A Novel Technique to Reconstruct the Z mass in WZ/ZZ Events with Lepton(s), Missing Transverse Energy and Three Jets at CDFII.

Marco Trovato and Caterina Vernieri[1,2,a]

[1] INFN sez. Pisa, Italy
[2] Scuola Normale Superiore Pisa, Italy

Abstract. Observing WZ/ZZ production at the Tevatron in a final state with a lepton, missing transverse energy and jets is extremely difficult because of the low signal rate and the huge background. In an attempt to increase the acceptance we study the sample where three high-energy jets are reconstructed, where about 1/3 of the diboson signal events are expected to end. Rather than choosing the two E_T-leading jets to detect a Z signal, we make use of the information carried by all jets.
To qualify the potential of our method, we estimate the probability of observing an inclusive diboson signal at the three standard deviations level ($P_{3\sigma}$) to be about four times larger than when using the two leading jets only. Aiming at applying the method to the search for the exclusive WZ/ZZ→ $\ell\nu q\bar{q}$ channel in the three jets sample, we analyzed separately the sample with at least one *b*-tagged jet and the sample with no tags. In WZ/ZZ→ $\ell\nu b\bar{b}$ search, we observe a modest improvement in sensitivity over the option of building the Z-mass from the two leading jets in E_T . Studies for improving the method further are on-going.

1 Motivations

The study of diboson production provides a test of the electroweak sector of the Standard Model (SM). In particular the predicted $W^\pm$, Z couplings (Trilinear Gauge Couplings) are sensitive to new physics.
The study of associated WZ boson production in the final state $\ell\nu b\bar{b}$ is important since the event topology of this process is the same as expected for WH associated production ($M_H \lesssim 135$ GeV) [1].
Observing this process at Tevatron is difficult since the event rate is extremely low. NLO calculations predict WZ production cross section to be ∼3.22 pb [2]. Thus, one expects a handful of events per fb^{-1} of integrated luminosity in the $\ell\nu q\bar{q}$ final state, after allowing for trigger and kinematical selection efficiency[1]. Furthermore, the signal to background ratio is very poor, due primarily to the production of W and associated jets. Since the preferred method used at CDF to disentangle the diboson signal from the backgrounds is a fit to the invariant mass of the two E_T-leading jets, an optimal resolution in jet systems mass is of utmost importance.

2 Three jets

In diboson analyses at CDF the standard kinematical cut requires two high energy jets in the candidate sample (*two jets region*). In order to increase signal acceptance, we investigate the sample where three high-E_T jets are found (*three-jets region*), which in simulations is predicted to contain about 33% of signal events.
Additional jets may be initiated by gluons radiated from the interacting partons (Initial State Radiation, ISR) or from the Z-decay quarks (Final State Radiation, FSR)[2].
The experimental signature involves the presence of a charged lepton (electron or muon), a neutrino (identified through the missing transverse energy, $\not\!\!E_T$) and large-E_T jets.
The sample we investigate is selected by the following cuts:

- exactly three jets[3] with $E_T(J1, J2, J3) > 25, 15, 15$ GeV and $|\eta(J1, J2, J3)| < 2, 2, 3.6$
- an isolated triggered electron or muon with $|\eta| < 1.1$ and $E_T > 20$ GeV
- $\not\!\!E_T > 20$ GeV
- Multi-jet QCD veto:
  - $M_T^W > 10$ (30) GeV if the triggered lepton is a muon (electron), M_T^W being the W-invariant mass in the transverse plane.
  - $\not\!\!E_T$-significance[4] > 1.8 if the triggered lepton is an electron.

In the sample where three jets are found MJ1J2 has a degraded resolution, and high mass and low mass tails due to wrong jet choices are present (see Fig. 1, top). Our goal is to resolve the combinatorics problem present in this region for building the Z mass and consequently improve the resolution of the invariant mass distribution. This work builds on the efforts of Ref. [3].

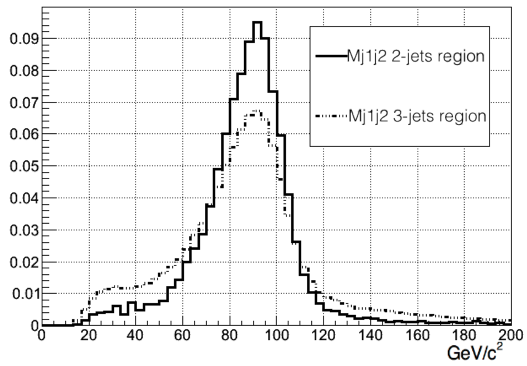
[a] e-mail: caterina.vernieri@pi.infn.it

[1] This statement remains valid even if the few accepted ZZ events with leptonic decay of one Z, where one lepton is not detected, are included.

[2] Extra-activity produced by spectator partons or by pile-up of events is negligible in our studies.

[3] Events with a fourth jet with $E_T > 10$ GeV are excluded.

[4] $\not\!\!E_T$-significance = $(-\log_{10}(\mathrm{P}(\not\!\!E_T^{fluct} > \not\!\!E_T)))$, where P is the probability and $\not\!\!E_T^{fluct}$ is the expected missing transverse energy arisen from fluctuations in the energy measurements.



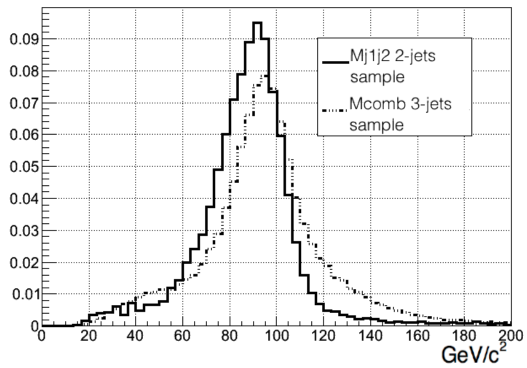


Fig. 1. Top, MJ1J2 in the *three jets region* (dotted) is compared to MJ1J2 in the *two jets region*. Bottom, MJ1J2 in the *two jets region* is compared to MJJ$_{COMB}$ (dotted) in the *three jets region*.

2.1 Composition of the selected events

The following processes would contribute to a data selected sample within our cuts:

- **Electroweak and top (EW)**: *WW*, *WZ*, *ZZ*, *Z*+jets, $t\bar{t}$, single-*top*. Each of these processes can mimick the signal signature, with one detected lepton, large $\not{E}_T$ and jets. The contamination of these processes in the selected data sample is estimated by using their accurately predicted cross sections [2]. The shapes (templates) of a number of observables are obtained from ALPGEN+Pythia [4], [5], Pythia MC [5] after the simulation of the CDF detector.
- **W(→ lν)+jets,** $l = e, \mu, \tau$. Due to the presence of real leptons and neutrinos, the $W + jets$ background is the hardest to be reduced. Templates are obtained from ALPGEN+Pythia MC, while the rate normalization is obtained from data.
- **QCD**: multi-jet production with a jet faking the lepton and fake $\not{E}_T$. Since the mechanism for a jet faking a lepton or for fake missing transverse energy is not expected to be well modeled in MC events, both rate normalization and templates are obtained from data.

In Table 1 we show the estimated number of events for each process contributing for the MJ1J2 distribution.

3 Adopted strategy

We started from studying the three jets sample in WZ MC in which jets are matched in direction to quarks from Z decay. Then, we investigate at generator level the origin of the not-matched jet (**NMJ**) in order to find the Right Jet Combination (**RJC**) which would give the Z mass.
In terms of the RJC frequency the selected sample is composed as follows:

1. NMJ = J3 is from ISR ↦ RJC = J1J2 - 35% of events
2. NMJ = J2 is from ISR ↦ RJC = J1J3 - 21% of events
3. NMJ = J1 is from ISR ↦ RJC = J2J3 - 10% of events
4. NMJ is from FSR ↦ RJC = J1J2J3 - 19% of events

15% of events cannot be allocated to any of these categories. This problem is a subject of further studies.
Four different Neural Networks (NNs) have been trained with the MLP method [7], (NN$_{12}$, NN$_{13}$, NN$_{23}$ and NN$_{123}$) in order to decide event by event which RJC should be used. Inputs to NNs are:

1. Kinematical variables: $d\eta_{j_i j_k}$, $dR_{j_i j_k}$, $dR_{j_i \ell}$, $dR_{j_k j_i, j_p}$, $dR_{j_1 j_2 j_3, j_k}$[5]
2. Variables related to 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' = $P_{T J1} + P_{T J2} - P_{T \ell}$- $\not{E}_T$
   - $\eta(j_i + j_k)/\eta(j_p)$, $p_T(j_i + j_k)/p_T(j_p)$
3. Some tools developed by CDF Collaboration for distinguishing gluon-like and *b*-like jets from light-flavored jets [6] [8].

In Fig. 2 some inputs are shown.

Combining by a set of subsequent optimal cuts[6] the information provided by the outputs of the four NNs, we build a "MJJ$_{COMB}$" Z-mass. Using MJJ$_{COMB}$ rather than MJ1J2, resolution improves by a factor ~2, see Fig. 1 and Table 2. 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. In Fig. 3 and in Table 2 can be noticed that MJJ$_{COMB}$ allows a better separation of the WZ/ZZ signal from background.

Process	Rate (Electrons)	Rate (Muons)
Signal	66.2 ± 0.9	69.5 ± 0.9
WW	386.2 ± 3.0	311.1 ± 3.1
$t\bar{t}$	333.0 ± 1.4	288.5 ± 1.2
single-top	68.9 ± 0.4	57.8 ± 0.3
Z+jets	350.0 ± 3.2	1167.8 ± 4.5
W+jets	10304.2 ± 29.6	8275 ± 22.8
QCD	1600.4 ± 60.0	352.3 ± 5.4
Total Observed	13109.0 ± 114.5	10522.0 ± 102.6

Table 1. Predicted and observed number of events in the notag sample. *W*+jets and QCD rates are estimated by fitting data. The expected rates are separated for different triggered lepton type. By construction the expected numbers are equal to the observed.

4 Tests of the method

To qualify the potential of the method we have studied an experimental data sample accepting events with an isolated large E$_T$ (p$_T$) lepton, large missing E$_T$ and three large

[5] i, k, p = 1; 2; 3. ℓ = highest E$_T$ lepton

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

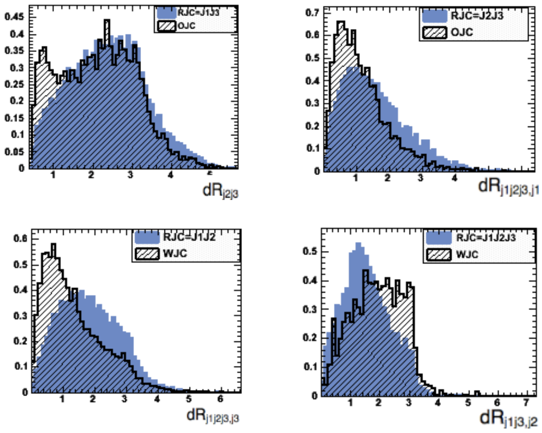


Fig. 2. Some distributions of the variables used as input to NNs, built for the RJC sample and for the complementary one (shaded).

	MJ1J2	MJJ_{COMB}
Acc	100%	72%
p	35%	64%
σ/μ	0.27	0.13

Table 2. *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

.

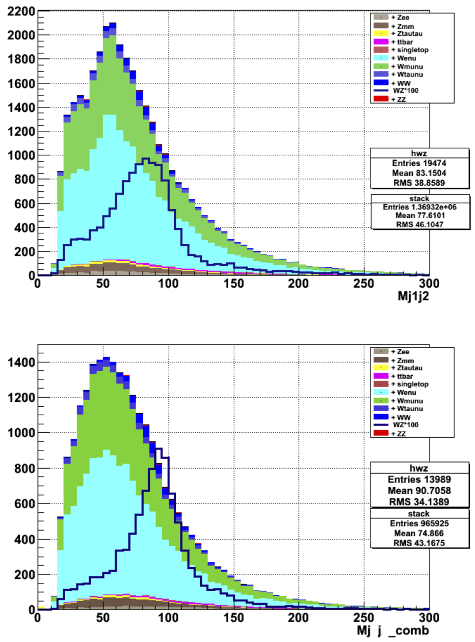


Fig. 3. Simulation of signal+background. Top, MJ1J2. Bottom, MJJ_{COMB}. The horizontal scale is in GeV/c^2. The signal is multiplied by 100.

transverse momentum jets. The selection cuts accept jets of all flavors (*pretag* sample), and all diboson events including WW besides WZ, ZZ may pass the cuts. We estimate the probability at three standard deviations level to extract an inclusive diboson signal. After our procedure for building the Z mass is applied, $P_{3\sigma}$ is about 4 times greater than

Fit Method	$P_{2\sigma}$	$P_{3\sigma}$
Fit signal *WZ/ZZ/WW* (pretag)		
- MJ1J2	51.2%	6.4%
- MJJ_{COMB}	66.7%	25.9%
	p-value	
Fit signal *WZ/ZZ* (notag+tag)		
- MJ1J2	0.35 σ	
- MJJ_{COMB}	0.45 σ	

Table 3. Sensitivity of the fits considering only the three jets region.

when building the Z mass "by default" with the two E_T leading jets.

This attempt represents just a check of our technique. Since diboson signal has been observed in CDF [9], it would be an useful test to understand if using only the three-jets sample a diboson[7] signal could be extracted.

In order to discriminate WZ against the WW contribution we apply our technique considering only WZ/ZZ as the signal. We decide to treat separately the *notag* and *tag* three jets regions and then combine the results in order to reach a greater sensitivity. The sensitivity increases when MJJ_{COMB} rather than the standard MJ1J2 is used: the expected *p*-value is about 30% greater in the former case.

In conclusion, our technique allows including the three jets sample in the WZ/ZZ search in order to increase acceptance and sensitivity in the search for the hadronically decaying Z-boson.

Improvements to this technique and other possible applications are being investigated.

We are grateful to Prof. Giorgio Bellettini, Dr. Giuseppe Latino and Dr. Vadim Rusu for many fruitful discussions and suggestions.

References

1. CDF Collaboration [arXiv:1112.4358v1] (2011)
2. J. M. Campbell and R. K. Ellis, Phys. Rev. D **65**, 113007 (2002).
3. M. Trovato, C. Vernieri, J Phys. Conf. Series **323**, 012014 (2011)
4. M. Mangano, M. Moretti M, F. Piccinini, R. Pittau and A. Polosa, J. High Energy Phys. **07**, 001 (2001)
5. T. Sjöstrand, P. Edén, C. Friberg, L. Lönnblad, G. Miu, S. Mrenna and E. Norrbin Computer Phys. Commun. **135**, 238 (2001).
6. J. Freeman, W. Ketchum, J.D. Lewis, S. Poprocki, A. Pronko, V. Rusu, P. Wittich, Nuclear Instruments and Methods in Physic Research Sec. A **663**, 37 (2012)
7. A. Hoecker et al. *TMVAUsersGuide* (http: tmva. sourceforge.net, 2009)
8. W. Ketchum, CDF Public Note **10643** (2011)
9. T. Aaltonen et al. , Phys. Rev. Lett. **104**, 101801 (2010)

[7] we expect the *ZZ* contribution to be negligible due to the requirement on $\not{E}_T$