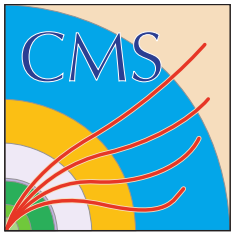
Universität Zürich[UZH]



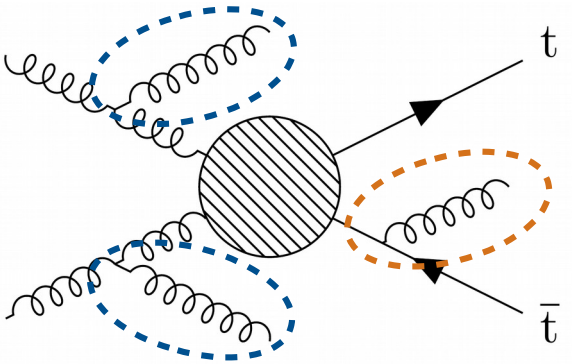
EPS-HEP 2019, Ghent, Belgium

12th July 2019

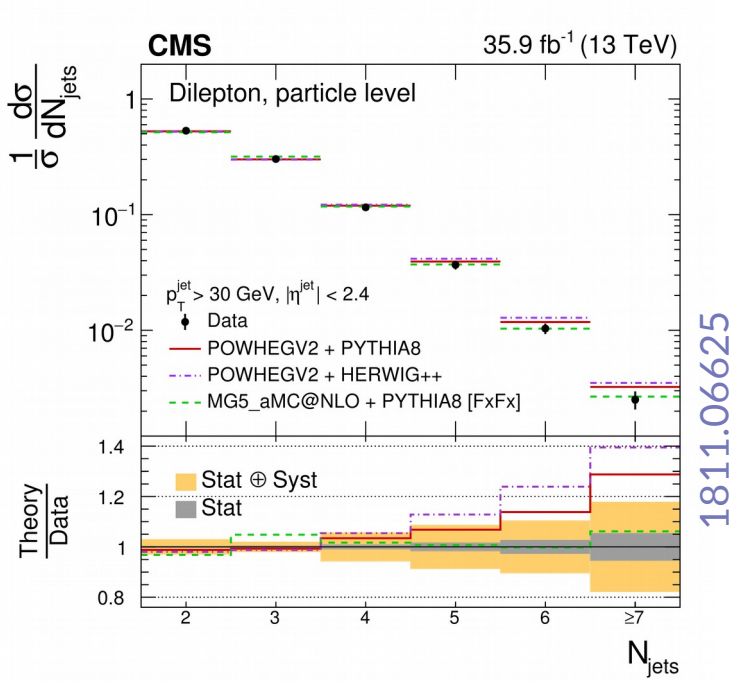
Measurements of $t\bar{t}$+jets, $t\bar{t}b\bar{b}$ and $t\bar{t}t\bar{t}$ production in CMS

Sébastien Wertz, for the CMS collaboration

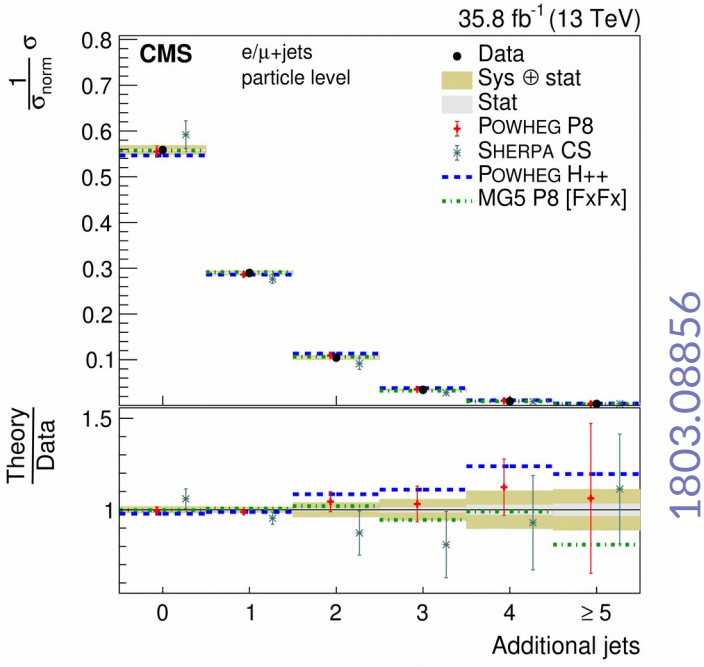
$t\bar{t}$+jets production: the basics



- Abundance of extra jets in $t\bar{t}$ production (ISR, FSR)
- Background to searches and measurements
  → important to model accurately & precisely
- Interplay of parton shower, multi-parton matrix elements, matching & merging schemes
  → need validation with measurements
- CMS measurements (2016 data, 36 fb^{-1}): jet multiplicity, additional jet properties



1811.06625



1803.08856

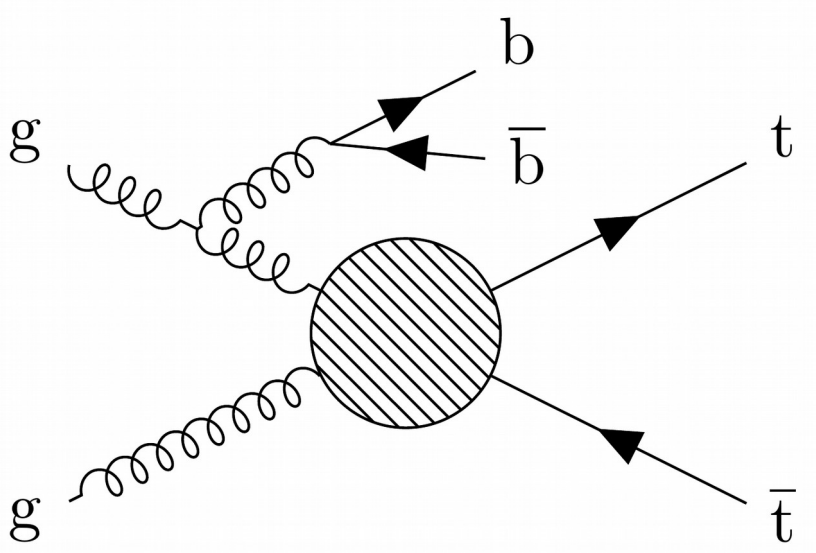
See also:
- 1803.03991
- Otto's talk
- Roberto's talk

- Trend of lower multiplicity in data (within uncertainties) than in Powheg+Pythia8, good agreement up to 3 extra jets

$t\bar{t}b\bar{b}$ production: the next level

- $t\bar{t}b\bar{b}$ not just a subset of $t\bar{t}jj$: unique modelling challenges!
  - Non-negligible b quark mass: multiple massive partons
  - Different scales in process ($t\bar{t}$ and $b\bar{b}$ systems)
- Different modelling approaches:
  - NLO $t\bar{t}$ ME + PS, 5FS
  - Merged NLO $t\bar{t}$ + 0/1/2 jets ME + PS, 5FS
  - NLO $t\bar{t}b\bar{b}$ ME + PS, 4FS

ME = Matrix Element, PS = Parton Shower
5FS = 5-flavour scheme, massless b quarks (inside PDFs)
4FS = 4-flavour scheme, massive b quarks



Models (mostly) in agreement,
but large uncertainties

→ Leading systematic for $t\bar{t}H(b\bar{b})$, $t\bar{t}t\bar{t}$

→ **Need measurements**

Previous 13 TeV CMS result: 1705.10141

- Dilepton final state, 2015 data (2.3 fb^{-1})
- Inclusive $t\bar{t}b\bar{b}$ cross section, $t\bar{t}b\bar{b}/t\bar{t}jj$ ratio
- 35% precision, $\sigma/\sigma_{th} \sim 1.25$

$t\bar{t}b\bar{b}$ cross section: all-jet final state

PAS-TOP-18-011

2016 data, 36 fb^{-1}

- Fully-hadronic channel: 8 jets, of which 4 b jets
  → largest branching fraction (45%) & fully reconstructible final state
- Select events with ≥ 8 jets, ≥ 2 b-tagged jets
- Suffers from:
  - Combinatorial self-background

<u>"Permutation" BDT:</u>

- Trained to identify jets from $t\bar{t}$ decays
- ~60% correct (if all jets reconstructed)
- Keep permutation with highest score

$t\bar{t}b\bar{b}$ cross section: all-jet final state

PAS-TOP-18-011

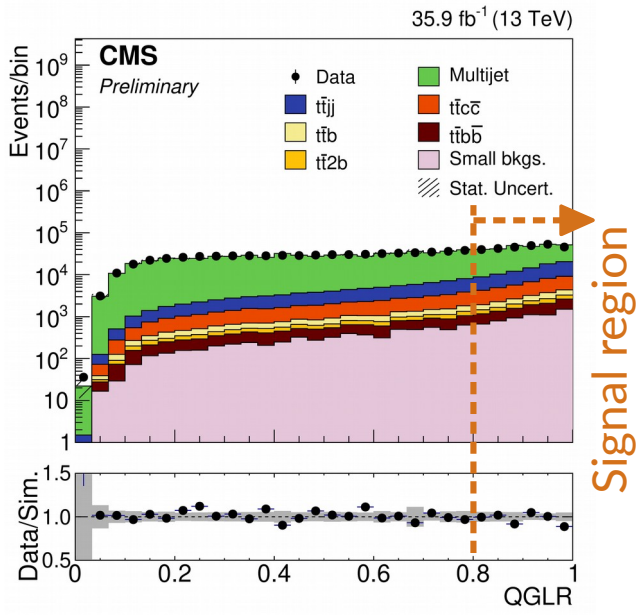
NEW!

2016 data, 36 fb^{-1}

- Fully-hadronic channel: 8 jets, of which 4 b jets
  → largest branching fraction (45%) & fully reconstructible final state
- Select events with ≥ 8 jets, ≥ 2 b-tagged jets
- Suffers from:
  - QCD multijet background

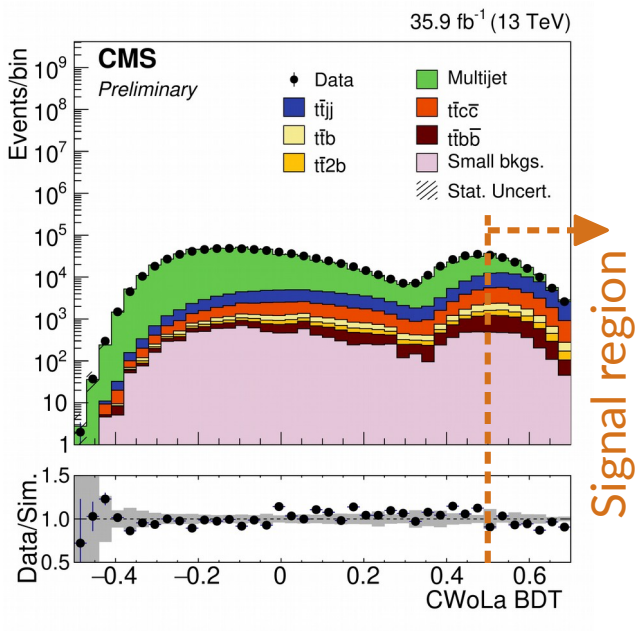
Quark-gluon likelihood ratio (QGLR):

- QGL: discriminates quark ↔ gluon jet
- QCD multijet: more gluon jets



QCD rejection BDT:

- Uses classification without labels (CWoLa)
  1708.02949
- Trained using data with =7 jets

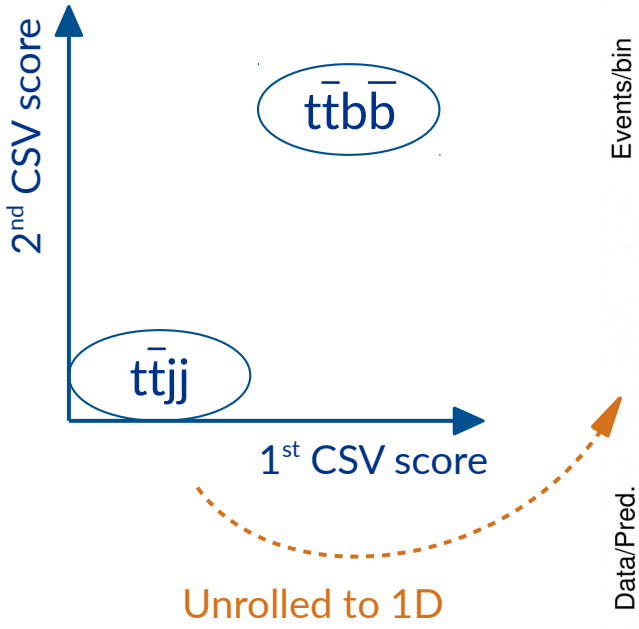


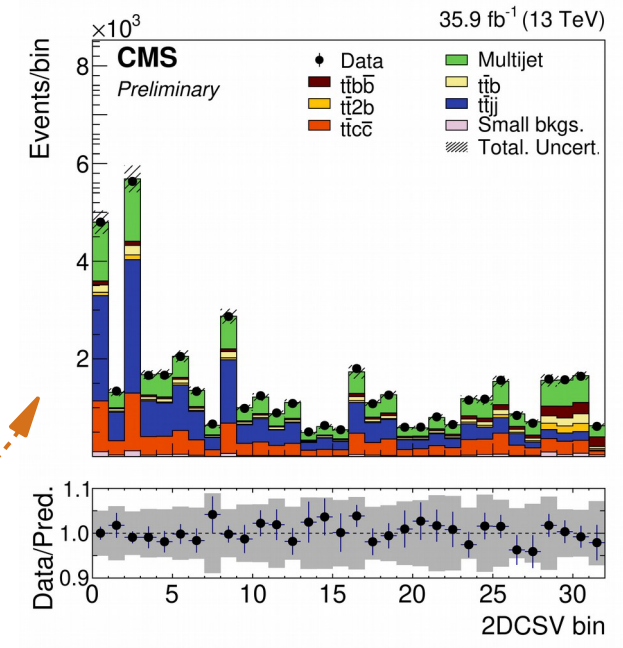
$t\bar{t}b\bar{b}$ cross section: all-jet final state

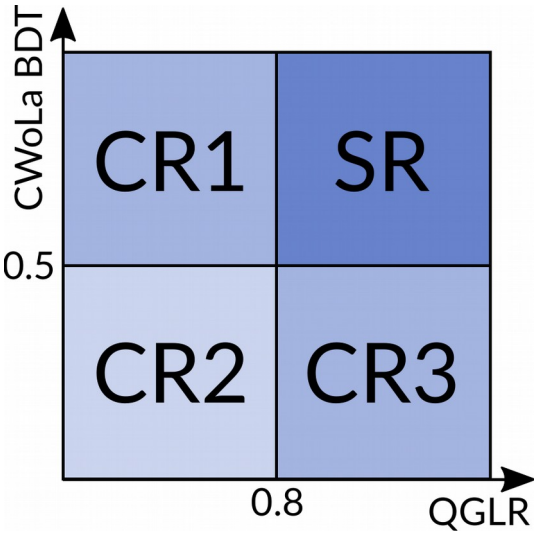
PAS-TOP-18-011

NEW!

- Extract signal using b tagging (CSV) discriminant of two extra jets with highest score:
- Estimate QCD contribution: "bin-wise ABCD"
- QGLR and CWoLa BDT uncorrelated
- For each bin: $N^{SR} = N^{CR3}\, N^{CR1}/N^{CR2}$





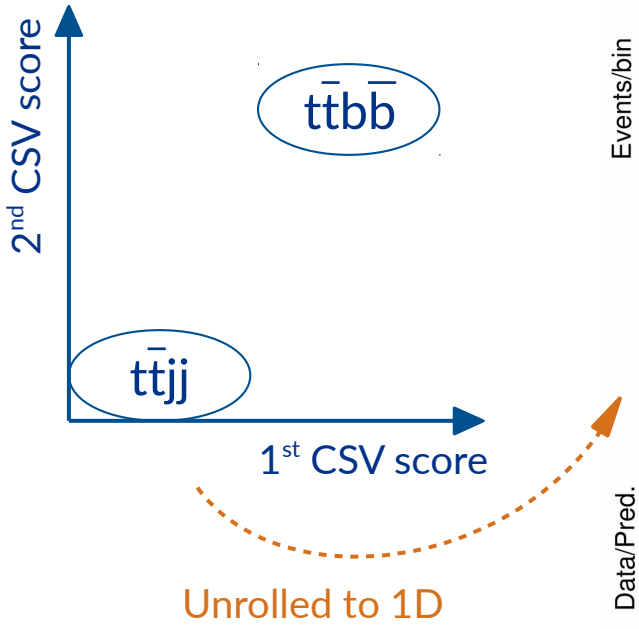


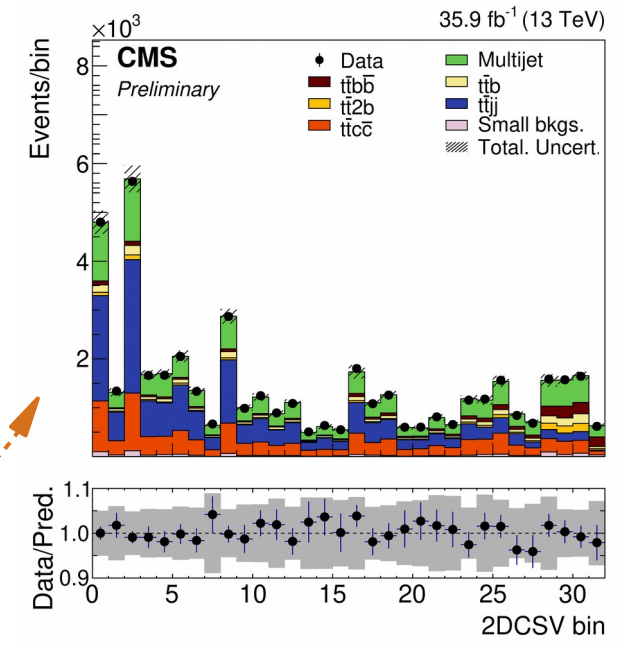
$t\bar{t}b\bar{b}$ cross section: all-jet final state

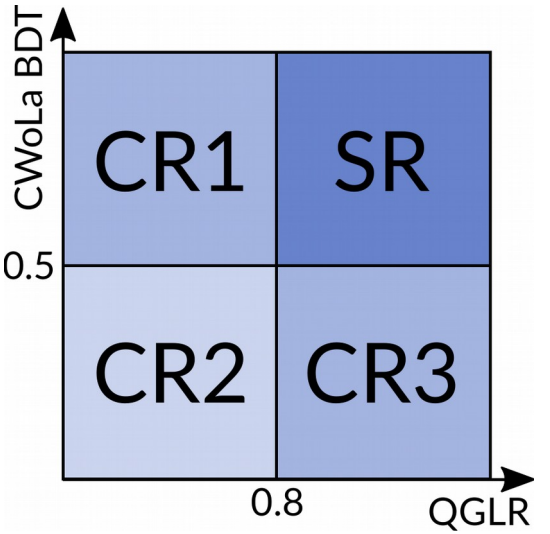
PAS-TOP-18-011

NEW!

- Extract signal using b tagging (CSV) discriminant of two extra jets with highest score:
- Estimate QCD contribution: "bin-wise ABCD"
- QGLR and CWoLa BDT uncorrelated
- For each bin: $N^{SR} = N^{CR3}\, N^{CR1}/N^{CR2}$







- Cross section measured for three phase spaces:

Visible, Parton Agnostic (PA):

- ≥ 8 jets
- ≥ 4 b jets

Visible, Parton Based (PB):

- ≥ 8 jets
- ≥ 4 b jets, of which ≥ 2 not coming from top decays

Full:

≥ 2 b jets not coming from top decays

(particle-level jets with p_T>20 GeV and |η|<2.4)

$t\bar{t}b\bar{b}$ cross section: all-jet final state

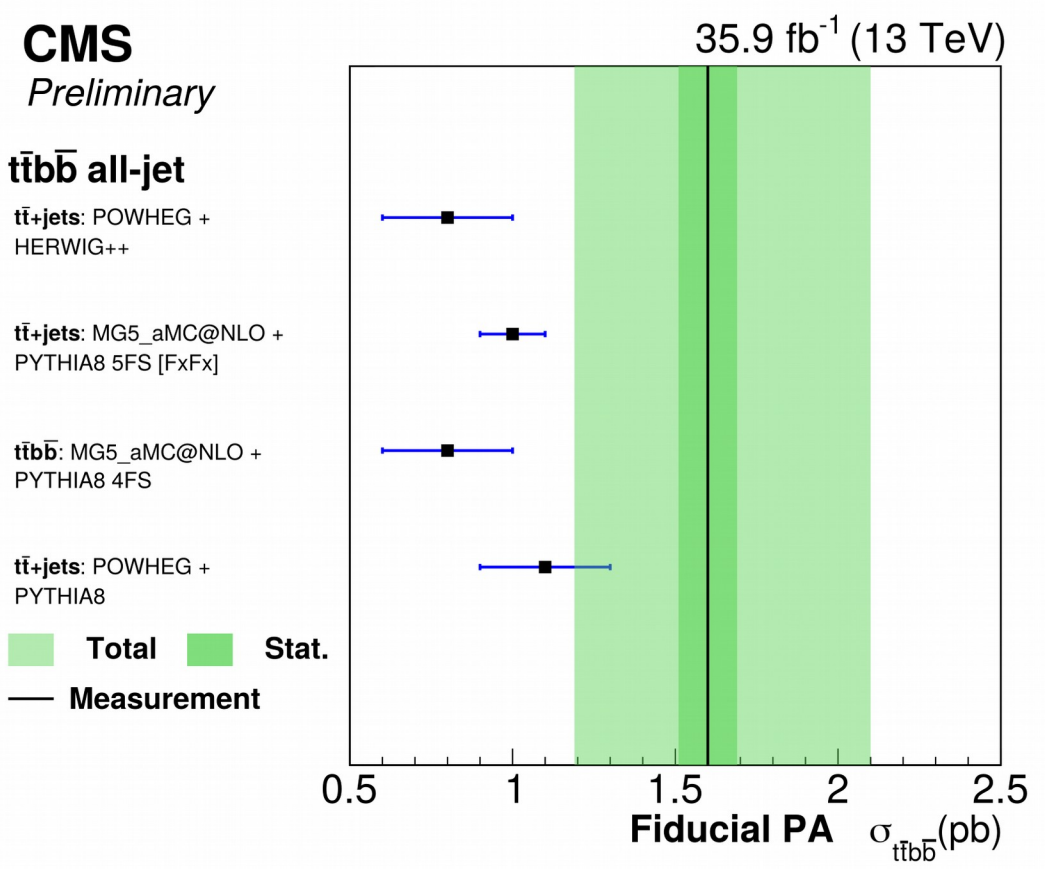
PAS-TOP-18-011

NEW!

- Systematic uncertainties (in %):

Source	VPS (PA)	
Simulated sample size	$+15/-11$	
Quark-gluon likelihood	$+13/-8$	
b tagging	± 10	
JES & JER	$+5.1/-5.2$	
Integrated luminosity	$+2.8/-2.2$	
Trigger efficiency	$+2.6/-2.1$	
Pileup	$+2.3/-2.0$	
μ_R and μ_F scales	$+13/-9$	Modelling
Parton shower scale	$+11/-8$	Modelling
UE tune	$+9.0/-5.3$	Modelling
Colour reconnection	± 7.2	Modelling
Shower matching (h_{damp})	$+4.3/-2.8$	
$t\bar{t}c\bar{c}$ normalisation	$+3.2/-4.4$	
Top quark p_T modelling	± 2.5	
PDFs	$+2.2/-2.0$	

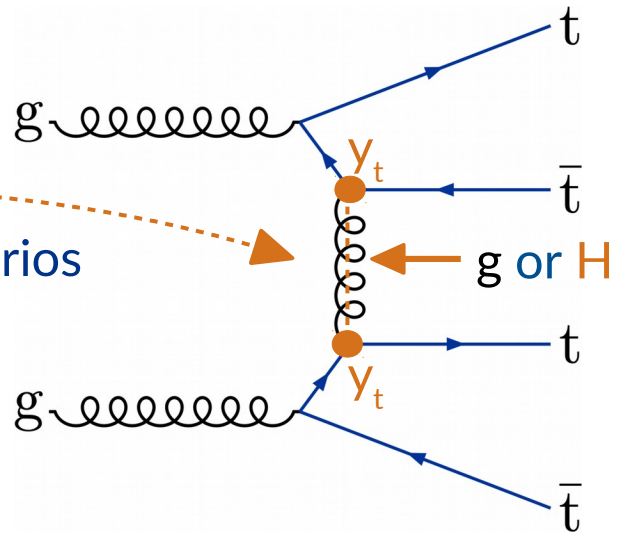
- Results:



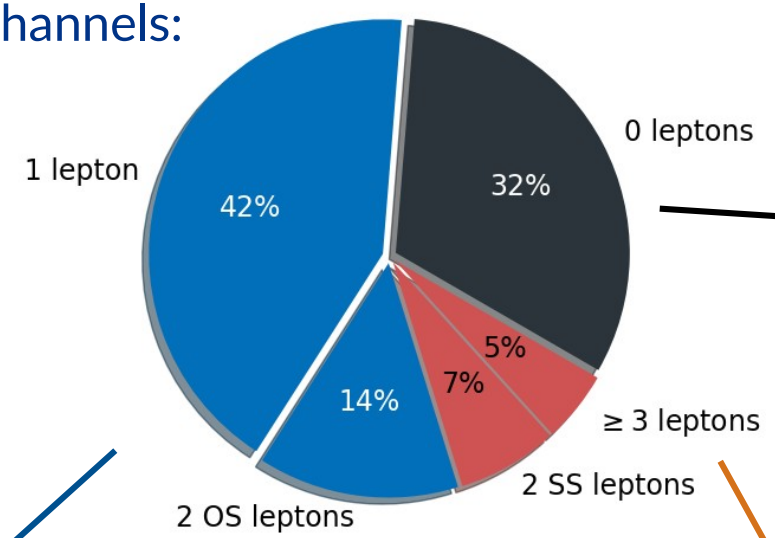
- **Under-predictions** of all generators by factor ~1.5
- Consistent with previous results
- Note: NNLO K-factor not applied (1.1-1.15)

$t\bar{t}t\bar{t}$ production: LHC goes to eleven

- Rare SM process: σ = 12±2 fb (NLO QCD + EWK)
- Probe top quark Yukawa coupling y_t
- Cross section can be enhanced in numerous BSM scenarios
  - 2HDMs, dark matter mediator, EFT, ...



- Variety of decay channels:



Not yet considered

- "High" rate, low purity
- Main background: $t\bar{t}$ + (b) jets

- Low rate, high purity
- Main backgrounds: $t\bar{t}W$, $t\bar{t}Z$ (+ b jets!), $t\bar{t}H$, nonprompt leptons

$t\bar{t}t\bar{t}$ production: 1, 2 (OS) leptons

1906.02805

NEW!

2016 data, 36 fb^{-1}

1 lepton:

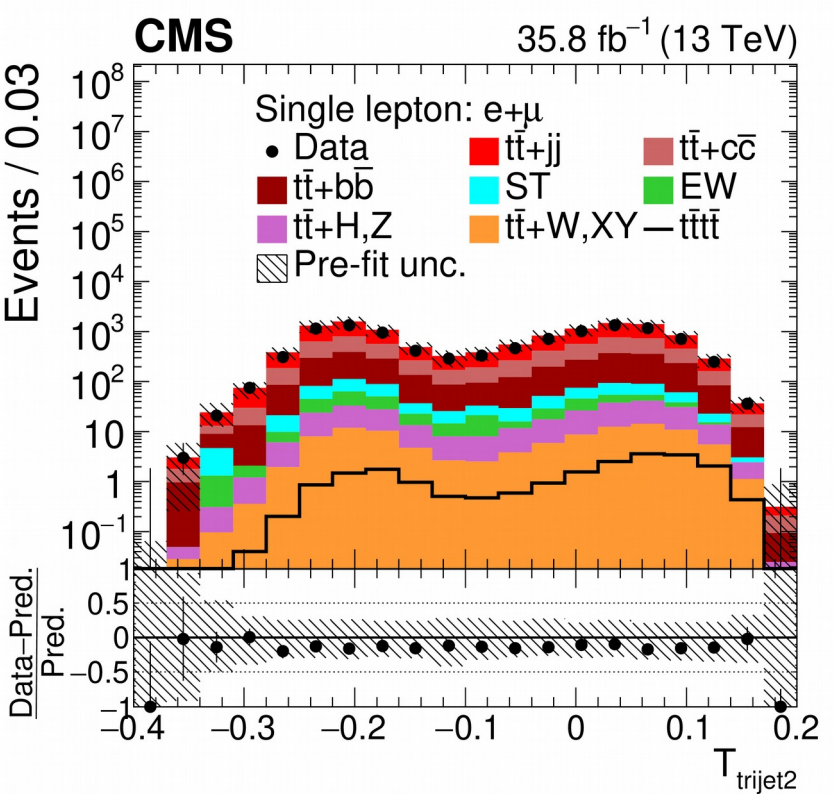
- ≥ 7 jets ($\mu^{\pm}$), ≥ 8 jets ($e^{\pm}$)
- ≥ 2 b-tagged jets

2 OS leptons:

- $\mu^{+}\mu^{-}$, $e^{+}e^{-}$: veto Z peak
- ≥ 4 jets, ≥ 2 b-tagged jets

Hadronic top quark BDT:

- Identify 3-jet combinations: top quark decays
- Uses 2- and 3-jet inv. masses, angles, b tagging
- Trained using $t\bar{t}$+jets simulation
- Use largest (2l), 2nd largest (1l) scores



$t\bar{t}t\bar{t}$ production: 1, 2 (OS) leptons

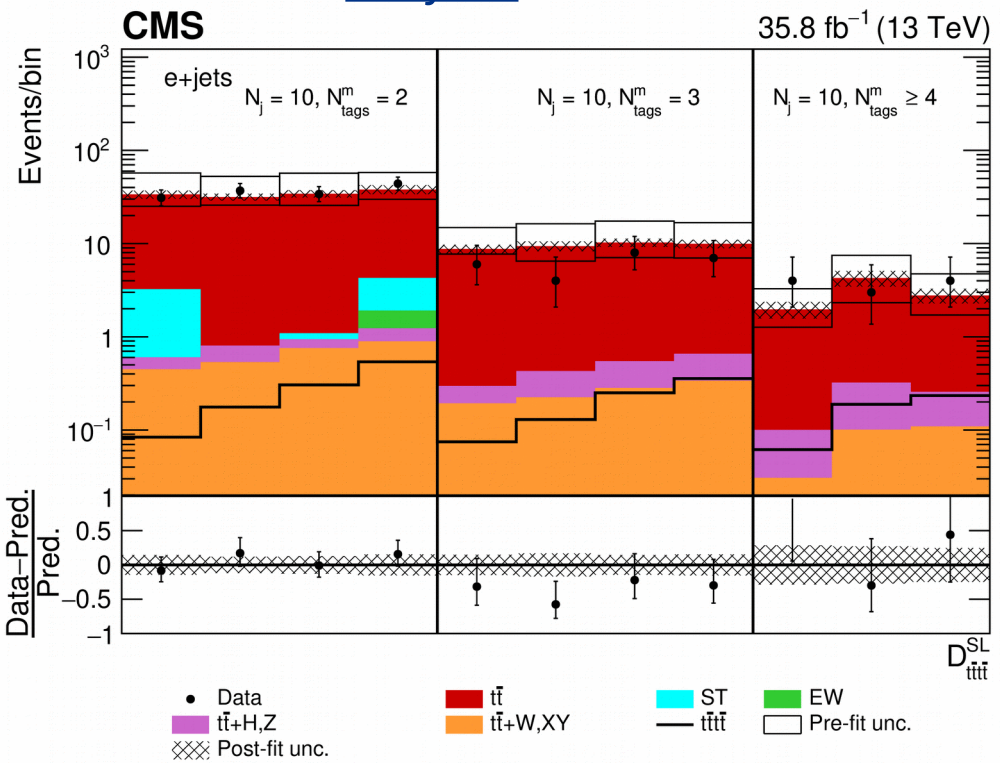
1906.02805

NEW!

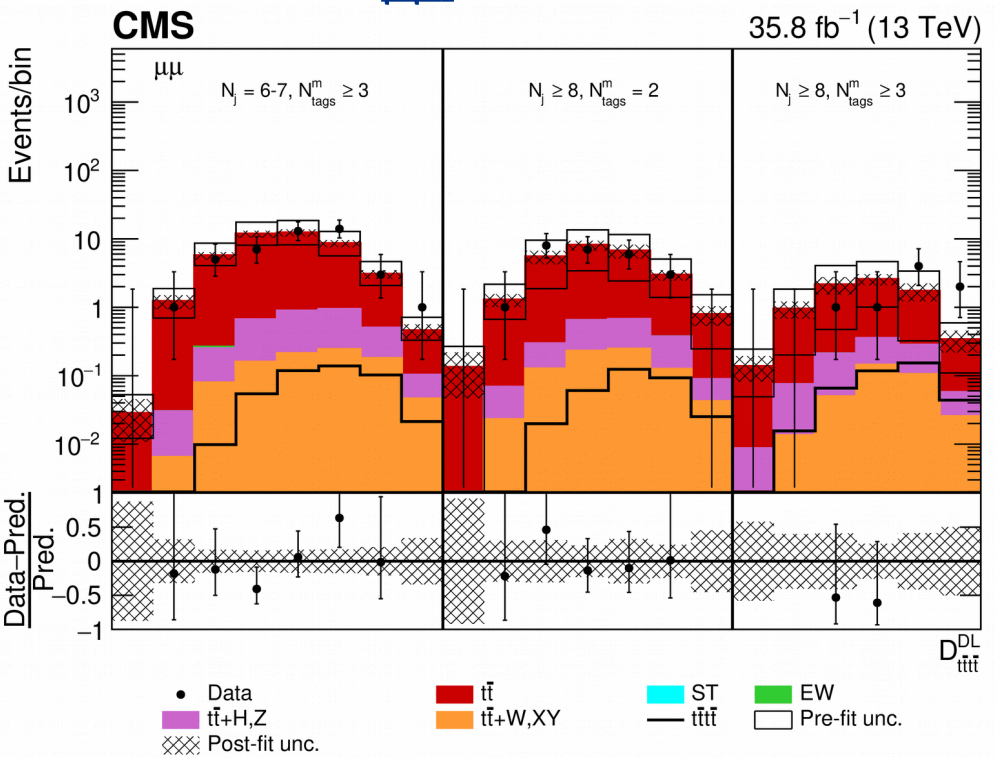
BDTs for sig/bkg discrimination:

- Use trijet discriminants, event hadronic activity, event topology, lepton kinematics…
- Several BDTs trained (lepton flavours, jet bins)
- Extract signal: BDT shape in jet and b-tag multiplicity bins

e + jets:



μμ:



$t\bar{t}t\bar{t}$ production: 1, 2 (OS) leptons

1906.02805

NEW!

Systematic uncertainties: dominated by **$t\bar{t}$+jets modelling**

- Jet multiplicity correction, from single-lepton channel
- Extra 35% uncertainty on $t\bar{t}b\bar{b}$ rate, taken from measurement (older dilepton CMS result)
- Parton shower scale, underlying event tune

Results:

	single lepton	dilepton	combined
Exp./obs. Significance (σ)*	0.21/0.36	0.36/0.00	0.40/0.00
Exp./obs. limit on signal strength*	9.4/10.6	7.3/6.9	5.7/5.2

- Sensitivity dominated by dilepton channels
- Comparable statistical and systematic uncertainties

Also EFT reinterpretation: see Kirill's talk!

* Note: based on $\sigma_{t\bar{t}t\bar{t}}$ = 9 fb (NLO QCD)

$t\bar{t}t\bar{t}$ production: 2 (SS), 3 leptons

PAS-TOP-18-003

Uses **full Run 2 dataset: 137 fb^{-1}**

Baseline selection:

- ≥ 2 jets, ≥ 2 b-tagged jets
- 2 SS leptons
- Advanced lepton identification & isolation criteria
- If present, 3rd lepton (low p_T)
  → to $t\bar{t}Z$ control region if same-flavour m(ll) close to Z peak

2 analyses

Cut based

- 14 Signal Regions (SRs), based on number of jets, leptons, b-tagged jets
- $t\bar{t}Z$ and $t\bar{t}W$-enriched control regions (CRZ, CRW)

BDT based

- Shape of BDT (17 SRs)
- $t\bar{t}Z$-enriched control region (CRZ)

$t\bar{t}t\bar{t}$ production: 2 (SS), 3 leptons

NEW!

PAS-TOP-18-003

Background estimation

Prompt leptons: $t\bar{t}W$, $t\bar{t}Z$, $t\bar{t}H(WW)$

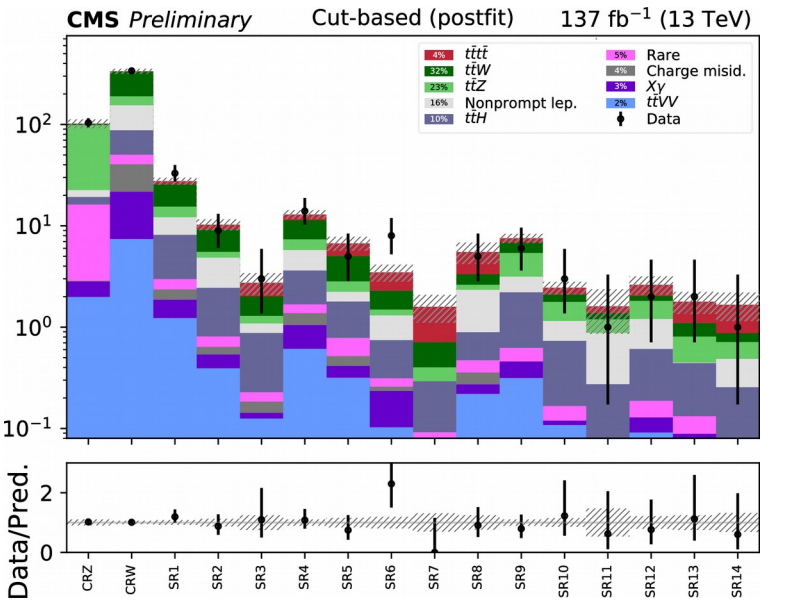
- Simulation + control regions (for $t\bar{t}V$)
- **Jet multiplicity** correction for $t\bar{t}V$ (from dilepton $t\bar{t}$ measurement)
- Correction for **additional b jets**, from measured $t\bar{t}b\bar{b}/t\bar{t}jj$ ratio (older dilepton CMS result)

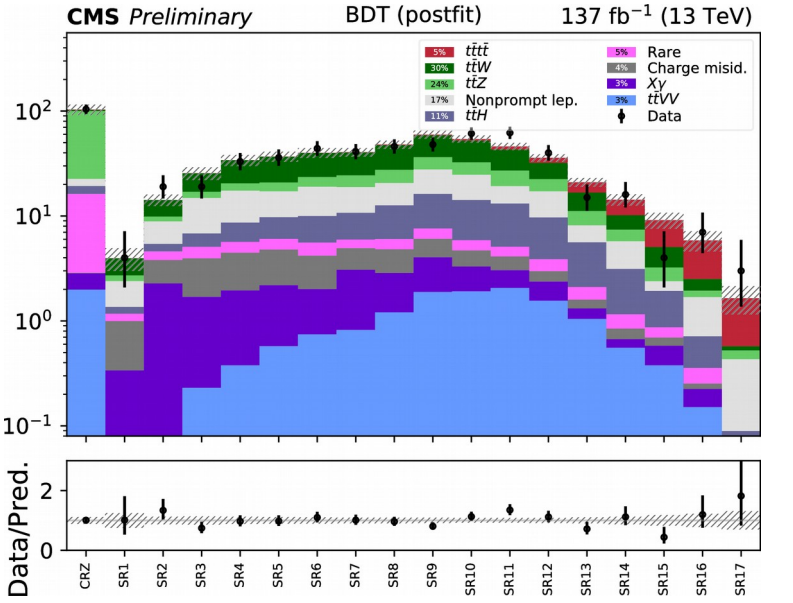
Nonprompt leptons:

- $t\bar{t}$+jets → nonprompts, γ conversions, jet mis-ID'd as lepton
- Invert ID & isolation
- Mistag rate measured in single-lepton sample

Lepton charge flips:

- Negligible for muons
- Electrons: simulation corrected using $Z(e^+e^-)$ events





$t\bar{t}t\bar{t}$ production: 2 (SS), 3 leptons

PAS-TOP-18-003

Dominant systematic uncertainties:

- **$t\bar{t}b\bar{b}$/$t\bar{t}jj$ ratio**
- Jet energy scale & resolution, b tagging
- $t\bar{t}H$ normalisation

Results:

	$\sigma_{t\bar{t}t\bar{t}}$ (fb)	Significance (exp/obs.)
Cut based	$9.4^{+6.2}_{-5.6}$	2.5/1.7
BDT based	$12.6^{+5.8}_{-5.2}$	2.7/2.6

$t\bar{t}t\bar{t}$ production: 2 (SS), 3 leptons

NEW!

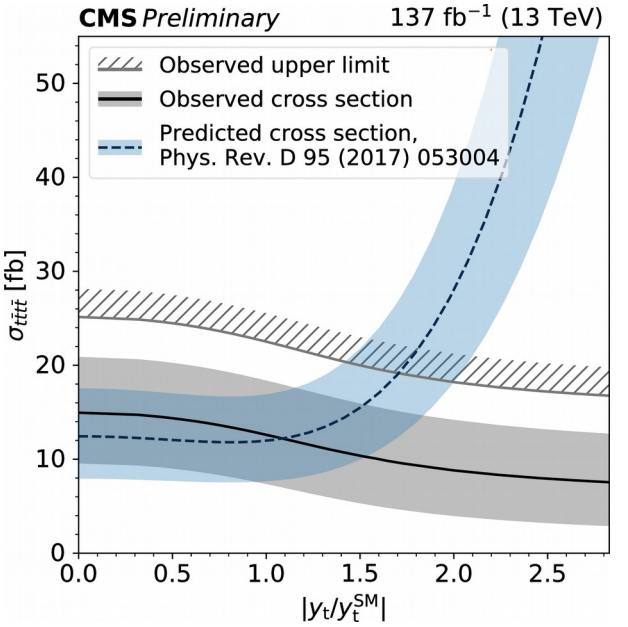
PAS-TOP-18-003

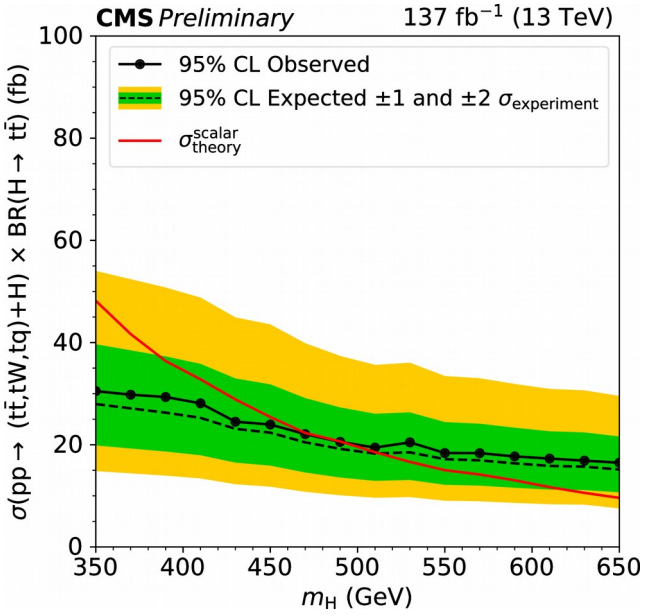
Dominant systematic uncertainties:

- **$t\bar{t}b\bar{b}$/$t\bar{t}jj$ ratio**
- Jet energy scale & resolution, b tagging
- $t\bar{t}H$ normalisation

Results:

	$\sigma_{t\bar{t}t\bar{t}}$ (fb)	Significance (exp/obs.)
Cut based	$9.4^{+6.2}_{-5.6}$	2.5/1.7
BDT based	$12.6^{+5.8}_{-5.2}$	2.7/2.6





Reinterpretation:

- Top quark Yukawa coupling: $|y_t/y_{SM}| < 1.7$ @ 95% CL
  → take into account effect on $t\bar{t}H$ background
- (Pseudo)scalar A/H → $t\bar{t}$

Conclusions

Extensive activity by CMS on $t\bar{t}$ production with additional coloured objects

- Important validation of additional jets modelling
- Increasing precision on $t\bar{t}b\bar{b}$ production
- Closing in on four top quark production

Feeds in

Differential measurements of $t\bar{t}b\bar{b}$ and $t\bar{t}t\bar{t}$ searches not yet systematics-limited

→ will profit from full Run 2 dataset, upcoming Run 3: stay tuned!

Conclusions

Extensive activity by CMS on $t\bar{t}$ production with additional coloured objects

- Important validation of additional jets modelling
- Increasing precision on $t\bar{t}b\bar{b}$ production
- Closing in on four top quark production

Feeds in

Differential measurements of $t\bar{t}b\bar{b}$ and $t\bar{t}t\bar{t}$ searches not yet systematics-limited

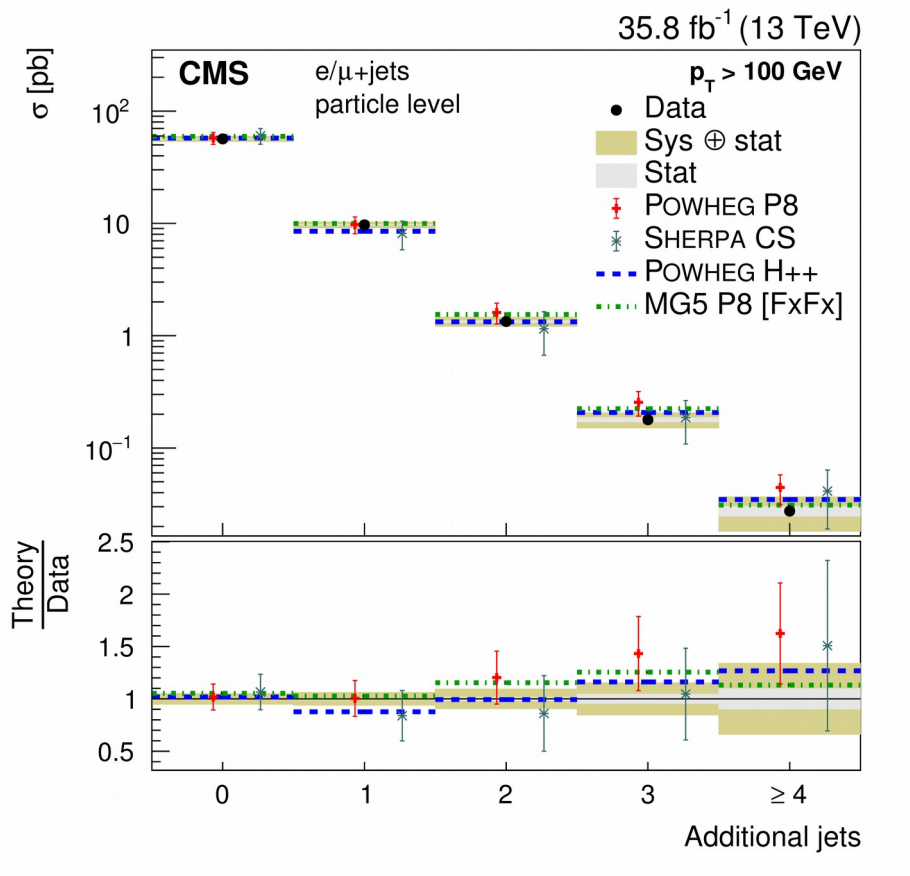
→ will profit from full Run 2 dataset, upcoming Run 3: stay tuned!

Thank you!

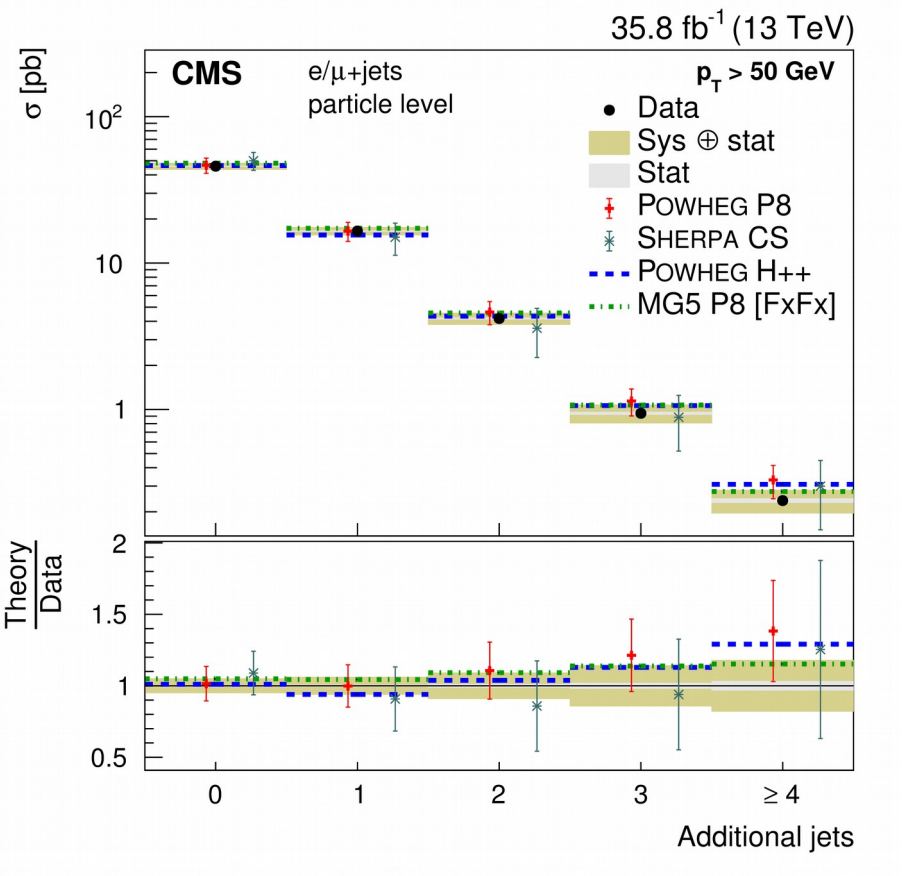
Back-up

$t\bar{t}$+jets production TOP-17-002

- Sherpa: 1st extra jet at NLO, up to 4 jets at LO + Catani-Seymour parton shower
- MG5_aMC@NLO: up to 2 jets at NLO, FxFx merging

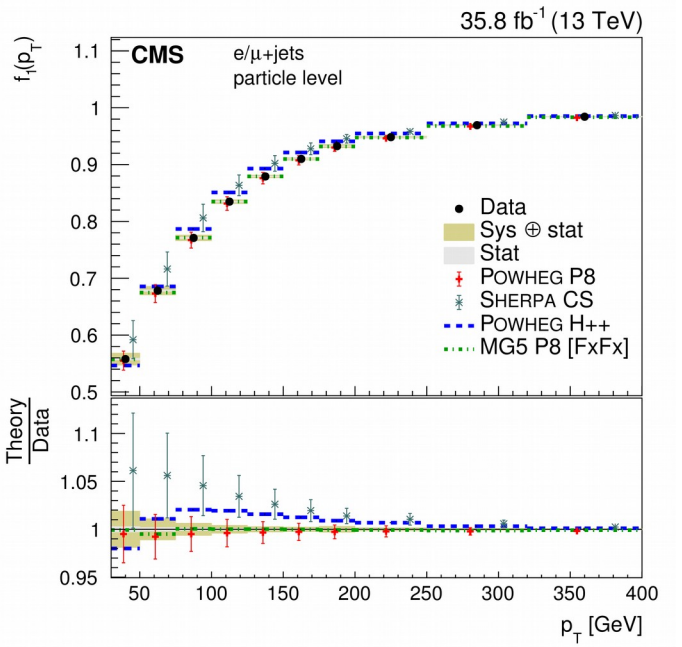


Jet multiplicity for jet p_T > 100 GeV



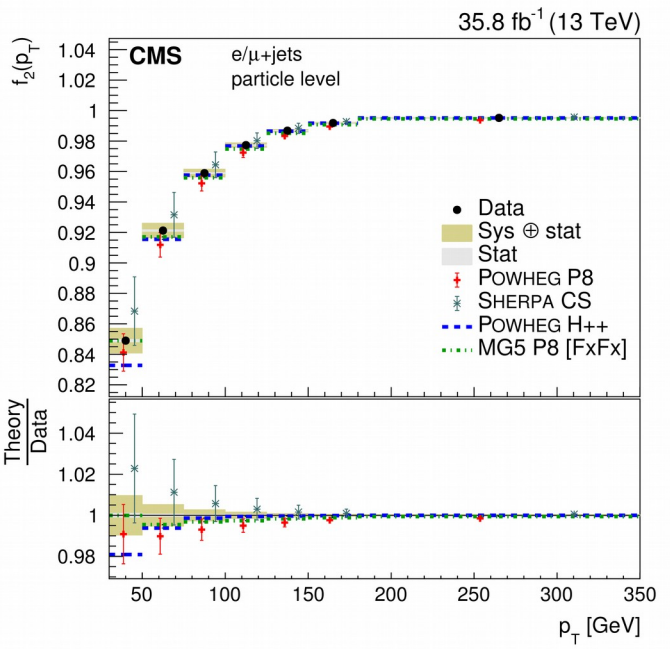
Jet multiplicity for jet p_T > 50 GeV

$t\bar{t}$+jets production TOP-17-002

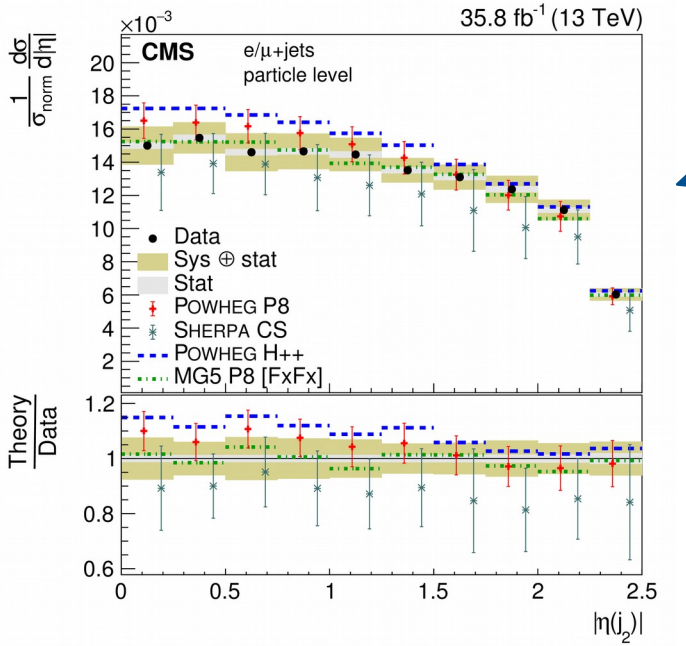


1st (left) and 2nd (right) gap fraction:

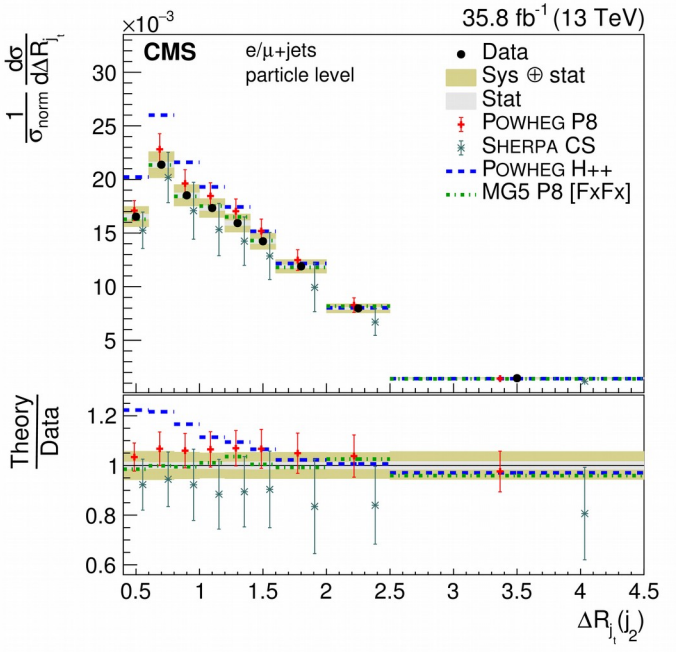
Fraction of events with less than n extra jets above pT threshold



2nd extra jet |η|



Angular separation between 2nd extra jet and closest jet from hadronic top quark decay



$t\bar{t}b\bar{b}$ cross section: all-jet final state

PAS-TOP-18-011

- Cross section measured for three phase spaces:

Visible, Parton Agnostic (PA):

- ≥ 8 jets with pT > 20 Gev, of which ≥ 6 with pT > 30 GeV
- ≥ 4 b jets

Visible, Parton Based (PB):

- ≥ 8 jets with pT > 20 Gev, of which ≥ 6 with pT > 30 GeV
- ≥ 4 b jets, of which ≥ 2 not coming from top decays

Full:

≥ 2 b jets with pT > 20 GeV, not coming from top decays

- Event selection:
  - ≥ 8 jets with pT > 30 Gev, of which ≥ 6 with pT > 40 GeV
  - ≥ 2 b-tagged jets
  - HT > 500 GeV
  - P(Chi2) < 1e-6 $\chi^2 = (M(j_1, j_3, j_4) - m_t)^2/\sigma_t^2 + (M(j_3, j_4) - m_W)^2/\sigma_W^2 + (M(j_2, j_5, j_6) - m_t)^2/\sigma_t^2 + (M(j_5, j_6) - m_W)^2/\sigma_W^2$

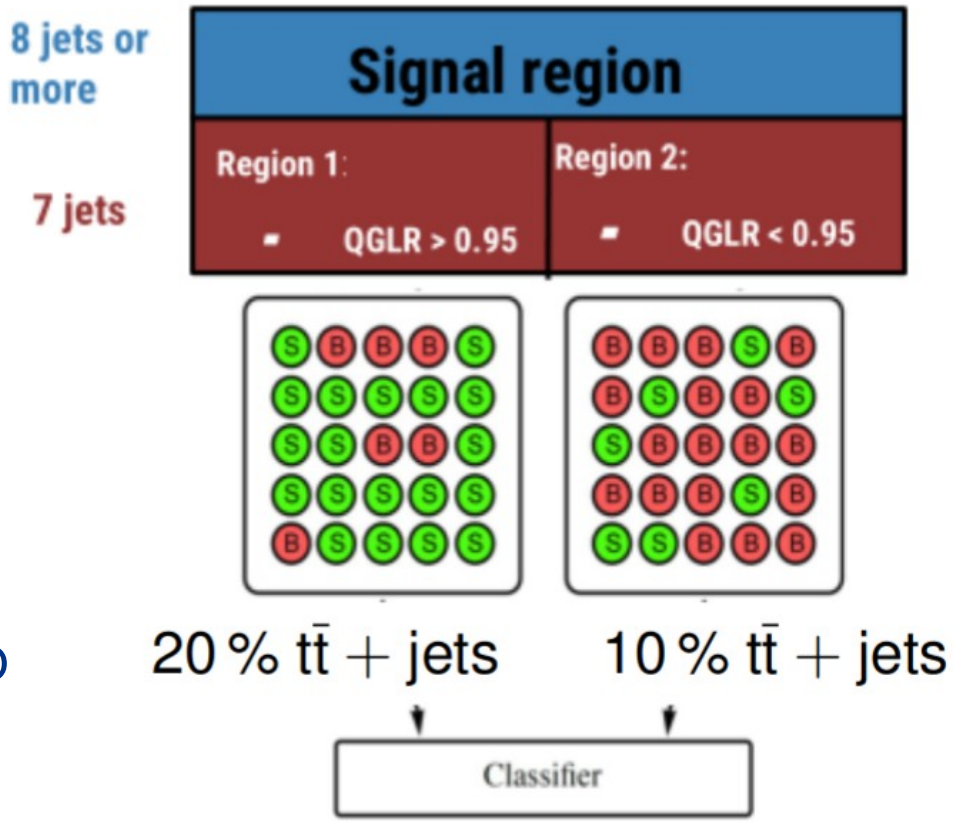
- QGLR: L(4, 0)/(L(4, 0) +L(0, 4)) ↔ on average 4 quark jets in tt events

$$L(N_q, N_g) = \sum_{\text{perm}} \left(\prod_{k=i_1}^{i_{N_q}} \prod_{k=i_{N_q+1}}^{i_{N_q+N_g}} f_q(\zeta_k) f_g(\zeta_m) \right)$$

$t\bar{t}b\bar{b}$ cross section: all-jet final state

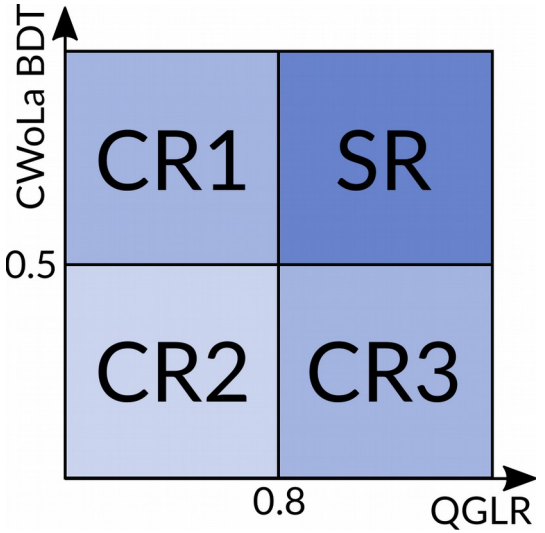
PAS-TOP-18-011

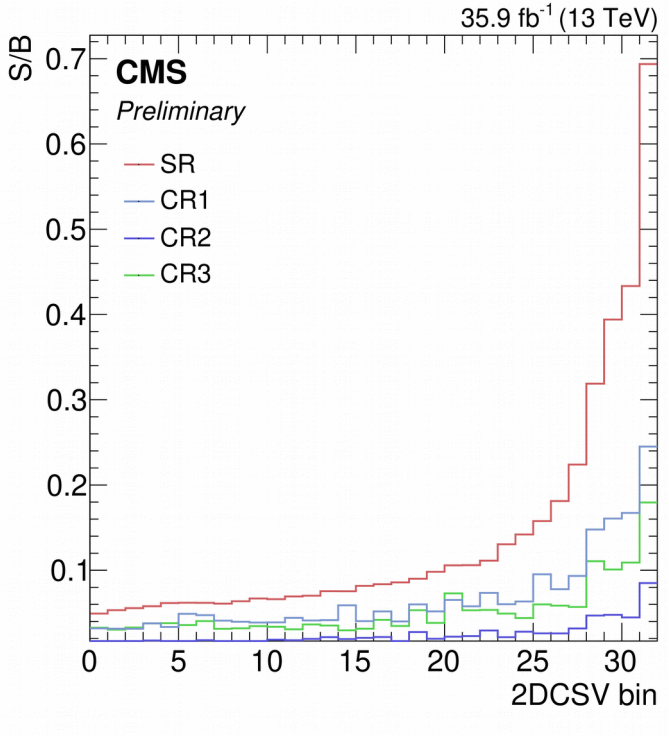
- CWoLa method: 1708.02949
  - 1 signal, 1 background
  - 2 regions with different signal purity
  - Treat 2 regions as “signal” and “background” in training
  - Classifier converges to discriminator for actual signal and background
  - Condition: region definition uncorrelated with input variables
  - Here: signal = tt+jets, background = QCD

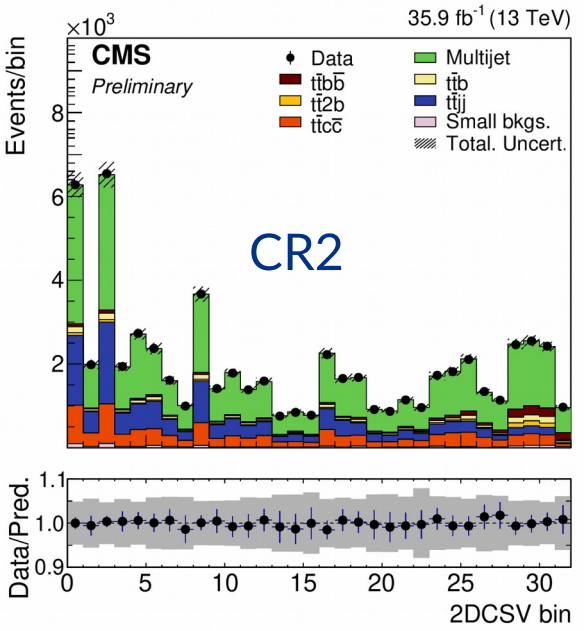


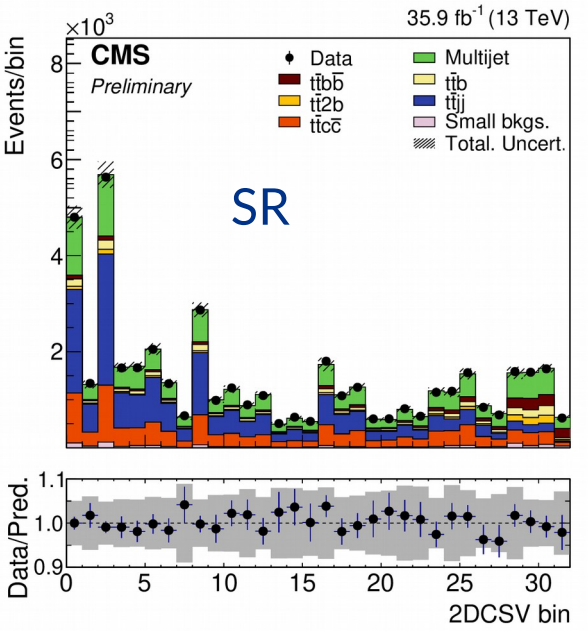
$t\bar{t}b\bar{b}$ cross section: all-jet final state

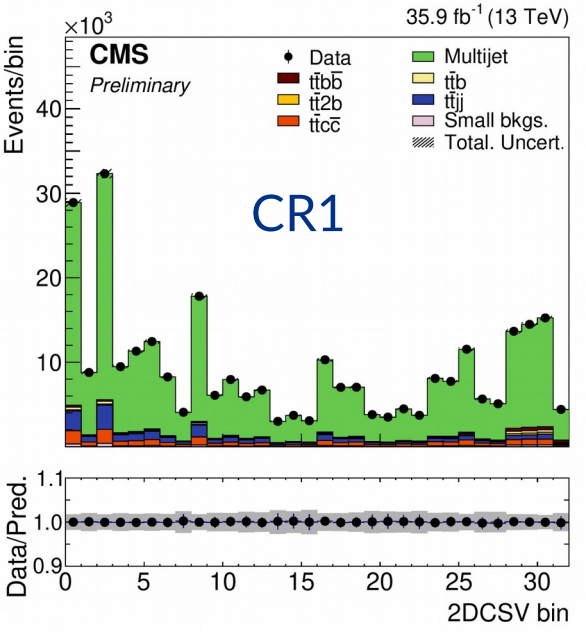
PAS-TOP-18-011

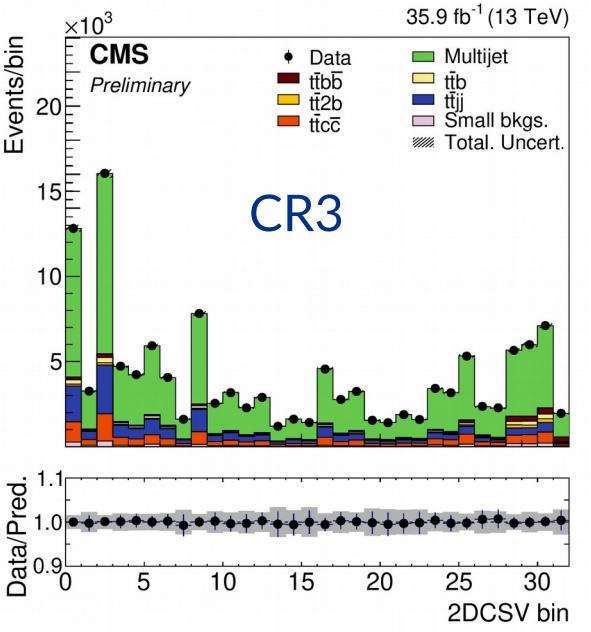






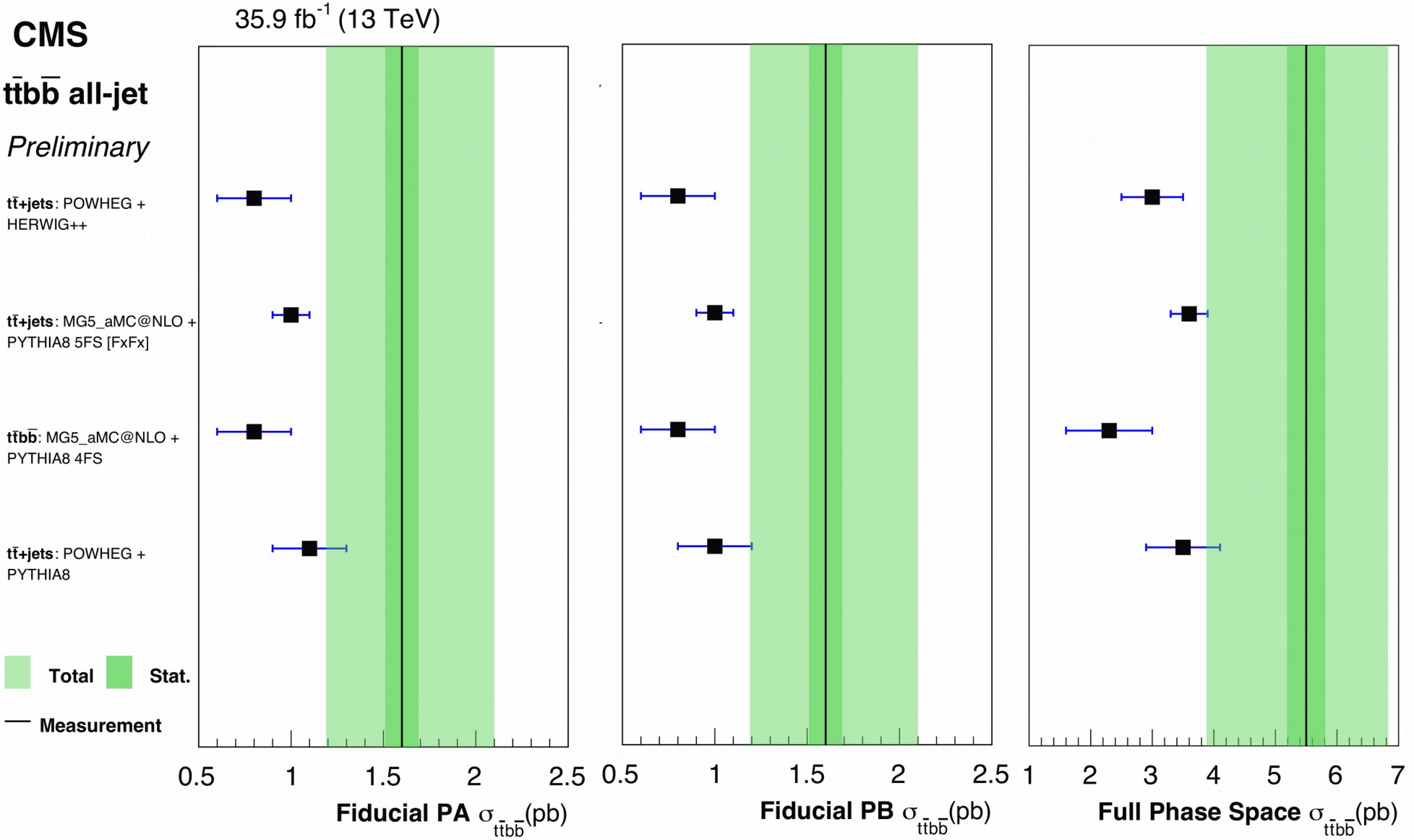






$t\bar{t}b\bar{b}$ cross section: all-jet final state

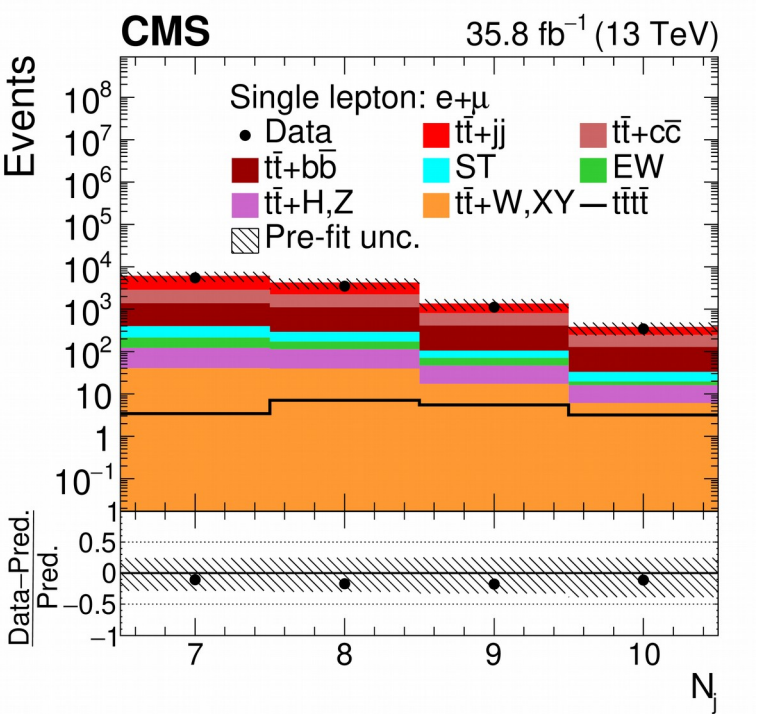
PAS-TOP-18-011

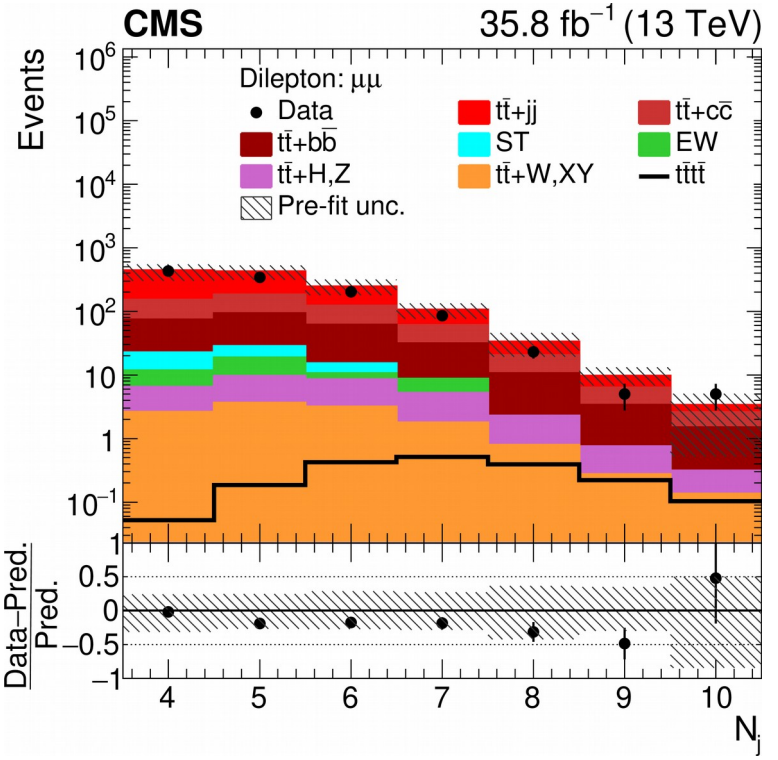


$t\bar{t}t\bar{t}$ production: 1, 2 (OS) leptons

TOP-17-019

Jet multiplicities after correction:





$t\bar{t}t\bar{t}$ production: 2 (SS), 3 leptons

PAS-TOP-18-003

Lepton identification crucial:

- Multivariate electron identification
- Advanced isolation requirements
- Reduction of charge flips

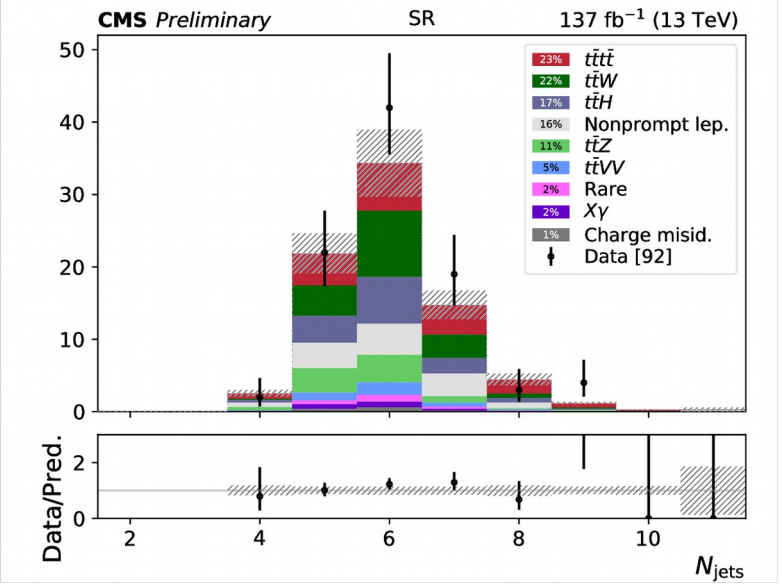
- Lepton p_T-dependent isolation cone
- Ratio lepton p_T / closest jet p_T
- Lepton p_T relative to direction of nearby hadronic activity

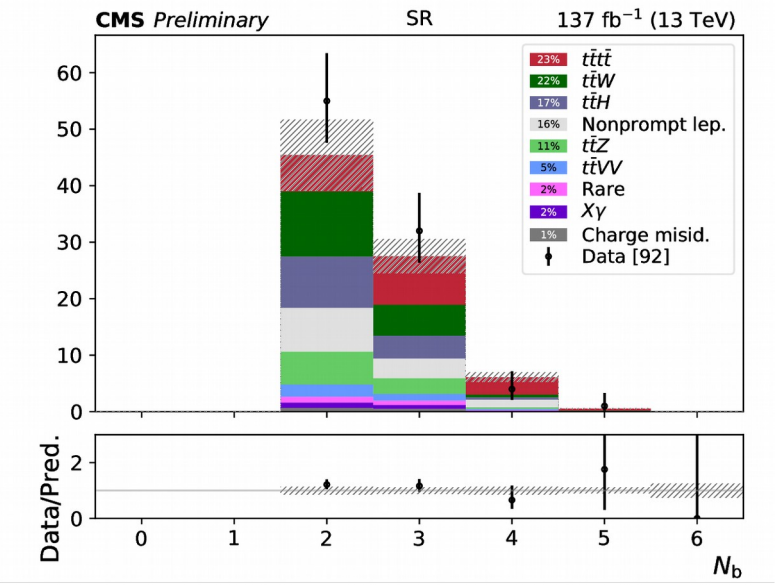
- Electrons: linear extrapolation of pixel hits to calorimeter, compare with position of energy cluster → should be consistent with track charge
  → charge mis-ID 10^{-3} – 10^{-5}
- Muons: require small p_T uncertainty → negligible charge mis-ID

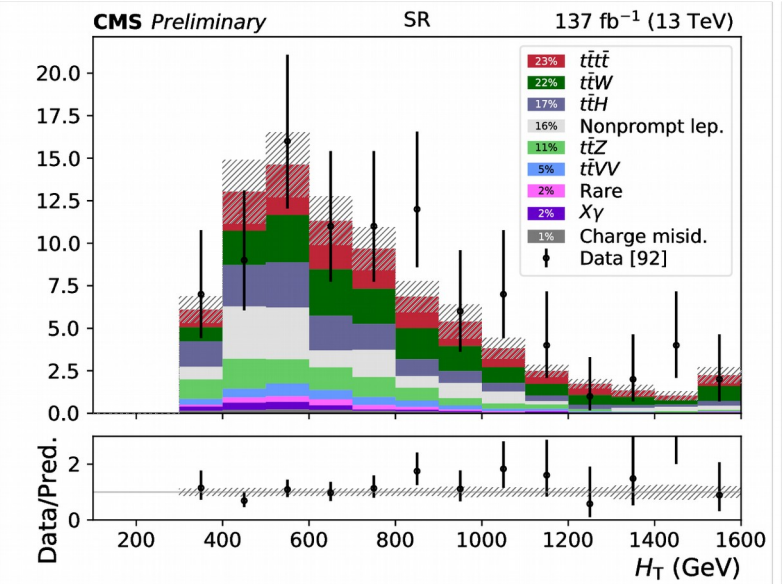
$t\bar{t}t\bar{t}$ production: 2 (SS), 3 leptons

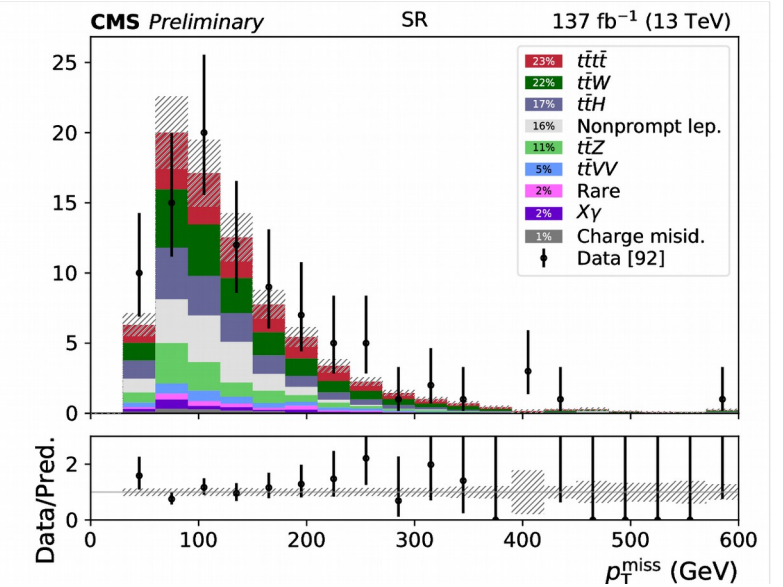
PAS-TOP-18-003

Sum of 14 Signal Regions (cut-based), pre-fit:





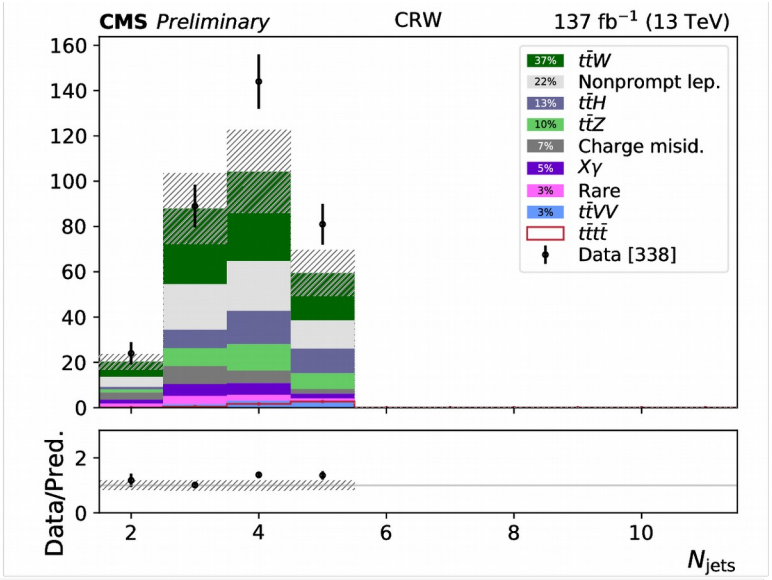


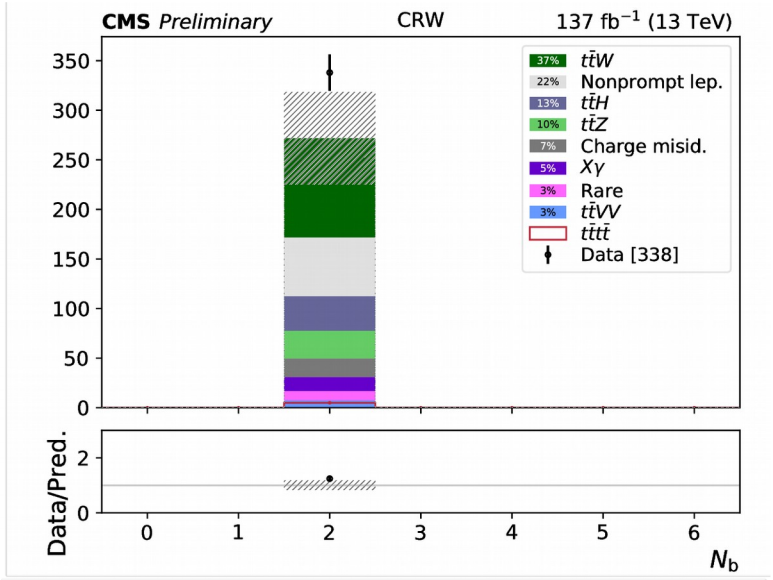


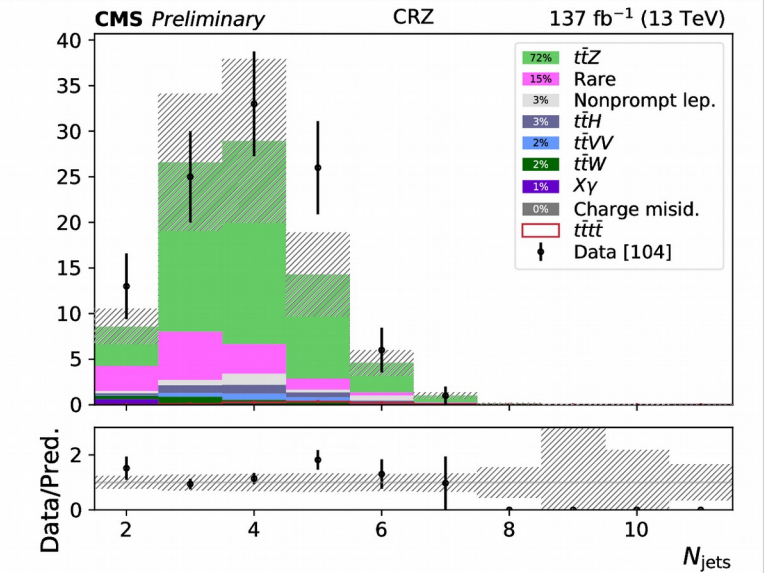
$t\bar{t}t\bar{t}$ production: 2 (SS), 3 leptons

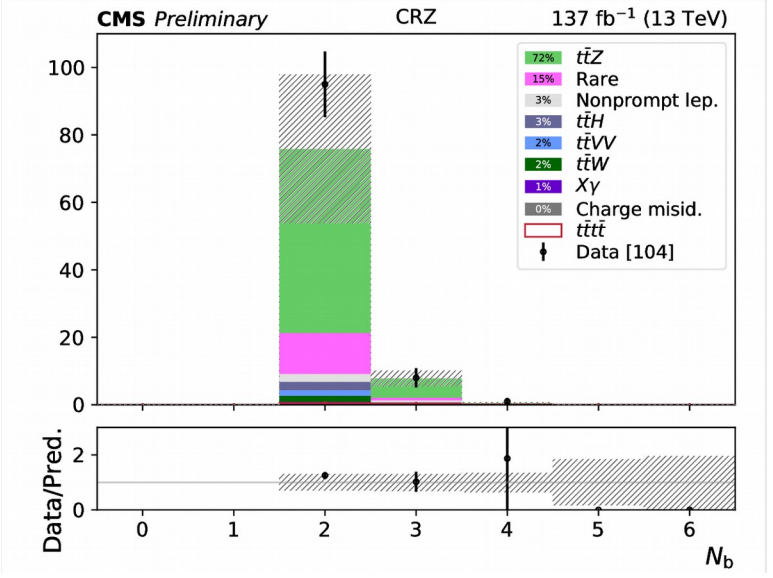
PAS-TOP-18-003

ttW, ttZ control regions (pre-fit):









Note: ttW and ttZ cross sections measured to be 23% and 17% larger than SM predictions (within uncertainties):

TOP-17-005

$t\bar{t}t\bar{t}$ production: 2 (SS), 3 leptons

PAS-TOP-18-003

Cut-based signal regions:

N_ℓ	N_b	N_{jets}	Region
2	2	≤ 5	CRW
		6	SR1
		7	SR2
		≥ 8	SR3
	3	5	SR4
		6	SR5
		7	SR6
		≥ 8	SR7
	≥ 4	≥ 5	SR8
≥ 3	2	5	SR9
		6	SR10
		≥ 7	SR11
	≥ 3	4	SR12
		5	SR13
		≥ 6	SR14
inverted Z-veto			CRZ

Systematic uncertainties:

Source	Uncertainty (%)	Impact on the $t\bar{t}t\bar{t}$ cross section (%)
Integrated luminosity	2.3–2.5	3
Pileup	0–5	1
Trigger efficiency	2–7	2
Lepton selection	2–10	2
Jet energy scale	1–15	9
Jet energy resolution	1–10	6
b tagging	1–15	6
Size of simulated sample	1–25	<1
Scale and PDF variations †	10–15	2
ISR/FSR (signal) †	5–15	2
$t\bar{t}H$ (normalization) †	25	5
Rare, $X\gamma$, $t\bar{t}VV$ (norm.) †	11-20	<1
$t\bar{t}Z$, $t\bar{t}W$ (norm.) †	40	3-4
Charge misidentification †	20	<1
Nonprompt leptons †	30–60	3
$N_{jets}^{ISR/FSR}$ †	1-30	2
$\sigma(t\bar{t}b\bar{b})/\sigma(t\bar{t}jj)$ †	35	11