Direct measurement of the W boson mass in $e^+ e^-$ collisions at LEP
Mulders, M.P.

Citation for published version (APA):
Mulders, M. P. (2001). Direct measurement of the W boson mass in $e^+ e^-$ collisions at LEP

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Chapter 5

Historical account

The aim of this chapter is to give a historical overview of the developments and key new ideas that led to the realization of the main analysis presented in this thesis. Starting from the ‘Yellow Book approach’ (section 5.1) several new ideas are introduced including a 5-jet treatment (section 5.2) and the Ideogram technique with its development from a 1D convolution in the 172 GeV analysis to a fast 2D convolution at 189 GeV (section 5.3). In the last section (5.6) a few words are spent on the role that Jackknife and MLBZs have played in the studies of systematic errors. A comprehensive description of the main analysis is given in the next chapter.

5.1 Yellow Book approach

The different methods for the direct measurement of the W mass as proposed at the 1995 ‘Physics at LEP2’ workshop [67] were all based on the following approach:

1. Event selection and jet clustering in 4 jets (q̅q̅q̅q̅ channel) or 2 jets + lepton (q̅qlν channel)
2. Full kinematic event reconstruction extracting 1 or 2 masses per event
3. A fit to the obtained mass spectrum, using one of the following four methods proposed:

   (a) fitting the mass spectrum with a ‘simple’ analytical function, followed by calibration using Monte Carlo simulation.

   (b) the convolution method, where the underlying physics function (or differential cross-section \( \frac{d^2\sigma(s;m_W,\Gamma_W)}{dm_1dm_2} \)) is used as a fitting function, taking into account the effects of the detector by convolution. The prediction of the reconstructed invariant masses \((\bar{m}_1, \bar{m}_2)\) is thus given by:

\[
\frac{d^2\sigma(s;m_W,\Gamma_W)}{dm_1dm_2} = \int dm_1 \int dm_2 G(s;\bar{m}_1,\bar{m}_2,m_1,m_2) \cdot \frac{d^2\sigma(s;m_W,\Gamma_W)}{dm_1dm_2} \tag{5.1}
\]

The yellow report does not give any guidelines of how to choose the Green’s function \(G(s;\bar{m}_1,\bar{m}_2,m_1,m_2)\), which is obviously the most involved part of this method. This approach in fact is a special case of method (a), with a slightly more restricted and therefore less arbitrary choice of the fitting function. Also in this method a calibration using Monte Carlo simulation is needed.
(c) **Monte Carlo interpolation.** Monte Carlo samples with different values of $m_{W}^{MC}$ and $\Gamma_{W}^{MC}$ are generated and processed with the same event selection and kinematic fit as the data. Then for each of the samples the compatibility of the invariant mass spectrum with the mass spectrum obtained with real data is determined (e.g. using a binned maximum likelihood fit). Interpolation of the likelihood (or $\chi^2$) in the $(m_{W}^{MC}, \Gamma_{W}^{MC})$ grid then gives the fitted value of $m_{W}$ or $\Gamma_{W}$ automatically correcting for all possible biases due to mass reconstruction and experimental cuts, provided they are described by the Monte Carlo.

(d) **Monte Carlo reweighting.** Same as (c), but using a Monte Carlo reweighting technique as described in section 4.3 to produce the samples with different values of $m_{W}^{MC}$ and $\Gamma_{W}^{MC}$ from a single (or just a few) samples of generated Monte Carlo events, which is more efficient and more flexible.

**Example of a Yellow Book analysis**

For illustrative purposes a simple implementation of such an analysis, based on method (d) and only including the $q\bar{q}q\bar{q}$ channel, is described in the following. This ‘reference’ Yellow Book analysis consists of:

- **Event selection:**
  Identical to the $q\bar{q}q\bar{q}$ selection of the main analysis described in chapter 6. The DURHAM jet clustering algorithm is used to cluster these events in 4 jets.

- **Full kinematic event reconstruction:**
  A 5C equal mass constrained fit is performed as described for the main analysis. Out of 3 possible jet pairings the one with lowest $\chi^2$ is chosen. Thus one mass per event is extracted and plotted as in Figure 5.1.

- **Fit to the obtained mass spectrum using Monte Carlo reweighting and a binned maximum likelihood fit, based on Poissonian statistics in each bin.** The Monte Carlo events are reweighted changing the W mass in steps of 0.1 GeV/c$^2$ and the likelihood given by

$$
\Delta \chi^2 = 2 \sum_{j} \left[ N_{j}^{MC} - N_{j}^{data} + N_{j}^{data} \cdot \ln\left( N_{j}^{data} / N_{j}^{MC} \right) \right]
$$

(5.2)

is calculated and plotted (see Figure 5.1), where $N_{j}^{data}$ and $N_{j}^{MC}$ are the number of selected data events and Monte Carlo events respectively in bin no.$j$. From the likelihood curve the mass and error on the mass are derived in the usual way by fitting a parabola as shown in Figure 5.1.

To good approximation this method is unbiased, since all known reconstruction biases (e.g. due to experimental cuts and detector efficiencies and resolution) are automatically corrected for by the Monte Carlo simulation.

This analysis is only a simplified approach, meant to quantify some of the effects discussed in the next sections.
**Figure 5.1:** Fully-hadronic mass spectrum showing the invariant mass from the ‘best’ (= lowest $\chi^2_{5C}$) jet pairing in each event, before (top left) and after (top right) a 5C constrained fit, illustrating the improvement in resolution. The bottom plots show the corresponding likelihood curves obtained from a binned maximum likelihood fit of reweighted Monte Carlo to the 189 GeV DELPHI data shown in Figure 5.1. The errors quoted are the statistical errors obtained from the parabolic fit to the likelihood curves shown.

**Other analyses**

All $m_W$ measurements based on 172-189 GeV data published to date by the other LEP experiments [64, 65] are improved variations of the Yellow Book analysis, in the sense that they are all based on a Monte Carlo reweighting fit of a mass spectrum to the data. Technical improvements
were made to the Monte Carlo reweighting fits and the performance of the measurements was improved by optimising event selection, the development of better jet pairing algorithms, and attempts to further improve the constrained fit. A few examples of such developments are:

- **ALEPH**: 4C fit + energy scaling instead of 5C fit, extracting two masses per event instead of one
- **L3**: neural network for event selection, improved handling of the binning in the global fit, and a separate fit of the second best jet pairing
- **OPAL**: improved event selection and choice of jet pairing using multivariate likelihood discriminants, and (at 189 GeV) a separate treatment for 4-jet and 5-jet events

The development of analysis methods is still ongoing: In the recent WW physics at LEP2 Workshops [68] results were presented not only on the study of systematic uncertainties, but also on the continuing effort to further improve the mass extraction methods. OPAL has been working on a convolution method similar to the one presented in this thesis, and L3 reported about an investigation of a separate treatment for 5-jet events, also pioneered by the work presented here. A totally different approach is being pursued by ALEPH, based on a 3-dimensional reweighting technique in the semi-leptonic channels, where the main challenge is to control the danger of instabilities in the fit when the limited Monte Carlo statistics is spread out in 3 dimensions.

### 5.2 5-jet events... a first attempt

The first attempt back in 1996 to improve the statistical treatment of the events beyond the Yellow Book approach, was to treat 5-jet events as 5-jet events. In about 30% to 50% of the selected WW events (depending on the jet resolution variable $y_{\text{cut}}$) more than 4 jets are visible, due to the radiation of final state gluons with high $k_T$ (see section 2.5). Obviously, as was the case with the determination of $\sqrt{s'}$ (section 4.2, page 60), by acknowledging the apparent energy-flow structure including the ‘natural’ number of jets one should be able to extract more detailed and correct statistical information from the event.

But here the situation is more complex: In addition to the advantage of having a more correct model of the event, also the energyflow separation $\epsilon_{\text{sep}}$ will improve, provided that an effective algorithm can be designed to choose the correct jet pairing. As the number of possible jet pairings increases from 3 in a 4-jet event to 10 in a 5-jet event, it becomes more challenging to find the correct jet combination. An optimal treatment has to balance two pieces of information:

1. the distance of the closest jets (equivalent to the DURHAM $y_{\text{cut}}$ used to distinguish a 4-jet from a 5-jet event)

2. a jet pairing criterion; in our simple example just the pairing with the lowest $\chi^2_{5C}$ (i.e. smallest difference of the two measured boson masses).

The traditional approach applies these two measures sequentially: first measure no.1 is used to reduce the number of jets from 5 to 4, followed by criterion no.2. In our analysis we chose to use the $1/k_T$ dependence as discussed in 2.5 as a natural measure to estimate the relative probability for each jet pairing that the fifth jet was radiated with the observed $k_T$ with respect
Figure 5.2: Fully-hadronic mass spectrum (left) showing the fitted mass from the ‘best’ (see text) jet pairing in each event with 4-jet events treated as 4-jet and 5-jet events treated as 5-jet and the corresponding likelihood curve and fitted mass (right) using the Monte Carlo reweighting fit described in section 5.1.

to its corresponding parent quark jets. The $p \propto 1/k_T$ probability was transformed to a $\chi^2$ using $\chi^2_{k_T} = -2 \cdot \ln(p)$, and the jet pairing with the lowest overall $\chi^2$ was then further used in the analysis:

$$\chi^2_{\text{tot}} = \chi^2_{5C} + \chi^2_{k_T}$$

(5.3)

thus combining both pieces of information on an equal footing. This straight-forward 5-jet treatment resulted in a visible improvement of a few percent as Figure 5.2 compared to 5.1 shows.

This 5-jet study proved that a further refinement of the Yellow Book approach could lead to modest improvements in the statistical sensitivity. More important was the fact that it highlighted as main limitation of the analysis the representation of each WW event by just one fitted mass, inevitably forcing a trade-off between a more detailed and correct description of the event ambiguities and the resulting increased difficulty in making the right ’choice’. The solution, not to make a choice at all, emerged from an entirely different analysis approach: the Ideogram technique.

### 5.3 The Ideogram technique

The basic idea of the Ideogram technique is to abandon the analysis paradigm based on a lineshape fit of the global invariant mass spectrum, and change to event-by-event likelihoods describing the full ambiguity of the mass information in each event as correctly as possible. By taking into account the full mass ambiguity the limitation of ‘choosing’ the correct solution is avoided. The full information is carried on to the combined likelihood of the overall event sample, where
ambiguities that could not be resolved on the event level become easy to solve, as the signal clearly stands out w.r.t. the background of ‘wrong’ solutions.

**Statistics background**

Though originally conceived on the basis of sheer common sense, the Ideogram analysis technique can be derived directly from Bayesian inference principles, starting from very basic statistics rules, namely the sum rule:

\[
p(X|I) + p(X'|I) = 1 \tag{5.4}
\]

and the product rule:

\[
p(X, Y|I) = p(X|Y, I) \cdot p(Y|I) \tag{5.5}
\]

where \(X, Y\) and \(I\) are Boolean variables or propositions that can either be true or false and \(p(X|Y)\) signifies the probability that \(X\) is true, given that \(Y\) is true. The comma stands for the logical conjunction ‘AND’, while \(X\) is the negation of \(X\), i.e. the proposition that \(X\) is false. Related to the sum rule equation (5.4) is the principle of marginalisation, either used in discrete form:

\[
p(X|I) = \sum_I p(X, H_i|I), \tag{5.6}
\]

where the different propositions \(H_i\) should be mutually exclusive and form a complete set, or as a continuous integration:

\[
p(X|I) = \int_{-\infty}^{\infty} p(X, a|I) da \tag{5.7}
\]

where \(a\) is a so-called nuisance parameter, i.e. a parameter that is not of primary interest for the measurement. A result that follows directly from the basic equations (5.4) and (5.5) is Bayes’ Theorem:

\[
p(m_w|\text{event}, I) = \frac{p(\text{event}|m_w, I) \cdot p(m_w|I)}{p(\text{event}|I)} \tag{5.8}
\]

where \(m_w\) stands for the parameter to be measured, \(I\) encompasses all underlying assumptions, and \(\text{event}\) stands for the observed data. This theorem turns out to be very useful to describe the process of scientific inference. If one is interested in parameter estimation the normalisation constant \(p(\text{event}|I)\) which does not depend on the parameter to be measured can be omitted. Also the ‘Bayesian prior’ \(p(m_w|I)\) is often chosen to be flat, which is certainly a good choice for a statistically well-behaved precision measurement like the W mass analysis. Technically this reduces the procedure to a standard maximum likelihood approach:

\[
p(m_w|\text{event}, I) \propto p(\text{event}|m_w, I) \tag{5.9}
\]

where the proportionality sign indicates that whereas the posterior probability density function \(p(m_w|\text{event}, I)\) is normalised the relative likelihood function \(p(\text{event}|m_w, I)\) is not. In practice this does not matter. In order to determine the W mass (or width) and the statistical uncertainty it is sufficient to obtain a relative likelihood curve (or likelihood ratio) which is the product of all relative event likelihood curves \(p(\text{event}|m_w, I)\) that are calculated for each event as a function of \(m_w\) (or \(\Gamma_W\)). The absolute WW cross-section is kept fixed which is a good approximation in the \(m_w\) range of interest, to be cross-checked later by doing a full Monte Carlo calibration. In principle \(\text{event}\) includes the complete set of observations connected to the event \(\text{event}\), but in
HISTORICAL ACCOUNT

practice the likelihood can only be evaluated for a limited number of observables. Moreover, the only way to take into account complicated jet fragmentation and detector acceptance and resolution effects with sufficient precision is to rely on Monte Carlo simulation; either directly or indirectly, as a final calibration of the analysis.

Limitations of the ‘black box’ Monte Carlo reweighting approach

As was shown in the previous sections (5.1 - 5.2) a straightforward way to obtain the likelihood function \( p(event|m_W, I) \) is to represent the event event just by its 5C fitted mass \( m_{fit}^{5C} \) and estimate the probability distribution of \( m_{fit}^{5C} \) using Monte Carlo simulation directly:

\[
p(event|m_W, I) = p_{MC}(m_{fit}^{5C}|m_W, I). \tag{5.10}
\]

In fact the Monte Carlo histograms shown in Figure 5.1 and 5.2 give precisely this likelihood as a function of \( m_{fit}^{5C} \) for the corresponding analyses and a given W mass of 80.35 GeV/c². The Monte Carlo reweighting technique can be used to determine the best available estimate of the relative likelihood for other values of \( m_W \). In fact this is exactly (an event-by-event version of) the ‘Monte Carlo reweighting’ analysis described before in section 5.1.

The main drawback of this method is that including more than one observable (e.g. also the error \( \sigma_{m_{fit}^{5C}} \) on the fitted mass \( m_{fit}^{5C} \)) would require the available Monte Carlo events to be distributed in more than one dimension in observable space. For an increasing number of dimensions this quickly reaches the limits of the available statistics. In practice above 2 or 3 dimensions technical problems arise because statistical fluctuations in the coverage of the observable space can no longer be neglected. When these statistical limitations start playing a significant role this will lead in most cases to incorrect results and possibly to an underestimation of the statistical error if this is not taken into account.

Ideogram construction of the event likelihood

An analytically constructed likelihood does not have this limitation. In principle the Ideogram approach allows for the inclusion of all available event information taking into account all observables that are believed to be relevant. To include more event specific information the Ideogram method relies on a physics and statistics model to analytically evaluate expression (5.9) further and construct the event likelihood without using Monte Carlo simulation. This can be done to varying degrees of sophistication.

The first insight exploited in the Ideogram analysis is the fact that the event likelihood \( p(event|m_W, I) \) consists to a good approximation of two independent parts that can essentially be factorised (Figure 5.3). In statistical language, this factorisation is done by marginalisation. Using equation (5.7) and the product rule (5.5) the event likelihood \( p(event|m_W, I) \) can be written as:

\[
p(event|m_W, I) = \int \int p(event, \overline{m}'|m_W, I) \, d\overline{m}' \\
= \int \int p(event|\overline{m}', m_W, I) \cdot p(\overline{m}'|m_W, I) \, d\overline{m}' \\
\approx \int \int p(event|\overline{m}', I) \cdot p(\overline{m}'|m_W, I) \, d\overline{m}'. \tag{5.11}
\]
Figure 5.3: A schematic representation of the generation of a single WW event in two phases: the production of two W bosons with masses $m_1$ and $m_2$, followed by their decay, detection and analysis resulting in a fitted mass $m_{\text{fit}}^{5C}$.

Here the integration has to be performed over the whole physically allowed range of values of $m'$, with $m'$ representing the 'true invariant masses ($m_1$, $m_2$) of the two objects in the event that are supposed to be W bosons'. The last step in equation (5.11) reveals the motivation for separating these two factors: to a good approximation the $m_W$ mass dependence is concentrated in the physics function $p(m'|m_W, \Gamma_W)$ while the other factor $p(\text{event}|m', \Gamma_W)$ only depends on the kinematics observed in the event.

The (QED + EW) physics part of the process $p(m'|m_W, \Gamma_W)$ describing the production of two W bosons with masses $m_1$ and $m_2$, is well defined. The differential cross-section is contained in equation (2.4) in chapter 2, and can be written to a good approximation as the product of two running-width Breit-Wigner functions (2.5) and a two-dimensional phase space function:

$$p(m'|m_W, \Gamma_W) \approx S(m'|m_W, \Gamma_W, \sqrt{s}) = \text{BW}(m_1|m_W, \Gamma_W) \cdot \text{BW}(m_2|m_W, \Gamma_W) \cdot \frac{1}{s} \sqrt{(s - m_1^2 - m_2^2)^2 - 4m_1^2m_2^2} \quad (5.12)$$

The evaluation of the experimental resolution function $p(\text{event}|m', \Gamma_W)$ relies on approximations, where the aim is to describe the main features of the likelihood as correctly as possible. The resulting description of $p(\text{event}|m', \Gamma_W)$ is what is referred to as 'Ideogram' throughout this thesis.

One indispensable ingredient in the calculation of the experimental Ideogram is the constrained fit. The likelihood to observe the jets (and possible lepton) seen in the event for a given pair of true invariant masses $m'$ is estimated by the goodness-of-fit probability of a 6C kinematic fit:

$$p(\text{event}|m', \Gamma_W) = p_{\text{GC}}^{\text{fit}}(\text{jets (+ lepton)}|m', \sqrt{s}) \quad (5.13)$$

where the two boson masses given by $m'$ are fixed by the constraints.

As an example, for a $q\bar{q}l\nu$ event $\text{event}$ represented by the 4-momenta of the observed jets and lepton this would result in the following likelihood expression:

$$p(\text{event}|m_W, \Gamma_W) = \int p_{\text{GC}}^{\text{fit}}(\text{jets + lepton} | m', \sqrt{s}) \cdot S(m'|m_W, \Gamma_W, \sqrt{s}) \, dm'.$$  \quad (5.14)
The integral can be evaluated in one or two dimensions for $\overline{m}'$. A one-dimensional model assuming equal masses of the two W bosons works well for most events, and often the resulting 6C goodness-of-fit as a function of the equal mass $m'$ even turns out to be close to a Gaussian with a maximum for $m' = m'_{6C}$ and a width equal to the error $\sigma_{m'_{6C}}$ estimated by the constrained fit. For some $q\bar{q}\nu\nu$ events, however, this is not at all the case as will be discussed later in section 5.5.

In the $qqq\bar{q}$ channel the situation is slightly more complicated: there a jet pairing must be chosen to define to which combination of jets the mass constraints $\overline{m}'$ in the 6C kinematic fit apply. As discussed before the correct jet pairing can never be identified with certainty. This lack of knowledge is a key feature of of the $qqq\bar{q}$ resolution function $p(event|\overline{m}', I)$ and can be taken into account using discrete marginalisation (equation (5.6)) to sum over all possible jet pairing hypotheses $H^\text{pair}_j$, and then applying the product rule (equation (5.5)) as was done before:

$$p(event|\overline{m}', I) = \sum_{j=1}^{n_{\text{pair}}} p(event, H^\text{pair}_j|\overline{m}', I) =$$

$$= \sum_{j=1}^{n_{\text{pair}}} p(H^\text{pair}_j|\overline{m}', I) \cdot p(event|H^\text{pair}_j, \overline{m}', I) =$$

$$= \sum_{j=1}^{n_{\text{pair}}} p_j \cdot p_{6C}(event|\overline{m}', H^\text{pair}_j, \sqrt{s}) \quad (5.15)$$

where the probabilities $p_j = p(H^\text{pair}_j|\overline{m}', I)$ of the different jet pairings can be determined to different levels of sophistication, as will be discussed in more detail later.

![Illustration of Ideograms](image)

Figure 5.4: Some examples of Ideograms are shown, for different hypothetical 4-jet events. For all events the 3-fold jet-pairing ambiguity is the same, but the resulting knowledge about the true boson mass $m'$ is fundamentally different.

In order to illustrate the non-trivial effect of jet-pairing ambiguities on the mass information in an event, Figure 5.4 shows the Ideograms for 3 hypothetical 4-jet events, using an equal-mass Gaussian approximation. For simplicity in this example the sigmas are all taken to be equal and the weights given to each of the three jet pairings is chosen equal to $1/3$. In each event the correct
pairing has a solution close to 80 GeV/c², while the wrong jet pairings have maxima for more-or-less random values of the mass \( m' \). The effect of the jet pairing ambiguity on the mass is fundamentally different for each of the 3 cases:

1. If all pairings happen to give the same mass, the jet-pairing uncertainty does not matter at all. As a consequence this event contains the same information about \( m_W \) as an event without any ambiguity.

2. When the masses are rather close, the approximate value of the correct mass is known, but the ambiguity cannot be resolved and the net effect is a deterioration of the mass resolution for this event.

3. If, however, the masses are well separated, a broad range of possible masses should be considered for this event. By retaining this ambiguity in the event likelihood curve, however, this ambiguity will automatically become irrelevant when the addition of other events will unambiguously reveal the approximate location of the real \( W \) mass among the (almost) uniformly distributed background of wrong jet pairings. Effectively, for an event like this, the jet pairing ambiguity does not affect the mass resolution but it does reduce the weight (typically \( \propto 1/(n_{pair}) \)) of the signal peak w.r.t. to the background hypothesis (to be introduced later).

Thus, by describing the full ambiguity, the Ideogram method is able to take into account all possible jet pairings, and the fact that only one pairing per event can be correct. When more jet pairings are included, the weight per jet pairing is decreased accordingly, making sure that the integral over the signal probability stays normalised w.r.t. the background.

The extension from 3 jet pairings in a 4-jet event to 10 jet pairings in a 5-jet event is natural; the sum over 3 pairings is simply replaced by a sum over 10 pairings. To improve the analysis further, additional information can be used to determine the relative probabilities of the jet pairings \( p_j \). Ambiguities of a different nature can be included by adding more hypotheses using discrete marginalisation as before.

**Including the estimated event purity**

One additional ambiguity that is included in all Ideogram analyses is the question whether the event originated from a \( WW \) signal interaction, or some kind of background physics process. For background processes the expected mass distribution \( p(\vec{m}'|m_W,I) \) of the two identified heavy objects in the event is not given by a Breit-Wigner, but rather by some background shape \( B(\vec{m}'|I) \) which does not depend on \( m_W \) and can be extracted from Monte Carlo simulation. Applying discrete marginalisation, formula (5.11) can be expanded to:

\[
p(event|m_W,I) \approx \int \int p(event|\vec{m}',I) \cdot \left[ p_{event} \cdot S(\vec{m}'|m_W,I) + (1 - p_{event}) \cdot B(\vec{m}'|I) \right] d\vec{m}'
\]  

(5.16)

where the event purity \( p_{event} \) signifies the probability that the event was a \( WW \) signal event, estimated for each event using observables that are — as much as possible — uncorrelated with the mass information \( \vec{m}' \) (or \( m_W \)). In this manner the analysis takes into account the fact that 'pure'
events are likely to contain useful W mass information, while this is unlikely to be the case for background-like events. The Ideogram function $p(\text{event}|m', I)$ is evaluated as described before, where the number of hypotheses to be included depends on the analysis channel. The evolution of the Ideogram analysis along these lines will be discussed further in section 5.4 and 5.5.

**Comparison with the Convolution method**

The integration over the invariant masses introduced in equation (5.11) appears similar to the Yellow Book convolution (formula (5.1)), but this similarity is misleading. One difference is that the Ideogram integration is performed over the 'true' invariant masses in the event instead of the reconstructed invariant mass. Furthermore, the Ideogram idea is to calculate event-by-event likelihoods rather than constructing a function to be fitted to the overall mass spectrum. Finally, a characteristic and distinctive feature of the Ideogram method is the sum over different hypotheses.

**Resulting event likelihood curves**

The posterior likelihood curves are multi-modal, showing typically more than one maximum (A few examples are shown in Figures 5.8 and 6.15.). They are numerically stored in logarithmic form:

$$L_{\text{event}}(m_W) \equiv -2 \cdot \ln(p(m_W|\text{event}, I))$$

(5.17)

and kept for further analysis involving standard maximum likelihood techniques.

**Conclusion**

The Ideogram technique presented here introduced a new approach in the direct measurement of the W mass based on the analytical construction of event-by-event likelihood curves, allowing for the inclusion of more information specific to each event.

Such an event-by-event approach is advantageous because W events come in different qualities. Some beautiful 4-jet events give a clear clustering, obvious jet pairing and good mass resolution. Those events give excellent information about the W mass. Other events are more ambiguous and effectively contain less reliable or less precise mass information. By taking this large event-by-event variation into account as correctly as possible, the final uncertainty on the W mass can be reduced significantly.

**5.4 Evolution of fully-hadronic Ideograms in DELPHI**

**5.4.1 1D Ideograms at 172 GeV**

The W mass measurement based on the 172 GeV DELPHI data [1] was the first published application of the Ideogram analysis, introducing a 5-jet treatment and taking into account

- All possible jet pairings. A 1-dimensional convolution was used. Each jet pairing was represented by a Gaussian resolution function, the mean being equal to the mass obtained from a 5C equal mass fit, with a sigma equal to the estimated error from the constrained fit. In principle each of the jet pairings has equal a priori probability. However, in this
1-dimensional approach the compatibility of each jet pairing with the equal mass hypothesis had to be put in via additional weights: the relative probabilities of the different jet pairings were derived from the mass difference determined with a 4C fit (without equal mass constraint). The theoretically expected distribution for the mass difference of the two W bosons was used to obtain the relative probability for each jet pairing. Additionally, for 5-jet events, each jet pairing was weighted according to an additional relative probability \( 1/k_T \) as described in section 5.2.

- The event purity. The expected mass distribution for the background was approximated by a flat function times a 1-dimensional phase space function. For each event the purity \( \mathcal{P}_{\text{event}} \) was estimated as a function of a discriminating variable \( D \equiv \theta_{\text{min}} \cdot E_{\text{jet}} \), the product of the smallest angle between 2 jets and the minimum jet energy in a 4-jet configuration. The Ideogram was calculated as the sum of the WW and the background hypothesis as in equation (5.16).

**5-jet treatment**

In this analysis a \( y_{\text{cut}} \) value of 0.004 was used, giving 5 or more jets in approximately 30% of the selected W events.

**Adding jet broadness to PUFITC**

This was also the first time that the improved error-parameterisation with transverse jet-broadness errors was used, as described in the previous chapter (see section 4.2).

### 5.4.2 2D Ideograms at 183 GeV

With two W bosons to be fitted, it is more natural and more convenient to do a 2-dimensional convolution. Equation (5.11) can easily be interpreted as such, by letting \( \overline{m}' \) represent the combination of the two invariant masses in the event \( m_1 \) and \( m_2 \). For each combination of masses \( (m_1, m_2) \) a 6C constrained fit is done, and the \( \chi^2 \) from the fit used to derive a goodness-of-fit probability \( \propto \exp(-\Delta \chi^2/2) \) as a function of the two masses. The sum of the probability distributions for these jet pairings (Figure 5.5) is then convoluted with a 2D Breit-Wigner of the two W bosons. The advantages of the 2-dimensional approach are:

- No equal mass assumption is needed.
- No relative jet pairing weights have to be used. The solutions in which the two masses prefer to be close to each other will automatically obtain the largest weight by the convolution with a finite width Breit-Wigner.
- The background distributions tend to be more flat in 2-dimensional phase space, which makes it more correct to make the assumption that the background is flat (multiplied by a 2D phase space function), which simplifies the likelihood expressions.

Of course in 5-jet events with 10 jet pairings the \( 1/k_T \) additional weights are still used to take into account the gluon radiation probability.

Further improvements that were developed for the 183 GeV publication [2] are the following:
Figure 5.5: Example of 2D Ideograms as used to analyse the 183 GeV data [2] for a simulated 4-jet event (left) and a 5-jet event (right).

**Improved ISR treatment in the event selection**

In order to have a more accurate treatment of ISR in the event selection, especially for photons inside the detector acceptance, SPRIME, a standard DELPHI package, was adopted in the analysis to identify ISR photons inside the detector. Soon it was realised, however, that the kinematic treatment in this package was far from optimal for events with more than two jets (as discussed already in section 4.2 on page 60). Therefore a new algorithm was written, based on a constrained fit with the 'natural' number of jets. The new algorithm was published [5] and used in this analysis as part of the updated DELPHI SPRIME(+) package.

**Jet charge**

As W bosons are produced preferably in the same direction as the electron with the same electrical charge, and the difference in charge between the $W^+$ and the $W^-$ is 2e, there is some experimental information in the measured jet charges that can be used to improve the jet-pairing. This will be explained in more detail in chapter 6.

**Clustering ambiguity treatment**

In the previous chapter (section 4.1) it was shown that different jet clustering algorithms, though giving similar clustering performance on a whole sample of events can give strikingly different clustering results on an event-by-event basis for those events where clustering is ambiguous.

The Ideogram method is well equipped to identify this kind of ambiguity and take it into account in the statistical analysis. By repeating the ideogram construction for three different clustering algorithms (DURHAM, Cambridge and DICLUS), and simply adding the three ideograms
with equal weight, a combined ideogram is obtained that contains information about the clustering ambiguity of the event. If an event has clearly resolved jets, all three algorithms will find the same jets, and hence give identical ideograms. In this case the combined ideogram will be identical to the original ‘DURHAM’ ideogram. Alternatively, if the jet clustering is ambiguous, the ideograms produced using the three different algorithms may differ and lead to a combined ideogram in which the invariant mass information is smeared out. This takes into account more correctly the mass reconstruction ambiguity in such events.

Effectively this identification of ambiguous events will give more weight to unambiguous events, and on average improve the extraction of mass information from the overall event sample. This simple but effective idea led to an improvement of 4±1% on the W mass resolution.

5.4.3 Fast 2D Ideograms at 189 GeV

The final improvements to the 2D Ideogram analysis, accepted for publication [3] as analysis of DELPHI 189 GeV data and presented in full in chapter 6, were the following:

Improved ISR treatment in mass reconstruction

With increasing centre-of-mass energy the Initial State Radiation increases almost linearly with the available phase space ($\propto \sqrt{s} - 2m_W$). In most cases the ISR photon escapes undetected down the beampipe, which leads to a z-momentum imbalance and a reduction of the effective centre-of-mass energy, not taken into account in the standard constraints. At 189 GeV this has a significant effect on both resolution and bias in the reconstructed mass in about 20% of the events.

Several attempts to take into account ISR on an event-by-event basis in the analysis failed due to the non-Gaussian nature of the radiated ISR photon energy (see for example Figure 5.6). It turned out to be too difficult to use a first principle approach based on semi-analytical integration combining the a priori probability with the detector resolution effects that are Gaussian around the origin, but not in the tails where ISR starts to have a noticeable effect.

In the end a more pragmatic, Monte Carlo-based approach was found to be successful: for those events with fitted $p_T^z$ more than 1.5 $\sigma_p$, away from zero, ideograms were reconstructed both for the hypothesis that there was significant ISR, and that there was no significant ISR. The relative probabilities of the two hypotheses to be true were extracted from the Monte Carlo simulation as a function of $|p_T^z|/\sigma_p$, and the weighted sum of the two ideograms was used in the analysis.

At 189 GeV this ISR treatment was thus applied to 16% of the events, leading to a 15% improvement in W mass resolution for those events. The other events remained unaffected. The overall positive mass bias caused by ISR was reduced from 353 to 290 MeV/c² (at 189 GeV). The treatment of ISR will become more important with increasing $\sqrt{s}$. One of the advantages of this event-by-event approach is that the correction automatically increases when there are more events with collinear ISR radiation. Therefore it will slightly reduce possible ISR systematic effects as well.
Figure 5.6: Events with a fitted $p_{z}$ far away from zero are either badly reconstructed, or had an energetic ISR photon collinear with the beam. In these plots the fitted $p_{z}$ is compared to the generated $p_{z}$ (from 189 GeV EXCALIBUR simulation), for selected WW events with a fitted $|p_{z}|/\sigma_{p_{z}}$ between 1.5 and 2.5 (left), 2.5 and 3.5 (middle) and more than 3.5 (right). The fraction of events with true ISR (encircled with the dashed lines) increases for higher values of $|p_{z}|/\sigma_{p_{z}}$.

**Faster Ideograms**

Perhaps the most significant improvement introduced was a technical one: a small sacrifice in resolution led to a gain in speed of the analysis of a factor 10, by

- reverting to 4C instead of 6C ideograms, using the correlation between the masses as found by the 4C constrained fit, and assuming a 2D Gaussian resolution function for each jet pairing. This reduced the number of fits from $O(30)$ to one kinematic fit per jet pairing, clustering algorithm and ISR hypothesis.

- introducing a cut on the event purity at 25%. This reduced the number of events to be analysed by 23%, while affecting mass resolution and bias by only 1% and less than 1 MeV/c$^2$ respectively.

This improvement played a crucial role in speeding up the further development and testing of the analysis. Furthermore it enabled the analysis of systematic effects on millions of simulated events, necessary for an optimal estimation of the systematic errors.
Figure 5.7: The statistical mass resolution is shown as a function of the $y_{\text{cut}}$ going from 5 to 4 jets in 4 bins of approximately equal statistics, either when the events are treated as 4-jet (solid squares) or as 5-jet (open circles). The resolution was determined using 189 GeV WW simulation for Poissonian samples with average benchmark size of 100 events. The statistical uncertainty is indicated by the error band. As expected the resolution deteriorates for increasing values of $y_{\text{cut}}$. This deterioration can be reduced significantly, however, by treating 5-jet events as 5-jet. The plot also shows that for clear 4-jet events the 5-jet treatment performs almost as well as the 4-jet treatment, which proves that also there the $1/k_{\text{T}}$ weight works satisfactorily (creating a smooth transition from the 5-jet to the 4-jet regime).

Other final ‘adjustments’ at 189 GeV

- Improved $D$ variable
  The $D$ variable (to be defined later, on page 90), used to discriminate the 4-fermion signal from 2-fermion background, was slightly improved by not only taking into account the lowest jet energy and smallest inter-jet angle, but also the second-smallest energy and angle, thus probing both the difference between 4-jet and 3-jet and the distinction between 3-jet and 2-jet topologies. This improved the optimum product of selection efficiency and purity by about 3% without mass bias.

- Smaller value of $y_{\text{cut}}$
  In preparations for the 1999 WW Crete Workshop — in response to inquiries made by OPAL — it was found that an additional gain in resolution would be possible by increasing the fraction of events fitted as 5-jet (see Figure 5.7). This was implemented by reducing the $y_{\text{cut}}$ value from 0.004 to 0.002. The fraction of 5-jet events thereby increased from 30% to 50% of the selected WW events. A further reduction of $y_{\text{cut}}$ would only increase the
amount of CPU time needed without improvement in the mass resolution.

- Improved background description
  Cross-checks of the shape of the $q\bar{q}\gamma$ background in the 2D reconstructed mass plane revealed that the shape is actually more flat than had been thought previously. In fact a completely flat description instead of a 2D phase space function turned out to be a better description and gave a small improvement in W mass resolution.

- Introduction of a soft anti-b-tag cut to reduce the background from heavy flavour ZZ events (by 17%) and the QCD backgrounds nearly without loss of WW events, exploiting the fact that the decay of $W \rightarrow b\bar{q}$ is Cabibbo suppressed (section 2.2, page 23).

A further description of this analysis and its results is given in chapter 6-8.

5.5 Application to the semi-leptonic channel

In the $q\bar{q}\ell\nu$ channel the mass information per event is less affected by ambiguities than in the $q\bar{q}qqq$ channel, as jet clustering and jet pairing do not play a role in separating the decay products of the two W bosons. However, the missing neutrino introduces a new challenge. Depending on the decay angles of the two W bosons the topology of the final state can differ. In some configurations the neutrino causes a larger uncertainty on the fitted mass than in other configurations. This results in a large spread from event to event in the W mass resolution. In order to take the event resolution into account, the published DELPHI results are based on a method that is similar to the 1D ideogram approach — using 1D Gaussian event resolutions and the event purity in calculating likelihood curves for each event [1, 2, 3].

Semi-Leptonic Ideograms

It was realised for the first time at the LEP2 workshop [67] that in some events the missing neutrino can even lead to double solutions in the constrained fit, visible as double minima in the $\chi^2$ as function of the mass in an equal mass fit.

Also here the Ideogram method is a natural way to take into account such non-Gaussian resolution functions. First studies with a 2D Ideogram method showed that the convolution of a 2D Breit-Wigner with such a double kinematic solution can even lead to a triple ambiguity in the W mass likelihood, as illustrated in Figure 5.8.

To demonstrate the power of 2D ideograms in the semi-leptonic channel in the following chapter also a semi-leptonic 2D Ideogram analysis is described. In this analysis an extra ambiguity is taken into account, viz. the hypotheses that the observed electron or muon either was a direct decay product of a W boson, or just one of the decay products of the tau in a $q\bar{q}\tau\nu$ event, in which case the constrained fit is slightly different due to the special treatment of the tau.
Figure 5.8: Example of a semi-leptonic 2D Ideogram (bottom left), and the event W mass likelihood curve derived from it (bottom right). In this event the ideogram has a clear non-Gaussian shape, and three maxima are visible in the likelihood curve. The maxima occur for the three values of $m_W$ for which the BW has the largest overlap with the boomerang shaped maximum in the Ideogram (illustrated in the top right plot). For those three cases the region of most significant overlap is indicated by a dashed circle. The boomerang shape is believed to originate from multiple solutions for the reconstruction of the neutrino due to the relatively large errors on the jet energies (top left).
5.6 Systematics, Jackknife and MLBZ method

All of the W mass analyses discussed in this chapter rely either directly or indirectly on Monte Carlo simulation to correct for detector effects, experimental cuts and statistical approximations made. Imperfections in the Monte Carlo simulation will therefore cause systematic errors in each of these analyses, and the different analyses are expected to have very similar sensitivity to most of these systematic effects.

The addition of the 189 GeV data brought the LEP combined statistical error on the W mass down to the level of the quoted systematic errors. Therefore, extensive studies have been made to better understand possible systematic effects, as described in detail in chapter 7. Regarding this systemsatics study, in this historical account, two aspects should be mentioned: the use of the Jackknife and the development of the MLBZ method.

Event likelihood curves and the Jackknife

For the W mass measurement systematic effects have to be studied and controlled at the level of a few MeV/c². Using the ‘brute force’ approach, for each possible systematic effect a large number of Monte Carlo events would have to be generated to obtain such a precision: with a typical W mass resolution of about 3 GeV/c² per event O(1 million) q̅q̅q̅q̅ events are needed to reach a precision of 3 MeV/c² per effect per channel. The precision of complicated systematic studies therefore easily becomes limited by statistics.

Fortunately, many systematics effects can be studied in a much more powerful and efficient way, applying the (small) systematic disturbance to a sample of generated Monte Carlo events, and taking the difference in fitted W mass with and without applying the effect. By keeping the 4-fermion configurations of the events fixed, this approach can give statistically very precise results. In fact, in most cases the statistical error turns out to be proportional to the size of the effect, which is a most desirable property. But the problem is how to determine the statistical precision, now that the samples with and without systematic effect have become correlated.

It turns out that the Jackknife method (see section 4.4) is an excellent solution to this problem. Especially in combination with event-by-event likelihood curves the Jackknife method provides a convenient and quick way to determine the statistical error on the shift in W mass between two samples containing the same events.

The ability to determine the statistical significance of results found in the systematics studies was a great advantage leading more quickly to more complete and more precise results with modest Monte Carlo requests. The Jackknife method was used in most of the numbers quoted in chapter 7 and played a crucial role to determine the precision of results obtained with the MLBZ method.

MLBZ

With the addition of the 189 GeV data the systematic uncertainty due to possible imperfections in the jet fragmentation modelling emerged as the dominating systematic error quoted on the combined LEP W mass measurement.

The MLBZ method was developed to study a possible systematic effect related to jet reconstruction directly from the data, using hadronic Z⁰ events.
The basic idea is to take pairs of hadronic $Z^0$ events from the $Z^0$ calibration runs that are taken every year, boost them in opposite direction and then superimpose them emulating a 4-jet hadronic WW event (see Figure 5.9). The W mass analysis is applied directly to these ‘Mixed Lorentz Boosted $Z^0$’ events in order to measure the ‘$Z$ boson mass’. This can be done on data and Monte Carlo simulation separately, and the difference of the measured mass (rescaled to the W mass scale) is a measure of the systematic error on the W mass due to jet reconstruction. A detailed example of an implementation of the MLBZ method is given in Appendix A and its results and interpretation are summarised in chapter 7.

Figure 5.9: Illustration of the basic principle of the MLBZ method.