Plenary sessions: **6, 9**

Measurement of the gluon splitting rate to $b\bar{b}$ in hadronic Z decays

The ALEPH Collaboration

Abstract

A measurement of the gluon splitting rate to a $b\bar{b}$ pair ($g_{b\bar{b}}$) is presented. The measurement uses a combination of topological cuts and a lifetime tag on the jets. The preliminary result is: $g_{b\bar{b}} = (2.57 \pm 0.40(\mathrm{stat}) \pm 0.87(\mathrm{syst})) \times 10^{-3}$.

Submitted to the 1997 lepton-photon conference, Hamburg

1 Introduction

The process of the splitting of a gluon into an heavy quark pair is relatively poorly known, both theoretically and experimentally, despite the fact that is one of the elementary processes in QCD.

The rate of events in which a gluon splits into a $b\bar{b}$ pair ($g_{b\bar{b}}$) is predicted to be almost an order of magnitude smaller than the corresponding one into a $c\bar{c}$ pair ($g_{c\bar{c}}$).

The rate of such events, relative to the hadronic cross section, has been predicted by Seymour[1] at different centre-of-mass energies, within the framework of perturbative QCD. The result is exact to leading order in α_s and resums large logarithms and next-to-leading logarithmic terms to all orders.

This rate is sensitive mainly to the Λ_{QCD} parameter, and can vary by almost 25% when Λ_{QCD} varies by a factor 2.

The theoretical prediction gives the values in table 1.

Table 1: Predicted values for the gluon splitting rates. The errors are due to the assumption in quark masses. For the c quark, the central mass is $m_c = 1.5$ GeV, the lower error is for $m_c = 1.8$ GeV and the upper for $m_c = 1.2$ GeV. For the b quark, the central mass is $m_b = 4.75$ GeV, the lower error is for $m_b = 5$ GeV and the upper for $m_b = 4.5$ GeV.

	$g_{c\bar{c}}(\%)$	$g_{b\bar{b}}(\%)$
$\Lambda_5^{MS} = 150\text{MeV}$	$1.35^{+0.48}_{-0.30}$	0.20 ± 0.02
$\Lambda_5^{MS} = 300\text{MeV}$	$1.85^{+0.69}_{-0.44}$	0.26 ± 0.03

The Parton Shower Monte Carlo[2, 3] provide description of the parton shower cascade accurate to leading log order, together with matching to the first-order result for the production of a hard gluon. The rate predicted by these Monte Carlo is in good agreement with the theoretical prediction.

The experimental measurements of $g_{c\bar{c}}$ have been performed by OPAL[4, 5] and recently by ALEPH[6] and show an evidence for an higher value with respect to the theoretical predictions.

The only measurement of $g_{b\bar{b}}$ has been reported by DELPHI[7] and is in agreement with the prediction, even if with large errors. The limited accuracy of such a measurement is one of the main errors in the measurement of the partial decay width of the Z to $b\bar{b}$ (R_b[8]).

This paper describes a measurement of $g_{b\bar{b}}$ making use of a lifetime tag and topological properties of the events.

2 Event selection

The measurement uses 3.7 million events collected from 1992 to 1995 which pass the standard ALEPH hadronic event selection[9] with the requirement that the vertex detector was fully operational. The analysis also uses 6 million simulated events produced with a generator based on JETSET 7.4 parton shower model[2].

Monte Carlo events were reweighted to take into account the latest LEPHF[10] recommendations. In particular, the rate of the gluon splitting in $c\bar{c}$ pairs was set to

$$g_{c\bar{c}} = (2.38 \pm 0.48)\%$$

while the rate of splitting to $b\bar{b}$ was set to

$$g_{b\bar{b}} = 0.3\%$$

A dedicated Monte Carlo production of $28,000$ $g_{b\bar{b}}$ and $90,000$ $g_{c\bar{c}}$ events has been used in order to better evaluate the efficiencies.

Besides the signal events, hereafter called $\mathcal{B}$, two additional categories of background events exist:

- events which don't contain any gluon splitting into heavy flavour at all, hereafter called $\mathcal{Q}$ events
- events in which a gluon splits to a charm quark pair, named $\mathcal{C}$ events.

3 Method

The analysis profits of the presence of two B hadrons with particular topology configuration.

After selecting 4–jet events the B jets candidates were tagged by a lifetime algorithm. The candidate pair was required to form a small angle. The other two jets originating from the hadronisation of the initial quarks were required to be in opposite hemispheres.

Jets were formed using charged tracks and neutral calorimeter objects using the DURHAM algorithm with $y_{cut} = 0.006$. Events with 5 or more jets (about 0.5% of the total in background and 2% for the signal) were forced to exactly 4 jets by varying y_{cut} to y_5.

The value of y_{cut} has been chosen such that the background events from $Z \to q\bar{q}$ show a 4–jet multiplicity of about 10%, while retaining a large fraction of the signal events.

After this requirement the 4–jet multiplicity in the Data is $(9.56\pm0.02)\%$, where the error is only statistical. In the Monte Carlo simulation the rate is $(9.42 \pm 0.01 \pm 0.10)\%$; the first error quoted is statistical, the second one is due to the uncertainty in $g_{c\bar{c}}$.

Jets containing B hadron tracks were searched for by making use of the information coming from the vertex detector, using the standard ALEPH b–tagging algorithm[12]. Therefore the 4–jet events were retained if each jet has at least one track with one VDET hit.

At this point the fraction of events that survive in Data and in Monte Carlo simulation is $(4.979\pm0.011)\%$ and $(4.908\pm0.009\pm0.054)\%$, respectively.

For all charged tracks in each jet we computed the confidence level for the *i-th* jet to originate from the primary vertex, $Pjet_i$. Then we define

$$probjet_i \equiv -\log_{10}(Pjet_i).$$

The four jets are then ordered in *probjet*, so that the first jet is more likely to have tracks from b decay than the others.

Figure 1 shows the *probjet* distributions for each jet in Data and Monte Carlo.

Monte Carlo shows a discrepancy with respect to data of about 10% for large values of the *probjet*, irrespectively of the jet ordering, as shown in figure 2(a). That region is mainly populated by $Z \to b\bar{b}$ events. In order to detect the cause of the problem, an enriched sample of b events was then obtained by searching for displaced secondary vertices in the event. The b purity of the sample as given by the Monte Carlo is about 80%. In this sample the *probjet* distributions of the data and simulation agreed as can be seen in fig. 2(b). Therefore the problem in the unbiased sample was due to a different heavy flavour composition in data and Monte Carlo simulation.

We then checked the heavy flavour composition of the sample as well as the jet tagging efficiencies. A double tag was performed by using a cut in $probjet > 3$ for the two most energetic jets in the event. The average jet tag efficiency in the Monte Carlo truth was predicted to be $22.36 \pm 0.13\%$ with a b purity of about 80%. The fraction of $Z \to b\bar{b}$ events (R_b) in 4-jets events in the Monte Carlo truth was $20.01 \pm 0.05\%$ while the $g \to b\bar{b}$ events amount to 3%. The correlation between the two jet efficiencies was taken

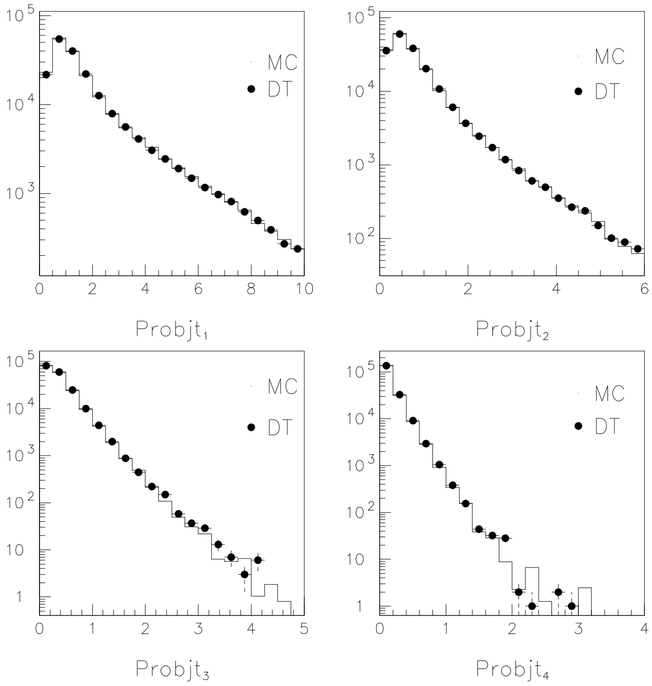

Figure 1: *Probjet* distributions for Data and Monte Carlo. Note both the log scale on the ordinate and the different scales on the abscissa.

into account and was $-7.5 \pm 1.0\%$. R_b was extracted from the fit to the data singly and double tagged events, and was $21.45 \pm 0.31 \pm 0.24\%$, where the first error comes from data statistics and the other from the uncertainties in the background efficiencies and b correlation. The average jet tagging efficiency in the data was found to be $22.82 \pm 0.40\%$, showing a very good agreement with the simulation.

Therefore we corrected R_b in the simulation to reproduce the one observed in the data. This correction was then taken into account when estimating the systematic errors.

The two jets with the largest $probjet$ were then considered as candidates for originating from the gluon splitting process $g \to b\bar{b}$.

Afterwards we considered the angle between the two candidate b jets. Normal $Z \to b\bar{b}$ events tend to have the b jets in two different hemispheres, while in the signal events they are more likely to form a smaller angle. Fig-

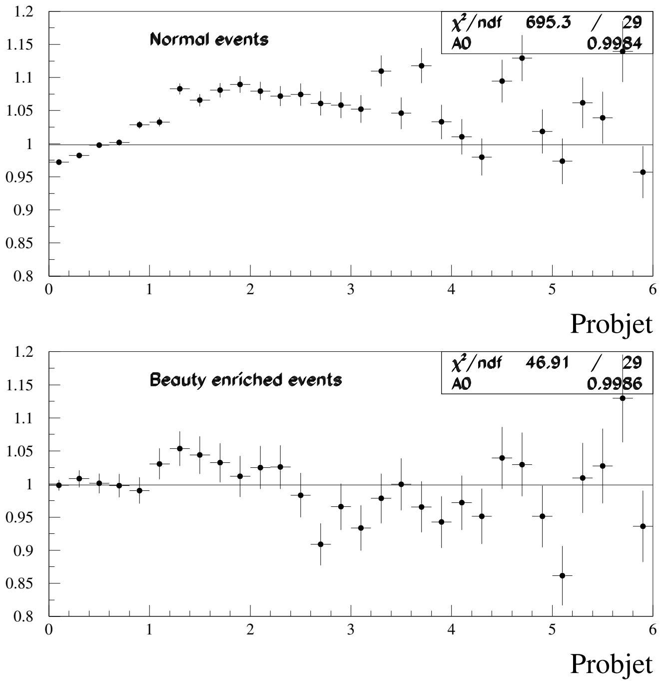

Figure 2: Ratio of the probjet distributions for Data and Monte Carlo in unbiased (upper plot) and $b\bar{b}$ enriched events (lower plot).

ure 3 shows the angular distributions of the candidate b jets (θ_{12}) in Data and in Monte Carlo simulation, as well as for signal and background events. To select $g_{b\bar{b}}$ events we have therefore requested $0.2 < \cos\theta_{12} < 0.9$.

The fraction of data events that passed this cut was in good agreement with the Monte Carlo prediction: $(1.135 \pm 0.006)\%$ in the data and $(1.115 \pm 0.005 \pm 0.013)\%$ in the simulation.

After the $\cos\theta_{12}$ cut the *probjet* distributions for signal and background show a clear difference as can be seen in figure 4.

The requirement

$$probjet_1 > 2.4 \quad , \quad probjet_2 > 1.7$$

selects an additional factor of 10 more signal than background events; at this point about 20% of the selected events are estimated to come from signal.

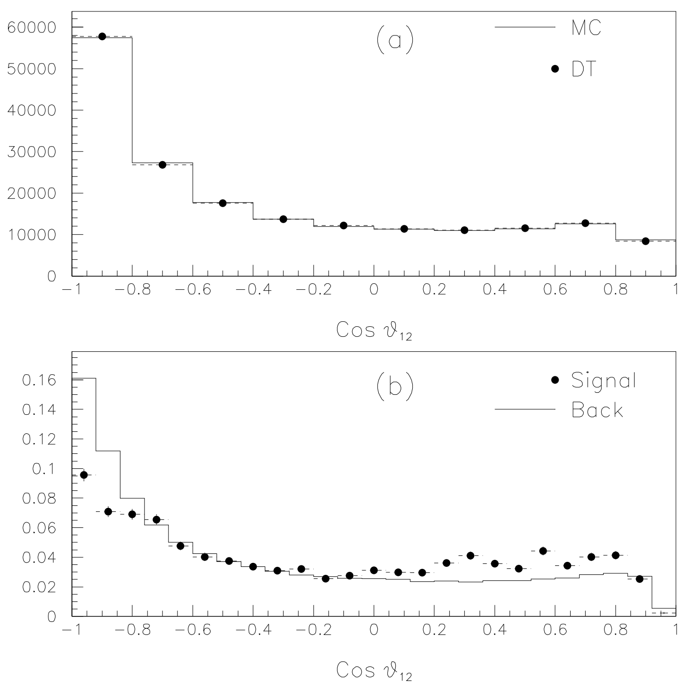

Figure 3: Cosine of the angle between the first two jets in Data and Monte Carlo (a) and signal and background (b).

We further noticed that for the simulated signal events the angle between the third and the fourth jet tends to be distributed at large values, while the background events populate the small angle region, as shown in Fig. 5. We required $-0.9 < \cos\theta_{34} < 0.1$.

Finally, the b jets coming from a gluon tend to have a softer momentum than the other two.

By requiring the momentum of the fourth jet to be greater than 11 GeV/c and the momentum of the first jet to be lower than 27 GeV/c we increase the purity of the signal to about 40%.

At this stage, about 70% of the background is coming from $Z \to b\bar{b}$ events.

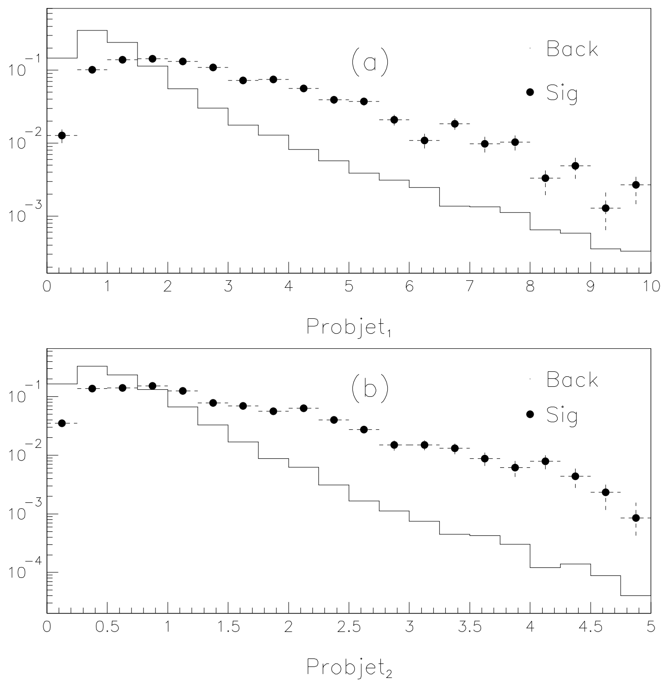

Figure 4: Probjet distributions for jet 1 (a) and jet 2 (b) in signal and background events after $\cos\theta_{12}$ cut.

4 Results

After all the above mentioned cuts 329 events are selected in the Data.

Table 2 shows the tagging efficiencies for the three categories of events, where the errors are statistical only.

From the above efficiencies and selected events one can extract the value of $g_{b\bar{b}}$:

$$g_{b\bar{b}} = \frac{f_d - (1 - g_{c\bar{c}})\epsilon_Q - g_{c\bar{c}}\epsilon_C}{\epsilon_B - \epsilon_Q}$$

where $f_d = (8.94 \pm 0.49) \times 10^{-5}$ is the fraction of events surviving the cuts in the data.

The measured value of the gluon splitting rate into $b\bar{b}$ pairs is

$$g_{b\bar{b}} = (2.57 \pm 0.40) \times 10^{-3}$$

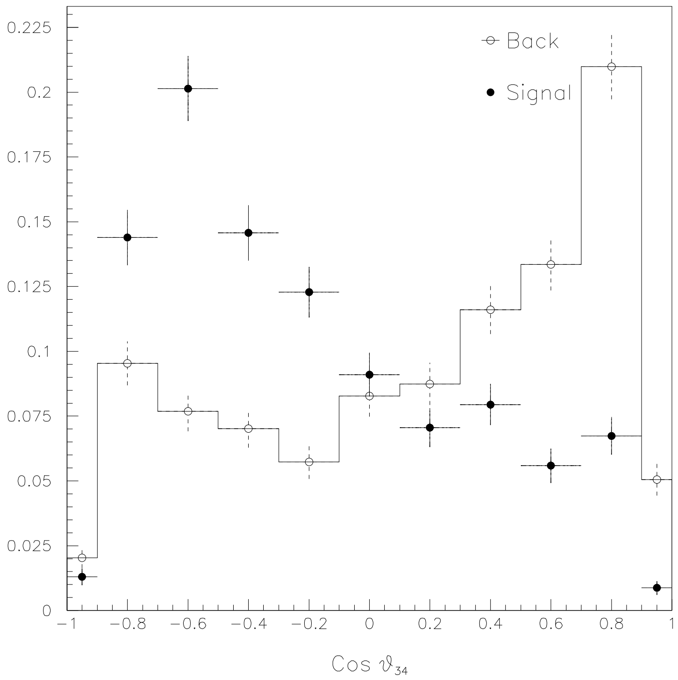

Figure 5: Cosine of the angle between the third and fourth jets in signal and background after the cuts in *probjet*.

where the error is statistical only.

5 Systematic errors

The efficiencies for the three event categories have been evaluated by Monte Carlo simulation. The inadequacy of the simulation in estimating these efficiencies reflects the uncertainty in the result.

The error due to the limited Monte Carlo statistics in the efficiency evaluation is $\Delta g_{b\bar{b}} = \pm 0.38 \times 10^{-3}$. This uncertainty mainly comes from the efficiency to tag $\mathcal{Q}$ events.

Since 70% of the $\mathcal{Q}$ background comes from $Z \to b\bar{b}$ events a source of systematic error is coming from the uncertainty in the knowledge of the physical processes in the simulation of heavy flavour production and decays.

Table 2: Efficiencies after all cuts for the three categories.

Source	Efficiency(%)
$\mathcal{B}$	1.24 ± 0.07
$\mathcal{C}$	0.048 ± 0.004
$\mathcal{Q}$	0.0047 ± 0.0004

We have varied all the physical simulation parameters within their allowed experimental ranges, as it was done in the R_b paper[8]. In particular, the B and C hadron lifetimes as well as production rates have been varied, following the latest LEPHF [10] recommendations. The charm topological BR and their rate into K^0_S have also been varied within the experimental limits[11].

Two additional sources of systematic uncertainties are the average charged multiplicity of B hadron decays (5.73 ± 0.35) and the rate of hard gluon events, in which the two original b quarks are pushed in the same hemisphere by a very energetic gluon ($2.2\% \pm 0.7\%$).

Table 3 resumes the main sources of physics uncertainties.

Table 3: Systematic errors in g_{bb} due to physics uncertainties.

Source	$\Delta g_{bb}(10^{-3})$
B hadron lifetimes	± 0.09
B hadron production	± 0.08
B hadron fragmentation	± 0.07
C hadron lifetimes	± 0.08
C hadron production	± 0.03
C hadron fragmentation	± 0.01
BR ($D \to K^0_s$)	± 0.13
D^+ charged multiplicity	± 0.09
D^0 charged multiplicity	± 0.20
D_s charged multiplicity	± 0.27
B hadron charged multiplicity	± 0.23
Hard Gluon	± 0.28
Total	± 0.54

We then considered the effect on the result due to the possible different angular distributions between data and simulation. We checked the jet relative angle distributions in data and Monte Carlo simulation before the

$\cos\theta_{12}$ cut. Jets were ordered in decreasing energy, in order not to be biased by the any *probjet* ordering discrepancies. A comparison was made in the angle between the two harder jets (α_{12}) and the two softer jets (α_{34}) in data and simulation, as shown in figure 6(a). The α_{12} distribution is reasonably well reproduced, while a disagreement of about 10% is found for α_{34}. Notice instead the better agreement for the θ_{12} and θ_{34} angles, shown in figure 6(b).

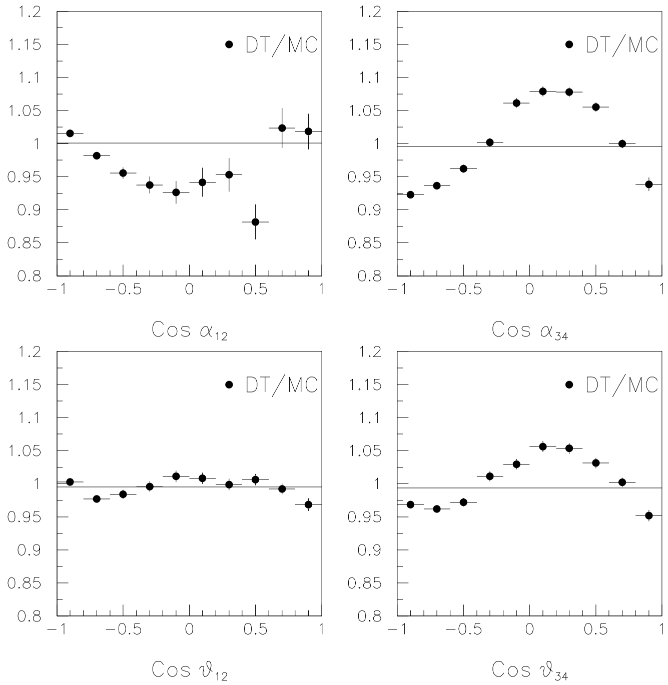

Figure 6: Angular distributions for $\cos\alpha_{12}$ and $\cos\alpha_{34}$ (upper plot) and $\cos\theta_{12}$ and $\cos\theta_{34}$ (lower plot) in the ratio Data and Monte Carlo simulation.

In order to evaluate the impact of this discrepancy in the analysis, the Monte Carlo background distribution in (α_{12},α_{34}) was forced to match the data before the cut in $\cos\theta_{12}$. The difference in the result was $\Delta g_{b\bar{b}} = 0.05 \times 10^{-3}$, and is taken as a systematic due to the topological cuts.

The simulation of the signal events is based on JETSET PS Monte Carlo, which is in good agreement with the theoretical predictions[1]. In order to estimate the uncertainty on this assumption we have produced 25,000 $g \to b\bar{b}$ events using the HERWIG 5.9 Monte Carlo[3] at the generator level.

The tagging efficiency depends on two uncorrelated variables: the invariant mass of the two B hadrons (m_g) and the decay angle in their centre-of-mass ($|\cos\theta^*|$).

The distribution of $|\cos\theta^*|$ differs substantially for the two Monte Carlo, as can be seen in figure 7, due to a different choice of the Q^2 in the gluon splitting kernel.

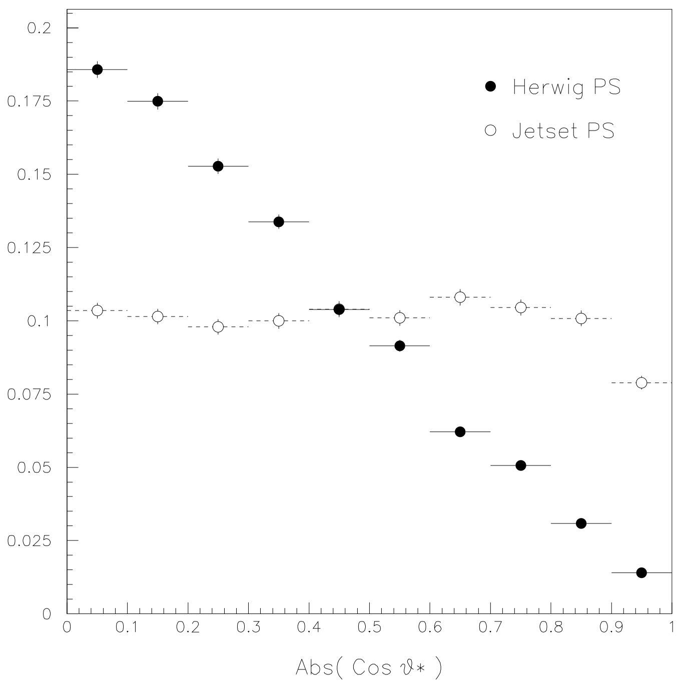

Figure 7: Distributions of $|\cos\theta^*|$ in JETSET and HERWIG signal simulations.

We then computed the differential signal efficiency in a (5×3) bins in the ($m_g, |\cos\theta^*|$) plane, $\epsilon_B(m_g, |\cos\theta^*|)$.

We have calculated the HERWIG average $\mathcal{B}$ efficiency applying the differential efficiency $\epsilon_B(m_g, |\cos\theta^*|)$ to the HERWIG spectrum. The efficiency increased by a factor 1.20 ± 0.12 with respect to JETSET. The difference in $g_{b\bar{b}}$ using this new efficiency is then taken as the systematic error on $g_{b\bar{b}}$: $\Delta g_{b\bar{b}} = \pm 0.51 \times 10^{-3}$.

The uncertainty in the rate of the $g \to c\bar{c}$ background events, ($\Delta g_{c\bar{c}} = \pm 0.48\%$), gives an error $\Delta g_{b\bar{b}} = \pm 0.17 \times 10^{-3}$.

Charged Monte Carlo tracks used by the lifetime tag have been smeared to better reproduce the data negative distribution of the impact parameter significance[12]. Uncertainties in the efficiencies due to this smearing have been assessed by evaluating the MC efficiencies without the smearing algorithm. Half of the difference in the $g_{b\bar{b}}$ result is then taken as systematic error. The error on $g_{b\bar{b}}$ due to the tracking is then estimated as $\Delta g_{b\bar{b}} = \pm 0.10 \times 10^{-3}$.

We evaluated the error coming from the correction made for R_b in $\mathcal{Q}$ background events (see page 3) by varying R_b by its total error. This gives $\Delta g_{b\bar{b}} = \pm 0.05 \times 10^{-3}$.

Finally, in order to evaluate any other error, due to possible differences in the variables used to tag the signal events, a comparison of the MC/Data fraction of events is computed as varying the different cuts.

Table 4 reports the efficiencies for the 3 categories of events, evaluated by the MC simulation, the MC and the Data fractions as well as their ratio as a function of different cuts (the MC $g_{b\bar{b}}$ rate has been set to 0.3%).

For all the cuts shown in table 4, the background reaches at maximum about 85% of the total number of selected events. Therefore discrepancies between Monte Carlo and Data test any inadequacy in the evaluation of the background efficiencies. The largest difference is taken conservatively as a systematic error on ϵ_Q. This gives $\Delta g_{b\bar{b}} = \pm 0.15 \times 10^{-3}$.

Table 5 resumes the different sources of systematic error on $g_{b\bar{b}}$.

The final result is

$$g_{b\bar{b}} = (2.57 \pm 0.40(\mathrm{stat}) \pm 0.87(\mathrm{syst})) \times 10^{-3}$$

6 Systematic checks

We have performed the following systematic checks on the measurement.

Figure 8 shows the angular distribution for the third and fourth jets in data and simulation, together with the expected composition. The contribution from the signal events cannot be replaced by any other background component.

Figure 9 shows the stability of the result as a function of the $probjet_1$ and $probjet_2$ cut, together with the uncorrelated statistical and systematic error.

We then checked the stability of the result as a function of the jet angular distributions in data and simulation, cutting in different slices of $\cos\theta_{12}$ and $\cos\theta_{34}$. Changing the cuts to $0.3 < \cos\theta_{12} < 1$ we get $\Delta g_{b\bar{b}} = (-0.37 \pm 0.35) \times$

	njet=4	$\oplus nch_i > 0$	$\oplus 0.2 < \cos\theta_{12} < 0.9$	$\oplus PROBJET_1 > 2.4$
$f_{\mathcal{B}}$ (%)	47.66 ± 0.25	27.83 ± 0.24	8.80 ± 0.16	4.31 ± 0.12
$f_{\mathcal{C}}$ (%)	28.64 ± 0.09	15.83 ± 0.07	3.74 ± 0.04	0.54 ± 0.01
$f_{\mathcal{Q}}$(%)	8.83 ± 0.01	4.57 ± 0.01	1.027 ± 0.004	0.098 ± 0.002
f_{MC} (%)	$9.42 \pm 0.01 \pm 0.10$	$4.91 \pm 0.01 \pm 0.05$	$1.115 \pm 0.005 \pm 0.013$	$0.121 \pm 0.002 \pm 0.002$
f_{DT}(%)	9.56 ± 0.02	4.98 ± 0.01	1.135 ± 0.006	0.126 ± 0.002
DT/MC	1.015 ± 0.010	1.014 ± 0.011	1.018 ± 0.013	1.041 ± 0.028

Table 4: Efficiencies for the $\mathcal{Q}$, $\mathcal{C}$ and $\mathcal{B}$ categories, as well as the total MC and Data fractions of selected events as a function of increasing cuts. The first error is statistical, the second is due to $g_{c\bar{c}}$ uncertainty.

Table 5: Systematic uncertainties on $g_{b\bar{b}}$.

Source	$\Delta g_{bb}(10^{-3})$
MC statistics	± 0.38
Beauty and Charm physics	± 0.54
$g_{c\bar{c}} = (2.38 \pm 0.48)\%$	± 0.17
JETSET – HERWIG comparison	± 0.51
Tracking uncertainty	± 0.10
Background topological cuts uncertainty	± 0.05
Data – Monte Carlo comparison	± 0.15
4-jet $R_b = 0.2145 \pm 0.0040$	± 0.05
Total	± 0.87

10^{-3}, where the error comes from uncorrelated statistics, while cutting in $-1 < \cos\theta_{34} < 0$ we get $\Delta g_{b\bar{b}} = (+0.20 \pm 0.22) \times 10^{-3}$.

Finally we repeated the analysis with a different value of the clustering parameter, $y_{cut} = 0.01$. The difference in the result is $\Delta g_{b\bar{b}} = (+0.03 \pm 0.39) \times 10^{-3}$, where the error takes into account the uncorrelated data and simulation statistics.

7 Acknowledgements

We wish to thank M. Ciafaloni, P. Nason, M. Seymour, T. Sjöstrand and B. Webber for useful discussions on some theoretical aspects of the measurement.

We also wish to thank our colleagues in the CERN accelerator divisions for the successful operation of the LEP storage ring. We also thank the engineers and technicians in all our institutions for their support in constructing and operating ALEPH. Those of us from non-member states thank CERN for its hospitality.

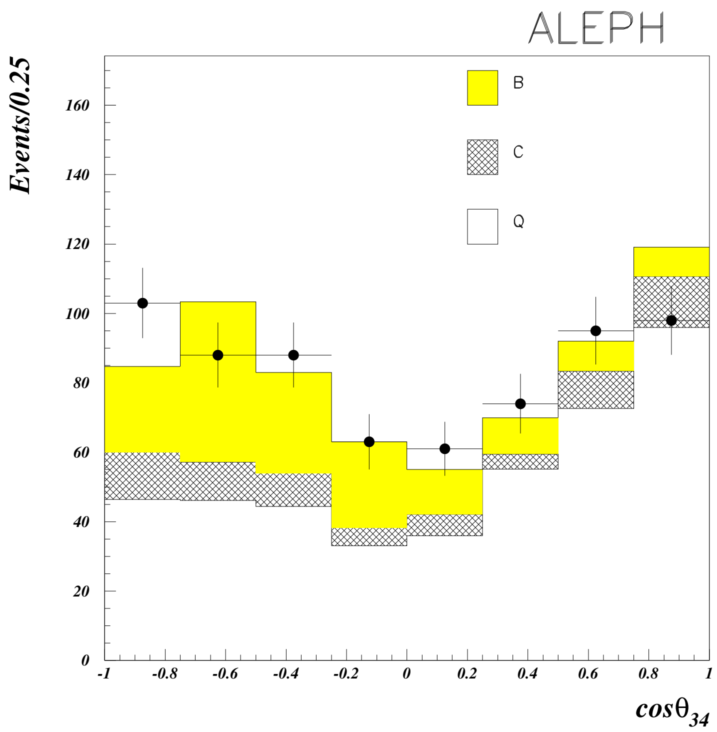

Figure 8: Angular distribution for the third and fourth jets in data and simulation, together with the expected composition. Monte Carlo events are normalised to data.

References

[1] M. H. Seymour, Nucl. Phys. **B** 436 (1995) 163

[2] T. Sjöstrand, Comp. Phys. Comm. **82** (1994) 74.

[3] G. Marchesini, B.R.Webber, G.Abbiendi, I.G.Knowles, M.H.Seymour, and L.Stanco, Comp. Phys. Comm. **67** (1992) 465.

[4] R. Akers *et al.* OPAL Collab., Phys. Lett. **B 353** (1995) 595

[5] J. Schwiening, *First Measurements on gluon splitting into heavy quarks in e^+e^- annihilation*, BONN-HE-95-03

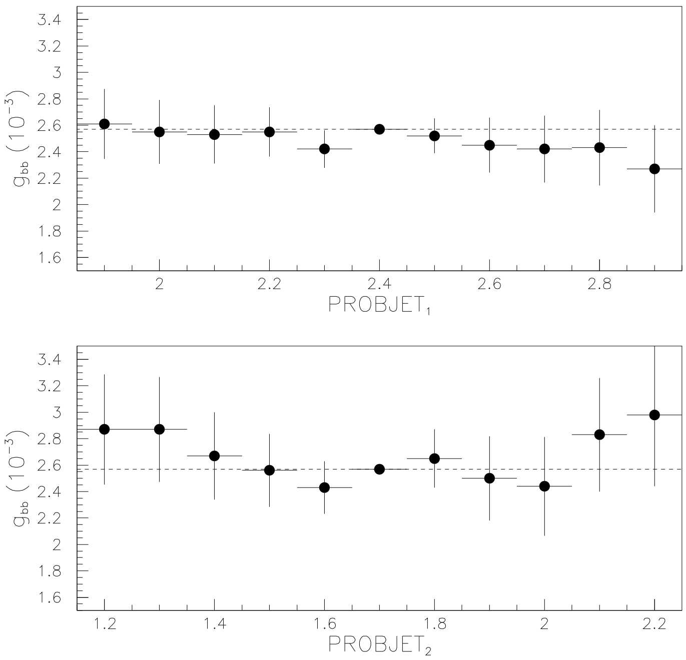

Figure 9: Stability plot of $g_{b\bar{b}}$ as a function of the $probjet_1$ and $probjet_2$ cut. The uncorrelated errors are the ones referred to the central cut at 2.4 and 1.7, respectively.

[6] D. Buskulic *et al.*, ALEPH Collab., *A Measurement of* $g \to c\bar{c}$ *using* D* *Mesons*, Contributed paper to the ICHEP96 Conference, paper PA05-065.

[7] P. Abreu *et al.* DELPHI Coll., CERN-PPE/97-39 (1997) To be submitted to Phys. Lett. **B**

[8] R. Barate *et al.* ALEPH Coll., Phys. Lett. **B 401** (1997) 163

[9] D. Buskulic *et al.*, ALEPH Collab., Z. Phys. **C60** (1993) 71.

[10] Presentation of LEP Electroweak Heavy Flavour Results for Summer 1996 Conferences, LEPHF/96-01, ALEPH note 96-099

[11] Particle Data Group, Phys. Rev. **D** 54 (1996)

[12] D. Buskulic *et al.*, ALEPH Collab., Phys. Lett. **B313** (1993) 535.