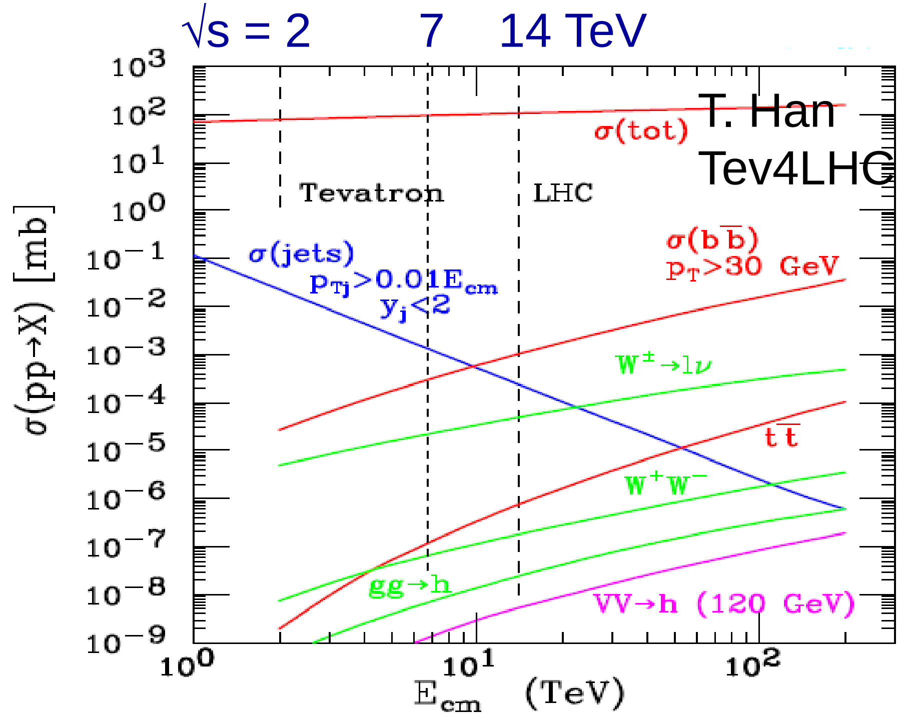
Top Quark Studies with CMS

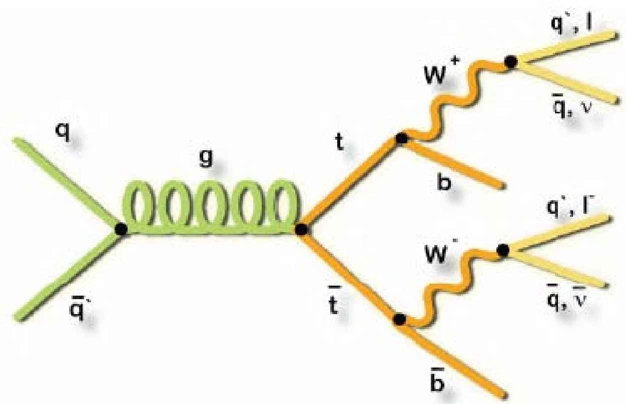
Frank-Peter Schilling (KIT)

on behalf of the CMS collaboration

Top Physics at the LHC

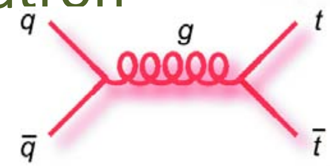
- Precise SM measurements
  - o Heaviest known elementary particle (large Yukawa coupling)
  - o Constraints on Higgs mass
  - o Unique window on bare quarks due to short lifetime
- A window to new physics
  - o New physics might couple preferentially to top
  - o New particles may decay to top
  - o Non-standard couplings
- In many new physics scenarios (e.g. SUSY) top is dominant BG
- Great tool to calibrate detector
  - o Jet energy scale, b-jet eff.

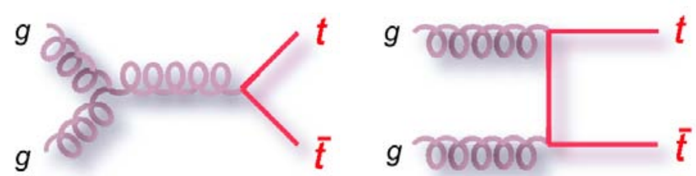


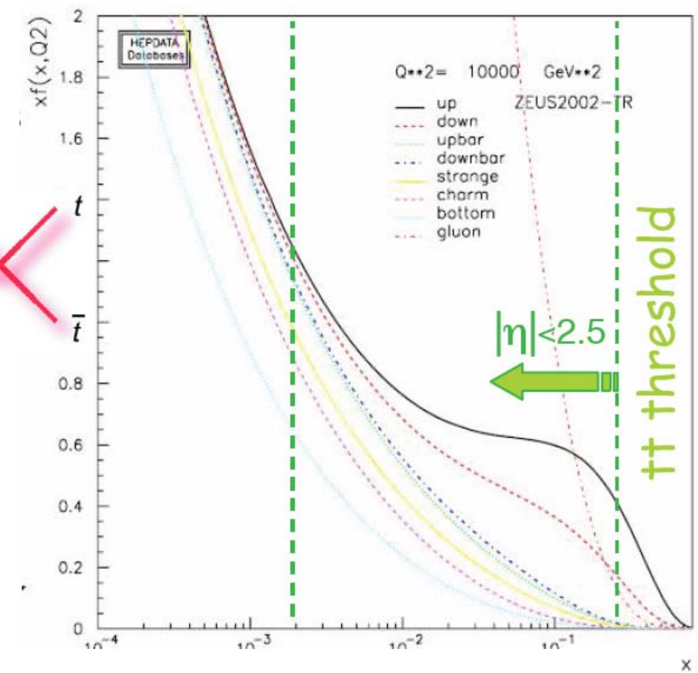


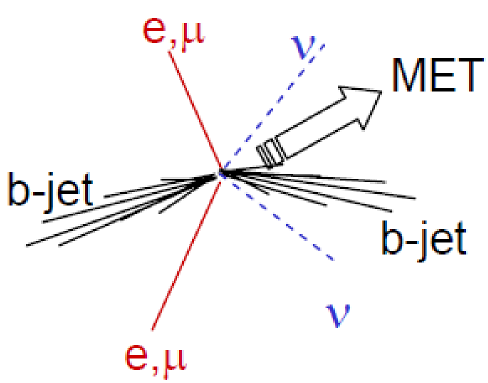
Top pair production and decay

- Pair production
  - o ~85% quark-induced at Tevatron
  - o ~85% gluon-induced at LHC

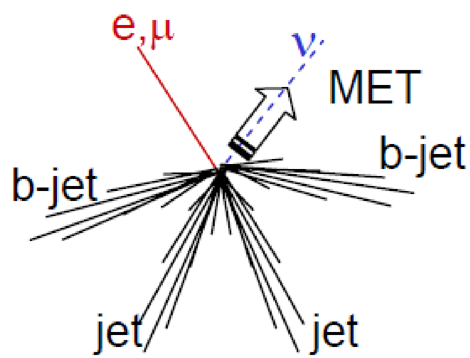




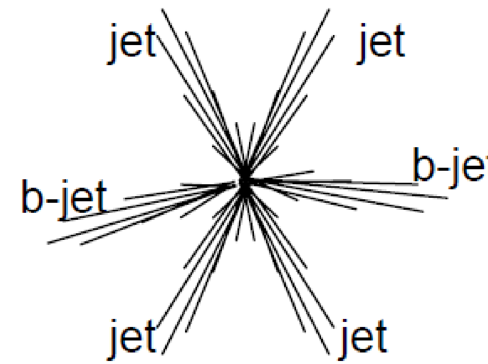




dileptons
(BR~5%)



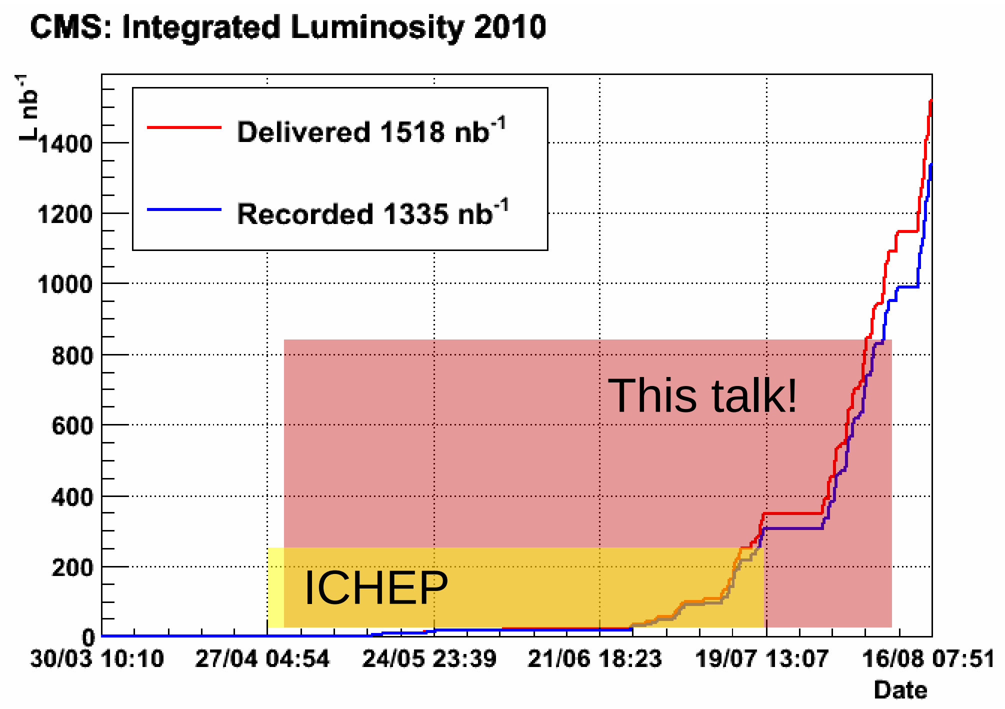
lepton + jets
(BR~30%)



all hadronic
(BR~46%)

Data and MC samples

- Dataset
  - L = 840nb-1
  - Data taken up to 11th August, certified on 13th August
  - Update of results of PAS TOP-10-004 and shown at ICHEP (up to 250nb-1)



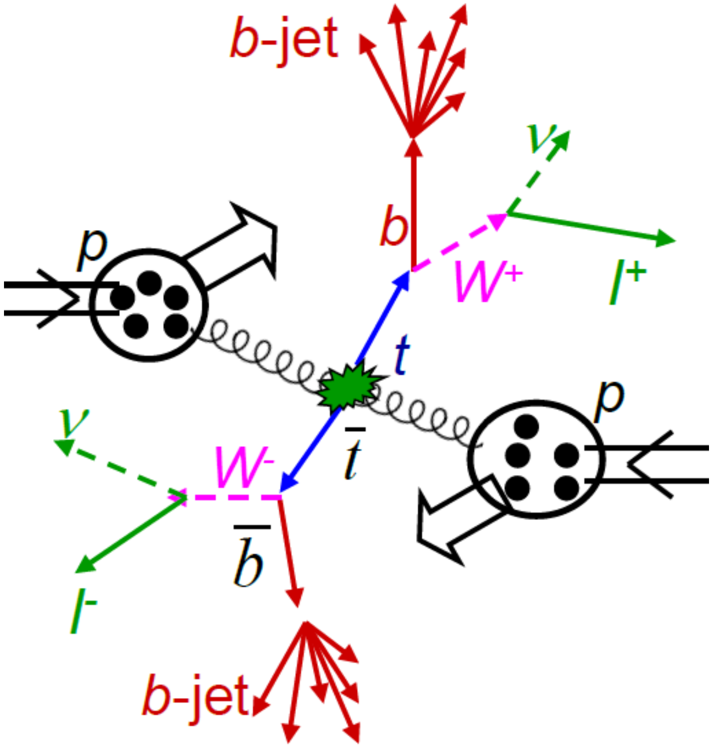
- Monte Carlo samples
  - ttbar+jets, W/Z+jets: Madgraph, matching ME with parton showers
    - V+bb/c(c)+jets matrix elements included
  - Cross sections normalized to inclusive (N)NLO cross sections
    - sigma(ttbar,NLO)=157pb (MCFM), mtop=172.5 GeV
    - sigma(W->lnu,NNLO)=31314pb (FEWZ)
  - QCD: PYTHIA (filtered at gen level)

Dilepton: Event selection

- Single lepton triggers
  - mu+X (Pt>9 GeV),e+X (Pt>15 GeV)
- Two isolated, opposite charge leptons (ee,mumu,emu)
  - Pt>20 GeV, |eta|<2.5(mu),2.4(e)
  - Rel. isolation < 0.15

$$\text{Rel.isol.} = \frac{\sum_{R<0.3} p_T^{\text{track}} + \sum_{R<0.3} p_T^{\text{ECAL}} + \sum_{R<0.3} p_T^{\text{HCAL}}}{p_T(\text{lepton})}$$

- Z-boson veto (ee,mumu)
  - |M(ll)-M(Z)|>15 GeV
- Missing Et (MET)
  - Using calorimeter & tracking
  - MET>30(20) GeV in ee,mumu (emu)



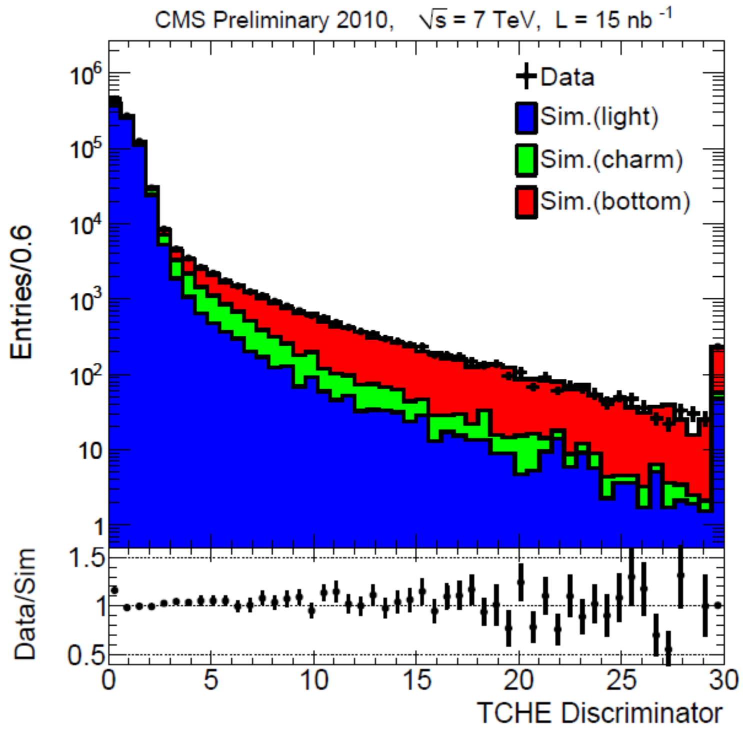
- Jets
  - Anti-Kt (R=0.5)
  - Using calorimeter & tracking
  - Pt>30 GeV, |eta|<2.4
  - Expect >=2 jets for ttbar

A word on b-tagging

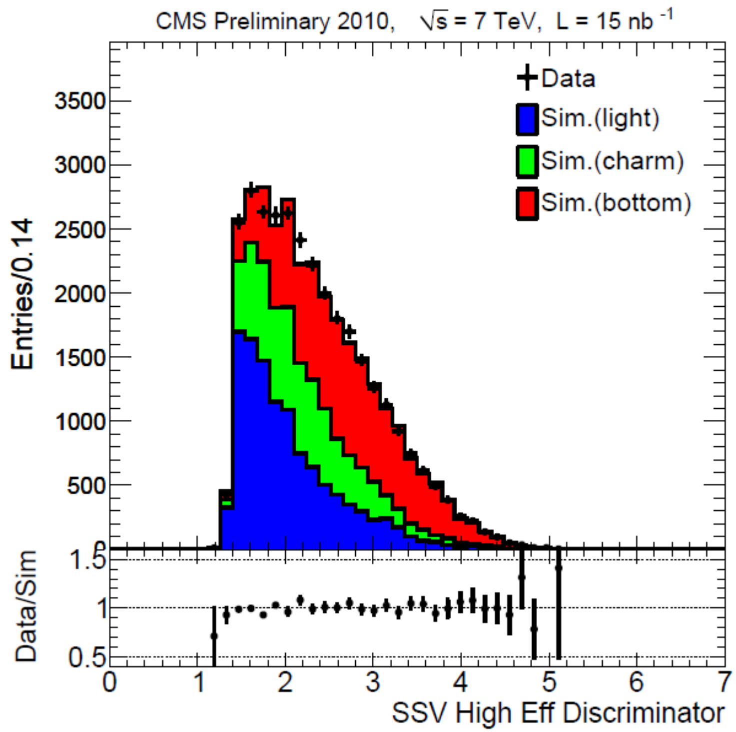
b-tagging: cruical ingredient for top physics
Use simple and robust tagging algorithms for early data

CMS PAS BTV-10-001

- "Track counting" tagger
  - Uses IP significance of n-th track as discriminator



- Secondary vertex tagger
  - Uses discriminator based on 3D flight distance



Scale factors Data/MC for efficiency / mistag rate close to one!

Dilepton Event Yields

- Relaxed selection:
  - No Z-veto, no MET, N(jets) requirements

L=0.84pb^{-1}

Process	ee	$\mu\mu$	$e\mu$
Dilepton $t\bar{t}$	0.84 ±0.13±0.16	0.94 ±0.14±0.17	1.75 ±0.26±0.33
VV	0.23 ± 0.12	0.25 ± 0.13	0.35 ± 0.18
Single top - tW	0.06 ± 0.03	0.07 ± 0.03	0.13 ± 0.07
Drell-Yan $\tau\tau$	0.6 ± 0.3	0.7 ± 0.4	1.3 ± 0.7
Drell-Yan $ee, \mu\mu$	298 ± 74	343 ± 86	0.1 ± 0.1
Non-dilepton $t\bar{t}$	0.02 ± 0.01	0.004 ± 0.002	0.03 ± 0.02
W+jets	0.3 ± 0.1	0.01 ± 0.01	0.3 ± 0.2
QCD multijets	$0\ ^{+10}_{-0}$	$0.00\ ^{+10}_{-0}$	$0\ ^{+10}_{-0}$
Total simulated	300 ± 74	345 ± 86	4.0 ±0.8
Data	305	294	6

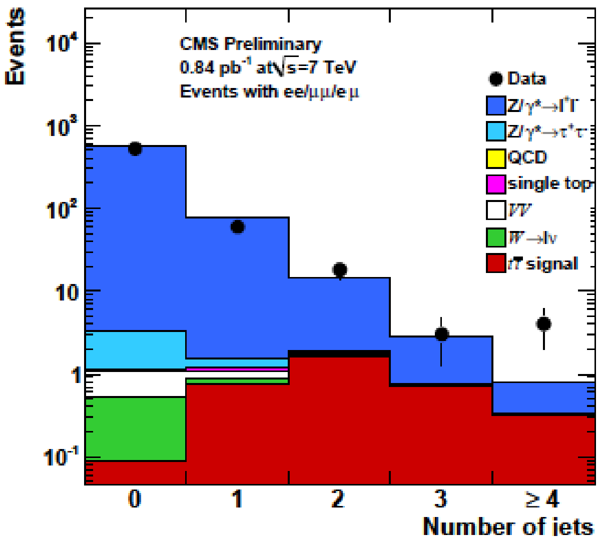
Systematics:
- Signal and DY: 15% acc*eff (conservative), 15% theory, 11% lumi
- Other backgrounds: 50% (conservative)

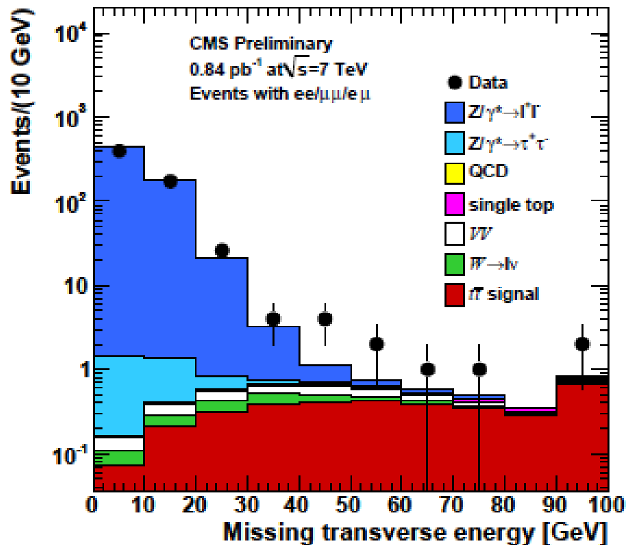
Good agreement observed!

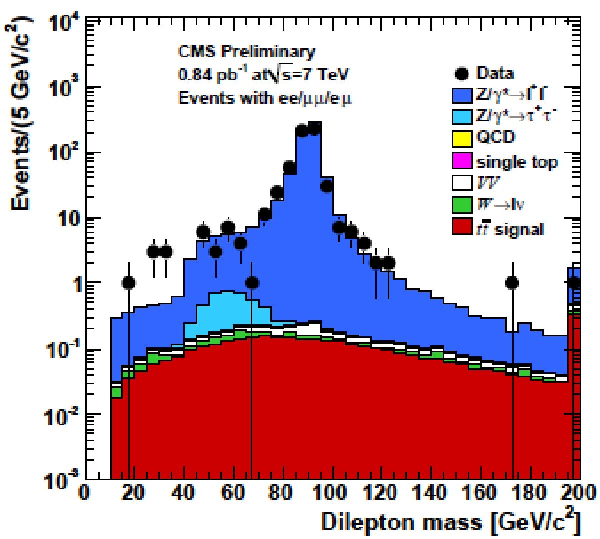
Data vs MC (relaxed selection)

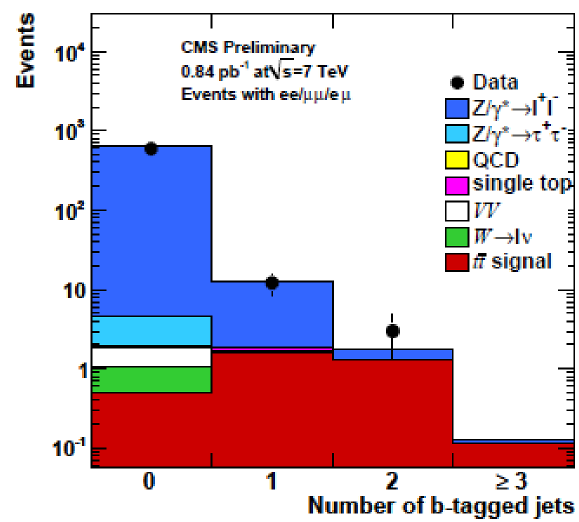
- No Z-veto, no MET, N(jets) requirements

L=0.84pb^{-1}









Rightmost bins contain overflow

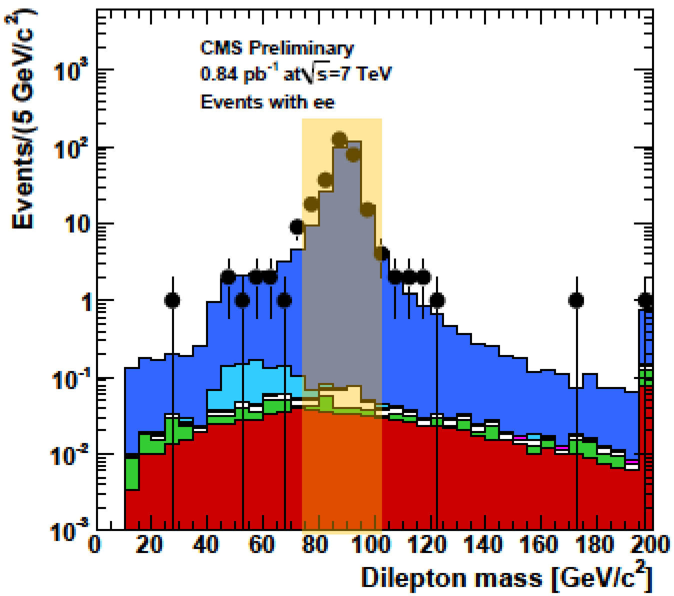
Data driven Drell-Yan Background

Estimate Drell-Yan background outside Z-veto region from events inside:

ratio outside/inside from *DY* simulation

$$N_{\text{out}}^{ee,\text{data}} = \overbrace{R_{\text{out/in}}^{ee}} \left(N_{\text{in}}^{ee,\text{data}} \underbrace{- 0.5 N_{\text{in}}^{e\mu,\text{data}} k_{ee}} \right)$$

correction for non-DY contribution in *Z*-veto region from eμ sample



Sample	ID, ISO, Z-veto	with $N_{\text{jet}} \geq 1$	with $N_{\text{jet}} \geq 2$ and $\not{E}_T > 30$ GeV
ee			
DY in simulation	26 ± 6	4.2 ± 1.1	0.04 ± 0.01
DY estimate in data	$26 \pm 1.6 \pm 13$	$3.6 \pm 0.6 \pm 1.8$	$0.4 \pm 0.2 \pm 0.2$
$\mu\mu$			
DY in simulation	31 ± 8	5.0 ± 1.2	0.07 ± 0.02
DY estimate in data	$27 \pm 1.6 \pm 13$	$4.3 \pm 0.7 \pm 2.1$	$0.21\,^{+0.23}_{-0.21} \pm 0.11$

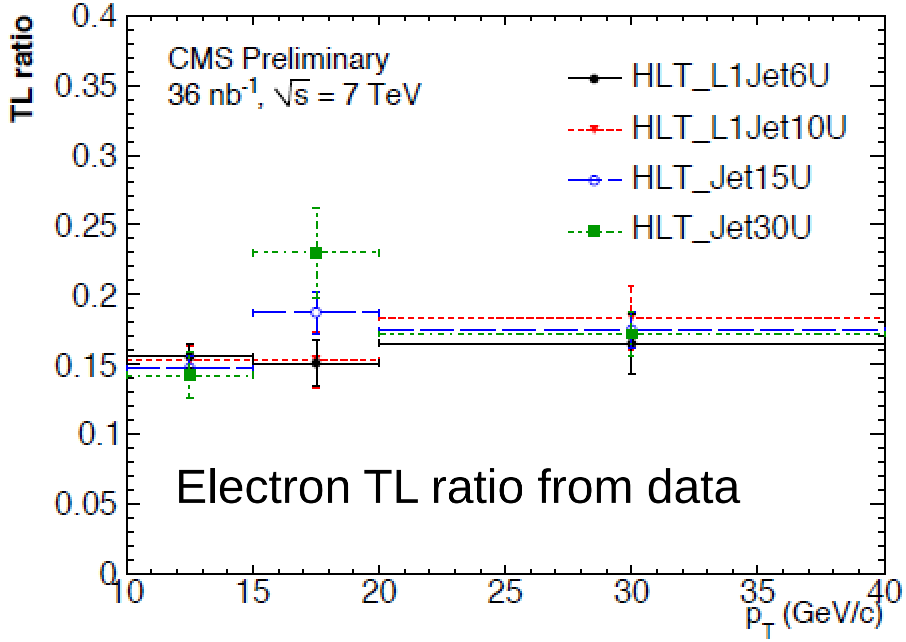
Agreement between simulation and data-driven estimate

Data-driven QCD, W+jets backgrounds

- "Fake" lepton backgrounds:
  - W+jets : one fake lepton
  - QCD: two fake leptons
- Determine a 'tight-to-lose ratio' (TL) in jet-triggered sample
- Apply to events where one (both) leptons pass loose, but fail tight lepton selection
- Weighed sum yields background estimate
- 50% systematics per "fake" lepton

$$N_{nn}^{QCD} = \sum_{i,j} \frac{TL_i TL_j}{(1-TL_i)(1-TL_j)} N_{nn}^{ij}$$

$$N_{nn}^{WJets} = \sum_{i,j} \frac{TL_i}{(1-TL_i)} N_{nn}^{ij}$$



Event yields: tighter selection

- Z-veto, N(jets)>=1

L=0.84pb-1

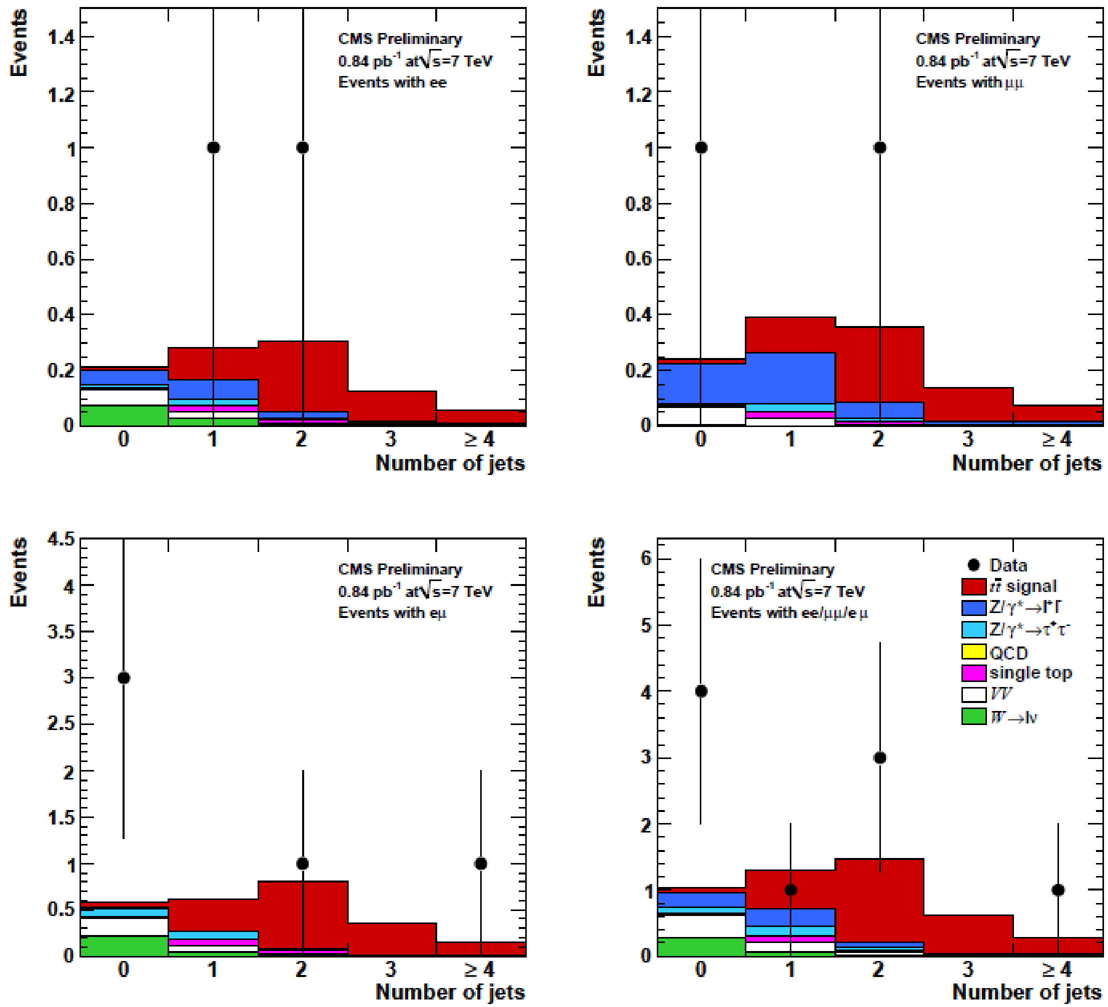
Sample	ee	$\mu\mu$	$e\mu$
Dilepton $t\bar{t}$	$0.63 \pm 0.09 \pm 0.12$	$0.70 \pm 0.11 \pm 0.13$	$1.70 \pm 0.26 \pm 0.32$
VV	0.05 ± 0.03	0.05 ± 0.03	0.12 ± 0.06
Single top - tW	0.04 ± 0.02	0.05 ± 0.03	0.12 ± 0.06
Drell-Yan $\tau\tau$	0.08 ± 0.04	0.13 ± 0.07	0.19 ± 0.09
Drell-Yan $ee, \mu\mu$	4.2 ± 1.1	5.0 ± 1.2	0.04 ± 0.02
Non-dilepton $t\bar{t}$	0.02 ± 0.01	0.003 ± 0.002	0.03 ± 0.02
W+jets	0.06 ± 0.03	$0.000\ ^{+0.002}_{-0.000}$	0.07 ± 0.04
QCD multijets	$0\ ^{+10}_{-0}$	$0\ ^{+10}_{-0}$	$0\ ^{+10}_{-0}$
Total simulated	5.1 ± 1.1	5.9 ± 1.2	2.3 ± 0.4
QCD data-driven	$0.0\ ^{+0.1}_{-0.0}\ ^{+0.1}_{-0.0}$	$0.0\ ^{+0.2}_{-0.0}\ ^{+0.2}_{-0.0}$	$0.0\ ^{+0.1}_{-0.0}\ ^{+0.1}_{-0.0}$
W+jets data-driven	$0.2\ ^{+0.2}_{-0.0}\ ^{+0.1}_{-0.0}$	$0.0\ ^{+0.4}_{-0.0}\ ^{+0.2}_{-0.0}$	$0.0\ ^{+0.4}_{-0.0}\ ^{+0.2}_{-0.0}$
Drell-Yan data-driven	$3.6 \pm 0.6 \pm 1.8$	$4.3 \pm 0.7 \pm 2.1$	N/A
Data	6	6	2

Systematics:
- Signal and DY: 15% acc*eff (conservative), 15% theory, 11% lumi
- Other backgrounds: 50% (conservative)
- Data-driven backgrounds: DY,Wjets: 50%; QCD: 100%

Good agreement observed!

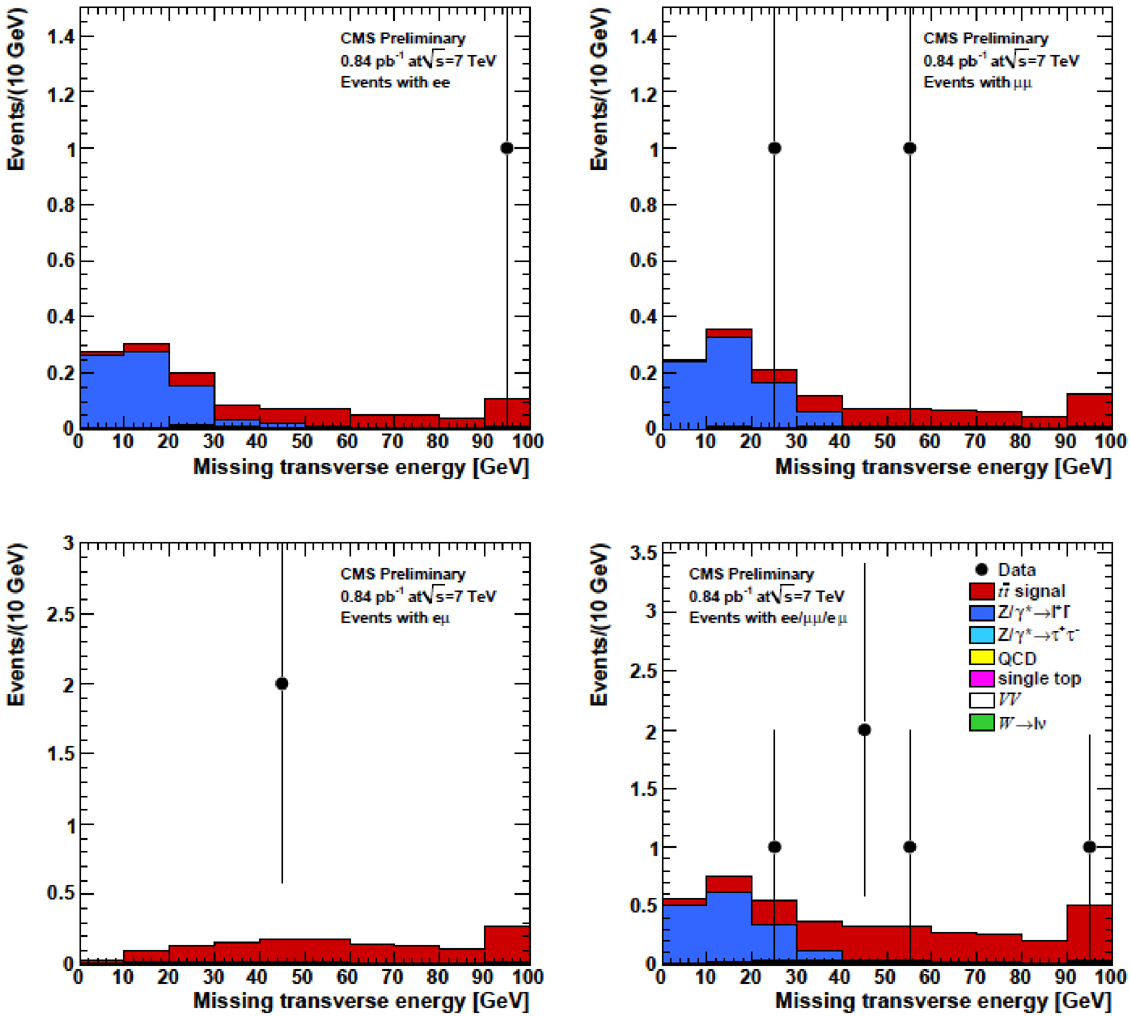
- N(jets), with Z-Veto, MET requirement applied

ee
mumu
emu
sum



Tightening the selection

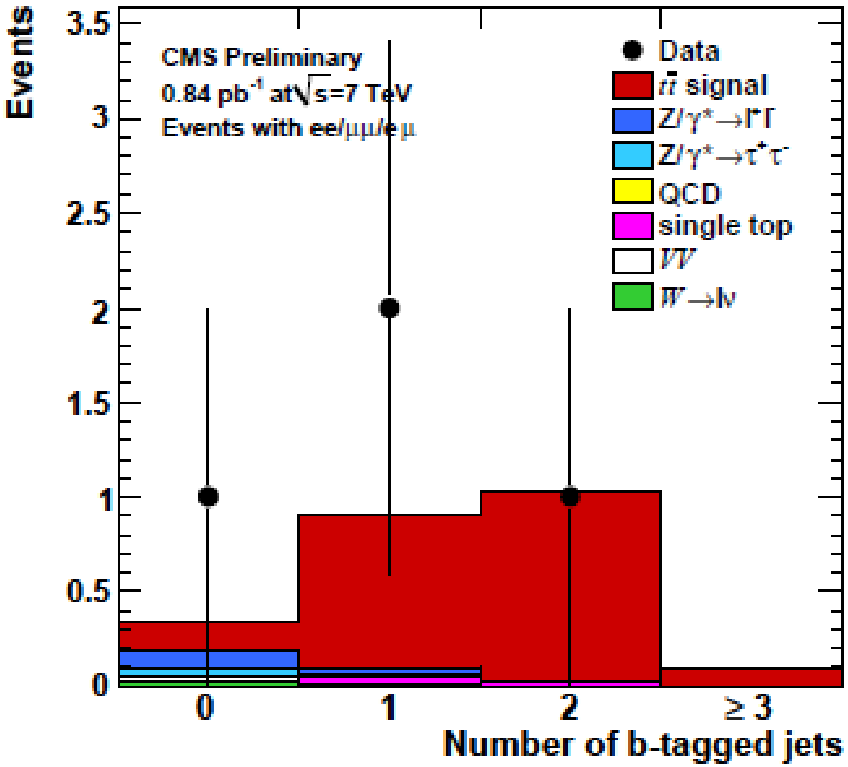
- MET, with Z-veto, N(jets)>=2 requirements applied

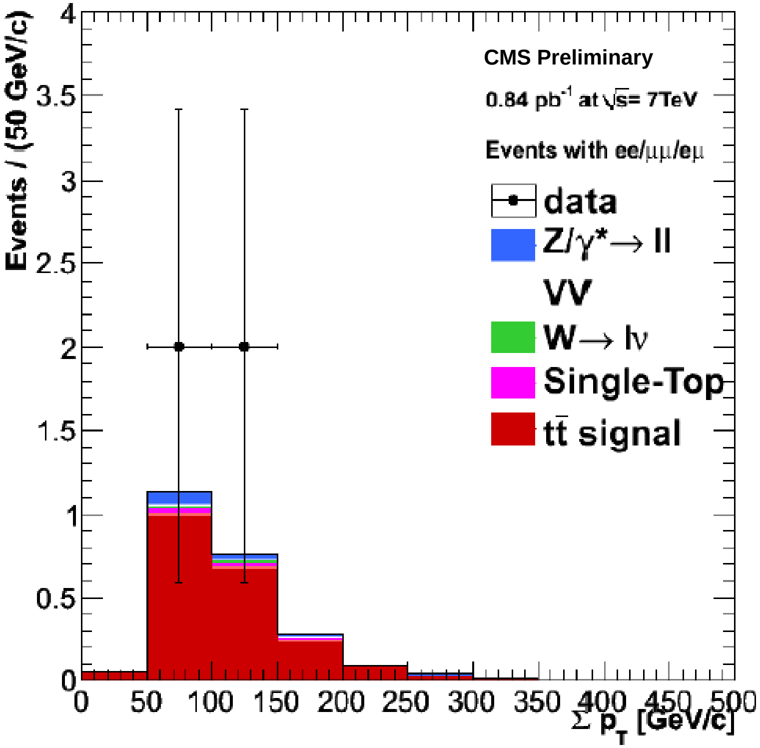


Dilepton: full selection

- All cuts applied: Z-Veto, MET, N(jets)>=2

L=0.84pb^{-1}



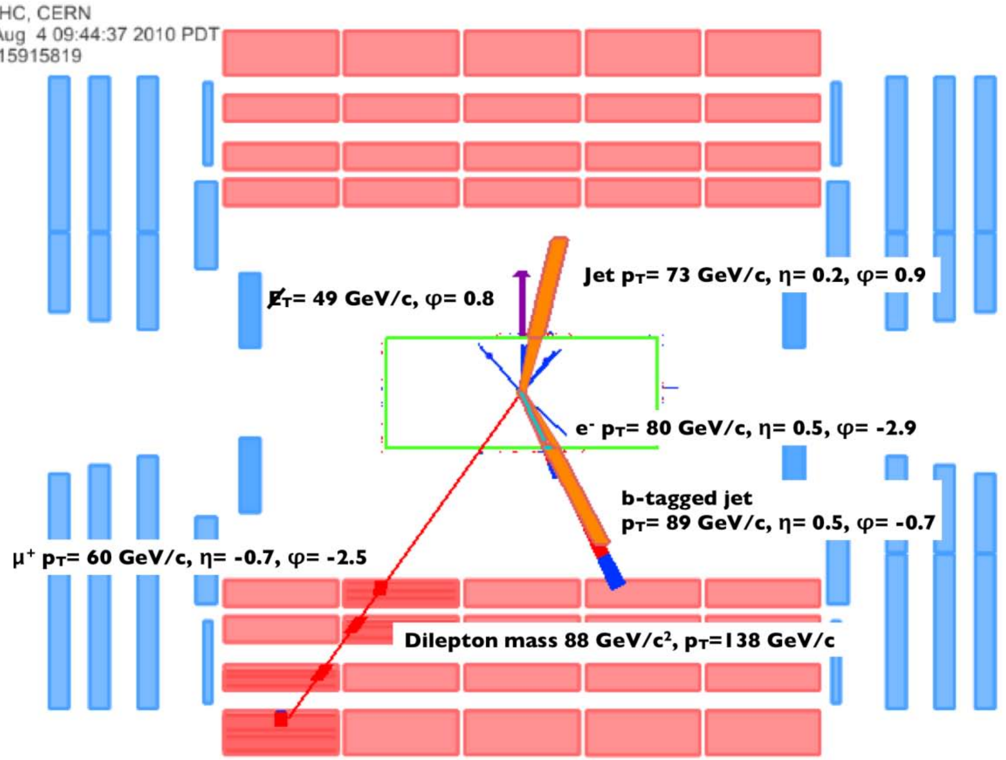


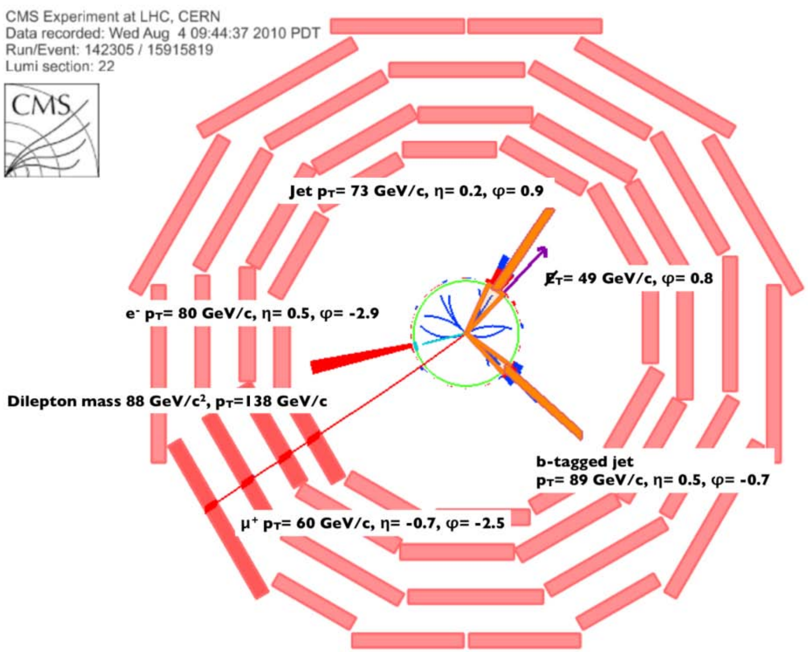
Track-counting tagger

4 clean ttbar candidates observed
~2.1 signal events expected

A beautiful e-mu candidate

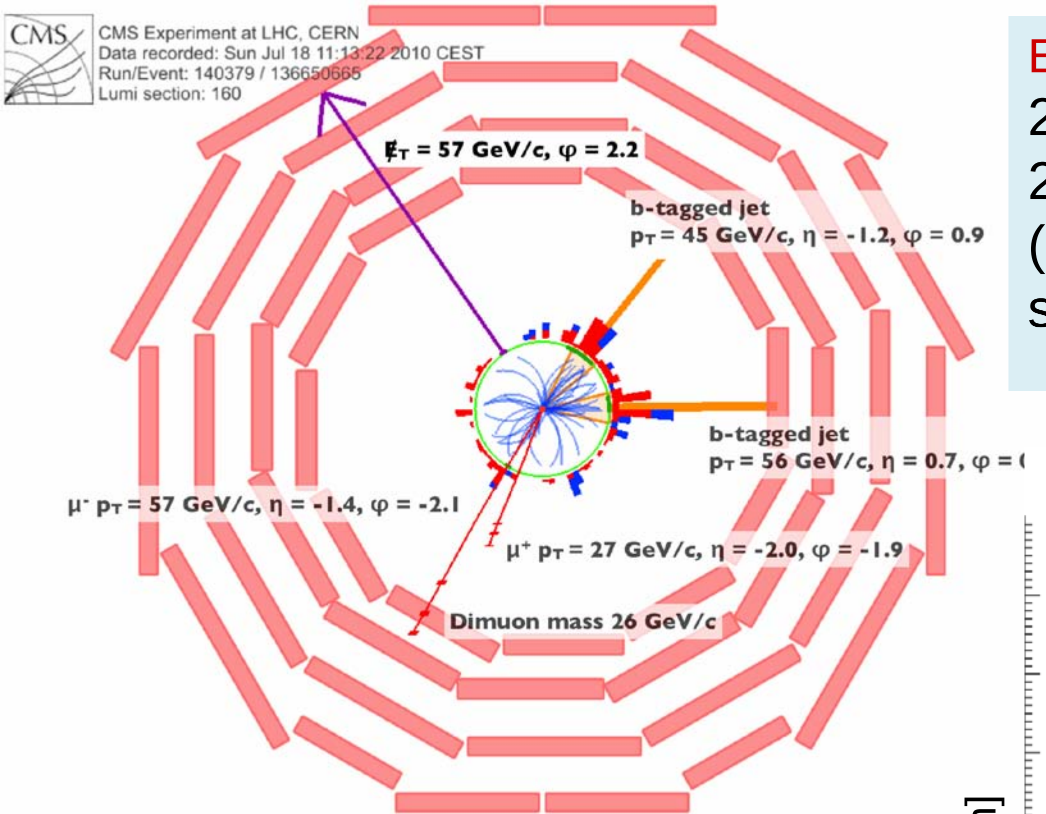
Muon Pt=60 GeV
Electron Pt=80 GeV
2 jets, 1 b-tag
MET=49 GeV





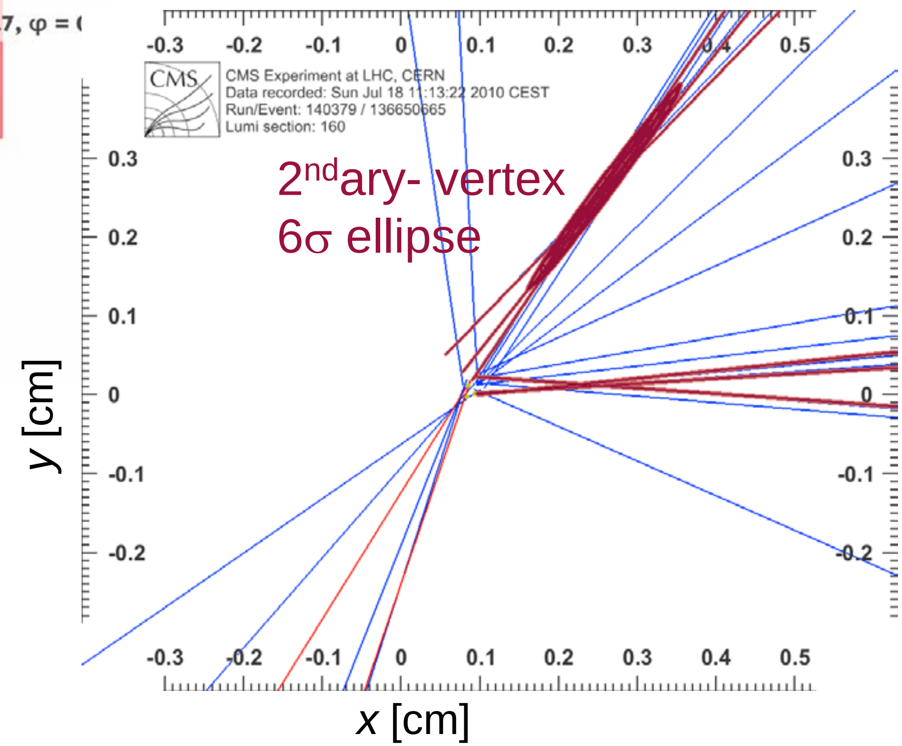
Mass hypothesis consistent with being a ttbar event

Dimuon event with 2 b-tags



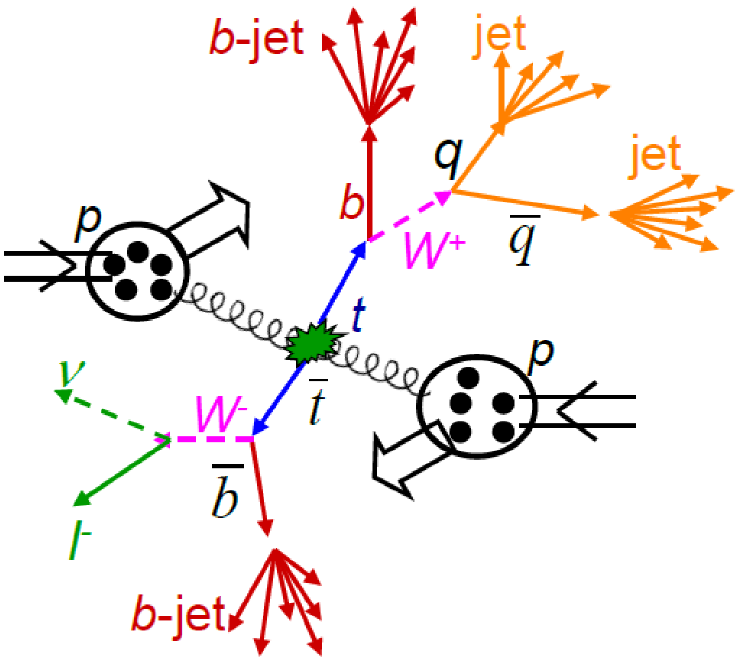
Event passes full selection:
2 muons with opposite charge
2 jets, both w/ good/clear *b*-tags
(and secondary vertices!)
significant MET (>50 GeV)

Preliminarily reconstr. mass in the range 160–220 GeV/c^2



Lepton+jets: Event selection

- Considered modes:
  - e+jets
  - mu+jets
- Single lepton triggers
- Exactly one isolated lepton
  - Muons: Pt>20 GeV,|eta|<2.1
    - Rel. Isolation < 0.05
  - Electrons: Pt>30 GeV, |eta|<2.4
    - Rel. Isolation, conversion veto
- Missing Et (MET)
  - Not used in event selection, but to reconstruct transverse Mass

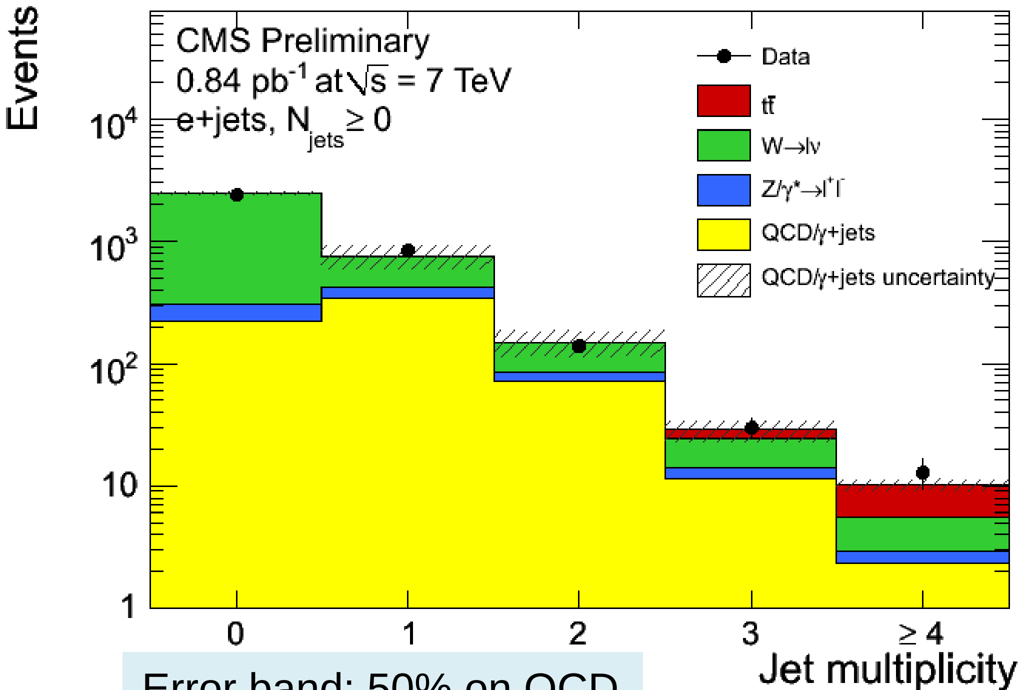


- Jets
  - Anti-Kt (R=0.5)
  - Pt>30 GeV, |eta|<2.4
  - Expect >=4 jets for ttbar
  - No b-tagging in baseline selection

Electron+Jets

No b-tagging, no MET cut applied

Jet multiplicity	ttbar	single top	W+jets	Z+jets	QCD	Sum MC	Data
$N_{jets} \geq 0$	12 ± 2	3.4 ± 0.4	2619 ± 317	180 ± 21	658 ± 73	3472 ± 326	3434
$N_{jets} \geq 1$	12 ± 2	3.1 ± 0.4	419 ± 77	92 ± 11	436 ± 62	962 ± 99	1022
$N_{jets} \geq 2$	11 ± 2	1.9 ± 0.3	74 ± 18	19 ± 5	85 ± 22	191 ± 29	183
$N_{jets} \geq 3$	8.9 ± 1.8	0.70 ± 0.14	13 ± 4	3.3 ± 1.0	14 ± 5	40 ± 7	43
$N_{jets} \geq 4$	4.8 ± 1.2	0.21 ± 0.06	2.6 ± 1.1	0.60 ± 0.23	2.3 ± 1.1	11 ± 2	13



Error band: 50% on QCD

MC Uncertainties (table):
- Jet energy scale (known to 10%)
- Luminosity (known to 11%)
- Cross section unc. (scale,PDF)

$L=0.84pb^{-1}$

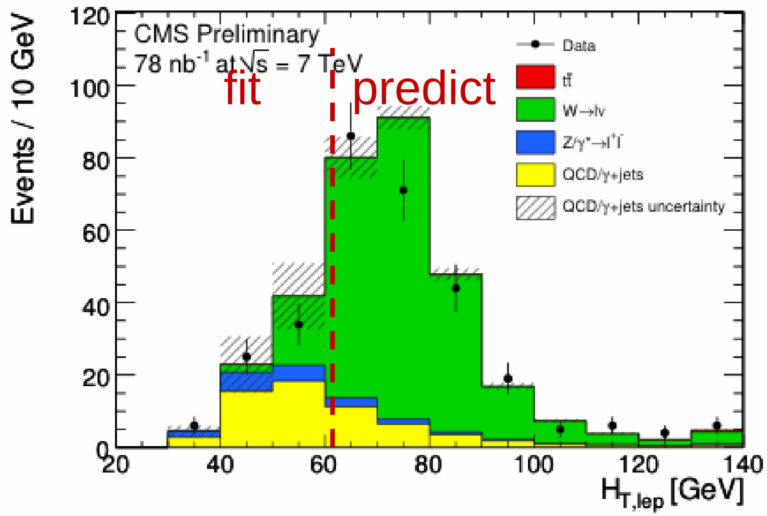
Good agreement observed in all Jet bins!

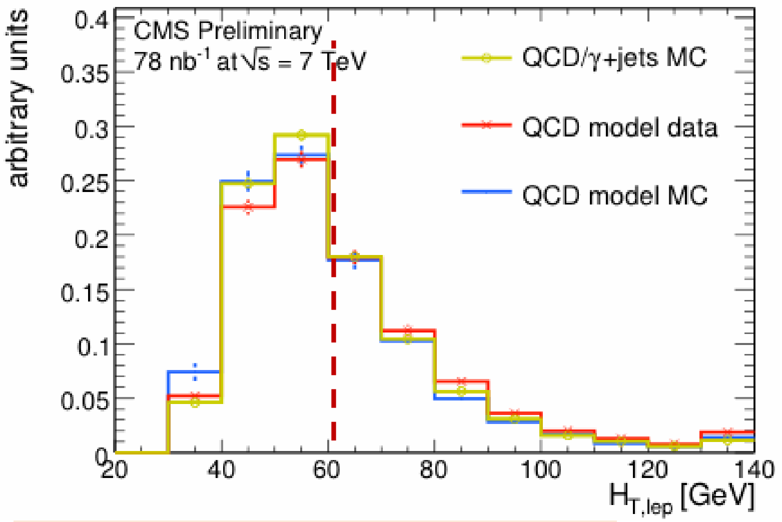
Data-driven QCD estimate: template fit

- First tests of DD methods in low N(jets) bins
- e+jets: Fit sum of templates in low MET or HT(lep) region
  - QCD template from multijet sample (near-miss electrons or large EM jets)
- Predict N(QCD) in signal region

N(jets)>=0	QCD MC	QCD estimate
MET>25 GeV	12.2+/-0.2	19+/-7
HT(lep)>60 GeV	26.0+/-0.3	39+/-11

N(jets)>=1	QCD MC	QCD estimate
MET>25 GeV	5.3+/-0.1	8+/-5
HT(lep)>60 GeV	12.4+/-0.2	10+/-4

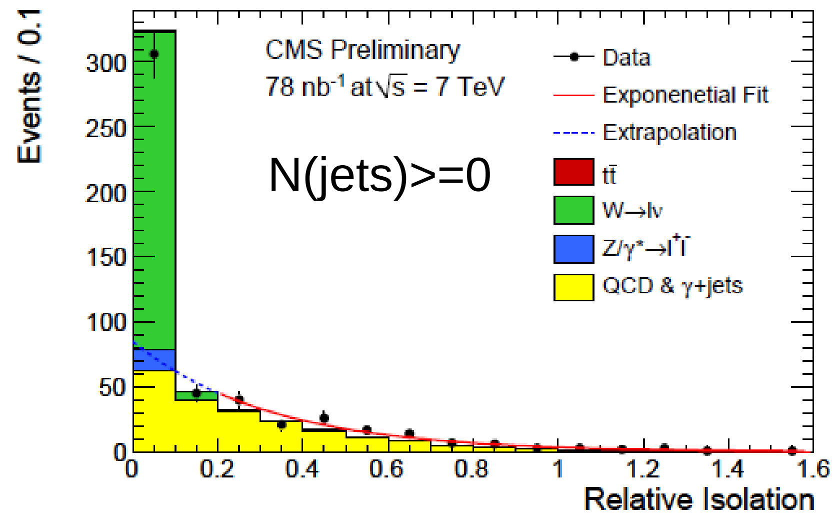




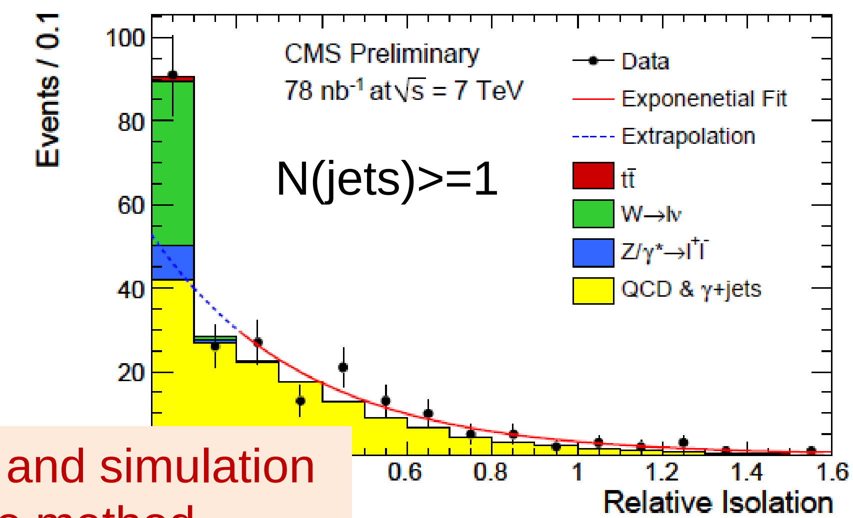
e+jets: consistent with MC
mu+jets: see later

Data-driven QCD II: Isolation extrapolation

- Fit function to isolation distribution in non-isolated (QCD dominated) region
- Extrapolate to isolated (W-like) region



Isolation extrapolation method (e+jets)		
Fit Range	$N^{est.}_{QCD}(\geq 0\text{-jet})$	$N^{est.}_{QCD}(\geq 1\text{-jet})$
0.1–1.6	67 ± 9	40 ± 6
0.2–1.6	73 ± 13	46 ± 9
0.3–1.6	71 ± 17	45 ± 12
Average $N^{est.}_{QCD}$	70 ± 35	44 ± 22
Prediction N^{MC}_{QCD}	63 ± 7	42 ± 6

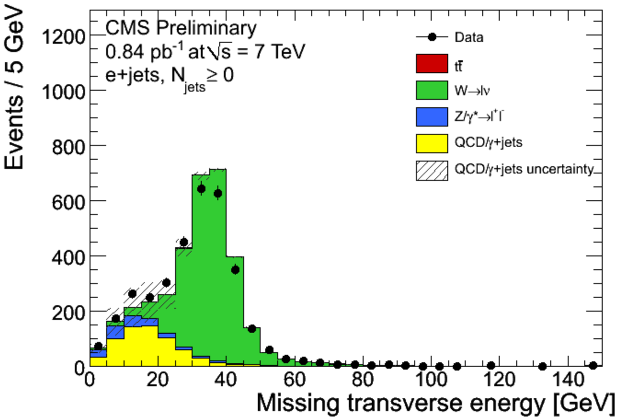


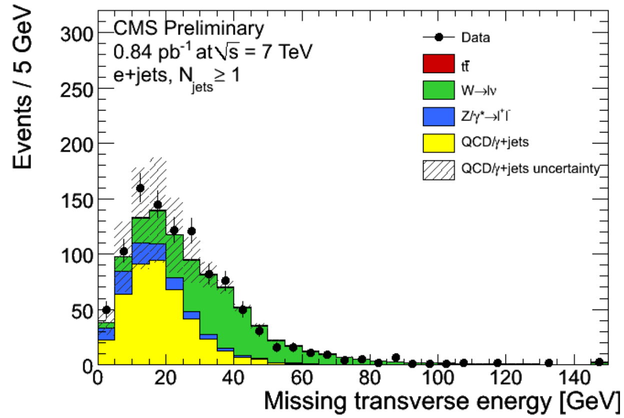
Agreement between data-driven QCD estimate and simulation
More importantly: result consistent with template method
(NB no MET,HT cuts applied here!)

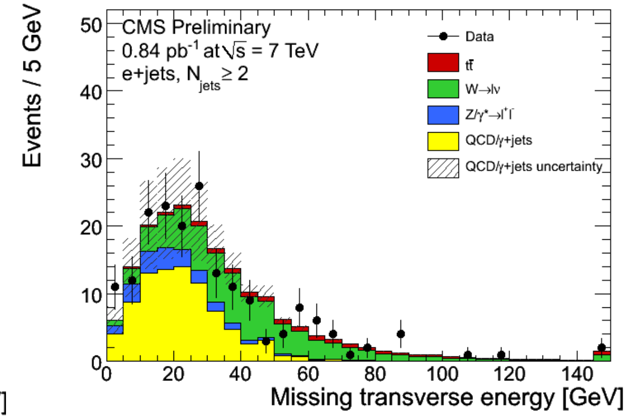
e+jets in various jet bins

$L=0.84pb^{-1}$

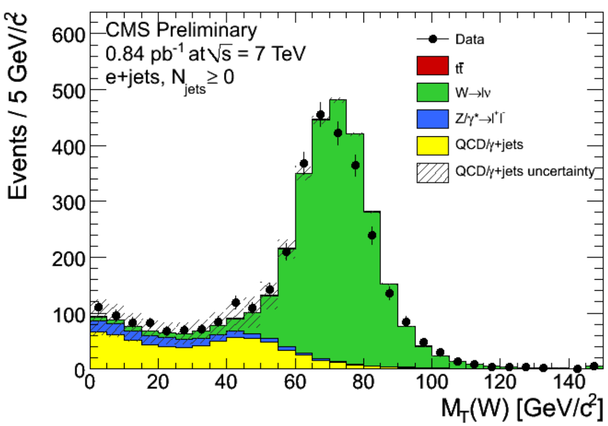
N(jets)>=0	N(jets)>=1	N(jets)>=2

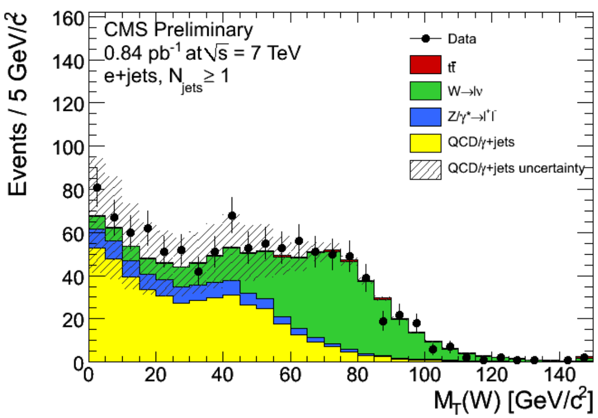


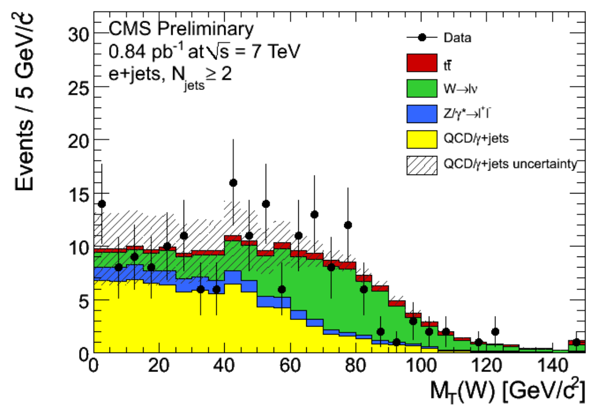




Missing ET (hard to get right, important for any top quark measurement)







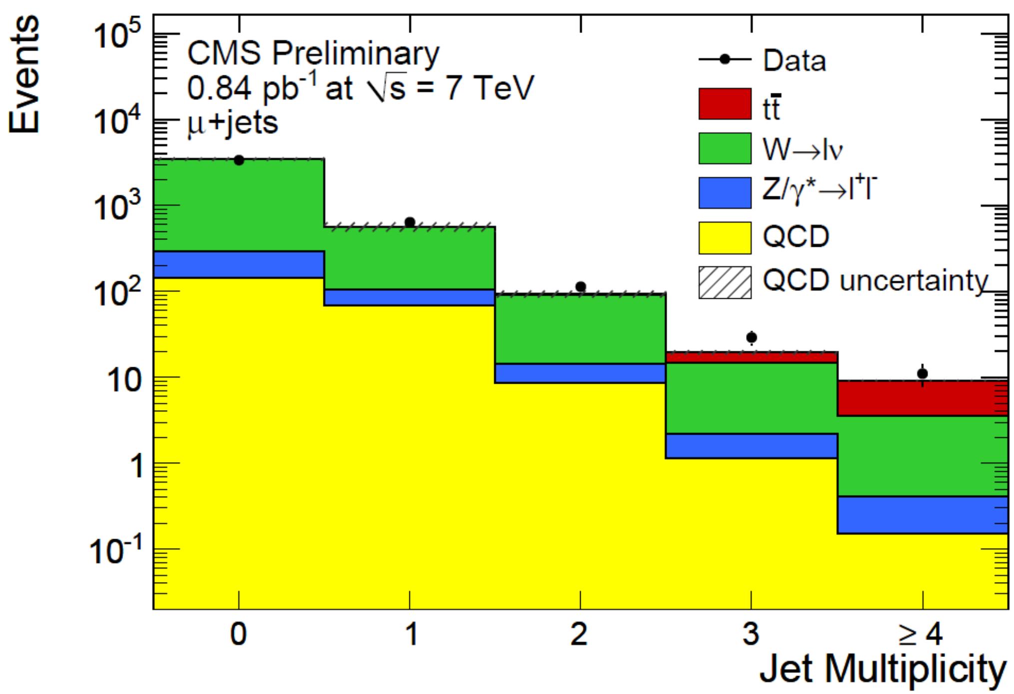
$M_T(W)$: transverse W mass (calculated from lepton+MET)

Good agreement Data-Simulation! QCD background important in e+jets!

Muon+jets

No b-tagging, no MET cut applied

Jet multiplicity	ttbar	single top	W+jets	Z+jets	QCD	Sum MC	Data
$N_{jets} \geq 0$	13 ± 3	4.2 ± 0.4	3708 ± 448	192 ± 29	223 ± 25	4140 ± 450	4142
$N_{jets} \geq 1$	13 ± 3	3.9 ± 0.4	552 ± 106	42 ± 12	79 ± 17	690 ± 108	789
$N_{jets} \geq 2$	13 ± 2	2.3 ± 0.3	92 ± 24	7.1 ± 4.4	10 ± 3	124 ± 25	153
$N_{jets} \geq 3$	10 ± 2	0.82 ± 0.15	16 ± 5	1.3 ± 0.9	1.3 ± 0.5	29 ± 5	40
$N_{jets} \geq 4$	5.6 ± 1.4	0.24 ± 0.06	3.1 ± 1.2	0.25 ± 0.18	0.15 ± 0.07	9.3 ±1.9	11

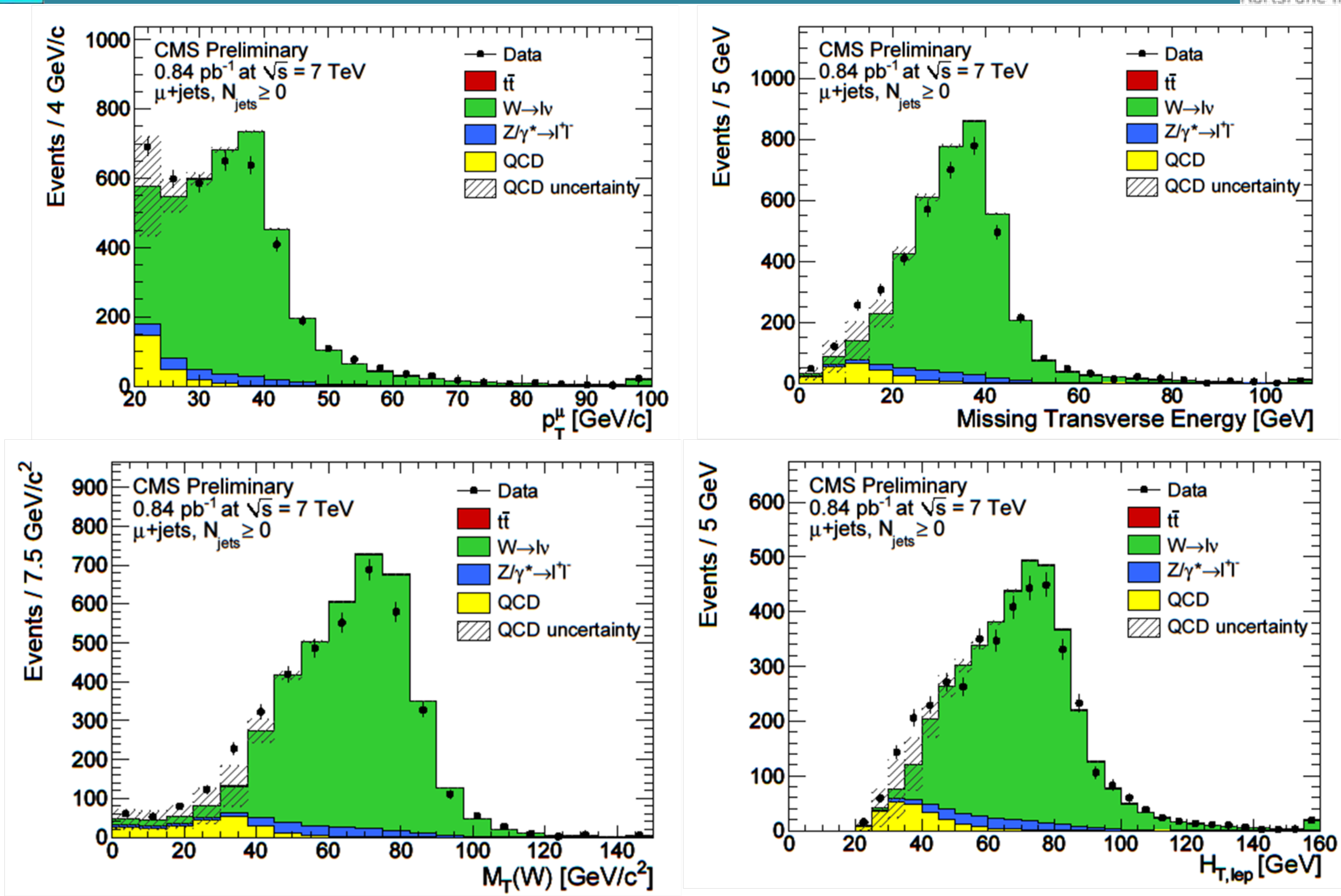


MC Uncertainties (table):
- Jet energy scale (known to 10%)
- Luminosity (known to 11%)
- Cross section unc. (scale,PDF)

L=0.84pb^{-1}

Good agreement observed in all Jet bins!

mu+jets, N(jets)>=0

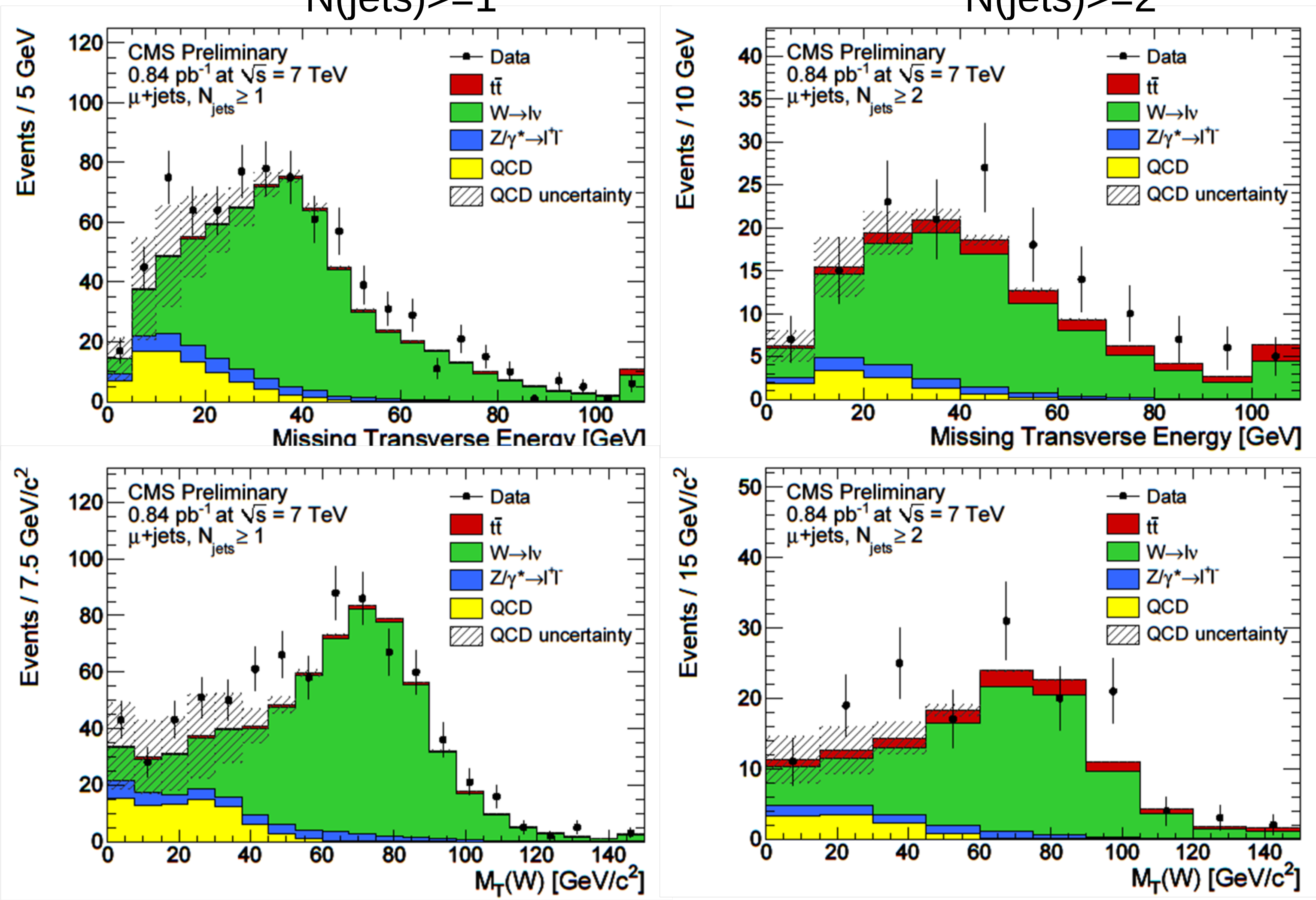


Excess observed in data at low Pt(mu), MET, MT and HT
Consistent with QCD MC being factor ~2 too low

Error band: 100% on QCD
(from data-driven methods)

mu+jets, N(jets)>=1,2

N(jets)>=1 | N(jets)>=2

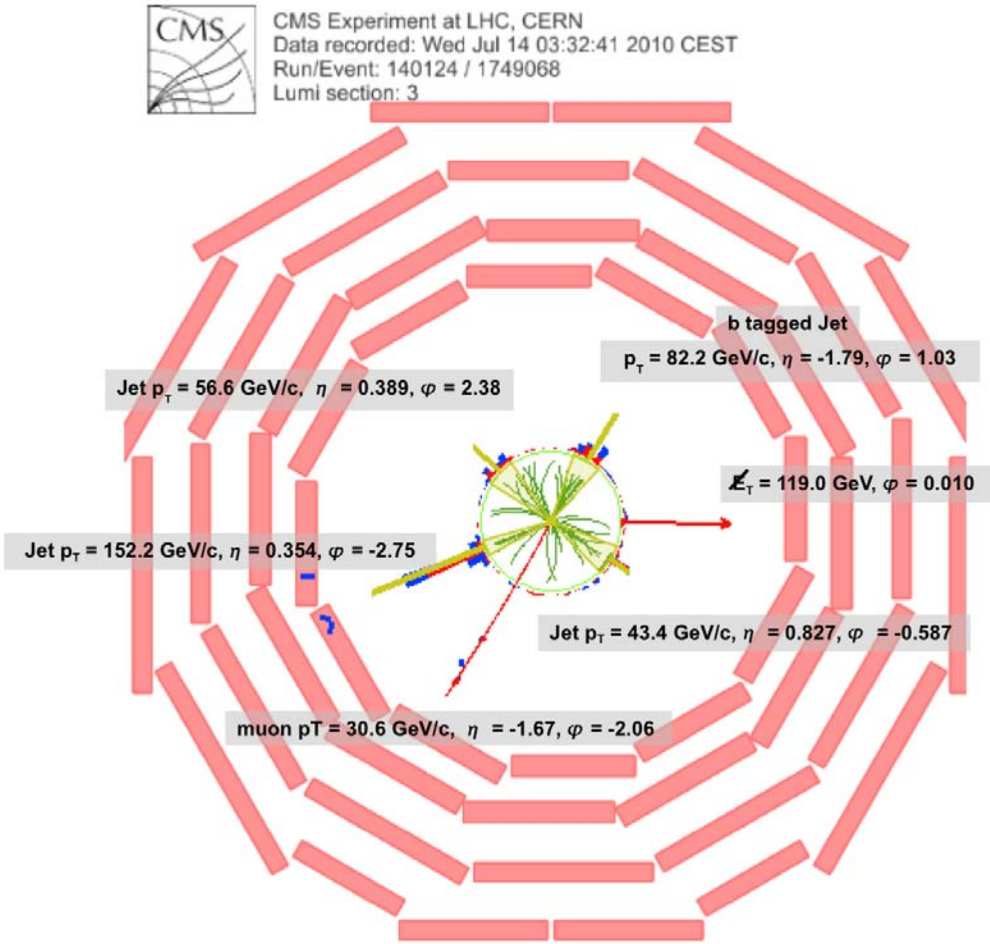


Consistent with QCD too low by factor ~2, indep. of N(jets)

Data slightly above MC also where QCD less important

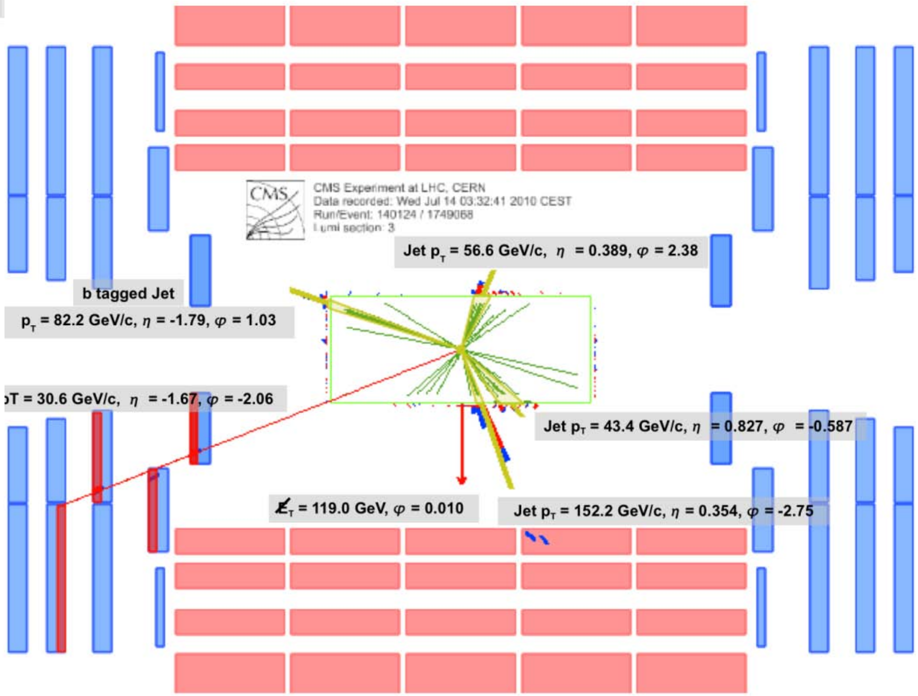
Note: expect significant JES & theory uncertainties (not incl. in error bars!)

A beautiful mu+jets candidate



Event passes all cuts of full selection

1 high-momentum muon
significant MET>100GeV
$m_T(W)$ = 104 GeV/c^2
4 high-p_T jets,
one of which with good *b*-tag



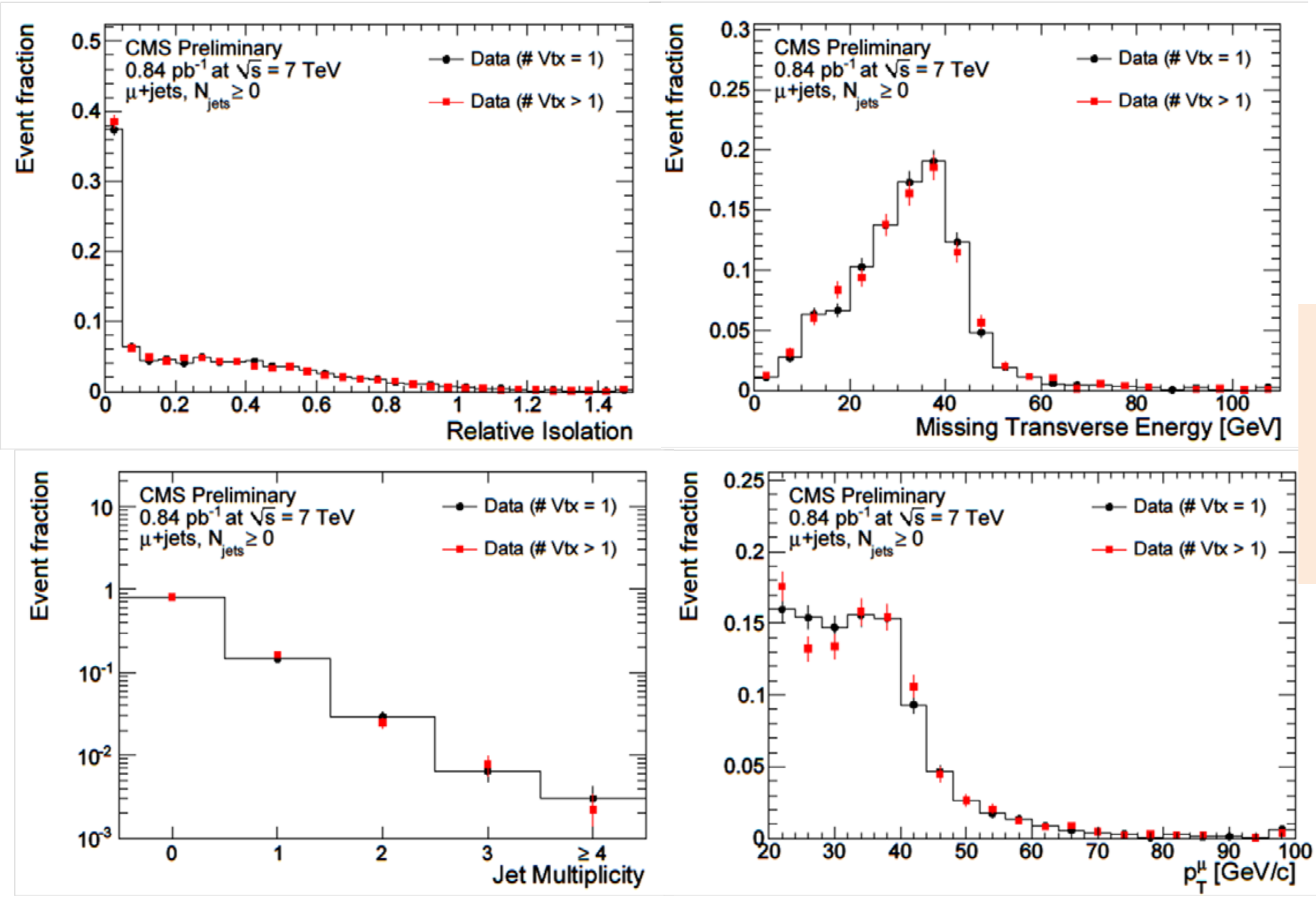
reconst. top mass around 210 GeV/c^2

masses of 2 untagged jets (3 possible comb.): 104, 105, 151 GeV/c^2

Are we sensitive to Pile-up?

Do have non-negligible pileup in recent data <N>~0.9
Simulation is without pileup
Compare data with one vertex vs data with >=1 vertex

L=0.84pb^{-1}



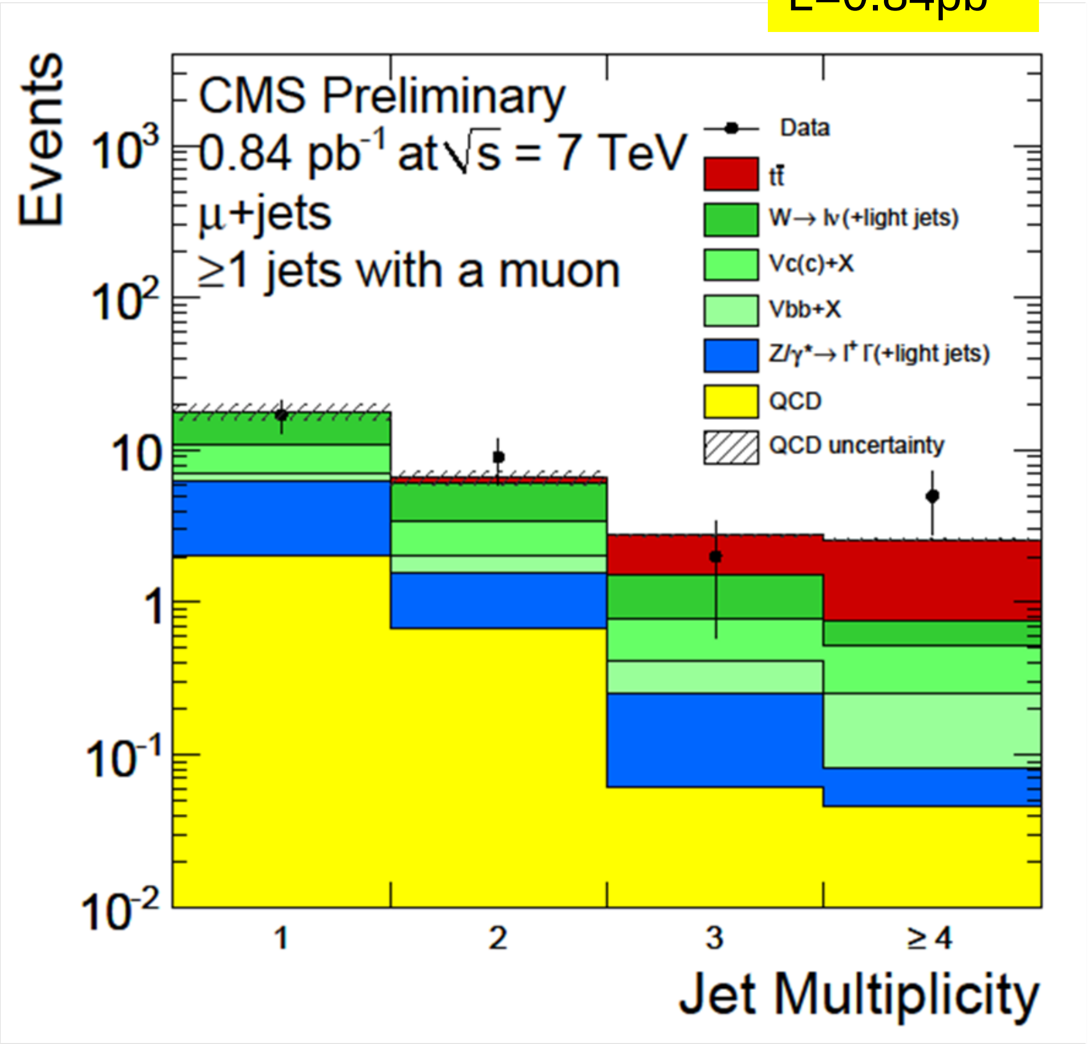
So far little effect on sensitive distributions (e.g. isolation, MET)

Enriching the b-content

- "Muon-in-jet"
  - mu+jets: request at least one jet associated with a muon within dR<0.4
  - Sensitive to semileptonic b-decays in jets

- For N(jets)>=3, observe 7 events, consistent with ttbar signal plus ~2.5 background events

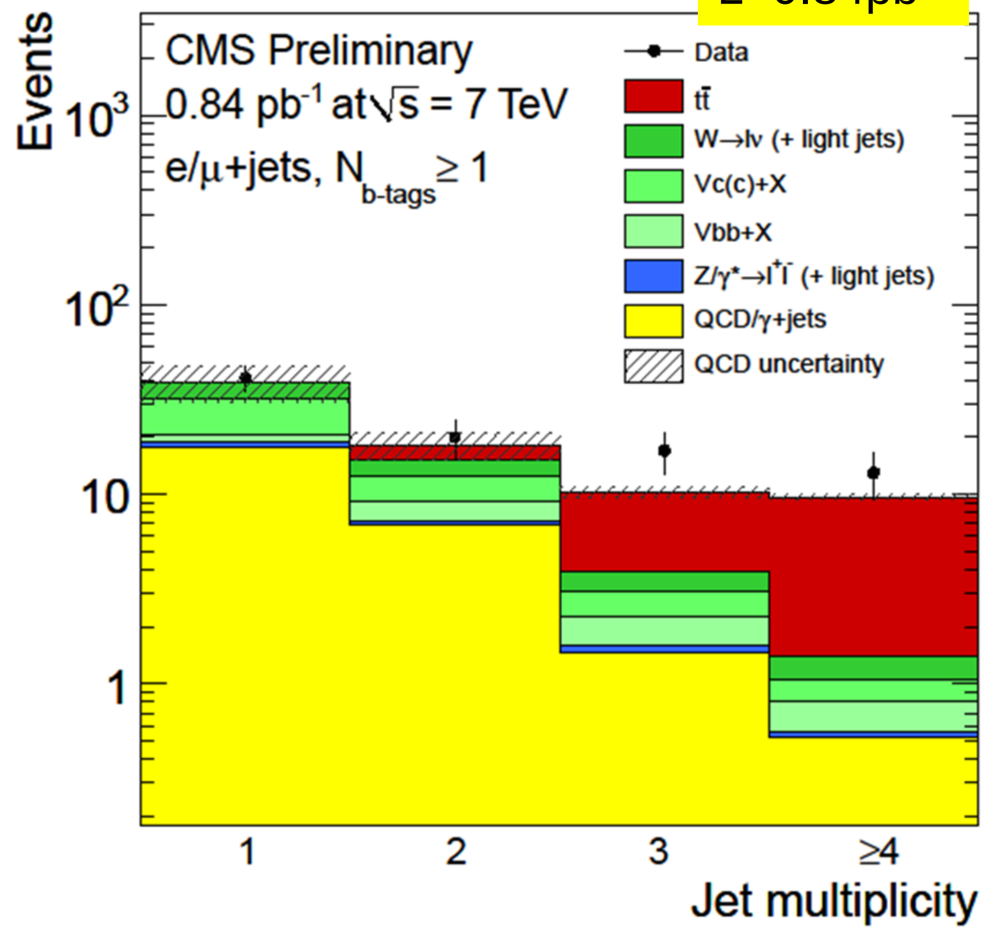
L=0.84pb-1



e/mu+jets with b-tagging

- e/mu+jets combined
- Secondary vertex tagger (working point with ~1% fake rate)
- For N(jets)>=3:
  - Observed N(data)=30
  - Prediced background N(BG,MC)=5.3
  - Predicted signal N(ttbar,MC)=15

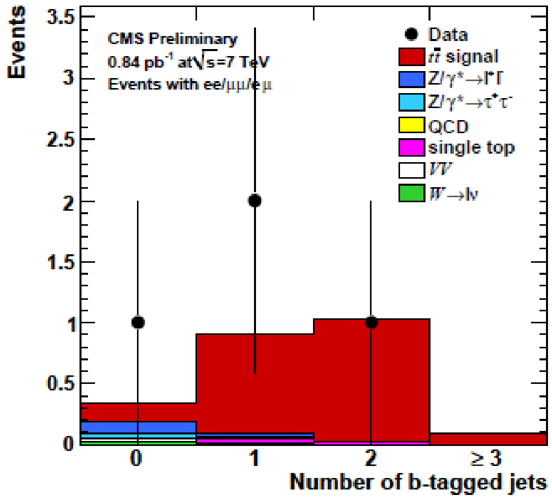
L=0.84pb-1

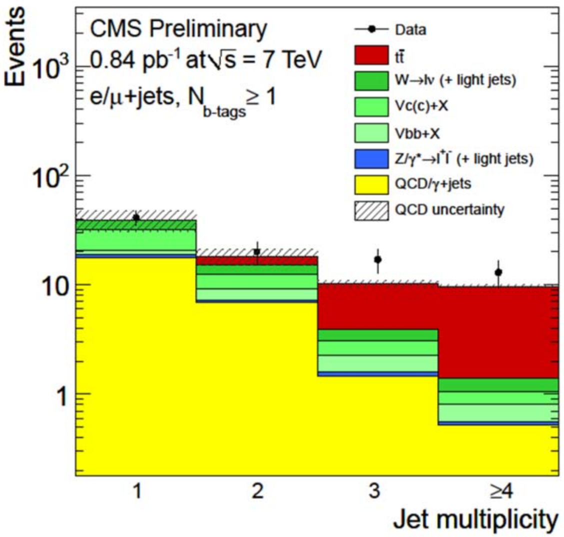


Seeing ttbar events at a rate roughly consistent with NLO cross section, considering experimental (JES,b-tagging) and theoretical (scale, PDF, HF modelling, …) uncertainties

Conclusions

- Analyzed first L=0.84pb-1 of 7TeV data
  - o Updated with respect to ICHEP results (L~250nb-1)
- Event yields in background dominated regions ~consistent with expectations, within uncertainties
  - o Tests of data-driven background estimation
- Enrich signal by going to high N(jets) and employing b-tagging
- Observed number of candidates approx. consistent with ttbar expectation, on top of small backgrounds
- Strong evidence for excellent performance of CMS detector (jets, MET, leptons, b-tagging)!

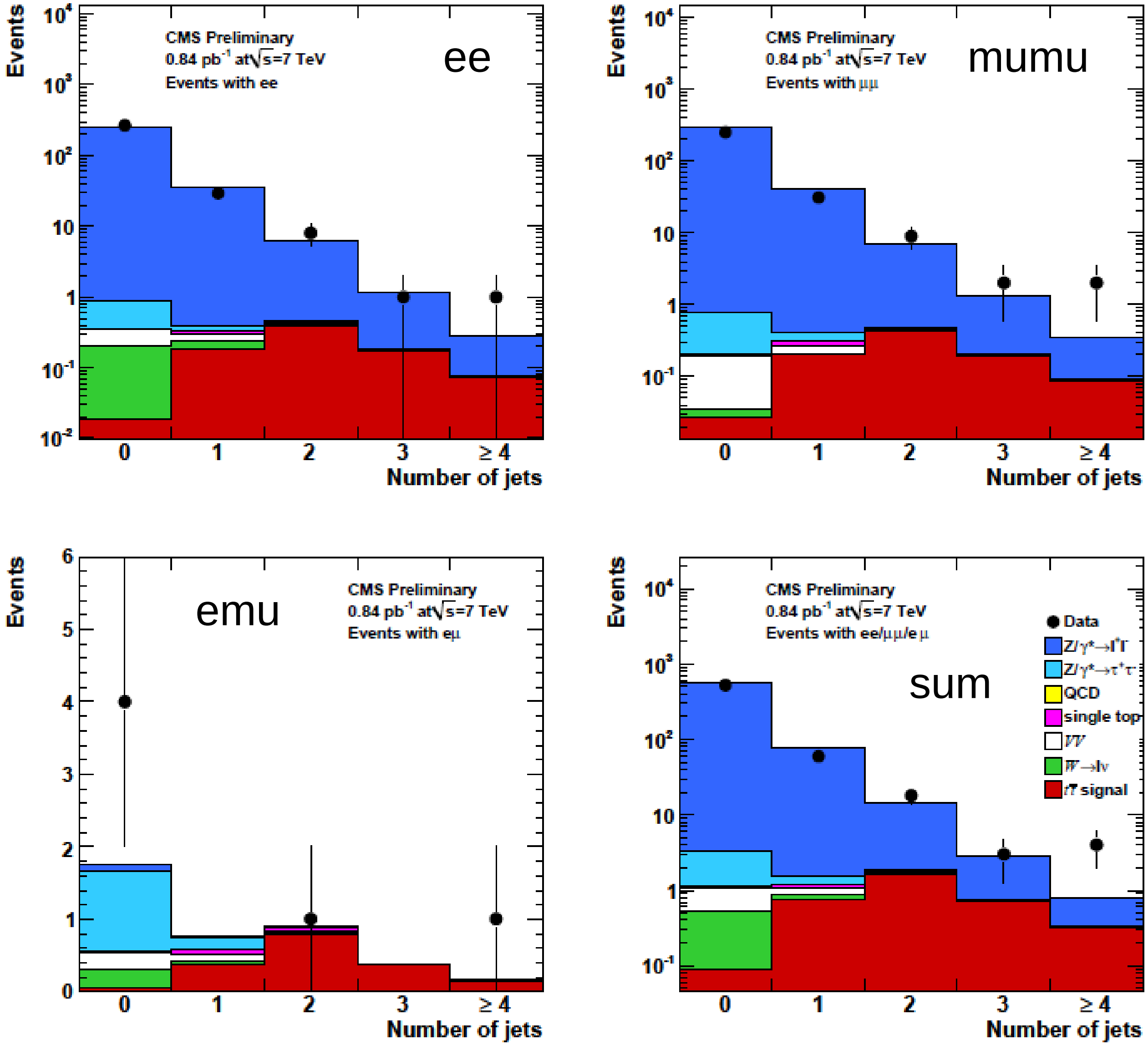


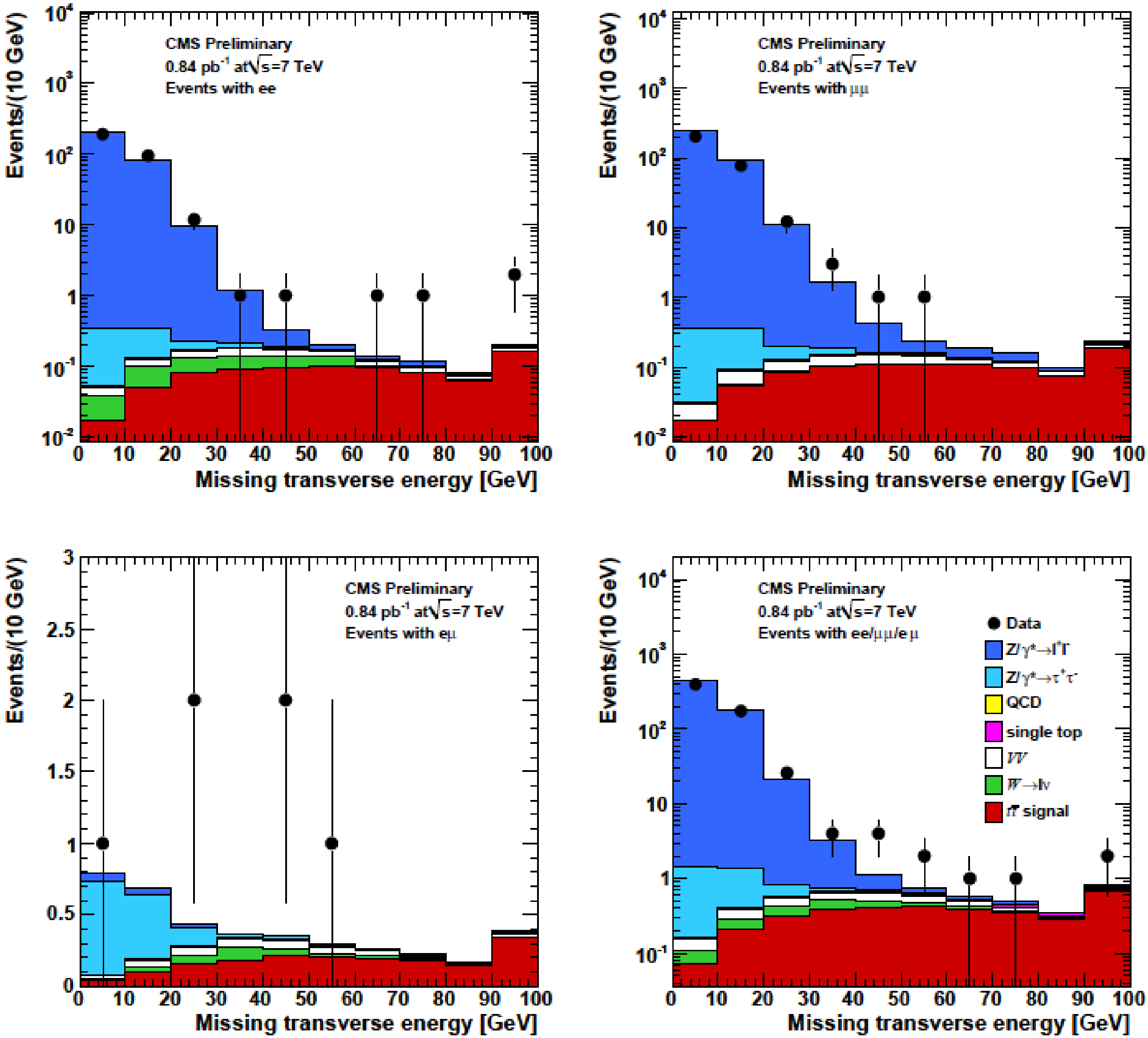


Established Top signal at LHC, first cross sections will come soon!

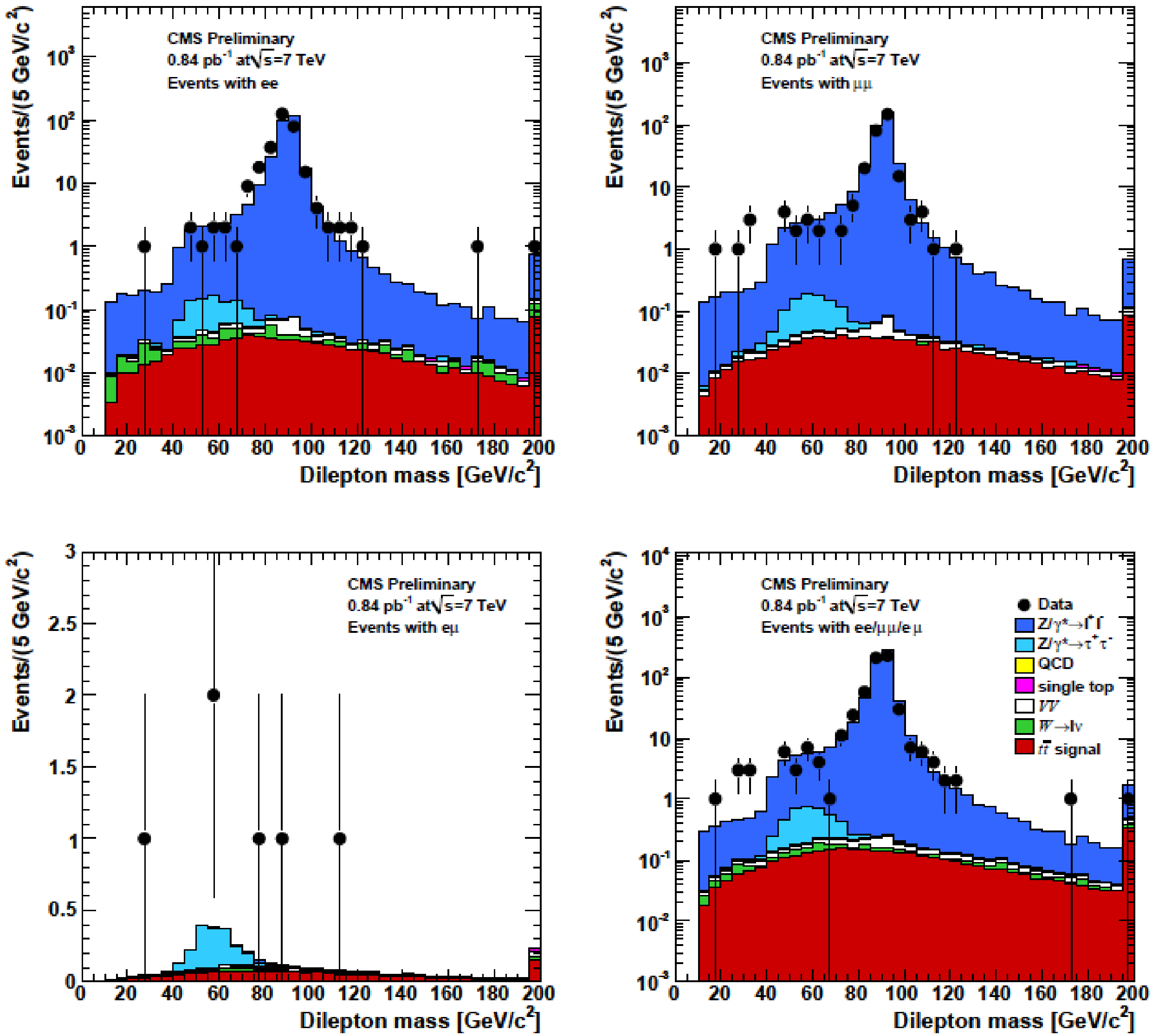
Backup

Jet multiplicity (relaxed selection)

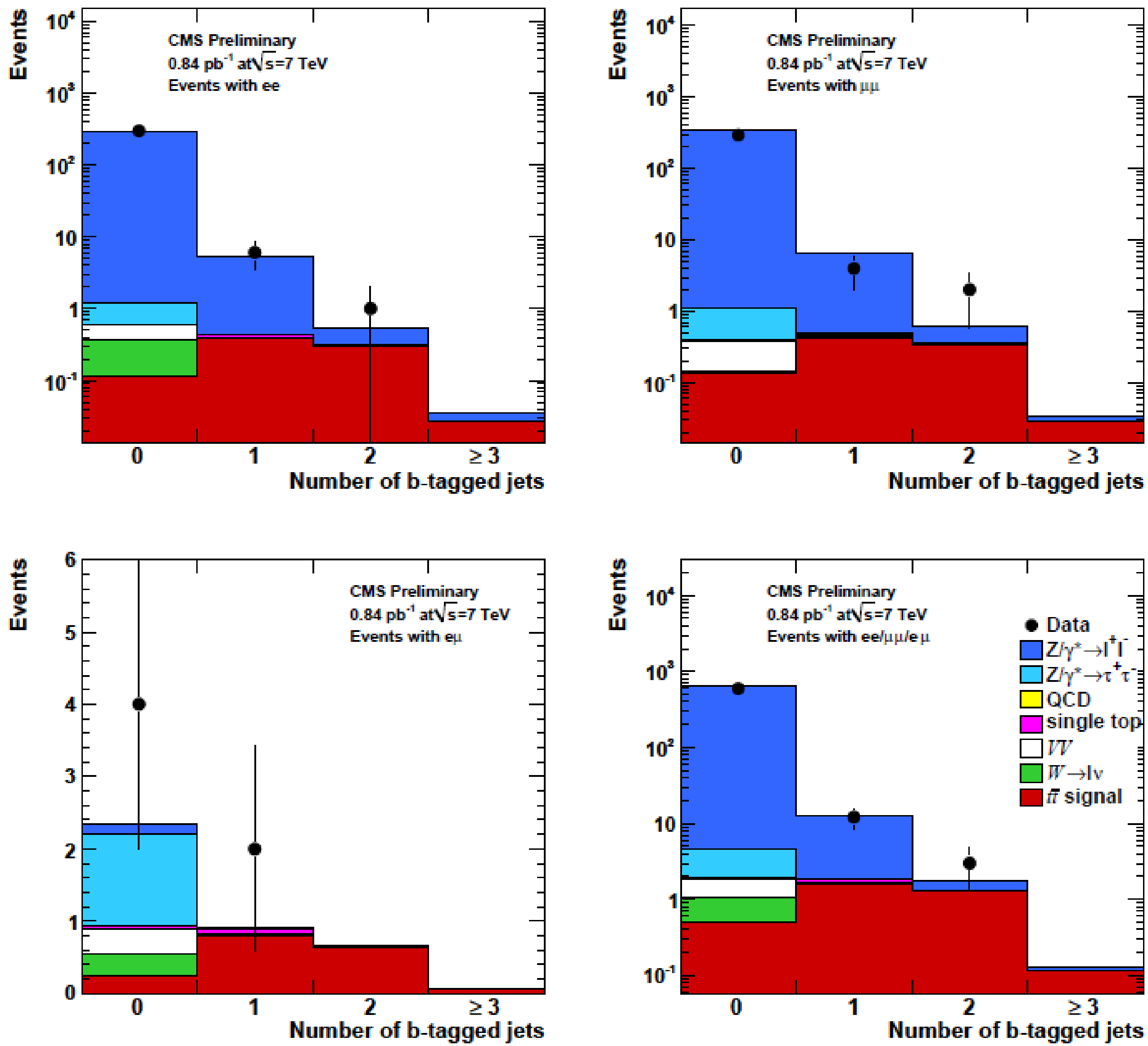



CMS Preliminary
0.84 pb⁻¹ at √s=7 TeV
Events with ee
CMS Preliminary
0.84 pb⁻¹ at √s=7 TeV
Events with μμ
CMS Preliminary
0.84 pb⁻¹ at √s=7 TeV
Events with eμ
CMS Preliminary
0.84 pb⁻¹ at √s=7 TeV
Events with ee/μμ/eμ
Data
Z/γ*→l⁺l⁻
Z/γ*→τ⁺τ⁻
QCD
single top
VV
W→lν
tt̄ signal
Events/(10 GeV)
Missing transverse energy [GeV]

Dilepton invariant mass (relaxed sel.)

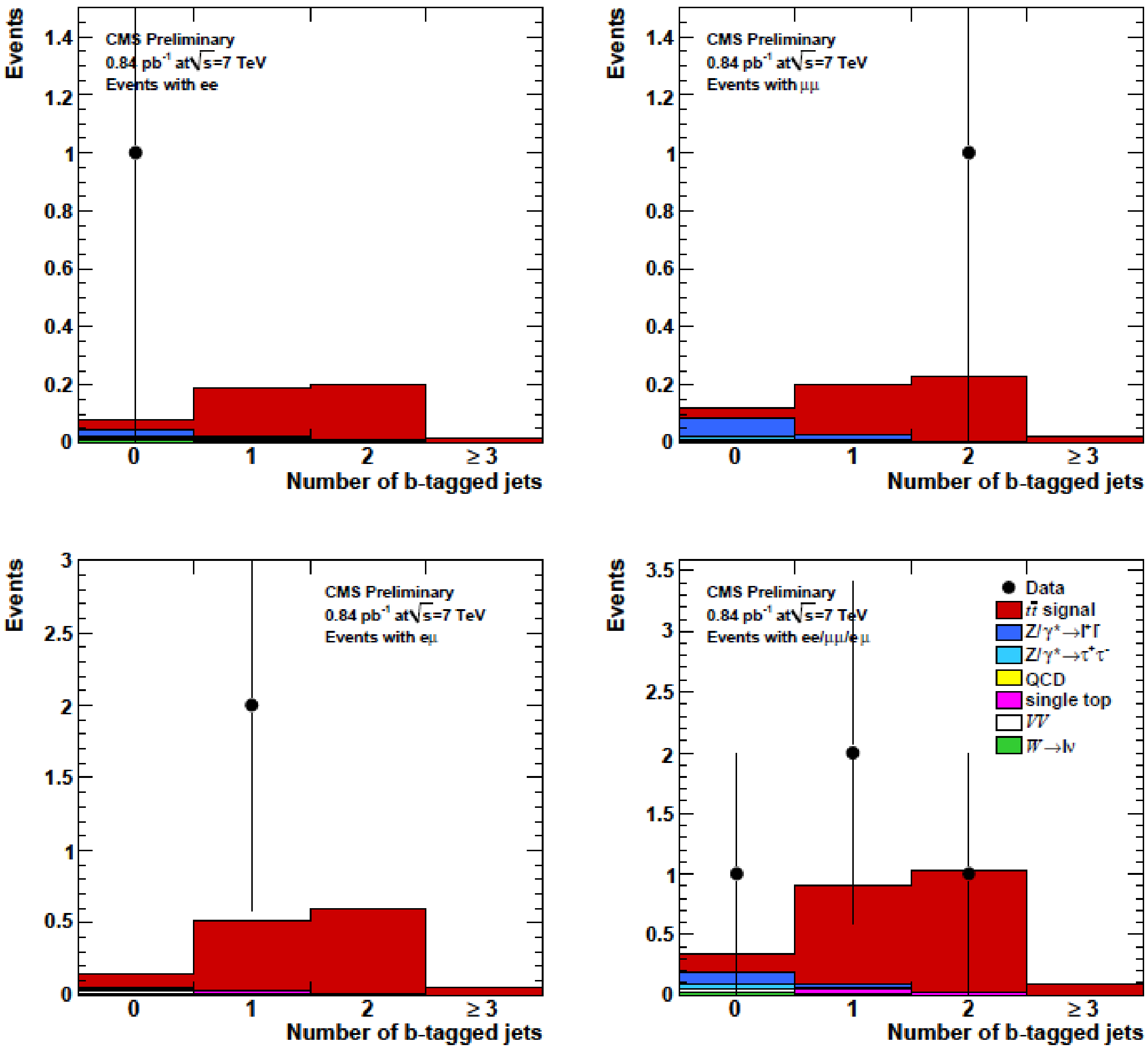


b-tag multiplicity (relaxed sel.)

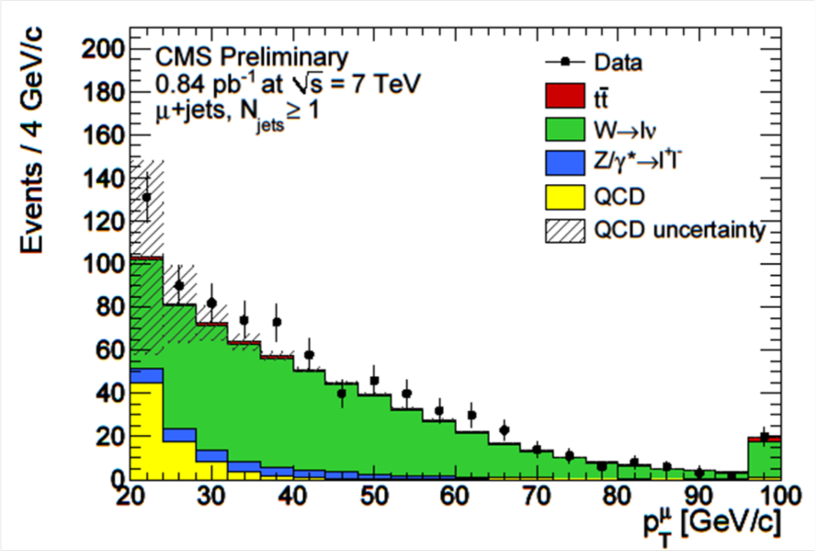


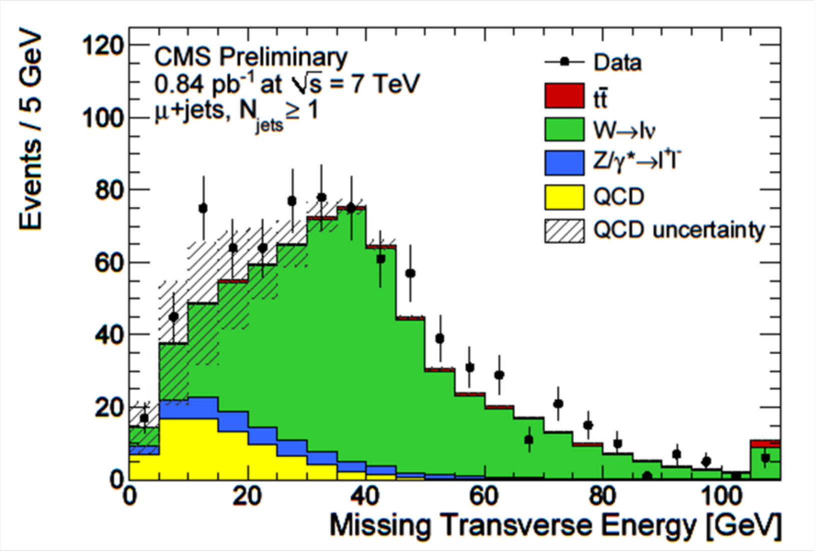
b-tag multiplicity (full selection)

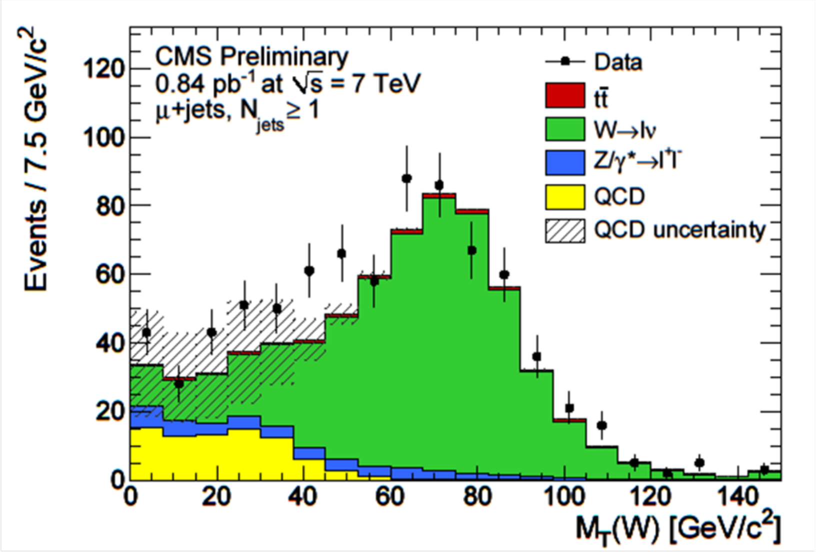
- With Z-veto, MET, N(jets)>=2 cuts applied

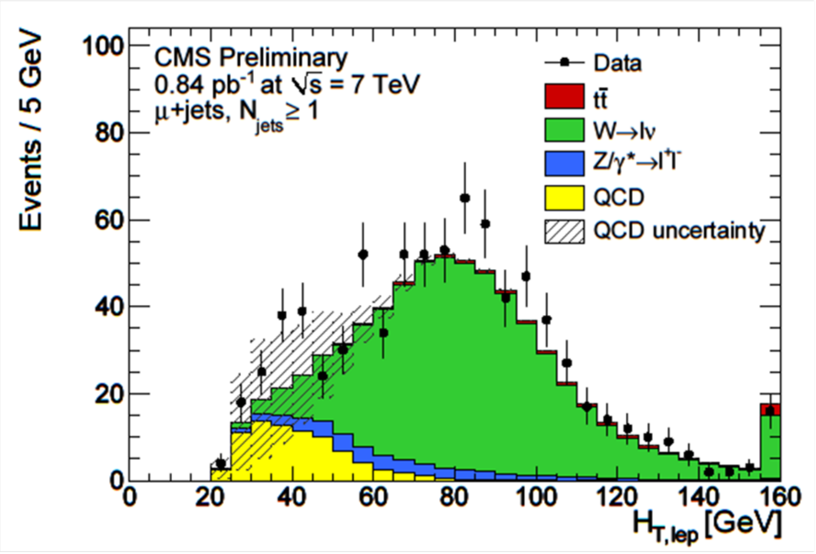


mu+Jets, >=1 jets

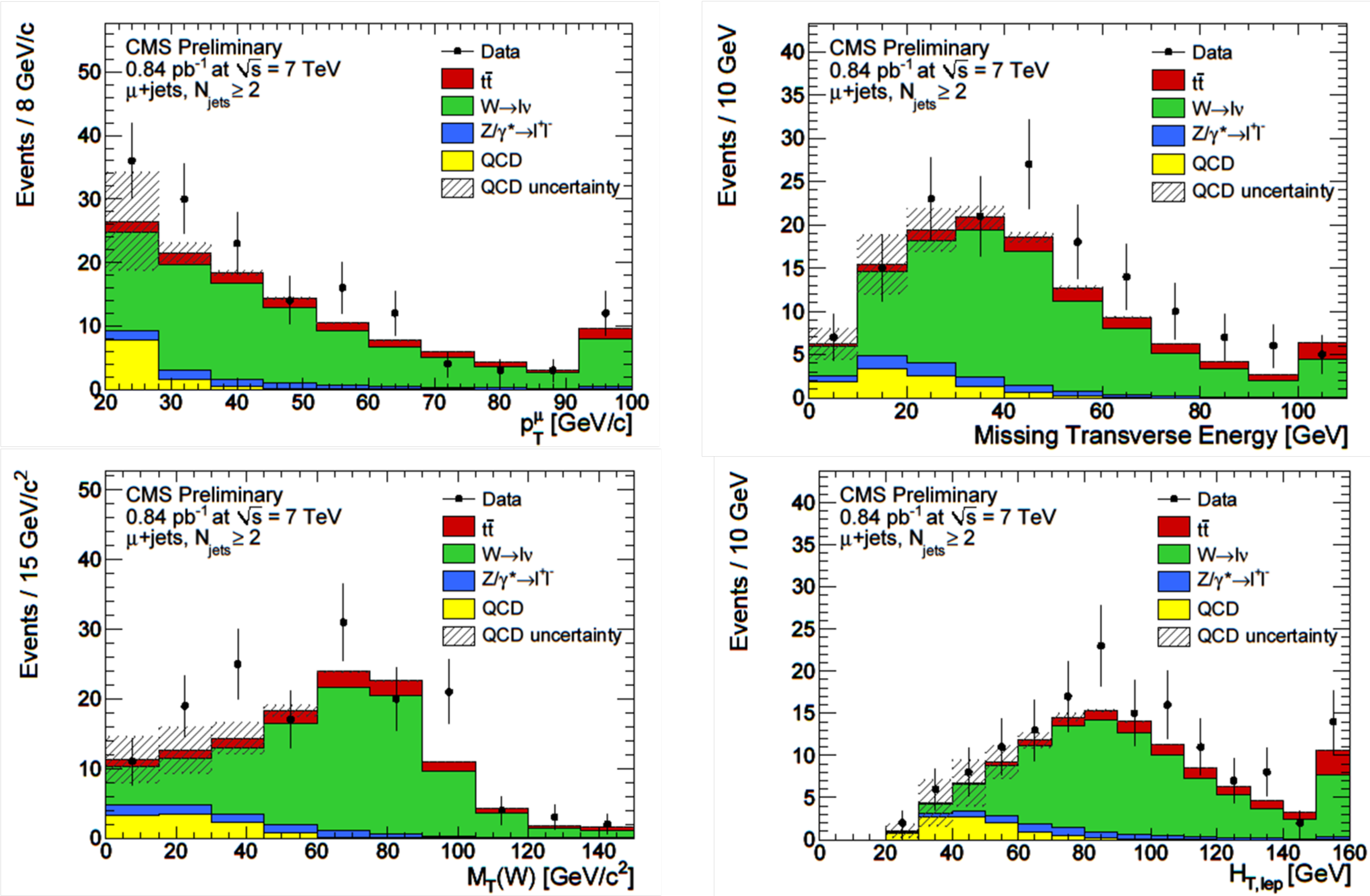






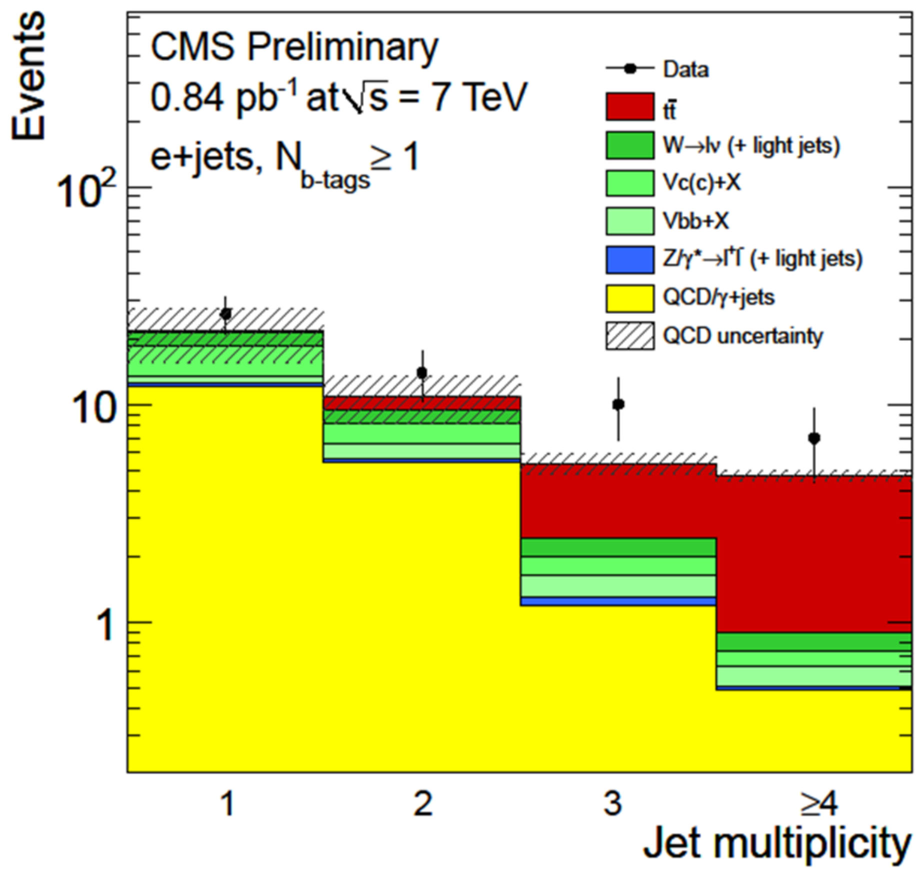


mu+Jets, >=2 jets

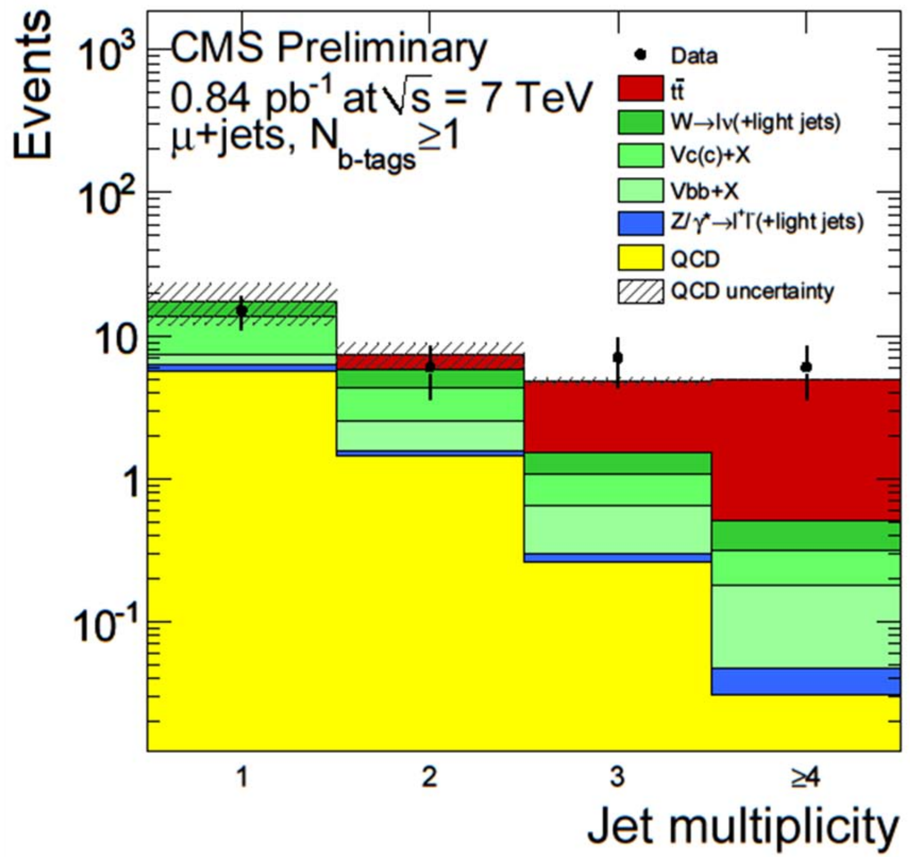


e/mu+jets+b-tag

- e+jets

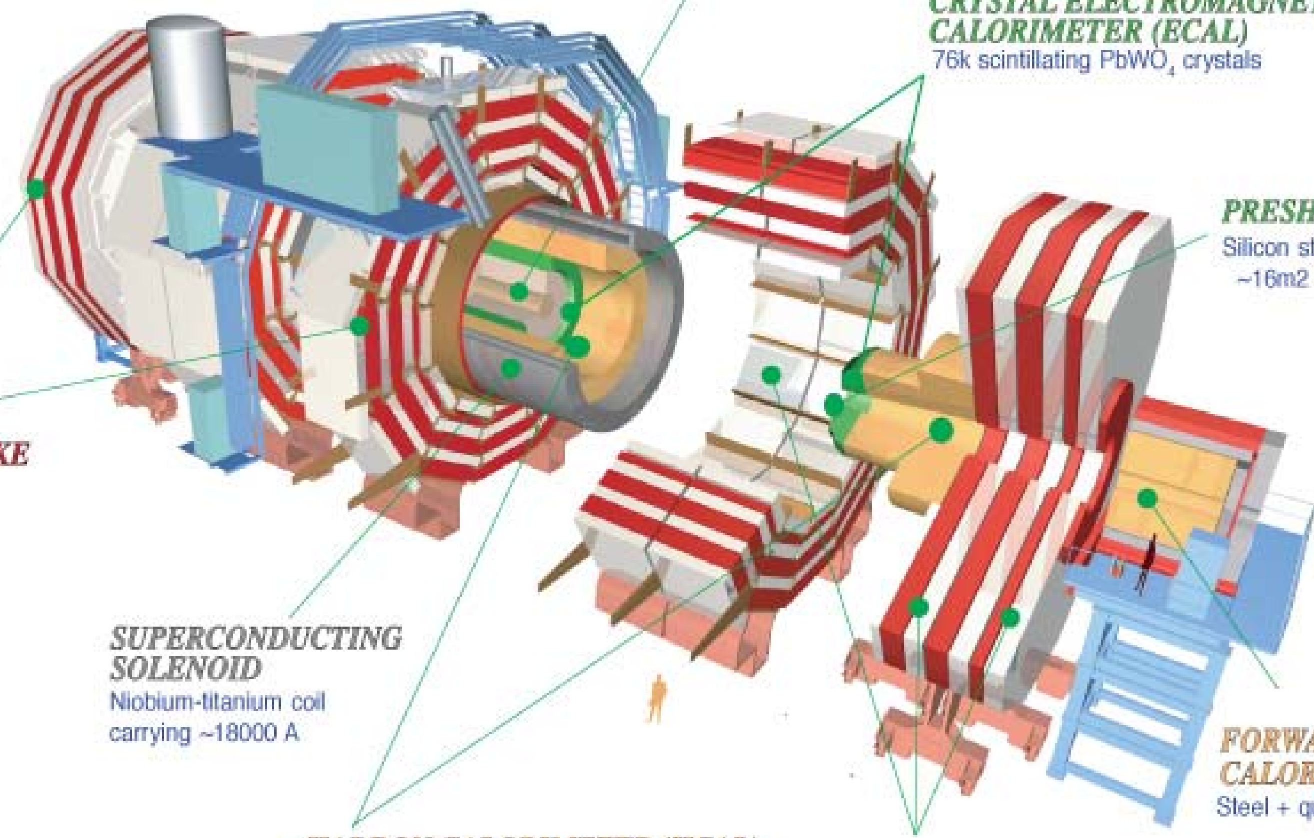


- mu+jets



CMS Detector

Pixels
Tracker
ECAL
HCAL
Solenoid
Steel Yoke
Muons



Total weight	**: 14000 tonnes**
Overall diameter	**: 15.0 m**
Overall length	**: 28.7 m**
Magnetic field	**: 3.8 T**