Experimental aspects of colour reconnection

M.F. Watson† and N.K. Watson‡

† CERN, European Organisation for Particle Physics, CH-1211 Geneva 23, Switzerland

‡ School of Physics and Space Research, University of Birmingham, Birmingham B15 2TT, UK

Abstract. This report summarises experimental aspects of the phenomena of colour reconnection in W^+W^- production, concentrating on charged multiplicity and event shapes, which were carried out as part of the "Phenomenology Workshop on LEP2 Physics, Oxford, Physics Department and Keble College", 14–18 April, 1997. The work includes new estimates of the systematic uncertainty which may be attributed to colour reconnection effects in experimental measurements of M_W.

Short title: Experimental aspects of colour reconnection

September 2, 1997

1. Introduction

Colour reconnection (also referred to as 'rearrangement' or 'recoupling') in W^+W^- decays has been the subject of many studies (e.g. [1, 2, 3]) and at present there is agreement that observable effects of interference between the colour singlets in the perturbative phase are expected to be small. In contrast, significant interference in the hadronisation process appears a viable prospect but, with our current lack of knowledge of non-perturbative QCD, such interference can only be estimated in the context of specific models [2, 4, 5, 6, 7, 8]. In the studies described below, experimentally accessible features of these models† are investigated, paying particular attention to the bias introduced to a typical measurement of M_W by direct reconstruction of the decay products.

Throughout this section reconnection effects were studied using: PYTHIA 5.722, type I and type II superconductor models (with the string overlap integral in type I case characterised by $\rho = 0.9$) [2, 3]; ARIADNE 4.08 allowing reconnection between the two W bosons; and HERWIG 5.9, in both its default reconnection model and also a 'colour octet' variant in which merging of partons to form clusters was performed on a nearest neighbour basis‡. In all cases, the tuning of the models was as used in reference [9].

2. Inclusive charged multiplicity

It has been suggested [2, 4] that simple observable quantities such as the charged multiplicity in restricted rapidity intervals may be sensitive to the effects of colour reconnection. More recently [8] it was suggested that the effect on the inclusive charged multiplicity itself may be larger than previously considered and that the mean hadronic multiplicity in $W^+W^-\to q\overline{q}q\overline{q}$ events, $\langle N_{ch}^{4q}\rangle$, may be as much as 10% smaller than twice the hadronic multiplicity in $W^+W^-\to q\overline{q}\ell\overline{\nu}_\ell$ events, $\langle N_{ch}^{qq\ell\nu}\rangle$. It was also reported during this workshop that the effects of Bose-Einstein correlations may increase $\langle N_{ch}^{4q}\rangle$ by $\sim$ 3–10% (see [10]).

The shifts in $\langle N_{ch}^{4q}\rangle$ at the hadron level predicted by the models studied thus far are given in table 1, where $\Delta\langle N_{ch}^{4q}\rangle$ is defined as the change in mean multiplicity relative to the 'no reconnection' scenario of each model. From these, it is clear that the multiplicities themselves and also the magnitude and sign of the predicted shifts are model dependent.

In this study, the precision with which such tests may be performed is quantified. As a starting point for such tests, it was first verified that in the absence of reconnection effects $\langle N_{ch}^{4q}\rangle = 2\langle N_{ch}^{qq\ell\nu}\rangle$ in the models PYTHIA and HERWIG. The statistical uncertainty of this test was $\mathcal{O}(0.1\%)$. Next, samples of 10^5 HERWIG and PYTHIA W^+W^- events were generated at $\sqrt{s} = 171$ GeV with a full simulation of the OPAL detector, and realistic event selections were applied for both $W^+W^-\to q\overline{q}q\overline{q}$ and $W^+W^-\to q\overline{q}\ell\overline{\nu}_\ell$ (ℓ = e, μ and τ). The efficiency in each case was $\sim$80%, while the purity is $\sim$ 80% for $W^+W^-\to q\overline{q}q\overline{q}$ and $\sim$ 88% for the $W^+W^-\to q\overline{q}\ell\overline{\nu}_\ell$ channel.

The resulting (uncorrected) charged multiplicity distributions for the hadronic and semi-leptonic channels are shown in Figs. 1(a) and 1(b), respectively. The simulated data correspond to an integrated luminosity of 10 pb^{-1} at $\sqrt{s} = 171$ GeV, i.e. that

† In studying these models, no retuning was performed when reconnection was enabled.
‡ This was suggested by B R Webber, as a partial emulation of the model of reference [8].

Table 1. Mean charged multiplicities, $\langle N_{\rm ch}^{4q} \rangle$, and predicted shifts for various models

model		$\langle N_{\rm ch}^{4q} \rangle$	$\Delta \langle N_{\rm ch}^{4q} \rangle$ (%)
PYTHIA	normal	38.64	
	type I	38.21	-1.1 ± 0.1
	type II	38.39	-0.7 ± 0.1
HERWIG	normal	37.07	
	reconnected ($P = \frac{1}{9}$)	37.25	$+0.5\pm0.1$
	reconnected ($P = 1$)	38.38	$+3.5\pm0.1$
ARIADNE	normal	38.14	
	reconnected	37.07	-2.8 ± 0.1

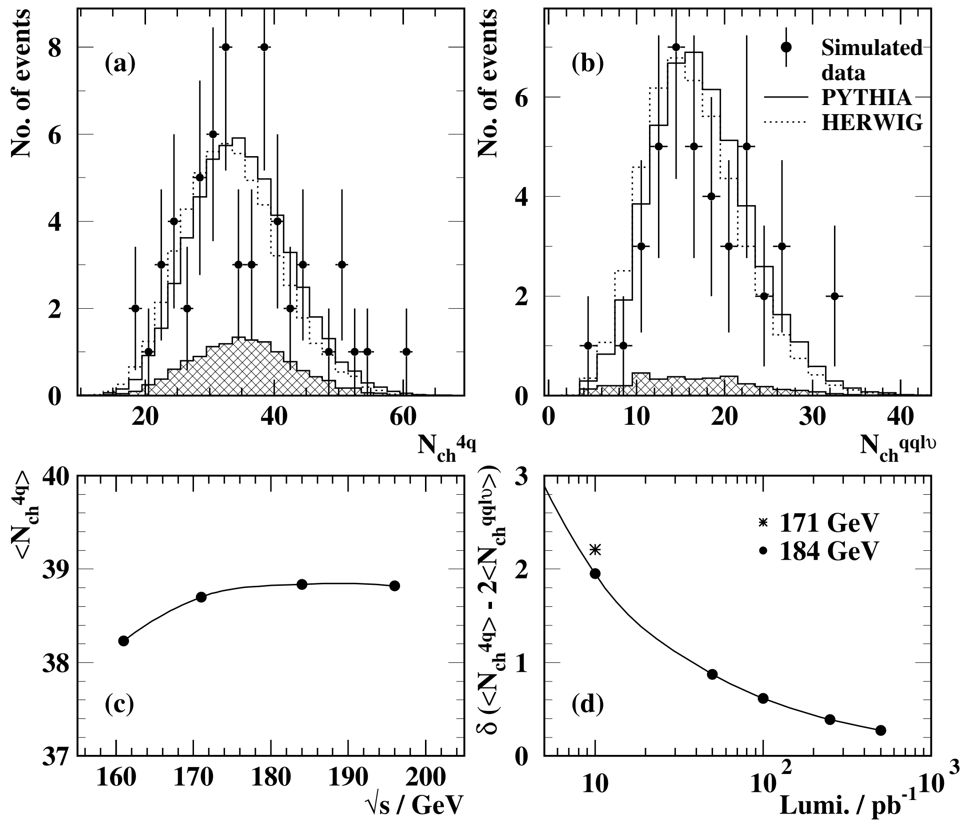

Figure 1. Inclusive charged multiplicity distributions with 10 pb^{-1} of fully simulated data, with background indicated hatched, at $\sqrt{s}$ = 171 GeV for (a) $\mathrm{W^+W^-} \to \mathrm{q\overline{q}q\overline{q}}$, and (b) $\mathrm{W^+W^-} \to \mathrm{q\overline{q}}\ell\overline{\nu}_\ell$ events. (c) Variation of $\langle N_{\rm ch}^{4q} \rangle$ with $\sqrt{s}$. (d) Luminosity dependence of the statistical uncertainty of $\langle N_{\rm ch}^{4q} \rangle - 2\langle N_{\rm ch}^{qq\ell\nu} \rangle$ (units of multiplicity).

delivered by LEP in 1997. In both distributions, the expected background is shown as a hatched histogram. The significant level of $Z^0/\gamma \to q\overline{q}$ background is apparent in the fully hadronic channel.

To extract the mean charged multiplicity at the hadron level at a fixed centre-of-mass energy from such distributions, one can apply a simple correction, based on Monte Carlo, to the observed mean value, after subtracting the expected background contribution. An alternative is to carry out a matrix-based unfolding procedure using the event-by-event correlation between the charged multiplicity at the hadron level and that observed in the detector after all analysis cuts have been performed. A separate correction for the effects of initial state radiation are necessary in this latter case. A third alternative is to integrate the fragmentation function but this is not discussed here.

Based on the the simulated data in Fig. 1(a) and (b), the expected statistical uncertainty on the difference $\langle N_{\rm ch}^{\rm 4q}\rangle - 2\langle N_{\rm ch}^{\rm qq\ell\nu}\rangle$ for an integrated luminosity of 10 pb^{-1} is 2.2 units, or 5.7% on $\langle N_{\rm ch}^{\rm 4q}\rangle$. The evolution of the precision of such difference measurements with more data is estimated using the following assumptions. Firstly, the distributions of $N_{\rm ch}^{\rm 4q}$ and $N_{\rm ch}^{\rm qq\ell\nu}$ are seen to be relatively insensitive to changes in centre-of-mass energy once away from the threshold region, as illustrated by the energy dependence of $\langle N_{\rm ch}^{\rm 4q}\rangle$ in Fig. 1(c). Therefore both the mean and the corresponding rms are assumed constant at their 184 GeV values. Secondly, above $\sqrt{s} = 184$ GeV the $\rm W^+W^-$ production cross-section is predicted to vary by less than 10% in the region up to $\sqrt{s} < 200$ GeV, and so a constant cross-section of 16 pb is assumed. Thirdly, it is assumed that the selection efficiency at 171 GeV may be maintained at higher energies. The expected background cross-section is not important as it is subtracted in performing the measurement. Given these assumptions, the dependence of the expected statistical error on the difference, $\delta(\langle N_{\rm ch}^{\rm 4q}\rangle - 2\langle N_{\rm ch}^{\rm qq\ell\nu}\rangle)$, is shown as a function of integrated luminosity in Fig. 1(d).

Typically in such multiplicity determinations, systematic effects become significant below a statistical precision of 0.5 units of multiplicity. Uncertainty in the modelling of 4-jet like $\rm Z^0/\gamma \to q\overline{q}$ background with parton shower Monte Carlos in the fully hadronic channel may become a significant systematic.

3. Event shapes

Global event shape variables have been considered in earlier studies as potential signatures for reconnection [2, 4, 8]. In most studies the predicted effects on such observables induced by reconnection has been sufficiently small that detection would be marginal, even with an integrated luminosity of 500 pb^{-1}.

The choice of a 'no reconnection' reference sample with which to compare data deserves some thought. In trying to find sensitive observables, using the models alone is ideal. However, once possible signatures have been developed, and one starts to search for effects in data, it will be invaluable to have a well defined 'no reconnection' reference sample in data to reduce model and tuning dependence. LEP 1 data provide a high statistics reference, but additional assumptions are necessary in either extrapolating energy scales, or in combining pairs of $\rm Z^0/\gamma \to q\overline{q}$ to emulate $\rm W^+W^- \to q\overline{q}q\overline{q}$ events without reconnection. It is also necessary to assume that data recorded and processed by the detectors before 1996 can be directly compared with those recorded near the end of the LEP 2 programme. For some signatures, the ideal reference data are $\rm W^+W^- \to q\overline{q}\ell\overline{\nu}_\ell$ events. However, this sample has only limited size and the comparison may require the association of pairs of jets with Ws in the fully hadronic channel, a procedure which experimentally introduces more uncertainty. In

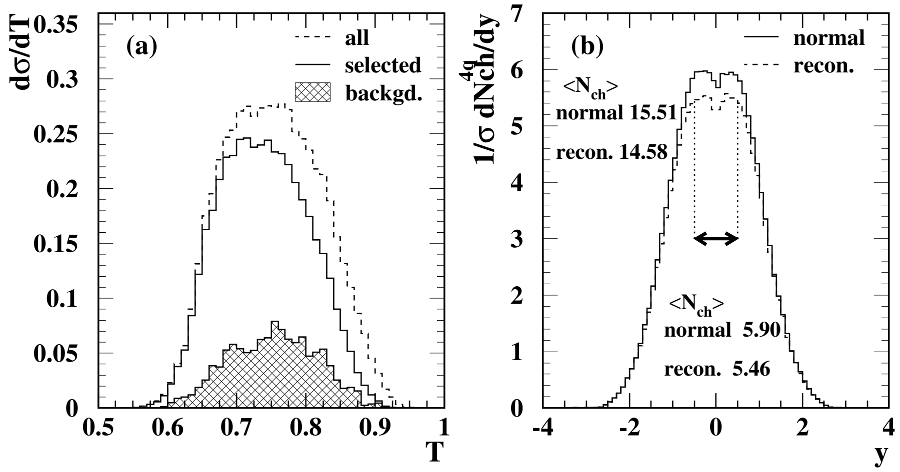

Figure 2. (a) Effect of typical experimental selection on thrust distribution, and (b) hadron level rapidity distribution in ARIADNE for $p < 1$ GeV.

the following, all changes are relative to the 'no reconnection' version of each Monte Carlo model and all samples are $\mathrm{W^+W^-}\!\to\!\mathrm{q\overline{q}q\overline{q}}$.

This study compares the differences in the rapidity distribution of charged particles, $\mathrm{d}N_{ch}/\mathrm{d}y$, relative to the thrust axis of each event, in the central region, $|y| < 0.5$ and for all y, as suggested in [1, 2, 4]. As the effects are expected to be more pronounced for softer particles, the distribution is studied for three momentum ranges, $p < 0.5$ GeV, $p < 1$ GeV and all momenta. It has been suggested [4, 8] that reconnection effects may be more pronounced in specific topologies where the quarks from different Ws are close to one another, therefore events are also studied for all thrust values and for $T > 0.76$. One aspect not considered in previous studies has been the effect of applying a realistic event selection, which is necessary to reduce the large background ($\sigma(\mathrm{Z^0}/\gamma \to \mathrm{q\overline{q}}) \sim 20\sigma(\mathrm{W^+W^-}\!\to\!\mathrm{q\overline{q}q\overline{q}})$). As this is dominated by two-jet like events, the efficiency for selecting $\mathrm{W^+W^-}\!\to\!\mathrm{q\overline{q}q\overline{q}}$ events in a similar configuration is relatively small, as illustrated in Fig. 2(a); $\sim 38\%$ of $\mathrm{W^+W^-}\!\to\!\mathrm{q\overline{q}q\overline{q}}$ events selected satisfy $T > 0.76$, falling to $\lesssim 0.05\%$ for $T > 0.92$.

In [8], the rapidity was studied relative to the axis bisecting the two di-jet axes, as a function of the angle separating these axes. Experimentally, without any reliable charge identification algorithm to separate quarks from anti-quarks, the specific angle proposed in [8] must at best be folded in experimental analyses, and also requires pairing of jets into Ws. While the reliability of associating the 'correct' jets together is possible with moderate efficiency using kinematic fits, selecting high thrust events was used in the current studies for expediency and simplicity. As the shifts in M_{W} expected are modest compared to the experimental mass resolution on an event-by-event basis, it is worth considering the use of kinematic fits in which our current knowledge of M_{W} is applied as a constraint, in a similar way to that used by experimental TGC analyses.

Hadronic events were generated using the models PYTHIA, HERWIG and ARIADNE, with and without a simulation of the OPAL detector, and $\mathrm{d}N_{ch}/\mathrm{d}y$ studied within the ranges of y, p and T described above. A smearing simulation of the OPAL detector, which is reliable for studies in the $\mathrm{W^+W^-}\to\mathrm{q\overline{q}q\overline{q}}$ channel and necessary to achieve the relatively high statistics required, was used herein and also to estimate shifts in M_W.

As an example of how the differences may be concentrated in restricted rapidity intervals, Fig. 2(b) shows the $\mathrm{d}N_{ch}/\mathrm{d}y$ distribution for $p < 1$ GeV in ARIADNE, for events with and without reconnection. Changes in charged multiplicity, $\Delta\langle N_\mathrm{ch}^\mathrm{4q}\rangle$, within given p and y intervals are summarised in Fig. 3(a) for each of the models introduced in table 1, without detector simulation. The left (right) hand side of the figure shows the percentage change in $\langle N_\mathrm{ch}^\mathrm{4q}\rangle$ for the three momentum ranges considered for all y ($|y| < 0.5$). The leftmost points in this figure correspond to the results of table 1. Fig. 3(b) gives the analogous results for $T > 0.76$. For illustration, statistical errors corresponding to an integrated luminosity of 500 pb^{-1} are given for the 'HERWIG colour octet' model.

It is seen that in all models the magnitude of the change increases when only low momentum particles are considered. Applying a thrust cut such as $T > 0.76$ rejects $\sim 40\%$ of events and may change $\langle N_\mathrm{ch}^\mathrm{4q}\rangle$ by up to two units, but differences relative to the 'no reconnection' scenarios are essentially unchanged, therefore the sensitivity is reduced. The predicted maximum statistical significance of $\Delta\langle N_\mathrm{ch}^\mathrm{4q}\rangle$, as well as its sign, depends strongly on the model, varying from $\sim 6\sigma$ for ARIADNE and the HERWIG 'colour octet' model, $\sim 3.5\sigma$ for PYTHIA type I, $\sim 2\sigma$ for PYTHIA type II, down to $\sim 0.8\sigma$ for HERWIG. The point of maximal sensitivity is indicated (square markers) for each model in the figure. Similar trends were observed in studies with detector simulation but typically $\Delta\langle N_\mathrm{ch}^\mathrm{4q}\rangle$ was found to be $\sim 50\%$ smaller.

It may be possible to increase the sensitivity to reconnection effects using charged multiplicity based methods, by considering particle distributions relative to the $\mathrm{W^+W^-}$ decay axis, as reconstructed using kinematic fits. In [3], an alternative multiplicity signature ('interjet multiplicity') was introduced, having similar sensitivity to integrating $\mathrm{d}N_{ch}/\mathrm{d}y$ for $|y| < 0.5$. This interjet multiplicity was similar in idea to methods normally used to quantify the 'string effect' in 3-jet $\mathrm{e^+e^-}$ events. It was suggested that this be studied further, using the shape of the particle density distribution as a function of the angular separation between jet pairs, rather than restricting the study to the integrated particle density in the fixed angular regions. However, the 4-jet case is somewhat more complex than the familiar 3-jet case, being non-planar, and so this was not pursued during the workshop.

4. Shifts in M_W

Extracting M_W from the decay products observed by experiments is non-trivial, requiring much attention to bias induced from effects such as initial state radiation, detector calibration, imperfect modelling of the underlying physics processes and of the apparatus. In comparison to this, estimating a shift which could result from the effects of reconnection phenomena is straightforward, as the value of interest is the relative shift between M_W determined in two different scenarios of the same model. The absolute value of "M_W" obtained is not central to these studies. However, there are still many uncertainties inherent in such studies, such as sensitivity of the method used to extract M_W to changes in $\sqrt{s}$, to tuning of the Monte Carlo models (e.g.

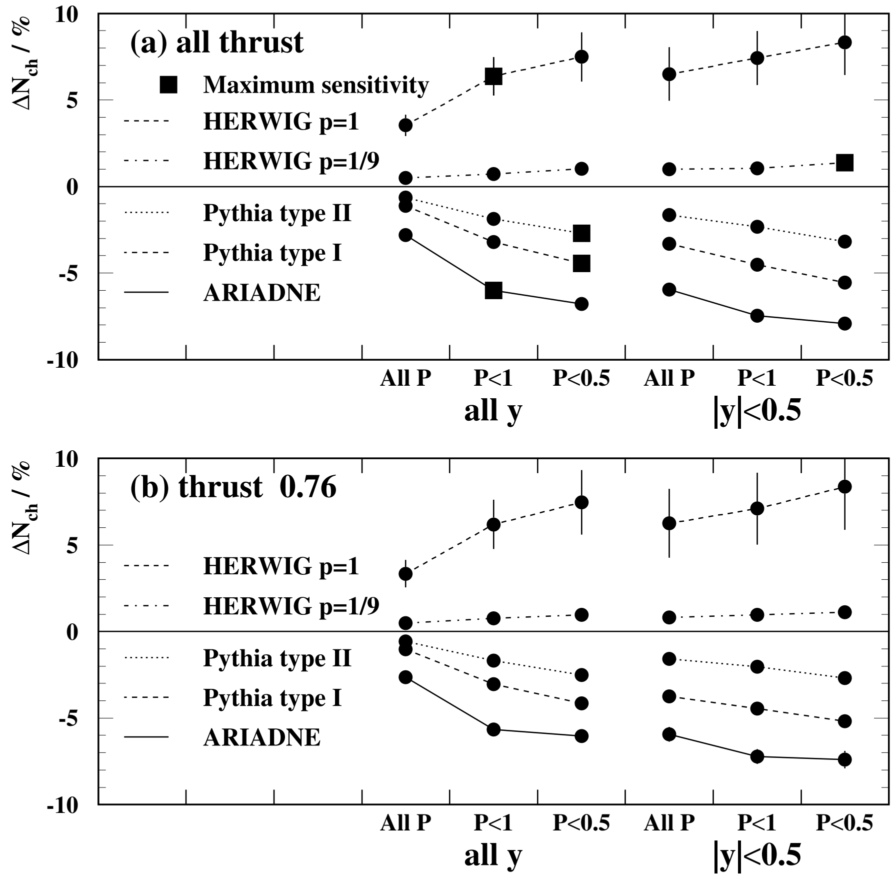

Figure 3. Fractional change in charged multiplicity as a function of maximum particle momentum, in two rapidity regions, for (a) all T, and (b) $T > 0.76$. See text for details.

virtuality cut-offs in the parton shower development), to treatment of combinatorial background and ambiguous jet-jet combinations, and the range over which fitting is performed to name but a few.

In these studies, the method used to extract M_{W} followed closely that used by OPAL for its preliminary M_{W} results using 172 GeV data. In this, events with detector simulation are first selected using the same procedure as noted earlier. Four jets are formed using the $k_{\perp}$ jet finder, corrected for double counting of energy within the apparatus, and a parametrisation of the errors on the measured jet 4-momenta is carried out. A five-constraint kinematic fit, in which the jet-jet pair masses are constrained to be equal, is performed for each of the three possible jet-jet pairings, event by event. A mass distribution is constructed using the mass from the combination having the highest probability from the kinematic fit in each event if this has probability greater than 1%. A second entry is also admitted if the second most

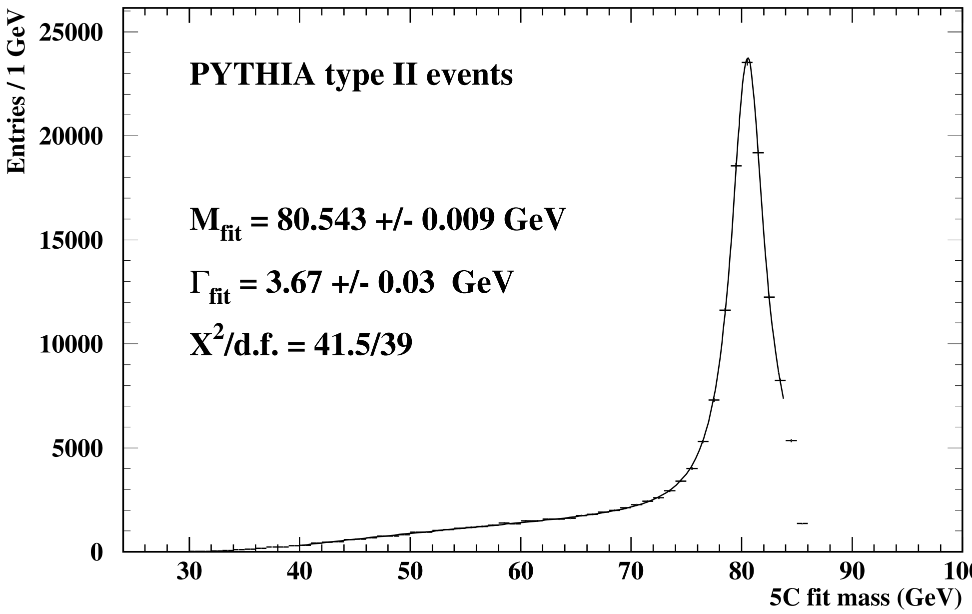

Figure 4. Typical mass distribution with fit results with detector simulation and event selection.

probable fit result has probability greater than 1% and within a factor of three of the highest probability combination. The aim of this is to include additional mass information for events in which the most probable fit combination is incorrect. In such events, these two masses are essentially uncorrelated. A typical mass distribution formed by this procedure is given in Fig. 4.

This method was applied to simulated events from each of the models in turn, and the shifts obtained are summarised in table 2, where uncertainties on these shifts are statistical. The ARIADNE model predicts a modest shift in mass of approximately 50 MeV. No significant shift is predicted by the models PYTHIA and HERWIG. In an earlier study, performed in a similar way, significant shifts were determined [3]. The PYTHIA and ARIADNE models considered in the present study were also included in [3], albeit with different model dependent parameters and looser event selection criteria.

Table 2. Table of shifts in M_{W} for each model.

model		$\langle \Delta M_{\mathrm{W}} \rangle$ (MeV) selected events ($\epsilon \simeq 80\%$)	all events
PYTHIA	type I	$+18\pm11$	$+11\pm11$
	type II	-13 ± 11	-19 ± 11
HERWIG	reconnected ($P = \frac{1}{9}$)	-16 ± 16	-19 ± 16
	reconnected ($P = 1$)	$+13\pm15$	$+8\pm14$
ARIADNE	reconnected	$+51\pm16$	$+51\pm15$

One quite plausible explanation proposed was that the difference was due to the significantly more stringent event selection currently used. It has been shown that the current selection preferentially rejects events having two-jet like characteristics, which is where reconnection effects may be expected to be prevalant. The rejection of these events does not appear to be the reason for small mass shifts, as a similarly small effect is observed when all events are selected, as seen in table 2.

Many possible sources for the difference were investigated in the context of the PYTHIA models. Neither changes in the tuning of PYTHIA/JETSET by OPAL† to improve the description of LEP 1 data, nor the different centre-of-mass energy ($\sqrt{s} = 175$ GeV in [3]) were found to be significant. The current analysis procedure is slightly different to that in [3]. However, significant shifts are still found when the current procedure is applied to the same simulated events used in [3]. Conversely, applying the former procedure of [3] to the samples herein does not induce a significantly larger mass shift.

One apparently significant effect was found to be the choice of mass assigned to jets in performing kinematic fits. As discussed in [3], this choice is not unique. In the analysis of [3], the hadronic jets were assumed massless whereas in the current studies, the measured jet mass was used. Re-analysing the same simulated events of [3] but assigning measured masses to the jets reduces the mass shifts estimated, e.g. shifts quoted in [3] of 130 ± 40 MeV (type I) and 50 ± 40 MeV (type II) become 70 ± 40 MeV and 30 ± 40 MeV, respectively. For comparison, a sample of 200 000 fully hadronic type I events were generated at $\sqrt{s} = 175$ GeV using *identical* model parameters and program versions, and analysed using the procedure of [3], also using measured jet masses. This yielded an estimated shift of 46 ± 16 MeV. It should be noted that fluctations due to finite Monte Carlo statistics have to be considered when comparing with the results of [3], in which samples sizes for the analoguous studies were 50 000 events.

Comparing the results for mass shifts in table 2 with multiplicity shifts in table 1, it can be seen that any relationship between them is model dependent. Furthermore, relatively large shifts in the charged multiplicity do not necessarily lead to a significant shift in M_{W}.

5. Future

The future for experimental studies of colour reconnection is quite open. There is clear model dependence in signatures and mass shifts may be smaller than earlier proposed [3], although there are other models available [7, 8] which were not tested in this study from which different conclusions may be drawn. A necessary condition for a model to be taken seriously is that it should describe the data, therefore tuning of models has to be addressed. With the current statistical precision of LEP 2 data, none of the models has been put to a stringent test. The effect of background cannot be ignored in the $\mathrm{W^+W^-} \rightarrow \mathrm{q\bar{q}q\bar{q}}$ channel as it proves difficult to remove. More sophisticated selections may be developed, but typically these make use of non-trivial correlations between observables, which may be poorly described by the models. A particular concern is the description of parton shower Monte Carlos to describe the hard, 4-jet like background which is selected. The remaining point of note is that given the model dependence inherent to such studies, it is most important to develop signatures which

† Among these, the cut-off parameter, Q_0, was increased from 1.0 GeV in the similar investigation of [3], to 1.9 GeV.

can be tested taking the 'no reconnection' scenario from data themselves.

References

[1] Gustafson G, Pettersson U and Zerwas P M, 1988 *Phys. Lett.* **209B** 90
[2] Sjöstrand T and Khoze V A 1993 *Z. Phys.* C **62** 281
[3] *Determination of the Mass of the W boson*, Kunszt Z *et al* , in 'Physics at LEP2', eds. Altarelli G, Sjöstrand T and Zwirner F, CERN Report 96-01, Vol.1, p.141-205 (1996)
[4] Gustafson G, Häkkinen J, 1994 *Z. Phys.* C **64** 659
[5] Program ARIADNE, Lönnblad L, 1992 *Comp. Phys. Commun.* **71** 15
[6] Program HERWIG, Marchesini G *et al* , 1992 *Comp. Phys. Commun.* **67** 465
[7] Todorova-Nová S, DELPHI 96-158 PHYS 651, *Colour Reconnection in String Model* (unpublished).
[8] Ellis J and Geiger K, preprint CERN-TH/97-46 (1997); 1996 *Phys. Rev.* D **54** 1967
[9] OPAL collaboration: Ackerstaff K *et al* , 1996 *Phys. Lett.* **389B** 416
[10] Kartvelishvili V, proceedings of "Phenomenology Workshop on LEP2 Physics, Oxford, Physics Department and Keble College", 14–18 April, 1997, to be published in J. Phys. G.