The NA62 GigaTracKer: a low mass high intensity beam 4D tracker with 65 ps time resolution on tracks

G. Aglieri Rinella[a] **D. Alvarez Feito**[a] **R. Arcidiacono**[e] **C. Biino**[e] **S. Bonacini**[a] **A. Ceccucci**[a] **S. Chiozzi**[c] **E. Cortina Gil**[b] **A. Cotta Ramusino**[c] **H. Danielsson**[a] **J. Degrange**[a] **M. Fiorini**[a,b,c,d] **L. Federici**[a] **E. Gamberini**[c,d,a] **A. Gianoli**[c] **J. Kaplon**[a] **A. Kleimenova**[b] **A. Kluge**[a] **R. Malaguti**[c,d] **A. Mapelli**[a] **F. Marchetto**[e] **E. Martín Albarrán**[a,b] **E. Migliore**[e] **E. Minucci**[b] **M. Morel**[a] **J. Noël**[a] **M. Noy**[a] **G. Nüessle**[b] **L. Perktold**[a] **M. Perrin-Terrin** [* a,b,1] **P. Petagna**[a] **F. Petrucci**[c,d] **K. Poltorak**[a] **G. Romagnoli**[a] **G. Ruggiero**[a,2] **B. Velghe**[b,3] **H. Wahl**[d]

[a]*CERN, Switzerland*
[b]*UCL Louvain, Belgium*
[c]*INFN Sezione di Ferrara,Italy*
[d]*Università di Ferrara, Italy*
[e]*INFN sezione di Torino, Italy*

E-mail: mathieu.perrin-terrin@cern.ch

ABSTRACT: The GigaTracKer (GTK) is the beam spectrometer of the CERN NA62 experiment. The detector features challenging design specifications, in particular a peak particle flux reaching up to $2.0\,\mathrm{MHz/mm^2}$, a single hit time resolution smaller than 200 ps and, a material budget of $0.5\%\,X_0$ per tracking plane. To fulfil these specifications, novel technologies were especially employed in the domain of silicon hybrid time-stamping pixel technology and micro-channel cooling. This article describes the detector design and reports on the achieved performance.

---

[1]present address: Aix Marseille Univ, CNRS/IN2P3, CPPM, Marseille, France

[2]present address: Lancaster University, UK

[3]present address: TRIUMF, Vancouver, British Columbia, Canada

[*]Corresponding author.

Contents

1 Introduction

The GigaTracKer (GTK) is the silicon time-resolving hybrid pixel beam spectrometer of the NA62 experiment. The NA62 experiment [1] is installed in one of the north area extraction lines, called P42, of the CERN Super Proton Synchrotron (SPS). The experiment main goal is to measure the branching fraction ($\mathcal{B}$) of the ultra rare $K^+ \rightarrow \pi^+ \nu \overline{\nu}$ decay,

$$\mathcal{B}(K^+ \rightarrow \pi^+ \nu \overline{\nu}) = (8.4 \pm 1) \times 10^{-11},$$

as predicted by the Standard Model (SM) [2], with a 10% precision [3]. This observable provides remarkably stringent tests for the SM and the physics models beyond it [4].

Measuring the branching fraction of an ultra rare decay in such a final state is experimentally challenging. In the previous experiments dedicated to this observable, E787 and E949 [5], a kaon beam was stopped in a target surrounded by detectors measuring the properties of the particles produced by the kaons decaying at rest. At NA62, a decay in flight technique is employed. A beam of $(75 \pm 1)\,\mathrm{GeV}/c$ particles, composed of positively charged $K^+(6\%)$, $\pi^+(70\%)$ and p (23%) with a total nominal rate of 750 MHz, is transported through a 270 m long evacuated tank in which the particles may decay. The tank is instrumented to measure the properties of the beam particles and the decay products, as shown in Figure 1. The momentum, direction, spacial and time coordinates of the beam particles are measured using the GTK. To allow for such a measurement, the beam is
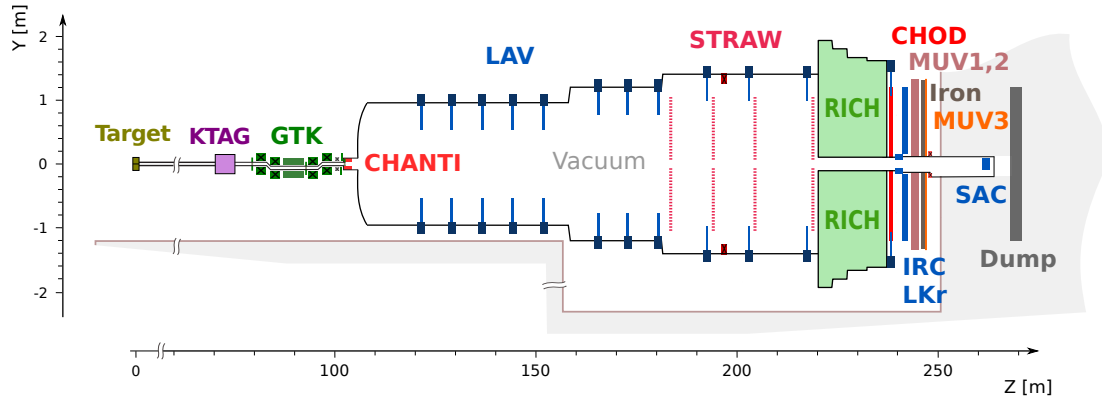
Figure 1. Schematic longitudinal view of the NA62 experiment. The SPS proton beam impinges on a target producing a secondary beam of π^+, proton and K^+. K^+ are identified by a differential Cerenkov detector (KTAG). Beam particles momentum and direction are measured by the GTK. A plastic scintillator based guard ring (CHANTI) is used to veto inelastic collisions occurring in the GTK last station. A fast plastic scintillator based detector (CHOD) is used to trigger the experiment readout when a charged decay product of a beam particle is detected. The decay products are reconstructed with a straw tube based magnetic spectrometer (STRAW), a set of electromagnetic calorimeters (LAV, LKr, IRC and SAC), a Cerenkov detector (RICH), a set of hadronic calorimeters (MUV1 and MUV2) and a plastic scintillator based muon veto (MUV3). See [1] for more details.

focussed and collimated to a transverse size of around $6 \times 3\,\mathrm{cm}^2$ and a maximum particle flux of $2.0\,\mathrm{MHz/mm}^2$.

The GTK is composed of three stations, each made of 18000 time resolving pixels of $300 \times 300\,\mathrm{\mu m}^2$, covering a total area of $60.8 \times 27\,\mathrm{mm}^2$. The stations are installed in the beam pipe in vacuum around two pairs of bending magnets as shown in Figure 2. The magnet pairs displace the beam particles vertically such that they describe, between the two magnet pairs, a trajectory parallel to the undeviated one, and, that they emerge, after the last magnet, as undeviated with respect to the original trajectory. A fifth magnet deflects the beam by 1.2 mrad in the XZ plane to compensate the horizontal displacement due to the STRAW spectrometer dipole magnet. The positions recorded in the first and third GTK stations allow the measurement of the particle direction. This information,
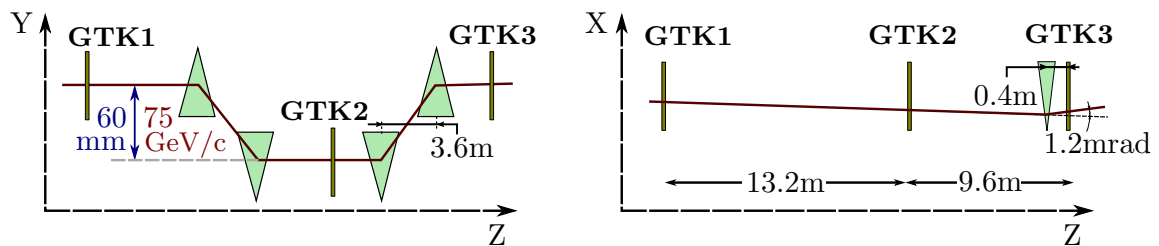

Figure 2. Schematic view of the GTK. The green triangles represent bending magnets. The beam is deflected by 1.2 mrad in the XZ plane to compensate the effect of the STRAW spectrometer dipole magnet.

together with the position measured in the second GTK station, allow the evaluation of the vertical displacement undergone between the magnet pairs. This displacement depends on the momentum rigidity and amounts to 60 mm for 75 GeV/c particles.

The beam particles direction and momentum are used to reconstruct the signal candidates squared missing mass,

$$m_m^2 = |\mathbf{p}_{K^+} - \mathbf{p}_{\pi^+}|^2, \tag{1.1}$$

where $\mathbf{p}_{K^+(\pi^+)}$ denotes the 4-momentum of the parent (decay) particles assumed to be a K^+ (π^+). The $K^+ \to \pi^+ \nu\overline{\nu}$ signal is searched in the two squared missing mass regions on either side of the peak of the $K^+ \to \pi^+\pi^0$ decay, displayed as dashed areas in Figure 3. With a branching fraction of $20.67 \pm 0.08\%$ [6], this decay is 9 to 10 orders of magnitude more abundant than the $K^+ \to \pi^+\nu\overline{\nu}$. Measuring the signal branching ratio with a 10% precision requires therefore a $K^+ \to \pi^+\pi^0$ suppression of 11 orders of magnitude. This suppression can be achieved with an efficient π^0 rejection and an event selection based on the decay squared missing mass. The former provides a background rejection of 7 orders of magnitude [1], requiring the latter to be of 4 orders of magnitude. To obtain this rejection, based on simulations, the resolution on the $K^+ \to \pi^+\pi^0$ squared missing mass is required to be smaller then $0.001 \, \text{GeV}^2/c^4$. This requirement corresponds to an angular and momentum resolution on the beam particle at the exit of the GTK of $\theta_{x,y} = p_{x,y}/p_z = 16 \, \mu\text{rad}$ and $\delta p/p = 0.2\%$.

This performance constrains the maximum amount of material crossed by the beam in the GTK as shown in Figure 5. Beside this loose constraint, the GTK material budget is, above all, critical to control the number of inelastically scattered beam particles. Indeed, while the four magnets are sweeping away fragments produced by particles inelastically scattered in the first two GTK stations, those produced in the third station may enter the acceptance of the other detectors and generate candidates mimicking the signal. Extrapolating from existing technologies [7], an approximative
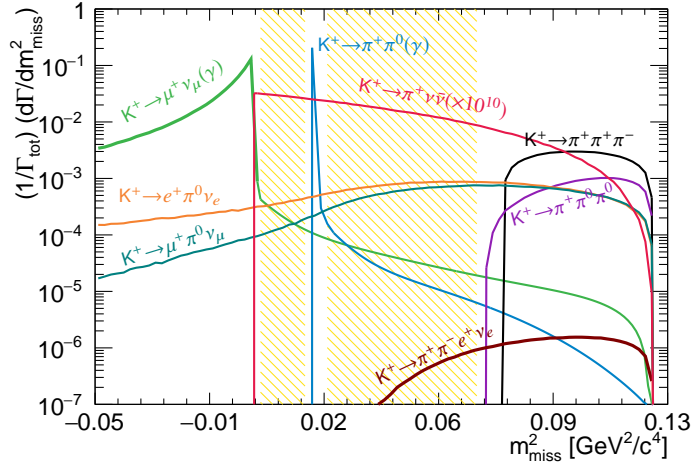

Figure 3. Squared missing mass distribution of the main kaon decay modes and $K^+ \to \pi^+\nu\overline{\nu}$ enhanced by 10 orders of magnitude. The two signal regions are represented by the dashed area.

limit on the amount of material crossed by the beam at each station of 0.5% of a radiation length (X_0) or 0.08% of a nuclear interaction length (λ_i) was considered achievable.

The pixel time resolution is essential to reconstruct the particle tracks inside the GTK. Indeed for each kaon decay considered as valuable by the trigger system, a GTK readout spanning over 75 ns is recorded. Considering the instantaneous nominal beam particle rate of about 0.75 particles/ns, a hit list from more than 50 particles is expected for each record. As the three-station configuration offers almost no geometrical redundancy, track pattern recognition relies mostly on hits time coordinate comparison. With 200 ps time resolution per plane, hits from two consecutive particles are separated by a duration larger than five times the hit resolution. As a result, the resolution on the beam particle arrival time is better than 150 ps which allows unambiguous association of the GTK track with the activities measured in other subdetectors in particular the RICH detector which feature a 90 ps time resolution.

Nonetheless, wrong associations are still possible if some particles are not reconstructed due to detector inefficiency. If such mis-associations occur for a $K^+ \to \pi^+ \pi^0$ decay, the squared missing mass obtained will be incorrect and ineffective to reject the event. The probability for these events to occur is the product of the GTK inefficiency by the probability to have a beam pion or proton forming a good vertex (i.e. with a tracks distance of closest approach and a time difference both within 3 times the detector resolution) with the K^+ decay product. The later being 0.3%, the GTK inefficiency has therefore to be below 3% (corresponding to a station efficiency above 99%) to reach a mis-association probability below 10^{-4} and thus compensating for the inefficiency of the squared missing rejection of $K^+ \to \pi^+ \pi^0$ (nominally 4 orders of magnitude).

Finally, at a beam particle rate of 750 MHz, for an average year of 200 days of operation with a duty cycle of the accelerator of 25%, the beam deposits in the beam center a peak (average) dose of 15.5 (4.5) MRad corresponding to a fluence of 4.5 (1.0) $\times 10^{14}$ 1MeVn_{eq}/cm^2. Despite the irradiation (larger than initially expected [8] due to revisions of the experimental conditions) the performance previously mentioned and reported in Table 1 has to be maintained.

Table 1. Specifications for the GigaTracKer at nominal beam intensity. At 65% nominal beam intensity, corresponding to the maximum operating intensity in 2016 and 2017, the specifications for the rate, flux and fluence are 65% of those reported in the table.

Beam Particle Rate	750 MHz
Peak Beam Particle Flux	2.0 MHz/mm^2
Average Fluence	1.0×10^{14} 1 MeV n_{eq}/cm^2/200 days
Peak Fluence	4.5×10^{14} 1 MeV n_{eq}/cm^2/200 days
Average Integrated Dose	3.5 MRad/200 days
Peak Integrated Dose	15.5 MRad/200 days
Station Efficiency	99%
Momentum Resolution	0.2%
Angular Resolution	16 µrad
Pixel Time Resolution	200 ps RMS
Station Material Budget	0.5% X_0

2 Detector Design

The GTK was designed to fulfil the specifications listed in Table 1. The irradiation level has strong implications on the overall design of the detector. Indeed, the GTK had to be made such that any tracking plane could be promptly replaced once the irradiation had induced significant performance degradation. This consideration imposed a tracking station concept that would be compact, standard and self-contained. Photographs of the station are shown in Figure 4.

(a)   (b)

Figure 4. Photograhs of a GTK station viewed from the sensor side (a) and from the cooling plate side (b).

The $60.8 \times 27\,\mathrm{mm}^2$ sensitive region of the detector is made of a $200\,\mu\mathrm{m}$ thick silicon sensor read out by two rows of five custom made application specific integrated circuit (ASIC) called TDCPix, thinned to $100\,\mu\mathrm{m}$. The choice of the chip pixel size results from a compromise between the number of channels to read out, the particle rate per channel and the angular and momentum resolution. Using $300 \times 300\,\mu\mathrm{m}^2$ pixels, 18000 channels are needed to cover the sensitive region and 180×10^3 particle/s cross the most central pixels. The expected kinematics performance with such pixels are within the specifications and shown as a function of the pixel size in Figure 5.

The hybrid pixel detector (HPD) made by the sensor bonded to the 10 chips is glued on a silicon plate serving both as mechanical support and heat exchanger. The plate is etched to form channels with a cross section of $200 \times 70\,\mu\mathrm{m}^2$ in which a coolant (liquid C_6F_{14}) is circulated to convect the heat produced by the ASICs. In order to minimise the amount of particles inelastically scattered in the detector, the cooling plate thickness was minimised. A thickness of $210\,\mu\mathrm{m}$ was considered technologically achievable. As such, the material budget of the assembly made by the HPD and the cooling plate is about $0.5\%\,X_0$. Figure 5 shows that such material budget allows the achievement of the required kinematics performance.

The assembly made of the HPD and cooling plate is inserted in a countersink in a printed circuit board (Carrier PCB) to which the plate is clamped as shown in Figure 6. The chips are electrically connected with aluminium wire bonds to the PCB to transfer power, clocks, configuration instructions and output data. The PCB is glued into a frame and a vacuum flange such that, once mounted in the beam pipe vessel, the board serves as vacuum feedthrough. In the region outside the vacuum vessel, the PCB is equipped with optical tranceivers communicating through fibers with readout boards 250 meters away in surface counting room. These boards provide clocks and configuration instructions and receive and process the data of all fired pixels. The readout boards
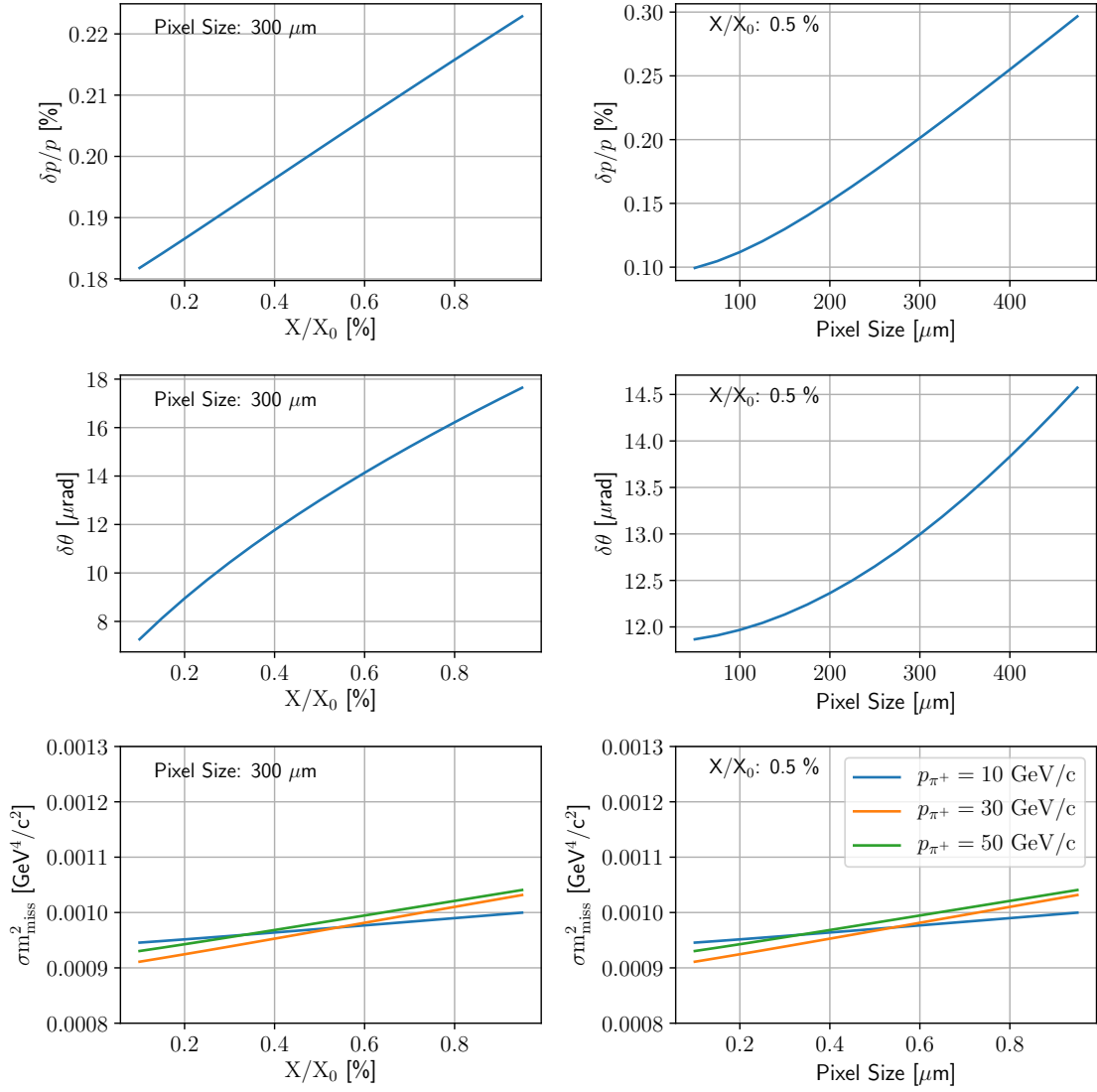
Figure 5. Expected momentum, angular and $K^+ \to \pi^+ \pi^0$ squared missing mass resolution as a function of the material budget and pixel size. The $K^+ \to \pi^+ \pi^0$ squared missing mass resolution is given for different π^+ momenta.
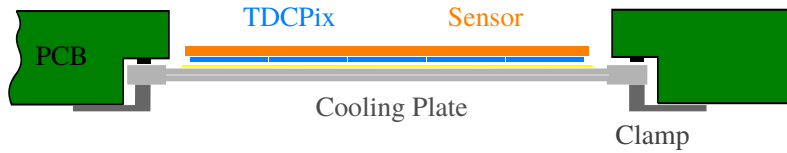

Figure 6. Schematics cross view of the assembly made of the HPD and cooling plate inserted in the PCB Carrier.

are connected to the data acquisition (DAQ) system of the experiment. Upon request of this system, the boards retrieve the information of all the pixels fired around a given reference time and send them to a computer farm which assembles these data. These computers finally send the assembled data packets to the NA62 central computer farm which, in turn, assembles information from all detectors. The next sections describe in details each of the elements constituting the GTK.

2.1 Sensor

The relevant constraints for the design of the GTK sensor are, first, the intense radiation exposition with a fluence reaching 2×10^{14} 1MeVn$_{eq}$/cm^2 in 100 days of run and second, the minimization of the station total material budget fixed to a maximum of 0.5 X$_0$.

As far as the latter is concerned, a sensor thickness of 200 μm was found to be optimal. Such a thickness limits the sensor contribution to the material budget to 0.2% X$_0$ and still provides a sufficiently large signal: 15000 electron-hole pairs corresponding to a most probable charge value of 2.4 fC for a 75 GeV/c particle.

The exposure to radiation has two main consequences: the increase of the sensor leakage current and the reduction of the sensor charge collection efficiency. Both effects can be mitigated by cooling the sensor. Moreover, the loss of charge collection efficiency can be recovered by increasing the bias voltage. At the sensor end-of-life, the bias voltage was foreseen to be raised up to 750 V. Finally, n-in-p type sensors are in principle more resistant to radiation as they can still operate under partial depletion. However, the use of these sensors was still uncertain at the time of the GTK design, hence both p-in-n and n-in-p sensor types were manufactured.

Prototype sensors made of 3×3 mm^2 n-in-p and p-in-n diodes were irradiated up to a fluence of 2×10^{14} 1MeVn$_{eq}$/cm^2, corresponding to 100 days of operation, and showed a satisfying
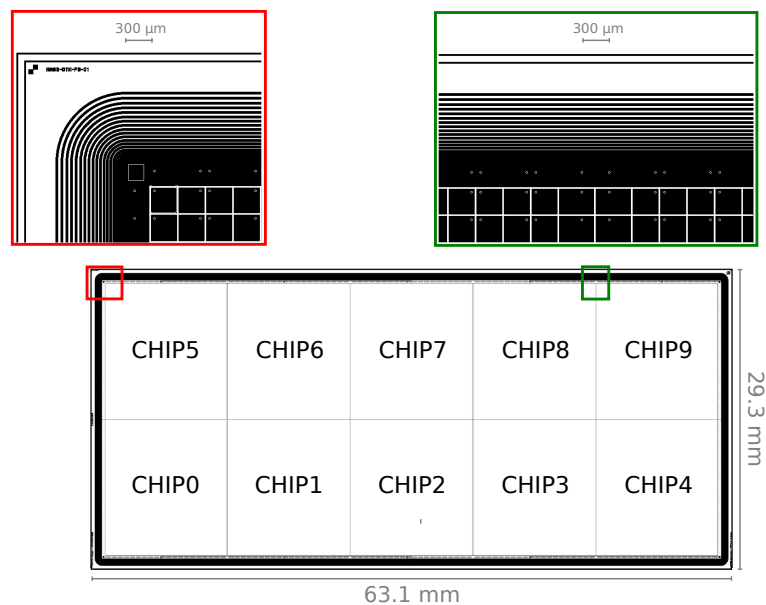

Figure 7. GTK sensor geometry: the top zoomed inlets show the guard rings and the pixel matrix at the corner (left, red) and at the transition between two TDCPix (right green) where enlarged pixels are placed.

performance. The sensor leakage currents and depletion voltages were measured before and after irradiation. The maximum full depletion voltage after 100 days of operation was measured to be around 300 V while the sensors were stably operated at 750 V. The leakage current measurements showed that with a sensor temperature of 5° or less, the sensors can safely be operated for 100 days.

The sensor geometry, shown in Figure 7, is designed to be bump-bonded to two rows of five TDCPix and features extra horizontal spacing between the bonding pads at the chip borders thus forming enlarged pixels ($400 \times 300 \, \mu m^2$ instead of $300 \times 300 \, \mu m^2$). The bump-bonding employs SnAg solder bumps with a diameter of $30 \, \mu m$ and a height of $25 \, \mu m$. A layer of $3 \, \mu m$ of benzocyclobutene was deposited on both the sensor and the chips to prevent discharges between these two elements.

2.2 TDCPix

A dedicated readout chip, the TDCPix [9–12], was designed for the GTK specifications. The chip main functionalities are the sensor hit signal amplification, discrimination, digitisation, time-stamping and the transmission of the resulting digitised data off chip. Most of the specifications reported in Table 1 apply directly to the chip design. The particle rate translates into a maximum pixel hit rate of 180 kHz and implies a careful handling of radiation induced single event upset (SEU).

The TDCPix was designed in a commercial 130 nm technology. The chip is organised in two main parts: first, the 40×45 pixel matrix functioning asynchronously where the sensor hit signals are digitised and second, the end-of-column (EoC) area where those digitised hits are time-stamped and serialised. As such, the analog logic is separated from the digital one which reduces digital switching noise in the amplifying and digitisation circuits. A schematics of the chip is shown in Figure 8(a).

The expected electron collection time for a sensor bias voltage between 100 V and 750 V ranges from 4 ns to 2 ns. At the pixel, the hit induced current on the electrode is shaped by an amplifier with a 5 ns peaking time and then followed by a Time-over-Threshold (ToT) edge discriminator whose threshold can be trimmed. Despite the variation of the electron collection time, the convoluted response of the amplifier and the input signal has a stable rise time (20-80%) ranging from 2.6 ns to 2.1 ns.

At the discriminator output, the digital signal associated to the hit has a length corresponding to the ToT duration. For a fixed hit arrival time, signals with larger amplitudes cross the discriminator threshold earlier than smaller signals, inducing a time-walk up to 2.2 ns, as shown in Figure 9. Being the ToT value directly related to the signal amplitude, it can thus be used to correct for the time-walk. The digital hit signals are shipped on individual differential lines to the EoC using the pre-emphasis technique. Furthermore, groups of five pixels are multiplexed into nine time-stamping units as shown in Figure 8(b). These units consist of two time-to-digital converters (TDC), registering the signal leading and trailing edges. Each TDC is composed of a fine and a coarse time counter of 100 ps and 3.125 ns precision, respectively, with a range of up to 6.4 μs[1]. The hit time-stamp and pixel address are encoded into a 48 bit word stored in a buffer before being shipped out by four

---

[1]The counters are incremented by clocks derived from a 320 MHz clock. For the fine time counter, the clock is divided by 32 corresponding a 97.65625 ps period.
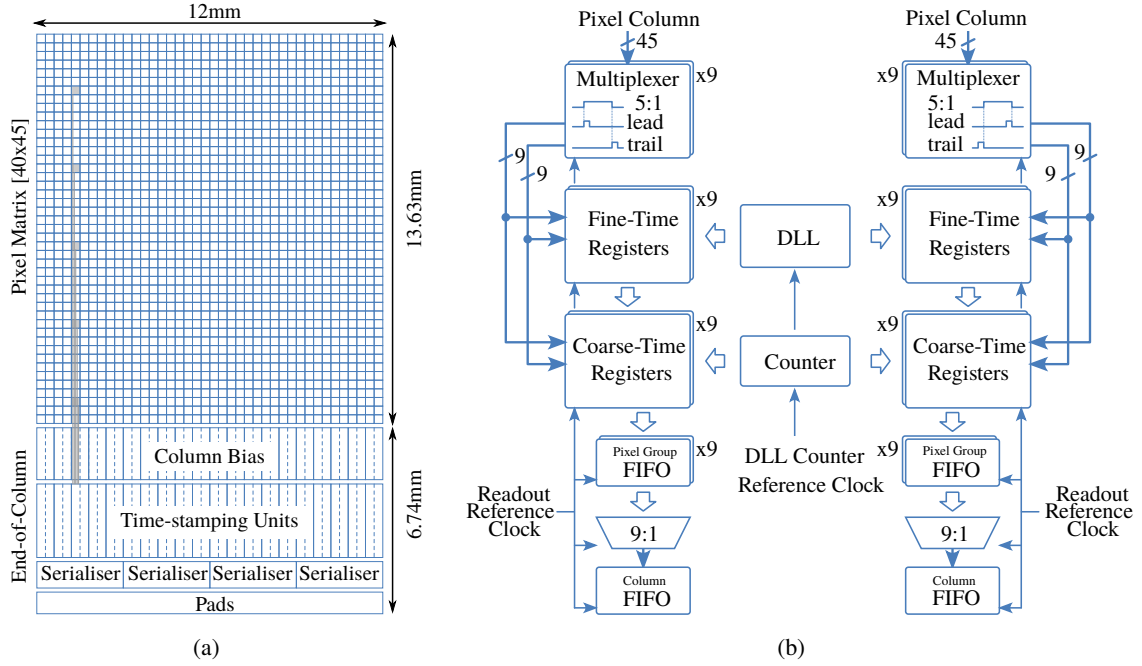
Figure 8. (a) The TDCPix architecture is separated in two parts, the pixel matrix and the end-of-column (EoC) regions. The output of five geometrically separated pixels are multiplexed in one time-stamping unit located in the EoC. The outputs of the time-stamping units are sent out by four serialisers. (b) Each time-stamping unit is fed by one clock shared amongst two pixel columns. This clock is used to increment a counter and to feed a delay-locked loop (DLL) dividing the clock period by 32. Hits leading and trailing edges are time-stamped by latching the counter and the DLL states in two registers whose content is encoded and multiplexed to be sent to the serialisers.
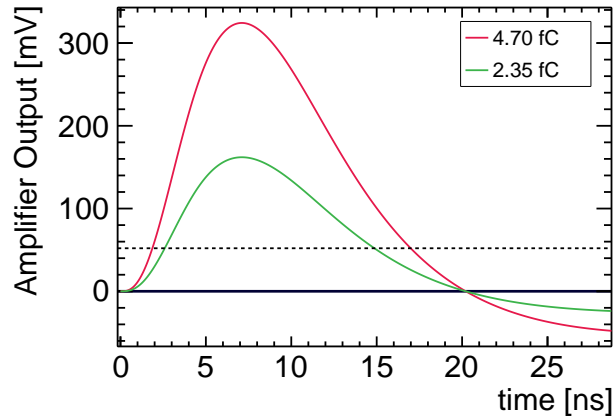

Figure 9. Pixel amplifier output for 100 V sensor bias, simulated using the WeightField2 software [13]. For a fixed hit arrival time, signals with larger amplitudes (red line) cross the discriminator threshold (dotted black line) earlier than smaller signals (green), inducing a time-walk.

3.2 Gb/s serialisers each serving 10 pixel columns. To further extend the time-stamp range to 1718 seconds, the serialiser intercalates every 6.4 µs a frame word containing an incrementing 27 bit counter. Hit collisions in the multiplexer are kept below the percent level by grouping non-adjacent pixels. In case of collision, the time-stamping unit inputs are locked by the first incoming hit until it is processed. The concurrent hits are not time-stamped but the hits pixel addresses are encoded in the word of the first incoming hit. The configuration registers are triplicated to mitigate SEU effects.

The chip power dissipation is 4.1 W and varies across the chip, with approximately 4.8 W/cm^2 in the EoC and 0.32 W/cm^2 in the pixel matrix[2]. As the detector is in vacuum, active cooling is therefore required.

2.3 Micro-channel Cooling Plate

The cooling system was designed to cool the front-end electronics and the sensor to less than 5 °C with a maximum temperature difference of less than 10 °C. In addition, the amount of material in the beam acceptance has to be as small as possible. To fulfil these requirements a micro-channel cooling device was developed [14]. The implementation of such a device is the world's first for high energy physics. Using this technology, the cooler – a silicon plate – also serves as mechanical interface between the HPD and the carrier board. This additional functionality adds a constraint on the plate planarity to be better than 30 µm. As the cooler is entirely made of silicon, the coefficient of thermal expansion of the plate is naturally matched to the HPD one.

The designed cooling device consists of a 210 µm thick 70×80 mm^2 silicon plate fabricated by bonding silicon wafers together, one of which etched to form 150 channels with a cross section of 200×70 µm^2 and arranged in two circuits as shown on Figures 10(a), 10(c) and 10(d). Such a device stands out from regular micro-channel designs [15] by the outstanding large area it covers.

Initially, the cooling plates were fabricated by etching microchannels into a bulk silicon substrate and bonding another bulk wafer on top of it to seal the microchannels. In order to thin the cooling plates, the bonded stack was etched in the central region on one or both sides using deep reactive ion etching. An accelerated etching was observed on the edges of the thinned regions. Thus, the cooling plates produced were thinner at the edges of the thinned region, but thicker in the most central part of it. As a result, a low production yield was reached, and the cooling plates equipping the first modules were thicker than the design value: 280 µm if both sides of the stack were etched and 380 µm if only one side was etched.

To overcome these issues, a process based on silicon on insulator (SOI) wafers was developed. The principle is illustrated in Figure 10(b). Two sets of SOI wafers (1-2) and (3-4) are bonded together such that the resulting stacks feature an oxide inner layer. The two silicon stacks are thinned to the desired thickness. Then channels are formed in stack (1-2) using plasma etching techniques. In parallel, an oxide layer is created on stack (3-4) using a wet oxidation technique. The two stacks are bonded together such that the micro-channels of stack (1-2) are closed by stack (3-4). Inlets and outlets are etched in the assembly. The resulting cooling plate is finally thinned using a plasma etching technique designed to be stopped by the inner oxide layers of the silicon stacks. This technique allows to obtain a planarity and a thickness matching the specifications and

---

[2] The chip operating voltages are 1.35 V for the EoC and 1.3 V for the pixel matrix.
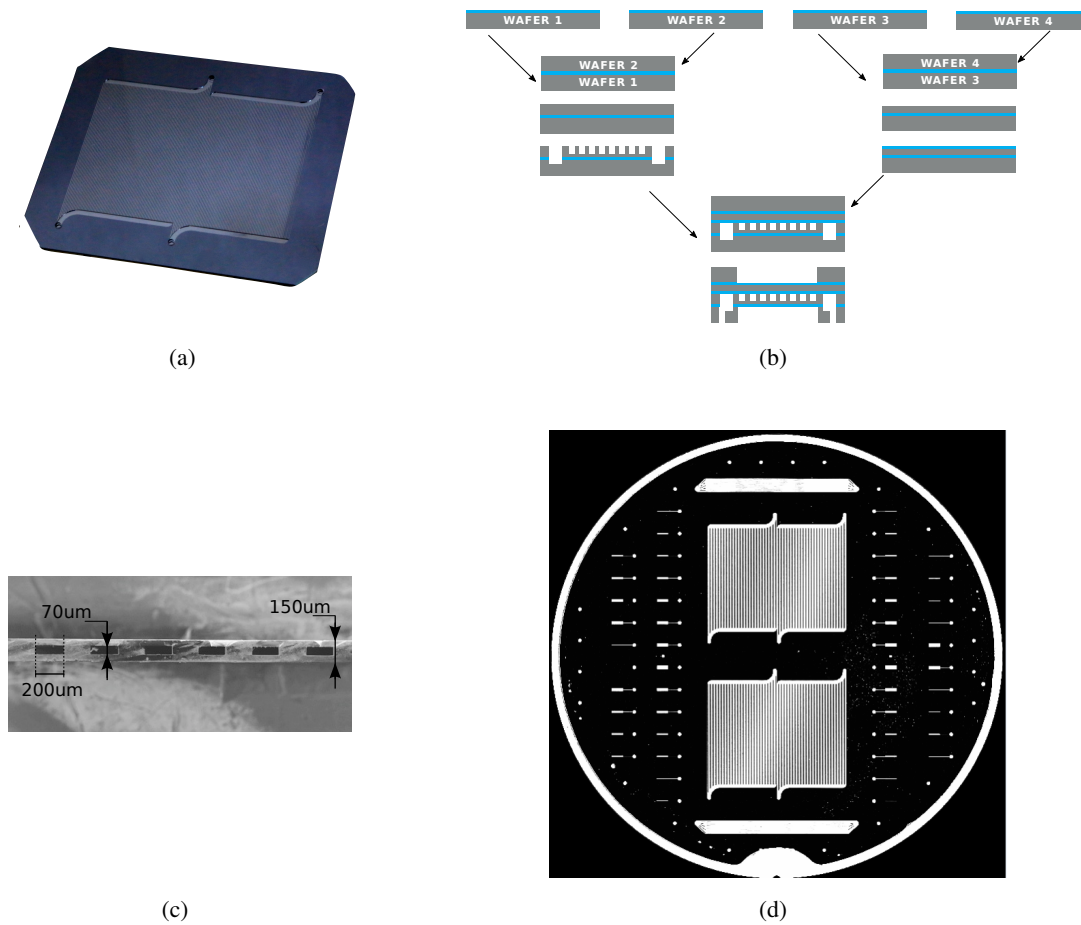
(a)

(b)

(c)

(d)

Figure 10. (a) Cooling plate wafer etched with micro-channels and closed with glass. (b) Schematics of the silicon wafer stack forming a cooling plate. The blue layers indicate the oxidised side of the wafers. See text for explanations. (c) Scanning electron microscope image of the cross-section of a 150 μm thick prototype cooling plate produced and characterised to evaluate the limit of the technology. (d) Acoustics imaging of a cooling plate wafer. Two cooling plates are placed in one wafer and are surrounded by testing devices. Each plate contains two cooling circuits.

the production yield was improved from less than 50% with the initial process to about 90% with the SOI-based process.

Each of the two plate circuits has an inlet and an outlet, as shown on Figure 11(b). Those openings are metallized such that KOVAR connectors welded with capillaries can be soldered on them as depicted on Figure 11(a). The capillaries are routed to exert no stress on the connectors and brazed in manifolds to the cooling distribution circuit.

2.4 Carrier Board

The assembly made of the HPD and the cooling plate is hosted in the carrier board which interfaces it to the systems installed outside the vacuum vessel. Hence, all services to the detector have to pass through the carrier, resulting in a dense and complex design. The carrier board is a T-shaped
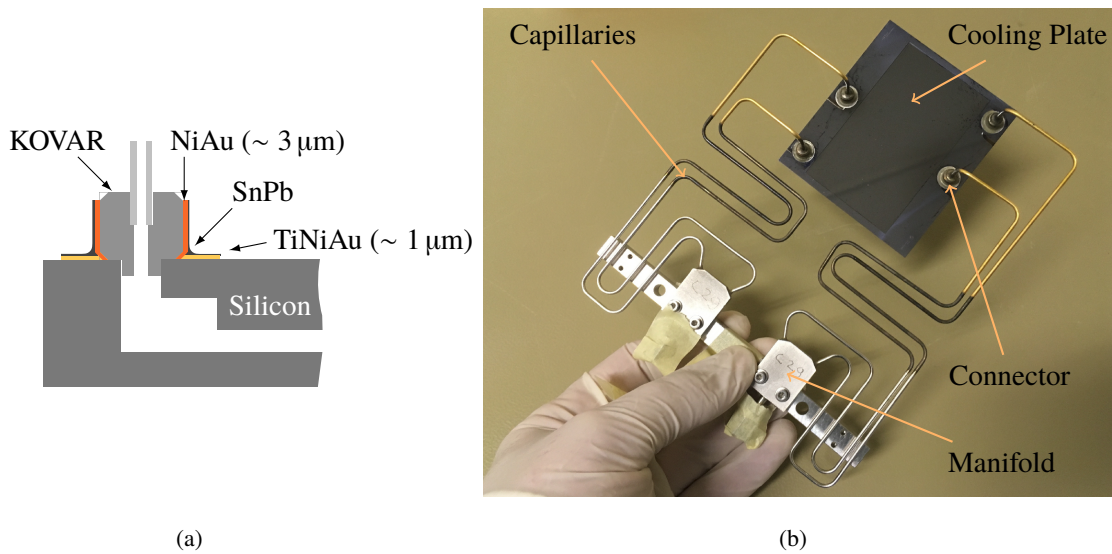
Figure 11. (a) Cross view of the connectors between the plate and the capillaries. (b) The two cooling plate circuits are connected to capillaries each by one set of inlet and outlet. The capillaries are then brazed on manifolds that are connected to the cooling distribution circuit.
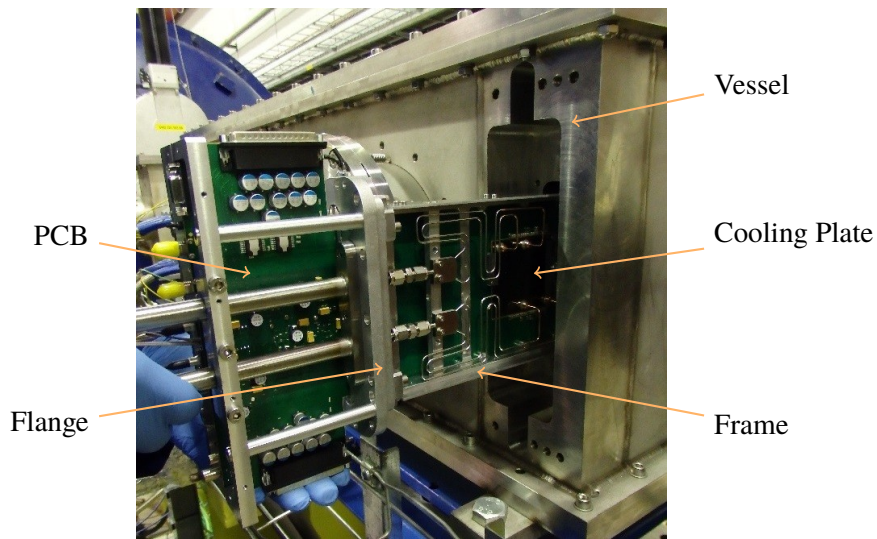

Figure 12. Photograph of a GTK module just before being inserted in the third vacuum vessel. The carrier board is equipped with the assembly made of the HPD and cooling plate and mounted in a frame and a flange acting as vacuum feed-through.

14-layer printed circuit board (PCB) with a rectangular countersink to receive the assembly made of the HPD and the cooling plate. The board is sealed in an aluminum frame which supports the capillaries manifolds and is equipped with vacuum tight joints. As such, the frame and board, shown on Figure 12, act as vacuum feed-through.
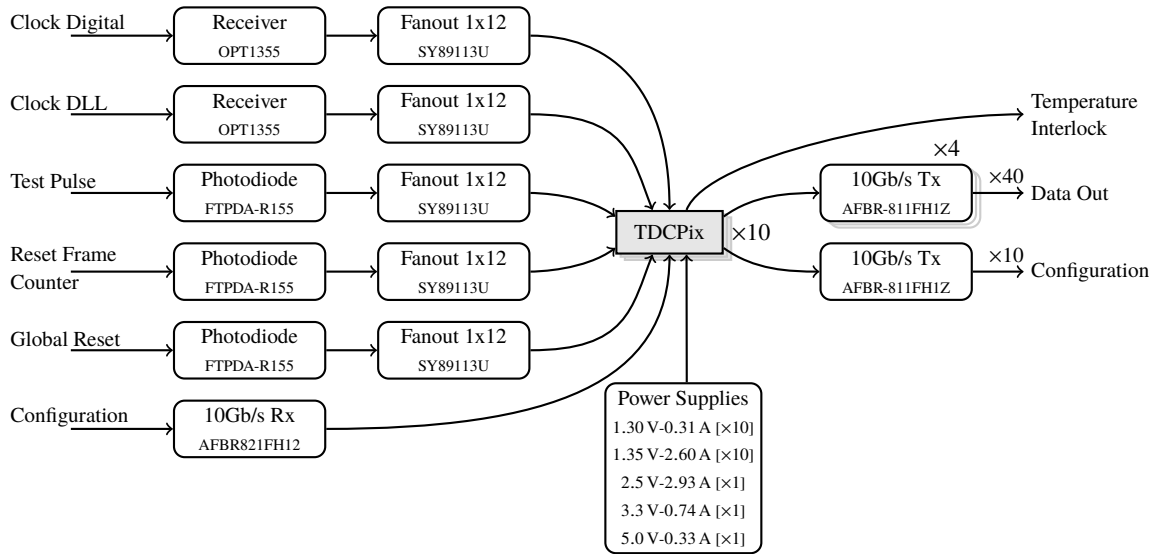
Figure 13. Simplified schematics of the carrier electronics.

The assembly made of the HPD and cooling plate is clamped from below in the countersink. The planarity around this opening is therefore required to be better than $100\,\mu m$. The top side of the opening is machined to form a staircase equipped with 1450 bonding pads spaced by $73\,\mu m$ onto which the chips are wire bonded. Power, configuration instructions, state information, interlock signals and hit data are transferred through those bonds to the board in which they are routed on $\sim 30\,cm$ long strips outside the vacuum vessel. The board is equipped with opto-electrical devices connected with optical fibres to the readout boards installed on surface. A simplified schematics of the carrier electronics is shown in Figure 13.

2.5   Readout System

The GTK readout system is the interface between the GTK stations and the rest of the NA62 experiment. The system provides an NA62 wide synchronised clock to the stations, it transfers, the configuration and control commands from the NA62 run control to the stations, and it receives the station data and extracts from them the hits matching the NA62 trigger requests. The latter task is the most challenging as all the data produced by the TDCPix ASICs are sent out to the readout system in a continuous unfiltered stream. As each chip sends data over four $3.2\,Gb/s$ serialisers, the data rate can reach up to $128\,Gb/s$ per station.

The GTK readout system consists of three sets, one per station, of ten custom electronics boards, serving one TDCPix each. Each set is connected to two computers. The boards comprise a motherboard onto which a daughter card is plugged. The mother board is equipped with:

- a field-programmable gate array (FPGA, Altera Stratix GX110) with two $1\,GB$ DDR2 SDRAM, mostly dedicated to the trigger matching operations,

- four $3.2\,Gb/s$ optical links dedicated to the TDCPix data,

- one 320 Mb/s optical link dedicated to the TDCPix configuration and

- two Gigabit Ethernet copper links to communicate with the computers.

The daughter board implements the TTC protocol [16] with which clock and trigger signals are handled. The board is also equipped with a spare Gigabit Ethernet copper link.

The computers are dual CPU servers running Linux. They are equipped with a least 24 GB of RAM and two 10 Gb/s Ethernet ports. One port is used to send the trigger matched data to the NA62 computer farm. The other one is used to receive the data from five boards multiplexed by an Ethernet switch.

The boards receive the NA62 asynchronous L0 trigger requests, with a rate up to 1 MHz, extract the corresponding pieces of input data and send them to the six computers. A simplified block diagram of the logic implemented in the FPGA to extract the hits matching a trigger request from the GTK data is shown in Figure 14. Each GTK readout is logically divided in two identical halves, each one handling the data coming from two of the four serializers of each TDCPix. Incoming data are first of all de-serialized, assembled in a segment of cache DPRAM spanning 6.4 µs and then copied to the DDR2 memory waiting for level-0 trigger decision (the maximum level-0 trigger latency is 1 ms). Two such segments (the even and odd ones) exist and are cycled between them so
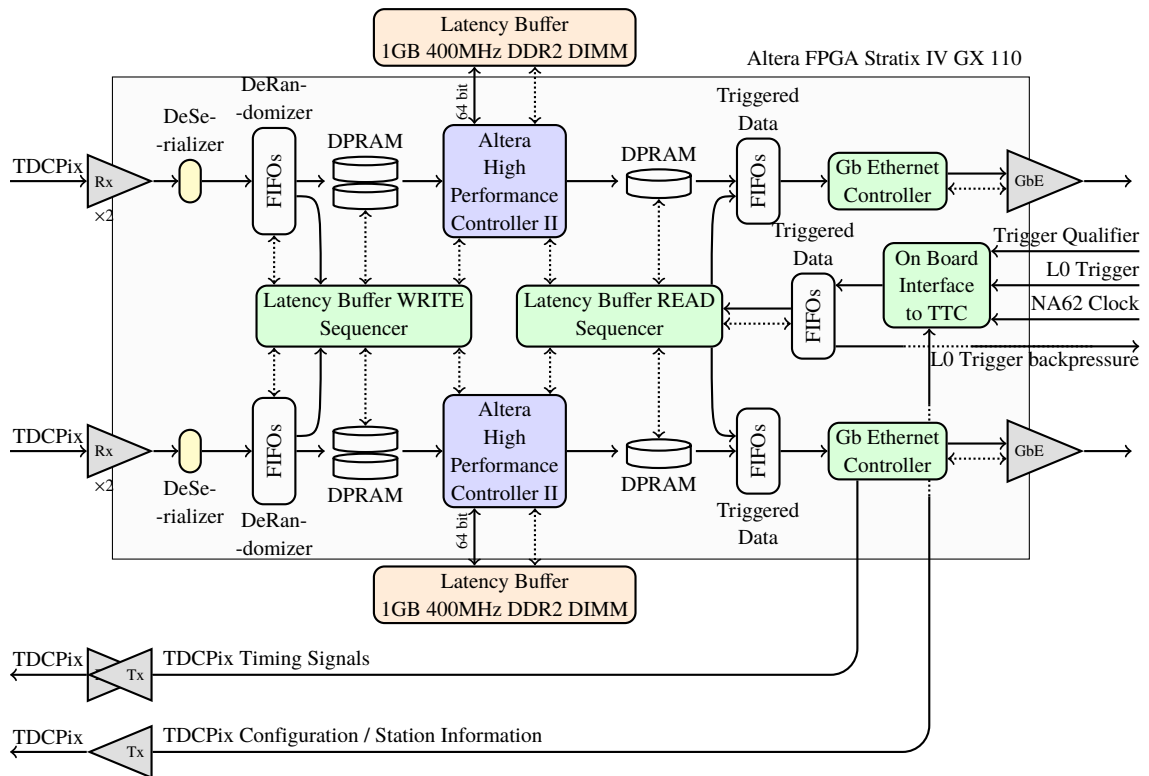

Figure 14. Simplified block diagram of the logic implemented in the FPGA to extract from the GTK data the hits matching a trigger request. See text for explanations.

that while the even one is receiving the data, the odd one is being copied to DDR2 (and vice-versa). To extract the hits matching a trigger request, data are copied from DDR2 memory to a different DPRAM segment, and data in a 75 ns time window around the trigger time-stamp are selected. These data are then assembled in UDP packets and sent to the readout servers. Upon asynchronous L1 request, with a rate up to 100 kHz, each server assembles the data in one packet and sends it to the NA62 computer farm.

2.6 Services

The operation of the GTK stations relies on the provision of refrigerated coolant, stable electrical power and vacuum.

Cooling Plant

A cooling plant was assembled to provide the three stations with refrigerated perfluorohexane - C_6F_{14}. To reduce the exposition to radiation, the plant is located 10 m away from the beam line. The coolant is provided to the stations through three lines of 10 m to 20 m lengths. The operating detector pressure is 3 bars with a coolant flow of 3 g/s. The coolant temperature at the detector inlet is adjusted with heat exchangers installed just before the stations and fed in C_6F_{14} by the cooling plant though dedicated high flow distribution lines.

The detector distribution lines are equipped with pressure regulating analogue valves to limit the input pressure. In addition, robust operating procedures were developed to ensure smooth pressure and temperature ramping at the detector during starting and stopping phases. To avoid risks related to condensation, the cooling plant is interlocked with the detector vacuum. All the programmable logic controller (PLC) electronics are located in a radiation protected area to avoid perturbations due to the accumulated dose.

Vacuum

The detectors are installed in three vacuum vessels operated at a pressure of $\sim 10^{-7}$ mbar. Each vessel can be isolated from the beam vacuum by valves at the input and the output. The third station vessel hosts as well the CHANTI detector. A coarse vacuum is provided in each vessel by a turbo-molecular pump. The nominal pressure is obtained opening all the valves and using the cryo-pumps installed in the downstream part of the beam pipe.

Power

The detector electronics power, except the sensor high voltage, is provided by radiation hard power supplies installed 10 m away from the beam line to reduce the radiation exposition. The power supplies cables are 10 m to 20 m long. Power cables are thus equipped with sense wires through which the voltage delivered at the detector is read out and adjusted with a feed-back loop. The power supplies are interlocked to the TDCPix over-temperature alarm and to the cooling system. The high voltage power supplies are installed in the surface counting room and connected to the detector with a 200 m long cable. These power supplies are interlocked to the same state-machines as the low voltage power supplies and also to the vacuum gauge to avoid electrical discharges.

3 Detector Performance

3.1 Detectors and Operations

Three stations were constructed for the 2016 data taking operations. They were all equipped with n-in-p sensors and mounted on cooling plates with thickness of 380 μm for GTK1 and GTK2 and and 280 μm for GTK3. The material budget in the beam acceptance is therefore 0.73% X_0 for GTK1 and GTK2 and and 0.62% X_0 for GTK3. The fraction of dead pixels is below 0.5% which can be seen in the detector hit maps shown in Figure 15.

The thresholds of the pixel discriminators were tuned in the laboratory. The expected charge distribution for minimum ionizing particles following a Landau distribution peaking at 2.4 fC. The pixel thresholds were trimmed to 0.8 fC. The threshold calibration was done using a ^{109}Cd source extrapolating from the 0.98 fC and 1.11 fC charge depositions of the low energy photons (22.1 keV and 24.9 keV) emitted.
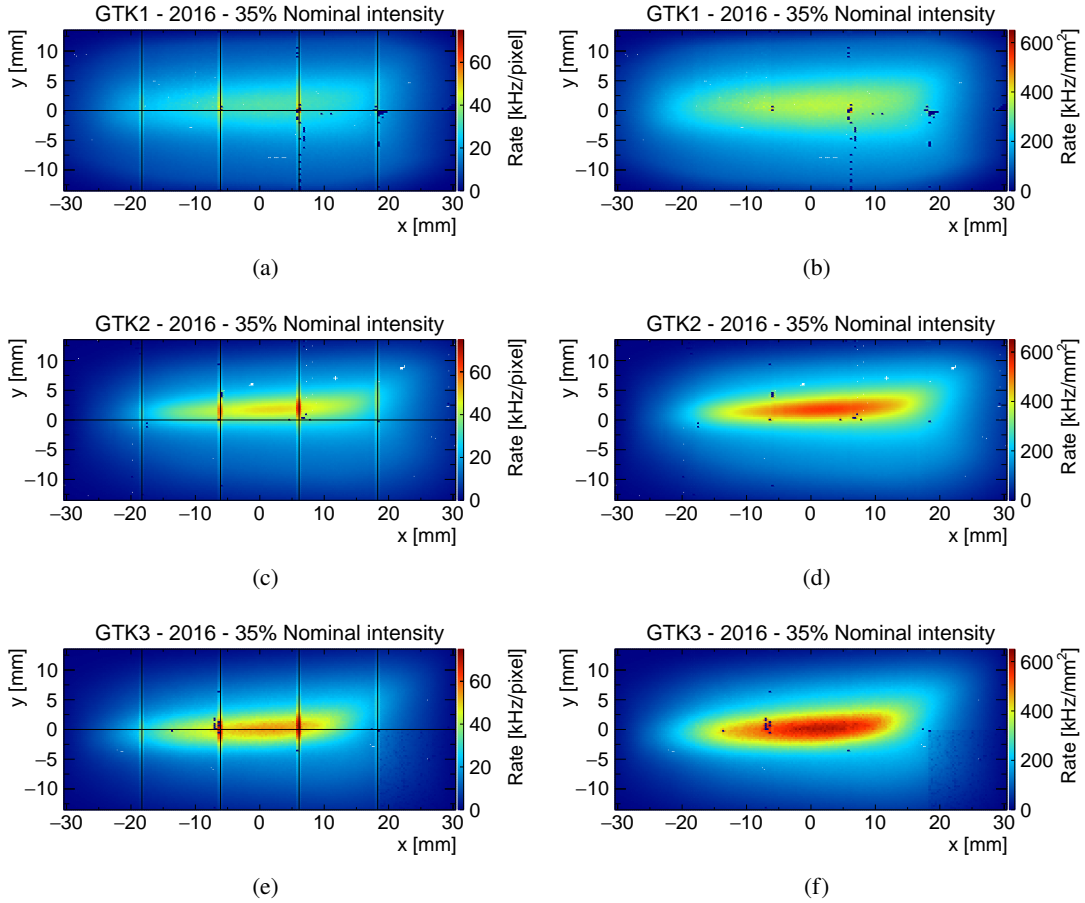

Figure 15. Hit rate per pixel (left column in kHz/pixel, right column in kHz/mm^2) in the three GTK modules recorded in 2016 at around 35% of nominal beam intensity. In (b), (d) and, (f) the rates in the elongated pixels is scaled down by 30% to correspond to a standard pixel. When present, the black lines indicate the readout chips.

The three GTK stations were fully commissioned by mid-September 2016 and took data until mid-November 2016. They were then dismounted and stored at $-21\,°C$. The same three stations were installed in the beam line in May 2017 and took data until mid-October 2017.

During data taking, the stations were operated at a bias voltage of $100\,V$ except for few days at the end of the 2016 data taking period when the bias voltage was increased by steps of $50\,V$ up to $250\,V$. As shown in Figure 16(a), the stations were kept cold most of the time and reheated only for short maintenance periods. The beam intensity was around 35% of the nominal intensity in 2016 and 60-65% in 2017. Figure 15 shows the hit rate recorded in each station in 2016. The sensors irradiation was monitored using a proportional chamber operated independently of NA62. The station's average integrated fluence is reported in Figure 16(b). At the end of the 2017 data taking period, a peak (average) fluence of $1.3\,(0.26) \times 10^{14}$ 1 MeV eq. n/cm^2 was integrated in the stations installed first (GTK1 and GTK3). This fluence corresponds to around 60 days of operation at full beam intensity.
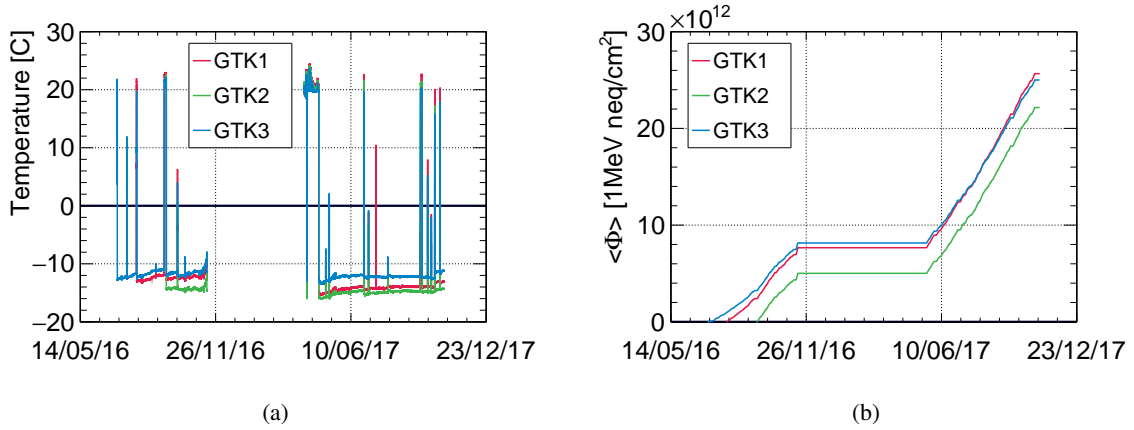

Figure 16. (a) Inlet temperature of the three GTK station when installed in the beam line. (b) Average integrated fluence of the three GTK stations.

3.2 Time Resolution

As explained in Section 2.2, the measurement of the hit time coordinate relies on off-line time-walk correction. For every chip of every station, time-walk corrections are derived in $100\,ps$ ToT bins between 0 and $30\,ns$. Above this ToT value, the time-walk measurements saturate.

Moreover, within a chip, pixel output signals are delayed due to the propagation time between the pixels and the EoC depending on the pixel position in the matrix. In addition, the digital signals at the output of the pixel discriminators degrade as they transit in the lines connecting the pixel to the EoC. Hence, the signal delay at the EoC increases with the distance travelled by the signal from the pixel. A time correction for each pixel is derived to compensate for these effects.

All the corrections are derived from data using the coincidence of the GTK tracks, formed of three single hit clusters, and a reference time provided by the KTAG or RICH time. The time resolution of these detectors are about $70\,ps$ and $100\,ps$, respectively. Example of corrections are shown in Figure 17.
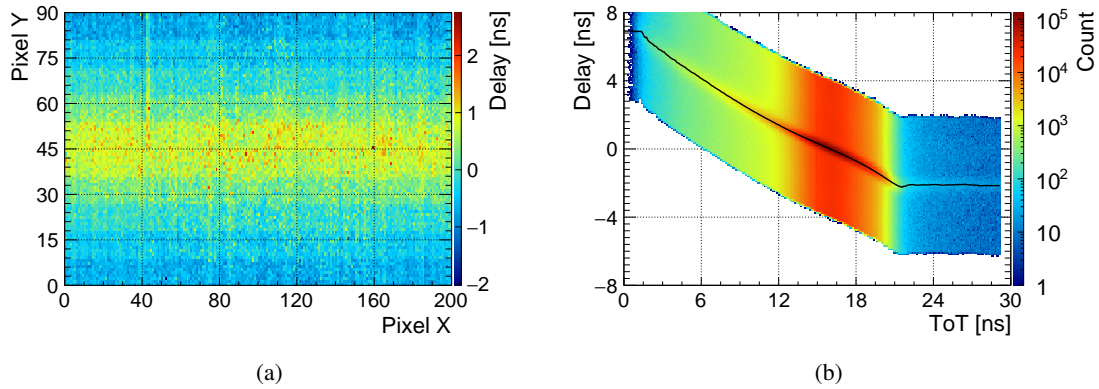
Figure 17. (a) Station pixel delays as function of pixels coordinates. Delays are offset such that the average delay of the station is null. The pixel delay decreases as function of the distance to the EoC. (b) Chip GTK to KTAG time difference distribution as function of GTK hit ToT, overlaid with fitted time-walk delay (black line).

The time resolution of the detector is measured using samples of $K^+ \to \pi^+\pi^+\pi^-$ collected during 8 to 10 consecutive hours of data taking. The K^+, formed by the three pions, is assigned two time coordinates: the average time coordinate of the three pion RICH candidates and the time coordinate of the closest in-time KTAG candidate. Finally, the K^+ is associated with a GTK track based on the track kinematics and direction correspondence and no time criteria is applied.

A standalone derivation of the GTK time resolution is performed comparing the three station to station time differences shown in Figure 18. The time resolution values obtained are 132.0 ps for GTK1, 127.1 ps for GTK2 and 129.2 ps for GTK3. Assuming that the track time coordinate is the average of the three hits time coordinates, the time resolution for the track is 74.7 ps.
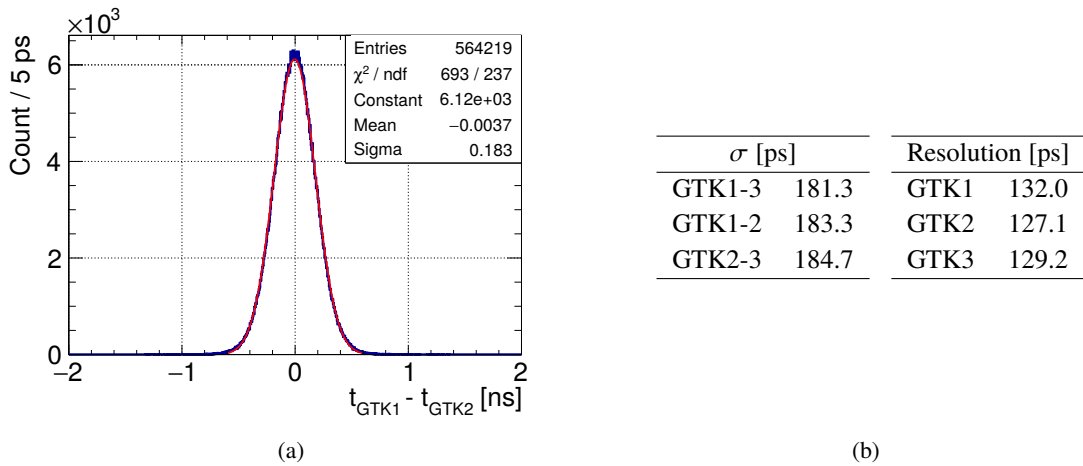

σ [ps]		Resolution [ps]	
GTK1-3	181.3	GTK1	132.0
GTK1-2	183.3	GTK2	127.1
GTK2-3	184.7	GTK3	129.2

Figure 18. (a) Time difference distribution between hits in the first and second GTK station corresponding to the same K^+ track. A Gaussian function is fitted to the distribution (red). (b) Standard deviation from the fits to the time difference distributions between the three GTK stations and the corresponding station time resolutions.

A NA62-wide time resolution is derived fitting Gaussian functions to the time difference distributions between the three GTK stations, the RICH, and the KTAG candidates split by RICH and KTAG candidate hit numbers. The resolutions obtained are compatible with those from the standalone methods and are 135.0 ps for GTK1, 128.0 ps for GTK2 and 132.4 ps for GTK3.

The time resolution was also measured at different sensor bias voltages during a ten day period at the end of the 2016 data taking period. The results are reported on Figure 19(a) and show a moderate but systematic improvement of the station time resolution of 20 ps increasing the sensor bias from 100 V to 250 V. Figure 19(a) also shows that for a bias voltage of 250 V a single station time resolution of 115 ps and a track resolution of 65 ps was achieved.

The time resolution was monitored over time as shown in Figure 19(b)–(d). The resolution degrades over time and reaches 146 − 158 ps at the end of the 2017 data taking period which is still better than the nominal specifications. Such a degradation was an expected effect of the detector irradiation, however no study rules out other concurrent factors.
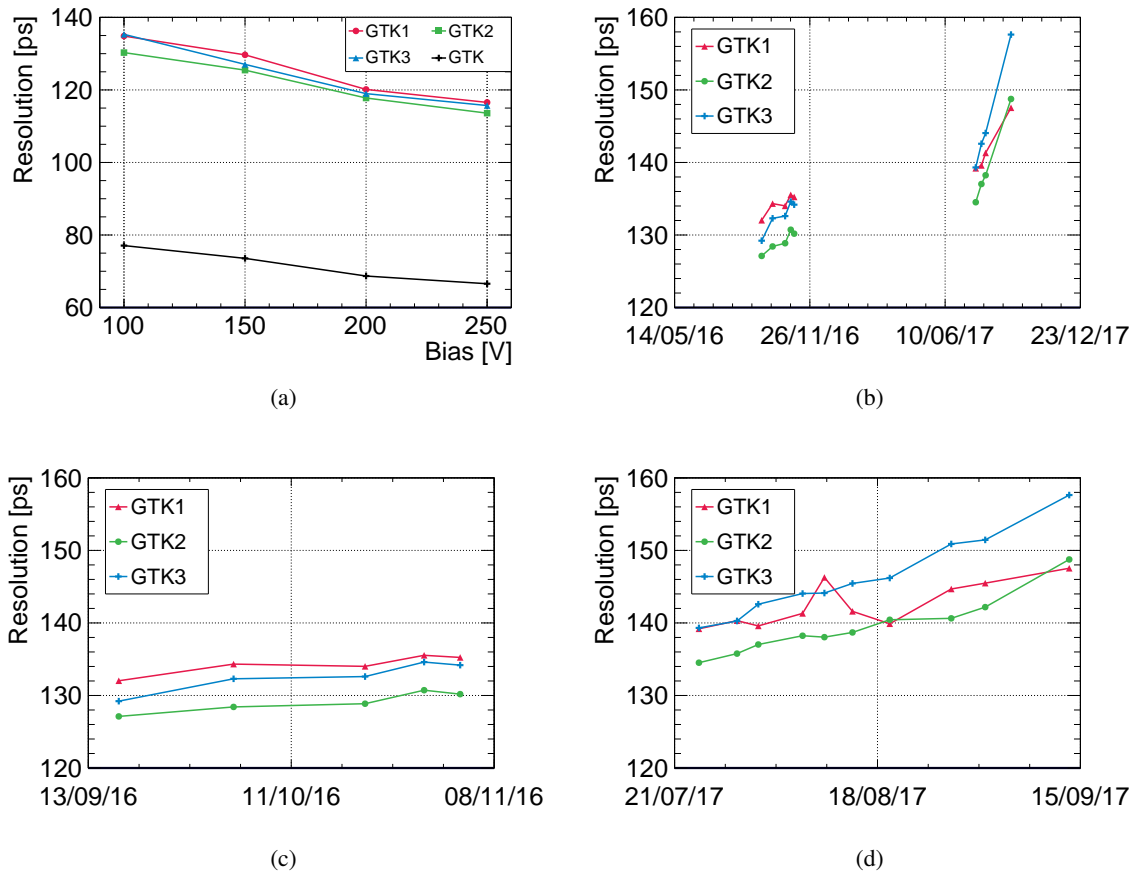

(a)

(b)

(c)

(d)

Figure 19. (a) GTK station time resolution and average time resolution as function of the bias voltage. (b)–(d) Evolution of the time resolution over time. Small ticks on the x-axis are spaced by four weeks for (b) and one week for (c) and (d).

3.3 Kinematics Performances and Material Budget

The kinematics performances of the GTK are of paramount relevance to reduce the background from $K^+ \to \pi^+\pi^0$ for the $K^+ \to \pi^+\nu\bar{\nu}$ branching ratio measurement as explained in Section 1. One of the key elements controlling these performances is the station's material budget as shown in Figure 5. Moreover, the material budget of the last GTK station is crucial to reduce the background arising from beam particles inelastically scattered in this station. Solid estimates of this quantity are thus needed.

A first estimate is obtained by measuring the thickness of the station sensor, chips and cooling plate. The thickness found correspond to a material budget of 0.73% X_0 for GTK1 and GTK2 and 0.62% X_0 for GTK3. These results can be cross checked for GTK2 using data by measuring the residual along the horizontal axis between the hit position in this station and the position extrapolated based on the hits in the two other stations. The results found are compatible within 10% with those based on the thickness measurement.

The figure of merit for the kinematics performance is the squared missing mass resolution for $K^+ \to \pi^+\pi^0$ as defined in Equation 1.1. This quantity is shown in Figure 20 as a function of the π^+ momentum for the 2016 data overlaid with the expectations based on nominal performance. The contribution of the π^+ and K^+ momentum and angular resolutions are also shown. The overall agreement between the squared missing mass resolution obtained with a simulation based on nominal performance and the measured ones indicates that the GTK momentum and angular resolutions are matching the specifications.
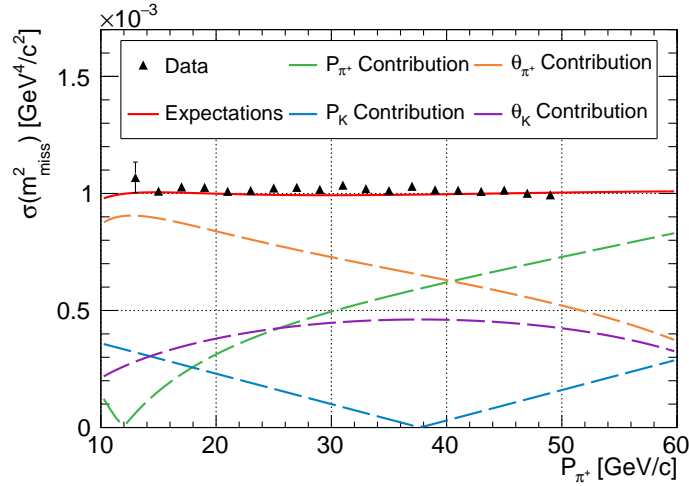

Figure 20. Squared missing mass resolution for $K^+ \to \pi^+\pi^0$ as function of the π^+ momentum for the 2016 data (black triangle) overlaid with the resolution expected based on nominal performance (plain red line) and the expected contributions (dashed line).

3.4 Efficiency

The $K^+ \to \pi^+\pi^0$ background rejection obtained using the decay kinematics becomes ineffective when a wrong track is assigned to the K^+. Such situations occur when K^+ hits are missing due to the limited GTK efficiency. The intrinsic GTK inefficiency is thus defined as the fraction of events in which hits from the K^+ triggering the detector readout are missing. Events for which K^+ hits are expected to be missing due to an identified readout inefficiency are considered as dead-time and not counted as inefficiency. Likewise, missing hits due to geometrical consideration are counted as geometrical acceptance loss. The cumulated loss due to dead-time and geometrical acceptance is about 5%.

The efficiency was measured using samples of $K^+ \to \pi^+\pi^+\pi^-$ selected and reconstructed in the same fashion as for the time resolution measurements. The efficiency is then derived from the fraction of reconstructed K^+ matched with a GTK track. A good match is realised when the GTK track and the K^+ have compatible momentum, direction, time and position at GTK3.

Using this method, the intrinsic GTK efficiency was found to be to be 97.0% which, assuming no correlation and similar behaviour of the three stations, corresponds to a 99.0% station efficiency. This performance matches the detector specifications.

Studies were pursued to identify the origin of the missing hits. Unmatched events were reprocessed considering GTK tracks formed by hits in two stations only. Sources were identified based on the expected position of the missing station hit and searching for hits offset in space or time. The main inefficiency sources found are dead pixels, local mis-calibration inducing signal loss on pixels set at higher thresholds, and synchronous firing of pixels multiplexed in the the same time-stamping unit inducing ambiguities in the hit pixel address. Last, few hits were lost in GTK2 due to the misalignment between the beam peak intensity region and the station centre as shown
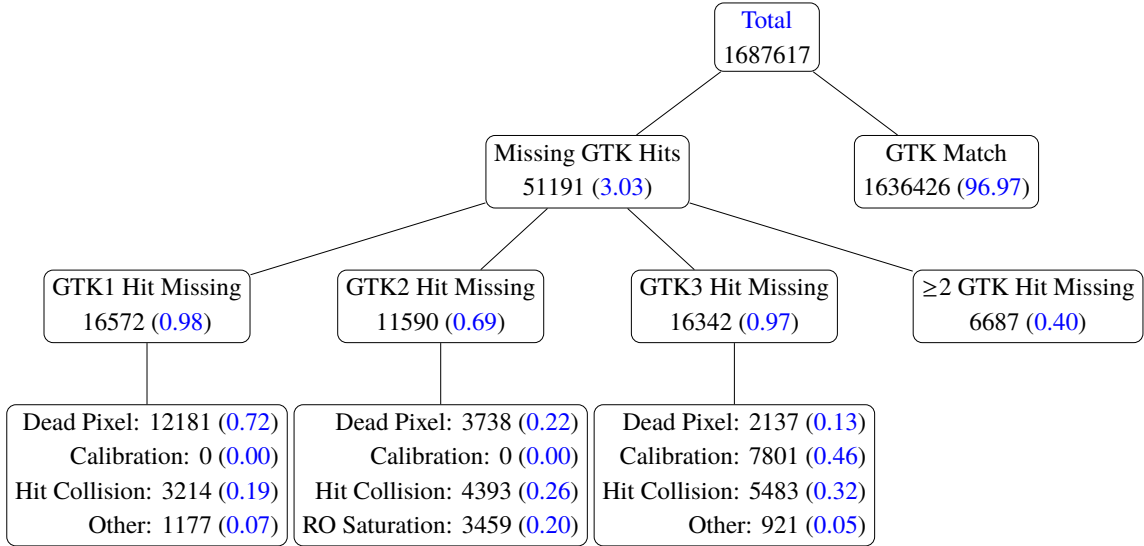

Figure 21. Partition of the $K^+ \to \pi^+\pi^+\pi^-$ collected during an eight hours data taking period in October 2017. The numbers denote the numbers of events in each category. The numbers in parenthesis are the percentage of the total number of $K^+ \to \pi^+\pi^+\pi^-$.

in Figure 15. This misalignment induces a much higher rate on few chips. At the 2016 beam intensity, this effect was irrelevant but at higher beam intensities, the readout electronics of the most illuminated chips started to saturate. The importance of each source varies depending on the station and is reported in Figure 21. The stations inefficiency are slightly correlated as the position of some dead pixel clusters are aligned between two stations.

4 Conclusions

The NA62 physics case requires the development of a beam spectrometer, the GigaTracKer, with unprecedented specifications in terms of particle rate, material budget and time resolution. To fulfil those requirements, a radically novel approach to tracking, based on particle time-stamping at the sub nanosecond level, was developed. A custom made chip, the TDCPix, and a carefully designed sensor allowed for this tour-de-force.

The time-stamp based approach enables the reduction of the number of tracking stations to three, the bare minimum for a magnetic spectrometer. The detector material budget is furthermore reduced by using silicon micro-channel cooling plates, a world first development in high energy physics.

The detector was successfully assembled and operated. Performance in most cases reach or exceed the specifications. The results achieved in terms of time resolution for the single hit are 115 ps corresponding to a track time resolution of 65 ps.

Acknowledgements

References

[1] E. Cortina Gil et al. The Beam and detector of the NA62 experiment at CERN. *J. Instrum.*, 12(05), 2017. https://arxiv.org/abs/1703.08501. http://dx.doi.org/10.1088/1748-0221/12/05/P05025.

[2] A. J. Buras, et al. $K^+ \to \pi^+ \nu \bar{\nu}$ and $K_L \to \pi^0 \nu \bar{\nu}$ in the Standard Model: status and perspectives. *J. High Energy Phys.*, 11, 2015. https://arxiv.org/abs/1503.02693. http://dx.doi.org/10.1007/JHEP11(2015)033.

[3] E. Cortina Gil et al. First search for $K \to \pi^+ \nu \bar{\nu}$ using the decay-in-flight technique. *Phys. Lett. B*, 791, 2019. http://www.sciencedirect.com/science/article/pii/S0370269319301121.

[4] A. J. Buras, et al. Can we reach the Zeptouniverse with rare K and $B_{s,d}$ decays? *J. High Energy Phys.*, 11, 2014. https://arxiv.org/abs/1408.0728. http://dx.doi.org/10.1007/JHEP11(2014)121.

[5] A. V. Artamonov et al. Study of the decay $K^+ \rightarrow \pi^+ \nu \bar{\nu}$ in the momentum region $140 < P_\pi < 199$ MeV/c. *Phys. Rev.*, D79, 2009. https://arxiv.org/abs/0903.0030. http://dx.doi.org/10.1103/PhysRevD.79.092004.

[6] M. Tanabashi et al. Review of Particle Physics. *Phys. Rev.*, D98(3), 2018. http://dx.doi.org/10.1103/PhysRevD.98.030001.

[7] P. Riedler et al. Recent test results of the ALICE silicon pixel detector. *Nucl. Instrum. Methods Phys. Res., Sect. A*, 549, 2005. http://dx.doi.org/10.1016/j.nima.2005.04.027.

[8] G. Anelli, et al. The P326 (NA48/3) Gigatracker: Requirements and design concept. *Nucl. Instrum. Methods Phys. Res., Sect. A*, 572(1), 2007. http://dx.doi.org/10.1016/j.nima.2006.10.369.

[9] G. Aglieri Rinella, et al. The TDCpix ASIC: High rate readout of hybrid pixels with Timing Resolution Better than 200 ps. In *IEEE*, 2013. http://dx.doi.org/10.1109/NSSMIC.2013.6829432.

[10] M. Noy, et al. Characterisation of the NA62 GigaTracker end of column readout ASIC. *J. Instrum.*, 6(01), 2011. http://dx.doi.org/10.1088/1748-0221/6/01/c01086.

[11] A. Kluge, et al. The TDCpix readout ASIC: A 75ps resolution timing front-end for the NA62 gigatracker hybrid pixel detector. *Nucl. Instrum. Methods Phys. Res., Sect. A*, 732, 2013. http://dx.doi.org/10.1016/j.nima.2013.06.089.

[12] G. A. Rinella, et al. Test-beam results of a silicon pixel detector with time-over-threshold read-out having ultra-precise time resolution. *J. Instrum.*, 10(12), 2015. http://stacks.iop.org/1748-0221/10/i=12/a=P12016.

[13] F. Cenna, et al. Weightfield2: A fast simulator for silicon and diamond solid state detector. *Nucl. Instrum. Methods Phys. Res., Sect. A*, 796, 2015. http://www.sciencedirect.com/science/article/pii/S0168900215004842.

[14] G. Romagnoli, et al. Silicon micro-fluidic cooling for NA62 GTK pixel detectors. *Microelectron. Eng.*, 145, 2015. http://dx.doi.org/10.1016/j.mee.2015.04.006.

[15] Y. Temiz, et al. Lab-on-a-chip devices: How to close and plug the lab? *Microelectron. Eng.*, 132, 2015. http://www.sciencedirect.com/science/article/pii/S0167931714004456.

[16] B. G. Taylor. TTC distribution for LHC detectors. *IEEE Trans. Nucl. Sci.*, 45, 1998. http://dx.doi.org/10.1109/23.682644.