

A combined fit to the Higgs branching ratios at the International Large Detector with backgrounds

Jonas Kunath^{a,†,*}

^a*Laboratoire Leprince-Ringuet, IN2P3-CNRS, École Polytechnique, Institut Polytechnique de Paris
Route de Saclay, 91120 Palaiseau, France*

E-mail: kunath@llr.in2p3.fr

The Higgsstrahlung process for the production of Higgs bosons at a lepton collider provides access to a high-purity Higgs boson sample. This sample can be constructed in such a way that it is impartial to the type of the Higgs decay. By allocating the events of the sample to classes defined within the sample, we create a binned distribution. These class counts are associated to the Higgs branching ratios through a model-independent fit. The fit provides an estimate for each of the Higgs branching ratios with the full matrix of statistical correlations between the channels. By design, these are pure branching ratio measurement, independent of any Higgs production mode cross section measurement.

We present a study on data simulated for the ILD concept at the International Linear Collider (ILC) at 250 GeV center-of-mass energy.

*** *The European Physical Society Conference on High Energy Physics (EPS-HEP2021)*, ***

*** *26-30 July 2021* ***

*** *Online conference, jointly organized by Universität Hamburg and the research center DESY* ***

*Speaker

†On behalf of the ILD concept group.

1. Introduction

The International Linear Collider (ILC) will initially collide polarized beams of electrons and positrons at center-of-mass energy $\sqrt{s} = 250$ GeV (ILC250). This energy will then be increased up to $\sqrt{s} = 500$ GeV [1]. The ILD concept group proposes the International Linear Detector [2, 3] as a detector for the e^+e^- interaction point. Its silicon trackers allow to measure impact parameters to less than 5 μm . With the help of a Time Projection Chamber, transverse momenta of charged particles can be measured with a resolution of $\sigma(1/p_T) = 2 \cdot 10^{-5} \text{GeV}^{-1}$. The highly granular electromagnetic and hadronic calorimeters are based on the concept of particle flow for particle reconstruction [4]. A jet energy resolution of 3-4% can be achieved. The surrounding coil provides a magnetic field of 3 T.

Here we focus on the ILC250. It will produce a large number of Higgs boson events. The dominant Higgs production mode is Higgsstrahlung: $e^+e^- \rightarrow ZH$. The background rates are low compared to hadron colliders.

2. Implementation

The samples used in this study were created by the ILD concept group since 2020. More than 400.000 events were simulated for each of the nine main Standard Model (SM) Higgs decay processes. As background, we consider SM processes with two or four fermions in the final state. All events are weighted according to their SM cross sections and to 2000 fb^{-1} integrated luminosity. While this is the integrated luminosity from the 20-year running scenario, the time would realistically be shared between left-polarized (80% left-polarized electron beam, 30% right-polarized positron beam) and right-polarized (80% right e^- , 30% left e^+) runs. Since this analysis is comparatively insensitive to beam polarization, we simplify the analysis by considering only the left-polarized run.

The events are simulated for the ILC250 using the SetA beam parameters [1]. Leading order event generation is performed by WHIZARD 2.8.5 [5, 6]. Initial State Radiation, Beamstrahlung and Final State Radiation (FSR) are included. The fragmentation and hadronization of final-state quarks and gluons is performed with PYTHIA 6.422 [7]. The ILD detector geometry is described with DD4hep [8] and simulated in GEANT4 [9]. Event reconstruction is performed with ILCSoft v02-02, which includes PandoraPFA [10] for the construction of particle flow objects and LCFIPlus [11] for flavor tagging.

3. Event selection

We focus on Higgsstrahlung events in which the primary Z boson decays leptonically, $Z \rightarrow e^+e^-$ or $Z \rightarrow \mu^+\mu^-$. Event variables can be reconstructed after the isolated lepton pair is identified and potentially combined with FSR photon candidates. Figure 1 shows the distribution of signal and background for $Z \rightarrow \mu^+\mu^-$ in the event variables that were built from the primary Z boson's decay products. The cuts from the previous steps are already applied in the subsequent steps. Lepton colliders have the advantage that the initial state kinematics are known. This is used when creating the recoil mass: $M_{\text{recoil}}^2 = s + M_Z^2 - 2\sqrt{s} \cdot E_Z$. For signal events, the distribution peaks

at $M_{\text{recoil}} = M_{\text{Higgs}}$. The presented cuts lead to a sample with about 50% purity at 50% signal efficiency. By construction, the efficiency is constant over Higgs decay modes and the resulting sample thus unbiased.

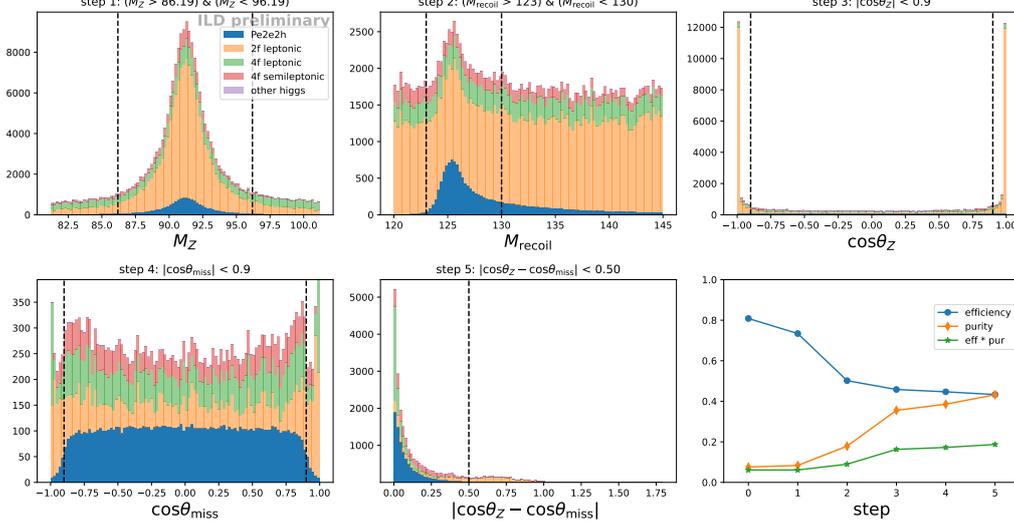


Figure 1: Cut flow for the event selection in the $Z \rightarrow \mu^+\mu^-$ sample. The last figure shows the efficiency and purity after each step.

We will now build a binned distribution from the sample, as shown in Figure 2. Each bin contains the events that fall into a specific class. The class definitions can be found in the appendix of [12]. A final bin contains all events that do not match any class definition. The relative sensitivity to a Higgs decay mode varies between bins. Next, we will present a fit that uses this variation in sensitivity to derive the Higgs branching ratios.

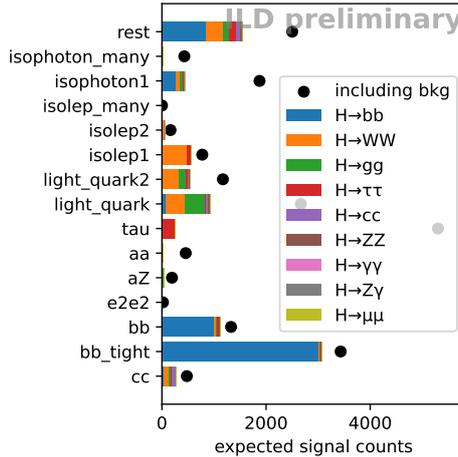


Figure 2: Expected contributions per class from each of the Higgs decay modes assuming the Standard Model branching ratios. The black dots additionally contain the background contribution to the $Z \rightarrow \mu^+\mu^-$ Higgsstrahlung sample after event selection. The class definitions are described in the appendix of [12].

4. Branching ratio uncertainties

4.1 Setting up the fit

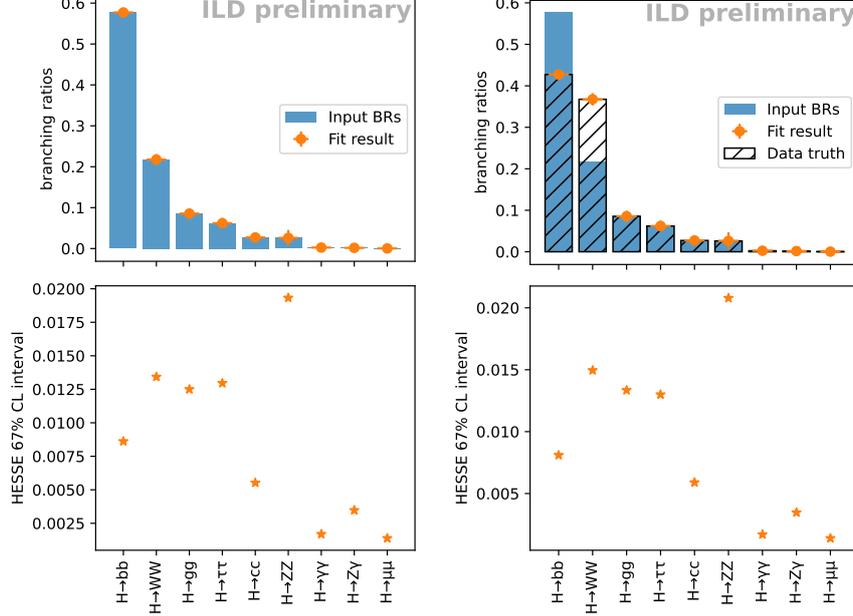


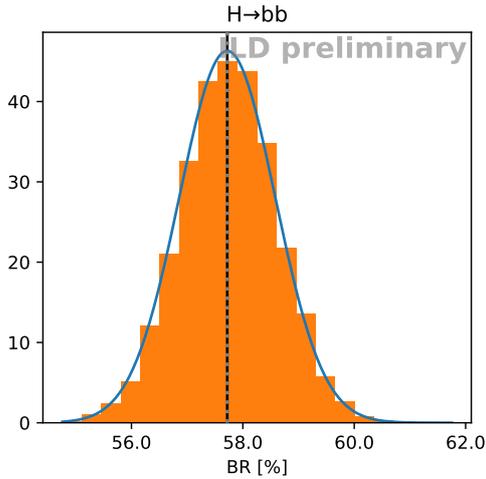
Figure 3: Top row: In blue, the BRs that were used in the simulation. The optimum of the fit and its uncertainty are indicated as orange error bars. The starting values of the fit are always the Standard Model Higgs BRs. The considered scenarios are the SM BRs (left) and a BSM scenario described in the text. Bottom row: The absolute statistical uncertainties of the fit per branching ratio.

The class probabilities for each Higgs decay mode or background process can be determined using the Monte Carlo simulated data.¹ They are combined into a matrix, where each column corresponds to one process. Applying the truth vector of the process counts to this matrix yields the class counts. Conversely, given the class counts, a fit can be performed to gain access to the process counts. Only the relative contribution of each Higgs BR is taken as a free parameter. With the nine BRs considered in the current analysis, the fit provides the prediction for eight free parameters and the corresponding 8×8 covariance matrix. They can be converted into mean values and uncertainties for each of the BRs.

4.2 Results

Figure 3 gives the resulting values for a combination of the $Z \rightarrow e^+e^-$ and $Z \rightarrow \mu^+\mu^-$ primary Z boson channels. The left part of the figure shows the result in the SM case. The right part of the figure supports the claim that the method is truly model independent. Here we assume a model with $BR(H \rightarrow b\bar{b})$ decreased by 15%, and $BR(H \rightarrow W^+W^-)$ increased by 15%. While this is an unrealistically large deviation from the SM, this shows that the procedure still works far away from the SM hypothesis.

¹The simulated events are needed both here and later on as a placeholder for the detector data. To prevent overfitting, we split the data set into two equal but statistically independent parts.



	SM BR	σ_{stat}
$H \rightarrow bb$	57.72	0.86
$H \rightarrow WW$	21.76	1.34
$H \rightarrow gg$	8.55	1.25
$H \rightarrow \tau\tau$	6.20	1.30
$H \rightarrow cc$	2.72	0.55
$H \rightarrow ZZ$	2.62	1.93
$H \rightarrow \gamma\gamma$	0.24	0.17
$H \rightarrow Z\gamma$	0.17	0.35
$H \rightarrow \mu\mu$	0.03	0.14

Figure 4: Left: Validation of the uncertainty from a fit to the expected event counts scenario using a toy study.

Right: Preliminary results of a MINUIT fit to the expected event counts. The table gives the fitted values of the Higgs BRs and their absolute statistical uncertainties. All numbers are in percentages.

The results are validated in a toy study. For this purpose we draw 10.000 class count vectors from a multinomial distribution centered on the expected counts for each class. The fit optimum for each toy is calculated and the corresponding BRs are stored in histograms. Figure 4 (left) displays the histogram for $H \rightarrow b\bar{b}$ in orange. The blue Gaussian has as its mean value the optimum of the fit on the expected event counts (EEC). Its standard deviation is obtained through the covariance matrix of the fit at the optimum. The blue Gaussian and the histogram agree well, indicating that the EEC fit is stable. The black line also indicates the position of the optimum of the EEC fit. As expected, it agrees with the BR value that was put into the simulation (dotted gray line).²

The (absolute) 1σ -uncertainties are listed in the table on the right side of Figure 4. These values assume BRs close to the SM expectation. Additional non-SM Higgs decays ($H \rightarrow$ invisible $H \rightarrow \mu^+\tau^-$, $H \rightarrow b\bar{c}$, ...) could be added to obtain explicit upper limits for them.

5. Conclusion

The Higgsstrahlung process at a lepton collider allows the measurement of all Higgs branching ratios (BRs) at once from within the same unbiased sample. At a Higgs factory such as the ILC250, the method presented in this contribution can be applied to perform precision measurements of the observed BRs and to obtain upper limits for further BRs. It is a pure BR measurement, independent of any production cross section measurement. The results can be combined with any analysis that does not use the $ZH \rightarrow (e^+e^-, \mu^+\mu^-)$ channels.

²Up to a small deviation due to the finite size of the simulated data sets.

Acknowledgments

The authors would like to thank the LCC generator working group and the ILD software working group for providing the simulation and reconstruction tools and producing the Monte Carlo samples used in this study. This work has benefited from computing services provided by the ILC Virtual Organization, supported by the national resource providers of the EGI Federation and the Open Science GRID.

References

- [1] LINEAR COLLIDER collaboration, *The International Linear Collider Machine Staging Report 2017*, [1711.00568](#).
- [2] H. Abramowicz et al., *The International Linear Collider Technical Design Report - Volume 4: Detectors*, [1306.6329](#).
- [3] ILD CONCEPT GROUP collaboration, *International Large Detector: Interim Design Report*, [2003.01116](#).
- [4] J.-C. Brient and H. Videau, *The Calorimetry at the future e^+e^- linear collider*, *eConf C010630* (2001) E3047 [[hep-ex/0202004](#)].
- [5] W. Kilian, T. Ohl and J. Reuter, *WHIZARD: Simulating Multi-Particle Processes at LHC and ILC*, *Eur. Phys. J. C* **71** (2011) 1742 [[0708.4233](#)].
- [6] M. Moretti, T. Ohl and J. Reuter, *O'Mega: An Optimizing matrix element generator*, [hep-ph/0102195](#).
- [7] S. Torbjorn, S. Mrenna and P.Z. Skands, *PYTHIA 6.4 Physics and Manual*, *JHEP* **05** (2006) 026 [[hep-ph/0603175](#)].
- [8] M. Frank, F. Gaede, M. Petric and A. Sailer, *Aidasoft/dd4hep: v01-16-01*, [DOI:10.5281/zenodo.4638724](#) (2021) .
- [9] GEANT4 COLLABORATION collaboration, *GEANT4—a simulation toolkit. GEANT4. A Simulation toolkit*, *Nucl. Instrum. Meth. A* **506** (2002) 250.
- [10] M. Thomson, *Particle flow calorimetry and the pandorapfa algorithm*, *Nucl. Instrum. Meth. A* **611** (2009) 25–40.
- [11] T. Suehara and T. Tanabe, *LCFIPlus: A Framework for Jet Analysis in Linear Collider Studies*, *Nucl. Instrum. Meth. A* **808** (2016) 109 [[1506.08371](#)].
- [12] J. Kunath et al., *A combined fit to the Higgs Branching Ratios at ILD*, in *International Workshop on Future Linear Colliders*, 5, 2021 [[2105.05718](#)].