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

MEASUREMENT OF THE HIGGS BRANCHING RATIO BR ($H \to \gamma\gamma$) AT 3 TeV CLIC

G. Kačarević[a,1], I. Božović-Jelisavčić[1], N. Vukašinović[1], G. Milutinović-Dumbelović[1], M. Radulović[2], J. Stevanović[2], I. Smiljanić[1], T. Agatonović-Jovin[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, Radoja Domanovića 12, Kragujevac, Serbia

Received: date / Accepted: date

Abstract In this paper we address the potential of a 3 TeV centre-of-mass energy Compact Linear Collider (CLIC) to measure the branching fraction of the Higgs boson decay to two photons, BR($H \to \gamma\gamma$). Since photons are massless, the Higgs boson coupling to photons is realized through higher order processes involving heavy particles either from the Standard Model or beyond. Any deviation of the measured BR($H \to \gamma\gamma$), and consequently of the Higgs coupling $g_{H\gamma\gamma}$ from the predictions of the Standard Model, may indicate New Physics. The Higgs decay to two photons is thus an interesting probe of the Higgs sector.
This study is performed using simulation of the detector for CLIC and by considering all relevant physics and beam-induced processes in a full reconstruction chain. It is shown that the product of the Higgs production cross-section in W^+W^- fusion and BR($H \to \gamma\gamma$) can be measured with a relative statistical uncertainty of 5.5%, assuming the integrated luminosity of 5 ab^{-1} and unpolarized beams.

1 Introduction

The Higgs boson decay to a pair of photons was one of the discovery channels at the LHC [1] and also a benchmark process that has shaped requirements for the electromagnetic calorimetry at ATLAS [2] and CMS [3]. This channel is also important at proposed e^+e^- colliders, both in terms of detector performance requirements and complementarity to the expected HL-LHC results [4].The combined HL-LHC and future e^+e^- collider measurements are expected to give a statistical uncertainty for the Higgs to photons coupling of ~1% [5].
The CLIC provides an excellent environment to study the properties of the Higgs boson, including its couplings, with a very high precision. Operation is expected to be staged at

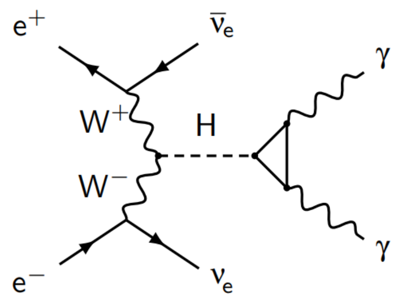

Fig. 1: Lowest order Feynman diagram of the Higgs production in WW-fusion and subsequent Higgs decay to a pair of photons.

three centre-of-mass energies: at 380 GeV, 1.5 TeV and 3 TeV. WW-fusion (Figure 1) as the dominant Higgs production mechanism at center-of-mass energies above ~500 GeV will produce large signal yields allowing rare processes such as $H \to \mu^+\mu^-$, $H \to Z\gamma$ and $H \to \gamma\gamma$ to be studied. For a Higgs mass of 126 GeV, the SM prediction for the branching fraction BR($H \to \gamma\gamma$) is 2.23×10^{-3} [6]. It is expected that 2×10^6 Higgs bosons will be produced at 3 TeV, assuming the nominal integrated luminosity of 5 ab^{-1} which will be used in this paper unless stated otherwise. The signal yield can be increased with the proposed beam polarization by a factor of 1.5 [7]. The high photon-identification efficiency and good photon energy resolution of a detector for CLIC enable excellent identification of $H \to \gamma\gamma$ decays.

This paper presents a comprehensive simulation of the experimental measurement of the Higgs production cross-section in WW-fusion $\sigma(e^+e^- \to H\nu\bar{\nu}) \times BR(H \to \gamma\gamma)$ at 3 TeV CLIC. The result of the study presented in this paper supersede the estimates based on 1.4 TeV studies given in [8].

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

The paper is structured as follows: Simulation and analysis tools are introduced in Section 2, the detector for CLIC is described in Section 3, while Sections 4 to 6 provide details on signal and background identification and separation, pseudo-experiments and uncertainties of the measurement.

2 Simulation and Analysis Tools

The Higgs production in WW-fusion is generated in WHIZARD 1.95 [9], where a Higgs mass of 126 GeV is assumed. Background processes are also generated in WHIZARD, using PYTHIA 6.4 [10] to simulate hadronisation and fragmentation processes. The CLIC luminosity spectrum and beam-induced effects are obtained using the GuineaPig 1.4.4 [11]. Interactions with the detector are simulated using the CLIC_ILD detector model [12] within the Mokka simulation package [13] using the GEANT4 framework [14]. Events are reconstructed using the Particle Flow approach (PFA) implemented in the Pandora algorithm [15]. Photons are reconstructed using PandoraPFA v02-04-00 photon processor [16]. Simulation, reconstruction and analysis are carried out using ILCDIRAC [17]. The TMVA package [18] is used for the multivariate analysis classification (MVA) of signal and background events using their kinematic properties.

3 Detector for CLIC

The CLIC_ILD model is based on the ILD detector proposed for ILC [19] and it has been modified to the CLIC experimental conditions. The vertex detector is closest to the interaction point to provide reconstruction of secondary vertices for accurate flavor tagging. The Time Projection Chamber is foreseen as the main tracking device providing single point resolution better than 100 μm in the plane transverse to the beam axis [12], together with a low material budget. The CLIC_ILD detector uses high-granularity electromagnetic (ECAL) and hadronic (HCAL) sampling calorimeters to reconstruct photons and neutral hadrons. The electromagnetic calorimeter is a Silicon-Tungsten calorimeter optimized for longitudinal containment and latelar separation of electromagnetic showers. High-granularity in combination with the information from the central tracker leads to an electron identification efficiency of 96%, while photon identification efficiency is 99% [20]. The hadronic calorimeter consists of 60 steel absorbers interleaved with scintillator tiles with a purpose to contain hadronic showers from neutral hadrons [12]. A more recent detector model CLICdet [21] improves the stohastic energy resolution term of the ECAL to 17% from 20% of CLIC_ILD. This has no significant impact on the conclusions of this paper [21].

4 Signal and background processes

The main Higgs production processes and backgrounds considered in this paper are summarised in Figure 2 and Table 1. Higgs boson production at 3 TeV is dominated by the WW-fusion process. Without beam polarization, the effective cross-section for the Higgs production is 415 fb, including Initial State Radiation (ISR) effects as well as a realistic CLIC luminosity spectrum. Taking into account that $BR(H \to \gamma\gamma)$ ~0.23%, 4750 signal events are expected with the nominal integrated luminosity. In order to describe fully the CLIC experimental environment, simulated Beamstrahlung photons producing hadrons ($\gamma_{BS}\gamma_{BS} \to$ hadrons) are overlaid on each event after the full simulation of the detector response and before digitization phase. Background processes are considered if two generated photons can be found in the central tracker acceptance with invariant mass of di-photon system between 100 GeV and 150 GeV. Backgrounds arising from mono-photon final states are considered as well if an auxiliary photon (from $\gamma_{BS}\gamma_{BS} \to$ hadrons overlay, final state radiation or false particle identification) can be found in the detector polar angle acceptance, complementing the final state photon to the invariant mass of di-photon system in the required window.

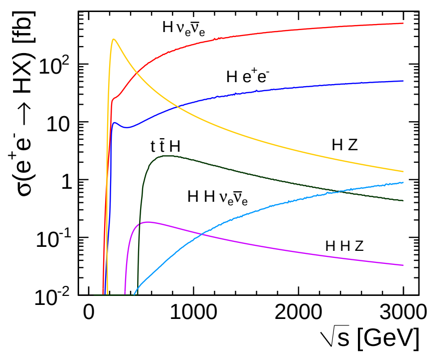

Fig. 2: Higgs production cross-sections at different centre-of-mass energies.

5 Event selection

5.1 Photon isolation and Higgs candidate definition

To ensure that Higgs candidate is found, only events with exactly two isolated photons with transverse momenta greater than 15 GeV are selected. Requirement that both photons have p_T above 15 GeV removes to a great extent photons in a signal event that do not originate from the Higgs decays, as illustrated in Figure 3. We define photon as isolated if the energy of all reconstructed particles within a 14

Signal process	$\sigma(fb)$	N@5 ab^{-1}	$N_{\text{simulated}}$
$e^+e^- \to H\nu\nu, H \to \gamma\gamma$	0.95	4750	24550
Background processes	$\sigma(fb)$		
$e^+e^- \to \gamma\gamma$	15.2	$7.6 \cdot 10^4$	$3 \cdot 10^4$
$e^+e^- \to e^+e^-\gamma$	335	$1.7 \cdot 10^6$	$3 \cdot 10^6$
$e^+e^- \to e^+e^-\gamma\gamma$	33	$1.6 \cdot 10^5$	$1.5 \cdot 10^5$
$e^+e^- \to \nu\bar{\nu}\gamma$	13	$6.6 \cdot 10^4$	$2 \cdot 10^5$
$e^+e^- \to \nu\bar{\nu}\gamma\gamma$	26	$1.3 \cdot 10^5$	$1.6 \cdot 10^5$
$e^+e^- \to q\bar{q}\gamma$	210	$1.1 \cdot 10^6$	$1.2 \cdot 10^6$
$e^+e^- \to q\bar{q}\gamma\gamma$	47	$2.3 \cdot 10^5$	$3 \cdot 10^5$

Table 1: Considered signal and background processes with the corresponding effective[b] cross-sections at 3 TeV centre-of-mass energy.

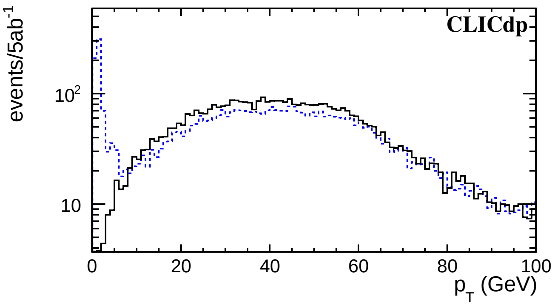

Fig. 3: The 2nd highest reconstructed photons p_T in a signal event (dashed) and the 2nd highest p_T photon generated in a Higgs decay (solid). The difference in the two distributions at low p_T values comes from the presence of Beamstrahlung photons at the reconstructed level (dashed).

mrad cone is less than 20 GeV. This isolation criteria reduces background processes (in particular $e^+e^- \to q\bar{q}\gamma$ and $e^+e^- \to q\bar{q}\gamma\gamma$) by 23%. Signal loss is negligible. Selection of events with exactly two isolated photons with $p_T > 15$ GeV results in 22.3% signal loss, as illustrated in Figure 4.

5.2 Preselection

Signal is separated from backgrounds in a two-stage selection process: preselection and MVA based selection. The preselection aims to suppresses high cross-section backgrounds like $e^+e^- \to e^+e^-\gamma$ and $e^+e^- \to e^+e^-\gamma\gamma$. Preselection variables are optimized as follows:

- Reconstructed di-photon invariant mass in the range from 110 GeV to 140 GeV, corresponding to the Higgs mass window,
- Reconstructed di-photon energy in the range between 100 GeV and 1000 GeV,

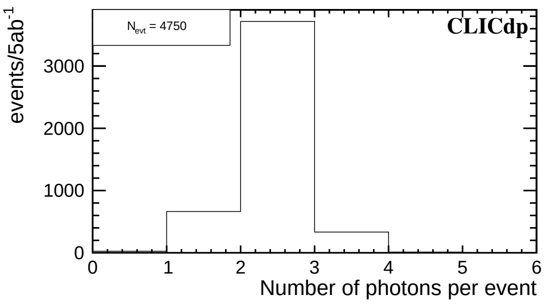

Fig. 4: Number of reconstructed isolated photons per signal event with $p_T(\gamma) > 15$ GeV.

- Reconstructed di-photon transverse momentum in the range between 20 GeV and 600 GeV.

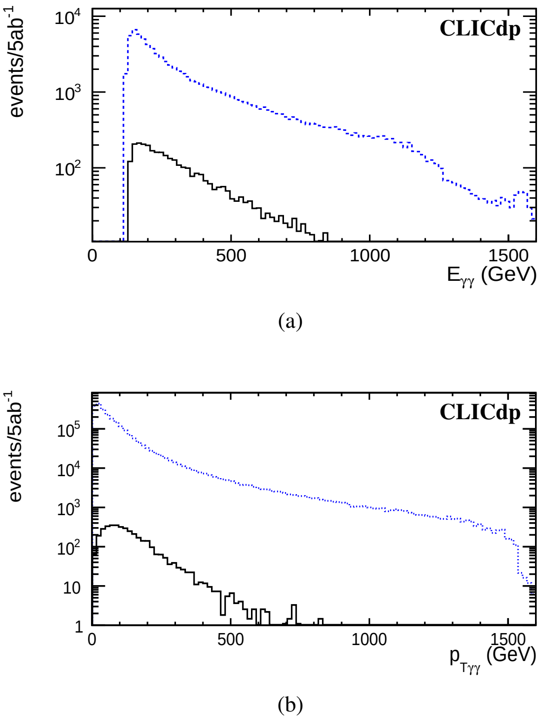

Fig. 5: Higgs candidate observables for signal and background: energy (a) and transverse momentum (b). Signal is represented with the solid line while background is represented as dashed.

Distributions of di-photon energy and transverse momentum are given in Figure 5a and Figure 5b respectively, illustrating the selection range. The signal and background di-photon invariant mass after preselection is given in Figure 6. Preselection efficiency for signal is 70% and background dominates over the signal by a factor of 25.

[b] The cross-sections are effective in a sense that condition 100 GeV < $m_{\gamma\gamma}$ < 150 GeV is applied to any di-photon system found in the central tracker.

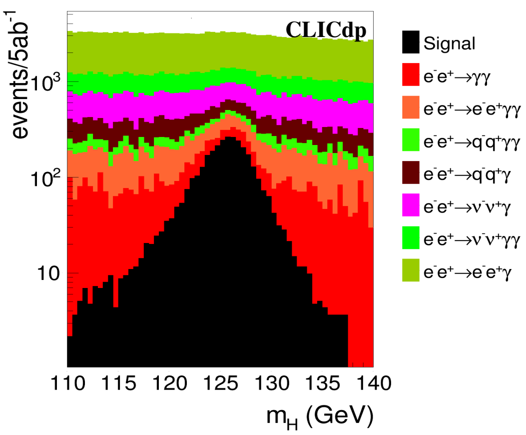

Fig. 6: Stacked histograms of Higgs mass distributions for signal and background after preselection.

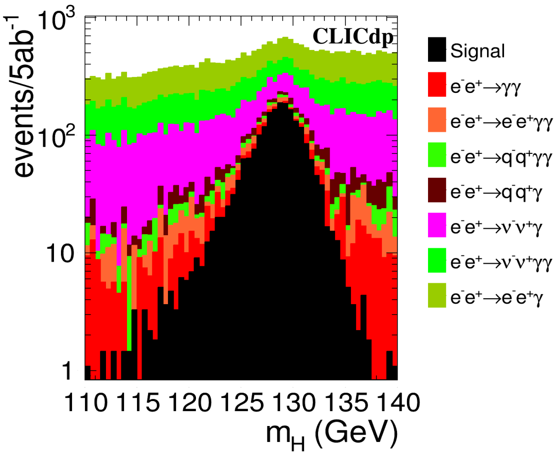

Fig. 8: Stacked histograms of Higgs mass distributions for signal and background after MVA selection .

5.3 Multivariate analysis

Preselected signal and background events are further separated using an MVA method based on the Gradient Boosted Decision Trees (BDTG). Twelve observables are used for classification of events: di-photon energy, di-photon transverse momentum, di-photon polar angle, cosine of the helicity angle, transverse momenta of photons, polar angle of photons, energy of photons, total ECAL energy per event and total HCAL energy per event. The optimal cut-off value of the BDTG output variable was found to be 0.34, what is illustrated in Figure 7. Variables are sufficiently uncorrelated for MVA to perform optimally.

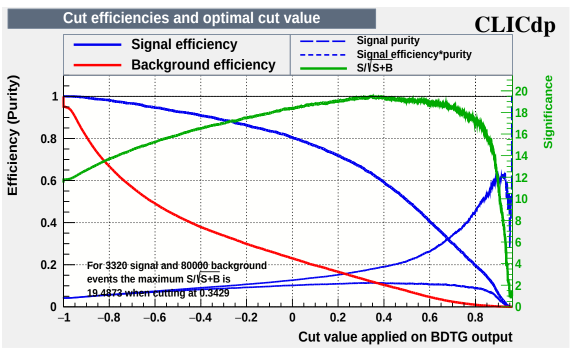

Fig. 7: BDTG performance in the training phase.

The classifier cut was selected to maximize statistical significance defined as:

$$S = \frac{N_s}{\sqrt{N_s + N_b}} \quad (1)$$

where N_s and N_b are number of signal and background events after the MVA selection. The MVA efficiency for signal is 62.7%, resulting in an overall signal selection efficiency of 43.7%, corresponding to a signal yield of 2080 selected Higgs candidates. The remaining background after the MVA application is ~10 times larger than the signal and originates mostly from the processes such as $e^+e^- \to \nu\bar{\nu}\gamma$ and $e^+e^- \to \nu\bar{\nu}\gamma\gamma$ or from a high cross-section process like $e^+e^- \to e^+e^-\gamma$. The Higgs candidate mass distribution after MVA selection is illustrated in Figure 8, giving the composition of the background.

6 Pseudo experiments

6.1 Method

The observable to be measured is a product of the Higgs production cross-section and a corresponding branching fraction for Higgs di-photon decay and it can be experimentally determined from the counted number of signal events N_s as:

$$\sigma(e^+e^- \to H\nu\bar{\nu}) \times BR(H \to \gamma\gamma) = \frac{N_s}{L \cdot \varepsilon_s} \quad (2)$$

where L represents the integrated luminosity, ε_s is the overall signal efficiency including detector acceptance, photon identification efficiency and signal selection efficiency. The number of signal events will be determined from combined fit of di-photon invariant mass distributions of selected simulated (or experimental) data with the function f:

$$f(m_{\gamma\gamma}) = N_s \cdot f_s(m_{\gamma\gamma}) + N_b \cdot f_b(m_{\gamma\gamma}) \quad (3)$$

where N_s and N_b are number of selected signal and background events, and f_s and f_b are the probability density functions (PDF) describing $m_{\gamma\gamma}$ for signal and background respectively. These PDFs are determined from simulated samples of signal and background data.

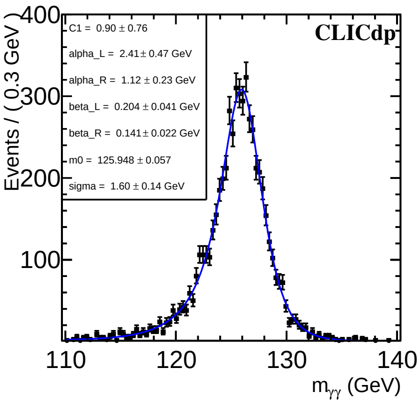

Fig. 9: Fit of di-photon invariant mass of the selected signal (points) and the fit function f_s (line).

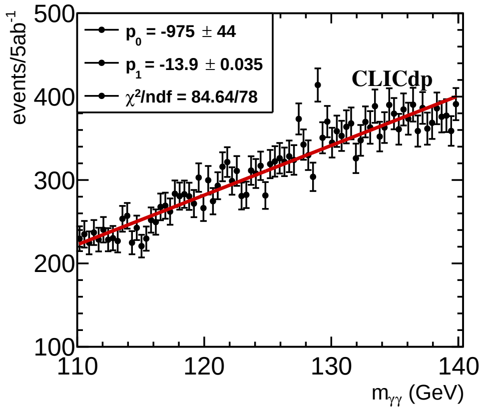

Fig. 10: Di-photon invariant mass $m_{\gamma\gamma}$ for the sum of all background processes remaining after event selection (points). The fit function given in Equation 5 is overlaid.

6.2 Signal and background PDF

Functions f_s and f_b from Eq. 3 are used to fit the fully simulated datasets of signal and background after the the MVA selection phase. The signal PDF consists of two Gaussian functions, one describing the tail (f_{flat}) and the other describing exponential part (f_{exp}) of di-photon mass distribution of the signal:

$$f_s = f_{flat} + C_1 \cdot f_{exp} \tag{4}$$

$$f_{flat} = \begin{cases} e^{-\frac{(m_{\gamma\gamma}-m_H)^2}{2\sigma^2+\beta_L(m_{\gamma\gamma}-m_H)^2}} & ,(m_{\gamma\gamma} < m_H) \\ e^{-\frac{(m_{\gamma\gamma}-m_H)^2}{2\sigma^2+\beta_R(m_{\gamma\gamma}-m_H)^2}} & ,(m_{\gamma\gamma} > m_H) \end{cases}$$

$$f_{exp} = \begin{cases} e^{-\frac{(m_{\gamma\gamma}-m_H)^2}{2\sigma^2+\alpha_L|m_{\gamma\gamma}-m_H|}} & ,(m_{\gamma\gamma} < m_H) \\ e^{-\frac{(m_{\gamma\gamma}-m_H)^2}{2\sigma^2+\alpha_R|m_{\gamma\gamma}-m_H|}} & ,(m_{\gamma\gamma} > m_H), \end{cases}$$

where $\sigma, C_1, \alpha_{L,R}, \beta_{L,R}$, as well as Higgs mass m_H are free parameters determined by the fit (Figure 9). The fit is performed using RooFit [22].

The di-photon mass distribution for background is fitted with a linear function f_b:

$$f_b = p_0 + p_1 \cdot m_{\gamma\gamma} \tag{5}$$

where p_0 and p_1 stand for free parameters of the fit. Fit of background di-photon invariant mass distribution is illustrated in Figure 10, and shows no sensitivity to the SM Higgs mass.

6.3 Pseudo-experiments

Pseudo-data distribution, combining both signal and background after MVA selection, is fitted with function f (Eq. 3), where N_s and N_b are set as free parameters. In this way the number of signal events is determined in the same way it would be on a set of experimental data. Such a measurement we call a pseudo-experiment. An example of one pseudo-experiment is shown in Figure 11. In order to estimate the statistical dissipation of the measured number of signal events, 5000 pseudo-experiments with 5 ab^{-1} of data were performed. Pseudo-data for signal is randomly picked from fully simulated signal sample, while $m_{\gamma\gamma}$ distribution for background is generated from background PDF by randomly changing parameters p_0 and p_1 from Eq.5. The RMS of the resulting pull distribution over all pseudo-experiments is taken as the estimate of the statistical uncertainty of the measurement (Figure 12). It reads that the statistical uncertainty of the extracted number of signal events is 5.5%.

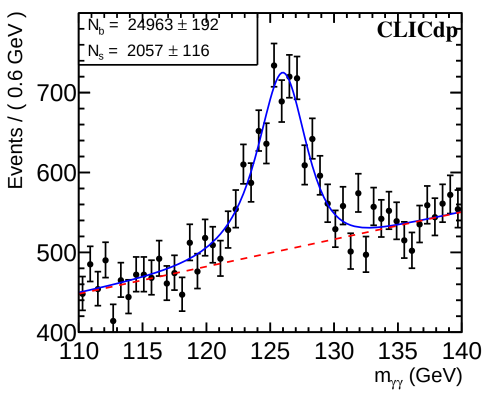

Fig. 11: Example of one pseudo-experiment, showing di-photon invariant mass of pseudo-data (black), corresponding fit with the function f from Eq. 3 (full line) and background fit with function f_b (dashed line) from Eq. 5.

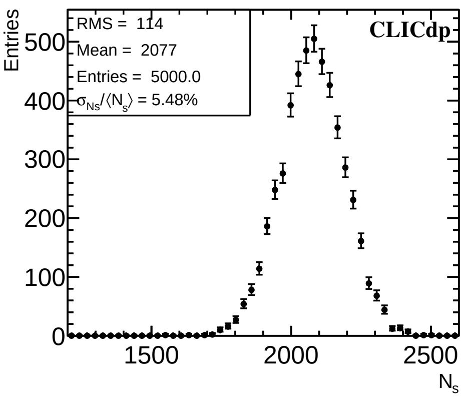

Fig. 12: Pull distribution of 5000 pseudo-experiments.

6.4 Systematic uncertainty

Several sources of systematic uncertainty of the measured observable are considered. Systematic uncertainty rising from the uncertainty of a single photon identification efficiency of 1% results in ~2% systematic effect in $BR(H \to \gamma\gamma)$ measurement. The relative uncertainty of the integrated luminosity, and hence of the measured cross-section, is expected to be of order of several permille at CLIC [23]. Another source of systematic uncertainty is due to uncertainty of the luminosity spectrum recontruction. In [24] it has been shown that the CLIC luminosity spectrum at 3 TeV centre-of-mass energy can be corrected better than 5% above 50% of the nominal centre-of-mass energy, while above 75% of the nominal centre-of-mass energy the corresponding uncertainty of the correction is at a permille level [25]. As discussed in [8], the impact of uncertainty of the luminosity spectrum reconstruction on $H\nu\bar{\nu}$ production at 3 TeV ($H \to b\bar{b}$) is found to be of order of several permille. The energy resolution of the ECAL also has the permille-level impact on preselection efficiency. The relative uncertainty of the ECAL sampling term of 2% leads to the uncertainty of reconstructed photon energy of ~ 40 MeV which has negligible effect on N_s determination. Similarly, the uncertainty of di-photon transverse momentum as a preselection variable hardly contributes to the systematic uncertainty of the measurement. With all considerations above, total systematic uncertainty is estimated to be ~2.4%, which is approximately two times smaller than the statistical one.

7 Summary

The accessibility of WW-fusion as a dominant Higgs production mechanism at energies of 500 GeV and above enable the Higgs rare decays at 3 TeV CLIC to be measured. Excellent performance of the electromagnetic calorimeter to identify high-energy photons together with the overall PFA reconstruction of physics processes enables the measurement of the loop induced Higgs decays to two photons at the percent level. In the full simulation of experimental measurement, we have shown that $\sigma(e^+e^- \to H\nu\bar{\nu}) \times BR(H \to \gamma\gamma)$ can be measured at 3 TeV CLIC with a relative statistical uncertainty of 5.5%, assuming 5 ab^{-1} of integrated luminosity and unpolarized beams. This result can be further improved with the proposed beam polarization scheme, which would increase the Higgs production cross-section by a factor of 1.5, due to a chiral nature of WW-fusion as a charged-current interaction. The estimated systematic uncertainty of ~2.4% is smaller than the statistical one. This analysis completes the set of Higgs to $\gamma\gamma$ measurements foreseen at CLIC energy stages above 1 TeV centre-of-mass energy.

8 Acknowledgements

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 the CLICdp members for their support. We are in particular grateful to the colleagues from the Analysis Working Group for useful discussions, to Philipp Roloff, for useful ideas exchanged in the course of the analysis and to Aleksander Filip Zarnecki who contributed to the paper by raising several important questions. We acknowledge the support received until 2020 from the Ministry of Education, Science and Technological Development of the Republic of Serbia within the project OI171012.

References

1. ATLAS Collaboration, Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC, Physics Letters B, Volume 716, Issue 1, 17 September 2012, Pages 1-29.
2. G. Unal and on behalf of the ATLAS Collaboration, The Role of the LAr Calorimeter in the Search for $H \to \gamma\gamma$ in ATLAS, 2012 J. Phys.: Conf. Ser. 404, https://iopscience.iop.org/article/10.1088/1742-6596/404/1/012001/pdf.
3. F. Cavallari, Performance of calorimeters at the LHC, 2011, J. Phys.: Conf. Ser. 293 012001.
4. M. Cepeda, S. Gori, P. Ilten, M. Kado, F. Riva, et al. Higgs Physics at the HL-LHC and HE-LHC, 2019, https://arxiv.org/abs/1902.00134.
5. Aidan Robson, Philipp Roloff, Jorge de Blas, CLIC Higgs coupling prospects with a longer first energy stage, CLICdp-Note-2020-001.

6. S. Dittmaier et al., Handbook of LHC Higgs Cross Sections: 2. Differential Distributions, arxiv.org/abs/1201.3084 (2012).
7. H. Abramowicz, N. Alipour Tehrani, D. Arominski et al., Top-quark physics at the CLIC electron-positron linear collider, JHEP 11 (2019) 003.
8. H. Abramowicz, A. Abusleme, K. Afanaciev et al. Higgs physics at the CLIC electron-positron linear collider, Eur. Phys. J. C 77, 475 (2017), arxiv.org/abs/1608.07538.
9. W. Kilian, T. Ohl, J. Reuter, WHIZARD: Simulating Multi-Particle Processes at LHC and ILC, Eur. Phys. J. C71, 1742, (2011), arXiv:0708.4233.
10. T. Sjostrand, S. Mrenna, P. Z. Skands, PYTHIA 6.4 Physics and Manual, JHEP 05, 026 (2006), arXiv:hep-ph/0603175.
11. D. Schulte, Beam-beam simulations with GUINEA-PIG, CERN-PS-99-014-LP (1999).
12. L. Linssen et al., 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.
13. P. Mora de Freitas, H. Videau, Detector Simulation with Mokka/Geant4 : Present and Future, International Workshop on Linear Colliders (2002), LC-TOOL-2003-010, JeJu Island, Korea.
14. S. Agostinelli et al., Geant4 - A Simulation Toolkit, Nucl. Instrum. Methods Phys. Res., Sect. A 506 3 (2003), info-science.epfl.ch/record/49909/files/Ago+03.pdf.
15. J.S. Marshall, A. Münnich, M. Thomson, Performance of Particle Flow Calorimetry at CLIC, Nucl. Instrum. Methods A700, 153 (2013), arXiv:1209.4039.
16. F. Gaede, Marlin and LCCD: Software tools for the ILC, 2006, Nucl.Instrum.Meth.A 559 (2006) 177-180, https://doi.org/10.1016/j.nima.2005.11.138.
17. C. Grefe, S. Poss, A. Sailer, ILCDIRAC, a DIRAC extension for the Linear Collider community, CLICdp-Conf 2013-003, CERN, Geneva (2013), cds.cern.ch/record/1626585/files/ChepProceedings.pdf.
18. A. Hocker et al., TMVA - Toolkit for multivariate data analysis (2009), arXiv:physics/0703039.
19. T. Abe et al., The International Large Detector: Letter of Intent (2010), arXiv:1006.3396.
20. D. Arominski et al., A detector for CLIC: main parameters and performance, 2018, https://arxiv.org/abs/1812.07337v1.
21. CLICdp Collaboration, CLICdet: The post-CDR CLIC detector model, 2017, https://cds.cern.ch/record/2254048. Kalinowski, J., Kotlarski, W., Sopicki, P. et al., Simulating hard photon production with WHIZARD, Eur. Phys. J. C 80, 634 (2020), https://arxiv.org/abs/2004.14486.
22. W. Verkerke, D. P. Kirkby, The RooFit toolkit for data modeling, arXiv:physics/0306116, 2003.
23. I. Bozovic-Jelisavcic, S. Lukic, G. Milutinovic Dumbelovic, M. Pandurovic, I. Smiljanic, Luminosity measurement at ILC, JINST 8 (2013) P08012 arXiv:1304.4082.
24. S. Poss, A. Sailer, Luminosity spectrum reconstruction at linear colliders, Eur. Phys. J. C 74, 2833 (2014), arxiv.org/abs/1309.0372.
25. S. Lukic, I. Bozovic-Jelisavcic, M. Pandurovic, I. Smiljanic, Correction of beam-beam effects in luminosity measurement in the forward region at CLIC, JINST 8 (2013) P05008, arXiv:1301.1449.