Neutrino energy corrections to the reconstruction of the Higgs mass at the International Linear Collider

Sukeerthi Dharani
RWTH Aachen
Germany

Summer Student Programme
Supervisor: Dr. Jenny List

September, 2019

Abstract

A method to add missing neutrino energy to the reconstruction of jets in a Higgsstrahlung process has been developed. The fully hadronic final state of the Higgsstrahlung process is considered in this study. When the missing energy due to the neutrinos could be worked out on a jet-by-jet basis, the jet energy resolution would be improved significantly. As a proof of principle, a correction for the neutrino energy is applied in events with one semileptonic decay of a b-quark.Since the neutrino correction has two solutions, a kinematic fit is employed in order to identify the correct solution. Since the neutrino correction has two solutions, a kinematic fit is employed in order to identify the correct solution.The uncertainties on the jet energy and direction are improved by utilizing the ParticleFlow reconstruction in order to derive an individual covariance matrix for each jet, depending on its composition. Future steps towards a realistic implementation of the neutrino correction and improvements to the error parametrization will be discussed.

Contents

1 Introduction

The Standard model of particle physics is the fundamental theory of particles and their interactions. The standard model comprises the electroweak theory, which describes interactions via the electromagnetic and weak forces, and quantum chromodynamics, the theory of the strong nuclear force. Even-though the Standard Model predicts the behavior of particles accurately, it incorporates only three out of the four fundamental forces i.e. electromagnetism, strong and weak forces but excludes gravity. There are open questions about dark matter, matter to antimatter ratio after the big bang, different mass scales of fermions, the origin of the Higgs boson and its mass etc., which are unanswered by the Standard Model. High energy experiments are needed to solve these mysteries of nature.

The elementary particles acquires mass by interacting with the Higgs field, through Brout-Englert-Higgs mechanism [11]. The quantized excitation of Higgs field is the Higgs boson which was discovered in CERN in 2012 [5, 9]. The Higgs couples to the associated particle to give the particle its mass. Deviation of the Higgs couplings from their SM predicted values, implies physics beyond standard model.

This report will focus on a Higgstrahlung process with a 4-jet final state i.e. $e^-e+ \rightarrow ZH \rightarrow q\bar{q}b\bar{b}$. Higgsstrahlung is one of the prominent Higgs production mechanisms in an electron-positron collider. This analysis is done for the International Large Detector [16] which is one of the detector concepts for in the International Linear Collider (ILC) [6]. ILC is a proposal for a future electron-positron collider with an initial center-of-mass energy of 250 GeV, upgradable to 1 TeV. The Higgs mass can be obtained by reconstructing the recoil mass from the Z boson. This provides lepton collider a unique possibility to measure Higgs parameters independently [6]. A hadron collider has limited sensitivity to reconstruct 4-jet events. Unlike hadron colliders, a lepton collider such as the ILC provides a cleaner initial state to measure fully hadronic final states.

The reconstruction of the Higgs in a 4-jet event is done by a kinematic fit. This study is conducted on events with one semileptonic B decay. Thus, the neutrino energy is found using kinematics of the B hadron′s decay. The energy solutions from the kinematics are then added to the jet which has the neutrino and then the kinematic fit is applied.

This report is organized as follows. Section 2 introduces the ILC and the ILD. The kinematic fit method of reconstructing the Higgs mass is explained in section 3.2.2. The calculation of neutrino energy, its implementation in the kinematic fit and the results are shown in section 3. The implementation of the accurate representation of errors is discussed in section 4. Finally, the outlook of this study is discussed in section 23

2 International Linear Collider

International Linear Collider (ILC) is a planned 200-500 GeV e^+ e^- linear collider and is upgradeable to reach energies up to 1 TeV. ILC is a e^+ e^- linear collider so energy loss due to synchrotron radiation, which is inversely proportional to the fourth power of the mass of the colliding particles, is minimum. Since protons are collided in Large Hadron Collider (LHC), there is a large QCD background due to quark content and gluonic interactions. Colliding e^+ e^- eliminates such background because electrons and positrons are elementary particles. Thus, we have a clean e^+ e^- collision event with clear input.
In this chapter, the layout of ILC and its physics goals are discussed, based on [6, 4]. The figure below shows the layout of ILC. The subsystems are explained in section 2.1 and 2.2.

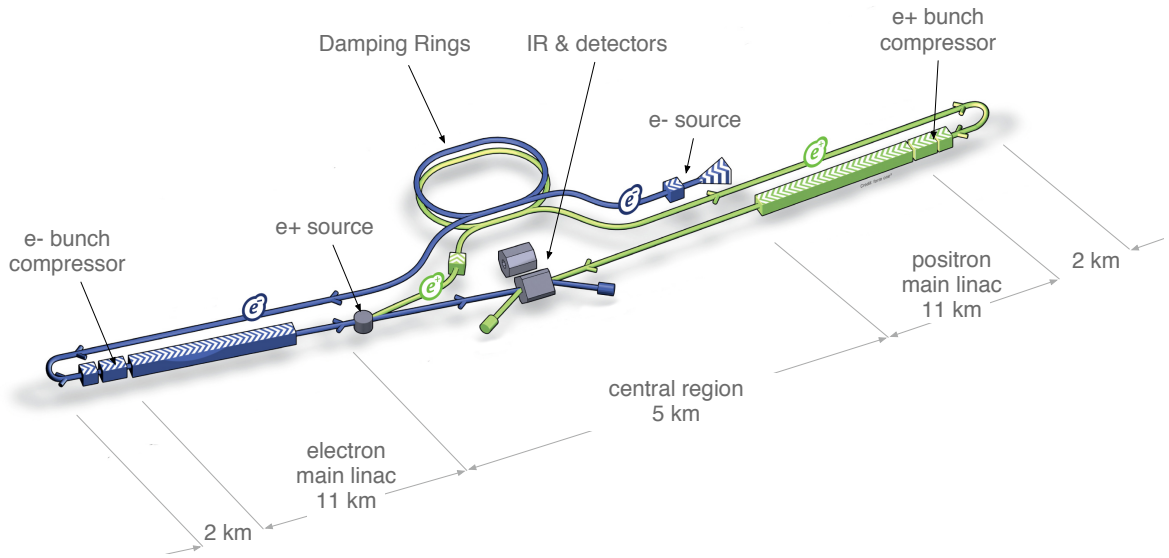


Figure 1: Schematic of ILC with subsystems [2]. Refer to section 2.1 and 2.2 for explanations.

2.1 The Accelerator

The ILC uses 1.3 GHz superconducting radio-frequency (SCRF) accelerating technology. The electrons are produced by 2 nanosecond laser light pulses targeted at a photo-cathode. The electrons are sent through damping rings with a circumference of 3.2 km to reduce the emittances. The beams from the damping rings are then injected into the main linacs to an energy of 5 GeV.
The high energy electrons are then passed through an undulator, producing synchrotron radiation. When the radiation is targeted at titanium alloy, it produces electron positron pairs. Then the positrons are collected and sent through damping rings before accelerated to 5 GeV in separate linac.

For a center-of-mass energy of 500 GeV, each of the linacs is about 11 km long. The two beam-delivery systems, each 2.2 km long, bring the beams into collision. More details can be found in [3].

2.2 The Detectors

There are two detector concepts proposed for the ILC namely the Silicon Detector (SiD) [14] and International Large Detector (ILD)[16].

SiD is a compact detector with a 5 T magnetic field, consisting of a silicon pixel vertex detector, silicon tracking, silicon tungsten electromagnetic calorimetry (ECAL) and highly segmented hadronic calorimetry (HCAL). SiD also incorporates a high-field solenoid, iron flux return, and a muon identification system [15].
SiD enables time-stamping on single bunch crossings to provide robust performance with respect to beam backgrounds or beam loss. The use of silicon sensors in the vertex, tracking and calorimetry enables a unique integrated tracking system ideally suited for the particle flow.

The ILD is based on silicon vertex detectors, a time projection chamber as the central tracker, silicon tungsten electromagnetic calorimetry and steel scintillator hadron calorimetry. The ILD which this study is based, has an axial magnetic field of 3.5 T.
The ILD is optimised for energy and momentum resolution, with flexibility for operation at energies up to the TeV range. The time projection chamber (TPC) in ILD provides continuous tracking for pattern recognition and particle identification. The ILD is explained further in Chapter 3.

The International Large Detector
All particles in an event, neutral or charged, have to be reconstructed individually in a detector. The International Large Detector (ILD) is a detector concept proposed for the ILC which is designed to perform high precision measurements using the particle flow concept to reconstruct multi-jet final states. The particle flow approach is to use the best suited detector to measure the individual particles [1]. Charged particles are reconstructed in the tracker and all neutral particles are determined in the calorimetric system such that jet is reconstructed by reconstructing individual particles. The software that is developed for this purpose is called Particle Flow Algorithm (PFA) [17].

The identification of heavy (charm and bottom) quarks and tau leptons is essential for the ILC physics programme [4]. The reconstruction of decay vertices of short lived particles such as D or B mesons, is accomplished by tracking back the trajectory of the short lived decay products. This requires a very precise vertex detector near the interaction point. The vertex detector (VTX) is a multi-layer pixel-vertex detector, having pure barrel geometry with an inner radius of 16mm which is the distance to the interaction point.

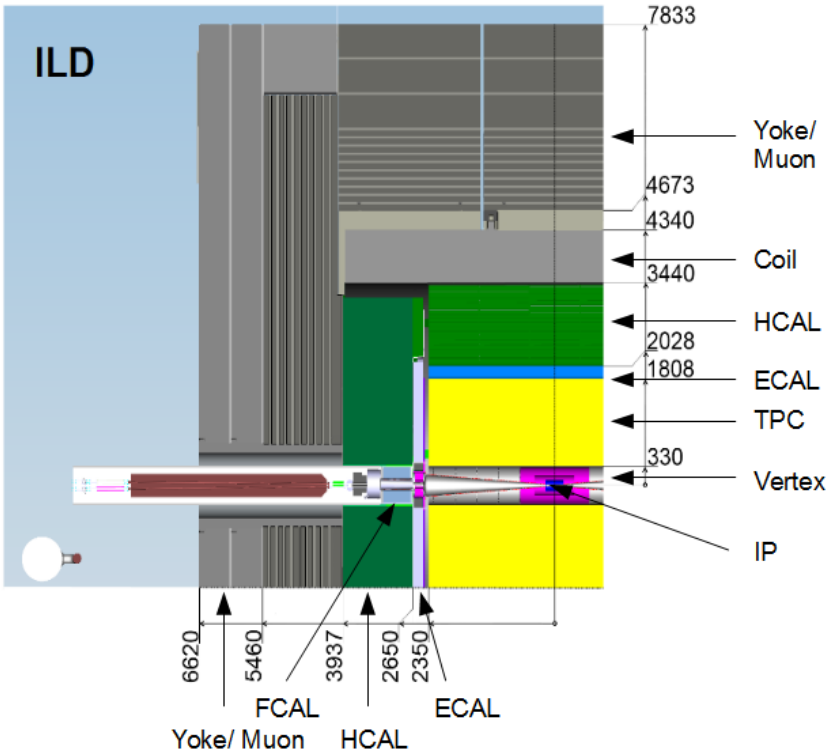


Figure 2: Octant view of the ILD with dimensions of each component.

The silicon pixel detector that surrounds the vertex detector, is called Silicon Inner Tracker (SIT). It consists of two layer of silicon strip detectors positioned between VTX and Time projection chamber (TPC). In the forward region, a system of two silicon-pixel disks and five silicon - strip disks (FTD) provides low angle tracking coverage. A time projection chamber (TPC) is a volume consisting of field cage, which is usually filled with noble gas (90%) and quencher (10%), and end plates. It typically provides a robust tracking with over 200 space points per track. Outside the TPC system, there is a layer of silicon strip detectors called Silicon External Tracker (SET). The SET system enhances the tracking by providing high precision space points and also provides a possibility of time stamping. The ILD concept incorporates two different technology options for both highly segmented calorimeters ECAL and HCAL. The ECAL technologies are a Silicon-Tungsten (SiW) calorimeter and a scintillator-Tungsten calorimeter [16]. The ECAL with an inner radius of 1808 mm, provides up to 30 samples in depth and small transverse cell size, split into a barrel and an end-cap system. The two HCAL technologies under consideration are analogue steel-scintillator hadron calorimeter (AHCAL) [12] and semi-digital calorimeter (DHCAL) [13]. A large volume superconducting coil surrounds the calorimeters, creating an axial B field of 3.5 Tesla. The return yoke acts as muon detector, a tail catcher and returns the magnetic flux of the coils.

2.3 ILC Physics

The ILC is designed to investigate various open questions of particle physics like the mechanism of electroweak symmetry and search for physics beyond the Standard Model like supersymmetry, dark matter, extra dimensions etc.
The major aspects of the ILC physics program [4] are given below.

- precision measurements of the properties of the Higgs boson which are discussed in detail in the next section.
- study of top quark pair production near threshold and at higher energies near the maximum of the cross section for $e^+e^- \rightarrow t\bar{t}$.
- searching for and potentially measuring the properties of new particles predicted by various extensions of the Standard Model which are accessible at the energy range of the ILC. This includes in particular also dark matter particles.

2.4 Higgs production mechanisms

In an electron-positron collider, there are three major Higgs production mechanisms. The first process is a Higgsstrahlung process $e^+e^- \rightarrow ZH$ in which a Higgs recoils against a Z boson. The second process is WW fusion $e^+e^- \rightarrow \nu_e\bar{\nu}_eH$ where two W boson fuse to form a Higgs boson. The third process is ZZ fusion $e^+e^- \rightarrow e^+e^-H$ in which two Z bosons fuses to create a Higgs boson. The Feynman diagrams of these processes are shown below.

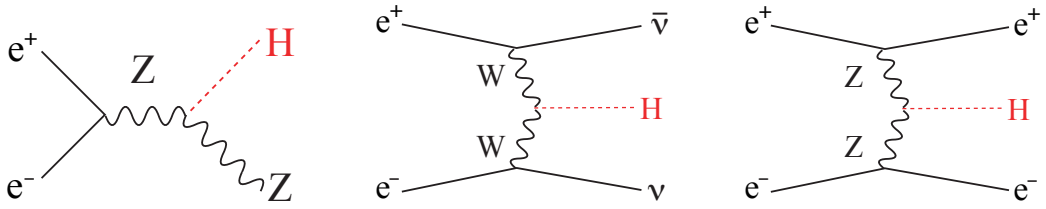


Figure 3: Feynamnn diagrams of the Higgsstrahlung, WW fusion, ZZ fusion [6]

Figure 4 shows the production cross section of the above processes as a function of the center of mass energy with the electron beam polarization of -80% and positron beam polarization of 30%. One can observe that the Higgsstrahlung has by far the largest cross section at the center of mass energy of 250 GeV and at the center of mass energy of 500 GeV, WW fusion is the dominating production mechanism. ZZ fusion is the least contributing Higgs production mechanism in an electron-positron collider.

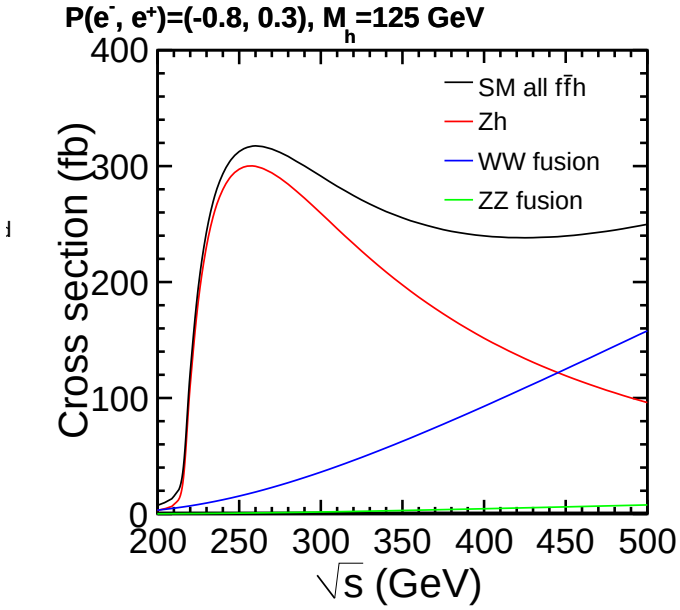


Figure 4: Cross sections of various Higgs production processes as a function of the center-of-mass energy at the ILC [6]

2.4.1 Mass of the Higgs boson

In the Standard Model, the mass of the Higgs boson is related to v and λ like $M_H = \sqrt{2\lambda v^2}$ where v is the vacuum expectation value and λ is Higgs self-coupling parameter. The only parameter unconstrained by the standard model is the Higgs boson's mass. It also plays a crucial role in the SM prediction of Higgs production and decay rates.
The coupling of the Higgs to the particles is described by a set of coupling constants. The coupling to a fermion is proportional to the mass of fermions and the coupling to a boson is proportional to the square of the boson's mass. Since couplings of fermions are proportional to the mass of the fermions, massive particles have higher branching ratios. Hence, $H \to b\bar{b}$ has the maximum branching ratio of 57.8%.

2.4.2 Higgsstrahlung with a fully hadronic final state

As mentioned above, Higgsstrahlung is one of the prominent productions of Higgs boson in small center of mass energies. The figure below shows the Feynman diagram of the Higgsstrahlung process in a fully hadronic decay channel $e^+e^- \to ZH \to q\bar{q}b\bar{b}$.

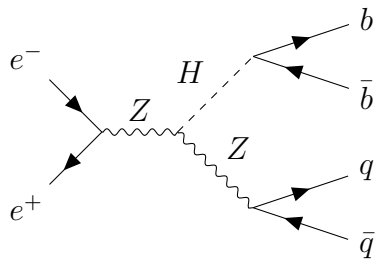


Figure 5: Feynman diagram of a Higgsstrahlung process with a fully hadronic final state at e^+e^- colliders.

This study is conducted on an event sample with the center of mass energy of 500 GeV with a 4-jet final state.

2.4.3 Invariant mass of a pair of jets

As quarks have a net color charge, they cannot exist freely due to color confinement. They immediately hadronize to form color neutral hadrons. This process is called hadronization. A quark can radiate a gluon in the same way a charged particle would emit a photon. Gluons also hadronize to color neutral hadrons, leading to the formation of jets. The outgoing quarks in the final state in Fig. 5 create four jets in the detector. The jet finding algorithm used in this study is an exclusive k_t algorithm which clusters the particles in an event and generates the jet parameters.
In a 4-jet final state, one of the jet pairs is produced by the Higgs boson and the other pair is produced by the Z boson. The invariant mass M_{ij} of two jets i, j is determined as follows:

$$M_{i,j}^2 = (P_i^\mu + P_j^\mu)^2 = m_i^\mu + m_j^\mu + 2\left(E_i E_j - \vec{p_i} \cdot \vec{p_j}\right) \tag{0.1}$$

In order to find the best jet pairing, a χ^2 minimization can be employed. The χ^2 is defined as

$$\chi^2 = \frac{(M_{ij} - M_Z)^2}{\sigma_{M_Z}^2} \tag{0.2}$$

where M_{ij} and M_{kl} are the invariant masses of jet pairs i, j and k, l, M_Z is the expected mass values of the Z boson and σ_{M_Z} is the assumed uncertainties of the mass of Z.
The jet pairs are permuted and the invariant mass of each jets pair is found. Thereby, a jet pairing is chosen such that the χ^2 defined in Eq. 0.2 is minimized. Figure 6 shows the reconstructed Higgs mass distribution from 22937 $ee \to ZH \to q\bar{q}b\bar{b}$ events after a full simulation of the ILD detector.

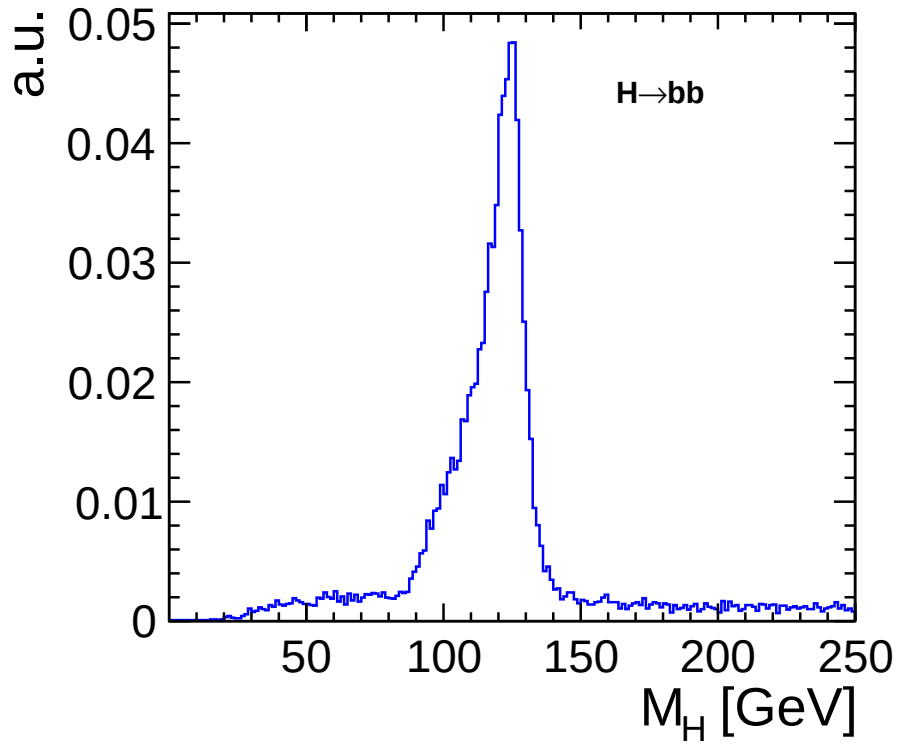


Figure 6: Reconstructed mass of $H \to b\bar{b}$ for the jet pairing giving the minimum χ^2.

Instead of assuming a fixed number of GeV's for σ_{M_Z}, the expected resolution of the Z mass is calculated for each event individually, assuming the following resolutions on the energy and the angles of the jets: $\sigma_{E_{jet}} = 120\%/\sqrt{E_{jet}}$ and $\sigma_{\theta_{jet}}$=0.1 rad, $\sigma_{\theta_{jet}}$=0.1 rad. A more accurate representation of the uncertainties is discussed in detail in section 4.

3 Neutrino Energy Corrections

A semileptonic decay is a weakly interacting hadron decay in which a lepton and its associated neutrino is produced along with one or many hadrons. One of the common bottom quark decays is the charged-current transition into a charm quark, where the W boson decays into two leptons. This process is shown in Fig. 7.

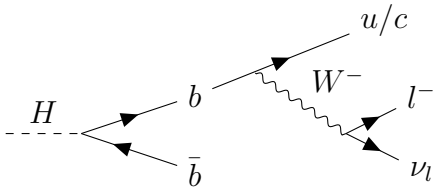
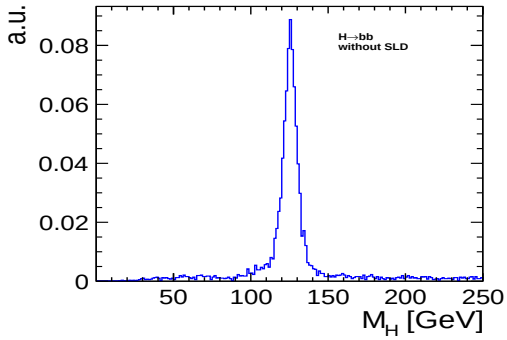
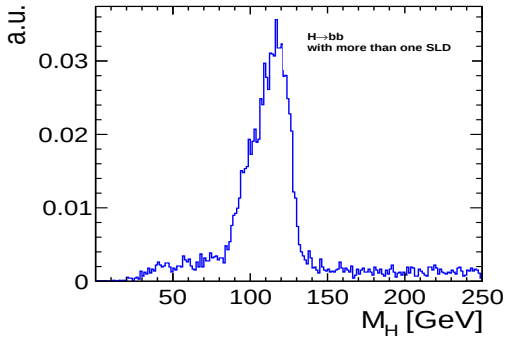


Figure 7: Feynman diagram of the semileptonic decay of a b-quark

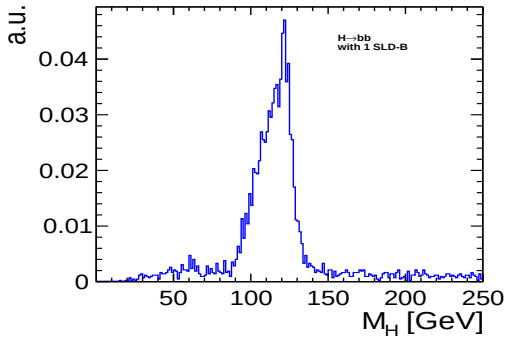
The neutrino produced in a semileptonic decay travels undetected in the detector i.e. the energy carried by the neutrinos is not reconstructed. The Fig.8 shows $H \to b\bar{b}$ with different cases of semileptonic decays.



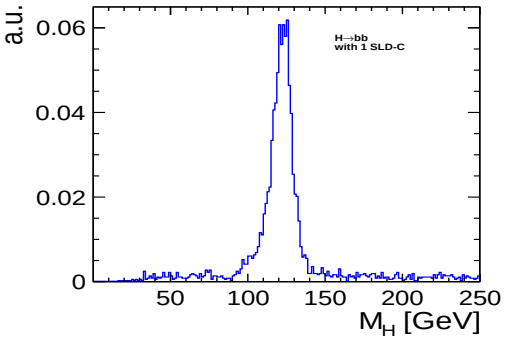
(a) $H \to bb$, without semileptonic decay



(b) $H \to bb$, with more than one semileptonic decay



(c) $H \to bb$, with one semileptonic decay of B



(d) $H \to bb$, with one semileptonic decay of C

Figure 8: Reconstructed mass of $H \to b\bar{b}$ for the jet pairing giving the minimum χ^2.

The shoulder in Fig.6 and Fig. 8b stems from semileptonic decay of B hadrons which is visible from Fig. 8c. The reconstructed energy of a B-jet with a semileptonic decay has less energy due to the missing neutrino energy. This leads to a reduced invariant mass of the jet pair. So, neutrino energy corrections are required to enhance the reconstruction the initial state. Therefore, it is interesting to study if the neutrino momentum can be determined from the decay kinematics. This will be discussed in the next section.

3.1 Neutrino Kinematics

In order to add the missing energy of the neutrino in a jet, one needs to calculate the decay kinematics of the process. In this section, neutrino momenta is calculated from the conservation laws and the momentum of the visible component of the decay.

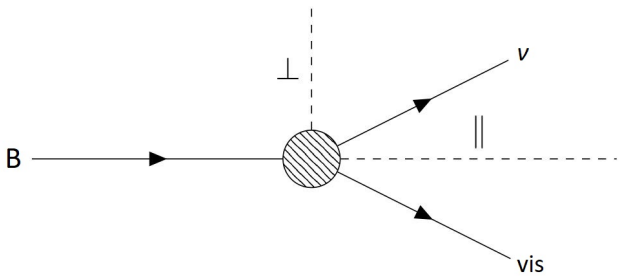


Figure 9: A schematic of a semileptonic decay of the B hadron

In a semileptonic decay of a B, there is a charged lepton in the jet. The charged leptons are identified and reconstructed to an excellent precision in the ILD using dE/dx in the TPC and shower shapes in the highly granular calorimeter and potentially also from its time of flight. It is assumed that the decaying hadron has the mass of a B^0 and its flight direction is given by the connecting from primary to secondary vertex. This can be achieving using excellent vertexing capabilities described in Sec 2.2.
In this study, it is assumed that the knowledge of particles created at the secondary vertex, and particles that belong to the rest of the jet are known from the generator level. In reality, it is relatively easier to find the vertex for the charged particle using the vertex detector but difficult for neutral particles.
In this study, the direction of the lepton which decayed from the semileptonic decay of the B hadron is found from the generator level.
Since neutrinos are massless:

$$m_\nu = 0 \tag{0.3}$$

The energy and momentum conservation leads to the following:

$$E_B = E_{vis} + E_\nu \tag{0.4}$$

$$p_{\nu,\perp} = -p_{vis,\perp} \tag{0.5}$$

$$E_B = E_{vis} + E_\nu \tag{0.6}$$

$$p_{B,\parallel} = p_{\nu,\parallel} + p_{vis,\parallel} \tag{0.7}$$

The solutions for the neutrino momentum is below. For detailed derivation see appendix 7.1

$$p_{\nu,\parallel} = \frac{1}{2D}\left(-A \pm \sqrt{A^2 - BD}\right) \tag{0.8}$$

$$\begin{aligned} \text{where } A &= p_{vis,\parallel}\left(2p_{vis,\perp}^2 + m_{vis}^2 - m_B^2\right) \\ B &= 4p_{vis,\perp}^2 E_{vis}^2 - \left(2p_{vis,\perp}^2 + m_{vis}^2 - m_B^2\right)^2 \\ D &= E_{vis}^2 - p_{vis,\parallel}^2 \end{aligned}$$

There are two solutions for the neutrino energy. In order to determine which energy solution is correct, both the solutions are added to the jet and a kinematic fit is performed. The solution that gives least χ^2 is picked. In principle, a reconstruction using χ^2 method is effective in reproducing the mass of the parent particle but it could also lead to picking wrong jet pairing leading to inefficient neutrino hypothesis. Using all the information available in the whole event, the conservation of energy and momentum can be maintained. This can be done by using a kinematic fit which explained in the Sec. 3.2.

3.2 Kinematic Fit

A kinematic fit is a fitting technique in which the known quantities and the physical laws of an event are set as constraints and the measured quantities are varied according to their measurement uncertainties such that the constraints are fulfilled. In addition, quantities which are not measured can be determined as along as a sufficient number of constraints is available. However this feature is not used in the following. Lagrange multipliers are used along with the χ^2 minimization to fulfill the constraints. In order to perform well and to provide a meaningful hypothesis test, the kinematic fit requires the best possible assessment of the measurement uncertainties.

3.2.1 Constraints

For N measured quantities, a χ^2 is defined as [10]

$$\chi^2(\eta) = (y-\eta)^T V^{-1}(y-\eta) \tag{0.9}$$

where y and η are a N-dimensional vectors of measurements a quantity and fit values of the quantity respectively, V is the covariance matrix.
For a constraint k to be fulfilled, the below relation should hold.

$$f_k(\eta_1, \eta_2, ..., \eta_N) = 0 \tag{0.10}$$

The constraints that one can apply in a Higgsstrahlung process are:
Center of mass energy constraint
Sum of energies of all the jets should add up to 500 GeV i.e. $\sum_{j=1}^{4} E_j = \sqrt{s}$.

$$f_{\sqrt{s}} = \sum_{i=1}^{4} E_i - \sqrt{s} \tag{0.11}$$

This constraint can't be fulfilled in case of the presence of Initial State Radiation (ISR). The ISR photon takes away a part of center of mass energy. The incorporation of ISR in a kinematic fit is discussed in section 3.2.3

Momentum constraint
The sum of the three momenta of the jets should add up to zero i.e. $\sum_{j=1}^{4} \vec{p_{x}}_j = 0$, $\sum_{j=1}^{4} \vec{p_{y}}_j = 0$,$\sum_{j=1}^{4} \vec{p_{z}}_j = 0$. Since the particles collide with a crossing angle θ_c of 7 mrad, the p_x is boosted to compensate the crossing angle.

$$f_{p_x} = \sum_{i=1}^{4} p_{x,i} - \sin(\theta_c/2)\sqrt{s} \tag{0.12}$$

$$f_{p_y} = 0 \tag{0.13}$$

$$f_{p_z} = 0 \tag{0.14}$$

Z mass constraint
From the previous HEP experiments, the mass of the Z boson is known to a great precision. From the event hypothesis, we know that two of the jets' invariant mass should be equal to mass of Z i.e. M_Z^{true} .

$$f_{M_Z} = M_{i,j} - M_Z^{true} \tag{0.15}$$

In this study, the Z mass is set as a hard constraint which implies that the invariant

mass of a jet pair should be exactly the mass of Z. Since the Z boson has a natural width, the invariant mass of two jets can not be exactly equal to the Z mass. The natural width of Z mass is 2.4952 GeV which is similar to the detector resolution of ILD. In the future, this study can be conducted with Z mass as a soft constraint.

3.2.2 Kinematic fit procedure

After clustering the event into four jets (c.f. Sec. 2.4.3), the kinematic fit is performed. The kinematic fit permutes the four jets and runs the fit hypothesis on 5 combinations. All the constraints are applied to each combination. The kinematic fit varies the four-momenta of the four jets such that all the constraints are fulfilled. The combination of jets which yields the smallest final χ^2 of the fit (or the largest fit probability) is chosen. The invariant mass of this combination of the jet pair is then plotted for H. The plot below shows the Higgs mass distribution with and without the kinematic fit.

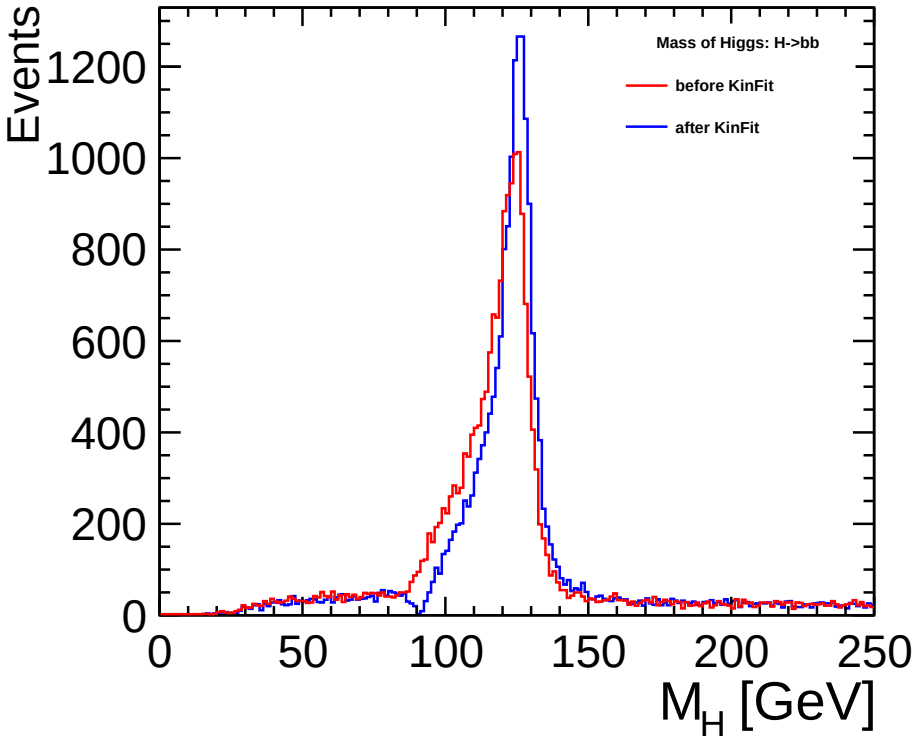
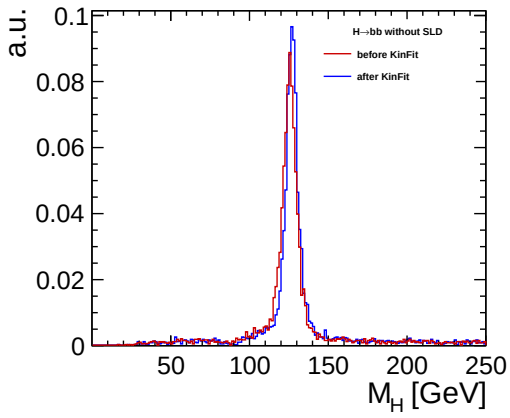
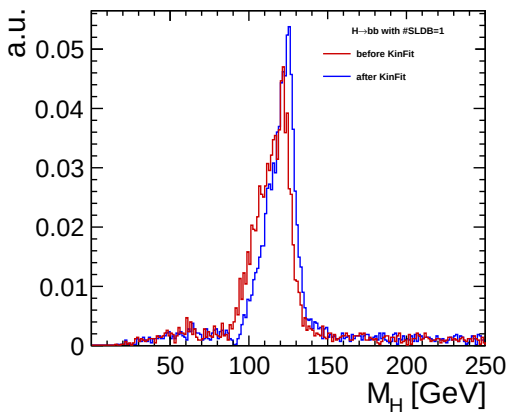


Figure 10: Recontructed Higgs mass before and after kinematic fit for $H \to b\bar{b}$ events (the red histogram is identical to Fig. 6)

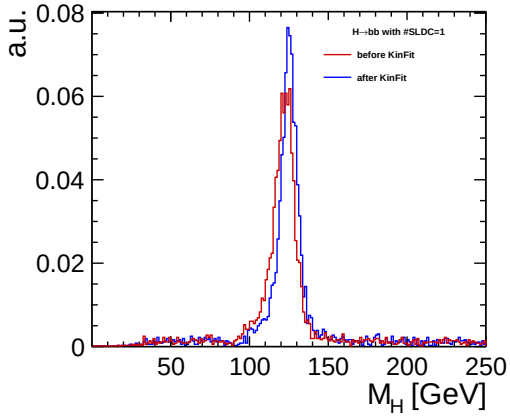
It is seen that the distribution without the kinematic fit is shifted to the left and the mean distribution with the kinematic fit approaches the Higgs mass of $125\,GeV$. The shoulder of the distribution still exists in both cases. It is to be noted that the reconstruction is done for any particle that recoiled against Z, decaying to $b\bar{b}$. This describes the potential



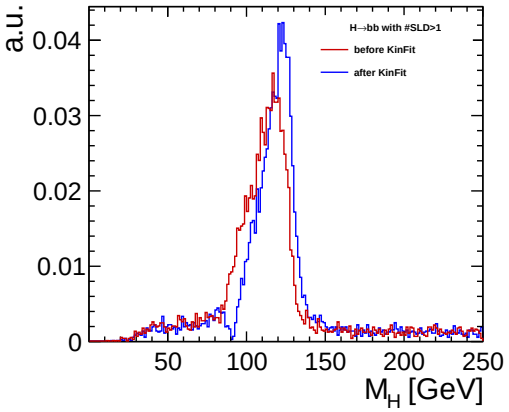
(a) $H \to bb$, without semileptonic decay



(b) $H \to bb$, with one semileptonic decay of B



(c) $H \to bb$, with one semileptonic decay of C



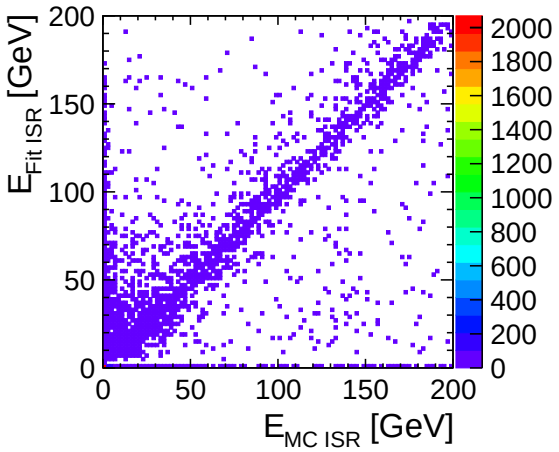
(d) $H \to bb$, with more than one semileptonic decay

Figure 11: Reconstructed Higgs mass for various types of events

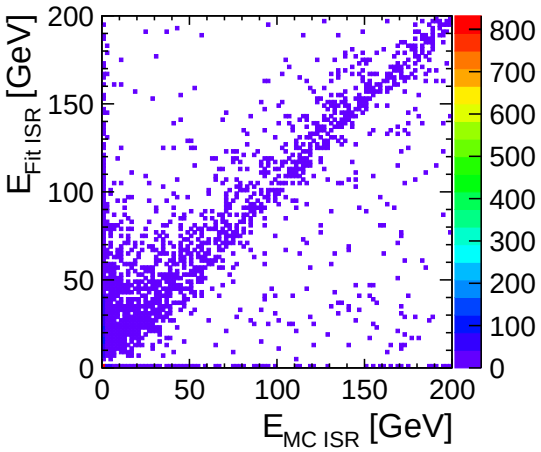
of the kinematic fit. The same procedure is been carried out for various cases of Higgs decays, in order to understand the M_H distribution.
Figure11 shows the reconstructed Higgs mass distribution with and without the kinematic fit for various subcategories in $H \to b\bar{b}$ events.
For events without SLD and events with one SLD-C, the kinematic fit performs comparatively well. For events with one SLD-B and events with more than one SLD, the shoulder around 115 GeV is not removed and also there is a dip around 90 GeV which is due to the incorrect picking of jet pairings by the kinematic fit. Missing energy in the SLD-B jet leads to invariant mass of the jet pair closer to M_Z. The kinematic fit prefers to vary jet momenta such that Z mass hard constraint, leading to the dip near 90 GeV. The invariant mass of a jet pair which without kinematic fit is closer to 90 GeV, but farther to 90 GeV than another pair, is pushed away from 90 GeV.

3.2.3 Initial State Radiation

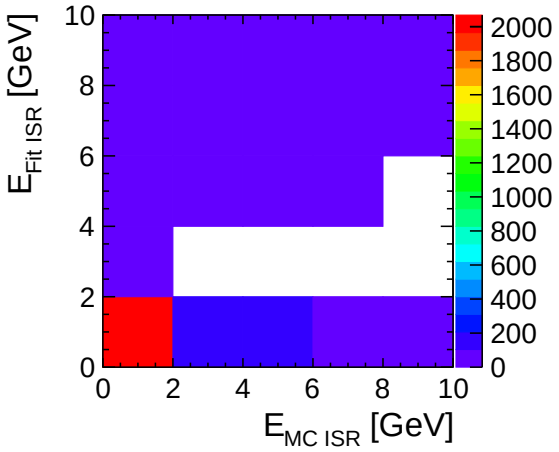
Initial State Radiation (ISR) is a QED effect causing an electron or a positron to emit a photon before the collision and thus escapes from detection through the beam pipe. In this case, the center-of-mass energy available for the hard sub-process is reduced. The direction of this photon is typically very forward. So, the ISR is parameterized by its longitudinal momentum p_z as the free parameter while its transverse component p_t is considered 0. The parametrization uses ISR energy spectrum which follows the power law $E_{ISR} \sim E^{b-1}$ where b is 0.1235 for the center of mass energy of 500 GeV [7]. In the kinematic fit, the residual longitudinal momentum from the reconstructed jets is assigned to the ISR. Section 3.2.2 uses this parametrization of ISR photon.
Figure 14 compares the correlation between fitted and true ISR energy for $H \to b\bar{b}$ events without SLD and with 1 SLD-B.



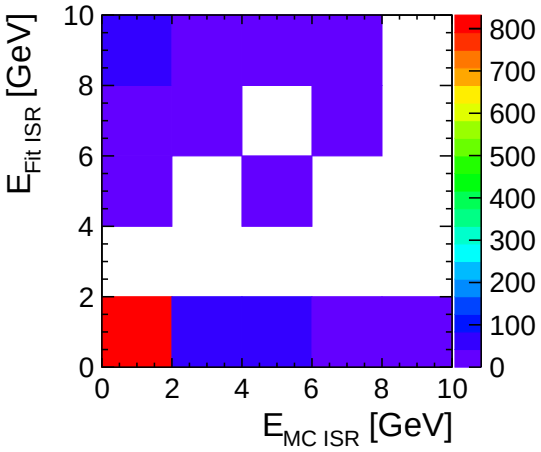
(a) without semileptonic decays



(b) with one semileptonic decay of B meson



(c) without semileptonic decays, axes rescaled to focus on the large number of entries at low energy



(d) with one semileptonic decay of B meson, axes rescaled to focus on the large number of entries at low energy

Figure 12: True energy of ISR vs reconstruction energy from the kinematic fit for $H \to bb$

One can see that ISR energies above $\sim$15 GeV are reconstructed rather well. Thereby the performance of the kinematic fit is slightly better on events without semileptonic decays. Nevertheless there is a significant source for overestimation of the energy of the ISR when it is below $\sim$15 GeV.

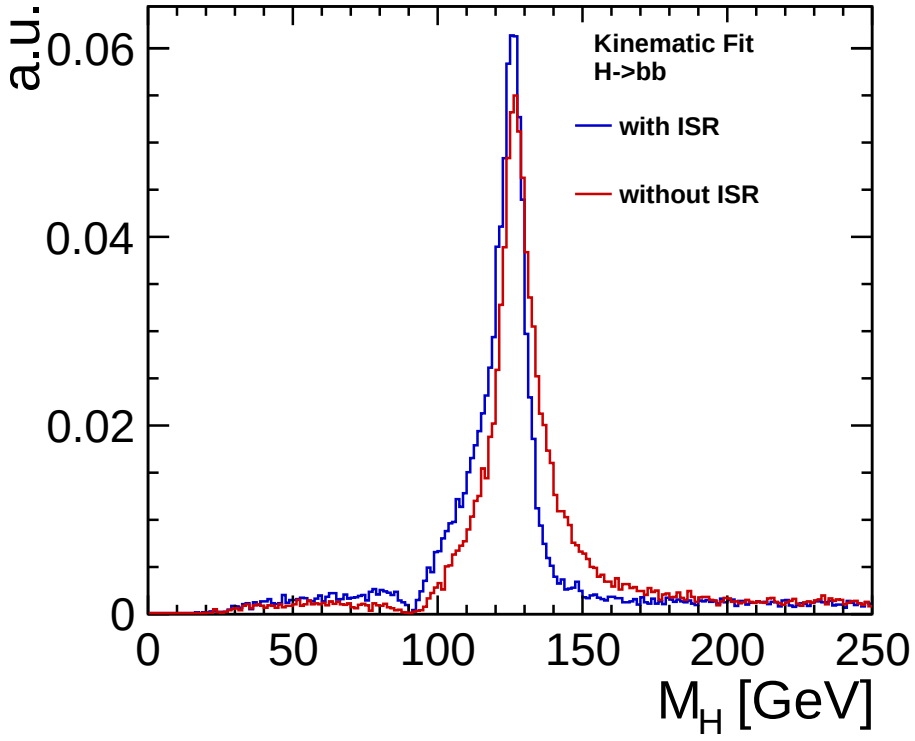
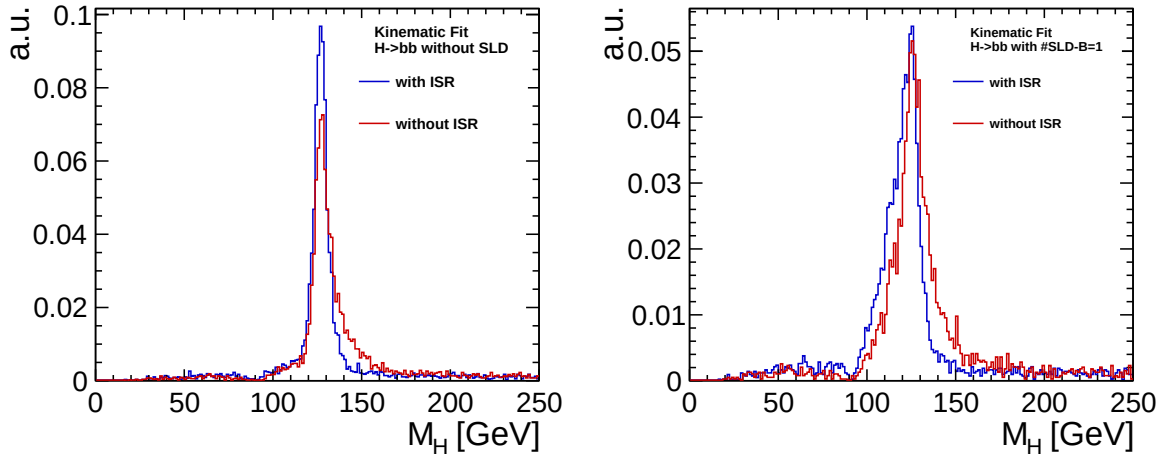


Figure 13: Reconstructed Higgs mass distribution with and without including the ISR parametrization in the fit. The blue histogram is identical to Fig. 10



(a) without semileptonic decays

(b) with one semileptonic decay of B meson

Figure 14: Reconstructed Higgs mass distribution with and without including the ISR parametrization in the fit for $H \to bb$

The effect of the ISR hypothesis in the reconstruction of the Higgs mass by the kinematic fit can be observed in Fig. 13. Applying the ISR hypothesis, the distribution is centered around 125 GeV but still has the shoulder around 115 GeV. This can be improved with the addition of neutrino energy corrections which is evident from Fig. 14. In this report, only one direction of ISR photon is considered. In the future, one can also try both directions of photon and also no ISR photon hypothesis and choose the best of the three cases.

3.3 Application of the neutrino correction

In this section, the kinematic fit will used to implement the neutrino energy correction by selecting the more probable one of the two solutions. The performance of the correction will be evaluated under the idealized conditions discussed in Sec. 3.1 as a proof of principle. The neutrino correction is implemented in the following way: First, the jet with the closest direction to the lepton is identified and the neutrino energy solution (c.f. Eq. 0.8) is added. In order to find the correct solution, both the solutions are tested with the help of the kinematic fit. The solution that yields the smaller χ^2 of the kinematic fit is chosen as the correct solution.

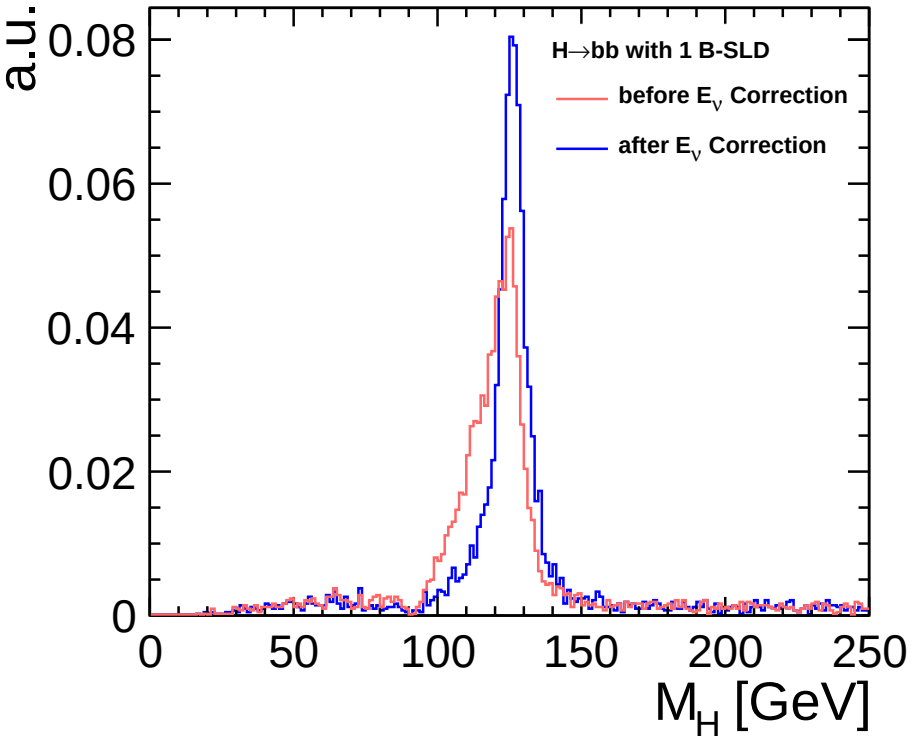


Figure 15: The reconstructed Higgs mass distribution before and after the missing neutrino energy correction, for events with exactly one semileptonic B decay

Figure 15 compares the distributions of the reconstructed Higgs mass in events with exactly one semileptonic B decay before and after applying the neutrino correction. A drastic improvement can be seen. The distribution after the E_ν correction is well-defined and the shoulder around 115 GeV is significantly reduced.

3.4 Fit evaluation

A pull quantifies how much the fit has to adjust the observables away from their measured values in relation to the assumed measurement uncertainties. The pull is defined as:

$$Pull = \frac{x_{fit} - x_{meas}}{\sigma_x} \tag{0.16}$$

Ideally, the pull distribution is centered at 0 and has a width of 1. A biased fit will have the peak shifted away from 0. If the assumed measurement uncertainties are too large, the pull distribution will be too narrow, and when the uncertainties are underestimated, it will be too wide.

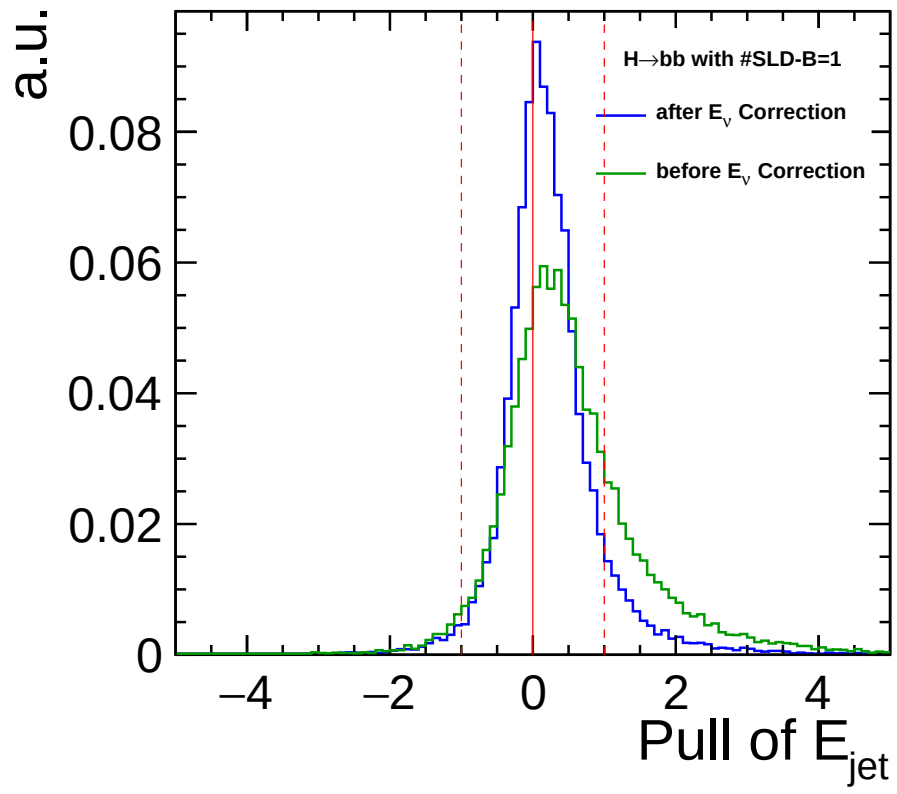


Figure 16: Pull distribution of energy of jets before and after the missing neutrino energy correction

Figure 16 shows the pull distribution of the jet energies for event with one semileptonic B decay with and without the neutrino correction. In the plot, one can observe that the pull on jet energy without the neutrino energy correction is shifted to the right and is positively biased. This means that the kinematic fit constantly adds energy to the jets. This is explained by the fact that kinematic fit has to add energy to the jet to compensate the missing energy of the neutrinos, in order to fulfill the energy conservation constraint.

Adding the neutrino energy corrections, the positive pull has considerably reduced but the peak is narrow with $\sigma = 0.465$ which is found by fitting the distribution using a Gaussian function. This tells us that the error assumed in the fit hypothesis is too large.

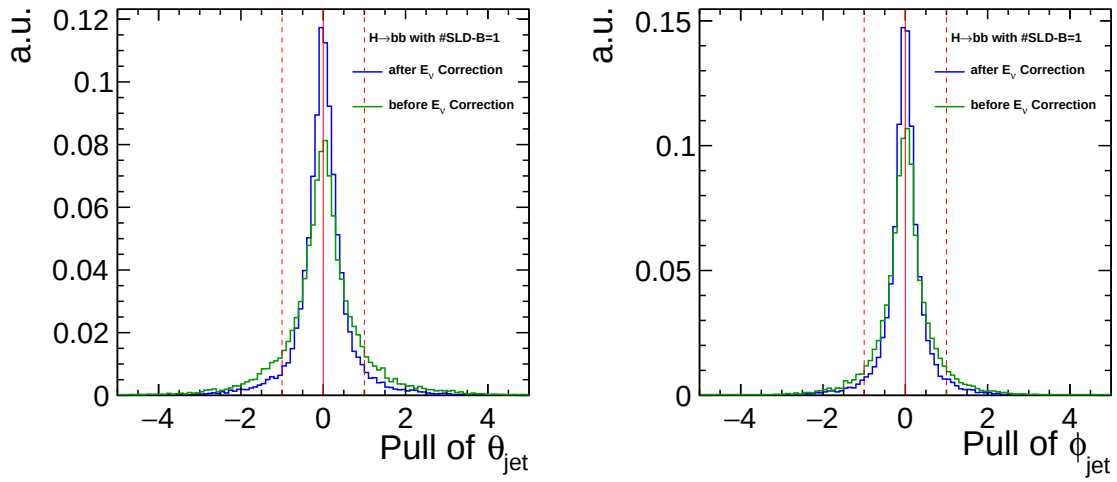


Figure 17: Pull distribution of jet direction before and after the missing neutrino energy corrections

The pull of the jet angles is not biased but the width of the distribution is too narrow. This indicates that a more accurate modeling of the jet uncertainties is required. This will be discussed in Sec 4.

For larger uncertainties, χ^2 decreases and fit probability increases. For small uncertainties, χ^2 increases and P has small value close to 0. This could also mean that the fit hypothesis is incorrect. For ideal case, one would expect a flat distribution.

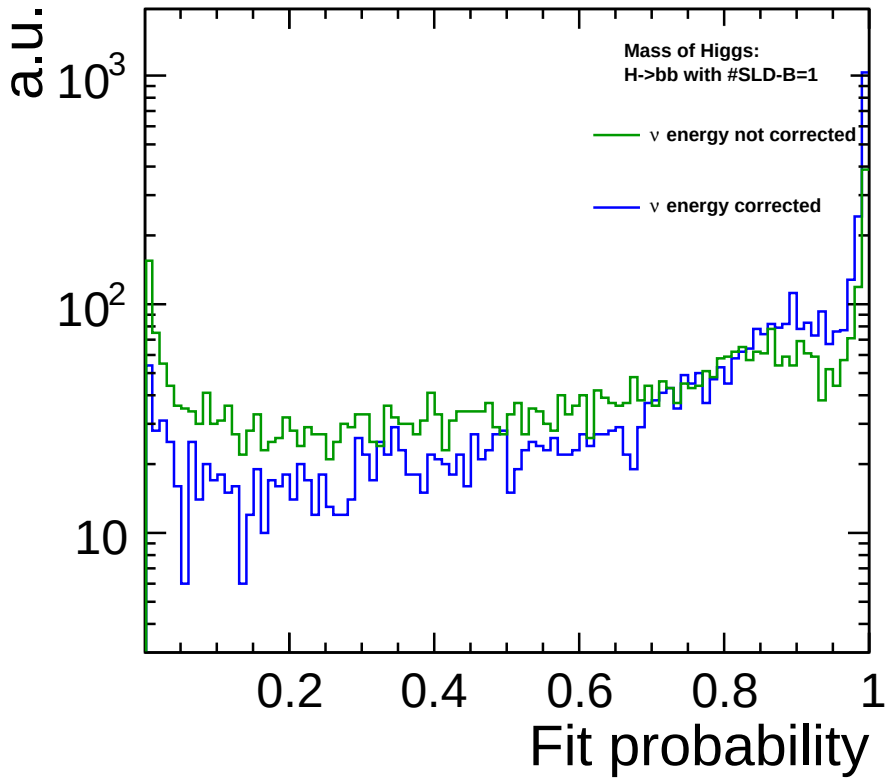


Figure 18: Fit probability of $H \to bb$ with one semileptonic B decay with and without the missing neutrino energy corrections

The number of entries at very low probabilities reduces with the neutrino correction, which is seen in Fig. 18. At the same time, there are too many entries at large fit probabilities, which indicates that with the neutrino correction in place, the jet energy resolution of $120\%/\sqrt{E}$ which was used in the kinematic fit previously is too large now.

4 Towards an improved uncertainty estimation

In the previous section we saw that the simple error parametrizationused before the neutrino correction overestimates the uncertainties. Previously, the error used on E_{jet} was $120\%/\sqrt{E_{jet}}$ and on θ_{jet} and ϕ_{jet} was 0.1 rad. In this section, ErrorFlow which is the more accurate representation of jet uncertainties, is implemented.

4.1 ErrorFlow

As discussed in Sec 2.2, the ILD detector has been designed as a particle flow detector. This means that the jet energy (and directions) is combined from the momentum measurement of the charged particles and the energy measurement of the neutral particles. This concept can be extended to the measurement uncertainties.
In principle, the measurement uncertainty can be obtained for each individual jet from the measurement errors provided by the track fit for the charged particle and the ECal and HCal resolutions for the neutral particles. In addition, an extra term for errors of the particle flow itself, the so-called confusion, has to be accounted for. This method is called ErrorFlow [10]. The output of the ErrorFlow algorithm is a covariance matrix V of each jet:

$$V = \begin{pmatrix} \sigma_{p_x}^2 & \sigma_{p_x}\sigma_{p_y} & \sigma_{p_x}\sigma_{p_z} & \sigma_{p_x}\sigma_E \\ \sigma_{p_y}\sigma_{p_x} & \sigma_{p_y}^2 & \sigma_{p_y}\sigma_{p_z} & \sigma_{p_y}\sigma_E \\ \sigma_{p_z}\sigma_{p_x} & \sigma_{p_z}\sigma_{p_y} & \sigma_{p_z}^2 & \sigma_{p_z}\sigma_E \\ \sigma_E\sigma_{p_x} & \sigma_E\sigma_{p_y} & \sigma_E\sigma_{p_z} & \sigma_E^2 \end{pmatrix} \tag{0.17}$$

4.2 Implementation of the ErrorFlow into the fit

In order to use the ErrorFlow in the kinematic fit, the uncertainties have to be expressed in terms of E, θ, ϕ. While the uncertainty of the jet energy is directly given by the last element of the covariance matrix, the uncertainties on the polar and azimuthal angle had to calculated by error propagation. Thereby, the polar angle θ and the azimuthal angle ϕ are given by:

$$\theta = \arccos\left(\frac{p_z}{\sqrt{p_x^2 + p_y^2 + p_z^2}}\right) = \arccos\left(\frac{p_z}{|\vec{p}|}\right) \tag{0.18}$$

$$\phi = \arctan\left(\frac{p_y}{p_x^2}\right) \tag{0.19}$$

Since θ and ϕ are functions of p_x, p_y and p_z, error propagation is correlated. From this follows:

$$\sigma_\theta = \sigma_{p_x}^2 \left(\frac{\partial\theta}{\partial p_x}\right)^2 + \sigma_{p_y}^2 \left(\frac{\partial\theta}{\partial p_y}\right)^2 + \sigma_{p_z}^2 \left(\frac{\partial\theta}{\partial p_z}\right)^2 + 2\left(\sigma_{p_x P_y}\left(\frac{\partial\theta}{\partial p_x}\right)\left(\frac{\partial\theta}{\partial p_y}\right) + \sigma_{p_y p_z}\left(\frac{\partial\theta}{\partial p_y}\right)\left(\frac{\partial\theta}{\partial p_z}\right) + \sigma_{p_x p_z}\left(\frac{\partial\theta}{\partial p_x}\right)\left(\frac{\partial\theta}{\partial p_z}\right)\right) \tag{0.20}$$

$$\frac{\partial\theta}{\partial p_x} = \frac{p_x p_z}{(|\vec{p}|^2)^{\frac{3}{2}}\sqrt{1-\frac{p_z^2}{|\vec{p}|^2}}} \tag{0.21}$$

$$\frac{\partial\theta}{\partial p_y} = \frac{p_y p_z}{(|\vec{p}|^2)^{\frac{3}{2}}\sqrt{1-\frac{p_z^2}{|\vec{p}|^2}}} \tag{0.22}$$

$$\frac{\partial\theta}{\partial p_z} = \frac{\frac{p_z^2}{(|\vec{p}|^2)^{\frac{3}{2}}} - \frac{1}{|\vec{p}|}}{\sqrt{1-\frac{p_z^2}{|\vec{p}|^2}}} \tag{0.23}$$

$$\begin{aligned}\sigma_\theta = \sigma_{p_x}^2 &\left(\frac{p_x p_z}{(|\vec{p}|^2)^{\frac{3}{2}}\sqrt{1-\frac{p_z^2}{|\vec{p}|^2}}}\right)^2 + \sigma_{p_y}^2\left(\frac{p_y p_z}{(|\vec{p}|^2)^{\frac{3}{2}}\sqrt{1-\frac{p_z^2}{|\vec{p}|^2}}}\right)^2 + \sigma_{p_z}^2\left(\frac{\frac{p_z^2}{(|\vec{p}|^2)^{\frac{3}{2}}} - \frac{1}{|\vec{p}|}}{\sqrt{1-\frac{p_z^2}{|\vec{p}|^2}}}\right)^2 \\ &+ 2\Bigg[\sigma_{p_x P_y}\left(\frac{p_x p_z}{(|\vec{p}|^2)^{\frac{3}{2}}\sqrt{1-\frac{p_z^2}{|\vec{p}|^2}}}\right)\left(\frac{p_y p_z}{(|\vec{p}|^2)^{\frac{3}{2}}\sqrt{1-\frac{p_z^2}{|\vec{p}|^2}}}\right) \\ &+ \sigma_{p_y p_z}\left(\frac{p_y p_z}{(|\vec{p}|^2)^{\frac{3}{2}}\sqrt{1-\frac{p_z^2}{|\vec{p}|^2}}}\right)\left(\frac{\frac{p_z^2}{(|\vec{p}|^2)^{\frac{3}{2}}} - \frac{1}{|\vec{p}|}}{\sqrt{1-\frac{p_z^2}{|\vec{p}|^2}}}\right) \\ &+ \sigma_{p_x p_z}\left(\frac{p_x p_z}{(|\vec{p}|^2)^{\frac{3}{2}}\sqrt{1-\frac{p_z^2}{|\vec{p}|^2}}}\right)\left(\frac{\frac{p_z^2}{(|\vec{p}|^2)^{\frac{3}{2}}} - \frac{1}{|\vec{p}|}}{\sqrt{1-\frac{p_z^2}{|\vec{p}|^2}}}\right)\Bigg]\end{aligned} \tag{0.24}$$

$$\sigma_\phi = \sigma_{p_x}^2 \left(\frac{\partial\phi}{\partial p_x}\right)^2 + \sigma_{p_y}^2 \left(\frac{\partial\phi}{\partial p_y}\right)^2 + 2\sigma_{p_x p_y}\left(\frac{\partial\phi}{\partial p_x}\right)\left(\frac{\partial\phi}{\partial p_y}\right) \tag{0.25}$$

$$\frac{\partial \phi}{\partial p_x} = \frac{-p_y}{p_x^2 + p_y^2} \tag{0.26}$$

$$\frac{\partial \phi}{\partial p_y} = \frac{p_x}{p_x^2 + p_y^2} \tag{0.27}$$

$$\sigma_\phi = \sigma_{p_x}^2 \left(\frac{-p_y}{p_x^2 + p_y^2} \right)^2 + \sigma_{p_y}^2 \left(\frac{p_x}{p_x^2 + p_y^2} \right)^2 + 2\sigma_{p_x p_y} \left(\frac{-p_y}{p_x^2 + p_y^2} \right) \left(\frac{p_x}{p_x^2 + p_y^2} \right) \tag{0.28}$$

The uncertainty of energy $\sigma_{E_{jet}}$ and direction $\sigma_{\theta_{jet}}$ and $\sigma_{\phi_{jet}}$ are implemented in the kinematic fit instead of the previous fixed values. Some of the covariance matrices of some of the jets have null entries except for σ_E^2. This happens about in $\sim$2% of the events. In this case, $\sigma_{\theta_{jet}}$ and $\sigma_{\phi_{jet}}$ cannot be found using the covariance matrix. In these cases, the previously assumed error of 0.1 for $\sigma_{\theta_{jet}}$ and $\sigma_{\phi_{jet}}$ is applied. The reason for the null entries in the covariance matrix is currently not known and needs to be clarified and fixed in the future.

4.3 Application to the Higgs mass reconstruction

The higgs mass is reconstructed for the events with one semileptonic B decay using the updated uncertainties. The distribution of the Higgs mass with the ErrorFlow and previously used uncertainties are compared in Fig. 19.

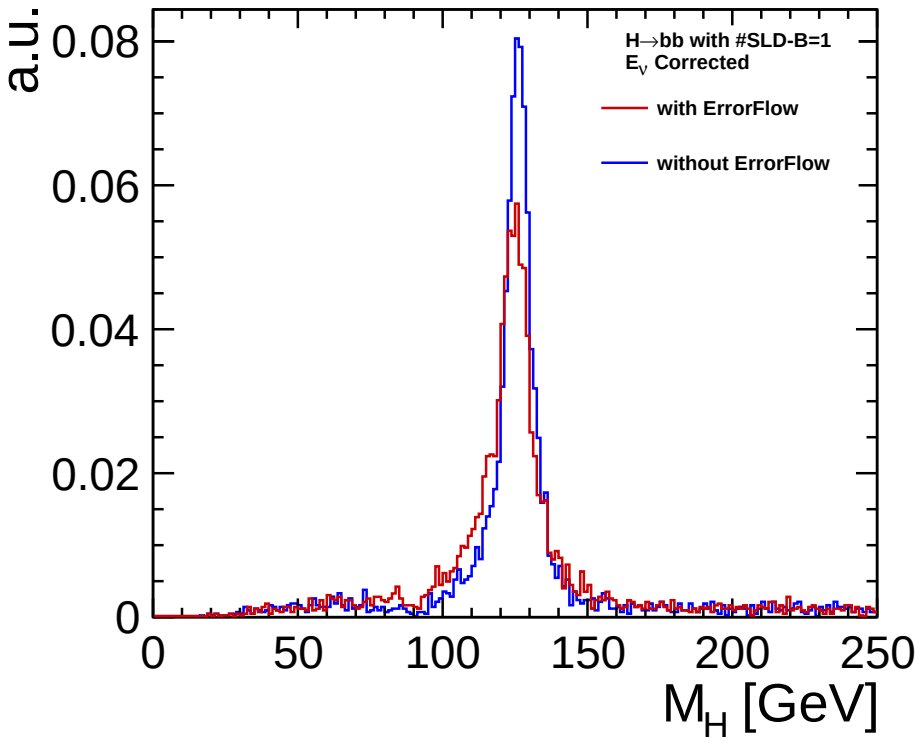


Figure 19: Reconstructed Higgs mass with the updated and previous error parametrizations with the missing neutrino energy corrected

After the ErrorFlow implementation, the shoulder at low energies reappear in the M_H distribution. This could be due to incorrect jet pair pairing or incorrect neutrino energy solution. distribution is wider and shifted to low energies. To investigate further, one can to look into the permutation of the jets in the kinematic fit and check whether the same combination is used for with and without the ErrorFlow.
Pull distributions of E_{jet}, θ_{jet} and ϕ_{jet} and pull distribution with and without the updated uncertainties are plotted and analyzed in this section.

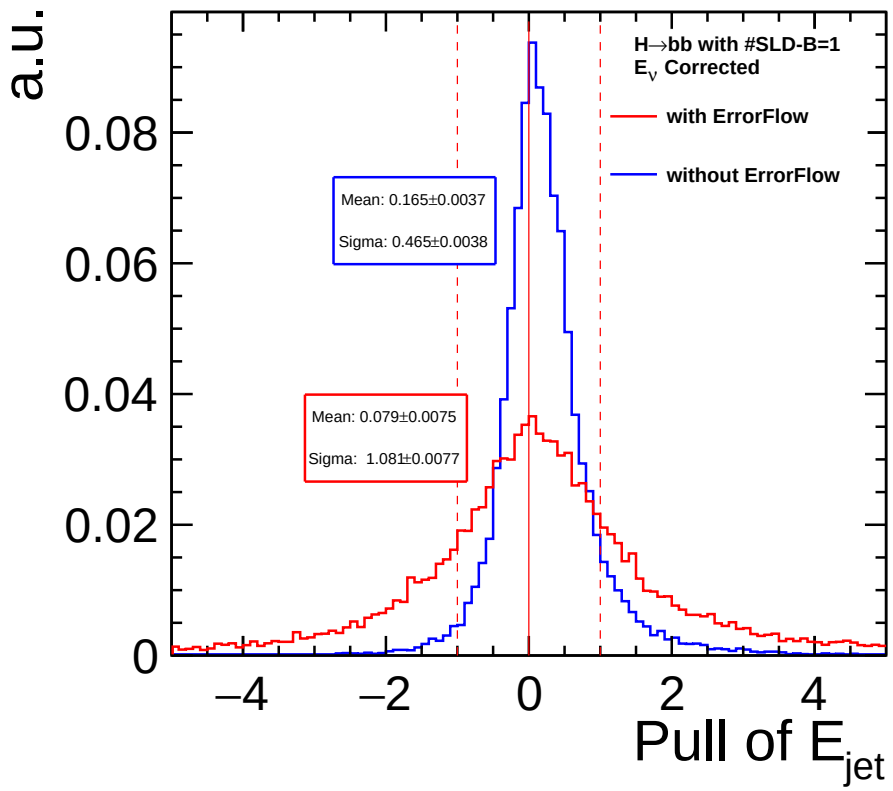


Figure 20: Pull distribution of jet energy with the updated and the previous error parametrization with the missing neutrino energy corrected

It is previously seen that the assumed error are larger without the ErrorFlow. The pull of E_{jet} has a mean ∼44 times its error above 0 and asymmetry of distribution is clearly visible by eye. The sigma of the pull distribution is ∼140 its error below 1. With the implementation of the ErrorFlow, the mean is consistent with 0 within uncertainties and sigma is ∼10 times its error above 1. The mean and sigma are closer to the ideal value, confirming that the implementation of the ErrorFlow functions gives better results.

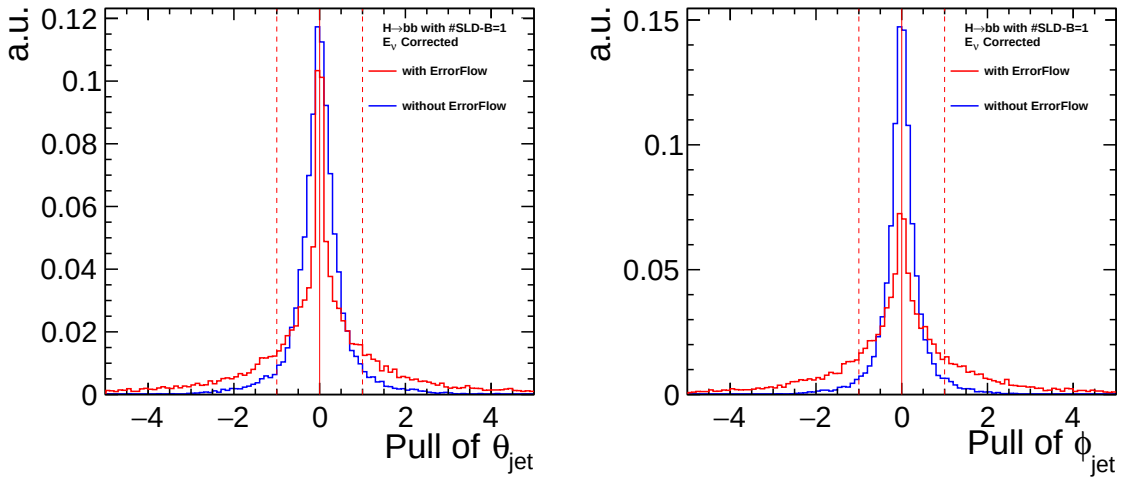


Figure 21: Pull distribution of jet direction with the updated and the previous error parametrization with the missing neutrino energy corrected

Though the sigma of $pull_{E_{jet}}$ is closer to 0, the distribution of the $pull_{\theta_{jet}}$ and $pull_{\phi_{jet}}$ after the ErrorFlow implementation is narrow and less Gaussian-like. One should also take into account that the momentum elements of the covariance matrix are null in some events. Figure suggests correction of jet angles with the neutrinos would be essential.

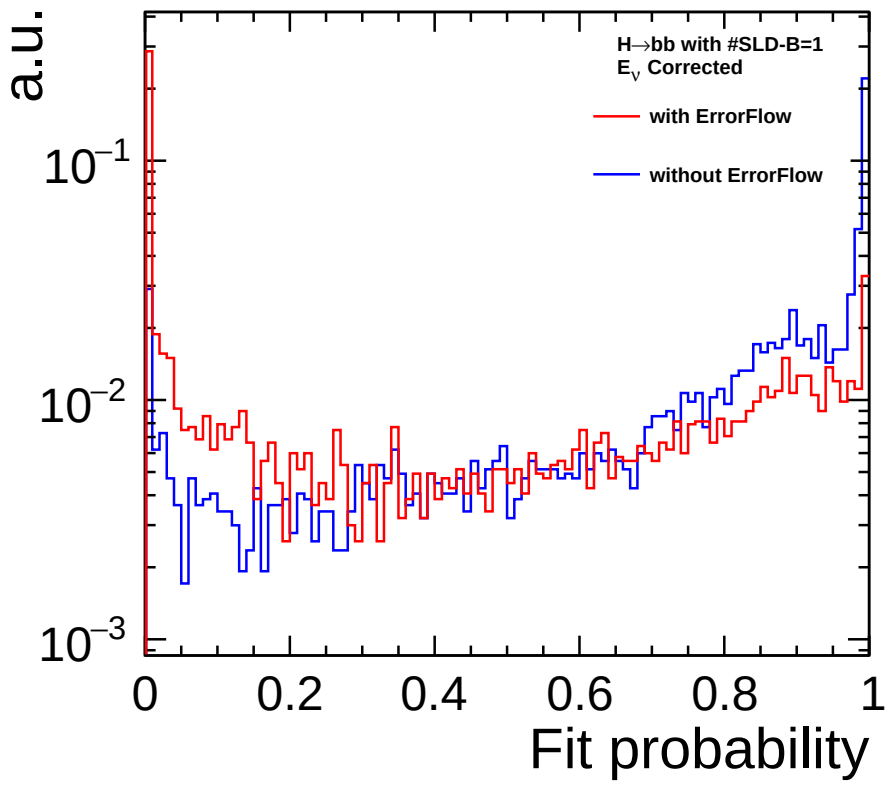
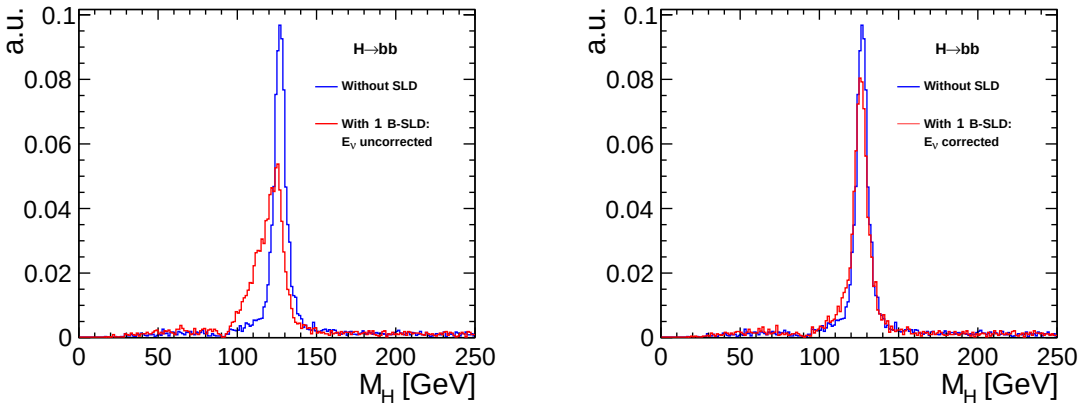


Figure 22: Distribution of the fit probability with the updated and the previous error parametrization with the missing neutrino energy corrected

Overall, the distribution of fit probability with the ErrorFlow implemented, looks flatter compared to the previously assumed uncertainties. In particular, the accumulation of event at large fit probabilities is significantly reduced. However at small fit probabilities, a new peak is appearing. Considering the respecting pull distributions, the next required step is most likely a correction of the jet angles.

5 Conclusion

Precision measurements on the Higgs properties especially its mass is one of the major physics case in the International Linear Collider which is a proposal for a future electron-positron collider with an initial center-of-mass energy of 250 GeV, upgradable up to 1 TeV.
This study quantified for the first time the impact of the number of semileptonic heavy quark decays on the reconstructed Higgs mass. Since in particular events in which at least one B hadron decays semileptonically show a significant tail towards lower Higgs masses, a first attempt has been made to derive a correction for the missing neutrino momentum.
In a proof-of-principle mode which cheats some of the required inputs from MC truth, a neutrino energy correction is calculated from the mass and direction of flight of the B hadron and the visible products of the semileptonic decay. In order to resolve the remaining two-fold ambiguity, a previously developed kinematic fit is employed.



(a) Before the neutrino energy correction

(b) After the neutrino energy correction

Figure 23: Reconstructed Higgs in $H \to bb$ with one semileptonic B decay

The success of this approach in summarized in Fig. refconclu, which shows that with the neutrino correction event with semileptonic b-decays can be nearly as well reconstructed as event without semileptonic decays.
The fit characteristics are investigated with pull distributions and fit probability. Using this, the effect of uncertainties on the jet are analyzed and improved using the implementation of the ErrorFlow in the kinematic fit. It is observed that fit characteristics with respect to energy is improved with the use of the ErrorFlow.
Outlook
This study is a proof of principle that the missing neutrino energy corrections are required in the reconstruction of Higgs in a hadronic channel.
This concept can now applied to the semileptonic decays of the remaining quarks in the

final state. One could also correct for the neutrinos from subsequent semileptonic decay of D hadrons.
The null momentum components in the covariance matrix can be investigated further. In the future, the use of MC true information can be replaced with reconstructed information. This method of implementing the neutrino energy corrections can be extended to the events with more than one semileptonic decays. As discussed in Sec. 4, the error parametrization needs further improvement. The jet angles has to be corrected for the momentum of the neutrinos in the future. After successful implementation of the neutrino energy corrections, the assignment of neutral PFOS to the secondary vertex is an next important thing to study in the future.

References

[1] Particle flow.

[2] The international linear collider, technical design report, volume 1: Executive summary, 2013.

[3] The international linear collider, technical design report, volume 3: Accelerator, 2013.

[4] The international linear collider, technical design report, volume 4: Detectors, 2013.

[5] G. Aad, T. Abajyan, B. Abbott, J. Abdallah, S. Abdel Khalek, A.A. Abdelalim, O. Abdinov, R. Aben, B. Abi, M. Abolins, and et al. Observation of a new particle in the search for the standard model higgs boson with the atlas detector at the lhc. *Physics Letters B*, 716(1):1–29, Sep 2012.

[6] Tim and Barklow, Keisuke Fujii, Yuanning Gao, Andre Hoang, Shinya Kanemura, Jenny List, Heather E Logan, Andrei Nomerotski, Maxim Perelstein, Michael E Peskin, Roman Pöschl, Jürgen Reuter, Sabine Riemann, Aurore Savoy-Navarro, Geraldine Servant, and Tim M P and Tait. The International Linear Collider Technical Design Report - Volume 2: Physics. Technical report, Geneva, Jun 2013. arXiv:1306.6352.

[7] Moritz Beckmann, Benno List, and Jenny List. Treatment of photon radiation in kinematic fits at future e+ e- colliders. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 624:184–191, 06 2010.

[8] J. Beyer. Measuring vector boson scattering at the future international linear collider, 2018.

[9] S. Chatrchyan, V. Khachatryan, A.M. Sirunyan, A. Tumasyan, W. Adam, E. Aguilo, T. Bergauer, M. Dragicevic, J. Erö, C. Fabjan, and et al. Observation of a new boson at a mass of 125 gev with the cms experiment at the lhc. *Physics Letters B*, 716(1):30–61, Sep 2012.

[10] Aliakbar Ebrahimi. Jet Energy Measurements at ILC:Calorimeter DAQ Requirements and Application in Higgs Boson Mass Measurements, 2017.

[11] F. Englert and R. Brout. Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett.*, 13:321–323, Aug 1964.

[12] C. Adloff et al. Construction and commissioning of the calice analog hadron calorimeter prototype, 2010. arXiv:1003.2662.

[13] G. Baulieu et al. Construction and commissioning of a technological prototype of a high-granularity semi-digital hadronic calorimeter, 2015. arXiv:1506.05316.

[14] H. Aihara et al. Sid letter of intent, 2009. arXiv:0911.0006.

[15] P.N. Burrows et al. Sid detailed baseline design, 2012. http://www.desy.de/ stanitz/DBD-pac101212.pdf.

[16] (ILD Concept Group) H. Stoeck et al. T. Bergauer et al. The international large detector: Letter of intent, 2010.

[17] M. A. Thomson J. S. Marshall. The pandora particle flow algorithm, 2013. arXiv:1308.4537.

6 Acknowledgments

I would like to thank the summer student organization team for this remarkable opportunity. Also, I would like to thank my supervisor Jenny for offering me a very interesting topic and for supervising me. Many thanks to Jakob for the guidance and the patience. I would like to thank Paul for all the quick help and insight and my fellow summer students Andreas, Lorenzo, Mattew and Peter for all the support and fun. Special thanks to Mattew for proofreading in the last minute ;) Thanks to Vikas for all the love and support during the summer school. Finally, I would like to thank the FLC group for providing an pleasant atmosphere to conduct my research.

List of Figures

7 Appendix

7.1 Neutrino Kinematics[8]

Momentum and energy conservation are necessary to calculate the neutrino momentum parallel to the direction of the initial particle.

$$E_B = E_{\text{vis}} + E_\nu \overset{(0.3)}{=} E_{\text{vis}} + |p_\nu| \tag{0.29a}$$
$$= E_{\text{vis}} + \sqrt{p_{\nu,\parallel}^2 - p_{\nu,\perp}^2} \tag{0.29b}$$
$$\overset{(0.5)}{=} E_{\text{vis}} + \sqrt{p_{\nu,\parallel}^2 + p_{\text{vis},\perp}^2} \tag{0.29c}$$

$$p_{B,\parallel} = |\vec{p_B}| = \sqrt{E_B^2 - m_B^2} \tag{0.30a}$$
$$\overset{(0.29c)}{=} \sqrt{\left(E_{\text{vis}} + \sqrt{p_{\nu,\parallel}^2 + p_{\text{vis},\perp}^2}\right)^2 - m_B^2} \tag{0.30b}$$
$$= p_{\nu,\parallel} + p_{\text{vis},\parallel} \tag{0.30c}$$

$$\overset{()^2}{\Longrightarrow} \quad p_{\nu,\parallel}^2 + 2p_{\nu,\parallel}p_{\text{vis},\parallel} + p_{\text{vis},\parallel}^2 = p_{\nu,\parallel}^2 + p_{\text{vis},\perp}^2 + 2\sqrt{p_{\nu,\parallel}^2 + p_{\text{vis},\perp}^2}E_{\text{vis}} + E_{\text{vis}}^2 - m_B^2 \tag{0.31a}$$
$$\Longrightarrow \quad 2\sqrt{p_{\nu,\parallel}^2 + p_{\text{vis},\perp}^2}E_{\text{vis}} = 2p_{\nu,\parallel}p_{\text{vis},\parallel} - 2p_{\text{vis},\perp}^2 - m_{\text{vis}}^2 + m_B^2 \tag{0.31b}$$

$$\overset{()^2}{\Longrightarrow} \quad 0 = \left(4E_{\text{vis}}^2 - 4p_{\text{vis},\parallel}^2\right)p_{\nu,\parallel}^2 + 4p_{\text{vis},\parallel}\left(2p_{\text{vis},\perp}^2 + m_{\text{vis}}^2 - m_B^2\right)p_{\nu,\parallel} \tag{0.32a}$$
$$+ 4p_{\text{vis},\perp}^2E_{\text{vis}}^2 - \left(2p_{\text{vis},\perp}^2 + m_{\text{vis}}^2 - m_B^2\right)^2 \tag{0.32b}$$
$$\equiv 4D \cdot p_{\nu,\parallel}^2 + 4A \cdot p_{\nu,\parallel} + B \tag{0.32c}$$

Here the shortcuts

$$A = p_{\text{vis},\parallel}\left(2p_{\text{vis},\perp}^2 + m_{\text{vis}}^2 - m_B^2\right) \tag{0.33a}$$
$$B = 4p_{\text{vis},\perp}^2E_{\text{vis}}^2 - \left(2p_{\text{vis},\perp}^2 + m_{\text{vis}}^2 - m_B^2\right)^2 \tag{0.33b}$$
$$D = E_{\text{vis}}^2 - p_{\text{vis},\parallel}^2 \tag{0.33c}$$

are introduced.
The quadratic equation (0.32) has two solutions.

$$p_{\nu,\parallel} = \frac{1}{2D}\left(-A \pm \sqrt{A^2 - BD}\right) \tag{0.34}$$

Equations (0.5), (0.3) and (0.34) determine the invisible four-momentum up to two solutions. Which one of these solutions is correct can not be determined from the available kinematics. Additional information e.g. about the scattering process, the specific decay or the decay products would be needed to infer the sign in eq. (0.34).