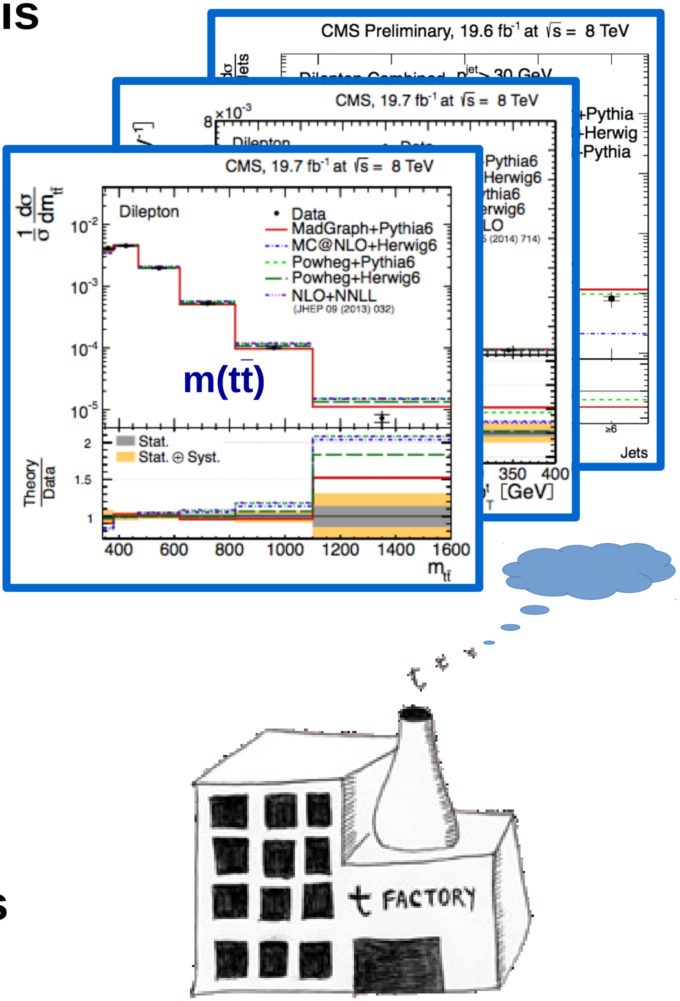
Differential cross sections for $t\bar{t}$ production at 8 and 13 TeV by CMS

Mykola Savitskyi
(DESY)

Why measure differentially?

- **Precise understanding of top quark distributions is crucial:**
  - Precision tests of perturbative QCD for top quark production at different phase space regions
  - Theory predictions and models need to be tuned and tested with measurements:
    → potential to reduce signal modelling systematics
  - Extract/use for PDF fits
  - Enhance sensitivity to New Physics
  - Background for Higgs, rare processes and many BSM searches

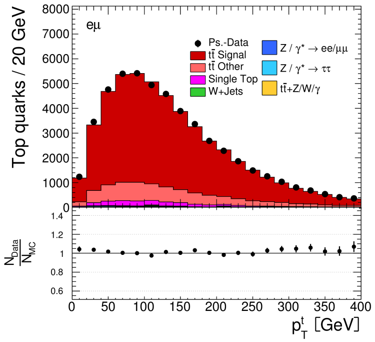


- **Large $t\bar{t}$ samples at the LHC allow measuring $\sigma(t\bar{t})$ as a function of many kinematic observables**

General analysis strategy

In this talk latest selected results by CMS:
at **8 TeV** → [arXiv:1505.04480] (dilepton & l+jets)
at **13 TeV** → PAS TOP-15-010 (dilepton)

- **Goal**: measure σ($t\bar{t}$) as a function of top quark, $t\bar{t}$ system, b-jet, lepton, lepton pair and event-level observables

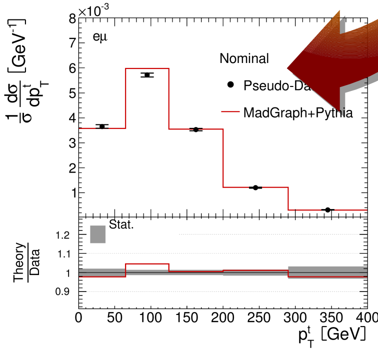


- **Main analysis ingredients**
  - > Event selection
  - > $t\bar{t}$ kinematic reconstruction
  - > Bin-wise cross section measurement
  - > Unfolding: correct for detector effects & acceptance to parton or particle level after background subtraction

Δ^i_X = bin width for variable X

$$\frac{1}{\sigma}\frac{d\sigma}{dX_i}=\frac{1}{\sigma}\frac{unfold\left(N^X_{data,i}-N^X_{BG,i}\right)}{\Delta^i_X\cdot\int\mathscr{L}\,dt}$$

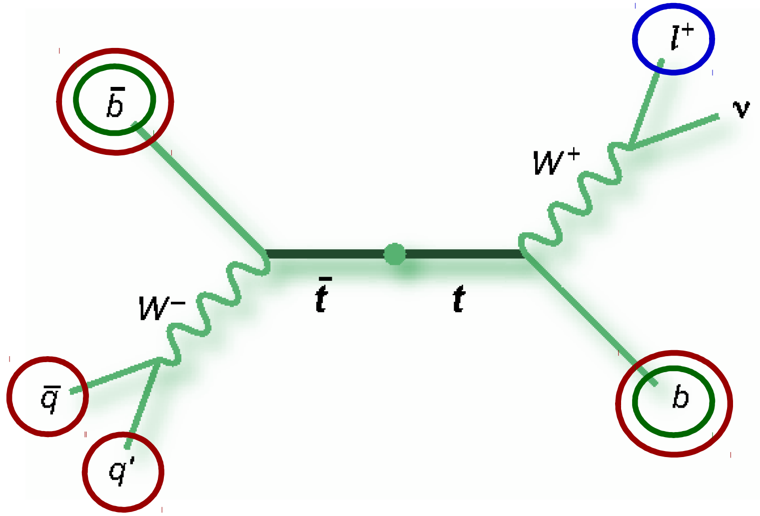
- **Differential $t\bar{t}$ cross sections**
  - > Normalize to in-situ measured σ($t\bar{t}$): mostly shape uncertainties contribute



Event selection

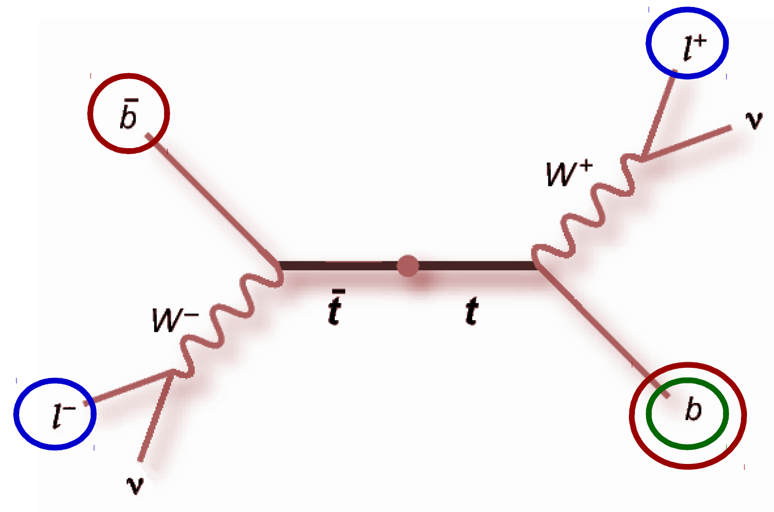
Lepton+jets:

- **Exactly 1 high-p_T isolated lepton (e or μ)**
  - $p_T > 33$ GeV, $|\eta| < 2.1$
- **≥ 4 jets**: $p_T > 30$ GeV, $|\eta| < 2.4$
- **≥ 2 b-tagged jets**



Dileptons:

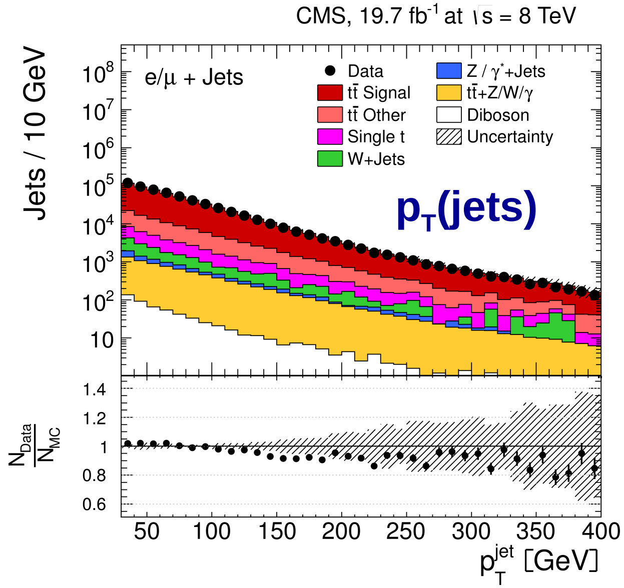
- **> 2 OS, high-p_T isolated leptons (ee, $\mu\mu$, μe)**
  - $p_T > 20$ GeV, $|\eta| < 2.4$
- **QCD veto**: $m_{ll} > 20$ GeV
- **≥ 2 jets**: $p_T > 30$ GeV, $|\eta| < 2.4$
- **≥ 1 b-tagged jets**
- **ee, $\mu\mu$ channels**: $E_T^{miss} > 40$ GeV
  Z veto: $|m_Z - m_{ll}| > 15$ GeV

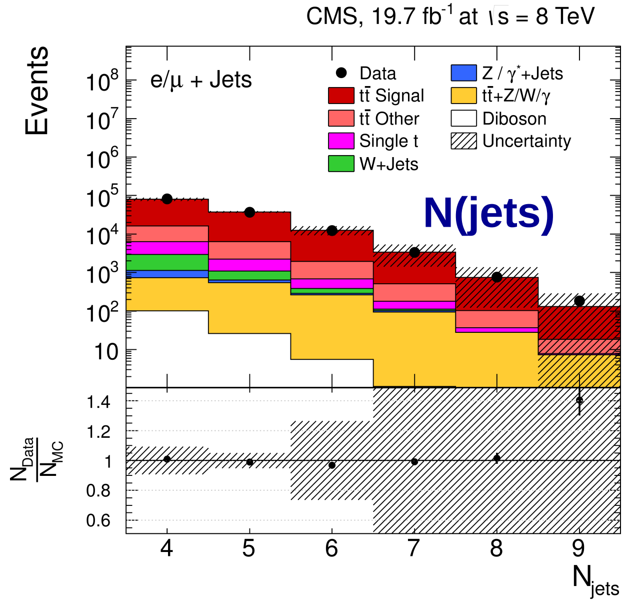


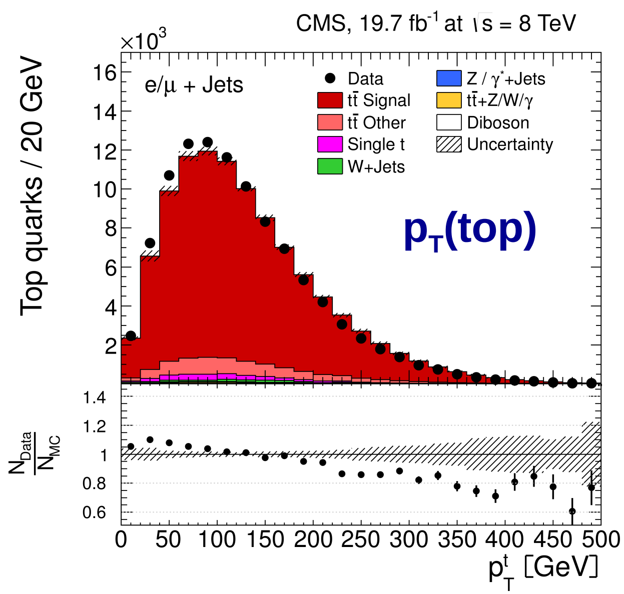
In addition: kinematic reconstruction of $t\bar{t}$ system

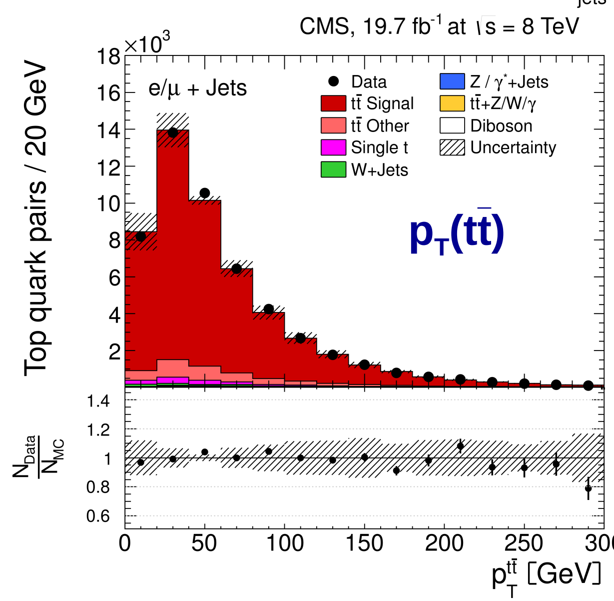
Kinematic distributions in l+jets at 8 TeV

- Very pure **$t\bar{t}$ signal** after full event selection & kinematic reconstruction (~80%)
- Dominant backgrounds: **$t\bar{t}$ other**, **single top**, **W+jets**
- **$t\bar{t}$ other** includes all non-l+jets decays
- Reference $t\bar{t}$ prediction: **MadGraph+Pythia6**
- Softer top p_T spectrum in data than one determined from simulation (same for dilepton)



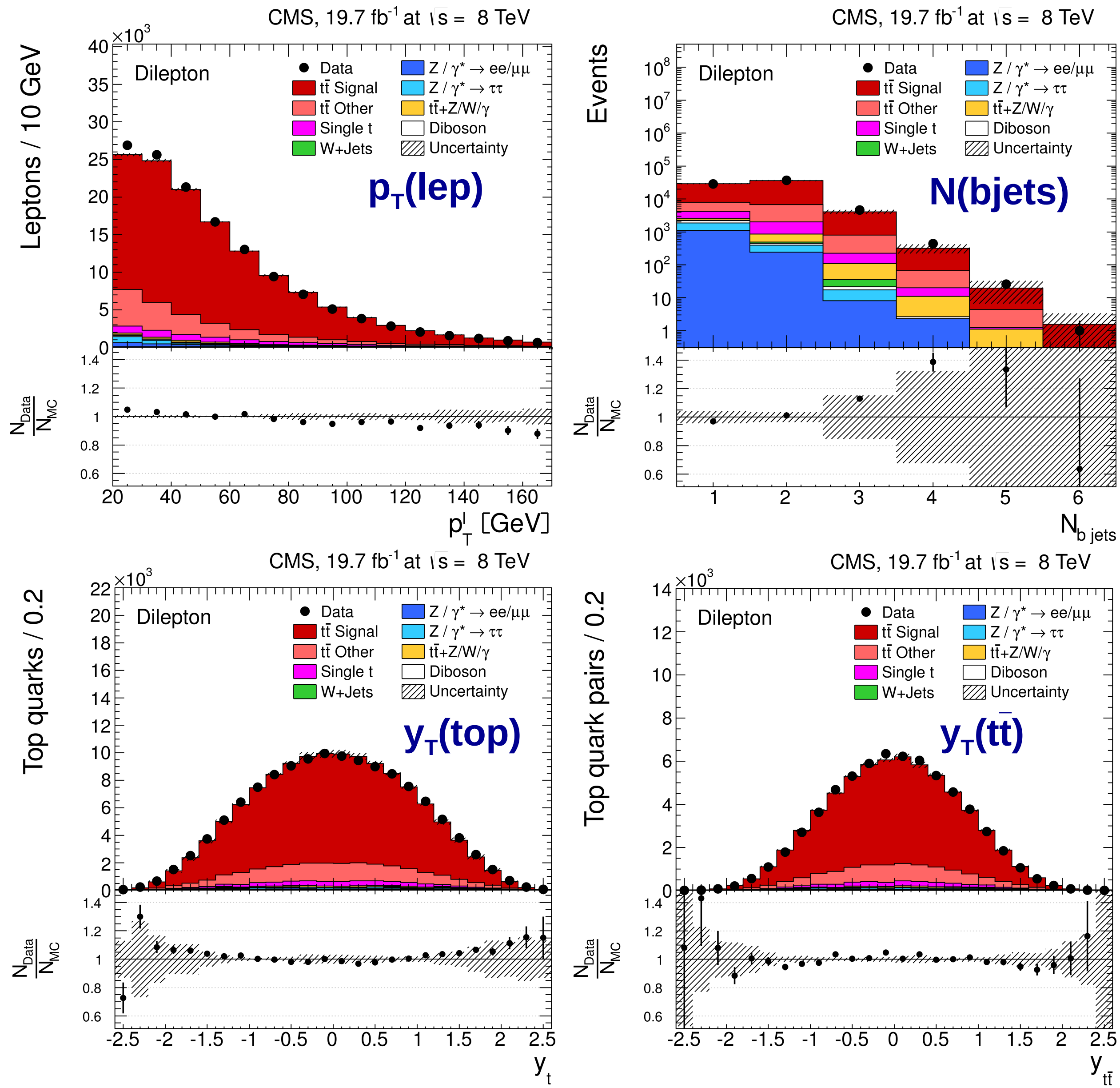






Kinematic distributions in dileptons at 8 TeV

- Very pure $t\bar{t}$ **signal** after full event selection & kinematic reconstruction (~80%)
- Dominant backgrounds: **$t\bar{t}$ other**, **single top**, **Z+jets**
- **$t\bar{t}$ other** includes all non-dilepton decays
- Reference $t\bar{t}$ prediction: **MadGraph+Pythia6**
- Lepton and jet p_T spectra feature similar behavior as in top p_T (same for l+jets)

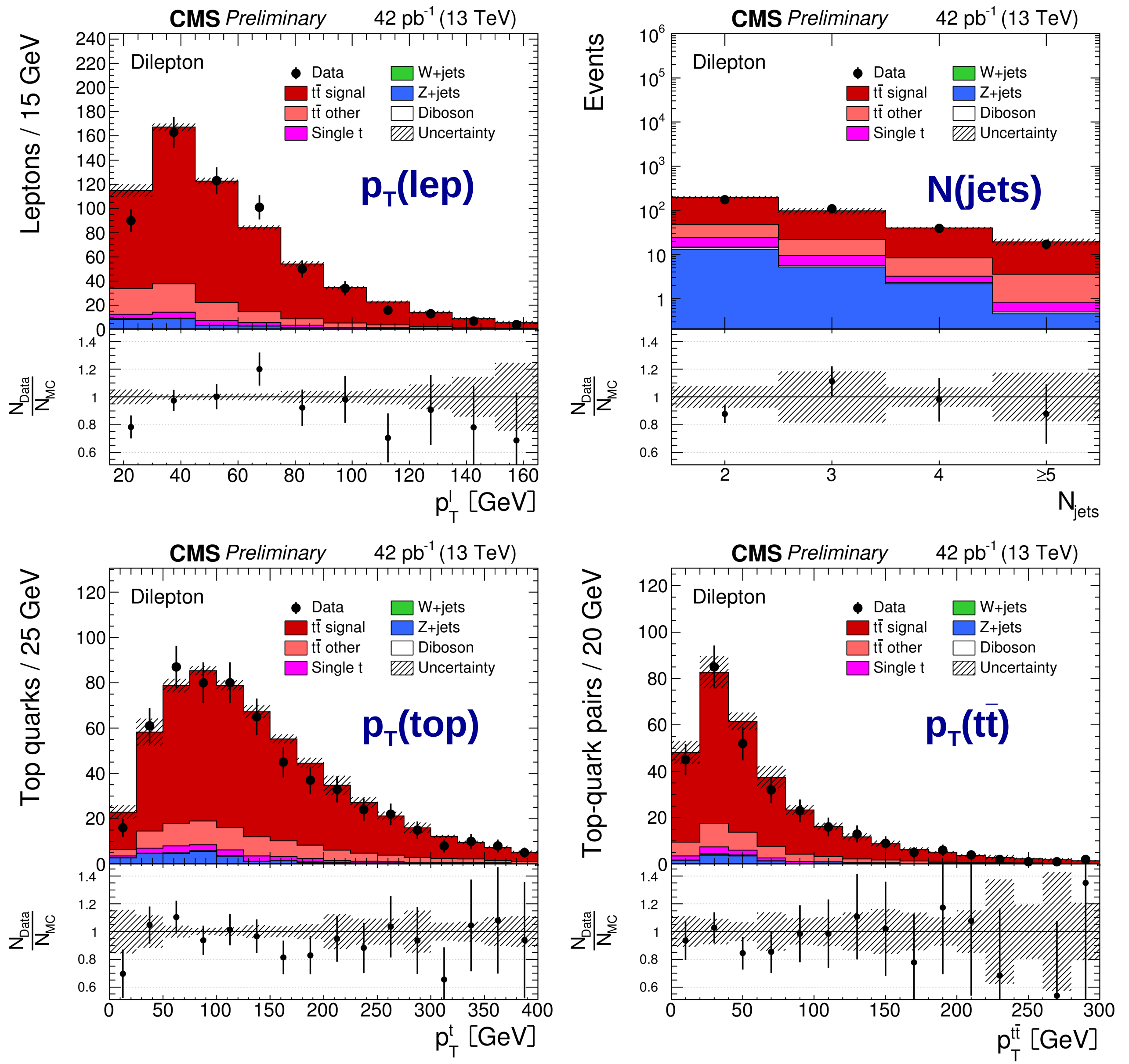


Kinematic distributions in dileptons at 13 TeV

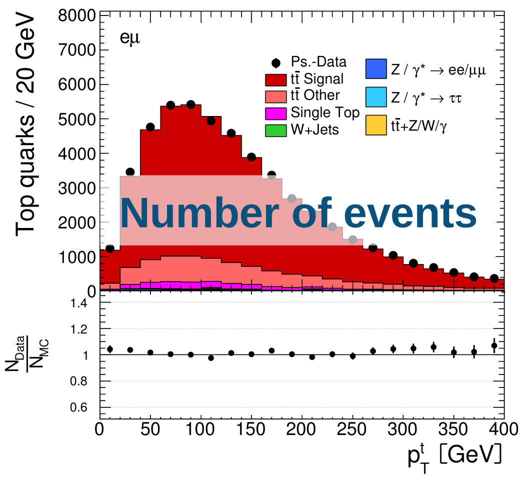
Using first L = 42 pb^{-1} of 13 TeV data

- Very pure **$t\bar{t}$ signal** after full event selection & kinematic reconstruction (~80%)
- Dominant backgrounds: **$t\bar{t}$ other**, **single top**, **Z+jets**
- **$t\bar{t}$ other** includes all non-dilepton decays
- Reference $t\bar{t}$ prediction: **Powheg+Pythia8**
- Dominated by statistical uncertainty

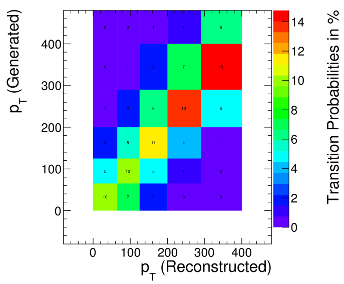
In general, good data-to-MC agreement

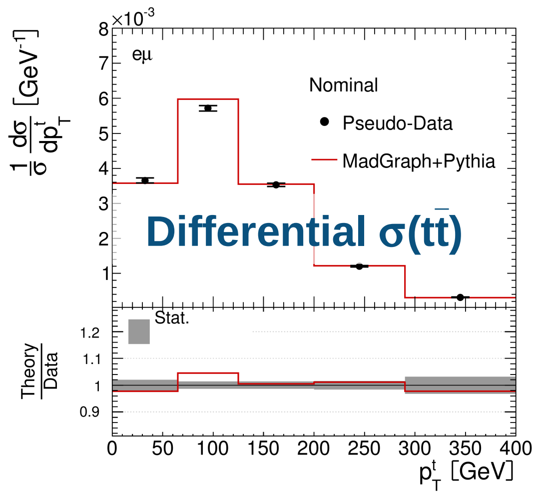


Normalized differential cross section



Response matrix A_{ij}





$$\frac{1}{\sigma}\frac{d\sigma}{dX_i}=\frac{1}{\sigma}\frac{unfold\left(N^X_{data,i}-N^X_{BG,i}\right)}{\Delta^i_X\cdot\int\mathscr{L}\,dt}$$

Binning

Chosen to limit migration effects in and out of bins:

- purity (p_i) & stability (s_i): ≥ 50%
- ≈ flat in all bins

$$p_i=\frac{N_i^{rec\,\&\,gen}}{N_i^{rec}} \qquad s_i=\frac{N_i^{rec\,\&\,gen}}{N_i^{gen}}$$

Regularized unfolding

- Basic unfolding - simple inversion of response matrix A_{ij}:

$$N_{i,unf}=A^{-1}_{ij}N_{j,measured}$$

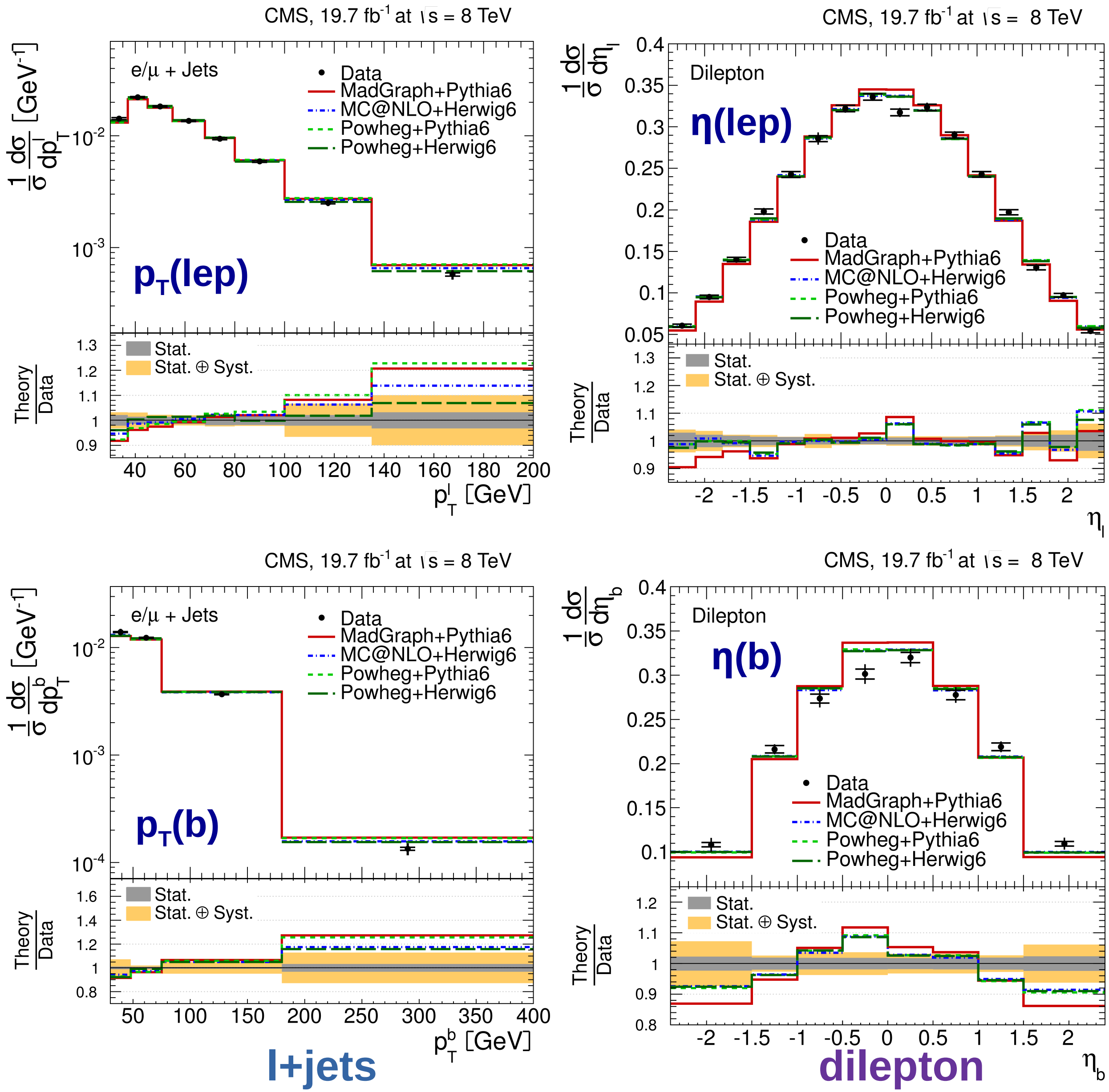
- Regularization used to remove large statistical fluctuations (SVD)

Phase space

- Correct back to parton or particle level in full or fiducial phase space
- Top quark definition: *before decay and after QCD radiation*
- Fiducial phase space: closely follows event selection

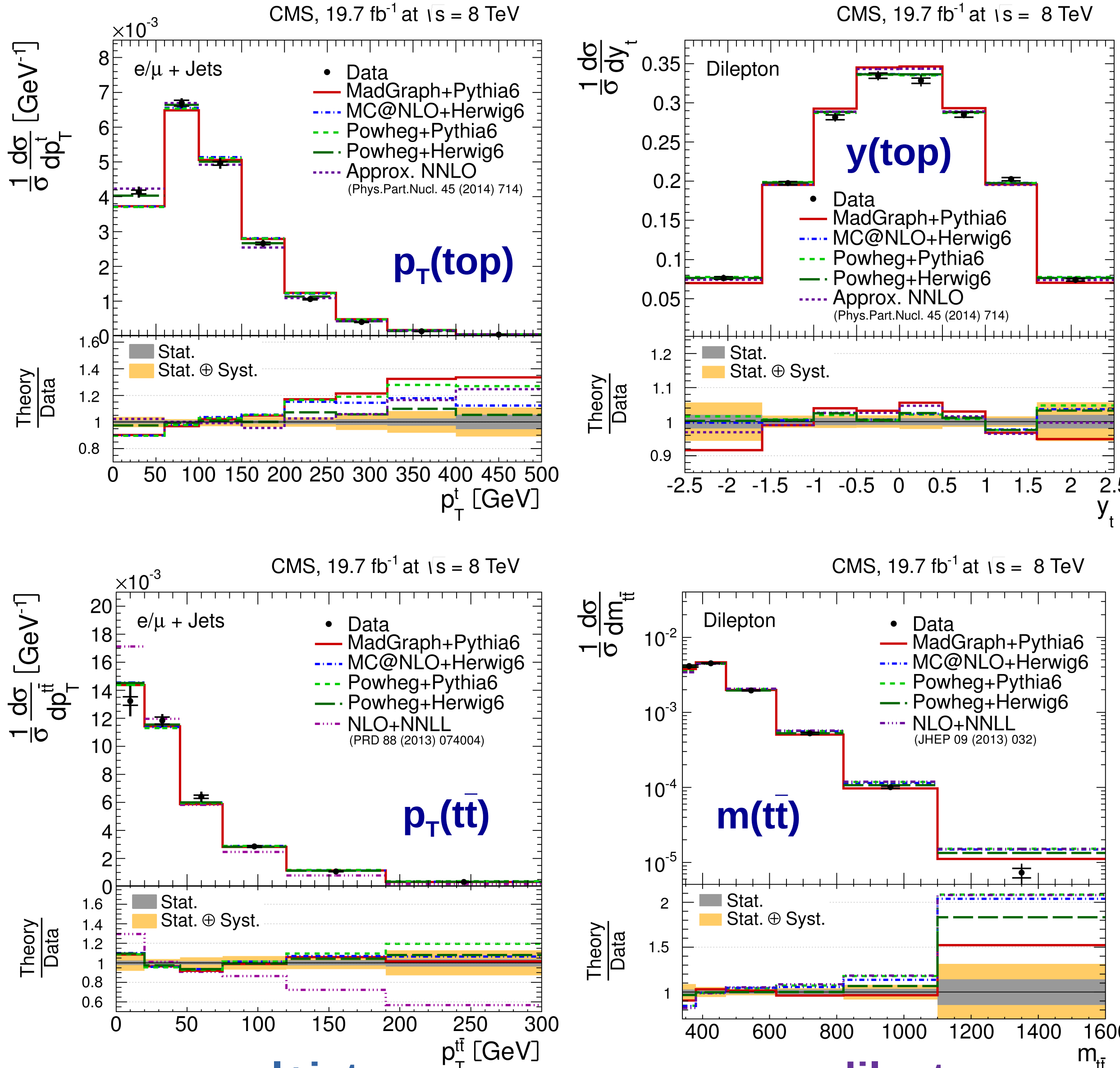
8 TeV results: leptons & b-jets

- Fiducial phase space, particle level
- Reference $t\bar{t}$ prediction used for unfolding: **MadGraph+Pythia6**
- Slightly softer p_T spectra and less centered η distributions in data
- Good agreement with data in all distributions: **Powheg+Herwig6**
- Consistent with 7 TeV results by CMS: [EPJ C73 (2013) 2339]



8 TeV results: top and $t\bar{t}$

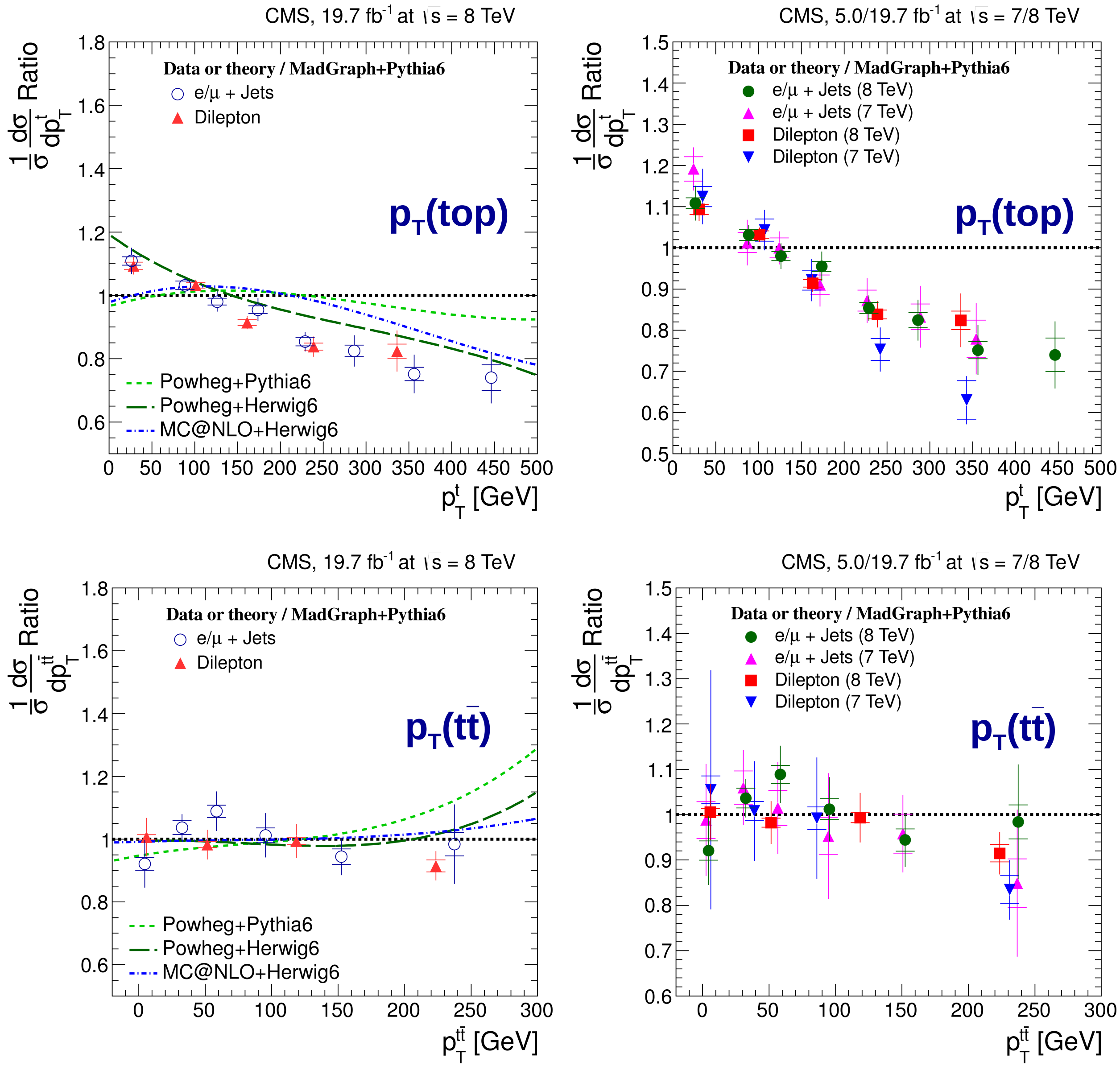
- Full phase space, parton level
- Reference $t\bar{t}$ prediction used for unfolding: **MadGraph+Pythia6**
- Best description of data by **Powheg+Herwig6**
- **p_T(top)**: softer in data
- **y(top)**: less central in data
- **$p_T(t\bar{t})$**: in agreement with all predictions, except NLO+NNLL calculations
- **$m(t\bar{t})$**: softer in data



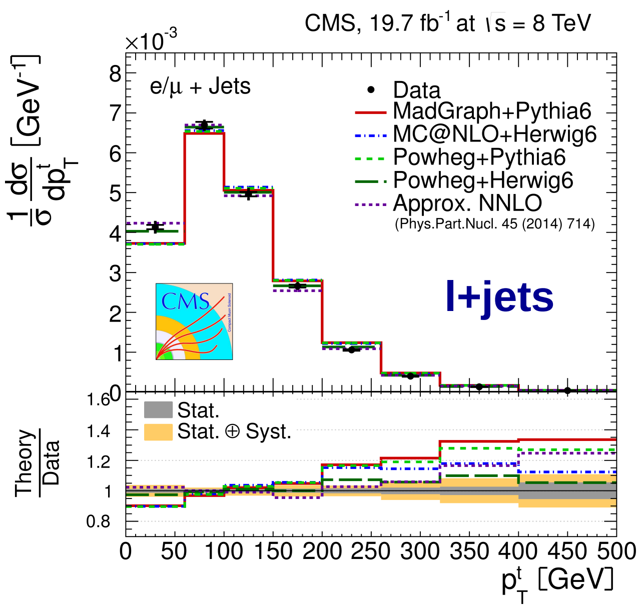
l+jets   dilepton

Results consistency: p_T(top), $p_T(t\bar{t})$

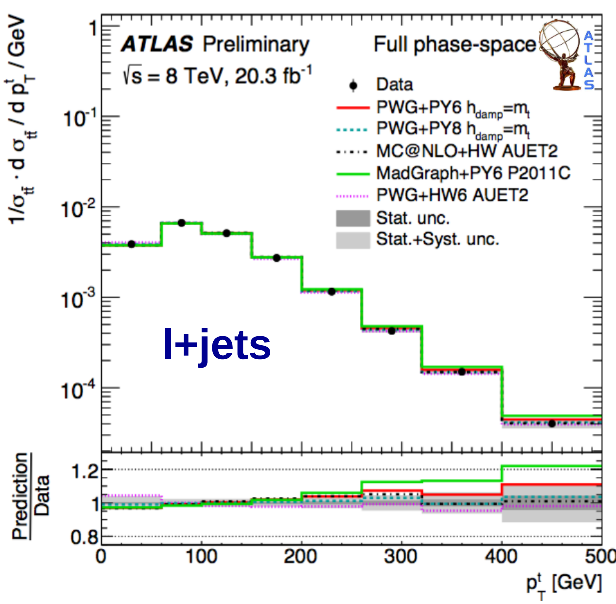
- **Left plots**: data vs theory predictions at 8 TeV
- **Right plots**: 7 TeV vs 8 TeV – results consistent
- All values relative to MadGraph+Pythia6
- Results consistent among all decay channels



The p_T(top) distribution at 8TeV



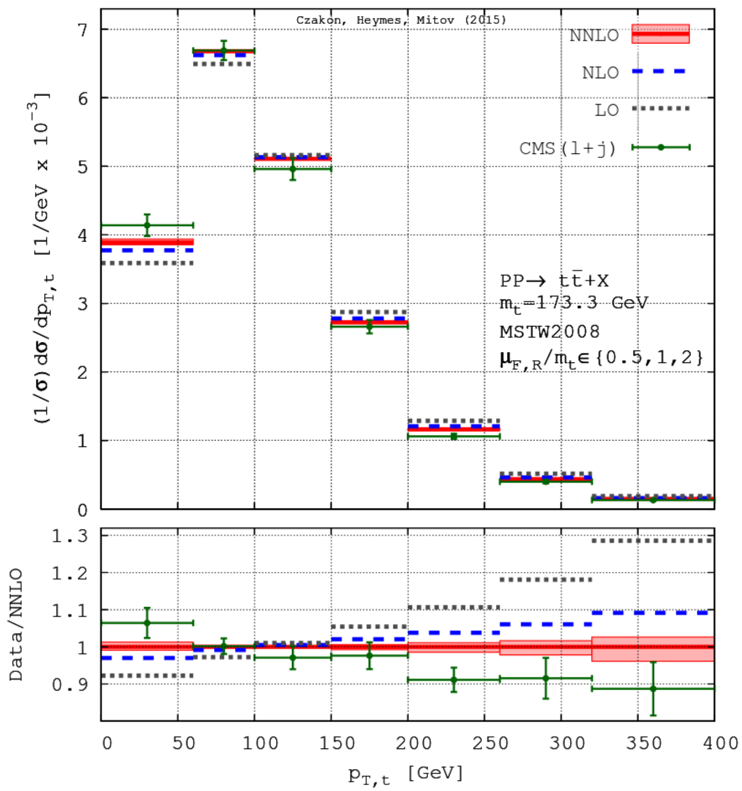
[arXiv:1505.04480], accepted by EPJC



TOPQ-2015-06

- p_T(top) spectrum softer in data (in particular at the tail):
  → potential impact on searches and $t\bar{t}$+H
- CMS: observed consistently in all channels at 7 & 8 TeV
- **ATLAS and CMS data appear in good agreement at 8 TeV**

Full **NNLO** calculation available!



Full NNLO “confirms” observed slope, in direction closer to data

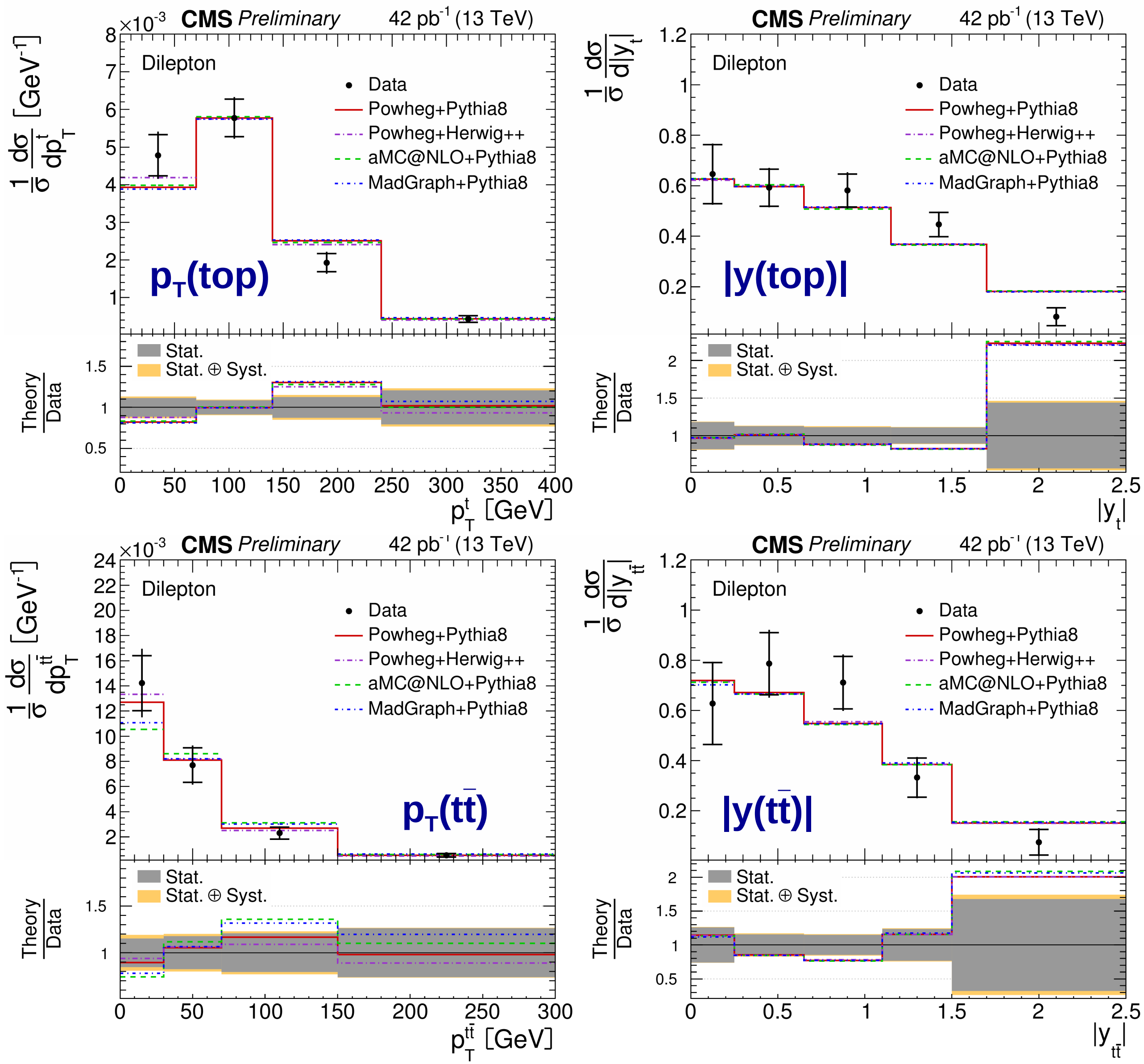
[arXiv:1511.00549]
M. Czakon, D. Heymes, A. Mitov

13 TeV results: p_T(top), |y(top)|, $p_T(t\bar{t})$, $|y(t\bar{t})|$

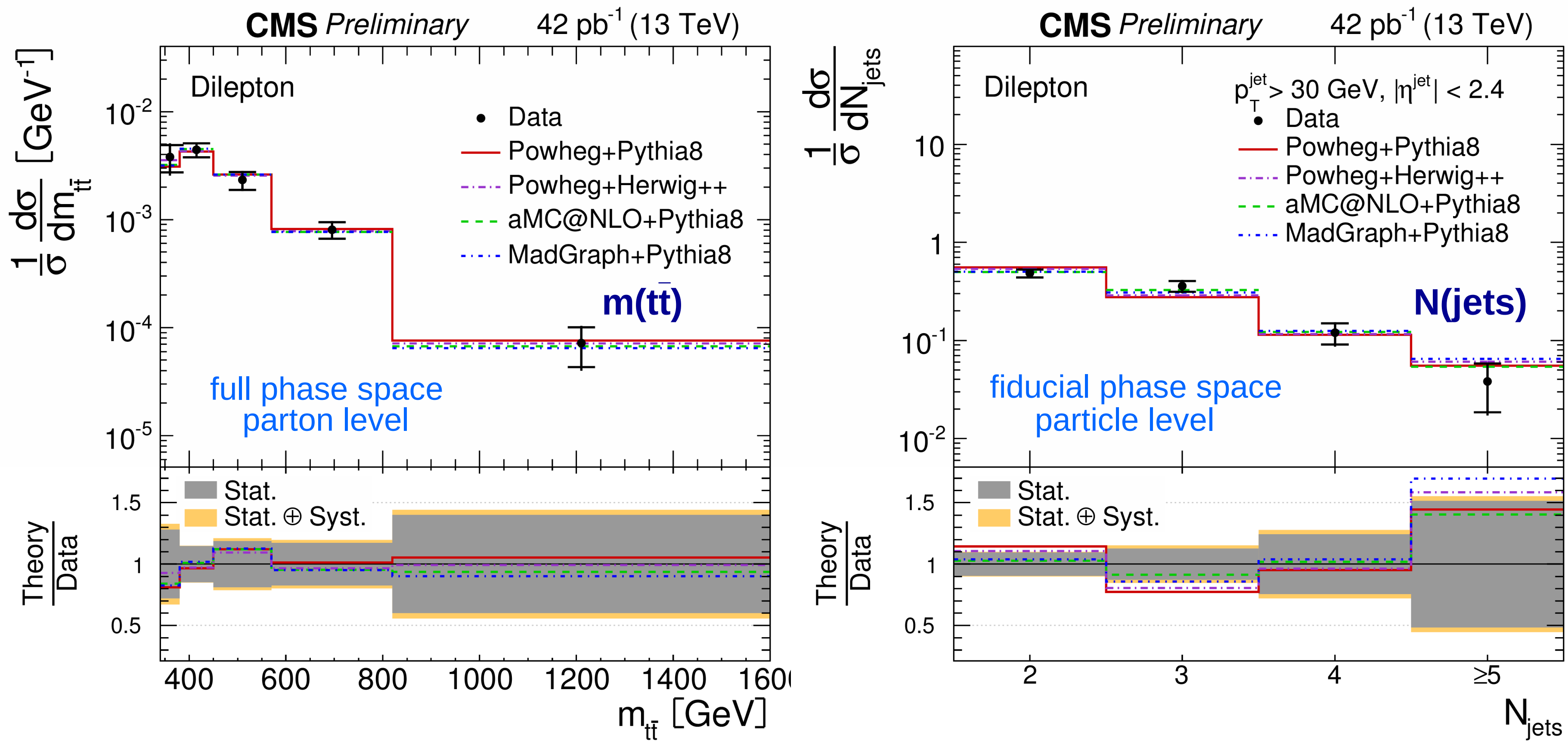
Using first L = 42 pb^{-1} of 13 TeV data

Dilepton channel

- Full phase space, parton level
- Reference $t\bar{t}$ prediction used for unfolding: **Powheg+Pythia8**
- Reasonable agreement between data and predictions
- Dominated by statistical uncertainty



13 TeV results: m($t\bar{t}$), N(jets)



- Reference $t\bar{t}$ prediction used for unfolding: **Powheg+Pythia8**
- Reasonable agreement between data and predictions
- Dominated by statistical uncertainty

Summary

Top quark pair differential cross section measurements:

- Essential for constraining the SM
- Ideal probe for looking for new physics beyond the SM

Latest dilepton and l+jets 8 TeV results from CMS (L = 19.7 fb^{-1}):

- Measurement dominated by systematical uncertainty: 3-10% precision
- Good agreement between data and predictions
- p_T(top): NNLO corrections bring SM predictions closer to data

Latest dilepton 13 TeV results from CMS (L = 42 pb^{-1}):

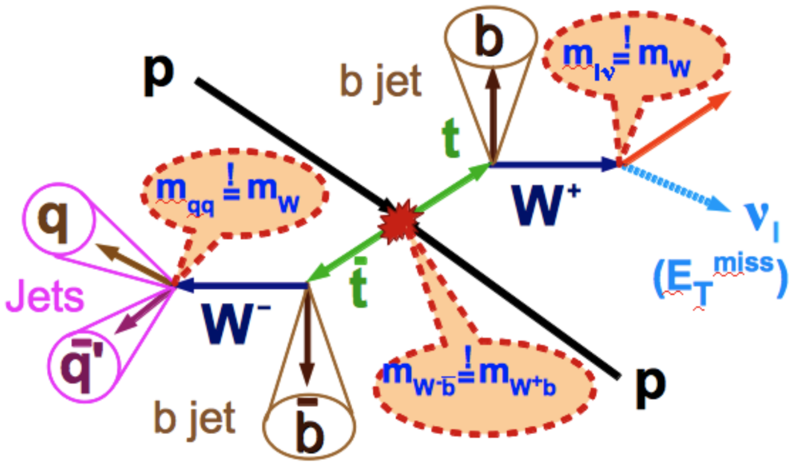
- Measurement dominated by statistical uncertainty
- In general, data described reasonably well by all MC predictions

Other recent differential cross section results by CMS:
8 TeV → PAS TOP-14-012 (l+jets: boosted topologies), [arXiv:1509.06076] (all-jets)
13 TeV → PAS TOP-15-005 (l+jets), PAS TOP-15-013 (l+jets: global event variables)

Backup

Kinematic reconstruction of $t\bar{t}$ in l+jets

- Vary 4-momenta of leptons, jets & neutrino within resolutions
- Constraints:
  - $m_{top} = m_{antitop}$
  - $m_{qq} = m_{l\nu} = m_W = 80.4$ GeV
- Limit permutations: consider 4/5 leading jets, use b-tag information
- Take 4-jet permutation with minimum χ^2
- "Trick":
  - first fit with $m_{top} = 172.5$ GeV → select best permutation
  - m_{top} free + fixed jet permutation → obtain kinematics for differential measurements
- Cut on χ^2 probability > 2% → increase correct jet permutations and signal purity

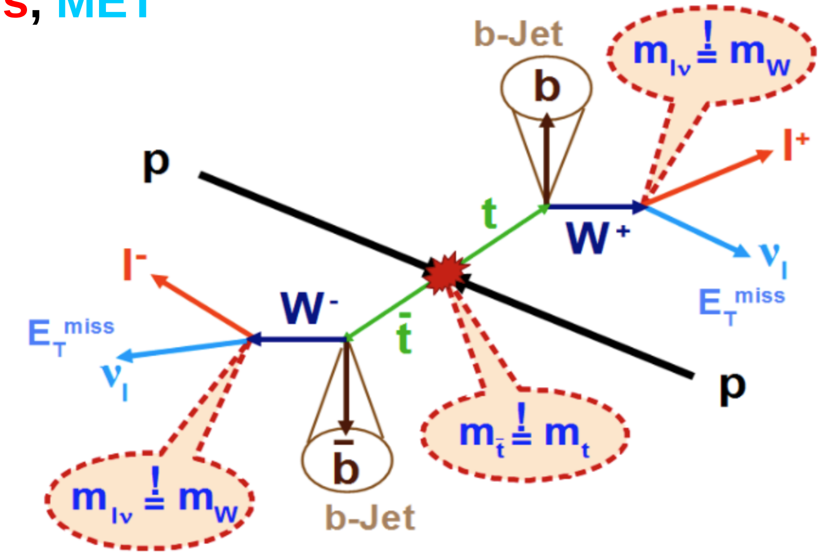


Kinematic reconstruction of $t\bar{t}$ in dileptons

- **Measured input: 2 jets, 2 leptons, MET**
- **Unknowns:** $\vec{p}_\nu, \vec{p}_{\bar{\nu}} \rightarrow 6$
- **Constraints:**
  > $m_t, m_{\bar{t}} \rightarrow 2$

  > $m_{W(+)}, m_{W(-)} \rightarrow 2$

  > $(\vec{p}_\nu + \vec{p}_{\bar{\nu}})_T = MET \rightarrow 2$



- Reconstructing each event 100 times and smearing inputs by their resolution:
  > top mass fixed to 172.5 GeV

  > W mass at RECO level smeared accordingly to W mass distribution

  > Jet and lepton energies are corrected for detector effects
- Consider weighted average of solutions for all smeared points:

$$p^{top}_{x,y,z} = \frac{1}{w} \sum_{i=0}^{100} w_i \cdot \left(p^{top}_{x,y,z}\right)_i$$

Phase space definitions

- **Top quarks and $t\bar{t}$ observables**: presented at **parton level, extrapolated to full phase space**
  > Allows for comparison with available highest order QCD calculations
  > Consistent top quark definition in ATLAS & CMS: *before decay and after QCD radiation*

- **Leptons, jets and b-jets**: presented at **particle level, fiducial phase space**

  - **Object definition** at generator level: based on stable particles after radiation and hadronization
    > Leptons: from W decay
    > Jets: anti-kT algorithm (as for reco jets), cluster all but prompt particles
    > b-jets: matched to the original b quark from top

  - **Phase space definition** closely follows the (detector level) event selection. In example, for dilepton channel:
    > 2 leptons, $p_T > 20$ GeV, $|\eta| < 2.4$
    > 2 b-jets from top, $p_T > 30$ GeV, $|\eta| < 2.4$
    > (if any) additional jets, $p_T > 30$ GeV, $|\eta| < 2.4$

Unfolding

- **Unfolding** techniques correct migrations between bins
- **Response matrix (A):** represents bin-by-bin correlations
- Unfolding problem is transformed to χ^2 - minimization problem:

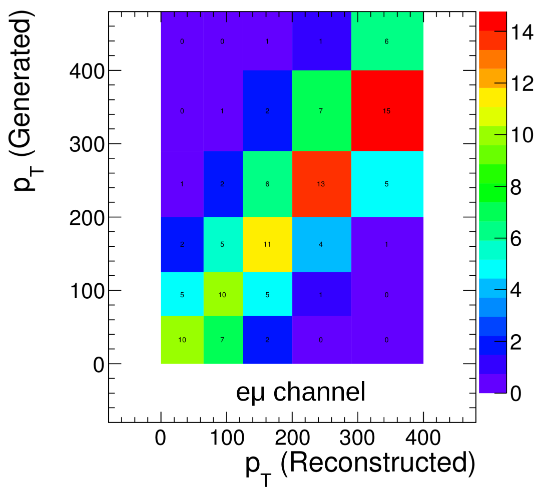
unfolding — regularization

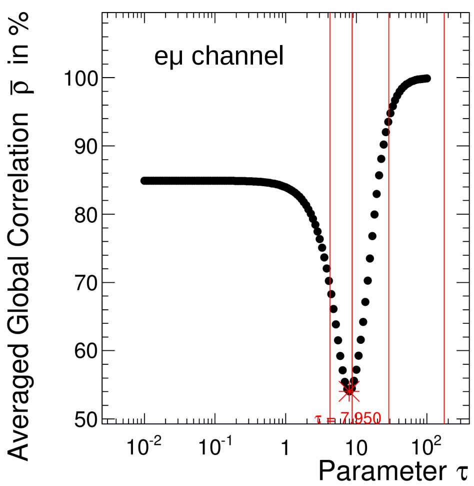
$$\chi^2 = \left(\vec{N} - A\cdot\vec{x}\right)^T COV_{\vec{N}}^{-1} \left(\vec{N} - A\cdot\vec{x}\right) - \tau^2 \cdot K\left(\vec{x}\right)$$

- **N:** BG corrected data
- **x:** unfolded result

- Non-physical fluctuations removed by means of the regularization:
  - > τ – continuous regularization parameter
  - > selected at minimum of average global correlation

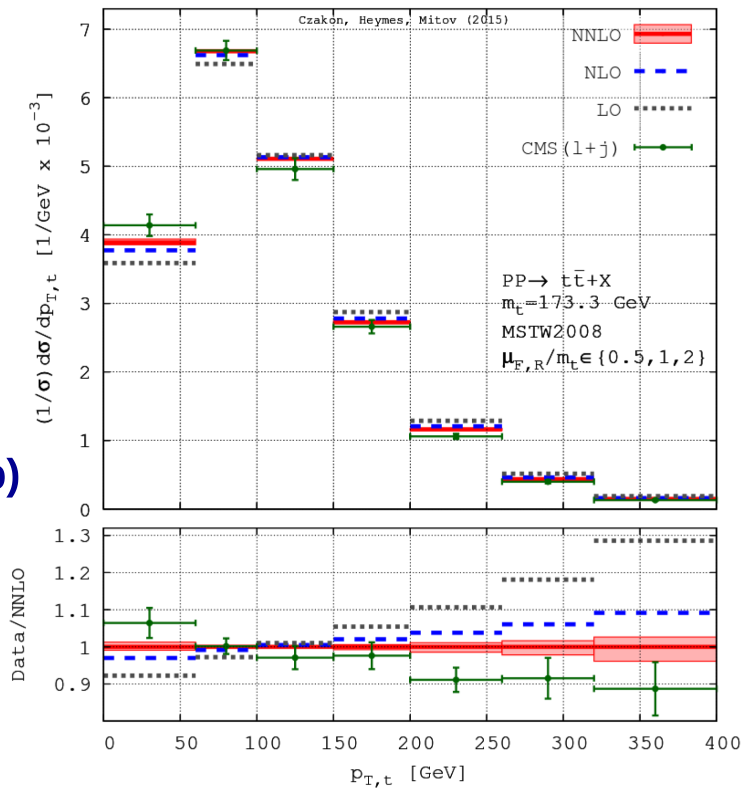
- Signal $t\bar{t}$ reference sample used for unfolding: **MadGraph+Pythia6** (**Powheg+Pythia8**) at **8** (**13**) TeV

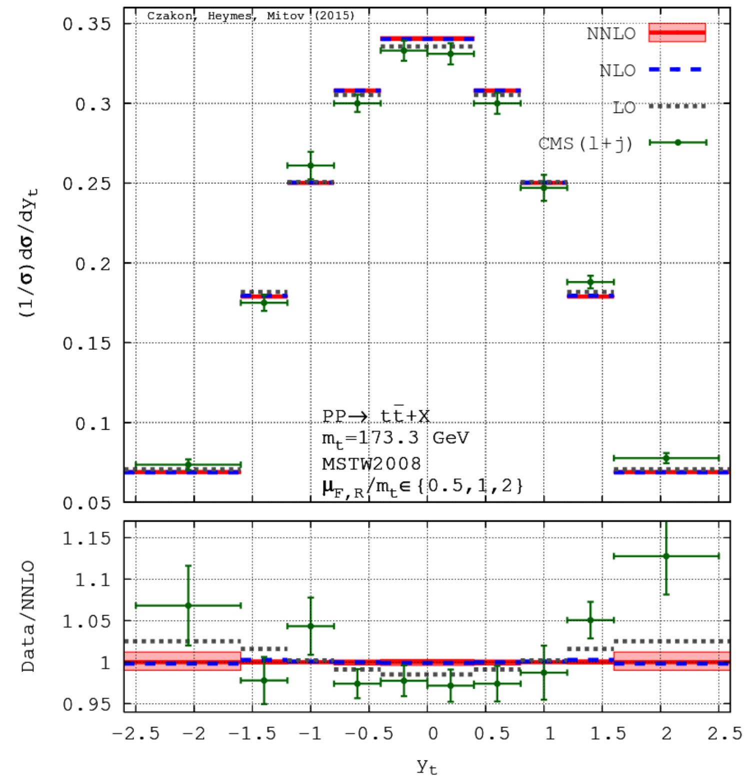




Full NNLO vs 8TeV CMS Data: p_T(top), y(top)

p_T(top)





y(top)

- First full **NNLO** calculations for top-quark pair production at 8 TeV LHC are available! [arXiv:1511.00549, by M. Czakon, D. Heymes, A. Mitov]
- Normalized top/antitop p_T and y theory distributions vs **CMS data** [arXiv:1505.04480]
- NNLO error band from scale variations only
- **NNLO QCD corrections bring SM predictions closer to CMS data in all bins of p_T(top)**
- **NLO** and **NNLO** looks almost identical for y(top) → looking forward for new measurements!

Overview of uncertainties at 8 TeV

Each uncertainty propagated through analysis chain individually

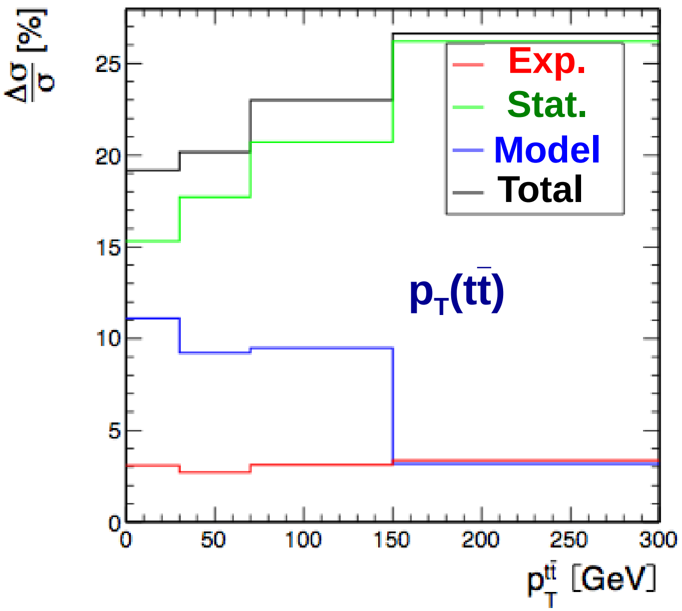
- For each source, the corresponding efficiency, resolution or scale is changed by its uncertainty or similar
- Systematic unc. per bin: difference of the changed result wrt nominal value

Normalization: only shape uncertainties contribute

Relative systematic uncertainty (%)				
Source	Lepton and b jet observables		Top quark and $t\bar{t}$ observables	
	ℓ+jets	dileptons	ℓ+jets	dileptons
Trigger eff. & lepton selec.	0.1	0.1	0.1	0.1
Jet energy scale	2.3	0.4	1.6	0.8
Jet energy resolution	0.4	0.2	0.5	0.3
Background (Z+jets)	—	0.2	—	0.1
Background (all other)	0.9	0.4	0.7	0.4
b tagging	0.7	0.1	0.6	0.2
Kinematic reconstruction	—	$<$0.1	—	$<$0.1
Pileup	0.2	0.1	0.3	0.1
Fact./renorm. scale	1.1	0.7	1.8	1.2
ME-PS threshold	0.8	0.5	1.3	0.8
Hadronization	2.7	1.4	1.9	1.1
Top quark mass	1.5	0.6	1.0	0.7
PDF choice	0.1	0.2	0.1	0.5

Overview of uncertainties at 13 TeV (dilepton)

- Measurement dominated by statistical uncertainty in all bins of each observable



- Hadronization: PowhegV2+Pythia8 vs PowhegV2+Herwig++
- Generator: PowhegV2+Pythia8 vs aMC@NLO(FxFx)+Pythia8

- **Typical dominant uncertainties:** medians of the distribution of uncertainties over all bins for rapidity (all other) observables

Source	Uncertainty (%)
Generator	3.4 (1.6)
Hadronization	2.3 (2.9)
PDF	1.5 (0.5)
JES	1.2 (1.2)
JER	0.7 (0.8)
b-tagging	0.6 (0.9)