UNIVERSITÀ
DEGLI STUDI
FIRENZE

DOTTORATO DI RICERCA IN
FISICA E ASTRONOMIA

CICLO XXXVII

Search for a dark matter mediator (X) in the $K^+ \to \mu^+\nu_\mu X,\ X \to \gamma\gamma$ decay from the high-intensity K^+ beam of the CERN NA62 experiment

Settore Scientifico Disciplinare
PHYS-01/A

Dottoranda
Dott.ssa Panichi Ilaria

Supervisori
Prof. Lenti Massimo

Dott.ssa Bucci Francesca

Coordinatore
Prof. Modugno Giovanni

Contents

Introduction

Dark matter (DM) is known to constitute approximately 27% of our Universe. However, the nature of DM particles is not described by the Standard Model (SM) of particle physics and still remains elusive. One paradigm that has recently become of great interest and is still largely unexplored, assumes that DM particles have a mass below the electroweak scale and are neutral with respect to the SM gauge interactions. To explain both the observed DM abundance and the absence of significant evidences for such light DM, the existence of a new interaction mediated by a light force carrier that feebly couples to SM particles can be hypothesized.

A dark matter mediator (X) with mass $< 1\,\mathrm{GeV}$, both a gauge boson vector or a CP-even scalar, neutral with respect to the SM interactions and coupling preferentially to muons, is not only motivated by the thermal freeze-out of dark matter but also by anomalies observed in muon physics. The longstanding discrepancy between the SM prediction and the experimental value of the muon anomalous magnetic moment, $a_\mu = (g-2)_\mu/2$, recently confirmed by the precise measurement of the Muon $g-2$ experiment at Fermilab [1][2], could be explained, for example, by the exchange of the X mediator.

Kaon decays are ideal to investigate the existence of such a light mediator. In the decay $K^+ \to \mu^+\nu_\mu$, for example, X could be radiated from the muon via the $\mu - X$ coupling if the X mass (m_X) is lower than the difference $m_K - m_\mu$, making $K^+ \to \mu^+\nu_\mu X$ a rare decay channel for the X search. Several New Physics (NP) models also predict the decay of the X mediator to both SM and DM particles, each decay mode constituting a different investigation channel. The unprecedented number of K^+ decays collected by the CERN NA62 experiment, that aims to measure the branching ratio of the ultra-rare $K^+ \to \pi^+\nu\bar{\nu}$ decay with an accuracy comparable to its theoretical prediction, could allow NA62 to probe the yet unexplored region of the $\mu - X$ coupling. The first search for $K^+ \to \mu^+\nu_\mu X_{inv}$ decays in the X mass range $10 - 370$ MeV/c^2 by the analysis of the data collected in 2016-2017-2018 (Run1) as already been published by NA62 [3]. X_{inv} can be both a long lived X that escapes the detector before decaying or decay into neutrinos or light DM states, hence the *inv* subscript.

This work focuses on the first search for a scalar mediator in the $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ process in the NA62 Run1 data. The $X \to \gamma\gamma$ is expected to be the main X visibly decaying scenario when $m_X < 2m_\mu$ and can produce an invisible signature if X is sufficiently long lived. Here the hypothesis of prompt $X \to \gamma\gamma$ decay is assumed, so that the decay occurs inside the NA62 detector volume and can be reconstructed, thus making the search complementary to the $K^+ \to \mu^+\nu_\mu X_{inv}$ search.

This thesis is organized as follows. In Chapter 1 more details are provided about the physics case of the X mediator. In Chapter 2 the NA62 experiment is presented and the relevant details about the detector are provided. In Chapter 3 the strategy adopted for the analysis is explained and the NA62 data samples and the Monte Carlo simulations, used to study both the signal and its backgrounds, are described. Chapter 4 and Chapter 5 are then dedicated to the report of the data analysis, with highlights on the main backgrounds and on the outline of the signal selection. Finally, the statistical method adopted for the search is presented in Chapter 6 and the preliminary results are reported in Chapter 7.

Chapter 1

The $K^+ \to \mu^+\nu_\mu X$ physics case

The Standard Model (SM) of particle physics provides a solid foundation for the present understanding of elementary particles and their interactions. Nevertheless, there are fundamental questions that are still open, suggesting that the SM is an approximation of a more fundamental theory. However, there is currently no theoretical indication about the scale the New Physics (NP) occurs.

The existence of the dark matter (DM) is one of the clearest evidences of NP beyond the Standard Model (BSM). While astrophysical measurements offer insights into the impact of dark matter (DM) at various scales, from the galactic to the primordial universe, its fundamental nature remains elusive [4][5]. The landscape of the possible masses spans several tenth of order of magnitude. Assuming the DM to be in thermal equilibrium with visible matter in the early universe [6], for example, DM with any mass from the $\mathcal{O}(\text{MeV/c})^2$ up to the multi TeV/c^2 range can achieve the observed abundance by thermal freeze-out.

1.1 The WIMP miracle

So far, most of the experimental efforts in DM searches have been focused on probing the existence of Weakly Interacting Massive Particles (WIMPs), new particles with mass in the 10 GeV/c^2 - 10 TeV/c^2 range, coupling to the visible matter with $\alpha \sim \alpha_{weak}$, where α_{weak} is the strength of the SM weak interaction. Other than being naturally produced as thermal relics of the big bang with the densities accounting for the observed dark matter abundance [7], WIMPs are highly compelling NP candidates. In several models addressing to the hierarchy problem of the Standard Model, the solution to the unnaturalness of the weak scale is the existence of NP at the TeV/c^2 scale, e.g. susy partners in Supersymmetry [8]. This often comes along with a predicted WIMP candidate if one of the new particle is protected against the decay by a symmetry [9]. The connection between the measured DM abundance and solutions to the hierarchy problem is often referred to as WIMP miracle. Additionally, WIMPs occur at an energy scale that can be probed by current experiments. In particular, WIMPs are investigated in several ways. They can be directly detected by their scattering off normal matter through processes like $DM + SM \to DM + SM$, exploited, for example, in underground nuclear-recoil experiments [10]. They can also be indirectly detected by the products of the annihilation $DM + DM \to SM + SM$, occurring somewhere in space and observed in space-based or ground-based experiments [11]. They can also be searched for at accelerators. The most powerful accelerator existing today, the Large Hadron Collider (LHC) at CERN [12], was designed with one of the main goals being the discovery or exclusion of the SM Higgs boson, and provides energies in the center of mass frame up to several TeV/c^2. Given the available energy in the collision and the WIMPs properties, if they exist they will likely be produced at LHC [13]. Up to now, however, no unambiguous WIMP signal has been found from direct detection, indirect detection or accelerator experiments.

1.2 Overview of feebly interacting particles

With the experimental program to search for WIMPs well underway, the possibility of DM candidates at smaller masses has recently became of great interest. Light thermal DM in the MeV/c^2 - GeV/c^2 range generically requires the existence of a new interactions to enable a sufficiently large annihilation rate to ensure the observed abundance [14]. Light new mediators with feeble couplings to the SM are also supported by possible explanations to other open questions in particle physics, as will be specified.

The paradigm that accounts for NP with mass below the electroweak (EW) scale and feebly interacting with the SM fields is usually referred to as a Feebly Interacting Particles (FIPs) scenario [15]. The assumption of feeble interactions opens the possibility to the existence of entire new dark sectors, with similar structures to the SM one but that are "hidden" due to the smallness of their coupling to the SM. In the following a brief overview of FIPs is provided.

Portal interactions. Light NP particles must be neutral under the SM gauge interactions, otherwise they would have been discovered in previous or present accelerator experiments, and couple the SM fields via mixing through low-dimensional operators. Let O_{SM} be an operator composed from the SM fields and O_{DS} the counterpart composed from the dark sector fields. The most general interaction (or portal) Lagrangian is:

$$\mathcal{L}_{portal} = \sum O_{SM} \times O_{DS},$$

where the sum goes over a variety of possible operators of different composition and dimension. According to the general logic of quantum field theory, the lowest dimension operators are the most important. As the SM renormalizable singlet operators of lowest dimension are only four, four different portal interactions can be identified, each associated to a different mediator. Table 1.1 reports the summary of the minimal portal interactions for every model, while in the following a brief description is provided:

- the vector portal model, where the interaction with the SM fields is mediated by a new gauge vector boson, resulting from extra-gauge symmetries of the hidden

Dark photon (A'_μ)	$-\frac{\epsilon}{2\cos\theta_W} F'_{\mu\nu} B^{\mu\nu}$ with F' (B) dark photon (SM hypercharge) field strength ϵ is called kinetic mixing parameter
Dark Higgs (S)	$(\mu S + \lambda S^2) H^\dagger H$ with H SM Higgs doublet, and μ and λ dimensionless and dimensional couplings
HNL (N)	$y_N LHN$ with y_N Yukawa coupling of the sterile neutrino and L one of the SM left-handed doublets
Axion-like particle (a)	$\frac{a}{f_a} F_{\mu\nu}\tilde{F}^{\mu\nu}$, $\frac{a}{f_a} G_{i,\mu\nu}\tilde{G}_i^{\mu\nu}$, $\frac{\partial a}{f_a}\bar{\psi}\gamma^\mu\gamma^5\psi$ with dimension-4 di-photon, di-gluon or di-fermions operators.

Table 1.1: Summary of minimal portal interaction in the FIPs scenario. Left: mediator. Right: Lagrangian describing the interaction between SM and hidden sector. For the axion-like particle the set of possible couplings is in principle very large, here only the flavor-diagonal subset is considered.

sector physics. The minimal vector portal model predicts the existence of a new $U'(1)$ gauge interaction, the corresponding gauge boson being, the dark photon (A').

- the scalar portal (or Higgs portal) model, where H the minimal extension of the SM predicts the existence of one extra singlet scalar (S). The interaction arises through the coupling of S and the Higgs boson (h) via the bilinear SM operator $H^\dagger H$, where H is the SM Higgs doublet. Two type of mixing are possible, the $h - S$ mixing, with coupling μ, and the $h - S^2$ mixing, with coupling λ.

- the neutrino portal, where the existence of one or more new fermions, called Heavy Neutral Leptons (HNLs), is assumed. The existence of HNLs is strongly motivated as they are possibly involved also in the origin of the neutrino mass, another open question of the SM.

- the pseudoscalar portal model, mediated by axion-like particles (ALP) coupling to SM gauge boson and fermion fields. The existence of ALPs was initially postulated to provide a solution to the strong CP problem, or apparent lack of CP violation in strong interactions.

Extensions of the minimal portals can be constructed by gauging accidental symmetries of the SM or individual flavor numbers. An example of non minimal portal interactions relevant for this work are the leptonic force mediators, that will be explained in Sec.1.3.

Signatures. Contrary to models with NP above the EW scale, FIPs scenarios are still largely unexplored. As a result of its feeble coupling to the SM fields, the experimental methods used for WIMPs searches, for example, are not suitable for light DM searches. However, FIPs are expected to be produced, for example, in rare meson decays or in hadronic interactions of protons with nucleons when high-intensity sources are available. Their experimental signatures are usually characterized by missing energy and momentum, as they typically escape the detector before decaying or they decay into invisible states (such as neutrinos or DM particles), but decays to visible final states are also allowed.

The detection of FIPs produced in proton-nucleon interactions typically relies on observing the decay of these particles into SM particles. Forward configurations are needed and long experimental setups are best suited for this type of search, as those of fixed-target experiments. In particular, the solid angle covered by the detector defines the kinematic range of the FIPs that can be studied. Examples of FIP decays include dark scalars or dark photon decays in lepton pairs $S/A' \to l^+l^-$ and heavy neutral lepton decays $N \to \pi^{\pm}\mu^{\mp}$.

In the associated production of FIPs with SM particles in mesons decays, the meson mass determines the FIPs mass range that can be probed. Kaon experiments have typically the best sensitivity to new light particles in the mass range from a few to few hundred MeV/c^2, where the lower bounds on the couplings are typically set by cosmological constraints. A comprehensive overview of observables in the kaon decays and of the possible search strategies have been recently published in [16]. Authors have identified the following as the most promising:

- searches for the $K \to \pi X_{inv}$ decay, where X_{inv} is an invisible particle, represent a unique probe into the parameter space of a dark scalar or ALP long lived or decaying into invisible particles. Searches for resonances in the $K \to \pi l^+l^-$ and $K \to \pi\gamma\gamma$ decay spectra can account for the ALPs decaying into visible states;
- searches for heavy neutral lepton (N) production in the $K^+ \to l^+N$ decay mode
- searches for a leptonic force mediator (X) in $K^+ \to \mu^+\nu_\mu X$ decays.

Of particular interest for this work is the production of the leptonic force mediator in the process $K^+ \to \mu^+\nu_\mu X$.

1.3 Leptonic forces

Leptonic forces are non-minimal portal interactions predicted in some hidden sector models. A leptonic force mediator is a new scalar (Φ) or vector (V) particle that couples predominantly to leptons and leads to signatures that are distinct from canonical hidden sector scenarios with dark scalars and dark photons, respectively.

The existence of leptonic forces is also strongly motivated in models not involving DM. It is well known that NP may also help to explain experimental anomalies and deviations from the SM. In particular, the contributions to the muon magnetic moment due to the coupling of the muon to new light particles remain as one of the last valuable solution to explain the longstanding discrepancy between the measured and the predicted value of the muon anomalous magnetic moment $a_\mu = (g-2)_\mu/2$ ([1], [2], [17]) with NP below the EW scale.

1.3.1 Flavor non-universal gauge boson model

Assuming the minimal vector model and the existence of the dark photon A', it has been demonstrated that the A' kinetic mixing with the SM photon through the interaction $(\epsilon/2)F'^{\mu\nu}F_{\mu\nu}$, with F' and F field strengths of A' and of the SM photon, respectively, could generate an appreciable shift of the a_μ prediction towards its experimental value (see for example [18]). However, the A' solution to the $(g-2)_\mu$ anomaly has been ruled out. The last remaining portion of the dark photon parameter space able to account for the discrepancy has been excluded by the NA48/2 experiment [19].

New low-mass vectors can couple directly to SM leptons in non-minimal models in which the lepton flavor symmetries are embedded in gauge symmetries. Example of new symmetries are the gauged differences between lepton flavor numbers (L_e, L_μ, L_τ), such as $L_\mu - L_\tau$ or $L_e - L_\mu$ ([20],[21]). These gauged differences are anomaly free, thus provide theories that are valid up to arbitrarily high energy scales (or UV-complete models). Since vectors coupled to electrons are generally well-constrained from B-factories searches and other experiments with electron initial states, $L_\mu - L_\tau$ can be considered as the representative leptonic force model with vector mediator explaining the $(g-2)_\mu$ discrepancy (as proposed in [22] and [23] for example), with only small deviations phenomenologically allowed, as shown in [24]. The portal interaction in these models can be written as (from [16]):

$$\mathcal{L} = -g_V V_\alpha(\bar{\mu}\gamma^\alpha\mu - \bar{\tau}\gamma^\alpha\tau + \bar{\nu}_{\mu_L}\gamma^\alpha\nu_{\mu_L} - \bar{\nu}_{\tau_L}\gamma^\alpha\nu_{\tau_L}). \tag{1.1}$$

where V is the field of the vector boson of the $L_\mu - L_\tau$ gauged symmetry. Independently from the mediator mass, thanks to the coupling to neutrinos in Eq.1.1, there is an appreciable branching ratio for V into undetected particles ($V \to invisible$). Between the di-muon and di-tau thresholds, V has an approximately equal branching fractions to muons and invisible products, while below the di-muon threshold the invisible branching fraction is approximately 100%. For masses $m_V \geq 210$ MeV, the model is constrained by direct searches at BaBar [25], and indirect constraints from neutrino trident processes [26]. Below the di-muon threshold, the model is relatively unconstrained. In this region, present kaon experiments have the best sensitivities to the V production via the process $K^+ \to \mu^+\nu_\mu V$. To be mentioned that, if dark matter fields are charged under the $L_\mu - L_\tau$ symmetry group and have a mass lower than the V mass, the decay of V into dark matter particles is allowed and the branching ratio of $V \to invisible$ could significantly change, especially below the di-muon threshold.

1.3.2 Scalar model

New light scalars coupled to leptons could also provide a solution to the $(g-2)_\mu$ discrepancy. In 1990 Kinoshita and Marciano [27] observed that a SM-like Higgs boson with a mass well below the muon mass could shift the muon anomalous magnetic moment prediction towards the scale of its experimental value, the shift being:

$$\Delta a_\mu = \frac{3}{16\pi^2} \times \left(\frac{m_\mu}{v}\right)^2 \sim 3.5 \times 10^{-9}, \tag{1.2}$$

with $v = 246\,\text{GeV}$, related to the vacuum expectation value of the SM Higgs doublet (H) by the equation $\langle H \rangle = v/\sqrt{2}$. Today this hypothesis is excluded by the discovery of the Higgs at the LHC [28] [29], however, it is still possible to assume that if a new light scalar particle couples to leptons with a coupling strength of the order of SM Yukawa couplings of the leptons ($m_\mu/v \sim 4 \times 10^{-4}$ for the muon), the $(g-2)_\mu$ discrepancy can be solved. Thus, the study of the effective Lagrangian:

$$\mathcal{L}_{eff} = \frac{1}{2}(\partial_\mu \Phi)^2 - \frac{1}{2} m_\Phi^2 \Phi^2 + \sum_{i,j=e,\,\mu,\,\tau} (-y_{ij}\bar{l}_i \Phi l_j + h.c.), \tag{1.3}$$

is well motivated. $\mathcal{L}_{eff}$ describes the interaction of an elementary scalar particle Φ with suppressed or absent quark couplings and dark Yukawa couplings to the SM leptons y_{ij}. To evade stringent constraints from charged lepton flavor violation experiments, only the flavor diagonal couplings $y_{ij} = y_i \delta_{ij}$ are considered. To be noted that, as Φ is not the SM Higgs boson, the Yukawa interactions of Eq.1.3 are forbidden by the SM gauge symmetry before the electroweak symmetry breaking. Hence, the effective Lagrangian may originate from an effective gauge-invariant operator, that provides the low-energy approximation of a theory that holds at higher energy scales (UV completion of the model). Different UV completion are possible, also leading to model-dependent effects. The couplings between the scalar field and the SM leptons can therefore be of different types:

- mass proportional ($y_i \propto m_i/v$): the scalar couples predominantly to τ leptons. This interaction occurs, for example, in UV-complete models including lepto-philic Higgs sectors, as in [30].
- muon-philic ($y_\mu \neq 0$, $y_e = 0 = y_\tau$): this coupling can arise in a model with a new weak-scale vector-like lepton (ψ') with the same gauge charge as the SM right-handed muon (μ_R) and mass $m_\psi > v$ [31]. The terms $\phi\bar{\psi}'\mu_R$ and $H^*\bar{\psi}'\mu_L$ generate the $\phi\mu\mu$ interaction when the vector-like leptons are integrated out. This scenario can account for $(g-2)_\mu$ discrepancy with the minimum number of couplings to other SM leptons.
- generic but with $y_e << y_\mu, y_\tau$ to avoid constraints from experiments with electron initial states. As an example, in [32] a dark sector $U(1)_d$ symmetry group is assumed, with an associated dark photon with mass $\leq$ GeV-scale, together with

more then one dark weak-scale vector-like lepton and one extra Higgs doublet with hypercharge. The low-mass scalar Φ is associated to the $U(1)_d$ symmetry breaking in the dark sector.

1.3.3 Signatures and benchmarks

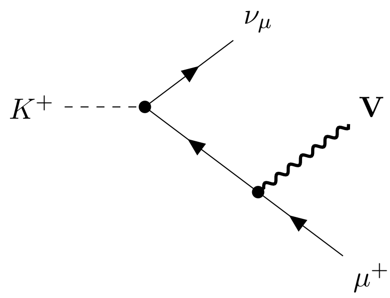
Let X be either a vector or a scalar leptonic force mediator. Thanks to the availability of large datasets, charged kaon decays are ideal processes for probing low-mass leptonic mediators. In particular, in the most abundant charged kaon decay mode, $K^+ \to \mu^+\nu_\mu$ (having a branching ratio $\sim 63\%$), X can be produced through final state radiation (FSR) of the muon or (when allowed) of the muon neutrino. The process of interest is therefore $K^+ \to \mu^+\nu_\mu X$, that allows to investigate the existence of X in the mass range $m_X < m_K - m_\mu$. The relevant Feynman diagram representing the process is shown in Figure 1.1 assuming a vector (left) or a scalar (right) mediator.

The branching ratio (BR) of the $K^+ \to \mu^+\nu_\mu X$ decay is mainly determined by the $X - \mu$ coupling. As the same coupling determines X contributions to the $(g-2)_\mu$ value, it is possible to map the models that account for $(g-2)_\mu$ to specific values of $BR(K^+ \to \mu^+\nu_\mu X)$. The only model dependence from the NP model arises in the spin of X, since the scalar and vector FSR rates are different. There is more ambiguity in mapping the DM parameter space to the kaon branching fraction. Even assuming the DM thermal freeze-out, the specific process that sets its relic abundance depends on the mediator/DM mass hierarchy. Thus, in determining the region of the $BR(K^+ \to \mu^+\nu_\mu X)$ vs m_X space consistent with DM, some hypotheses for the nature of DM and for the coupling of the DM to the mediator are needed. In [16] the target branching fractions motivated by the $(g-2)_\mu$ and by the DM (under specific assumptions, not reported here) are evaluated. An experimental sensitivity to branching fractions of $\mathcal{O}(10^{-9})$ would almost cover the whole interesting parameter space for both the vector and scalar mediator. For masses smaller than 10 MeV/c^2, thermal DM models can run into constraints from the big bang nucleosynthesis [14] while for mediator masses near the kinematic endpoint $(m_K - m_\mu)$, B factories can provide complementary sensitivity to kaon factories.

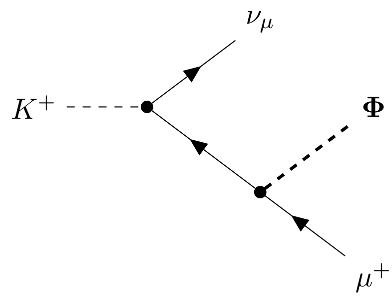
The $K^+ \to \mu^+\nu_\mu X$ process can be searched for in different channels. The experimental signatures are determined by the decay mode of X, that can be either to SM or to hidden sector particles. The best-motivated decay modes are:

- $X \to \mu^+\mu^-$: this decay mode is always present and becomes the dominant one in models where X is lighter than all the other hidden sector particles and $2m_\mu \leq m_X$. For the range of couplings accessible at the present kaon factories, $X \to \mu^+\mu^-$ is always prompt.

- $X \to e^+e^-$: this decay dominates below the di-muon threshold in models with mass-proportional $X - e$ coupling, and is completely absent if a null X coupling to electrons is assumed. Several constraints has already been set through this channel. For the range of couplings accessible at the present kaon factories, $X \to e^+e^-$ both the prompt and displaced searches are of interest.



(a)



(b)

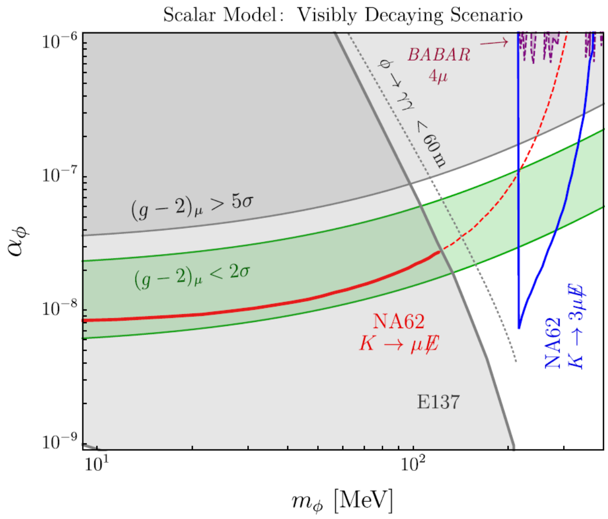
Figure 1.1: Representative Feynman diagrams of the production of a leptonic mediator X as muon final state radiation in the decay $K^+ \to \mu^+\nu_\mu X$, when X is a flavor non-universal gauge boson (V) on the left or a lepto-philic scalar (Φ) on the right. To be noted that a second diagram exists for the vector model, in which V radiates off from the neutrino line. The diagram is taken into account for calculations. Diagrams are created by using [33].

- $X \to \gamma\gamma$: this is a loop-induced decay mode that dominates if $m_X < 2m_\mu$ in the absence of other interactions. The partial width depends on the UV completion of the model and can lead to prompt or displaced di-photon decays.
- $X \to$ *invisible*: this decay mode is present if X couples to neutrinos; the decay width for $X \to$ *invisible* can largely increase if the X decay into invisible hidden sector states is allowed. As the $X \to \gamma\gamma$, this decay mode is relevant below the di-muon threshold if X does not couple to electrons and $X \to \gamma\gamma$ lifetime is long enough for X to decay outside of the detector volume, becoming itself invisible.

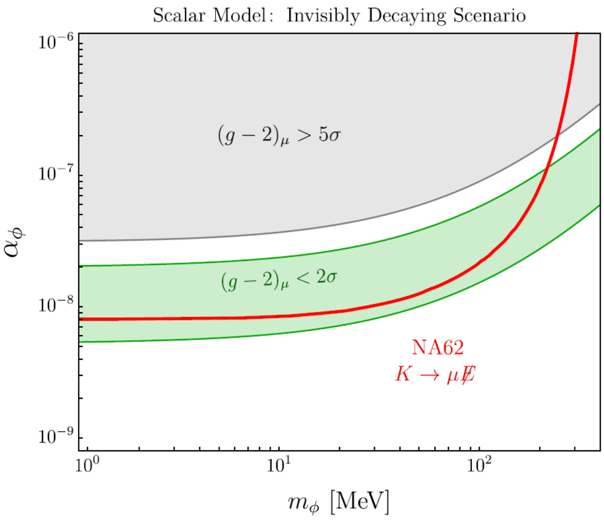
1.4 Searches for muon-philic forces at the CERN NA62 experiment

The present kaon factories, such as the NA62 experiment at CERN (see Chapter 2), could have the leading sensitivity to leptonic forces with enhanced coupling to the muons. A theoretical study of the NA62 sensitivity to the $K^+ \to \mu^+\nu_\mu X$ process [31], shows that, thanks to the unprecedented number of collected K^+ decays, NA62 may considerably improve the existing limits on muon-philic mediators for both the vector and the scalar models.

Considering the $X \to \mu^+\mu^-$, the search for the prompt $K^+ \to \mu^+\nu_\mu X$, $X \to \mu^+\mu^-$ mode is promising due to the low background expected. Assuming the total background to be of the same order of magnitude as the contribution from the rare $K^+ \to \mu^+\nu_\mu\mu^+\mu^-$ decay (which is supported by a preliminary study of the data [34]), the analysis of [31] showed that through a di-muon search in a dataset of 10^{13} collected K^+ decays, NA62 would completely cover parameter space accommodating the experimental $(g-2)_\mu$ value for most masses between the di-muon threshold and the kinematic endpoint for the X production. The NA62 potential probing power would complement the BaBar



(a)



(b)

Figure 1.2: From reference [31]: Parameter space of a muon-philic scalar mediator Φ, defined by the mass of Φ and by $\alpha_\Phi \equiv y_\mu^2/4\pi$, where y_μ is the Yukawa coupling $\Phi - \mu$ from Eq.1.4. NA62 projections are also shown, assuming 10^{13} K^+ decays and searches statistics dominated; other details on their computation can be found in the supplemental material of the paper. Top: projection for an NA62 search for $K^+ \to \mu^+\nu_\mu\Phi$ with Φ visibly decaying into $\mu^+\mu^-$ or $\gamma\gamma$. $\Phi \to \mu^+\mu^-$ is dominant above the di-muon threshold (blue line). $\Phi \to \gamma\gamma$ is favored below the di-muon threshold; on the left of the dashed gray line the lifetime of Φ is such that all the decays occur outside the NA62 detector. The search for $K^+ \to \mu^+\not{E}$ is therefore considered (red line), where $\not{E}$ stands for every undetected particle produced in the process, neutrino and Φ long-lived. The light green band represents the region of the parameter space that is favored, within 2σ, by the experimental value of $(g-2)_\mu$ (updated at the time the paper was published); the gray band on the top part of the parameter space represents excluded regions leading to a 5σ discrepancy with respect to the measured $(g-2)_\mu$ value. Exclusion regions are also shown from secondary muon production at E137 [35] and from the BABAR search for $e^+e^- \to \mu^+\mu^-\mu^+\mu^-$ [25]. Bottom: projection for an NA62 search for $K^+ \to \mu^+\nu_\mu\Phi$ with Φ decaying to undetected particles (e.g. DM) only. The green and gray bands are defined as on top plot.

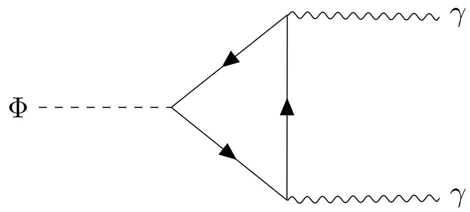


Figure 1.3: Representative Feynman diagram of the contribution to the $\Phi \to \gamma\gamma$ decay for Φ scalar muon-phylinc force mediator, assuming the effective interaction in Eq. 1.4. Assuming that y_μ is the only non-zero coupling between Φ and leptons, the only fermions exchanged in the loop are virtual muons.

exclusion region [25], which rules out the $(g-2)_\mu$ preferred parameter space for di-muon decays of vectors but not of scalars.

However, for X produced in the $K^+ \to \mu^+\nu_\mu X$ decay, most of the favored parameter space is below the di-muon threshold, as shown in Figure 1.2 for the scalar mediator. In particular, given the muon-philic hypothesis, the effective lagrangian of Eq.1.3 reads (ϕ is a real scalar particle):

$$\mathcal{L}_{eff}^{\mu-philic} = \frac{1}{2}(\partial_\mu \Phi)^2 - \frac{1}{2}m_\Phi^2\Phi^2 - y_\mu \Phi\bar{\mu}\mu \tag{1.4}$$

and $\Phi \to \gamma\gamma$, that dominates when $m_\Phi < 2m_\mu$, occurs through the muon loop reported in Figure 1.3.

The corresponding proper decay length of $\Phi \to \gamma\gamma$ is in the range $1\,cm - 1\,m$, assuming for the coupling $\alpha_\Phi \equiv y_\mu^2/4$ the unexcluded parameter range below the di-muon threshold in Figure 1.2a. Then, if Φ is produced in the decay of an NA62 beam K^+, the decay length in the laboratory frame is:

$$l_{\Phi\to\gamma\gamma} \sim (60\,m)\left(\frac{3\times 10^{-6}}{\alpha_\Phi}\right)\left(\frac{50\,\text{MeV}}{m_\Phi}\right)^4\left(\frac{E_\Phi}{75\,\text{GeV}}\right) \tag{1.5}$$

where the m_Φ^{-4} scaling accounts for the boost factor, 60 m is the typical length of the NA62 fiducial volume for the reconstruction of the K^+ decays and 75 GeV/c is the typical K^+ momentum. If Φ has energy $E_\Phi \sim 75\,\text{GeV}$, nearly all the Φ decays occur outside the NA62 detector and mimic the missing energy signature of the invisibly decaying scenario, in which Φ goes to undetected particles (e.g. dark matter). However, depending on the completion model and for some choice of parameters, further UV contributions can introduce new diagrams and their contribution can lead to a prediction of the partial width for $\Phi \to \gamma\gamma$ photons sensibly different from the one including only the loop in Figure 1.3. As different choice of parameters can lead to either prompt or displaced decays to photons, experimental searches for the invisible final state and for di-photon resonances are both necessary.

This work aims to search the $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ process with X scalar muon-philic mediator, through a bump-hunt in the distribution of the reconstructed di-photon invariant mass, by the analysis of the NA62 data collected in 2016-2017-2018 (Run1). The prompt $X \to \gamma\gamma$ decay hypothesis is assumed, ensuring the decay to be visible in the detector. This search considers the whole kinematic accessible range and is complementary to the first X search in the $X \to$ *invisible* channel published by NA62 in 2021 [3].

NA62 searched for $K^+ \to \mu^+\nu_\mu X$, $X \to$ *invisible* decays in the minimum bias data collected during Run1, the main background sources coming from the SM $K^+ \to \mu^+\nu_\mu\gamma$ and $K^+ \to \mu^+\nu_\mu\pi^0$ decay modes when the final state photons go undetected. The background estimate relies on simulations and is affected by large systematic uncertainties due to the accuracy of non-gaussian tails modeling of the discriminating variable of the study, of the photon veto inefficiency and of the photon conversions. X mass hypotheses in the range $10-370\,\mathrm{MeV/c^2}$ were tested and upper limits at 90% C.L. of the $K^+ \to \mu^+\nu_\mu X$ branching ratio were set for both the vector and the scalar mediator. The branching ratio upper limits range from $\mathcal{O}(10^{-5})$ for low m_X to $\mathcal{O}(10^{-7})$ for high m_X, much weaker than the values accounting for the $(g-2)_\mu$ discrepancy. The reduction of systematic uncertainties in the background estimate in order to make progress towards the region of interest of the parameter space is very challenging. To be mentioned that, by the same dataset, the first upper limit on the decay rate of the SM $K^+ \to \mu^+\nu_\mu\nu\bar{\nu}$ mode was also established: $BR(K^+ \to \mu^+\nu_\mu\nu\bar{\nu}) < 1.0 \times 10^{-6}$ at 90% C.L. This decay mode may constitute an irreducible background to the $K^+ \to \mu^+\nu_\mu X$, $X \to$ *invisible* search, but was foud to be negligible.

In this Chapter, the physics case of a new dark matter mediator (X) neutral with respect to the SM interactions and with enhanced coupling to the muons, that can be radiatively produced in rare kaon decays, has been introduced, together with the aim of this thesis, being the first experimental search for the $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ process at NA62. The data analysis will be detailed in the next Chapters, after the description of the NA62 experiment.

Chapter 2

The NA62 experiment at CERN

NA62 is a fixed-target experiment at the CERN Super Proton Synchrotron (SPS) dedicated to the K^+ meson physics. The main goal of the NA62 experiment is the measurement of the branching ratio of the ultra rare $K^+ \rightarrow \pi^+\nu\bar{\nu}$ decay with an accuracy comparable to its very precise SM prediction (Sec.2.1), but a much broader physics program can be pursued (Sec.2.2) thanks to the availability of an high intensity K^+ beam (Sec.2.3) and to the experimental setup (Sec.2.4 and Sec.2.5).

NA62 collected physics data in 2016, 2017 and 2018 (Run1), then the NA62 Run2 restarted after the second long shutdown (LS2) of the CERN accelerator complex, with data taking periods in 2021, 2022, 2023 and 2024.

2.1 $K^+ \rightarrow \pi^+\nu\bar{\nu}$ in a nutshell

The main focus of the NA62 experiment is the measurement of $BR(K^+ \rightarrow \pi^+\nu\bar{\nu})$, which also determines the experimental setup. A detailed overview of the $K^+ \rightarrow \pi^+\nu\bar{\nu}$ physics case is beyond the aim of this work, however, few highlights are provided in this paragraph.

Within the SM, the $K^+ \rightarrow \pi^+\nu\bar{\nu}$ decay is a flavor Changing Neutral Current (FCNC) process involving the $s \rightarrow d$ quark transition, that is forbidden at tree level. In particular, the decay proceeds at the leading order through the box and penguin diagrams reported in Figure 2.1. In each diagram, all the three up-like anti-quarks ($\bar{u}$, $\bar{c}$, $\bar{t}$) can be exchanged and the amplitude of the contribution for each of them is of the form:

$$A_{q_u} \sim \frac{m_{q_u}^2}{m_W^2} V_{q_u s}^* V_{q_u d}, \tag{2.1}$$

with $V_{q_u d}$ element of the Cabibbo-Kobayashi-Maskawa (CKM) matrix (introduced for the first time in 1973 [36]) determining the probability of transition between the quark flavor $q_u = (u, c, t)$ and the d flavor.

Due to the quadratic GIM mechanism [37] and taking into account the experimental values of the CKM elements, Eq.2.1 shows that:

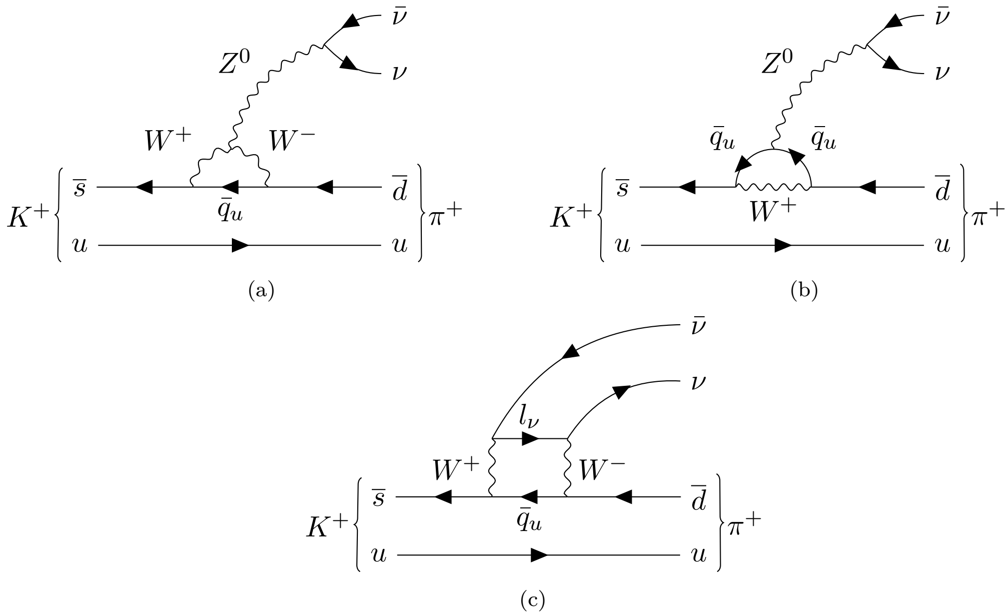


Figure 2.1: Leading order representative Feynman diagrams of the $K^+ \to \pi^+\nu\bar{\nu}$ decay in the SM description. The process is forbidden at tree level and occurs at loop level via the penguins diagrams in (a) and (b) or the box diagram in (c). In the loops, q_u stands for every possible quark of $u-$type while l_ν is the charged lepton forming the left-handed SM doublet together with ν, which can be either ν_e, ν_μ or ν_τ. As virtual $(\bar{u}, \bar{c}, \bar{t})$ can appear in the loops, each diagram amplitude is the sum of three different contributions for every fixed lepton flavor.

- the total amplitude of each of the possible diagrams is dominated by the top exchange, with the charm providing a small but non-negligible contribution;
- the strong CKM-suppression of the $t \to d$ transition places the process among the rarest meson decay in the SM.

In particular, as the up-quark contribution is quadratically GIM-suppressed and as the $\bar{u}$ exchange is the only involved process in which the non-perturbative QCD effects (or long-distance contributions) can be non-negligible in kaon decays, the $K^+ \to \pi^+\nu\bar{\nu}$ process results dominated by short-distance contributions. Moreover, the $BR(K^+ \to \pi^+\nu\bar{\nu})$ can be directly related to the precisely measured $K^+ \to \pi^0 e^+\nu_e$ branching ratio by means of the isospin symmetry. As a consequence, by the normalization to the $BR(K^+ \to \pi^0 e^+\nu_e)$, the contribution to uncertainties coming from the knowledge of the hadronic matrix element of the $K - \pi$ transition, which is usually the biggest one, largely cancels, allowing for a very clean $BR(K^+ \to \pi^+\nu\bar{\nu})$ theoretical prediction. The $BR(K^+ \to \pi^+\nu\bar{\nu})$ can then be written in terms of the CKM matrix parameters as [38]:

$$BR(K^+ \to \nu\bar{\nu}) = (8.39 \pm 0.30) \times 10^{-11} \left(\frac{|V_{cb}|}{0.0407}\right)^{2.8} \left(\frac{\gamma}{73.2°}\right)^{0.74} \tag{2.2}$$

where the precision of the measurements of $|V_{cb}|$ and $\gamma = arg[(-V_{ub}V^*_{ub})/(V_{cd}V^*_{cb})]$ provide the main contribution (known as parametric uncertainty) to the remaining uncertainties. By using the values of the CKM parameters as external inputs it is possible to get the prediction for the branching ratio. The prediction from [38] has been recently updated by using the most recent values for the CKM and other input parameters involved in the calculation of the short-distance contributions [39], providing:

$$BR_{SM}(K^+ \to \pi^+ \nu\bar{\nu}) = (7.86 \pm 0.61) \times 10^{-11}.$$

A different strategy has also been pursued, eliminating the dependence in $BR(K^+ \to \pi^+ \nu\bar{\nu})$ from the CKM parameters by taking suitable ratios of various observables in K and B physics [40]:

$$BR_{SM}(K^+ \to \pi^+ \nu\bar{\nu}) = (8.60 \pm 0.42) \times 10^{-11}.$$

In view of these facts, the $K^+ \to \pi^+ \nu\bar{\nu}$ decay offers unique opportunities for testing the Standard Model and deepening our knowledge of the CKM matrix, in a way that is independent and complementary to B decays, that were historically used to measure the CKM parameters. Furthermore, thanks to its extremely suppressed branching ratio, the $K^+ \to \pi^+ \nu\bar{\nu}$ decay can be very sensitive to the exchange of possible new particles in the loops, probing the highest mass scales among the rare meson decays. The largest deviations from SM are expected in models with new sources of flavor violation [41, 42]. Depending on new physics models, correlation patterns between possible deviations from the SM prediction of $K^+ \to \pi^+ \nu\bar{\nu}$ and $K_L \to \pi^0 \nu\bar{\nu}$ decay modes can be expected [43]. Present experimental constraints limit the range of variation within supersymmetric models [44, 45, 46]. The $K^+ \to \pi^+ \nu\bar{\nu}$ decay can also be sensitive to effects of lepton flavor non-universality [47] and can constrain leptoquark models [48]. Another recent hypothesis includes BSM effects to explain the tension between theoretical expectations and experimental measurements on ϵ'/ϵ and $\Delta M_K = M_{K_L} - M_{K_S}$, with new CP-violation sources that could affect also $BR(K \to \pi\nu\bar{\nu})$ [49].

From an experimental point of view, the smallness of the expected BR and the presence of two undetected neutrinos in the final state make the measurement of this mode very challenging. The first one was performed by the E787 and E949 experiments at the Brookhaven National Laboratory (BNL) using a kaon decay-at-rest technique, where a kaon beam was stopped in a target surrounded by detectors measuring the properties of the particles produced by kaons decaying at rest. The obtained final result is based on 7 $K^+ \to \pi^+ \nu\bar{\nu}$ candidates with a very broad statistical uncertainty ([50],[51]):

$$BR_{BNL}(K^+ \to \pi^+ \nu\bar{\nu}) = (1.73^{+1.15}_{-1.05} \pm 0.29) \times 10^{-10}.$$

Data collected by NA62 during Run1 allowed significant progress in the study of the $K^+ \to \pi^+ \nu\bar{\nu}$ decay. In particular, the analysis of 2016 data [52] demonstrated the feasibility of measuring $BR(K^+ \to \pi^+ \nu\bar{\nu})$ using a completely different technique with respect to previous experiments, the decay-in-flight technique, that allows K^+ decays in the vacuum, avoiding kaon interactions with matter. From data collected in 2017, NA62 observed 2 $K^+ \to \pi^+ \nu\bar{\nu}$ candidates with an expected background of 1.5 events, ensuring the robustness of the measurement strategy and background suppression capabilities [53]. From data collected in 2018, 17 $K^+ \to \pi^+\nu\bar{\nu}$ candidates were observed with an expected background of 5.3 events, allowing for the measurement [54]:

$$BR_{NA62\,Run1}(K^+ \to \pi^+\nu\bar{\nu}) = (10.6^{+4.0}_{-3.4} \pm 0.9) \times 10^{-11} \text{ at } 68\% \text{ CL.}$$

based on a total of 20 signal candidates with an expected background of 7.0 events, from the combination of the results from the analysis of the total data sample collected from 2016 to 2018. This result provides an evidence for the $K^+ \to \pi^+\nu\bar{\nu}$ decay with a significance of 3.4σ. The analysis of the data collected starting from 2021 is now in progress. It can profit from several upgrades of the detector and from an higher beam intensity, and will further improve the world leading precision of the NA62 study of the $K^+ \to \pi^+\nu\bar{\nu}$ decay.

2.2 The NA62 physics program

The employ of the decay-in-flight was motivated by the availability of an high momentum and high intensity K^+ beam, derived from the protons accelerated by the SPS. The high momentum is necessary to boost the decay product to high energies, where their detection and identification can be efficient. The boost folds the decay products into a small angular region close to the beam axis, allowing for a classic fixed target geometry.

This experimental setup, together with the high intensity of the K^+ beam, allows for the study of many processes beyond the $K^+ \to \pi^+\nu\bar{\nu}$, including other precision tests of the SM, searches for decays forbidden by SM symmetries and searches for hidden-sector particles produced in K^+ decays. In particular, thanks to different trigger lines running in parallel to the $K^+ \to \pi^+\nu\bar{\nu}$ one, a total of 6×10^{12} K^+ decays to be analyzed have been collected during NA62 Run 1 [55]. This number provides NA62 with the leading sensitivities to a variety of K^+ (and pion) rare or forbidden decays, with the possibility to both greatly improve on kaon observables measured in the past, or to search for them for the first time. Some example of results recently published by the analysis of Run 1 data (that are only a small part of the Run 1 program outcomes) are: the precision measurements of the $K^+ \to \pi^+\mu^+\mu^-$ [56], $K^+ \to \pi^0 e^+\nu_e\gamma$ [57] and $K^+ \to \pi^+\gamma\gamma$ [58] decays, which are test of the Chiral Perturbation Theory describing the low-energy effects of QCD, the first search for the $K^+ \to \pi^+e^+e^-e^+e^-$ [59] decay, interpreted both as SM decay or as the final state of a process involving FIPs that then decaying in e^+e^- pairs, or the searches for the lepton flavor violating $K^+ \to \pi^+\mu^-e^+$ and $\pi^0 \to \mu^-e^+$ decays and for the lepton number violating $K^+ \to \pi^-\mu^+e^+$ mode [60].

The NA62 experiment also operates in a beam-dump configuration, providing the opportunity to seek decays of hidden-sector particles with masses above the K^+ mass. A dataset corresponding to $(1.4 \pm 0.28) \times 10^{17}$ protons on target (POT) was collected in beam-dump configuration in 2021 and, by the analysis of this dataset, the NA62 searches for a dark photon in the di-lepton final state were published in [61], [62]. Searches in other final states of the dark photons and for heavy neutral leptons are ongoing.

2.3 Beam and experimental requirements

The measurement of the branching fraction of an ultra rare decay with only a charged track and two neutrinos in the final state with the desired precision is experimentally challenging. The NA62 experimental setup is hence designed and optimized to cope with the requirements of the $BR(K^+ \to \pi^+ \nu\bar{\nu})$ measurement, which are summarized in the following.

As a figure of merit, assuming the SM $BR(K^+ \to \pi^+ \nu\bar{\nu})$, an acceptance of $\mathcal{O}(10\%)$ and $\mathcal{O}(100)$ $K^+ \to \pi^+ \nu\bar{\nu}$ candidates observed, an integrated flux of $\mathcal{O}(10^{13})$ kaon decays and a $\mathcal{O}(10^{12})$ rejection factor of the backgrounds coming from the main K^+ decay modes are necessary to achieve the $\mathcal{O}(10\%)$ experimental accuracy. The necessary flux is provided by the kaon beam that can be obtained from SPS protons. A high intensity 400 GeV/c proton beam extracted from the SPS impinges on a 40 cm long beryllium rod (in spills of 3 s effective duration) to produce a high energy secondary beam. The secondary beam particles are selected with positive unitary electric charge and 75 GeV/c nominal momentum (1% rms momentum bite), necessary to employ the decay-in-flight technique to fully reconstruct K^+ decays. Due to their high momentum, protons and pions produced together with kaons on proton-target interactions cannot be separated from them. As a consequence beam detectors measuring kaons are exposed to an hadron beam composed of $(6\%)K^+$, $(23\%)p$ and $(70\%)\pi^+$, producing a particle flux 16 times larger than the kaon flux and a nominal rate of 750 MHz at the entrance of the NA62 detector. Then, to cope with this high rate environment and considering that the experimental signature of $K^+ \to \pi^+ \nu\bar{\nu}$ events consists of one incoming charged track identified as a kaon and a single outgoing charged track identified as a pion, with missing energy and momentum due to the undetected neutrinos, the NA62 detectors must ensure time matching at the level of $\mathcal{O}(100\,\mathrm{ps})$ when measuring the incoming kaon and the outgoing pion.

To suppress the backgrounds to the $K^+ \to \pi^+ \nu\bar{\nu}$ mode, kinematic rejection and particle identification (PID) are used. The kinematic rejection relies on the definition of two signal regions in the distribution of a discriminating variable, the squared missing mass $m^2_{miss}(K-\pi) \equiv (P_K - P_\pi)^2$, where P_K and P_π are the reconstructed four-momenta of the parent and daughter particles assumed to be a K^+ and a π^+, respectively. A suppression of $\mathcal{O}(10^4)$ can be obtained, depending critically on the precision of the tracking systems measuring the three-momenta. The most abundant K^+ decay mode, the $K^+ \to \mu^+ \nu_\mu$ $(63.56 \pm 0.11)\%$ [63]), has both only one charged track and missing energy and momentum in the final state, making a further $\mathcal{O}(10^7)$ rejection factor from PID mandatory to accomplish the $K^+ \to \pi^+ \nu\bar{\nu}$ measurement. To distinguish muons from charged pions, their different stopping power in calorimeters and a Ring Imaging CHerenkov (RICH) detector are used, together with a muon detector vetoing on the muon passage. The $K^+ \to \pi^+ \pi^0$ decay ($(20.67 \pm 0.08)\%$ [63]) provides another main source of background as it can mimic $K^+ \to \pi^+ \nu\bar{\nu}$ events if the photons from the $\pi^0 \to \gamma\gamma$ decay go undetected. Therefore, a $\mathcal{O}(10^7)$ π^0-rejection is also necessary and is obtained by an hermetic system of photon veto.

2.4 Beam line and detector layout

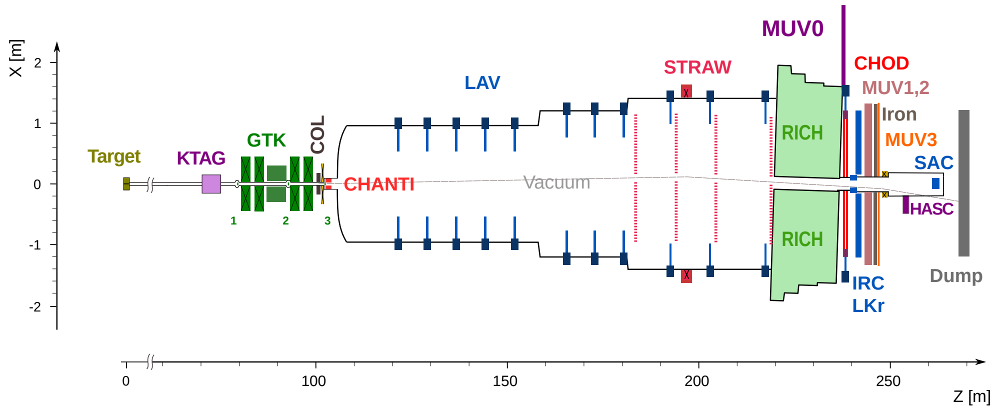


Figure 2.2: Schematic top view of the NA62 detector [64] in the xz plane, with the coordinate system defined as in the text. Dipole magnets are displayed as boxes with superimposed crosses. The label CHOD refers both to the old CHOD detector of the NA48 experiment and the New-CHOD detector designed for NA62.

The layout of the NA62 detector [64] is sketched in Figure 2.2. The beam line defines the Z-axis of the experiment's right-handed laboratory coordinate system, the origin is the kaon production target. Beam particles travel in the positive z direction. The y axis is vertical (positive up), and the x axis is horizontal (positive left).

Beam kaons are characterized by a differential Cherenkov counter (KTAG) and a three station silicon pixel detector (GigaTracker or GTK). The last GTK station (GTK3) is immediately preceded by a collimator, which serves as a shield against hadrons produced by upstream K^+ decays; a six stations plastic scintillator guard ring detector (CHANTI) is used to veto inelastic collisions occurring in GTK3. GTK3 marks the beginning of a 117 m long vacuum tank: the first 80 m of the tank denote a volume in which about 13% of the kaons decay. The K^+ decays used at analysis level are typically detected in a region along the z axis that can start from 105 m and can extend up to 180 m, depending on the analysis. This range defines a decay region referred to as fiducial volume. Downstream of the fiducial volume, the momentum and direction of charged decay products are measured by a magnetic spectrometer (STRAW); a ring-imaging Cherenkov counter (RICH) identifies pions, muons and electrons, and measures the track time, together with two scintillator hodoscopes (CHOD and New-CHOD). Other systems of sub-detectors provide signals used for vetoing the presence of photons (LAV, LKr, IRC and SAC) or multiple charged particles (MUV0, HASC) in the final state, while two hadronic sampling calorimeters (MUV1, MUV2) and an array of scintillator tiles located behind 80 cm of iron (MUV3) supplement the π/μ identification system.

The GTK, all detectors surrounding the fiducial volume and the STRAW spectrometer are placed in vacuum to obtain improved resolution for the measured kinematic

KTAG	$\sigma_t = \mathcal{O}(70\text{ ps})$
GTK	$\sigma_t = \mathcal{O}(100\text{ ps})$, $\sigma_\theta = \mathcal{O}(16\,\mu\text{rad})$, $\sigma_p = \mathcal{O}(0.15\text{ GeV/c})$
STRAW	$\sigma_p/p = 0.3\%$, $\sigma_\theta = 60\,\mu$rad at 20 GeV/c $\sigma_p/p = 0.4\%$, $\sigma_\theta = 20\,\mu$rad at 50 GeV/c
RICH	$\sigma_t = \mathcal{O}(70\text{ ps})$
CHOD	$\sigma_t = \mathcal{O}(200\text{ ps})$
New-CHOD	$\sigma_t = \mathcal{O}(1\text{ ns})$
LKr	$\sigma_E/E = 1.4\%$ at 25 GeV $\sigma_t = 0.5$ ns up to 1 ns, depending on the energy deposit

Table 2.1: Key experimental features of the NA62 sub-detectors relevant for this thesis work.

quantities. The active volume of the RICH and of the detectors downstream with respect to it are traversed by a beam pipe. The evacuated passage surrounding the beam trajectory allows the intense flux of un-decayed beam particles to pass through the detectors without interacting with their material and to get to the dump.

2.5 Detectors

In this section the NA62 subdetectors introduced in Sec.2.4 are described in more detail. The key experimental features of the detectors relevant for this thesis work are summarized in Table 2.1 for quick reference.

2.5.1 K^+ tracking and identification

KTAG. The K^+ mesons from the 75 GeV/c secondary beam are identified by a CEDAR detector, developed at CERN in the late 1970 [65], equipped with the KTAG, a bespoke photon-detection system of 8 arrays (also referred to as sectors) of photomultiplier tubes and front-end electronics [66]. A CEDAR detector is a differential Cherenkov counter with achromatic ring focus inside its vessel, designed to discriminate kaons, pions and protons in unseparated beams extracted from the CERN SPS. It exists in two variants, North (CEDAR-N) and West (CEDAR-W), coming with photon-detectors and readout electronics that cannot sustain the nominal 45 MHz kaon rate of the NA62 beam nor provide the $\mathcal{O}(100\,\text{ps})$ timing resolution required to match the kaon track to the tracks of its charged decay products. The KTAG was therefore developed to meet these requirements [66]. During Run1, NA62 installed a CEDAR-W type vessel, extending from 70 m to 75 m downstream of the beryllium target, filled with nitrogen gas (N_2) at 1.71 bar at room temperature. Since the beginning of Run2, hydrogen (H_2) is used as radiator gas, in order to minimize the amount of material in the path of the beam and to reduce multiple scattering. As no suitable CEDAR existed, a new detector, named CEDAR-H, was developed [67] by refitting a CEDAR-N with optics designed specifically for the usage of H_2. Filling the CEDAR-H vessel with H_2 at 3.85 bar at room temperature allows to obtain similar Cherenkov angles with respect to the

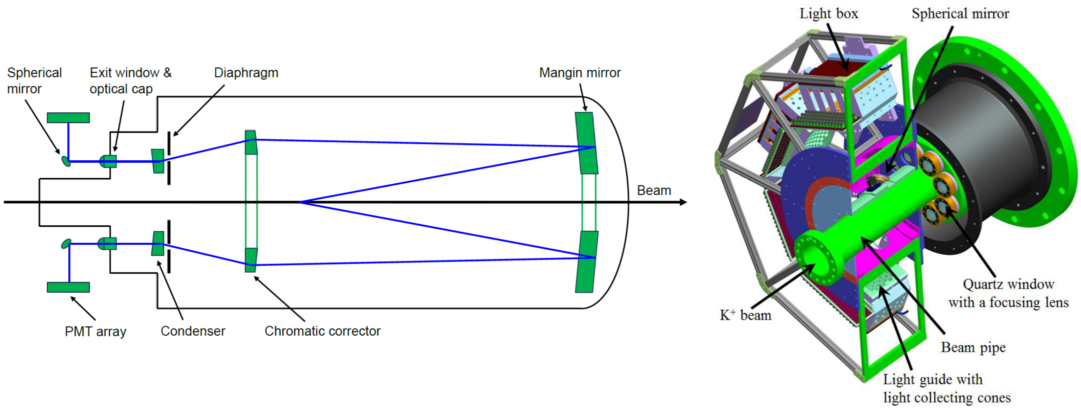


Figure 2.3: Left: Sketch of the CEDAR-H internal optical system (not to scale) [67], constituted by a mangin mirror, a chromatic corrector, lenses and a diaphragm. The diaphragm aperture is fixed depending on the desired kaon/pion separation and on the pressure of the gas in the CEDAR vessel. It allows to collect only light coming from K^+ with 75 GeV/c momentum and it is (almost) blind to pions and protons with the same momentum. The Cherenkov light (blue lines) is collected, reflected and steered onto 8 quartz exit windows in the upstream end, equally spaced around the circumference of the diaphragm. The photon-detection system collects photons exiting the quartz windows and focuses them onto spherical mirrors, which reflect them onto 384 photomultipliers tubes (PMTs). Right: mechanical design of the enclosure housing the KTAG photon-detection system and front-end electronics [64]; the enclosure is divided in two halves (4 octants of PMTs each) to enable installation around the beam pipe. The quartz windows for Cherenkov light collection are also shown.

ones produced in N_2 at 1.71 bar, and to continue to use the KTAG photon-detection system. As an example, a sketch of the CEDAR-H optical system is shown in the left hand side of Figure 2.3, while the right hand side reports the KTAG mechanical design.

Tagging kaons within the unseparated monochromatic beam of charged hadrons entering the NA62 detector allows to reject events in which the final state particles come from beam pions interacting with the residual material in the decay volume. Incoming kaons are tagged with a $\mathcal{O}(70\text{ ps})$ time resolution and with an efficiency above 95%, and the kaon/pion separation is better than 10^4, meaning that for each mis-identified pion there are more than 10^4 correctly identified kaons.

GTK. The GTK or GigaTracKer is the beam spectrometer and measures the momentum, the direction of flight and the time of beam particles, with experimental resolutions of 0.15 GeV/c, 16 μrad and 100 ps, respectively. It is located inside the vacuum pipe immediately upstream of the fiducial volume for the reconstruction of K^+ decays, and is composed of three stations (GTK1, GTK2, GTK3) of a hybrid silicon pixel matrix installed around four dipole magnets arranged as an achromat, as shown in Figure 2.4. Starting from Run2, a fourth station as been added in the upper harm of the achromat, close to GTK1, allowing to improve the efficiency of the track reconstruction and to reduce fake tracks. Each station constitutes a tracking plan with 18000 pixels of area $300 \times 300\ \mu\text{m}^2$ and a total thickness of 300 μm. To be noted that, regardless of its energy, the initial trajectory of a charged particle that crosses an achromat is unchanged when it exits the magnetic system. The achromat, therefore, allows measurements to be made on the NA62 beam while maintaining its divergence. Achromat dipoles deflect trajectories in the yz plane and the particle momentum can be derived from the vertical displacement (Δy) of its trajectory between GTK1 and GTK2. For K^+ of 75 GeV/c momentum (p) traveling along the z axis, the kick on the transversal momentum component (p_T) given by the magnetic field corresponds to a vertical displacement of $\Delta y \propto p_T/p = -60\,\text{mm}$. For Run1 data, the particle direction of flight is derived from the (x, y) impact points in GTK1 and GTK3 with no redundancy; during Run2, GTK0 measurements provide extra information for the reconstruction of the momentum.

The detector features very challenging design specifications. The 750MHz $\sim$ GHz nominal rate of beam particles (hence the name of GigaTracKer) requires a nominal hit time resolution $\sigma_t < 200\,\text{ps}$ [68]. Moreover, the material budget must be limited to $0,5\% X_0 \sim 500\,\mu\text{m}$ of silicon per tracking plane to reduce the probability of multiple scattering and obtain the required momentum and angular resolutions, and to reduce inelastic interactions in GTK3. Also, because of an operating environment characterized by the vacuum and by a high level of radiation, each station must be equipped with an efficient cooling system that carries away the heat produced by the readout electronics and mitigate radiation damage effects without substantially increase the material in the region crossed by the beam. To fulfill these specifications, innovative technologies were employed in the domain of silicon hybrid pixel technology and micro-channel cooling [69].

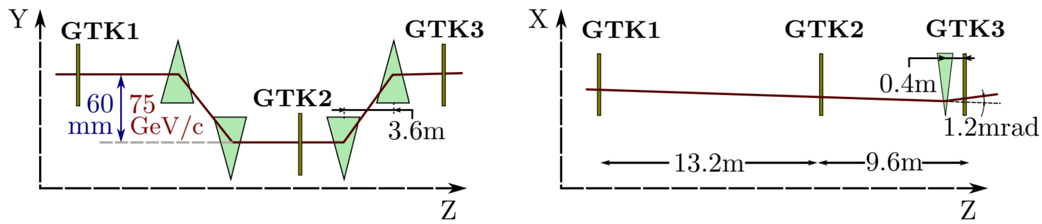


Figure 2.4: Schematic view of the beam spectrometer setup adopted during Run1 [64], in the yz (right) and xz (left) planes. Before the beginning of Run2 a fourth station was installed close to GTK1. Green triangles represent bending magnets: in the yz plane, an achromat system allows the measure the magnitude of the beam particles momentum via the vertical deflection of their trajectory; in the xz plane, the beam is deflected by 1.2 mrad by the TRIM5 magnet to compensate the effect of the STRAW spectrometer dipole magnet.

COLLIMATOR. A steel collimator precedes the GTK3 to absorb background particles surrounding the beam. In 2016, 2017 and during the first part of the 2018 run period, a 1 m thick collimator was used with inner aperture of $66 \times 33\ \text{mm}^2$ and outer dimensions of about 15 cm, and it served as partial shield against hadrons produced by upstream K^+ decays. Shortly after the start of the 2018 run period, the collimator was replaced by a 1.2 m thick collimator, with outer dimensions 1.7×1.8 m^2 and a central aperture of 76×40 mm^2, designed to completely absorb upstream decay hadrons passing outside his aperture.

CHANTI. The CHarged ANTI-coincidence detector (CHANTI) is installed immediately after the GTK3, at the beginning of the fiducial volume, and provides veto for charged hadrons resulting from inelastic interactions in GTK3 that enters the acceptance of the detector. Another main source of activity in the CHANTI comes from the particle halo close to the beam, that is mainly due to muons from the decay to $\mu^+\nu_\mu$ of beam K^+ or π^+ while traveling from the target to the detector. The CHANTI is composed of six stations of 300×300 mm^2 with a 95×65 mm^2 central hole for beam passage. All the CHANTI stations are located, together with the GTK3 station, inside the same vacuum vessel; their distances along the z axis are such that the angular region between 49 mrad and 1.34 rad is hermetically covered for particles generated in GTK3. Each station constitutes an hodoscope made of 48 scintillator bars of triangular cross section and has two readout planes, with the bars oriented vertically and horizontally to form x and y views. The bars are read out with fast wavelength-shifting (WLS) fibers coupled to silicon hotomultipliers (SiPMs). The hit position resolution is about 2.5 mm, the hit time resolution is 1 ns and the CHANTI efficiency is about 99%.

2.5.2 Charged decay products tracking and identification

STRAW. The STRAW spectrometer tracks the charged products of the K^+ decays with a momentum resolution $\sigma(p)/p$ between 0.3% and 0.4% and an angular resolution ranging from 60 μrad at 10 GeV/c to 20 μrad at 50 GeV/c. It is located 20 m downstream of the decay volume and consists of four straw chambers, two on each side of a large aperture dipole magnet (MNP33). To minimize multiple scattering, the straw chambers are realized with ultra-light material and are installed in vacuum. The magnetic field mainly points toward negative y: the integrated field is ~ 0.9 Tm for the vertical component and a factor 10^{-3} smaller for the complementary component.

Each straw chamber is composed of two modules and four views: one module contains the x and y views, measuring the (x, y) coordinates of the track passage in the chamber, the other contains the u and v views, measuring the (u, v) coordinates rotated around z by $(-45°)$ with respect to (x, y) (Figure 2.5 on the left). Every chamber contains 1792 straws made of 36 μm thick polyethylene terephthalate coated with 50 nm of copper outside and 20 nm of gold inside. Each straw has a 9.82 mm diameter, is 2160 mm long and has a gold-plated tungsten anode wire of 30 μm diameter. The gas inside the straw tubes is a mixture of 70% Ar and 30% CO_2 at atmospheric pressure. The straws geometry is based on two double layers per view with sufficient overlap to guarantee at least two hits per view (Figure 2.5 on the right), as needed to solve the left-right ambiguity. Near the middle of each view, few straws are left out, forming an octagon shaped free passage of 6 cm apothem for the beam passage after overlaying the four of them. As the beam has an angle of $+1.2$ mrad and -3.6 mrad in the horizontal

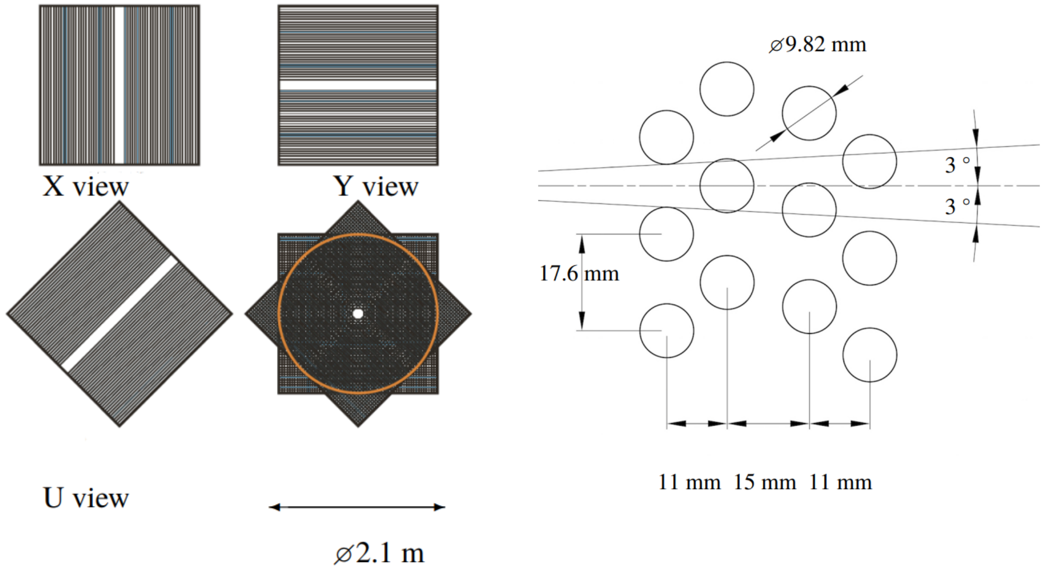


Figure 2.5: Left: a straw chamber and its four views (x, y, u, v). Right: the straws assembly geometry; the 3° angle corresponds to the angular range of tracks produced in kaon decays and detected within the spectrometer's geometrical acceptance.

plane, upstream and downstream of the magnet, respectively, this hole is not centered on the z axis, but has offsets along the x direction for each chamber.

RICH. The NA62 Ring Imaging CHerenkov detector (or RICH) is designed to separate pions from muons in the momentum range between 15 to 35 GeV/c and to provide a muon suppression factor of at least $\mathcal{O}(10^2)$ as part of the $\mathcal{O}(10^7)$ overall rejection factor from PID, needed to suppress the $K^+ \to \mu^+\nu_\mu$ mode in the $K^+ \to \pi^+\nu\bar{\nu}$ measurement. Moreover, the RICH provides a measurement of the time of charged particles crossing it with a precision of $\mathcal{O}(70\,\text{ps})$ and provides one of the main signals to trigger K^+ decays with a single downstream charged track.

As shown in Figure 2.6, the RICH is constituted by a 17.5 m long cylindrical vessel, made of ferro-pearlitic steel, filled with neon gas and crossed by an evacuated beam pipe. The beam pipe allows undecayed beam particles to travel without interaction with the active volume of the detector. Neon gas is chosen as Cherenkov radiator because, at atmospheric pressure and room temperature, it has a refractive index of $(n-1) = 62 \times 10^{-6}$, corresponding to a pion threshold momentum for Cherenkov emission of 12.5 GeV/c. With this choices, the RICH detector is fully efficient for 15 GeV/c momentum pions. Moreover, neon has good UV transparency and low chromatic dispersion. The vessel has four sections of gradually decreasing diameter and different lengths. The downstream end has 3.2 m diameter and houses a mosaic of spherical mirrors with their support panel. The RICH mirror system [70] reflects towards the

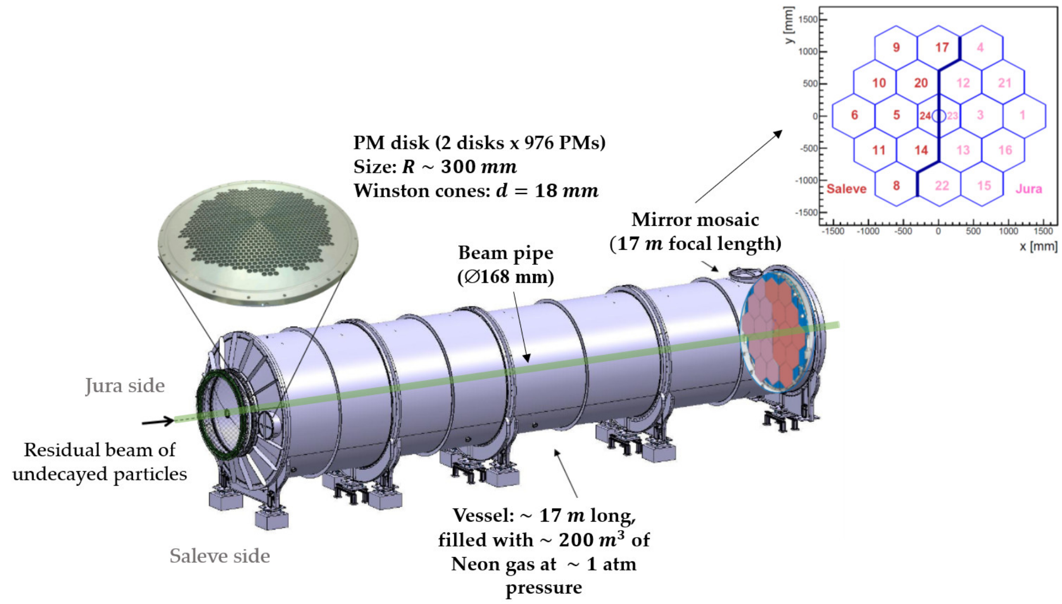


Figure 2.6: Schematic of the RICH detector and of the mirror system. The mirrors are divided into two groups, the Jura (in light pink) and the Saleve ones (in dark pink), pointing to the left and to the right of the beam pipe, which crosses the center of the mosaic.

upstream end the Cherenkov light cones produced by charged particles traversing the radiator. The mosaic contains 18 hexagonal mirrors (350 mm side) and 2 semi-hexagonal mirrors, that can be aligned by using a system of piezoelectric motors with a precision of $\mathcal{O}(30\,\mu\text{rad})$ in terms of mirror orientation [71]. The two semi-hexagonal mirrors are adjacent to the beam pipe and have a circular opening to accommodate it. The mirrors have a nominal radius of curvature of 34 m, hence a focal length of 17 m. To avoid absorption of reflected light by the beam pipe, the mirrors are divided into two spherical surfaces: one with the center of curvature to the left and one to the right of the beam pipe. The upstream end of the vessel is about 4.2 m wide to accommodate, outside the active area of the detector, on the focal plane of the two mirrors spherical surfaces, two disks instrumented with photomultiplier tubes (PMTs) for Cherenkov light detection [72]. To optimize the angular resolution, a good granularity of the photon detection system is needed; as a reasonable compromise between the number of sensors, the photon acceptance and the cost, the two disks contain a total of 1952 single anode PMTs with UV glass entrance. The photon detection sensitivity starts at a wave length of 185 nm and peaks at a wave length of 420 nm.

Charged Hodoscopes. The CHOD and the New-CHOD detectors are located downstream and upstream of the LAV12 detector, respectively, and are operated simultaneously and independently. The CHOD exploits a high granularity design that provides track timing with $\mathcal{O}(200\,\text{ps})$ precision, while the expected time resolution for

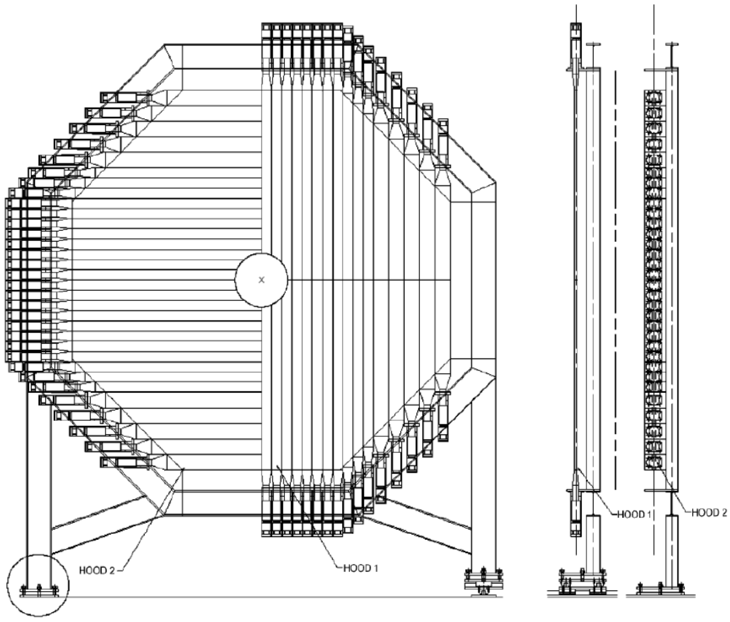


Figure 2.7: Schematic of the CHOD detector [64]. Half sections of the horizontal and vertical planes are shown, together with the light detection system. The center of the detector houses the beam pipe, with the beam traversing the detector from the vertical plane towards the horizontal plane.

New-CHOD signals is of $\mathcal{O}(1\,\text{ns})$.

The CHOD was used in the former kaon experiment NA48 [73]. As shown in Figure 2.7, it is composed of two consecutive orthogonal planes of 64 vertical and 64 horizontal BC408 plastic scintillator slabs, respectively. Slabs are 20 mm ($0.10\,X_0$) thick and are assembled into 4 quadrants forming an octagon of 1210 mm apothem; each slab is read out by a photomultiplier through a fishtail-shaped Plexiglas light guide. Two independent time measurements are provided by the CHOD for each charged particle crossing a vertical and an horizontal slab, reducing possible tails in the event time distribution due to out-of-time accidental activity, interactions in the upstream material and back-scattering from the LKr calorimeter.

The New-CHOD is a newly constructed charged hodoscope optimized for the high intensity conditions of NA62. It has a single plane of 152 plastic scintillator tiles of 30 mm thickness and a finer tile configuration in the high occupancy area near the beam axis, as shown in Figure 2.8. The overall lateral dimensions of the New-CHOD are those of an octagonal box with 1550 mm apothem, but its active area is an annulus with inner and outer radii of 140 mm and 1070 mm, respectively. The tile centers are spaced vertically resulting in a 1 mm overlap; this displacement is obtained mounting rows of tiles alternatively front and back of a thin central support panel. The resulting total thickness of the New-CHOD detector in the active area is $0.13\,X_0$. The matrix of tiles is read out by SiPMs, the scintillation light is collected and transmitted by 1 mm diameter wavelength shifting fibers.

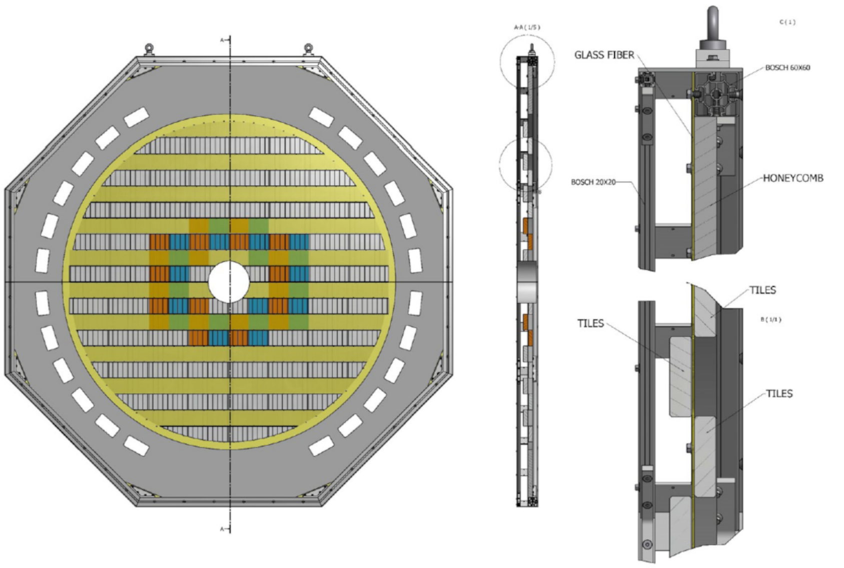


Figure 2.8: Schematic of the New-CHOD detector [64]. It is mounted on the front face of LAV12 (left). The scintillator tiles are mounted front and back of a thin G10 support panel (bottom right). The structure is stiffened at the periphery with honeycomb and aluminum construction profiles (top right).

2.5.3 Photon veto system.

A system of photon-veto detectors provides hermetic coverage for photons produced in K^+ decays occurred in the fiducial volume and propagating at angles up to about 50 mrad with respect to the detector axis.

Ring-shaped Large Angle Veto (LAV) detectors are arranged around the perimeter of the sensitive volume: 11 stations are placed in vacuum along the fiducial volume while a 12th station is located about 3 m upstream of the LKr calorimeter and is operated in air. Detector positions are chosen in order to achieve full acceptance for photons emitted at polar angles between 10 and 50 mrad. Each LAV station is an electromagnetic calorimeter (EC) made of 37 cm long Schott SF57 lead-glass blocks (density $\rho = 5.5\,\mathrm{g/cm^3}$ and a radiation length $X_0 = 1.50\,\mathrm{cm}$) recycled from the OPAL EC barrel [74]. Electromagnetic showers in the lead glass are detected via the Cherenkov light produced by the e^+/e^- composing the shower and by the ionization energy loss. The blocks are arranged around the perimeter of the sensitive volume of the experiment, with the blocks aligned radially to form an inward-facing ring. Multiple rings and a staggering scheme between the block in each ring are used in each station in order to provide the depth required for the efficient detection of incident particles. As a result, particles incident on any station are intercepted by blocks in at least three rings (at least 21 X_0), or by four or more blocks (at least 27 X_0) in most of the stations. The activity in the LAV stations is typically used for vetoing the presence of particles emitted at large angles, in-time with the signals of interest. At nominal beam intensity, the rates on the LAV detectors are dominated by the muon halo; for this accidental rate to be kept to the percent level with a $\pm 5\sigma$ coincidence window, the LAVs must have a time resolution of $\mathcal{O}(1\,\text{ns})$ for 1 GeV photons. An energy resolution of at least 10% for 1 GeV photons is needed for the definition of the veto threshold.

The NA48 liquid krypton EC (LKr) [73] detects photons emitted at polar angles between 1 and 10 mrad, providing energy, position and time measurements for their energy deposits. In order to be able to satisfy the demanding rate requirements in NA62, however, the external components of the cryogenic system and the auxiliary parts of the readout system were modified [64]. The LKr energy resolution in the NA62 conditions is $\sigma(E)/E = 1.4\%$ for energy deposits of 25 GeV; its spatial and time resolutions are 1 mm and between 0.5 and 1 ns respectively, depending on the amount and type of energy released. The LKr is a quasi-homogeneous calorimeter filled with about 9000 liters of liquid krypton at 120 K, housed inside a cryostat. The calorimeter extends from the beam pipe ($r \sim 8\,\text{cm}$) to a radius of 128 cm; its depth is 127 cm, corresponding to 27 X_0. The sensitive area is divided into 13248 longitudinal cells with a cross section of about 2×2 cm^2. The cells are formed by $Cu - Be$ electrodes aligned along the longitudinal axis of the experiment, and have a zig-zag shape to avoid inefficiencies when the particle shower is very close to the anode. The signal produced by a particle crossing the LKr is collected by preamplifiers inside the cryostat, directly attached to the calorimeter strips.

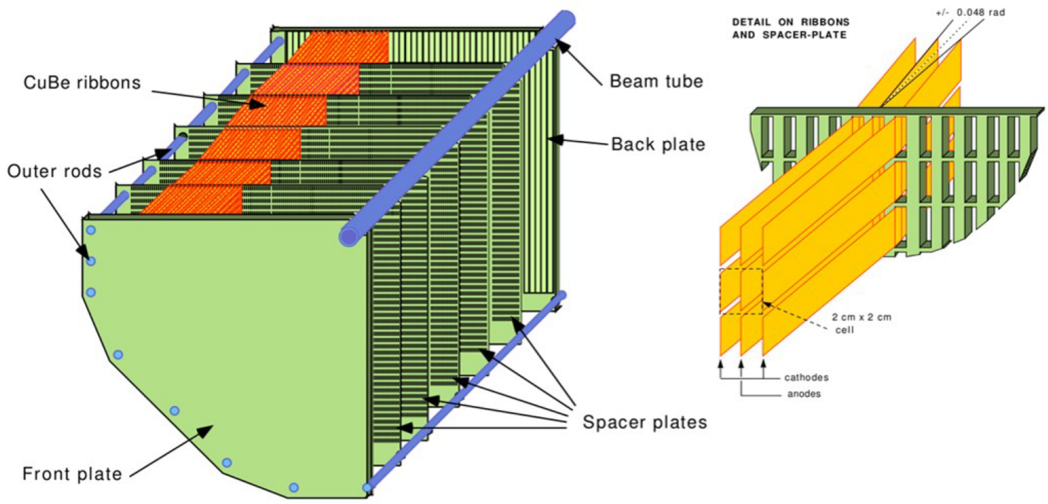


Figure 2.9: Sketch of the LKr detector [64]. On the left hand side, the schematic of the calorimeter structure (one quadrant) is reported. The detail of the calorimeter cells is shown on the right.

The Intermediate Ring Calorimeter (IRC) and the Small Angle Calorimeter (SAC) constitute the small angle veto system for the detection of photons emitted at angles between 0 and 1 mrad. The IRC is located in front of the LKr and is divided into two longitudinal modules, upstream and downstream module. The upstream (downstream) module is 89 mm (154 mm) long and contain 25 (45) ring-shaped layers forming a sampling calorimeter. Each layer consists of a 1.5 mm thick lead absorber alternating with 1.5 mm thick scintillator plate. The total depth of the calorimeter is 19 X_0. The SAC detector is installed inside the beam vacuum just upstream of the beam dump. Similarly to the IRC detector, it is a lead/scintillator sampling calorimeter consisting of 70 plates of lead and 70 plates of plastic scintillator, both of 1.5 mm thick, for a total depth of 19 X_0.

2.5.4 Muon veto system

The MuonVeto System (MUV) complements the pion/muon identification system and consists of an hadronic sampling calorimeter, placed immediately after the LKr and formed by two independent detectors (MUV1 and MUV2), an 80 cm filtering iron wall and a fast array of scintillator tiles located behind it (MUV3).

The MUV1 detector was realized for the NA62 experiment with a fine transversal segmentation, that allows to better disentangle the hadronic and electromagnetic shower components, while the MUV2 detector, was inherited from the NA48 experiment. Both the MUV1 and MUV2 use iron plates as absorber layers, interleaved with scintillator strips sensitive layers, with consecutive scintillators layers alternately aligned in the horizontal and vertical direction. The two detectors assembly corresponds to about 8 interaction lengths, while the whole calorimetric system (LKr+MUV1+MUV2) provides 14 interaction lengths to be traversed by pions and muons from K^+ decays.

The MUV3 detector, because of its position, detects charged particles traversing the whole calorimetric system (LKr, MUV1 and MUV2) and the iron wall. At typical NA62 daughter particles energy scales, only muons have still enough energy after calorimeters to cross the iron wall and hit the MUV3 detector, so that signals from MUV3 are used to identify muons. Moreover, the detector provides fast signals useful at trigger level. The MUV3 has a transverse size of 2640×2640 mm^2 and is composed of 50 mm thick scintillator tiles, including 140 regular tiles of 220×220 mm^2 transverse dimensions and 8 smaller tiles adjacent to the beam pipe, as required by the high particle rate near the beam. The total rate of muons traversing the MUV3 at the nominal beam intensity is 13 MHz, with a typical 3.2 MHz hottest tile close to the beam pipe, dominated by muons from beam pion decays. The MUV3 time resolution is of $\mathcal{O}(400\,\text{ps})$, and its muon identification efficiency exceeds 99.5% for muon momenta above 15 GeV/c.

2.5.5 Additional veto detectors

Two additional counters, MUV0 and HASC, are installed at optimized locations to provide nearly hermetic coverage for charged particles produced in multi-track kaon decays escaping the acceptance of the STRAW chambers. The MUV0 detector is a scintillator hodoscope designed to detect π^- emitted in $K^+ \to \pi^+\pi^-\pi^+$ decays, with momentum below 10 GeV/c, deflected towards positive x by the spectrometer magnet and leaving the lateral acceptance near the RICH. It is mounted on the downstream flange of the RICH, covering the periphery of the lateral acceptance at positive x, and consists of two layers of 48 plastic scintillator tiles with dimensions of $200 \times 200 \times 10$ mm^3. The HAdronic Sampling Calorimeter (HASC) is used for the detection of π^+ emitted in $K^+ \to \pi^+\pi^-\pi^+$ decays with momentum above 50 GeV/c and propagating through the beam holes in the centers of the STRAW chambers. The detector is located downstream of the MUV3 and to the dipole magnet used to dump the undecayed beam particles, which sweeps these pions out of the K^+ beam towards negative x. The HASC is constructed of 9 identical modules and the active element of a module is a sandwich of 60 lead plates alternated with 60 scintillators plates of section 100×100 mm^2 and 4 mm thickness. In 2018, the $K^+ \to \pi^+\nu\bar{\nu}$ analysis revealed that the HASC is also effective at mitigating backgrounds from certain configurations of the $K^+ \to \pi^+\pi^0$ decay. Therefore, for the Run2, a second HASC station was placed on the opposite side of the beam axis with respect to the first one, in order to double the π^0-rejection.

For the Run2, three new veto counter stations (VetoCounter) and a new hodoscope (ANTI-0) were also installed in the upstream part of the beam line, both of them playing a key role for the detection of K^+ decays occurring upstream of the fiducial volume, that were found to provide one the dominant background to the $K^+ \to \pi^+\nu\bar{\nu}$ signal from the analysis of Run1 data. The VetoCounter was designed to detect the products of K^+ decays occurring within the GTK, typically charged pions and photons, before they are absorbed by the collimator. It surrounds the beam pipe, with two stations upstream and one downstream of the collimator. The second plane is preceded by a layer of lead, that acts as photon converter, with the products of the interaction detected by the second layer. The first and third planes allows to identify minimum ionizing particles and help to distinguish between kaon decay products and muon halo. Each station is a rectangular plane of plastic scintillator tiles; the scintillation light is collected on both sides of each tile by fast photomultipliers coupled through light-guides. The ANTI-0 detector [75] is a plane of scintillator tiles, each of area 124×124 mm^2 and 10 mm thickness, arranged in an octagonal shape of 1082 mm apothem. It is installed at the entrance of the fiducial volume and detects charged particles up to ~ 1 m far from the beam line.

2.6 Trigger and Data acquisition systems

The intense flux of the NA62 beam requires a high-performance triggering and data acquisition system, which must minimize dead time while maximizing data collection reliability. Moreover, the broad physics program demands a selective and flexible trigger. For most of the detectors, a unified trigger and data acquisition (TDAQ) system was designed to meet these requirements, that will be briefly described in the following considering the Run1 operating conditions.

The NA62 trigger system is divided in two levels. The first one (called L0) is an hardware trigger, based on the inputs coming from a small set of fast detectors [76]. The aim of the L0 trigger is to reduce the $\mathcal{O}(10\,\text{MHz})$ input rate, the estimated rate of decays in the NA62 detector, down to $\mathcal{O}(1\,\text{MHz})$, with a fixed latency of 1 ms. The second level, called L1, is a software trigger based on the input sent from a different subset of detectors after the response of the L0 [55]. The L1 must reduce the data rate from the $\mathcal{O}(1\,\mathrm{MHz})$ of the L0 output down to the $\mathcal{O}(100\,\mathrm{kHz})$ that can be saved on disk. At L0 there are 10 different trigger lines, suitable for different physics analysis. The goal of each line is to select certain topology of events, such as decays to single-track or multi-track events with or without muons or positrons, with or without energy release in the LKr, and so on. The same structure is propagated at L1, where different algorithms run, depending on the topology selected from the L0. The main line is the one specifically setup for the $K^+ \to \pi^+\nu\bar{\nu}$ analysis, called PNN trigger, that is described below as an example. The PNN L0 trigger exploits the logical signals (primitives) produced by the RICH, New-CHOD, LKr, and MUV3 detectors. The temporal coincidence of signals from at least tree RICH PMTs in a 25 ns time window defines the trigger time and tags a charged particle. The trigger time is used as a reference to define a coincidence within 6.3 ns of: a signal in one up to four New-CHOD tiles; no signals in opposite New-CHOD quadrants to suppress multi-track K^+ decays; no signals in the MUV3 to reject $K^+ \to \mu^+\nu_\mu$ decays; less than 30 GeV deposited energy and no more than one cluster in the LKr, to reject $K^+ \to \pi^+\pi^0$ decays. A dedicated board combines the primitives to build and dispatch trigger decisions to the detectors for data readout. If all the conditions that defines the L0 trigger are satisfied, all the subdetectors answering to the L0 trigger send data to the NA62 PC-Farm, where the L1 software algorithms are running. The PNN L1 trigger exploits quantity that are available at software level, coming from KTAG, STRAW, MUV3 and LAV, aiming to discard events with multitracks or events with a single track with momentum outside the range (8, 50) GeV/c or coming from K^+ decays occurred outside the fiducial decay volume. If all the conditions are satisfied, the trigger is sent to all the other detectors and, when all the answers are received by the NA62 PC-farm, all the information coming from the L0 and from the L1 are merged and stored on disk.

For this work, the line of interest is the minimum-bias one, also referred to as Control (CTRL) trigger, a very loose trigger selection collecting a wide and representative sample of events. In particular, the CTRL trigger selects events with at least one charged track. Based on primitives coming from the CHOD, it requires the presence of at least 2 hits firing in a time window of 12.5 μs at L0 and no further conditions is applied at L1. Due to the high rate of such a trigger, a downscaling factor of 400 is needed in order to not saturate the global L0 bandwidth. Minimum-bias data collected by the CTRL trigger are typically used to determine the K^+ flux, to measure trigger efficiencies and to estimate backgrounds. Also, CTRL trigger data may be further filtered offline and used for physics analysis.

Chapter 3

Analysis strategy and dataset

This work is an experimental study of the $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ process at the NA62 K^+ factory, to search for the radiative production in K^+ decays of a new force mediator (X) with mass below the EW scale, neutral with respect to the SM and with enhanced coupling to muons. The existence of such a particle could play a key role, for example, in explaining the measured value of the muon anomalous magnetic moment, as described in Chapter 1.

In particular, X is assumed to be a scalar particle that decays promptly into a photons pair. This hypothesis allows to constrain the photons from $X \to \gamma\gamma$ to the K^+ decay vertex, as will be explained in the following. From the theory, X could be also long-lived, but the search for displaced $X \to \gamma\gamma$ is in principle not feasible at NA62, as the decay point or the mass of a neutral particle must be known to reconstruct it starting from photons. If the X vertex is displaced with respect to the K^+ vertex, nor the X decay point neither its mass can be reconstructed, hence X is not detectable through the di-photon final state. It has to be mentioned, however, that the prompt search can be sensitive to displaced $X \to \gamma\gamma$ decays with decay lengths comparable to the K^+ vertex resolution, of $\mathcal{O}(1\ \mathrm{m})$. The focus of this work will be the definition of a selection for the $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ signal and the reduction of the main backgrounds to this mode for prompt X decays; the extension of the study to the displaced decays compatible with the decay vertex resolution is left for future studies.

In this Chapter, the strategy adopted for the analysis is explained and the NA62 data samples and the used Monte Carlo (MC) simulations are described.

3.1 $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ analysis strategy

The $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ prompt decay chain is experimentally characterized by a single charged track identified as a μ^+ and two electromagnetic energy deposits (clusters) in the LKr calorimeter produced by the two photons. The clusters must be spatially unmatched with the charged track but associated in time with it. Also, in the hypothesis that X decays promptly, the photons can be assumed to originate from the same position of the K^+ decay vertex.

Let $P_{\gamma_i} \equiv (E_{\gamma i}, \overrightarrow{p}_{\gamma i})$ be the four-momentum of the photon i ($i = 1, 2$) and

$$m_{\gamma\gamma} \equiv \sqrt{(P_{\gamma_1} + P_{\gamma_2})^2} = \sqrt{2(E_{\gamma_1}E_{\gamma_2} - \overrightarrow{p}_{\gamma_1} \cdot \overrightarrow{p}_{\gamma_2})} \tag{3.1}$$

the invariant mass of the two photons, where $E_{\gamma i}$ is the energy of the LKr cluster associated to the photon i; $\overrightarrow{p}_{\gamma i} = E_{\gamma_i}\hat{n}_{\gamma_i}$ is the three-momentum of the photon i, with the direction of flight $\hat{n}_{\gamma_i}$ determined by using the measured position at the LKr front plane of the cluster i and the reconstructed K^+ decay vertex position. As a resonance is expected in the invariant mass spectrum at the X mass value (m_X), $m_{\gamma\gamma}$ is chosen as the main discriminating variable for the $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ signal search. To exclude the expected $\pi^0 \to \gamma\gamma$ resonance, two regions sensitive to the X production are defined in the $m_{\gamma\gamma}$ distribution (also referred to as signal regions here on):

$$\text{R1: } 0 < m_{\gamma\gamma} < 110\,\text{MeV/c}^2 \tag{3.2}$$

$$\text{R2: } 160 < m_{\gamma\gamma} < 380\,\text{MeV/c}^2 \tag{3.3}$$

where the upper limit of R2 is the kinematic bound $m_X < (m_K - m_\mu)$ from the $K^+ \to \mu^+\nu_\mu X$ decay. The signal regions are expected to be also populated by the following processes:

- $K^+ \to \mu^+\nu_\mu\pi^0$ **decay** ($K_{\mu3}$). Due to the prompt $\pi^0 \to \gamma\gamma$ decay, it is expected to have the same experimental signature as signal if $m_X = m_{\pi^0}$. As a consequence, $K_{\mu3}$ events satisfy the same conditions selecting signal events. If well reconstructed, they populate the π^0 resonance in the $m_{\gamma\gamma}$ spectrum; otherwise they mimic signal events with $m_X \neq m_{\pi^0}$. As the $K_{\mu3}$ has a branching ratio of $\sim 3.3\%$, this mode is expected to be one of the dominant backgrounds both in R1 and in R2.

- $K^+ \to \pi^+\pi^0$ **decay** ($K_{2\pi}$). It can contribute to the background in two different ways: if the π^+ is misidentified as a muon or if the π^+ decays-in-flight to $\mu^+\nu_\mu$. While the first contribution can be suppressed by the particle identification conditions requested for the track, the second one provides, in principle, an irreducible background.

- $K^+ \to \pi^+\pi^0\pi^0$ **decay** ($K_{3\pi0}$). In R1 and R2 it is expected to produce an almost uniform distribution in $m_{\gamma\gamma}$ due to the combinatorial probability of choosing LKr clusters not coming from the same π^0 and to photons emitted at very large angles going undetected. As for the $K_{2\pi}$ decay, different contributions come from the π^+ mis-identification as μ^+ and from the $\pi^+ \to \mu^+\nu_\mu$ decay-in-flight.

- $K^+ \to \mu^+\nu_\mu$ **decay** ($K_{\mu2}$). Neglecting radiative corrections, there are no final state photons in this decay, so its experimental signature is different from the signal one. However, even this channel can mimic the signal due to accidental photons in time with the muon and constitutes an important source of background because of its BR ($\sim 63\%$).

- $K^+ \to \pi^+\pi^+\pi^-$ **decay** ($K_{3\pi}$). It passes the signal selection due to: one π^+ mis-identified as a μ^+ or decaying-in-flight; the remaining $\pi^+\pi^-$ pair, not detected by the STRAW spectrometer, producing clusters in the LKr.

- **Photon conversions in e^+e^- pairs**. When a photon coming from a K^+ decay crosses the downstream detectors, it can interact with their materials and convert in e^+e^- pair before getting to the LKr calorimeter. Depending on the z coordinate of the photon conversion, different scenarios may occur, not all of them producing a well reconstructed π^0 or X.

- **Upstream decays**. The K^+ decays of interest for physics analyses are the ones occurring in the fiducial volume of the detector (see Chapter 2.4). However, K^+ decays occurring in the region upstream of the fiducial volume may constitute a non-negligible background component.

A cut-based $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ signal selection is defined, with specific conditions for the reduction of the different backgrounds studied by using Monte Carlo simulations. A peak-search procedure is then adopted to get a model-independent measurement of $BR(K^+ \to \mu^+\nu_\mu X,\, X \to \gamma\gamma)$ for the X mass hypotheses included in R1 and R2:

$$BR(K^+ \to \mu^+\nu_\mu X,\, X \to \gamma\gamma) = \frac{N_{sig}(m_X)\, D_{trigger}}{N_K\, A(m_X)\, \epsilon_{trigger}} \tag{3.4}$$

where, for a fixed X mass hypothesis, $N_{sig}(m_X)$ is the observed number of signal events and $A(m_X)$ is the acceptance of the signal selection; N_K is the number of K^+ decays occurring in the considered fiducial volume; $\epsilon_{trigger}$ is the trigger efficiency and $D_{trigger}$ the trigger downscaling.

As the branching ratio in Eq.3.4 depends on the flux of K^+ decays, which is not known a priori, a normalization procedure has to be defined. N_K is estimated selecting a clean sample of events of a known SM K^+ decay with a topology similar to the signal, referred to as normalization channel. For this search, the $K^+ \to \mu^+\nu_\mu\pi^0$ decay is chosen as normalization channel, as it is an abundant mode having exactly the same topology as the signal (as mentioned above), that is collected by the same trigger line and that can be selected by a normalization selection defined by slightly changing the signal one. Given the $K_{\mu3}$ acceptance ($A^{K_{\mu3}}_{norm}$) after the normalization selection, estimated from simulations, and the number of observed data events in the normalization sample (N^{data}_{norm}):

$$N_K = \frac{N^{data}_{norm}\, D_{trigger}}{BR(K^+ \to \mu^+\nu_\mu\pi^0) BR(\pi^0 \to \gamma\gamma) A^{K_{\mu3}}_{norm}\, \epsilon_{trigger}}. \tag{3.5}$$

In this approach the signal decay rate is measured with respect to the normalization channel and systematic effects due to residual detector inefficiencies not fully accounted for in simulations, as well as trigger inefficiencies and random veto losses, common to signal and normalization modes, cancel at the first order.

3.2 NA62 data sample

For the $K^+ \to \mu^+ \nu_\mu X$, $X \to \gamma\gamma$ search, NA62 data collected during Run1 by using the minimum bias trigger line (see 2.6) and filtered offline are used. One of the available NA62 filtering algorithms, the *REStricted MUon* (RESMU) filter, is considered. The RESMU filter is a set of loose conditions selecting events that are likely to include K^+ decays to a final state that includes a muon.

An event collected by the CTRL trigger satisfies the RESMU filter if at least one out of its reconstructed tracks:

- is in the acceptance of each of the four STRAW spectrometer chambers, with the number of straw chambers used to build the track equal to four. The χ^2 of the fit used for the track reconstruction is required to be < 40;
- it is in the LKr acceptance;
- it has positive and unitary electric charge;
- it has a measured momentum $p < 70\,\mathrm{GeV/c}$ and $E/p < 0.25$, with E energy of the LKr cluster associated to the track.
- it has at least one spatially associated candidate in the MUV3 muon detector. The geometrical association of a MUV3 signal with the track will be detailed in Sec.4.1.
- has a closest distance of approach to the nominal beam axis $CDA_{nom} < 60\,\mathrm{mm}$ and Z_{nom} included in the range (105, 180) m, with Z_{nom} coordinate along the z axis of the middle point between the track and the nominal beam axis at their closest distance of approach. This ensures, very loosely and considering the information available for the daughter track only, the compatibility of the track with a charged product from the decay of a beam particle.

Table 3.1 reports the details about the total size of the filtered dataset used, providing $\sim 11.4 \times 10^9$ events to be analyzed.

	events in RESMU sample	days of data taking	mean beam intensity [protons per spill]
2016	0.4×10^9	30	11×10^{11}
2017	4×10^9	161	16×10^{11}
2018	7×10^9	217	20×10^{11}

Table 3.1: Dimension of the NA62 data samples analyzed for the $K^+ \to \mu^+ \nu_\mu X$, $X \to \gamma\gamma$ search. Data were collected using the CTRL trigger line and were filtered by the RESMU filter described in Sec.3.2. During 2017 and 2018, the beam intensity gradually increased over the time; the value reported in the table is a representative mean value.

3.3 Monte Carlo samples

The acceptance of the $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ signal and the suppression of the backgrounds sources listed in Sec.3.1 are studied by Monte Carlo (MC) simulations centrally produced on the NA62 Grid [77]. The simulation framework is based on the CERN ROOT [78] and GEANT4 [79] toolkits.

To take into account the pile-up of the K^+ decay of interest with other beam particles and their interactions, all the MC samples used for this analysis are produced with the NA62 overlay technique, described in detail in Appendix A. When the physics event of interest (main event) is generated, additional MC events are overlaid to it (overlaid events); the number and the topology of the overlaid events is determined according to the beam intensity and the beam composition as measured from data. The main event is generated according to the K^+ mean lifetime, but only decays occurring in the z range between 102.5 and 180 m are saved. This range defines the standard decay region and starts just after the last station of the GTK and ends right before the first STRAW chamber.

Specific features, described later, have also been implemented to boost the statistical power of the simulations.

3.3.1 $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ signal samples

$K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ events are generated in two steps. First, the $K^+ \to \mu^+\nu_\mu X$ decay is simulated. Then, aiming to a model-independent $BR(K^+ \to \mu^+\nu_\mu X,\ X \to \gamma\gamma)$ with a prompt X decay, for each event, X is forced to decay into a $\gamma\gamma$ pair, assuming a $BR(X \to \gamma\gamma) = 100\%$ and a null proper mean lifetime ($\tau_X = 0$).

The $K^+ \to \mu^+\nu_\mu X$ decay is simulated as follows. Assuming the same conventions as in [31] and after the integration over angles (see [63]), the partial width for the $K^+ \to \mu^+\nu_\mu X$ decay can be generically written as:

$$\Gamma(K^+ \to \mu^+\nu_\mu X) = \frac{1}{256\,\pi^3 m_K^3} \int |\mathcal{M}|^2 dm_{12}^2 dm_{23}^2 \tag{3.6}$$

where

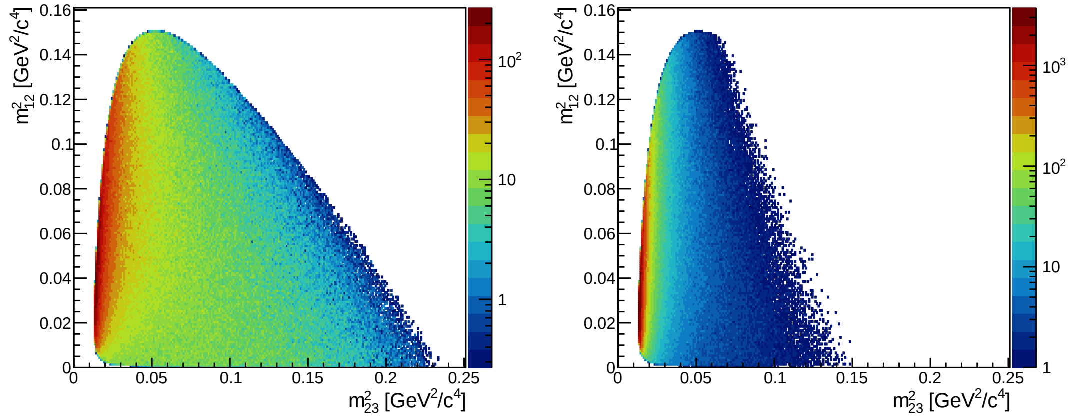
$$m_{12}^2 = (P_{\nu_\mu} + P_X)^2 \tag{3.7}$$

and

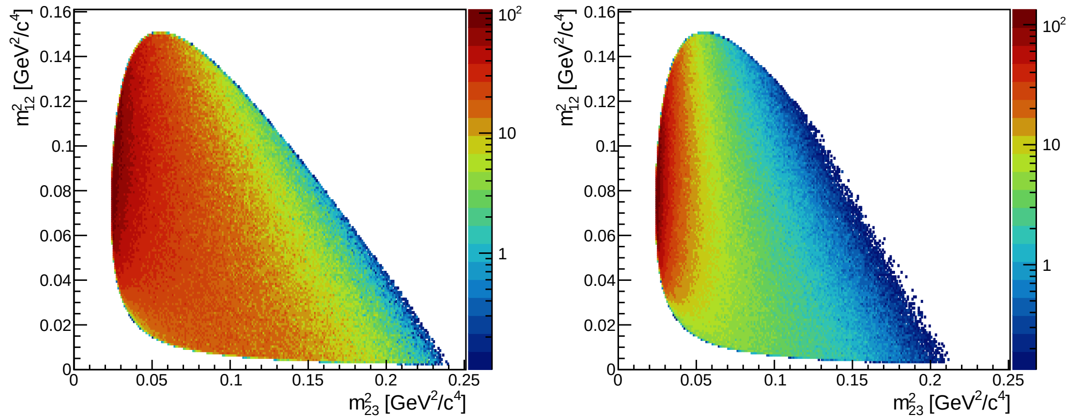
$$m_{23}^2 = (P_X + P_\mu)^2 \tag{3.8}$$

with P_i four-momentum of the particle in the subscript, are the two Lorentz invariant variables describing the available phase space of the process, and $\mathcal{M}$ is the matrix element describing its dynamics. The final state particles are first generated in the K^+ rest frame, by using the CERN ROOT *TGenPhaseSpace* class, assuming the phase space of a three body decay. Subsequently, the event is weighted by the probability of X to be radiated by the muon, as provided by the matrix element.

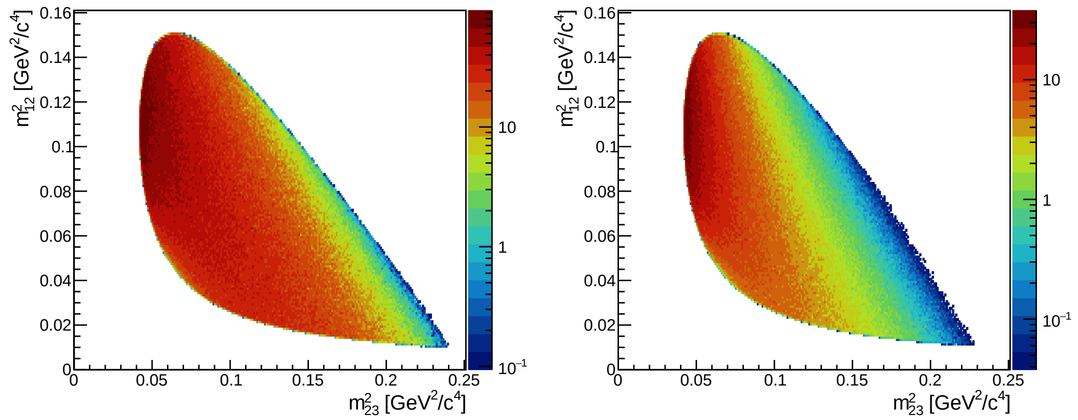
Figure 3.1 and Figure 3.2 show the phase space distributions for different X mass hypotheses included in R1 and R2, respectively, obtained by running a toy MC



(a) $m_X = 10\text{MeV}/\text{c}^2$

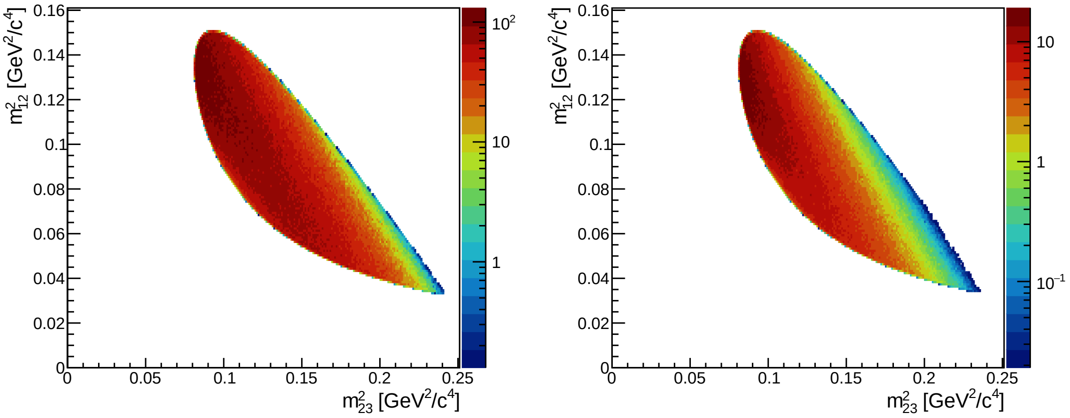


(b) $m_X = 50\text{MeV}/\text{c}^2$

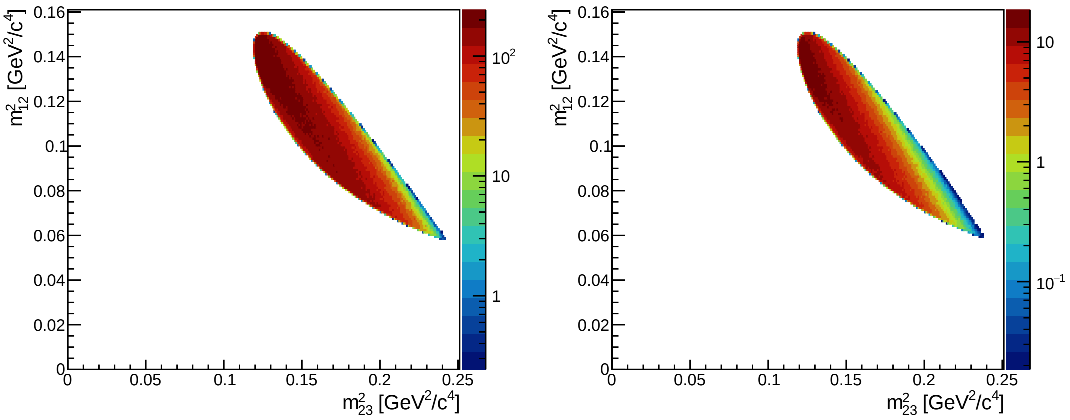


(c) $m_X = 100\text{MeV}/\text{c}^2$

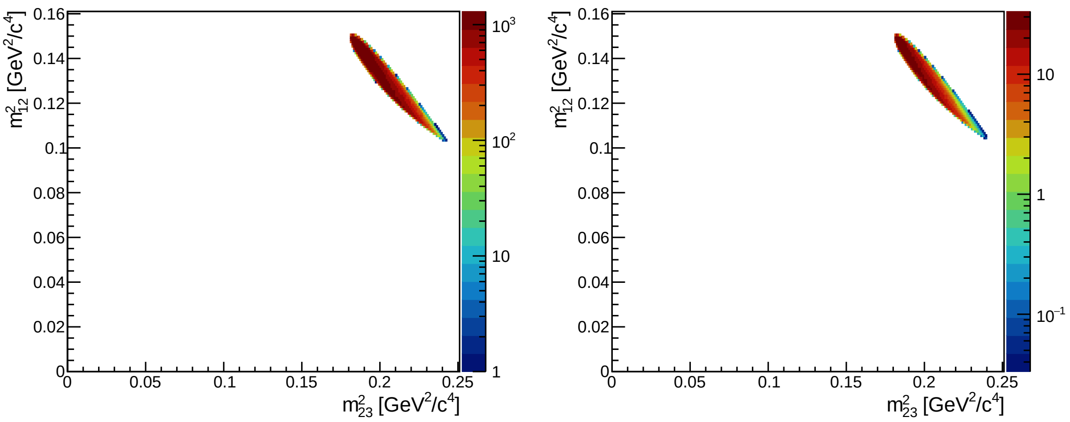
Figure 3.1: Dalitz plots obtained requiring 10^6 simulated $K^+ \to \mu^+\nu_\mu X$ events for different X mass hypotheses included in R1 ($m_X < 110\,\text{MeV}/\text{c}^2$). For each mass, the plot on the left hand side reports the phase space distribution as obtained by considering the kinematic of the three body decay only. The right plot shows how the distribution on the left is modified by the dynamics of the process (as described by the matrix element in Eq.3.11).



(a) $m_X = 180 \mathrm{MeV/c^2}$



(b) $m_X = 240 \mathrm{MeV/c^2}$



(c) $m_X = 320 \mathrm{MeV/c^2}$

Figure 3.2: Dalitz plots obtained requiring 10^6 simulated $K^+ \to \mu^+ \nu_\mu X$ events for different X mass hypotheses included in R2 ($160 < m_X < 380\,\mathrm{MeV/c^2}$). For each mass, the plot on the left hand side reports the phase space distribution as obtained by considering the kinematic of the three body decay only. The right plot shows how the distribution on the left is modified by the dynamics of the process (as described by the matrix element in Eq.3.11).

simulation with 10^6 events requested for each mass value. In the left column, the phase space distributions as obtained by the kinematic of the three body decay are shown. The boundaries are determined by the conservation of the total four-momentum:

$$(m_{12}^2)_{min} = m_X^2; \quad (m_{12}^2)_{max} = (m_K - m_\mu)^2. \tag{3.9}$$

For fixed m_{12}, the range of m_{23} is determined by its values when the three-momenta of X and μ^+ are parallel or anti-parallel:

$$(m_{23}^2)_{max}^{min} = (E_2^* + E_3^*)^2 - \left(\sqrt{E_2^{*2} - m_X^2} \pm \sqrt{E_3^{*2} - m_\mu^2}\right)^2 \tag{3.10}$$

with $E_2^* = (m_{12}^2 + m_X^2)/2m_{12}$ and $E_3^* = (m_K^2 - m_{12}^2 - m_\mu^2)/2m_{12}$ energies of X and μ^+ in the $X - \nu_\mu$ rest frame, respectively. Therefore, the total number of available states decreases as a function of m_X, decreasing the overall decay probability. The right column of the plots shows how the phase space on the left is modified due to the dynamics of the process. As expected for radiative processes, the favored states for X are shifted towards lower m_{23}^2 with respect to the pure three body decay. The squared matrix element describing the K^+ decay into a muon, neutrino and a scalar particle radiated from the muon, can be written as:

$$\begin{aligned} |\mathcal{M}|^2 = &\frac{\lambda_\mu^2 y_\mu^2}{2m_\mu^2(m_{23}^2 - m_\mu^2)^2}[m_K^2(m_{23}^2 + m_\mu^2)^2 \\ &- m_{23}^2((m_{23}^2 - m_\mu^2)^2 + m_{12}^2(m_{23}^2 - m_\mu^2)) \\ &+ m_X^2(m_{23}^2 - m_\mu^2 m_K^2)]. \end{aligned} \tag{3.11}$$

In Eq. 3.11, λ_μ is the effective coupling arising in the width of the SM $K^+ \to \mu^+\nu_\mu$ decay ($\lambda_\mu \equiv 2G_F f_K m_\mu V_{us} \sim 8.7 \times 10^{-8}$, as in [31]). For an explicit calculation see, for example, Reference [80]. To be noted that the unknown quantity of interest is the $X - \mu$ coupling (y_μ), that determines the strength of the Yukawa interaction in Eq.1.4. As it defines a scale factor for a fixed X mass, it can be absorbed in the constant value $\lambda_\mu^2 y_\mu^2/2m_\mu^2$, that enters both the nominator and denominator of the signal acceptance $A(m_X)$ defined in Eq.3.4, and can therefore be neglected in the simulation.

To evaluate the acceptance $A(m_X)$, signal samples for different m_X were centrally produced using a custom feature, whereby a set of masses is uniformly generated within a single MC production. The overall dimension requested for the sample is 10M; the number of generated X masses is 38, from 10 MeV/c^2 up to the kinematic end point $m_K - m_\mu$, with a step of 10 MeV/c^2. Assuming $A(m_X)$ of $\mathcal{O}(1\%)$, the dimension of each sample, of $\mathcal{O}(32 \times 10^4)$, allows statistical uncertainties due to the size of the simulation of $\mathcal{O}(1\%)$. Figure 3.3 reports, as a function of the simulated m_X, the spectrum of the true muon momentum (p_μ) and the distributions of the true angle between the K^+ and X directions (θ_{KX}) and between the two photons from the X decay (θ_{12}) in the laboratory frame. In the top and central plots, kinematic boundaries are evident, strongly related to the phase space distributions reported in Figures 3.1 and 3.2. The

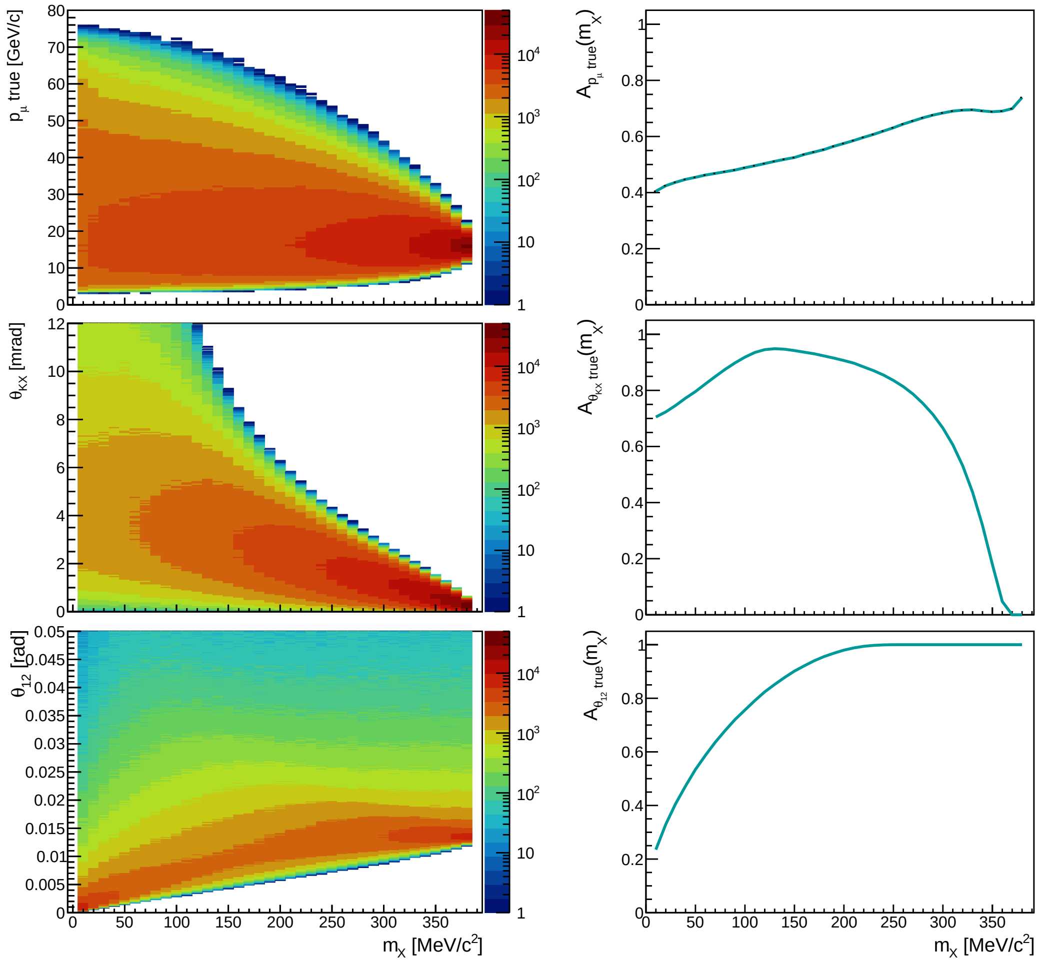


Figure 3.3: Left: the distributions of the true values for the muon momentum (top), for the angle of emission of X with respect to the K^+ direction of flight (center) and for the angle of aperture of the $\gamma\gamma$ pair (bottom) are reported as a function of the simulated mass m_X (from 10 MeV/c^2 to 380 MeV/c^2). The horizontal axis is the same for every plot and is reported only for the bottom one. Right: signal acceptances as a function of m_X assuming selection intervals for the corresponding variable on the left based on the detector characteristics. Details about the cuts are reported in the text.

bottom plot clearly shows the correlation $\theta_{12} \propto m_X^{-1}$ defining the minimum aperture of the photons pair, that comes from the four-momentum conservation $P_X = P_{\gamma 1} + P_{\gamma 2}$ in the small angle limit. These variables are the most affected by the characteristics of the NA62 detector. In particular, assuming $15 < p_\mu < 35\,\mathrm{GeV/c}$, $1 < \theta_{KX} < 10\,\mathrm{mrad}$ and $\theta_{12} > 0.0075\,\mathrm{rad}$, the expected signal acceptance is reported in the right column of the Figure 3.3. The momentum interval accounts for the range where the RICH provides the best π/μ separation. The interval for θ_{KX} comes from the LKr inner and outer geometrical acceptances. The range of the photon's aperture depends on the LKr clusters reconstruction, which is not able to properly separate energy deposits closer in space more than 150 mm. Figure 3.4 shows the $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ acceptance after requiring all these three conditions and taking into account the correlations among them. Because of the geometry and characteristics of the detector, the signal acceptance is largely reduced at very low and high masses, especially due to the angular cuts. To be noted that the acceptances reported in Figures 3.3 and 3.4 are based on conditions applied on the true values of p_μ, θ_{KX} and θ_{12}, therefore their trends should be considered as representative of the expected shapes only. The real acceptance values strongly depends on the reconstruction efficiency of momentum and directions. As an example, the distributions of the reconstructed di-photon invariant mass defined in Eq.3.1 is reported in Figure 3.5 for the simulated signal samples, after selecting a track identified as a muon and two clusters in the LKr identified as the $\gamma\gamma$ pair from the X decay. Different colors corresponds to a different simulated value of m_X, with a lower number of events reconstructed at very low and high masses and the detector resolution deteriorating the mass resolution with increasing mass.

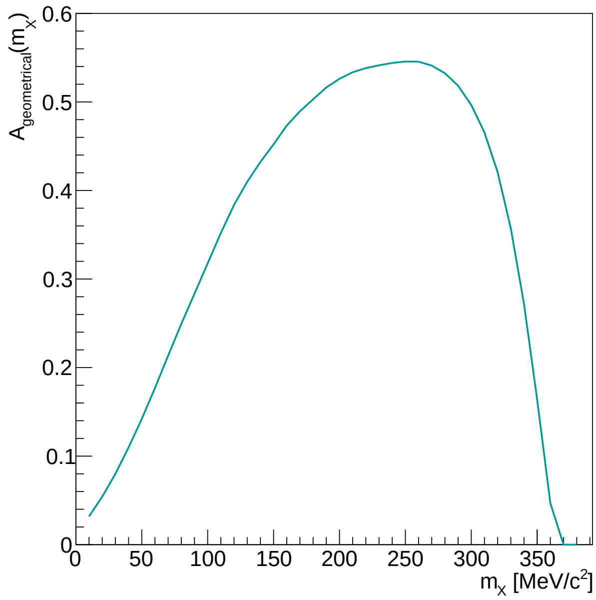


Figure 3.4: $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ acceptance after the conditions $15 < p_\mu < 35\,\mathrm{GeV/c}$, $1 < \theta_{KX} < 10\,\mathrm{mrad}$ and $\theta_{12} > 0.0075\,\mathrm{rad}$ are required on the true muon momentum, on the true angle of emission of X with respect to the K^+ direction of flight and on the true angle of aperture of the $\gamma\gamma$ pair, respectively. This trend provides information about the expected reduction of the signal acceptance due to the geometry and characteristics of the NA62 detector, but should be considered as representative only, as it is based on the true values of the simulated variables. The real signal acceptance strongly depends on the reconstruction efficiency of momentum and directions.

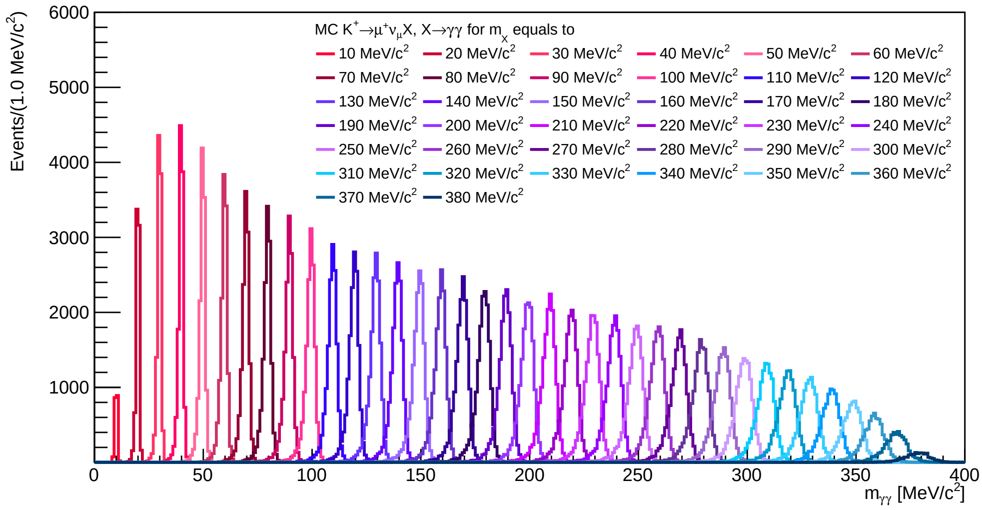


Figure 3.5: Distributions of the reconstructed $\gamma\gamma$ invariant mass for the simulated samples after selecting a track identified as a muon and two clusters in the LKr identified as the $\gamma\gamma$ pair from the X decay. The invariant mass is defined in Eq.3.1. The selection will be described in Chapter 4.

3.3.2 Backgroud samples

The suppression of the backgrounds to the $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ search is studied by using MC simulations. In particular, samples centrally produced of all the SM K^+ decay modes listed in Sec.3.1 are included.

To study the contributions coming from the $\pi^+ \to \mu^+\nu_\mu$ decay-in-flight, MC samples of the $K_{2\pi}$ and $K_{3\pi 0}$ decay modes were produced, in which the π^+ is forced to decay, as described in Appendix B. These samples are referred to as "capped" samples.

To study the contributions from upstream events that are mis-reconstructed with a vertex included the fiducial decay volume, a special production was requested of $K_{\mu 3}$, $K_{2\pi}$ capped and $K_{3\pi 0}$ capped samples in which events are saved if the z coordinate of the true K^+ decay point ($Z_{vtx\,true}$) is in the interval starting from a z included between the third and fourth bends of the GTK achromat and extending up to the GTK3 station ($[96.950, 102.425]$ m).

Radiative corrections are taken into account. Different models were used, depending on the decay mode. For the hadronic K^+ modes with three pions in the final states included in Table 3.2 (both standard and capped), the PHOTOS MC generator [81] is used to simulate QED corrections. For the $K^+ \to \pi^+\pi^0$ decay and the leptonic and semi-leptonic modes, the KLOE MC generator [82] is used instead.

A summary of the different samples, together with the dimension of each sample, is reported in Table 3.2.

	branching ratio	sample dimension	tag
$102.425 < Z_{vtx\,true} < 180\,\text{m}$ (standard K^+ decay region)			
$K^+ \to \mu^+\nu_\mu$	$(63.56 \pm 0.11)\%$ [63]	200M	$K_{\mu 2}$
$K^+ \to \pi^+\pi^0$	$(20.67 \pm 0.08)\%$ [63]	100M	$K_{2\pi}$
$K^+ \to \pi^+\pi^+\pi^-$	$(5.583 \pm 0.0024)\%$ [63]	400M	$K_{3\pi}$
$K^+ \to \pi^0 e^+\nu_e$	$(5.07 \pm 0.04)\%$ [63]	500M	K_{e3}
$K^+ \to \pi^0\mu^+\nu_\mu$	$(3.352 \pm 0.033)\%$ [63]	900M	$K_{\mu 3}$
$K^+ \to \pi^+\pi^0\pi^0$	$(1.760 \pm 0.023)\%$ [63]	100M	$K_{3\pi 0}$
$K^+ \to \pi^0\pi^0\mu^+\nu_\mu$	$(3.45 \pm 0.16) \times 10^{-6}$ [83]	10M	$K_{\mu 4}$
$K^+ \to \pi^+\pi^0$ with forced π^+ decay		400M	$K_{2\pi}$ capped
$K^+ \to \pi^+\pi^0\pi^0$ with forced π^+ decay		50M	$K_{3\pi 0}$ capped
$96.950 < Z_{vtx\,true} < 102.425\,\text{m}$ (upstream K^+ decay region)			
$K^+ \to \pi^0\mu^+\nu_\mu$		200M	$K_{\mu 3}$ ups
$K^+ \to \pi^+\pi^0$ with forced π^+ decay		11M	$K_{2\pi}$ ups capped
$K^+ \to \pi^+\pi^0\pi^0$ with forced π^+ decay		11M	$K_{3\pi 0}$ ups capped

Table 3.2: Details about Monte Carlo simulations of the K^+ decay modes providing background sources to the $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ search that are considered for the analysis. For the standard samples (top part of the table), the corresponding branching ratio are also reported. Specific samples of $K^+ \to \mu^+\nu_\mu\pi^0$, $K^+ \to \pi^+\pi^0$ and $K^+ \to \pi^+\pi^0\pi^0$ events were produced for the evaluation of background contribution coming from in-flight π^+ decays and decays occurring upstream of the standard decay region.

.

Chapter 4

$K\mu\gamma\gamma$ events selection

The $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ signal selection and the $K^+ \to \mu^+\nu_\mu\pi^0$ normalization selection are now introduced; they are both applied (separately) on the RESMU data sample described in Sec.3.2.

Events with only one reconstructed track in the STRAW spectrometer (downstream track) and at least two energy clusters in the LKr are considered for the analysis. The first requirement reduces the sample of about 50% but ensures a cleaner sample, excluding on top of the analysis events with multiple tracks in the final state.

The selection conditions can be grouped in three different categories, listed below.

$K\mu\gamma\gamma$ events selection. Given an event with only one downstream track, the quality of the track is checked and particle identification (PID) criteria are applied to distinguish muon tracks with respect to the pion and electron ones. Then, among the tracks reconstructed from the beam spectrometer (upstream tracks), the K^+ candidate that best matches in time and space to the muon track is chosen. Finally, two concurrent LKr clusters are associated to the $\gamma\gamma$ pair, satisfying good quality conditions and best matching in time with the μ^+ and the K^+ tracks, but not geometrically matching with the muon track at the LKr front plane. The selected track and clusters form a triplet.

Photon conversion rejection conditions. The presence in the detectors downstream of the fiducial volume of extra activity in-time with the triplet and correlated with at least one of the selected cluster is exploited to identify possible photon conversions ($\gamma \to e^+e^-$) in the detector material.

Kinematic constraints and photons veto. The conservation of the total four momentum is used to define a set of kinematic cuts reducing backgrounds from different K^+ decay modes. The presence of in-time activity in the photon veto system is used to discard events with extra photons.

All the conditions in the three categories define a common selection to signal and normalization, with the only exception being the cut on the $m^2_{miss}(K^+ - \pi^0)$, that is tighter for the normalization with respect to the signal selection, as will be specified in Sec.5.2.3. In addition to the common selection, the normalization one requires $110 < m_{\gamma\gamma} < 160\,\mathrm{MeV/c^2}$. This interval includes the π^0 mass and is complementary to the regions sensitive to signal defined in Eq.3.2 and 3.3 used in the signal selection.

In this Chapter, the triplet selection is described, while the study of the photon conversions and of the kinematic constraints and vetoes is reported in Chapter 5.

4.1 Muon track selection

Track good quality conditions. For the selected downstream track, some standard NA62 good quality conditions are required, ensuring: the quality of the track reconstruction, the compatibility with the assumption that it comes from the decay of a beam K^+; that the track is included in the geometrical acceptance of the downstream detectors. A good track is defined as follows:

- the track is in the acceptance of each of the four STRAW chambers, with the number of straw chambers used to build the track equal to four, and in the acceptance of the LKr, RICH, CHOD, New-CHOD and MUV3 detectors;
- it has positive and unitary electric charge;
- it has a reconstructed momentum $p_{track} < 70\,\mathrm{GeV/c}$, the χ^2 goodness of the fit for the track reconstruction < 20 and $|\overrightarrow{p}_{track} - \overrightarrow{p}_{track,\,before\,fit}| < 20\,\mathrm{GeV/c}$;
- has $CDA_{nom} < 60$ mm and Z_{nom} included in the range $(105,\,180)\,\mathrm{m}$, where CDA_{nom} and Z_{nom} are defined as in Sec.3.2.

It can be noted that this requirements strongly overlap with the RESMU filter described in Sec.3.2. The filter is always satisfied in data when asking for a single downstream reconstructed track. The same conditions need to be applied to simulated events to have exactly the same selection for both data and MC samples. However, as the muon track selection as a whole (good quality and PID conditions) has the same or more stringent requirements with respect to the filter, it is not necessary to apply it on simulations.

The reconstructed momentum of the good track is then required to be between 15 and 35 GeV/c, which is the optimal range for the particle identification with the NA62 RICH. As the track is required to be associated with a reconstructed Cherenkov ring in the RICH, the track reference time is the associated RICH ring time (t_{RICH}). Thanks to the $\mathcal{O}(100\,\mathrm{ps})$ RICH time resolution, this time measurement constitutes the most precise among the ones available for the charged track.

Particle identification. As some of the major background sources come from SM K^+ decays with a π^+ in the final state, the π/μ separation plays a key role in the $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ search. In this analysis, the muon PID relies on RICH, LKr and MUV3 signals.

- **MUV3 matching.** The selected STRAW track must have a time and space matching with a signal in the MUV3 muon detector (MUV3 candidate). The geometrical matching is determined by extrapolating the track to the MUV3 front plane and checking if the MUV3 candidate is within a search radius around the track impact point. The search radius is inversely proportional to the track momentum to account for multiple scattering effects on the track direction across the iron wall placed in front of the detector. The time matching is defined in a time window 6 ns large around the track time: $|t_{MUV3} - t_{RICH}| < 3$ ns. The distribution of the time difference $t_{MUV3} - t_{RICH}$ is reported in Figure 4.1 for a subsample of the NA62 Run 1 data.

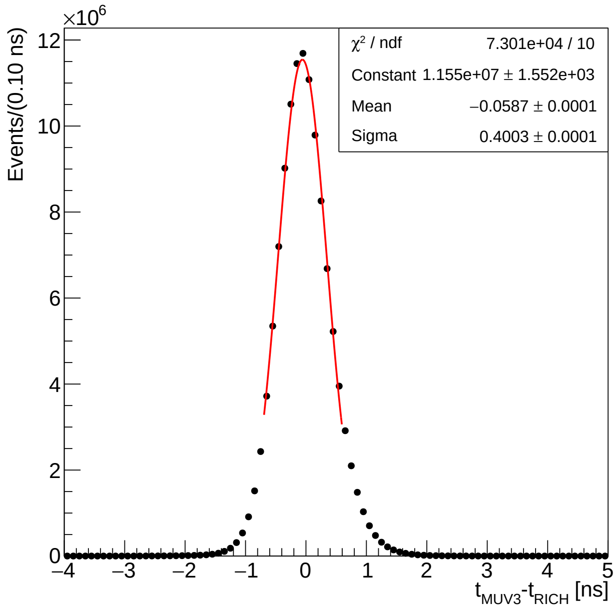


Figure 4.1: Distribution of the time difference $t_{MUV3} - t_{RICH}$ for a subsample of NA62 Run 1 data. t_{MUV3} is the time of the closest in time MUV3 candidate geometrically associated to the downstream track, t_{RICH} is the track reference time. A Gaussian fit to the peak is also reported.

- **RICH PID.** The procedure to associate a RICH mass hypothesis to a selected track is the following. The track is extrapolated to the RICH mirrors plane. Depending on the track direction, momentum and impact point in the mirror system, the expected position of the ring center at PMs plane and the expected ring radii for the electron, muon, pion, and kaon mass hypotheses are computed, together with the expected number of illuminated PMs (hits). For each mass hypothesis, the expected ring parameters and the position of the observed hits are used to compute a likelihood function. Five different numbers are then provided for the track, corresponding to the likelihoods of the track to be an electron ($\hat{L}_e$), a muon ($\hat{L}_\mu$), a pion ($\hat{L}_\pi$), a kaon ($\hat{L}_K$) or background ($\hat{L}_b$), where the background hypothesis can be interpreted as an heavy particle below the Cherenkov threshold. All the likelihoods are then divided by the maximum one ($L_i = \hat{L}_i/\max(\hat{L}_e, \hat{L}_\mu, \hat{L}_\pi, \hat{L}_K, \hat{L}_b)$), so that the normalized likelihood for the hypothesis with the highest likelihood is equal to 1 by construction. To be noted that the sum over the likelihoods is not 1.

In order to identify muons while efficiently rejecting the other mass hypotheses, the maximum not muon likelihood is defined:

$$\max L_{!\mu} = \max(L_e, L_\pi, L_K, L_b) \tag{4.1}$$

and the conditions:

$$\begin{cases} L_\mu = 1 \\ \max L_{!\mu} < 0.2 \end{cases} \tag{4.2}$$

are required. Due to the definition of L_i, these conditions are equivalent to a maximum likelihood ratio test of the muon hypothesis. The cut at 0.2 ensure a probability to mis-identify a pion as a muon smaller than 3% with an efficiency for the identification of muons greater than 80% for each selected track momentum, as shown in Figure 4.2.

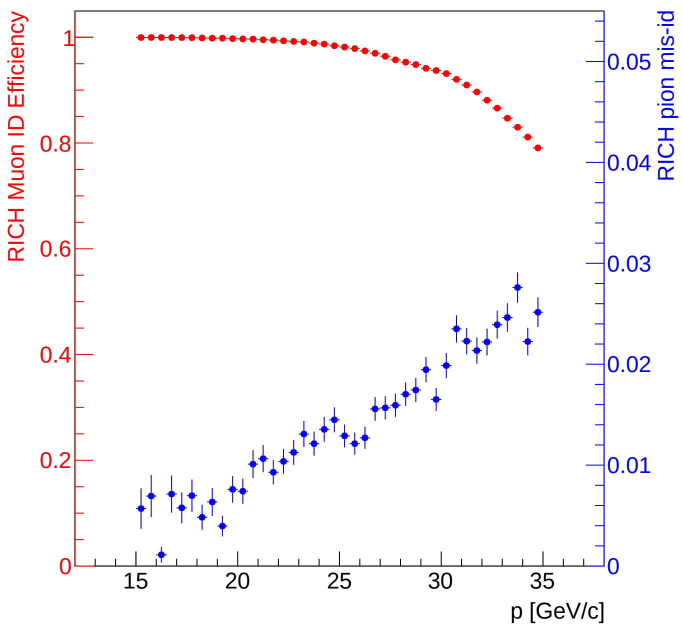


Figure 4.2: RICH muon identification efficiency (red) and pion mis-identification probability (blue) as a function of the track momentum for events with $L_\mu = 1$ and $\max L_{!\mu} < 0.2$ as measured from a sample of $K^+ \to \mu^+\nu_\mu$ and $K^+ \to \pi^+\pi^0$ data events. The observed trend of the pion mis-identification probability is related to the algorithm computing the likelihood and is amplified by the different scale with respect to the one of the muon identification probability.

- **E/p.** To identify muons, electrons and pions, the E/p ratio is considered, where E is the energy deposited from the charged particle in the LKr and p is the track momentum measured in the STRAW spectrometer. Its distribution is reported in Figure 4.3. Muons and electrons interact with matter only electromagnetically, loosing energy through ionization or generating electromagnetic showers, depending on their energy. Electrons with momentum in the range $15-35\,\mathrm{GeV/c}$ deposit all their energy in the LKr calorimeter, hence their E/p peaks at one, with a widening at larger values due to resolution effects. In the same momentum range, muons are minimum ionizing particles (MIP), and their E/p distribution peaks at zero. Nevertheless, a few percent of muons undergo Bremsstrahlung and the produced photons can start an electromagnetic shower, generating a tail at positive E/p values. High energy pions mostly produce hadronic showers dominated by QCD and nuclear processes. As the LKr calorimeter is not designed to fully contain hadronic showers, it often constitutes a pre-shower for pions, that are then absorbed from the hadronic calorimeters MUV1 and MUV2. The E/p distribution of pions therefore spreads all over the range (0, 1), depending on the position where the hadronic shower starts in the LKr; $\mathcal{O}(20-30\%)$ of pions start the shower in the hadronic calorimeters or produce an energy deposit in the LKr too small to be reconstructed as a cluster, generating an E/p close to zero.

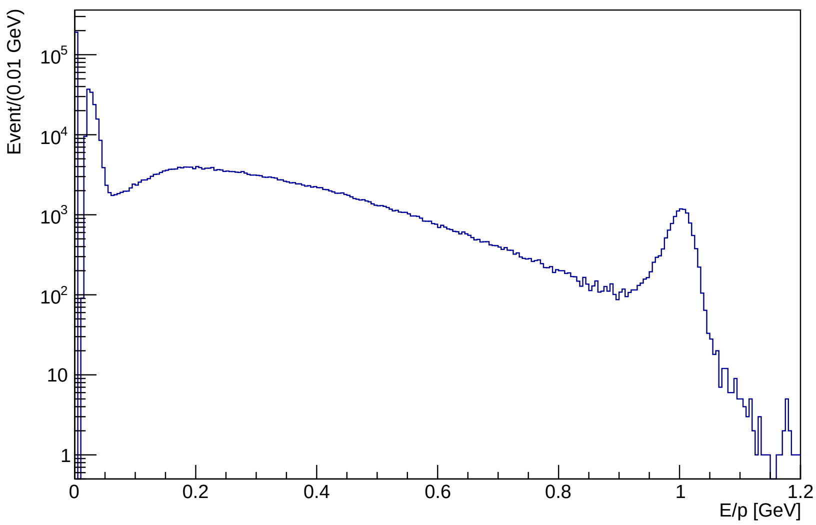


Figure 4.3: E/p distribution for tracks in data events selected by using a $K^+ \to \pi^+\pi^0$ selection before applying the PID conditions.

Muons can be efficiently identified against electrons by selecting small E/p values, with a residual contamination due to mis-identified pions. By requiring $E/p < 0.05$, the probability to mis-identify pions as muons is of $\mathcal{O}(60\%)$ at low track momentum and of $\mathcal{O}(40\%)$ at higher p, with a muon identification efficiency always higher than 87%, as shown in Figure 4.4. The measured pion mis-identification probability takes into account also pions without an associated LKr cluster.

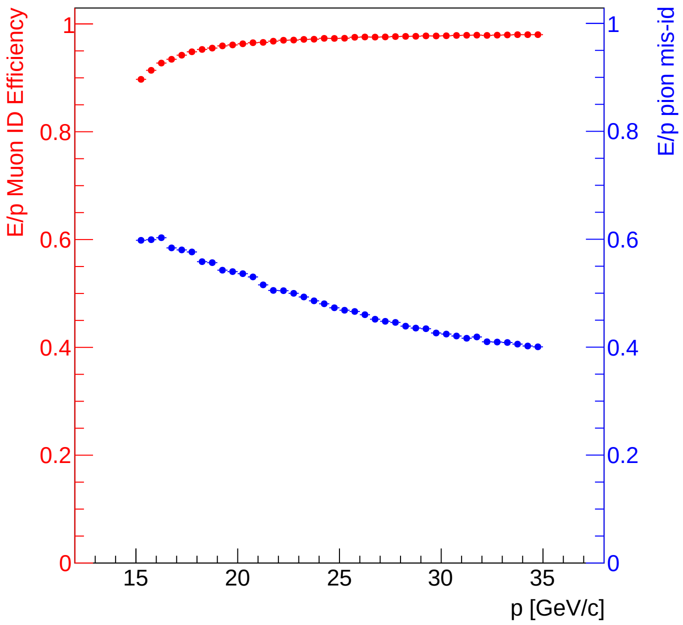


Figure 4.4: E/p muon identification efficiency (red) and pion mis-identification probability (blue) as a function of the track momentum for events with $E/p < 0.05$ as measured from a sample of $K^+ \to \mu^+\nu_\mu$ and $K^+ \to \pi^+\pi^0$ data events, respectively.

4.2 Matching between the μ^+ and K^+ candidate tracks

The track of the muon candidate is matched with an upstream kaon candidate to reconstruct the decay. Beam kaons are tagged via the KTAG and their direction and momentum are measured by the GTK beam spectrometer.

KTAG matching. A kaon is identified in the KTAG if coincident light in multiple sectors of its photon-detection system is detected. A 5-fold sector coincidence has been the standard kaon tagging requirement for the $K^+ \to \pi^+\nu\bar{\nu}$ analysis since 2016 and ensures a K^+ identification efficiency above 98%, while providing a separation between kaons and pions better than 10^4. The 5-fold condition is also used for this search. The time condition $|t_{KTAG} - t_{RICH}| < 0.6\,\text{ns}$, where t_{RICH} is the muon track reference time and t_{KTAG} is the time of a KTAG kaon candidate, is then required for the matching between the K^+ candidate and the selected downstream track. The distribution of the difference $t_{KTAG} - t_{RICH}$ is shown in Figure 4.5 for a subsample of the N62 Run1 data.

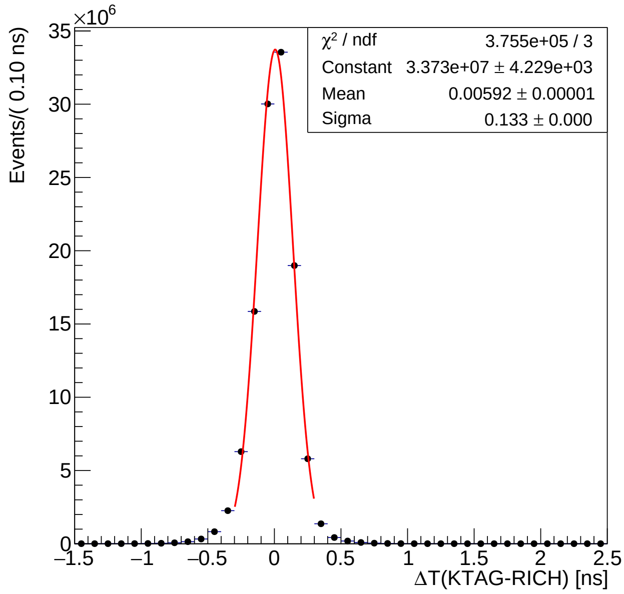


Figure 4.5: Distribution of the time difference $\Delta T(KTAG - RICH) \equiv t_{KTAG} - t_{RICH}$ for a subsample of NA62 Run 1 data. t_{KTAG} is the time of the KTAG candidate closest in time to the selected muon track and t_{RICH} is the muon track reference time. A Gaussian fit to the peak is also performed.

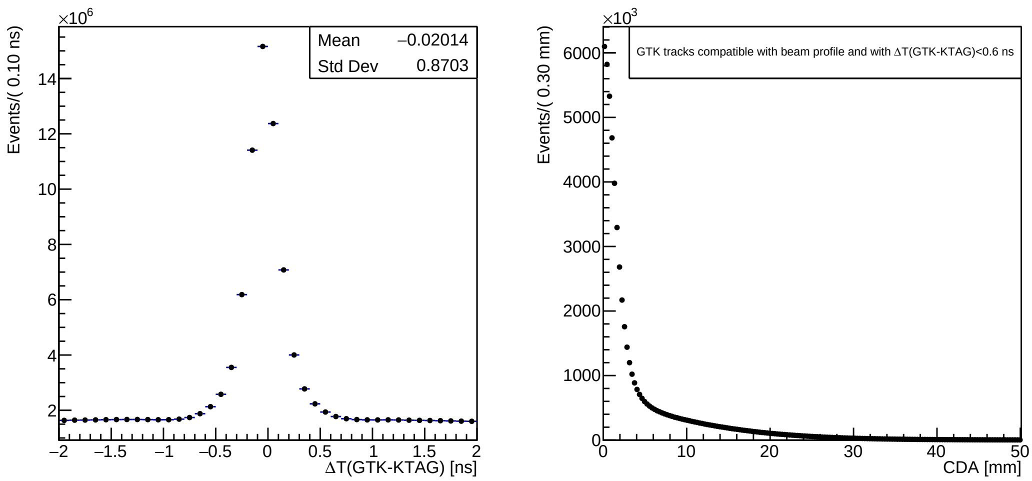


Figure 4.6: Left: Distribution of the time difference $\Delta T(GTK - KTAG) \equiv t_{GTK} - t_{KTAG}$ in a subsample of NA62 Run 1 data. Right: Distribution in the same sample of the CDA between the selected muon track and the GTK candidates compatible with the beam profile and $|t_{GTK} - t_{KTAG}| < 0.6$ ns.

GTK matching To match the reconstructed muon track with a GTK track, the procedure described below is followed.

First, the consistency of one GTK reconstructed track with the beam profile (in terms of momentum and direction) is required. Subsequently, the time coincidence between the GTK track and the previously identified KTAG candidate is required in a time window 0.6 ns large: $|t_{GTK} - t_{KTAG}| < 0.6$ ns. The distribution of $t_{GTK} - t_{KTAG}$ is reported in Figure 4.6 (left). The right hand side of Figure 4.6 shows the distribution of the closest distance of approach (CDA) between the selected muon track and the GTK candidates compatible with the beam profile and in time with the KTAG candidate.

The two variables, $\Delta T \equiv (t_{GTK} - t_{KTAG})$ and CDA, are then used to define a discriminant for the matching between the μ^+ and the K^+ candidates:

$$D(CDA, \Delta T) = [1 - p(CDA)] \cdot [1 - p(\Delta T)], \tag{4.3}$$

with $p(CDA)$ and $p(\Delta T)$ probabilities to observe the current value of CDA and ΔT, respectively:

$$p(CDA) = \frac{1}{N_{CDA}^{norm}} \int_0^{CDA} f(CDA')\, dCDA' \tag{4.4}$$

$$p(\Delta T) = \frac{1}{N_{\Delta T}^{norm}} \int_{-\Delta T}^{+\Delta T} f(\Delta T')\, d\Delta T' \tag{4.5}$$

The probability density functions (p.d.f.) for CDA ($f(CDA')$) and ΔT ($f(\Delta T')$) are template functions derived from a $K^+ \to \pi^+\pi^+\pi^-$ analysis, as described in reference [53], and are shown in Figure 4.7; N_{CDA}^{norm} and $N_{\Delta T}^{norm}$ are normalization factors.

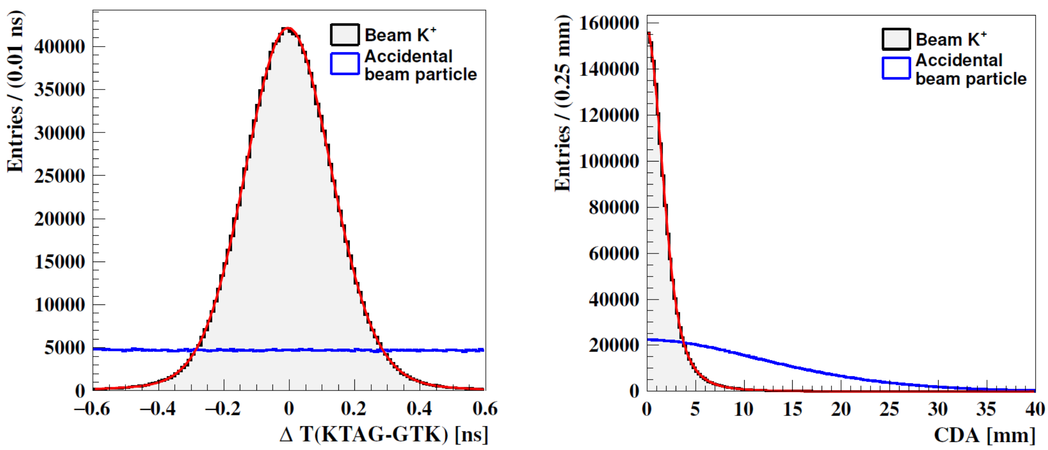


Figure 4.7: From [53]: distributions of $\Delta T = (t_{GTK} - t_{KTAG})$ (left) and CDA (right) for beam K^+ tracks (shaded gray histogram) and for accidental beam particle tracks (empty histogram). Beam K^+ tracks are tagged from fully reconstructed $K^+ \to \pi^+\pi^+\pi^-$ decays in the data. Pile-up GTK tracks are reconstructed in the same $K^+ \to \pi^+\pi^+\pi^-$ analysis as for beam K^+ track using out-of-time GTK signals. The red curves superimposed on the histograms describe the functions templates used to model the time and CDA distributions of the beam K^+. More details are reported in the paper.

By construction, the largest the discriminant, the higher the probability that the GTK track matches with the parent K^+, therefore GTK tracks compatible with the beam and satisfying the timing condition are sorted for increasing $D(CDA,\, \Delta T)$. The GTK track with the highest $D(CDA,\, \Delta T)$ is referred to as the most probable matched kaon track. However, because of the high particle rate in the beam tracker, a wrong association to the muon track may occur when a pile-up track leads to a discriminant value larger than the one of the parent K^+ track or when the parent K^+ track is not reconstructed in the GTK. Hence, when the selected GTK tracks have similar discriminant values, their probability to be a pile-up track (p_{PU}) is also assessed and the event is then discarded if the most probable matched kaon track is also the best matched pile-up track or if it is not possible to distinguish unambiguously if the track is a beam K^+ or an accidental particle. Finally, for the most probable matched kaon track, an additional discriminant is defined $D(CDA,\, \Delta T_{RICH})$, with $\Delta T_{RICH} \equiv (t_{GTK} - t_{RICH})$ and D as in Eq.4.3 with the appropriate p.d.f., $f(\Delta T_{RICH})$, to reinforce the coincidence between the GTK track and the muon track. The most probable kaon track matches the selected muon track if the following conditions are met:

- $\Delta T_{RICH} < 0.6\,$ns, to enforce the upstream-downstream time matching
- at least one between $D(CDA,\, \Delta T)$ and $D(CDA,\, \Delta T_{RICH}) > 0.03$
- $D(CDA,\, \Delta T) > 0.005$ and $D(CDA,\, \Delta T_{RICH}) > 0.005$
- CDA $<$ 7 mm

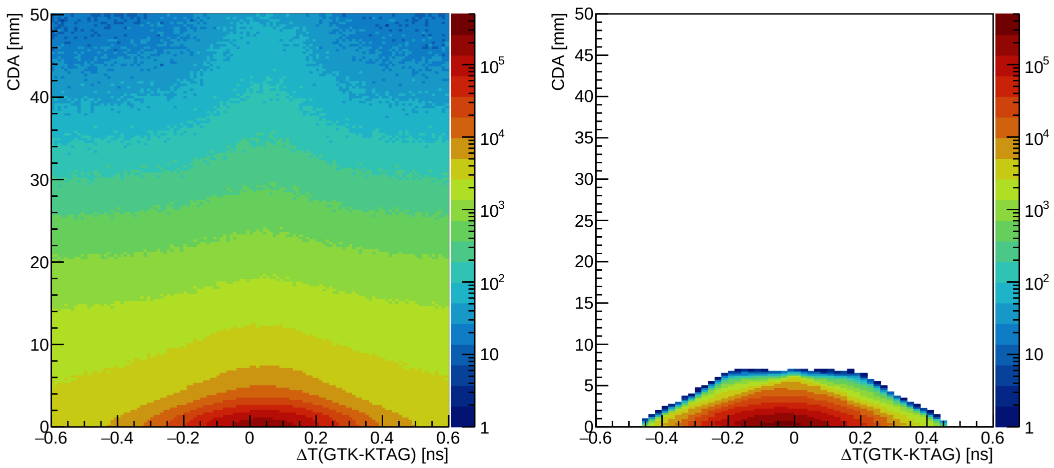


Figure 4.8: Left: CDA vs $\Delta T(GTK - KTAG)$ for a subsample of the NA62 Run1 data. The GTK reconstructed tracks considered for the plot are compatible with the beam profile and in time with the identified KTAG kaon candidate within 0.6 ns. Right: same distribution as on the left hand side after applying the matching conditions described in the text.

The last two conditions are equivalent to a trapezoidal cut in the $CDA - \Delta T$ plane, for both the $\Delta T(GTK - KTAG)$ and $\Delta T(GTK - RICH)$, as evident in Figure 4.8 for $\Delta T(GTK - KTAG)$.

Figure 4.9 reports the distribution of each component of the measured momentum for the selected K^+ track and the z coordinate of the reconstructed $K^+ - \mu^+$ charged vertex. The component along the z-axis of the K^+ momentum is around $75\,\mathrm{GeV/c}$, as expected for the particles forming the selected secondary hadron beam of NA62. The component along the x-axis is compatible with the kick provided in xz plane by the TRIM magnet (see Figure 2.4), while in the yz plane the achromat system preserves the beam direction, leading to an y component distributed around zero. The charged vertex is defined as the mid-point of the closest distance of approach between the selected K^+ and μ^+ tracks. Figure 4.9 on the bottom right shows that the reconstructed z component spreads all over the fiducial volume of the experiment for the selected tracks.

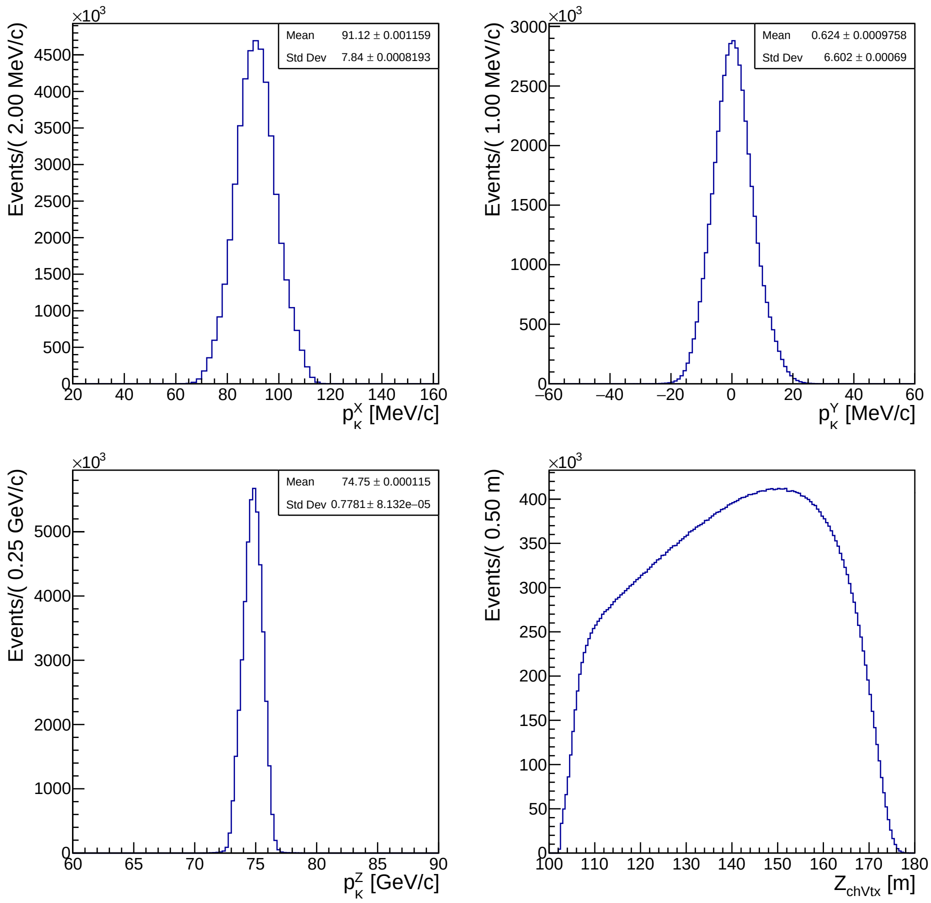


Figure 4.9: Distributions of the measured momentum of the selected K^+ track; the components p_K^X, p_K^Y and p_K^Z, are reported in the top left, top right and bottom left, respectively. The distribution of the z coordinate of the reconstructed $K^+ - \mu^+$ vertex is also shown (bottom right).

4.3 Photon clusters selection

After excluding the cluster associated to the selected track (if present), the couple of LKr clusters forming the best triplet with the track is selected, with the best triplet defined in the following. Some standard conditions are required for each cluster on top of the triplet choice to ensure good quality clusters to be used. A good quality cluster:

- is in the standard LKr acceptance. The transverse surface of the LKr has the shape of an octagon with a central internal circular hole crossed by the beam pipe (as shown in Figure 2.9). Let (x, y) be the cluster position in the xy plane and $R = \sqrt{x^2 + y^2}$ its distance with respect to the origin of the plane. The standard inner and outer LKr acceptances are defined as $R > 150\,\text{mm}$ and as $|x|, |y| < 1130\,\text{mm}$, $|x| + |y| < \sqrt{2} \cdot 1130\,\text{mm}$, respectively;

- it has energy $> 3\,\text{GeV}$, to ensure a reliable measurement of the energy deposit;

- it has a distance from the nearest LKr dead cell > 20 mm, to avoid energy losses due to the overlap of the shower with a not working cell;

- it is geometrically isolated with respect to any other in-time cluster, to reduce mis-reconstruction effects coming from the merging of energy deposits produced from different showers. The standard NA62 isolation condition requires that if another cluster exists within $\pm 7\,\text{ns}$, the distance between the selected cluster and the second one should be at least $150\,\text{mm}$. The distance depends on the reconstruction of the clusters; the width of the time window depends on the reconstruction and on the LKr time resolution, that varies with the energy.

For the clusters satisfying the good quality conditions, the time matching with the track is required. Let t_{cls} be the time of the cluster and E_{cls} its energy, the condition $|t_{RICH} - t_{cls}| < 3\sigma_{t\,LKr}$ must be satisfied, with $\sigma_{t\,LKr}$ the time resolution of the LKr calorimeter. The LKr time resolution can be parameterized [84] as a function of the cluster energy as:

$$\sigma_{t\,LKr}[\text{ns}] = 0.56 + \frac{1.53}{E_{cls}[\text{GeV}]} - \frac{0.233}{\sqrt{E_{cls}[\text{GeV}]}}, \tag{4.6}$$

The effect of the time matching condition defined with an energy dependent width is shown in Figure 4.10. To be noted that, due to the standard NA62 isolation condition, the good clusters in-time with the track are spatially separated by at least 150 mm from the cluster associated to the track. However, as for the muon track the energy deposit can be so low that it is not always possible to associate an LKr cluster to the track, a distance between the in-time good cluster and the extrapolation of the track to the LKr front plane $d(track - cls)$ is required to be $d(track - cls) > 150\,\text{mm}$, in order to reinforce the geometrical isolation of the cluster from the track.

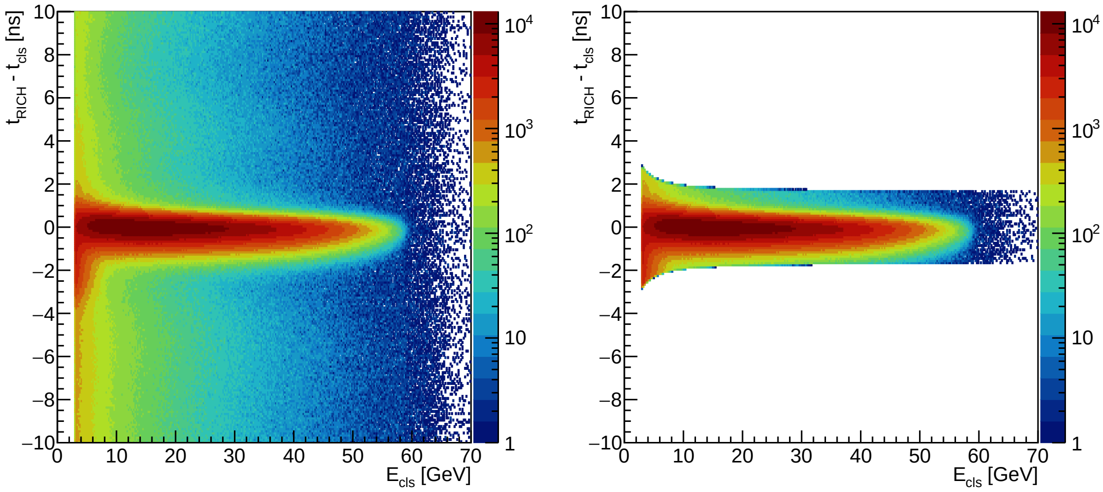


Figure 4.10: Left: Distribution of the time difference between the selected track reference time (t_{RICH}) and the time of a good cluster (t_{cls}) as a function of the cluster energy (E_{cls}). The good cluster is defined as mentioned in Sec.4.3; the condition on the minimum cluster energy generates the cutoff at 3 GeV/c in the plot. Right: effect of the time matching condition $|t_{RICH} - t_{cls}| < 3\sigma_{t\,LKr}$, with $\sigma_{t\,LKr}$ LKr time resolution. The shape of the distribution depends on the energy dependence of $\sigma_{t\,LKr}$ from the cluster energy, reported in Eq.4.6.

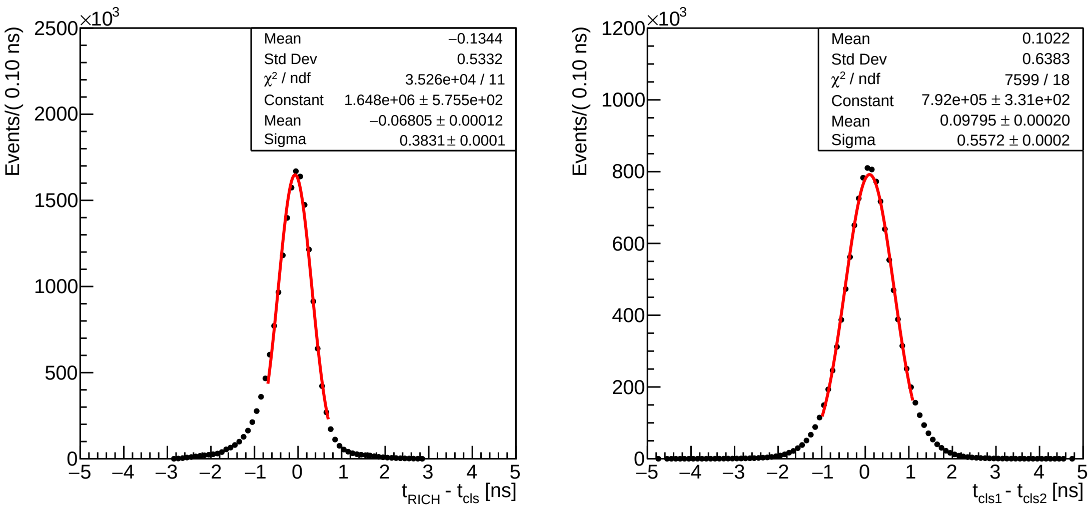


Figure 4.11: Left: distribution of the time difference between the selected track and a good cluster in-time with it according to $|t_{RICH} - t_{cls}| < 3\sigma_{t\,LKr}$. This distribution is the projection along the vertical axis of the one in the right hand side of Figure 4.10. Right: distribution of the time difference between two of the good clusters in-time with the track. Gaussian fits to the peaks of the distributions are also reported.

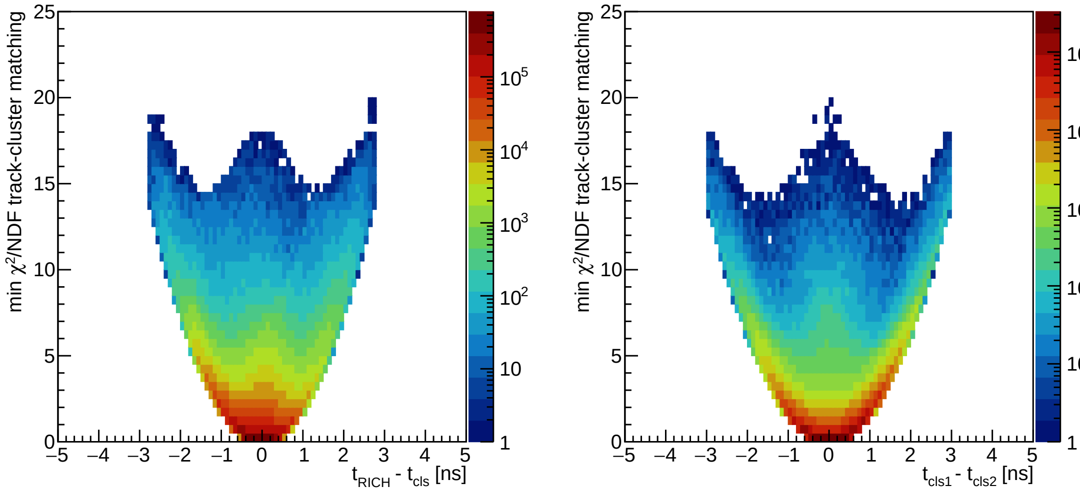


Figure 4.12: The distribution of the minimum of the χ^2 defined in Eq.4.7 is normalized to the number of degrees of freedom (NDF) and reported as a function of ($t_{RICH} - t_{cls}$) and ($t_{cls1} - t_{cls2}$) in the plot on the left and on the right, respectively.

The couple of good clusters in-time with the selected track forming the best triplet with it is chosen considering also the time matching between the two clusters. In particular, the best triplet is the one providing the minimum value of the χ^2 defined by the three time differences $(t_{cls1} - t_{cls2})$ and $(t_{RICH} - t_{cls_i})$ with $i = 1, 2$:

$$\chi^2 = \frac{(t_{RICH} - t_{cls_1})^2 + (t_{RICH} - t_{cls_2})^2}{(0.5\,\mathrm{ns})^2} + \frac{(t_{cls_1} - t_{cls_2})^2}{(0.6\,\mathrm{ns})^2}. \tag{4.7}$$

By choosing the triplet with the minimum χ^2 value, the selection is less sensitive to intensity variations and pile-up effects in the calorimeter, that increase the number of simultaneous clusters in the LKr. The denominator of the first term is a conservative estimate of the resolution of $(t_{RICH} - t_{cls})$, shown in the left hand side of Figure 4.11, taking into account the asymmetry of the distribution and the strong energy dependence at low cluster energies. The denominator of the second term approximates the resolution of $t_{cls1} - t_{cls2}$ (Figure 4.11 right). The distribution of the minimum χ^2, normalized to the number of degrees of freedom (NDF), is reported as a function of $t_{RICH} - t_{cls}$ and $t_{cls1} - t_{cls2}$ in Figure 4.12 left and right, respectively. The event is selected if the condition $\chi^2/NDF < 5$ is satisfied.

4.4 $m_{\gamma\gamma}$ spectrum after the $K\mu\gamma\gamma$ selection

Figure 4.13 reports the distribution of the invariant mass $m_{\gamma\gamma}$ defined in Eq.3.1 for the whole NA62 Run1 dataset, after the selection of the triplet and assuming the two clusters to originate from the reconstructed $K^+ - \mu^+$ charged vertex. A fit to the peak with a double-sided Crystal Ball function [85] points out the presence of the π^0 resonance in the $m_{\gamma\gamma}$ distribution, that validates the goodness of the triplet choice. The regions sensitive to the $K^+ \to \mu^+\nu_\mu X,\ X \to \gamma\gamma$ process ($m_{\gamma\gamma} < 110\,\mathrm{MeV/c^2}$ and $m_{\gamma\gamma} > 160\,\mathrm{MeV/c^2}$) are populated by the continuous distribution surrounding the peak, that is produced from the different backgrounds listed in Sec.3.1. The study of the main backgrounds are reported in the next Chapter, together with the conditions defined to reduce them.

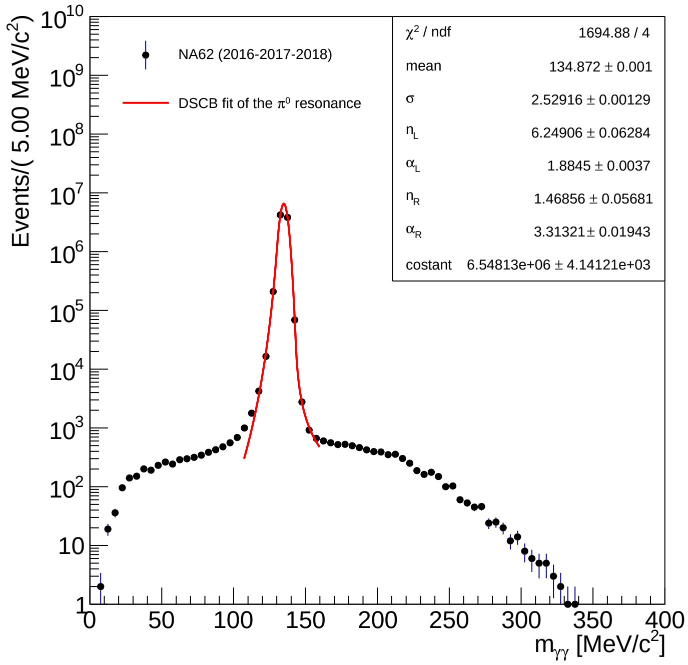


Figure 4.13: Distribution of the reconstructed di-photon invariant mass defined in Eq.3.1 after applying the triplet selection on the NA62 Run1 dataset. A double-sided Crystal Ball fit to data shows that the central peak corresponds to the $\pi^0 \to \gamma\gamma$ decay, validating the triplet selection. The double-sided Crystal Ball function used for the fit is defined as in Eq.6.1 of [85].

Chapter 5

Backgrounds suppression study

In this Chapter, the study of the different sources of background populating the regions sensitive to the signal in the $m_{\gamma\gamma}$ distribution is reported and the strategies defined to reduce them are described.

5.1 Photon conversion rejection

One of the main contributions to the tails of the $m_{\gamma\gamma}$ distribution is due to early conversions of the photons from the $\pi^0 \to \gamma\gamma$ decay in e^+e^- pairs. To study the topology of events with photon conversions that enters the triplet selection, the $K^+ \to \mu^+\nu_\mu\pi^0$ sample is used. Depending on the number of photons and leptons that are absorbed in the LKr and on the z coordinate of the photon conversion, different scenarios can arise when the clusters forming the triplet are chosen, leading to well or badly reconstructed $m_{\gamma\gamma}$ (assuming the muon and kaon track to be well matched).

If both the photons from the $\pi^0 \to \gamma\gamma$ are absorbed in the LKr (Figure 5.1) and their corresponding clusters are selected, the reconstructed $m_{\gamma\gamma}$ is expected to be compatible with the π^0 mass.

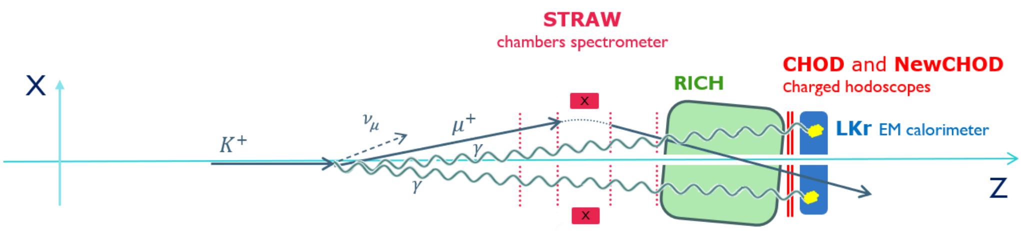
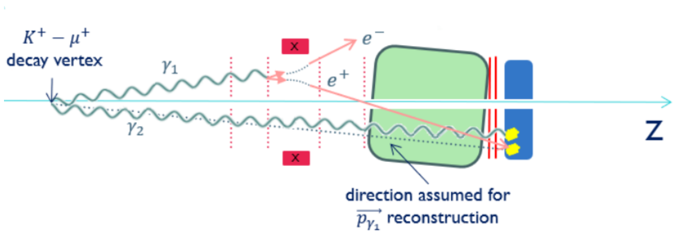
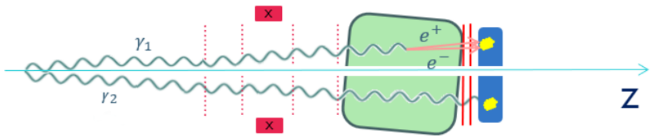


Figure 5.1: In this sketch, the photons from the $\pi^0 \to \gamma\gamma$ decay are absorbed in the LKr and their energies are reconstructed in the clusters chosen for the triplet. The particles generating the selected clusters are thus invisible to the downstream detectors but the LKr (only the relevant downstream detectors are reported in the sketch).



(a)



(b)

Figure 5.2: One of the two photons converts before getting to the LKr. In case (5.2a) the conversion occurs upstream of the STRAW magnet and one of the leptons is deflected out of acceptance. In case (5.2b) the conversion occurs downstream of the STRAW magnet. In both cases some activity other then the muon one is expected in the detectors downstream of the e^+e^- production position, correlated (both in time and position) with one of the selected clusters and in time with the track.

If one out of the two photons converts in an e^+e^- pair before the LKr, different $m_{\gamma\gamma}$ distributions are expected, depending on where the conversion occurs with respect to the dipole magnet (MNP33 magnet) of the STRAW spectrometer:

1. if the conversion occurs in one of the first two STRAW chambers, upstream of the MNP33 magnet, the dipole bends the trajectory of the $e^\pm$ in opposite directions, widening their angle of emission, and one of the two leptons can also be deflected out of the acceptance of the downstream detectors (Figure 5.2a). Assuming that the triplet with the muon track contains one cluster due to only one of the conversion tracks and the other from a not converted photon, the reconstructed $m_{\gamma\gamma}$ can in principle assume any kinematically allowed value, due to both the fraction of energy carried by the second lepton, not included in the selected clusters, and the angle between the clusters, which is different from the angle between the two original photons. The back-extrapolation from the LKr front plane to the reconstructed vertex leads to a wrong direction assumed for the three-momentum associated to the lepton cluster (dotted line in Figure 5.2a) as it does not take into account the momentum kick provided from the magnet to the lepton.

2. if the photon conversion occurs after the MNP33 magnet (Figure 5.2.b), the e^+e^- pair is emitted almost collinear to the photon direction and the energy deposits from the two electrons are expected to be reconstructed as part of the same cluster. As a consequence, $m_{\gamma\gamma} \sim m_{\pi^0}$ if this cluster and the one from the second photon are selected. Nevertheless, if the photon converts in the beam pipe

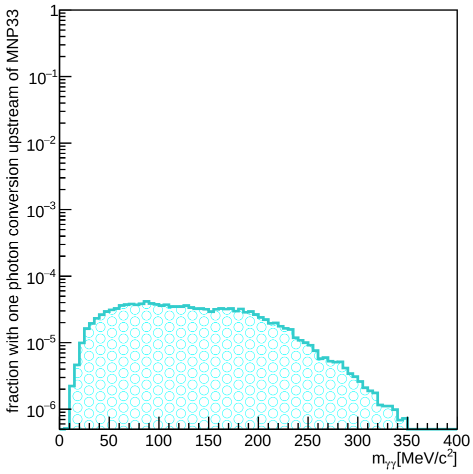


Figure 5.3: Reconstructed $m_{\gamma\gamma}$ distribution for $K^+ \to \mu^+\nu_\mu\pi^0$ simulated events. Left: one of the gamma from the π^0 decay is absorbed in the LKr, the second converts in an e^+e^- pair upstream of the MNP33 magnet, in the first or second STRAW chamber. Right: one of the gamma from the π^0 decay is absorbed in the LKr, the second converts in an e^+e^- pair after the magnet, in the third or fourth STRAW chamber or after. The shoulder below 100 MeV/c^2 is mainly due to photons that convert interacting with the beam pipe, with at least one $e^\pm$ traveling inside it.

crossing the downstream detectors, at least one out of the two conversion track can travel inside the beam pipe and its energy gets lost to the LKr detection.

Differences in the reconstructed $m_{\gamma\gamma}$ between scenario 1 and 2 are shown in Figure 5.3 for $K^+ \to \mu^+\nu_\mu\pi^0$ simulated events. On the left, one of the selected clusters comes from a gamma from the π^0 decay, while the second comes from an $e^\pm$ produced in a photon conversion upstream of the MNP33 magnet. The result is an uniform distribution of $m_{\gamma\gamma}$ with no peak at the π^0 mass. On the right, one of the selected clusters comes from a gamma and the second from an $e^\pm$ produced in a photon conversion downstream of the MNP33 magnet. The result is a clear peak at the π^0 mass value with a shoulder below 100 MeV/c^2, mostly due to photons that convert in the beam pipe with at least one $e^\pm$ traveling inside the beam pipe.

If both the photons from the π^0 convert into e^+e^- pairs before reaching the electromagnetic calorimeter, different combinations of the scenarios 1 and 2 may occur.

When both the selected clusters are due to photons, the muon is the only particle of the final state that produces signals in the downstream detectors preceding the LKr. On the contrary, if one of the photons converts, the charged $e^\pm$ generates in the detectors downstream with respect to the conversion position some extra activity in addition to the muon signals, in time with the muon track and both in-time and geometrically matched with the LKr cluster due to the $e^\pm$. Thus, the presence of extra-in time activity correlated to the cluster in the downstream detectors can be exploited to identify events with photon conversions.

5.1.1 Small angle extra-activity

A large fraction of the conversions happens in the beam pipe crossing the downstream detectors. The energy of at least one of the $e^{\pm}$ produced is usually lost because the particle travels inside the beam pipe. Signals in the IRC and SAC detectors (see Sec.2.5.3) that are in time with the triplet reference time, $t_{ref} = (t_{RICH} + t_{cls1} + t_{cls2})/3$, are used to veto these events.

A tighter LKr inner acceptance condition is also introduced with respect to the standard one (described in Sec.4.3) to reinforce the IRC and SAC veto. In particular, the left hand side of Figure 5.4 shows the coordinates of the end position of photons from $K^+ \to \mu^+\nu_\mu\pi^0$ simulated events that convert in e^+e^- at the RICH entrance. As the events under study are expected to populate especially the tails of the $m_{\gamma\gamma}$ distribution, only events with a reconstructed mass $< 110\,\mathrm{MeV/c^2}$ or $> 160\,\mathrm{MeV/c^2}$ are considered. The right hand side of the same figure shows the position at the LKr front plane of the selected clusters for the same events as in left hand side. Events with a photon conversion in the beam pipe and at least one $e^{\pm}$ crossing the detector active volume are likely to generate clusters close to the beam pipe. The condition $R > 180\,\mathrm{mm}$ instead of $R > 150\,\mathrm{mm}$ is therefore required to reduce these events.

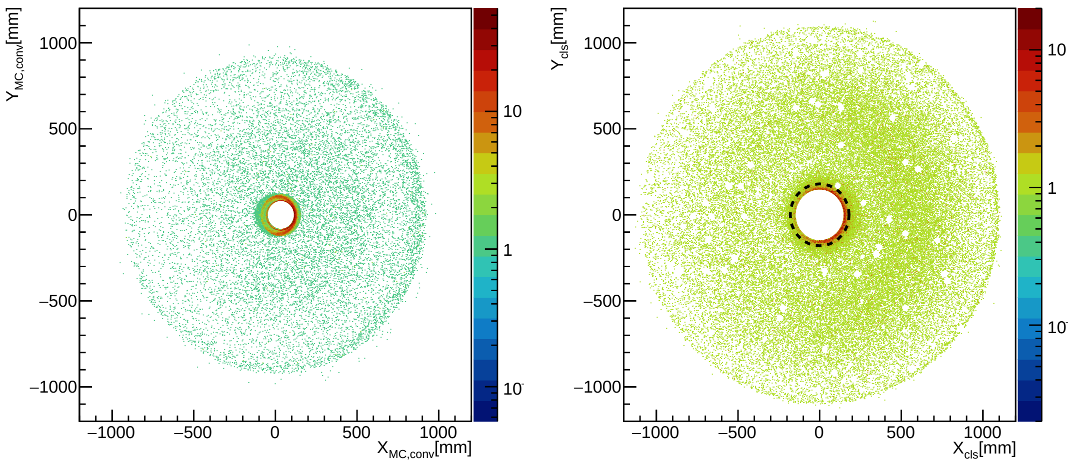


Figure 5.4: Selected $K^+ \to \mu^+\nu_\mu\pi^0$ simulated events with $m_{\gamma\gamma}$ in one of the region sensitive to the $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ signal ($m_{\gamma\gamma} < 110\,\mathrm{MeV/c^2}$ or $m_{\gamma\gamma} > 160\,\mathrm{MeV/c^2}$). Left: coordinates of the end position of one photon that converts in e^+e^- pairs at the RICH entrance. The structures in the central part represent photon conversions due to the interaction of the gamma with the beam pipe in a 0.5 m long z interval. The central hole is asymmetrical with respect to zero in the x direction because of the tilt of the RICH volume in the xz plane. Right: position at the LKr front plane of the selected clusters. The tighter inner LKr acceptance condition ($\sqrt{x_{cls}^2 + y_{cls}^2} > 180\,\mathrm{mm}$) with respect to the standard ($> 150\,\mathrm{mm}$) is represented by the dashed black lines.

5.1.2 CHOD and New-CHOD extra-activity veto

As the hodoscopes are located in front of the LKr, the e^+e^- pair from a photon conversion is expected to generate some activity in the CHOD or New-CHOD detector, independently from the position in z the conversion occurs at. Therefore, a "cluster-charged particle" association can be defined between one of the selected clusters and CHOD or New-CHOD signals, pointing to the possibility that the cluster is due to the energy deposit of a charged particle instead of a photon. In particular, for a selected cluster, a "cluster-charged particle" association exists if there is a "cluster-CHOD association" or a "cluster-New-CHOD association".

"Cluster - New-CHOD" association. Let a New-CHOD candidate be a reconstructed signal from a New-CHOD tile. There is activity in the New-CHOD associated to a LKr cluster if at least one New-CHOD candidate exists:

- in time with the track in a time window 10 ns wide: $|t_{\text{extra New-CHOD cand}} - t_{RICH}| \leq 5\,\text{ns}$. The distribution of this time difference is reported on the left hand side of Figure 5.5. The New-CHOD candidate associated to the track is excluded a priori.
- geometrically matched with the cluster. The geometrical matching is defined as: $|x_{cls} - x_{\text{extra New-CHOD cand}}| < 200\,\text{mm}$ and $|y_{cls} - y_{\text{extra New-CHOD cand}}| < 100\,\text{mm}$, where the expected cluster position at the New-CHOD plane is assumed equal to the measured cluster position at the LKr front plane. This approximation is justified by the small distance along the z direction between the LKr and New-CHOD detectors, $\mathcal{O}(1\,\text{m})$, with respect to the $\mathcal{O}(10^2\,\text{m})$ distance between the LKr and the K^+ vertex position, that should be used to back-extrapolate from the LKr the cluster position at the New-CHOD plane. Figure 5.5 right shows the matching variables in the two dimensional space. Their distribution strongly depends on the geometry of the New-CHOD tiles (see Sec.2.5.2).

New-CHOD signals satisfying these two conditions are in time with the cluster in $\pm 5\,\text{ns}$.

"Cluster - CHOD" association. The association between the cluster and some activity in the CHOD detector must take into account that the measured times of CHOD hits need to be corrected for the time of propagation of the scintillation light through a slab and that the corrections strongly depend on the expected position of the hit in the CHOD plane. For this reason the association is defined in three steps.

First, after excluding the hits associated to the selected muon track, extra CHOD hits in time with the track are selected with $|t_{CHOD\,extra\,hit} - 7\,\text{ns} - t_{RICH}| \leq 7\,\text{ns}$, where $(t_{CHOD\,extra\,hit} - 7\,\text{ns})$ roughly takes into account the maximum time of propagation of the scintillation light through a slab. The corresponding time difference distribution is shown in Figure 5.6 on the left.

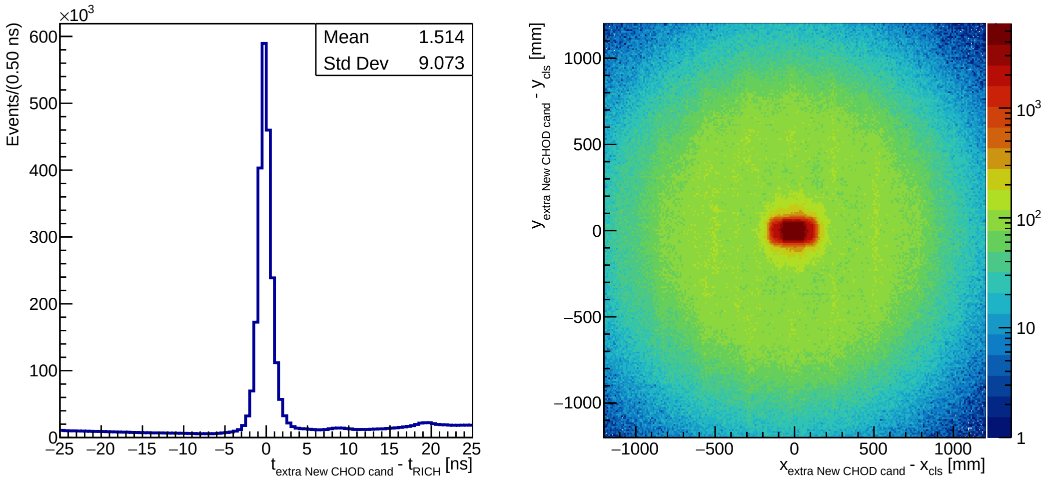


Figure 5.5: Subsample of the NA62 Run1 data. Left: distribution of the difference between the time of extra New-CHOD candidates not associated to the track and the track reference time (t_{RICH}). The condition $|t_{\text{New-CHOD cand}} - t_{RICH}| \leq 5\,\text{ns}$ defines extra in-time New-CHOD candidates. Right: distribution in the (x, y) plane of the difference between the position of extra in-time New-CHOD candidates with respect to the expected cluster position at the New-CHOD plane.

Subsequently, among the selected hits, pairs of simultaneous signals in vertical and horizontal slabs in the same quadrant are looked for. The right hand side of Figure 5.6 reports the distribution in the CHOD (x, y) plane of the coordinates of the obtained CHOD pair (or cross) compared to the expected coordinate of the LKr cluster in the same plane. As for the definition of "cluster - New-CHOD" association, the back-extrapolated position of the cluster at the CHOD plane is approximated with the measured position of the cluster at the LKr front plane. The presence of the central cross is due to the fact that, without taking into account more precise timing conditions, for a given hit in a vertical slab due to the event of interest, the paired hit can be in each of all the 16 horizontal slabs of the same quadrant (see Sec.2.5.2). The time of the CHOD cross is defined as: $t_{CHOD\,cross} \equiv (t_{CHOD\,hit,\,\text{V}} + t_{CHOD\,hit,\,\text{H}})/2$, where $t_{CHOD\,hit,\,\text{V}}$ and $t_{CHOD\,hit,\,\text{H}}$ are the measured hit times with time corrections applied. Time corrections are based on the expected cluster position at the CHOD plane.

Finally, the CHOD extra in-time activity is associated to the selected cluster if:

- $|t_{CHOD\,cross} - t_{cls}| < 5\,\text{ns}$
- $D(CHOD\,cross - cls) < 100\,\text{mm}$, with $D(CHOD\,cross - cls)$ distance in the CHOD (x, y) plane between the CHOD cross and the expected position of the cluster.

The distribution of $D(CHOD\,cross - cls)$ as a function of $|t_{CHOD\,cross} - t_{cls}|$ is reported in Figure 5.7.

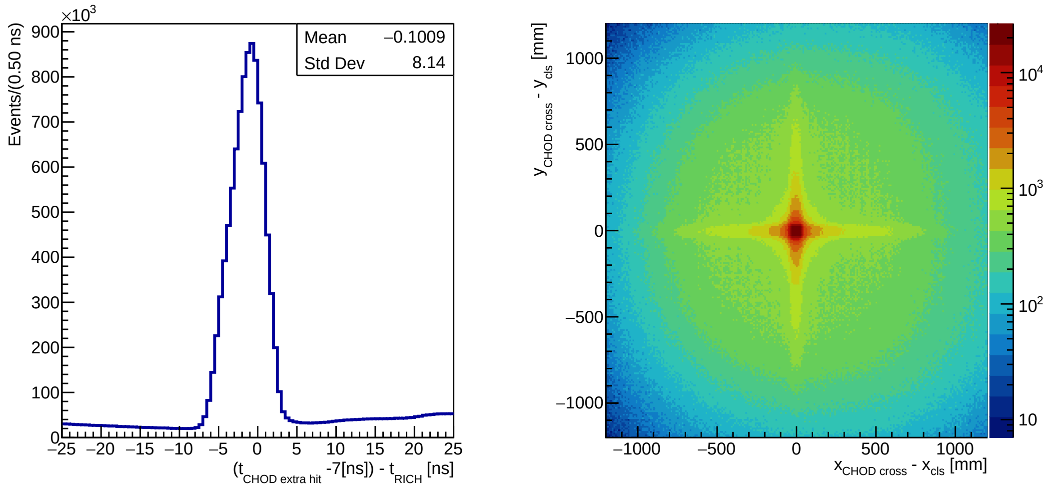


Figure 5.6: Subsample of the NA62 Run1 data. Left: distribution of the difference between the measured time of CHOD hits not associated to the track and the track reference time (t_{RICH}). The time of the hits is corrected for a rough estimate of the propagation time of the scintillation light through a CHOD slab, as explained in the text. The condition $|t_{\text{CHOD extra hit}} - 7\,\text{ns} - t_{RICH}| \leq 7\,\text{ns}$ defines extra in-time CHOD hits. Right: distribution in the CHOD (x, y) plane of the difference between the position of the pairs (or crosses) of horizontal and vertical extra in-time CHOD hits belonging to the same quadrant with respect to the expected cluster position at the CHOD plane.

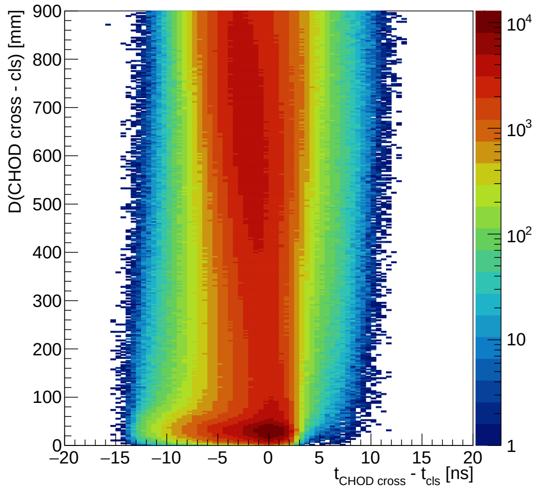


Figure 5.7: Distribution of the distance between the selected CHOD cross and the LKr cluster at the CHOD plane as a function of their time difference for the same subsample of the NA62 Run1 data as in Figure 5.6. The times of the CHOD cross hits are corrected assuming the cluster position as expected position for the cross.

Once the CHOD and NewCHOD associations are defined, three classes of events can be identified according to the "cluster - charged particle" association:

Events with 0 "cluster-charged particle" associations: none of the two selected clusters has a CHOD or New CHOD associated signal. These are the events of interest as both the clusters are likely to be associated to photons.

Events with 1 "cluster-charged particle" association: one of the two selected clusters is associated to signals in at least one of the hodoscopes and it is likely to be the cluster due to one or both the tracks from a photon conversion. At this level, it is not possible to distinguish between these two possibilities, corresponding to the scenarios 1 and 2 described in Sec.5.1, respectively.

Events with 2 "cluster-charged particle" associations: both the selected clusters have an association to signals in at least one of the hodoscopes.

Figure 5.8 shows the performance of the "cluster-charged particle" association algorithm as measured from the sample of simulated $K^+ \to \mu^+\nu_\mu\pi^0$ decays. For this study, events are selected, by using the MC truth, with both the photons from the π^0 decay absorbed in the calorimeter or with one or both of them converted in e^+e^- pair before it, in a z range starting from the first chamber of the STRAW spectrometer. The matrix on the left shows the efficiencies and the purity for events with conversions up to the New CHOD: events with both photons absorbed in the LKr are identified with an efficiency $> 99\%$ and the purity of the sample with 0 "cluster-charged particle" associations is $> 99\%$ as well. If also the conversions in the New-CHOD (CHOD) planes, are included, the purity of the sample is $> 97\%$ (0.84%) due to the inefficiencies of the detectors to the production of the charged particles in their own volumes. The matrix on the right shows the performance of the algorithm including the whole range from the STRAW 1 to the LKr. The decrease of the sample purity is due to events with conversions that occur after the CHOD, that cannot be detected by the hodoscopes by definition.

5.1.3 STRAW extra-activity

To enforce the photon conversions veto when photons convert in one of the STRAW chambers, the presence of extra-activity in the STRAW in addition to the track one is considered. In particular, the standard NA62 algorithm for the reconstruction of segments in the spectrometer is used. This tool can identify events with segments, STRAW hits compatible with partial (not reconstructed) tracks originating from a user-provided vertex, both in the chambers upstream and in those downstream of the MNP33 magnet. Here the reconstructed $K^+ - \mu^+$ vertex is provided, together with the selected muon track time, t_{RICH}, as reference time for the matching. From the $K^+ \to \mu^+\nu_\mu\pi^0$ simulation, in the sample selected with 0 "cluster-charged particle" associations, events without photon conversions in one of the STRAW chambers are identified with an efficiency $> 97\%$; the residual contribution to the sample from conversions in one of the chambers is $< 0.1\%$.

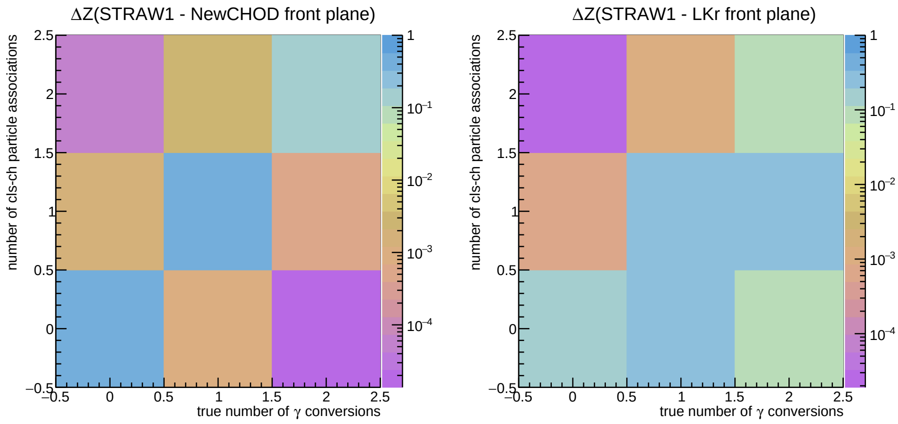


Figure 5.8: Performance of the "cluster-charged particle" association algorithm measured by $K^+ \to \mu^+\nu_\mu\pi^0$ simulated decays. Left: events are selected with both the photons absorbed in the LKr, one absorbed in the LKr and the other converted in e^+e^- pair in the z range from STRAW1 up to the New-CHOD front plane, or both the photons converted in the range specified. The true number of converted photon is reported in the horizontal axis. The vertical axis corresponds to the output of the algorithm for each event. Each bin corresponds to a fraction of events with respect to the total. Events with the $\gamma\gamma$ pair absorbed in the LKr, with one or both converted photons are identified with an efficiency of $\mathcal{O}(99.3\%)$, $\mathcal{O}(98.9\%)$ and $\mathcal{O}(99.7\%)$, respectively. The sample with null "cluster-charged particle" associations has a purity $> 99\%$. Right: same as on the left but including events with photon conversions up to the LKr front plane. The matrix represents how efficiencies and purity change due to the inefficiencies of CHOD and New-CHOD to conversions occurring in their own volume (not included on the left plot) and to conversions after the hodoscopes.

Figure 5.9 reports the $m_{\gamma\gamma}$ distribution of the NA62 data before (blue dots) and after applying each of the conditions defined to reduce the background due to photon conversions.

5.2 Kinematic constraints for background suppression in signal regions

The four-momentum conservation is exploited to reduce backgrounds coming from SM K^+ decay modes in the di-photon invariant mass regions sensitive to the $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ signal defined in Eq.3.2 and 3.3:

$$R1 : 0 < m_{\gamma\gamma} < 110\,\mathrm{MeV/c^2}$$

$$R2 : 160 < m_{\gamma\gamma} < 380\,\mathrm{MeV/c^2}$$

As the majority of the backgrounds coming from K^+ decays are common between the normalization channel ($K^+ \to \mu^+\nu_\mu\pi^0$ decay mode) and the signal process, the kinematic conditions defined for the signal selection (or more stringent ones) are also required in the normalization selection. In the following, each section specifies the kinematic conditions introduced for the reduction of a specific background.

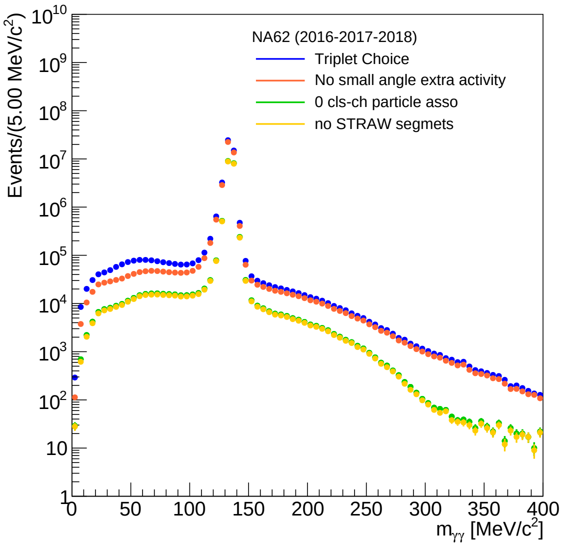


Figure 5.9: Di-photon invariant mass distribution of NA62 Run1 data after the triplet selection described in Chapter 4 and after each of the conditions against the photon conversions described in Sec. 5.1.

5.2.1 $K^+ \to \mu^+\nu_\mu$ suppression

The $K^+ \to \mu^+\nu_\mu$ ($K_{\mu 2}$) decay has an experimental signature very different from the signal and normalization ones. However, this K^+ decay mode enters both the selections due to accidental photons in time with the muon and, because of its branching ratio ($\sim 63\%$), constitutes an important source of background. To reduce it, the squared missing mass:

$$m^2_{miss}(K-\mu) = (P_K - P_\mu)^2 \tag{5.1}$$

is defined, with P_K and P_μ the kaon four-momentum and the downstream track four-momentum in the muon mass hypothesis, respectively. The left hand side of Figure 5.10 shows the data $m^2_{miss}(K-\mu)$ distribution as a function of $m_{\gamma\gamma}$ after the triplet selection and the conditions to reject photon conversions. The $m^2_{miss}(K-\mu)$ peak at zero mainly comes from the missing energy and momentum of the neutrino from the $K^+ \to \mu^+\nu_\mu$ decay, and corresponds to $m_{\gamma\gamma}$ values spread all over the range of interest $(0 < m_{\gamma\gamma} < (m_K - m_\mu))$, with a higher density at lower masses. Low $m_{\gamma\gamma}$ are expected, for example, for triplets formed by a cluster coming from a radiative photon of the $K^+ \to \mu^+\nu_\mu$ decay in time coincidence with an accidental photon. Figure 5.10 right reports the distribution of $m^2_{miss}(K-\mu)$ for the $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ MC sample as a function of the true mass of X, after the triplet selection and the conditions to reject photon conversions.

The neutrino peak in the data distribution can be completely rejected by the requirement $m^2_{miss}(K-\mu) > 0.015\,\mathrm{GeV}^2/\mathrm{c}^4$, without significant losses of signal.

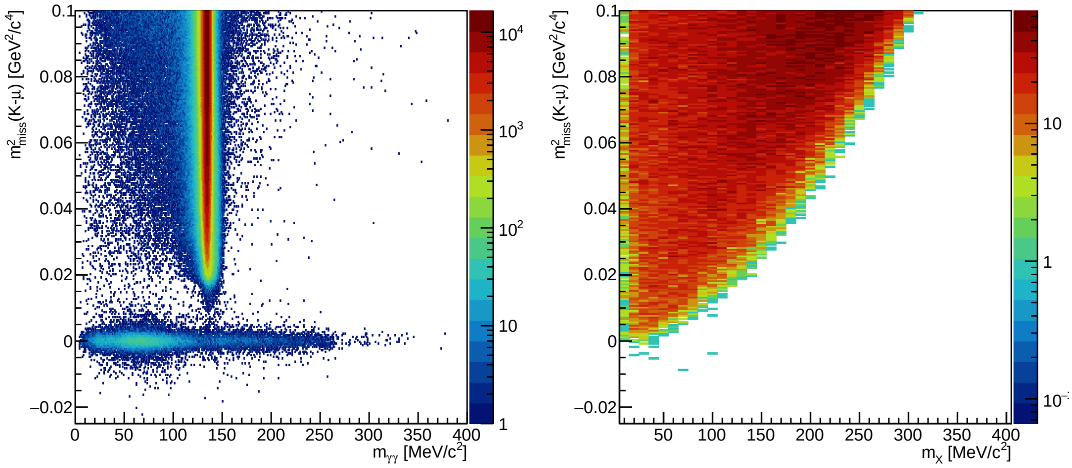


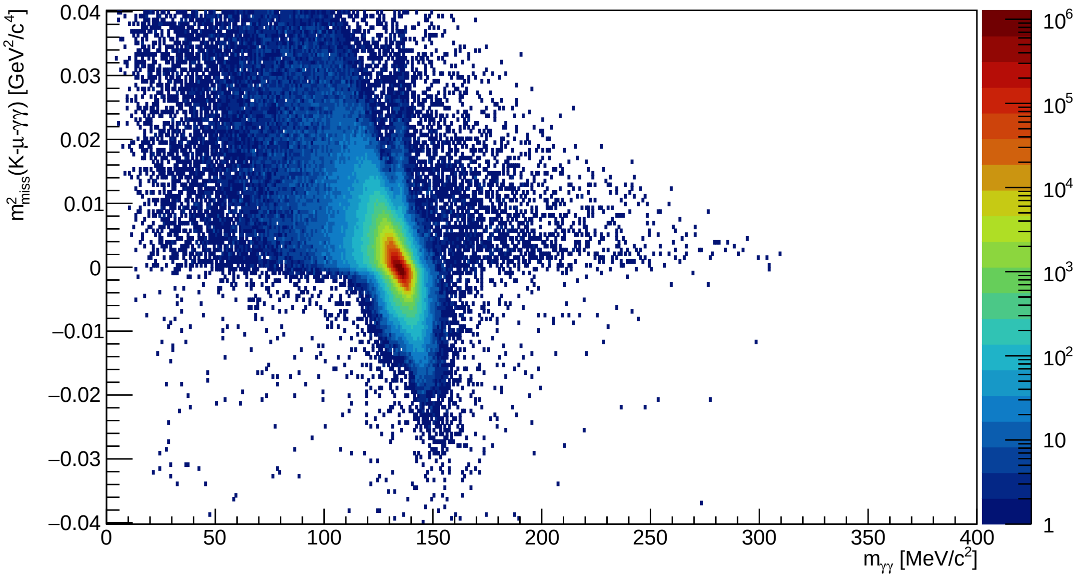
Figure 5.10: Distribution of the squared missing mass $m^2_{miss}(K-\mu)$ as a function of $m_{\gamma\gamma}$ for the NA62 Run1 data (left) and as a function of the true X mass for the $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ signal samples (right), after the triplet selection and the conditions to reject photon conversions. $m^2_{miss}(K-\mu)$ is defined according to Eq.5.1.

5.2.2 $K^+ \to \mu^+\nu_\mu\pi^0$ tails suppression

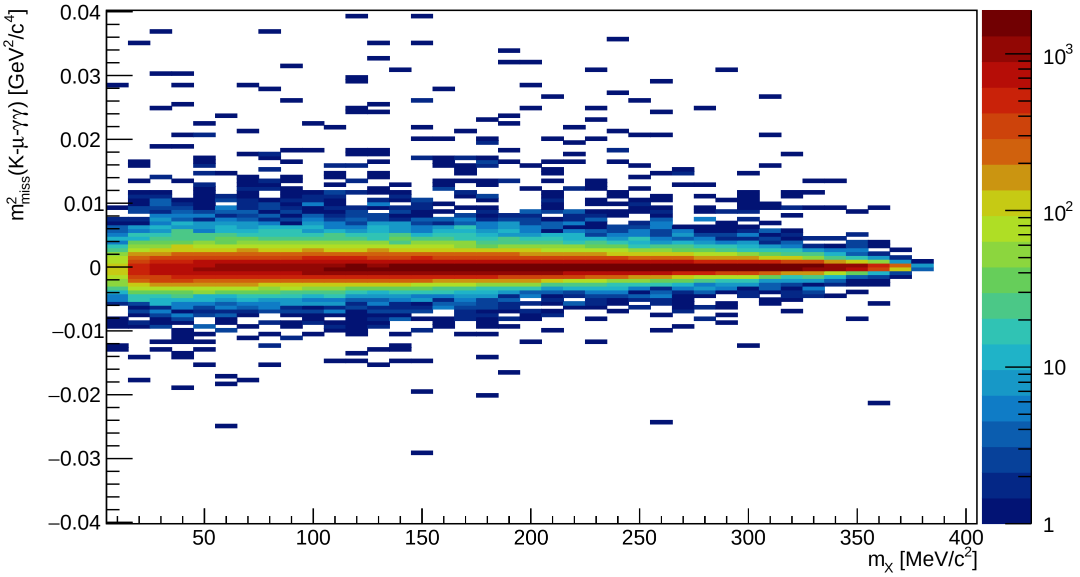
$K^+ \to \mu^+\nu_\mu\pi^0$ ($K_{\mu 3}$) events with badly reconstructed π^0, mainly due to residual photon conversion contributions and $K^+ - \mu^+$ mismatch, populate the tails of the $m_{\gamma\gamma}$ distribution. These events enter as background in the regions sensitive to the $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ decay in the signal selection and degrade the quality of the events in the normalization sample. To reduce them, the squared missing mass:

$$m^2_{miss}(K-\mu-\gamma\gamma) \equiv (P_K - P_\mu - P_{\gamma 1} - P_{\gamma 2})^2, \tag{5.2}$$

is introduced, where P_K and P_μ are defined as in Eq.5.1, $P_{\gamma 1,2}$ are the four-momenta of the photons computed as explained in Sec.3.1. For processes like $K_{\mu 3}$ and $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$, $m^2_{miss}(K-\mu-\gamma\gamma)$ represents the squared neutrino mass and its distribution peaks at zero for well reconstructed events. Figure 5.11 reports the distribution of $m^2_{miss}(K-\mu-\gamma\gamma)$ as a function of $m_{\gamma\gamma}$ for simulated $K_{\mu 3}$ events (top) and as a function of m_X for simulated signal events (bottom), after the triplet selection, the conditions to reject photon conversions and the cut defined to suppress the $K^+ \to \mu^+\nu_\mu$ background are applied. These distributions clearly show that $K_{\mu 3}$ events in which the π^0 is badly reconstructed can be strongly reduced by selecting the neutrino mass peak in the squared missing mass, while preserving signal events for every tested X mass. Therefore, $0.004 < m^2_{miss}(K-\mu-\gamma\gamma) < 0.006\,\mathrm{GeV^2/c^4}$ is required.



(a) $K^+ \to \mu^+\nu_\mu\pi^0$ MC sample. $m^2_{miss}(K-\mu-\gamma\gamma)$ vs $m_{\gamma\gamma}$ after the triplet selection and the conditions to reject photon conversions and the $K^+ \to \mu^+\nu_\mu$ background. The squared missing mass is defined as in Eq.5.2.



(b) $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ MC sample. $m^2_{miss}(K-\mu-\gamma\gamma)$ vs m_X after the triplet selection, the conditions to reject photon conversions and the $K^+ \to \mu^+\nu_\mu$ background. The squared missing mass is defined as in Eq.5.2.

Figure 5.11: Top: correlation between $m^2_{miss}(K-\mu-\gamma\gamma)$ and $m_{\gamma\gamma}$ for selected $K^+ \to \mu^+\nu_\mu\pi^0$ simulated events. The clear structure at positive missing mass values and $m_{\gamma\gamma} \sim m_{\pi^0}$ is due to the emission of radiative photons. Other events with positive missing mass are due to residual photon conversions. Bottom: $m^2_{miss}(K-\mu-\gamma\gamma)$ for selected MC $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ events as a function of the simulated X mass hypothesis (m_X).

5.2.3 $K^+ \to \pi^+\pi^0$ suppression

The $K^+ \to \pi^+\pi^0$ decay mode ($K_{2\pi}$) enters the signal and the normalization selections due to a π^+ wrongly identified as a μ^+ (mis-identification contribution) or because of the π^+ decay-in-flight to the $\mu^+\nu_\mu$ final state. The mis-identification contribution becomes negligible thanks to the particle identification conditions (see Sec.4.1), while events with the $\pi^+ \to \mu^+\nu_\mu$ decay constitute one of the dominant backgrounds. In particular, when the charged pion decays while travelling along the volume of the NA62 detector, two different categories of events enter the triplet selection, depending on the position where the $\pi^+ \to \mu^+\nu_\mu$ decay occurs. Assuming that the clusters considered in the triplet are produced by the photons from the decay of the neutral pion in the $K_{2\pi}$ event, two possibilities arise:

1. if the $\pi^+ \to \mu^+\nu_\mu$ decay occurs downstream of the STRAW spectrometer, the spectrometer measures the pion track, while the detectors used for PID detect the muon track. As a consequence, the pion track is matched with the upstream kaon track, ideally providing the right K^+ decay vertex. The corresponding reconstructed di-photon invariant mass is expected close to the π^0 mass.

2. if the $\pi^+ \to \mu^+\nu_\mu$ decay occurs upstream of the STRAW spectrometer, the spectrometer measures the muon track instead of the pion one. As the muon track is matched with the track of the mother kaon, the reconstructed $K^+ - \mu^+$ vertex is shifted forward with respect to the real position of the K^+ vertex. Therefore, the directions of the photons three-momenta are always mis-reconstructed. However, if the π^+ vertex is sufficiently close to the K^+ one, the reconstructed photon three-momenta can also be sufficiently close to the true ones, causing a widening of π^0 mass peak observed for case 1.

As the assumed K^+ vertex position determines the reconstructed direction of flight of photons, the presence of the final state π^0 can be exploited to define a decay vertex determination independent with respect to the downstream track, that allows to reduce the number of events coming from the $K^+ \to \pi^+\pi^0\,(\pi^+ \to \mu^+\nu_\mu)$ decay chain. The neutral $K^+ - \pi^0$ decay vertex is therefore used to determine the alternative photons three-momenta $\overrightarrow{p}_{\gamma 1,2\,neut}$. The calculation of the $K^+ - \pi^0$ vertex and $\overrightarrow{p}_{\gamma 1,2\,neut}$ is explained in Appendix C. The squared missing mass $m^2_{miss}(K - \pi^0)$ computed by using the alternative photon three-momenta is used for background suppression. To enhance the rejection of $K_{2\pi}$ events, the squared missing mass $m^2_{miss}(K - \pi^+)$ and the total reconstructed downstream momentum are also considered, both built starting from $\overrightarrow{p}_{\gamma 1,2}$ as defined in Sec.3.1.

Total downstream momentum condition. The total downstream momentum is defined as $p_{\mu\gamma\gamma} = |\overrightarrow{p}_{track} + \overrightarrow{p}_{\gamma 1} + \overrightarrow{p}_{\gamma 2}|$. For well reconstructed signal and normalization decays it must be consistent with the $\sim 75\,\text{GeV/c}$ momentum of the NA62 K^+ beam and, at the same time, with the missing momentum carried by the final state neutrino.

On the other hand, for $K_{2\pi}$ decays in which the pion momentum and the clusters from the π^0 are measured, $p_{\mu\gamma\gamma}$ peaks at the beam momentum.

The condition $p_{\mu\gamma\gamma} < 72\,\mathrm{GeV/c}$ is therefore introduced. This requirement allows missing momenta greater than $\mathcal{O}(3\,\mathrm{GeV/c})$ and ensures an overall reduction of $K_{2\pi}$ events of $\mathcal{O}(60\%)$. For $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ events, as the available phase space is smaller when the X mass hypothesis is closer to the kinematic bound $m_K - m_\mu$, the $p_{\mu\gamma\gamma}$ condition reduces the acceptance at higher X masses. The efficiency of the $p_{\mu\gamma\gamma}$ condition is reported as a function of m_X in Figure 5.12.

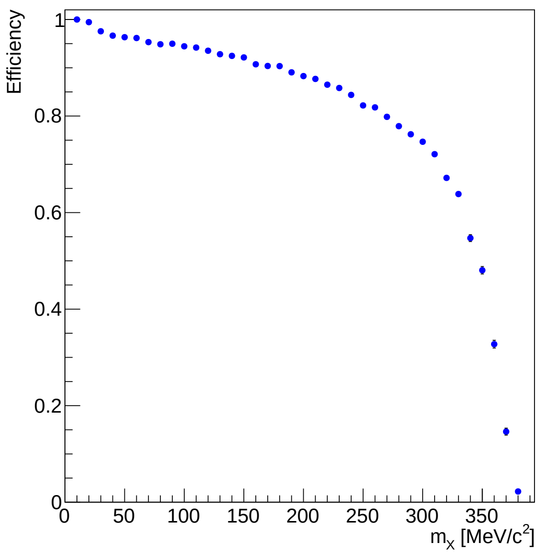


Figure 5.12: Efficiency of the $p_{\mu\gamma\gamma} < 72\,\mathrm{GeV/c}$ cut as a function of the simulated X mass hypothesis for MC $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ events.

Squared missing masses $m^2_{miss}(K-\pi^0)$ and $m^2_{miss}(K-\pi^+)$. The squared missing mass:

$$m^2_{miss}(K-\pi^0) \equiv (P_K - P_{\pi^0})^2 \tag{5.3}$$

is introduced, where P_K is the reconstructed kaon four-momentum previously defined and $P_{\pi^0} \equiv (E_{neut}, \overrightarrow{p}_{\gamma 1,neut} + \overrightarrow{p}_{\gamma 2\,neut})$, with $E_{neut} = \sqrt{m^2_{\pi^0} + |\overrightarrow{p}_{\gamma 1,neut} + \overrightarrow{p}_{\gamma 2\,neut}|^2}$ and $\overrightarrow{p}_{\gamma_i\,neut}$ photon three-momenta evaluated assuming the $K^+ - \pi^0$ neutral decay vertex, as in Eq.C.4 in Appendix C. To be noted that the squared missing mass in Eq.5.3 is the only quantity in the analysis evaluated by using the neutral decay vertex to compute the photon direction of flight. It strongly depends on the π^0 mass assumption for the neutral particle generating the photons, thus is not suitable for the $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ search, and has a worst resolution with respect to the $K^+ - \mu^+$ reconstructed vertex. Every kinematic variable but $m^2_{miss}(K-\pi^0)$ in this analysis is computed by using the $K^+ - \mu^+$ charged vertex, which depends only on the reconstructed direction of the selected upstream and downstream tracks. In particular,

$$m^2_{miss}(K - \pi^+) \equiv (P_K - P_{\pi^+})^2 \tag{5.4}$$

is defined by the difference between the kaon four-momentum P_K defined above and the four-momentum of the selected downstream track in the hypothesis of pion mass. Figure 5.13a reports the correlation between the two squared missing masses defined by Eq.5.3 and Eq.5.4 for $K^+ \to \pi^+\pi^0$ events with $\pi^+ \to \mu^+\nu_\mu$ decay. The squared mass $m^2_{miss}(K - \pi^0)$ is, by construction, independent by the charged decay vertex and it equals the π^+ squared mass if the selected K^+ track and the two LKr clusters come from a $K_{2\pi}$ event, whether the π^+ decays or not. This reflects into a uniform distribution of this variable as a function of $m^2_{miss}(K - \pi^+)$, centered at the π^+ squared mass in the $m^2_{miss}(K - \pi^0)$ axis. The $m^2_{miss}(K - \pi^+)$ distribution is expected to peak at the π^0 squared mass for $K_{2\pi}$ events if the measured track is the π^+ one, and to be a continuous distribution if the measured track is the one of the μ^+ from the π^+ decay, as shown in the plot. The π^0 and π^+ squared mass peaks have different resolutions due to a wider resolution of the z coordinate of the neutral $K^+ - \pi^0$ vertex with respect to the charged vertex, as shown in Figure C.2 in Appendix C.

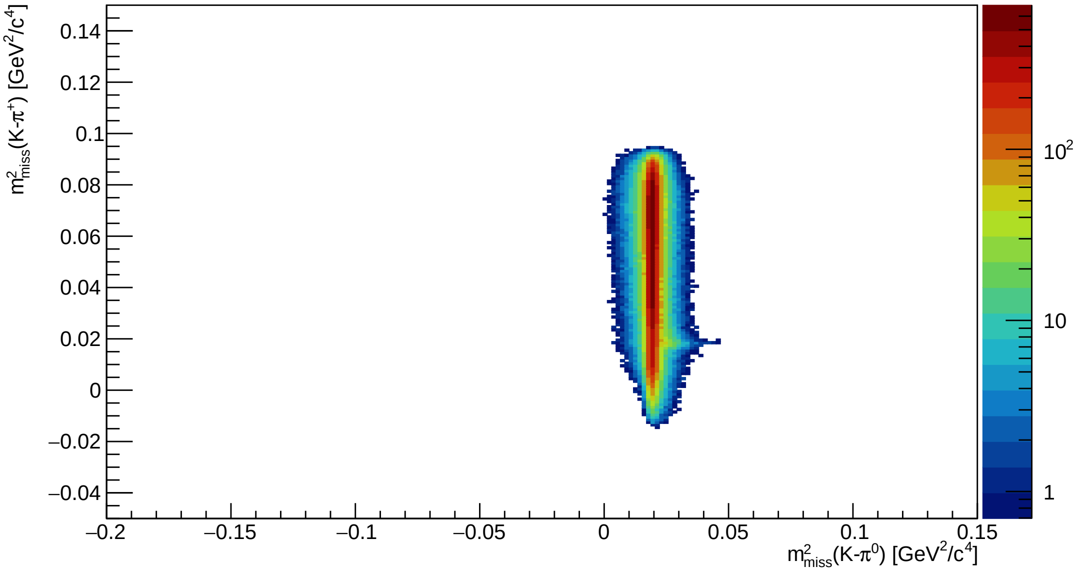
The condition $m^2_{miss}(K - \pi^0) \geq 0.06\,\mathrm{GeV}^2/\mathrm{c}^4$ would reject $K^+ \to \pi^+\pi^0$ events with $\pi^+ \to \mu^+\nu_\mu$ decay-in-flight almost completely. However, even if this condition is suitable for the normalization selection, it would strongly reduce the signal acceptance for every X mass hypothesis below $100\,\mathrm{MeV/c}^2$, as clearly shown by the distribution of $m^2_{miss}(K - \pi^0)$ as a function of the simulated m_X in Figure 5.13b. The following conditions are therefore introduced:

$m^2_{miss}(K - \pi^0) < 0.01\,\mathrm{GeV}^2/\mathrm{c}^4$ or $> 0.03\,\mathrm{GeV}^2/\mathrm{c}^4$ in the signal selection (loose cut)

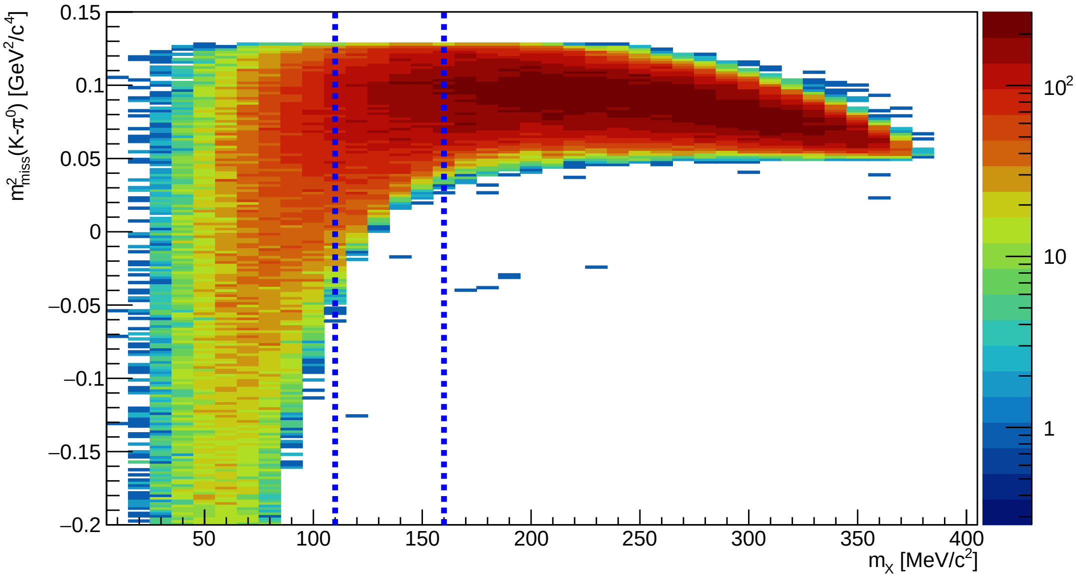
$m^2_{miss}(K - \pi^0) > 0.06\,\mathrm{GeV}^2/\mathrm{c}^4$ in the normalization selection (tight cut)

$m^2_{miss}(K - \pi^+) > 0.04\,\mathrm{GeV}^2/\mathrm{c}^4$ in both signal and normalization selection.

The minimum selected $m^2_{miss}(K - \pi^+)$ value has been chosen to exclude background events populating R1, as shown in Figure 5.14a. Those events are $K^+ \to \pi^+\pi^0$ decays in which at least one of the $\pi^0 \to \gamma\gamma$ photons undergoes a conversion into the e^+e^- pair but enters the 0 "cluster-charged particle" associations sample (see Sec.5.1). For the very specific case of the two body $K_{2\pi}$ decay with well or almost well reconstructed charged vertex, residual photon conversions can be excluded by ruling out the π^0 peak in the $m^2_{miss}(K - \pi^+)$ distribution. Figure 5.14b shows the $m^2_{miss}(K - \pi^+)$ distribution for selected MC signal events as a function of the m_X.

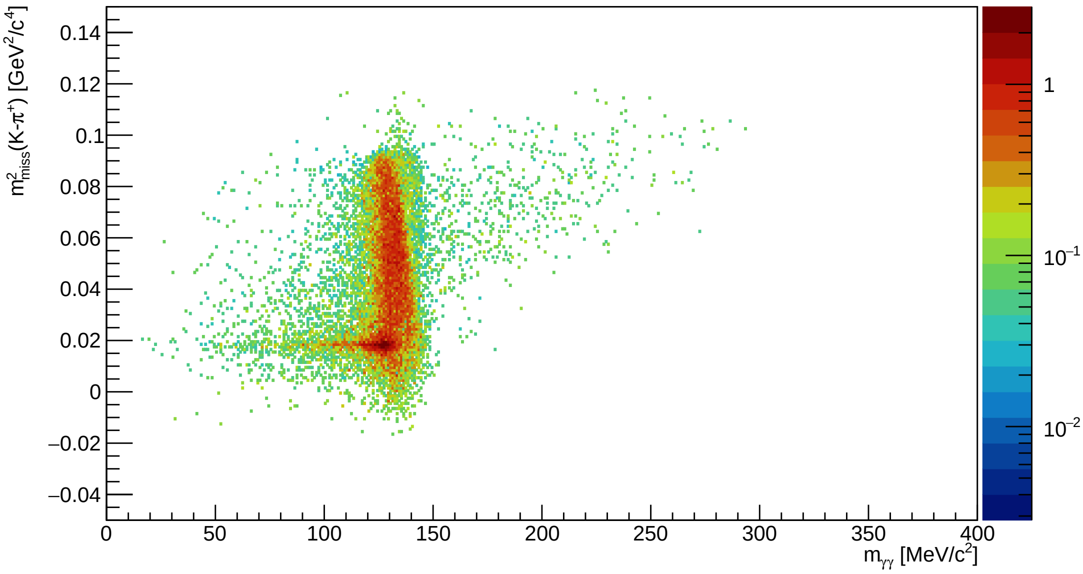


(a) $K^+ \to \pi^+\pi^0$, $\pi^+ \to \mu^+\nu_\mu$ MC sample. $m^2_{miss}(K-\pi^+)$ vs $m^2_{miss}(K-\pi^0)$.

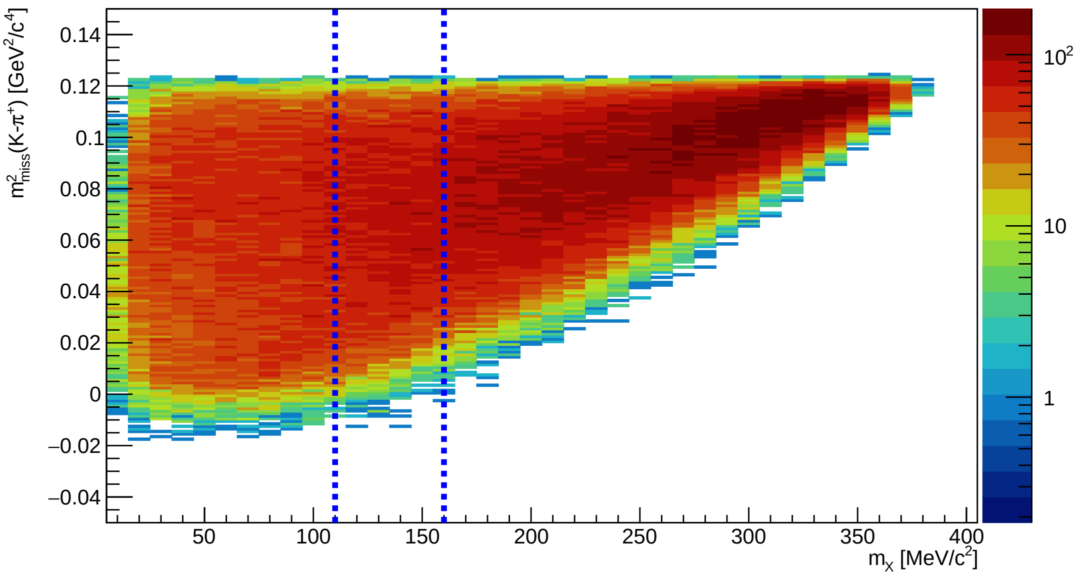


(b) $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ MC sample. $m^2_{miss}(K-\pi^0)$ vs m_X. Blue lines indicate the edges of the X mass intervals included in the regions sensitive to the signal in the $m_{\gamma\gamma}$ distribution ($m_X < 110\,\mathrm{MeV/c^2}$ or $m_X > 160\,\mathrm{MeV/c^2}$).

Figure 5.13: Top: correlation between $m^2_{miss}(K-\pi^+)$ and $m^2_{miss}(K-\pi^0)$ for MC $K^+ \to \pi^+\pi^0$ events with $\pi^+ \to \mu^+\nu_\mu$ decay entering the $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ selection. The squared missing masses are defined as in Eq.5.4 and Eq.5.3, respectively. Bottom: $m^2_{miss}(K-\pi^0)$ for selected MC $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ events as a function of the simulated X mass hypothesis (m_X).



(a) $K^+ \to \pi^+\pi^0$, $\pi^+ \to \mu^+\nu_\mu$ MC sample. $m^2_{miss}(K-\pi^+)$ vs $m_{\gamma\gamma}$. The squared missing mass is defined as in Eq.5.4.



(b) $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ MC sample. $m^2_{miss}(K-\pi^+)$ vs m_X. Blue lines indicate the edges of the X mass intervals included in the regions sensitive to the signal in the $m_{\gamma\gamma}$ distribution ($m_X < 110\,\mathrm{MeV/c^2}$ or $m_X > 160\,\mathrm{MeV/c^2}$). The squared missing mass is defined as in Eq.5.4.

Figure 5.14: Top: correlation between $m^2_{miss}(K-\pi^+)$ and $m_{\gamma\gamma}$ for MC $K^+ \to \pi^+\pi^0$ events with $\pi^+ \to \mu^+\nu_\mu$ decay passing the $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ selection after the loose $m^2_{miss}(K-\pi^0)$ condition. Bottom: $m^2_{miss}(K-\pi^+)$ for selected MC $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ events as a function of the simulated X mass hypothesis (m_X) after the loose $m^2_{miss}(K-\pi^0)$ condition.

5.3 Photon vetoes for $K^+ \to \pi^+\pi^0\pi^0$ rejection

In the same way as for the $K^+ \to \pi^+\pi^0$ decay, $K^+ \to \pi^+\pi^0\pi^0$ events enter both the signal and the normalization selection due to the $\pi^+ \to \mu^+\nu_\mu$ decay-in-flight and to the mis-identification of the pion as a muon. The presence of two neutral pions in the final state, however, makes the phenomenology of the mis-reconstructed $K^+ \to \pi^+\pi^0\pi^0$ decays wider than the $K^+ \to \pi^+\pi^0$ one. In fact, for this decay mode, an almost uniform $m_{\gamma\gamma}$ distribution is observed because of the combinatorial background of selected photons coming from different π^0's, dominating over a small component of well reconstructed π^0's. The uniform component is affected by vertex mis-reconstruction and photon conversions as well.

To reduce the background from the $K^+ \to \pi^+\pi^0\pi^0$ decay mode, events with extra in-time activity in the LKr or with activity in the LAV system in time with the selected track-clusters triplet are rejected.

5.3.1 LKr veto

Due to the presence of two neutral pions in the final state, for $K^+ \to \pi^+\pi^0\pi^0$ decays a higher multiplicity of in-time electromagnetic clusters in the LKr is expected compared, for example, to the K^+ decays with only one π^0. For this reason, events with additional clusters in the LKr beyond those associated to the photons and to the track (if present) in the triplet, are vetoed. In particular there must be no extra clusters with energy greater than 1 GeV in a time window of ± 7 ns around the track time (t_{RICH}). The half width of the window accounts for the maximum time difference between one cluster and the downstream track if they both originate from the same K^+, and it is obtained from data as follows. A loose set of conditions is applied on minimum bias data selecting a good downstream track matched with a K^+ according to the same conditions used in the signal and normalization selections. Then, a couple of clusters is selected. The first cluster in required in time with the track within ± 3 ns ($|t_{cls1} - t_{RICH}| < 3$ ns); the second cluster is searched in different time intervals with respect to the track time: $|t_{cls2} - t_{RICH}| < 3\,\mathrm{ns}$, $3 < |t_{cls2} - t_{RICH}| < 5\,\mathrm{ns}, 5 < |t_{cls2} - t_{RICH}| < 6\,\mathrm{ns}, 6 < |t_{cls2} - t_{RICH}| < 7\,\mathrm{ns}$, $|t_{cls2} - t_{RICH}| > 7\,\mathrm{ns}$. The di-photon invariant mass of the selected clusters is computed for each time interval between the track and the second cluster and is reported in Figure 5.15. The π^0 mass peak is evident with the second cluster up to 6 ns far from the track time, while the $m_{\gamma\gamma}$ distribution becomes smooth if it is 7 ns or more distant from the track.

5.3.2 LAV veto

The kinematics of the $K^+ \to \pi^+\pi^0\pi^0$ decay is such that a π^0 can be emitted at large angles with respect to the K^+ direction of flight and the photons coming from the π^0 can be detected by one of the LAV stations surrounding the volume of the detector (see Sec.2.5.3).

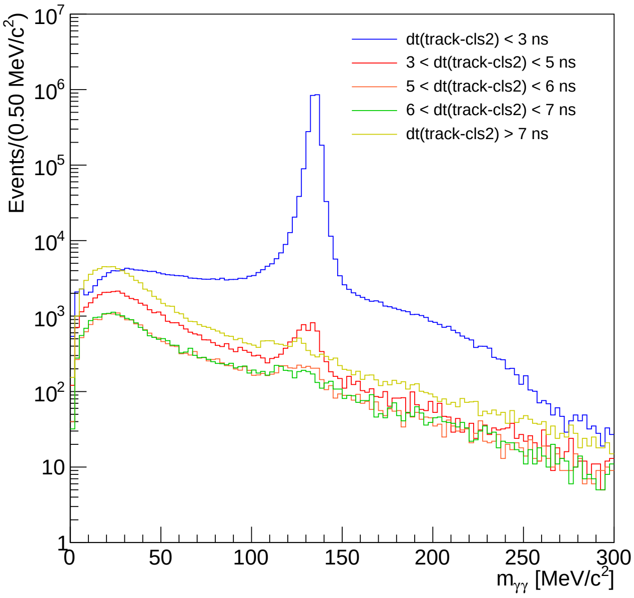


Figure 5.15: $m_{\gamma\gamma}$ distribution in minimum bias data after the triplet selection where one photon is in time with the track within $\pm 3\,$ns ($|t_{cls1} - t_{RICH}| < 3\,$ns) while the other is searched in five different exclusive time intervals: $|t_{cls2} - t_{RICH}| < 3\,$ns (blue), $3 < |t_{cls2} - t_{RICH}| < 5\,$ns (red), $5 < |t_{cls2} - t_{RICH}| < 6\,$ns (orange), $6 < |t_{cls2} - t_{RICH}| < 7\,$ns (green) and $|t_{cls2} - t_{RICH}| > 7\,$ns (yellow). The π^0 mass peak is evident as long as the second photon is within 7 ns from the track.

To check in the NA62 data for the presence of in-time activity in the LAV stations compatible with the reconstructed decay, the standard NA62 LAV matching algorithm is used. For the matching, the triplet reference time $t_{ref} = (t_{RICH} + t_{cls1} + t_{cls2})/3$ and a time window of $\pm 3\,$ns around t_{ref} is considered. To reduce random veto effects coming from accidental signals in the LAVs, mainly due to the muon halo of beam particles decays, only the stations located downstream of the z coordinate of the reconstructed vertex (Z_{chVtx}) are considered for each event.

The LAV matching algorithm, however, is not suitable for MC simulations. Since the Geant4 based NA62 simulation does not provide sufficient accuracy in the modeling of photo-nuclear interactions, the photon veto response, especially for low energy photons, is not sufficiently well simulated. Therefore, the LAV photon detection inefficiency has been measured in bins of the photon energy from a sample of $K^+ \to \pi^+\pi^0$ events [86] and the obtained values are used to correct the LAV photon veto inefficiency for each MC event. By using the information of true production point and three-momentum of a simulated photon, its direction is extrapolated forward to check if it is or not in the LAV geometrical acceptance. If the photon intercepts one of the LAV stations, the corresponding measured inefficiency, which depends on the photon true energy, is used as weight to assign to the MC event. A photon is assumed to be in the LAV geometrical acceptance if it intercepts either the front or the back plane of at least one LAV station within the nominal transverse acceptance of the station. Photon conversions are included in this algorithm as follows. If the photon converts in an e^+e^- pair in one of the STRAW chambers, the $e^\pm$ direction of flight are extrapolated instead of the photon one, taking into account the effect of the spectrometer magnetic field. The detection inefficiencies for electrons and positrons are assumed to be equal to the inefficiencies measured for photons of the same energy.

To align the algorithm used in data with that used for simulation, only photons (or $e^\pm$) with time included in the interval ($t_{ref} - 3$ns, $t_{ref} + 3$ ns) and the LAV stations downstream of the reconstructed vertex Z_{chVtx} are considered. However, the algorithm still doesn't take into account several effects present in the data. As an example, the algorithm, by construction, is not applied to random veto, that is mainly due to halo muons.

5.4 Upstream decays suppression

The decays of the beam K^+ are expected to be well reconstructed if they occur in the z range (102.425 m, 180 m). However, events with K^+ decays occurring upstream to this range can pass the selection due to a mis-reconstructed vertex. Of particular interest for this analysis are K^+ that decay in the range (96.950 m, 102.425 m), between the third and fourth bending magnets of the GTK achromat and up to the GTK3 station.

As an example, Figure 5.16 reports the distribution of the reconstructed di-photon invariant mass as a function of the z coordinate of the reconstructed vertex (Z_{chVtx}) for $K^+ \to \mu^+\nu_\mu\pi^0$ decays from the normalization sample, after whole the conditions defined above (triplet selection, photon conversion rejection, kinematic constraints with tight $m^2_{miss}(K - \pi^0)$ cut and photon vetoes) are applied. In the plot on the left hand side, K^+ decays are simulated in the standard decay region (102.425 m, 180 m); $m_{\gamma\gamma}$ peaks at the π^0 mass value independently from the value of Z_{chVtx}, as expected for well reconstructed $K^+ \to \mu^+\nu_\mu\pi^0$ decays. On the right hand side, K^+ decays are simulated in the upstream region. Due to the kick provided to the muon by the TRIM magnet, the vertex, which is reconstructed as mid-point of the closest distance of approach of the measured muon and kaon directions, is shifted towards higher z with respect to the its true value. As a consequence, $m_{\gamma\gamma}$ is linearly correlated with Z_{chVtx} and shifted towards higher values with respect to the π^0 mass. The shift is shown in Figure 5.17, where the projections along the $m_{\gamma\gamma}$ axis of the distributions in Figure 5.16 are plotted, each normalized to unitary area for comparison.

To reduce upstream events, that pass both the signal and the normalization selections, 110 m $< Z_{chVtx}$ is required. To ensure the quality of the events, the condition $Z_{chVtx} < 170$ m must also be satisfied.

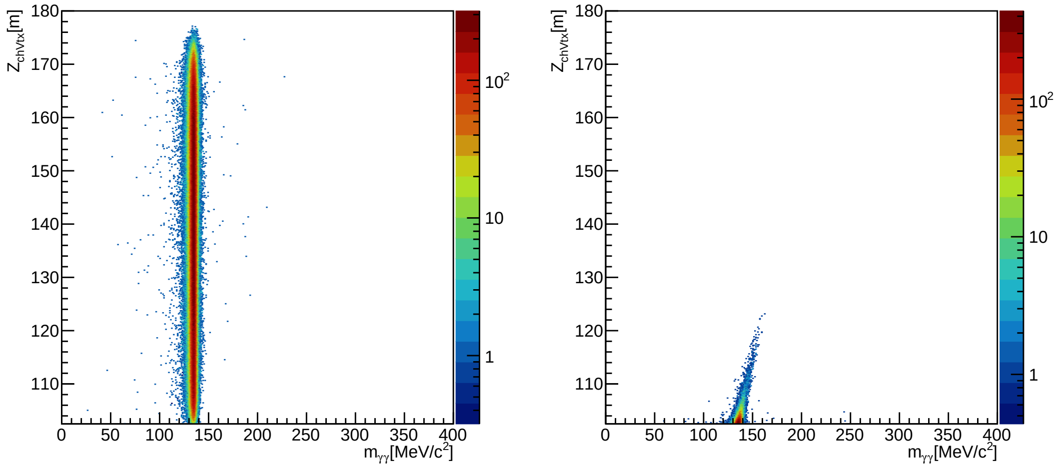


Figure 5.16: Correlation between the z coordinate of the reconstructed vertex and $m_{\gamma\gamma}$ for $K^+ \to \mu^+\nu_\mu\pi^0$ decays simulated in the standard decay region ($102.425\,\mathrm{m} < Z_{\mathrm{true\ vertex}} < 180\,\mathrm{m}$), on the left, and in the upstream region ($96.950\,\mathrm{m} < Z_{\mathrm{true\ vertex}} < 102.425\,\mathrm{m}$), on the right, and reconstructed in the standard decay region. The distributions are obtained after the normalization selection is applied but $110 < m_{\gamma\gamma} < 160\,\mathrm{MeV/c^2}$ (see Chapter 4) and are not normalized to data.

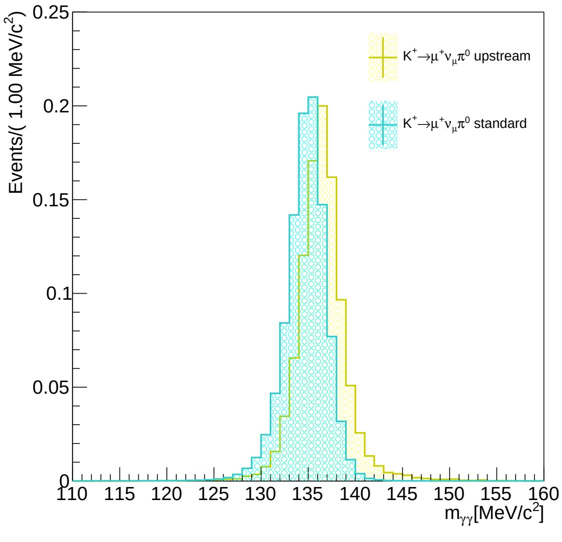


Figure 5.17: Di-photon invariant mass distribution after the normalization selection but $110 < m_{\gamma\gamma} < 160\,\mathrm{MeV/c^2}$. $K^+ \to \mu^+\nu_\mu\pi^0$ decays are simulated in the standard decay region ($102.425\,\mathrm{m} < Z_{\mathrm{true\ vertex}} < 180\,\mathrm{m}$) and in the upstream decay region ($96.950\,\mathrm{m} < Z_{\mathrm{true\ vertex}} < 102.425\,\mathrm{m}$), and their vertex is reconstructed in standard region. Both the distributions are normalized to unitary area for comparison. A shift of the $m_{\gamma\gamma}$ distribution with respect to the π^0 mass value is evident for the decays occurring in the upstream region.

Chapter 6

Statistical analysis

6.1 Number of K^+ decays

In the NA62 experiment the flux of K^+ decays is not known a priori. Thus, as explained in Sec.3.1, the effective number of K^+ decays (N_K) occurred in the z range considered in the $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ search, $110 - 170$ m, is estimated selecting a clean sample of $K^+ \to \mu^+\nu_\mu\pi^0$ decays, which have a topology similar to the signal, a well known branching ratio and are collected by the same trigger line as signal events.

To select well reconstructed $K^+ \to \mu^+\nu_\mu\pi^0$ ($K_{\mu3}$) candidates, the normalization selection (see Chapter 4 for a summary) is applied on NA62 data and MC samples, as shown in Figure 6.1. The mass range $130 < m_{\gamma\gamma} < 140\,\mathrm{MeV/c^2}$ is used to determine the acceptance of the normalization selection ($A^{K_{\mu3}}_{norm}$); the resolution of the π^0 mass peak is $\sigma_{m_{\gamma\gamma}} \sim 2$ MeV/c^2. $A^{K_{\mu3}}_{norm}$ is measured from the $K_{\mu3}$ simulated sample:

$$A^{K_{\mu3}}_{norm} = (8.247 \pm 0.004) \times 10^{-3}. \tag{6.1}$$

Given the observed number of normalization events in data ($N^{data}_{norm} = 3.6521 \times 10^6$), the $BR(K^+ \to \mu^+\nu_\mu\pi^0) = (3.352 \pm 0.033)\%$ [63] and the $BR(\pi^0 \to \gamma\gamma) = (98.823 \pm 0.034)\%$ [63], N_K is computed according to Eq.3.5, neglecting both the trigger efficiency and the downscaling factor, as they have the same values for signal and normalization. The result is:

$$N_K = (1.3368 \pm 0.0039) \times 10^{10} \tag{6.2}$$

Possible contamination to the normalization sample due to other SM processes has been found to be negligible ($\mathcal{O}(10^{-6})$ or less, depending on the K^+ decay mode).

The error budget of N_K is reported in Table 6.1. The first value (stat) accounts for the statistical error due to the Poissonian fluctuation of the number of observed $K^+ \to \mu^+\nu_\mu\pi^0$ events in the normalization data sample. The second (MC) is associated to the dimension of the MC sample used to evaluate the acceptance $A^{K_{\mu3}}_{norm}$. The dominant uncertainties are the external one (ext), coming from the knowledge of the SM branching ratios, and the systematic one (sys). The quoted systematic contribution accounts for the variation of N_K due to the not ideal data/MC agreement above

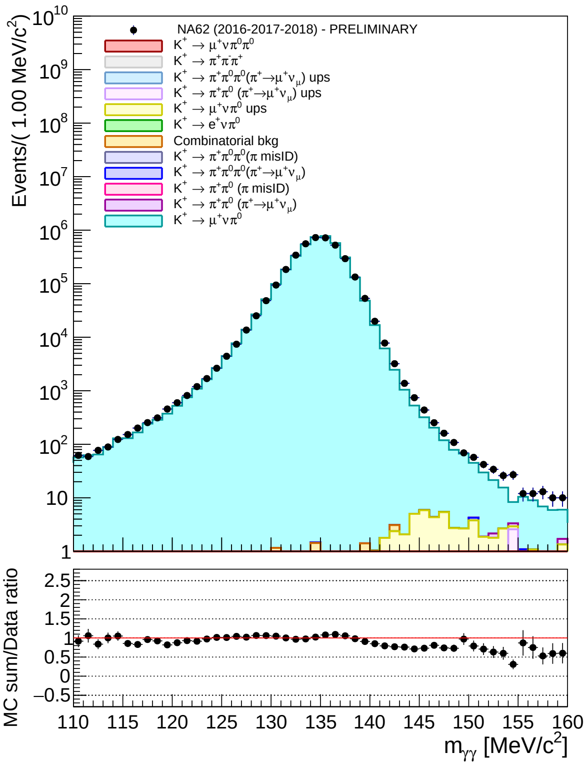
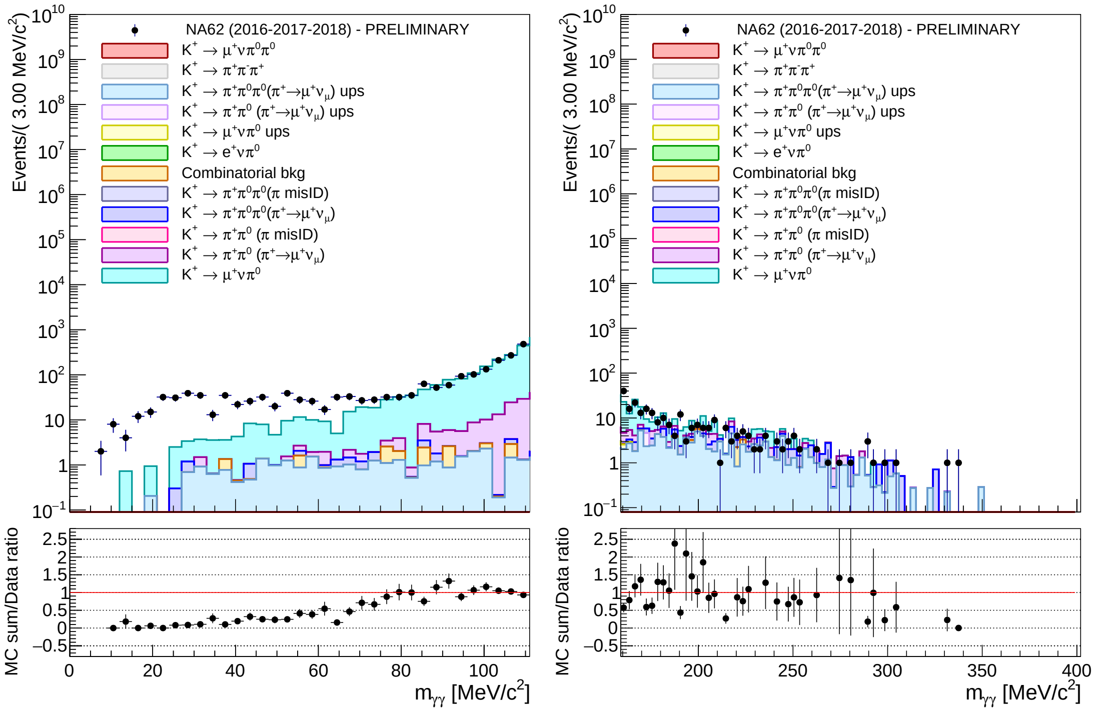


Figure 6.1: $m_{\gamma\gamma}$ distribution after the normalization selection. NA62 Run1 data (dots) are superimposed to stacked MC samples, each normalized to data according to the value of N_K in Eq.6.2. All the MC samples listed in Table 3.2 are included. In the legend, *ups* specifies samples in which the decay is simulated in the upstream region (see Sec.3.3.2), while the *Combinatorial bkg* includes events where the selected track and one or both the clusters come from different decays. The bottom part of the plot shows the data-MC agreement.

	absolute [$\times 10^{10}$]	relative [$\times 10^{-4}$]
stat	0.0007	5.16
MC	0.0006	4.29
ext	0.0014	10.43
sys	0.0035	26.18
tot	0.0039	28.97

Table 6.1: The different contributions to the uncertainty on the N_K measurement. The external contribution is due to the knowledge of the SM branching ratios used to compute N_K.

137 MeV/c^2, shown in Figure 6.1, and has been evaluated by varying the bounds of the $m_{\gamma\gamma}$ interval used to compute N_K. As a first order cancellation is expected when measuring $BR(K^+ \to \mu^+\nu_\mu X, X \to \gamma\gamma)$ with respect to the normalization channel, for the aim of this work, other systematic uncertainties are neglected. Nevertheless, further investigations of systematic uncertainties are needed, together with a more accurate estimate of random veto effects and trigger efficiency losses.



(a) $R1 : 0 < m_{\gamma\gamma} < 110\,\mathrm{MeV/c^2}$

(b) $R2 : 160 < m_{\gamma\gamma} < 380\,\mathrm{MeV/c^2}$

Figure 6.2: $m_{\gamma\gamma}$ distribution after the signal selection; only the regions sensitive to the $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ process are shown. NA62 Run1 data (dots) are superimposed to stacked MC samples, each normalized to data according to the value of N_K in Eq.6.2. All the MC samples listed in Table 3.2 are included. In the legend, *ups* specifies samples in which the decay is simulated in the upstream region (see Sec.3.3.2), while the *Combinatorial bkg* includes events where the selected track and one or both the clusters come from different decays. The bottom part of the plot shows the data-MC agreement.

6.2 Di-photon invariant mass distribution in R1 and R2

Figure 6.2a and 6.2b show the $m_{\gamma\gamma}$ data distribution after the signal selection for R1 and R2, respectively. The sum of the different background MC events passing the selection is superimposed to data; each contribution is normalized according to its branching ratio and to the measured N_K (Eq. 6.2). The bottom part of each plot shows the agreement between data and the sum of the considered SM backgrounds.

In R1, a good agreement is achieved for $m_{\gamma\gamma}$ values higher than $80\,\mathrm{MeV/c^2}$, where the dominant contributions are residual badly reconstructed $K^+ \to \mu^+\nu_\mu\pi^0$ events and $K^+ \to \pi^+\pi^0, \pi^+ \to \mu^+\nu_\mu$ events. Nevertheless, a discrepancy is present below $80\,\mathrm{MeV/c^2}$, that needs to be further investigated and that is probably due to SM backgrounds not included or not well reproduced in the considered simulations. In R2, the sum of the MC samples agrees with the data distribution throughout the mass range, with a lot of fluctuations present due to the very low statistics in data.

Excluding the range $m_{\gamma\gamma} < 80\,\mathrm{MeV/c^2}$, a χ^2 test comparing the data histogram with the sum of the normalized MC samples, performed by using the ROOT method *Chi2Test*, provides a $\chi^2/NDF = 11.05/10$ and $p - value = 0.35$ in R1 and $\chi^2/NDF = 15.79/12$ and $p - value = 0.20$ in R2. Note that in R2 the test only includes bins up to 310 MeV/c^2, where the statistics is enough in both data and summed MC samples; a rebin is eventually needed for the test to be meaningful. As the data distribution is not significantly different from the SM expectation, upper limits for the signal branching ratios are established.

6.3 $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ search

The search for the $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ process is conducted through a mass scan. For the moment, only the simulated X mass hypotheses included in R1 or R2 are considered, meaning that the scan is performed in mass steps of 10 MeV/c^2 between $10 - 100$ MeV/c^2 and $170 - 380$ MeV/c^2, excluding the mass hypotheses at the edges of the regions (110 MeV/c^2 and 160 MeV/c^2).

As the decay is characterized by a peak in the reconstructed di-photon invariant mass distribution, a signal region:

$$|m_{\gamma\gamma} - m_X| < 2\sigma_m, \tag{6.3}$$

can be defined for each X mass hypothesis to test, where σ_m is the resolution of the reconstructed $X \to \gamma\gamma$ resonance, evaluated with signal simulations. For a fixed mass m_X, the signal search is here considered as a counting experiment with a single observation channel. Since in a counting experiment the only observable is the total number of selected events (N), that is modeled as a Poisson distributed variable, the background model is completely defined once the estimate of the expected number of background events (b) in the signal region ($m_X \pm 2\sigma_m$) is obtained, together with its uncertainty db. In R2, although the contributions to the tail in data are quite well understood, the precision of a data-driven background estimate is expected to be better than a MC-driven one. Hence, a data-driven approach is adopted to determine the expected background in each signal region, using sideband data events in the reconstructed $m_{\gamma\gamma}$ spectrum. This method is only valid provided there are no other peaking background structures in the regions sensitive to the signal. Therefore, as a uniform distribution is observed also below 80 MeV/c^2, this method can also be used to set upper limits for X masses included in this range, regardless of the data/MC agreement.

6.4 Mass resolution and signal regions

The mass resolution σ_m is, by definition, the width of the reconstructed di-photon invariant mass distribution obtained for each simulated X mass at the end of the signal selection. The reconstructed $m_{\gamma\gamma}$ distributions are reported in Figure 6.3 for all the available signal samples, each color corresponding to a different m_X value. The width

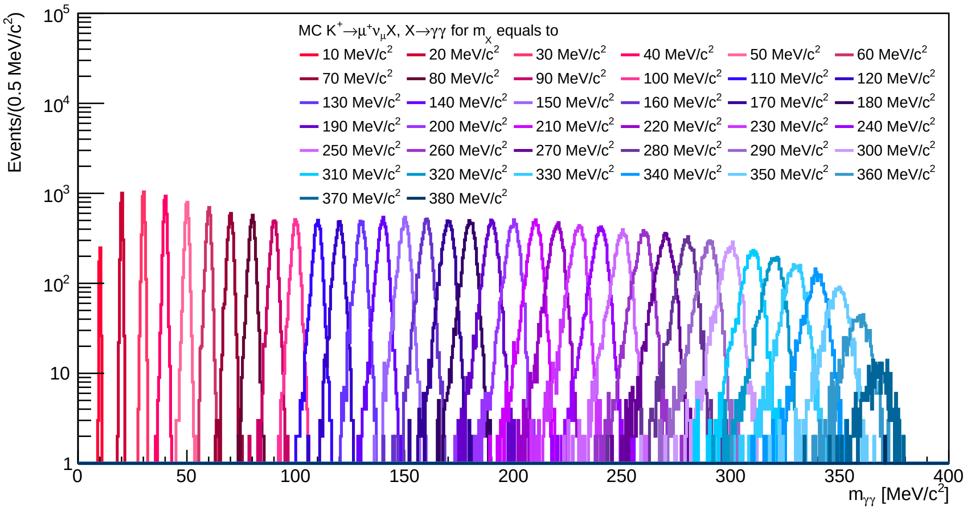


Figure 6.3: $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ MC samples. Reconstructed di-photon invariant mass after the signal selection for each simulated mass of X (corresponding to a different color in the plot).

of the mass peaks is measured here through Gaussian fits to the peaks; some examples are reported in Figure 6.4.

For increasing X mass, $m_{\gamma\gamma}$ present asymmetrical tails, with the left tail probably due to residual photon conversion contributions in the sample or to a not complete containment of the showers in the calorimeter. As a consequence, the mean value from the Gaussian fit (μ_X) underestimates the true X mass value for increasing m_X.

Figure 6.5a reports the relative difference between μ_X and m_X as a function of the X mass. A negative trend is present, that can be parameterized as a linear function (red fit in the same Figure), with some deviations from the trend for m_X lower than 30 MeV/c^2 and higher than 350 MeV/c^2. The mass hypotheses $m_X = 370\,\mathrm{MeV/c^2}$ and $m_X = 380\,\mathrm{MeV/c^2}$ are not included in the plot as, due to the very low number of selected events, the fits to their distributions are not reliable. It has to be noted, however, that μ_X is consistent with the true value of m_X within few per mil relative for each mass, and that the presence of the negative trend is not relevant for this study as long as the resolution of the mass resonances is well reproduced from the Gaussian width. The obtained resolutions for the same mass hypotheses included in Figure 6.5a are reported in Figure 6.5b as a function of m_X. The increase of σ_m with the mass is well described by a third order polynomial (red fit in the same Figure). The values of the resolution provided from the polynomial fit ($\sigma_m(m_X)$) deviate from the Gaussian sigma values only for those m_X for which the Gaussian fit shows some inconsistencies; the two values are summarized in Table 6.2. The mass resolution provided from the polynomial fit is therefore assumed to be reliable even when the Gaussian fit to the $X \to \gamma\gamma$ resonance is not and is used to define the limits of the signal regions $|m_{\gamma\gamma} - m_X| < 2\sigma_m(m_X)$ for every simulated mass up to $380\,\mathrm{MeV/c^2}$.

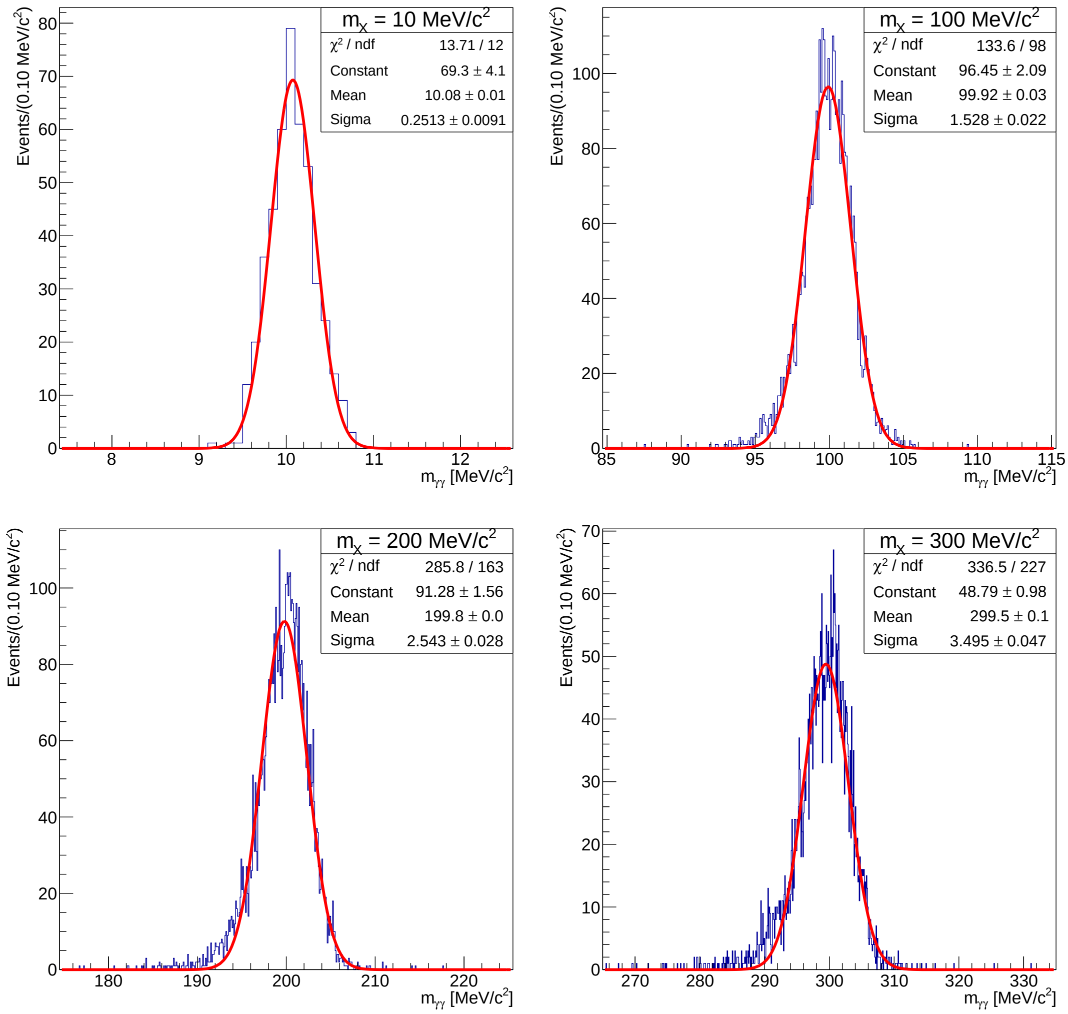
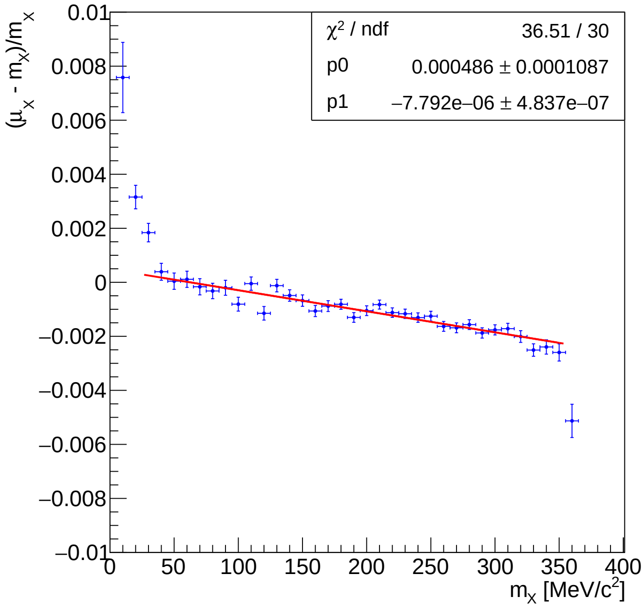
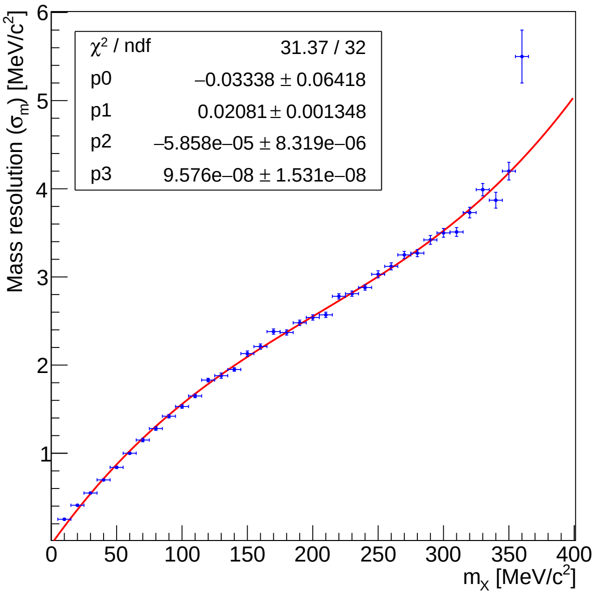


Figure 6.4: Di-photon invariant mass distribution for signal MC samples with $m_X = 10, 100, 200, 300$ MeV/c^2 after all the $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ selection conditions are applied. Gaussian fits (in red) are performed in order to obtain mass resolutions.



(a) Relative difference $(\mu_X - m_X)/m_X$ vs m_X



(b) Mass resolution σ_m vs m_X

Figure 6.5: Relative difference $(\mu_X - m_X)/m_X$ (top) and mass resolution σ_m (bottom) as a function of the simulated X mass for the signal MC samples (blue dots) and polynomial fits describing their trends (red function). μ_X and σ_m are the mean and the sigma parameters defining the Gaussian function fitting the reconstructed $m_{\gamma\gamma}$ distribution for a given simulated m_X (examples reported in Figure 6.4).

m_X [MeV/c^2]	resolution from Gaussian fit [MeV/c^2]	$\sigma_m(m_X)$ from polynomial fit [MeV/c^2]
10	0.251 ± 0.009	0.17 ± 0.05
20	0.410 ± 0.006	0.36 ± 0.04
30	0.549 ± 0.009	0.54 ± 0.04
40	0.698 ± 0.01	0.71 ± 0.03
50	0.84 ± 0.01	0.87 ± 0.02
60	1.00 ± 0.01	1.03 ± 0.02
70	1.15 ± 0.02	1.17 ± 0.02
80	1.28 ± 0.02	1.31 ± 0.02
90	1.42 ± 0.02	1.43 ± 0.02
100	1.53 ± 0.02	1.56 ± 0.02
110	1.65 ± 0.02	1.67 ± 0.02
120	1.83 ± 0.02	1.79 ± 0.02
130	1.88 ± 0.03	1.89 ± 0.02
140	1.95 ± 0.02	1.99 ± 0.02
150	2.13 ± 0.03	2.09 ± 0.02
160	2.21 ± 0.03	2.19 ± 0.02
170	2.38 ± 0.03	2.28 ± 0.02
180	2.37 ± 0.03	2.37 ± 0.02
190	2.48 ± 0.03	2.46 ± 0.02
200	2.54 ± 0.03	2.55 ± 0.02
210	2.57 ± 0.03	2.64 ± 0.02
220	2.78 ± 0.03	2.73 ± 0.02
230	2.81 ± 0.03	2.82 ± 0.02
240	2.88 ± 0.03	2.91 ± 0.02
250	3.03 ± 0.04	3.00 ± 0.02
260	3.12 ± 0.04	3.10 ± 0.02
270	3.25 ± 0.04	3.20 ± 0.02
280	3.27 ± 0.04	3.30 ± 0.02
290	3.42 ± 0.05	3.41 ± 0.02
300	3.5 ± 0.05	3.52 ± 0.02
310	3.51 ± 0.05	3.64 ± 0.03
320	3.73 ± 0.06	3.77 ± 0.03
330	3.99 ± 0.07	3.90 ± 0.04
340	3.87 ± 0.09	4.03 ± 0.05
350	4.2 ± 0.1	4.18 ± 0.06
360	5.5 ± 0.3	4.34 ± 0.07

Table 6.2: Mass resolution for each m_X value obtained from $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ MC events. The second column reports the value of the resolution as measured from the Gaussian fit to the corresponding $m_{\gamma\gamma}$ distribution. The third column reports the value of the resolution as provided by the polynomial fit in Figure 6.5b ($\sigma_m(m_X)$).

6.5 Signal acceptance

Figure 6.6a shows the $(m_{\gamma\gamma} - m_X)$ distribution as a function of m_X. The two red lines, given by $m_{\gamma\gamma} = m_X \pm 2\sigma_m(m_X)$, delimit the signal search interval for each m_X value.

To determine the branching ratio of the process, the $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ acceptance relative to the requirement $|m_{\gamma\gamma} - m_X| < 2\sigma_m(m_X)$ has to be provided as input to Eq.3.4. The acceptance is evaluated from simulations as:

$$A(m_X) = \frac{N_{sel}^{s.i.}(m_X)}{N_{gen}(m_X)} \tag{6.4}$$

with error:

$$\sigma_A = \sqrt{\frac{A(m_X)(1 - A(m_X))}{N_{gen}(m_X)}} \tag{6.5}$$

where N_{gen} is the total number of generated events with X mass equal to m_X and $N_{sel}^{s.i.}$ is the number of events passing the signal selection and included in the signal region. The uncertainty is evaluated assuming for $N_{sel}^{s.i.}$ the binomial probability distribution. The total signal selection acceptance can also be evaluated and is defined as:

$$A_{tot}(m_X) = \frac{N_{sel}^{tot}(m_X)}{N_{gen}(m_X)} \tag{6.6}$$

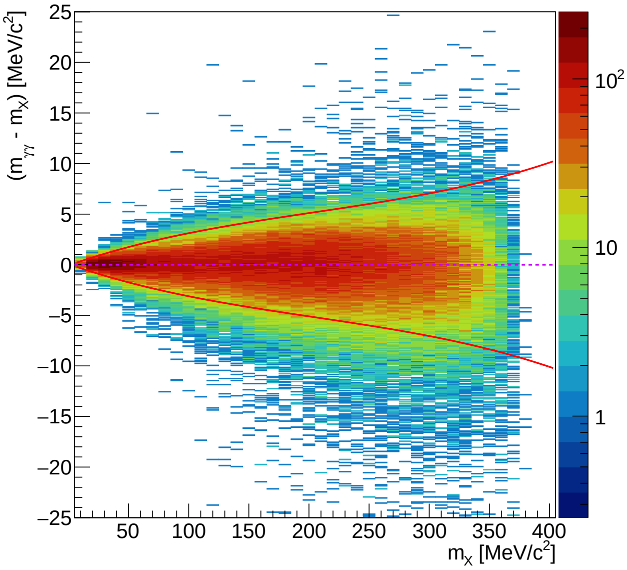
$$\sigma_{A_{tot}} = \sqrt{\frac{A_{tot}(m_X)(1 - A_{tot}(m_X))}{N_{gen}(m_X)}}, \tag{6.7}$$

where N_{sel}^{tot} is the total number of selected events before $|m_{\gamma\gamma} - m_X| < 2\sigma_m(m_X)$ is required. Table 6.3 reports the values of both $A_{tot}(m_X)$ and $A(m_X)$ for every m_X simulated up to $380\,\mathrm{MeV/c^2}$.

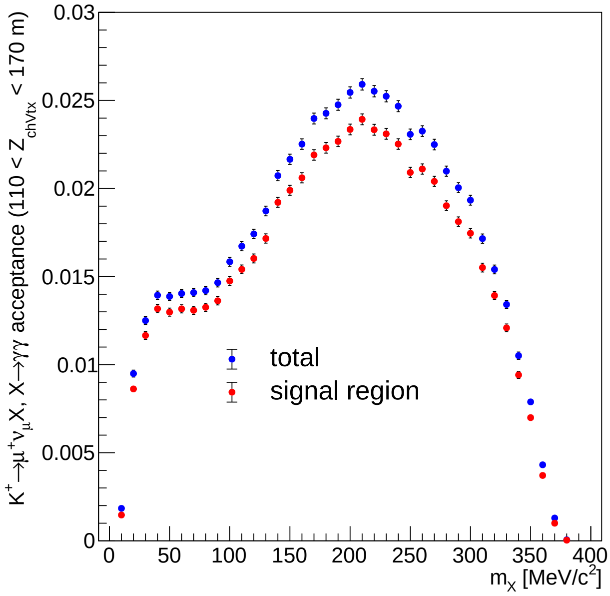
As expected from the preliminary study of the acceptance (see Sec.3.3.1), A_{tot} decreases for $m_X < 50\,\mathrm{MeV/c^2}$ and the closer m_X is to the kinematic end point. The effects due to the detector geometry are here convoluted with the effects of the reconstruction and of the selection. In particular, the drop at low masses is mainly due to the selection conditions reducing the $K^+ \to \pi^+\pi^0, \pi^+ \to \mu^+\nu_\mu$ background. The $4\sigma_m(m_X)$ wide mass window defined for the signal search selects good quality signal events, with a loss with respect to the total fraction of selected events of few percent up to $\mathcal{O}(10\%)$ for different masses, as shown in Figure 6.6b.

6.6 Data-driven background estimate

For the data-driven background estimate, in each X mass hypothesis considered, sidebands are defined on both sides of the search interval $m_X \pm 2\sigma_m(m_X)$ in the reconstructed $m_{\gamma\gamma}$ spectrum. The expected background in the signal region is estimated from local polynomial fits to the sideband data. The width of the sideband varies for different X masses, depending on the mass resolution and on the available statistics



(a) Signal region $|m_X - m_{\gamma\gamma}| < 2\sigma_m(m_X)$.



(b) Signal acceptance vs m_X

Figure 6.6: $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ MC samples after the signal selection. Top: difference between the reconstructed di-photon invariant mass and the simulated mass of X as a function of m_X. The red lines delimit the signal region $|m_X - m_{\gamma\gamma}| < 2\sigma_m(m_X)$ for each m_X value, with $\sigma_m(m_X)$ value of the resolution from the polynomial fit in Figure 6.5b. Bottom: $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ acceptance as a function of m_X. In blue, the total selection efficiency defined by Eq.6.6 is reported; in red, the search interval selection efficiency defined by Eq.6.4 is shown.

m_X [MeV/c^2]	$A_{tot}(m_X)$	$A(m_X)$
10	0.00184 ± 0.00009	0.00147 ± 0.00008
20	0.0095 ± 0.0002	0.0086 ± 0.0002
30	0.0125 ± 0.0002	0.0117 ± 0.0002
40	0.0139 ± 0.0002	0.0132 ± 0.0002
50	0.0139 ± 0.0002	0.0130 ± 0.0002
60	0.0140 ± 0.0002	0.0132 ± 0.0002
70	0.0141 ± 0.0002	0.0131 ± 0.0002
80	0.0142 ± 0.0002	0.0133 ± 0.0002
90	0.01467 ± 0.0002	0.0136 ± 0.0002
100	0.0158 ± 0.0003	0.0147 ± 0.0002
110	0.0167 ± 0.0003	0.0154 ± 0.0003
120	0.0174 ± 0.0003	0.0160 ± 0.0003
130	0.0187 ± 0.0003	0.0172 ± 0.0003
140	0.0207 ± 0.0003	0.0192 ± 0.0003
150	0.0217 ± 0.0003	0.0199 ± 0.0003
160	0.0225 ± 0.0003	0.0206 ± 0.0003
170	0.0240 ± 0.0003	0.0219 ± 0.0003
180	0.0243 ± 0.0003	0.0223 ± 0.0003
190	0.0247 ± 0.0003	0.0227 ± 0.0003
200	0.0255 ± 0.0003	0.0234 ± 0.0003
210	0.0259 ± 0.0003	0.0239 ± 0.0003
220	0.0255 ± 0.0003	0.0233 ± 0.0003
230	0.0252 ± 0.0003	0.0231 ± 0.0003
240	0.0247 ± 0.0003	0.0225 ± 0.0003
250	0.0231 ± 0.0003	0.0209 ± 0.0003
260	0.0233 ± 0.0003	0.0211 ± 0.0003
270	0.0225 ± 0.0003	0.0204 ± 0.0003
280	0.0210 ± 0.0003	0.0190 ± 0.0003
290	0.0200 ± 0.0003	0.0181 ± 0.0003
300	0.0193 ± 0.0003	0.0175 ± 0.0003
310	0.0172± 0.0003	0.0155 ± 0.0003
320	0.0154 ± 0.0003	0.0139 ± 0.0002
330	0.0134 ± 0.0002	0.0121 ± 0.0002
340	0.0105 ± 0.0002	0.0094 ± 0.0002
350	0.0079 ± 0.0002	0.0070 ± 0.0002
360	0.0043 ± 0.0001	0.0037 ± 0.0001

Table 6.3: $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ acceptance. The first column reports the signal selection acceptance (A_{tot}) defined in Eq.6.6. The second column reports the acceptance relative to the search interval (A) defined in Eq.6.4.

in data. For $m_{\gamma\gamma} < 20\,\mathrm{MeV/c^2}$ or $> 320\,\mathrm{MeV/c^2}$ it is not possible to perform local fits due to the lack of statistic; these mass hypotheses are therefore neglected in the following. A different method would be needed to account also for them.

Different functions are considered for different mass regions, depending on the shape of the data $m_{\gamma\gamma}$ distribution (examples of fits are reported in Figure 6.7):

- first order polynomial if $m_X < 80\,\mathrm{MeV/c^2}$ or $> 190\,\mathrm{MeV/c^2}$, where $m_{\gamma\gamma}$ is almost flat;
- second order polynomial if $80 \leq m_X \leq 90\,\mathrm{MeV/c^2}$ or $180 \leq m_X \leq 190\,\mathrm{MeV/c^2}$;
- third order polynomial if $90 < m_X < 110\,\mathrm{MeV/c^2}$ or $160 < m_X < 180\,\mathrm{MeV/c^2}$, that are the mass regions closest to the π^0 resonance.

The expected number of background events (b) is then evaluated as the integral of the fitting function (f) over the signal region divided by the bin width of the fitted histogram:

$$b = \frac{1}{w_{bin}(m_X)} \int_{m_X - 2\sigma_m(m_X)}^{m_X + 2\sigma_m(m_X)} f(m_{\gamma\gamma}) dm_{\gamma\gamma}. \tag{6.8}$$

For each tested mass, the histogram is built with bin width $w_{bin}(m_X) = 0.5\sigma_m(m_X)$. The uncertainty on the number of expected background events (db) is then evaluated including a statistical (db_{fit}) and a systematic (db_{sys}) component. The former comes from the propagation of the statistical errors on the parameters of the fitting function and is quoted by using the ROOT *TF1::IntegralError* method, that takes into account the covariance matrix of the statistical errors in the fit parameters. The latter is estimated as half of the maximum difference between the central values b_i obtained both from the reference fit function and from fits using polynomials of order $n-1$ and $n+1$ with respect to the reference fit function, and varying the width of the sidebands.

All the obtained b values with their uncertainties are listed in Table 6.4. The systematic uncertainty is the dominant one the closer m_X is to the π^0 resonance, otherwise, it is comparable to the uncertainty from the fit.

6.7 Statistical framework

Once $b \pm db$ is known, the experimental sensitivity to the $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ process is evaluated first:

$$BR_{sensitivity}(m_X, N_K) = BR_{SES}(m_X, N_K) \cdot Exp.\,UL(m_X), \tag{6.9}$$

where $BR_{SES}(m_X, N_K)$ is the branching ratio that can be measured if a single signal event is observed and $Exp.\,UL(m_X)$ is the expected upper limit at the 90% confidence level (C.L.) of the number of signal events. The $Exp.\,UL(m_X)$ at a fixed C.L. is the limit that can be set in the absence of signal, due to the fluctuations of the background, as will be specified in the following.

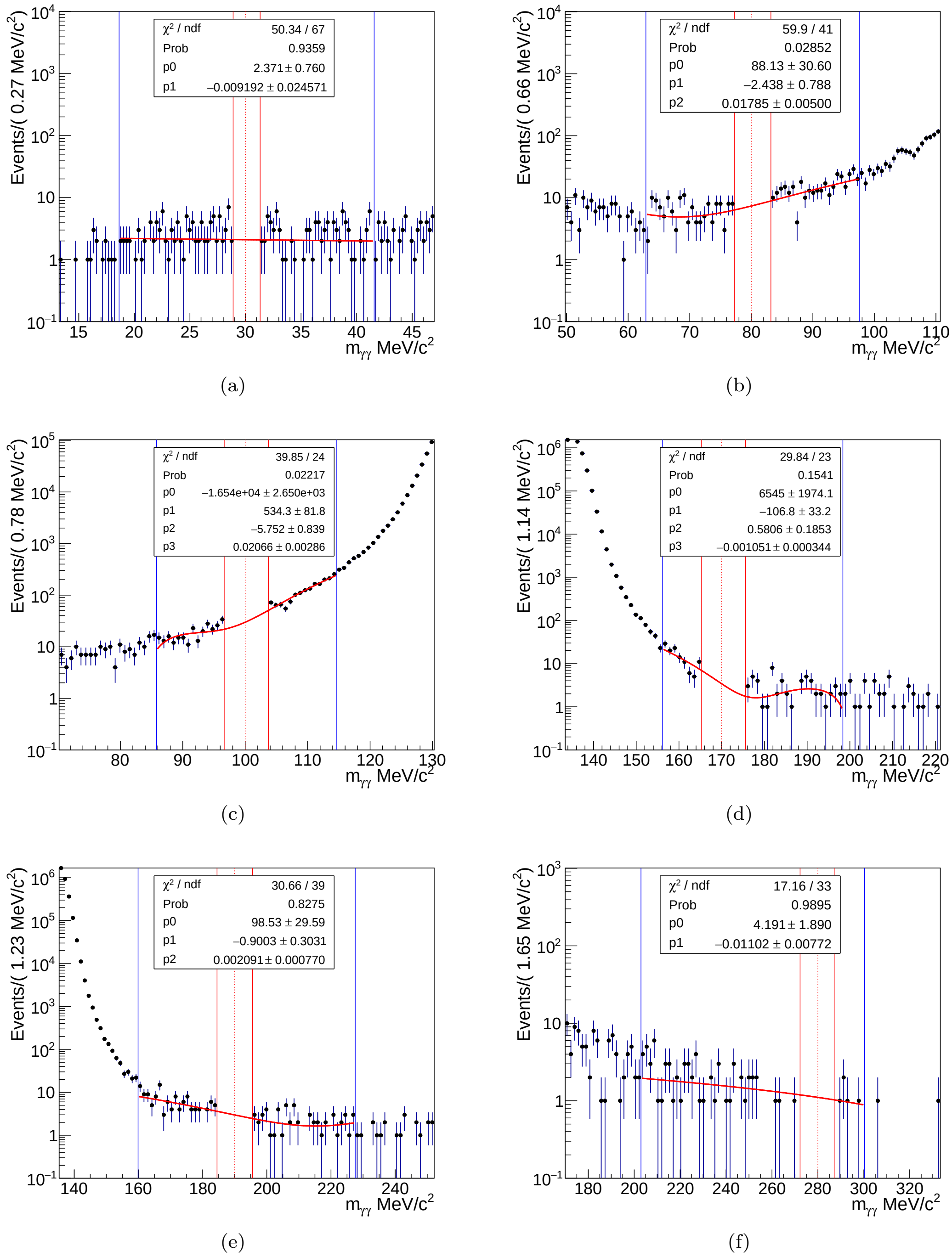
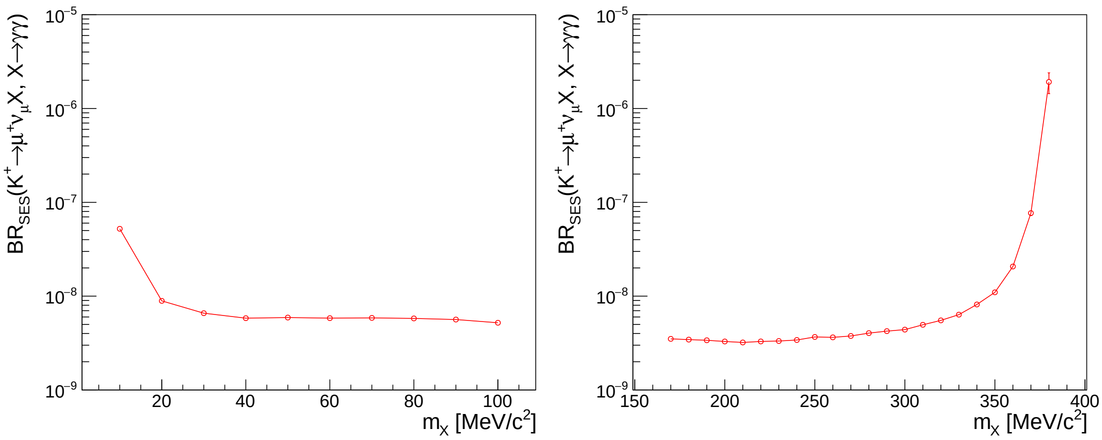


Figure 6.7: Example of the local polynomial fits to data used for the estimate of the expected background in the signal regions defined for the $m_X = 30, 80, 100, 170, 190$ and $280\,\mathrm{MeV/c^2}$, from the top left to the bottom right, respectively. For each tested m_X, the data $m_{\gamma\gamma}$ spectrum is built with bin width $w_{bin}(m_X) = \sigma_m(m_X)/2$. The limits of the sidebands are represented by the blue lines. In each plot, the bins included in the signal regions (vertical red lines) defined around the X mass to test (vertical dashed red line) are empty as they are kept blind to the fit.

m_X [MeV/c^2]	$b \pm db_{fit}$	db_{sys}	$db = \sqrt{db_{sys}^2 + db_{fit}^2}$
20	13 ± 2	2	3
30	19 ± 2	3	4
40	20 ± 2	2	3
50	27 ± 2	5	5
60	36 ± 3	8	9
70	43 ± 3	8	9
80	67 ± 6	11	13
90	133 ± 12	14	18
100	294 ± 23	66	70
170	32 ± 9	17	19
180	31 ± 4	8	9
190	27 ± 4	4	6
200	23 ± 2	5	5
210	16 ± 2	3	4
220	16 ± 2	1	2
230	14 ± 2	1	2
240	15 ± 3	3	4
250	11 ± 3	3	4
260	9 ± 3	4	5
270	9 ± 4	3	5
280	10 ± 3	2	4
290	9 ± 4	3	5
300	9 ± 4	3	5
310	9 ± 4	2	4
320	9 ± 4	1	4

Table 6.4: Expected number of background events in the signal region $|m_{\gamma\gamma} - m_X| < 2\sigma_m(m_X)$, together with its statistical (db_{fit}) and systematic (db_{sys}) uncertainties, estimated from fits to the sideband data events as explained in Sec.6.6. The total uncertainty (db) is also reported. The expected background for the simulated m_X included in $R1$ or $R2$ is reported in the top and bottom panel, respectively.



(a) $R1 : 0 < m_{\gamma\gamma} < 110\,\mathrm{MeV/c^2}$ (b) $R2 : 160 < m_{\gamma\gamma} < 380\,\mathrm{MeV/c^2}$

Figure 6.8: $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ single event sensitivity for simulated X mass hypotheses included in $R1$ (left) or $R2$ (right); $m_X = 110$ MeV/c^2 ($m_X = 160$ MeV/c^2) is neglected here because the signal region $|m_{\gamma\gamma} - m_X| < 2\sigma_m(m_X)$ is not fully included in R1 (R2). Errors are also drawn but are not visible in this scale.

Single event sensitivity. The single event sensitivity is defined as:

$$BR_{SES}(m_X, N_K) = \frac{1}{A(m_X)N_K}. \tag{6.10}$$

Figure 6.8a and 6.8b reports the values obtained for $BR_{SES}(m_X, N_K)$ to the production of X with mass included in $R1$ and $R2$, respectively, by using N_K reported in Eq.6.2 and $A(m_X)$ from Table 6.3. As $BR_{SES}(m_X, N_K)$ does not depend on the background estimate, it is possible to evaluate it also for the X mass hypotheses for which the sidebands can not be defined. It is of $\mathcal{O}(10^{-8})$ for $20 \le m_X < 100\,\mathrm{MeV/c^2}$ and of $\mathcal{O}(5 \times 10^{-9})$ for $170 \le m_X \le 320\,\mathrm{MeV/c^2}$, with a loss of sensitivity at very low and high masses due to the loss of acceptance.

UL computation. As mentioned in Sec.6.3, the $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ search is here considered as a counting experiment with a single observation channel for each X mass hypothesis tested. For a fixed value of m_X, let s be the expected number of signal events in the signal region $|m_{\gamma\gamma} - m_X| < 2\sigma_m(m_X)$. The CL_s method [87] is here adopted to set an upper limit on s. This method has been widely used for the last twenty years in searches for new signal processes to set upper limits on signals near to the sensitivity limits of the experiments and is recommended by the Particle Data Group ([63], Statistics review).

The two models of interest for the signal search are: the signal+background model (or $s + b$ model), assuming for the total number of events in the signal region (N) the Poisson distribution with mean $s + b$, where b is the expected number of background events defined is Sec.6.6; the background only model (or b *only*), assuming $s = 0$ and N to be a Poisson distributed variable with mean b.

Let's assume, for the moment, that the background b is exactly known (that is, $\mathrm{d}b = 0$) and indicate the observed N value as N_{obs}. As both the $s+b$ and b *only* models are completely defined once b and s are fixed, according to the Neyman-Pearson lemma, the optimal choice of test statistic is the (log) likelihood ratio $\lambda(N)$ of the likelihoods of the $s+b$ and b *only* models. For fixed size $\alpha \equiv (1 - C.L.)$, with $C.L.$ confidence level, of the test to discriminate between the two hypotheses $s \neq 0$ and $s = 0$, $\lambda(N)$ maximizes the power of the test. The likelihood functions are:

$$L_{s+b} \equiv \frac{(s+b)^N e^{-(s+b)}}{N!} \tag{6.11}$$

and

$$L_b \equiv \frac{(b)^N e^{-b}}{N!} \tag{6.12}$$

for the $s+b$ and b *only* model, respectively, and their ratio is:

$$\Lambda(N) = \frac{L_{s+b}}{L_b} = e^{-s}\left(1 + \frac{s}{b}\right)^N \tag{6.13}$$

that becomes:

$$\lambda(N) = \log \Lambda(N) = -s + N \log\left(1 + \frac{s}{b}\right). \tag{6.14}$$

To use the CL_s method, two probabilities are then evaluated ($\lambda_{obs} \equiv \lambda(N_{obs})$):

$$CL_{s+b} \equiv P\left(\lambda(N) < \lambda_{obs} | s+b\right) \tag{6.15}$$

$$CL_b \equiv P\left(\lambda(N) < \lambda_{obs} | b\right) \tag{6.16}$$

CL_{s+b} represents the *p-value* for the $s+b$ model (p_{s+b}), while CL_b provides $(1 - p_b)$, where p_b is the *p-value* for the b *only* model. Finally:

$$CL_s \equiv \frac{CL_{s+b}}{CL_b} = \frac{p_{s+b}}{1 - p_b} \tag{6.17}$$

is evaluated and the condition:

$$CL_s \leq 1 - C.L. \tag{6.18}$$

is used to set upper limits of s at 90% C.L. For each tested s value, the probability density functions of the test statistic $\lambda(N)$ in the $s+b$ or b *only* hypothesis are needed to compute CL_{s+b} and CL_b in Eq.6.15 and 6.16, respectively, but are not known a priori. The computation is here performed by a Monte-Carlo integration technique involving multiple trials for each tested s. In each trial, the value of N (N_{toy}) is sampled according to the $s+b$ model and the value of $\lambda(N_{toy})$ is evaluated according to Eq.6.14, then it is compared to λ_{obs}; CL_{s+b} is the total count of the number of times

$\lambda(N_{toy}) \leq \lambda_{obs}$. CL_b is obtained in a similar way by sampling N_{toy} according to the $b\,only$ model.

To take into account the uncertainty on the knowledge of the background, the number of background events is considered as a nuisance parameter (b') and a Gaussian $p.d.f$ is assumed to model it, having mean equal to b and width $\sigma = db$. The likelihood functions 6.11 and 6.12 become then:

$$L_{s+b} = \int_0^\infty \frac{(s+b')^N e^{-(s+b')}}{N!} \cdot \frac{e^{-(b-b')^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}} db' \tag{6.19}$$

and

$$L_b = \int_0^\infty \frac{(b')^N e^{-b'}}{N!} \cdot \frac{e^{-(b-b')^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}} db', \tag{6.20}$$

where the convolution with the Gaussian distribution is introduced to average over all the possible values of the background.

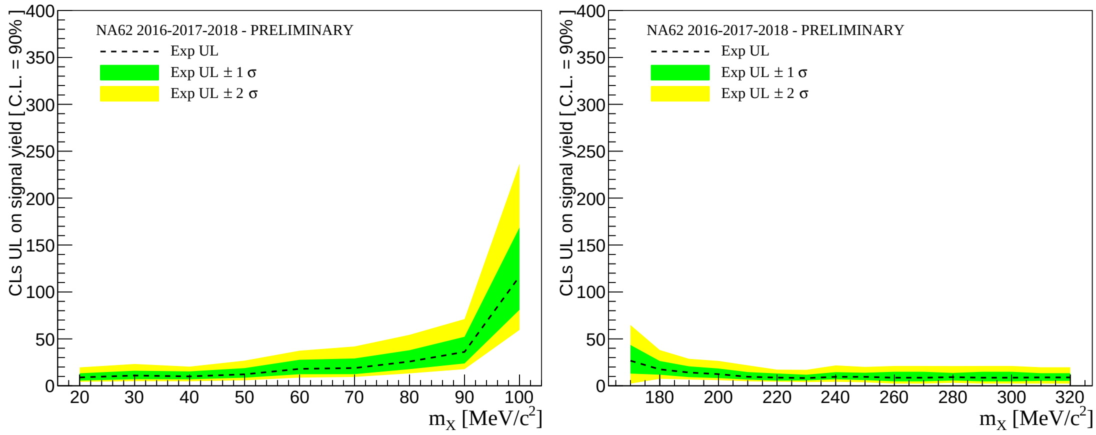
In the CL_s computation, Eq.6.19 and 6.20 are taken into account introducing a gaussian smearing of the number of expected background events for fixed s, b and db. For each trial: b' is sampled according to $f(b'|b,\sigma) = \frac{e^{-(b-b')^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}}$; to evaluate the probabilities CL_{s+b} and CL_b, N_{toy} is sampled according to a Poisson distribution with mean $s + b'$ and b' instead of $s + b$ and b, respectively.

$\mathbf{CL_s}$ Expected UL By definition, the expected UL is the maximum signal yield that could be observed due to the fluctuations of the background if, in reality, the signal is not present in the sample. As an observed limit will differ from the expected limit even if the signal is absent, the fluctuation of the expected UL must also be quantified in order to quote how large these differences are.

In this analysis, for each available m_X included in one of the regions sensitive to the signal, the CL_s expected UL at 90% CL is computed through 1000 fake experiment outcomes N, sampled according to the only background model with expected number of background events $b \pm db$ fixed. For each fake N, the upper limit on s is computed as explained above and used to build a distribution. The mean value of the distribution provides the expected UL. The intervals corresponding to $Exp.UL \pm 1\sigma$ and $Exp.UL \pm 2\sigma$ are quoted considering the quantiles that give the same probability content for a Gaussian distribution ($[0.1587; 0.8413]$ for $\pm 1\sigma$ and $[0.02275; 0.97725]$ for $\pm 2\sigma$). The obtained limits are reported in Figure 6.9.

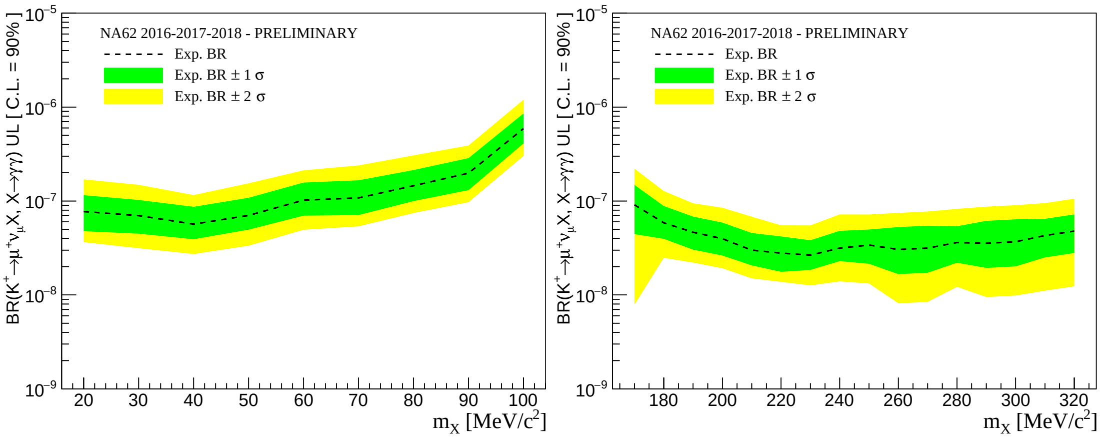
6.8 Branching ratio sensitivity

Figure 6.10 reports the $BR_{sensitivity}(m_X, N_K)$ defined in Eq.6.9 for X masses in $R1$ and $R2$ on the left and right hand side, respectively. It ranges from $\mathcal{O}(8 \times 10^{-8})$ up to $\mathcal{O}(6 \times 10^{-7})$ in $R1$ and is of $\mathcal{O}(5.5 \times 10^{-8})$, with the lower value of $\mathcal{O}(3 \times 10^{-8})$, in $R2$. The bands included in the plots are the expected fluctuations at $\pm 1\sigma$ and $\pm 2\sigma$. The uncertainties due to the knowledge of N_K and $A(m_X)$ are completely negligible with respect to the fluctuation of the $Exp.UL$ and are therefore not included here.



(a) (b)

Figure 6.9: CL_s expected upper limit at 90% C.L. on the signal yield for the simulated X mass hypotheses included in $0 < m_{\gamma\gamma} < 110\,\mathrm{MeV/c^2}$ (left) or $160 < m_{\gamma\gamma} < 380\,\mathrm{MeV/c^2}$ (right), evaluated as explained in Sec.6.7. The lack of statistics in the $m_{\gamma\gamma}$ distribution of the NA62 Run1 data below 20 MeV/c^2 (Fig.6.2a) and above 320 MeV/c^2 (Fig.6.2b) prevents from a data-driven background estimate, therefore a different method is needed to account for this mass values, that are not included here.



(a) (b)

Figure 6.10: NA62 expected branching ratio sensitivity to the $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ process from the analysis of the Run1 data, for the simulated X mass hypotheses included in $0 < m_{\gamma\gamma} < 110\,\mathrm{MeV/c^2}$ (left) or $160 < m_{\gamma\gamma} < 380\,\mathrm{MeV/c^2}$ (right), based on the single event sensitivity in Figure 6.8 and on the $Exp.UL$ at 90% C.L. in Figure 6.9.

Chapter 7

Results

In this Chapter the preliminary results of the mass scan for the $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ search in the NA62 data collected during Run1 (2016-2017-2018) are reported.

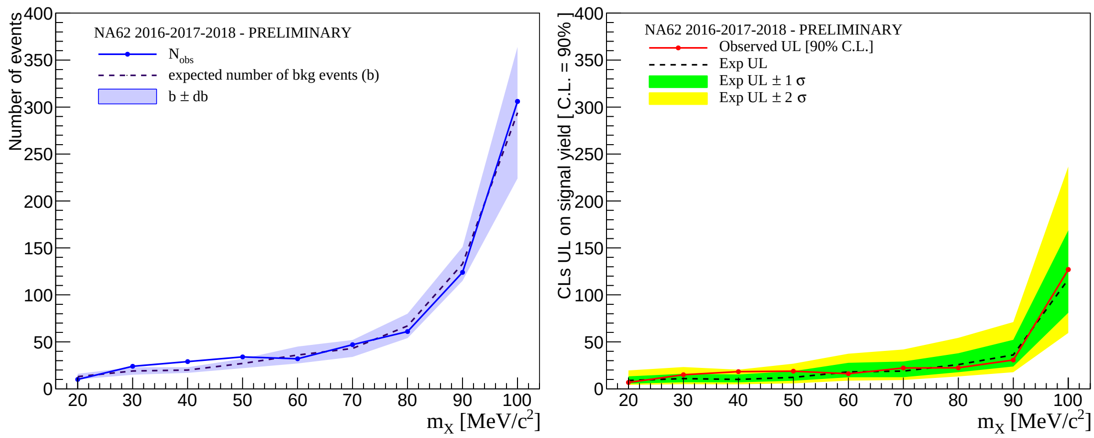
Assuming the prompt X decay hypothesis, the search has been performed in the mass ranges $10 - 110\,\mathrm{MeV/c^2}$ and $160 - 380\,\mathrm{MeV/c^2}$. Due to the low statistics in the data after the signal selection (Figures 6.2a and 6.2b), below $20\,\mathrm{MeV/c^2}$ and above $320\,\mathrm{MeV/c^2}$ the method adopted for the data-driven estimate of the expected number of the background events (described in Sec.6.6) is not suitable. Thus, although they are kinematically allowed, X masses included in these intervals are not considered in the scan. However, even if a different procedure for the background estimate could allow to account also for them, the experimental sensitivity for these mass values are expected to be at least one order of magnitude worst than for the tested masses, as evident from the single event sensitivities in Figure 6.8.

Figures 7.1a and 7.2a report the number of data events observed in the signal region $|m_X - m_{\gamma\gamma}| < 2\sigma_m(m_X)$ for each tested X mass hypothesis included in $R1$ or $R2$, respectively. The expected number of background events in each signal region is also reported (dashed lines), together with its total uncertainty (light-blue band). The uncertainty accounts for the statistical error from the fit to the sideband events to the signal region and for the systematic uncertainty, evaluated by changing the fitting function and the width of the sidebands. Figures 7.1b and 7.2b show the upper limit at the 90% C.L. of the number of signal events ($UL(m_X)$) estimated with the CL_s method, as explained in Sec. 6.7, compared to the upper limit expected in the background only hypothesis. No significant excess is observed. Also, the lower background in the mass interval above 160 MeV/c^2 with respect to the one below 110 MeV/c^2 allows to set stringent upper limits in $R2$ than in $R1$, but they are more sensitive to statistical fluctuations. The same holds for the model defined to describe the background.

Figures 7.3a and 7.3b report the observed upper limits of the $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ branching ratio in the hypothesis of prompt $X \to \gamma\gamma$ decay, compared to the expected sensitivity from Figure 6.10. $BR(K^+ \to \mu^+\nu_\mu X, X \to \gamma\gamma) < UL(m_X) \cdot BR_{SES}(m_X, N_K)$, where $UL(m_X)$ is the observed upper limit on the signal yield in Figure 7.1b or 7.2b, depending on the X mass, and $BR_{SES}(m_X, N_K)$ is the single event sensitivity in Figure 6.8. The obtained values range from 5.9×10^{-8} up to 6.4×10^{-7} in $R1$ and from 1.6×10^{-7} down to 1.2×10^{-8} in $R2$. The recast of these results in the parameter space of Figure 1.2 is beyond the aim of this work. However, it can be noted that, assuming $BR(X \to \gamma\gamma) = 100\%$, the obtained values are only one order of magnitude far from the benchmark branching ratio of $\mathcal{O}(10^{-9})$ probing the existence of a new scalar light DM mediator accounting for the observed value of muon anomalous magnetic moment and for the DM thermal freeze out.

Major improvements to the search are expected including the NA62 Run2 dataset. As the search is statistics dominated, by considering the K^+ collected from 2021 until today, an increase of at least a factor 2 is expected for the sensitivity to $K^+ \to \mu^+\nu_\mu X$ with prompt $X \to \gamma\gamma$ process. New samples will also become available with the data taking periods foreseen until the third long shutdown of the CERN accelerator complex. However, the Run2 sample profits from a minimum bias trigger line with different conditions with respect to Run1, from an higher beam intensity and from the information provided from the new detectors installed (see Chapter 2), therefore a dedicated study must be performed to use both Run1 and Run2 collected data.

Moreover further investigation are required for masses below $80\,\mathrm{MeV/c^2}$. The observed discrepancy in the data $m_{\gamma\gamma}$ spectrum are probably due to a background from the $\pi^+ \to \mu^+\nu_\mu$ decay-in-flight that is not included or well reproduced in the simulated samples considered. A deeper understanding of the contributions to the low $m_{\gamma\gamma}$ tail could lead to a further reduction of the background in this region, further increasing the experimental sensitivity to the signal.

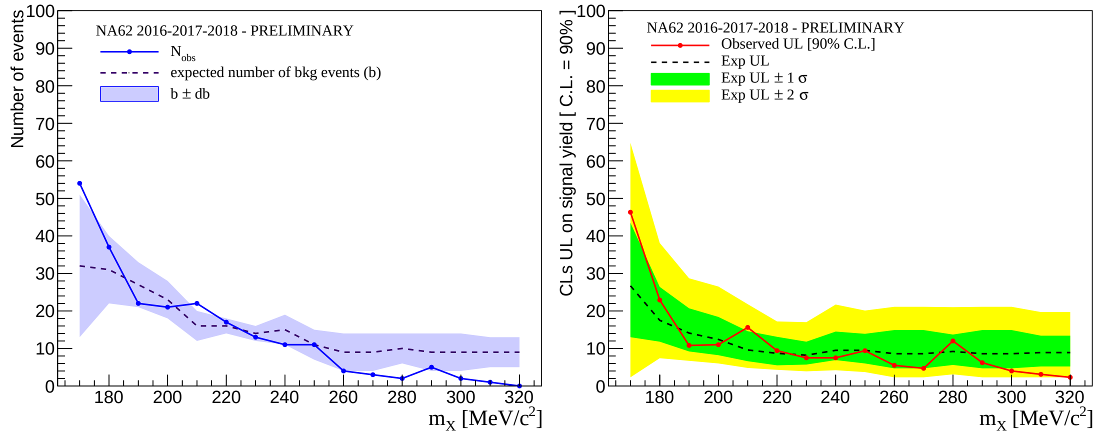


(a)

(b)

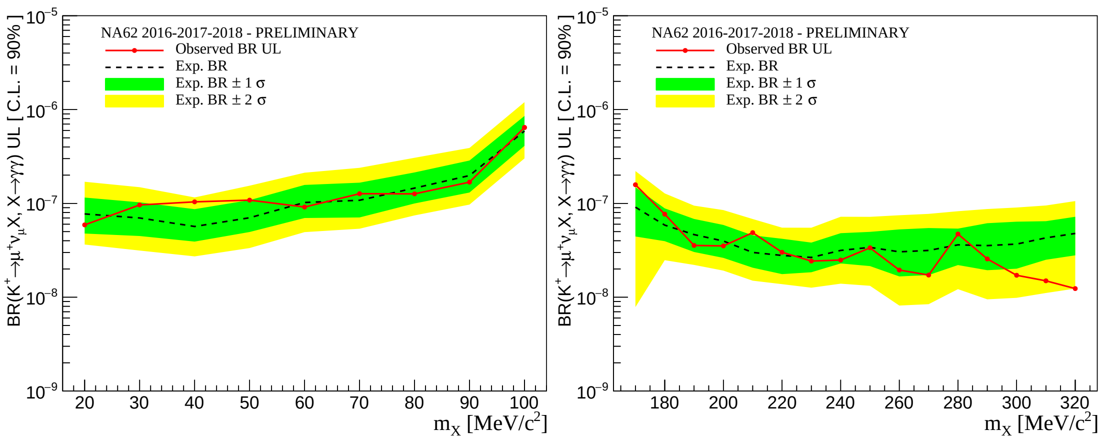
Figure 7.1: Left: number of observed events in data (N_{obs}) and expected number of events in the background only hypothesis (with its uncertainty) for each tested X mass value in R1 ($m_X <$ 110 MeV/c^2). The uncertainty is $db = \sqrt{db_{fit}^2 + db_{sys}^2}$, with db_{fit} and db_{sys} defined in Sec.6.6. Right: upper limits at 90% C.L. of the number of signal events, with their expected $\pm 1\sigma$ and $\pm 2\sigma$ bands in the background only hypothesis, based on N_{obs} and $b \pm db$ on the left-hand side. The upper limits are evaluated as explained in Sec.6.7

.



(a) (b)

Figure 7.2: Left: number of observed events in data (N_{obs}) and expected number of events in the background only hypothesis (with its uncertainty) for each tested X mass value in R2 ($m_X >$ 160 MeV/c^2). The uncertainty is $db = \sqrt{db_{fit}^2 + db_{sys}^2}$, with db_{fit} and db_{sys} defined in Sec.6.6. Right: upper limits at 90% C.L. of the number of signal events, with their expected $\pm 1\sigma$ and $\pm 2\sigma$ bands in the background only hypothesis, based on N_{obs} and $b \pm db$ on the left-hand side. The upper limits are evaluated as explained in Sec.6.7. To be noted that the vertical scale is a factor four smaller with respect to Figure 7.1.



(a) (b)

Figure 7.3: Upper limits at 90% C.L. of $BR(K^+ \to \mu^+\nu_\mu X,\, X \to \gamma\gamma)$ in the prompt $X \to \gamma\gamma$ decay assumption, with their expectation in the background only hypothesis, for X mass hypothesis < 110 MeV/c^2 (left) or > 160 MeV/c^2.

Conclusions

This thesis work focuses on the first experimental search for the $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ process, with X produced radiatively from the muon, in the decay of the K^+ from the NA62 beam. X is assumed to be a new scalar particle, neutral with respect to the SM and coupling preferentially to muons. The existence of such a particle provides a possible and one of the last valuable solution to solve the longstanding discrepancy between the experimental value and the SM prediction of the muon $(g-2)_\mu$ with new physics below the electroweak scale. Moreover, X could represent the mediator of the interaction responsible for the DM thermal freeze out, assuming the DM has a mass below the electroweak scale and exhibits feeble couplings to the SM.

For this search, the whole data sample collected by NA62 through the minimum bias trigger line during Run1 (2016-2017-2018) is considered. The $X \to \gamma\gamma$ prompt decay hypothesis is assumed, that ensures to be able to reconstruct the X decay in NA62 detector. The di-photon invariant mass, $m_{\gamma\gamma} \equiv \sqrt{(P_{\gamma_1} + P_{\gamma_2})^2}$, with P_{γ_i} reconstructed four-momentum of the photon i and $i = 1, 2$, is chosen as the main discriminating variable of the study and a peak-search procedure is adopted to get a model-independent measurement of $BR(K^+ \to \mu^+\nu_\mu X,\, X \to \gamma\gamma)$. Moreover, since to compute the branching ratio the flux of K^+ decays is needed but is not known a priori, a normalization procedure is defined, measuring the $BR(K^+ \to \mu^+\nu_\mu X,\, X \to \gamma\gamma)$ with respect to a normalization channel, a known SM decay with the same topology as the signal. In this approach, a first order cancellation of systematic effects common between the signal and the normalization channel is expected. Thus, the effective number of K^+ decays in the analyzed data, $N_K = (1.3368 \pm 0.0039) \times 10^{10}$, is estimated selecting a clean sample of $K^+ \to \mu^+\nu_\mu\pi^0$ events ($BR(K^+ \to \mu^+\nu_\mu\pi^0) = (3.352 \pm 0.033)\%$).

Two regions sensitive to the signal are identified in the $m_{\gamma\gamma}$ distribution, surrounding the $\pi^0 \to \gamma\gamma$ resonance: $m_{\gamma\gamma} < 110\,\mathrm{MeV/c^2}$ ($R1$) and $160 < m_{\gamma\gamma} < 380\,\mathrm{MeV/c^2}$ ($R2$), where the upper bound of $R2$ is due to the kinematic constraint $m_X < m_K - m_\mu$. A cut-based $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ signal selection is defined, with specific conditions to reduce the different backgrounds populating $R1$ and $R2$ studied by using Monte Carlo simulations. The dominant backgrounds are due to:

1. mis-reconstructed $K^+ \to \mu^+\nu_\mu\pi^0$ events and $K^+ \to \pi^+\pi^0$ events with $\pi^+ \to \mu^+\nu_\mu$ decay-in-flight;

2. the $K^+ \to \pi^+\pi^0\pi^0$ decays with $\pi^+ \to \mu^+\nu_\mu$ decay when some of the photons

from the decay of the neutral pions are not detected or when the selected photons come from a different π^0;

3. the conversion into e^+e^- pairs of one or both the photons from the decay of a π^0 in one of the K^+ decays mentioned above.

After the selection, an overall good agreement is observed between the $m_{\gamma\gamma}$ distributions in data and the sum of the MC samples describing the residual SM contributions, normalized to data according to the measured N_K. However, in $R1$, a discrepancy is present for $m_{\gamma\gamma} < 80\,\mathrm{MeV/c^2}$. This discrepancy is probably due to some background source not included in the simulations, and further investigations are necessary to asses its nature. Since the $m_{\gamma\gamma}$ data distribution is not significantly different from the SM expectation (excluding the low mass region), a mass scan is performed and upper limits of $BR(K^+ \to \mu^+\nu_\mu X,\, X \to \gamma\gamma)$ are set. Signal samples are simulated from 10 up to 380 Mev/c^2 in step of 10 Mev/c^2; for the time being, only the X mass hypotheses included in $R1$ or $R2$ are included in the scan. Also, exploiting the peaking nature of the signal, signal regions are defined as $|m_{\gamma\gamma} - m_X| < 2\sigma_m(m_X)$ for each tested X mass, with $\sigma_m(m_X)$ signal mass resolution. The resolution is estimated from simulation and increases from 0.2 MeV/c^2 at $m_X = 20\,\mathrm{MeV/c^2}$ to 4.2 MeV/c^2 at $m_X = 360\,\mathrm{MeV/c^2}$.

For each X mass value, the signal search is considered as a counting experiment with a single observation channel. The expected number of background events in each signal region is determined by local polynomial fits to the sideband data events surrounding each signal region in the data $m_{\gamma\gamma}$ spectrum. This data-driven method is only valid provided that there are no peaking background structures in the regions sensitive to the signal, and has the advantage to be also suitable below $80\,\mathrm{MeV/c^2}$, regardless of the data/MC agreement. Upper limits (UL) at 90% C.L. are then established for the number of signal events by using the CL_s method and exploited to compute branching ratio upper limits as $BR(K^+ \to \mu^+\nu_\mu X,\, X \to \gamma\gamma) < UL/N_K A(m_X)$, with $A(m_X)$ acceptance of the signal selection, for X masses from 20 up to 100 MeV/c^2 and from 170 up to 320 MeV/c^2. The signal acceptance goes from $\mathcal{O}(1\%)$ up to $\mathcal{O}(1.6\%)$ for masses in $R1$ and from $\mathcal{O}(2.4\%)$ down to $\mathcal{O}(1.5\%)$ in $R2$. Due to the low statistics in the data after the signal selection, below $20\,\mathrm{MeV/c^2}$ and above $320\,\mathrm{MeV/c^2}$ the method adopted for the background estimate is not suitable.

The preliminary NA62 upper limits of $BR(K^+ \to \mu^+\nu_\mu X,\, X \to \gamma\gamma)$, for prompt X decay and $BR(X \to \gamma\gamma) = 100\%$ range from 5.9×10^{-8} at 20 MeV/c^2 up to 6.4×10^{-7} at 100 MeV/c^2 and from 1.6×10^{-7} at 170 MeV/c^2 down to 1.2×10^{-8} at 320 MeV/c^2, only one order of magnitude higher with respect to the benchmark branching ratio of $\mathcal{O}(10^{-9})$ covering the interesting parameter space of X scalar muon-phylic DM mediator. Improvements are expected both from the addition of the data collected during NA62 Run2 and from a further reduction of background in R1, provided a better understanding of the discrepancy below 80 MeV/c^2 is achieved.

Appendix A

NA62 overlay technique for pile-up simulation

Let a single beam particle decay, along with its secondary interactions, constitute an event. By definition, pile-up refers to additional events that occur in a small time window around an interesting physics event. Given the high rate of the NA62 environment, the simulation of pile-up events becomes crucial to get a proper data description and is strongly dependent upon the beam intensity. The higher the beam intensity is, the higher the amount of interesting physics events is, as well as the number of pile-up events surrounding it. Then, the developed overlay technique relies on intensity templates obtained from the measured beam intensity in data.

For the collected data, the instantaneous intensity of the NA62 beam is estimated by counting the number of hits in each GTK station during a given out-of-time window. The out-of-time window is defined as $(-25, -2.5)$ ns and $(2.5, 25)$ ns around the trigger time associated to a beam particle decay, and is $\Delta t = 45\,\mathrm{ns}$ large. The central part of the window, $(-2.5, 2.5)\,\mathrm{ns}$, is excluded to avoid hits from the physics event, that would bias the estimate. The number of hits detected in the out-of-time window by a given GTK station (n_i) is converted to an intensity value (I_i):

$$I_i = \frac{n_i}{\Delta t} \tag{A.1}$$

This value is then averaged across all the GTK stations to get a measure of the mean intensity of the beam (I_{mean}). However, due to detector and DAQ inefficiencies, I_{mean} is not the real one. Hence, in order to get an unbiased intensity estimate to be used for the overlay, an unfolding procedure [88] is used, taking into account the simulated detector response (resolutions, acceptances and inefficiencies). In this way a representative intensity template is obtained from data, which is then sampled for every generated main event by the overlay procedure, to randomly select a mean intensity value to be simulated ($I_{sampled}$). It has to be noted that this procedure strongly relies on the simulation of the GTK.

As the intensity distribution and detector layout might change for different data-taking years, multiple templates are required, obtained from specific good quality reference runs. For the $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ search, NA62 Run1 data are included and the reference intensity template used for the whole period is reported in Figure A.1, together with the measured intensity distribution. For the simulation of the Run2 conditions other samples are available.

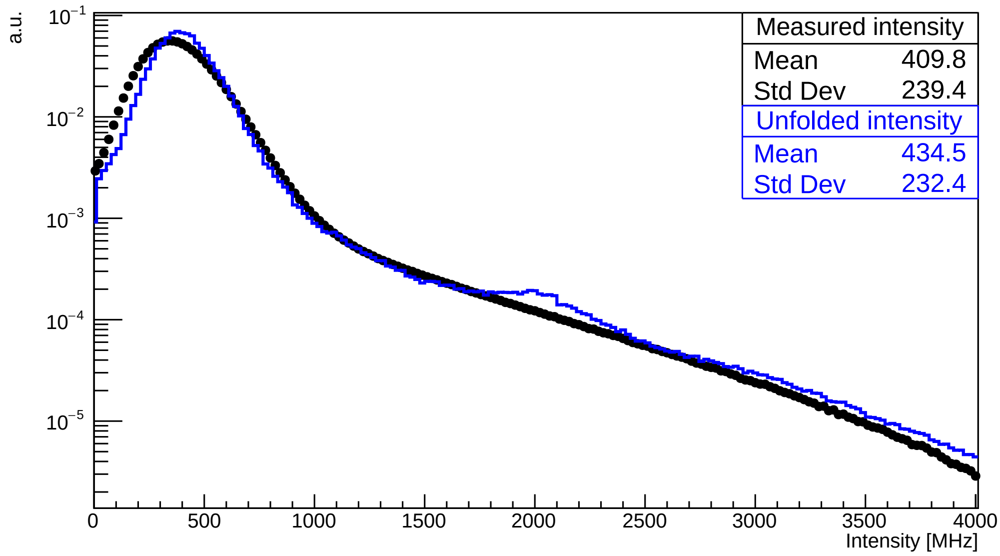


Figure A.1: Measured intensity distribution (black dot) of the reference run for the NA62 Run1 data taking conditions. The corresponding unfolded intensity template, to be used for the sampling of the intensity to be simulated, is also reported (blue). Both distributions are normalized to unitary area and have a bin width of 22.2 MHz. The bin width reflects the granularity of the intensity estimator in Eq.A.1, and depends on the width of the out-of-time window used to count GTK hits. Below 1500 MHz, the measured and the unfolded distributions are in good agreement, compatible with a GTK inefficiency of few percent. The overlap of hits from different beam particles in the same GTK pixel produces a drop of the GTK efficiency around 10%, which causes the measured intensity to be underestimated with respect to the unfolded intensity.

For a sampled intensity value, the number of pile-up events to overlay to the main one is determined as follows. By assuming that a beam particle will leave a single hit in each GTK station, the total number of beam particles (N_{beam}) may be written as:

$$N_{beam} = \sum_{j=K^+,\pi^+,p} N_{beam,j} = \sum_{j=K^+,\pi^+,p} I_{sampled} \times \Delta t \times w_i. \tag{A.2}$$

with w_i weight taking into account the fraction of the species $i = K^+, \pi^+, p$ present in the beam at the entrance of the NA62 detector. In particular, weights are evaluated to match the nominal beam composition at GTK3 considering the losses due to K/π decays between KTAG and GTK3:

$$w_K = 0.064, w_p = 0.705, w_p = 0.24. \tag{A.3}$$

The number of events that shall be overlaid to the main event is then determined by applying Poisson statistics, with a distribution having N_{beam} as mean value. Overlaid pions and kaons are allowed to decay according to their mean life and branching ratio to different modes, and are saved whatever the z of the decay is.

Overlaid event times are generated uniformly in a time window of $\pm\,200\,\mathrm{ns}$ around the time of the main event, such that an approximately uniform intensity across the given time window is observed.

Appendix B

NA62 technique for forced π^+ decay simulation

The probability for a pion produced from a beam K^+ to decay inside the NA62 detector is strongly reduced due to its mean life time ($\tau_\pi = 2.6 \times 10^8$ s) and to the typical momentum of $\mathcal{O}(20\,\mathrm{GeV/c})$, As a consequence, a huge amount of generated K^+ decays is needed to get accurate estimates of the backgrounds to the $K^+ \to \mu^+\nu_\mu X$, $X \to \gamma\gamma$ search due to the $\pi^+ \to \mu^+\nu_\mu$ in K^+ decays with charged pions. The forced π^+ decay simulation mode is then used instead of the standard one. This mode restricts to the scenario with a single charged pion the multiple pion decay simulation developed for the NA62 search for the $K^+ \to \pi^-\mu^+\mu^+$ lepton number violating decay [89].

The generation of a sample proceeds as follows. After the kaon is tracked by GEANT4 [79] and decays at Z_{vtx}, a random variable (t) is sampled according to the distribution:

$$f(t) = \frac{1}{\tau_\pi} \cdot \exp(-t/\tau_\pi) \tag{B.1}$$

where τ_π is the pion mean lifetime. The sampled variable is accepted as "new" pion proper lifetime if the condition:

$$t < t_{max} = 1.1 \times (Z_{max} - Z_{vtx})/\gamma_\pi c\beta \tag{B.2}$$

is satisfied. Here Z_{max} is the coordinate of the face of the iron wall located in front of the MUV3 and $Z_{max} - Z_{vtx}$ defines the range covered by the NA62 detector, γ_π is the Lorentz factor of the pion and $c\beta$ is the speed of the pion in the NA62 reference frame. The constant 1.1 is a safety factor accounting for the fact that the pion trajectory is not a straight line parallel to the z axis. Once a sampled proper lifetime is assigned to the daughter pion, the simulation of the event proceeds.

This procedure modifies the space of the pion life time, increasing the rate of pion decays in the detector volume and biasing the acceptance for the K^+ decay mode simulated. Therefore, each event generated with a forced pion decay is weighted with:

$$w = 1 - \exp(-t_{max}/\tau_\pi). \tag{B.3}$$

to take into account the real probability of the π^+ to decay before the MUV3. Equation B.3 is an exact expression not involving any approximations. The mean value of the w distribution quantifies the speed-up factor achieved. Figure B.1 shows the distribution of the weights for the capped samples used for this analysis. As an example, for the $K^+ \to \pi^+\pi^0$ sample, the mean value is ~ 0.05, and for the $K^+ \to \pi^+\pi^0\pi^0$ is ~ 0.08. Given the dimension of the generated samples of 400×10^6 and 50×10^6, respectively, the simulations have the same statistical power of 8×10^9 and 625×10^6 large standard samples. The mean value of this distributions is lower for the ones reported for the same process simulated in the upstream region, reported in the right column of the same Figure. This is expected for the lower Z_{vtx} values allowed, corresponding to a higher decay probability for the π^+.

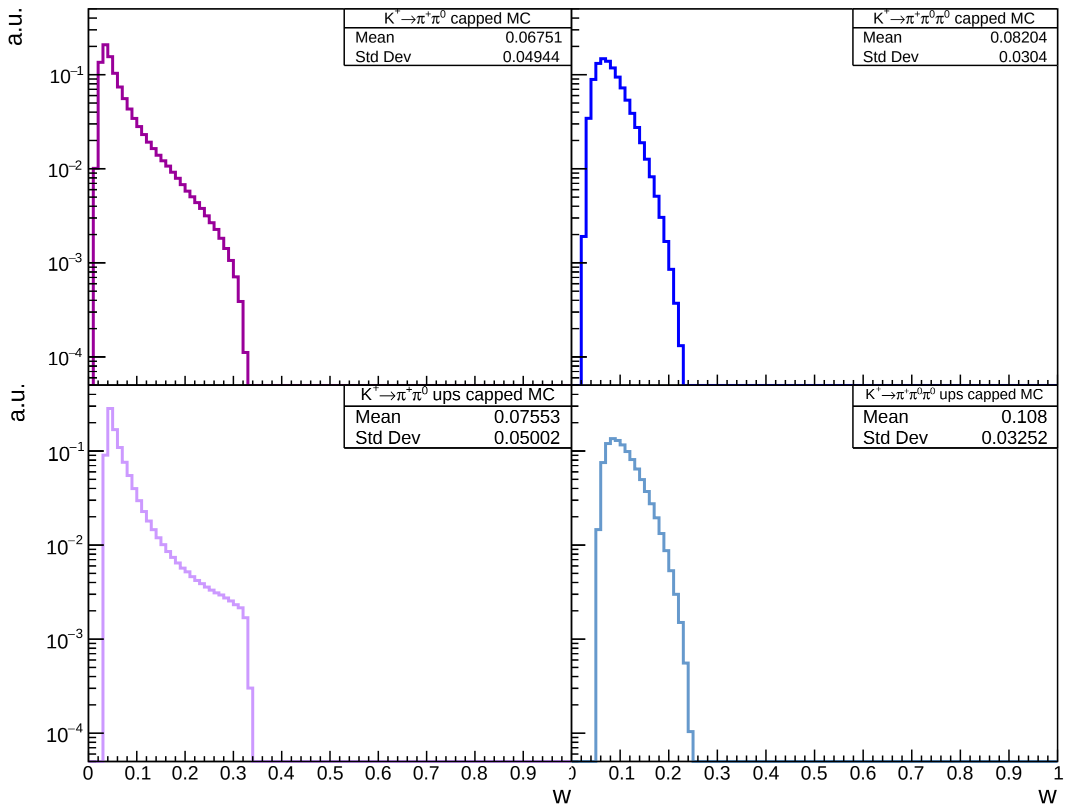


Figure B.1: Distribution of weights evaluated according to Eq.B.3 for Monte Carlo simulations of the $K^+ \to \pi^+\pi^0$ (top) and $K^+ \to \pi^+\pi^0\pi^0$ (bottom) modes with forced π^+ decay, with the K^+ decay generated in standard region (left) on in the upstream region (right), defined in Sec.3.3.2

Appendix C

Neutral vertex computation

As the position assumed as K^+ decay vertex determines the reconstructed direction of flight of the photons $\gamma's$, the presence of a final state π^0 can be exploited to define a decay vertex independent with respect to the selected downstream track. As the neutral vertex is useful to identify $K^+ \to \pi^+\pi^0$ events with $\pi^+ \to \mu^+\nu_\mu$ decay-in-flight, this decay mode is considered in the following.

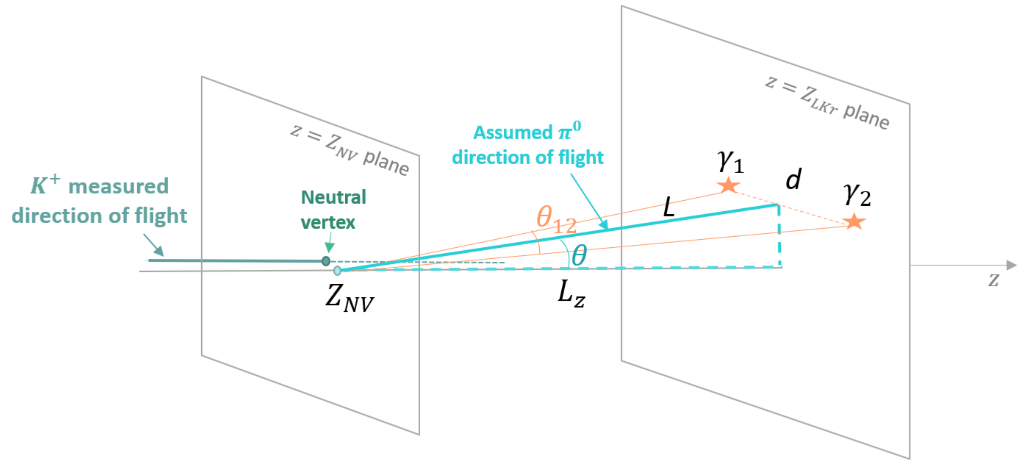


Figure C.1: Simplified schematic view of the geometry considered for the neutral decay vertex computation.

Neutral decay vertex computation. The $K^+ - \pi^0$ vertex, or neutral decay vertex, is computed by the back extrapolation of the trajectory of the π^0 from the LKr front plane and by the intersection with the selected K^+ track.

The z coordinate of the vertex (Z_{NV}) is determined first. In Figure C.1 a simplified schematic view of the geometry is shown:

- Z_{LKr} is the coordinate of the LKr front plane, where the LKr clusters are reconstructed; their measured position in the plane are (x_i, y_i), $i = 1, 2$.
- $d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$ is the distance between the two clusters position at the LKr front plane.
- θ_{12} is the angle between the two clusters and θ is the polar angle between the z axis and the segment joining the midpoint of d and the vertex.

Let L be the distance between the vertex and the midpoint of d. The projection of L along the z axis is $L_Z = Z_{LKr} - Z_{NV}$, and $L_Z = L\cos\theta \sim L$ in the small angle approximation. By requiring the four-momentum conservation in the $\pi^0 \to \gamma\gamma$ decay, assuming P_{π^0} and $P_{\gamma i}$ with $i = 1, 2$ the four-momenta of the particles in the corresponding subscript:

$$m_{\pi^0}^2 = P_{\pi^0}^2 = (P_{\gamma 1} + P_{\gamma 2})^2 = 2E_{\gamma 1}E_{\gamma 2}(1 - cos\theta_{12}). \tag{C.1}$$

As $(1 - cos\theta_{12}) = 2\sin^2(\theta_{12}/2) = 2(d/2L)^2 = d^2/2L^2$, Eq.C.1 becomes:

$$m_{\pi^0}^2 = 2E_{\gamma 1}E_{\gamma 2}\frac{d^2}{2L^2} = E_{\gamma 1}E_{\gamma 2}\frac{d^2}{L^2}, \tag{C.2}$$

and L_Z reads:

$$L_Z \sim L = \sqrt{E_{\gamma 1}E_{\gamma 2}\frac{d^2}{m_{\pi^0}^2}}, \tag{C.3}$$

so that $Z_{NV} = Z_{LKr} - L_Z$ can be calculated.

The neutral vertex position in the (x, y) plane, (X_{NV}, Y_{NV}), is computed as the intersection between the reconstructed K^+ track and the plane $z = Z_{NV}$. Since the K^+ momentum, after the GTK, is not parallel to the Z axis, the measured direction of flight is used to extrapolate forward from the measured K^+ position at GTK3 to the plane $z = Z_{NV}$. As an example, Figure C.2 shows the comparison between the distributions of z coordinate of the charged and neutral reconstructed vertices with respect to the true vertex position (ΔZ) for $K^+ \to \pi^+\pi^0$ simulated events. Both the reconstructed z are compatible with the true coordinate of the K^+ vertex although the resolution of the neutral vertex is wider with respect to the resolution of the charged one, that depends only on the GTK and STRAW resolutions for the momentum measurement. The tail at large positive ΔZ arise when $L_Z \sim L$ underestimates L_z, shifting Z_{NV} towards higher z with respect to the true vertex position.

Figure C.3 shows the comparison between the reconstructed charged and neutral vertices for simulated $K^+ \to \pi^+\pi^0$ events with and without the $\pi^+ \to \mu^+\nu_\mu$ decay-in-flight. As for Figure C.2, the difference with respect to the true K^+ vertex is considered. Due to the kick of the TRIM magnet, the K^+ direction of flight is deflected in the xz plane. Due to this kick, when the track of the μ^+ from the $\pi^+ \to \mu^+\nu_\mu$ decay is measured and matched with the K^+ track, a widening of the charged vertex distribution is observed along the x and z directions (Figure C.3a and C.3c left), while the neutral decay vertex remains unchanged, as expected (Figure C.3a and C.3c right).

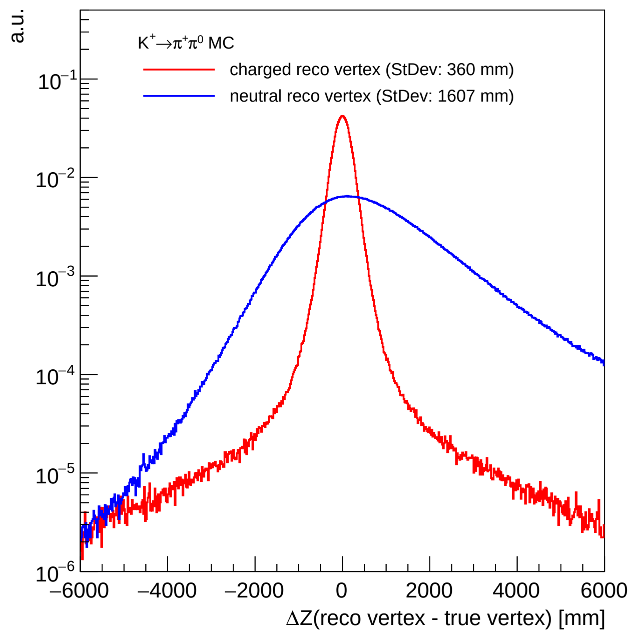
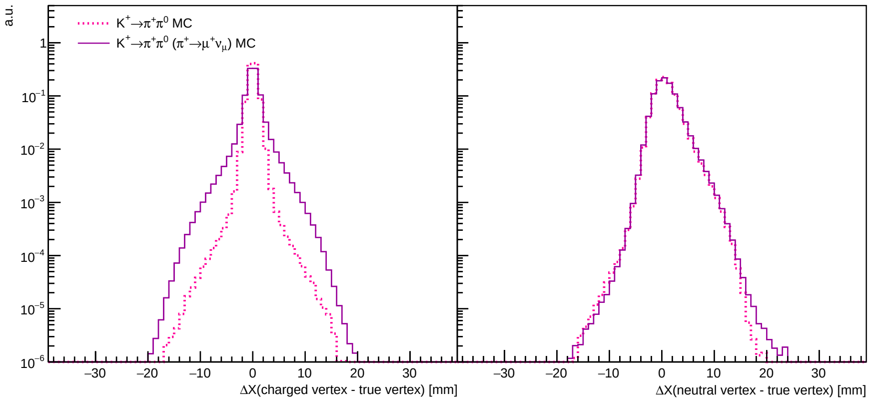


Figure C.2: Difference between the two considered reconstructed vertices (charged in red and neutral in blue) with respect to the true vertex position along the z axis for simulated $K^+ \to \pi^+\pi^0$ events.

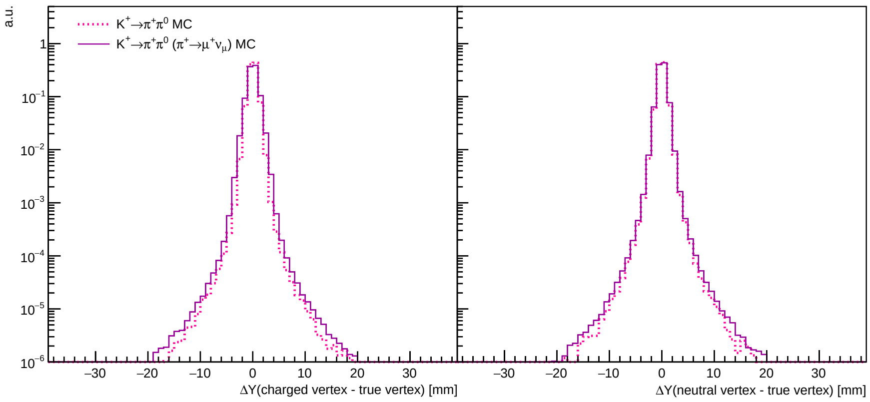
$\gamma's$ three-momentum assuming the neutral vertex. The neutral decay vertex is then used to determine the four-momentum of each of the two gamma. Let $(\hat{x}, \hat{y}, \hat{z})$ be the versors of the NA62 reference frame axes, $\vec{r_i} = x_i\hat{x} + y_i\hat{y} + z_i\hat{z}$ vector of the cluster i position of the cluster and $\vec{r}_{NV} = X_{NV}\hat{x} + Y_{NV}\hat{y} + Z_{NV}\hat{z}$ of vector of the neutral vertex position in the reference frame, respectively. The three-momentum $\vec{p}_{\gamma i\, neut}$ is defined as:

$$\vec{p}_{\gamma i\, neut} = \frac{E_{\gamma i}}{c}\left(\frac{\vec{r_i} - \vec{r}_{NV}}{|\vec{r_i} - \overrightarrow{r_{NV}}|}\right) = \frac{E_{\gamma i}}{c}\left(\frac{\vec{x_i} - \vec{X}_{NV}}{L}\hat{x} + \frac{\vec{y_i} - \vec{Y}_{NV}}{L}\hat{y} + \frac{\vec{z_i} - \vec{Z}_{NV}}{L}\hat{z}\right) \tag{C.4}$$

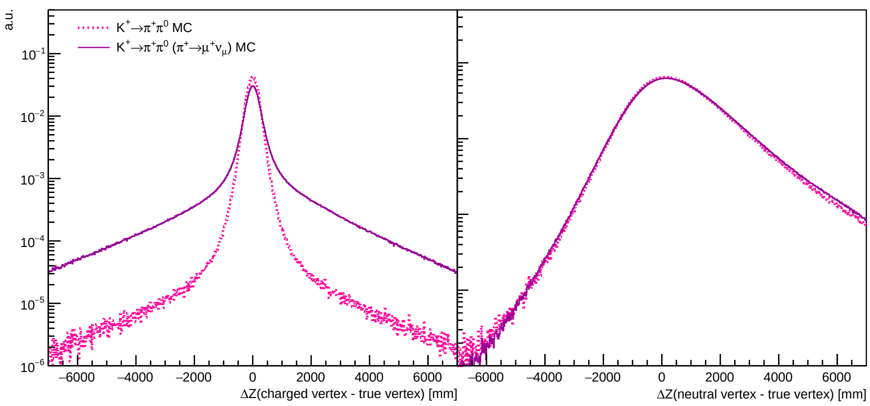
where $\frac{\vec{x_i} - \vec{X}_{NV}}{L}, \frac{\vec{y_i} - \vec{Y}_{NV}}{L} << 1$ and $\frac{\vec{z_i} - \vec{Z}_{NV}}{L} \sim \mathcal{O}(1)$ for K^+ in the fiducial volume.



(a)



(b)



(c)

Figure C.3: Difference between the charged (left) and the neutral (right) vertex with respect to the true vertex position along the x (top), y (center) and z (bottom) axis for simulated $K^+ \to \pi^+\pi^0$ events without (pink) and with (purple) the $\pi^+ \to \mu^+\nu_\mu$ decay-in-flight. Each distribution is normalized to unitary integral.

Bibliography

[1] B.Abi *et al.* (The Muon $g-2$ Collaboration), *Measurement of the Positive Muon Anomalous Magnetic Moment to 0.46ppm.* Phys. Rev. Lett. **126**, 141801

[2] D.P.Aguillard *et al.* (The Muon $g-2$ Collaboration), *Measurement of the Positive Muon Anomalous Magnetic Moment to 0.20ppm.* Phys. Rev. Lett. **131**, 161802

[3] E.Cortina Gil *et al.* (The NA62 Collaboration), *Search for K^+ decays to a muon and invisible particles.* Phys. Lett. B **816** (2021) 136259

[4] G. Bertone and D. Hooper *History of dark matter.* Rev. Mod. Phys. **90** (2018) 045002

[5] A. Arbey, F. Mahmoudi, *Dark matter and the early Universe: A review.* Progress in Particle and Nuclear Physics **119** (2021) 103865

[6] Y. Gouttenoire, (2022) *Thermal Dark Matter.* In: Beyond the Standard Model Cocktail. Springer Theses. Springer, Cham. https://doi.org/10.1007/978-3-031-11862-3 4

[7] R. J. Scherrer and M. S. Turner *On the relic, cosmic abundance of stable, weakly interacting massive particles.* Phys. Rev. D **33** (1986) 1585

[8] J. L. Feng, *Naturalness and the Status of Supersymmetry.* Annu. Rev. Nucl. Part. Sci. 2013.63:35182

[9] L. Roszkowski, E. M. Sessolo and S. Trojanowski, *WIMP dark matter candidates and searches—current status and future prospects.* Rept. Prog. Phys. **81** (2018) no.6, 066201

[10] M. Kuźniak, *Status of direct detection searches of heavy dark matter.* PoS **TAUP2023** (2023), 018

[11] C. Pérez de los Heros, *Status, Challenges and Directions in Indirect Dark Matter Searches*, Symmetry 2020, 12(10), 1648

[12] L. Evans and P. Bryant, *LHC Machine.* JINST **3** (2008) S08001.

[13] A. De Roeck, *Dark matter searches at accelerators.* Nucl. Phys. B **1003** (2024), 116480

[14] G. Krnjaic *et al.*, *A Snowmass Whitepaper: Dark Matter Production at Intensity-Frontier Experiments.*[arXiv:2207.00597 [hep-ph]]

[15] C. Antel *et al. Feebly-interacting particles: FIPs 2022 Workshop Report.* Eur. Phys. J. C **83** (2023) no.12, 1122

[16] E. Goudzovski *et al.*, *New Physics Searches at Kaon and Hyperon Factories.* 2023 Rep. Prog. Phys. **86** 016201

[17] T. Aoyama *et al.*, *The anomalous magnetic moment of the muon in the Standard Model.* Physics Reports **887** (2020) 1166

[18] M. Pospelov, *Secluded $U(1)$ below the weak scale.* Physics Review D **80** (2009) 095002

[19] J.R. Batley *et al.* (NA48/2 Collaboration), *Search for the dark photon in π^0 decays.* Physics Letters **B 746** (2015) 178-185

[20] X.G. He, G.C. Joshi, H. Lew and R. R. Volkas, *New-Z' phenomenology.* Phys. Rev. D **43** (1991) R22(R)

[21] Xiao-Gang He, G.C. Joshi, H. Lew, and R. R. Volkas, *Simplest Z' model.* Phys. Rev. D **44** (1991) 2118

[22] S. Baek, N.G. Deshpande, X.G. He and P. Ko, *Muon anomalous $g-2$ and gauged $L_\mu - L_\tau$ models.* Phys. Rev. D **64** (2001) 055006

[23] K. Harigaya, T. Igari, M.M. Nojiri, M. Takeuchi and K. Tobe, *Muon $g-2$ and LHC phenomenology in the $L_\mu - L_\tau$ gauge symmetric model.* JHEP **03** (2014) 105

[24] A. Greljo, Y. Soreq, P. Stangl, A.E. Thomsen and J. Zupan, *Muonic force behind flavor anomalies.* JHEP **04** (2022) 151

[25] J. P. Lees *et al.* (BaBar Collaboration), *Search for a muonic dark force at BABAR.* Phys. Rev. D **94** (2016) 011102(R)

[26] W. Altmannshofer, S. Gori, M. Pospelov, and I. Yavin, *Neutrino Trident Production: A Powerful Probe of New Physics with Neutrino Beams.* Phys. Rev. Lett. **113** (2014) 091801

[27] T. Kinoshita and W. J. Marciano, *Theory of the Muon Anomalous Magnetic Moment.* Adv. Ser. Direct. High Energy Phys. **7** (1990) 419

[28] G. Aad *et al.* (ATLAS collaboration), *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC.* Phys. Lett. B **716** (2012) 129

[29] S. Chatrchyan *et al.* (CMS collaboration), *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC.* Phys. Lett. B **716** (2012) 3061

[30] B. Batell, N. Lange, D. McKeen, M. Pospelov and A. Ritz, *Muon anomalous magnetic moment through the leptonic Higgs portal.* Phys. Rev. D **95**, 075003

[31] G.Krnjaic *et al.*, *Probing Muonphilic Force Carriers and Dark Matter at Kaon Factories.* Phys. Rev. Lett. **124**, 041802

[32] C.-Y. Chen, H. Davoudiasl, W. J. Marciano and C. Zhang, *Implications of a light "dark Higgs" solution to the* $g_\mu - 2$ *discrepancy.* Phys. Rev. D **93**, 035006

[33] J. Ellis, *TikZ-Feynman: Feynman diagrams with TikZ* (2016) arXiv:1601.05437 [hep-ph]

[34] M. Boretto, PhD thesis, *First observation of the rare decay mode* $K^+ \rightarrow \mu^+\nu_\mu\mu^+\mu^-$ *in the NA62 experiment at CERN SPS.* https://cds.cern.ch/record/2716696. Presented 15 Jan 2020.

[35] L. Marsicano *et al.*, *Probing leptophilic dark sectors at electron beam-dump facilities.* Phys. Rev. D **98** (2018) 115022

[36] M. Kobayashi and T. Maskawa, *CP-Violation in the Renormalizable Theory of Weak Interaction.* Prog. Theor. Phys. **49** (1973), pp. 652657

[37] S.L. Glashow, J. Iliopoulos, and L. Maiani, *Weak Interactions with Lepton-Hadron Symmetry.* Phys. Rev. D **2** (1970), pp. 12851292

[38] A.J. Buras *et al.*, $K^+ \rightarrow \pi^+\nu\bar{\nu}$ *and* $K_L \rightarrow \pi^0\nu\bar{\nu}$ *in the Standard Model: Status and Perspectives.* JHEP 11 (2015) 033

[39] G. D'Ambrosio, A.M. Iyer, F. Mahmoudi *et al.*, *Anatomy of kaon decays and prospects for lepton flavour universality violation.* JHEP **09** (2022) 148

[40] A.J. Buras, E. Venturini, *The exclusive vision of rare K and B decays and of the quark mixing in the standard model*, Eur. Phys. J. C **82** (2022) 615

[41] M. Blanke, A.J. Buras and S. Recksiegel, *Quark flavour observables in the Littlest Higgs model with T-parity after LHC Run 1* Eur. Phys. J. C **76** (2016) 182

[42] M. Blanke, A.J. Buras, B. Duling, K. Gemmler and S. Gori, *Rare K and B Decays in a warped extra dimension with custodial protection.* JHEP **03** (2009) 108

[43] A.J Buras, D. Buttazzo and R. Knegjens, $K \rightarrow \pi\nu\bar{\nu}$ *and* ϵ'/ϵ *in simplified new physics models.* JHEP **11** (2015) 166

[44] G. Isidori, F. Mescia, P. Paradisi, C. Smith and S. Trine, *Exploring the flavour structure of the MSSM with rare K decays.* JHEP **08** (2006) 064

[45] T. Blazek and P. Matak, *Non-minimal flavour violation, light pseudoscalar Higgs and evidence for SUSY in $B_s^0 \to \mu^+\mu^-$ decays.* Nucl. Phys. Proc. Suppl. **198** (2010) 216

[46] M. Tanimoto and K. Yamamoto, *Probing SUSY with 10 TeV stop mass in rare decays and CP violation of kaon.* PTEP **2016** (2016) 12, 123B02

[47] M. Bordone, D. Buttazzo, G. Isidori and J. Monnard, *Probing lepton-flavour universality with $K \to \pi\nu\bar{\nu}$ decays.* Eur. Phys. J. C **77** (2017) 618

[48] C. Bobeth and A.J. Buras, *Leptoquarks meet ϵ'/ϵ and rare Kaon processes.* JHEP **02** (2018) 101

[49] J. Aebischer, A.J. Buras and J. Kumar, *Another SMEFT story: Z' facing new results on ϵ'/ϵ, ΔM_K and $K \to \pi\nu\bar{\nu}$.* JHEP **12** (2020) 097

[50] S. Adler *et al.*, *Measurement of the $K^+ \to \pi^+\nu\bar{\nu}$ branching ratio.* Phys. Rev. D **77** (2008) 052003

[51] A. V. Artamonov *et al.* (E949 Collaboration), *Study of the decay $K^+ \to \pi^+\nu\bar{\nu}$ in the momentum region* $140 < P_\pi < 199\,MeV/c$. Phys. Rev. D **79** (2009) 092004

[52] E. Cortina Gil *et al.* (The NA62 Collaboration), *First search for $K^+ \to \pi^+\nu\bar{\nu}$ decay using the decay-in-flight technique.* Phys. Lett. **B791** (2019) 156

[53] E. Cortina Gil *et al.* (The NA62 Collaboration), *An investigation of the very rare $K^+ \to \pi^+\nu\bar{\nu}$ decay,* JHEP **11** (2020) 42

[54] E. Cortina Gil *et al.* (The NA62 Collaboration), *Measurement of the very rare $K^+ \to \pi^+\nu\bar{\nu}$ decay.* JHEP **06** (2021) 093

[55] E.Cortina Gil *et al.* (The NA62 Collaboration), *Performance of the NA62 trigger system.* JHEP **03** (2023) 122

[56] E.Cortina Gil *et al.* (NA62 Collaboration), *A measurement of the $K^+ \to \pi^+\mu^+\mu^-$ decay.* JHEP **11** (2022) 011

[57] E.Cortina Gil *et al.* (NA62 Collaboration), *A study of the $K^+ \to \pi^0 e^+\nu\gamma$ decay.* JHEP **09** (2023) 040

[58] E.Cortina Gil *et al.* (NA62 Collaboration), *Measurement of the $K^+ \to \pi^+\gamma\gamma$ decay.* Phys. Lett. **B850** (2024) 138513

[59] E.Cortina Gil *et al.* (NA62 Collaboration), *Search for the K^+ decays into the $\pi^+e^+e^-e^+e^-$ final state* Phys. Lett. **B846** (2023) 138193

[60] E.Cortina Gil *et al.* (NA62 Collaboration), *Search for lepton number and flavour violation in K^+ and π^0 decays* Phys. Rev. Lett. **127** (2021) 131802

[61] E.Cortina Gil *et al.* (NA62 Collaboration), *Search for leptonic decays of the dark photon at NA62.* arXiv.2312.12055[hep-ex], accepted by Phys. Rev. Letters on May 2024

[62] E.Cortina Gil *et al.* (NA62 Collaboration), *Search for dark photon decays to $\mu^+\mu^-$ at NA62.* JHEP **09** (2023) 035

[63] S. Navas *et al.* (Particle Data Group), *Review of particle physics.* Phys. Rev. D **110** (2024) 030001

[64] E.Cortina Gil *et al.* (NA62 Collaboration), *The beam and detector of the NA62 experiment at CERN.* JINST **12** (2017) P05025

[65] C. Bovet, R. Maleyran, L. Piemontese, A. Placci and M. Placidi, *The CEDAR Counters for Particle Identification in the SPS Secondary Beams.* CERN Report: CERN 82 (1982)

[66] E. Goudzovski *et al.*, *Development of the kaon tagging system for the NA62 experiment at CERN.* Nucl. Instrum. Meth. A **801** (2015) 86

[67] A. Bethani *et al.* (NA62 Collaboration), *Development of a new CEDAR for kaon identification at the NA62 experiment at CERN.* JINST **19** (2024) P05005

[68] M. Fiorini *et al.*, *High rate particle tracking and ultra-fast timing with a thin hybrid silicon pixel detector.* Nucl. Instrum. Meth. A **718** (2013) 270

[69] G. Nössle *et al.*, *NA62 gigatracker cooling with silicon micro channels.* Proc. ICATPP Conf. **8** (2013) 525

[70] D. Aisa *et al.*, *Mirror system of the RICH detector of the NA62 experiment.* JINST **12** (2017) P12017.

[71] G. Anzivino *et al.*, *Precise mirror alignment and basic performance of the RICH detector of the NA62 experiment at CERN.* JINST **13** (2018) P07012

[72] G. Anzivino *et al.*, *Light detection system and time resolution of the NA62 RICH.* JINST **15** (2020) P10025

[73] V. Fanti *et al.* (NA48 collaboration), *The beam and detector for the NA48 neutral kaon CP violations experiment at CERN.* Nucl. Instrum. Meth. A **574** (2007) 433

[74] OPAL Collaboration, *The OPAL detector at LEP.* Nucl. Instrum. Meth. A **305** (1991) 275

[75] H. Danielsson *et al.*, *New veto hodoscope ANTI-0 for the NA62 experiment at CERN.* JINST 15 (2020) 07, C07007

[76] R. Ammendola *et al.*, *The integrated low-level trigger and readout system of the CERN NA62 experiment* Nucl. Instrum. Meth. A **929** (2019) 1-22

[77] D. Protopopescu *et al.*, *NA62 Grid Monte Carlo Production Tools.* NA62-13-12.2013

[78] Rene Brun and Fons Rademakers, ROOT - An Object Oriented Data Analysis Framework, Proceedings AIHENP'96 Workshop, Lausanne, Sep. 1996, Nucl. Inst. & Meth. in Phys. Res. A 389 (1997) 81-86.

[79] GEANT4 Collaboration, *GEANT4: A Simulation toolkit.* Nucl.Instrum.Meth.A **506** (2003) 250-303

[80] C. E. Carlson, B. C. Rislow, *New physics and the proton radius problem.* Phys. Rev. D **86** (2012) 035013

[81] E. Barberio, Z. Was, *PHOTOS: A Universal Monte Carlo for QED radiative corrections. Version 2.0.* Comput. Phys. Commun. 79 (1994) 291-308

[82] C. Gatti, *Monte Carlo simulation for radiative kaon decays.* Eur. Phys. J. C **45**(2006)417

[83] J.R.Batley *et al* (NA48/2 Collaboration), *First observation and study of the $K^+ \to \pi^0\pi^0\mu^\pm\nu$ decay.* arXiv:2310.20295v1 [hep-ex] accepted for publication by JHEP

[84] R.Marchevski and G.Ruggiero, *$K^+ \to \nu\bar{\nu}$ Selection.* NA62 internal note NA62-18-04 dated 05.06.2018

[85] Aaboud, M., Aad, G. *et al.* (The ATLAS Collaboration), *Search for resonances in diphoton events at TeV with the ATLAS detector.* JHEP **09** (2016) 001.

[86] E. Cortina Gil *et al.* (The NA62 Collaboration), *Search for π^0 decays to invisible particles.* JHEP **02** (2021) 201.

[87] A. L. Read, *Presentation of search results: the CL_s technique.* J. Phys. G **28** (2002) 2693-704

[88] J. Henshaw, PhD thesis, *Beam intensity effects on the $K^+ \to \pi^+\nu\bar{\nu}$ measurement at NA62.* http://www.ep.ph.bham.ac.uk/publications/thesis/jackhenshaw.pdf, December 2023.

[89] E. Cortina Gil *et al.* (The NA62 Collaboration), *Searches for lepton number violating K+ decays.* Phys. Lett. B **797**(2019)134794

List of Figures

List of Tables