Top quark event modeling and generators at CMS

Silvano Tosi

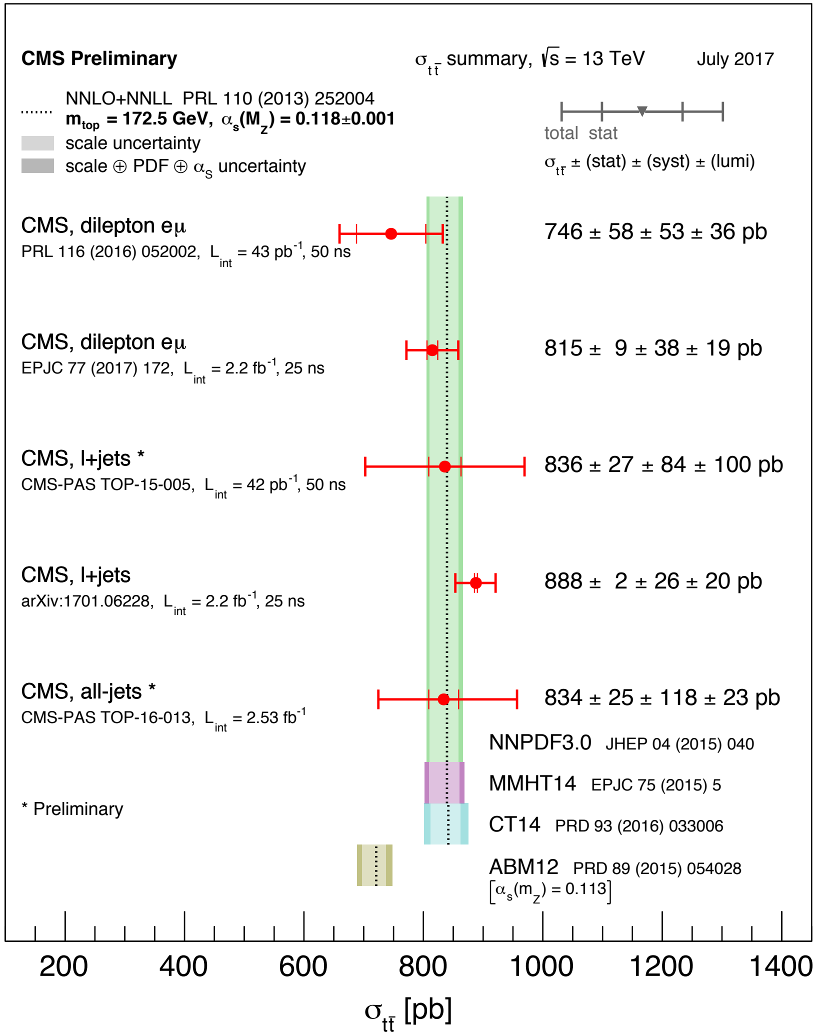
Università & INFN Genova

Introduction and motivations

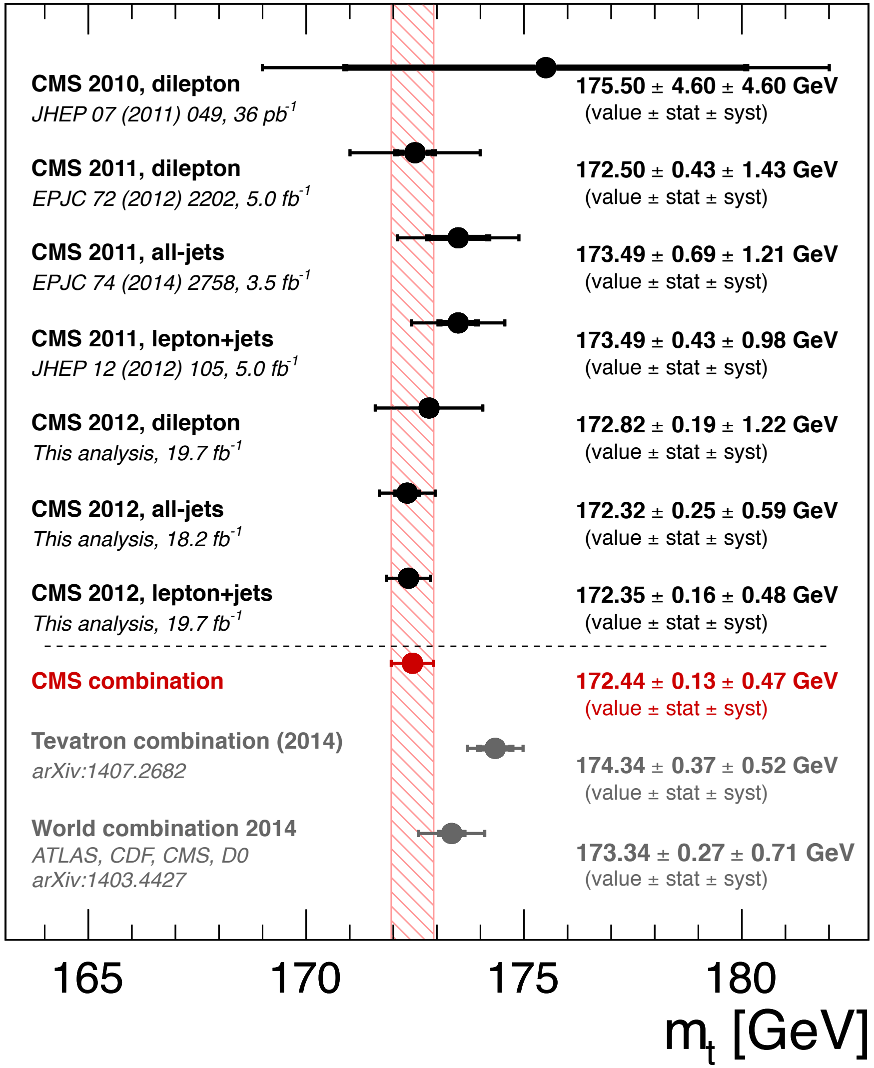
- With the accumulation of data, top quark physics has become precision physics
- Top quark measurements allow to obtain precision tests of the SM and in particular perturbative QCD predictions
  - Determine PDF more accurately
  - Derive fundamental parameters like m(top) or α_s
  - Improve accuracy of SM predictions and so have a better prediction in extreme phase-space regions used in NP searches
- This demands for improved experimental and theoretical uncertainties
  - Experimental uncertainties will improve with time (e.g. statistics, lepton ID, jet energy scale…)
  - This talk focuses on efforts to improve theoretical predictions

https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsTOP

CMS inclusive tt xsec measurements



CMS direct top mass measurements



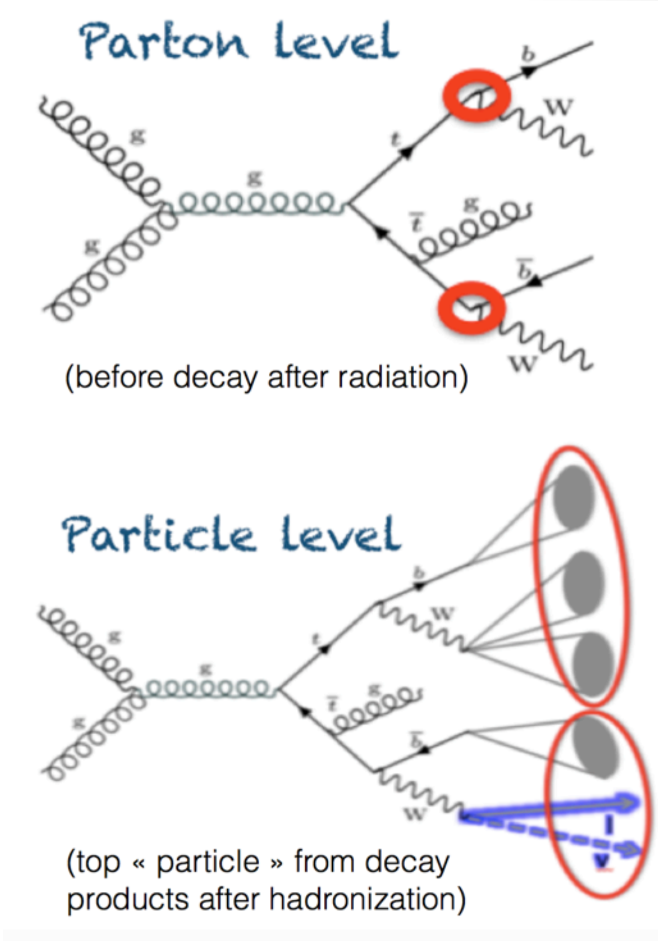
- Top physics has already reached a precision level, and cross sections are comparable to NNLO calculations!
  - Important systematic sources: b-quark hadronization modeling, modeling of hard scattering, top p_T modeling

CMS contributions

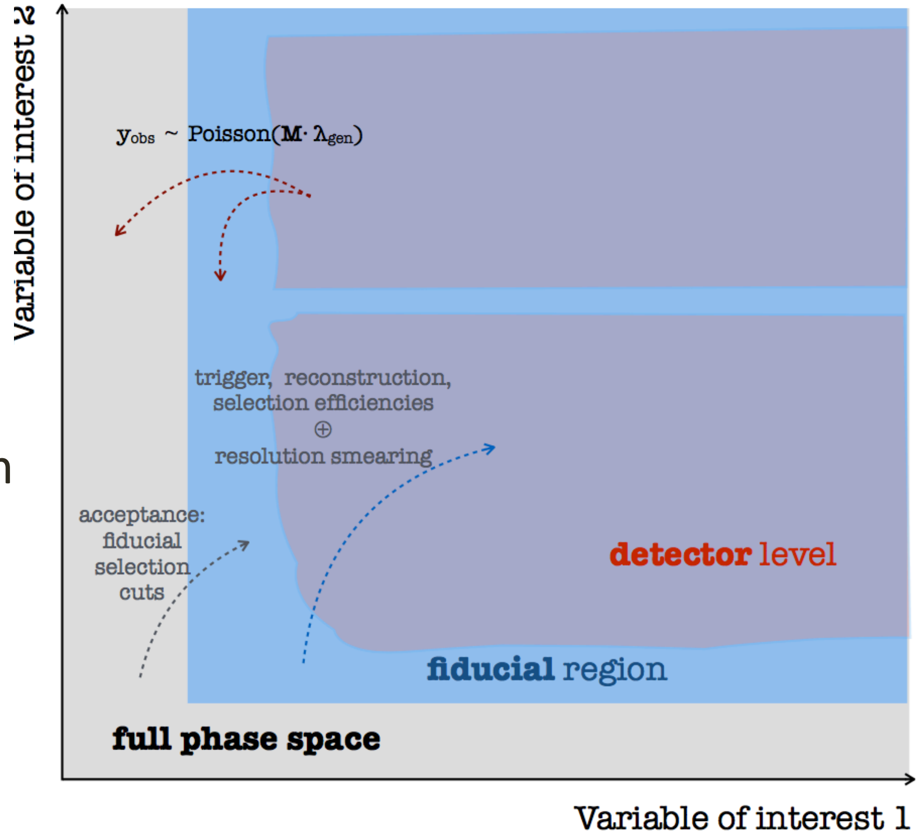
- Theoretical uncertainties can be (partially) tested and improved with measurements: e.g. hadronization, modelling of top kinematic quantities, extra jets etc
- Results presented in this talk are mostly based on 3 recent CMS works
- Object definitions for top quark analyses at the particle level (CMS Note2017_004)
- Underlying event measurement with tt+X events with p-p collision data at √s = 13 TeV (TOP-15-017)
- Investigations of the impact of the parton shower tuning in Pythia8 in the modelling of tt at √s = 8 and 13 TeV (TOP-16-021)
- These works use several CMS measurements of both Run1 and Run2

Object definition at particle level

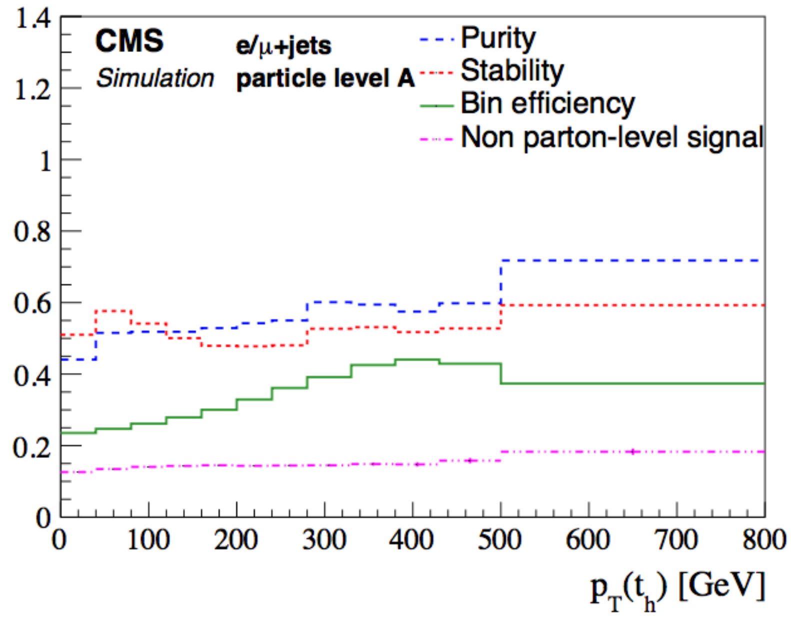
- Current MC generators
  - NLO matched to parton shower and take into account the finite width of the top
  - Finite width effects are needed to accurately predict off-shell production of top and interference with background
  - On-shell definition is ill defined: better use observables at particle level

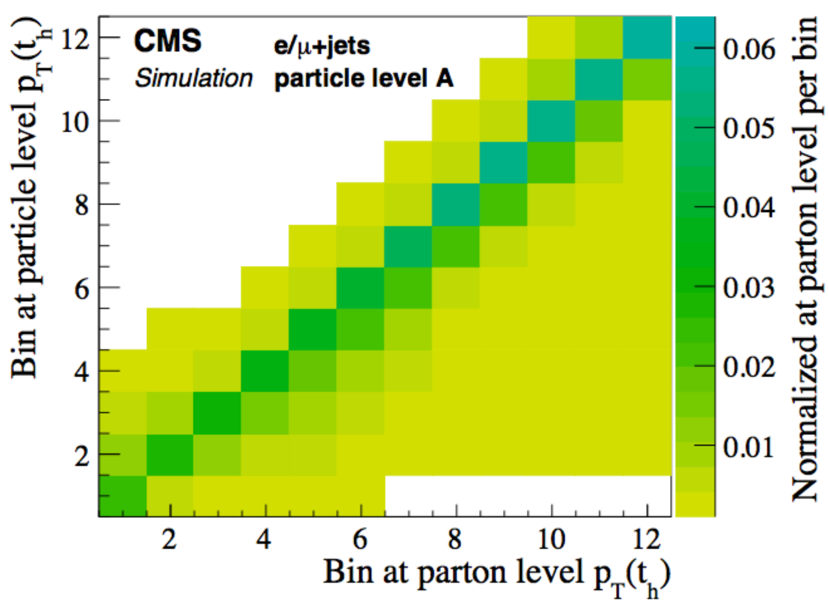


- From detector level quantities to parton/particle level quantities
  - Unfolding: Poisson regression model using a matrix derived from MC simulation
  - If the matrix is constructed from parton-level predictions it will be sensible to all theoretical effects (parton showering, factorization, hadronization, underlying event….)
  - Particle-level is less prone to theo uncertainties



- An important point for all future measurements of top differential cross-sections:
  - contruct tops only from observed final state (particle-level top)

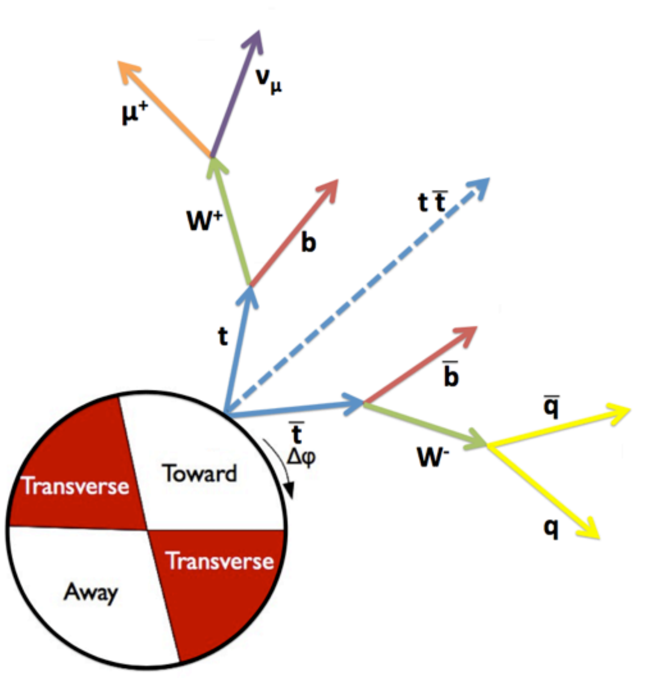




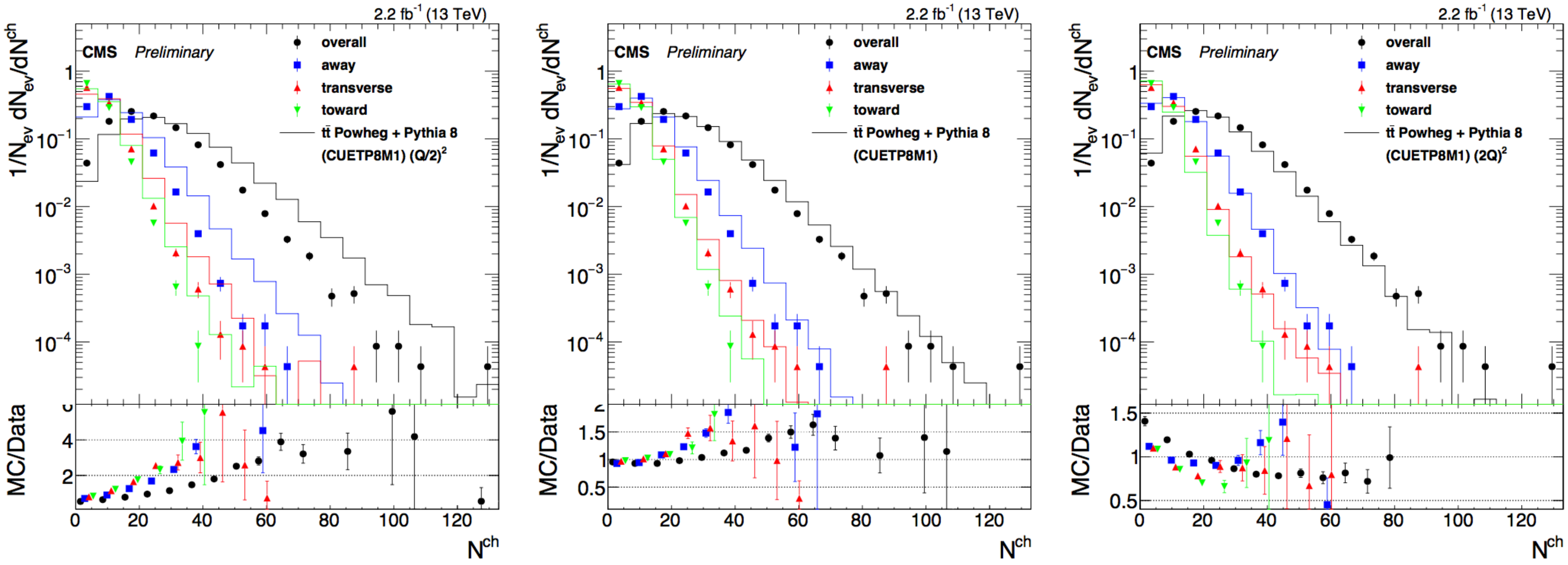
- Adoption of RIVET in the official CMS reco code for definition of particle-level objects
- Separate optimization of "pseudo-top" quarks to maximize sensitivity of future differential x-sec measurements
- Work in progress: future improvements expected

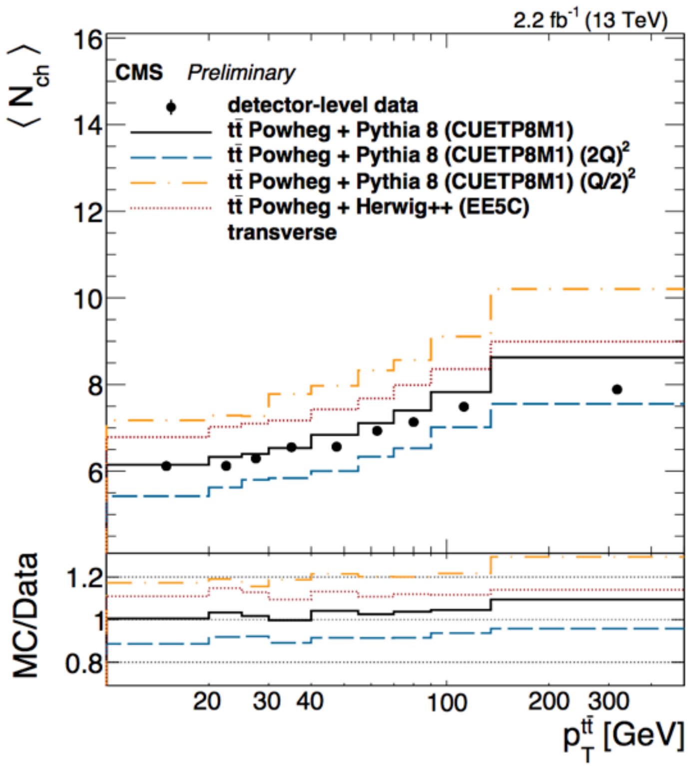
Underlying Event in tt+X

- UE: beam-beam remnants and multi-particle interaction. The relevant observables also have contributions from ISR and FSR
- Experimentally, UE is studied using topological structure of hard collisions
- The transverse regions are typically sensible to UE



- MC generators have sets of parameters to control the event modelling ("tune").



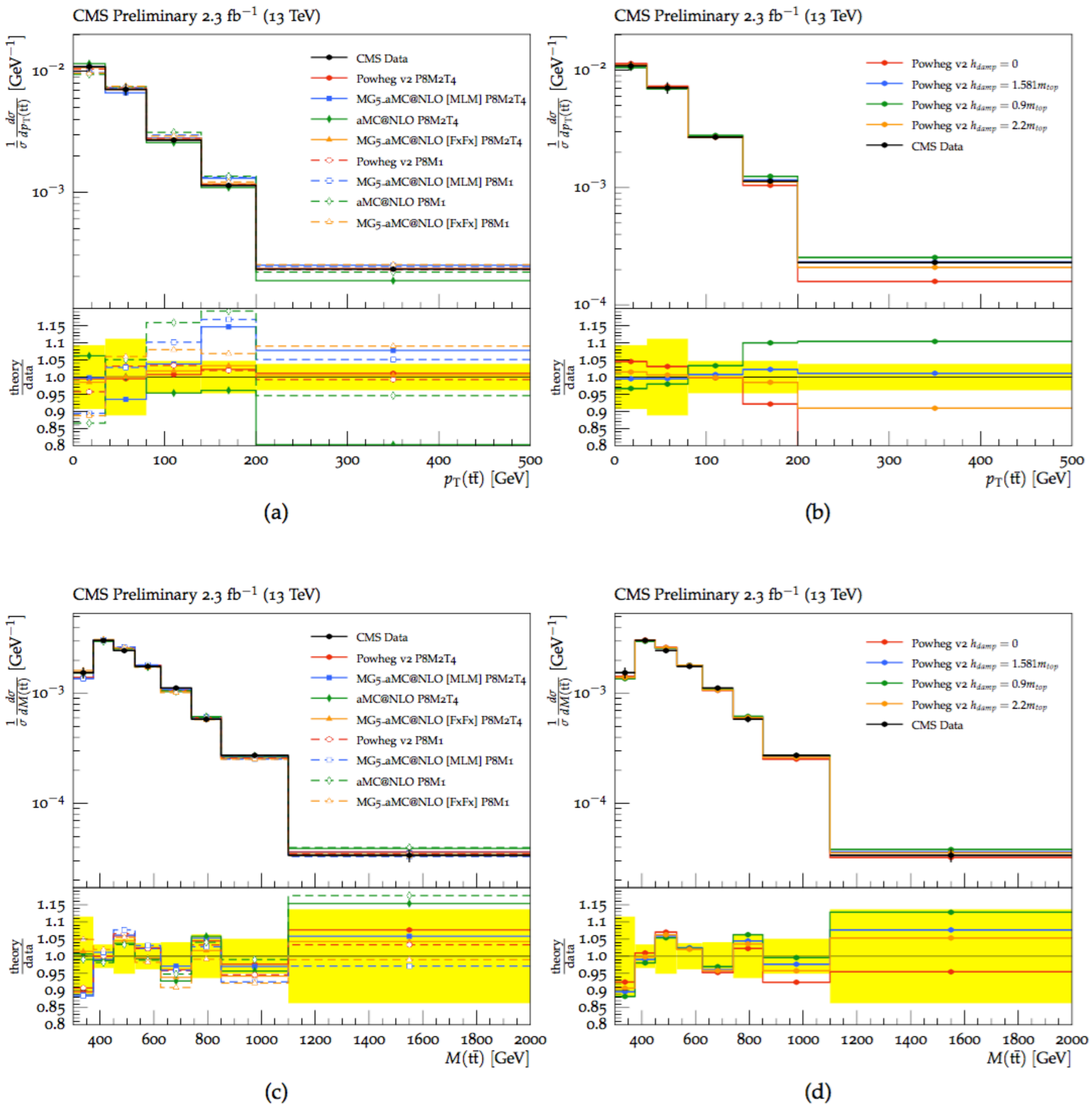


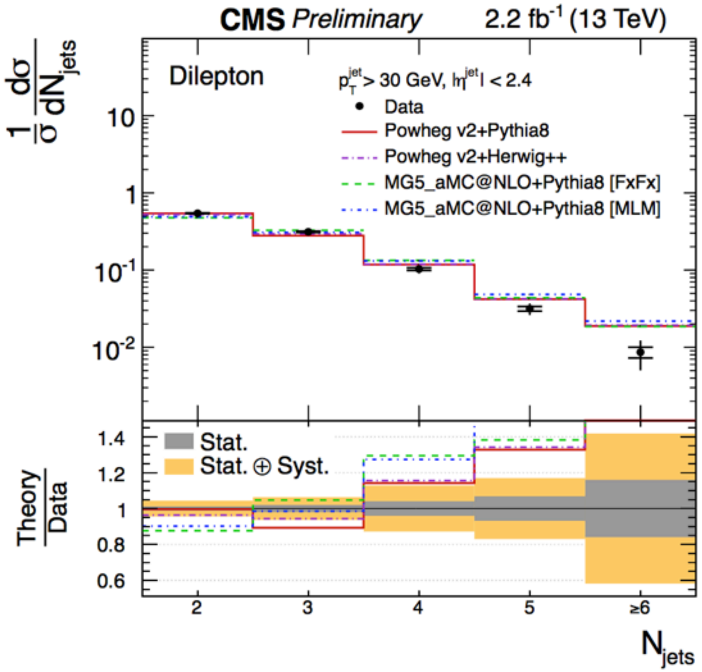
- UE activity measured using charged particle properties in tt events
- Results show there is no need for a specificheavy quark UE tune

PS tuning in Pythia8 for tt

- In Run1 CMS used LO MadGraph 5.1.3.30 interfaced with Pythia 6.426 for PS using the Z2* tune
  - Reasonably good description of most differential distribution, except for p_T(top).
- New simulations use NLO ME generators interfaced with new PS codes and updated tunes: better description expected as well reduction of theoretical uncertainties
- Uncertainties on renormalization and factorization scales at ME level are observed to provide good coverage of the discrepancies
  - Notable exception: modelling of extra jet multiplicity in tt
  - Important not only for top physics but also for Higgs physics and searches for NP

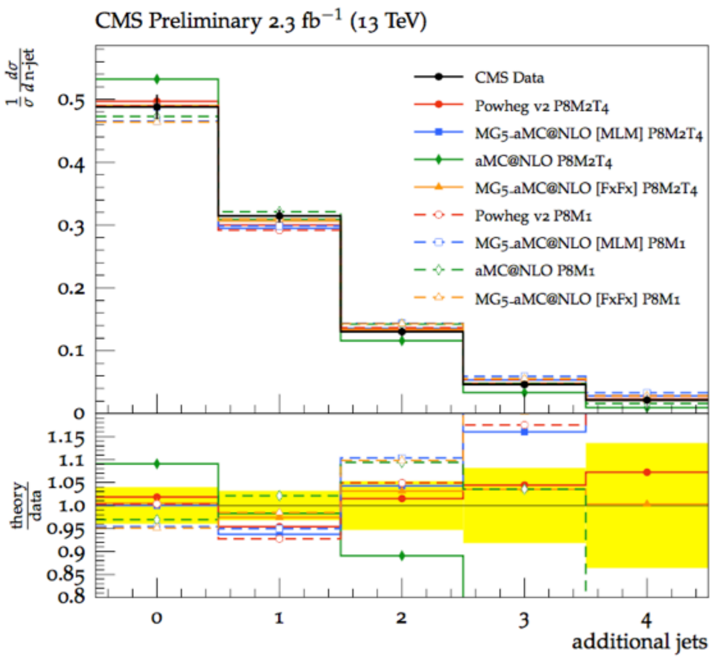
- Approach:
- Jet activity constrains parameters of probability of parton emission and interplay of hard and soft parton emission
  - No big constraint on UE
- A set of parameters is used to derive a new tune, that is tested on both Run 1 and 2 data
- This is meant to become the nominal setup to calibrate top physics in Run 2
  - New Pythia8 tune with a lower α_s value + a new h_{damp} value for PowHeg improves the description of tt kinematic variables and overall description
  - Caveat for NP searches: no bias is expected for searches based on missing E_T but searches based on other variables should proceed carefully

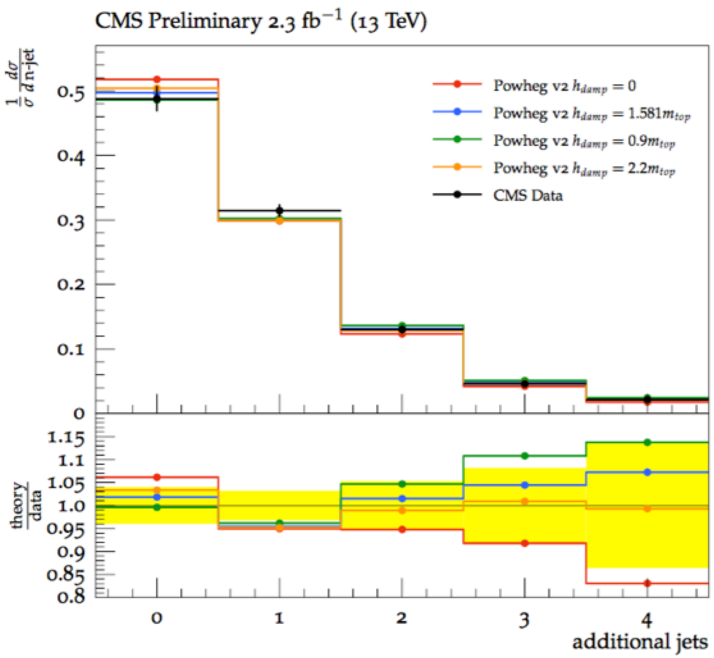

CMS Preliminary 2.3 fb^{-1} (13 TeV)
CMS Data
Powheg v2 P8M2T4
MG5_aMC@NLO [MLM] P8M2T4
aMC@NLO P8M2T4
MG5_aMC@NLO [FxFx] P8M2T4
Powheg v2 P8M1
MG5_aMC@NLO [MLM] P8M1
aMC@NLO P8M1
MG5_aMC@NLO [FxFx] P8M1
Powheg v2 $h_{damp} = 0$
Powheg v2 $h_{damp} = 1.581 m_{top}$
Powheg v2 $h_{damp} = 0.9 m_{top}$
Powheg v2 $h_{damp} = 2.2 m_{top}$
$\frac{1}{\sigma}\frac{d\sigma}{dp_T(t\bar{t})}$ [GeV^{-1}]
$\frac{1}{\sigma}\frac{d\sigma}{dM(t\bar{t})}$ [GeV^{-1}]
$\frac{\text{theory}}{\text{data}}$
$p_T(t\bar{t})$ [GeV]
$M(t\bar{t})$ [GeV]
(a)
(b)
(c)
(d)



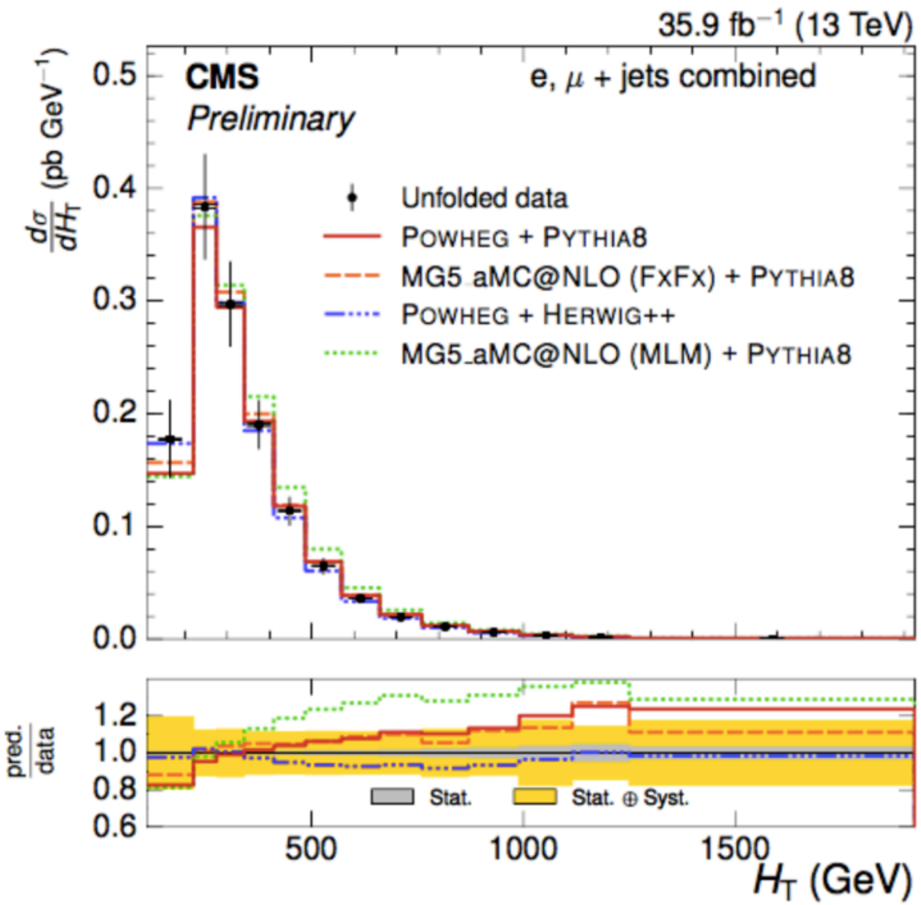
Predictions are larger than data at high multiplicity with old tune parameters (α_s=0.1365)

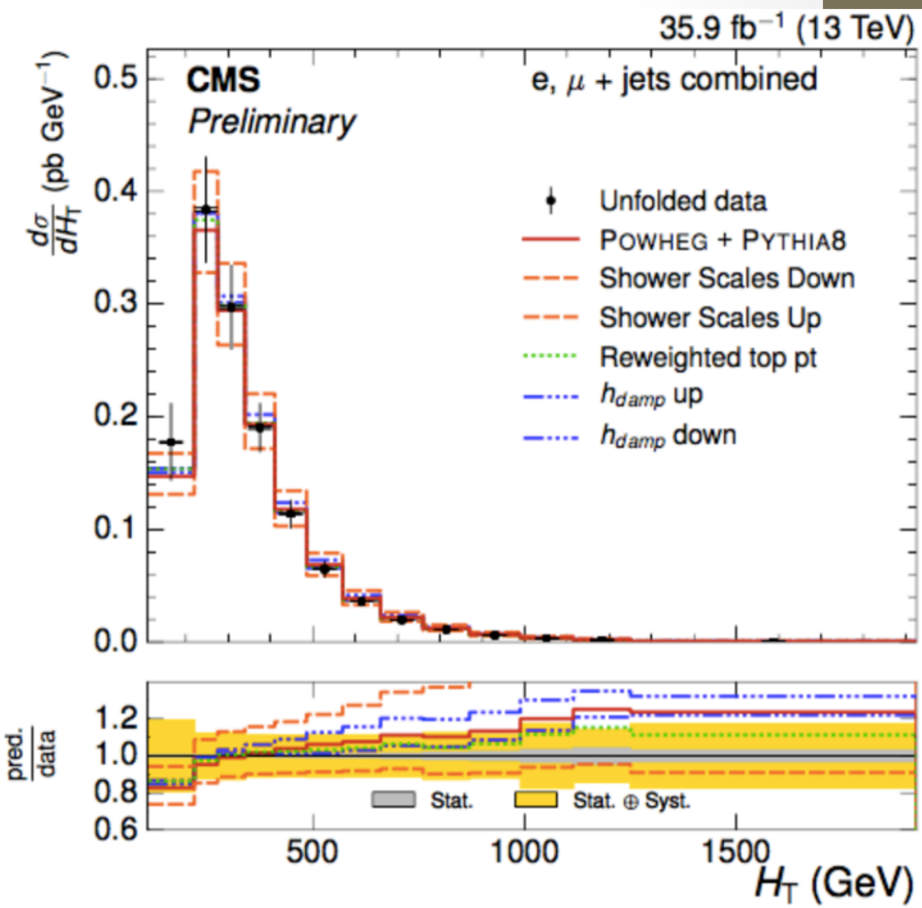
α_s and h_{damp} retuned using data: significantly lower value describes data well, α_s=0.1108 and h_{damp}=1.581 x m(top)





Performance of the new tune on 13 TeV data





TOP-16-014

Conclusions

- Top physics plays a very important role in precision tests of SM and NP searches
  - And more data is still accumulating!
- Inclusive xsections are already better than NNLO accuracy and the mass is determined to a 0.5 GeV uncertainty!
  - Need to attack all sources of uncertainties, including theory-related uncertainties
- Dedicated measurements and theory studies are needed to improve modelling and reduce uncertainties
  - CMS has undertaken several efforts