Università degli Studi di Torino

Scuola di Dottorato in Scienza ed Alta Tecnologia

The FPGA-based first level trigger for the NA62 Experiment at CERN SPS.

May 2017

Dario Soldi

Università degli Studi di Torino

Scuola di Dottorato in Scienza ed Alta Tecnologia

Indirizzo di Fisica ed Astrofisica

The FPGA-based first level trigger for the NA62 Experiment at CERN SPS.

May 2017

Dario Soldi

Supervisor: Ezio Menichetti

Co-Supervisor: Flavio Marchetto

Acknowledgements

It is my pleasure to thank those who made this thesis possible. First of all, I would like to express my special appreciation and thanks to Flavio Marchetto, you have been a tremendous mentor for me. You were and remain my best role model for a scientist. Your positive attitude and understanding really deserve an everlasting gratitude.

I am also very grateful to Ezio Menichetti for his scientific advice and his knowledge spent in many insightful discussions, suggestions and academic support. To Cristina Biino, for her motivational drive and practical suggestions during the entire PhD period.

I am thankful to Marco Sozzi, who believed in this project, allowing its implementation and use in the official NA62 data taking.

Furthermore, this project would never have been possible without the support and guidance of various people of the NA62 Collaboration, and in particular Francesco Gonnella, Mattia Barbanera, Nicolas Lurkin, Chris Parkinson, Giuseppe Ruggiero, Riccardo Fantechi, Gianluca Lamanna, Enrico Gamberini and Roberto Piandani. Thanks for your advice and for your precious help.

Thanks to Marco Boretto, a travel mate.

A special thanks to my family. Words cannot express how grateful I am to my mother and my father for all of the sacrifices that they have made on my behalf. To my brother, to attend to my complaints calmly and for his advice as older brother.

Thanks to all my old friends, for the good times spent by arguing about everything and nothing.

Lastly I would like to thank Irene, for her love, support, patience and advice more valuable than you could ever imagine.

Abstract

The main purpose of the NA62 experiment is to measure the branching ratio of the ultra-rare decay $K^+ \to \pi^+ \nu\bar{\nu}$ with a precision of 10%. The intense flux of particles in a rare-event experiment requires a high-performance trigger and data acquisition system, ensuring a high acceptance for the signal events together with a high rejection of the decays accounting for most of the rate. In this work the trigger processor, called Level-0 Trigger Processor (L0TP), of the first level of the NA62 trigger logic is presented. It has been designed to select candidate events at $\sim$ 10 MHz input rate coming from seven different sources of data and to reduce the output trigger rate to 1 MHz. It is the only NA62 trigger level implemented in hardware, by the use of programmable logic (FPGA).

Data collected by detectors are sent via Gigabit Ethernet connections to the L0TP board. The logic of firmware compares different event configurations with pre-selected masks, sending the result of the computation to all sub-systems. Performances and limitations of the entire system are reported, including a detailed study of the efficiency of detectors participating to the trigger generation.

The development of the L0TP has been performed concurrently with the construction of the experiment, growing in complexity from a preliminary version commissioned during the first pilot run in 2014 to the state of the art system used in the 2016 data taking which is presented in this work.

Table of contents

Introduction

NA62 is a fixed target experiment located at CERN, operating on the 400 GeV/c proton beam supplied by the CERN Super-Proton-Synchrotron (SPS). The NA62 experiment aims to measure the branching ratio of the (ultra)rare decay $K^+ \to \pi^+ \nu\bar{\nu}$ in order to check the Standard Model prediction. This decay is significant since its branching ratio can be computed with a good accuracy and it is very sensitive to new physics, being strongly suppressed by GIM mechanism. The experiment is designed to collect of the order of one hundred $K^+ \to \pi^+ \nu\bar{\nu}$ events in about two years of data taking, with a signal to background ratio of 10 : 1.

Two experiments at BNL, E787 and E949 [1], have collected only seven candidates using kaon decays at rest and obtaining a branching ratio value of $1.73^{1.15}_{-1.05} \times 10^{-10}$, affected by large statistical uncertainties.

The NA62 experiment has been designed to collect in flight decays of K^+ at high energy with momentum $P_{K^+} = 75 GeV/c$, produced by a proton beam impinging on a Beryllium target. The signature of the $K^+ \to \pi^+ \nu\bar{\nu}$ event is defined by a sigle pion charged track, since it is not possible to detect the neutrinos coming from the decay. The most challenging part of the experiment is to reject all the background up to the needed sensitivity with an adequate veto and particle identification (PID) system.

In order to achieve the measurement, NA62 should be able to:

- identify K^+ in a high incident flux of particles, with minimal accidental mis-tagging, and track incoming particles at a rate of about 1 GHz, with a time resolution better of 200 ps;
- separate μ from π particles up to a maximum momentum of 35 GeV/c, achieving a muon rejection of at least 10^5 using a magnetized iron/scintillator detector;
- create a hermetic veto system able to detect photons and the charged particles coming from three- and four-body kaon decays.

One of the central challenges to for high rate experiments is the design of the trigger and data acquisition system. Searches of ultra-rare processes such as $K^+ \to \pi\nu\bar{\nu}$, which has a

predicted branching ratio of around 10^{-11}, demand an high instantaneous beam intensity simply to collect a sizable sample of the signal events. The main challenge of the trigger system is therefore to guarantee a high acceptance for the signal events, keeping at the same time a high rejection of kaon decays accounting for most of the rate.
In this work the trigger processor, called Level-0 Trigger Processor (L0TP), of the first level of the NA62 trigger logic is presented. It has been designed to select candidate events at $\sim$ 10 MHz input rate coming from seven different sources of data and to reduce the output trigger rate to 1 MHz. It is the only trigger level implemented in hardware, by the use of programmable logic (FPGA). Two higher trigger levels based on software algorithms follow it.

The development of the L0TP design includes the following points:

- FPGA firmware implementation of the selection logic;
- setting up of the test-bench where the system has been developed;
- C++ simulations and data analysis.

The development of the L0TP has been performed concurrently with the construction of the experiment, growing in complexity from a preliminary version commissioned during the first pilot run in 2014 to the state of the art system used in the 2016 data taking which is presented in this work.

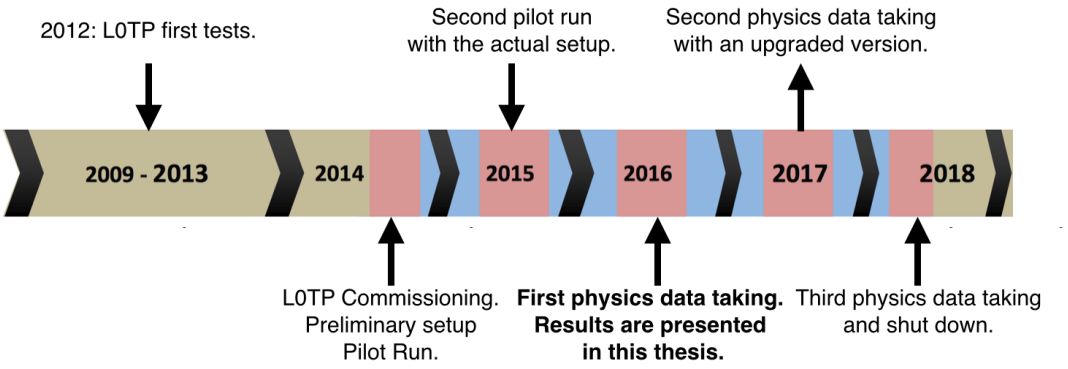


Figure 1. Timeline of the L0 Trigger Processor project development.

This thesis is structured in five chapters.
The first chapter contains a brief theoretical introduction to the $K^+ \to \pi\nu\bar{\nu}$ decay mode. The second chapter describes the NA62 experiment in detail, including both detectors and

trigger system used for the 2016 data taking. The third chapter is dedicated to a detailed description of the Level-0 Trigger Processor, including the hardware of the system and the logic driving the firmware. The performance of the L0 trigger, in terms of rates and efficiencies, as measured by NA62 in 2016 at a beam intensity of $\sim 35\%$ of the design value are presented in chapters four and five.

Chapter 1

Theoretical framework

1.1 Introduction

The Standard Model (SM) of fundamental constituents and interactions provides a successful theoretical description of the core properties of matter, yielding a coherent and detailed explanation of all elementary processes which can be observed in the experiments at the large accelerators and elsewhere, including interactions of high energy cosmic rays and cosmic gamma rays and neutrinos. At the same time, it is widely acknowledged that this is unlikely to be our final picture of the Universe. On the one hand, gravity is not included in the SM, due to the well-known difficulties one encounters in making a proper quantum version of it: while this is usually not taken as a serious problem for a quantitative description of high energy reactions, in view of the smallness of the gravitational coupling, it is clearly unsatisfactory for a full model of the Universe, and may indicate we are still missing some fundamental point. On the other hand, several sectors of the SM are not fully understood, either because of the huge range of the constants involved (like the fermion masses, whose observed values range over fourteen orders of magnitude), or because the physical principles underlying the formalism are still lacking (like quark and lepton mixing and CP violation, well described by the CKM paradigm but whose origin is unclear).

For these reasons one is led to believe there is new physics beyond the SM, whose evidence can be searched for by experiment and observation. In search for new physics beyond the SM, two approaches can be taken. The first approach is to search directly for new, heavy particles by producing them at increasingly powerful accelerators or by observation of the interactions of very-high-energy cosmic rays. The second approach is to measure, to very high precision, quantities predicted by the SM in the attempt to find discrepancies between the measured and predicted quantities. One example of this second approach is the

measurement of the branching ratio of the ultra-rare decay $K^+ \to \pi^+ \nu\bar{\nu}$, which is the main goal of the NA62 experiment at the CERN SPS.

The Branching Ratio of the $K^+ \to \pi^+ \nu\bar{\nu}$ process can be computed to an unusually high degree of precision, being the pure theoretical uncertainties below 4%. Such a precision is achievable as the transition is dominated by perturbatively computable short-range processes, namely Feynman diagrams with a top quark loop, for which the hadronic matrix elements can be derived from the well measured $K^+ \to \pi^0 e^+ \nu$ branching ratio. Thanks to its theoretical cleanness, this decay is one of the best probes for new physics effects complementary to direct searches, especially within non Minimal Flavor Violation models. As a Standard Model test, the Branching Ratio provides a measurement of V_{td} free from any hadronic uncertainty and independent from the one obtained from B meson decays.

1.2 The CKM Matrix

In the Standard Model the six flavors of quarks are grouped into three generations or families, each generation consisting of a quark doublet: up (u) and down (d), charm (c) and strange (s), and top (t) and bottom (b). For each of these doublets, the first quark has an electric charge of $+2/3e$ and the second has an electric charge of $-1/3e$, e being the positron charge. Through weak transitions these quarks can transform from one to another. The charged-current processes, mediated by the charged intermediate W^+ or W^- bosons, allow a quark to transform into another quark having a different charge, such as $d \to u + W^-$. Conversely, the neutral-current process, mediated by the neutral Z^0 boson, cannot directly transform a quark into another quark of different flavor but same charge, i.e. $s \to d$. This type of transformation, known as a flavor-changing neutral current process (FCNC), is forbidden at tree level and suppressed at one-loop level due to the Glashow-Iliopoulos-Maiani (GIM) mechanism. The weak charge-changing interactions lead primarily to the transitions $u \leftrightarrow d$, $c \leftrightarrow s$, $t \leftrightarrow$b between left-handed quarks (u, c, t) of charge 2/3 and those (d, s, b) of charge -1/3. However, as noted by Cabibbo for two families of quarks and by Kobayashi and Maskawa for three, additional transitions of lesser strength can be incorporated into this framework in a universal manner. The charge-changing transitions then connect u, c, t not with d, s, b but with a rotated set (d', s', b') = V (d, s, b), where V is a unitary 3×3 matrix now known as the Cabibbo-Kobayashi-Maskawa (CKM) matrix. In the Standard Model the elements of V and the quark masses are intimately connected, since the V matrix arises as a result of diagonalization of the quark mass matrices. By restricting to two generations, a single angle θ_c is sufficient to describe mixing:

$$\begin{pmatrix} d' \\ s' \end{pmatrix} = \begin{pmatrix} cos\theta_c & sin\theta_c \\ -sin\theta_c & cos\theta_c \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix}$$

where θ_c is the Cabibbo angle, whose experimental value is $\simeq 13^\circ$. The mixing of three generations requires three angles, also allowing for a single complex phase able to originate CP violation as observed in some weak processes, like the decays of neutral kaons. A standard representation of the CKM matrix is the following

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} \begin{pmatrix} c_{12}c_{13} & s_{12}s_{13} & s_{13}e^{-i\delta_{CP}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{CP}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{CP}} & s_{23}s_{13} \\ s_{12}s_{23} - c_{12}c_{13}s_{23}e^{i\delta_{CP}} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta_{CP}} & c_{23}c_{13} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{1.1}$$

where $s_{ij} \equiv \sin(\theta_{ij})$, $c_{ij} \equiv \cos(\theta_{ij})$ and δ_{CP} denote the CP-violating phase, while θ_{ij} is the mixing angle between i and j quark-families.
Experimentally, since s_{13} and s_{23} are very small ($\approx O(10^{-3})$ and $\approx O(10^{-4})$ respectively), both c_{13} and $c_{23} \approx 1$ as a result of the unitarity. The independent parameters are in this case $s_{12} =\mid V_{us} \mid, s_{13} =\mid V_{ub} \mid, s_{23} =\mid V_{cb} \mid$ and δ. They can be evaluated from tree-level decay in which there are the transitions $s \to u, u \to b$ and $c \to b$. This parametrization shows immediately that the CP phase is always multiplied by s_{13}. As a consequence, the CP violation is quite suppressed regardless the value of δ.
Mixing and flavor changing do not take place at tree level in neutral current processes, just because of the flavor structure of neutral current:

$$\begin{aligned} J^0 &= \bar{u}u + \bar{c}c + \bar{d}'d' + \bar{s}'s', \\ &= \bar{u}u + \bar{c}c + (\bar{d}d + \bar{s}s)cos^2\theta_c + (\bar{d}d + \bar{s}s)sin^2\theta_c \\ &\quad + (\bar{d}s + \bar{s}d - \bar{d}s - \bar{s}d)sin\theta_c cos\theta_c \\ &= \bar{u}u + \bar{c}c + \bar{d}d + \bar{s}s \end{aligned} \tag{1.2}$$

The strong suppression of flavor changing neutral current processes also at one loop can be understood by making reference to the GIM mechanism, where almost exact CKM cancellation of diagrams with different virtual quark internal lines strongly reduces the decay amplitude.

A second parametrization of CKM was suggested by L. Wolfenstein. Expanding in powers of the parameter $\lambda =\mid V_{us} \mid \approx 0.23$ every term of equation (1.1), by setting

$$s_{12} = \lambda \ , s_{23} = A\lambda^2, s_{13}e^{-i\delta} = A\lambda^3\rho - i\eta$$

one obtains

$$\begin{pmatrix} 1-\frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho-i\eta+i\eta\frac{\lambda^2}{2}) \\ -\lambda & 1-\frac{\lambda^2}{2}-i\eta A^2\lambda^4 & A\lambda^2(1+i\eta\lambda^2) \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} \tag{1.3}$$

The current experimental values of the matrix elements are [2]:

$$\begin{pmatrix} |V_{ud}| = 0.97425 \pm 0.00022 & |V_{us}| = 0.2252 \pm 0.0009 & |V_{ub}| = (4.15 \pm 0.49) \times 10^{-3} \\ |V_{cd}| = 0.230 \pm 0.011 & |V_{cs}| = 1.006 \pm 0.023 & |V_{cb}| = (40.9 \pm 1.1) \times 10^{-3} \\ |V_{td}| = (8.4 \pm 0.6) \times 10^{-3} & |V_{ts}| = (42.9 \times 2.6) \times 10^{-3} & |V_{tb}| = 0.89 \pm 0.07 \end{pmatrix}$$

It is evident that diagonal elements are dominant and represent the most favored quark transitions, while off-diagonal elements describe suppressed transitions. The unitarity of the CKM matrix $VV^+ = 1$ results into six unitarity conditions for its elements, which can be represented in the form of six triangles in the complex plane, all with the same area

$$Area = \frac{1}{2}J_{CP} = s_{12}^2 s_{23} s_{13} c_{12} c_{23} c_{13} sin\delta \approx \frac{1}{2}A^2\lambda^6\eta \tag{1.4}$$

which is a measure of CP violation in the quark sector. Leaving aside charm decays, where CP violation is expected and observed as a very small effect, K and B decays provide an interesting test-bed for precise checks of the CKM framework. They also allow for sensitive search of physics beyond the Standard Model, by accurate comparison of independent estimates of a particular unitarity triangle. Indeed, the unitarity condition most relevant for the $K \to \pi\nu\bar{\nu}$ decay can be written as

$$\begin{aligned} V_{ub}^* V_{ud} + V_{cb}^* V_{cd} + V_{tb}^* V_{td} = 0 \\ \to V_{ub}^* - \lambda V_{cb}^* + V_{td} \approx 0 \end{aligned} \tag{1.5}$$

where the approximations $V_{ud} \sim V_{tb}^* \sim 1$ and $V_{cd} \sim -\lambda$ have been made. The corresponding unitarity triangle is depicted in figure 1.1. In figure 1.1 all sides have been scaled by the

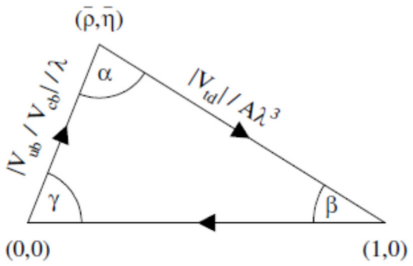


Figure 1.1. Unitary triangle corresponding to the equation 1.5.

length of the baseline. In this normalized unitarity triangle, the apex is given by the two Wolfenstein parameters

$$\bar{\rho} = \rho(1 - \frac{\lambda^2}{2}), \bar{\eta} = \eta(1 - \frac{\lambda^2}{2}) \tag{1.6}$$

As can be seen in figure 1.2.

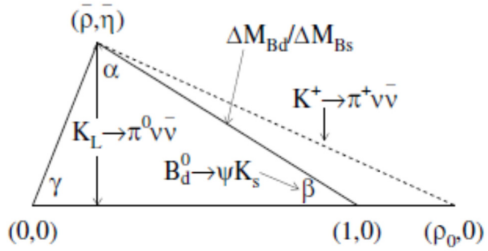


Figure 1.2. Possible measurements to determine ρ and η.

independent measurements of ρ and η from B mixing and decay asymmetries and K decays to $K^+ \to \pi\nu\bar{\nu}$ can be used either to validate Standard Model predictions of CP violation coming from a single CKM phase, or to signal new physics.

1.3 Ultra-rare K decays

The decay $K^+ \to \pi^+\nu\bar{\nu}$ is a FCNC process, forbidden at the three level but allowed at one loop. The diagrams relevant to the process at the leading level are shown in figure 1.3. The

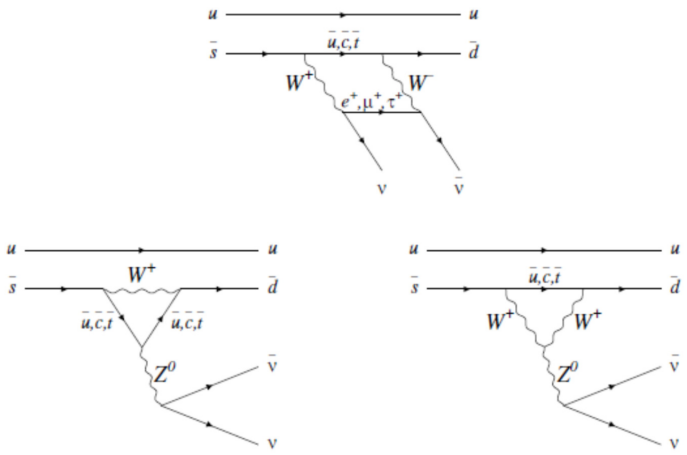


Figure 1.3. $K^+ \to \pi^+\nu\bar{\nu}$ diagrams.

first one is a *box*, while the following two are *penguin* diagrams. The standard method to compute transition amplitudes in light flavor weak decays is to use an effective theory, namely a simplified, low energy version of the full electroweak interaction where a large masses are integrated out. The whole procedure results into an effective Hamiltonian, to be used to compute amplitudes. A well known example is the Fermi theory of beta decay, a current-current effective interaction which correctly describes many charged current processes in the limit of momentum transfer q^2 small as compared to M_W^2. For the $K^+ \to \pi^+ \nu\bar{\nu}$ decay, as for similar semileptonic decays, the effective Hamiltonian is written as

$$H_{eff} = \sum_{l=e,\mu,\tau} \frac{G_F}{\sqrt{2}} \frac{\alpha}{2\pi sin^2\theta_W} (V *_{cs} V_{cd} X^l + V *_{ts} V_{td} X(x_t))(\bar{s}d)_{V-A} (\bar{\nu}_l \nu_l)_{V-A} \tag{1.7}$$

where

- G_F is the Fermi coupling constant;
- α is the fine-structure constant;
- θ_W is the Weinberg angle;
- V_{ij} are the CKM matrix elements;
- $x_t = \frac{m_t^2}{M_W^2}$ where m_t is the top-quark mass and M_W is the W mass;
- X^l with $l = e, \mu, \tau$ are three functions describing the charm contribution;
- $X(x_t)$ is a function encoding the dominant top contribution;
- $(\bar{s}d)$ and $(\bar{\nu}\nu)$ are the quark and lepton neutral weak currents with vector - axial (V - A) structure:

$$\left[\bar{s}\gamma_\mu(1-\gamma_5)d\right]\left[\bar{\nu}\gamma_\mu(1-\gamma_5)\nu\right]. \tag{1.8}$$

The sum over lepton flavors includes standard current-current operator multiplied by CKM and kinematical factors, which are lepton flavor independent for t loops, while somewhat lepton flavor dependent for c loops , due to tau mass effects in box diagrams. Surprisingly, the up quark contribution to H_{eff} is negligible, in spite of the relative CKM factor being large as compared to those of c and t contributions. The reason is the 'hard' GIM mechanism suppression (i.e. quadratic in the mass of the internal quark for $m_q/M_W << 1$) which is active in this transition, as a result of box and penguin diagrams reduction. As

shown in figure 1.4, the net effect is make the top contribution dominant, in spite of the tiny CKM factor. Charm contribution cannot be neglected, while the u quark term is negligible. The effective Hamiltonian above is written in terms of partonic currents: In order to get

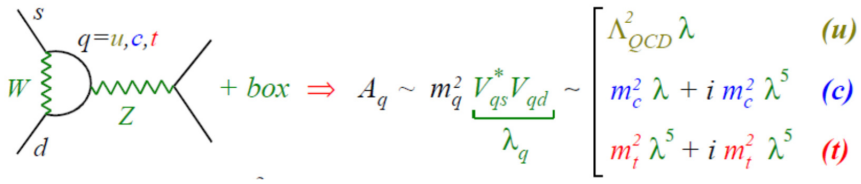


Figure 1.4. $K^+ \to \pi^+ \nu\bar{\nu}$ transition contributions.

amplitudes at the hadron level, one has to take matrix elements of the quark current operators between hadron states.

$$\langle \pi | (sd)_{V-A} | K \rangle \tag{1.9}$$

Given the low energy and the light quarks involved, no perturbative approach is available for this step. However, as mentioned above top (and also charm) quark dominance implies the process is essentially driven by short distance dynamics, so that the relation between partonic and hadronic amplitudes is unusually accurate. The matrix element of the charged current $\bar{s}\gamma^\mu d$ between a kaon and a pion state is related by isospin symmetry to those entering Kl3 decays, which are very well measured. Indeed, one has for Ke3 H_{eff}

$$H_{eff}(K^+ \to \pi^0 e^+ \nu) = \frac{G_F}{\sqrt{2}} V_{us}^* (\bar{s}u)_{V-A} (\bar{\nu}_e e)_{V-A} \tag{1.10}$$

So, by isospin symmetry:

$$\langle \pi^+ | (\bar{s}d)_{V-A} | K^+ \rangle = \sqrt{2} \langle \pi^0 | (\bar{s}u)_{V-A} | K^+ \rangle \tag{1.11}$$

and to a good approximation:

$$\frac{B(K^+ \to \pi^+ \nu\bar{\nu})}{B(K^+ \to \pi^0 e^+ \nu)} = \frac{\alpha^2}{|V_{us}|^2 2\pi^2 sin^4\theta_W} \sum_{l=e,\mu,\tau} |V_{cs}^* V_{cd} X_{NL}^l + V_{ts}^* V_{td} X(x_t)|^2 \tag{1.12}$$

The equation above can be written summing over three neutrino flavors [3]

$$BR(K^+ \to \pi^+ \nu\bar{\nu}) = \kappa_+ (1 - \Delta_{EM}) \left[\left(\frac{Im\lambda_t}{\lambda^5} X(x_t) \right)^2 + \left(\frac{Re\lambda_c}{\lambda} P_c(X) + \frac{Re\lambda_t}{\lambda^5} X(x_t) \right)^2 \right] \tag{1.13}$$

where $\lambda = |V_{us}|$ and $\lambda_q = V *_{qs} V_{qd}$ contain the parameters of the CKM matrix. The hadronic matrix element of the effective Hamiltonian can be extracted from the well-measured semi-leptonic process $K \to \pi l \nu$, including isospin breaking, long-distance effects and QED radiative corrections. These contributions are contained in the parameters k_+, including Next to Leading Order (NLO) and partially Next Next to Leading Order (NNLO) corrections obtained in chiral perturbation theory, and Δ_{EM}. The $kappa_+$ parameter can be written as:

$$\kappa_+ = r_{K^+} \frac{3\alpha^2 BR(K^+ \to e^+ \pi^0 \nu_e)}{2\pi^2 sin^4 \theta_W} \lambda^8 \tag{1.14}$$

The parameter $r_K^+ = 0.901$ describes three isospin breaking corrections in relating $K^+ \to \pi^+ \nu \bar{\nu}$ to $K^+ \to e^+ \pi^0 \nu_e$. Other numerical parameters are [4]

- BR($K^+ \to e^+ \pi^0 \nu_e$) = $(5.07 \pm 0.04)\%$;
- $\alpha = \frac{1}{127.9}$;
- $\sin^2 \theta_W = 0.231$.

The final numerical value is [5]

$$\kappa_+ = (5.173 \pm 0.025) \dot{1} 0^{-11} \left[\frac{\lambda}{0.225} \right]^8 \tag{1.15}$$

where the main source of uncertainty regarding κ_+ derives from the BR($K^+ \to e^+ \pi^0 \nu_e$).

The fully inclusive correction is $\Delta_{EM} = -0.003$ [3].

The top quark term can be expanded as

$$X(x_t) = X^{(0)}(x_t) + \frac{\alpha_s(m_t)}{4\pi} X^{(1)}(x_t) + \frac{\alpha}{4\pi} X^{EM}(x_t) \tag{1.16}$$

with α_s the QCD coupling constant, while $X^{(0)}$ represents the leading-order (LO) contribution [6], $X^{(1)}$ is the NLO QCD correction [7], and the last term corresponds to the recently computed two-loop electroweak correction [8]. The current value for the function relevant to the top quark is[9][10]:

$$X_t = 1.469 \pm 0.017 \pm 0.002$$

where the first error quantifies the remaining scale uncertainty of the QCD corrections, and the second error corresponds to the uncertainty of the electroweak corrections. The dominant contribution to the error comes from the experimental error on the top quark mass ($\sim 3\%$).

The charm quark contribution is summarized through the parameter $P_c(X)$ which is written as

$$P_c(X) = P^{SD}(X) + \delta P_{c,u} \tag{1.17}$$

The short-distance part is given by $P_c^{SD}(X) = \frac{1}{\lambda^4}\left[\frac{2}{3}X^e + \frac{1}{3}X^\tau\right]$ The distinction between charged lepton flavours in the box diagrams is irrelevant in the top contribution since $m_\tau >> m_l$, but is relevant in the charm contribution as $m_\tau > m_c$.
The current value for the short-distance contribution is $P_c^S D(X) = 0.404 \pm 0.024$ Using all the above numerical inputs, the current theoretical prediction for the Branching Ratio of the $K^+ \to \pi^+ \nu\bar{\nu}$ decay is

$$BR(K^+ \to \pi^+ \nu\bar{\nu}) = (8.4 \pm 1.0) \times 10^{-11} \tag{1.18}$$

The attempts to observe this ultra-rare decay mode of the K^+ have a long history, as shown in figure 1.5. The BNL E787/E949 experimental result[1], based on seven observed

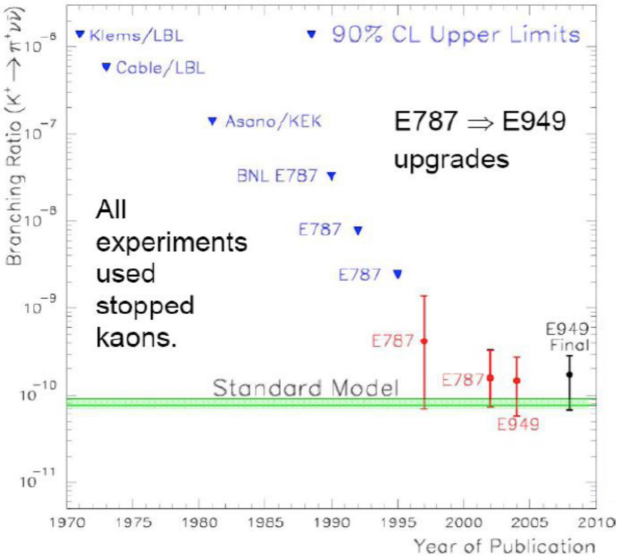


Figure 1.5. Different measurements of the $\pi\nu\bar{\nu}$ branching ratio.

events is

$$BR(K^+ \to \pi\nu\bar{\nu}) = 17.3^{11.5}_{-10.5} \times 10^{-11}$$

Even though the obtained value is compatible with the Standard Model prediction, a significant gap exists between the theoretical precision and the present experimental error.

The NA62 experiment at CERN aims to improve the result by a factor 10 bringing the experimental error close to the theoretical one.

Chapter 2

The NA62 Apparatus

The NA62 experiment is located at the CERN Super Proton Synchrotron accelerator. The following scheme shows the entire setup (not in scale), with the hadron beam coming from the left.

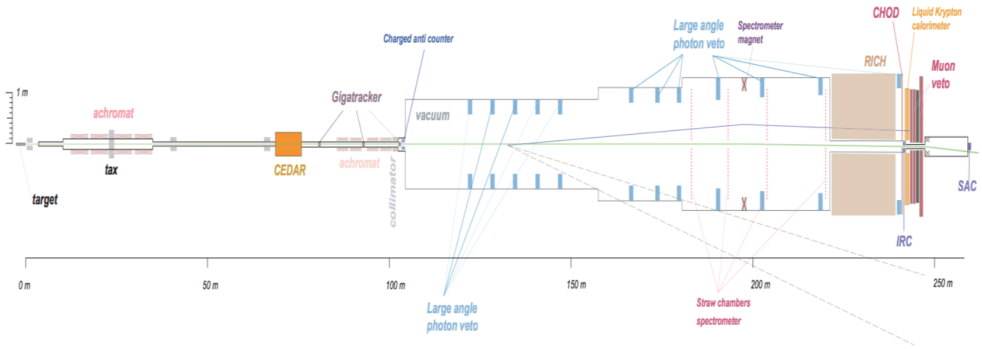


Figure 2.1. The NA62 experiment layout.

NA62 uses a decay-in-flight technique to study kaon decays, well matched to the characteristics of the CERN SPS.

In general it is possible to divide the apparatus in three macro-regions. The first is the *upstream region* where the beam is tracked. The second is the *fiducial volume*, a 60 m long vacuum region, which is the decay acceptance region. Finally, in the *downstream region* the properties of the decayed particles are measured. The largest detectors have an approximately cylindrical shape around the beam axis with a radius up to about 2 m and down to 10 cm in order to let the very intense flux of undecayed beam particles pass-through without affecting the active area.

Timing, spatial and angular information are needed to completely match kaons with the tracks of the decay products. To fulfill this request, both the beam particle and the secondary charged products must be accurately measured using spectrometers. Electromagnetic and hadronic calorimeters detect photons, positrons and pions. Furthermore, a powerful particle identification system to distinguish π^+ from μ^+ and e^+ complements the tracking and calorimeter detectors to reach the required sensitivity, also guaranteeing redundancy. A muon veto system completes the apparatus.

In the following part of this chapter, a brief description of all the systems is reported.

2.1 Beam

The incoming beam [11] contains $75\frac{GeV}{c}$ positive particles derived from a primary 400 $\frac{GeV}{c}$ proton beam extracted from the CERN SPS accelerator and impinged on a beryllium target. The high intensity beam ($750 \times 10^6 \frac{particles}{second}$) and the high energy of the particles do not allow to purify the beam before entering in the detector area. The beam composition is on average: 71% pions, 23% protons and 6% K^+. By choosing a negative beam the kaon fraction would be even smaller being the ratio $\frac{\frac{K^+}{\pi^+}}{\frac{K^-}{\pi^-}} \approx 1.2$. Finally, in order to maximize the ratio between $K^+ \to \pi^+ \nu\bar{\nu}$ and the incident flux, the selected beam momentum of 75 GeV/c is the best compromise between the kaon production and the number of decays in a 50 meters region.

2.2 Kaon Tag

Being based on an unseparated beam, the experiment requires a method to tag the kaons among the ≈17 times more intense particle flux. Kaon tagging is performed by a Cherenkov counter detector, called CEDAR/KTAG. An adjustable diaphragm can be set up to only let

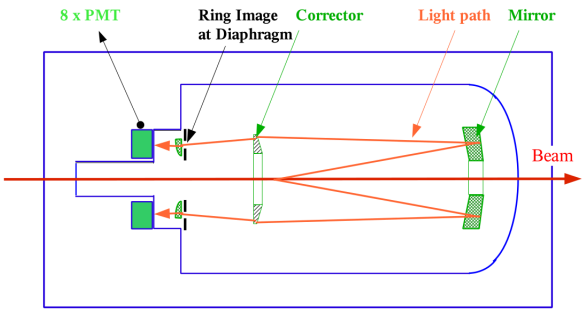


Figure 2.2. The CEDAR schematic layout.

through the Cerenkov light emitted by 75 GeV/c kaons. The CEDAR/KTAG detector reaches a time resolution of 70 ps[12], and an efficiency of about 95% with a incident kaon flux that is approximately 50 MHz. The good time resolution performance is obtained by averaging the arrival time of the signal given by about 100 photons produced by each kaon.

2.3 Gigatracker

The beam spectrometer, called Gigatracker (GTK), which provides with high precision time, direction and momentum of all incoming particles, consists of three stations of hybrid silicon pixel detectors installed around four dipole magnets.

The stringent requirement of low material budget has led[13] to the choice of a sensor thickness of $200\mu m$, which is a good compromise between thickness, corresponding to 0.22% radiation length X_0 and the signal pairs yield, with an expected average number of ≈ 15000 e-h pairs per minimum ionizing particle.

For the readout[14], ten Application Specific Integrated Circuit (ASIC) chips are bump bounded to the sensor, arranged in two rows of five chips each. The size of each sensor is

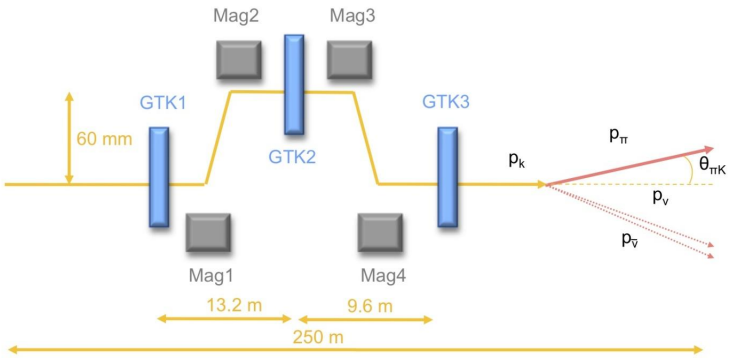


Figure 2.3. The GigaTracker layout.

$63.1 \times 29.3 mm^2$, composed by 18000 $300 \times 300 \mu m^2$ pixels arranged in a matrix of 90×200 elements. In order to reduce the rate of beam interactions with the sensor, a limit of $0.5\% X_0$ for the material traversed by the beam was set. To satisfy this request the thickness of the read out chip is thinned to $100\mu m$. Furthermore the material budget allowed to the cooling system is about 100-150 micron of Silicon. The GTK is designed to operate in vacuum and the space between GTK 1 and GTK 3 is occupied by four dipole magnets deflecting incoming particles. The first two magnets deflect the beam off-axis, whilst the last two redirect the beam back on-axis. The GTK 2 station, which is shifted $60mm$ vertically provides measurement of the particle displacement, yielding a momentum precision of about $dp/p = 0.2\%$.

The GTK measurements are matched with the ones obtained by the downstream spectrometer (Straw Tracker - see section 2.6). The expected beam angular resolution is $\sigma_\theta = 16\mu rad$ and the momentum relative resolution is $\frac{\sigma_{P_K}}{P_K} = 0.2\%$.

2.4 Charged Anticounters

A positive pion, produced by the beam impacting GTK and detected by the apparatus of NA62 with the origin erroneously reconstructed in the fiducial region, may mimic a signal event. To reject this source of background, the Charged Anticounter [13] (or CHANTI) detector is placed after the last station of the GTK. Furthermore the CHANTI can tag beam halo muons in the region close to the beam.

The CHANTI is composed by six rings extending from 5 to 15 cm radius plus an extra ring from 18 through 60 cm. Each ring consists of two planes of scintillators, 1cm thick.

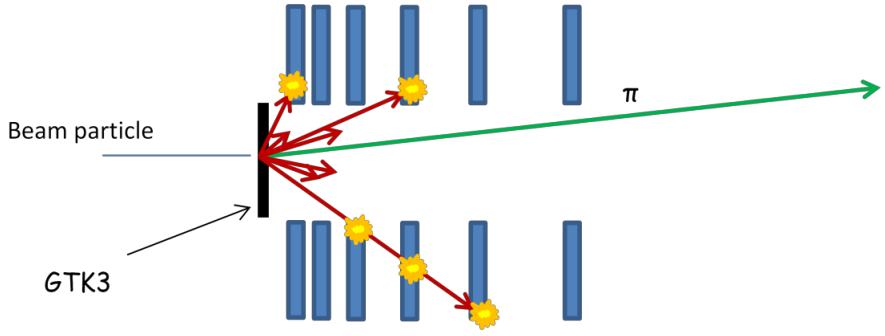


Figure 2.4. The Chanti layout.

2.5 Photon Veto Detectors

In order to reach the sensitivity required to observe the $\pi\nu\bar{\nu}$ signal, the background coming from the decay $K^+ \to \pi^+\pi^0$ should be suppressed with an inefficiency of $\sim 10^{-8}$ or smaller. The photon veto system provides a hermetic coverage at angles up to 50 mrad for photons generated from decays in the fiducial region. Monte Carlo simulations show that in this configuration only 0.2% of events has one photon emitted at larger angles [13].

The photon veto system is composed by three different detectors:

- Large Angle Vetoes (LAV), covering the angles between 8.5 and 50 mrad;
- Liquid Krypton Calorimeter (LKr), covering the angles between 1 and 8,5 mrad;

- small angles vetoes, covering the region down to zero degrees (SAC detector) and the zone inside the inner part of the Ring Imaging CHerenkov detector (RICH) and LKr calorimeter (IRC detector).

2.5.1 Large Angle Veto Detector

The LAV detector consists [13] of twelve stations located between 120 and 140 m along the beam direction. Eleven of them are placed around the vacuum thank, while the twelfth is inserted between LKr and RICH.

Each module is made of a circular ring of active-material surrounding the beam pipe. A particle generated by decays with angle between 8.5 and 50 mrad crosses the detector. The basic building blocks of these detectors are lead glass crystals read out by photomultipliers (PMT) from the former OPAL electromagnetic calorimeter.

Each LAV station is made up of four or five rings where all the counters lie in one plane which is perpendicular to the beam line. Each lead glass crystal has a square-based truncated

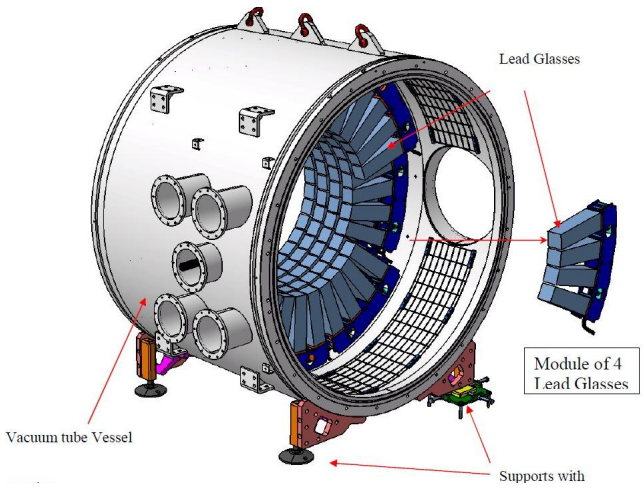


Figure 2.5. The LAV layout.

pyramid shape, the larger base being of $11 \times 11 cm^2$ and the smaller base of $9 \times 9 cm^2$, with a height of 37 cm. The lead glass composition is dominated by a 74% of PbO. A cylindrical 2 cm long light-guide made of plexiglass connects the counter to the PMT.

The system detects photons or muons/pions in the beam halo achieving a time measurement with a $\approx$1 ns resolution and an energy measurement with a precision of the order of 10%.

2.5.2 Liquid Krypton

The Liquid Krypton Calorimeter (LKr) is the same detector used in the experiment NA48. It is a key element for vetoing photons, given its inefficiency less than 10^{-5} for photons with an energy greater than 35 GeV. The LKr also provides trigger signals based on energy deposition for the first level of trigger (see chapter 4).

The Liquid krypton was adopted in a quasi-homogeneous configuration, allowing for the full ($> 99\%$) development of electromagnetic showers within a cylinder of $5.3m^2$ cross-section and 127 cm depth, corresponding to 27 radiation lengths ($X_0 = 4.7cm$). In order to obtain a good transverse granularity, the LKr is segmented in 13212 cells of size 2×2 cm^2 [15]. It has also a 16 cm diameter hole to allow the passage of the non-decayed particles. To maintain Krypton in a liquid phase,the vessel is cooled to $(121 \pm 0.1\%)K$. This system, vessel and cryostat, introduces $\sim 0.65X_0$ of passive material, equivalent to 50 MeV energy loss for a minimum ionizing particle.

Because the NA62 beam rate is two orders of magnitude higher than NA48, which was designed for rates up to 10 kHz, the LKr electronics has been completely redesigned to cope with the higher rate. The new system will use commercial components such as standard DDR2 memories, gigabit ethernet, network switches and standard PCs.

2.6 Straw Tracker

The Straw Tracker detector is the downstream spectrometer meant to measure direction and momentum of charged products generated by kaon decays. To reduce the multiple scattering of the decay-generated particles, the tracker is operated in vacuum. The detector[13] consists of four straw chambers, the first two sitting if front of a dipole magnet, while two more placed downstream. The typical magnetic field is 0.36 T, giving a momentum kick of 270 MeV/c (see figure 2.6).

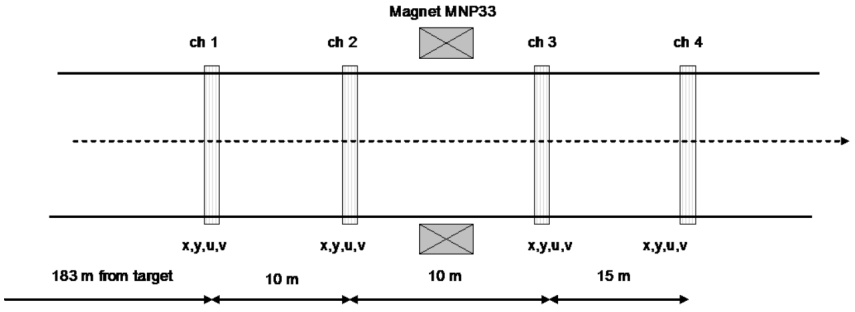


Figure 2.6. The magnetic spectrometer layout.

Each chamber is made out of 1792 straw tubes arranged in four planes to yield the typical four views X,Y,U, and V. Each view consists of 448 ultra-light straw tubes 2.1 m long and 9.8 mm in diameter filled with a gas mixture of CO2 (90%), Isobutan C4H10 (5%) and CF4 (5%).

The main requirements for the detector are:

- spatial resolution $\leq 130\mu m$ per view and $\leq 80\mu m$ per space point;
- thickness $\leq 0.5\%$ of a radiation length for each chamber;
- sustainable particle flux up to 40 kHz/cm, and up to 500 kHz/Straw.

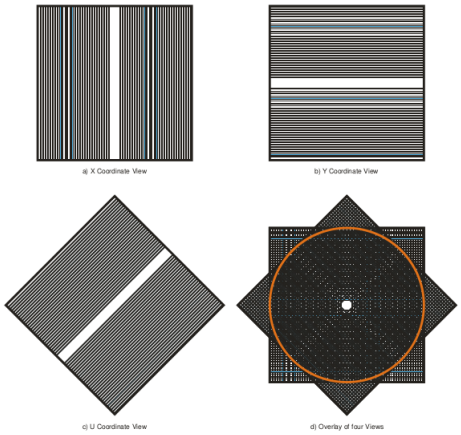


Figure 2.7. Schematic layout of the four view chamber.

2.7 Ring Imaging Cherenkov Detector

The Ring Imaging Cherenkov counter (RICH)[13] is placed between the forth Straw Chamber and the LKr calorimeter. The RICH detector is designed to separate muon from pion tracks in the $15 \div 35$ GeV/c momentum range. The detector is filled with atmospheric-pressure neon gas ($\sim 5.6\%X_0$) as radiating medium, contained in a cylindrical vessel 18 m long and 2.5 m wide around the beam pipe.

In a RICH detector, the Cherenkov light is reflected by a mosaic of mirrors, placed at the downstream end of the gas volume, onto the focal plane, located at the upstream end of the gas container. There are two regions (to avoid shadowing the beam pipe) in which ~ 1000 photomultipliers are arranged, collecting Cherenkov photons. The muon contamination in the pion sample is $< 1\%$ between 15 and 35 GeV/c[16]. Thus, when coupled to the Muon Veto

System (MUV), the muon inefficiency of the system is reduced by two orders of magnitude. Furthermore the RICH provides the particle crossing time measurement with a resolution better than 100 ps[16].

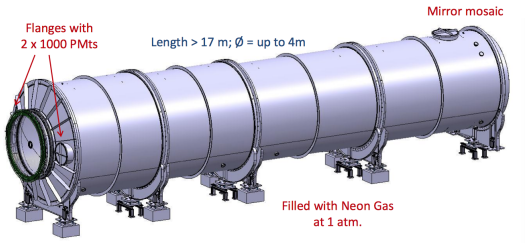


Figure 2.8. The RICH detector layout.

The high time resolution of the RICH detector drove the choice to use this system as time reference for the first level of the NA62 trigger, so that the coincidences between the different trigger data sources are computed within a time window centered on the RICH reference time. Going a bit into detail, the firmware that generates the trigger data, based on the number of signals in time-coincidence coming from the PMTs, does not count these hits directly from all PMTs. Instead, a dedicated TEL62 read out board (see section 2.13.1) computes the output of 8 channels digital adders (called *SuperCells*). We will refer to the SuperCells when we will discuss the RICH efficiencies in generating L0 trigger data in chapter 5.

2.8 Charged Hodoscope

The Charged Hodoscope (CHOD) is the same detector used by the NA48 experiment, made of a system of scintillation counters with high granularity. Two planes of scintillators (64 in horizontal and 64 in vertical position) 2 cm thick are placed between the RICH and the LKr detectors.

The CHOD is useful to determine time and position of charged particles, in order to provide trigger signals to select events with one or more charged tracks and to associate tracks in time coincidence with kaons tracks in the GTK detector. Given the high rate in the GTK, the CHOD must have a reliable efficiency and a time resolution of 100 ps or better.

Such a resolution can be achieved only with offline analysis, just because the counter length of 2 m spoils the resolution due to the light propagation delay. By associating horizontal and vertical counters, one can correct for the light propagation delay.

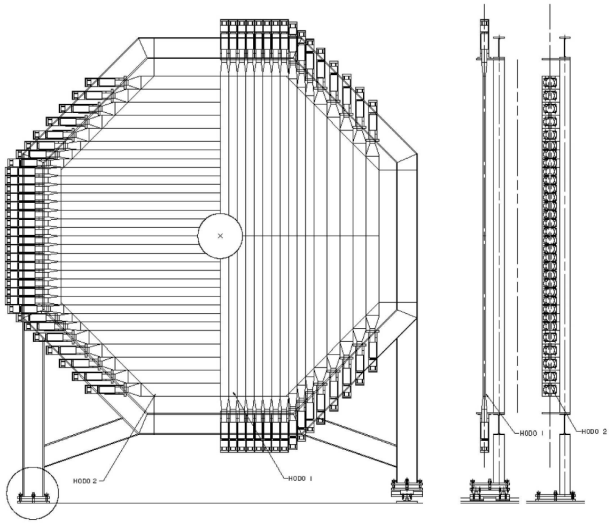


Figure 2.9. The CHOD layout.

2.9 New Charged Hodoscope

NA62 has been equipped also with a new charged hodoscope (NewCHOD), used for the first time during the 2016 run together with the old CHOD. The main reason to build the NewCHOD is the high hit rate on the long slabs (1 m) of the CHOD. The intrinsic dead time and the light transit time inside the scintillator are not compatible with the expected total rate on the detector (above 10 MHz). The NewCHOD is a two-dimensional array of 152 scintillator tiles installed after the LAV12. In each quadrant, a 30 mm thick plastic scintillator is divided into 38 tiles (see figure 2.10). Light is collected and transmitted by 1 mm diameter wavelength shifting fibres connected to arrays of 3 X 3 mm^2 silicon photomultipliers (SiPMs) on dedicated mother-boards located on the periphery of the detector. A maximum rate of the order of 500 kHz is expected on the tiles close to the beam pipe.

2.10 Muon Veto System

The muon veto system (MUV) consists of three different detectors (MUV1, MUV2 and MUV3) placed after the LKr calorimeter. MUV1 and MUV2 are hadronic calorimeters, used to measure the deposited energies and shower shapes of incident particles. On the other hand, MUV3 detects the emerging particles, which are mostly muons. MUV2 is the front module of the former NA48 hadron calorimeter (HAC), but turned by 180°, while MUV1 is a completely new detector. Both of them are sampling calorimeters, with either 24 (MUV1) or 22 (MUV2) planes of scintillator strips oriented alternatively along the vertical and horizontal

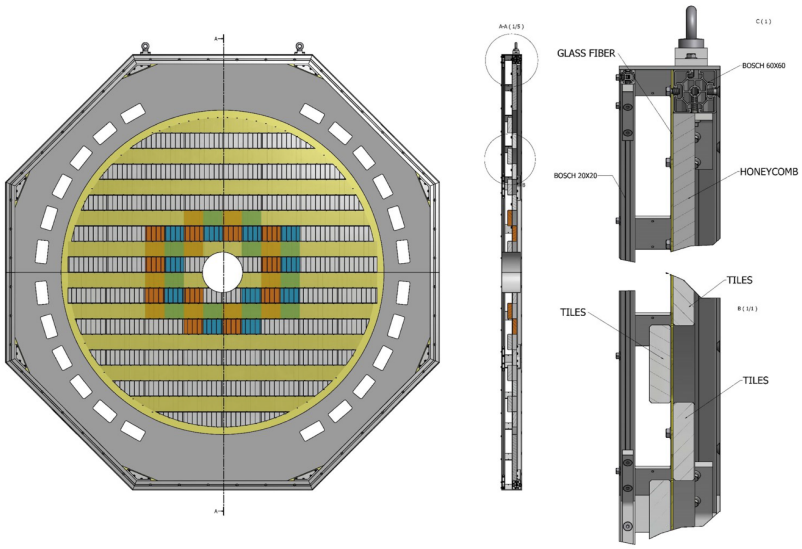


Figure 2.10. New CHOD Sketch, front and side views.

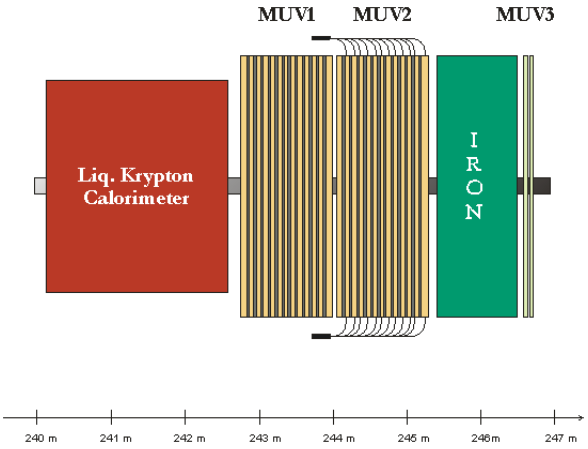


Figure 2.11. The muon veto system layout.

directions, arranged in a sandwich with either 25 (MUV1) either 23 (MUV2) iron layers. The probability that a pion reaches the MUV3 is 1.1×10^{-7} [17].

MUV3 - or Fast Muon Veto - is placed after MUV1, MUV2 and a 80 cm thick iron wall. It is the key element to detect non-showering muons and serves as fast muon veto at the lowest trigger level (see chapter 4). MUV3 is made of plastic scintillators slabs, the light is collected by two PMT for each slab, both viewing the scintillator without light-guide. PMT's are positioned about 20 cm downstream the slab.

2.11 Small Angle Vetoes

The Small Angle Vetoes (SAV) consist of a Intermediate Ring Calorimeter (IRC) and of a calorimeter for small angles (SAC). The first one is placed directly on the beam pipe, just before the LKr detector, in order to veto the beam-halo particles. The latter is placed at the end of the experimental apparatus to detect only neutral particles, after a magnet dipole deflecting off charged particles.

2.12 NA62 Trigger and Data Acquisition System

2.12.1 Introduction

The NA62 experiment is characterized by a total flux of incoming particles of about 750 MHz and a total event rate of $\sim$ 10 MHz. For this reason it requires an high performance Trigger and Data Acquisition (TDAQ) system, in order to collect a reasonable amount of data, minimizing the dead time and random veto effects. The system has been designed to be very reliable, with an adequate bandwidth to cover the kaon-decay rate without introducing biases to measurement. The trigger logic is based on 3 different levels, as shown in figure 2.12:

- **Level Zero Trigger (L0)**: it is implemented with a commercial board, receiving the digitized information from a predetermined set of detectors. L0 compares the primitives to at least one of a pre-defined series of masks, sending the result of the computation to all sub-systems.
- **Level One Trigger (L1)**: it is software-based and checks first data quality. As a second step it requires simple correlations between conditions computed by individual detectors. As an example of the requests one can mention: a single charged pion in the detector with the momentum in a selected range, an odd number of tracks identical in the front and back of the STRAW, hit patterns in the RICH, further rejection from SAV calorimeters and MUV1-2 hadronic calorimeters.
- **Level Two Trigger (L2)**: as L1, it is software-based and handles the information coming from all detectors to make final cuts. It is applied to the final sample before storing it on the Central Data Recorder (CDR)[1]

[1] During the 2016 run, no L2-cuts have been implemented. It simply re-builds the event merging all the detector information after L0 and L1 trigger levels.

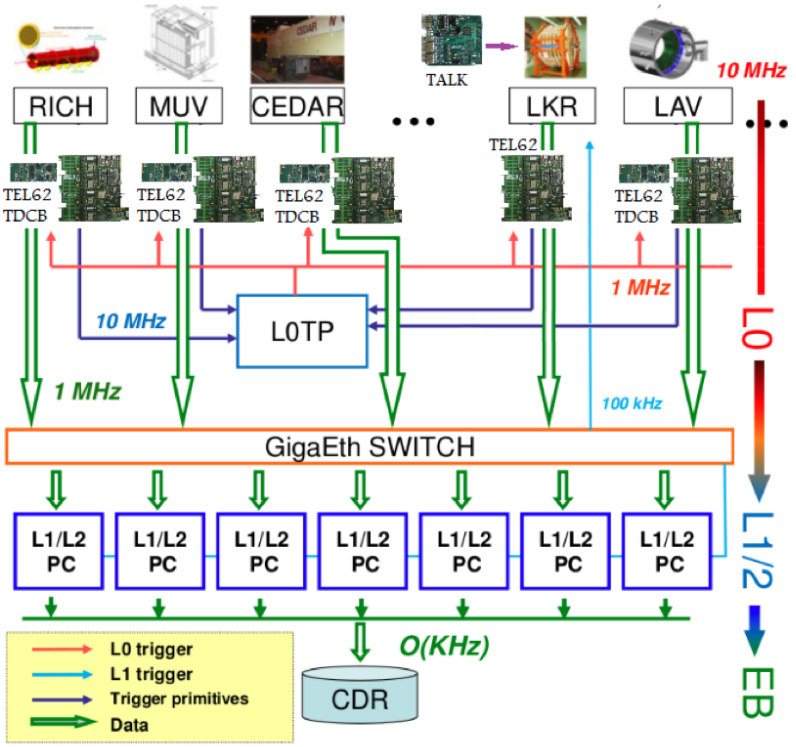


Figure 2.12. Schematic view of the NA62 TDAQ system.

Each individual detector locally elaborates event data and generates 64 bit words called **primitives** containing the information of time and encoded characteristics of the events.

The detectors that are part of the L0 trigger share the following type of data: calorimetetric energy (measured by LKr, MUV1, MUV2) and number of clusters or hits in the LKr and RICH, New CHOD, CHOD LAV and MUV3. The definition of cluster/hit depends on the on the specific firmware of single detectors. In what follows, the specific detectors contributing to the L0 trigger are reviewed with some detail.

Calorimeters : it includes the LKr and the MUV1 and 2, part of the veto system. They tag events with no more than a single cluster, allowing to veto events with electromagnetic shower, being the energy released above a given threshold.

CHOD : it tags decays into charged particles, allowing for selection of events with at least one prong. A hit is defined by a pair of signals occurring both in the vertical and horizontal planes within a given time interval;

MUV3 : it is used to veto events with muons and the muon halo components from π and K decays upstream of the final collimator;

RICH : it contributes to the reduction of the background generated by charged particles. The aim of the RICH as for L0 logic is to produce clusters of groups of hits belonging to the same Cherenkov ring. However, no spatial information is used and only clusters produced by hits in a certain time window are considered;

LAV : it tags the crossing of photons from $K^+ \to \pi^+\pi^0$. Being the energy threshold below the value of the energy released by a minimum ionizing particle, the LAV system can also identify and tag muons from the beam halo;

New CHOD : it improves the CHOD track identification. Thanks to the fact that the geometry of New CHOD is similar to the MUV3 one, the same firmware to generate events at L0 is used.

Data generated from the above detectors are sent via ethernet to the L0TP, with User Data Protocol (UDP) format for the transmission.
The UDP format has been chosen by taking into account the 270 m total length of the experiment. A commercial switch could be easily plugged to extend in space any possible other connection with detectors distant more than 100 m from the L0TP board without any modification of the protocol. The maximum latency of L0TP allows for absorbing the delay added by the switch. The detector read out should be able to store events for 1 ms before clearing their memories.
As already stated, with the nominal beam intensity the expected rate in the detector is quite high. Thus even at the trigger level a good time resolution is necessary. Indeed a typical value that has been achieved is roughly 1 ns. With such a figure the selection can be considered rather efficient, limiting as much as possible both fake triggers due to overlapping events and signal rejection due to random veto. For this reason, the collaboration was led to use a high-performance TDC for all the detectors. Being of paramount importance the time definition, a brief description of the different time measurement systems follows. Most of the detectors (CEDAR, CHANTI, LAV, RICH, CHOD, NEWCHOD and MUV3) adopted for the time measurement a TDC and FPGA-based common readout scheme, made of the TEL62 carrier board and the TDCB daughter-card. A few detectors, due to the specific application, adopted a different TDC system or an ADC based readout instead and these read-out systems are briefly described in the rest of this section.

The readout of the Gigatracker is based on ten TDCPix chips providing the hit time and the time-over-threshold of the signal generated by a particle crossing [14]. The TDCPix time bin size is about 100 ps, suitable to the expected rate of about 750 MHits/s. The digitized output is transmitted to the readout by four 3.2 Gbit/s serial links. The 4 output serial links send data to the carrier GTK-RO VME 6U board where they are temporary stored waiting for

a L0 trigger decision. After the reception of a L0 trigger signal the data inside a 75 ns time window are sent to the GTK-Farm, made of six different PCs, two per station. To reduce the rate, data are sent to the PC-Farm only after the L1 trigger signal.

The STRAW spectrometer electronics[18] is based on a 8-channels analogue front-end circuit, the CARIOCA chip[19], containing for each channel a fast pre-amplifier, a semi-Gaussian shaper, a tail cancellation circuitry, a base-line restorer and a discriminator. Two CARIOCA chips are integrated in a custom COVER board[20] together with an Altera Cyclone III FPGA [21]. The COVER board houses 16 pairs of TDC implemented within the FPGA with de-randomizers and an output link serializer. The full system is composed of 14236 TDCs producing a data rate of 2 GB/s. The back-end VME 9U Straw Readout Board (SRB) receives data from 16 COVER boards for processing (in FPGAs) and storage in DDR3 memories. A TTC interface (TTCex) [22] is also present, to optically receive the clock and the L0 trigger signal. After receiving a L0 trigger, data are extracted from the buffers and sent to the PC farm.

The LKr, the MUV1 and MUV2 use a Flash ADC based readout, whose basic element is the Calorimeter Readout Module (CREAM) [23]. It is a VME 6U board able to digitize 32 LKr channels at 40 MHz using four 8-channel, 14-bit ADCs. 414 such boards are needed to read out the 13248 calorimeter cells. After signal digitization, samples are stored in a circular buffer built inside a DDR3 module, waiting for the L0 trigger signal. Upon reception of such a signal through a custom backplane, data are moved to the L0 buffer, also built in the same DDR3 module. In order reduce the total amount of data, only triggers confirmed by the L1 are then sent to the PC-Farm.

When such signal is received, the corresponding data set is finally sent to the PC farm. The CREAM module also computes digital sums of 4×4 channels, called Super-cells, for L0 trigger purposes: two Super-cells are read out by each CREAM, and digital sums are sent every 25 ns to a system based on 36 TEL62s and custom interface mezzanine boards, where LKr L0 trigger primitives, based on energy deposit and cluster identification, are generated.

Both IRC and SAC small angle photon veto detectors present two different readout system: the TDCB-TEL62 and the CREAM board systems. Both IRC and SAC have only 4 channels, Due to the transverse size of an electromagnetic shower, a greater segmentation of these detectors would not have diminished the single channel rate.

2.12.2 Timing and Trigger Distribution

The system providing the time reference and the trigger distribution has been carefully designed, in order to guarantee a safe and adequate timing to the whole experiment. Referring to figure 2.13, one can identify a common, free-running, high-stability oscillator with a

frequency of 40.079 MHz, representing the clock reference for time measurement. We refer to this clock as the 40 MHz clock.

As expected for a fixed target experiment, kaons are not generated in short bunches, rather the beam delivery lasts for the entire burst. The burst is delimited by a Start of Burst (SOB) signal which is delivered by the SPS one second before the actual extraction of the beam, and by an End Of Burst (EOB), ~ 0.5 seconds after the end of the particle extraction. The entire duration of the burst is approximately 6 seconds, while the inter-burst period between two consecutive extractions is around 10 seconds.
A system composed by the Logic Trigger Unit (LTU)[24] + TTCex is used for encoding and delivering both the clock and the trigger data to the read-out systems of the experiment.

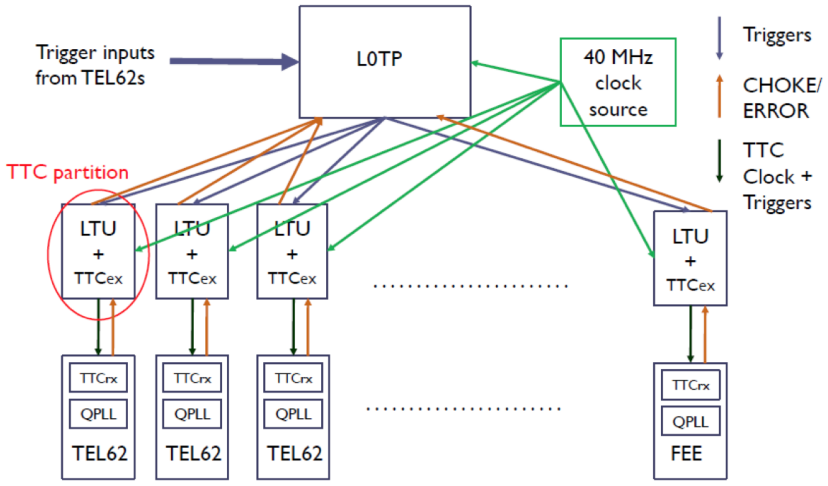


Figure 2.13. Schematic view of the Timing, Trigger Control.

The LTU is an upgraded version of the module developed by the ALICE experiment, a 6U VME module with on-board FPGA. It receives signals of SOB and EOB from the SPS and dispatches them to the L0TP. It also receives the trigger data generated by the L0TP via an 8-line bus and encodes them as optical inputs sent to the detectors. Each sub-system is connected to a LTU module in daisy chain. Finally LTUs are also used to send asynchronous error signals to L0TP and freeze the data-taking when the error signal is on.
On the detector side, all the read-out boards mount a TTCex mezzanine board to convert the optical signals into electric. The board contains an on-board QPLL that can be used as a jitter filter for the TTC clock or any external signal operating at the frequency of 40 MHz.
The fundamental unit of time is the inverse of the clock frequency: 24.950741 ns, namely *timestamp*. The time information coming from any individual channel is given with this unit

of time, available to each sub-system through a TTC receiver. The counting of timestamp starts at each burst with the SOB, thus any event within a burst is uniquely identified by a timestamp. Furthermore, in different detectors, the same event is time separated by a fixed number (± 1 to account for the edge effects) of timestamps. The timestamp counter is 32-bit long for a full time range of 53.6 s. To allow for a better time resolution, all the TDCs split the time unit in 8 bit, corresponding to an overall time unit of 97.466 ps.

2.13 Data acquisition

2.13.1 TEL62 Board

The TEL62 board[25] is a multi-purpose FPGA-based motherboard; it is a upgraded version of the TELL1 board designed by EPFL Lausanne for the LHCb experiment at CERN. NA62 uses the TEL62 both for trigger primitive generation and for data acquisition of most detectors. The TEL62 carrier board hosts the TDC Board (TDCB) as a mezzanine daughter card.

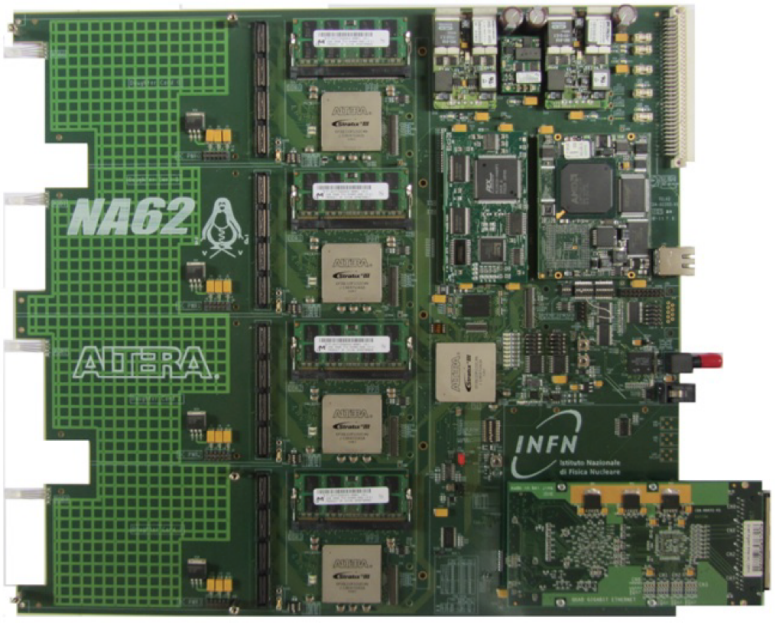


Figure 2.14. TEL62 read-out board.

The design (see figure 2.15) integrates a high number of channels, leaving only the analog front end electronics on the detector side. The implementation of the interfaces has been kept quite general to be easily adapted to different sub-systems. The TDCB mounts four High Performance Time To Digital Converter (HPTDC) developed at CERN [26]. Each HPTDC

provides 32 TDC channels with a LSB of 97.466 ps resolution, so that a TDCB contains 128 TDC channels, measuring both leading and trailing edges generated by a particle crossing a detector. One can also obtain analog pulse-height information by the time-over-threshold method. The chip HPTDC has been designed to digitize the arrival time of both signal edges. Finally HPTDC can pack both digitizations into a single word by stripping out the MSBs of the trailing edge measurement. The TDCB houses a dedicated Altera Cyclone III ®EP3C120

Figure 2.15. TDC Board mezzanine for the TEL62 board.

FPGA, programmed as a TDC Controller, which can handle the configuration of the four HPTDCs and store the data acquired. A 2 MB external static RAM block is also available and used for online data monitoring purposes and low-level checks on data quality.
Moreover, the TEL62 board is equipped with 2 GB Double Data Rate-II (DDR2) memory, an optical TTC link to distribute the main 40 MHz clock and the L0 triggers, an on-board QPLL to reduce the signal jitter of the input clock to a few tens of ps.
TDCs can handle a few tens of MB/s per channel, depending on detector structure. Data are organized in packets, each one related to time frames of defined duration of 6.4 μs. A first set of four FPGAs (PP-FPGAs) receives and merges the data and then stores them on the DDR2, where each memory-page is related to a well defined 25 ns time window. Upon a L0 trigger request, TEL62 moves the data within a programmable number of 25 ns time windows around the trigger timestamp from the PP-FPGAs to the central one (SL-FPGA). On the SL-FPGA data are merged, synchronized and sent out in packets via four Gigabit Ethernet links hosted on a custom daughter-card to the PC-farm for the next levels of triggers. The logic of the board is entirely defined by the configuration of the five FPGAs, which also control the other on-board devices (such as TTCex, RAM and GBE). All the four PP-FPGAs implement the same firmware. Data are first received via four buses linked to the TDCB mezzanine cards. Data packets are then compressed to be stored in the DDR2 memory. The firmware elaborates the trigger request, reading the corresponding data from the DDR2. The SL-FPGA decodes the trigger type and the trigger timestamp. The trigger type can be either physics or special purpose. For physics triggers (32 possible different types), a request is forwarded to the DDR2 in order to read a programmable number of 25 ns time slots around

the trigger timestamp. The trigger requests received from the TTC interface are dispatched to the PP-FPGAs where the corresponding data are read. Data belonging to a trigger can be packed into a single Multi-Event Packet (MEP) to optimize the output link bandwidth. After MEP assembly the packets are temporarily stored in a external 1Mbit QDR memory and finally sent to the PC-farm.

Detectors involved in the L0 trigger developed Primitive-Generator modules to elaborate data for the L0TP. The trigger primitives follow a path similar to the data. The primitives coming from the PPs are merged by a primitive merger block before undergoing further processing stages, like multiplicity counting in the RICH or cluster identification in the LAV. Triggers, like data, are packed in the SL-FPGA in Multi-Trigger Packets (MTPs), formatted as UDP packets and sent via dedicated Ethernet ports of the GbE to the L0 Trigger Processor.

2.13.2 L0 Data Format

The key point of the transmission of the trigger information by NA62 detectors to the L0TP is performed via gigabit ethernet connection. Data are bundled in UDP-packets and sorted in little-endian format. A variable number of L0 primitives are collected in a single MTP for transferring them to the L0 Trigger Processor.

One UDP packet contains a single MTP with a maximum size of 65500 bytes.

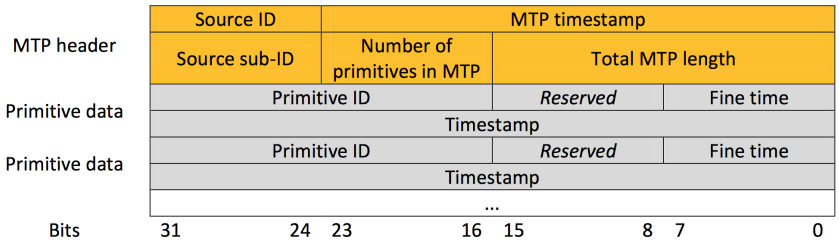


Figure 2.16. The Multi-Trigger-Packet.

Figure 2.16 shows the L0 primitives sent from detectors to the L0TP: First 64 bits are the MTP header, where some general information are located. **Source ID** is a 8-bit identifier for the detector or sub-system which generates data, according with table 2.17. **MTP timestamp** is the detector time in unit of 25 ns, indicating when the packet has been delivered by the TEL62. It can be used for debug purpose, monitoring the delay between the generation of the primitive data and their delivering. **Source Sub ID** is a identifier of the individual DAQ board (or part of it) within a detector or sub-system generating the data. **Number of Primitives in MTP** is the number of primitives encapsulated in the packet, possibly different for different MTPs. Total MTP length is the total length of the MTP in bytes, including the header. After

CEDAR	0x04
GTK	0x08
CHANTI	0x0C
LAV	0x10
STRAW	0x14
CHOD	0x18
RICH	0x1C
IRC	0x20
LKr	0x24
MUV1	0x28
MUV2	0x2C
MUV3	0x30
SAC	0x34
New CHOD	0x38
HASC	0x3C
L0 Trigger Processor	0x40
L1 Trigger Processor node	0x44
L2 Event Building node	0x48
End-of-Burst non-standard block	0x4C
Reserved	0x50-0xFF

Figure 2.17. List of Source IDs for different sub-systems.

the 64-bit header, data are divided in two parts. A first part consists of the **Primitive ID** which identifies the conditions satisfied by the event, the **Reserved** field which shall not be used for any purpose, as it might be used in the future, and the **fine time** which represents a more precise measurement of the event time. The precision could be different for each detectors, and the bits exceeding the precision are set to zero. Finally the field Timestamp contains the number of 40 MHz counts at which the present primitive occurred. Data are transmitted on a private network. Each detector sends a single MTP/MEP on a single Gigabit Ethernet cable to a single IP address with the correct (statically defined) MAC address.

MTP packets are sent periodically, even in case of zero primitives, the default period being 6.4 μs (MTP Period). It means that the primitives should be merged in frames of 6.4 μs. Considering 10 MHz of incoming primitives at maximum rate, a mean of 64 primitives are stored in a MTP. A primitive is 8 byte long, so 512 byte are occupied, plus 8 of the header of the packet, while the maximum payload size is 1472 byte. A frame structure is almost mandatory in order to have a first rough synchronization, constraining the time in which a primitive could arrive and avoiding to wait infinite time before starting the coincidence procedure on the L0TP, as described in the next chapter.

Chapter 3

Level Zero Trigger Processor

3.1 Introduction

The L0 Trigger Processor (L0TP) is the lowest level system of the NA62 trigger chain. It is the only component of the trigger hardware-implemented, using programmable logic (FPGA). The L0TP selects events that are interesting for the analysis and determines the event rate sent to the NA62-Farm hosting the higher trigger levels before the storing data on disk.

The L0TP architecture is completely new for a high-energy physics experiment, based on standard Gigabit Ethernet communication between the detectors and the L0TP Board.

The L0TP is responsible for the Level-0 trigger decision as a result of the inputs of six detectors (CHOD, MUV3, LKr, RICH, LAV, New CHOD), with the possibility to extend this number to seven [13]. It is integrated in the NA62 Run Control System in order to receive the proper run configuration as other detectors.

A block diagram of the L0TP is shown in figure 3.1.

The L0TP has been designed to check the time alignment of the primitives which are asynchronously generated by detectors; when, for a given timestamp, the primitives match with a predetermined configuration (*mask*), the trigger is validated and it is delivered to detectors in phase with the 40 MHz clock after a programmable fixed latency.

There are many other features that characterize the design: it implements individual downscaling for each mask, handles CHOKE and ERROR signals coming from detectors and consequently generates special triggers to stop/re-enable the data acquisition. L0TP can also receive special primitives meant for calibration purpose, or NIM pulses for the LKr calibration. Periodic trigger generators have been implemented for monitor, debug and test purposes, and control triggers are performed to measure the trigger efficiency or study particular trigger conditions without encoding them in a mask. Finally, the L0TP sends to the PC-Farm all the information useful to reproduce offline the trigger conditions.

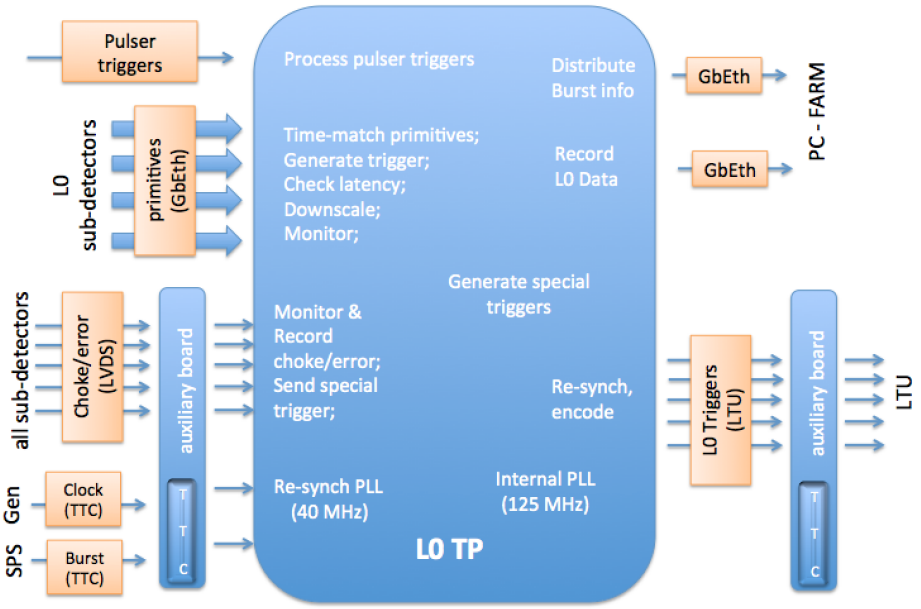


Figure 3.1. Sketch of Level Zero Trigger Processor. Input/Outputs are illustrated with the characteristics of the project.

Commissioning of the L0TP started in 2013. A large sample of data has been collected with a first version of data generators in 2014 (NA62 pilot run), in order to gain an initial understanding of the trigger performance. The first data taking with the actual primitive generators has been performed in 2015. As the NA62 instantaneous beam intensity increased over the course of 2015-2016 runs, L0TP firmware evolved in order to cope with the increasing trigger rates.

This chapter illustrates how the firmware and the other auxiliary components for L0TP have been implemented. In section 3.2 are given the details of the hardware. Some basic principles of the simulation are introduced in section 3.3. In section 3.4 the main blocks and logic elements of the L0TP are described. Finally, in section 3.5 some results obtained with tests in laboratory are presented.
In the rest of this chapter, I will refer to the components that generate data as detectors or sub-systems equally to indicate the logical components which send primitive to the L0TP.

3.2 Hardware

L0TP is based on Terasic ®DE4 development board (shown in figure 3.2), a commercial board hosting an ALTERA STRATIX4 EP4SGX530KH40C2 FPGA. The main characteristics of the FPGA module are:

- maximum number of user I/Os: 736;
- number of registers: 424960;
- fabrication technology: 40nm;
- RAM Bits: 27376K;
- device Logic Cells: 531200.

The board provides also:

- two DDR2 SO-DIMM Sockets with the possibility to have up to 8 GB of external memory clocked at 400 MHz;
- two different oscillators (50MHz and 100MHz);
- SMA connectors for external clock;
- four Serial ATA Ports;
- four gigabit ethernet ports;
- PCI Express x8 Edge Connector;
- two 172-pins High Speed Mezzanine Card (HSMC);
- two 40-pin Expansion Headers;
- USB Host/Slave Controller.

The choice of the DE4 board among several commercial boards has been made for several reasons. The guideline was to minimize custom developments using off-the-shelf electronics. The DE4 features relevant to the project were the possibility to expand from four to eight the number of gigabit ethernet links using the HSMC connectors and the large internal RAM of the FPGA, allowing to work with the only internal static memory without the need of implementing the communication with a external dynamic DDR2. This is a big simplification of the firmware.

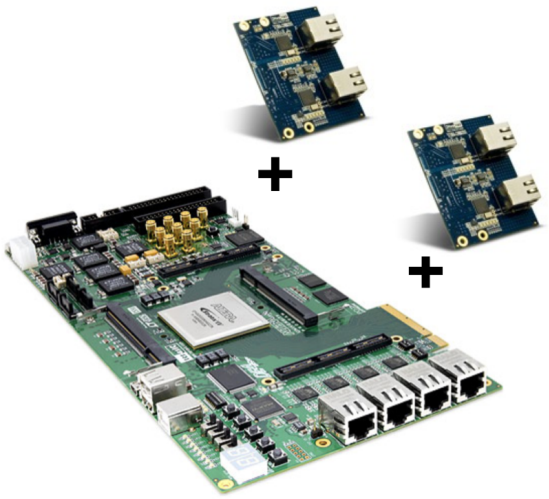


Figure 3.2. Terasic DE4 Board and two HSMC-Gigabit Ethernet extensions

A USB-blaster provides the Joint Test Action Group (JTAG) configuration through on-board USB-to-JTAG configuration logic. Connecting the board with a PC using the USB-blaster gives the possibility to download directly programs developed with Altera Quartus II ®(version 13.1) software.

Connecting the board to a PC via PCI-express also opens the possibility of developing a parallel Level Zero Trigger Processor project based on software algorithms running on a standard PC [27][28]. This project has been already tested both in laboratory and during the 2015-2016 data taking. The NA62 setup allows also to run the two systems in parallel in order to compare the performances. The PC-based project is developed by the INFN-Ferrara group and it is taken in account as an option for future upgrades of the current L0TP.

The 8 gigabit ethernet links allow to have 7 inputs to receive data from detectors while the last is reserved to send information to the PC-Farm.

A L0TP Auxiliary Board (L0TPAB) has been designed to interface the L0TP with the master clock, detectors, the LTU, and the SPS control signals. The L0TPAB hosts a TTCex module which receives the triggers from L0TP and forwards them to the LTU. It also receives the clock signal via optical fiber and translates it to LVDS standard. Similarly SOB and EOB controls from SPS, which are NIM signals, are converted to LVDS. The board also provides LVDS connections for the CHOKE/ERROR signals coming from all the NA62 detectors.

3.3 L0TP Behavioral Simulation

The behavior of the logical blocks described below has been obtained through an extensive simulation, both of the individual sub-section that make up the system, and of the system as a whole, using Altera Modelsim ®Simulator. Simulations are controlled using test benches, which are additional VHDL modules used to generate the appropriate waveforms on the input ports of the module under test, in order to test the functionality of that module. Input primitives are read from files generated by real data taken in 2015. Multiple configurations are available by changing the run control parameter settings. While the ethernet interface has not been simulated, the test bench mimics the receiving algorithm actually implemented in the L0TP firmware.

A Finite State Machine (FSM) waits for a programmable starting time and then creates fake ethernet payloads reading binary files. On the other side, another fake FSM simulates the behavior of ethernet interface, extracting one byte per clock period. Propagation of signals from the USB to the input registers have not been simulated and the correctness was implicitly checked by verifying the input data in the simulation. Only two sources, CHOD and MUV3, are used to produce coincidences between primitives in simulation. The same text file has been used as input both to the simulation and to a firmware-implemented test bench (see section 3.5). In this way an immediate cross check between the Modelsim outputs and test bench results is possible.

3.4 Description of the L0TP Firmware

In this section the main elements and blocks of the firmware are described in some details. Figure 3.3 shows a sketch of the firmware implemented. In summary, the L0TP contains:

- two PLLs to generate the clock domains driving the logic;
- a TTC block, handling asynchronous signals;
- 7 ethernet interfaces, receiving primitives from detectors;
- a Delay Generator, aligning in time different ethernet packets;
- 7 different circular buffers storing the incoming primitves;
- a Memory Management Unit (MMU);
- an Associative Memory Module (AMM);

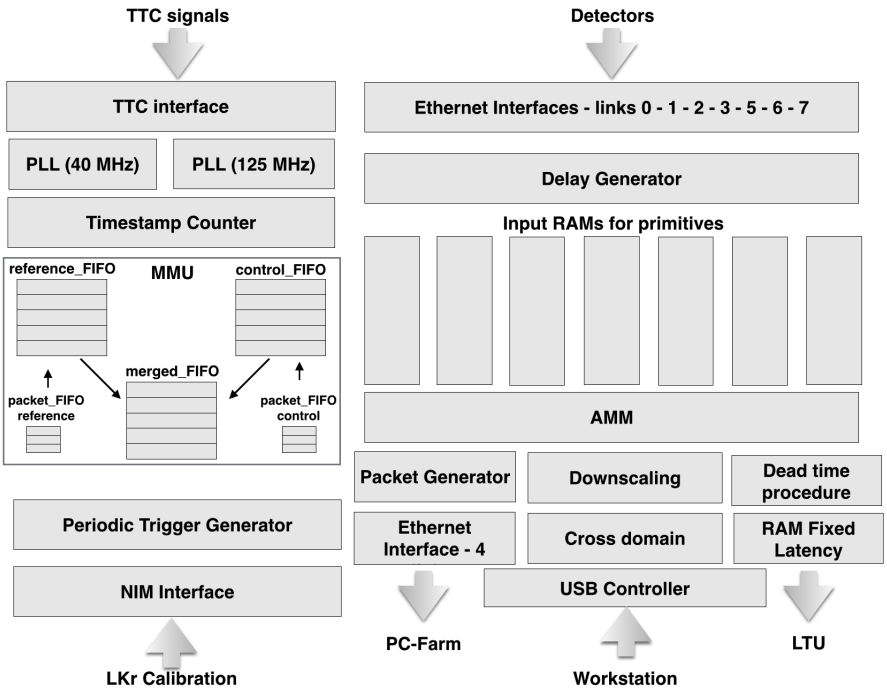


Figure 3.3. Sketch of L0TP firmware

- special Trigger Generators;
- a Fixed Latency Generator sending triggers to detectors;
- a MEP Generator sending packets to the PC-Farm;
- a USB Controller to read/write parameters from/on L0TP registers.

Other components implemented via software, complementing the firmware, are:

- a Round Robin Software to deliver MEPs to the correct IP of the PC-Farm computer;
- a USB Controller Software to communicate with the L0TP;
- a Server connecting L0TP to the NA62-Run Control System.

3.4.1 Clock Domains

Two clock domains drive the logic in the L0TP:

- 40 MHz clock received from the common NA62 time distribution system;
- 125 MHz clock, driving the trigger-algorithm logic and the ethernet communication.

The 40 MHz clock is used to set the time to the internal events in phase with the other parts of NA62. It is received through the L0TPAB. All the output are synchronized with this clock. It is managed by a Phase Locked Loop (PLL) circuitry, a feedback control system that automatically adjusts the phase of a locally generated signal to match the phase of the input. A green led on the board indicates that the PLL is correctly locking to the clock .
The 125 MHz clock comes from a PLL fed with the 50 MHz DE4 internal oscillator input. Common resets for the different clock domains are also provided to clear all the buffers and reset all the finite state machines and counters on the board.

3.4.2 Timing and Trigger Control block

The Timing and Trigger Control (TTC) block monitors the external asynchronous information coming from the SPS and the Choke/Error signals from 14 detectors to generate a group of signals to be used by L0TP to control the basic operations. It manages the SOB, EOB and CHOKE and ERROR signals, receiving the inputs (CHOKE, ERROR, BCRST, ECRST pins) from the L0TPAB via the 40 pin flat cable. The common reset, the start of burst signal and the 40 MHz clock are also sent to the module. Only when L0TP is in RUN mode data are

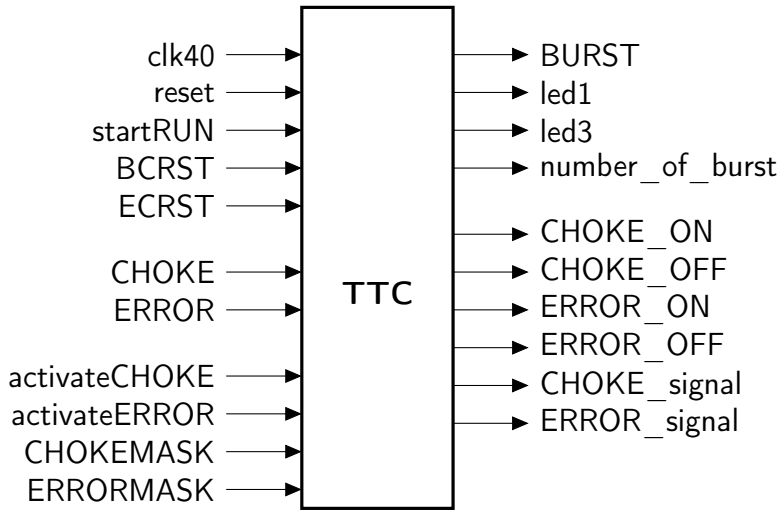


Figure 3.4. Entity of TTC

analyzed. A START_RUN signal comes from the Run Control block. Only when the data acquisition is in RUN mode, the L0TP makes an internal level (BURST) available. When it is high, the entire logic is active and ready to handle primitives or special triggers.
The start of burst is labeled by two signals called Event Counter Reset (ECRST) and Burst Counter Reset (BCRST). They are both set to 1 when the SPS control sends the Warning Of Extraction message. It means that the system is considered in Burst state and the level

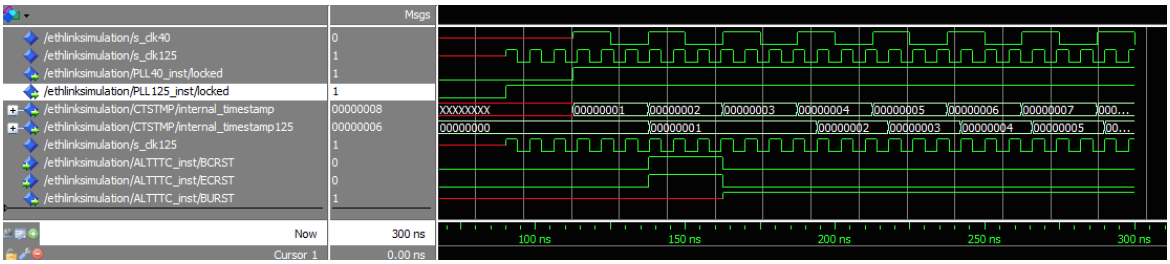


Figure 3.5. Simulation of SOB signal. When ECRST and BCRST levels are high, the BURST level is set to one and L0TP can manage input/output signals.

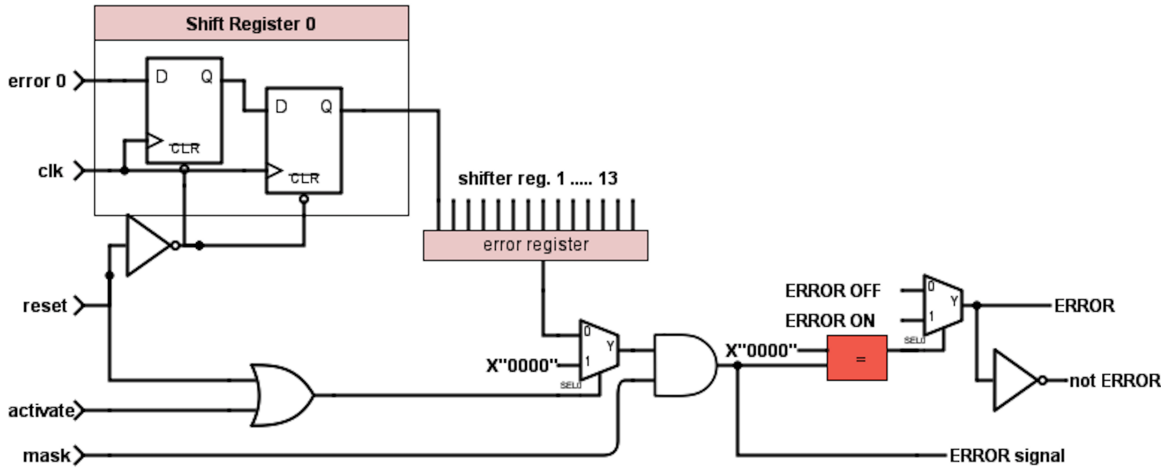


Figure 3.6. Schematic view of the error synchronization logic. A shift register is implemented for each ERROR signal associated to a detector. The status information is stored in a 14-bit register. If the output is activated, then the AND operation with a pre-selected mask is performed in order to accept only the bits that should be monitored.

BURST is set. The signal BURST is then propagated to rest of the logic.
As the burst ends, namely the SPS has stopped to send protons on the target, ECRST is set to one and BCRST is set to zero. In this conditions, the EOB signal is set whilst SOB is cleared and L0TP clear memories and resets counters.

Two LEDs indicate when the board is IN or OUT of burst. A counter counts the number of bursts from the start of run. Because ECRST and BCRST are asynchronous signals, the logic should make them synchronous and in phase with respect to the internal 40 MHz clock. The synchronization is done with a two-stage shift-register clocked by the target domain's clock. Each CHOKE (ERROR) input pin is monitored in the same way and a register stores the information of the actual global status of the CHOKE (ERROR). Figure 3.6 shows the scheme of the synchronization logic for the CHOKE/ERROR.

When the user enables the output, a logical AND operation between the CHOKE (ERROR) register and a pre-selected mask is performed in order to monitor only the selected inputs. In addition to the 14-bit ERROR signal output, a differential pair called CHOKE_ON/CHOKE_OFF (ERROR_ON/ERROR_OFF) is distributed through the firmware in order to implement the CHOKE/ERROR functionality in the trigger dispatching logic.

Another critical feature is the dispatching of the timestamp to the different firmware blocks. It refers to 40 MHz clock domain, but it is also used to mark special events in the logic clocked at 125 MHz. A dual clock FIFO is used to input the timestamp values with a 40 MHz clock. The readout from the same FIFO is done at 125 MHz. As shown in figure 3.5, this operation takes two 40MHz clock periods to be performed, due to the FIFO writing/reading operation.

3.4.3 Communication Between Detectors and L0TP

The starting point for the development of the communication block between detectors and L0TP was the ethernet test-interface implemented by INFN-Ferrara group for the Gigatracker readout. The project was an optimization of the Altera ethernet interface for the IP/UDP datagrams. As described in chapter 2, the incoming primitives are encapsulated in UDP datagrams together with the checksum.
In this description we can assume the Media Access Controller (MAC) as a black box which contains the logic of transmission or reception and interfacing to the ethernet ports. The MAC addresses of the different ethernet links are defined by a register set by hand via an external dip switch.
A variable number of primitives can be merged into a MTP and then sent to the L0 Trigger Processor. MTP are sent in variable length frames with a period of 6.4 μs. The frame is transmitted even if it is empty. The fixed time structure of the transmission simplifies the task of the L0TP, which can then proceed to search for coincidences after a fixed delay, absorbing any possible delay of the inputs.
In fact there are two contributions to delays from the primitive production:

- the intrinsic detector average processing time;
- the random fluctuation of the processing time due to the event complexity.

In the first case, an offset could be present between two different systems because of the different time intervals that firmwares spend to generate primitives. As a result, a primitive relating to a certain time could be stored in the packet N for a faster detector, while it could be in the packet M for a slower one, where $N < M$. This delay is measurable and can be

compensated by the L0TP. In the second case, L0TP should be able to absorb time differences between events coming from the same sub-system. In fact the detector firmware logic will spend larger time in analyzing more complex hit patterns. As a consequence, a primitive belonging to the time covered by the frame X could be written in one of the next packets.

Receiving logic The algorithm to read data from the ethernet buffer (rx fifo) extracts one byte every 8 ns, exploiting the maximum bandwidth of the gigabit ethernet communication. Being the primitive 8-byte long and considering two-clock-period, the overhead time to store the primitive in a memory is 80 ns. This is a limitation for primitive transfer, which cannot be more that 12.5 MHz on average. Ethernet buffers can manage instantaneous intensity peaks from the beam.
Two different modules share the same logic, but the interface to communicate with the ethernet links is different. One works for links directly mounted on the DE4 board with the Serial Gigabit Media-Independent Interface (SGMII) while the other works for the two daughter cards with the Reduced Gigabit Media-Independent Interface (RGMII).
Starting from the header of a MTP, a Finite State Machine extracts all the primitives contained in the packet adding one more flag word ($0xFFFFFFFF$) which indicates the end of the packets. Another flag-bit tags the header of the packet. This additional information limits the packet structure that can be maintained during the other steps of the firmware.
All the 7-ethernet links are processed in parallel.

The extracted data are stored in a Delay Generator component (the main module of this component is an array of FIFOs called *DeayFIFOs*) to generate a programmable delay triggered from the SOB for absorbing any potential fixed offset between detectors.

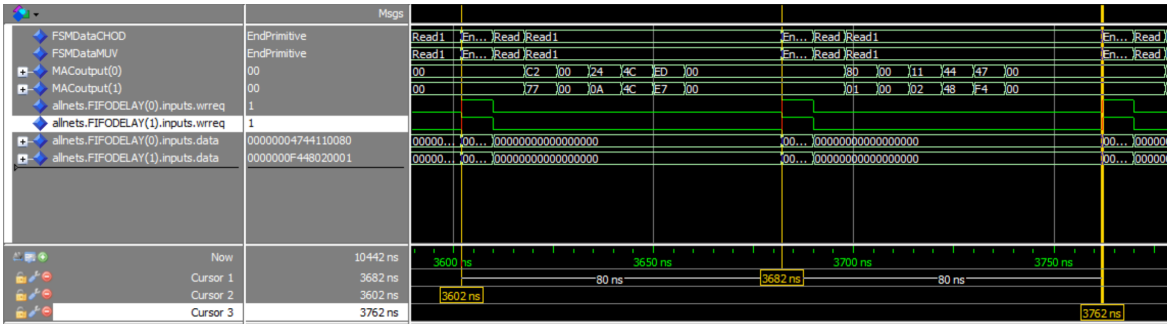


Figure 3.7. Simulation of one primitive extraction. FSMDataCHOD and FSMDataMUV are the Finite State Machine to simulate the ethernet input (from detector to L0TP). MACoutput(i) simulates the output of the ethernet buffer of the link i on the DE4. FIFODelay(i) is the main component of the delay generator, in which primitives are written. As discussed, to extract one single primitive, L0TP spends 80 ns.

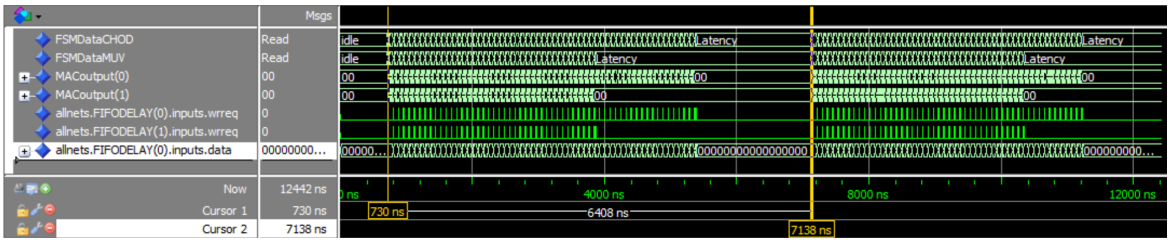


Figure 3.8. Simulation of a MTP extraction. All the primitives are passed to FIFODelay(i), which is the main component of the Delay Generator. The figure shows a packet frame of 6.4 μs.

3.4.4 Delay Generator Block

As mentioned above, primitives related to the same physics event are not delivered at the same time by all the sub-systems involved in the trigger, due to the different processing time of each detector logic. In other words, it is the necessary to realign the primitives according to their intrinsic timestamp rather than the arrival time. The realignment must account for a fixed delay, which basically depends on the detector firmware, but also for a delay component depending on each individual primitive. The latter is related to the pattern of the signals in the primitive generator and consequently to the path followed in the firmware.
The idea for realigning the packets and absorbing the fixed offset consists on saving the primitives coming from faster detectors in dedicated buffers, and skipping a programmable number of frames of the slower ones. This mechanism is performed in the Delay Generator procedure which is sketched in figure 3.9.

Assuming A being the slow primitive-producer detector and B the fast one, by skipping, as an example, the first two empty packets form A one can align packet 2 of A to packet 0 of B, each packet being received every 6.4 μs. This is shown in figure 3.9 on the right side (reading), where one can see the shift of A with respect to B.

In this way, the Delay Generator works as a shift register based on FIFOs. Writing a packet in a buffer and reading it when the next arrives introduces a delay of 6.4 μs.

In general, with this mechanism one can realign the packet N with the packet $N + offset$, where *offset* is a number of packets that can be determined for each sub-system and is fixed for a given firmware version of that sub-system. Being the SOB signal delivered from SPS 0.5 seconds before the beam extraction, the skipped packets are in general empty. All input primitives pass through the DelayFIFOs, and if the delay of all of them is set to zero they are immediately read and passed to the rest of the logic to be aligned. The length of the FIFOs is 8192 64-bit words. If the FIFOs are full, the incoming primitives are skipped and an error message appears on the USB monitor.

Figure 3.10 shows a simulation of the Delay Generator operation. In the simulated example

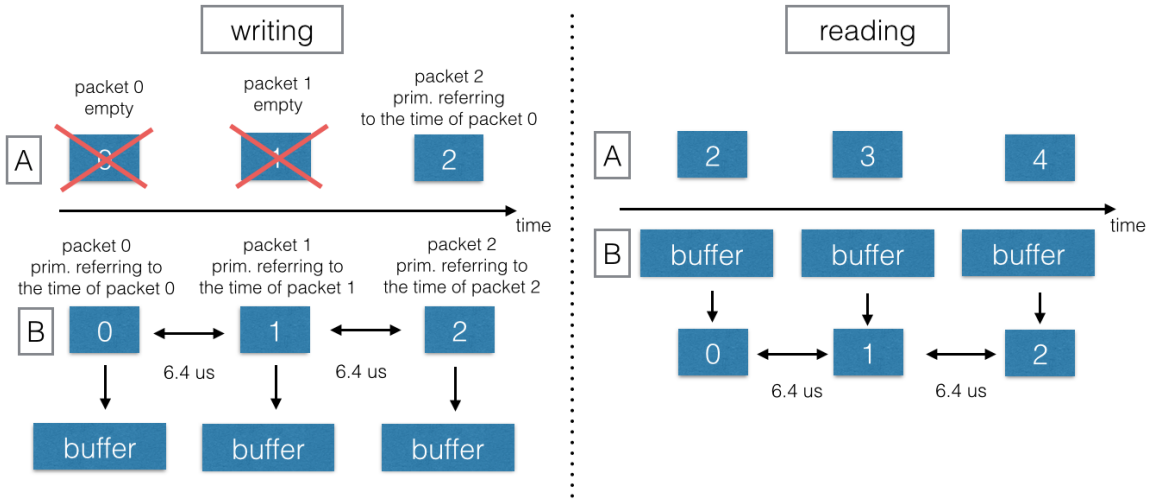


Figure 3.9. Delay Generator module.

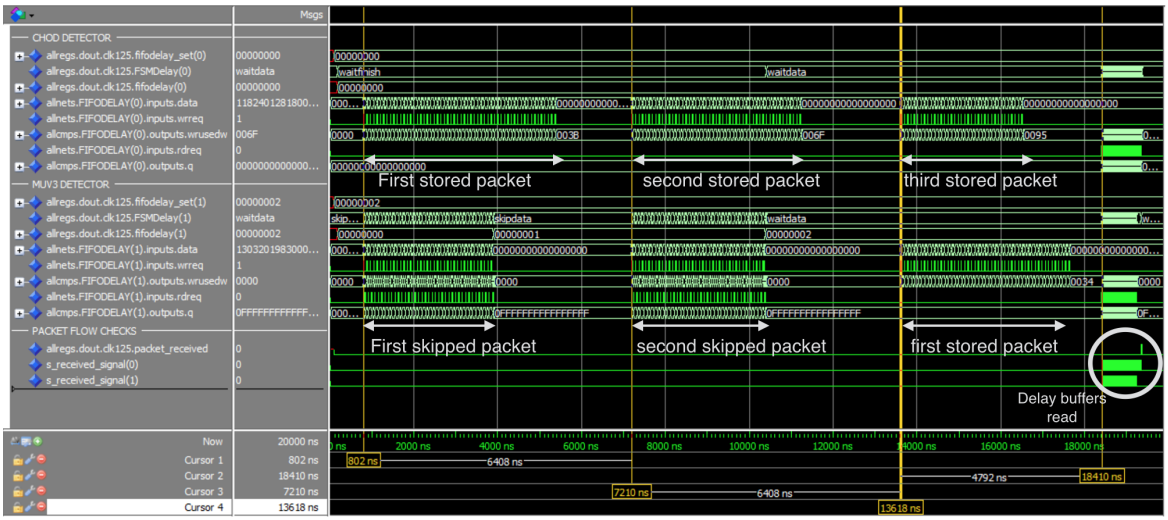


Figure 3.10. Simulation of the Delay Generator logic.

the request was to skip two MUV3 packets (index 1). In this particular case, first packets are not empty, but this is not relevant to the simulation of the logic. The parameter *FifoDelaySet* gives the number of packets to be skipped by each detector. MUV3 primitives belonging to the first two packets are written in the delay buffers, which are immediately read out (for the read request, see line *FIFODELAY(1).inputs.rdreq*). CHOD instead keeps primitives in the buffer, and start the read out of the first packet after MUV3 has skipped the second packet (line *allregs.dout.clk125.fifodelay(1)*). One can observe the FSM which driving the logic (*FSMDelay(i)*) of the Delay Generator: after two skipped packets, it goes in *waitdata* state, meaning it is ready to pass data to the time alignment logic.
If the FSM is waiting for a new packet from detectors which does not arrive before a programmable latency (up 1000 clock cycles, 8 μs), but the delay buffers are not empty, they are read anyway. This allows to read the last packet of the burst, so that it does not remain stuck in the memory.

3.4.5 Time Alignment logic

Writing primitives in memories When primitives are moved from the delay module to the logic for the time alignment, they are stored in RAMs, one for each source, using the timestamp and the fine time information to generate the address. Before writing the primitive data on the RAMs, the system applies the time-consistency check:

$$Internal\ timestamp\ L0TP - Primitive\ timestamp < output\ latency \tag{3.1}$$

If the relationship 3.1 is not satisfied, it means that the time at which the trigger should be transmitted is already larger than the maximum allowed. In this case, data are rejected and an error message appears on the USB monitor. This kind of error can happen due to very complex events requiring exceptionally long time to generate the primitive.
The RAMs for the alignment have 16384 locations of 56 bits. The data addresses are generated using the lowest significant bits of the timestamp and the most significant bits of the fine time.

The number of fine time bits used in the address is an parameter that can be set externally, matching a RAM location to a time interval. The time interval corresponding to a RAM location is called *granularity*. The lowest granularity of the system is 3.125 ns, generated by 11 LSBs of timestamp and 3 bits of fine time. In this case the depth of the buffer is equivalent to 51.2μs. In other words, generating the address with the primitive time immediately aligns them up to the granularity of the RAMs. This is the first rough alignment performed by

L0TP.
Writing data using the time of the primitives is critical, because it generates a fragmentation of the RAMs. In order to read the fragmented memories without loosing time performance, a Memory Management Unit (MMU) has been implemented.
MMU requires a reference detector, which can be selected by the user, to determine which location of the RAM is not empty. A FIFO, named *reference_FIFO*, is filled during the loading of the RAMs with the fine time and the timestamp of this reference detector. Only non-empty locations around the reference detector time will be read during the procedure to determine the finite alignment of the primitives, skipping the holes in the reference memory.

The price to pay introducing a reference detector is a bias in the coincidence process. In fact, if only the addresses with the reference detector data are read, it means that it should be present in each mask of the trigger, and it is not possible to measure its efficiency.
The problem is solved adding a second detector, namely called "control detector" - associated to a *control_FIFO* in the MMU.

Reading primitives from memories - MMU behavior Before reading primitives from the RAMs dedicated to the time alignment, the L0TP waits for a programmable number of packets being stored, avoiding to miss coincidences between data written in different packets due to delayed primitives or edge effect between packets.

When a packet is received, the packet-counter is increased by one. The number of primitives encapsulated in each packet is written in a small FIFO (named *packet_FIFO)*. At the end of the burst, a timeout signal triggers the complete readout of the last packets. When the packet-counter arrives to the set value, the L0TP starts to manage the first packet received using the MMU block. It extracts primitives from the two internal FIFOs, reference_FIFO and control_FIFO, reading always the one with the lower address in output. In this way addresses are sorted in timestamp between the control_FIFO and the reference one.

The addresses are then stored in a third FIFO (named *merged_FIFO*) merging the lists of reference_FIFO and control_FIFO, also adding two extra bits to tag the data source. If the two FIFOs yield the same address, both are read. Obviously the rate between reference detector and control detector can be different, generating a different number of primitives written in the two FIFOs. For this reason the MMU requires a control system to read and check all the primitives up to the end-packet word from both the FIFOs. This is done with a Finite State Machine procedure.
The addresses of the alignment RAM are set using the output of the MMU, which is essentially the output of the merged_FIFO. Depending on how many bits of the fine time are used, the read pointer of the RAM goes to the location hosting the reference detector or the control

detector data, or both. Finally all the RAMs are read out in parallel.
With this procedure is possible to create two different sample of events, one driven by the reference detector and the other by the control one. The two samples are uncorrelated streams of data in the FPGA, but they are correlated from the point of view of physics. Using the control sample it is possible to measure the efficiency of each trigger mask.
During the reading process, L0TP does not clear the alignment RAM (indeed, they are used as circular buffers) and there is no possibility to discriminate between an actual value from an old value. To avoid reprocessing previous primitives, the MSBs of the timestamps are compared. Only if the MSBs coincide then the primitives are kept, otherwise they are discarded. In this way the L0TP prevents to read more than one time the same data.

To avoid edge effects due to the time-granularity of the RAMs, L0TP reads not only the address corresponding to the time of the reference (or control) of all detectors, but also the previous location and the next one.

If a position is not empty and it is not old data, a 7-bit register points to the RAM that has the information compliant for the matching procedure. Each bit of this register corresponds to one source detector. If the bit N is set, it means that source N has valid data, and only valid primitive IDs are transferred to the Associative Memory Module (AMM).

Before being analyzed by the AMM, the primitives are checked against a more selective time coincidence. A tighter timing cut, based on the full fine time is applied:

$$Abs\left(time\ primitive_i - time\ primitive_{reference\ detector}\right) < cut_i \tag{3.2}$$

where i is the index of the input link, namely the i-th detector associated to that link. Only if the time difference between the reference detector and the other source is smaller than a pre-selected cut, the primitive is transferred to the AMM.

Logic for matching primitives Two parts compose the block comparing detector conditions. The first part is a shift register storing the information read out from three RAM locations: the corresponding address, the preceding and the following. At each clock, the system reads one RAM position and moves forward by one slot the data in the shift-register. An OR operation between the Primitive IDs of the three positions is implemented. Then, the result of the OR is passed to the Associative Memory to determine if one of the masks is satisfied.

This module works like a PROM, user programmable with a list of masks to select which data should pass the trigger selection. A mask is the combination of the primitive IDs that should be satisfied. Eight masks are available, each mask being enabled or not by using a dedicated register.

The implementation has been done by making a copy of the incoming configuration of detectors for each mask. Then these copies are compared in the same clock cycle with the corresponding mask.

It is possible to set a mask bit in 3 states: 0, 1 or don't care. Each bit is supposed to have some meaning for physics. To implement the "don't care" logic, a "don't care" register is used. This register has the bits set to 1 if they are supposed to be in don't care. The AMM module performs the OR between the primitive ID and the "don't care" register. The logical request can be explicitly written as

$$(incoming\ configuration_i\ OR\ dontcare) = mask_i \tag{3.3}$$

where i is the number of the mask and the corresponding copy of the incoming configuration. It means that the "don't care" bits of the incoming configuration are always set to 1 after this operation. Setting at the same bits of the masks to 1 , the corresponding primitive ID bits always satisfy the requests.

The output bus of the AMM brings all the information useful to reconstruct offline the condition generating the trigger: fine time and primitive ID of all detectors, timestamp and the information of which trigger has been satisfied. This last value is called Trigger Flag and it is composed by the same number of bits of the maximum number of masks potentially enabled. When the bit N is set to 1 it means that the mask number N has been satisfied. The AMM generates also a Trigger Word. It is a 6+2 bit word, which is sent via LTU to the detectors after a predefined fixed latency. 6 bits depends on the trigger, and they are encoded as in the table 3.1. The last two bits are reserved to encode SOB/EOB signals.

In summary, each mask is used to select a given physics event stream. Triggers belonging to a stream can be then downscaled by a factor N, meaning that only one out of N accepted triggers of that stream is transmitted to the output. This flexibility on the trigger handling is required to allow L0TP to acquire all the event types which would otherwise saturate the data acquisition bandwidth.

When the output of the AMM is different from zero, a downscaling-system counts how many times a specific trigger has been satisfied during the burst. If a user requires downscaling, the module skips that triggers until the downscaling register reach the pre-selected value. Another counter records how many times each trigger satisfies the downscaling, or, equivalently, how many physics or control triggers have been produced.

L0 Trigger Word [MSBs]	Trigger	Sub-Detector Action
0b0xxxxx	Physics trigger	Send data frame
0b100000	Synchronisation	Send special frame
0b100001		Reserved
0b100010	Start of Burst	Enable data-taking, send special frame
0b100011	End of burst	Disable data-taking, send end of burst data
0b100100	Choke on	Send special frame
0b100101	Choke off	Send special frame
0b100110	Error on	Send special frame
0b100111	Error off	Send special frame
0b101000	Monitoring	Send monitoring data frame
0b101001		Reserved
0b10101x		Reserved
0b10110x	Random	Send data frame
0b10111x		Reserved
0b11xxxx	Calibration	Send empty data frame

Table 3.1 Trigger Word encoding for different types of events. Reserved field are not yet used.

3.4.6 Special Triggers

Physics and control triggers are not the only triggers that the system can generate. Other types of events tagged as "special" are used, for example, to calibrate the system or to measure pedestals. A parameter available in the analysis framework allows to distinguish each data type from the others (see figure 3.11)

Periodic Triggers L0TP can generate periodic triggers for test and monitoring purposes. of the detectors. This function is implemented in the Pulser Module. The Pulser Module has

Physics	0x01
Random	0x02
Calibration	0x04
Synchronization	0x10
Monitoring	0x20
Choke ON	0xA1
Choke OFF	0xA0
Error ON	0xC1
Error OFF	0xC0

Figure 3.11. Different data types coming from L0TP.

two output lines and can generate two samples of periodic triggers with two different periods tagged by different Trigger Words. It is also possible to set a precise starting time different from the SOB, and ending time different from the EOB.
During the 2016 data taking the periodic triggers have been used for two purposes:

- monitoring noisy pixels of GigaTracker;
- acquiring pedestals from LKr.

The precision of the pulses is $\pm$ 1 timestamp, due to the fact that they are generated by the 125 MHz clock, and then shared with 40 MHz domain using again a dual port FIFO. Because the ratio of the two frequencies is not an integer number, a non compensated phase arises, creating this small fluctuation. The fact that they are generated in 125 MHz domain excludes timing violations in the FPGA. Being precise enough for the experiment purpose, they are left non optimized. The outputs of the Pulser Module are finally transmitted to the ethernet module, and then follow the pattern as any other trigger.

Choke and Error The L0TP has been designed to handle two kinds of anomalous activities:

- the CHOKE, which is a level raised by a detector that cannot respond to all the triggers sent by L0TP;
- the ERROR, which is a level raised by a detector when the data acquisition failed due to a detector dependent problem.

These signals come through the TTC interface, as described above (section 3.4.2). The behavior of L0TP in these cases is the same: it stops to deliver triggers, recording the time slot in which the choke (error) level has been raised. A special trigger is generated with a special TriggerWord and the internal timestamp, which is latched to the event. A special TriggerFlag shows which detector was in a problematic state. Choke and error triggers are delivered without latency, in order to act immediately after the recognition of the problem. Two triggers are delivered: one when L0TP stops delivering triggers, the other when it resumes sending them normally. The difference of the timestamps of this pair of triggers provides the dead time of the data acquisition.

Calibration with Primitives Several detectors need special runs for calibration purpose, as LKr to determine the baseline, the Straws to determine the time-offsets.

L0TP can handle these special triggers. When a certain detector sends primitives flagged with the most significant bit (bit 15) equal to 1, L0TP accepts them independently on the

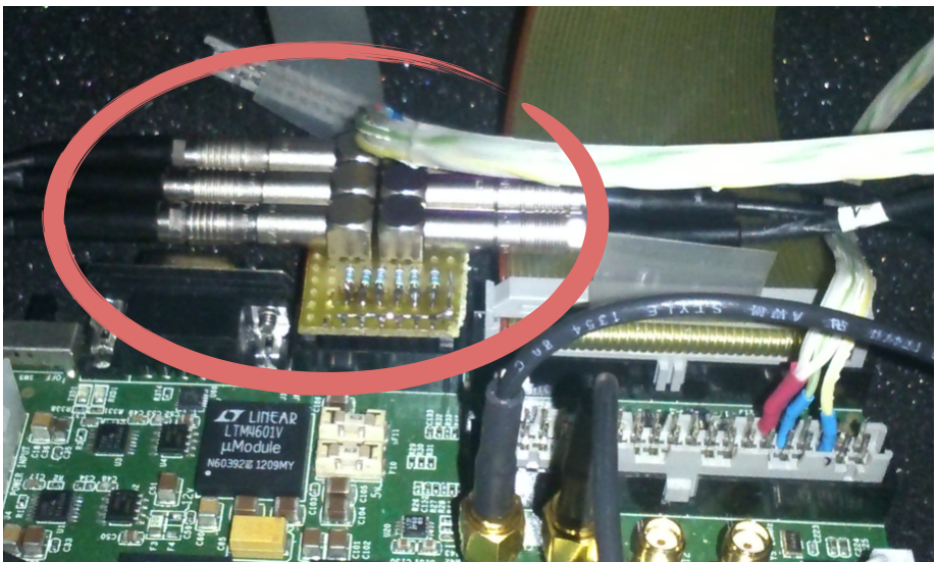


Figure 3.12. LEMO connectors added to 14-pin socket of DE4 board.

configuration of the other systems. For these events, the primitive information is written in a FIFO (named *FIFOCALIB*) as soon as they are extracted from the delay generator. Then, a FSM checks if the FIFO is empty, and if it is not, a trigger is generated and sent directly to the output stage.

Calibration with NIM signals At an early stage of the experiment, namely in the first NA62 run (September-December 2014 [29]), it has been quite convenient to build part of the trigger with separate NIM modules. Then L0TP was adapted to handle NIM signals as inputs, with an hardware patch shown in figure 3.12. This feature has been maintained mainly to deal with LKr calibration runs. When the L0TP receives a NIM signal indicating a calibration event in the Calorimeter, it creates an output string latching the internal timestamp to the trigger and generating the relative Trigger Word. This event is then propagated to the output stage and written in the memories in order to be dispatched to detectors and PC-farm after the fixed latency.

3.4.7 Triggers to Detectors

Whenever a trigger satisfies one of the masks, or a primitive is labeled as special and for this reason bypasses the AMM checks, a signal has to be sent to all detectors via optical fiber in phase with the external 40 MHz clock. The pattern of the signal is the following: an output cable goes from the DE4 to the auxiliary board, and then it is connected in daisy chain with

all the LTU serving all the detectors. Finally, the LTU converts the electrical signals into optical.

The L0TP delivers the triggers after a programmable fixed latency of the order of some hundred microseconds. To generate the latency, a circular buffer (detector-buffer) is implemented, 820 μs deep (32768 memory slots). It is generated using a dual clock RAM. Triggers are stored in this buffer at an address obtained from the most significant 16 bits of the timestamp. The circular buffer contains:

- the TriggerWord of the event;
- the address of the RAM containing the PC-farm information (described in more details in section 3.4.8).

The mechanism to implement a fixed latency between the event timestamp and the the time at which the trigger is issued to the detectors is obtained in the following way: a 40 MHz counter starts at the arrival time of the SOB signal, while the read address pointer of the circular buffer stays set to zero. When the counter reaches the latency value LV, the read address pointer is incremented by one at each clock period. This very simple trick allows to read the position X at the time X + LV.

However some L0TP output signals, as triggers of choke and error or SOB and EOB are distributed to the detectors without any latency. For this kind of triggers the circular buffer is bypassed and the signal is immediately sent to the detectors.

More than one trigger can occur together, in the same timestamp. To deal with this superposition, a FSM checks for each trigger whether the location in the circular buffer is already occupied by some other trigger, and when it is, a logic OR is applied to the TriggerWord.
A minimum dead time between two consecutive triggers is required due to the transmission logic of the TTC. For this reason the L0TP memorizes the last trigger delivered and checks the inter-time between two consecutive output-signals.

3.4.8 Trigger To PC-Farm

The NA62 PC-Farm is a network of 30 PCs whose purpose is reconstructing full events using the information coming from each detector (event fragment). It expects to receive event fragments from each sub-detector for every L0 trigger. From the PC-Farm point of view, L0TP is like any other sub-detector, sending packets of data with the complete information used for generating the trigger. For each trigger, a L0 packet is generated according to the format shown in figure 3.13. These packets are stored sequentially in a RAM composed

bits	31-24/63-56	23-16/55-48	15-8/47-40	7-0/39-32
MEP Header	Source ID	First event n		
	Source subID	# ev in mep	Total mep length	
Event Header	Event Flags	First ev n LSB	First ev length	
	First timestamp			
Data_0	Prim ID A N		Data type	Fine time
	Prim ID C N		Prim ID B N	
Data_1	Prim ID E N		Prim ID D N	
	Prim ID G N		Prim ID F N	
Data 2	Old timestamp			
	TriggerFlag		Old trigger word	Triggerword
Data 3	RESERVED			
Data 4	FT D N	FT C N	FT B N	FT A N
	reserved	FT G N	FT F N	FT E N
Data 5	PRIM ID A N -1		reserved	
	PRIM ID C N-1		PRIM ID B N -1	
Data 6	PRIM ID E N-1		PRIM ID D N -1	
	PRIM ID G N-1		PRIM ID F N-1	
Data 7	FT D N-1	FT C N-1	FT B N-1	FT A N-1
	reserved	FT G N-1	FT F N-1	FT E N-1
Data 8	PRIM ID A N + 1		reserved	
	PRIM ID C N + 1		PRIM ID B N + 1	
Data 9	PRIM ID E N + 1		PRIM ID D N + 1	
	PRIM ID G N + 1		PRIM ID F N + 1	
Data 10	FT D N+1	FT C N+1	FT B N+1	FT A N+1
	reserved	FT G N+1	FT F N+1	FT E N+1

Figure 3.13. Transmission L0TP to PC-Farm: data packet format

of 1024 locations. The memory address is increased by one for each new trigger. The synchronization of the trigger signal sent to the detectors and of the data transmitted to the PC-Farm is done by construction reading the address from the detector-buffer. There is an intrinsic problem in the transmission: the RAM is addressed in the 40 MHz domain, while the transmission to the PC-Farm is performed with a 125 MHz clock. Furthermore the RAM address for reading is propagated with the 40 MHz, but then it has to be applied in the 125 MHz domain. To do the time domain crossing, a dual clock FIFO is used. To take advantage of the UDP format, more events are packed in Multi-Event-Packets (MEPs). The number of events (fragments) merged into a single MEP and sent to the PC-Farm is programmable. This number ("MEP packing factor") can vary between 1 and the maximum permitted by the bandwidth (in general it is set to 8). Also for the data packet, a FSM checks if another event occurred in the same timestamp and, in the case of positive result, the logic OR of the TriggerWord, TriggerFlag and DataType is performed.

EOB signals When the EOB signal is received from SPS, the L0 stops managing inputs from detectors. Without any delay, L0TP sends the EOB trigger and the EOB packet to the farm, with the format shown in 3.14. Memories are reinitialized at the default values and all buffers cleared. A summary of the burst activity is sent to the USB buffer for monitoring and debugging.

3.4.9 Round Robin Software

Each computer of the PC-Farm waits for a programmable number of events from each detector. If this number is equal to the MEP factor, only one packet containing this pre-selected number should be sent to a PC. Then the sub-system have to change the destination IP in order to change the PC that receives the ethernet packet.

In the L0 Trigger Processor the mechanism to change the IP is obtained using an auxiliary external-PC. A loop procedure written in $C++$ language sends primitives to the farm continuously.

Two ethernet UDP sockets are opened, one for receiving from the DE4, the other to send the received packets to the right destination. The program needs a list of IPs belonging to the active PC-Farm computers to start the round robin. To introduce the necessary flexibility, the Run Control transmits them as argument through DIM interface. To ensure the selection of the correct IP address and the data transmission integrity a consistency check of the event number has been implemented. The algorithm to extract the correct index from the IP table

Latency		Data type	*Reserved*
Reserved	Sec ref. detector	Pri ref. detector	Fine time bits
Reserved			
Reserved			
Previous timestamp			
Trigger flags		Previous trig type	Trigger type
Number of primitives from sub-detector A			
Number of primitives from sub-detector B			
Number of primitives from sub-detector C			
Number of primitives from sub-detector D			
Number of primitives from sub-detector E			
Number of primitives from sub-detector F			
Number of primitives from sub-detector G			
Number of CHOKE ON/OFF triggers			
Number of ERROR ON/OFF triggers			
Number of random triggers			
Number of calibration triggers			
Reserved			
Reserved			0x00
Downscaling factor for trigger mask 0			
Number of mask 0 triggers sent			
Number of mask 0 triggers generated			
Don't care primitive bit mask det. A		Required primitive bit mask det. A	
Don't care primitive bit mask det. B		Required primitive bit mask det. B	
Don't care primitive bit mask det. C		Required primitive bit mask det. C	
Don't care primitive bit mask det. D		Required primitive bit mask det. D	
Don't care primitive bit mask det. E		Required primitive bit mask det. E	
Don't care primitive bit mask det. F		Required primitive bit mask det. F	
Don't care primitive bit mask det. G		Required primitive bit mask det. G	
Reserved			

Figure 3.14. End of Burst packet sent by L0TP to the PC-Farm

is the following:

$$IP = (FirstEventNumber/MEPfactor)\ \%number\ of\ PCs \tag{3.4}$$

where *First Event Number* is extracted from the MEP header. It is also easy to forecast the first event number expected at the next MEP, which should be a multiple of the MEP factor.

$$First\ Event\ Number\ Expected = First\ Event\ Number + MEP\ factor. \tag{3.5}$$

If the First Event Number is greater than the one expected due to lost packets, the program jumps to the right IP, reporting an error on the L0TP server, also calculating how many IPs it has to skip:

$$MEPjumped = First\ Event\ Number - First\ Event\ Number\ Expected; \tag{3.6}$$

$$\begin{aligned} &Number\ of\ IPs\ To\ Skip = \\ &(First\ Event\ Number - First\ Event\ Number\ Expected)/MEPfactor; \end{aligned} \tag{3.7}$$

3.4.10 USB controller - Firmware

The L0TP input interface is done via the USB On-The-Go (OTG) module connected via a micro-AB receptacle. During the data taking, this module connects the DE4 to the Run Control System, sending the global setting of the run, while during maintenance or development, the board can be connected to any other linux running PC.
The main package used to activate the communication has been developed at the Edwards S. Rogers Sr. Department of Electrical and Computer Engineering at the University of Toronto[30]. This package allows a program running on a Linux or UNIX workstation to communicate with a user circuit in an FPGA development board. With an automatic procedure, a component connected with two network TCP stream sockets is generated, accepting data to be transferred to or from the L0 board (see section 3.4.11). The L0TP has a USB port multiplexer (*usb_portmux*) connected to the signals coming from the USB. A USB Interface Module, connected to the *usb_portmux*, manages these signals. It uses two in and two out buses to communicate with the outside world. Their functions are:

- move the writing address to the requested register;
- write the correct number in the register;

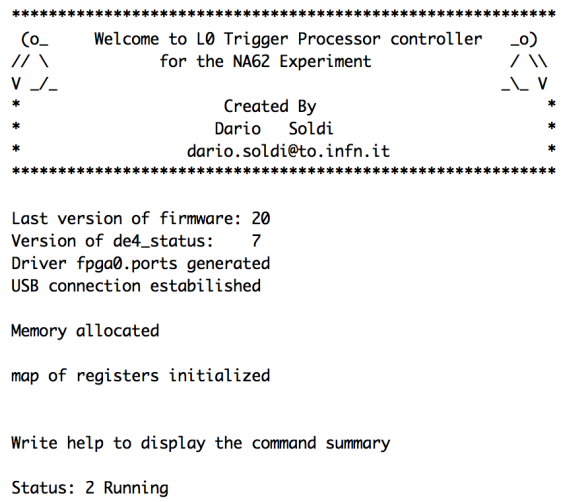


Figure 3.15. Stand alone program to communicate with DE4 board.

- move the reading address to the requested register;
- read the corresponding value from the register.

L0TP handles the following possible states, matching the Run Control states:

- **Idle**: this is the initial state after start or reset the FPGA;
- **Initialized**: the device has been configured and it is ready to start data taking;
- **Ready**: the device is in data taking state and waits that triggers are enabled;
- **Running**: the device is dispatching triggers.

When the board is in idle state, all registers are set to 0. When it receives the initialize command, it enables the procedure to set the register values via the USB 32-bit input bus. If the set-address command is passed through, the Finite State Machine moves into a state where the address value of the register is expected. Then it moves the pointer to the correct memory-location, and then allows to set the value. The initialization command brings back the state to the initialized one. Through another bus it is possible to know which is the actual state of the FSM. The entire procedure works with the 125 MHz internal clock. The same module permits to monitor the board activities, as the number of primitives received from each sub detector, the number of trigger produced, the number of trigger sent to the detector and to the PC-Farm etc. The FSM driving this procedure is very similar to the previously described one. It is possible to read a debug information written in a dedicated memory just by sending the address of the requested-output at the end of burst. When the board is in ready state, it is possible to enable/disable triggers, moving it to the running mode.

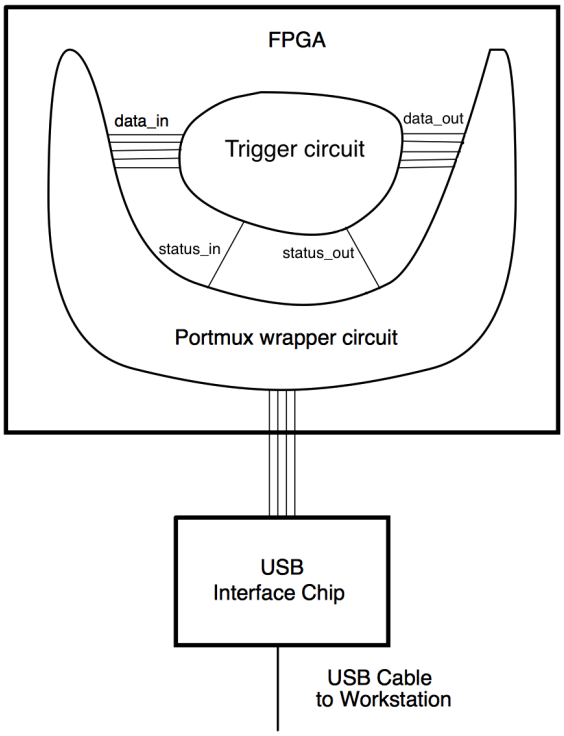


Figure 3.16. Sketch of USB controller

3.4.11 USB controller - Software

In this sub-section it is described the Linux package to communicate with the L0TP board through a daemon process, *tmumon*, acting as front-end to the DE4 board communication. Any Linux program, running *tmumon*, can communicate over the network with the DE4. Tmumon opens two network TCP stream sockets, and waits for connections to those sockets. The first socket (the housekeeping) accepts requests to change or report the state of the DE4 board. The second socket (the ports) accepts data to be transferred to or from a user's design inside the DE4 board. Data are sent to this ports socket as packets, each packet consisting of a 32 bit packet header, followed by optional data bytes. Each header specifies a chip and port number where the data are to be transferred. Among some other features, tmumon also checks that the number of bytes in reception or transmission matches the one reported in the header. A flag word (-1) is transmitted if an error occurred. The transmission unit is 32-bit word. The routines look for a port description file in the current directory, and use it to translate port names into specific port numbers.
Other functions used to establish a correct connection are:

- tm_init(char *hostname): initialize the package, and talk to the board that is physically attached to the given hostname. If hostname is an empty string, a default host will be contacted;
- int tm_open(char *portname, char *mode): open a particular port with a mode of "r" for reading (data flows from the circuit to the workstation) or "w" for writing (data flows from the workstation to the circuit). Returns an integer port descriptor to be passed to the other routines, or -1 on error;
- tm_close(int p): closes the given port.

A stand alone program to communicate with the board and the server for the run control system (l0Server -see section 3.4.12) of the experiment has been developed within the NA62 framework to set the L0TP parameters. This program maps the address of the parameter memory on the L0TP board to a user-friendly name. Taking the parameter values from a text files it uploads them on the firmware. It can read back the values and generates an .xml file used by the l0Server. Parameters to be read/written can be addressed indifferently by number or name.

3.4.12 l0Server

During the data taking, L0TP is managed by the Run Control System exactly as other detectors. This function is performed by a package known as *l0Server*. The link between l0Server and the central Run Control Server is established by the distributed information management system (DIM)[31], transmitting the commands from the RunControl to the device. In this section, the features of l0Server are briefly described. Needless to say this set of features can be improved as required by the experiment. To simplify the operations, the RunControl states are mapped in the DE4. Moreover, the Run Control allows to load the firmware on the L0TP and then it passes to the board *xml* files to setup all the programmable registers. As the RunControl is not informed of the internal operations of any given device, the commands are generic: the device is expected to execute the appropriate sequence of actions specific to itself and implemented in its own server. After the execution of the required action, the device is expected to notify to RunControl the success or failure of the action by updating/refreshing its state parameters. Regarding L0TP, the server, in addition to set parameters and move from one state to another, extracts, with the RoundRobin Software, the list of active PCs among the PC-Farm full list, which is then shared with the "auxiliary PC". From the list of the active PCs, the destination of any given ethernet packet is established by means of a round-robin algorithm.

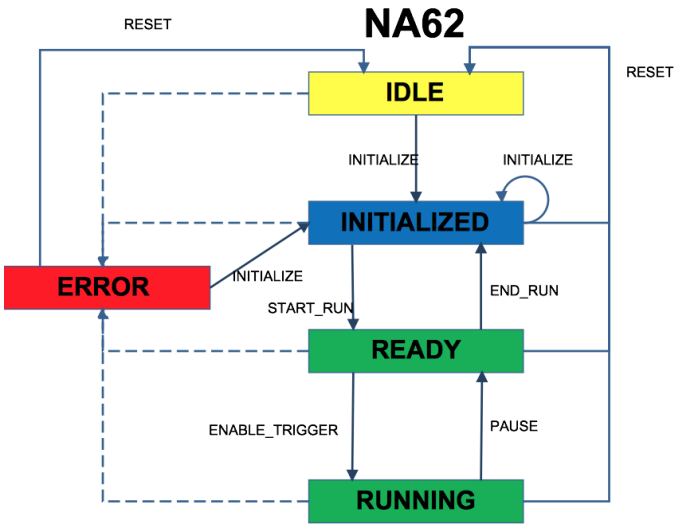


Figure 3.17. FSM diagram representing the global state of the L0TP in the Run Control System

Finally l0Server publishes on DIM all the debug information which can be monitored burst by burst.

3.4.13 Primitive parallel acquisition system

The last component of the L0TP system is the primitive parallel acquisition system. This system is formed by a series of ethernet switches (D-Link Gigabit Ethernet EasySmart ®Switch DGS-1100) placed upstream of the L0TP, duplicating trigger-primitive streams going from the detectors to the L0TP. Linux-daemons run on two PCs and collect all the trigger-primitive data, storing them on the NA62 data merger. The primitive parallel acquisition system is used for monitoring the incoming primitives and check the beam quality. It allows also for the time alignment of detectors before data are sent to L0TP. A sketch of the connection between the detectors and the L0TP, with the primitive parallel acquisition system included, is shown in figure 3.18. The system allows also to connect the L0 PC-based system to compare the performances of the current L0TP with a future new project.

3.5 Verification and Testing

The functional verification of the design is one of the most important tasks in the life cycle of a digital project, because, when the design is built and programmed, it is almost impossible to peer inside it to see if it behaves as intended. For this reason, before commissioning at the

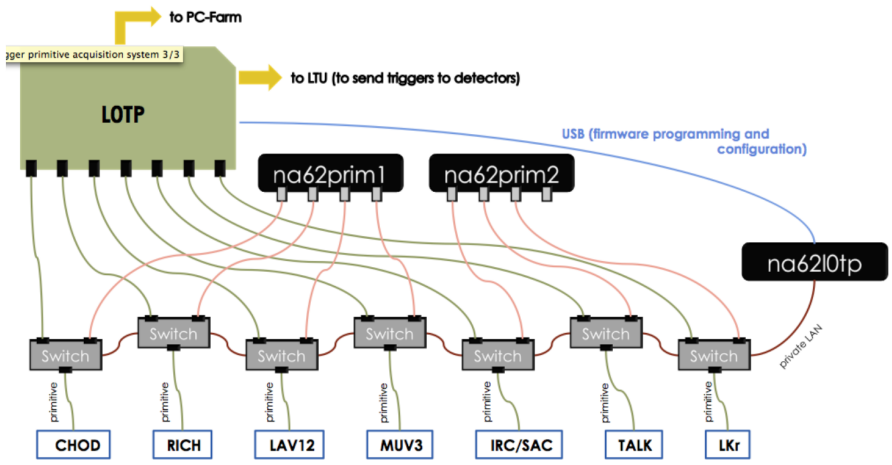


Figure 3.18. Connection between detectors and L0TP. The switches used for the primitive parallel acquisition system are connected with the PC where the l0Server runs (na62l0tp). The primitives coming from detector are dispatched both to the L0TP board and to the PCs for the parallel acquisition (na62prm1 and na62prim2)

experiment, the system as described in the previous sections has been extensively tested in laboratory, performing a first debug of the project before the start of the data taking.

The test bench was realized by connecting two DE4 boards, the DE4 hosting the L0TP firmware being connected to an auxiliary board which emulates the experimental environment surrounding the L0TP. The auxiliary board is identical to the one where the L0TP runs, but it mounts a different FPGA (STRATIX IV ®EP4SGX230KF40C2). Figure 3.19 shows the two boards connected together.

In general, the auxiliary board mimics the behavior of:

- the 40 MHz external clock generator;
- the LTU board;
- Choke/Error signals;
- the detector primitive generators.

Moreover the auxiliary board receives the triggers back from the L0TP.

To check the correctness of the triggers received, they are time-stamped with the 40 MHz clock and results are delivered to a workstation where ethernet packets are monitored. Figure 3.20 illustrates the different components of the auxiliary board logic.

This section describes the additional firmware, developed to reproduce the behavior of the environment surrounding the L0TP. Some results are reported in subsection 3.5.2.

Figure 3.19. Test bench environment. Two DE4 boards connected together. One with the L0TP firmware, while the other runs the firmware to emulate the environment surrounding the L0TP.

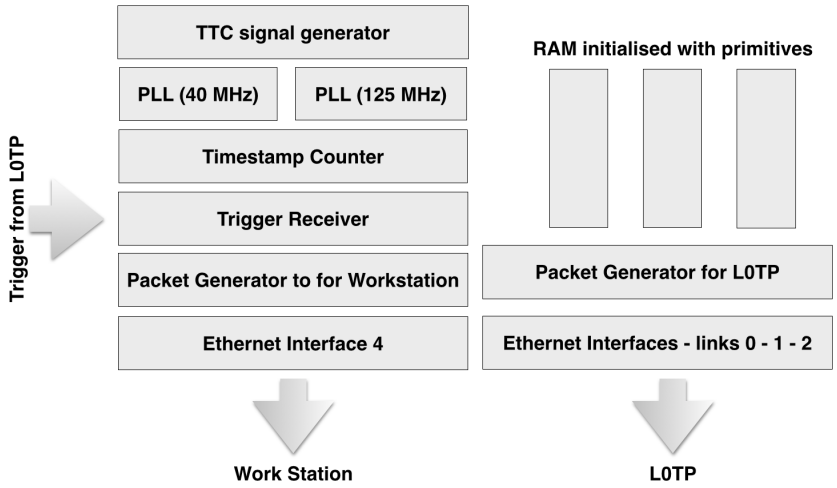


Figure 3.20. Block diagram of the auxiliary board logic components.

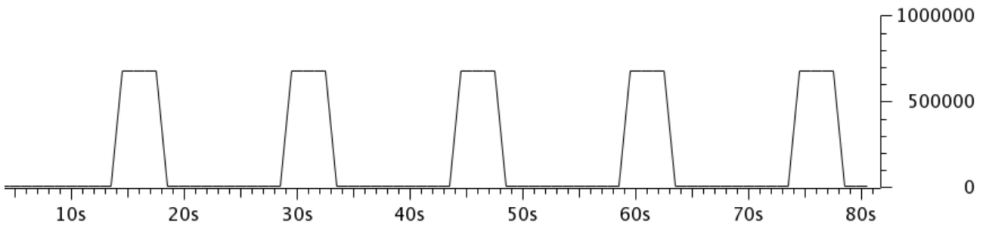


Figure 3.21. Number of bytes sent to L0TP as a function of time. It is evident the simulated burst structure.

3.5.1 Test Bench Firmware

Three different parts compose the firmware of the test bench:

- the clock generation and the implementation of the burst structure;
- the logic to send primitives to L0TP simulating the behavior of detectors;
- the acquisition of triggers by the L0TP.

Clock generation In order to reproduce the behavior of the experimental environment, the 40 MHz external clock and the burst structure are replicated. The 40 MHz clock is generated by a PLL fed with the DE4 internal oscillator at 50 MHz. The generated clock is sent to the L0TP board with two SMA cables as LVDS signals. An internal counter is used to increment the number of timestamp of the burst.
A dedicated process generates the burst sequence, composed by 5 seconds in which the burst is active, separated by 10 seconds of inter-burst. Figure 3.21 shows the number of bytes received by L0TP as a function of time, where the burst structure is observed.

When the firmware is in *burst state*, the logic to send the MTP via ethernet is enabled. The MTPs contain the primitives needed to test the L0TP. The ethernet data transmission is driven by a 125 MHz clock, generated by a second PLL.

Primitive logic The ethernet module for sending the MTPs to the L0TP is identical to the one implemented in the L0TP firmware. In the test bench, for simplicity, only the SGMII interfaces are used, so that only four ethernet ports out of eight are available.
In particular:

- three ports are used to send MTPs to the L0TP;

- one port sends packets containing the triggers received from L0TP to a workstation.

Primitives to fill MTPs are stored in RAMs, one for each ethernet port, and initialized at the compilation time. Each RAM stores up to 32768 primitives. Primitives used for tests can be software-random generated or real primitives collected during the NA62 data taking.
A FSM reads the primitives from the RAMs and copies the ones belonging to a frame of 6.4 μs into a FIFO behaving as a buffer for the generation of the final MTP. Frames of 6.4 μs reproduce the MTP period characterizing the data transmission from detectors to the L0TP. When the FSM has copied all the primitives of a frame into the FIFO, they are ready to be delivered to the L0TP. Another FSM completes the MTP adding the header of the packet and fills the ethernet output buffer reading one primitive form the FIFO every two clock periods. At an average rate of 10 MHz, there are 64 primitives per packet, so that the test bench spends $\sim 1.3\mu$s to generate the packet associated to the frame. Then it waits for the end of the frame time. After 6.4 μs the transmission is enabled and data are sent to L0TP.
Because of the limited size of the RAMs which contain the primitives, a loop mechanism has been implemented: when the RAM pointer reaches the last location it is reset adding the correspondent ram-cycle-offset to the timestamp of the primitive. Also the fine time can be modified to generate random coincidences between the different sources.
In order to simulate the complete experimental environment, a choke/error generator has been implemented. Choke/error signals are sent as NIM levels via a flat cable form the auxiliary board to the L0TP, allowing to test the response of the L0TP to asynchronous signals arriving during the burst.

Acquisition of the triggers from L0TP The third part of the firmware is dedicated to receive from L0TP the words summarizing the response related to each accepted trigger. From this information it is possible to check the arrival time and the TriggerWord reconstructed and added by L0TP, and finally the consistency between data sent to PC-Farm and detectors.
A process sampling the trigger signals is enabled at the start of burst: as soon as a trigger is received, a dual port FIFO is written with a string containing all the information, namely the timestamp, the TriggerWord and the number of the triggers received. The dual port FIFO is used as a temporary buffer to change the clock domain from 40 MHz to 125 MHz. Results are sent via ethernet to a dedicated workstation where they are stored on disk for later analysis.
In summary, two ethernet socket programs run on the workstation and record the data coming from the L0TP and the auxiliary board for further studies.

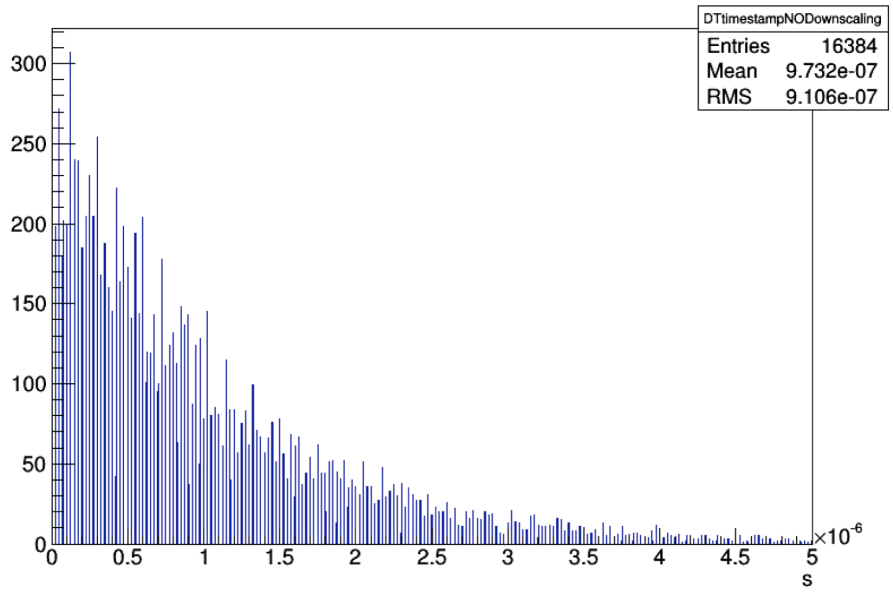


Figure 3.22. Primitive generated with an exponential-random distance between two consecutive timestamps.

3.5.2 Results from Test Bench

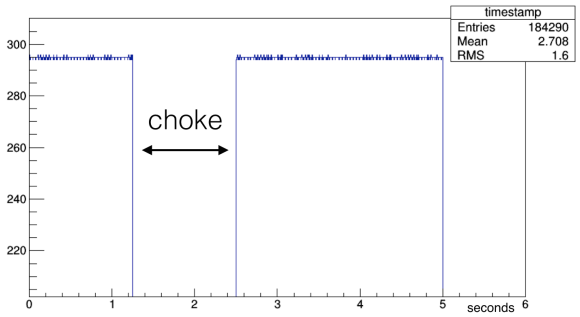
The tests executed in laboratory included the analysis of the following implementations:

- correctness in finding coincidences between two sources;
- correctness of data transmission to both PC-Farm and Detectors;
- correctness of the round robin mechanism.

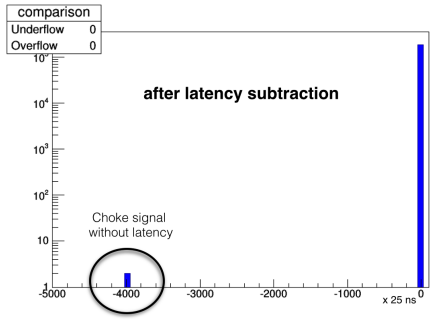
First tests with simulated primitives have been performed in 2014, before the first NA62 data taking. Given the value of the (N-1)th timestamp, the next timestamp, Nth, has been generated with an exponential distribution as in equation 3.8.

$$Timestamp_N = Timestamp_{N-1} + (-frequency \times log(1-(finetime)); \qquad (3.8)$$

where *frequency* is the average frequency for the primitive generation and the finetime has been generated with a random flat distribution. The result of the distribution is shown in figure 3.22. With this sample, the L0TP has been tested up 10 MHz of primitive input rate. The behavior of the system to handle the choke/error signals is shown in figure 3.23a. Performing this test, it was verified that as a choke (or error) signal was received, L0TP



(a) Choke (error) behavior during burst.



(b) Difference between timestamps of data sent to PC-Farm and detectors after latency subtraction. Choke (error) signals are delivered without any latency.

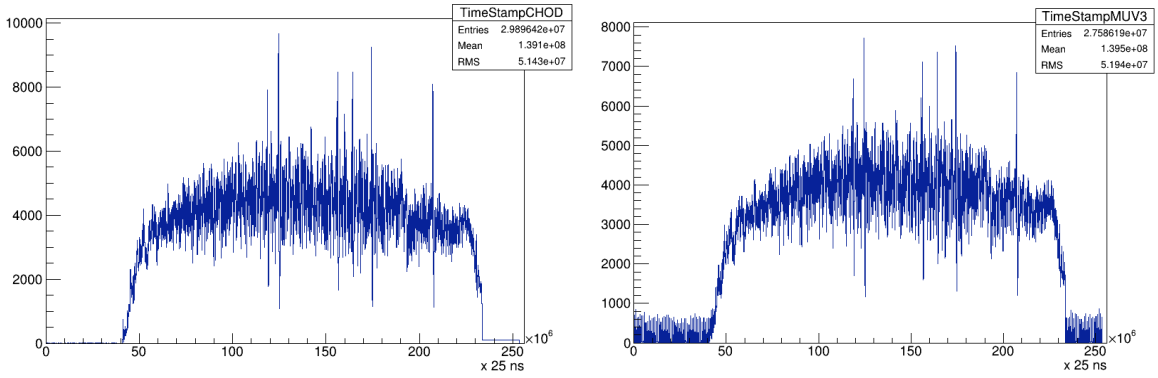
Figure 3.23. L0TP behavior after a choke (error) signal.

dispatched a special trigger to indicate the timestamp in which the choke occurred. Then the logic to send triggers to the PC-Farm was inhibited, and immediately restarted after the CHOKE_OFF (ERROR_OFF) signal has been received.

Analyzing data offline and knowing the timestamp of choke (error) signal, it is possible to reduce the length of the burst to the effective time in which triggers are delivered. Special triggers as choke and errors are sent without delay in order to act immediately in reducing the pressure of data requests on detector side.
Figure 3.23 shows the difference of the timestamp between triggers sent to the PC-farm, and triggers sent to the detectors after the latency subtraction: all triggers are correctly handled by the L0TP, and there are no differences between the times sent to PCs and detectors. Only two events (CHOKE_ON and CHOKE_OFF) are located in bin -4000, which corresponds to 100 μs, correctly delivered without latency.

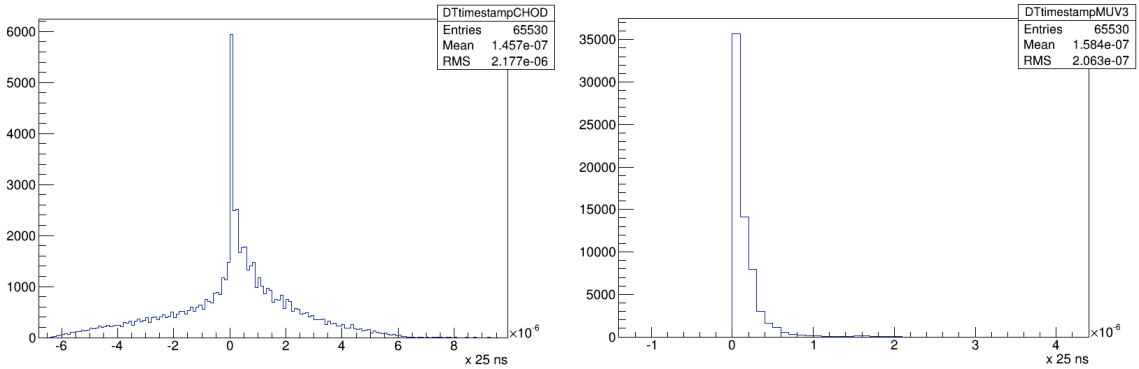
After the installation of the primitive parallel acquisition system (see section 3.4.13), it has been possible to record the primitives generated by detectors and replicate part of bursts in laboratory. One of the characteristics of the first version (deployed in 2015) of the primitive-generator firmware of some detector was to deliver primitives not ordered in time. This was due to the fact that primitives are processed in parallel, and the time to end the generating time depends on the complexity of an event. If a sorting algorithm is not implemented, it is possible to have time-unsorted primitives.

The L0TP was therefore designed to realign the outputs, even if the input primitives are not sorted in time. The primitives used for testing purpose have been collected in 2015 with a beam intensity at 35% of the nominal one. Figure 3.24 shows the timestamp profiles of the primitives generated by CHOD and MUV3, used to initialize the RAMs of the auxiliary



(a) CHOD primitive timestamp used for tests. **(b)** MUV3 primitive timestamp used for tests.

Figure 3.24. Timestamp profiles of primitives as input for the test.



(a) CHOD detector **(b)** MUV3 detector

Figure 3.25. Time difference between trigger N and trigger N-1 for run 3754

board. Because the RAMs can contain only 32k primitives, only a fraction of the total number of primitives collected during the entire burst has been selected. The sample refers to data collected in the central part of the burst, roughly $\sim$2 s from the SOB. Figure 3.25 shows the difference between two consecutive timestamps both for CHOD and for MUV3 samples. In order to check the capability of L0TP in finding coincidences, a software program has been written to mimic the logic of the FPGA. Then a comparison with the coincidences found by L0TP has been done. The coincidence time window has been set to $\pm$25 ns. The CHOD was used as reference detector by the L0TP. The sample contained 32768 primitives and the coincidences were 883.

The sample of triggers identified by the L0TP process and the one obtained with the software check implementation are fully overlapping. From the limited statistics that have been analyzed one can evaluate an upper limit to the inefficiency of L0TP logic. For this we

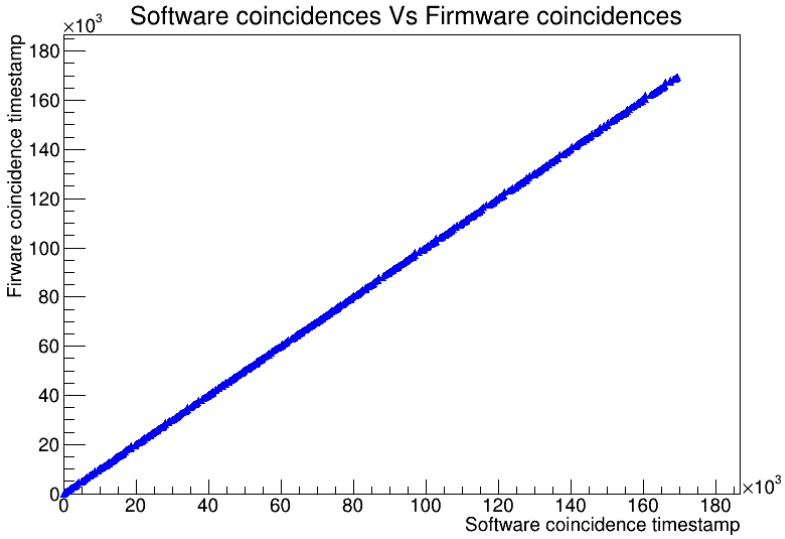


Figure 3.26. Timestamps found by L0TP firmware vs timestamps found by software simulation of the L0TP algorithm.

consider that the number of events that have been lost by L0TP is 0 out of 883, corresponding to an upper limit at 1.1 events with 32% confidence level (CL)[1]. The resulting upper limit of the inefficiency is then:

$$L0TP\ inefficiency < \frac{1.1}{883} = 1.2 \times 10^{-3}. \tag{3.9}$$

To further verify the correctness of the L0TP process, a similar test has been done to check the consistency of the timestamp sent to the detectors with respect to the one sent to the PC-Farm. They should be equal after the latency subtraction on the detector side. The comparison between the two samples of timestamps showed that 883 triggers were correctly dispatched to the detectors and to the PC-farm without any corruption. The upper limit for corruption probability has been derived using the same arguments as in the previous calculation, which gave an upper limit of 1.2×10^{-3}.
With the same sample of primitives, a further check on the downscaling procedure has been performed. Two masks have been used:

- mask0: *CHOD not MUV3*;
- mask1: *CHOD and MUV3*.

[1] Incidentally we observe that the number of wrongly recognized as good triggers by L0 is also 0, out of 883. However this possible category of events would not affect the efficiency, but only increase the throughput fraction.

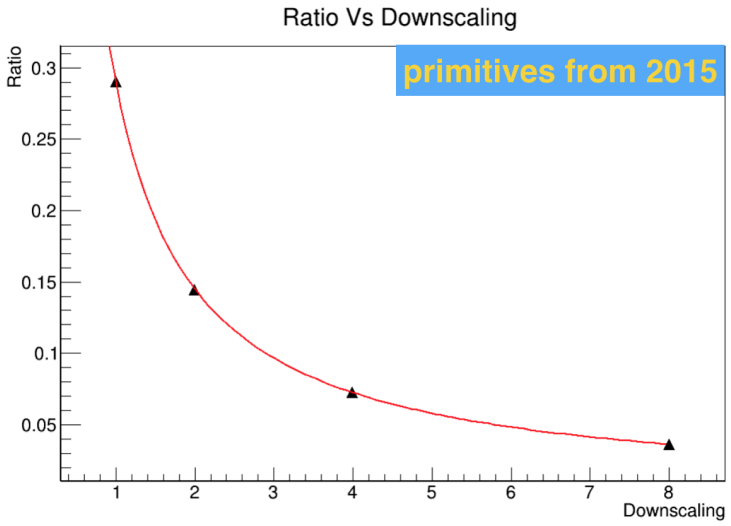


Figure 3.27. Ratio between mask *CHOD and MUV3* and mask *CHOD not MUV3*

Coincidences (or anticoincidences) have been searched in $\pm$ 50 ns around the reference detector (CHOD). The downscaling applied to triggers satisfying mask1 is maintained fix to one (without any downscaling applied), while the downscaling for mask0 is incremented, at each check, in steps of a factor 2. Figure 3.27 shows the results, confirming that the behavior of the output after the application of the downscaling is the one expected.

3.5.3 Test of Round Robin Program

In order to test the round robin program developed for sending data to the PC-Farm, the test bench has been used in laboratory to simulate the experimental conditions.

- test Bench sends primitives to L0TP;
- L0TP sends primitives to a workstation running the round robin program;
- the round robin program sends trigger packets to another computer where three different IPs have been added as aliases;
- WireShark[32] monitors the round robin.

The test has been done up to 1 MHz without any failure, while above 1 MHz L0TP starts to lose data because the ethernet buffers are full. Figure 3.29 shows the round robin mechanism, observed with WireShark. The destination IPs of the received packet, are monitored packet by packet. As an example, in figure 3.29 the IP 10.194.12.14 represents the transmitting IP

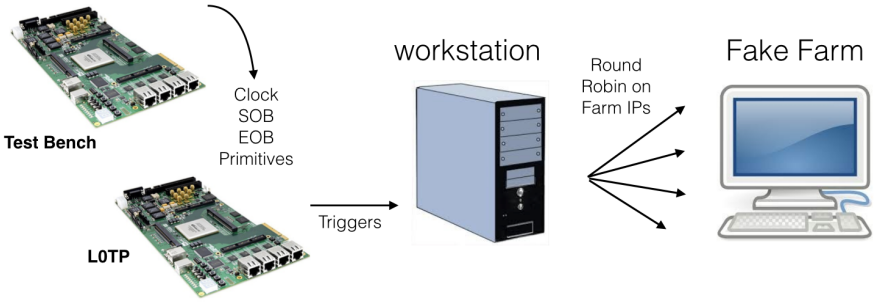


Figure 3.28. Illustration of the test bench to check the round robin behavior of the l0Server.

and the packets are sent to 10.194.20.46, then to 194.194.20.40, then to 10.194.20.41, to start again with the first IP address.

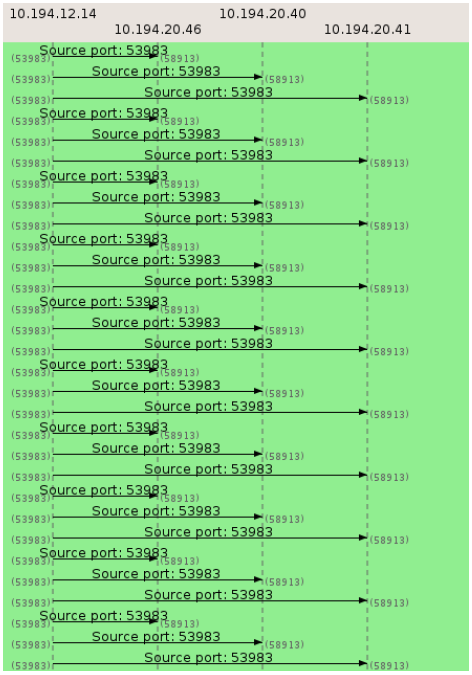


Figure 3.29. Round robin test: output page of WhiteShark for several transmission rounds.

3.6 Conclusions

In this chapter the main features of the last version L0TP firmware have been presented, together with the software tools needed to the correct operation of the trigger processor. The L0TP has been used since the first pilot run of the NA62 experiment, increasing in complexity and performances during 2015 and 2016 data taking period, in order to cover

all the requirements to maximize the quality of the data collected. The design has been tested and proven to behave according to specification. Further upgrades are still needed for future NA62 data taking runs, as the possibility to sustain the high latency of a primitive generator GPU-based to be included in the read out system of the RICH detector for the 2017 run[33]. Severe limitations have been experienced related to an extremely high instantaneous beam intensity, requiring a new version of the firmware, now scheduled to be available at the start-up of the 2017 data taking.

Chapter 4

2016 Data Taking

4.1 Introduction

This chapter describes the L0TP performance during the 2016 data taking, at a time when all the features of the firmware presented in the previous chapter were fully implemented. In particular, some remarks regarding the choice of the L0TP parameters and the rates sustained will be also discussed.

The data taking started in April and ended at the beginning of November 2016. The main goal of the 2016 data taking was to maximize the collection of the triggers for the $K^+ \to \pi^+ \nu\bar{\nu}$ decay mode. Nevertheless the size of the L0 bandwidth (1 MHz) allows implementing other trigger masks in parallel, in order to complete the NA62 tasks.

At the beginning of the data taking the beam intensity was set at 5% of the design intensity, foreseen at 750×10^6 particles/s (in jargon 750 MHz), and it was later increased step by step. Stable conditions have been obtained up to 35% of the nominal, which means a beam intensity of $\sim$260 MHz. Nonetheless, the choice of running at 35% of the nominal intensity for 2016 allows to collect good quality data for physics analysis, hopefully reaching the Standard Model sensitivity for the $K^+ \to \pi^+ \nu\bar{\nu}$ branching ratio. Figure 4.1 shows the number of triggers collected up to November 2016.

Tests with intensity higher than 35% have been performed, experiencing some instabilities of the TDAQ system, mainly due to the time structure of the beam itself. Figure 4.2 shows the time profile during a burst as measured with the RICH detector at 50% of the nominal beam intensity. Due to the high rate of primitives generated by the RICH, a reduction factor of 10 has been applied, collecting one primitive out of ten. Scaling the rate (reported on the y-axis) by a factor 10, it is possible to read directly the reached value.

Primitives generated by CHOD and MUV3 are characterized by similar time structure and values. We can identify two main issues possibly originating data acquisition instabilities:

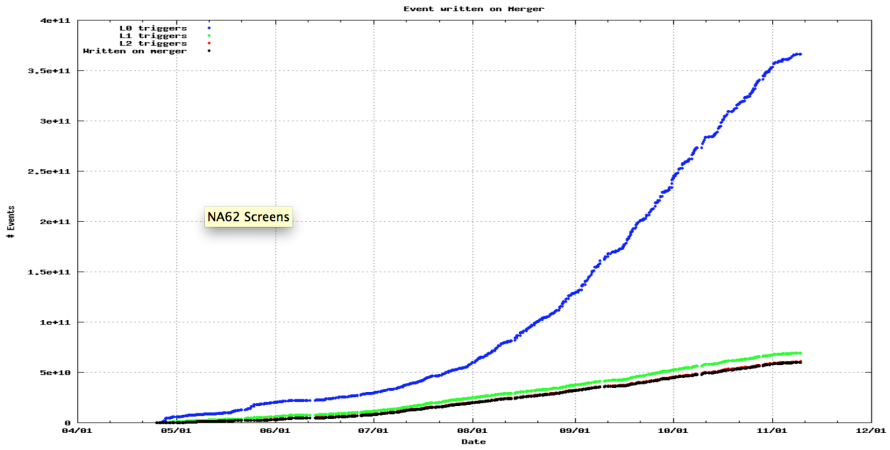


Figure 4.1. Number of triggers collected up to November 2016. The blue line shows the number of L0 triggers, the green one is the number of L1, the red is the number of L2. Finally the black one represents the number of triggers written on disk.

- a prominent peak at the beginning of the beam extraction (see figure 4.2);
- intensity modulation of the beam due to the 50 Hz component (see figure 4.3).

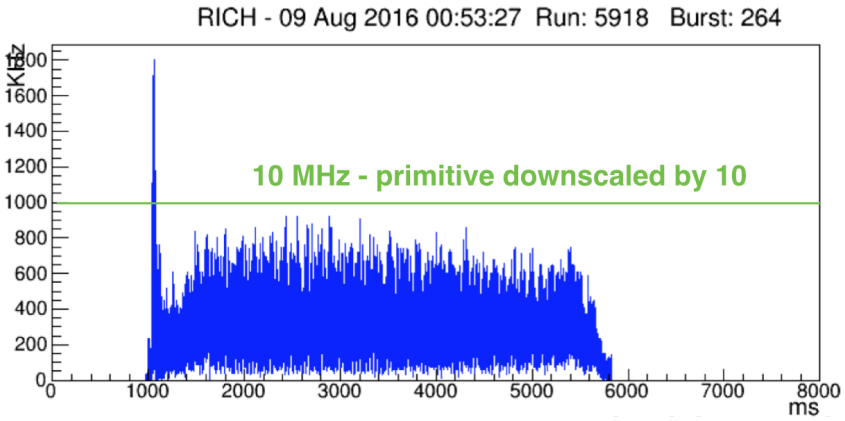


Figure 4.2. Time structure recorded with the RICH detector of a typical burst at 50% of the nominal beam intensity.

As previously discussed the TDAQ is designed to work up to 10-12 MHz of primitives. Spikes at the beginning of the extraction can seriously affect the data taking because the instantaneous intensity reaches ∼18 MHz. As a consequence, the input buffers of the data acquisition systems are filled, causing sizeable loss of events.
The second issue is the modulation of the intensity with 50 Hz harmonics. While the origin of the beam intensity fluctuations is not understood, the accelerator crew tried to mitigate this effect, without a complete success. For this reason, already at 35% of the nominal flux,

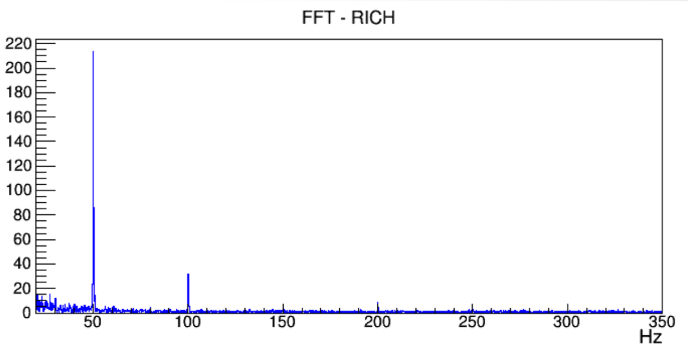


Figure 4.3. Fast Fourier Transform of the beam profile.

these modulations often generate peaks of intensity higher than 10 MHz. These issues were not expected at the beginning of the data taking and the development of new firmware for L0TP and detectors to run at full intensity is presently ongoing.

With the present conditions, running at 35% of the nominal beam intensity was the best compromise between the number of kaon decays collected and the capability of sustain the rate.

To maximize the NA62 physics output, seven different masks and a control trigger have been run in parallel to select the triggers, as reported in table 4.1.

mask number	name	configuration	downscaling
0	$e\nu_H$	RICH * NewCHOD * !MUV3	300
1	$\pi^+\nu\bar{\nu}$	RICH * !QX * UTMC * !MUV3 * !LAV12 * ! LKr> 20 GeV	1
2	$\mu\mu$ exotic	RICH*Q2*MO2*!LKr>20 GeV	1
3	$\pi\mu$	RICH*QX*MO1	5
4	$\pi\mu$ exotic	RICH*Q2*MO1	10
5	multi-tracks	RICH * QX	50
6	$\mu\mu$	RICH*QX*MO2	1
Control	control trigger	CHOD	400

Table 4.1 Trigger masks used for the 2016 data taking.

The different trigger components forming the masks are listed in table 4.2.

- **mask 0**: to select $K^+ \to electrons$ and collect data to search the heavy neutrino production in the $K^+ \to e^+\nu_H$ decay mode.
- **mask 1**: to select $\pi^+\nu\bar{\nu}$; it requires RICH hits, only one quadrant of the hodoscope firing with a cut on the multiplicity of 6 hits, no muon signals in MUV3, no hits in LAV12 and the energy in LKr lower than 20 GeV to reject $\pi^+\pi^0$;

Name	Description
RICH	Clusters of at least 3 hits in time from RICH
!MUV3	Veto signal from at least one PM of a pad of MUV3
!Qx	Veto signal from two hits in opposite quadrants of NewCHOD
!LAV12	No clusters in LAV12
UTMC	Multiplicity of hits < 6 in NewCHOD
Q2	Two hits in NewCHOD
M01	Outer muon in MUV3
M02	Outer dimuon in MUV3
!LKr>20	Veto on the LKr energy greater that 20 GeV

Table 4.2 Description of the single parts of trigger configurations used in 2016.

- **mask 2**: to select 3-body decay as $K^+ \to \pi^-\mu^+\mu^+$; Q2 requires 2 or more tracks in the CHOD without topology constraint. At least one track has to be confirmed by the RICH and two muons from MUV3. π^0 are then rejected by vetoing events with a deposit of energy higher than 20 GeV in the LKr.

- **mask 3**: it is designed to select 3 body decay a single muon in final state; these triggers are collected for searching decays as $K^+ \to \pi^+\mu^+e^-$ or $K^+ \to \pi^+\mu^-e^+$;

- **mask 4**: it is identical to mask 3, but the the request of hits in opposite quadrants is dropped;

- **mask 5**: to select multi-track events with hits in opposite quadrants. The sample is used to normalize the $\pi^+\nu\bar{\nu}$ on 3π decays;

- **mask 6**: to select two muons in the final state with hits in opposite quadrants topology.

Other parameters characterizing the L0TP and set for the standard run condition are reported in table 4.3. A very brief description of the meaning of terms in table 4.3 is following. For full details please refer to chapter 3.

- *Reference Detector*: the primitives of the reference detector contain the time against which primitive times of the other detectors are compared. It is used by the MMU to avoid the fragmentation in the RAM for the primitive alignment during the L0TP data processing;

- *Control Detector*: it is used to collect a minimum bias sample of events useful to study trigger efficiencies;

Parameter	**State**
Reference Detector	RICH
Control Detector	CHOD
Time Granularity	12.5 ns
Time coincidence with the reference detector	$\pm$ 10 ns
Dead Time between triggers	75 ns
Latency	200 μs
Detector with the maximum delay	LKr
Maximum delay in number of 6.4μs	8

Table 4.3 Parameters used in 2016 standard run conditions.

- *Time granularity*: set to 12.5 ns, it is the time-length of one cell of the RAM for the time alignment. As described in the previous chapter, when primitives are received and unpacked by the L0TP, they are aligned with a precision equal to the granularity of the RAMs, using the time as memory address;

- *Time coincidence with reference detector*: a time window cut of $\pm$ 10 ns, it is applied to a detector, required by the trigger mask, with respect to the reference detector time.

- *Dead time between trigger*:, set to 75 ns, it is the minimum interval between two consecutive triggers. This interval is mandatory being strictly required by the TTCex logic. It implies that, at full trigger rate, 7.5% of the triggers are rejected. All trigger types are equally affected and this avoids a bias of the final sample.

- *Latency*: set to 200 μs, it is the elapsed time between the time of the primitive of the reference detector in input and the trigger time in output. It has to be long enough to accommodate the L0TP processes and to manage all data before delivering triggers, ensuring a complete matching between all the sources composing the trigger selection;

- *Detector with the maximum delay*: indicates the system which generates the primitives with the largest delay. L0TP does not start to process events if this detector has not sent data. Meanwhile, packets from the other sources are stored in the MMU buffers. The maximum delay detector can be evaluated from data collected with the primitive parallel acquisition system (described in section 3.4.13), using the sending timestamp written in the header of the MTP as shown in figure 4.4.

- *Maximum delay in number of packets*: set to 8, it is the number of packets that the delay generator module has to count before starting to process events. It has been

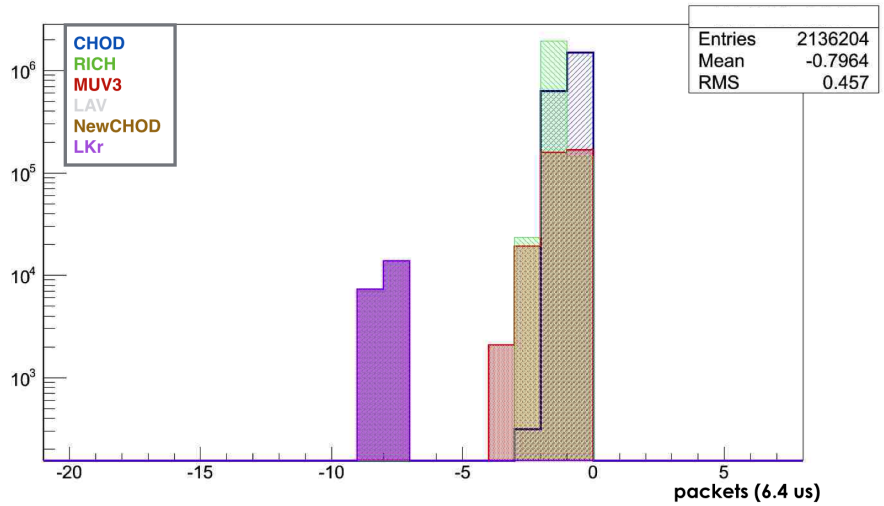


Figure 4.4. Difference between MTP timestamp and primitive timestamp in MTP.

computed as

$$Number\ of\ delayed\ packets = MTP\ timestamp - timestamp\ of\ primitives\ in\ MTP \tag{4.1}$$

4.2 Intrinsic Inefficiencies due to parameter selection

The choice of some of the parameters described in the previous section as the reference detector, the granularity and the time coincidence window has important consequences on the L0TP efficiency in finding the correct trigger pattern.
The L0TP reference detector, used in the MMU logic, is the RICH detector. It means that a primitive from the RICH is required to be present in all the masks. The RICH system has been chosen for the good time resolution, similar to the KTAG, but with the relevant advantage that the rate is much lower [1]. The distribution of the time difference between RICH primitives and KTAG hits is shown in figure 4.5. When fitted with a Gaussian function, the distribution has a sigma $\sigma \approx 320$ ps.

The granularity of the RAMs for the alignment of the primitives in the L0TP mechanism has been set to 12.5 ns. If the granularity of RAMs is X, a time cut greater than X between the reference detector and the other sources has no sense, because the RAM resolution dominates

[1] The KTAG detects all the kaons in the beam, whilst the RICH only a relevant fraction of those which decayed.

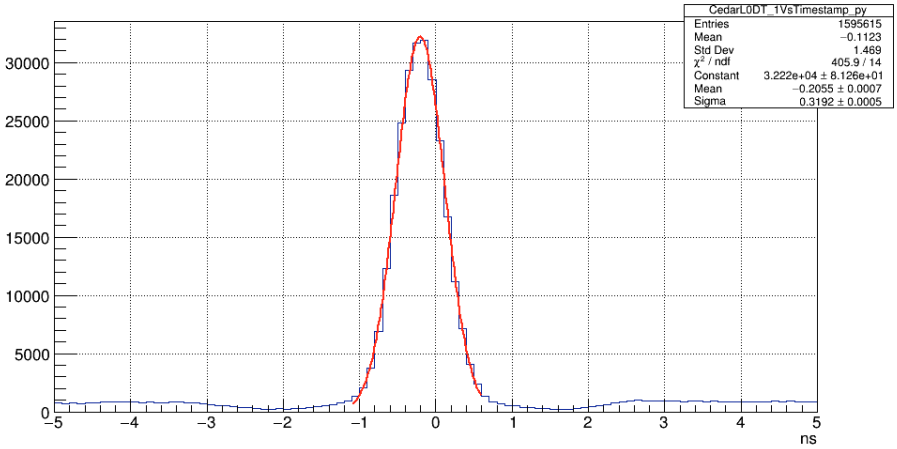
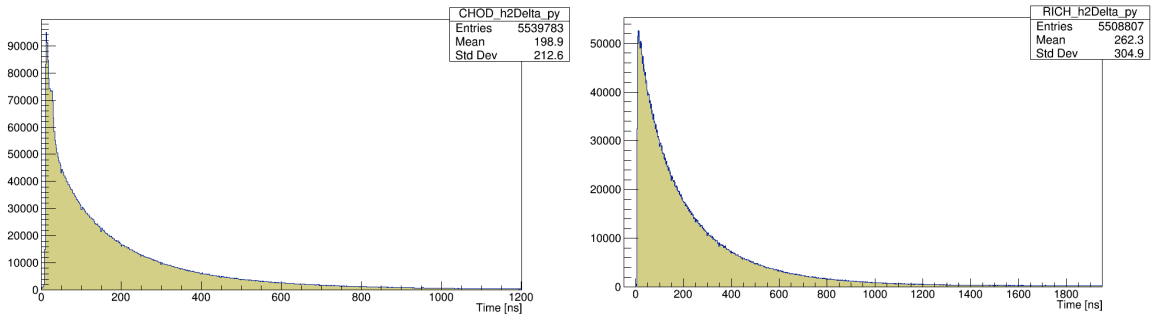


Figure 4.5. Time difference between KTAG and RICH primitives. The resolution is ∼320 ps.

the precision of the cut itself. To select the value of the granularity, the time distribution of the primitives with respect to the reference detector has been considered. They are shown in figure 4.8. For the Calorimeter primitives, the standard deviation of the distribution is 6.8 ns, the largest compared to the other detectors. The granularity of the RAMs has been chosen equal to approximately two standard deviations of this distribution.



(a) CHOD **(b)** RICH

Figure 4.6. Difference between two consecutive primitives acquired with the parallel acquisition system for CHOD and RICH detectors.

The price to pay for this choice is that primitives coming from the same detector which are closer than 12.5 ns could overwrite the previous ones in the alignment RAMs, and only the last primitive stays written and taken in the trigger generation process. The overwriting primitives modify the previous primitives in the specific RAM location, changing:

- Primitive ID;

- Fine time.

Concerning the effect of the overwriting of primitive IDs, masks associated to a trigger which does not require any specific ID are not affected by the modification. On the other hand by changing the fine time, a primitive, originally on time, can be shifted out of time with respect to the reference detector. In this case, the trigger is lost, yielding an intrinsic L0TP inefficiency.
Figure 4.6 shows the distribution of the time interval between two consecutive primitives in RICH and CHOD detectors. Table 4.4 reports the probability that a primitive stored in the L0TP input RAM is overwritten by a successive one, measured at $\sim$35% of the nominal beam intensity. The measurement has been done using data coming from the primitive parallel acquisition system. The statistical errors on the figures reported in table 4.4 are completely negligible being based on very large data samples.

Figure 4.7 shows the probability of having a primitive overwritten in the L0TP RAMs as a function of the intensity. In this plot, only detectors with a significant rate have been taken into account, thus neglecting LAV and Calorimeter primitives.

CHOD	0.5 %
RICH	1.0 %
LAV	0.02%
MUV3	1.7 %
NewCHOD	1.0 %
Calorimeters	0.02%

Table 4.4 Probability of having a primitive overwritten in the L0TP memories due to the RAM granularity, measured at 35% of the nominal beam intensity.

As shown in figure 4.7, a trend of the overwriting process with the intensity is observed for all the systems. Nevertheless, the probability of having a change of the PrimitiveID or a time shift is the only one on which we are interested. The probability of having a PrimitiveID modification for consecutive events closer in time than the granularity of the L0TP RAMs can be measured with the data collected. Results for the PrimitiveID configurations used in 2016 for the 35% of the nominal beam intensity are reported in table 4.5. Again, due to the large data samples on which are based the measurements, the statistical errors are negligible.

Trigger Type	percentage
Qx	0.5%
Q2	0.5 %
UTMC	0.4 %
MO1	0.7 %
MO2	0.5 %

Table 4.5 Percentage of events that change the primitive ID due to the overwriting process in the L0TP memories as measured at 35% of the nominal beam intensity.

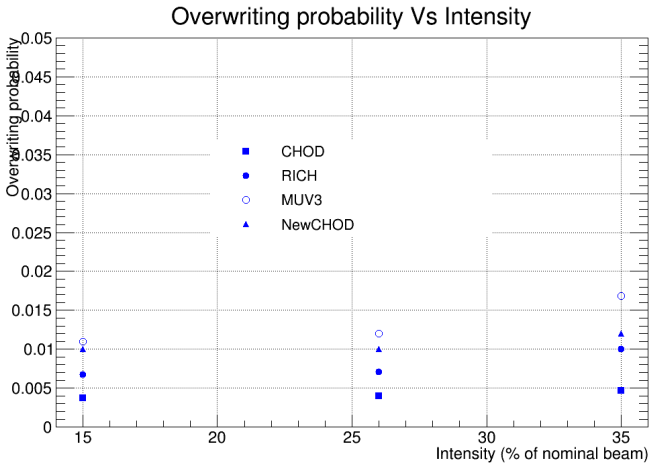


Figure 4.7. Probability of the overwriting process in the L0TP input memories as a function of the beam intensity.

The probability that a second primitive overwrites a previous one, changing the fine time and shifting the event out of time can be calculated as:

$$P = RateOfOverwriting \times \\ ([granularity] - [coincidence\ window])/[granularity] \tag{4.2}$$

where the*granularity* is the interval in which the fine time can fill the RAM cells ($\pm$ 12.5 ns), while the coincidence window is the window set in the configuration file of L0TP ($\pm$ 10 ns). This is certainly an overestimation of the true value. It does not include the correlation between primitives that exists because they are related to the physics and aligned in time. Table 4.6 reports the results of the estimation and the figures should be read as upper limits rather than central values.

Detector	percentage
CHOD	0.1 %
RICH	0.2 %
LAV	0.004%
MUV3	0.3 %
NewCHOD	0.2 %
Calorimeters	0.004%

Table 4.6 Probability of losing a good events due to the overwriting process on L0TP.

Looking at the $\pi\nu\bar{\nu}$ trigger configuration, the probability to lose a primitive that was in time is the sum of the probability of each sub-system included in the mask. The probability to be shifted out of the coincidence window or with a modified primitive ID is correlated because the primitive arrives with both the information. We take into account only the higher inefficiency.

$$\begin{aligned} P_{Losing\ Primitives} = P_{RICH}(0) + P_{LAV}(0.004\%) + P_{NewCHOD}(0.5\%) + \\ + P_{MUV3}(0.3\%) + P_{Calorimeters}(0.004\%) = 0.88\% \end{aligned} \tag{4.3}$$

where we took RICH without any probability to fail because of its reference role in the mask. By construction a RICH primitive is always requested in the L0TP masks and the coincidence window is centered on the RICH primitive time. The fine time fluctuations are kept in account requiring a fine time span window of $\pm$ 12.5 ns.
Summarizing the results, the global L0TP inefficiency due to primitive overlapping for 2016 data taking for the $\pi\nu\bar{\nu}$ trigger is:

$$1 - \varepsilon_{L0TP}^{\pi\nu\bar{\nu}} = 0.88\% \tag{4.4}$$

while for the other triggers the same calculation have been performed and results are reported in table 4.7.

Trigger	inefficiency of L0TP
$e\nu_H$	0.8 %
$\mu\mu$ exotic	0.84%
$\pi\mu$	0.8 %
$\pi\mu$ exotic	0.8 %
multi-tracks	0.5 %
$\mu\mu$	0.8 %

Table 4.7 L0TP inefficiency for each mask.

Detector	Standard deviation
CHOD	(1.65 $\pm$ 0.01) ns
LAV12	(1.99 $\pm$ 0.04) ns
MUV3	(0.66 $\pm$ 0.01) ns
NewCHOD	(1.24 $\pm$ 0.01) ns
Calorimeters	(3.20 $\pm$ 0.01) ns

Table 4.8 Standard deviation of the time alignment with respect to the reference detector (RICH).

This is not the global inefficiency of the L0 trigger, but only the losses of the L0TP processing the triggers. Here in fact we are not taking into account either the efficiency of detectors for producing primitives and delivering them correctly or the efficiency of the detector itself. This evaluation aims to show the L0TP contribution to the global trigger efficiency, which is reported with all the details in chapter 5.
Even though the inefficiencies are below 1%, there is room for improvement. On the L0TP side, for example, there is the possibility to decrease the granularity down to 6.25 ns, decreasing the probability of overwriting events. A reduction of the granularity applied to all the detectors to values smaller than 6.25 ns would imply an efficiency loss at least for the Calorimeters which have a resolution of $\sim$6.25 ns. A possible upgrade of the firmware would require the implementation of different resolution of the RAMs for the alignment for the different sources.

4.3 Time resolution of primitives and coincidences at L0

L0TP sends to the PC-Farm all the information regarding the memory slot in which the trigger has been found. It is possible, by analyzing the data, to check the time alignment of the primitives with respect to the reference time around which the cut has been applied. Having adjusted the offset of the detectors, the data confirm that the alignment is within $\pm$ 200 ps. Thus the efficiency of the cuts depends only on the spread of the time measurement.

In table 4.3 is reported the standard deviations between different detectors participating to the L0 trigger selection with respect to the RICH, which is the time-reference system. The results have been obtained considering the standard deviation of a gaussian-fit of the time difference distribution between detectors and reference system.

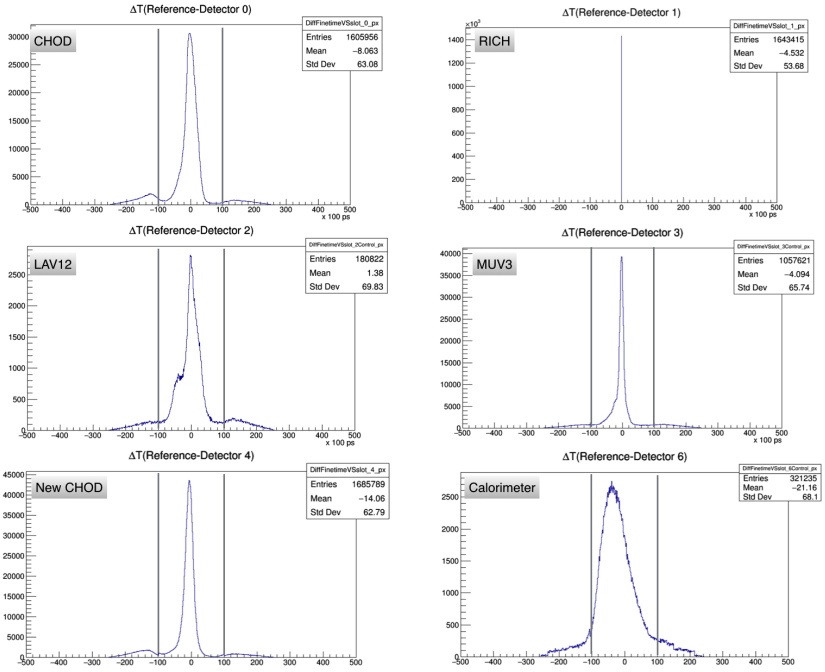


Figure 4.8. Time difference of detector primitives respect to the reference time (RICH).

All detectors have been aligned within 2 ns or better (except for the Calorimeter primitive). Thus for any given detector the acceptance time window depends mainly on the spread of that detector. The Calorimeter primitive time distribution is quite large due to its intrinsic primitive time resolution. For 2016 data taking, it has been applied a time cut around the time-coincidence peaks considering a window of about 3σ of the worst distribution. This justify the cut of $\pm$ 10 ns around the peak. The RICH distribution shows a tight spike at zero due to the fact that it is used as reference time. For detectors in veto, the difference between their primitives and reference time shows the cut applied on L0TP (see figure 4.9).

4.4 Rates

The stability of the trigger system is one of the required conditions for the success of a data taking. This can be monitored by looking at the rates and ratios of the detector primitives.

In the EOB data of L0TP all the information regarding the internal counters of the L0TP firmware are stored. This information can be used to monitor the rates both of the triggers dispatched from the detectors to L0TP and the number of triggers for each mask. Table 4.9 shows the average number of primitives read by L0TP, integrated for the burst length at 35% of the nominal intensity.

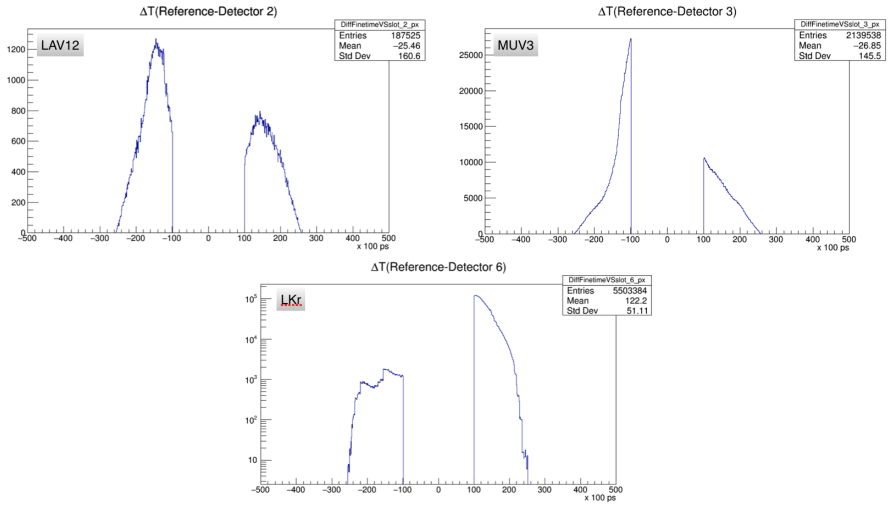


Figure 4.9. Time difference of detector primitives respect to the reference time (RICH) for detector in veto.

Detector	Average number of primitives per burst
CHOD	32.0 M
RICH	26.2 M
LAV	4.5 M
MUV3	29.9 M
NewCHOD	30.6 M
Calorimeters	5 M

Table 4.9 Average number of primitives received from L0TP per burst.

Rate is evaluated by dividing the number of primitives for the *effective spill length* of the burst. The effective spill length is an indicator of the non-uniformity of the arrival times of the beam particles. In the optimal situation the accelerator provides a uniform distribution of N particles over the total spill length T. In that case, the spill duty factor is defined as

$$DF = \frac{Effective\ Spill\ Lenth}{Total\ Spill\ Length} \tag{4.5}$$

is equal to one. The effect of deviations from uniformity is has the consequence that, for DF < 1, the same number of particles N are arriving over shorter time – the effective spill. Moreover, the data-taking capabilities of the experiment may be reduced due to the fact that any particle that arrives during the detector dead time is effectively lost. In the case of primitive generators, for example, when too many primitives are produced in 6.4μs, fulling

the ethernet packet size, the excess has to be discarded. Deviations from uniformity occur at various frequencies, and as we have seen, they are particularly strong at multiples of 50 Hz due to modulation of the magnet currents. Being the F(t) the beam current as a function of time, the effective spill length, T_{eff} is given by [34]

$$T_{eff} = \frac{[\int_0^T F(t)dt]^2}{\int_0^T F(t)^2 dt} \tag{4.6}$$

The instantaneous numbers of particles, F(t), can be well approximated by the number of primitives generated by the detectors. Results from effective spill length calculated with primitives generated by several detectors are reported in table 4.10. The difference among the effective spill lengths are due to the primitive rates used to approximate F(t). The result obtained with the CHOD, due to its higher number of primitives generated, is the closest to the actual one and for this reason the CHOD value is the one chosen for the following considerations.

Detector	Effective Spill Length (seconds)
CHOD	3.8
LAV	3.5
MUV3	3.6

Table 4.10 Effective Spill Length.

In table 4.11 the rates of primitives normalized to the effective spill rate of measured by the CHOD are reported as measured at 35% of the nominal beam intensity.

Detector	Average Rate of primitives per detector
CHOD	8.4 MHz
RICH	6.8 MHz
LAV	1.2 MHz
MUV3	7.9 MHz
NewCHOD	8.1 MHz
Calorimeters	1.3 MHz

Table 4.11 Average rate of primitives per detector.

Recalling that the designed values of incoming primitives are 10 MHz per source, the L0TP is working at $\sim$ 80% of its capabilities for CHOD, MUV3 and NewCHOD systems.

Moreover, it is possible, in this case, to determine how the number of primitives per burst scales with the beam intensity. The intensity of the beam is measured using the Argonion detector [35]. Argonion detector is a multiwire proportional chamber operating in ionization

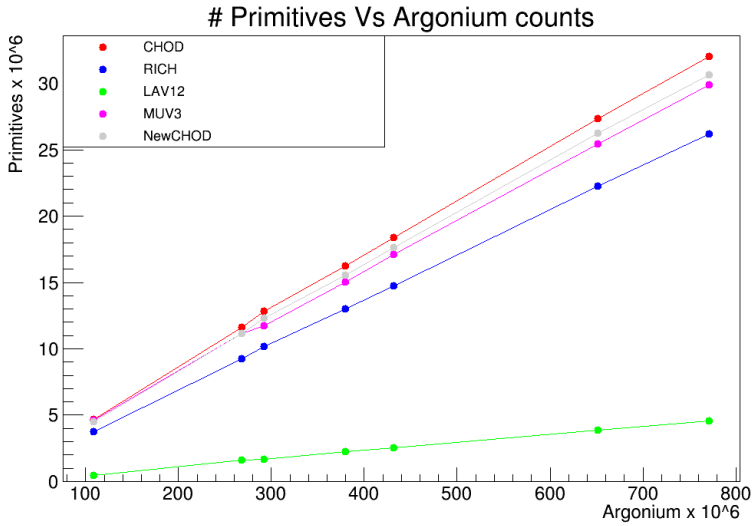


Figure 4.10. Number of primitives versus the Argonion counts.

mode and placed at the end of the NA62 experiment to measure the beam integrated intensity. The proportional chamber integrates the number of particles received during the entire burst length.

Figure 4.10 shows the results obtained: the number of primitives scales linearly with the intensity. With a linear fit to the points, it is possible to extrapolate the rates that occur at the nominal beam intensity, considering again the effective spill length equal to 3.8 seconds. Results are reported in table 4.12.

Detector	Average rate at the nominal beam intensity
CHOD	21.7 MHz
RICH	17.8 MHz
LAV	3.2 MHz
MUV3	20.2 MHz
NewCHOD	20.8 MHz
Calorimeters	3.5 MHz

Table 4.12 Average rate of primitives at the nominal beam intensity.

As the table 4.12 shows, at the nominal beam intensity, the rate is more or less double with respect to the one foreseen in the 2010-technical design report [13].

The reasons are still not completely understood.

Summarizing these considerations, an upgrade of the entire L0 trigger system seems to be mandatory to handle the primitive rates and perform a full intensity run in 2017.

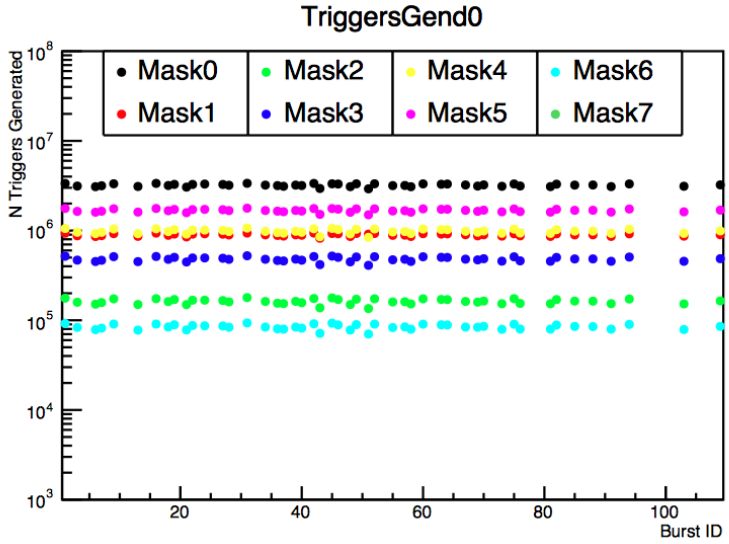


Figure 4.11. Number of triggers generated per burst.

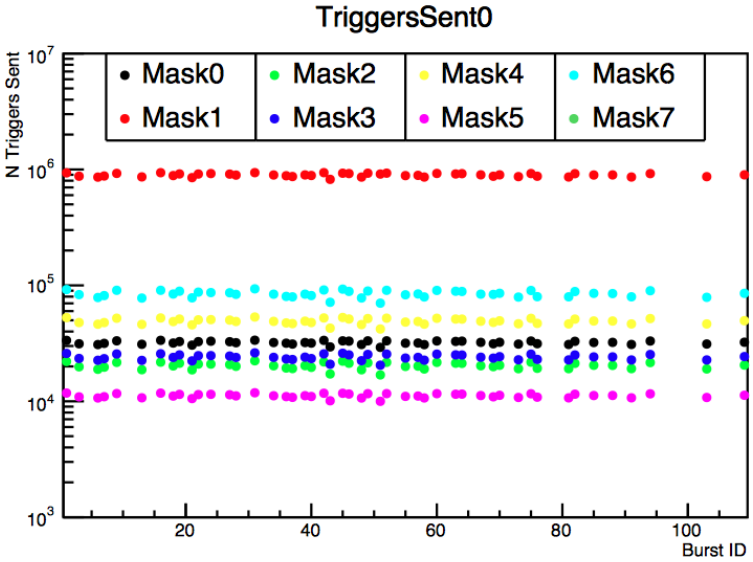


Figure 4.12. Number of trigger dispatched per burst, after applying downscaling.

On the output side of L0TP, the maximum trigger rate by design is 1 MHz. In table 4.13 is reported the number of triggers generated for each mask before applying any downscaling, after the application of the downscaling and the rate per second considering the effective spill length. In figure 4.11 and 4.12 triggers generated and dispatched per burst are reported.

Type	Number of generated triggers	rate	downscaling	Number of dispatched triggers
eV_H	4 M	1.05 MHz	200	13 k
$\pi^+\nu\bar{\nu}$	1.2 M	0.31 MHz	1	1.2 M
$\mu\mu$ exotic	280 k	0.07 MHz	1	280 k
$\pi\mu$	750 k	0.2 MHz	5	150 k
$\pi\mu$ exotic	1.5 M	0.4 MHz	4	150 k
multi-tracks	2.3 M	0.6 MHz	50	46 k
$\mu\mu$	150 k	0.04 MHz	1	150 k
Control Trigger	24 MHz	6.7 MHz	400	75 k

Table 4.13 Output triggers per burst.

The total output from L0TP integrated on the burst duration for stable data taking is about 1.65 M of triggers. Dividing this number for the effective spill rate, it returns the number of L0 trigger per second which is

$$L0\ Trigger\ Rate_{2016} \sim 435kHz \tag{4.7}$$

The rate of the L0TP output can also be measured from the distribution of the time difference between two consecutive triggers. It is reported in figure 4.13. Fitting the distribution with an exponential function, one can extract the rate from the slope parameter of the fit. The slope has to be divided by the bin-width of the distribution, which is in timestamp units (25 ns). The output rate measured with this procedure is $\sim 440kHz$, close to the one obtained dividing the number of triggers integrated in the burst time for the effective spill length. It means also that the effective spill length measured with the CHOD primitives is a good approximation of the real value.

The L0TP stability during the data taking can be monitored by comparing the number of triggers to the output of the Argonion detector (see figure 4.14).

In the analysis, bursts for which the ratio between the number of triggers and the Argonion counts deviates from the mean value can be flagged as bad bursts and discarded.

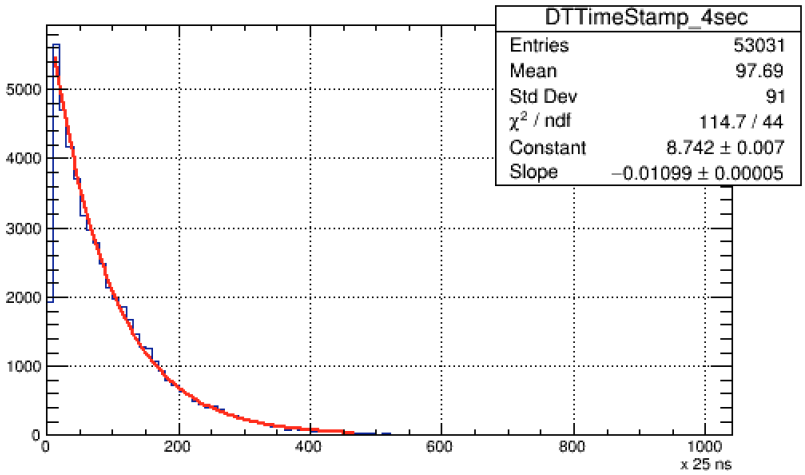


Figure 4.13. Time difference between two consecutive triggers.

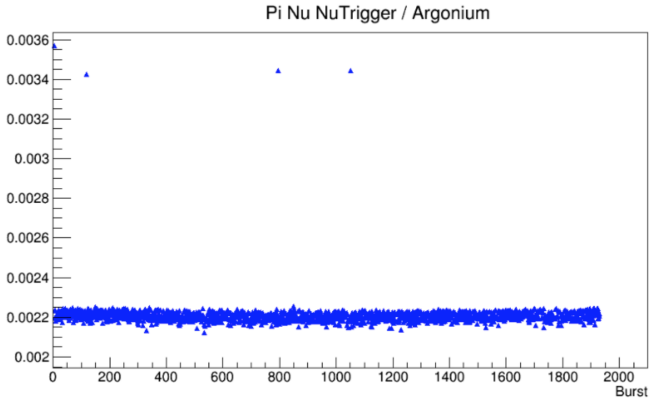


Figure 4.14. Number of $\pi\nu\bar{\nu}$ triggers divided by Argonion counts.

4.5 Summary and perspectives

The NA62 L0TP has been successfully commissioned and validated in the challenging environment of a high intensity beam, resulting in a highly efficient and stable system. It delivered events for physics analysis throughout the first runs of NA62 operations from 2014 to 2016 enabling the start-up of NA62 physics program. After two pilot runs in 2014 and 2015, the physics data taking was started in 2016 with a beam intensity up to $\sim$35% of the nominal design value. The incoming primitive rates from detectors to the L0TP, already at this intensity, is close to the design value, at least for CHOD, New CHOD and MUV3 detectors. To reach the designed intensity a new improved version of the L0TP has to be implemented. Output rates are close to 50% of the nominal values, resulting in a smooth data collection. The L0TP shows highly uniform and stable performances.

The development of the upgraded firmware for the 2017 data taking is ongoing. In particular, in the new firmware a new data format will be used to label a primitive, reducing the primitive size from 64 to 32 bits. With this trick the primitive input that can be sustained is doubled, without changing the structure of the trigger logic. On the output side, an auto-CHOKE system will allow to stop the data taking if the buffers in the L0TP are close to be full, avoiding loss events during the burst. Finally, reorganization of the buffer sizes and latencies has to be performed to allow the L0TP system to receive data generated by a primitive generator GPU-based, as the RICH read-out upgrade foresee. These modifications are scheduled to be ready for April 2017.

Chapter 5

L0 Trigger Efficiencies

5.1 Introduction

The reduction of the amount of data acquired by a high-rate experiment which looks for rare events is mandatory in order to collect only events with the required characteristics. In the NA62 experiment in particular, the presence of muons in the final state characterizes roughly 70% of the total kaon decay rate. On the other hand, $K^+ \to \pi^+\pi^0$ mode accounts for 20% of the total kaon decays. To reach the sensitivity level of $K^+ \to \pi^+\nu\bar{\nu}$, it is essential to strongly reject these backgrounds already at the first trigger level.

The evaluation of the efficiencies of the L0 trigger is fundamental to control the data quality, in order to maximize the rejection of the backgrounds and minimize the loss of the signal. For this reason, the L0 trigger configuration for the main data stream has to exploit the characteristics of the signal signature: the presence of a π^+ of energy up to 35 GeV and missing energy above 40 GeV, without any muon in time-coincidence. The goal of this analysis is to quantify the capability to fit the design requests at the first trigger level.

The performance of the NA62 L0 trigger system has been evaluated with the data collected in 2016 at the beam intensity of $\approx 35\%$ of the nominal.

In order to measure the efficiencies of the detectors involved in the L0 trigger generation, an unbiased sample, where no specific masks are applied, is needed. For this reason events collected with minimum bias control trigger have been used. Control triggers are recorded requiring a time-coincidence between horizontal and vertical slabs in the CHOD detector. This sample is downscaled of a factor 400, recording almost 60k triggers per burst.

In the analysis framework, specific conditions are applied offline to reproduce the trigger behavior, looking at the L0TP data in which both the primitive ID and the time of all the primitives sent for each trigger are stored.

The analysis has been performed considering the following decay modes:

- $K^+ \to \mu^+ \nu$ (in jargon $K_{\mu 2}$, BR: 63.56 %);
- $K^+ \to \pi^+ \pi^0$ (in jargon $K_{\pi 2}$, BR: 20.67 %);

An estimated value for each individual rejection factor has been calculated, and the efficiency in detecting a single pion detection in the signal region has been quantified.

The analysis can be split to in different blocks:

- selection of good tracks from the spectrometer (One Track Selection);
- selection of good kaons;
- selection of different samples of $K_{\mu 2}$ and $K_{\pi 2}$;
- calculation of the efficiencies.

The selection of the samples cannot be the same for each detector. Different samples of the same particles are needed to exclude from the selection those detectors on which the efficiency is measured.

The values of the different cuts here reported have been validated with Monte Carlo simulation. For the sake of brevity, Monte Carlo results are not reported in this thesis. As reference, detailed studies can be found in [36] and in [37].

To calculate the efficiencies, two histograms are filled with the number of passed (k) and total (N) events, where passed events are the ones for which a trigger has been confirmed by the L0 system under study.

The expectation value of the number of passed events is given by the true efficiency times the total number of events. One can estimate the efficiency by replacing the expected number of passed events by the number of passed events.

$$k = \varepsilon \times N \to < \varepsilon > = \frac{k}{N}$$

This analysis reports the efficiencies taking in account the inefficiency of the detector itself, of the primitive firmware generators and of the L0TP coincidence algorithm. In some cases it is possible to separate the different components, useful only for debug purposes.

The chapter is structured as follows: In the first section a selection to isolate good tracks and good kaons is presented. Then we focus our attention on the detector behavior for the $K_{\mu 2}$ sample, quantifying the rejection power of the L0-trigger for this kind of events.

A similar analysis on $K_{\pi 2}$ follows adding the calculation of the single pion detection efficiency.

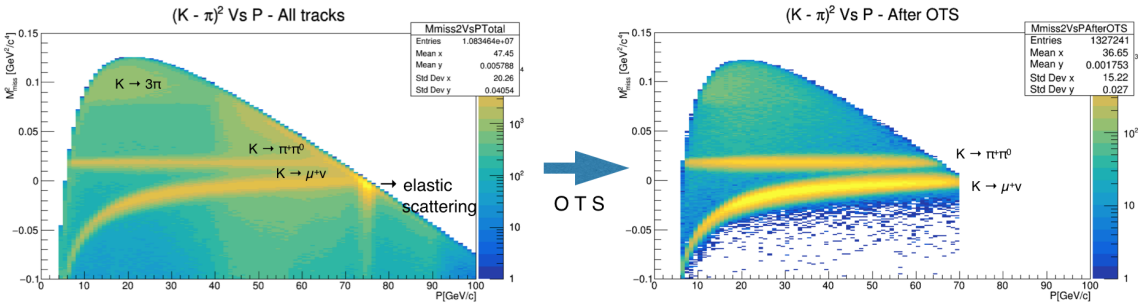


Figure 5.1. Squared missing mass under π^+ hypothesis as a function of the momentum of the track measured in the STRAW spectrometer. On the left no selection has been applied, while on the right the One Track Selection has been performed.

5.2 One Track Selection

The first step of the analysis is the identification of the tracks properly reconstructed by the STRAW-spectrometer, which will be referred to as One Track Selection (OTS).

The OTS requires:

- positive charge;
- χ^2 of reconstruction of the track lower than 20;
- a hit in each of the four straws;
- track in the geometrical acceptance of all chambers;
- coordinate z of the reconstructed vertex in the fiducial region: 130 m < Z < 165 m;
- measured momentum: 6 GeV/c < P < 70 GeV/c.

The vertex has been calculated using the spectrometer track and the kaon track as inputs of the closest distance of approach (cda) algorithm. The cda for a valid vertex has to be less than 4 cm. The requirement of momentum lower than 70 GeV/c excludes all the beam particles that are deflected by elastic interaction in the GTK detector or in the beam collimators. These particles are basically pions and kaons with momentum close to 75 GeV/c. The main variable used for the event selection is the squared missing mass (m^2_{miss}) between the incoming kaon and the decayed particle. For this reason, in-time kaons have to be matched with the tracks as briefly described in the section 5.3. Figure 5.1 shows the missing mass before and after the one track selection.

5.3 Selection of kaons

In order to characterize completely the kaon decays, the properties of the incoming particles have to be measured. However, for sake of simplicity, in this analysis the GTK detector is not used to determine momentum and position of the kaons. Instead, the nominal beam values have been assumed for all the beam particles:

- $dx/dz = 0.0012$;
- $dy/dz = 0$;
- (x,y) at Z=101.8 m (GTK3 position);
- total momentum $P = 75$ GeV/c.

The identification of a time coincidence between an incoming kaon and a spectrometer track is performed with the Cedar detector. To identify a good kaon, the following constraints are required:

- hits formed in more than 4 Cedar sectors;
- time coincidence within 2 ns between the track in the the STRAWs and the signal in the Cedar detector.

The track time is selected considering the time of the hits associated to the track in the CHOD detector. In this context, the CHOD has been selected due to the good time resolution. When the CHOD is not reconstructed, due to geometrical acceptance or detector inefficiencies, the track time is taken from the spectrometer. Figure 5.2 shows the time difference between the Cedar time measurement and the track time measurement as from the CHOD or STRAW whenever the CHOD time is missing.

A time cut is performed to consider only events with a time difference better than 2 ns.

5.4 Analysis of a $K_{\mu 2}$ sample

Generally speaking, a muon passing through the NA62 detector crosses the RICH detector, charged hodoscopes (New CHOD and CHOD), LKr, hadronic calorimeter and MUV3 detector. In the LKr a muon behaves mostly as a minimum ionizing particle (MIP - with $\frac{E}{P} < 0.2$). This is true also for the hadron calorimeter system (MUV1 and MUV2), where, almost in all cases, a muon leaves an energy deposition smaller than 2 GeV.

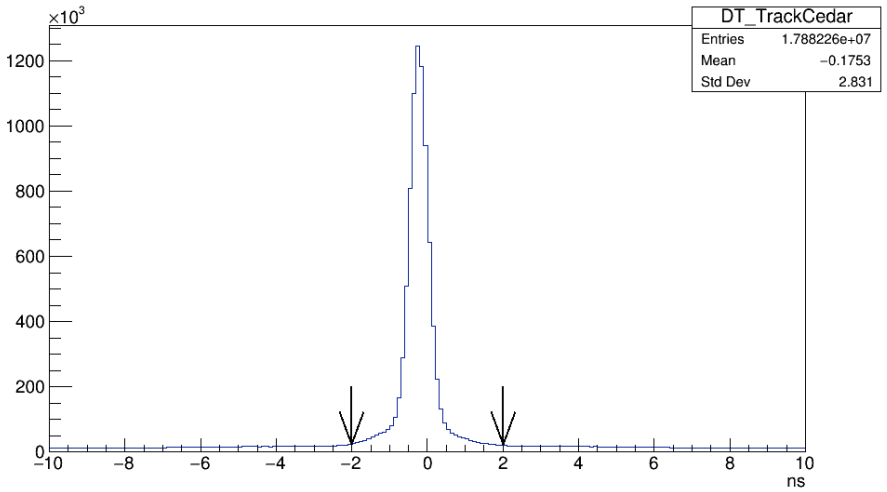


Figure 5.2. Time difference between kaons and tracks of decay products.

To evaluate the response of the different systems to an incoming muon, a $K_{\mu 2}$ sample has been selected. As mentioned above, the selection is specific for different detectors, avoiding to apply cuts based on the detector under study. The starting point for all the system is to require the squared missing mass measured with the spectrometer close to the neutrino mass, considered null. A window around zero takes into account the detector precision: $0.01\ GeV^2 < m^2_{miss_{K-\mu^+}} < 0.01\ GeV^2$. Then, particles in space-time coincidence with the charged track are looked for in the MUV3. The LAV system is required in veto. Finally, a muon-like particle in the LKr is required. The $K_{\mu 2}$ selection is over-constrained and, as stated above, one can measure the efficiency of a given detector by relaxing the requests of the detector under study. For the MUV3 system, in order to compensate the impossibility to use it in the analysis, a muon-like particle in the hadron calorimeter has been required. This allows to purify the sample from a possible pion contamination. The results are reported in the next sub-sections.

RICH

Figure 5.3 shows the RICH efficiency as a function of momentum for muons at 35% of the nominal beam intensity. The coincidence window with respect the time of the CHOD primitive has been set at $\pm$ 10 ns. The momentum threshold required to generate Cherenkov light is well in agreement within the value theoretically calculated for muons in neon at atmospheric pressure, namely P = 9.49 GeV/c.

The efficiency curve turns on sharply around the threshold, reaching a plateau around P $\approx$ 15 GeV. In order to quantitatively evaluate the efficiency as a function of momentum the

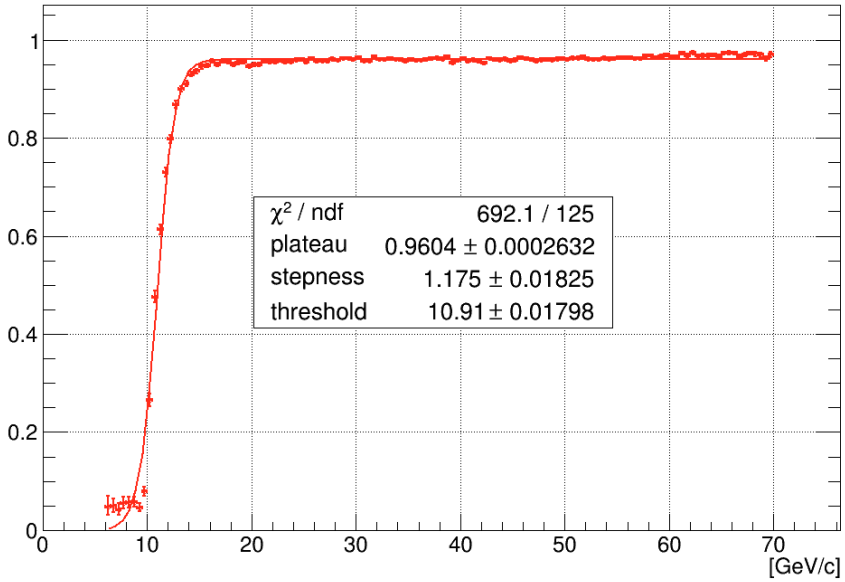


Figure 5.3. Efficiency of RICH primitive generation as measured from a $K_{\mu 2}$ sample.

efficiency was fitted with the Fermi function $f(P)$:

$$f(P) = \frac{a}{1 + e^{b(c-P)}} \tag{5.1}$$

where a indicates the plateau value of the efficiency, b the steepness of the turn-on slope and c is the threshold value. The fit yields an efficiency of a=0.9604 $\pm$ 0.0002, for muon momenta higher than 15 GeV/c. The Cherenkov threshold, defined as the momentum value corresponding to 50% efficiency, measured from data is c=10.91 $\pm$ 0.018 GeV/c while the steepness is b=1.175 $\pm$ 0.018 (GeV/c)$^{-1}$.

In order to find the inefficiency of the primitive-generator firmware as distinct from the detector inefficiency, a specific analysis was performed based on the different requirements to define a primitive and a particle. Given that a threshold on the number of photons in time coincidence is implemented in the RICH primitive generator firmware, the detector generates primitives only if 3 or more hits are produced. As previously described, the hits are not processed directly from the PMT readout, but they are processed by a dedicated TEL62 board which add different channels into a SuperCell. Figure 5.4 shows the number of hits counted by the SuperCells versus the number counted by the PMT read out.

In spite of figure 5.4 showing a strong correlation between the two systems, looking at the events without a primitive, we find that 41.6% of them have less of 3 hits in the SuperCell read-out (as shown in figure 5.5), so that a primitive cannot be generated. It means that about

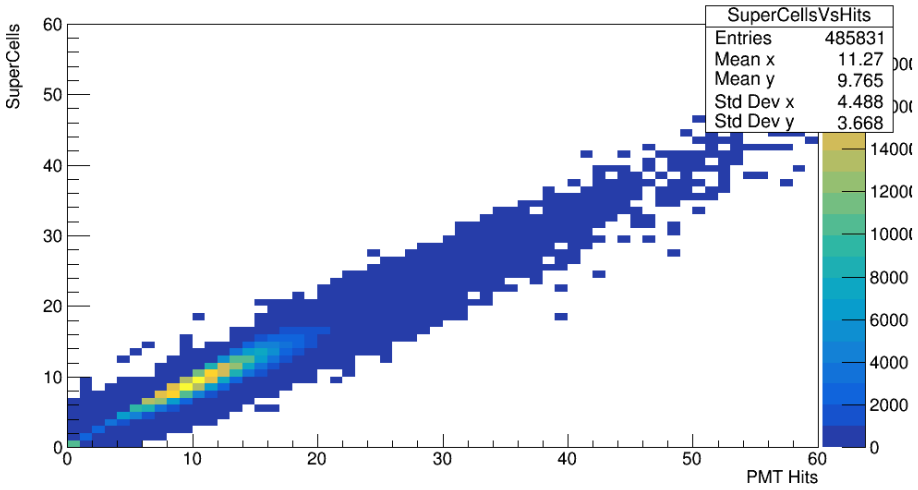


Figure 5.4. Number of hits in SuperCells versus number of hits counted by PMTs.

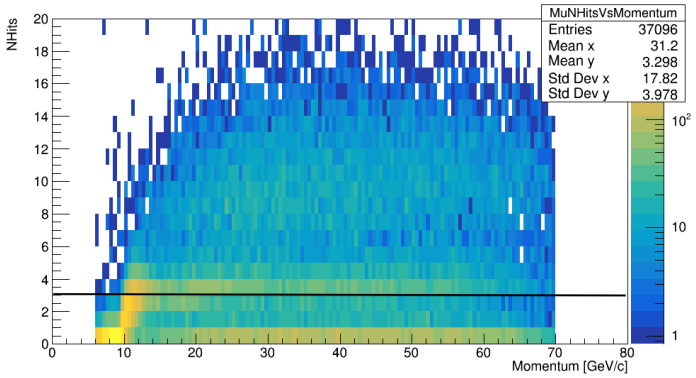


Figure 5.5. Number of hits in the SuperCell-read out system for the events without any primitive associated.

1.64% of the global inefficiency is related to SuperCell-read-out system, while the remaining 2.36% of inefficiency is originated by the L0 trigger chain.

To check the spatial distribution, the plot of the track projected in the XY plane has been reported in figure 5.6. Two regions, enclosed by squares, show higher inefficiency as compared to the other regions of the detector, due to the geometrical complexity of the RICH, requiring for the two central mirrors a read out split into two different boards rather than one. As a consequence, the number of hits that generates a SuperCells is, on average, the half respect to the other parts of the detector, resulting in a smaller efficiency of the primitive generator.

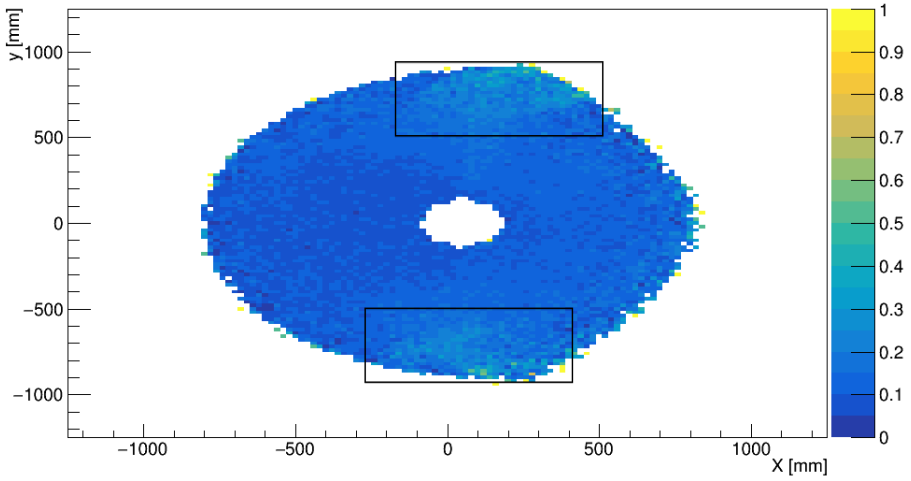


Figure 5.6. Map of the $K_{\mu 2}$ tracks without associated RICH primitives

In summary one can conclude that the RICH inefficiency measured ad 35% of the nominal beam intensity for $K_{\mu 2}$ events is 4%, of which 41% is due to the SuperCell read out and the 59% to the trigger chain.

LAV12

LAV12 is the last station of the LAV detector. The twelfth station is the only one included in the L0 trigger during the 2016 data taking. It has been designed to veto particles generated at large angle with respect to the beam direction.
Looking at the distribution of the energy deposited per LAV station in a $K_{\mu 2}$ events (figure 5.7), the energy deposited in LAV12 is much larger than in the other modules, justifying the choice of including only the last station in the L0 trigger. For the $K_{\mu 2}$ trigger the LAV12 is in veto with a window width of $\pm$ 10 ns around the RICH time.

Figure 5.8 shows the distribution of the random veto due to a primitive of LAV12 as a function of the muon momentum. A small enhancement is clearly visible in the range of 15 $\div$ 25 GeV/c. This is due to the fact that the corresponding angle of particles in this range of momentum belongs to tracks in the LAV12 acceptance.

It can be better visualized by plotting the tracks measured by the spectrometer after the magnet with a LAV12 primitive associated (figure 5.9). The half moon where the acceptance of the LAV is maximum is clearly evident.

At 35% of the nominal beam intensity the random veto due to LAV12 is 1.10% for momenta larger than 25 GeV/c. The figure climbs to 1.6% for momenta between 8 and 25 GeV/c. The increment in this momentum range is likely due to a small overlap in the

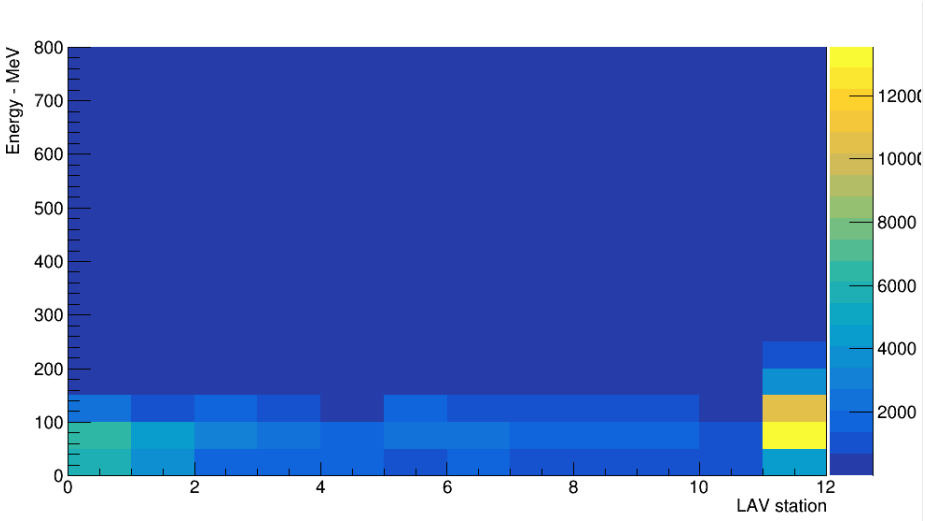


Figure 5.7. Energy deposited in the LAV stations for the $K_{\mu 2}$ sample

.

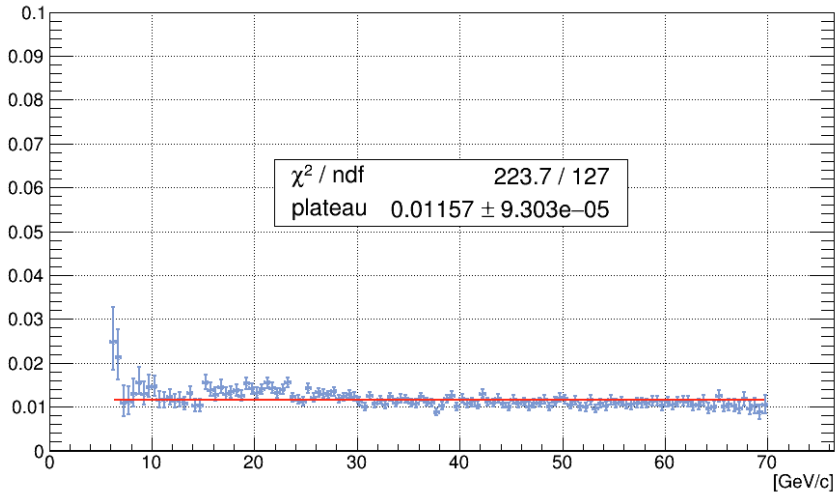


Figure 5.8. Efficiency of LAV12 primitive generation as measured from a $K_{\mu 2}$ sample

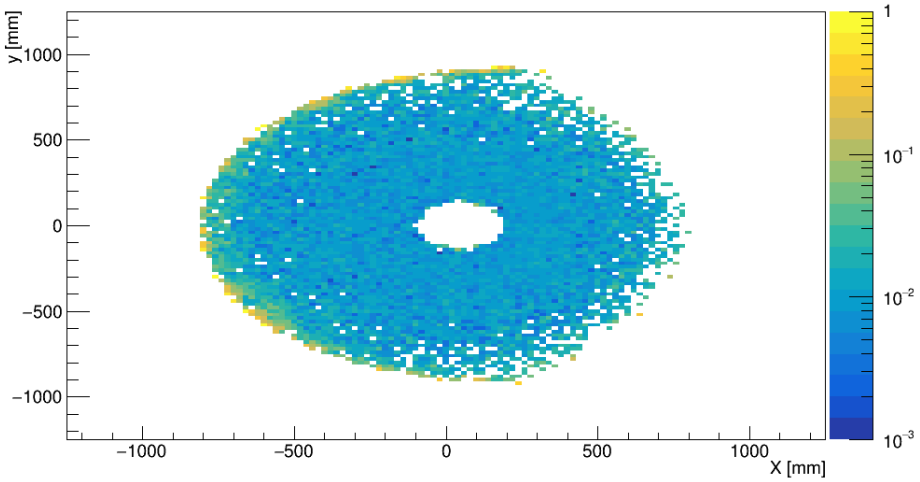


Figure 5.9. Map of the $K_{\mu 2}$ tracks with a LAV primitive associated

acceptance. Nevertheless we are interested in the mean value of the primitive production percentage in function of momentum. Because the increase factor between 15 and 25 GeV/c is small, the distribution has been fitted with a flat function.

The result obtained by the global fit is 1.16% of events with a primitive in time from LAV12.

New CHOD

The main purpose of the NA62 New Charged Hodoscope is to provide a basic element for the L0 trigger when at least one charged particle crosses the scintillator tiles. The evaluation of the efficiency in triggering muons is relevant due to the fact that, from the point of view of the New CHOD, a $K_{\mu 2}$ decay behaves exactly as $\pi\nu\bar{\nu}$ decay. In particular they are characterized by only one signal in a single quadrant, whose released ionization corresponds to a minimum ionizing particle. The efficiency of muon trigger has been evaluated in 4 different conditions:

- without any particular request on multiplicity and geometry;
- by vetoing events with hits in two opposite quadrants;
- by vetoing events with multiplicity greater than 6 hits (UTMC);
- by vetoing events with multiplicity greater than 6 hits and events with hits in two opposite quadrants.

Figure 5.10 shows the NewCHOD efficiency for generating primitive without any request either on the geometry of the decay nor on the hit multiplicity (black markers), applying

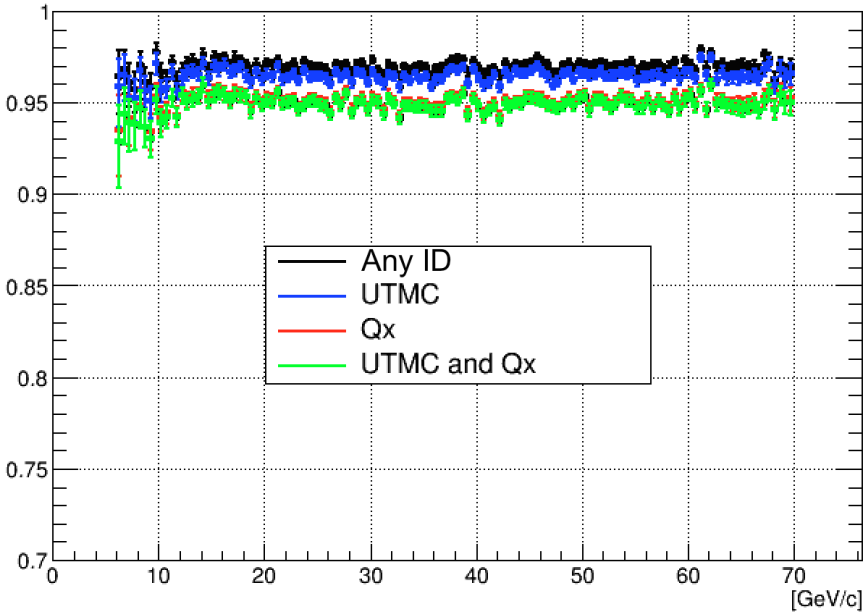


Figure 5.10. Efficiency of NewCHOD primitive generation as measured from a $K_{\mu 2}$ sample at 35% of the nominal beam intensity.

Qx (red markers), applying the cut on the multiplicity (blue markers) and applying all cuts (green markers). The hit multiplicity for a $K_{\mu 2}$ does not contribute significantly because of the presence of only one charged track. In table 5.1 the results obtained at 35% of the nominal beam intensity are reported, for the different conditions applied on the NewCHOD. The time coincidence window between New CHOD and RICH has been set at ± 10 ns.

Configuration	**Value**
All sample	0.970 ± 0.001
UTMC	0.965 ± 0.001
Qx	0.950 ± 0.001
Qx and UTMC	0.948 ± 0.001

Table 5.1 New CHOD efficiency.

When Qx condition is satisfied it means that at least two charged tracks pass through the New CHOD detector, generating more hits. This explains the very small decrease observed when applying the UTMC cut after Qx. A two tracks event can be included in the $K_{\mu 2}$ selection due to the background from the muon halo, which can generate an hit in the coincidence time window of the primitives, overlapping the $K_{\mu 2}$ signal.

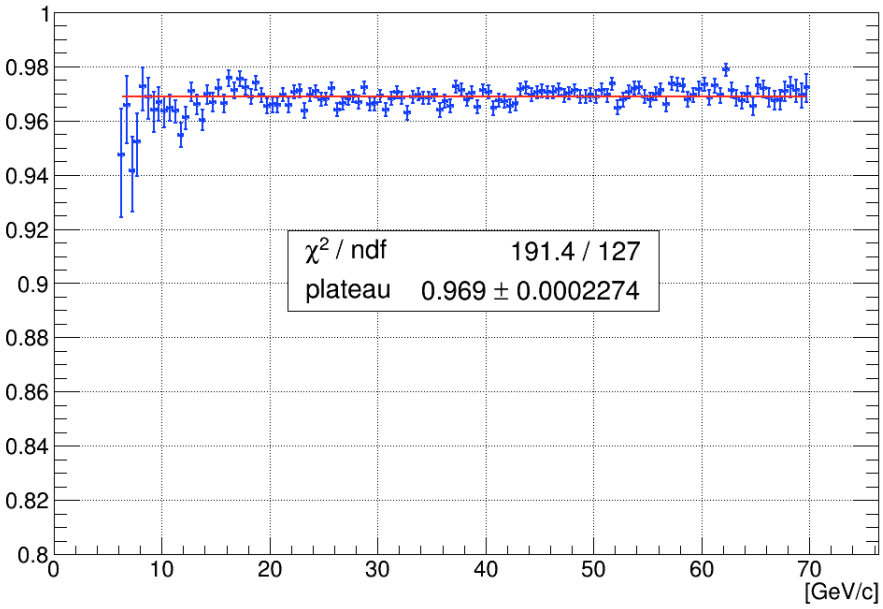


Figure 5.11. Efficiency of MUV3 primitive generation as measured from a $K_{\mu 2}$ sample at 35% of the nominal beam intensity

MUV3

The MUV3 detector is the most important muon-veto at L0 and it is used to reject muons at the first level of the trigger. It has to veto kaon decays into muons, but also the muon halo, which includes:

- muons generated at the target;
- muons generated by beam interactions with material;
- muons from decays of beam particles ($\pi^+ \to \mu\nu$);

Considering kaon branching ratios and the muons of the halo, MUV3 vetoes more than 70% of the events.
Figure 5.11 shows the MUV3 primitive efficiency for the selected $K_{\mu 2}$ sample, as a function of the muon momentum, fitted with a flat distribution to quantify the efficiency. The result of the fit shows an efficiency in detecting muons of 96.9%. In other terms, being the MUV3 set as veto in the $\pi\nu\bar{\nu}$ trigger mask, the rejection power is about 33.

In order to account for the efficiency of the detector as distinct from the one of the data acquisition system, the same analysis has been processed requiring explicitly a muon hit associated in space-time with the track. In this second case, the efficiency rises to 0.979 ± 0.001.

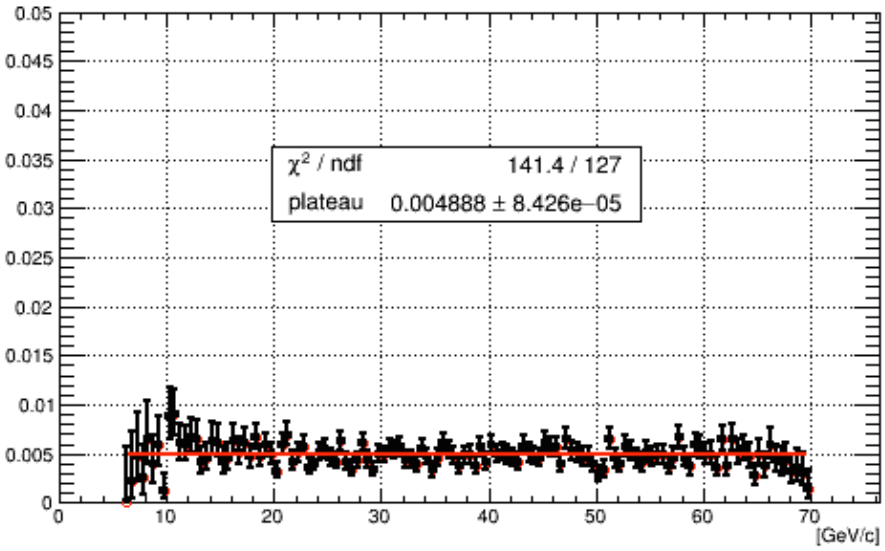


Figure 5.12. Probability of a primitive generated by LKr as measured from $K_{\mu 2}$ sample

We can conclude that about the 33% of the global inefficiency of the MUV3 in detecting the sample of $K_{\mu 2}$ selected is intrinsic to the detector, while about the 66% is due to the inefficiency of the L0 chain.

Calorimeters

In 2016 data taking the calorimeter configuration has been set to send a primitive when the energy deposited in the LKr is larger than 20 GeV, in order to veto the π^0 mainly coming from $K^+ \to \pi^+\pi^0$. The total energy deposited in the LKr for a muon is normally less than 2 GeV. For this reason primitives from $K_{\mu 2}$ events are not expected except for those events with secondaries generated by catastrophic Bremsstrahlung or events with a radiative gamma. The probability to have a catastrophic bremsstrahlung event in the LKr is of the order of 10^{-5}, negligible for this kind of analysis, while the $K^+ \to \mu^+\nu\gamma$, is characterized by a branching ratio of about 6.2 $\times 10^{-3}$. Requiring the MUV3 as positive in the data selection, we are also excluding from the sample all the kaons and pions decaying into positrons before reaching the calorimeter.

In order to select muons at the first trigger level, such events can be considered as veto, and we quoted the value considering a time coincidence window between the primitives from RICH and LKr of $\pm$ 10 ns. Figure 5.12 shows the result obtained with a flat function fit. The probability to have one primitive generated by the LKr in time with our muon selection is about 0.49%.

$K_{\mu 2}$ Analysis Conclusions

For the detectors which are relevant for the muon rejection (namely CHOD, RICH and MUV3), a check of the stability in the L0 efficiency has been performed considering different runs. In particular three different intensities of the beam have been chosen to study the dependency of the efficiency as a function of the intensity.

As discussed above, the quality of the data is strongly dependent on the beam conditions. Already at 35% of the intensity, some instabilities can affect the data taking. A catastrophic example of how the instability of the data acquisition can affect the efficiency can be seen when the instantaneous rate goes higher than 14 MHz, with the gigabit ethernet stopping its data stream. As a consequence, the primitives related to such events are lost, increasing immediately the inefficiency. In order to exclude the bursts affected by these anomalies, a cut on the number of primitives sent by each sub-detector has been applied, requiring that the ratio between the number of primitives sent and the argonium counts is within a window of $\pm$ 20% with respect to the mean value.

Figure 5.13 shows the analysis outputs for different runs performed at 35% of the nominal intensity. The differences between the various results are within 0.5%. Figure 5.14 shows instead the results obtained for the different values of intensity selected: 5%, 17% and 35% of the nominal one. The LKr cannot be included due to different firmware versions being used at different intensities. A dependence on the intensity is noticeable: a decrease of the efficiency of 1.2% affects the RICH going from 5% of beam intensity to 35%, less than 1% affects the MUV3, and 1.8% the New CHOD.

Being the run still ongoing during the writing of this thesis, a quality reprocessing for the 2016 data has not been performed yet. A more stable sample of data will be ready at the end of the data taking period.

Having studied the individual contributions of the detectors comprised in the trigger when prompted with $K_{\mu 2}$ decays, one can finally derive the rejection capability of L0 for such events.

The $K_{\mu 2}$ events that remain after the L0 selection due to the inefficiencies of detectors can be evaluated considering the responses of RICH, New CHOD and MUV3 systems as mutually independent. At 35% of the nominal beam intensity we have that the L0 selection for $K_{\mu 2}$ decays is given by:

$$K_{\mu 2}\ after\ L0 = (\varepsilon_{RICH} * \varepsilon_{NewCHOD} * \varepsilon_{MUV3}) \simeq 92\%. \tag{5.2}$$

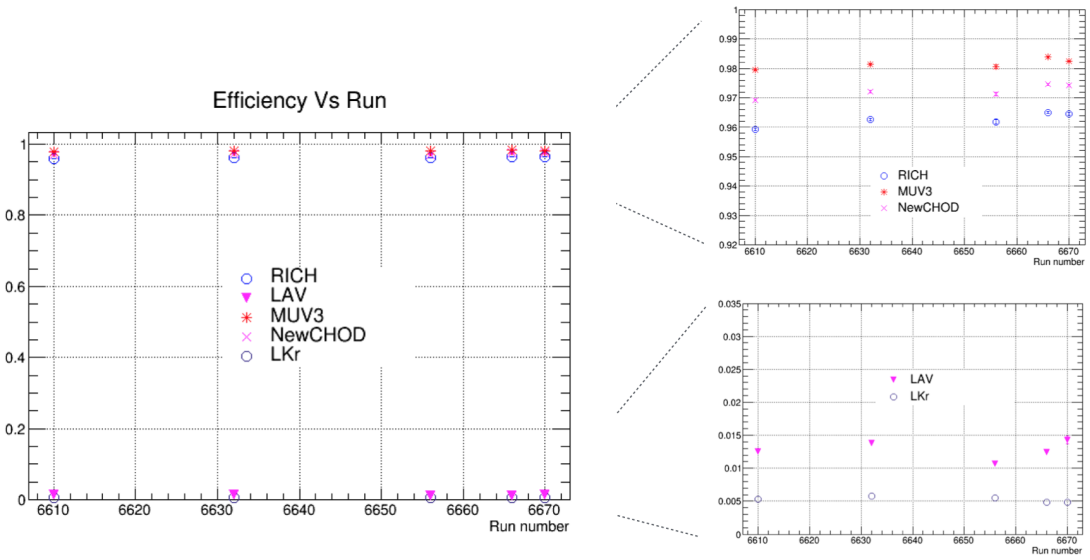


Figure 5.13. Efficiency values obtained for different runs at 35% of the nominal intensity.

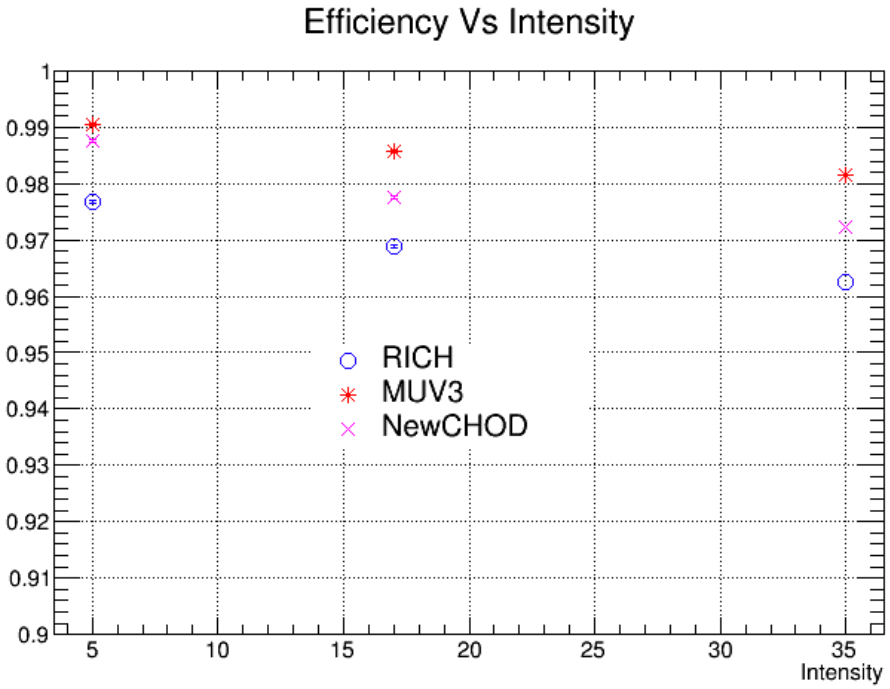


Figure 5.14. Efficiency values obtained for different intensirties of the beam.

The uncertainty is mainly due to systematic effects related to the average beam intensity and the fluctuations. The instantaneous value affects the efficiency more severely as the average is higher.

Depending on the trigger mask, the random veto due to LAV12 and Calorimeter should be accounted for. For both detectors the contribution is almost stable at 1% in the entire range of momentum considered. However, the L0 rejection power against $K_{\mu 2}$ decays is not affected by the random veto and one can conclude that 92% of the $K_{\mu 2}$ decays are detected at the L0 trigger and rejected. This value takes into account the efficiency of the L0TP in making the coincidences between the different sources. It has been calculated in chapter 4, and the result obtained is 0.88%. Another 1% is coming from the intrinsic MUV3 efficiency, which depends on the detector behavior.

5.5 Analysis of a $K_{\pi 2}$ sample

Following a very similar procedure, we can now focus our attention on the detector behavior when a $K_{\pi 2}$ event occurs. The analysis of the $K_{\pi 2}$ sample has two main purposes: a) to determine the L0 trigger efficiency in order to recognize a pion single track (π^+) in the momentum range between 15 and 35 GeV/c; b) to determine the rejection power of $\pi^+\pi^0$ decays. As discussed above, $K_{\pi 2}$ is the second source of background after $K_{\mu 2}$, in terms of branching ratio. The difference respect to the main signal $K^+ \to \pi^+ \nu\bar{\nu}$ is the presence of the π^0. Being the mean life-time of the π^0 equal to 8.4×10^{-17} seconds, it decays immediately in gammas or electrons, with the branching ratios reported in table 5.1

Decay Type	**Branching ratio**
$\pi^0 \to \gamma\gamma$	0.98823
$\pi^0 \to \gamma e^+ e^-$	0.01174
$\pi^0 \to e^+ e^- e^+ e^-$	3.34×10^{-5}

Table 5.2 Main branching ratio of π^0

The presence of gammas or electrons in the $K_{\pi 2}$ sample can be easily detected, allowing to reject these events while keeping a good efficiency for the signal. In particular

- Electrons are detected by the New CHOD together with the π^+;
- if photons or electrons are accepted by the LKr detector, they release extra-energy as compared to events with only π^+s;
- low energy gammas produce showers in LAV12;

- high energy gamma rays are detected by the small angle veto systems (IRC/SAC);

Moreover gamma rays passing through the RICH mirrors can generate electron-positron pairs, which are again vetoed by the hodoscope. Events with electromagnetic showers in the LKr are rejected at the first trigger level by requiring the energy deposited in the detector to be above a certain threshold.

For all the systems, the one-track selection has been applied. Except for the LKr analysis, the selection of $K_{\pi 2}$ has been performed by reconstructing the π^0 in the LKr: after a pre-selection of good electromagnetic clusters, the π^0 vertex and 4-momentum are obtained by using only the positions and energies of the two clusters.

The vertex is reconstructed as

$$Z_{Vertex} = Z_{LKr} - d_{1,2}\frac{\sqrt{E_{\gamma_1}E_{\gamma_2}}}{m_{\pi^0}} \tag{5.3}$$

where Z_{LKr} is the position of the LKr, E_{γ_1}, E_{γ_2} are the energies of the two electromagnetic clusters, and $d_{1,2}$ is the distance between the two clusters. The π^0 4-momentum is obtained by summing the 4-momenta of the two photons. The energy component of the 4-momentum is already known from the cluster energy, while the direction of the 3-momentum for each gamma is found from the positions of the clusters relative to the decay vertex. An additional cut on the vertex of the π^0 is applied, such that it is inside the fiducial region (110 m $< Z_{Vertex} <$ 170 m). Moreover, the invariant mass of the two pions is expected to be equal to the kaon mass:

$$(P_{\pi^+} + P_{\pi^0})^2 = P_{K^+}^2 = M_{K^+}^2 \tag{5.4}$$

To account for the experimental errors the kaon mass has to be in the region $480MeV < M_{K^+}^2 < 510MeV$. The strong request to have both gammas from a π^0 in the LKr acceptance limits the statistics to events with π^+ momentum lower than 50 GeV/c. Events with higher π^+ momentum, in fact, have low energy gammas, that are outside the LKr acceptance, or their energy is too low to reconstruct a π^0.
Depending on the system under study, additional cuts are applied. They include the squared missing mass measured with the spectrometer to be compatible with the π^0 mass, no hits in time in the MUV3 or in the LAVs detectors and the association between the charged track of the π^+ with a cluster in the LKr.

When studying the Calorimeter performance, the π^0 identification step cannot be applied, as it requires the LKr cluster reconstruction. Thus, a different selection has been applied. The first requirement is the squared missing mass close to the π^0 mass:

$$M_{\pi^0}^2 - 0.01\ GeV^2 < m_{miss_{K-\pi^+}}^2 < M_{\pi^0}^2 + 0.01\ GeV^2. \tag{5.5}$$

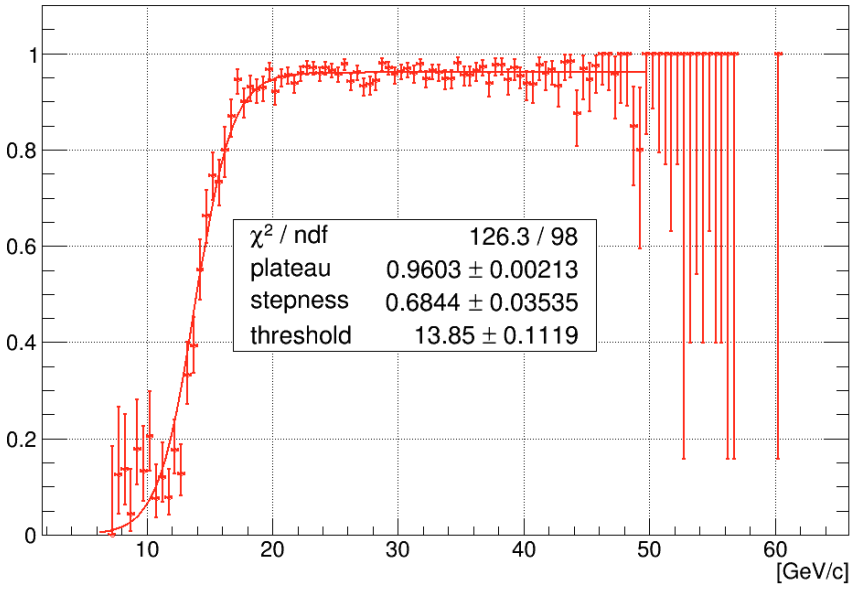


Figure 5.15. Efficiency of RICH primitive generation as measured from a $K_{\mu 2}$ sample.

In addition, LAVs and MUV3 are required in veto.

RICH

In case of pions passing through the RICH detector, the theoretical threshold of momentum required to generate Cherenkov light is $P_t = 12.53$ GeV/c. Thus, in the range between 0 and 12.5 GeV/c the RICH is sub-threshold being unable to detect π^+s.

The distribution of the RICH efficiency in generating primitives is reported in figure 5.15. The coincidence window between CHOD and RICH primitive has been set, as for muons, equal to $\pm$ 10 ns. The distribution has been fit with the Fermi function (5.4). Results show a plateau value with an efficiency of 0.960, with a measured threshold of 13.85 GeV/c. Comparing the RICH performances as from $K_{\mu 2}$ and $K_{\pi 2}$ sample in the same run, it is possible to see (figure 5.16) a good agreement of the efficiency in the region of momentum between 15 and 50 GeV/c.

LAV12

LAV12 is designed to veto large angle photons, from 8.5 mrad up to 50 mrad. Photons generated by π^0 decays can be emitted inside the LKr acceptance, or one of the two can be inside the LAV acceptance. By taking the events with a π^0 fully reconstructed in the LKr, it is

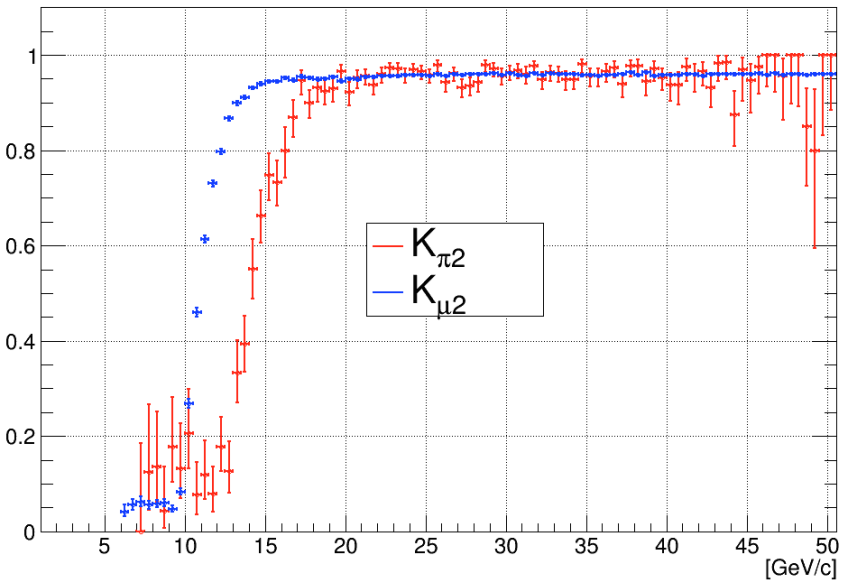


Figure 5.16. Comparison between RICH efficiencies as extracted from a $K_{\mu 2}$ and a $K_{\pi 2}$ sample in the region of momentum $15 \div 50$ GeV/c.

possible to evaluate the probability of random veto of the detector. The coincidence window between the trigger time and the primitive from LAV has been chosen, as for muons, at $\pm$ 10 ns. This is shown in figure 5.17, which reports the distribution of the probability to have a primitive from LAV when the $K_{\pi 2}$ has been selected with both gammas of the π^0 in the LKr. The probability of random veto has been evaluated in the region of momentum between 15 and 35 GeV/c, which is the region foreseen for the $\pi\nu\bar{\nu}$ analysis. The result obtained shows that 1.95% of the events have a LAV primitive in time, higher than the random veto rate quoted for the $K_{\mu 2}$ sample. This difference is still not completely understood. Nevertheless it is definitely true that the pion sample can be affected by delta rays produced in the crossing materials before the LAV12 or by photons generated by external bremsstrahlung releasing enough energy, and thus generating a primitive.

NewCHOD

Generally speaking, the behavior of the NewCHOD for $K_{\pi 2}$ is very similar to the one obtained for $K_{\mu 2}$, due to the presence of only one charged track. Nevertheless, the π^0 is expected to influence the topological and multiplicity responses (Qx and UTMC PrimitiveIDs). In particular, gamma rays generated from the π^0 decay can convert into pairs passing through the RICH mirror. The RICH mirror thickness is 25 mm, which is almost 1/4 of the mean free

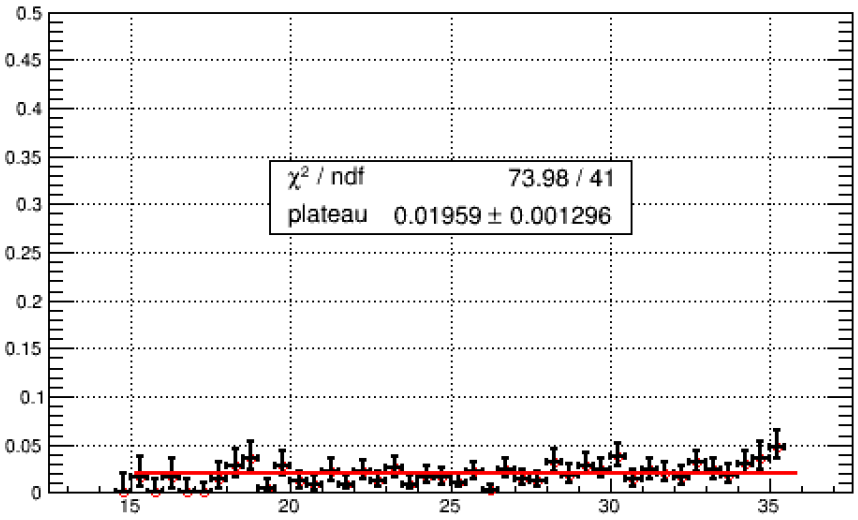


Figure 5.17. LAV random veto for $\pi^+\pi^0$ events with a π^0 well reconstructed in the LKr

path of a gamma in glass. With two photons crossing the mirrors, the probability to have at least one gamma converted is roughly 50%, thus increasing the complexity of the event topology. Once fixed a time window for the coincidence a primitive at $\pm$ 10 ns, the efficiency of NewCHOD to generate primitives with any ID is reported in figure 5.18. The efficiency has been calculated as a function of momentum. The result obtained by fitting the values with a flat distribution returns an average efficiency of 96.7%. Then, the requirement of events with a number of hits lower than 6 (UTMC request) has been applied. This is a quite weak cut, the efficiency decreasing from 96.7% to 94.2%. Requiring instead a Qx topology in veto, the number of primitives from the $K_{\pi 2}$ sample passing the condition decreases to 60.6%. This confirms that the impact of the photon conversion is large. Finally, applying both Qx and UTMC in veto, the number of events passing the cuts is 60.5 %.

In table 5.3 the results obtained at 35% of the nominal beam intensity are summarized, for the different conditions. As expected, the percentage of events detected without any topological request is the same for muon and pion sample.

Configuration	Value
any ID	0.967 ± 0.002
UTMC	0.942 ± 0.003
!Qx	0.606 ± 0.006
!Qx and UTMC	0.605 ± 0.007

Table 5.3 New CHOD efficiency.

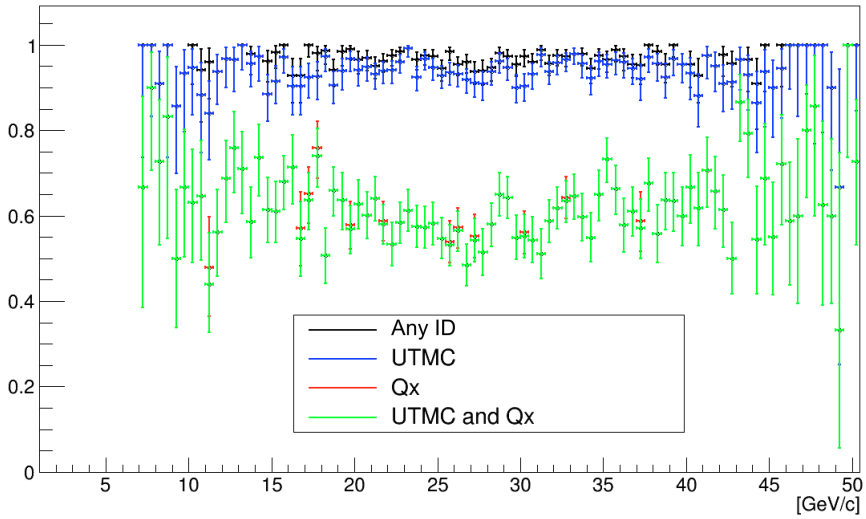


Figure 5.18. Efficiency of NewCHOD primitive generation as measured from a $K_{\pi 2}$ sample at 35% of the nominal beam intensity

MUV3

For a pure sample of $\pi^+\pi^0$, there are three main effects that generate a signal in the MUV3. The first is due to the π^+ decay into $\mu\nu$. Then, the beam muon halo can generate hits in the MUV3. Finally there is a probability for a pion to be not completely absorbed by the hadron calorimeter, thereby being detected by MUV3. In general, the probability to have a MUV3 primitive in the selected $K_{\pi 2}$ sample should be considered as random veto. This random veto will also affect the efficiency of the trigger on the main signal, being the presence of the π^0 completely irrelevant. As figure 5.19 shows, about 6% of the $\pi^+\pi^0$ have a primitive in time generated from the MUV3 detector. This result has been obtained at 35% of the intensity.

It is not straightforward how to separate the three contributions. In fact the beam muon halo it is not random-generated and could be correlated to the beam particles, so that an estimate of the contribution is not easily to evaluate. The same for the leakage: to evaluate such probability would require a complete simulation of the shower development through the calorimeters with weak reliability for the very end of the shower. On the other hand the contribution due to the π^+ into muon decay can be estimated.

Considering the mean lifetime of a pion being equal to 26 ns, by boosting it for the Lorentz factor $\gamma \simeq 323$ (taking into account pions with 45 GeV/c momentum) we can calculate how many π^+ decays in the 45 m long fiducial region used for this analysis. One obtains that about 1.6% of pions decay in the fiducial region. This component, which corresponds to 26% of the total, is irreducible. Looking at different runs, a trend as a function of the intensity is

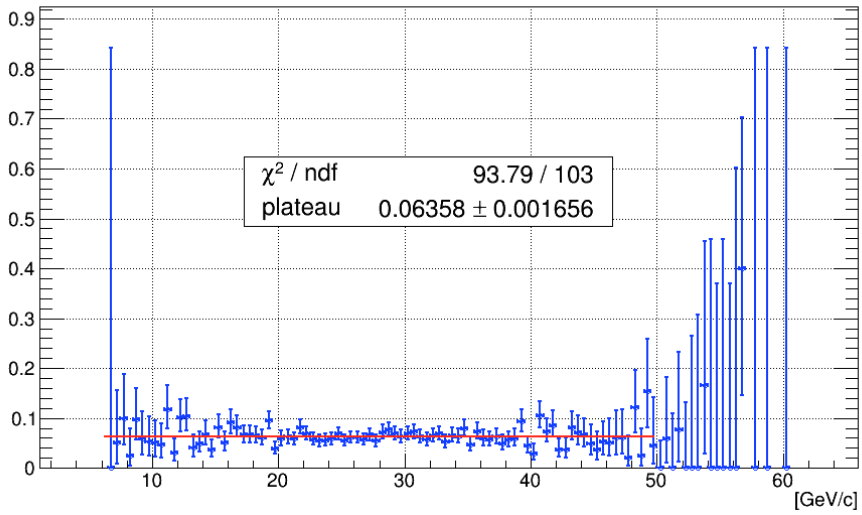


Figure 5.19. Probability to have a primitive from MUV3 in time with the $K_{\pi 2}$ sample.

noticeable, as reported in figure 5.20. In particular it is possible to see that at low intensity (17% of the nominal) the only component to the veto fraction is the $\sim 1.6\%$ coming from the pion decay.

Calorimeters

The calorimeter system is the most relevant veto for events with π^0. In particular in 2016 data taking, only LKr has been used in the trigger masks. The selection of the sample requires the LAVs being in veto in order to maximize events with the gammas in the LKr acceptance. From the $K_{\pi 2}$ reconstructed sample, the analysis has been performed considering the primitives generated from the calorimetric system in a window of $\pm$ 20 ns with respect to the CHOD primitive time. Since the criteria to generate a primitive are based on the energy deposited in the detector, the efficiency has been evaluated as a function of the energy, obtained by counting the number of events with an associated LKr primitive as compared to the events in the $K_{\pi 2}$ sample.

Figure 5.21 shows the result of the analysis. The plot can be fit with the Fermi function. The fit gives an efficiency for the events on the plateau of 97.8% with a threshold around 22.2 GeV. The steepness in this case is 0.42 $(\text{GeV}/c)^{-1}$. The same analysis can be performed requiring the small angle detectors, IRC and SAC, in veto. The efficiency in this second case grows up to 98.2%, losing statistics of events with low energy in the LKr. In this case in fact both gammas are inside the LKr, increasing the energy deposited. Finally, it is possible to repeat the analysis by requiring explicitly a cluster associated with the track as measured by

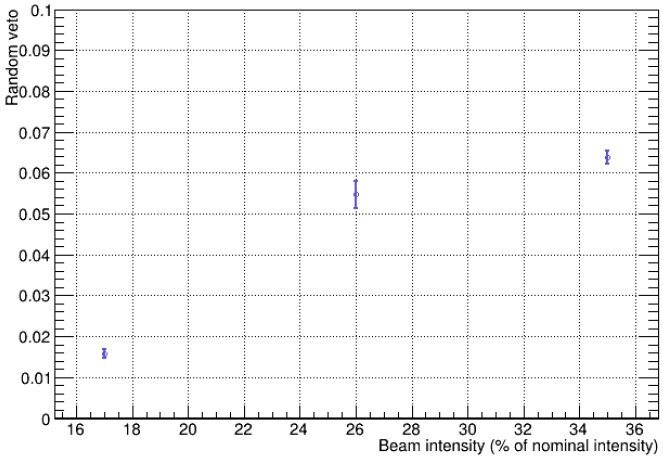


Figure 5.20. MUV3 veto as a function of the beam intensity

the STRAW spectrometer. With this requirement one can factorize the impact of the detector from the primitive generation inefficiency. With this selection we find again an efficiency of 98.2% on the plateau, equal to the one obtained requiring the IRC/SAC in veto. We can conclude that the trigger efficiency of the Calorimetric L0 system for a sample of $K_{\pi 2}$ with two gammas in the LKr acceptance and with a deposited energy greater than 20 GeV is 98.2%. Without requiring the IRC/SAC in veto, and therefore allowing for a gamma in the small angle veto detectors, the efficiency is 97.8%.

Changing the time coincidence window between RICH and Calorimeter primitives from $\pm$ 20 ns to $\pm$ 10 ns, the efficiency drops to 88.1% for events with energy higher than 20 GeV. This is due to the firmware used in the last period of the data taking was affected by a known issue, namely generating primitives with two different temporal distributions with respect to the RICH, one peaked at zero, the second peaked at +12.5 ns, as show in figure 5.22.

So far for what concerns the $K_{\pi 2}$ rejection power at the L0 trigger level. However, by applying the cut ELKr > 20 GeV there is a not negligible probability for the π^+ signal to release an energy larger than the cut. An investigation related to the efficiency of the LKr thus has been performed. The question is then: given a π^+ in the momentum range between 15 and 35 GeV/c, what is the probability that the LKr vetoes it?

To evaluate the veto effect of the Calorimeter system on the signal, one can select events with a sigle pion in the LKr and then divide the signal region in two parts: 1) a region with π^+ momentum between 15 GeV/c and 35 GeV/c and energy in the LKr lower than 20 GeV; 2) a region with π^+ momentum between 15 GeV/c and 35 GeV/c and energy in the LKr higher than 20 GeV. In the region 1) the energy in the LKr is sub-threshold and no primitives should be produced. Nevertheless, it could happen that a single pion releases more than 20

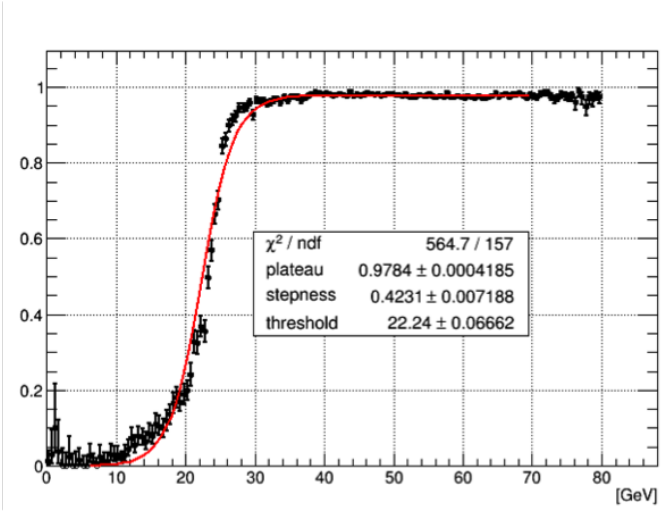


Figure 5.21. Calorimeter efficiency for the $K_{\pi 2}$ sample

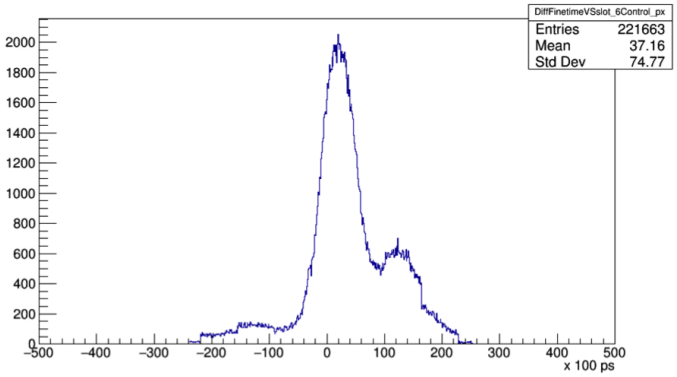
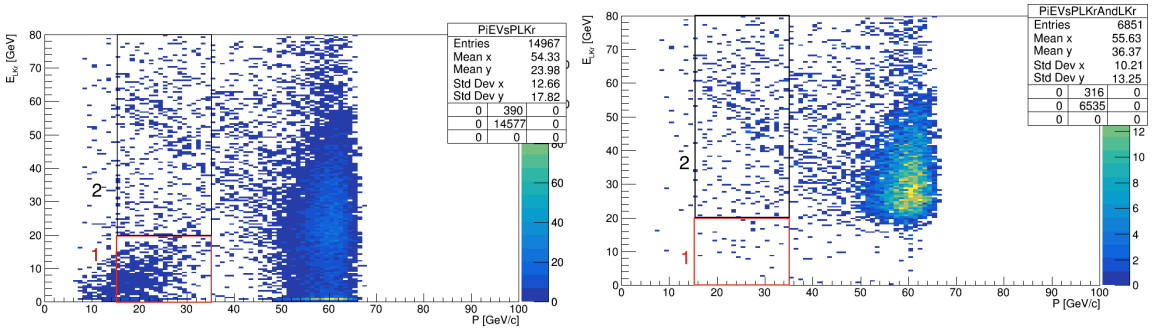


Figure 5.22. Time difference between RICH and LKr primitives: two peaks are noticeable in the distribution, one peaked at 0, the other peaked at +12.5 ns. The issue is under study.



(a) All events.

(b) Events with a primitive from the Calorimeter system associated.

Figure 5.23. Energy deposited in the LKr versus π^+ momentum for events with only one cluster in the LKr associated to the track in a time window of $\pm$ 5 ns.

GeV in the LKr. To select a single pion, the explicitly request to have one and only one cluster in the LKr has been done. The cluster has to be in time in a coincidence window of $\pm$ 5 ns with respect to the track time measured with the CHOD. MUV3 is set as veto, but we relaxed the constraint on the LAVs and IRC/SAC, allowing gammas to be emitted at large or small angles. The $K_{\pi 2}$ sample has been selected by requiring the $(K^+ - \pi^+)$ missing mass to be M_{π^0} +/- 0.1 MeV/c^2. A final cut as been applied to exclude electrons from the selection, requiring explicitly the ratio between energy deposited and momentum of the π^+ to be less than 0.9.

Figure 5.23 shows the scatter plot of the energy in the LKr vs the π^+-momentum for all the events (5.23a) and for events with a primitive fired by the Calorimeter system (5.23b). The two squares show the two regions above mentioned.

This means that given a total of 1377 tracks related to pions in the 15 to 35 GeV/c momentum range, 54.7% deposit less than 20 GeV, while the remaining 45.3% deposit more. Taking a time window of 10 ns, events in the region 1) have a primitive in the 2.25% of the cases. Enlarging the time window to 20 ns, the events with a primitive associated become 4.5%. This is due to a bug in the Calorimeter trigger algorithm, which overestimates the energy with respect the offline value.
Considering the events in region 2, they deposit more than 20 GeV in the LKr, so they are expected to be vetoed. The analysis shows that about the 73.3% of the events selected have a primitive associated and are expected to be rejected. This is of course a source of inefficiency at the trigger level, which, given the value of the cut, looks to be not irreducible. Obviously it can be increased by optimizing the value of the cut.

One should remark that the pattern of the energy deposition by hadron in the LKr, which is spread in different cells, makes the online cluster reconstruction more complex.

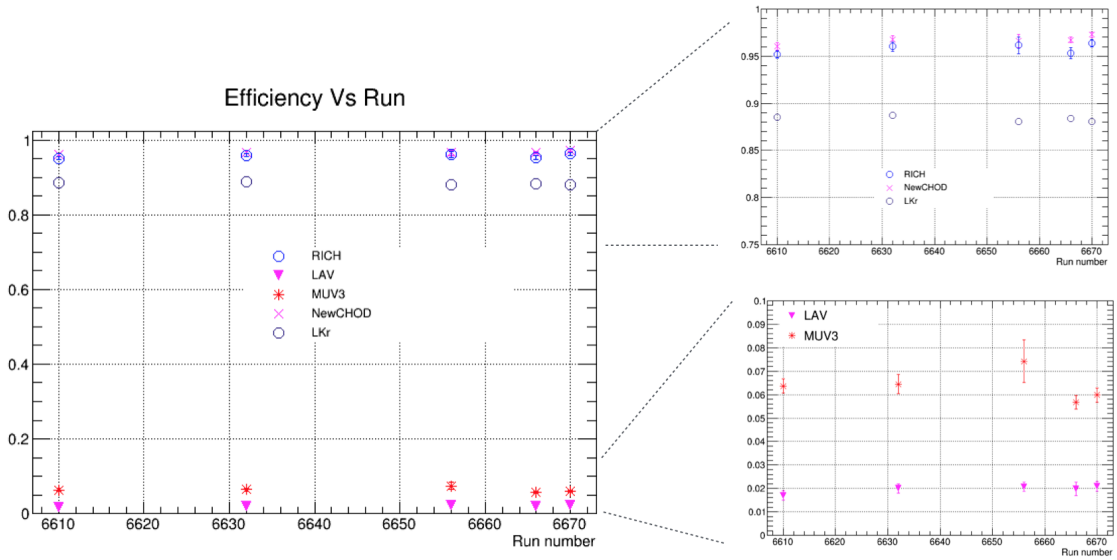


Figure 5.24. Efficiency values obtained for different runs at 35% of the nominal intensity.

The overall cut inefficiency is 13.3%, obtained as the sum of the inefficiency of the Calorimetric system in the two regions weighted by the fraction of tracks in each region.

Table 5.5 summarizes the results obtained for the analysis performed on the Calorimeter system:

Type	$\pm$ 10 ns coinc. window	$\pm$ 20 ns coinc. window
Efficiency (E_{LKr}>20 GeV)	88.1%	97.8%
Efficiency (E_{LKr}>20 GeV - IRC/SAC as veto)	88.7%	98.2%
Veto in the signal region (15 GeV/c<P_{π^+}<35 GeV/c, E_{LKr}<20 GeV)	2.25%	4.5%

Table 5.4 LKr efficiency

5.6 $K_{\pi 2}$ Analysis conclusions

Proceeding in the same way as for the $K_{\mu 2}$ sample, the stability of the $K_{\pi 2}$ analysis has been checked for different runs. The result is reported in figure 5.24. In this plot the Calorimeter system efficiency has been calculated using a coincidence time window with respect to the RICH of $\pm$ 10 ns. The analysis values are highly stable for the different runs considered. In figure 5.25 is reported the efficiency as a function of the intensity for the RICH and NewCHOD. The LKr efficiency cannot be included in the analysis due to the different firmware versions used in the different intensity conditions.

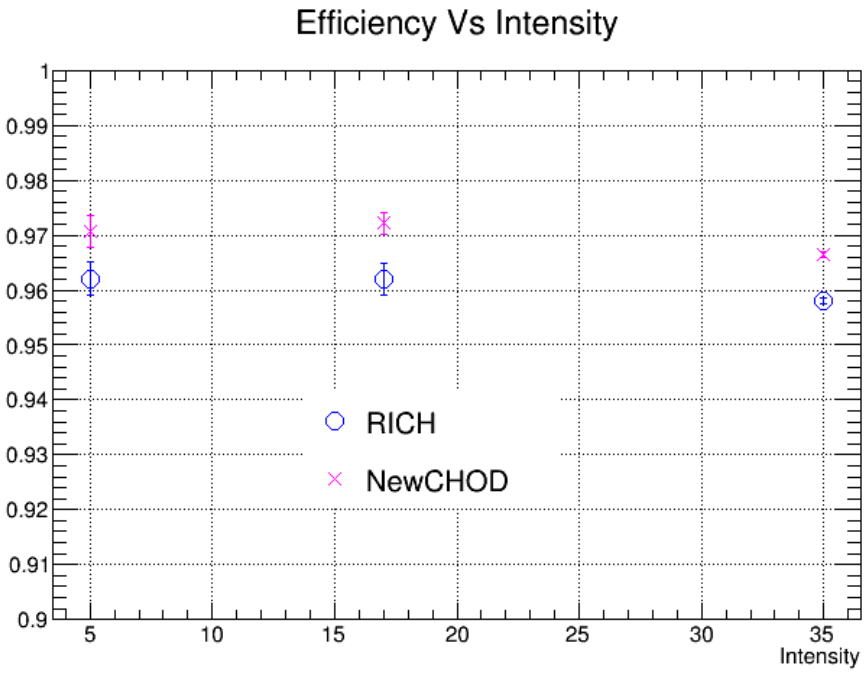


Figure 5.25. Efficiency values obtained for different beam intensity.

Summarizing the results obtained from the $K_{\pi 2}$ selection, we can now estimate the L0-efficiency in detecting a single pion at 35% of the nominal beam intensity in a range of 15 ÷ 35 GeV/c, when the pion release less than 20 GeV in the LKr. The value can be calculated considering the different contribution of detectors as mutually independent, as we did for the muon sample analysis.

$$\varepsilon_{RICH} * \varepsilon_{NewCHOD} * (1 - \varepsilon_{MUV3}) * (1 - \varepsilon_{LAV12}) * (1 - \varepsilon_{Calorimeters}) = \simeq 84\%. \tag{5.6}$$

As for the muon sample, the uncertainty is due to systematic effects related to the beam intensity fluctuations. Nevertheless, this result is in a good agreement with a Monte Carlo study performed in 2014[36] giving an estimation of detection efficiency of about 74.85% for the signal evaluated at 100% of the intensity. The MC studies were performed by adding a cut on the total hadronic calorimeter energy, required to be higher than 8 GeV, and an additional electromagnetic energy cut requiring energy higher than 1.5 GeV, in order to exclude muons from the selection. The random veto component for both the MUV3 and the LAV systems will increase with the increasing of the rate, contributing to reduce the efficiency in the pion detection down to the MC prediction.

Finally, it is possible to calculate the rejection factor for the $K_{\pi 2}$ decay mode. In this case the Calorimeter system takes the part of the MUV3 for the $K_{\mu 2}$, being the most important veto system at the first trigger level.

The result obtained for the rejection is:

$$\varepsilon_{RICH} * \varepsilon_{NewCHOD} * (1 - \varepsilon_{MUV3}) * (1 - \varepsilon_{LAV12}) * (\varepsilon_{Calorimeters}) = \simeq 85.0\%. \quad (5.7)$$

In both cases ($K_{\pi 2}$ rejection and single pion detection), the most critical contribution to the inefficiency is the random veto of the MUV3, due to the large flux of muons coming from kaon and pion decays and from the beam halo.

A new Monte Carlo study where the simulation of the state-of-art of trigger generation algorithms is implemented is under development, driven by results obtained with data collected in 2016. It will be mandatory to validate the future upgrades of the system with the simulation, maximizing the trigger performances for the next data taking.

Conclusions

The NA62 experiment has started the data taking in 2014, with the first pilot run, with a preliminary read out system. A second test run took place in 2015, when the final read out has been used. In 2016 the first physics data taking aims to reach the standard model sensitivity for the ultra rare decay $K^+ \to \pi^+ \nu\bar{\nu}$, whose branching ratio, at level of 10^{-11}, requires a high performing trigger system in order to reject the huge amount of background. The work presented in this thesis focuses on two aspects of the Level Zero Trigger system of the NA62 experiment. The first part of the work concerned the design, development, testing and commissioning of the Level Zero Trigger Processor. The algorithm implemented on the FPGA and the environment set up for the proper trigger operation have been presented, and the capability to achieve an efficient data taking in the challenging environment of very high particle flux of the NA62 beam have been reported. The system has been commissioned and extensively tested during the 2014-2015 runs, and adapted to scale with the increased intensity of the beam for the first physics data taking in 2016. The performance of the L0TP with a beam intensity of about 260×10^6 particles per second (about 35% of the nominal intensity) has been presented. The system almost reached the design input rates (10 MHz from each source), requiring some upgrades for the next run. The L0-output rate in the 2016 run reached $\sim 50\%$ of the nominal value (500 kHz). Smooth data taking was guaranteed by the stability of the system.

The second part of the work focused on the study of the global L0 trigger efficiencies. The sample selection has been performed at 35% of the nominal beam intensity usinf data collected by the L0-control trigger, without higher level triggers applied. The efficiency to detect $K_{\mu 2}$ and $K_{\pi 2}$ events and therefore the capability to reject them have been studied, after obtaining a pure over-constrained selection.

A study of the efficiency in detecting a single pion in a momentum range of $15 \div 35$ GeV/c is also reported.

The present studies also show the limits of the present system and give suggestions for future improvements, in order to be able to sustain the 100% of the beam intensity foreseen to be, $\sim 750 \times 10^6$ particles per second.

References

[1] A. Artamonov *et al.*, "New measurement of the $k^+ \to \pi^+ \nu\bar{\nu}$ branching ratio," *Phys. Rev. Lett.*, vol. 101, p. 191802, 2008.

[2] J. Beringer *et al.*, "Review of particle physics*," *Phys. Rev. D*, vol. 86, p. 010001, Jul 2012.

[3] A. Buras, D. Buttazzo, and R. Knegjens, "$k^+ \to \pi\nu\bar{\nu}$ and $\frac{\varepsilon'}{\varepsilon}$ in simplified new physics models," *JHEP*, vol. 1511, p. 166, 2015.

[4] K. Olive and al., "Particle data group," *Chin. Phys. C*, vol. 38, p. 090001, 2014.

[5] F. Mescia and C. Smith, "Improved estimates of rare k decay matrix elements from $K_{\ell 3}$ decays," *Phys. Rev. D*, vol. 76, p. 034017, Aug 2007.

[6] T. Inami and C. Lim *Prog. Theor. Phys.*, vol. 1772(E), p. 65, 1981.

[7] M. Misiak and J. Urban, "Qcd corrections to fcnc decays mediated by z-penguins and w-boxes," *Phys. Lett. B*, vol. 451, p. 161, 1999.

[8] J. Brod, M. Gorbahn, and E. Stamou, "Two-loop electroweak corrections for the $k \to \pi\nu\overline{\nu}$ decays," *Phys. Rev. D*, vol. 83, p. 034030, Feb 2011.

[9] J. Brod, M. Gorbahn, and E. Stamou, "Two-loop electroweak corrections for the k $\to$ pi nu anti-nu decays," *Phys.Rev. Lett.*, vol. D83, p. 034030, 2011.

[10] A. Buras, D. Buttazzo, and R. Knegjens, "$k^+ \to \pi\nu\bar{\nu}$ $k_L \to \pi^0\nu\bar{\nu}$ in the standard model: Status and perspectives," *JHEP*, vol. 1511, p. 033, 2015.

[11] CERN-SPSC-2005-013, *Proposal to Measure the Rare Decay $K^+ \to \pi^+\nu\bar{\nu}$ at the CERN SPS*, 2005.

[12] A. Romano, "The kaon identification system in the NA62 experiment at CERN," in *Proceedings, 4th International Conference on Advancements in Nuclear Instrumentation Measurement Methods and their Applications (ANIMMA 2015): Lisbon, Portugal, April 20-24, 2015*, pp. 1–8, 2015.

[13] NA62-Collaboration, *Tenchnical Design Document*, 2010.

[14] A. Kluge *et al.*, "The TDCpix readout asic: A 75 ps resolution timing front-end for the NA62 gigatracker hybrid pixel detector.," *Nuclear Instruments and Methods in Physics Research Section A*, vol. 732, p. 511–514, 2013.

[15] J. Ocariz, "The na48 liquid krypton calorimeter description and performances," 2008.

[16] V. Duk, "The rich detector of the na62 experiment," *International Journal of Modern Physics: Conference Series*, vol. 44, p. 1660230, 2016.

[17] K. Massri, *Studio di alcuni processi di background nell'esperimento NA62*, 2010. Tesi di Laurea Specialistica.

[18] N. Azorskiy *et al.*, "The NA62 spectrometer acquisition system," *Journal of Instrumenation*, p. p. C02064, 2016.

[19] D. Moraes *et al.*, "The carioca front end chip for the lhcb muon chambers," *LHCb-2003-009*, 2003.

[20] P. Lichard *et al.*, "Performance evaluation of multiple (32 channels) sub-nanosecond tdc implemented in low-cost fpga," *Journal of Instrumentation*, vol. 14, p. p. C03013, 2014.

[21] Altera, "Cyclone iii device handbook,"

[22] B. Taylor, *TTC laser transmitter (TTCex, TTCtx, TTCmx) User Manual*, 2012.

[23] A. Ceccucci *et al.*, "The na62 liquid krypton calorimeter's new readout system," *Journal of Instrumentation*, vol. 9, p. p. C01047, 2014.

[24] P. Jovanovic, *Local Trigger Unit Preliminary Design Review*, 2003.

[25] B. Angelucci, E. Pedreschi, M. Sozzi, and F. Spinella, "Tel62: an integrated trigger and data acquisition board," *Journal of Instrumentation*, vol. 7, no. 02, p. C02046, 2012.

[26] E. Pedreschi, B. Angelucci, G. Lamanna, G. Magazzu', J. Pinzino, R. Piandani, M. Sozzi, F. Spinella, and S. Venditti, "A high-resolution tdc-based board for a fully digital trigger and data acquisition system in the na62 experiment at cern," in *Real Time Conference (RT), 2014 19th IEEE-NPSS*, pp. 1–1, 2014.

[27] S. Chiozzi, E. Gamberini, A. Gianoli, G. Mila, I. Neri, F. Petrucci, and D. Soldi, "Level zero trigger processor for the ultra rare kaon decay experiment—na62," *Journal of Instrumentation*, vol. 11, no. 02, p. C02037, 2016.

[28] S. Chiozzi, E. Gamberini, A. Gianoli, G. Mila, I. Neri, F. Petrucci, and D. Soldi, "The Level 0 Trigger Processor for the NA62 experiment," *Nuclear Instruments and Methods in Physics Research A*, vol. 824, pp. 324–325, July 2016.

[29] D. Soldi *et al.*, *Results from the NA62 2014 Commissioning Run*, 2014.

[30] "High speed communications package between host computer and fpga." http://www.eecg.toronto.edu/~tm4/tmu/.

[31] C. Gaspar *et al.*, *Distributed Information Management System*, 2009.

[32] "Wireshark." https://www.wireshark.org.

[33] R. Ammendola *et al.*, "Graphics processors in hep low-level trigger systems," *EPJ Web Conf.* **127** *00011.*, 2016.

[34] H. Weisberg, "Effective spill length monitor," *AGS Division Technical Note*, vol. 163, 1980.

[35] V. Agoritsas, "A sealed metal argon ionization chamber (argonion)," *IEEE Transactions on Nuclear Science*, vol. 28, pp. 2243–2245, June 1981.

[36] G. Ruggiero, "Monte carlo studies of a l0 trigger scheme for $k^+ \to \pi^+ \nu\bar{\nu}$," *NA62 internal note*, 2014.

[37] E. Goudzovski and C. Parkinson, "Studies of the l0 trigger for rare decays," *NA62 internal note*, 2014.