The Higgs Self Coupling Analysis Using The Events Containing H→WW* Decay

Masakazu Kurata, Tomohiko Tanabe
The University of Tokyo
Junping Tian, Keisuke Fujii
High Energy Accelerator Research Organization(KEK)
Taikan Suehara
Tohoku University

Abstract

We report the analysis to measure Higgs self-coupling in double Higgs events at the International Linear Collider (ILC). Both signal and background event samples are generated by a full detector simulation based on the International Large Detector (ILD). To gain the signal acceptance for the improvement of the final significance, the analysis is focused on the events, in which one of the Higgs boson decays into W boson pair. To reject the background events effectively, Multi Variate Analysis(MVA) is introduced instead of cut based event selection. At 500 GeV, assuming an integrated luminosity of 2 ab^{-1} and the Higgs mass of 125 GeV/c^2, the total signal significance of $e^+e^-\to ZHH\to Z(b\bar{b})(WW^*)$ process reaches 1.91σ.

1 Introduction

Higgs-like boson has been discovered at the Large Hadron Collider(LHC) in 2012. And the elementary particle physics will go to the next stage of precise measurement. It is very important to verify the Higgs boson condenses in the vacuum and gives masses to all the standard model particles. Higgs self-coupling is just the force which makes the Higgs boson condense in the vacuum. Hence to measure the coupling is one of the most essential tests which reveal the nature of the Higgs sector.

In the standard model, the Higgs potential after the electroweak symmetry breaking is given as

$$V(H) = \lambda v^2 H^2 + \lambda v H^3 + \frac{1}{4}\lambda H^4 \tag{1}$$

where H is the physical Higgs field, v is the vacuum expectation value of $v \sim 246\text{GeV}$, and λ is the strength of Higgs self-coupling. There are three terms in this potential, the first is the Higgs mass term, with the mass $M_H = \sqrt{2\lambda v^2}$; the second term is a trilinear Higgs self interaction, with the trilinear self-coupling $\lambda_{HHH} = 6\lambda v$; the third term is a quartic Higgs self interaction, with the quartic Higgs self-coupling $\lambda_{HHHH} = 6\lambda$. It is necessary to verify all these three terms respectively. However, the quartic Higgs self-coupling measurement is very difficult due to the very small cross section of three Higgs boson production process (less than 0.001fb). Therefore, it becomes crucial to investigate the feasibilities of measuring the trilinear Higgs self-coupling.

2 Signal and Backgrounds

2.1 Signal

As the signal event for measuring trilinear Higgs self-coupling, double Higgs-strahlung process, $e^+e^- \to ZHH$[1][2][3], is considered. Figure 1 shows the cross sections of these two processes as a function of the center-of-mass energy. The double Higgs-strahlung process is expected to be a dominant process at around the center-of-mass energy of 500 GeV and to be superseded by the WW fusion process, $e^+e^- \to \nu\overline{\nu}HH$, at higher energy around 1 TeV. Signal events are generated with **WHIZARD**, and full detector simulation is based on the International Large Detector (ILD). The latest detector configuration of ILD is adopted. The constraint of Z decay is imposed for the detailed analysis(see section3.1). Efficient number of events($\sim 5.0 \times 10^5$) are generated to obtain stable results of Multi Variate Analysis training. It is assumed that the polarizations of electron(denote as e^-) and positron(denote as e^+) are $P(e^-, e^+) = (-0.8, +0.3)$. The Higgs bosons decay into all the modes. The events having $H \to WW^*$ would be selected for the analysis. Higgs mass is set at $m_H = 125\text{GeV/c}^2$.

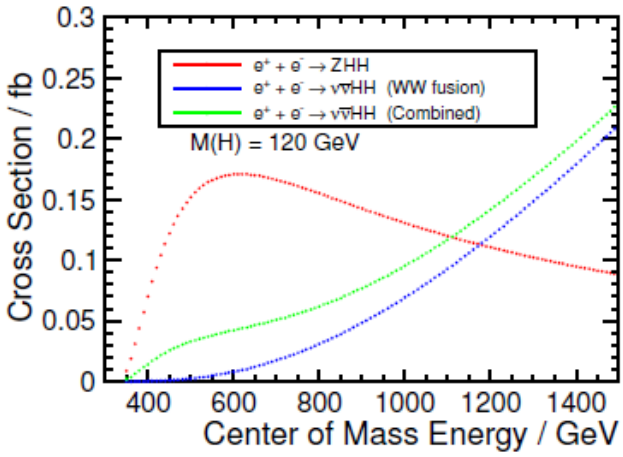

Figure 1: The separate and combined production cross sections for the ZHH and $\nu\bar{\nu}$HH processes as a function of the center-of-mass energy assuming the Higgs mass of 120 GeV. The red line is for the ZHH process, the blue line is for the $\nu\bar{\nu}$HH fusion process and the green line is for the combined result.

2.2 Backgrounds

- $t\bar{t}$ and ZWW events are the main backgrounds due to those huge cross sections. However, flavor tagger can reduce those backgrounds drastically because those events have fewer number of b quarks in the final states. In addition, by using the difference of kinematic topology, those backgrounds can be suppressed at enough degree.

- The events which have many b quarks in the final state can be one of the backgrounds because it is expected that many events would survive after using flavor tagger. Therefore, those backgrounds should be reduced using the difference of kinematic topology. Those backgrounds include ZZ, Zγ, Z+$b\bar{b}$ and triple boson events(ZZZ and ZZH).

The signal-to-background ratio is expected to be $\sim$ 1/3500 at 500GeV of the center mass of energy. Therefore, it is essential to reduce the backgrounds effectively for the good sensitivity. All the background events are generated with the same configuration as the signal event simulation.

	$WW \to (q\overline{q})(q\overline{q})$	$WW \to (l\nu)(q\overline{q})$
$Z \to b\overline{b}$	8jets	lepton + 6jets
$Z \to c\overline{c}$	8jets	lepton + 6jets
$Z \to l^+l^-$	dilepton + 6jets	trilepton + 4jets

Table 1: Summary of analyzed events

3 Analysis

3.1 Analysis strategy

Signal events can be classified on the basis of Z boson and W boson decays. Table 1 summarizes the process types considered for the analysis. So far, flavor tagging is an essential tool to suppress the huge number of backgrounds to enough extent. Therefore the events that Z decays into heavy quark pair can be mainly available. In addition to such decays, very clean Z mass distribution can also suppress the backgrounds effectively. Lepton pair from Z boson will meet such condition.
Analyzed processes are as following:

- All the Z and W bosons decay into quark pair. This process has the largest cross section among the signal events. All the final state objects form jets.
- Z decays into quark pair and one of the W bosons decays into lepton and neutrino while the other W boson decays into quark pair. This process is easier than
- Z decays into lepton pair, and all the W bosons decay into quarks.
- Z decays into lepton pair, and one of the W boson decays into lepton and neutrino while the other W boson decays into quarks. It is expected that the process can suppress backgrounds extremely though the cross section is quite small.

3.2 Event Selection and Making samples

Event selection proceeds according to the 3 steps. Figure 2 shows the analysis flow. First, events are collected on the basis of lepton, jets, and missing momentum selection. The selection is performed to have the event samples for more detailed analysis. And then the background rejection is performed with multi variate technique using the part of pre-selected event samples as training samples. And the final signal and background events are obtained after imposing optimized b-tagging constraint.

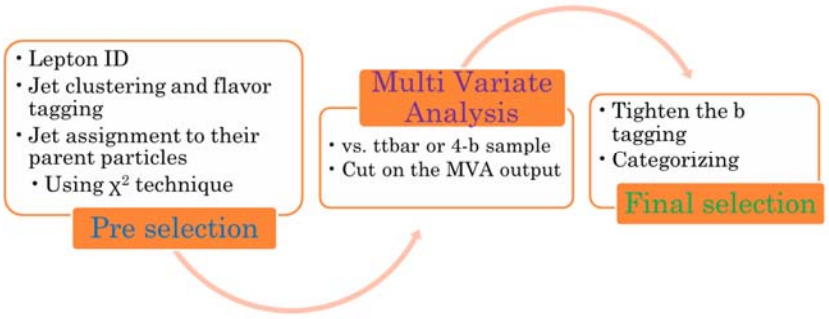

Figure 2: Schematic image of the analysis flow

Lepton	Electron	Muon
Cut	0.65<E/p<1.25	E/p<min(0.5, 10.0/E)
	E(EM)/(E(EM)+E(HAD))>0.90	E(EM)/(E(EM)+E(HAD))<0.45
	$\|d0\|$<0.02	$\|d0\|$<0.02
	$\|z0\|$<0.02	$\|z0\|$<0.02
	Cone Energy<61.10-0.28p	Cone Energy<52.45-0.28p

Table 2: Lepton selection criteria. Isolated electrons and muons coming from primary vertex are selected.

3.2.1 Preselection

Preselection consists of lepton selection, jet selection, missing momentum cut, and the baseline of b-tagging requirement.

Lepton selection Leptons should be daughters of Z boson or W bosons. Therefore, leptons should be coming from the primary vertex. Table 2 shows a summary of isolated lepton selection. Selection is based on the variables of energy/momentum ratio, and the ratio of the deposit energies on hadron/electromagnetic calorimeter. Impact parameter constraint is imposed for leptons coming from primary vertex. The tracks should isolate from other high momentum tracks. The selection can find tight elections or muons positively. Using above selection, over 98% of the signal events with a tight lepton can be collected, and more than 93% of the events without tight leptons can be suppressed. Lepton energy cut of E(lep)> 15GeV is required to reject soft leptons coming from all the hadronic events.

Jet selection and b-tagging Jet clustering should be applied to obtain the jets coming from final state objects. Jet clustering and b-tagging are based on the LCFIPlus module,

which is optimized for the latest detector configuration. Correct number of jets should be required for each event sample. And then minimum jet energy cut is imposed to reject trivial background events. Minimum jet energy cuts vary over required number of jets. over 99% of signal events will survive after imposing this cut. After the jet clustering, b jets are loosely tagged with the b-likeliness>0.4, which is the output of the optimized heavy flavor tagging algorithm in the LCFIPlus module.

Jet pairing Jets should be assigned to their parent particles correctly because it is important to obtain good kinematic variables, which are used to reject background events. Jet paring is based on the χ^2 technique. For this analysis, there are 2 cases which need to solve jet assignment problem.

- 4 b-jet candidates should be assigned to Z boson or Higgs boson. In this case, χ^2 is defined as

$$\chi^2 = \frac{(m_1 - m_Z)^2}{\sigma_Z^2} + \frac{(m_2 - m_H)^2}{\sigma_H^2}, \tag{2}$$

  where m_1, m_2 are the b jet pair masses, m_Z, m_H are Z boson and Higgs boson masses, and σ_Z, σ_H are mass width of Z boson and Higgs boson respectively.

- 4 non-b jet candidates should be assigned to 2 W bosons from Higgs boson decay. Figure 3 shows the distributions of 2 W bosons. For this figure, matching between reconstructed jets and generator level quarks is performed. In Figure 3, The red histogram indicates W bosons with largest energy non-b jet candidate. This shows that W bosons with largest energy jet tend to be on-shell W bosons. Hence W mass constraint is imposed on jet pair with largest energy jet. In this case Breit-Wigner distribution is assumed for mass distribution shape because it is taken into account that both W bosons are off-shell. Therefore, χ^2 is defined as

$$\chi^2 = -2\log(\mathrm{BW}(m_1 | m_W, \sigma_W)), \tag{3}$$

  where m_1, is the jet pair mass with largest energy jet, m_W is W boson mass, and σ_W, is the decay width of W boson. m_W, σ_W is taken from PDG[4].

Missing momentum selection Missing momentum constraint should be imposed on the basis of whether the events have neutrino or not in the final state. Table 3 summarizes the missing momentum constraints. In the case of the event sample with 3 leptons, lower bound cut will be afraid of losing the signal acceptance though background events cannot be suppressed. Therefore, only upper bound cut will be imposed to reject the backgrounds with very large momentum neutrino.

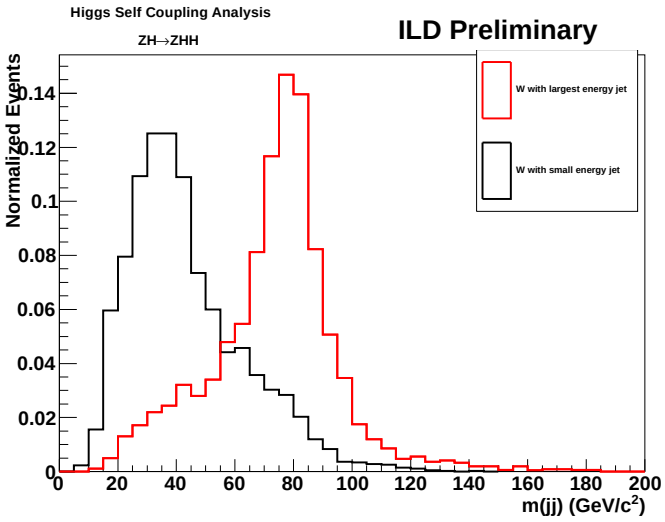

Figure 3: W mass distribution. Jet matching between reconstructed jets and generator level quarks is performed. Red and black histograms are normalized to unity.

lepton number	cut
0	P(Miss)<80 GeV/c
1	P(Miss)>15 GeV/c
2	P(Miss)<80 GeV/c
3	P(Miss)<150 GeV/c

Table 3: Lepton selection

3.2.2 Making Event Samples

With the number of leptons, all the events are classified into the orthogonal samples. The samples have from 0 to 3 leptons. In the case of the event samples with 2 or 3 leptons, lepton pair from Z boson should be required. The leptons are regarded as Z boson daughters when they have same flavor(electrons or muons) and the opposite charge each other.

3.3 Soft Jet Discrimination

The analysis has the disadvantage of b-tagging over the events with $HH \to (b\bar{b})(b\bar{b})$ because one of the Higgs boson decays not into b quark pair but into 2 W bosons. Hence background events cannot be suppressed to excellent degree even if b-tagging constraint is imposed. It is necessary to introduce a new idea which compensates this disadvantage. Final decay objects of top pair production are expected to be 6 fermions. If the jet clustering requirement is

8, it is expected that the top pair production have 6 jets from quarks of top decay, and 2 extra jets. These 2 jets would be coming from gluon emission or small jet split from larger jets. On the other hand, signal events have 8 fermion final objects and all the jets would be coming from hard process quarks. In hadron collider analyses[5][6], it is shown that tracks in the gluon jets spread wider than those in the quark jets. We checked the jet substructure of quark jet candidates and gluon jet candidates. The traditional jet shape can be compared to the analyses of hadron collider experiments. The traditional jet shape, $\Psi(\cos\theta)$, can be defined as

$$\Psi(\cos\theta) = \int_0^{\cos\theta} \frac{p(r)}{p_{jet}} dr \tag{4}$$

where $\cos\theta$ is the defined cone angle of a jet, $p(r)$ is the momentum of the track inside the cone, and p_{jet} is the jet momentum. The traditional jet shape indicates how much the tracks inside the cone contribute the jet. Figure 4 shows mean of the event ensemble along the cone. This figure indicates tracks in gluon jets spread wider than those of quark jets, which has the same tendency as that in hadron collider analyses. Using the feature mentioned and some basic variables, such as number of tracks in the jet, number of neutral particles in the jet, energy deposits on hadron/electromagnetic calorimeter, and maximum transverse momentum of charged or neutral track in the jet, hard jet likeliness is formed. The likeliness is obtained using multi variate analysis. Figure 5 shows the distribution of hard jet likeliness. The hard jet and the soft jets samples are divide into 2 categories, which are whether the jet is far from its nearest jet or not. The jets of $\cos\theta_{jet} < 0.75$, where $\cos\theta_{jet}$ is the jet angle with the nearest jet, are relatively isolated from the other jets, and the jets of $\cos\theta_{jet} > 0.75$ are collinear with those nearest jets, and have large shared area with them. The shared area with the nearest jet could make the difference of the traditional jet shape between quark jets and gluon jets deteriorate. It is difficult to use these 2 kinds of hard jet likeliness on the same scale because those classifiers are obtained using different jet samples. It is necessary to re-defined the hard jet likeliness Therefore, the final hard jet likeliness for all jets is defined as

$$d(hardjet) = \frac{d(\cos\theta_{jet} < 0.75) + d(\cos\theta_{jet} > 0.75)}{2}, \tag{5}$$

where $d(hardjet)$ is the final hard jet likeliness, $d(\cos\theta_{jet} < 0.75)$ is a hard jet likeliness for isolated jets, and $d(\cos\theta_{jet} > 0.75)$ is a hard jet likeliness for collinear jets. Figure 6 shows the distributions of the final hard jet likeliness of the jets with small values in the event. This figure indicates the events with tight leptons have smaller value of the hard jet likeliness because those events have fewer number of hard quarks in the event. The hard jet likeliness can be used for one of the input variables for Multi Variate Analysis.

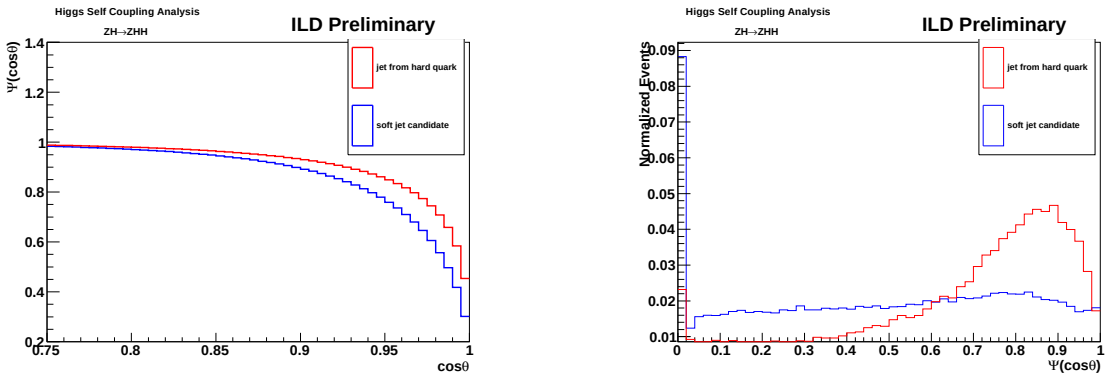

Figure 4: Comparison of the traditional jet shape between jets formed by quarks and soft jet candidates. Left figure shows mean of jets ensemble along the defined cone angle mentioned in the text. Right figure shows the distribution on the third bin from the right side in the right figure. This corresponds to the cone angle of $0.925 \leq \cos\theta \leq 0.95$.

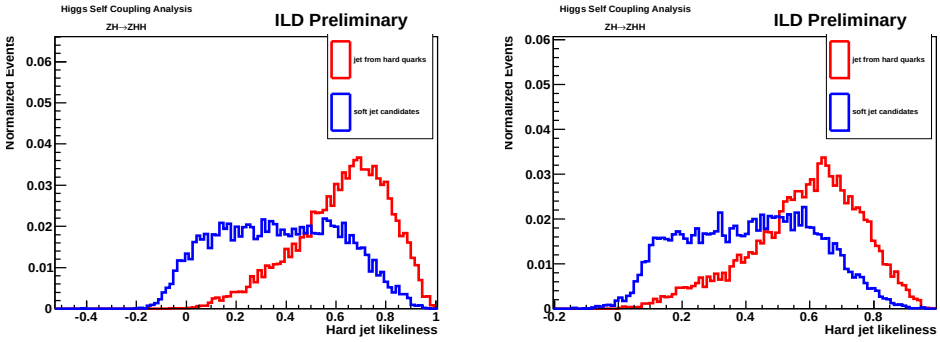

Figure 5: Hard jet likeliness after the training. Left figure is for the isolated jet, which is $\cos\theta_{jet} < 0.75$ and right figure is for the collinear jet, which is $\cos\theta_{jet} > 0.75$

3.4 Background Rejection Using Multi Variate Analysis

It is difficult to reject backgrounds based on the cut based technique because all the physics variables correlate each other. Signal has small cross section and effective background rejection is essential. Therefore, Multi Variate Analysis(MVA) using TMVA framework is used to reject backgrounds. Background rejection should center on the separation of main background. Hence the analysis will proceed firstly to small background rejection, and then into main background rejection, imposing the cut on MVA output for small background rejection because it is necessary to tighten the variable space when separating main background effectively. For example, Figure 7 shows the schematic picture of analysis flow for the case of background rejection all hadronic (number of leptons is 0) case. In this case, ttbar all hadronic events are the main background. For the MVA training, many input variables are

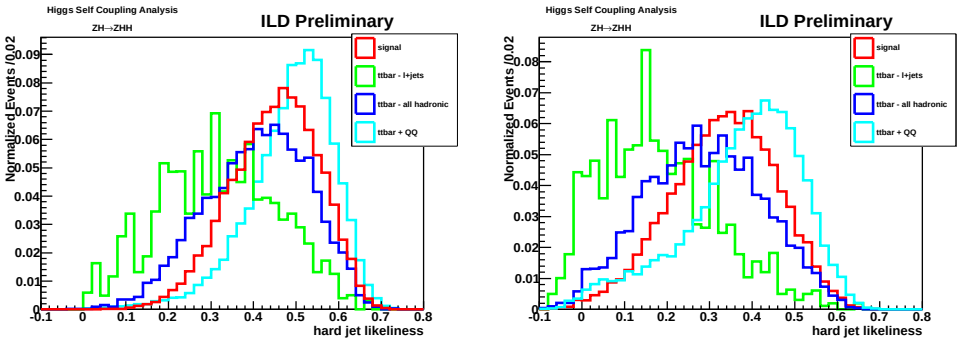

Figure 6: Distributions of the hard jet likeliness of the jets with small values in the event. Left figure is for the jets with the second smallest value and right figure is for the jets with the smallest value. Signal and each background component are normalized to unity.

tested and variable sets are determined to reject backgrounds most effectively. Figure 8 shows are the distribution of some kinematic variables. Powerful variable for background separation is expected to be Higgs mass distribution decaying into W boson pair, $m(jjjj)$ or $m(l\nu jj)$. For the analysis of the events with $Z \to l\bar{l}$, lepton pair mass distribution is a good variable for main background rejection($t\bar{t}$) while the background events with Z boson cannot be suppressed.

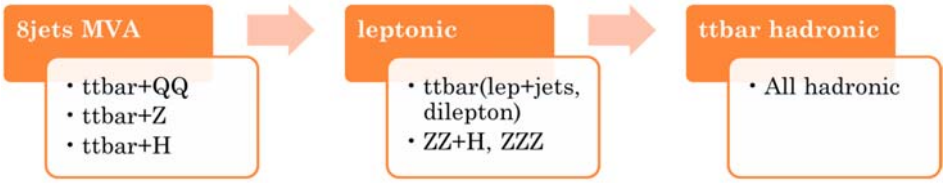

Figure 7: Schematic image of the procedure for classifier training. This is the case of all hadronic events. In this case $t\bar{t}$ which decays into quarks is the main background. $t\bar{t}$ with leptonic decay has large cross section. However it is relatively easy to suppress it.

3.4.1 Variables which Characterize the Process

In addition to basic kinematic variables, some variables which characterize the event feature are introduced as the input variables for MVA.

Sphericity and Aplanarity Sphericity and aplanarity[7][8] are the eigenvalues' combination of the sphericity tensor:

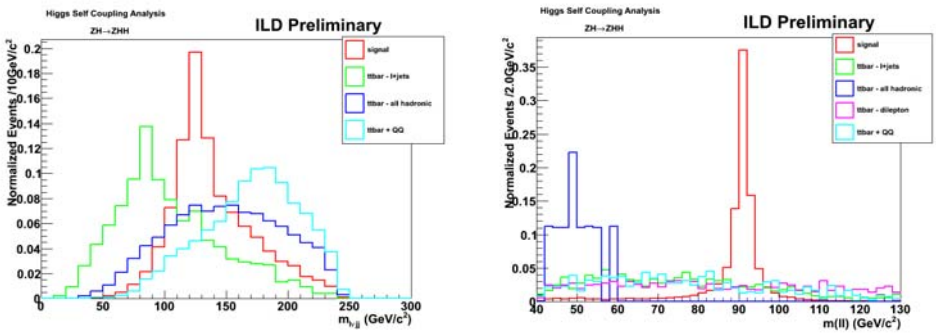

Figure 8: Some kinematic distributions. Left figure is $m(jjjj)$ distribution for all hadronic event analysis. Right figure shows the lepton pair distribution for dilepton sample. These kinematic variables have good power for background discrimination. .

$$S^{\alpha\beta} = \frac{\sum_i p_i^\alpha p_i^\beta}{\sum_i |\mathbf{p}|_i^2} \tag{6}$$

where α, β=1,2,3 correspond to the x, y, and z components. By standard diagonalization of $S^{\alpha\beta}$, three eigenvalues of $\lambda_1 > \lambda_2 > \lambda_3$ can be found. The sphericity S is then defined as

$$S = \frac{3}{2}(\lambda_2 + \lambda_3) \tag{7}$$

so that $0 \leq S \leq 1$. Sphericity is essentially a measure of the sum of the transverse momentum in the event with respect to the event axis. 2-jets event corresponds to $S \sim 0$, and isotropic event to $S \sim 1$. The aplanarity is defined as $A = \frac{3}{2}\lambda_3$, and is constrained the range of $0 \leq A \leq \frac{1}{2}$. It measures the transverse momentum component out of the event plane. A planar event corresponds to $A \sim 0$ and an isotropic event to $A \sim \frac{1}{2}$. Figure 9 shows the sample distribution of the sphericity and aplanarity.

Fox-Wolfram moments The Fox-Wolfram moments ,H_l, $l = 0, 1, 2, ...$, are defined by

$$H_l = \sum_{i,j} \frac{|p_i||p_j|}{E_{vis}^2} P_l(\cos\theta_{ij}), \tag{8}$$

where θ_{ij} is the opening angle between hadrons i and j and E_{vis} is the total visible energy of the event. Note that also autocorrelations, $i = j$, are included. The P_l are the Legendre polynomials, If the total momentum of the event is balanced, H_0 will be 0. Especially H_2 can characterize 3-jet event structure. Figure 9 shows the sample distributions of the Fox-Wolfram moments up to 5th order of the Legendre polynomial.

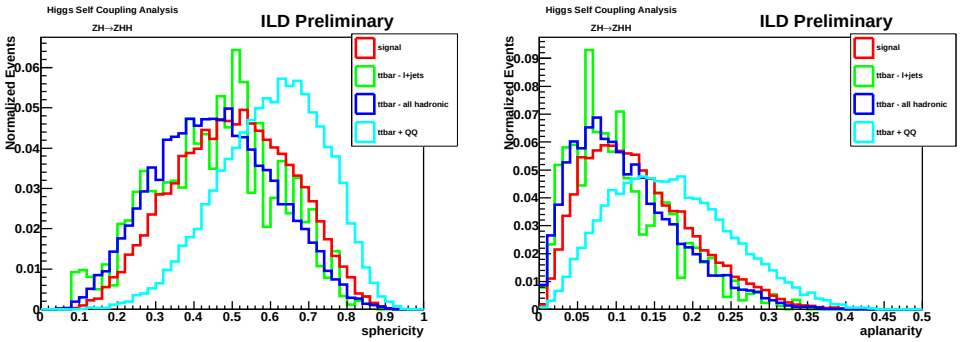

Figure 9: Distributions of sphericity and aplanarity. Signal and each background component are normalized to unity.

3.4.2 MVA Training and Output

To form the classifier, MVA training is performed. Training samples are part of the full samples, and they are exclusive samples from final samples. Figure 11 shows the distribution examples of MVA output.

3.5 Final selection

After the MVA output cut, b likeliness cut is optimized to make the final sensitivity improved. Table 4 summarizes the cuts on the b likeliness for each event sample. On each event, all the jets are sorted into descending order of b likeliness and then the cut is imposed. As for the events with $Z \to b\bar{b}$, all hadronic event samples are classified into 2 b-tagging categories according to the number of b jets in the event. 4-btag events, which have 4 jets with b-likeliness$>$ 0.4, have better sensitivity than the 3-btag events, which have 3 jets with b-likeliness$>$ 0.4. In the case of c tagging, c likeliness is used as one of the input variables for Multi Variate Analysis instead of imposing the c likeliness cut.

4 Results

Signal and background events are estimated after imposing all the selection. It is assumed that an integrated luminosity is 2ab^{-1} with the polarization of $P(e^-, e^+) = (-0.8, +0.3)$. Table 5 summarizes the number of events expected after the selection for each category. And the significance, $S/\sqrt{S+B}$ is estimated.

Finally, the combined significance in the process of $HH \to (b\bar{b})(WW^*)$ is 1.91σ, assuming the Higgs mass of 125 GeV/c^2 and the integrated luminosity of 2ab^{-1} with the polarization $P(e^-, e^+) = (-0.8, +0.3)$. This analysis just starts. Therefore there is much

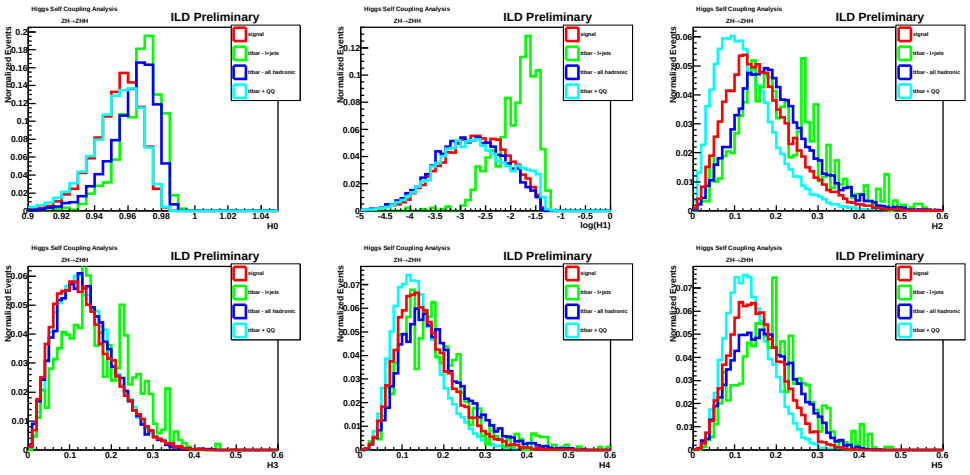

Figure 10: Distributions of Fox-Wolfram moments. Signal and each background component are normalized to unity. The order of the Legendre polynomial increases from top left to bottom right figure.

room to improve. Especially, b-tagging strategy can add the signal acceptance to gain the final significance.

5 Summary

We are analyzing the measurement of the Higgs self-coupling at ILC at center-of-mass energy of 500GeV. The Higgs strahlung process with the decay of $HH \rightarrow (b\bar{b})(WW)$ are being analyzed to give the contribution to the final sensitivity of the Higgs self-coupling. At 500 GeV, assuming an integrated luminosity of 2 ab^{-1} and the Higgs mass of 125 GeV/c^2, the signal significance of $e^+e^- \rightarrow ZHH \rightarrow Z(b\bar{b})(WW)$ process reaches 1.91σ.

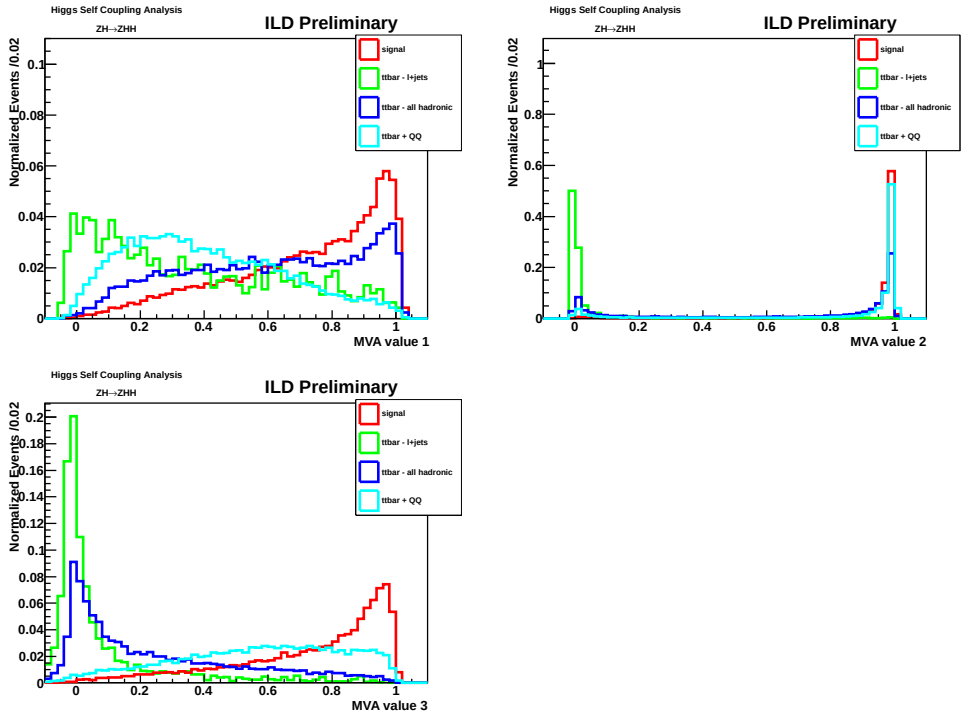

Figure 11: Distributions of the MVA output. Upper left figure shows the classifier to reject small backgrounds ($t\bar{t}+QQ$, $t\bar{t}+Z$, and $t\bar{t}+H$). Upper right figure shows the classifier to reject the backgrounds with leptonic decay ($t\bar{t}$ with leptonic decay, $ZZ+H$, and ZZZ), and lower left figure is the classifier for hadronic backgrounds rejection. These MVA classifiers are used for all hadronic mode analysis.

Category	All hadronic		
	$Z\to b\bar{b}$ 4-btag	$Z\to b\bar{b}$ 3-btag	$Z\to c\bar{c}$
Cut	btag(3^{rd}) >0.86	btag(2^{nd}) >0.92	N/A
Category	Lepton + jets		Dilepton + jets
	$Z\to b\bar{b}$	$Z\to c\bar{c}$	
Cut	btag(2^{nd}) >0.80 btag(3^{rd}) >0.66	btag(3^{rd}) >0.86	N/A
Category	Trilepton + jets		
Cut	N/A		

Table 4: Cut for final b-tagging. In the case of all hadronic events with $Z\to b\bar{b}$, events are classified with the number of b tagged jets in the event. Left b-tag cut condition is for 4-btag events, and right b-tag cut condition is for 3-btag events.

Energy(GeV)	Modes	Z decay	Signal	Background	Significance
500	All hadronic	$Z \to b\bar{b}$ 4-btag	15.20	87.52	1.50σ
		$Z \to b\bar{b}$ 3-btag	19.43	3099.49	0.35σ
		$Z \to c\bar{c}$	11.29	366.13	0.58σ
500	Lepton + jets	$Z \to b\bar{b}$	1.65	17.62	0.38σ
500		$Z \to c\bar{c}$	0.88	146.09	0.04σ
500	Dilepton	$Z \to l\bar{l}$	2.24	8.44	0.69σ
500	Trilepton	$Z \to l\bar{l}$	1.05	2.60	0.55σ
combined					1.91σ

Table 5: Summary of the results of $HH \to (b\bar{b})(WW^*)$ analysis at 500 GeV of center-of-mass energy.

References

[1] G. Gounaris, D. Schildknecht, and F. Renard: Phys.Lett. **B83**, 191 (1979)

[2] A. Djouadi, H. Haber, and P.Zerwas: Phys.Lett. **B375**, 203 (1996)

[3] V.Ilyin, A. Pukhov, Y.Kurihara, Y. Shimizu, and T.Kaneko: Phys.Rev. **D54**, 6717 (1996)

[4] http://pdg.lbl.gov/

[5] Jason Gallicchio, and Matthew D. Schwartz: arXiv: 1211.7038, (2012)

[6] Jason Gallicchio, and Matthew D. Schwartz: arXiv: 1001.5027, (2010)

[7] Michele Faucci Giannelli: arXiv: 0901.4895, (2009)

[8] http://cepa.fnal.gov/psm/simulation/mcgen/lund/pythia_manual /pythia6.3/pythia6301/node2.html