Search for Light Scalars Produced in Association with Muon Pairs for $\sqrt{s} = 250$ GeV at the ILC

YAN WANG$^{1)a,b}$, JENNY LIST$^a$, MIKAEL BERGGREN$^a$

$^a$DESY, Notkestraße 85, 22607 Hamburg, GERMANY, $^b$IHEP, Yuquan Road 19B, 100049 Beijing, CHINA

Abstract

In many models with extended Higgs sectors, e.g. 2HDM, NMSSM and the Randall Sundrum model, there exists a light scalar $h$, lighter than the Standard Model (SM) like Higgs, and the coupling of $hZZ$ can be very small, as expected from the likeness of the 125 GeV Higgs boson measured at the LHC to the SM Higgs boson. Such a light scalar with suppressed couplings to the $Z$ boson would have escaped detection at LEP due to its limited luminosity. With a factor of 1000 higher luminosity and polarized beams, the International Linear Collider (ILC) is expected to have substantial discovery potential for such states. Furthermore, searches for additional scalars at LEP and LHC are usually dependent on the model details, such as decay channels. Thus, it is necessary to have a more general analysis with model-independent assumptions.

In this work, we perform a search for a light higgs boson produced in association with $Z$ boson at the ILC with a center-of-mass energy of 250 GeV, using the full Geant4-based simulation of the ILD detector concept. In order to be as model-independent as possible, the analysis is performed using the recoil technique, in particular with the $Z$ boson decaying into a pair of muons. As a preliminary result, the ILC’s exclusion limits will be shown for different higgs masses between 30 and 115 GeV.


This work was carried out in the framework of the ILD detector concept group.

$^{1)$speaker
Search for Light Scalars Produced in Association with Muon Pairs for $\sqrt{s} = 250$ GeV at the ILC

YAN WANG$^{1,2}$, JENNY LIST$^1$, and MIKAEL BERGGREN$^1$

$^1$DESY, Notkestraße 85, 22607 Hamburg, GERMANY
$^2$IHEP, Yuquan Road 19B, 100049 Beijing, CHINA

February 13, 2018

ABSTRACT

In many models with extended Higgs sectors, e.g. 2HDM, NMSSM and the Randall Sundrum model, there exists a light scalar $h$, lighter than the Standard Model (SM) like Higgs, and the coupling of $hZZ$ can be very small, as expected from the likeness of the 125 GeV Higgs boson measured at the LHC to the SM Higgs boson. Such a light scalar with suppressed couplings to the Z boson would have escaped detection at LEP due to its limited luminosity. With a factor of 1000 higher luminosity and polarized beams, the International Linear Collider (ILC) is expected to have substantial discovery potential for such states. Furthermore, searches for additional scalars at LEP and LHC are usually dependent on the model details, such as decay channels. Thus, it is necessary to have a more general analysis with model-independent assumptions.

In this work, we perform a search for a light higgs boson produced in association with Z boson at the ILC with a center-of-mass energy of 250 GeV, using the full Geant4-based simulation of the ILD detector concept. In order to be as model-independent as possible, the analysis is performed using the recoil technique, in particular with the Z boson decaying into a pair of muons. As a preliminary result, the ILC’s exclusion limits will be shown for different higgs masses between 30 and 115 GeV.$^\dagger$

$^\dagger$yan.wang@desy.de
1 Introduction

Many new physics models predict one or more extra scalars. For example, in Two Higgs Doublet Model (2HDM), the additional neutral particles are two scalars and a pseudoscalar [1]. In the Next-to-Minimal Supersymmetric Standard Model (NMSSM), it has three scalars and two pseudoscalars [3, 4, 5, 6, 7]. In the Randall Sundrum model, it predicts the presence of scalar particles, the so-called radion [Randall:1999ee]. In these models, a scalar lighter than 125 GeV is well motivated [Ellwanger:2011sk, Angelescu:2017ijy, Richard:2017kot, 2]. However, the 125 GeV Higgs boson measured at the LHC is rather Standard Model (SM) like [8, 9]. Thus, if such a new light higgs exists, its coupling to the Z boson will be suppressed [10]. The LEP searches for the SM Higgs boson can be used to constrain these additional scalars, but in most cases only when their properties, especially decay profiles, are similar to those of the SM Higgs boson [11, 12, 13]. Furthermore, LEP/LHC constraints on the extra scalars rely on the model details, e.g. CP properties, mass hierarchy, couplings, etc. [14]. Therefore, it is necessary to have a more general analysis with as few assumptions as possible. The OPAL collaboration has searched for light scalars in a model-independent way at LEP, but the results are limited by luminosity [15].

The International Linear Collider (ILC) is an electron-positron linear collider, with a center-of-mass energy of 250 GeV at its first stage. In Table 1, there is a brief comparison between LEP and ILC when searching for scalars [11, 15, 16]. Comparing with LEP, ILC has 1000 times higher luminosity, which makes the recoil mass technique more accurate, and it can provide more observables with polarized beams, such as angular correlations, which will help to distinguish unique signals of the scalar production [17, 18]. Thus, ILC will be sensitive to light scalars with a very weak interaction of Z boson using model-independent analysis.

This paper is structured as follows: Section 2 introduces the signal and background processes as well as the ILD detector concept and the detector simulation tools; Section 3 presents the methods of data selection; Section 4 shows the exclusion limits for $hZZ$ coupling; Section 5 concludes the paper.

2 Event Generation and Detector Simulation

2.1 Signal and background processes

We consider here the production of a light scalar $h$ in association with a Z boson where the Z boson decays to a muon pair. The $\mu\mu h$ signal Monte Carlo (MC) samples have been generated using the Whizard 1.95 event generator [19], at a center-of-mass energy of 250 GeV, for 100% left-handed and right-handed beam polarization configurations. Five benchmark points are chosen for the signal MC

*In this paper, "higgs" is used for a scalar lighter than 125 GeV, while "Higgs" stands for the SM/125 GeV Higgs boson.
Table 1: Comparison of LEP and ILC characteristics when searching scalars. The traditional method for discriminating signal and background at LEP is identifying the decay modes of scalars [11]. The recoil method at LEP refers to the analysis with recoil technique by the OPAL collaboration at LEP [15].

samples, $M_h = 30, 50, 70, 90, 115$ GeV. The decay branching ratios are assumed to be the same as a 125 GeV SM Higgs boson, but no use would be made of this fact.

The scalar mass can be measured by the recoil technique due to the high luminosity at the ILC. Figure 1 shows the recoil mass distributions ($M_{\text{rec}}$) for different scalar masses, where the peak of recoil mass distributions are very sharp for each mass, which can be used to distinguish the signal and background.

As backgrounds, we use samples which have been generated in the context of the Detailed Baseline Design document (ILC TDR, Vol. 4 [16, 20]). They are generated at tree level and grouped by lepton numbers in the final state as follows:

- 2-fermion leptonic ($2f_l$), main channel: $e^+e^- \rightarrow Z/\gamma^* \rightarrow l^+l^-/\nu\nu$.
- 4-fermion leptonic ($4f_l$), main channel: $e^+e^- \rightarrow ZZ/WW \rightarrow 4l$.
- 4-fermion semi-leptonic ($4f_{sl}$), main channel: $e^+e^- \rightarrow ZZ/WW \rightarrow 2l2\nu/2\nu2q$.
- 4(2)-fermion hadronic ($4f_h / 2f_h$), main channel: $e^+e^- \rightarrow ZZ/WW \rightarrow 4q$ or $e^+e^- \rightarrow Z \rightarrow 2q$. These events are almost fully rejected when identifying isolated muon pairs.
Bremsstrahlung and initial state radiation (ISR) are explicitly taken into account for all events. Pythia 6.4 is used for the parton shower and hadronization [21]. The samples are reweighted to the polarization of $P(e^-, e^+) = (-80\%, +30\%)$, and an integrated luminosity of 2000 fb$^{-1}$.

2.2 Detector simulation

The generated events have been simulated with the full Geant4-based [22] simulation of the ILD detector concept [16], more specifically the ILD_o1_v05 detector modle in mokka [23]. The simulated events have been reconstructed with the standard tools in ILCSoft v01-16 [24]. The ILD detector is designed for optimal particle-flow performance [25]. It has a vertex detector consisting of three double-sided layers of silicon pixel sensors, a hybrid tracking system, which is realized with a time projection chamber and a combination of silicon tracking, and a calorimeter system. These systems are surrounded by a solenoid producing a 3.5 T magnetic field, and an iron flux return yoke. Event reconstruction has been performed using the PandoraPFA algorithm [26] to reconstruct individual particles within the Marlin framework [27]. The beam crossing angle of 14 mrad has been also taken into account [16].

3 Analysis

In this section, we describe the analysis strategy following the cross section measurement of the SM Higgs boson at the ILC [18]. The signal is selected by firstly identifying a pair of isolated and oppositely charged muons. Then final state radiation (FSR) and bremsstrahlung photons are recovered. Finally, background events are rejected with several kinematic cuts.
3.1 Selection of the best muon pair

Isolated muons are identified with the following criteria.

- Muons are required to have sufficient track momentum: \( p_{\text{track}} > 5 \) GeV, where \( p_{\text{track}} \) is the measured track momentum.

- Muon ionization is minimal when passing the ECAL and HCAL: \( E_{\text{CAL,tot}}/p_{\text{track}} < 0.3 \), \( E_{\text{yoke}} > 1.2 \) GeV, where \( E_{\text{CAL,tot}} \) and \( E_{\text{yoke}} \) is the energy deposit in ECAL plus HCAL and inside the muon detector.

- Suppression of the muons from \( \tau \) decay or \( b/c \) quark jets by requiring the uncertainties of \( d_0 \) and \( z_0 \), \( |d_0/\delta d_0| < 5 \), \( |z_0/\delta z_0| < 5 \), where \( d_0 \) and \( z_0 \) are the impact parameters in the transverse and longitudinal direction and \( \delta d_0 \) and \( \delta z_0 \) are their uncertainties.

Then a multi variate method is used for further identifying isolated muons [27, 28].

For the signal, at least one pair of isolated oppositely charged muons is selected. However, there may be more muons in the events, for example, the muons produced in the higgs decay process \( h \rightarrow WW^* \) and \( h \rightarrow ZZ^* \). In order to find the correct muon pair, the invariant mass \( M_{\mu^+\mu^-} \) of the correct muon pair should be close to the Z boson mass \( M_Z=91.2 \) GeV. Meanwhile, with the recoil technique, the recoil mass of the correct muon pair system can be calculated from Eq. (1). It should be close to higgs boson mass \( M_h = 30, 50, 70, 90, \) or 115 GeV, respectively, for each higgs mass benchmark point in this study:

\[
M_{\text{rec}}^2 = (\sqrt{s} - E_{\mu^+\mu^-})^2 - |\vec{p}_{\mu^+\mu^-}|^2.
\]  

(1)

Thus the best muon pair candidate is selected with the following criteria: First, the basic criteria \( |M_{\mu^+\mu^-} - M_Z| < 40 \) GeV is applied. Then the muon pair should minimize the following \( \chi^2 \) function:

\[
\chi^2(M_{\mu^+\mu^-}, M_{\text{rec}}) = \frac{(M_{\mu^+\mu^-} - M_Z)^2}{\sigma_{M_{\mu^+\mu^-}}^2} + \frac{(M_{\text{rec}} - M_h)^2}{\sigma_{M_{\text{rec}}}^2}.
\]  

(2)

where \( \sigma_{M_{\mu^+\mu^-}} \) and \( \sigma_{M_{\text{rec}}} \) are determined by a Gaussian fit to the generator-level distributions of \( M_{\mu^+\mu^-} \) and \( M_{\text{rec}} \).

After identifying the best muon pair, the bremsstrahlung and FSR photons from the muons are identified and added to the muons. A photon is identified as a FSR or bremsstrahlung photon when its cosine of the polar angle with respect to the isolated muons exceeds 0.99. Then the four momentum of the photon is combined with that muon.

3.2 Background rejection

Background events are rejected by considering kinematic variables as described below, while the specific cut values are adjusted for each higgs mass.
• Since $M_{\mu^+\mu^-}$ should be close to the Z boson mass, a criterion is imposed as $M_{\mu^+\mu^-} \in [73, 120]$ GeV for each higgs mass. Figure 2 compares the $M_{\mu^+\mu^-}$ distribution of signal and background processes when $M_h = 50$ GeV.

![Invariant mass distribution of muon pair system for signal $M_h = 50$ GeV and backgrounds.](image)

**Figure 2:** Invariant mass distribution of muon pair system for signal $M_h = 50$ GeV and backgrounds.

• The transverse momentum of the muon pair system $P_{\mu^+\mu^-}^T$ in the 2$f_l$ channel tends to have very small values, in contrast to the signal, which should have a peak at larger values which is determined by kinematics. This motivates $P_{\mu^+\mu^-}^T > 10$ GeV. In addition, an upper limit is also needed to reduce other backgrounds with large transverse momentum, mainly the events from 2$f_l$ processes. In order to maximize the sensitivity, the upper limit of $P_{\mu^+\mu^-}^T$ for each higgs mass benchmark point is chosen as the maximum $P_{\mu^+\mu^-}^T$ value in the region $|M_h - M_h^{\text{benchmark}}| < 10$ GeV. The maximum $P_{\mu^+\mu^-}^T$ is calculated as a function of the higgs mass, as shown in Figure 3 (a), and the $P_{\mu^+\mu^-}^T$ upper limit cuts are shown in Table 2. When $M_h < 50$ GeV, the maximum $P_{\mu^+\mu^-}^T$ cut is not necessary. As an example, the Figure 3 (b) compares the $P_{\mu^+\mu^-}^T$ distribution of the signal and major background processes when $M_h = 50$ GeV.

<table>
<thead>
<tr>
<th>$M_h$ (GeV)</th>
<th>30</th>
<th>50</th>
<th>70</th>
<th>90</th>
<th>115</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max $P_{\mu^+\mu^-}^T$ cut (GeV)</td>
<td>120</td>
<td>115</td>
<td>110</td>
<td>100</td>
<td>90</td>
</tr>
</tbody>
</table>

Table 2: Transverse momentum uplimits of muon pair system for different higgs masses.

• The 2$f_l$ processes contain a large number of ISR photons. If photons escape through the beam pipe, they contribute to the missing momentum in longitudinal direction. Thus, the cut for $\cos \theta_{\text{mis}} < 0.98$ can reject many 2$f_l$
Figure 3: (a) Maximum value for transverse momentum of muon pair system as a function of higgs mass; (b) transverse momentum distribution of muon pair system for $M_h = 50$ GeV.

background events. However, when the higgs mass is small, more ISR photons will escape from the detector. Thus, the $|\cos \theta_{miss}| < 0.98$ cut is applied only for $M_h > 50$ GeV.

- The signatures of the $ZZ$ background with one $Z$ boson decay to a muon pair are harder to distinguish from the signal. Therefore, a multi-variate analysis (MVA) based on the Gradient Boosted Decision Tree (BDTG) method [29], which is included in TMVA package [30] in ROOT [31], is used for further background rejection. The input variables for the BDTG are $\cos \theta_{\mu^+\mu^-}$, $\cos \theta_{\mu^-\mu^+}$, $\cos \theta_{\text{track}^+}$, $\cos \theta_{\text{track}^-}$ and $M_{\mu^+\mu^-}$. Here, $\theta_{\mu^+\mu^-}$ is the polar angle of the $Z$ boson, $\theta_{\mu^-\mu^+}$ is the angle between the muons, and $\theta_{\text{track}^+/\text{track}^-}$ is the polar angle of $\mu^+/\mu^-$ track. Figure 4 shows distribution of these variables for $M_h = 50$ GeV.

The BDTG is trained by using half of simulated signal and background events. Then the BDTG response is calculated with the other half of the events. In most $M_h$ cases, there are no obvious changes in final sensitivity for BDTG cuts between (-0.3, 0.3). The BDTG cut is chosen to be larger than 0 for all higgs masses. Figure 5 shows as an example the BDTG response for $M_h = 50$ GeV.

- The unique characteristic of the signal is the peak of the recoil mass close to the higgs mass. This motivates the cut $[(M_h - 20), 160]$ GeV. In Figure 6, taking $M_h = 50$ GeV for example, it shows the recoil mass distribution for signal and backgrounds.

According to the recoil mass distribution, the main backgrounds are distinguishing for different higgs boson masses. The signal-to-background ratio is clearer in $125 \text{GeV} > M_h > M_Z$ and $M_Z > M_h > 50$ GeV regions, because no background processes have the same feature as the signal. While in the $M_h \sim M_Z$ region, the $ZZ$ process with one $Z$ boson decaying to a muon pair is an irreducible background; and in $M_h < 50$ GeV region, the $2f_i$ background $e^+e^- \rightarrow Z \rightarrow \mu^+\mu^-$ with energetic ISR photon is overwhelming the signal.
Figure 4: Example distributions of the variables $\cos\theta_{\mu^+\mu^-}$, $\cos\theta_{\mu^-\mu^+}$, $\cos\theta_{\text{track}^+}$ and $\cos\theta_{\text{track}^-}$ for $M_h = 50$ GeV, which is used for the BDTG training input variables.

# 4 Results

In Table 3, there are the number of remaining signal and background as well as the signal efficiency and significance after all cuts for different higgs masses. The significance is defined as $S = \frac{S}{\sqrt{S+B}}$, and $S = \kappa_{hZZ}^2 \times \sigma_{h\mu\mu}^{M_h} \times \int L dt$, where $\kappa_{hZZ} = 1$. The cross section $\sigma_{h\mu\mu}^{M_h}$ increases when the scalar mass becomes small due to kinematics, which leads to a higher significance.

<table>
<thead>
<tr>
<th>$M_h$ (GeV)</th>
<th>light higgs</th>
<th>$4f_l$</th>
<th>$4f_{sl}$</th>
<th>$2f_l$</th>
<th>total bk</th>
<th>cut efficiency</th>
<th>significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>115</td>
<td>17420</td>
<td>61034</td>
<td>53869</td>
<td>13878</td>
<td>128781</td>
<td>0.67</td>
<td>45.56</td>
</tr>
<tr>
<td>90</td>
<td>22198</td>
<td>63211</td>
<td>74563</td>
<td>18514</td>
<td>156288</td>
<td>0.59</td>
<td>52.54</td>
</tr>
<tr>
<td>70</td>
<td>26841</td>
<td>51672</td>
<td>60358</td>
<td>37167</td>
<td>149196</td>
<td>0.57</td>
<td>63.97</td>
</tr>
<tr>
<td>50</td>
<td>30494</td>
<td>46128</td>
<td>54373</td>
<td>80074</td>
<td>180575</td>
<td>0.54</td>
<td>66.37</td>
</tr>
<tr>
<td>30</td>
<td>33844</td>
<td>51207</td>
<td>55743</td>
<td>213184</td>
<td>320134</td>
<td>0.49</td>
<td>56.88</td>
</tr>
</tbody>
</table>

Table 3: The number of events left after all kinematic cuts for $P(e^-, e^+) = (-80\%, +30\%)$, $\int L dt = 2000$ fb$^{-1}$ and $\sqrt{s} = 250$ GeV. Also given are the efficiency and signal significance (defined as significance = $S = \frac{S}{\sqrt{S+B}}$, and $S = \kappa_{hZZ}^2 \times \sigma_{h\mu\mu}^{M_h} \times \int L dt$, where $\kappa_{hZZ} = 1$).

A likelihood analysis is applied for calculating 2σ expected exclusion limits on the coupling $\kappa_{hZZ}$ with a bin-by-bin comparison between the signal and back-
Figure 5: The distribution of the BDTG response for signal and background for $M_h = 50$ GeV.

Figure 6: The recoil mass distribution of the signal and background for $M_h = 50$ GeV.

ground recoil mass histograms. Two hypotheses are provided, the background-only hypothesis which assumes no new higgs in the investigated mass range, while the signal-plus-background hypothesis which assumes the new higgs is produced in the mass range. Then a global test-statistic $Q(M_h) = \mathcal{L}_{s+b}(s(M_h))/\mathcal{L}_b(0)$ is constructed to discriminate signal and background, where the $\mathcal{L}_{s+b}$ ($\mathcal{L}_b$) is the likelihood function for the signal plus background (only background) hypothesis. The distribution of $Q(M_h)$ is normalized so that it becomes a probability density function, which is integrated to provide the confidence levels $CL_b(M_h)$ and $CL_{s+b}(M_h)$. The ratio $CL_s(M_h) = CL_{s+b}(M_h)/CL_b(M_h)$ is used as the final confidence level.

Finally, the 95% confidence level upper bounds on $\kappa_{hZZ}$ ($\kappa_{hZZ}^{95}$) is calculated for five higgs mass benchmark points, which is shown in Figure 7 with the black points. Comparing the $\kappa_{hZZ}^{95}$ in $M_h \sim M_Z$ and $40$ GeV $> M_h$ regions, the coupling magnitudes reach the valley bottom in $125$ GeV $> M_h > M_Z$ and $M_Z > M_h >$
40 GeV regions.

Figure 7: The 2\(\sigma\) exclusion limits for the coupling \(\kappa_{hZZ}\) for different higgs masses. The black points are the results at ILC with polarization \(P(e^-, e^+) = (-80\%, +30\%)\), \(\int L dt = 2000 \text{ fb}^{-1}\) and \(\sqrt{s} = 250\text{ GeV}\). The green line is the results with recoil technique at LEP, while the red line use traditional methods.

Figure 7 also shows the results at LEP for comparison. The red line is \(\kappa_{hZZ}^{95}\) measured at LEP by combining the data of the four LEP collaborations, ALEPH, DELPHI, L3 and OPAL [11]. In each of the four LEP experiments, the data analysis is done with traditional methods, i.e. identifying the decay modes of the higgs bosons in the discrimination between signal and background. The green line presents \(\kappa_{hZZ}^{95}\) obtained with the recoil technique by OPAL Collaboration [15] at LEP. In this search, the scalar masses have been measured down to the lowest generated signal mass of 1 keV, which is the only model-independent higgs search at LEP due to limited luminosity. The \(\kappa_{hZZ}^{95}\) is independent of the decay modes of the higgs boson, which is \(\kappa_{hZZ}^{95} < 1\) when \(M_h < 81\text{ GeV}\). When extrapolating these results to the ILC with \(\sqrt{s} = 250\text{ GeV}\), \(P(e^-, e^+) = (-80\%, +30\%)\) and \(\int L dt = 2000 \text{ fb}^{-1}\), the \(\kappa_{hZZ}^{95}\) is estimated to be [0.055 – 0.071] [32]. These extrapolated exclusion limits are about a half of the values given in this analysis, especially in the low mass region. The main reason is that the OPAL analysis has been split into visible and invisible decay modes of the higgs boson, and assume that the signal events rejected in the visible decay mode analysis (by photon rejection) can be recovered in the invisible decay study (including mono-photon production in signal decay channel)[15]. This may slightly sacrifice model independence. The current ILC results are completely model-independent. If the slight model dependence can be accommodated, the ILC can also reach the extrapolated exclusion limits. Another reason is the center-of-mass energy of ILC is larger than LEP. Thus, when higgs mass is small, the recoil mass distribution is duller than that in LEP, which makes it harder to reject the backgrounds by recoil mass cut.
5 Conclusions

Many BSM models favor light scalars. By applying the recoil technique, the potential of the ILC to search for scalars has been investigated at $\sqrt{s} = 250$ GeV, with the full simulation of the ILD detector concept. We have optimized the methods of signal selection and background rejection to be independent of the scalar decay modes. Preliminary $2\sigma$ expected exclusion limits for scale factor $r_{hZZ}^{05}$ of the $hZZ$ coupling are shown for five scalar mass benchmark points. The analysis shows two irreducible backgrounds affect the final results, two fermion backgrounds will be dominant in the low higgs mass region, while the $ZZ$ process with one $Z$ boson decaying to a muon pair is irreducible when $M_h$ is close to $Z$ boson mass. In future, we plan to split the decay mode of higgs boson into visible and invisible to improve the background rejection and extend our analysis to more mass points and check the model independency for each higgs decay modes.

ACKNOWLEDGEMENTS

We would like to thank the LCC generator working group and the ILD software working group for providing the simulation and reconstruction tools and producing the Monte Carlo samples used in this study. This work has benefited from computing services provided by the ILC Virtual Organization, supported by the national resource providers of the EGI Federation and the Open Science GRID. We are grateful for the support from Collaborative Research Center SFB676 of the Deutsche Forschungsgemeinschaft (DFG), Particles, Strings and the Early Universe, project B1. Y.W. is supported by the China Postdoctoral Science Foundation under Grant No. 2016M601134, and an International Postdoctoral Exchange Fellowship Program between the Office of the National Administrative Committee of Postdoctoral Researchers of China (ONACPR) and DESY.

References


