

Study of the Higgs Direct Reconstruction in $ZH \rightarrow q\bar{q}H$ for ILC

Hiroaki Ono*

Nippon Dental University School of Life Dentistry at Niigata, Niigata, Japan

Precise measurement of the Higgs boson properties is an important issue of the International Linear Collider (ILC) experiment. We studied the accuracy of the Higgs mass reconstruction in the $ZH \rightarrow q\bar{q}H$ multi-jet process with the Higgs mass of $M_H = 120$ GeV at $\sqrt{s} = 250$ GeV with the ILD detector model. In this study, we obtained the reconstructed Higgs mass of $M_H = 120.79 \pm 0.089$ GeV and 5.3% measurement accuracy of the cross-section for $ZH \rightarrow q\bar{q}b\bar{b}$ with the integrated luminosity of $\mathcal{L} = 250 \text{ fb}^{-1}$ data samples.

1 Introduction

International Linear Collider (ILC) [1] is a future e^+e^- collider experiment for the precise measurement and the validation of the Standard Model (SM) physics, especially for the measurement of the Higgs boson property, even the discovery of the Higgs boson will be realized in Large Hadron Collider (LHC) experiment. In the SM, light Higgs boson mass (M_H) is predicted around the $114.4 \text{ GeV} \leq M_H \leq 160 \text{ GeV}$ from the study in LEP [2] and Tevatron [3] experiment. The largest production cross-section for SM Higgs boson is obtained through the Higgs-strahlung ($e^+e^- \rightarrow Z^* \rightarrow ZH$) process which associated with the Z boson and the Z mainly decays to $q\bar{q}$ pair, as shown in Fig. 1, around the ZH production threshold energy shown in Fig. 2 (a).

Since Higgs boson mainly decays to $b\bar{b}$ pair at the Higgs mass below 140 GeV region as shown in Fig. 2 (b), the final state of the $ZH \rightarrow q\bar{q}H$ process forms the four-jet. In ILC experiment, the most of interesting physics processes including ZH process form the multi-jets final state from the decay of gage bosons (W, Z) and heavy flavor quarks (b, c), thus ILC detectors are required to have the good jet energy resolution for the precise measurement. There are three detector concepts, SiD, ILD and 4th for the ILC detector, and ILD is the merged concepts of the previous GLD [4] (Asian group) and LDC [5] (European group) models for the Letter of Intent (LOI) submission [6]. In order to achieve the best jet energy resolution, ILD adopt the Particle Flow Algorithm (PFA) suited detector design. Since the PFA performance is degraded by the cluster overlapping and the double-counting of the particles energy in the calorimeter, particles separation in the calorimeter is an important key for better PFA performance.

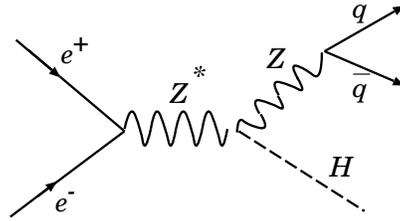


Figure 1: Higgs boson production via Higgs-strahlung (ZH) process and Z mainly decay to $q\bar{q}$.

*TEL:+81-25-267-1500-(537), MAIL:ono@ngt.ndu.ac.jp

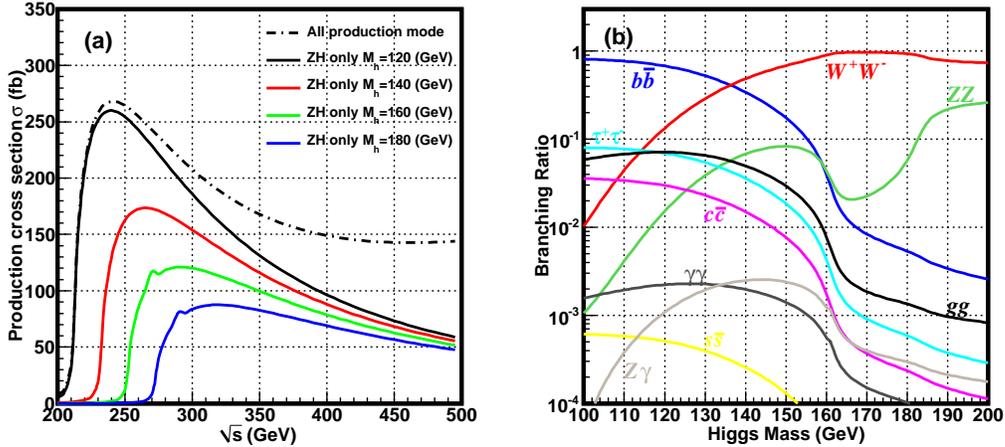


Figure 2: (a). Production cross-section of the Higgs boson as a function of center-of-mass energy (\sqrt{s}) and (b). branching ratio of the Higgs decay as a function of the Higgs mass.

The figure-of-merit of the PFA performance from each detector parameter relating to the particles separation in the calorimeter is described as $F.O.M. = BR^2/\sqrt{\sigma^2 + R_M^2}$, where B is a magnetic field, R is a detector radius, σ is a segmentation of the calorimeter and R_M is an effective Moliere radius of the calorimeter. In order to maximize the $F.O.M.$, ILD detector adopts the large radius tracker and high granularity calorimeter with 3.5 T magnetic field. In this analysis, we study the direct reconstruction of the Higgs boson mass with the full detector simulation for $ZH \rightarrow q\bar{q}H$, $H \rightarrow b\bar{b}$ four-jet mode with the ILD detector model.

2 Simulation tools

For full detector simulation study, we use the ILD detector model based Monte Carlo (MC) full simulation package called Mokka, which is based on the MC simulation package Geant4 [7]. Generated MC hits are reconstructed and smeared in the reconstruction package called MarineReco which includes the PFA package called PandoraPFA [8]. Since $\sqrt{s} = 250$ GeV reconstructed and skimmed signal and background samples called DST files are generated for the LOI physics analysis in ILD group, we use these DST data samples saved in the linear collider common data format called LCIO. For the DST data sample analysis, we use the useful analysis package library called Anlib for the event shape analysis and jets reconstruction, and analysis process is handled through the Root [9] based analysis framework called JSF [10]. For the comparison of the PFA performance between realistic PFA and perfect-clustering PFA, we also use the GLD detector model MC full simulator called Jupiter [11] with the generating the signal and background events by PYTHIA, and reconstruction package called Satellites [12] based on Root, both of them are also controlled in the JSF framework. From the comparison of the $ZH \rightarrow q\bar{q}H$ in GLD detector model, shown in Fig 2, PandoraPFA reconstruction performance (a) achieve the comparable perfor-

mance with perfect-clustering PFA (b) in terms of the reconstructed Higgs mass distribution width of σ which corresponds to the jet energy resolution even only the $ZZ \rightarrow q\bar{q}q'\bar{q}'$ background is considered. Therefore, we shift to the full SM background analysis with common DST data.

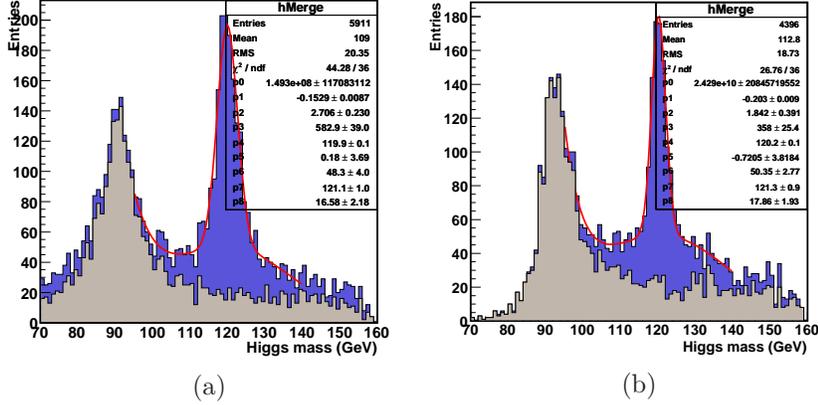


Figure 3: Comparison of the reconstructed Higgs mass distribution for $ZH \rightarrow q\bar{q}b\bar{b}$ only with ZZ background in GLD detector model with the different PFA clustering of (a) realistic PandoraPFA and (b) perfect clustering PFA.

3 Analysis Procedure of $ZH \rightarrow q\bar{q}H$ mode

3.1 MC samples

The SM Higgs boson is mainly produced through the Higgs-strahlung $e^+e^- \rightarrow ZH$ process around the production threshold center-of-mass energy ($\sqrt{s} \sim 230$ GeV). Since the main decay mode at $M_H < 2M_W$, Higgs boson mainly decays to $b\bar{b}$ pair, thus largest production cross-section is obtained from the $ZH \rightarrow q\bar{q}b\bar{b}$ process, which forms four-jet final state and both Z and H can be reconstructed directly. Fig. 4 shows the typical event display of the $ZH \rightarrow q\bar{q}H$ in JSF. In this analysis, we assume the center-of-mass energy as the ZH production threshold of $\sqrt{s} = 250$ GeV and the light Higgs mass of $M_H = 120$ GeV. Each DST data samples is scaled to the integrated luminosity of $\mathcal{L} = 250 \text{ fb}^{-1}$ and the beam polarization to $P(e^+, e^-) = (30\%, -80\%)$. The main backgrounds for $ZH \rightarrow q\bar{q}b\bar{b}$ are considered as following processes: $ZH \rightarrow Z^*/\gamma \rightarrow q\bar{q}$, $e^+e^- \rightarrow WW/ZZ \rightarrow$

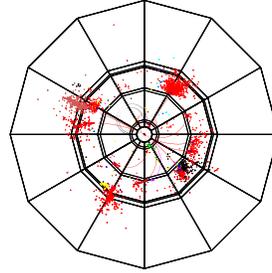


Figure 4: Typical event display of the $ZH \rightarrow q\bar{q}H$ four-jet final state.

$qq'q''q'''$ or $q\bar{q}q'\bar{q}'$, $e^+e^- \rightarrow WW \rightarrow \nu lqq'$ and $e^+e^- \rightarrow ZZ \rightarrow \ell\ell\ell$. Generated signal and background MC samples which scaled to be $\mathcal{L} = 250 \text{ fb}^{-1}$ are summarized in Table. 1.

MC samples ($\mathcal{L} = 250 \text{ fb}^{-1}$)	$ZH \rightarrow qqH$ (sig)	$qqqq$	νlqq	$\ell\ell\ell$	qq
Number of generated events	51763	814163	302807	98127	2529928

Table 1: Generated signal and background MC data samples scaled with $\mathcal{L} = 250 \text{ fb}^{-1}$.

In order to correct the escape energy from the heavy quark decay including neutrinos, kinematic five constraint (5C) fit is applied, which consists of the four constraints (4C) of momentum balance ($\sum P_{x,y,z_i} = 0$) and jets energy balance ($\sum E_i - \sqrt{s} = 0$) of the four-jet and one Z mass constraints for Z candidate di-jet. For the kinematic fitting, jet energies (E_j) and jet angles (θ, ϕ) of each jet are used as measured variables. Finally, reconstructed Higgs mass distribution is fitted with the Gaussian convoluted with Gaussian function for the signal and exponential function for the contribution from background events which remain after the Higgs boson selections.

3.2 Jet Reconstruction

Since the final state of the $ZH \rightarrow q\bar{q}H$ mode forms four-jet, after the PandoraPFA clustering, forced four-jet clustering based on Durham jet-clustering algorithm has applied. In order to select the best jet pair combination from the four-jet, following χ^2 value is evaluated,

$$\chi^2 = \left(\frac{M_{12} - M_Z}{\sigma_{M_Z}} \right)^2 + \left(\frac{MissM_{34} - M_Z}{\sigma_{MM_H}} \right)^2 \quad (1)$$

where M_{12} is Z candidate di-jet mass, $MissM_{34}$ is a missing mass of the remaining Higgs candidate di-jet, M_Z is the Z boson mass (91.2 GeV), and σ_{M_Z} and $\sigma_{MissM_{34}}$ are sigma of distribution of the reconstructed Z boson mass and the missing mass of the Higgs candidate jets, respectively. In order to select the best jets pair combination, $\chi^2 < 10$ is required for the reconstructed jets pair.

3.3 Event selection

After the χ^2 cut to select the best jet pair combination, following event selections are applied for background rejection:

- (a) visible energy : $200 \leq E_{vis} \leq 270 \text{ GeV}$;
- (b) Longitudinal momentum of the Z : $|P_{\ell Z}| < 70 \text{ GeV}$ to reduce ZZ background;
- (c) Higgs production angle : $|\cos \theta_H| < 0.85$ to reduce the ZZ background;
- (d) thrust angle : $thrust < 0.9$;
- (e) Number of particles: $N_{particle} > 40$ to suppress the $\ell\ell\ell$ background;
- (f) Maximum and minimum jet energy fraction: $E_{min}/E_{max} > 0.25$;
- (g) Maximum momentum of jet: $P_{j_{max}} < 100 \text{ GeV}$;
- (h) Y Plus : $YPlus > 0.0001$;

- (i) Y Minus : $Y_{Minus} > 0.001$;
- (j) Minimum angle of Z - H jets : $20 < \theta_{ZHj_{min}} < 135$;
- (k) Maximum angle of Z - H jets : $110 < \theta_{ZHj_{max}}$;
- (l) b -tagging : $P_{btag} > 0.5$ from LCFIVTX package.

The distribution and its cut positions for each selection variable are shown in Fig. 5. Since the W/Z generated in the WW/ZZ background event are relatively boosted compare to the Z generated in ZH signal event, longitudinal momentum of Z ($P_{\ell Z}$) and maximum momentum in jets ($P_{j_{max}}$) are higher in WW/ZZ background event than in signal event. None jet-like background events are reduced by the number of particles (N_{PFO}) cut. Y Plus and Y Minus values are threshold Y-values used in the jet clustering topology which reconstructed from four-jet to five-jet or three-jet, respectively. Minimum and maximum angles between Z and H candidate jets are also used for the separation by the event shape difference between ZH event and backgrounds.

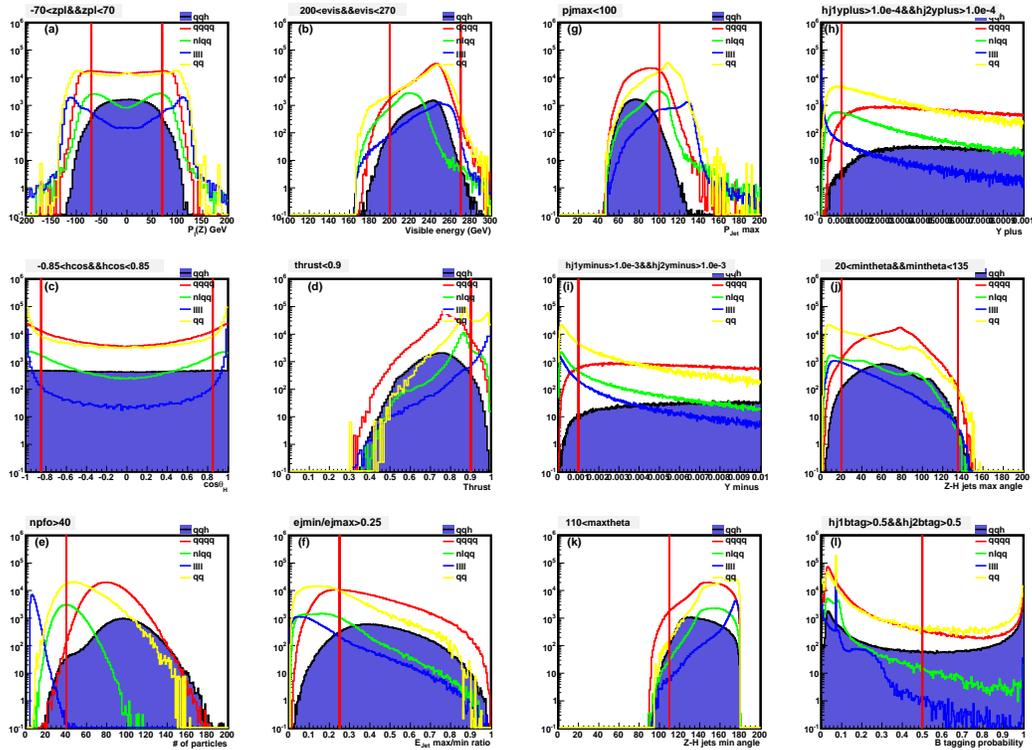


Figure 5: Distribution of each selection variable and its cut positions to select $ZH \rightarrow q\bar{q}b\bar{b}$ event.

Finally, we apply the vertex tagging selection for the neural net output of the b -likeness analyzed in the vertexing package called LCFIVTX in ilcsoft. The reduction summary in each event selection is listed in the Table 2.

Selections	$ZH \rightarrow q\bar{q}H(\text{Sig})$	$qqqq$	$\nu\ell qq$	$\ell\ell\ell$	qq
no cuts	51745	814162	302807	98127	2529928
χ^2	36748 (71.02 %)	688703 (84.59 %)	19043 (6.29 %)	25375 (25.86 %)	541852 (21.42 %)
$ P_{1Z} $	34952 (67.55 %)	479403 (58.88 %)	12832 (4.24 %)	5565 (5.67 %)	293883 (11.62 %)
E_{vis}	34924 (67.49 %)	477994 (58.71 %)	12457 (4.11 %)	5335 (5.44 %)	287324 (11.36 %)
$ \cos\theta_H $	30451 (58.85 %)	397270 (48.79 %)	9934 (3.28 %)	2167 (2.21 %)	223873 (8.85 %)
$thrust$	29916 (57.81 %)	389703 (47.87 %)	8312 (2.75 %)	1422 (1.45 %)	103283 (4.08 %)
$N_{particles}$	29820 (57.63 %)	389514 (47.84 %)	4353 (1.44 %)	0 (0.00 %)	87022 (3.44 %)
$E_{i\min}/E_{i\max}$	27843 (53.81 %)	297580 (36.55 %)	1603 (0.53 %)	0 (0.00 %)	40880 (1.62 %)
$p_{j\max}$	27622 (53.38 %)	289490 (35.56 %)	1500 (0.50 %)	0 (0.00 %)	31382 (1.24 %)
Y_{plus}	27607 (53.35 %)	288421 (35.43 %)	1465 (0.48 %)	0 (0.00 %)	30773 (1.22 %)
Y_{minus}	27559 (53.26 %)	287825 (35.35 %)	1354 (0.45 %)	0 (0.00 %)	27250 (1.08 %)
$\theta_{Z-Hi\min}$	27311 (52.78 %)	285704 (35.09 %)	1284 (0.42 %)	0 (0.00 %)	24601 (0.97 %)
$\theta_{Z-Hi\max}$	27031 (52.24 %)	277203 (34.05 %)	1263 (0.42 %)	0 (0.00 %)	24280 (0.96 %)
b -tagging	5972 (11.54 %)	4732 (0.58 %)	0 (0.00 %)	0 (0.00 %)	458 (0.02 %)

Table 2: Backgrounds reduction summary in each selection for $ZH \rightarrow q\bar{q}b\bar{b}$.

From the reduction summary of Table. 2, $\ell\ell\ell$ four-leptonic background can be suppressed completely by number of particles cut ($N_{PFOs} < 40$) and the remaining backgrounds are $qqqq$ and qq which including b -quarks event after applying the b -tagging.

4 Results

Reconstructed Higgs mass distribution after the selection of $ZH \rightarrow q\bar{q}b\bar{b}$ is fitted with the function of Gaussian convoluted Gaussian with the exponential function assuming the background, as shown in Fig. 6. Fitted results of the reconstructed $ZH \rightarrow q\bar{q}b\bar{b}$ Higgs mass distribution are summarized in the Table. 3. From the fitted results, Higgs mass ($M_H = 120$ GeV at MC) is reconstructed as $M_H = 120.79 \pm 0.089$ GeV and the measurement accuracy of cross-section to $ZH \rightarrow q\bar{q}b\bar{b}$ is obtained as $\delta\sigma/\sigma = 5.3\%$.

5 Conclusion

Simulation study of the direct reconstruction of the Higgs boson in $ZH \rightarrow q\bar{q}b\bar{b}$ four-jet mode with the Higgs mass of 120 GeV at the $\sqrt{s} = 250$ GeV and the integrated luminosity of $\mathcal{L} = 250 fb^{-1}$ has performed for the ILD detector model considering with

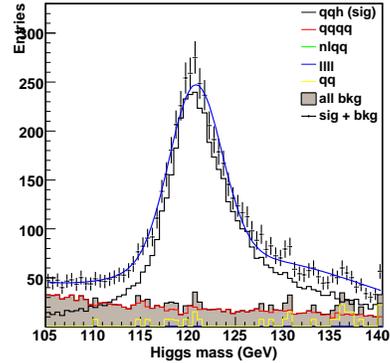


Figure 6: Reconstructed Higgs mass distribution of $ZH \rightarrow q\bar{q}b\bar{b}$.

Higgs mass ($M_H = 120$ GeV at MC)	$M_H = 120.79$ (GeV)
Measurement accuracy of M_H	$\delta M_H = 89$ (MeV)
Measurement accuracy of $\sigma(ZH \rightarrow q\bar{q}b\bar{b})$	$\delta\sigma/\sigma = 5.3\%$

Table 3: Fitted results for the reconstructed Higgs mass distribution.

the $qqqq$, νlqq $\ell\ell\ell\ell$, qq background processes. From the study, measurement accuracy of the reconstructed Higgs mass is estimated as 87 MeV and the measurement accuracy of the cross-section of $ZH \rightarrow q\bar{q}b\bar{b}$ mode is obtained as $\delta\sigma/\sigma = 5.3\%$.

Acknowledgment

I would like to thank to everyone who join the ILC physics WG subgroup [14] for useful discussion of this work and to ILD optimization group members who maintain the softwares and MC samples. This study is supported in part by the Creative Scientific Research Grant No. 18GS0202 of the Japan Society for Promotion of Science and promotion.

References

- [1] ILC Reference Design Report (RDR) <http://www.linearcollider.org/rdr/>
- [2] The LEP Electroweak Working Group, arXiv:0811.4682 [hep-ex] (November 2008).
- [3] CDF Collaboration and D0 Collaboration, arXiv:0903.4001 [hep-ex].
- [4] GLD Detector Outline Document (DOD), arXiv:physics/0607154v1 [physics.ins-det]
- [5] <http://ilcldc.org/documents/dod>
- [6] <http://www.ilcild.org/>
- [7] GEANT4 Collaboration: S Agostinelli et al, Nucl. Instrum. Methods A506, 250 (2003).
- [8] <http://ilcsoft.desy.de/portal/>
- [9] <http://root.cern.ch/>
- [10] <http://acfahep.kek.jp/subg/sim/simtools/>
- [11] ACFA Linear Collider Working Group, KEK Report 2001-11, August, 2001.
- [12] Proceedings of the APPI Winter Institute, KEK Proceedings 2002-08, July (2002).
- [13] S. Yamamoto, K. Fujii and A. Miyamoto, arXiv:0809.4111 [physics.comp-ph].
- [14] <http://www-jlc.kek.jp/subg/physics/ilcphys/>