Eur. Phys. J. C manuscript No.
(will be inserted by the editor)

Measurement of the H to ZZ branching fraction at 350 GeV and 3 TeV CLIC

N. Vukašinović[a,1], I. Božović-Jelisavčić[1], G. Kačarević[1], M. Radulović[2], J. Stevanović[2], G. Milutinović-Dumbelović[1], T. Agatonović-Jovin[1], I. Smiljanić[1]

[1]"VINČA" Institute of Nuclear Sciences - National Institute of the Republic of Serbia, University of Belgrade, Belgrade, Serbia
[2]University of Kragujevac, Faculty of Science, Kragujevac, Serbia

Received: date / Accepted: date

Abstract In this paper we investigate the prospects for measuring the Standard Model Higgs boson decay into a pair of Z bosons at the future Compact Linear Collider (CLIC) at 350 GeV and 3 TeV centre-of-mass energies. Studies are done with full GEANT4 based simulation of detector for CLIC, taking into consideration all relevant physics and beam-induced background processes. It is shown that the product of the Higgs production cross-section and the branching fraction BR($H \to ZZ^*$) can be measured with a relative statistical uncertainty of 20% and 3.0% at centre-of-mass energies of 350 GeV and 3 TeV, respectively, in the semileptonic final states. Integrated luminosity of 1 ab^{-1} (5 ab^{-1}) is assumed at 350 GeV (3 TeV) centre-of-mass energies.

1 Introduction

As a staged e^+e^- collider, CLIC can provide a comprehensive physics program of the measurements in the Higgs sector. Cumulative statistics of data collected from all energy stages enables precision measurements of the Higgs couplings, mass and width. Higher centre-of-mass energies (i.e. above 1 TeV) enhance sensitivity to measure the Higgs self-coupling as well as to probe Beyond the Standard Model physics (BSM) in the Higgs sector. Designed to operate at the highest centre-of-mass energies among proposed e^+e^- collider projects, together with the possibility to polarise electron beam up to 80%, CLIC offers extensive set of physics measurements needed for understanding of the Higgs sector.

In general, it is important to measure the Higgs couplings with the highest possible precision. Most of the BSM models predict Higgs couplings to electroweak bosons to deviate from the Standard Model (SM) predictions at the order of a percent [1]. As discussed in [2], a global fit to the full set of data from all energy stages allows extraction of the Higgs couplings with the required precision.

So far, only the result of BR($H \to ZZ^*$) measurement obtained at 1.4 TeV centre-of-mass energy CLIC has been published in [2], while an estimation is made for 3 TeV centre-of-mass energy on the basis of luminosity scaling of the 1.4 TeV result. Though inferior in precision in comparison to high-energy measurements, the 350 GeV measurement completes the set of Higgs branching fraction measurements at CLIC, serving as input to a global fit of the Higgs couplings, i.e. the Effective Field Theory (EFT) approach [3].

In this paper we determine the CLIC statistical precision to measure $H \to ZZ^*$ branching ratio at 350 GeV and 3 TeV centre-of-mass energies in the semileptonic final state, in the full simulation of experimental conditions. The semileptonic final state is chosen due to the fact that it is less contaminated with irreducible background than the hadronic final state.

The paper is organized as follows: A detector for CLIC is described in Section 2; Section 3 lists possible Higgs production mechanisms at CLIC, while Sections 4 to 6 provide details on event samples, method of the analyses and statistical precision of the measurements.

2 The CLIC_ILD detector model

The CLIC_ILD detector [4], based on the ILD detector concept for ILC [5], has been modified to the experimental conditions at CLIC. More recently, CLICdet detector concept [6] has been developed. Both detector concepts use fine-grained electromagnetic and hadronic calorimeters (ECAL and HCAL) optimised for the Particle Flow Algorithm (PFA) employed in event reconstruction [7]. Excellent muon track-momentum resolution of $\sigma(p_t/p_t^2) \sim 2 \cdot 10^{-5}$ GeV^{-1} [4] enables precise reconstruction of leptonic Z decays, while the jet-energy resolution ranges between $3.5\% - 5\%$ depend-

[a]e-mail: nvukasinovic@vin.bg.ac.rs

ing on the jet energy [4]. The later is considered crucial for separation of close-by jets from Higgs, W and Z bosons. Assumption on detector model however has no impact on statistical precision of the measurements discussed in this paper.

3 Higgs production mechanisms at CLIC

CLIC is foreseen to run at three energy stages: 380 (350) GeV, 1.5 TeV and 3 TeV centre-of-mass energies. The first stage enables precision measurement both in the Higgs and top-quark sector. Taking into account beam polarisation, CLIC will produce about $4.5 \cdot 10^6$ Higgs bosons at all stages [8]. As illustrated at Figure 1 [2], the main Higgs production mechanism in the first stage is Higgsstrahlung (ZH), while at around 500 GeV centre-of-mass energy $WW-$fusion starts to dominate. The cross-section for the Higgsstrahlung process (ZH) at 350 GeV, where the primary Z decays hadronically, is 93.44 fb, while at 3 TeV the cross-section for Higgs production in $WW-$fusion is 415.05 fb. The branching fraction for the $H \to ZZ^*$ decay is 2.89% [9]. Z boson branching fractions to hadrons and leptons are 69.91% and 3.37%, respectively [10]. The expected number of ZH events (where the primary Z decays hadronically) is around 93,000 in 1 ab^{-1} of unpolarized data. The expected number of $H\nu_e\bar{\nu}_e$ events is around $2 \cdot 10^6$ in 5 ab^{-1} of unpolarized data. The above estimates assume a realistic CLIC luminosity spectrum with Initial State Radiation (ISR) included. The CLIC accelerator baseline design foresees sharing of the running time for -80 % and $+80$% e^- polarization in the ratio 80:20 at 1.5 TeV and 3 TeV, while the ratio 50:50 is assumed at 380 GeV, with no e^+ polarization at all stages [11]. Such polarisation scheme will be assumed as *beam polarisation* throughout the text. Due to the chiral nature of the charged-current interaction, WW-fusion is much more affected by the e^- polarization than the Higgsstrahlung. With the proposed polarization scheme, the cross-section for WW-fusion will increase by a factor of 1.48 [11]. Impact of the beam polarisation on statistical precision of $\sigma(H\nu\bar{\nu}) \times BR(H \to ZZ^*)$ measurement at 3 TeV centre-of-mass energy is discussed in Sec. 6.

4 Event samples and preselection

4.1 Event samples

Signal and background events are simulated in Whizard 1.95 [12]. The processes of hadronization and fragmentation of final-state quarks and gluons are simulated in Pythia 6.4 [13]. The Higgs boson mass is assumed to be 126 GeV at the generator level. The CLIC luminosity spectrum and interactions between beams are obtained using GuineaPig 1.4.4

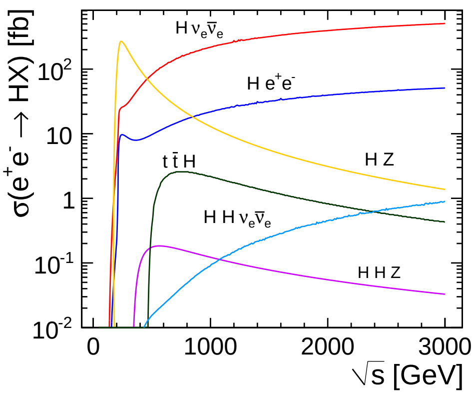

Fig. 1: Unpolarized cross-sections as a function of centre-of-mass energy for the main Higgs production processes at an e^+e^- collider, assuming a Higgs boson mass of 126 GeV.

[14], while hadron production from Beamstrahlung photons is simulated with Pythia 6.4. The full list of signal and background processes with the corresponding cross-sections is given in Table 1a and Table 1b, at 350 GeV and 3 TeV, respectively. One may note that $\gamma_{BS}\gamma_{BS}$ and $e^{\pm}\gamma_{BS}$ processes involving photons from Beamstrahlung (γ_{BS}) are not considered explicitly as background at 350 GeV (Table 1a) due to the fact that these processes are much less pronounced at lower centre-of-mass energies and thus contribute negligibly to this study [4]. Although in order to simulate realistic experimental environment at CLIC, the $\gamma_{BS}\gamma_{BS} \to hadrons$ events are overlaid before the digitisation phase on the reconstructed signal and background events at all centre-of-mass energies. At 3 TeV (Table 1b), simulation of the background process $e^-e^+ \to q\bar{q}l^+l^-\nu\bar{\nu}$ was available only at the generator level. Approximately 99.8% of these events can be removed by considering optimised intervals of the Higgs mass and off-shell Z mass. It is estimated that less than 30 $q\bar{q}l^+l^-\nu\bar{\nu}$ events will remain in 5 ab^{-1} of data, which does not poses relevant influence on the statistical uncertainty of the branching fraction measurement.

Interactions with the detector are simulated using the CLIC_ILD detector model within the Mokka simulation package [15] based on GEANT4 [16]. Event reconstruction is based on the Particle Flow Algorithm (PFA) [17] that is implemented in the Pandora algorithm. For the jet clustering, the k_T algorithm [18] is used in the exclusive mode, implemented in the Fast Jet processor [19]. The Isolated Lepton Finder Marlin processor [20] is used for isolated lepton (e, μ) identification. Tagging of beauty and charm jets is performed with the LCFIPlus processor [21]. The TMVA package [22] is used for the multivariate classification (MVA) of signal and background events using their kinematic properties. The simulation, reconstruction and analyses are done with the ILCDIRAC [23].

Table 1: List of considered processes with the corresponding cross-sections, expected number of events and simulated number of events at 350 GeV (a) and 3 TeV (b) centre-of-mass energies. N_{sim} stands for the number of simulated events, with exactly two truth-linked[b] leptons (electrons or muons) in case of the signal.

(a)

Signal process	$\sigma(fb)$	$N@1ab^{-1}$	N_{sim}
$e^-e^+ \to HZ; Z \to q\bar{q}, H \to ZZ^*, ZZ^* \to q\bar{q}l^+l^- (l = e, \mu)$	0.24	240	17721
Background processes	$\sigma(fb)$	$N@1ab^{-1}$	N_{sim}
$e^-e^+ \to HZ; Z \to q\bar{q}, H \to others$	7.0	$7 \cdot 10^3$	$77 \cdot 10^3$
$e^-e^+ \to HZ; Z \to q\bar{q}, H \to \mathrm{WW} \to 4q$	10.5	$10.5 \cdot 10^3$	$12 \cdot 10^3$
$e^-e^+ \to HZ; Z \to \mu^+\mu^-, H \to others$	2.3	$2.3 \cdot 10^3$	$85 \cdot 10^3$
$e^-e^+ \to HZ; Z \to e^+e^-, H \to others$	2.3	$2.3 \cdot 10^3$	$85 \cdot 10^3$
$e^-e^+ \to HZ; Z \to \mu^+\mu^-, H \to \mathrm{WW} \to 4q$	0.7	$0.7 \cdot 10^3$	$14 \cdot 10^3$
$e^-e^+ \to HZ; Z \to e^+e^-, H \to \mathrm{WW} \to 4q$	0.7	$0.7 \cdot 10^3$	$14 \cdot 10^3$
$e^-e^+ \to q\bar{q}q\bar{q}l^+l^-$	4.5	$4.5 \cdot 10^3$	$44 \cdot 10^3$
$e^-e^+ \to qqqq$	5847	$5.8 \cdot 10^6$	$191 \cdot 10^3$
$e^-e^+ \to q\bar{q}l^+l^-$	1704	$1.7 \cdot 10^6$	$746 \cdot 10^3$

(b)

Signal process	$\sigma(fb)$	$N@5ab^{-1}$	N_{sim}
$e^-e^+ \to H\nu\bar{\nu}; H \to ZZ^*, ZZ^* \to q\bar{q}ll, (l = e, \mu)$	1.13	5650	16752
Background processes	$\sigma(fb)$	$N@5ab^{-1}$	N_{sim}
$e^-e^+ \to H\nu\bar{\nu}; H \to \mathrm{WW}, \mathrm{WW} \to 4q$	43	$218 \cdot 10^3$	$219 \cdot 10^3$
$e^-e^+ \to H\nu\bar{\nu}; H \to b\bar{b}$	233	$1.2 \cdot 10^6$	$1.1 \cdot 10^6$
$e^-e^+ \to H\nu\bar{\nu}; H \to c\bar{c}$	11.7	$58.5 \cdot 10^3$	$52 \cdot 10^3$
$e^-e^+ \to H\nu\bar{\nu}; H \to gg$	35.2	$176 \cdot 10^3$	$128 \cdot 10^3$
$e^-e^+ \to H\nu\bar{\nu}; H \to others$	91	$452 \cdot 10^3$	$465 \cdot 10^3$
$e^-e^+ \to q\bar{q}l^+l^-$	3320	$16.6 \cdot 10^6$	$2 \cdot 10^6$
$e^-e^+ \to qql\nu$	5561	$27.8 \cdot 10^6$	$3.1 \cdot 10^6$
$e^-e^+ \to q\bar{q}\nu\bar{\nu}$	1317	$6.6 \cdot 10^6$	$569 \cdot 10^3$
$\gamma\gamma \to q\bar{q}l^+l^-$	20293	$135.7 \cdot 10^6$	$2.5 \cdot 10^6$
$\gamma\gamma \to q\bar{q}$	112039	$517.4 \cdot 10^6$	$1 \cdot 10^6$
$e^\pm\gamma \to q\bar{q}e$	20661	$60.3 \cdot 10^6$	$462 \cdot 10^3$
$e^\pm\gamma \to qq\nu$	36832	$138.3 \cdot 10^6$	$692 \cdot 10^3$
$e^-e^+ \to q\bar{q}l^-l^-\nu\bar{\nu}$	3.4	$17 \cdot 10^3$	$10 \cdot 10^3$

[b] *Truth − linked* refers to the reconstructed leptons that can be matched with the corresponding Monte Carlo (MC) information confirming that reconstructed lepton is a descendant of a Higgs boson.

4.2 Preselection

The method of the analyses uses staged selection of signal and background events, through preselection and MVA based selection. The main goal of the preselection is to reduce backgrounds with large cross-sections. Preselection requirement for measurements at both 350 GeV and 3 TeV is that two isolated leptons (electrons or muons) are found per event. Lepton isolation is optimized with the Isolated Lepton Finder according to track energy, ratio of energies deposited in the electromagnetic and hadronic calorimeters as well as impact parameters of the lepton tracks.

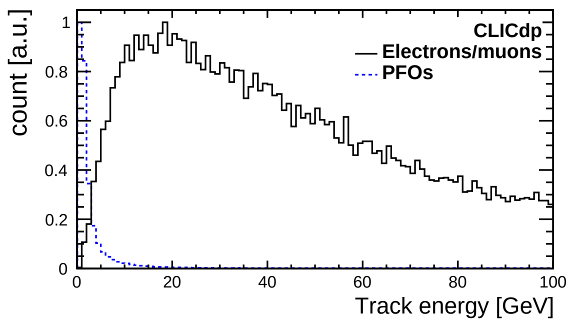

Fig. 2: Energy of the reconstructed signal leptons (solid) and other reconstructed PFO objects (dashed) in signal events, at 3 TeV centre-of-mass energy.

Electrons and muons originating from ZZ^* decays have energies that are much higher than the energy of a typical particle in a jet (PFO), as illustrated in Figure 2. Selection is optimized in a way that muons and electrons are required to have track energy of at least 5 GeV (6 GeV) at 350 GeV (3 TeV).

Since the final state leptons from a signal event will behave in the detector as if they originate from the primary vertex, the range of impact parameter components has been optimized in addition. The impact parameter of a track describes the distance between the track and the primary vertex (PV) at the point of track's closest approach to the PV. It can be decomposed into longitudinal (z_0) and transverse (d_0) components that are combined into R_0:

$$R_0 = \sqrt{{z_0}^2 + {d_0}^2} \tag{1}$$

In a signal event, electrons and muons will have significantly smaller impact parameters than other reconstructed particles. Thus it is required: $d_0 < 0.02$ mm and $z_0 < 0.02$ mm at 350 GeV and $d_0 < 0.02$ mm and $R_0 < 0.03$ mm at 3 TeV centre-of-mass energies.

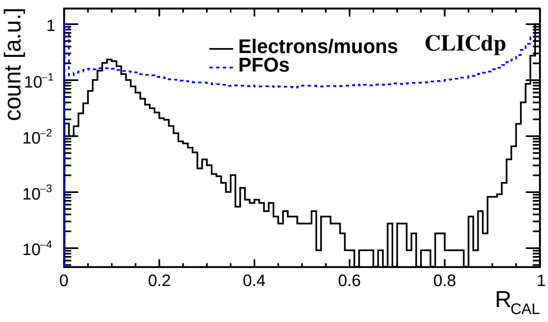

Fig. 3: Calorimeter energy ratio R_{CAL} (Eq.2) of reconstructed electrons and muons (solid) and other reconstructed particles (dashed) for signal, at 3 TeV centre-of-mass energy.

Muons can be distinguished from electrons using the ratio R_{CAL} of energy deposits in ECAL and HCAL:

$$R_{CAL} = E_{ECAL}/(E_{ECAL} + E_{HCAL}) \tag{2}$$

Because electrons are contained within the ECAL, they peak at R_{CAL} = 1. Muons deposit a minimal amount of energy throughout the calorimeters and have a peak at R_{CAL} = 0.1. This is illustrated in Figure 3, for reconstructed signal at 3 TeV centre-of-mass energy. In order to remove particles which do not behave as electrons or muons in the calorimeters, the calorimeter energy ratio R_{CAL} is required to be: ($R_{CAL} > 0.9$) or ($R_{CAL} < 0.35$) at 350 GeV and ($R_{CAL} > 0.94$) or ($0.02 < R_{CAL} < 0.35$) at 3 TeV centre-of-mass energies.

Finally, the leptons from the signal are required to be isolated within event. In order to be considered as isolated, lepton tracks are required to pass an optimized two-dimensional cut on cone energy vs. track energy, where the cone energy sums up all particle energies, without the isolated lepton candidate, in a cone size of approximately 6°around the isolated lepton track. The isolation requirement is:

$$E^2_{cone} < B \cdot E_{trk} + C \tag{3}$$

where E_{trk} and E_{cone} are track energy and cone energy, respectively, while the constants B and C optimized to achieve efficient isolation of signal leptons are found to be: B = 48 GeV and C = 16 GeV^2 at 350 GeV and B = 20 GeV and C = -20 GeV^2 at 3 TeV centre-of-mass energies. Figure 4 shows the energy within a cone size of 6°around a lepton's track as a function of a track energy, at 350 GeV (Figure 4a) and 3 TeV (Figure 4b).

Particles which are not identified as isolated leptons are clustered into jets. This is achieved using the FastJet implementation of the k_T algorithm. Events are forced into 4 (2) jets at 350 GeV (3 TeV) centre-of-mass energy. The distance parameter R corresponding to the effective jet width is chosen to be 1.1 at 350 GeV and 0.7 at 3 TeV. Reconstructed leptons and jets are combined to give the Z boson candidates. At 3 TeV, there are two fermion pairs (qq and ll) in the final state and fermionic pair with the higher invariant mass is considered to originate from on-shell Z boson (Z), while the other pair of fermions comes from the off-shell Z boson (Z^*) decay. Since the signal signature at 350 GeV is 6 fermions in the final state ($qqll$ from ZZ^* decays plus a qq pair from the primary Z decay) reconstruction of a Higgs boson candidate is different than at 3 TeV. Invariant mass of a combination of di-jet and di-lepton pairs that is closest to the simulated Higgs boson mass (126 GeV) is taken as a Higgs candidate mass. The other pair of quarks is then considered to come form a primary Z boson decay.

With the criteria described above, preselection efficiencies for the signal are 77% and 67% at 350 GeV and 3 TeV, respectively. Preselection efficiences for signal and background processes are given in Table 2a at 350 GeV and in Table 2b at 3 TeV. Signal efficiencies of the isolation curves are 93% and 86% at 350 GeV and 3 TeV centre-of-mass energies.

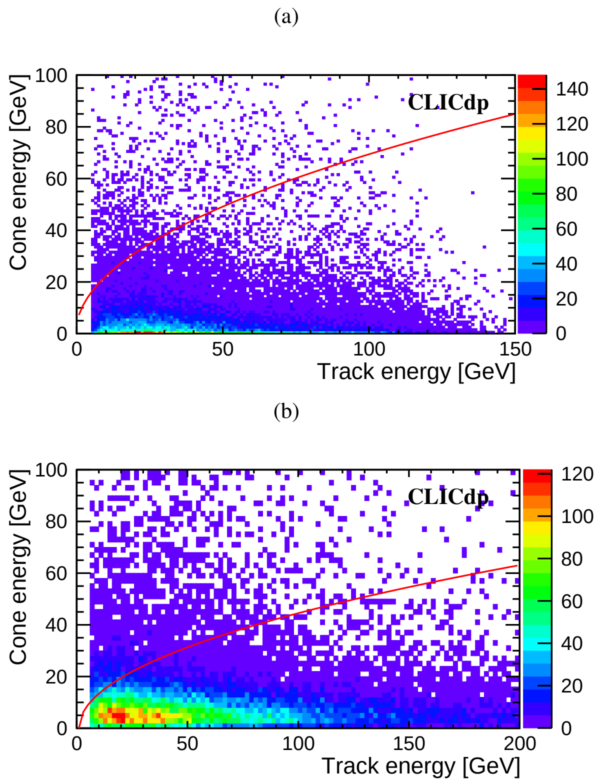

Fig. 4: Cone energy as a function of track energy of reconstructed electrons and muons at 350 GeV (a) and at 3 TeV centre-of-mass energy (b). The red line represents the polynomial distribution from Eq.3, separating isolated lepton tracks from other particles in a jet.

In order to recover for Bremsstrahlung of the final state leptons, four-vectors of photons in a cone of 3° around lepton candidate are added to the lepton's four-vector. This correction (lepton dressing) is done prior to any preselection step. Bremsstrahlung correction does not have a significant impact on preselection efficiencies, while it improves the resolution of the Z reconstruction and consequently of the MVA performance.

In Figure 5, stacked histograms for signal and background are given for preselected events. Figures 5a and 5b show the Higgs mass distributions from the reconstructed Z bosons at 350 GeV and 3 TeV, respectively. With the described preselection, background dominates the signal up to 1000 times at both centre-of-mass energies. Background rejection rates are around 3% and 3‰ at 350 GeV and 3 TeV centre-of-mass energies.

Table 2: Summary of preselection efficiencies for the signal and background with number of events that pass preselection (N_{presel}) in the simulated samples and with expected integrated luminosities at 350 GeV (a) and 3 TeV (b).

(a)

Signal process	$\varepsilon_{\mathrm{presel}}(\%)$	$N_{\mathrm{presel}}@1\mathrm{ab}^{-1}$	N_{presel}
$e^-e^+ \to HZ; Z \to q\bar{q}, H \to ZZ^*, ZZ^* \to q\bar{q}l^+l^- (l = e, \mu)$	77	185	13645
Background processes	$\varepsilon_{\mathrm{presel}}(\%)$	$N_{\mathrm{presel}}@1\mathrm{ab}^{-1}$	N_{presel}
$e^-e^+ \to HZ; Z \to q\bar{q}, H \to others$	0.37	26	285
$e^-e^+ \to HZ; Z \to q\bar{q}, H \to \mathrm{WW} \to 4q$	0.42	44	50
$e^-e^+ \to HZ; Z \to \mu^+\mu^-, H \to others$	61	1421	51850
$e^-e^+ \to HZ; Z \to e^+e^-, H \to others$	62	1445	52700
$e^-e^+ \to HZ; Z \to \mu^+\mu^-, H \to \mathrm{WW} \to 4q$	60	434	8400
$e^-e^+ \to HZ; Z \to e^+e^-, H \to \mathrm{WW} \to 4q$	60	434	8400
$e^-e^+ \to q\bar{q}q\bar{q}l^+l^-$	21	939	9240
$e^-e^+ \to qqqq$	0.32	18560	611
$e^-e^+ \to q\bar{q}l^+l^-$	11.4	193800	85044

(b)

Signal processes	$\varepsilon_{\mathrm{presel}}(\%)$	$N_{\mathrm{presel}}@5\mathrm{ab}^{-1}$	N_{presel}
$e^-e^+ \to H\nu\bar{\nu}; H \to ZZ^*, ZZ^* \to q\bar{q}ll, (l = e, \mu)$	67	3788	11224
Background process	$\varepsilon_{\mathrm{presel}}$(‰)	$N_{\mathrm{presel}}@5\mathrm{ab}^{-1}$	N_{presel}
$e^-e^+ \to H\nu\bar{\nu}; H \to \mathrm{WW}, \mathrm{WW} \to 4q$	1.7	371	372
$e^-e^+ \to H\nu\bar{\nu}; H \to b\bar{b}$	0.6	720	660
$e^-e^+ \to H\nu\bar{\nu}; H \to c\bar{c}$	0.6	35	31
$e^-e^+ \to H\nu\bar{\nu}; H \to gg$	0.9	158	115
$e^-e^+ \to H\nu\bar{\nu}; H \to others$	45	20340	20925
$e^-e^+ \to q\bar{q}l^+l^-$	7.5	124500	15000
$e^-e^+ \to qql\nu$	3	83400	9300
$e^-e^+ \to q\bar{q}\nu\bar{\nu}$	0.7	4620	398
$\gamma\gamma \to q\bar{q}l^+l^-$	11	1500000	27500
$\gamma\gamma \to q\bar{q}$	1	517400	1000
$e^\pm\gamma \to q\bar{q}e$	8.8	530640	4066
$e^\pm\gamma \to qq\nu$	1.4	193620	968

5 Multivariate analysis

5.1 MVA at 350 GeV

Separation of signal from background is done by employing a multivariate analysis based on the Boosted Decision Trees (BDT) classifier [22]. At 350 GeV, an MVA is trained with 22 observables: mass of the on-shell Z boson, mass of the off-shell Z boson, mass of the primary Z, invariant mass of two selected leptons, invariant mass of two reconstructed jets, mass of a Higgs candidate, visible energy in the event, difference between the visible energy and the Higgs energy, polar angle of a Higgs candidate, angle between on-shell and off-shell Z bosons in the plane perpendicular to the beam axis, number of all PFO objects in an event, jet transition variables ($-\log y_{12}$,$-\log y_{23}$ and $-\log y_{45}$), b-tag and c-tag probabilities of jets sorted by decreasing transverse momentum of a jet, transverse momenta and energies of isolated leptons. Individual leptons are sorted in a way that the higher transverse momentum lepton has index 1. Higgs mass is constrained in the interval ($50\,\mathrm{GeV} < m_H < 170\,\mathrm{GeV}$). The three most sensitive observables in the BDT training phase are found to be: energy of the reconstructed lepton with the highest p_{T}, jet transition variable ($-\log y_{23}$) and mass of the reconstructed primary Z.

The BDT output variable cut-off value is chosen to maximize the statistical significance S:

$$S = N_S/\sqrt{N_S + N_B} \tag{4}$$

where $N_{S,B}$ denotes the number of selected signal and background events. The optimal BDT cut is found to be 0.20 corresponding to a statistical significance of 5. The overall efficiency of the signal including preselection and MVA selection is found to be approximately 19%, due to the relatively low MVA efficiency of approximately 25%. Since the signal is relatively rare in Nature, with less than 200 preselected events in 1 ab^{-1}, the smallest loss of signal in the optimisation of S (Eq.4) by the MVA method, leads to a significant loss of signal efficiency. The residual number of background events after MVA suppression is found to be: $N \pm \Delta N$ = 39±14, in 1 ab^{-1} of data, where ΔN stands

(a)

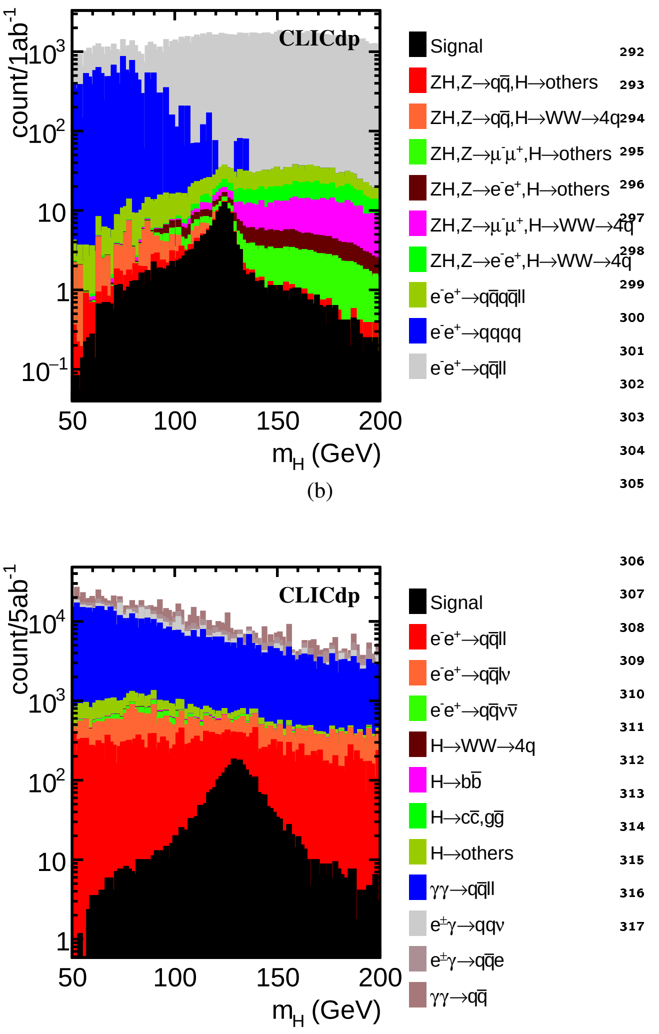

Fig. 5: Stacked histograms of the Higgs mass distributions after preselection phase, at 350 GeV (a) and 3 TeV (b).

for the Poisson variance of the number of selected background events. Processes with the scaling factor[c] larger than 1 ($e^-e^+ \to q\bar{q}l^+l^-$, $e^-e^+ \to HZ; Z \to q\bar{q}, H \to WW \to 4q$) dominantly contribute to the value of ΔN. The uncertainty of the estimated number of signal and background events in 1 ab^{-1} of data leads to the 2% uncertainty of our estimate of δ (δ = (20 ± 2)%). Stacked histograms of the Higgs mass distributions for signal and background after MVA are given in Figure 6a.

[c]The scaling factor is the ratio of number of events expected in 1 ab^{-1} of data w.r.t the size of available simulated sample. It can be read from Table 1a.

5.2 MVA at 3 TeV

At 3 TeV centre-of-mass energy, the MVA is trained with 16 observables: mass of the on-shell Z boson, mass of the off-shell Z boson, invariant mass of two selected leptons, invariant mass of two reconstructed jets, mass of a Higgs candidate, visible energy in an event, difference between the visible energy and the Higgs energy, polar angle of a Higgs candidate, missing transverse momentum per event, number of all PFO objects in an event, jet transition variables ($-\log y_{12}$ and $-\log y_{23}$), b-tag and c-tag probabilities of jets sorted by decreasing transverse momentum of a jet. The Higgs candidate mass is limited to the interval (75 GeV $< m_H <$ 175 GeV). The three most sensitive observables are found to be masses of Higgs and off-shell bosons and polar angle of the reconstructed Higgs boson.

The optimal BDT cut is found to be 0.11, corresponding to the statistical significance of 33. Overall efficiency of signal selection including preselection and MVA selection is found to be about 36%. This corresponds to the MVA signal selection efficiency of approximately 53%. Figure 6b gives stacked plots of Higgs mass distributions for signal and background after MVA selection. BDT background efficiency is on average at the permille level and Table 3 gives the composition of irreducible background. The uncertainty of the number of background events is obtained in the same way as discussed in Sec. 5.1. With 5 ab^{-1} of data, uncertainty of our estimate of δ is 0.1% (δ = (3.0 ± 0.1)%).

Table 3: Preselection and MVA selection efficiences for signal and irreducible background processes and number of selected events N_{BDT}, at 3 TeV centre-of-mass energy, in 5 ab^{-1} of data.

Process	$\varepsilon_{\mathrm{presel}}$	$\varepsilon_{\mathrm{BDT}}$	N_{BDT}
Signal @ 3 TeV	67%	53%	2020
Background processes			
$\gamma\gamma \to q\bar{q}l^+l^-$	11‰	0.3‰	438
$e^-e^+ \to qql\nu$	3‰	4‰	322
$e^-e^+ \to H\nu\bar{\nu}; H \to others$	45‰	1.3%	259
$e^{\pm}\gamma \to qq\nu$	8.8‰	1.3‰	252
processes with N_{BDT}<100	5.3‰	1.1‰	140

(a)

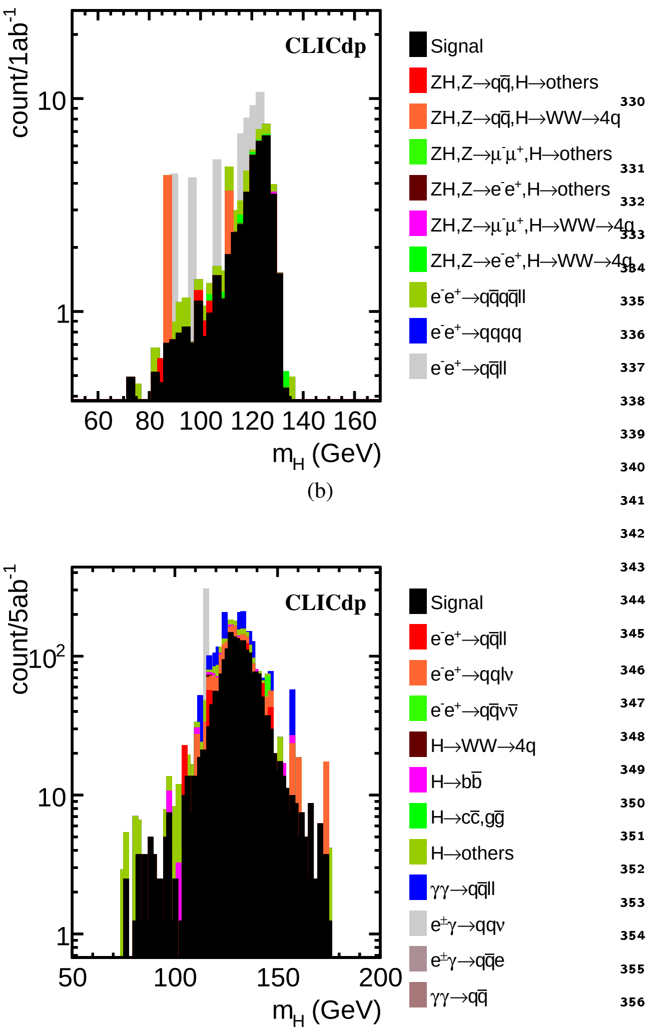

Fig. 6: Stacked histograms of the Higgs mass distributions after MVA, at 350 GeV (a) and 3 TeV (b) centre-of-mass energies.

6 Statistical uncertainties

Relative statistical uncertainty of the $\sigma(H\nu\bar{\nu}) \times BR(H \to ZZ^*)$ measurement is determined as:

$$\delta = \frac{1}{S} \qquad (5)$$

where S is the statistical significance defined in Eq.4. Assuming unpolarized beams, we found statistical uncertainties to be: 20% at 350 GeV and 3.0% at 3 TeV centre-of-mass energies. The high-energy result can be further improved by the beam polarization due to a chiral nature of $WW-$fusion. Assuming the beam polarization scheme discussed in Sec. 3, the statistical uncertainty of the 3 TeV measurement can be conservatively decreased by a factor $\sqrt{1.48}$ [11].

7 Conclusions

Statistical precision of $\sigma(H\nu\bar{\nu}) \times BR(H \to ZZ^*)$ measurement at 350 GeV and 3 TeV centre-of-mass energies CLIC is determined on the basis of a full simulation of physics processes and detector response. Both measurements are simulated for the semi-leptonic final states of the signal. From the obtained statistical significance of signal to background separation, the relative statistical uncertainty of the observable $\sigma(H\nu\bar{\nu}) \times BR(H \to ZZ^*)$ is found to be 20% at 350 GeV and 3.0% at 3 TeV, assuming integrated luminosities of 1 ab^{-1} and 5 ab^{-1}, respectively. The statistical uncertainty at 3 TeV is inline with the expectations from [2] based on luminosity scaling of the precision of 1.4 TeV measurement. Statistical uncertainty of the high-energy result can be further reduced through statistical enhancement of the signal with the proposed beam polarization scheme.

Acknowledgments The work presented in this paper has been carried out in the framework of the CLIC detector and physics study (CLICdp) collaboration and the authors would like to thank CLICdp members for their support, in particular to the colleagues from the Analysis Working Group for useful discussions. We are particularly grateful to Aleksander Filip Żarnecki and Philipp Roloff, for useful ideas exchanged in the course of the analysis. We acknowledge the support received until 2020, from the Ministry of Education, Science and Technological Development of the Republic of Serbia within the national project OI171012.

References

1. R. S. Gupta, H. Rzehak, J. D. Wells, How well do we need to measure Higgs boson couplings, Phys. Rev. D 86 095001 (2012), arXiv:1206.3560
2. H. Abramowicz et al. [CLICdp Collaboration], Higgs physics at the CLIC Electron-Positron Linear Collider, Eur. Phys. J. C 77, 475 (2017), arXiv:1608.07538
3. J. Ellis, P. Roloff, V. Sanz and T. You, Dimension-6 Operator Analysis of the CLIC Sensitivity to New Physics, KCL-PH-TH/2017-04, CERN-PH-TH/2017-009, Cavendish-HEP-17/01, DAMTP-2017-01 (2017), arXiv:1701.04804
4. L. Linssen, A. Miyamoto, M. Stanitzki, H. Weerts (eds.), Physics and Detectors at CLIC: CLIC Conceptual Design Report, ANL-HEP-TR-12-01, CERN-2012-003, DESY 12-008, KEK Report 2011-7 (2012), arXiv:1202.5940

5. T. Abe et al., The International Large Detector: Letter of Intent, DESY-2009-87, FERMILAB-PUB-09-682-E, KEK-REPORT-2009-6 (2010), arXiv:1006.3396
6. N. Alipour Tehrani et al., CLICdet: The post-CDR CLIC detector model, CLICdp-Note-2017-001 (2017), https://cds.cern.ch/record/2254048
7. M. A. Thomson, Particle Flow Calorimetry and the Pandora PFA Algorithm, Nucl. Instrum. Methods A 611, 25 (2009), arXiv:0907.3577
8. P. Roloff, R. Franceschini, U. Schnoor, A.Wulzer (eds.), The Compact Linear e^+e^- Collider (CLIC): Physics Potential, Input to the European Particle Physics Strategy Update on behalf of the CLIC and CLICdp Collaborations (2018), arXiv:1812.07986
9. S. Dittmaier et al., Handbook of LHC Higgs Cross Sections: 2. Differential Distributions, CERN-2012-002 (2012), arXiv:1201.3084
10. P. A. Zylaet et al. [Particle Data Group], Prog. Theor. Exp. Phys. 2020, 083C01 (2020), DOI: 10.1093/ptep/ptaa104
11. A. Robson and P. Roloff, Updated CLIC luminosity staging baseline and Higgs coupling prospects, CLICdp-Note-2018-002 (2018), arXiv:1812.01644
12. W. Kilian, T. Ohl, J. Reuter, WHIZARD: Simulating Multi-Particle Processes at LHC and ILC, Eur. Phys. J. C 71, 1742 (2011), arXiv:0708.4233
13. T. Sjostrand, S. Mrenna, P. Z. Skands, PYTHIA 6.4 Physics and Manual, JHEP 05, 026 (2006), arXiv:hep-ph/0603175
14. D. Schulte, Beam-beam simulations with GUINEA-PIG, (1999), CERN-PS-99-014-LP
15. P. Mora de Freitas, H. Videau, Detector Simulation with Mokka/Geant4: Present and Future, International Workshop on Linear Colliders, JeJu Island, Korea (2002), LC-TOOL-2003-010
16. S. Agostinelli et al., Geant4 - A Simulation Toolkit, Nucl. Instrum. Methods Phys. Res., Sect. A 506, 3 (2003)
17. J. Marshall, A. Münnich, M. Thomson, Performance of Particle Flow Calorimetry at CLIC, Nucl. Instrum. Methods A 700, 153 (2013), arXiv:1209.4039
18. S. Catani et al., Longitudinally-invariant $k_\perp-$clustering algorithms for hadron-hadron collisions, Nucl. Phys. B 406, 187 (1993)
19. G. S. M. Cacciari, G. Soyez, FastJet User Manual, Eur. Phys. J. C 72, 1896 (2012), arXiv:1111.6097
20. O. Wendt, F. Gaede, T. Kramer, Event reconstruction with MarlinReco at the ILC, Pramana 69 (2007), arXiv:physics/0702171
21. T. Suehara, T. Tanabe, LCFIPlus: A Framework for Jet Analysis in Linear Collider Studies, Nucl. Instrum. Meth. A 808, 109-116 (2016)
22. A. Höcker et al., TMVA - Toolkit for multivariate data analysis (2009), arXiv:physics/0703039
23. C. G. et al., ILCDIRAC, a DIRAC extension for the Linear Collider community, CLICdp-Conf 2013-003, CERN, Geneva (2013), https://cds.cern.ch/record/1626585/