Matthias Mozer for the CMS experiment

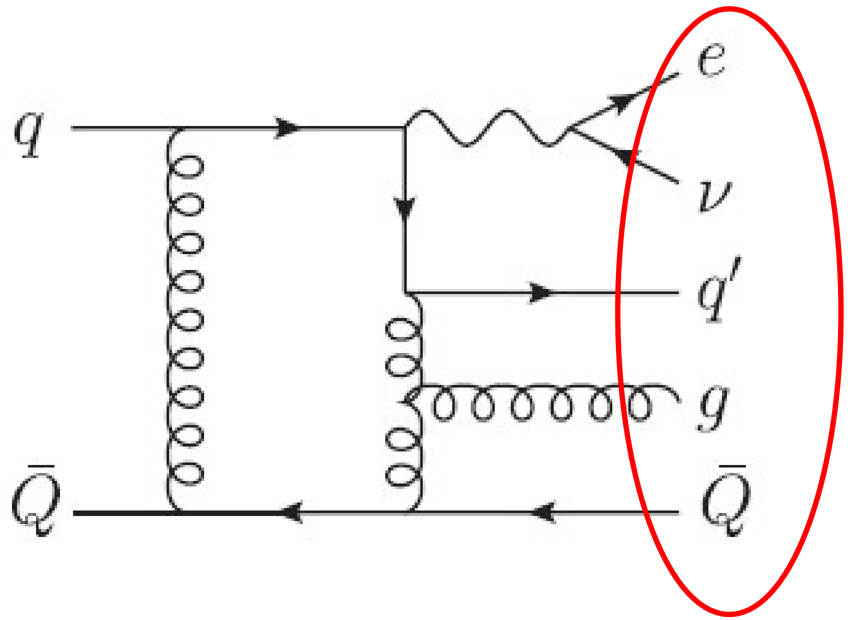
Z/W + jets in CMS

Introduction

- W/Z + jets: precision probe into QCD
- Important background on searches

- Three studies presented here:
  - W/Z + jets
  - Z + b jets
  - Boosted W polarization
- All on full 2010 data (36 pb^{-1})

W/Z + Jets

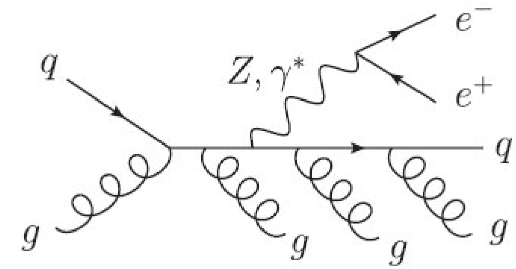


- Basic QCD + EWK
- But contains:
  - Jets
  - Leptons
  - MET
- Prominent background in searches

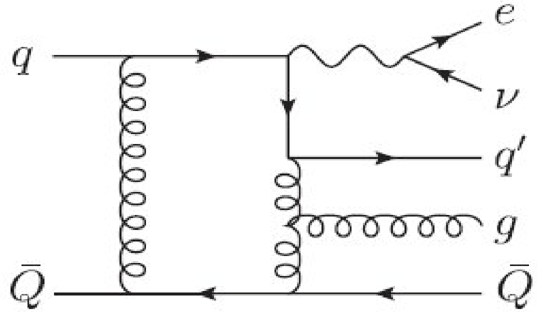
- How well do we understand these processes?
- How much can we rely on simulation?

W/Z + Jets

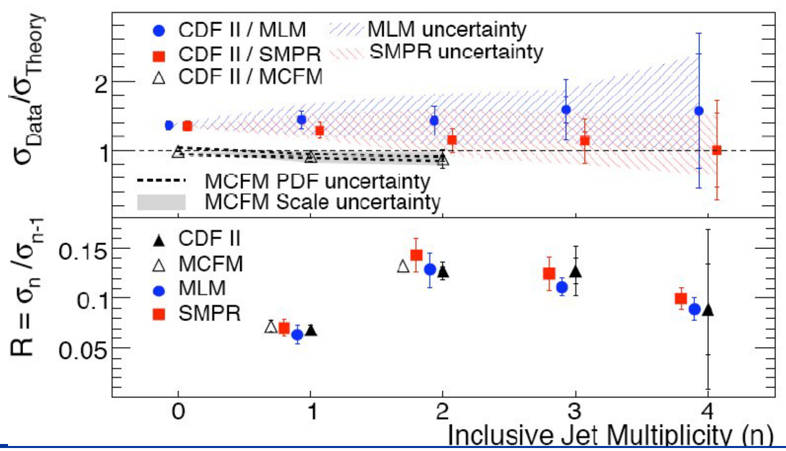
- Most used for background studies: ME+PS simulation:
  - Tree level only
  - Includes non-perturbative corrections



- Proper NLO calculations
  - Recent development
  - Not yet widely used



- Good results from the Tevatron
  - Can we improve?



Selection: e

- Require single electron trigger (thresholds <17 GeV)
- Require offline reconstructed electron with
  - P_t > 20 GeV, |η| < 2.5 , 1.44 < |η| < 1.57 excluded
  - Matches trigger primitive
  - Tight isolation, cluster shape, track matching, conversion rejection → ~80% efficiency
- Search for second electron with:
  - P_t > 10 GeV, |η| < 2.5 , 1.44 < |η| < 1.57 excluded
  - $60 < M_{ee} < 120$
- If second e passes loose (~95% efficiency) identification => Z sample
- No second electron => W sample if
  - M_T > 20 GeV
  - no muon with p_t >10 GeV (top veto)

Selection: e

- Require single electron trigger (thresholds <17 GeV)
- Require offline reconstructed electron with
  - **P_t > 20 GeV, |η| < 2.5 , 1.44 < |η| < 1.57 excluded**
  - Matches trigger primitive
  - Tight isolation, cluster shape, track matching, conversion rejection → ~80% efficiency
- Search for second electron with:
  - **P_t > 10 GeV, |η| < 2.5 , 1.44 < |η| < 1.57 excluded**
  - **60 < M_{ee} < 120**
- If second e passes loose (~95% efficiency) identification => Z sample

Results quoted in this acceptance

- No second electron => W sample if
  - M_T > 20 GeV
  - no muon with p_t >10 GeV (top veto)

Selection: μ

- Require single muon trigger (threshold < 15 GeV)
- Require one muon with
  - $P_t > 20$ GeV, $|\eta| < 2.1$
  - Matches trigger primitve
  - Isolation, good track fit quality
- Search second muon with
  - $P_t > 10$ GeV, $|\eta| < 2.4$
  - $60 < M_{\mu\mu} < 120$
- If second muon is found
  => Z sample
- No second muon => W sample if
  - $M_T > 20$ GeV

Selection: μ

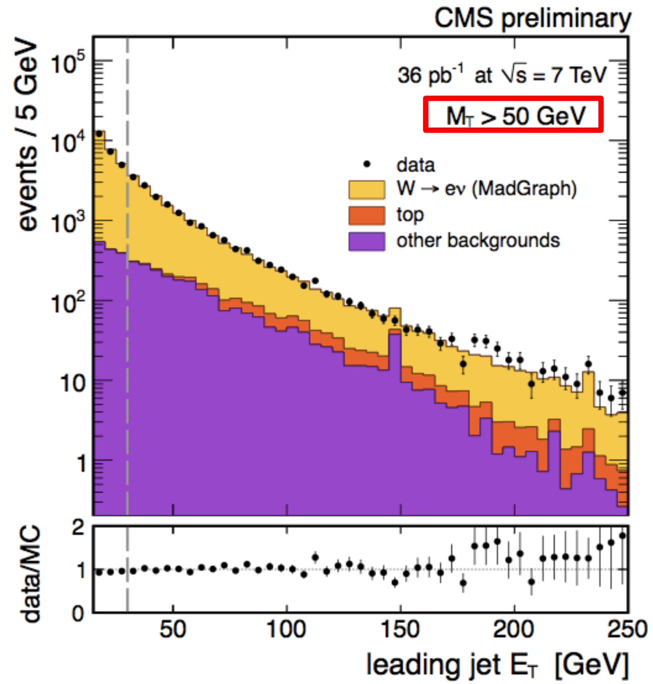
- Require single muon trigger (threshold < 15 GeV)
- Require one muon with
  - P_t > 20 GeV, |η| < 2.1
  - Matches trigger primitve
  - Isolation, good track fit quality
- Search second muon with
  - P_t > 10 GeV, |η| < 2.4
  - 60 < $M_{\mu\mu}$ < 120
- If second muon is found
  => Z sample
- No second muon => W sample if
  - M_T > 20 GeV

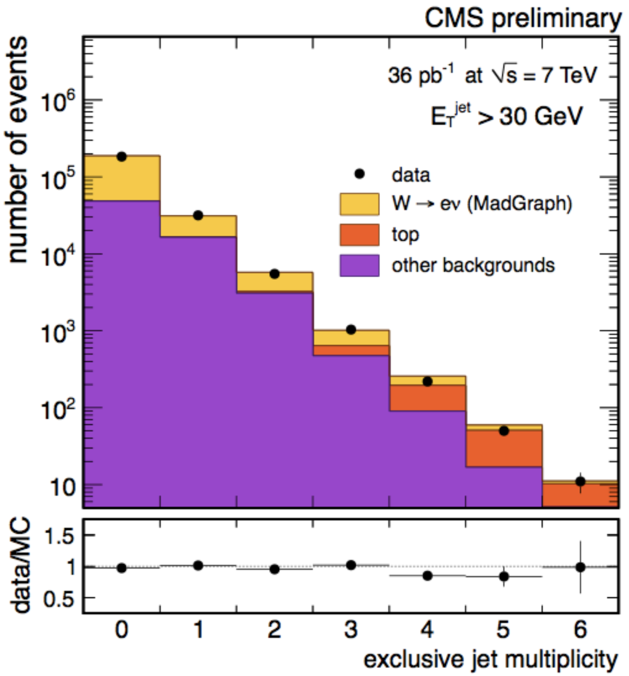
Results quoted in this acceptance

Jets

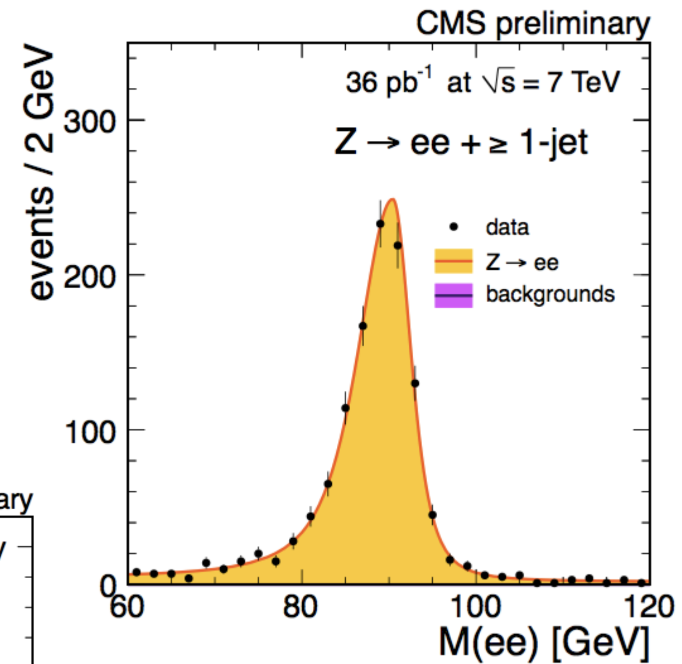
- Anti-kt algorithm (ΔR = 0.5) using "Particle Flow" objects
  - |η|<2.4 (tracker acceptance), E_T>30GeV
  - Data driven jet energy calibration
  - Pile-up: remove energy offset with FastJet
  - Muons: removed from particle list before clustering
    Electrons: veto jets within ΔR < 0.3 of W/Z decay electrons

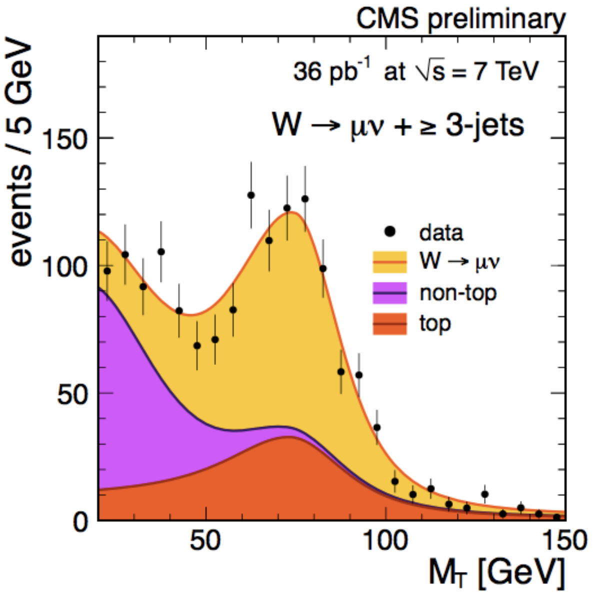
W→eν

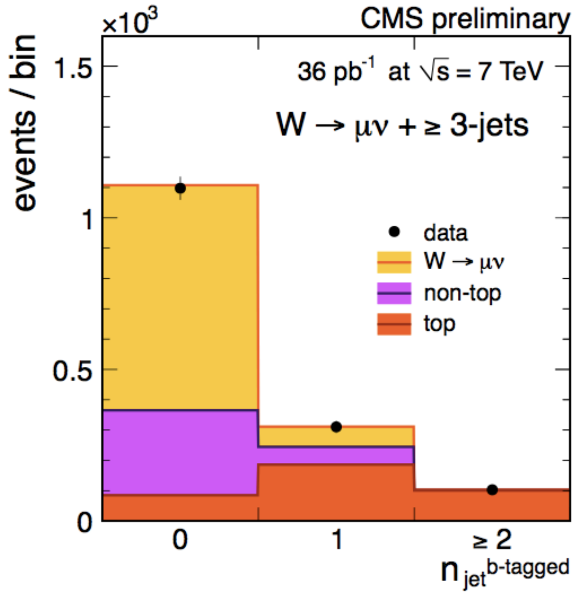




- Unbinned maximum likelihood fits to
  Z: M_{ll}
  W: M_T and n-btag (to control top)
- Only left tail and non-top bkg shape fixed to MC/control sample







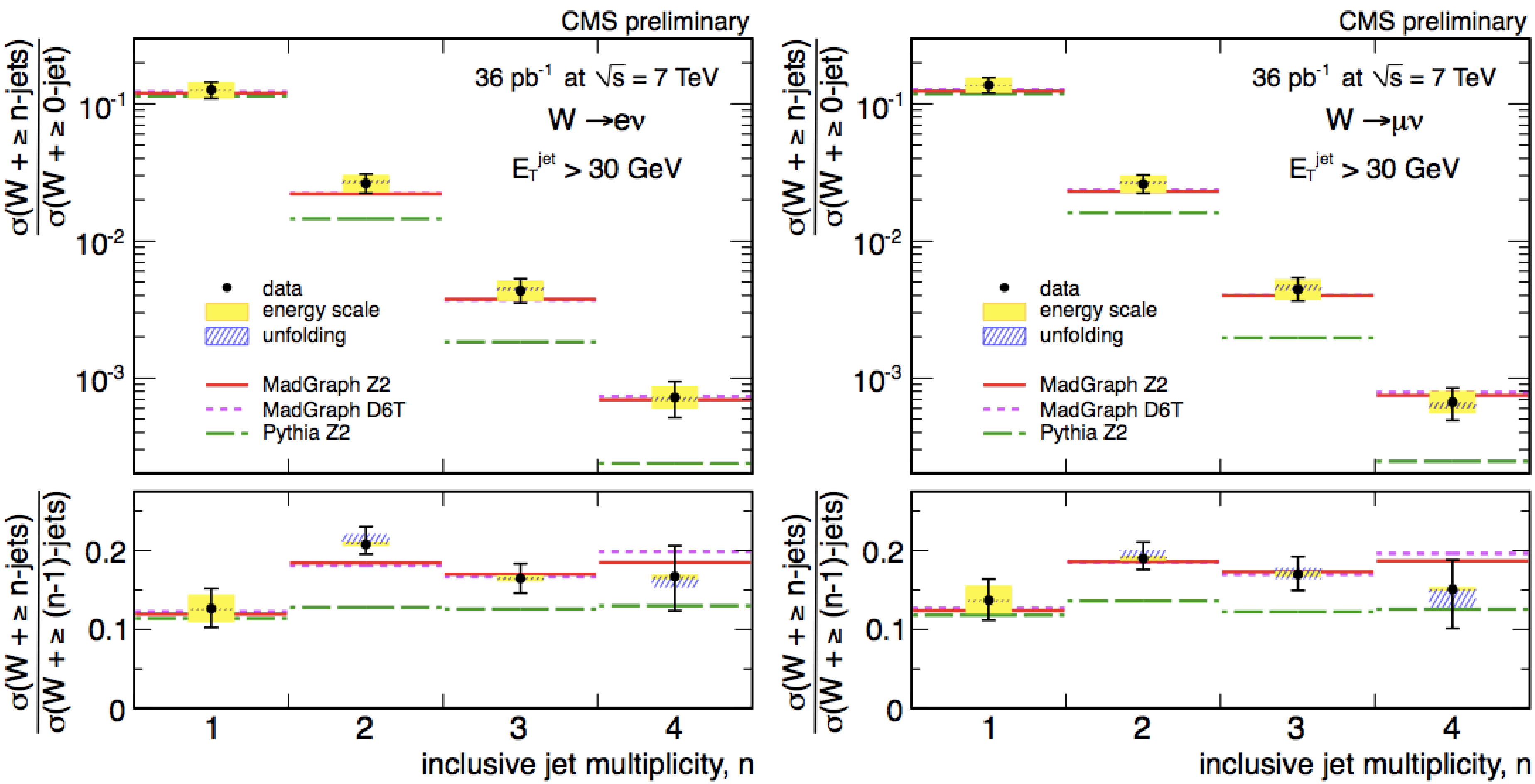
Corrections

- Efficiency: dependence on N_{jet} most important
  - Study with Tag & Probe (muons), MC (electrons), factorize as:
    - reconstruction (cluster → ele / track → muon): no N_{jet} dependence
    - identification: N_{jet} dependence due to isolation cuts
    - trigger (leading leg only): no N_{jet} dependence
- Migrations between jet bins
  - Extract migration matrix $R(n^{RECO}, n^{GEN})$ from MC
  - Use singular value decomposition (SVD) to “unsmear” N_{jet} distribution
- Measure

$$\frac{\sigma(\mathrm{V}+\geq n\mathrm{-jets})}{\sigma(\mathrm{V}+\geq 0\mathrm{-jets})} \qquad \frac{\sigma(\mathrm{V}+\geq n\mathrm{-jets})}{\sigma(\mathrm{V}+\geq (n-1)\mathrm{-jets})}$$

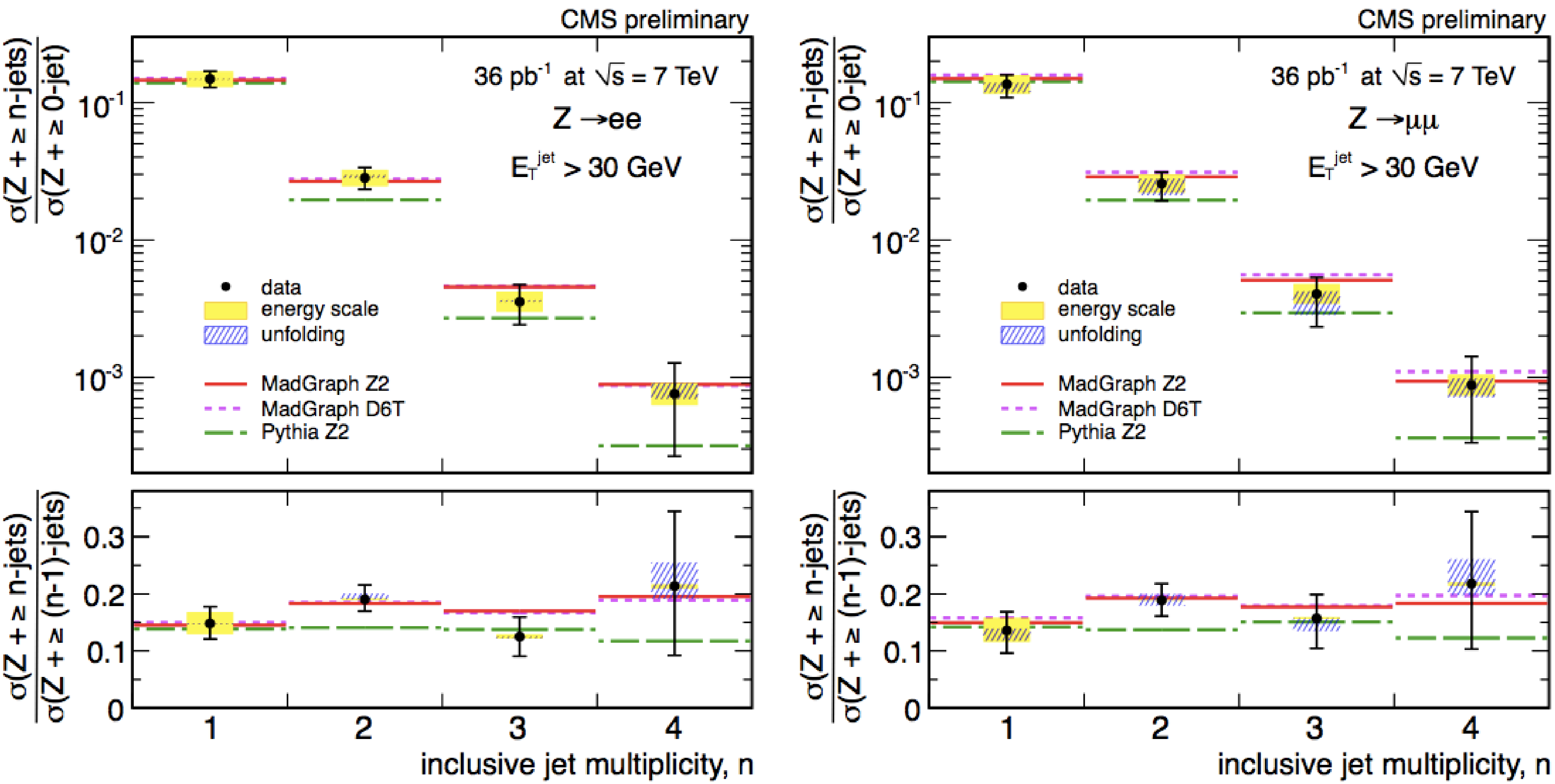
to reduce systematic uncertainties (lepton id, jet energy scale, lumi …)

Results: Rates W



Very good agreement with predictions from ME+PS simulation, while PS alone starts to fail for $n_{jet} \geq 2$

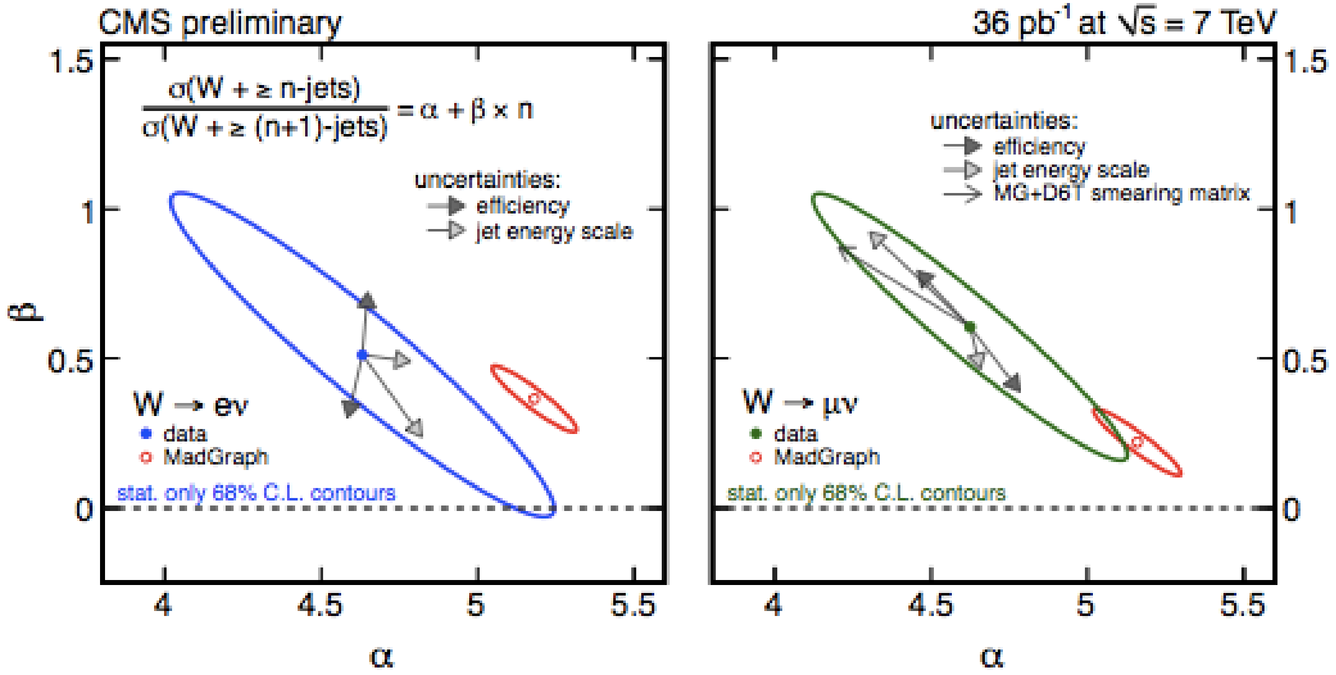
Results: Rates Z



Excellent agreement with ME+PS, but PS alone also compatible

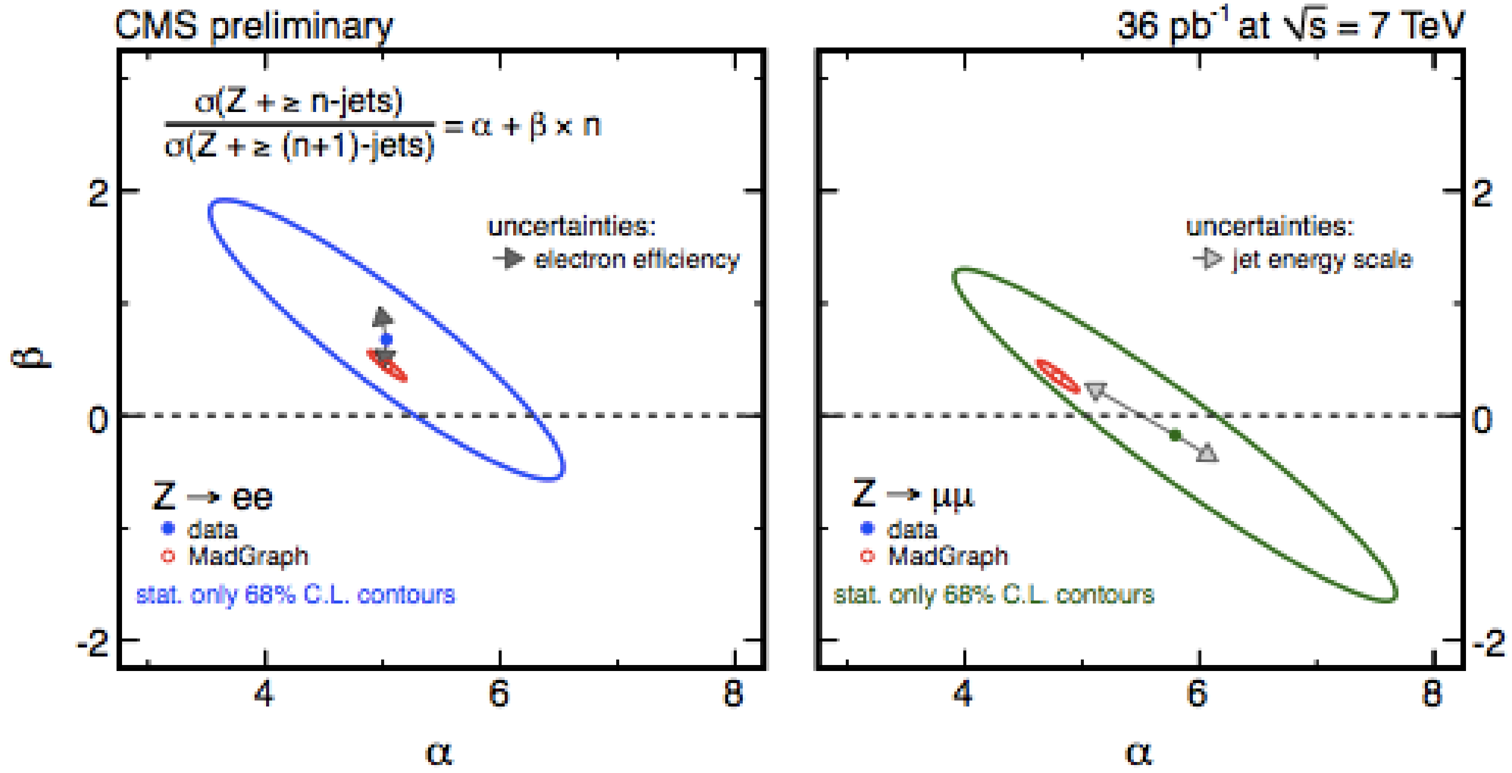
Results: Berends-Giele Scaling

- Expect $C_n = \frac{\sigma_n}{\sigma_{n+1}}$ to be ~ constant for n ≥1
- Test scaling by fitting $C_n = \alpha + \beta n$
- Taking into account correlations between σ_n
- Taking into account migrations between jet bins



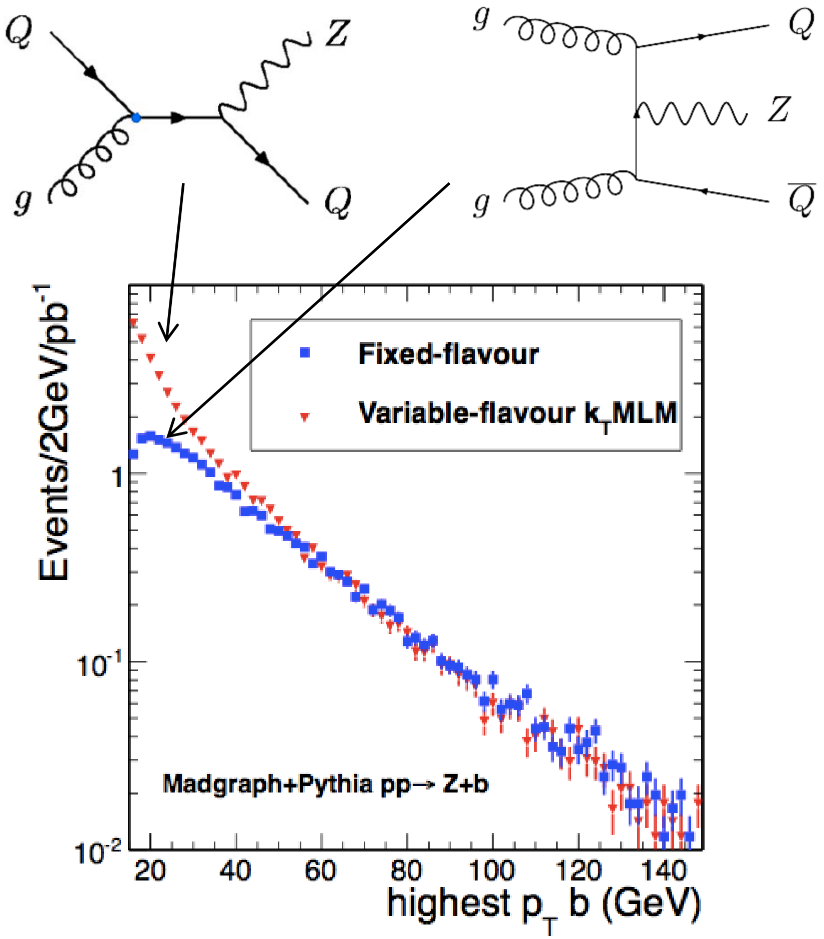
Results: Berends-Giele Scaling

- Reasonable agreement to ME+PS expectation for W and Z, e and μ

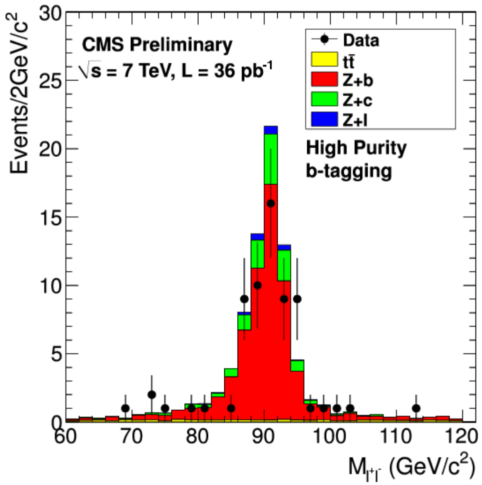


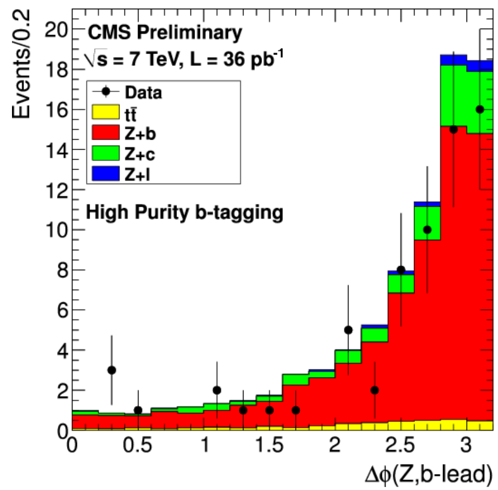
Z + b jets

- Benchmark channel for MSSM Higgs searches
- Fixed vs variable flavour number schemes (LO only)
- Select Events with
  - At least one Z
  - At least one jet (Et >25 GeV)
  - At least one secondary vertex in the jet
  - Met < 40 GeV (top rejection)
- Two b-jet selections:
  - High purity
  - High efficiency

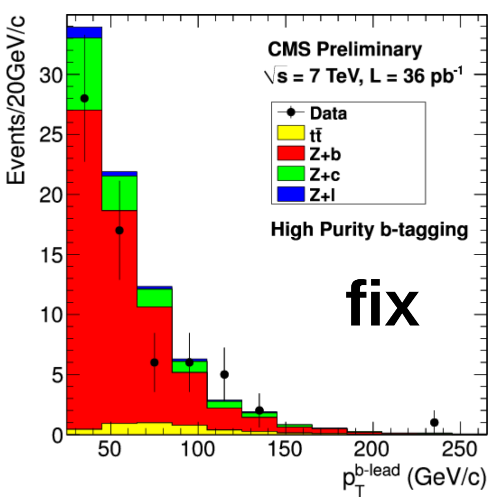


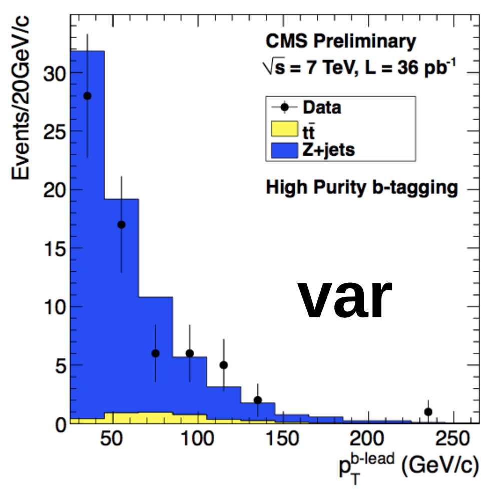
Z + b jets





Good agreement with ME+PS





No clear distinction between fixed and variable flavour number schemes:
Data sample mostly in the kinematic Domain where both agree

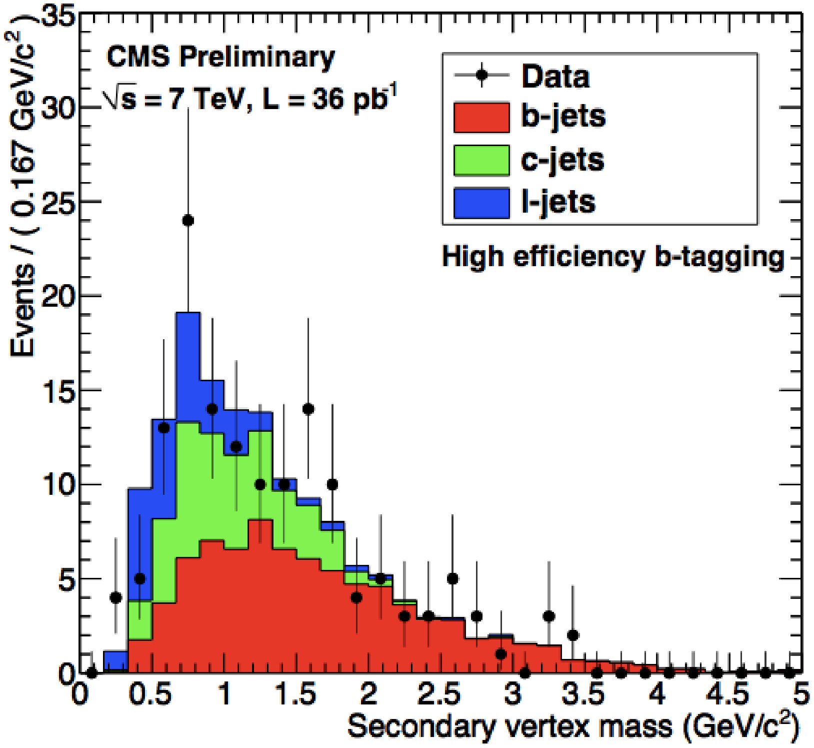
Z+b/Z+jet ratio

- Z+b purity is extracted from fit to the secondary vertex mass

Purity (%)	SSVHE	SSVHP
data	55±9	88±11
MC	57±3	82±4

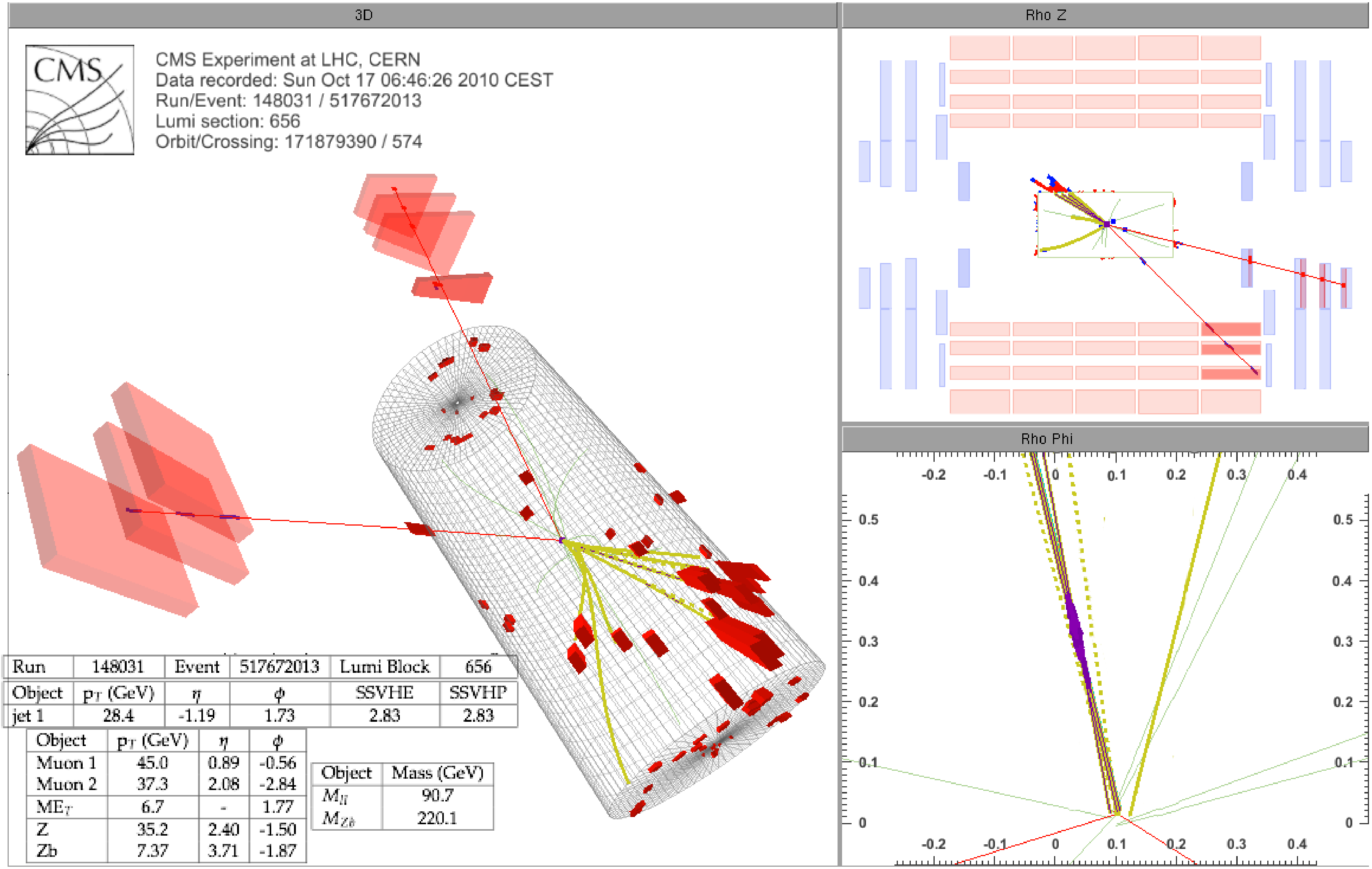
- Results are compatible with MadGraph(*) and MCFM NLO calculations

- (*) Z+b and Z+c with $p_{T,jet}$ >15 GeV scaled to corresponding MCFM x-sec



Sample	$\mathcal{R} = \frac{\sigma(pp\to Z+b+X)}{\sigma(pp\to Z+j+X)}$ (%), $p_T^e > 25$ GeV, $\lvert\eta^e\rvert < 2.5$	$\mathcal{R} = \frac{\sigma(pp\to Z+b+X)}{\sigma(pp\to Z+j+X)}$ (%), $p_T^\mu > 20$ GeV, $\lvert\eta^\mu\rvert < 2.1$
Data HE	$4.3 \pm 0.6(stat) \pm 1.1(syst)$	$5.1 \pm 0.6(stat) \pm 1.3(syst)$
Data HP	$5.4 \pm 1.0(stat) \pm 1.2(syst)$	$4.6 \pm 0.8(stat) \pm 1.1(syst)$
MADGRAPH	$5.1 \pm 0.2(stat) \pm 0.2(syst) \pm 0.6(theory)$	$5.3 \pm 0.1(stat) \pm 0.2(syst) \pm 0.6(theory)$
MCFM	$4.3 \pm 0.5(theory)$	$4.7 \pm 0.5(theory)$

Z + b: Example Event



Run	148031	Event	517672013	Lumi Block	656
Object	p_T (GeV)	η	ϕ	SSVHE	SSVHP
jet 1	28.4	-1.19	1.73	2.83	2.83

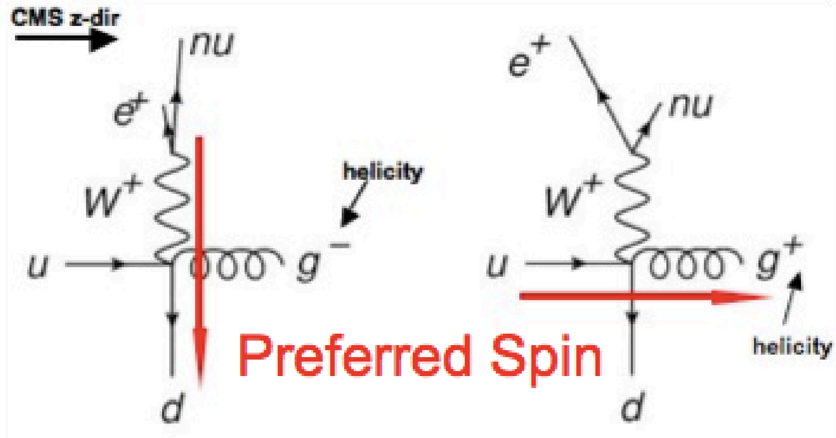
Object	p_T (GeV)	η	ϕ
Muon 1	45.0	0.89	-0.56
Muon 2	37.3	2.08	-2.84
ME_T	6.7	-	1.77
Z	35.2	2.40	-1.50
Zb	7.37	3.71	-1.87

Object	Mass (GeV)
M_{ll}	90.7
M_{Zb}	220.1

Boosted W polarization

Production of high p_T W-bosons (p_T > 50 GeV)

- 7 TeV+high p_T dominant production valence quark w/gluon



W+jet dominant production mechanisms at LHC (not a Feynman diagram)

- **Strong polarization effects in transverse plane**
- SM: Predominant left handedness for + and -
- Unlike tevatron ($p\bar{p}$)
  - No CP counterparts
  - Cause for left handedness
- Robust over jet multiplicity

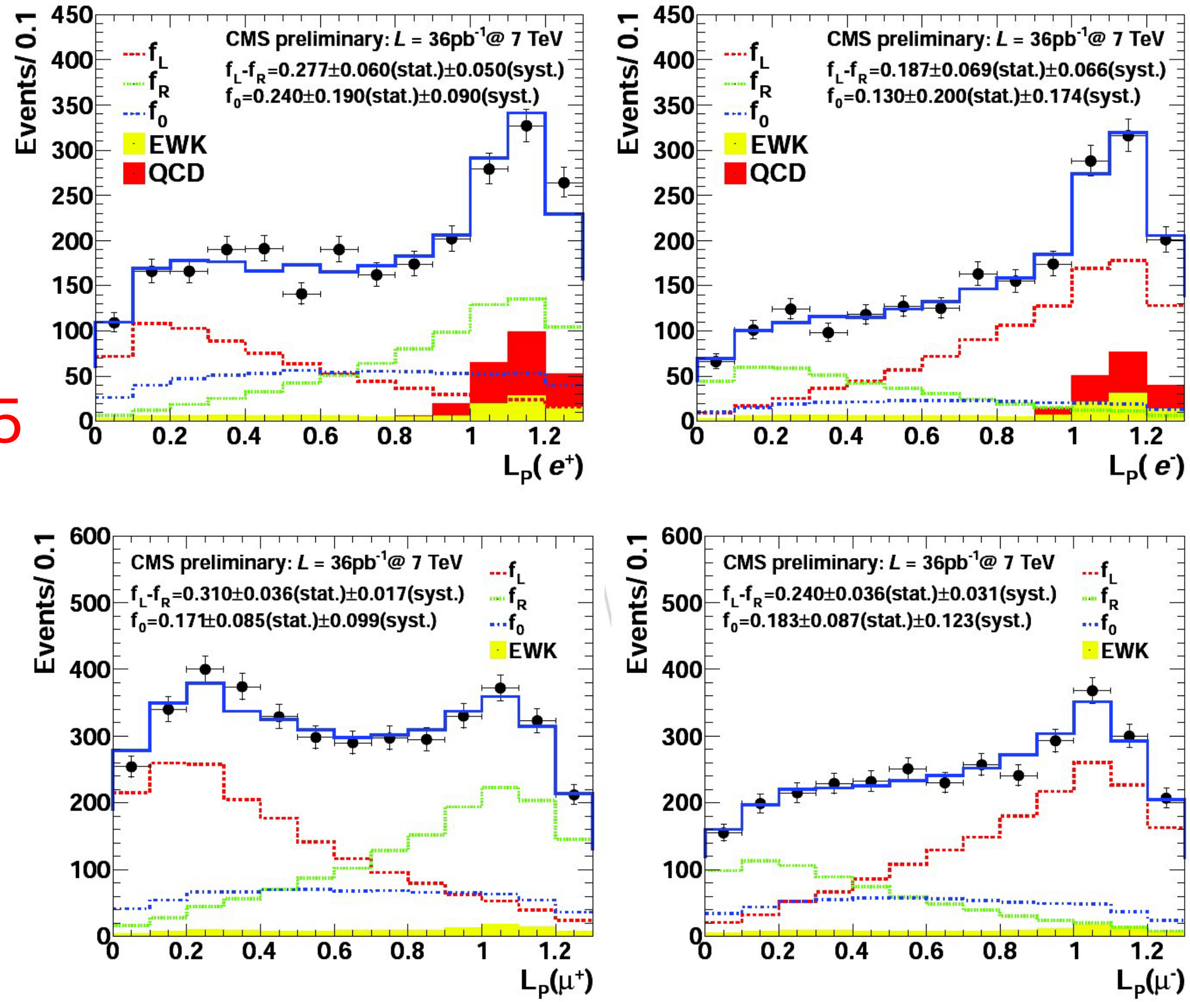
Expect left right polarization asymmetry in a pp collider

Signal Extraction

- Require W canidate: $P_T(W)=P_T(l)+MET > 50$ GeV
- ν not measured:
  => ϑ^* undetermined
- Use proxy instead:

$$\mathrm{LP} = \frac{\vec{p}_T(\ell)\cdot\vec{p}_T(W)}{|\vec{p}_T(W)|^2}$$

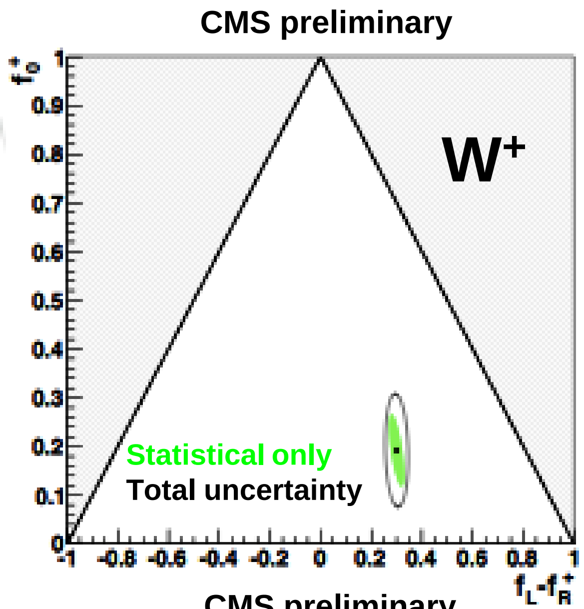
  $\mathrm{LP} \approx 0.5\cos(\vartheta^*)+0.5$
- Extract polarization with template fit

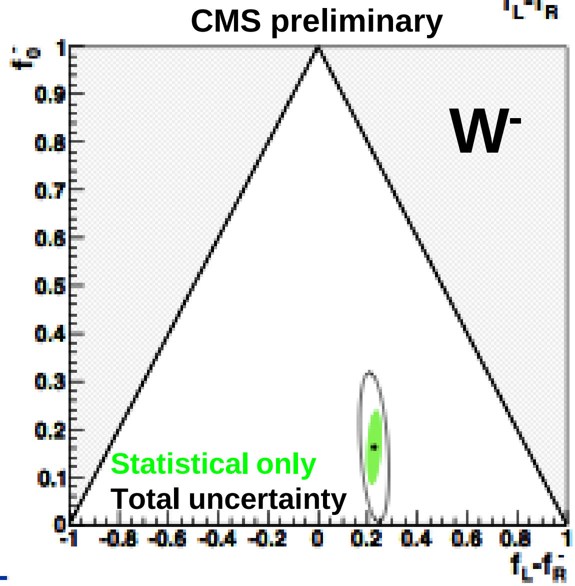


Boosted W polarization results

- Systematics dominated by MET uncertainty
- $f_L - f_R > 0$ => mostly left-handed

	Combined Results
$(f_L - f_R)^-$	0.226 ± 0.031 (stat) ± 0.050 (syst)
f_0^-	0.162 ± 0.078 (stat) ± 0.136 (syst)
$(f_L - f_R)^+$	0.300 ± 0.031 (stat) ± 0.034 (syst)
f_0^+	0.192 ± 0.075 (stat) ± 0.089 (syst)

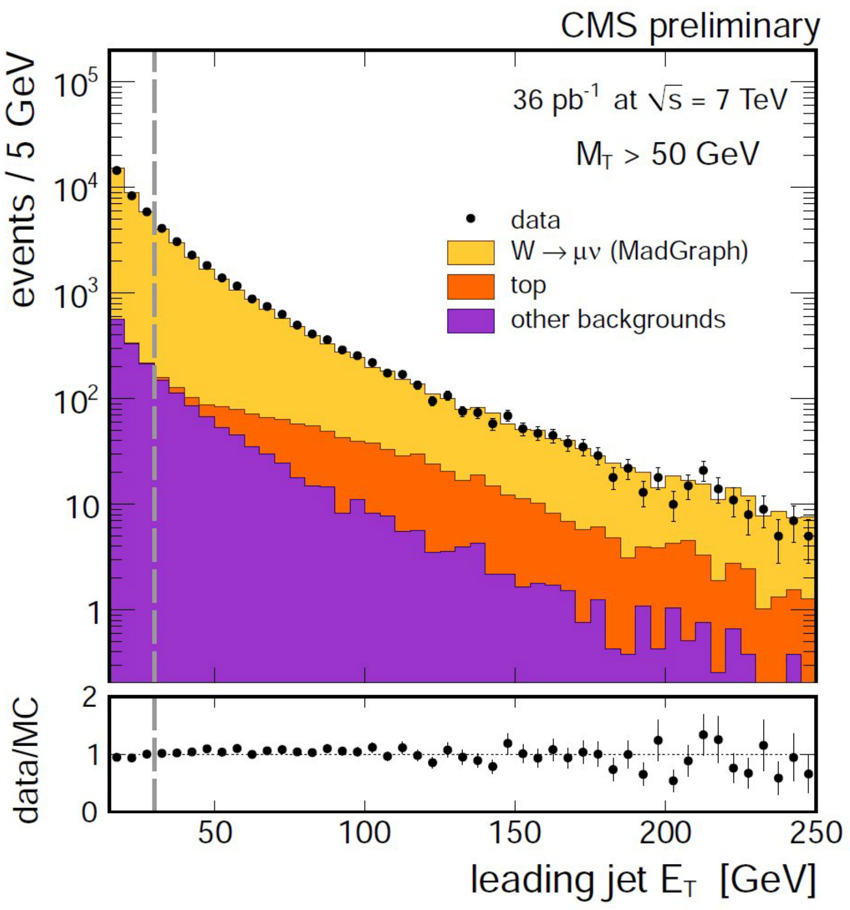


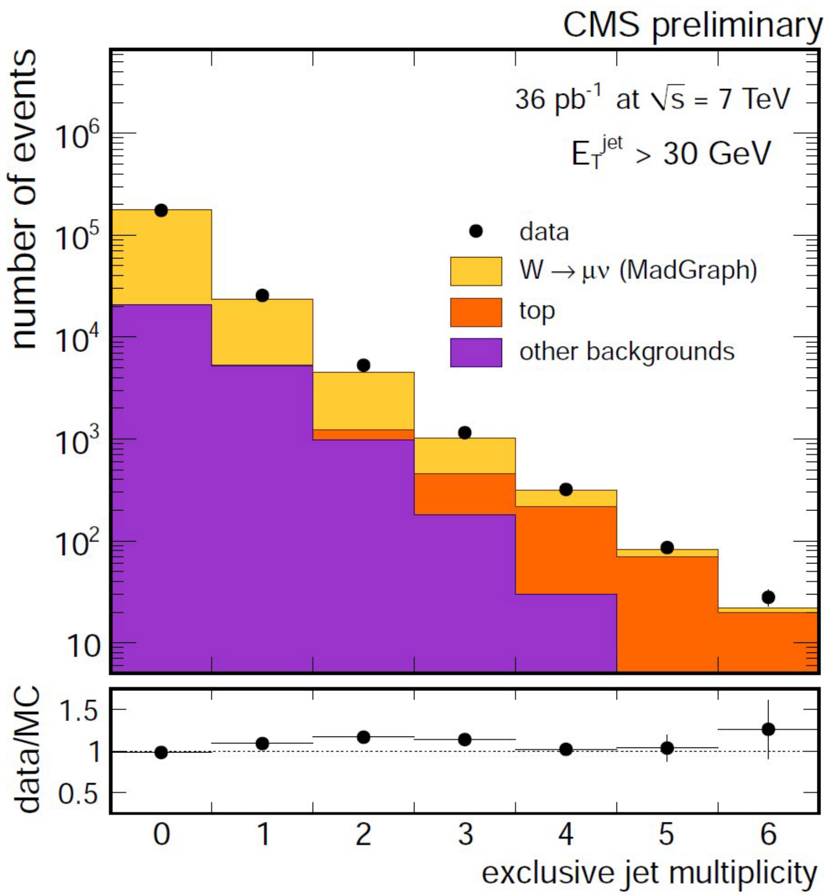


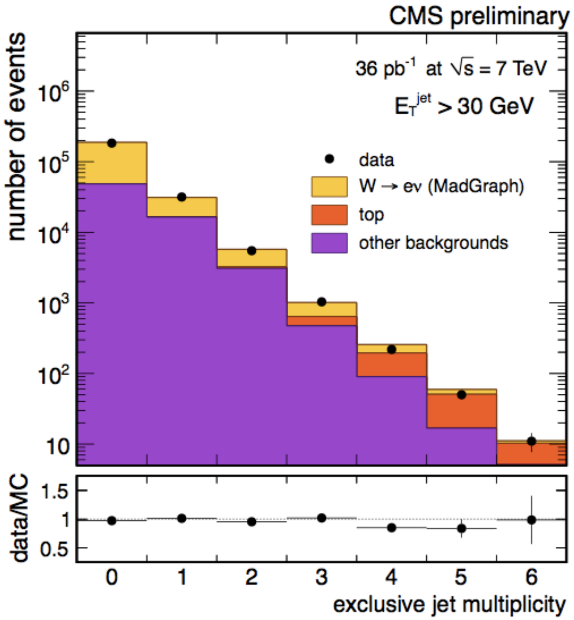
Summary

- Comprehensive set of measurements on full 2010 data ($36\ pb^{-1}$)
- Jet rates for $E_t > 30$ GeV in agreement with ME+PS
- Direct measurement of Berends-Giele scaling agrees with expectations
- Measured significant polarization of boosted W
- Observation of Z + b and ratio Z + b / Z + jets agrees well with NLO calculation

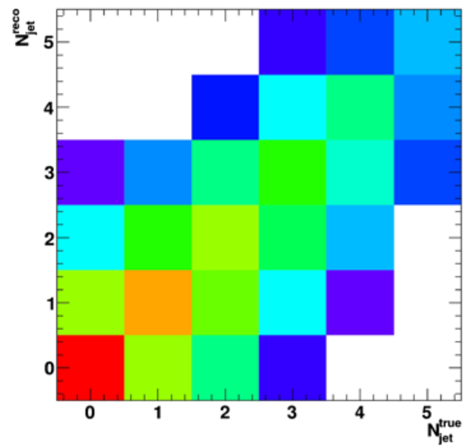
Backup


CMS preliminary
36 pb-1 at √s = 7 TeV
M_T > 50 GeV
data
W → μν (MadGraph)
top
other backgrounds
events / 5 GeV
data/MC
leading jet E_T [GeV]


CMS preliminary
36 pb-1 at √s = 7 TeV
E_T^jet > 30 GeV
data
W → μν (MadGraph)
top
other backgrounds
number of events
data/MC
exclusive jet multiplicity



=



N_0

N_1

$N_2 = N_1/(\alpha+2\beta)$

$N_3 = N_2/(\alpha+3\beta)$

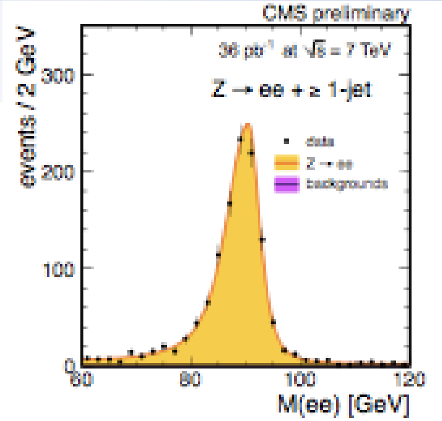
....

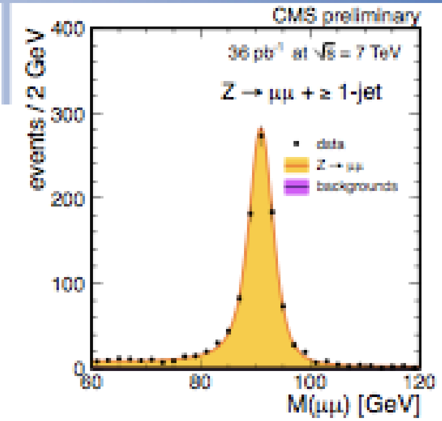
muons

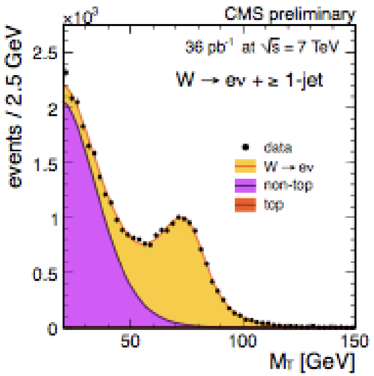
		data	stat	JES MC	$\epsilon(\ell)$	D6T tune	Theory
Z	α	5.8	$\pm$ 1.2	$\pm$0.6	$\pm$0.1	+0.3	4.8 ± 0.1
	β	-0.2	$\pm$ 1.0	$\pm$0.3	$\pm$0.1	-0.0	0.35 ± 0.09
W	α	4.3	$\pm$0.3	$\pm$0.2	$\pm$0.2	-0.4	5.16 ± 0.09
	β	0.7	$\pm$0.3	$\pm$0.2	$\pm$0.3	+0.3	0.22 ± 0.06

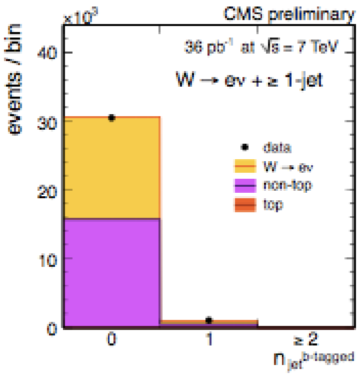
electrons

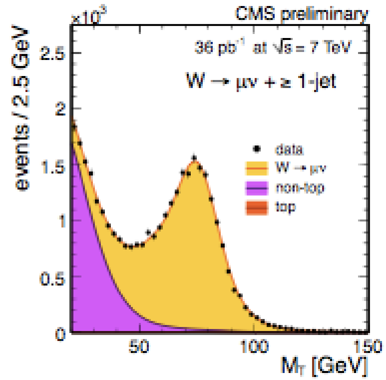
		data	stat	JES	$\epsilon(\ell)$	Theory
Z	α	5.0	$\pm$1.0	$^{+0.1}_{-0.0}$	$^{+0.00}_{-0.06}$	5.04 ± 0.10
	β	0.7	$\pm$0.8	$^{+0.08}_{-0.04}$	$^{+0.3}_{-0.6}$	0.45 ± 0.08
W	α	4.6	$\pm$0.4	$^{+0.2}_{-0.0}$	$^{-0.05}_{+0.02}$	5.18 ± 0.09
	β	0.5	$\pm$0.4	$^{+0.0}_{-0.3}$	$\pm$0.2	0.36 ± 0.07

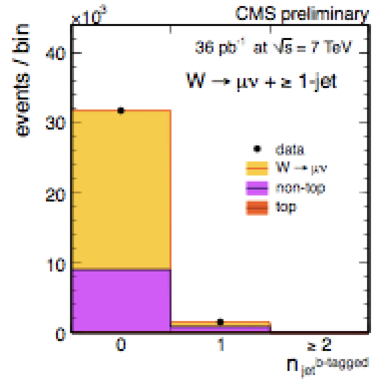

CMS preliminary
36 pb⁻¹ at √s = 7 TeV
Z → ee + ≥ 1-jet
data
Z → ee
backgrounds
events / 2 GeV
M(ee) [GeV]

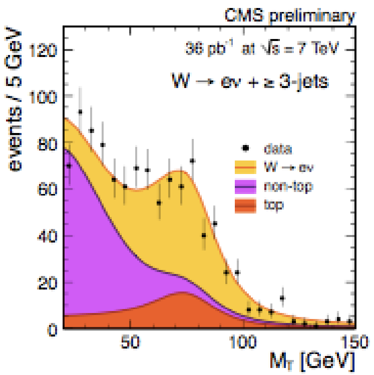

CMS preliminary
36 pb⁻¹ at √s = 7 TeV
Z → μμ + ≥ 1-jet
data
Z → μμ
backgrounds
events / 2 GeV
M(μμ) [GeV]

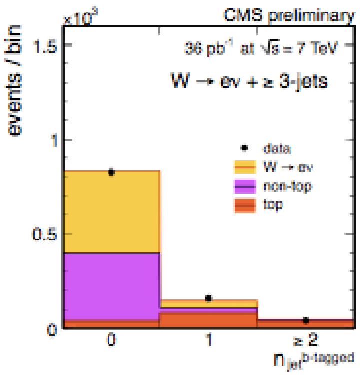

CMS preliminary
36 pb⁻¹ at √s = 7 TeV
W → eν + ≥ 1-jet
data
W → eν
non-top
top
events / 2.5 GeV
M_T [GeV]

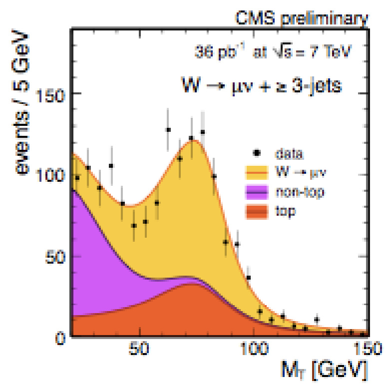

CMS preliminary
36 pb⁻¹ at √s = 7 TeV
W → eν + ≥ 1-jet
data
W → eν
non-top
top
events / bin
n_jet^b-tagged

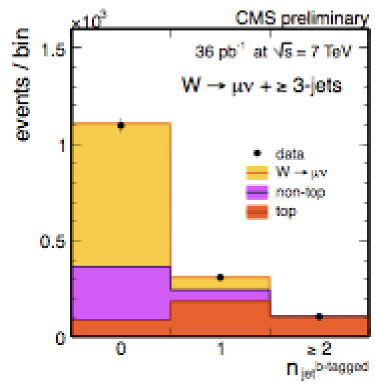

CMS preliminary
36 pb⁻¹ at √s = 7 TeV
W → μν + ≥ 1-jet
data
W → μν
non-top
top
events / 2.5 GeV
M_T [GeV]


CMS preliminary
36 pb⁻¹ at √s = 7 TeV
W → μν + ≥ 1-jet
data
W → μν
non-top
top
events / bin
n_jet^b-tagged


CMS preliminary
36 pb⁻¹ at √s = 7 TeV
W → eν + ≥ 3-jets
data
W → eν
non-top
top
events / 5 GeV
M_T [GeV]


CMS preliminary
36 pb⁻¹ at √s = 7 TeV
W → eν + ≥ 3-jets
data
W → eν
non-top
top
events / bin
n_jet^b-tagged


CMS preliminary
36 pb⁻¹ at √s = 7 TeV
W → μν + ≥ 3-jets
data
W → μν
non-top
top
events / 5 GeV
M_T [GeV]


CMS preliminary
36 pb⁻¹ at √s = 7 TeV
W → μν + ≥ 3-jets
data
W → μν
non-top
top
events / bin
n_jet^b-tagged

Jet energy scale uncertainty

- Systematics obtained from uncertainty on Jet Energy Corrections (JEC) including:
  - Corrections from data
  - JEC flavour dependence (estimated from MC)
  - UE offset subtraction by FastJet (500 MeV on each jet in MC)
- In addition we considered:
  - effects on MET were studied on a fit to M_T on data
  - jet energy resolution
  - pile-up residual effect on the jet rate after subtraction

Systematic uncertainty on jet counting [%]					
Jet multiplicity	0	1	2	3	≥ 4
Jet Energy Scale	∓ 1	± 6	$^{+9}_{-8}$	$^{+12}_{-11}$	$^{+14}_{-13}$
$\not{E}_T$ (W only)	$^{+0.6}_{-0.7}$	$^{+3.5}_{-3.1}$	$^{+4.5}_{-3.9}$	$^{+5.2}_{-4.5}$	$^{+6}_{-5}$
Jet Energy Resolution		$^{+0.6}_{-0.5}$	$^{+0.8}_{-0.7}$	$^{+1.0}_{-0.9}$	$^{+1.1}_{-1.0}$
Pile-up	∓ 5	± 5	± 5	± 5	± 5
Total in W events	∓ 5	± 8	$^{+11}_{-10}$	$^{+14}_{-12}$	$^{+16}_{-15}$
Total in Z events	∓ 5	± 8	± 10	$^{+13}_{-12}$	$^{+15}_{-14}$

Systematic uncertainties

- Systematic uncertainties on the exclusive rates - *after efficiency correction* - are shown
- Largest systematics due to jet reconstruction and efficiency
- Errors on the efficiency are largerly uncorrelated (statistical error on T&P in each multiplicity bins)
- Errors due to jet counting are instead fully correlated among channels and jet multiplicity

Uncertainties on jet rate in $W \to e\nu$ events [%]					
Jet multiplicity	0	1	2	3	≥ 4
Jet counting	∓5	±8	$^{+11}_{-10}$	$^{+14}_{-12}$	$^{+16}_{-15}$
Lepton efficiency	±3	$^{+6}_{-5}$	$^{+7}_{-6}$	±10	$^{+24}_{-12}$
Signal extraction		±0.1	±0.4	±2.9	±8.5
Total systematics	±6	±10	$^{+13}_{-12}$	$^{+18}_{-16}$	$^{+30}_{-21}$
Statistical uncertainty	±0.3	±1.0	±2.4	±7.5	±22
Uncertainties on jet rate in $W \to \mu\nu$ events [%]					
Jet multiplicity	0	1	2	3	≥ 4
Jet counting	∓5	±8	$^{+11}_{-10}$	$^{+14}_{-12}$	$^{+16}_{-15}$
Lepton efficiency	±3	±6	±4	±10	±17
Signal extraction		±0.1	±0.4	±2.9	±8.5
Total systematics	±6	±10	$^{+13}_{-12}$	$^{+19}_{-17}$	±26
Statistical uncertainty	±0.2	±0.8	±2.3	±6.5	±27

Uncertainties on jet rate in $Z \to e^+e^-$ events [%]					
Jet multiplicity	0	1	2	3	≥ 4
Jet counting	∓5	±8	$^{+11}_{-10}$	$^{+14}_{-12}$	$^{+16}_{-15}$
Efficiency	±3	$^{+6}_{-5}$	$^{+7}_{-6}$	±10	$^{+24}_{-12}$
Total systematics	±6	±10	$^{+13}_{-12}$	$^{+18}_{-16}$	$^{+30}_{-21}$
Statistical uncertainty	±1.0	±3.0	±8.0	±20	±47
Uncertainties on jet rate in $Z \to \mu^+\mu^-$ events [%]					
Jet multiplicity	0	1	2	3	≥ 4
Jet counting	∓5	±8	$^{+11}_{-10}$	$^{+14}_{-12}$	$^{+16}_{-15}$
Efficiency	±3	$^{+6}_{-5}$	$^{+7}_{-6}$	±10	$^{+24}_{-12}$
Total systematics	±6	±10	$^{+13}_{-12}$	$^{+18}_{-16}$	$^{+30}_{-21}$
Statistical uncertainty	±1.1	±2.7	±5.2	±18	±35

W signal extraction: top discrimination

- Simple PDF with 2 params: b-tag and mistag eff
- Probability for n_j^{tagged} in case of n_{bj} b-jets and n_j jets:

$$P(n_j^{tagged}|n_j, n_{bj}, \epsilon_{nob}, \epsilon_b) =$$

- $$\begin{cases} (1-\epsilon_{nob})^{n_j-n_{bj}} \cdot (1-\epsilon_b)^{n_{bj}} & n_j^{tagged} = 0 \\ (1-\epsilon_{nob})^{n_j-n_{bj}-1} \cdot \epsilon_{nob} \cdot (n_j - n_{bj}) \cdot (1-\epsilon_b)^{n_{bj}} + \\ \quad (1-\epsilon_{nob})^{n_j-n_{bj}} \cdot (1-\epsilon_b)^{n_{bj}-1} \cdot (\epsilon_b) \cdot n_{bj} & n_j^{tagged} = 1 \\ 1 - P(0) - P(1) & n_j^{tagged} \geq 2 \end{cases}$$

  - Signal corresponds to events with 0 b-jets (n_{b0}), top to events with 1 and 2 b-jets (n_{b1} + n_{b2})
  - Top with 0 b-jets is fixed to MC yield (very small)
  - Other BKG is determined from the M_T component of the likelihood
  - mistag eff: 2.42 $\pm$ 0.03(stat) $\pm$ 0.5(syst)%
  - b-tag eff: 63 $\pm$ 6.3%
- Cruijff function: $f(x; m, \sigma_L, \sigma_R, \alpha_L, \alpha_R) = N_s \cdot e^{-\frac{(x-m)^2}{2\sigma^2 + \alpha(x-m)^2}}$

where $\sigma = \sigma_L(\sigma_R)$ for $x < m (x > m)$ and $\alpha = \alpha_L(\alpha_R)$ for $x < m (x > m)$