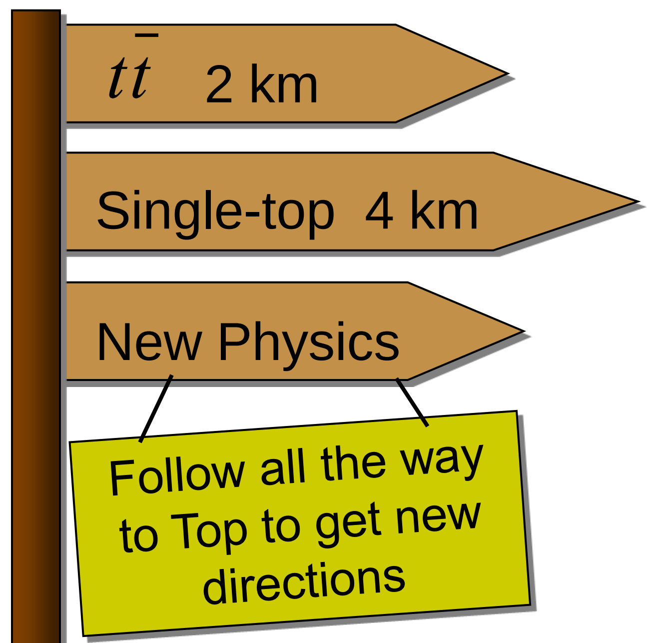
Top-Quark Studies at CMS

Tim Christiansen (CERN)
on behalf of the CMS Collaboration

ICHEP 2010, Paris

35th International Conference on High-Energy Physics

22–28 July 2010

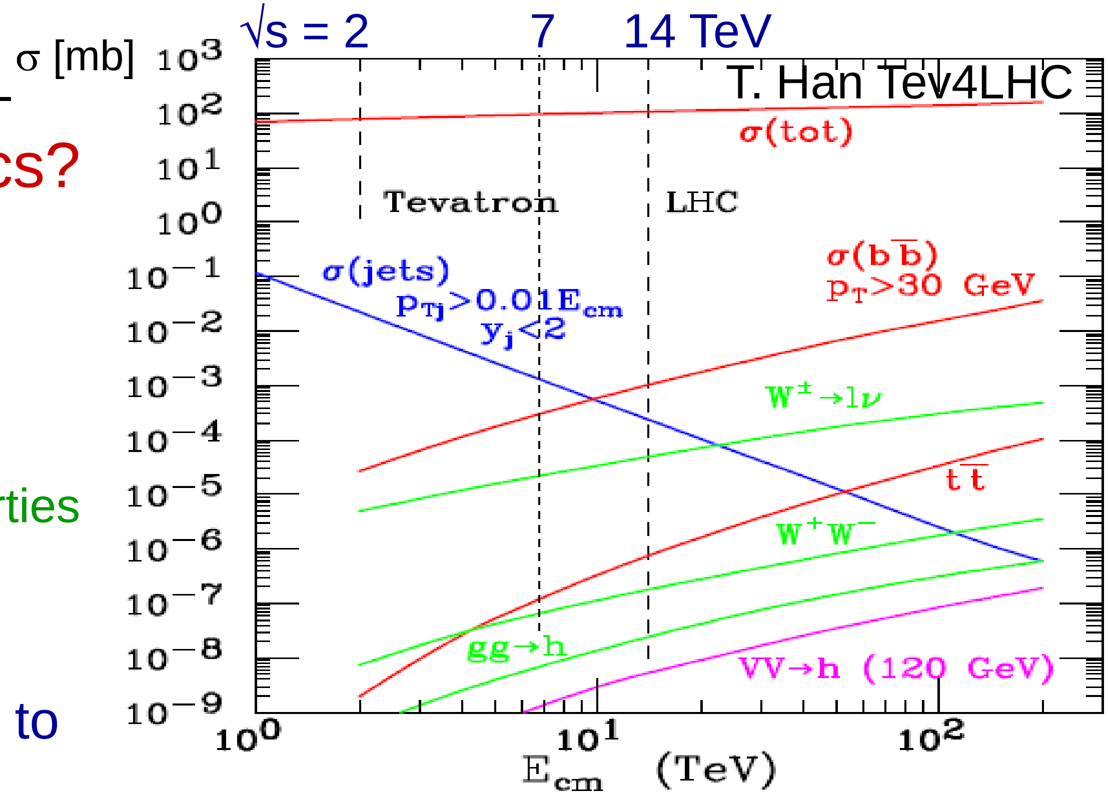


Outline

- Introduction
- Examining the top-quark selection on first 78 nb^{-1} of pp collision data at $\sqrt{s}$ = 7 TeV
  - In the dilepton and lepton+jets channels (l=e,μ)
  - Aimed (eventually) at first cross-section measurements
  - Not enough data to see any top-quark signal, but a good test of CMS's tools
    - event reconstruction in first data, comparison with simulation
    - testing of data-driven methods for background estimation
- Latest news from the most recent data
  - What we see in 0.25 pb^{-1} …
- Summary

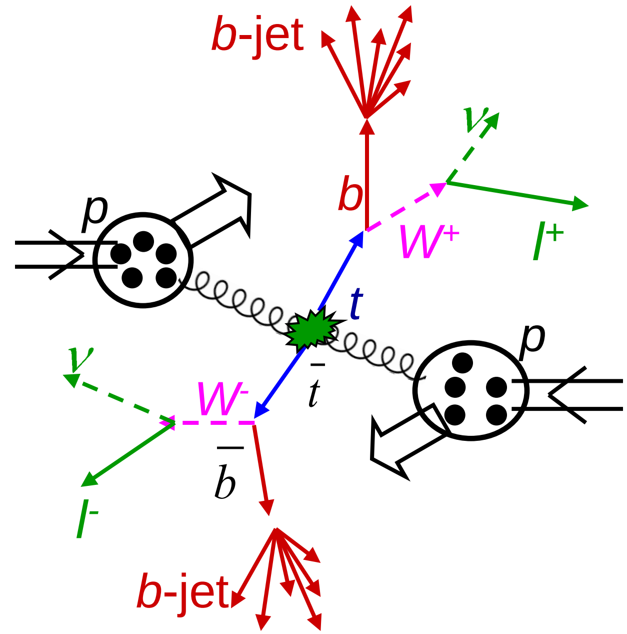
"The Truth (Quark) is ~~rarely~~ pure and never simple."
(Oscar Wilde)

Introduction



- Why top-quark physics?
  - Test of not-so-well explored area of the standard model
    - σ, couplings, rare decays, prod. properties
  - The most exotic of all known elem. particles
  - Important background to many searches, etc …
- Contains/tests ~all major features of the reconstruction
  - leptons of all kinds, light and heavy jets (*b*-tagging), missing transverse energy (MET), … + an on-shell W boson in its decay
  - Excellent tool for the calibration of the experiment
    - jet-energy scale for light & heavy quarks, b-tagging efficiency, MET

Dilepton+X Selection



- Dilepton channels: *ee*, *μμ*, *eμ*
  - Triggers: μ+X (p_T > 9 GeV/c) or e/γ+X (E_T > 15 GeV)
  - 2 isolated, prompt, oppositely charged leptons (l = e,μ) of good quality
    - $p_T(l) > 20$ GeV/c
    - $|\eta_\mu| < 2.5$, $|\eta_e| < 2.4$
    - Relative isolation:

Detected energy around lepton

$$\text{Rel.isol.} = \frac{\overbrace{\sum_{R<0.3} p_T^{\text{track}} + \sum_{R<0.3} p_T^{\text{ECAL}} + \sum_{R<0.3} p_T^{\text{HCAL}}}^{\text{Detected energy around lepton}}}{p_T(\text{lepton})} < 15\ \%$$

  - Missing transverse energy (MET)
    - using calorimeter⊕tracking
    - MET > 30 (20) GeV (in *eμ*+*X*)

- *Z*-boson veto:
  - 76 < $M_{ee,\mu\mu}$ < 106 GeV/c^2
- Count additional jets:
  - anti-k_T jets, *R* = 0.5
  - using calorimeter⊕tracking info
  - $|\eta| < 2.4$, $p_T > 30$ GeV/c

≥ 2 jets typical for ttbar

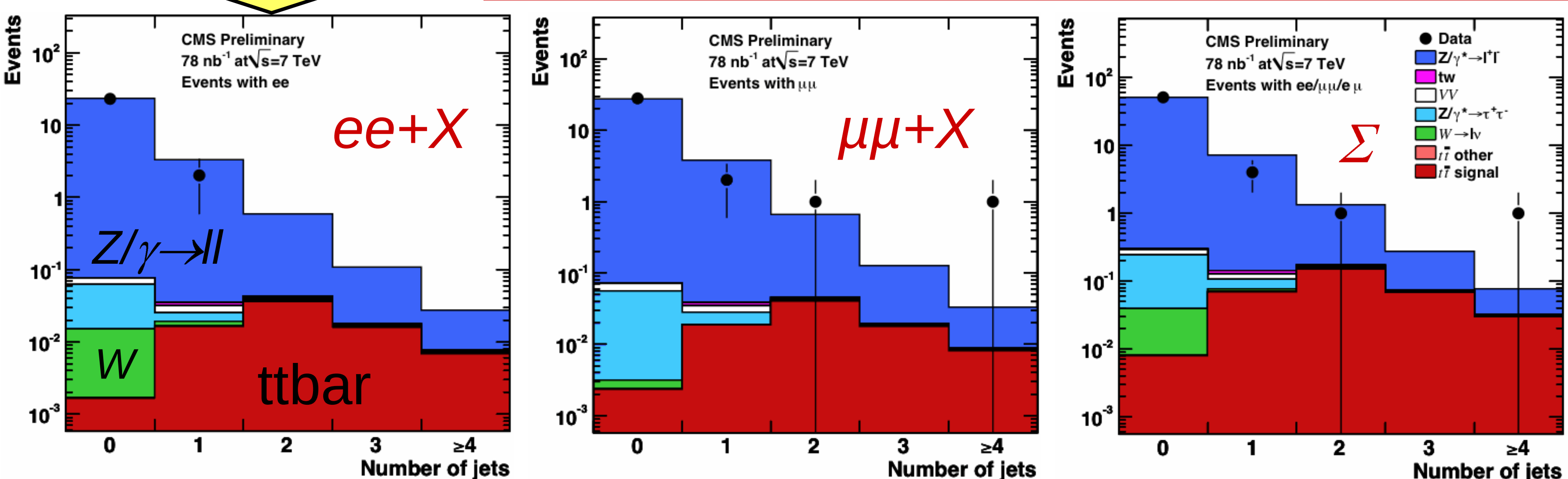
Dilepton Channel in 78 nb^{-1}

Insufficient data for signal just yet ⇒ background studies in relaxed selection & test of data-driven methods for background estimation

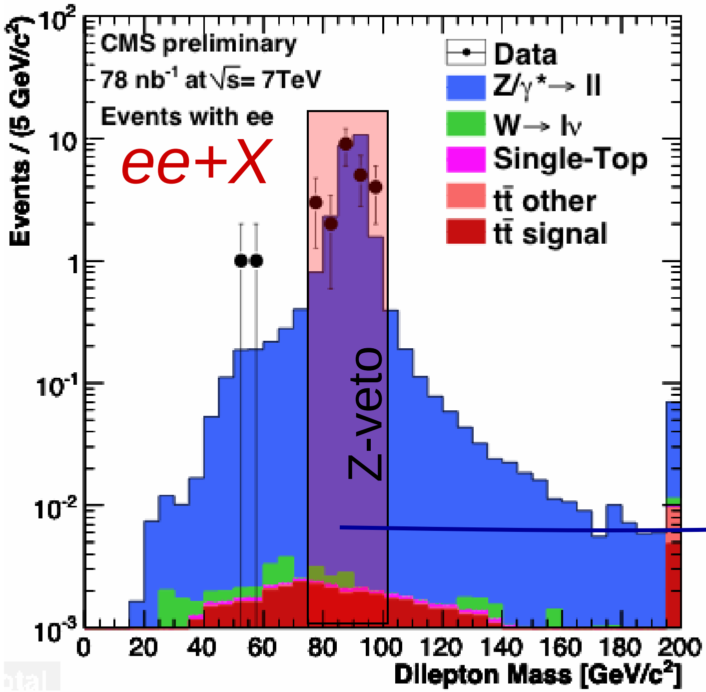
Z-veto, $N_{jet} \geq 1$, no MET cut

Sample	ee	$\mu\mu$	$e\mu$
Dilepton $t\bar{t}$	0.058 ± 0.015±0.010	0.065 ±0.010±0.011	0.157 ±0.031±0.025
Dibosons - VV	0.0045 ± 0.0023	0.0043 ±0.0022	0.011 ±0.006
Single top - tW	0.0040 ± 0.0020	0.0043 ±0.0022	0.011 ±0.005
Drell-Yan $\tau\tau$	0.008 ± 0.004	0.012 ±0.006	0.016 ±0.008
Drell-Yan $ee, \mu\mu$	0.41 ± 0.13	0.48 ±0.12	0.003 ±0.001
Non-dilepton $t\bar{t}$	0.002 ± 0.001	0.0003 ±0.0001	0.003 ±0.001
W+jets	0.003 ± 0.001	$0.000^{+0.001}_{-0.000}$	0.005 ±0.002
QCD multijets	$0.0^{+1.2}_{-0.0}$	$0.0^{+1.2}_{-0.0}$	$0.0^{+1.2}_{-0.0}$
Total simulation	$0.45^{+1.2}_{-0.11}$	$0.52^{+1.2}_{-0.11}$	$0.21^{+1.2}_{-0.03}$
Data	0	1	0

Relaxed selection: w/o Z-veto, no MET cut



Background Estimation in Dileptons



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

ratio outside/inside from *DY* simulation

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

correction for non-DY contribution in *Z*-veto region from $e\mu$ sample

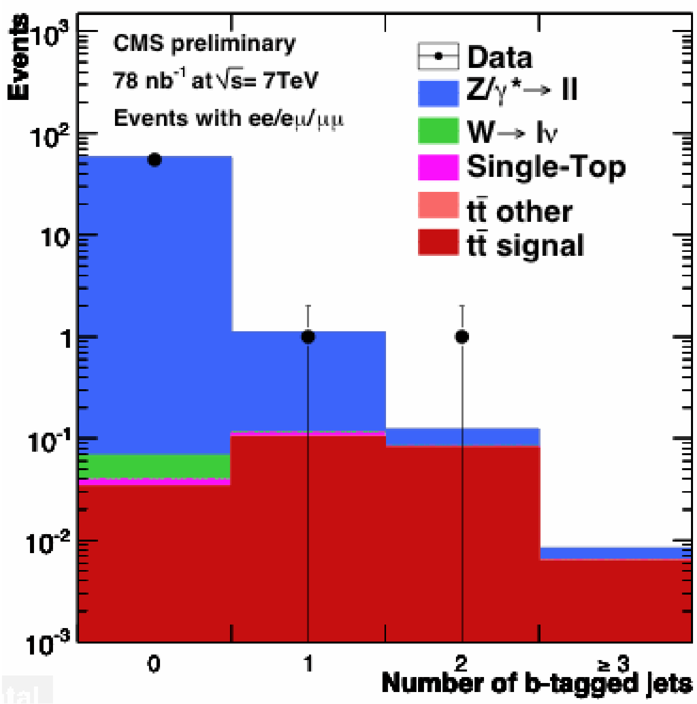
Excellent agreement on loosened selection!

Sample	ID, ISO, Z-veto	with $N_{\text{jet}} \geq 1$	with $\not\!\!E_T$
ee			
DY in simulation	2.4 ± 0.7	0.41 ± 0.13	0.008 ± 0.002
DY estimate in data	$2.2 \pm 0.5 \pm 1.1$	$0.23 \pm 0.16 \pm 0.12$	$0.00^{+0.16}_{-0.00}\,^{+0.08}_{-0.00}$
$\mu\mu$			
DY in simulation	2.9 ± 0.7	0.48 ± 0.12	0.019 ± 0.005
DY estimate in data	$2.9 \pm 0.6 \pm 1.5$	$0.36 \pm 0.21 \pm 0.18$	$0.00^{+0.25}_{-0.00}\,^{+0.12}_{-0.00}$

Also: QCD multijet & *W*+*X* contributions with additional fake prompt and isolated lepton(s) is estimated my measuring "fake rates" in data

Bottom-Jet Identification

- Relaxed selection: Simulation describes the data very well!
- Jets from *b*-quarks are a telling feature of top-quark decays
  - Not (yet) applied in the selection, but we start to observe the heavy-quark content of the event sample
  - Example: Loose b-tagger using the **impact parameter significance (IPsig)** of the tracks associated with the jet (loose cut ≡ ~10% fake rate/jet)



Simple and robust tagger: IPsig of track with 2nd-highest IPsig in the jet > loose threshold

Chosen here, as it allows to see any tagged jets in the selected sample already

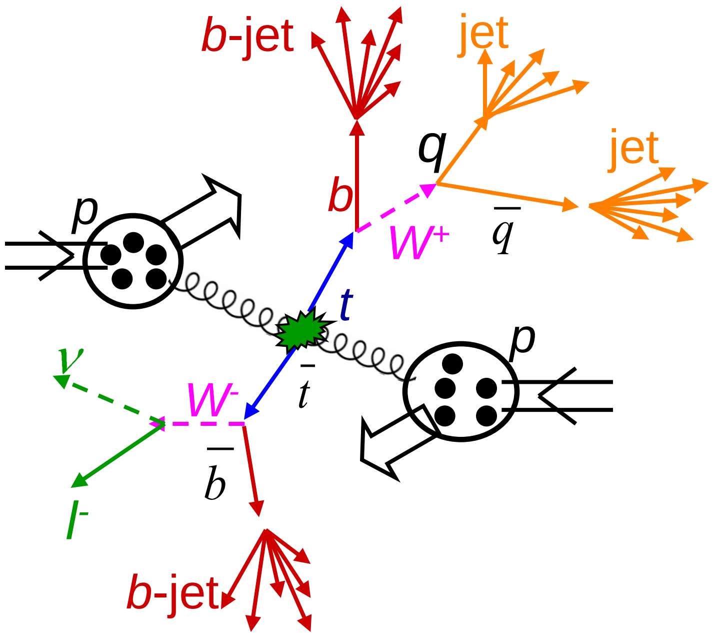
To be tightened as luminosity increases to improve purity of b-tagged jets

Relaxed selection here: without Z-veto, no MET requirement

Lepton+Jets Selection (simplified)

- Channels: *e*+jets, *μ*+jets
  - Ask for exactly 1 prompt, isolated electron (muon) of good quality
  - Very similar selection of *e*,*μ* as before, but
    - tightened ID requirements and isolation:
      - Rel.isol. < 10%(*e*), 5%(*μ*)

      due to larger backgrounds
    - p_T(e) > 30 GeV/c
    - $p_T(\mu)$ > 20 GeV/c, $|\eta_\mu| < 2.1$
  - Do not apply (yet) any requirement of significant missing transverse energy (MET)



- Count additional jets
  - anti-k_T jets, R = 0.5
  - using calorimeter info
  - $|\eta| < 2.4$, p_T > 30 GeV/c

≥ 4 jets is typical for ttbar

L+Jets Channel in 78 nb^{-1}

78 ± 9 nb^{-1}

e+jets channel, L = 78 nb^{-1}

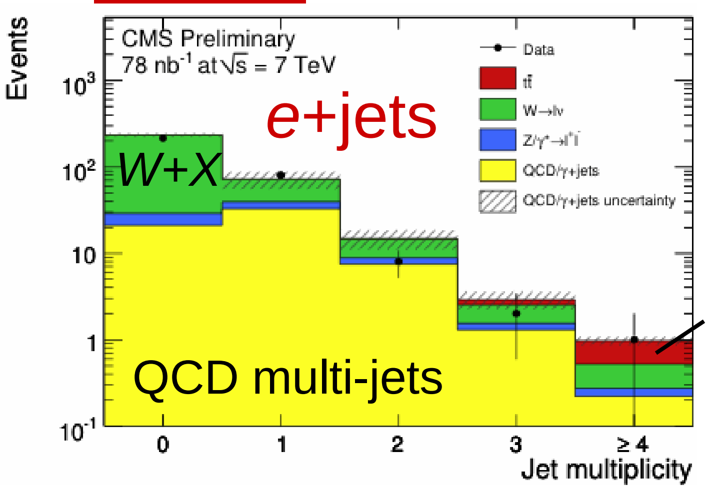
Jets	$t\bar{t}$	single top	W+Jets	Z+Jets	QCD	Sum MC	Data
≥ 0	1.1 ± 0.2	0.32 ± 0.04	244 ± 29	17 ± 2	63 ± 7	325 ± 30	306
≥ 1	1.1 ± 0.2	0.29 ± 0.03	39 ± 7	8.4 ± 1.1	42 ± 6	91 ± 10	91
≥ 2	1.1 ± 0.2	0.17 ± 0.01	7.0 ± 1.7	1.7 ± 0.4	9.0 ± 2.3	19 ± 3	11
≥ 3	0.83 ± 0.17	0.07 ± 0.01	1.2 ± 0.4	0.30 ± 0.09	1.5 ± 0.5	3.9 ± 0.7	3
≥ 4	0.45 ± 0.11	0.02 ± 0.01	0.25 ± 0.10	0.05 ± 0.02	0.22 ± 0.10	0.99 ± 0.18	1

μ+jets channel, L = 78 nb^{-1}

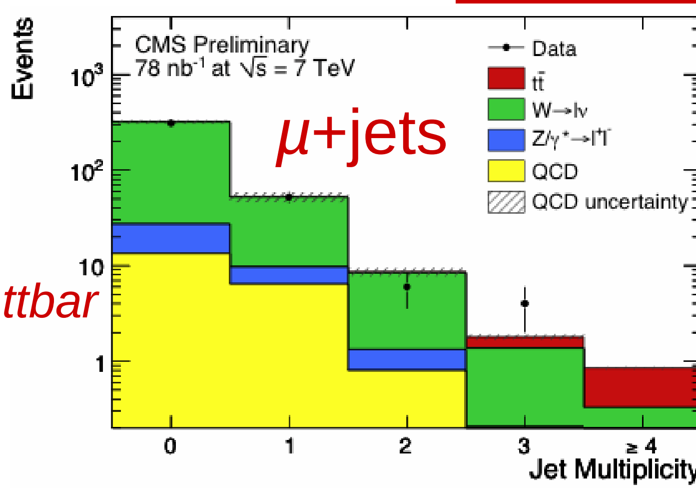
Jets	$t\bar{t}$	single top	W+Jets	Z+Jets	QCD	Sum MC	Data
≥ 0	1.3 ± 0.2	0.39 ± 0.04	344 ± 42	18 ± 3	21 ± 2	385 ± 42	371
≥ 1	1.2 ± 0.2	0.36 ± 0.04	51 ± 10	3.9 ± 1.1	7.3 ± 1.6	64 ± 10	62
≥ 2	1.2 ± 0.2	0.21 ± 0.02	8.5 ± 2.2	0.66 ± 0.41	0.92 ± 0.27	12 ± 2	10
≥ 3	0.95 ± 0.19	0.08 ± 0.01	1.5 ± 0.5	0.12 ± 0.08	0.12 ± 0.05	2.7 ± 0.6	4
≥ 4	0.52 ± 0.13	0.02 ± 0.01	0.29 ± 0.11	0.02 ± 0.02	0.01 ± 0.01	1.0 ± 0.1	0

No signal in 78 nb^{-1} yet

⇒ BG studies in low jet multiplicity bins



ttbar



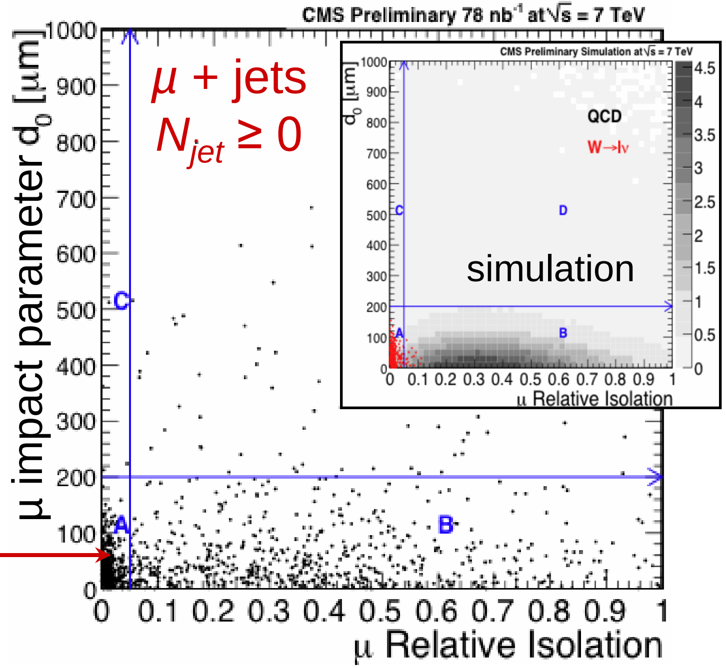
Test data-driven estimation of BG contribution

Background Estimation in L+Jets

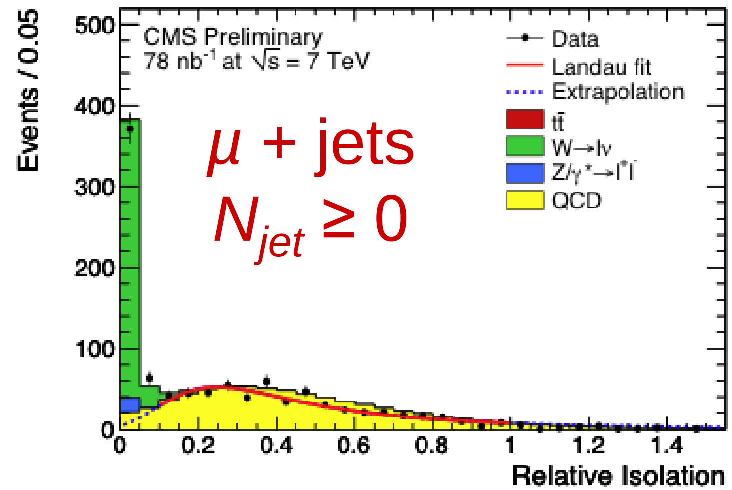
Estimate of backgrounds with fake isolated and prompt leptons from data:

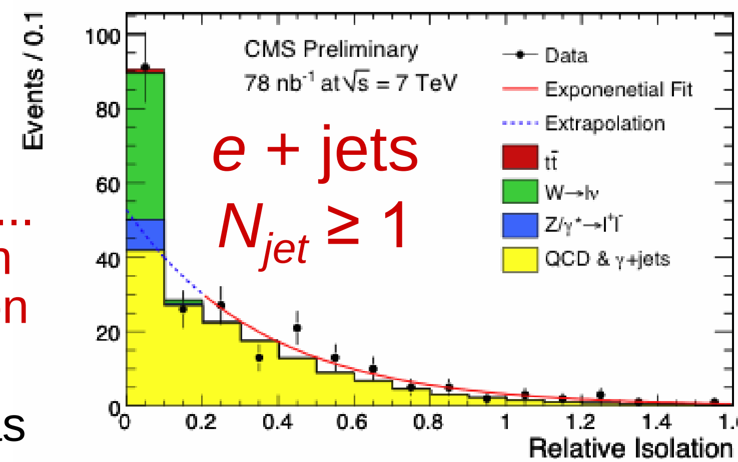
"ABCD" method with 2 ~ uncorrel. variables

signal region A



Fit & extrapolation to signal region:





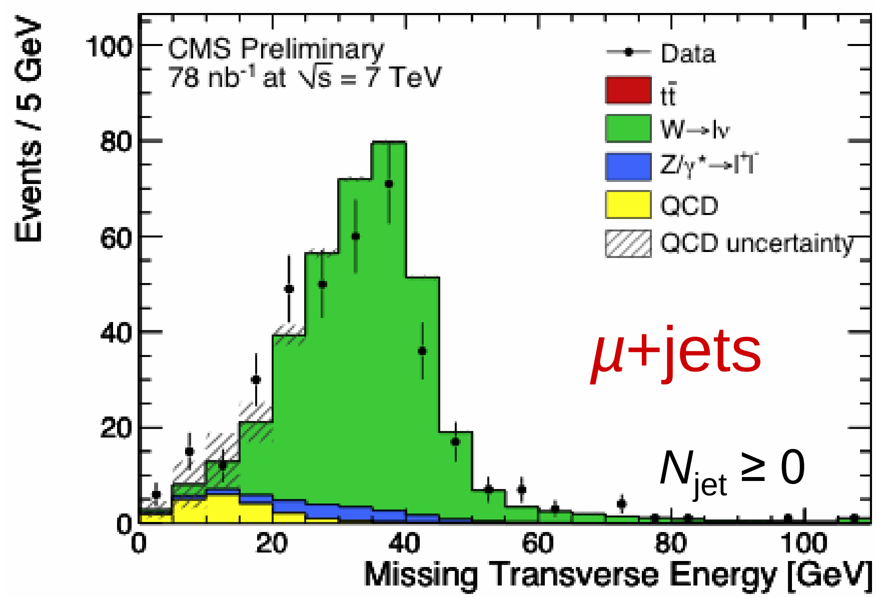
μ + jets	$N_{jet} \geq 0$	$N_{jet} \geq 1$
ABCD	57 ± 29	30 ± 15
Rel. isol. fit	7 ± 4	6 ± 3
Simulation	21 ± 2	7 ± 2

} disagreement currently u.i. with more data... Now, we assign 100% uncert. on N_{QCD} in μ+jets

e + jets	$N_{jet} \geq 0$	$N_{jet} \geq 1$
Rel. isol. fit	70 ± 35	44 ± 22
Simulation	63 ± 7	42 ± 6

(100% is OK, as BG small in $N_j \geq 4$)

L+Jets Channel in 78 nb^{-1}



Missing transverse energy:

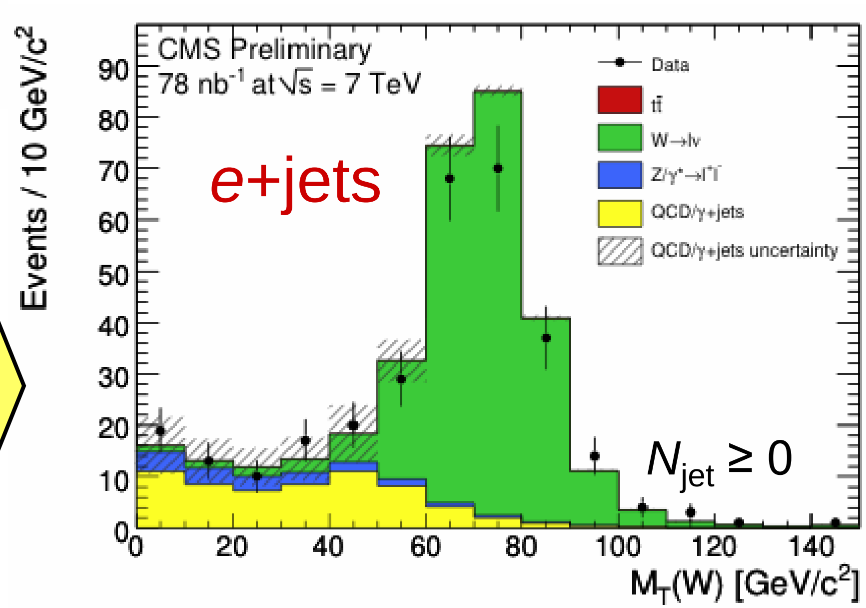
hard to get right, but important ingredient of any top-quark analysis in the future!

Currently dominated by *W*+*X* without jet requirement

Transverse mass $M_T(W)$:

sensitive to the W-boson mass in the event

calculated from lepton p_T & missing transverse energy



(In the figures: QCD BG is taken from simulation, with 50 /100 % uncertainty assigned for *μ*/*e*+jets)

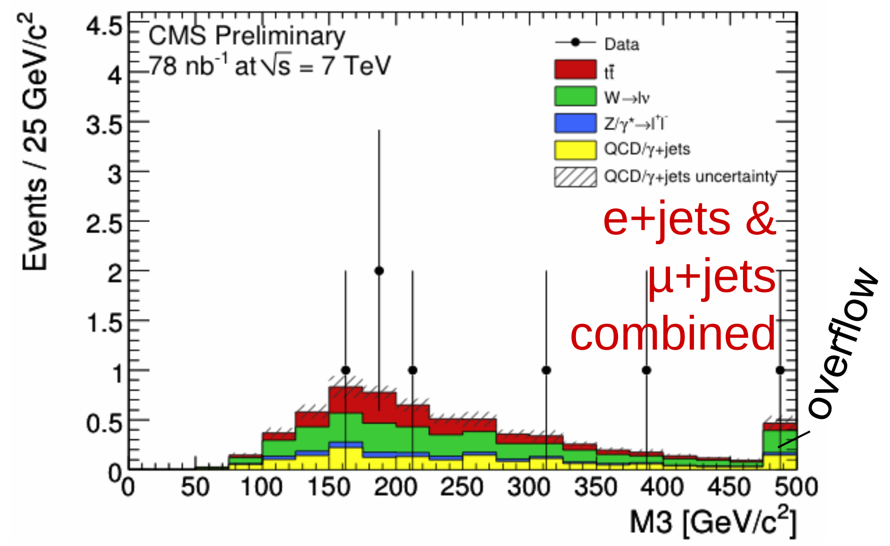
Telling L+Jets Distributions

M3: M_{jjj} of the three-jet system

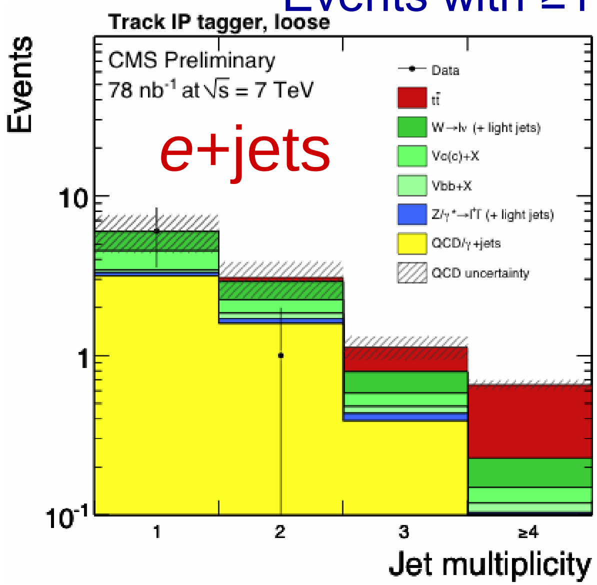
that maximizes $\left|\sum_{jjj} \vec{p}_T(\text{jet})\right|$

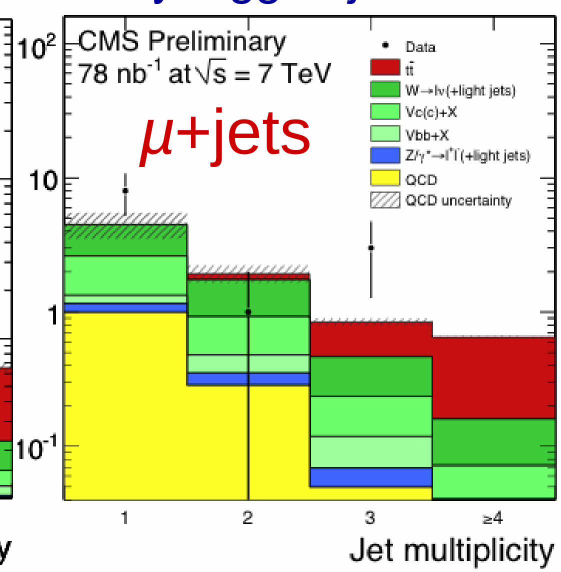
Jets with displaced vertices and/or soft muons inside will be a hint for bottom-jets in the event → typical for ttbar!

(here: same b-tagger as for dileptons)

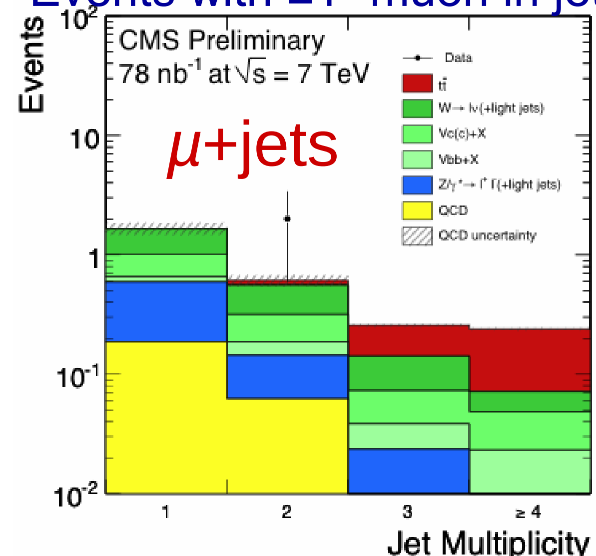


Events with ≥1 loosely tagged jet:





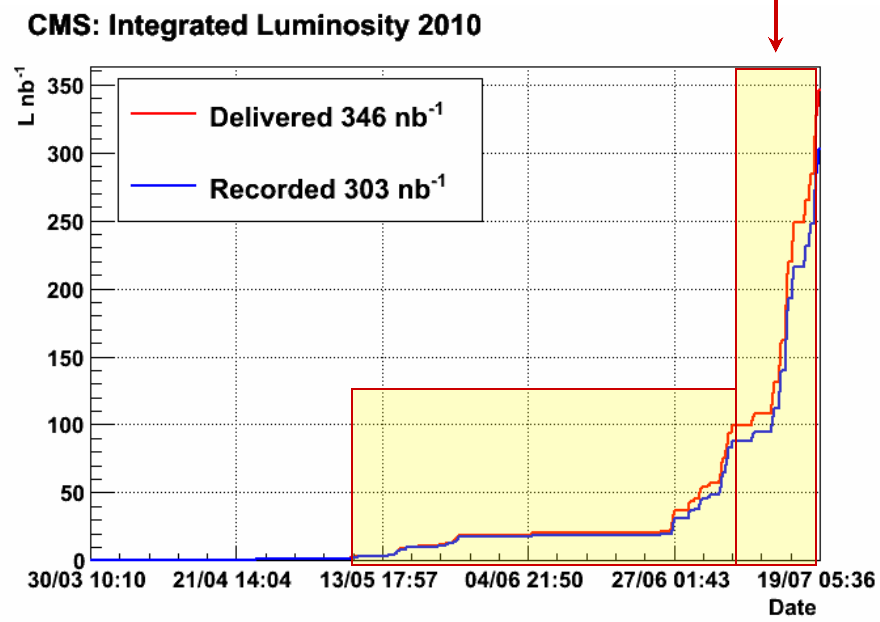
Events with ≥1 "muon in jet"



So do we see the Top-Quark?

Not in the 78 nb^{-1} of data presented so far

(in agreement with the expectation/predictions)



However … there a few interesting events that pass our selection in the more recent data:
Some of the first top-quark candidates observed outside Fermilab!

μ+Jets Candidate Event *(from July 14!)*

Event passes all cuts of full selection

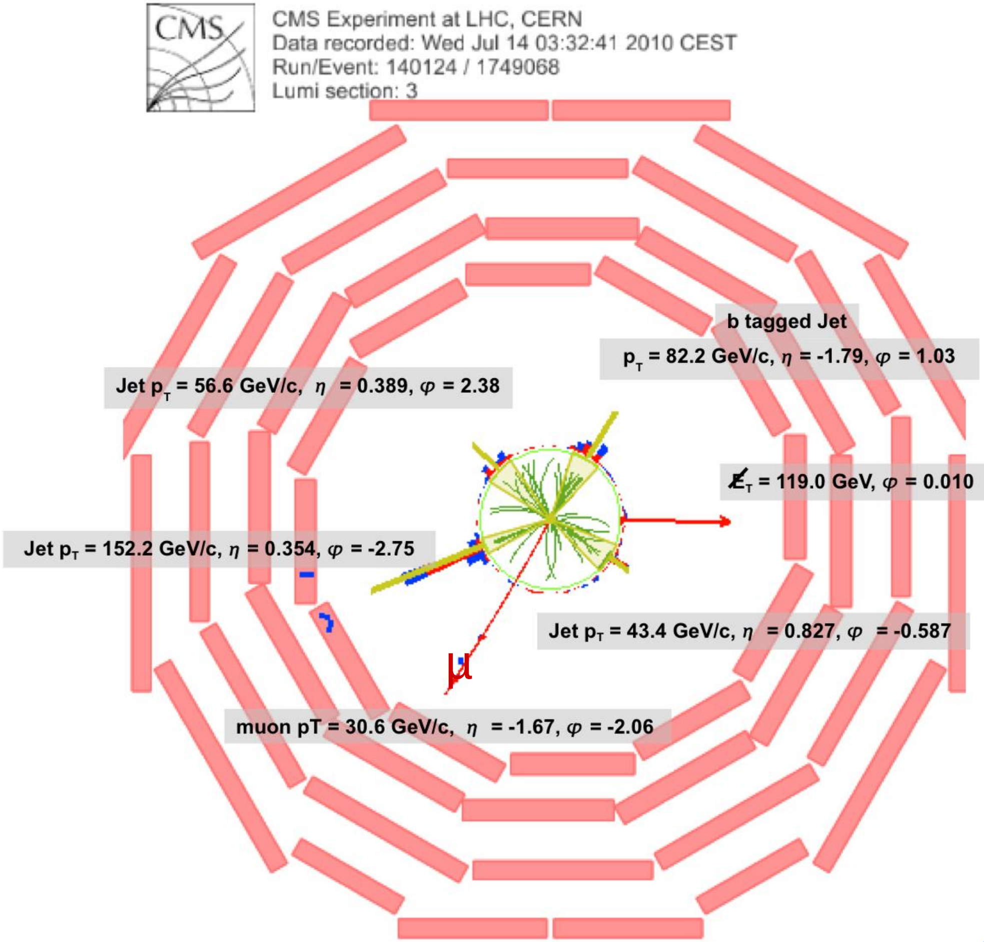
1 high-momentum muon

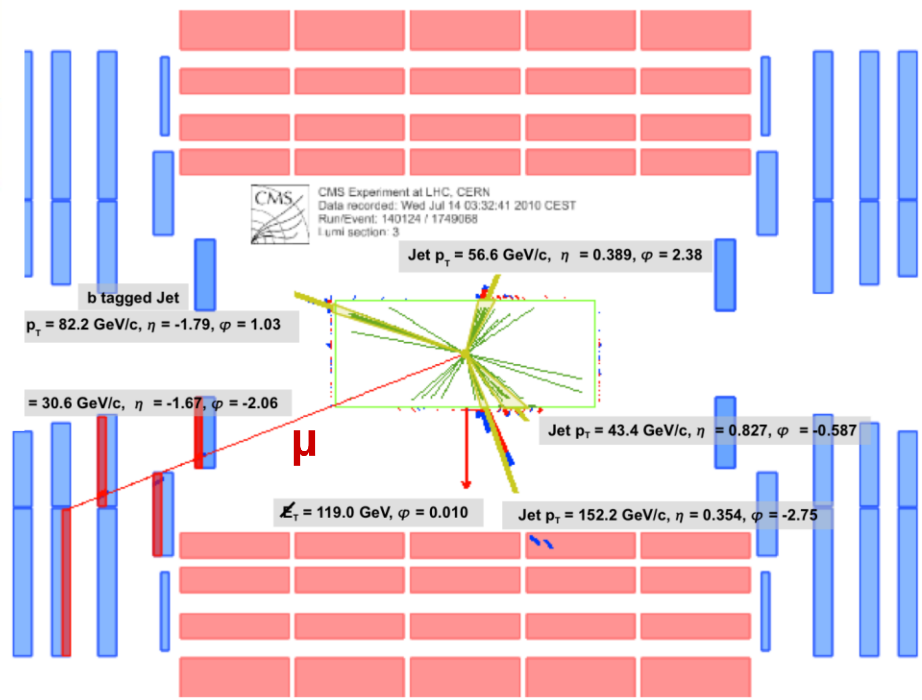
significant MET > 100

$m_T(W)$ = 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

e+Jets Candidate Event *(from July 18)*

Event passes all cuts:

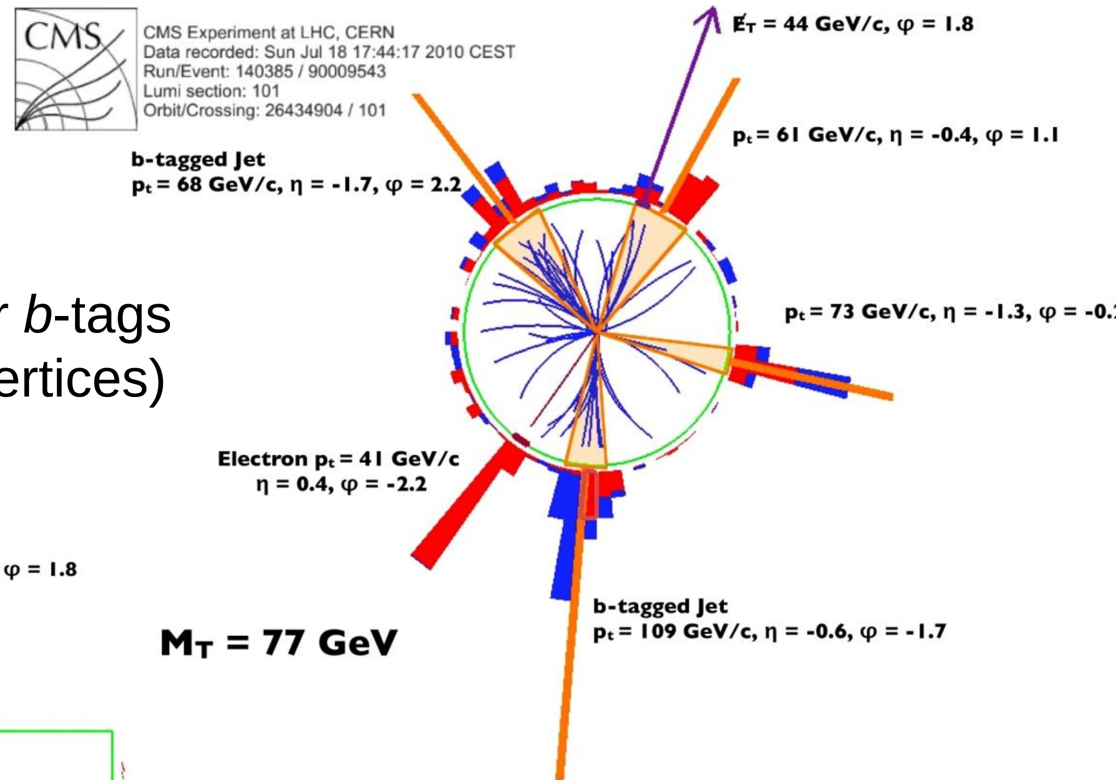
1 high-momentum electron

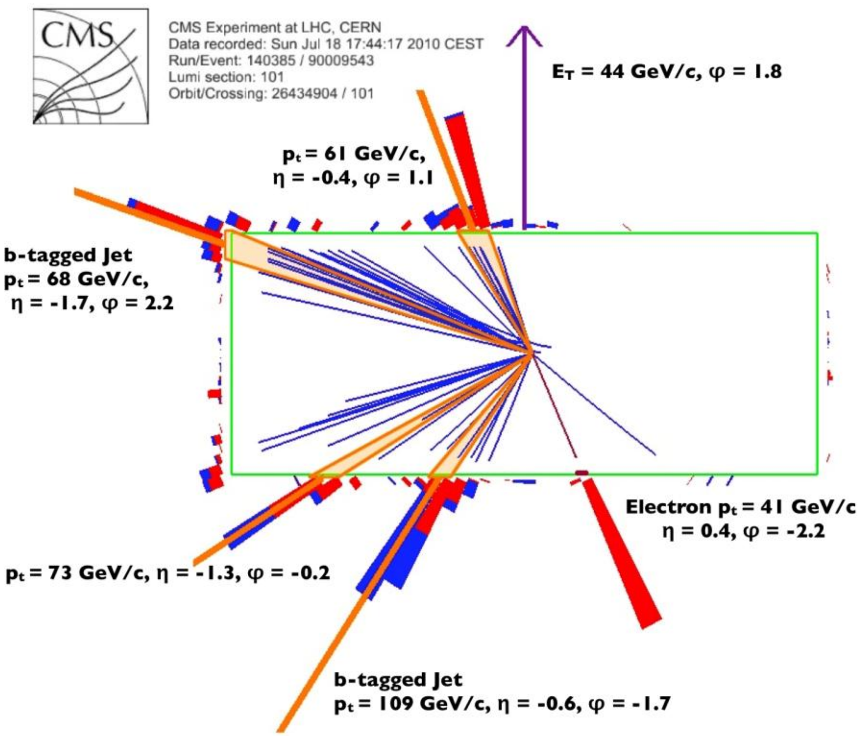
significant MET ≈ 44 GeV

4 high-p_T jets,

two of which with good/clear *b*-tags

(with reconstructed 2nd ary vertices)



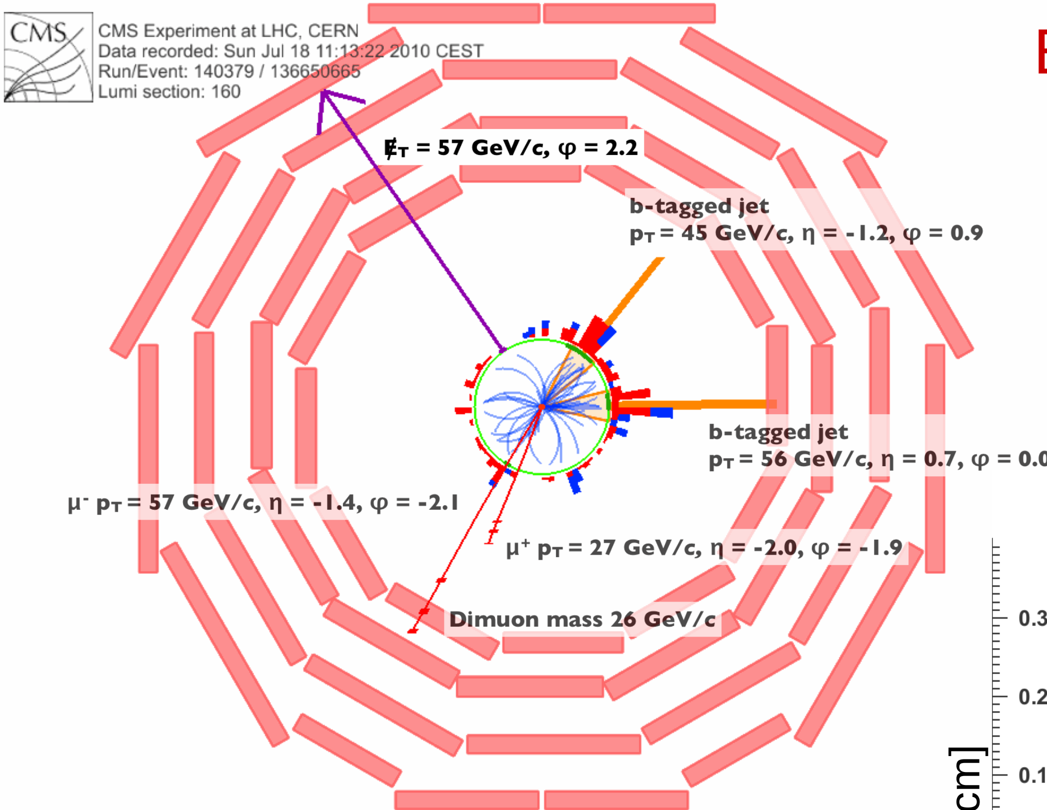


$m_T(W) \approx 77$ GeV/c^2

Mass of 2 untagged jets ≈ 102 GeV/c^2

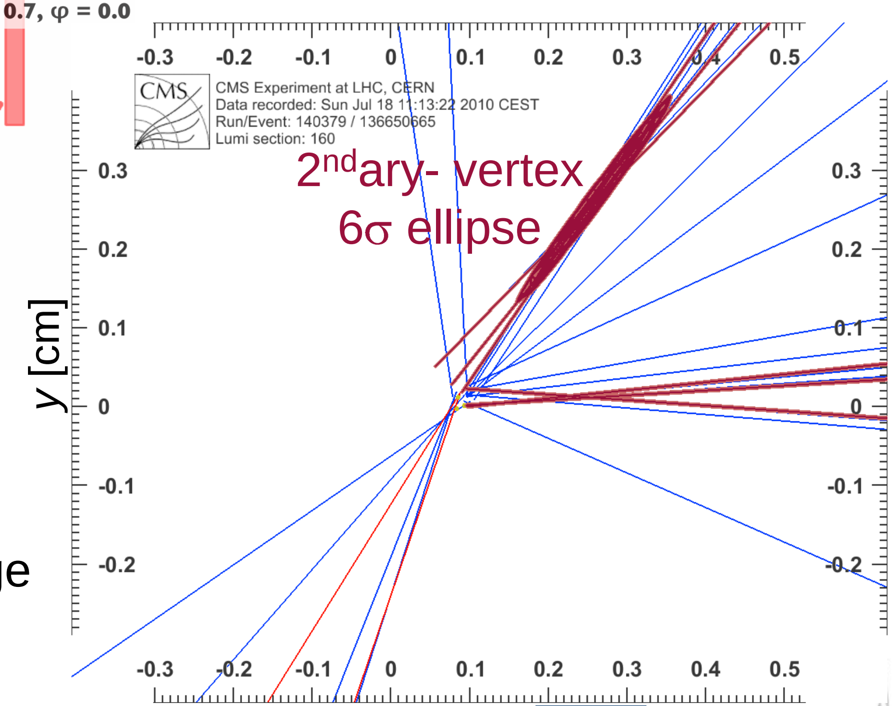
m(jjj) ≈ 208, 232 GeV/c^2
(for the two 3-jet combinations)

μμ +Jets Candidate Event (from July 18)



Event passes all cuts of full selection:

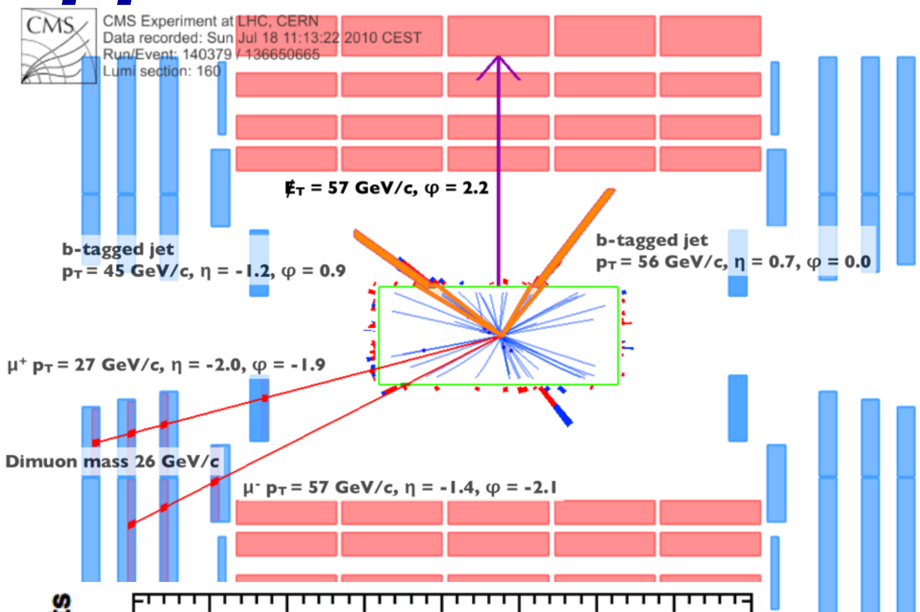
2 muons with opposite charge
2 jets, both w/ good/clear *b*-tags
(and secondary vertices!)
significant MET (>50 GeV)



$m(\mu\mu) = 26\ GeV/c^2$

Preliminarily reconstr. mass is in the range 160–220 GeV/c^2 (consistent with m_{top})

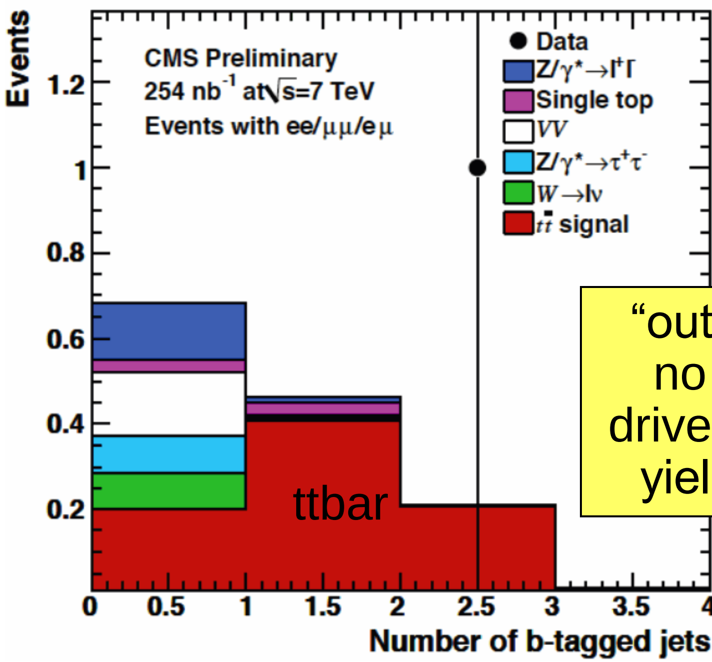
μμ +Jets Candidate … cont'd



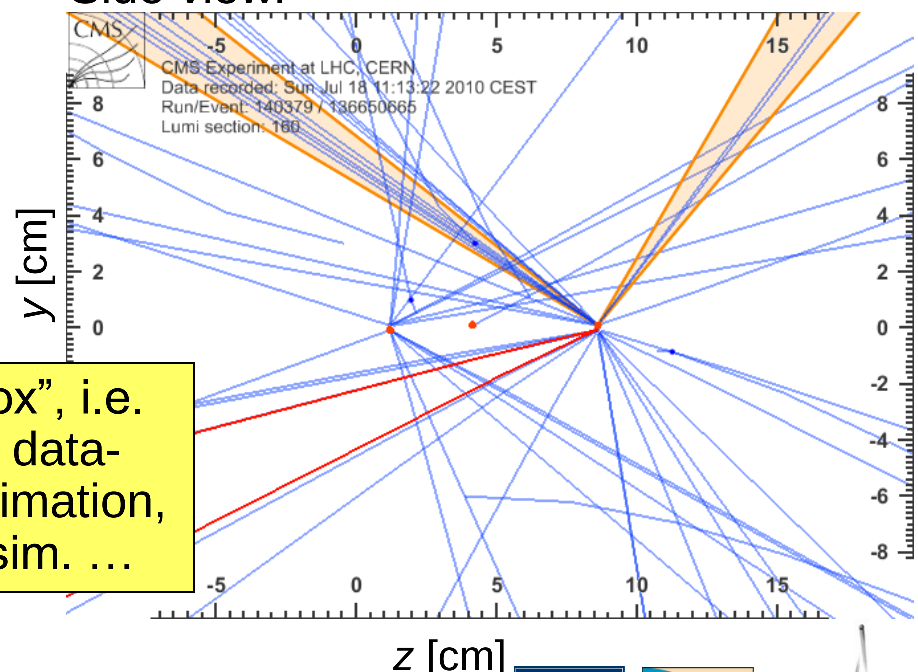
Very clean candidate with little expected background!

Multiple primary vertices → multiple *pp* collisions ("pile-up")

Jets & muons originate from same primary vertex



Side view:

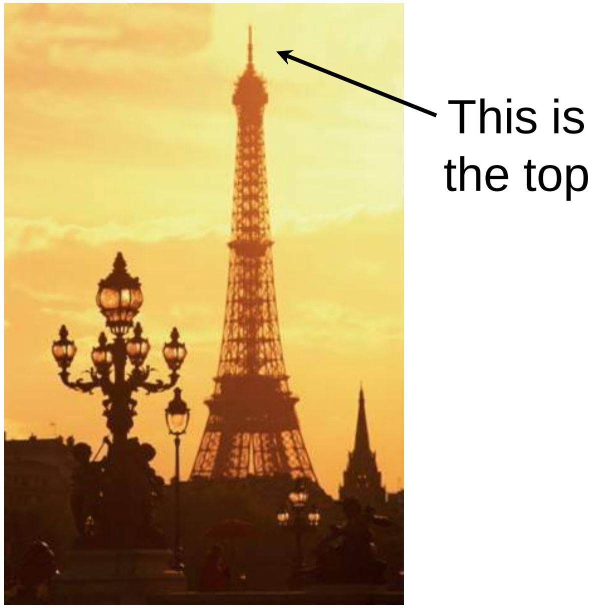


"out of the box", i.e. no syst., no data-driven BG estimation, yields from sim. …

Summary

- 78 ± 9 nb^{-1} of *pp* collisions at 7 TeV fully analyzed
  - Fast turn-around time (data is only ≥2 weeks old)
  - No signal, as expected
  - Tools are ready for first measurements:
    - Simulation describes the data very well
    - Successfully tested background estimations
- Currently establishing first top-quark signals
  - Exciting top candidates in the most recent data in 3 channels in 0.25 pb^{-1}!

- Very, very exciting times are ahead!
  - we've been saying this for years, but it's nothing but the Truth (quark)



The signal is starting to rise from the background

Thanks!

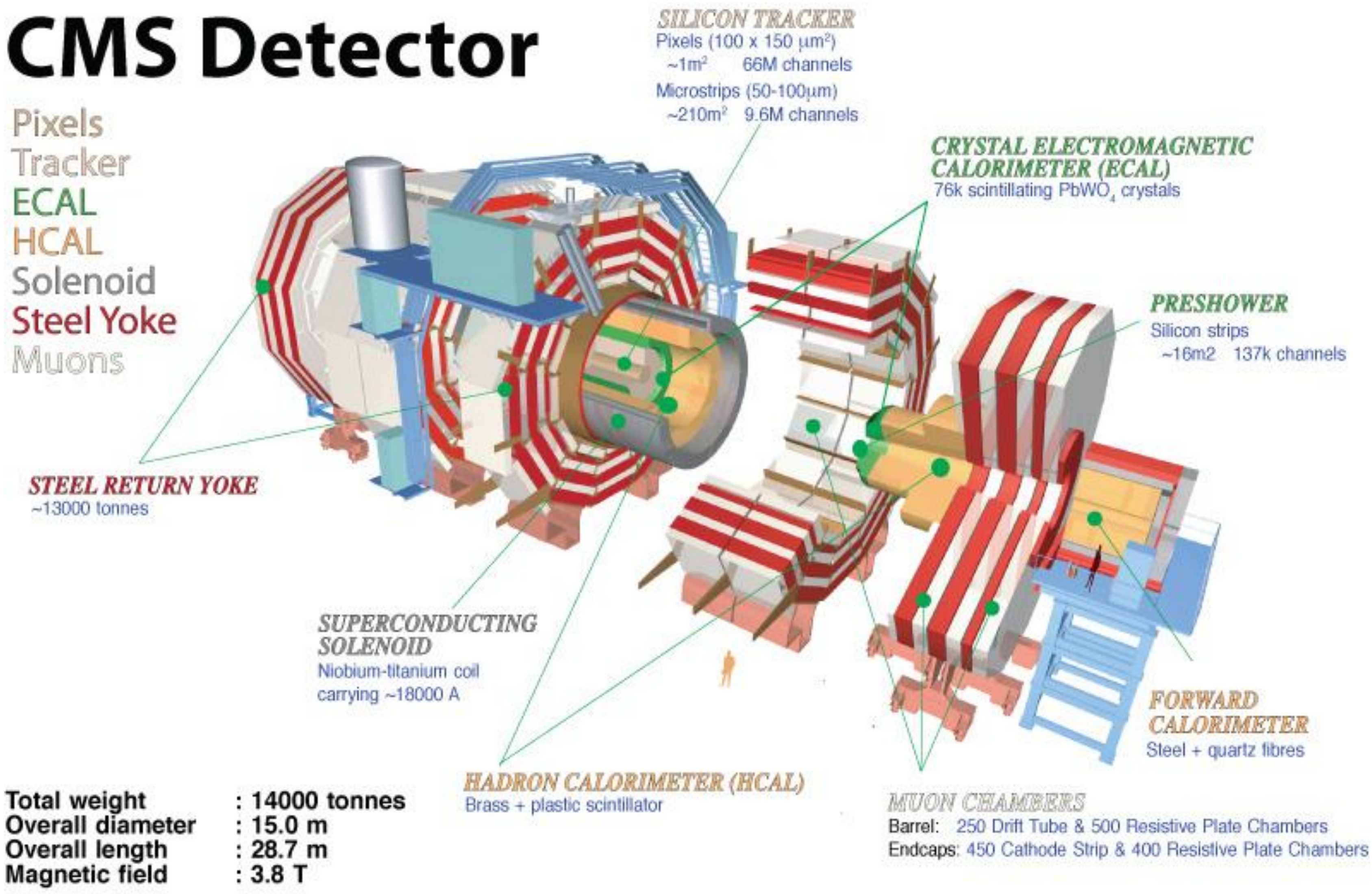
To you for the attention, to the organizing committees of ICHEP, and to the LHC, the CMS collaboration and all the analysts for their dedicated work to make first top-quark physics in *pp* collisions possible!

Reference:

"Selection of Top-Like Events in the Dilepton & Lepton+Jets Channels in Early 7 TeV Data," CMS Physics Analysis Summary, CMS-PAS-TOP-10-004 (2010)

Backup Slides …

Compact Muon Solenoid

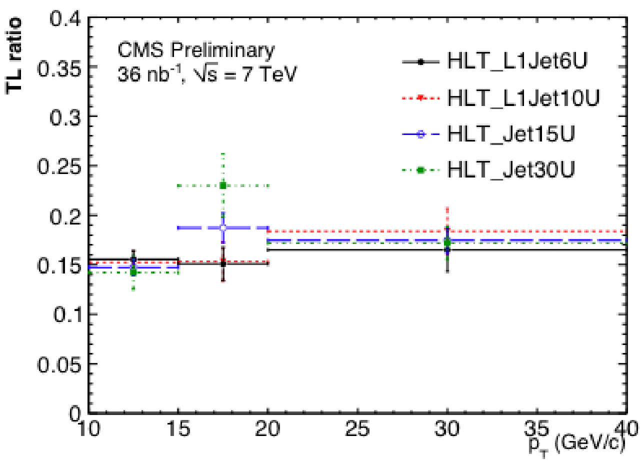


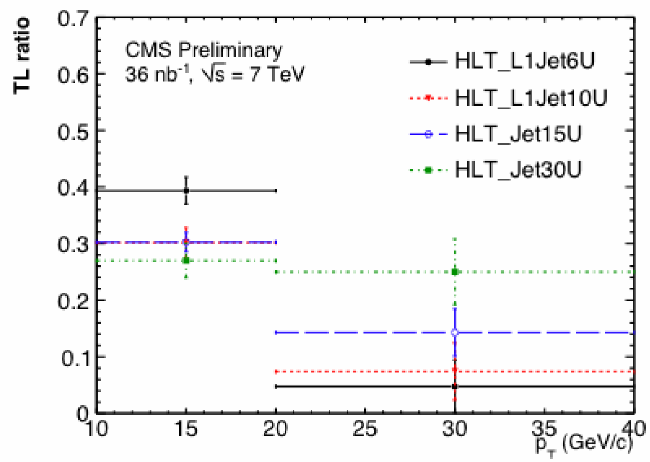
Dilepton Channel: Fake Rate & Yields

Sample	ee	$\mu\mu$	$e\mu$
Dilepton $t\bar{t}$	$0.058 \pm 0.015 \pm 0.010$	$0.065 \pm 0.010 \pm 0.011$	$0.157 \pm 0.031 \pm 0.025$
Dibosons - VV	0.0045 ± 0.0023	0.0043 ± 0.0022	0.011 ± 0.006
Single top - tW	0.0040 ± 0.0020	0.0043 ± 0.0022	0.011 ± 0.005
Drell-Yan $\tau\tau$	0.008 ± 0.004	0.012 ± 0.006	0.016 ± 0.008
Drell-Yan ee, $\mu\mu$	0.41 ± 0.13	0.48 ± 0.12	0.003 ± 0.001
Non-dilepton $t\bar{t}$	0.002 ± 0.001	0.0003 ± 0.0001	0.003 ± 0.001
W+jets	0.003 ± 0.001	$0.000\,^{+0.001}_{-0.000}$	0.005 ± 0.002
QCD multijets	$0.0\,^{+1.2}_{-0.0}$	$0.0\,^{+1.2}_{-0.0}$	$0.0\,^{+1.2}_{-0.0}$
Total simulation	$0.45\,^{+1.2}_{-0.11}$	$0.52\,^{+1.2}_{-0.11}$	$0.21\,^{+1.2}_{-0.03}$
QCD data-driven	$0.00\,^{+0.06}_{-0.00}\,^{+0.06}_{-0.00}$	$0.0\,^{+0.2}_{-0.0}\,^{+0.2}_{-0.0}$	$0.0\,^{+0.1}_{-0.0}\,^{+0.1}_{-0.0}$
W+jets data-driven	$0.0\,^{+0.2}_{-0.0}\,^{+0.1}_{-0.0}$	$0.0\,^{+0.4}_{-0.0}\,^{+0.2}_{-0.0}$	$0.0\,^{+0.4}_{-0.0}\,^{+0.2}_{-0.0}$
Drell-Yan data-driven	$0.23 \pm 0.16 \pm 0.12$	$0.35 \pm 0.21 \pm 0.18$	N/A
Data	0	1	0

Z-veto, $N_{jet} \geq 1$,

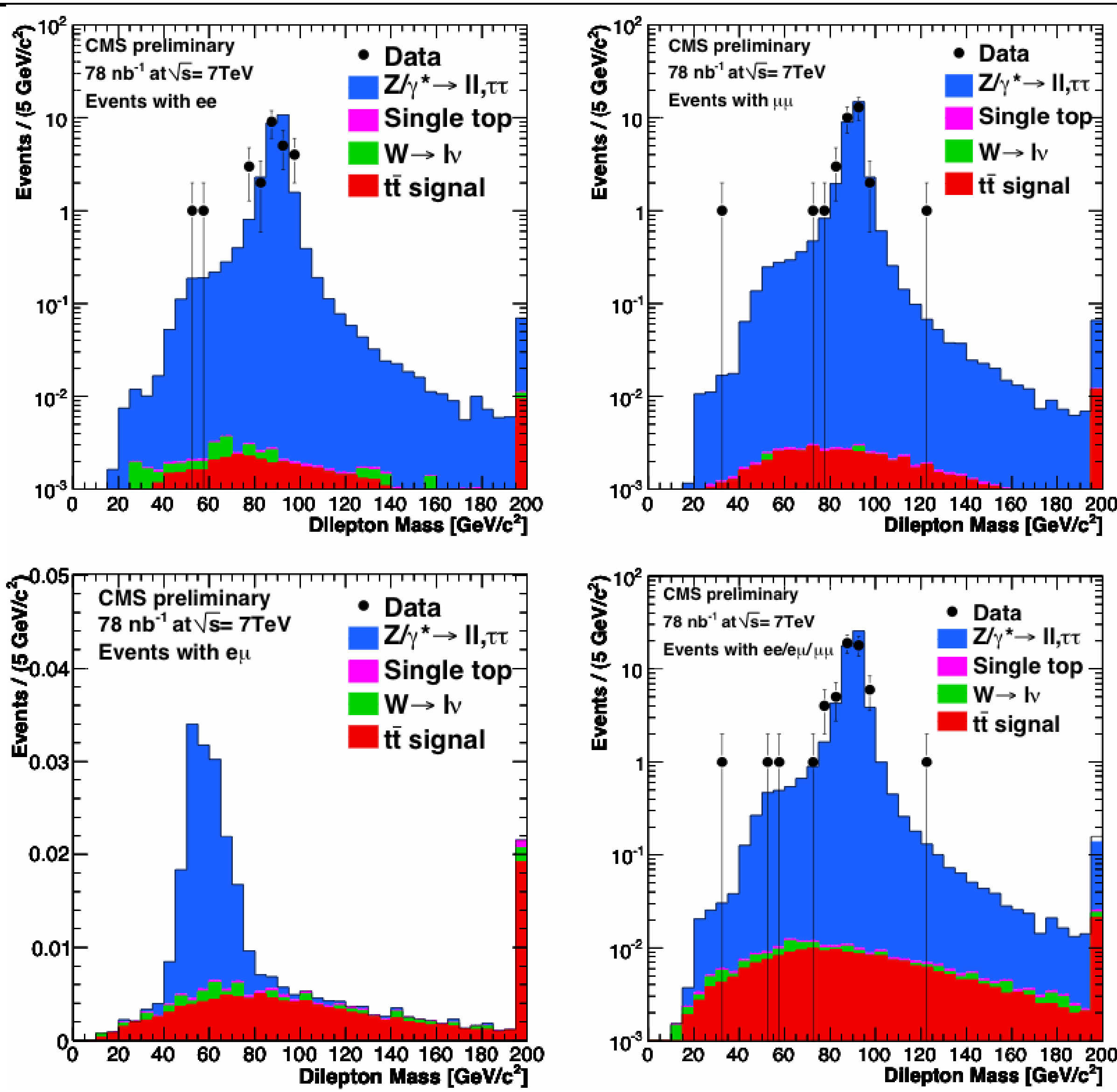
but no MET cut





"fake rate" measurements with jet-triggered data

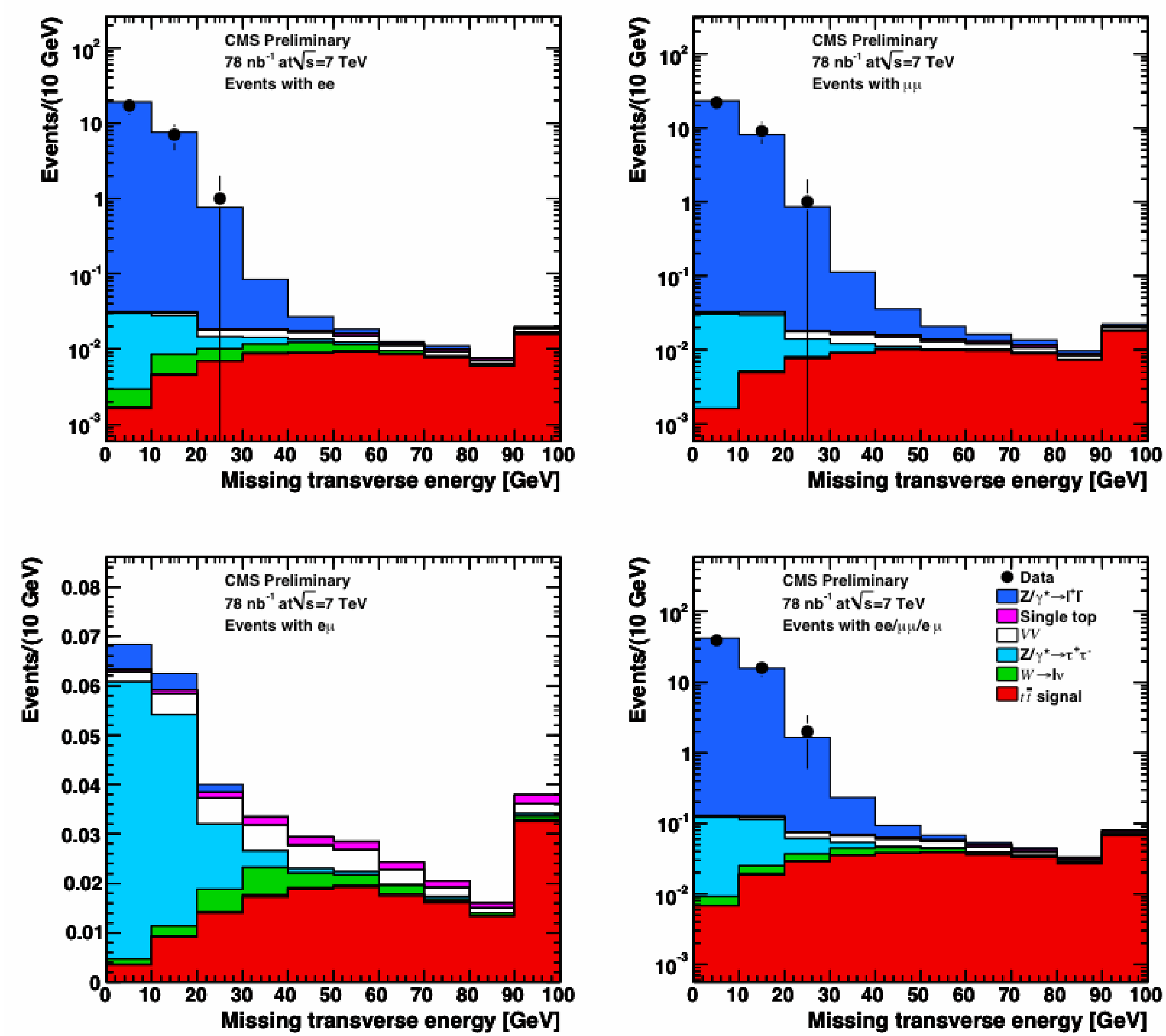
Dilepton Channel



Dilepton invariant mass

Relaxed selection: no Z-veto, $N_{jet} \geq 0$, no MET cut

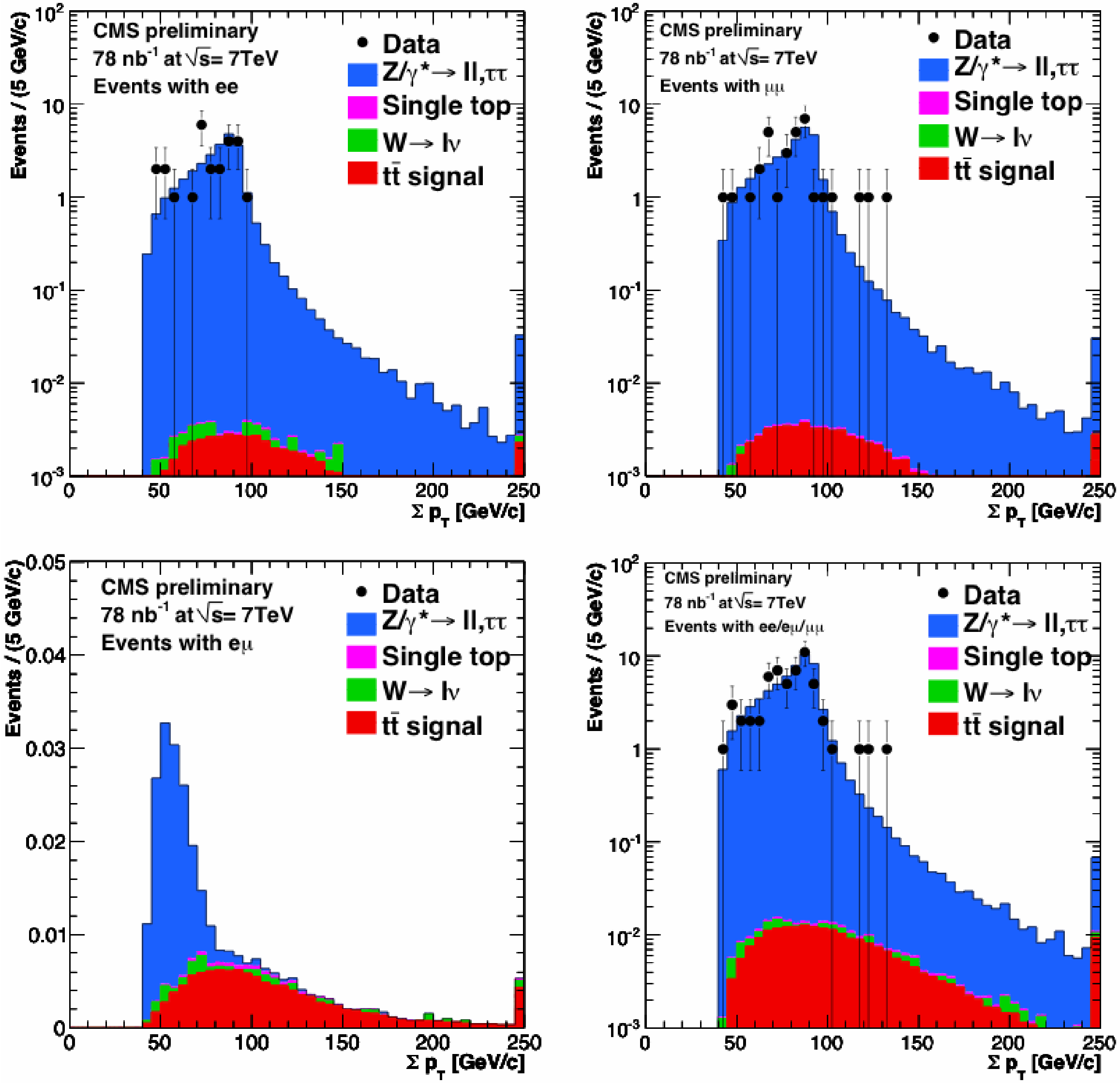
Dilepton Channel



Missing transverse energy

Relaxed selection: no Z-veto, $N_{jet} \geq 0$, no MET cut

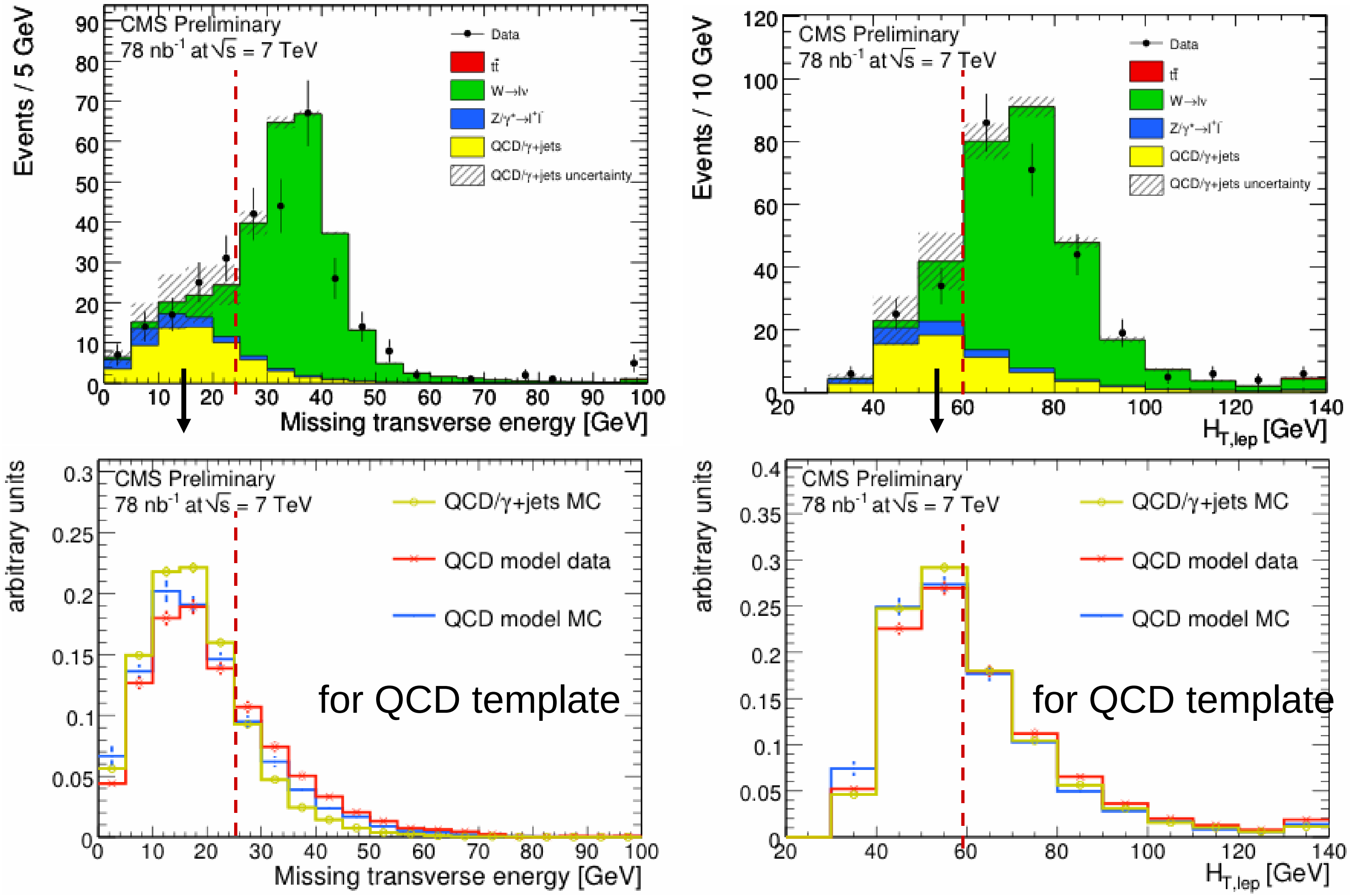
Dilepton Channel



Sum (p_T) of the 2 lepton transverse momenta

Relaxed selection: no Z-veto, $N_{jet} \geq 0$, no MET cut

e+Jets: Fits in $H_{T,lep}$ & MET



e+Jets: Fits in $H_{T,lep}$ & MET, Results

Fit templates in background region → extract N(QCD)

2 complementary templates for QCD: electron candidates that nearly pass e-ID criteria and jets with large electro-magnetic fraction (*W* template from simulation)

Table 5: *e*+jets: Results of the QCD estimation using templates for events without any jet requirement.

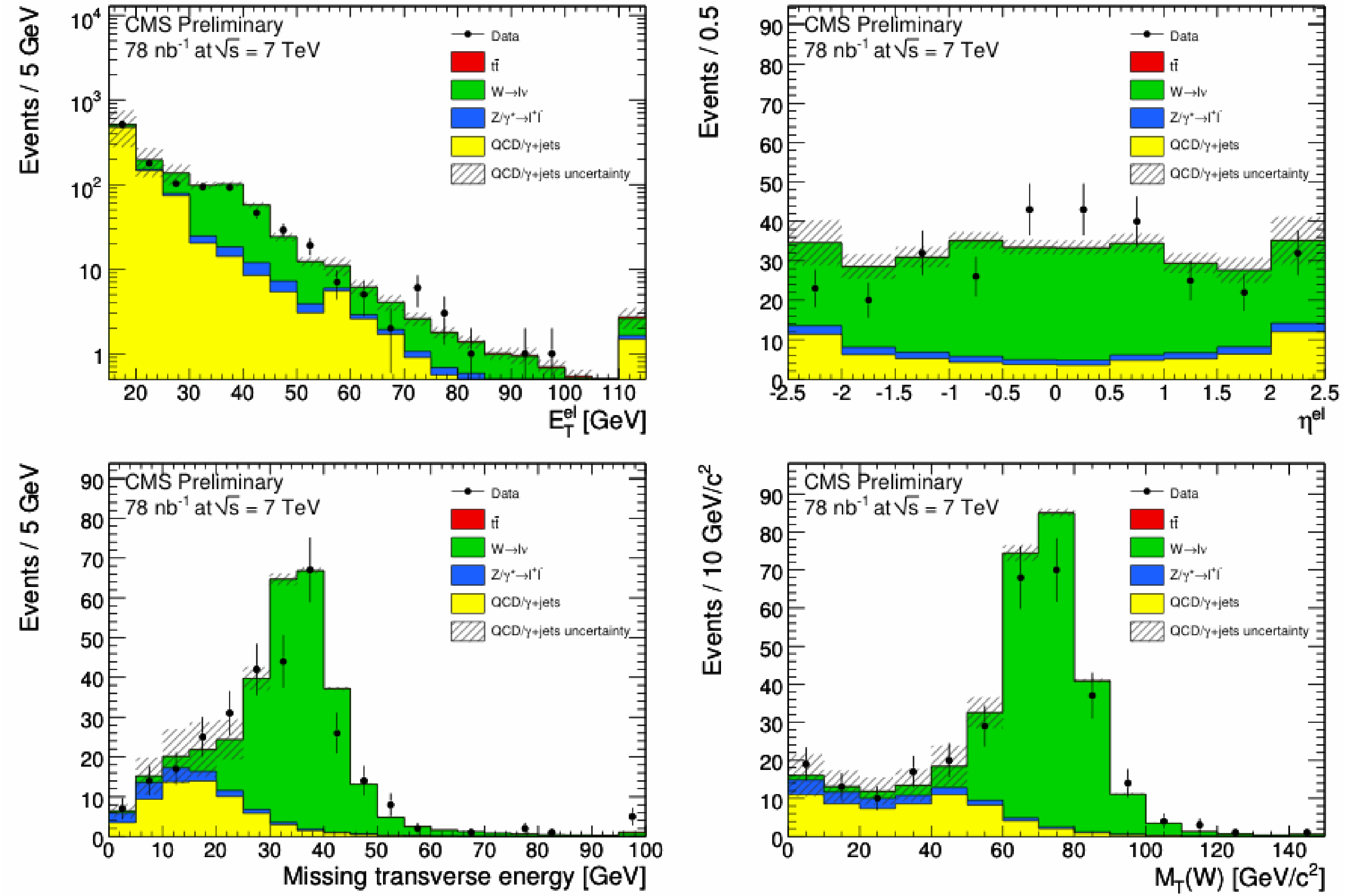
Variable	Template	QCD in bkg. region	QCD in sig. region	Whole dataset
$\not\!\!E_T$ (>25)	QCD model	41 ± 15	19 ± 7	60 ± 23
	prediction (sim.)	50.5 ± 0.5	12.2 ± 0.2	62.7 ± 0.5
$H_{T,lep}$(>60)	QCD model	47 ± 13	39 ± 11	86 ± 24
	prediction (sim.)	36.7 ± 0.4	26.0 ± 0.3	62.7 ± 0.5

Table 6: *e*+jets: Results of the QCD estimation using templates for event with at least one jet.

Variable	Template	QCD in bkg. region	QCD in sig. region	Whole dataset
$\not\!\!E_T$ (>25)	QCD model	28 ± 17	8 ± 5	37 ± 21
	prediction (sim.)	36.3 ± 0.4	5.3 ± 0.1	41.6 ± 0.4
$H_{T,lep}$(>60)	QCD model	26 ± 10	10 ± 4	36 ± 14
	prediction (sim.)	29.2 ± 0.4	12.4 ± 0.2	41.6 ± 0.4

Simulation describes the data as well
as the BG estimate very well!

e+Jets Kinematical Distributions



μ+Jets Kinematical Distributions