Finding a Z→2jets signal in W+3jets events at CDF

Caterina Vernieri
(Scuola Normale Superiore & INFN, Pisa)

On Behalf of the CDF Collaboration

XXVI Rencontres de Physique de la Valle dAoste, Feb 26th-Mar 3rd 2012, La Thuile, Italy

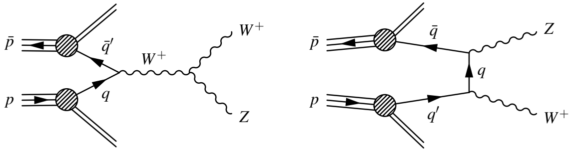
The importance of WZ

Measurements of diboson associated production are **fundamental tests of the EWK sector of SM**

1. Production cross sections are well known by theory at NLO and they are related to TGC^2

   $\mapsto$The virtual W propagator diagram in WZ production is sensitive to WWZ coupling

2 $W^{\pm}$, Z coupling (TGC) is sensitive to new physics



3 $WZ \to \ell\nu q\bar{q}$ is a preliminar step towards $WH \to \ell\nu b\bar{b}$

$\Rightarrow WZ \to l\nu q\bar{q}$ allows to develop a number of techniques to be used in the Higgs search

Why 3 jets?

Observing $WZ \to \ell\nu q\bar{q}$ is extremely difficult for two main reasons.

1^{st} The event rate is extremely low:

- $\sigma(p\bar{p} \to WZ) \cdot B.R.(WZ \to \ell\nu q\bar{q}) \sim 0.7$ pb $\Longrightarrow \sim 4.5 \cdot 10^3$ events for $\mathcal{L} = 6.6$ fb^{-1}
- trigger + kinematical selection efficiency $\sim$few %

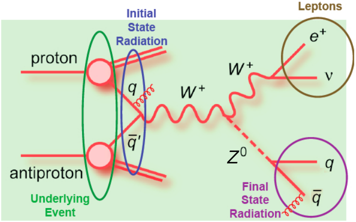
2^{nd} Signal to Background ratio is very poor.

- A preferred discriminant used at CDF to separate the background from the diboson signal is the invariant mass of the two E_T-leading jets.

It would be extremely important be able to search for the signal also in events with more than two high energy jets.

Origin of the Extra Jet

So we look at the MC sample of WZ$\to \ell\nu q\bar{q}$ events with 3 jets where about **33%** of the signal events lie.



1. Radiation from interacting partons (ISR)
2. Radiation from Z-decay products (FSR)
3. Extra-activity produced by spectator partons or by pile-up of events (negligible)

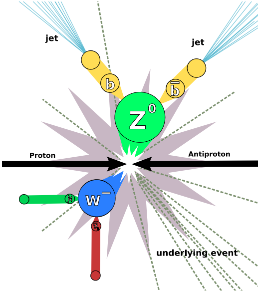
We look for stable hadrons within the jet cone, and ask that the total hadron energy originating from a single parton is >50% than the jet energy.

- Rate of matching is ∼99%
- For each jet we are able to state if is

FSR : all three hadron showers originate from the Z
ISR : one shower comes from a primary beam parton

Event Selection

3 jets $E_T > 25, 15, 15$ GeV



ν

- $\not{E}_T = -\sum_i \vec{E}^i_T >$ 20 GeV

QCD veto

- Cut on M^W_T
- Cut on $\not{E}_T$ along the jets

2 different regions:

- T A G : two jets are b-tagged
- N O T A G : no jet is b-tagged

Isolated lepton triggering the event:

- Central Electron : $E_T > 20$ GeV
- Central muon : $p_T > 20$ GeV

State of the art

- The extra jet due to initial (ISR) or final state radiation (FSR) confuses the choice of the jet system to be attributed to Z decay.
- Which jet combination should be used to build the mass?
  → a wrong choice spoils the resolution

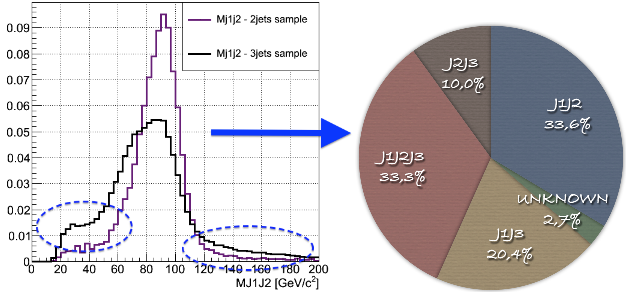


Figure: **MJ1J2** is dijet mass built with the two leading jets. **MJ1J2** mass in the tight dijet sample (no additional jets with E_T >15 GeV and $|\eta|$ <3.6).

The best we can do with three jets

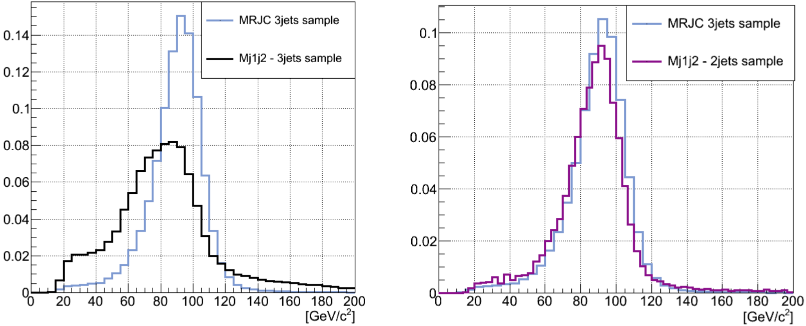


Figure: **MJJ$_{RJC}$** for the three jets sample. **Left**: **MJ1J2** is dijet mass built with the two leading jets. **Right**: **MJ1J2** mass in the tight dijet sample (no additional jets with E_T >15 GeV and $|\eta|$ <3.6).

Strategy

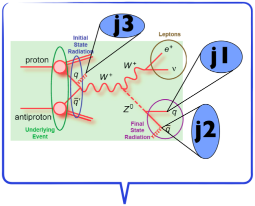
- Jets are ordered in decreasing "b-likeness" (Tag) or E_T (NoTag)
  - Jets are matched in direction to hadrons
  - Investigate at generator level the origin of the not-matched jet (**NMJ**) in order to find the Right Jet Combination (**RJC**)
    1. NMJ = J3 is from ISR $\mapsto$ RJC = J1J2
    2. NMJ = J2 is from ISR $\mapsto$ RJC = J1J3
    3. NMJ = J1 is from ISR $\mapsto$ RJC = J2J3
    4. NMJ = J1 or J2 or J3 is from FSR $\mapsto$ RJC = J1J2J3
- We decide to use 4 different NNs for selecting the 4 different combinations
- We build a "MJJ_{COMB}" Z-mass, combining by a set of subsequent optimal cuts[1] the information provided by the outputs of the four NNs.

[1] Cuts have been optimized against the sensitivity of the measurement.

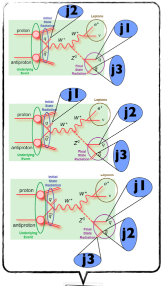
How do we select events with RJC=J1J2?

We train NN_{12} on WZ sample in order to separate:

1. **RIGHT** : events in which J1 and J2 are ‘Z-jets’
2. **WRONG** : events in which J1 or J2 is not a ‘Z-jet’



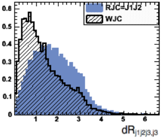
RJC = j1j2

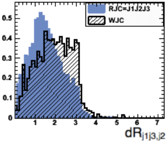


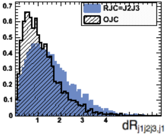
OJC

We want to select in data as many **RIGHT** events as possible.

Our Technique







A list of the variables used as input for the NNs.

1 Kinematical Variables

- $\mathrm{d}\eta_{j_i j_k}$, $\mathrm{dR}_{j_i j_k}$, $dR_{j_i \ell}$
- $dR_{j_k j_l, j_p}$, $dR_{j_1 j_2 j_3, j_k}$

where i, k, p $= 1, 2, 3$; $\ell =$ highest E_T lepton

2 Variables related with the jet systems

- $m_{j_i j_k} / m_{j_1 j_2 j_3}$
- $\gamma_{j_i j_k} = (E_{j_i} + E_{j_k}) / m_{j_i j_k}$
- $\gamma_{jjj} = (E_{j_1} + E_{j_2} + E_{j_3}) / m_{j_1 j_2 j_3}$
- 'pt-imbalance': $P_{T\,J1} + P_{T\,J2} - P_{T\,\ell} -$ MET
- $\eta(j_i + j_k)/\eta(j_p)$, $p_T(j_i + j_k)/p_T(j_p)$

3 CDF tools for discriminating b and gluon jets from light jets

Results - MC Signal

Using MJJ_{COMB} rather than MJ1J2, resolution improves by a factor $\sim$2.

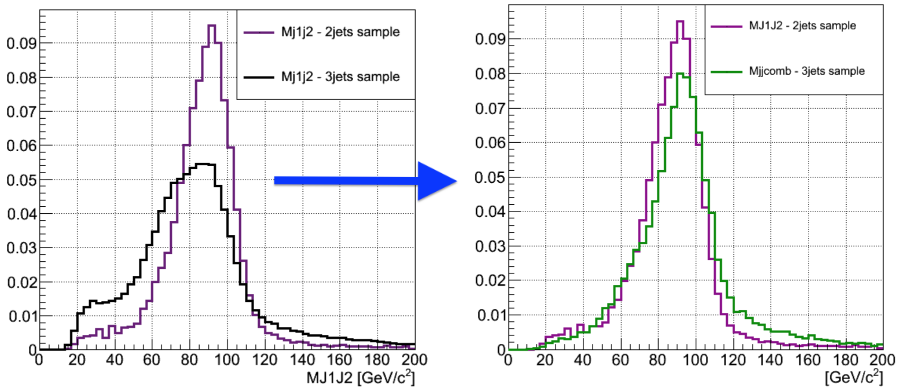
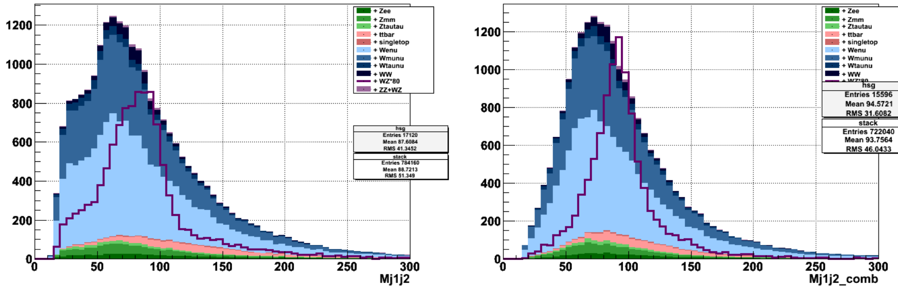


Figure: **MJ1J2** is the invariant mass in the tight dijet sample **SX**: **MJ1J2** is the invariant mass in the three jets sample built using the E_T-leading jets. **DX**: $\mathbf{MJJ_{COMB}}$.

Results - MC

We apply the method also to the major sources of background of a typical diboson analysis at CDF (W+jets, Z+ jets, $t\bar{t}$ and single top) and compare the result to WZ events.



- MJJ_{COMB} has a resolution comparable with the one of MJ1J2 in the 2-jets sample
- MJJ_{COMB} allows a better separation of the WZ/ZZ signal from background

Test on real data

Studied two samples to test the method: MJ1J2 VS MJJ$_{COMB}$

- We analysed a simulated *pretag* sample of WW/WZ/ZZ events.
- We estimate the probability at three standard deviations level to extract an inclusive diboson signal (P3σ).
- P3σ is $\sim$ 4 times greater when fitting MJJ$_{COMB}$ rather than the standard MJ1J2.

Fit Method	**P2σ**	**P3σ**
$WZ/ZZ/WW$ pretag		
- MJ1J2	51.2%	6.4%
- MJJ$_{COMB}$	74.9%	26.5%

- In order to discriminate against the WW contribution we apply our technique considering only WZ/ZZ as the signal.

  $\mapsto$ Information of the *notag* and *tag* channels is exploited.

- The expected p-value is about 22% greater when MJJ_{COMB} is used rather than the standard MJ1J2.

Fit Method	p**-value**
WZ/ZZ notag+tag	
- MJ1J2	$0.44\ \sigma$
- MJJ_{COMB}	$0.54\ \sigma$

Conclusions

- We show a procedure to reconstruct Z in diboson production with large $\not{E}_T$, lepton and 3 jets final state.
- **When WZ/ZZ/WW are considered as signal, the 3σ evidence (P3σ) increases by a factor 4**
- Fitting on MJJ$_{COMB}$ rather than MJ1J2 enhances the sensitivity of extracting the WZ/ZZ signal.
- Adding the 3-jets to 2-jets sample in the WZ/ZZ analyses increases the expected p-value by about 10%

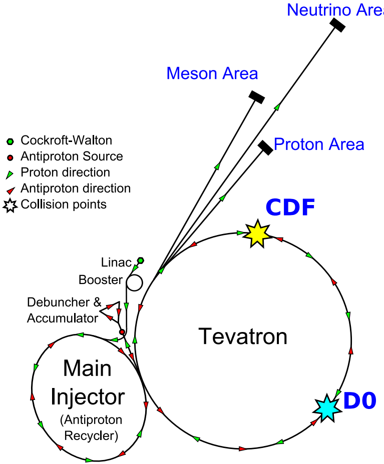
BACKUP

	MJ1J2	MJJ_{COMB}
Acc	100%	90%
p	33%	67%
σ/μ	0.27	0.13

Table: Acc is the acceptance; p is the purity and it is defined as the fraction of events where the corrected jets are selected; σ and μ are estimated by a Gaussian fit in the mass window [70,110] GeV/c^2

.

Tevatron Collider (Run II)

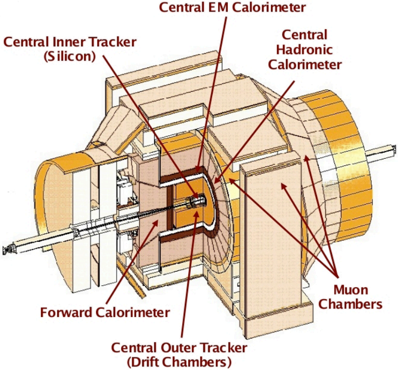


- $p\overline{p}$ collisions at $\sqrt{s} = 1.96$ TeV
- Run II collected data from 2001 until 30/09/2011
- Delivered luminosity ~ 12 fb^{-1}
- Instantaneous luminosity record$\sim 4.3 \cdot 10^{32}$ cm^{-2}s^{-1}
- 2 experiments: CDF, DØ

CDF II

Main Components

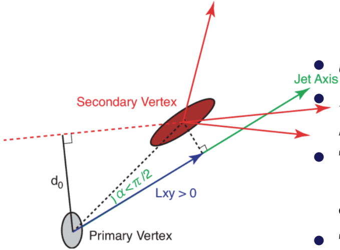
- Tracking system with B∼1.4 T
- Electromagnetic and hadronic calorimeters segmented in projective towers that point to the interaction region
- Muon chambers

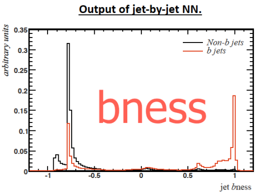


Coordinate System

- r, ϕ, $\eta \equiv -\ln(\tan\theta/2)$
- $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$
- $E_T = E \cdot \sin\theta$

b-jets at CDF: *b*ness





- *b*-jets are produced by *b*-quarks hadronization
- *B*-hadrons have long proper lifetime $\rightarrow$ $\beta\gamma c\tau \sim$ mm
- The vertex where the decay products originate (secondary) may be enough displaced to be distinguishable from the primary one.
- The *b*ness is a multivariate, neural network (NN) based tagger
- It provides an output which indicates how *b*-like a jet appears to be
  - output range is $[-1, 1]$
  - $\rightarrow +1$: *b*-like
  - $\rightarrow$ -1: *q*-jet