Towards a comparison and combination of differential distributions ATLAS and CMS

O. Hindrichs
On behalf of the ATLAS and CMS collaborations

University of Rochester

LHCTopWG open meeting

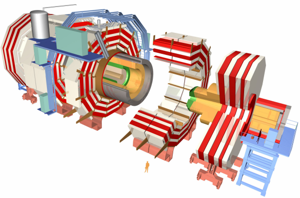
06.06.2017

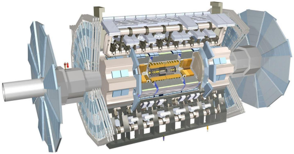
Why compare?

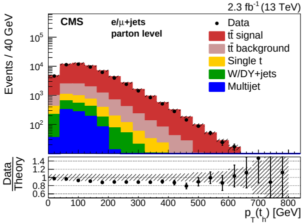
- Chance to check analyses and to spot problems.
- Good compatibility increases reliability.

Why combine?

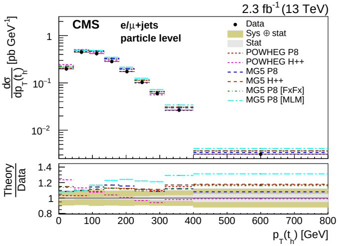
- Reduce uncertainties →**provide unbiased and most precise experimental statement** for comparisons with theory.

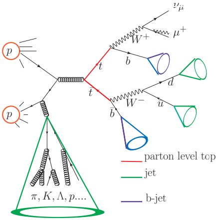


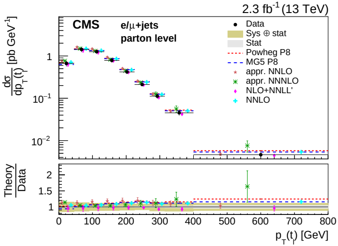




- Select $t\bar{t}$ enriched sample.
- Reconstruction of $t\bar{t}$ system. Depends on final state and type of measurement.
- Subtraction of backgrounds
- Unfolding: correct for detector effects and extrapolate to particle or parton level

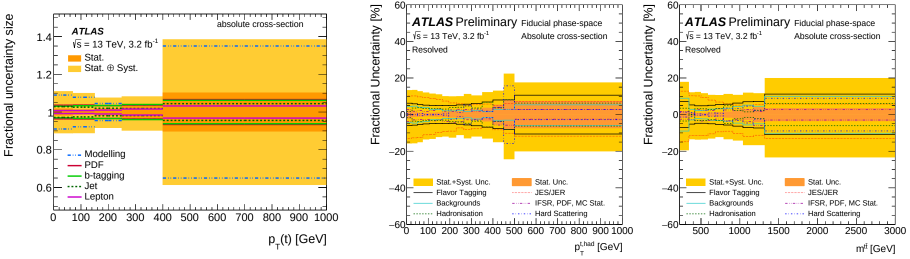


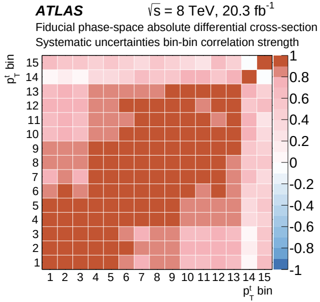




[PRD 95(2017) 092001]

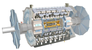
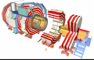
[EPJC 77(2017) 299, CONF-2016-040]



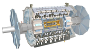
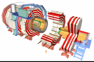


[EPJC 76(2016) 538]

- Typical uncertainties: 10–20%.
- Uncertainties are usually correlated among bins →never compare bin by bin to theory.
- For comparisons/combinations correlations between experiments to be considered.

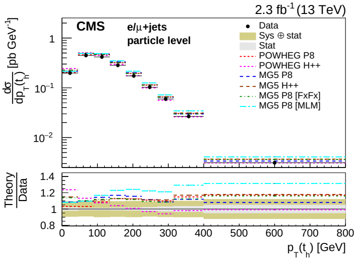
		ATLAS 	CMS 
dilepton	parton	PRD 94(2016) 092003 *7/8 TeV*	EPJC 73(2013) 2339 *7 TeV* EPJC 75 (2015) 542 TOP-14-013 (Sub. to EPJC) *double differential*
	particle	JHEP 1609 (2016) 074 *jet activity*	
ℓ+jets	parton	PRD 90 (2014) 072004 *7 TeV* EPJC 76(2016) 538 PRD 93(2016) 032009 *boosted*	EPJC 73(2013) 2339 *7 TeV* EPJC 75 (2015) 542 PRD 94(2016) 072002 *boosted*
	particle	EPJC 76(2016) 538 JHEP 01(2015)020 *jet activity* PRD 93(2016) 032009 *boosted*	PRD 94(2016) 052006 *event variables(no $t\bar{t}$ reco.)* PRD 94(2016) 072002 *boosted*
allhadronic	parton/particle		EPJC 76(2016) 128

Same binning

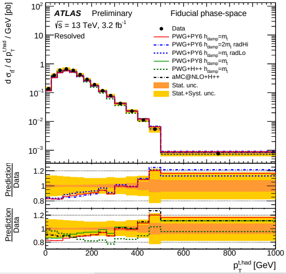
		ATLAS 	CMS 
dilepton	parton		PAS-TOP-16-011
	particle	EPJC 77(2017) 299 EPJC 77(2017) 220 *jet activity*	PAS-TOP-16-007
ℓ+jets	parton		PRD 95(2017) 092001 *1D/2D*
	particle	CONF-2016-040 *resolved/boosted*	PRD 95(2017) 092001 *1D/2D*
allhadronic	parton		PAS-TOP-16-013 *resolved/boosted*
	particle	CONF-2016-100 *boosted*	PAS-TOP-16-013 *resolved/boosted*

- All using 2015 data with 2–3 fb^{-1} luminosity.
- Analyses based on 2016 data (about 35 fb^{-1}) in pipeline.

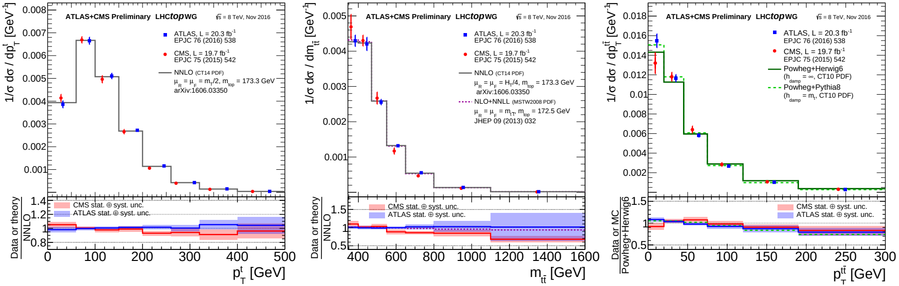
- **Particle-level definitions:**
  - Objects: lepton dressing (anti-k_{T} clustering vs summation in cone), jets (with vs without neutrinos)
  - Pseudo-tops: dilepton: differences in required number of b jets ..., ℓ+jets: do not follow the LHCTopWG recommendation.
- **Parton-level definitions:**
  - Run I: the parton-levels definitions were compared and agreed.
  - Run II: experiments use new MC setups. Parton level depends on parton shower and tuning could introduce bias in unfolding and acceptance correction (a priori smaller than modeling uncertainties). $\rightarrow$can be tested with common MC.
- Uncertainties and correlations for each experiment in compatible granularity and sources.
- Reasonable estimate of inter-experiment correlations.
- Measurements are done with different selections of binning.



$\stackrel{?}{=}$



At 8 TeV we have one pair of analyses with the same binning and phase space (parton level): ℓ+jets CMS (EPJC 75 (2015) 542) and ATLAS (EPJC 76(2016) 538)



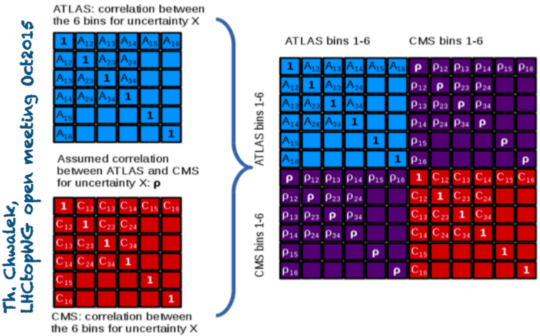
$p_{\mathrm{T}}^{\mathrm{t}}$ and $m_{\mathrm{t}\bar{\mathrm{t}}}$ CMS spectra seem to be softer than ATLAS. Is this significant?

Combination ongoing.

Combination strategy at 8 TeV

Method: use special case of χ^2 fit considering full covariance. Fit model: one parameter per bin, **BLUE** best linear unbiased estimator [NIM A270 (1988) 110, NIM A500 (2003) 391]

ATLAS and CMS uncertainties come with different granularities and have to be mapped into common groups. For each group correlations have to be estimated:



- Correlations among bins determined by experiments (A), (C).
- Overall correlation between experiments (ρ) set to values used in inclusive cross section combinations.
- Off-diagonal (ρ_{ij}) needs to be estimated, e.g.: $\rho_{ij} = \rho \cdot \frac{1}{2}(A_{ij} + C_{ij})$

Alternative method: the new framework Convino

Preliminary breakdown of uncertainties

	ATLAS	CMS	$\rho_{ATLAS\text{-}CMS}$
Detector	Lepton trigger & reco eff	Trigger eff & lepton selection	0
Detector	JES (22 sources combined in 1)	JES	0.5 *
Detector	JER	JER	0
Detector	Jet reconstruction	--	--
Detector	btagging (3 sources combined in 1)	btagging	0 *
Detector	MET	--	--
Detector	--	Pileup	--
	BG from data (W+jets, fake leptons)	--	--
	BG from MC	**BG from MC**	1
Signal model	Parton shower	Hadronization and parton shower	0.5 *
Signal model	Generator Radiation (ISR/FSR)	Fact & renorm scale ME-PS threshold	0.5 *
Signal model	PDF	PDF	1
Signal model	--	Top mass	--

F. Spanò, M. Aldaya

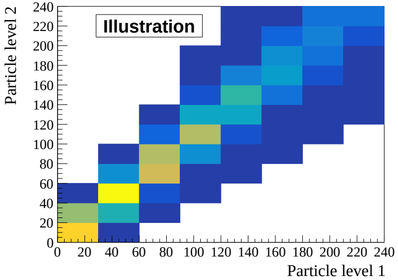
* Foresee scan of values from 0 to 1 to test stability

- Most experimental uncertainties can be assumed as uncorrelated. Some correlations remain since similar calibration methods are used.
- Unc. in background from MC and PDF are the same for both experiments →fully correlated.

In the end ensure that results do not strongly depend on the choice of correlation.

How to compare measurements with (slightly) different particle-level definitions (PL1, PL2) and different binnings?

- From MC calculate migration matrix between PL1 and PL2 and efficiencies with respect to each other, e.g., using RIVET routines.
- Project measurement at PL1 on PL2 and vice versa.
  In practice: subtract fraction that is not considered as signal in the other definition, multiply with (transposed) migration matrix where each column is normalized to efficiency.
- Uncertainty propagation is straight forward.



BUT

- Introduces new MC dependence $\rightarrow$new uncertainty to be estimated using MC variations.
- Possible gain of precision in combination to be studied distribution by distribution.

We can compare the same things.

If there is only a different binning other methods have been proposed: Heraverager

- Many degrees of freedom in the choice of binnings and definitions of the measured quantities make comparisons difficult.
- 8 TeV: combination of most precise measurements is advancing steadily.
- 13 TeV: ramping up, combination of 2015 analyses is rather exercise, since 2016 data has 15 times more statistics.

Things could be easier!

- Define common binning.
- Define common particle level, e.g., in form of RIVET routine. (In CMS physics object from RIVET routines can be integrated in analysis work flow. See talk by Markus Seidel).
- Uncertainties and correlations for each experiment in compatible granularity and sources (Modeling? See talk by Andrea Knue)
- No need to synchronize analyses completely, experiments need freedom to optimize their analyses, but a set of basic variables ($p_{\mathrm{T}}^{\mathrm{t}}$, $p_{\mathrm{T}}^{\mathrm{t\bar{t}}}$, and $m_{\mathrm{t\bar{t}}}$ ...) useful for important consistency test.

Can we have a common MC (POWHEG+PY8)? Different work flows under discussion:

- exchange of HepMC: no standard work flow for both experiments, possible in CMS for "private" production
- exchange LHE and Pythia8 parameters: work flow okay, but different version ...

Could be asymmetric, e.g.: ATLAS provides their HepMC to CMS.

Backup