Study of di-boson WW, WZ production in W(→lν)+jj events

Jake Anderson, Jeffrey Berryhill, Dan Green,
Kalanand Mishra, Fan Yang, Francisco Yumiceva (Fermilab)
Ricardo Eusebi, Alexx Perloff (Texas A&M)
Andre Sznajder (UERJ, Brazil & LPC)
Pratima Jindal (Nebraska)
Kristina Krylova (Buffalo)

June 8, 2011

Goal of this analysis

Ultimately three things:

1. Establish Standard Model electroweak WW+WZ(→lvjj) production
   - with both 0 and 1 extra jet − if feasible
     - but focus on 0 extra jet for now

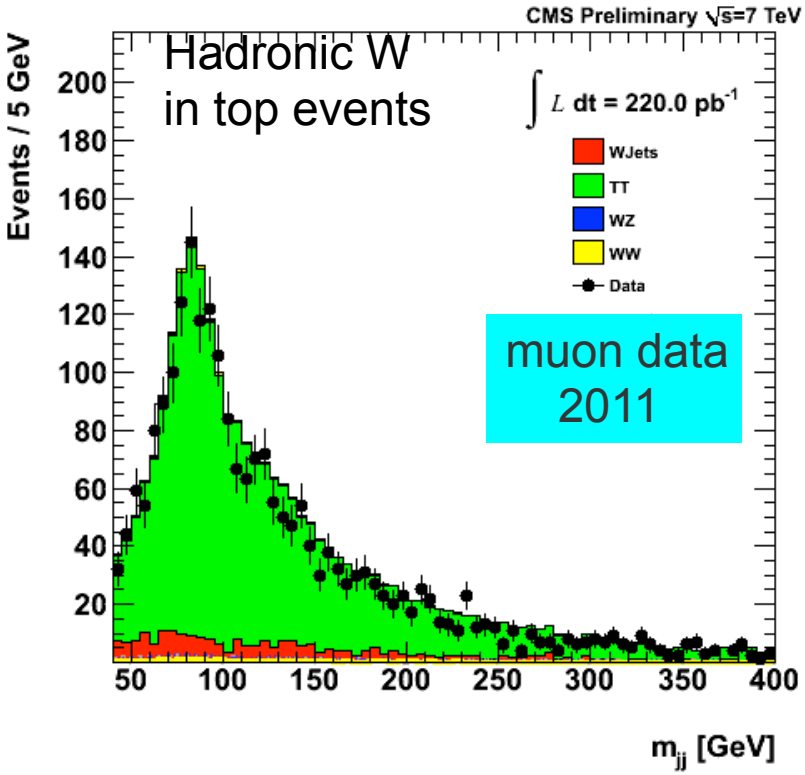
2. Measure cross section for di-boson production with this decay mode
   - crucial input for H→WW (lvjj) analysis and Wjj bump hunt

3. Understand basic features of di-boson production
   - set limit on anomalous gauge coupling

For EPS (summer conferences) focus on #1 but also try preliminary result for #2. By August aim to get #3 done and complete the analysis in time for Fall conferences.

Calibrating hadronic W from top events in data

In top events reconstruct clear W peak "out-of-box" with good resolution



Just require

- ≥4 jets above p_T 25 GeV
- 2 b-tags and
- one leptonic W in the event
  - -mu p_T > 25 GeV OR ele E_T > 30 GeV
  - -MET > 25 GeV

Plot m_{jj} of those two jets which are not b-tagged. This is a pure W→jj sample.

Resolution in hadronic W shape is dominated by jet resolution

Are we triggering on the events we need ?

Yes, but life is a little complicated ....

✦For 2010 data (36 pb^{-1}) use single lepton triggers with p_T > 17 GeV (or lower)

✦For 2011A data before technical stop (~200 pb^{-1}) still rely on single lepton triggers: Mu_24 (non isolated) and Ele_27 (CaloId, TightIso)

- So, have to go to offline cuts: mu p_T > 25 GeV, electron E_T > 30 GeV

✦For 2011 data after technical stop there is dedicated "Ele17_CentralJet30_CentralJet25_MHT20" trigger. For muon still rely on Mu_24

- go to following offline cuts to be minimally tighter than trigger:

  pfMET>20 GeV, W transverse mass > 50 GeV

  leading pf jet pt > 30 GeV, second jet pt > 25 GeV.

✦ Now also have an "inclusive" W trigger in the menu for electron: keeps electron E_T > 25 GeV (or 30 if needed), pf MET > 25 GeV, W m_T > 40 GeV.

✦Submitted proposal for TSG's consideration for Mu+jj trigger

For electron may use: (Ele+MHT+jj) || (W_inclusive). For muons: Mu_24 for now.

Data samples

Already have about 700 pb^{-1} on tape

We will analyze both muon and electron data. We already have a significant fraction of integrated luminosity to be used for summer analysis on tape.

Data 2010: Apr21 re−Reco with 4_2_X

/Mu/Run2010A-Apr21ReReco-v1/AOD
/Mu/Run2010B-Apr21ReReco-v1/AOD
/EG/Run2010A-Apr21ReReco-v1/AOD
/Electron/Run2010B-Apr21ReReco-v1/AOD

Use Apr21 re-Reco JSON
Integrated luminosity ≈ 40 pb^{-1}

Data 2011 before May technical stop: May10 re−Reco with 4_2_X

/SingleMu/Run2011A-May10ReReco-v1/AOD
/SingleElectron/Run2011A-May10ReReco-v1/AOD

Use May20 prompt JSON
Integrated luminosity ≈ 200 pb^{-1}

Data 2011 after May technical stop: prompt Reco with 4_2_X

/SingleMu/Run2011A-PromptReco-v4/AOD
/SingleElectron/Run2011A-PromptReco-v4/AOD
/ElectronHad/Run2011A-PromptReco-v4/AOD

Use latest prompt JSON
Integrated luminosity ≈ 500 pb^{-1}

MC samples

Need to re-weight the MC to match PU in data

4.2.X Summer11 samples needed for this analysis are not produced yet. Physics management recommends using 4.1.X Spring11 samples for summer analysis.

W+ jets (MadGraph)

/WJetsToLNu_TuneZ2_7TeV-madgraph-tauola/Spring11-PU_S1_START311_V1G1-v1/AODSIM

WW, WZ (Pythia6)

/WWtoAnything_TuneZ2_7TeV-pythia6-tauola/Spring11-PU_S1_START311_V1G1-v1/AODSIM
/WZtoAnything_TuneZ2_7TeV-pythia6-tauola/Spring11-PU_S1_START311_V1G1-v1/AODSIM

Top (Powheg, MadGraph)

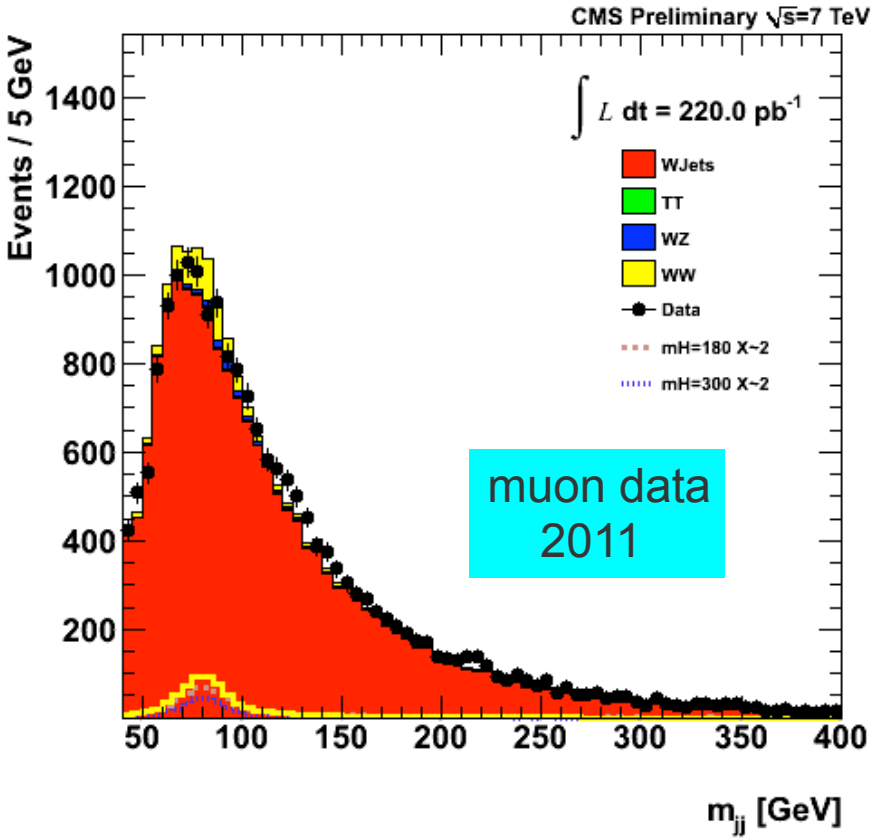
/TTToLNu2Q2B_7TeV-powheg-pythia6/Spring11-PU_S1_START311_V1G1-v1/AODSIM
/TTJets_TuneZ2_7TeV-madgraph-tauola/Spring11-PU_S1_START311_V1G1-v1/AODSIM
/TToBLNu_TuneZ2_s-channel_7TeV-madgraph/Spring11-PU_S1_START311_V1G1-v1/AODSIM
/TToBLNu_TuneZ2_t-channel_7TeV-madgraph/Spring11-PU_S1_START311_V1G1-v1/AODSIM
/TToBLNu_TuneZ2_tW-channel_7TeV-madgraph/Spring11-PU_S1_START311_V1G1-v1/AODSIM

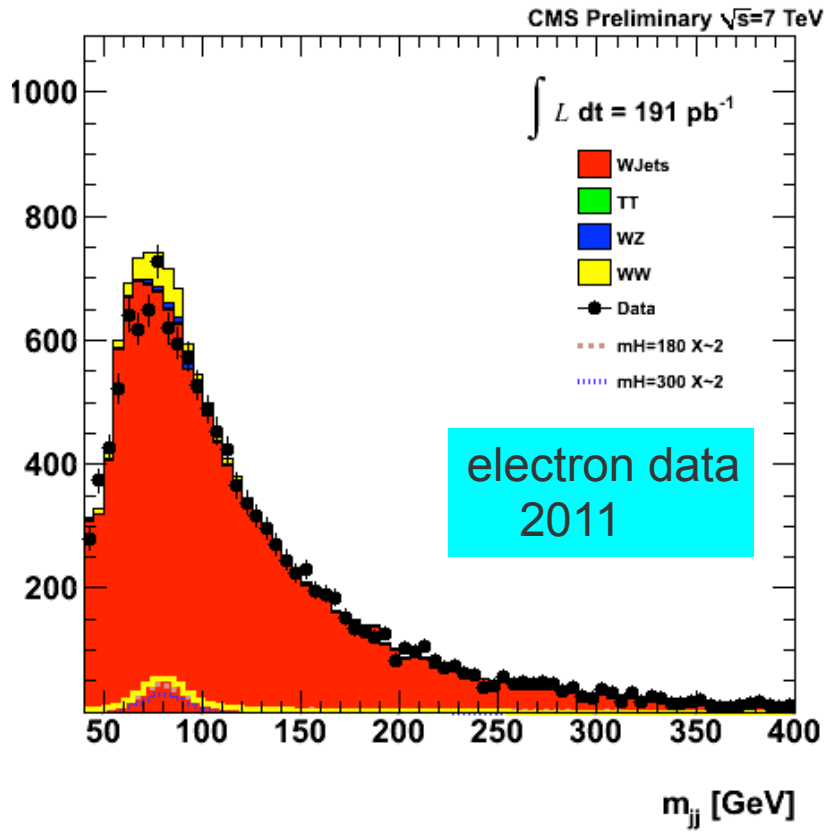
QCD multi-jets (Pythia6)

/QCD_Pt-20to30_EMEnriched_TuneZ2_7TeV-pythia6/Spring11-PU_S1_START311_V1G1-v1/AODSIM
/QCD_Pt-80to170_EMEnriched_TuneZ2_7TeV-pythia6/Spring11-PU_S1_START311_V1G1-v1/AODSIM
/QCD_Pt-20_MuEnrichedPt-15_TuneZ2_7TeV-pythia6/Spring11-PU_S1_START311_V1G1-v1/AODSIM
/QCD_Pt-20to30_BCtoE_TuneZ2_7TeV-pythia6/Spring11-PU_S1_START311_V1G1-v1/AODSIM
/QCD_Pt-30to80_BCtoE_TuneZ2_7TeV-pythia6/Spring11-PU_S1_START311_V1G1-v1/AODSIM

First look at 2011 data (I)

Dijet invariant mass in W+jj events out-of-box. Just require a tight lepton (μ or e passing top tight isolation / WP80), pf MET > 20 GeV, exactly 2 pf jets > 25 GeV.

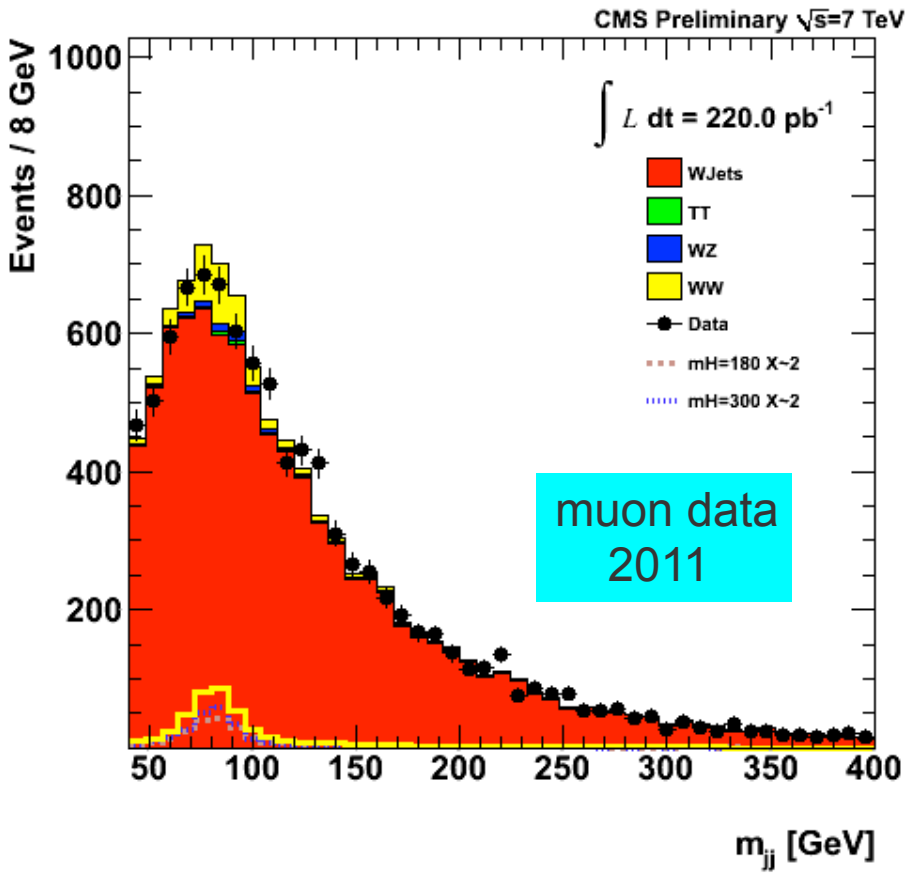


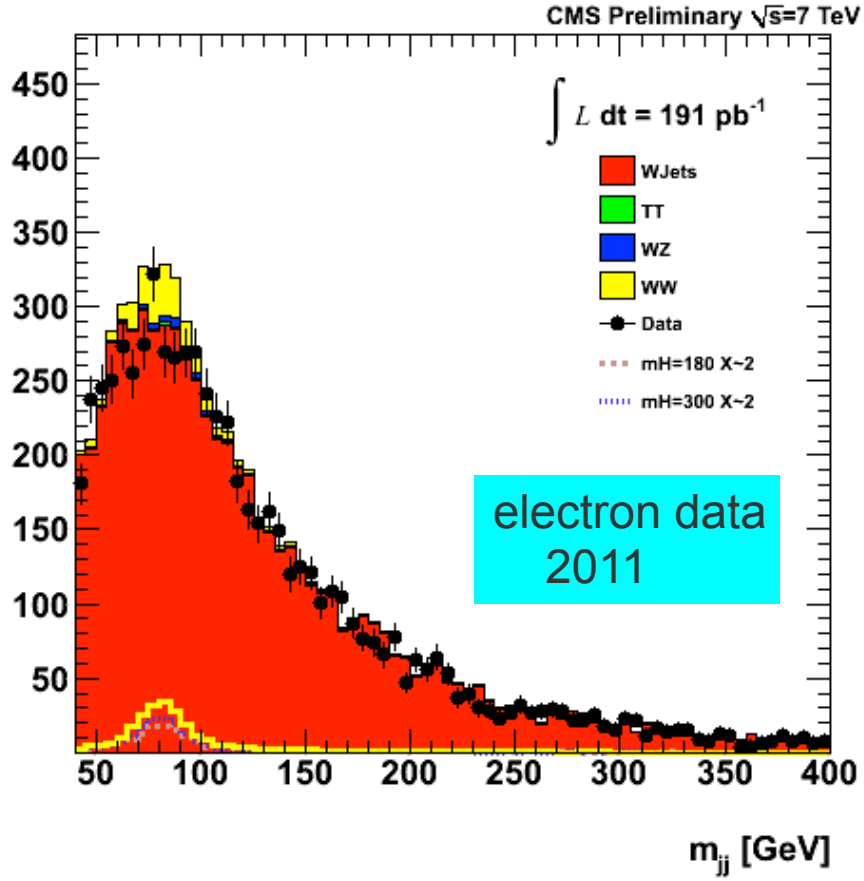


First look at 2011 data (II)

Dijet invariant mass in W+jj events after a simple $|\Delta\phi\ (W_{l\nu}, W_{jj}) - \pi| < 0.2$ cut

Start seeing the emergence of hadronic W peak from the W+jets bkg.





Analysis steps: task list

step 0: Use baseline W selection using recipe from top PAG

step 1: Use pfJet and pfMET as recommended by JetMET POG
- charge hadron subtraction: get rid of jets generated from PU
- fastJet area subtraction: remove contribution to jet energy from PU
- default “relative”, “absolute”, and “residual” jet corrections

step 2: Establish di-boson production using appropriate selection criteria

step 3: Converge on reasonable selection to enhance S/B if needed
step 4: Compute efficiency and acceptance for the selection used
step 5: Perform signal extraction, compute cross section

step 6: Set limit on anomalous gauge couplings

There are some inbuilt synergies, e.g., step 0 ensures that we can use lepton efficiency computed by top PAG in step 4. We can skip step 3 if we can reliably extract signal with low S/B.

Data & selection used in the following slides

Acceptance

- Tight lepton selection from top PAG
- Exactly two jets with p_T > 25 GeV (using PF2PAT cleaning)
- pf MET > 25 GeV
- W transverse mass > 50 GeV

Additionally apply b-veto on jets to reject top (SSV-HE-Medium)

4 kinematic cuts to suppress W+jets:

With just a single cut the S/B is low enough that the fit runs into instability. With three additional cuts S/B ~1/5.

- $|\Delta\phi(W, W) - \pi| < 0.3$ (our original single cut)
- $\Delta\eta(j1, j2) < 1.8$
- $\Delta\phi$ (j1, muon) > 2.1, $\Delta\phi$ (j1, electron) > 1.8
- $\Delta\phi$ (j2, muon) > 2.3, $\Delta\phi$ (j2, electron) > 2.0

These cuts are not necessarily optimal or final. Alexx Perloff and KM are working on this.

See Alexx's talk in Wjj forum

Processed ~ 350 pb^{-1} of data so far (340 pb^{-1} for electron, 360 pb^{-1} for muon). Still use 4.1.X MC. Big concern: MadGraph W+jets MC is only about 0.4 fb^{-1}, observe same statistical jittering in MC as in data. Hard to get good template.

Signal extraction procedure

✦Currently take W+jets background shape from MC: float absolute normalization in the fit to data. Insufficient MC statistics is the biggest hurdle. Not an ideal situation.

✦Plan to extract W+jets background shape from data-driven technique: either by inverting the selection criteria or by relaxing them. We need to try both. Float the absolute normalization in data fit.

✦Take di-boson shape from MC. EWK di-boson shapes are well computed in MC, so not a problem. We have further validated the resolution using top events. The absolute normalization will come from fit to data.

Will fine tune the technique as we go along

We take m_{jj} shape from MC

Problem

We do not have large enough W+jets MC sample to make a good template. The MadGraph sample corresponds to 700 pb^{-1} which is only ~ 2 times larger than our data size. Once we process full 0.6 fb^{-1}, the MC and data will have about the same statistics. This creates large statistical jitter if one takes shape from a simple uniformly-binned histogram of MC events.

Current solution

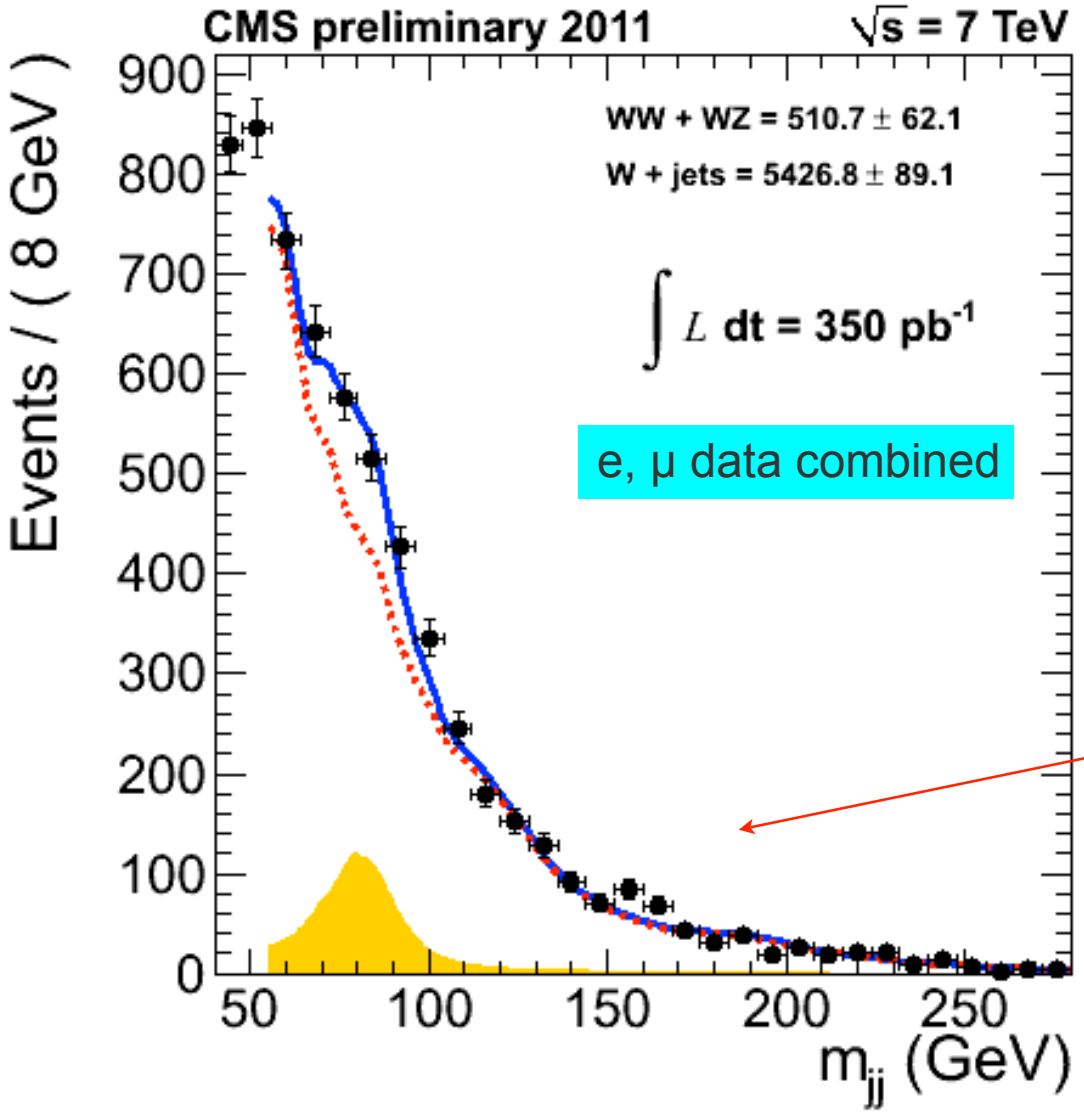
Instead of using fixed bin histograms to derive templates, I use a ROOT functionality called 'RooKeysPdf'. This class is useful if one has to deal with histograms with poor statistics and the trade-offs between having too large bins and having spikes in the plots. It's a class that behaves like a histogram, but internally saves the un-binned events and finally produces a smooth histogram.

Documentation of RooKeysPdf: http://root.cern.ch/root/htmldoc/RooKeysPdf.html

CMS Higgs combination group also uses this class for templates

see for example: HiggsAnalysis/CombinedLimit/interface/TH1Keys.h

Template fit to m_{jj} in W+2 jet events



MC predicts > ~ 400 WW + WZ events

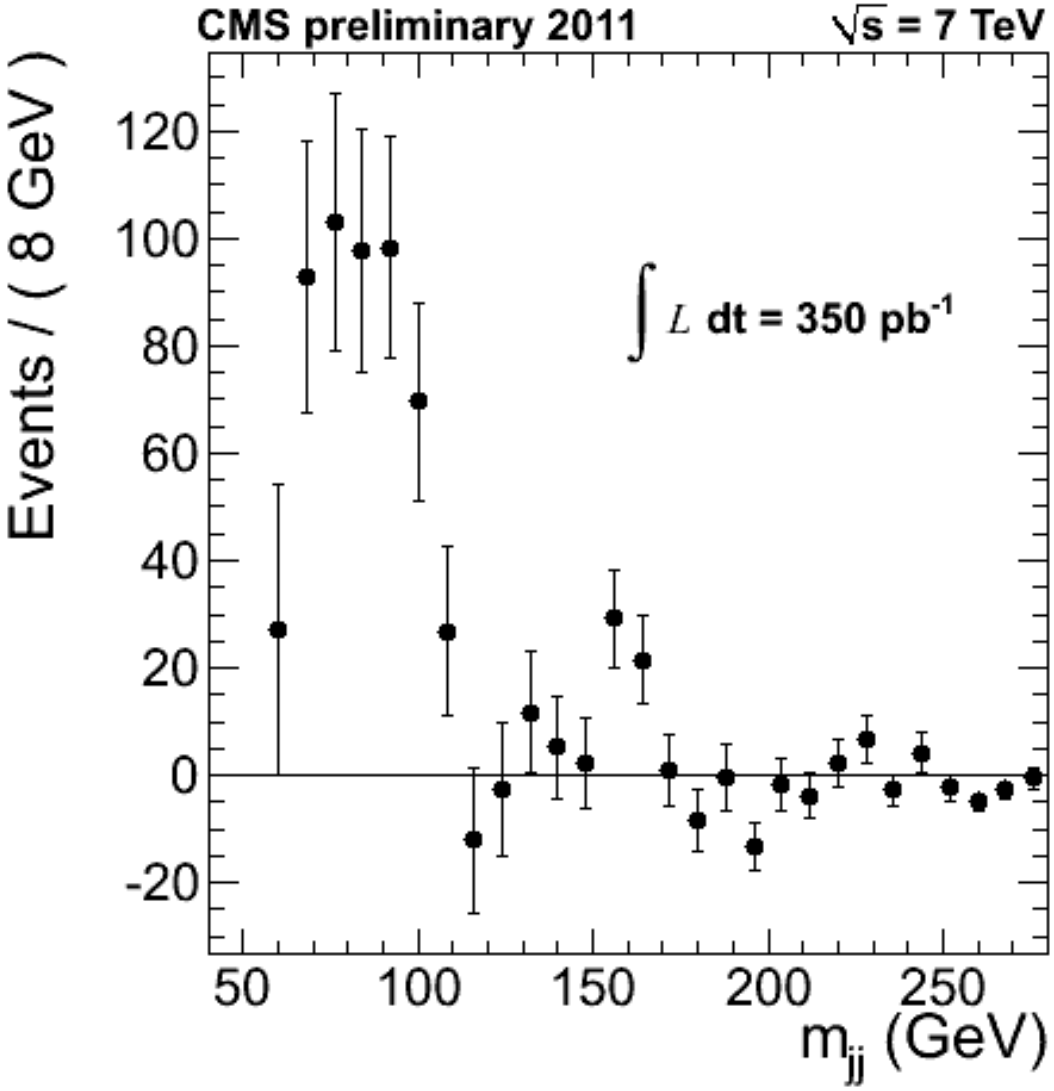
Take the shape from MC. Just fit for the normalization. Blue curve shows the fit to data, dashed red curve is W+jets component, shaded area in orange is di-boson component.

The fit is completely unbinned. Only the plot has binning.

Haven't included single top events in the fit which peak around 150 GeV. Working on it.

Background subtracted distribution on the next slide

m_{jj} in W+2 jet events after bkg subtraction

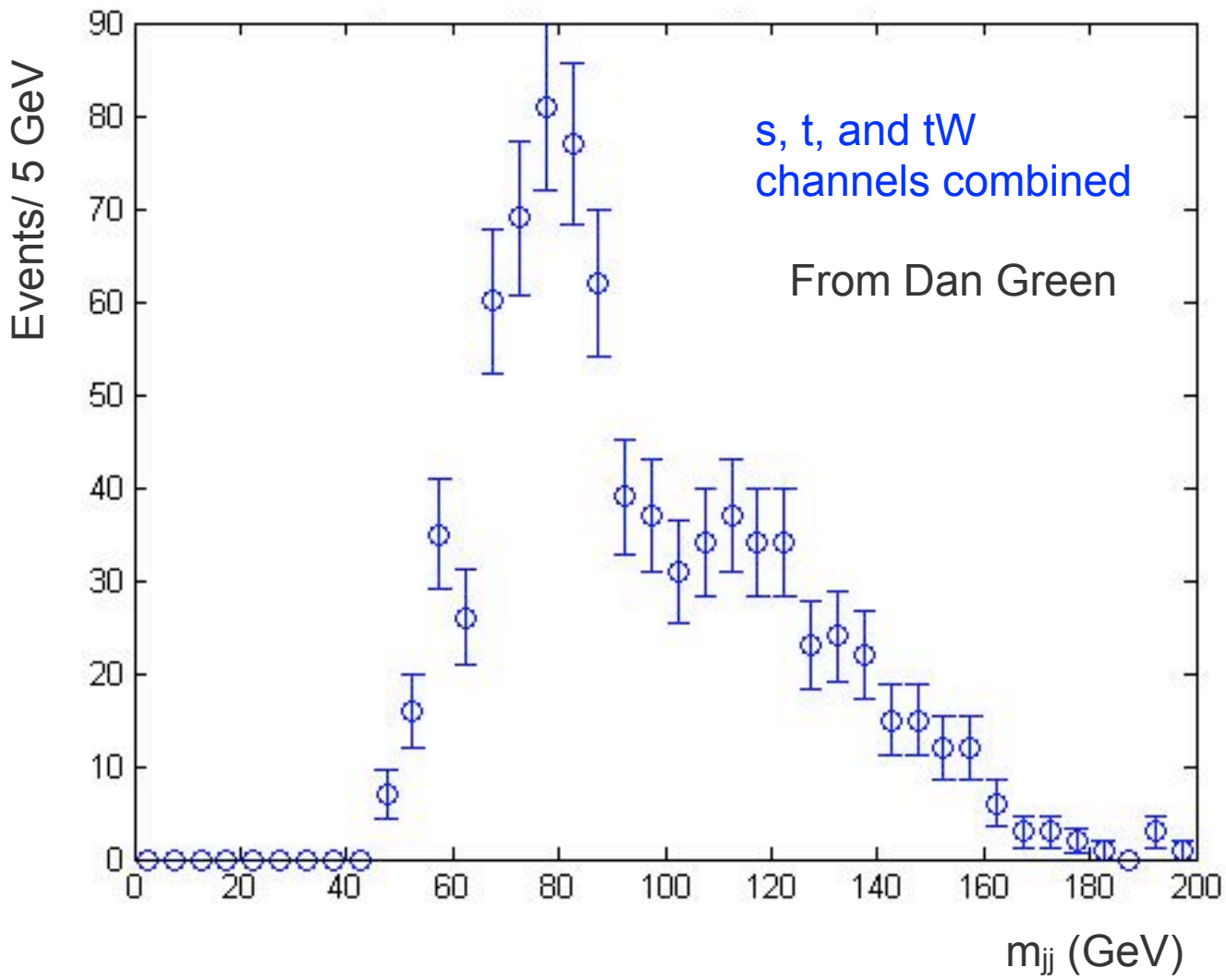


In the W mass window
$65 < m_{jj} < 95$ GeV we get:

338 di-boson events
1702 W+jets events

S/B ~ 1/5

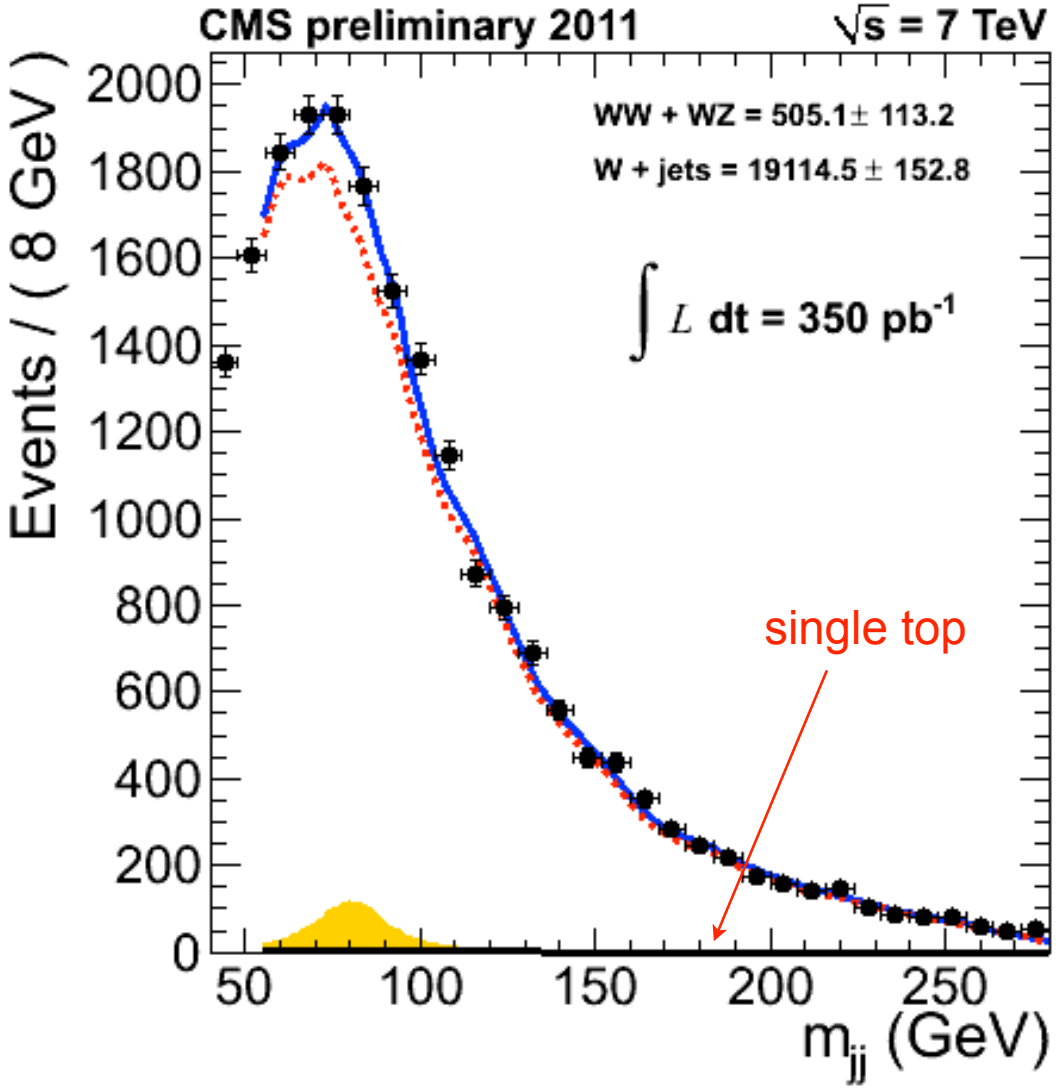
Single top has a broad peak in 100–180 GeV



But small cross section

It affects the template fit only in the tails of m_{jj}

Signal yield is pretty robust w.r.t. selection cuts



Removed all cuts but $\Delta\phi(W, W)$ and $\Delta\eta(j1, j2)$ cuts.

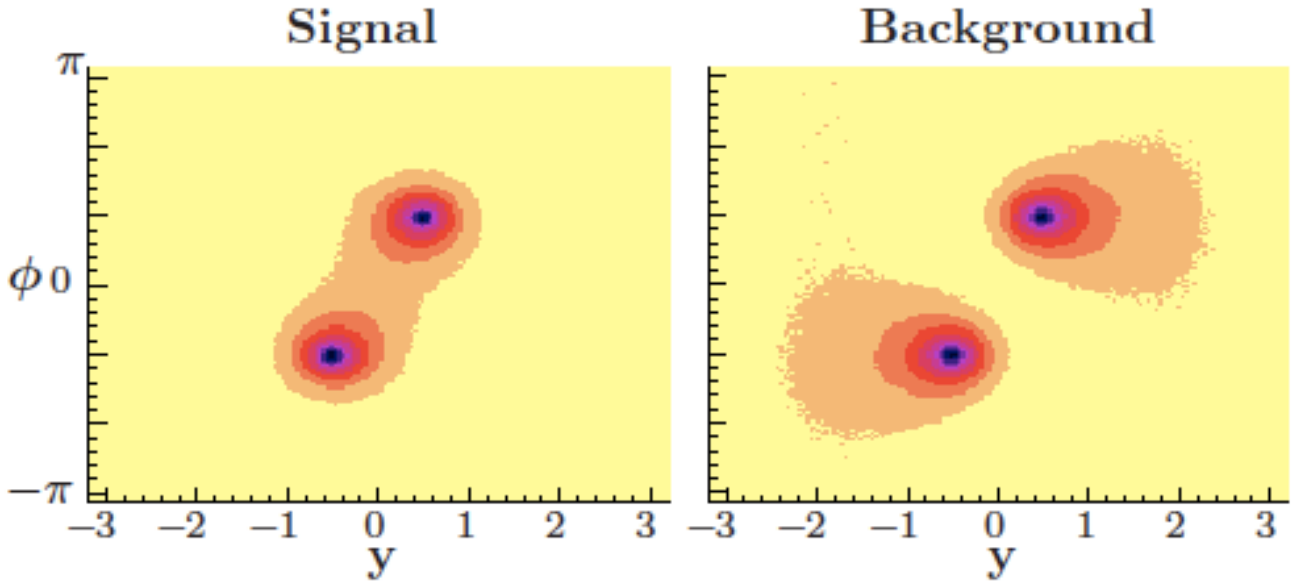
More background sneaks in but signal yield remains same.

Some future improvements

- Use information about color correlation between between the two jets

  Should give us another ~20 % more discrimination.

- Will see if selection optimization helps us any further



arXiv:1001.5027

color pull:

$$\vec{t} = \sum_{i \in \text{jet}} \frac{p_T^i \, |r_i|}{p_T^{\text{jet}}} \vec{r}_i \, .$$

Next steps

1. Converge on the m_{jj} fit: try W+jets shape from data, functional forms motivated by MC but fit on data.

2. Compute efficiency and acceptance

3. Include systematics in the likelihood
   - JES/JER are easy to include
   - Uncertainty in template due to NLO effect, Q^2 variation
   - Propagate uncertainty from efficiency, acceptance

Besides, need to have AN and PAS (even with place holders) written soon. Aim for a preliminary result and documentation by June 20.

Summary

- Di-boson signal in W(lν)+jj channel established
  - basic selection in place, working on further optimization
  - template fit is working, further fine-tuning in technique possible
  - working on a reliable description of W+jets bkgd from data
- Most ingredients in place to be able to compute cross section
  - aim to have a preliminary result soon
- Lagging behind in documentation
  - will try to push hard

BACKUP SLIDES