Event reconstruction and background suppression for a measurement of the $t\bar{t}Z$ cross section at ATLAS

Florian Fischer

Ludwig-Maximilians-Universität München
– LS Schaile –

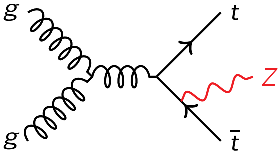
TAE 2019 - COST International Training School on High Energy Physics

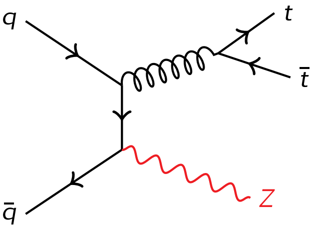
– Benasque, September 12th, 2019 –

Physics Motivation:

- $t\bar{t}Z$ a rare Standard Model (SM) process
  - Measurement of cross section a stringest test of Standard Model
  - Sensitivity to physics beyond Standard Model (BSM)
- $t\bar{t}Z$ an important background for several LHC analyses
  - $t\bar{t}H$, $t\bar{t}t\bar{t}$ + other rare SM processes
  - Several BSM searches including variety of SUSY signatures, flavour-changing neutral-current processes (FCNC), ...

⇒ **Improved knowledge of $t\bar{t}Z$ process can lead to improvements in many analyses**



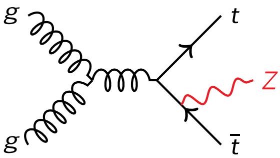


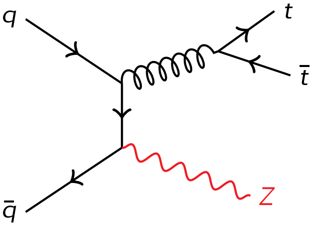
Analysis Strategy:

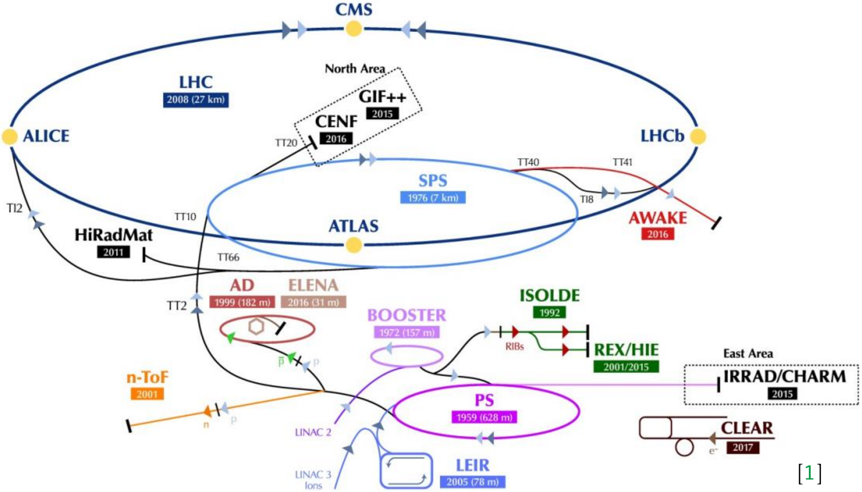
- Inclusive $t\bar{t}Z$ cross section measurement for decay channels with 3 and 4 leptons in final state
  - Successful measurements of inclusive cross section already performed by ATLAS & CMS experiments at centre-of-mass energy $\sqrt{s} = 13\,\text{TeV}$
  - Very good agreement with theoretical prediction (cf. 1609.01599, 1711.02547, 1901.03584)
- Perform unfolded differential cross section measurements as functions of kinematic variables
  - Will allow to further probe the top-Z coupling of the Standard Model
  - Unfolding: extrapolate measurements to fiducial detector volume and full generator phase space

  $\Rightarrow$ **Correct for detector effects as well as signal acceptance & efficiency w.r.t given phase space region**

  $\Rightarrow$ **Directly compare measurements with theory and other experiments**
- Employing the LHC 2015-2018 "Run 2" dataset taken at ATLAS experiment corresponding to $139\,\text{fb}^{-1}$

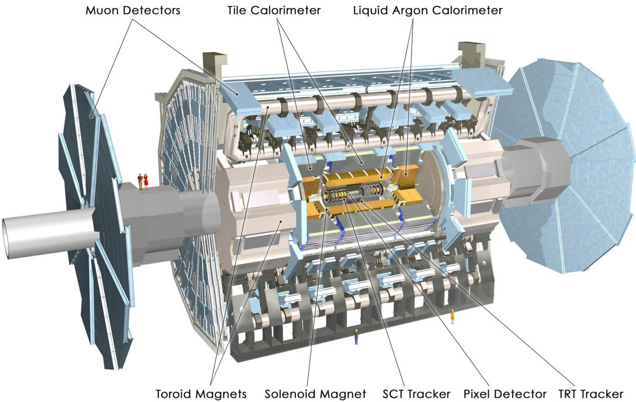





CMS
North Area
LHC
2008 (27 km)
GIF++
2015
CENF
2016
ALICE
LHCb
TT20
TT40
TT41
SPS
1976 (7 km)
TI8
TI2
TT10
AWAKE
2016
ATLAS
HiRadMat
2011
TT66
AD
1999 (182 m)
ELENA
2016 (31 m)
ISOLDE
1992
TT2
BOOSTER
1972 (157 m)
RIBs
REX/HIE
2001/2015
East Area
IRRAD/CHARM
2015
n-ToF
2001
PS
1959 (628 m)
CLEAR
2017
LINAC 2
LINAC 3
Ions
LEIR
2005 (78 m)

[1]

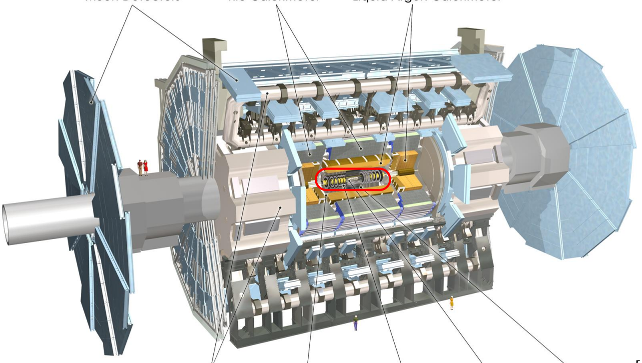
A Toroidal LHC ApparatuS



[2]

ATLAS

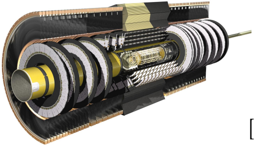
- Largest detector ever constructed at a particle collider
- General-/multi-purpose detector
- Many-layer design with nearly cylindrical symmetry



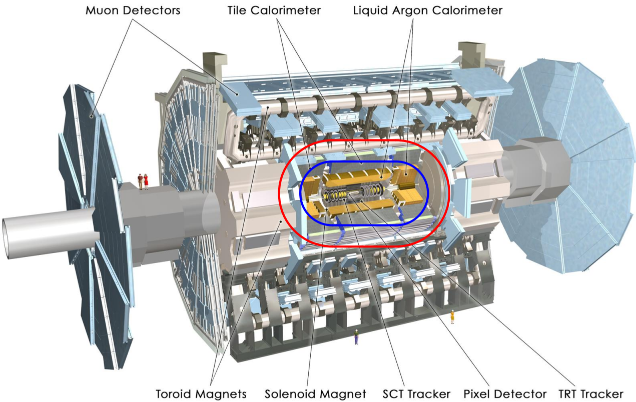
[2]

Inner Detector

- Charged particle track measurement
  → Charge & momentum
- Layers of silicon pixels and strips
- Tiny gas-filled drift tubes



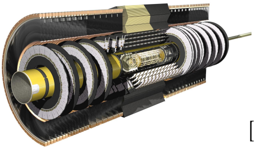
[3]



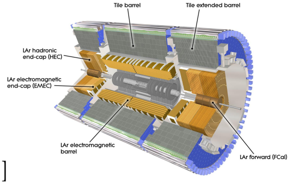
[2]

Calorimeters

- Particle energy measurement
- 2 sampling calorimeters:
  → electromagnetic
  → hadronic

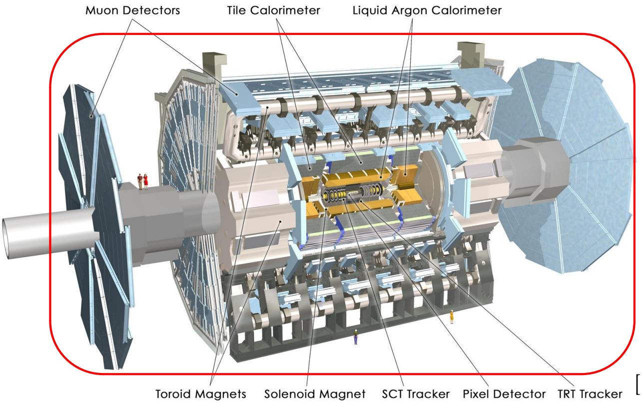


[3]



[4]

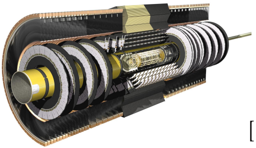
A Toroidal LHC ApparatuS



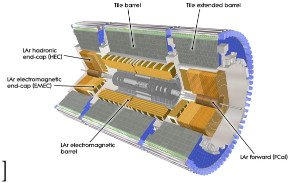
[2]

Muon Spectrometer

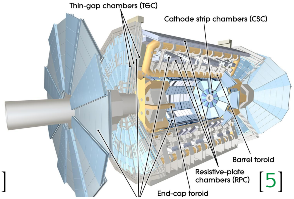
- Muon track measurement
  → Charge & momentum
- Contains 4 different detector types
- Same principle as for inner detector



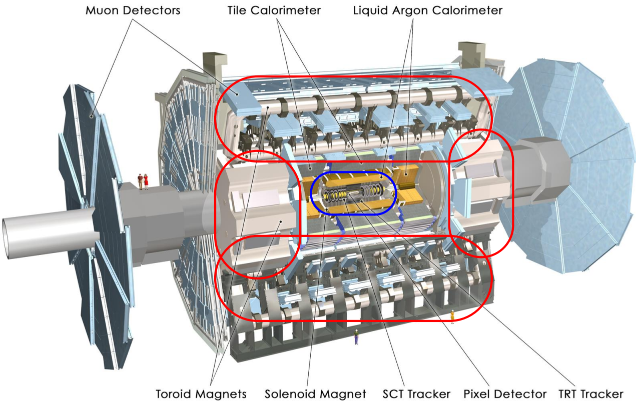
[3]



[4]



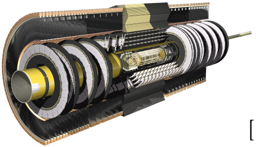
[5]



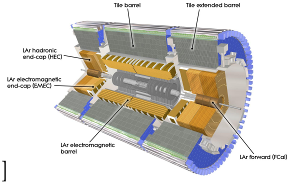
[2]

Magnet System

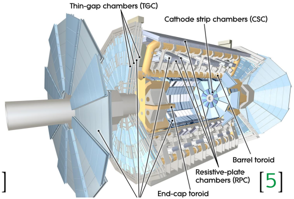
- 2 magnetic fields with different orientation, created by
  → Solenoid magnet for inner detector
  → Toroid magnet for muon measurement



[3]



[4]



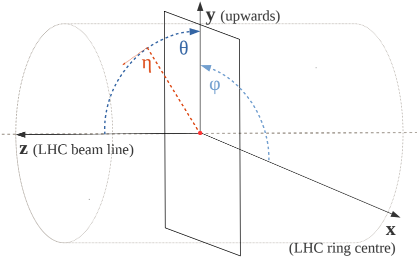
[5]

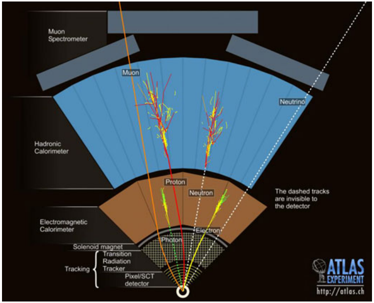
A Toroidal LHC ApparatuS

Coordinate System

- Right-handed, interaction point as centre
- Azimuthal angle φ in transverse plane
- Separation from beam line via polar angle θ
- Instead, often use pseudorapidity

$$\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right]$$





[6]

Object Features

- Leptons appear as distinct objects (tracks & calorimeter clusters)
- Colour-charged particles hadronise to cone-shaped showers of uncoloured particles $\Rightarrow$ **Jets** (reconstructed by anti-k_t algorithm [7])
- Neutrinos leave detector unseen, only reconstructable via energy- and momentum conservation in transverse plane

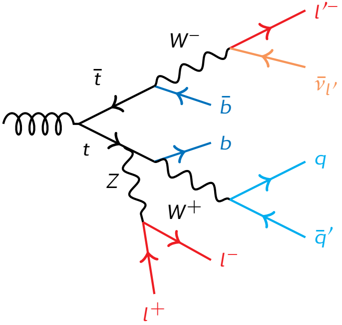
$$E_T^{\text{miss}} = \left| -\sum_i \vec{p}_{T,i} \right|$$

Signature:

- 2 jets from *b*-quarks ("*b*-jets")
- 2 jets from light quarks (u, d, c, s)
- 3 charged light leptons (e, μ)
- 1 neutrino

Dominant backgrounds:
WZ, tWZ, tZq, fake leptons

Branching: 2.3 %



$\Rightarrow$ **Requiring ≥ 3 jets, ≥ 2 b-jets and exactly 3 leptons**

Event Reconstruction

Motivation:

- Better probe of underlying top quark kinematics
- Allow for several sensitive differential variables

Strategy:

- Only **partial reconstruction** of signal signature
- Focus on top quark with **subsequent leptonic *W*-decay** ("leptonic-side")

Leptonic-Side Top Quark Reconstruction

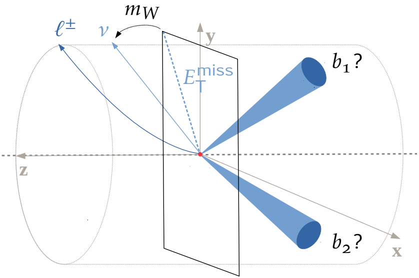
- Require only ≥ 2 jets
- Consider 2 leptons of opposite charge and same flavour with an invariant mass best consistent with Z boson mass to be Z decay products
- Assume neutrino from W-decay to be predominant source of missing **transverse** energy
- Apply W mass constraint: $\boxed{m_{l^\pm \nu} \equiv m_W}$

  → 2 neutrino p_z solutions from quadr. eq.
- Consider both solutions and combine with either of 2 b-jets
- Assign output weight to reconstructed top quark candidates based on interpolated value from:

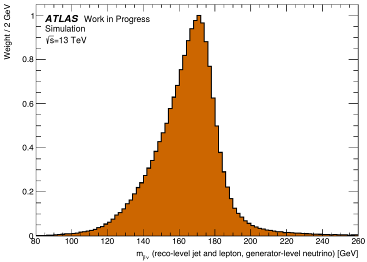
$m_{jl\nu}$ distribution

- Invariant mass distribution of correctly reconstructed top quarks in simulated $t\bar{t}Z$ events
- MC generator-level neutrino
- Reconstructed lepton and jet (MC)

⇒ Only consider top quark candidate with maximum output weight per event

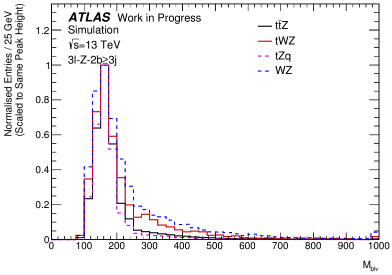


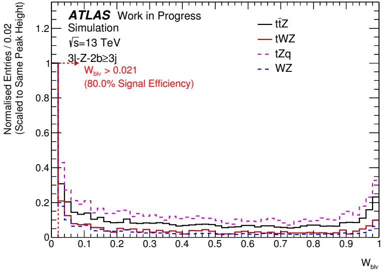
$m_{jl\nu}$ distribution



Plot kindly provided by T. McCarthy (MPI Munich)

Leptonic-Side Top Quark Reconstruction





- Use output weight to suppress non-fake backgrounds by setting cut on resulting distribution, e.g. at 80 % signal efficiency
- Focus on minimisation of signal and background uncertainties as they enter both the inclusive and the differential cross section measurement
- **Note:** inclusive cross section measured with fit, use calculation to estimate effect of cut

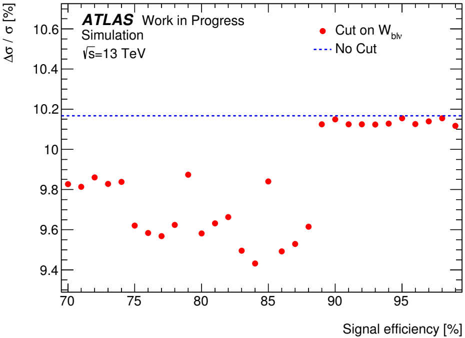
Cross Section Calculation

- Estimate uncertainties with cross section formula

$$\sigma^{\mathrm{obs}}_{t\bar{t}Z} = \frac{N^{\mathrm{obs}}_{t\bar{t}Z}}{\mathcal{L}\cdot\mathcal{BR}\cdot\varepsilon_{t\bar{t}Z}}\,, \qquad N^{\mathrm{obs}}_{t\bar{t}Z} \equiv N_{\mathrm{Data}} - N^{\mathrm{MC}}_{\mathrm{Bkg}}$$

- With $\varepsilon_{t\bar{t}Z}$ as the fraction of selected reconstructed MC events $N^{\mathrm{MC}}_{t\bar{t}Z}$ among the total number of generated events, the formula for the observed cross section becomes

$$\sigma^{\mathrm{obs}}_{t\bar{t}Z} = \frac{N^{\mathrm{obs}}_{t\bar{t}Z}}{N^{\mathrm{MC}}_{t\bar{t}Z}}\cdot\sigma_{\mathrm{MC}} = \sigma^{\mathrm{obs}}_{t\bar{t}Z}(N_{\mathrm{Data}}, N^{\mathrm{MC}}_{\mathrm{Bkg}}, N^{\mathrm{MC}}_{t\bar{t}Z})$$



- Full systematic (detector, theory) and statistical uncertainties applied
- Minimum uncertainty on cross section at signal efficiency of 84 %

⇒ Most dominant *WZ* background reduced by 42 %

- Values scattered in range of only 0.8 %
- However, certain trend towards minimum noticable

Relative cross section uncertainty:

$$\frac{\Delta\sigma}{\sigma} = \frac{1}{N_{\mathrm{Data}} - N_{\mathrm{Bkg}}^{\mathrm{MC}}} \cdot \sqrt{\left(\Delta N_{\mathrm{Data}}\right)^2 + \left(\Delta N_{\mathrm{Bkg}}^{\mathrm{MC}}\right)^2 + \left(\frac{N_{\mathrm{Data}} - N_{\mathrm{Bkg}}^{\mathrm{MC}}}{N_{t\bar{t}Z}^{\mathrm{MC}}}\right)^2 \left(\Delta N_{t\bar{t}Z}^{\mathrm{MC}}\right)^2}$$

Conclusions

- Inclusive and differential measurement of $t\bar{t}Z$ cross section with full LHC Run 2 dataset taken by ATLAS
- Performed partial reconstruction of $t\bar{t}$ system to exploit top quark kinematics restricting to top quark with subsequent leptonic W decay
- Used relative cross section uncertainty as measure to find optimum cut value on reconstruction output weight
- Minimised uncertainties entering inclusive $t\bar{t}Z$ cross section measurement by setting a cut at certain signal efficiency
- For differential measurement applying cut might get very harmful to precision due to reduced data statistics

Backup

Using four-momentum conservation,

$$(p_\ell + p_\nu)^2 = m_W^2$$

Expanding the left-hand side

$$\begin{aligned}
p_\ell^2 + p_\nu^2 + 2p_\ell \cdot p_\nu &= m_W^2 \\
E_\ell E_\nu - \vec{p}_\ell \cdot \vec{p}_\nu &= \frac{m_W^2 - m_\ell^2}{2} \\
E_\ell E_\nu - p_{\ell x} p_{\nu x} - p_{\ell y} p_{\nu y} - p_{\ell z} p_{\nu z} &= \frac{m_W^2 - m_\ell^2}{2}
\end{aligned}$$

Treating the neutrino as massless ($E^2 = p_\mathrm{T}^2 + p_z^2$), and with polar coordinates

$$\sqrt{p_{\mathrm{T}\nu}^2 + p_{\nu z}^2}\, E_\ell = \frac{m_W^2 - m_\ell^2}{2} + p_{\mathrm{T}\nu}\left(p_{\ell x}\cos\phi_\nu + p_{\ell y}\sin\phi_\nu\right) + p_{\ell z} p_{\nu z}$$

Squaring and defining $\alpha \equiv \frac{m_W^2 - m_\ell^2}{2} + p_{\mathrm{T}\nu}\left(p_{\ell x}\cos\phi_\nu + p_{\ell y}\sin\phi_\nu\right)$

$$\left(p_{\mathrm{T}\nu}^2 + p_{\nu z}^2\right) E_\ell^2 = \alpha^2 + 2\alpha p_{\ell Z} p_{\nu z} + p_{\ell z}^2 p_{\nu z}^2$$

Collecting the terms in powers of $p_{\nu z}$

$$\boxed{Ap_{\nu z}^2 + Bp_{\nu z} + C = 0}$$

With

$$\begin{aligned} A &= E_\ell^2 - p_{\ell z}^2 \\ B &= -2\alpha p_{\ell z} \\ C &= p_{\mathrm{T}\nu}^2 E_\ell^2 - \alpha^2 \end{aligned}$$

If discriminant $B^2 - 4AC > 0$, two real solutions can be evaluated

$$p_{\nu z,1/2} = \frac{-B \pm \sqrt{B^2 - 4AC}}{2A}$$

In case of no real solution, two options exist:

- Smearing $E_{\mathrm{T}}^{\mathrm{miss}}$ until discriminat becomes positive again
- Find value of $p_{\mathrm{T}\nu}$ such that $B^2 = 4AC$

Cross Section Calculation

The measurable cross section is given by

$$\sigma_{t\bar{t}Z}^{\text{obs}} = \frac{N_{\text{Data}} - N_{\text{Bkg}}^{\text{MC}}}{\mathcal{L} \cdot \mathcal{BR} \cdot \varepsilon_{t\bar{t}Z}}$$

where the product of signal acceptance and efficiency can be written as

$$\varepsilon_{t\bar{t}Z} = \frac{N_{t\bar{t}Z}^{\text{MC}}}{\sigma_{\text{MC}} \cdot \mathcal{L} \cdot \mathcal{BR}}$$

The Monte Carlo have been produced inclusively in $t\bar{t}$ decay so that only the Z branching fraction matters. The Monte Carlo production cross section σ_{MC} included also higher order calculations.

Together with the number of observed events $N_{t\bar{t}Z}^{\text{obs}} \equiv N_{\text{Data}} - N_{\text{Bkg}}^{\text{MC}}$ the cross section formula can be rewritten:

$$\sigma_{t\bar{t}Z}^{\text{obs}} = \underbrace{\frac{N_{t\bar{t}Z}^{\text{obs}}}{N_{t\bar{t}Z}^{\text{MC}}}}_{\mu_{t\bar{t}Z}} \cdot \sigma_{\text{MC}}$$

With Gaussian error propagation

$$\Delta y(x_i) = \sqrt{\sum_i \left(\frac{\partial y}{\partial x_i} \Delta x_i \right)^2}$$

the error on the cross section evaluates to

$$\begin{aligned}
\Delta\sigma &= \sqrt{\left(\frac{\sigma_{\mathrm{MC}}}{N_{t\bar{t}Z}^{\mathrm{MC}}} \Delta N_{\mathrm{Data}} \right)^2 + \left(\frac{\sigma_{\mathrm{MC}}}{N_{t\bar{t}Z}^{\mathrm{MC}}} \Delta N_{\mathrm{Bkg}}^{\mathrm{MC}} \right)^2 + \left(\frac{N_{\mathrm{Data}}}{\left(N_{t\bar{t}Z}^{\mathrm{MC}}\right)^2} \sigma_{\mathrm{MC}} \right)^2 \left(\Delta N_{t\bar{t}Z}^{\mathrm{MC}} \right)^2} \\
&= \frac{\sigma_{\mathrm{MC}}}{N_{t\bar{t}Z}^{\mathrm{MC}}} \sqrt{(\Delta N_{\mathrm{Data}})^2 + \left(\Delta N_{\mathrm{Bkg}}^{\mathrm{MC}} \right)^2 + \left(\frac{N_{\mathrm{Data}} - N_{\mathrm{Bkg}}^{\mathrm{MC}}}{N_{t\bar{t}Z}^{\mathrm{MC}}} \right)^2 \left(\Delta N_{t\bar{t}Z}^{\mathrm{MC}} \right)^2}
\end{aligned}$$

Hence, the relative uncertainty on the cross section can be written as

$$\frac{\Delta\sigma}{\sigma} = \frac{\sqrt{(\Delta N_{\mathrm{Data}})^2 + \left(\Delta N_{\mathrm{Bkg}}^{\mathrm{MC}} \right)^2 + \left(\frac{N_{\mathrm{Data}} - N_{\mathrm{Bkg}}^{\mathrm{MC}}}{N_{t\bar{t}Z}^{\mathrm{MC}}} \right)^2 \left(\Delta N_{t\bar{t}Z}^{\mathrm{MC}} \right)^2}}{N_{\mathrm{Data}} - N_{\mathrm{Bkg}}^{\mathrm{MC}}}$$

Differential Measurement and Unfolding

Inclusive Approach:

- Select events featuring characterisitcs of signal signature
- Possibly perform reconstruction to better probe underlying kinematics
- Get total cross section inclusive in by
  - Calculating according to a certain formula
  - Performing a fit (single or multiple bins) to extract signal strength

Differential Approach:

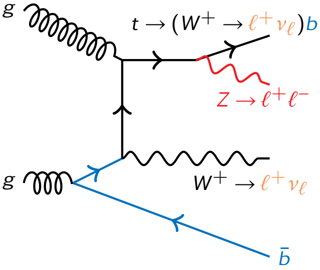
- Measure cross sections as functions of certain variables
- Unfold detector-level distributions of variables to particle- and parton level, respectively
- Correction of detector effects, as well as signal acceptance and efficiency w.r.t. given phase space region

$$\frac{\mathrm{d}\sigma}{\mathrm{d}X^i} = \frac{1}{\mathcal{L} \cdot \mathcal{BR} \cdot \Delta X^i \cdot f^i_{\mathrm{acc}}} \sum_j R^{-1}_{ij} \cdot \epsilon^j_{\mathrm{eff}} \cdot \left(N^j_{\mathrm{Data}} - N^j_{\mathrm{Bkg}}\right)$$

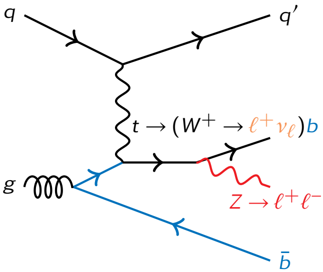
  with i the bin index of observable X^i with bin width ΔX^i. Background contribution estimated from Monte Carlo, N^j_{Bkg}, is subtracted from observed data, N^j_{Data}, providing estimated observed signal in bin j.

- Correction terms: $\epsilon_{\mathrm{eff}} = N^{\mathrm{reco}\wedge\mathrm{truth}}/N^{\mathrm{reco}}, \quad f_{\mathrm{acc}} = N^{\mathrm{reco}\wedge\mathrm{truth}}/N^{\mathrm{truth}}$
- Inversion of response matrix R done iteratively by applying repeatedly Bayes′ theorem:
  - Fold: $x_i^{(k)} = \sum_j R_{ij} \mu_j^{(k)}$
  - Unfold: $\mu_j^{(k+1)} = \frac{1}{\epsilon_i} \sum_i R_{ij} \mu_j^{(k)} \left(\frac{x_i}{x_i^{(k)}} \right)$
- Updating guess by applying Bayes:
  $\hat{\mu}_i = \frac{1}{\epsilon_i} \sum_j P(\text{generated in bin} i | \text{observed in bin} j) n_j = \frac{1}{\epsilon_i} \sum_j \left(\frac{R_{ij} p_i}{\sum_k R_{jk} p_k} \right) n_j$

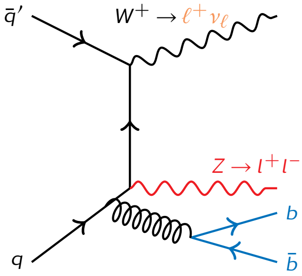
tWZ:


g
g
$t \rightarrow (W^+ \rightarrow \ell^+ \nu_\ell) b$
$Z \rightarrow \ell^+ \ell^-$
$W^+ \rightarrow \ell^+ \nu_\ell$
$\bar{b}$

tZq:


q
q'
g
$t \rightarrow (W^+ \rightarrow \ell^+ \nu_\ell) b$
$Z \rightarrow \ell^+ \ell^-$
$\bar{b}$

WZ:


$\bar{q}'$
$W^+ \rightarrow \ell^+ \nu_\ell$
$Z \rightarrow l^+ l^-$
q
b
$\bar{b}$

References

[1] Esma Mobs. "The CERN accelerator complex - August 2018. Complexe des accélérateurs du CERN - Août 2018". General Photo. 2018. URL: http://cds.cern.ch/record/2636343.

[2] Joao Pequenao. "Computer generated image of the whole ATLAS detector". 2008. URL: http://cds.cern.ch/record/1095924.

[3] Joao Pequenao. "Computer generated image of the ATLAS inner detector". 2008. URL: http://cds.cern.ch/record/1095926.

[4] Joao Pequenao. "Computer Generated image of the ATLAS calorimeter". 2008. URL: http://cds.cern.ch/record/1095927.

[5] Joao Pequenao. "Computer generated image of the ATLAS Muons subsystem". 2008. URL: http://cds.cern.ch/record/1095929.

[6] Joao Pequenao. "Event Cross Section in a computer generated image of the ATLAS detector." 2008. URL: https://cds.cern.ch/record/1096081.

[7] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "The anti-k_t jet clustering algorithm". In: *JHEP* 04 (2008), p. 063. DOI: 10.1088/1126-6708/2008/04/063. arXiv: 0802.1189 [hep-ph].