Università degli Studi di Modena e Reggio Emilia

Dipartimento di Scienze Fisiche Informatiche e Matematiche
MASTER DEGREE IN PHYSICS

Theoretical analysis and sensitivity study for the NA62 experiment of the decay of the K meson into two pions and a pseudoscalar particle

Relatore:

Prof. ANDREA BIZZETI

Correlatore:

Dott. DIEGO REDIGOLO

Laureando:

PRISCO LO CHIATTO

ACCADEMIC YEAR 2020-2021

Contents

Chapter 1

Axion

The QCD Axion is an hypothethical particle, that was first postulated in 1977 by Roberto Peccei and Helen Quinn [1], Franck Wilczek[2], and Steven Weinberg [3] as an answer to the strong CP problem The QCD Axion is a pseudo Nambu–Goldstone bosons of a chiral $U(1)$ symmetry, explicitly broken by a mixed 't Hooft anomaly which solves the strong CP problem.
Experimental evidence, including flavour experiments, bound arising from supernova collapse and direct detection experiments, set their mass to be low ($\lesssim 1\,\mu$eV). It was quickly realised that axions could serve as Dark Matter (DM) candidates and that sharp predictions about their properties could be made by requiring them to saturate the DM relic abundance today; most current experimental searches focus on the hypothesis of the QCD axion being a DM candidate.
Axion-like particles (ALP) are a generalisation of this concept: they are still pseudo Nambu–Goldstone bosons of a chiral $U(1)$, but the (dominant) explicit breaking of said symmetry is due to a generic operator, not the mixed anomaly. ALP do not, in general, provide a solution to the strong CP problem: their main motivation is to be a DM candidate, but they are also a generic prediction of many UV completion of the Standard Model, e.g. arising from extra-dimension compactification in String theory [4].
This chapter summarises some important properties of the axion.

1.1 Preliminaries: broken symmetries

1.1.1 Nambu-Goldstone Bosons

Given a conserved charge $\mathcal{Q}_a$ and an operator $\mathcal{O}$ such that their commutator has non-vanishing vacuum expectation value $\langle 0|\,[\mathcal{Q}_a, \mathcal{O}]\,|0\rangle = v \neq 0$, there exists a massless state $|\pi\rangle$ that couples non-vanishingly to both $\mathcal{O}$ and the current j^a associated with the conserved charge.

$$\langle 0|\,\mathcal{O}\,|\pi\rangle\,\langle\pi|\,j^a\,|0\rangle \neq 0 \tag{1.1}$$

This somewhat technical-sounding statement has a clear operational meaning: if an operator in the Lagrangian does not commute with a conserved charge we have massless states in the spectrum of the theory. These states are created from the vacuum by acting on it with said operator or with the conserved current. $v = \langle 0|\,[\mathcal{Q}_a, \mathcal{O}]\,|0\rangle$ is called the *order parameter* of SSB, since it measures the departure form the Wigner–Weyl realisation (which is an unbroken phase in the language of statistical physics).

An often-used toy model can help clarify the focal points. Consider the Lagrangian

$$L_\sigma = \frac{1}{4}Tr(\partial^\mu\Sigma^\dagger\partial_\mu\Sigma) - \frac{\lambda}{16}[Tr(\Sigma^\dagger\Sigma) - 2v^2]^2, \tag{1.2}$$

where $\Sigma(x)$ is a real 2x2 matrix. A (global) transformation according to the chiral group $SU(2)_L \otimes SU(2)_R$:

$$\Sigma' = g_L\Sigma g_R^\dagger \qquad g_{L/R} \in SU(2)_{L/R} \tag{1.3}$$

leaves the Lagrangian invariant.

To find the vacuum we minimise the potential: $Tr(\Sigma^\dagger\Sigma) = 2v^2 \Rightarrow \langle 0|\,\Sigma\,|0\rangle = vI_2$. This is not left invariant by the transformation above unless $g_L = g_R$, thus selecting the diagonal subgroup $SU(2)_V = SU(2)_{L+R}$. The broken group is $SU(2)_A = SU(2)_{L-R}$ which is 3-dimensional, suggesting the polar parametrisation of Σ using one scalar and three pseudoscalar fields:

$$\Sigma(x) = [v + S(x)]U(\vec{\pi}) \qquad U(\vec{\pi}) = \exp\{i\sigma_a\frac{\pi^a}{v}\} \tag{1.4}$$

with σ_a the three Pauli matrices.

Under a group transformation we have that:

$$\Sigma \to \Sigma' = g_L[v + S(x)]U(\vec{\pi})g_R^\dagger, \tag{1.5}$$

that is

$$S(x) \to S(x), U(\vec{\pi}) \to g_L U(\vec{\pi}) g_R^\dagger. \tag{1.6}$$

In this parametrisation the Lagrangian attains a particularly illuminating form:

$$\mathcal{L}_\sigma = \frac{v^2}{4}(1+\frac{S}{v})^2 Tr(\partial^\mu U^\dagger\partial_\mu U) + \frac{1}{2}(\partial^\mu S\partial_\mu S - 4\lambda^2v^4S^2) - 2\lambda^2v^3S^3 - \frac{\lambda^2v^2}{2}S^4, \tag{1.7}$$

the most important features for our purposes is that $\vec{\pi}$ only appears in the Lagrangian through its derivatives. This is in accord with the fact that Goldstone Bosons are constrained to be massless and have purely derivative couplings. The massless-ness is the statement of Goldstone theorem, and having purely derivative interaction protects the mass from higher-order correction which would spoil the vanishing mass condition at different energy scales. The symmetry that protects this is called shift-symmetry. Much can and has been said about this,[1] for our purposes it suffices to stress two facts:

1. Goldstone bosons enjoy a shift symmetry $\pi \to \pi + \alpha$, which force their interactions to be derivative and protects the mass to be vanishing

2. Purely derivative interactions imply that on-shell amplitudes vanish when one external momentum goes to zero. This is known as Adler's zero condition[7]

Pseudo Nambu–Goldstone Bosons

If the symmetry that undergoes SSB is also explicitly broken, many of the same results from the previous section apply, albeit approximately. Consider a theory whose Lagrangian L_0 enjoys invariance under the action of some group G:

$$\varphi \to D(g)\varphi \tag{1.8}$$

where $D(g)$ is the representation of the group G under which φ transforms.
Let's add a non-invariant operator to the Lagrangian $\Delta L = \varepsilon_a \mathcal{O}^a[\varphi]$, where $\varepsilon_a \ll 1$ are small real constants[2] and $\mathcal{O}^a[\varphi]$ are non-invariant operators that transform under some (non necessarily reducible) representation $\tilde{D}(g)$. Under the action of G

$$\mathcal{L} \to \mathcal{L} + \varepsilon_a \tilde{D}(g)^a_b \mathcal{O}^b, \tag{1.9}$$

hence the symmetry is (explicitly) broken.
We can employ a useful trick to account for the non-invariance: the spurion method. It works by noticing that if we replaced the constants[3] ε by fields transforming as

$$\varepsilon \to \tilde{D}^{-1}(g)\varepsilon \tag{1.10}$$

that is, if we introduced a *spurious* transformation law for them, we recover invariance, thus rapidly accounting for the pieces that fail to be invariant.

[1] Historically this belongs to the area of so-called soft pion theorems [5]. For a modern perspective see e.g. [6]

[2] Of course, if the coefficients are dimensional one must compare them to typical scales of the theory instead.

[3] We suppress indices from now now.

Using this method one can quickly recover the behaviour of matrix elements, dividing them into those that are unaffected by the symmetry breaking and those that instead vanish when $\varepsilon \to 0$. Then, a perturbative expansion is made in terms of the small parameters ε.
Turning back to Nambu–Goldstone bosons, explicit symmetry breaking spoils the shift-symmetry, hence the property of having exclusively derivative couplings. It follows that they acquire a mass: consider the above toy model, and add a scalar operator breaking the symmetry:

$$\mathcal{L}' = \mathcal{L}_\sigma + \varepsilon U + h.c. \tag{1.11}$$

Expanding up to second order in π, one finds a mass term for the bosons that is proportional to ε. Since we assumed the symmetry-breaking term to be small, we can account for its effect as a small perturbation: this means that we can describe the pions as approximate Nambu–Goldstone bosons, their properties approaching those of actual Nambu–Goldstone bosons as the symmetry breaking parameter vanishes. Such particles are usually called pseudo- Nambu–Goldstone boson. Pseudo-Nambu–Goldstone bosons have masses and non-derivative interactions that are small with respect to typical scales of the Lagrangian, being only in part protected by the shift symmetry.

1.2 The Strong CP Problem and the Axion Solution

1.2.1 The Chiral Anomaly

When quantising a classical field theory, one lifts the symmetry of the classical theory to the quantised one. In some cases, this procedure fails, the symmetry is broken and is said to be anomalous. The Standard Model has two source of anomalous breaking of the chiral symmetry (which relates left- and right-handed quarks) due to the coupling of the chiral coupling of the quarks to the gauge fields A_μ and G^a_μ. One particularly elegant way of understanding their origin is due to Fujikawa [8]: in the derivation of Ward identities, one makes the assumption that the path integral measure of integration is also invariant; this is not necessarily true, due to the Graßman nature of fermions: indeed consider the Lagrangian

$$\mathcal{L} = i\bar{\psi}\gamma^\mu(\partial_\mu - iA_\mu)\psi, \tag{1.12}$$

which is invariant under local $U(1)$ rotations $\psi \to \exp(i\alpha(x))\psi$ and global chiral $U(1)_A$ rotations $\psi \to \exp(i\gamma_5\alpha)\psi$. However, under an infinitesimal chiral rotation

$$\psi \to \psi + i\varepsilon\gamma_5\psi \qquad \bar{\psi} \to \bar{\psi} + i\varepsilon\bar{\psi}\gamma_5 \tag{1.13}$$

the measure picks up a Jacobian factor

$$\int D\psi D\bar{\psi} \to \int D\psi D\bar{\psi}\exp(-\frac{ie^2}{16\pi^2}\int d^4x F_{\mu\nu}\tilde{F}^{\mu\nu}), \tag{1.14}$$

where $F_{\mu\nu}$ is the field strength tensor and $\tilde{F}^{\mu\nu} = \frac{1}{2}\epsilon^{\mu\nu\alpha\beta}F_{\alpha\beta}$ is its dual. This means that it is not possible to gauge the chiral symmetry and the axial current $j^\mu_A = i\psi\gamma_\mu\gamma_5\psi$ is not conserved any more,

$$\partial_\mu j^\mu_A = \frac{e^2}{16\pi^2}F_{\mu\nu}\tilde{F}^{\mu\nu} \tag{1.15}$$

this is called the Adler–Bell–Jackiw current.

A different point of view on the anomaly can be obtained by computing it

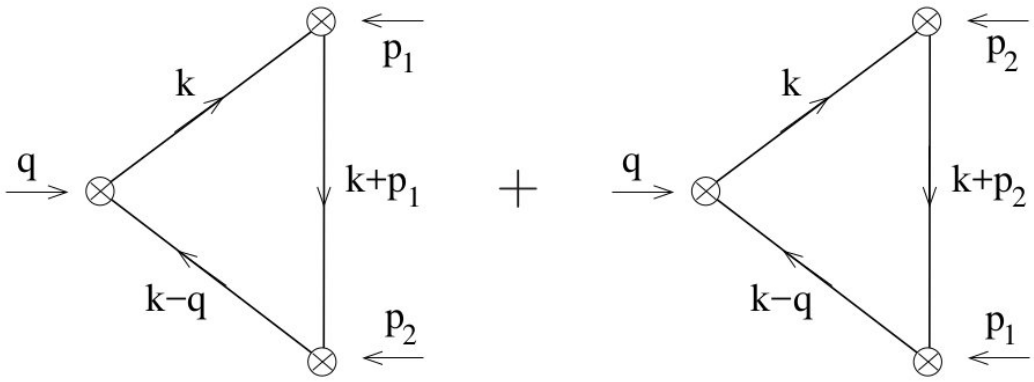


Figure 1.1: Triangle diagrams for the ABJ anomaly. At eacch vertex there is a current insertion. Picture from "Anomalies" chapter of David Tong's Lecture on Gauge Theories.

in terms of so-called triangle diagrams (see figure 1.1), where we introduce an arbitrary shift k in the loop momenta, which naively has no physical consequence. We can insist on this checking that the Ward identities for the two vector currents are respected. One finds (following the convention in figure 1.1 for the momenta):

$$\begin{aligned} ip_{1\mu}\Gamma^{\mu\nu\alpha}(p_1,p_2,q) &= \frac{1}{8\pi^2}\epsilon^{\nu\alpha\mu\beta}p_{1\mu}(k-p_2)_\beta \\ ip_{2\nu}\Gamma^{\mu\nu\alpha}(p_1,p_2,q) &= -\frac{1}{8\pi^2}\epsilon^{\nu\alpha\mu\beta}p_{1\nu}(k+p_1)_\beta \end{aligned} \tag{1.16}$$

so we get the Ward identity if $k = p_1 - p_2$. However, the Ward identity for the axial current also depends on k:

$$iq_\alpha \Gamma^{\mu\nu\alpha} = \frac{1}{8\pi^2}\epsilon^{\nu\alpha\mu\beta}[2p_{1\mu}p_{2\beta} + (p_{1\mu} + p_{2\mu})k_\beta] = \frac{1}{2\pi^2}\epsilon^{\nu\alpha\mu\beta}p_{1\mu}p_{2\beta} \tag{1.17}$$

Which can be rather easily shown to agree with the current in equation (1.15). This shows that the anomaly can in principle be shifted from the axial to the vector current, choosing k to satisfy the Ward identity for the axial current and making the vector symmetry anomalous. However, preventing the gauging of the vector symmetry would not allow the coupling of the fermion to photons, so we are forced to make the axial symmetry anomalous.
The anomaly described here is said to be *mixed*, since it is related to a triangle diagram including different currents, and of the 't Hooft type, since it is an obstruction to the gauging of a global symmetry. The Standard Model has two mixed 't Hooft anomalies, one involving QED currents and one for QCD currents, where the anomalous global current is a $U(1)_A$ current. They are called "chiral" anomalies.
The abelian case is the one in equation (1.15), while for QCD we get:

$$\partial_\mu j^\mu_A = \frac{g_s^2}{32\pi^2}G^a_{\mu\nu}\tilde{F}^{\mu\nu}_a \tag{1.18}$$

In both cases, one can account for the anomaly by adding the correspondent term in the Lagrangian, such that it is by default considered in the action.
This term is equal to the divergence of the current, up to a group-theoretic multiplicative factor called the anomaly coefficient.

1.2.2 Strong CP Problem

The Strong CP Problem is, at its heart, a naturalness problem; as such and despite its name it is better characterised as a puzzle, in the sense that nothing is fundamentally wrong with the theory, but it hints at deeper mechanisms at play.
Due to the chiral anomaly, the QCD Lagrangian contains a term of the form:

$$-\bar{\theta}\frac{g_s^2}{32\pi^2}G^a_{\mu\nu}\tilde{G}^{\mu\nu}_a, \tag{1.19}$$

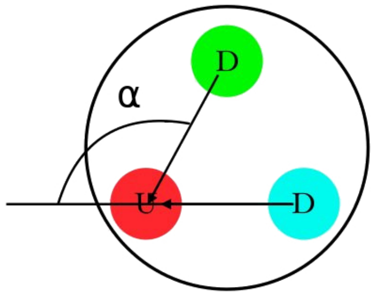
where $\bar{\theta}$ is a real constrant which definition is clarified in section 1.3
This term is both P and T odd[4], thus it can serve as an additional source of

[4]To get a feeling for this, consider the analogue in QED:$F_{\mu\nu}\tilde{F}^{\mu\nu} = -4E \cdot B$, which is both P and T odd since the electric field is even under time parity and odd under spatial parity, while the opposite is true for the magnetic field.

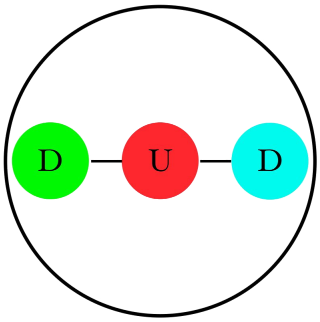
CP violation in the Standard Model; since the electroweak interaction already breaks this symmetry, there is no reason to think that the strong interaction should not.
CP violation in the strong sector has a clear experimental effect: the existence of an electrical dipole moment for the neutron d_n. This can be understood in a classical setting, clarifying the nature of the puzzle [9]. Classically, the neutron is made up from three quarks: $n = udd$, which form a triangle (see figure 1.2a). The direction of d_n is the same as the direction of the neutron spin, while its magnitude can be calculated since each pair of quark has a dipole moment $\sim er_n$, with r_n the radius of the nucleus:

$$|\mathbf{d_n}| = \frac{e\,r_n}{3}\sqrt{1-\cos\alpha} \approx 10^{-13}\sqrt{1-\cos\alpha}\; e\,\mathrm{cm} \tag{1.20}$$

Extremely high precision experiments have been performed, but none of them has been able to measure d_n, setting the stringent bound[10] of $\mathrm{d}_n < 10^{-26}e$ cm, which in turn implies[5] $\alpha < 10^{-13}$. This means that the quarks need to be aligned, or at least very nearly so (see figure 1.2b).



(a) Naive picture of the relative position of quarks.



(b) Updated picture after non observation of d_n.

Figure 1.2: Classical neutron. Picture from [9].

1.2.3 Peccei–Quinn Solution

Returning to QFT, the very strict experimental upper bound for d_n puts an upper bound of $\bar{\theta} < 10^{-10}$[9].
The axion is the field we get by making $\bar{\theta}$ dynamical:

$$-\bar{\theta}\frac{g_s^2}{32\pi^2}G^a_{\mu\nu}\tilde{G}^a_{\mu\nu} \rightarrow \frac{a}{f_a}\frac{g_s^2}{32\pi^2}G^a_{\mu\nu}\tilde{G}^a_{\mu\nu}. \tag{1.21}$$

[5]Note that $\alpha \neq \bar{\theta}$.

Then, if $\langle a \rangle = 0$, we have found an elegant way to forbid the existence of CP violation in the strong sector.
The simplest model we can think of that follows this prescription is to add an global $U(1)_{PQ}$ symmetry to the SM that has a mixed 't Hooft anomaly, such that a term of the type above is guaranteed. Then, since this symmetry is not present in the SM, we spontaneously break it, yielding the Nambu–Goldstone boson a, i.e. the dynamical theta angle. This boson, which we christen axion, does not enjoy a shift symmetry, because $U(1)_{PQ}$ is by construction broken by the anomaly, giving the axion a mass.
The important point is that this effectively shifts the value of $\bar{\theta}$: indeed it is customary to absorb it into the definition of the axion field $a + \bar{\theta} \to a$.
One then can calculate the potential induced by the QCD vacuum structure on the axion: a general argument, which follows from a theorem due to Vafa and Witten [11], guarantees us that the $\langle a \rangle$ = 0 is in general a minumum. This yields a mechanism that sets the theta angle to zero in a natural way.[6] It is worth noting that if we did set $\bar{\theta}$ to zero at the classical level, quantum correction due to (Standard Model) CP violating operator are expecte to mantain $\bar{\theta}$ extremely close to zero [12].

1.3 Definition of $\bar{\theta}$

Perhaps the strongest argument for the puzzling nature of a small value for $\bar{\theta}$ can be obtained by looking at its precise definition. If we look at the strong sector only, we have the following Lagrangian, including the now familiar CP violating term (note the subscript):

$$\mathcal{L}_{QCD} \supset -\theta_0 \frac{g_s^2}{32\pi^2} G^a_{\mu\nu}\tilde{G}^a_{\mu\nu} + i\bar{q}^i_L \gamma^\mu D_\mu q^i_L + i\bar{q}^i_R \gamma^\mu D_\mu q^i_R - (\bar{q}^i_L M^{ij} e^{i\theta_q} q^j_R + h.c.) \quad (1.22)$$

Where we introduced the SM chiral Weyl fields

$$q_{L/R} = \frac{1 \pm \gamma_5}{2} q \quad (1.23)$$

and M is a mass matrix generated through the Higgs mechanism, from which we pulled out a common phase.
The θ_0 term appearing here is unphysical: we can do a chiral U(1) rotation[9]:

$$q_{L/R} \to q_{L/R} e^{\pm i\alpha}$$

which, through the chiral anomaly, effectively induces an additional term $6\alpha G^a_{\mu\nu} G^{a\mu\nu}$ so that $\theta_0 \to \theta_0 - 6\alpha$.

[6] In section 1.4 we will perform an explicit computation to confirm this.

If this was the full story, we would always be able to redefine θ_0 away, so that it could not have any physical effect. It is not the case: the mass matrix also gets a correction when doing chiral rotations, picking up a phase $M_q e^{i\theta_q} \to M' = U^\dagger M U$, so that $det(M') = e^{6i\alpha} det(M)$. This lets us find a new quantity that is not changed under rotation and hence is physical:

$$\bar{\theta} = \theta_0 + \arg\det(M), \tag{1.24}$$

which is indeed invariant since the two terms rotate with opposite direction and equal magnitude.
As advertised at the start of this section, this makes a small value for the physical angle $\bar{\theta}$ all the more puzzling: why should two quantities coming from very different sectors in the SM (i.e. QCD non-perturbative physics and the Higgs mechanism) be equal to at least the 10th decimal place?

1.4 Properties of the Axion

The Goldstone nature of the axion, along with the anomalous breaking, fixes much of the properties of the axion. The main piece of the Lagrangian is the now-familiar CP violating term:

$$\mathcal{L}_a \supset \frac{1}{2}(\partial_\mu a)^2 + \frac{a}{f_a}\frac{g_s^2}{32\pi^2} G^a_{\mu\nu}\tilde{G}^{\mu\nu}_a \tag{1.25}$$

The decay constant f_a is related to the scale of PQ breaking Λ_{PQ} by the usual relation $\Lambda_{PQ} = 4\pi f_a$.

1.4.1 Axion Mass

To get the axion mass, we can rely on the potential calculated in the framework of χPT (see chapter 4). Since our goal is to perform a low-energy matching to meson fields, we rotate away the gluon coupling in favour of the quark bilinears. This is obtained by the axion-dependent rotation:

$$q \to \exp(\frac{i}{2}\frac{a}{f_a}\delta^A \gamma_5) q \tag{1.26}$$

where δ^A is a matrix in flavour space with unit trace. As observed in section 1.3, this induces a rotation on the quark mass matrix:

$$M_q^0 \bar{q} q \to \exp(i\frac{a}{f_a}\gamma_5 \delta^A) M_q^0 \bar{q} q, \tag{1.27}$$

that is, we can define a new mass matrix:

$$M_q = \exp(i\frac{a}{f}\gamma_5\delta_A)M_q^0\bar{q}q \tag{1.28}$$

We now use the $n_f = 2$ χPT Lagrangian at $\mathcal{O}(p^2)$:

$$\mathcal{L}^{(2)}_{CHPT} = \frac{F_0^2}{4}Tr(D_\mu U D^\mu U^\dagger) + \frac{F_0 B_0}{2}Tr(M_q^\dagger U + M_q U^\dagger), \tag{1.29}$$

where

$$U = \exp(\frac{i\sqrt{2}\Phi}{F_0}) \qquad \Phi = \sigma_a\phi^a(x) = \begin{pmatrix} \frac{\pi^0}{\sqrt{2}} & \pi^+ \\ \pi^- & -\frac{\pi^0}{\sqrt{2}} \end{pmatrix} \tag{1.30}$$

The axion potential is (the charged pion needs to have vanishing vev to respect $U(1)_{EM}$ gauge invariance):

$$\begin{gathered} V(\pi, a) = -B_0 F_0^2 \Re(Tr(UM_q^\dagger) = \\ -B_0F_0^2[m_u\cos(\frac{\pi^0}{F_0} + \frac{a}{f_a}) + m_d\cos(\frac{\pi^0}{F_0} - \frac{a}{f_a})] \end{gathered} \tag{1.31}$$

setting the π^0 vev to be vanishing and using $m_\pi^2 = B_0(m_u + m_d)$ we find:

$$V(a) = -m_\pi F_0^2\sqrt{1 - \frac{4m_u m_d}{(m_u+m_d)^2}\sin^2(\frac{a}{f_a})} \tag{1.32}$$

This is a crucial result. We have found that:

- $\langle a\rangle = 0$ minimises the potential, solving the CP problem.
- $m_a = V''(a) \sim m_\pi\frac{F_0}{f_a}$

1.4.2 Axion Coupling

Gauge and matter coupling are discussed in more detail in section 4.5, here we list some properties that will be relevant in the following section.
The coupling are in general dependent from the chosen UV completion, since they are dimension-5 operator. The most relevant for experimental searches is the coupling to photons, generated due to interplay between the QED and QCD anomaly:

$$\frac{1}{4}g_{a\gamma\gamma}F_{\mu\nu}\tilde{F}^{\mu\nu} \qquad g_{a\gamma\gamma} \propto \frac{\alpha}{2\pi f_a}, \tag{1.33}$$

where α is the fine structure constant.
Couplings to $SU(2)$ gauge bosons are also present in general, but can be

usually neglected since the energy scale that is relevant in most axion searches (especially DM searches) is extremely low.
Finally, there are coupling to matter currents. To respect the shift symmetry, this coupling needs to be derivative:

$$\frac{i}{2f_a}\partial_\mu a J^\mu = \sum_i \frac{\partial_\mu a}{F^A_{ii}}\bar{f}_i\gamma^\mu\gamma^5 f_i + \sum_{i\neq j}\partial_\mu a\bar{f}_i\gamma^\mu\left(\frac{1}{F^V_{ij}} + \frac{\gamma^5}{F^A_{ij}}\right)f_j, \tag{1.34}$$

where f_i are SM fermion fields and $F_{A/V} = 2f_a/C_{A/V}$ are effective decay constants, depending on the $\mathcal{O}(1)$ model-dependent constants $C_{A/V}$.
It is important to notice that there is no vectorial flavour-conservng coupling, because it is a total derivative[7].

1.5 Axion-Like Particles

A more general family of models can be obtained by considering cases in which the axion mass is not set by QCD dynamics only, but instead is set by some other UV mechanism. Since the $m_a \sim m_\pi f_\pi f_a^{-1}$ relation comes from QCD dynamics but $f_a^{-1} \sim g_{a\gamma\gamma}$ does not, these models have two free parameters (m_a and either f_a or $g_{a\gamma\gamma}$) instead of one.
Minimal models do not even contain the $G\tilde{G}$ coupling, but all models retain the anomalous coupling to the photon.

$$\mathcal{L} = \frac{1}{2}(\partial_\mu a)^2 + \frac{1}{2}m_a^2 a + \frac{g_{a\gamma\gamma}}{4}F_{\mu\nu}\tilde{F}^{\mu\nu} + \ldots \tag{1.35}$$

These models where the mass is not dictated by QCD effects go by the name of Axion-Like Particles (ALP). They are historically motivated by DM searches.
When it is necessary to distinguish them, the term QCD axion is used for the models where the properties are set by QCD dynamics and that necessarily solve the Strong CP problem.

1.6 The axion as a Dark Matter Candidate

For certain regions of the parameter space, axions can serve as a Cold Dark Matter (CDM) candidate.
First off, to have a DM candidate, we need to impose that the axion is stable

[7]To see this integrate by parts then use the equation of motion to get a total proportional to $m_i - m_i = 0$ and a total derivative.

against $a \to \gamma\gamma$ on cosmological timescales, i.e. $\tau_a \gtrsim \tau_U \approx 13.8\,Gyr$.
The decay rate of this process is:

$$\Gamma = \frac{m_a^3}{4\pi}(\frac{g_{a\gamma\gamma}}{4})^2 \propto \frac{m_a^3}{4\pi}\frac{\alpha}{16\pi^2 f_a^2} \Rightarrow \tau_a \approx \frac{2^8\pi^3 f_a^2}{\alpha m_a^3} \tag{1.36}$$

For the QCD axion we can use $f_a \approx \frac{m_\pi f_\pi}{m_a}$, obtaining an upper bound for m_a[8]:

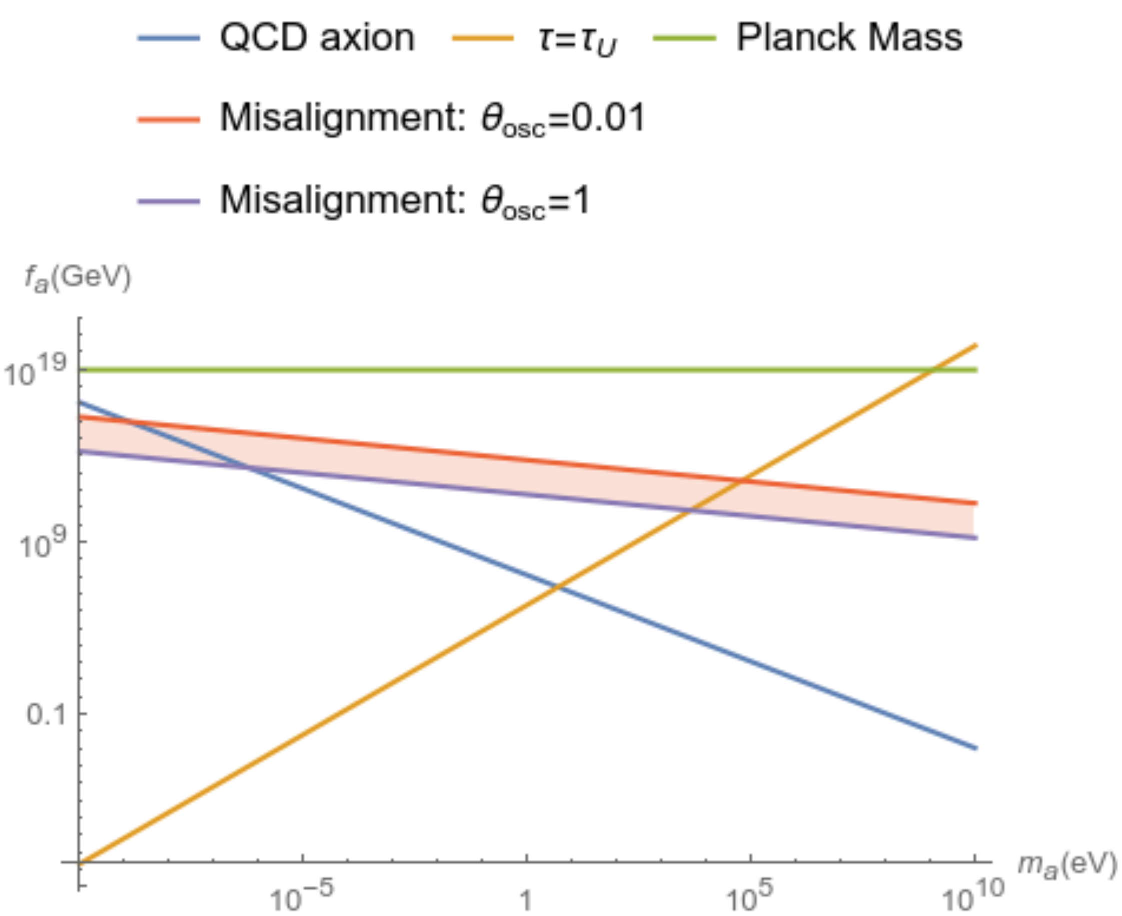


Figure 1.3: $f_a - m_a$ exclusion on the basis of cosmological stability for the ALP: any point above the mustard line is sufficiently stable. For QCD axions, models have to lie on the blue line.
Bound on the mass for QCD axions/ALP can be read from the intersection of the mustard line with the blue/green one (see text).
The shaded region shows the misalignment window, which is obtained by plugging $0.01 < \theta_{osc} < 1$ *into equation* (1.50).

$$\tau_U \lesssim \tau_a \approx \frac{1}{m_a^5}\frac{2^8\pi^3 f_\pi^2 m_\pi^2}{\alpha} \Rightarrow m_a \lesssim 10\,\mathrm{eV}. \tag{1.37}$$

For ALP, we can in general only exclude part of the $f_a - m_a$ plane. We can still make an order of magnitude prediction about an upper bound for the mass using the following argument: if f_a exceeds M_{Pl} the symmetry breaking happens in a regime where standard QFT cannot be trusted, so any effective theory for the ALP is not to be trusted, either. By requiring $f_a < M_{Pl}$, we get $m_{ALP} \lesssim 1$ GeV, provided the ALP decays into photons. The bounds are shown in figure 1.3.
Before continuing, it should be emphasised that these bounds are the weakest bound we can set for the axion to be a DM candidate. For example, imposing that the photon spectrum is not strongly modified by axion-produced photons can set more stringent bounds.
In any case, the low mass suggests that a non-thermal production mechanism is preferred in order to have a candidate for CDM.

1.6.1 Thermal production

To drive this point home, let's consider the thermal production of QCD axions via the process $gg \to ga$: the thermal production rate has been calculated to be [13]:

$$\Gamma \approx 10^{-5}\frac{T^3}{f_a^2}. \tag{1.38}$$

The usual criterion $\Gamma \lesssim H$ implies, in radiation domination (see appendix A), a decoupling temperature of:

$$T \gtrsim \frac{1}{10^{13}\,\mathrm{GeV}} f_a^2 = \frac{10^{-17}\,\mathrm{GeV}^3}{m_a^2}. \tag{1.39}$$

Axions decouple as non-relativistic particles if $T \lesssim m_a$, or equivalently

$$m_a^3 \gtrsim 10^{-17}\mathrm{GeV}^3 \Rightarrow m_a \gtrsim 1\,\mathrm{keV}, \tag{1.40}$$

which is indeed too heavy for the axion to be stable on cosmological scales.

[8] As a finer point one should consider the proportionality constant, which is related to the QED and QCD anomalies and could in general vanish, bypassing this argument.

1.6.2 Misalignment Mechanism

The misalignment mechanism is a non-thermal mechanism to produce a population of axions[14, 15]. It is based on the axion field being initially *misaligned* from the minimum potential value, then oscillating towards the minimum with frequency $\propto m_a$, producing particles in the zero mode. The oscillation can then be characterised as being that of a Bose–Einstein condensate.

The action for this system is, very generally:

$$\mathcal{S} = -\int d^4x\sqrt{-g}[\frac{1}{2}(\partial_\mu a)^2 + V(a)]. \tag{1.41}$$

Varying with respect to a one gets:

$$\Box a + V'(a) = 0. \tag{1.42}$$

Using the expression for the Dalembertian $\Box$ of a scalar field in FRW metric(appendix A) we get the equation of motion. For the zero-mode it reads:

$$\ddot{a}_0 + 3H(t)\dot{a}_0 + V'(a_0;t) = 0. \tag{1.43}$$

We are going to work in the approximation of small deviations from the minimum $V'(a_0;t) = m_a^2(t)a_0$. This also decouples different k-modes, as can be seen in Fourier space.

There are two relevant regimes (see figure 1.4):

- Over damped regime ($H(t) \gg m_a^2(t)$)
  $a_0 \equiv a_{osc}$
- Oscillating regime ($H(t) \approx m_a^2(t)$)

The small oscillation approximation can now be better understood as $\theta_{osc} = a_{osc}/f_a \lesssim 1$.

In between, we observe a transient regime in which the oscillator is approximately described as critically damped and as such quickly decays to the minimum, but overshoots it.

From now on we trade the time dependence of the parameters for a temperature dependence (see appendix A).

T independent mass For constant mass (this is usually the case for the ALP, where the mass is set at an extremely high scale and as such is for all intents and purposes constant) the equation becomes simpler [17]:

$$\ddot{a}_0 + 3H(t)\dot{a}_0 + m_a^2 a_0 = 0 \tag{1.44}$$

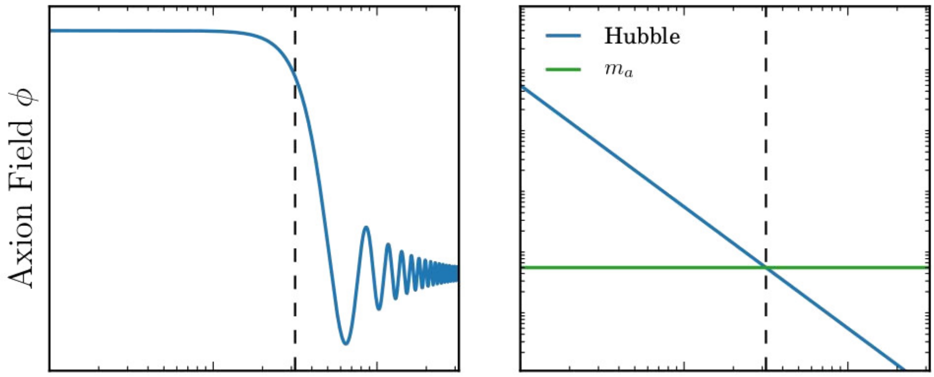


*Figure 1.4: Evolution of the zero-mode of the axion field (**left**) compared to the evolution of the Hubble parameter (**right**) in a simple model with constant axion mass and oscillation onset at $H = m_a$. Adapted from: [16]*

Let's define the energy density $\rho_a = \frac{1}{2}(\dot{a}_0^2 + m_a^2 a_0^2)$ and the number density $n_a = \rho_a/m_a$. Its time derivative is

$$\dot{\rho}_a = \ddot{a}_0 \dot{a}_0 + m_a^2 a_0 \dot{a}_0 = -3H(t)\dot{a}_0^2 \tag{1.45}$$

Where in the second equality we plugged in the previous equation.
Since we expect oscillations that are rapid with respect to the scale of H, we can time average over a period and substitute $< \dot{a}_0^2 >= m_a^2 < a_0^2 >$.
Using the same symbols in an abuse of notation, we get an equation for the time-average energy density:

$$\dot{\rho}_a = -3H\rho_a = -3\frac{\dot{R}}{R}\rho_a \tag{1.46}$$

with the solution $\rho_a \propto m_a R^{-3}$. The energy density redshifts as matter, as expected.
Let's call T_{osc} the temperature at which the oscillations begin; we also define the number density at the start of the oscillation

$$n_{osc} = n(T_{osc}) = \frac{1}{2m_a}(m_a^2 a_{osc}^2) = \frac{1}{2} m_a a_{osc}^2. \tag{1.47}$$

Let's now assume that the evolution of the universe is adiabatic. The yield $Y = n_a/s$ is then conserved; this lets us calculate the number density today from $n_a(T_0) = n_{osc}\frac{s(T_0)}{s(T_{osc})}$

$$\Omega_a = \frac{\rho_a(T_0)}{\rho_c} = \frac{m_a n_{osc} s(T_0)}{\rho_c s(T_{osc})} = \frac{m_a^2 a_{osc}^2}{2\rho_c}\left(\frac{T_0}{T_{osc}}\right)^3 \frac{g_{*S}(T_0)}{g_{*S}(T_{osc})} \tag{1.48}$$

where the expression for the entropy density in standard cosmology (see appendix A) was used in the last equation and ρ_c is the critical density.
Different cosmological scenarios result in different values for T_{osc} and n_{osc}, and theories with different particle content can have different values for the effective degrees of freedom.
Only radiation domination is considered in this document, and no other d.o.f. is added to the SM ones other than the axion.

Radiation domination Using the standard cosmology expression for the Hubble parameter (see appendix A), we impose $H(T_{osc}) = m_a$. Then we have

$$\frac{1}{T_{osc}^3} = \left(\frac{\pi^2}{90M_{Pl}^2}\frac{g_{*S}(T_{osc})}{m_a^2}\right)^{\frac{3}{4}} \tag{1.49}$$

We have then reached an expression for the relic abundance:

$$\begin{aligned} \Omega_a h^2 &= \frac{1}{2}\frac{h^2}{\rho_c}\left(\frac{\pi^2}{90M_{Pl}^2}\right)^{\frac{3}{4}} T_0^3 m_a^{\frac{1}{2}} a_{osc}^2 \left(\frac{g_{*S}(T_0)}{g_{*S}(T_{osc})}\right)^{\frac{1}{4}} \\ &\approx 0.12\left(\frac{f_a\theta_{osc}}{1.9\times 10^{13}\,\mathrm{GeV}}\right)^2\left(\frac{m_a}{1\,\mu\mathrm{eV}}\right)^{\frac{1}{2}}\left(\frac{90}{g_{*S}(T_{osc})}\right)^{\frac{1}{4}} \end{aligned} \tag{1.50}$$

Where in the last equality we swapped the initial value of the field a_{osc} for $f_a\theta_{osc}$, with θ_{osc} expected to be $\mathcal{O}(1)$ by naturalness (and to comply with our small-angles approximation).
For the QCD axion this expression only contains two parameter (f_a and θ_{osc}) so assuming natural values for θ_{osc} defines the so-called "axion window" (see the shaded band in figure 1.3).

1.7 Experimental Searches

As seen in section 1.4, all axion models are characterised by a $a\gamma\gamma$ coupling:

$$\mathcal{L} \supset g_{a\gamma\gamma} aF\tilde{F} = -4g_{a\gamma\gamma} a\mathbf{E}\cdot\mathbf{B} \tag{1.51}$$

most experimental searches exploit this coupling, in the production and/or in the detection of axions. In the presence of an intense magnetic field, this term means that axions and photons mix.

$$\begin{pmatrix}\gamma' \\ a'\end{pmatrix} = M\begin{pmatrix}\gamma \\ a\end{pmatrix}. \tag{1.52}$$

So that the physical, propagating fields are related to the $U(1)_{em}$ gauge boson and the pNGB by rotating through some matrix M having non-empty off-diagonal entries.
For production, we can consider axions generated inside of dense nuclear medium, such as stars. This is called Primakoff effect, named after the analogue mechanism for pseudoscalar mesons. Its inverse process turns axions into photons in the presence of strong magnetic field, and it is used for indirect detection.

*Figure 1.5: Primakoff effect for the production of axions (**left**) and its inverse (**right**), which can be used for detection. Taken from [18]*

1.7.1 Helioscopes

Helioscopes (from *Helios*, the Ancient Greek word for Sun) look for Sun-produced axions, using powerful magnets to observe $\gamma \to a$ conversion.
Let's look in a bit more detail at the mixing: consider axions and photons propagating along the x directions, in the presence of a magnetic field along the z direction, B_z.
The linearised mixing matrix is [19]:

$$M = \mathbb{1}_2(p_x + \omega) + \begin{pmatrix} \frac{C}{2\omega} & \frac{1}{2}g_{a\gamma\gamma}B_z \\ \frac{1}{2}g_{a\gamma\gamma}B_z & -\frac{m_a^2}{2\omega} \end{pmatrix}, \tag{1.53}$$

where C acts as an effective photon mass, which is in most cases negligible.
After diagonalising by rotating through the angle $\theta_{a\gamma} \approx \frac{g_{a\gamma\gamma}B_z\omega}{m_a^2}$, we can extract the conversion probability as a function of the distance travelled inside the detector L:

$$P_{a\to\gamma}(L) = \left(\frac{g_{a\gamma\gamma}B_zL}{2}\right) sinc^2\left(\frac{m_a^2 L}{4\omega}\right), \tag{1.54}$$

where $sinc(x) = \frac{\sin(x)}{x}$ has two regimes: when its argument is less than one we approximate $sinc(x) = 1$ and so:

$$P_{a\to\gamma} \approx \left(\frac{g_{a\gamma\gamma} B_z L}{2}\right)^2 . \tag{1.55}$$

On the other hand, the rapid oscillations at high values of m_a^2 and/or L can be averaged, giving:

$$P_{a\to\gamma} \approx 2\left(\frac{g_{a\gamma\gamma} B_z \omega}{m_a^2}\right)^2 . \tag{1.56}$$

The CERN Axion Solar Telescope (CAST [20]) used a decommissioned LHC magnet with a magnetic field of 8 T and a length of 10 m, meaning the change in regimes for the $sinc$ function is around $m_a \sim 10$ meV.
For masses lower than this value, the bound it puts on $g_{a\gamma\gamma}$ is independent of the mass, while for higher masses it weakens as m_a^2. The stronger bound it obtained is [20]:

$$g_{a\gamma\gamma} \lesssim 10^{10}\,\mathrm{GeV}^{-1} @\, m_a \lesssim 10\,\mathrm{meV} \tag{1.57}$$

1.7.2 Light Shining Through Walls

A similar experiment — which goes under the colourful name of "light shining through walls" — can be performed by using powerful lasers instead of the sun as an axion source.
A powerful laser is shone inside a powerful magnetic field, so that a part of the beam converts to axions. A wall is put in the path of the beam, so that none of the photons and all of the axions can pass through it; finally the now fully axionic beam passes again through a powerful magnetic field, thus partially converting into photons. Since the original beam of photons is stopped by the wall, any photon detected on the other side of it needs to come from a regeneration process.
The probability for this process to happen is:

$$P_{\gamma\to a\to\gamma} = P_{a\to\gamma}^2 \approx \left(\frac{g_{a\gamma\gamma}|\mathbf{B}|L}{2}\right)^4 . \tag{1.58}$$

Current bounds come from the DESY experiment ALPS and the CERN experiment OSQAR [21]:

$$g_{a\gamma\gamma} \lesssim 10^{-8}\mathrm{GeV}^{-1} \;\; @\, m_a \lesssim 1\,\mathrm{meV} \tag{1.59}$$

The bound is worse than the one from Helioscope, but the physical setting is under better control.

1.7.3 Haloscopes

Haloscopes are a DM-specific search: indeed as the name suggest they look for axionic wind coming from the halo of DM in our galaxy.
The experiment consists of a microwave cavity with high magnetic field: through inverse Primakoff effect the axions of the halo can interact inside it and resonantly produce photons.
As seen from the discussion on the misalignment mechanism, the axion field can be treated as a classical background field, so that we can think of it as a spatially constant source term for the electric field. The EOM for the electric field in frequency space reads:

$$(\omega^2 - \omega_j^2)E_{\omega,j} = g_{a\gamma\gamma}\omega^2 a_\omega B_{0,j}, \tag{1.60}$$

where j indexes the resonant frequencies for the cavity ω_j. The solutions of these equations take a familiar Lorentzian form:

$$E_{\omega,j} = g_{a\gamma\gamma}a_\omega B_{0,j}\frac{\omega^2}{\omega^2 - \omega_j^2 + i\omega\Gamma}. \tag{1.61}$$

Then by sweeping the cavity resonant frequency over some range, we would get a peak at the axion mass.
This kind of experiment is plagued by poor signal to noise ratio: current and next generation experiments use SQUID antenna to reduce thermal noise. Better signal to noise ratio means having to spend less time at a said frequency to receive the same amount of signal, thus being able to sweep a larger window. Current measurement exclude axions with $f_a \sim 10^{12}$ GeV [22].

1.7.4 SN1987A

In 1987, the Kamiokande experiment (among others) registered a long-lasting (~ 10 s) burst of neutrinos coming from the centre of the SN1987A supernova.
This validated the theory of supernova collapse, and as such put limit on new phenomena. In particular, it was quickly realised [23, 24] that axion flux would decrease the duration of the burst dramatically unless some conditions are met. The idea is that, had the axion carried away too much energy from the core, the collapse would have been much less long-lasting than observed.
Indeed, we can make a very simple argument to constrain the total luminosity of the flux: the gravitational energy of the supernova core, which had a mass of roughly $M_\odot$, is 3×10^{53} erg, so for the axion flux not to impact too much the duration of the collapse we need the luminosity of the flux of axion to respect

$L_a \lesssim 10^{52}$ erg/s.
The dominant process from which axions can carry energy away from the supernova is nucleon brehmsstrahlung $N + N \to N + N + a$. After being produced the axion can either escape the core, if the coupling to nucleons is weak enough that the mean free path is more than the radius of the core, or be trapped if the coupling is sufficiently strong. In the first case the luminosity is proportional to the coupling squared g^2_{aNN}; in the second case we can consider an "axionsphere" of radius r_x made up from trapped axions, radiating via blackbody emission, so the luminosity is $\propto r_x^2 T^4(r_x)$, which rapidly decreases as the coupling gets bigger (hence the radius gets smaller).
While both weak and strong coupling can provide an acceptable luminosity,

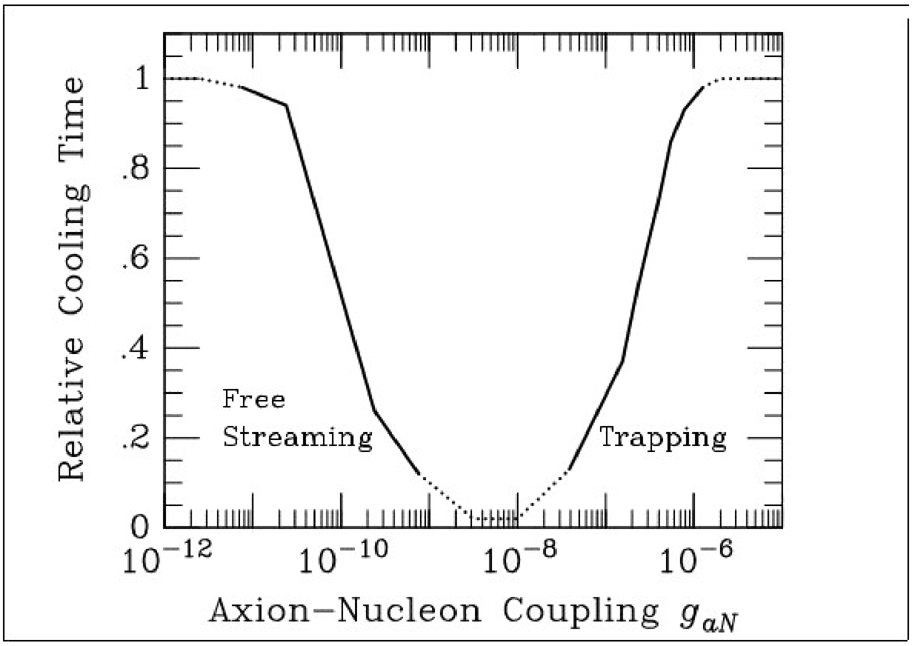


Figure 1.6: Duration of the neutrino burst of SN1987A relative to the measured one vs. axion-nucleon coupling. As can be seen both a low and a high value for the coupling can comply with observation. The solid line comes from numerical simulation, while the dotted one is a continuation to guide the eyes. Taken from [25]

the latter is to be excluded for other reasons (not least of which, we would have observed such strongly interacting axions in precision experiments), so we only consider weak coupling, which for QCD axions can be turned into constraint on the mass $m_a \lesssim 0.054\,\mathrm{eV}$.
For models where the coupling to nucleon is zero (nucleophobic axions) we have to consider Primakoff effect as the cause for energy loss, which leads to the bound $g_{a\gamma\gamma} \lesssim 5 \times 10^{-10}\,\mathrm{GeV}^{-1}$ [26].

1.7.5 Kaon Physics

Flavour physics has historically played an important role in axion physics: the original PQWW model, with the scale of breaking set to the electroweak scale, was excluded by the non observation of the $K \to \pi a$ decay.
For realistic axion models, current experiments have trouble setting limits that can compete with astrophysical bounds. This is because the axion is usually thought to only have flavour diagonal coupling (i.e. flavour off-diagonal terms in equation (1.34) is set to zero).
In this case, the flavour changing neutral current would have to be supplied by two vertices involving a W boson, as in the decay $K \to 2\pi$; then the mixing of axions with neutral mesons would provide the needed vertex. The branching ratio would be:

$$BR(K \to \pi a) = BR(K \to 2\pi) BR(\pi \to a) \approx 0.2 \left(\frac{f_\pi}{f_a} \right)^2$$

So that the current bound $BR(K \to \pi a) < 7.3 \times 10^{-11}$ gives $f_a \gtrsim 10^5$ GeV, 6 orders of magnitude below the bound put by stability.
There is a class of QCD axion model [27, 28] where FCNC can happen at tree level (i.e. we don't suppress the flavour off-diagonal terms). They are theoretically motivated by being a solution to the Flavour problem; their existence could then provide a simultaneous solution to three of the problems plaguing the Standard Model.
If FCNC is allowed at tree level, we have

$$BR(K \to \pi a) = \frac{m_K^3}{64\pi\Gamma_K} \left(F_{sd}^V\right)^2$$

with F_{sd}^V being the one appearing in equation (1.34).
The bound coming from these models would be $f_a \gtrsim 7.5 \times 10^{10}$ GeV, right in the middle of the axion window. Moreover, current generation Kaon factories could improve this limit by 1-2 orders of magnitude.

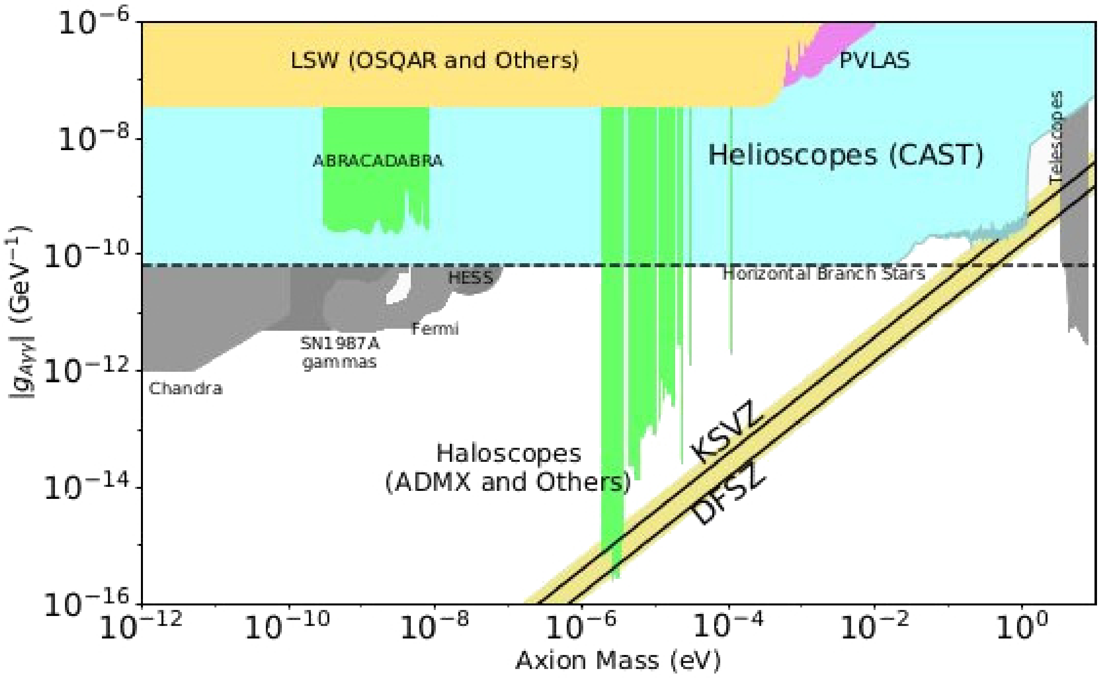


Figure 1.7: Summary of $m_a/g_{a\gamma\gamma}$ bounds [29].
We can recognise the ones referring to LSW (section 1.7.2), Helioscopes (section 1.7.1), Haloscopes (section 1.7.3) and the supernova bound (section 1.7.4). The lines labelled "KSVZ" and "DFSZ" refer to two realistic UV completion for the QCD axion, which are not discussed in this thesis.

Chapter 2

Axiflavons

Axiflavons are a particular kind of axion model where the axions has a flavour structure, i.e. it mixes with flavour observables. They are motivated by providing a solution to the flavour problem, on top of the strong CP problem and being a candidate for dark matter

2.1 The Flavour Problem and Froggatt–Nielsen mechanism

The flavour problem is the puzzle of why there are such different fermion masses. Since fermion masses are generated by the Higgs mechanism, this amounts to asking why the Yukawa couplings of different generation of fermions span so many different orders of magnitude, in contrast with the naturalness prescription that all coupling not enforced by symmetry principles should be of order one.
Froggatt and Nielsen [30] devised a mechanism to explain the hierarchy in the masses of fermions, as well as in the mixing angles. The set-up is extremely simple: one adds a $U(1)_{FN}$ symmetry to the SM, under which left and right-handed quarks have opposite charges, forbidding fermion mass terms. Then, the symmetry is spontaneously broken through the vacuum expectation value of a scalar field called the "flavon" Φ. The final ingredient is some heavy fermions, called Froggatt–Nielsen messengers, that are also charged under $U(1)_{FN}$, and thus can connect the standard model fermions through so-called "spaghetti diagrams", see figure 2.1. Since neither the flavon nor the messenger fermions have been observed, one sets both the vev of the flavon $\langle\Phi\rangle$ and the masses of the messengers M_{FN} to high scales, such that they can be safely integrated out. Indeed, they could very well be near the Planck scale, as long as $\langle\Phi\rangle/M_{FN} < 1$.

For simplicity, we show a simple realisation in a two-family model: the

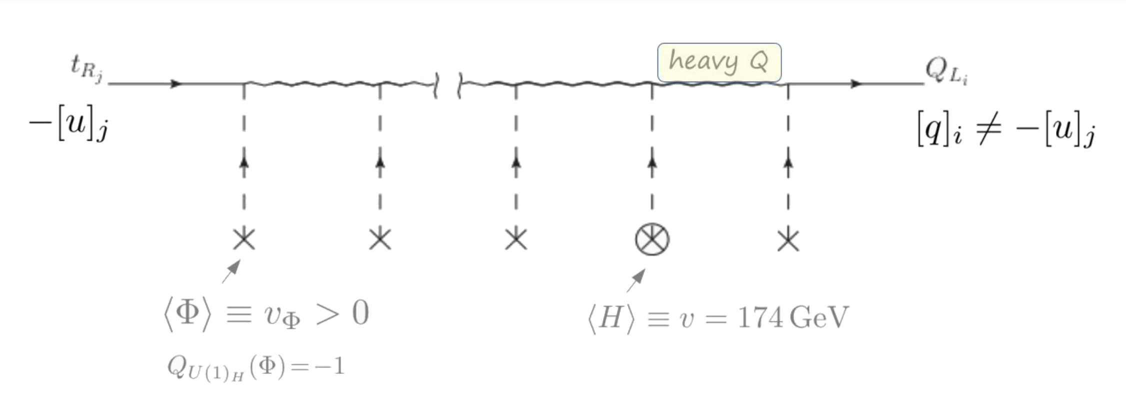


Figure 2.1: A "spaghetti diagram" for the Froggatt–Nielsen mechanism. One can see how different insertion of the flavon (as well as one Higgs insertion) are necessary to connect left and right-handed quarks. After integrating out the heavy messenger quarks, this gives an effective Yukawa interaction. Credits: Florian Goertz

Yukawa sector of the SM is (we suppress flavour indices):

$$\mathcal{L}^Y = y^u \bar{q}_L \tilde{H} u + y^d \bar{q}_L H d, \tag{2.1}$$

where q_L is the left-handed doublet, d and u the right-handed singlets, H the Higgs field and $\tilde{H} = i\sigma_2 H^*$.
The $U(1)_{FN}$ acts as $f \to e^{i\theta X_f} f$, where X_f the charge of the field f. We impose $X_{q_L} \neq X_{u/d}$.
We then add the flavon with charge $X_\Phi = 1$. As said before, we also add heavy messenger fields with masses around M_{FN}, and immediately integrate them out. The corresponding EFT contains these terms consistent with gauge symmetry and $U(1)_{FN}$:

$$\mathcal{L}_{FN} = c^u \bar{q_L} \tilde{H} u (\frac{\Phi}{M_{FN}})^{|X_{q_L} - X_u|} + c^d \bar{q_L} H d (\frac{\Phi}{M_{FN}})^{|X_{q_L} - X_d|} + h.c. \tag{2.2}$$

Finally, we give the flavon a vev $\langle\Phi\rangle$. The EFT lagrangian now looks like

$$\mathcal{L}_{FN} = c^u \varepsilon^{|X_{q_L} - X_u|} \bar{q_L} \tilde{H} u + c^d \varepsilon^{|X_{q_L} - X_d|} \bar{q_L} H d + h.c., \tag{2.3}$$

the "spaghetti diagrams" of the full theory now look like Yukawa terms, suppressed by powers of $\varepsilon = \langle\Phi\rangle / M_{FN}$. The Yukawa coefficients $c^{u/d}$ can now very well be $\mathcal{O}(1)$, the hierarchy fully reproduced by the different powers of the breaking parameter.

2.2 Flavoured Axions

Since both the Peccei–Quinn and the Froggat–Nielsen extension to the SM require adding a spontaneously broken $U(1)$ symmetry to the standard model, one can envision adding a single, anomalous $U(1)_{FN=PQ}$ symmetry that encapsulates both solutions [31]. This symmetry is then broken by the vev of a field S

$$S = \frac{1}{\sqrt{2}}(v_{\S} + \Phi)e^{i\frac{a}{v_S}}, \tag{2.4}$$

containing both the flavon and the axion. The flavon would then be expected to be heavy w.r.t to standard model energies ($\mathcal{O}(v_S)$), so it can be safely integrated out, leaving the axion whose interaction with SM fields are fixed by the (now common) $U(1)$ charge assignment.
This model is dubbed “axiflavon” and is part of a larger class of models which feature Peccei–Quinn symmetry as part of an enlarged symmetry in the flavour sector.

2.2.1 Bounds on Axiflavons

Axiflavons are interesting from a phenomenological perspective because they allow for tree-level flavour-changing neutral currents (FCNC) mediated by the axion. The Froggatt–Nielsen mechanism requires different $U(1)_{FN=PQ}$ charges for the different fermion generations, hence the charge matrix trivially does not commute with the fermions mass matrix. FCNC are greatly suppressed in the SM[1] and constrained experimentally, so they pose great challenges to any theory that predicts them. We can turn this challenge into an opportunity to test axion physics at flavour factories.
The best constraints for FCNC involving an axion comes from $K \to \pi a$ [27]. In a rather rough way, one can think about the decay in a 2-step fashion: first a SM penguin diagram induces the $K \to \pi\pi^0$ decay, then the neutral pion-axion mixing gives the final state. This means the rate is proportional to the ratio G_F^2/f_a. Current best measurements place the upper limit on the branching ratio of this decay around 10^{-11}, giving a lower bound $f_a \gtrsim 10^5$ GeV, which is extremely poor when compared to astrophysical bounds. However, if the FCNC happens at tree level, the rate is proportional to simply $1/f_a$, giving a lower bound of order 10^{11} GeV. If the axiflavon model is realised in nature, kaon factories can then probe axion models on scales higher than astrophysics, which is rather remarkable.
If we look at the axion-fermion couplings in equation (1.34), we can see that

[1] they are disallowed at tree-level and suppressed by the GIM mechanism at one-loop

they split nicely into a vector and an axial part. The $K \to \pi a$ decay is sensitive only to the vector coupling because of Wigner–Eckart theorem, so the axial part goes unprobed; since in general models for flavoured axions they can be not trivially related, it would make sense to probe them separately. The decay that does so is $K \to \pi\pi a$: the best limits on this decay come are at least a decade old, so the aim of this thesis is to study the sensitivity of current-generation experiments.

Chapter 3

Differential rates

3.1 Fully differential rate for $K^+ \to \pi^+\pi^0 a$

To be able to perform the sensitivity study, we will need the differential rate for the decay. In the spirit of having the most general search possible, we will not restrict ourselves to any particular model, calculating instead the rate for $K^+ \to \pi^+\pi^0 a$, where a — which we will continue calling an axion for short — is any pseudoscalar particle that interacts derivatively. We also will not impose any hypothesis on the mass. By doing so, our search will be as model-independent as possible.

3.1.1 Notation

4-vectors are denoted as a^μ; 3-vectors as $\vec{v}$, and its modulus is V. Quantities related to the Kaon have no index. Non-invariant quantities are denoted by an asterisk (*) when calculated in the 2-3 centre of mass frame.

3.1.2 Phase Space

In the decaying particle's rest frame, where $p^\mu = (M, \vec{0})$, the term to be simplified is:

$$(2\pi)^4\delta^4(p^\mu - \Sigma_{i=1}^3 p_i^\mu)\Pi_{i=1}^3 d^3p_i. \tag{3.1}$$

The delta function constrains the number of independent variables to be 3*3-4; moreover, since the decaying particle is scalar, we can eliminate three more variables, ending up with 3*3-7 = 2 independent variables.
We can eliminate $\vec{p_3}$ by dropping $\delta^3(\vec{p} - \Sigma_{i=1}^3 \vec{p_i})$, and writing all expressions with the understanding that $\vec{p_3} = -(\vec{p_1} + \vec{p_2})$. We also write $d^3p_i = P_i^2 dp_i d(\cos\theta)d\varphi$.

We are now left with

$$\delta(M-\sqrt{P_1^2+m_1^2}-\sqrt{P_2^2+m_2^2}-\sqrt{P_1^2+P_2^2+P_1P_2\cos\theta_{12}+m_3^2})P_1^2P_2^2\times$$
$$dp_1dp_2d(\cos\theta_1)d\varphi_1d(\cos\theta_{12})d\varphi_{12}.$$

Using the known property of the delta function

$$f(x_0)=0\wedge f'(x_0)\neq 0\Rightarrow \delta(f(x))=\frac{\delta(x-x_0)}{|f'(x_0)|}, \tag{3.2}$$

we can drop the delta gaining a factor of $\frac{E_3}{P_1P_2}$. Integrating over $d(\cos\theta_1)d\varphi_1d\varphi_{12}$ we get

$$(2\pi)^4\delta^4(p^\mu-\Sigma_{i=1}^3p_i^\mu)\Pi_{i=1}^3d^3p_i\rightarrow 8\pi^2P_1P_2E_3dp_1dp_2=8\pi^2E_1E_2E_3dE_1dE_2, \tag{3.3}$$

where the dispersion relation has been used in the last equality. We can now express the Lorentz-invariant phase space as

$$LIPS^{(3)}=(2\pi)^4\delta^4(p^\mu-\Sigma_{i=1}^3p_i^\mu)\Pi_{i=1}^3\frac{d^3p_i}{(2\pi)^32E_i}=\frac{dE_1dE_2}{4(2\pi)^3}. \tag{3.4}$$

We want to express this in term of invariant variables. We know that $m_{23}^2=(p_2+p_3)^\mu(p_2+p_3)_\mu=(p-p_1)^\mu(p-p_1)_\mu$ and that $p^\mu=(M,\vec{0})$. Then $m_{23}^2=M^2+m_1^2-2ME_1$. So $E_1=\frac{M^2+m_3^2-m_{23}^2}{2M}$. Hence

$$dLIPS^{(3)}=\frac{dm_{23}^2dm_{12}^2}{16M^2(2\pi)^3}. \tag{3.5}$$

We now want to go to the 2-3 COM system, where $(p_2+p_3)^\mu\overset{DEF}{=}(m_{23},\vec{0})$ and $\vec{p}=\vec{p_1}$. To get there we massage the expression for $m_{23}^2=(p_2+p_3)^\mu(p_2+p_3)_\mu$:

$$m_{23}^2=2E^{*2}-m_1^2+m_3^2+2E_2^*E_3^*\Rightarrow E_2^*=\frac{m_{23}^2+m_2^2-m_3^2}{2m_{23}} \tag{3.6}$$

and similarly we can get $E^*=\frac{M^2+m_{23}^2-m_1^2}{2m_{23}}$, $E_1^*=\frac{M^2-m_{23}^2-m_1^2}{2m_{23}}$.
From there we get

$$P_2^{*2}(=P_3^{*2})=E_2^2-m_2^2=\frac{m_{23}^4+m_2^4+m_3^4-2m_{23}^2m_2^2-2m_{23}^2m_3^2-2m_2^2m_3^2}{4m_{23}^2}=$$
$$=\frac{(m_{23}^2-(m_2+m_3)^2)(m_{23}^2-(m_2-m_3)^2)}{4m_{23}^2}\overset{DEF}{=}\frac{\lambda(m_{23},m_2,m_3)}{4m_{23}^2}$$

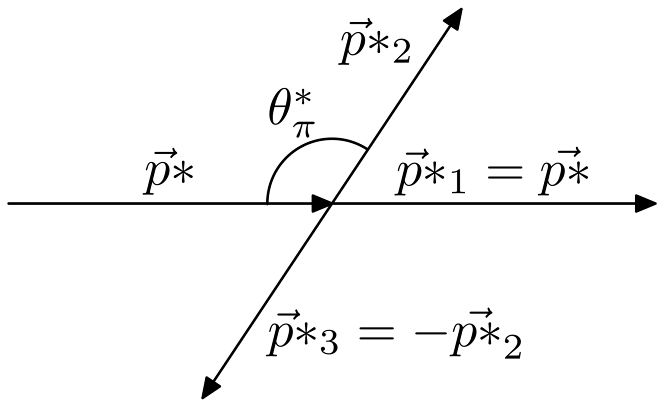


Figure 3.1: Kinematics in the 2-3 COM rest frame

and

$$P^{*2}(= P_1^{*2}) = \frac{(M^2 - (m_{23} + m_1)^2)(M^2 - (m_{23} - m_1)^2}{4m_{23}^2} \overset{DEF}{=} \frac{\lambda(M, m_{23}, m_1)}{4m_{23}^2}. \quad (3.7)$$

To end, let's collect all the possible scalar products:

$$\begin{cases} p \cdot p_1 = \frac{1}{2}(m_1^2 - m_{23}^2 + M^2) \\ p \cdot p_2 = \frac{1}{2}(m_{12}^2 + m_{23}^2 - m_\pi^2 - m_1^2) \\ p \cdot p_3 = \frac{1}{2}(M^2 + m_\pi^2 - m_{12}^2) \\ p_1 \cdot p_2 = \frac{1}{2}(m_{12}^2 - m_\pi^2 - m_1^2) \\ p_1 \cdot p_3 = \frac{1}{2}(M^2 + m_\pi^2 - m_{23}^2 - m_{12}^2) \\ p_2 \cdot p_3 = \frac{1}{2}(m_{23}^2 - 2m_\pi^2) \end{cases} . \quad (3.8)$$

All this work was needed to get the expression for the angle between 1 and 2 in the 2-3 rest frame, which can be read off of the expression for m_{12}

$$\cos(\theta_{12}^*) = \frac{(M^2 - m_{23}^2 - m_1^2)(m_{23}^2 + m_2^2 - m_3^2) + 2m_{23}^2(m_1^2 + m_2^2 - m_{12}^2)}{\sqrt{(M^2 - (m_{23} + m_1)^2)(M^2 - (m_{23} - m_1)^2(m_{23}^2 - (m_2 + m_3)^2)(m_{23}^2 - (m_2 - m_3)^2)}}. \quad (3.9)$$

The variable change $dm_{12}^2 \to d\cos\theta_{12}$ adds the multiplicative factor

$$\frac{\sqrt{(M^2 - (m_{23} + m_1)^2)(M^2 - (m_{23} - m_1)^2(m_{23}^2 - (m_2 + m_3)^2)(m_{23}^2 - (m_2 - m_3)^2)}}{-2m_{23}^2}. \quad (3.10)$$

Up until now the treatment was completely general. Let's now use the (approximate) isospin symmetry, where $m_{\pi^+} = m_{\pi^0} \overset{DEF}{=} m_\pi$.

We also define the angle $\pi - \theta^*_{12} \overset{DEF}{=} \theta_\pi$.

$$\cos(\theta^*_\pi) = \frac{-M^2 - 2m_\pi^2 + 2m_{12}^2 + m_{23}^2 - m_1^2}{\beta\sqrt{(M^2 - (m_{23}+m_1)^2)(M^2 - (m_{23}-m_1)^2)}} = \frac{-M^2 - 2m_\pi^2 + 2m_{12}^2 + m_{23}^2 - m_1^2}{\beta\lambda^{\frac{1}{2}}(M, m_{23}, m_1)}, \tag{3.11}$$

where we defined $\beta = \sqrt{1 - \frac{4m_\pi^2}{m_{23}^2}}$. So we have

$$dLIPS^{(3)} = \frac{\beta\lambda^{\frac{1}{2}}(M, m_{23}, m_1)}{32M^2(2\pi)^3} dm_{23}^2 d\cos\theta_\pi \tag{3.12}$$

if we also set $m_1 = 0$ (massless axion case), we get:

$$\cos(\theta^*_\pi) = \frac{-M^2 - 2m_\pi^2 + 2m_{12}^2 + m_{23}^2}{\beta(M^2 - m_{23}^2)} \tag{3.13}$$

$$dLIPS^{(3)} = \frac{\beta(M^2 - m_{23}^2)}{32M^2(2\pi)^3} dm_{23}^2 d\cos\theta_\pi \tag{3.14}$$

3.1.3 Amplitude

In the isospin approximation, we consider the pions to be identical particles, distinguished only by the I_3 label. Relying on the $|\Delta I = \frac{1}{2}|$ rule, we conclude that the accessible final states have either I = 0 or I = 1. Since $I_3 = 1$ for the $\pi^+\pi^0$ system, it has to be in an isospin triplet. This forces the spatial wavefunction to be anti-symmetric under exchange.
The starting point is the K_{l4} axial vector current matrix element:

$$< \pi^+(p_2)\pi^-(p_3)|\bar{s}\gamma^\mu\gamma^5 u|K^+(p) >= \frac{-i}{M}(F(p_2+p_3)^\mu + G(p_2-p_3)^\mu + Rp_1^\mu), \tag{3.15}$$

where $p_1 = p - p_2 - p_3$ will be interpreted as the axion 4-momentum in our case. This matrix element has to be contracted with $-ip_1^\mu$ because of the derivative coupling of the axion; this means that the last term gives a vanishing contribution in the massless axion approximation.
To connect this matrix element to the one we need for $K \to \pi^+\pi^0 a$, we exploit isospin simmetry.

$$\langle pi^+(p_2)\pi^0(p_3)| \, |\bar{s}\gamma^\mu\gamma^5 d| \, |K^+(p)\rangle = -\sqrt{2}\, \langle(\pi^+(p_2)\pi^-(p_3))_{I=1}| \, |\bar{s}\gamma^\mu\gamma^5 u| \, |K^+(p)\rangle \, , \tag{3.16}$$

the subscript meaning we have to project the dipion system in the triplet isospin configuration.

Let's now take a look at the form factors: we first do a partial wave expansion in the di-pion system, truncated at p-wave and neglecting the s-wave term because it is symmetric under exchange of the pions. $F \approx F_p \cos\theta_\pi e^{-i\delta}, G \approx G_p e^{-i\delta}$. We now Taylor expand at first order around $p_1^2 = 0$ and $\bar{m}_{23}^2 \stackrel{DEF}{=} (m_{23}^2/4m_\pi^2 - 1) = 0$, obtaining

$$F_p = f_p$$
$$G_p = g_p + g_p' \bar{m}_{23}^2$$

Where the lowercase coefficients are to be determined experimentally.
For the third form factor, we cannot rely on experimentally measured value, so we need a theory calculation. According to [32, 33][1]

$$R = \frac{1}{2}(1 + \frac{m_{23}^2 - \beta\lambda^{\frac{1}{2}}(M, m_{23}, P_1)}{p_1^2 - M^2}). \tag{3.17}$$

After dividing by the effective coupling constant F_{sd}^A, we are now ready to square the amplitude. The expressions for two scalar product that we need to do this can be found using the relations we calculated before for the 2-3(i.e. dipion) rest frame.

$$(p_2 + p_3)^\mu p_{1_\mu} = m_{23}E_1 = \frac{M^2 - m_{23}^2 - m_1^2}{2} \xrightarrow{m_1=0} \frac{M^2 - m_{23}^2}{2}$$
$$(p_2 - p_3)^\mu p_{1_\mu} = -2P_2P_1\cos(\pi - \theta_\pi) = \frac{\lambda^{\frac{1}{2}}(M, m_{23}, m_1)\beta}{2}\cos\theta_\pi \xrightarrow{m_1=0} \frac{M^2 - m_{23}^2}{2}\beta\cos\theta_\pi$$

Finally:

$$|\mathcal{A}_{massive}|^2 = \frac{1}{2M^2|F_{sd}^A|^2}[f_p(M^2 - m_{23}^2 - m_1^2)\cos\theta_\pi + \lambda^{\frac{1}{2}}(M, m_{23}, m_1)(g_p + g_p'\bar{m}_{23}^2)\beta\cos\theta_\pi + $$
$$+\frac{1}{2}m_1^2(1 + \frac{m_{23}^2 - \beta\lambda^{\frac{1}{2}}(M, m_{23}, P_1)}{p_1^2 - M^2})]^2 \tag{3.18}$$

$$|\mathcal{A}_{massless}|^2 = \frac{\{(M^2 - m_{23}^2)[f_p + (g_p + g_p'\bar{m}_{23}^2)\beta]\cos\theta_\pi\}^2}{2M^2|F_{sd}^A|^2} \tag{3.19}$$

[1]There seem to be contrasting results in literature: two other published papers ([34] and [35]), each reporting a different result, were found in literature. It should be noted, however, that any difference in the prediction of R will affect the decay rate by a quantity proportional to the small ratio p_1^2/M^2. For this reason, and since the objective of the thesis is a sensitivity study, we found it acceptable to stick with one of the published results, confident that the final result will not be sizeably influenced by this.

3.1.4 Putting it all together

Using equation (3.12) and equation (3.18) into the golden rule we get, for the massive case:

$$\frac{d\Gamma_{massive}}{dm_{23}^2 d(\cos\theta_\pi)} = \frac{\beta\lambda^{\frac{1}{2}}(M, m_{23}, m_1)}{2^{10}\pi^3 M^5 |F_{sd}^A|^2}[f_p(M^2 - m_{23}^2 - m_1^2)\cos\theta_\pi +$$
$$+\lambda^{\frac{1}{2}}(M, m_{23}, m_1)(g_p + g'_p\bar{m}_{23}^2)\beta\cos\theta_\pi + m_1^2(1 + \frac{m_{23}^2 - \beta\lambda^{\frac{1}{2}}(M, m_{23}, P_1)}{p_1^2 - M^2})]^2 \tag{3.20}$$

and for the massless case, using equation (3.14) and equation (3.19)

$$\frac{d\Gamma_{massless}}{dm_{23}^2 d(\cos\theta_\pi)} = \frac{(M^2 - m_{23}^2)^3 \beta\cos^2\theta_\pi [f_p + (g_p + g'_p\bar{m}_{23}^2)\beta]^2}{2^{10}\pi^3 M^5 |F_{sd}^A|^2} \tag{3.21}$$

3.2 Grossman–Nir

A model-independent relation between the decay rates $K^+ \to \pi^+\nu\bar{\nu}$ and $K_L \to \pi^0\nu\bar{\nu}$

$$\frac{BR(K_L \to \pi^0\nu\bar{\nu})}{BR(K^+ \to \pi^+\nu\bar{\nu})} < 4.4 \tag{3.22}$$

was found by Grossman and Nir [36]. This relation only assumes SM isospin breaking, so it can be used to put bounds on any new physics model that maintains the (experimentally well-verified) approximate isospin symmetry. Using the same kind of arguments as in the original paper, we can calculate the analogue relation for the three body decay we are considering: for simplicity let's stick to the massless case.
The main difference between $K^+ \to \pi^+\pi^0 a$ and $K_L \to \pi^0\pi^0 a$ is the fact that in the neutral case the pions are in a I = 0 state, which means that the s-wave form factors have to be used:

$$F_s = f_s + f'_s\bar{m}_{23}^2 + f''_s(\bar{m}_{23}^2)^2$$
$$G_s = 0$$

We also have to take into consideration the different isospin relation: in this case no factor is needed since the Clebsch–Gordan factors cancel out.
The differential rate for $K_L \to \pi^0\pi^0 a$ is (note the lack of dependence on $\cos\theta_\pi$):

$$\frac{d\Gamma}{dm_{23}^2 d(\cos\theta_\pi)} = \frac{(M_L^2 - m_{23}^2)^3 \beta[f_s + f'_s\bar{m}_{23}^2 + f''_s(\bar{m}_{23}^2)^2]^2}{2^{11}\pi^3 M_L^5\ Re(F_{sd}^A)^2} \tag{3.23}$$

The ratio between the partial decay widths is:

$$\frac{\int_{4m_\pi^2}^{M_+^2} d\Gamma_{K_L\pi^0\pi^0 a}(m_{23}^2)}{\int_{4m_\pi^2}^{M_L^2} d\Gamma_{K^+\pi^+\pi^0 a}(m_{23}^2)} =$$

$$= \frac{3M_+^5}{2M_L^5} \frac{\int_{4m_\pi^2}^{M_L^2} dm_{23}^2 (M_L^2 - m_{23}^2)^3 \beta F_s^2}{\int_{4m_\pi^2}^{M_+^2} dm_{23}^2 (M_+^2 - m_{23}^2)^3 \beta (F_p + G_p\beta)^2} \frac{Re(F_{sd}^A)^2}{|F_{sd}^A|^2} < 7.25$$

Which, together with the ratio $\Gamma_+/\Gamma_L \approx 4.17$, gives us:

$$\frac{BR(K_L \to \pi^0\pi^0 a)}{BR(K^+ \to \pi^+\pi^0 a)} < 31 \tag{3.24}$$

This relation gives us a motivation to consider charged kaon decays instead of K_L decays: a measurement/upper limit on the BR of the charged decay puts automatically an upper limit on the BR of the correspondent K_L decay, while the converse is not true.

Chapter 4

Low-Energy Limit of QCD

The theory of the strong interaction as an asymptotically-free SU(3) gauge theory implies the non-validity of perturbation theory at low energy. Indeed the β function for the strong coupling constant is positive, and we even have the existence of a Landau pole at finite, low energies, meaning the complete breakdown of perturbative methods. This means we have to resort to different schemes to describe the dynamics of the strong interaction at low (i.e. comparable or lower to ~ 1 GeV) energies. One way to do this is in the framework of Effective Field Theories (EFT).
In this chapter, we will see the low-energy limit of QCD, and how the axion is added to this effective theory. This is one possible framework used to calculate decay rates of Kaons; indeed it was used to derive the form factor in equation (3.17).

4.1 Effective Field Theories

The basis of this approach is the Wilsonian renormalisation, which assures us that - in a given energy range - one can consider only a subset of the degrees of freedom (*dof* for short) of the full theory, and integrate out the remaining dof of the theory, their dynamics frozen into the coefficient of local operators. The experimentally well-known spectrum of QCD at low energies ensures us that such an approach is sensible: there are only a few particles, an octet of light pseudoscalars called mesons, which interact weakly among themselves and with nucleons. In a similar way to the method of "quasiparticles" for non-relativistic theories, we may be able to get a perturbatively tractable potential if we consider the bound states as elementary, i.e. we integrate out the fast dof of the elementary particles of the full theory, remaining only with the slow dof of the bound states. We can then envision a theory where the mesons are the

dof entering the Lagrangian; this theory should be perturbatively tractable even at low energies, because of the weak interaction between its degrees of freedom. Chiral Perturbation Theory (χPT) is the theory that does this. Since we know that at high energies QCD is perturbative, hence the correct degrees of freedom to consider are partons, this theory must necessarily have a cut-off in energy, in accord with the fact that it is a low-energy effective description of a more general theory.
The general prescription of an EFT is to write down a Lagrangian containing all possible terms compatible with the (low-energy) symmetries of the underlying theory[37]. This means writing down an infinite number of terms and their respective coefficients, of growing mass dimension. They are organised in term of powers of a small parameter, allowing perturbative description. The kind of field theories most apt for a description like this are those with a wide gap in their energy spectrum: according to renormalisation group wisdom, we can then divide the dof in low-energy and high-energies one, the latter delimited by a cut-off energy Λ. If we are interested in describing processes at energies much lower than Λ, we can use the p/Λ as a perturbative parameter, since it is guaranteed to be small. What we now need is a method to organise all terms appearing in the Lagrangian according to powers of p/Λ, including terms not explicitly featuring power of momentum. Being a perturbative theory, we can calculate observable in terms of Feynman diagrams: we then need to assign a power counting to them, too. Once we have done this, we can calculate observables up to a precision depending on the power n at which we stop the perturbative series (for short, we will talk about $\mathcal{O}(p^n)$ precision). The perturbative series is now two-fold: going from one order of perturbation to the next, we need to consider both the usual series in term of loops, and the fact that new operators appear in the Lagrangian at each order, too.
This program is rendered much easier in theories where the low-energy description is highly constrained by symmetry, as for theories which undergo Spontaneous Symmetry Breaking (SSB).

4.2 Spontaneous Chiral Symmetry Breaking

The properties of Goldstone bosons (reviewed in section 1.1.1), together with the experimental evidence on the spectrum of QCD, suggest to construct a theory of mesons where they are Goldstone bosons of a $SU(3)_L \otimes SU(3)_R \to SU(3)_V$ symmetry breaking. Indeed, describing the pseudoscalar mesons as Goldstone bosons easily accounts for their lightness and the size of their interaction, while having the chiral group be broken explains the absence of scalar mesons at the same energy scale. At this stage, the number 3 is chosen

purely because we need 8 broken generators, which is exactly the number of generator of $SU(3)$.
This proposition is made even more interesting by the observation that the massless QCD Lagrangian in the light sector ($q^T = (u, d, s)$):

$$\mathcal{L}_Q^0 CD = -\frac{1}{4} G^a_{\mu\nu} G^{\mu\nu}_a + i\bar{q}_L \gamma^\mu D_\mu q_L + i\bar{q}_R \gamma^\mu D_\mu q_R, \tag{4.1}$$

where $q_{L/R} = \frac{1}{2}(1 \pm \gamma_5)q$ are the left/right handed quark fields, is indeed invariant under the group transformation $q_L \to U_L q_L$, $q_R \to q_R U_R$. Since it is known that the c, b, t quarks have much higher masses than the remaining three, this is the right Lagrangian to start from when considering the low-energy limit of QCD, neglecting the light quark masses and considering the heavy quarks as frozen. More confirmation that we are on the right path comes from the fact that there is no light octet of scalar mesons: if the low-energy theory was also — at least approximately — invariant under the full symmetry group, we would expect to see it. The plausible explanation to these facts is that spontaneous symmetry breaking occurred.
We can now set to find the operators $\mathcal{O}^a$ appearing in the statement of Goldstone theorem: we know that the conserved current and the $\mathcal{O}^a$ need to share quantum numbers with the Goldstone bosons, otherwise they could not have non-vanishing vacuum expectation value. The easiest set of 8 pseudoscalar operators we can construct is $\mathcal{O}^a = \bar{q}\gamma_5 \lambda^a q$, with λ^a the Gell-Mann matrices, i.e. the generators $SU(3)$, making it a rather natural choice. The order parameter is[1]

$$\langle 0 | \left[\lambda^a_L - \lambda^a_R, \mathcal{O}^b \right] |0\rangle = \langle 0 | [\gamma_5 \frac{\lambda^a}{2}, \bar{q}\gamma_g \lambda^a q] |0\rangle = \frac{2}{3}\delta^{ab} \langle 0 | \bar{q}q |0\rangle \tag{4.2}$$

and is called the (scalar) *quark condensate*.
This choice is not the only possible one [33]: other order parameters give rise to slightly different realisation of low-energy QCD. However in so-called "standard" formulation of the chiral Lagrangian this is the usual choice, which is justified by a range of considerations and evidence pointing to quark condensate being, if nothing else, the dominant contribution to symmetry breaking [38].

4.3 The coset formalism

Before going further, let's introduce an importance piece of formalism which will allow us to effortlessly derive the correct Lagrangian describing the

[1]The γ_5 in the definition of $\lambda^a_R - \lambda^a_L$ can be understood looking at the action of $SU(2)_R \otimes SU(2)_L$ on the $q_{L/R}$ fields

Nambu-Goldstone bosons at low energy: to do so we use a slight generalisation of the toy model above. Consider the O(N)-invariant Lagrangian

$$\mathcal{L} = \frac{1}{2}\partial^\mu \Phi \partial_\mu \Phi^T - \frac{\lambda}{4}(\Phi\Phi^T - v^2)^2, \tag{4.3}$$

where $\Phi(x)^T = (\varphi_1, \dots, \varphi_N)$.
It has vacuum configurations satisfying $\Phi\Phi^T = v^2$, i.e. the vacua lie on a N-1 dimensional sphere S^{N-1}. We can then freely choose $\Phi_0^i = \delta^{iN} v$, for example, which explicitly shows the invariance under N-1-dimensional rotation. Thus we have $O(N) \to O(N-1)$ SSB, and N-1 corresponding broken generator. We can again take polar coordinates:

$$\Phi(x) = (1 + \frac{S(x)}{v})U(x)\Phi_0, \tag{4.4}$$

the meaning of which should by now be clear by comparison with the 3-dimensional model: the $S(x)$ dof is a massive radial excitation and $U(x)$ contains the Nambu-Goldstone bosons, which can be thought of as co-ordinates on S^N and which are massless. This can be visualised as rotations on the sphere with no energy cost, while $S(x)$ feels a potential giving it a mass $M^2 = 2\lambda v^2$.
We take $U(x) = exp(i\tau^a \varphi_a(x)/v)$ in exponential form and consider what happens under a $O(N)$ transformation: we get a new point on S^{N-1}, but we have to account for the rotation of the N-th axis, too:

$$gU(x) = U'(x)h(g, U) \tag{4.5}$$

with $U'(x)$ in the canonical exponential form and $h(g, U)$ a compensating transformation, which is non-linear and belongs in $O(N-1)$. gU and U' then describe the same configuration of Nambu-Goldstone fields, but with a different choice of co-ordinate over the sphere (equivalently, with a rotated vacuum).
We can now create a mapping which takes a member g of $O(N)$ and a field configuration $U(x)$ and outputs another field configuration, corresponding to U after the action of g. Let's call the mapping $\mathcal{F}(g, U)$: under group transformations we have

$$U(x) \to \mathcal{F}(g, U(x)). \tag{4.6}$$

We require that $\mathcal{F}$ satisfies group axioms (closure, existence of identity and inverse). We also require that for any $h \in O(N-1)$ the vacuum is left invariant, as it should.
Let's now consider the action on elements the left coset

$$gO(N-1) = gh|h \in O(N-1) : \tag{4.7}$$

we have that all elements of the same coset map the vacuum on the same field configuration

$$\mathcal{F}(gh, \Phi_0) = \mathcal{F}(g, h\Phi_0) = \mathcal{F}(g, \Phi_0), \tag{4.8}$$

which is true because of group axioms. Finally, the mapping is injective with respects to the cosets: consider $g \in O(N)$ and $g' \not\in gO(N)$ assume $\mathcal{F}(g, \Phi_0) = \mathcal{F}(g', \Phi_0)$ we need

$$\Phi_0 = \mathcal{F}(g^{-1}g, \Phi_0) = \mathcal{F}(g^{-1}, \mathcal{F}(g, \Phi_0) = \mathcal{F}(g^{-1}, \mathcal{F}(g', \Phi_0), \tag{4.9}$$

which implies $g' \in gH$, in contradiction.
What we found is then a isomorphism between the space of cosets: $O(N)/O(N-1) = gO(N-1)|g \in O(N)$. We can then trade group transformation on fields for group transformations on members of the quotient space. Then equation (4.5) is just the statement that under group transformation acting on a coset representative we do not necessarily obtain another coset representative, but we need a corrections by a non-linear function.
We can now apply this to the case of $SU(3)_L \otimes SU(3)_L \to SU(3)_V$: the left cosets are represented by $U = (1, RL^\dagger)$, with R/L belonging to $SU(R/L)$: under the action of $(\hat{R}, \hat{L}) \in SU(3) \otimes SU(3)$ it gets transformed to

$$U' = \hat{R}U\hat{L}. \tag{4.10}$$

The correspondent fields are parametrised in polar coordinates:

$$U = e^{i\frac{\sqrt{2}\Phi}{F_0}} \quad \Phi = \lambda_a\phi^a(x) = \begin{pmatrix} \frac{\pi^0}{\sqrt{2}} + \frac{\eta}{\sqrt{6}} & \pi^+ & K^+ \\ \pi^- & -\frac{\pi^0}{\sqrt{2}} + \frac{\eta}{\sqrt{6}} & K^0 \\ K^- & \bar{K}^0 & -\frac{2\eta}{\sqrt{6}} \end{pmatrix}, \tag{4.11}$$

where F_0 is a constant with dimension of energy, since scalar fields have energy dimension of 1.
The vacuum is given by $\Phi_0 = 0 \to U_0 = 1$. The action of the unbroken subgroup $SU(3)_V = \{(V, V)|V \in SU(3)\}$ leaves it invariant, as expected. Meanwhile, the broken subgroup is $SU(3)_A = \{(A, A^+)|A \in SU(3)\}$, whose action does not leave the vacuum invariant, also as expected. Finally, we can see the transformation behaviour of the Goldstone bosons under the unbroken group:

$$U = 1 + i\frac{\sqrt{2}\Phi}{F_0} - \frac{\Phi^2}{F_0^2} + \ldots \to 1 + i\frac{V\Phi V^\dagger}{F_0} - \frac{V\Phi^2 V^\dagger}{2F_0^2} + \ldots \tag{4.12}$$

or, parametrising V as $\exp(-i\theta_a\lambda^2/2)$

$$\Phi = \lambda_a\phi^a \to \Phi - i\theta_a[\frac{\lambda_a}{2}, \lambda_b\phi^b] + \ldots = \Phi + f_{abc}\theta_a\phi_b\lambda_c + \ldots \tag{4.13}$$

where f_{abc} are SU(3) structure constant. This means that the Goldstone bosons transform as an octet.
If one were to try the same with $SU(3)_A$ action, one would find a complicated, non-linear function, which is consistent with SSB.

4.4 The Chiral Lagrangian

We are now in a position to write down the low-energy Lagrangian of 3-flavour massless QCD. To do so, we organise the Lagrangian in increasing order of momenta: owing to parity conservation of QCD we can disregard odd power of momenta:

$$\mathcal{L}_\chi^0 = \sum_n \mathcal{L}_{2n}. \tag{4.14}$$

The 0-th order Lagrangian can only amount to a constant: the Goldstone field is unitary by virtue of it being identified with a member of $SU(3)_V$, so the only hermitian we can construct is $Tr(U^\dagger U) \propto 1$. Going up to 2 derivatives, there is only one Lorentz invariant Hermitian operator we can construct that is invariant under $SU(3)_V$: $Tr(\partial^\mu U \partial_\mu U^\dagger)$. Then we have[39]

$$\mathcal{L}_\chi^0 = \frac{F_0^2}{4} Tr(\partial^\mu U \partial_\mu U^\dagger) + \mathcal{O}(\frac{p^4}{\Lambda_\chi^4}), \tag{4.15}$$

The coefficient in front fixed by requiring canonical kinetic term $\frac{1}{2}\partial^\mu \Phi \partial_\mu \Phi$.
We can now describe, to an accuracy $\mathcal{O}(p^2)$, all processes involving the Goldstone bosons, up to an unknown constant F_0. We will return to the determination of F_0 later, once we have the machinery to do. First, we have a more pressing matter at hand: physical mesons are not in fact massless particles. This can be traced back to the fact we neglected quark masses, which break chiral symmetry.

4.4.1 Explicit Chiral Symmetry Breaking

The full QCD Lagrangian can be divided into the sum of a chiral-conserving and a chiral-breaking term:

$$\mathcal{L}_{QCD} = \mathcal{L}_{QCD}^0 + \mathcal{L}_{QCD}^{\not\chi}, \tag{4.16}$$

the chiral breaking part containing the quark masses $\bar{q}\mathcal{M}q$ and the coupling to EW gauge fields.
A powerful method to account for symmetry breaking terms is the method of

external fields: we write the chiral-breaking term using classical (i.e. only appearing in external legs), non-dynamical fields:

$$\mathcal{L}^{\not\chi}_{QCD} = \bar{q}\gamma^{\mu}(v_{\mu} + \gamma_5 a_{\mu})q - \bar{q}(s - i\gamma p)q, \tag{4.17}$$

where (v_{μ}, a_{μ}, s, p) transform as a vector/axial vector/scalar/pseudoscalar under Lorentz transformation. The SM is then formally recovered through the following identifications:

$$\begin{cases} -e\mathcal{Q}A_{\mu} = v_{\mu} + a_{\mu} \stackrel{DEF}{=} r_{\mu} \\ -e\mathcal{Q}A_{\mu} - \frac{e}{\sqrt{2}\sin\theta_W}(W^{\dagger}_{\mu}T_{+} + h.c.) = v_{\mu} - a_{\mu} \stackrel{DEF}{=} l_{\mu} \\ \mathcal{M} = s \end{cases} \tag{4.18}$$

where all the symbols have their usual meaning: $\mathcal{Q}$ is diagonal and contains the quark charges in unit of e, A_{μ} , W_{μ} are gauge boson fields, and θ_W is the Weinberg angle. As for T_{+}, it incorporates the quark mixing factors:

$$T_{+} \begin{pmatrix} 0 & V_{ud} & V_{us} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{4.19}$$

We are neglecting Z-boson and Higgs coupling, which would enter l_{μ} and s, respectively, as they are not relevant to our purposes.
The next piece is to enforce (formal) invariance under gauged chiral group transformation, by having the external fields transform as:

$$\begin{cases} s + ip \to g_R(s + ip)g_L^{\dagger} \\ l_{\mu} \to g_L l_{\mu} g_L^{\dagger} + ig_L\partial_{\mu}g_L{}^{\dagger} \\ r_{\mu} \to g_R r_{\mu} g_R^{\dagger} + ig_R\partial_{\mu}g_R{}^{\dagger} \end{cases} \tag{4.20}$$

in the spirit of the spurion method described in section 1.1.1. There is of course little physical content in, for example, having the quark mass matrix transform under gauge transformation, this is merely a bookkeeping device. Continuing to enforce local invariance, we construct the covariant derivatives: this will quickly give us the correct coupling of matter and gauge fields:

$$D_{\mu}U = \partial_{\mu}U + iUl_{\mu} - ir_{\mu}U. \tag{4.21}$$

One could also construct the field strength tensor, but they do not appear in the low-energy at $\mathcal{O}(p^2)$.
Following the same prescription as above, we construct the LO Lagrangian with explicit chiral symmetry breaking. We can keep the derivative term,

augmenting the derivative to the covariant one. Looking for other operators satisfying Lorentz and Chiral invariance, we find only $(s - ip)U$ and its hermitian conjugate.

$$\mathcal{L}_\chi = \frac{F_0^2}{4} Tr(D^\mu U D_\mu U^\dagger) + \frac{F^2}{2} Tr(U^\dagger \mathcal{M} + \mathcal{M} U), \tag{4.22}$$

where we already made use of the definition of s we gave above. This Lagrangian, when the spurions are frozen at their physical value, breaks of course the chiral symmetry explicitly, but it does so in exactly the same way $\mathcal{L}_{QCD}$ does.
The coefficient in front of the new term was chosen with a bit of foresight: if we expand U up to second order in the field, we can read the mass term for the mesons:

$$- B_0 Tr(\mathcal{M}\Phi^2), \tag{4.23}$$

so that the meson masses are linked to the quark masses by the relations:

$$\begin{aligned} M_\pi^2 &= 2\hat{m} B_0 \\ M_K^2 &= (\hat{m} + m_s) B_0 \\ M_\eta^2 &= \frac{2}{3}(\hat{m} + 2m_s) B_0 \end{aligned} \tag{4.24}$$

where we assumed isospin symmetry for simplicity and set $\hat{m}$ as the mean between m_u and m_d.
Before going further it is crucial to notice that we are making one crucial assumption, closely related to the assumption of "standard" χPT covered above. As stressed before, when adding external fields to an EFT one must specify their p scaling: here we tacitly assumed that the vector fields count as $\mathcal{O}(p)$ and the scalar field counts as $\mathcal{O}(p^2)$. The first one follows from the definition of the covariant derivative, but the second one is not on as secure footing: why should the mass matrix count as order p^2? The obvious reason is the relations we derived above for the meson masses, along with the on-shell condition $p^2 = M^2$; there is however the proportionality constant B_0: taking the first derivative with respect to quark masses in both $\mathcal{L}_{QCD}$ and $\mathcal{L}_\chi$, we find

$$B_0 = -\frac{1}{3F_0^2} \langle 0| \bar{q}q |0\rangle \,. \tag{4.25}$$

So this power counting hinges on the hypothesis that the quark condensate is non-vanishing, and that it provides the (dominant) mechanism for SSB of the chiral symmetry. If this is not realised in nature, the dominant contribution for meson masses would not be linear in the quark masses. At any rate, within

this framework we can verify all pre-QCD relations between mesons masses, e.g. Gell-Mann–Okubo $4M_K^2 = 3M_\eta^2 + M_\pi^2$, independent of the value of B_0. This lends credence to the program.
In principle, one could calculate the quark condensate from first principle QCD (e.g. using lattice gauge theories) and fix the value of B_0. In practice, the measured values of the meson masses are used as inputs and to check internal consistency.
After clarifying the physical origin of B_0, we return to the other parameter, F_0.

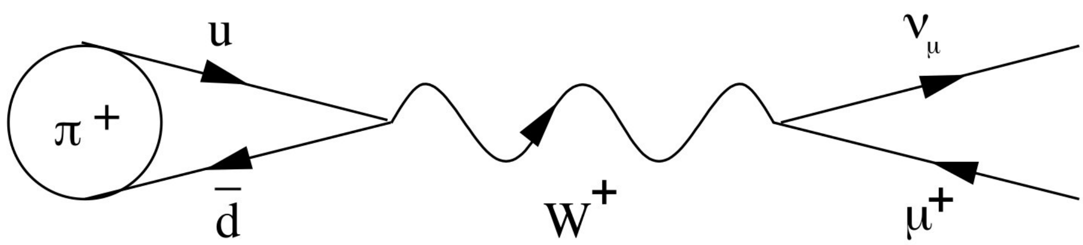


Figure 4.1: Feynman Diagram for $\pi^+ \to \mu^+\nu$ decay. Picture from [40]

It should also in principle be calculable from the full theory, but this is exactly the kind of calculation we want to avoid. Instead, we use one well-known process to fix the value of F_0; once we have done this we can use it to predict any other process involving mesons and (classical) gauge bosons, checking the internal consistency of the theory.
The process we can consider is pion decay $\pi^+ \to \mu^+\nu_\mu$. The corresponding Feynman diagram figure 4.1 can be factorised as $\pi^+ \to W^+ \to \mu^+\nu\mu$, with virtual W^+. Only the first vertex is related to QCD: to find the Feynman rule we expand equation (4.22) and write down the interaction between one positive pion and one W^+:

$$\mathcal{L}_\chi \supset \frac{-g}{\sqrt{2}}\frac{F_0}{2}\langle W^{+\mu}T^+\partial_\mu\Phi\rangle \supset -g\frac{F_0}{2}V_{ud}\partial_\mu\pi^-. \tag{4.26}$$

This needs to be contracted with the W^+ propagator (which gets simplified since $p \ll M_W^2$ in order for our approach to make sense) and the $W^+ \to \mu\nu_\mu$ Feynman rule, yielding the decay rate (see e.g. [41] for the - rather standard - derivation)

$$\Gamma = \frac{G_F^2|V_{ud}|^2}{4\pi}F_0^2 M_\pi m_\mu^2(1-\frac{m_\mu^2}{M_\pi^2})^2. \tag{4.27}$$

The measured value for the rate yields $F_0 = 92.4$ MeV. F_0 is then called the pion decay constant in the chiral limit: we performed the calculation at $O(p^2)$, but this is the same value it would have in the chiral limit, since F_0 does not get modified by the presence of external fields.

4.4.2 Power Counting

One of the cornerstones of the EFT approach is the fact that we can organise all terms appearing in calculations in powers of p/Λ: part of this was achieved in last section, when we determined the power counting properties of the external field. We also need to study the properties of Feynman diagrams; the standard argument is due to Weinberg [37], who considered the scaling of loop diagrams under momentum rescaling. The resulting theorem is that a connected diagram has a scaling dimension of p^{2d_Γ}, d_γ defined as:

$$d_\Gamma = 1 + N_L + \sum_{n=1}^{n=+\infty} (n-1) N_{2n}, \tag{4.28}$$

where N_L is the number of (independent) loops and N_{2n} is the number of vertices originating from $\mathcal{L}_{2n}$.
This means that at order $\mathcal{O}(p^2)$ we only need to consider tree-level diagrams, while at $\mathcal{O}(p^4)$ we have both tree-level diagrams involving vertices from $\mathcal{L}_4$ and loop diagrams with vertices of the lower order.
This argument proves that it is indeed possible to consistently perform a perturbative expansion in orders of p/Λ_χ. The final ingredient we need is the appropriate scale for χPT.

4.4.3 Estimating Λ_χ

Armed with the knowledge of how loop diagrams work, we can now reproduce an argument due to Manohar and Georgi[39][2] for what ought to be the scale we employ in χPT.
Consider $\pi - \pi$ scattering at one-loop level: expanding $\mathcal{L}_\chi$ we find the needed 4-meson vertices. One particular vertex has derivative acting on all the 4 external momenta (each vertex contributing $\sim p^2/F_0^2$). The amplitude is then

$$\mathcal{A} \sim \frac{p^4}{F_0^4} \int_\mu^{\Lambda_\chi} \frac{d^4k}{(2\pi)^4} \frac{1}{k^4} = \frac{p^4}{F_0^4} \ln(\frac{\Lambda_\chi^2}{\mu^2}), \tag{4.29}$$

where we are evaluating the integral using Λ_χ as a cut-off, in the spirit of EFT, and μ^2 is some renormalisation scale.

[2]For a more modern perspective on the matter, see [42]

The argument then follows noticing that a similar term is produced by the $\mathcal{O}(p^4)$ term

$$\frac{F_0^2}{\Lambda_\chi} Tr(\partial_\mu \Sigma \partial^\nu \Sigma \partial^\mu \Sigma^\dagger \partial_\nu \Sigma^\dagger), \tag{4.30}$$

but then order-by-order renormalisability requires the coefficient in front of this to be larger than the coefficient we would get by changing the renormalisation scale μ in the integral above:

$$\Lambda_\chi \lesssim 4\pi F_0. \tag{4.31}$$

The assumption of naive dimensional analysis is that this inequality is in fact saturated, yielding $\Lambda_\chi \approx 1.2$ GeV.
There is also another scale we should keep in mind when making calculations in χPT: as usual in any QFT, resonances can induce large anomalous dimensions through s-channel exchange. The first particle in the QCD spectrum with the correct quantum numbers is the ρ meson, with a mass of approximately 1 GeV. Thus χPT needs corrections to account for this, even before reaching Λ_χ; the theory that does so is Resonance chiral perturbation Theory (RχT)[43]. RχT is by now well developed, but way outside the scope of this thesis.

4.5 Chiral Lagrangian with the axion as light degree of freedom.

We start from the most general interaction of a pseudoscalar pNGB with QCD fields, retaining only operators of dimension ≤ 5:

$$\mathcal{L}_a = \frac{1}{2}\partial_\mu a \partial^\mu a - \frac{{m_a^0}^2}{2}a^2 + c_{GG}^0 \frac{\alpha_s}{4\pi}\frac{a}{f}G_{\mu\nu}\tilde{G}^{\mu\nu} + c_{\gamma\gamma}\frac{\alpha}{4}\frac{a}{f}F_{\mu\nu}\tilde{F}^{\mu\nu} + c_{GG}^0\frac{\partial_\mu a}{f}\bar{q}^i\gamma^\mu({C^{V}}^{0ij} + \gamma_5 {C^{A}}^{0ij})q^j. \tag{4.32}$$

The decay constant f is related to the scale of PQ breaking by the usual relation $\Lambda_{PQ} = 4\pi f$. Note that elsewhere the definitions $f_a = -\frac{f}{2c_{GG}}$ and $g_{a\gamma\gamma} = \frac{c_{\gamma\gamma}}{\pi f}$ are used for the gauge boson couplings.
Since our goal is to perform a low-energy matching to meson fields, we rotate away the gluon coupling in favour of the quark bilinears. This is obtained by the axion-dependent rotation:

$$q \to \exp(-ic_{GG}^0\frac{a}{f}(\delta_V + \gamma_5\delta_A))q, \tag{4.33}$$

where $\delta_{V/A}$ are matrices in flavour space. Through a Fujikawa analysis this can be shown to also transform the couplings to [44]:

$$\begin{aligned} c_{GG} &= 0 \\ c_{\gamma\gamma} &= c_{\gamma\gamma}^0 - 6c_{GG}^0 Tr(\delta_A Q^2) \end{aligned} \tag{4.34}$$

Where Q is the diagonal matrix containing the electrical charges (in unit of e) of the quarks and the first equation is valid as long as $< \delta_A >= 1$.
The quark bilinears also get modified:

$$\begin{aligned} \bar{q}\gamma^\mu({C^V}^0 + \gamma_5 {C^A}^0)q &\to \bar{q}\gamma^\mu e^{ic_{GG}^0\frac{a}{f}(\delta_V+\gamma_5\delta_A)}({C^V}^0 + \gamma_5{C^A}^0)e^{-ic_{GG}^0\frac{a}{f}(\delta_V+\gamma_5\delta_A)}q \\ i\bar{q}\gamma^\mu\partial_\mu q &\to i\bar{q}\gamma^\mu\partial_\mu q + c_{GG}^0\frac{\partial_\mu a}{f}\bar{q}\gamma^\mu(\delta_V + \gamma_5\delta_A) \\ M_q^0\bar{q}q &\to e^{-2i\frac{a}{f}\gamma_5\delta_A c_{GG}^0}M_q^0\bar{q}q \end{aligned} \tag{4.35}$$

That is, we can define new coupling:

$$\begin{aligned} C^V &= e^{ic_{GG}^0\frac{a}{f}(\delta_V+\gamma_5\delta_A)}{C^V}^0 e^{-ic_{GG}^0\frac{a}{f}(\delta_V+\gamma_5\delta_A)} + \delta_V c_{GG}^0 \\ C^A &= e^{ic_{GG}^0\frac{a}{f}(\delta_V+\gamma_5\delta_A)}{C^A}^0 e^{-ic_{GG}^0\frac{a}{f}(\delta_V+\gamma_5\delta_A)} + \delta_A c_{GG}^0 \\ M_q &= e^{-2i\frac{a}{f}\gamma_5\delta_A c_{GG}^0}M_q^0 \end{aligned} \tag{4.36}$$

To find the effective Lagrangian below the Chiral scale, we start from the χPT Lagrangian at $\mathcal{O}(p^2)$:

$$\mathcal{L}^{(2)}_{\chi PT} = \frac{F_0^2}{4} Tr(D_\mu U D^\mu U^\dagger) + \frac{F_0^2 B_0}{2} Tr({M_q}^\dagger U + M_q U^\dagger). \tag{4.37}$$

The rotation on the quarks introduces tree-level mixing between neutral mesons and the axion, which could be removed by choosing $\delta_A M_q \propto 1$. Together with the trace condition this fixes δ_A to be diag($\kappa_u, \kappa_d, \kappa_s$)[45]:

$$\begin{aligned}\kappa_u &= \frac{m_d m_s}{m_u m_d + m_u m_s + m_d m_s} \approx \frac{m_d}{m_u + m_d}\\ \kappa_d &= \frac{m_u m_s}{m_u m_d + m_u m_s + m_d m_s} \approx \frac{m_u}{m_u + m_d}\\ \kappa_s &= \frac{m_u m_d}{m_u m_d + m_u m_s + m_d m_s} \approx \frac{m_u m_d}{m_s(m_u + m_d)}\end{aligned} \tag{4.38}$$

Which also fixes the photon coupling.

$$\hat{c}_{\gamma\gamma} = c^0_{\gamma\gamma} - \frac{2}{3}\frac{4m_d m_s + m_u(m_s + m_d)}{m_u m_d + m_u m_s + m_d m_s} c_{GG} \approx c^0_{\gamma\gamma} - 2c^0_{GG}. \tag{4.39}$$

Higher order calculation change the coefficient in front of c^0_{GG} from 2 to $\sim$1.92[46]. The photon coupling is then generally expected to be non-vanishing in most models, barring fine-tuning.
One can also avoid making this choice, keeping the rotation matrices free to show explicitly how they cancel when computing observables.
We also need the definition of the covariant derivative on the meson matrix. To be consistent, other than the usual term for the electromagnetic and charged weak interaction, we need to add a term for the axion, as well[47]:

$$D_\mu U = \partial_\mu U - ieA_\mu[Q, U] - iUl_\mu - i\frac{\partial_\mu a}{f}([C_V, U] + \{C_A, U\}) \tag{4.40}$$

Where

$$\begin{aligned} l_\mu &= \frac{-g}{\sqrt{2}}(W^+_\mu T^+ + W^-_\mu T^-)\\ T^+ = T^{-T} &= \begin{pmatrix} 0 & V_{ud} & V_{us}\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}\end{aligned}$$

The third term causes the appearance in the Lagrangian of a term proportional to a left-handed $J^\mu_L = J^\mu_V - J^\mu_A$ current which can be found e.g. with the external fields method of section 4.4.1 [41]:

$$J_V^\mu = \frac{iF_0^2}{2}(D^\mu U U^\dagger + D^\mu U^\dagger U) = \frac{iF_0^2}{2}[D^\mu U, U^\dagger] = i[D^\mu\Phi, \Phi] - \frac{i}{6F_0^2}[[[\partial^\mu\Phi, \Phi], \Phi], \Phi] + \mathcal{O}(\frac{\Phi^6}{F_0^4})$$
$$J_A^\mu = \frac{iF_0^2}{2}\{D_\mu U, U^\dagger\} = -\sqrt{2}F_0 D^\mu\Phi + \frac{\sqrt{2}}{3F_0}\left[[D^\mu\Phi, \Phi], \Phi\right] + O(\frac{\Phi^5}{F_0^3}) \tag{4.41}$$

The second equality comes from $\partial^\mu(UU^\dagger) = 0 \Rightarrow \partial^\mu U^\dagger U = -U^\dagger\partial^\mu U$ and the third from expanding U. Notice that the Vector current only has terms that are even in the number of Goldstone bosons, while the axial current contains only odd terms.
Similarly, the third induces a term proportional to the both the vector and the axial-vector current[48].
There is still one term left to consider, related to flavour-violation in loop diagrams.

Loop-Induced Flavour Violation Since the goal is to describe $|\Delta S| = 1$ transitions, we need to introduce weak currents as well. For (semi-)leptonic decays, where the W^+ is contracted only once with the hadronic operators, we can treat it as an external field and it has already been taken care of by the covariant derivative.
For processes where the weak current is contracted with two quark currents (i.e. a loop process, at the quark level), the situation gets complicated by the presence of QCD operators renormalising the weak vertices. At leading order, this amount to adding gluon exchange to the Fermi interactions Feynman diagram see figure 1.6. RG equations are then run from $\mathcal{O}(M_W)$, to $\mathcal{O}(1GeV)$, the threshold for perturbative QCD. Finally, applying the standard EFT procedure of matching high-energy operators to low-energy one, one gets the effective theory of $|\Delta S| = 1$ transitions.
The result of this program at $\mathcal{O}(p^2)$ is to add the following interaction lagrangian [49]:

$$\mathcal{L}_W = F_0^4(G_8 W_8 + G_{27} W_{27} + G_{\underline{8}} W_{\underline{8}}) \tag{4.42}$$

Where the subscripts indicate the representation under $SU(3)_L \times SU(3)_R$: respectively $(8, 1)$, $(27, 1)$ and $(8, 8)$ [50]. Indeed for example a method to get the shape of W_8 is a spurion analysis of the QCD and χPT lagrangian under left-handed rotations, with the fields transforming as vectors ($q \to U_l q$, $\Phi \to U_L\Phi$ respectively).

$$\begin{aligned} W_8 &= Tr(\Lambda D^\mu U D_\mu U^\dagger) \\ W_{27} &= T^{ij}_{kl}(UD_\mu U^\dagger)^k_i(UD^\mu U^\dagger)^l_j \\ W_{\underline{8}} &= Tr(\Lambda UQU^\dagger) \end{aligned} \tag{4.43}$$

Where for the purpose of $s \to d$ transitions $\Lambda_{ij} = \delta_{is}\delta_{ju}$ and $T^{ds}_{uu} = 1, T^{du}_{us} = 2/3$. The last operator turns out to have negligible effect [51] so we neglect it. The coefficients of the remaining two operators can be extracted from $K \to 2\pi$ decays: a usual parametrisation in terms of dimensionless coefficient is [52]

$$G_{8/27} = -\frac{G_F}{\sqrt{2}} V_{ud} V^*_{us} g_{8/27}. \tag{4.44}$$

Matching with data one gets $g_8 \approx 5.0$, $g_{27} \approx 0.29$. The mismatch between the octet and the 27-plet is known as $\Delta I = 1/2$ rule because the octet induces only this kind of transition, while the 27-plet can also induce $\Delta I = 3/2$ transitions. Because of this rule, we will ignore the 27-plet, too. It's important to note that this can't be always done, e.g. when calculating the $K^+ \to \pi^+\pi^0$ rate, which proceeds uniquely through the $|\Delta I = 3/2|$ operator in the approximation of equal pion masses. However, for the off-shell decay $K^+ \to \pi^+(\pi^0 \to a)$, the octect current is non-vanishing also in the isospin limit, as we are going to explicitly calculate. Hence the two-body decay with the axion is enhanced with respect to the SM one [48, 53].
The operator W_8 is the chiral representation of an effective 4-quark interaction, so we need to take into account the quark rotations, just as for the mass term[48]:

$$(W_8)^{ij} = \frac{F_0^4}{16} e^{2ic^0_{GG}\left(\delta^i_V - \delta^j_V - \gamma_5(\delta^i_A - \delta^j_A)\right)\frac{a}{f}} (D^\mu U D_\mu U^\dagger)^{ij} \stackrel{def}{=} \frac{F_0^4}{16} e^{2ic^0_{GG}(\Delta^{ij}_V - \gamma_5\Delta^{ij}_A)\frac{a}{f}} (D^\mu U D_\mu U^\dagger)^{ij}. \tag{4.45}$$

The effective Lagrangian including interaction with the axion then is

$$\begin{aligned}
\mathcal{L}^{(2)}_{\chi PT-a} &= Tr(\frac{F_0^2}{4} D_\mu U D^\mu U^\dagger + \frac{F_0^2 B_0}{2}(M_q^\dagger U + U^\dagger M_q) + G_8 F_0^4 W_8) = \\
&= Tr(\left(\frac{F_0^2}{4} + G_8 F_0^4 \Lambda \left(1 + i\frac{c^0_{GG}}{2}(\Delta_V - \gamma_5\Delta_A)\frac{a}{f})\right)\right) \partial^\mu U \partial_\mu U^\dagger + \\
&\quad + F_0^2 B_0 M_q^0 \left(\Re(U) + 2c^0_{GG}\gamma_5 \frac{a}{f}\Im(U)\delta_A\right) \\
&+ \left(1 + 4G_8 F_0^2 \Lambda\right) \frac{\partial_\mu a}{f} \left((C^0_V + c^0_{GG}\delta_V) J^\mu_V - (C^0_A + c^0_{GG}\delta_A) J^\mu_A\right) + \frac{iF_0}{2} l_\mu J^\mu_L) + \mathcal{O}(1/f^2), \mathcal{O}(p^4)
\end{aligned} \tag{4.46}$$

Where the second-to-last term follows by using the cyclic property of the trace and the definition of the currents. The mass term also follows from the cyclic property along with the fact that δ_A and M_q^0 are diagonal matrices, so they commute with each other.
The full lagrangian reads

$$\mathcal{L} = \frac{1}{2}\partial_\mu a \partial^\mu a - \frac{m_a^{0\,2}}{2} a^2 + c_{\gamma\gamma}\frac{\alpha}{4}\frac{a}{f} F_{\mu\nu}\tilde{F}^{\mu\nu} + L^{(2)}_{\chi PT-a}. \tag{4.47}$$

Chapter 5

NA62

NA62 is a fixed-target experiment located in the P42 North-Area extraction line of the CERN Super Proton Synchrotron (SPS). It was built to measure rare kaon decays with the novel decay-in-flight method: the kaon decays with high momentum in the laboratory frame (75 GeV/c nominally), allowing for precise measurement of the decay products' kinematic quantities. The main physics objective of the experiment is a precise measurement of the branching ratio (BR) of the ultra rare $K^+ \to \pi^+ \nu\bar{\nu}$ decay with statistical errors better than 10%. To allow for this, along with precise measurement of the kinematic quantities, hermetical coverage is needed to veto decays with different signatures. This makes NA62 a perfect candidate to look for decays where the experimental signature involves missing energy, such as the one described in this thesis. The next paragraph describes the general structure of the NA62 experiment, whereas section 5.2 goes into more detail for the detectors relevant to the present analysis. More detailed description of the experiment and its detector can be found in the NA62 detector paper [55].

The schematic layout of the experiment is shown in figure 5.1. The 400 GeV/c proton beam of the SPS is extracted and impinges on a beryllium target, creating a secondary beam with a particle rate of 750 MHz; 6% of the beam is comprised of positively charged Kaons. After being collimated, the positively charged beam goes through a differential Čerenkov counter (KTAG), which identifies the K^+ particles. Three silicon pixel stations make up the Gigatracker (GTK), which is the tracking system for beam particles. The last GTK station is preceded by a collimator, shielding from upstream decays; further downstream a six-station plastic scintillator ring detector (CHANTI) vetoes against inelastic collisions occurring in the last station of the Gigatracker. The last station of the GTK marks the beginning of a 117 m long tank, the first 65 m of which is the fiducial decay region. Downstream of that, charged decay products are registered by a series of detectors, starting

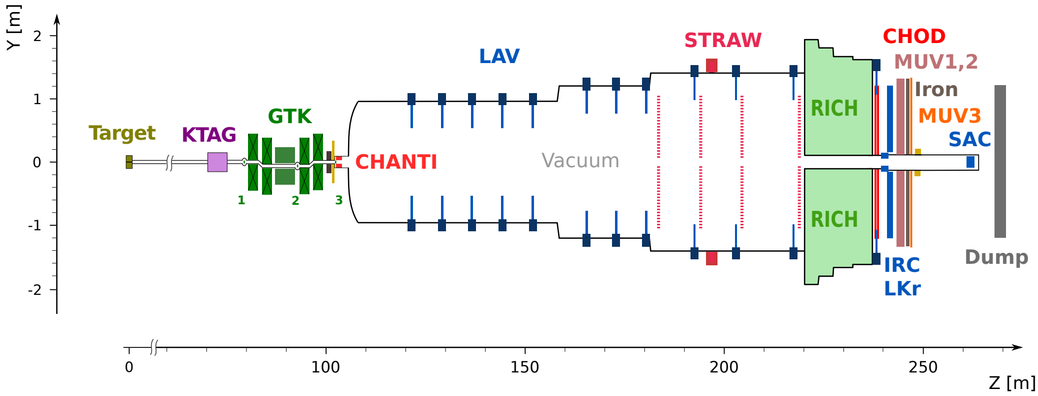


Figure 5.1: Schematic YZ (the Z axis is aligned to the beam pipe and the Y axis is vertical) view of the NA62 Experiment. Note the different scales on the Z and Y axis. Credits to the NA62 collaboration.

with a tracking system (STRAWS) which also measures the 3-momentum, a ring-imaging Čerenkov (RICH) counter for particle identification and two scintillator hodoscopes to provide timing. Other detectors are used as veto for photons (LAV, LKr, IRC and SAC), hadrons (MUV0, MUV1, MUV2, HASC), and muons (MUV3, located after a slab of iron).
The beam trajectory is surrounded by an evacuated pipe, which allows the flux of un-decayed particles to reach the beam dump without interaction with matter. All detectors between the KTAG and the RICH (excluded) are operated in vacuum for better resolution of kinematic quantities.

5.1 Analysis Principle

As said before, the same characteristics that allows NA62 to have a precise measurement of the $K^+ \to \pi^+\nu\bar{\nu}$ decay, make it a good candidate to study any kaon decay with missing 4-momentum as its experimental signature. It is then worth looking at how NA62 manages to detect this ultra-rare decay. To obtain a 10% statistical uncertainty on BR($K^+ \to \pi\nu\bar{\nu}$), $\mathcal{O}(100)$ decays need to be observed in this channel. Assuming the SM prediction $BR(K^+ \to \pi^+\nu\bar{\nu} = (9.11 \pm 0.72) \times 10^{-11}$[56] for the decay, one needs $\mathcal{O}(10^{13})$ K^+ decays, with $\mathcal{O}(10^{12})$ rejection for the background. The experimental signature is an incoming K^+ track, and an outgoing π^+ track with different 4-momentum (i.e. missing 4-momentum) and no signal from the veto system. The tools to reject the background and identify the wanted decays are particle identification (PID), kinematic rejection and photon veto.

PID provides a suppression of $\mathcal{O}(10^7)$ of muons with respect to pions, and is crucial to reject the most abundant K^+ decay, $K^+ \to \mu^+\nu$, which also has the experimental signature of an outgoing positively charged track and missing 4-momentum. PID is provided by the Ring Imaging Čerenkov Counter (RICH) and by the calorimeters (LKr,MUV1,MUV2).
Kinematic rejections is performed by cutting on the $m^2_{miss}(K^+ - \pi^+) = (P_{K^+} - P_{\pi^+})^2$ variable, where P_{K^+/π^+} is the 4-momentum of the incoming/outgoing track, where the particle energy is calculated under the hypothesis that it is a charged kaon/pion. Kinematic rejection is responsible of a suppression factor of $\mathcal{O}(10^4)$.
Photon Veto (PV) at the level of $\mathcal{O}(10^8)$ is needed to reject e.g. $K^+ \to \pi^+\pi^0(\pi^0 \to \gamma\gamma)$, which is the second-most abundant kaon decay. This is performed by three different types of calorimeter, together covering polar angles up to 50 mrad.

5.2 Detectors

5.2.1 K^+ tracking and Identification

Kaon Tagger (KTAG)

K^+ mesons constitute only 6% of the hadron beam coming out of the Beryllium target. They are identified using the Kaon Tagger (KTAG), which contains a differential Čerenkov counter, that is a Čerenkov detector accepting only light in a narrow range of angles; since the beam is at known momentum, this effectively selects a narrow range around a chosen mass. The particular design used for NA62 is called CERN W-type CEDAR, and it was designed in the late 1970s. The CEDAR radiator volume is filled with 0.94 m^3 of N_2 at 1.75 bar and room temperature: this was chosen so that only Čerenkov light from a K^+ can pass through the KTAG's annular diaphragm, which has a fixed central radius and varying radial aperture.
Light is then focused using spherical mirrors onto photomultipliers. The KTAG detector enjoys a time resolution of around 70 ps.

Beam spectrometer (GTK)

The GigaTracker (GTK) has been designed to accurately measure the trajectory and momentum of each beam particle. For the 75 GeV/c beam, it has 0.2% precision on magnitude of the momentum and 16 μrad of precision for the direction. The experimental requirement of a time resolution better than 200 ps was achieved using a three-station design. The beam is deflected vertically after the first station, and then returned to its nominal direction between the second station, where momentum is measured, and the third. This design also has the advantage of sweeping away positively charged muons. Each of the three stations is a pixel silicon detector, made up of 18 000 square pixels with a side of 300 μm, arranged in a 200 $\times$ 90 matrix. The stations were built so that their thickness is less than 0.5% radiation length each: this reduces multiple Coulomb scattering, giving the needed angular resolution and reducing spurious background (scattering in the last GTK station are particularly vicious, since the products are not swept away by the magnetic field).

Collimator

A collimator is located before the third GTK station.This is a 1 m thick steel slab with inner aperture of 66 x 22 mm^2 and outer dimensions of 15 cm. The purpouse of this collimator is to absorb hadrons produced in K^+ decays upstream of the fiducial decay volume.
In 2018 data-taking conditions, this was upgraded to the so-called Rainbow Collimator, with higher outer dimension which allow better absorption of hadrons from upstream K^+ decay.

Charged Anti-Coincidence Counter (CHANTI)

The Charged Anti-Coincidence detector (CHANTI) is installed right after the GTK, marking the beginning of the 65 m fiducial decay volume. It is designed to veto charged hadrons resulting from interactions in the last GTK station, or from muon resulting from decays of hadrons before the decay region. The CHANTI is composed of six hodoscope stations, each 300$\times$300 mm^2 in cross section, with a hole in the middle for the beam to pass. Each station consists of 48 scintillator bars of triangular cross section, half of them oriented vertically and half horizontally. Since it shares vacuum vessel with the GTK3, all hadrons originating in an angular region between 49 mrad and 1.34 rad are detected, with a time resolution of 1 ns and an efficiency of about 99%.

5.2.2 Charged Decay Products Tracking and Identification

STRAW

The STRAW spectrometer tracks charged decay products, with a momentum resolution better than 0.3% and an angular resolution ranging from 20 to

60 μrad (worsening with lowering momentum). It is located 20 m downstream of the end of the fiducial decay volume and is made up of four straw chambers, each of those composed by straw tubes arranged in four views, guaranteeing at least two hits per view. A magnetic field is provided by a large aperture dipole magnet, located between station two and three.

RICH

The Ring Imaging Čerenkov detector (RICH) is located right after the STRAW spectrometer. It is designed to separate pion from muons, suppressing muons at least of a factor 500. It also supplies the experiment with precise timing (better than 100 ps), and as such it is used as time reference detector in the trigger for data acquisition. The RICH is a 17.5 m long cylindrical vessel, with gradually decreasing diameter, filled with Neon gas at atmospheric pressure and room temperature. This allows for the RICH to be efficient in the region of momenta between 15 and 35 GeV/c: the lower threshold is given by the threshold for Čerenkov emission for charged pions (12.5 GeV/c), while the higher threshold is due to decreasing π/μ separation as the momentum increases.

NA62-CHOD and NA48-CHOD

Two Charged Hodoscopes, the NA62-CHOD and NA48-CHOD are used to tag charged particles. They are located upstream and downstream of the 12th LAV station, respectively, and operate independently of each other.
The NA48-CHOD was inherited by the NA48 experiment. It has a timing resolution of 200 ps, and as such serves as one of the main timing systems for charged tracks (the others being the RICH and the calorimeters). It is composed of 2 layers of plastic scintillator slabs, 64 slabs disposed vertically and 64 horizontally. Each slab is 20 mm thick - roughly one tenth of the radiation length. The slabs are assembled into 4 quadrants, forming an octagon of 1210 mm apothem, with a 128 mm hole in the middle for the beam pipe to pass. The light coming from each slab is read by a photomultiplier through a Plexiglas light guide.
The NA62-CHOD was built specifically for the NA62 experiment and it enjoys a $\mathcal{O}(1\,\mathrm{ns})$ time precision. It has an active area made up of 152 plastic scintillator tiles of 30 mm thickness, arranged in an annulus with inner/outer radius of 140/1070 mm. The light coming out of the scintillators is brought to SiPMs by wavelength shifting fibres.

Muon Veto System

The MUon Veto System (MUV), is part of the PID, helping to reach the needed $\mathcal{O}(10^7)$ rejection for $K^+ \to \mu^+\nu$. Perhaps confusingly, part of what is called MUV is actually two independent hadronic calorimeters (MUV1, MUV2), used to disentangle hadronic and electromagnetic showers for better PID. The reason for this is because MUV2 was inherited from the NA48 experiment, where indeed it was used for muon detection.
The MUV1 detector is made up of 24 layers of 26.8 mm thick steel plates, that is 4.1 radiation lengths, interleaved with 23 layers of 9 mm thick scintillator strips, alternating in the horizontal and vertical direction. The MUV2 detector has a similar "sandwich" layout, and was refurbished from NA62 predecessor in the study of Kaons at CERN SPS, NA48.
The Fast Muon Veto (MUV3) is located downstream of the hadron calorimeters, behind 80 cm thick iron wall. It detects particles traversing the whole calorimetric system plus the iron wall, which at NA62 energies means exclusively muons. As such, it is used for muon identification and as trigger for data taking, thanks to its time resolution of 400 ps. It is composed of 50 mm thick scintillator tiles: 140 tiles are of 220×220 mm^2 transverse dimension and 8 smaller tiles are located near the beam pipe, to sustain the much higher particle rate there.

5.2.3 Photon Veto System

Hermetic coverage for photons produced in decays occurring in the decay region at angles up to 50 mrad with respect to the beam axis is provided by the Photon Veto (PV) sytem. It is comprised of three systems: the Liquid Kripton (LKr) calorimeter, the Large-Angle Veto (LAV) and the Small-Angle Veto (SAV)

Liquid Krypton (LKr)

The Liquid Krypton (LKr) electromagnetic calorimeter is reused from the NA48 experiment, with some part of the readout system modified to sustain the new requirements. It is a quasi-homogeneous calorimeter filled with about 9 m^3 of Krypton at 120 K, kept inside a cryostat. It has a sensitive area of about 7m^3 divided into 13248 cells, arranged in a zig-zag shape to avoid inefficiencies due to aligned particle showers. In total, it corresponds to 27 radiation lengths.
It detects photons with polar angles between 1 and 10 mrad, with an energy resolution around the percent level for energy deposit of ≈ 20GeV and spatial

and time resolution around 1 mm and 1 ns respectively, for energy deposit of the same magnitude.

Large-Angle Veto (LAV)

The LAV system is comprised of 12 stations: one operated in air upstream of the LKr calorimeter, and other 11 positioned along the decay region vacuum volume, providing rejection for photons emitted at angles between 8.5 and 50 mrad. Each station is ring-shaped, made up of 37 cm long lead glass modules, recycled from the LEP experiment OPAL, arranged radially. The Čerenkov light produced inside each lead-glass block is detected by a 76 mm photomultiplier. Each station is comprised of multiple rings to ensure maximum radiation length: the total minimum effective depth is of 21 radiation lengths.
The time resolution for a particle of 1 GeV energy is 1 ns, with energy resolution better than 10%. This allows for the accidental rate from muons to be kept at $\mathcal{O}(1\%)$.

Small-Angle Veto (SAV)

The Intermediate Ring Calorimeter (IRC) and the Small-Angle Calorimeter system (SAC) are collectively known as Small-Angle Veto system (SAV).
The IRC is a lead/scintillator shashlyk calorimeter, shaped as an eccentric cylinder with outer diameter of 0.29 m and surrounding the beam pipe just upstream of the LKr. The IRC is comprised of two longitudinal modules, containing 25 and 45 ring-shaped layers, respectively; each layer is a 1.5 mm thick lead absorber and a 1.5 mm thick scintillator plate. Both modules have a total thickness of 19 radiation lengths.
The SAC is installed after the dipole magnet at the end of the beam pipe. It is a lead/scintillator sampling calorimeter, consisting of 70 plates of lead and 70 plates of plastic scintillator, both with a thickness of 1.5 mm. The SAC also has a total radiation depth of 19 radiation lengths.

5.3 Triggering and Data Acquisition (TDAQ)

Triggering and Data Acquisition (TDAQ) refers to all the techniques used to bring down the amount of data which is kept for storage and analysis. With an estimated rate of decays of 10 MHz, to be compared to a rate of roughly 10 kHz dictated by the available bandwidth for data-storage to tape, the importance of TDAQ is clear.
NA62 employs one level of hardware triggering (L0), which is based on inputs

from detectors with great time resolution, and analysed by FPGA boards. Various L0 triggers exist, the events being filtered based on informations from CHOD, RICH, MUV3, LKr and the 12th station of LAV. For the main analysis ($K^+ \to \pi\nu\bar{\nu}$), for example, loose requirement are applied to select events with only one positively charged particle in the final state. Events passing the L0 trigger are then sent to a PC farm for further analysis through the L1 and L2 trigger software.
For the present analysis, the relevant L0 trigger is the "Control" trigger, which selects all events with in-time coincidences of horizontal and vertical planes of the CHOD, thus selecting all events with (at least) one charged particle. The events are then downscaled by a factor of 400 to have a reasonable amount of data. Events passing the control trigger are generally used at analysis level to measure efficiencies, but for this analysis we will employ them, after having filtered them with an off-line filter, described in section 6.3.1.

5.4 Reconstruction

In order for a physics analysis to be performed, the quantities measured by the detectors need to be combined to get the information we care about, such as the particles' identities, their 4-momenta and trajectories, and decay vertices. A series of algorithms are used to do so, in a process called reconstruction. In this section we go over a few aspects that will be relevant to the present analysis. Before doing so, we introduce a piece of jargon: quantities related to the beam particles (in our case, positively charged kaons) are said to be "upstream", while quantities related to daughter particles are said to be "downstream", since they are measured by detectors located upstream and downstream, respectively, of the fiducial decay region.

5.4.1 Particle Identification

Upstream particle identification is straightforward: if the KTAG records a signal, the particle is identified as a K^+.
For downstream charged particles, the identification is performed by a series of algorithms that use the information gathered by either the RICH or the calorimeters. The algorithm for the RICH reconstructs the ring on which the Čerenkov photons lie in the photomultiplier plane which, together with the STRAWS-measured 3-momentum, is used to calculate the likelihood of the particle being a $\pi^+/\mu^+/e^+$. The calorimeters' algorithm reconstructs the showers - which leave so-called "clusters" of signals in the detector - distinguishing hadronic from electromagnetic showers. A probability is assigned by

the algorithm to each identification hypothesis. Muons are further identified by a signal in the MUV3. The efficiency of these algorithms can be seen in figure 5.2.
Photons also deposit an energy cluster in the LKr, which should be recognised as electromagnetic in nature.

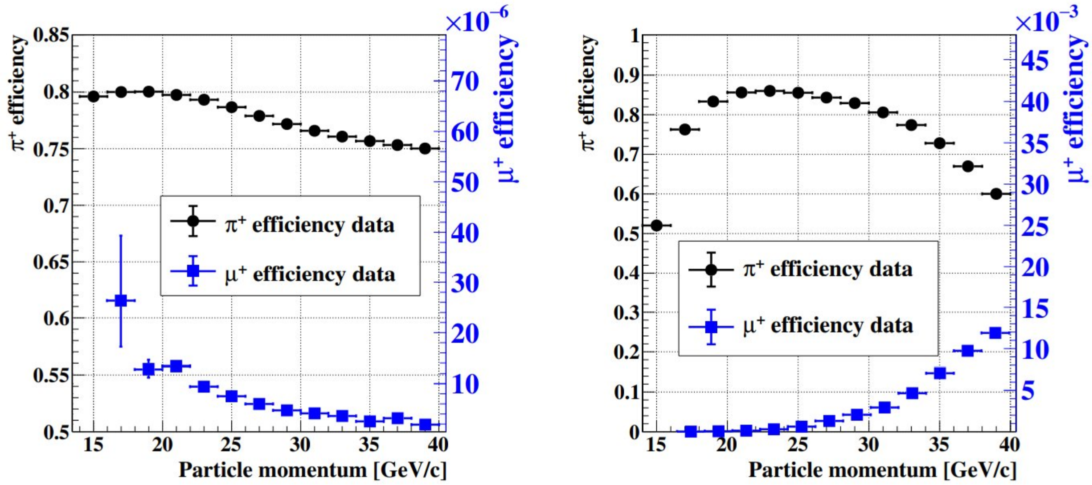


*Figure 5.2: Performance for π^+ identification of the calorimeters (**left**) and RICH (**right**). On each plot, the measured π^+/μ^+ efficiency, i.e. the fraction of π^+/μ^+ passing the π^+ selection of the $K_{\pi\nu\bar{\nu}}$ analysis, is plotted. Note the different scales. Note also that in the present analysis, slightly different selection criteria are applied to select pions: in particular we use the same cuts on the PID algorithm outputs for all downstream track momenta, while the efficiency shown here are calculated using optimised cuts that are different for each momentum bin, possibly yielding better performance than in our analysis. Taken from [57]*

5.4.2 Upstream Track Reconstruction

The information gathered by GTK is used to reconstruct the K^+'s kinematical quantities: the GTK's 3 stations measure the upstream 3-momentum $\vec{p}_K$and the position and timing at the third station. This gives the full trajectory in the decay region, after correcting for magnetic fields). On the hypothesis that the upstream track particle is a K^+, we can calculate the 4-momentum as:

$$p_K = (\sqrt{m_K^2 c^2 + |\vec{p}_K|^2}, \vec{p}_K) \tag{5.1}$$

The KTAG's efficiency for Kaons - i.e. the fraction of beam Kaons correctly identified - is not less than 98%. As for wrong identification of a beam particle as a kaon, the most common case is misidentification of a pion as a kaon. We need to distinguish two cases: isolated pion are misidentified as Kaons with a rate of $\mathcal{O}(10^{-4})$, while pions that are close in time to a kaon can be misidentified with a rate up to a few percent. All things considered, we have a rather certain identification of the upstream particle.

5.4.3 Downstream Track Reconstruction

Charged particles downstream of the fiducial decay region leave signal in the STRAWS and in the detectors used for PID. When a particle passes, its position in each STRAWS station plus the timing is recorded and called a "hit". The hits are fitted to yield the 3-momentum $\vec{p}_{track}$ of the downstream track. In a way that is completely analogously to what is done for kaons, one uses a mass hypothesis to calculate the daughter particles' 4-momenta:

$$p_{track} = (\sqrt{m_{daughter}^2 c^2 + |\vec{p}_{track}|^2}, \vec{p}_{track}), \tag{5.2}$$

where $m_{daughter}$ is the mass hypothesis for the charged track. This shows the importance of PID: since the calculated energy depends on the mass hypothesis, a wrong guess for the nature of the track leads to a wrong value. PID can be used to avoid that; however, downstream track PID enjoys a high but non-negligibly different from 100% efficiency, which is moreover dependent on the track momentum.

5.4.4 Vertex Reconstruction

Using the reconstructed tracks, one can determine the decay vertex, i.e. the place where the kaon decayed, as the point where the tracks' trajectory meet. Since two lines in 3D space do not generally meet, and because the measured tracks differ from the actual trajectories due to experimental resolution, the point of minimum distance between the tracks is used instead.

5.4.5 Neutral Pion Reconstruction

The π^0 decays with probability close to 1 into two photons without moving by any detectable amount from the vertex, since $c\tau \approx 10\mu$. This is used to reconstruct a π^0 from the detection of two electromagnetic clusters in the LKr. For each photon one calculates the versor $\hat{p}_{\gamma i}$ $i = 1/2$ directed from the decay vertex to the reconstructed impact point on the LKr entrance place. Together with the measured energy, this gives the estimate of the 4-momentum $p_{\gamma i}$ in the hypothesis that it is a photon originating from a π^0 decay that happened at the decay vertex. So finally one can reconstruct the π^0 momentum p_{π^0}:

$$p_{\pi^0} = p_{\gamma 1} + p_{\gamma 2} \qquad p_{\gamma i} = (\frac{E_{\gamma i}}{c}, \frac{E_{\gamma i}}{c}\hat{p}_{\gamma i}). \tag{5.3}$$

The determination of the correct value for p_{π^0} clearly depends on correct vertex and cluster reconstruction, as well as the elimination of accidental photons, since the photons can in principle come from different sources and still give rise to a π^0 being reconstructed by this procedure.

Chapter 6

Data Analysis

The sensitivity estimate of NA62 to the $K^+ \to \pi^+\pi^0 a$ decay is described in this chapter. The experimental signature of the decay is a single charged track identified as a π^+ and two electromagnetic clusters in the LKr calorimeter identified as the photons coming from $\pi^0 \to \gamma\gamma$; the clusters should be spatially unmatched with the charged track to prevent accidental misidentification, but they should be associated in time with it. There should be missing energy and 3-momentum, i.e. $p_{miss} = (E_{miss}, \vec{p}_{miss})$ should not be consistent with zero, where

$$\begin{cases} \vec{p}_{miss} = \vec{p}_K - \vec{p}_{\pi^+} - \vec{p}_{\pi^0} \\ E_{miss} = E_K - E_{\pi^+} - E_{\pi^0} \end{cases} \tag{6.1}$$

the 4-momenta of the measured particles measured/calculated as explained in section 5.4.

Two kinematic quantities will be studied to find the signal: the missing mass squared,

$$m^2_{miss} = (p_K - p_{\pi^+} - p_{\pi^0})^2, \tag{6.2}$$

which is equal to the square of the mass of the axion and the square of the transverse component of the missing 3-momentum:

$$p_T^2 = |\vec{p}_{miss} - \hat{p_K}(\vec{p}_{miss} \cdot \hat{p_K})|^2 \qquad \hat{p}_K = \frac{\vec{p}_K}{|\vec{p_K}|}. \tag{6.3}$$

The presence of a particle other than the two pions is deduced by the two quantities in equation (6.2) and equation (6.3) being not consistent with zero. The particle should be invisible, so we will make use of the NA62 veto system to make sure there are no visible particles in the final state other than the two pions.

The SM decays that can simulate the signal's signature are:

- $K^+ \to \pi^+\pi^0(\gamma)$ ($K_{2\pi}$), with mis-measurement of the kinematics quantities and/or an undetected γ simulating an axion.

- $K^+ \to \pi^+\pi^0\pi^0$ ($K_{3\pi 0}$) with one of the π^0 undetected.

- $K^+ \to \pi^+\pi^+\pi^-$($K_{3\pi C}$) with two charged pions undetected and accidental photons simulating the neutral pion.

- $K^+ \to l^+\pi^0\nu$ (K_{l3}) with the lepton $l^+ = e^+/\mu^+$ misidentified as a pion and the neutrino simulating the axion

- $K^+ \to \mu^+\nu$ ($K_{\mu 2}$) with the muon misidentified as a pion, accidental photons simulating a π^0 and the neutrino simulating the axion

To distinguish between the desired decay and these sources of background, Monte Carlo samples have been used, generated using the NA62MC framework of the NA62 collaboration, with particle-detector interactions simulated using the NA62Reconstruction framework, based on Geant4. The background samples were already available, while the signal samples were generated using a newly-written event generator (section 6.1.1).
A series of cuts on variables measured by the detectors were implemented, reducing as much of the background while retaining as much of the signal as possible. The MC samples are essential in this procedure, so that during the analysis one can see the effect of each cut on the various channels.
After a satisfactory set of cuts was found, the BR sensitivity of NA62 was estimated using the modified-frequentist CL_s method (chapter 7), yielding preliminary result on the upper limits of signal yield.

6.1 Data and Monte Carlo Samples

The data sample recorded during the full 2017B run period was used, as for the MC samples the following samples are used:

- $K_{2\pi}$ produced using v.2.1.4 of NA62MC and reconstructed using v.2.1.4 of NA62Reconstruction

- $K_{3\pi 0}$, $K_{\mu 3}$ produced using v.2.0.3 of NA62MC and reconstructed using v.2.2.2 of NA62Reconstruction

- $K_{3\pi C}$, K_{e3} produced using v.2.2.1 of NA62MC and reconstructed using v.2.2.1 of NA62Reconstruction

- $K_{\mu 2}$ produced using v.2.1.6 of NA62MC and reconstructed using v.2.1.6 of NA62Reconstruction

- Signal MC produced using v3.0 of NA62MC and reconstructed using v.3.0 of NA62Reconstrucion, for a masses ranging from 0 to 200 MeV/c^2 in steps of 20 MeV/c^2.

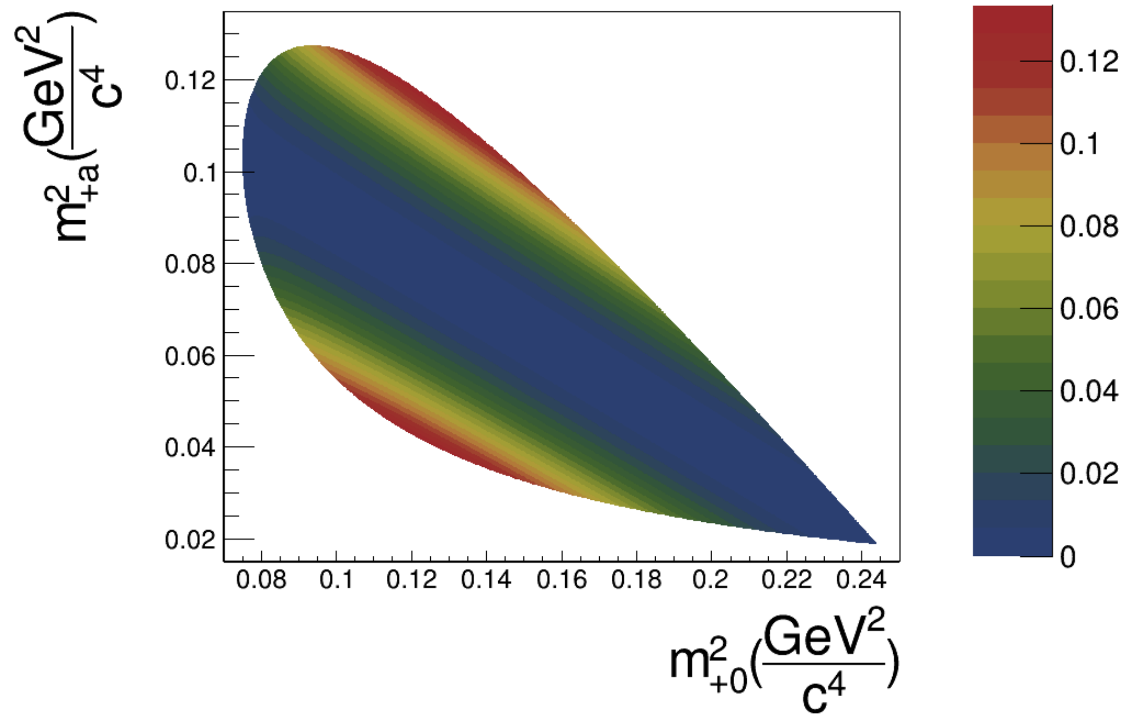
All MC samples were reconstructed in the "overlaid" mode, to simulate accidentals. This means that extra particles are added to each event, just like in actual data-taking each event can be contain extra particles on top one K^+ and its decay products.

6.1.1 Signal Monte Carlo Samples

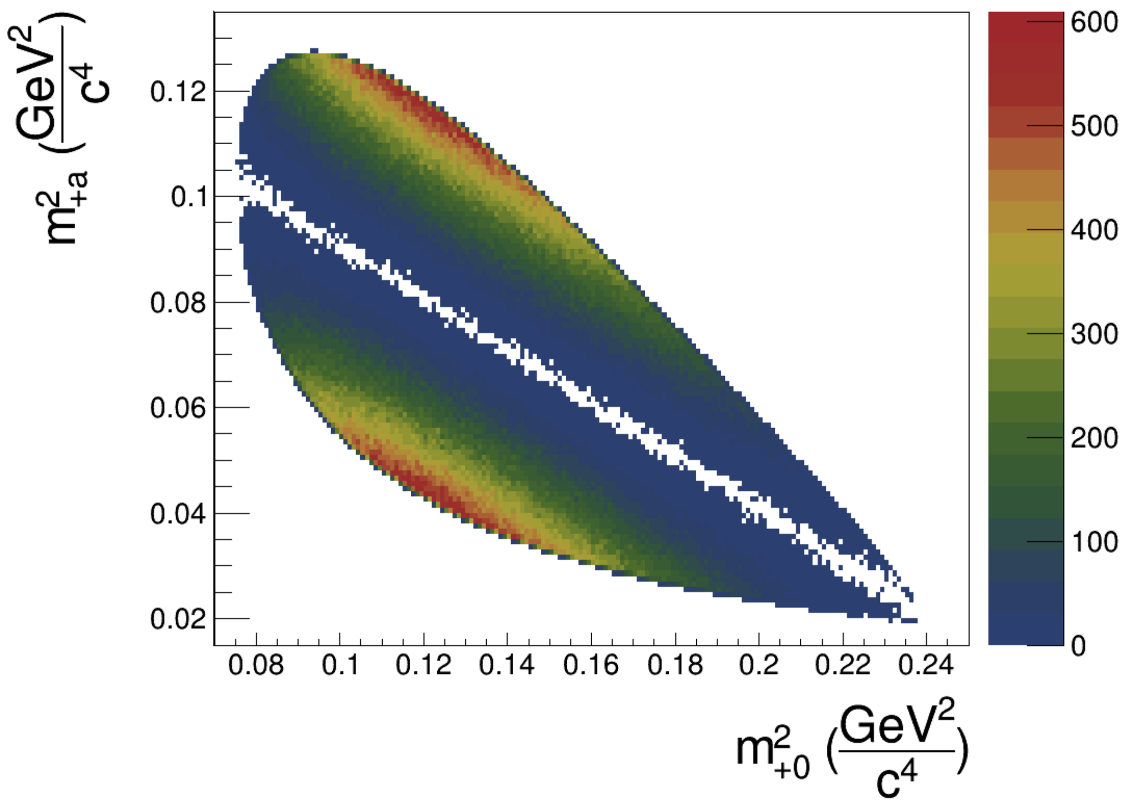
While all the samples for the background channels were already available to the collaboration, the MC sample for the signal was not, since this is the first time this particular channel is considered. An event generator was written using the differential rate calculated in chapter 3; it is now available to the collaboration for future studies. The sample was validated against the theoretical expression by checking that it produced the correct Dalitz plot (figure 6.1).
11 samples were produced and reconstructed using v.3.0 of NA62Reconstruction in the overlaid configuration. Each sample was generated under a different axion mass hypothesis, ranging from 0 to 200 MeV/c^2 in steps of 20 MeV/c^2.

6.2 Normalisation

The number of kaon decays N_K^{Data} that occurred in the fiducial during the period of data-taking is not, in principle, known. Moreover, we need to compare the data with MC samples, each corresponding to a different number of kaon decays. This means that some procedure needs to be applied to the MC samples to be able to a) sum all the MC samples and b) compare the sum of the MC samples with the data. This is called a normalisation procedure. It consists in applying a weight to each MC sample so that in the end they all represent the same number of kaon decays as the data sample. To calculate the weight for a given sample, we need to know the number of kaon decays in it, plus an estimate of the N_K^{Data}
For each sample i, it is easy to estimate the number of kaon decays N_K^i, since the number of generated events N_{gen}^i is known. Then using the PDG [29] value



(a) Analytical expression.



(b) Generated MC sample.

Figure 6.1: Dalitz Plot of the theoretical expression (a), and of the MC sample (b) for the $m_a = 0$ case.

BR^i of the decay we get[1]:

$$N_K^i = \frac{N_{gen}^i}{BR^i}. \tag{6.4}$$

For the data sample we cannot know the number of K decays *a priori*, so a more sophisticated technique is required. If the MC is a good description of the data, one can rely on it to estimate the number of kaon decays in the data sample by implementing a series of cuts, called the "normalisation selection" on both data and MC and measuring the efficiency ε_{sel}^{MC} of the normalisation selection:

$$\varepsilon_{sel}^{MC} = \frac{N_{K;\,sel}^{MC}}{N_K^{MC}}, \tag{6.5}$$

where $N_{K;\,sel}^{MC}$ is the number of event passing the normalisation selection in the MC.

Then, the number of kaon decays in the data is recovered by dividing the number of event that passed the selection in the data by the efficiency, assuming the selection has the same efficiency on the data.

The normalisation selection usually consists in a series of cuts designed to select only one decay channel. One decay among those with the highest BR of the channels constituting background is chosen, that mimics as closely as possible the signal. Choosing a decay with large BR assures us that the uncertainty in the BR is as low as possible, while the latter requirement is necessary since we are assuming that the normalisation selection has the same efficiency for the signal and the one channel we consider. For the present analysis, the decay chosen is $K_{2\pi}$, because the final-state particles are the same of the signal channel.

We now set to estimate the number of kaon decays passing the L0 Control trigger N_{norm}:

$$N_{norm} = \frac{N_K^{2017B}\varepsilon_{L0}}{D} \tag{6.6}$$

where N_K^{2017B} is the number of kaon decays that occurred in the fiducial region in the 2017B sample, ε_{L0} is the efficiency of the L0 Control trigger[2] and $D = 400$ is a downscaling factor applied by the L0 trigger to comply with bandwith restrictions of the L1 trigger.

After finding a set of cuts, whose precise description we put off until section 6.3.3, that lets only events from a $K_{2\pi}$ decay pass, the efficiency of the

[1] The BR of the signal we are looking for is, of course, not known. To have a reference when plotting distributions, we use $BR(K^+ \to \pi^+\pi^0 a) \sim 10^{-5}$, in accord with the UL set by the ISTRA+ collaboration[58].

[2] Here we are assuming the trigger has the same efficiency on all decay channel, which is very close to being true

selection is calculated:

$$\varepsilon^{K_{2\pi}} = \frac{N^{K_{2\pi}}_{K_{2\pi}-like}}{N^{K_{2\pi}}_{gen}} \tag{6.7}$$

with $N^{K_{2\pi}}_{K_{2\pi}-like}$ the number of events in the $K_{2\pi}$ MC sample passing the normalisation selection. This has a relative uncertainty equal to

$$\frac{\delta\varepsilon^{K_{2\pi}}}{\varepsilon^{K_{2\pi}}} = \sqrt{\frac{1-\varepsilon^{K_{2\pi}}}{N^{K_{2\pi}}_{K_{2\pi}-like}}}. \tag{6.8}$$

N_{norm} is then estimated as

$$N_{Norm} = \frac{N^{Data}_{K_{2\pi}-like}D'}{BR^{K_{2\pi}}\varepsilon^{K_{2\pi}}} \tag{6.9}$$

where D' is the downscaling factor of the data, if applicable.
The number found using this procedure is informally called the "Normalisation" and is subject to a relative uncertainty that we can set equal to the sum of the relative uncertainties of the factors in its definition

$$\frac{\delta N_{Norm}}{N_{Norm}} = \frac{1}{\sqrt{N^{Data}_{K_{2\pi}-like}}} + \sqrt{\frac{1-\varepsilon^{K_{2\pi}}}{N^{K_{2\pi}}_{K_{2\pi}-like}}} + \frac{\delta BR^{K_{2\pi}}}{BR^{K_{2\pi}}} \tag{6.10}$$

Finally, each MC sample is normalised to N_{Norm}, by weighting the events of the sample i by N_{Norm}/N^i_K. This defines what we call the "Number of Events" in the following: for the data it is directly the number of event satistying a certain cut, while for the MC samples it's corrected by the weights calculated in this section.
This does not introduce significant error: the dominant relative error in this factor is the one due to $BR^{K_{2\pi}}$ which is around 4‰, so we would need around 10.000 (or more) events per bin to have the uncertainty due to counting be comparable (or smaller). Since we are mainly interested on the uncertainties on the number of events toward the end of the analysis, when only a handful of events are present per bin, we will neglect the uncertainty due to the normalisation procedure.

6.3 Event selection

In this section, the cuts made to select data are described. First, a set of cuts is applied to reduce disk space usage while applying rather loose cuts (section 6.3.1). Then, after inspection of the distribution of variables for various MC samples, a series of cuts is chosen (section 6.3.2), this is called the "preselection". After that, two different selection are employed : the first one selects $K_{2\pi}$ event to calculate the normalisation (section 6.3.3), the second one selects instead signal-like events (section 6.3.4).

6.3.1 Data reduction

To reduce the amount of data for the analysis, an off-line filter, called OnePionTwoClusters (1P2C), is used. It selects events with one positively charged track and two energy cluster deposit in the LKr, also eliminating events that of dubious nature due to mismeasurements. To pass this filter, it is required that:

- The event passed the L0 Control trigger.
- There is one positively charged track, in the geometrical acceptance of all 4 STRAWS stations.
- The fit of the STRAWS hits in the track reconstruction yields a χ^2 less than 40.
- The track is in time with the trigger within a window of $\pm$ 20 ns.
- The decay vertex (section 5.4) is within the fiducial decay region.
- The minimum distance (in 3D space) between the reconstructed track and beam axis is less than 100 mm.
- The track momentum, as measured by the STRAWS, is less than 70 GeV/c.
- The ratio E/p between the energy deposited by the charged particle in the LKr and its momentum does not exceed 0.95 c.
- There are two in-time energy cluster in the LKr with energy bigger than 3 GeV.
- The clusters' and the track's impact points on the LKr are separated by at least 50 mm.

The combination of a positively charged track and two clusters selected by the algorithm is called a triplet.

The filter also flags events as “K2PI” or “Rare”, according to two variables: the $K^+ - \pi^+$ missing mass squared[3]

$$m^2_{K^+-\pi^+} = (p_K - p_{\pi^+})^2, \tag{6.11}$$

whose distribution can be seen in figure 6.2, and the sum of track and cluster

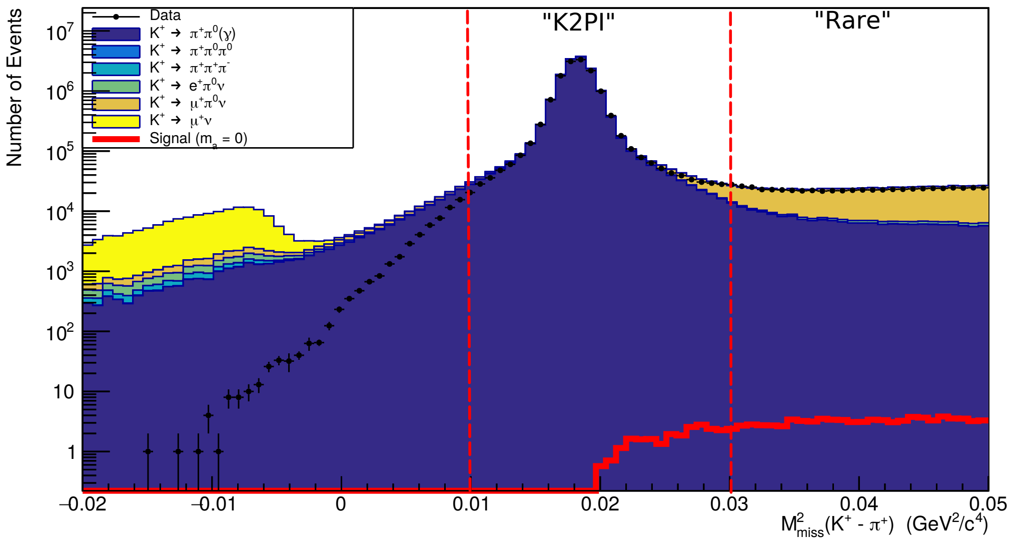


Figure 6.2: Distribution of $m^2_{K^+-\pi^+}$, after applying the 1P2C filter. The MC samples for the background channels are stacked on one another, while the data and signal samples are drawn by themselves. Note how the “K2PI” region is clearly dominated by $K_{2\pi}$ events, while the regions at higher or lower values of the missing mass also displays an abundance of other channels: in particular the signal is mostly found in the “Rare” region.

3-momenta

$$P_{\pi^+ + \pi^0} = |\vec{p}_{\pi^+} + \vec{p}_{\pi^0}| = |\vec{p}_{\pi^+} + \vec{p}_{\gamma_1} + \vec{p}_{\gamma_2}|, \tag{6.12}$$

whose distribution can be seen in figure 6.3. Both these quantities are defined in terms of the kinematical quantities in equation (5.2) and equation (5.3). The K2PI flag is designed to select $K_{2\pi}$-like events. It asks that

- 65 < $P_{\pi^+ + \pi^0}$ < 80 GeV/c
- 0.01 < $m^2_{K^+-\pi^+}$ < 0.03 GeV$^2/c^4$.

[3]NB: this is not the missing mass we consider in the rest of our analysis

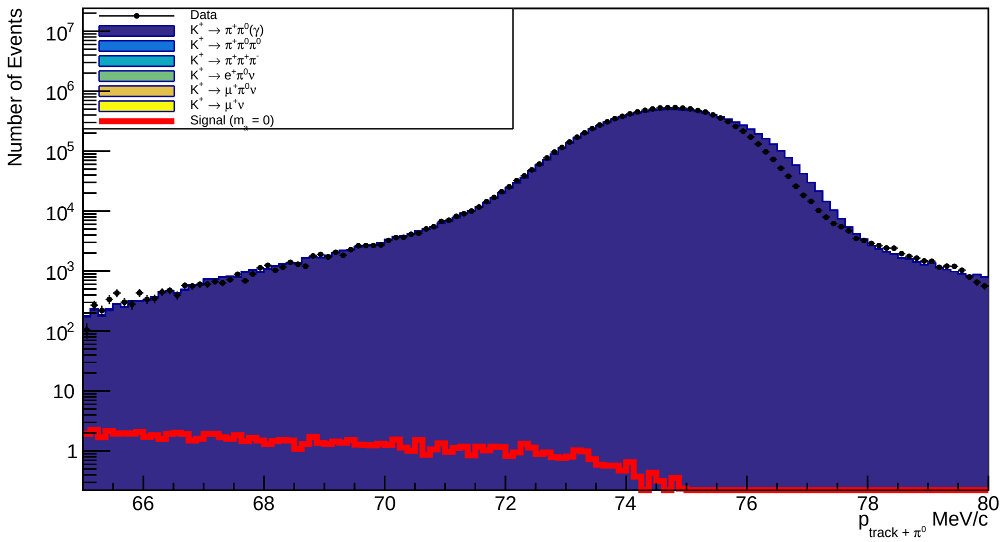


Figure 6.3: Distribution of $P_{\pi^+ + \pi^0}$ after the 1P2C filter. The MC samples for the background channels are stacked on one another, while the data and signal samples are drawn by themselves.

the first requirements is done to have the sum of the two pion momentum loosely consistent with the nominal $|\vec{p}_{K^+}| = 75$ GeV/c, while the second requirement selects a π^0 candidate ($m^2_{\pi^0} \approx 0.02\,\mathrm{GeV}^2/c^4$). Since this flag is generally used for normalisation studies, a downscaling factor of $D' = 10$ is applied on this selection to reduce disk space usage, so we will have to apply a correction for it by weighting the K2PI-flagged events with D'.
The Rare flag is instead designed to select rare and/or hypothetical decays of the Kaon containing other (possibly invisible) particles in the final state on top of a $\pi^+\pi^0$ pair. An event is flagged as “Rare” if

- $m^2_{K^+ - \pi^+}$ > 0.03 GeV$^2/c^4$.

To further reduce the amount of data while retaining only interesting events for the analysis, further requests were made, using a private analysis code: we define a triplet to be “good” if it passes the 1P2C filter and

- The time difference between track and the clusters, and between the clusters, does not exceed 10 ns.
- All 4 STRAW stations recorded a signal.
- The track is in the geometrical acceptance for the LKr .

- There is no in-time activity in the LAV, SAC or IRC within $\pm$ 5 ns.

We then require the event to have one and only one “good” triplet.

6.3.2 Preselection

As explained before, we will perform a normalisation selection, which should let only $K_{2\pi}$ events pass, and a signal selection, which should let most signal event pass while reducing the background. Since both of them have a π^+ and a π^0 in the final state, it is only natural that they share a common set of cuts. Indeed, we would like the two selection to be as close as possible, because we assume the efficiency of the normalisation selection to be the same on the signal and on the $K_{2\pi}$ channel. We call this set the “Preselection” cuts. We first ask that

- The Kaon momentum is measured by the GTK to be between 72 and 78 GeV/c.
- The Charged track momentum is measured by the STRAWS to be between 15 and 45 GeV/c

so that we have a reliable measure for the kaon momentum and the charged track momentum is in the region where PID is efficient.
Next, we require only one charged track in the event, weeding out events with accidental charged tracks and $K_{3\pi C}$ events. The charged track is required to be identified as a pion. This means that:

- The RICH algorithm reconstructs it as a π^+ and assigns a likelihood of less than 0.5 that it is a positron or a muon.
- The Calorimeter algorithm reconstructs it as π^+ and assigns probability greater than 0.8 to this identification.
- No in-time activity is registered in the MUV3

these requirements weed out decays containing leptons (see figure 6.8 to figure 6.7).

Finally, some requirements are made on clusters and more stringent cuts on timing: this is to weed out accidentals, and has the effect of greatly reducing the number of $K_{3\pi C}$ decays (see figure 6.9, figure 6.10 and figure 6.11).

- The clusters impact points are separated from those of the charged track by at least 100 mm

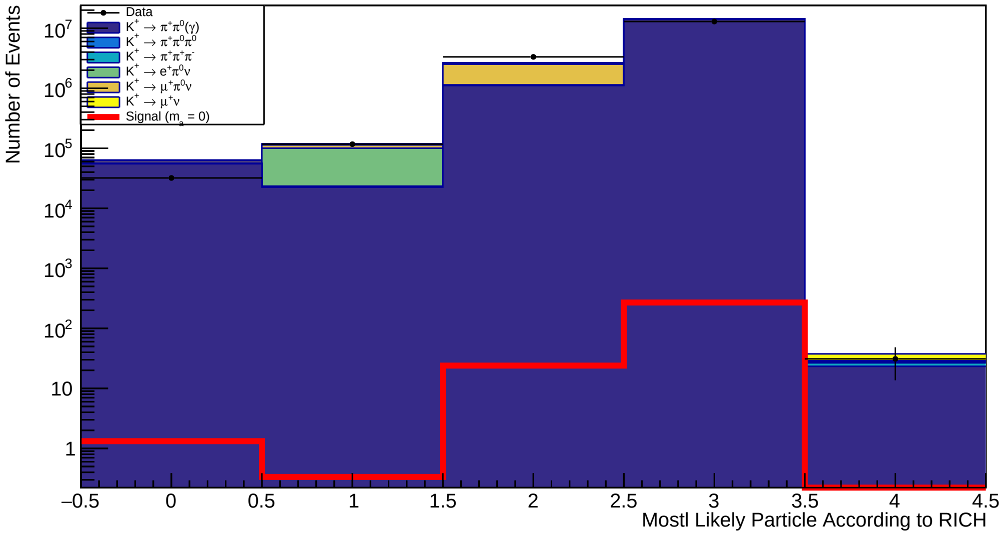


Figure 6.4: Distribution of the particle with the highest RICH likelihood after applying the 1P2C filter. The different hypothesis are: 0 = noise, 1 = positron, 2 = μ^+, 3 = π^+ and 4 = K^+. The MC samples for the background channels are stacked on one another, while the data and signal samples are drawn by themselves.

- The clusters impact points are separated from those of any other cluster in the LKr by at least 150 mm
- The clusters are identified by the LKr as electromagnetic.
- The time difference between track and any one cluster, and between one cluster and the other does not exceed 3 ns.

After applying these cuts, we will use the selected events for two purposes: to select $K_{2\pi}$-like events to estimate the normalisation and to select signal-like events for the sensitivity estimate, using the selections described in the two following subsections.

6.3.3 Normalisation Selection

The $K_{2\pi}$ selection used for normalisation was largely borrowed from the $K^+ \to \pi^+\nu\bar{\nu}$ analysis [59]. On top of the preselection, we require the “K2PI” flag of the 1P2C filter, which indeed selects $K_{2\pi}$-like events. We further require that the invariant mass of the two clusters $m_{\gamma\gamma}$ is loosely compatible with the pion mass, $110 < m_{\gamma\gamma} < 160$ MeV/c^2. As we can see in figure 6.14, this

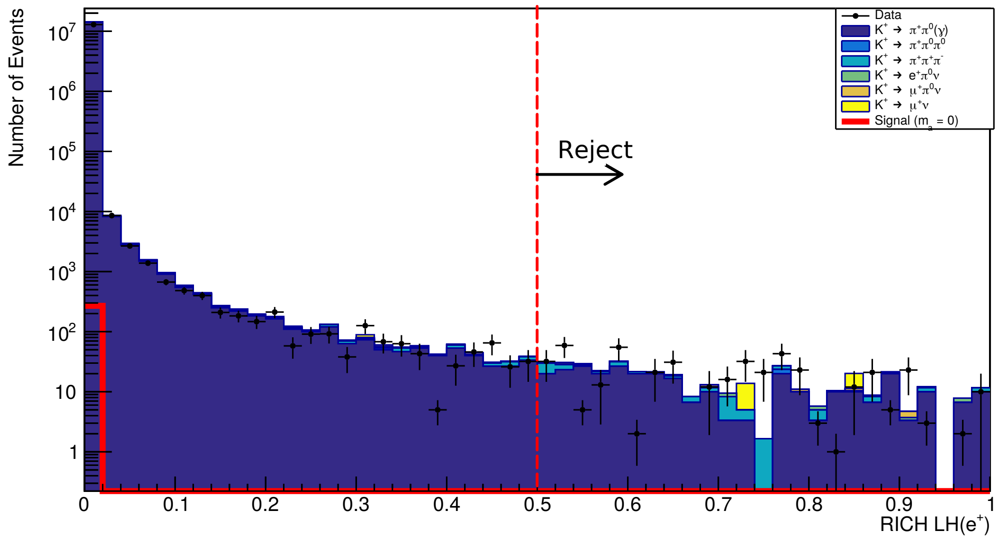


Figure 6.5: Distribution of the RICH likelihood of the track being a positron after applying the 1P2C filter ***and requiring the most likely particle to be a*** π^+*, the red dotted line representing the cut value. The MC samples for the background channels are stacked on one another, while the data and signal samples are drawn by themselves.*

selection does indeed let only $K_{2\pi}$ event pass. We can also observe that the MC reproduces rather closely not only the normalisation of the peak (which we imposed by hand), but also its shape and the tail region.

6.3.4 Signal Selection

To select signal-like events, on top of the Preselection we require the “Rare” flag of the 1P2C filter, as well as a more restrictive cut on $m_{\gamma\gamma}$: we ask that 125 < $m_{\gamma\gamma}$ < 145 MeV/c^2, as suggested by studying figure 6.15. The effect of this two cuts can be seen in figure 6.12 and figure 6.13, respectively. Applying the Rare flag weeds out $K_{2\pi}$ events, as it should, and we expect the $K_{2\pi}$ event passing this selection to mostly be event with one undetected γ. As for the $m_{\gamma\gamma}$ cut, it weeds out $K_{3\pi 0}$ and $K_{3\pi C}$ events; this is most likely because the events from these channels passing previous cuts have a either a fake π^0 candidate or accidental photons, leading to a wrong mass for the π^0.

After applying these cuts we can turn to the 2D-distributions in the plane p_T^2 vs m_a^2, see figure 6.16. The signal appears approximately as a Gaussian in the x direction and has a rather flat distribution in the y direction, the maximum value of p_T^2 lowering as the mass increases. On the other hand, the

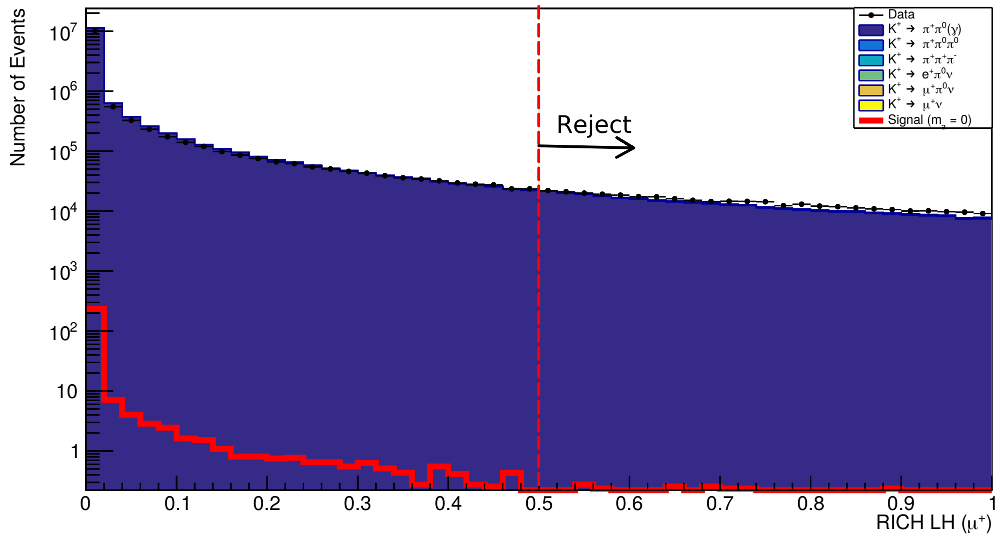


Figure 6.6: Distribution of the RICH likelihood of the track being a muon after applying the 1P2C filter ***and requiring the most likely particle to be a*** π^+*, the red dotted line representing the cut value. The MC samples for the background channels are stacked on one another, while the data and signal samples are drawn by themselves.*

$K_{2\pi}$ distribution is almost entirely located at low p_T^2 or at negative m^2_{miss}. An optimisation algorithm was ran to look for the best cut of the type $p_T^2 > p^2_{T\,min}$ to reject $K_{2\pi}$ while retaining as much of the signal as possible. The quantity to optimise is

$$\frac{N^{m_a=0}_{K_{2\pi a}-like}}{\sqrt{N^{K_{2\pi}}_{K_{2\pi a}-like}}}, \tag{6.13}$$

$N^{K_{2\pi}/m_a=0}_{K_{2\pi a}-like}$ being the number of events from the $K_{2\pi}$ MC sample and the signal MC sample in the massless case, respectively, passing the signal selection. The square root in the denominator is to account for statistical fluctuation of the background; the algorithm was also run optimising

$$\frac{N^{m_a=0}_{K_{2\pi a}-like}}{N^{K_{2\pi}}_{K_{2\pi a}-like}} \tag{6.14}$$

instead. We obtained comparable results, which gives us a check that the result is not extremely dependent on the quantity we decide to optimise for.
We obtain $p^2_{T\,min} = 6.5 \times 10^{-3}\mathrm{GeV}^2/c^4$.
A flat cut on p_T^2 works well for the massless hypothesis; however, since it serves

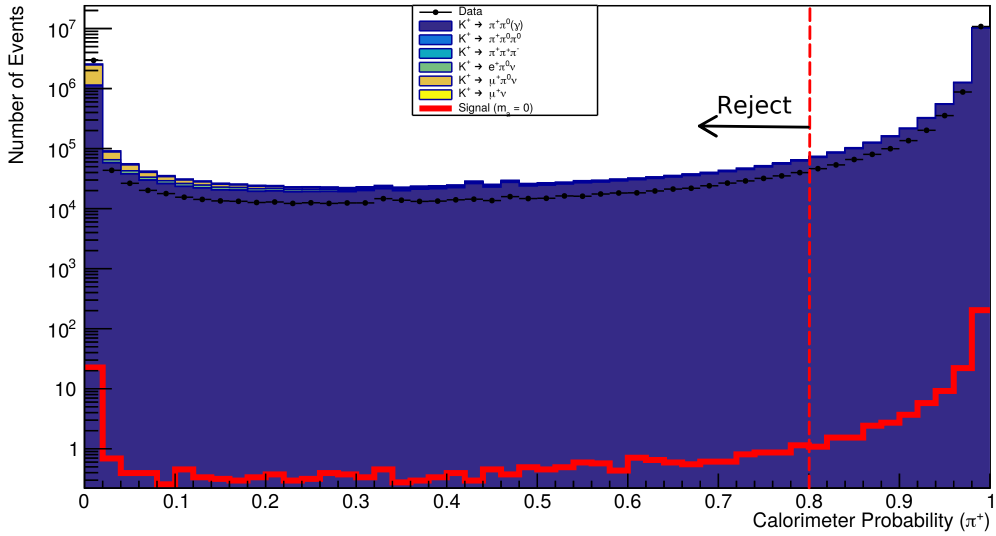


Figure 6.7: Distribution of the Calorimeter probability of the track being a pion after applying the 1P2C filter, the rejected region to the right of the red dotted line, the red dotted line representing the cut value. The MC samples for the background channels are stacked on one another, while the data and signal samples are drawn by themselves.

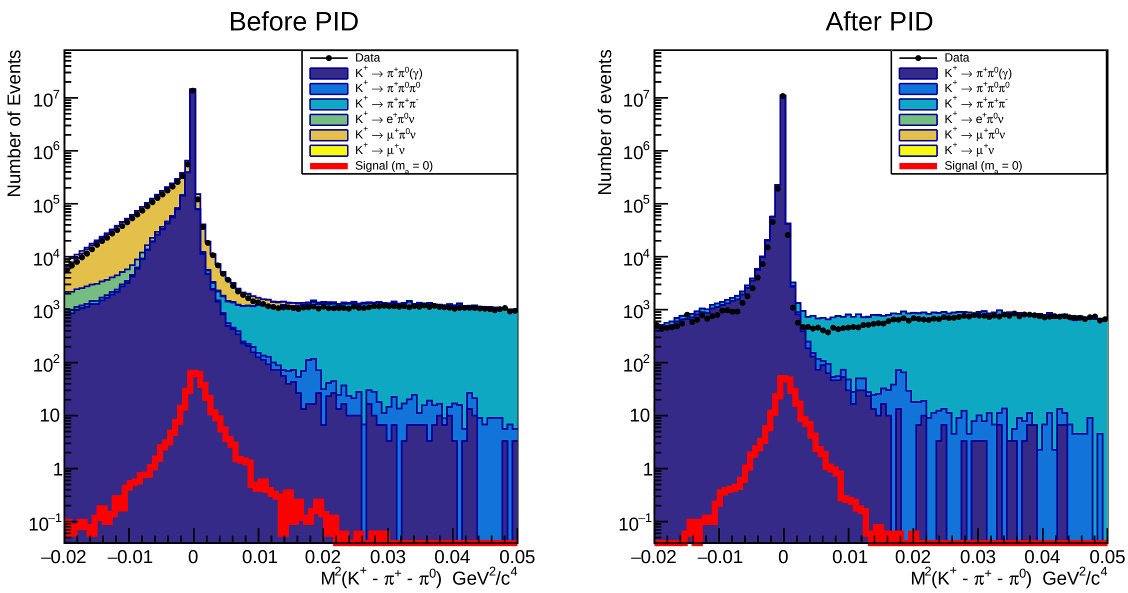


*Figure 6.8: m^2_{miss} distribution before (**left**) and after (**right**) PID cuts. The MC samples for the background channels are stacked on one another, while the data and signal samples are drawn by themselves.*

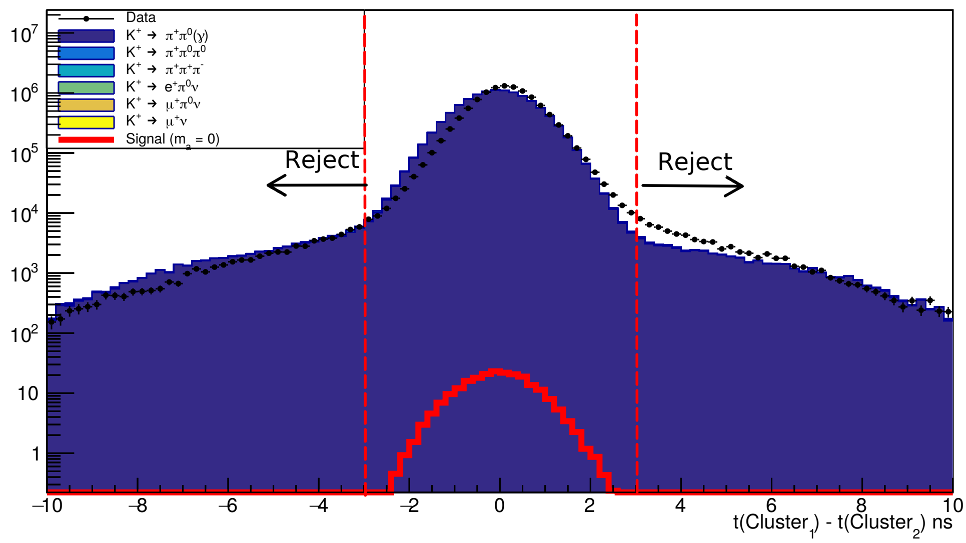


Figure 6.9: Distribution of the time difference between the two clusters after applying the 1P2C filter, the red dotted lines representing the cut values. The MC samples for the background channels are stacked on one another, while the data and signal samples are drawn by themselves. An offset can be observed between the MC samples and the Data: this is not an issue for our analysis since we are employing a cut that is rather large in comparison to the offset.

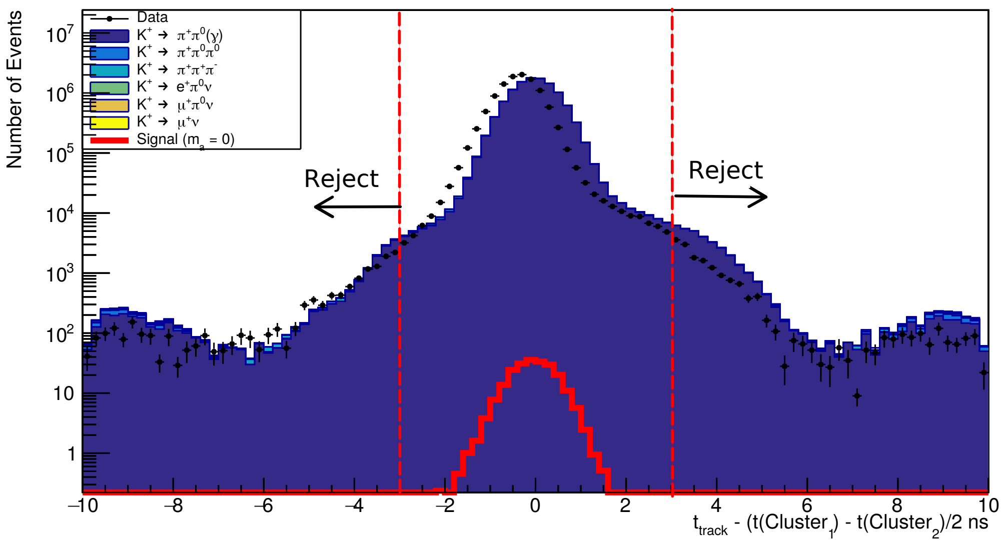


Figure 6.10: Distribution of the difference between the track timing and the mean of the time of the two clusters after applying the 1P2C filter, the red dotted lines representing the cut value. The MC samples for the background channels are stacked on one another, while the data and signal samples are drawn by themselves. An offset can be observed between the MC samples and the Data: this is not an issue for our analysis since we are employing a cut that is rather large in comparison to the offset.

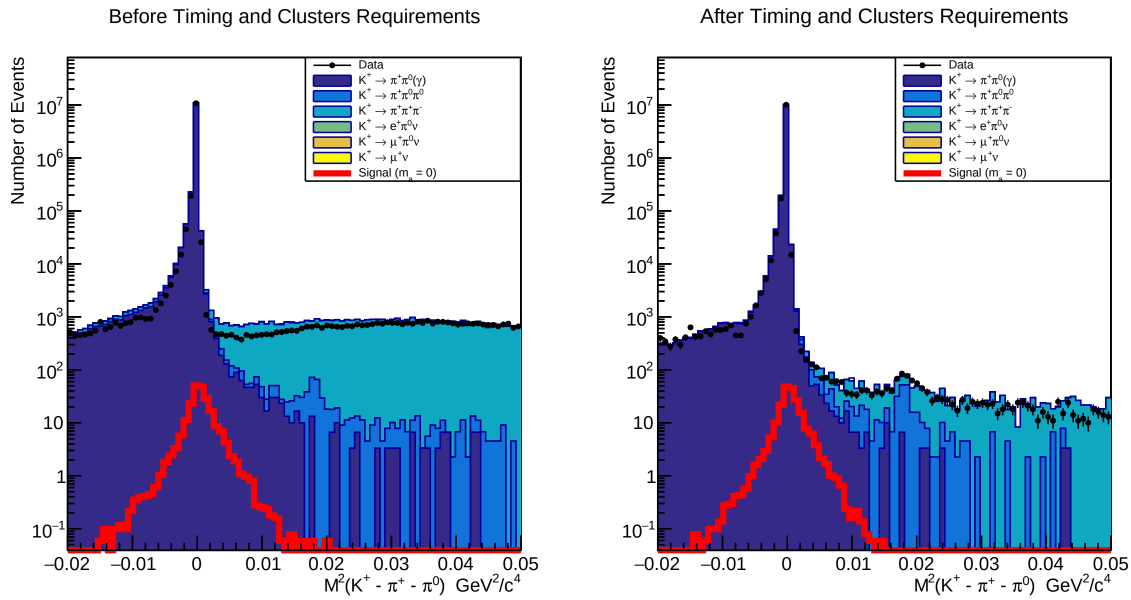


*Figure 6.11: m^2_{miss} distribution before (**left**) and after (**right**) timing and cluster cuts. The MC samples for the background channels are stacked on one another, while the data and signal samples are drawn by themselves.*

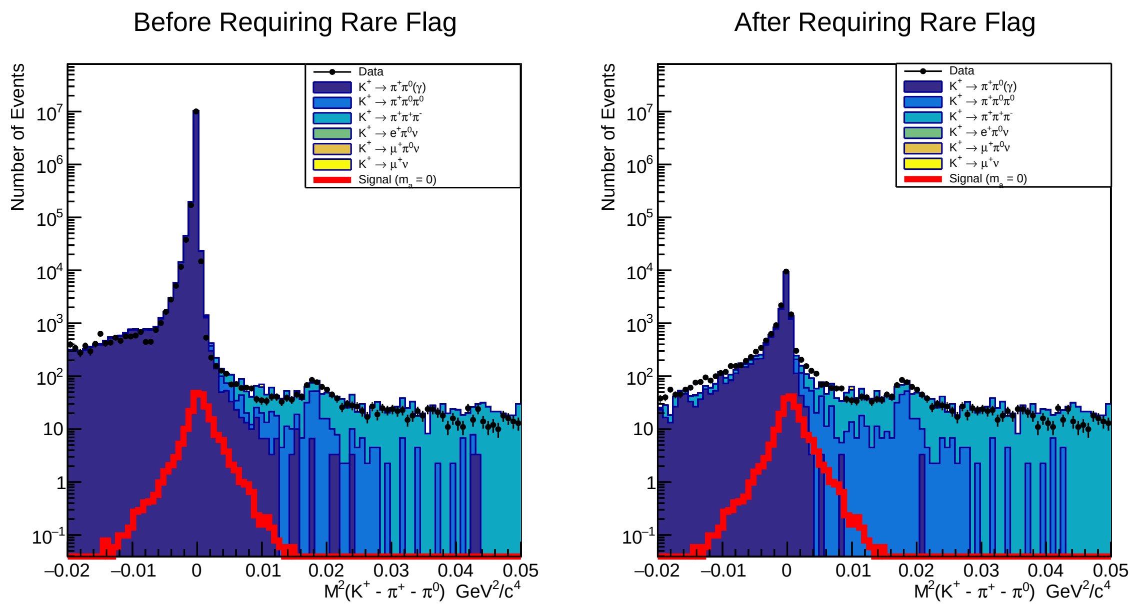


*Figure 6.12: m^2_{miss} distribution before (**left**) and after (**right**) requiring the "Rare" flag. The MC samples for the background channels are stacked on one another, while the data and signal samples are drawn by themselves.*

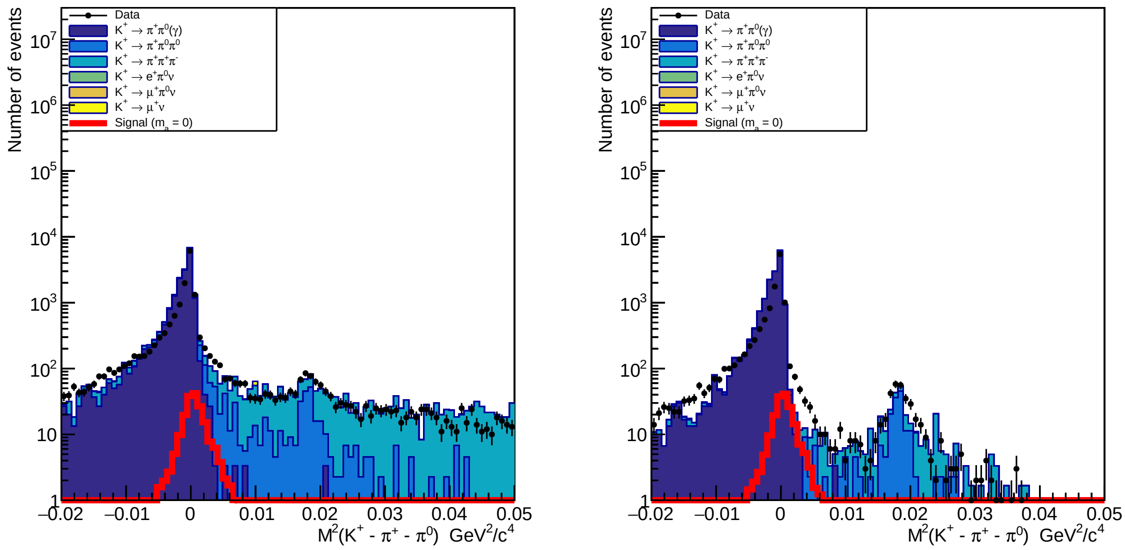


*Figure 6.13: m^2_{miss} distribution before (**left**) and after (**right**) the more stringent $m_{\gamma\gamma}$ cut. The MC samples for the background channels are stacked on one another, while the data and signal samples are drawn by themselves.*

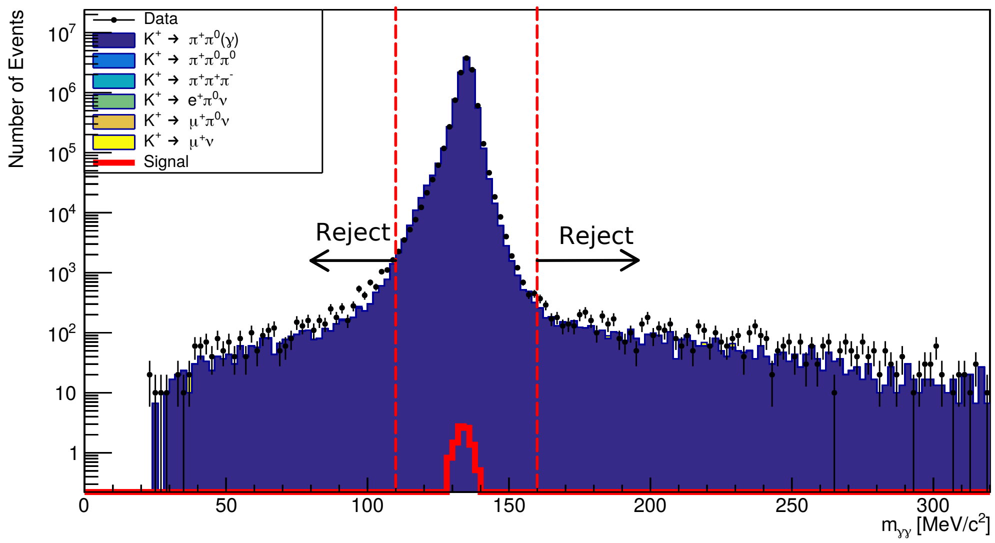


Figure 6.14: $m_{\gamma\gamma}$ distribution after the selection used for normalisation, the red dotted line representing the cut value. The MC samples for the background channels are stacked on one another, while the data and signal samples are drawn by themselves. Note how only $K_{2\pi}$ events surive the selection.

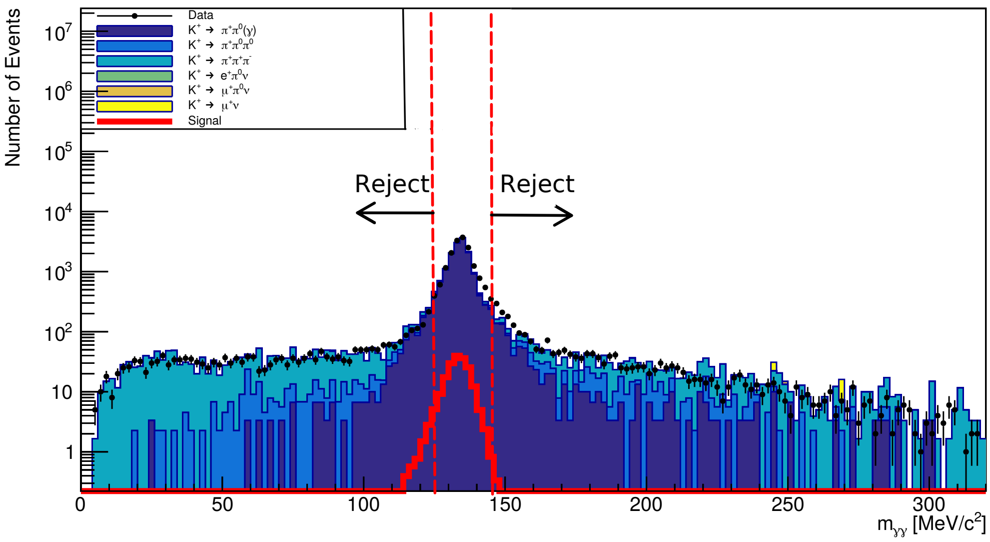
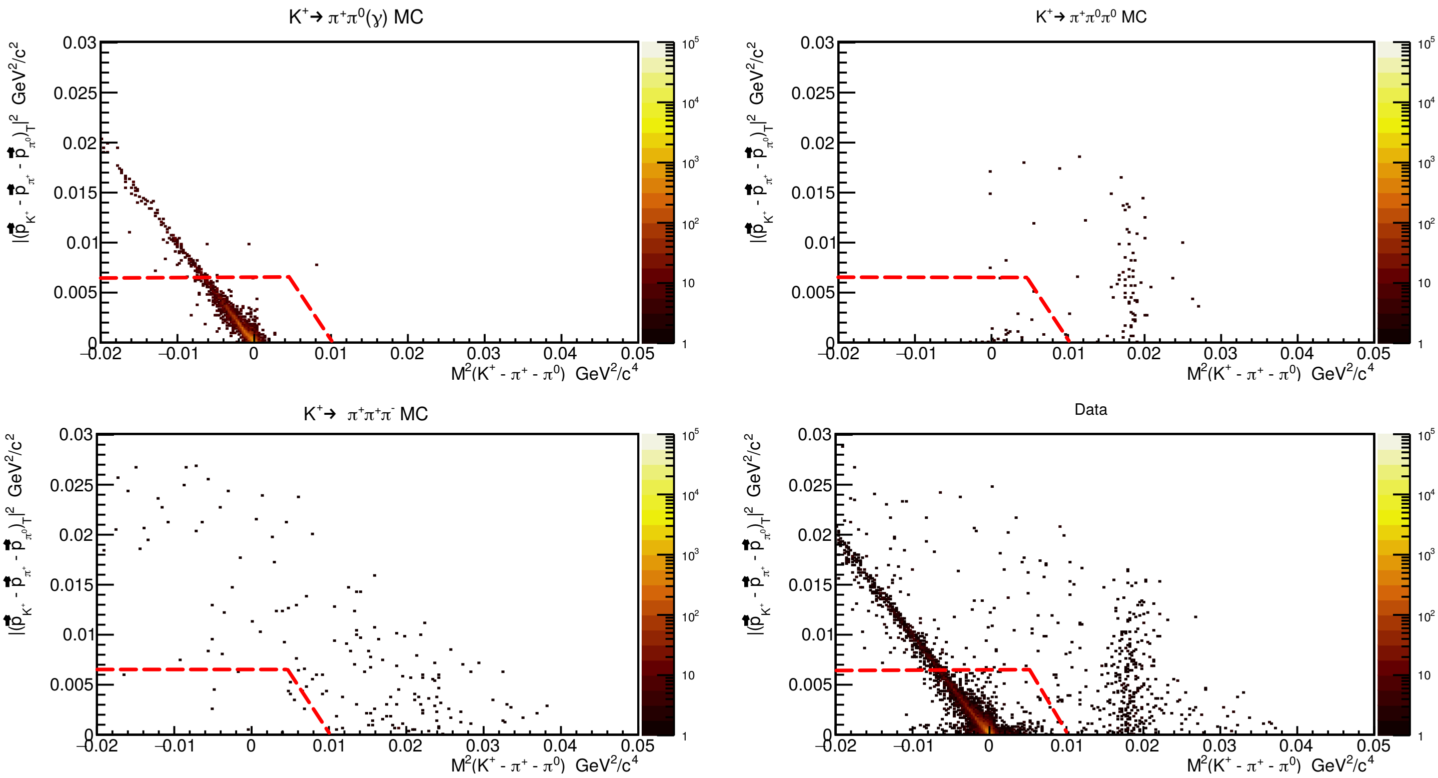


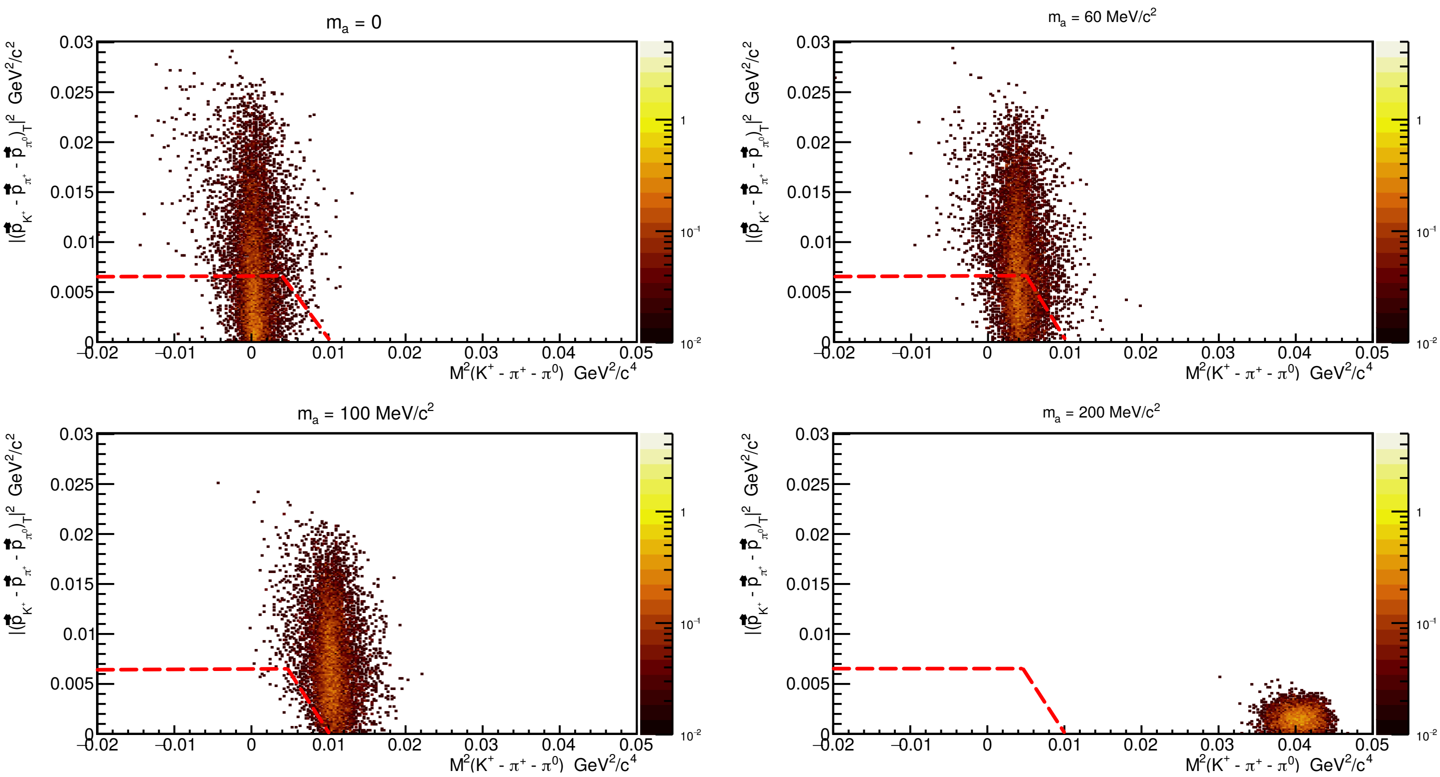
Figure 6.15: $m_{\gamma\gamma}$ distribution after requiring the "Rare" flag, the red dotted line representing the cut value. The MC samples for the background channels are stacked on one another, while the data and signal samples are drawn by themselves.

the purpose of eliminating $K_{2\pi}$, which are almost entirely found at low or even negative m^2_{miss}, it is suboptimal for larger axion mass. Indeed, at the end we will perform the search in a region of m^2_{miss} near the mass hypothesis, meaning that for high masses the $K_{2\pi}$ are automatically excluded. Because of this considerations, a trapezoidal cut was chosen: see figure 6.16. This cut was chosen as a compromise between ease of implementation and the result of an optimisation algorithm similar to the one above.
The result of all the signal selection cuts can be seen in figure 6.17.

The cuts on the variables p^2_T and m^2_{miss} are the last cuts we employ in our analysis. As said before, the signal search is performed in the m^2_{miss} variable: to do this we study the shape of the distribution for the signal and choose a region containing most of the signal. The sum of two Gaussian curves with equal mean approximates well the data (see figure 6.18), so for an axion of mass m_a the search region is defined as $[m_a - 2\sigma, m_a + 2\sigma]$, with σ the weighted mean of the standard deviation of the two Gaussians, using the normalisation as weight.
We are finally ready to make an estimate for the sensitivity of the NA62 experiment to this decay.

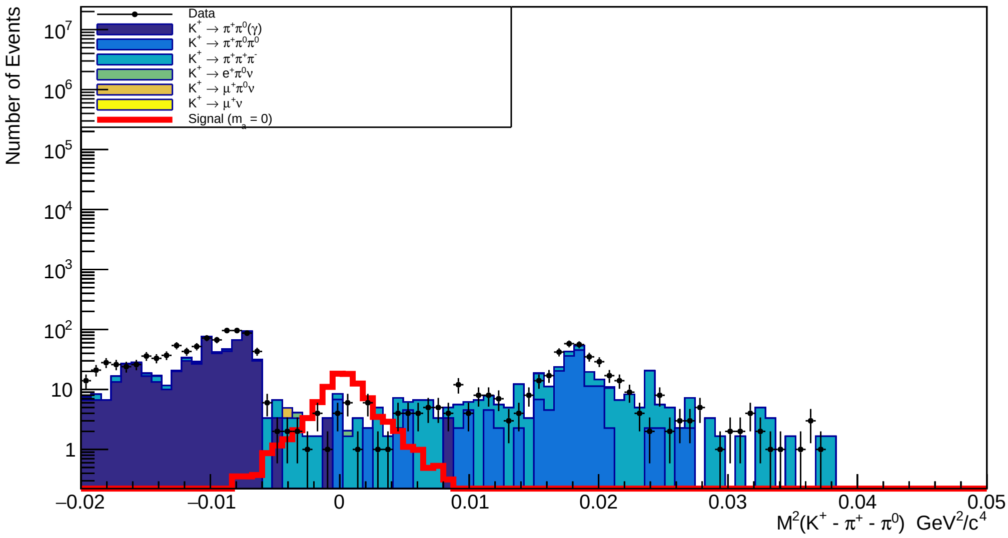


(a) Dominant backgrounds and data

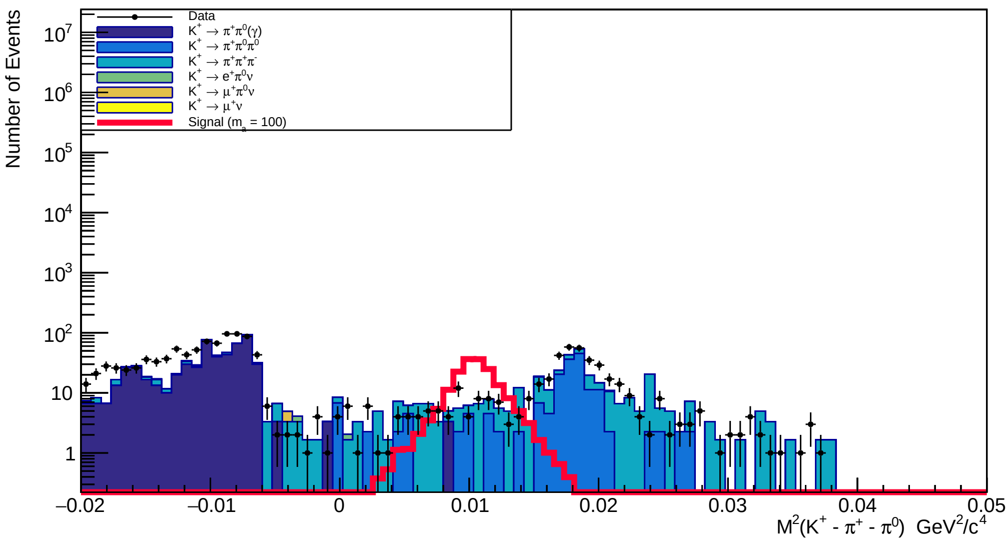


(b) Signal

Figure 6.16: p_T^2 vs m_{miss}^2 distribution for the MC samples of dominant background contributions, the data (section 6.3.4), and the signal under 4 different mass hypothesis (section 6.3.4), after the signal selection cuts are applied. The region below the dotted line is rejected.

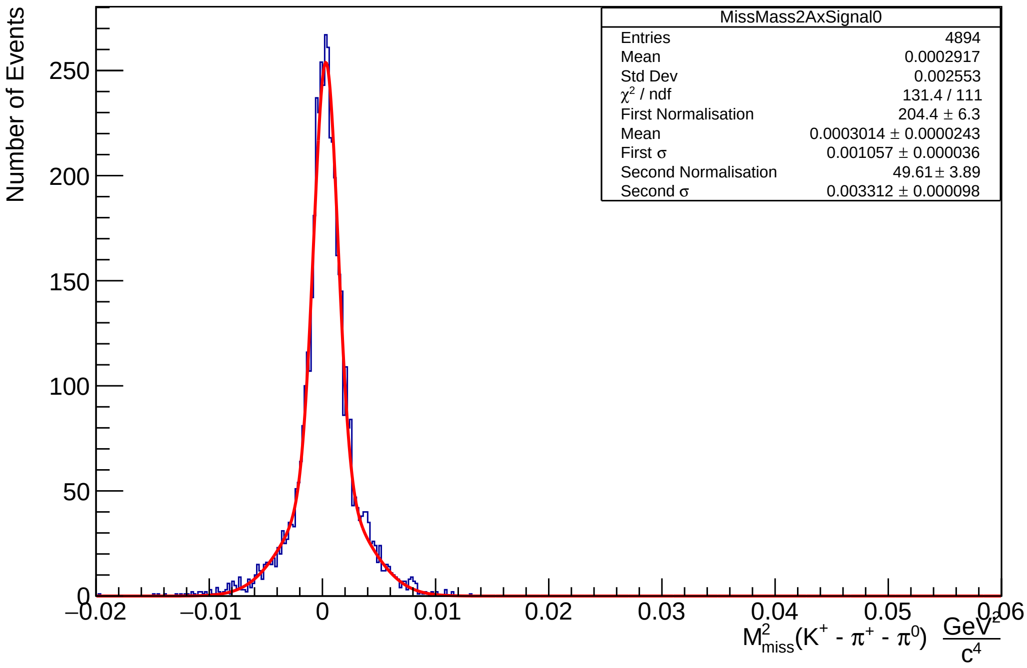


(a) $m_a = 0$

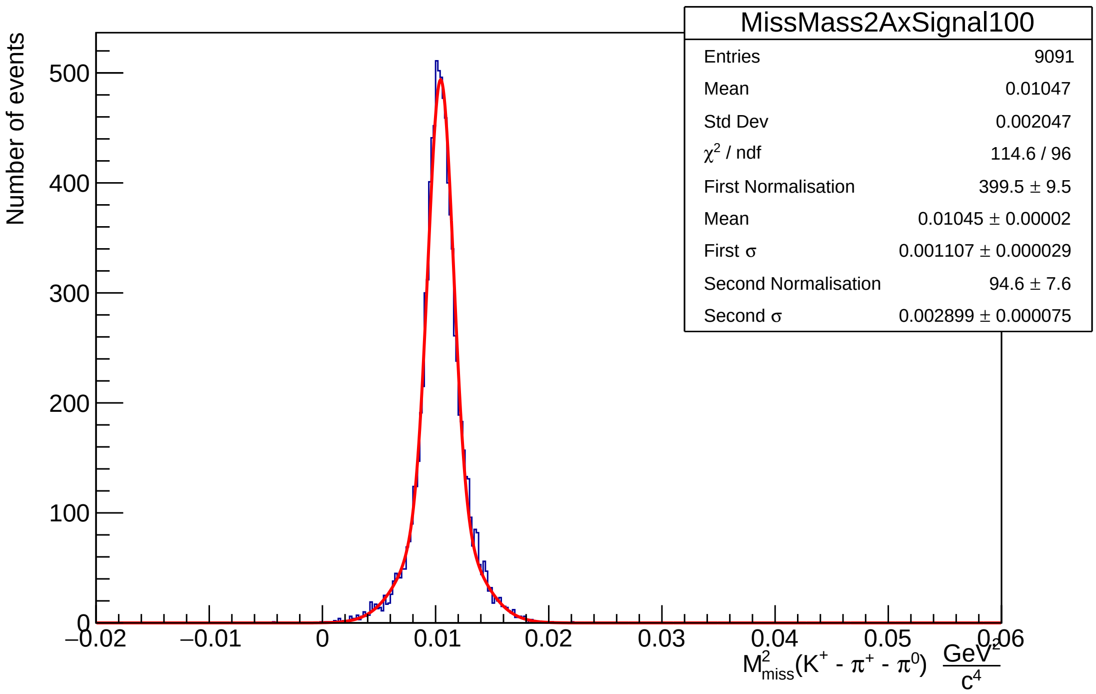


(b) $m_a = 100$ *MeV*$/c^2$

Figure 6.17: m^2_{miss} *distribution after all the cuts for the signal selection have applied. The MC samples for the background channels are stacked on one another, while the data and signal samples (for* m_a *= 0 in figure 6.17a and* $m_a = 100$ *MeV*$/c^2$ *in figure 6.17b) are drawn by themselves. Note how virtually all* $K_{2\pi}$ *events in the signal region have been eliminated.*



(a) $m_a = 0$



(b) $m_a = 100$ MeV/c^2

Figure 6.18: Double Gauss fit of the m^2_{miss} distribution for the two mass hypothesis ($m_a = 0$ in figure 6.18a and $m_a = 100$ MeV/c^2 in figure 6.18b).

Chapter 7

NA62 Sensitivity Estimate

In this chapter, the sensitivity of the NA62 experiment to the $K^+ \to \pi^+\pi^0 a$ decay is estimated. It is defined as the minimum BR which would be statistically distinguishable from fluctuations of the background-only case. First, we calculate the UL on the number of signal-like events in the background-only hypothesis using the CL_s method. We will also perform an estimate on the efficiency our selection has for signal events. Putting these two together yields the median UL on the BR, hence the sensitivity estimate.
The treatment in this chapter is based on the "Statistics" chapter of the PDG Review of Particle Physics [29] and references therein.

7.1 CL_s Method

Often, one uses the outcome of an experiment to draw conclusions on parameters of a model: the way to do so is to construct a function called the "estimator" which, given the data as an input, outputs an estimate of the parameter under consideration. In many cases, a point estimate for the parameter along with its expected error is sufficient; however, if the probability distribution function of the estimator is not Gaussian, or if there are physical boundaries on the parameters, this is not acceptable and an interval estimator is better suited. In a frequentist interpretation of probability, one constructs a so-called confidence interval, which is defined as the interval that includes (covers) the true value of the parameter with a probability α, called coverage probability or confidence level. The confidence interval is a function of both the data and the assumed model, so that repeated experiments will result in generally different confidence intervals. Indeed the confidence level is defined as the percentage of intervals that include the true parameter, in the limit of large number of repeated experiments.

For background-dominated experiments, this philosophy has the problem of excluding parameters outside the sensitivity of the experiment: this is in fact not an issue of the specific method used to define the interval, but it follows from the definition of coverage probability. By construction, one excludes true hypothesis with a probability of $1-\alpha$, leading to spurious exclusion of values that lie outside the sensitivity of the experimental design.
A solution to this problem was thought of in the late '90s in the context Higgs

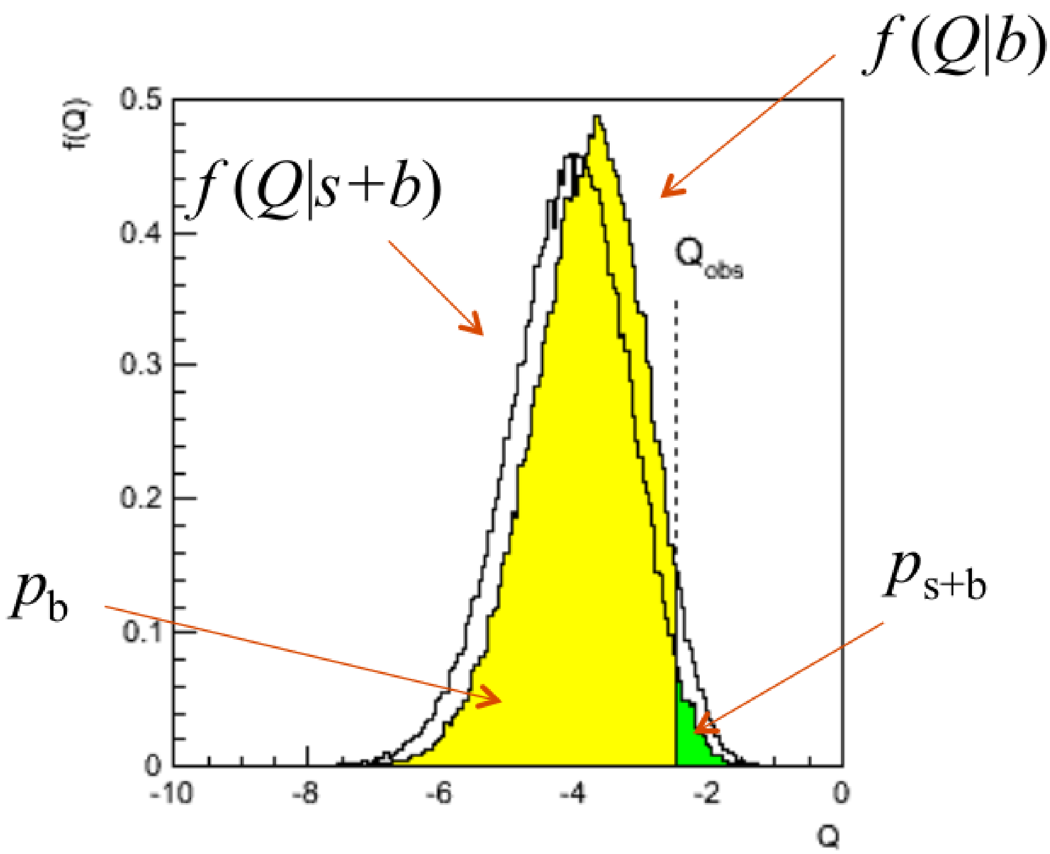


Figure 7.1: Example of great overlap between test statistic distribution of b and $s+b$ hypothesis: the CLs method avoids spurious exclusion in this cases. Credits: Klaus Reygers

searches at LEP, ultimately leading to a method which takes experimental limits into account by design, reducing spurious exclusion. This is the CL_s method.
The name is a bit of a misnomer, since it requires abandoning strict frequentist interpretation, thus a properly-defined confidence level. It should be noted

that, while this method is not on the strongest footing in statistical foundations, it has been used for more than two decades as the standard in new physics searches, producing sensible results. The deviation from the frequentist framework comes from the fact that it is designed to have over-coverage in the region where the experiment is known to be poorly sensitive. Because of this, it is usually called a modified frequentist method.
The basis of the CL_s procedure is the standard p-value test: take a test statistic Q and consider the probability distribution function for the background-only model $f(Q|b)$ and for the signal+background model $f(Q|s+b)$. The $s+b$ hypothesis is rejected if its p-value is less than the confidence level[1]:

$$p_{s+b} \stackrel{DEF}{=} \int_{Q_{obs}}^{\infty} dQ f(Q|s+b) < \alpha, \tag{7.1}$$

where Q_{obs} is the value of the test statistic calculated for the data set at hand. In a background-dominated experiment, however, $f(Q|b)$ and $f(Q|s+b)$ are not well separated, and thus we run into the problem described above. According to the CL_s method one defines an "effective" p-value

$$CL_s \stackrel{DEF}{=} \frac{p_{s+b}}{\int_{-\infty}^{Q_{obs}} dQ f(Q|b)} \stackrel{DEF}{=} \frac{p_{s+b}}{1-p_b} \tag{7.2}$$

and rejects the signal hypothesis when $CL_s < \alpha$. This has little effect when the distributions are well separated, so that we have $1 - p_b \approx 1$ for high values of p_{s+b}, but we get greater coverage with respect to the usual p-value when the distributions overlap greatly (see figure 7.1). This consciously-introduced over-coverage prevents spurious exclusion.
For our analysis, the log likelihood

$$Q = -2 \ln \frac{L(x|s+b)}{L(x|b)} \tag{7.3}$$

was chosen as the test statistic, in accordance with the Neyman–Pearson lemma, which guarantees that it is the most powerful. To compute p_{s+b} and p_b, Monte Carlo integration can be used: a number of toy experiment is performed where both the background and the signal are sampled according to a Poisson distribution. The background is modelled by a sum of Poissonian, one for every MC background sample in section 6.1, each with mean equal to the number of event from said MC sample passing the signal selection. For the

[1]NB: in standard treatments of the p-value method, one is concerned about rejecting the null — i.e. background-only — hypothesis. In that case, p_b is considered; the requirement for a discovery then reads $p_b < p$ for some previously-specified p.

signal+background model a poissonian with mean s is added to the background. $CL_s(s)$ is then calculated by performing 500000 toy experiment with the same value of s,. Finally, the procedure is repeated for all values of s we want to test; the (numeric) inverse of the equation

$$CL_s(s) = 1 - \alpha \tag{7.4}$$

gives the UL on the number of signal events s with confidence level α; in accordance to most new physics searches, we choose $\alpha = 0.9$. The results are shown in figure 7.2, the coloured bands showing the $\pm 1\sigma$ and $\pm 2\sigma$ regions, defined as the quantiles containing 68% and 95%, respectively, of the data points from the toy MC experiments.

7.2 Branching Ratio Sensitivity Estimate

The other quantity needed to estimate the BR sensitivity is the signal efficiency. It is calculated as:

$$\varepsilon_{sel} = \frac{N_{sel}}{N_{gen}} \tag{7.5}$$

with N_{sel} the number of event passing the signal selection, and N_{gen} the number of generated events, both number are measured on the signal MC sample.
The estimate for each mass hypothesis for the axion is shown in figure 7.3. Finally, we can estimate the BR sensitivity, using the two quantities described in this chapter and N_{Norm}, as estimated in section 6.2:

$$\text{BR sensitivity} = \frac{UL}{\varepsilon_{sel} N_{Norm}}. \tag{7.6}$$

The results are shown in figure 7.4. They ought to be compared to the results already present in literature, shown in figure 7.5: the comparison shows that NA62 is expected have much better sensitivity than these older experiments.

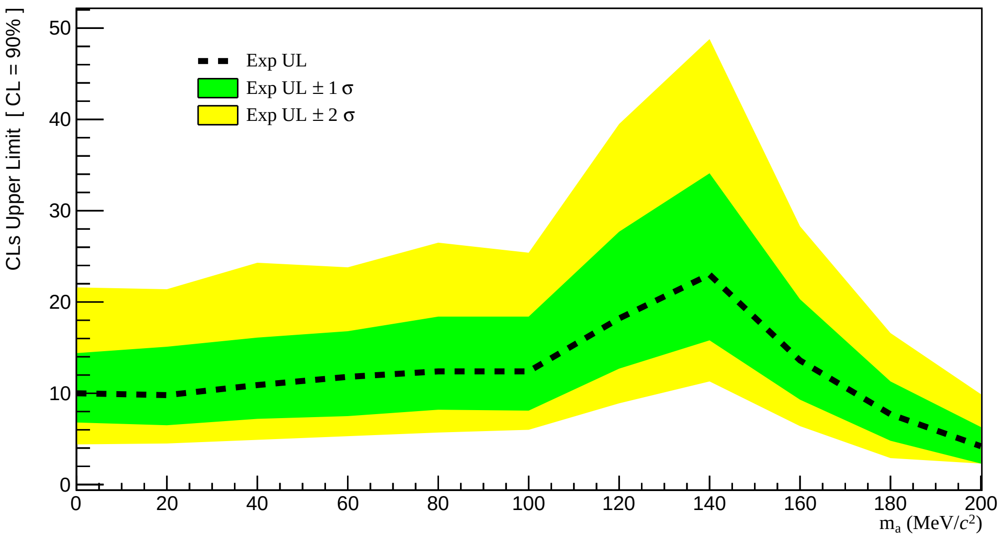


Figure 7.2: UL on the number of events as a function of m_a.

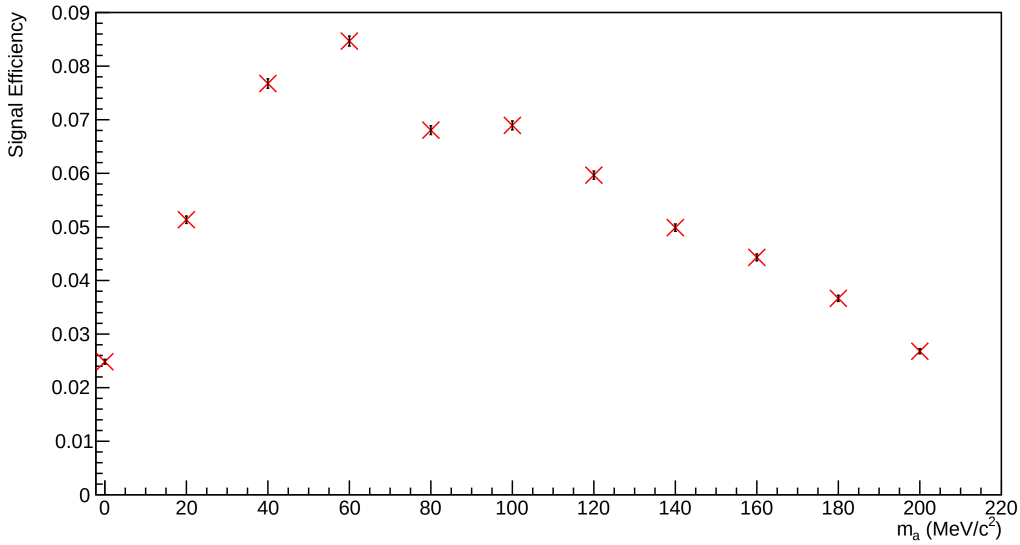


Figure 7.3: Signal Efficiency as a function of m_a

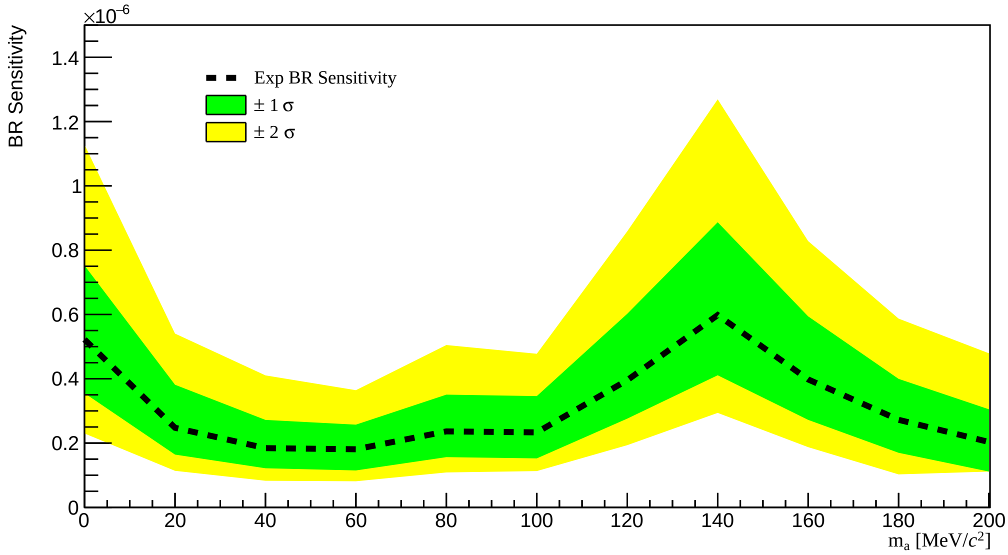


Figure 7.4: BR sensitivity estimate as a function of m_a

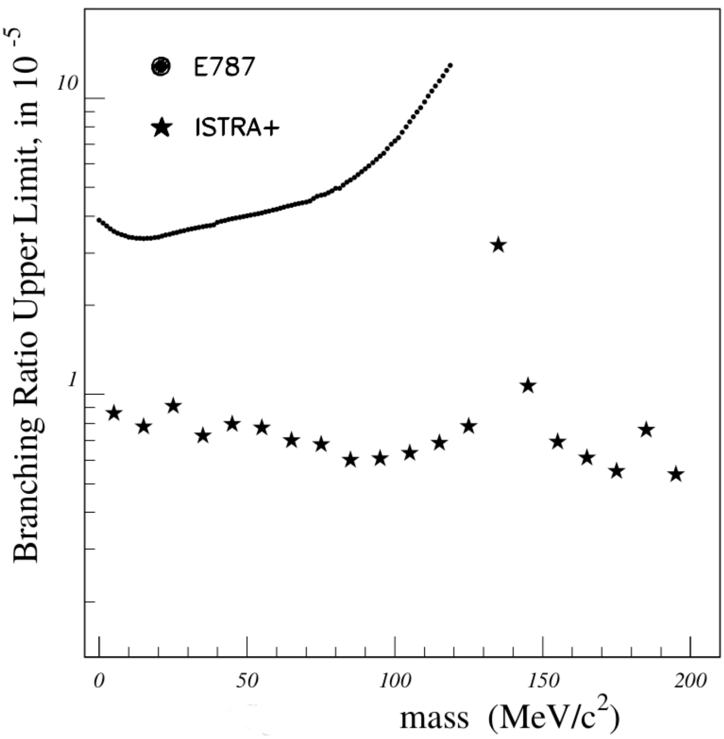


Figure 7.5: Upper limits on $K^+ \to \pi^+\pi^0 a$ decays from the E787 and ISTRA+ experiments. Taken from [58].

Chapter 8

Conclusions

The axion is widely thought to be among the best-motivated additions to the SM, providing a solution to the strong CP problem and being a possible DM candidate. In this thesis, we give a review of its properties — re-deriving some of them in the framework of χPT — and of the most common experimental techniques used in axion searches. Afterwards, we focus our attention on a specific family of models, dubbed axiflavon. Combining the QCD axion with a Froggatt–Nielsen flavon, they provide a solution to the flavour model, on top of being QCD axion models. For general axion models, the discovery potential of flavour physics experiments is in general not competitive with respect to astrophysical probes; the opposite is true for axiflavon models. Indeed, for the vectorial coupling of axiflavon to quarks, the Kaon factory NA62 can probe the DM window, a unique feat among particle physics experiments. However, an experimental effort to measure the corresponding axial-vector coupling at current experiments is, to date, lacking.
Motivated by this, the second and main part of this thesis is dedicated to the estimation of the sensitivity the NA62 experiment can have for the $K^+ \to \pi^+\pi^0 a$ decay, with a being any stable pseudoscalar. This is done using MC simulations, employing also a newly-written MC event generator, and the data from the 2017B data-taking period, corresponding to $\approx 8 \times 10^8$ K^+ decays in the fiducial volume. Our statistical analysis, which employs the CL_s method, yields an estimate of the BR sensitivity of $\mathcal{O}(10^{-7})$ in the mass range $[0, 200]$ MeV/c^2. This should be taken as a preliminary result, with areas of possible improvement already under examination in the NA62 collaboration. Nonetheless, it shows a sizeable margin of improvement (roughly 2 orders of magnitude) from current limits.

Appendix A

Useful results in Cosmology

In this appendix we state some results [60] from cosmology that are useful for present purposes. The Friedmann-Robertson-Walker metric is defined as:

$$ds^2 = dt^2 - R^2(t)d\mathbf{x}^2 \tag{A.1}$$

or, equivalently for cartesian coordinates:

$$g_{\mu\nu} = diag(1, -R^2(t), -R^2(t), -R^2(t)) \tag{A.2}$$

This implies that

$$\begin{aligned} \Box a = \frac{1}{\sqrt{-g}}\partial_\mu(\sqrt{-g}g^{\mu\nu}\partial_\nu a) = \\ = \frac{1}{R^3}[\partial_t(R^3\partial^t a) + \partial_i(-R^5\partial^i a)] = \ddot{a} + 3\frac{\dot{R}}{R}\dot{a} - R^2\nabla^2 a \end{aligned}$$

From the scale factor R we can define the Hubble parameter $H = \dot{R}/R$; in standard cosmology we have that:

$$H = \frac{1}{M_{PL}}\sqrt{\frac{8\pi\rho}{3}}$$
$$\rho = \frac{g_* \pi^2 T^4}{30}$$
$$s = \frac{2\pi^2}{45}g_{*S}T^3$$

where g_* is the number of effective degrees of freedom and s is the entropy per unit volume.

Solving the Friedmann equation can let us find the time-dependence of these parameter, hence enabling us to freely exchange t and T dependence.

$$H \propto \frac{1}{t}$$
$$\rho \propto \frac{1}{\sqrt{t}}$$

A.0.1 Boltzmann equation

The equation governing the thermal evolution of a particle phase space distribution f is the Boltzmann Equation, which in its most general form reads [61, 60]

$$\mathbb{L}[f] = \mathbb{C}[f] \tag{A.3}$$

where the LHS is the Liouville operator, while the RHS is the Collision operator. They are aptly named, as the Liouville operator is the one that enters Liouville's Theorem and thus describes the internal evolution of the phase space distribution, while the Collision operator describes the evolution due to scattering between particles. Both of them can be understood as the variation in particle number per unit time and covariant volume: $d^2N/dVdT$
In curved space, the Liouville operator is written as

$$\mathbb{L} = p_\mu \partial_x^\mu - \Gamma_\alpha^{\beta\gamma} p_\beta p_\gamma \partial_p^\alpha \overset{FRW}{=} E\dot{f} - \frac{\dot{R}}{R}|\vec{p}|^2 \frac{\partial f}{\partial E} \tag{A.4}$$

Where $\Gamma_\alpha^{\beta\gamma}$ is the Christoffel symbol and in the last equality we used the FRW metric.
Armed with this, we can cast Boltzmann equation into an equation for the number density $n = \frac{g}{(2\pi)^3} \int d^3p\, f(|\vec{p}|, t)$:

$$\frac{dn}{dt} + 3\frac{\dot{R}}{R}n = \frac{g}{(2\pi)^3} \int d^3p\, \mathbb{C}[f] \tag{A.5}$$

We can then see that if the Collision term can be neglected w.r.t to H , the number density per comoving volume is constant (i.e. $n \propto R^{-3}$).

A.0.2 Freeze-out mechanism

The freeze-out mechanism explains how massive, nonrelativistic species can have non-negligible present abundance: for a specie that is in contact with

the thermal bath one expects that the Boltzmann factor $e^{-\frac{m}{T}}$ would wash out any particle with mass $\gg T_0 \approx 2 \times 10^{-4} eV$. If the particle is not coupled to the bath, this argument fails and the abundance can be higher.
The intuition is that for interaction rates $\Gamma < H$, the particles cannot "find each other" since the universe is expanding too rapidly, and as such the interaction becomes too weak for thermal equilibrium to be mantained; the interaction is said to freeze-out. If this freeze-out happens when $\frac{m}{T}$ is close enough to 1, the relic abundance can be high enough to match experimental observations.

A.0.3 d'Alembert operator in FRW metric

For a scalar field the dalembertian reads:

$$\Box\phi = \frac{1}{\sqrt{-g}}\partial_\mu(\sqrt{-g}g^{\mu\nu}\partial_\nu\phi) \tag{A.6}$$

using the FRW metric we get

$$\Box\phi = \frac{1}{R^3}[\partial_0(R^3\partial^0\phi) - \partial_i(R^5\partial^i\phi)] = 3\frac{\dot{R}}{R}\dot{\phi} + \ddot{\phi} - R^2\nabla^2\phi \tag{A.7}$$

Bibliography

[1] R. D. Peccei and H. R. Quinn, “CP Conservation in the Presence of Instantons,” *Phys. Rev. Lett.*, vol. 38, pp. 1440–1443, 1977.

[2] F. Wilczek, “Problem of Strong P and T Invariance in the Presence of Instantons,” *Phys. Rev. Lett.*, vol. 40, pp. 279–282, 1978.

[3] S. Weinberg, “A New Light Boson?,” *Phys. Rev. Lett.*, vol. 40, pp. 223–226, 1978.

[4] A. Arvanitaki, S. Dimopoulos, S. Dubovsky, N. Kaloper, and J. March-Russell, “String Axiverse,” *Phys. Rev. D*, vol. 81, p. 123530, 2010.

[5] S. B. Treiman, E. Witten, R. Jackiw, and B. Zumino, *Current Algebra and Anomalies*. Princeton University Press, 1986.

[6] I. Low and Z. Yin, “The Infrared Structure of Nambu-Goldstone Bosons,” *JHEP*, vol. 10, p. 078, 2018.

[7] S. L. Adler, “Consistency conditions on the strong interactions implied by a partially conserved axial vector current,” *Phys. Rev.*, vol. 137, pp. B1022–B1033, 1965.

[8] K. Fujikawa, “Evaluation of the chiral anomaly in gauge theories with γ_5 couplings,” *Phys. Rev. D*, vol. 29, pp. 285–292, Jan 1984.

[9] A. Hook, “TASI Lectures on the Strong CP Problem and Axions,” *PoS*, vol. TASI2018, p. 004, 2019.

[10] J. M. Pendlebury *et al.*, “Revised experimental upper limit on the electric dipole moment of the neutron,” *Phys. Rev. D*, vol. 92, no. 9, p. 092003, 2015.

[11] C. Vafa and E. Witten, “Restrictions on symmetry breaking in vector-like gauge theories,” *Nucl. Phys. B*, vol. 234, no. 1, pp. 173–188, 1984.

[12] J. R. Ellis and M. K. Gaillard, “Strong and Weak CP Violation,” *Nucl. Phys. B*, vol. 150, pp. 141–162, 1979.

[13] A. Salvio, A. Strumia, and W. Xue, “Thermal axion production,” *JCAP*, vol. 2014, pp. 011–011, jan 2014.

[14] J. Preskill, M. B. Wise, and F. Wilczek, “Cosmology of the Invisible Axion,” *Phys. Lett. B*, vol. 120, pp. 127–132, 1983.

[15] M. Dine and W. Fischler, “The Not So Harmless Axion,” *Phys. Lett. B*, vol. 120, pp. 137–141, 1983.

[16] D. J. E. Marsh, “Axion Cosmology,” *Phys. Rep.*, vol. 643, pp. 1–79, 2016.

[17] N. Blinov, M. J. Dolan, P. Draper, and J. Kozaczuk, “Dark matter targets for axionlike particle searches,” *Phys. Rev. D*, vol. 100, p. 015049, Jul 2019.

[18] K. Zioutas *et al.*, “Axion searches with helioscopes and astrophysical signatures for axion(-like) particles,” *New Journal of Physics*, vol. 11, 03 2009.

[19] G. G. Raffelt, “Astrophysical methods to constrain axions and other novel particle phenomena,” *Phys. Rep.*, vol. 198, pp. 1–113, 1990.

[20] V. Anastassopoulos *et al.*, “New CAST Limit on the Axion-Photon Interaction,” *Nature Phys.*, vol. 13, pp. 584–590, 2017.

[21] R. Ballou *et al.*, “New exclusion limits on scalar and pseudoscalar axion-like particles from light shining through a wall,” *Phys. Rev. D*, vol. 92, no. 9, p. 092002, 2015.

[22] S. J. Asztalos *et al.*, “A SQUID-based microwave cavity search for dark-matter axions,” *Phys. Rev. Lett.*, vol. 104, p. 041301, 2010.

[23] M. S. Turner, “Axions from SN 1987a,” *Phys. Rev. Lett.*, vol. 60, p. 1797, 1988.

[24] G. Raffelt and D. Seckel, “Bounds on Exotic Particle Interactions from SN 1987a,” *Phys. Rev. Lett.*, vol. 60, p. 1793, 1988.

[25] G. G. Raffelt, “Astrophysical axion bounds,” *Lect. Notes Phys.*, vol. 741, pp. 51–71, 2008.

[26] F. Calore, P. Carenza, M. Giannotti, J. Jaeckel, and A. Mirizzi, "Bounds on axionlike particles from the diffuse supernova flux," *Phys. Rev. D*, vol. 102, p. 123005, Dec 2020.

[27] J. Martin Camalich, M. Pospelov, P. N. H. Vuong, R. Ziegler, and J. Zupan, "Quark Flavor Phenomenology of the QCD Axion," *Phys. Rev. D*, vol. 102, no. 1, p. 015023, 2020.

[28] R. Ziegler, "Flavored Axions," *arXiv e-prints*, p. arXiv:1905.01084, May 2019.

[29] P. Zyla *et al.*, "Review of Particle Physics," *PTEP*, vol. 2020, no. 8, p. 083C01, 2020.

[30] C. D. Froggatt and H. B. Nielsen, "Hierarchy of Quark Masses, Cabibbo Angles and CP Violation," *Nucl. Phys. B*, vol. 147, pp. 277–298, 1979.

[31] L. Calibbi, F. Goertz, D. Redigolo, R. Ziegler, and J. Zupan, "Minimal axion model from flavor," *Phys. Rev. D*, vol. 95, no. 9, p. 095009, 2017.

[32] J. Bijnens, G. Colangelo, and J. Gasser, "K_{l4} decays beyond one loop," *Nucl. Phys. B*, vol. 427, pp. 427–454, 1994.

[33] M. Knecht, H. Sazdjian, J. Stern, and N. H. Fuchs, "A Possible experimental determination of $m_s/\hat{m}$ from Kmu4 decays," *Phys. Lett. B*, vol. 313, pp. 229–233, 1993.

[34] J. Bijnens, "K_{l4} decays and the low-energy expansion," *Nucl. Phys. B*, vol. 337, no. 3, pp. 635–651, 1990.

[35] S. Weinberg, "Current Commutator Calculation of the K_{l4} form Factors," *Phys. Rev. Lett.*, vol. 17, pp. 336–340, 1966. [Erratum: Phys.Rev.Lett. 18, 1178 (1967)].

[36] Y. Grossman and Y. Nir, "$K_L \to \pi^0\nu\nu$ beyond the standard model," *Phys. Lett. B*, vol. 398, no. 1, pp. 163–168, 1997.

[37] S. Weinberg, "Phenomenological lagrangians," *Physica A: Statistical Mechanics and its Applications*, vol. 96, no. 1, pp. 327–340, 1979.

[38] G. Ecker, "Chiral perturbation theory," *Prog. Part. Nucl. Phys.*, vol. 35, pp. 1–80, 1995.

[39] A. Manohar and H. Georgi, "Chiral Quarks and the Nonrelativistic Quark Model," *Nucl. Phys. B*, vol. 234, pp. 189–212, 1984.

[40] S. Scherer, “Introduction to chiral perturbation theory,” *Adv. Nucl. Phys.*, vol. 27, p. 277, 2003.

[41] A. Pich, “Chiral perturbation theory,” *Reports on Progress in Physics*, vol. 58, pp. 563–609, jun 1995.

[42] G. Buchalla, O. Catá, and C. Krause, “On the Power Counting in Effective Field Theories,” *Phys. Lett. B*, vol. 731, pp. 80–86, 2014.

[43] G. Ecker, J. Gasser, H. Leutwyler, A. Pich, and E. de Rafael, “Chiral lagrangians for massive spin 1 fields,” *Phys. Lett. B*, vol. 223, pp. 425–432, 1989.

[44] H. Georgi, D. B. Kaplan, and L. Randall, “Manifesting the Invisible Axion at Low-energies,” *Phys. Lett. B*, vol. 169, pp. 73–78, 1986.

[45] M. Bauer, M. Neubert, S. Renner, M. Schnubel, and A. Thamm, “The low-energy effective theory of axions and ALPs,” *JHEP*, vol. 2021, Apr 2021.

[46] G. G. di Cortona, E. Hardy, J. P. Vega, and G. Villadoro, “The QCD axion, precisely,” *JHEP*, vol. 2016, Jan 2016.

[47] J. Gasser and H. Leutwyler, “Chiral perturbation theory: Expansions in the mass of the strange quark,” *Nucl. Phys. B*, vol. 250, no. 1, pp. 465–516, 1985.

[48] M. Bauer, M. Neubert, S. Renner, M. Schnubel, and A. Thamm, “Consistent treatment of axions in the weak chiral lagrangian,” 2021.

[49] B. Grinstein, S.-J. Rey, and M. B. Wise, “CP violation in charged-kaon decay,” *Phys. Rev. D*, vol. 33, pp. 1495–1498, Mar 1986.

[50] C. Bernard, T. Draper, A. Soni, H. D. Politzer, and M. B. Wise, “Application of chiral perturbation theory to K$\rightarrow 2\pi$ decays,” *Phys. Rev. D*, vol. 32, pp. 2343–2347, Nov 1985.

[51] G. Colangelo and G. Isidori, “An Introduction to ChPT,” *Frascati Phys. Ser.*, vol. 18, pp. 333–376, 2000.

[52] V. Cirigliano, G. Ecker, H. Neufeld, A. Pich, and J. Portolés, “Kaon decays in the standard model,” *Rev. Mod. Phys.*, vol. 84, p. 399–447, Mar 2012.

[53] I. Antoniadis and T. N. Truong, “Lower bound for branching ratio of $K^+ \rightarrow \pi^+$ axion and non-existence of Peccei-Quinn axion,” *Physics Letters B*, vol. 109, no. 1, pp. 67–72, 1982.

[54] M. Shifman, A. Vainshtein, and V. Zakharov, "Can confinement ensure natural CP invariance of strong interactions?," *Nucl. Phys. B*, vol. 166, no. 3, pp. 493–506, 1980.

[55] The NA62 Collaboration (E. Cortina Gil et al.), "The beam and detector of the NA62 experiment at CERN," *JIST*, vol. 12, pp. P05025–P05025, May 2017.

[56] A. J. Buras, D. Buttazzo, J. Girrbach-Noe, and R. Knegjens, " $K^+ \to \pi^+ \nu\overline{\nu}$ and $K_L \to \pi^0 \nu\overline{\nu}$ in the Standard Model: status and perspectives," *JHEP*, vol. 2015, Nov 2015.

[57] The NA62 Collaboration (E. Cortina Gil et al.), "An investigation of the very rare $K^+ \to \pi^+ \nu\overline{\nu}$ decay," *JHEP*, vol. 11, p. 042, 2020.

[58] O. G. Tchikilev *et al.*, "Search for light pseudoscalar sgoldstino in K-decays," *Phys. Lett. B*, vol. 602, pp. 149–156, 2004.

[59] R. I. Marchevski, *First measurement of the $K^+ \to \pi^+ \nu\overline{\nu}$ decay with the NA62 experiment at CERN*. PhD thesis, Jul 2018. Presented 19 Feb 2019.

[60] E. Kolb, *The Early Universe (Frontiers in Physics)*. Westview Press, Jan 1994.

[61] K. Huang, *Statistical Mechanics, 2nd Edition*. John Wiley & Sons, May 1987.