Influence of timing cuts in Testbeam data and Simulation

Dissertation

zur Erlangung des Doktorgrades

des Department Physik

der Universität Hamburg

vorgelegt von

ELDWAN BRIANNE

aus Saint-Malo, Frankreich

Hamburg

2017

Abstract

Etiam pede massa, dapibus vitae, rhoncus in, placerat posuere, odio. Vestibulum luctus commodo lacus. Morbi lacus dui, tempor sed, euismod eget, condimentum at, tortor. Phasellus aliquet odio ac lacus tempor faucibus. Praesent sed sem. Praesent iaculis. Cras rhoncus tellus sed justo ullamcorper sagittis. Donec quis orci. Sed ut tortor quis tellus euismod tincidunt. Suspendisse congue nisl eu elit. Aliquam tortor diam, tempus id, tristique eget, sodales vel, nulla. Praesent tellus mi, condimentum sed, viverra at, consectetuer quis, lectus. In auctor vehicula orci. Sed pede sapien, euismod in, suscipit in, pharetra placerat, metus. Vivamus commodo dui non odio. Donec et felis.

Zusammenfassung

Donec et nisl id sapien blandit mattis. Aenean dictum odio sit amet risus. Morbi purus. Nulla a est sit amet purus venenatis iaculis. Vivamus viverra purus vel magna. Donec in justo sed odio malesuada dapibus. Nunc ultrices aliquam nunc. Vivamus facilisis pellentesque velit. Nulla nunc velit, vulputate dapibus, vulputate id, mattis ac, justo. Nam mattis elit dapibus purus. Quisque enim risus, congue non, elementum ut, mattis quis, sem. Quisque elit.

Contents

Introduction

Chapter 1

ILD detector simulation studies

Simulation of detector response is an essential part in high energy physics experiment. In Early stage of a project, simulations are done in order to explore and understand the possibilities of a detector design as well as its limitations. Simulation can be use as a way to determine requirements of an experiment to reach certain goals. During data-taking and afterwards, simulations are used model physics processes to compare the expected value from theory to a measured value for various processes.

In this chapter, software tools will be briefly introduced. The ILCSOFT framework used for this analysis will be described in **??**. The chain starts by the simulation of single kaons (K_L^0) interaction with the ILD detector model based on GEANT 4. Then simulated events undergoes the full chain reconstruction as explained in **??**. The procedure of the analysis (based on Marlin) and its conclusions will be presented in **??**. Finally, a benchmark of a fast simulation software (SGV) against the ILD full simulation will be described in **??**, particularly focussing on particle-flow performance aspects.

1.1   Simulation and software framework

1.1.1   ILCSOFT software framework

Various tools developed by the Linear Collider community is regrouped in a common software framework called ILCSOFT [1]. It provides a complete framework that can be used for Monte-Carlo studies and experiments. As an example, physics studies, ILD detector optimisation and performance for the ILC are performed under the ILCSOFT framework.

Most of the tools in the framework use an Event Data Model (EMD) named Linear Collider I/O (LCIO) which provides a reliable and performant solution for simulation and analysis studies [5]. With this tool, various detector concepts and analysis can be shared.

The ILCSOFT framework provides a modular C++ framework named MARLIN for reconstruction and analysis of physics events [4]. MARLIN uses LCIO seamlessly and is configured using XML steering files. MARLIN enables users to develop custom modules for their own and run it along other already existing modules.

The reconstruction and analysis tools used in this analysis are mostly part of ILCSOFT. For this thesis, ILCSOFT v01-17-11 was used for simulation, reconstruction and analysis.

1.1.2   ILD Detector Simulation

The following analysis is using one of the generic ILD detector model (ILD_o1_v05) as describe in **??** within the MOKKA framework. Many other models are also considered for ILD as shown in Table 1.1. MOKKA is a front-end to GEANT 4 and provides a realistic geometry of the ILD detector. The MOKKA version used is v08-05 and the GEANT 4 version is 10.01. The simulation is performed by simply using the particle gun provided in GEANT 4 to shoot particles (π^- or K_L^0) in different regions of the detector by randomly variating the angles θ and ϕ of the gun. To model hadronic showers, the QGSP_BERT physics list was

<div align="center">

Table 1.1 – Considered ILD detector options.

Option	ECAL Technology	HCAL Technology
ILD_o1_v05	SiW-ECAL	AHCAL
ILD_o2_v05	SiW-ECAL	SDHCAL
ILD_o3_v05	Sc-ECAL	AHCAL

</div>

used. The output of the simulation provides a lcio file containing collections of the tracking hits and simulated calorimeter hits. This file is then reconstructed within MARLIN.

1.2   Reconstruction chain

The reconstruction is done on simulated data in order to implement detector effects. For example, calorimeter hits need to be digitized by implementing threshold and readout effects.

1.2.1   Tracking

The tracking reconstruction is performed on each individual tracking detector. Track segments are identified by pattern recognition algorithms.

Track fitting is performed using the track segments with an inversed Kalman filter to identify trajectories of charged particles. Each tracks contains origin, direction, charge and momentum of the particle [3].

1.2.2   Calorimeter digitization

The calorimeter digitisation is performed on simulated calorimeter hits as part of ILD-CaloDigi processor [2]. It takes account for threshold effect from the electronics, sampling fraction of the calorimeter and the readout technology used. In the considered model of ILD, the SiW-ECAL and AHCAL are used.

In both cases, it uses a silicon-pixel based technology. The digitisation then takes into account the finite number of pixels that can be fired as well as the statistical fluctuations related to pixel readout [6].

Concerning time, it uses a simple digitisation. For a simulated hit, all contributions are looped over and only adds contributions under a certain timing cut (default is 100 ns). This modelisation of timing is very simplified as in reality the electronics are shaping the signal with a certain shapping time and register the time of the first contribution over the threshold (default is 0.5 MIP) **??**.

1.2.3   Pandora PFA

PandoraPFA [7] is the Particle Flow algorithm used for Linear Colliders as explained in **??**. It uses as input tracks and calorimeter hits to form Particle Flow Objects (PFO). It uses a complex multi-stage process but basically, calorimeter hits are clustered and associated to tracks (if any) then the energy of a cluster can be corrected to improve the energy resolution. If the right criterium are matched, it forms a PFO which contains information about the reconstructed objects.

1.3   Influence of time cuts on hadronic showers

In this section, a study of timing cuts on hadronic shower is performed. The goal of this study is to assess the influence of timing cuts on the properties of hadronics showers as for example the width of the shower as well as the needed time resolution. The study will be divided in 2 parts, the first part assuming a perfect time resolution and the second part assuming time resolution for different cases.

1.3.1   Modification of timing window in ILDCaloDigi

Timing of hits is registered in a very simplyfied way as explained in **??**. The modification of the time window (ranging from 1 ns to 100 ns) is performed during the reconstruction for different simulated K_L^0 energies (ranging from 5 to 90 GeV).

1.3.2   Effects of calibration constants and Pandora constants

Before studying the effect of timing on hadrons showers, a check was performed on the initial provided calibration constants of the ILD detector. Several constants are used for the digitisation and reconstruction (GeV to MIP, sampling, Pandora EM/Had constants...) in order to get the correct reconstructed energy. The plots below are selecting events with only one PFO and a $\cos \theta$ cut on the reconstructed particle of 0.7.

The figures 1.1a and 1.1b show the linearity and resolution curves for different sets of calibration constants at the cluster hit level i.e. looking at all the hits in a PFO cluster. Thus this enables to understand the effects of the digitisation constants in ILDDigiCalo though small clustering effects are present.

One can see that the blue and black curves are very similar due to the fact that no constants were changed in theses sets. Moreover the linearity is not perfect and varies between -15% and 5% also the curve crosses the line $x = y$ which if corrected would degrade the energy resolution. The green and red lines are similar as they have the same constants (the non-linearity correction is only applied to PFOs). The linearity fluctuates between -15% and -5% but does not cross the line $x = y$.

Concerning the energy resolution all the curves are very similar and are as expected. The green and red curves are slightly better due to the improvement of the calibration constants.

Another option is to look at the PFO level as shown in figures 1.2a and 1.2b. This
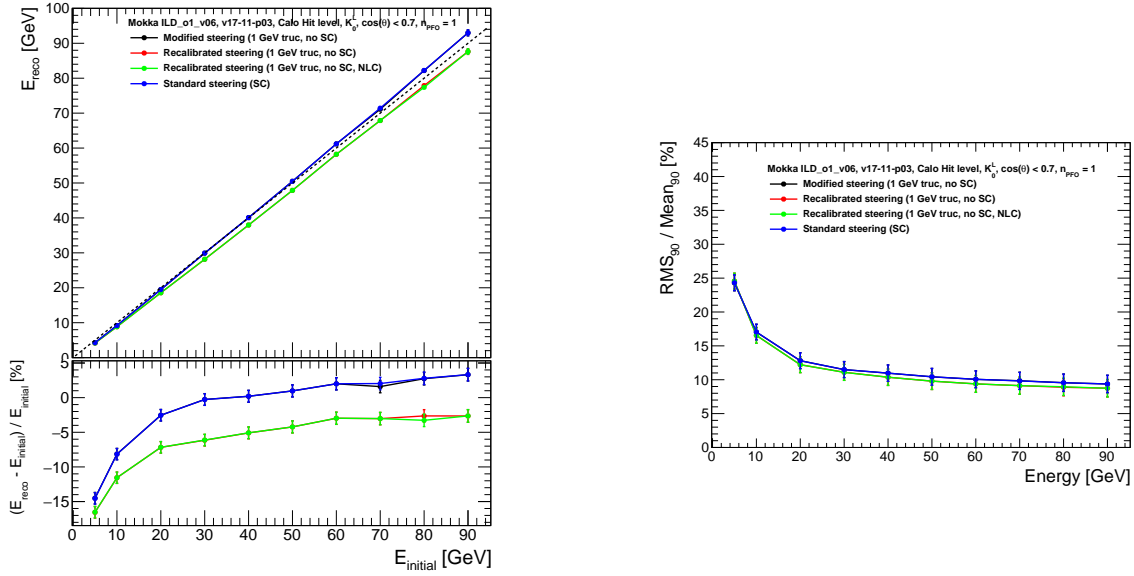
(a) Cluster hit linearity curve

(b) Cluster hit resolution curve

Figure 1.1 – a) The top plot shows the mean reconstructed energy E_{reco} for 5 to 90 GeV K_L^0 function of the simulated energy $E_{initial}$ for different constant parameters used in the reconstruction at the cluster hit level. The bottom plot shows the relative difference of the different curves to the line $x = y$. b) The plot shows the relative resolution $RMS_{90}/Mean_{90}$ for different constant parameters used in the reconstruction function of the energy. The blue curve uses the standard calibration, the black curve uses a modified set of parameters using energy truncation and no SC, the red curve uses constant parameters after recalibration and the green curve uses the same parameters as the red curve with non-linearity correction. The error bars represent statistical uncertainties.

permits to understand the effects of the calibration constants in PandoraPFA.

The plots show a different picture. For the linearity curve, the red and black line are quite similar and show a non-lineariy especially at high energies between -10% and 2%. The green line is nicely linear thanks to the non-linearity correction. Then the blue line is off by around 10%, this is believed to be due to the weights of the software conpensation that are not optimal for this model.

For the resolution curves, one can see a rise of the resolution at high energies for the red, green and black lines certainly due to the non-linearity. The blue curve present a bump after 50 GeV changing suddenly the slope of the curve due to the over-correction of the energy.
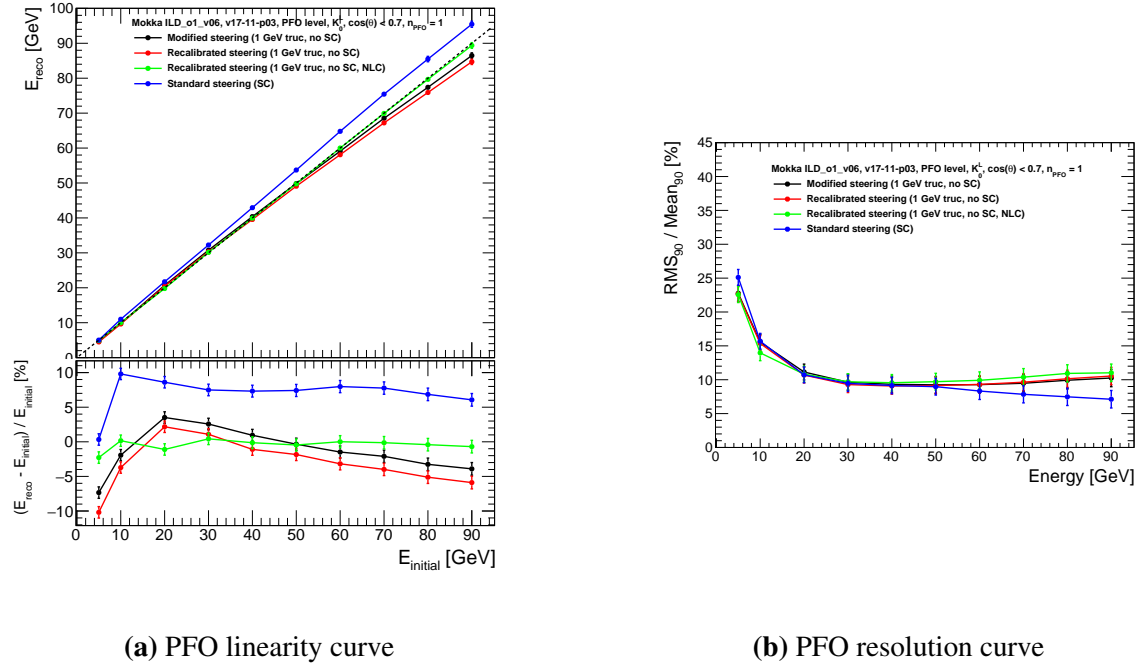
(a) PFO linearity curve

(b) PFO resolution curve

Figure 1.2 – a) The top plot shows the mean reconstructed energy E_{reco} for 5 to 90 GeV K_L^0 function of the simulated energy $E_{initial}$ for different constant parameters used in the reconstruction at the PFO level. The bottom plot shows the relative difference of the different curves to the line $x = y$. b) The plot shows the relative resolution $RMS_{90}/Mean_{90}$ for different constant parameters used in the reconstruction function of the energy. The blue curve uses the standard calibration, the black curve uses a modified set of parameters using energy truncation and no SC, the red curve uses constant parameters after recalibration and the green curve uses the same parameters as the red curve with non-linearity correction. The error bars represent statistical uncertainties.

Through the linearity is not perfect over all energies, the most regarded observable is the jet energy resolution. As explained in **??**, jets are mostly composed of charged particles of around 60%. In this case, the energy of the PFO is coming from the track. Moreover neutral particles are counting in general for around 30% of the contribution in a jet.

As shown in figure 1.3a, for jets representative of heavy boson decay near production threshold, the momentum spectrum is dominant at around 10 GeV as for heavy boosted jets with a more complex event topology, the momentum distribution is still dominant to low energies but present a tail to much higher energies. Thus the non-linearity has only little
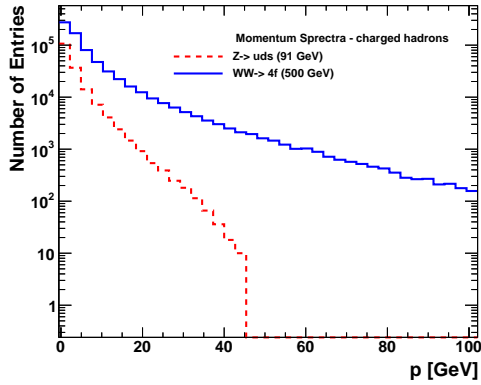
(a) Momentum spectrum of charged hadrons.
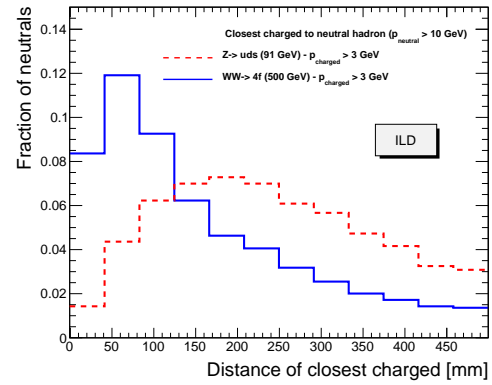
(b) Minimal distance between a charged and neutral particle at the front face of the ECAL.

Figure 1.3 – a) Momentum distribution for charged particles in simulated $e^+e^- \to Z/\gamma \to q\bar{q}$ with $q = u,d,s$ at $\sqrt{s} = 91$ GeV and $e^+e^- \to W^+W^- \to q\bar{q}q\bar{q}$ where q is a quark at $\sqrt{s} = 500$ GeV . b) Distribution of distances to the closest charged track for neutral particles produced in $Z/\gamma \to q\bar{q}$ and $W^+W^- \to q\bar{q}q\bar{q}$ processes measured at the front face of the electromagnetic calorimeter in the ILD detector.

effect there. It is still relevant to understand the different effects of the reconstruction at single particle level.

A complementary study was to look at the minimal distance between a charged and neutral particle for theses different physics processes. The figure 1.3b shows that for low energy jets the mean minimal distance (measured at the front face of the SiW-ECAL) between a charged and neutral hadron is around 180 mm thus in this context, showers are well separated. But at higher energies where density is higher, typical distances of 50 mm need to be resolved. This situation can become relevant in the contribution of confusion to the jet energy resolution. In this case, the use of timing information could help to separate nearby showers and improve the pattern recognition.

1.3.3 Timing cut effects on hadronic showers in ILD detector

In this section, the effect of timing cuts on hadronic showers is investigated. The study was performed using the ILCSOFT framework for reconstruction and a personal MARLIN

processor for analysis. In order to study the effect of timing on hadronic shower properties, the initial study was performed assuming a perfect timing resolution (i.e. the timing information is the Monte-Carlo truth). In a following step, several timing resolution were used to assess different scenarios. The smearing of the time was done by randomly sampling a normalised gaussian centered in 0 ns with a timing resolution denoted $\sigma_t = 0.4$, 1 and 8 ns.

1.3.3.1 Without time smearing

The following section present results of timing cuts assuming a perfect time resolution. To avoid any effects of clustering and Pandora, the study was performed on the calorimeter hit level. Several shower observables were looked as function of the time cut for energies from 5 GeV to 90 GeV K_L^0.

1.3.3.2 With time smearing

1.4 Benchmarking of fast simulation

1.4.1 SGV: fast simulation software

1.4.2 Particle Flow parametrisation

1.4.3 Benchmarking against full ILD Simulation

References

[1] ILCSoft web portal. URL http://ilcsoft.desy.de/portal.

[2] Oskar Hartbrich Daniel Jeans. Realistic calorimeter hit digitisation in the ILDCaloDigi processor. 2015.

[3] R. Fruhwirth. Application of Kalman filtering to track and vertex fitting. *Nucl. Instrum. Meth.*, A262:444–450, 1987. doi: 10.1016/0168-9002(87)90887-4.

[4] F. Gaede. Marlin and LCCD: Software tools for the ILC. *Nucl. Instrum. Meth.*, A559: 177–180, 2006. doi: 10.1016/j.nima.2005.11.138.

[5] Frank Gaede, Ties Behnke, Norman Graf, and Tony Johnson. LCIO: A Persistency framework for linear collider simulation studies. *eConf*, C0303241:TUKT001, 2003.

[6] Oskar Hartbrich. *Scintillator Calorimeters for a Future Linear Collider Experiment*. PhD thesis, Hasylab, DESY, Hamburg, 2016.

[7] M. A. Thomson. Particle Flow Calorimetry and the PandoraPFA Algorithm. *Nucl. Instrum. Meth.*, A611:25–40, 2009. doi: 10.1016/j.nima.2009.09.009.

Acknowledgments