IMPROVED MEASUREMENT OF THE PROBABILITY FOR GLUON SPLITTING INTO $b\bar{b}$ IN $Z^0$ DECAYS

The SLD Collaboration

Stanford Linear Accelerator Center
Stanford University, Stanford, CA 94309

ABSTRACT

We present a preliminary measurement of the rate of gluon splitting into bottom quarks, $g \rightarrow b\bar{b}$, in hadronic $Z^0$ decays collected by SLD between 1996 and 1998. The analysis was performed by looking for secondary bottom production in 4-jet events of any primary flavor. 4-jet events were identified, and a topological vertex-mass technique was applied to each jet in order to identify $b$ or $\bar{b}$ jets. The upgraded CCD based vertex detector gives very high $B$-tagging efficiency, especially for $B$ hadrons of the low energies typical of this process. The two most nearly collinear $b/\bar{b}$ jets were tagged as originating from $g \rightarrow b\bar{b}$. We measured the rate of secondary $b/\bar{b}$ production per hadronic event, $g_{b\bar{b}}$, to be $(2.84 \pm 0.61({\text{stat.}}) \pm 0.59({\text{syst.}})) \times 10^{-3}$ (preliminary).

Contributed to the XXX International Conference on High Energy Physics, July 27–August 2 2000, Osaka, Japan; Ref. 691.

* Work supported by Department of Energy contract DE-AC03-76SF00515 (SLAC).
1 Introduction

The process of the splitting of a gluon into a heavy-quark pair is one of the elementary processes in QCD but is poorly known, both theoretically and experimentally.

The rate \( g_{b\bar{b}} \) is defined as the fraction of \( e^+e^- \rightarrow \text{hadrons} \) events in which a gluon splits into a \( b\bar{b} \) pair, \( e^+e^- \rightarrow q\bar{q}g \rightarrow q\bar{q}b\bar{b} \). The value of \( g_{b\bar{b}} \) is an infrared finite quantity, because the \( b \)-quark mass provides a natural cutoff, hence it can be safely computed in the framework of perturbative QCD \cite{1}. However \( g_{b\bar{b}} \) is sensitive to \( \alpha_s \) and to the \( b \)-quark mass, which results in a substantial theoretical uncertainty in the calculation. The limited accuracy of the \( g_{b\bar{b}} \) prediction is one of the main sources of uncertainty in the measurement of the partial decay width \( R_b = \Gamma(Z^0 \rightarrow b\bar{b})/\Gamma(Z^0 \rightarrow q\bar{q}) \) \cite{2}. In addition, about 50\% of the \( B \) hadrons produced at the Tevatron are due to the gluon splitting process, and a larger fraction is expected to contribute at the LHC. A better knowledge of this process can improve theoretical predictions of heavy-flavor production at hadron colliders.

This measurement is difficult experimentally; \( g_{b\bar{b}} \) is very small even at the \( Z^0 \) resonance, since the gluon must have sufficient mass to produce the bottom-quark pair. There are huge backgrounds from \( Z^0 \rightarrow b\bar{b} \) events, which occur \( \sim 100 \) times more frequently than the \( Z^0 \rightarrow q\bar{q}g \rightarrow q\bar{q}b\bar{b} \) process. Moreover the \( B \) hadrons from \( g \rightarrow b\bar{b} \) have relatively low energy and short flight distance and are more difficult to distinguish using standard tagging techniques. So far, measurements of \( g_{b\bar{b}} \) have been reported by DELPHI, ALEPH and OPAL \cite{3}.

Here we present an improved measurement of \( g_{b\bar{b}} \) based on a 400k \( Z^0 \)-decay data sample taken in 1996-98 at the Stanford Linear Collider (SLC), with the SLC Large Detector (SLD). In this period, \( Z^0 \) decays were collected with an upgraded vertex detector with wider acceptance and better impact parameter resolution, thus improving considerably the \( b \)-tagging performance. This measurement supercedes our previous preliminary result \cite{4}.
2 The SLD Detector

A description of the SLD is given elsewhere [5]. Only the details most relevant to this analysis are mentioned here. SLD was well-suited for the measurement of $g \rightarrow b\bar{b}$ due to two unique features. The first is that the SLC, the first linear collider, provided a very small and stable beam spot. The SLC interaction point was reconstructed from tracks in sets of approximately thirty sequential hadronic $Z^0$ decays with an uncertainty of only 3.2$\mu m$ transverse to the beam axis and 17$\mu m$ (for $b\bar{b}$ events) along the beam axis. Second is the upgraded vertex detector (VXD3) [6], a pixel-based CCD vertex detector. VXD3 consists of 3 layers, each layer being only 0.36% of a radiation length thick, with a total of 307M pixels. The measured $r_\phi$ ($rz$) track impact-parameter resolution approaches 7.7$\mu m$ (9.6$\mu m$) for high momentum tracks, while multiple scattering contributions are $29\mu m/(p\sin^{3/2}\theta)$ in both projections ($z$ is the coordinate parallel to the beam axis) [7]. With these features, topological vertex finding [8] gives excellent $b$-tagging efficiency and purity. In particular, the efficiency is good even at low $B$-hadron energies, which is especially important for detecting $g \rightarrow b\bar{b}$.

The measurement used 400k $Z^0 \rightarrow$ hadron events collected between 1996 and 1998 with the requirement that VXD3 was fully operational.

3 Flavor Tagging

Topologically reconstructed secondary vertices [8] are used in many analyses at the SLD for heavy-quark tagging. To reconstruct the secondary vertices, the space points where track density functions overlap are found in 3-dimensions. Only the vertices that are significantly displaced from the primary vertex (PV) are considered to be possible $B$- or $D$-hadron decay vertices. The mass of the secondary vertex is calculated using the tracks that are associated with the vertex. Since the heavy-hadron decays
are frequently accompanied by neutral particles, the reconstructed mass is corrected to account for this fact. By using kinematic information from the vertex flight path and the momentum sum of the tracks associated with the secondary vertex, we calculate the $P_T$-corrected mass $M_{P_T}$ by adding a component of missing momentum to the invariant mass, as follows:

$$M_{P_T} = \sqrt{M_{VTX}^2 + P_T^2 + |P_T|}.$$

Here $M_{VTX}$ is the invariant mass of the tracks associated with the reconstructed secondary vertex and $P_T$ is the transverse momentum of the charged tracks with respect to the B flight direction. In this correction, vertexing resolution as well as the PV resolution are crucial. Due to the small and stable interaction point at the SLC and the excellent vertexing resolution from VXD3, this technique has so far only been successfully applied at the SLD.

4 Monte Carlo and data Samples

For the purpose of estimating the efficiency and purity of the $g \rightarrow b \bar{b}$ selection procedure, we made use of a detailed Monte-Carlo simulation of the detector. The JETSET 7.4 [9] event generator was used, with parameter values tuned to hadronic $e^+e^-$ annihilation data [10], combined with a simulation of B hadron decays tuned to $\Upsilon(4S)$ data [11] and a simulation of the SLD based on GEANT 3.21 [12]. Inclusive distributions of single-particle and event-topology observables in hadronic events were found to be well described by the simulations [13]. Uncertainties in the simulation were taken into account in the systematic errors (Section 7).

Monte-Carlo events were reweighted to take into account current estimates for gluon splitting into heavy-quark pairs [3, 14]. JETSET with the SLD parameters predicts $g_{b\bar{b}} = 0.14\%$ and $g_{c\bar{c}} = 1.36\%$, and we reweighted them so that $g_{b\bar{b}} = 0.247\%$ and $g_{c\bar{c}} = 3.07\%$ [15]. Monte-Carlo samples of about 1900k $Z \rightarrow q\bar{q}$ events, 1900k $Z \rightarrow b\bar{b}$
events, 1090k $Z \rightarrow c\bar{c}$ events and 60k $g \rightarrow b\bar{b}$ events were used in order to evaluate the efficiencies.

Besides the signal events, hereafter called ‘B events’, two categories of background events exist:

- Events which do not contain any gluon splitting into heavy quarks at all, hereafter called ‘Q events’; and

- Events in which a gluon splits to a charm quark pair, called ‘C events’.

5 Analysis

The two B hadrons coming from the gluon tend to be produced in a particular topological configuration, which allows one to discriminate the signal from background. We select $g \rightarrow b\bar{b}$ events as follows:

- Require 4 jets in the events;

- Require $b$ tags in two jets selected in a particular configuration; and

- Apply additional topological selections to improve the signal/background ratio.

Jets are formed by applying the Durham jet-finding algorithm [16] to charged tracks with $y_{\text{cut}} = 0.009$, chosen to minimize the statistical error. The overall 4-jet rate in the data is $(6.631 \pm 0.046)\%$, where the error is statistical only. In the Monte-Carlo simulation the 4-jet rate is $(6.173 \pm 0.017 \pm 0.065)\%$ where the first error is statistical and the second is due to the uncertainty in the simulation of heavy-quark physics. The 4-jet rates for the B, C and Q events predicted by the simulation are about 36%, 19% and 5.6%, respectively. The two jets forming the smallest angle in the event are considered as candidates for originating from the gluon splitting process $g \rightarrow b\bar{b}$. The
selected jets are labeled as jet 1 and jet 2, where jet 1 is more energetic than jet 2. The other two jets in the event are labeled as jets 3 and 4, where jet 3 is more energetic than jet 4.

Jets containing B-hadron decay products are then searched for by making use of the information coming from the vertex detector, using the topological vertex method. We required both jet 1 and jet 2 to contain a secondary vertex with a 3D decay length greater than 300µm. No tag was applied to jet 3 and jet 4. After topological vertexing, 547 events were selected. The selection efficiency for $g \rightarrow b\bar{b}$ is expected from the Monte Carlo simulation to be 9.7% while the signal/background ratio is $\sim 1/5$. 63% of the background comes from $Z \rightarrow b\bar{b}$ events, 24% from $g \rightarrow c\bar{c}$ events and remaining 13% from $Z \rightarrow q\bar{q} (q \neq b)$ events.

In order to improve the signal/background ratio, we used a neural network technique. We chose the following 9 observables as inputs to the neural network; each observable was scaled to correspond to a range between 0 and 1.

1. The larger of the $P_T$-corrected mass of jets 1 and 2. $b$ jets have higher $P_T$-corrected mass ($M_{P_T}$) than $c/uds$ jets (Figure 1).

2. The smaller of the $P_T$-corrected mass of jets 1 and 2 (Figure 2).

3. The angle $\theta_{12}$ between the vertex axes of jets 1 and 2 (Figure 3). Many $Z^0 \rightarrow b\bar{b}$ background events have one $b$-jet which was split by the jet-finder into 2 jets so that the two found vertices are from different decay products from the same B decay. The two vertex axes tend to be collinear.

4. The angle $\theta_{34}$ between the jet axes of jets 3 and 4 (Figure 4). This tends to be distributed at large values, while background events populate the small angle region.

5. $15M_{P_T} - P_{VTX}$ for the larger of the $P_T$-corrected masses of jets 1 and 2, where $P_{VTX}$ is the vertex momentum (Figure 5). This observable tends to be large
for $b$ jets since $B$ decay vertices typically have higher mass than those from charm decays, and vertices resulting from $B \rightarrow D$ cascade decays have a lower momentum than those from primary $D$ hadrons.

6. $15M_{P_T} - P_{VTX}$ for the smaller of the $P_T$-corrected masses of jets 1 and 2 (Figure 6). The observable also has nice discrimination power between signal and background events.

7. The energy sum of jets 1 and 2 (Figure 7). The $b$ jets coming from a gluon tend to have lower energy than the other two jets in the event.

8. The energy sum of jets 3 and 4 (Figure 8).

9. The cosine of the angle $\alpha_{1234}$ between the plane $\Pi_{12}$ formed by jets 1 and 2 and the plane $\Pi_{34}$ formed by jets 3 and 4 (Figure 9). This variable is similar to the Bengtsson-Zerwas angle [17], and is useful to separate $g \rightarrow b\bar{b}$ events because the radiated virtual gluon in the process $Z^0 \rightarrow q\bar{q}g$ is polarized in the plane of the three-parton event, and this is reflected in its subsequent splitting, by strongly favoring $g \rightarrow q\bar{q}$ emission out of this plane.

Data and MC agree well for these input observables. We trained the neural network using Monte-Carlo samples of about 1800k $Z \rightarrow q\bar{q}$ events, 1200k $Z \rightarrow b\bar{b}$ events, 780k $Z \rightarrow c\bar{c}$ events and 50k $g \rightarrow b\bar{b}$ events. These samples were independent of the ones used for the selection efficiency and background studies. Figure 10 shows the output distribution of the neural network. We retained events for which the output variable was greater than 0.6.

6 Result

After requiring all the above mentioned cuts, 79 events were selected in the data. The number of background events was estimated, using the Monte Carlo simulation, to be
<table>
<thead>
<tr>
<th>Source</th>
<th>Efficiency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>4.99 ± 0.10</td>
</tr>
<tr>
<td>C</td>
<td>0.171 ± 0.017</td>
</tr>
<tr>
<td>Q</td>
<td>0.0081 ± 0.0004</td>
</tr>
</tbody>
</table>

Table 1: Efficiencies after all cuts for the three categories. Errors are statistical only.

37.8, where 56% of the background comes from \( Z \rightarrow b\bar{b} \) events, 40% from \( g \rightarrow c\bar{c} \) events and the remaining 4% from \( Z \rightarrow q\bar{q} \) \((q \neq b)\) events. Table 1 shows the tagging efficiencies for the three categories of events, where the errors are statistical only. From these efficiencies and the fraction of events selected in the data \( f_d = (2.72 ± 0.30) \times 10^{-1} \), the value of \( g_{dF} \) can be determined:

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

We obtain

\[
g_{b\bar{b}} = (2.84 ± 0.61) \times 10^{-3},
\]

where the error is statistical only.

7 Systematic Errors

The efficiencies for the three event categories were evaluated using the Monte-Carlo simulation. The limitations of the simulation in estimating these efficiencies lead to an uncertainty on the result. The error due to limited Monte-Carlo statistics in the efficiency evaluation was \( \Delta g_{dF} = ±0.14 \times 10^{-3} \).

A large fraction of events remaining after the selection cuts contain \( B \) and \( D \) hadrons. The uncertainty in the knowledge of the physical processes in the simulation of heavy-flavor production and decays constitutes a source of systematic error. All the physical simulation parameters were varied within their allowed experimental
ranges. In particular, the B and D hadron lifetimes as well as their production rates were varied, following the recommendations of the LEP Heavy Flavour Working Group [18]. The uncertainties are summarized in Table 2.

The simulation of the signal events is based on the JETSET parton shower Monte Carlo, which is in good agreement with the theoretical predictions [1]. In order to estimate the uncertainty on this assumption, we have produced 50,000 $g \rightarrow b\bar{b}$ events using HERWIG [19] at the generator level. The signal tagging efficiency, $\epsilon_B$, mainly depends on the description of the split gluon: its energy $E_g$, its mass $m_g$ and the decay angle, $\theta^*$, of the two B hadrons in their center-of-mass frame relative to the gluon direction. This efficiency function, computed with JETSET, is reweighted by the ratio of the HERWIG to JETSET initial distributions to obtain the average efficiency. A systematic error of $\Delta g_{b\bar{b}} = \pm 0.49 \times 10^{-3}$ is estimated from the difference in efficiency between the two Monte-Carlo models.

The dependence of the efficiency on the $b$-quark mass has also been investigated at the generator level. Events are generated using the GRC4F Monte Carlo [20], which is based on a matrix element calculation including $b$-quark masses. The variation of $\epsilon_B$ was computed for $b$-quark masses between 4.7 and 5.3 GeV/$c^2$. The resultant uncertainty is estimated to be $\Delta g_{b\bar{b}} = \pm 0.10 \times 10^{-3}$.

The uncertainty in the production fraction of $g \rightarrow c\bar{c}$ background events, $\Delta g_{c\bar{c}} = \pm 0.40\%$, gives an error $\Delta g_{b\bar{b}} = \pm 0.13 \times 10^{-3}$.

There is about a 7% discrepancy in the 4-jet rate between data and Monte Carlo at our $y_{cut}$ value. The uncertainty due to the discrepancy is estimated by increasing the number of $Z \rightarrow q\bar{q}$ background events in the Monte Carlo and is found to be $\Delta g_{b\bar{b}} = \pm 0.12 \times 10^{-3}$.

In the Monte-Carlo simulation charged tracks used in the topological vertex tag were rejected to reproduce better the distributions in the data. Uncertainties in the efficiencies due to this rejection were assessed by evaluating the Monte-Carlo efficiencies without the rejection algorithm. The difference in the $g_{b\bar{b}}$ result was taken as a
Monte Carlo statistics $\pm 0.14$

$B$ hadron lifetimes $\pm 0.02$

$B$ hadron production $\pm 0.02$

$B$ hadron fragmentation $\pm 0.13$

$B$ hadron charged multiplicities $\pm 0.03$

$D$ hadron lifetimes $\pm 0.01$

$D$ hadron production $\pm 0.02$

$D$ hadron charged multiplicities $\pm 0.02$

Energy distribution of $g \rightarrow b\bar{b}$ $\pm 0.49$

$b$ quark mass $\pm 0.10$

$g_{e\bar{e}}$ $\pm 0.13$

4-jet rate discrepancy $\pm 0.12$

Tracking efficiency $\pm 0.17$

Total (Preliminary) $\pm 0.59$

<table>
<thead>
<tr>
<th>Source</th>
<th>$\Delta g_{b\bar{b}} (10^{-3})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monte Carlo statistics</td>
<td>$\pm 0.14$</td>
</tr>
<tr>
<td>$B$ hadron lifetimes</td>
<td>$\pm 0.02$</td>
</tr>
<tr>
<td>$B$ hadron production</td>
<td>$\pm 0.02$</td>
</tr>
<tr>
<td>$B$ hadron fragmentation</td>
<td>$\pm 0.13$</td>
</tr>
<tr>
<td>$B$ hadron charged multiplicities</td>
<td>$\pm 0.03$</td>
</tr>
<tr>
<td>$D$ hadron lifetimes</td>
<td>$\pm 0.01$</td>
</tr>
<tr>
<td>$D$ hadron production</td>
<td>$\pm 0.02$</td>
</tr>
<tr>
<td>$D$ hadron charged multiplicities</td>
<td>$\pm 0.02$</td>
</tr>
<tr>
<td>Energy distribution of $g \rightarrow b\bar{b}$</td>
<td>$\pm 0.49$</td>
</tr>
<tr>
<td>$b$ quark mass</td>
<td>$\pm 0.10$</td>
</tr>
<tr>
<td>$g_{e\bar{e}}$</td>
<td>$\pm 0.13$</td>
</tr>
<tr>
<td>4-jet rate discrepancy</td>
<td>$\pm 0.12$</td>
</tr>
<tr>
<td>Tracking efficiency</td>
<td>$\pm 0.17$</td>
</tr>
<tr>
<td>Total (Preliminary)</td>
<td>$\pm 0.59$</td>
</tr>
</tbody>
</table>

Table 2: Systematic uncertainties on $g_{b\bar{b}}$.

symmetric systematic error, $\Delta g_{b\bar{b}} = \pm 0.17 \times 10^{-3}$.

Table 2 summarizes the different sources of systematic error on $g_{b\bar{b}}$. The total systematic error is estimated to be the sum in quadrature, $0.59 \times 10^{-3}$.

8 Summary

A preliminary measurement of the gluon splitting rate to a $b\bar{b}$ pair in hadronic $Z^0$ decays collected by SLD has been presented. Excellent SLC and VXD3 performance provides advantages not only for the $b$-tag efficiency but also for the topological
selections. The result is

\[ g_{bb} = (2.84 \pm 0.61\text{(stat.)} \pm 0.59\text{(syst.)}) \times 10^{-3}\text{(preliminary)}. \]

where the first error is statistical and the second includes systematic effects.

**Acknowledgments**

We thank Mike Seymour for helpful conversations. We thank the personnel of the SLAC accelerator department and the technical staffs of our collaborating institutions for their outstanding efforts on our behalf. This work was supported by the U.S. Department of Energy and National Science Foundation, the UK Particle Physics and Astronomy Research Council, the Istituto Nazionale di Fisica Nucleare of Italy, the Japan-US Cooperative Research Project on High Energy Physics, and the Korea Research Foundation (Soongsil, 1997).

**List of Authors**


(1) Aomori University, Aomori, 030 Japan
(2) University of Bristol, Bristol, United Kingdom
(3) Brunel University, Uxbridge, Middlesex, UB8 3PH United Kingdom
(4) Boston University, Boston, Massachusetts 02215
(5) University of Colorado, Boulder, Colorado 80309
(6) Colorado State University, Ft. Collins, Colorado 80523
(7) INFN Sezione di Ferrara and Universita di Ferrara, I-44100 Ferrara, Italy
(8) INFN Laboratori Nazionali di Frascati, I-00044 Frascati, Italy
(9) Johns Hopkins University, Baltimore, Maryland 21218-2686
(10) Lawrence Berkeley Laboratory, University of California, Berkeley, California 94720
(11) University of Massachusetts, Amherst, Massachusetts 01003
(12) University of Mississippi, University, Mississippi 38677
(13) Massachusetts Institute of Technology, Cambridge, Massachusetts 02139
(14) Institute of Nuclear Physics, Moscow State University, 119899 Moscow, Russia
(15) Nagoya University, Chikusa-ku, Nagoya, 464 Japan
(16) University of Oregon, Eugene, Oregon 97403
(17) Oxford University, Oxford, OX1 3RH, United Kingdom
(18) INFN Sezione di Perugia and Universita di Perugia, I-06100 Perugia, Italy
(19) Rutherford Appleton Laboratory, Chilton, Didcot, Oxon OX11 0QX United Kingdom
(20) Rutgers University, Piscataway, New Jersey 08855
(21) Stanford Linear Accelerator Center, Stanford University, Stanford, California 94309
(22) Soongsil University, Seoul, Korea 156-743
(23) University of Tennessee, Knoxville, Tennessee 37996
References


R. Barate et al. [ALEPH Collaboration], Phys. Lett. B401, 163 (1997);
P. Abreu et al. [DELPHI Collaboration], Z. Phys. C70, 531 (1996);
P. Abreu et al. [DELPHI Collaboration], Z. Phys. C66, 323 (1995);
O. Adriani et al. [L3 Collaboration], Phys. Lett. B307, 237 (1993);
G. Abbiendi et al. [OPAL Collaboration], Eur. Phys. J. C8, 217 (1999);
K. Abe et al. [SLD Collaboration], Phys. Rev. Lett. 80, 660 (1998);

P. Abreu et al. [DELPHI Collaboration], Phys. Lett. B462, 425 (1999);
P. Abreu et al. [DELPHI Collaboration], Phys. Lett. B405, 202 (1997);
G. Abbiendi et al. [OPAL Collaboration], OPAL note PN383.


    M. Acciarri et al. [L3 Collaboration], Phys. Lett. B476, 243 (2000);


Figure 1: Distribution of the larger of the $P_T$-corrected mass of jets 1 and 2. Points indicate data, open box simulation, grey boxes are signal.

Figure 2: Distribution of the smaller of the $P_T$-corrected mass of jets 1 and 2. Points indicate data, open box simulation, grey boxes are signal.
Figure 3: Distribution of the angle between the vertex axes of jets 1 and 2 (points). The simulated distribution is shown as a histogram.

Figure 4: Distribution of the angle between the vertex axes of jets 3 and 4 (points). The simulated distribution is shown as a histogram.
Figure 5: Distribution of $15 M_{P_T} - P_{VTX}$ for the larger of the $P_T$-corrected masses of jets 1 and 2. Points indicate data, open box simulation, grey boxes are signal.

Figure 6: Distribution of $15 M_{P_T} - P_{VTX}$ for the smaller of the $P_T$-corrected masses of jets 1 and 2. Points indicate data, open box simulation, grey boxes are signal.
Figure 7: The distribution of the energy sum of jets 1 and 2.

Figure 8: The distribution of the energy sum of jets 3 and 4.
Figure 9: The distribution of the cosine of the angle between the plane $\Pi_{12}$ formed by jets 1 and 2 and the plane $\Pi_{34}$ formed by jets 3 and 4, for data (points) and Monte Carlo simulation (histogram).

Figure 10: The distribution of the neural network output for data (points) and Monte Carlo simulation (histogram).