PRAMANA — journal of physics | © Indian Academy of Sciences | Vol. 69, No. 6 December 2007 pp. 1109–1114

Event reconstruction with MarlinReco at the International Linear Collider

O WENDT*, F GAEDE and T KRÄMER
DESY, 22603 Hamburg, Germany
*E-mail: oliver.wendt@desy.de; frank.gaede@desy.de; thomas.kraemer@desy.de

Abstract. After an overview of the modular analysis and reconstruction framework Marlin an introduction on the functionality of the Marlin-based reconstruction package MarlinReco is given. This package includes a full set of modules for event reconstruction based on the particle flow approach. The status of the software is reviewed and recent results using this software package for event reconstruction are presented.

Keywords. Linear collider; simulation; software tools; event reconstruction; particle flow.

PACS Nos 07.05.Kf; 07.05.Tp; 29.40.Vj; 29.85.+c

1. Introduction

The International Linear Collider (ILC) will be the next machine beyond the LHC. It allows to explore the physics at the 500 GeV to 1 TeV energy scale with high precision. Sophisticated simulation and reconstruction software supports the ongoing development and optimisation of a detector for the ILC. Figure 1a shows a schematic overview of the core software chain used for the studies of the large detector concept (LDC), one of the four current detector proposals for the ILC [1]. This chain consists of two major parts:

1. The Geant4-based simulation of the detector response, Mokka [2,3].
2. The digitisation of simulated data, event reconstruction as well as analysis provided by different modules of the Marlin framework [3,4].

The event data is shipped through the software chain using LCIO [5] between different programs. Geometry-related data needed by the reconstruction is provided by the full detector simulation and can be accessed via Gear [3].

2. Marlin and MarlinReco

Marlin is a modular C++ application framework based on LCIO for the analysis and reconstruction of ILC data. Marlin provides the main program with the event

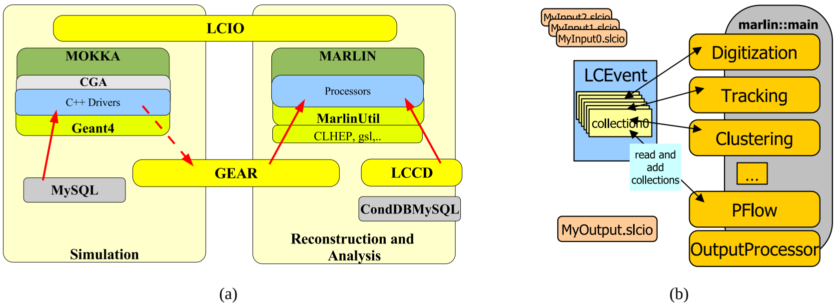

Figure 1. (**a**) LDC simulation and reconstruction framework, (**b**) structure of Marlin.

loop and a mechanism to call modules, so-called processors, to carry out specific tasks. These tasks can be as simple as filling histograms or as complex as a full event reconstruction. As an example, the full chain of an event reconstruction with processors for digitisation, tracking etc. is shown in figure 1b. The framework is reading data event by event creating an *LCEvent* which is used to transfer the data from processor to processor during a Marlin run. An *LCEvent* consists of a set of collections holding specific objects like hits, tracks, cluster etc. Processors only have the permission to read and add information to ensure the consistency of the data. Program steering is done via an XML file allowing to hand over processor parameters, specifying the order of the processors or exchanging processors without recompilation. The package MarlinReco is a specific set of processors for a complete event reconstruction system, based on the particle flow concept. Version 00-02 contains the following processors:

Tracker hit digitisation: For the vertex system there are two different digitisers available. A simple digitiser translates simulated tracker hits into tracker hits, without modifications. A more sophisticated digitiser takes the deposition and transfer of charge in silicon as well as the geometry into account [6]. In the time projection chamber (TPC) a Gaussian smearing of the simulated hit positions in r-ϕ and z is done to account for the intrinsic chamber resolution. Their parameters are obtained from Gear.

Calorimeter hit digitisation: There are two different digitisers for the electromagnetic and hadronic calorimeters (ECAL and HCAL). The first provides calibration, low energy hit rejection and various sampling fractions for the different regions of the calorimeter. The second has the capability to merge neighboring cells into larger cells. This feature allows the variation of the cell size in a simple way. Both digitisers are able to treat hits in analogue and digital calorimeters.

Tracking: There are two tracking processors. The first is based on algorithms taken from LEP providing full tracking in the TPC with energy loss and multiple scattering [7]. The hits of the inner silicon detectors can be included in the track fit, using the tracks reconstructed in the TPC as seeds. The second processor provides a stand-alone pattern recognition procedure for the vertex detector [6].

Clustering: One of the central parts in the particle flow approach is a sophisticated procedure to assign calorimeter hits to the proper reconstructed particle and to minimise the 'confusion' between adjacent particles. This so-called clustering is done by the 'Trackwise Clustering' [8] algorithm. It only relies on spatial information of the calorimeter hits with minimal dependence on the detector geometry. It is applicable to digital and analogue calorimeters as well as to different detector designs.

Particle flow: MarlinReco's particle flow processor 'Wolf' extrapolates tracks into the calorimeter and matches them to clusters by a proximity cut taking into account the detector geometry. In addition, a simple particle identification is done by calculating the fraction of energy in the ECAL and the HCAL. After that, a collection of reconstructed particles is created where the four momenta of charged particles are determined by the track parameters. The four momenta of neutral particles are calculated from the clusters only.

Track and cluster cheater: Processors allowing the assignment of hits to tracks and clusters, using Monte Carlo information only, are provided. To obtain the track parameters, either a simple helix hypothesis is fitted to the tracker hits or the information is taken from the Monte Carlo directly.

Analysis: There are processors to calculate the thrust axis and value (Tasso and JETSET algorithm [9,10]) as well as the sphericity and aplanarity of an event. In addition, a multi-algorithm jet finding processor is available [11].

Calibration: This processor calculates the calibration constants for the calorimeter by the method proposed in [12]. It is based on the energy conservation law giving an upper limit for the energy sum in all the cells of the calorimeter.

MarlinReco is based on the package MarlinUtil combining utility and helper classes and by the client-server based event display system CEDViewer. RAIDA, a ROOT implementation of the AIDA interface, is available [13,14]. Due to the modular structure and the well-defined data structures, alternative algorithms (Magic, PandoraPFA [15,16]) can easily be included in MarlinReco. All software packages as well as more detailed documentation can be accessed via [3].

3. Event reconstruction

Here the MarlinReco-based event reconstruction is tested and the dependence of the performance of the particle flow on basic geometric properties of the detector is studied. For this purpose the full detector simulation using Mokka v05.04 and event reconstruction with MarlinReco is done with four classes ($\gamma/Z^0 \rightarrow q\bar{q}$, WW and $Zh \rightarrow 4$ jets as well as $t\bar{t} \rightarrow 6$ jets) of events at four different center-of-mass energies (91.2, 360, 500 and 1000 GeV). Four different layouts of the LDC and two values of the magnetic field have been chosen to optimise the detector. For the variation of the detector geometry two detector models, LDC00Sc and LDC01Sc provided by Mokka, with different sampling structures in the ECAL are chosen. For each model two sizes of the TPC, determined by their outer radius (R_{TPC}) and length from the nominal IP to the end plane of the TPC (L_{TPC}), are constructed ((A)

Table 1. Layouts of the LDC simulated with Mokka v05.04. The four detector geometries are available with a magnetic field of 3 and 4 T.

Model variation	LDC01Sc		LDC00Sc	
	(A)	(B)	(A)	(B)
R_{TPC} (mm)	1380	1580	1690	1890
L_{TPC} (mm)	2000	2200	2730	2930

and (B) in table 1). This results, together with the two values of the magnetic field (3, 4 T), in eight detector lay-outs. Table 1 summarises the available geometries. Computing and data storage for simulation and reconstruction has been done using GRID resources. Meta information about the simulated data as well as the logical names to access the files are available through a database [3]. For this study Monte Carlo information has been used to perform the pattern recognition in the tracking system. A helical fit is applied afterwards to obtain the track parameters.

In figure 2a the distribution of the invariant mass of Z^0 measured with $Z^0 \rightarrow uds$ events at $\sqrt{s} = m_Z$ for the detector LDC00Sc (A) with a magnetic field of 4 T is shown. Due to the difference of the tails compared to a Gaussian distribution the root-mean-square (RMS) is not an appropriate measure of the width of the peak and therefore of the performance of the reconstruction. Hence, (1) the RMS is calculated with the bins around the maximum bin containing 90% of the events (RMS_{90}) [16] and (2) the sum of two Gaussian functions, one for the central part and one for tails, is fitted to the mass distribution. The width of the central Gaussian (σ_{c}) is the measure of the width of the peak (see figure 2a). The results of both methods are quoted to show the performance of the reconstruction. They also act as an indicator in the process of detector optimisation. The results for the geometry listed in table 1 are shown in figure 2b. To study the performance at higher energies, a simple analysis of $t\bar{t} \rightarrow 6$ jets at $\sqrt{s} = 500$ GeV has been performed by calculating $\Delta E_{\text{reco}} := \sum_i E^i_{\text{reco}} - \sum_i E^i_{\text{avail}}$ for each event. In the first part of ΔE_{reco} the energies of all reconstructed particles are added up, while in the second part the energy sum of all Monte Carlo particles which pass the acceptance cut $\theta > 0.1$ and which are not neutrinos is calculated. This results in $\Delta E_{\text{reco}} = 25.2$ GeV which is about a factor of two larger than the pure calorimeter resolution given by $\Delta E_{\text{calo}} := \sum_i E^i_{\text{calo}} - \sum_i E^i_{\text{avail}} = 12.6$ GeV, where the first part of ΔE_{calo} adds up the energy of all calorimeter cells [12]. One reason for this decrease of performance compared to $Z^0 \rightarrow uds$ at $\sqrt{s} = m_Z$ is the misassignment of hits due to overlaps of showers in the calorimeter. The Fortran-based simulation and reconstruction package Brahms has shown that it is possible to reach energy resolutions of about 9 GeV for $t\bar{t}$-events at $\sqrt{s} = 500$ GeV following the particle flow concept [3,17].

4. Conclusions

For the determination of σ_{c} for $Z^0 \rightarrow uds$ at $\sqrt{s} = m_Z$ MarlinReco comes close to the performance goal of the jet energy resolution at the ILC ($\sigma_E/E = 0.30/\sqrt{E}$

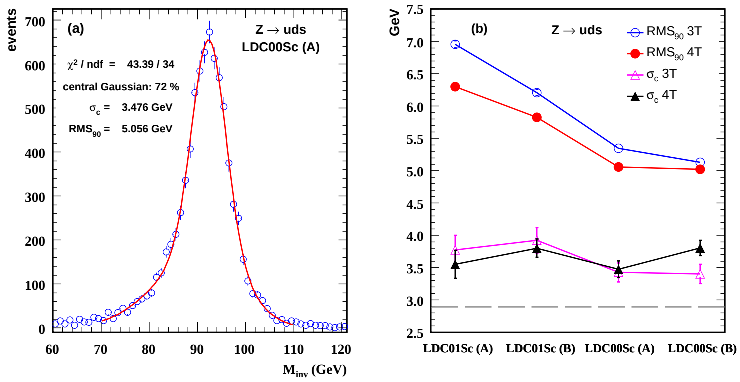

Figure 2. Reconstructed invariant mass of $Z^0 \to uds$ at $\sqrt{s} = m_Z$ fitted with a sum of two Gaussians (**a**), performance of MarlinReco (RMS_{90} and σ_c) for different detector geometries (see table 1) and magnetic fields (**b**).

corresponding to $\sigma_E = 2.9$ GeV at $\sqrt{s} = m_Z$) but no significant dependence on the detector geometry is observed. The results of the RMS_{90}-method are considerably larger but are showing a clear dependence on the detector geometry. In addition, this dependence follows the expectation, i.e. the resolution increases with a larger detector and a larger magnetic field (see figure 2b). The analysis of $t\bar{t} \to 6$ jets at $\sqrt{s} = 500$ GeV shows that improvements in MarlinReco are necessary, especially for high center-of-mass energies. Nevertheless, MarlinReco provides the full chain of event reconstruction following the particle flow concept.

Acknowledgments

We like to thank all members of the DESY-FLC software group: H Albrecht, S Aplin, T Behnke, P Krstonosic, D Martsch, V Morgunov, J Samson and A Vogel.

References

[1] LDC web page, http://www.ilcldc.org
[2] Geant4 Collaboration: S Agostinelli *et al*, *Nucl. Instrum. Methods* A506, 3 (2003)
Geant4 Collaboration: J Allison *et al*, *IEEE Trans. Nucl. Sci.* **53**, 1 (2006)
[3] ILC software portal, http://ilcsoft.desy.de
[4] F Gaede, *Nucl. Instrum. Methods* **A559**, 177 (2006)
[5] F Gaede, T Behnke, N Graf and T Johnson, eConf **C0303241**, TUKT001 (2003)
[6] A Raspereza, Simulation studies of VXD performance, to appear in the *Proceedings of LCWS06*, Bangalore, March 2006
[7] T Behnke *et al*, LC-DET-2001-029 *In *2nd ECFA/DESY Study 1998-2001* 1758-1787*
[8] See http://www.desy.de/~rasp/Raspereza_pfa.pdf for more information

[9] S Brandt, C Peyrou, R Sosnowski and A Wroblewski, *Phys. Lett.* **12**, 57 (1964)
[10] T Sjostrand, *Comput. Phys. Commun.* **82**, 74 (1994)
[11] S Yamashita, OPAL Technical Note TN579, 3rd November 1998
[12] V Morgunov, Calorimeter energy calibration using the energy conservation law, to appear in the *Proceedings of LCWS06*, Bangalore, March 2006
[13] ROOT web page, http://root.cern.ch/
[14] AIDA web page, http://aida.freehep.org/
[15] C G Ainsley, LC-TOOL-2004-015 *Prepared for International Conference on Linear Colliders (LCWS 04)*, Paris, France, 19–24 Apr. 2004
[16] Mark A Thomson, Particle flow calorimetry at the ILC, to appear in the *Proceedings of LCWS06*, Bangalore, March 2006
[17] S V Chekanov and V L Morgunov, *Phys. Rev.* **D67**, 074011 (2003)