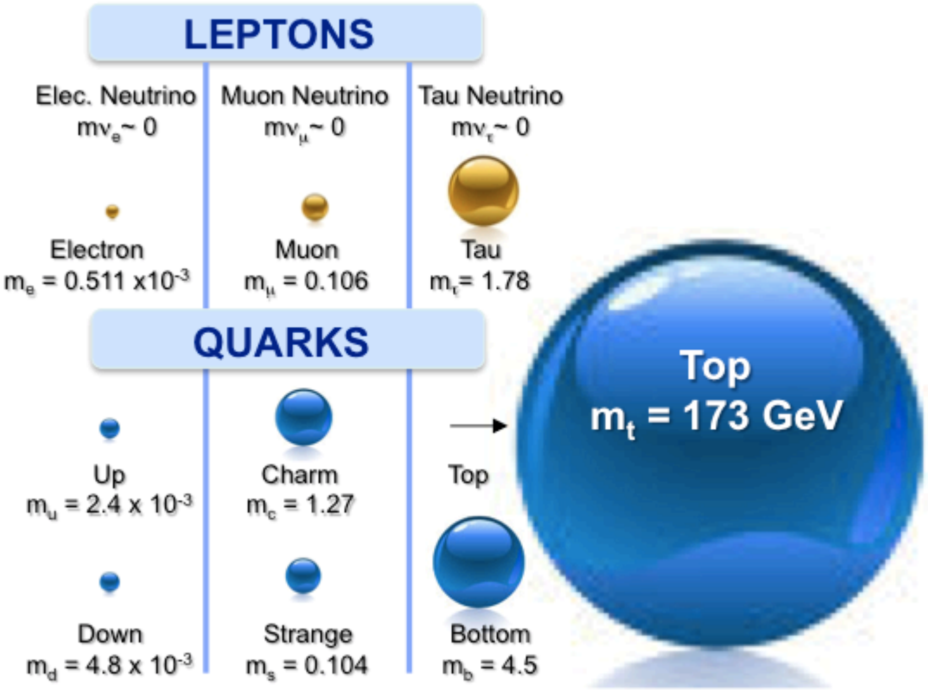
Measurement of top quark pair differential cross sections at 8 TeV using CMS

María Aldaya (DESY)

LHC Discussions, DESY, 13 July 2015

Top: key to QCD, EW and new physics



- Heaviest elementary particle known
  - top: largest Yukawa coupling to Higgs

$$y_t = \frac{\sqrt{2} m_t}{v} \approx 1$$

- Several open questions:
  - Role in EW symmetry breaking?
  - Role in beyond SM (BSM) physics?
  - Are the couplings affected?

- Only quark that decays before hadronizing

  → No bound states, study the properties of a 'bare' quark

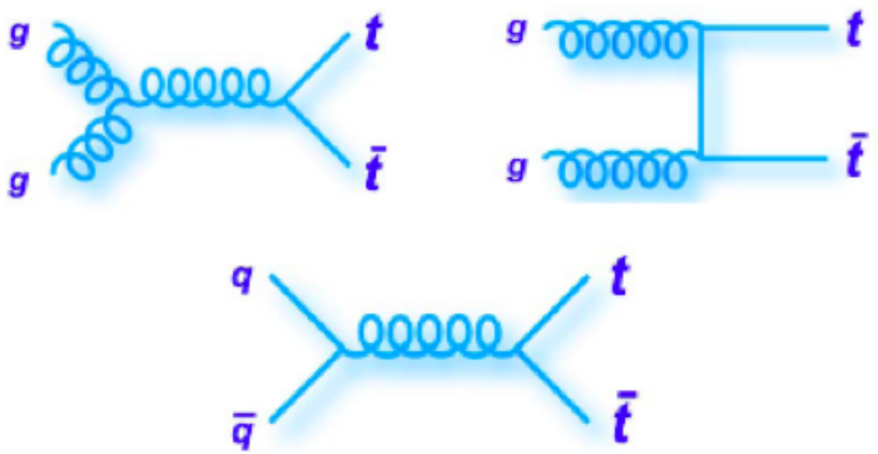
- Main background for Higgs and many searches for new physics

- Top quark measurements may provide insight into BSM physics

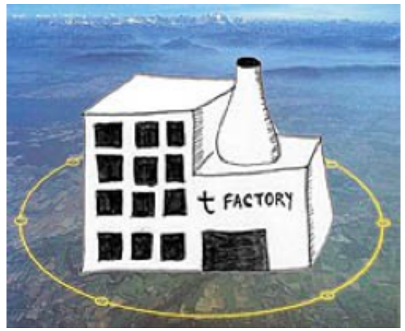
Precise understanding of top quark production is crucial

Top quark pair production …

- Top quark pair ($t\bar{t}$) production mainly by gluon-gluon fusion at LHC (~80% at 7,8 TeV)



- LHC Run-I:
- 7 TeV: ~ 1 M $t\bar{t}$ pairs
- 8 TeV: ~ 5.5 M $t\bar{t}$ pairs



Full NNLO+NNLL calculation[1] [Czakon, Fiedler, Mitov, arXiv:1303.6254]

$\sqrt{s}$ [TeV]	$\sigma_{t\bar{t}}$(NNLO+NNLL)[2] [pb] (172.5 GeV)	Scale uncert. [pb]	PDF+α_s[3] uncert. [pb]	Mass uncert. [pb]
7	177.3	+4.6 -6.0	+9.0 -9.0	+5.4 -5.3
8	252.9	+6.4 -8.6	+11.7 -11.7	+7.6 -7.3
13	831.8	+19.8 -29.2	+35.1 -35.1	+23.2 -22.5

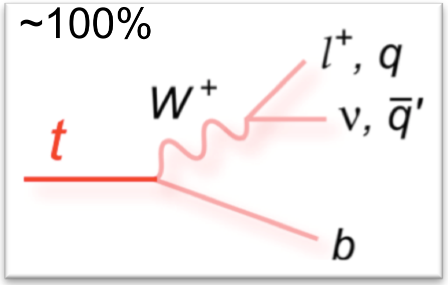
[1] https://twiki.cern.ch/twiki/bin/view/LHCPhysics/TtbarNNLO

[2] calculated using Top^{++}(v2.0)

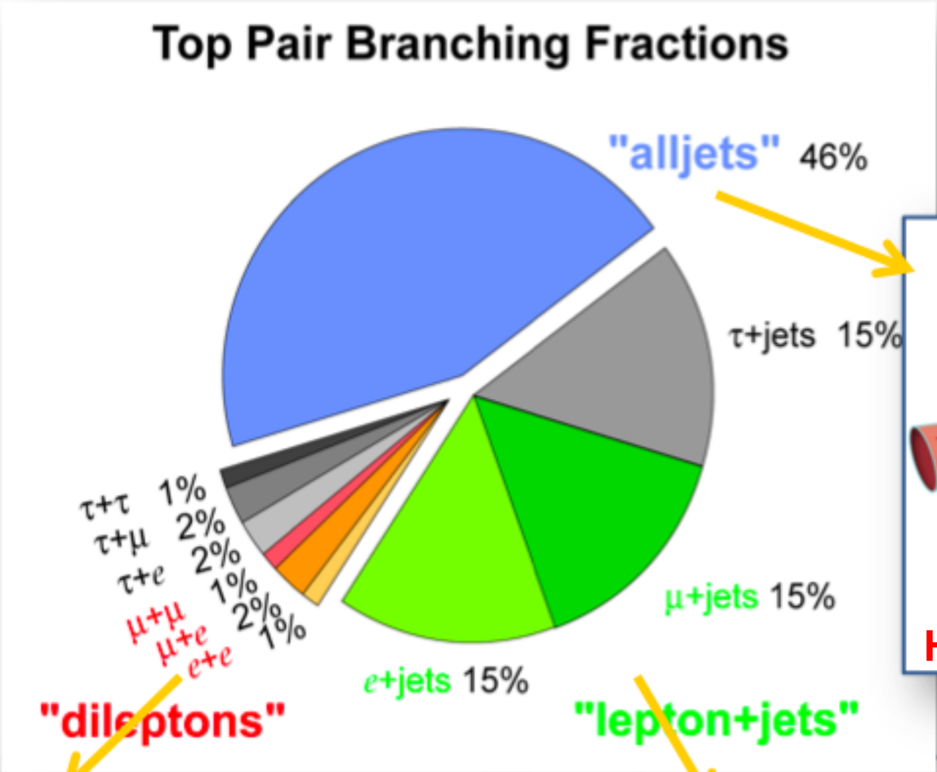
[3] calculated following PDF4LHC prescription

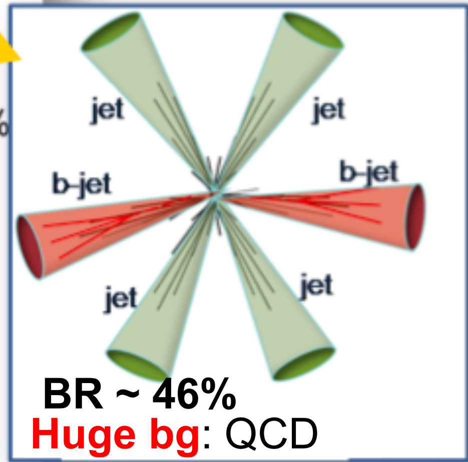
… and decay

In Standard Model:

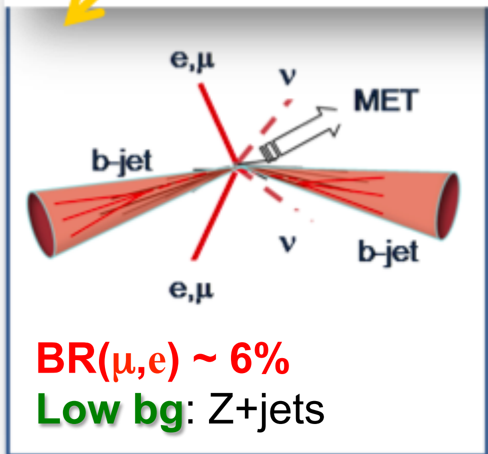


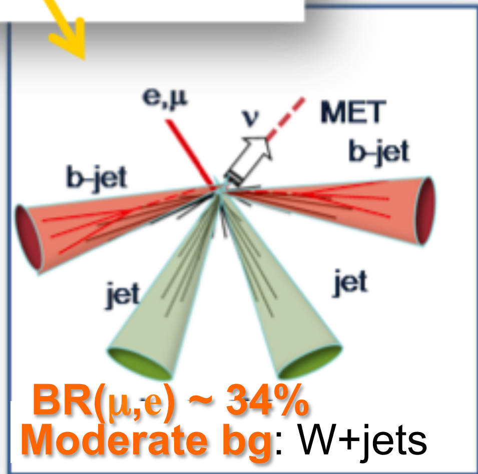
W decay modes
define
top final states





Focus on
l+jets and dilepton
final states





Why measure differentially?

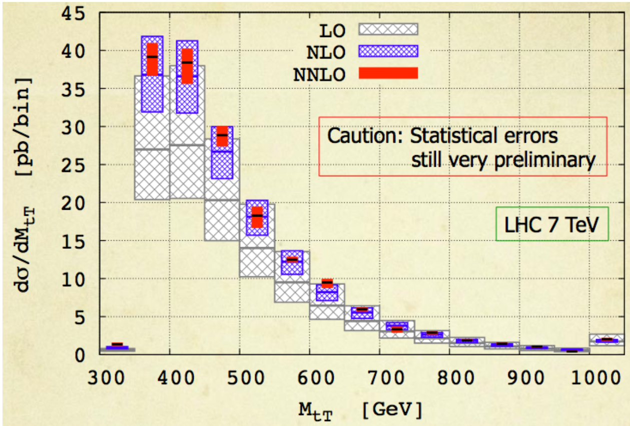
- The large LHC samples allow measuring $t\bar{t}$ production as a function of many kinematic observables
  - Precise tests of QCD for top quark production in different regions of the phase space
  - Theory predictions and models need to be tuned and tested with measurements
  - Extract/use for PDF fits
  - Enhance sensitivity to new physics
  - Background for Higgs, rare processes and many BSM searches

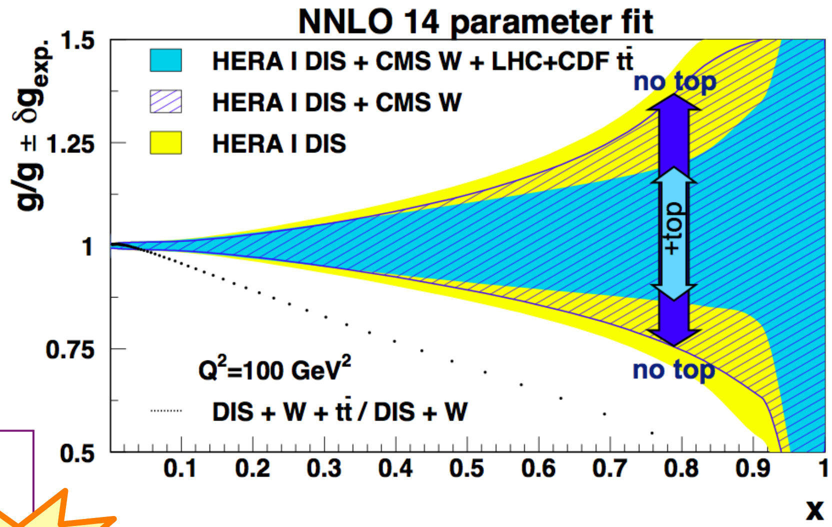
Important for LHC Run-II !

This talk: latest results from CMS at 8 TeV, performed by UHH & DESY (arXiv:1505.04480, submitted to EPJC)

May'15

A.Mitov et al, TOP2015, http://indico.cern.ch/event/290408/session/1/contribution/27/material/slides/0.pdf



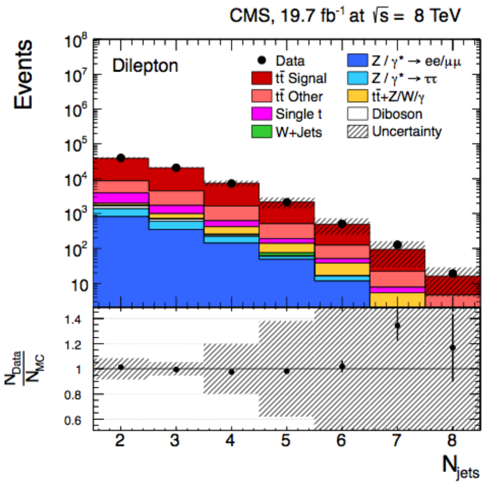


M.Guzzi, K.Lipka, S.Moch, arXiv:1406.0386

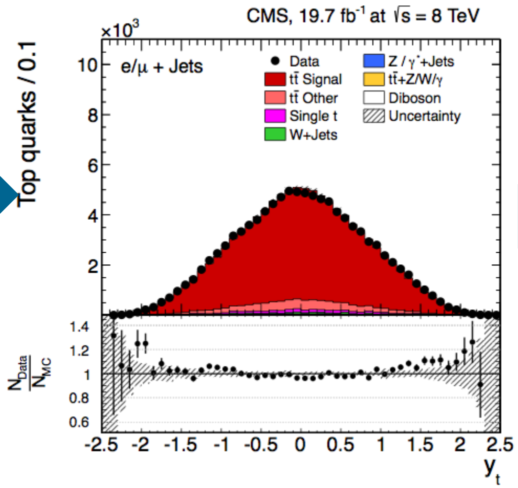
General analysis strategy

Measure cross section $\sigma(t\bar{t})$ as a function of kinematic distributions of
top, top pairs, b-jets, leptons, and lepton pairs

(1) Event selection



(2) $t\bar{t}$ kinematic reconstruction



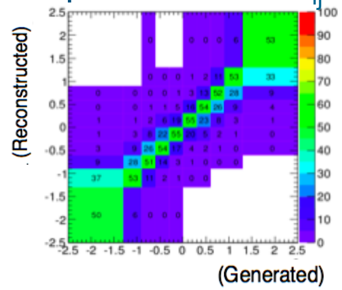
(3) Bin-wise cross section measurement

- Subtract background
- Unfolding: correct for detector effects & acceptance

Δ^i_X = bin width for variable X

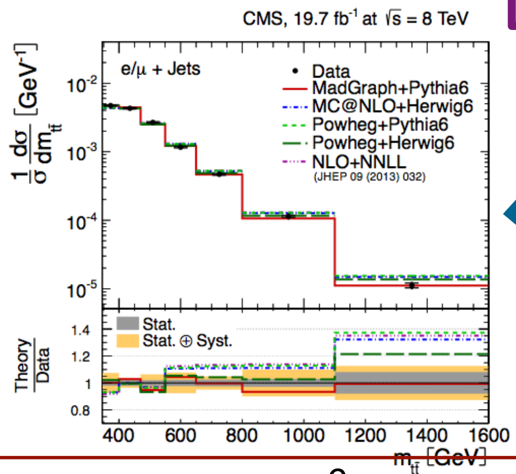
$$\frac{1}{\sigma}\frac{d\sigma}{dX_i} = \frac{1}{\sigma}\frac{\sum_j A^{-1}_{ij}\left[(N_{\mathrm{data},j} - N_{\mathrm{BG},j})\right]}{\Delta_{X_i}\cdot L}$$

Response matrix A_{ij}



(4) Differential $t\bar{t}$ cross sections

- Fiducial or extrapolated to full phase space
- Compare with theory
- Normalize to inclusive $\sigma(t\bar{t})$: many systematics cancel, only shape uncertainties contribute

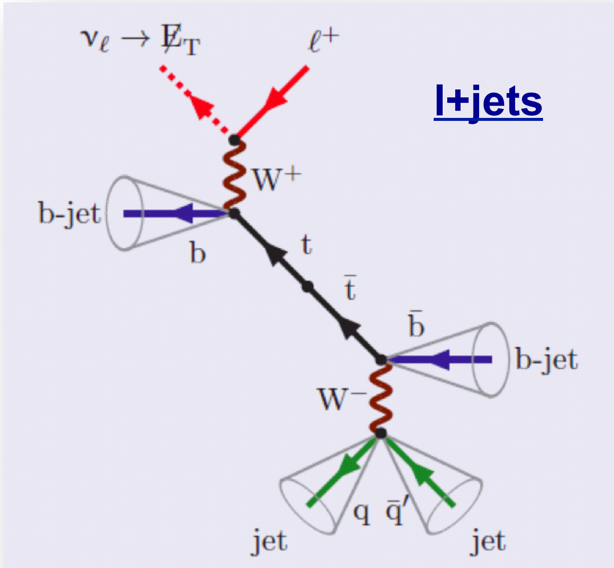


Event selection

- Goal: clean $t\bar{t}$ sample (<20% background)

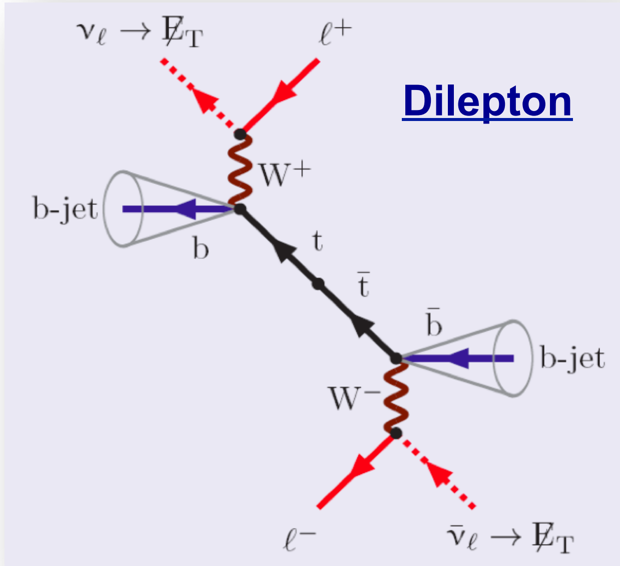
Lepton+jets:

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



Dileptons:

- 2 opp.-sign, 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, $\mu\mu$: $E_T^{miss} > 40$ GeV; Z veto: $|m_Z - m_{ll}| > 15$ GeV



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

Kinematic distributions – ℓ+jets

- Pure $t\bar{t}$ samples after event selection:

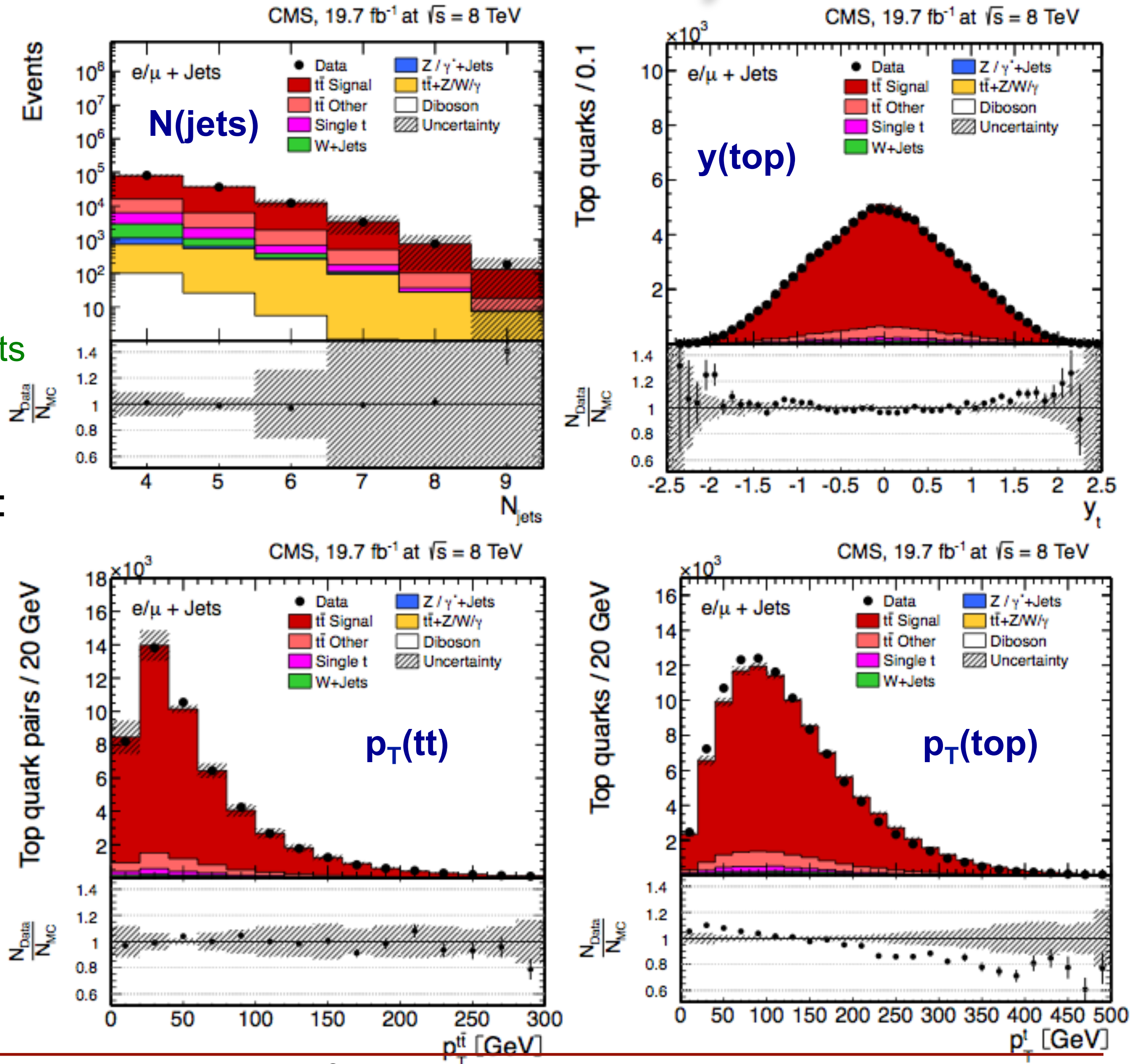
  ~ 80% $t\bar{t}$

- Main backgrounds:

  $t\bar{t}$(other), single top, W+jets

- Reference $t\bar{t}$ prediction:

  MadGraph+Pythia6



Top p_T spectrum tends to lower p_T values in data than in simulation, both for l+jets and dileptons

Kinematic distributions – dileptons

- Pure $t\bar{t}$ samples after event selection:

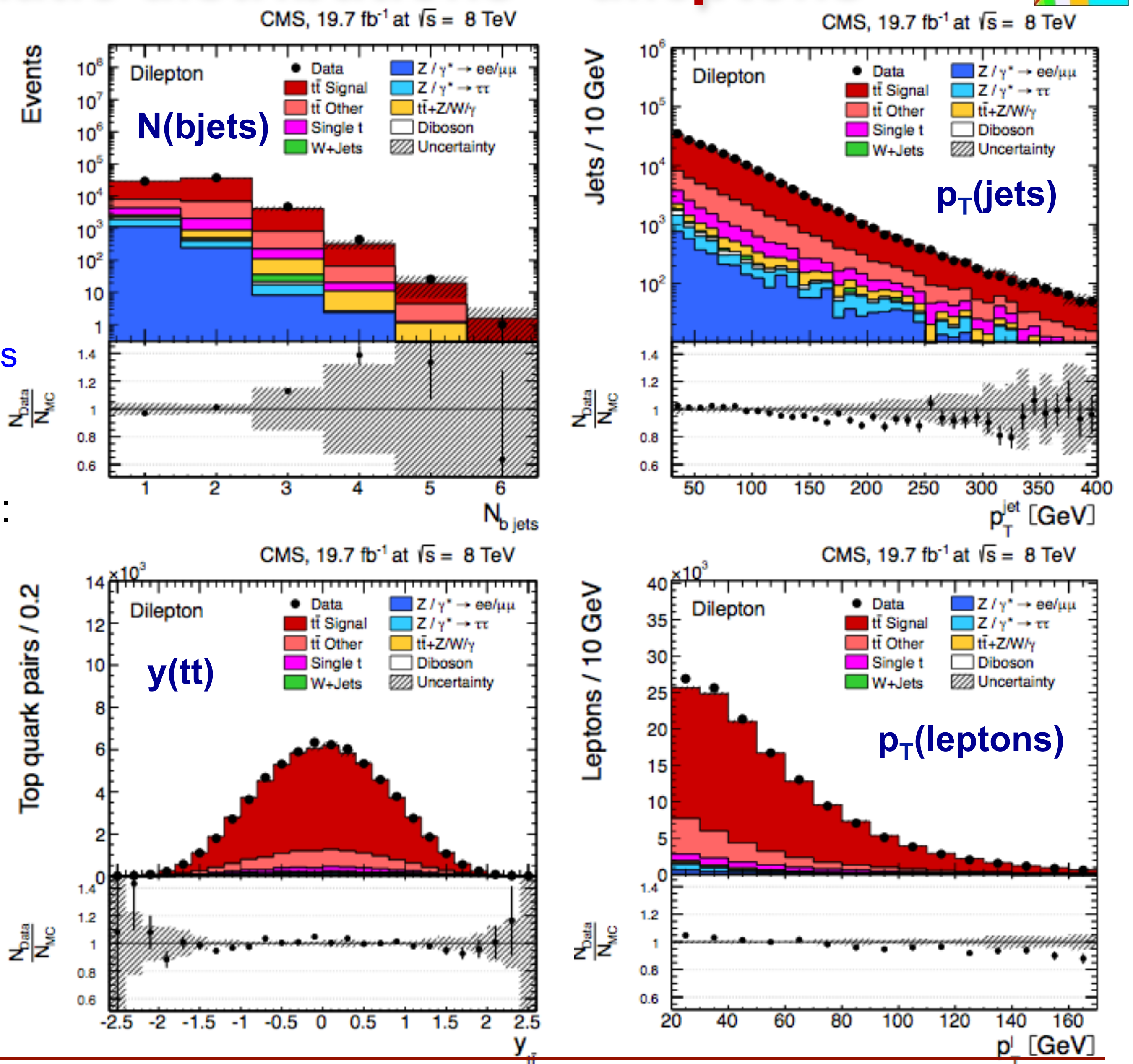
  ~ 80% $t\bar{t}$

- Main backgrounds:

  $t\bar{t}$(other), single top, Z+jets

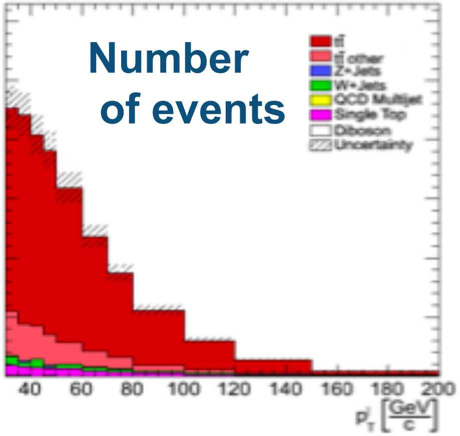
- Reference $t\bar{t}$ prediction:

  MadGraph+Pythia6

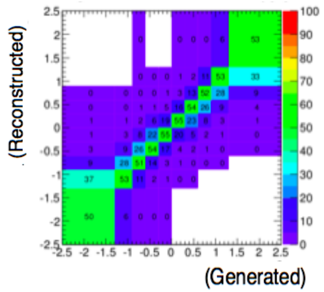


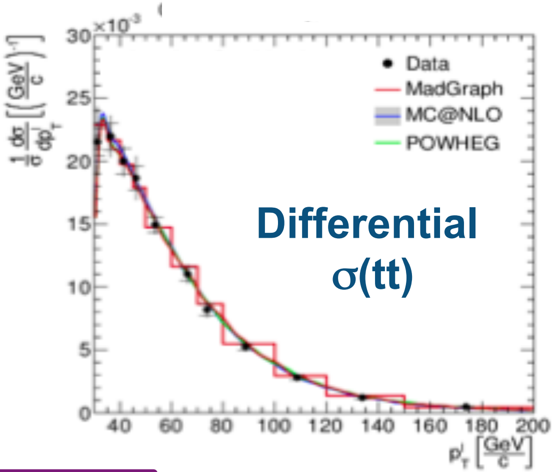
Leptons and jets p_T spectra shows similar trend as top $p_{T,}$ both for l+jets and dileptons

Normalized differential cross section



Response matrix A_{ij}





$$\frac{1}{\sigma}\frac{d\sigma}{dX_i} = \frac{1}{\sigma}\frac{\sum_j A_{ij}^{-1}\left[(N_{\rm data,j} - N_{\rm BG,j})\right]}{\Delta_{X_i}\cdot L}$$

Binning

Chosen to limit migration effects, quantify with:

purity (p^i) & stability (s^i): ≥ 60%

$$p^i = \frac{N^i_{rec\&gen}}{N^i_{rec}} \qquad s^i = \frac{N^i_{rec\&gen}}{N^i_{gen}}$$

Regularized unfolding

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

$$N_{i,\rm unf} = A_{ij}^{-1} N_{j,\rm meas}$$

- Regularization used to remove large statistical fluctuations (SVD)

Phase space

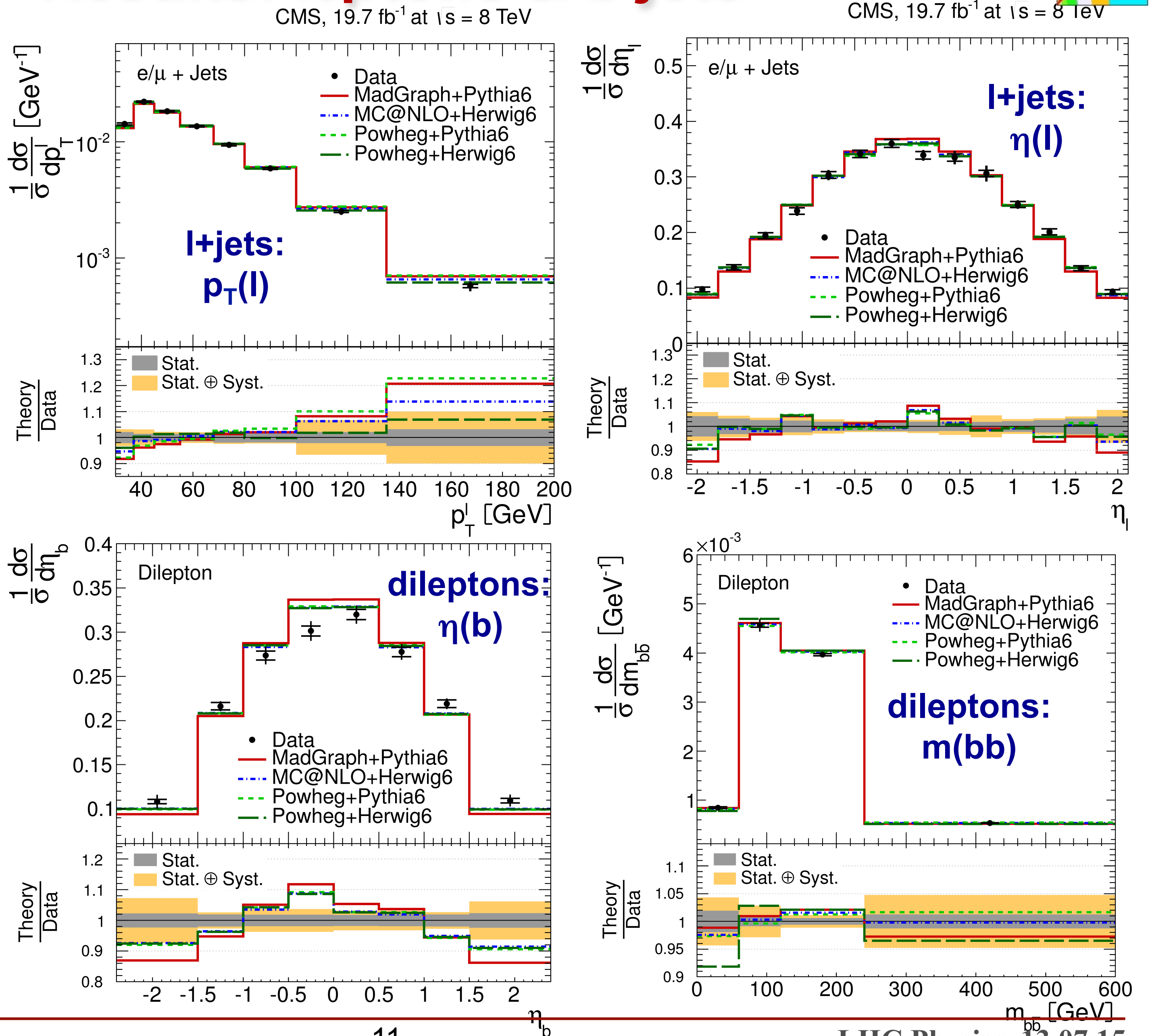
- **Leptons and (b)jets**: correct to **particle level** in **fiducial phase space**
  - leptons: p_T > 33 (20) GeV, $|\eta|$ < 2.1 (2.4) for l+jets (dilep)
  - jets: p_T > 30 GeV, $|\eta|$ < 2.4
- **Top and tt** observables: correct to **parton level** in **full phase space**

→ Compare with highest QCD calculations

Results: leptons & b-jets

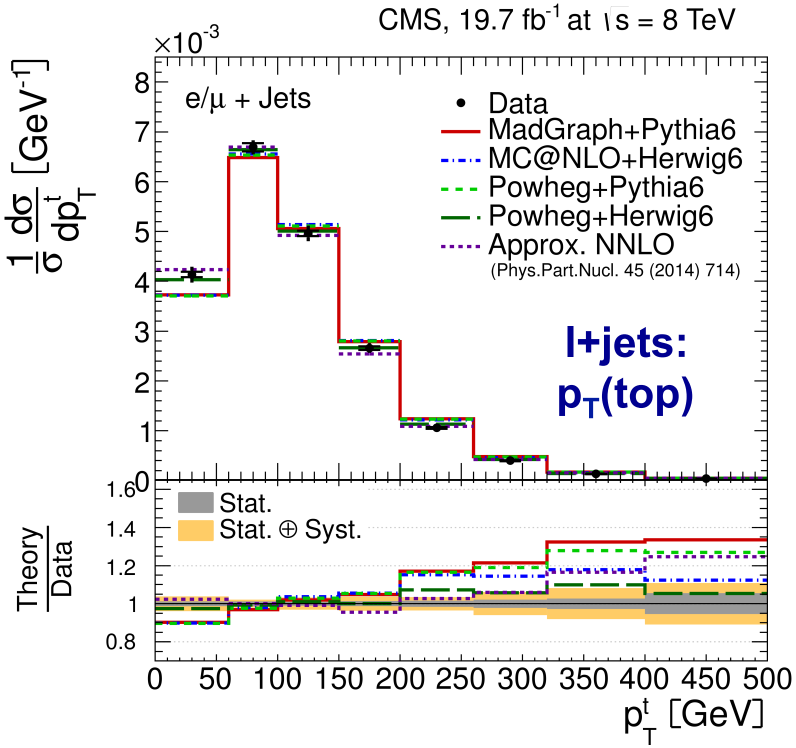
- Measured in fiducial phase space
- p_T spectra slightly softer in data
- η(b) distributions slightly less central in data
- Good description of all distributions: Powheg+Herwig6

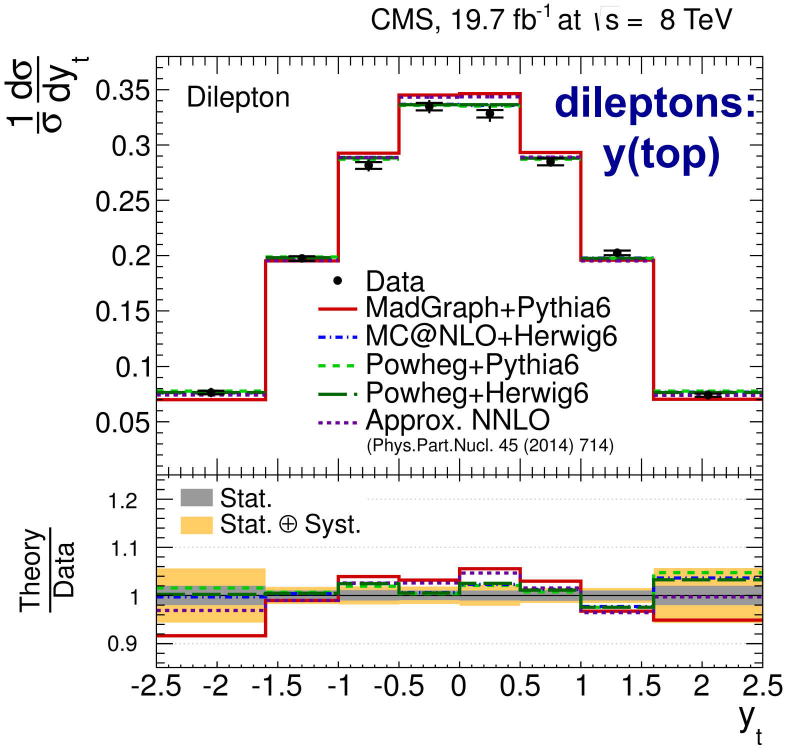
Consistent with 7 TeV results at CMS
[EPJ C73 (2013) 2339]



Results: top quarks

- Full phase space, parton level





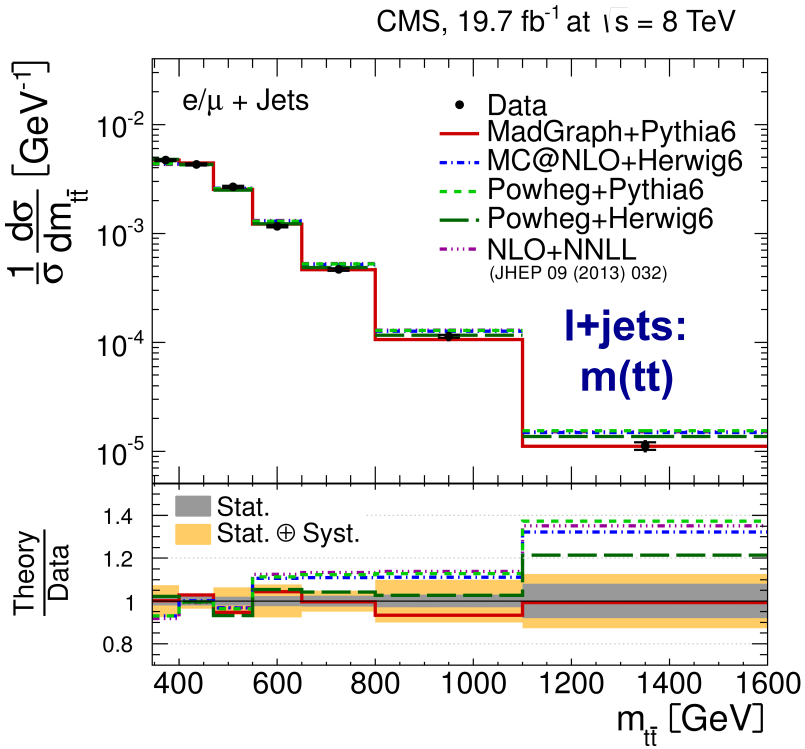
- For all distributions, data generally better described by Powheg+Herwig6
- **p_T(top)**: - Softer spectrum in data, in particular at the tails (in agreement with ATLAS)

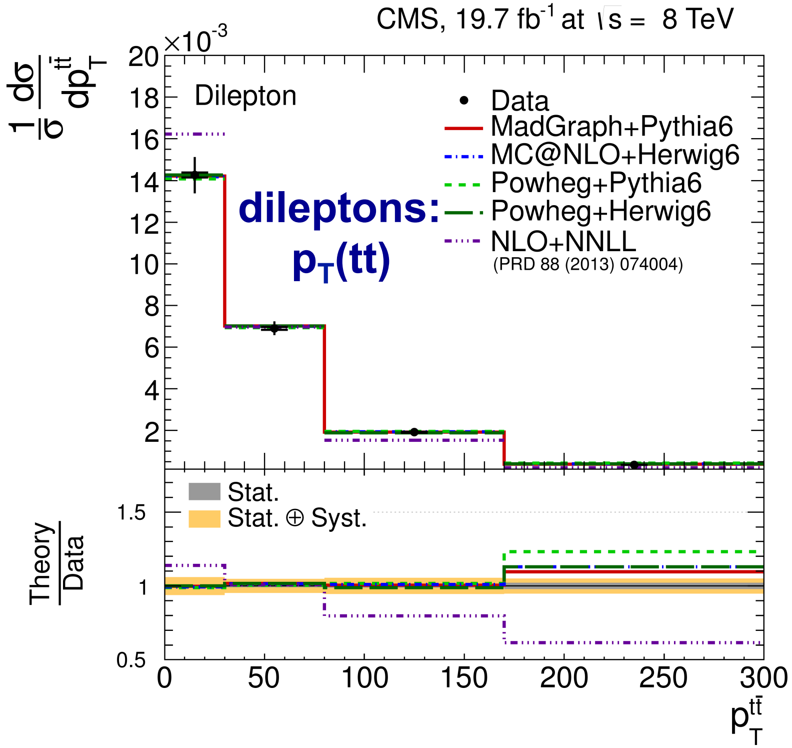
  - Good agreement with approx. NNLO calculation over full p_T range

- **y(top)**: - Slightly less central in data than in MC, in particular for MadGraph+Pythia6 and approx. NNLO

Results: $t\bar{t}$ system

- Full phase space, parton level

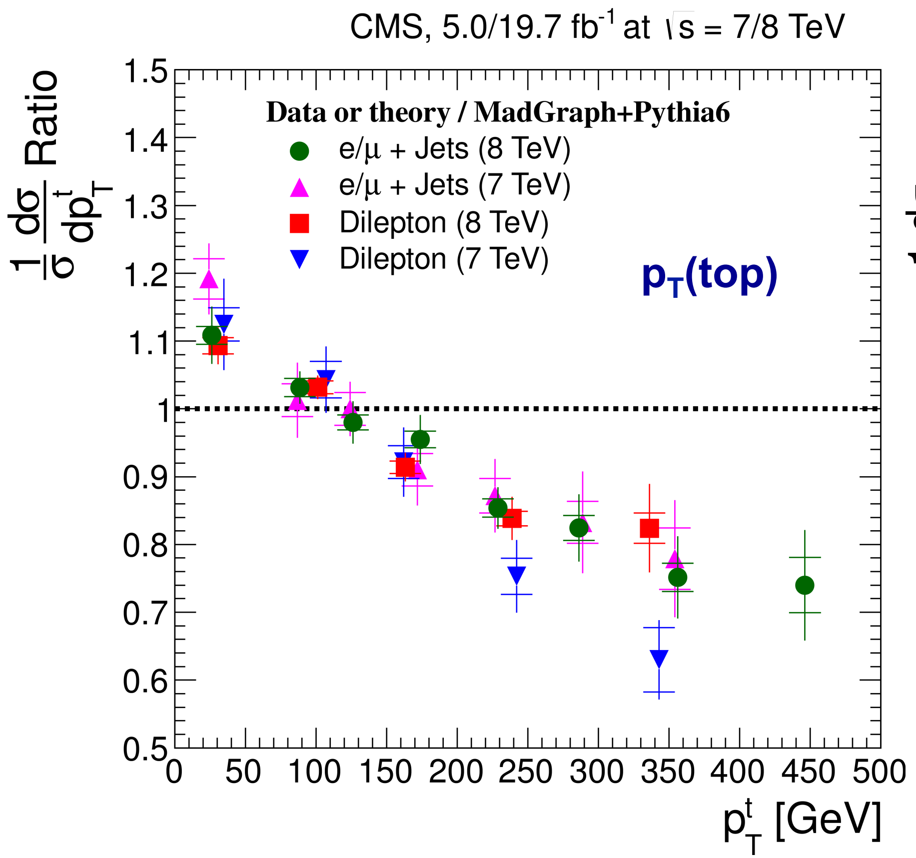


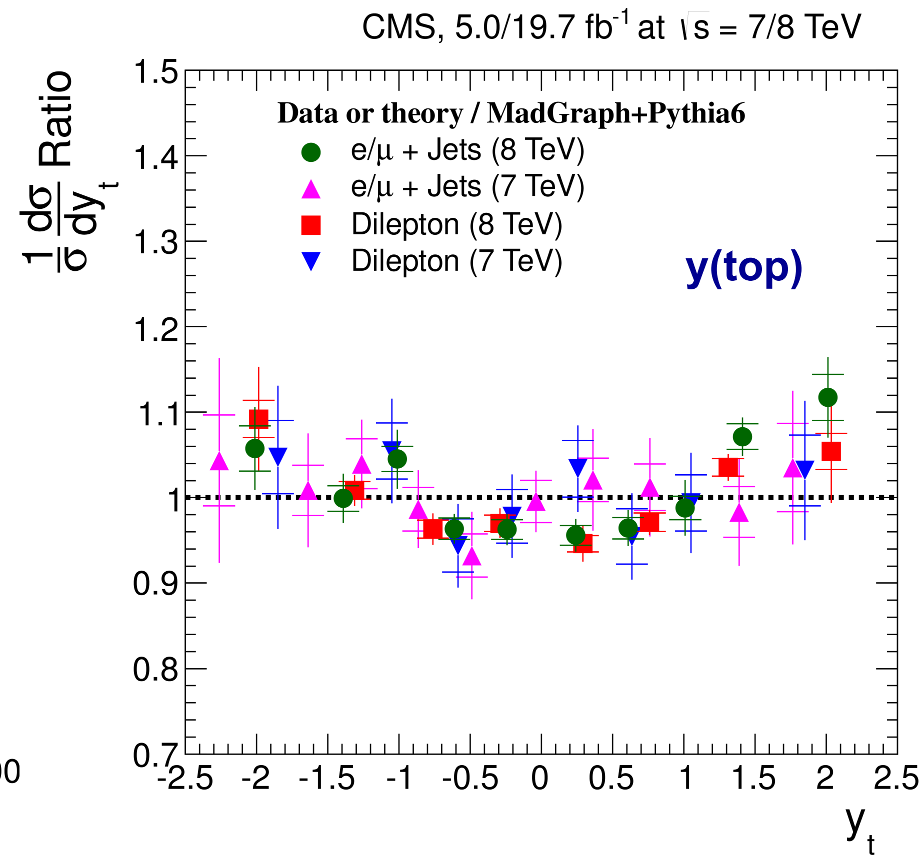


- For all distributions, data generally better described by Powheg+Herwig6
- **m(tt)**: Tails in data lower than predictions (in agreement with ATLAS)
- **p_T(tt):** Well described by all predictions, except NLO+NNLL (in agreement with ATLAS)

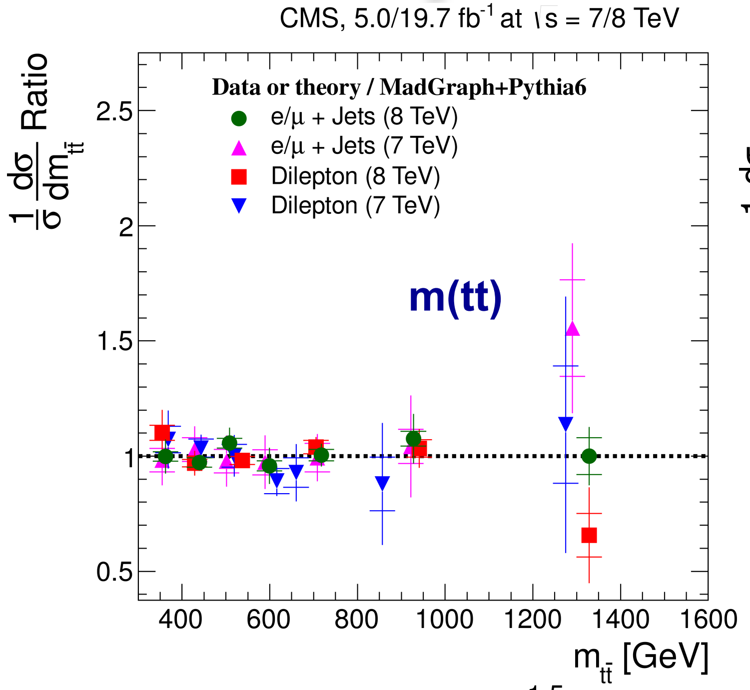
Consistency of results in CMS: top quarks

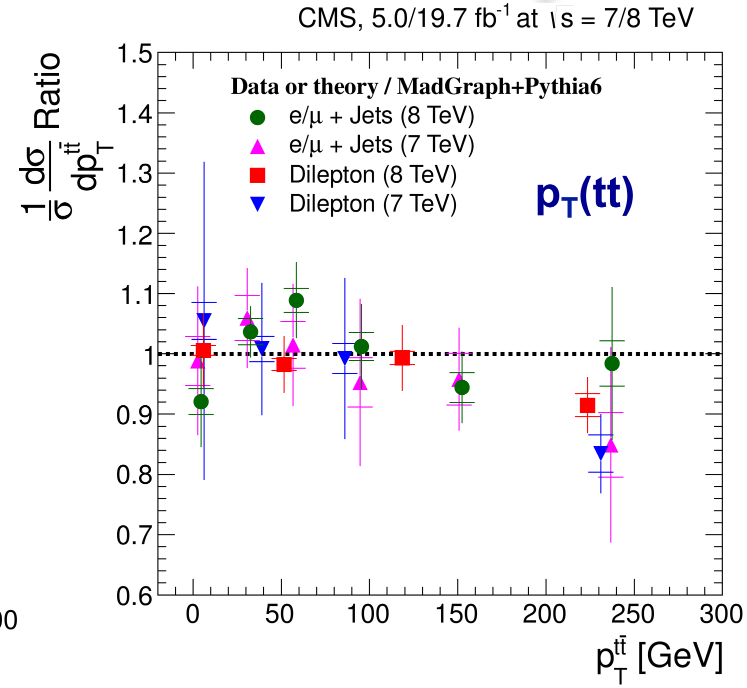
- Results are consistent:
  - l+jets vs dileptons
  - 7 TeV vs 8 TeV



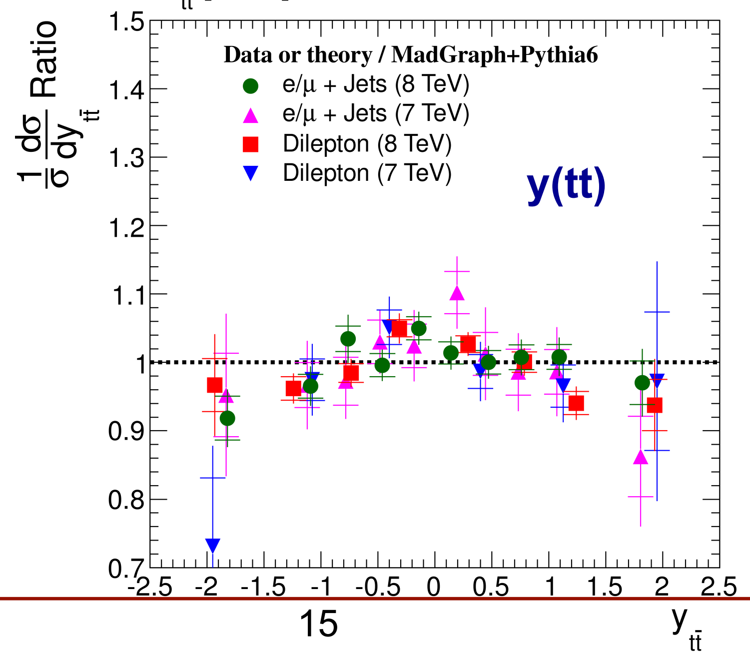


Consistency of results in CMS: $t\bar{t}$ system





- Results are consistent:
  - l+jets vs dileptons
  - 7 TeV vs 8 TeV



Summary

- **Top quark differential cross sections: next step in precision physics**
  - Essential for constraining the SM
  - Ideal probe for looking for new physics beyond the SM
- **Presented latest results from CMS from Run-I**
  - Top quarks, $t\bar{t}$, leptons, lepton pairs, b-jets
  - Normalization to inclusive cross section: precision 3 % – 10 % (syst. dominated)
- **General good agreement with SM predictions**
  - Gaining sensitivity to model differences
  - Measurements are typically lower than predicted with increasing p_T(top) and m(tt)
  - Comparison between ATLAS & CMS (7 TeV): consistent definition of top quark, default generators are compatible → work within TOPLHCWG
- **More results in the pipeline at 8 TeV**
  - Can be used for more quantitative studies to investigate the compatibility between ATLAS & CMS data

Looking forward to LHC Run-II !

Additional information

Systematic uncertainties

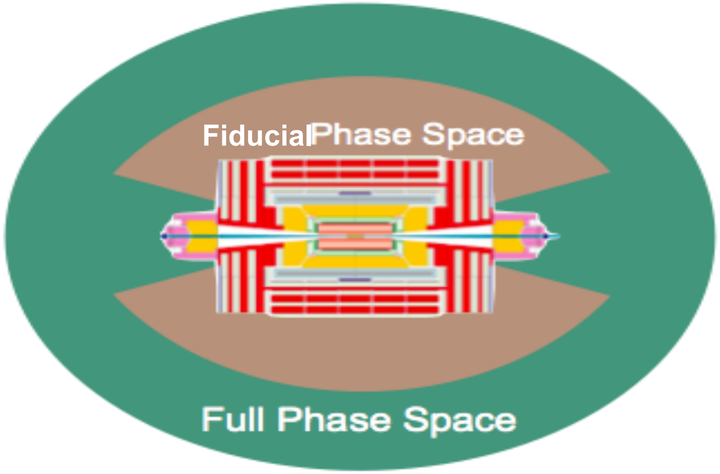
- Determined individually for each bin of the measurement
- General strategy: propagate uncertainties through full analysis
  - Experimental uncertainties: efficiencies, resolutions, …
  - Modelling uncertainties: scale choices, hadronization, PDFs, …
- Normalization: only shape uncertainties contribute

Typical values per bin

	Source	Relative systematic uncertainty (%)			
		Lepton and b jet observables		Top quark and $t\bar{t}$ observables	
		ℓ+jets	dileptons	ℓ+jets	dileptons
Experimental	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
Modeling	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

Phase space definitions and observables

- **Acceptance** corrections:
  - Extrapolation from limited detector acceptance to full phase space with theory or MC simulation
  - Measurement of differential cross sections in fiducial phase space → reduced dependence of measurement on signal/background modelling



[Picture from U. Husemann]

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

- **leptons, (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 (i.e, ν, μ from hadron decays are inside jets)

b-jets: contains any of the decay products of a B-hadron

  - Phase space definition closely follows the (detector-level) event selection

MC $t\bar{t}$ samples: parameters & tunes

Default samples

ATLAS

Matrix element	Shower & Hadronization	PDF	Tune
MC@NLO v4	Herwig 6.5 + Jimmy 4.31	cteq66 or CT10	AUET1/2
Powheg	Pythia 6	cteq66 (7 TeV) or CT10 (8 TeV)	Perugia 2011 C
Alpgen	Herwig 6.5 + Jimmy 4.31	cteq6ll	AUET2

- Powheg+Herwig: NLO PDF CT10, AUET2 Herwig 6.5 tune

CMS

Matrix element	Shower & Hadronization	PDF	Tune
MadGraph v5	Pythia 6	cteq6l	Z2 (7 TeV) Z2* (8 TeV)
Powheg	Pythia 6	cteq6m (7 TeV) CT10 (8 TeV)	Z2 (7 TeV) Z2* (8 TeV)
MC@NLO v3.4	Herwig 6 + Jimmy	cteq6m	default tune

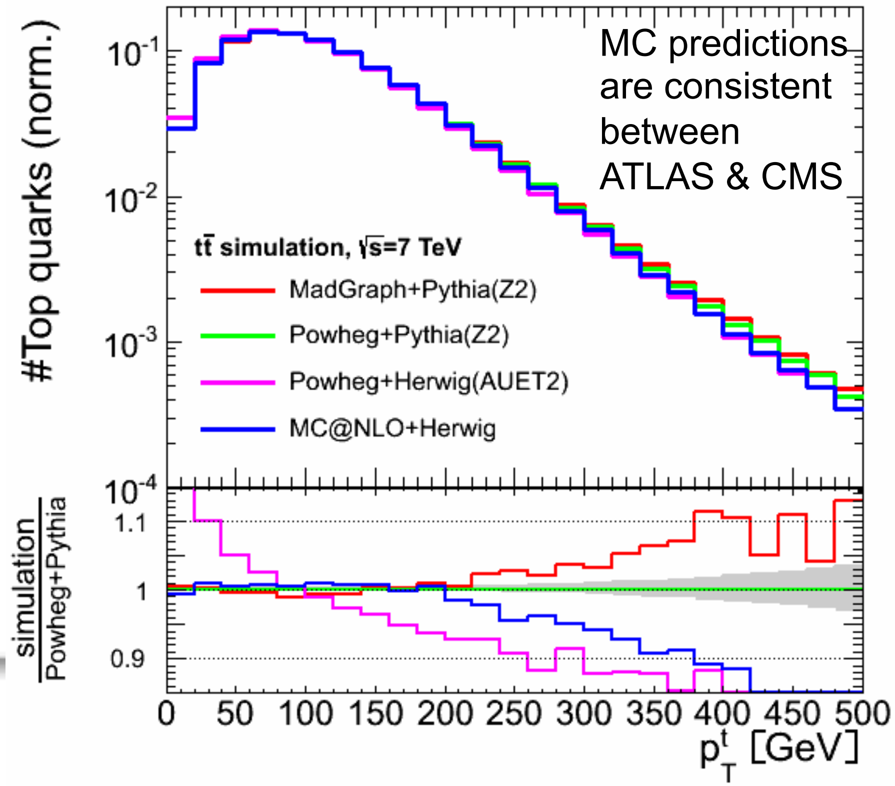
- Powheg+Herwig: NLO PDF CT10 (8 TeV), AUET2 Herwig6 tune
  CTEQ6M (7 TeV)

ATLAS & CMS generator comparison (7 TeV)

- ATLAS and CMS have consistent definition of the top quark
- Compatible behaviour in corresponding sample pairs: same differences between generator and parton shower schemes

Monte Carlo simulations

- Generators:
  - ◇ ME at tree level (Alpgen, MadGraph)
  - ◇ NLO (POWHEG, MC@NLO)
- Showering:
  - ◇ Pythia (transverse-momentum-ordered evolution scale)
  - ◇ HERWIG (angular-ordered)



- Powheg+Herwig provides reasonable description of the data for both experiments (different treatment of the hardest ISR than Pythia)

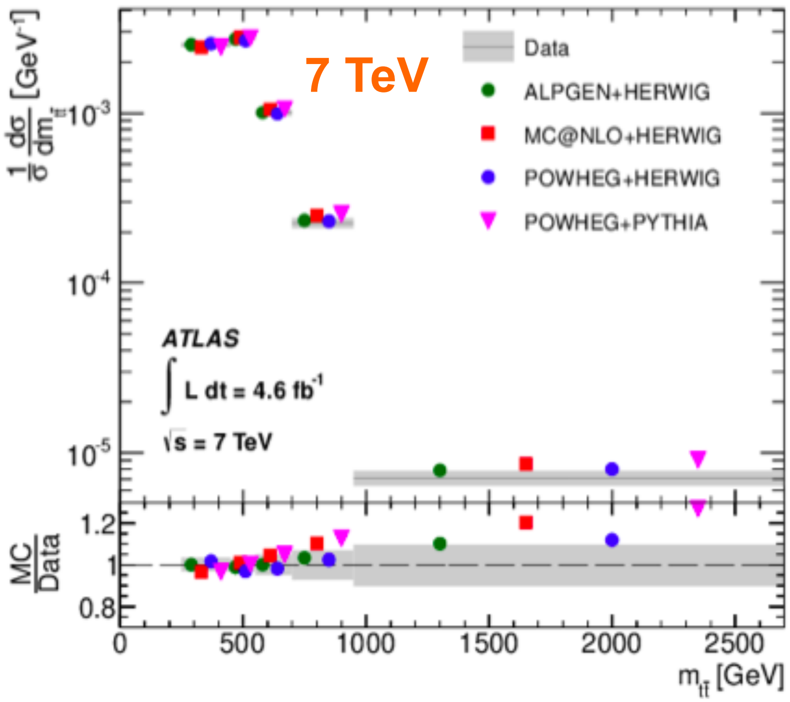
Ongoing work within the TOPLHCWG

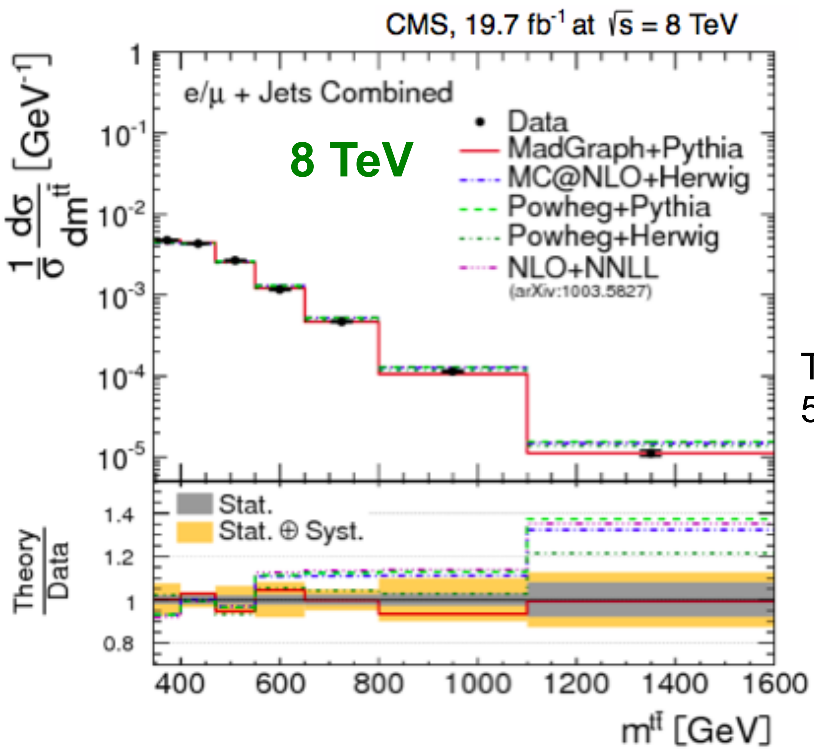
ATLAS (7 TeV) & CMS (8 TeV), l+jets – m(tt̄)

- Parton level, extrapolated to full PS
- Consistent top quark definition in ATLAS & CMS: *before decay and after QCD radiation*
- Compatible generators between ATLAS & CMS

  (See TOPLHCWG meeting Nov13: https://indico.cern.ch/event/280522/session/2/contribution/7/material/slides/0.pdf)

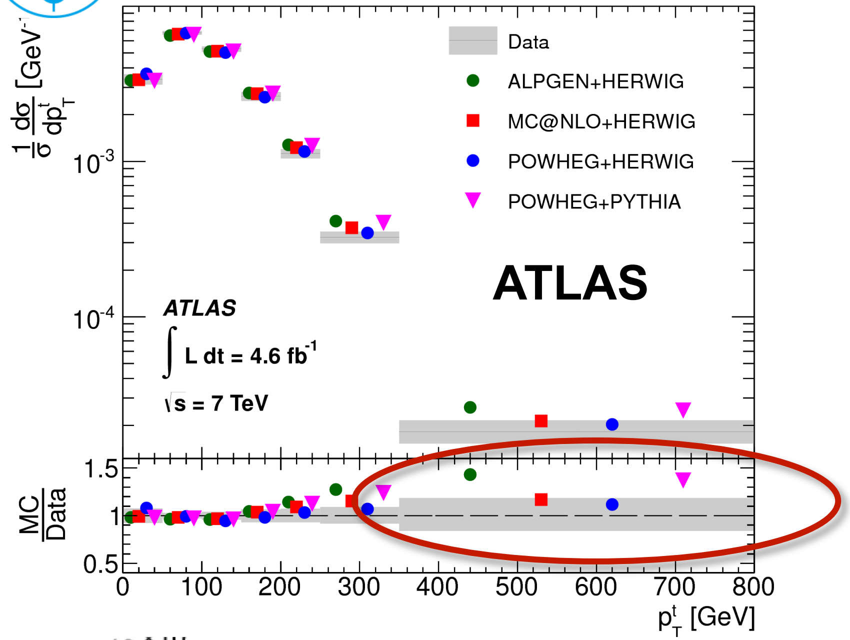
- In general: predictions are larger than data for ATLAS & CMS, in particular at high m(tt̄) values
- CMS: Similar behaviour for dileptons, both at 7 and 8 TeV

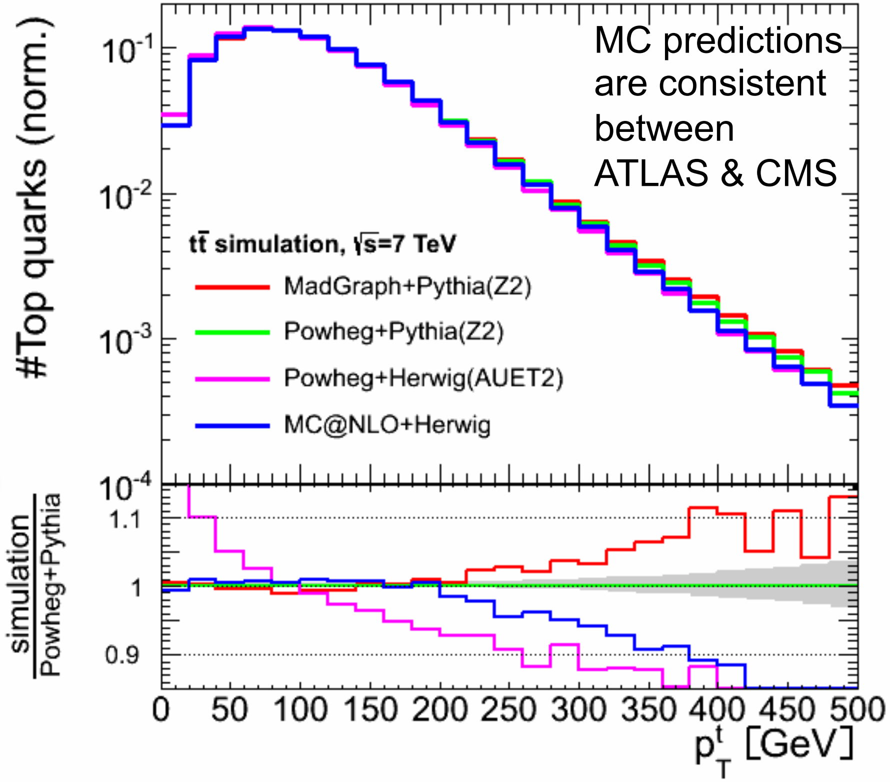


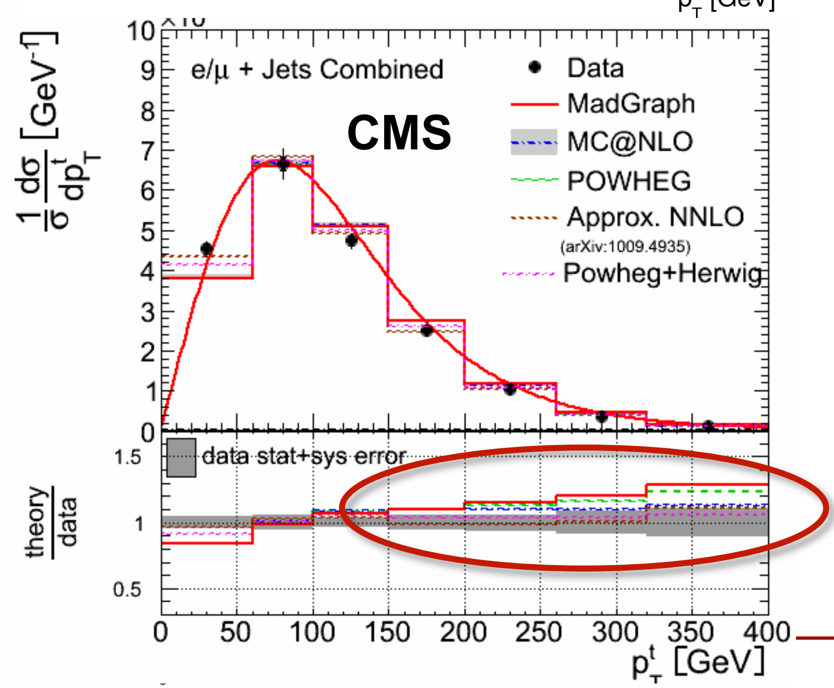


Total uncert: 5% – 10%

ATLAS & CMS (7 TeV, l+jets) – p_T(top)



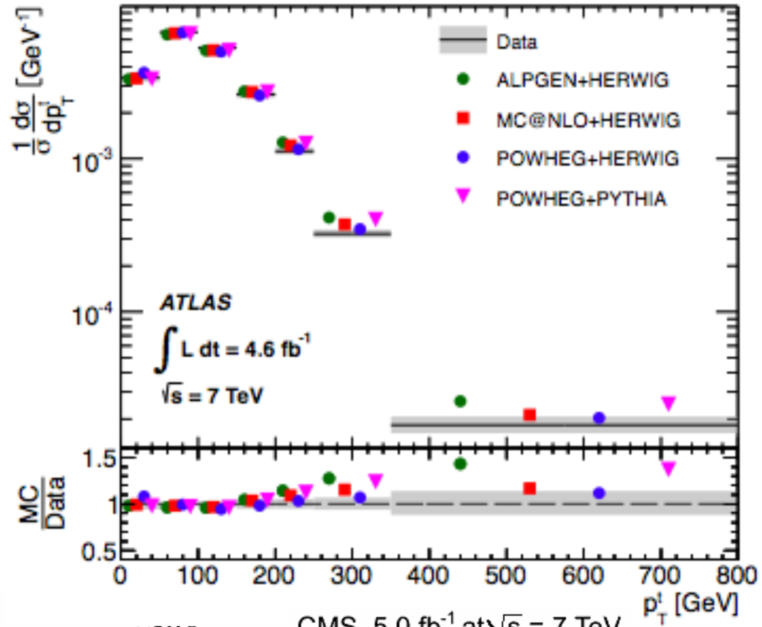


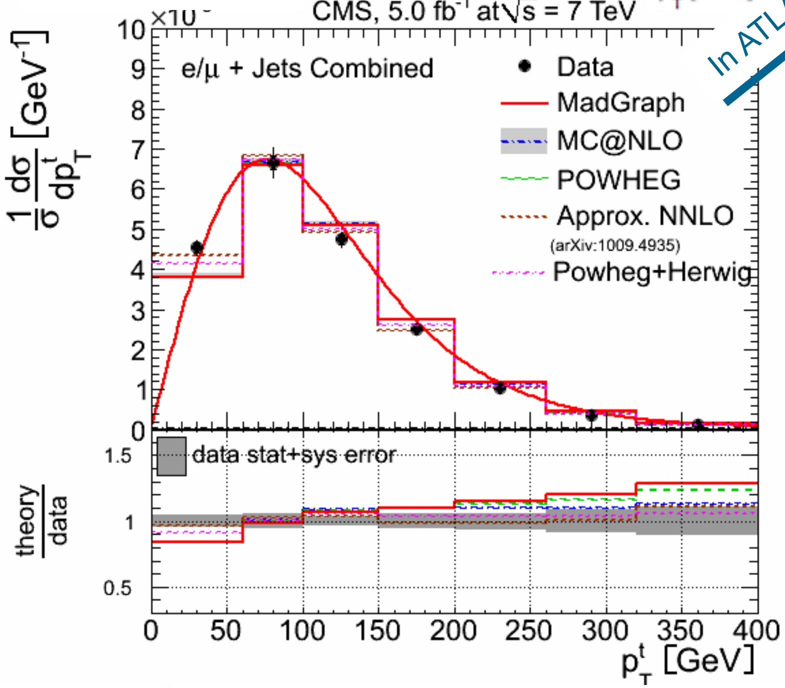


- Significant differences between MC models
- Data not well described by MC at large p_T(top)

→ Needs to be understood: crucial also for BSM searches

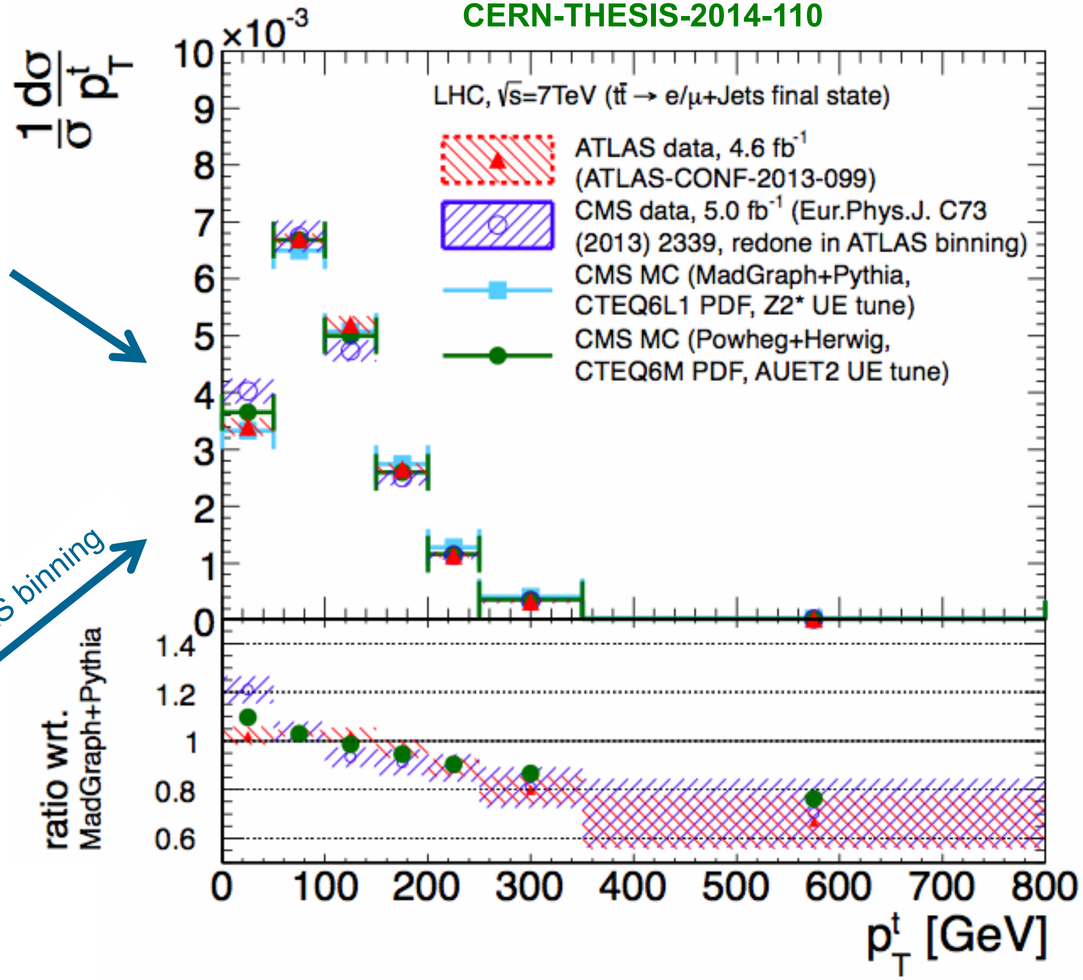
ATLAS & CMS (7 TeV, l+jets) – p_T(top)





In ATLAS binning

CERN-THESIS-2014-110



- Powheg+Herwig is consistent between ATLAS & CMS
- Powheg+Herwig is consistent with ATLAS & CMS data

Unfolding

- **Regularized** SVD **unfolding** (Höcker, Kartvelishvili 1996; Blobel 1984, 2002)
  - Smearing of true distribution **f** by detector effects → reconstructed distribution **g**

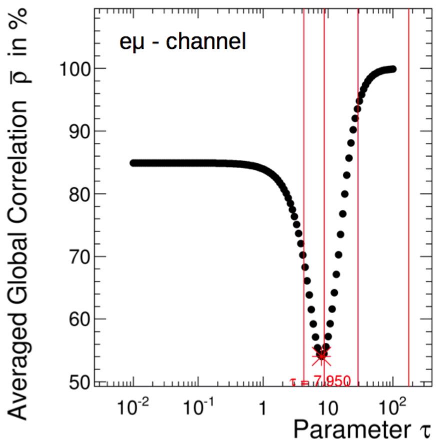
$$g_i(\mathbf{f}) = \sum_j A_{ij} f_j \qquad \text{(response matrix A from MC simulation)}$$

  - Obtain estimator for f by **inverting A** via minimization of regularized χ^2 function

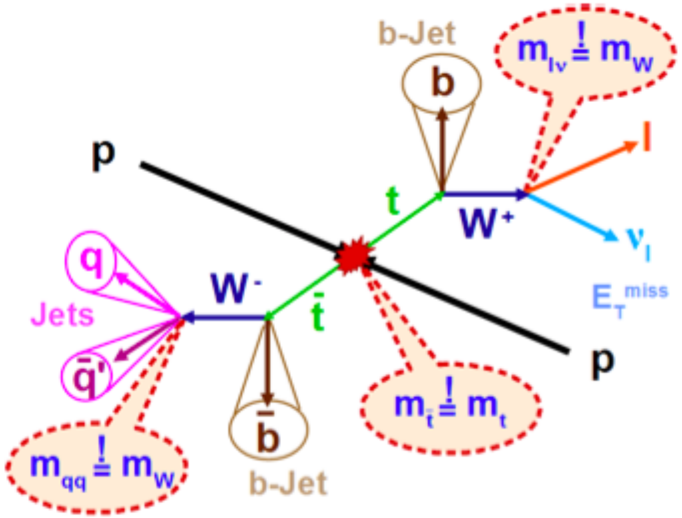
$$R(\mathbf{f}, \tau) = \sum_i g_i(\mathbf{f}) - x_i \ln g_i(\mathbf{f}) + \tau C(\mathbf{f})$$

  - Regularization parameter τ optimized for each distribution: minimize **averaged global correlation** among bins

$$\overline{\rho}(\tau) = \frac{1}{n}\sqrt{\sum_i \rho_i(\tau)^2} \text{ with } \rho_i(\tau) = \max\left(\rho(f_i, \sum_{j\neq i} \alpha_j f_j)\right)$$



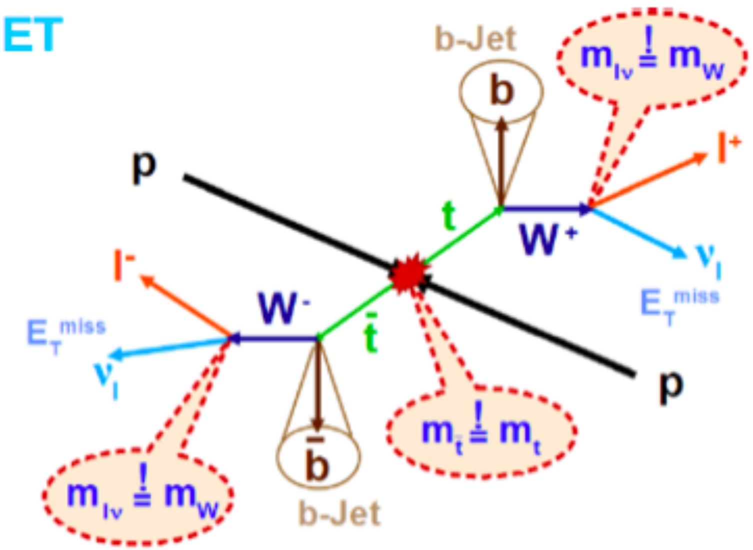
Kinematic reconstruction – ℓ+jets



- 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 – dileptons

- **Measured input: 2 jets, 2 leptons, MET**
- **Unknowns:** $\bar{p}_\nu, \bar{p}_{\bar{\nu}} \rightarrow 6$
- **Constraints:**
  - > $m_t, m_{\bar{t}} \rightarrow 2$
  - > $m_{W(+)}, m_{W(-)} \rightarrow 2$
  - > $(\bar{p}_\nu + \bar{p}_{\bar{\nu}})_T = \text{MET} \rightarrow 2$



- Reconstruct neutrino momenta from **p_T conservation** and known **W boson and top quark masses** → up to four-fold ambiguity, take solution with lowest $t\bar{t}$ mass
- Reconstruct event **100 times** with **varied b-jet** and **lepton four-momenta**
  → recover events without neutrino solution
  → top kinematics from **average**, weighted with true M_{lb} distribution
- Take jet-parton assignment with **largest sum of weights**