Eur Phys J C **33**, s01, s1038–s1040 (2004)
Digital Object Identifier (DOI) 10.1140/epjcd/s2004-03-1805-1

EPJ C direct
electronic only

ATLAS reconstruction software

David Rousseau, on behalf of the ATLAS collaboration

Laboratoire de l'Accélérateur Linéaire, Orsay

Received: 28 November 2003 / Accepted: 10 March 2004 /

Abstract. An overview of the atlas reconstruction algorithms is given, with their typical physics applications and the future development path. The developers environment is also briefly described.

Introduction

With the completion of the first data challenge, which meant processing several Terabytes of simulated data in worldwide distributed sites, ATLAS reconstruction software has reached a mature stage. Through years of development first in Fortran, now migrated to C++ in the flexible Gaudi framework, the algorithms are now fairly complete. This contribution describes the main algorithms with typical figure-of-merit in bench mark physics channels as well as paths of future development.

Although this will not be described here, the raw data flow is very detailed, from the front-end board byte-stream to pre-reconstructed quantities. In the same framework and within the same job besides reconstruction detailed trigger simulation and fast detector simulation can be performed, allowing correlations to be studied.

The Atlas detector (described in details elsewhere [1, 2]) is a multipurpose experiment optimized for high energy 14 TeV center of mass energy proton proton collision at the LHC, with data taking start due in 2007. It has overall cylindrical barrel and end-caps geometry. The Inner Detector immersed in a 2 Tesla solenoidal magnetic field provides tracking up to $|\eta|$ =2.5 [3]. Beyond the solenoid coil the liquid argon calorimeter [4] provides electromagnetic calorimetry in the barrel and both electromagnetic and hadronic calorimetry in the end-cap. The electromagnetic calorimetry extend to $|\eta|$ =2.5 and the overall calorimetry to $|\eta|$ =5. Hadronic calorimetry in the barrel is provided by the iron tile scintillator calorimeter. These elements are surrounded by the air toroids of the muon spectrometer [5, 6]. At the design luminosity, $L=10^{34}\text{cm}^2\text{s}^{-1}$, each bunch crossing every 25 ns yields in average 23 minimum bias pp collisions, producing some 1000 very low P_T charged particles (and a similar number of neutrals) in the tracker acceptance. The data taking rate is expected to be 100 Hz, totaling 1 PByte of data per year.

1 Software framework and work model

ATLAS reconstruction is now being developed in the Athena framework, based on Gaudi [7], initially developed by LHCb. Briefly, Athena relies on the separation between data (handled in the Transient Event Store, TES) and algorithms. An algorithm read data (e.g. calorimeter cells) in the TES, process it and write data (e.g. calorimeter clusters). Converters read/write data from/to persistency to/from the TES. The sequencing and configuration of Algorithms are specified at run time through an ASCII "jobOption" file. Binary libraries are loaded at run time as well. This makes the configuration of a job very flexible, so that one can e.g. easily switch between a very detailed simulation of the data flow and a coarser one, or switch between (or even run simultaneously) two different calorimeter clustering algorithms. It also allows an independent development of the circa 300 packages involved in reconstruction, by a hundred people world-wide. The development follows a six month cycle for major validated releases to be used by physicists and Data Challenges. There are also developer releases every three weeks which facilitates the integration effort.

2 Tracking

Track reconstruction is described in details elsewhere [3]. Tracking is performed up to pseudo rapidity 2.5 with three layers of pixels and four layers of stereo silicon strip detectors, followed by a straw tracker providing typically 30 drift time measurements per track, in addition to transition radiation detection capability. The transverse impact parameter resolution at high P_T approaches 10 microns, degraded in the 1-10GeV/c range by multiple scattering. The momentum resolution is multiple scattering dominated at 1.5% up to 20 GeV/c. The addition of high-luminosity pile-up does not degrade significantly the tracking performance (except speed) as the tracker has been optimized to resolve tracks in high P_T jets where the track density is larger.

3 Electron and photon identification

Electromagnetic clusters are searched for with a sliding window algorithms. Rectangle clusters of typically 3 cells in eta (granularity 0.025) and 5 in phi (granularity 0.025) are preferred, for their robustness against pile-up, underlying event and material effects. The larger size in phi allows to recuperate photon conversion and electron bremsstrahlung, it can also be increased if a conversion is reconstructed.

High $\mathrm{P_T}$ electron/photon identification relies on the lateral and longitudinal shower shapes. The finer granularity of the front sample (1/8 in η) allows the rejection of jets with leading neutral pion. Typical jet rejection of 3000 is obtained for a photon efficiency of 80%, This is sufficient for the fake photon contribution not to be dominant for low mass Higgs search in the important $\mathrm{H}\rightarrow\gamma\gamma$ channel.

For electron identification, an electron track (fitted allowing for bremsstrahlung energy loss) is searched for and matched in E/p. Transition Radiation (TR) hits in the straw tracker add a rejection factor of 100 for 95% efficiency to reach overall a rejection of almost 100000 for 70% efficiency. Additional particle isolation requirements using tracking and calorimeter isolation can also be used but need to be tuned according to the channel of interest. For low $\mathrm{P_T}$ (below 7 GeV/c) electron reconstruction, tracks satisfying TR hit identification are extrapolated to the e.m. calorimeter, where the energy deposition around the track impact is tested against the electron shower hypothesis.

The precise energy measurement of the electro magnetic object is complicated by the material (between one third and one radiation length) in front of the calorimeter. Early photon conversion and electron bremsstrahlung cause energy to be deposited in the material or soft electron to be deviated by the magnetic field outside the cluster. The first order effect can be corrected by weighting the energy in the presampler and, for photon, by an explicit reconstruction of the conversion. Typical energy resolution of 1.5% at $\mathrm{P_T}$ =50 GeV/c is obtained, with weak dependance on pseudo-rapidity and luminosity.

The 5 cm RMS longitudinal beam spread blurs the photon pseudo-rapidity estimate. The longitudinal segmentation of the electromagnetic calorimeter allows a standalone estimation of the true dip angle of the particle yielding a $\mathrm{H}\rightarrow\gamma\gamma$ mass resolution of 1.3 GeV/c^2. An independent determination of the z vertex with the tracks, which may be difficult at high luminosity with relatively low $\mathrm{P_T}$ $\mathrm{H}\rightarrow\gamma\gamma$, given the low track multiplicity of the main event, improves the mass resolution by 10%.

4 Muon identification and reconstruction

The muon spectrometer provides a standalone muon identification and measurement from typically three stations in the toroids (fitted with tracking detectors using four different technologies), each capable of reconstructing a 3D segment of the muon trajectory. The efficiency is typically 95%, due to holes for detector support and services. The efficiency drops at very high $\mathrm{P_T}$ (above 500 GeV/c) due to catastrophic energy loss in the calorimeters, for which electromagnetic showering disturbs the pattern recognition. Below 6 GeV/c, the muon energy loss in the calorimeter is of order of its initial energy so that it is not possible anymore to follow the muon in the inhomogeneous magnetic field.

The reconstructed muon is backtracked to the interaction point through the calorimeter, corrected for its estimated energy loss, and combined with its inner detector track in order to improve the momentum resolution for $\mathrm{P_T}$ up to 20 GeV/c. Typical resonances mass resolution obtained is 45 MeV/c^2 for $\mathrm{J}/\psi \rightarrow \mu^+\mu^-$, 2.9 GeV/c^2 for $\mathrm{Z}\rightarrow\ \mu^+\mu^-$ (dominated by the natural width of the Z), 1.5 GeV/c^2 for $\mathrm{H}\rightarrow\ \mathrm{Z}^*(\mu^+\mu^{)}\mathrm{Z}(\mu^+\mu^-)$ with a Higgs boson mass of 130 GeV/c^2.

A complementary approach under study is to extrapolate the inner detector track into the muon spectrometer: it would allow (i) to decrease the identification threshold to 4 GeV/c by using the inner station only (ii) to help sorting out the pattern recognition for very high $\mathrm{P_T}$ muon. Additional identification information from m.i.p. energy deposit in the calorimeter is also investigated.

5 Jet reconstruction

Jet reconstruction is done with the usual cone algorithm, but also with the Kt algorithm, which is less classical in hadron collision experiments. Cells are summed in towers of coarser granularity, possible negative energy from noise fluctuation for which kT would be ill-defined is eliminated by clustering the nearest towers. To optimize the jet resolution, non-linear weights are fitted which uses the longitudinal segmentation of the calorimeter. The best resolution is obtained by using seven sets of weights of the form $w_i(|\eta|) = a_i(|\eta|) + b_i(|\eta|)/E + c_i(|\eta|)\log(E)$, where the index runs on the layers of the electromagnetic and hadronic calorimeters. The $|\eta|$ dependence allow to take into account the varying effective length of the calorimeters, the material distribution and the degraded measurements in the barrel end-cap transition region. The typical resolution of $\sigma_E/E = 65\%/\sqrt(E) \bigoplus 2\%$ is obtained at low pseudo-rapidity $|\eta| \leq 0.8$.

Work has started on an energy flow algorithm which, after building clusters in the calorimeters with a nearest neighbor approach, attempt to classify them as electromagnetic or hadronic and bring them to the correct scale to take into account the non-compensating nature of the calorimeter (by default anything in the LAr e.m calorimeter is brought to the e.m scale). Identified electrons and muons are properly subtracted from the calorimeter energy, and track measurements are used as well. The goal is to improve on the jet resolution by a proper particle hypothesis assignment.

6 Missing transverse energy

The reconstruction of the missing transverse energy is essential for the reconstruction of channels with a neutrino or other hypothetical invisible particles. The current algorithm is to compute the visible transverse momentum in the calorimeter from all cells above a threshold expressed in sigma of the expected noise. The resolution scales with the square root of the visible transverse energy: $\sigma_{P_x} = 0.48\sqrt{\sum E_\mathrm{T}}$ and is fairly independent of the physics channel (e.g $\mathrm{W}\rightarrow\ell\nu_\ell$ or $\mathrm{A}^0 \rightarrow\tau^+\tau^-$) provided this dependence is taken into account. Under study is the calorimeter layer weighting similar to the jet, and the possibility to compute the visible transverse energy from the energy flow reconstruction.

The missing transverse momentum can also be used as a vector to reconstruct events with two unseen particles under certain conditions. A typical example is $\mathrm{A}^0 \rightarrow\tau^+\tau^-$ where assuming that each neutrino is collinear to the τ decay products and using the missing transverse momentum leads to a sufficiently constrained system. A typical A^0 mass resolution of 10 $\mathrm{GeV/c}^2$ is obtained for $\mathrm{b\bar{b}A}^0(\tau^+\tau^-)$ with $\mathrm{m_A}$ =450 $\mathrm{GeV/c}^2$.

7 b-jet tagging

b-jet tagging is an important tool for low mass Higgs search but also in the long list of decay channels where b quarks are involved. b-jet tagging relies on the combination of the transverse and longitudinal track impact parameter (IP) of the tracks of a jet after severe quality requirements to avoid fake displacements. Typical u-jet rejection of 100 for $\mathrm{H}\rightarrow \mathrm{b\bar{b}}$ with $\mathrm{m_H}$ =120 $\mathrm{GeV/c}^2$ and 60% b-jet efficiency is obtained. The rejection degrades at larger pseudo-rapidity mainly because of material effects, leading to deterioration of impact parameter resolution and producing more high IP tracks from conversion and hadronic interaction. The rejection is optimum for jets in the $\mathrm{P_T}$ range 50 to 150 GeV/c: multiple scattering dominates below, while above, the narrowing of the jet affects pattern recognition, the fraction of B-track is lower and the fraction of fake high IP track larger. When the $|\eta|$ and $\mathrm{P_T}$ dependence are taken into account it is found that the topology of the event has little effect on the b-tagging performance, so that b-tagging performance is similar on complex events such as $\mathrm{t\bar{t}\,H}\rightarrow \mathrm{W}(\ell\nu_\ell)\mathrm{bW}(\mathrm{q\bar{q}})\mathrm{\bar{b}H}(\mathrm{b\bar{b}})$.

Although heavy ion physics was not considered in the design of the tracker, some b-tagging capability (light jet rejection 35 at 50 % b efficiency) is preserved despite the high track density: some 10000 tracks in the tracker acceptance, 10 times more than high luminosity pile-up, but a density still comparable to high $\mathrm{P_T}$ jet.

8 τ identification

Hadronic τ decay identification is crucial in the pseudo-scalar super-symmetric Higgs discovery channel $\mathrm{A}^0 \rightarrow \tau^+\tau^-$. The τ being the heaviest lepton often appear in the most copious decay modes of other super-symmetric particles. Hadronic tau decay appear as a very narrow isolated jet (which can be estimated thanks to the fine granularity of the front electromagnetic calorimeter sampling) with small track multiplicity. The $\mathrm{P_T}$ dependence of the identification is very strong, given that the width of a QCD jet increases with energy when it decreases for a τ given the higher relativistic boost. Using optimum variable combination allows to choose easily an optimal working point depending on the analysis, for example at 50 GeV/c, a rejection of 1000 at 30% efficiency or 100 at 60% efficiency.

9 Conclusion

A wide variety of algorithms are being developed in order to extract the best physics from the Atlas detector. Studies are so far based on a very detailed Monte-Carlo simulation, with the performances of each detector carefully tuned on test beams data. In 2004, a barrel wedge with all the ATLAS detectors (tracking and calorimetry) and a calorimeter end-cap wedge will be put in test beam. Data will be analyzed with an evolution of the software described in this contribution so that (i) it is made robust against real data peculiarities (mis-calibration and mis-alignments, noisy or dead channels...) (ii) algorithms can be tested in a real environment. A repackaging is on-going so that all algorithms commonalities (“tools”) are factored out, and that algorithms performing the same tasks with different strategies share the same interface. This will ease algorithms improvement and new algorithms development, and their optimization for a variety of physics channels and running conditions. High level trigger algorithms will also share a large amount of tools with the offline reconstruction.

Acknowledgements. I'd like to thank all my ATLAS colleagues who develop the reconstruction software, and in particular for the preparation of this contribution : Nektarios Benekos, Frédéric Derue, Ambreesh Gupta, Michael Heldmann, Anna Kaczmarska, Jean-Francois Laporte, Jessica Levèque, Pavel Nevski, Frank Paige, Gilbert Poulard, Jean-Baptiste de Vivie, Silvia Resconi, Francesco Tartarelli and Monika Wielers.

References

1. ATLAS Collaboration: *ATLAS Detector and Physics Performance Technical Design Report* Vol 1, CERN/LHCC/99-14, 25 May 1999
2. Daniel Fournier: these proceedings
3. Alessia Tricomi: these proceedings
4. Lucia di Ciaccio: these proceedings
5. Susanne Mohrdiek: these proceedings
6. Christoph Amelung: these proceedings
7. M. Cattaneo et al.: in *Proceedings of 2001 Conference for Computing in High Energy and Nuclear Physics, Beijing*