Higgs measurement in the four jets final state at a e+e- collider

Prepared by Gaoyuan Wang

University of Goettingen

Supervisor: Jike Wang, Krisztian Peters, DESY

07. Sept. 2016

Contents

1 Introduction

The Standard Model of elementary particles (SM) as an effective theory describes the elementary particles and their interactions. The are four fundamental interactions between elementary particles, namely gravitation, electromagnetism, weak interactions and strong interactions. The electromagnetism is described via QED, the mediation of the interaction is the gauge boson photon. The $W$and Z boson take over the role of mediation in the weak interaction. And the QCD describing strong interactions has the gluon as mediation (see also [1]). In the Standard Model (SM), the elementary particles gain their mass through the coupling with higgs boson. This mechanism is called the higgs mechanism. The higgs mechanism is the way that the W and Z bosons acquire mass without breaking the local gauge symmetry of the Standard Model. Without it, the SM is not a consistent theory, for only theories with local gauge invariance are renormalisable. higgs mechanism is the spontaneous symmetry breaking of a complex scalar field. Embedding it in a theory with a local gauge symmetry leads to extra polarizations freedoms instead of massless Goldstone Bosons. For complex scalar field:

$$\Phi = \frac{1}{\sqrt{2}}(\Phi_1 + i\Phi_2)\,, \tag{1}$$

we choose the vacuum at $(\Phi_1, \Phi_2) = (v, 0)$. Then the perburbation of Φ around the vacuum is:

$$\Phi_1(x) = \eta(x) + v\,, \tag{2}$$

$$\Phi_1(x) = \xi(x)\,. \tag{3}$$

The Lagrangian now reads:

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\eta)(\partial^\mu\eta) - \frac{1}{2}m_\eta^2\eta^2 + \frac{1}{2}(\partial_\mu\xi)(\partial^\mu\xi) - V_{int}(\eta,\xi) \tag{4}$$

$$with \tag{5}$$

$$V_{int} = \lambda v\eta^3 + \frac{1}{4}\lambda\eta^4 + \frac{1}{4}\lambda\xi^4 + \lambda v\eta\xi^2 + \frac{1}{2}\lambda\eta^2\xi^2\,, \tag{6}$$

$$m_\eta = \sqrt{2\lambda v^2}\,. \tag{7}$$

One can see that this Lagrangian also contain a massless Goldstone Boson ξ. In order to keep the Lagrangian invariant under the transformation

$$\Phi(x) \to \Phi'(x) = \exp(ig\chi(x))\Phi(x)\,, \tag{8}$$

we need a new gauge field $B_\mu(\partial_\mu \to D_\mu = \partial_\mu + igB_\mu)$. Now the Lagrangian becomes:

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\eta)(\partial^\mu\eta) - \frac{1}{2}m_\eta^2\eta^2 + \frac{1}{2}(\partial_\mu\xi)(\partial^\mu\xi) - V_{int}(\eta,\xi) \tag{9}$$

$$- \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}g^2v^2B_\mu B^\mu + gvB_\mu(\partial^\mu\xi)\,, \tag{10}$$

and the gauge field also becomes massive. Despite its success, SM can't explain all the problems. For instance he smallness of higgs mass (e.g. compare to the Planck scale M_{pl}) requires fine-tuning [2] and leads to the hierarchy problem of higgs mass. And also the nature of the mysterious dark matter particle is unknown. There are extended theories of SM like Supersymmetry (SUSY) proposed to deal with these unsolved problems, which predicts the existence of more than one higgs boson. This makes a precision higgs measurement interesting. International linear collider (ILC) and future linear collider (FCC) are the proposed next generation e^+e^- collider, which could serve as a higgs factory.

2 Theoretical and experimental background

2.1 LHC and ATLAS

The Large Hadron Collider (LHC) is the current world's largest particle accelerator, one of its designed purpose is the search for higgs boson. The first search for the higgs boson was conducted at the Large Electron-Positron Collider (LEP) at CERN in the 1990s. The first research run at LHC started in February 2012, ended in May 2013, with $\sqrt{s} = 7 - 8$ TeV. After a break for upgrades, the second run starts in may 2015 with $\sqrt{s} = 13$ TeV. The operating temperature for the magnets, which guides the beams and keep them focused is 1.9K [3]. The design luminosity of the LHC is $10^{34}\text{cm}^{-2}\text{s}^{-1}$. The higgs boson was found at LHC Juli 2012.
There are 7 detectors at CERN, the two general purpose detectors are ATLAS and CMS. Both of them are looking for new physics including sign for extra dimension and dark matter, one of their major task is also to study higgs Boson. Although they have the same scientific goals, they use different detection technologies.
ATLAS is 46 meter long and has a height of around 25 meter. It has 4 major components, inner detector, calorimeter, muon spectrometer and magnet system. The inner detector measures the momentum of each charged particle. The electromagnetic and hadronic calorimeter measure the energies of the corresponding particles. The muon spectrometer is used to identify and measure the momentum of muon. The magnet system bends charged particles for momentum measurement. Also trigger-devices can be found around calorimeters and muon chamber, this enable a first level trigger to select interesting events.
ATLAS also has a powerful data analysing system containing the trigger system, the data acquisition system and the computing system. The trigger system selects 100 interesting events out of 1000 million in a second. The data acquisition system channels data from detector to the storage and the computing system analyses 1000 million events per year.

2.2 FCC-ee and ILC

The FCC-ee is proposed as the circular e^+e^- accelerator for the post-LHC era with high precision and high luminosity. The ILC is a proposed linear e^+e^- accelerator with the

planned collision energy of 500 GeV initially. They could be the window to answer the questions: What is the nature of dark matter? Does SUSY particles really exist? How can the hierarchy problem be explained?
To serve as a higgs factory, FCC-ee or ILC should have at least centre-of-mass energy around 240 GeV (350 GeV for top pairs). e^+e^- collision enables a clean experimental conditions, which is important for measurement of properties of the Z, W, higgs and top particles with high accuracy. The measurement of the invisible or exotic decays of the higgs and Z bosons could be sensitive dark matter or other new physics [12] [11].

2.3 Higgs production and decay channel

A higgs boson can be produced at a e^+e^- collider through higgs radiation by a Z namely $e^+e^- \to Zh$. Where the higgs and the Z decay further. We are interested in the full hadronic channel $h \to b\bar{b}$ and $Z \to q\bar{q}$ where the q stands for light quarks [5]. Around 60% of the higgs boson would decay into a pair of b and anti-b quarks. 2.1% of the higgs boson would decay to two W bosons, and 2.5% to two Z bosons [4]. The detector records decay product of Z and h boson, which forms the signature of the event.
This process has several backgrounds, For instance the Z pair can also decay into four jets. And also two W bosons has fullhadronic channel, in which case both W bosons decay into $q\bar{q}'$.

2.4 Event sampling

We use three tools to sample the events and simulate the detector response: WHIZARD, PYTHIA and DELPHES.

- WHIZARD is a program system designed for the efficient calculation of multi-particle scattering cross sections and simulated event samples. In WHIZARD, Tree-level matrix elements are generated automatically for arbitrary partonic processes. These events are then written to file in Les Houches event format (LHEF), and can then be hadronized [7].

- PYTHIA is a program for the generation of high-energy physics events. It can simulate hard and soft interactions, parton distributions, initial- and final-state parton showers, multiparton interactions, fragmentation and decay. We use PYTHIA to simulate the decay of Z, W and h, their further fragmentation and hadronization of their decay products [8].

- Delphes is a C++ framework simulating detector response. By editing the DELPHES card one can include a tracking system, embedded into a magnetic field, electromagnetic and hadronic calorimeters and a muon system. The input file for this framework is standard file formats (e.g. Les Houches Event File or HepMC) and it outputs observables such as isolated leptons, transverse momentum and collection of jets which can be used for dedicated analyses in a ROOT file. The simulation of the detector response takes into account the effect of magnetic field, the efficiency, the granularity of the calorimeters and sub-detector resolutions [9].

For the process $e^+e^- \to Zh$ on the reconstructed level, we first use WHIZARD to generate the process $e^+e^- \to Zh$ and write the file in LHEF. Then input this file to PYTHIA to generate the decay of Z and h. Then we use DELPHES to simulate the signatures in the ILD (International Large Detector), which is then analysed.
For each processes, we generate one million events (0.2 million for the process $e^+e^- \to q\bar{q}b\bar{b}$ due to time constrain). Due to some unknown failure of PYTHIA, we loose about 20% of the events, so that at the end we have around 8×10^5 events. We don't know if the events we lost in PYTHIA are random or if they share some common feature, which could effect the result of analysis.

3 Analysis

3.1 Study of the four jets channel

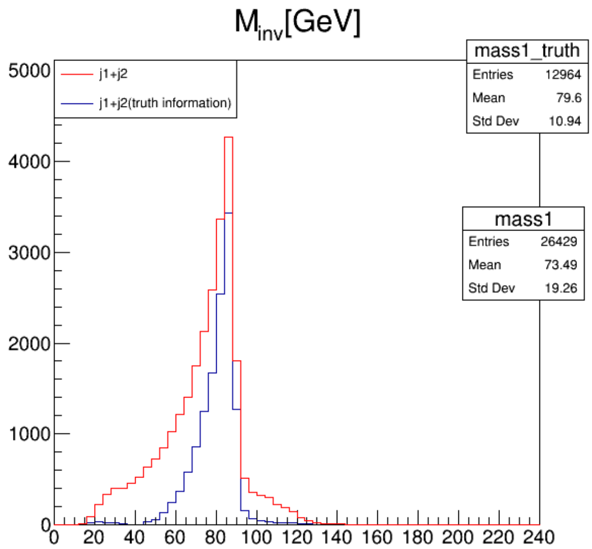
To avoid mismatching, we just look at events with two b jets and two light q jets (i.e. u, d, c or s jets.). Since the branching ratio of $h \to b\bar{b}$ is 60%, we will in the most cases obtain $h \to b\bar{b}$ and $Z \to q\bar{q}$. About 10% of the events will be $Z \to b\bar{b}$ and $h \to q\bar{q}$.
We first study the invariant mass of the two quarks pairs at generator level, where we can look into the truth information.
Then we move to the detector level and study the reconstructed jets (rec jets). Here we first use the truth information at generator level and do a ΔR matching for the rec jet. This means we firth find the parton which has as mother particle a h or Z, we then search for rec jets within a cone with ΔR parameter $= 0.5$ around the parton. For h we require the rec jets to be b-tagged while for Z not b-tagged.
To understand this channel better, we compare the analysis with and without truth-matching, as shown in Fig. 1.
As we can see, some events didn't pass the truth matching, these are for instance b, τ faking the signature of $Z \to jj$ and τ, c, gluon faking the signature $h \to b\bar{b}$.
The major backgrounds [13] that fake the $bbjj$ signature in a e^+e^- collider are:

- $e^+e^- \to ZZ$
- $e^+e^- \to W^+W^-$
- $e^+e^- \to b\bar{b}q\bar{q}$.

The cross-section of each process is calculated automatically in WHIZARD when generating the events. The cross-section and the event size are listed in Tab. 1.
The weight of each process is defined as

$$w = \frac{\text{Cross-section}}{\text{Event size}}\,. \tag{11}$$

We apply here the following cuts according to [13]:



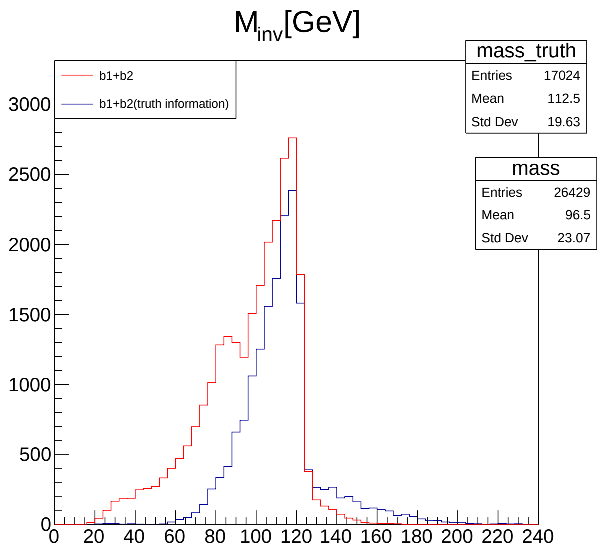


Figure 1: The comparison of the results with and without truthmatching. The cone size for truthmatching is set to $\Delta R = 0.5$. The difference between blue and red curve shows there are particles faking the signal.

Process	Event size	Cross-section
$e^+e^- \to Zh$	804860	2.0304×10^2
$e^+e^- \to ZZ$	808028	1.09×10^3
$e^+e^- \to W^+W^-$ 2	807126	1.6682×10^4
$e^+e^- \to bbqq$	260766	2.359×10^4

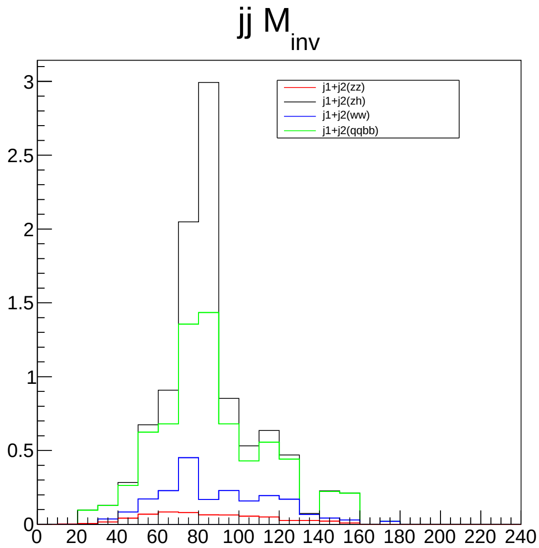
Table 1: Event size and cross section for each process.

- All the jets should have $p_t > 20$ GeV.
- For all jets, at least 5 reconstructed particles should be found, at least one of them should be charged.
- The sum of visible mass should be larger than 180 GeV.

We drop the truth information matching and study events with exact two b jets and two light jets ($bbjj$), and plot the inv. mass. The result is shown in Fig. 2.
Instead of being rather flat, the background process $ee \to qqbb$ shows a peak at Z and h pole. This is not unexpected for two reasons. The first reason is that when applying the cuts we mentioned earlier, we tend to select events which has invariant mass at Z and h pole. The second reason is due to time constrain, we didn't have time to generate enough samples, so more fluctuation is expected for this process. Another interesting feature is that the peak in Fig. 2 is slightly shifted towards the lower end, this is not caused by jet radiation but rather the settings of the calorimeter. We will return to this point in Sec. 3.2.

3.2 Rescaled jet energy

Electrons are not protons, they do not interact strongly, this means there's no pile-up collisions. Thus are e^+e^- colliders the tool of precision measurement. Also electrons are leptons, there is thus no underlying event, the final states have known energy and momentum $(\sqrt{s}, 0, 0, 0)$ [10]. For instance in the four jet channel, the total energy and momentum are conserved, and jet directions are very well measured. We can thus determine the jet energies analytically using the measured three momentum. It turns out that this allows more accurate analysis.
In Fig. 3, we can see the difference in the reconstructed invariant mass using the energy measured in the calorimeter and calculated using energy and momentum conservation.
We now study again the invariant mass of $b\bar{b}$ and jj using the calculated rescaled jet energy for both the signal and background. The result is shown in Fig. 4. As we can see in Fig. 4, the peak of the invariant mass is now at h/Z pole compare to Fig. 2.



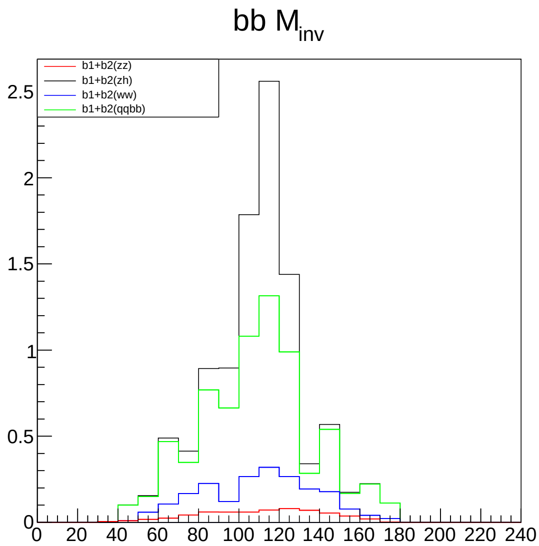
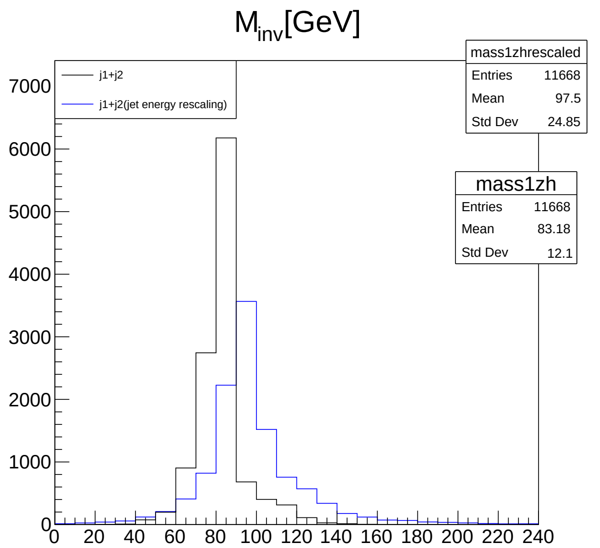


Figure 2: The invariant mass for signal and background.



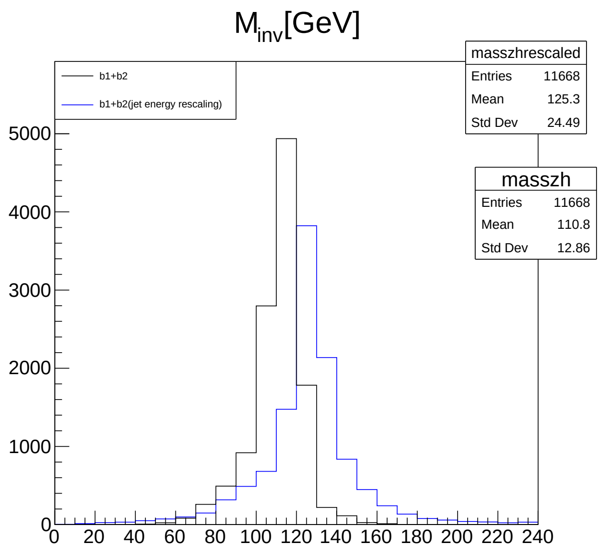


Figure 3: The comparison between invariant mass calculated using energy measured in the calorimeter and using rescaled jet energy.

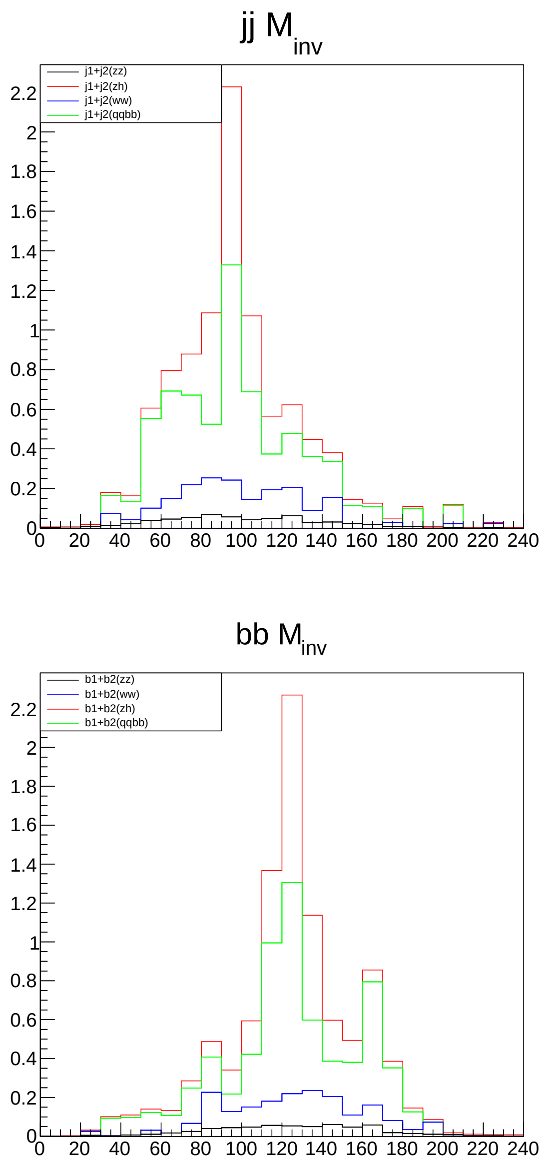


Figure 4: The invariant mass for signal and background using rescaled jet energy.

4 Conclusion

In this 8 weeks project, we study the higgs measurement in the four jets channel and the sensitivity depending on the detector parameters at a e^+e^- collider. The higgs is produced via higgs radiation $e^+e^- \rightarrow Zh$. We use three tools: WHIZARD, PYTHIA and DELPHES to generate event samples and simulate detector response. We noticed a failure in PYTHIA and we loose around 20% of the events, and we can't fix the problem. We use in DELPHES a ILD (international large collider) like card with tunable parameters.
With the event samples, we study the four jet ($bbjj$) channel. We also compare the result using the truth information and the result without truth information and thus find several particles which fake the signal. Furthermore we study the major backgrounds of the event in Sec. 3.1.
We then take the advantage of a e^+e^- collider that the four momentum of the initial state is known to calculate the jet energy using the measured three momentum. We further compare the result obtained this way with the one using energy measured in the calorimeter in Sec. 3.2.
The next step is to study how the sensitivity depends on the detector parameter in the DELPHES card and use this as indication for detector design study.

References

[1] M. Thomson, “Modern particle physics”

[2] N. Seiberg, Phys. Lett. B **318** (1993) 469 doi:10.1016/0370-2693(93)91541-T [hep-ph/9309335].

[3] ”http://home.cern/about”

[4] R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, R. Pittau and P. Torrielli, JHEP **1109** (2011) 061 doi:10.1007/JHEP09(2011)061 [arXiv:1106.6019 [hep-ph]].

[5] ”https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CrossSections”

[6] ”Cards provided by the DELPHES collaboration”

[7] ”WHIZARD website: https://whizard.hepforge.org/”

[8] ”PYTHIA website: http://home.thep.lu.se/ torbjorn/Pythia.html”

[9] ”DELPHES website: https://cp3.irmp.ucl.ac.be/projects/delphes”

[10] ”Patrick Janot, Lecture note: Physics at Future Colliders, 28-29 July 2016”

[11] ”FCCee website:http://tlep.web.cern.ch/”

[12] ”ILC website: https://www.linearcollider.org/ILC”

[13] P. Azzi, C. Bernet, C. Botta, P. Janot, M. Klute, P. Lenzi, L. Malgeri and M. Zanetti, arXiv:1208.1662 [hep-ex].