Study of Higgs self-coupling at the ILC based on the full detector simulation at $\sqrt{s} = 500$ GeV and $\sqrt{s} = 1$ TeV

Junping Tian

High Energy Accelerator Research Organization (KEK), Tsukuba, Japan

(Dated: April 19, 2013)

In this analysis we investigated the feasibilities of the measurement of Higgs self-coupling at the International Linear Collider (ILC) during its two phases of operation with centre-of-mass energy of 500 GeV and 1 TeV. Three combinations of the decay modes of double Higgs strahlung process $e^+e^- \to ZHH$, where $Z \to l^+l^-$, $Z \to \nu\bar{\nu}$ and $Z \to q\bar{q}$ accompanying with both Higgs decay into $b\bar{b}$, were analyzed together at 500 GeV. The decay mode of WW fusion process $e^+e^- \to \nu\bar{\nu}HH$, where both Higgs decay into $b\bar{b}$ was analyzed at 1 TeV. Both the signal and background event samples are generated by a full detector simulation based on the International Large Detector (ILD). At 500 GeV, assuming an integrated luminosity of 2 ab^{-1} and the Higgs mass of 120 GeV, an excess of the $e^+e^- \to ZHH$ events with a statistical significance of 5.0σ is expected to be observed in case of the polarized electron and positron beams, $P(e^-, e^+) = (-0.8, +0.3)$. The cross section of $e^+e^- \to ZHH$ can be measured to the precision of 27%, corresponding to the precision of 44% on the Higgs trilinear self-coupling. At 1 TeV, in case of $P(e^-, e^+) = (-0.8, +0.2)$, we can expect the precision of self-coupling to be 18%.

Contents

I. Introduction 2

II. Measuring the Higgs self-coupling at the ILC 2

III. Simulation Framework 3

A. Accelerator and Detector 5

B. Event Generator 5

C. Simulation and Reconstruction 5

IV. Analyses of $e^+e^- \to ZHH$ at 500 GeV 5

A. Analysis of the mode $ZHH \to l^+l^-HH \to l^+l^-b\bar{b}b\bar{b}$ at 500 GeV 5

1. Pre-selction 5

2. Final Selection 9

3. Results 15

4. Summary of the $llHH$ mode 16

B. Analysis of $e^+ + e^- \to \nu\bar{\nu}HH \to \nu\bar{\nu}b\bar{b}b\bar{b}$ at 500 GeV 17

1. Summary of the $\nu\nu HH$ mode 18

C. Analysis of $e^+ + e^- \to q\bar{q}HH \to q\bar{q}b\bar{b}b\bar{b}$ at 500 GeV 18

1. Summary of the $qqHH$ mode 19

V. Combined Result of $e^+e^- \to ZHH$ at 500 GeV 19

A. Statistical independence of the three modes 20

B. Combined ZHH excess significance 20

C. Extracting the Cross Section of ZHH 21

VI. Analysis of $e^+e^- \to \nu\bar{\nu}HH$ at 1 TeV 21

1. Summary of the $\nu\nu HH$ at 1 TeV 22

VII. Summary 23

Acknowledgments 23

References 23

I. INTRODUCTION

The Higgs sector is the piece of Standard Model which is responsible for the spontaneous breaking of electroweak symmetry and offers the source of mass generation for both the gauge bosons and fermions; the expected Higgs boson is the last particle of the Standard Model to be found by experiment. Once a Higgs-like boson is discovered, we need to verify that it is indeed the Higgs boson that condenses in the vacuum and gives masses to all the standard model particles. Higgs self-coupling is just the force that makes the Higgs boson condense in the vacuum, therefore probe of this coupling is one of the most decisive tests of the Higgs sector.

In the Standard Model, after the electroweak symmetry breaking, the Higgs potential, 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, $v \approx 246$ GeV is the vacuum expectation value of the neutral component of Higgs field and λ is the Higgs self-coupling, is uniquely determined by the 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 $\lambda_{HHHH} = 6\lambda$. Considering that all the interactions discovered up to now are gauge interactions, the second and third terms predict non-gauge interactions, which would be a completely new type of interaction. To fully verify the shape of Higgs potential, we need to measure these three terms respectively. The mass term is possible to first be measured at the Tevatron and the LHC and then precisely determined at the ILC. The quartic Higgs self-coupling turns out to be very difficult to be measured at the Tevatron, LHC and the ILC due to the very small cross section of three Higgs bosons production (less than 0.001 fb). Therefore, it becomes crucial to investigate the feasibilities of measuring the trilinear Higgs self-coupling.

On the other hand, alternatively to the Higgs Sector in the standard model, which is the simplest way spontaneously breaking the electroweak symmetry, there are several extended Higgs theories. To reveal these new physics models, the Higgs self-coupling is one of the most important discriminative quantities. One of the latest articles [1] gives a clear conclusion from the theoretical calculation in those extended theories that the Higgs self-coupling measurement is quite useful to explore new physics. In addition, studies of the Higgs self-coupling in the framework of the Minimal Supersymmetric extension of the Standard Model (MSSM) or the general Two Higgs Doublet Model (THDM), where typically there are five Higgs bosons instead of one, can be found in these references [1–4], showing that the Higgs self-coupling can significantly deviate from the standard model value, at the level of order 100% in some scenarios. The Higgs self-coupling can also be a common signature of the Higgs sector with the strong first order electroweak phase transition [5–10], which is required for a successful scenario of electroweak baryogenesis [11–13]. The new physics model with the sequential fourth generation fermions also predict large one loop contributions to the Higgs self-coupling [14]. Therefore, the precision measurement of trilinear Higgs self-coupling could directly reveal the nature of the extended Higgs sector.

II. MEASURING THE HIGGS SELF-COUPLING AT THE ILC

The measurement of trilinear Higgs self-coupling can be carried out at the ILC through two leading processes: double Higgs-strahlung process $e^+e^- \to ZHH$ [15–17] and WW fusion process $e^+e^- \to \nu\bar{\nu}HH$ [18, 19]. 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 dominant at around the center of mass energy of 500 GeV and to be taken over by the WW fusion process at higher energy at around 1 TeV. Their tree-level Feynman diagrams are respectively shown in Figure 2 and Figure 3. However, in both cases, there exist the irreducible Feynman diagrams which have the same final-state particles but don't concern the Higgs self-coupling. The interferences between the interested Higgs self-coupling related diagrams and these irreducible diagrams make the measurement of the Higgs self-coupling more complicated. As a result of the interferences, the cross sections (σ) of $e^+e^- \to ZHH$ and $e^+e^- \to \nu\bar{\nu}HH$, as a function of the Higgs self-coupling (λ), can be formulated as $\sigma = a\lambda^2 + b\lambda + c$, where constant a comes from the contribution of Higgs self-coupling diagram, c comes from the contribution of the irreducible diagrams and b comes from the contribution of the interference between them. For a particular value of the Higgs mass of $M_H = 120$ GeV, Figure 4 shows these functions, by which we can infer the Higgs self-coupling from the cross sections of the two processes. And at the value of the standard model, the precision of the Higgs self-coupling ($\frac{\delta\lambda}{\lambda}$) is determined to be 1.8 times of the precision of the cross section of $e^+e^- \to ZHH$ ($\frac{\delta\sigma}{\sigma}$) at 500 GeV,

$$\frac{\delta\lambda}{\lambda} = 1.8\frac{\delta\sigma}{\sigma}. \tag{2}$$

In case of $e^+e^- \to \nu\bar{\nu}HH$ at 1 TeV, the factor will be 0.85,

$$\frac{\delta\lambda}{\lambda} = 0.85\frac{\delta\sigma}{\sigma}. \tag{3}$$

Here we see the complication caused by the interference, without which the factor will always be 0.5. A new weighting method developed recently [20] shows we can enhance the coupling sensitivity, as a result of which the above factors can be improved correspondingly to 1.66 and 0.76

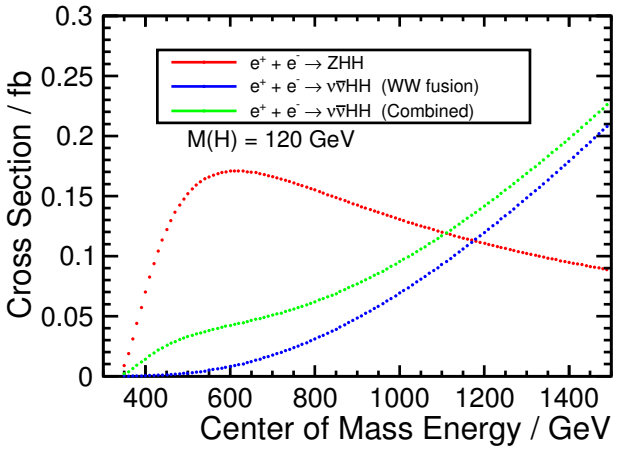

FIG. 1: The separate and combined production cross sections for the ZHH and $\nu\bar{\nu}$HH processes as a function of the center of mass energy assuming the Higgs mass of 120 GeV. The red line is for the ZHH process, the blue line is for the $\nu\bar{\nu}$HH fusion process and the green line is for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

III. SIMULATION FRAMEWORK

[**This Part is to be added later, which nevertheless is common for all the DBD benchmark analyses. In this analysis the $\gamma\gamma$ to low p_t hadrons background has not been overlaid.**]

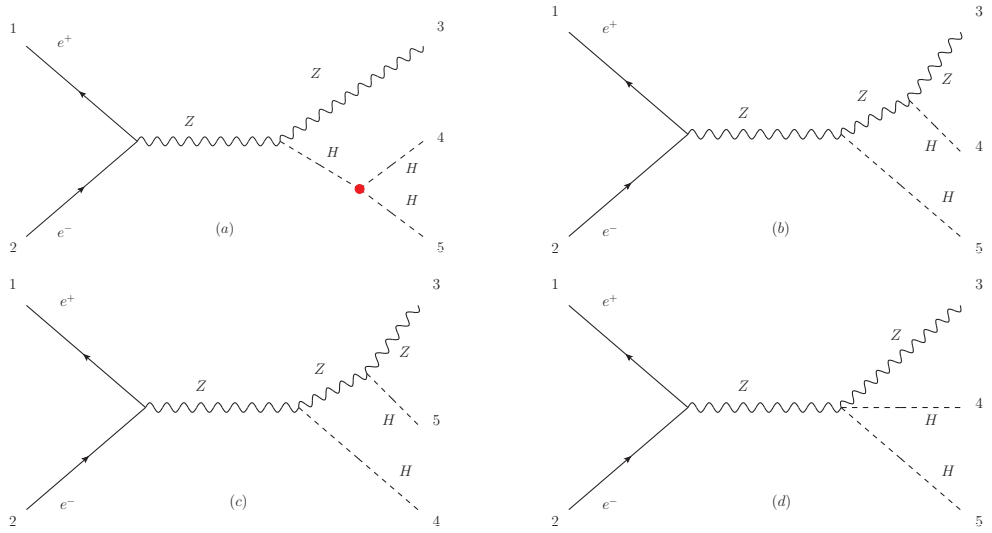

FIG. 2: Feynman diagrams for the double Higgs strahlung process $e^+e^- \to ZHH$. (a): involving trilinear Higgs self-coupling; (b), (c), (d): the irreducible diagrams.

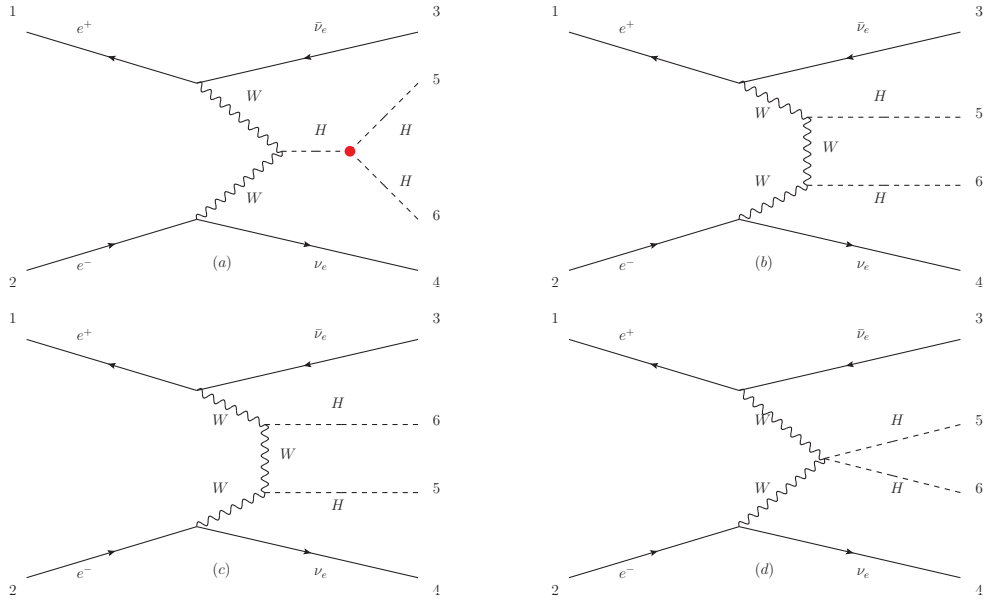

FIG. 3: Feynman diagrams for the WW fusion process $e^+e^- \to \nu\bar{\nu}HH$. (a): involving trilinear Higgs self-coupling; (b), (c), (d): the irreducible diagrams.

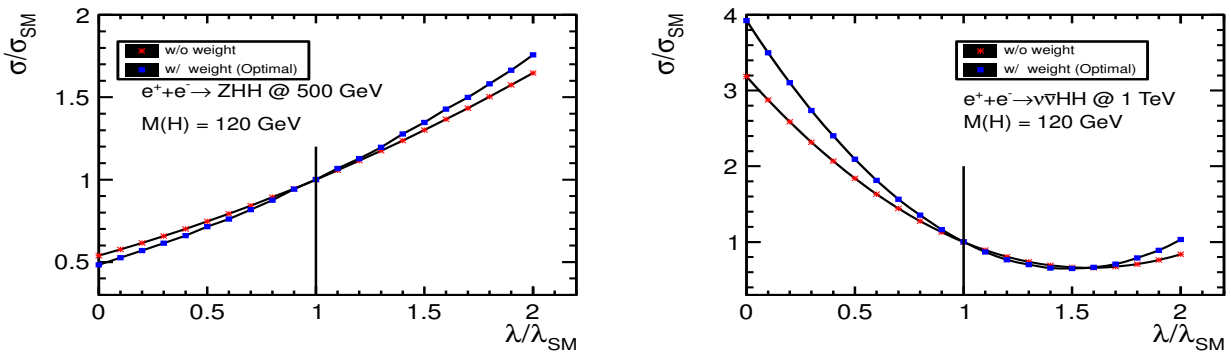

FIG. 4: The evaluation of cross section as a function of the Higgs self-coupling. left: for $e^+e^- \to ZHH$ at 500 GeV; right: for $e^+e^- \to \nu\bar{\nu}HH$ at 1 TeV. The λ_{HHH} stands for the Higgs trilinear self-coupling and $\lambda_{HHH}(SM)$ stands for the standard model value, which is indicated by the vertical line. The red ones are without weighting method and the blue ones are with weighting method.

A. Accelerator and Detector

B. Event Generator

C. Simulation and Reconstruction

IV. ANALYSES OF $e^+e^- \to ZHH$ AT 500 GEV

A. Analysis of the mode $ZHH \to l^+l^-HH \to l^+l^-b\bar{b}b\bar{b}$ at 500 GeV

1. *Pre-selction*

In this search mode, the final state of a candidate signal event contains two isolated charged leptons and four b quarks segmenting into four jets. For the pre-selection, we first require there are two isolated oppositely charged leptons and then force all the particles other than the two selected leptons to four jets and pair the four jets to two Higgs boson candidates.

a. Isolated Lepton Selection Isolated electrons and muons are identified from all of the PFOs. Each PFO contains the information from different sub-detectors, such as energies deposited in the ECAL and HCAL. An electron deposits almost all the energy in ECAL while a muon deposits very small fraction of its energy in both ECAL and HCAL. Other charged particles, being mainly hadrons, deposit most of their energy in HCAL. These pieces of information are used for electron and muon identification. The following two quantities are checked for each PFO. One is the $\frac{E(\text{ecal})}{E(\text{total})}$ ratio, the other is the $\frac{E(\text{total})}{P}$ ratio, where $E(\text{total}) = E(\text{ecal}) + E(\text{hcal})$, $E(\text{ecal})$ and $E(\text{hcal})$ are the energies deposited in ECAL and HCAL and P is the momentum. Some constraints are added to the vertex position which can reduce the selection of the leptons from B-hadrons.

- For electron identification, two samples of PFOs from the process $e^+ + e^- \to e^+e^-HH$ are investigated by using MC truth information. One is the real prompt charged electrons, and the other is all the charged PFOs other than the two prompt charged leptons. Figure 5 shows the distributions of the above two quantities for these two samples, where the red histogram is for the prompt electrons and the blue is for non-prompt charged PFOs. Since the red and blue distributions are very different it is straightforward to add requirements on these two quantities

$$e: \begin{cases} \frac{E(\text{ecal})}{E(\text{total})} > 0.9 \\ 0.8 < \frac{E(\text{total})}{P} < 1.2 \end{cases} \tag{4}$$

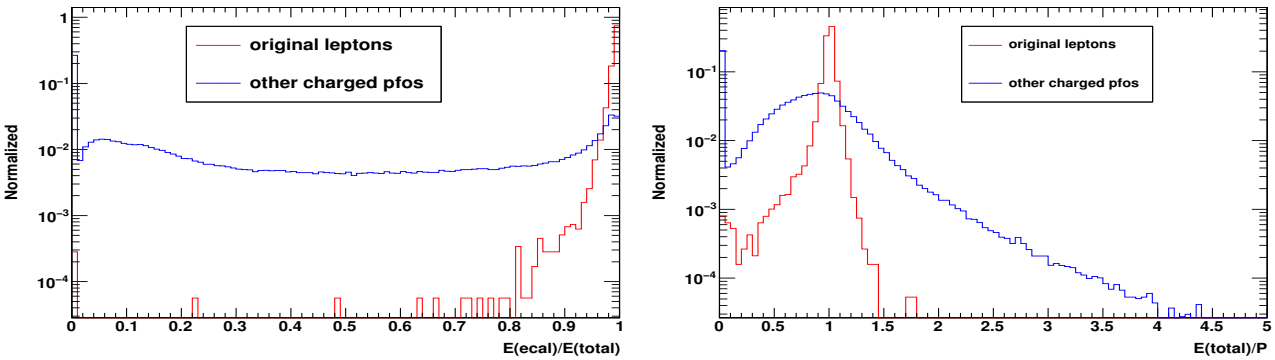

FIG. 5: The distribution of $\frac{E(\text{ecal})}{E(\text{total})}$ (left) and $\frac{E(\text{total})}{P}$ (right) for PFOs in sample $e^+ + e^- \to e^+e^-HH$. The red histogram is for prompt electrons and the blue one is for other charged non-original PFOs.

After these requirements, the mis-identified electrons are mainly from: (i) charged pions, which become neutral pions through charge exchange interaction with the nuclei inside ECAL, decaying into photons which deposit almost all of their energies in ECAL; (ii) electrons from weak decays of b or c quarks, such as $b \to cW^- \to ce^-\bar{\nu}_e$; (iii) electrons from a Higgs boson decaying into WW^* followed by a semi-leptonic W decay. Type (i) and (ii) mis-identified PFOs usually have smaller momenta and more PFOs around them due to parton showering and fragmentation than the prompt electrons. It is hence possible to further reduce the mis-identification by using the cone energy. For each PFO, define a cone with angle θ, around the momentum of that PFO as shown in Figure 6, and sum up the energies of the other PFOs which are inside this cone. This energy sum is called the cone energy. If only the charged PFOs are considered, then the sum is called the charged cone energy. The effect of bremsstrahlung tends to give the prompt electrons a sizable cone energy. This makes the charged cone energy more discriminative to separate the prompt electrons from the other PFOs. Figure 7 shows a scatter plot of charged cone energy versus momentum of the PFOs from these two samples, where the red points denote prompt electrons and the blue points denote the remaining mis-identified non-prompt charged PFOs. By using Fisher classification, we decided to impose

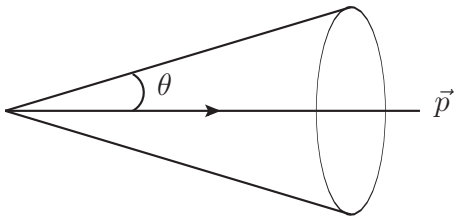

FIG. 6: Schematic view of a cone around a particle with momentum $\vec{p}$. The cone angle is θ.

$$P - 0.25E_{\text{cone}} > 12.6 \text{ GeV} \tag{5}$$

- For muon identification, the strategy is very similar to the electron identification, plus that we require the energy deposited in Yoke to be larger than 1.2 GeV. Samples for prompt muons and the other charged PFOs are from the $e^+ + e^- \to \mu^+\mu^- HH$ process. The distributions of $\frac{E(\text{ecal})}{E(\text{total})}$ and $\frac{E(\text{total})}{P}$ are shown in Figure 8. The requirements to these two quantities are

$$\mu: \begin{cases} \frac{E(\text{ecal})}{E(\text{total})} < 0.5 \\ \frac{E(\text{total})}{P} < 0.3 \end{cases} \tag{6}$$

In this case, the mis-identified muons are mainly from: (i) charge pions which have small momentum and do not reach HCAL, thereby having small energy deposits in ECAL and HCAL; (ii) and (iii) are similar to the electron case, namely from weak decays of b, c quarks and from Higgs decaying into WW^*. Also, the charged cone energy and momentum can be used to further reduce the mis-identification. A scatter plot of the charged

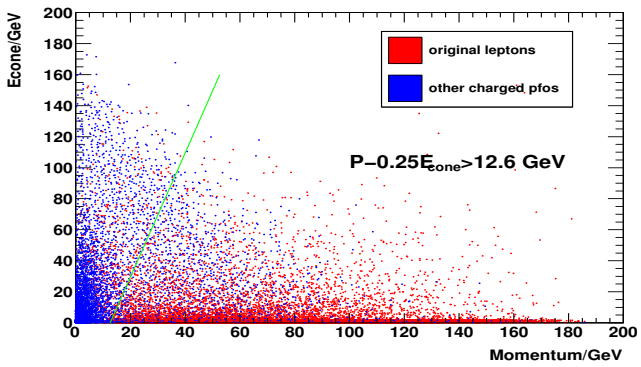

FIG. 7: Scatter plot of charged cone energy versus momentum for PFOs in sample $e^+ + e^- \to e^+e^-HH$. Red points denote original electrons and blue ones denote other charged PFOs.

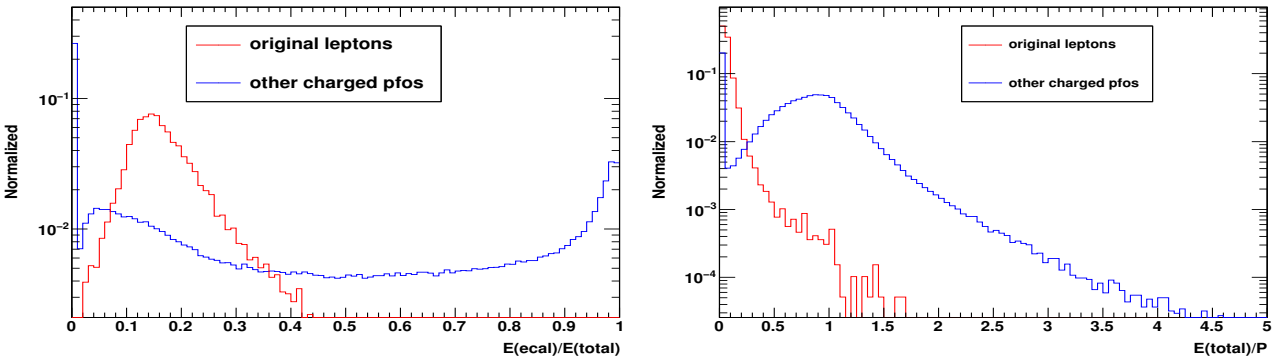

FIG. 8: The distribution of $\frac{E(\text{ecal})}{E(\text{total})}$ (left) and $\frac{E(\text{total})}{P}$ (right) for PFOs in the $e^+ + e^- \to \mu^+\mu^-HH$ sample. The red histogram is for prompt muons and the blue one is for the other charged PFOs.

cone energy versus momentum for the samples of prompt muons and the other non-prompt PFOs are shown in Figure 9. We require

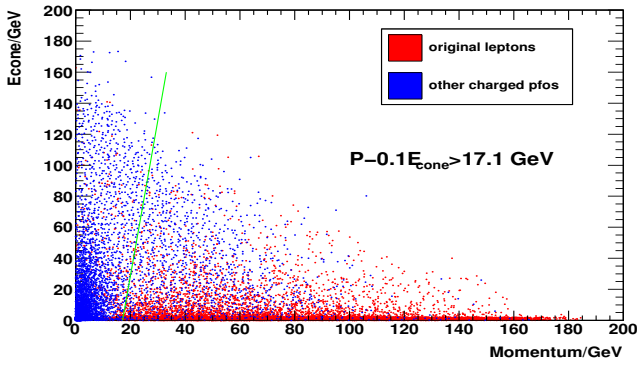

FIG. 9: Scatter plot of charged cone energy versus momentum for PFOs in the $e^+ + e^- \to \mu^+\mu^-HH$ sample. Red points denote prompt muons and blue ones denote the other charged PFOs.

$$P - 0.1E_{\text{cone}} > 17.1 \text{ GeV}. \tag{7}$$

The angle of the cone in Figure 6 is expected to affect the performance of charged lepton selection. To minimize the mis-identification, the value of the cone angle θ is scanned from $\cos\theta = 0.8$ to $\cos\theta = 1$. At each value, while fixing the efficiency for the prompt lepton identification to 98%, we looked at the efficiency of the other charged

PFOs being identified. The result is shown in Figure 10. The optimized cone angle $\cos\theta = 0.98$ giving the minimal mis-identification efficiency is adopted.

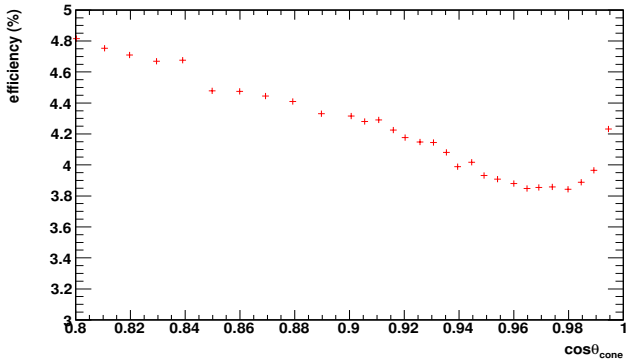

FIG. 10: Optimization of cone angle using PFOs in the $e^+ + e^- \to e^+e^-HH$ sample.

For each event, at least two oppositely charged PFOs are required to be both identified as electron or muon. If there are more than two PFOs identified, we look at all the pairs which have opposite charge. The pair of which the invariant mass is the nearest to the mass of Z, $M(Z) = 91.18$ GeV, is selected as the two prompt charged leptons, effectively suppressing type (iii) mis-identification. As a loose requirement, the invariant mass of the two selected charged leptons $M(l^+l^-)$ should satisfy

$$|M(l^+l^-) - M(Z)| < 40 \text{ GeV}. \tag{8}$$

The distribution of $M(l^+l^-)$ for the signal events after the above selection is shown in Figure 11, where the bremsstrahlung and FSR effects are recovered by using algorithms in ZFinder.

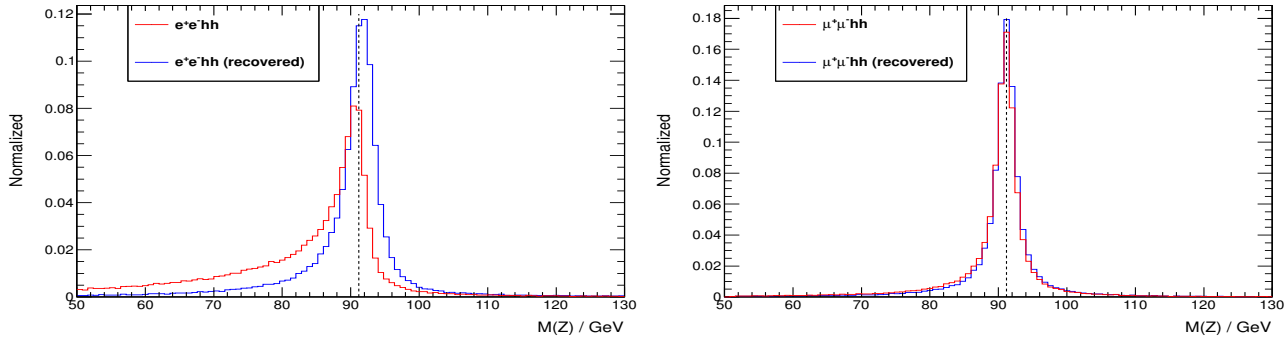

FIG. 11: Invariant mass of the two selected charged leptons. Left one is for the electron mode of the signal $e^+ + e^- \to e^+e^-HH$ process; Right one is for the muon mode of the signal $e^+ + e^- \to \mu^+\mu^-HH$ process.

b. Jet Clustering and Jet Pairing After the two charged leptons are selected, all the other PFOs are forced to four jets by using the Durham jet algorithm. Then the four jets j_1, j_2, j_3 and j_4 are combined to two pairs, each of which contains two jets. Among all the possible combinations, the one which minimizes the χ^2 is selected. The χ^2 is defined by

$$\chi^2 = \frac{(M(j_1,j_2) - M(H))^2}{\sigma_H^2} + \frac{(M(j_3,j_4) - M(H))^2}{\sigma_H^2} \tag{9}$$

where $M(j_1,j_2)$ is the invariant mass of jets j_1 and j_2, $M(H)$ is the nominal Higgs mass, and σ_H is the Higgs mass resolution, which doesn't affect the combination here. The two jets pairs are reconstructed as two Higgs bosons. The order of the two Higgs bosons are determined by the order of jets output from the jet clustering algorithm. Usually a jet output earlier has a relatively larger momentum. While pairing, $M(j_1,j_2)$ and $M(j_3,j_4)$ are required to satisfy loose cuts:

$$|M(j_1,j_2) - M(H)| < 80 \text{ GeV}, \quad |M(j_3,j_4) - M(H)| < 80 \text{ GeV}. \tag{10}$$

2. Final Selection

The remaining signal and background events can be grouped into four: first one, called full hadronic background, such as *bbcsdu*, *bbuddu*, *bbcssc* and *bbbb*, without leptons in the parton level final states; second one, called jets-poor background, such as *llbb*, only two partons with two leptons in the parton level final states; third one, called semi-leptonic background, such as $l\nu bbqq$, with one charged lepton, one missing neutrino and four partons in the parton level final states; the last one, called the most signal-like background, such *llbbbb* and *llbb*H, with two charged leptons and four partons in the parton level final states. Since the event topologies and the amounts of contamination from these four groups are very different, it is not very efficient if we use only one multivariate classification for the whole backgrounds. Actually it is almost impossible to find any global minimum if we put together the backgrounds with very different topologies and very different weights, considering the limited MC statistics. Instead of one multivariate classification, the strategy adopted is to use a separate multivariate classification to suppress the backgrounds in each group.

a. Full Hadronic Backgrounds The full hadronic backgrounds, which mainly come from $e^+ + e^- \to t\bar{t}$, WWZ and ZZ, are significantly suppressed by the pre-selection mainly due to the requirement of two charged leptons. Even though, the number of remaining events is still much larger than that of the signal. For these backgrounds, the selected charged leptons must have come from the hadronization and decay, as a result of which the momenta of the charged leptons should relatively be smaller and the cone energy relatively larger. A natural strategy to further suppress them is to apply tighter cuts on the cone energy and momenta of the two selected charged leptons. Figure 12 shows the scatter plot of the total cone energy ($E_{\rm cone}12$) versus the total charged cone energy ($E_{\rm cone}$Charge12) of the two leptons, and the distribution of the total momentum ($p\text{Lep}_1 + p\text{Lep}_2$). Tighter requirements are imposed to further suppress the full hadronic backgrounds, which is denoted as Cut1:

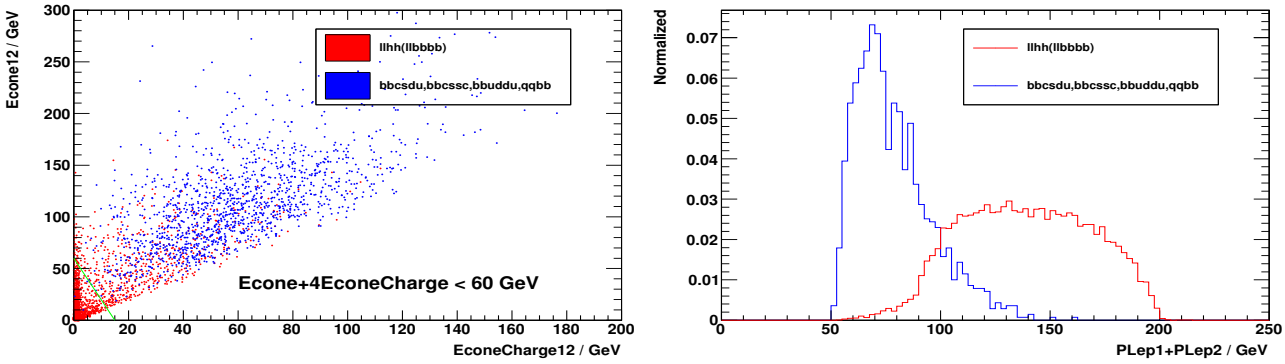

FIG. 12: The scatter plot of $E_{\rm cone}12$ versus $E_{\rm cone}$Charge12(left) and the distribution of $p\text{Lep}_1 + p\text{Lep}_2$ (right). Red denotes signal and blue denotes full hadronic backgrounds.

$$\text{Cut1}: \begin{cases} E_{\rm cone}12 + 4E_{\rm cone}\text{Charge12} < 60 \text{ GeV} \\ p\text{Lep}_1 + p\text{Lep}_2 > 80 \text{ GeV} \end{cases} \tag{11}$$

After these tighter requirements, the full hadronic backgrounds are almost completely eliminated.

b. Jets-Poor Backgrounds The jets poor backgrounds *llbb* and *llcc*, which mainly come from ZZ, ZZ^*, $b\bar{b}Z$ and l^+l^-Z, are the dominant backgrounds after the pre-selection. Though there are only two partons in their parton level final states, due to the gluon to qq and the imperfection of the jet clustering algorithm, they can be clustered to four jets and some of them survived the mass constraints in Eqn. 10. To suppress these backgrounds while keeping as many signal events as possible, one of the multivariate data analysis methods, neural-net is used. The following discriminative quantities are included for the neural-net training:

- Y value, which is given by the jet clustering algorithm. Because there are only two partons for these jets poor backgrounds, their Y values are relatively smaller than that of the signal events. Among all the Y values, $Y_{4\to3}$ and $Y_{3\to2}$ turned out to be the most discriminative. The distributions of $Y_{4\to3}$ and $Y_{3\to2}$ are shown in Figure 13, respectively denoted by "yminus"and "yplus2".

- Thrust, which is derived from the quantity

$$p = \frac{\sum_i |\vec{p}_i \cdot \vec{n}|}{\sum_i |\vec{p}_i|} \tag{12}$$

  where $\vec{p}_i$ is the momentum of a PFO, $\vec{n}$ is any possible unit vector $|\vec{n}| = 1$ and the summation is over all the PFOs in each event. The thrust is defined to be the maximum of p, and the corresponding $\vec{n}$ is called the axis of the event. The thrust value reflects the anisotropy of an event, indicating if there is any special direction favored by this event. Because these jets poor backgrounds are mainly from two-body t-channel processes, most of the PFOs in each event are very forward or backward. Their thrust is much closer to 1 than that of the signal, which is from a three-body process. The axis of these backgrounds is much closer to the beam direction than that of the signal. The distributions of the thrust value and the polar angle of the thrust axis are shown in Figure 13, respectively denoted by "pthrust" and "cosaxis".

- Reconstructed Z mass. Some of these backgrounds are from ZZ, Zγ, or $\gamma\gamma$ fusion processes and from s-channel processes, where two charged leptons in the final states are not from a Z decay. In this case, the reconstructed Z mass does not peak at the nominal Z mass, as indicated by the flat part in the distribution of the reconstructed Z mass in Figure 13, denoted by "mz".

- The total number of PFOs. For this background, the total number of PFOs is much smaller than that of the signal, because there are only two partons. The distribution of this quantity is shown in Figure 14, denoted by "npfos".

- The smallest number of PFOs in a jet. For the same reason, the smallest number of PFOs in a jet is much smaller than that of the signal. The distribution of this quantity is shown in Figure 14, denoted by "npfomin".

- The largest jet momentum when reconstructed as two jets. If we force the PFOs other than the two selected charged leptons to two jets, the momenta of these jets for the background will be relatively larger than signal. The Distribution of this quantity is shown in Figure 14, denoted by "pjmaxjets2".

- The largest angle between the reconstructed Z and the other two jets. Some of these backgrounds come from $e^+ + e^- \to b\bar{b}Z$, where Z is radiated from one of the two b partons. In this case, the Z is very close to one of the two b jets. The distribution of this quantity is shown in Figure 13, denoted by "cosjzmax2".

These quantities are used as input variables by the MLP method in the TMVA package [21]. A neural-net is trained for the signal and the *llbb* background. For the neural-net training, additional statistically independent signal e^+e^-HH, $\mu^+\mu^-HH$ and background *llbb* samples are used. The weights for different processes are normalized to the corresponding cross sections. The statistics of the training samples are higher than 2 ab^{-1} for both the signal and the background. The neural-net outputs and cut efficiencies for the signal and the background are shown in Figure 15. The *llbb* background is well separated by the neural-net output (MLP$_{llbb}$). A cut, MLP$_{llbb} > 0.56$, is imposed to suppress the *llbb* background, denoted by MLP1. Though the neural-net is trained for the *llbb* background, another jets-poor background *llcc* is also significantly suppressed by this cut.

c. Semi-leptonic Backgrounds The semi-leptonic backgrounds such as $e^-\bar{\nu}b\bar{b}c\bar{s}$, $e^-\bar{\nu}b\bar{b}u\bar{d}$, and their corresponding muon or tau modes, together with their conjugate modes, are mainly from $t\bar{t}$ and W^+W^-Z. After the pre-selection, they are the second dominant backgrounds, being hundreds times more than the signal events. Unlike the jets-poor backgrounds, these semi-leptonic backgrounds have four quarks, but only one prompt charged leptons. We trained another neural-net to suppress these backgrounds by using the following quantities:

- Visible energy and missing P_t. Because there's one prompt neutrino in the backgrounds, the visible energy is smaller and the missing P_t is larger for the backgrounds than for the signal. The distributions of these two quantities are shown in Figure 16, respectively, denoted by "evis" and "mpt".

- Cone energy and momentum of the lower momentum selected charged lepton. Because there's only one prompt charged lepton in the backgrounds, the other selected charged lepton must have originated from hadronization and decay, which has larger charged cone energy and smaller momentum. The distributions of these two quantities are shown in Figure 16, respectively, denoted by "econec2" and "plmin".

- Reconstructed Z mass. The invariant mass of the two selected charged leptons should be very different for the backgrounds, as indicated in Figure 17, denoted by "mz".

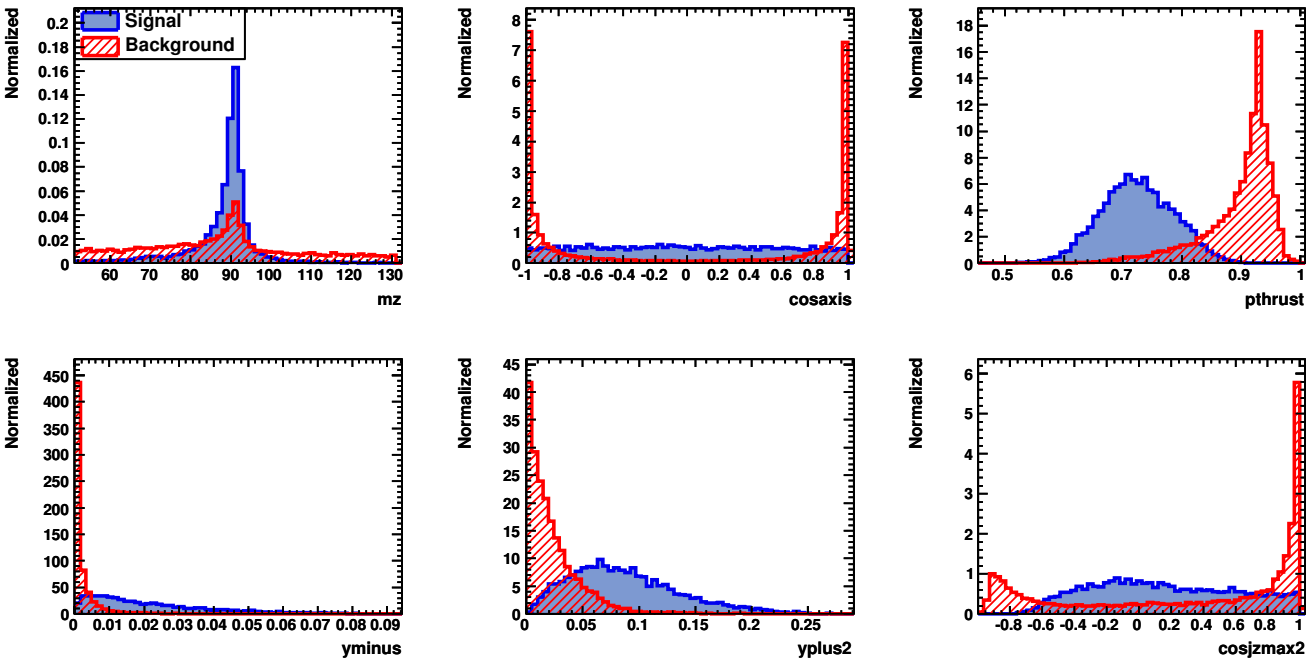

FIG. 13: Discriminative quantities for the signal (blue) and the jets-poor backgrounds *llbb* (red). The variable names are explained in the text.

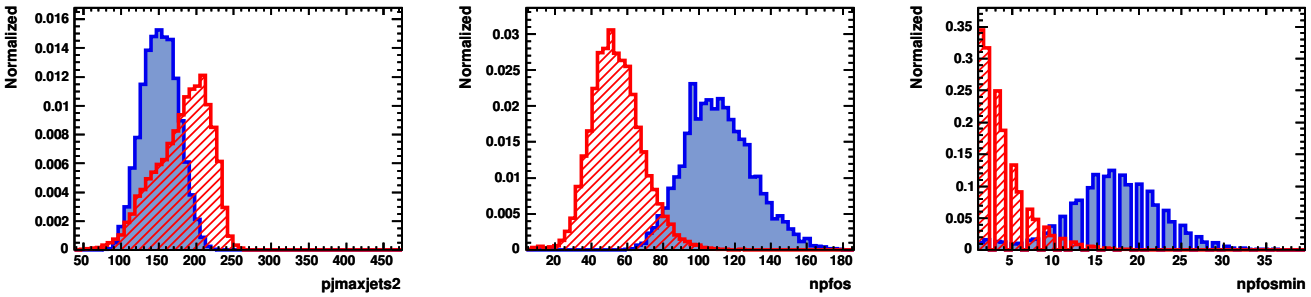

FIG. 14: Discriminative quantities for the signal (blue) and the jets-poor backgrounds *llbb* (red). The variable names are explained in the text.

- The total number of PFOs. During parton showering, a b quark usually results in more particles than light quarks. So the total number of PFOs for the backgrounds is smaller than that of the signal. This information is independent of the b tagging algorithm. The distribution of this quantity is shown in Figure 17, denoted by "npfos".
- Reconstructed W mass. The four jets are ordered from the largest b-likeness to the smallest. The backgrounds contain two b quark jets and two light quark jets. The two light quarks are from a W decay. The invariant mass of the 3rd and 4th jets are reconstructed as the W mass. The distribution of this quantity is shown in Figure 16, denoted by "massb34".
- Angle between two b jets. A large fraction of these backgrounds come from $t\bar{t}$, where the angle between two prompt b jets is relatively large. The angle between 1st and 2nd jets are reconstructed as the angle between the two prompt b jets. The distribution of this quantity is shown in Figure 16, denoted by "cosbmax".

Statistically independent $llHH$ signal sample and the $lvbbqq$ background samples are used for the neural-net training. The statistics are higher than 2 ab^{-1} for both the signal and the background. The neural-net outputs and cut efficiencies for signal and background are shown in Figure 18. The $lvbbqq$ background can well be separated by the

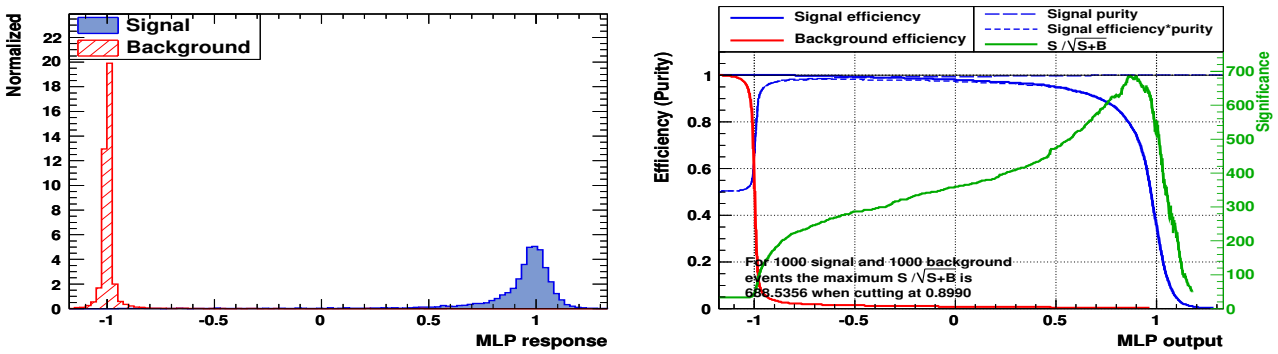

FIG. 15: (left): The neural-net output for the signal versus the *llbb* background, where the blue histogram is for the signal and the red one is for *llbb*. (right): The cut efficiencies for the signal and background at different cut values on neural-net output, where the solid blue curve is for the signal and the red one is for *llbb*.

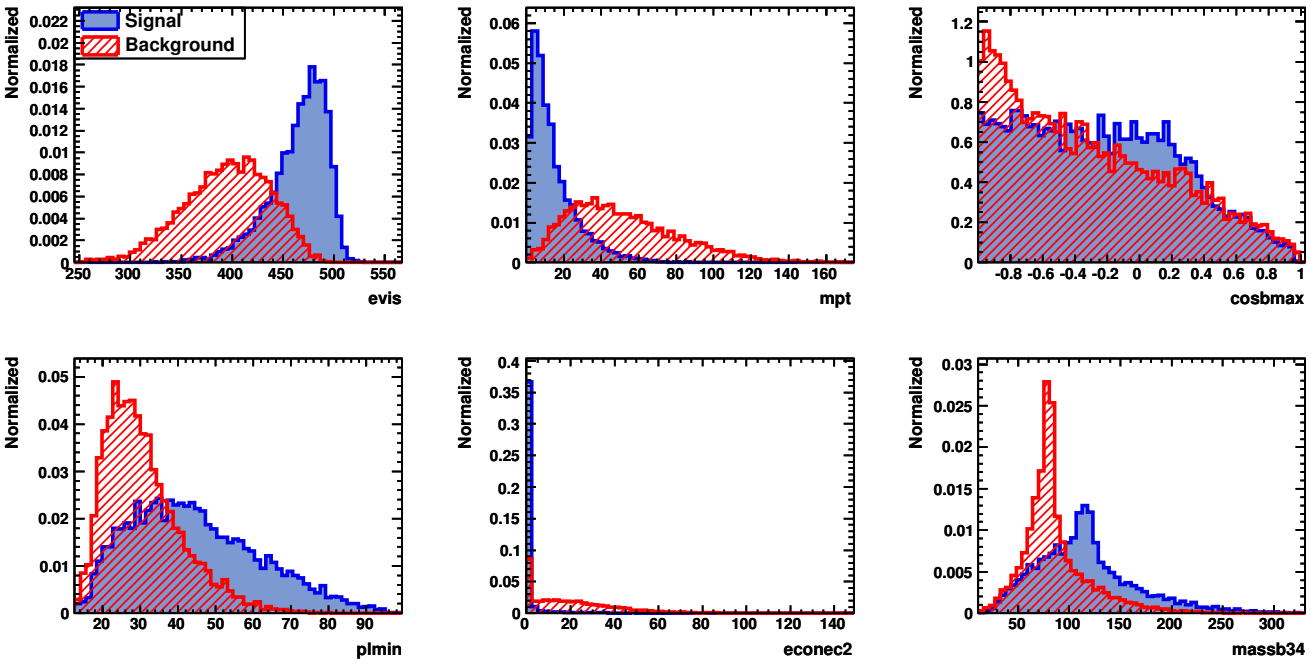

FIG. 16: Discriminative quantities for the signal (blue) and the semi-leptonic backgrounds *lvbbqq* (red). The variable names are explained in the text.

neural-net output (MLP_{lvbbqq}). A cut, $\mathrm{MLP}_{lvbbqq} > 0.81$, is applied to suppress the *lvbbqq* background, denoted by MLP2.

d. Backgrounds with Same Final States These backgrounds including *llbbbb* and *llbbH* mainly come from *ZZZ* and *ZZH*. Though their cross sections are not as large as the previous backgrounds, they have the same parton level final states as the signal, and, consequently, are more difficult to suppress. The quantities used in the previous neural-nets are of little use, requiring quantities related to the invariant mass and angular distributions to suppress them. For this purpose yet another neural-net is trained using the following quantities:

- Reconstructed Higgs mass. The two Higgs bosons masses should be the most discriminative to separate these backgrounds. The distributions of these two quantities are shown in Figure 20, respectively, denoted by "mh1" and "mh2".
- Reconstructed Z, H and Z, Z masses. In order to take maximal use of the mass information, in addition to

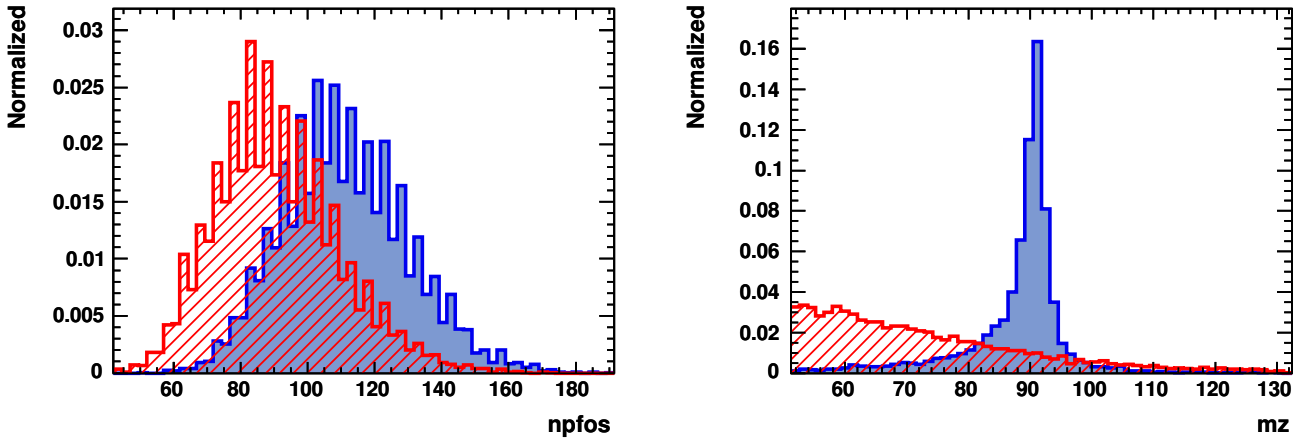

FIG. 17: Discriminative quantities for the signal (blue) and the semi-leptonic backgrounds $lvbbqq$ (red). The variable names are explained in the text.

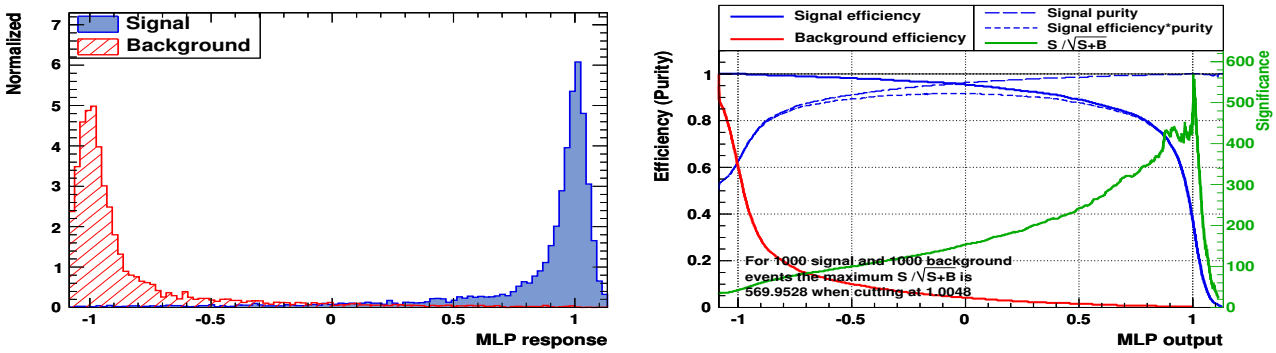

FIG. 18: (left): The neural-net output for the signal versus the $lvbbqq$ background, where the blue histogram is for the signal and the red one is for $lvbbqq$. (right): The cut efficiencies of the signal and background at different cut values of the neural-net output, where the solid blue curve is for the signal and the red one is for $lvbbqq$ background.

the two Higgs boson masses reconstructed as from the signal process, the four jets are also paired as from the l^+l^-ZH and l^+l^-ZZ processes. The reconstructed Z and Higgs masses in the case of l^+l^-ZH pairing, though correlated with the two Higgs masses in the case of l^+l^-HH pairing, can offer some additional discriminative power to suppress the $llbbH$ background. The distributions of these reconstructed Z and Higgs masses are shown in Figure 20, respectively, denoted by "mzzh" and "mhzh". Similarly, the reconstructed two Z masses in the case of l^+l^-ZZ pairing are useful to suppress the $llbbbb$ background, distributions of which are shown in Figure 20, respectively, denoted by "mz1zz" and "mz2zz".

- t-channel characteristics. The processes $e^+ + e^- \to ZZZ$ and $e^+ + e^- \to ZZH$ are dominated by diagrams stem from the t-channel process $e^+ + e^- \to ZZ$, with one more Z boson strahlung from the electron line or one more Higgs strahlung from a Z, as shown in Figure 19. This feature makes the two Z bosons emitted from the electron line move very fast and very forward. To effectively use this information, each event is re-reconstructed as from ZZZ or ZZH, and the boson candidate with the largest momentum among the three is identified for the both hypotheses. The largest momentum and its polar angle reflect the t-channel characteristics, distributions of which are shown in Figure 21, respectively, denoted by "p1zzz" and "cos1zzz"in case of ZZZ, "p1zzh" and "cos1zzh"in case of ZZH.

For the neural-net training, statistically independent $llHH$ signal samples and the $llbbbb$ and $llbbH$ background samples are used, with each sample having statistics higher than 2 ab^{-1}. The neural-net outputs and cut efficiencies for the signal and backgrounds are shown in Figure 22. The $llbbbb$ and $llbbH$ backgrounds are not as well separated as

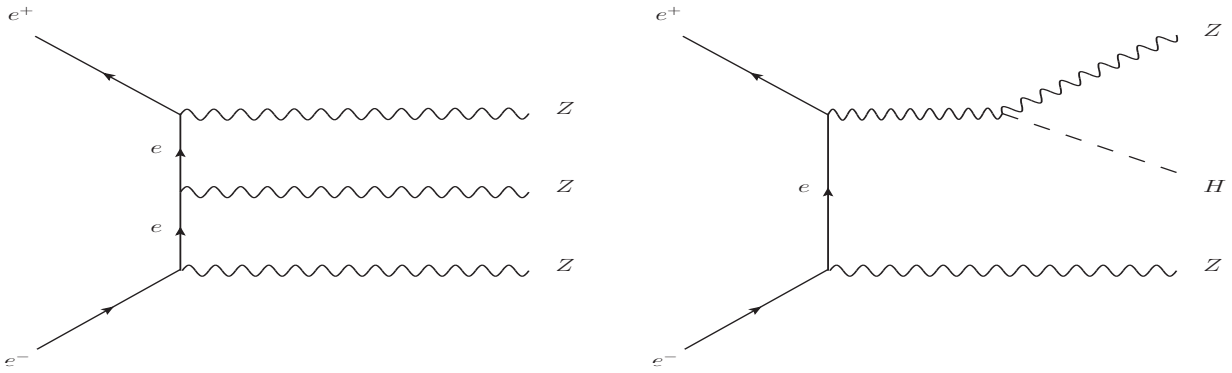

FIG. 19: Typical Feynman diagrams for $e^+ + e^- \to ZZZ$ (left) and $e^+ + e^- \to ZZH$(right).

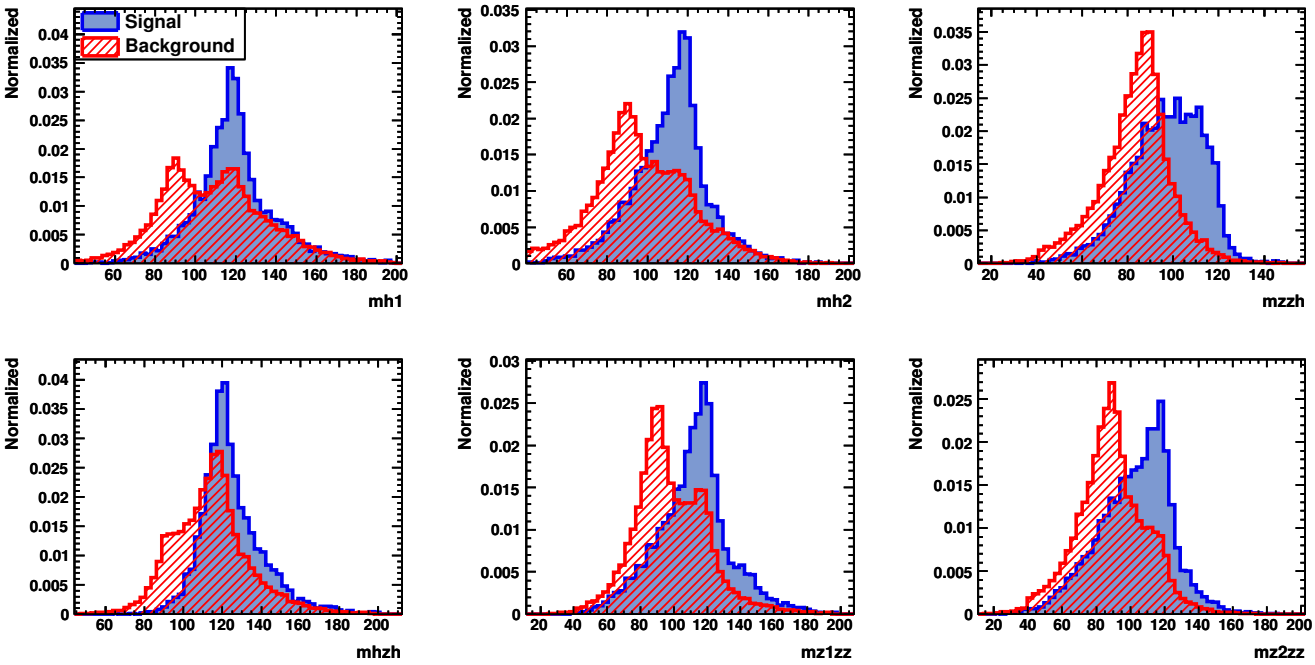

FIG. 20: The discriminative quantities for the signal (blue) and the same final states backgrounds *llbbbb*, *llbbH* (red). The variable names are explained in the text.

the previous two backgrounds. A requirement to the neural-net output $\mathrm{MLP}_{llbbbb} > -0.5$ is imposed to suppress the same final states background, denoted by MLP3. Because the two Higgs masses are the most discriminative quantities in this neural-net, this cut is also effective to suppress all the other backgrounds.

e. B tagging Though the jets-poor backgrounds and the semi-leptonic backgrounds are significantly suppressed by the neural-net, the number of remaining background events is still much larger than that of the signal. On the other hand, so far we have only considered the backgrounds which contain at least two b quarks in the parton level final states. Information of flavour tagging can be used to eliminate the backgrounds with less than two b quarks, and further suppress the jets-poor and the semi-leptonic backgrounds.

For each jet, three outputs (b-likeness, c-likeness and bc-likeness) are calculated. The signal mode is supposed to have four b jets. The b-likeness of the four jets are investigated. To make the difference between the signal and the background more significant, the four jets are ordered by the b-likeness from the largest to the smallest. The distributions of the four b-likeness values are shown in Figure 23, denoted by Bmax1, Bmax2, Bmax3 and Bmax4, where Bmax1 > Bmax2 > Bmax3 > Bmax4. One can see that Bmax1 is usually large and Bmax4 is usually small for both the signal and the backgrounds. Bmax3 turn out to have the most discriminative power. The following cut on

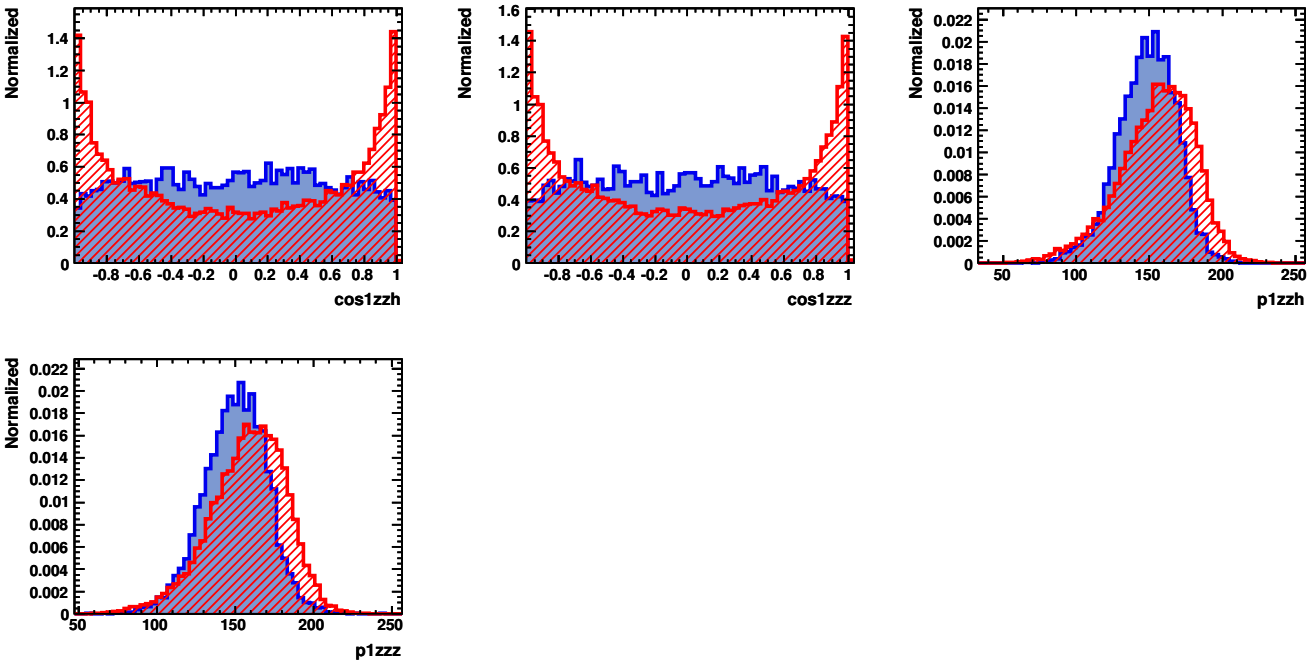

FIG. 21: The discriminative quantities for the signal (blue) and the same final states backgrounds *llbbbb*, *llbbH* (red). The variable names are explained in the text.

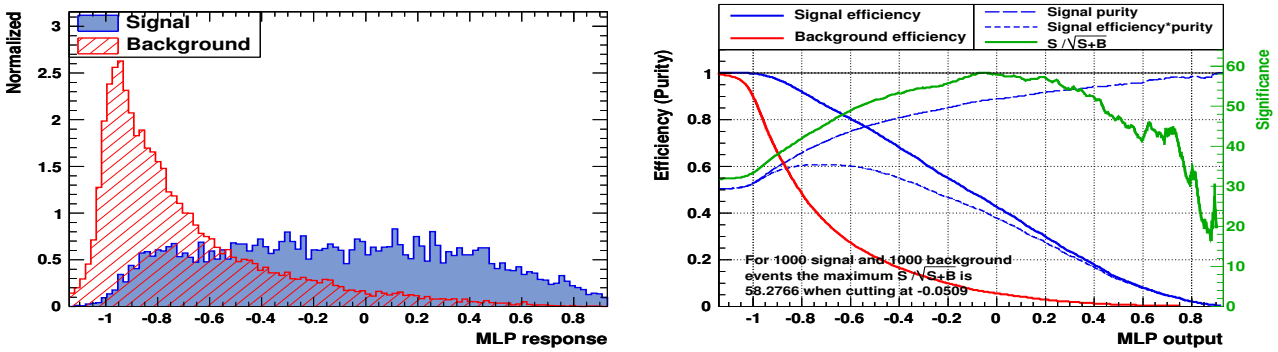

FIG. 22: (left): The neural-net outputs for the signal and the *llbbbb*, *llbbH* backgrounds, where the blue histogram is for the signal and the red one is for *llbbbb* and *llbbH*. (right): The cut efficiencies for the signal and the background at different cut values on the neural-net output, where the solid blue curve is for the signal and the red one is for *llbbbb*, *llbbH*.

Bmax3, denoted by B tagging, is applied to suppress the backgrounds:

$$\text{Bmax3} > 0.19. \tag{13}$$

3. Results

The number of the signal and background events remained after the final selection are shown in the reduction table I and II, where all the events are separated into two categories, electron-type and muon-type. The finally cuts in these two categories are summarized as following:

- For electron-type category, the final cuts are

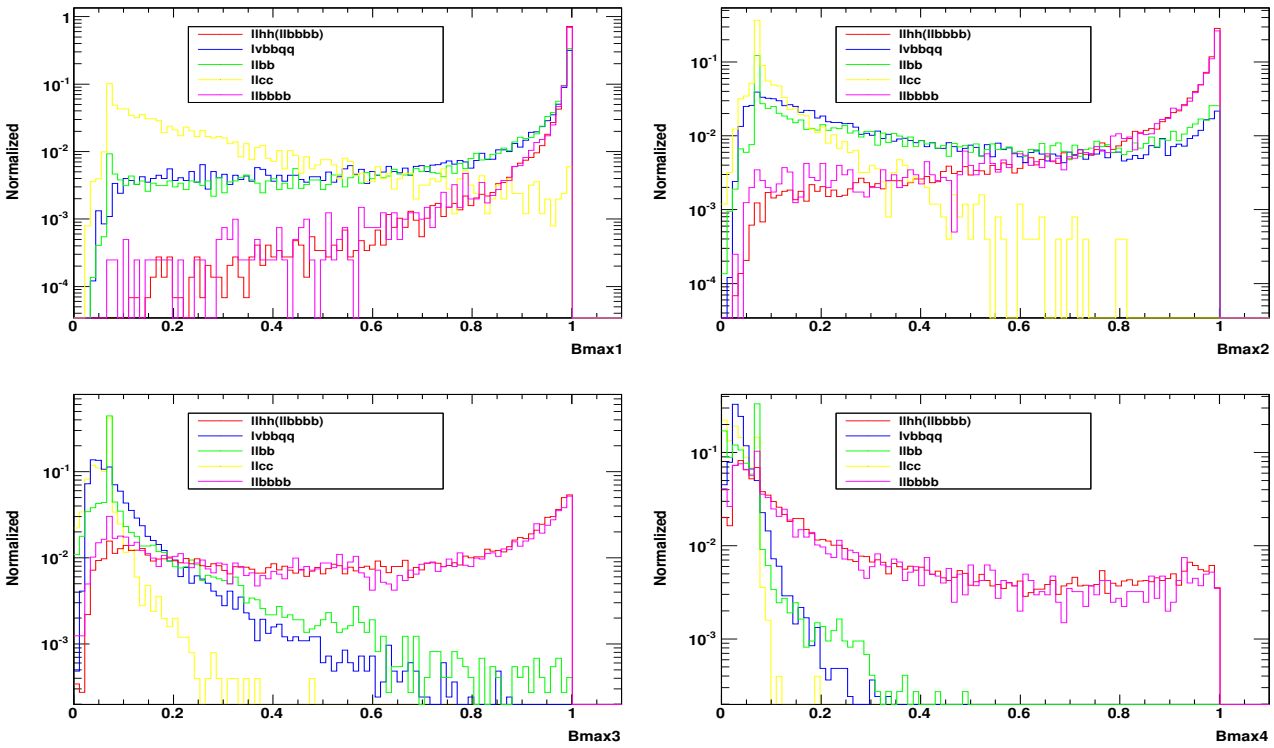

FIG. 23: Distributions of the four b-likeness values, top left for Bmax1, top right for Bmax2, bottom left for Bmax3 and bottom right for Bmax4. Red histograms are for the signal, blue ones for the semi-leptonic background $lvbbqq$, the green and the yellow ones for the jets-poor backgrounds $llbb$ and $llcc$, and pink ones for the same final states background $llbbbb$.

1. Cut1: $E_{\rm cone}12 + 4E_{\rm cone}{\rm Charge12} < 60$ GeV and $|M(ll) - M(Z)| < 32$ GeV.
2. Cut2: ${\rm MLP}_{llbb} > 0.56$.
3. Cut3: ${\rm MLP}_{lvbbqq} > 0.81$.
4. Cut4: $Bmax3 > 0.19$.
5. Cut5: ${\rm MLP}_{llbbbb} > 0.5$.

- For muon-type category, the final cuts are

1. Cut1: $E_{\rm cone}12 + 4E_{\rm cone}{\rm Charge12} < 60$ GeV, $p{\rm Lep}_1 + p{\rm Lep}_2 > 80$ GeV and $|M(ll) - M(Z)| < 27$ GeV.
2. Cut2: ${\rm MLP}_{llbb} > 0.53$.
3. Cut3: ${\rm MLP}_{lvbbqq} > 0.2$.
4. Cut4: $Bmax3 > 0.16$.
5. Cut5: ${\rm MLP}_{llbbbb} > 0.52$.

4. Summary of the llHH mode

In the search mode $e^+e^- \to l^+l^-HH$, assuming a Higgs mass of 120 GeV and an integrated luminosity of 2 ab^{-1} with the beam polarization $P(e^-, e^+) = (-0.8, +0.3)$, category a) it is expected to observe 3.7 signal events with 4.3 backgrounds events, corresponding to a ZHH excess significance of 1.5σ and a ZHH cross section measurement significance of 1.1σ; category b) it is expected to observe 4.5 signal events with 6.0 backgrounds events, corresponding to a ZHH excess significance of 1.5σ and a ZHH cross section measurement significance of 1.2σ. [the definition of excess significance and measurement significance can be found in the appendix.]

TABLE I: The reduction table for the signal and backgrounds after the final selection for the electron-type category of $llHH$ mode, together with the number of expected events and generated events. The cuts names are explained in text.

Process	expected	generated	pre-selection	electron-type	Cut1	Cut2	Cut3	Cut4	Cut5
$llHH$	46.5	3.88×10^5	26.5	13.1	12.3	10.1	8.6	4.64	3.73
$eebb$	2.84×10^5	4.18×10^6	3950	3950	2762	75.4	57.8	3.88	0.81
$\mu\mu bb$	4.96×10^4	1.00×10^6	1944	0.74	0.10	0	0	0	0
$e\nu bbqq$	2.48×10^5	1.51×10^6	2437	2437	928	675	25.7	1.93	0.46
$\mu\nu bbqq$	2.46×10^5	1.48×10^6	239	24.5	0.52	0.36	0	0	0
$\tau\nu bbqq$	2.46×10^5	1.35×10^6	156	148	38.6	30.3	1.50	0.25	0
$bbqqqq$	6.24×10^5	3.90×10^6	107	106	3.93	3.93	1.04	0.16	0.16
$bbbb$	4.02×10^4	1.02×10^6	5.84	5.76	0.10	0	0	0	0
$llbbbb$	69.5	1.06×10^5	15.0	7.42	6.69	5.44	4.68	4.18	0.97
$llqqh$	157	6.30×10^4	138	68.1	65.0	51.1	46.9	9.92	1.93
BG	1.74×10^6	1.46×10^7	8992	6748	3806	842	138	20.3	4.32

TABLE II: The reduction table for the signal and backgrounds after the final selection for the muon-type category of $llHH$ mode, together with the number of expected events and generated events. The cuts names are explained in text.

Process	expected	generated	pre-selection	electron-type	Cut1	Cut2	Cut3	Cut4	Cut5
$llHH$	46.5	3.88×10^5	26.5	13.3	13.0	10.6	10.4	5.76	4.47
$eebb$	2.84×10^5	4.18×10^6	3950	0	0	0	0	0	0
$\mu\mu bb$	4.96×10^4	1.00×10^6	1944	1943	1750	73.3	72.8	7.28	2.33
$e\nu bbqq$	2.48×10^5	1.51×10^6	2437	0	0	0	0	0	0
$\mu\nu bbqq$	2.46×10^5	1.48×10^6	239	215	95.7	65.7	33.3	2.78	0
$\tau\nu bbqq$	2.46×10^5	1.35×10^6	156	7.76	2.62	1.82	0.8	0	0
$bbqqqq$	6.24×10^5	3.90×10^6	107	1.09	0	0	0	0	0
$bbbb$	4.02×10^4	1.02×10^6	5.84	0.08	0	0	0	0	0
$llbbbb$	69.5	1.06×10^5	15.0	7.57	7.10	5.92	5.90	5.38	1.29
$llqqh$	157	6.30×10^4	138	69.7	68.4	54.3	54.0	12.8	2.36
BG	1.74×10^6	1.46×10^7	8992	2244	1924	201	167	28.2	5.97

B. Analysis of $e^+ + e^- \to \nu\bar{\nu}HH \to \nu\bar{\nu}b\bar{b}b\bar{b}$ at 500 GeV

In this search mode, the final state of a candidate signal event contains two missing neutrinos and four b quarks fragmenting into four jets. The three types of neutrinos ν_e, ν_μ and ν_τ are considered together. The analysis strategy is quite similar with the lepton pair mode. For pre-selection, we reject the isolated lepton and cluster the events to four jets and pair them by minimizing the χ^2 of two Higgs masses, loose mass cut and b-tagging cut are added. In the final selection, first we use missing energy and missing pt to suppress the full hadronic background. Then we train three neural-nets to suppress the dominant $bbbb$, $l\nu bbqq$ and $\nu\nu bbbb$ backgrounds. And eventually tight b-tagging is added. The reduction table of this searching mode is shown in Table III and the all the final cuts are summarized as following:

- Pre-selection: $|M(bb) - M(H)| < 80$ and $Bmax3 > 0.2$
- Final-selection:
  1. Cut1: $Evis - 0.83MissPt < 360$ GeV, $MissMass > 60$ GeV.
  2. Cut2: $NpfosMin >= 8$, 92 GeV $< M(H_1) <$ 136 GeV, 94 GeV $< M(H_2) <$ 130 GeV, $M(HH) >$ 200 GeV.
  3. Cut3: $\text{MLP}_{bbbb} > 0.83$.
  4. Cut4: $\text{MLP}_{l\nu bbqq} > 0.56$.
  5. Cut5: $\text{MLP}_{\nu\nu bbbb} > 0.61$.
  6. Cut6: $Bmax3 + Bmax4 > 1.14$.

TABLE III: The reduction table for the signal and backgrounds after the final selection for $\nu\nu HH$ mode, together with the number of expected events and generated events. The cuts names are explained in text.

Process	expected	generated	pre-selection	Cut1	Cut2	Cut3	Cut4	Cut5	Cut6
$\nu\nu HH$	103	7.06×10^5	45.0	43.6	26.0	22.7	20.6	17.1	8.47
$\nu\nu bb$	2.73×10^5	4.79×10^5	861	758	9.17	4.25	4.25	3.02	0
$e\nu bbqq$	2.48×10^5	1.51×10^6	3884	2126	504	451	72.6	54.9	0
$\mu\nu bbqq$	2.46×10^5	1.48×10^6	1637	951	22.3	195	72.8	52.1	0
$\tau\nu bbqq$	2.46×10^5	1.59×10^6	37440	24728	2591	3890	959	724	2.07
$bbqqqq$	6.24×10^5	3.88×10^6	58457	1212	178	71.5	38.6	37.2	0
$bbbb$	4.02×10^4	7.06×10^5	30826	3684	350	13.2	9.82	7.87	2.99
$\nu\nu bbbb$	97.1	8.22×10^4	82.1	80.5	10.1	6.90	5.66	2.03	0.87
$\nu\nu qqh$	469	7.41×10^4	82.1	79.0	21.5	17.5	13.0	5.86	1.93
BG			1.33×10^5	33619	5887	4650	1176	887	7.86

1. *Summary of the $\nu\nu HH$ mode*

In the $e^+ + e^- \to \nu\bar{\nu}HH$ search mode, assuming the Higgs mass of 120 GeV, the integrated luminosity of 2 ab^{-1}, and the beam polarization $P(e^-, e^+) = (-0.8, +0.3)$, it is expected to observe 8.5 signal and 7.9 backgrounds events, corresponding to a ZHH excess significance of 2.5σ and a ZHH cross section measurement significance of 2.1σ.

C. Analysis of $e^+ + e^- \to q\bar{q}HH \to q\bar{q}b\bar{b}b\bar{b}$ at 500 GeV

In this search mode, the final state of a candidate signal event contains four of six b quarks each fragmenting into a b jet. In the pre-selction, we require no isolated lepton and cluster the particles to six jets, which are then paired to form two Higgs bosons and one Z boson. The third larges b-likeness of the four jets from two Higgs is required to be larger than 0.16 in the pre-selection. In the final selection, all the events are separated into two categories according to the flavor tagging of the two jets from Z decay, which are correspondingly bbHH dominant and light qqHH dominant. The sum of b-likeness of the two jets from Z decay is used to achieve the separation. The dominant background in this analysis are $bbbb$ from $ZZ(bbZ)$, full hadronic $bbqqqq$ from $t\bar{t}$, $qqbbbb$ from ZZZ and ZZH, each of which is suppressed with a neural-net. The reduction table of this searching mode is shown in Table IV and V, and the all the final cuts are summarized as following:

- $bbHH$ dominant category
  1. Cut1: sum of b-likeness of the two jets from Z > 0.54.
  2. Cut2: $MissPt < 60$ GeV, $Npfos < 245$, 30 GeV $< M(Z) <$ 139 GeV, 73 GeV $< M(H_1) <$ 170 GeV, 73 GeV $< M(H_2) <$ 148 GeV.
  3. Cut3: MLP$_{bbbb} > 0.47$.
  4. Cut4: MLP$_{bbqqqq} > 0.33$.
  5. Cut5: MLP$_{qqbbbb} > 0.16$.
  6. Cut6: $Bmax3 + Bmax4 > 1.17$.
- light $qqHH$ dominant
  1. Cut1: sum of b-likeness of the two jets from Z < 0.54.
  2. Cut2: $MissPt < 60$ GeV, $Npfos < 245$, 60 GeV $< M(Z) <$ 131 GeV, 97 GeV $< M(H_1) <$ 133 GeV, 84 GeV $< M(H_2) <$ 136 GeV.
  3. Cut3: MLP$_{bbbb} > 0.48$.
  4. Cut4: MLP$_{bbqqqq} > 0.51$.
  5. Cut5: MLP$_{qqbbbb} > 0.09$.
  6. Cut6: $Bmax3 > 0.85$, $Bmax3 + Bmax4 > 1.21$.

TABLE IV: The reduction table for the signal and backgrounds after the final selection for $bbHH$ dominant category, together with the number of expected events and generated events. The cuts names are explained in text.

Process	expected	generated	pre-selection	Cut1	Cut2	Cut3	Cut4	Cut5	Cut6
$qqHH$	310	3.73×10^5	111	26.9	25.1	23.0	22.4	21.1	13.6
$l\nu bbqq$	7.40×10^5	3.56×10^6	17240	363	103	18.7	15.9	12.8	0.03
$bbuddu$	1.56×10^5	8.87×10^5	565	11.4	11.3	10.0	7.65	6.92	0.55
$bbcsdu$	3.12×10^5	1.26×10^6	6109	89.0	78.4	67.6	51.2	45.1	1.01
$bbcssc$	1.56×10^5	1.17×10^6	12456	263	246	212	147	129	3.69
$bbbb$	4.02×10^4	7.19×10^5	22889	2319	733	16.5	15.0	11.8	5.25
$qqbbbb$	140	1.23×10^5	82.9	13.9	12.7	9.80	9.19	5.78	3.03
$qqqqh$	818	5.98×10^4	154	27.5	25.4	22.5	21.6	18.5	10.9
ttz	2.20×10^3	8.49×10^4	172	17.2	13.6	12.5	12.3	11.4	2.88
$ttbb$	2.11×10^3	8.25×10^4	450	47.8	29.9	26.0	24.5	22.6	3.40
BG			60119	3152	1253	395	304	264	30.7

TABLE V: The reduction table for the signal and backgrounds after the final selection for light $qqHH$ dominant category, together with the number of expected events and generated events. The cuts names are explained in text.

Process	expected	generated	pre-selection	Cut1	Cut2	Cut3	Cut4	Cut5	Cut6
$qqHH$	310	3.73×10^5	111	84.0	36.9	34.2	31.0	30.8	18.8
$l\nu bbqq$	7.40×10^5	3.56×10^6	17240	16877	408	147	74.0	73.2	1.07
$bbuddu$	1.56×10^5	8.87×10^5	565	554	102	96.7	48.4	47.9	5.93
$bbcsdu$	3.12×10^5	1.26×10^6	6109	6020	1200	1094	501	492	15.7
$bbcssc$	1.56×10^5	1.17×10^6	12456	12193	2308	2111	848	829	16.0
$bbbb$	4.02×10^4	7.19×10^5	22889	20570	273	22.0	18.1	17.2	10.0
$qqbbbb$	140	1.23×10^5	82.9	68.9	11.1	9.49	7.92	6.95	4.07
$qqqqh$	818	5.98×10^4	154	126	37.8	34.0	30.5	29.9	16.1
ttz	2.20×10^3	8.49×10^4	172	155	30.3	29.4	25.7	25.5	7.74
$ttbb$	2.11×10^3	8.25×10^4	450	402	62.4	59.3	49.0	48.6	14.0
BG			60119	56967	4433	3603	1603	1570	90.6

1. *Summary of the qqHH mode*

In this $e^+ + e^- \to q\bar{q}HH$ search mode, assuming the Higgs mass of 120 GeV and the integrated luminosity of 2 ab^{-1}, with the beam polarization $P(e^-, e^+) = (-0.8, +0.3)$, in $bbHH$ dominant category, it is expected to observe 13.6 signal events with 30.7 backgrounds events, corresponding to a ZHH excess significance of 2.2σ and a ZHH cross section measurement significance of 2.0σ; in light $qqHH$ dominant category, it is expected to observe 18.8 signal events with 90.6 backgrounds events, corresponding to a ZHH excess significance of 1.9σ and a ZHH cross section measurement significance of 1.8σ. .

V. COMBINED RESULT OF $e^+e^- \to ZHH$ AT 500 GEV

The results of the three searching modes of $e^+e^- \to ZHH$ are shown in Table VI for the beam polarization $P(e^-, e^+) = (-0.8, +0.3)$, which is favored benefiting with higher cross section. The ZHH excess significance (i) and the measurement significance (ii) are also shown there. Notice that there are two independent parts in the $q\bar{q}HH$ mode. In this section, we will combine these results and try to answer the following two crucial questions:

- Can we observe the ZHH events? How much is the combined ZHH excess significance?
- Can we observe the trilinear Higgs self-interaction? How precisely can we measure the trilinear Higgs self-coupling?

TABLE VI: The numbers of the remaining signal and background events in each search mode of the $e^+e^- \to ZHH$ analysis based on the full detector simulation at 500 GeV, with the beam polarization $P(e^-, e^+) = (-0.8, +0.3)$. The last two columns are ZHH excess significance (i) and cross section measurement significance (ii). The $qqHH$ mode and $llHH$ mode are both separated into two categories: (a) $bbHH$ dominant, (b) light $qqHH$ dominant, (c) electron-type $llHH$, (d) muon-type $llHH$.

Search Mode	Signal	Background	Significance (i)	Significance (ii)
$qqHH$ (a)	13.6	30.7	2.2σ	2.0σ
$qqHH$ (b)	18.8	90.6	1.9σ	1.8σ
$\nu\nu HH$	8.5	7.9	2.5σ	2.1σ
$llHH$ (c)	3.7	4.3	1.5σ	1.1σ
$llHH$ (d)	4.5	6.0	1.5σ	1.2σ

A. Statistical independence of the three modes

Before deriving the combined result, it is necessary to check the statistical independence of the three modes.

- Due to the very different visible energy requirement in the $\nu\nu HH$ mode and the other two modes, events selected for the $\nu\nu HH$ mode will not satisfy the selection criteria for the other two modes. Thus the $\nu\nu HH$ mode is statistically independent of the $llHH$ and $qqHH$ modes.
- Due to the very energetic isolated lepton requirement for the $llHH$ mode, all fully hadronic events will not be selected, so that the $llHH$ mode is statistically independent of the $qqHH$ mode.

Thus we conclude that all the three modes are statistically independent.

B. Combined ZHH excess significance

A hypothesis test is used to calculate the combined ZHH excess significance. Define the null hypothesis:

$$\mathrm{H}_0: \quad \text{there is only background (B).} \tag{14}$$

and the alternative hypothesis:

$$\mathrm{H}_1: \quad \text{there are ZHH signal and background (S+B).} \tag{15}$$

Then define the test variable

$$\chi^2 \equiv -2\ln\frac{L_{s+b}}{L_b} \tag{16}$$

where the likelihood L_{s+b} is defined as

$$L_{s+b} = \prod_i \frac{e^{-(s_i+b_i)}(s_i+b_i)^{n_i}}{n_i!} \tag{17}$$

and the L_b is defined as

$$L_b = \prod_i \frac{e^{-b_i} b_i^{n_i}}{n_i!}. \tag{18}$$

The s_i and b_i are the expected numbers of remaining signal and background events in mode i (search modes $i = 1, .., 4$). The n_i is the total number of observed events in mode i, which is a Poisson random variable, with mean value $s_i + b_i$ under hypothesis H_1, and with mean value b_i under hypothesis H_0.

Figure 24 shows the distributions of the χ^2 test variable under hypothesis H_0, denoted by blue line, and under hypothesis H_1, denoted by red line, produced using a Toy Monte-Carlo. The black line shows the observed value of the test χ^2. The significance of the observed value under no signal hypothesis is obtained to be $s\sigma = 5.0\sigma$, meaning that a statistical significance of 5.0σ is expected to observe the excess of ZHH events.

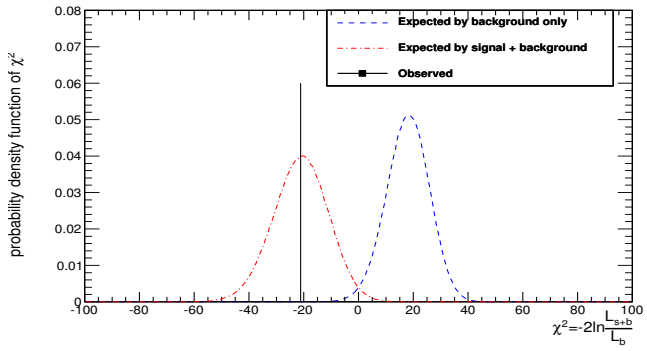

FIG. 24: The distribution of the test χ^2 under the background only hypothesis (blue) and the signal + background hypothesis (red). The black vertical line denotes the observed value of the test χ^2 variable.

C. Extracting the Cross Section of ZHH

The precision of the Higgs self-coupling is determined by the precision of the ZHH cross section, as introduced in Eqn 2. The cross section measurement can be carried out by parameter estimation through Maximum Likelihood. Define the combined likelihood

$$L_{s+b} = \prod_i \frac{e^{-(s_i+b_i)}(s_i+b_i)^{n_i}}{n_i!}, \tag{19}$$

where b_i is the expected number of background events, which is known from MC simulations; n_i is the number of observed events, which is known from the measurement; s_i is related to the cross section σ_{ZHH}, which is the unique parameter. The relation between s_i and σ_{ZHH} is

$$s_i = (\sigma_{ZHH} + \sigma_i) \cdot \text{Lumi} \cdot \text{Br}_i \cdot \text{Eff}_i \tag{20}$$

where Lumi is the integrated luminosity; Br_i is the branch ratio of mode i; Eff_i is the selection efficiency of mode i; σ_i is the fusion contribution for mode i, which is negligible at 500 GeV. The Likelihood hence contains only one parameter σ_{ZHH}. The minimization of $\chi^2 = -2\ln\frac{L}{L_{max}}$ is shown in Figure 25. The result is

$$\sigma_{ZHH} \cdot \text{Lumi} = 443^{+122}_{-115}. \tag{21}$$

For the integrated luminosity of 2 ab^{-1}, we then have

$$\sigma_{ZHH} = 0.22 \pm 0.06 \text{ fb}. \tag{22}$$

The precision of the cross section is 26.7%. Recalling the sensitivity of Higgs self-coupling to the cross section in Figure 4 (left), the Higgs self-coupling can be measured to the precision of 48% in case of without weighting. If we use the weighting method, the precision on Higgs self-coupling would be further improved to 44%.

VI. ANALYSIS OF $e^+e^- \to \nu\bar{\nu}HH$ AT 1 TEV

In this searching mode, the strategy is quite similar as that in $\nu\nu HH$ at 500 GeV. The dominant background here are from the semi-leptonic decay of $t\bar{t}$ and $\nu\nu ZH$ from WW fusion process, which are suppressed by two neural-nets. The pre-selection and final selection are summarized as following and the reduction table is shown in Table VII.

- Pre-selection:
  - no isolated lepton.
  - cluster all pros to 4 jets, each at least with 7 pfos and the third largest b-likeness to be larger than 0.2.

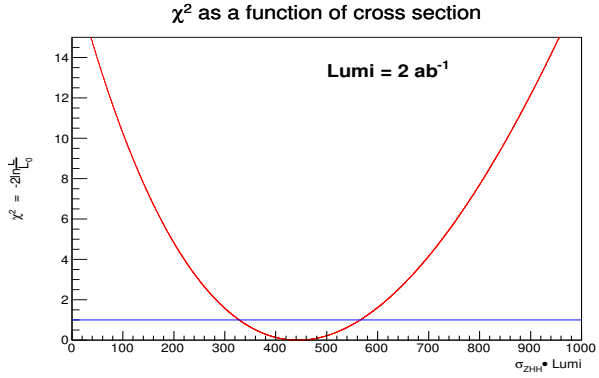

FIG. 25: The χ^2 as a function of $\sigma_{ZHH} \cdot \text{Lumi}$.

– visible energy less than 900 GeV, missing p_t large than 5 GeV, and missing mass larger than 0.

- Final-selection:

  1. Cut1: $Evis < 700 + 5MissPt$ GeV.
  2. Cut2: $\text{MLP}_{l\nu bbqq} > 0.84$.
  3. Cut3: $\text{MLP}_{\nu\nu bbbb} > 0.36$.
  4. Cut4: $Bmax3 + Bmax4 > 0.71$.

TABLE VII: The reduction table for the signal and backgrounds after the final selection for $\nu\nu HH$ at 1 TeV mode, together with the number of expected events and generated events. The cuts names are explained in text.

Process	expected	generated	pre-selection	Cut1	Cut2	Cut3	Cut4
$\nu\nu HH\ (fusion)$	272	1.05×10^5	127	107	77.2	47.6	35.7
$\nu\nu HH\ (ZHH)$	74.0	2.85×10^5	32.7	19.7	6.68	4.88	3.88
$yyxye\nu$	1.50×10^5	6.21×10^5	812	424	44.4	11.0	0.73
$yyxyl\nu$	2.57×10^5	1.17×10^6	13457	4975	202	84.5	4.86
$yyxyyx$	3.74×10^5	1.64×10^6	18951	4422	38.5	26.7	1.83
$\nu\nu bbbb$	650	2.87×10^5	553	505	146	6.21	4.62
$\nu\nu ccbb$	1070	1.76×10^5	269	242	63.3	2.69	0.19
$\nu\nu qqh$	3125	7.56×10^4	522	467	257	30.6	17.6
BG	7.86×10^5		34597	11054	758	167	33.7

1. *Summary of the $\nu\nu HH$ at 1 TeV*

In this $e^+ + e^- \to \nu\bar{\nu}HH$ searching mode, assuming the Higgs mass of 120 GeV and the integrated luminosity of 2 ab^{-1}, with the beam polarization $P(e^-, e^+) = (-0.8, +0.2)$, it is expected to observe 35.7 signal events with 33.7 backgrounds events, expecting the measurement significance of 4.3σ. The cross section of $\nu\nu HH$ from fusion can be measured to the precision of 23%, corresponding to the precision of 20% on the Higg self-coupling according to the sensitivity in Figure 4 (right). And with the weighting method, the precision on Higgs self-coupling would be further improved to 18%. Another important information from this analysis is that the double Higgs production excess with a statistical significance of 7.2σ is expected to be observed.

VII. SUMMARY

Acknowledgments

We would like to thanks all the members of the ILC physics subgroup for useful discussions. This study is supported in part by KEK, Center of High Energy Physics, Tsinghua University and the JSPS Core University Program.

[1] E. Asakawa, D. Harada, S. Kanemura, Y. Okada, and K. Tsumura, Phys.Rev. **D82**, 115002 (2010), 1009.4670.
[2] S. Kanemura, S. Kiyoura, Y. Okada, E. Senaha, and C. Yuan, Phys.Lett. **B558**, 157 (2003), hep-ph/0211308.
[3] S. Kanemura, Y. Okada, E. Senaha, and C.-P. Yuan, Phys.Rev. **D70**, 115002 (2004), hep-ph/0408364.
[4] A. Djouadi, W. Kilian, M. Muhlleitner, and P. M. Zerwas, Eur. Phys. J. **C10**, 27 (1999), hep-ph/9903229.
[5] C. Grojean, G. Servant, and J. D. Wells, Phys.Rev. **D71**, 036001 (2005), hep-ph/0407019.
[6] S. Kanemura, Y. Okada, and E. Senaha, Phys.Lett. **B606**, 361 (2005), hep-ph/0411354.
[7] S. Ham and S. Oh, (2005), hep-ph/0502116.
[8] M. Aoki, S. Kanemura, and O. Seto, Phys.Rev.Lett. **102**, 051805 (2009), 0807.0361.
[9] M. Aoki, S. Kanemura, and O. Seto, Phys.Rev. **D80**, 033007 (2009), 0904.3829.
[10] A. Noble and M. Perelstein, Phys.Rev. **D78**, 063518 (2008), 0711.3018.
[11] A. G. Cohen, D. Kaplan, and A. Nelson, Ann.Rev.Nucl.Part.Sci. **43**, 27 (1993), hep-ph/9302210.
[12] M. Quiros, Helv.Phys.Acta **67**, 451 (1994).
[13] V. Rubakov and M. Shaposhnikov, Usp.Fiz.Nauk **166**, 493 (1996), hep-ph/9603208.
[14] G. D. Kribs, T. Plehn, M. Spannowsky, and T. M. Tait, Phys.Rev. **D76**, 075016 (2007), 0706.3718.
[15] G. Gounaris, D. Schildknecht, and F. Renard, Phys.Lett. **B83**, 191 (1979).
[16] A. Djouadi, H. Haber, and P. Zerwas, Phys.Lett. **B375**, 203 (1996), hep-ph/9602234.
[17] V. Ilyin, A. Pukhov, Y. Kurihara, Y. Shimizu, and T. Kaneko, Phys.Rev. **D54**, 6717 (1996), hep-ph/9506326.
[18] V. D. Barger and T. Han, Mod.Phys.Lett. **A5**, 667 (1990).
[19] F. Boudjema and E. Chopin, Z.Phys. **C73**, 85 (1996), hep-ph/9507396.
[20] J. Tian, (2012), Presentation for LCWS12.
[21] P. Speckmayer, A. Hocker, J. Stelzer, and H. Voss, J.Phys.Conf.Ser. **219**, 032057 (2010).

Appendix A: Definition of Significance

(i) excess significance. Assuming there's only background, the p value is defined as the probability of observing events equal to or more than the number of the expected events, $N_S + N_B$:

$$p = \int_{N_S+N_B}^{\infty} f(x; N_B)\mathrm{d}x \tag{A1}$$

where $f(x; N_B)$ is the probability density function for the number of observed events when only the background exists, with the expected number N_B. Here, the number of observed events is a Poisson random variable $f(n; N_B) = \frac{e^{-N_B} N_B^n}{n!}$, the p value is calculated as

$$p = \sum_{n=N_S+N_B}^{\infty} f(n; N_B) \tag{A2}$$

corresponding to the significance:

$$1 - p = \int_{-\infty}^{s\sigma} N(x; 0, 1)\mathrm{d}x \tag{A3}$$

where $N(x; 0, 1)$ is the normal gaussian probability density function. The significance $s\sigma$ is defined as the excess significance. In the large statistics limit where if $f(x)$ becomes gaussian, this definition leads to the familiar significance formula $\frac{N_S}{\sqrt{N_B}}$.

(ii) measurement significance. Assuming both signal and background exist, the p value is defined as the probability of observing events equal to or less than the expected number of background events:

$$p = \int_{-\infty}^{N_B} f(x; N_B + N_S)\mathrm{d}x. \tag{A4}$$

This definition of significance is called "measurement significance". In the large statistics limit where if $f(x)$ becomes gaussian, this definition leads to the familiar significance formula $\frac{N_S}{\sqrt{N_S+N_B}}$.