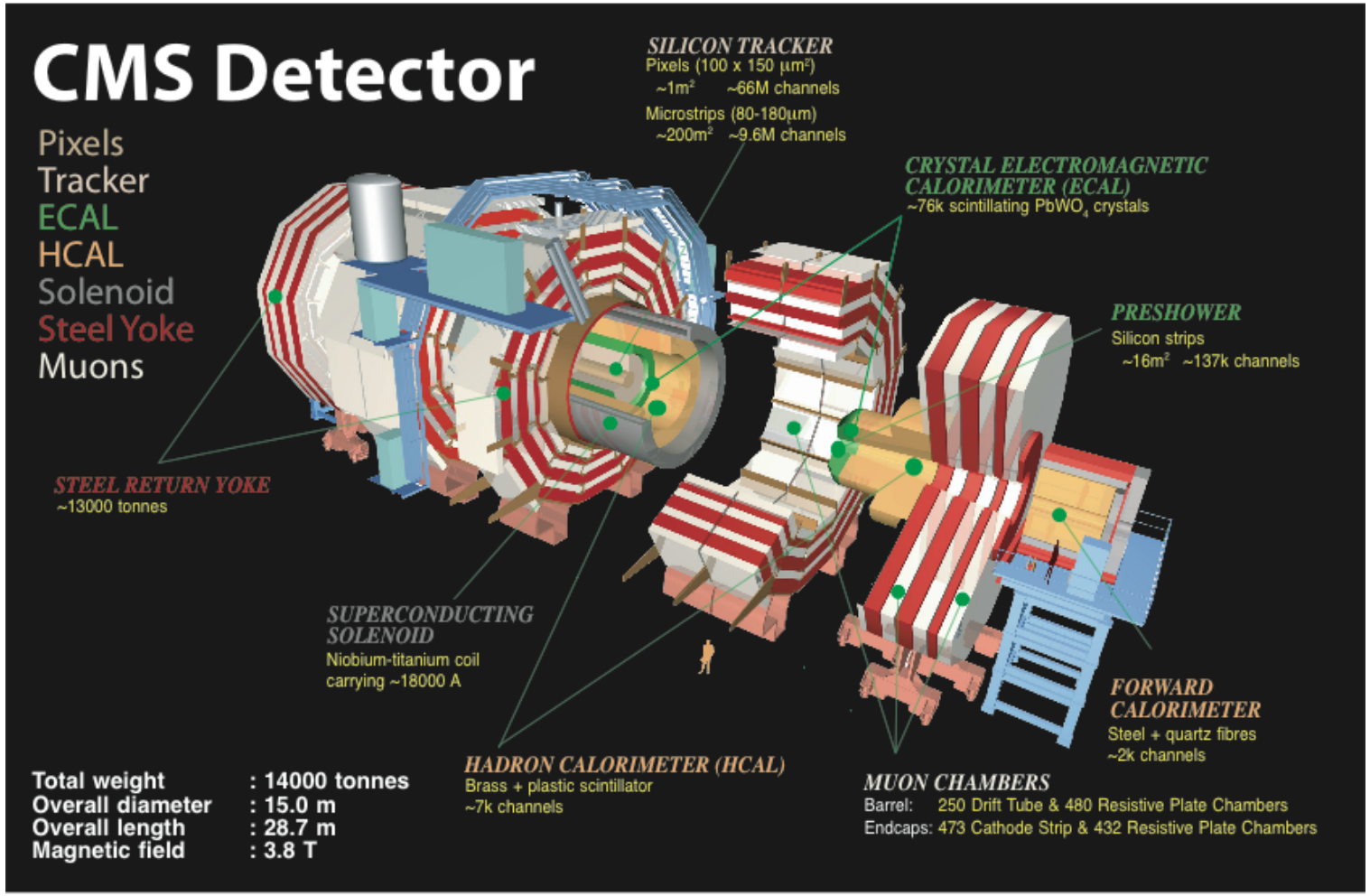
Top quark pair production with additional jets (heavy flavor) from CMS

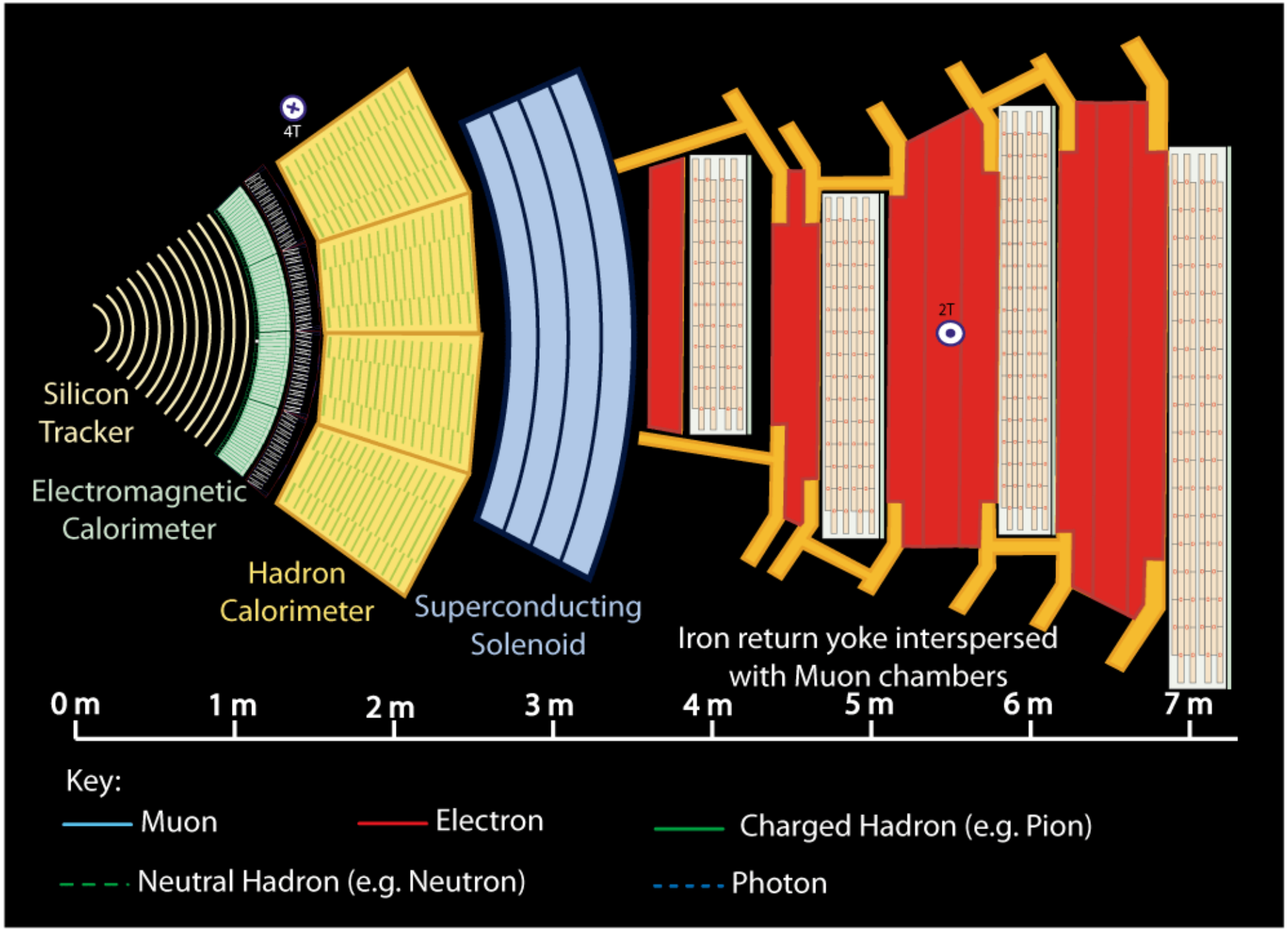
Tae Jeong Kim (Hanyang University)

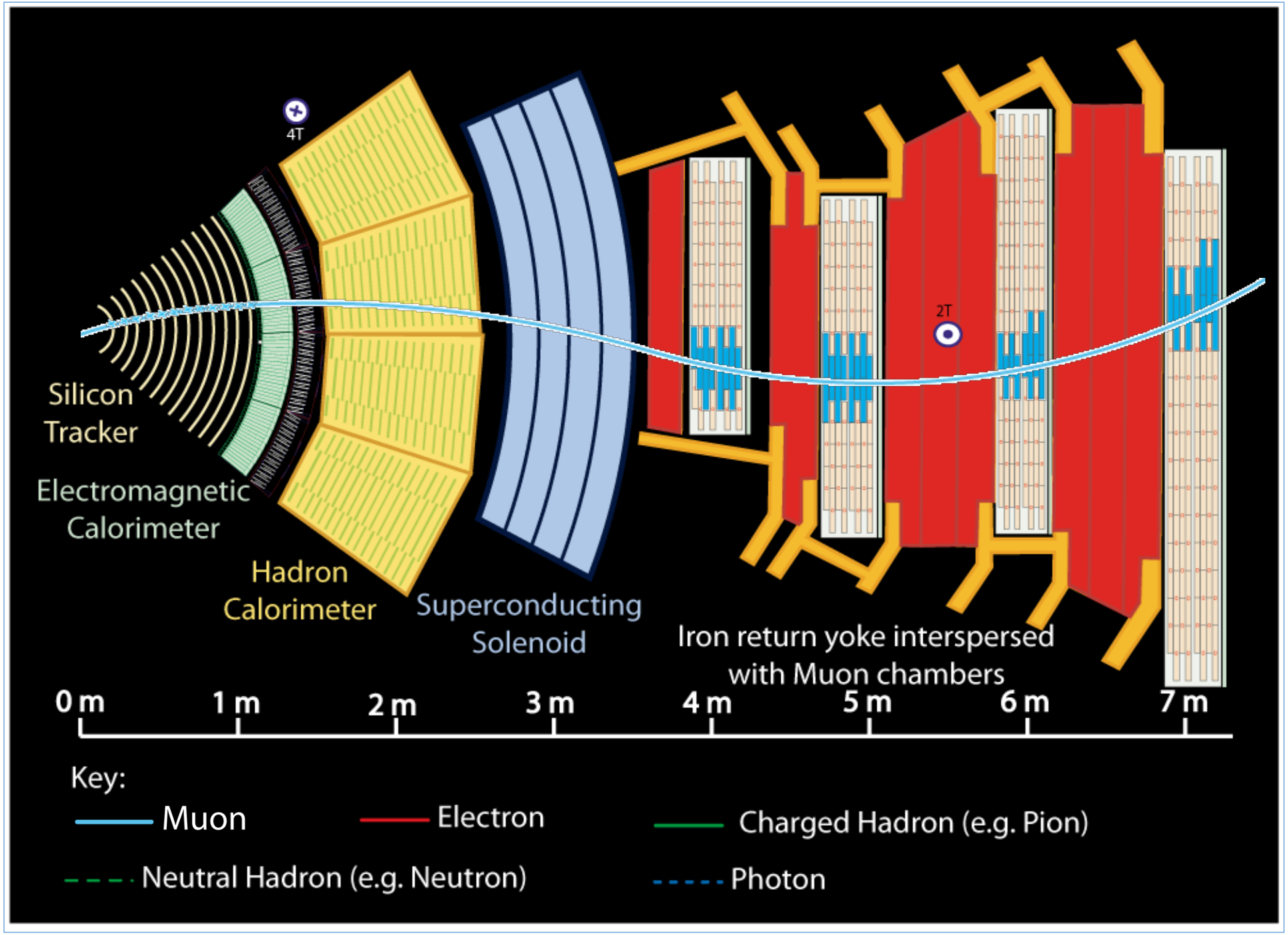
For the first MadAnalysis 5 workshop on LHC recasting in Korea 22/August/2017, at High 1

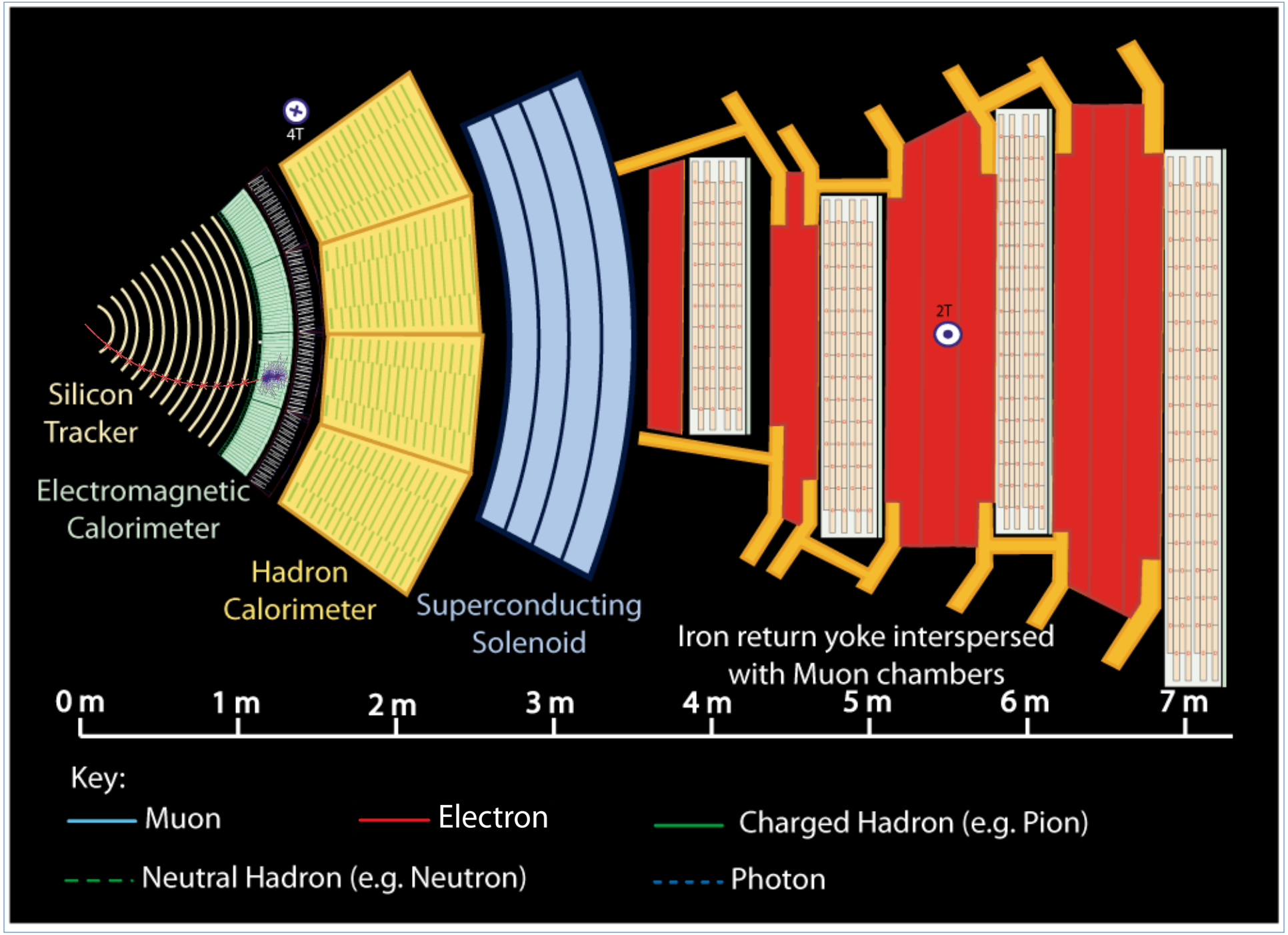
CMS detector

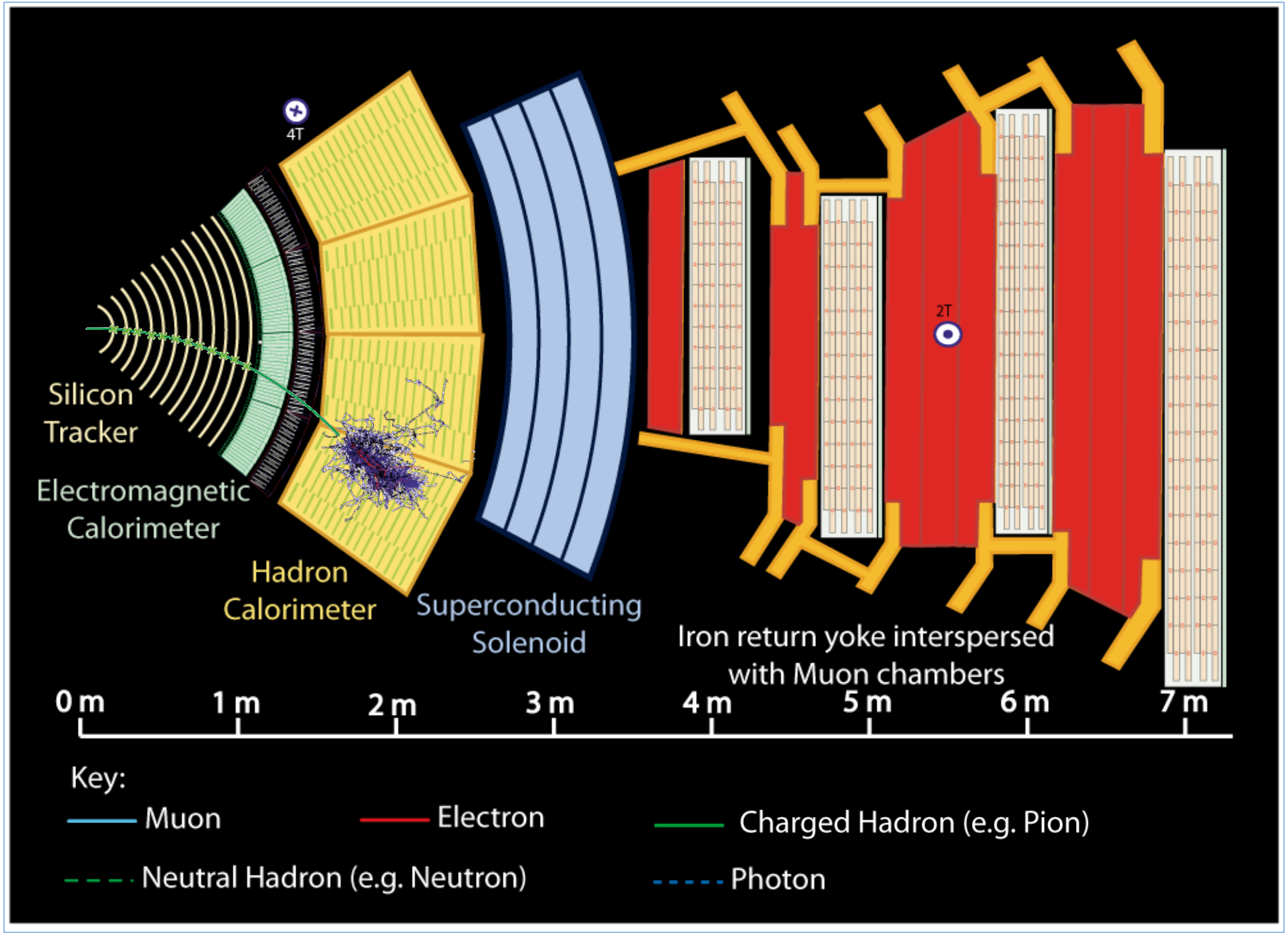
- Mass 14000 tonnes
- Height = 15 m
- Length = 28.7 m
- measure the direction and energy of the particles
- Select 100 events per second among billion events

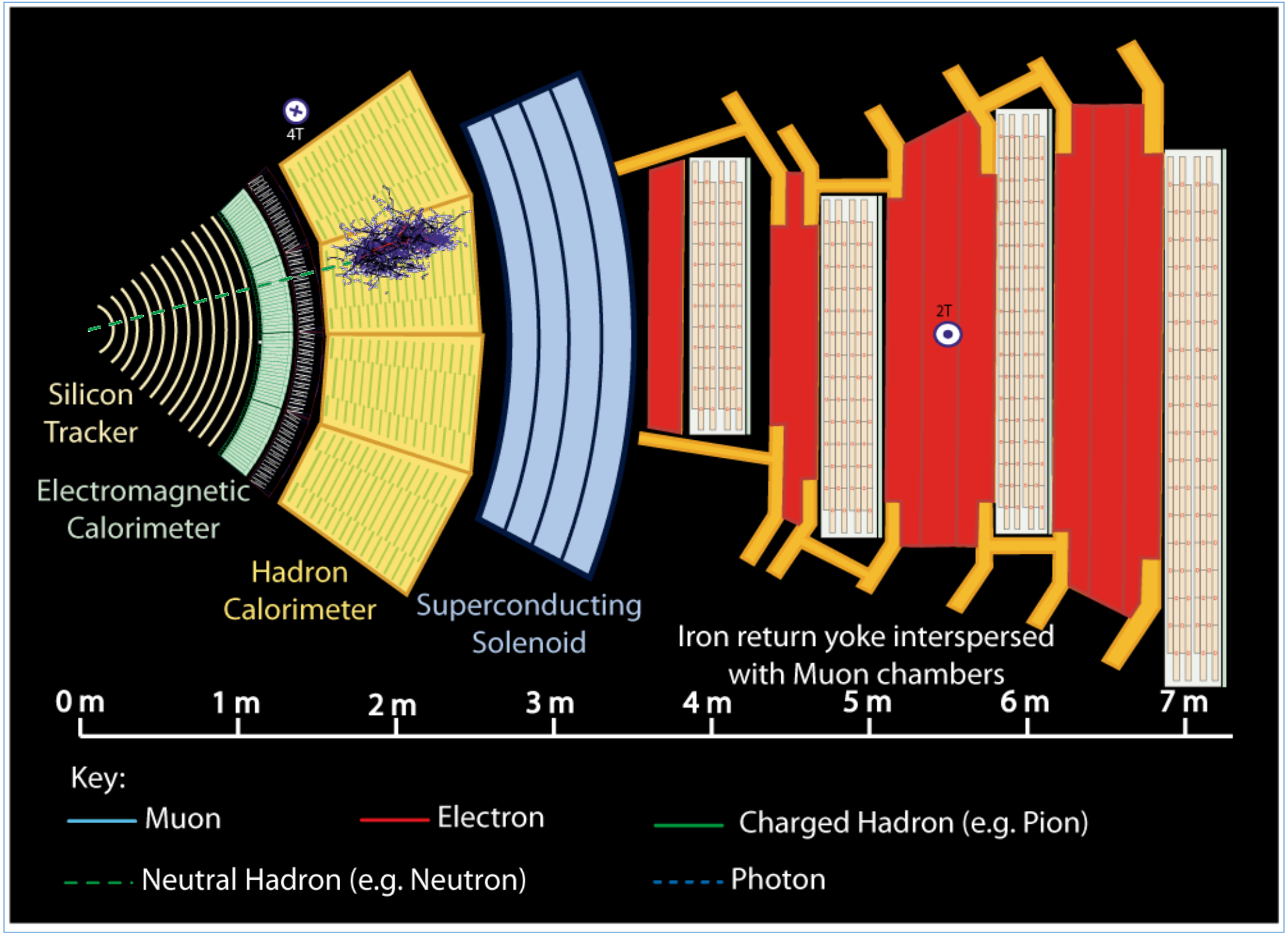


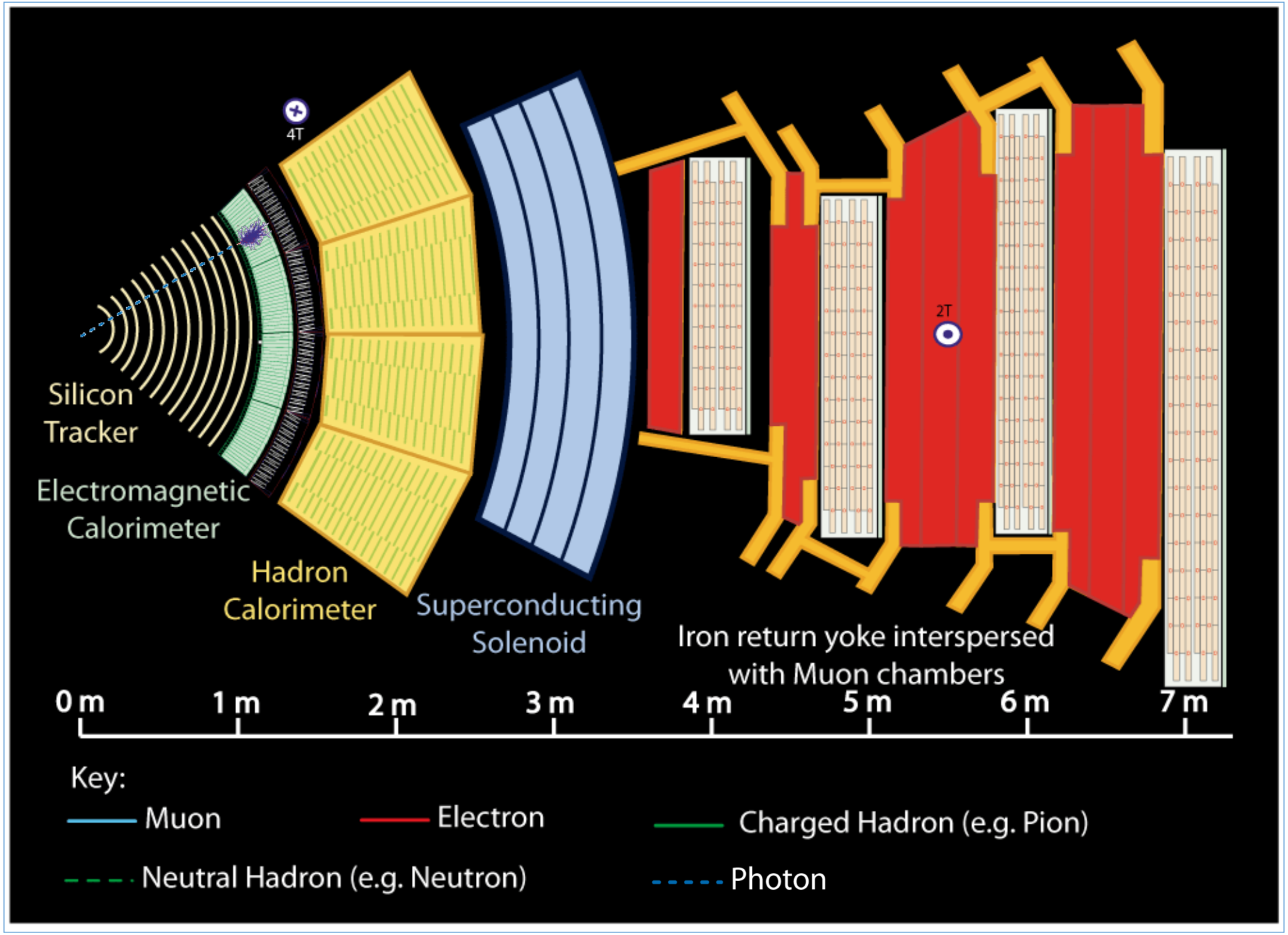

4T
2T
Silicon Tracker
Electromagnetic Calorimeter
Hadron Calorimeter
Superconducting Solenoid
Iron return yoke interspersed with Muon chambers
0 m
1 m
2 m
3 m
4 m
5 m
6 m
7 m
Key:
Muon
Electron
Charged Hadron (e.g. Pion)
Neutral Hadron (e.g. Neutron)
Photon


4T
2T
Silicon Tracker
Electromagnetic Calorimeter
Hadron Calorimeter
Superconducting Solenoid
Iron return yoke interspersed with Muon chambers
0 m
1 m
2 m
3 m
4 m
5 m
6 m
7 m
Key:
Muon
Electron
Charged Hadron (e.g. Pion)
Neutral Hadron (e.g. Neutron)
Photon


4T
2T
Silicon Tracker
Electromagnetic Calorimeter
Hadron Calorimeter
Superconducting Solenoid
Iron return yoke interspersed with Muon chambers
0 m
1 m
2 m
3 m
4 m
5 m
6 m
7 m
Key:
Muon
Electron
Charged Hadron (e.g. Pion)
Neutral Hadron (e.g. Neutron)
Photon

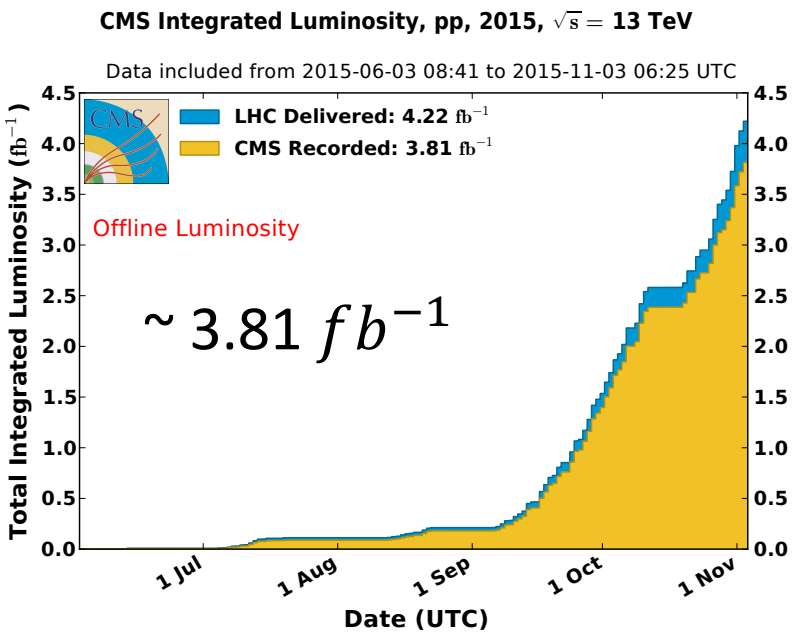

4T
2T
Silicon Tracker
Electromagnetic Calorimeter
Hadron Calorimeter
Superconducting Solenoid
Iron return yoke interspersed with Muon chambers
0 m
1 m
2 m
3 m
4 m
5 m
6 m
7 m
Key:
Muon
Electron
Charged Hadron (e.g. Pion)
Neutral Hadron (e.g. Neutron)
Photon


4T
2T
Silicon Tracker
Electromagnetic Calorimeter
Hadron Calorimeter
Superconducting Solenoid
Iron return yoke interspersed with Muon chambers
0 m
1 m
2 m
3 m
4 m
5 m
6 m
7 m
Key:
Muon
Electron
Charged Hadron (e.g. Pion)
Neutral Hadron (e.g. Neutron)
Photon

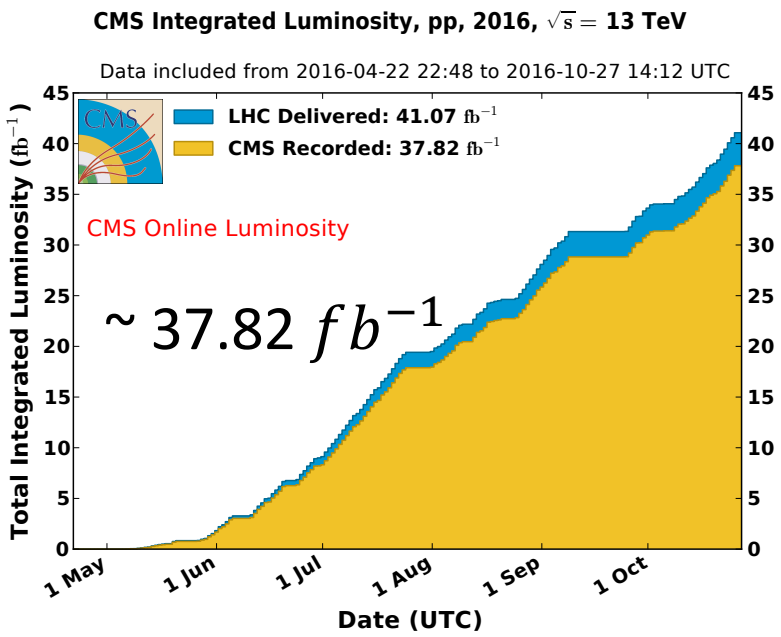

4T
2T
Silicon Tracker
Electromagnetic Calorimeter
Hadron Calorimeter
Superconducting Solenoid
Iron return yoke interspersed with Muon chambers
0 m
1 m
2 m
3 m
4 m
5 m
6 m
7 m
Key:
Muon
Electron
Charged Hadron (e.g. Pion)
Neutral Hadron (e.g. Neutron)
Photon

Integrated luminosity

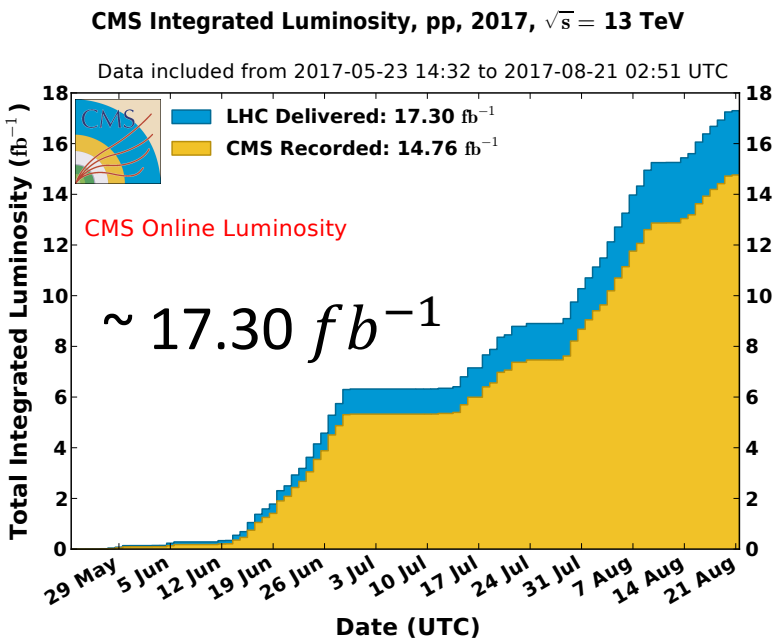
Integrated luminosity = $\int_0^t L_{int} dt$ collected at 13 TeV



2015



2016



2017 as of today

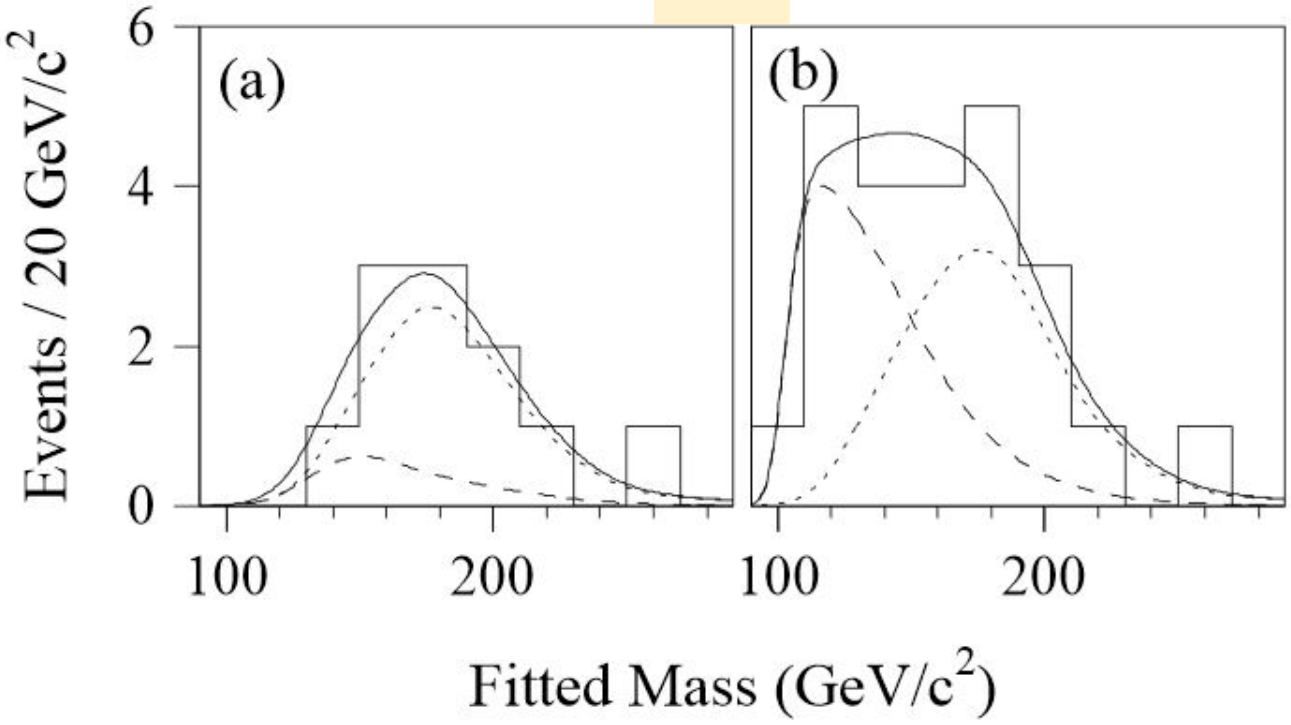
Introduction to top quark

- Predicted as isospin partner of bottom quark (bottom quark was discovered in 1977)
- The top quark is the most massive of the quarks ~ 173 GeV in the standard model : m(t) > m(H) > m(Z) > m(W)
- Because of its mass, it was not observed directly at the LEP
- Discovered in 1995 by CDF and D0 experiment
- Lifetime is very short: $\tau = \frac{1}{\Gamma} = 4\times10^{-25}s, \Gamma = 1.5$ GeV
- It decays before it can form strongly interacting bound states
- The top quark properties can be studied without non-perturbative effects

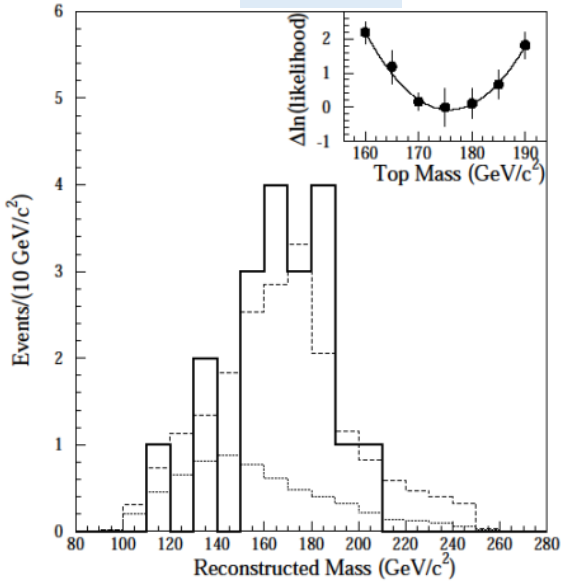
Top quark discovery

- Discovered in 1995 by CDF and D0 with only 17 events (D0) and 19 events (CDF)
- Produced 1000s events at Tevatron

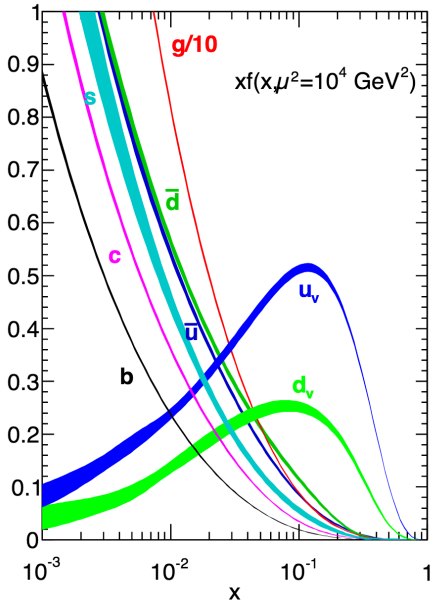
D0



CDF

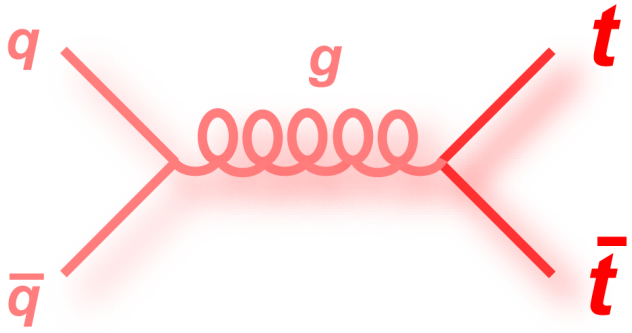


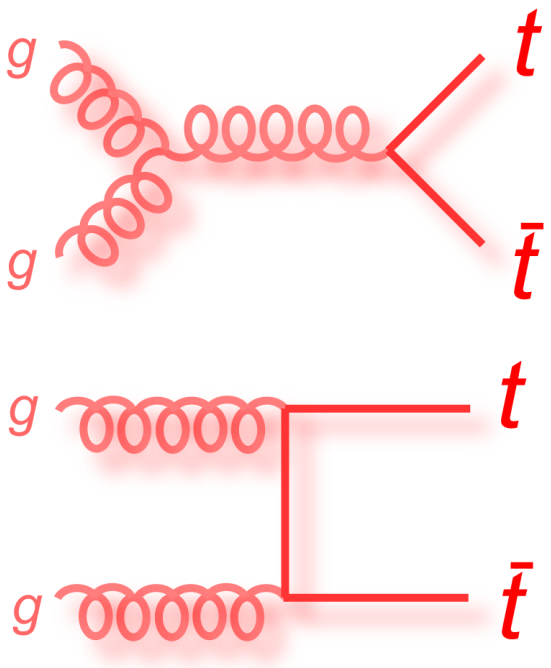
top quark pair production at the LHC



13 TeV

Quark annihilation ~10% Gluon-gluon fusion ~90%





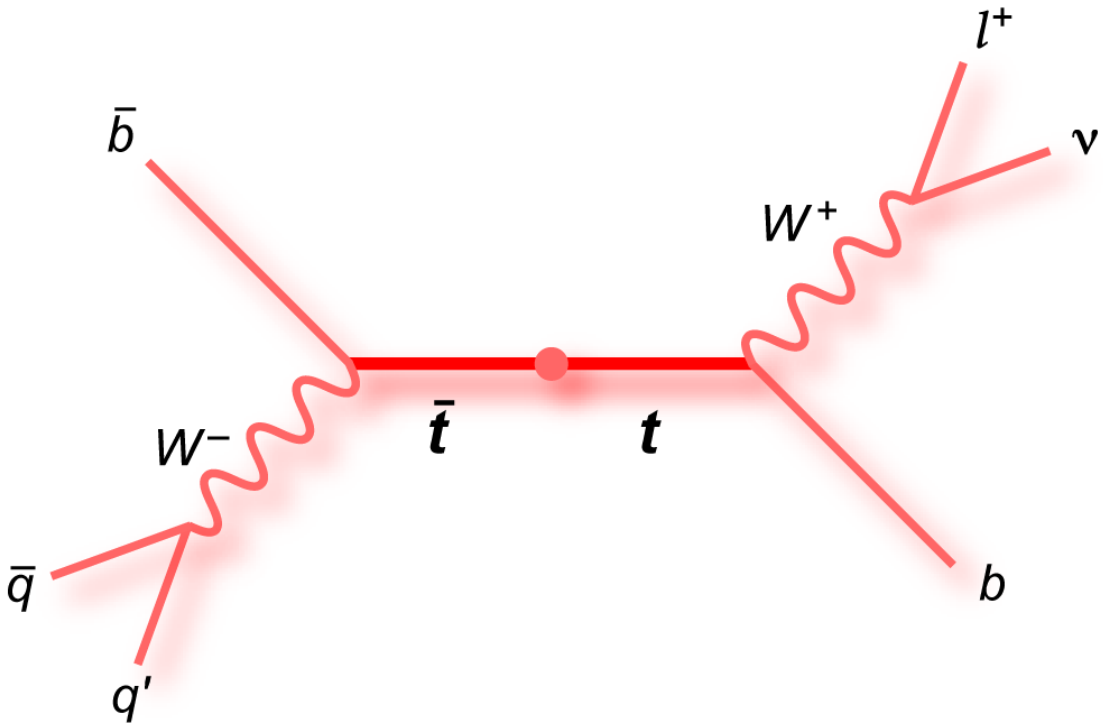
Top++, Comput.Phys.Commun. 185 (2014) 2930

Top mass	Central (pb)	Scale uncert.	PDF+α_S uncert.	Mass uncert.
172.5 GeV	831.76	+19.77-29.20	+35.06-35.06	+23.18-22.45
173.2 GeV	815.96	+19.37-28.61	+34.38-34.38	+22.67-21.95

- With Run 2 data in 2016 only corresponding to around 36 fb^{-1}, LHC already produced ~million events after the final selection.

Top quark decay

- ~100% top decays to bW



- High p_T isolated leptons
- missing transverse energy (ν)
- Jets, b-jets

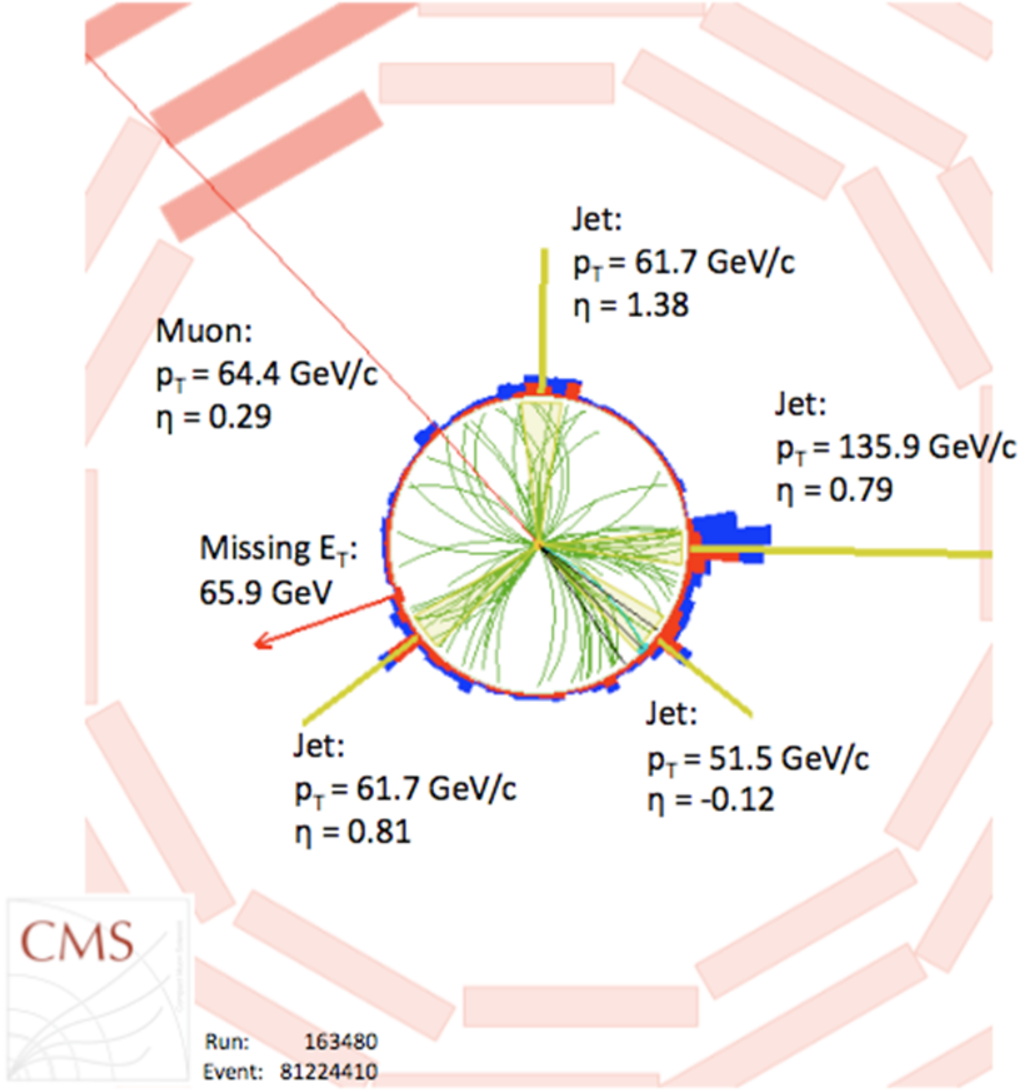
Top Pair Decay Channels

W decay	e^+	μ^+	τ^+	$u\bar{d}$	$c\bar{s}$
e^-	ee	$e\mu$	$e\tau$	electron+jets	
μ^-	$e\mu$	$\mu\mu$	$\mu\tau$	muon+jets	
τ^-	$e\tau$	$\mu\tau$	$\tau\tau$	tau+jets	
$\bar{u}d$	electron+jets	muon+jets	tau+jets	all-hadronic	
$\bar{c}s$					

dileptons

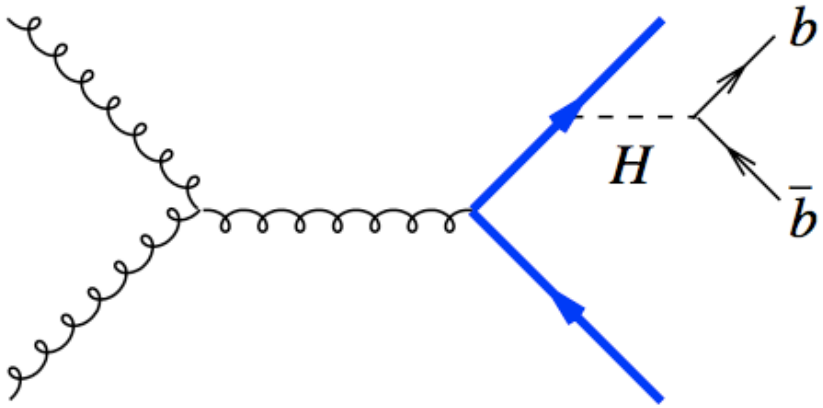
Dilepton (e/μ) ~ 6.7%
Lepton+jets ~ 35%
All hadronic ~ 45%

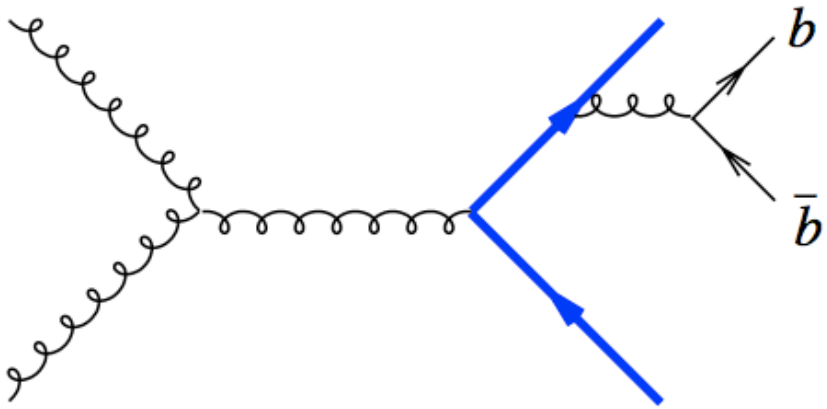
Top quark candidate



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
- $t\bar{t}jj(c\bar{c})$ is the reducible background faking 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.





MC samples from CMS

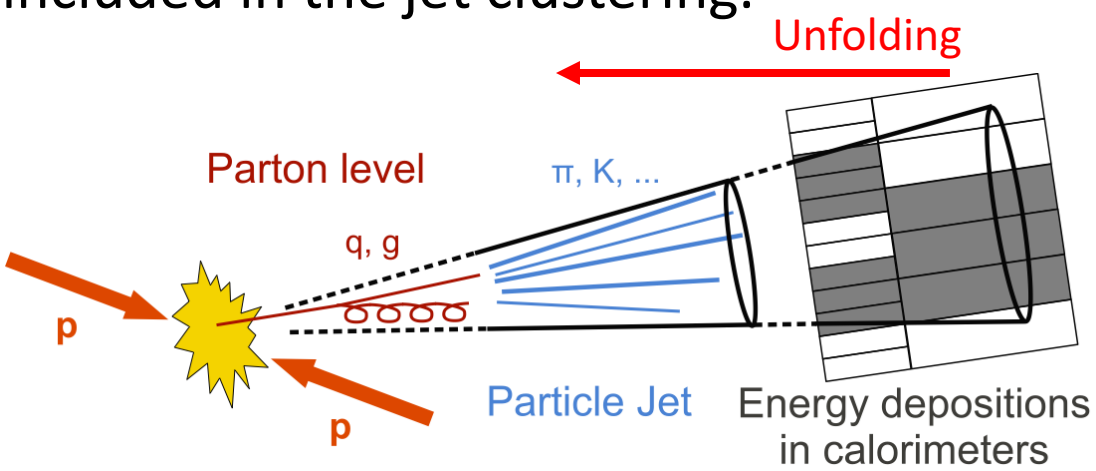
- For Run 1, the LO MadGraph with up to three additional partons interfaced with PYTHIA 6 is reference $t\bar{t}$ sample.
  - It described most of kinematic distributions except top quark p_T spectrum.
- For Run 2,

Event generator	Parton shower	
POWHEG	Pythia 8	Default for Run 2
MadGraph5_aMC@NLO	Pythia 8 / Herwig++	
MadGraph5_aMC@NLO	Pythia 8 with FxFx /MLM merging	
POWHEG	Herwig 7/ Herwig++	

Pileup reweighting done : MC samples are reweighted to match the distribution of the average number of interactions in data

Particle-level objects

- Particle-level objects (based on the stable particle)
  - 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.



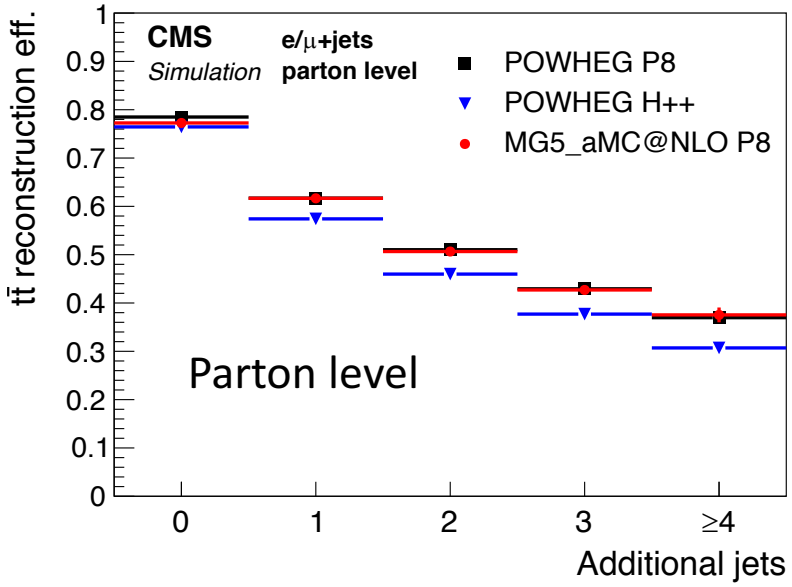
Particle level top quarks

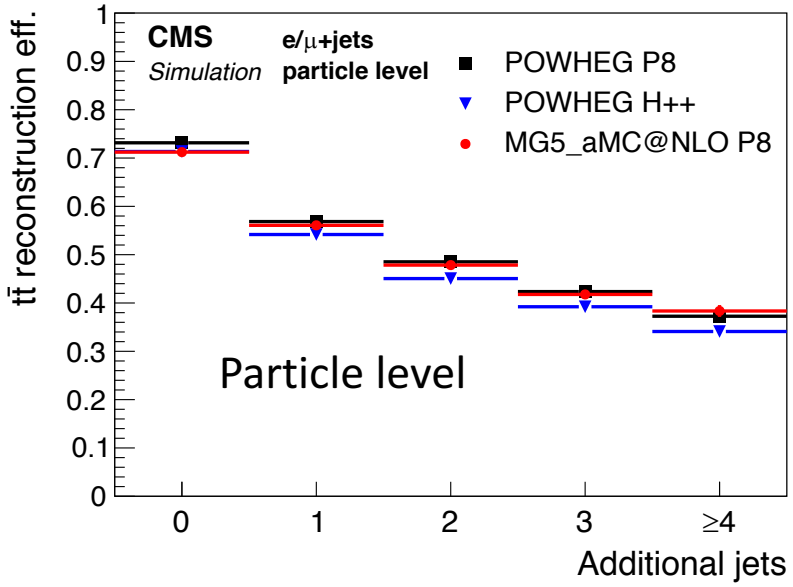
- In the lepton + jets, take the jet permutation by minimizing following quantity

$$
\begin{aligned}
K^2 = {} & [M(p_\nu + p_\ell + p_{b_\ell}) - m_t]^2 \\
& + [M(p_{\mathrm{j}_1} + p_{\mathrm{j}_2}) - m_W]^2 \\
& + [M(p_{\mathrm{j}_1} + p_{\mathrm{j}_2} + p_{b_\mathrm{h}}) - m_t]^2
\end{aligned}
$$

- b-tagged jets are assigned first with higher priority

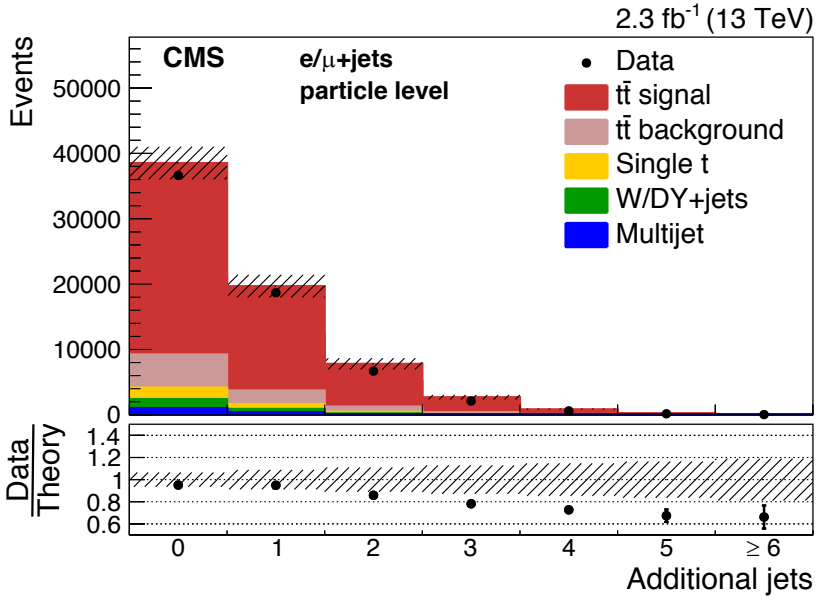
lepton + jets at 13 TeV (CMS)





PRD 95, 092001 (2017)

Number of permutation increases drastically → more likely select a wrong permutation.

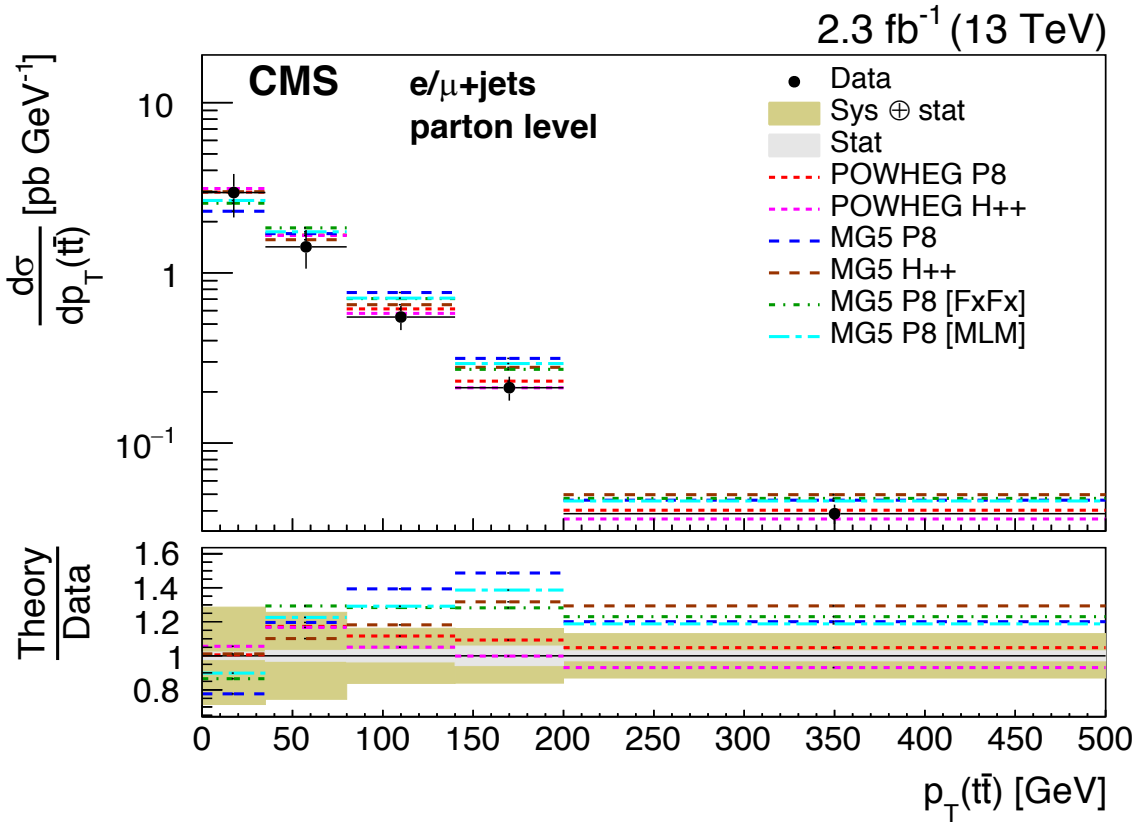


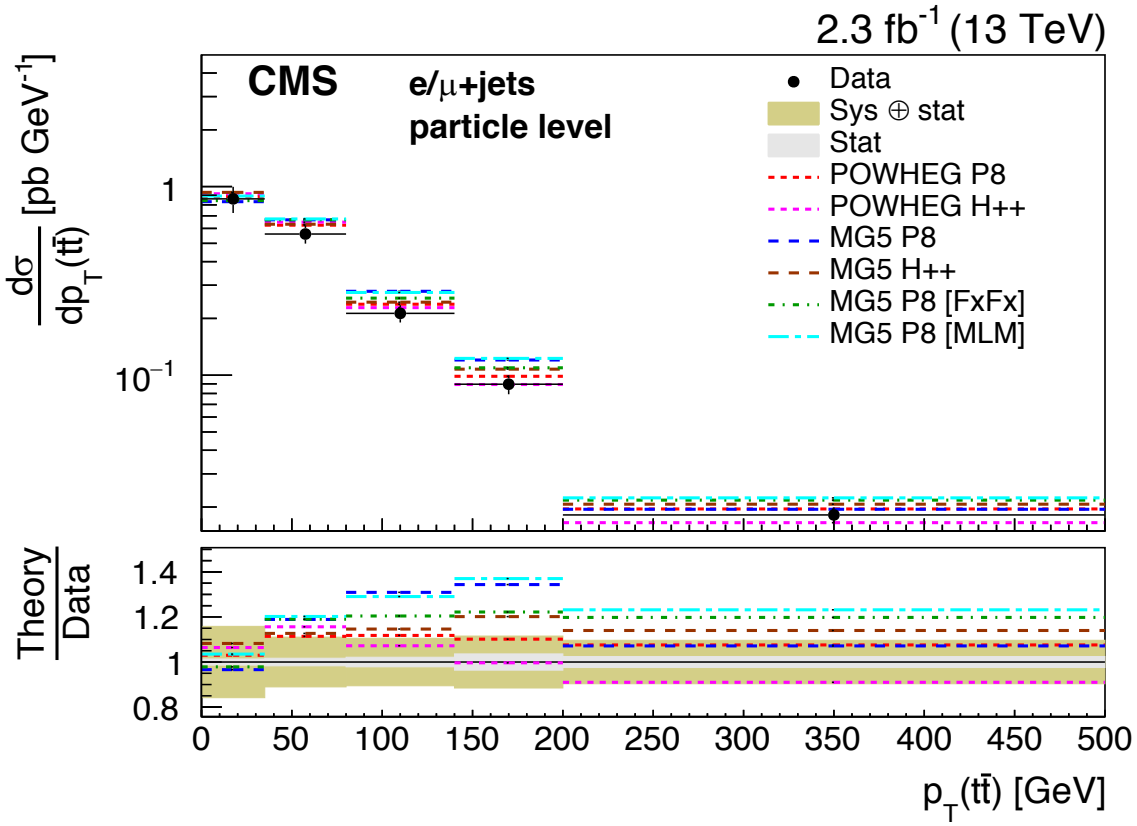
- Only one electron or muon
  $p_T > 30$ GeV, $|\eta| < 2.1$
- No additional lepton
  $p_T > 15$ GeV, $|\eta| < 2.4$
- At least four jets
  $p_T > 30$ GeV, $|\eta| < 2.4$
- At least two b-tagged jets

lepton + jets at 13 TeV (CMS)

Recoiling objects depend on the transverse momentum of $t\bar{t}$

PRD 95, 092001 (2017)

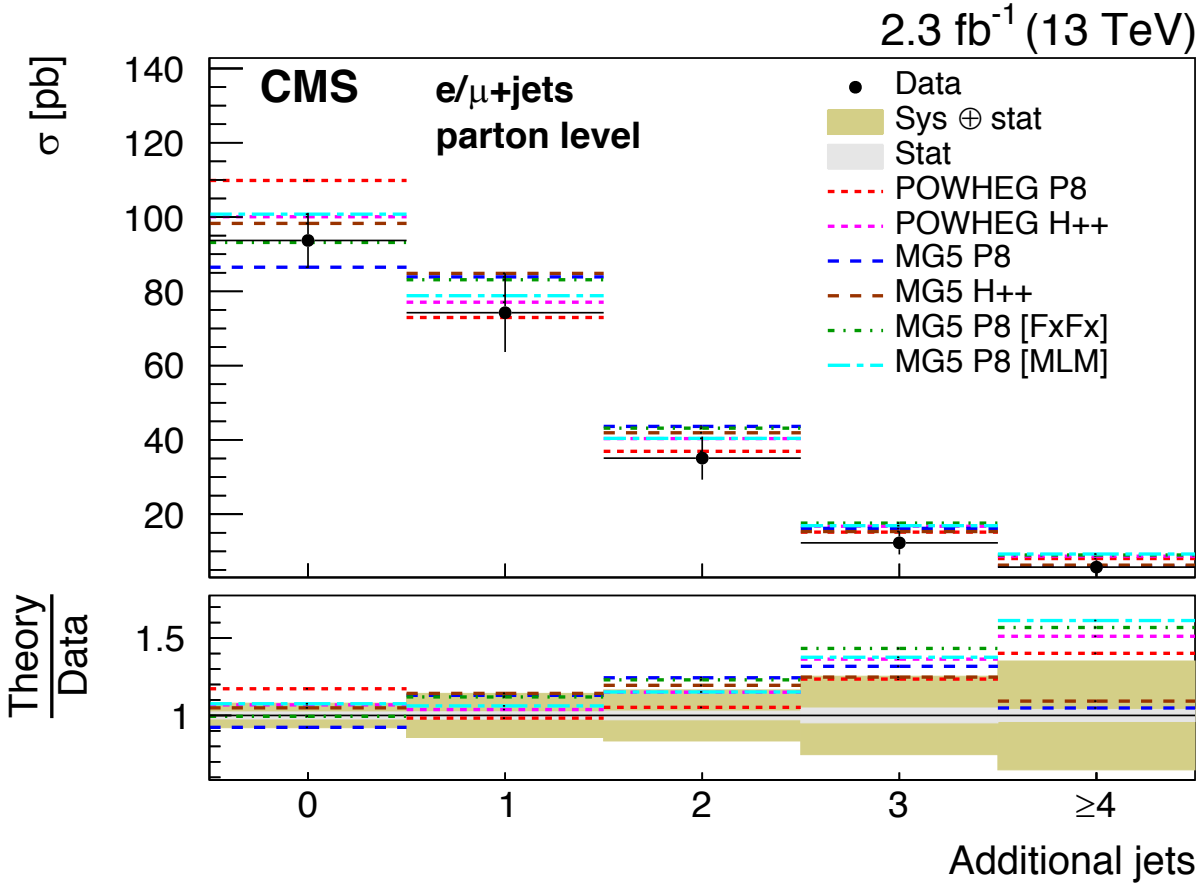


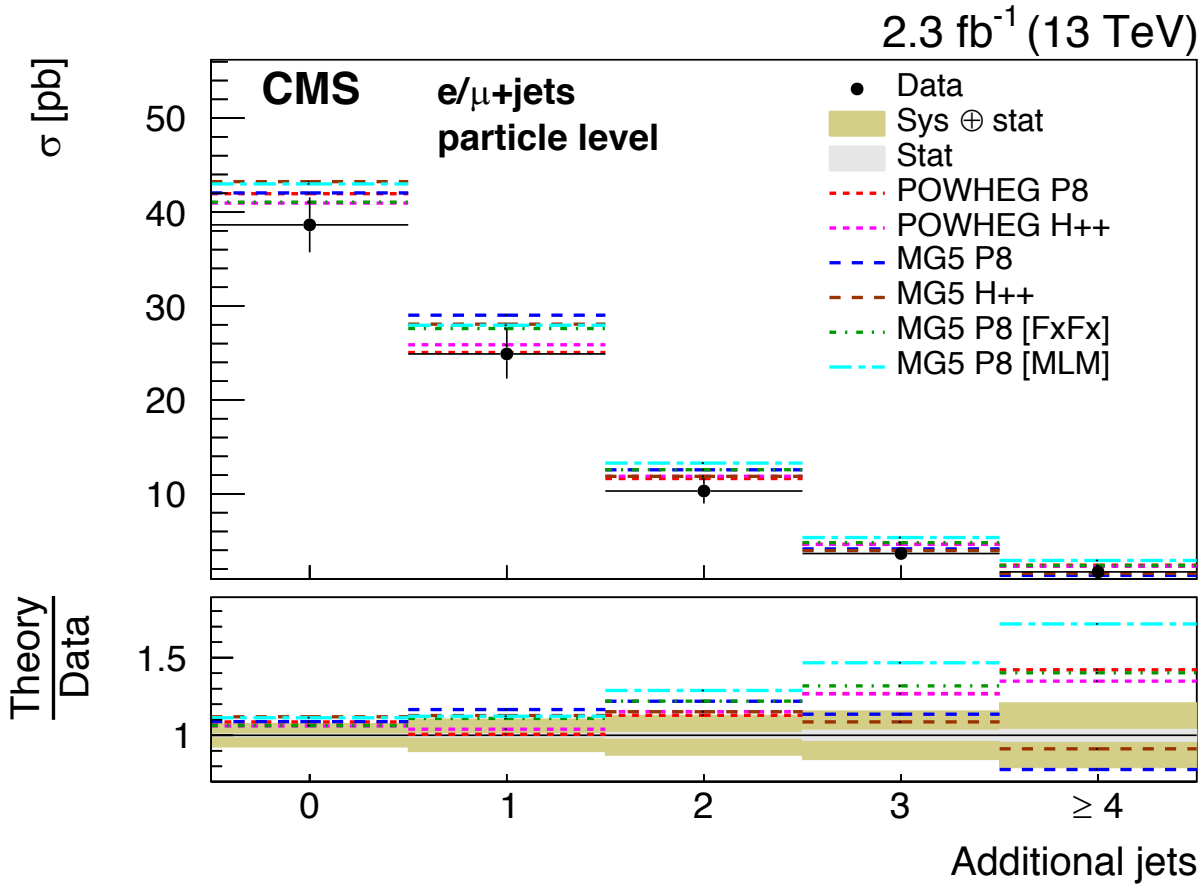


- Data has softer $t\bar{t}$ system p_T spectrum
- Theoretical uncertainties are reduced in the particle-level measurements as these have less dependency on theory-based extrapolation

lepton + jets at 13 TeV (CMS)

PRD 95, 092001 (2017)

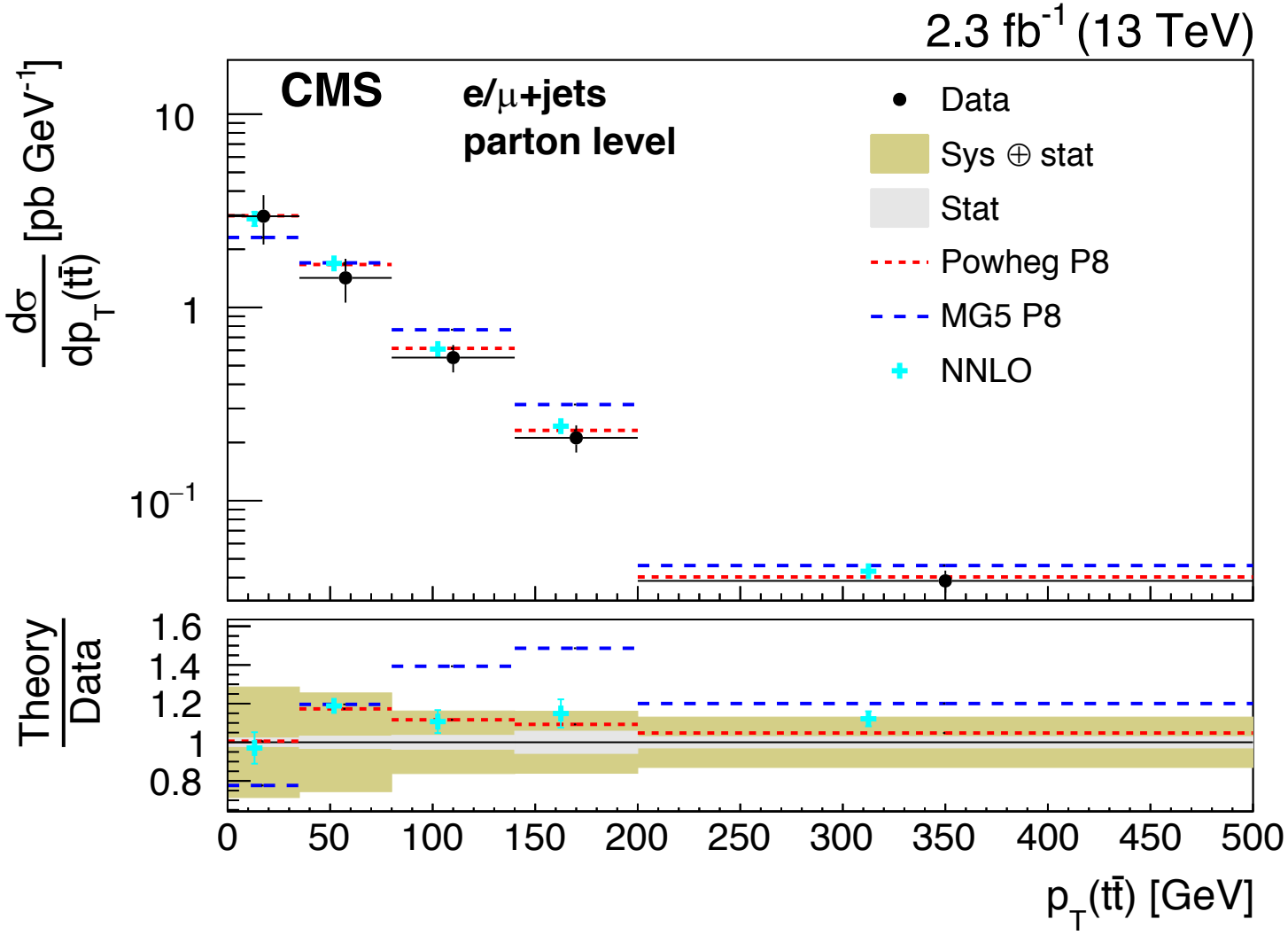




Most of MC predictions overshoot data

lepton + jets at 13 TeV (CMS)

PRD 95, 092001 (2017)



NNLO has a better agreement with data

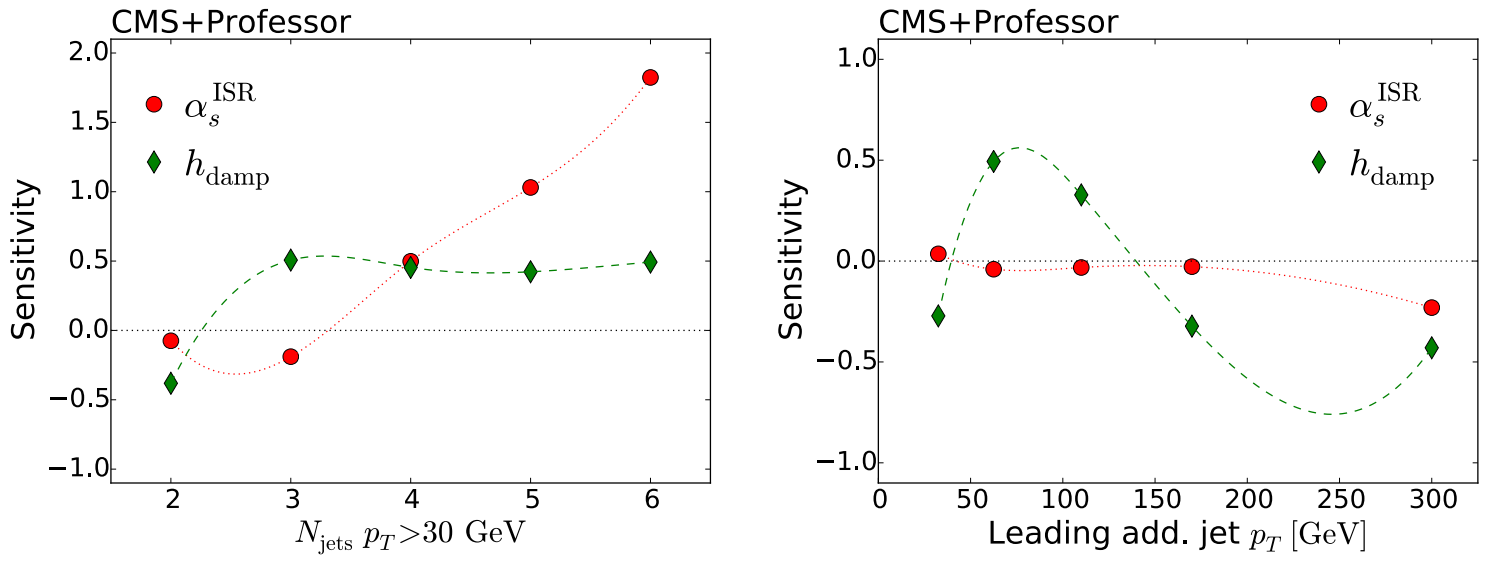
Tuning with α_S^{ISR} and h_{damp}

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)



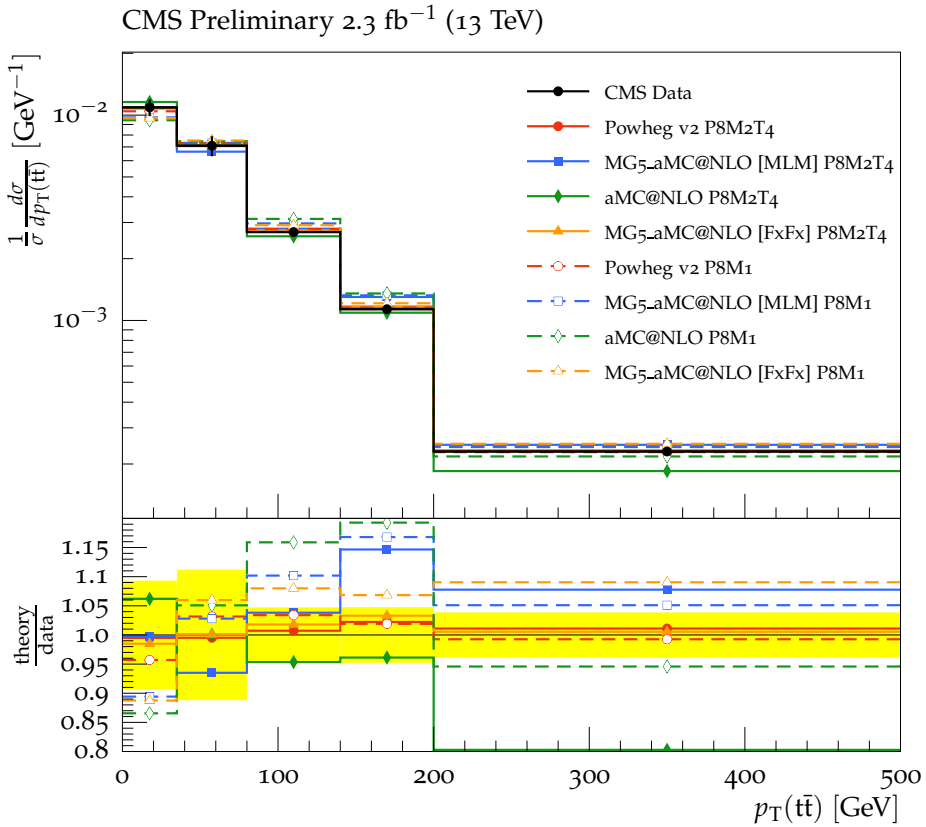
(a) (b)

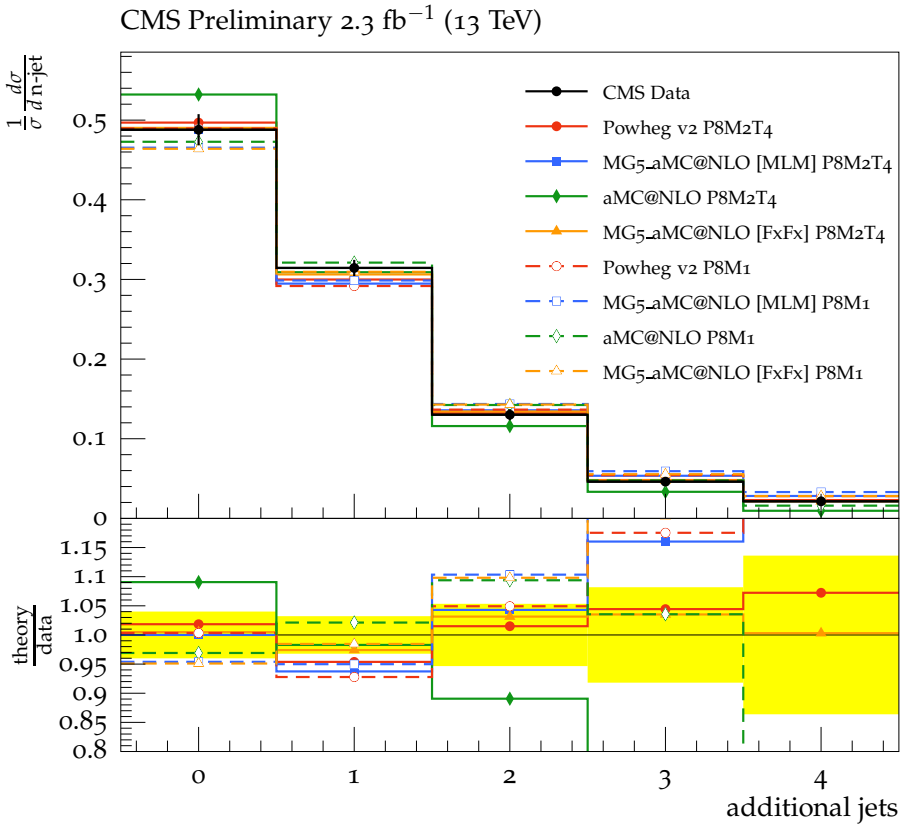
Tune using the 8 TeV data

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

$t\bar{t}$ p_T spectrum and jet multiplicity

CMS-PAS-TOP-16-021





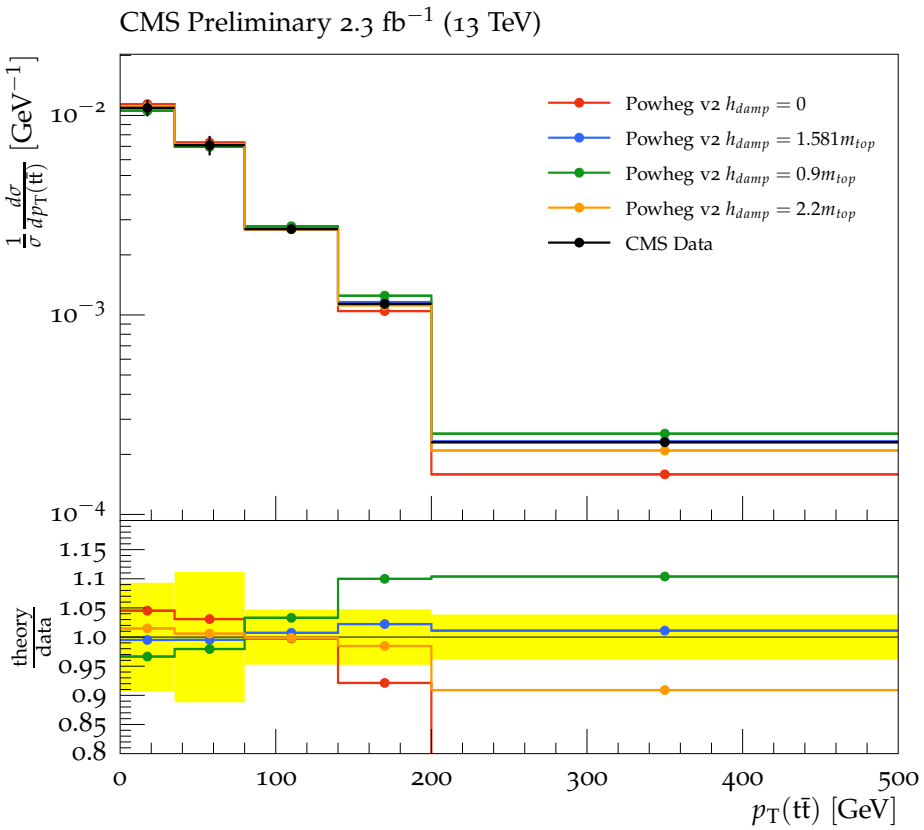
CUETP8M2T4 is new tune
CUETP8M1 is old tune

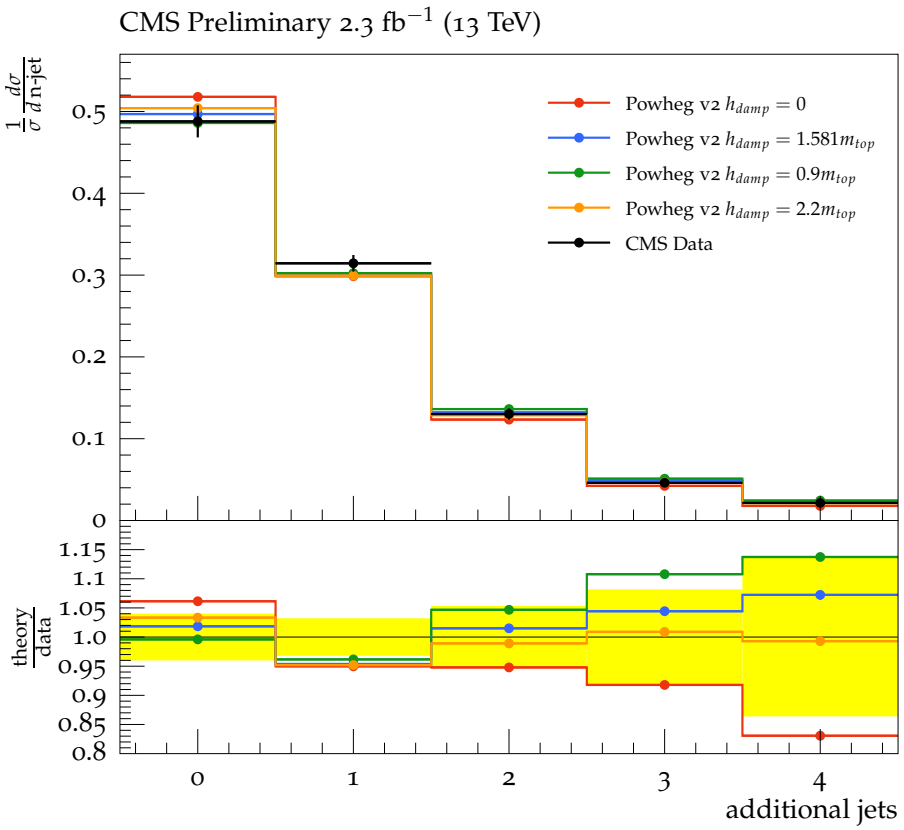
NLO generators agree with data within uncertainty
LO order of MG5_aMC@NLO (MLM configuration) and aMC@NLO do not agree with data

$t\bar{t}$ p_T spectrum and jet multiplicity

- Comparisons with different h_{damp} with h_{damp}=1.581m_{top} being the central value
- Data disfavors vanishing h_{damp}

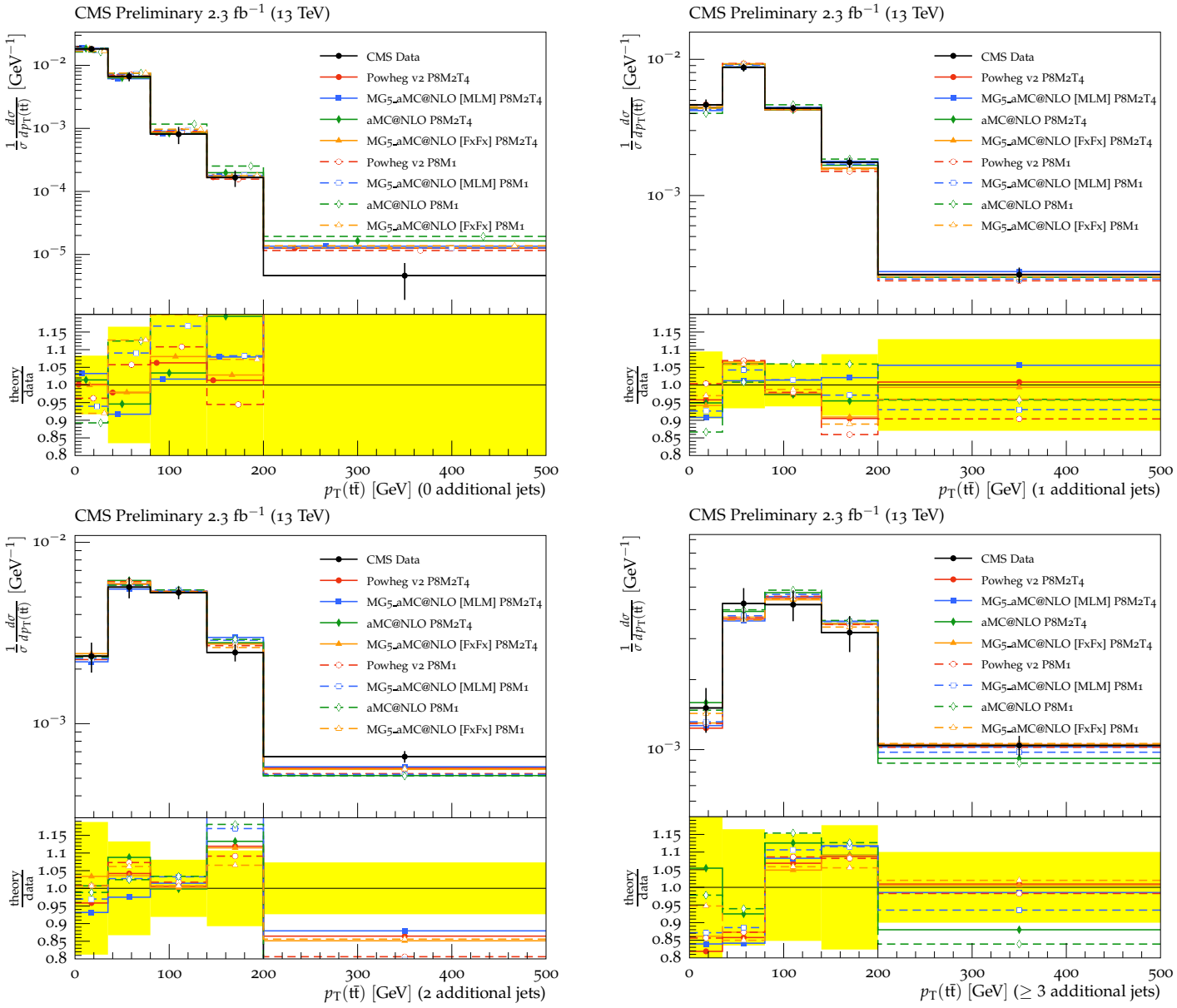
CMS-PAS-TOP-16-021





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

CMS-PAS-TOP-16-021



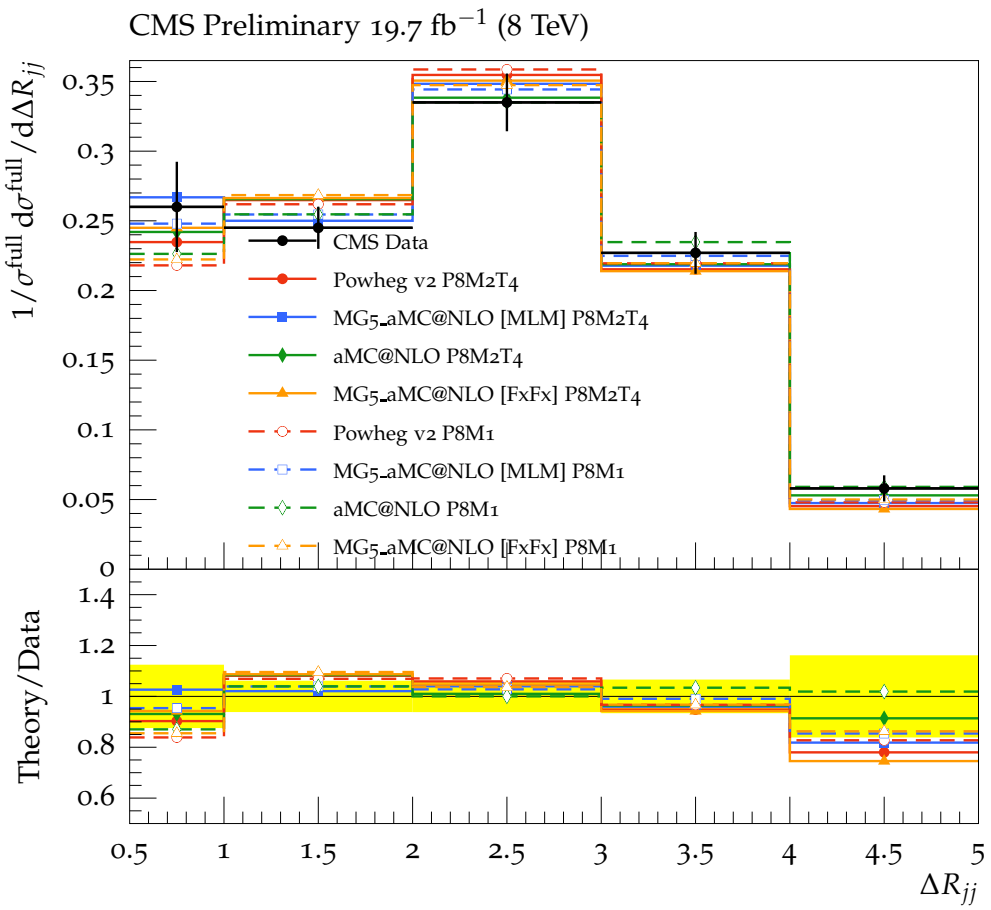
The 0-additional jet case has the worst agreement data and theory predictions.

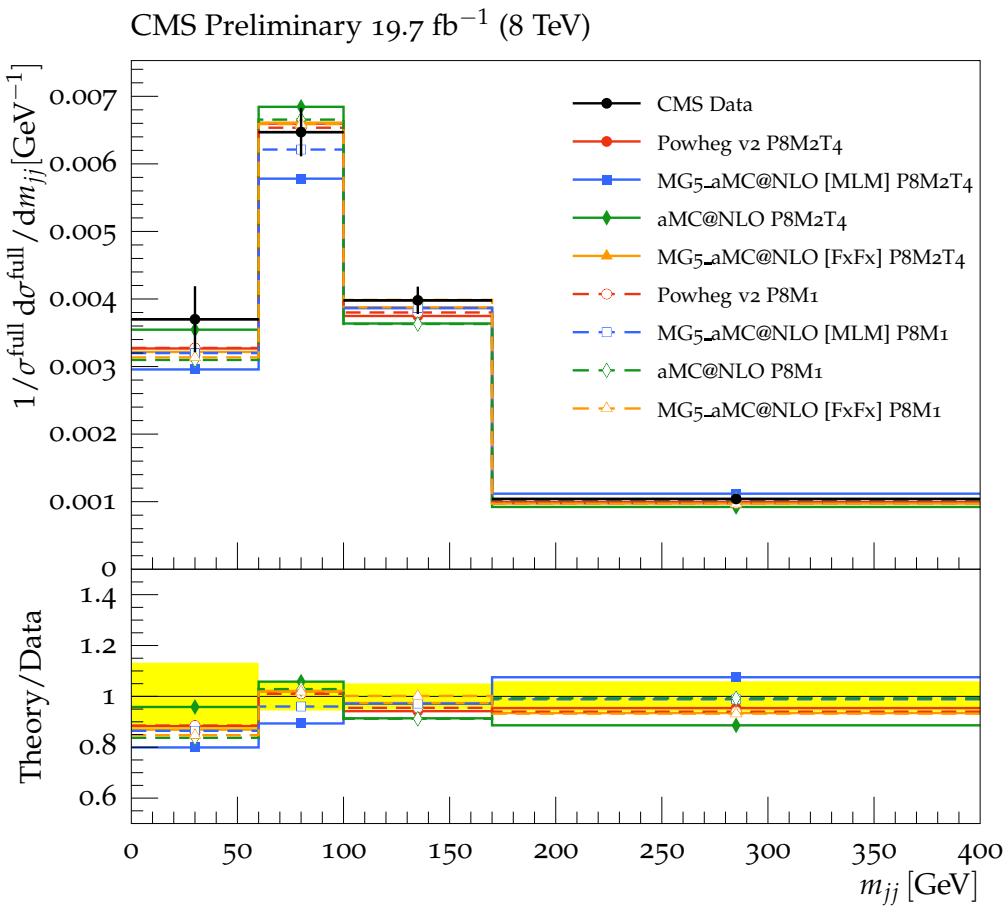
Additional jet distribution from CMS

8 TeV

CMS-PAS-TOP-16-021

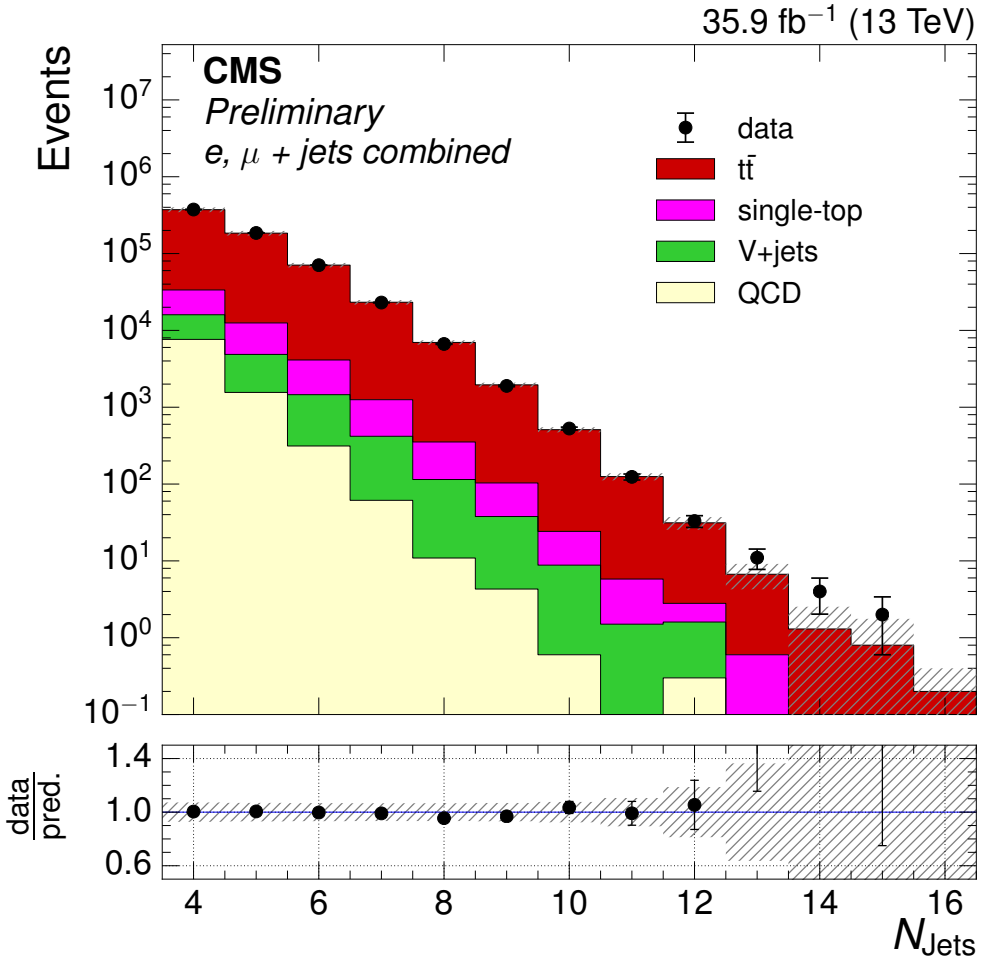
Data is from Eur. Phys. J. C 76 (2016) 379





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

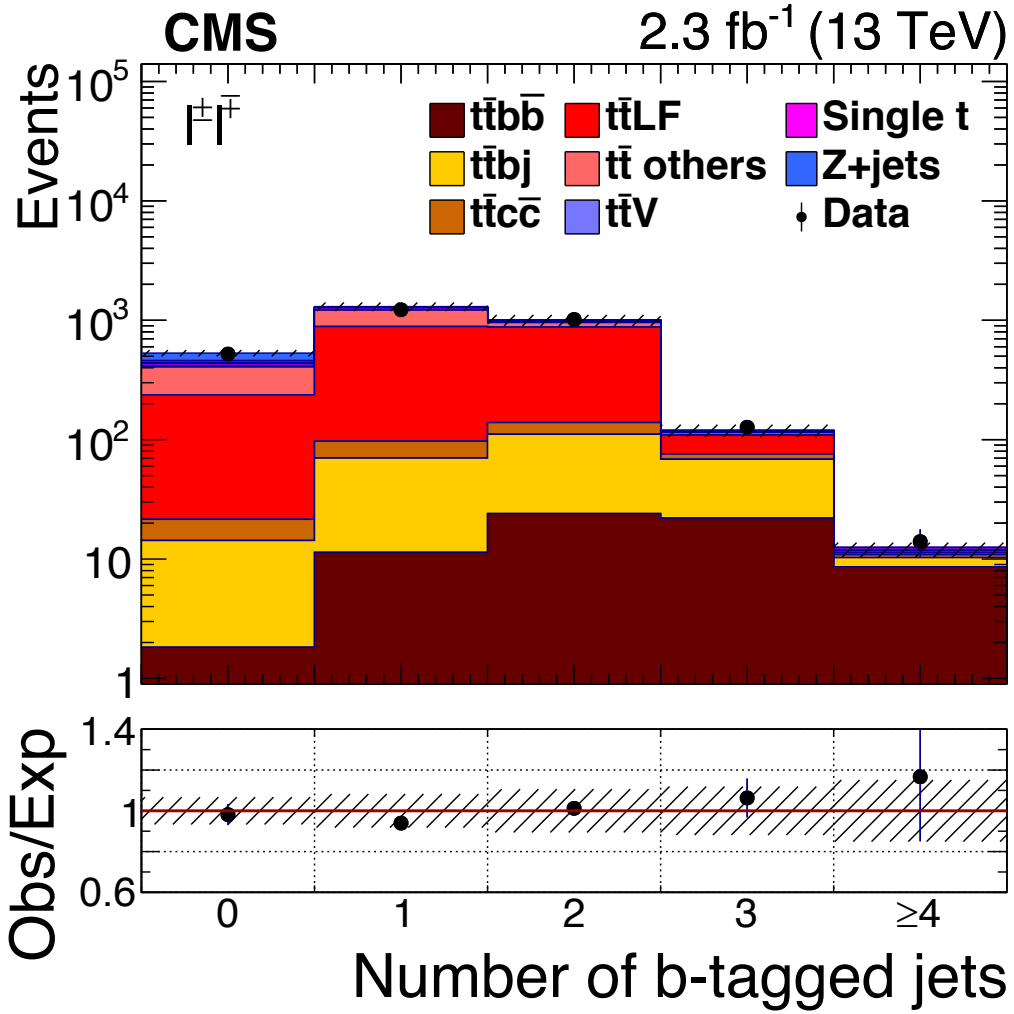
CMS-PAS-TOP-16-014



- 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 $\sqrt{s}$

Cut flow for $t\bar{t}b\bar{b}$ in dilepton decay mode

arXiv:1705.10141

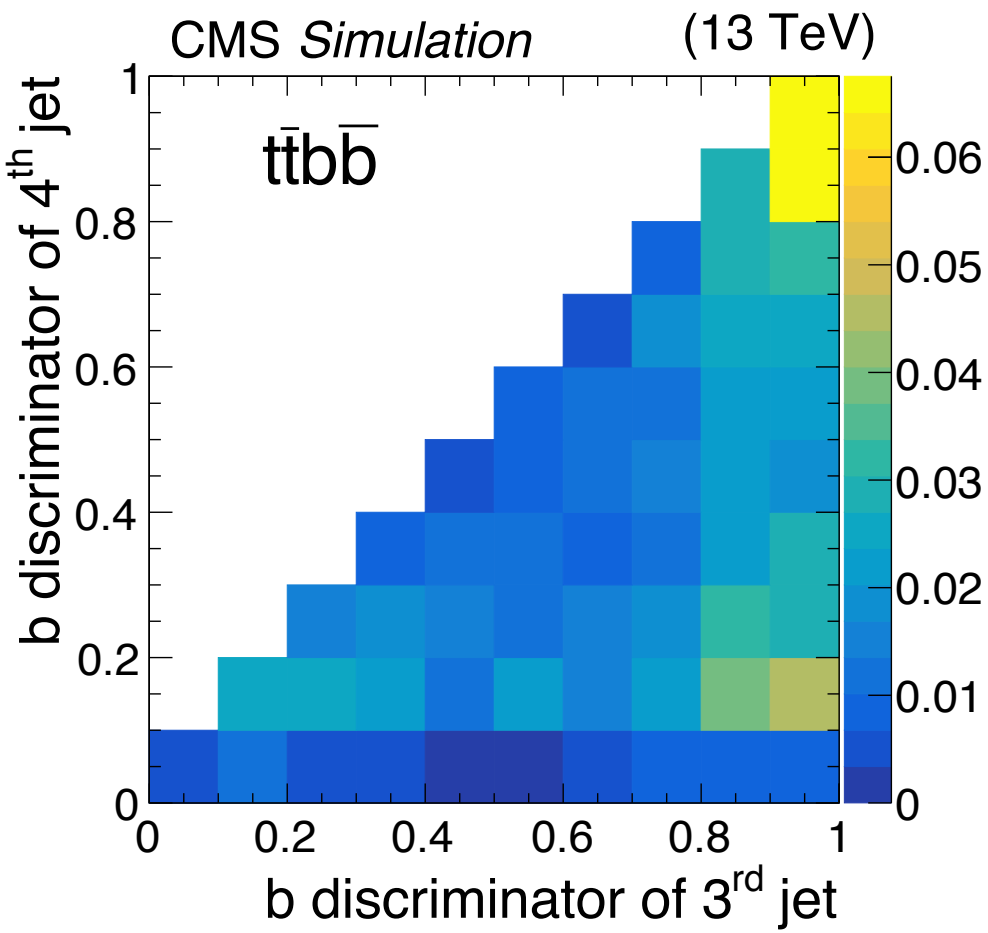


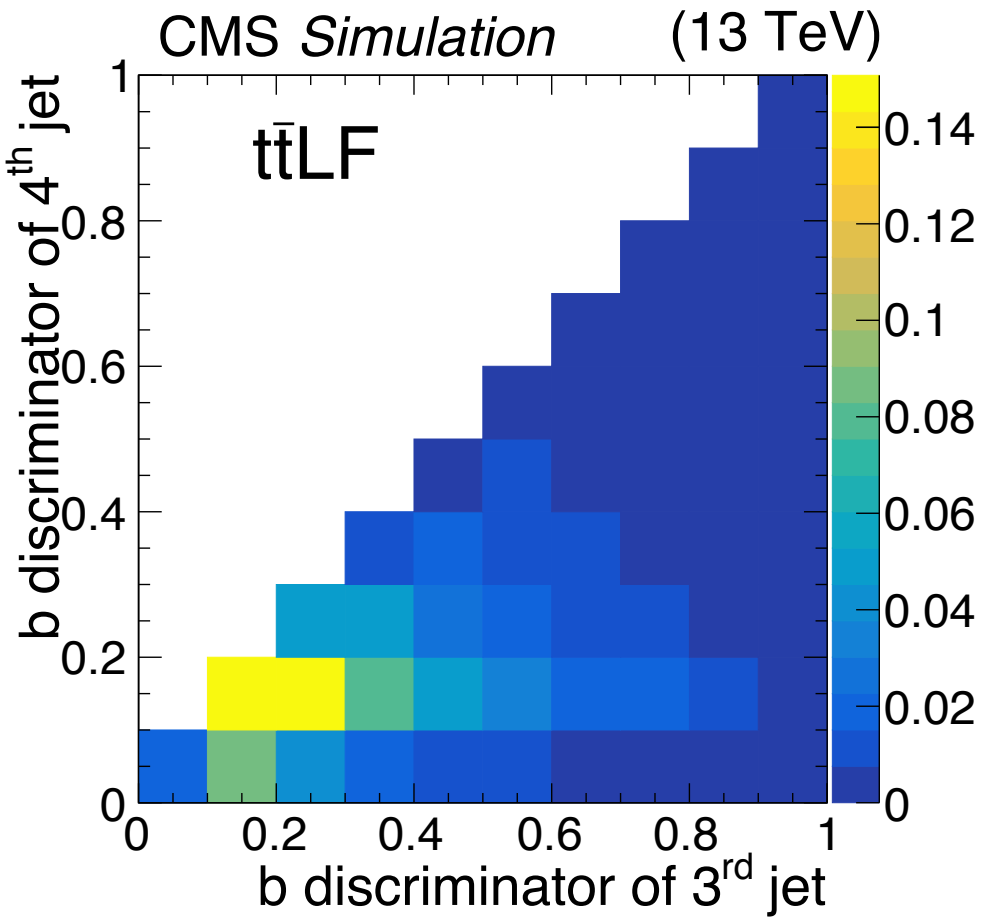
Process	e^+e^-	$\mu^+\mu^-$	$e^\pm\mu^\mp$	All
$t\bar{t}b\bar{b}$	6.3 ± 0.4	8.6 ± 0.4	24 ± 1	39 ± 1
$t\bar{t}bj$	16 ± 1	21 ± 1	57 ± 2	95 ± 2
$t\bar{t}c\bar{c}$	7.7 ± 0.4	11 ± 1	27 ± 1	46 ± 1
$t\bar{t}LF$	157 ± 2	220 ± 2	596 ± 3	972 ± 4
$t\bar{t}$ others	18 ± 1	19 ± 1	61 ± 1	99 ± 1
$t\bar{t}$ V	2.5 ± 0.1	3.2 ± 0.2	7.3 ± 0.2	14 ± 1
Single t	6.6 ± 0.8	8.4 ± 0.8	23 ± 2	39 ± 2
Z+jets	$0.8^{+1.0}_{-0.8}$	5.4 ± 1.5	0.6 ± 0.5	6.8 ± 1.9
Total	215 ± 2	297 ± 3	796 ± 4	1311 ± 6
Data	186	288	682	1156

$t\bar{t}$ + Heavy flavor from CMS

arXiv:1705.10141

- Rearrange jets in b-tagging algorithm discriminator
- Using b-tagging algorithm discriminator of third and fourth jets





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

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

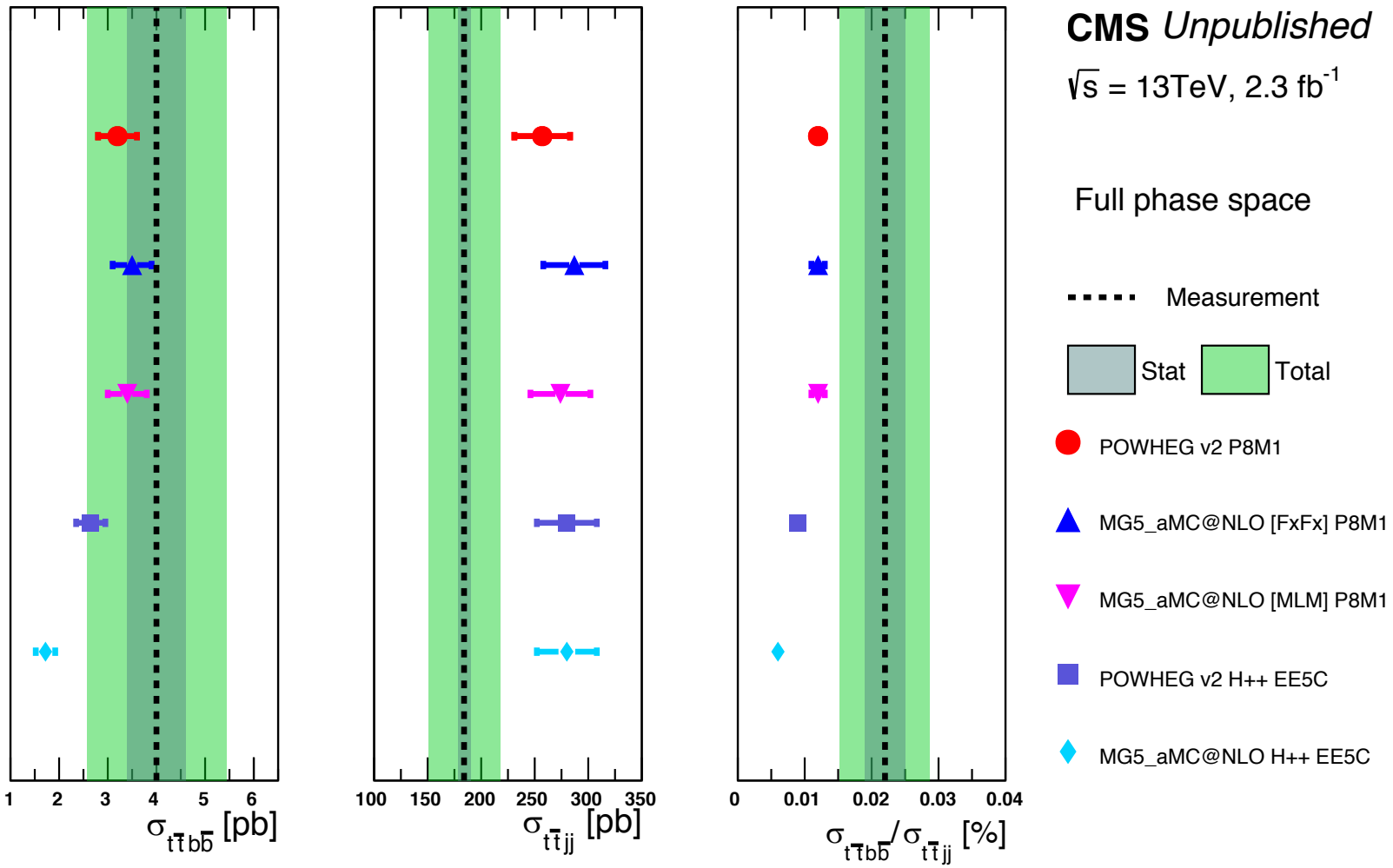
Phase space		$\sigma_{t\bar{t}b\bar{b}}$ [pb]	$\sigma_{t\bar{t}jj}$ [pb]	$\sigma_{t\bar{t}b\bar{b}}/\sigma_{t\bar{t}jj}$
Visible	Measurement	$0.088 \pm 0.012 \pm 0.029$	$3.7 \pm 0.1 \pm 0.7$	$0.024 \pm 0.003 \pm 0.007$
	SM (POWHEG)	0.070 ± 0.009	5.1 ± 0.5	0.014 ± 0.001
Full	Measurement	$4.0 \pm 0.6 \pm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

- Theoretical ratios are lower than the measured values
- But consistent within two standard deviations

Comparisons with various MC predictions

arXiv:1705.10141

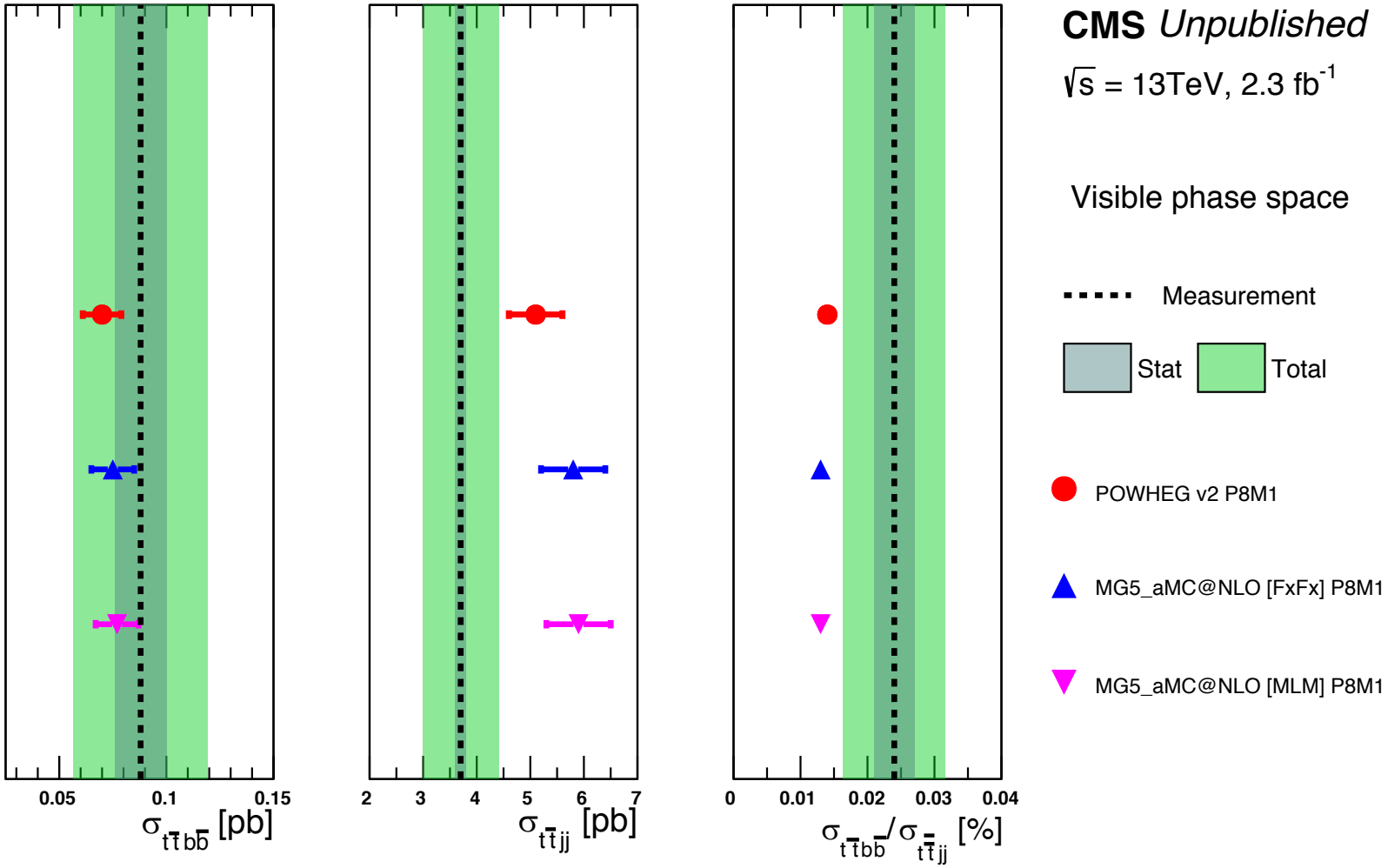
Full phase space



Comparisons with various MC predictions

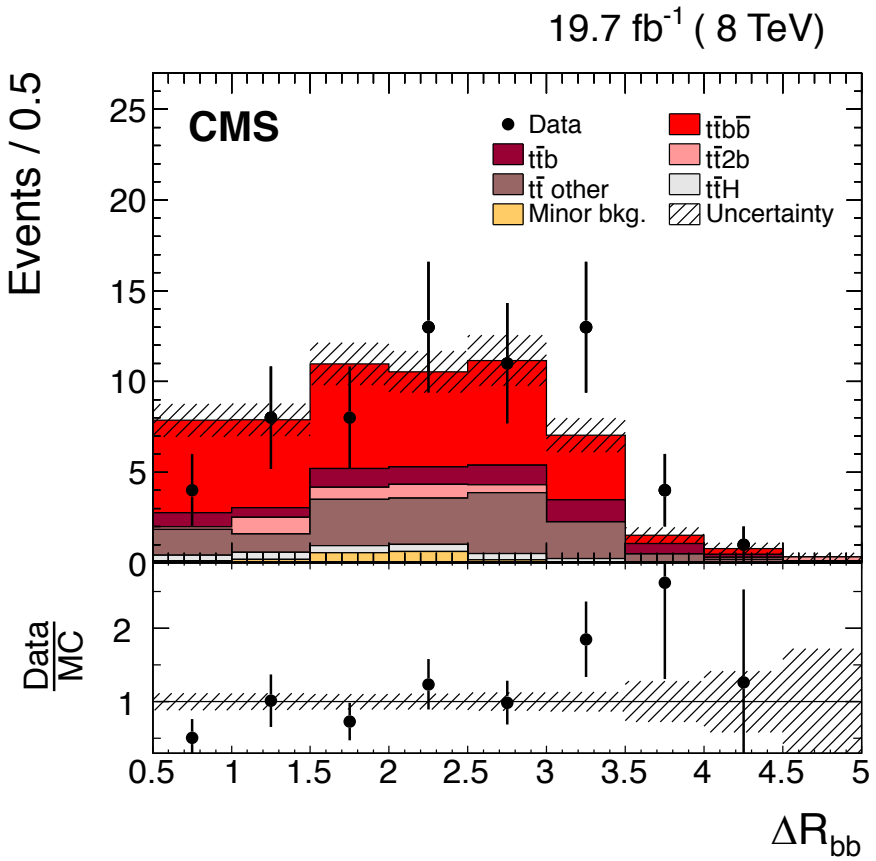
Visible phase space

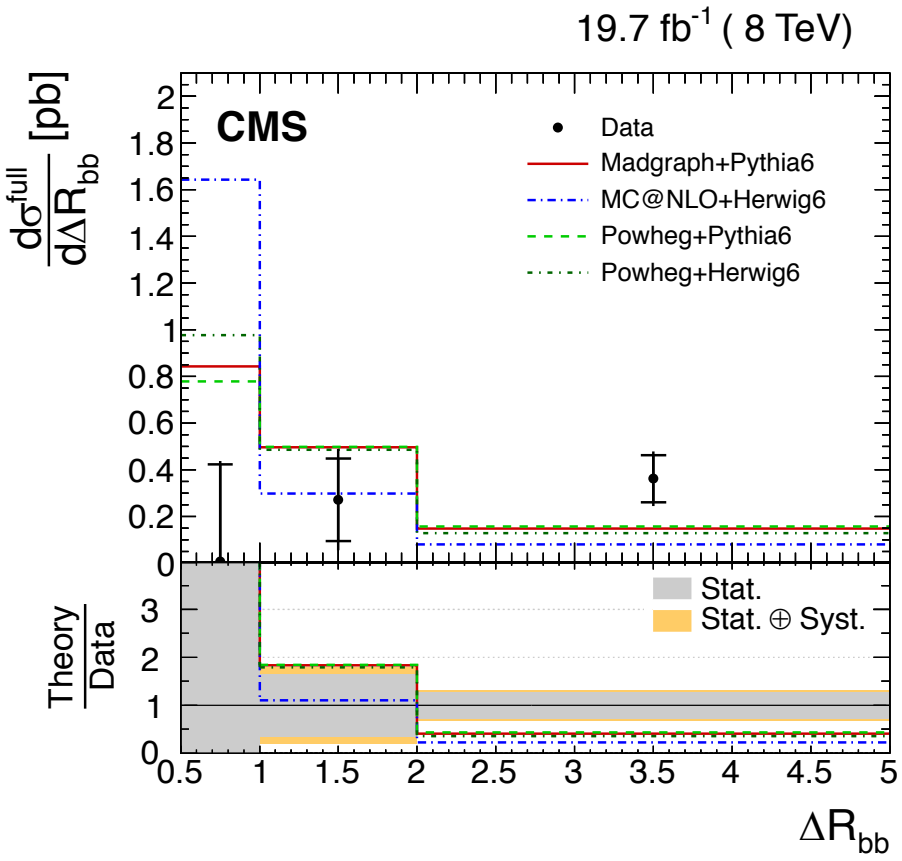
arXiv:1705.10141



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

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



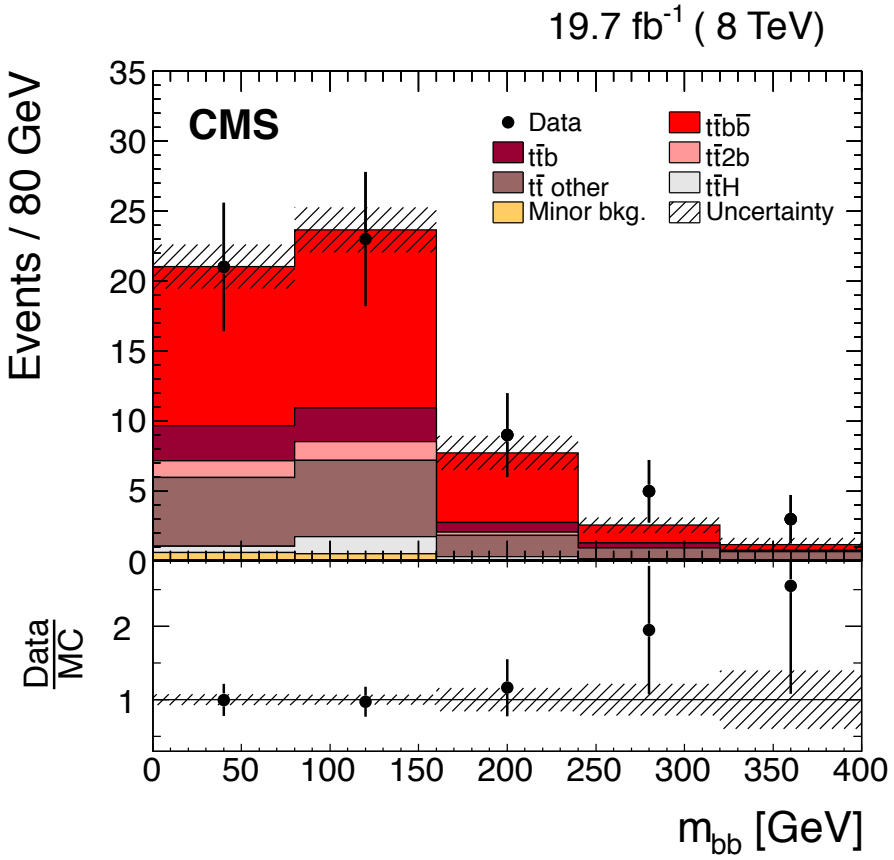


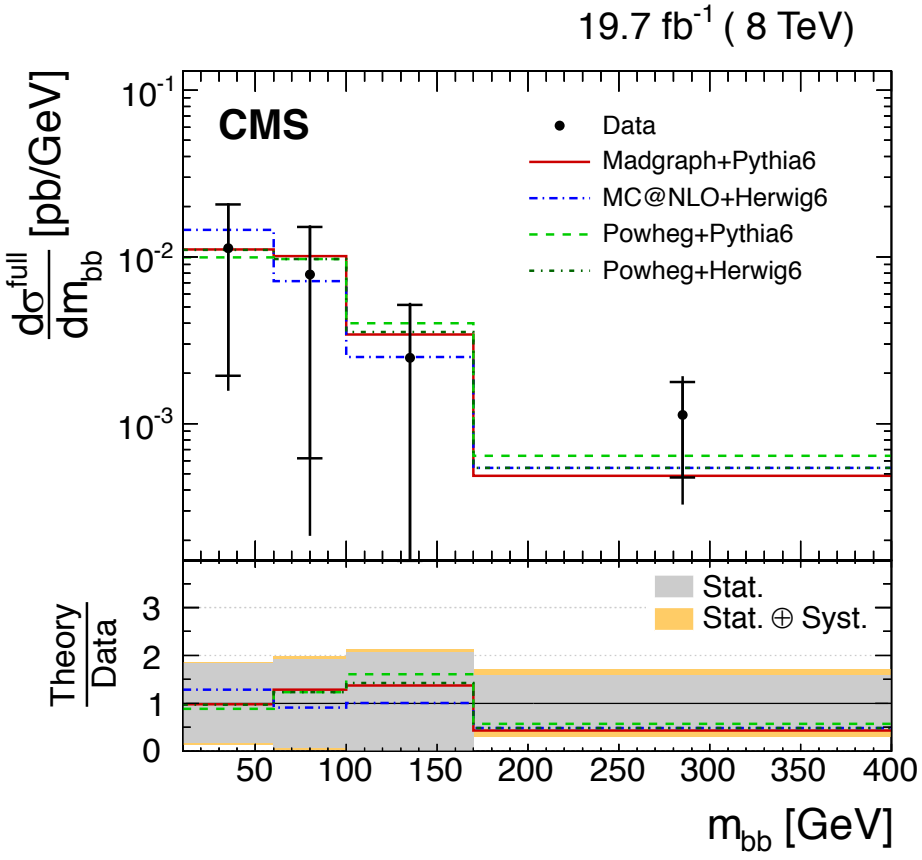
$$\Delta R = \sqrt{\Delta\phi(b,b)^2 + \Delta\eta(b,b)^2}$$

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

8 TeV

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





Conclusion

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