Universidad de Oviedo
Universidá d'Uviéu
University of Oviedo

ICTEA
Instituto Universitario de Ciencias y Tecnologías Espaciales de Asturias

CMS

04-11-2020

4th Red LHC Workshop

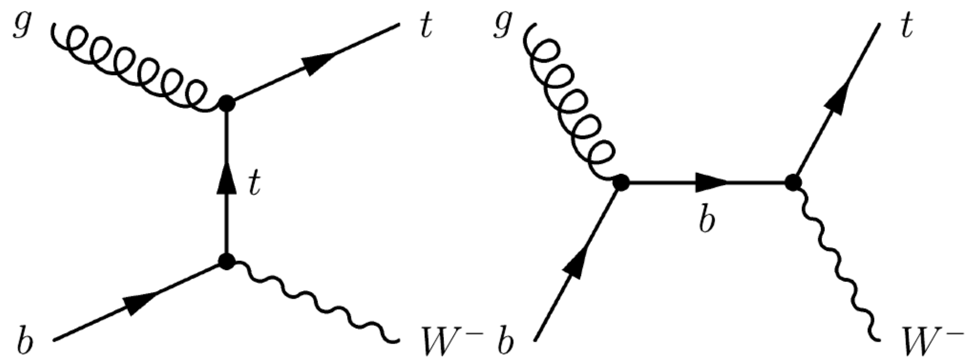
Measurement of differential cross sections for single top quark production in association with a W boson at √s=13 TeV

Víctor Rodríguez Bouza (on behalf of the CMS Collaboration)

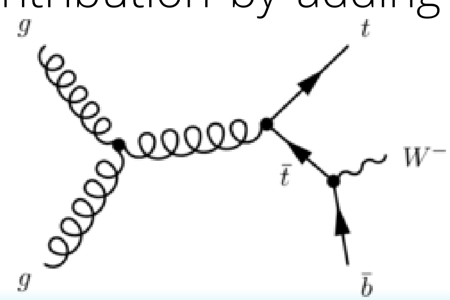
Table of contents

Analysis overview

- **Aim**: measurement of tW differential cross-section with 2016 data using dilepton final states.



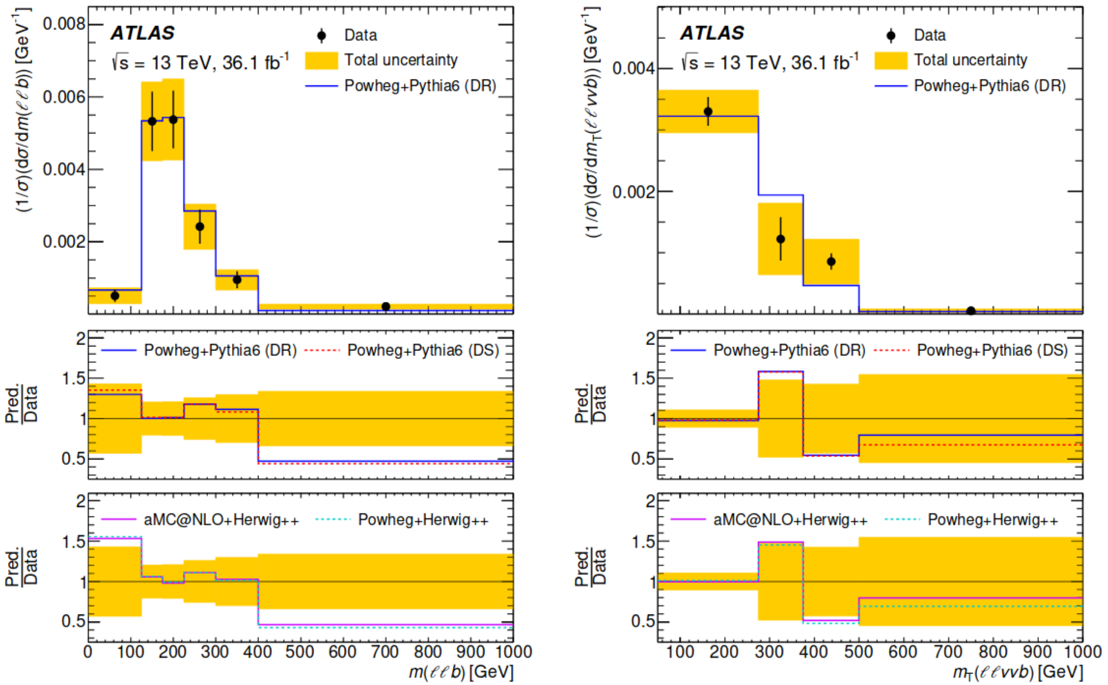
- Main challenge: **background** *largely* **dominates signal**, being the most important $t\bar{t}$.
- This is a consequence of its similarity with the most important way of producing top quarks at p-p collisions: pair production. Actually, at NLO, **they interfere**. In order to resolve both processes' definitions, and avoid double counting issues, two approaches (JHEP 07 (2008) 029) are used to obtain the tW simulation samples.
  - Diagram Removal (**DR**): we remove Feynman diagrams that might present two on-shell tops (also called *double resonant*).
  - Diagram Subtraction (**DS**): we remove locally the pair-production contribution by adding an artificial term in the calculation.



- Its relationship with $t\bar{t}$ is key in order to perform the analysis.

Analysis overview – Previous measurements

- The differential tW analysis with 2016 data has already been **published by ATLAS**: *"Measurement of differential cross-sections of a single top quark produced in association with a W boson at √s = 13 TeV with ATLAS"*. **Eur. Phys. J. C 78 (2018) 186**.
  - Uses a BDT to enhance S/B discrimination, obtaining a pure signal region.



- Our approach is based on CMS' **tW inclusive cross section measurement analysis** (**10.1007/JHEP10(2018)117**).
  - Not using MVA's: we are free to unfold any kind of distribution.
  - S/B ratio enhanced by optimising event selection.

- Recently, CMS has published another single-top **differential measurement**, in the **t-channel**: *"Measurement of differential cross sections and charge ratios for t-channel single top quark production in proton-proton collisions at √s= 13 TeV"*, **Eur. Phys. J. C 80 (2020) 370**.

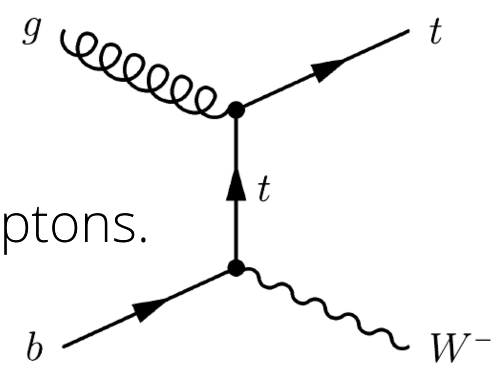
Analysis overview - Selected distributions

Global schema of the analysis is the following.

1) Apply an event selection to a **signal-enriched region**.
2) **Unfold** chosen variables to a particle level region.
3) Calculate the normalised differential cross section depending on those observables.

The selected observables to be unfold are the following (here ℓ_1, ℓ_2 refer to the event's leptons ordered by p_T).

- $p_T(\ell_1)$: provides kinematic information of the process.
- $p_T(j)$: provides kinematic information of the process.
- $\Delta\varphi(\ell_1, \ell_2)$: sensible to spin features and correlations of both leptons.
- $m(\ell_1, \ell_2, j)$: explore correlations between objects.
- $m_T(\ell_1, \ell_2, j, \not{E}_T)$: explore correlations between objects (in the transversal plane).
- $p_Z(\ell_1, \ell_2, j)$: probe for the boost of the system.



Data samples and Monte Carlo simulations

- The analysis is performed using the complete 2016 dataset (35.9 fb^{-1}).
- The trigger strategy uses a combination of single and double triggers to maximise efficiency.
- To get particle level information, the RIVET pseudoparticle producer is used as recommended by "*Object definitions for top quark analyses at the particle level*" (**CMS-NOTE-2017-007**).
- Signal MC samples:
  - DR: NLO `POWHEGv1 + Pythia8` (nominal)
  - DS: NLO `POWHEGv1 + Pythia8`
  - DR: NLO `aMC@NLOv2.2.2 + Pythia8`
- Background MC samples:
  - $t\bar{t}$: NLO `POWHEGv2 + Pythia8`
  - **DY**: NLO M50 & M10to50 `aMC@NLOv2.2.2FxFx + Pythia8`
  - **W+jets**: NLO `aMC@NLOv2.2.2FxFx + Pythia8`
  - **VV +** $t\bar{t}$**V**: LO `MLM + Pythia8` and NLO `aMC@NLOv2.2.2 + Pythia8`

Object identification

- After the **Particle Flow** algorithm reconstructs and identifies the candidates from each event, other requirements are imposed over them (summary).

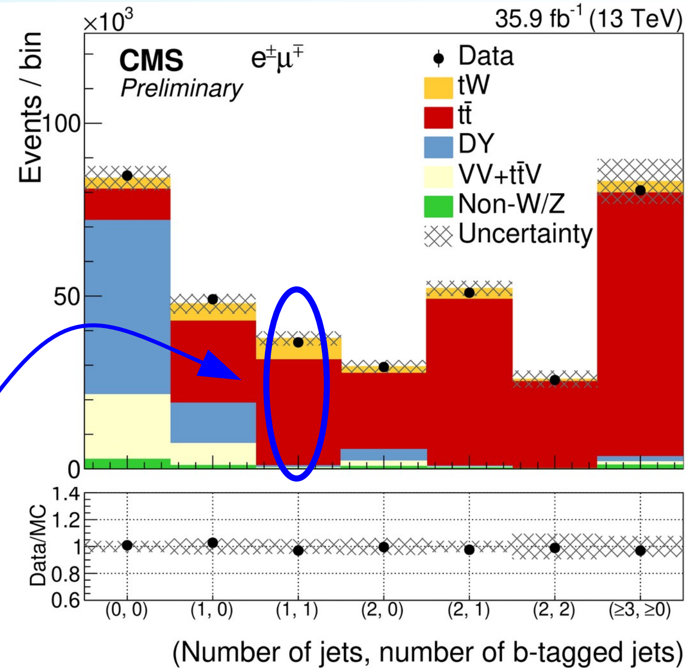
- **Electrons**
  - $p_T > 20$ GeV
  - $|\eta| < 2.4$ (plus veto of transition region: $1.4442 < |\eta| < 1.5660$).
  - Isolation <~ 0.05 (small change depending on pseudorapidity).

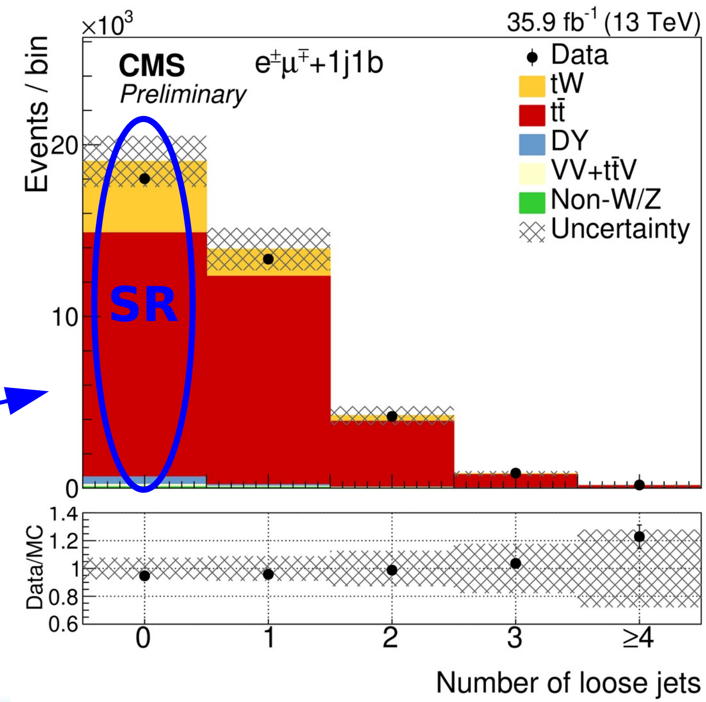
- **Muons**
  - $p_T > 20$ GeV
  - $|\eta| < 2.4$.
  - Isolation < 0.15.

- **Jets**
  - $p_T > 30$ GeV .
  - $|\eta| < 2.4$.
  - Jets are identified as coming from a b quark using a MVA algorithm for those candidates (CSVv2).
    - WP with ~70% tagging eff. and mistagging prob. of ~1% for light-flavour jets and 15% for c-jets.
  - Loose jets are defined as those that fulfil $p_T > 20$ GeV and the other previous requirements.

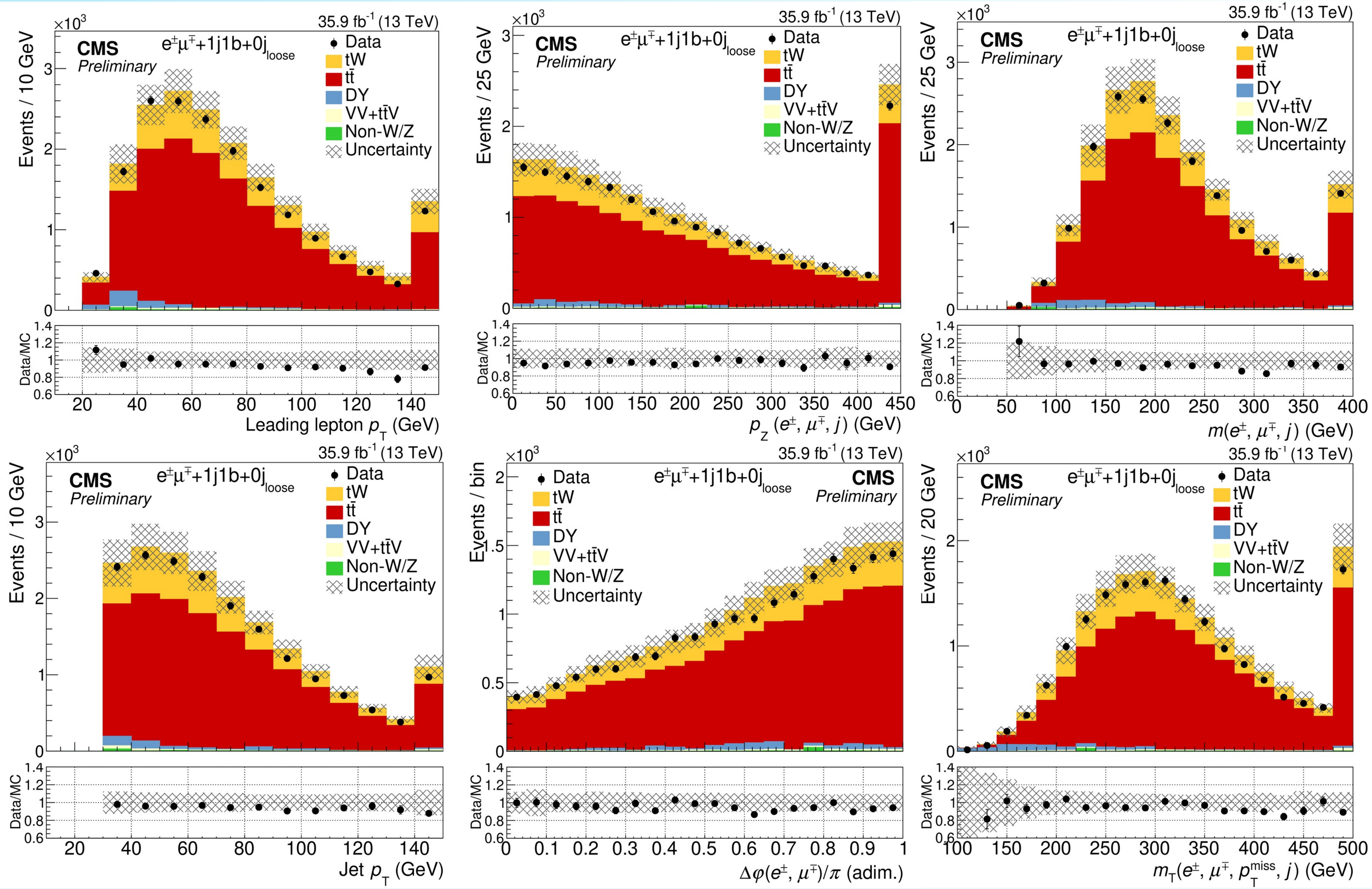
Event selection

- At least two identified leptons.
- One of them must fulfil $p_T > 25$ GeV.
- The two first (in terms of p_T) must have opposite charge…
- … and be an electron and a muon ($e\mu$ channel)…
- …whose invariant mass satisfy $m(e, \mu) > 20$ GeV.
- Exactly 1 jet that must be b-tagged.
- Exactly zero loose jets.





Event selection - Monte Carlo/data comparison



Signal extraction and unfolding

- Signal extraction is done by subtracting the background (as estimated from MC) from the data.
- Unfolding is implemented using `TUnfold` (JINST 7 (2012) T10003), allowing us to do the "jump" from detector to particle level. The fiducial region is defined as follows over particle level objects.
  - At least two leptons with $p_T > 20$ GeV and...
    - ...$|\eta| < 2.4$ (muons).
    - ...$|\eta| < 1.4442$ or $|\eta| > 1.566$ and always $|\eta| < 2.4$ (electrons)
  - One of them, with at least $p_T > 25$ GeV.
  - The two leptons with largest p_T must be an electron and a muon...
  - ...with opposite charge.
  - In addition, their invariant mass must be over 20 GeV.
  - There must be only one jet (with $p_T > 30$ GeV and separated with a cone of at least $\Delta R > 0.4$ of any of the selected leptons with the previous properties ($p_T > 20$ GeV and the η requirements)...
  - ...and it must come from a b quark.
  - No other jet with $p_T > 20$ GeV is allowed to be present in the event.

Uncertainties – I

Experimental

- **Jet energy scale and resolution**: varying both (all sources together) within its p_T and η bin uncertainties.
- **B–tagging** and **mistagging**: varying the data-to-simulations SF by their uncertainties.
- **Pile–up**: varying ±4.6% the p-p inelastic cross section.
- **Lepton ID** and **trigger**: varying the data-to-simulations SF by their uncertainties.
- **Luminosity**: 2.5% flat.

Background normalisation of MC samples

- $t\bar{t}$: 4%, from Eur. Phys. J. C. 79 (2019) 368 .
- **DY**: 50% from *.
- **Non-W/Z**: 50% from *.
- **VV+$t\bar{t}$ V**: 50% from *.

* 10.1007/JHEP10(2018)117

Statistical

- Size of data samples.
- Size of simulation samples.
- Size of the simulation samples used for obtaining the response matrices.

Uncertainties – II

<u>Modeling-related</u>

- **Underlying event** ($t\bar{t}$): using dedicated samples that vary the Pythia parameters that tune the measurements to the UE (details in the backup; CMS-PAS-TOP-16-021).
- **ME/PS matching** ($t\bar{t}$): using dedicated samples that vary the Powheg h_{damp}=1.58·m_t parameter by its own uncertainty (+0.66, -0.59; CMS-PAS-TOP-16-021).
- **Initial and final states radiation** (tW, $t\bar{t}$): using dedicated samples that vary the PS scale by a factor of two.
- Comparison with the **DS schema** (tW): using dedicated samples, we take the difference w.r.t. nominal (i.e. DR) values.
- **Colour reconnection** ($t\bar{t}$): using various different models, we take the envelope of their variations for each bin as its uncertainty.
- **Factorisation/renormalisation scales** (tW, $t\bar{t}$): we take the difference w.r.t. scaling μ_F and μ_R by 2 and 0.5 relative to their common nominal value.
- **PDF** ($t\bar{t}$): we take the root-mean-square of the difference between nominal values with the 100 NNPDF3.0 replicas as an uncertainty.
- **m_{top}** (tW, $t\bar{t}$): using varied samples obtained with modified m_{top} values by ±1 GeV.
- **Top quark p_T mismodeling** ($t\bar{t}$): as it is applied to the nominal background estimation of the $t\bar{t}$ sample, it is considered by taking the difference with the values w/o applying it. Phys. Rev. D 95, 092001, EPJC 75 542 (2015), EPJC 76 128 (2016), CMS-PAS-TOP-16-011.

Results – Considerations

- For all variables, **no regularisation nor area constraint was needed**, as the response matrices are mostly diagonal. Thus, direct response matrix (R) inversion is done, after signal is extracted by subtracting the background:

$$N_i^{\text{sig}} = N_i - N_i^{\text{bkg}} = \sum_{j=1} R_{ij} N_j^{\text{sig, unf}},$$

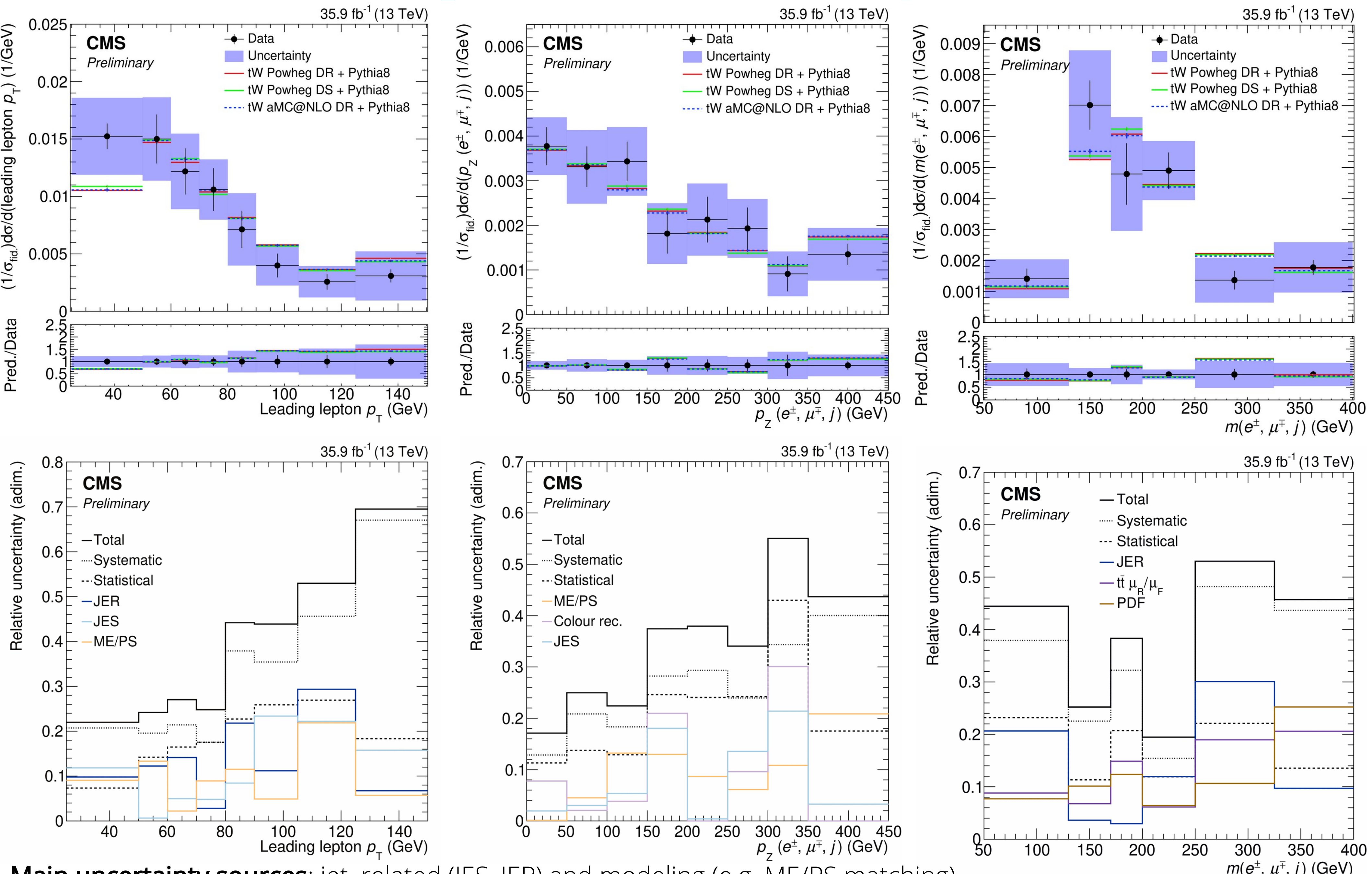
- The results are calculated obtaining first (where Δ_j is the j bin width)

$$\left(\frac{d\sigma}{dX}\right)_j = \frac{1}{\mathcal{L}} \frac{N_j^{\text{sig, unf}}}{\Delta_j}$$ *(for each particle level bin **j**)*

  and then dividing those values by the fiducial cross section.

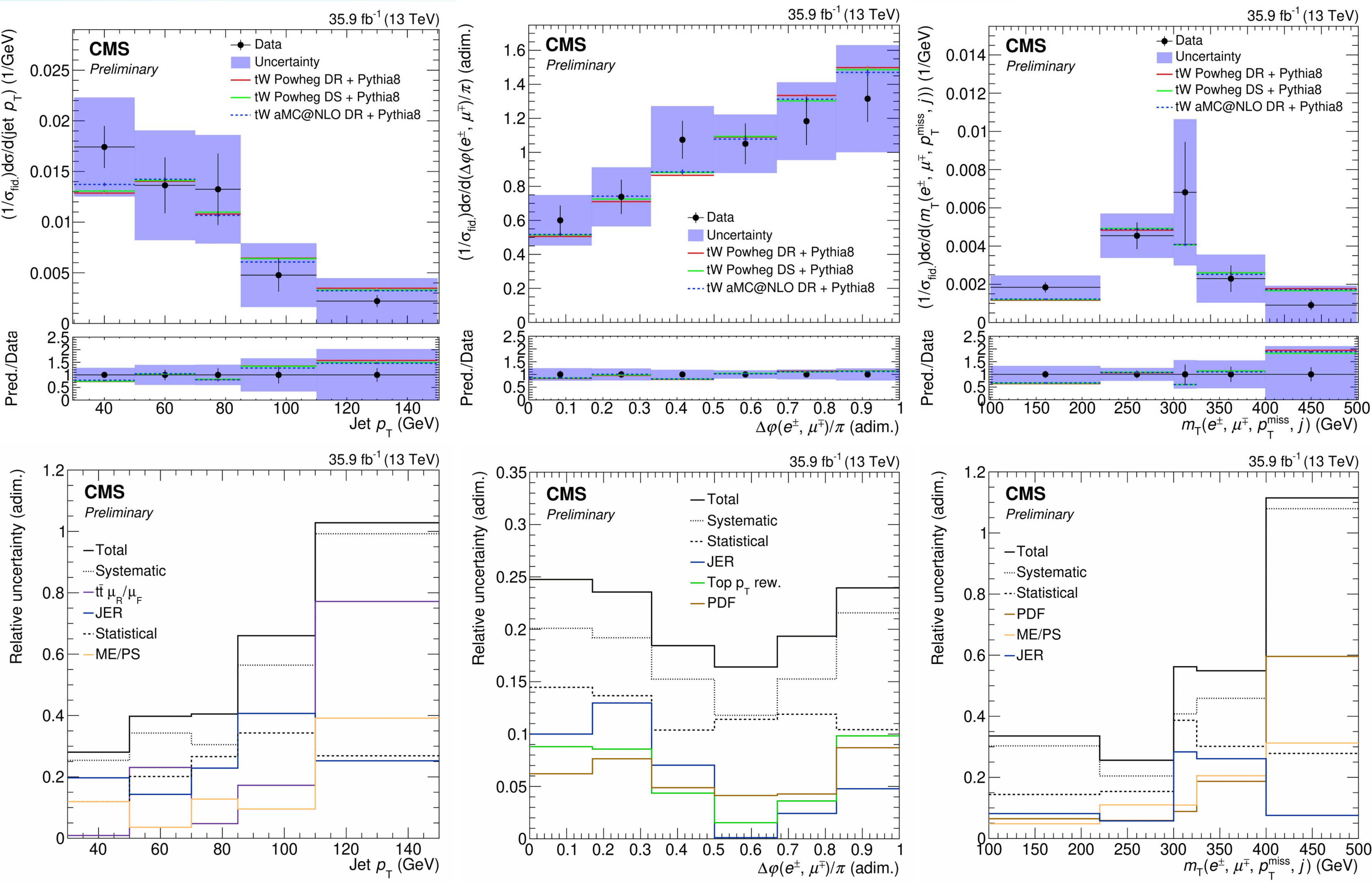
- All systematic uncertainties are **propagated** through unfolding by performing a separated deconvolution for each variation and afterwards taking the difference with respect to the nominal values.
  - Special attention is payed to correlations when obtaining the results.
  - Uncertainties can affect the background, but also the response matrix: in those cases, a varied response matrix is calculated for considering them.
  - They are shown symmetrically, by taking the average of the two variations.

Results - I



Main uncertainty sources: jet–related (JES, JER) and modeling (e.g. ME/PS matching).

Results - II


35.9 fb^{-1} (13 TeV)
CMS
Preliminary
Data
Uncertainty
tW Powheg DR + Pythia8
tW Powheg DS + Pythia8
tW aMC@NLO DR + Pythia8
$(1/\sigma_{fid.})d\sigma/d(\text{jet } p_T)$ (1/GeV)
Pred./Data
Jet p_T (GeV)
$(1/\sigma_{fid.})d\sigma/d(\Delta\varphi(e^{\pm}, \mu^{\mp})/\pi)$ (adim.)
$\Delta\varphi(e^{\pm}, \mu^{\mp})/\pi$ (adim.)
$(1/\sigma_{fid.})d\sigma/d(m_T(e^{\pm}, \mu^{\mp}, p_T^{miss}, j))$ (1/GeV)
$m_T(e^{\pm}, \mu^{\mp}, p_T^{miss}, j)$ (GeV)
Relative uncertainty (adim.)
Total
Systematic
$t\bar{t}$ μ_R/μ_F
JER
Statistical
ME/PS
Top p_T rew.
PDF

Summary

- A measurement of the differential cross section in the tW process using CMS 2016 data has been made depending on six observables of the events (**CMS-PAS-TOP-19-003** , **twiki**).
  - **Agreement** between data and expectations is **fairly good**.

- Analysis **largely dominated by systematic sources of uncertainties**, whose main origin is the overwhelming $t\bar{t}$ **background**.
  - Main sources: jet–related uncs. (e.g. JES, JER) and modeling (e.g. ME/PS matching).
  - Depending on the bin and distribution, varying from ~15-40% (bulk of distributions) up to ~25-100% in the tails.

- In addition, the result shows **compatible agreement** for the DR and DS schemes of the signal process.

Thanks for your attention

4th Red LHC Workshop

Backup slides

Other differential measurements

- **7 TeV**
  - (ATLAS Collab.) *"Comprehensive measurements of **t-channel** single top-quark production cross sections at √s = 7 TeV with the ATLAS detector". Phys. Rev. D. 90, 112006 (2014)*, arXiv:1406.7844.

- **8 TeV**
  - (CMS Collab.) *"Measurement of top quark polarisation in **t-channel** single top quark production". JHEP 04 (2016) 073*, arXiv:1511.02138.
  - (ATLAS Collab.) *"Fiducial, total and differential cross-section measurements of **t-channel** single top-quark production in pp collisions at 8 TeV using data collected by the ATLAS detector". Eur. Phys. J. C 77 (2017) 531*, arXiv:1702.02859.

Underlying event uncertainty source details

- The following table shows the nominal values for the parameters taken as nominal for the underlying event properties of the simulations. The "up" variations are shown right and the "down", left.

	CUETP8M2T4		
Tune		pp 14	
Tune		ee 7	
MultipartonInteractions ecmPow		0.2521	
SpaceShower:alphaSvalue		0.1108	
PDF pSet LHAPDF6		NNPDF30_lo_as_0130	
MultipartonInteractions:pT0Ref	2.13	2.20	2.26
MultipartonInteractions:expPow	1.71	1.6	1.56
ColourReconnection:range	6.5	6.6	8.7

CMS-PAS-TOP-16-021