Observation of single top quark production in association with a Z boson

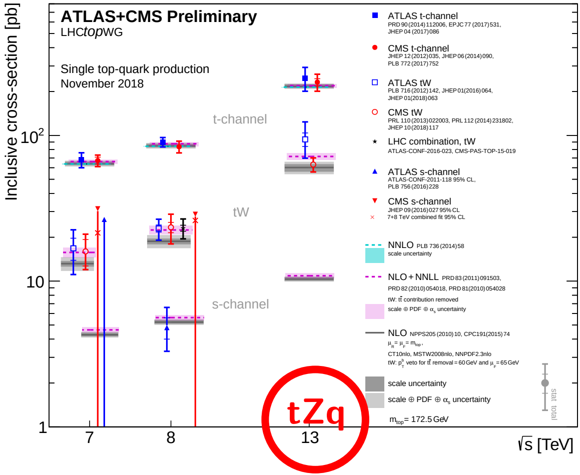
Willem Verbeke

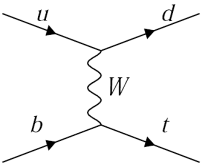
Ghent University

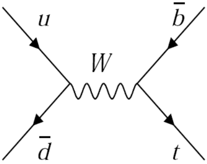
December 20, 2018

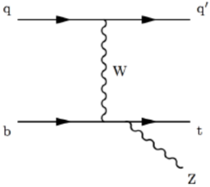
arXiv:1812.05900

Single top quark production at the LHC

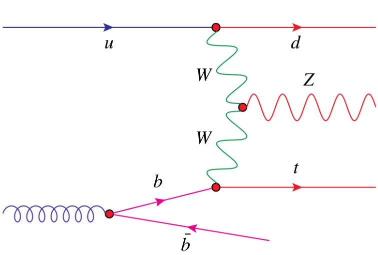


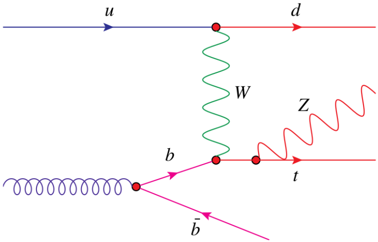






The tZq process





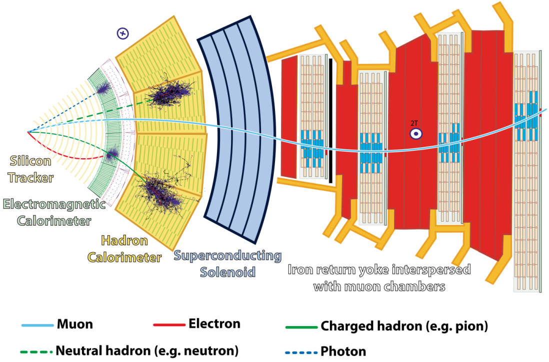
- small cross section: σ(tZq) $\approx$ 1 pb
- target **leptonic decays** of Z and top (**3% branching fraction, but experimentally cleanest**)
- final state with **3ℓ + b jet + forward jet**

- remained undiscovered with 2016 data:
  - CMS: 3.1 s.d. exp., **3.7 s.d.** obs.
  - ATLAS: 5.4 s.d. exp., **4.2 s.d.** obs.

Designed new analysis aimed at discovery

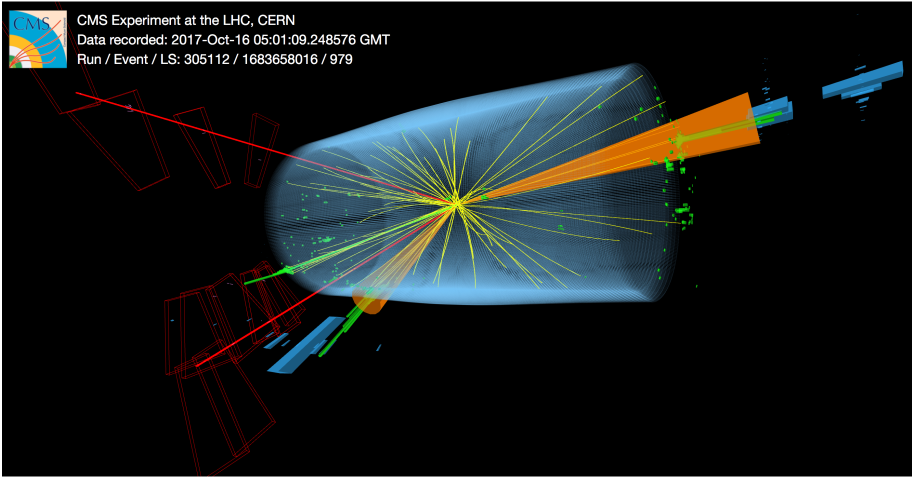
CMS particle reconstruction

- Muon: **tracker** track, small energy deposit in calorimeters, hits in outer **muon system**
- Electron: **tracker** track, **ECAL** cluster, brehmstrahlung
- charged hadron : **tracker** track, **HCAL** energy deposit
- photon : **ECAL** energy deposit
- neutral hadron: **HCAL** energy deposit



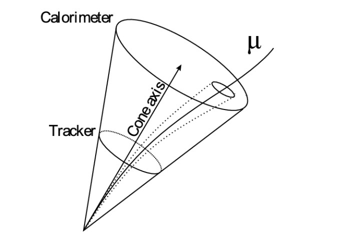
reconstruct particles using an **optimized combination of all subdetectors** (particle-flow)

tZq event display

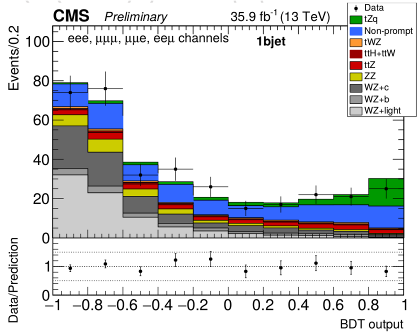


Lepton MVA

- **large nonprompt lepton background** in previous CMS search
- **tigh cut-based ID** + **relative isolation** was used
- switch to machine learning based **lepton MVA** (similar to **ttH observation**)



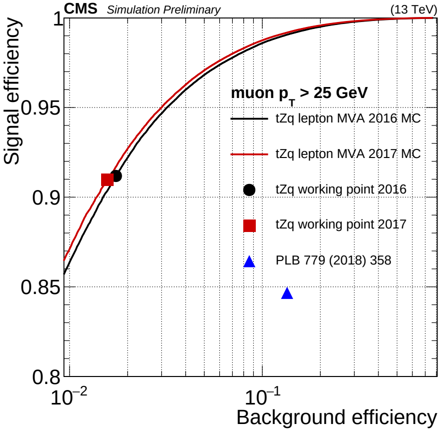
previous CMS tZq search:

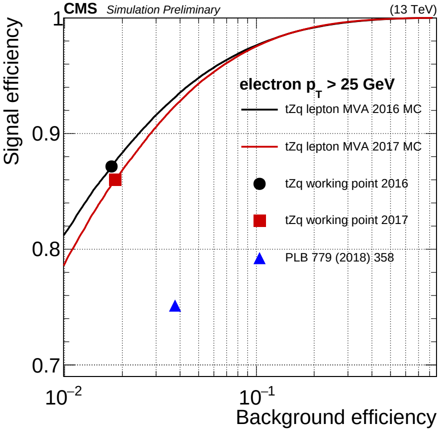


- train multivariate discriminant
- use properties of closest jet (Deep CSV, $\frac{P_T^{jet}}{P_T^{\ell}}$, P^{jet} orthogonal to lepton axis
- trained and optimized **gradient boosted forest** (BDT) in TMVA
- trained and optimized **densely connected deep neural network** in Tensorflow (with Keras)

TMVA K + TensorFlow

Lepton MVA performance





- muons: prompt **efficiency increases by 8%**, while reducing nonprompt lepton efficiency by **factor** $\sim$ **8**
- electrons: prompt **efficiency increases by 12%**, while reducing nonprompt lepton efficiency by **factor** $\sim$ **2**

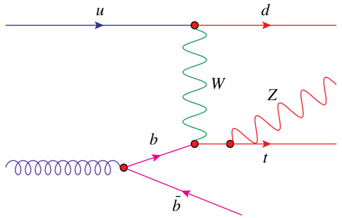
Outline of analysis strategy

categorize events according to jet content:

- 2-3 jets, 1 b-tagged: **most sensitive category**
- $\geq$ 4 jets, 1 b-tagged: good sensitivity, **not considered previously**
- 2 b jets: good sensitivity
- events with less jets have very little sensitivity to tZq

require presence of Z-candidate:

- **Z candidate present** $\rightarrow$ **most sensitivity** because of $Z \rightarrow \ell\ell$ in signal
- **no Z candidate** $\rightarrow$ very small fraction of signal $\rightarrow$ **do not consider for analysis**



discriminate signal from background:

- even in signal enriched regions **signal much smaller than backgrounds**
- exploit **tendency of tZq to have a forward jet** $\rightarrow$ jet with high $|\eta|$ value and large dijet mass
- combine with other kinematic differences into **gradient boosted forest (BDT)**

3 signal enriched event categories, BDT trained in each one

Backgrounds

$\mathbf{WZ} \rightarrow 3\ell\nu$, $\mathbf{ZZ} \rightarrow 4\ell$

- signal-like leptons
- tends to have few jets

$\mathrm{t\bar{t}Z}$

- signal-like leptons
- **several jets and b jets**

Nonprompt e/μ

- Drell-Yan $\rightarrow 2\ell$ and $\mathrm{t\bar{t}} \rightarrow 2\ell$
- third **lepton from jet-fragmentation**

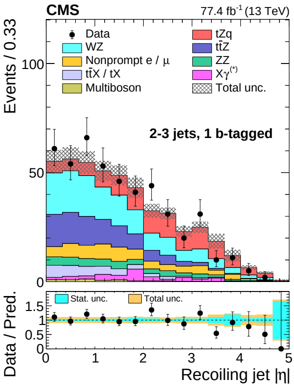
Conversions

- dominated by $\mathrm{Z}\gamma$, $\mathrm{t\bar{t}} + \gamma$ with $\gamma^{(*)} \rightarrow 2\ell$

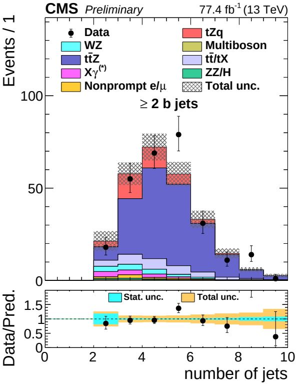
rare processes

- $\mathrm{t\bar{t}} + \mathrm{H}$, triboson, tWZ, ...
- **very small**

2/3 jets, 1 b-tagged:



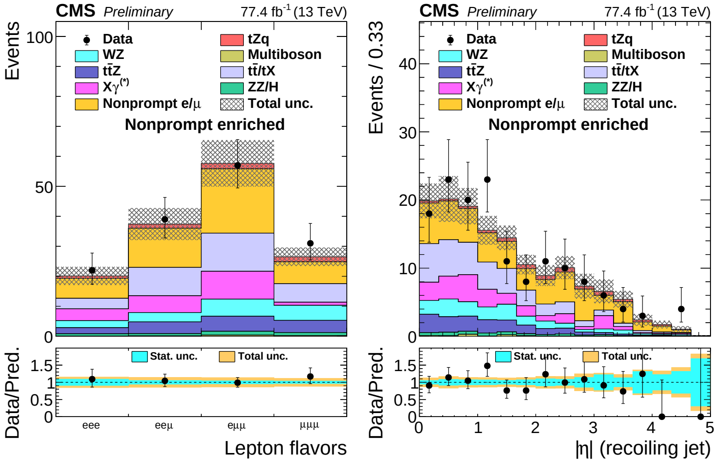
$\geq$ 2 b jets:



- **WZ** is dominant background in **2/3 jets, 1 b-tagged** category
- $\mathbf{t\bar{t}Z}$ is dominant background in $\geq$ **4 jets, 1 b-tagged and** $\geq$ **2 b jets** categories
- nonprompt background largely killed by lepton MVA

Data-driven prediction of nonprompt lepton background

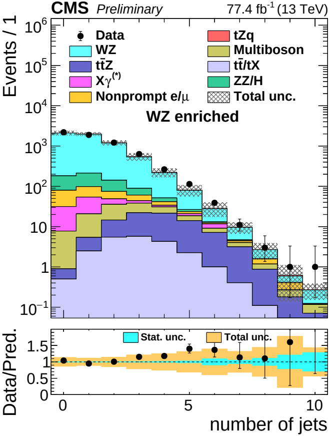
- measure **probability that nonprompt lepton passing a loose lepton selection also passes full selection**
- apply this probability to events where one or more leptons fail full selection

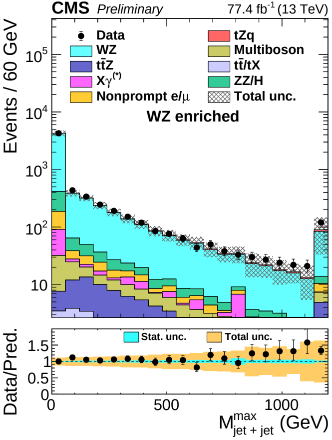


Control regions (simulated backgrounds)

predictions of search variables checked in **3 control regions**:

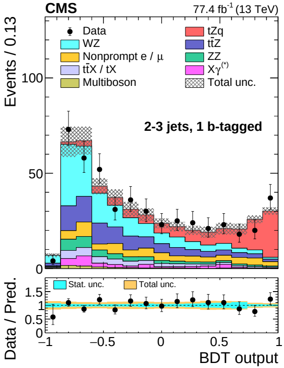
- WZ : 3ℓ, OSSF pair, Z candidate present, 0 b jets (orthogonal to SR), $E_T^{miss} > 50$ GeV
- ZZ : 4ℓ, 2 OSSF pairs, both forming Z candidate
- $X\gamma$: 3ℓ, OSSF pair, no dilepton Z candidate, trilepton mass compatible with Z mass



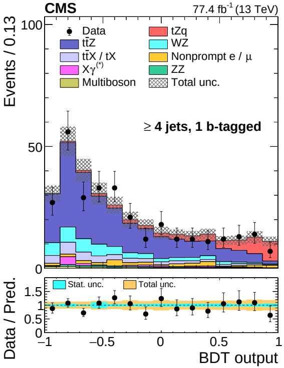


Results

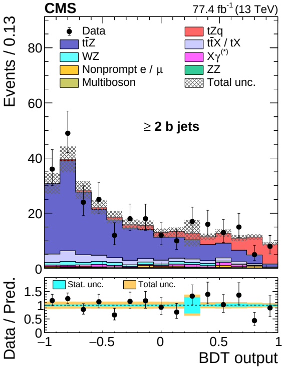
2/3 jets, 1 b-tagged:



≥ 4 jets, 1 b-tagged:



≥ 2 b jets:



good agreement between data and expectation in three categories

large excess of events over background-only hypothesis

Signal significance

2016 Data :

- Observed (expected) significance of **7.2 (5.6) s.d.**

$$\mu = 1.36\ ^{+0.22}_{-0.20}\,(\text{stat})\ ^{+0.14}_{-0.12}\,(\text{syst})$$

2017 Data :

- Observed (expected) significance of **5.4 (6.0) s.d.**

$$\mu = 1.03\ ^{+0.18}_{-0.17}\,(\text{stat})\ ^{+0.14}_{-0.12}\,(\text{syst})$$

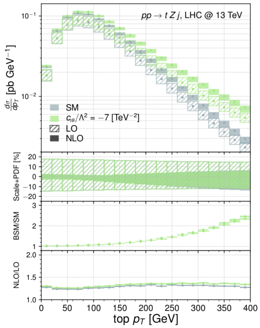
Observation of tZq in both 2016 and 2017 datasets!
total observed (expected) significance of **8.2 (7.7) s.d.**

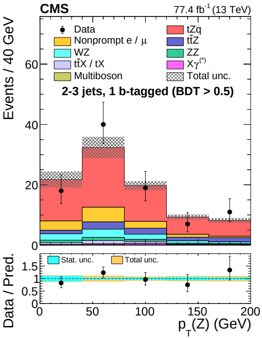
$$\frac{\sigma(\text{tZq} \to \text{t}\ell\ell\text{q})}{\sigma^{SM}(\text{tZq} \to \text{t}\ell\ell\text{q})} = \mu = 1.18\ ^{+0.14}_{-0.13}\,(\text{stat})\ ^{+0.11}_{-0.10}\,(\text{syst})\ ^{+0.04}_{-0.04}\,(\text{theo})$$

$$\sigma(\text{tZq} \to \text{t}\ell\ell\text{q}) = 111 \pm 13\ (\text{stat})^{+11}_{-9}\ (\text{syst})\ \text{fb}$$

(all significances computed in asymptotic approximation of test statistic)

Future plans





arXiv:1804.07773

tZq kinematics are **uniquely sensitive to new physics:**

- tZq is sensitive to a **large number of SM interactions** (WWZ coupling, tbW and ttZ vertices, bW → tZ amplitude)
- several cancellations in SM
- modified interactions can lead to **anomalous energy growth**
- would be seen in for instance P_T^{top} and P_T^Z

plans for full Run II result (nearly double the data volume):

- **differential measurement** in variables most sensitive to new physics
- **EFT** interpretation

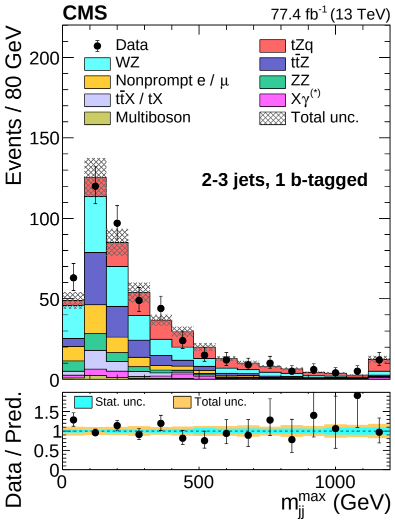
Conclusions

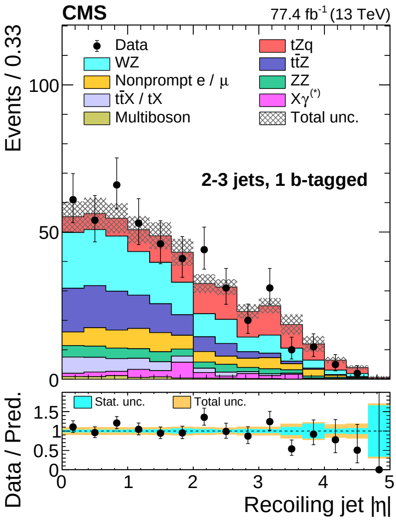
- new search for tZq was designed and carried out on the 2016 and 2017 data sets
- nonprompt lepton background, which limited earlier CMS searches, severely reduced by usage of machine learning for lepton identification
- new analysis strategy designed from the bottom up

tZq is observed with a **significance of 8.2 s.d.** (7.7 s.d. expected)

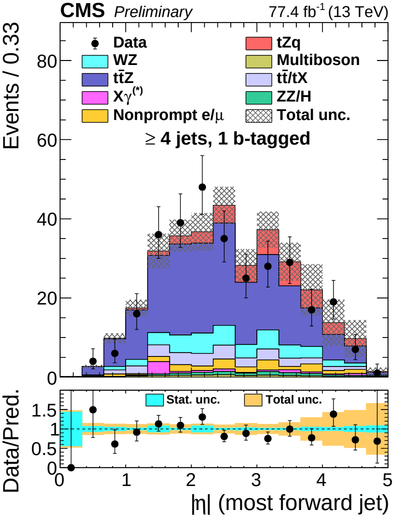
$$\sigma(\mathrm{tZq} \to \mathrm{t}\ell^+\ell^-\mathrm{q}) = 111^{+13}_{-13}\ (\mathrm{stat})^{+11}_{-9}\ (\mathrm{syst})\ \mathrm{fb}$$

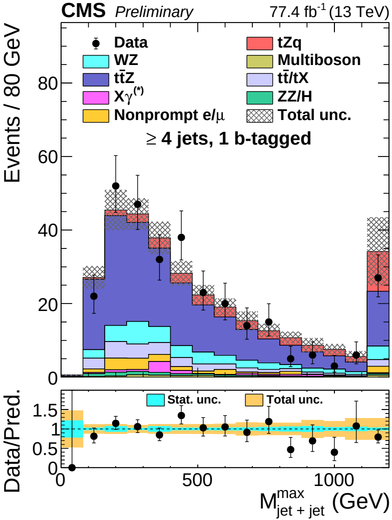
Backup: Discriminating variables (2-3 jets, 1 b-tagged)



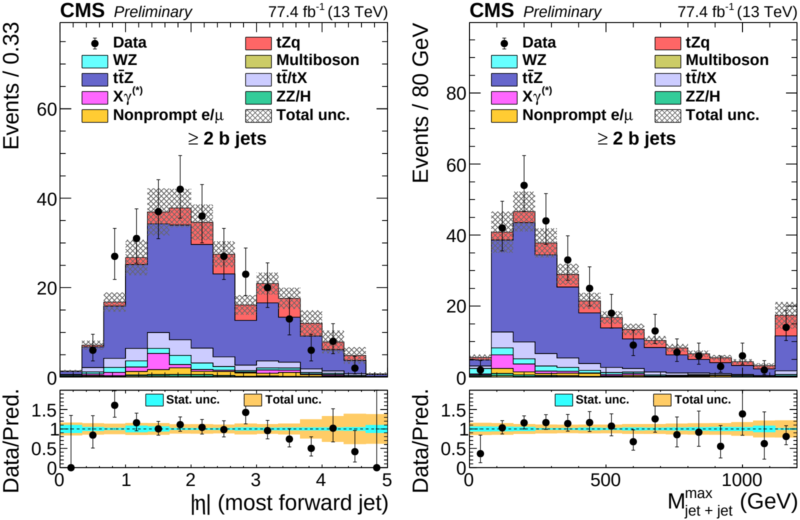


Backup: Discriminating variables (4 jets, 1 b-tagged)



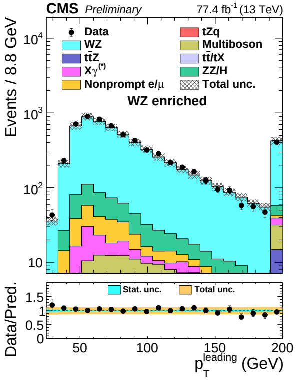


Backup: Discriminating variables (2 b jets)

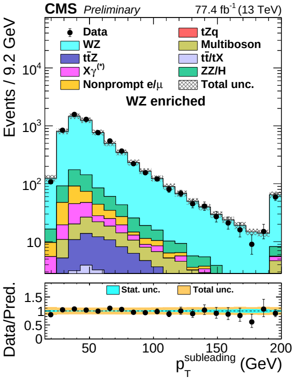


Backup: Lepton p_T spectra (WZ enriched region)

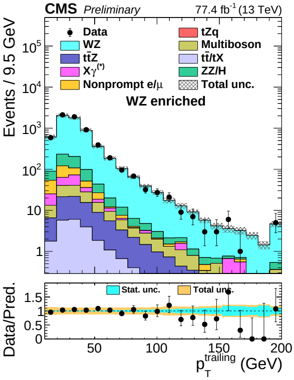
highest p_T:



second highest p_T:

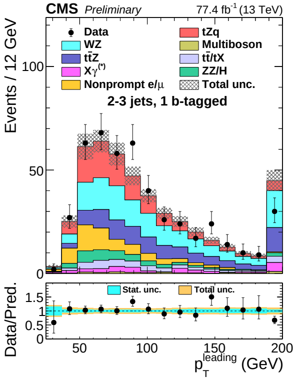


lowest p_T:

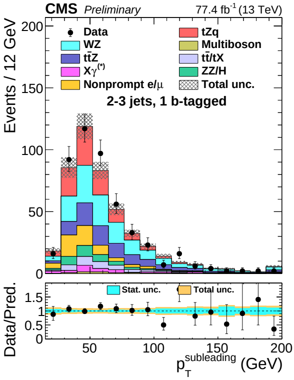


Backup: Lepton p_T spectra (2-3 jets 1 b-tagged)

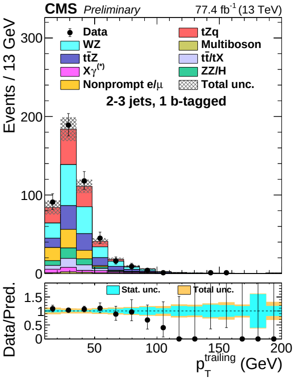
highest p_T:



second highest p_T:

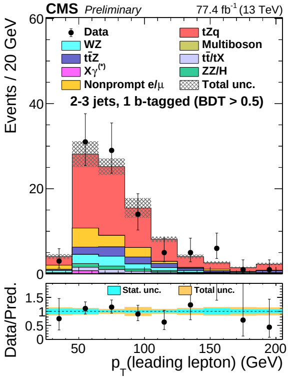


lowest p_T:

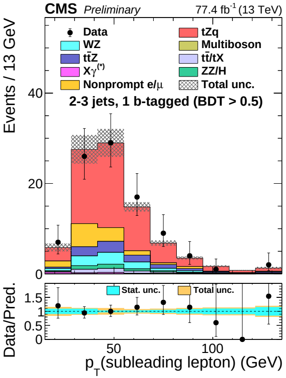


Backup: Lepton p_T spectra (high purity region)

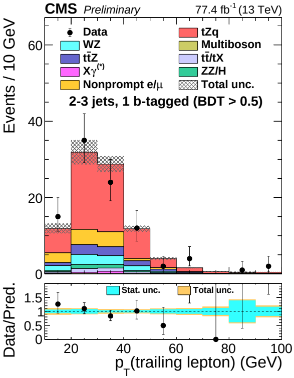
highest p_T:



second highest p_T:

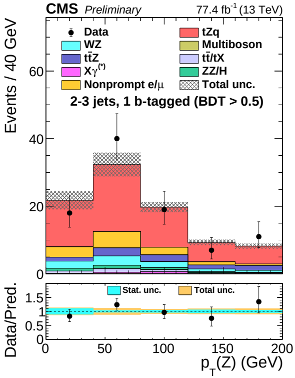


lowest p_T:

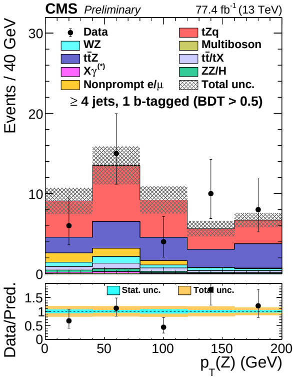


Backup: Z boson p_T (High purity region)

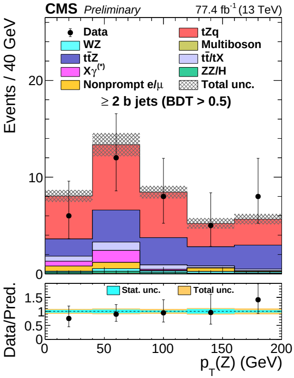
highest p_T:



second highest p_T:

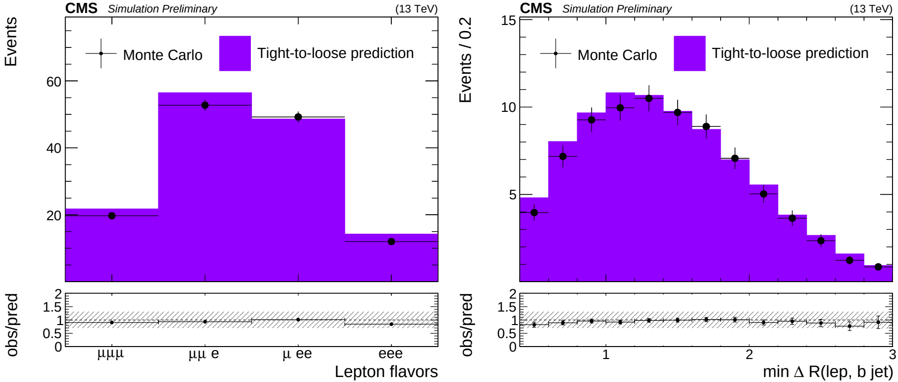


lowest p_T:



Backup:Nonprompt closure in MC

- measure "fake-rate" in QCD enriched region in data
- nonprompt background comes from Drell-Yan and $t\bar{t}$
- verify that fake-rate measured in QCD MC can predict Drell-Yan and $t\bar{t}$ backgrounds
- closure tests in $t\bar{t}$ MC (fake-rate prediction VS direct MC prediction):



Largest systematic uncertainties

Uncertainty	Impact (%)
Experimental	
lepton selection	3.2
trigger efficiency 2016 (2017)	1.0 (1.1)
jet energy scale 2016 (2017)	0.9 (3.1)
b-tagging efficiency 2016 (2017)	0.7 (1.2)
nonprompt normalization	4.1
$t\bar{t}Z$ normalization	1.0
luminosity 2016 (2017)	1.2 (1.3)
pileup	1.9
other	1.3
Theoretical	
final-state radiation	2.0
tZq QCD scale	2.0
$t\bar{t}Z$ QCD scale	1.4