FERMILAB-PUB-14-222-E

Exploiting the full information carried by jets for reconstructing the mass of the hadronically decaying Z in WZ/ZZ events with a lepton, missing transverse energy and 3 jets at CDF.

G. Bellettini[b,c], G. Latino[b], V. Rusu[d], M. Trovato[d], G. Velev[d], C. Vernieri[a,b]

[a]*Scuola Normale Superiore, Pisa, Italy*
[b]*INFN Sez. Pisa*
[c]*University of Pisa, Italy*
[d]*Fermilab, Batavia, IL*

Abstract

Observing WZ/ZZ production at the Tevatron in the final state with a lepton, missing transverse energy and two jets is extremely difficult because of the low signal rate and the very large background. In the attempt to increase the acceptance in the analysis of the data collected by the CDF experiment, we study the sample with three high-energy jets, where according to simulations about 1/3 of the diboson events are expected to be. Rather than choosing always the two jets of largest transverse energy (E_T) to reconstruct the Z mass, we make use of the information carried by all jets. We describe in detail how to combine the jet information optimally, and introduce a method of interest in every experiment searching for hadronic resonances in the W/Z + jets channel, including measurements of Higgs boson production associated with a W or Z.

Keywords:

1. Introduction

The study of diboson (WZ/ZZ) production at hadron colliders provides a test of the electroweak sector of the Standard Model (SM) of particle physics. In particular, any deviation from the predicted WWZ and ZZZ couplings (TGC, Trilinear Gauge Couplings) would be indicative of new physics. The WZ identification in the final state with a lepton, missing transverse energy and two b-tagged jets is particularly important since the event topology is the same as expected for WH associated production.

At the 1.96 TeV center-of-mass energy of the Fermilab proton-antiproton Tevatron collider the process WH→ W$b\bar{b}$, for $m_H \simeq 125$ GeV/c^2[1] has an expected cross section times branching ratio ($\sigma \cdot BR$) lower by about eight times than WZ → W$b\bar{b}$. Therefore, observing the WZ→ W$b\bar{b}$ process would be a benchmark for the even more difficult Higgs measurement in the WH→ W$b\bar{b}$ channel. NLO calculations predict a WZ production cross section of 3.22 pb [2]. Given such a small cross section, in the $\ell\nu b\bar{b}$ final state after accounting for trigger and kinematical selection efficiency, only a handful of events per fb^{-1} of integrated luminosity are expected. 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.

Consequently, the observation of this process is difficult. Furthermore, the signal to background ratio is very poor, due primarily to the production of W and associated jets. Since the main feature to be exploited for disentangling signal from background is the invariant mass of Z-decay jets, a correct selection of the jets to be assigned to Z decay and an optimal resolution in di-jet mass are of utmost importance.
A standard kinematical cut in the analysis of the WZ process requires exactly two large transverse energy jets (i.e. $E_T > 20$ GeV) in the candidate sample. Since simulations show that if a third high energy jet is allowed the signal acceptance is increased by about 1/3, it would be important to be able to detect the Z signal also in events with more than two high energy jets.

However, the situation is complicated by (higher-order) sub-processes where an additional jet is produced in association to the W and Z bosons. This work presents a method to overcome this difficulty by making use of the full information on the diboson event in the sample with three jets included. Extra-activity

This manuscript has been authored by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy, Office of Science, Office of High Energy Physics.

Element	Coverage
Tracker	
Silicon Detector (L00, SVX, ISL)	$\lvert\eta\rvert < 2$
COT	$\lvert\eta\rvert < 1$
Calorimeter	
CEM	$\lvert\eta\rvert < 1.1$
CHA	$\lvert\eta\rvert < 0.9$
PEM, PHA	$1.1 < \lvert\eta\rvert < 3.6$
WHA	$0.6 < \lvert\eta\rvert < 0.3$
Muon chambers	
CMU, CMP	$\lvert\eta\rvert < 0.6$
CMX	$0.6 < \lvert\eta\rvert < 1.1$
IMU	$1.1 < \lvert\eta\rvert < 1.5$

Table 1: Coverage of CDF II detector elements (CEM and CHA have an insensitive area at $\eta < 0.1$).

produced by by multiple interactions are negligible in our studies.

2. CDF Run II detector

CDF II is a general purpose apparatus designed to study a wide range of physical processes produced in $p\bar{p}$ collisions at the Tevatron. The detector, which is described in details elsewhere [3] consists of a central *tracking system* embedded in a superconducting solenoid providing a uniform 1.4 T magnetic field parallel to the beam; a *calorimetric system* outside the solenoid at large angles and in the forward regions; a *muon detector*, external to the calorimeters surrounding the detector at all angles down to a few degrees from the beam. Table 1 shows the coverage of the tracker elements, of the calorimeters and of the muon system, and defines the acronyms used in the text.[1]

3. The 3 Jets CDF Data Sample

The experimental signature of the process being investigated involves the presence of a charged lepton (electron or muon), a neutrino (identified as missing transverse energy, $\not{E}_T$[2]) and large-E_T hadron jets. The off-line selection of the data events identifies leptons, jets and $\not{E}_T$ with similar criteria as in Higgs and top quark CDF studies [4]. In particular, electrons are identified as isolated electromagnetic energy clusters that match with a reconstructed track; muons are identified as isolated tracks in the COT which extrapolate to track segments in a muon detector element; jets are reconstructed using the JETCLU cone algorithm [5] with radius[3] 0.4 and properly corrected for detector and physics effects, as described in Ref. [6].

The data sample (”pretag”) that we investigated was selected, in the data as well as in the simulation, with the following cuts:

- exactly three jets[4] with $E_T(J1, J2, J3) > 25, 15, 15$ GeV and $\lvert\eta(J1, J2, J3)\rvert < 2, 2, 3.6$
- An isolated electron (muon) with $\lvert\eta\rvert < 1.1$ and E_T (p_T)> 20 GeV (GeV/c).
- $\not{E}_T > 20$ GeV
- Multi-jet QCD veto:
  - $M_T^W > 10$ (30) GeV if the triggered lepton is a muon (electron)[5].
  - $\not{E}_T$-significance> 1.8 [6] [7].

Two different subsamples corresponding to an integrated luminosity of 6.6 fb^{-1} are studied separately. One, the ”tag” sample, in which we require two jets

[1] CDF uses a cylindrical coordinate system with z along the proton beam direction and the origin at the center of the detector. The polar angle θ is measured with respect to the positive z axis, the pseudorapidity being defined as $\eta = -\log[\tan(\theta/2)]$. The transverse energy and momentum are defined as $E_T = E\cdot\sin\theta$ and $P_T = P\cdot\sin\theta$, respectively. CDF also uses a cartesian coordinate system, with the x axis pointing radially outside on the horizontal plan and the y axis pointing upwards.

[2] $\not{\vec{E}}_T$ is defined as the modulus of the opposite of the vector sum of the transverse energies of all the calorimeter towers:

$$\not{\vec{E}}_T = -\sum_i \vec{E}_T^i \quad (1)$$

where i indexes each tower. The sum involves all the towers with total energy above 0.1 GeV in the region $\lvert\eta\rvert < 3.6$. Tower energy is corrected for known detector and physics effects. When muons are identified $\not{\vec{E}}_T$ is properly modified to account for the muon momentum rather than its associated calorimetric deposit. This vector gives the energy and the direction in the transverse plane of a (massless) undetected particle.

[3] The radius of a cone around a detector position is defined as the distance in the η-ϕ space, i.e. $\sqrt{(\Delta\phi^2 + \Delta\eta^2)}$

[4] Events with a fourth jet with $E_T > 10$ GeV and $\lvert\eta\rvert < 3.6$ are rejected.

[5] M_T^W being the W-invariant mass in the transverse plane:

$$M_T^W = \sqrt{2 \cdot E_T^\ell \cdot \not{E}_T \cdot (1 - \cos\Delta\phi_{\ell\nu})} \quad (2)$$

[6] This parameter is defined as $-\log_{10}(P\not{E}_T^{fluct} > \not{E}_T)$, where P is the probability and $\not{E}_T^{fluct}$ is the expected missing transverse energy due to fluctuations in the energy measurements.

originated by a b quark (b-jets), represented the golden channel for the light SM Higgs boson search at Tevatron ($WH \to Wb\bar{b}$). In this analysis the identification of a b-jet is based on the "bness" [8] b-tagger, which is a multivariate, neural network (NN) based tagger exploiting in particular the high precision track measurement performed by the innermost silicon detectors of the CDF II tracker. It provides an output value which serves as a figure of merit to indicate how b-like a jet appears to be.

The second, the "notag" sample is the "pretag" sample where the tag sample has been removed. The two samples are made independent of each other in order to be able to combine the results.

In Fig. 1 M(J_1J_2), the invariant mass built using the two E_T leading jets, for WZ events in the "two jets region"[7] is compared with the same distribution built in the "three jets region".

In the three jets region M(J_1J_2) has a degraded resolution since the third jet confuses the assignment of two jets to the Z decay: high mass and low mass tails due to wrong combinations are present. Simulations show that by properly handling initial and final state gluon radiation the resolution on the Z-mass would be greatly improved (Fig.2).

3.1. Simulated composition of the data samples

Besides WZ and ZZ, the following background processes contribute to the data samples selected with our cuts:

- Electroweak and top quark production processes: i.e. WW, Z+jets, $t\bar{t}$, single-top. Each of these processes can mimic 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 theoretical cross sections [2]. The distributions ("templates") of a number of observables are obtained from the Pythia (diboson, $t\bar{t}$, single-top) [9] and Alpgen+Pythia (Z+jets) generators [9, 10] after simulation of the CDF detector and reconstruction with the CDF reconstruction software.

- W($\to l\nu$)+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, while the rate normalization is obtained from the data.

[7] Events with a third jet with $E_T > 15$ GeV are rejected in this region.

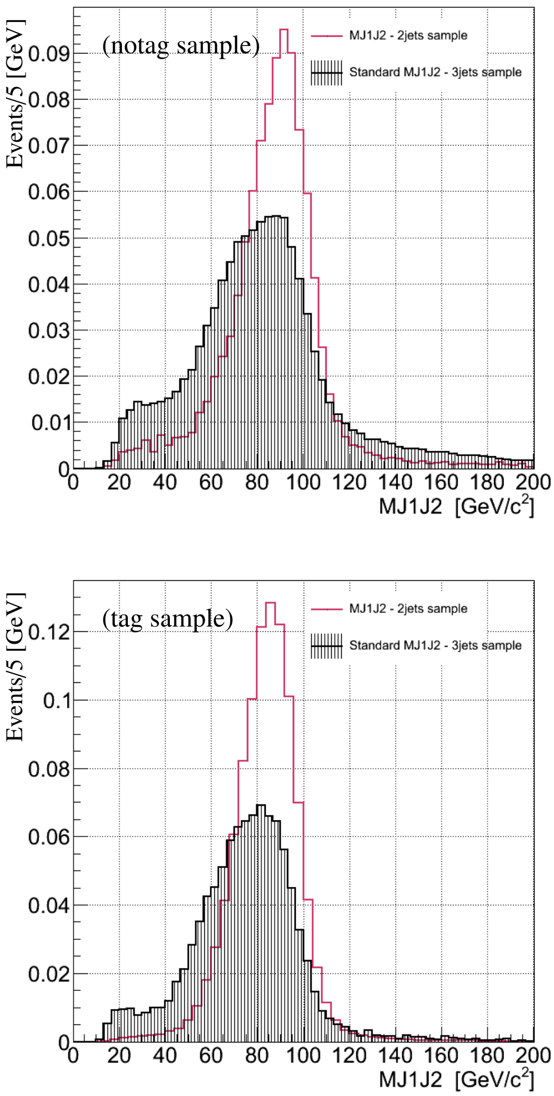

Figure 1: Pythia simulation, top: notag sample, bottom: tag sample. The black dashed-filled distributions are the dijet mass built with the two leading jets in the three jets sample. For comparison, the histograms show the dijet mass in exclusive two-jets events. This comparison shows that using in all cases the mass of the leading jets in three jets events causes a heavy loss of information.

- QCD: multi-jet production with a jet faking a lepton and mismeasurements of the jet energies leading to large missing transverse energy. Both rate normalization and templates are obtained from the data.

In Table 2 we show the estimated number of events for each process contributing to the M(J_1J_2) distribution in the notag and tag samples.

NOTAG	Process	Rate (Electrons)	Rate (Muons)
	Signal (WZ/ZZ)	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

TAG	Process	Rate (Electrons)	Rate (Muons)
	Signal (WZ/ZZ)	3.5 ± 0.2	3.6 ± 0.2
	WW	6.2 ± 0.4	4.7 ± 0.3
	$t\bar{t}$	146.4 ± 0.9	127.9 ± 0.8
	single-top	22.5 ± 0.2	18.7 ±0.2
	Z+jets	8.0 ± 0.4	23.6 ± 0.6
	W+jets	212.0 ± 3.9	189.9 ± 3.2
	QCD	32.5 ± 0.3	5.7 ± 0.0
	Total observed	431.0 ± 20.8	374.0 ± 19.3

Table 2: Predicted number of events in the notag and tag samples. W+jets and QCD rates are estimated by fitting the data, as explained in the text. The expected rates are separated for different triggered lepton type. Since the W+jets contribution is obtained by fitting the simulated template to the data, the total expected rates are equal by construction to the observed ones.

4. Adopted strategy

In order to find a strategy for building properly the Z boson invariant mass in the three jets sample, we analyzed a simulated WZ sample selected as described in Sec.3. This sample was obtained using the Alpgen interfaced to the Pythia to include parton showering and hadronization.

Jets are ordered in decreasing E_T in the notag sample and in decreasing *b*ness in the tag sample[8].

4.1. Matching jets to stable hadrons

Requiring two jets to be matched to the primary $q/\bar{q}$ from Z decay selects a limited event sample (66%) where the NNs could be trained. We found that in the events where this direct matching is unsuccessful the distribution of primaries is confused, with some parton missing or piling-up with others.

In order to reach a greater NN efficiency we adopt a different matching algorithm, which searches for hadrons rather than for primary quarks in the jet cone. The hadrons are traced back to their point of origin in order to identify if they come from a primary beam parton (initial state radiation, ISR) or if they are Z prongs or are radiated by the Z prongs (final state radiation, FSR). Both the Z prongs and the radiation by the Z prongs are named FSR in our classification.

Next, we look for stable hadrons within the jet cone, and for each of the three jets in the event we ask that the total hadron energy originating from a single parton is greater than 50% of the jet energy. When this is possible we label jets as ISR or FSR.

In the rare cases when more than one jet satisfies the condition, the jet with the highest fraction of shared energy is chosen. We checked that when direct matching to primary partons is possible the same jet- to parton matching is achieved as with this indirect method.

In this way the rate of matching reaches ∼ 99% and we can train NNs with a set of events with no kinematic biases due to the matching algorithm. The residual small matching inefficiency of 1% could originate because bending in the CDF magnetic field is not accounted for when tracing the hadrons back to their origin.

By construction we expect at least two jets from FSR (i.e. jets originated by $q/\bar{q}$ from Z). But, in 2.7% of cases our association method fails and we are not able to find them. We investigate these events and we see that one of the two Z-jets is not reconstructed. This interpretation is supported by the observation that in these events also searching for a primary quark in the jet cone fails. The reason could be the calorimeter cracks in $\eta \sim |1|$ region, or the coarser calorimeter granularity for $\eta \sim |2|$ region. We neglect these few events in the NN training.

Once the origin of each jet is well understood we know event-by-event which jet combination should be used to reconstruct the Z mass. We named this "the right jet combination", (RJC). In terms of RJC frequency the notag (tag) sample is composed as follows:

1. J3 is from ISR, J1 and J2 from FSR ↦ RJC = J1J2 33.6% (52.7%)
2. J2 is from ISR, J1 and J3 from FSR ↦ RJC = J1J3 20.4% (9.5%)
3. J1 is from ISR, J2 and J3 from FSR ↦ RJC = J2J3 10% (4.9%)
4. J1, J2, J3 are from FSR ↦ RJC = J1J2J3 33.3% (30.2%)

Notice that in the tag sample the RJC rate of J1J2 is 52.7%. This large rate is because in this sample J1, J2 are defined as the two leading jets in *b*ness. therefore in the tag sample J1 and J2 are already a good bet on which the Z decay jets. Still, even in this sample a better

[8]In the tagged sample, J1, J2 would be the two jets with highest *b*ness, J3 the one with highest E_T among the others.

combination than J1J2 can be searched for in 45% of events.

The best resolution that could in principle be reached in the three jets region is shown in Figs. 2 and 3, where we compare the invariant mass built using the proper RJC for each event with the distribution built with the two E_T leading jets (Fig. 2) and with the dijet mass in the tight dijet sample (Fig. 3). One sees that choosing the RJC in the three jets sample recovers a resolution as good as in the tight dijet sample. The low and high mass tails affecting the M(J_1J_2) distribution are much reduced by choosing the correct combination.

5. Neural Networks

Four different Neural Networks (NNs) have been trained, using the multi layer perceptron method (MLP) [11], in simulated WZ events to describe the signal in four distributions of interest: NN(J_1J_2), NN(J_1J_3), NN(J_2J_3) and NN($J_1J_2J_3$). These NNs combine kinematical information and some tools developed by the CDF Collaboration for discriminating gluon-like and b-like jets from light-flavored jets [8, 12] in MC samples. Inputs to NNs are:

1. Angular variables:
   - d$\eta_{J_iJ_k} = |\eta_{J_i} - \eta_{J_k}|$
   - dR$_{J_iJ_k} = \sqrt{d\eta^2_{J_iJ_k} + d\phi^2_{J_iJ_k}}$
   - dR$_{J_i\ell}$[9]
   - dR$_{J_kJ_l,J_p}$
   - dR$_{J_1J_2J_3,J_k}$
2. Jet kinematics:
   - $m_{J_iJ_k}/m_{J_1J_2J_3}$
   - $\gamma_{J_iJ_k} = (E_{J_i} + E_{J_k})/m_{J_iJ_k}$
   - $\gamma_{JJJ} = (E_{J_1} + E_{J_2} + E_{J_3})/m_{J_1J_2J_3}$
   - 'pt-imbalance' $= P_{T\,J1} + P_{T\,J2} - P_{T\,\ell}$- $\not{E}_T$
   - $\eta(J_i + J_k)/\eta(J_p)$
   - $p_T(J_i + J_k)/p_T(J_p)$
   - Z_{p_t} which is the modulus of the vectorial sum of $P_{T\,J1}$ and $P_{T\,J2}$
   - H$_T^{-1}$ which is the total transverse hadron energy after excluding J1.
3. b/light quark discriminant, quark/gluon discriminant.

With a multiple trial method we scan the NNs outputs and apply cuts to the response of the four NNs determining the most appropriate jet combination for building the Z mass for each event. The method chooses a different combination from J1J2 in about 65% (45%) of cases in the notag (tag) sample. Thanks to the good performance of the algorithm which identifies the origin of

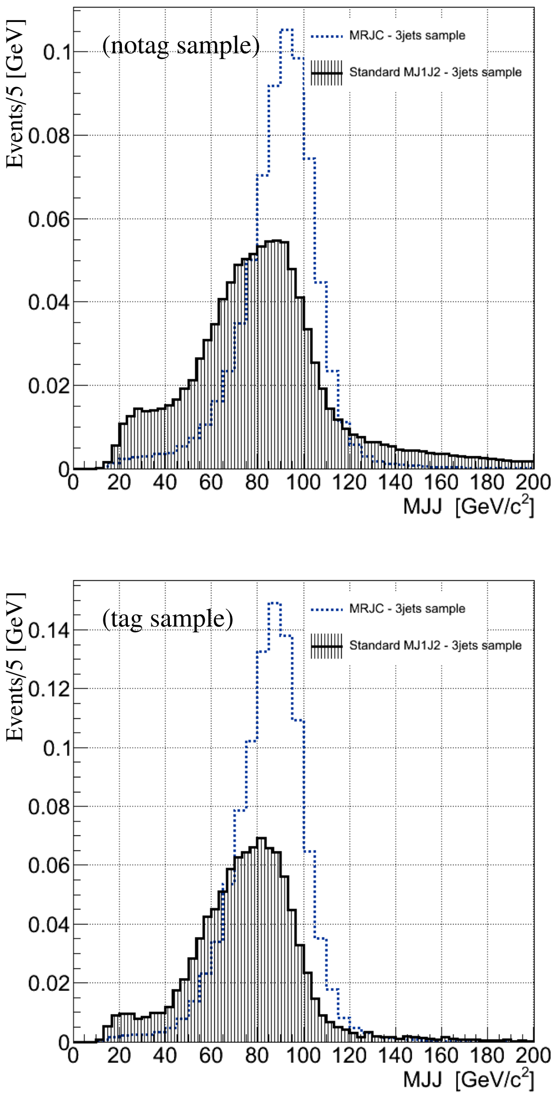

Figure 2: Pythia Simulation, top: notag sample, bottom: tag sample. In the top plot, the histogram is the invariant mass of the two Z-jets when the third jet is from ISR, or of the three jets combined if one jet is FSR. In the bottom plot, the histogram is the invariant mass of the two highest bness jets. The black dashed filled distributions are the dijet mass built with the two E$_T$-highest jets in the samples. The top figure shows that in the notag sample a much improved resolution is obtained if the correct jet combination is chosen. The resolution would approximately be the same as for the tagged dijets in the tag sample (bottom figure). Comparing with Fig. 1, one observes that if jets are correctly assigned the optimal resolution (histograms in Fig. 1) can be approximately achieved also in the three jets sample.

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

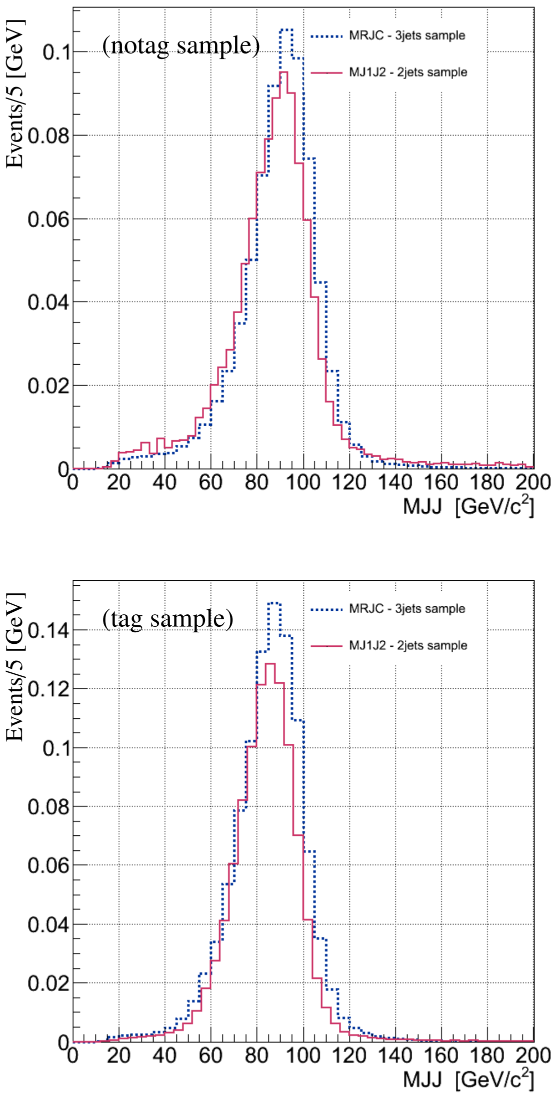

Figure 3: Pythia simulation. The dashed line represents the M_{RJC} distribution; the reference histogram is the invariant mass in the tight dijet sample, in the notag (top) and tag (bottom) sample.

each jet, we are able to identify almost all the cases in which J1J2 is not the appropriate combination.

5.1. *The Novel Technique in the notag sample*

We describe here the method and how the number of cases when J1J2 is the RJC in the notag sample was determined.

5.1.1. *Exploiting the two leading jets: NN(J_1J_2)*

In order to isolate events when RJC = J1J2 we analyze differences of some variables in two subsamples:

- The sample where RJC = J1J2
- The "other jet combination" sample (OJC), where RJC= J1J3, J2J3, J1J2J3

The list of the variables used (see also Fig. A.10) can be found in Appendix A.

In order to avoid a spurious peaking of the invariant mass distribution for the background (W+jets, $t\bar{t}$, etc...) within the signal windows, the input variables are weighted. Weights are such that the M(J_1J_2) distribution in the OJC sample becomes approximately the same as in the RJC = J1J2 sample. By reweighting one ensures that any difference in the NN output distribution in the OJC sample is a reflection only of differences in the kinematical parameters other than the Z-mass to be reconstructed.

The weighted variables are used for training a Neural Network with the MLP method. The NN(J_1J_2) responses in the two samples are shown in Fig. 4

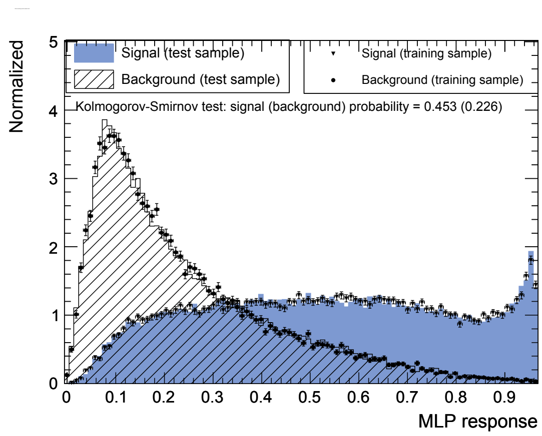

Figure 4: NN(J_1J_2) MLP response for the RJC sample (solid) and for the OJC sample (dashed).

5.1.2. *Criterion for the notag sample*

For combining the information provided by the outputs of the four NNs, a criterion for building the invariant mass has been developed. We started with a requirement on NN(J_1J_2) and selected 33.5% of the sample where M(J_1J_2) is chosen for reconstructing the *Z*. Next we applied a requirement on NN($J_1J_2J_3$) in order to select the subsample where M($J_1J_2J_3$) would be used. Lastly, we applied a cut on the NN(J_1J_3) and NN(J_2J_3)

outputs. The values for the cuts for each NNs have been chosen in order to obtain for each combination a number of events equal to the expected frequency of the RJC in the selected sample (see sec. 4.1)[10]. As a final step, the cuts were varied slightly around these values in order to optimize them against the sensitivity of the WZ/ZZ cross section measurement.

The selected cuts, applied sequentially, (see Table 3) allow for calculating the appropriate MJJ_{COMB} for each 3-jets event. The resulting distribution is shown in Fig. 5 compared with the $\mathrm{M(J_1J_2)}$ distribution in the three jets region. With MJJ_{COMB} a clear improvement in resolution is obtained, with a distribution in the exclusive two jets region.

In order to understand quantitatively its impact on

NN	MJJ_{COMB}
$\mathrm{NN(J_1J_2)} > 0.5$	$\mathrm{M(J_1J_2)}$
$\mathrm{NN(J_1J_2J_3)}>0.3$	$\mathrm{M(J_1J_2J_3)}$
$\mathrm{NN(J_1J_3)}>0.55$	$\mathrm{M(J_1J_3)}$
$\mathrm{NN(J_2J_3)}>0.55$	$\mathrm{M(J_2J_3)}$

Table 3: Cuts used for calculating MJJ_{COMB} in the notag sample.

the sensitivity of the measurement we also apply the method to the major sources of background (W+jets, Z+jets, $t\bar{t}$ and single top) and compare the result to WZ events. In Fig. 6, $\mathrm{M(J_1J_2)}$ and MJJ_{COMB} are shown in signal and background events. The overlaid distribution represents the signal multiplied by 80 to allow for visual comparison.

One observes that the slight background change does not compromise the improvement in the reconstructed Z-mass distribution obtained with MJJ_{COMB}. We also note that the resolution is still not good enough to distinguish from each other the contributions of the WW$\rightarrow \ell\nu$jj and WZ-$\rightarrow \ell\nu$jj processes, which are shown in different color codes in the figures.

The acceptance A, the purity p, the standard deviation over mean ratio (σ/μ), and the FWHM.[11] of the Z-peak for the standard $\mathrm{M(J_1J_2)}$ and for the optimized MJJ_{COMB} are compared in Table 4

$$A = \frac{Evt^{sel}}{Evt^{tot}} \quad p = \frac{MJJ^{RIGHT}}{Evt^{sel}} \tag{3}$$

where Evt^{sel} is the number of selected signal events out of the total number of events in the sample Evt^{tot}, and

[10]with a perfect NN optimization we would expect that exactly all RJC would be selected.

[11]FWHM and μ are estimated by a Gaussian fit in the mass window [70,110] GeV/c^2.

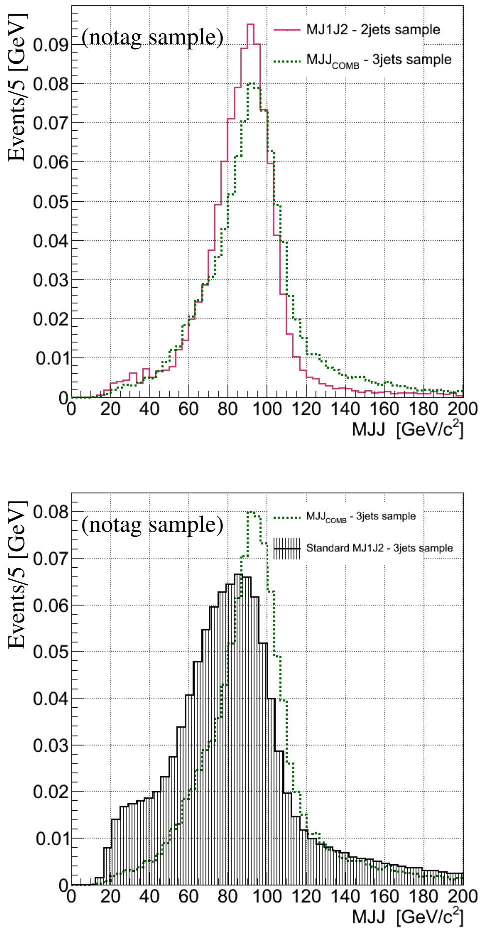

Figure 5: Pythia simulation, notag sample. $\mathrm{M(J_1J_2)}$ in the two jets region (plain, upper plot) and $\mathrm{M(J_1J_2)}$ in the three jets region (dashed, lower plot) are compared with MJJ_{COMB} in the three jets region (dashed line, both plots). The top plot shows that by using MJJ_{COMB} in the three jets sample one approaches the resolution obtained in the exclusive dijet sample. The lower plot shows the large progress made by using MJJ_{COMB} rather than $\mathrm{M(J_1J_2)}$ in the 3-jets sample.

MJJ^{RIGHT} is the number of the signal events in which the correct RJC is found by the applied NN cuts.

The excellent performance of the decorrelation proce-

	std	if **criteria**
A	100%	90%
p	35%	65%
σ/μ	0.25	0.13
FWHM [GeV/c^2]	48.3	29.9

Table 4: Parameters (see text) assessing the performance of MJJ_{COMB} in the notag sample

dure is illustrated in Fig. 5.1.2, where MJJ_{COMB} is com-

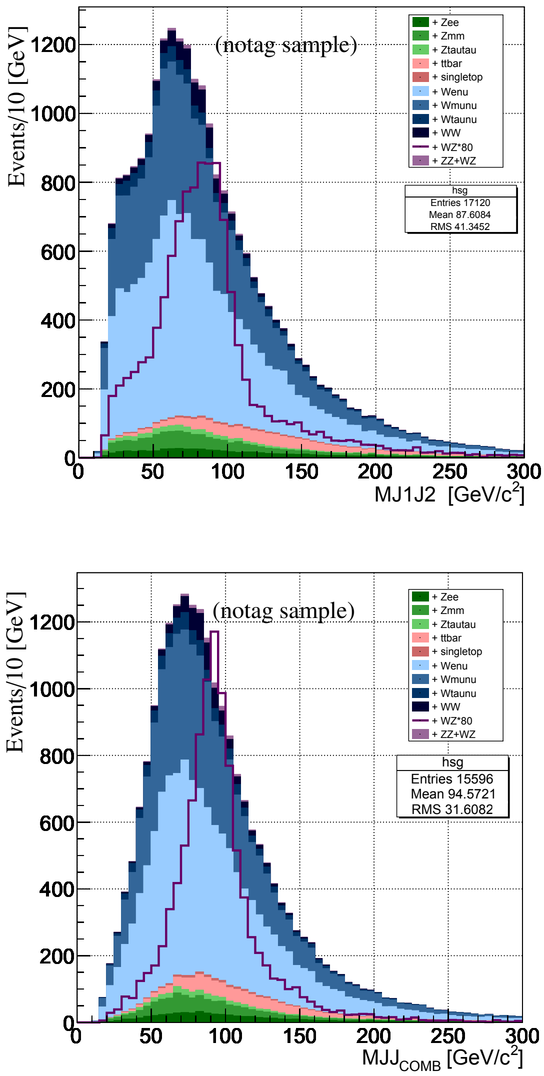

Figure 6: Notag Sample. Simulation of signal+background. Upper plot, M(J_1J_2). Lower plot, MJJ$_{COMB}$. In both figures the background rate is normalized to the CDF data in a sample of integrated luminosity 6.6 fb^{-1}, while the overlaid signal is multiplied by 80.

pared to M(J_1J_2) for a sample of the dominant W+jets background. The MJJ$_{COMB}$ distribution is narrower than M(J_1J_2), but its mean is only slightly changed.

5.2. *The Novel Technique in the tag sample*

In the tag sample we used a very similar technique. Differences from the criterion developed in the notag sample are mainly in the variables used for training each NN.

Since we expect two *b*-jets in this sample we also use *b*ness information in our NNs. The parameters used to train the NNs in the tag sample are listed in Appendix B. Here we only mention the adopted criterion and present the results.

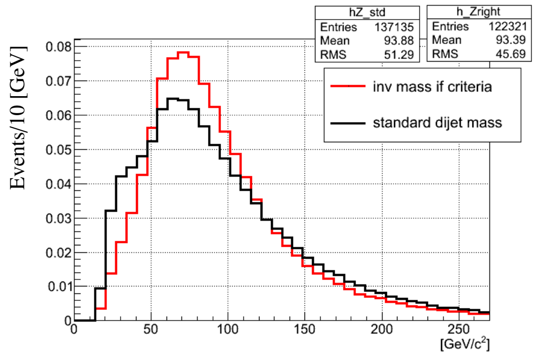

Figure 7: Notag Sample. MJJ$_{COMB}$ is compared to M(J_1J_2) for W+jets events.

NN	MJJ$_{COMB}$
NN(J_1J_2)> 0.4	M(J_1J_2)
NN($J_1J_2J_3$)>0.3	M($J_1J_2J_3$)
NN(J_1J_3)> 0.6	M(J_1J_3)
NN(J_2J_3)> 0.6	M(J_2J_3)

Table 5: Cuts used for calculating MJJ$_{COMB}$ in the tag sample.

5.2.1. *Application of the criterion to the tag sample*

The developed criterion allows for choosing the NN's cuts (see Table 5) calculating a MJJ$_{COMB}$ for the three jets region, as appropriate for the tag sample. The derived mass distribution is compared with M(J_1J_2) in the two jets region in Fig. 8.

Although in this sample the resolution of M(J_1J_2) is already quite better than in the notag sample, a significant improvement in resolution is obtained as is manifested for example by the significant reduction of the σ/μ ratio (Table 6). In order to understand the impact of this method on a real measurement, also in this sample we compute MJJ$_{COMB}$ for the main sources of background (W+jets, Z+jets, $t\bar{t}$ and single top) and compare it to WZ events. In Fig. 9, M(J_1J_2) and MJJ$_{COMB}$ distributions are shown for signal and combined background events. The overlaid signal distribution is multiplied by 40 to allow for visual comparison.

We note that, even in this sample, the invariant mass distribution for the background events does not vary significantly, as a consequence of the adopted decorrelation procedure.

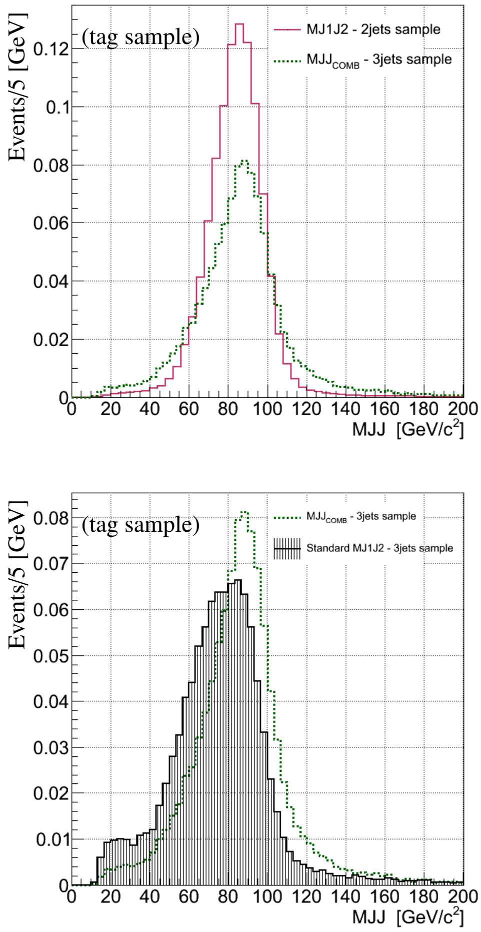

Figure 8: Tag sample. M(J_1J_2) in the two jets region (plain, upper plot) and in the three jets region (dashed, lower plot) are compared with MJJ$_{COMB}$ in the three jets region (dashed line).

	std	if **criteria**
A	100%	92%
p	53%	72%
σ/μ	0.22	0.14
FWHM [GeV/c^2]	40.9	30.0

Table 6: Performance of MJJ$_{COMB}$ in the tag sample: acceptance, purity and resolution parameters.

6. Sensitivity and Optimization

We estimate the increased probability of observing a signal with a 2 σ or a 3 σ significance by applying our method to a set of simulated experiments (pseudoexperiments, PE). We generate about 100,000 PE's.

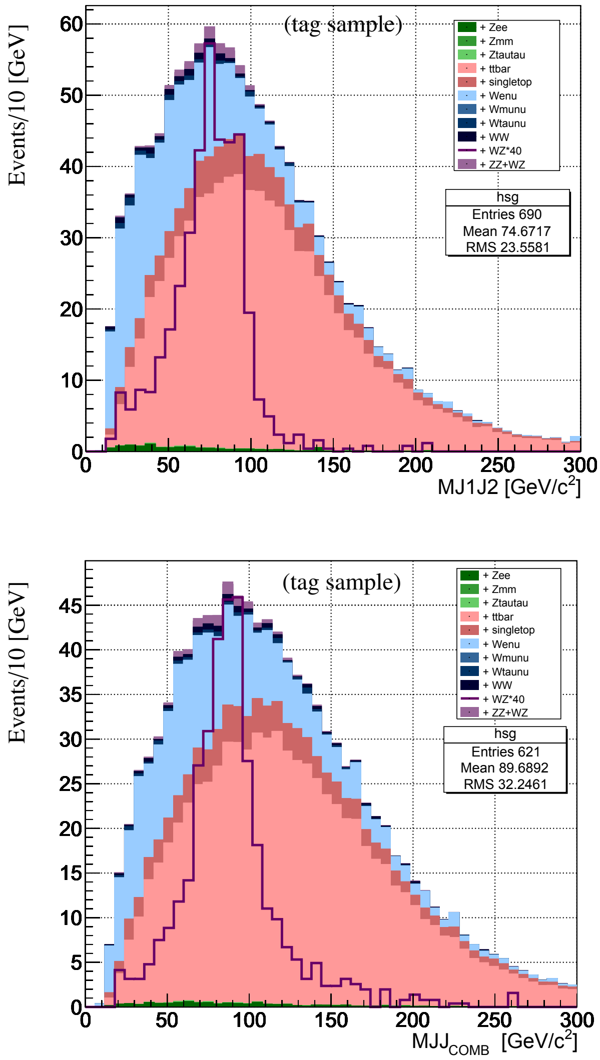

Figure 9: Simulation of signal+overall background in the tag sample. Upper plot, M(J_1J_2). Lower plot, MJJ$_{COMB}$.

6.1. WZ/ZZ/WW as a signal

It is of interest to estimate the probability at two and three standard deviations level to extract an inclusive (WW+WZ+ZZ) diboson signal in the exclusive 3-jets pretag sample ($P_{2\sigma}$, $P_{3\sigma}$). For this estimate, systematic uncertainties were not included for generating PE's and for the fits to the pseudo-data. By applying our technique the resolution improves by 22%, see Table 7. After our procedure for calculating the Z mass is applied, $P_{3\sigma}$ is about 4 times greater than when building the Z mass "by default" with the two E_T leading jets (see Table 8).

	std	if **criteria**
σ/μ	0.21	0.16
FWHM [GeV/c^2]	39.4	32.7

Table 7: Resolution parameters of MJJ_{COMB} and $M(J_1J_2)$ in the pretag sample for the diboson signal.

6.2. WZ/ZZ signal in tag and no-tag samples

We estimate the expected p-value to extract the WZ/ZZ signal in the 3-jets sample by the combined information of the notag and tag channels. An increased sensitivity is expected by combining the results in the tag and notag samples if they are analized separately. The results are given in Table 8.

After applying our technique to calculate the Z mass in the three jets region, the $P_{2\sigma}$ sensitivity increases from 0.35 to 0.45. We also estimate the expected p-value to extract a WZ/ZZ signal by combining the information of the notag and tag channels and of the 2-jets and 3-jets samples. The expected p-value estimated considering only the two jets region is 0.75 σ [13]. We estimate a p-value = 1.05 σ including the three jets region. This is the main result of this study. By combining the two samples one one expects to achieve a reduced statistical error.

Fit Method	$P_{2\sigma}$	$P_{3\sigma}$
Fit signal $WZ/ZZ/WW$ (pretag)		
- $M(J_1J_2)$	51.2%	6.4%
- MJJ_{COMB}	66.7%	25.9%
	p-value	
Fit signal WZ/ZZ:		
(combined tag and notag analyses)		
- $M(J_1J_2)$	0.35 σ	
- MJJ_{COMB}	0.45 σ	

Table 8: Sensitivity of the fits in the exclusive 3 jets sample.

7. A test on data

To check the potential of the method in the real world we have applied it to a CDF data sample, accepting events with a leptonically decaying W and three large transverse momentum jets, as in the previously-described simulation studies. The selection cuts accepted jets of all flavors (pretag sample), and a fraction of all diboson events including WW besides WZ, ZZ can pass the cuts. After our procedure for calculating the Z mass is applied, $P_{3\sigma}$ was found to be about 4 times larger than when calculating the Z mass from the two leading-E_T jets only.

This finding confirms the expectation that the impact of the statistical errors on the search for a signal is significantly reduced.

The sensitivity of this measurement is primarily limited by the large statistical errors. With respect to the Z mass calculated with the two E_T leading jets, no additional systematic uncertainties need to be introduced when applying the new technique. However, we have estimated the impact of the systematic error induced by the uncertainty on the jet energy scale, which is the largest one in such jet studies. We have evaluated that this systematic uncertainty affects the $P_{3\sigma}$ of the two methods in the same way. We then conclude our result is robust against systematic uncertainties.

8. Conclusion

The sensitivity of this measurement is primarily limited by the large statistical errors. We have shown that by including the three jets sample in the WZ/ZZ analyses one would increase significantly the acceptance and thereby the sensitivity in a search for the hadronically decaying Z-boson in associated WZ production.

The aim of this work was to determine the gain in sensitivity achievable at CDF by exploiting the $\ell\nu$+3 jets in the search for associated WZ production, with $Z\to jj$. The significant gain obtained with our method in this particular search, over an analysis where only events with 2 exclusive jets are accepted, suggests that a similar method could profitably be implemented in any search for rare dijet resonances produced in hadron collisions. This would naturally be applicable to the studies of the Higgs boson produced in association with a W or Z in the process W(Z)H→ $\ell\nu(\ell\ell, \nu\nu)$jj, but might in general allow reaching a significant progress in searches of structures in more inclusive jet distributions. In such searches the potential of the quark-versus-gluon jet discriminant, that could be further refined, could play a major role.

Appendix A. Neural Networks input in notag sample

In Fig.A.10 the distributions of the kinematic parameters adopted as input variables to NN(J_1J_2) are shown. The input variables to the other three neural networks trained in the notag sample to isolate events when RJC = J1J3, J2J3, J1J2J3, are listed in table A.9.

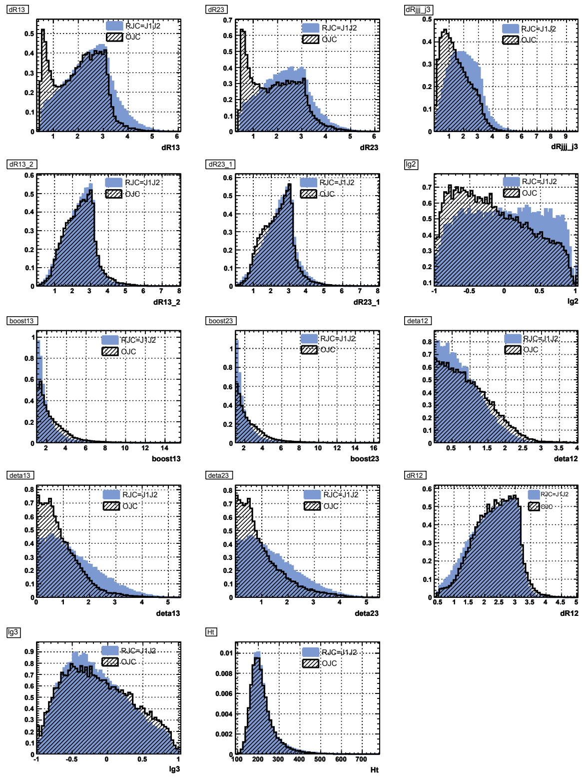

Figure A.10: Distributions of the NN(J_1J_2) input variables after weighting. For each variable a comparison is made between the distribution in the signal subsample where J1J2 is the RJC (plain distribution), and the one where the RJC is different from J1J2, (dashed).

Appendix B. Neural Networks input in tag sample

The variables used in the pretag and in the tag samples for training the four NN‘s described in the text are listed in Table B.10.

NN(J_1J_2)
$m_{JJ'}/m_{J_1J_2J_3}$[12]
$\gamma_{JJ'} = (E_J + E_{J'})/m_{JJ'}$
$d\eta_{JJ'}$
$dR_{JJ'}$
$dR_{J_1J_2,J_3}$
$dR_{J_1J_2J_3,J_3}$
Z_{p_t}
”pt-imbalance”
H_T
Quark Gluon Discriminant for J2, J3
NN(J_1J_3)
$m_{JJ'}/m_{J_1J_2J_3}$
$\gamma_{JJ'} = (E_J + E_{J'})/m_{JJ'}$
$d\eta_{J_1J_2}$
$d\eta_{J_2J_3}$
$dR_{J_2J_3}$
$\eta(J_1 + J_3)/\eta(J_2)$
$dR_{J_1J_2,J_3}$
$dR_{J_1J_2J_3,J_2}$
$dR_{J_2,\ell}$
Quark Gluon Discriminant for J3, J1
H_T
$Z_{p_T,13}$ related to the jet system, 13.
NN(J_2J_3)
$m_{JJ'}/m_{J_1J_2J_3}$
$\gamma_{JJ'} = (E_J + E_{J'})/m_{JJ'}$
$d\eta_{J_1J_2}$
$p_T(J_2 + J_3)/p_T(J_1)$
$dR_{J_1J_3}$
$dR_{J_1J_2,J_3}$
$dR_{J_1J_2J_3,J_1}$
$dR_{J_2,\ell}$
H_T
$Z_{p_T,23}$ related to the jet system, 23.
Quark Gluon Discriminant for J2, J3
NN($J_1J_2J_3$)
$\gamma_{JJ'} = (E_J + E_{J'})/m_{JJ'}$
$\gamma = (E_{J_1} + E_{J_2} + E_{J_3})/\mathrm{M}(J_1J_2J_3)$
“pt-imbalance” : $p_{T\,J_1} + p_{T\,J_2} + p_{T\,J_3} - p_{T\,\ell}$−MET
$d\eta_{J_1J_3}$
$dR_{J_2J_3}$
$dR_{J_1J_3,J_2}$
$dR_{J_2J_3,J_1}$
$dR_{J_3\ell}$
Z_{p_T}
Quark Gluon Discriminant for J2, J3

Table A.9: Input variables used to train the Neural Networks of the notag sample

NN(J_1J_2)
$m_{JJ'}/m_{J_1J_2J_3}$[13]
$\gamma_{JJ'} = (E_J + E_{J'})/m_{JJ'}$
$d\eta_{J_1J_3}$
$d\eta_{J_2J_3}$
$\eta(J_1 + J_2)/\eta(J_3)$
$dR_{J_1J_2,J_3}$
$dR_{J_2J_3,J_1}$
$dR_{J_1J_2J_3,J_3}$
*b*ness for J1, J2
Quark Gluon Discriminant for J3
NN(J_1J_3)
$m_{JJ'}/m_{J_1J_2J_3}$
$\gamma_{JJ'} = (E_J + E_{J'})/m_{JJ'}$
$d\eta_{J_1J_2}$
$d\eta_{J_2J_3}$
“pt-imbalance” : $p_{T\,J_1} + p_{T\,J_3} - p_{T\,\ell}$−MET
$dR_{J_1J_2,J_3}$
$dR_{J_1J_2J_3,J_2}$
$dR_{J_3,\ell}$
EMfr for J2[14]
*b*ness for J2
Quark Gluon Discriminant for J3, J1
NN(J_2J_3)
$m_{JJ'}/m_{J_1J_2J_3}$
$\gamma_{JJ'} = (E_J + E_{J'})/m_{JJ'}$
$d\eta_{J_1J_2}$
$d\eta_{J_1J_3}$
“pt-imbalance” : $p_{T\,J_2} + p_{T\,J_3} - p_{T\,\ell}$−MET
$p_T(J_2 + J_3)/p_T(J_1)$
$dR_{J_1J_2,J_3}$
$dR_{J_1J_3,J_2}$
$dR_{J_1J_2J_3,J_1}$
*b*ness for J1, J3
Quark Gluon Discriminant for J1, J3
NN($J_1J_2J_3$)
$\gamma_{JJ'} = (E_J + E_{J'})/m_{JJ'}$
$\gamma = (E_{J_1} + E_{J_2} + E_{J_3})/\mathrm{M}(J_1J_2J_3)$
“pt-imbalance” : $p_{T\,J_1} + p_{T\,J_2} + p_{T\,J_3} - p_{T\,\ell}$− MET
$d\eta_{J_1J_3}$
$dR_{J_1J_3,J_2}$
$dR_{J_2J_3,J_1}$
$dR_{J_1J_2J_3,J_3}$
$dR_{J_3\ell}$
EMfr for J2, J3
*b*ness J2, J3
Quark Gluon Discriminant for J2, J3

Table B.10: Input variables used to train the Neural Networks of the tag sample.

References

[1] A. Denner, S. Heinemeyer, I. Puljak, D. Rebuzzi, and M. Spira, Eur. Phys. J. C **71** 1753 (2011).

[2] J. M. Campbell and R. K. Ellis Update on Vector Boson Pair Production at Hadron Colliders Phys. Rev. D **65** (2002) 113007.

[3] A. Abulencia et al. (CDF Collaboration), Journal of Physics G **34**, 2457 (2007).

[4] Aaltonen et al. (CDF Collaboration), Phys.Rev.Lett. **109** (2012) 111804.

[5] G. C. Blazey et al., Run II Jet Physics, arXiv:hep-ex/0005012v2 (2000).

[6] A. Bhatti et al., Nucl. Instrum. Meth.**A566**, 375 (2006).

[7] R. Culbertson et al *Search for Anomalous Production of di-photon+MET Events in 2 fb^{-1} of Data* Phys.Rev.D **82** (2010) 052005.

[8] J. Freeman et. al., Nucl. Instrum. Meth.**A663**, 37 (2012).

[9] T. Sjöstrand et al Computer Phys. Commun. **135** (2001) 238

[10] M. L. Mangano et al J. High Energy Phys. **07** (2001) 001

[11] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss TMVA - Toolkit for Multivariate Data Analysis (2007) arXiv:physics/0703039

[12] W. Ketchum, V. Rusu, Y.K. Kim Search for WZ/ZZ Production in leptons+jets channel CDF Public Note **10601** (2011)

[13] G. Bellettini, G. Latino,V. Rusu, M. Trovato, G. Velev, C. Vernieri Search for WZ/ZZ production in events with lepton(s) plus jets plus missing transverse energy CDF Public Note **10838** (2011)