Part I: W/Z + Jets results from CMS

Lukas Vanelderen (Universiteit Gent)
on behalf of the CMS Collaboration

April 13, 2011

introduction

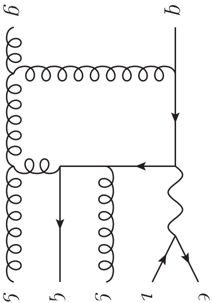
W/Z + jets

- clean environment to probe QCD
- important bkg to searches

studies on the 36pb^{-1} from 2010 presented here:

- W/Z + jets
- Z+b
- boosted W polarization

W/Z + Jets



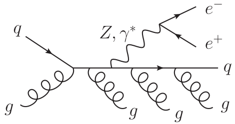
- combination of QCD and EWK
- interesting topology
  - jets
  - leptons
  - MET
- $\Rightarrow$ important BKG in searches

- how well are these processes understood?
- how much can we rely on simulation?

W/Z + Jets

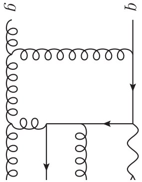
ME + PS simulation:

- tree level only
- includes non-perturbative corrections
- most common method



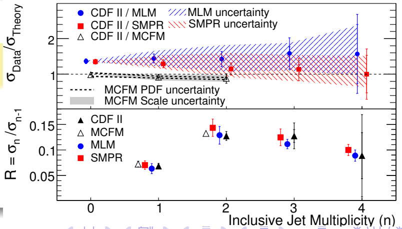
NLO calculations:

- recent development
- not yet widely used



Tevatron results:

- can we do better?
  (shown: W + jets)



W/Z candidate selection

single electron or muon trigger

first lepton

electron / muon

- tight selection
- high purity
- caused trigger

second lepton

same flavor

- loose selection
- high efficiency

1 lepton $\Rightarrow$ W candidate
2 leptons $\Rightarrow$ Z candidate

electron selection

- **single electron trigger** (threshold below 17 GeV)
- **find first electron**
  - $p_T > 20\text{GeV}$, $|\eta| < 2.5$, $1.44 < |\eta| < 1.57$ excluded
  - matches trigger primitive
  - tight isolation, ID, conversion rejection (80% efficiency)
- **search second electron**
  - $p_T > 10\text{GeV}$, $|\eta| < 2.5$, $1.44 < |\eta| < 1.57$ excluded
  - $60 < M_{ee} < 120\text{GeV}$
- **W candidates**
  - no muon with $p_T > 10\text{GeV}$ (top veto)

electron selection

- **single electron trigger** (threshold below 17 GeV)
- **find first electron**
  - $p_T > 20\text{GeV}$, $|\eta| < 2.5$, $1.44 < |\eta| < 1.57$ excluded
  - matches trigger primitive
  - tight isolation, ID, conversion rejection (80% efficiency)
- **search second electron**
  - $p_T > 10\text{GeV}$, $|\eta| < 2.5$, $1.44 < |\eta| < 1.57$ excluded
  - $60 < M_{ee} < 120\text{GeV}$
- **W candidates**
  - no muon with $p_T > 10\text{GeV}$ (top veto)

results are quoted in this acceptance

muon selection

- **single muon trigger** (threshold below 15 GeV)
- **find first muon**
  - $p_T > 20\text{GeV}$, $|\eta| < 2.1$
  - matches trigger primitive
  - tight isolation, ID
- **search second muon**
  - $p_T > 10\text{GeV}$, $|\eta| < 2.4$
  - $60 < M_{\mu\mu} < 120\text{GeV}$

muon selection

- **single muon trigger** (threshold below 15 GeV)
- **find first muon**
  - $p_T > 20\text{GeV}$, $|\eta| < 2.1$
  - matches trigger primitive
  - tight isolation, ID
- **search second muon**
  - $p_T > 10\text{GeV}$, $|\eta| < 2.4$
  - $60 < M_{ee} < 120\text{GeV}$

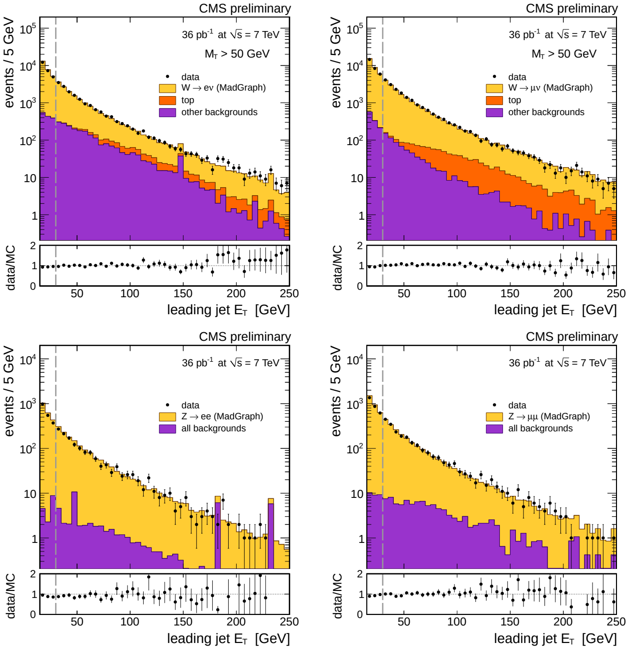
results are quoted in this acceptance

Jets

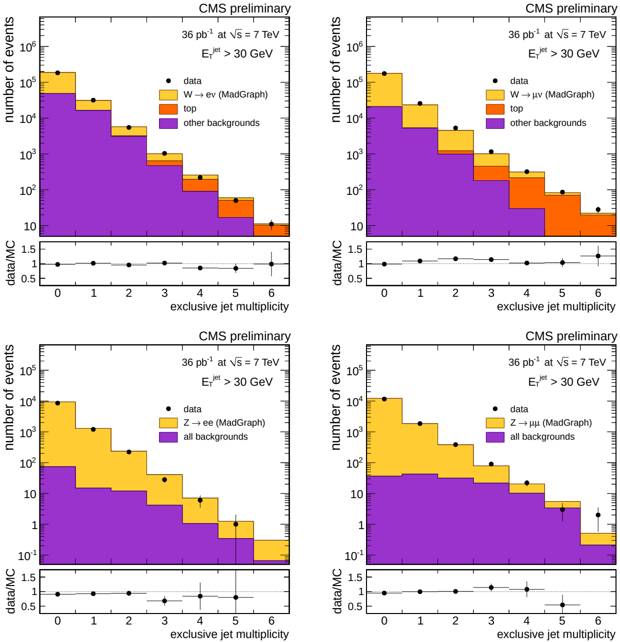
cluster "Particle Flow" objects, Anti-kt, $\Delta R = 0.5$

- $|\eta| < 2.4$ (tracker acceptance), $E_T > 30$GeV
- data driven jet energy calibration
- pile-up removal with FastJet
- muons from W/Z candidates:
  removed from particle list before clustering
- electrons from W/Z candidates:
  veto jets within $\Delta R < 0.3$

Jets: data vs MC



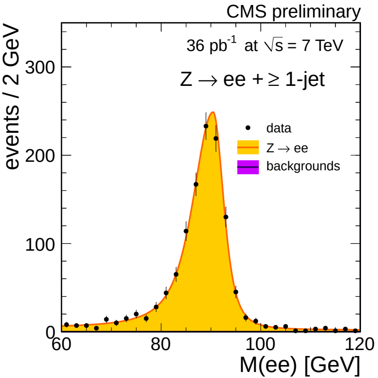
Jets: data vs MC



signal extraction: Z

Unbinned Maximum Likelihood (UML) Fit on M_{ll}

- functional forms
- parametrization in events with 0, 1, 2, 3 and $>=$ 4 jets
- BKG params floated
- signal params floated, kept equal for all jet multiplicities



signal extraction: W

UML Fit on m_T and $n_{jet}^{b-tagged}$

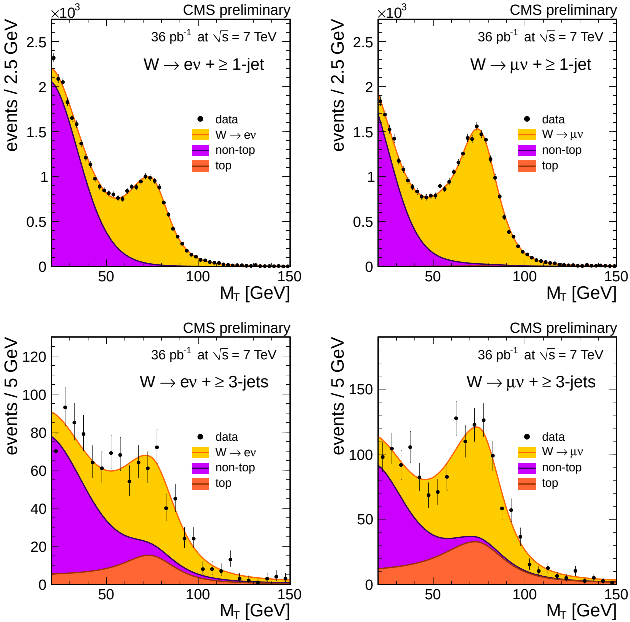
- functional forms
- parametrization in events with 0, 1, 2, 3 and >= 4 jets

signal vs QCD

- $m_T = \sqrt{2p_T^l \cdot MET \cos(\Delta\Phi(l, MET)},$
- MET from "Particle Flow" objects
- fit in range $20 < m_T < 150$
- 0 jets: high stats, important parameters floated

  ...
- 4 jets: low stats, few params floated (electron/muon)

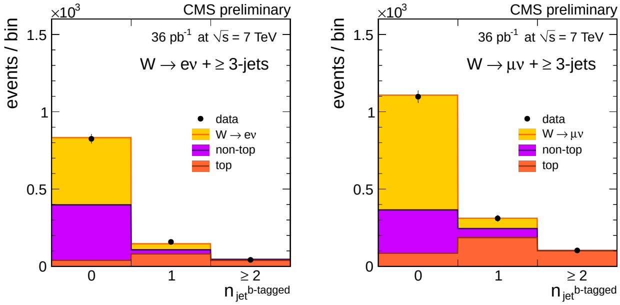
signal extraction: W



signal extraction: W

signal vs top

- $n_{jet}^{b-tagged}$
- b-tag and mistag eff measured in control samples
- top contribution floated!



corrections

lepton efficiency

measured with tag and probe on Z events, factorizes as:

- reconstruction
  (cluster $\rightarrow$ electron, track $\rightarrow$ muon)
- selection
  (differs between 1st and 2nd lepton, n_{jet} dependence)
- trigger (only for 1st lepton)

fit range, $m_T > 20$

from MC, verified on data in Z events

unfolding jet multiplicity spectrum

- extract migration matrix R(n^{RECO} ,n^{GEN}) from MC
- singular value decomposition (SVD) to "unsmear" n_{jet} distribution

corrections

measure ratios

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

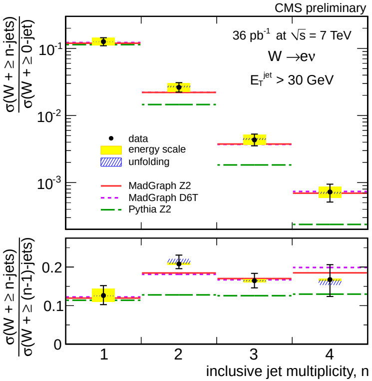
$\Rightarrow$ reduce systematics

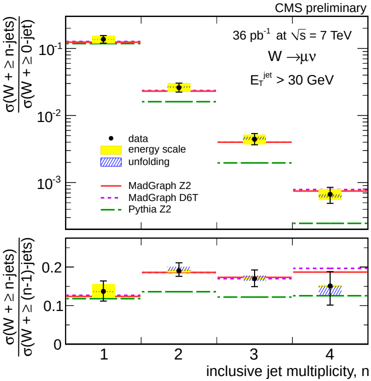
- partial cancelation JES
- full cancelation luminosity
- only lepton efficiency vs n_{jets}

...

results for W

$p_T > 30\text{GeV}$

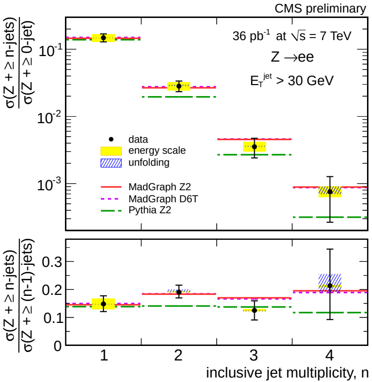


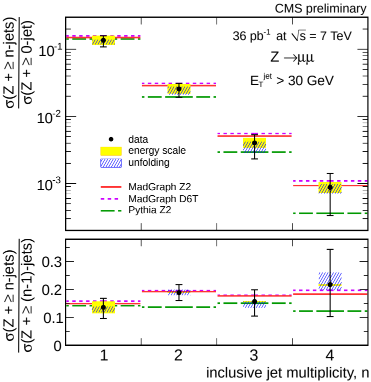


- very good agreement with predictions from ME+PS
- negligible difference between tunes for $p_T > 30\text{GeV}$
- PS alone starts to fail for $n_{jet} \geq 2$

results for Z

$p_T > 30\text{GeV}$



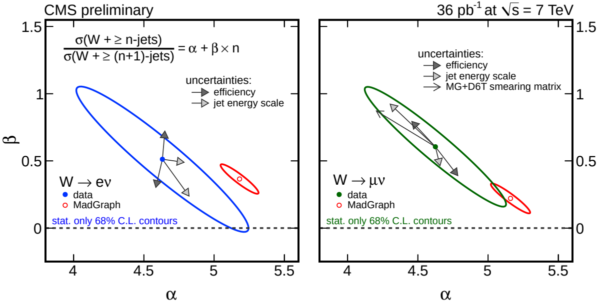


- excelent agreement with expectations from ME+PS
- PS alone is also compatible with data

Berends-Giele scaling

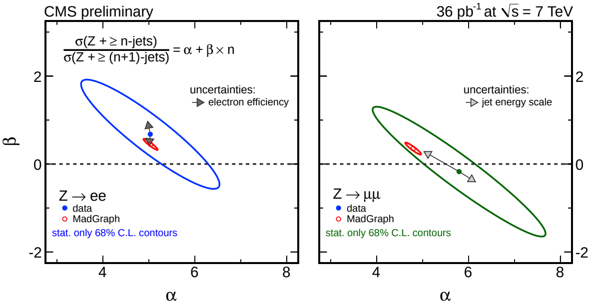
- test scaling by fitting $C_n = \frac{\sigma_n}{\sigma_{n+1}} = \alpha + \beta n$
- take into account correlations between σ_n
- take into account migration between jet bins

W



Berends-Giele scaling

Z



reasonable agreement between ME + PS expectation

Z + b-jets

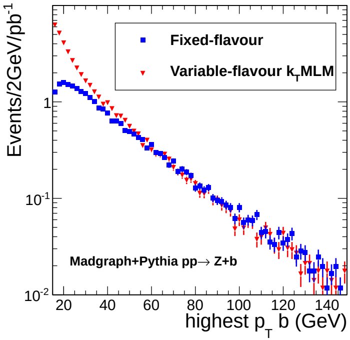
- benchmark channel for MSSM Higgs searches
- fixed vs variable flavour number schemes (LO only)

selection

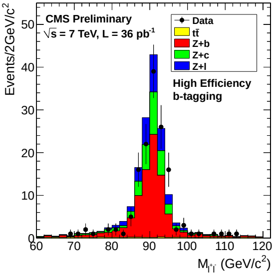
- At least one Z
- At least one >25GeV jet
- $\geq$ 1 secondary vertex in jet
- MET < 40GeV (top rejection)

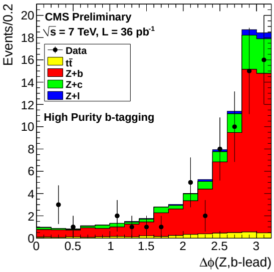
2 ways of b-tagging

- high purity
- high efficiency



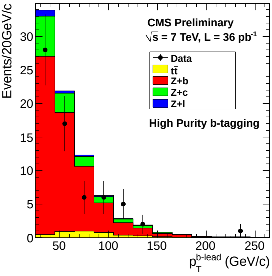
Z + b-jets

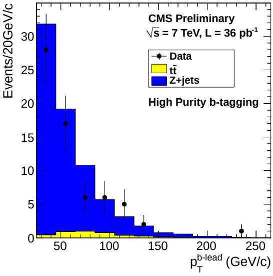




fixed flavor

variable flavor



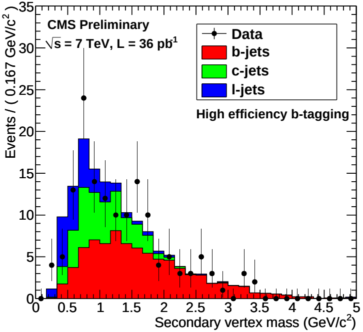


- Good agreement data vs ME+PS

- no distinction between schemes
- low stats were schemes disagree

Z+b/Z+jet ratio

- Z+b purity extracted from fit on secondary vertex mass
- results compatible with MadGraph(*) and MCFM NLO calculations
- (*) Z+b and Z+c with p_T,jet > 15 GeV scaled to corresponding MCFM x-sec



$$\text{Ratio } \mathcal{R} = \frac{\sigma(pp \to Z+b+X)}{\sigma(pp \to Z+j+X)}$$

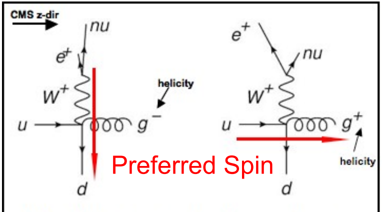
Sample	$\mathcal{R}(Z \to ee)$ (%), $p_T^e > 25$ GeV, $\lvert\eta^e\rvert < 2.5$	$\mathcal{R}(Z \to \mu\mu)$ (%), $p_T^\mu > 20$ GeV, $\lvert\eta^\mu\rvert < 2.1$
Data HE	$4.3 \pm 0.6(stat) \pm 1.1(syst)$	$5.1 \pm 0.6(stat) \pm 1.3(syst)$
Data HP	$5.4 \pm 1.0(stat) \pm 1.2(syst)$	$4.6 \pm 0.8(stat) \pm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

boosted W polarization

production of high p_T W bosons ($p_T > 50\text{GeV}$)

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

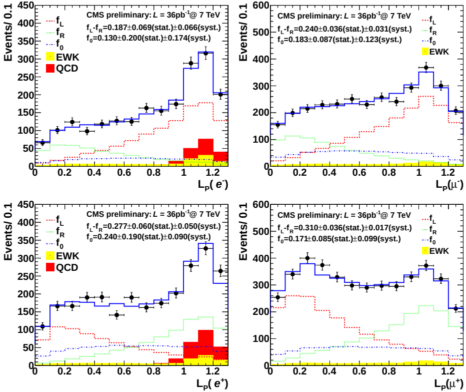
Expect left right polarization asymmetry in a pp collider

boosted W polarization

- z component ν not measured
  θ^* not available
- use instead

$$L_P = \frac{\vec{p}_T(l)\cdot\vec{p}_T(W)}{|\vec{p}_T(W)|^2} \approx 0.5cos(\theta^*) + 0.5$$

- template fit to extract polarization



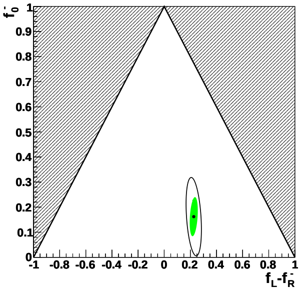
W^-

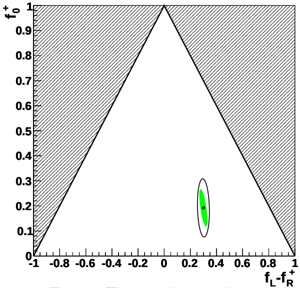
W^+

boosted W polarization results

- systematics dominated by MET uncertainty
- $f_L - f_R > 0 \Rightarrow$ mostly left-handed

	Combined Results
$(f_L\text{-}f_R)^-$	0.226 ± 0.031 (stat) ± 0.050 (syst)
f_0^-	0.162 ± 0.078 (stat) ± 0.136 (syst)
$(f_L\text{-}f_R)^+$	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
- Observation of Z + b and ratio Z + b / Z + jets agrees well with NLO calculation
- Measured significant polarization of boosted W

PART II:
V+jets as bkg in searches in CMS

Lukas Vanelderen (Universiteit Gent)
on behalf of the CMS collaboration

Aprils 7, 2011

CMS has developed and uses a large variety of methods to estimate V+jets bkg in searches

picking out a few of the most interesting examples from

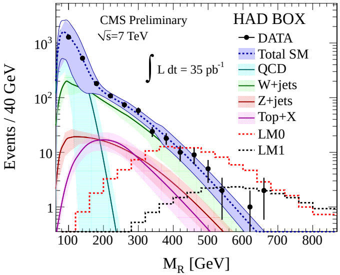
- "razor" analysis
- multi jet + MET analysis
- single lepton + jets + MET analysis

Razor Analysis

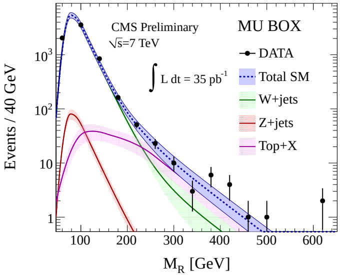
inclusive search for heavy squark pair production based on 2 variables (arXiv:1006.2727) :

- M_R: "search variable"
  $\sim$ per event estimate for $\frac{M_{\tilde{q}}^2 - M_{\tilde{\chi}}^2}{M_{\tilde{q}}}$
- R: cleans up bulk of QCD
  related to mass scale for signal, and to MET
  $[0, \sim 1]$, low value $\rightarrow$ fake/low MET

hadronic box, $R > 0.50$



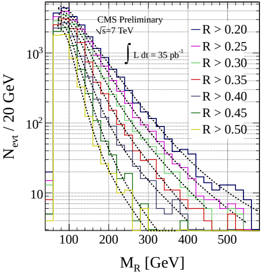
muon box, $R > 0.45$



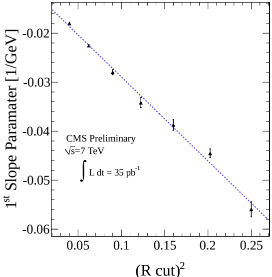
Razor Analysis

? shape W + jets given cut on R ?

- $W \to \nu$ from “lepton boxes”
- bkg dominated: low R, M_R region
- fit 2 exponentials
  - exponential parameters from fit
  - contributions from fit



- slopes are linear function of R^2
- $\to$ parameters a, b
  for 1st and 2nd slope



Razor Analysis

- **$W \to l\nu$ with l in acceptance OK**
- **?$W \to l\nu$ with l out acceptance?**
  calculate a MC $\to$ data conversion factor in "lepton box"

$$\rho(a)_1^{\mathrm{DATA/MC}} = 0.97 \pm 0.02$$
$$\rho(b)_1^{\mathrm{DATA/MC}} = 0.97 \pm 0.02$$

- **?$Z \to \nu\nu$?**
  Correct slopes in MC to slopes measured in "lepton box" with lepton treated as neutrino

multi jet + MET analysis

- ≥ 3 jets ($p_T > 50$GeV, $|\eta| < 2.5$)
- $HT > 300$GeV
  scalar sum jet p_T
- $MHT > 150$GeV
  magnitude vectorial sum jets,
  $p_T > 30$GeV, $|\eta| < 5$
- muon and electron veto

V+jets bkg

- $Z \to \nu\nu$
- $W \to e\nu$, $W \to \mu\nu$, $W \to \tau\nu \to l\nu\nu$
  with leptons out of acceptance
- $W \to \tau\nu \to had + \nu\nu$

multi jet + MET analysis

$Z \to \nu\nu$

from $Z \to ll$

clean, but low statistics

- nominal selection + lepton id + $60\mathrm{GeV} < M_{ll} < 120\mathrm{GeV}$
- observed 1(2) events, muons(electrons)
- **too few statistics**

multi jet + MET analysis

$Z \to \nu\nu$

from γ + jets

Z and γ production similar in many ways for high boson p_T

- measure γ + jets
  nominal selection + hard isolated photon
- derive γ to Z conversion factor
  - Z/γ correction
    MADGRAPH γ + jets and Z + jets
    uncertainty: LO vs NLO MC
  - fragmentation
    contribution strongly suppressed by γ isolation
    remainder calculated with NLO JetPhox
  - secondary photons (neutral pions and η meson) data driven

multi jet + MET analysis

$Z \to \nu\nu$

from γ + jets

Z and γ production similar in many ways for high boson p_T

- measure γ + jets
- derive γ to Z conversion factor

		Baseline selection		High-MHT selection		High-HT selection	
Z/γ correction	±theory	0.41	±6 %	0.48	±6 %	0.44	±4 %
	±acceptance		±5 %		±5 %		±5 %
	±MC stat		±7 %		±13 %		±13 %
Fragmentation		0.95	±1 %	0.95	±1 %	0.95	±1 %
Secondary photons		0.94	±9 %	0.97	±10 %	0.90	±9 %
Photon mistag		1.00	±1 %	1.00	±1 %	1.00	±1 %
ID data/MC ratio		1.01	±2 %	1.01	±2 %	1.01	±2 %
Total correction		0.37	±14 %	0.45	±18 %	0.38	±17 %

multi jet + MET analysis

$Z \to \nu\nu$

from γ + jets

Z and γ production similar in many ways for high boson p_T

- measure γ + jets
- derive γ to Z conversion factor

	# events in γ+jets data sample	# $Z \to \nu\nu$ events predicted	# $Z \to \nu\nu$ events from simulation
Baseline selection	72	$26.3 \pm 3.2(\mathrm{stat.}) \pm 3.6(\mathrm{syst.})$	21.2 ± 1.4
High-MHT selection	16	$7.1 \pm 1.8(\mathrm{stat.}) \pm 1.3(\mathrm{syst.})$	6.3 ± 0.8
High-HT selection	22	$8.4 \pm 1.8(\mathrm{stat.}) \pm 1.4(\mathrm{syst.})$	5.8 ± 0.7

multi jet + MET analysis

$W \to e\nu$, $W \to \mu\nu$, $W \to \tau\nu \to l\nu\nu$

treated together with bkg from top

some leptons pass veto

- outside acceptance
- fail trigger
- fail reconstruction, iso, ID

- nominal selection + μ, Control Selection (CS)
- reweight to events with "lost lepton"

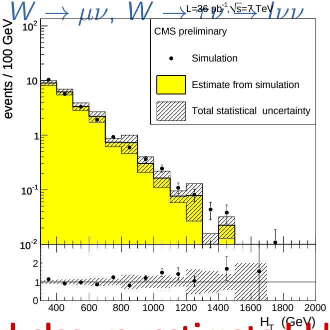
multi jet + MET analysis

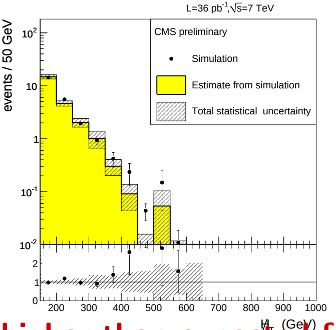
$W \to e\nu$, $W \to \mu\nu$, $W \to \tau\nu \to l\nu\nu$

- identified electrons/muons failing isolation
  $!ISO^{e,\mu} = CS \cdot \frac{1-\epsilon_{ISO}^{e,\mu}}{\epsilon_{ISO}^{\mu}}$
  from TnP on Z,
- electrons or muons in acceptance, failing id
  $!ID^{e,\mu} = CS \cdot \frac{1}{\epsilon_{ISO}^{\mu}} \frac{1-\epsilon_{ID}^{e,\mu}}{\epsilon_{ID}^{\mu}}$
- $\epsilon_{ISO}^{e,\mu}$ and $\epsilon_{ID}^{e,\mu}$ from TnP on Z
  $\epsilon_{ISO}^{e,\mu}$ vs $p_T^{e,\mu}$ and $\eta^{e,\mu}$
  $\epsilon_{ID}^{e,\mu}$ vs $p_T^{e,\mu}$ and $\Delta R(jet)^{e,\mu}$
  impact other differences in η and p_T spectrum small
  evaluated on MC, $\to$ systematics
- muon or electrons outside acceptance
  ratio R_{accept} from MC
  same control sample corrected for isolation and ID eff

multi jet + MET analysis

$W \to e\nu$, $W \to \mu\nu$, $W \to \tau\nu \to l\nu\nu$





good closure, estimated bkg higher than expected from MC

not so far off within uncertainties

Method	Baseline selection	(stat.)	(syst.)	High-MHT selection			High-HT selection		
Estimate from data	33.0	± 5.5	$^{+6.0}_{-5.7}$	4.8	± 1.8	$^{+0.8}_{-0.6}$	10.9	± 3.0	$^{+1.7}_{-1.7}$
Estimate (PYTHIA)	22.9	± 1.3	$^{+2.7}_{-2.6}$	3.2	± 0.4	$^{+0.5}_{-0.5}$	7.2	± 0.7	$^{+1.1}_{-1.1}$
MC Truth (PYTHIA)	23.6	± 1.0		3.6	± 0.3		7.8	± 0.5	
Estimate (MADGRAPH)	20.4	± 1.5	$^{+2.6}_{-2.5}$	2.4	± 0.3	$^{+0.3}_{-0.3}$	4.8	± 0.4	$^{+0.6}_{-0.5}$
MC Truth (MADGRAPH)	21.4	± 0.7		3.0	± 0.3		5.9	± 0.4	

multi jet + MET analysis

$W \to e\nu$, $W \to \mu\nu$, $W \to \tau\nu \to l\nu\nu$

uncertainties dominated by statistics

	Relative size		# events	
Statistics of control-sample	−17%	+17%	−5.5	+5.5
Iso- & id- efficiencies (statistical)	−13%	+14%	−4.1	+4.7
Kinematic differences $t\bar{t}$, W, Z-samples	−10%	+10%	−3.3	+3.3
SM background in control-region	−3%	+0%	−1.0	+0
MC use for acceptance calculation	−5%	+5%	−1.7	+1.7
Total	−24%	+25%	−7.9	+8.1

multi jet + MET analysis

$W \to \tau\nu \to had + \nu\nu$

treated together with bkg from top

- same muon control sample
- replace muon by simulated tau jet

corrections

- kinematic and geometrical acceptance
  apply procedure on MC with and without muon selection
- muon trigger, reconstruction and isolation eff
  same procedure as for "lost leptons" method
- relative branching ratio
  $W \to \mu\nu$ vs $W \to \tau\nu \to had + \nu\nu$

multi jet + MET analysis

$W \to \tau\nu \to had + \nu\nu$

Systematic uncertainties

	Baseline selection	High-MHT selection	High-HT selection
τ response template	2%	2%	2%
Acceptance	+6%,-5%	+6%,-5%	+6%,-5%
Muon efficiency on data	1%	1%	1%
SM backgr. subtraction	5%	5%	5%

Results

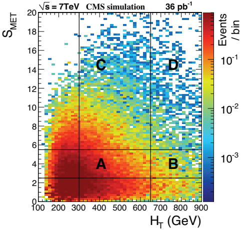
	Predicted W / $t\bar{t} \to \tau_{\mathrm{hadr}}$
Baseline selection	22.3 $\pm$4.0 (stat.) $\pm$2.2 (syst.)
High-MHT selection	6.7 $\pm$2.1 (stat.) $\pm$0.5 (syst.)
High-HT selection	8.5 $\pm$2.5 (stat.) $\pm$0.7 (syst.)

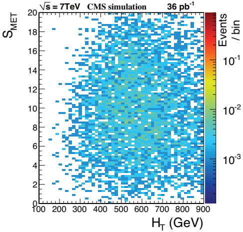
uncertainties dominated by statistics

single lepton + jets + MET analysis

- main bkg: V+jets and top
- 2 uncorrelated variables: $\mathcal{S}_{MET} = MET/\sqrt{(H_T)}$ and H_T
- predict bkg in signal region with ABCD method

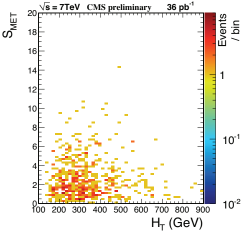
MC: total SM in μ channel

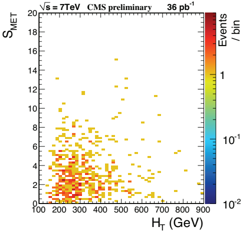




MC: μ channel SUSY LM1

Data: e channel





Data: μ channel

correlation effects studied in MC, non observed, systematic

Summary

- wide diversity of methods
- wherever possible based on measurements
- quantities measured in control region converted to signal region
  - conversion factors from measurements
  - conversion factors from MC
- predictions from MC scaled
  - scaling factors from data

assumptions based on MC, taken into account in systematics
uncertainty in methods still dominated by statistics