AB

Institut
de Physique
Nucléaire
de Lyon

Université Claude Bernard

IN2P3 - CNRS

LYCEN 9869
DELPHI 98-150 PHYS 794
September 1998

Measurement of the W mass in $WW \to q\bar{q}q\bar{q}$ at 183 GeV with neural network

A. Duperrin

IPN Lyon IN2P3/CNRS, Université Claude Bernard,
F-69622 Villeurbanne Cedex, France

43, Boulevard du 11 Novembre 1918 - 69622 VILLEURBANNE Cedex - France

DELPHI Collaboration

DELPHI 98-150 PHYS 794
13 August, 1998

Measurement of the W Mass in $WW \to q\bar{q}q\bar{q}$ at 183 GeV with neural network

A. Duperrin
Université Claude Bernard de Lyon, IPNL, IN2P3-CNRS,
F-69622 Villeurbanne Cedex, France

Abstract

The mass of the W boson is obtained from the reconstructed invariant mass distribution in the fully hadronic channel. The sample of WW pairs is selected from 54 pb^{-1} collected with the DELPHI detector during the 183 GeV run in 1997. The selection of the events, and the assignment of jets are performed with two neural networks. The distribution of the two invariant masses is then fitted by a binned likelihood method to obtain the mass of the W boson

$$m_W = 80.126 \pm 0.183(stat.) \pm 0.050(syst.) \pm 0.100(BE/CR) \pm 0.030(LEP)\ \mathrm{GeV}/c^2$$

1 Introduction

This note describes a new measurement of the W mass in the WW $\rightarrow$ $q\bar{q}q\bar{q}$ channel from the data collected at 183 GeV in 1997 by the detector DELPHI. The integrated luminosity was 53.5 pb^{-1}. The note is organised as follows : a brief description is first given of the Monte Carlo samples used in the analysis. Then the event selection and the jet assignment with neural net methods are described. Finally, the W mass is extracted from a 2-dimensional likelihood fit of reweighted Monte Carlo events.

2 Monte Carlo samples

The simulated events allow to predict for each W mass the expected mass distribution which will be used in the fit. We have used 410000 fully simulated W events in this analysis. Among these events, 370K were generated with EXCALIBUR [1], including all four-fermion diagrams, and 40K with CC03 PYTHIA [2]. The EXCALIBUR sample is divided into 5 subsamples with different values of the generated W mass m_{gen}: 45K at 79.35, 60K at 79.95, 160K at 80.35, 60K at 80.75, and 45K at 81.35. These 5 samples are reweighted to the same mass in the analysis, as described in section 6.
To train the neural network algorithm, 55K WW $\rightarrow$ $q\bar{q}q\bar{q}$ events, restricted to the generator level, have been generated with PYTHIA (see subsection 5.3).
In addition to the signal events, 1 million $e^+e^- \rightarrow q\bar{q}(\gamma)$ background interactions were generated with the program PYTHIA.

3 WW $\rightarrow$ $q\bar{q}q\bar{q}$ event selection

Using the Monte Carlo sample generated at m_W = 80.35 with PYTHIA, the expected observable cross sections for the signal and the background are summarized in Table 1. They are calculated from the numbers of events surviving the selection cuts described below and the mass reconstruction procedures described in Section 5.

Process	σ_{cuts} (pb)
$e^+e^- \rightarrow WW \rightarrow q\bar{q}q\bar{q}$	5.830 $\pm$ 0.087
$e^+e^- \rightarrow q\bar{q}(\gamma)$	1.348 $\pm$ 0.013
$e^+e^- \rightarrow ZZ$	0.178 $\pm$ 0.008

Table 1: Cross sections of signal and background processes after analysis cuts.

A neural network [3] is used to tag the signal at 183 GeV since it is clearly more efficient than the simple sequential cuts. It allows an improvement of the purity by 14%, when compared to a standard procedure using the same variables. The efficiency and the purity were estimated to be 85% and 80% respectively. The number of events selected from the data was 401, which corresponds to the expected numbers of events from the simulation with 316 WW, 9.6 ZZ, and 72.8 $q\bar{q}(\gamma)$. A preselection is applied with the following cuts :

- The visible energy is above 105 GeV and the neutral energy less than 110 GeV.
- There are at least 20 charged tracks.
- The effective center of mass energy is larger than 150 GeV.
- The event is forced into 4 jets using DURHAM [4] and each jet is requested to have at least 4 particles. The number of jets will be unfrozen later on for the mass reconstruction.

A 29-29-1 neural network pattern is then trained on the full simulation. The 29 variables used for the learning and the computation of the neural network output are the variables of the preselection, the electromagnetic and hadronic neutral energy, jet masses, the particle multiplicity in jets, the number of reconstructed jets with a $y_{\rm cut} = 0.001$, all the $y_{\rm cut}$ to form 2 to 8 jets, the quark-quark-gluon-gluon matrix element [5], the Fox Wolfram moments, the thrust, the oblateness. Figure 1 shows the neural network output for the signal WW $\rightarrow$ $q\bar{q}q\bar{q}$ events, generated with PYTHIA, at $m_W = 80.35$ GeV/c^2, and the $q\bar{q}$ and ZZ background contribution. The overall agreement between data and Monte Carlo is good. The neural network output is required to be larger than 0.5 for the selected W hadronic events.

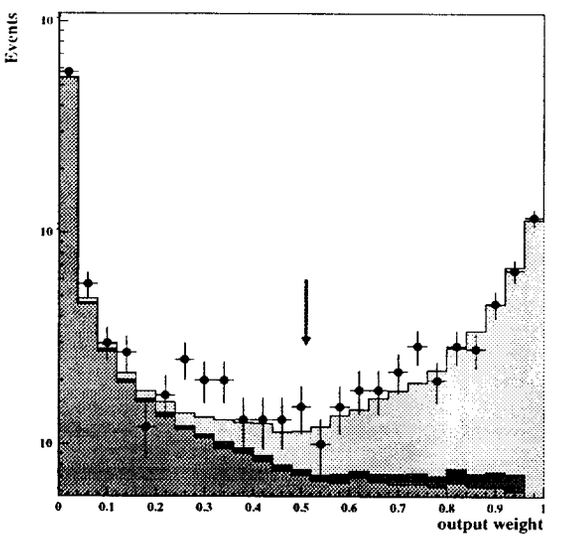

Figure 1: The neural network output: the black circles are the data points, the light shaded histogram the signal events, and the dark-shaded histogram the $q\bar{q}(\gamma)$ and ZZ background. The arrow indicates the selected range. The simulated histograms are normalized to 53.5 pb^{-1}.

4 Kinematical fit

In order to improve the mass resolution, a kinematical fit, called MULTIJETS [6] has been developed for multijet final states to impose the constraint of energy-momentum conservation at the jet level in the final state. We assume that the quality of the tracking and the granularity of the calorimeters are good enough, so that the direction of the jets is not altered in the fit, while the absolute value of the momentum is modified. It can be checked from the simulation that 77% of the reconstructed jets are within 5 degrees of the generated direction. A χ^2 is defined taking into account the individual momentum measurement errors of the jets σ_i which are parametrised as function of the polar angle. The momentum and energy conservation (4C fit) are imposed by Lagrange multipliers. A fifth constraint can be added. The 4C fit has been chosen since it gives informations on two masses for each events and moreover no improvement has been obtained using a 5C fit. Details of the constrained fit method are given in reference [6], with a minor modification. The masses of the jets, fixed in the previous reference, are now rescaled as the energy :

$$\frac{\Delta M_i}{M_i} = \frac{\Delta P_i}{P_i} = \frac{\Delta E_i}{E_i}$$

where M_i, P_i, and E_i are respectively the mass, the momentun modulus, and the energy of the measured jet i. The corrected momentum is then

$$\overrightarrow{P_i'} = \overrightarrow{P_i}(1 + \frac{\Delta P_i}{P_i})$$

with

$$\Delta P_i = {\sigma_i}^2 \left[\alpha \frac{E_i}{P_i} + \overrightarrow{\beta} . \frac{\overrightarrow{P_i}}{P_i} + \gamma z_i \left(E_A \frac{E_i}{P_i} - \overrightarrow{P_A} . \frac{\overrightarrow{P_i}}{P_i} \right) \right]$$

where α, $\overrightarrow{\beta}$ and γ are the Lagrange multipliers. E_A is the sum of jets energy belonging to one W , $z_i = 1$ for these jets and $z_i = 0$ for the jets belonging to the opposite W.

5 Jet pairing

The jet multiplicity distribution is sensitive to the choice of the parameter defining their size in the DURHAM algorithm. It was found advantageous to chose a small jet 'size' with $y_{\rm cut} = 0.001$ in order to isolate a very pure 4-jet sample. We see in figure 2 that the simulated and reconstructed jet multiplicities with this definition are in agreement. When the final particle configuration is not enforced into 4 jets, the possibility of a better assignment of the jets after they have been kinematically corrected is left open, but the combinatorial background quickly becomes severe.

The events are separated into two classes with different neural net parameters after the 4C fit : the 4-jet events, and the others.

5.1 4-jet associations

The chosen value of $y_{\rm cut} = 0.001$ yields a statistics of 107 WW $\rightarrow$ q$\bar{\rm q}$q$\bar{\rm q}$candidates events. The jets are coupled in three different ways. The output layer of the neural net includes 4

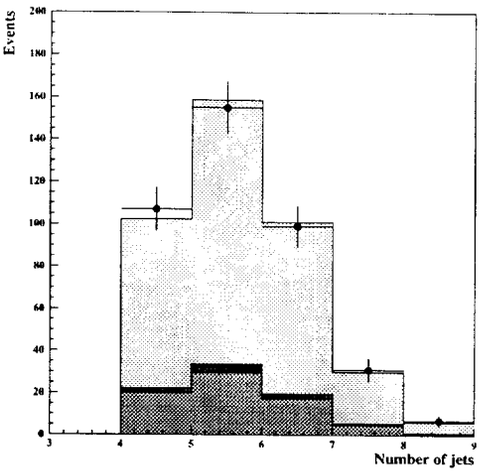

Figure 2: Distribution of the number of reconstruted jets with $y_{\rm cut} = 0.001$ using DURHAM. The black circles are the data points, the light shaded histogram is the signal, and the dark-shaded histogram the $q\bar{q}(\gamma)$ and ZZ background.

nodes : one for each of the 3 previous combinations (node i is equal to 1 if combination i is correct), and one for the background (node 4 is trained with 1 for background events). For each combination, two input variables are used to discriminate between the different solutions, the sum of the jet opening angles (interjet angles) $\theta_{ij} + \theta_{kl}$, and the difference between the two reconstructed masses $|M_{ij} - M_{kl}|$. The hidden layer includes 6 nodes (as many as the input variables). The distributions of these variables are shown in figure 3 for the three possibilities and for the solution chosen by the neural net. The W bosons have a small boost and the pair of quarks is almost back to back, resulting into an angular sum in the range of 200 to 300 degrees. The difference of the two masses is peaked around zero and follows a gaussian-like distribution, as after the 4C fit, the masses of the two W are close to each other. It is also seen in figure 3 that the third combination, which pairs the most energetic jets with the weakest, the good one in more than 70% of the cases.

5.2 5-jet associations

When the final state configuration is reconstructed with 5 jets the combinatorial background increases to ten differents combinations. The jets are coupled into two di-jets for one W and three jets for the opposite W. When the jets are reordered as a function of their energy, only seven combinations are retained out of the ten possibilities, as the bi-jets combinations 3-4, 3-5, 4-5, occur in 2% of the cases only.

Three variables are used to discriminate among the different combinations :

- The opening angle between the two jets in the bi-jet side $\theta_{ij} + \theta_{kl}$

- The absolute difference between the two associated masses of the 'candidate' W's : $|M_{ij} - M_{klm}|$ (as with the 4 jets).

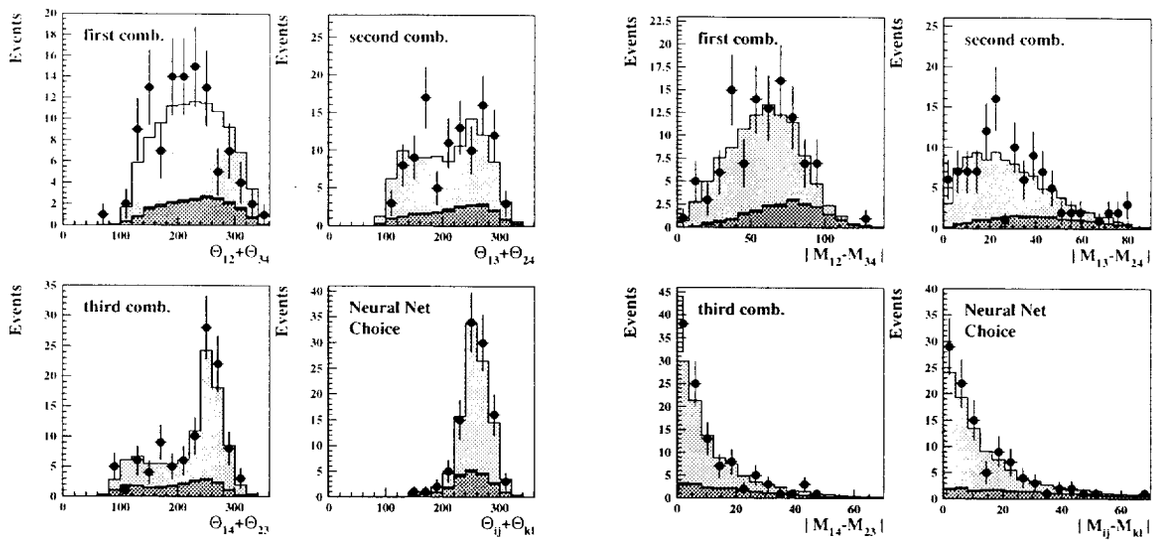

Figure 3: Distributions of the sum of the jet opening angles (left) and the bi-jet invariant mass difference (right) for the three differents pairing possibilities in the 4 jets events. The fourth histogram shows the selected combination. The black circles are the data points, the light shaded histogram is the signal, and the dark-shaded histogram the $q\bar{q}(\gamma)$ and ZZ background.

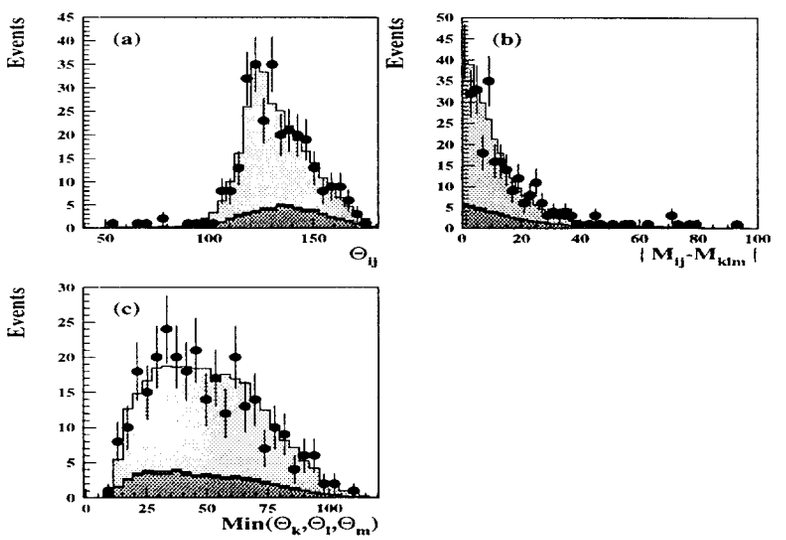

Figure 4: Selection variables of 5-jets events. The jet opening angles sum (a), the bi-jet invariant mass difference (b), and the minimal angle between the jets in the three-jet side for the selectect combination by the neural network. The black circles are the data points. the light shaded histogram is the signal, and the dark-shaded histogram the $q\bar{q}(\gamma)$ and ZZ background.

- The minimal angle between the jets in the three-jet side. We then favour solutions where a gluon is emitted with a softer jet at low angle.

The 6 and more jet events are re-clustered as 5 jets for the pairing. The hidden layer consists of 21 nodes and the output layer includes 8 nodes, one for each of the selected combinations, and 1 for the background. The distributions of the three discriminating inputs are shown for the combination selected by the neural network in figure 4. It is seen once more that these variables are well described by the simulation.

5.3 Training of the neural network

The training of the neural network to the recognition of the correct jet association will be performed with the generated particles (i.e. before the simulation of the detector). to avoid the ambiguities arising in the assignment of a particle to a given W at the detector level. A generated jet is then assigned to one W if the energy fraction from this W is larger than 50%. It turns out that 85% of the jets have more than 90% of their energy coming from a single W. The matching of the generated and detector jets allows to monitor the performance of the neural network. Two jets are matched according to the smallest angle between their directions. It is seen in figure 5 that although the neural network was trained with the generated hadrons, the energy distributions of the generated and detected jets (after the 4C fit) are very similar. At this stage, the invariant mass

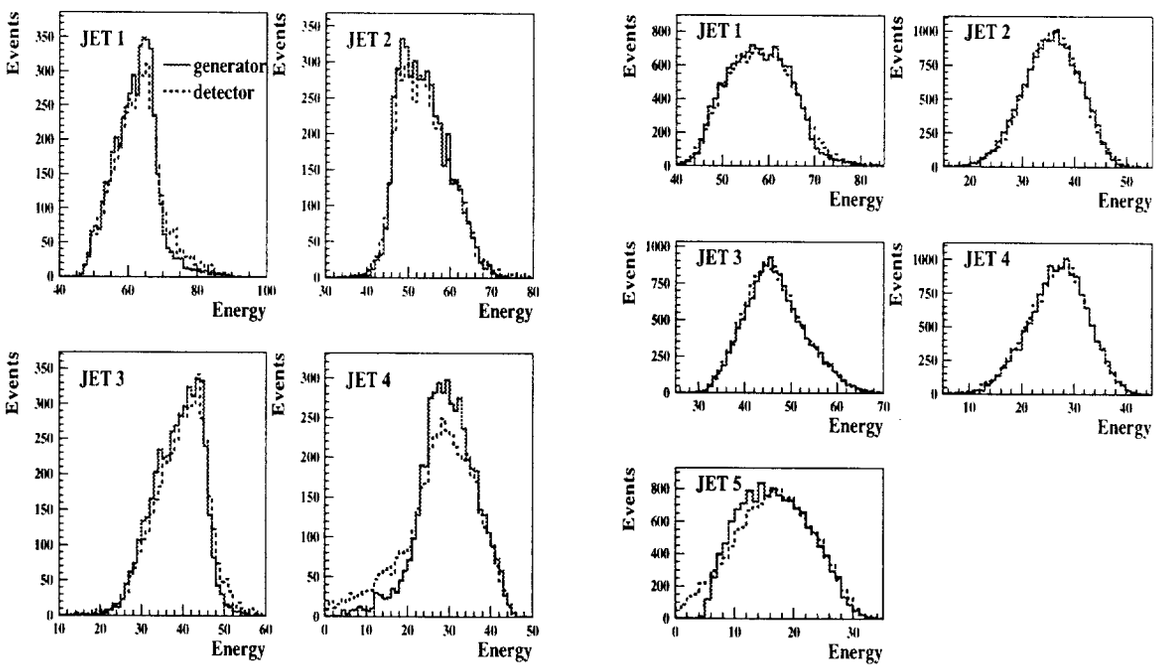

Figure 5: Jet energy comparison for the 4-jet (left) and for the 5-jet (right) between generated jets (solid line) and detected jets (dotted line) after a 4C fit.

of the generated jets belonging to the same W is computed, and called "true W" mass. A comparison between the true W mass distribution and the one found by the neural

network at the detector level is shown in figure 6. It is seen that the quality and the efficiency of the neural network training on the generated particles is quite adequate.

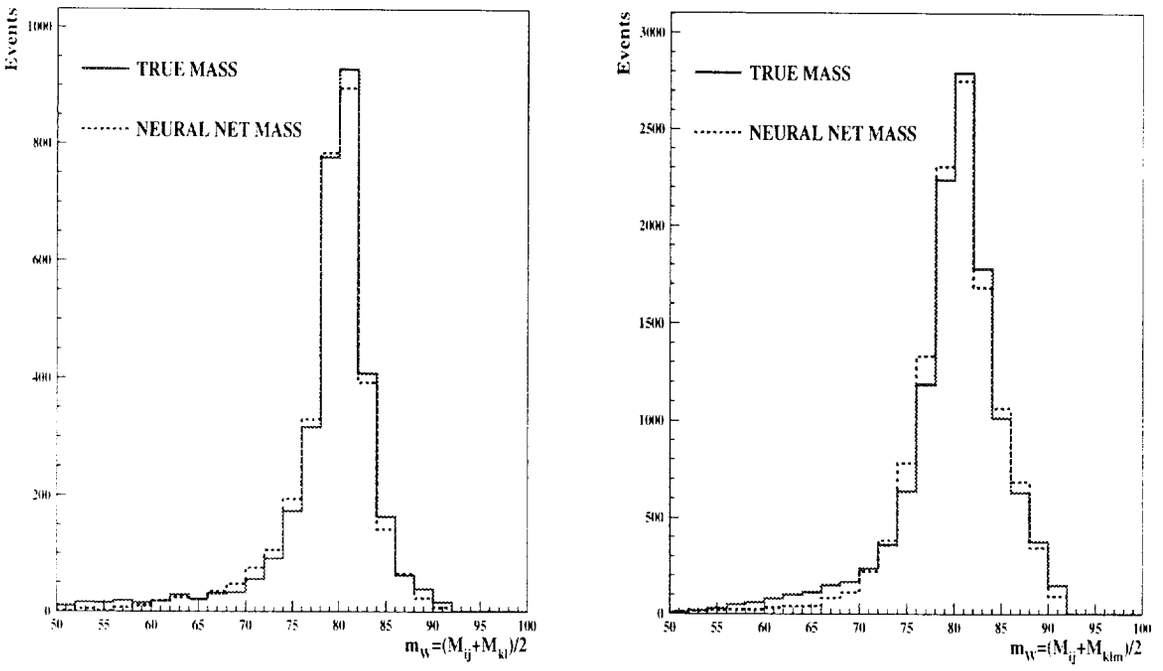

Figure 6: Comparison between the true W mass and the reconstructed one by the neural network for the 4 jets (right) and the 5 jets (left).

To avoid the bias arising from a single generation mass, the neural network has in fact been trained with 55k W^+W^-events with generated masses spread uniformly in the range $75 < m_W < 86$ GeV/c^2.

6 The 2-dimensional likelihood fit

When the two reconstructed W masses of an event differ, the reconstruction accuracy can generally be expected to be poor. The difference of the two masses, after a 4C fit, is then an indicator of the reconstruction quality of W^+W^- events. The W boson mass is extracted from a likelihood fit to the 2-dimensional plot formed by the average and the difference of the two W-masses, using the distribution predicted by the full simulation. In order to obtain the Monte Carlo spectrum for arbitrary values of m_W, the Monte Carlo reweighting technique is used as in [7].

The background distributions are normalised to the expected number of background events, and are assumed to be independent of m_W. A binned log-likelihood fit to the data is performed. The fit is restricted to the range $50 < \frac{M_1+M_2}{2} < 89$ GeV/$c^2$and $0 < |M_1 - M_2| < 80$ GeV/c^2. For an average mass above 65 GeV, and a difference of masses lower than 40 GeV, there are 7×21 bin uniformly distributed, and 153 bins are employed

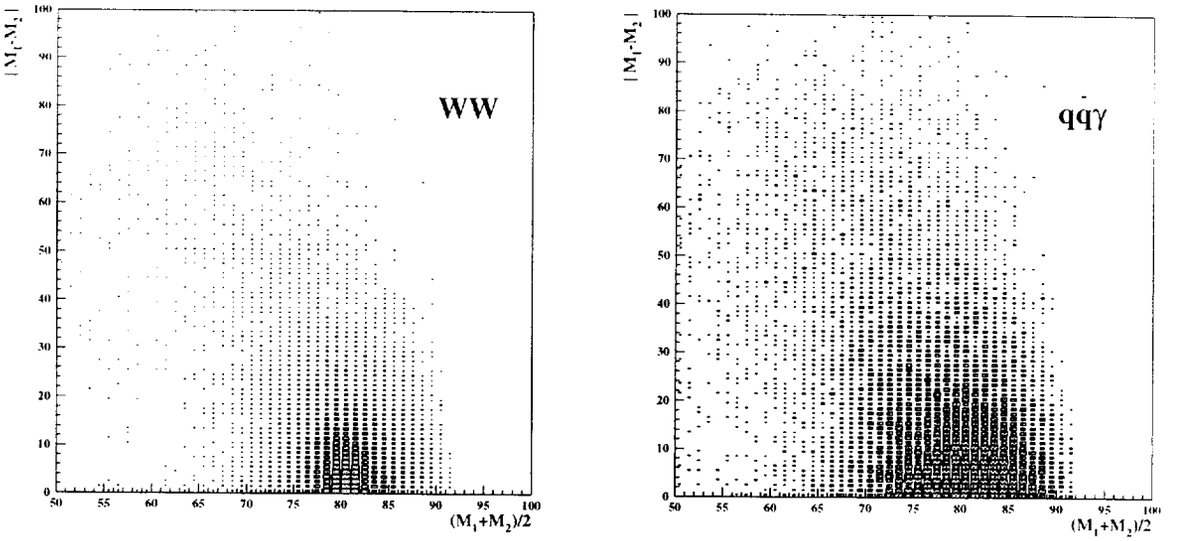

Figure 7: The mass difference versus the average mass for the W^+W^-(left) and the background (right). The relative normalisation is arbitrary. A 2-dimensional binned likelihood fit is applied to these distributions to extract the W mass.

to cover the full 2-dimensional domain. The log-likelihood is formed as

$$\mathcal{L}(\mathrm{m_W}) = -2. \sum_{i}^{\mathrm{N^i_{bins}}} \sum_{j}^{\mathrm{N^j_{bins}}} \mathrm{N^{data}_{i,j}} log(\frac{\mathrm{N^{MC}_{i,j}}}{\mathrm{N^{MC}_{Tot}}})$$

where $\mathrm{N}_{i,j}$ is the number of observed events in the (i,j)th bin and N^{MC}_{Tot} is the total number of Monte Carlo events assuming a W boson mass of $\mathrm{m_W}$. The minimum of $\mathcal{L}$ will be reached when the simulated expected $\mathrm{m_W}$distribution reproduces the data distribution. This method includes all the effects from the detector and the jet-clustering, and avoids the analytical parametrisations of the shapes, needed in an unbinned likelihood. In order to limit the statistical fluctuations from large weights, the 5 Monte Carlo samples with generated masses between 79.35 and 81.35 have all been used to reweigh each distribution to a new W mass $\mathrm{m_W}$, leading to an improved Monte Carlo statistics. For each sample with a given generated W mass m_{gen}, the weight $w(i)$ of event i will be:

$$w(i) = \frac{|\mathcal{M}_{\mathrm{m_W}}(i)|^2}{|\mathcal{M}_{m_{gen}}(i)|^2}$$

where the matrix elements $\mathcal{M}_{\mathrm{m_W}}$, or $\mathcal{M}_{m_{gen}}$ computed by the EXCALIBUR programme take into account all the 4 fermion contributions. The effective statistics is $\mathrm{N_{eff}}$ events :

$$\mathrm{N_{eff}} = \frac{\left(\sum_i^{\mathrm{N}} w(i)\right)^2}{\sum_i^{\mathrm{N}} w(i)^2}$$

where N is the number of Monte Carlo events. The different samples of generated masses are then combined with the proper statistical significance.

As the final states with jet multiplicities of 4, 5, and above, have very different mass resolutions, the overall likelihood used in the fit is the sum of the three independent

likelihoods, evaluated with the 2 dimensional Monte Carlo probabilities pertinent to each class of final states.

7 Results

The mass value and the statistical error obtained from a 2-dimensional likelihood fit to the data using our neural network are

$$m_W = 80.126 \pm 0.183(stat.)$$

The expected statistical error from the Monte Carlo is ± 0.215 GeV/c^2. Figure 8 shows the mass distribution for the 183 GeV data compared to the reweighted Monte Carlo prediction at the fitted mass.

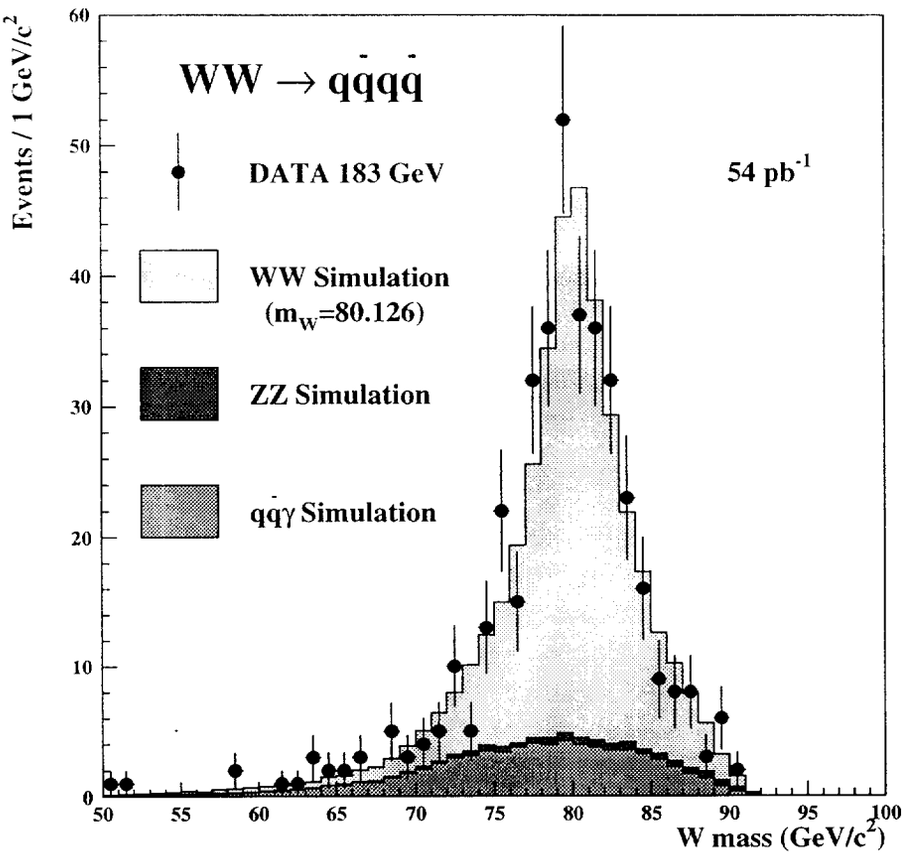

Figure 8: Reconstructed mass distributions for the data (points with error bars) compared to the reweighted Monte Carlo expectation (solid histogram) corresponding to the fitted mass using a neural network analysis for the process WW $\rightarrow$ q$\bar{\text{q}}$q$\bar{\text{q}}$.

8 Consistency and stability checks

8.1 Linearity of the reweighting technique

The consistency of the reweighting method has been checked on the independent Monte Carlo samples. We find with the fitting programme :

$$m_{fit} - m_{data} = \text{slope} * (m_{gen} - m_{data}) + \text{shift}$$

slope	1.05 ± 0.07
shift (GeV/c^2)	0.01 ± 0.03

Table 2: test calibration of the algorithm.

The slope is consistent with one, and the shift with zero, which confirms the validity of the reweighting technique.

8.2 Event selection

The events are selected from the data and Monte Carlo by requiring the neural network output to be larger than 0.5. This cut was varied from 0.4 to 0.6 and the measured mass changed by less than 4 MeV. The statistical error remains the same.

8.3 Shapes

No apparent discrepancy are observed between the shapes of the data and simulated events for all the variables used in the neural network algorithm.

8.4 Clusterisation Algorithm

Changing the value of y_{cut} around 0.001 does not modify the previous result on the W mass. In order to investigate the effect of the clusterisation algorithm, all the analysis has been repeated with the CAMBRIDGE jet algorithm [8] instead of DURHAM. The statistical error is respectively 5 MeV/$c^2$and 20 MeV/c^2 greater with the CAMBRIDGE jet algorithm for the Monte Carlo and the data. A downward shift of the mass by 75 MeV/c^2 is however found with this algorithm. Part of this shift at least originates from statistical fluctuations in the sample of jets, and should not be fully included as a systematic error. This large shift has to be investigated with more Monte Carlo on farther studies.

8.5 Background contamination

The expected background remaining after the selection is 18%. Changing the level of the background by $\pm 10\%$ introduces a systematic uncertainty of $\Delta m_W = 5$ MeV/c^2.

8.6 Preliminary sources of systematic errors

The analysis of all the systematic uncertainties has not been performed in this note and we borrow some numbers in reference [9]. More profound studies are in progress.

Source of systematic error	Error (MeV/c^2)
Fragmentation	30
Calorimeter calibrations	25
Jet energies	20
Background level	5
I.S.R.	20
LEP energy	30
Reference MC statistics	10
Colour reconnection and Bose Einstein corr.	100

Table 3: Contributions to the preliminary systematic error

9 Result for $e^+e^- \rightarrow WW \rightarrow q\bar{q}q\bar{q}$

This neural network analysis on the fully hadronic channel gave the following result

$$m_W = 80.126 \pm 0.183(stat.) \pm 0.050(syst.) \pm 0.100(BE/CR) \pm 0.030(LEP) \text{ GeV}/c^2$$

where "stat" denotes statistical error, "LEP" the uncertainty on the beam energy, "BE/CR" comes from the Bose-Einstein and Colour Reconnection systematics.

References

[1] F.A. Berends, R. Kleiss and R. Pittau, Nucl. Phys. B424 (1994) 308;
Nucl. Phys. B426 (1994) 344;
Nucl. Phys. (Proc. Suppl.) B37B (1994) 163-168;
R. Pittau, Phys. Lett. B335 (1994) 490-493;
R. Kleiss and R. Pittau, Comp. Phys. Comm. 83 (1994) 141.

[2] T. Sjöstrand, PYTHIA/JETSET, CERN-TH. 7112/93

[3] Stuttgart Neural Network Simulator,
http://www.informatik.uni-stuttgart.de/ipvr/bv/projekte/snns/snns.html

[4] S. Bethke, *Jets in Z0 decays*, OPAL-CR084, and HD-PY 92/12 (AACHEN QCD WORKSHOP 1992)

[5] A. Ballestrero, V.A. Khoze, E. Maina, S. Moretti, and W.J. Stirling, *Perturbative rates and colour rearrangement effetects in four-jet events at LEP2*, Cavendish-HEP-95/15, DFT 04/95, DTP/95/10 (1996)

[6] A. Duperrin, *W mass reconstruction using MULTIJETS*, DELPHI 97-170 PHYS 745

[7] F. Cossuti, A. Ouraou, *Reweighting of EXCALIBUR events for 4 fermion physics at LEP200*, DELPHI 98-9 PROG 231

[8] S. Bentvelsen, I. Meyer, *The Cambridge jet algorithm features and application.* CERN-EP/98-043, march 1998,
http://wwwcn1.cern.ch/ stanb/ckern/ckern.html

[9] W mass team, *Measurement of the W boson mass and width in* e^+e^- *collisions at* $\sqrt{s}$ *= 183 GeV*, DELPHI 98-85 CONF 153