Top quark pair production with additional jets (heavy flavor) from CMS and ATLAS

Tae Jeong Kim (Hanyang University)

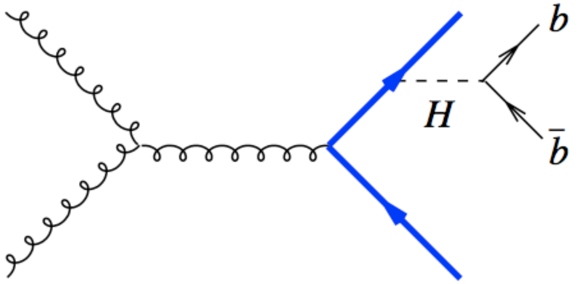
For Heavy Flavour Production workshop

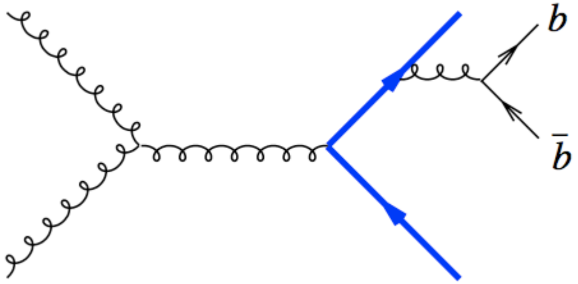
at the LHC (HF@LHC) workshop at IPPP, Durham, UK

September 6 in 2017

Motivation of $t\bar{t}$ + additional jets

- $t\bar{t}b\bar{b}$ is the main irreducible background for $t\bar{t}H(b\bar{b})$ searches
  - NLO calculation still has an uncertainty of around 30%.
- $t\bar{t}jj(c\bar{c})$ is the reducible background faking b jets
- It is crucial to understand precisely the $t\bar{t}jj$ and $t\bar{t}b\bar{b}$ processes as these are also the main background for most of new physics searches such as four top search, FCNC, SUSY,…
- In particular, differential distributions will allow us to check the validity of the QCD calculation involving top quark pair plus additional quarks or gluons. (two different scales)





MC samples at 13 TeV

- ATLAS

Event generator	Parton shower	
POWHEG (v2)	Pythia 6 / Pythia 8	Default
POWHEG	Herwig++/Herwig7	
MadGraph5_aMC@NLO	Pythia 8 / Herwig++/Herwig7	
SHERPA	SHERPA	

- CMS

Event generator	Parton shower	
POWHEG (v2)	Pythia 8	Default
POWHEG	Herwig++	
MadGraph5_aMC@NLO	Pythia 8 with FxFx ($t\bar{t}$ + 0, 1, 2j)	
MadGraph5 (LO)	Pythia 8 with MLM ($t\bar{t}$ + 0, 1, 2, 3j)	
MadGraph5_aMC@NLO	Herwig++	

$t\bar{t}$+heavy flavour events are extracted from inclusive $t\bar{t}$ sample

Particle-level objects

- Decrease MC uncertainty from extrapolating to unmeasurable phase space
- Particle-level objects

CMS NOTE-2017/004
based on LHCTopWG

  - Electrons and muons : p_T > 25 GeV and $|\eta|$ < 2.5
    - not originate from a hadron, Adding the four-momentum of all photons within $\Delta R = 0.1$
  - Jets : p_T > 25 GeV and $|\eta|$ < 2.5
    - clustering all stable particle except the selected e, μ and radiated photons as well as neutrinos using the anti-k_t algorithm with R=0.4.
    - Neutrinos from hadron decay are included
  - b jets : ghost matching technique - b hadron momentum is scaled down to a negligible value and included in the jet clustering.

- Particle level top quark
  - Take the jet permutation by minimizing following quantity in the lepton + jets mode.

$$K^2 = [M(p_\nu + p_\ell + p_{b_\ell}) - m_t]^2 + [M(p_{j_1} + p_{j_2}) - m_W]^2 + [M(p_{j_1} + p_{j_2} + p_{b_h}) - m_t]^2$$

Event categorization

- Particle level signal definition for $t\bar{t}$+heavy flavor

CMS event categorization (dilepton)

Visible phase space

$t\bar{t}jj$: $n_{leptons} = 2, n_{b-jets} \geq 2$ and $n_{jets} \geq 4$

$t\bar{t}b\bar{b}$: $t\bar{t}jj + n_{b-jets} \geq 4$

$t\bar{t}bj$: $t\bar{t}jj + n_{b-jets} = 3$

$t\bar{t}c\bar{c}$: $t\bar{t}jj + n_{c-jets} \geq 2$

$t\bar{t}$LF : $t\bar{t}jj - t\bar{t}b\bar{b} - t\bar{t}bj - t\bar{t}c\bar{c}$

Full phase space

$t\bar{t}jj$: $n_{jets\ not\ from\ top} \geq 2$

$t\bar{t}b\bar{b}$: $t\bar{t}jj + n_{b-jets\ not\ from\ top} \geq 2$

ATLAS event categorization

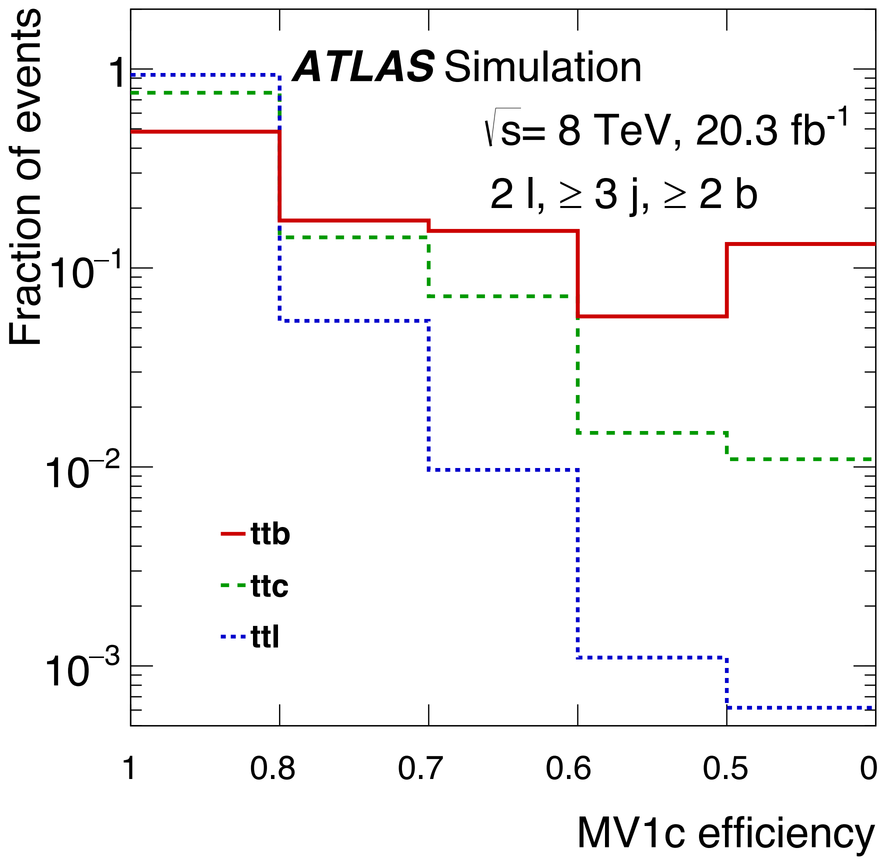
Shorthand notation for the templates	Particle-level event requirements
	ttb lepton-plus-jets
ttb	$n_{\text{leptons}} = 1$, $n_{\text{jets}} \geq 5$ and $n_{b-\text{jets}} \geq 3$
ttc	$n_{\text{leptons}} = 1$, $n_{\text{jets}} \geq 5$ and $n_{b-\text{jets}} = 2$ and $n_{c-\text{jets}} \geq 1$
ttl	other events
	ttb $e\mu$
ttb	$n_{\text{jets}} \geq 3$ and $n_{b-\text{jets}} \geq 3$
ttc	$n_{\text{jets}} \geq 3$ and $n_{b-\text{jets}} \leq 2$ and $n_{c-\text{jets}} \geq 1$
ttl	other events
	ttbb dilepton fit-based
ttbb	$n_{\text{jets}} \geq 4$ and $n_{b-\text{jets}} \geq 4$
ttbX	$n_{b-\text{jets}} = 3$
ttcX	$n_{b-\text{jets}} = 2$ and $n_{c-\text{jets}} \geq 1$
ttlX	other events

has to rely on MC mother particle information

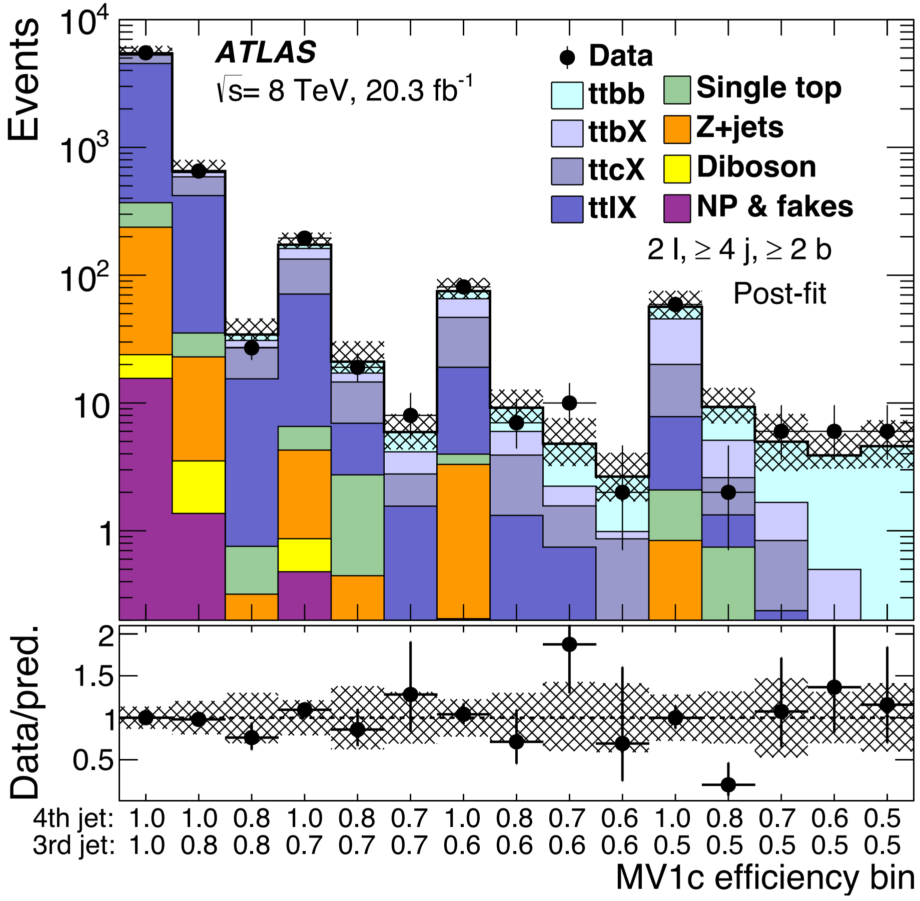
$t\bar{t}$ + Heavy flavor measurement (ATLAS)

Eur. Phys. J. C76 (2016) 11

Third highest MV1c jet



Post-fit

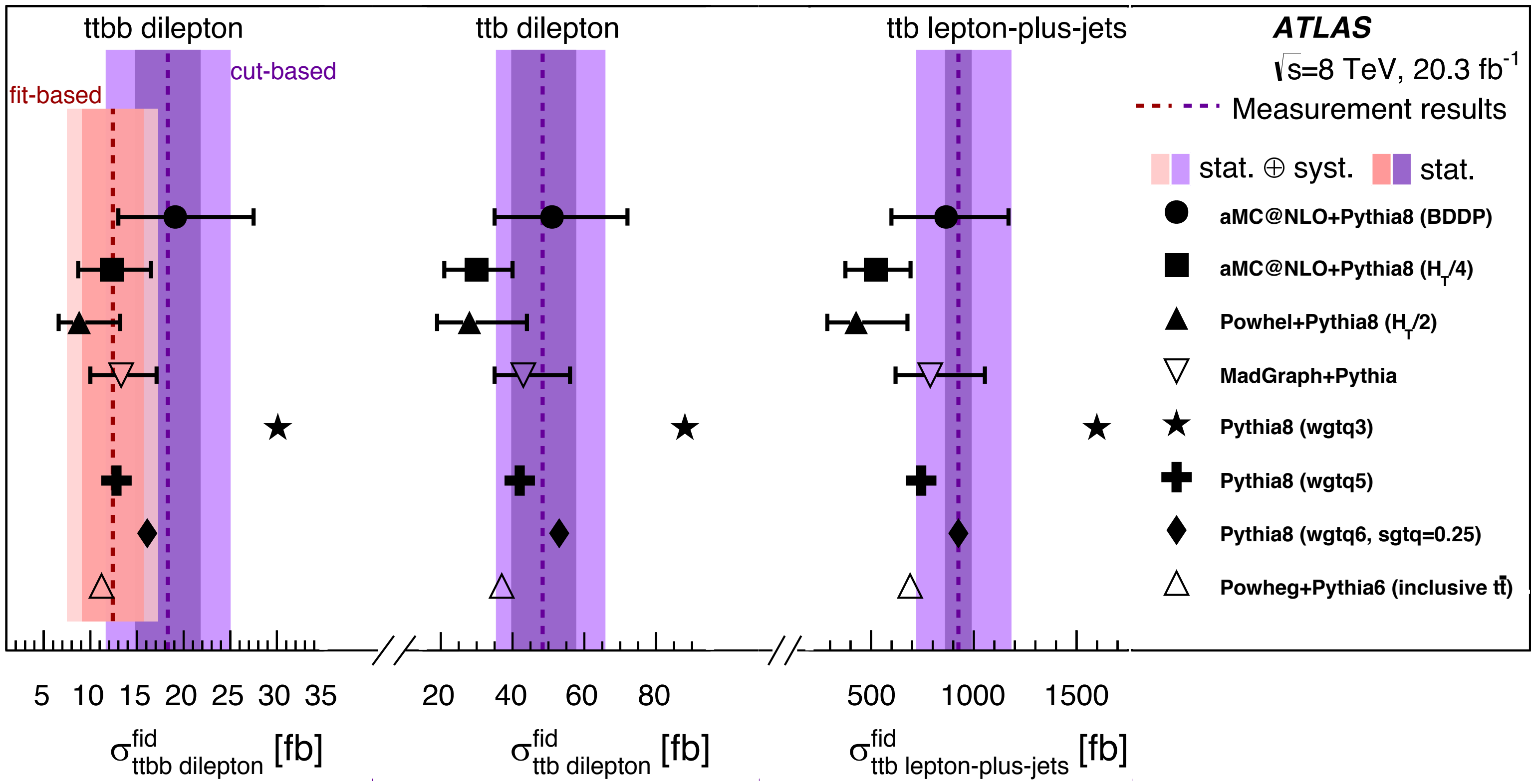


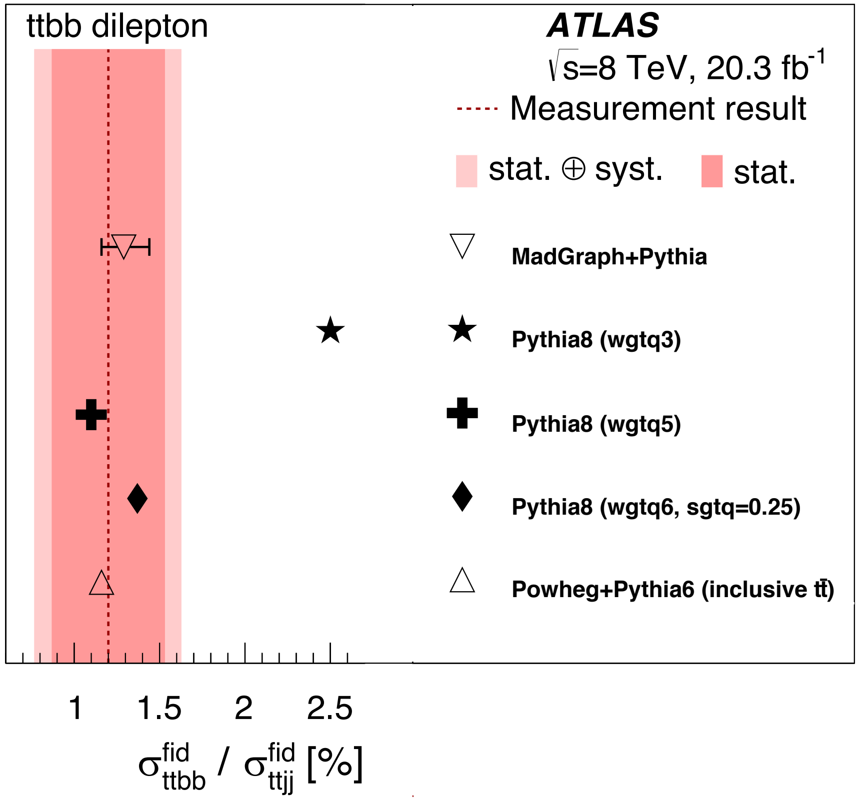
- Using the third and fourth highest MV1c b tagging discriminant labelled with the upper edge of the efficiency, it has the significant shape differences between the $t\bar{t}$ components.

$t\bar{t}$ + Heavy flavor measurement (ATLAS)

Eur. Phys. J. C76 (2016) 11

- Compared with various samples in the massive 4F scheme (MG5_aMC@NLO) and 5F scheme (POWHEL)
- Fiducial phase space regions
- $t\bar{t}H$ and $t\bar{t}V$ contributions are removed for direct comparison

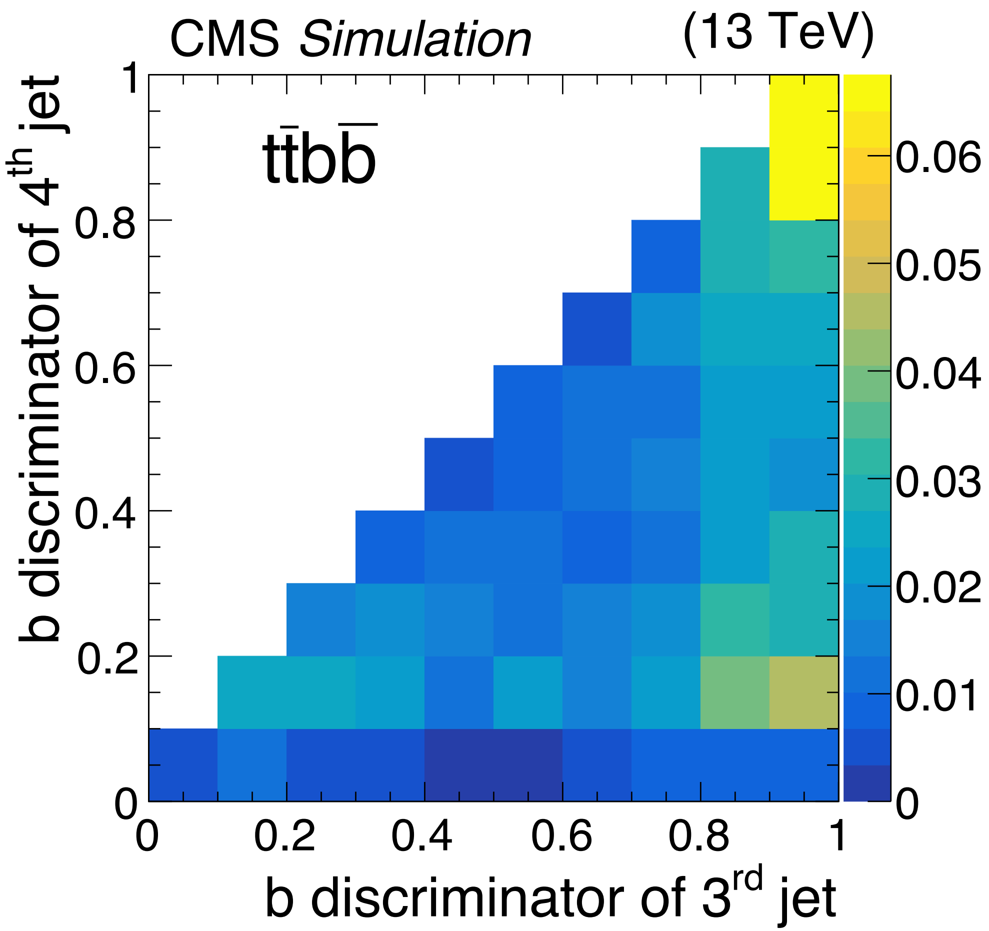


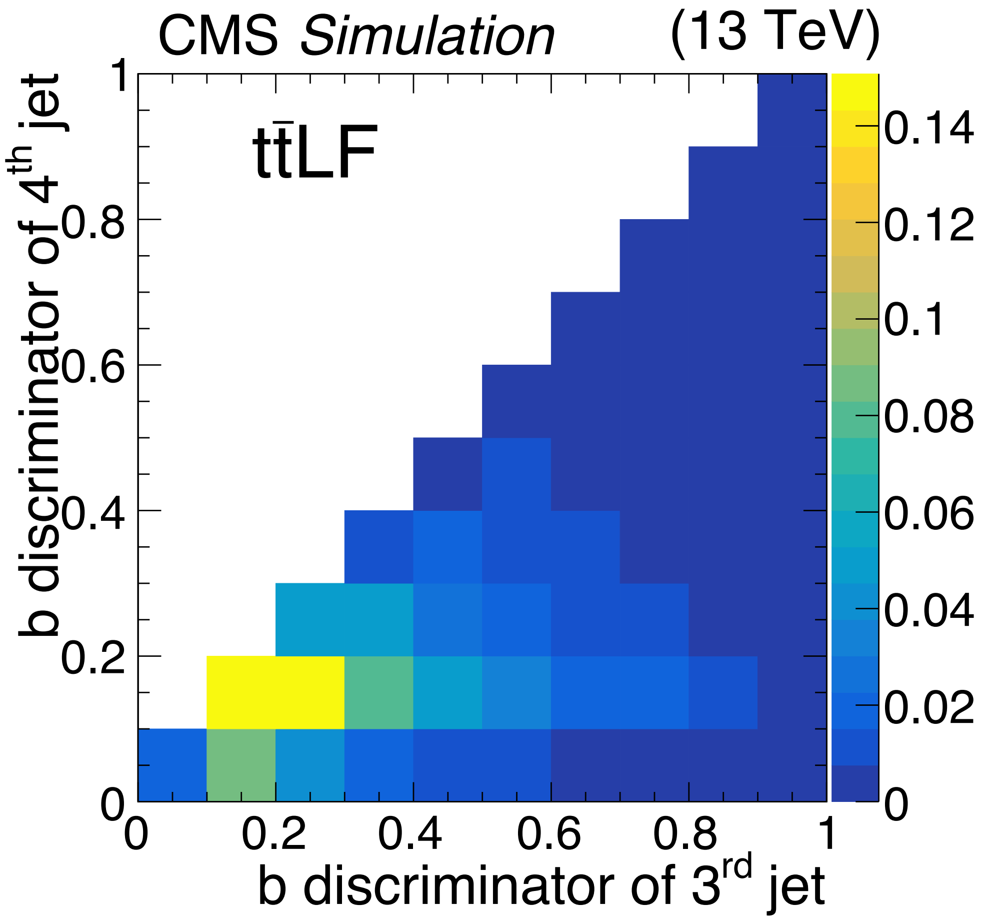


$t\bar{t}$ + Heavy flavor measurement (CMS)

arXiv:1705.10141

- Rearrange jets in b-tagging algorithm discriminator
- Using b-tagging algorithm discriminator of third and fourth jets in 2D fitting.



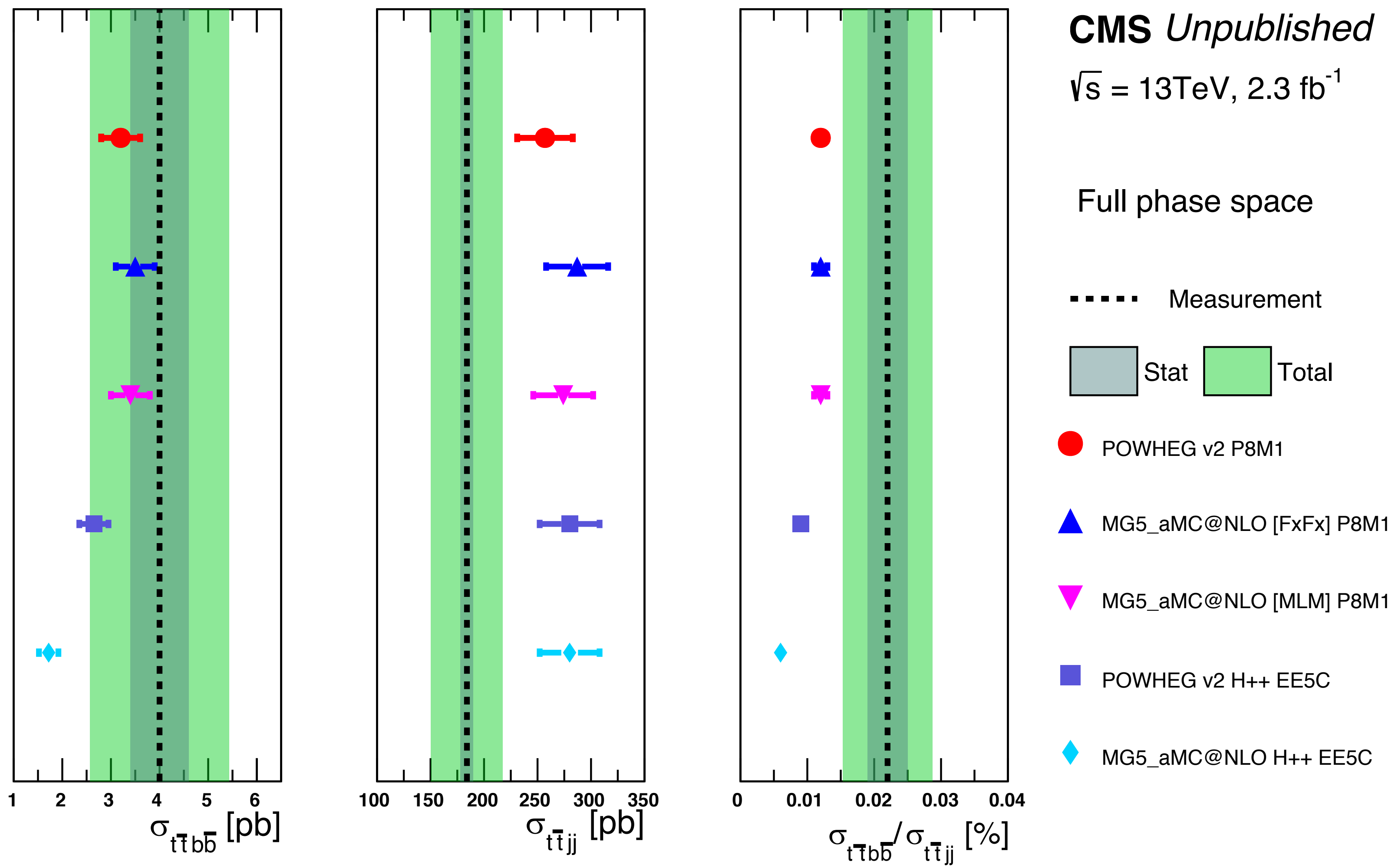


$t\bar{t}$ + Heavy flavor measurement (CMS)

- Full phase space

no cut on the top quark decay products and additional b jet p_T > 20 GeV

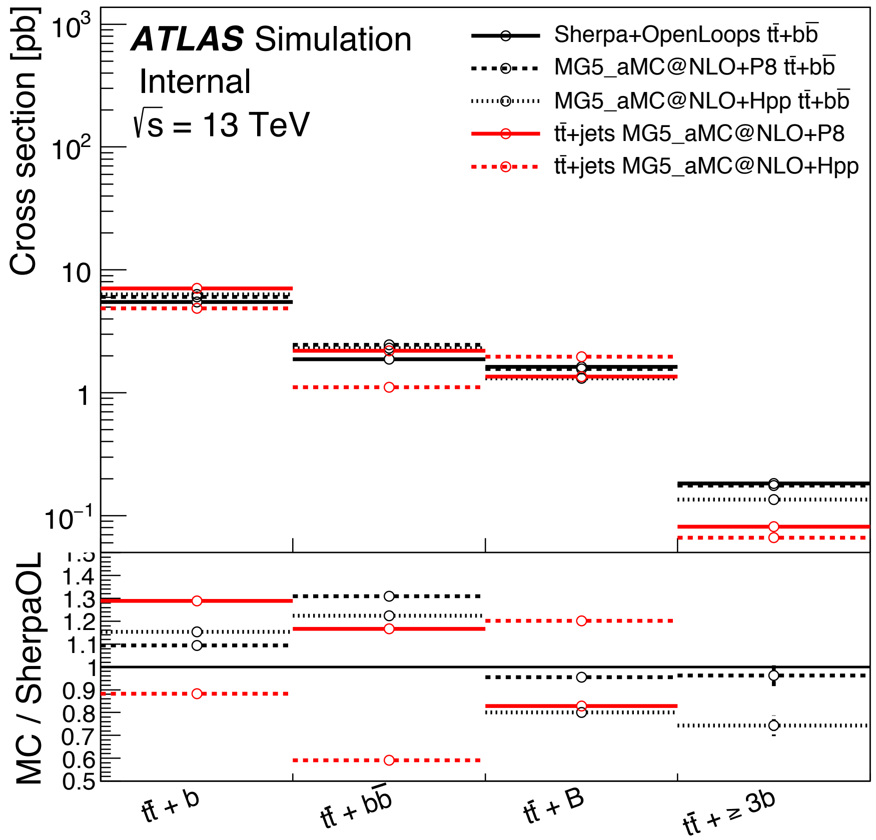
arXiv:1705.10141

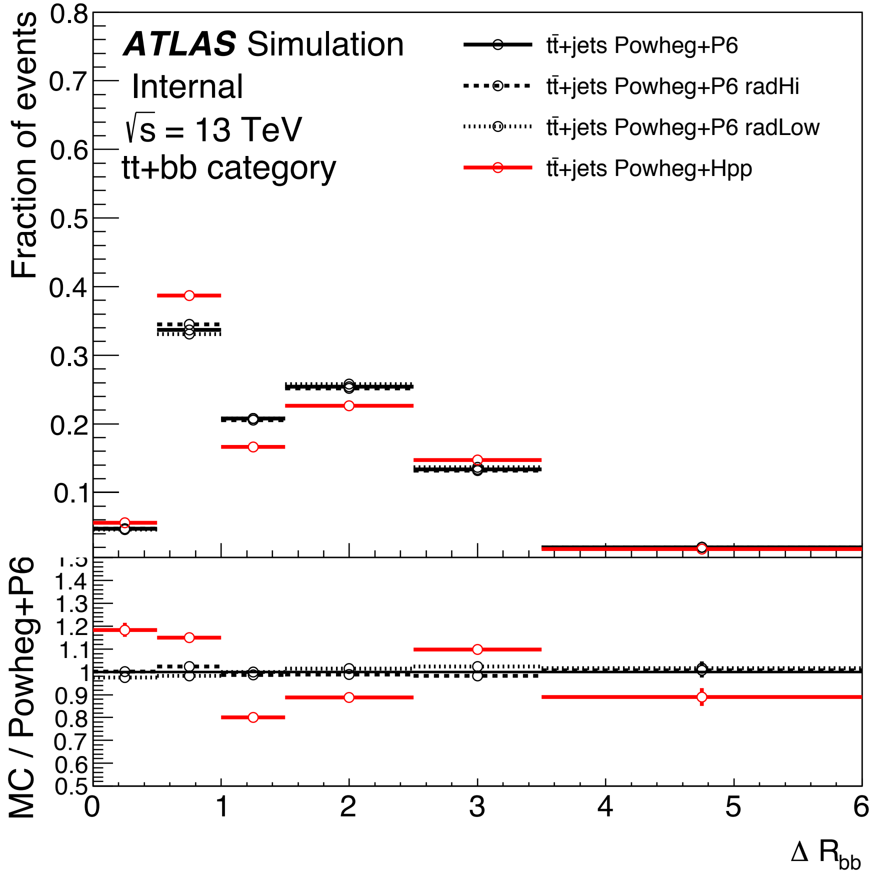


- Theoretical ratios are lower than the measured values
- But consistent within two standard deviations
- $t\bar{t}H$ are not removed

Differential $t\bar{t}b\bar{b}$ cross sections (ATLAS)

- b jets are identified using $\Delta R < 0.4$ matching with a B-hadron
- NLO predictions with 4F scheme (massive b-quarks) are compared
  - Sherpa + OpenLoops, MG5_aMC@NLO + Herwig++/Pythia8

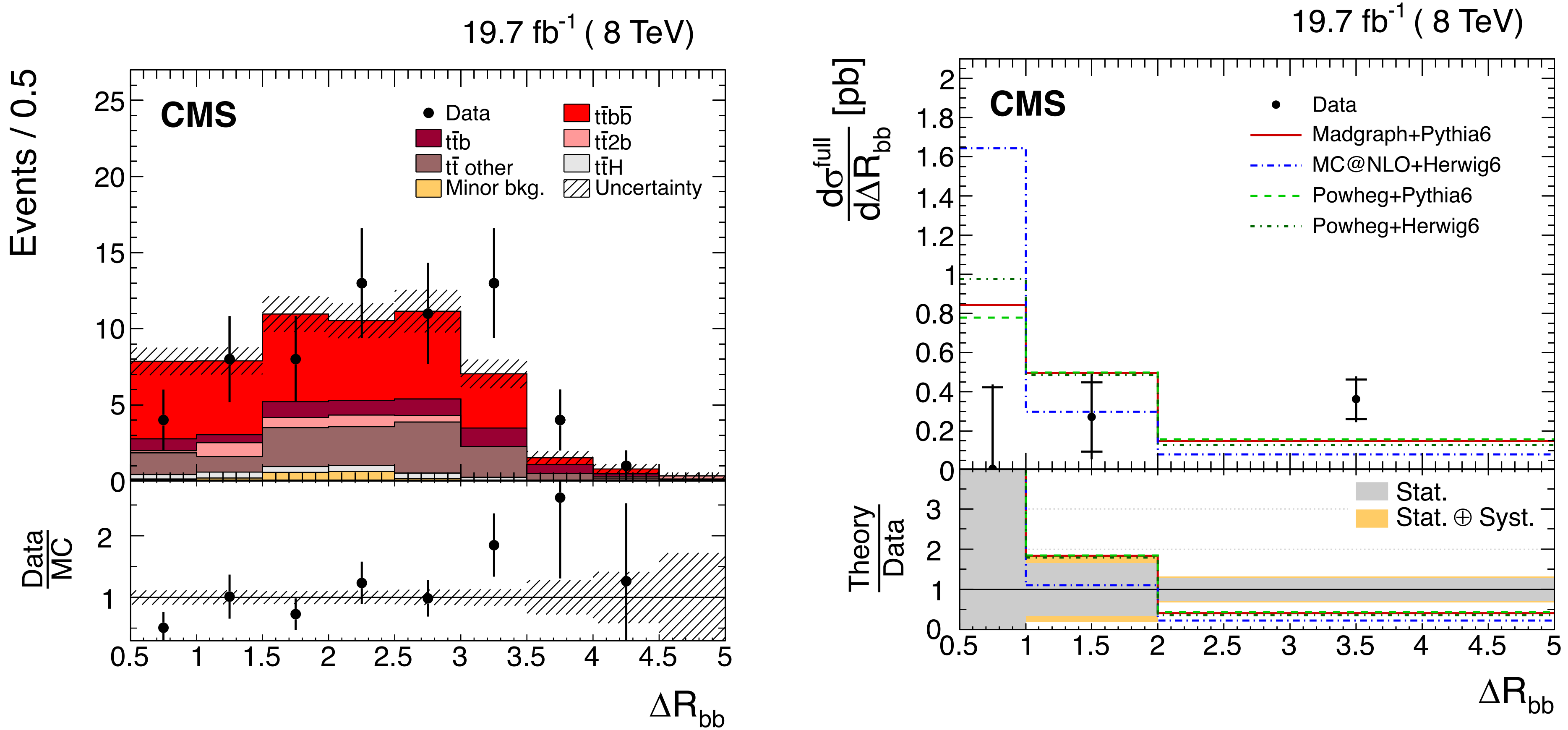




Differential $t\bar{t}b\bar{b}$ cross sections (CMS)

$$\Delta R = \sqrt{\Delta\phi(b,b)^2 + \Delta\eta(b,b)^2}$$

8 TeV | Eur. Phys. J. C 76 (2016) 379



- Correct assignment is difficult in case of 4 b jets
  - typically ~50% correct assignment in case of 4 b jets
- Need more statistics

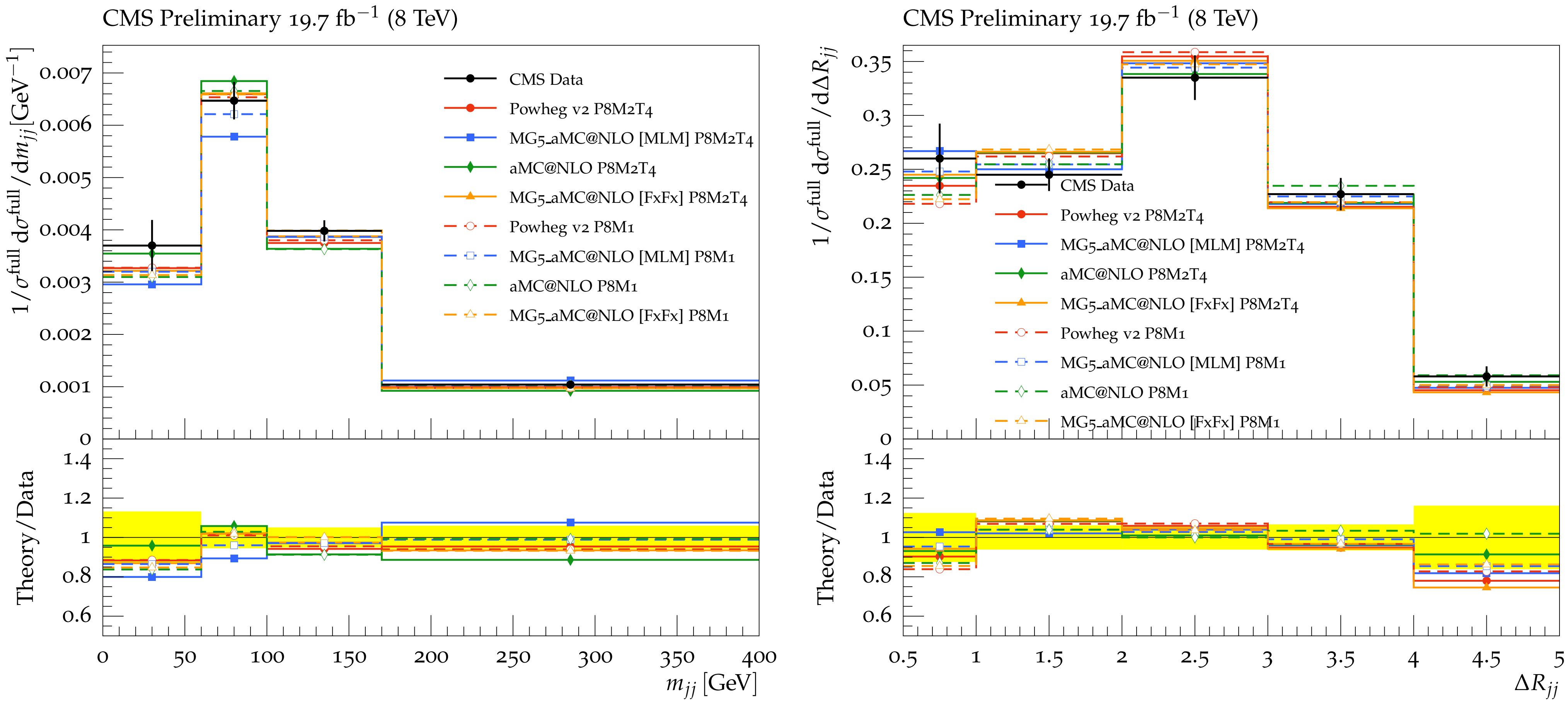
ΔR_{jj} and m_{jj} distribution of additional jets (CMS)

- Normalized $t\bar{t}$ cross sections in bins of invariant mass and the angular distance of the leading and second leading additional jets

8 TeV

CMS-PAS-TOP-16-021

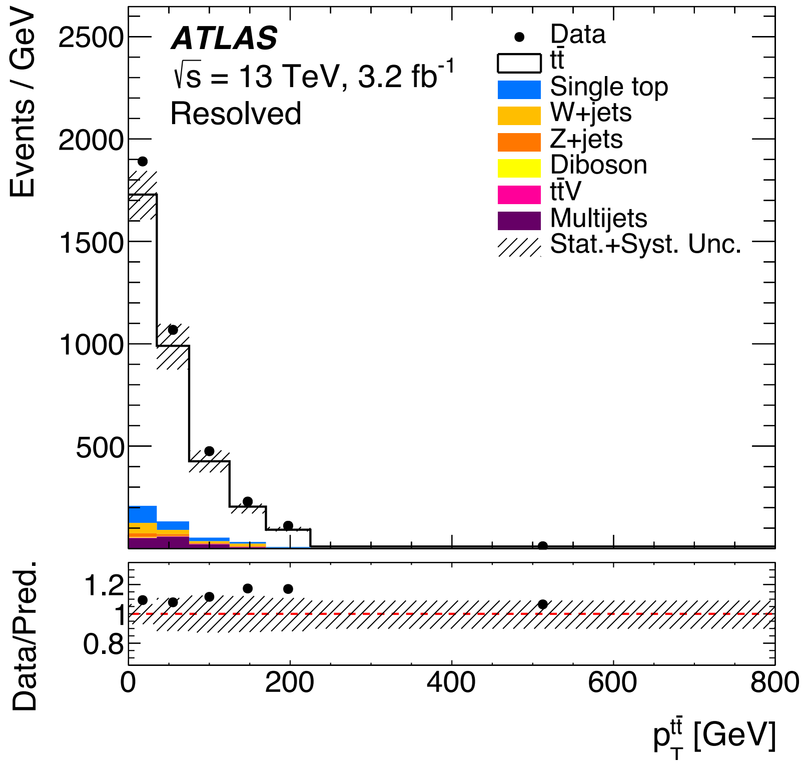
Data is from Eur. Phys. J. C 76 (2016) 379

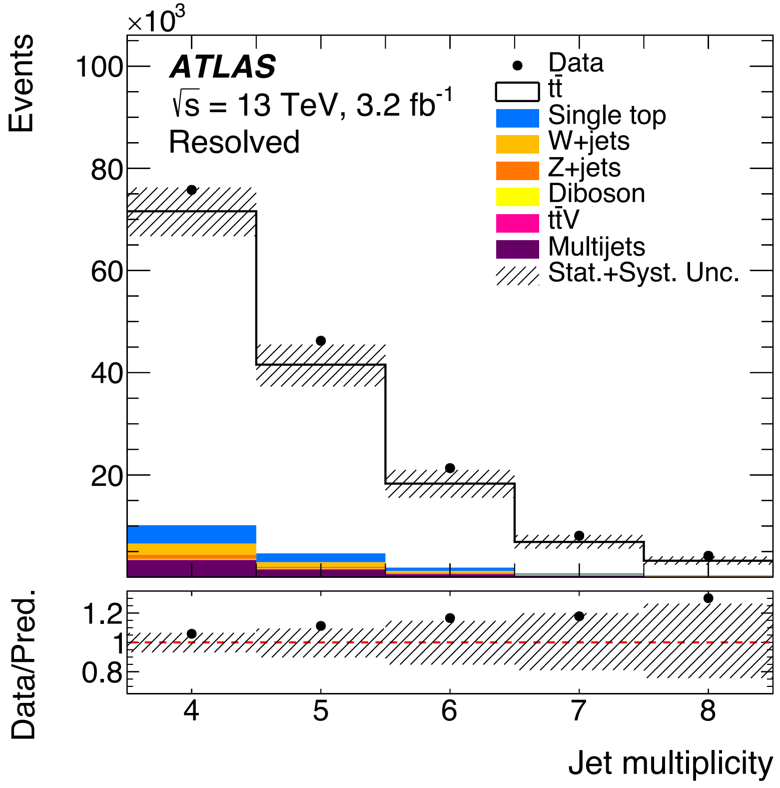


$t\bar{t}$ + additional jets at 13 TeV (ATLAS)

arXiv:1708.00727

- Recoiling objects depend on the transverse momentum of $t\bar{t}$ system
- Resolved topology in the combined *l+jets* channel at detector level





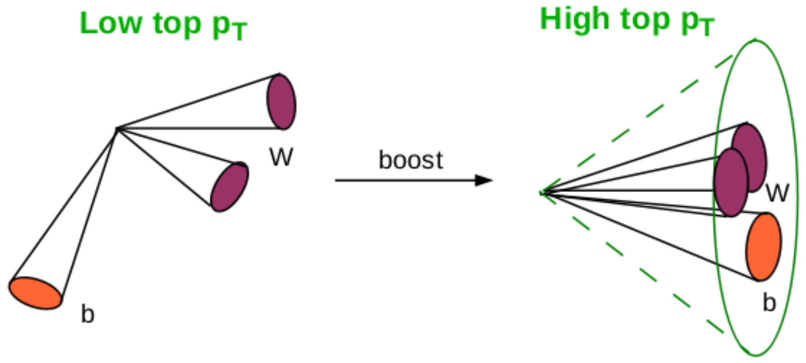
- Overall prediction underestimates data
- The largest uncertainty comes from JES/JER and flavour tagging

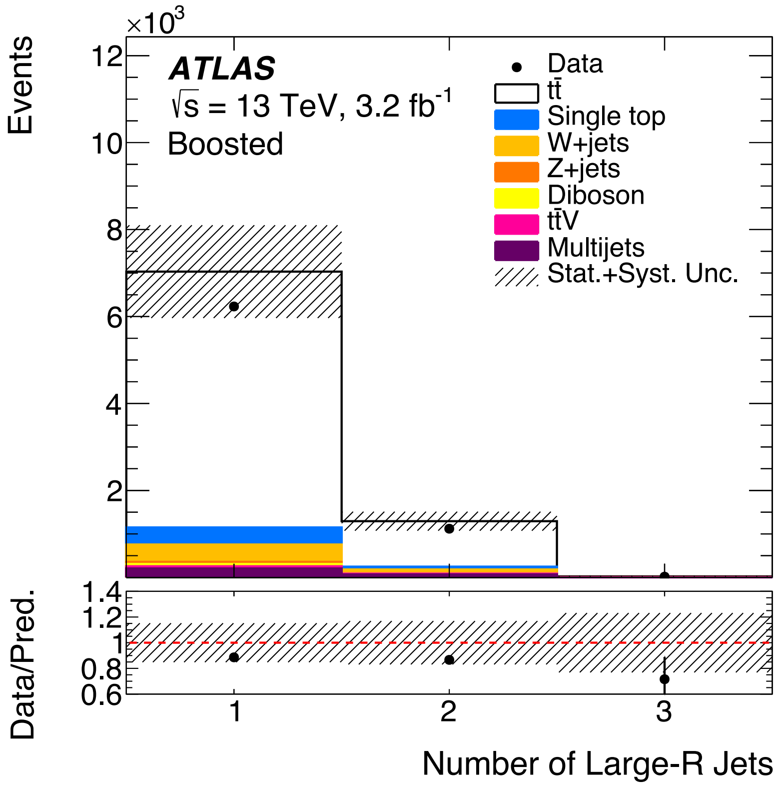
Highly boosted top in hadronic channel (ATLAS)

Boosted objects

large-R jet : 300 GeV < p_T < 1500 GeV, m > 50 GeV and $|\eta|$<2.

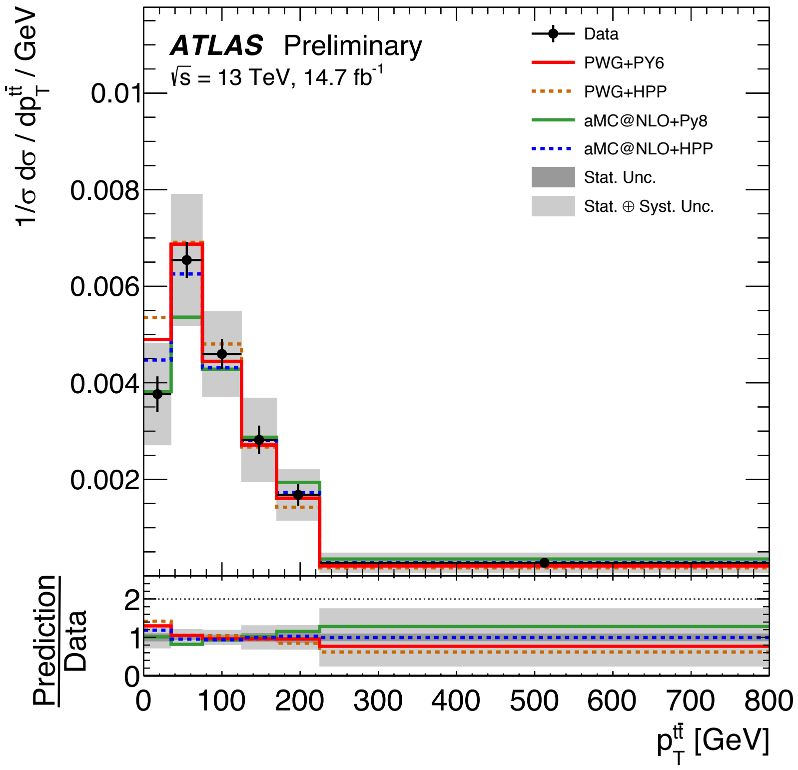
top-tagged jets : if m(large-R jet) > 100 GeV, τ_{32} < 0.75





arXiv:1708.00727

l+jets



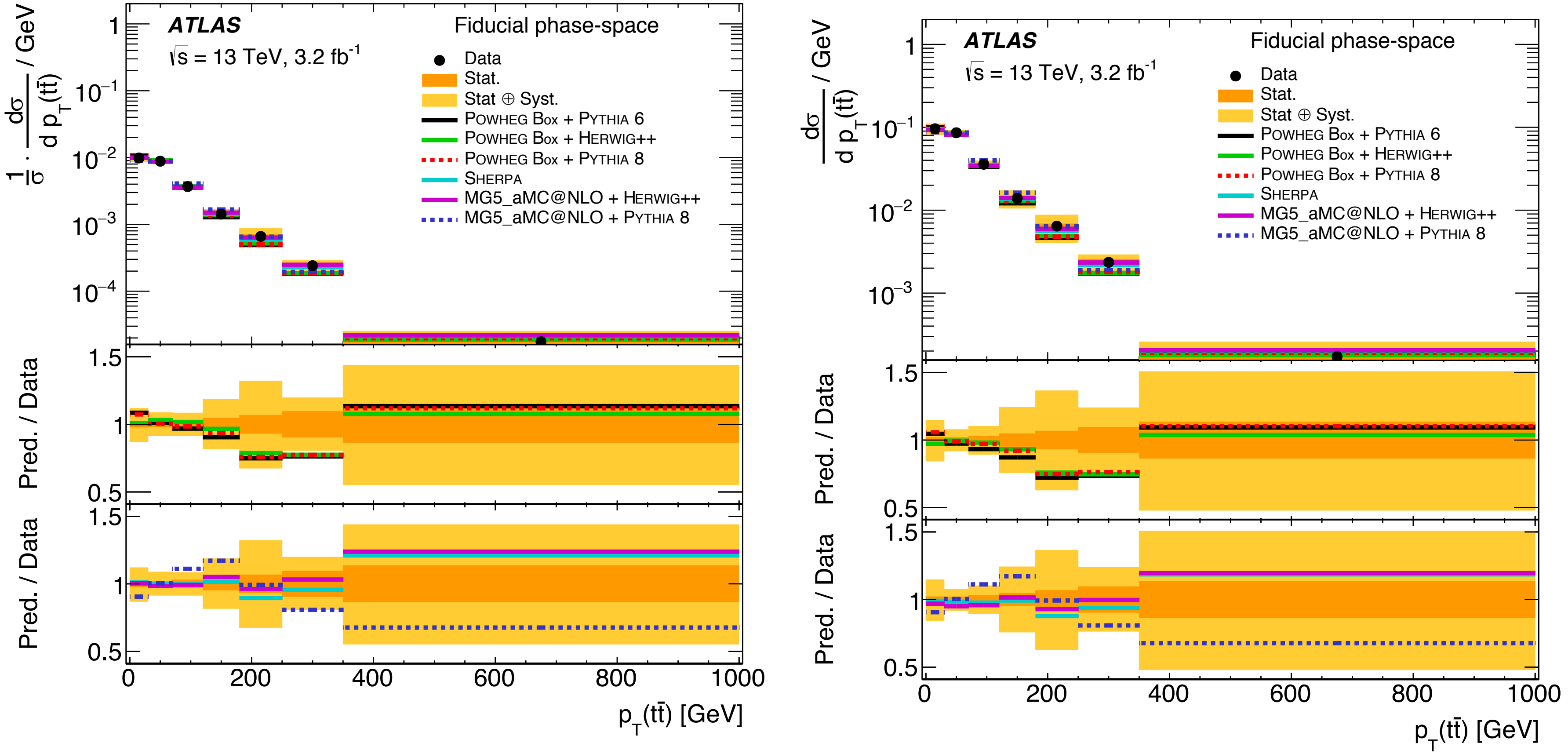
ATLAS-CONF-2016-100

- All hadronic channel
- p_T^1 > 500 GeV, p_T^2 > 350 GeV
- Boosted top quark S/B ~ 3
- Conceivable to pursue more detailed study for high p_T SM

$p_T(t\bar{t})$ spectrum comparisons (ATLAS)

- electron and muon of opposite sign, at least two jets, one b jet

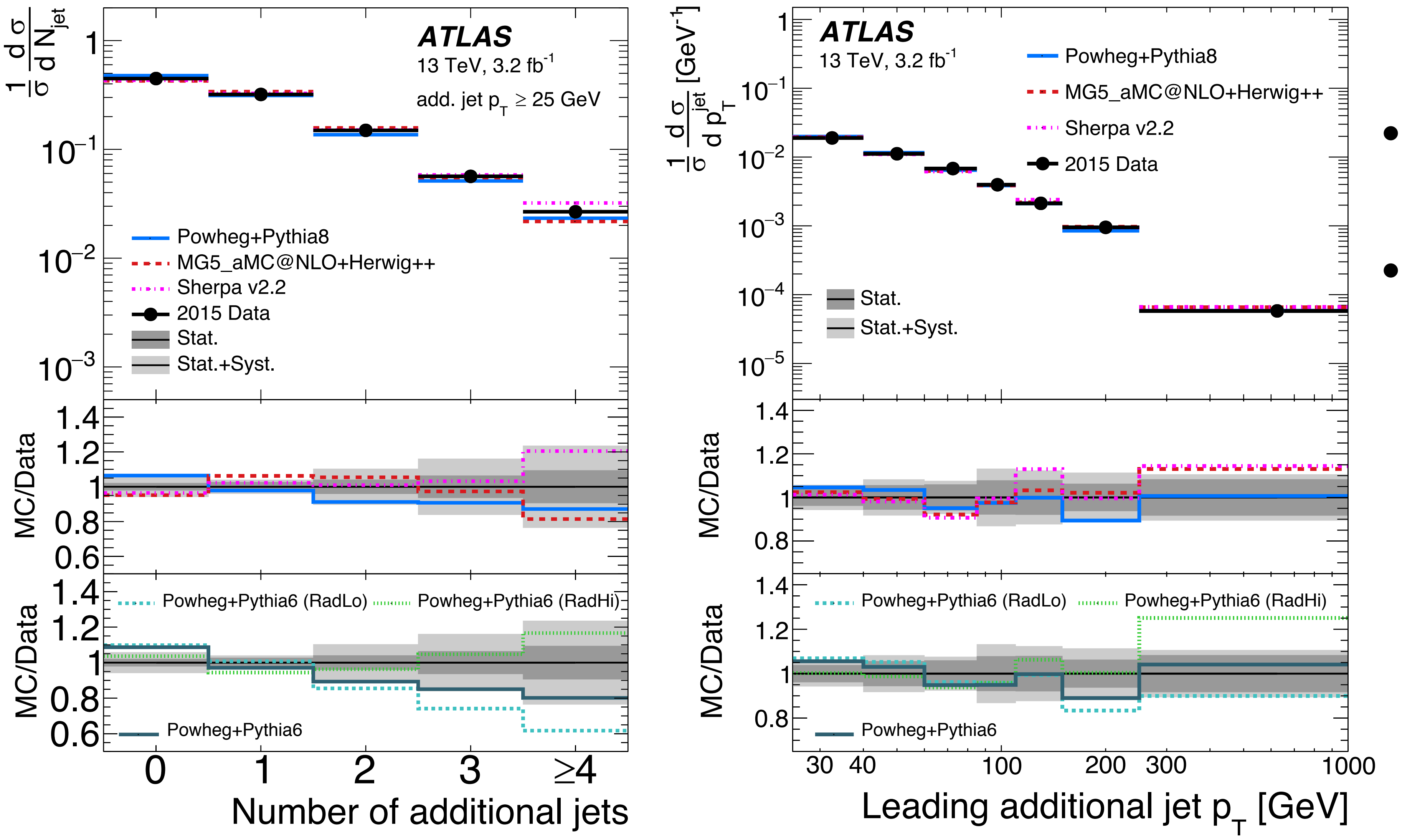
Eur. Phys. J. C77 (2017) 299



- POWHEG + Pythia 6 and Pythia 8 predict a softer spectrum than the data but consistent within the experimental uncertainties
- MG5_aMC@NLO, POWHEG + Herwig++ do not describe data well.

$t\bar{t}$ + additional jets at 13 TeV (ATLAS)

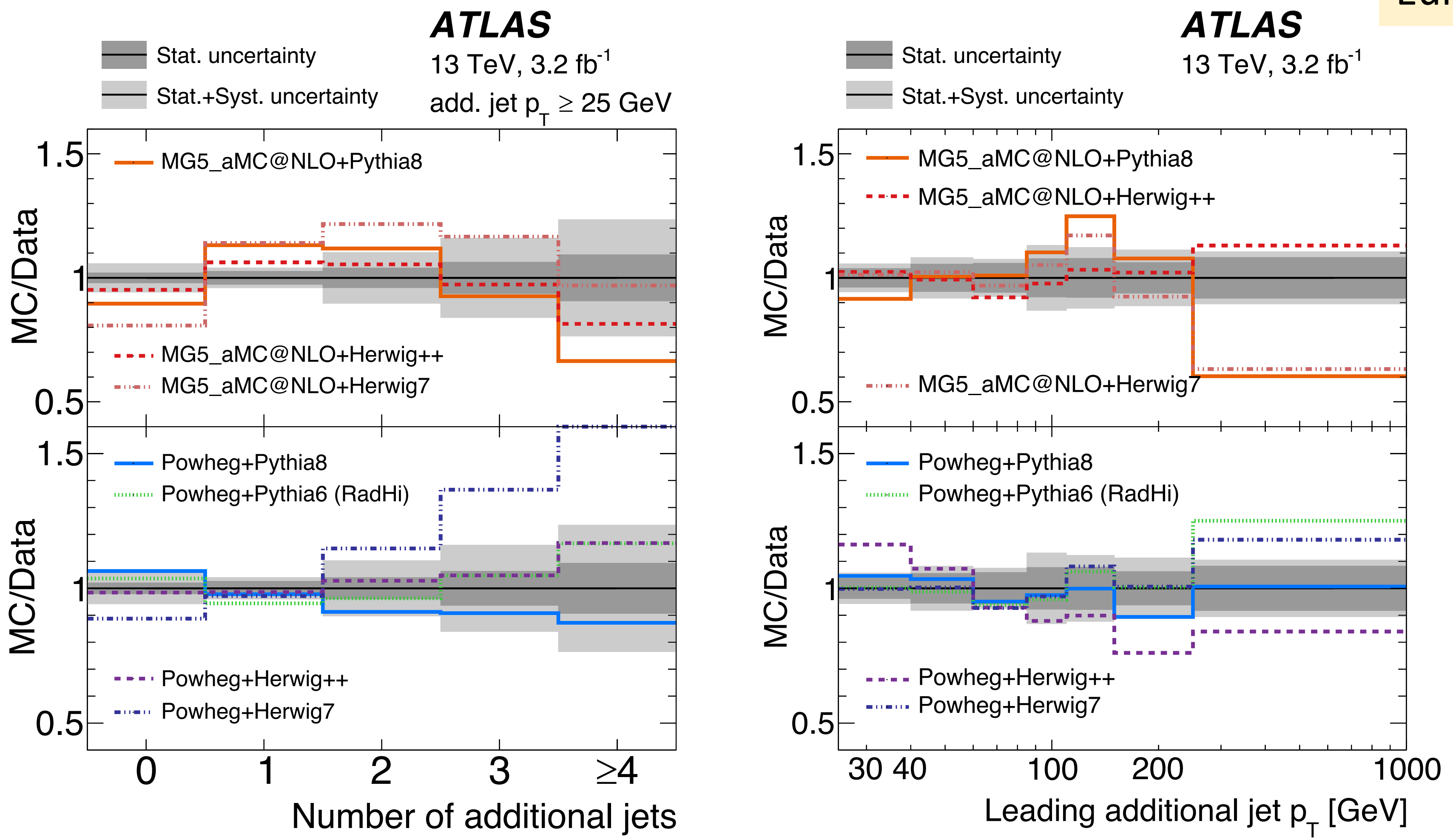
Eur. Phys. J. C77 (2017) 220



- Sherpa overshoots data
- Powheg+Pythia6 (low radiation variation of the Perugia 2012 tune) is a way lower than data in high multiplicity

$t\bar{t}$ + additional jets at 13 TeV (ATLAS)

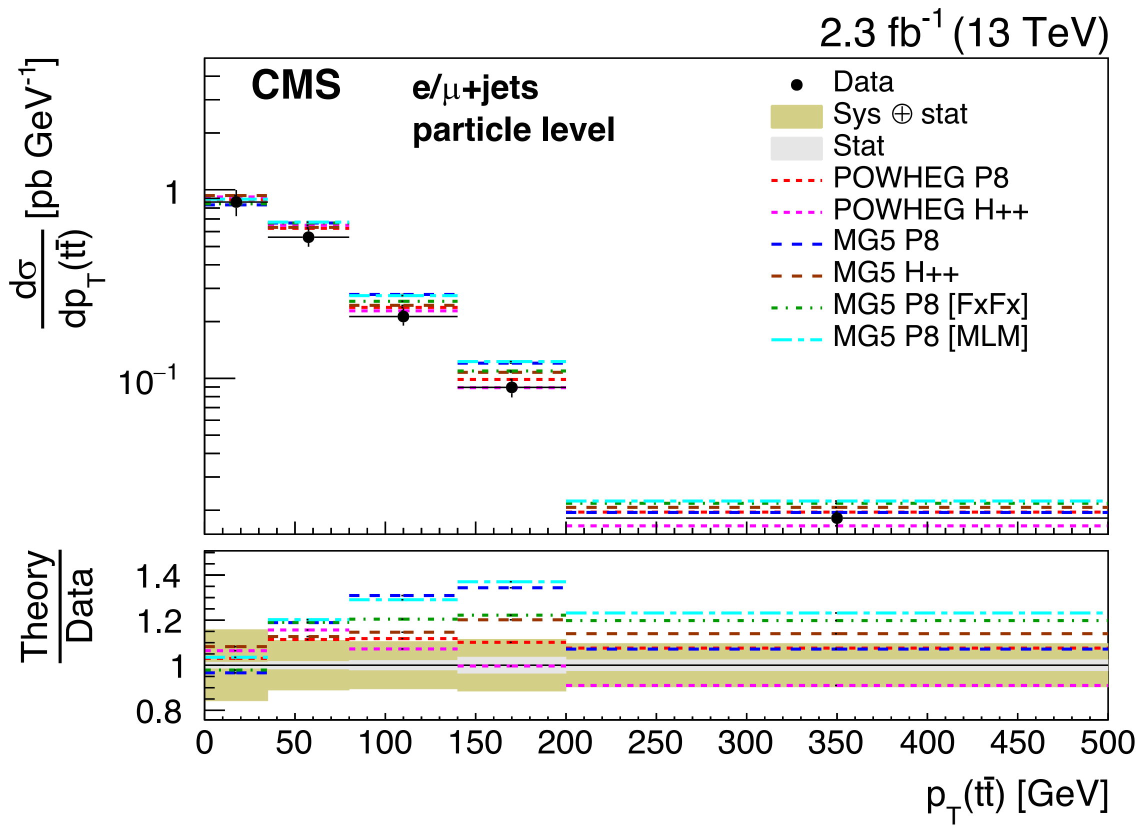
Eur. Phys. J. C77 (2017) 220

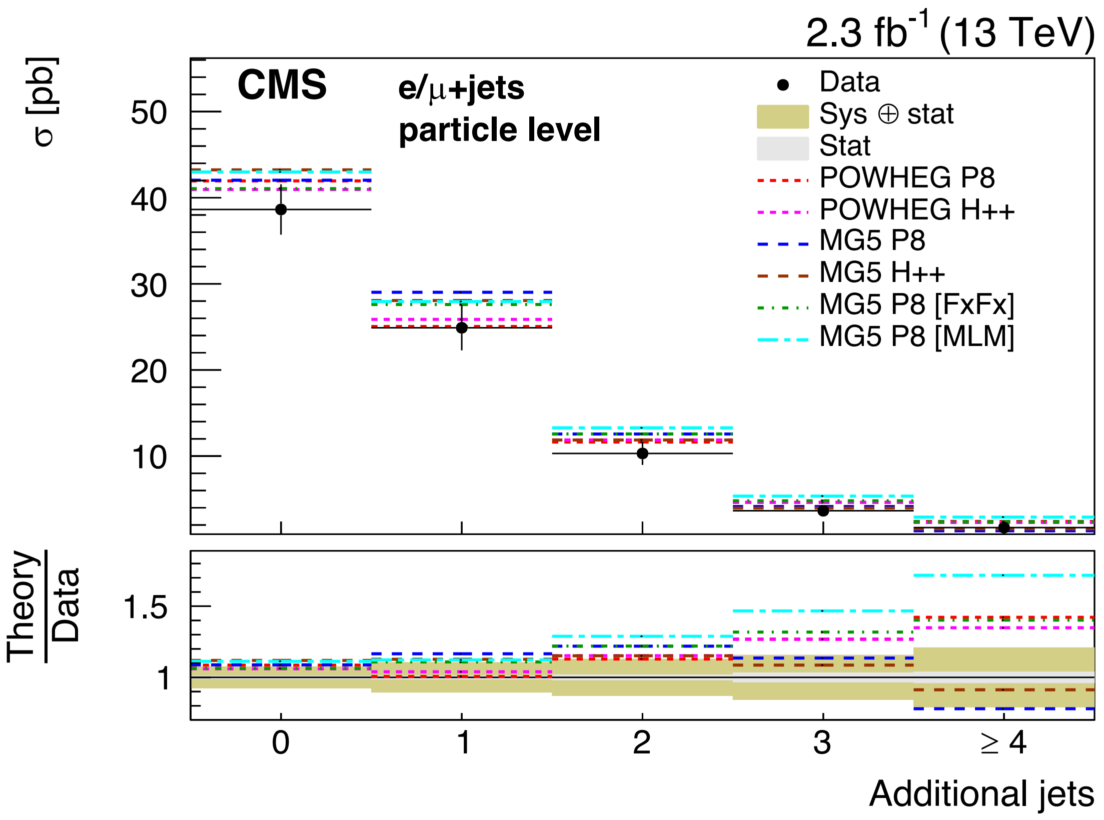


- Herwig7 does not really describe the jet multiplicity
- MG5_aMC@NLO requires more work

$t\bar{t}$ + additional jets at 13 TeV (CMS)

PRD 95, 092001 (2017)





- Data has softer $t\bar{t}$ system p_T spectrum
- Most of MC predictions overshoot data

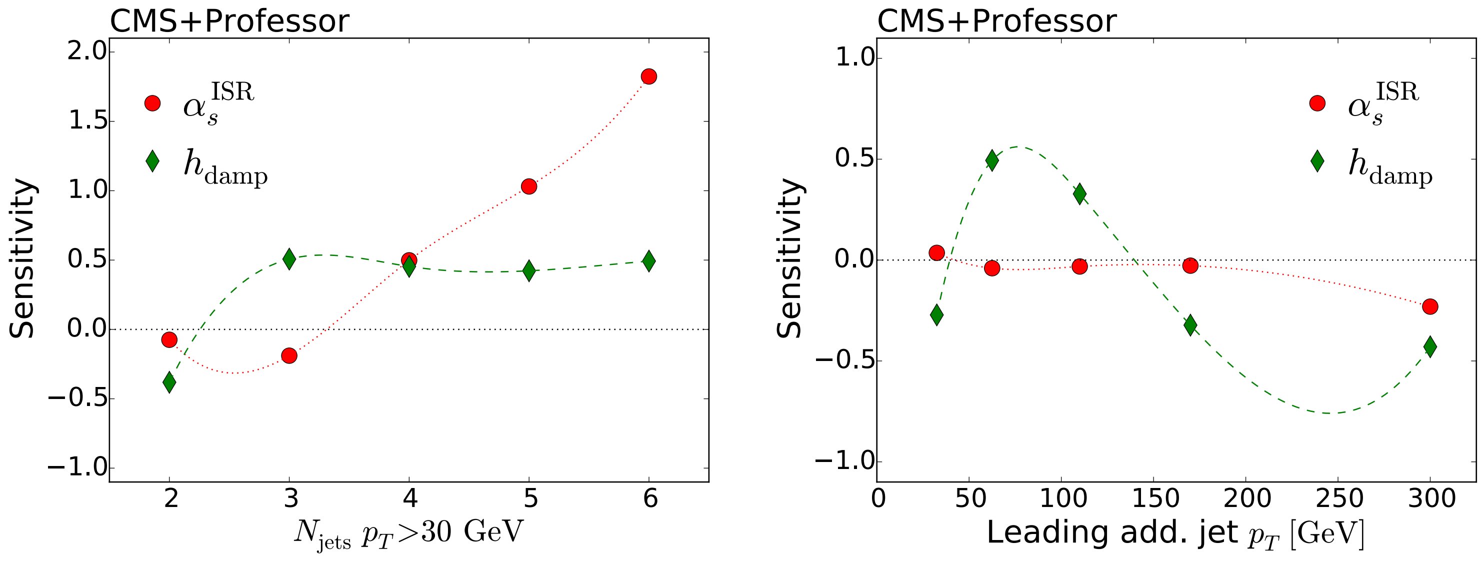
Tuning with α_S^{ISR} and h_{damp}

CMS-PAS-TOP-16-021

POWHEG : h_{damp} = controls of the p_T of the first additional emission beyond the Born configuration (default is top quark mass 172.5 GeV)
→ regulate the high-p_T emission against top quark pair system recoils

damping real emission generated by POWHEG with a factor $\frac{h_{damp}^2}{(p_T^2+h_{damp}^2)}$

PYTHIA 8: α_S^{ISR} is the value of the strong coupling at m_Z (default is 0.1365)



(a) (b)

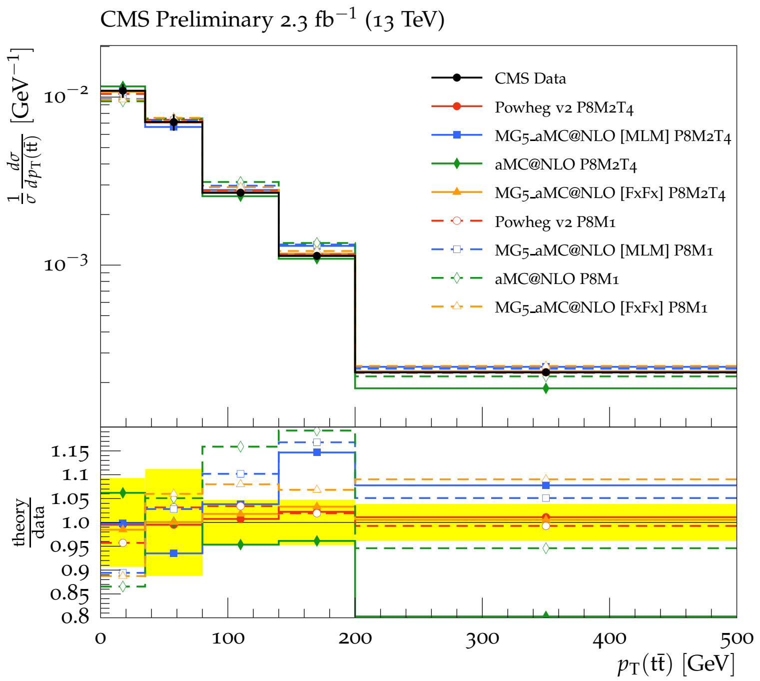
Tune using the 8 TeV data

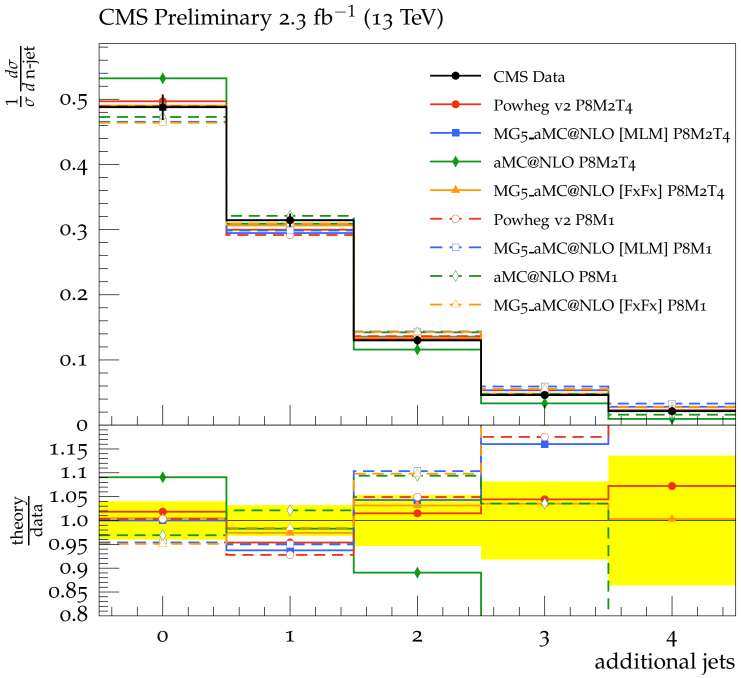
$$h_{damp} = 1.581^{+0.658}_{-0.585} \times m_t, \qquad \alpha_s^{ISR} = 0.1108^{+0.0145}_{-0.0142}$$

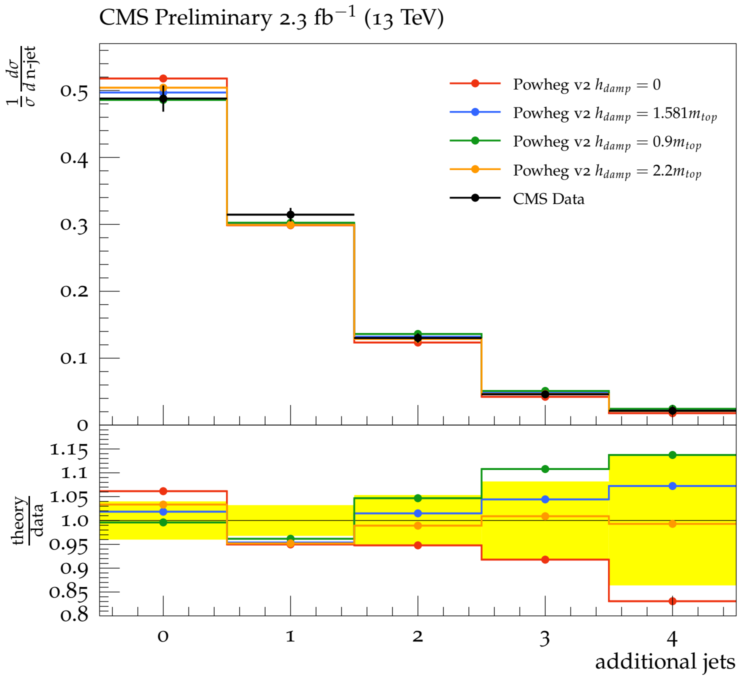
$t\bar{t}$ p_T spectrum and jet multiplicity (CMS)

CUETP8M2T4 is new tune
CUETP8M1 is old tune

CMS-PAS-TOP-16-021



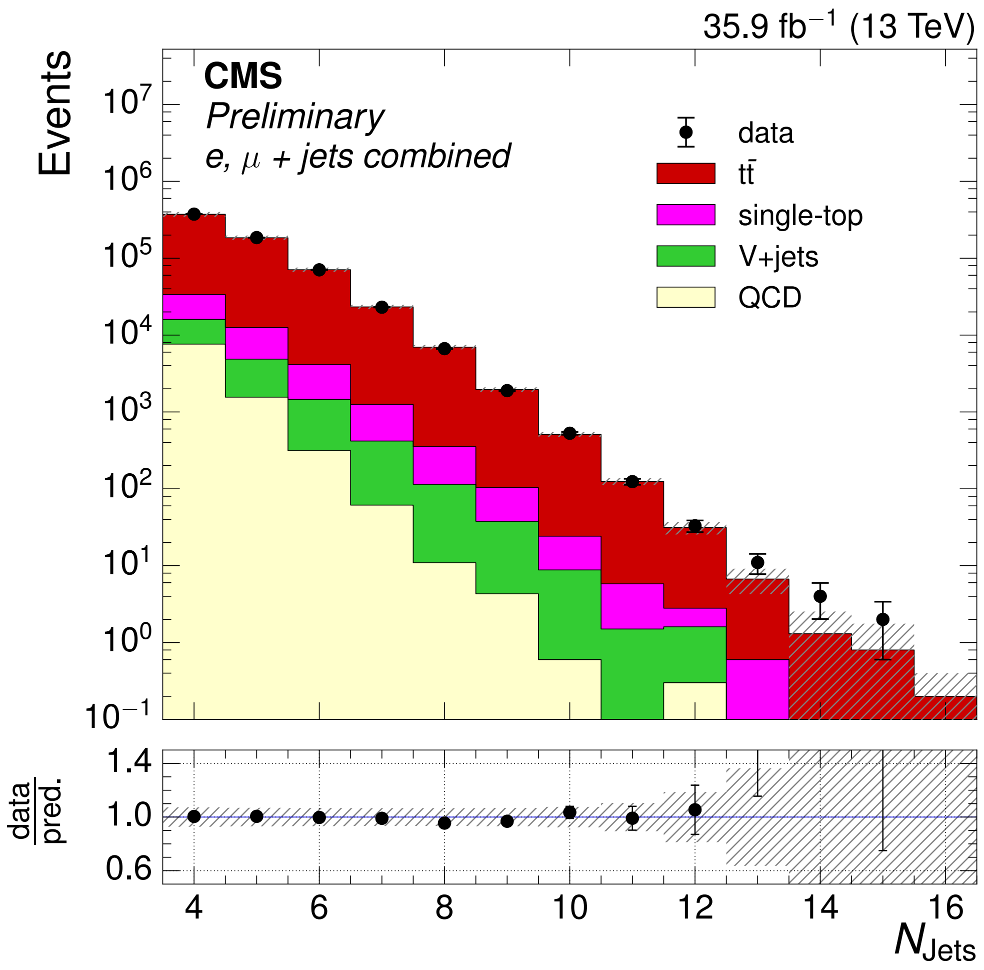




- NLO generators agree with data within uncertainty
- LO order of MG5_aMC@NLO (MLM configuration) and aMC@NLO do not agree with data
- Data disfavors vanishing h_{damp}

Jet multiplicity in lepton + jets with 36 fb^{-1}

CMS-PAS-TOP-16-014



- POWHEG+Pythia8 prediction of the jet multiplicity is consistent with data
- The jet multiplicity from previous 8 TeV measurements was used for CUETP8M2T4 tune
- The tune accurately described the jet multiplicity on a larger dataset with a higher $\sqrt{s}$

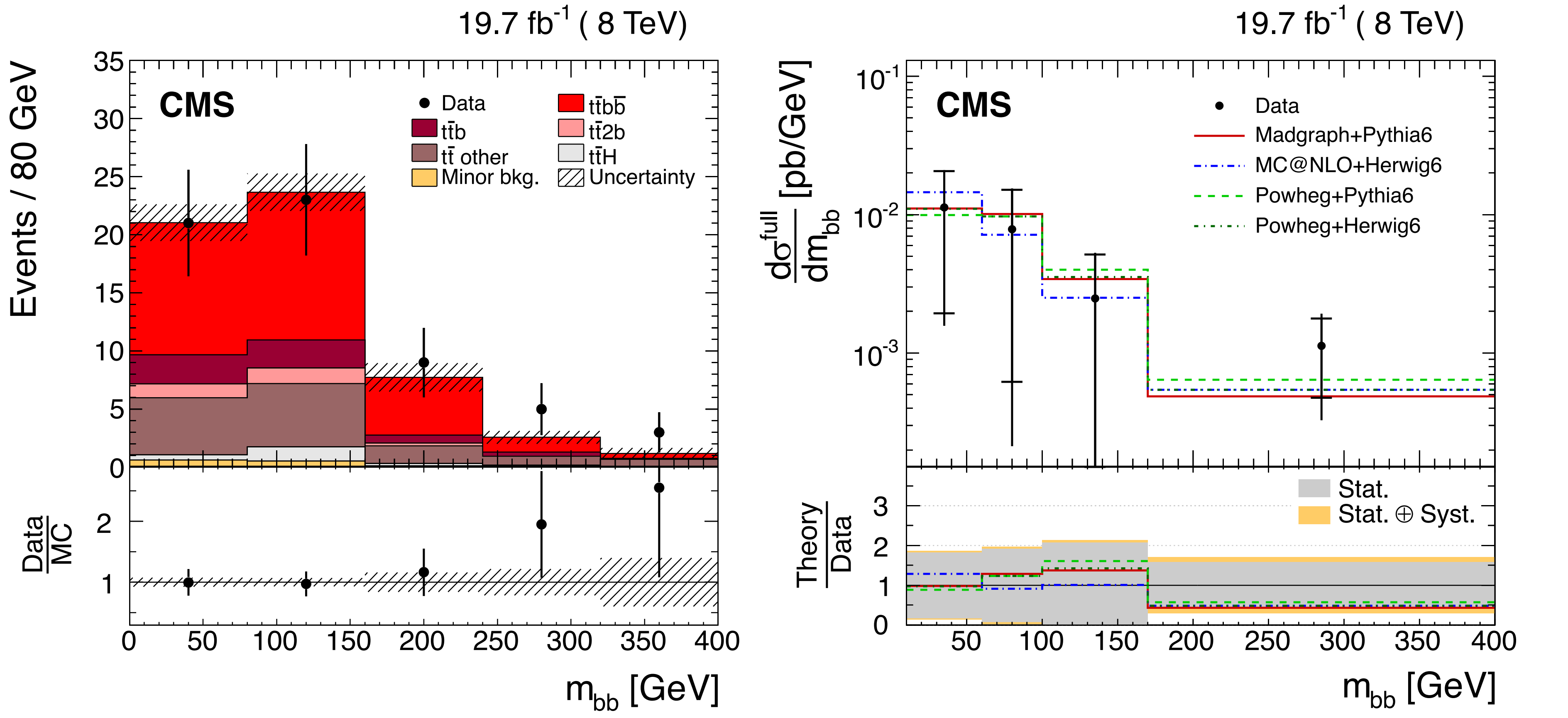
Conclusion

- We are towards very precision measurement in top quark measurement, in particular, using differential distributions.
- $t\bar{t}$+X is the main background for many new physics searches. It is crucial to control experimental and theoretical systematic uncertainties.
- More interaction between theory and experiments is required.
- A bunch of ATLAS and CMS differential measurements are available and need to be compared in details.
- Differential $t\bar{t}b\bar{b}$ cross section measurement with more data of 36 fb^{-1} is in progress.

Differential $t\bar{t}b\bar{b}$ cross sections

8 TeV

Eur. Phys. J. C 76 (2016) 379



$t\bar{t}$ + Heavy flavor (ATLAS)

Eur. Phys. J. C76 (2016) 11

	$ttbb$ [fb]	ttb Lepton-plus-jets [fb]	ttb $e\mu$ [fb]	R_{ttbb} (%)
Observed	(cut-based) 18.2 ±3.5 ±5.7 (fit-based) 12.4 ±3.3 ±3.6	$930 \pm 70\ ^{+240}_{-190}$	$48 \pm 10\ ^{+15}_{-10}$	1.20 ±0.33 ±0.28
MADGRAPH5_aMC@NLO (μ_{BDDP})	$19.1^{+8.4}_{-6.1}$	870^{+300}_{-270}	51^{+21}_{-16}	–
MADGRAPH5_aMC@NLO ($\mu_{H_{\mathrm{T}}/4}$)	$12.3^{+4.2}_{-3.6}$	520^{+170}_{-150}	30^{+10}_{-9}	–
POWHEL	$8.8^{+4.4}_{-2.2}$	430^{+250}_{-140}	28^{+16}_{-9}	–
MADGRAPH5+PYTHIA 6	$13.3^{+3.8}_{-3.3}$	790^{+270}_{-170}	43^{+13}_{-8}	$1.29^{+0.15}_{-0.13}$
PYTHIA 8 (wgtq=3)	30.1	1600	88	2.50
PYTHIA 8 (wgtq=5)	12.8	740	42	1.10
PYTHIA 8 (wgtq=6,sgtq=0.25)	16.1	930	53	1.37
POWHEG+PYTHIA 6 (HDAMP=m_{top})	11.2	690	37	1.16

$t\bar{t}b\bar{b}(jj)$ cross sections (CMS)

arXiv:1705.10141

Full phase space vs Visible phase space

Phase Space (PS)	Parton level	Particle level
Visible PS	–	4 (b) jets and 2 leptons (e, μ)
Full PS	t, $\bar{t}$ and 2 (b) jets (not from t or $\bar{t}$)	–

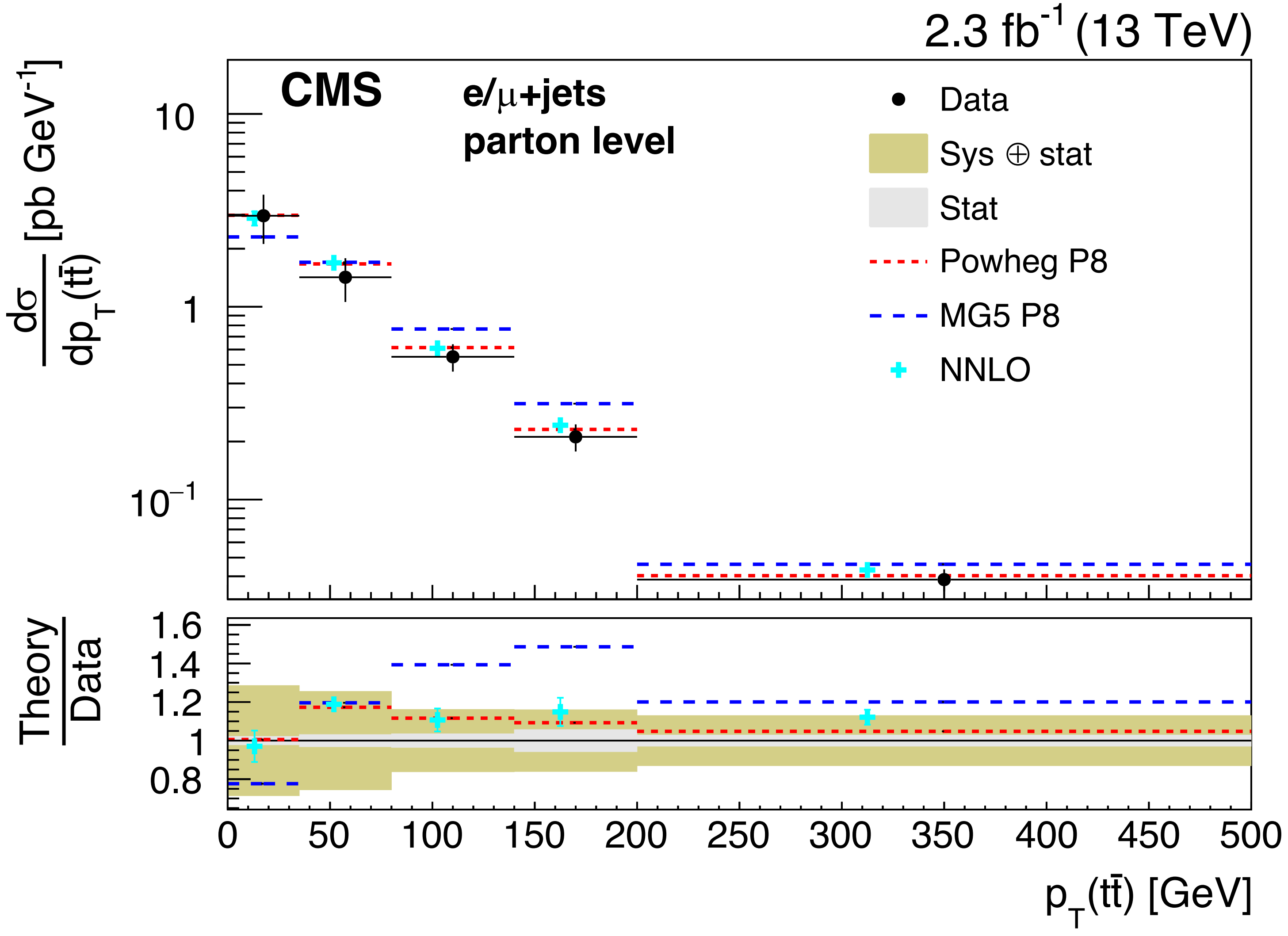
leptons : p_T > 20 GeV, $|\eta|$ < 2.4 / Jets : p_T > 20 GeV, $|\eta|$ < 2.5

Results

Phase space		$\sigma_{t\bar{t}b\bar{b}}$ [pb]	$\sigma_{t\bar{t}jj}$ [pb]	$\sigma_{t\bar{t}b\bar{b}}/\sigma_{t\bar{t}jj}$
Visible	Measurement	$0.088 \pm 0.012 \pm 0.029$	$3.7 \pm 0.1 \pm 0.7$	$0.024 \pm 0.003 \pm 0.007$
	SM (POWHEG)	0.070 ± 0.009	5.1 ± 0.5	0.014 ± 0.001
Full	Measurement	$4.0 \pm 0.6 \pm 1.3$	$184 \pm 6 \pm 33$	$0.022 \pm 0.003 \pm 0.006$
	SM (POWHEG)	3.2 ± 0.4	257 ± 26	0.012 ± 0.001

lepton + jets at 13 TeV (CMS)

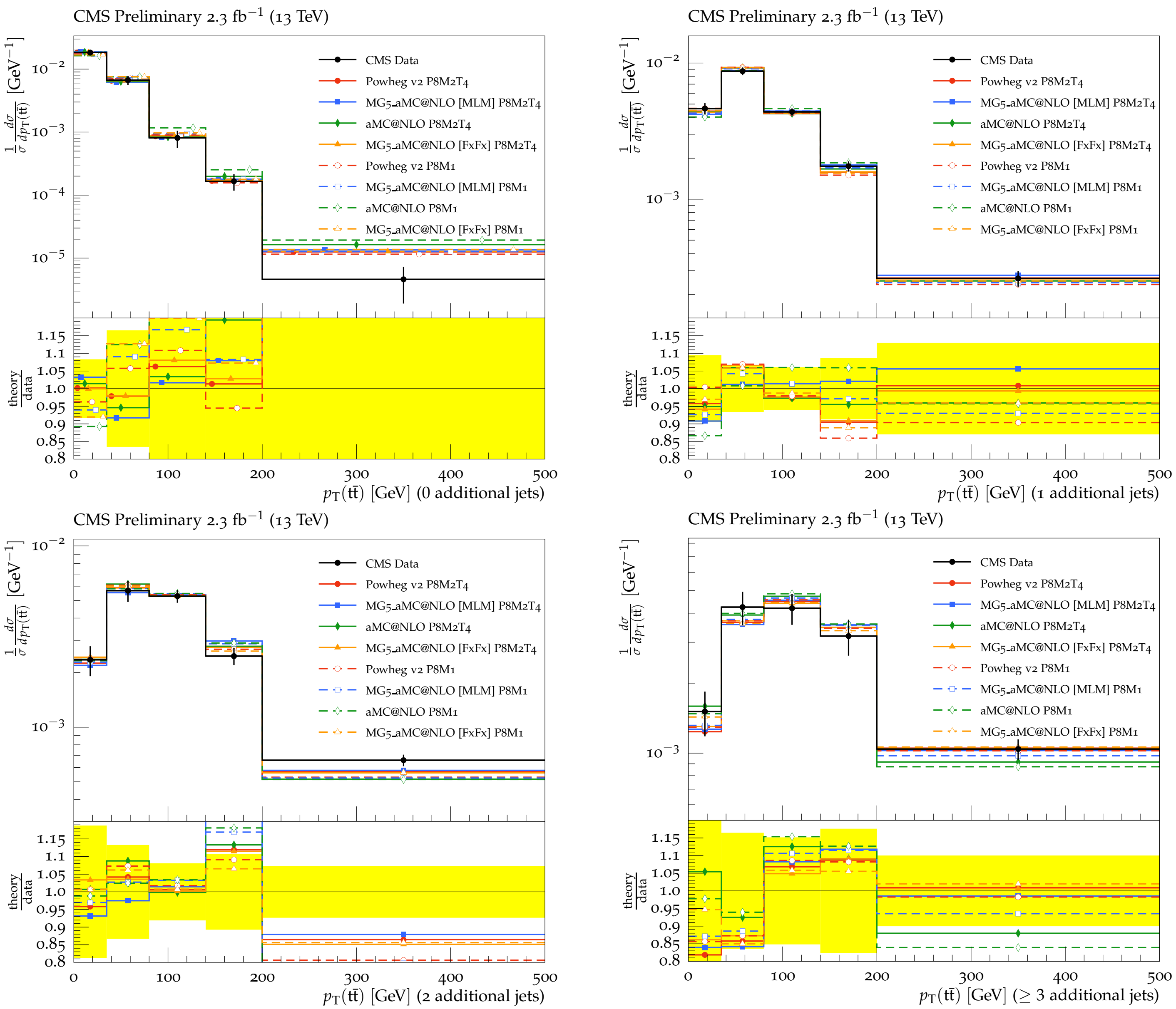
PRD 95, 092001 (2017)



NNLO has a better agreement with data

$t\bar{t}\ p_T$ spectrum in 2D (CMS)

CMS-PAS-TOP-16-021



The 0-additional jet case has the worst agreement data and theory predictions.