Tests of Vector Boson plus Jets Production at the LHC

CMS VBF Z+2 jet candidate

CTEQ workshop
November 14, 2013
Jeffrey Berryhill (Fermilab)
On behalf of the CMS and ATLAS collaborations
Photon plus Jets production
Inclusive photon cross section

80 GeV photon trigger, with 2 M photons in the 100-1000 GeV range for $|\eta| < 2.37$, with 4.6/fb at 7 TeV

Tight shower shape selection plus isolation ET < 7 GeV in a 0.4 cone

Background isolation rejection estimated from non-tight photons (vs. ET)

>93% purity for ~80% efficiency

6-7% systematic uncertainty is dominant for total cross section estimate > 100 GeV (~25% > 800 GeV)
Inclusive photon cross section

Comparison with JETPHOX 1.3:
MC/Data ratio dominated by MC scale uncertainty, 12-20% (except at highest ET where PDFs matter too)

Below 500 GeV data are 2X more precise than NLO theory

Shape described well, normalization agrees within scale uncertainty

Comparison with LO+PS:
Pythia and Herwig shapes agree well.

Herwig norm undershoots data by 10-20%
Photons selected in the 40-300 GeV range for $|\eta| < 2.5$, jet ET $> 30$ GeV with $|\eta| < 2.5$, with 2.1/fb at 7 TeV

>20-70% purity for (not strictly isolated) photons at ~70-90% efficiency

~5-8% systematic uncertainty is dominant for cross section estimates near 100 GeV
Comparison with JETPHOX 1.2.2:
Shape and norm within 10-20% scale uncertainty band
Data ~2x more precise than NLO MC

Comparison with SHERPA 1.3.1 (up to 3 jets):
Similar level of agreement to JETPHOX
Prediction defined w.r.t. 5 GeV hadronic isolation in a cone of 0.4
Photon+jet dynamics

JETPHOX, HERWIG fail to describe shape of the $\Delta \phi < 2$ region at the 20% level.

PYTHIA, SHERPA succeed.

$\cos \theta$, mass distributions are in good agreement.

arxiv:1307.6795
**Photon+jets: Mini-prospectus**

**7 TeV conclusions:** LO+PS and NLO are doing mostly OK at the 20% scale uncertainty level compared to the 7 TeV photon data. There are some exceptional phase space regions.

7 TeV data are at <10% precision so 2X theory improvements can be readily confronted (NLO+PS, NNLO?)

**With 8 TeV data:** TeV scale cross sections will ~double in precision to the 10% level, and extend differentially beyond 1 TeV

At lower scales multiply differential distributions can be more completely explored.

**Unexplored so far at LHC:**
- Photon + multijet production
- Photon + heavy flavor
- Diphoton or V+photon + multijet
- VBF photon, diphoton, or V+photon + 2 jet
W, Z plus
Jets
production
Z+jets diff. cross section

\~500k Z candidates selected in 4.6/fb at 7 TeV with \( \geq 1 \) AK4 jet with PT>30 GeV, \( |y| < 4.4 \); **up to 7 jets observed**

\( \tt \tt \) background at 20% level for 6 jets (est. from emu)

JES systematic uncertainty dominates (8% for \( N \geq 1 \), 20% for \( N \geq 4 \))

Scale uncertainty dominates NLO error (4-13% for \( N=1-4 \))

**Data error is 2X NLO theory**

NLO and LO+PS describe data well where applicable (\( N=4 \) and 5, resp.)
Z+jets diff. cross section

Leading Jet PT probed in bins out to 700 GeV

Data unc. < NLO MC unc. for PT>= 100 GeV

BlackHAT+Sherpa describe data well.

ALPGEN, SHERPA show 20% discrepancies

Insufficient generated partons leads to worse disagreement at higher PT (MC@NLO Z+1jet)

NLO EWK is also a factor at highest PT
Z+jets diff. cross section

BlackHat+Sherpa fails to describe HT shape.

LO+PS does a somewhat better job.

Which of the following would improve the prediction the most?:

More partons
NLO+PS
NLO EWK
NNLO QCD
Jet properties explored with **Two leading jets in a “VBF” configuration:**

$\text{M}_{jj} > 350 \text{ GeV}$

$|\Delta y_{jj}| > 3$

This is still predominantly QCD Z+2 jets, so a background study for future VBF Higgs and VBS analysis

NLO and LO+PS describe data well in this regime, at the 25-50% level.
Z+jets diff. cross section

Efficiency of 3rd central jet veto as a function of 3rd jet PT threshold for VBF Z+>=2 jet

20 (7)% inefficiency observed at 30 (50) GeV central jet threshold

Agrees with LO+PS at <5% level

\[ \int L dt = 4.6 \text{ fb}^{-1} \]
\[ Z/\gamma^*(\rightarrow \mu^+\mu^-) + \geq 3 \text{ jets} \]
\[ p_T^{\text{jet}} > 30 \text{ GeV}, \, |y^{\text{jet}}| < 4.4, \]
\[ m^l > 350 \text{ GeV}, \, |y^{l\ell}| > 3.0 \]
\[ |y^{\text{jet}} (3\text{rd leading jet})| < 2.4, \]

MC / Data

min \( p_T^{\text{jet}} \) (3rd leading jet) [GeV]
With 5/fb at 7 TeV, azimuthal Z-jet and jet-jet angular shapes and transverse thrust explored, inclusively and for Z PT > 150 GeV

Madgraph, Sherpa describe data well at low and high ZPT for Njet >= 1-3

POWHEG, PYTHIA have too few partons and so have limited applicability

$DF \sim p, \ln \tau T$ $\sim -\infty$  
$\Delta \Phi \sim \pi, \ln \tau T \sim -\infty$  
$\Delta \Phi \ll \pi, \ln \tau T \sim -1$
Z+jets event shape

Madgraph describes log transverse thrust well, Sherpa shows 10-20% discrepancies

$\Delta \Phi \sim \pi$, $\ln \tau T \sim -\infty$

$\Delta \Phi \ll \pi$, $\ln \tau T \sim -1$
$Z^+=1$ jet rapidity shape

$y_Z$ and $y_{jet}$ well described by LO+PS

BUT $|y_Z + y_{jet}|/2$ and $|y_Z - y_{jet}|/2$ exhibit large discrepancies at large values.

SHERPA and Madgraph diverge from data in different directions. Attributed to parton-shower matching differences.

MCFM somewhat better in $y_{DIF}$ but still with poor Sherpa-like $y_{SUM}$
W+jets diff. cross section

W+light jets only examined in 2010 data thus far.

Similar level of agreement to Z+jets

Larger (tt) backgrounds at high Njet

Larger comparable reach in leading jet PT and HT
Towards VBF/VBS: VBF Z production

Comprehensive study of Z+forward dijet production at 7 and 8 TeV.

VBF Z one of the interfering amplitudes

Z+2jet events selected with “VBF topology”: large dijet mass, large dijet Δy

Small S/B enhanced with BDT selection exploiting all Z+2jet kinematics

5 sigma signal for electroweak Z+jet production observed, fully consistent with SM

TGC potential under study

σEWK = 226 ± 26 stat ± 35 syst fb

σVBFNLO = 239 fb

CMS-PAS-FSQ-12-035
JHEP10(2013)101
Towards VBF/VBS: VBF Z production

Multijet properties explored in EWK-enriched subsample ($M_{jj}>1250$ GeV)

Madgraph w/K factor describes central jet multiplicity well

Third jet PT also well described.
**7 TeV conclusions**: LO+PS and NLO are doing mostly OK at the 10-20% scale uncertainty level compared to the 7 TeV data. There are some exceptional phase space regions, especially in angular patterns of radiation and high HT. Data precision leads theory precision at highest PT.

**VBF-like V+2 jet production is now a detailed topic of study in its own right**

**With 8 TeV data**: Close to probing TeV PT scale and very large jet multiplicities (8 or more)

At lower scales multiply differential distributions can be more completely explored.

NLO+PS and NLO EWK effects can be tested.

**W+jets** has been neglected but has stronger sensitivity to high PT/VBF phenomena.
W, Z plus Heavy Flavor production
Z+b, bb production

12k Z+1 b-tag and 500 Z+2 b-tag events expected in 5/fb at 7 TeV.

$t\bar{t}$ suppressed by Z mass and MET significance cut, Z+light/charm jets rejected by large secondary vertex mass (MSV).

Z+b (bb) extracted from 1D (2D) template fit to MSV (MSV1, MSV2)

Exclusive 1,2-tag cross section estimated after N-tag-wise unfolding of MET, lepton, JES, and b-tag response
Z+b, bb production

Exclusive cross sections agree with MadGraph 4F and 5F predictions.

B-tag efficiency and mistag uncertainty dominate total cross sections.

Z PT in 2b case is somewhat harder than MadGraph.

Mbb and other variables in good agreement.
Using a tracker-driven inclusive vertex reconstruction technique, B-hadron pairs can be identified with excellent angular resolution→can explore very collinear production from e.g. gluon splitting.

Total cross sections predicted by MadGraph 4F/5F, ALPGEN, aMC@NLO about 15% low, worse at high ZPT.

Madgraph, aMC@NLO underpredicting lowest $\Delta R$. 

arxiv:1310.1349
W candidates with =1 or 2 jets and = 1 b-tag selected from 4.6/ fb at 7 TeV

Two different taggers with complementary info combined into an ANN discriminant against light/charm jets. 40-60% of tags retained for signal extraction via MLH template fit of ANN.

tt contribution constrained by 4-jet 1-tag sample; single top constrained by m(Wb) distribution
Exclusive cross sections measured to 20% precision (dominated by JES syst.)

W+b + single top results also presented

MCFM NLO and LO+PS predictions consistent with data at 1.5σ

Data and MC precision comparable

Diff. b PT cross sections a bit harder than MCFM/ALPGEN
W candidates with =2 jets and =2 btags selected from 5/fb at 7 TeV

W+c,cc reduced by combined cut on the MSV of the two b-tags

tt normalized by =4jet,2-tag sample

Remaining top discrimination from leading jet PT templates.

MCFM in good agreement with measured cross section.

MadGraph agrees with Mbb distribution

MCFM

0.52 ± 0.03 pb

CMS data

0.53 ± 0.05 (stat.) ± 0.09 (syst.) ± 0.06 (theo.) ± 0.01 (lum.) pb
W+charm cross section

Leading order W+c directly probes **strange quark PDF**

Strange and anti-strange probed independently by W+, W-

W and c are opposite sign

Higher-order W+cc, W+bb, top are **same-sign/opposite-sign symmetric** → subtract with same-sign data

(semi-)exclusive **charm hadron reconstruction** gives high-purity, self-charge-tagged W+c samples

---

[Diagram showing W+charm cross section]

[Graph showing mass difference with p_T > 25 GeV]

[Graph showing secondary vertex mass]

[Graph showing p_T in jet]
W+charm cross section

- Measure cross section and charged ratio vs lepton $|\eta|$
- Consistent across three different hadron reco methods
- Leading syst. are JES, charm branching fractions
- Consistent with NLO MCFM predictions

[Graphs showing data and predictions]

arXiv:1310.1138
W+charm cross section

- Data consistent with strange content of pre-LHC PDFs (neutrino fixed-target), approaching good precision
- Data consistent with charge symmetric strange PDF

arXiv:1310.1138
ATLAS has a very similar analysis with somewhat different selected phase space and cross section definitions.

Cross sections measured for charm hadrons, not partons

Compared with aMC@NLO: Minimum PT(D) > 8 GeV, |ηD|<2.2
PTl > 25 GeV, MET>25 GeV, MT> 40 GeV

Favors somewhat higher s-sea

Measures also PTD diff. cross section as well as η(lepton)
W+charm hadron cross section

CMS-ATLAS difference is between 1-2 sigma, roughly, since CMS ~ CT10 with similar error.
**W,Z+heavy flavor: Mini-prospectus**

**7 TeV conclusions:** LO+PS and NLO are doing mostly OK at the 20% scale uncertainty level compared to the 7 TeV data. There are some exceptional phase space regions. Interesting sensitivity to sea quark PDFs.

**With 8 TeV data:** differential cross sections to be explored in more detail and in regions more relevant for Higgs/searches

**Unexplored so far at LHC:**
- $Z+c(\text{cc})$
- $W+cc$
- $Z+b$ angular distributions
- Comprehensive $W+N\text{jet}+\text{Mtags}$
Summary

• V+jets production measurement has been explored in most all relevant areas with the 7 TeV data. But almost no 8 TeV results available yet!

• NLO and LO+PS are successfully describing the data at the advertised level of accuracy in a wide variety of situations. But there are exceptional distributions everywhere. Can the next generation of predictions resolve those exceptions?

• V+jets is not the only precision laboratory for multijet QCD going forward. Also look at jet dynamics:
  • In VV and VBF, relevant for precision Higgs physics
  • In phase space relevant for high PT searches