CERN - European Organization for Nuclear Research

LCD-Note-2012-019

Measurement of $\sigma(\mathrm{e^+e^-} \rightarrow \mathrm{H}\nu\overline{\nu})\times\mathbf{BR}(\mathrm{H} \rightarrow \tau\tau)$ at CLIC @ 350 GeV

A. Münnich*

* *CERN, CH-1211 Geneva 23, Switzerland*

April 29, 2013

Abstract

This detector benchmark study evaluates the statistical precision with which the $\mathrm{H} \rightarrow \tau\tau$ branching ratio times cross section can be measured at CLIC running at $\sqrt{s} = 350$ GeV. Only the hadronic decay of τs are considered. Results for $M_{\mathrm{H}} = 126$ GeV and 500 fb^{-1} of integrated luminosity are obtained using full detector simulation and including beam-induced backgrounds resulting in a statistical accuracy of cross section times branching ratio of 6.2%.

1 Introduction

The study reported in this note is carried out for a centre-of-mass energy of 350 GeV. The dominant production process for a Standard Model Higgs at this energy is $e^+e^- \to HZ$ as shown in Figure 1. The goal is to measure the cross section for the Higgs decay into τs. This requires the reconstruction of τ leptons at high energies and in the presence of machine-induced backgrounds. The investigated final state has two τ leptons and two quark jets from the Z decay. The corresponding branching ratios and cross section are listed in Table 1. Only hadronic decays of the τ are considered.

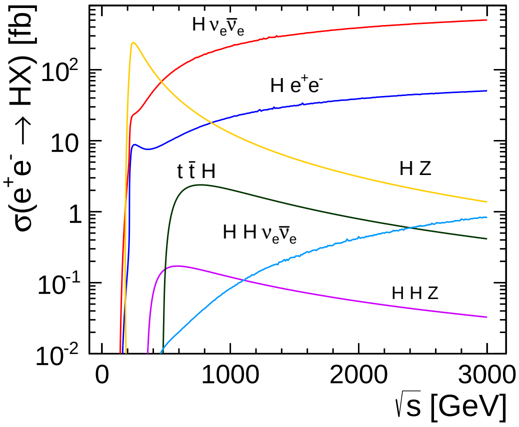

Figure 1: Cross sections for different Higgs production channels as a function of $\sqrt{s}$ for $M_H = 125$ GeV.

Table 1: The Higgs mass, final state, branching ratios and cross section used in this study.

Process:	$e^+e^- \to HZ$
Particle mass:	$m_H = 126$ GeV
Final state:	$\tau\tau qq$
Branching ratios:	$H \to \tau\tau$ (6.15%)
	$\tau \to$ hadrons (64.8%), $\tau \to e\nu\bar{\nu}$ (17.8%), $\tau \to \mu\nu\bar{\nu}$ (17.4%)
Cross section:	$\sigma = 93.5$ fb $\times$ $0.06 = 5.75$ fb

2 Monte Carlo production

The physics events used for the study presented here were produced with the same machinery used for the CLIC CDR [1]. Events were generated using the `WHIZARD 1.95` [2] program. Initial and final state radiation (ISR and FSR) were enabled during the event generation. The luminosity spectrum expected at CLIC was used during the event generation [3]. The hadronization of final state partons was simulated using `Pythia` [4]. The generated events were subsequently passed through the detector simulation program Mokka [5] which is based on the Geant4 [6] package. The CLIC_ILD [7] detector geometry model was used.

Events were overlayed with pileup from $\gamma\gamma \to$ hadrons interactions corresponding to 300 bunch crossings with 0.0464 events per bunch crossing [8]. The reconstruction chain included an improved version [9] of the PandoraPFA [10] algorithm to reconstruct particle flow objects.

An overview of all produced Monte Carlo (MC) samples is given in Table 2.

Table 2: Cross sections and integrated luminosities of the available Monte Carlo samples for the $H \to \tau\tau$ study and the relevant backgrounds. γe represents both γe^+ and γe^- production.

Process	Cross section [fb]	Luminosity [ab^{-1}]
$ee \to HZ \quad (H \to \tau\tau, Z \to qq)$	5.8	6.6
$ee \to HZ \quad (H \to X, Z \to \tau\tau)$	4.6	2.4
$ee \to qq\tau\tau$ (no Higgs)	70.0	1.2
$ee \to qq\tau\tau\nu\nu$	1.6	5.0
$ee \to qqqq$	5900	0.1
$\gamma\gamma \to qq\tau\tau$	4.5	0.6
$\gamma\gamma \to qqqq$	84.0	0.6
$\gamma e \to qq\tau\tau e$	1.1	2.7
$\gamma e \to qqqqe$	52.6	0.5

3 Event reconstruction

The steps to reconstruct events with two τs and two qs from particle flow objects (PFOs) are described in this section. The presence of pileup from the process $\gamma\gamma \to$ hadrons increases the number of reconstructed PFOs which are emitted mostly in the forward direction.

A large fraction of the background can be rejected using combined timing and momentum cuts [1]. At 350 GeV the backgrounds are less severe and the specific choice of the selection cuts does not influence the shape of the reconstructed τ energy distribution nor the τ reconstruction efficiency. Purely for consistency reasons with the corresponding analysis done at 1.4 TeV [11] the “tight selected PFOs” were chosen for this study as well.

Table 3: Efficiency for the different steps of the analysis.

Step	Efficiency or Branching Ratio [%]
τ reconstruction	Eff(τ) = 0.7*0.7
hadronic τ decay	BR($\tau \to$ hadrons) = 0.64*0.64
pre-selection cut	Eff(PreSel) = 0.9

The reconstruction of τ leptons is done with the `TauFinder` [12] which is essentially a seeded cone based jet clustering algorithm. Via steering parameters like the opening angle of the search and isolation cones the reconstruction algorithm can be optimized for a given τ signature. For this study a scan of the algorithm's parameters was carried out on events with four quarks to evaluate the fake rate, meaning mistaking a quark jet for a τ jet.

The following parameter set was chosen for the reconstruction of τ leptons:

- Minimum p_T for τ seed: 5 GeV
- Maximum for invariant mass of τ candidate: 2.5 GeV
- Opening angle of search cone: 0.1 rad
- Opening angle of isolation cone (relative to search cone): 0.3 rad
- Maximum energy allowed in isolation cone: 2.0 GeV

With these settings the efficiency to reconstruct a τ in signal events is 73%. The fake rate of 4.7% to mistake a quark for a τ is rather high but acceptable in this analysis as the background from quarks can be reduced significantly in the final event selection.

After reconstructing all τ candidates in the event the remaining particles are given to the jet reconstruction. The `kt` algorithm from the FastJet library [13] is used with an R value of 0.7 in exclusive mode forcing the formation of 2 jets.

4 Event selection

The selection of $H \to \tau\tau$ signal events is performed in two steps. First, a cut-based pre-selection is applied to select the hadronic 1 and 3 prong τ decays by requiring that no lepton is part of the reconstructed τ and that it has either 1 or 3 tracks. To distinguish between signal and background events further, the Toolkit for Multivariate Analysis (TMVA) [14] is used. Boosted decision trees (BDT) proved to be the most efficient classifiers for this analysis. For training purposes, 30% of the available events for each process are used. These events are not considered in the analysis to measure the cross section.

Event classification

The boosted decision trees base their selection on 17 variables describing the event topology and describing kinematic quantities of the reconstructed τ candidates as well as the quark system:

- Missing transverse momentum $p_{\mathsf{T,miss}}$
- Thrust of the full event
- Thrust and oblateness of the τ system and the quark system
- Sum of the transverse momenta of both τ candidates and quark jets
- $\cos\theta_{\tau,1}$ and $\cos\theta_{\tau,2}$
- Invariant mass of the τ system and the quark system
- Angle between the two τ candidates
- Angle between the two quark jets
- θ^{miss}, where θ^{miss} is the polar angle of the missing momentum
- $\Delta\phi$ between the two τs
- $\Delta\phi$ between the two quarks
- Visible energy in the event

As examples for the input variables, the thrust of the event (left) and the angle between the two τ candidates in the event for the signal and all backgrounds are shown in Figure 2.

Using the input variables described above, the classifier response for each event is computed which is referred to as BDT in the following. Figure 3 shows the distribution of the BDT for the signal and the different backgrounds. Signal events tend to be at higher BDT values than the backgrounds. Additionally, the signal efficiency, purity and significance for events passing the pre-selection as a function of the BDT cut value are shown. The highest significance is reached for a BDT of 0.08 which is then chosen for the analysis giving an efficiency of 59% and a purity of 67% for this selection. Combined with the efficiency from the reconstruction and the pre-selection this amounts to a signal efficiency of 9% of the produced data. The two main background contributions are from the four quark events due to the huge cross section of this process and the process ee $\rightarrow$ qq$\tau\tau$ which has a very similiar signature as the signal events. After the event selection 76% of the remaining background comes from the process ee $\rightarrow$ qq$\tau\tau$.

The purely statistical uncertainty on the cross section can be calculated as $\sqrt{S+B}/S$. Figure 4 shows the dependence of this statistical uncertainty of the cross section on the selection efficiency and the BDT value. The result is summarized in Table 4.

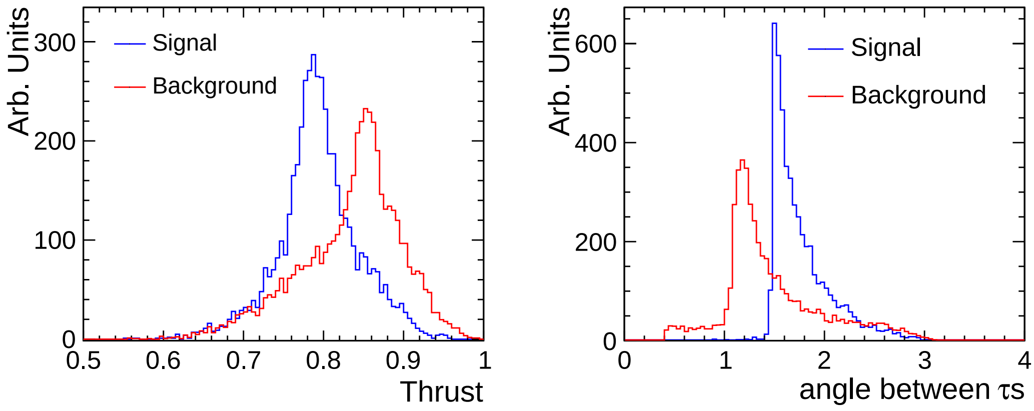

Figure 2: Thrust of the event (left) and the angle between the two τ candidates (right) for the signal and all backgrounds. The distributions are normalization to illustrate the different shapes.

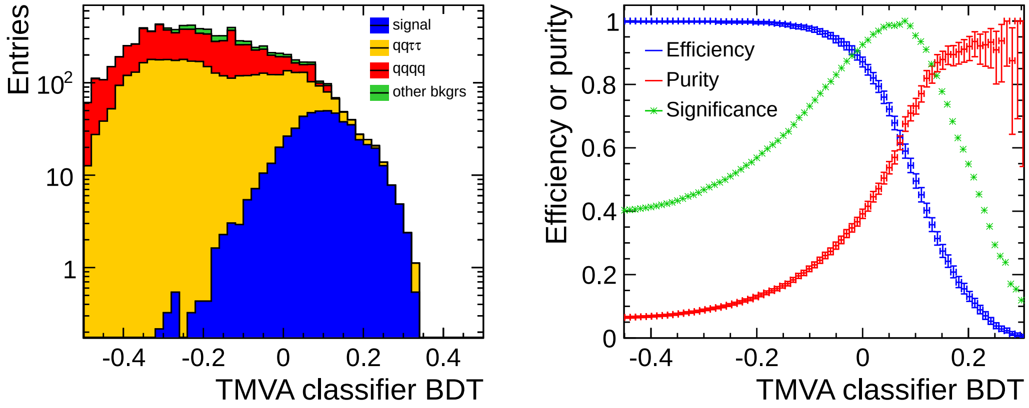

Figure 3: Distribution of the BDT values for the signal and the backgrounds for 0.5 ab^{-1} (left) and the selection efficiency, purity and significance in dependence on the chosen BDT cut value (right).

Since this analysis is merely a counting experiment it is not necessary to limit the calculation of the statistical error by actually choosing a BDT cut. Each BDT bin with signal can be used to

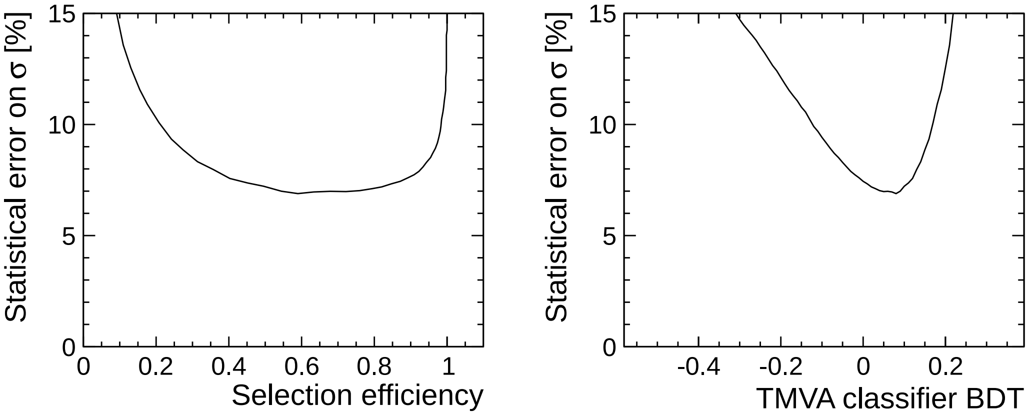

Figure 4: Statistical uncertainty of the cross section in dependence on the selection efficiency (left) and the chosen BDT cut value (right).

determine a value for $\Delta\sigma/\sigma$. These individual errors can be summed quadratically:

$$x_i = \frac{\sqrt{S_i + B_i}}{S_i}$$

$$\frac{\Delta\sigma}{\sigma} = \sqrt{\frac{1}{\sum_i \frac{1}{x_i^2}}}$$

This methods improves the statistical uncertainty for the cross section slightly from 6.9% to 6.2%. It is very stable and does not depend on the number of bins choosen. Reducing the number of bins from 200 to 25 the obtained result changes from 6.16% to 6.22%. The result is summarized in Table 4.

Table 4: Event selection performance and measured statistical uncertainty of the cross section for H → ττ decays.

Total signal events	529
Signal events (above BDT cut)	312
Background events (above BDT cut)	150
Signal efficiency	59%
Signal purity	67%
Statistical uncertainty cross section × branching ratio	6.2%

5 Conclusion and Summary

The cross section for $\mathrm{H} \rightarrow \tau\tau$ decays was measured for the producion via $\mathrm{ee} \rightarrow \mathrm{HZ}$ reconstructing hadronic decays of the two τs. The study was performed using full simulation and considering pileup from $\gamma\gamma \rightarrow$ hadrons. A center-of-mass energy of 350 GeV and an integrated luminosity of 500 fb^{-1} was used. For $M_{\mathrm{H}} = 126$ GeV the statistical uncertainty of the cross section times branching ratio is 6.2%.

References

[1] L. Linssen et al., *Physics and Detectors at CLIC: CLIC Conceptual Design Report*, CERN-2012-03

[2] W. Kilian, T. Ohl and J. Reuter, *WHIZARD: Simulating Multi-Particle Processes at LHC and ILC*, arXiv:0708.4233 (2007); M. Moretti et al., *O'Mega: An Optimizing Matrix Element Generator*, LC-TOOL-2001-040, also arXiv:hep-ph/0102195 (2001).

[3] B. Dalena, J. Esberg and D. Schulte, *Beam-induced backgrounds in the CLIC 3 TeV CM energy interaction region*, Proceedings of LCWS11, arXiv:1202.0563 (2012).

[4] T. Sjöstrand, S. Mrenna and P. Skands, *PYTHIA 6.4 physics and manual* , JHEP 05 (2006) 026.

[5] Simulator for the Linear Collider (MOKKA), `http://www.lcsim.org/software/slic/`.

[6] S. Agostinelli et al., *GEANT4 - a simulation toolkit*, Nucl. Instr. Meth. A 506 (2003) 250; J. Allison et al., *Geant4 Developments and Applications*, IEEE Trans. Nucl. Sci. 53 (2006) 270.

[7] A. Münnich and A. Sailer, *The CLIC ILD CDR Geometry for the CDR Monte Carlo Mass Production*, LCD-Note-2011-002 (2011).

[8] P. Schade, A. Lucaci-Timoce, *Description of the signal and background event mixing as implemented in the Marlin processor OverlayTiming*, LCD-Note-2011-006 (2011).

[9] J. Marshall and M.A. Thomson, *Redesign of the Pandora Particle Flow algorithm*, Report at IWLC2010 (2010).

[10] M.A. Thomson, *Particle Flow Calorimetry and the PandoraPFA Algorithm*, Nucl. Inst. Meth. A 611 (2009) 25.

[11] A. Münnich, *Measurement of* $\sigma(\mathrm{e}^+\mathrm{e}^- \rightarrow \mathrm{H}\nu\bar{\nu}) \times BR(\mathrm{H} \rightarrow \tau\tau)$ *at CLIC @ 1.4 TeV*, LCD-Note-2012-010 (2012).

[12] A. Münnich, *`TauFinder`: A Reconstruction Algorithm for τ Leptons at Linear Colliders*, , LCD-Note-2010-009 (2010).

[13] M. Cacciari, G.P. Salam, *Dispelling the N3 myth for the kt jet-finder* Phys. Lett.B 641 (2006) 57

[14] A. Hoecker et al., *TMVA 4 - Toolkit for Multivariate Data Analysis with ROOT*, arXiv:physics/0703039, also CERN-OPEN-2007-007 (2009).