Evaluation of measurement accuracies of the Higgs boson branching fractions in the International Linear Collider

H. Ono\(^1\)\(^*\) and A. Miyamoto\(^2\)

\(^1\)Nippon Dental University School of Life Dentistry at Niigata
\(^2\)High Energy Accelerator Research Organization

(Dated: May 2, 2014)

Precise measurement of Higgs boson couplings is an important task for International Linear Collider (ILC) experiments and will facilitate the understanding of the particle mass generation mechanism. In this study, the measurement accuracies of the Higgs boson branching fractions to the \(b\) and \(c\) quarks and gluons, \(\Delta Br(H \to bb, c\bar{c}, gg)/Br\), were evaluated with the full International Large Detector model (ILD) for the Higgs mass of 120 GeV at the center-of-mass (CM) energies of 250 and 350 GeV using neutrino, hadronic and leptonic channels and assuming an integrated luminosity of 250 \(fb^{-1}\), and an electron (positron) beam polarization of \(-80\%\) \((+30\%)\). We obtained the following measurement accuracies of the Higgs cross section times branching fraction \((\Delta \sigma \cdot Br)/\sigma \cdot Br\) for decay of the Higgs into \(bb\), \(c\bar{c}\), and \(gg\); as 1.0\%, 6.9\%, and 8.5\% at a CM energy of 250 GeV and 1.0\%, 6.2\%, and 7.3\% at 350 GeV, respectively. After the measurement accuracy of the cross section \((\Delta \sigma/\sigma)\) was corrected using the results of studies at 250 GeV and their extrapolation to 350 GeV, the derived measurement accuracies of the branching fractions \((\Delta Br/Br)\) to \(bb\), \(c\bar{c}\), and \(gg\) were 2.7\%, 7.3\%, and 8.9\% at a CM energy of 250 GeV and 3.6\%, 7.2\%, and 8.1\% at 350 GeV, respectively.

PACS numbers: 13.66.Fg, 29.20.Ej

Keywords: ILC, Higgs boson, Branching ratio

I. INTRODUCTION

Precise measurement of the Higgs boson branching ratios (BRs) is an important task for the International Linear Collider (ILC) program. It is also crucial for the understanding of the nature of electro-weak symmetry breaking and provides a window to investigate physics beyond the standard model (SM). The relatively low background and well-defined initial state of the ILC experiments allow precise, model-independent study of the Higgs boson, which is not an easy task for Large

\(^*\)Electronic address: ono@ngt.ndu.ac.jp
Hadron Collider experiments [1, 2]. Measurements of the Higgs BRs to $b\bar{b}$ and $c\bar{c}$ decays at an $e^{+}e^{-}$ linear collider were reported in Refs. [3–7]. In this study, we investigate the accuracies of BRs of the Higgs to $b\bar{b}$, $c\bar{c}$, and $gg$ using Geant4 [8] based realistic simulation implemented with a proposed International Large Detector (ILD) [9].

In this study, we assume a Higgs mass of 120 GeV/c$^2$ and an integrated luminosity of 250 fb$^{-1}$, and estimate the accuracies of the BRs at center-of-mass (CM) energies of 250 and 350 GeV. The former value is close to the threshold of Higgs production and thus is considered as initial target of ILC experiments. The latter is close to the threshold of top quark pair production; therefore, Higgs data can be corrected simultaneously with a top threshold study. The difference between kinematical conditions at 250 and 350 GeV could yield different detection efficiencies and thus different BR accuracies. The accuracies at 250 and 350 GeV under the same conditions are studied and compared.

The experimental conditions for this study are described in section II. We selected the Higgs events in three channels: neutrino, hadronic and leptonic. The event selection and background suppression processes are described in the section III. The derivation of the BRs is presented in the section IV, and the conclusion is given in the last section.

II. HIGGS PHYSICS IN THE ILC EXPERIMENT

A. ILC experiment and Higgs production

The ILC is a future electron-positron ($e^{-}e^{+}$) linear collider for experiments at an initial center-of-mass (CM) energy ($\sqrt{s}$) up to 500 GeV, which can be extended to 1 TeV. The production cross section of the Higgs boson is shown in Fig. 1(a) as a function of the CM energy for a Higgs mass of 120 GeV. At a low CM energy, the Higgs boson is produced primarily through the Higgs-strahlung $e^{+}e^{-} \rightarrow ZH$ process, which has a maximum around 250 GeV when the effect of the initial state radiation is considered. This is about 20 GeV higher than that without the initial state radiation. At the $\sqrt{s} = 350$ GeV, the total cross section is reduced, although the contribution of W/Z fusion is greater than that at 250 GeV. The decay BRs of the Higgs boson in the SM are shown as a function of its mass in Fig. 1(b). The Higgs decays mainly to $b\bar{b}$ if its mass is below 140 GeV and to $WW^*$ in the case of a mass of above 140 GeV.

Higgs analysis modes are categorized in terms of the three $Z$ boson decay channels: $Z \rightarrow \nu\bar{\nu}$ (neutrino), $q\bar{q}$ (hadronic), and $\ell^{+}\ell^{-}$ (leptonic), as shown in Fig. 2. We assumed the $-80\%$ and
FIG. 1: (a) Production cross section of Higgs boson at a Higgs mass of 120 GeV through the Higgsstrahlung ($ZH$) (solid) and all $ffH$ (dashed) processes assuming the -80% electron and +30% positron beam polarization. The cross section is calculated including the initial state radiation by Whizard [10]. (b) SM Higgs branching fractions as a function of Higgs mass with PYTHIA [11].

+30% polarization of the initial electrons and positrons, respectively, in order to enhance the Higgs signals.

FIG. 2: Higgs boson production diagrams categorized according to the final states: (a) neutrino ($\nu\bar{\nu}H$), (b) hadronic ($qqH$), and (c) leptonic ($\ell^+\ell^-H$) channels. Each channel is produced mainly through the Higgs-strahlung ($ZH$) process at low CM energies, although the neutrino and leptonic channels also include the $WW$ and $ZZ$ fusion processes, respectively.
B. ILD concept

We used the ILD [9] model for this study. The ILD, which is the validated detector concept for the ILC, is equipped with a highly segmented calorimeter and a hybrid tracking system consisting of gaseous, silicon-strip, and silicon-pixel trackers. They provide an excellent jet energy resolution by particle flow analysis, as well as excellent momentum resolution and vertex flavor tagging capability, which are necessary for measuring multi-jet final states in the ILC energy region. All sub-detector components of the ILD are shown in Fig. 3; which consists of silicon-pixel vertex detectors (VTX), silicon inner and outer detectors (SIT, SET), a time projection chamber (TPC), high-granularity electromagnetic and hadron calorimeters (ECAL, HCAL), a super-conducting solenoid magnet with a 3.5 T magnetic field, and an iron return yoke with a muon detector. In addition, forward silicon trackers (FTD, ETD) and beam/luminosity calorimeters (LCAL, LHCAL and BCAL) are installed in the forward region.

![Fig. 3: Schematic view of the ILD detector.](image)

The VTX system consists of three double layers of silicon pixel sensors with a 2.8 \( \mu \text{m} \) point resolution located at radii between 16 mm and 60 mm, the total radiation length being 0.74%. The impact parameter resolution \( (\sigma_{IP}) \) of the VTX system is 5 \( \mu \text{m} \oplus 10 \mu \text{m} \cdot \text{GeV/c/p}\sin^{3/2}{\theta} \). The TPC occupies a volume up to a radius of 1.8 m and a half-length in Z of 2.3 m, providing a stand-alone momentum resolution of \( \sigma_{1/\text{P}_T} \sim 9 \times 10^{-5} \text{ GeV}^{-1} \). The SIT and SET are placed at the inner and outer sides of the TPC with 7 and 50 \( \mu \text{m} \) point resolutions in the \( R - \phi \) and \( z \) directions, respectively. The overall momentum resolution of the tracking system \( (\sigma_{1/\text{P}_T}) \) is
$2 \times 10^{-5}$ GeV$^{-1} \oplus 1 \times 10^{-3}/P_T \sin \theta$ for the momentum range 1-200 GeV [9]. The ECAL consists of 24 $X_0$ tungsten absorbers with highly segmented ($5 \times 5$ mm$^2$) readouts. The HCAL consists of 5.5 $\lambda_I$ steel absorbers with a $3 \times 3$ mm$^2$ scintillator tile readout. With the ILD particle flow algorithm package, PandoraPFA [12], a dijet energy resolution of $25\%/\sqrt{E \text{ (GeV)}}$ has been achieved for a 45-GeV dijet, which corresponds to a single-jet energy resolution of $\sigma_{E_j}/E_j = 3.7\%$ [9].

C. Analysis framework and Monte Carlo samples

Monte Carlo (MC) generator samples for the physics study were produced using the Whizard [10], and fragmentation and hadronization processes were simulated by PYTHIA [11]. The SM Higgs branching fractions in PYTHIA are 65.7%, 3.6%, and 5.5% for $b\bar{b}$, $c\bar{c}$, and $gg$, respectively. The generated particles were passed through the Geant4 [8] based detector simulator Mokka [13] with the ILD model. The simulated hits were digitized and then reconstructed by the MarlinReco package; then, the resulting skinned data were analyzed. The statistics of the simulated Higgs signal samples were 500 fb$^{-1}$ for both CM energies of both 250 and 350 GeV, whereas those for background processes varied with the signal-to-noise ratio (S/N). They are scaled in the analysis in order to obtain results corresponding to an integrated luminosity of 250 fb$^{-1}$. The major SM background processes for the $e^+e^- \rightarrow ZH$ analysis are $e^+e^- \rightarrow ZZ$ and $W^+W^-$; thus we considered final sample states of $\nu\bar{\nu}q\bar{q}$, $\nu\ell\bar{q}$, $\ell\ell\ell\ell$, $q\bar{q}q\bar{q}$ and $\ell\ell\ell\ell$. In addition, the $q\bar{q}$ and $t\bar{t}$ backgrounds were also considered for the neutrino and hadronic channels (but only for $\sqrt{s} = 350$ GeV because we used a top mass of 174.9 GeV/c$^2$). In the leptonic channel, most of the multi-jet backgrounds are well suppressed if dilepton identification is required; thus, only the $\ell\ell\bar{q}\bar{q}$ and $\nu\ell\bar{q}q$ backgrounds were considered. We used the 250-GeV samples produced for the ILD letter of intent (LOI) studies [9]; thus, their beam parameters correspond to those defined in the ILC Reference Design Report [3]. On the other hand, the 350-GeV samples were newly produced for this study using the updated beam parameter SB2009 [14]. The instantaneous luminosities were 0.75 and $1 \times (10^{34}$ cm$^{-2}$s$^{-1}$) for 250 and 350 GeV, which yield integrated luminosities of 188 and 250 fb$^{-1}$, respectively, for about 3 years at 100 days of operation per year.

III. EVENT RECONSTRUCTION AND BACKGROUND SUPPRESSION

Depending on the $Z$ decay mode, the analysis channels are categorized as the neutrino (dijet), hadronic (four-jets) and leptonic channels (dileptons + dijets), which are described in the following
subsections.

A. Neutrino channel ($\nu\bar{\nu}H$)

For neutrino channel analysis, particles in the event are first forcibly clustered into two jets by the Durham jet-finding algorithm. After the dijet clustering, background reductions are applied according to the selection criteria in Table I. At a CM energy of 250 GeV, the Higgs is produced almost at rest because it is close to the production threshold, whereas it is boosted at 350 GeV. Thus, the cut conditions are optimized to obtain the best S/N at each energy. In this channel, $Z$ boson decays invisibly ($\nu\bar{\nu}$); thus, the $\nu\bar{\nu}q\bar{q}$ and $\nu\ellqq$ processes in the SM are the main backgrounds.

To reduce them, a cut on the missing mass ($M_{\text{miss}}$) is applied. Although this cut decreases the Higgs signal from the WW fusion process, the $\nu\ellqq$, $\ell\ellq\bar{q}$ and $q\bar{q}q\bar{q}$ backgrounds are effectively reduced. $q\bar{q}$ background is reduced by the following kinematical cuts: the transverse momentum ($P_t$), longitudinal momentum ($P_l$), and maximum momentum ($P_{\text{max}}$). The $\ell\ell\ell\ell$ background is well reduced by a cut on the number of charged tracks in an event ($N_{\text{chd}}$). In addition, the $\nu\ellqq$ background reduction is improved by the $Y_{12}$ and $Y_{23}$ cuts. $Y_{12}$ and $Y_{23}$ are the maximum and the minimum of $y$ values (scaled jet masses), respectively, required to cluster the event into two jets.

The background reductions for each cut are summarized in Table I for each CM energy. After all selection criteria are met, an additional likelihood ratio ($LR$) cut is applied to improve the background reduction. The $LR$ is defined using the following variables: $M_{\text{miss}}$, the number of particles ($N_{\text{PFO}}$), $Y_{12}$, $P_{\text{max}}$, $P_t$, and $M_{jj}$. The likelihood cut positions are optimized to maximize signal significance and $LR > 0.165$ and $LR > 0.395$ are selected for CM energies of 250 and 350 GeV, respectively. The signal significance ($S/\sqrt{S+B}$) after all background reductions is also listed in Table with its efficiency, where $S$ and $B$ are the numbers of Higgs signal and background entries, respectively, after all cuts are applied. The remaining backgrounds are $\nu\ellqq$ (60%), $\nu\bar{\nu}q\bar{q}$ (20%), and $q\bar{q}$ (10%) at both 250 and 350 GeV.

B. Hadronic channel ($q\bar{q}H$)

For hadronic channel analysis, particles in the event are first forcibly clustered into four jets. Next, a Higgs and $Z$ candidate dijet pair that minimize the following $\chi^2$ formula are selected from the four jets:

$$\chi^2 = \left(\frac{M_{jj2} - M_Z}{\sigma_Z}\right)^2 + \left(\frac{M_{jj4} - M_H}{\sigma_H}\right)^2,$$

\[1\]
TABLE I: Summary of the $\nu\bar{\nu}H$ channel background reduction assuming $\mathcal{L} = 250$ fb$^{-1}$ with $P(e^-, e^+) = (-0.8, +0.3)$.

| CM energy (GeV) | Cut names | Condition | Sig. | Bkg. | Condition | Sig. | Bkg. |
|-----------------|-----------|-----------|------|------|-----------|------|------|
|                 | Generated |           | 19360 | 44827100 |           | 26307 | 20855900 |
|                 | Missing mass (GeV) | $80 < M_{miss} < 140$ | 15466 | 6214050 | $50 < M_{miss} < 240$ | 23292 | 5627040 |
|                 | Transverse $P$ (GeV) | $20 < P_T < 70$ | 13727 | 549340 | $10 < P_T < 140$ | 22648 | 2271090 |
|                 | Longitudinal $P$ (GeV) | $|P_L| < 60$ | 13342 | 392401 | $|P_L| < 130$ | 22459 | 2051010 |
|                 | # of charged tracks | $N_{chd} > 10$ | 12936 | 374877 | $N_{chd} > 10$ | 21270 | 1936220 |
|                 | Maximum $P$ (GeV) | $P_{max} < 30$ | 11743 | 205038 | $P_{max} < 60$ | 20556 | 1167050 |
|                 | $Y_{23}$ value | $Y_{23} < 0.02$ | 7775 | 74439 | $Y_{23} < 0.02$ | 14992 | 465461 |
|                 | $Y_{12}$ value | $0.2 < Y_{12} < 0.8$ | 7438 | 62584 | $0.2 < Y_{12} < 0.8$ | 14500 | 413762 |
|                 | Di-jet mass (GeV) | $100 < M_{jj} < 130$ | 6691 | 19061 | $100 < M_{jj} < 130$ | 12334 | 71918 |
|                 | Likelihood ratio | $LR > 0.165$ | 6293 | 10940 | $LR > 0.395$ | 9543 | 11092 |
|                 | Significance (Efficiency) | $S/\sqrt{S + B}$ | 47.9 (32.5%) | | | 66.4 (36.3%) | |

where $M_{j1j2/j3j4}$, $M_{Z/H}$ represent the dijet invariant masses paired from the four jets ($j_{1-4}$) and the $Z$ and Higgs masses, respectively. Here $\sigma_Z = 4.7$ GeV and $\sigma_H = 4.4$ GeV are used for $\sqrt{s} = 250$ GeV and $\sigma_{Z/H} = 3.8$ GeV for $\sqrt{s} = 350$ GeV. They are determined from the dijet mass distribution reconstructed from the true MC information. After the jet pairing, background reductions are applied.

To select the four-jet-like events, cuts on the number of charged tracks $N_{\text{charged}}$ and jet clustering parameter $Y_{34}$ are applied. $Y_{34}$ is the minimum scaled jet mass $y$ required for four-jet clustering. The leptonic backgrounds ($\ell\ell\ell\ell, \ell\ell q\bar{q}$) are reduced effectively by these selections. In addition, cuts on the thrust and thrust angle are applied to reduce the $ZZ$ background, utilizing the difference between the event shape of the signal (spherical) and $ZZ, q\bar{q}$ (back-to-back). The numbers of $q\bar{q}q\bar{q}$ and $q\bar{q}$ background events are reduced by a cut on the angle between the Higgs candidate jets ($\theta_H$). The $WW$ and $ZZ$ backgrounds are further suppressed by cuts on the Higgs and $Z$ candidates after the kinematical constraint fit is applied to the four-jet system as follows. Each jet is parameterized by $E_{ji}$, $\theta_i$, and $\phi_i$ ($i = 1 - 4$) and fitted with constraints on the total energy ($\sum_i E_{ji} = \sqrt{s}$), the total momentum ($\sum_i \vec{P}_{ji} = 0$), and Higgs and $Z$ mass difference ($|M_{j1j2} - M_{j3j4}| = |M_H - M_Z|$), where $E_{ji}$, $P_{ji}$, $\theta_i$, and $\phi_i$ are the energy, momentum, and theta and phi angles of the $i$-th jet, respectively. After these cuts are applied, an additional cut is applied on the $LR$ derived from...
the following input variables: thrust, $\cos\theta_{\text{thrust}}$, minimum angle between all jets ($\theta_{\text{min}}$), number of particles in Higgs candidate jets, fitted $Z$ mass, and fitted Higgs mass. The likelihood cut position is selected to maximize signal significance; $LR > 0.375$ for 250 GeV and $LR > 0.15$ for 350 GeV. All background reduction procedures are summarized in Table II. The background fractions after all cuts are 80% $q\bar{q}q\bar{q}$ and 20% $q\bar{q}$ at 250 GeV and 60% $q\bar{q}q\bar{q}$, 30% $q\bar{q}$ and 10% $t\bar{t}$ at 350 GeV.

TABLE II: Summary of $q\bar{q}H$ channel background reduction assuming $\mathcal{L} = 250 \text{ fb}^{-1}$ with $P(e^-, e^+) = (-0.8, +0.3)$.

| CM energy (GeV) | 250          | 350          |
|-----------------|--------------|--------------|
| Cut names       | condition    | Sig.  | Bkg.  | condition    | Sig.  | Bkg.  |
| Generated       | $\chi^2 < 10$ | 32447 | 2608980 | $\chi^2 < 10$ | 20207 | 1034810 |
| # of charged tracks | $N_{chd} > 4$ | 25281 | 1120950 | $N_{chd} > 4$ | 14900 | 305649 |
| $Y_{34}$ value  | $-\log(Y_{34}) > 2.7$ | 25065 | 1002125 | $-\log(Y_{34}) > 2.7$ | 14543 | 250995 |
| thrust          | thrust < 0.9 | 24688 | 935950 | thrust < 0.85 | 13522 | 144560 |
| thrust angle    | $|\cos\theta_{\text{thrust}}| < 0.9$ | 21892 | 696201 | $|\cos\theta_{\text{thrust}}| < 0.9$ | 12523 | 107025 |
| Higgs jets angle | $105^\circ < \theta_H < 160^\circ$ | 20062 | 622143 | $70^\circ < \theta_H < 120^\circ$ | 11185 | 77659 |
| $Z$ di-jet mass (GeV) | $80 < M_Z < 100$ | 16359 | 411863 | $80 < M_Z < 100$ | 9468 | 45671 |
| $H$ di-jet mass (GeV) | $105 < M_H < 130$ | 16359 | 411863 | $105 < M_H < 130$ | 9451 | 44399 |
| Likelihood ratio | $LR > 0.375$ | 13726 | 166807 | $LR > 0.15$ | 8686 | 25393 |
| Significance (Efficiency) | $S/\sqrt{S+B}$ | 32.3 (26.1%) | $S/\sqrt{S+B}$ | 47.1 (24.1%) |

C. Leptonic channel ($\ell^+\ell^- H$)

For leptonic channel analysis, we considered the cases where the lepton is an electron or a muon. We considered only the $\ell\ell qq$ and $\ell\nuqq$ background processes. First, the following cuts were applied to selected isolated leptons:

- Lepton isolation: $E_{cone} < 20$ GeV (cone angle: $10^\circ$),
- Lepton track momentum:
  - $10 < E_{lep} < 90$ GeV at $\sqrt{s} = 250$ GeV,
  - $10 < E_{lep} < 160$ GeV at $\sqrt{s} = 350$ GeV,
where $E_{\text{cone}}$ is the energy sum for particles within $10^\circ$ of the lepton. The prompt lepton has a smaller $E_{\text{cone}}$ than nonprompt leptons. Electrons and muons are identified from their charged tracks as follows:

- **Electron ID:** $\frac{E_{\text{ECAL}}}{E_{\text{Total}}} > 0.9$, $0.7 < \frac{E_{\text{Total}}}{P} < 1.2$
- **Muon ID:** $\frac{E_{\text{ECAL}}}{E_{\text{Total}}} < 0.5$, $E_{\text{Total}}P < 0.4$,

where $E_{\text{ECAL}}$, $E_{\text{Total}}$ and $P$ denote the ECAL energy associated with a track, total energy deposited in the ECAL and HCAL, and track momentum, respectively. If there are more than two isolated lepton candidates after the electron or muon identification, a pair whose invariant mass is closest to $Z$ is selected. After dilepton identification, forced two-jets clustering is applied to the remaining particles and the following selections are applied. First, a dilepton mass ($M_{\ell\ell}$) cut, which should be consistent with the $Z$ mass, is applied: $70 < M_{\ell\ell} < 110$ GeV for electrons and $70 < M_{\ell\ell} < 100$ GeV for muons. Because the $ZZ$ or $WW$ backgrounds are boosted to the forward region compared to the signal, a cut on the $Z$ direction is applied: $|\cos \theta_Z| < 0.8$. Finally, cuts on dijet mass ($M_{jj}$) and a mass recoil to the lepton pair ($M_{\text{rec}}$) are applied to select the Higgs signal: $100 < M_{jj} < 140$ GeV and $70 < M_{\text{rec}} < 140$ GeV for electrons; $115 < M_{jj} < 140$ GeV and $70 < M_{\text{rec}} < 140$ GeV for muons. The background reduction procedures for the leptonic channel are summarized in Table III. After all cuts were applied, the background was dominated by the $\ell\ell q\bar{q}$ whereas the $\nu\ell q\bar{q}$ was well suppressed.

**IV. BRANCHING RATIO MEASUREMENT**

After event selection, the measurement accuracies of the Higgs BRs to $b\bar{b}$, $c\bar{c}$, and $gg$ are evaluated on the basis of a template fitting to the flavor likeness of the Higgs dijets obtained by using the LCFIVertexing package [15]. The probabilities of $b$ and $c$ quarks for each jet [$b_i$, $c_i$ ($i = 1, 2$)] are calculated in LCFIVertex using neural net training with a $Z \rightarrow q\bar{q}$ sample at the $Z$-pole. In addition, another $c$ probability ($bc_{1,2}$) is also calculated whose neural-net is trained only with $Z \rightarrow b\bar{b}$ sample as the background. For Higgs dijets, we define the flavor likeness $X$ ($X = b$, $c$, $bc$) as follows from the $x_i$ [$x_i = b_i$, $c_i$, $bc_i$ ($i = 1, 2$)] flavor probability of each jet:

$$X = \frac{x_1x_2}{x_1x_2 + (1-x_1)(1-x_2)}.$$  \hspace{1cm} (2)

The flavor tagging performance in the $ZZ \rightarrow \nu\nu q\bar{q}$ sample at the $\sqrt{s} = 250$ and 350 GeV is shown in Fig. 4. The $ZZ \rightarrow \nu\nu q\bar{q}$ samples are compared for each CM energy because they form the
TABLE III: Summary of background reduction in the $eeH$ and $\mu\mu H$ channels assuming $\mathcal{L} = 250$ fb$^{-1}$ with $P(e^-, e^+ = (-0.8, +0.3))$.

| CM energy (GeV) | 250 | 350 |
|----------------|-----|-----|
| Cut names      |     |     |
| Generated      |     |     |
| e/\mu condition|     |     |
| Sig.           | 3137 | 2740 |
| Bkg.           | 4512520 | 3822410 |
| Generated      |     |     |
| e/\mu condition|     |     |
| Sig.           | 2917 | 1789 |
| Bkg.           | 4512520 | 3822410 |
| # of e/\mu track ID |     |     |
| e/\mu condition|     |     |
| Sig.           | 2717 | 2270 |
| Bkg.           | 204403 | 179580 |
| # of e/\mu track ID |     |     |
| e/\mu condition|     |     |
| Sig.           | 2270 | 1631 |
| Bkg.           | 179580 | 23598 |
| Di-lepton mass (GeV) |     |     |
| e/\mu condition|     |     |
| Sig.           | 2208 | 1425 |
| Bkg.           | 34162 | 51436 |
| Z direction     |     |     |
| e/\mu condition|     |     |
| Sig.           | 1797 | 1192 |
| Bkg.           | 21600 | 20874 |
| Di-jet mass (GeV) |     |     |
| e/\mu condition|     |     |
| Sig.           | 1889 | 1203 |
| Bkg.           | 8036 | 6250 |
| Recoil mass (GeV) |     |     |
| e/\mu condition|     |     |
| Sig.           | 1365 | 638 |
| Bkg.           | 983 | 465 |
| Significance (Efficiency) |     |     |
| e/\mu condition|     |     |
| Sig.           | 22.4 | 16.7 |
| Bkg.           | 37.8% | 20.7% |
| Significance (Efficiency) |     |     |
| e/\mu condition|     |     |
| Sig.           | 28.2 | 19.2 |
| Bkg.           | 46.8% | 35.7% |

same final state as $Z \rightarrow q\bar{q}$, which was used to train the flavor tagging neural network. Figure 4 shows that no significant difference in the flavor tagging performance at $\sqrt{s} = 250$ and 350 GeV is observed for any of the flavors.

To evaluate the measurement accuracy of the BRs, the $b$-, $c$-, and $bc$-likenesses of the selected events were binned in a three-dimensional histogram and fitted with those of the template samples, which consist of $H \rightarrow b\bar{b}$, $c\bar{c}$, and $gg$ and other background processes. Figure 5 shows the three-dimensional histogram projected to the two-dimensional $b$- and $c$-likeness axes for the hadronic channel. The probability of entries in each template sample bin is expected to be given by the Poisson statistics:

$$P_{ijk} = \frac{\mu^{n}e^{-\mu}}{n!} \left( n \equiv N_{ijk}^{\text{data}}, \mu \equiv N_{ijk}^{\text{template}} \right),$$

where $P_{ijk}$ and $N_{ijk}^{\text{data}}$ are the probability of entries and the number of data entries at the $(i, j, k)$
FIG. 4: Flavor tagging performance at CM energies of 250 and 350 GeV in $ZZ \rightarrow \nu\bar{\nu}q\bar{q}$ sample. Horizontal axis shows the efficiency of $b/c$ jets; vertical axis shows the purity of tagged $b/c$ jets.

$N^\text{template}_{ijk}$ is given by

$$N^\text{template}_{ijk} = \sum_{s=bb,cc,gg} r_s \cdot N^s_{ijk} + N^b_{ijk},$$

(4)

where $N^s_{ijk}$ is the number of entries at the $(i,j,k)$ bin in each $H \rightarrow bb$, $cc$, and $gg$ template; $N^b_{ijk}$ is the number of entries in the background template sample, which is the sum of the SM background events and the Higgs-to-nonhadronic decay events. Furthermore, $r_{bb}$, $r_{cc}$, and $r_{gg}$ are the parameters to be determined by the template fitting. They are defined as the Higgs branching fractions to $H \rightarrow bb$, $cc$ and $gg$, respectively, normalized by that of the SM,

$$r_s = \frac{\sigma \cdot Br(H \rightarrow s)}{\sigma_{SM} \cdot Br(H \rightarrow s)^{SM}} (s = bb, cc, gg).$$

(5)

Here $\sigma$ is the Higgs production cross section and $\sigma_{SM}$ and $Br(H \rightarrow s)^{SM}$ are the cross section and branching fraction in the SM, respectively. From Eq. (5), the measurement accuracies of $\sigma \cdot Br$ are obtained as follows;

$$\frac{\Delta (\sigma \cdot Br)(H \rightarrow s)}{\sigma \cdot Br}(H \rightarrow s) = \frac{\Delta r_s}{r_s} (s = bb, cc, gg).$$

The $r_s$’s values were determined by a binned log likelihood fitting, where each bin probability is given by Eq. (3).

On the basis of the three-dimensional (3D) histogram, 5000 toy MC events were generated using the Poisson distribution function for each bin, which were fitted to obtain $r_{bb}$, $r_{cc}$, and $r_{gg}$. The
number of bins in the 3D histogram were optimized to minimize the statistical fluctuation in the fitted results caused by low-statistic bins. Bins with fewer than one entry were not used for the fitting. The distributions of $r_{bb}$, $r_{cc}$, and $r_{gg}$ for template fitting to 1000 toy MC events are shown in Fig. 6. The error in $r_s$ is determined by the Gaussian fittings to these distributions, which are shown in Tables IV and V for CM energies of 250 and 350 GeV, respectively.

The tables also show the accuracies after correction of the total cross section. From a study of the recoil mass in the process of $e^+e^- \rightarrow eeH$ and $\mu\mu H$, the accuracy of the total cross section ($\Delta\sigma/\sigma$) was estimated to be 2.5% at 250 GeV [9, 16]. For 350 GeV, we assumed an accuracy of 3.5% because the recoil mass measurement relies on the $ZH$ process, whose cross section is inversely proportional to the square of the CM energy; thus, the accuracy of the total cross section measurement would be inversely proportional to the CM energy.

From Tables IV and V, we see that the Higgs cross section times branching ratio can be measured at about 1% for $H \rightarrow bb$ and 7 to 9% for $H \rightarrow cc$ and $gg$. The measurement is approximately 10 – 20% better at 350 GeV than at 250 GeV. The instantaneous luminosity at 350 GeV is 25% greater than that at 250 GeV according to the ILC beam parameters. Thus, for an equal running time, measurements at 350 GeV will give us about 20 – 30% better accuracy than those at 250 GeV. On the other hand, the accuracy of the BR to $bb$, $\Delta Br/Br(H \rightarrow bb)$, is limited by the total cross section ambiguity; thus, measurement at 250 GeV gives us better results than that at 350 GeV. In the other decay channels, comparable BR measurements are possible even if the same
The measurement accuracy of the Higgs branching fractions, $H \rightarrow b \bar{b}$, $c \bar{c}$, and $gg$, were evaluated at $\sqrt{s} = 250$ GeV and 350 GeV. In terms of signal significance, $\sqrt{s} = 350$ GeV yields better background suppression than $\sqrt{s} = 250$ GeV for each channel. The combined results for measurement accuracies of the Higgs cross section times BRs ($\Delta (\sigma \cdot Br)/\sigma \cdot Br$) to $H \rightarrow b \bar{b}$, $c \bar{c}$, and $gg$ are 1.0%, 6.2%, and 7.3% at CM energies of 250 GeV and 1.0%, 6.2%, and 7.3% at 350 GeV, respectively, assuming the same integrated luminosity of 250 fb$^{-1}$. At the ILC, the total Higgs ...
TABLE V: Summary of template fitting results $r_s$ and accuracies of $(\sigma \cdot Br)$ and $Br$ after correcting $\sigma$ for an accuracy of 3.5% at $\sqrt{s} = 350$ GeV assuming $L = 250$ fb$^{-1}$ with $(e^-, e^+) = (-0.8, +0.3)$.

|                  | $\nu\bar{\nu}H$ | $q\bar{q}H$ | $e^+e^-H$ | $\mu^+\mu^-H$ | comb. |
|------------------|------------------|--------------|------------|----------------|-------|
| $r_{bb}$         | 1.00±0.01        | 1.00±0.02    | 1.00±0.05  | 1.00±0.05      | 1.00±0.01 |
| $r_{cc}$         | 1.02±0.11        | 1.01±0.10    | 1.02±0.31  | 1.04±0.32      | 1.01±0.06 |
| $r_{gg}$         | 1.02±0.14        | 1.04±0.14    | 1.04±0.37  | 1.03±0.34      | 1.02±0.07 |
| $\Delta(\sigma Br)/(H \to bb)$ (%) | 1.4 | 1.5 | 5.3 | 5.1 | 1.0 |
| $\Delta(\sigma Br)/(H \to cc)$ (%) | 8.6 | 10.1 | 30.5 | 30.9 | 6.2 |
| $\Delta(\sigma Br)/(H \to gg)$ (%) | 9.2 | 13.7 | 35.8 | 33.0 | 7.3 |
| $\Delta Br/Br(H \to bb)$ (%) | 3.8 | 3.8 | 6.4 | 6.2 | 3.6 |
| $\Delta Br/Br(H \to cc)$ (%) | 9.2 | 10.6 | 30.7 | 31.1 | 7.2 |
| $\Delta Br/Br(H \to gg)$ (%) | 9.8 | 14.1 | 36.0 | 33.2 | 8.1 |

cross-section $\sigma$ is measured using the $Z$ recoil mass process. Using $\Delta \sigma/\sigma = 2.5\%$ for 250 GeV and assuming it is 3.5% at 350 GeV, Higgs BRs $(\Delta Br/Br)$ to $bb$, $cc$, and $gg$ are derived as 2.7%, 7.3%, and 8.9% at CM energies of 250 GeV and as 3.6%, 7.2%, and 8.1% at 350 GeV. Therefore, we conclude that the Higgs cross section times BR $(Br \times \sigma)$ can be measured better at 350 GeV than at 250 GeV owing to the higher S/N at the higher energy. However, when the accuracy of the total cross section measurement by recoil mass measurement is considered, BR of $H \to bb$ can be measured better at 250 GeV, even if the integrated luminosity is the same at both energies.

Acknowledgment

The authors thank the members of the ILC physics subgroup for useful discussions of this work and those of the ILD software and optimization group, who maintain the software and MC samples used in this work. This work was supported in part by Creative Scientific Research Grant No. 18GS0202 from the Japan Society for Promotion of Science (JSPS), the JSPS Core University Program, and JSPS Grant-in-Aid for Scientific Research No. 22244031.

[1] The ATLAS Collaboration, Phys. Lett. B 710, 49 (2012).
[2] The CMS Collaboration, Phys. Lett. B 710, 26 (2012).
[3] A. Djouadi et al., International Linear Collider Reference Design Report Volume 2: PHYSICS AT THE ILC, arXiv:0709.1893v1 [hep-ph].
[4] Y. Banda, T. Lastovicka, and A. Nomerotski, Phys. Rev. D 82, 033013 (2010).
[5] M. Battaglia, arXiv:9910271 [hep-ph].
[6] M. Carena, H. Haber, and H. Logan, and S. Mrenna, arXiv:0106116 [hep-ph].
[7] T. Kuhl and K. Desch, LC-PHSM-2007-001.
[8] GEANT4 Collaboration: S Agostinelli et al., Nucl. Instrum. Methods A 506, 250 (2003).
[9] ILD Concept Group, The International Large Detector: Letter of Intent, KEK Report 2009-6.
[10] W. Kilian et al., arXiv:0708.4233 [hep-ph]
    M. Moretti et al., arXiv: 0102195v1 [hep-ph].
[11] T. Sjöstrand, S. Mrenna, and P. Skands, JHEP 0605, 026 (2006).
[12] M. A. Thomson, Nucl. Instrum. Methods A 611, 1 (2009).
[13] P. Mora de Freitas and H. Videau, LC-TOOL-2003-010, Prepared for LCWS 2002., Jeju Island, Korea, 26-30 Aug 2002.
[14] J. Brau et al., The International Linear Collider Interim report Volume 2, Physics and Detectors 2011 status report, KEK Report 2011-5.
[15] D. Bailey et al., Nucl. Instrum. Methods A 610, 2 (2009).
[16] H.Li, arXiv:1007.2999v1 [hep-ex].