Mixed Lorentz Boosted Z^0's

N. Kjær
CERN

M. Mulders
NIKHEF

Abstract

A novel technique is proposed to study systematic errors on jet reconstruction in W physics measurements at LEP2 with high statistical precision. The method is based on the emulation of W pair events using Mixed Lorentz Boosted Z^0 events. The scope and merits of the method and its statistical accuracy are discussed in the context of the DELPHI W mass measurement in the fully hadronic channel. The numbers presented are preliminary in the sense that they do not constitute the final DELPHI systematic errors.

1 Introduction

The measurement of the W boson mass is one of the major topics in LEP2 research. It provides a precision test of the Standard Model and a way to further constrain the predictions for the Higgs boson mass. With a final expected statistical error on the W mass of 20 MeV/c^2 for the four LEP experiments combined, it is very important to control systematic uncertainties to a level of 10 MeV/c^2 or -where possible- even lower. This paper focuses on systematic errors related to jet reconstruction, which have proven to be some of the most challenging systematic errors at the moment. Typical errors quoted for the W mass analyses of the four LEP collaborations at a centre-of-mass energy of 183 GeV [1-5], vary between 20 and 60 MeV/c^2 for jet fragmentation, and between 20 and 35 MeV/c^2 for detector effects.

Our knowledge about jet reconstruction errors is almost entirely based on Z^0 events. Events recorded at LEP1, or during the calibration runs at the Z^0 peak energy in between the high energy runs in 1997, 1998 and 1999 are used as 'template' to tune and check the detector alignment and calibration, efficiency and resolution, and to test and tune models describing jet fragmentation. The same models and detector simulation are then ported to the description of W pair events.

The conventional way to estimate errors on this description follows a similar approach: a 'realistic' uncertainty in the simulation of Z^0 data is derived from a comparison between Z^0 data and Monte Carlo and translated into a realistic 'shaking' of the high energy WW simulation (detuning fragmentation parameters, using different models or varying the description of the detector, etc.). The shift in the W mass obtained from the Monte Carlo sample is used as an estimate for this source of systematic error. This is repeated for different possible sources of shaking, and the systematic errors are added in quadrature or somehow combined taking into account the correlations between the different estimates. The limitations to this approach are:

- Each of the individual systematic errors is typically close to statistical sensitivity. This means that we often have to include statistical errors on the individual errors in our estimate of the systematic error.
- To take into account a 'complete' set of systematic effects, a considerable number of different shifts has to be determined and combined.
- It is not always clear how different estimates of systematic effects are correlated. The question often is how complementary and how realistic different shakings are.

This necessarily leads to conservative estimates of the systematic error.

The philosophy of the Mixed Lorentz Boosted Z^0 (MLBZ) method presented in this paper is to use Z^0 events from data and Monte Carlo simulation (MC) to emulate WW events *first*, and *then* do a comparison of the reconstructed W mass, applying the analysis directly on the MLBZ events (see figure 1). In this way a large subset of all possible detector effects and fragmentation model imperfections is probed at once, following a well-defined procedure that simplifies the intricate task of composing a complete, realistic list of effects and correctly handling the internal correlations to good approximation. It should be stressed already here, however, that not all effects are covered. A breakdown of the effects that are considered to be covered by the MLBZ method is given in table 2 and discussed further in section 4.1.

The effective statistical precision on a systematic shift in the W mass obtained with the MLBZ method with the example analysis described in this paper is around 300 MeV/c^2 per generated or detected Z^0 event [1], compared to a typical mass resolution of 3 GeV/c^2 per generated WW MC event [2] used in the conventional approach.

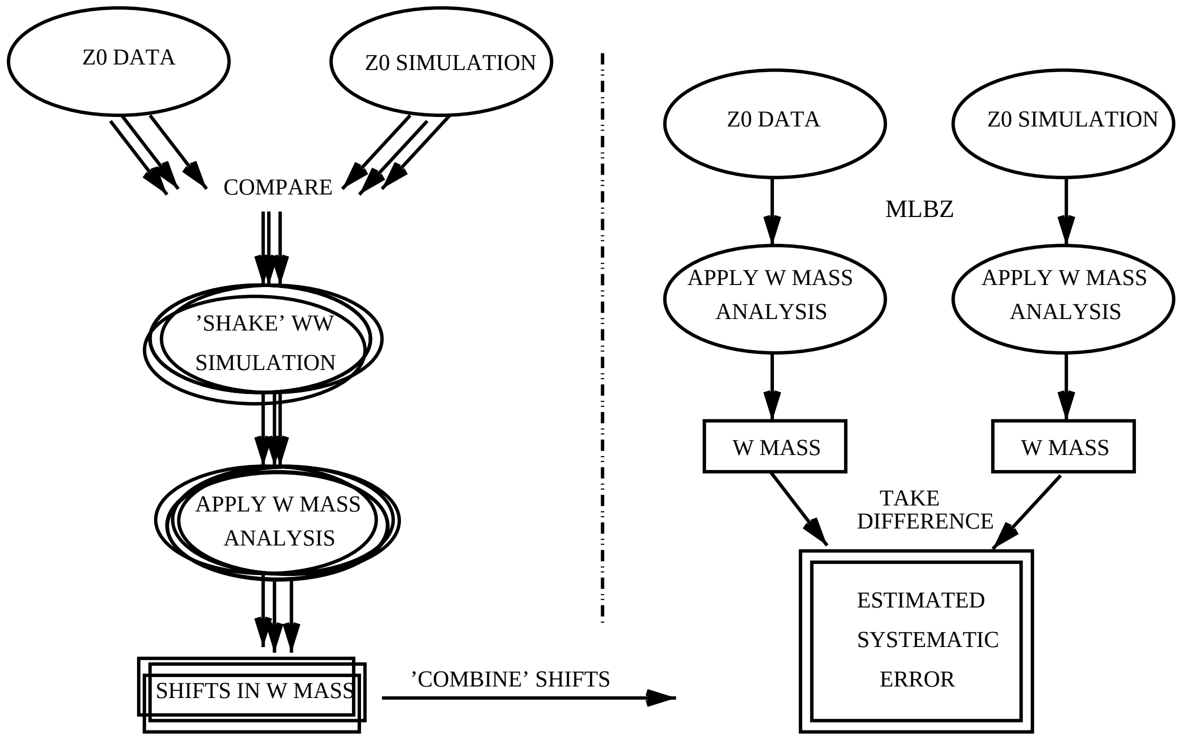

Figure 1: Conventional approach to estimate systematic errors vs. MLBZ

The aim of this paper is not to give definitive answers to all questions related to this new method, but merely to give a comprehensive description of a first implementation and present the first results. Hopefully this will serve as a basis for a fruitful discussion about this new and potentially very useful technique.

Section 2 gives a description of the MLBZ method itself and appendix A the technique used to determine the statistical precision. In section 3 the first results are presented, followed by a discussion of the possible limitations of the method in section 4. The final two sections contain an outlook and conclusions.

2 Description of the MLBZ method

The general outline of the MLBZ method [3] is as follows:

1. Select hadronic Z^0 events taken during the Z^0 calibration runs

[1] A better resolution can be obtained by increasing the number of mixed pairs per Z^0 event (i.e. increasing the sample size) or the number of boosts for each mixed Z^0 event pair. For the analysis described here available CPU time was a limiting factor.

[2] The statistical resolution of 3 GeV/c^2 is a convolution of the Breit Wigner width of the W ($\approx$2 GeV/c^2) and the average resolution per event ($\approx$2 GeV/c^2).

[3] Here only the emulation of fully hadronic WW events is mentioned. See section 5.1 for other channels.

2. Superimpose the measured 4-momenta of the particles of two Z^0 bosons after Lorentz boosting them in opposite direction with a boost typical for W bosons in the high energy run, thus emulating the topology of a fully hadronic WW event.

3. To fully exploit the information in the calibration data, Z^0 events should be used more than once by mixing them more than once and using more than one (isotropically distributed) boost direction per mixed event pair.

4. Apply the WW analysis of which the systematic effects are to be studied to the MLBZ events thus created.

5. Do this both on Z^0 events from data and from MC simulation and study the observed differences to draw conclusions about systematic errors.

2.1 Details of event selection and MLBZ procedure

For this paper the DELPHI W mass measurement at 183 GeV [3] in the fully hadronic channel was used as an example. The Z^0 events were mixed and Lorentz boosted in the following way:

1. The Z^0 events were collected at the beginning of 1997 with a beam energy of $E_{\mathrm{beam}}(t)$, varying slightly as a function of time t. It is not necessary to know $E_{\mathrm{beam}}(t)$ to a high precision. Instead a fixed beam energy of $E_{\mathrm{bZ}^0} = 45.625$ GeV was assumed.

2. On applying the runquality selection as used in [3], 1.6 pb^{-1} of Z^0 data remained.

3. The Z^0 candidates were selected with hadronic cuts giving a purity exceeding 99%:

   - at least 8 charged particles
   - carrying at least 15 GeV total energy
   - of which at least 3 GeV per hemisphere

   The measured angular distribution of jets in W^+W^- events is almost uniform (see figure 2), while Z^0 jets are distributed according to $1 + (\cos\theta)^2$, where $\cos\theta$ is the polar angle with the LEP beam. In order to have the same angular distribution of jets in the Z^0 candidates and the W^+W^- events, Z^0 events were randomly discarded according to the value of the polar angle of their thrust axis. 31557 events were selected in 157 samples of 201 events each, using approximately 94% of the data available. The sample size of 201 events was chosen as a compromise to have reasonable statistics but still be able to perform the full analysis (as described below) for one sample in a single 8 CPU-hour batch job. From Pythia/Jetset Z^0 MC 161 samples of the same size were selected. Some distributions of Z^0 event variables for data and MC are shown in figure 3.

4. The 4-momenta [4] of the measured particles in the Z^0 event were Lorentz boosted in a single random direction with a boost given by $\gamma_{\mathrm{boson}} = \frac{\sqrt{s_{\mathrm{high}}}}{2 m_{\mathrm{W}}}$, where $\sqrt{s_{\mathrm{high}}} =$ 182.7 GeV is an approximation to the LEP centre-of-mass energy of the high energy data (at 183 GeV) and $m_{\mathrm{W}} = 80.35\ \mathrm{GeV}/c^2$ is a nominal value of the W mass.

[4] Assuming pion masses for charged particles and photon masses for neutrals.

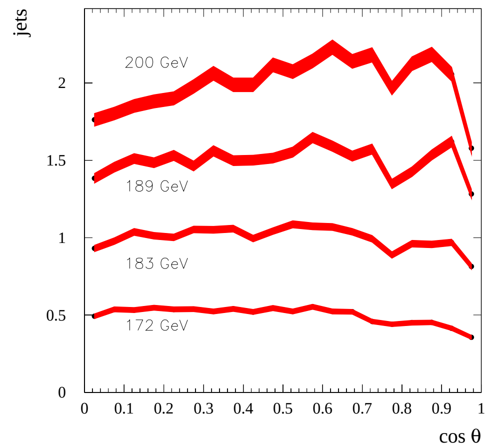

Figure 2: The distribution of the measured jets in simulated WW events as a function of the cosine of the polar angle with the beam, with arbitrary normalisation for different values of the centre-of-mass energy. The thickness of the line indicates the statistical uncertainty.

5. Another Z^0 candidate was then treated in the same way except the opposite direction of the boost was chosen. The particles of both boosted Z^0 events were mixed and a new event created. Each event was used 400 times by mixing it k times (k different isotropically distributed boost directions) with every other event from the same sample of 201 events. k was chosen to be 2, giving 40200 MLBZ events per sample.

6. These MLBZ events have nearly the same kinematical properties as W^+W^- events. The main differences come from the overall energy scale and Initial State Radiation (ISR). The MLBZ events were then treated as W^+W^- events and a fitted mass $m_{\mathrm{MLBZ}}^{\mathrm{sim}}$ extracted from simulated Z^0s and $m_{\mathrm{MLBZ}}^{\mathrm{data}}$ from the Z^0 data, using the 183 GeV W mass analysis [3], assuming a centre-of-mass energy $\sqrt{s_{\mathrm{MLBZ}}} = 4\gamma_{\mathrm{boson}}\mathrm{E}_{\mathrm{bZ}^0}$ in the constrained fits.

7. The difference of the fitted mass with respect to $2E_{\mathrm{bZ}^0}$ is interpreted as a measure for the experimental bias at the scale of $2E_{\mathrm{bZ}^0}$, the approximated Z^0 mass. Thanks to the use of relative errors on the jet energies in the constrained fit [6], the fit is largely invariant under a scale transformation of all energies and masses. Therefore the measured experimental bias is related to the experimental bias at the scale of the W mass by a factor $\frac{m_{\mathrm{W}}}{2E_{\mathrm{bZ}^0}} \approx \frac{m_{\mathrm{W}}}{m_{\mathrm{Z}}}$.

8. If the simulation of jet reconstruction in W^+W^- events is affected by systematic errors most of these will be the same for simulated MLBZ events. Thus the difference between the fitted mass on MLBZ data and MLBZ simulation, rescaled to the W mass scale,
$$\Delta m_{\mathrm{W}}^{\mathrm{MLBZ}} = \left(m_{\mathrm{MLBZ}}^{\mathrm{data}} - m_{\mathrm{MLBZ}}^{\mathrm{sim}}\right) \cdot \frac{m_{\mathrm{W}}}{m_{\mathrm{Z}^0}} \tag{1}$$
is a measure of the systematic error from non-perfect simulation on the W mass.

This measure can be improved by taking into account the difference in flavour composition between Z^0 and W boson decay products, as will be discussed in section 3.3.

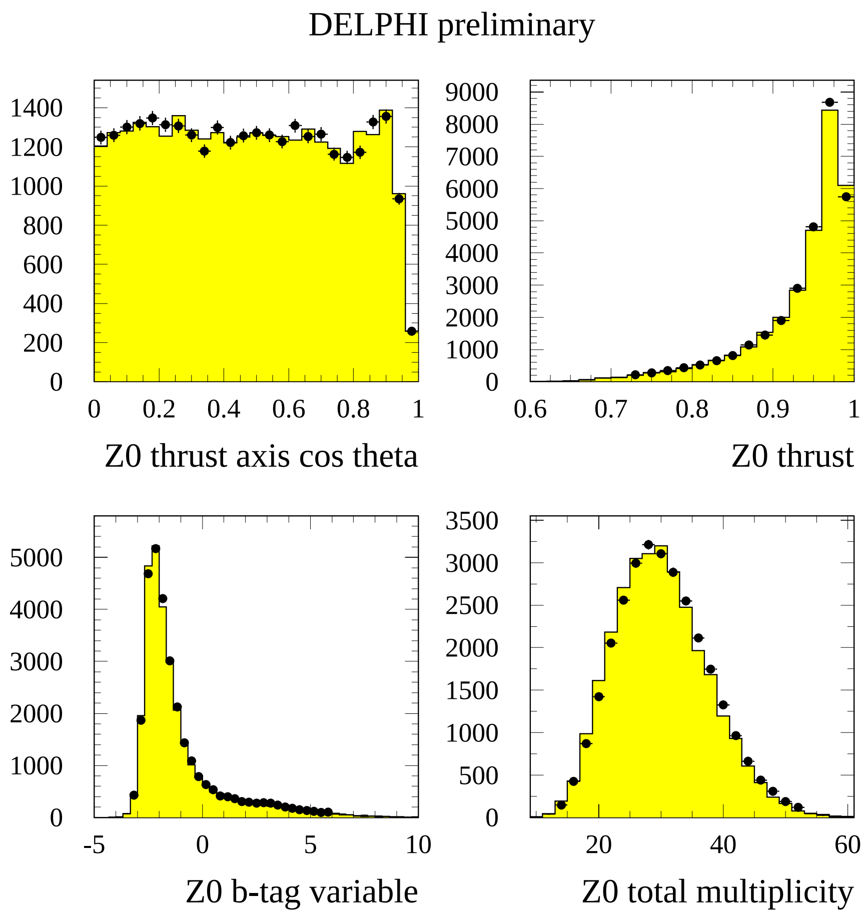

Figure 3: Comparison between 1997 data (the points with error bars) and MC (the filled histogram) for different event variables of the selected Z^0 events. The total Z^0 multiplicity includes both charged and neutral particles. The MC plots are normalised to the number of data events.

2.2 Statistical accuracy

When applying the standard analysis to the MLBZ events as described in the previous subsection, the statistical error on the measurement can not be determined in the usual way assuming that events are uncorrelated observations of the quantity that is to be measured. The Z^0 events are independent, but the MLBZ events are not.

Therefore a resampling technique know as the 'Jackknife' method [7, 8, 9, 10] was used. A full description of its implementation is given in appendix A. As the data was treated in such a way that many samples of 201 events are independent, a comparison with the traditional RMS estimate can serve as a cross-check of the Jackknife method at the level of the measured MLBZ mass per sample. This comparison is shown in table 1 and confirms that the Jackknife procedure used estimates the statistical errors correctly to within 10%.

	Z^0 data	Z^0 simulation
number of samples n_{samp}	157	161
uncertainty on mass per sample (MeV/c^2)		
average Jackknife estimate $\langle\sigma^{\text{MLBZ}}_{sample}\rangle$	23.7 ± 0.2	22.3 ± 0.2
cross-check: RMS of sample masses	22.0 ± 1.3	20.3 ± 1.1
all samples combined; uncertainty on mass (MeV/c^2)		
Jackknife estimate $\sigma^{\text{MLBZ}}_{all}$	1.9	1.8
cross-check: RMS of samples $/\sqrt{n_{\text{samp}}}$	1.8	1.6

Table 1: Comparison of different estimates of the statistical uncertainty on the mass measured with MLBZ events.

For the example analysis discussed here the statistical precision on the fitted MLBZ mass turns out to be $1/\sqrt{n_{Z^0}} \cdot 300$ MeV/c^2 at the scale of the W mass, where n_{Z^0} is the number of selected Z^0 events. This allows a precise determination of various effects to a precision of typically order of 2 MeV/c^2 or better.

3 Results

In this section the expected experimental bias on the W mass is compared to the observed bias in data and simulation, and the results are briefly discussed.

3.1 Expected reconstruction bias

The – average – reconstructed W mass obtained through kinematical reconstruction is not equal to the average 'true' or generated W mass. The difference of the two will from now on be called reconstruction bias:

$$b_{\text{rec}} = m_{\text{reconstructed}} - m_{'\text{true}'} \tag{2}$$

This bias is caused by a range of many more or less correlated effects. It is known from MC studies that the following two dominating effects give the largest contribution to the reconstruction bias:

- A positive bias due to ISR photons that are lost inside the beam pipe and not taken into account in the kinematical fit. At 183 GeV this causes an average positive shift of the order of 300 MeV/c^2 on the reconstructed mass.
- A negative bias due to the imperfect reconstruction of jets, which smears the masses preferably downwards. As will be shown later in this section this bias turns out to be of the order of -200 MeV/c^2 at 183 GeV.

The treatment of ISR falls outside the scope of this paper, as its effects cannot be studied with the MLBZ method, and will not be discussed further. The negative bias due to jet misreconstruction, on the other hand, can be studied very well using MLBZs, and is the main subject of this paper.

3.1.1 Reconstructing a boosted Z^0

The origin of the negative reconstruction bias can already be demonstrated with the simple case of a single boosted Z^0. When a Z^0 boson, produced at rest, decays hadronically and is detected by the DELPHI detector, the main error on the measured direction and energy of the jets is caused by missing particles. In this way

- on average the measured energy is 10-20 % lower than the Z^0 boson mass
- when clustered in two jets, the jets are not exactly back-to-back which means that the measured boson is moving in the laboratory frame with a 'measured' velocity (or boost) $\beta_{\rm meas}$.

This means that the invariant mass of the detected particles is typically 10-20% smaller than the 'true' Z^0 mass. When the Z^0 boson is boosted with a certain boost $\beta_{\rm boson}$ before or after the measurement, the detected invariant mass will remain the same [5]. This is no longer true, however, when applying a constrained fit to improve the measurement of the invariant mass beyond the detector resolution, irrespective of the specific analysis technique that is used. The DELPHI convolution method [3, 6] gives a bias in the reconstructed mass which is corrected for by calibration curves, while Monte Carlo reweighting methods [2, 4, 5] automatically correct for the bias in the procedure using simulation. The two methods behave very similar to deficiencies in the simulation, and are therefore affected by systematics in a comparable way.

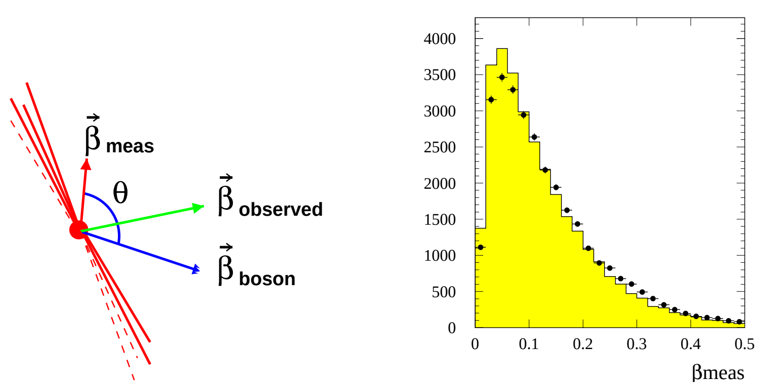

Figure 4: The observed boost $\beta_{\rm observed}$ is a combination of the true boost of the boson $\beta_{\rm boson}$ and the mismeasurement of the event $\beta_{\rm meas}$. The plot on the right shows the distribution of $\beta_{\rm meas}$ in data and MC.

To approximate the effect of a constrained fit, all particle momenta in our boosted Z^0 event are rescaled with a rescaling factor so that the total energy equals the 'expected'

[5] This is only perfectly true when the same particles are missed before and after the boost (uniformity of the detector), and assuming that the lack of Lorentz invariance of other resolution effects like momentum resolution on the reconstructed tracks is negligible.

energy $\gamma_{\rm boson} m_{Z^0}$, with $\gamma_{\rm boson} = 1/\sqrt{1-\beta^2_{\rm boson}}$. In this overall rescaling procedure the directions of the observed particle momenta and therefore the ratio $|\mathbf{p}|/E = \beta_{\rm observed}$ remain constant. The final observed invariant mass is then given by:

$$m_{\rm fit} = \frac{E_{\rm observed}}{\gamma_{\rm observed}} = \frac{\gamma_{\rm boson} m_{Z^0}}{\gamma_{\rm observed}} = \gamma_{\rm boson} m_{Z^0} \sqrt{1-\beta^2_{\rm observed}} \tag{3}$$

which is thus fully determined by the observed boost $\gamma_{\rm observed}$ or equivalently $\beta_{\rm observed}$ which is the combination of the artificial boost $\beta_{\rm meas}$ due to the resolution of the detector and the true boost $\beta_{\rm boson}$ given to the boson (see figure 4). The following expressions for $\gamma_{\rm observed}$ can be derived:

$$\gamma_{\rm observed} = \gamma_{\rm boson}\gamma_{\rm meas}(1+\beta_{\rm boson}\beta_{\rm meas}\cos\theta) \tag{4}$$

and substitution of $\gamma_{\rm observed}$ in equation 3 yields

$$m_{\rm fit} = \frac{m_{Z^0}}{\gamma_{\rm meas}(1+\beta_{\rm boson}\beta_{\rm meas}\cos\theta)} \tag{5}$$

where θ is the angle between the two boosts in the laboratory frame. For two MC events the distribution of measured masses is shown in figure 5, applying 100,000 boosts corresponding to different centre-of-mass energies in random directions (uniformly distributed in the laboratory frame). The peculiar 'box' shape is easily understood from equation 5,

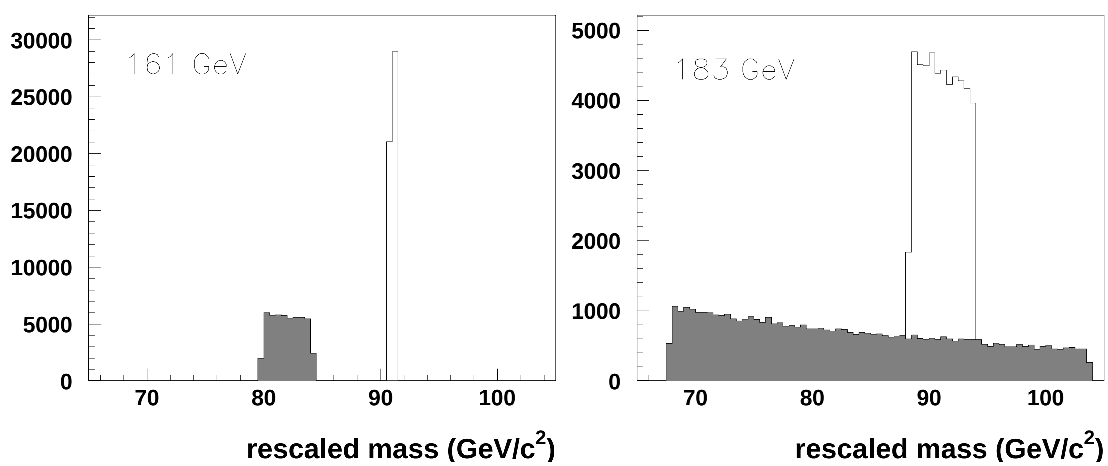

Figure 5: For two MC Z^0 events the distribution of rescaled masses is shown after boosting each event a 100,000 times in random directions. This is done for a boost corresponding to a WW event at $\sqrt{s}$=161 GeV (left) and $\sqrt{s}$=183 GeV (right). One event is measured with large $\beta_{\rm meas}$ (shaded histogram); the other with small $\beta_{\rm meas}$ (open histogram).

bearing in mind that $\vec{\beta}_{\rm boson}$ has a uniform angular distribution, corresponding to a flat distribution in $\cos\theta$.

For a perfectly measured event $\beta_{\rm meas} = 0$ and $\gamma_{\rm meas} = 1$ so that equation 5 reduces to $m_{\rm fit} = m_{Z^0}$. In other cases the average bias $m_{\rm fit} - m_{Z^0}$ and the RMS of this bias can be calculated analytically:

$$m_{\rm fit} - m_{\rm Z^0} = \frac{m_{\rm Z^0}}{2\alpha\gamma_{\rm meas}}\ln(\frac{1+\alpha}{1-\alpha}) - m_{\rm Z^0} = m_{\rm Z^0}\left[(\frac{\beta^2_{\rm boson}}{3} - \frac{1}{2})\beta^2_{\rm meas} + \mathcal{O}(\beta^4_{\rm meas})\right] \tag{6}$$

and

$$\sigma^2_{m_{\rm fit}} = \frac{m^2_{\rm Z^0}}{\gamma^2_{\rm meas}}\left[\frac{1}{1-\alpha^2} - \left(\frac{1}{2\alpha}\ln\left(\frac{1+\alpha}{1-\alpha}\right)\right)^2\right] = \frac{m^2_{\rm Z^0}}{\gamma^2_{\rm meas}}\left[\frac{\beta^2_{\rm boson}}{3}\beta^2_{\rm meas} + \mathcal{O}(\beta^4_{\rm meas})\right] \tag{7}$$

where $\alpha = \beta_{\rm meas}\beta_{\rm boson}$.

These formulas are in good agreement with the plots shown in figure 6 generated by boosting 500 $\rm Z^0$ events from MC 50,000 times each. They show an *increasing negative bias* for *increasing* values of $\beta_{\rm meas}$ (corresponding to worse jet reconstruction), and they also show that this effect becomes *less pronounced* for *increasing* values of the boson boost $\beta_{\rm boson}$. When extracting information about jet reconstruction systematics, the interesting

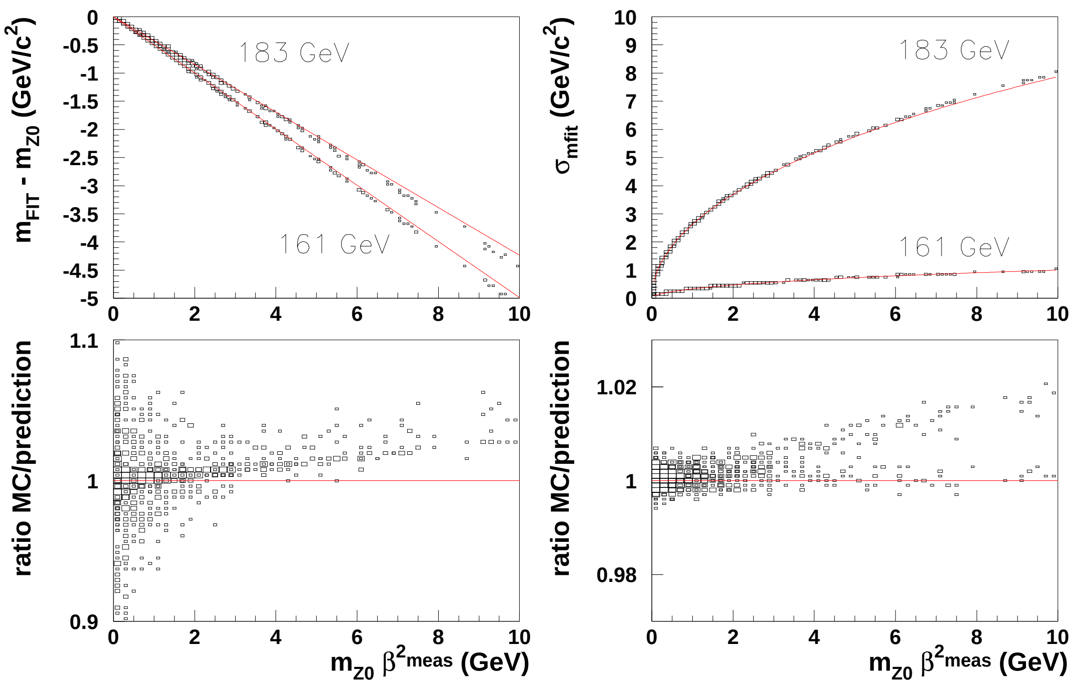

Figure 6: The bias (left) and error on the bias (right) as a function of the measurement boost, compared to the analytical predictions up to $\mathcal{O}(\beta^2_{\rm meas})$

quantity to obtain from a $\rm Z^0$ event is the average bias; not the measured mass obtained for a single boost. For a boost corresponding to 183 GeV, 400 boosts per $\rm Z^0$ event and a typical value of $m_{\rm Z^0}\beta^2_{\rm meas}$ of 0.5 GeV (see figure 4) the average bias is -210 MeV, and the relative precision per event on this bias is ≈50%. It can be seen from figure 4 that the peak in the distribution of $\beta_{\rm meas}$ in data is shifted to slightly higher values compared to simulation. The shift of the peak corresponds to an average shift in $m_{\rm Z^0}\beta^2_{\rm meas}$ of ≈ 0.05 GeV, which in turn would be equivalent to a systematic shift of -21 MeV on the rescaled $m_{\rm Z^0}$ mass.

This simplified experiment is only an approximation of the full MLBZ analysis, as in a real constrained fit the rescaling factor is not the same for all jets, and jet masses and transverse errors on the jet momenta are taken into account. Furthermore additional statistical effects from the interaction of the two mixed events (e.g. jet clustering ambiguities) and more realistic analysis details play a role.

3.1.2 Reconstructing a W pair event

It is evident that at WW production threshold, when the true jets of the W bosons are strictly back-to-back, any imperfection in the mass reconstruction will lead to a negative mass bias (analogous to the reconstruction of the Z^0 bosons described previously). At a centre-of-mass energy of 183 GeV, still a negative shift remains.

That the negative bias from jet reconstruction depends both on the W boost and on the quality of the jet measurement can easily be checked using a simplified WW simulation. In figure 7 the average bias is shown for 161 and 183 GeV, and jets generated as WW events with Breit-Wigner but without ISR. The measurement errors where assumed to be Gaussian according to the parameterisation as used in the fit of the 172 GeV analysis [6]. The transverse jet errors where multiplied with an additional factor X, and the bias plotted as a function of this factor (see figure 7). For the most realistic value of X=1, the bias at 183 GeV was found to be ≈ -200 MeV/c^2, in agreement with the bias seen in MLBZ events (see section 3.2).

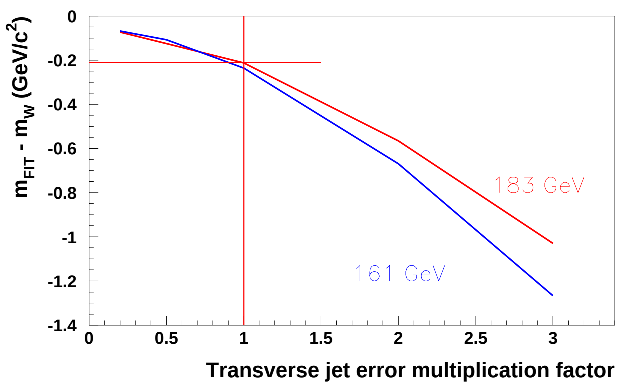

Figure 7: Bias as a function of a transverse jet error scaling factor in a simplified WW simulation, for centre-of-mass energies of 161 and 183 GeV.

Again it is seen that the negative bias

- Is smaller for higher centre-of-mass energies (183 GeV compared to 161 GeV)
- Becomes larger when the jet reconstruction becomes worse.

3.2 Observed reconstruction bias

An interesting feature of the MLBZ method is that the mass bias can be studied as a function of properties of the *individual* bosons, rather than the usual *pairs* of W bosons.

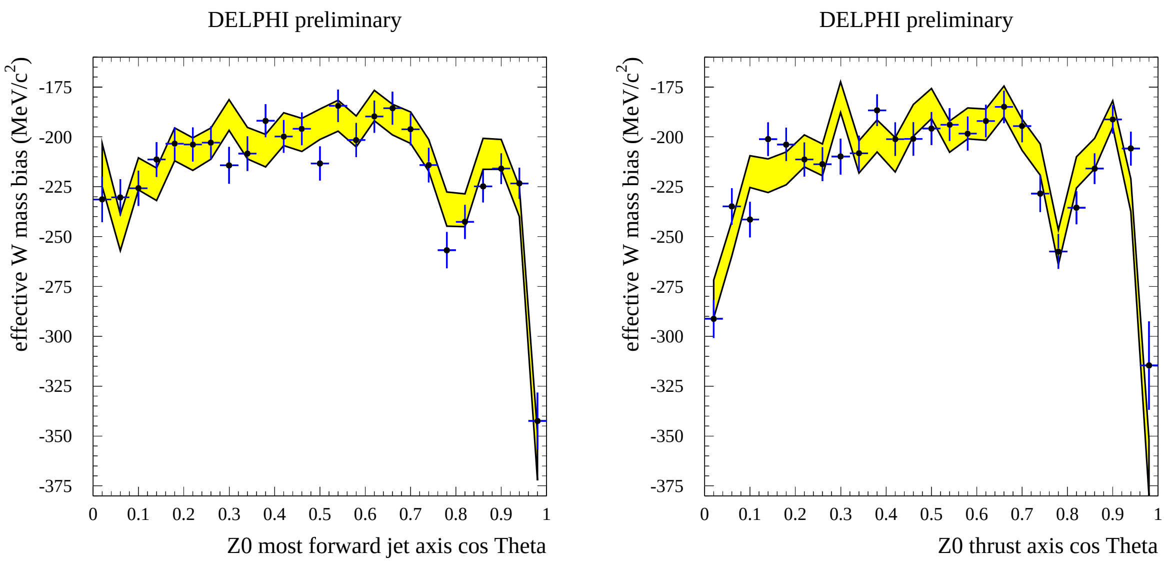

Figure 8: Effective W mass bias as a function of the polar angle with the beam, for 1997 simulation (light shaded error band) and data (points with error bars).

It can be shown (see appendix A and B) that the MLBZ bias is a linear sum of the reconstruction biases of the individual Z bosons:

$$b_{\mathrm{MLBZ}}^{kl} \equiv \frac{b_{\mathrm{Z}}^{k} + b_{\mathrm{Z}}^{l}}{2} \tag{8}$$

Hence one can use the relation

$$b_{\mathrm{rec}}^{i} = \frac{\langle b_{\mathrm{Z}} \rangle^{i} + \langle b_{\mathrm{Z}} \rangle}{2} \tag{9}$$

to define an individual boson bias $\langle b_{\mathrm{Z}} \rangle^{i}$ in terms of measurable quantities:

$$\langle b_{\mathrm{Z}} \rangle^{i} = 2b_{\mathrm{rec}}^{i} - b_{\mathrm{rec}} \tag{10}$$

where b_{rec} (2) is the overall MLBZ reconstruction bias and b_{rec}^{i} the reconstruction bias obtained when mixing only Z^0 events with a certain property i with all other Z^0 events (including the ones with property i). $\langle b_{\mathrm{Z}} \rangle^{i}$ should be interpreted as the average individual boson bias of all Z^0 events in bin i, while $\langle b_{\mathrm{Z}} \rangle$ is the average of this bias over all bins and has to be equal to b_{rec} by definition.

In figures 8 to 10 $\langle b_{\mathrm{Z}} \rangle^{i}$ is plotted as a function of individual Z^0 event variables. In all cases the bias has been rescaled to the W mass scale as in equation (1) and is denoted as 'effective W mass bias'.

In figure 8 the bias is shown as a function of the cosine of the polar angle with the beam. It is clearly visible that the reconstruction of jets in the forward region ($|\cos(\theta)| > .7$) is worse than in the barrel region, corresponding to a larger negative bias. The large negative bias for the most central bin ($|\cos(\theta)|$ close to 0) was unexpected and has not yet been understood. It is comforting, however, that all features in the data are described very well by the MC simulation, taking into account many effects of alignment, energy calibration, acceptance and detection efficiencies for all particle types integrated over the whole momentum range.

The reconstruction of jets is not only hampered by detector effects, but also by the broadening of jets due to soft and hard gluon radiation. Broader jets will be detected with larger uncertainties on the jet direction and therefore cause larger negative shifts on the mass. In addition broader jets will cause more confusion in the jet clustering effectively leading to a further deterioration of the jet reconstruction. As shown in figure 9 the reconstruction bias depends very strongly on the Z^0 event thrust and particle multiplicity. The negative bias increases with almost 20 MeV/c^2 per extra particle. There is a positive correlation between the amount of gluon radiation and the number of particles, giving broader jets (i.e. a lower thrust value) for larger multiplicities. Again the behaviour of the data is excellently described by the simulation, except for a significant discrepancy for very low thrust events which however corresponds to an overall shift in the W mass of only 2 MeV/c^2.

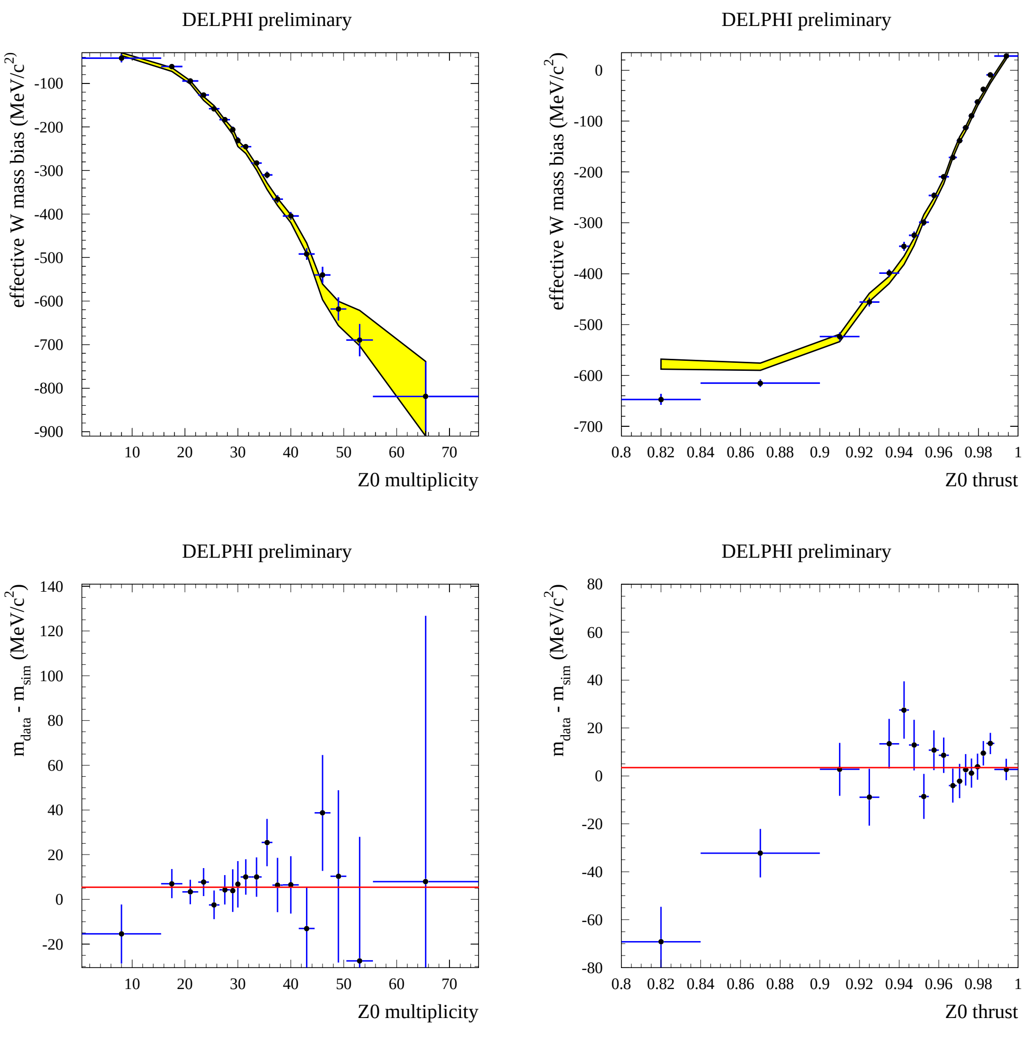

Figure 9: Effective W mass bias as a function of the multiplicity (charged + neutral) and the thrust of the Z^0 event, for 1997 simulation (light shaded error band) and data (points with error bars).

Another important variable is the DELPHI combined b-tag variable that will prove to be useful in studying light-quark jets and heavy-quark jets separately. Again the dependence visible in the data is followed nicely by the simulated event bias (see figure 10). Extended study revealed that the dependence of $\langle b_Z \rangle^i$ on the b-tag should not be ascribed to the true b-quark content, but rather to indirect correlations through the strong dependence on average multiplicity and thrust. When the expected bias is calculated from the average multiplicity of all the events in the corresponding b-tag bin using the dependence shown in figure 9, the main features of the shape of the curve are reproduced well in all bins except for the lowest and the highest b-tag bins, where the higher value of $\langle b_Z \rangle^i$ is correlated with the lower than average fraction of low-thrust Z^0 events.

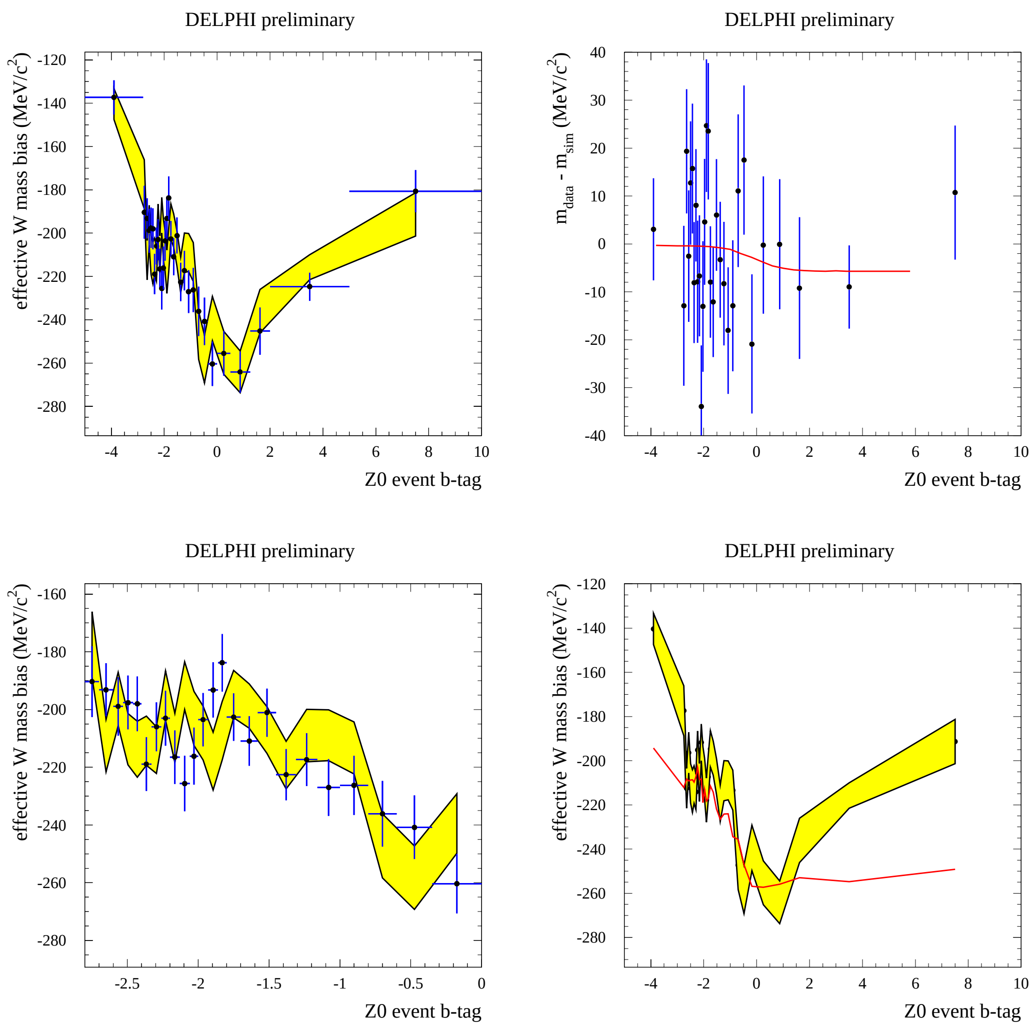

Figure 10: Effective W mass bias as a function of the b-tag variable, for 1997 simulation (light shaded error band) and data (points with error bars) on the left, the difference between simulation and data (top right), and the expectation from the average multiplicity in each bin superimposed on the MC result (bottom right). The bottom left plot zooms in on central part of the top left plot.

3.3 Comparison between data and Monte Carlo

As was shown in this section, the behaviour of the individual boson mass reconstruction bias $\langle b_{\rm Z}\rangle^i$ (10) in MLBZ events has been understood qualitatively, and is quantitatively described by the simulation to excellent precision.

Relevant for the final systematic error on the W mass is the average difference in reconstruction bias between data and MC simulation. Combining all analysed 1997 MLBZ data the overall difference $b_{\rm rec}^{\rm data} - b_{\rm rec}^{\rm sim} = \Delta m_{\rm W}^{\rm MLBZ}$ is equal to :

$$\Delta m_{\rm W}^{\rm MLBZ} = \left(m_{\rm MLBZ}^{\rm data} - m_{\rm MLBZ}^{\rm sim}\right) \cdot \frac{m_{\rm W}}{m_{\rm Z^0}} = -1.9 \pm 2.6 \ \ {\rm MeV/c^2} \tag{11}$$

This value has been scaled to the W mass scale as in equation (1).

As W bosons hardly ever decay into b quarks, it is interesting to determine $\Delta m_{\rm W}^{\rm MLBZ}|_{\rm b}$ for b-quark jets and $\Delta m_{\rm W}^{\rm MLBZ}|_{\rm udsc}$ for light-quark (u,d,s,c) jets separately. The following model, based on a linear dependence on the b-quark purity $P_{\rm b}$ was used:

$$\Delta m_{\rm W}^{\rm MLBZ}(P_{\rm b}) = P_{\rm b} \cdot \Delta m_{\rm W}^{\rm MLBZ}|_{\rm b} + (1 - P_{\rm b}) \cdot \Delta m_{\rm W}^{\rm MLBZ}|_{\rm udsc} \tag{12}$$

The dependence of $P_{\rm b}$ as a function of b-tag as known from MC simulation was used to fit this model to the difference between data and simulation of $b_{\rm rec}$ as a function of the b-tag (see figure 10) giving the following values for the heavy-quark and light-quark systematic MLBZ shifts:

$$\Delta m_{\rm W}^{\rm MLBZ}|_{\rm b} = -5.7 \pm 5.7 \ \ {\rm MeV/c^2} \tag{13}$$

and

$$\Delta m_{\rm W}^{\rm MLBZ}|_{\rm udsc} = -0.3 \pm 2.8 \ \ {\rm MeV/c^2} \tag{14}$$

The latter number is to be used as 'best MLBZ estimate' of the systematic bias on the W mass, even though the difference in systematic bias between light-quark and heavy-quark jets is not statistically significant.

4 Coverage and possible limitations

This section will start with some comments on the coverage of the MLBZ method as presented in table 2, then concentrate on possible limitations and finally propose a scheme for a complete treatment of the systematic errors.

4.1 Coverage of the MLBZ method

As listed in table 2 the MLBZ technique is expected to give relevant information about the modelling of jet fragmentation and detection. The method constitutes a stringent test on many aspects of the simulation of jets that may influence the W physics measurement. The basic idea is that imperfections in the WW simulation that bias our measurement will also be present in the $\rm Z^0$ simulation and thus give a measurable difference between MLBZ data and Monte Carlo.

As long as the emulation of the WW event topologies is reasonable, with a realistic coverage of phase space, this difference is believed to represent the actual systematic error to first order, automatically including

Fragmentation effects covered	'fully'	partly	not at all
Hard gluon radiation	X		
Soft gluon radiation	X		
Fragmentation functions	X		
(+ detector response)			
2-particle correlations inside W's/Z^0's	X		
(+ detector response)			
FSI between W's/Z^0's			X
Detector effects covered	**'fully'**	**partly**	**not at all**
Jet energy scale at 45 GeV	X		
Jet energy versus (θ,ϕ)		X	
Jet energy non-linearity			
below 45 GeV		X	
above 45 GeV			X
Jet direction syst. bias			
asymmetric		X	
back-to-back symmetric			X
Track density		X	
Most other effects		X	

Table 2: Summary of systematic effects covered by the MLBZ method.

- practically all fragmentation effects with their internal correlations and the bulk of 'known' detector systematics
- and possible 'unknown' additional systematic effects that are not covered by traditional error estimates and would otherwise have escaped attention.

The advantage of such an 'inclusive' approach is that with one well-defined measurement the combined systematic effect is determined with excellent precision. In addition to this inclusive measurement different contributions to the systematic error can be studied as a function of relevant event variables (as discussed in section 3.2), providing a highly sensitive test to disentangle more exclusive effects. This is important to

- improve our understanding of the different contributions that play a role
- to spot hypothetical large systematic discrepancies that accidentally cancel in the inclusive measurement but could render the result unstable for imperfections in the WW emulation by MLBZ events.

Systematic effects that are obviously NOT covered include the LEP beam energy calibration, the description of the ISR spectrum in WW events, Final State Interference (FSI) effects between particles from different W bosons and the description of background from non-WW physics processes. Those effects have to be taken into account separately.

4.2 Limitations

The MLBZ events contain the full information of any systematic bias in the description of the fragmentation, thanks to the fact that practically all processes involved are Lorentz invariant.

This is less true for the detector response for which, however, the main features are still expected to be the same in MLBZ and W^+W^- events except for higher order (non-linear) corrections.

To estimate how much a certain imperfection in the MLBZ description might affect the MLBZ measurement we will use the following approach: if there is a systematic deficiency in the detector description it will turn up in an (independent) data/MC comparison when this effect is larger than x_{local}. For effects smaller than x_{local} we have to rely on the MLBZ correction, so it has to be investigated to what precision p% of the effect the linear model is precise on the range up to $x = x_{\text{local}}$. This will then give a maximal possible contribution of $y = p \cdot x_{\text{sum}}$ MeV/c^2 to the W mass, where x_{sum} is the linear sum of the local deficiencies x_{local}, again known from independent study. The x, p and y have to be determined for the individual problems.

Detector-related limitations:

- *Back-to-back detector holes*
  The Z^0 events are produced back-to-back unlike the jets from W pair production. This could lead to correlations in MLBZ events not present in WW decays. This effect was studied by applying possible deteriorations to the simulation of the Z^0 events before and after they have been mixed and Lorentz boosted. We have put a hole in $|\cos(\theta)|$ varying the hole from 0.795-0.800 (realistic) to 0.700-0.800 (unrealistic). Applying the hole before Lorentz boosting corresponds to a discrepancy in the description of the detector during the Z^0 data taking, while applying the hole after Lorentz boosting corresponds to the effect of this additional hole on the kinematical reconstruction of WW events. As shown in figure 11 the estimated systematic shift agrees better than $p = 15\%$ up to an additional artificial hole of 0.02 in $\cos\theta$ for the two cases. The overall acceptance of DELPHI is known from independent studies to an accuracy of $x_{\text{sum}} = 0.01$ (conservative), giving a maximal effect of these correlations of the order of $y = 0.9$ MeV/c^2.

- Another consequence of the back-to-back topology of the Z^0 events is that the MLBZ result is highly insensitive to certain systematic biases in the reconstruction of the jet direction only rotating the thrust axis direction (e.g. always 1 degree away from the beampipe). This kind of biases on the direction of the jets has to be investigated separately.

- *Track density and detector occupancy*
  The tracking efficiency depends on the track density. This effect is covered by MLBZ events, but only for particles originating from the same vector boson. It was seen by analysing the bias as a function of the thrust of the individual Z^0 events used in the MLBZ events (see previous chapter) that the bias depends highly on this quantity since events with gluon radiation have a low thrust value and give a much larger negative bias than events where the jets are slim. This was shown to be correctly described by the simulation. To study the effect of increased particle density for a 4-jet event with boosted jets, an additional particle inefficiency was artificially introduced as a function of the track density. This was studied on 1998 Z^0 simulation. Tracks that are close to each other were discarded according to a Gaussian with width of 5 mrad in $R\phi$, leading to a total loss of particles up to

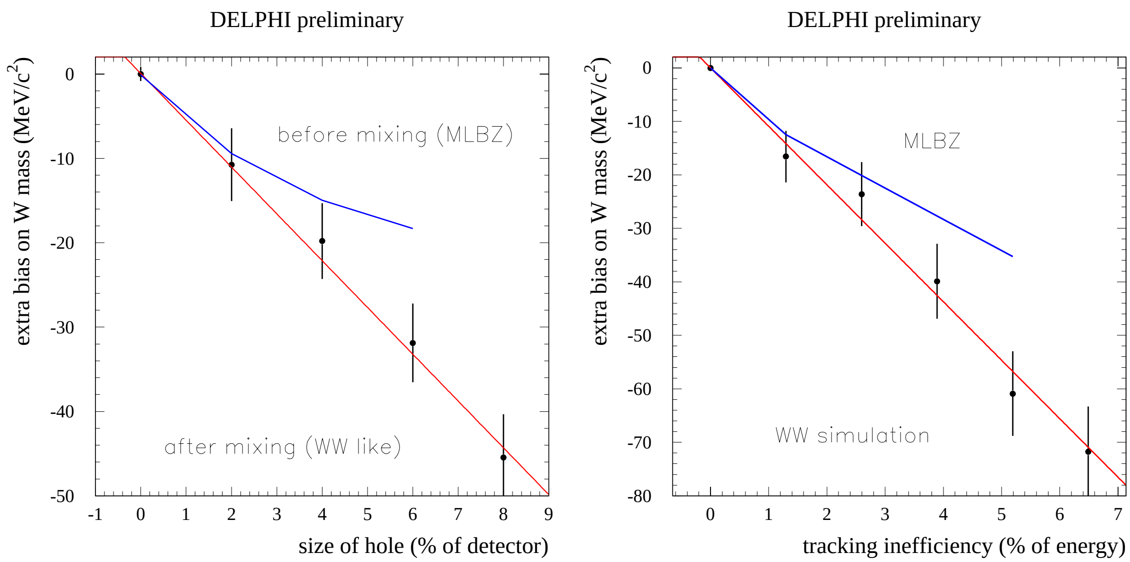

Figure 11: The left plot shows the effect of an additional back-to-back symmetric hole in the detector, applying the hole before and after Lorentz boosting. The right plot compares the extra bias induced in MLBZ events and WW events from standard simulation introducing an additional track-density dependent tracking inefficiency.

7%. This change is applied both in MLBZ events before boosting, and in fully simulated WW events as shown in figure 13. A linearity agreement of $p \approx 20\%$ can be derived, giving a maximum additional systematic of $y = 2.4$ MeV/c^2 for a track density discrepancy of $x_{\mathrm{sum}} = 1.0\%$ (conservative).

- The energy spectrum of jets in W^+W^- events is quite different from Z^0 decays. Any non-linearity in the energy response could become a systematic error in m_W. This effect is partly covered for jet energies below 45 GeV in 3-jet (low-thrust) Z^0 events but not at all for jet energies above 45 GeV. This requires separate study.

- The distribution and correlation of jet directions are not completely identical. As a cross-check the bias was estimated as a function of the polar angle of the Z^0 thrust (see section 3.2) with excellent agreement between simulation and data.

- A final detector-related point is the time dependency of the bias. By definition the data used for the MLBZ measurement is taken during the Z^0 calibration runs outside the high energy data taking periods. Therefore the conclusions about detector performance and description have to be extrapolated and/or interpolated in time. In 1997 on-pole Z^0 events were only recorded at the start of data taking while in 1998 Z^0 data were taken both at the beginning and towards the end of the year. In order to illustrate both the short term and the long term stability 1998 data was analysed and the MLBZ mass plotted as a function of time in figure 12. The measured overall mass differences between data and Monte Carlo are listed in table 3. From these numbers and the fact that the stability plot in figure 12 is compatible with a fully stable detector, a preliminary estimate for this effect is 5 MeV/c^2. By studying the

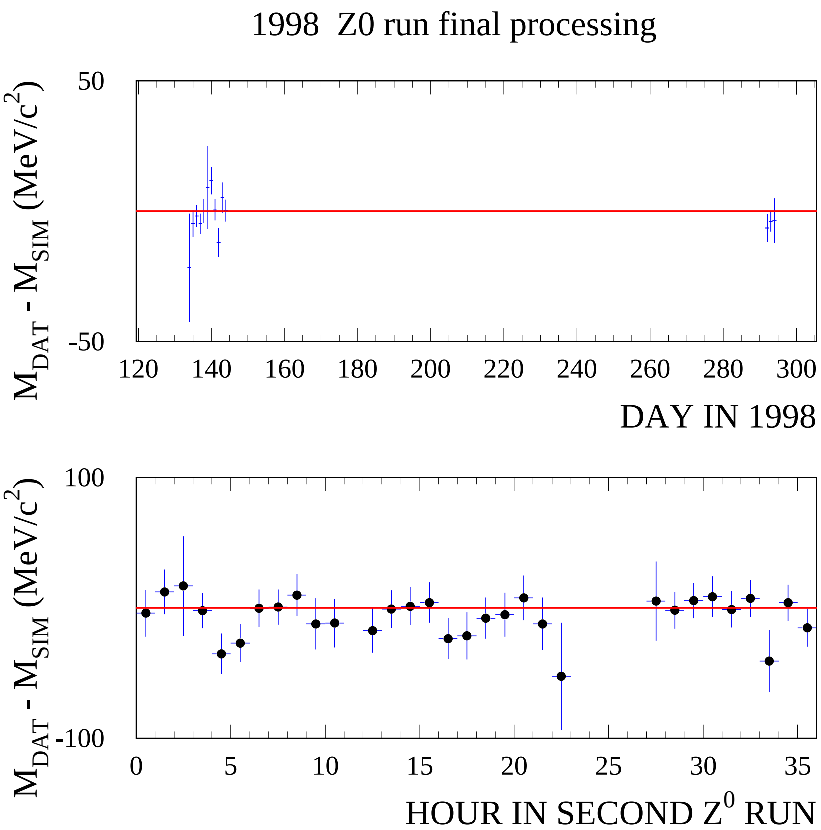

Figure 12: Experimental stability of the systematic difference in W mass bias between data and simulation for 1998 Z^0 data.

detector stability over the years taking into account also the calibration data of 1999 and with additional studies (e.g. using Z^0s from radiative returns during the high energy data taking) one can probably draw more firm conclusions.

Fragmentation-related limitations:

- The flavour composition is different in Z^0 and W decays. A non-perfect simulation of the b decays would then lead to wrong estimation of the bias. As a function of the event b-tag variable the simulation describes the data perfectly and an upper bound of 1 MeV/c^2 is estimated.
- The fragmentation of Z^0 events happens at a scale 13% larger than W^+W^- decays. This has negligible impact on the bias since the simulation is adequately able to describe this energy evolution.
- When boosting the particles in the MLBZ procedure the particle masses are not precisely known. The approximation used (no mass for neutral particles, pion mass for charged particles) was compared to the all-photons and all-kaons hypothesis, giving a maximal effect of xx MeV on the W mass.

Z^0 data taking period	$\Delta m_{\mathrm{W}}^{\mathrm{MLBZ}}\|_{\mathrm{udsc}}$ (MeV/c^2)
1997	-0.3 ± 2.8
1998 P1	-1.0 ± 1.5
1998 P2	-4.7 ± 2.9

Table 3: Measured systematic mass difference between data and Monte Carlo simulation for the different Z^0 data taking periods in 1997 and 1998.

General limitations:

- *No-width approximation (reducible)*
  The MLBZ events produced in the analysis reported here have no width of the boson masses. This is a 'reducible' limitation, as it can be solved by using a more complete algorithm (see section 5.2). By using only simulated WW decays where the two masses are within 2 GeV/c^2 of m_{W}, it was verified that the bias from undescribed (tracking) inefficiencies of the detector was the same for WW event samples with the nominal W width and samples with the low width within the statistical error (see figure 13). With a precision p equal to 15% over a large range, this approximation leads to an estimated systematic error of 1.8 MeV/c^2 for a conservative $x_{\mathrm{sum}} = 1.0\%$.

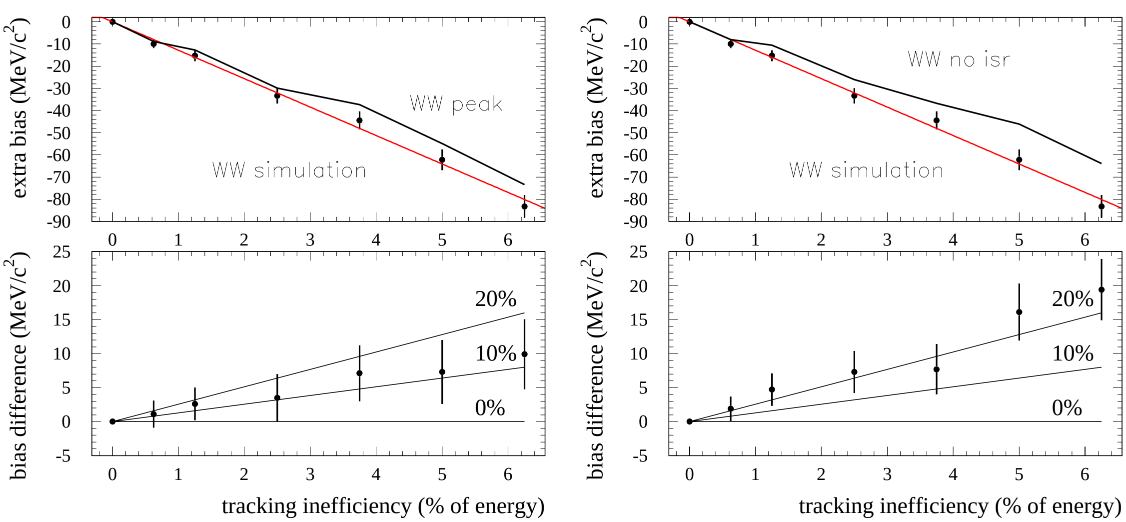

Figure 13: The effect of an additional track-density dependent tracking inefficiency is shown for WW events from standard simulation compared to WW events with both generated masses less than 2 GeV/c^2 away from m_{W}(left) and events without ISR radiation (right). The different points are highly correlated within each plot.

- *No-ISR approximation (reducible)*
  Unlike WW events MLBZ events contain hardly any ISR. Like the previous point

this problem can be solved using the algorithm described in section 5.2. It was verified that WW events without ISR responded in the same manner as the full sample to additional tracking inefficiencies, requiring a generated effective centre-of-mass energy less than 0.5 GeV away from 2 times the LEP beam energy. As shown in figure 13 the deviation y for a discrepancy $x_{\mathrm{sum}} = 1.0\%$ and a precision $p = 20\%$ is again of the order of 2 MeV/c^2.

- The true LEP beam energy during Z^0 data taking is not exactly the same in simulation as in data. The only way this difference enters in the MLBZ measurement is through the measured Z^0 energy. As the energy uncertainty (10-20%) is very large compared to the fluctuations and the calibration uncertainty on the beam energy, this influence is hardly significant. By varying the assumed LEP beam energy (E_{bZ^0}) in the MLBZ analysis we have found that this effect amounts to 0.5 ± 0.2 MeV/c^2 using a conservative uncertainty of 100 MeV on the average LEP beam energy during the Z^0 calibration runs.

4.3 Towards a complete estimate of the systematic error

From this study and other studies that are ongoing at the moment it can be concluded that the MLBZ method *can* be used to measure the combined systematic error from jet fragmentation and detector description, provided that the following additional effects are taken into account separately:

- non-linearity in jet energy response
- systematic bias in the jet directions
- stability in time

It is recommended to use the full 4-fermion MLBZ emulation (see section 5.2) if CPU time allows that. Otherwise the effect of neglecting ISR and the W decay width have to be re-investigated for other analyses. When there is reason to believe that another detector or analysis will have a significantly different response to the back-to-back correlations present in Z^0 events, the cross-checks as presented in this paper have to be repeated.

5 Outlook

5.1 Semileptonic channel

MLBZs can also be applied to the semileptonic channel [6]. There the jets can be taken from hadronic Z^0 events, while the lepton can either be constructed by taking half a leptonic Z^0, or can be generated artificially, as the only information in the lepton is its energy and direction. By using artificially generated leptons one can statistically optimise the study of the systematics of the hadronic part, and later fold in the knowledge of the full energy and momentum spectrum of the lepton.

For the semileptonic channel systematic shifts from the reconstruction of the hadronic part of the event are expected to be larger than in the fully hadronic channel, because

[6] The method can also be of use for other 4 fermion final states like ZZ or HZ.

mass shifts are less strictly controlled by the constraints in the constrained fit. The final statistical sensitivity should be of the same order of magnitude, as it depends purely on the number of hadronic Z^0 events available and the resolution of the boost in the Z^0 events.

5.2 Improved emulation

The method described in this paper can be extended in order to improve the emulation of WW events. In DELPHI an algorithm is being developed (by Chris Parkes [11]) in which the Z^0 events can be rotated, rescaled and boosted reproducing the 4-fermions taken from WW simulation, thus including the effects of Initial State Radiation, a Breit-Wigner with the correct Γ_W, helicity structure and proper energy scale in the MLBZ topology [7]. This will emulate WW events closer to the truth, thus further improving the reliability of the measured systematic shift. The only drawback is that more mixings and boosts will be needed to obtain the same statistical precision because of the Breit-Wigner mass distribution and ISR.

The greatest virtue of such an improved emulation is that it will allow the use of analyses that are not easily scalable with energy, are sensitive to the helicity structure (TGC analyses) or e.g. contain neural networks that have been trained with WW simulation.

This was not needed for the DELPHI analysis used in this paper as it has a very simple cut based selection and kinematic fits that are very well scalable with energy.

6 Conclusion

A new technique to measure systematic errors from jet reconstruction in W physics measurements at LEP2 has been presented: the MLBZ method. A full description of the method itself and the procedure to determine the statistical precision has been given.

Some results relating to the W mass measurement on 1997 DELPHI data in the fully hadronic channel were shown and discussed. They give a consistent picture and show that the combined systematic error on jet fragmentation and a large fraction of detector effects in the fully hadronic channel is around 5 MeV/c^2, which is a factor 4 smaller than so far quoted by DELPHI [3] for fragmentation only (20 MeV/c^2). For a precise estimate of the jet energy response error more work is needed, as outlined in this paper.

The coverage and limitations of the MLBZ method were discussed, leading to the conclusion that it can be used for a complete estimate of the systematic error due to jet reconstruction effects, provided the method is complemented with separate studies of non-linearity in the jet response, back-to-back symmetric systematic biases in reconstructed jet directions, and a sound estimate of time stability of the detector.

One of the main worries, being the back-to-back correlation of Z^0 event topologies with the detector symmetry, was shown to have a negligible influence for realistic uncertainties on the DELPHI detector simulation.

Thus the method seems to be a very promising candidate to replace a number of existing methods thanks to its better precision, greater coverage and ease of use and definition.

[7]The package already includes the possibility to emulate ZZ and semileptonic WW final states.

Our understanding of the MLBZ method and its merits is rapidly improving, and will certainly benefit when it is further tested, used and improved in other analyses, including W physics analyses other than the W mass measurements. In particular it would be interesting to compare results obtained for different LEP experiments.

Acknowledgement

We would like to thank Ahmimed Ouraou, Chris Parkes, Ivo van Vulpen, Jan Timmermans, Louis Lyons and Peter Renton for their help and many valuable suggestions. This work would not have been possible without the data and infrastructure provided by the DELPHI collaboration.

References

[1] LEP EW working group note, LEPEWWG/WW/99-01, revised 15/04/99.

[2] ALEPH Collaboration, Phys. Lett. **B453**(1999) 121.

[3] DELPHI Collaboration, Phys. Lett. **B462**(1999) 410.

[4] L3 Collaboration, Phys. Lett. **B454**(1999) 386.

[5] OPAL Collaboration, Phys. Lett. **B453**(1999) 138.

[6] DELPHI Collaboration, Eur. Phys. J. **C2**(1998) 581.

[7] B. Efron, Computers and the Theory of Statistics, SIAM Rev. **21**(1979) 460.

[8] P. Diaconis and B. Efron, "Computer-Intensive Methods in Statistics", Scientific American **248:5**(1983), 96.

[9] B. Efron, R.J. Tibshirani, "An Introduction to the Bootstrap", Chapman & Hall 1993. ISBN 0-412-04231-2

[10] A.C. Davison, D.V. Hinkley, "Bootstrap Methods and Their Application", Cambridge University Press, 1997. Hardback ISBN 0-521-57391-2 Paperback ISBN 0-521-57471-4

[11] A manual of the program can be found at `file:/afs/cern.ch/delphi/tasks/wphysics/mlbz/mlbz_manual/mlbz_manual.html` and more information can be obtained from Chris Parkes (Chris.Parkes@cern.ch).

A The Jackknife method

The statistical errors on the measured MLBZ masses were estimated using the 'Jackknife' resampling method [7, 8, 9, 10]. It is a non-parametric statistical method, i.e. a technique that can be used to estimate a statistical quantity without assuming knowledge about the underlying probability distribution. These methods have become more and more popular with the advance of modern computers. A well-known example of a rather advanced (and CPU time consuming) non-parametric technique is the Monte Carlo technique.

The Jackknife is a method for estimating the bias and standard error of an estimate. Here we are interested in the standard error. For a sample $\mathbf{x} = (x_1, x_2, ...x_n)$ and an estimator $\hat{\theta} = f(\mathbf{x})$, the method focuses on samples that *leave out one observation at a time*:

$$\mathbf{x}_{(i)} = (x_1, x_2, ..., x_{i-1}, x_{i+1}, ..., x_n)$$

for $i = 1, 2, ..., n$. These samples are called Jackknife samples. The ith Jackknife sample consists of the data set where the ith observation is removed. Now $\hat{\theta}_{(i)} = f(\mathbf{x}_{(i)})$, is called the ith Jackknife replication of θ, with a replication average

$$\hat{\theta}_{(\cdot)} = \sum_{i=1,n} \hat{\theta}_{(i)}/n$$

The Jackknife estimate of the standard error on $\hat{\theta}$ is then defined as

$$\hat{s}_{Jack} = \sqrt{\frac{n-1}{n} \sum_{i=1,n} (\hat{\theta}_{(i)} - \hat{\theta}_{(\cdot)})^2} \tag{15}$$

and is known to be a reliable estimator of the standard deviation provided that the distribution of $\hat{\theta}_{(i)}$ is smooth (in our case it turns out to be a Gaussian distribution). For non-smooth statistics, other methods like the Bootstrap [9, 10] can be used. In the central limit theorem the Jackknife error is shown to be equal to the Minimum Variance Bound for uncorrelated measurements.

The main building block for the technical implementation of the Jackknife method in the MLBZ analysis is the summed log-likelihood curve

$$\mathrm{L}_{j,k}^{\mathrm{mlbz}}(m_{\mathrm{W}}) = \sum_{m=1,n_{\mathrm{boost}}} \mathrm{L}_{j,k,m}^{\mathrm{mlbz}}(m_{\mathrm{W}})$$

where

$$\mathrm{L}_{j,k,m}^{\mathrm{mlbz}}(m_{\mathrm{W}}) = -2 \cdot \log(\mathcal{L}_{j,k,m}^{\mathrm{mlbz}}(m_{\mathrm{W}}))$$

with the event likelihood curve $\mathcal{L}_{j,k,m}^{\mathrm{mlbz}}(m_{\mathrm{W}})$ calculated for the MLBZ event consisting of Z^0 event no. j mixed with Z^0 event no. k using random boost no. m. In order to reduce the number of possible combinations to a managable level the data is divided in independent samples of $n = 201$ Z^0 events, and each event is mixed and Lorentz boosted $n_{\mathrm{boost}} = 2$ times with all other events from the same sample.

Taking the sum over all MLBZ events that contain event j, one obtains the Z^0 event likelihood curve

$$\mathrm{L}_{j}^{\mathrm{Z}^0}(m_{\mathrm{W}}) = \sum_{k=1,j-1} \mathrm{L}_{k,j}^{\mathrm{mlbz}}(m_{\mathrm{W}}) + \sum_{k=j+1,n} \mathrm{L}_{j,k}^{\mathrm{mlbz}}(m_{\mathrm{W}})$$

and combining all MLBZ events from the whole samples gives the overall sample likelihood curve

$$\mathrm{L}^{\mathrm{sample}}(m_{\mathrm{W}}) = \sum_{j=1,n} \sum_{k=j+1,n} \mathrm{L}^{\mathrm{mlbz}}_{j,k}(m_{\mathrm{W}}) = \sum_{j=1,n} \mathrm{L}^{\mathrm{Z}^0}_{j}(m_{\mathrm{W}}) \cdot \frac{1}{2}$$

For each sample the mass can be extracted by finding the minimum of the overall sample likelihood curve.

To calculate for each sample the Jackknife estimate of the statistical error on the fitted mass m_{fit}, we need the Jackknife replications $m_{\mathrm{fit(i)}}$ obtained by minimising $\mathrm{L}^{\mathrm{sample}}(m_{\mathrm{W}})_{(i)}$ given by

$$\mathrm{L}^{\mathrm{sample}}(m_{\mathrm{W}})_{(i)} = \mathrm{L}^{\mathrm{sample}}(m_{\mathrm{W}}) - \mathrm{L}^{\mathrm{Z}^0}_{i}(m_{\mathrm{W}})$$

thus totally removing event no. i from the sample.

As the shape of the summed log-likelihood curve $\mathrm{L}^{\mathrm{sample}}(m_{\mathrm{W}})$ around the minimum is very close to a parabola, the influence of event i on the fitted mass $m_{\mathrm{fit(i)}} - m_{\mathrm{fit}}$ can be approximated by

$$\Delta_i = \left. \frac{\frac{\delta \mathrm{L}^{\mathrm{Z}^0}_{i}(m_W)}{\delta m_{\mathrm{W}}}}{\frac{\delta^2 \mathrm{L}^{\mathrm{sample}}(m_W)}{\delta m_W^2}} \right|_{m_W = m_{\mathrm{fit}}} = \left. \frac{\delta \mathrm{L}^{\mathrm{Z}^0}_{i}(m_W)}{\delta m_{\mathrm{W}}} \right|_{m_W = m_{\mathrm{fit}}} \cdot \frac{1}{2} \sigma^2_{m_{\mathrm{fit}}} \tag{16}$$

where $\sigma_{m_{\mathrm{fit}}}$ is the standard likelihood error on the fitted sample mass. The average Jackknife replication $m_{\mathrm{fit}(\cdot)}$ is then given by

$$m_{\mathrm{fit}(\cdot)} = \sum_{i=1,n} m_{\mathrm{fit(i)}}/n = m_{\mathrm{fit}} + \left. \frac{\sum_{i=1,n} \frac{\delta \mathrm{L}^{\mathrm{Z}^0}_{i}(m_W)}{\delta m_W}}{n \cdot \frac{\delta^2 \mathrm{L}^{\mathrm{sample}}}{\delta m_W^2}} \right|_{m_W = m_{\mathrm{fit}}} = m_{\mathrm{fit}}$$

as

$$\sum_{i=1,n} \mathrm{L}^{\mathrm{Z}^0}_{i}(m_W) = 2 \cdot \mathrm{L}^{\mathrm{sample}}(m_W)$$

giving

$$\sum_{i=1,n} \left. \frac{\delta \mathrm{L}^{\mathrm{Z}^0}_{i}(m_W)}{\delta m_W} \right|_{m_W = m_{\mathrm{fit}}} = 2 \cdot \left. \frac{\delta \mathrm{L}^{\mathrm{sample}}(m_W)}{\delta m_W} \right|_{m_W = m_{\mathrm{fit}}} = 0$$

by definition. This means that we can substitute the *Jackknife influence value* $\hat{\theta}_{(i)} - \hat{\theta}_{(\cdot)}$ in equation 15 by Δ_i and use the following very simple formula

$$\sigma_{\mathrm{MLBZ}} = \sqrt{\sum_{i=1,n} \Delta_i^2} = \sqrt{\frac{n}{n-1}} \cdot \hat{s}_{\mathrm{Jack}}$$

as an excellent approximation of the Jackknife estimate 15 of the standard error on the fitted MLBZ mass for reasonably large values of n.

The same procedure applied to the whole data set of n_{samp} independent samples l used should give consistent results:

$$\sigma^{\mathrm{MLBZ}}_{all} = \sqrt{\sum_{l=1,n_{\mathrm{samp}}} \sum_{i=1,n} \Delta^2_{i,l}} \approx \frac{\langle \sigma^{\mathrm{MLBZ}}_{l} \rangle}{\sqrt{n_{\mathrm{samp}}}}$$

where $\Delta_{i,l}$ is calculated according to equation 16 replacing the sample error $\sigma_{m_{\mathrm{fit}}}$ by the overall likelihood error on the fitted mass when combining all samples.

In table 1 a comparison of the different error estimates is given. It shows that for the number of events used the Jackknife method estimates the statistical error correctly to within 10%.

B Linearity

Similarly to the Z^0 event influence Δ_i (16) one can define the MLBZ event influence Δ_{ij} as the change in the fitted sample mass when removing the MLBZ events containing the mixed pair of Z^0 events i and j:

$$\Delta_{ij} = \left.\frac{\delta \mathrm{L}^{mlbz}_{i,j}(m_W)}{\delta m_{\mathrm{W}}}\right|_{m_W = m_{\mathrm{fit}}} \cdot \frac{1}{2}\sigma^2_{m_{\mathrm{fit}}} \tag{17}$$

In figure 14 the average Δ_{ij} is plotted as a function of the sum of the individual boson influences Δ_i and Δ_j of the constituent Z^0 events. The dependence is quite linear over the whole range, which means that linearity is conserved during the whole procedure of mixing and boosting the bosons, jet clustering, applying a constrained fit for each jet pairing and extracting the mass.

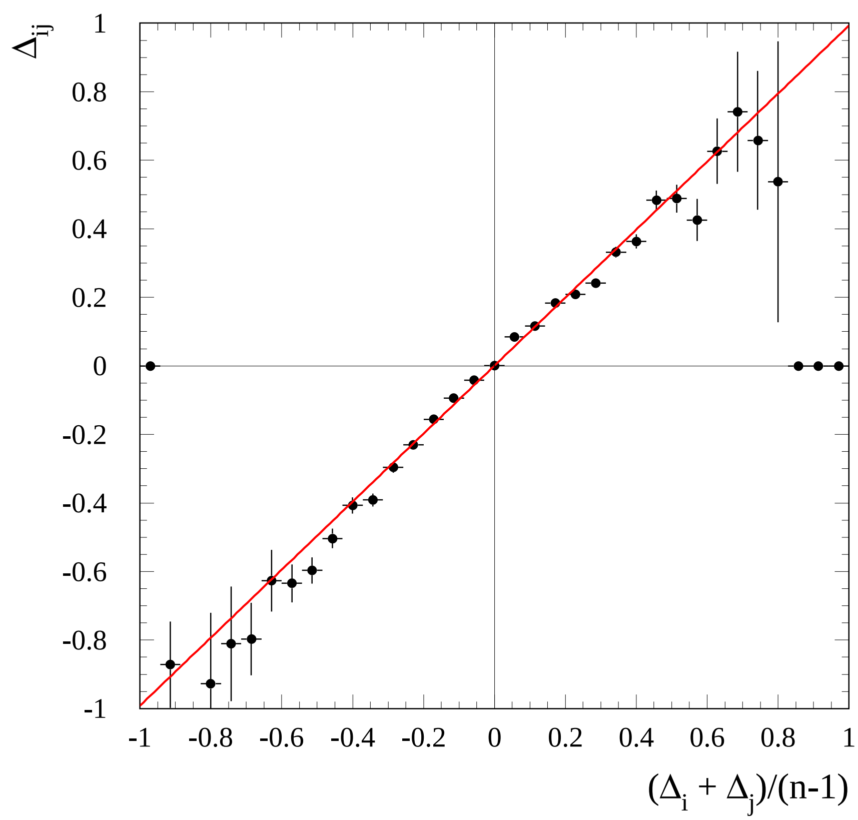

Figure 14: The average MLBZ event influence as a function of the sum of the individual Z^0 event influences, in units of $\frac{1}{2}\sigma^2_{m_{\mathrm{fit}}}$.