the asymmetry depending on proton internal structure

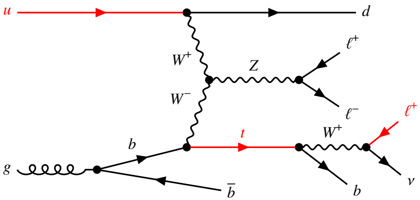


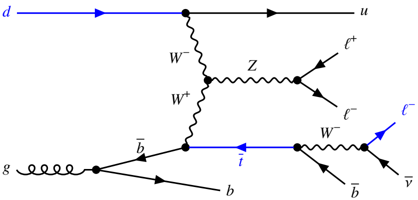
Ratio of *top* and *antitop* cross sections in tZq production

Motivation

- tZq process can occur with two different initial state quarks:
  - "up" type (u, c) $\Rightarrow$ *top* (t) $\Rightarrow$ ℓ^+
  - "down" type (d, s) $\Rightarrow$ *antitop* ($\bar{t}$) $\Rightarrow$ ℓ^-
- In proton, *up* more present than *down* $\Rightarrow$ expected asymmetry, with the size of the asymmetry depending on proton internal structure
- Asymmetry measured in other channels but never in tZq







Ratio of *top* and *antitop* cross sections in tZq production - results

MLE for the POIs

- $\mu(tZq, t) = 1.00\ ^{+0.14}_{-0.13} = 1.00\ \pm 0.11(\text{stat})\ \pm 0.07(\text{syst})$
- $\mu(tZq, \bar{t}) = 1.00\ ^{+0.20}_{-0.18} = 1.00\ ^{+0.17}_{-0.16}(\text{stat})\ \pm 0.09(\text{syst})$
- $R(t/\bar{t}) = 1.00\ ^{+0.23}_{-0.18} = 1.00\ ^{+0.22}_{-0.17}(\text{stat})\ \pm 0.06(\text{syst})$

Ratio of *top* and *antitop* cross sections in tZq production - results

MLE for the POIs

- $\mu(tZq, t) = 1.00\ ^{+0.14}_{-0.13} = 1.00\ \pm 0.11(\text{stat})\ \pm 0.07(\text{syst})$
- $\mu(tZq, \bar{t}) = 1.00\ ^{+0.20}_{-0.18} = 1.00\ ^{+0.17}_{-0.16}(\text{stat})\ \pm 0.09(\text{syst})$
- $R(t/\bar{t}) = 1.00\ ^{+0.23}_{-0.18} = 1.00\ ^{+0.22}_{-0.17}(\text{stat})\ \pm 0.06(\text{syst})$

Results

- These three measurements are the most precise in the tZq channel (uncertainties lower than CMS)

Summary

- This thesis aimed at improving ATLAS tZq cross section measurement by:

Summary

- This thesis aimed at improving ATLAS tZq cross section measurement by:
    - Increasing signal acceptance by improving event selection
    - Better controlling of the backgrounds, increasing separation and improving modelling

Summary

- This thesis aimed at improving ATLAS tZq cross section measurement by:
  - Increasing signal acceptance by improving event selection
  - Better controlling of the backgrounds, increasing separation and improving modelling
- Cross section extracted by a *binned profile Likelihood* fit setting the signal strength as POI

Summary

- This thesis aimed at improving ATLAS tZq cross section measurement by:
  - Increasing signal acceptance by improving event selection
  - Better controlling of the backgrounds, increasing separation and improving modelling
- Cross section extracted by a *binned profile Likelihood* fit setting the signal strength as POI
- Obtained expected $\pm$ 12% uncertainty on inclusive cross section $\Rightarrow$ 20% improvement, using same dataset

Summary

- This thesis aimed at improving ATLAS tZq cross section measurement by:
  - Increasing signal acceptance by improving event selection
  - Better controlling of the backgrounds, increasing separation and improving modelling
- Cross section extracted by a *binned profile Likelihood* fit setting the signal strength as POI
- Obtained expected $\pm$ 12% uncertainty on inclusive cross section $\Rightarrow$ 20% improvement, using same dataset
  - Expected to be ATLAS most precise tZq cross section measurement with data

Summary

- This thesis aimed at improving ATLAS tZq cross section measurement by:
  - Increasing signal acceptance by improving event selection
  - Better controlling of the backgrounds, increasing separation and improving modelling
- Cross section extracted by a *binned profile Likelihood* fit setting the signal strength as POI
- Obtained expected $\pm$ 12% uncertainty on inclusive cross section $\Rightarrow$ 20% improvement, using same dataset
  - Expected to be ATLAS most precise tZq cross section measurement with data
- Moreover, presented ATLAS first measurements of *top* and *antitop* separate inclusive cross sections and their ratio in tZq channel:

Summary

- This thesis aimed at improving ATLAS tZq cross section measurement by:
  - Increasing signal acceptance by improving event selection
  - Better controlling of the backgrounds, increasing separation and improving modelling
- Cross section extracted by a *binned profile Likelihood* fit setting the signal strength as POI
- Obtained expected $\pm$ 12% uncertainty on inclusive cross section $\Rightarrow$ 20% improvement, using same dataset
  - Expected to be ATLAS most precise tZq cross section measurement with data
- Moreover, presented ATLAS first measurements of *top* and *antitop* separate inclusive cross sections and their ratio in tZq channel:
  - Expected to be most precise for the tZq process

Summary

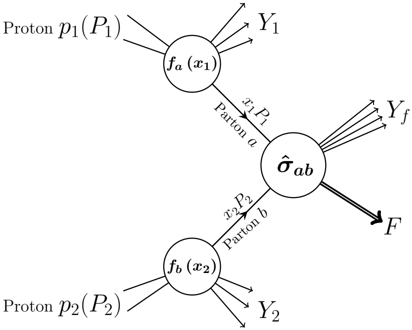
- This thesis aimed at improving ATLAS tZq cross section measurement by:
  - Increasing signal acceptance by improving event selection
  - Better controlling of the backgrounds, increasing separation and improving modelling
- Cross section extracted by a *binned profile Likelihood* fit setting the signal strength as POI
- Obtained expected $\pm$ 12% uncertainty on inclusive cross section $\Rightarrow$ 20% improvement, using same dataset
  - Expected to be ATLAS most precise tZq cross section measurement with data
- Moreover, presented ATLAS first measurements of *top* and *antitop* separate inclusive cross sections and their ratio in tZq channel:
  - Expected to be most precise for the tZq process

Future implications

- This selection and this MVA are currently being used in order to public ATLAS first tZq differential cross section measurement

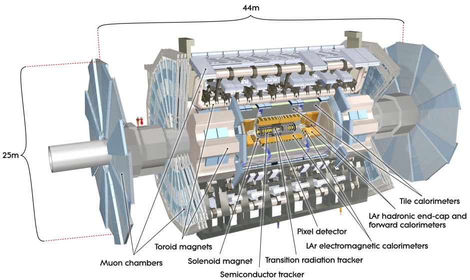
Thank you for your attention
Alberto Plebani

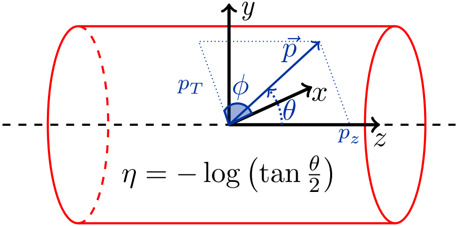
Proton-proton inelastic cross section



$$\sigma(P_1, P_2; Q^2) = \sum_{a,b=q,\bar{q},g} \int_0^1 dx_1 dx_2 f_{a/p_1}\left(x_1, \mu_F^2\right) f_{b/p_2}\left(x_2, \mu_F^2\right) \hat{\sigma}_{ab}\left(x_1 P_1, x_2 P_2; \alpha_S(\mu_F^2), \mu_F^2, \mu_R^2, Q^2\right)$$

The ATLAS experiment





Multipurpose detector

- Higgs discovery, SM precision measurements
- BSM (SUSY, Dark Matter) searches

Systems

- Magnetic: Toroidal and solenoidal
- Inner detector (ID)
- Electromagnetic Calorimeter (ECal)
- Hadronic Calorimeter (HadCal)
- Muon Spectrometer (MS)

Data collected in Run2

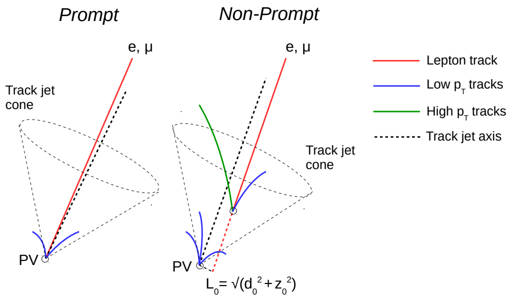
- $140\ \text{fb}^{-1}$ of proton-proton inelastic collisions
- Centre-of-mass energy: $\sqrt{s} = 13$ TeV

Final States Branching Ratios

- 3 charged leptons ($Z \to \ell^+\ell^-$, $W \to \ell\nu_\ell$):
  - Without hadronic τ leptons: 1.8%
  - With hadronic τ leptons: 3.2%
- 2 charged leptons ($Z \to \ell^+\ell^-$, $W \to q\bar{q}'$):
  - Without hadronic τ leptons: 4.6%
  - With hadronic τ leptons: 6.8%
- 1 charged lepton ($Z \to q\bar{q}$, $W \to \ell\nu_\ell$):
  - Without hadronic τ leptons: 22.8%
  - With hadronic τ leptons: 16.9%
- 0 charged leptons ($Z \to q\bar{q}$ or $Z \to \nu_\ell\bar{\nu}_\ell$, $W \to q\bar{q}'$):
  - Inclusive: 60.2%
  - Only Z visible: 44.8%
  - Only Z invisible: 13.4%

Prompt vs non-prompt leptons

- Prompt leptons are the ones generated from W or Z boson decay
- Non-prompt leptons are instead generated inside a hadronic jet
- Therefore, prompt leptons are isolated, whereas non-prompt leptons are not



Comparison between prompt and non-prompt lepton

Event Yields in different jet multiplicity regions

	SR 2j1b	SR 3j1b	SR 4j1b	SR 5j1b	SR 3j2b	SR 4j2b	SR 5j2b
tZq	101.0 ± 12.2	87.5 ± 8.9	42.8 ± 5.5	13.5 ± 2.2	35.1 ± 3.5	32.1 ± 3.9	14.3 ± 2.1
$t\bar{t}Z$	39.6 ± 3.0	84.3 ± 4.9	91.1 ± 3.4	63.8 ± 3.0	73.1 ± 5.6	109.1 ± 5.6	88.8 ± 2.3
Diboson	517.3 ± 7.0	347.5 ± 4.9	185.3 ± 2.7	83.6 ± 1.8	32.0 ± 0.5	24.6 ± 0.5	14.9 ± 0.2
Fakes	400.6 ± 11.9	196.6 ± 6.2	74.3 ± 2.0	31.8 ± 1.1	28.2 ± 0.9	18.6 ± 0.6	7.2 ± 0.2
$t\bar{t}W$	5.2 ± 0.1	3.8 ± 0.1	2.0 ± 0.0	1.1 ± 0.0	5.1 ± 0.1	3.0 ± 0.1	1.6 ± 0.0
tWZ	16.2 ± 1.7	24.7 ± 2.9	19.4 ± 2.8	10.4 ± 1.7	8.0 ± 1.2	10.5 ± 2.3	8.0 ± 2.3
$t\bar{t}H$	1.6 ± 0.0	2.6 ± 0.1	2.6 ± 0.1	1.8 ± 0.0	2.8 ± 0.1	3.3 ± 0.1	2.8 ± 0.1
Other	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.1 ± 0.0	0.0 ± 0.0	0.1 ± 0.0	0.1 ± 0.0
Total	1081.7 ± 23.3	746.9 ± 15.7	417.4 ± 9.4	206.0 ± 5.9	184.3 ± 6.2	201.3 ± 6.4	137.6 ± 3.6

- Other is 3 and 4 *top* quark productions
- $njmb$ notation refers to events with n jets, of which m b-jets

Binary vs Multi Class

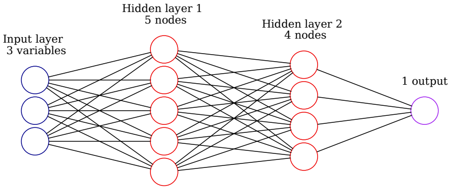
Binary

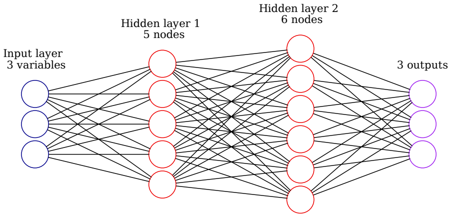
- 2 targets: signal trained versus all remaining backgrounds
- 1 output, with $O \in [0, 1]$
- Greater separation
- Fit made using one region, in which the output was fitted
- Training only similar events improved the fit

Multi Class (MC)

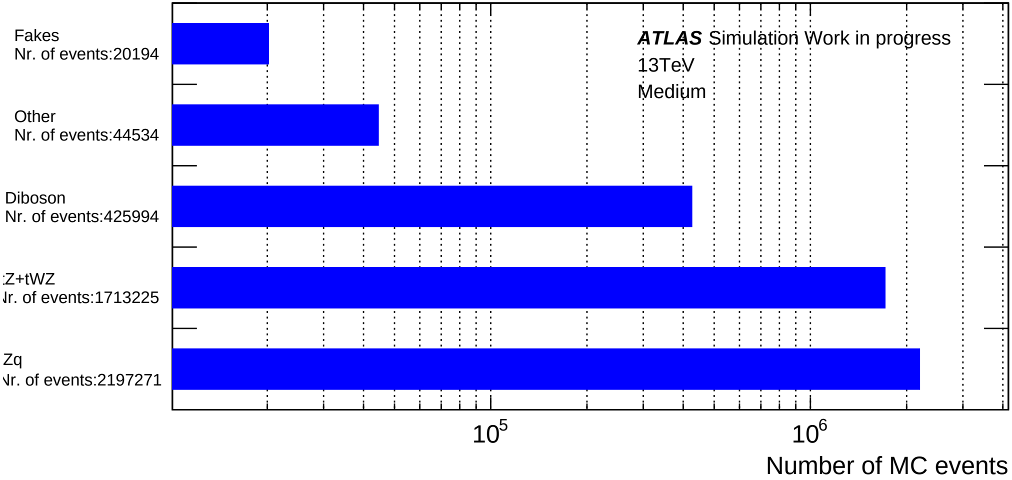
- 3 targets: signal, $t\bar{t}Z$ or Diboson, all remaining backgrounds
- 3 outputs ($O_i \leq 1$), with $\sum O_i = 1$ for each event
- Lower separation
- Useful for the fits: allow to define different regions in order to maximise signal significance, and have background-enhanced regions
- Training only similar events improved the fit

Binary vs Multi Class

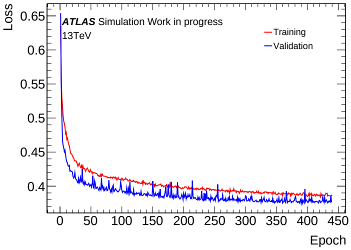


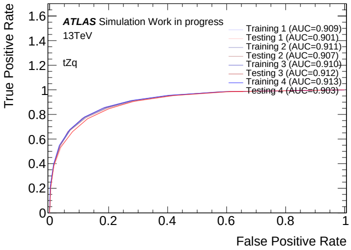


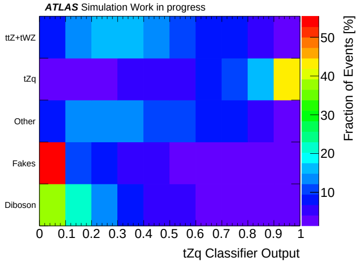
Training dataset size

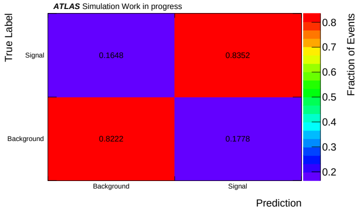


NN specifics







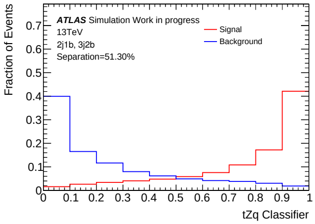


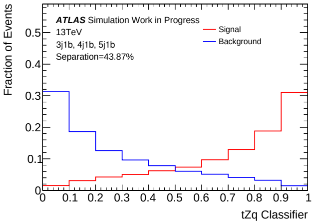
Binary training: splitting similar jet multiplicity events

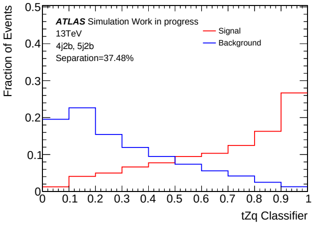
Best training

- Three Binary NNs, each trained on similar jet multiplicity events: $2j1b$ and $3j2b$; $3j1b$, $4j1b$, and $5j1b$; $4j2b$ and $5j2b$
- We can also use radiation jet information in the latter two trainings (not present in $2j1b$ and $3j2b$)

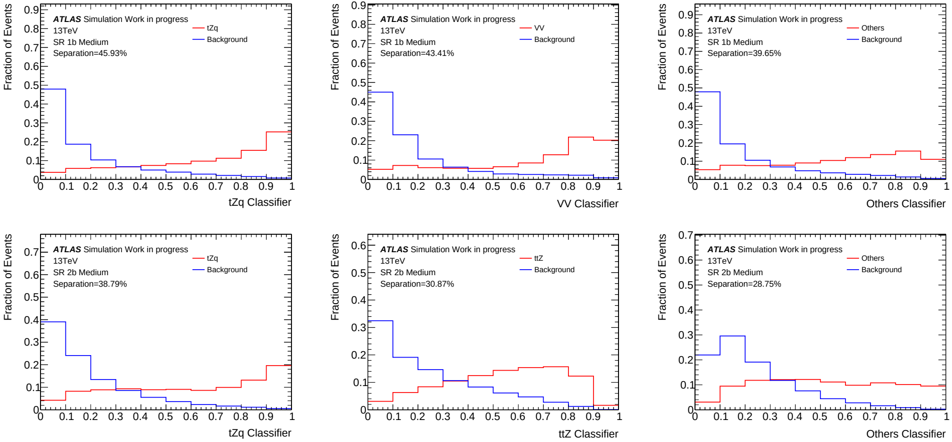
	$2j1b$, $3j2b$	$3j1b$, $4j1b$, $5j1b$	$4j2b$, $5j2b$
tZq	136.2 ± 29.8	143.7 ± 21.2	46.4 ± 10.9
$VV+b$	115.3 ± 4.1	131.7 ± 4.4	21.9 ± 0.8
$VV+c$	242.1 ± 8.4	268.8 ± 8.8	13.4 ± 0.5
$VV+l$	191.9 ± 6.6	215.8 ± 7.4	4.1 ± 0.2
$t\bar{t}Z$	112.7 ± 10.5	239.2 ± 10.6	197.9 ± 10.0
$t\bar{t}W$	10.3 ± 0.4	7.0 ± 0.3	4.6 ± 0.2
tWZ	24.2 ± 3.0	54.4 ± 7.8	18.5 ± 4.9
$t\bar{t}H$	4.4 ± 0.2	7.0 ± 0.2	6.1 ± 0.2
Fakes	428.9 ± 29.1	302.7 ± 18.7	25.7 ± 2.4
Other	0.0 ± 0.0	0.1 ± 0.0	0.2 ± 0.0
Total	1266.0 ± 58.4	1370.3 ± 52.3	338.9 ± 18.2







Multi Class training: splitting different b-jet multiplicity events

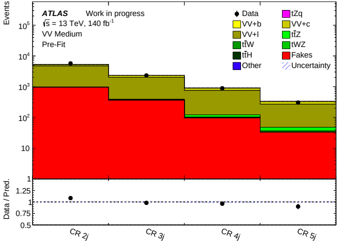


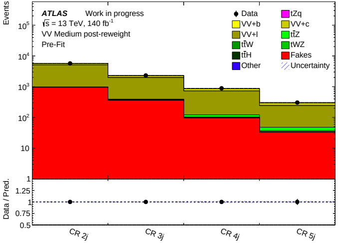
Diboson background reweighting

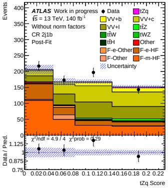
Strategy

- Extract normalisation factors (NF) for Diboson using 0-b-jets events (orthogonal selection wrt event selection)
- One NF for number of jets
- Validation: applying the NF to the Diboson MC simulated events to see if the Data/MC agreement improved

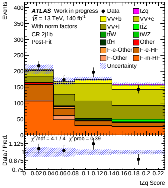
	$2j$	$3j$	$4j$	$5j$
NF	1.10	0.97	0.95	0.88



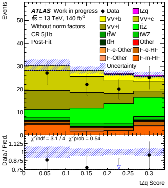




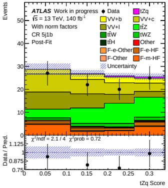
Before applying NFs



After applying NFs



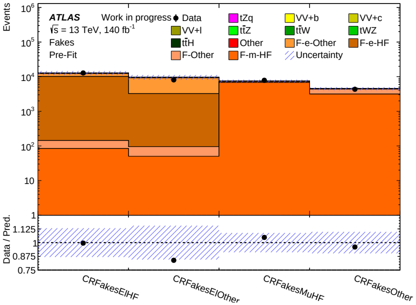
Before applying NFs

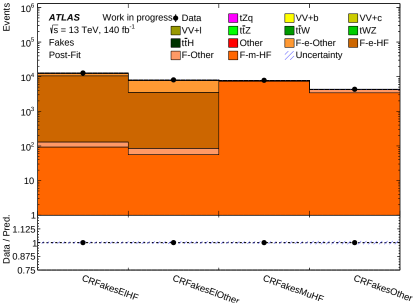


After applying NFs

Fake background reweighting

Regions definition using *Medium* selection				
SR	CR Fakes El HF	CR Fakes El Other	CR Fakes μ HF	CR Fakes Other
3 tight leptons (e or μ)	Exactly 2 tight leptons (e or μ) Loose non tight lepton is electron		Loose non tight lepton is muon	
2-5 jets 1-2 b-jets ($2j2b$ excluded)	2-5 jets ≥ 1 b-jets	0 b-jets	≥ 1 b-jets	0 b-jets
$\lvert m_{\ell\ell} - m_Z \rvert < 10\,\text{GeV}$ -	$\lvert m_{\ell\ell} - m_Z \rvert > 10\,\text{GeV}$ -	$\lvert m_{\ell\ell} - m_Z \rvert < 10\,\text{GeV}$ $m_T(W) < 20\,\text{GeV}$	$\lvert m_{\ell\ell} - m_Z \rvert > 10\,\text{GeV}$ -	$\lvert m_{\ell\ell} - m_Z \rvert < 10\,\text{GeV}$ $m_T(W) < 20\,\text{GeV}$
-	Fit on number of jets	Fit on $p_T(\ell_3)$	Fit on number of jets	Fit on $p_T(\ell_3)$





List of uncertainties

Theoretical uncertainties

- Parton Distribution Functions (PDF): uncertainty on proton's internal structure
- Scale variation: uncertainty on the QCD coupling constant, depending on renormalisation and factorisation scales
- Showering: uncertainty on the parton shower generated after the collision, sensitive to jet multiplicity
- Initial state radiation: uncertainty on QCD radiation prior to the collision

Instrumental uncertainties

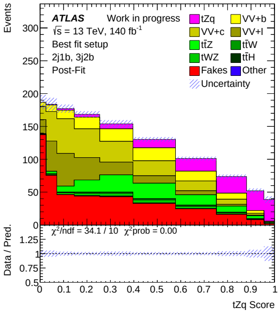
- Luminosity: uncertainty on number of total events
- Pile-up reweighting: uncertainty on additional interactions occurring in the same time window of the main collision
- Jets: uncertainty on hadronic jets identification and reconstruction
- b-tagging: uncertainty on the jets flavour tagging
- Leptons: uncertainties on the identification, isolation and reconstruction of electrons, muons, and photons

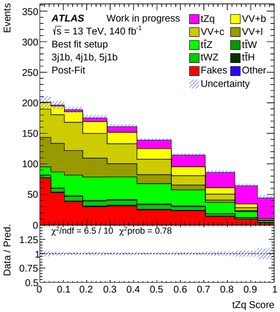
Group	Sub-group	Nuisance Parameters	Group	Sub-group	Nuisance Parameters
Theory (28)	tZq	4 (PDF, PS, ISR, scale)	Instrumental (158)	Leptons	23
	$t\bar{t}Z$	4 (PDF, PS, ISR, scale)		Jets	48
	tWZ	2 (Normalisation, diagram removal)		b-tagging	85
	VV	6 (PDF, scale; 2 for each flavour)		Pile-Up modelling	1
	$t\bar{t}$	12 (PS, Generator, ISR; 3 for each flavour)		Luminosity	1
Total = 186 Nuisance Parameters for systematics uncertainties					

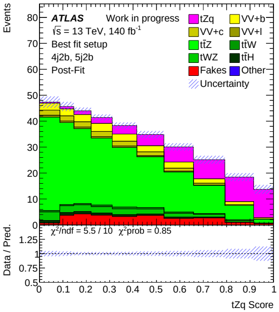
Fitting Binary NN trained on similar jet multiplicity events

Fit setup

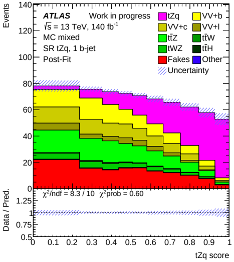
- Asimov fit is blind ⇒ data not used for the fit, and not visible on plots
- Signal Regions (SR) defined by grouping events the same way they were for the training (page 15): $2j1b$ and $3j2b$; $3j1b$, $4j1b$, and $5j1b$; $4j2b$ and $5j2b$
- In each group the fit was made to the output of the Binary NN trained on those events

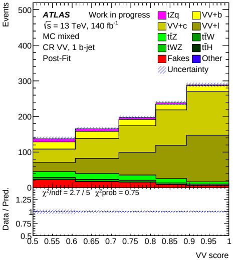


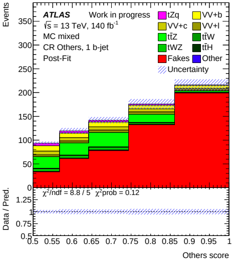


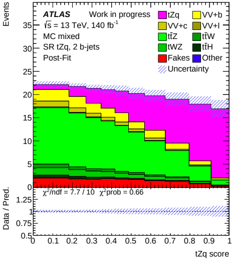


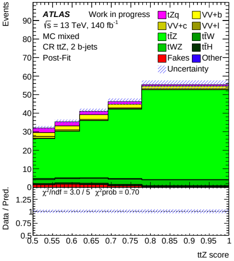
Fit using Multi Class training, splitting different b-jet multiplicity events

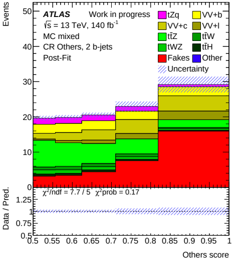












Ratio of *top* and *antitop* cross sections in tZq production - 2

Methods

- As aforementioned, the difference between *top* and *antitop* is the charge of ℓ_W
- We split the dataset based on that charge, and obtained the following values
- We observe the asymmetry in simulated signal events, with a ratio of 1.8
- Diboson asymmetric due to WZ
- $t\bar{t}Z$ and Fakes symmetric because generated by gluon-gluon fusion, and thus they do not depend on the PDFs

	$t,\ Q(\ell_W) = +1$	$\bar{t},\ Q(\ell_W) = -1$	Ratio $t/\bar{t}$
tZq	207.9 ± 20.8	118.4 ± 11.8	1.8
$VV + b$	156.7 ± 3.2	112.2 ± 2.3	1.4
$VV + c$	279.3 ± 5.6	245.0 ± 5.2	1.1
$VV + l$	244.5 ± 5.4	167.3 ± 3.9	1.5
$t\bar{t}Z$	281.5 ± 10.8	268.1 ± 9.4	1.0
$t\bar{t}W$	14.3 ± 0.3	7.6 ± 0.2	2.0
tWZ	48.6 ± 7.4	48.5 ± 7.1	1.0
$t\bar{t}H$	8.8 ± 0.2	8.7 ± 0.2	1.0
Fakes	375.5 ± 10.2	381.8 ± 10.2	1.0
Other	0.2 ± 0.0	0.2 ± 0.0	1.0
Total	1617.4 ± 28.9	1357.8 ± 24.4	1.2

Ratio of *top* and *antitop* cross sections in tZq production - results

MLE for the POIs

- $\mu(tZq,t) = 1.00\ ^{+0.14}_{-0.13} = 1.00\ \pm 0.11(\text{stat})\ \pm 0.07(\text{syst})$
- $\mu(tZq,\bar{t}) = 1.00\ ^{+0.20}_{-0.18} = 1.00\ ^{+0.17}_{-0.16}(\text{stat})\ \pm 0.09(\text{syst})$
- $R(t/\bar{t}) = 1.00\ ^{+0.23}_{-0.18} = 1.00\ ^{+0.22}_{-0.17}(\text{stat})\ \pm 0.06(\text{syst})$

Considerations

- Systematics uncertainties tend to cancel in the ratio $\Rightarrow$ much dominated by statistics
- The three cross sections are dominated by theoretical uncertainties, particularly signal modelling uncertainty
- *Top*, *Antitop* and ratio uncertainties represent the most precise measurements in the tZq channel

Systematic	Inclusive	*Top*	*Antitop*	Ratio
tZq modelling	5.0%	4.8%	5.7%	1.0%
Diboson+$t\bar{t}Z$ modelling	3.3%	3.0%	4.2%	1.7%
Leptons	3.3%	3.2%	3.6%	0.5%
Jets	1.6%	1.6%	2.2%	1.6%
Fakes modelling	1.6%	2.5%	2.9%	4.6%
b-tagging	1.3%	1.4%	1.4%	0.8%
MC statistics	1.0%	1.3%	2.4%	2.7%
Luminosity	0.8%	0.8 %	0.8%	0.1%
Pileup	0.8%	0.9%	0.8%	0.1%
tZq PDF	0.3%	0.4%	0.6%	0.2%
Theory	6.1%	6.1%	7.5%	5.1%
Instrumental	4.1%	4.2%	5.0%	2.9%
Full systematics	7.4%	7.4%	9.0%	5.9%
Statistical	9.5%	11.0%	16.5%	19.5%
Total	12.0%	13.5%	19.0%	20.5%