Estimating W+Jets Background from Data

Sascha Mehlhase

Deutsches Elektronen-Synchrotron DESY

Zeuthen Site

"Physics at the Terascale"

3rd Annual Workshop

November 11th 2009

Method overview

- W+jets important background for top and new physics
- MC predictions (e.g. ALPGEN) not very precise

- Use data driven approach, with following assumptions / facts
  - W+Jets background distributed over wide rapidity range and dominant at low jet multiplicities
  - $t\bar{t}$ instead dominates in central rapidity region and at high multiplicities

- First approach
  - Event selection – exactly one lepton (e, μ) and $|m_{T,inv}^{l+\nu} - m_W^{PDG}| < 25$ GeV
  - Extrapolate W+Jets contribution from low jet multiplicities and large rapidity to high jet multiplicity and low rapidity (signal region)
  - For now: signal = $t\bar{t}_{lep}$, background = W+Jets
  - Test method by measuring $t\bar{t}_{lep}$ cross-section

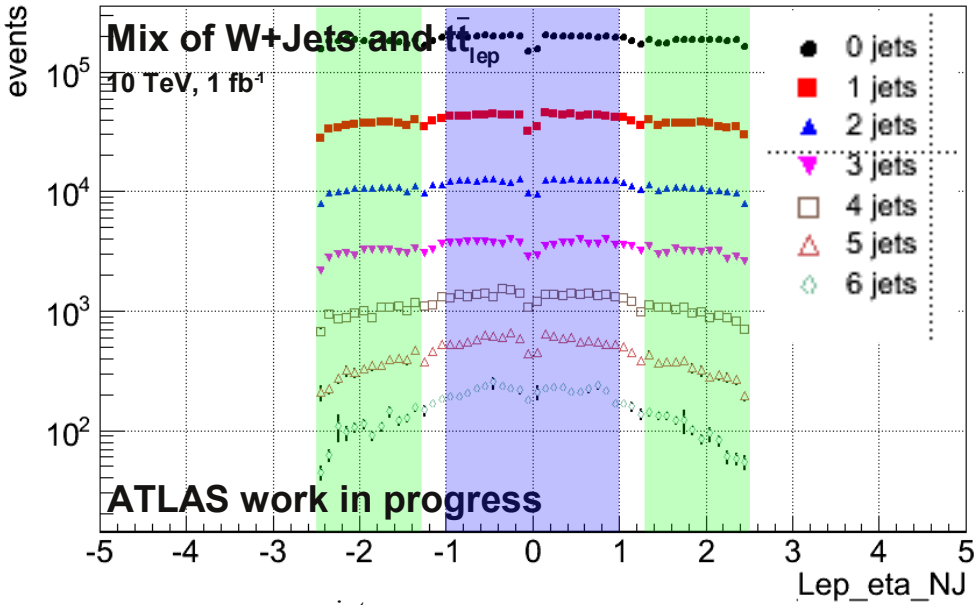
- Based on method by V. Pavlunin and D. Stuart (Phys.Rev.D78:035012,2008)

Method walk-through

- Determine $R_{N^{jet}} \equiv Y^{central}_{N^{jet}} / (Y^{central}_{N^{jet}} + Y^{forward}_{N^{jet}})$

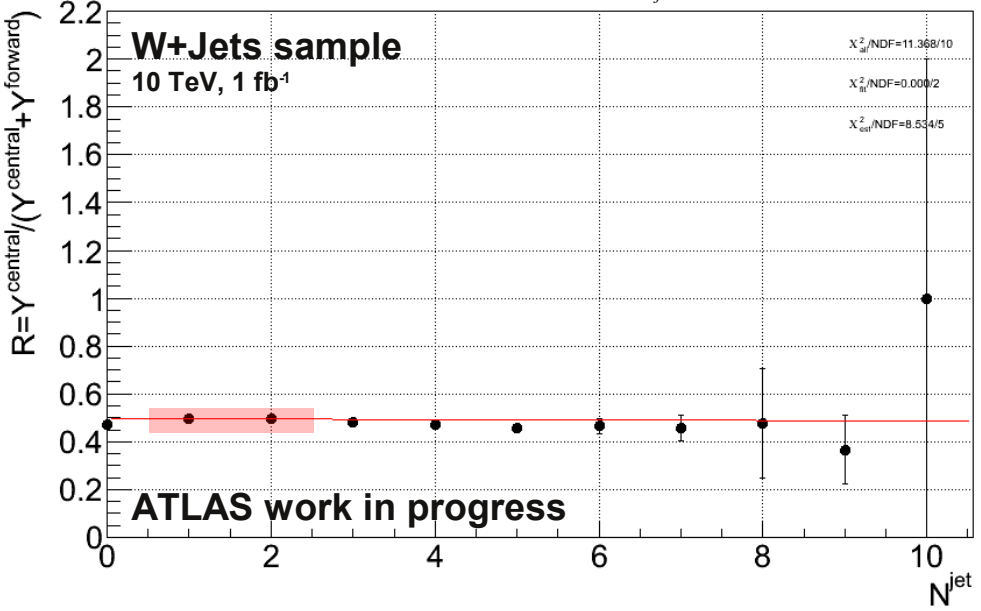
$$Y^{central}_{N^{jet}} \equiv N^{|\eta_{lepton}| \leq 1}_{events}$$

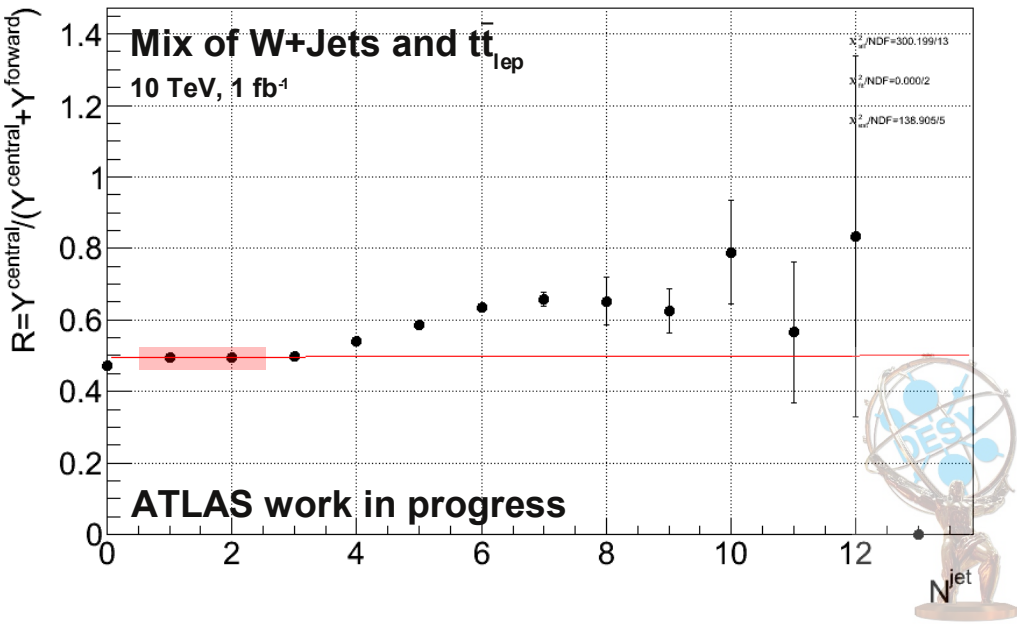
$$Y^{forward}_{N^{jet}} \equiv N^{1.3 \leq |\eta_{lepton}| \leq 2.5}_{events}$$



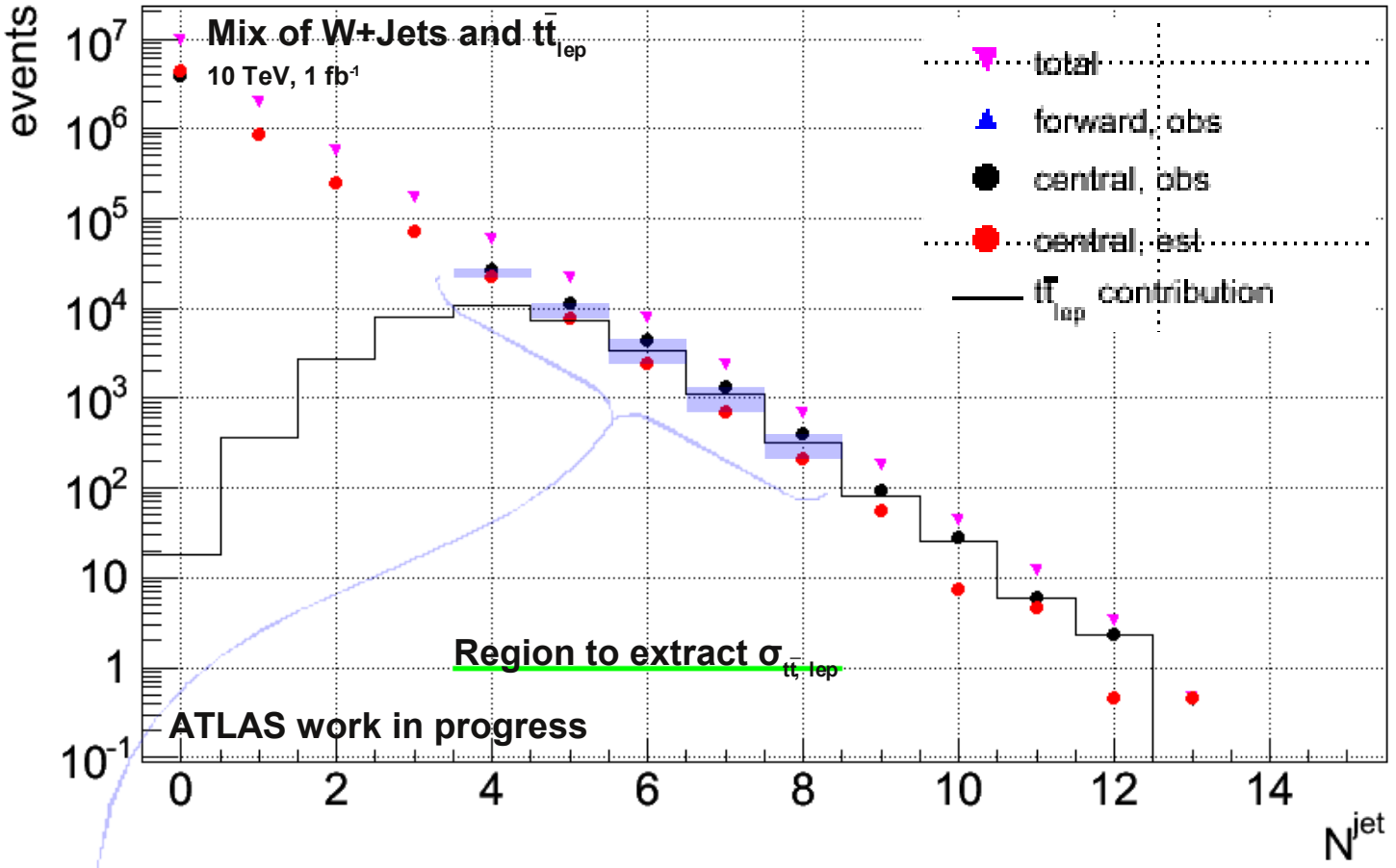
- Fit $R = R(N^{jet})$ with linear function at low multiplicities $1 \leq N^{jet} \leq 2$

- Extract $Y^{central,\,est}_{N^{jet}} = \frac{R_{N_{jet}}}{1 - R_{N_{jet}}} Y^{forward}_{N^{jet}}$ at high multiplicities $4 \leq N^{jet} \leq 8$





Method walk-through



- From shaded area one can extract the $t\bar{t}$ cross-section as a consistency check

$$\sigma_{t\bar{t}_{lep}} = \frac{N_{obs} - N_{est}}{L \cdot \varepsilon^{total}_{t\bar{t}}}$$

$$\varepsilon^{total}_{t\bar{t}_{lep}} = \underbrace{\varepsilon^{selection}_{t\bar{t}_{lep}} \cdot \varepsilon^{central}_{t\bar{t}_{lep}} \cdot \varepsilon^{Njet}_{t\bar{t}_{lep}}}_{\text{from MC}}$$

$$\varepsilon^{selection}_{t\bar{t}_{lep}} = N^{selected}_{t\bar{t}_{lep}} / N^{total}_{t\bar{t}_{lep}}$$

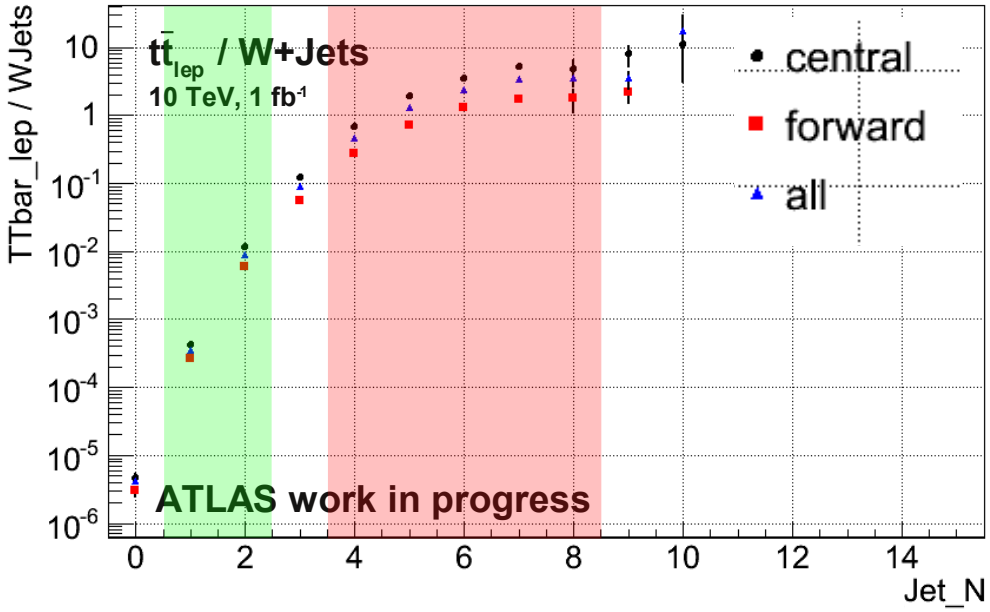
$$\varepsilon^{central}_{t\bar{t}_{lep}} = N^{|\eta_{lep}|<1}_{t\bar{t}_{lep}} / N^{total}_{t\bar{t}_{lep}}$$

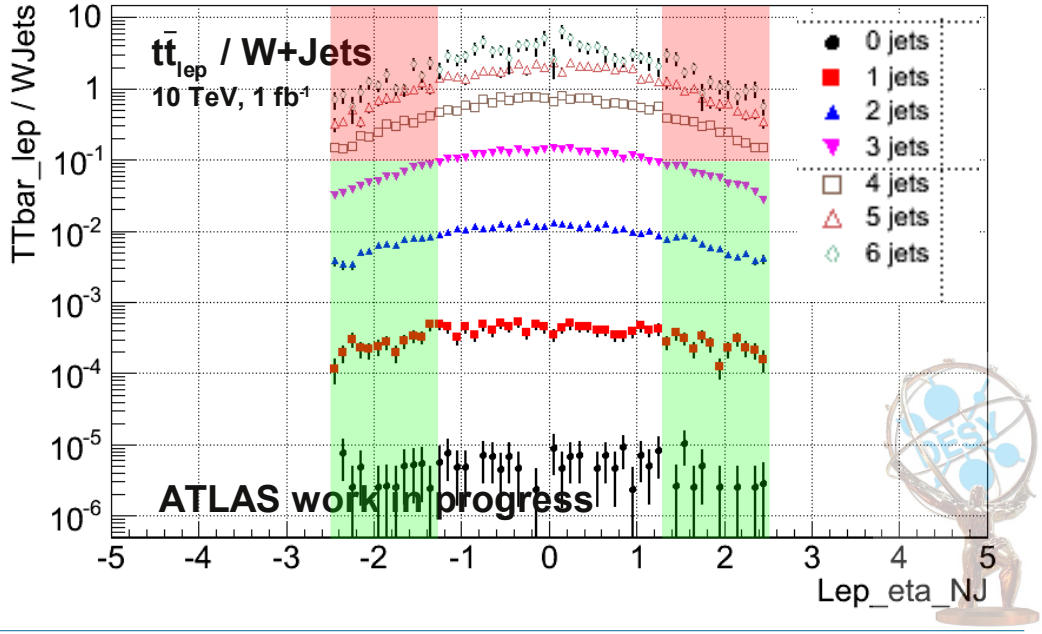
$$\varepsilon^{N_{jet}}_{t\bar{t}_{lep}} = N^{4\le N_{jet}\le 8}_{t\bar{t}_{lep}} / N^{total}_{t\bar{t}_{lep}}$$

Limitation of first approach

- Followed paper up to here
- To cross-check with the $t\bar{t}_{lep}$ cross-section measurement the method has to be extended
- To calculate $Y_{N^{jet}}^{central,\,est}$ we assumed $Y_{N^{jet}}^{forward}$ to be (see page 3)
  - Free of background at large multiplicities
  - Free of signal at low multiplicities
  - Assumption exact for $t\bar{t}$? » no, too much signal in forward region at high multiplicity!

    » simple counting (see page 4) doesn't work ...





Modified approach

- … the overlap has to be taken into account for

$$\sigma_{t\bar{t}} = \frac{Y(Mix)^{central} - \frac{R_{N_{jet}}}{1-R_{N_{jet}}} Y(Mix)^{forward}}{\left(\varepsilon_{t\bar{t}_{lep}}^{central} - \frac{R_{N_{jet}}}{1-R_{N_{jet}}} \varepsilon_{t\bar{t}_{lep}}^{forward} \right) \cdot \varepsilon_{t\bar{t}_{lep}}^{N_{jet}} \cdot \varepsilon_{t\bar{t}_{lep}}^{sel} \cdot L}$$

from MC

  - Efficiency has to be corrected for $t\bar{t}$-contamination in forward region

- Which yields $\sigma_{t\bar{t}_{lep}} = \left(136.07 \pm 2.43^{stat}\right) pb$ with $\Delta_{\varepsilon_{t\bar{t}_{lep}}^{central}} = \Delta_{\varepsilon_{t\bar{t}_{lep}}^{forward}} = \Delta_{\varepsilon_{t\bar{t}_{lep}}^{N_{jet}}} = \Delta_{\varepsilon_{t\bar{t}_{lep}}^{sel}} = \Delta_L = 0$ for now
  - including 10 % luminosity error $\sigma_{t\bar{t}_{lep}} = \left(136.07 \pm 13.82^{stat+lumi}\right) pb$

- Expected cross-section is $\sigma_{t\bar{t}_{lep}} \approx 120\ pb$

Checks for stability of result

- Range of N_{jet} bins to determine $t\bar{t}$ cross-section changed
  - Default is 4 to 8, also tested other ranges
  - Obtained cross-section should be stable within errors

    4 to 6 146.51 ± 2.84 pb

    4 to 7 138.73 ± 2.52 pb

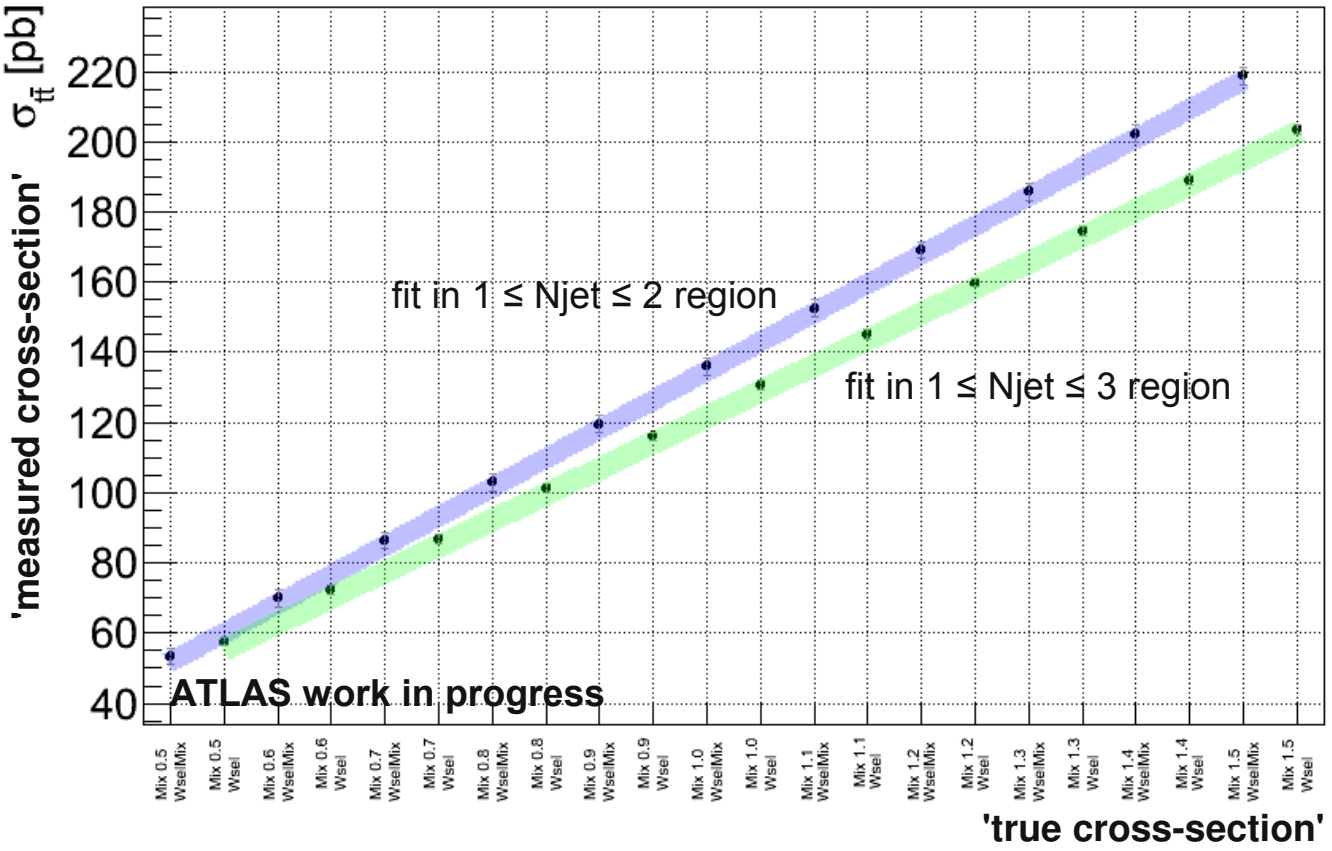
    5 to 8 108.46 ± 1.12 pb

    4 to 8 136.07 ± 2.43 pb
- Range of fit on R could be extended to $N_{jet} = 3$ bin
  - Pro: fit becomes more stable
  - Con: already quite high $t\bar{t}$-contribution

Mix	4 to 6	140.28 ± 1.51 pb	W+Jets	4 to 6	-17.96 ± 1.42 pb
Mix	4 to 7	133.15 ± 1.34 pb	W+Jets	4 to 7	-16.07 ± 1.26 pb
Mix	5 to 8	106.18 ± 0.62 pb	W+Jets	5 to 8	-6.52 ± 0.59 pb
Mix	4 to 8	130.67 ± 1.29 pb	W+Jets	4 to 8	-15.44 ± 1.21 pb

Linearity check

- $t\bar{t}$ cross-section as a function of tt content in mixture with W+Jets
  - $t\bar{t}$ event scale with factors from 0.5 to 1.5
  - Linear dependency of cross-section expected and seen
  - Allows for 'simple' correction in case the cross-section is not know precisely
    - e.g. correct for too high estimated $t\bar{t}_{lep}$ cross-section (see page 6)



Summary & Conclusion

- First test of original and modified method on fully simulated events in ATLAS
- First results look promising, further cross-checks needed
- Next steps
  - Include uncertainties on efficiencies
  - Include other backgrounds (already started)
    - » QCD - *tricky* due to too low MC statistics and resulting large event weights
    - » single top, Z+Jets, $t\bar{t}_{had}$
  - Look at other systematics
    - » Jet energy scale
    - » influence of MET (already started)

Modified approach

$$Y(Mix)^{central}_{N^{jet}} = Y(W+Jets)^{central}_{N^{jet}} + Y(t\bar{t})^{central}_{N^{jet}}$$

$$Y(Mix)^{forward}_{N^{jet}} = Y(W+Jets)^{forward}_{N^{jet}} + Y(t\bar{t})^{forward}_{N^{jet}}$$

$$Y(W+Jets)^{central\,,\,est}_{N^{jet}} = \frac{R_{N_{jet}}}{1-R_{N_{jet}}} Y(W+Jets)^{forward}_{N^{jet}}$$

$$Y(t\bar{t})^{forward} = \sigma_{t\bar{t}} \cdot L \cdot \varepsilon^{forward}_{t\bar{t}_{lep}}$$

$$Y(t\bar{t})^{central} = \sigma_{t\bar{t}} \cdot L \cdot \varepsilon^{central}_{t\bar{t}_{lep}}$$

$$\sigma_{t\bar{t}} = \frac{Y(Mix)^{central} - \frac{R_{N_{jet}}}{1-R_{N_{jet}}} Y(Mix)^{forward}}{\left(\varepsilon^{central}_{t\bar{t}_{lep}} - \frac{R_{N_{jet}}}{1-R_{N_{jet}}} \varepsilon^{forward}_{t\bar{t}_{lep}}\right) \cdot \varepsilon^{N_{jet}}_{t\bar{t}_{lep}} \cdot \varepsilon^{sel}_{t\bar{t}_{lep}} \cdot L}$$

from MC