A study of the measurement precision of the Higgs boson decaying into tau pairs at the ILC

Shin-ichi Kawada¹, Keisuke Fujii², Taikan Suehara³, Tohru Takahashi¹, Tomohiko Tanabe⁴

¹ Advanced Sciences of Matter (AdSM), Hiroshima University, 1-3-1, Kagamiyama, Higashi-Hiroshima, Hiroshima 739-8530, Japan
² High Energy Accelerator Research Organization (KEK), 1-1, Oho, Tsukuba, Ibaraki 305-0801, Japan
³ Graduate School of Science, Kyushu University, 6-10-1, Hakozaki, Higashi-ku, Fukuoka 812-8581, Japan
⁴ International Center for Elementary Particle Physics (ICEPP), The University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

Abstract We evaluate the measurement precision of the production cross section times the branching ratio of the Higgs boson decaying into tau lepton pairs at the International Linear Collider (ILC). We analyze various final states associated with the main production mechanisms of the Higgs boson, the Higgs-strahlung and WW-fusion processes. The statistical precision of the production cross section times the branching ratio is estimated to be 2.6 and 6.9 % for the Higgs-strahlung and WW-fusion processes, respectively, with the nominal integrated luminosities assumed in the ILC Technical Design Report; the precision improves to 1.0 and 3.4 % with the running scenario including possible luminosity upgrades. The study provides a reference performance of the ILC for future phenomenological analyses.

1 Introduction

After the discovery of the Higgs boson by the ATLAS and the CMS experiments, who reported a combined signal yield consistent with the SM expectation, with a combined observed significance at the level of $5.5\sigma$ [4–6]. The purpose of this study is to estimate the projected ILC capabilities of measuring the $h \rightarrow \tau^+\tau^-$ decay mode in final states resulting from the main Higgs boson production mechanisms in $e^+e^-$ collisions. Existing studies on $h \rightarrow \tau^+\tau^-$ decays at $e^+e^-$ collisions [7,8] did not take into account some of the relevant background processes or were based on a Higgs boson mass hypothesis which differs from the observed value, both of which are addressed in this study. We assume the ILC capabilities for the accelerator and the detector as documented in the ILC Technical Design Report (TDR) [9–13] together with its running scenario published recently [14,15]. The results presented in this paper will be useful for future phenomenological studies.

The contents of this paper are organized as follows. In Sect. 2, we describe the ILC and the ILD detector concept, and the analysis setup. The event reconstruction and selection at center-of-mass energies of 250 and 500 GeV are discussed in Sects. 3 and 4. Section 5 describes the prospects for improving the measurement precision with various ILC running scenarios. We summarize our results in Sect. 6.

2 Analysis conditions

2.1 International linear collider

The ILC is a next-generation electron–positron linear collider. It covers a center-of-mass energy ($\sqrt{s}$) in the range of 250–500 GeV and can be extended to $\sqrt{s} = 1$ TeV. In the ILC design, both the electron and positron beams can be polarized, which allow precise measurements of the proper-
Table 1  Typical integrated luminosities $L$ and center-of-mass energies $\sqrt{s}$ of the ILC [9,14,15]

| Scenario              | $\sqrt{s}$ (GeV) | $L$ (fb$^{-1}$) |
|-----------------------|------------------|-----------------|
| Nominal               | 250              | 250             |
|                       | 500              | 500             |
| Luminosity upgrade    | 250              | 2000            |
|                       | 500              | 4000            |

Fig. 1  Diagrams of the main production mechanisms of the Higgs boson in $e^+e^-$ collisions. Top Higgs-strahlung (Zh) process. Middle $WW$-fusion process. Bottom $ZZ$-fusion process.

Fig. 2  Cross sections of the Higgs boson production as a function of $\sqrt{s}$ in the electron–positron interaction.

2.2 Production and decay of the Higgs boson

Figure 1 shows the diagrams for the main production mechanisms of the Higgs boson in $e^+e^-$ collisions. The cross sections of Higgs boson production calculated by WHIZARD [16] with a Higgs mass of 125 GeV are shown in Fig. 2, where polarizations of $-80$ and $+30$ % for the electron and positron beams are assumed, and initial state radiation is taken into account.

For the calculation of the production cross section and the subsequent decay of the signal processes of $e^+e^- \rightarrow f\bar{f}h \rightarrow f\bar{f}\tau^+\tau^-$, where $f$ denotes a fermion, we use an event generator based on GRACE [17,18]. The effect of beamstrahlung is implemented according to the calculation by GuineaPig [19], which simulates $e^+e^-$ beam–beam interactions, with the beam parameters described in the TDR [20]. Initial state radiation is incorporated following the prescription developed by the ILC Event Generator Working Group [13,21]. To handle the spin correlation of tau pairs from the Higgs boson decay, GRACE is interfaced with TAUOLA [22–25]. The decays of other short-lived particles and the hadronization of quarks and gluons are handled by PYTHIA [26].

2.3 Background processes

For background processes, we use common Monte-Carlo (MC) samples for SM processes previously prepared for the studies presented in the ILC TDR [13]. The event samples include $e^+e^- \rightarrow 2f$, $e^+e^- \rightarrow 4f$, $e^+e^- \rightarrow 6f$, and $e^+e^- \rightarrow f\bar{f}h$. The event generation of these processes is performed with WHIZARD [16], in which beamstrahlung, initial state radiation, decay of short-lived particles, and hadronization are taken into account in the same way as described in the previous section for the signal process. The background processes from $\gamma\gamma$ interactions with hadronic final states, in which photons are produced by beam–beam interactions, are generated on the basis of the cross section model in Ref. [27]. We find that the interactions between electron or positron beams and beamstrahlung photons, i.e., $e^\pm\gamma \rightarrow e^\pm\gamma$, $e^\pm\gamma \rightarrow 3f$, and $e^\pm\gamma \rightarrow 5f$, have negligible contributions to background.
2.4 Detector model

The detector model used in this analysis is the International Large Detector (ILD), which is one of the two detector concepts described in the ILC TDR. It is a general-purpose 4π detector designed for particle flow analysis, aiming at best possible jet energy resolution.

The ILD model consists of layers of sub-detectors surrounding the interaction point. One finds, from the innermost to the outer layers, a vertex detector (VTX), a silicon inner tracker (SIT), a time projection chamber (TPC), a silicon envelope tracker (SET), an electromagnetic calorimeter (ECAL), and a hadron calorimeter (HCAL), all of which are put inside a solenoidal magnet providing a magnetic field of 3.5 T. The return yoke of the solenoidal magnet has a built-in muon system. The ILD design has not yet been finalized. In this analysis, we assume the following configurations and performance. The VTX consists of three double layers of silicon pixel detectors with radii at 1.6, 3.7 and 6 cm. Each silicon pixel layer provides a point resolution of 2.8 μm. The TPC provides up to 224 points per track over a tracking volume with inner and outer radii of 0.33 and 1.8 m. The SIT and SET are used to improve the track momentum resolution by adding precise position measurements just inside and outside of TPC. The ECAL consists of layers of tungsten absorbers interleaved with silicon layers segmented into $5 \times 5$ mm$^2$ cells, has an inner radius of 1.8 m, and has a total thickness of 20 cm corresponding to 24 radiation length. The HCAL consists of layers of steel absorbers interleaved with scintillator layers segmented into $3 \times 3$ cm$^2$ cells and has an outer radius of 3.4 m corresponding to 6 interaction length. Additional silicon trackers and calorimeters are located in the forward region to assure hermetic coverage down to 5 mrad from the beam line. The key detector performance of the ILD model is summarized in Table 2. Details of the ILD model and the particle flow algorithm are found in Refs. [13,28].

2.5 Detector simulation and event reconstruction

In this study, we assume a Higgs boson mass of 125 GeV, a branching ratio of the Higgs boson decay into tau pairs $\left( BR(h \rightarrow \tau^+\tau^-) \right)$ of 6.32 % [29], and beam polarizations of $-80$ and $+30$ % for the electron and the positron beams, respectively.

We perform a detector simulation with Mokka [30], a Geant4-based [31] full detector simulator, with the ILD model for all signal and background processes, with the exception of the $e^+e^- \rightarrow e^+e^- + 2f$ process at $\sqrt{s} = 500$ GeV, for which SGV fast simulation [32] is used. The event reconstruction and physics analysis are performed within the MARLIN software framework [33], in which events are reconstructed using track finding and fitting algorithms, followed by a particle flow analysis using the PandoraPFA package [28].

### Table 2

| Name                        | Value                   |
|-----------------------------|-------------------------|
| Impact parameter resolution | $5 \oplus \frac{10}{p} \sin^{3/2} \theta$ μm |
| Momentum resolution         | $2 \times 10^{-5} \oplus \frac{1}{2} \times 10^{-3}$ GeV/c |
| Jet energy resolution       | $30 \sqrt{E(\text{GeV})}$ % |

3 Analysis at the center-of-mass energy of 250 GeV

At $\sqrt{s} = 250$ GeV, the Higgs-strahlung ($e^+e^- \rightarrow Zh$) process dominates the SM Higgs production, as shown in Fig. 2. The $WW$-fusion and $ZZ$-fusion cross sections are negligible at this energy. We take into account $e^+e^- \rightarrow f \bar{f}h$ (excluding the $h \rightarrow \tau^+\tau^-$ signal), $e^+e^- \rightarrow 2f$, and $e^+e^- \rightarrow 4f$ for the background estimation. The $\gamma\gamma$ hadrons background is overlaid onto the MC samples with an average of 0.4 events per bunch crossing [27]. An integrated luminosity of 250 fb$^{-1}$ is assumed for the results in this section.

There are four main signal modes: $e^+e^- \rightarrow q\bar{q}h$, $e^+e^- \rightarrow e^+e^- h$, $e^+e^- \rightarrow \mu^+\mu^- h$, and $e^+e^- \rightarrow \nu\bar{\nu}h$. For our $\sqrt{s} = 250$ GeV results, we report on the first three of these modes. We do not quote the results for the $\nu\bar{\nu}h$ mode as we find that it suffers from background processes with neutrinos in the final state. We do not analyze the $e^+e^- \rightarrow \tau^+\tau^- h$ mode in this study.

3.1 $e^+e^- \rightarrow q\bar{q}h$

3.1.1 Reconstruction of isolated tau leptons and the $Z \rightarrow q\bar{q}$ decay

For the $q\bar{q}h$ mode, we first identify the tau leptons using a dedicated algorithm developed for this topology. The algorithm proceeds as follows.

1. The charged particle with the highest energy is chosen as a working tau candidate.
2. The tau candidate is combined with the most energetic particle (charged or neutral) satisfying the following two conditions: the angle $\theta_l$ between the particle and the tau...
candidate satisfies $\cos \theta_t > 0.99$; and the combined mass, calculated from the sum of the four momenta of the particle and the tau candidate, does not exceed 2 GeV. The four momentum of this particle is then added to that of the tau candidate.

3. Step 2 is repeated until there are no more particles left to combine. The resulting tau candidate is then set aside.

4. The algorithm is repeated from Step 1 until there are no more charged particles left.

A tau candidate is accepted if the number of charged particles with track energy greater than 2 GeV is equal to one or three, the net charge is equal to $\pm 1$, and the total energy is greater than 3 GeV. Furthermore, an isolation requirement is applied as follows. A cone of half-angle $\theta_c$, with $\cos \theta_c = 0.95$, is defined around the direction of the tau momentum. The tau candidate is accepted if the energy sum of all particles inside the cone (excluding those forming the tau candidate) does not exceed 10% of the tau candidate energy. We require exactly two final tau candidates with opposite charges. This results in a selection efficiency of 49.3% for the $q\bar{q}\tau^+\tau^-$ signal events.

After the tau candidates are identified, the neutrino energy is recovered by using the collinear approximation [34]. Because tau leptons from a Higgs boson decay are highly boosted, it is reasonable to assume that the tau momentum and the neutrino momentum are nearly parallel. Under this assumption, the energy of the two neutrinos, one from each tau decay, can be solved by requiring that the overall transverse momentum of the event is balanced in two orthogonal directions. The neutrino reconstructed in this way is added to the tau candidate. Figure 3 shows the invariant mass distributions of the tau pairs without ($M_{\tau+\tau-}^{\text{col}}$) and with ($M_{\tau+\tau-}^{\text{col}}$) the collinear approximation for the events containing two tau lepton candidates with opposite charges. With the collinear approximation, a clear peak is visible at 125 GeV for signal events. The $M_{\text{col}}$ distribution for background events with the same criteria is also shown.

The Durham jet clustering algorithm [35] is applied to the remaining particles to reconstruct the two jets from the $Z$ boson decay.

3.1.2 Event selection

We perform a pre-selection over the reconstructed events, followed by a multivariate analysis. The pre-selection is designed to reduce background while keeping most of the signal. The events are pre-selected according to the following criteria. The $Z \rightarrow q\bar{q}$ candidate and the $h \rightarrow \tau^+\tau^-$ candidate are successfully reconstructed. The total number of charged particles is at least 9. The visible energy of the event, $E_{\text{vis}}$, lies in the range of 105 GeV < $E_{\text{vis}}$ < 255 GeV. The visible mass of the event, $M_{\text{vis}}$, is greater than 95 GeV. The sum of the magnitude of the transverse momentum of all visible particles, $P_t, \text{sum}$, is greater than 40 GeV. The thrust of the event is less than 0.97. The Z candidate dijet has an energy, $E_Z$, in the range of 60 GeV < $E_Z$ < 175 GeV and has an invariant mass, $M_Z$, in the range of 35 GeV < $M_Z$ < 160 GeV. The angle between the two jets, $\theta_{jj}$, satisfies $\cos \theta_{jj} < 0.5$. The recoil mass against the $Z$ boson, computed as $M_{\text{recoil}} = \sqrt{\sqrt{s} - E_Z}^2 - |\mathbf{p}_Z|^2$, is in the range of 65 GeV < $M_{\text{recoil}}$ < 185 GeV. The Higgs candidate tau pair before the collinear approximation has an energy, $E_{\tau+\tau-}^{\text{col}}$, less than 140 GeV and an invariant mass, $M_{\tau+\tau-}^{\text{col}}$, in the range of 5 GeV < $M_{\tau+\tau-}^{\text{col}}$ < 125 GeV. The angle between the two tau candidates, $\theta_{\tau+\tau-}$, satisfies $\cos \theta_{\tau+\tau-} < -0.1$. The tau pair after the collinear approximation has an energy, $E_{\tau+\tau-}^{\text{col}}$, in the range of 30 GeV < $E_{\tau+\tau-}^{\text{col}}$ < 270 GeV and an invariant mass, $M_{\tau+\tau-}^{\text{col}}$, in the range of 15 GeV < $M_{\tau+\tau-}^{\text{col}}$ < 240 GeV.

We use a multivariate analysis using Boosted Decision Trees (BDTs) as implemented in the Toolkit for Multivariate Data Analysis [36] of the ROOT framework [37]. The input variables are

- $E_{\text{vis}}, P_{t,\text{vis}}, \cos \theta_{\text{miss}}$, where $P_{t,\text{vis}}$ is the magnitude of the visible transverse momentum and $\theta_{\text{miss}}$ is the angle of the missing momentum with respect to the beam axis;
- $M_Z, \cos \theta_{jj}, M_{\text{recoil}}, \cos \theta_Z$, where $\theta_Z$ is the angle of the $Z$ candidate momentum with respect to the beam axis;
- $M_{\tau+\tau-}^{\text{col}}, E_{\tau+\tau-}^{\text{col}}, \cos \theta_{\tau+\tau-}^{\text{col}}, \cos \theta_{\text{acop}}$, where $\theta_{\text{acop}}$ is the acoplanarity angle between the two tau candidates;
- $\sum_{\tau^+\tau^-} \log_{10}|d_0/\sigma_{d_0}|, \sum_{\tau^+\tau^-} \log_{10}|z_0/\sigma_{z_0}|$, where $d_0/\sigma_{d_0}$ and $z_0/\sigma_{z_0}$ are respectively the transverse and longitudinal impact parameters of the most energetic track in the tau candidate divided by their respective uncertainty estimated from the track fit;

![Fig. 3](image-url) Distributions of the invariant mass of the reconstructed tau lepton pairs at $\sqrt{s} = 250$ GeV for the $e^+e^- \rightarrow q\bar{q}h$ mode. $M_{\tau+\tau-}^{\text{col}}$ and $M_{\tau+\tau-}^{\text{col}}$ stand for the tau pair masses before and after the collinear approximation, respectively, for the signal. $M_{\tau+\tau-}^{\text{col}}$(bkg) is the tau pair mass with the collinear approximation for the background.
Pre-selected 1451 3526 2316 6
No cut 3318 7

Boosted Decision Trees for the Distributions of the multivariate discriminant from training

For the particles that are identified as electrons or positrons, we further require that $|d_{0}/\sigma_{d_{0}}| < 6$ and $|z_{0}/\sigma_{z_{0}}| < 3$, to reduce the electrons from secondary decays such as the tau lepton decays from the Higgs boson. We also require the track energy to be greater than 10 GeV, which removes the contamination from the $\gamma\gamma \rightarrow$ hadron background. The $e^{+}e^{-}$ pair whose combined mass is closest to the $Z$ boson mass is selected as the $Z$ boson candidate. To improve the mass and energy resolutions, the momenta of nearly neutral particles are added to that of the $Z$ candidate if their angle $\theta$ measured from at least one of the $e^{+}e^{-}$ satisfies $\cos \theta > 0.999$. The fraction of $e^{+}e^{-}\tau^{+}\tau^{-}$ signal events that survive the $Z$ boson selection is 61%.

We apply a tau finding algorithm to the remaining particles. Compared with the $q\bar{q}h$ mode, the algorithm is simpler due to the absence of hadronic jet activities aside from the tau decays. Starting with the charged particle with the highest energy as a working tau candidate, we define a cone around its momentum vector with a half-angle of $\theta_{c}=1.0$ rad. Particles inside the cone are combined with the tau candidate if the combined mass remains smaller than 2 GeV. The tau candidate is then set aside, and the tau finding is repeated until there are no more charged particles left. The tau candidates are then separated into two categories according to its charge. Within each category, the tau candidate with the highest energy is selected. The chosen $\tau^{\pm}\tau^{-}$ pair forms the Higgs candidate. Finally, the collinear approximation is applied to the selected tau candidates.

### 3.2.2 Event selection

A pre-selection is applied with the following requirements before proceeding with the multivariate analysis. The $Z \rightarrow e^{+}e^{-}$ candidate and the $h \rightarrow \tau^{+}\tau^{-}$ candidate are successfully reconstructed. The total number of charged tracks is 8 or fewer, which ensures statistical independence from the $q\bar{q}h$ mode. The visible energy is in the range of 100 GeV < $E_{vis}$ < 280 GeV. The visible mass is in the range of 85 GeV < $M_{vis}$ < 275 GeV. The sum of the magnitude of the transverse momentum of all visible particles, $P_{t,\text{sum}}$, is greater than 35 GeV. The $Z \rightarrow e^{+}e^{-}$ candidate has an energy in the range of 40 GeV < $E_{Z}$ < 160 GeV and an invariant mass in the range of 10 GeV < $M_{Z}$ < 145 GeV. The recoil mass against the $Z$ boson, $M_{\text{recoil}}$, is greater than 50 GeV.

- $M_{\text{col}}$, $E_{\text{col}}$, and $\cos \theta_{\text{col}}$, where $\theta_{\text{col}}$ is the angle of the Higgs candidate momentum with the collinear approximation measured from the beam axis.

The BDTs are trained using a set of statistically independent signal and background samples. The distribution of the resulting multivariate discriminant is shown in Fig. 4. We apply a final selection on the multivariate discriminant that maximizes the signal significance defined as $S/\sqrt{S+B}$, where $S$ and $B$ are the number of signal and background events, respectively. The final selected sample consists of 1232 signal and 543 background events. The estimated event yields before and after the selection are summarized in Table 3. The signal selection efficiency is 37% with a signal significance of 29, which corresponds to a statistical precision of $\Delta (\sigma \times BR) / (\sigma \times BR) = 3.4\%$. 

### 3.2 $e^{+}e^{-} \rightarrow e^{+}e^{-}h$

#### 3.2.1 $Z$ boson and tau lepton reconstruction

For the $e^{+}e^{-}h$ mode, we first reconstruct the $e^{+}e^{-}$ pair that forms a $Z$ boson candidate. A reconstructed particle is identified as an electron or a positron if its track momentum ($P_{\text{track}}$) and its associated energy deposits in the ECAL ($E_{\text{ECAL}}$) and HCAL ($E_{\text{HCAL}}$) satisfy the following criteria:

$$E_{\text{ECAL}}/(E_{\text{ECAL}} + E_{\text{HCAL}}) > 0.96,$$

$$(E_{\text{ECAL}} + E_{\text{HCAL}})/P_{\text{track}} > 0.6.$$
the following input variables:

\[ M_{\text{vis}}, E_{\text{vis}}, \cos \theta_{\text{miss}} - \cos \theta_{\text{thrust}}, \] where \( \theta_{\text{thrust}} \) is the angle of the thrust axis with respect to the beam axis;

\[ M_{Z}, M_{\text{recoil}}; \]

\[ M_{\tau^{+}\tau^{-}}, \cos \theta_{\tau^{+}\tau^{-}}, \cos \theta_{\text{acop}}; \]

\[ \sum_{\tau^{+},\tau^{-}} \log_{10} |d_{0}/\sigma_{d_{0}}|, \text{ and } \sum_{\tau^{+},\tau^{-}} \log_{10} |z_{0}/\sigma_{z_{0}}|. \]

A final selection on the multivariate discriminant is applied to maximize the signal significance, giving 76.3 signal and 44 background events. The final signal selection efficiency is estimated to be 70%, corresponding to a statistical precision of \( \Delta(\sigma \times BR)/(\sigma \times BR) = 14.4\% \).

### 3.3 \( e^{+}e^{-} \rightarrow \mu^{+}\mu^{-}h \)

#### 3.3.1 \( Z \) boson and tau lepton reconstruction

The reconstruction procedure of this mode is similar to that of the \( e^{+}e^{-}h \) mode, with the electron identification replaced by the muon identification. The muons are identified by requiring

\[ E_{\text{ECAL}}/(E_{\text{ECAL}} + E_{\text{HCAL}}) < 0.5, \]

\[ (E_{\text{ECAL}} + E_{\text{HCAL}})/P_{\text{trk}} < 0.6. \]

We additionally require the identified muons to satisfy \( |d_{0}/\sigma_{d_{0}}| < 3 \) and \( |z_{0}/\sigma_{z_{0}}| < 3 \), and to have a track energy greater than 20 GeV. The efficiency for selecting such muon pairs in \( \mu^{+}\mu^{-}\tau^{+}\tau^{-} \) signal events is 92%. The tau lepton reconstruction is the same as in the \( e^{+}e^{-}h \) mode.

#### 3.3.2 Event selection

The following pre-selection requirements are applied before proceeding with the multivariate analysis. The \( Z \rightarrow \mu^{+}\mu^{-} \) candidate and the \( h \rightarrow \tau^{+}\tau^{-} \) are successfully reconstructed. The total number of charged tracks is 8 or fewer. The visible energy is in the range of 105 GeV < \( E_{\text{vis}} < 280 \) GeV. The visible mass is in the range of 85 GeV < \( M_{\text{vis}} < 275 \) GeV.

The sum of the magnitude of the transverse momentum of all visible particles, \( P_{t,\text{sum}} \), is greater than 35 GeV. The \( Z \rightarrow \mu^{+}\mu^{-} \) candidate has an energy in the range of 45 GeV < \( E_{Z} < 145 \) GeV and an invariant mass in the range of 25 GeV < \( M_{Z} < 125 \) GeV. The recoil mass against the \( Z \) boson, \( M_{\text{recoil}} \), is greater than 75 GeV. The invariant mass of the tau pair system before the collinear approximation, \( M_{\tau^{+}\tau^{-}} \), is smaller than 170 GeV.

A multivariate analysis with BDTs is applied to the pre-selected events using the following input variables:

\[ M_{\text{vis}}, E_{\text{vis}}, P_{t,\text{vis}}; \]

\[ M_{Z}, \cos \theta_{Z}, M_{\text{recoil}}; \]

\[ M_{\tau^{+}\tau^{-}}, E_{\tau^{+}\tau^{-}}, \cos \theta_{\tau^{+}\tau^{-}}, M_{\text{coll}}; \]

\[ \sum_{\tau^{+},\tau^{-}} \log_{10} |d_{0}/\sigma_{d_{0}}|, \text{ and } \sum_{\tau^{+},\tau^{-}} \log_{10} |z_{0}/\sigma_{z_{0}}|. \]

We apply a final selection on the multivariate discriminant to maximize the signal significance, and obtain 101.9 signal and 31 background events. The final signal selection efficiency is 62%. The estimated event yields before and after the event selection are shown in Table 5. The signal significance is estimated to be 8.8, corresponding to a statistical precision of \( \Delta(\sigma \times BR)/(\sigma \times BR) = 11.3\% \).

### 4 Analysis at the center-of-mass energy of 500 GeV

At \( \sqrt{s} = 500 \) GeV, both the \( WW \)-fusion and the Higgs-strahlung processes have sizable contributions to the total signal cross section. We take into account the \( e^{+}e^{-} \rightarrow f \bar{f}h \) (except \( h \rightarrow \tau^{+}\tau^{-} \), \( e^{+}e^{-} \rightarrow 2f, e^{+}e^{-} \rightarrow 4f \), and \( e^{+}e^{-} \rightarrow 6f \) processes as backgrounds. The \( \gamma \gamma \rightarrow h \) hadron background is overlaid onto the signal and background MC samples, assuming an average rate of 1.7 events per bunch crossing [27]. The analysis in this section assumes an integrated luminosity of 500 fb\(^{-1}\). We report our results on the \( e^{+}e^{-} \rightarrow q\bar{q}h \) and \( e^{+}e^{-} \rightarrow \nu\bar{\nu}h \) modes. We do not give results for the \( e^{+}e^{-} \rightarrow e^{+}e^{-}h \) and \( e^{+}e^{-} \rightarrow \mu^{+}\mu^{-}h \) modes, as they do not contribute significantly to the overall sensitivity due to their small cross sections.
4.1 $e^+ e^- \rightarrow q \bar{q} h$

4.1.1 Reconstruction of isolated tau leptons and the $Z \rightarrow q \bar{q}$ decay

We start with the tau finding, following the same procedure described in Sect. 3.1.1. We additionally require the tau candidate to have an energy greater than 4 GeV. The energy of the neutrino from tau decays is corrected using the collinear approximation as before, resulting in a clear peak around the Higgs boson mass as can be seen in Fig. 5. We find that 54% of $q \bar{q} \tau^+ \tau^-$ signal events survive the requirement of finding exactly one pair of $\tau^+ \tau^-$. The invariant mass of all the particles, except those belonging to the two identified tau candidates, should be consistent with the $Z$ boson mass; however, a shift to a higher-mass value is observed, due to the presence of non-negligible background particles from $\gamma \gamma \rightarrow$ hadron events contaminating signal events. In order to mitigate the effect of these background particles, we use the $k_t$ clustering algorithm implemented in the FastJet package with a generalized jet radius of $R = 0.9$. The jets that are formed along the beam axis are then discarded. The remaining particles are clustered into two jets by using the Durham clustering algorithm to reconstruct the $Z$ boson decay.

4.1.2 Event selection

To facilitate the multivariate analysis, we impose the following pre-selections. The $Z \rightarrow q \bar{q}$ candidate and the $h \rightarrow \tau^+ \tau^-$ candidate are successfully reconstructed. The total number of charged tracks is between 8 and 70. The visible energy of the event is in the range of $140 \text{ GeV} < E_{\text{vis}} < 580 \text{ GeV}$. The visible mass of the event is in the range of $120 \text{ GeV} < M_{\text{vis}} < 575 \text{ GeV}$. The sum of the magnitude of all visible particles, $P_{\text{vis}}$, is greater than 70 GeV. The thrust of the event is less than 0.98. The $Z$ candidate dijet has an energy in the range of $50 \text{ GeV} < E_Z < 380 \text{ GeV}$ and has an invariant mass in the range of $5 \text{ GeV} < M_Z < 350 \text{ GeV}$. The recoil mass against the $Z$ boson is in the range of $40 \text{ GeV} < M_{\text{recoil}} < 430 \text{ GeV}$. The Higgs candidate tau pair before the collinear approximation has an energy, $E_{\tau^+ \tau^-}$, less than $270 \text{ GeV}$ and an invariant mass, $M_{\tau^+ \tau^-}$, less than $180 \text{ GeV}$, and the angle between the two tau candidates satisfies $\cos \theta_{\tau^+ \tau^-} < 0.7$. The tau pair after the collinear approximation has an energy in the range of $40 \text{ GeV} < E_{\text{col}} < 430 \text{ GeV}$ and an invariant mass, $M_{\text{col}}$, which is less than $280 \text{ GeV}$.

A multivariate analysis with BDTs is applied using the following input variables:

- $E_{\text{vis}}$, $P_{\text{vis}}$, $P_{\text{vis}}$, where $P_{\text{vis}}$ is the magnitude of the visible momentum;
- $M_Z$, $E_Z$, $\cos \theta_{\tau \tau}$, $M_{\text{recoil}}$;
- $M_{\tau^+ \tau^-}$, $\cos \theta_{\tau^+ \tau^-}$, $M_{\text{col}}$, $E_{\text{col}}$;
- $\sum |\vec{p}_{\tau^+ \tau^-} \cdot |\vec{p}_{\text{vis}}|$, and $\sum |\vec{p}_{\tau^+ \tau^-} \cdot |\vec{p}_{\text{vis}}|$.

After choosing the optimum threshold on the multivariate discriminant to maximize the signal significance, we are left with 782 signal and 335 background events. The final signal selection efficiency is 37%. The event yields before and after the selection are summarized in Table 6. The signal significance is found to be 23.4, corresponding to a statistical precision of $\Delta(\sigma \times \text{BR})/(\sigma \times \text{BR}) = 4.3 \%$.

4.2 $e^+ e^- \rightarrow \nu \bar{\nu} h$

4.2.1 Tau pair reconstruction

The tau finding algorithm proceeds in the same way as described for the $e^+ e^- h$ mode in Sect. 3.2.1, except that the half-angle of the cone $\theta_c$ around the most energetic track is modified to 0.76 rad. The most energetic positively and negatively charged tau candidates are combined to form a Higgs boson candidate.

4.2.2 Event selection

For the $\nu \bar{\nu} h$ mode, it is necessary to suppress the large background coming from the $e^+ e^- \rightarrow e^+ e^- + 2 f$ processes. We apply the following requirements to mitigate this background. A tau lepton pair $\tau^+ \tau^-$ is successfully reconstructed. The total number of tracks is less than 10. There is at least one charged track with a transverse momentum greater than 3 GeV and at least one charged track with an energy greater than 5 GeV. The missing momentum angle with respect to the beam axis satisfies $|\cos \theta_{\text{miss}}| < 0.98$. The acoplanarity angle between the two tau candidates satisfies $\cos \theta_{\text{acop}} < \ldots$
0.98. At this point, 94 % of the \( e^+e^- \rightarrow e^+e^- + 2f \) background is eliminated, while retaining 85 % of the signal events.

The following additional pre-selections are applied before the multivariate analysis. The visible energy is in the range of 10 GeV < \( E_{\text{vis}} \) < 265 GeV. The visible mass is in the range of 5 GeV < \( M_{\text{vis}} \) < 235 GeV. The missing mass, \( M_{\text{miss}} \), is greater than 135 GeV. The sum of the magnitude of the transverse momentum of all visible particles, \( P_t,_{\text{sum}} \), is greater than 10 GeV. The Higgs candidate tau pair before the collinear approximation has an energy, \( E_{\tau+\tau^-} \), less than 240 GeV and an invariant mass, \( M_{\tau+\tau^-} \), of less than 130 GeV. The angle between the two tau candidates satisfies \( \cos \theta_{\tau+\tau^-} < 0.8 \). A requirement on the transverse impact parameter of the tau candidate which gives a smaller value of the two is applied, such that \( \min |d_{0}/\sigma_{d_{0}}| > 0.01 \).

A multivariate analysis with BDTs is applied using the following input variables:

- Number of tracks with energy greater than 5 GeV;
- Number of tracks with transverse momentum greater than 5 GeV;
- \( M_{\tau}, E_{\tau}, P_{t,\tau}, P_{t,\text{sum}}, \cos \theta_{\tau,\text{thrust}}, \cos \theta_{\tau,\text{miss}} \);
- \( M_{\tau+\tau^-}, E_{\tau+\tau^-}, \cos \theta_{\tau+\tau^-}, \cos \theta_{\tau,\text{acop}} \);
- \( \log_{10} \min |d_{0}/\sigma_{d_{0}}| \).

We obtain 1642 signal and \( 1.11 \times 10^4 \) background events after optimizing the selection on the multivariate discriminant. The final signal selection efficiency is 30 %. The event yields before and after the selection are summarized in Table 7. The signal significance is 14.5, corresponding to a statistical precision of \( \Delta(\sigma \times \text{BR})/(\sigma \times \text{BR}) = 6.9 \% \).

In this mode, the \( e^+e^- \rightarrow Zh \rightarrow \nu\bar{\nu}h \) process and the \( e^+e^- \rightarrow \nu\bar{\nu}h \) process via WW-fusion are expected to be the dominant contributions. The effect of the interference between these two processes is studied using the distribution of the invariant mass of the neutrino pair \( M_{\nu}\nu \) computed from event generator information as shown in Fig. 6. A clear peak around the Z boson mass is visible, with a small contribution underneath it coming from the tail from higher masses, indicating that the interference of the \( e^+e^- \rightarrow Zh \rightarrow \nu\bar{\nu}h \) process and the WW-fusion is small. We hence split the events into two categories based on this generator-level variable, and define events with \( M_{\nu}\nu \leq 120 \text{ GeV} \) as “\( e^+e^- \rightarrow Zh \)” events, and those with \( M_{\nu}\nu > 120 \text{ GeV} \) as “WW-fusion” events. We find that the 1642 signal events after the final selection is composed of 13 % \( e^+e^- \rightarrow Zh \) events and 87 % WW-fusion events. The selection efficiencies for \( e^+e^- \rightarrow Zh \) and WW-fusion events are 33.5 and 29.2 %, respectively.

5 Discussion

5.1 Precision with the ILC running scenarios

We now discuss the prospects of the measurement precision with the ILC running scenarios proposed in Refs. [14,15] by

---

**Table 6** Event yields estimated for the \( e^+e^- \rightarrow q\bar{q}h \) mode at \( \sqrt{s} = 500 \text{ GeV} \), assuming an integrated luminosity of 500 fb\(^{-1}\) and beam polarizations of \( P(e^-, e^+) = (-0.8, +0.3) \). Refer to Table 3 for the row definitions.

| Signal       | \( f/\sqrt{s} \) | \( 2f \)   | \( 4f \)   | \( 6f \)   |
|--------------|------------------|----------|----------|----------|
| No cut       | 2131             | \( 1.266 \times 10^5 \) | \( 1.320 \times 10^7 \) | \( 9.989 \times 10^8 \) | \( 6.929 \times 10^9 \) |
| Pre-selected | 1088             | 2889     | \( 3.013 \times 10^4 \) | \( 1.144 \times 10^5 \) | \( 1.737 \times 10^4 \) |
| Final        | 782.1            | 17.6     | 1.5      | 275      | 41       |

**Table 7** Event yields estimated for the \( e^+e^- \rightarrow \nu\bar{\nu}h \) mode at \( \sqrt{s} = 500 \text{ GeV} \), assuming an integrated luminosity of 500 fb\(^{-1}\) and beam polarizations of \( P(e^-, e^+) = (-0.8, +0.3) \). Refer to Table 3 for the row definitions.

| Signal       | \( f/\sqrt{s} \) | \( 2f \)   | \( 4f \)   | \( 6f \)   |
|--------------|------------------|----------|----------|----------|
| No cut       | 5534             | \( 1.232 \times 10^5 \) | \( 1.320 \times 10^7 \) | \( 9.989 \times 10^8 \) | \( 6.929 \times 10^9 \) |
| Pre-selected | 3623             | 1543     | \( 5.957 \times 10^4 \) | \( 1.756 \times 10^7 \) | 990.8     |
| Final        | 1642             | 65.5     | 379      | \( 1.043 \times 10^4 \) | 238       |

---

Fig. 6 Distribution of neutrino-pair invariant mass, \( M_{\nu}\nu \), using event generator information for the \( e^+e^- \rightarrow \nu\bar{\nu}h \) mode at \( \sqrt{s} = 500 \text{ GeV} \). The effect of detector simulation is not applied.
extrapolating the results presented in the previous sections. Table 8 summarizes the integrated luminosities for various center-of-mass energies and beam polarizations for three different scenarios we consider.

In order to estimate the statistical precision of the cross section times the branching ratio measurements with electron and positron beam polarizations other than (−0.8, +0.3) used in the previous sections, we need to know the corresponding selection efficiencies for the signal and background processes. In the following, we assume the same selection efficiencies obtained in the previous sections for all of these beam polarizations, although in principle the angular distributions of the final states may depend on the beam polarizations. This assumption is nevertheless justified as follows. The $e^+e^-\rightarrow Zh$ process is mediated by the $s$-channel $Z$ boson exchange with the vector or the axial vector coupling, which forbids the same-sign helicity states $(\pm1, \pm1)$, while giving more or less the same angular distributions for the opposite-sign helicity states $(\mp1, \mp1)$. On the other hand, the $WW$-fusion process proceeds only through the left-right helicity states $(-1, +1)$, since the $W$ boson couples only to the left-handed $e^-$ and the right-handed $e^+$. For the signal processes, therefore, their angular distributions stay the same for the active (i.e. opposite) helicity states, independently of the choice of beam polarizations. The same reasoning applies to the background processes with the $s$-channel $\gamma/Z$ exchange or those involving $W$ bosons coupled to the initial state $e^+$ or $e^-$. On the other hand, the processes involving $t$-channel photon exchange or photon-photon interactions do not forbid the same-sign helicity states. However, since the probability of finding an electron and a positron in the same-sign helicity states is the same for both the $(-0.8, +0.3)$ and $(+0.8, -0.3)$ beam polarizations, the efficiencies for such background processes with the same-sign helicity states should also be the same. In our estimation, we do not use the results of the $(\pm0.8, \pm0.3)$ beam polarizations, since the signal cross sections are small and the integrated luminosities collected at these beam polarizations are foreseen to be small. Under these assumptions, the selection efficiencies will not depend on the choice of beam polarizations. We can then estimate the projected statistical precision for other scenarios by calculating the number of signal and background events with the production cross sections and the integrated luminosities for individual beam polarizations, according to the running scenarios. The result from this estimation is summarized in Table 9.

### Table 8: Integrated luminosity (in $\text{fb}^{-1}$) for various beam polarizations $(P(e^-), P(e^+))$ at $\sqrt{s} = 250$ and 500 GeV for three running scenarios. “Nominal” is the scenario described in the TDR [9] and used in the main results of this paper. “Initial” and “Full” are the scenarios proposed in Refs. [14,15]

| Scenario | $\sqrt{s}$ (GeV) | $(-80, +30)$ | $(+80, -30)$ | $(-80, -30)$ | $(+80, +30)$ | Total |
|----------|-----------------|-------------|-------------|-------------|-------------|------|
| Nominal  | 250             | 250         | 0           | 0           | 0           | 250  |
|          | 500             | 500         | 0           | 0           | 0           | 500  |
| Initial  | 250             | 337.5       | 112.5       | 25          | 25          | 500  |
|          | 500             | 200         | 200         | 50          | 50          | 500  |
| Full     | 250             | 1350        | 450         | 100         | 100         | 2000 |
|          | 500             | 1600        | 1600        | 400         | 400         | 4000 |

5.2 Precision of the $h \rightarrow \tau^+\tau^-$ branching ratio

So far, we discussed the precision of the production cross section times the branching ratio, which is the primary information we will obtain from the experiments. Here, we discuss the prospects for measuring the branching ratio itself. At the ILC, the production cross section for the Higgs-strahlung process can be separately measured using the recoil mass technique [10,13]. The cross section for the $WW$-fusion process can also be determined by using the branching ratio for the $h \rightarrow b\bar{b}$ decay [10]. The obtained cross section values allow us to derive the branching ratio for the $h \rightarrow \tau^+\tau^-$ decay.

At $\sqrt{s} = 250$ GeV, the contributions of the $WW$-fusion and $ZZ$-fusion processes are negligible. Therefore, we can use the Higgs-strahlung cross section to derive the branching ratio. The Higgs-strahlung cross section $\sigma_{Zh}$ can be measured to a statistical precision of $\Delta\sigma_{Zh}/\sigma_{Zh} = 2.5\%$ with the nominal TDR running scenario [13]. This improves to a subpercent level with the full running scenario [14].

At $\sqrt{s} = 500$ GeV, both the Higgs-strahlung and the $WW$-fusion processes contribute to the Higgs boson production, whereas the contribution of the $ZZ$-fusion process is negligible. For the $e^+e^-\rightarrow \nu\bar{\nu}h$ mode, in which both processes are present, it is in principle possible to estimate the contributions from the Higgs-strahlung and the $WW$-fusion processes separately, as discussed in Sect. 4.2.2. However, we do not use this mode here for the estimate. The expected statistical precision of the branching ratio after combining all the modes except the $\nu\bar{\nu}h$ mode is 3.6 % for the nominal running scenario. This improves to 1.4 % with the full running scenario, where we assume $\Delta\sigma_{Zh}/\sigma_{Zh} = 1.0\%$. 

\[ \]
5.3 Systematic uncertainties

The MC statistical uncertainties are found to have negligible impact on the results. The systematic uncertainty in the luminosity measurement has been estimated to be 0.1% or better for the ILC [41] and is not expected to be a significant source of systematic errors. The uncertainties in the selection criteria, such as those caused by the uncertainty in the momentum/energy resolutions and tracking efficiencies, are not included in this analysis, since they are beyond the scope of this paper.

6 Summary

We have evaluated the measurement precision of the Higgs boson production cross section times the branching ratio of decay into tau leptons at the ILC. The study is based on the full detector simulation of the ILD model. The dominant Higgs boson production mechanisms were studied at the center-of-mass energies of 250 and 500 GeV, assuming the nominal luminosity scenario presented in the ILC TDR. The analysis results are then scaled up to the running scenarios taking into account realistic running periods and a possible luminosity upgrade.

The results for the various modes and scenarios are summarized in Table 9. In short, the cross section times the branching ratio can be measured with a statistical precision of $\Delta(\sigma \times BR)/(\sigma \times BR) = 2.6$ and 1.0% for the nominal and full running scenarios, respectively. We evaluate the statistical precision of $BR(h \rightarrow \tau^+\tau^-)$ to be 3.6% for the nominal TDR integrated luminosity and 1.4% for the full running scenario, respectively. These results serve to provide primary information on the expected precision of measuring Higgs decays to tau leptons at the ILC, which will be useful for future phenomenological studies on physics beyond the SM.

Acknowledgments The authors would like to thank all the members of the ILC Physics Working Group. We thank H. Yokoyama for providing the code for the signal event generator interfacing GRACE and TAUOLA. This work has been partially supported by JSPS Grants-in-Aid for Science Research No. 22244031 and the JSPS Specially Promoted Research No. 23000002.
