Data Processing and Event Reconstruction for the ATIC Balloon Payload

O. Ganel[1], J.H. Adams Jr.[2], J. Chang[3],T.G. Guzik[4], J. Isbert[4], H.J. Kim[5], S.K. Kim[5], I.M. Koo[5], Y. Kwon[5],B. Price[4],W. Schmidt[3], E.S. Seo[1], R. Sina[1], J.Z. Wang[1],J.P. Wefel[4], E.I. Won[5], and J. Wu[1]
for the ATIC Collaboration
[1] *IPST, University of Maryland, College Park, MD, USA*
[2] *Naval Research Laboratory, Washington, D.C., USA*
[3] *Max Planck Institut fur Aeronomie, Katlenburg-Lindau, Germany*
[4] *Louisiana State University, Baton Rouge, LA, USA*
[5] *Seoul National University, Seoul, S. Korea*

Abstract

The Advanced Thin Ionization Calorimeter (ATIC) experiment is planned to fly on Long Duration Balloon (LDB) flights starting in 2000. ATIC is comprised of a fully active 400-crystal, 1200-channel Bismuth Germanate (BGO) calorimeter, 202 scintillator strips in 3 hodoscopes interleaved with a graphite target and read out by 808 channels, and a 4480-pixel silicon matrix charge detector. In this paper we describe a potential ROOT-based data processing scheme designed to handle the 50 Gbytes of data expected from each LDB flight. We detail the trigger scheme and some event reconstruction algorithms, such as track and charge reconstruction. We also present the data rates, efficiencies and resolutions expected, based on GEANT simulations.

1 Introduction:

Cosmic ray particles are the only proxy we have for studying extra-solar matter. A detailed study of cosmic ray composition and energy spectra can yield hints as to the source of these particles, their acceleration mechanism, and their propagation through the inter-galactic and galactic medium. The all-particle energy spectrum of cosmic rays obeys a rapidly falling power-law over an extremely broad range of energies, from about 10^{10} eV (10 GeV) up to the highest energies measured by ground-based experiments at over 10^{20} eV (Wiebel-Sooth and Biermann, 1998). This spectrum has secondary features, most notably it has been observed to become steeper above about 10^{15} eV - the so-called 'knee'. In the supernova remnant shock-wave acceleration model this is attributed to a sharp cutoff at a magnetic rigidity of about $Z \times 10^{14}$ V, predicting a yet-to-be-verified break in the proton spectrum at 10^{14} eV.

The ATIC experiment (Guzik et al., 1999; Seo et al., 1996) shown in Fig. 1 is a payload under construction, to be flown on LDB flights starting Dec. 2000. The ATIC collaboration intends to investigate the spectral shapes of individual elements from hydrogen to iron, from about 10 GeV to 100 TeV, with a statistical accuracy better than 30% for protons at the highest energy. ATIC will allow a single-instrument measurement of cosmic ray spectra over almost 4 decades in energy, with an expected energy resolution of about 40% for vertically incident protons interacting in the target (Seo et al., 1996). ATIC is comprised of a 25 cm deep BGO calorimeter preceded by a 30 cm thick graphite target with three interleaved scintillator strip hodoscopes and a silicon matrix charge detector, with a total of 6488 readout channels.

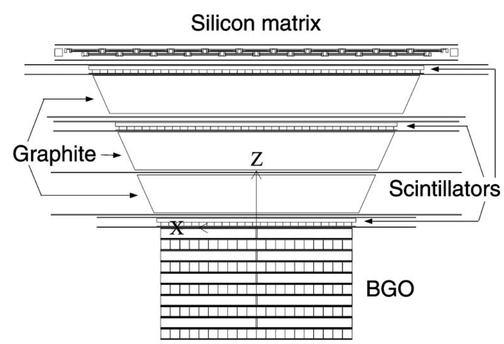

Figure 1: Schematic of the ATIC detector.

2 Storage Capacity, Data Rates and the ATIC Trigger:

Most triggered ATIC events will have showers under 100 GeV and will thus have only pedestal counts in most channels. To minimize storage requirement such channels will be sparsified, reducing the average event size from 30 Kbytes (including header, etc.) to about 1.25 Kbytes. ATIC will carry two 50 Gbyte disks for on-board data storage and backup. LDB flights range in duration from 8 to 20 days, with a nominal flight being 10 days. ATIC is thus limited to about 50 Hz average data-taking rate. The backup disk will be overwritten with new data if and when the primary disk fills up (e.g. as a result of a longer-than-average flight duration).

With the steep drop of the cosmic ray spectrum, collecting 10 events above 100 TeV will require a sophisticated trigger to avoid swamping the disk with low-energy and/or background events (out-of-geometry, late- or non-interacting, etc.). Such a trigger must be highly efficient, have strong background rejection, and provide a random sampling of lower-energy events. The efficiency must be as energy-independent as possible to avoid biasing spectral-index measurements. To achieve this, the ATIC trigger system will combine several algorithms. Each of these will require at least one hit ($>$0.5 MeV) in each of the three hodoscopes. A low energy trigger (LET) will provide high efficiency starting almost as low as 10 GeV, by additionally requiring only 6 consecutive layers with at least one crystal showing a deposit $>$30 MeV (greater than expected from a single minimum ionizing particle). To avoid the large volume of such events expected, this trigger is planned to be randomly reduced by a factor of (e.g.) 100. For high-energy showers, ATIC will utilize a combination of a vertical high-energy trigger (VHT) for particles traversing the detector from top to bottom, and an angled high-energy trigger (AHT) for showers exiting the calorimeter through the side of the four lower BGO layers. The VHT will require an energy deposit $>$30 MeV in at least one crystal in each of the ten BGO layers. The AHT will require a significant energy deposit ($>$1.1 GeV) in at least one crystal of each of four consecutive layers. To reduce background from non-interacting light nuclei, the VHT and the AHT will each require that at least one crystal in the upper half of the calorimeter to register at least 0.5 GeV. Most heavy nuclei will interact in the target and, if in the geometric aperture, will constitute good events.

Additional triggers have been defined for calibration purposes. These will include LED flashes and random triggers for pedestal and aliveness testing, as well as pass-through (non-interacting protons and helium nuclei) events for inter-crystal and absolute energy calibration. These triggers will be collected during dedicated (e.g. 6-minute) calibration runs planned to take place (e.g.) once every 1 - 2 hours.

3 Data Processing:

The ATIC Flight Data System will record events in a raw data format. This data will be a series of event records of various types (science, calibration, environment, etc.), each with its own header, sizes, addresses, etc. This data will have to be separated into the different data types and processed before analysis can begin. Raw data will be examined for media integrity (identify/correct bad blocks) and files will be time-ordered and merged, streaming different data-types into separate Level 0 data files. Environmental information and LED events included in the data will be used to verify data quality (e.g. identify any dead and/or noisy channels). Level 0 data will then be used to generate time-dependent pedestal subtraction and calibration functions. These functions will be used to translate ADC counts (possibly in 2 or 3 ranges for the same crystal or scintillator strip) to energy deposits in each crystal, strip and pixel. This corrected data will be considered Level 1 data, and will be used to verify the pedestal subtraction and calibration. The event selection and reconstruction algorithms described in Sec. 4 will be used to determine event parameters (e.g. energy, charge, trajectory, etc.) which will be added to each event record, generating a Level 2 dataset for science analysis.

ROOT (Brun and Rademakers, 1997) is a powerful object-oriented data analysis and presentation tool developed at CERN. Written in C++ and incorporating a C++ interpreter, ROOT allows users to define objects and classes appropriate to their experiment. Large scripts can be compiled and dynamically linked in. A powerful graphic user interface allows a variety of functions such as zooming, 3-D rotations, etc. Each user-defined object can be selected on-screen and studied. Analysis tools include multi-dimensional histogramming, fitting, minimization, etc. User routines can be used to read in and write out any data format. For ATIC, a possible set of defined classes would be a master class *event*, and sub-classes *header*, *silicon*, *scintillator*, *bgo* and *track*. A preliminary version of this application has been developed and used to process simulated events. Running on a 400 MHz Pentium II with a Linux operating system, events have been processed at about 170 Hz. Processing a full LDB dataset with a single PC should thus take about 30% of flight duration (up to 7 days).

4 Offline Event Selection and Reconstruction:

While the trigger must be highly efficient, its purity, defined as the fraction of triggered events in geometry and interacting in the target, is not a driving concern so long as the rate is acceptable. An offline event selection will be utilized to reduce background to levels acceptable for analysis. The efficiency of this selection and its purity are important considerations, as is their energy dependence. A preliminary selection algorithm has been developed with this in mind. Since the high energy trigger efficiency is energy dependent below 100 GeV, ATIC will use the LET there (above 97% efficient for 50 GeV and higher). The integrated flux for 10 GeV - 100 GeV is about 50-fold that above 100 GeV, making this possible despite the large reduction factor. With lower energies, containment is also less of a concern, and ATIC may analyze particles interacting in the top two layers of the BGO, in addition to those interacting in the target. Figure 2a shows the trigger and event selection efficiencies as a function of energy for the high energy algorithm. From 100 GeV, trigger efficiency

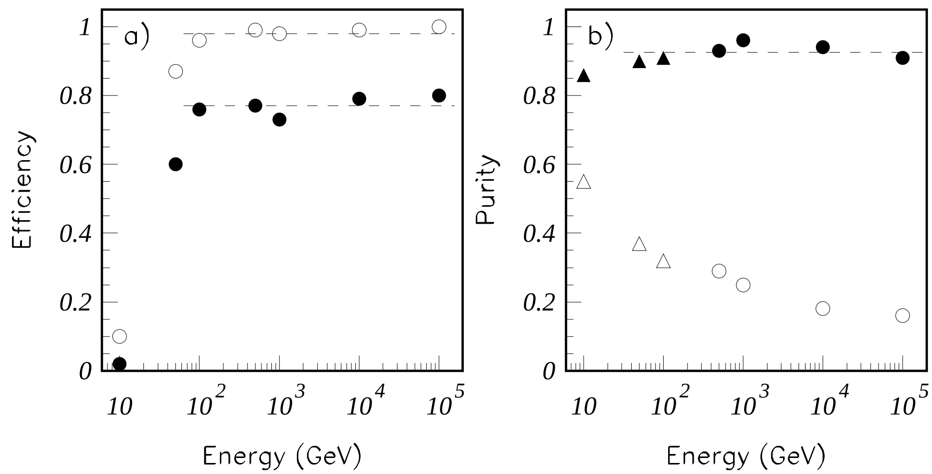

Figure 2: a) Trigger (open circles) and event selection (full circles) efficiencies. b) Trigger (open markers) and event selection (full markers) purities, for low energy algorithms (triangles) and high energy algorithms (circles). Statistical uncertainties are $\sim$1% and smaller than the markers.

is above 96% and nearly energy-independent. Selection efficiency, defined as the fraction of good events triggered and selected, is above 76% and nearly energy-independent from 100 GeV. The efficiency of (pre-reduction) LET and low-energy selection (including events interacting in the top of the BGO) ranges from 72% - 79% for 50 GeV and higher. The two ranges can be inter-calibrated in an overlap region around 100 GeV. As shown in Fig. 2b, while trigger purity drops from about 55% to less than 20% as incident energy increases from 10 GeV to 100 TeV, selection purity is nearly energy-independent from 500 GeV averaging about 94%. Adding in the low-energy algorithm, purity is about 93% from 50 GeV.

After triggered data is recorded and offline event selection is applied, reconstruction algorithms will be used to determine the charge, trajectory and energy of each event. These quantities can then be used to reproduce the incident spectra of each element, allowing determination of spectral indices, relative abundances, etc. The thickness of the ATIC target section (about 0.75 nuclear interaction lengths) and the depth of the calorimeter (about 1.15 nuclear interaction lengths or 22 radiation lengths) are such that the energy deposit in the BGO is a nearly linear measure of the incident particle's energy. GEANT/FLUKA 3.21 (Brun et al., 1984; Arino et al., 1987) simulations show that the response (i.e. the ratio of measured energy to incident energy) is nearly Gaussian in form with a mean decreasing by only 2% (from 0.43 to 0.42) when incident energy increases from 100 GeV to 100 TeV. This can easily be corrected using simulations. The energy resolution is about 33% and almost energy independent for isotropically incident, triggered and selected protons, from 10 GeV to 100 TeV, as shown in Fig. 3. To reconstruct particle trajectory, the shower core is localized using the energy deposit centroid in each BGO layer, providing up to 5 points each in X and Y. Fitting straight lines through these points, the shower axis is extrapolated upwards to the silicon matrix. The silicon hit nearest to the extrapolated track is added to the fit. Hodoscopes showing a significant total signal are deemed to be below the first interaction and are treated as were the BGO layers. Hodoscopes above the first interaction include hits from the primary track and possibly a collection of back-scattered particles. A combinatorial optimization procedure is used to select from among these the best collection of points (Ganel and Seo, 1998). This procedure has been shown to improve on 'BGO only' tracking resolution significantly in about 60% of simulated events. In events with at least two non-BGO points in the track, of which one is in the top hodoscope, the extrapolation resolution to the top of the device is accurate to within 7 mm. In remaining events the full reconstruction does not improve systematically on the 'BGO-only' track, and the resolution is about 5 cm at 100 GeV, improving to about 2 cm from 500 GeV. Defining a circle with a radius three times the extrapolation accuracy around the reconstructed track, we look for the largest signal in any silicon pixel. This is deemed to be the signal from the primary with a possible overlay of back-scattered secondaries. Simulations show that this correctly identifies the incident charge in all but 1% - 2% of events even at the highest energies (Ganel et al., 1999).

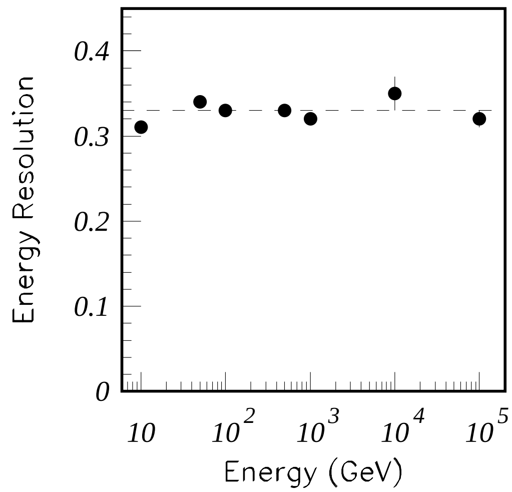

Figure 3: Energy resolution for simulated isotropically incident, triggered and selected protons.

5 Summary:

ATIC is a balloon-borne, calorimeter-based cosmic ray experiment with a finely segmented silicon charge detector, intended to study the energy spectra of individual elements at energies up to 100 TeV with data from four LDB flights. Its trigger rate will be kept under about 50 Hz, with a total of about 50 Gbytes of data expected per LDB flight. Trigger efficiency is above 95% and almost energy independent above 100 GeV. Low energy trigger efficiency is almost energy independent from just above 10 GeV. Event selection efficiency is about 75% and almost energy independent above 100 GeV. Low energy selection efficiency is almost energy independent above 50 GeV. Trigger purity is good enough to accept all interesting events within the maximal rate capability. Selection purity is about 94% above 500 GeV, and somewhat lower up to 100 GeV. For the energy range up to 100 GeV, a randomly reduced LET trigger and selection provide improved energy independence. Fast data processing may be handled with C++ routines through ROOT, from raw data to fully processed and reconstructed Level 2 event records.

References

Arino, P.A. et al. 1987, FLUKA User's Guide, CERN, TIS-RP-190
Brun, R. and Rademakers, F. 1997, Nucl. Instr. Meth. A389, 81
Brun, R. et al. 1984, GEANT User's Guide, CERN DD/EE/84-1
Ganel, O. and Seo, E.S. 1998, Adv. Space Res., in press
Ganel, O. et al. 1999, STAIF-99 Space Tech. & Appl. Int. Forum AIP 1-56396-846-0/99, 272
Guzik, T.G., et al. 1999, 26^{th} ICRC , OG.4.1.03
Seo, E.S. et al. 1996, SPIE Int. Symp. on Opt. Sci., Eng. and Instr. (Denver) 2806, 134
Wiebel-Sooth, B. and Biermann P.L. 1998, Cosmic Rays, in Landolt-Boernstein New Series VI 3C