

The Associative Memory system for the FTK processor at ATLAS

Ricardo Cipriani^{ab}, Saverio Citraro^{*ab}, Simone Donati^{ab}, Paola Giannetti^b, Agostino Lanza^c, Pierluigi Luciano^{ab}, Daniel Magalotti^d, Marco Piendibene^{ab}.

^aUniversity of Pisa, Largo B. Pontecorvo, 3 56127 Pisa, Italy

^bSezione di Pisa INFN, Largo Bruno Pontecorvo 3, 56127 Pisa, Italy

^cSezione di Pavia INFN, Via Agostino Bassi, 6 - 27100 Pavia, Italy

^dUniversity of Modena and Reggio Emilia, Via Università' 4, 41121 Modena, Italy

E-mail: r.cipriani@studenti.unipi.it, saverio.citraro@pi.infn.it,
simone.donati@pi.infn.it, paola.giannetti@pi.infn.it,
agostino.lanza@pv.infn.it, p.luciano@studenti.unipi.it,
daniel.magalotti@pg.infn.it, marco.piendibene@pi.infn.it

Experiments at the LHC hadron collider search for extremely rare processes hidden in much larger background levels. As the experiment complexity, the accelerator backgrounds and instantaneous luminosity grow, increasingly complex and exclusive event selections are necessary. We present results and performances of a new prototype of the Associative Memory (AM) system, the core of the Fast Tracker processor (FTK). FTK is a real time tracking device for the ATLAS experiment trigger upgrade. The AM system provides massive computing power to minimize the online execution time of complex tracking algorithms. The time consuming pattern recognition problem, generally referred to as the "combinatorial challenge", is solved by the AM technology by exploiting parallelism to the maximum level. The Associative Memory compares the event to all precalculated "expectations" or "patterns" (pattern matching) at once and looks for candidate tracks called "roads". Pattern recognition is complete by the time the data have been loaded into the AM devices. We report on the tests of the integrated AM system, boards and chips. The prototype has a large network of long high-speed serial links which has been successfully tested. We report on cooling and the expectations for the system's power consumption.

11th International Conference on Large Scale Applications and Radiation Hardness of Semiconductor Detectors

3-5 July 2013

Auditorium Cassa di Risparmio di Firenze, via Folco Portinari 5, Florence, Italy

*Speaker.

1. FTK system overview

The trigger system at hadron colliders must maintain high trigger efficiency for the physics we are most interested in, while suppressing the enormous QCD backgrounds. A multilevel trigger[1] is an effective solution for this task. The ATLAS trigger system[2, 3] consists of three levels. The hardware Level-1 Trigger quickly locates the regions of interest in the calorimeter and the muon system, operating with output rates up to 100 kHz. The subsequent trigger levels, Level-2 and the Event Filter (EF), are collectively known as the high-level trigger (HLT). They consist of software algorithms running on a farm of commercial CPUs. The Level-2 algorithms may request track information in a Level-1 region of interest while the EF has access to information throughout the entire detector. The final EF output rate is limited to 300-500 Hz.

The FastTracker processor (FTK)[4] will be an important element in triggering and is designed to process information from the silicon tracking detectors of the ATLAS experiment[5]. FTK provides massive computing power to minimize the on-line execution time of complex tracking algorithms. The FTK is highly parallel, with the detector segmented into $\eta - \phi$ towers, each with its own tracking processors. Each processor covers one sixteenth of the detector in ϕ , 22.5° , plus 10° overlap to maintain high efficiency. The η range of each region is divided into four overlapping intervals, for a total of 64 $\eta - \phi$ towers. Consequently, a tower receives only a fraction of the silicon hits, and the Processing Units (PUs) executing track reconstruction have substantially fewer candidates to process. Within each tower, we distribute the high luminosity data on 12 parallel buses at the full 100 kHz Level-1 Trigger rate.

The pattern recognition inside each detector tower is executed by two PUs working in parallel. The time consuming pattern recognition problem, generally referred to as the “combinatorial challenge”, can be solved by the Associative Memory (AM) technology[6] exploiting parallelism to the maximum level: it compares the clusters found in the event (“hits”) to all pre-calculated “expectations” or “patterns” (pattern matching) at once, searching for candidate tracks called “roads”. This approach reduces the typical exponential complexity of the CPU based algorithms into a linear problem.

Figure 1 shows the architecture of the FTK processor for the pixel and microstrip trackers (SCT) of the ATLAS detector. The FTK architecture is organized in a pipeline, with each functional block wait for the data from the previous block and sending the results to the next block. The pixel and SCT data are transmitted from the front end ReadOut Drivers (RODs) to the Data Formatters (DFs) which perform cluster finding. The DFs organize the detector data into the FTK tower structure for output to the core crates, taking the needed overlap into account. The cluster centroids in each logical layer are sent to the Data Organizers (DOs). The barrel layers and the forward disks are grouped into logical layers so that there are 12 layers over the full rapidity range. The DO boards are smart databases, where full resolution hits are stored in a format that allows fast access based on the pattern recognition road identifier, and then retrieved when the AM finds roads with the requisite number of hits. In addition to storing hits at full resolution, the DO also converts them to a coarser resolution, referred to as super-strips (SS), appropriate for pattern recognition in the AM. The AM boards contain a very large number of preloaded patterns, corresponding to the possible combinations for real tracks passing through a SS in each detector layer. The AM is a massively parallel system, as it compares each hit with all patterns nearly simultaneously. When

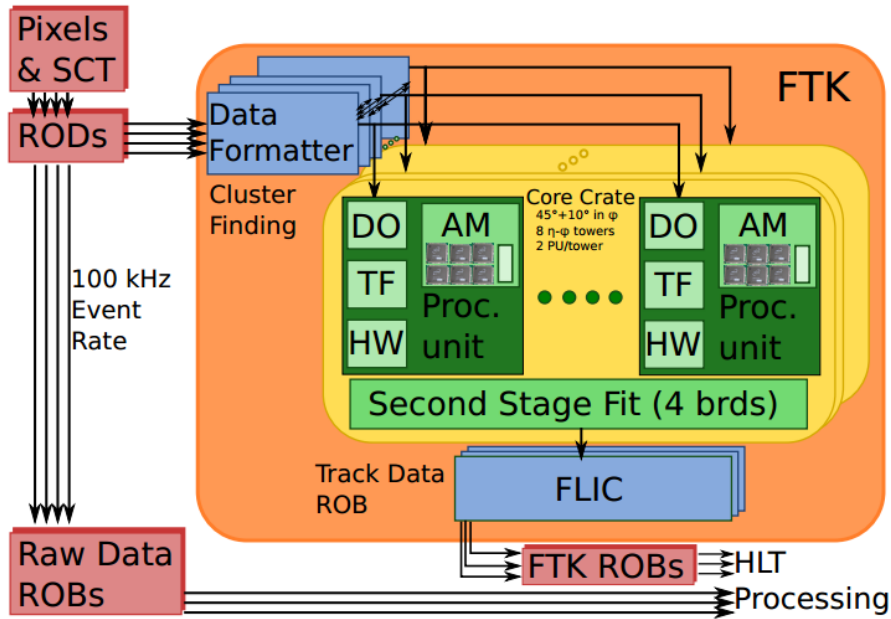


Figure 1: Architecture of the FastTracker processor

a pattern has been found with the requisite number of hit layers, it is then labelled as a road, and the AM sends the road back to the DOs. They immediately fetch the associated full resolution hits and send them and the road to the Track Fitter (TF). Because each road is quite narrow, the TF can provide high resolution helix parameters using the average parameters across the relevant tracking modules and applying corrections that are linear in the actual hit position in each layer. Fitting a track is thus extremely fast since it consists of a series of multiply-and-accumulate steps. In a modern FPGA, approximately 10^9 track candidates can be fit per second. Following fitting, duplicate track removal (the Hit Warrior or HW function) is carried out among those tracks that pass the χ^2 cut. All these functions are executed in a pipeline inside FTK. In summary: FTK has a very large number of devices organized in pipelines connected by thousands of serial links; there are 8200 dedicated custom chips (AM chips) that perform pattern matching and 2000 FPGAs for all other functions.

2. Associative Memory System

The associative memory system carries out pattern recognition at the high silicon detector readout rate by comparing hits at reduced resolution with a very large number of prestored patterns simultaneously. The AM system consists of the Associative Memory chip (AMchip), an ASIC designed and optimized for this particular application, and two types of boards, a VME board (AMB) on which are mounted local associative memory boards (LAMB), a mezzanine that hosts the AMchips[7]. The AMBFTK is a 9U VME board on which 4 LAMBFTKs are mounted. Figure 2 shows the AMBFTK layout, highlighting one of the LAMBs (in yellow). A network of high speed serial links characterizes the bus distribution on the AMBFTK: 12 input serial links (in red) that carry the silicon hits from the high frequency ERNI P3 (in green) connector to the LAMBs,

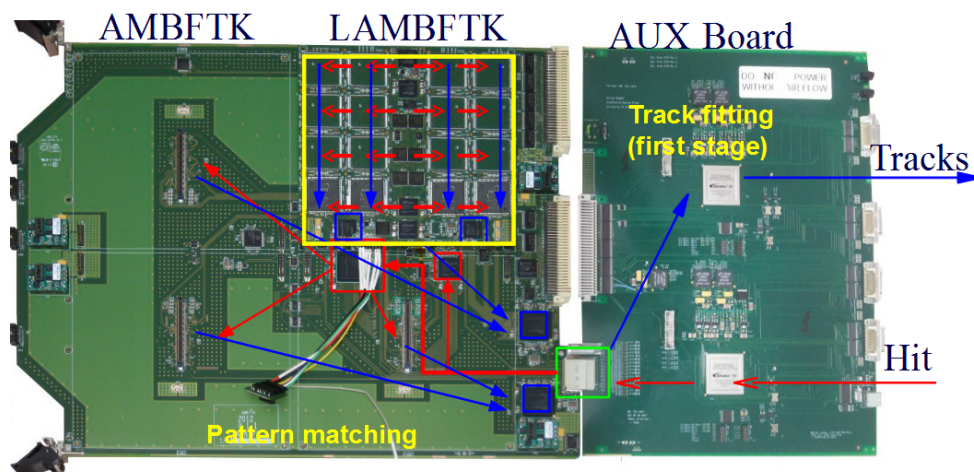


Figure 2: The AMBFTK board showing the data paths to (blue) and from (red) the AUX card that sits behind it. The large yellow square is one of the 4 LAMBs that almost totally cover the AMBFTK.

and 16 output serial links (each blue arrow represents 4 links) that carry the road numbers from the LAMBs to P3. The data rate is up to 2 Gb/s on each serial link. Thus the AMBFTK has to handle a challenging data rate: a huge number of silicon hits must be distributed at high rate (24 Gb/s) with very large fan-out to all AMchips (8 million patterns will be located on 128 AMchips on a single AMBFTK) and a similarly large number of roads must be collected and sent back to the AUX (32 Gb/s). The main AMBFTK logic functions are fanouts for the input data-paths (Figure 2, red squares) and collecting the output from the AMchips (blue squares). There are also diagnostic and test functions; through the VME interface we can spy the dataflow or simulate the arrival of silicon hits in the input[7]. All these functions are configured in 6 Xilinx FPGAs. They are Xilinx Spartan6 FPGAs which have Low-Power Gigabit Transceivers (GTP) that provide ultra-fast data transmission (2Gb/s). The incoming hits are received by the GTPs in the two input FPGAs (red boxes) and saved in large derandomizing FIFOs that are 4k words deep per link. Outgoing road IDs from the LAMBs (4 links/LAMB, final bandwidth of 8Gb/s) are sent to the FPGAs (in the blue boxes). The LAMBFTK and the AMBFTK communicate through an SMD connector placed in the center of the LAMB. Each LAMB contains 32 AMchips, 16 on each side of the board, the roads are collected in a pipeline of 4 chips, as shown in Figure 2 (blue arrows in yellow square). Each LAMB also contains 2 FPGAs named GLUE (blue squares); one GLUE collects the roads from 4 pipelines of AMchips. The FPGAs convert data from parallel format to serial and transfer it to the motherboard through the SMD.

In the AM system most connections between FPGAs are provided by high speed serial links. On each link the AM information is transmitted in words (AM words) whose format depends on the kind of information being processed. The protocol used for the transmission is a simple pipeline transfer driven by control words, for example idle words that are transmitted when no valid data is available. Input valid words in each processing step of the pipeline are pushed into a FIFO dual-clock buffer, needed to synchronize the transmitter and receiver asynchronous devices. All the words that are not identified as control words are pushed into the FIFO (write-enable signal asserted to the FIFO). The transmitter and the receiver use a different clock source, for this reason

GTP interfaces need to be aligned. To provide this we use an alignment word sent roughly every 100 words. To maximize speed, no handshake is implemented on a word-by-word basis. A hold signal (HOLD) is used instead as a loose handshake to prevent loss of data when the destination is busy. If the destination processor does not keep up with the incoming data, the FIFO produces an Almost Full signal that is sent back to the source as the HOLD signal. The source responds to the HOLD signal by suspending data flow. Using Almost Full instead of Full gives the source enough time to stop. Since the source is not required to wait for an acknowledge signal from the destination device before sending the next data word, data can flow at the maximum rate compatible with the link bandwidth even when transit times are long. The standard AM clock frequency is 100 MHz for 16-bit words, which corresponds to 2 Gb/s for serial transmission. An 8b/10b encoding is used in the serial data stream in order to provide effective error detection, i.e. a 16-bit word is transmitted as 20 bits.

3. Tests results

The first AMBFTK prototype and its mezzanine were tested before soldering in AMchip04 prototypes. This was an important test to verify the VME communication and the correct FPGA configuration. For testing the high speed serial link we used an useful Xilinx tool: “Pattern Checker”, in which the GTP transmitter generates a Pseudo-random bit sequences (PRBS) that are commonly used to test the signal integrity of high-speed links. These sequences appear random but have specific properties that can be used to measure the quality of a link. The GTP receiver includes a built-in PRBS checker. The expected pattern is generated from the previous incoming data. The checker counts the number of word (20 bits per word) errors and increments the word error counter by 1 when an error is found in the incoming parallel data. This means that the word error counter may not match the actual number of bit errors if the incoming parallel data contains two or more bit errors. In this way we calculate the receiver’s symbol error rate, which is an estimation of the bit error rate (BER). BER is widely used to measure high speed link performance. The BER represents the ratio of the error bit received to the total bits transmitted. BER can be measured using a bit error rate tester that is composed of a pattern generator and an error detector. In our tests it was lower than 10^{-14} , below the threshold usually required by a modern communication system. After the serial link tests, a few AMchip04s were mounted on the LAMB. To test the complete system, we sent fake hits compatible with the downloaded pattern bank. Sequences of hits were downloaded through VME into the AMBFTK input FIFOs. When the FIFOs were full, their outputs were enabled and the data were sent at 2 Gb/s to the LAMB on the serial links. After the pattern match function was completed inside the AMchips we read the output roads from the AMFTK output FIFOs through VME and saved them in an output file. In parallel we executed a logical simulation of the system with the same fake hits in input. At the end we compared the hardware output with the expectation from simulation. Again the system worked correctly at 100 MHz, as expected.

4. Power and Cooling

Large current must be provided for the core of the AM chips. The AMchip and AMboard

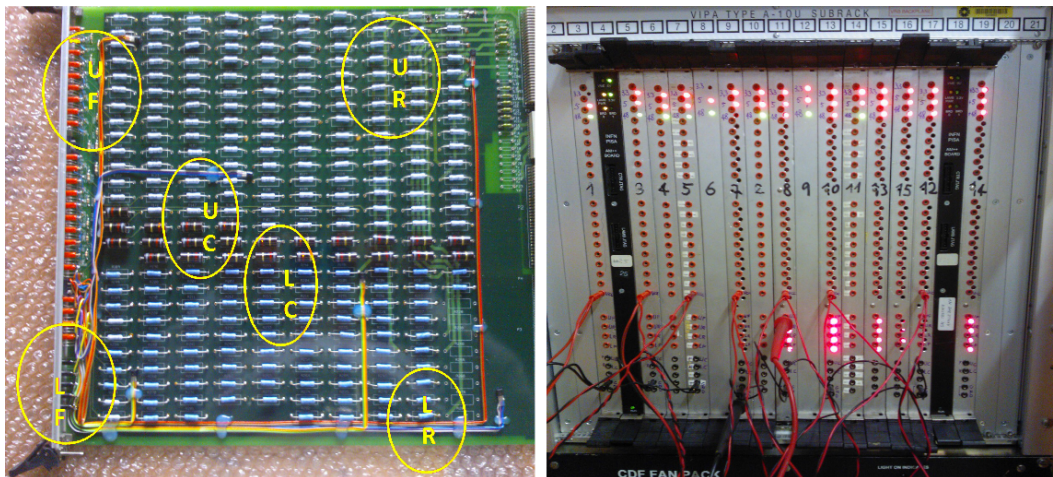


Figure 3: The test stand used for cooling tests: on the right the crate is shown; on the left the positions of thermal sensors are shown. UF, LF monitor the front in the up and low regions. UC, LC and UR, LR do the same for the central and rear part of the board.

design have the power consumption goal of roughly 1 W per chip core (128 W per board) and 120 W for the serialized I/O of the AMchips and fanout chips, for a total of ~ 250 W per board and 4 kw in a 16 AMboard crate. Additional, ~ 500 W is expected from the other boards in the rear crate compartment. We assembled a stand for cooling tests as shown in Fig.3. The front of the crate contains 9U boards with enough resistors to produce a load of 4.5 kW. The uniform distribution of resistors on the boards reproduces the power distribution of the AMBFTK which is uniformly covered by AMchips and fanout chips. Temperature sensors were mounted on these boards. We placed in the crate two old AM boards (black front panel in the figure) containing 4 LAMBFTKs filled as usually to check the cold air flow produced by powerful fans placed below the crate. Two fan packs were tested, one from Wiener and one from the CDF experiment. The results are shown in Fig.4. The Wiener device gives generally worse results, showing higher temperatures in the upper part of the crate, while performing similarly to the CDF fans in the lower part. Until now, the AM boards have been turned on but without receiving input data, so the AM chips were consuming relatively little power due to the clock distribution. The boards contain CDF AMchip03's, whose core consumption is 1.8 W, similar to what we expect from low consumption final chips under design (less than 2 W). In comparison to the final chip, the AMchip03 package thickness is large (> 3 mm), substantially decreasing the air flow. Thus the tests are conservative. The CDF fans appear to coll adequately with a power consumption of 4.5 kW in the front of the crate. When the final AMchip is ready, the test will be repeated with inputs to a final AMboard running at full speed. The final AMchip will have a heat slug to enhance heat exchange with the air. However if this proves insufficient, we will use standard external techniques to increase thermal contact.

5. Conclusion

The design of the first prototype of the Processing Unit for the FTK processor had to deal with the most challenging aspects of this technology. A huge number of detector clusters (“hits”) is

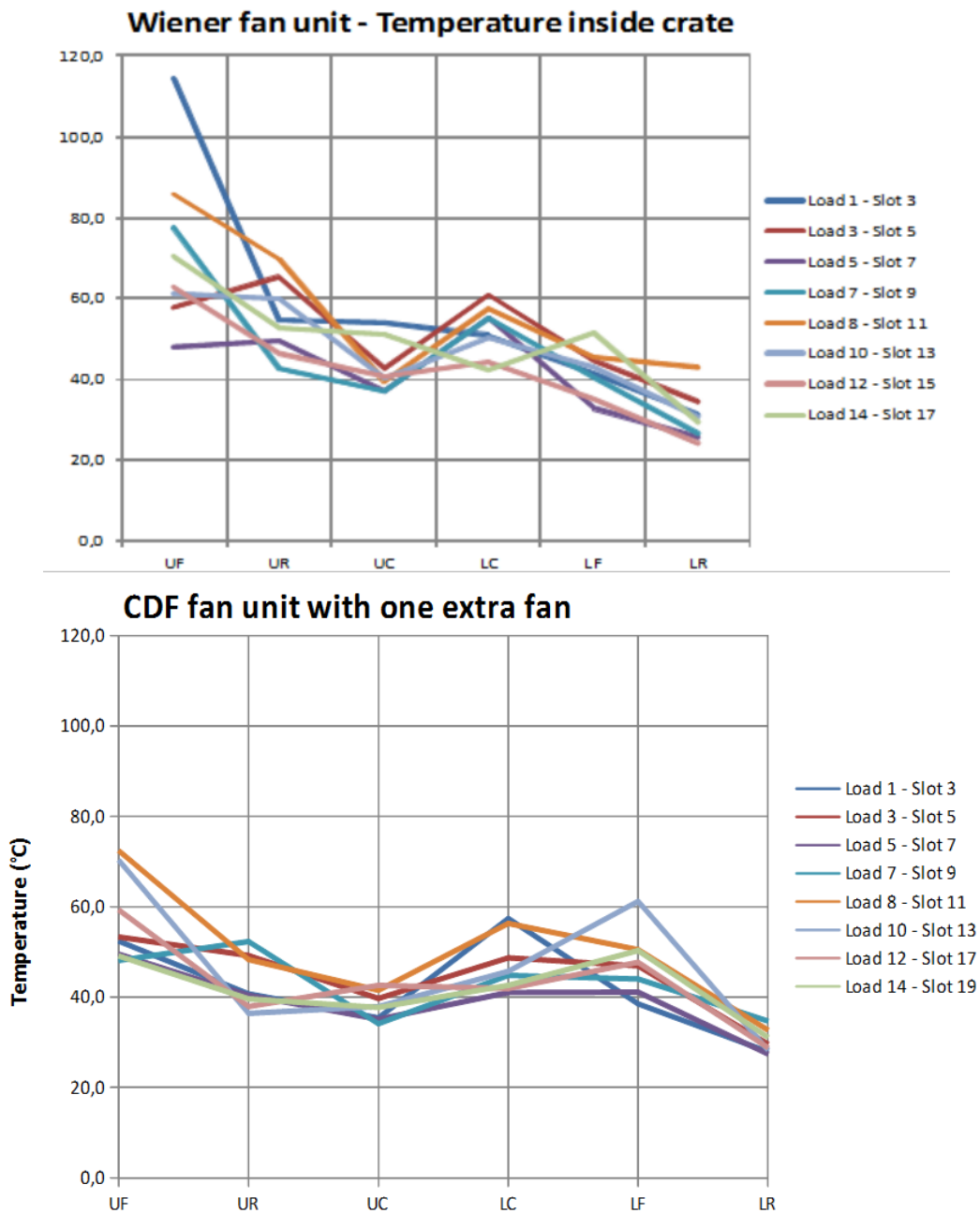


Figure 4: Results of cooling tests with the Wiener and CDF fan units. Each line is associated to a board in a specific crate slot. For each board on the X axis all the sensor are reported.

distributed at high rate and with large fan-out to 8 million patterns located on 128 chips on a single board, and a large number of roads is collected and sent back to the FTK post-pattern recognition functions. The network of high speed serial links used to solve the data distribution problem has been experimentally verified. Cooling tests show that a crate with a 4.5 kW consumption uniformly distributed can be cooled. In addition we have set up a test stand where the final logic and final mechanics can be tested before production.

Acknowledgments

The Fast Tracker project receives support from Istituto Nazionale di Fisica Nucleare; the US National Science Foundation and Department of Energy; Grant-in-Aid for Scientific Research from the Japan Society for the Promotion of Science and MEXT, Japan; the Bundesministerium für Bildung und Forschung, FRG; the Swiss National Science Foundation; and the European community FP7 Marie Curie OIF and IAPP programs.

References

- [1] W. Smith, “Triggering at LHC Experiments”, 2002 Nucl. Instr. and Meth. A, vol. 478, pp. 62–67
- [2] ATLAS Collaboration, “The ATLAS Experiment at the CERN Large Hadron Collider”, JINST, vol.3, pp.437, 2008
- [3] ATLAS Collaboration, “Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics”, arXiv:0901.0512 [hep-ex], pp. 549, 2008.
- [4] A. Andreani et al., “The FastTracker Real Time Processor and Its Impact on Muon Isolation, Tau and b-Jet Online Selections at ATLAS”, 2012 TNS Vol.: 59 , Issue:2, pp, 348 – 357
- [5] ATLAS Collaboration, “The ATLAS experiment at the CERN Large Hadron Collider”, IOP J. Instr. 3 (2008) S08003.
- [6] M. Dell’Orso and L. Ristori, “VLSI structures for track finding”, Nucl. Instr. and Meth. in Phys. Res. – Sect. A 278 (1989) 436–440.
- [7] A. Andreani et al., “The AMchip04 and the processing unit prototype for the FastTracker”, IOP J. Instr. 7, C08007 (2012).