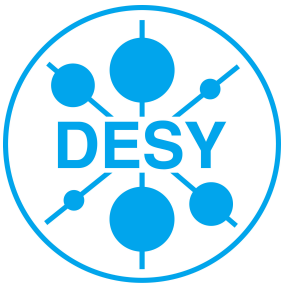


Search for ZH events with below SM Higgs masses at the 250GeV ILC in the $Z \to q\bar{q}$ channel

Jakob Beyer

Technische Universität Dresden

Germany

September 6, 2017

Abstract

The well-known initial conditions in e^+e^- colliders enable searches for the Higgs boson which are independent of the Higgs decay. Higgsstrahlung events ($e^+e^- \to ZH$) can be searched for by reconstructing the Z boson and finding the H through momentum conservation. Because the specific H decay (ideally) has no influence on this search it promises a near model-independence of the resulting parameters. In this report such a search is investigated for the future International Linear Collider with $\sqrt{s} = 250\text{GeV}$ in the $Z \to q\bar{q}$ channel for Standard Model like Higgs bosons with masses $m_H = \{30, 50, 70, 90, 115\}\text{GeV}$. Using the analysis developed by Thomson [1] for the $m_H = 125\text{GeV}$ case cuts are placed to suppress specific SM backgrounds. Cuts to exclude WW and ZZ backgrounds as well as the recoil mass cut are investigated separately since they leave detector signatures similar to the ZH signal. They are optimized by maximizing the signal significance. While the signal significance is found to increase with lower m_H the model-independence and the sharpness of the recoil mass peak decreases. Large significances show that this channel should be studied in more detail as it promises strong exclusion limits on the coupling parameters of any additional H.

Contents

1. Introduction

One simply cannot talk about particle physics without talking about the Standard Model (SM). This has good reason as it has been largely successful in predicting and describing the results of modern high energy physics experiments. However, the SM is not the theory of everything. Gravity, Dark Matter and Dark Energy are examples of experimental facts which cannot be explained by the SM. Theories such as Supersymmetry attempt to solve parts of these problems. Tests for those theories can be done e.g. by searching for predicted new elementary particles or particle substructures. Here the mass region of a few TeV or less is of special interest since resonances in it may suggest additional, possibly broken, symmetries.
This energy scale can be directly probed in modern particle colliders as is currently done at the LHC at CERN. With its pp collisions at center of mass energies of 13TeV direct searches for resonances are possible with high statistics. But the not well-known initial state conditions in a pp collider set a lower limit on the measurement uncertainties.
e^+e^- colliders are the complementary precision tools. Both collision particles are elementary so that the initial state is well-defined and does not smear the measurement results. This can be used to introduce model-independent measurements for particles recoiling against a well-measurable particle such as the Z.
In the present report this recoil mass method is investigated for the search for an additional Higgs boson with lower-than-SM mass at the $\sqrt{s} = 250$GeV ILC.

1.1. Additional Higgs bosons

The first question to be answered is why this search is important. Since a $m_H = 125$GeV Higgs boson was discovered at the ATLAS and CMS detectors at the LHC in 2012 the Standard Model itself does not require another lighter Higgs. Extensions to the Standard Model however can include such lower mass bosons. Theories such as Supersymmetry and Axion Models are motivated by both quantum field theoretical background and experimental need, e.g. for Dark Matter. In the examples of SUSY and Axion Models, new symmetries are introduced which turn out to require additional Higgs doublets. The coupling and mass of the non-SM Higgs may deviate from its SM counterpart, making a direct detection harder if the coupling is weak. Searching for additional lower mass Higgs and setting restrictions on its parameter space is therefore an important step in testing which - if any - of the proposed theories can explain the discrepancies of the Standard Model. More detailed motivations for light Higgs searches can be found in [2], [3].

1.2. The ILC

To test the Standard Model down to the percent level and search for Beyond the Standard Model physics up to several TeV a precision measurement tool such as the proposed International Linear Collider (ILC) is needed. Since its first outline of the physics goal in 2003 the international community worked on the detailed physics requirements and

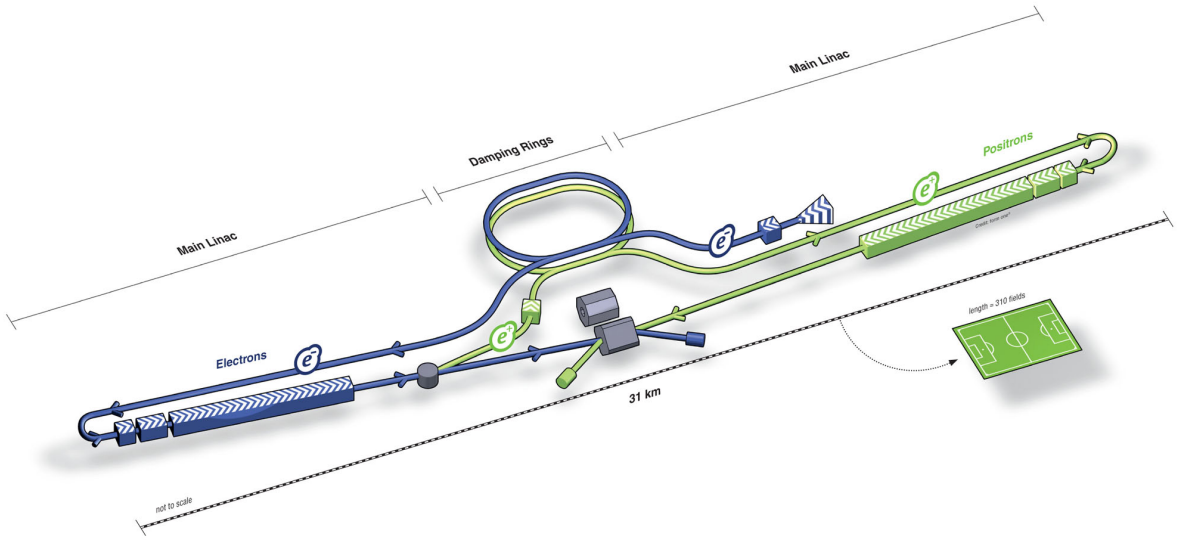


Figure 1: Schematic overview of the ILC experiment. [4]

possible technical implementations, culminating in the release of the Technical Design Report in 2013 [5, 6, 7, 8, 9] .
The design intends colliding electrons and positrons at center-of-mass energies in the range of 250GeV to 500GeV, with possibility of adjustment up- or downwards if needed. Superconducting radio-frequency cavities, such as currently in use at the European XFEL [10], are used as the major source of acceleration. At the interaction point the two beamlines meet at a crossing angle of 14mrad where one detector takes measurements while another is pushed into a hall next to the collisions. The two detectors can be switched using a "push-pull" method.
Current proposals for the detectors are the International Large Detector (ILD) and the Silicon Detector (SiD). Both detectors implement Particle Flow Algorithms (PFA). In PFAs the information from all detector-subsystems is taken into account when reconstructing the particle tracks rather than reconstructing in each subdetector separately. For the ILD optimal PFA performance is the design priority. A Time Projection Chamber (TPC) was developed for both track and energy reconstruction, surrounded by highly segmented calorimeters. TPC and calorimeters are placed in a 3.5T magnet using the return yoke for muon detection.

1.3. Recoil mass method

With the well-known initial state of the collision and the precise reconstruction by the ILC detectors a nearly model-independent measurement of the Higgs boson parameters is possible. The core of this measurement is the recoil mass method applied to the Higgsstrahlung process $e^+e^- \to HZ$ shown in fig. 2.
Instead of investigating the Higgs decay products the recoil mass method uses only information from the Z decay. The Z boson is identified by requiring a decay candidate

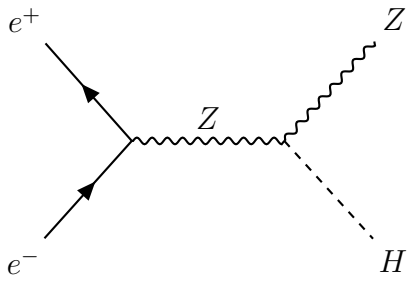


Figure 2: Feynman diagram of a Higgsstrahlungs process. This process is used in recoil mass measurement to measure the Higgs.

- e.g. $Z \to l^+l^-$ or $Z \to q\bar{q}$ - with an invariant mass close to m_Z. Taking into account the crossing angle at the interaction point the (ideal) 4-momentum of the initial state is known

$$(p^{\mu}_{init}) = \left(p^0, p^x, p^y, p^z\right) = \left(\sqrt{s}, \sqrt{s} \cdot \sin\left(\frac{0.014}{2}\right), 0, 0\right) . \tag{1}$$

Having measured the 4-momentum p_Z of the Z before its decay to leptons or quarks the 4-momentum of the recoil p_{rec} and the corresponding mass m_{rec} can be calculated.

$$p_{rec} = p_{init} - p_Z \tag{2}$$

$$\Rightarrow\ m^2_{rec} = (p_{rec})^2 = p^2_{init} - 2p_{init} \cdot p_Z - p^2_Z \tag{3}$$

In case of Higgsstrahlung the mass of the recoil is expected to be close to the Higgs boson mass. The recoil mass can therefore be used to distinguish such events from background e.g. where particles of other mass recoil against the Z.

1.3.1. Model-independence

To obtain the coupling parameter g_{HZZ} the total cross-section σ_{ZH} of the Higgsstrahlung process must be measured. σ_{ZH} includes all H decay modes. A deviation of the Higgs decay branching ratios from their Standard Model values should not influence this measurement.
The influence of the branching ratios B_i can be seen in the formula describing the number of measured events

$$N_{meas} = \sum_i \sigma_{ZH} B_{Z\to f\bar{f}} L B_i \epsilon_i \tag{4}$$

where $B_{Z\to f\bar{f}}$ is the branching ratio of the investigated Z decay, L is the total luminosity and ϵ_i is the efficiency of the search selection for the H decay channel i.

- **ideal case:**
  The detector and analysis do not introduce any biases towards specific Higgs decay modes.

$$\epsilon_i \equiv \epsilon = \text{const.} \tag{5}$$

  In this ideal case with

$$\sum_i B_i = 1 \tag{6}$$

  the measured number of events does not depend on the branching ratios. Assuming $B_{Z\to f\bar{f}}$ and ϵ are known the cross-section is then easily calculated as:

$$\sigma_{ZH} = \frac{N_{meas}}{B_{Z\to f\bar{f}} L\epsilon}\,. \tag{7}$$

- **realistic case:**
  If the efficiencies of the analysis depends on the H decay mode the branching ratios can not be eliminated from the cross-section calculation.

$$\sigma_{ZH} = \frac{N_{meas}}{B_{Z\to f\bar{f}} L \sum_i B_i\epsilon_i} \equiv \frac{N_{meas}}{B_{Z\to f\bar{f}} L\bar{\epsilon}} \tag{8}$$

$$\bar{\epsilon} = \sum_i B_i\epsilon_i = \sum_{\text{known}} B_i\epsilon_i + \sum_{\text{unknown}} B_i\epsilon_i \tag{9}$$

  For modes with known branching ratios B_i an average efficiency can be defined.

$$\epsilon_0 = \frac{\sum_{\text{known}} B_i\epsilon_i}{\sum_{\text{known}} B_i} \tag{10}$$

  By writing the efficiency of the unknown modes as

$$\epsilon_i = \epsilon_0 + \delta\epsilon_i \tag{11}$$

  the weighted efficiencies $\bar{\epsilon}$ can expressed as

$$\bar{\epsilon} = \epsilon_0 + \sum_{\text{unknown}} B_i \times \delta\epsilon_i\,. \tag{12}$$

  This shows that the influence of the H decay branching ratio scales with the deviation of the decay-channel specific efficiency.

To minimize the influence of the unknown as well as the known branching ratios, a uniform selection efficiency amongst the Higgs decay channels must be achieved.

1.4. Z boson decays

In this report the recoil mass method is investigated for the hadronic Z boson decay $Z \to q\bar{q}$. The quarks in the Z decay are measured as jets in the detector. Correctly reconstructing those jets and identifying the Z needs more effort than finding the high-energy muons of the leptonic Z decay. However, visible leptonic Z decay modes in total have a branching ratio of $\sim 10\%$. Almost 70% of the decays are hadronic [11].
In hadron colliders such as the LHC the strong hadronic background does not allow precision measurements in the hadronic decay modes. Here lepton colliders have the advantage of a well-known background due to the elementary nature of the colliding particles.
For the ILC it is therefore possible and important to investigate hadronic Z to increase the statistical significance of the experimental results.

2. Method

2.1. MC samples

To study the signal of Higgs bosons with lower-than-SM masses of $m_H = \{30, 50, 70, 90, 115\}$GeV Monte Carlo samples of the $e^+e^- \to HZ$, $Z \to q\bar{q}$ process are used. For the non-Standard Model H in those samples SM-like properties were assumed. As the goal of analysis is to be optimally model-independent taking the SM branching ratios of the H ideally should not affect the results.
All samples had been generated using Whizard version 1.95 [12] assuming center-of-mass energy of 250 GeV. The processes as $e^+e^- \to 2f$, $4f$, and $f\bar{f}H$ with the 125GeV SM H are used as the backgrounds in this analysis. Effects of beamstrahlung are simulated by GuineaPig [13] and included in the samples. The effects of initial state radiation are estimated as well [14]. Hadronization, fragmentation, and particle decays are handled by Pythia [15]. The tau lepton decay is simulated by Tauola [16],[17],[18]. Mokka is used for a Geant4-based detailed detector simulation [19],[20]. $\gamma\gamma \to$hadron backgrounds (överlay background) are added to the samples. Marlin is used for reconstruction and as analysis framework [21]. PFA reconstruction of particles is performed using the PandoraPFA package [22].
The samples are generated for four possible beam polarizations ($e^-_L e^+_L$, $e^-_R e^+_L$, $e^-_L e^+_R$, $e^-_R e^+_R$). For each polarization the signal sample contains 20,000 events.
Investigated was a 250fb^{-1} run of the $\sqrt{s} = 250$GeV ILC with detection using the International Large Detector.

2.2. Durham jet clustering

Quarks and gluons in the final state of the hard scattering process form jets. Jets are clusters of hadrons created due to confinement in the strong force. The individual hadrons can be measured in the detector and must then be clustered by an algorithm to regain the structure of the scattering process at the interaction point.

An example of such an algorithm is Durham jet clustering [23]. In it an energy-scaled distance parameter $y(i,j)$ between two particles is defined.

$$y(i,j) = \frac{2E_iE_j\,(1-\cos\theta_{ij})}{E_{\text{vis}}^2} \tag{13}$$

Here $E_{i/j}$ is the energy of particle i/j, θ_{ij} is the angle between the tracks of the particles and E_{vis} is the total visible energy in the detector.
The two particles with the lowest $y(i,j)$ are clustered to a jet and their 4-momenta are added. After they are clustered they are treated as one particle. This process is repeated until the breaking condition is reached. Each jet is then assumed to correspond to a particle in the final state of the hard process. In order to reconstruct the correct number of jets in an event a maximum clustering distance y_{cut} can be defined. When the minimal distance parameter between to particles surpasses this cut value the particles/jets are assumed to be flying in defined well-directions which are to different to belong to the same initial particle.
Alternatively, as used in this report, the algorithm can be iterated until a certain number of jets is reached. The minimal distance parameter between the resulting clusters then gives insight into whether the required jet topology is likely to be correct. If the distance parameter is small compared to 1 the true topology is likely to contain less jets. Large parameters can mean that either the topology is likely to be correct or the event was forced into a topology with too few jets.
In the following the notation y_{ij} stands for the distance parameter between the two clusters which where combined to force the j-jet topology into the i-jet topology.

2.3. Analysis framework

The analysis and selection of the events was done using the *Marlin* framework [21] provided in *ilcsoft* [24] together with code in *C++* and *Python*, both implementing *ROOT* [25]. Settings for the Marlin processors were written in *xml* files.

2.4. Preselection

The selection procedure used in this report is based on the work done in [1]. Jet clustering is done using the Durham clustering algorithm implemented in Marlin which is described in 2.2. Isolated leptons are identified using the multi-variate based Isolated Lepton Tagging [26] implemented in Marlin.
Both the visibility categorization and the recoil mass calculation below require jets containing a minimum number of charged particles. This intends to assure that the jets are of hadronic origin (quark, gluon) since hadronic jets contain a high number of charged particles.

2.4.1. Higgs decay visibility

Events are first categorized into visible and invisible H decays. This contradicts the philosophy of the recoil mass measurement to stay model-independent by not analysing

the H decay products. But because Higgs decay which are visible in the detector can lead to mis-clustering between the Z and H decay products such a distinction is necessary. Near model-independence could be restored by recombining the results of visible and invisible category.
The procedure to distinguish invisible from visible H decays takes the following steps.

- The event is clustered into two jets. If they do not each contain more than two charged particles the event is discarded.
- An event is marked as invisible candidate if $\log_{10}(y_{23}) < -2.0$ and $\log_{10}(y_{34}) < -3.0$. Such small distance parameters indicate an event with less than three jets. This includes invisible H decays in which only the two jets from $Z \to q\bar{q}$ are seen.
- Invisible candidate events with $\log_{10}(y_{23}) > -2.5$ and $\log_{10}(y_{34}) > -3.5$ are clustered into three jets and categorized as visible if the lowest-energy jet contains less than four tracks are found or a $E > 5$GeV $e^{\pm}$ or $\mu^{\pm}$. This prevents H decays with low-energy charged leptons from being clustered as invisible.

After the preselection the distinction between visible and invisible categorized events is not considered in this report. Due to the importance for achieving near model-independence this should be done in future research. It is ignored here because the Monte Carlo samples contain only around 100 invisible Higgs decays due to their small branching ratio in the Standard Model.

2.4.2. m_{rec} reconstruction

For events categorized as invisible the two jets are identified as the $Z \to q\bar{q}$ candidate and the mass of the recoil against the Z is calculated.
In the visible category 4- and 5-jet topologies are considered. First the event is forced into a 4-jet topology. To find the best 2-jet $Z \to q\bar{q}$ candidate $|m_{qq} - m_Z|$ is minimized with m_{qq} being the mass of the combined 4-momentum. The recoil mass against this Z candidate is then calculated. Five-(or higher-)jet-multiplicities are considered by clustering the event to five-jets and searching for the best Z candidate as above. The higher-multiplicity result is kept if $\log_{10}(y_{45}) < -3.5$ and $m_{q\bar{q}}$ and m_{rec} are closer to m_Z and the investigated m_H respectively.

2.4.3. Preselection cuts

The goal of the preselection as described in [1] is to exclude specific Standard Model backgrounds while keeping the selection efficiencies as uniform as possible to achieve model-independence. Because the signal of the Higgs boson depends on its mass those cuts are not optimal for lighter Higgs searches. Due to time constraints on this work only three cuts are investigated separately. For this the $WW \to q\bar{q}q\bar{q}$ and $ZZ \to q\bar{q}q\bar{q}$ exclusions and the recoil mass cut are chosen. They are expected to have a strongly differing influence on the result when the new Higgs mass is varied near the W and Z mass.
An overview of the cuts which where kept identical to [1] is shown in tab. 1.

Table 1: Preselection cuts adopted from [1], with information whether they are applied to the visible or invisible candidates and with their intended purpose. These cuts are executed universally for all m_H.

cut description	applied to (vis./inv.)	intention
$70\text{GeV} < m_{q\bar{q}} < 110\text{GeV}$	both	best Z candidate in event must have mass close to m_Z
reject if $p_T^{\text{net}} < 20\text{GeV}$ and $\log_{10}(y_{34}) > 2.5$	both	suppress $e^+e^- \to q\bar{q}$, small y_{34} indicating lower jet multiplicity
reject if $\lvert\cos(\theta_{\text{miss}})\rvert > 0.7$, θ_{miss} is the polar angle of the missing momentum	invis.	suppress $e^+e^- \to q\bar{q}$ with unobserved initial state radiation seen in missing momentum along beam axis
reject if $p_T^{\text{net}} < 20\text{GeV}$ and $\lvert\cos(\theta_{\text{miss}})\rvert > 0.9$	both	suppress $e^+e^- \to q\bar{q}$ with unobserved ISR and possible radiative return to Z resonance
reject if isolated $e^\pm$ or $\mu^\pm$ with energy $E_{e/\mu} > 10\text{GeV}$ can be found	invis.	suppress $e^+e^- \to WW \to q\bar{q}l\nu$

2.5. WW and ZZ exclusion

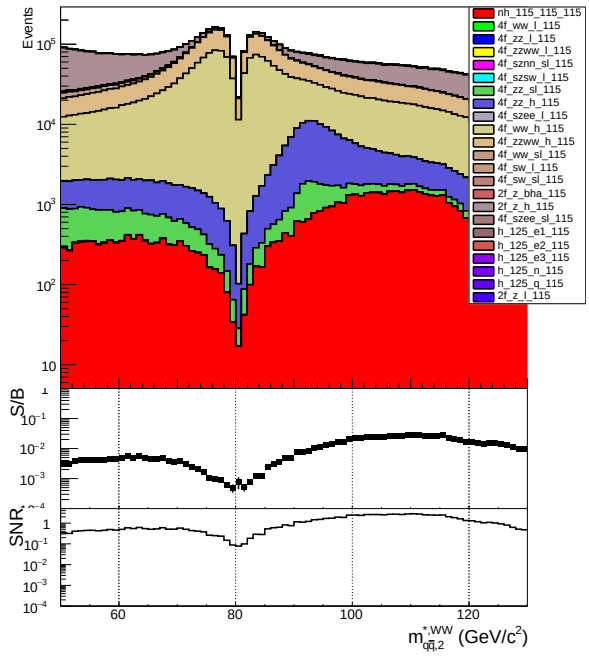
With the preselection results a cut-based analysis is performed. This analysis is solely based on maximizing the signal significance S.

$$S = \frac{\#\text{ Signal events}}{\sqrt{\#\text{ Signal events} + \#\text{ Background events}}} \tag{14}$$

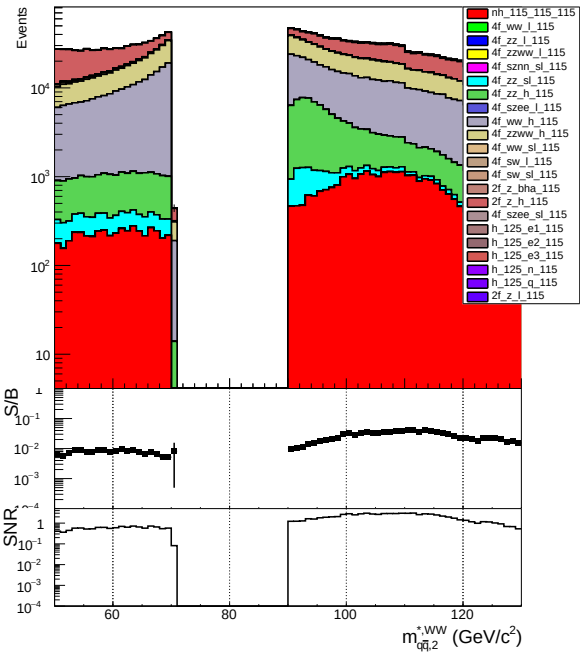
First the $WW \to q\bar{q}q\bar{q}$ and $ZZ \to q\bar{q}q\bar{q}$ rejections are optimized.
To find $WW \to q\bar{q}q\bar{q}$ like events the particles are forced into four jets. The 2-jet combinations most similar to WW are chosen by minimizing

$$\chi^2 = (m_{q\bar{q},1} - m_W)^2 + (m_{q\bar{q},2} - m_W)^2 \tag{15}$$

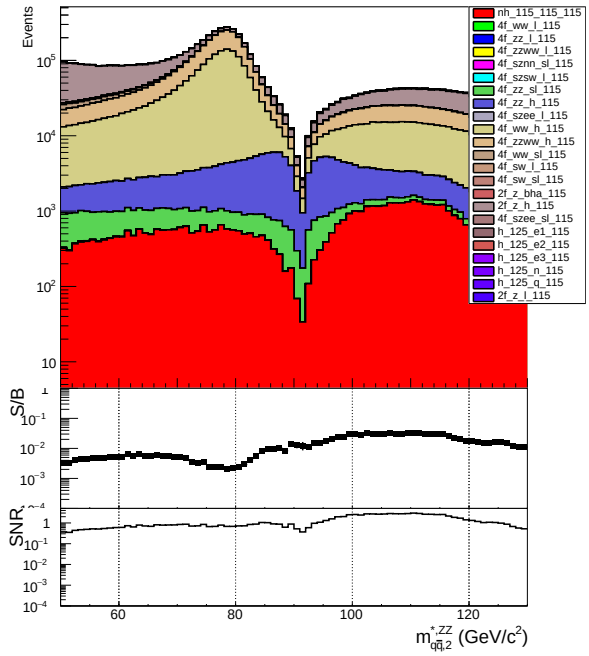
where $m_{q\bar{q},i}$ is the mass corresponding to the combined 4-momenta of the two jets. If both $m_{q\bar{q},1}$ and $m_{q\bar{q},2}$ are within the determined cut range (min, max) the event is discarded. Rejecting $ZZ \to q\bar{q}q\bar{q}$ like events is done analogous to the WW case using the m_Z mass.
For the cut optimization the $m_{q\bar{q},i}$ with larger distance to the respective boson mass $|m_{q\bar{q},i} - m_X|$ is filled into a histogram. The cuts are chosen symmetrically with a midpoint approximating the respective boson mass. As a consequence their influence can be tested by simply analysing the number of events outside the cut range in this histogram. The second mass $m_{q\bar{q},j\neq i}$ will by construction for most cases also fall within the cut range if $m_{q\bar{q},i}$ does.
Examples for such $m_{q\bar{q},i}$ spectra on which the cuts are tested can be found in fig. 3



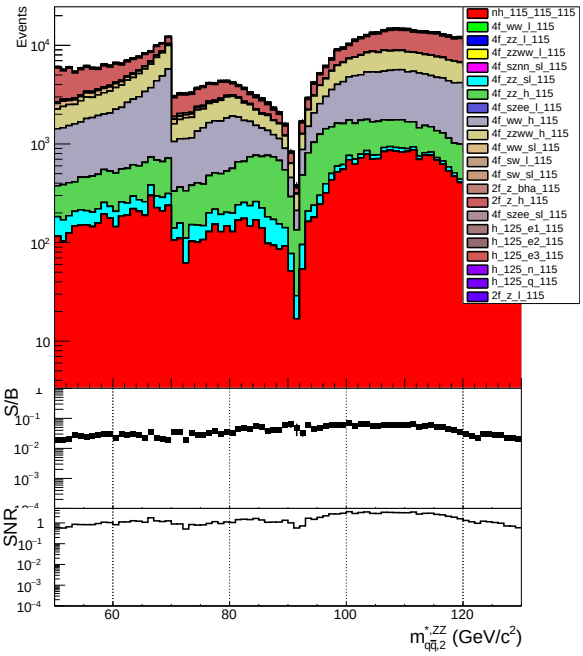
(a) After preselection. No $WW/ZZ \to q\bar{q}q\bar{q}$ rejection.



(b) After preselection and $WW/ZZ \to q\bar{q}q\bar{q}$ rejection.



(c) After preselection. No $WW/ZZ \to q\bar{q}q\bar{q}$ rejection.



(d) After preselection and $WW/ZZ \to q\bar{q}q\bar{q}$ rejection.

Figure 3: 4-jet topology is clustered to two 2-jets such that both have mass as close as possible to m_W and m_Z, respectively. Shown here are spectra for the 2-jet mass of the both which is further apart from the respective boson mass. Cut on this is almost identical to $XX \to q\bar{q}q\bar{q}$ rejection and is used for cut analysis. Diplayed for $m_H = 115$GeV, before and after rejection of both $WW \to q\bar{q}q\bar{q}$ and $ZZ \to q\bar{q}q\bar{q}$. No m_{rec} cuts are applied.

Table 2: Limits of the $WW \to q\bar{q}q\bar{q}$ and $ZZ \to q\bar{q}q\bar{q}$ exclusion cuts for the different signal Higgs masses m_H. Events are forced into a 4-jet topology and the best WW (ZZ) pair combination is found. If the mass of both W (Z) is within the given range the event is rejected.

	m_H:	30GeV	50GeV	70GeV	90GeV	115GeV
$WW \to q\bar{q}q\bar{q}$ exclusion	min	65	65	75	75	70
	max	95	95	85	85	90
$ZZ \to q\bar{q}q\bar{q}$ exclusion	min	85	-	-	-	-
	max	95	-	-	-	-

and in the Appendix in fig. 11 and fig. 12 (therein $m_{q\bar{q},i}$ is called $m^{WW}_{q\bar{q},2}/m^{ZZ}_{q\bar{q},2}$). The spectrum can be understood as a combination of mass spectrum of the visible recoil (such as a Z or W peak) and a drop towards the mass of the investigated boson $m_{W/Z}$. This drop originates from the definition of this mass spectrum. Shown is the mass which is further away from the boson mass $m_{W/Z}$. Therefore the combinatorial probability for it increases as it moves further from $m_{W/Z}$ because the other mass has a wider range to fall into.
After minimizing χ^2 for the WW cut an additional ZZ rejection cut is found to increase the significance only for the $m_H = 30$GeV case. This may be because m_H is closer to m_Z than in the Standard Model case. The H could be off-shell inside the cut range. Alternatively by minimizing χ^2 the jets may be wrongly combined such that both resulting masses fall into the cut range.
If the ZZ cut is performed first it shows a small increase in significance compared to the WW cut. When placing an additional ZZ cut on the remaining events the significance does not reach the level achieved by only applying the WW cut. An exception is the $m_H = 30$GeV case where the cut order has no influence on the optimization. This decrease in significance due to the ZZ cut is attributed to the overlap of the two cuts. Because they are performed using the same algorithms with boson masses just ≈ 10GeV apart the cut with the larger exclusion interval is expected to also exclude many of the events falling into the exclusion region of the other cut.
The results are shown in tab. 2.

2.6. m_{rec} cut placement

With the recoil mass calculated as described in section 2.4.2 a simple cut analysis is performed. As in the previous optimization only the signal significance S (eq. (14)) is considered. To find the cut that maximizes S the rejection interval is enforced on the m_{rec} spectrum and the resulting significance is calculated. The resulting cut intervals are shown in tab. 3.

Table 3: Recoil mass cuts for the different lighter Higgs masses m_H. Events that fall outside the given interval are rejected.

	m_H	30GeV	50GeV	70GeV	90GeV	115GeV
m_{rec}	min	0	0	40	75	100
	max	160	160	160	155	155

3. Results

3.1. Recoil mass spectra

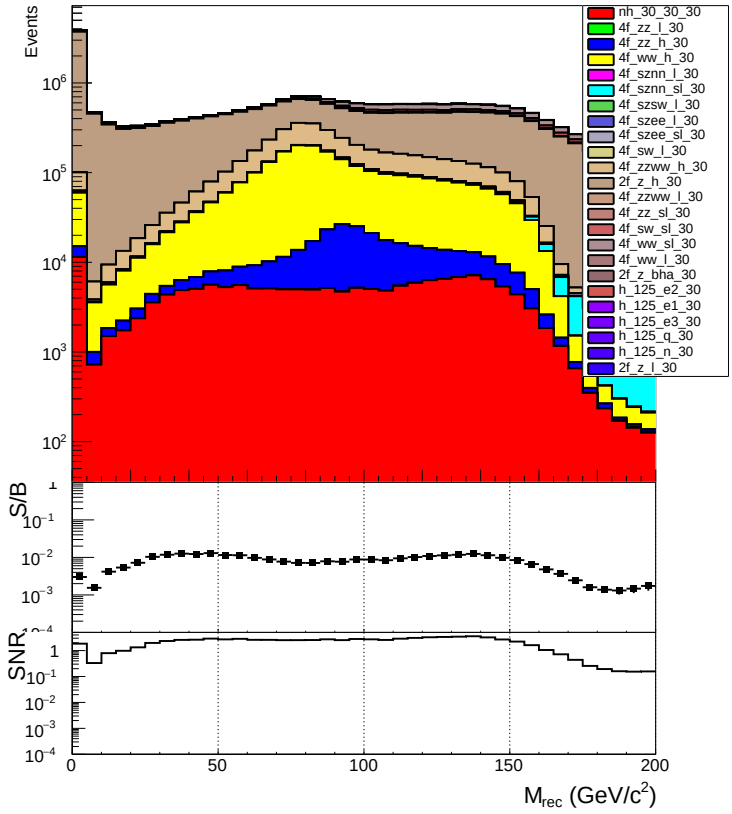


Figure 4: Recoil mass spectrum for $m_H = 30$GeV before any cuts are applied. The lower plots show the signal-to-background and signal-to-noise ratios.

The recoil mass spectra for all investigated Higgs masses and for different stages in the selection process are shown in the appendix in figures 5-9. As an example the m_{rec} spectrum for $m_H = 30$GeV before any selection cuts is displayed in fig. 4. Its properties are discussed specifically as they show some phenomena emerging when going to lower masses.
The expected peak at 30GeV is observed. It is however not as sharp as for the Standard Model case. Additionally, above around 90GeV the spectrum turns into a second peak

with a maximum near 140GeV. This peak may be explained by unobserved initial state radiation (ISR) wherein the initial $e^{\pm}$ radiates of a high energy photon in beam direction. Because the photon can not be detected the center-of-mass energy $\sqrt{s}$ of the observed e^+e^- scattering process is reduced. The effect on the calculated recoil mass can be seen by rewriting eq. (3), ignoring the crossing angle for simplicity.

$$m_{rec}^2 = s - 2\sqrt{s}E_{q\bar{q}} + m_{q\bar{q}}^2 \tag{16}$$

If $\sqrt{s}$ is overestimated (as in the case where ISR reduces the effective $\sqrt{s}$) m_{rec} will be also be overestimated. In an event in which just enough energy effective is left for a $e^+e^- \to ZH$ process to occur the total momentum of the Z candidate will be small. Then its energy is $E_{q\bar{q}} \approx m_{q\bar{q}} \approx m_Z$ and the recoil mass for an assumed $\sqrt{s} = 250$GeV yields $m_{rec} \approx 160$GeV. At this point in the recoil mass spectrum a corresponding drop is observed. Even higher recoil masses can then be explained by the width of the m_Z and m_H spectra.
Another deviation from the recoil mass spectrum of the SM H is the peak at $m_{rec} = 0$. The peak strongly increases with decreasing m_H and is yet not explained. However, the fact that this peak occurs at exactly $\sqrt{s}$ even when the resolution is increased hints that this may be a numerical or algorithmic error in the code.

3.2. Signal significance and efficiency

Signal significances and efficiencies for the selection steps described in section 2 are summarised in tab. 4.
The largest increase in significance and decrease in efficiency are observed when applying the preselection. Several SM backgrounds are rejected by some specific signature. In the signal sample this background rejection leads to a decrease of efficiency with decreasing m_H. Meanwhile the significance increases with decreasing m_H due to the higher total cross-section. After the other cuts this behaviour of the significance remains.
After the $WW \to q\bar{q}q\bar{q}$ and $ZZ \to q\bar{q}q\bar{q}$ rejections the efficiency still decreases with decreasing m_H. When the recoil mass cut is applied this changes. The highest efficiency is found for $m_H = 70$GeV, decreasing towards both sides. The decrease for higher masses can be explained by the sharper recoil mass peak which allows a narrower cut. While this leads to a stronger decrease in efficiency the signal significance also shows a higher increase.
As the preselection does the strongest cut on the signal it should be further investigated. This is especially the case for lower m_H where the decrease in efficiency is strongest.

3.3. H decay channel dependence

As the one of the motivations for using recoil mass measurements is their model-independence the decay-channel-dependent efficiency needs to be investigated. For the analysis presented here those efficiencies for selected decay channels after all cuts are presented in tab. 5.

Table 4: Signal significances and efficiencies with respect to the raw reconstructed MC sample for all new Higgs masses and analysis steps.

	m_H	30GeV	50GeV	70GeV	90GeV	115GeV
no Cuts	S	27.42	22.80	18.92	15.18	10.36
Preselection	S	42.20	39.96	35.80	30.64	21.54
	ϵ	53.68%	61.18%	66.08%	70.50%	72.67%
after $m_{W/Z}$ cut	S	48.92	46.83	38.40	32.32	23.97
	ϵ	48.96%	56.43%	64.91%	68.11%	67.81%
all cuts	S	49.05	46.99	39.28	35.95	28.97
	ϵ	48.54%	56.02%	62.97%	61.03%	52.93%

Table 5: Efficiencies for different decay channels of the additional Higgs with respect to the raw reconstructed MC sample for all new Higgs masses.

m_H:	30GeV	50GeV	70GeV	90GeV	115GeV
H decay mode	ϵ				
$q\bar{q}$	49.10%	56.32%	64.25%	62.28%	52.40%
gg	44.25%	50.53%	58.78%	57.86%	48.47%
$WW \to q\bar{q}q\bar{q}$	45.60%	54.54%	58.81%	54.72%	46.87%
$W\bar{W} \to l\nu l\nu$	39.27%	40.17%	51.66%	53.50%	56.30%
$\tau\tau$	40.71%	47.55%	50.83%	55.22%	54.30%
$\nu\nu$	41.61%	43.52%	47.00%	25.27%	39.86%
$\gamma\gamma$	5.86%	11.71%	22.87%	31.22%	38.33%
all	48.54%	56.02%	62.97%	61.03%	52.93%

In none of the m_H cases near model-independence is found. This is due to procedure used in this report. While the analysis done by Thomson [1] achieves near model-independence for the SM Higgs their analysis is not optimized for other mass cases. By using the same preselection cut the same decay channel efficiencies cannot be expected. Model-independence is further not taken into account in the cut-based analysis.
In this report model-independence was not prioritized due to the limited time frame and the small number of events in the signal sample. For an optimization and alignment of the decay channel efficiencies the signal samples need to include statistically sufficient amounts of all major expected H decays.

4. Conclusion

For the search for additional lower mass Higgs bosons at the 250fb^{-1} the SM case search by Thomson [1] has been implemented and modified. While most of the cuts done in the analysis were kept identical a cut-based analysis was performed for the $WW \to q\bar{q}q\bar{q}$ and $ZZ \to q\bar{q}q\bar{q}$ rejection and recoil mass selection. This cut was solely optimized with the goal of maximizing the signal significance.

After the complete selection signal significances ranging from almost 30 for higher H masses ($m_H = 115\text{GeV}$) to almost 50 for lower masses ($m_H = 30\text{GeV}$) were achieved.
However, achieving model-independence was not in the scope of this report. Therefore large differences of over 40% are observed in the efficiencies for different H decay channels. As model-independence is one of the strongest motivations to use recoil mass measurements further studies should be done using higher-statistics samples to implement a model-independent version of this analysis.
For lower m_H additional features were found in the recoil mass spectrum. A peak at energies around 140GeV could be explainable by initial-state radiation. At 0GeV an additional yet unexplained peak was observed.
Due to the limited time frame for this report several additional features could not be investigated which are of importance for the final analysis. This includes the detailed influence of the preselection cuts, separate cut optimization for visible and invisible candidates, higher jet multiplicities and additional cuts on the Z candidate.
While they remain to be studied this report found that a search for additional low mass Higgs bosons at the ILC in the $e^+e^- \to ZH$, $Z \to q\bar{q}$ channel promises high precision measurements.

5. Acknowledgements

Through lots of work, fun, almost sleepless nights and restless days I can say with confidence that I have had an incredible and wonderful experience here at DESY. For that I want to give my deep gratitude to the organizers of this summer student program, especially Olaf Behnke as he made sure that we would not find one second of boredom while we were here.
I would like to thank Shin-ichi Kawada who, as my supervisor, had to deal with all my beginner's mistakes, showed patience with me despite all of them and helpfully guided me to a solution whenever I got too stuck to continue. My appreciation goes out as well to Yan Wang for laying the groundwork of my studies by providing me with his code and knowledge, and Junping Tian from the University of Tokyo for generating samples for me no matter how short notice - even when almost on his way to his vacation.
Last but most certainly not least I want to thank all the other summies who worked, danced, attended lectures and movies, ate lunch and had something resembling a social life with me. Especially my fellow FLC nerds Alina and Chris and the whole rest of the Hostel7a+ crew. All our wonderfully chaotic tours and card games shall never be forgotten.

Thank you,
Jakob

6. References

[1] M. A. Thomson. "Model-independent measurement of the $e^+e^- \to HZ$ cross section at a future e^+e^- linear collider using hadronic Z decays". In: *The European Physical Journal C* 76.2 (Feb. 2016), p. 72. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-016-3911-5. URL: https://doi.org/10.1140/epjc/s10052-016-3911-5.

[2] G.C. Branco, P.M. Ferreira, L. Lavoura, M.N. Rebelo, Marc Sher, and João P. Silva. "Theory and phenomenology of two-Higgs-doublet models". In: *Physics Reports* 516.1 (2012). Theory and phenomenology of two-Higgs-doublet models, pp. 1–102. ISSN: 0370-1573. DOI: http://dx.doi.org/10.1016/j.physrep.2012.02.002. URL: http://www.sciencedirect.com/science/article/pii/S0370157312000695.

[3] Alexander Belyaev, Renato Guedes, Stefano Moretti, and Rui Santos. "Very light Higgs bosons in extended models at the LHC". In: *Phys. Rev. D* 81 (9 May 2010), p. 095006. DOI: 10.1103/PhysRevD.81.095006. URL: https://link.aps.org/doi/10.1103/PhysRevD.81.095006.

[4] *ILC schematic*. Last checked: 2017-08-28. URL: https://www.linearcollider.org/images/ILC_schematic.jpg.

[5] Ties Behnke, James E. Brau, Brian Foster, Juan Fuster, Mike Harrison, James McEwan Paterson, Michael Peskin, Marcel Stanitzki, Nicholas Walker, and Hitoshi Yamamoto. "The International Linear Collider Technical Design Report - Volume 1: Executive Summary". In: (2013). arXiv: 1306.6327 [physics.acc-ph].

[6] Howard Baer, Tim Barklow, Keisuke Fujii, Yuanning Gao, Andre Hoang, Shinya Kanemura, Jenny List, Heather E. Logan, Andrei Nomerotski, Maxim Perelstein, et al. "The International Linear Collider Technical Design Report - Volume 2: Physics". In: (2013). arXiv: 1306.6352 [hep-ph].

[7] Chris Adolphsen, Maura Barone, Barry Barish, Karsten Buesser, Philip Burrows, John Carwardine, Jeffrey Clark, Hélène Mainaud Durand, Gerry Dugan, Eckhard Elsen, Atsushi Enomoto, Brian Foster, Shigeki Fukuda, Wei Gai, Martin Gastal, Rongli Geng, Camille Ginsburg, Susanna Guiducci, Mike Harrison, Hitoshi Hayano, Keith Kershaw, Kiyoshi Kubo, Victor Kuchler, Benno List, Wanming Liu, Shinichiro Michizono, Christopher Nantista, John Osborne, Mark Palmer, James McEwan Paterson, Thomas Peterson, Nan Phinney, Paolo Pierini, Marc Ross, David Rubin, Andrei Seryi, John Sheppard, Nikolay Solyak, Steinar Stapnes, Toshiaki Tauchi, Nobu Toge, Nicholas Walker, Akira Yamamoto, and Kaoru Yokoya. *The International Linear Collider Technical Design Report - Volume 3.I: Accelerator RD in the Technical Design Phase*. Tech. rep. arXiv:1306.6353. ANL-HEP-TR-13-20. BNL-100603-2013-IR. IRFU-13-59. CERN-ATS-2013-037. Cockcroft-13-10. CLNS-13-2085. DESY-13-062. FERMILAB-TM-2554. IHEP-AC-ILC-2013-001. ILC-REPORT-2013-040. INFN-13-04-LNF. JAI-2013-001. JINR-E9-2013-35. JLAB-R-2013-01. KEK-Report-2013-1. KNU-CHEP-ILC-2013-1. LLNL-TR-635539. SLAC-R-1004. ILC-HiGrade-Report-2013-003. Comments: See also http://www.linearcol-

lider.org/ILC/TDR . The full list of signatories is inside the Report. Geneva, June 2013. URL: https://cds.cern.ch/record/1541315.

[8] Chris Adolphsen, Maura Barone, Barry Barish, Karsten Buesser, Philip Burrows, John Carwardine, Jeffrey Clark, Hélène Mainaud Durand, Gerry Dugan, Eckhard Elsen, et al. "The International Linear Collider Technical Design Report - Volume 3.II: Accelerator Baseline Design". In: (2013). arXiv: 1306.6328 [physics.acc-ph].

[9] Halina Abramowicz et al. "The International Linear Collider Technical Design Report - Volume 4: Detectors". In: (2013). Ed. by Ties Behnke, James E. Brau, Philip N. Burrows, Juan Fuster, Michael Peskin, Marcel Stanitzki, Yasuhiro Sugimoto, Sakue Yamada, and Hitoshi Yamamoto. arXiv: 1306.6329 [physics.ins-det].

[10] *European XFEL - Accelerator*. Last checked: 2017-08-31. URL: https://www.xfel.eu/facility/accelerator/index_eng.html.

[11] C. Patrignani and Particle Data Group. "Review of Particle Physics". In: *Chinese Physics C* 40.10 (2016), p. 100001. URL: http://stacks.iop.org/1674-1137/40/i=10/a=100001.

[12] Wolfgang Kilian, Thorsten Ohl, and Jürgen Reuter. "WHIZARD—simulating multi-particle processes at LHC and ILC". In: *The European Physical Journal C* 71.9 (Sept. 2011), p. 1742. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-011-1742-y. URL: https://doi.org/10.1140/epjc/s10052-011-1742-y.

[13] Daniel Schulte. "Study of Electromagnetic and Hadronic Background in the Interaction Region of the TESLA Collider". PhD thesis. DESY, 1997. URL: http://inspirehep.net/record/888433/files/shulte.pdf.

[14] M. Skrzypek and S. Jadach. "Exact and approximate solutions for the electron nonsinglet structure function in QED". In: *Zeitschrift für Physik C Particles and Fields* 49.4 (Dec. 1991), pp. 577–584. ISSN: 1431-5858. DOI: 10.1007/BF01483573. URL: https://doi.org/10.1007/BF01483573.

[15] T. Sjöstrand, S. Mrenna, and P. Skands. "PYTHIA 6.4 physics and manual". In: *Journal of High Energy Physics* 5, 026 (May 2006), p. 026. DOI: 10.1088/1126-6708/2006/05/026. eprint: hep-ph/0603175.

[16] Stanislaw Jadach, Johann H. Kuhn, and Zbigniew Was. "TAUOLA: A Library of Monte Carlo programs to simulate decays of polarized tau leptons". In: *Comput. Phys. Commun.* 64 (1990), pp. 275–299. DOI: 10.1016/0010-4655(91)90038-M.

[17] P. Golonka, B. Kersevan, T. Pierzchala, E. Richter-Was, Z. Was, and M. Worek. "The Tauola photos F environment for the TAUOLA and PHOTOS packages: Release. 2." In: *Comput. Phys. Commun.* 174 (2006), pp. 818–835. DOI: 10.1016/j.cpc.2005.12.018. arXiv: hep-ph/0312240 [hep-ph].

[18] N. Davidson, G. Nanava, T. Przedzinski, E. Richter-Was, and Z. Was. "Universal Interface of TAUOLA Technical and Physics Documentation". In: *Comput. Phys. Commun.* 183 (2012), pp. 821–843. DOI: 10.1016/j.cpc.2011.12.009. arXiv: 1002.0543 [hep-ph].

[19] P. Mora de Freitas and H. Videau. "Detector simulation with MOKKA / GEANT4: Present and future". In: *Linear colliders. Proceedings, International Workshop on physics and experiments with future electron-positron linear colliders, LCWS 2002, Seogwipo, Jeju Island, Korea, August 26-30, 2002*. 2002, pp. 623–627. URL: http://www-library.desy.de/cgi-bin/showprep.pl?lc-tool03-010.

[20] S. Agostinelli et al. "Geant4—a simulation toolkit". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3 (2003), pp. 250–303. ISSN: 0168-9002. DOI: http://dx.doi.org/10.1016/S0168-9002(03)01368-8. URL: http://www.sciencedirect.com/science/article/pii/S0168900203013688.

[21] F. Gaede. "Marlin and LCCD: Software tools for the ILC". In: *Nucl. Instrum. Meth.* A559 (2006), pp. 177–180. DOI: 10.1016/j.nima.2005.11.138.

[22] M. A. Thomson. "Particle flow calorimetry and the PandoraPFA algorithm". In: *Nuclear Instruments and Methods in Physics Research A* 611 (Nov. 2009), pp. 25–40. DOI: 10.1016/j.nima.2009.09.009. arXiv: 0907.3577 [physics.ins-det].

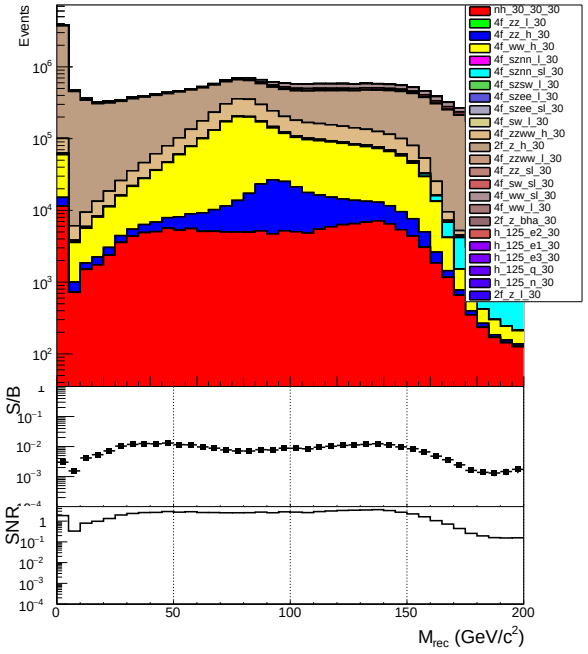
[23] S. Catani, Yu.L. Dokshitzer, M. Olsson, G. Turnock, and B.R. Webber. "New clustering algorithm for multijet cross sections in e+e− annihilation". In: *Physics Letters B* 269.3 (1991), pp. 432–438. ISSN: 0370-2693. DOI: http://dx.doi.org/10.1016/0370-2693(91)90196-W. URL: http://www.sciencedirect.com/science/article/pii/037026939190196W.

[24] *ilcsoft Software package*. Last checked: 2017-09-01. URL: http://ilcsoft.desy.de/portal/.

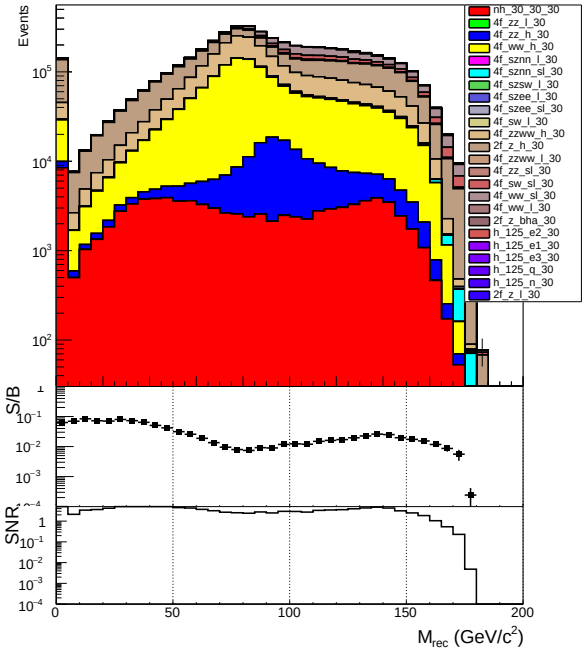
[25] Rene Brun and Fons Rademakers. "ROOT — An object oriented data analysis framework". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 389.1 (1997). New Computing Techniques in Physics Research V, pp. 81–86. ISSN: 0168-9002. DOI: http://dx.doi.org/10.1016/S0168-9002(97)00048-X. URL: http://www.sciencedirect.com/science/article/pii/S016890029700048X.

[26] Claude Dürig Junping Tian. *Isolated Lepton Finder*. Last checked: 2017-09-05. URL: https://agenda.linearcollider.org/event/6787/contributions/33415/attachments/27509/41775/IsoLep_HLRec2016.pdf.

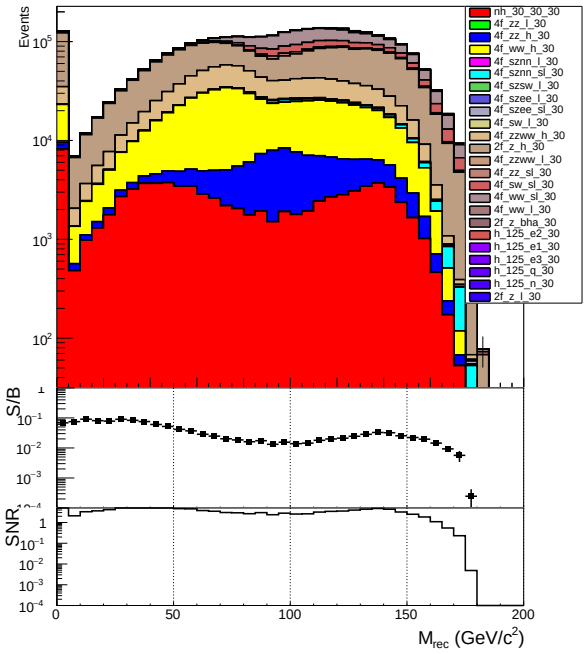
A. Appendix



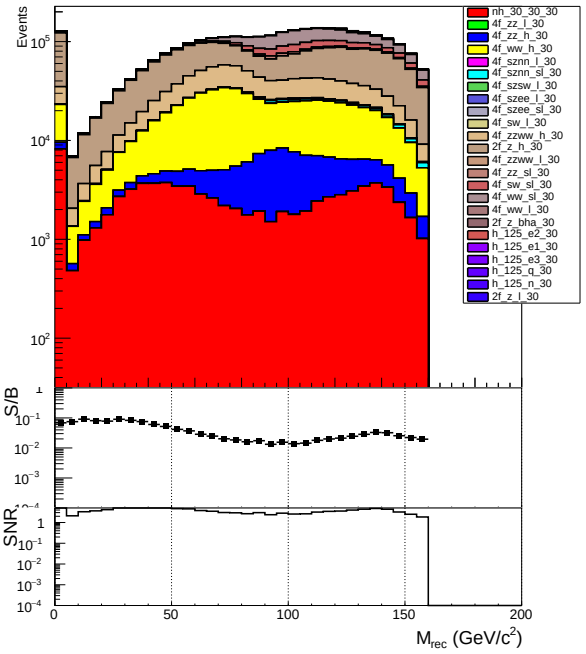
(a) Before any cuts.



(b) After preselection. No $WW/ZZ \to q\bar{q}q\bar{q}$ rejection or m_{rec} cut.

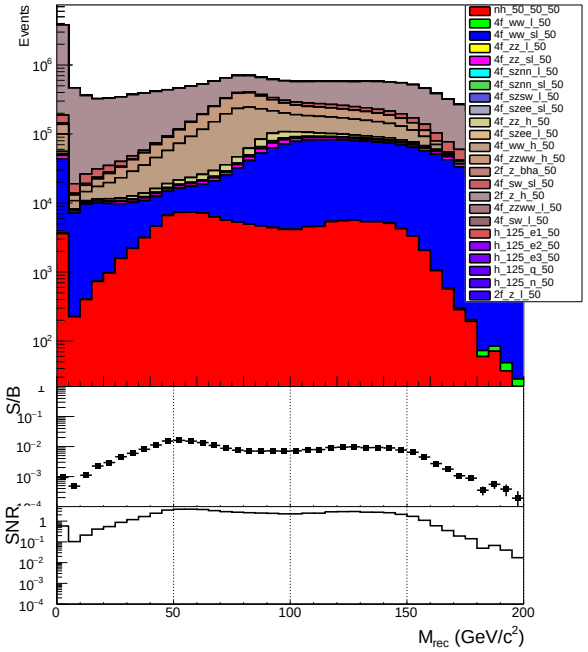


(c) After preselection and $WW/ZZ \to q\bar{q}q\bar{q}$ rejection, no m_{rec} cut applied.

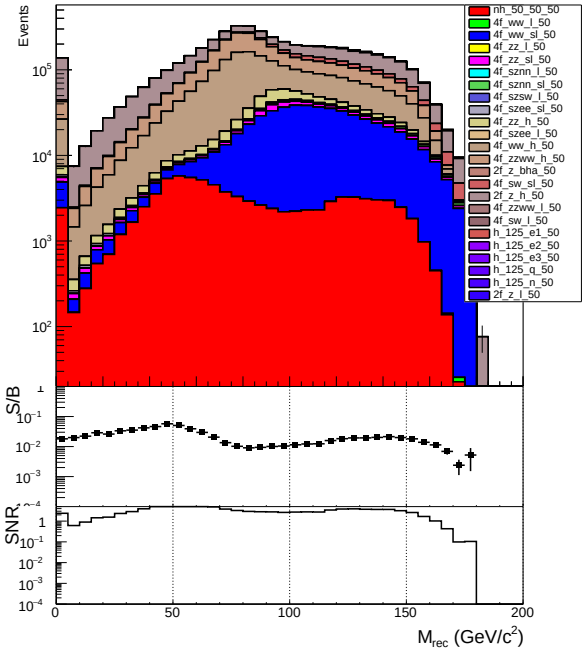


(d) After all cuts applied.

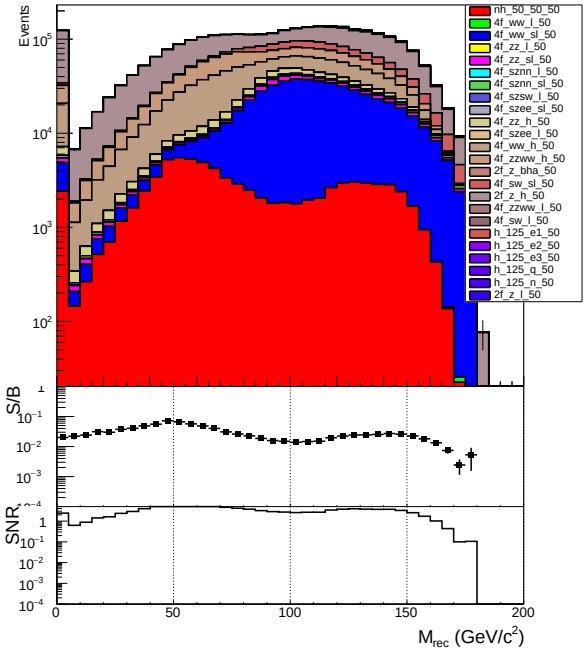
Figure 5: Recoil mass spectra for an additional Higgs boson signal with $m_H = 30$GeV at the different selection stages. The signal is coloured in red. Below the plot with the weighted number of events the signal to background ratio (S/B) and the signal-to-noise ratio (SNR=S/ΔB) are shown.



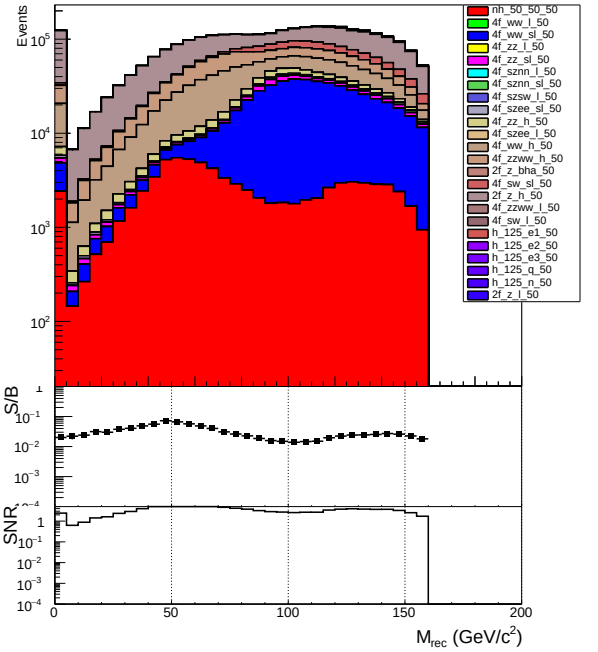
(a) Before any cuts.



(b) After preselection. No $WW/ZZ \to q\bar{q}q\bar{q}$ rejection or m_{rec} cut.

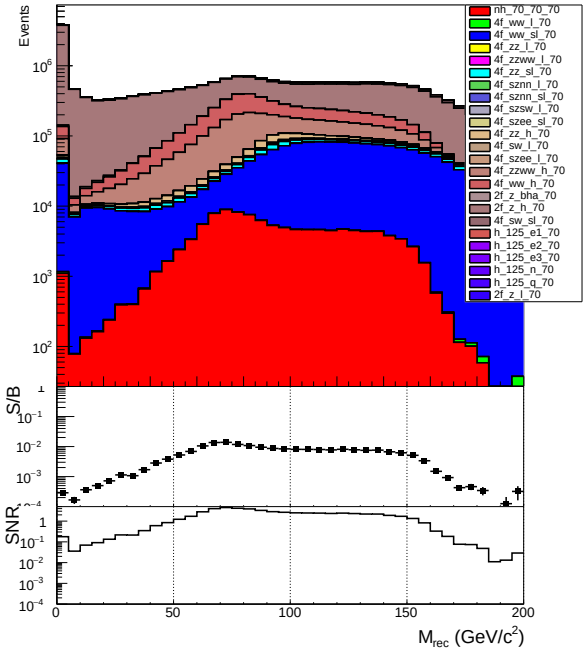


(c) After preselection and $WW/ZZ \to q\bar{q}q\bar{q}$ rejection, no m_{rec} cut applied.

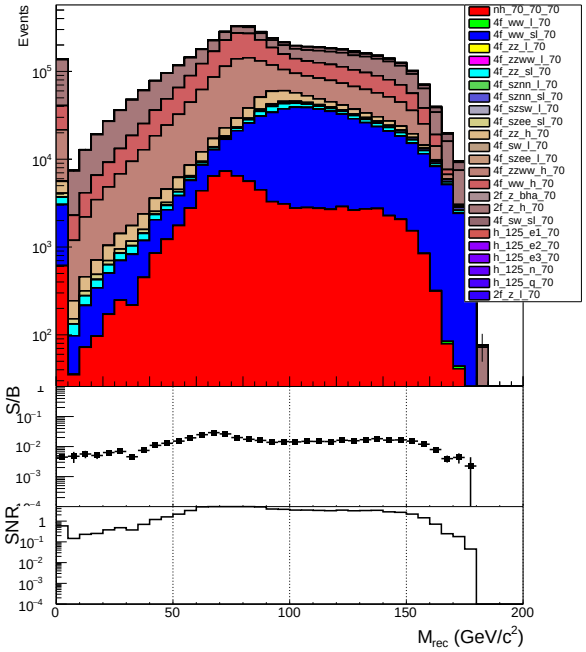


(d) After all cuts applied.

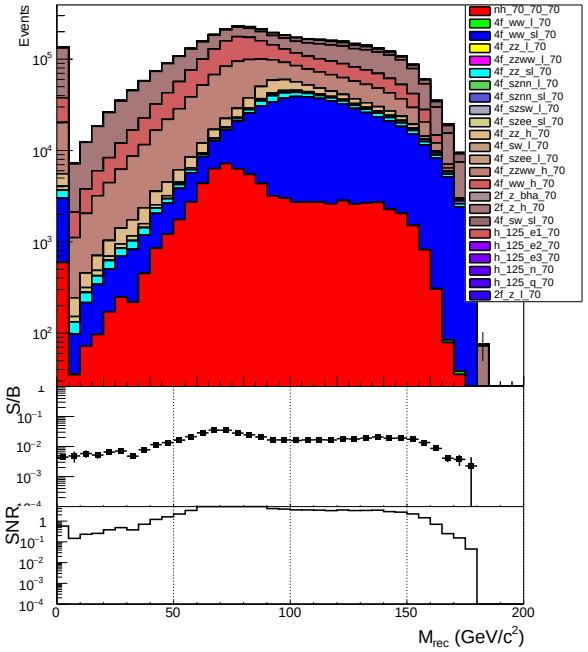
Figure 6: Recoil mass spectra for an additional Higgs boson signal with $m_H = 50$GeV at the different selection stages. The signal is coloured in red. Below the plot with the weighted number of events the signal to background ratio (S/B) and the signal-to-noise ratio (SNR=S/ΔB) are shown.



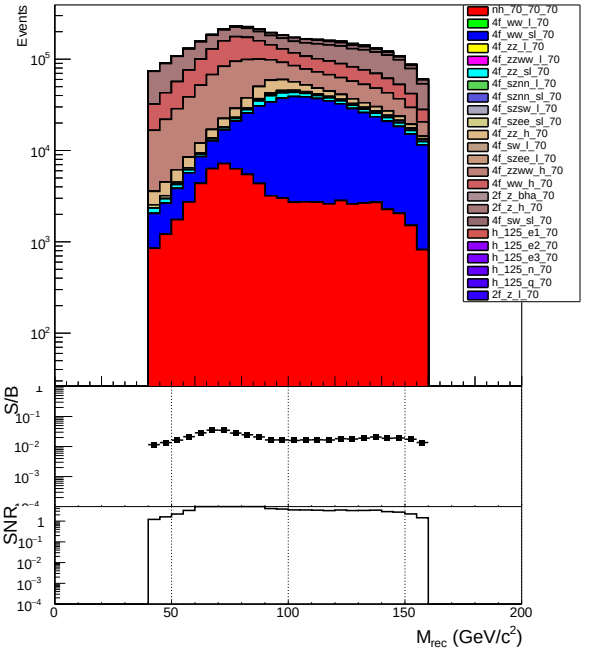
(a) Before any cuts.



(b) After preselection. No $WW/ZZ \to q\bar{q}q\bar{q}$ rejection or m_{rec} cut.

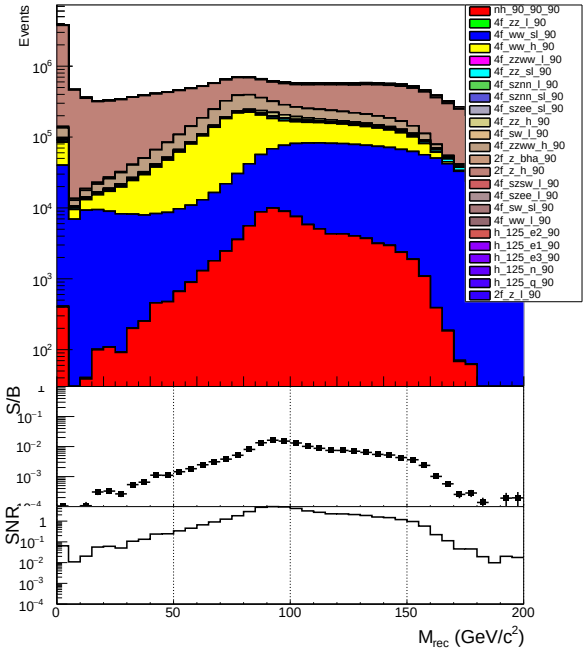


(c) After preselection and $WW/ZZ \to q\bar{q}q\bar{q}$ rejection, no m_{rec} cut applied.

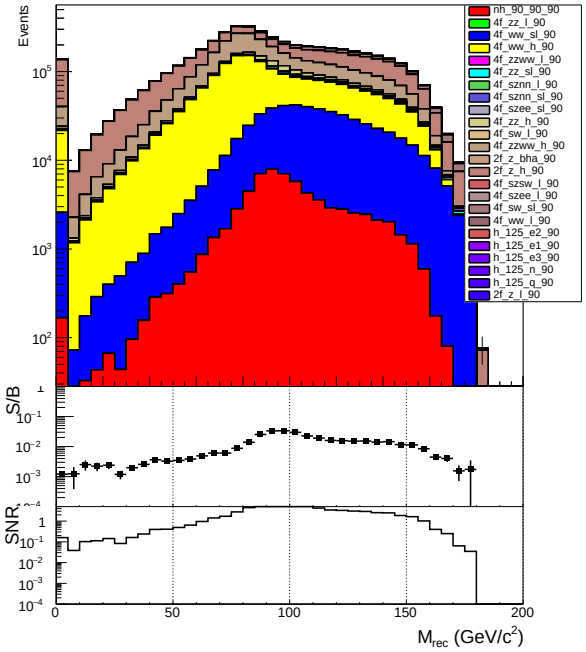


(d) After all cuts applied.

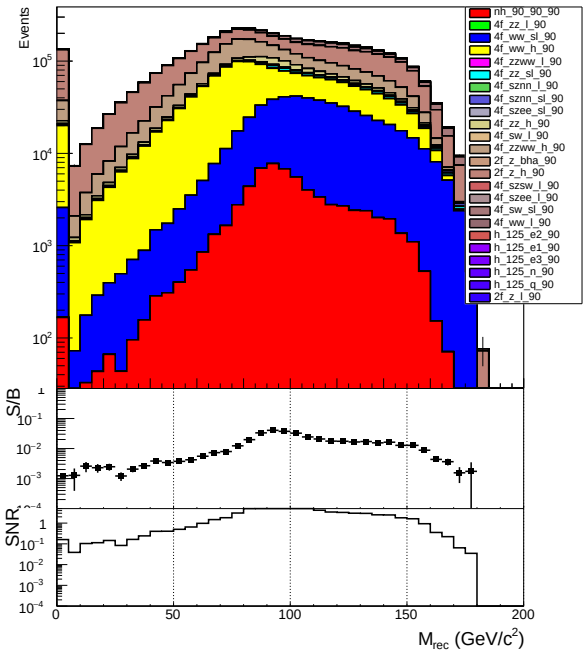
Figure 7: Recoil mass spectra for an additional Higgs boson signal with $m_H = 70$GeV at the different selection stages. The signal is coloured in red. Below the plot with the weighted number of events the signal to background ratio (S/B) and the signal-to-noise ratio (SNR=S/ΔB) are shown.



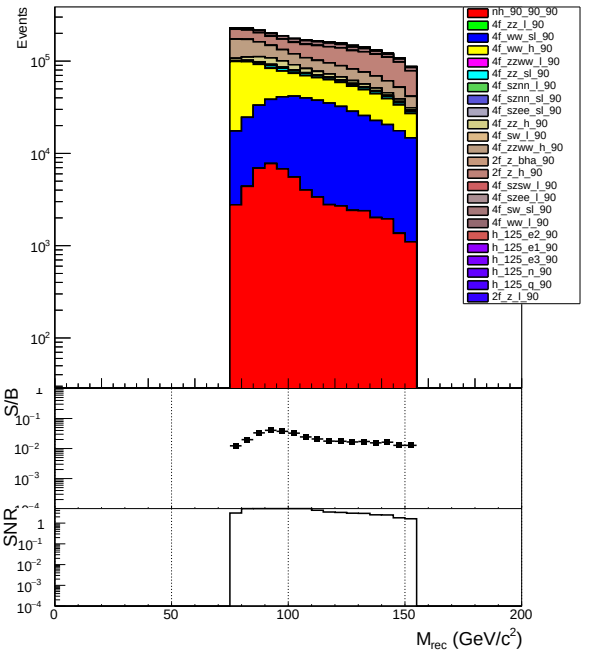
(a) Before any cuts.



(b) After preselection. No $WW/ZZ \to q\bar{q}q\bar{q}$ rejection or m_{rec} cut.

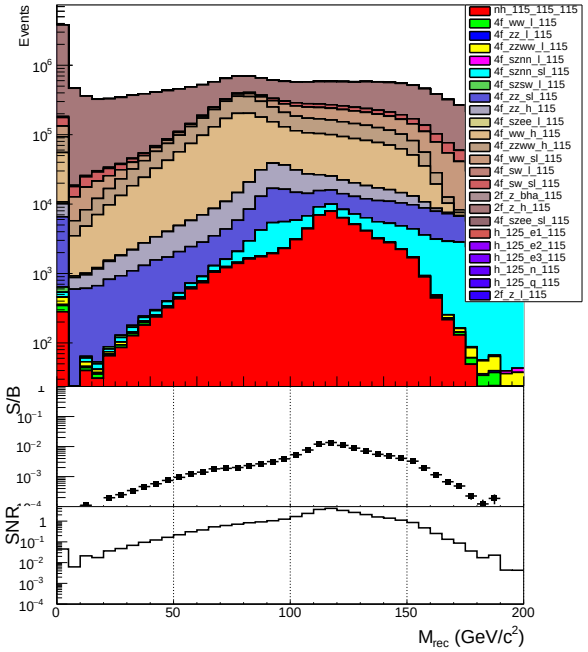


(c) After preselection and $WW/ZZ \to q\bar{q}q\bar{q}$ rejection, no m_{rec} cut applied.

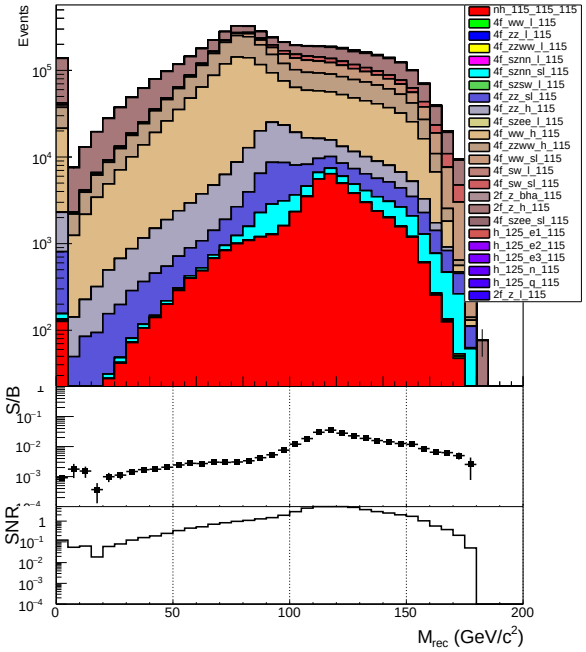


(d) After all cuts applied.

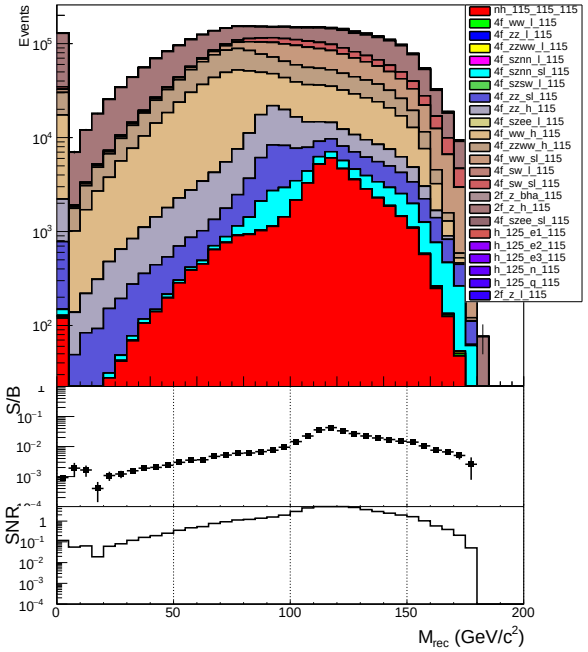
Figure 8: Recoil mass spectra for an additional Higgs boson signal with $m_H = 90$GeV at the different selection stages. The signal is coloured in red. Below the plot with the weighted number of events the signal to background ratio (S/B) and the signal-to-noise ratio (SNR=S/ΔB) are shown.



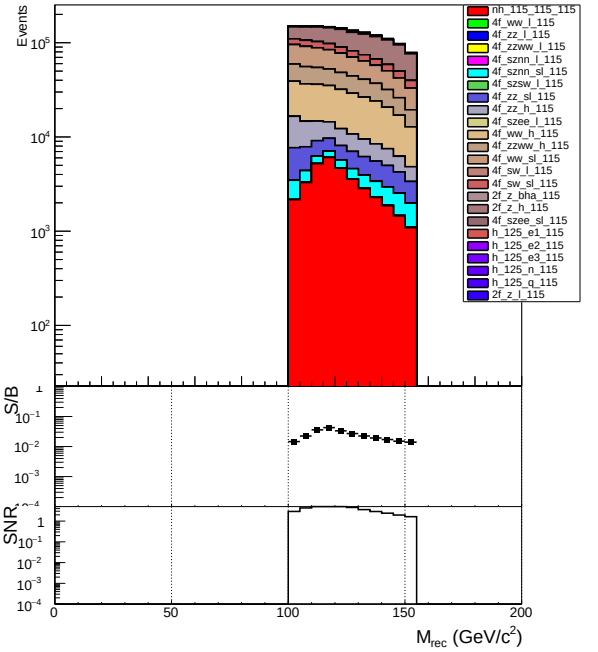
(a) Before any cuts.



(b) After preselection. No $WW/ZZ \to q\bar{q}q\bar{q}$ rejection or m_{rec} cut.

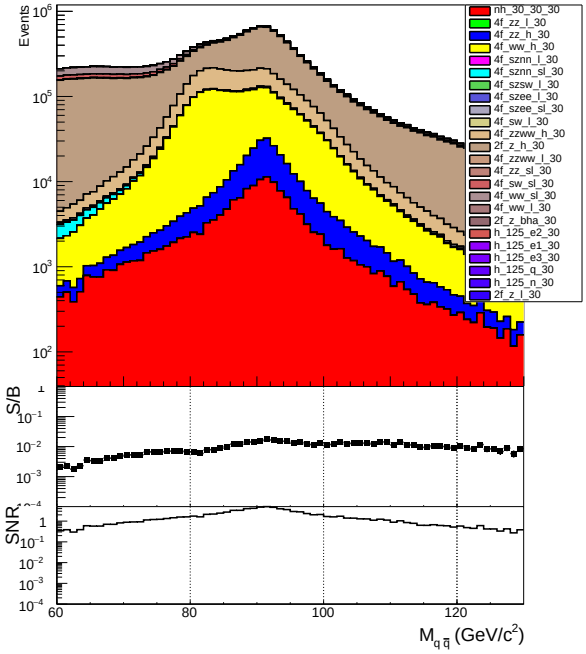


(c) After preselection and $WW/ZZ \to q\bar{q}q\bar{q}$ rejection, no m_{rec} cut applied.

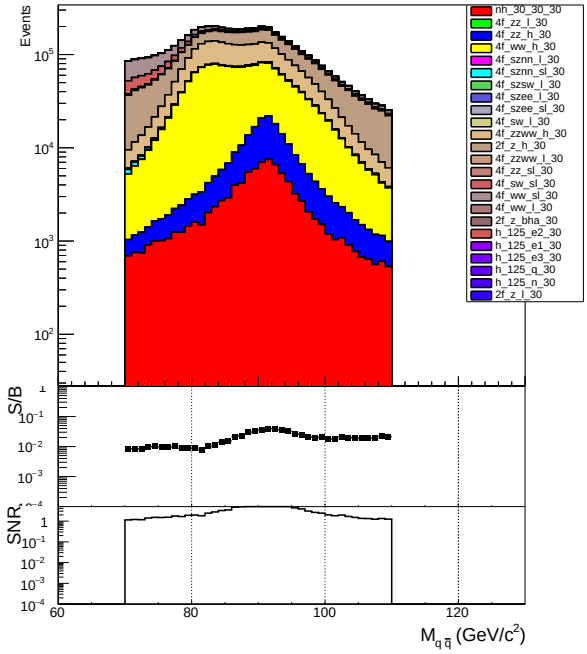


(d) After all cuts applied.

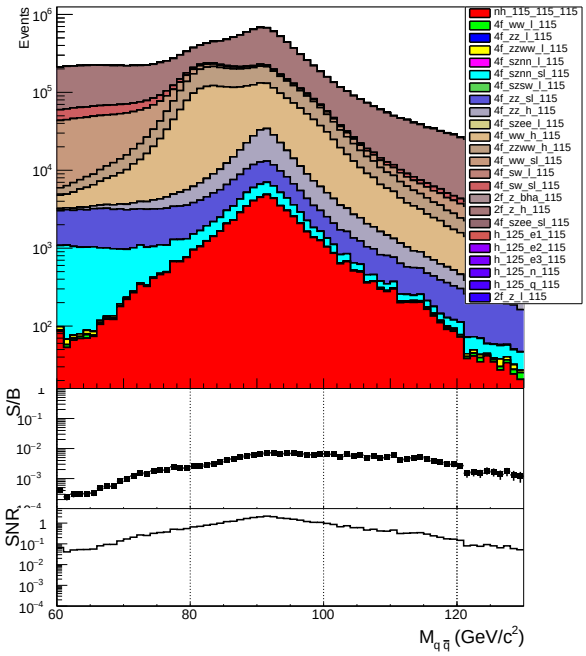
Figure 9: Recoil mass spectra for an additional Higgs boson signal with $m_H = 115$GeV at the different selection stages. The signal is coloured in red. Below the plot with the weighted number of events the signal to background ratio (S/B) and the signal-to-noise ratio (SNR=S/ΔB) are shown.



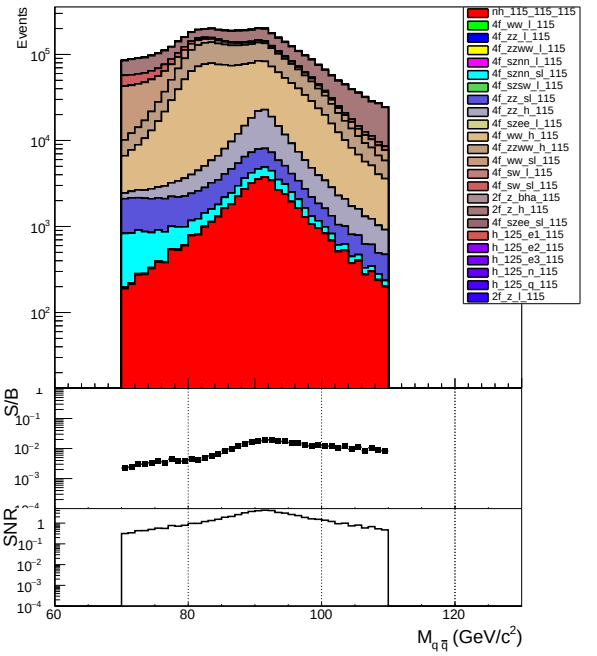
(a) $m_H = 30$GeV. Before any cuts.



(b) $m_H = 30$GeV. After preselection. No $WW/ZZ \to q\bar{q}q\bar{q}$ rejection or m_{rec} cut.

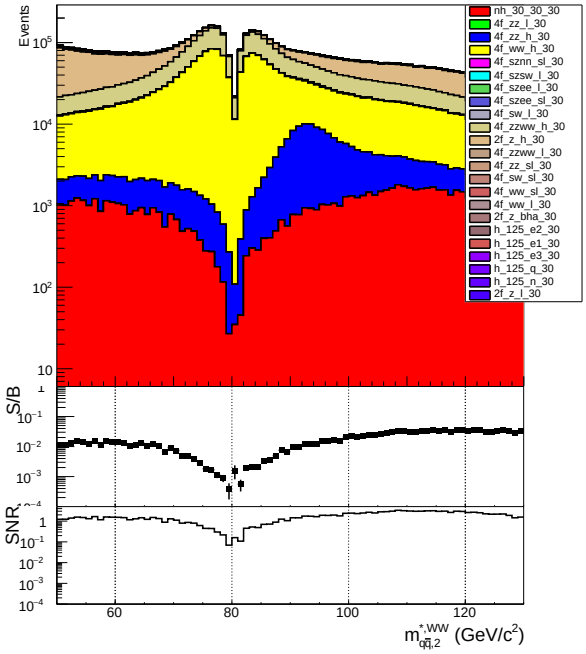


(c) $m_H = 115$GeV. Before any cuts.

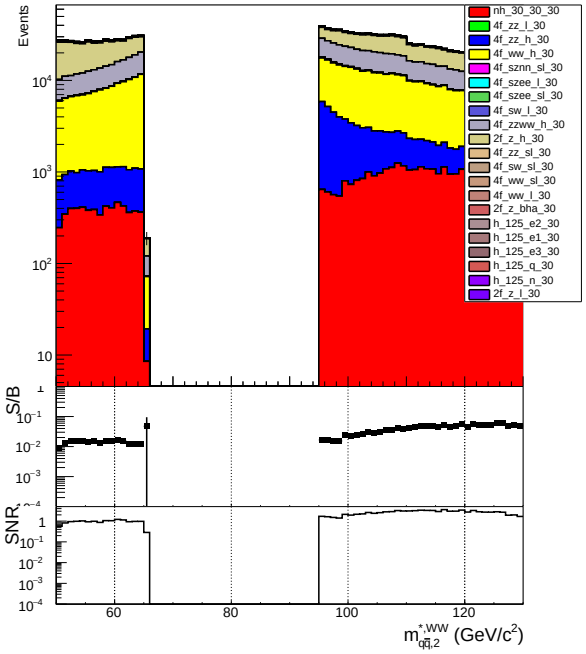


(d) $m_H = 115$GeV. After preselection. No $WW/ZZ \to q\bar{q}q\bar{q}$ rejection or m_{rec} cut.

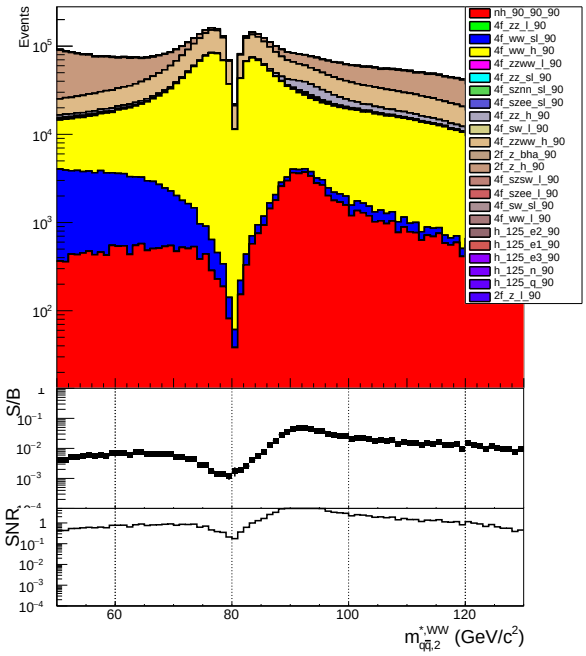
Figure 10: Z candidate pair mass spectra for an additional Higgs boson signal with $m_H = 30$GeV and $m_H = 115$GeV, before and after preselection, including signal-to-background ratio (S/B) and signal-to-noise ratio (SNR) . The signal is coloured in red.



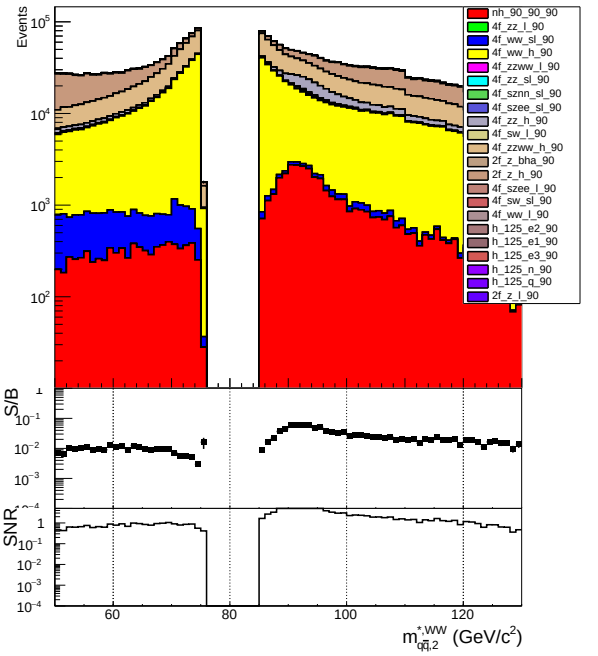
(a) $m_H = 30$GeV. After preselection. No $WW/ZZ \to q\bar{q}q\bar{q}$ rejection.



(b) $m_H = 30$GeV. After preselection and $WW/ZZ \to q\bar{q}q\bar{q}$ rejection.

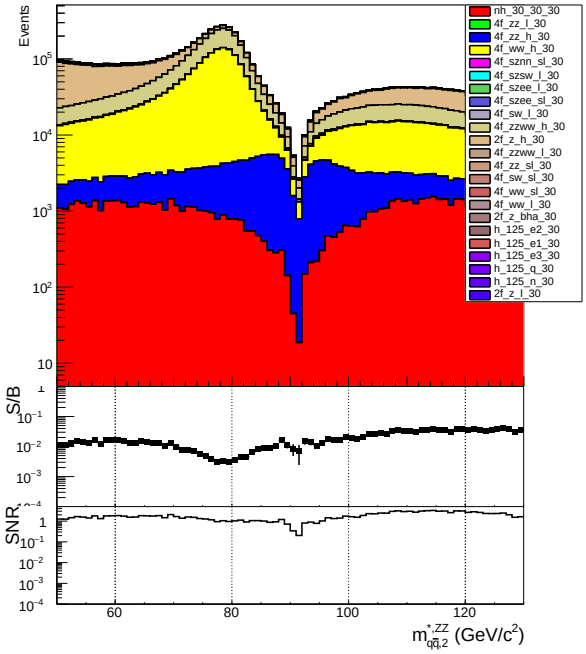


(c) $m_H = 90$GeV. After preselection. No $WW/ZZ \to q\bar{q}q\bar{q}$ rejection.

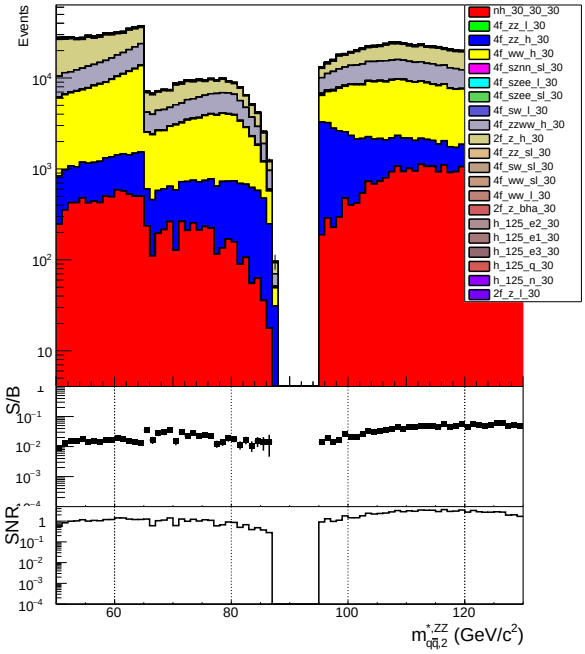


(d) $m_H = 90$GeV. After preselection and $WW/ZZ \to q\bar{q}q\bar{q}$ rejection.

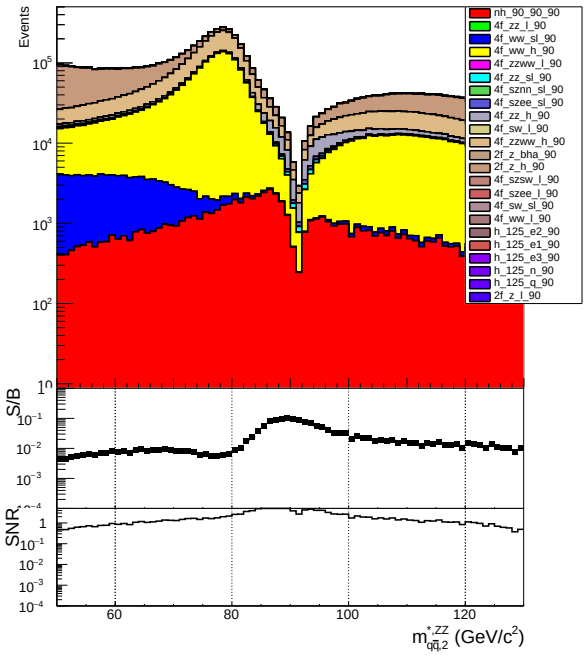
Figure 11: 4-jet topology is clustered to two 2-jets such that both have mass as close as possible to m_W. Shown here are spectra for the 2-jet mass of the both which is further apart from m_W. Cut on this is almost identical to $WW \to q\bar{q}q\bar{q}$ rejection and is used for cut analysis. Diplayed for both $m_H = 30$GeV and $m_H = 90$GeV, before and after $WW/ZZ \to q\bar{q}q\bar{q}$ rejection. No m_{rec} cuts are applied.



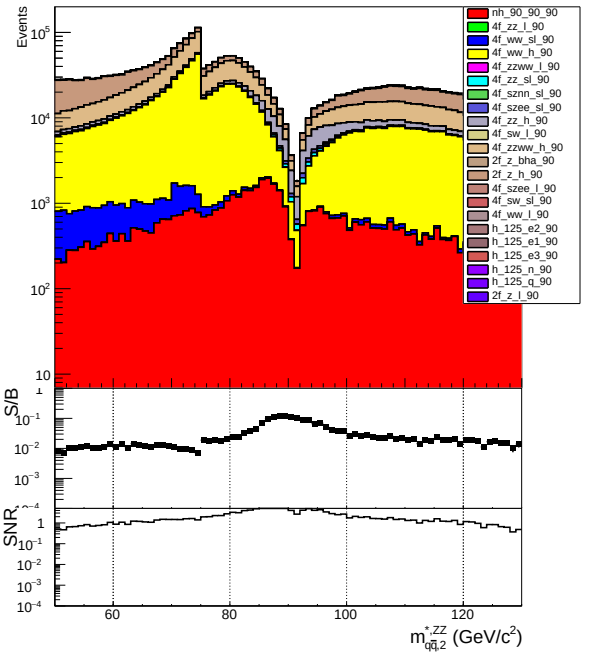
(a) $m_H = 30$GeV. After preselection. No $WW/ZZ \to q\bar{q}q\bar{q}$ rejection.



(b) $m_H = 30$GeV. After preselection and $WW/ZZ \to q\bar{q}q\bar{q}$ rejection.



(c) $m_H = 90$GeV. After preselection. No $WW/ZZ \to q\bar{q}q\bar{q}$ rejection.



(d) $m_H = 90$GeV. After preselection and $WW/ZZ \to q\bar{q}q\bar{q}$ rejection.

Figure 12: 4-jet topology is clustered to two 2-jets such that both have mass as close as possible to m_Z. Shown here are spectra for the 2-jet mass of the both which is further apart from m_Z. Cut on this is almost identical to $ZZ \to q\bar{q}q\bar{q}$ rejection and is used for cut analysis. Diplayed for both $m_H = 30$GeV and $m_H = 90$GeV, before and after $WW/ZZ \to q\bar{q}q\bar{q}$ rejection. No m_{rec} cuts are applied.