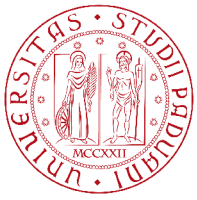


Sede Amministrativa: Università degli Studi di Padova
Dipartimento di Fisica ed Astronomia "G. Galilei"

CORSO DI DOTTORATO DI RICERCA IN: FISICA

CICLO XXX

Measurement of angular parameters from the decay $B^0 \to K^{*0}\mu^+\mu^-$ in proton-proton collisions at $\sqrt{s} = 8$ TeV

Tesi redatta con il contributo finanziario
dell'Istituto Nazionale di Fisica Nucleare

Coordinatore: Ch.mo Prof. Gianguido Dall'Agata

Supervisore: Dr. Paolo Ronchese

Co-Supervisore: Dr. Stefano Lacaprara

Dottorando: Alessio Boletti

Contents

Abstract

Angular distributions of the decay $B^0 \to K^{*0}\mu^+\mu^-$ are studied using a sample of proton-proton collisions at $\sqrt{s} = 8$ TeV collected with the CMS detector at the LHC, corresponding to an integrated luminosity of 20.5 fb^{-1}. An angular analysis is performed to determine the P_1 and P_5' parameters, where the P_5' parameter is of particular interest because of recent measurements that indicate a potential discrepancy with the standard model predictions. Based on a sample of 1397 signal events, the P_1 and P_5' parameters are determined as a function of the dimuon invariant mass squared. The measurements are in agreement with predictions based on the standard model.

Chapter 1

Analysis introduction

In this thesis, I will present the angular analysis of the $B^0 \to K^{*0}\mu^+\mu^-$ decay, performed with the data collected by the Compact Muon Solenoid (CMS) Experiment in the 2012 run of proton-proton (pp) collisions at a centre-of-mass energy $\sqrt{s} = 8\,\text{TeV}$, and corresponding to an integrated luminosity of $20.5\,\text{fb}^{-1}$.

Two analyses of the $B^0 \to K^{*0}\mu^+\mu^-$ decay have been performed by the CMS Collaboration, using the data described above. The first analysis [25] was aiming to measure the branching fraction and to perform a partial measurement of the angular distribution. The K^{*0} longitudinal polarisation fraction, F_L, and the forward-backward asymmetry of the dimuon system, A_{FB}, were estimated as a function of the dimuon invariant mass squared, q^2. A combination with the results obtained with pp collision data collected in 2011, corresponding to an integrated luminosity of $5.2\,\text{fb}^{-1}$ at $\sqrt{s} = 7\,\text{TeV}$, was also performed. The results of this analysis are in good agreement with the Standard Model (SM) predictions and with the results from previous experiments: BaBar [14], Belle [46], CDF [8], and LHCb [2].

The first complete angular analysis was performed by the LHCb Collaboration [3], with the data collected in 2011 pp collisions at $\sqrt{s} = 7\,\text{TeV}$. In this analysis the measurement has been extended to the full set of angular parameters, which will be defined in Section 1.2.1. A tension with respect to the SM predictions has been found in the measurement of one of these parameters, P_5'. The significance of this tension, without taking into account the look-elsewhere effect, is reported to be 3.7 standard deviations.

The LHCb collaboration published also a complete angular analysis [7], combining the data collected in 2011 and 2012 pp collisions for an overall integrated luminosity of $3\,\text{fb}^{-1}$. The results for the P_5' parameter were still showing a discrepancy with respect to the predictions, as shown in Figure 1.1. The tension is localised in the q^2 region between $4\,\text{GeV}^2$ and $8\,\text{GeV}^2$, and its significance is evaluated, using the full set of parameters, to be of about 3.4 standard deviations.

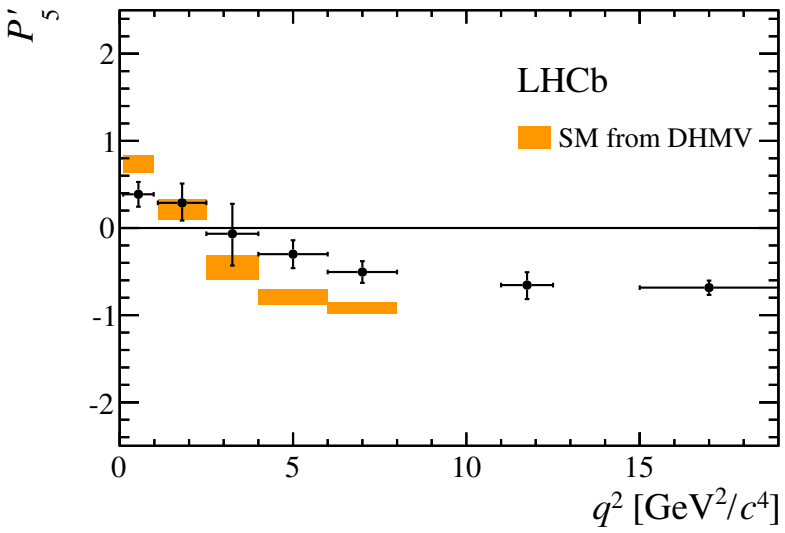


Figure 1.1: Results of P'_5 angular parameter in bins of q^2 from LHCb Run 1 data, compared with SM-based predictions. Image from Reference [7].

After these results, the Belle Collaboration performed an extension [10] of its original analysis, to measure the full set of angular parameters, in particular P'_5. The results are in agreement with the measurements of LHCb, but with a larger uncertainty that prevent to obtain a significant discrepancy of the P'_5 results with respect to the SM predictions.

The second analysis from the CMS Collaboration [42] is an extension of the first analysis, with the goal of measuring the P'_5 angular parameter to provide a new independent experimental result. The expected precision of this measurement is not as good as the latest result of the LHCb Collaboration, given the lower number of signal events and worse signal over background ratio. Anyway, it will have a smaller uncertainty and a q^2 bin structure more similar to the LHCb one, with respect to the result from Belle Collaboration. For this reason, these results could have an important role in the process of shedding light on the matter.

In this thesis I will describes the details of this second analysis of the CMS Collaboration. Chapter 1 will continue with the description of the theoretical framework used to described the decays like $B^0 \to K^{*0}\mu^+\mu^-$, and with the formulation of the angular distribution and few considerations about it. In Chapter 2 the Large Hadron Collider (LHC) and the CMS experiment will be described, focusing in particular on the aspects more related to the analyses of Flavour Physics. The selection criteria used to reduce the contamination from background events will be presented in Chapter 3, while in Chapter 4 the procedure used to extract the angular parameters from the selected data will be explained in details. Chapter 5 will be dedicated to the methods used

to describe the efficiency of the selection criteria on signal events. Several tests have been performed to validate the analysis procedure; they will be described in Chapter 6. The sources of systematic uncertainty that have been investigated will be listed in Chapter 7, while the study of the statistical uncertainty and the results of the analysis will be reported in Chapter 8. Finally, I will conclude in Chapter 9 by describing the future prospects for this analysis and focusing in particular on the ongoing efforts to improve it for the data from LHC Run 2 collisions.

1.1 Flavour Physics

The sector of the Standard Model of particle physics in which the precision of the experimental probes reached higher levels is the Flavour Physics.

In particular, it is the realm of the Cabibbo-Kobayashi-Maskawa (CKM) matrix, i.e. the 3×3 unitary matrix that rules the charge current weak interactions between quarks. This matrix was proposed by Kobayashi and Maskawa about 45 years ago, extending the Cabibbo 2×2 matrix to allow a source of CP-violation to be included in the model through a non-vanishing phase (that is not possible in a 2×2 unitary matrix). It allowed to foresee the existence of the three heavier quarks before discovering them in the experimental facilities. Its parameters, that link the basis of weak eigenstates with the one of Yukawa eigenstates, have been determined in the past decades with great precision.

The unitarity of the CKM matrix is the real testing ground of the completeness of the theory. An observation of non-unitary behaviours in the CKM parameters would open the gates to scenarios beyond the SM, suggesting the presence of a forth generation of quarks. However, after many years of B-factory probes, and few years of analyses at LHC, no cracks have been found in the paradigm. This robust long-standing structure is not promising for searches of phenomena of New Physics (NP).

On the other hand, few recent results from the LHCb Collaboration raised the interest for the world of the so-called rare decays. One result have been already mentioned before: the tension of the P_5' parameter measured in the angular analysis of the flavour-changing neutral current (FCNC) $\mathrm{B}^0 \to \mathrm{K}^{*0}\mu^+\mu^-$ decay. Another, very interesting, result that is showing a discrepancy with respect to the predictions is the measurement [6] of $R(\mathrm{K}^{(*)})$. These parameters are defined as the branching-fraction ratios:

$$\begin{aligned} R(\mathrm{K}^*) &= \frac{\mathcal{B}r(\mathrm{B}^0 \to \mathrm{K}^{*0}\mu^+\mu^-)}{\mathcal{B}r(\mathrm{B}^0 \to \mathrm{K}^{*0}e^+e^-)} \\ R(\mathrm{K}) &= \frac{\mathcal{B}r(\mathrm{B}^+ \to \mathrm{K}^{+}\mu^+\mu^-)}{\mathcal{B}r(\mathrm{B}^+ \to \mathrm{K}^{+}e^+e^-)} \end{aligned} \tag{1.1}$$

expressed as a function of the dimuon invariant mass squared, q^2.

In the latest LHCb results for both these ratios, a hint for tension with respect to the SM is present. Even if this discrepancy is less significant than the one on P_5', the theoretical prediction for this variables are not affected at all by hadronic uncertainties, and it makes them very robust. One very interesting aspect of these discrepancies is that they could be related to a common cause, affecting one of the coefficients of the effective Hamiltonian, the so-called Wilson coefficients.

In the following section, a brief description of the theoretical framework used to described these FCNC decays is given.

1.1.1 Flavour-changing neutral currents $b \to s\ell^+\ell^-$

The penguin-mediated FCNC $b \to s\ell^+\ell^-$ are a set of b-hadron semileptonic decays with a pair of non-resonant charged leptons in the final state. These decay channels are doubly suppressed within the SM:

- they are forbidden at tree-level, since there is no neutral current in SM allowed to violate the flavour, and the leading-order Feynman diagram that mediates them is a weak penguin loop;
- the leading-order diagram is Cabibbo suppressed, since it is proportional to the CKM elements $|V_{ts}V_{tb}| \sim 10^{-2}$.

These causes of suppression result in small branching fractions and they create an ideal environment for NP searches. Any potential contribution from physics beyond the SM that enters in the loop diagram could produce sizeable effects in the decay branching fractions or in their angular distributions.

On the other hand, the theoretical predictions of the angular distributions and branching fractions of these decays are quite susceptible to hadronic uncertainties due to long-distance Quantum Chromodynamic (QCD) processes. In order to separate, in the theoretical framework, the effect from long-distance QCD and from short-distance effects, which are related to QCD and electroweak interactions but also to NP processes, an effective Hamiltonian is defined as follow:

$$\mathcal{H}_{\text{eff}} = -\frac{4\,G_F}{\sqrt{2}}\left(\lambda_t \mathcal{H}_{\text{eff}}^{(t)} + \lambda_u \mathcal{H}_{\text{eff}}^{(u)}\right) \tag{1.2}$$

where $\lambda_i = V_{ib}V_{is}^*$ and

$$\begin{aligned}
\mathcal{H}_{\text{eff}}^{(t)} &= C_1\mathcal{O}_1^c + C_2\mathcal{O}_2^c + \sum_{i=3}^{6} C_i\mathcal{O}_i + \sum_{i=7,8,9,10,P,S} (C_i\mathcal{O}_i + C_i'\mathcal{O}_i')\,, \\
\mathcal{H}_{\text{eff}}^{(u)} &= C_1(\mathcal{O}_1^c - \mathcal{O}_1^u) + C_2(\mathcal{O}_2^c - \mathcal{O}_2^u)\,.
\end{aligned}$$

The operators $\mathcal{O}_i^{(a)}$ for $i < 7$ are equal to the P_i defined in the Reference [21], while for $i \geq 7$ they are defined accordingly to Reference [12]:

$$\mathcal{O}_7 = \frac{e}{g^2} m_b (\bar{s}\sigma_{\mu\nu} P_R b) F^{\mu\nu} \qquad \mathcal{O}'_7 = \frac{e}{g^2} m_b (\bar{s}\sigma_{\mu\nu} P_L b) F^{\mu\nu} \tag{1.3}$$

$$\mathcal{O}_8 = \frac{1}{g} m_b (\bar{s}\sigma_{\mu\nu} T^a P_R b) G^{\mu\nu\, a} \qquad \mathcal{O}'_8 = \frac{1}{g} m_b (\bar{s}\sigma_{\mu\nu} T^a P_L b) G^{\mu\nu\, a} \tag{1.4}$$

$$\mathcal{O}_9 = \frac{e^2}{g^2} (\bar{s}\gamma_\mu P_L b)(\bar{\mu}\gamma^\mu \mu) \qquad \mathcal{O}'_9 = \frac{e^2}{g^2} (\bar{s}\gamma_\mu P_R b)(\bar{\mu}\gamma^\mu \mu) \tag{1.5}$$

$$\mathcal{O}_{10} = \frac{e^2}{g^2} (\bar{s}\gamma_\mu P_L b)(\bar{\mu}\gamma^\mu \gamma_5 \mu) \qquad \mathcal{O}'_{10} = \frac{e^2}{g^2} (\bar{s}\gamma_\mu P_R b)(\bar{\mu}\gamma^\mu \gamma_5 \mu) \tag{1.6}$$

$$\mathcal{O}_S = \frac{e^2}{16\pi^2} m_b (\bar{s} P_R b)(\bar{\mu}\mu) \qquad \mathcal{O}'_S = \frac{e^2}{16\pi^2} m_b (\bar{s} P_L b)(\bar{\mu}\mu) \tag{1.7}$$

$$\mathcal{O}_P = \frac{e^2}{16\pi^2} m_b (\bar{s} P_R b)(\bar{\mu}\gamma_5 \mu) \qquad \mathcal{O}'_P = \frac{e^2}{16\pi^2} m_b (\bar{s} P_L b)(\bar{\mu}\gamma_5 \mu) \tag{1.8}$$

where g is the strong coupling constant, m_b is the running b quark mass in the $\overline{\text{MS}}$ scheme, and $P_{L,R} = (1 \mp \gamma_5)/2$ are the chirality projectors.

The set of Wilson coefficients $C_i^{(\prime)}$ encodes the contribution from short-distance physics and could contain NP effects. These coefficients can be expanded and calculated in perturbation theory in powers of $\alpha_s \equiv g^2/4\pi$:

$$C_i = C_i^{(0)} + \frac{\alpha_s}{4\pi} C_i^{(1)} + \left(\frac{\alpha_s}{4\pi}\right)^2 C_i^{(2)} + O(\alpha_s^3)\,, \tag{1.9}$$

Their predicted values of these expanded coefficients, both within SM and for NP scenarios, are calculated at the scale $\mu = m_W$ and then evolved down to $\mu \sim m_b$.

According to the SM, several of the terms in the effective Hamiltonian are expected to vanish or be highly suppressed; this is the case for all the primed operators, as well as for the scalar and pseudoscalar ones, $\mathcal{O}_S$ and $\mathcal{O}_P$.

If we restrict the study to a $\mathrm{B}^{(0,\pm)} \to \mathrm{K}^{*(0,\pm)} \ell^+ \ell^-$ decay, the matrix elements of the effective Hamiltonian operators $\mathcal{O}^{(\prime)}_{7,9,10,S,P}$ can be described as a function of seven form factors, $A_i(q^2)$, $V(q^2)$, and $T_j(q^2)$, where $0 \leq i \leq 2$ and $1 \leq j \leq 2$, and q^2 is the squared momentum carried by the pair of leptons. In literature the form factor $A_3(q^2)$, which is a linear combination of $A_1(q^2)$ and $A_2(q^2)$, is also used.

The theoretical predictions of the form factors are computed using the QCD sum rules on the light-cone (LCSRs) [37]. With this technique, the form factors are expanded in powers of $m_b m_{\mathrm{K}^*}/(m_b^2 - q^2)$. This expansion works well for values of q^2 lower than about $8\,\mathrm{GeV}^2$, but for higher values the high-order terms become more relevant and the uncertainty associated to the approximated prediction grows accordingly.

In the $B^{(0,\pm)} \to K^{*(0,\pm)}\ell^+\ell^-$ decay amplitude, not only the terms proportional to the form factors are contributing. It contains additional "non-factorisable effects", which are related to the matrix elements of the operators not considered when defining the form factors: the purely hadronic operators $\mathcal{O}_{1,2,3,4,5,6}$ and the chromomagnetic operators $\mathcal{O}_8^{(\prime)}$, where an additional virtual photon emission is needed to produce the lepton pair.

The contribution of these effects cannot be expressed as a function of a new set of form factors. In the combined heavy-quark and large-energy limits, they can be calculated using the QCD factorisation technique [18–20].

1.2 The decay $B_d^0 \to K^{*0}\mu^+\mu^-$

The FCNC decay $B^0 \to K^{*0}(K^+\pi^-)\mu^+\mu^-$ is an optimal laboratory to probe the flavour sector of the SM. From the experimental point of view, it has a fully charged final state, with two muons, which are easy to identify in a multi-purpose particle detector, and two charged hadronic particles. Furthermore, the charges of the hadrons in the final state determine the CP-state of the decay, i.e. whether it is $B^0 \to K^{*0}(K^+\pi^-)\mu^+\mu^-$ or $\overline{B}^0 \to \overline{K}^{*0}(K^-\pi^+)\mu^+\mu^-$. From the phenomenological point of view, the angular analysis of this decay can provide information on several components of the effective Hamiltonian described in Section 1.1.1: the electromagnetic and semileptonic operators, $\mathcal{O}_{7,9,10}$, and their chirality-flipped counterparts, $\mathcal{O}_{7',9',10'}$, together with scalar and pseudoscalar operators, $\mathcal{O}_{S,P,S',P'}$.

The decay in the four-body final state is completely described by a set of four kinematical variables, θ_l, θ_K, ϕ, and q^2. The variable θ_l is defined as the angle between the momentum of the μ^+ (μ^-) and the direction opposite to the B^0 ($\overline{B}^0$) momentum, in the dimuon rest frame:

$$\begin{aligned} \cos\theta_l &= \left(p_{\mu^+}^{(\mu^+\mu^-)}\right)\cdot\left(-p_{B^0}^{(\mu^+\mu^-)}\right) \quad \text{for } B^0 \text{ decay} \\ \cos\theta_l &= \left(p_{\mu^-}^{(\mu^+\mu^-)}\right)\cdot\left(-p_{\overline{B}^0}^{(\mu^+\mu^-)}\right) \quad \text{for } \overline{B}^0 \text{ decay} \end{aligned} \tag{1.10}$$

The variable θ_K is defined as the angle between the direction of the kaon and the direction opposite that of the B^0 ($\overline{B}^0$), in the K^{*0} ($\overline{K}^{*0}$) rest frame

$$\begin{aligned} \cos\theta_K &= \left(p_{K^+}^{(K^{*0})}\right)\cdot\left(-p_{B^0}^{(K^{*0})}\right) \quad \text{for } B^0 \text{ decay;} \\ \cos\theta_K &= \left(p_{K^-}^{(\overline{K}^{*0})}\right)\cdot\left(-p_{\overline{B}^0}^{(\overline{K}^{*0})}\right) \quad \text{for } \overline{B}^0 \text{ decay.} \end{aligned} \tag{1.11}$$

The variable ϕ is defined as the angle between the plane containing the μ^+ and μ^- momenta and the plane containing the momenta of the kaon and

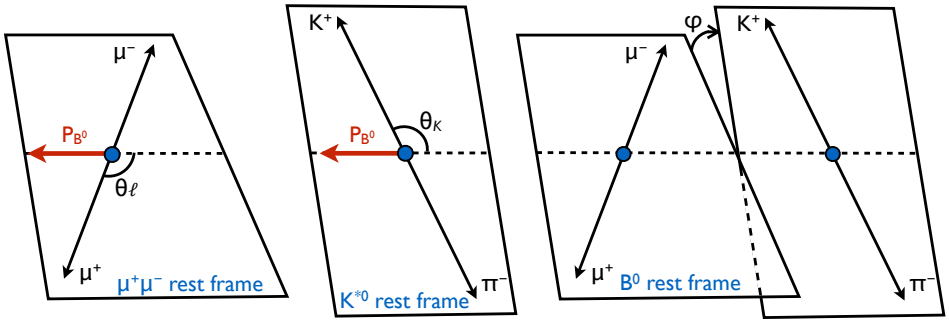


Figure 1.2: Illustration of the angular variables θ_l (left), θ_K (middle), and ϕ (right) for the decay $\mathrm{B}^0 \to \mathrm{K}^{*0}(\mathrm{K}^+\pi^-)\mu^+\mu^-$.

the pion, in the B^0 rest frame, with the following sign conventions:

$$
\begin{aligned}
\cos(\phi) &= \left(p_{\mu^+}^{(\mathrm{B}^0)} \times p_{\mu^-}^{(\mathrm{B}^0)}\right) \cdot \left(p_{\mathrm{K}^+}^{(\mathrm{B}^0)} \times p_{\pi^-}^{(\mathrm{B}^0)}\right) && \text{and} \\
\sin(\phi) &= \left[\left(p_{\mu^+}^{(\mathrm{B}^0)} \times p_{\mu^-}^{(\mathrm{B}^0)}\right) \times \left(p_{\mathrm{K}^+}^{(\mathrm{B}^0)} \times p_{\pi^-}^{(\mathrm{B}^0)}\right)\right] \cdot p_{\mathrm{K}^{*0}}^{(\mathrm{B}^0)} && \text{for } \mathrm{B}^0 \text{ decay;} \\
\cos(\phi) &= \left(p_{\mu^-}^{(\overline{\mathrm{B}}^0)} \times p_{\mu^+}^{(\overline{\mathrm{B}}^0)}\right) \cdot \left(p_{\mathrm{K}^-}^{(\overline{\mathrm{B}}^0)} \times p_{\pi^+}^{(\overline{\mathrm{B}}^0)}\right) && \text{and} \\
\sin(\phi) &= -\left[\left(p_{\mu^-}^{(\overline{\mathrm{B}}^0)} \times p_{\mu^+}^{(\overline{\mathrm{B}}^0)}\right) \times \left(p_{\mathrm{K}^-}^{(\overline{\mathrm{B}}^0)} \times p_{\pi^+}^{(\overline{\mathrm{B}}^0)}\right)\right] \cdot p_{\overline{\mathrm{K}}^{*0}}^{(\overline{\mathrm{B}}^0)} && \text{for } \overline{\mathrm{B}}^0 \text{ decay.}
\end{aligned}
\tag{1.12}
$$

Finally, the variable q^2 is defined as the invariant mass squared of the dimuon system. In the Equations 1.10, 1.11, 1.12 the notation $p_a^{(b)}$ indicates the momentum of the particle a in the rest frame of the particle b.

The definition of the kinematical variables is coherent with the one used in the previous experimental analyses performed on this decay channel. The definition of θ_l differs from the one used in some phenomenological papers. A graphic representation of the definition of the three angular variables is shown in Figure 1.2.

The range of definition of the angular variables are $[0, \pi]$, for θ_l and θ_K, and $[-\pi, \pi]$, for ϕ.

1.2.1 The angular decay rate

The matrix element for the $B^0 \to K^{*0}(K^+\pi^-)\mu^+\mu^-$, obtained from the effective Hamiltonian in Equation 1.2, is:

$$\begin{aligned}\mathcal{M} =& \frac{G_F\alpha}{\sqrt{2}\pi}V_{tb}V_{ts}^*\Bigg\{\Bigg[\langle K\pi|\bar{s}\gamma^\mu(C_9^{\text{eff}}P_L + C_9^{\prime\text{eff}}P_R)b|\bar{B}\rangle \\ &- \frac{2m_b}{q^2}\langle K\pi|\bar{s}i\sigma^{\mu\nu}q_\nu(C_7^{\text{eff}}P_R + C_7^{\prime\text{eff}}P_L)b|\bar{B}\rangle\Bigg](\bar{\mu}\gamma_\mu\mu) \\ &+ \langle K\pi|\bar{s}\gamma^\mu(C_{10}^{\text{eff}}P_L + C_{10}^{\prime\text{eff}}P_R)b|\bar{B}\rangle\,(\bar{\mu}\gamma_\mu\gamma_5\mu) \\ &+ \langle K\pi|\bar{s}(C_SP_R + C_S'P_L)b|\bar{B}\rangle\,(\bar{\mu}\mu) + \langle K\pi|\bar{s}(C_PP_R + C_P'P_L)b|\bar{B}\rangle\,(\bar{\mu}\gamma_5\mu)\Bigg\}.\end{aligned} \tag{1.13}$$

where the naive factorisation has been applied, by ignoring the non-factorisable terms, and a set of effective Wilson coefficients, C_i^{eff}, is used. These effective coefficients are linear combinations of the set defined in Section 1.1.1, and their definition is described in Reference [12].

The differential decay distribution as a function of the kinematical variables can be obtained by squaring Equation 1.13. Projecting the result on a basis of combinations of spherical harmonics of the angular variables, one obtains the following expression for the $\overline{B}^0 \to \overline{K}^{*0}(K^-\pi^+)\mu^+\mu^-$ decay:

$$\begin{aligned}\frac{d^4\Gamma}{dq^2\,d\cos\theta_l\,d\cos\theta_K\,d\phi} = \frac{9}{32\pi}\Bigg[&I_1^s\sin^2\theta_K + I_1^c\cos^2\theta_K \\ &+ (I_2^s\sin^2\theta_K + I_2^c\cos^2\theta_K)\cos 2\theta_l + I_3\sin^2\theta_K\sin^2\theta_l\cos 2\phi \\ &+ I_4\sin 2\theta_K\sin 2\theta_l\cos\phi + I_5\sin 2\theta_K\sin\theta_l\cos\phi \\ &+ (I_6^s\sin^2\theta_K + I_6^c\cos^2\theta_K)\cos\theta_l + I_7\sin 2\theta_K\sin\theta_l\sin\phi \\ &+ I_8\sin 2\theta_K\sin 2\theta_l\sin\phi + I_9\sin^2\theta_K\sin^2\theta_l\sin 2\phi\Bigg]\end{aligned} \tag{1.14}$$

where the complex angular coefficients $I_i^{(a)}$ depend only on q^2. The differential distribution of the opposite CP-state decay, $B^0 \to K^{*0}(K^+\pi^-)\mu^+\mu^-$, has the same expression, but it is function of the weak-phase conjugated coefficients $\bar{I}_i^{(a)}$.

The set of angular coefficients $I_i^{(a)}$ has been expressed as a function of the effective Wilson coefficients and of the form factors, under the assumptions of naive factorisation of the matrix element and adding corrections to describe the effect of non-factorisable terms. In literature they are also expressed as a function of eight K^{*0} transversity amplitudes.

Unlike Wilson coefficients, form factors, and transversity amplitudes, the angular coefficients $I_i^{(a)}$ are physical observable and they can be experimentally measured by the angular analysis of the $B^0 \to K^{*0}(K^+\pi^-)\mu^+\mu^-$ decay. However, the theoretical predictions of these coefficients are prone to

hadronic uncertainties, derived from their strong dependency on the form factors. For this reasons, more sophisticated bases of angular coefficients have been defined in order to be independent from form factors at leading order of the effective-theory expansion.

Firstly, two sets of coefficients are defined, the CP-averages:

$$S_i^{(a)} = \frac{I_i^{(a)} + \bar{I}_i^{(a)}}{\frac{d\Gamma}{dq^2} + \frac{d\bar{\Gamma}}{dq^2}} \tag{1.15}$$

and the CP asymmetries:

$$A_i^{(a)} = \frac{I_i^{(a)} - \bar{I}_i^{(a)}}{\frac{d\Gamma}{dq^2} + \frac{d\bar{\Gamma}}{dq^2}} \tag{1.16}$$

that are useful to disentangle potential NP effects that introduce new sources of CP-violation from the others. Since this analysis is aiming the measurement of two CP-averaged angular parameters, I will not spend time describing the basis of CP-violating coefficients but I will focus on the set of $S_i^{(a)}$ parameters.

The twelve real angular coefficients in this set can be reduced to eight, in the approximation of negligible lepton mass with respect to q. This is true in this analysis, since only candidates with q value greater than 1 GeV are used, as described in Section 3.3. In the massless-muon limit, the following conditions are valid:

$$\begin{gathered} S_1^s = 3S_2^s \\ S_1^c = -S_2^c \\ \frac{3}{4}\left(2S_1^s + S_1^c\right) - \frac{1}{4}\left(2S_2^s + S_2^c\right) = 1 \\ S_6^c = 0 \end{gathered} \tag{1.17}$$

where the latter condition is always true within the SM, even without the massless-muon limit, since the S_6^c term is generated by the scalar operator, which does not exist in the SM.

The S_1^c coefficient corresponds to the fraction of K^{*0} produced with longitudinal polarisation, and it is usually referred to as F_L. In the same way, the coefficient S_6^s is proportional to the forward-backward asymmetry of the muon system, thus the parameter $A_{FB} = \frac{3}{4}S_6^s$ is used.

The so-called P-primed basis of angular parameters, clean from leading-

order hadronic uncertainties, is defined as:

$$\begin{aligned} P_1 &= \frac{2S_3}{1-F_L} \\ P_2 &= \frac{2}{3}\frac{A_{FB}}{1-F_L} \\ P_3 &= \frac{-S_9}{1-F_L} \\ P'_{4,5,8} &= \frac{S_{4,5,8}}{\sqrt{F_L(1-F_L)}} \\ P'_6 &= \frac{S_7}{\sqrt{F_L(1-F_L)}} \end{aligned} \tag{1.18}$$

Using the $P_i^{(\prime)}$ basis, with F_L and its complementary $F_T = 1 - F_L$, the differential angular distribution can be written as:

$$\begin{aligned} \frac{1}{\mathrm{d}\Gamma/\mathrm{d}q^2}\frac{\mathrm{d}^4\Gamma}{\mathrm{d}q^2\mathrm{d}\cos\theta_l\mathrm{d}\cos\theta_\mathrm{K}\mathrm{d}\phi} = \frac{9}{32\pi}\Bigg[&\frac{3}{4}F_\mathrm{T}\sin^2\theta_\mathrm{K} + F_\mathrm{L}\cos^2\theta_\mathrm{K} \\ &+ \left(\frac{1}{4}F_\mathrm{T}\sin^2\theta_\mathrm{K} - F_\mathrm{L}\cos^2\theta_\mathrm{K}\right)\cos 2\theta_l + \frac{1}{2}P_1 F_\mathrm{T}\sin^2\theta_\mathrm{K}\sin^2\theta_l\cos 2\phi \\ &+ \sqrt{F_\mathrm{T}F_\mathrm{L}}\left(\frac{1}{2}P'_4\sin 2\theta_\mathrm{K}\sin 2\theta_l\cos\phi + P'_5\sin 2\theta_\mathrm{K}\sin\theta_l\cos\phi\right) \\ &- \sqrt{F_\mathrm{T}F_\mathrm{L}}\left(P'_6\sin 2\theta_\mathrm{K}\sin\theta_l\sin\phi - \frac{1}{2}P'_8\sin 2\theta_\mathrm{K}\sin 2\theta_l\sin\phi\right) \\ &+ 2P_2 F_\mathrm{T}\sin^2\theta_\mathrm{K}\cos\theta_l - P_3 F_\mathrm{T}\sin^2\theta_\mathrm{K}\sin^2\theta_l\sin 2\phi\Bigg] \end{aligned} \tag{1.19}$$

S-wave contamination

Although the $K^+\pi^-$ invariant mass must be consistent with a K^{*0}, there can be contributions from a spinless (S-wave) $K^+\pi^-$ combination. The presence of a $K^+\pi^-$ system in an S-wave configuration, due to a non-resonant contribution or to feed through from $K^+\pi^-$ scalar resonances, results in additional terms in the different decay rate.

Denoting the right-hand side of Equation 1.19 by W_p, the differential decay rate takes the form:

$$\frac{1}{\mathrm{d}\Gamma/\mathrm{d}q^2}\frac{\mathrm{d}^4\Gamma}{\mathrm{d}q^2\mathrm{d}\cos\theta_l\mathrm{d}\cos\theta_\mathrm{K}\mathrm{d}\phi}\bigg|_{S+P} = (1-F_\mathrm{S})W_p + (W_s + W_{sp}) \tag{1.20}$$

where

$$W_s = \frac{3}{16\pi}F_\mathrm{S}\sin^2\theta_l \tag{1.21}$$

and W_{sp} is:

$$W_{sp} = \frac{3}{16\pi}\left[A_{\rm S}\sin^2\theta_l\cos\theta_{\rm K} + A_{\rm S}^4\sin\theta_{\rm K}\sin 2\theta_l\cos\phi + A_{\rm S}^5\sin\theta_{\rm K}\sin\theta_l\cos\phi + A_{\rm S}^7\sin\theta_{\rm K}\sin\theta_l\sin\phi + A_{\rm S}^8\sin\theta_{\rm K}\sin 2\theta_l\sin\phi\right] \tag{1.22}$$

where $F_{\rm S}$ is the fraction of the S-wave component in the $\rm K^{*0}$ mass window, and $A_{\rm S}^i$ are the interference amplitudes between the S-wave and the P-wave decays [40].

1.2.2 The angular folding

The differential decay distributions, described above, contains 8 parameters for the P-wave component, and 6 parameters for the S-wave and interference contamination. As mentioned in the introduction, the primary goal of this angular analysis is to measure the angular parameter P_5'. In order to perform this measurement in the best possible environment, it has been decided to simplify the angular analysis by reducing the number of measured parameters. This is done with the *angular folding* method.

In the following, I will use the expression "folding a variable x around a" referring to the redefinition of that variable in the following way:

$$\begin{aligned} x &\to x && \text{for } x \geq a \\ x &\to 2a - x && \text{for } x < a \end{aligned} \tag{1.23}$$

The key of this method is to exploit the even and odd symmetries of the terms forming the differential distribution, around some values of the angular variables. By folding a variable around one of these values, all the terms with an even symmetry will be unchanged after the transformation, while all the terms with an odd symmetry will be become equal to zero. Any folding applied around a value for which there is no symmetry, in at least one term of the distribution, would modify the expression of that term in a non-trivial way. In this analysis, only foldings corresponding to symmetries of the angular decay rate are applied.

In order to preserve the term proportional to P_5' in the angular distribution and to reduce as much as possible the total number of parameters, two foldings have been applied. The ϕ variable has been folded around 0, reducing its range from $[-\pi, \pi]$ to $[0, \pi]$. Then, the θ_l variable has been folded around $\pi/2$, reducing its range from $[0, \pi]$ to $[0, \pi/2]$.

Note that, according to the definition in Equation 1.23, the resulting range of θ_l should have been $[\pi/2, \pi]$. However, for convenience this second folding has been applied in the other way:

$$\begin{aligned} x &\to 2a - x && \text{for } x \geq a \\ x &\to x && \text{for } x < a \end{aligned} \tag{1.24}$$

The effect of this alternative definition on the differential distribution is identical to the original one.

After the application of the two foldings, the differential decay distribution can be written as:

$$\begin{aligned}\frac{1}{\mathrm{d}\Gamma/\mathrm{d}q^2}\frac{\mathrm{d}^4\Gamma}{\mathrm{d}q^2\mathrm{d}\cos\theta_l\mathrm{d}\cos\theta_\mathrm{K}\mathrm{d}\phi} = \\ \frac{9}{8\pi}\Bigg\{\frac{2}{3}\left[(F_\mathrm{S}+A_\mathrm{S}\cos\theta_\mathrm{K})\left(1-\cos^2\theta_l\right)+A_\mathrm{S}^5\sqrt{1-\cos^2\theta_\mathrm{K}}\sqrt{1-\cos^2\theta_l}\cos\phi\right] \\ +(1-F_\mathrm{S})\Bigg[2F_\mathrm{L}\cos^2\theta_\mathrm{K}\left(1-\cos^2\theta_l\right)+\frac{1}{2}\left(1-F_\mathrm{L}\right)\left(1-\cos^2\theta_\mathrm{K}\right)\left(1+\cos^2\theta_l\right) \\ +\frac{1}{2}P_1(1-F_\mathrm{L})(1-\cos^2\theta_\mathrm{K})(1-\cos^2\theta_l)\cos 2\phi \\ +2P_5'\cos\theta_\mathrm{K}\sqrt{F_\mathrm{L}\left(1-F_\mathrm{L}\right)}\sqrt{1-\cos^2\theta_\mathrm{K}}\sqrt{1-\cos^2\theta_l}\cos\phi\Bigg]\Bigg\}\end{aligned} \tag{1.25}$$

Now it contains 6 angular parameters: F_L, P_1, and P_5' for the P-wave component, F_S for the S-wave component, A_S and A_S^5 for the interference component.

1.2.3 Parameter constrains

Range of definition of interference terms

Due to their nature, the values of the interference terms A_s and A_s^5 are limited by the amplitude of the pure P-wave and S-wave components [40]. Their allowed ranges are the following:

$$|A_s| < 2\sqrt{3}\sqrt{F_S(1-F_S)F_L} * F_{theo} \tag{1.26}$$

$$|A_s^5| < \sqrt{3}\sqrt{F_S(1-F_S)F_T(1+P_1)} * F_{theo} \tag{1.27}$$

where F_{theo} is a constant factor that depends on the selection cuts applied to the Kπ system mass. For the selection criteria used in this analysis, as described in Section 3.2, according to the theoretical prescriptions [40] the value used for F_{theo} is 0.89.

In order to make sure that the fitted value of the A_s^5 parameter is contained in this range, it has been substituted in the `pdf` by:

$$A_s^5 \to f\sqrt{3}\sqrt{F_S(1-F_S)F_T(1+P_1)} * F_{theo} \tag{1.28}$$

where f is a placeholder parameter defined in the range [-1;1]. The fit will be performed with respect to f, constrained in his range of validity, and the resulting values of A_s^5 will be obtained from f by reversing Equation 1.28. In this thesis this additional step will be kept implicit: when it will be stated that the likelihood is maximised as a function of A_s^5, it will be intended that it is maximised as a function of f.

`pdf` validity in the parameter space

Using the `pdf` parameterisation described above it is not guaranteed that it is physical, i.e. positive in the whole ($\cos\theta_K$,$\cos\theta_l$,ϕ) space.

In order to have a working fit sequence and reliable results, the values of the angular parameters that allow the `pdf` to be physical need to be identified. As explained in Section 4.3, the angular parameters that are free to float in the fit to the data are P_1, P_5', and $A_{\rm S}^5$, while the others, $F_{\rm L}$, F_S, and $A_{\rm S}$, are kept fixed to the results of the previous CMS analysis. Thus, the region of validity of the `pdf` is computed only in the three-dimensional space of the floating parameters, assuming the values of the others fixed like in the fit procedure.

The analytic boundaries for the pure P-wave component can be found in literature [38] and, after the reduction to the P_1, P_5' parameter space, they are:

$$(P_5')^2 - 1 < P_1 < 1 \tag{1.29}$$

without any dependence on the value of $F_{\rm L}$.

The pure S-wave component is always positive, as can be derived from Equation 1.21 and from the range of definition of F_S, from 0 to 1. In general, the interference terms can be also negative, so additional constraints are needed to guarantee the positiveness of the full `pdf`. Since these constraints are not available in literature in an analytic form, a numerical computation is needed to describe them.

For each bin, eight different physical boundaries in the (P_1,P_5') parameter space have been computed:

- one boundary is obtained by requiring that only the P-wave component is positive; this is just a cross-check of the validity of Equation 1.29, since the same result is expected;
- seven boundaries are obtained by requiring the whole `pdf` to be positive, for seven different values of A_s^5; according the convention defined in Section 1.2.3, the seven values of the placeholder parameter f are chosen to be [-1, -2/3, -1/3, 0, 1/3, 2/3, 1].

To compute numerically each of these physical boundaries, the P_1,P_5' space has been scanned with a grid of step 0.01 in both directions. For each point of this *parameter* grid, the values of $\cos\theta_K$, $\cos\theta_l$ and ϕ are moved on a three-dimensional grid with step 0.02 in each direction; if the `pdf` is positive for all of the points of this *angle* grid, the point in the P_1,P_5' space is defined to be inside the physical region, otherwise it is outside.

The boundaries of the resulting physical regions, plotted in the negative P_5' sector only, are shown from Figure 1.3 to Figure 1.9. The physical region is the one with P_5' values smaller, in module, than the plotted boundaries, and it is included between them and their projections in the positive P_5'

sector. The P-wave boundary is symmetrically reflected with respect to $P_5' = 0$; the boundaries assigned to a value of A_s^5 can be reflected to the positive P_5' sector, but the reflected boundary refers to the opposite value of A_s^5, as can be derived from the symmetries in Equation 1.25.

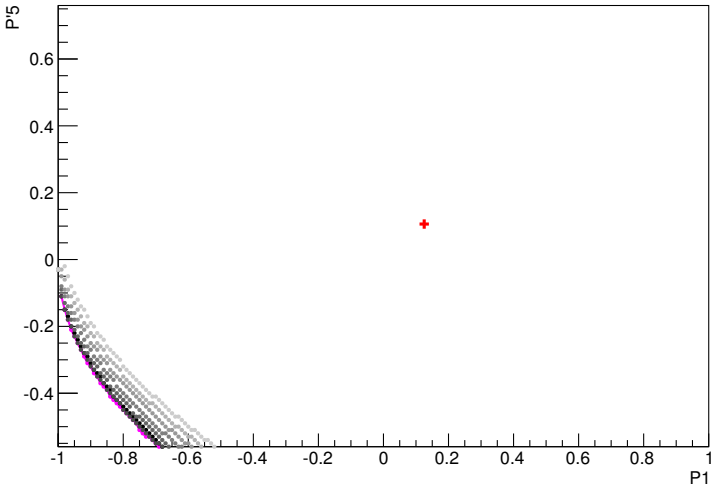


Figure 1.3: Physical boundaries of the negative P_5' sector of q^2 bin 0. Accordingly to the description in Section 1.2.3, the magenta line is the boundary of the P-wave physical region and the set of grey-scale lines are the boundaries of the total-PDF physical region, for different A_s^5 values (black for $f = -1$, lightest grey for $f = 1$).

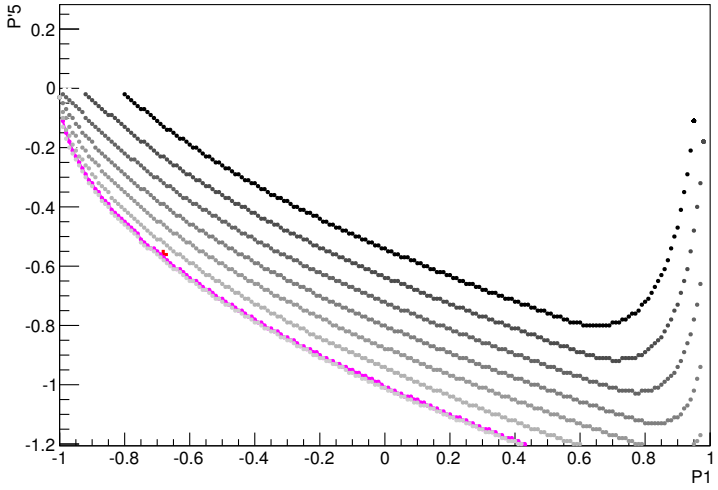


Figure 1.4: Physical boundaries of the negative P_5' sector of q^2 bin 1. Accordingly to the description in Sec. 1.2.3, the magenta line is the boundary of the P-wave physical region and the set of grey-scale lines are the boundaries of the total-PDF physical region, for different A_s^5 values (black for $f = -1$, lightest grey for $f = 1$).

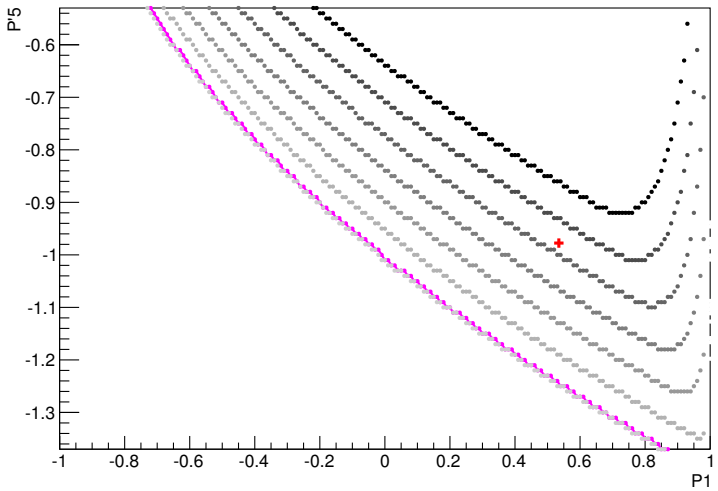


Figure 1.5: Physical boundaries of the negative P_5' sector of q^2 bin 2. Accordingly to the description in Sec. 1.2.3, the magenta line is the boundary of the P-wave physical region and the set of grey-scale lines are the boundaries of the total-PDF physical region, for different A_s^5 values (black for $f = -1$, lightest grey for $f = 1$).

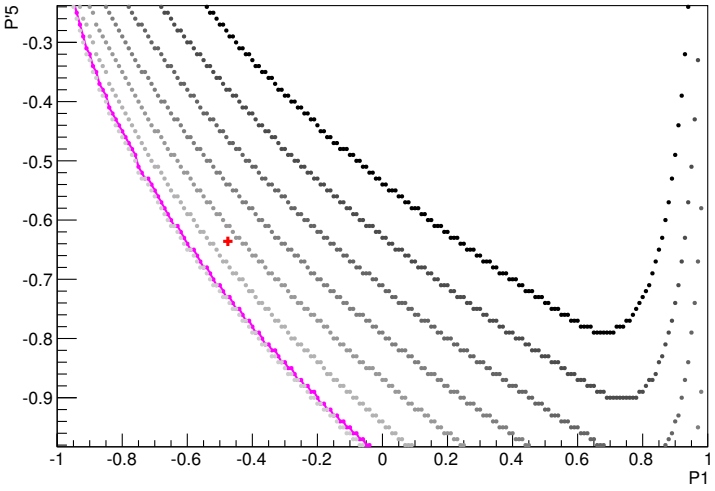


Figure 1.6: Physical boundaries of the negative P_5' sector of q^2 bin 3. Accordingly to the description in Sec. 1.2.3, the magenta line is the boundary of the P-wave physical region and the set of grey-scale lines are the boundaries of the total-PDF physical region, for different A_s^5 values (black for $f = -1$, lightest grey for $f = 1$).

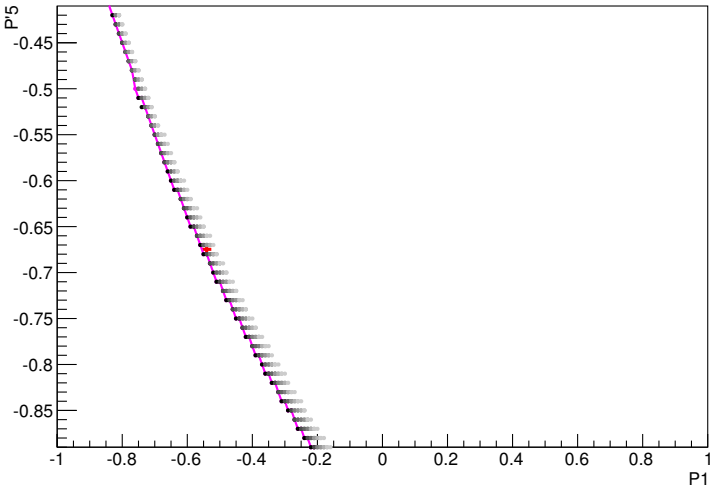


Figure 1.7: Physical boundaries of the negative P_5' sector of q^2 bin 5. Accordingly to the description in Sec. 1.2.3, the magenta line is the boundary of the P-wave physical region and the set of grey-scale lines are the boundaries of the total-PDF physical region, for different A_s^5 values (black for $f = -1$, lightest grey for $f = 1$).

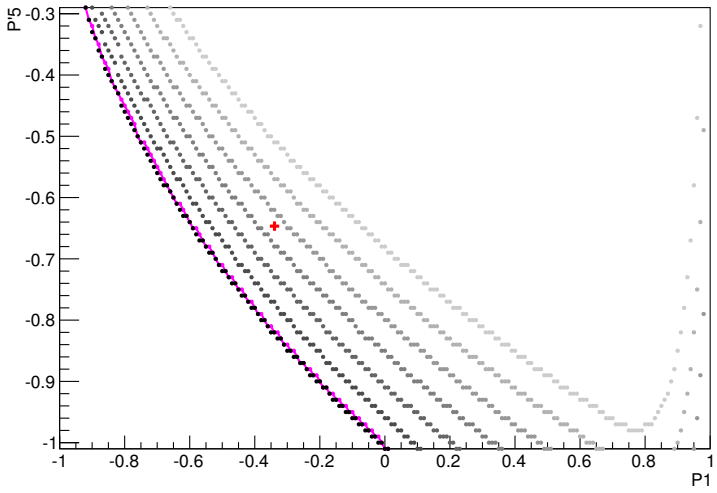


Figure 1.8: Physical boundaries of the negative P_5' sector of q^2 bin 7. Accordingly to the description in Sec. 1.2.3, the magenta line is the boundary of the P-wave physical region and the set of grey-scale lines are the boundaries of the total-PDF physical region, for different A_s^5 values (black for $f = -1$, lightest grey for $f = 1$).

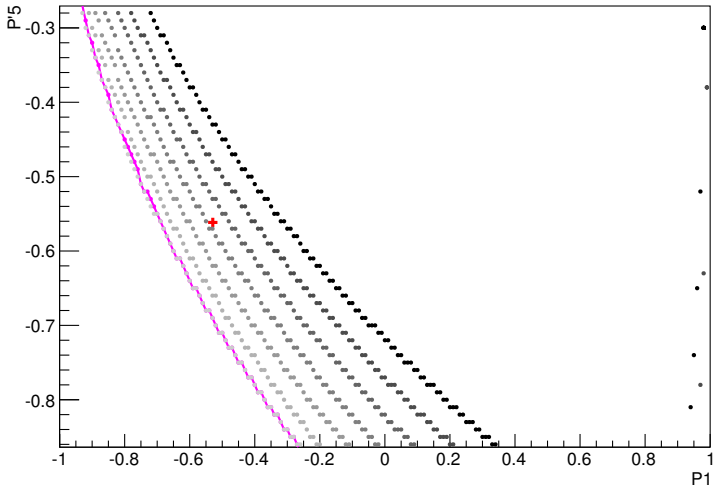


Figure 1.9: Physical boundaries of the negative P_5' sector of q^2 bin 8. Accordingly to the description in Sec. 1.2.3, the magenta line is the boundary of the P-wave physical region and the set of grey-scale lines are the boundaries of the total-PDF physical region, for different A_s^5 values (black for $f = -1$, lightest grey for $f = 1$).

Chapter 2

The Large Hadron Collider and the Compact Muon Solenoid Experiment

The Large Hadron Collider (LHC) [22] is an accelerator located at the European Laboratory for Particle Physics Research (CERN) in Geneva. It has been conceived to collide proton beams at a centre-of-mass energy of $\sqrt{s} = 14$ TeV and a nominal instantaneous luminosity of $\mathcal{L} = 10^{34}$ cm^{-2} s^{-1}, representing a seven-fold increase in energy and a hundred-fold increase in integrated luminosity over the previous hadron collider experiments. Its main purpose is to search for rare processes like the production of Higgs or new particles with mass of 1 TeV and beyond. Two experiments have been installed around the LHC to pursue these results: ATLAS [1] and CMS [23]. Furthermore, the LHCb [13] experiment studies the properties of charm and beauty hadrons produced with large cross sections in the forward region in collisions at the LHC, and the ALICE [9] experiment analyses the data from relativistic heavy ion collisions to study the hadronic matter in extreme temperature and density conditions (i.e. high quark-gluon density).

2.1 The Large Hadron Collider

The LHC has been installed in the same tunnel which hosted the e^+e^- collider LEP (Large Electron-Positron collider). Accelerated electrons and positrons suffer large energy loss due to the synchrotron radiation, which is proportional to $E^4/(Rm^4)$, where E is the electron energy, m its mass and R the accelerator radius. To obtain energies of the order of TeV, at the fixed accelerator radius, only massive charged particles could have been used: protons and heavy nuclei. The energy loss is reduced by a factor $(2000)^4$ for a given fixed energy E for protons, respect to electrons. Another important aspect of the LHC is the collision rate. To produce a sufficient number of

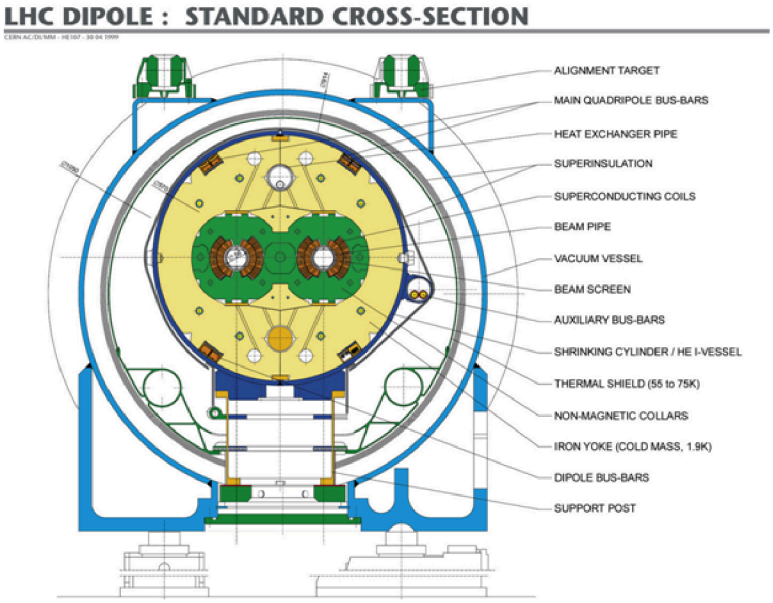


Figure 2.1: LHC dipole magnet section scheme.

rare processes, the collision rate needs to be very high. Beam protons are collected in packets called bunches. The collision rate is proportional to the instantaneous luminosity of the accelerator, defined as:

$$\mathcal{L} = \frac{f k n_p^2}{4\pi\sigma_x\sigma_y} \quad ,$$

where f is the bunch revolution frequency, k the number of bunches, n_p the number of protons per bunch, σ_x and σ_y their transverse dispersion along the x and y axis. At the nominal 14 TeV LHC conditions ($\mathcal{L} = 10^{34}$ cm^{-2} s^{-1}) the parameter values are: $k = 2808$, $n_p = 1.5{\cdot}10^{11}$ and $\sigma_x\sigma_y = 16.6\,\mu\text{m}^2$ (with $\sigma_z = 7.6$ cm along the beam). The integrated luminosity is defined as $L = \int \mathcal{L}dt$. For comparison we can consider the Tevatron accelerator at Fermilab, which produced proton-antiproton collisions since 1992. Its centre of mass energy was 1.8 TeV until 1998 and 1.96 TeV since 2001. To increase $\mathcal{L}$ by two orders of magnitude, protons are injected in both LHC beams. The antiprotons, in fact, are obtained by steering proton beams onto a nickel target and represent only a small fraction of the wide range of secondary particles produced in this interactions, thus have a production rate lower than the proton one.

The LHC is composed by 1232 super-conducting dipole magnets each 15 m long, providing a 8.3 T magnetic field to let the beams circulate inside their trajectories along the 27 km circumference. Two vacuum pipes are used to let beams circulate in opposite directions. A scheme representing

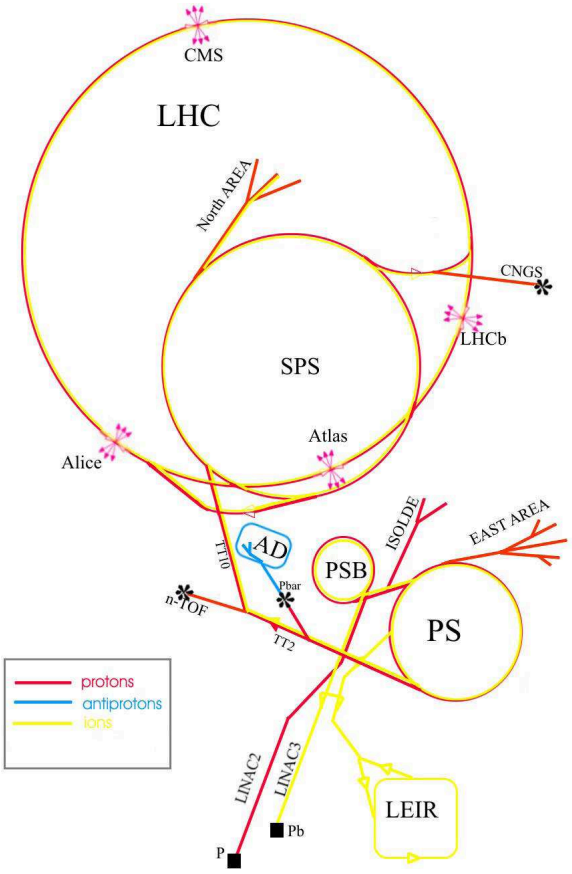


Figure 2.2: Scheme representing the CERN accelerator complex.

the transverse dipole magnet section is represented in figure 2.1. More than 8000 other magnets are utilized for the beam injection, their collimation, trajectory correction, crossing. All the magnets are kept cool by superfluid helium at 1.9 K temperature. The beams are accelerated from 450 GeV (the injection energy from the SPS) to 7 TeV with 16 Radio Frequency cavities (8 per beam) which raise the beam energy by 16 MeV each round with an electric field of 5 MV/m oscillating at 400 MHz frequency.
Before the injection into the LHC, the beams are produced and accelerated by different components of the CERN accelerator complex. Being produced from ionized hydrogen atoms, protons are accelerated by the linear accelerator LINAC, Booster and the Proton Synchrotron (PS) up to 26 GeV energy, the bunches being separated by 25 ns each. The beams are then injected into the Super Proton Synchrotron (SPS) where they are accelerated up to 450 GeV. They are then finally transferred to the LHC and accelerated up to 7 TeV energy per beam. The CERN accelerator complex is illustrated in figure 2.2.

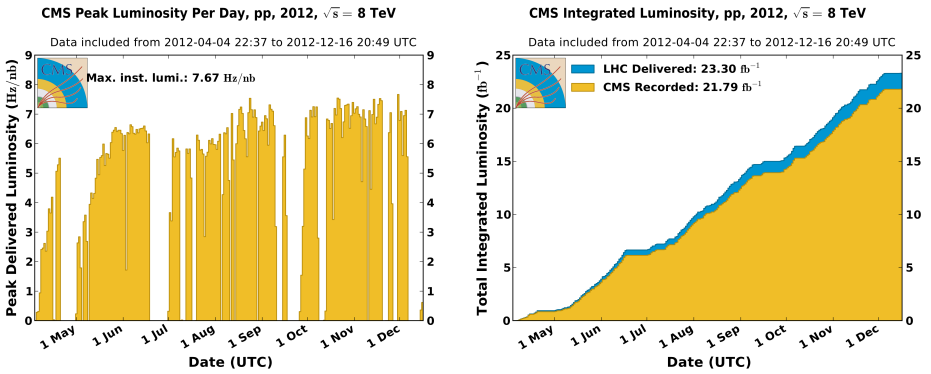


Figure 2.3: LHC performance in 2012. Left: CMS detected peak luminosity; right: CMS detected integrated luminosity.

The LHC started its operations in December 2009 with centre of mass energy for the proton-proton collision $\sqrt{s} = 0.9$ TeV. The centre of mass energy was set to $\sqrt{s} = 7$ TeV in the 2010 and 2011 runs and raised to $\sqrt{s} = 8$ TeV in the 2012 runs. Here are reported the CMS detected peak and integrated luminosities for proton-proton runs. In 2010 the peak luminosity reached $\mathcal{L} = 203.80\,\mathrm{Hz}/\mu\mathrm{b}$ and the integrated luminosity has been $L = 40.76\,\mathrm{pb}^{-1}$, while during 2011 the peak luminosity increased to $\mathcal{L} = 4.02\,\mathrm{Hz/nb}$ and the integrated luminosity has been $L = 5.55\,\mathrm{fb}^{-1}$. In the 2012 runs the peak luminosity reached $\mathcal{L} = 7.67\,\mathrm{Hz/nb}$ and the integrated luminosity has been $L = 21.79\,\mathrm{fb}^{-1}$, as graphically summarized in figure 2.3.

2.2 CMS Experiment

The Compact Muon Solenoid [23] is a general purpose detector situated at interaction point 5 of the CERN Large Hadron Collider. It is designed around a 4 T solenoidal magnetic field provided by the largest superconducting solenoid ever built. The structure of CMS is shown in figure 2.4, where particular emphasis is put on the volumes of the different subsystems: the Silicon Pixel Detector, the Silicon Strip Tracker, the Electromagnetic and Hadronic Calorimeters, and Muon Detectors.

We can briefly summarize the aims of the CMS detector [16]. They are mainly:

- search for SM and MSSM Higgs boson decaying into photons, b quarks, τ leptons, W and Z bosons,

- search for additional heavy neutral gauge bosons predicted in many superstring-inspired theories or Great Unification Theories and decaying to muon pairs,

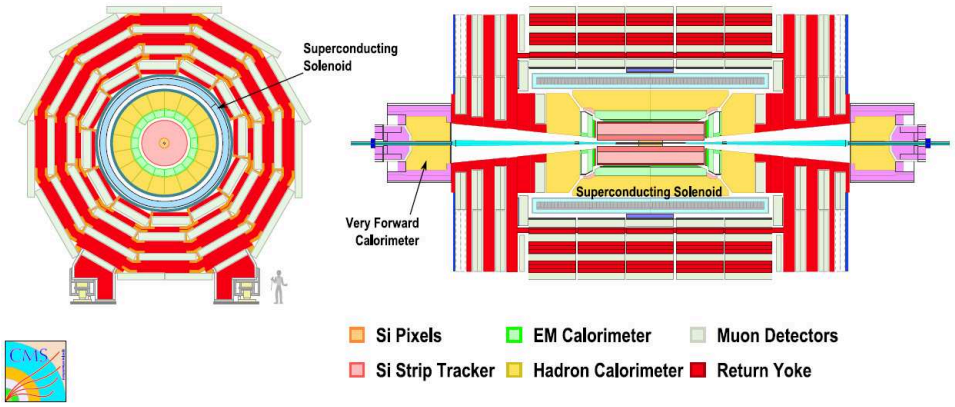


Figure 2.4: Transverse (left) and longitudinal (right) cross sections of the CMS detector showing the volumes of the different detector subsystems. The transverse cross section is drawn for the central barrel, coaxial with the beam line, while complementary end-caps are shown in the longitudinal view.

- search for new Physics in various topologies: multilepton events, multijet events, events with missing transverse energy[1] or momentum, any combination of the three above,
- study of the B-hadron rare decay channels (like $B^0_{(s)} \to \mu\mu$) and of CP violation in the decay of the B mesons (like $B^0_s \to J/\psi\phi \to \mu^+\mu^- K^+ K^-$),
- search for $B^0 \to \mu^+\mu^-$ decays,
- study of QCD and jet physics at the TeV scale,
- study of top quark and EW physics.

CMS has been therefore designed as a multipurpose experiment, with particular focus on muon, photon, and displaced tracks reconstruction. Superb performances have been achieved overall, in particular in:

- primary and secondary vertex localization,
- charged particle momentum resolution and reconstruction efficiency in the tracking volume,
- electromagnetic energy resolution,

[1]Missing transverse energy mE_T is the amount of energy which must be added to balance the modulus of the vector sum of the projections of the track momenta and calorimeter clusters in the plane perpendicular to beam axis. Its direction is opposite to this vector sum directions.

- isolation of leptons and photons at high luminosities,
- measurement of the direction of photons, rejection of $\pi^0 \to \gamma\gamma$,
- diphoton and dielectron mass resolution $\sim 1\%$ at 100GeV,
- measurement of the missing transverse energy mE_T and dijet mass with high resolution,
- muon identification over a wide range of momenta,
- dimuon mass resolution $\sim 1\%$ at 100 GeV,
- unambiguously determining the charge of muons with p_T up to 1 TeV,
- triggering and offline tagging of τ leptons and b jets.

The reference frame used to describe the CMS detector and the collected events has its origin in the geometrical centre of the solenoid. Different types of global coordinates measured with respect to the origin[2] are used:

- cartesian coordinate system, $\hat{x}$ axis points towards the centre of LHC, $\hat{y}$ points upwards, perpendicular to LHC plane, while $\hat{z}$ completes the right-handed reference,
- polar coordinate system, directions are defined with an azimuthal angle $\tan\phi = y/x$ and a polar angle $\tan\theta = \rho/z$, where $\rho = \sqrt{x^2 + y^2}$,
- polar coordinate system, with instead of the polar angle the rapidity y and the pseudorapidity η, obtained for any particle from

$$y = \frac{1}{2}\ln\left(\frac{E + p_z}{E - p_z}\right) \quad ,$$

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right) \quad ,$$

  where E is the particle energy and p_z the component of its momentum along the beam direction.

[2]Global coordinates are measured in the CMS reference frame while local coordinates are measured in the reference frame of a specific sub-detector or sensitive element.

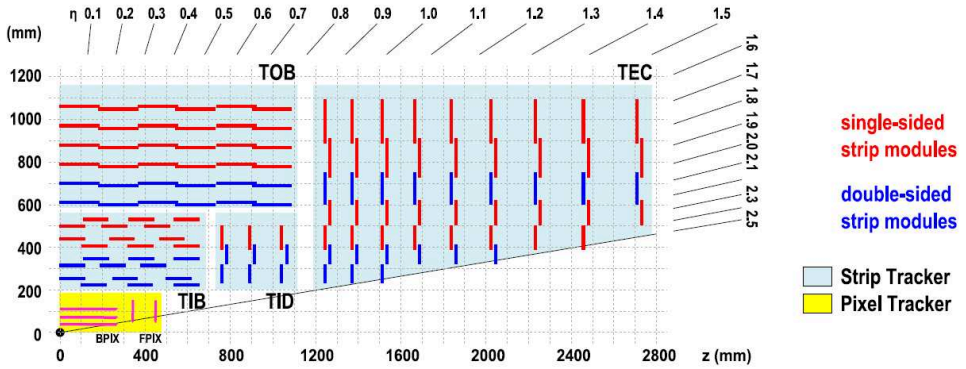


Figure 2.5: Layout of the CMS silicon tracker showing the relative position of hybrid pixels, single-sided strips and double-sided strips. Figure from [23].

2.2.1 Magnet

The whole CMS detector is designed around a $\sim$ 4 T superconducting solenoid [30] 12.5 m long and with inner radius of 3 m. The solenoid thickness is 3.9 radiation lengths and it can store up to 2.6 GJ of energy.

The field is closed by a 10^4 t iron return yoke made of five barrels and two end-caps, composed of three layers each. The yoke is instrumented with four layers of muon stations. The coil is cooled down to 4.8 K by a helium refrigeration plant, while insulation is given by two pumping stations providing vacuum on the 40 m^3 of the cryostat volume.

The magnet was designed in order to reach precise measurement of muon momenta. A high magnetic field is required to keep a compact spectrometer capable to measure 100 GeV track momentum with percent precision. A solenoidal field was chosen because it keeps the bending in the transverse plane, where an accuracy better than 20 μm is achieved in vertex position measurements. The size of the solenoid allows efficient track reconstruction up to a pseudorapidity of 2.4. The inner radius is large enough to accommodate both the Silicon Tracking System and the calorimeters. During the 2012 acquisitions the magnet was operated at 3.8 T.

2.2.2 Tracking System

The core of CMS is a Silicon Tracking System [35] with 2.5 m diameter and 5.8 m length, designed to provide a precise and efficient measurement of the trajectories of charged particles emerging from LHC collisions and reconstruction of secondary vertices.

The CMS Tracking System is composed of both silicon Pixel and Strip Detectors, as shown in figure 2.5. The Pixel Detector consists of 1440 pixel modules arranged in three barrel layers and two disks in each end-cap as in

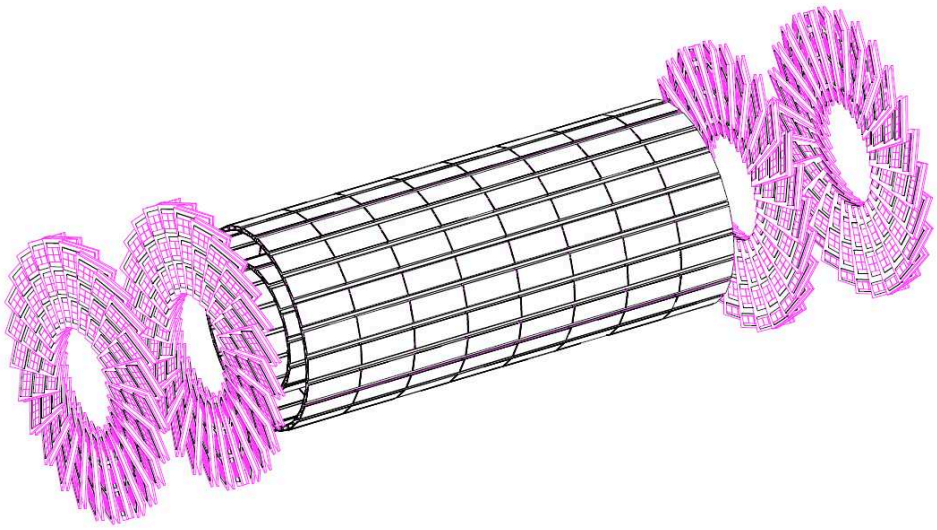


Figure 2.6: Layout of the current CMS Pixel Detector. Figure from [35].

figure 2.6. The Strip Detector consists of an inner tracker with four barrel layers and three end-cap disks and an outer tracker with six barrel layers and nine end-cap disks, housing a total amount of 15148 strip modules of both single-sided and double-sided types. Its active silicon surface of about 200 m^2 makes the CMS tracker the largest silicon tracker ever built.

The LHC physics programme requires high reliability, efficiency and precision in reconstructing the trajectories of charged particles with transverse momentum larger than 1 GeV in the pseudorapidity range $|\eta| < 2.5$. Heavy quark flavours can be produced in many of the interesting channels and a precise measurement of secondary vertices is therefore needed. The tracker completes the functionalities of ECAL and Muon System to identify electrons and muons. Also hadronic decays of tau leptons need robust tracking to be identified in both the one-prong and three-prongs topologies. Tracker information is heavily used in the High Level Trigger of CMS to help reducing the event collection rate from the 40 MHz of bunch crossing to the 100 Hz of mass storage.

Silicon Pixel Detector

The large number of particles produced in 25 pile-up events[3], at nominal LHC luminosity, results into a hit rate density of 1 MHz mm^{-2} at 4 cm from the beamline, decreasing down to 3 kHz mm^{-2} at a radius of 115 cm. Pixel detectors are used at radii below 10 cm to keep the occupancy below 1%. The chosen size for pixels, $0.100 \times 0.150\,mm^2$ in the transverse and

[3]Events that occur in the same bunch crossing.

longitudinal directions respectively, leads to an occupancy of the order of 10^{-4}. The layout of the Pixel Detector consists of a barrel region (BPIX), with three barrels at radii of 4.4, 7.3 and 10.2 cm, complemented by two disks on each side (FPIX), at 34.5 and 46.5 cm from the nominal interaction point. This layout provides about 66 million pixels covering a total area of about 1 m^2 and measuring three high precision points on each charged particle trajectory up to $|\eta| = 2.5$. Detectors in FPIX disks are tilted by 20° in a turbine-like geometry to induce charge sharing and achieve a spatial resolution of about 20 μm.

Silicon Strip Tracker

In the inner Strip Tracker, which is housed between radii of 20 and 55 cm, the reduced particle flux allows a typical cell size of $0.080 \times 100\,mm^2$, resulting in a 2% occupancy per strip at design luminosity. In the outer region, the strip pitch is increased to $0.180 \times 250\,mm^2$ together with the sensor thickness which scales from 0.320 mm to 0.500 mm. This choice compensates the larger capacitance of the strip and the corresponding larger noise with the possibility to achieve a larger depletion of the sensitive volume and a higher charge signal.

The Tracker Inner Barrel and Disks (TIB and TID) deliver up to 4 (r, ϕ) measurements on a trajectory using 0.320 mm thick silicon strip sensors with strips parallel to the beamline. The strip pitch is 0.080 mm in the first two layers and 0.120 mm in the other two layers, while in the TID the mean pitch varies from 0.100 mm to 0.141 mm. Single point resolution in the TIB is 0.023 mm with the finer pitch and 0.035 mm with the coarser one. The Tracker Outer Barrel (TOB) surrounds the TIB/TID and provides up to 6 $r - \phi$ measurements on a trajectory using 0.500 mm thick sensors. The strip pitch varies from 0.183 mm in the four innermost layers to 0.122 mm in the outermost two layers, corresponding to a resolution of 0.053 mm and 0.035 mm respectively. Tracker End-Caps (TEC) enclose the previous sub-detectors at 124cm $< |z| <$ 282cm with 9 disks carrying 7 rings of microstrips, 4 of them are 0.320 mm thick while the remaining 3 are 0.500 mm thick. TEC strips are radially oriented and their pitch varies from 0.097 mm to 0.184 mm.

As shown in figure 2.5, the first two layers and rings of TIB, TID and TOB, as well as three out of the TEC rings, carry strips on both sides with a stereo angle of 100 milliradians to measure the other coordinate: z in barrels and r in rings. This layout ensures 9 hits in the silicon Strip Tracker in the full acceptance range $|\eta| < 2.4$, and at least four of them are two-dimensional. The total area of Strip Tracker is about 198 m^2 read out by 9.3 million channels.

Trajectory Reconstruction

Due to the magnetic field charged particles travel through the tracking detectors on a helical trajectory which is described by 5 parameters: the curvature κ, the track azimuthal angle ϕ, the pseudorapidity η, the signed transverse impact parameter d_0 and the longitudinal impact parameter z_0. The transverse (longitudinal) impact parameter of a track is defined as the transverse (longitudinal) distance of closest approach of the track to the primary vertex. The main standard algorithm used in CMS for track reconstruction is the Combinatorial Track Finder (CFT) algorithm [11] which uses the reconstructed positions of the passage of charged particles in the silicon detectors to determine the track parameters. The CFT algorithm proceeds in three stages: track seeding, track finding and track fitting. Track candidates are best seeded from hits in the pixel detector because of the low occupancy, the high efficiency and the unambiguous two-dimensional position information. The track finding stage is based on a standard Kalman filter pattern recognition approach which starts with the seed parameters. The trajectory is extrapolated to the next tracker layer and compatible hits are assigned to the track on the basis of the χ^2 between the predicted and measured positions. At each stage the Kalman filter updates the track parameters with the new hits.

The tracks are assigned a quality based on the χ^2 and the number of missing hits and only the best quality tracks are kept for further propagation. Ambiguities between tracks are resolved during and after track finding. In case two tracks share more than 50% of their hits, the lower quality track is discarded. For each trajectory the finding stage results in an estimate of the track parameters. However, since the full information is only available at the last hit and constraints applied during trajectory building can bias the estimate of the track parameters, all valid tracks are refitted with a standard Kalman filter and a second filter (smoother) running from the exterior towards the beam line. The expected performance of the track reconstruction is shown in figure 2.7 for muons, pions and hadrons. The track reconstruction efficiency for high energy muons is about 99% and drops at $|\eta| > 2.1$ due to the reduced coverage of the forward pixel detector. For pions and hadrons the efficiency is in general lower because of interactions with the material in the tracker.

The material budget is shown in figure 2.8 as a function of pseudorapidity, with the different contributions of sub-detectors and services.

The performance of the Silicon Tracker in terms of track reconstruction efficiency and resolution, of vertex and momentum measurement, are shown in figure 2.7 and figure 2.9 respectively. The first one, in particular, shows the difference in reconstruction efficiency for muons and pions, due to the larger interaction cross section of pions, which cannot be assumed to be minimum-ionizing particles and therefore are much more degraded by the

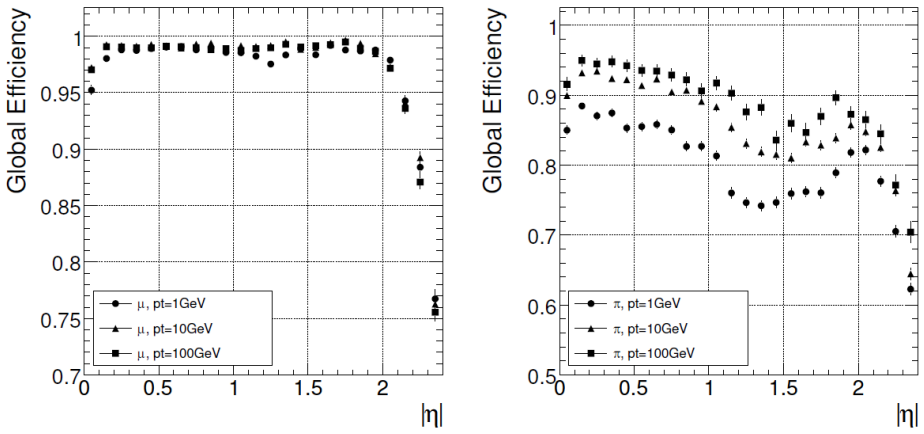


Figure 2.7: Global track reconstruction effciency as a function of track pseudorapidity for muons (left) and pions (right) of transverse momenta of 1, 10 and 100 GeV. Figures from [23].

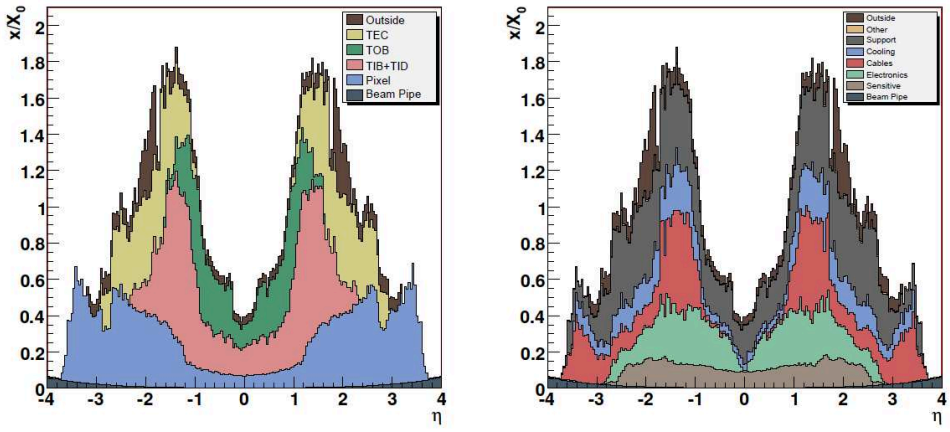


Figure 2.8: Material budget of the current CMS Tracker in units of radiation length X_0 as a function of the pseudorapidity, showing the different contribution of sub-detectors (left) and functionalities (right). Figures from [23].

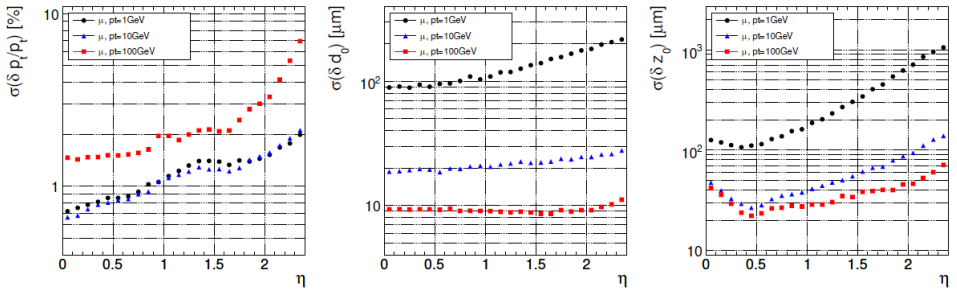


Figure 2.9: Resolution of several track parameters as a function of track pseudorapidity for single muons with transverse momenta of 1, 10 and 100 GeV: transverse momentum (left), transverse impact parameter (middle) and longitudinal impact parameter (right). Figures from [23].

amount of material.

Vertex Reconstruction

The reconstruction of interaction vertices allows CMS to reject tracks coming from pile-up events. The primary vertex reconstruction is a two-step process. Firstly the reconstructed tracks are grouped in vertex candidates and their z coordinates at the beam closest approach point are evaluated, retaining only tracks with impact parameter respect to the vertex candidate less than 3 cm. Vertices are then reconstructed through a recursive method for parameter estimation through a Kalman filter [44] algorithm. For a given event, the primary vertices are ordered according to the total transverse momentum of the associated tracks, $\sum p_T$. The vertex reconstruction efficiency is very close to 100% and the position resolution is of the order of $\mathcal{O}(10)\,\mu$m in all directions.

It is also possible to reconstruct the secondary vertices, for example those from b-quark decays. The secondary vertex reconstruction uses tracks associated to jets applying further selection cuts: the transverse impact parameter of the tracks must be greater than 100 μm, to avoid tracks coming from the primary vertex, and the longitudinal impact parameter below 2 cm, to avoid tracks from pile-up events.

2.2.3 Muon Spectrometer

Detection of muons at CMS exploits different technologies and is performed by a "Muon System" rather than a single detector [31]. Muons are the only particles able to reach the external muon chambers with a minimal energy loss when traversing the calorimeters, the solenoid and the magnetic field return yoke. Muons can provide strong indication of interesting signal events and are natural candidates for triggering purposes. The CMS Muon

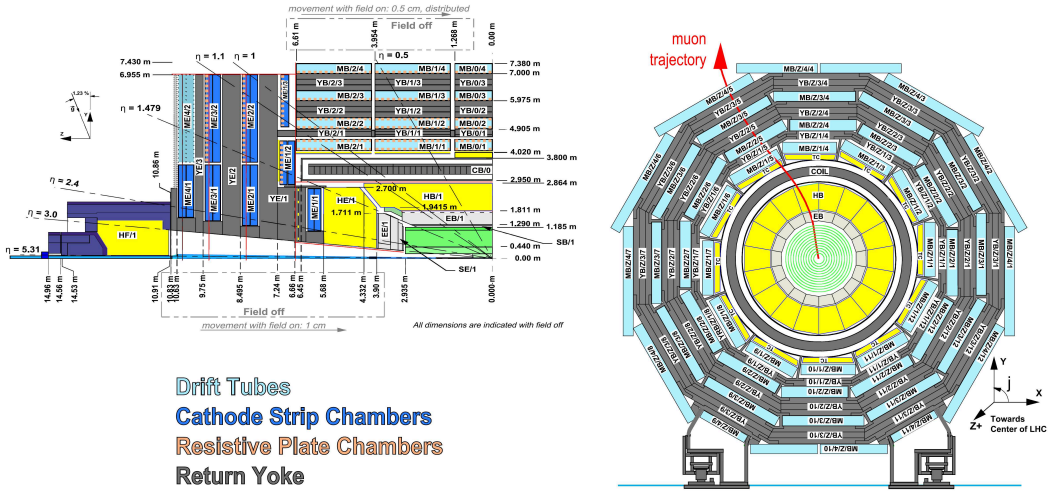


Figure 2.10: Transverse and longitudinal cross sections of the CMS detector showing the Muon System with particular emphasis on the different technologies used for detectors; the ME/4/2 CSC layers in the end-cap were included in the design but are not currently installed. Figures from [23].

System was designed to cope with three major functions: robust and fast identification of muons, good resolution of momentum measurement and triggering.

The Muon System is composed of three types of gaseous detectors, located inside the empty volumes of the iron yoke, and therefore arranged in barrel and end-cap sections. The coverage of Muon System is shown in figure 2.10.

In the barrel region the neutron-induced background is small and the muon rate is low; moreover, the field is uniform and contained in the yoke. For these reasons, standard drift chambers with rectangular cells are used. The barrel Drift Tubes (DT) cover the $|\eta| < 1.2$ region, are divided in five wheels in the beam direction and are organized in four stations housed among the yoke layers. The first three stations contain 12 cell planes, arranged in two superlayers providing measurement along $r\phi$ and one superlayerlayer measuring along z, each of them containing four layers. The fourth station provides measurement only in the transverse plane.

Both the muon rates and backgrounds are high in the forward region, where the magnetic field is large and non uniform. The choice for muon detectors fell upon cathode strip chambers (CSC) because of their fast response time, fine segmentation and radiation tolerance. Each end-cap is equipped with four stations of CSCs. The CSCs cover the $0.9 < |\eta| < 2.4$ pseudorapidity range. The cathode strips are oriented radially and provide precise measurement in the bending plane, the anode wires run approximately per-

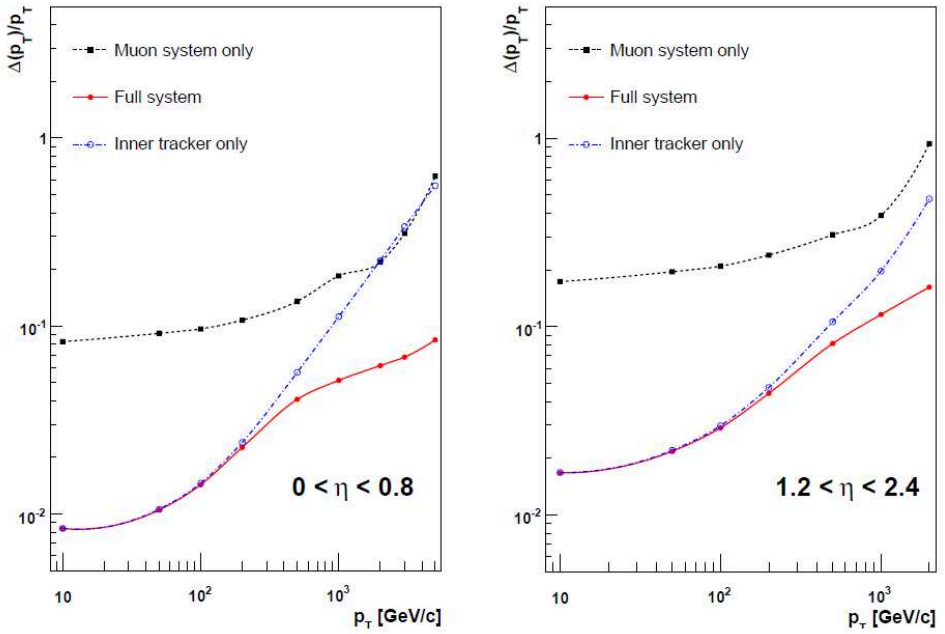


Figure 2.11: Resolution on p_T measurement of muons with the Muon System, the Silicon Tracker or both, in the barrel (left) and end-caps (right). Figures from [23].

pendicular to the strips and are read out to measure the pseudorapidity and the beam-crossing time of a muon. The muon reconstruction efficiency is typically $95 - 99\%$ except for the regions between two barrel DT wheels or at the transition between DTs and CSCs, where the efficiency drops.

Both the DTs and CSCs can trigger on muons with a Level 1 p_T (see section 2.2.5) resolution of 15% and 25%, respectively. Additional trigger-dedicated muon detectors were added to help measured the correct beam-crossing time. These are Resistive Plate Chambers (RPC), gaseous detector operated in the avalanche mode, which can provide independent and fast trigger with high segmentation and sharp p_T threshold over a large portion of the pseudorapidity range. The overall p_T resolution on muons is shown in figure 2.11, with emphasis on the different contribution from the Muon System and the Silicon Tracker.

Muon Reconstruction

Muon detection and reconstruction play a key role in the CMS physics program, both for the discovery of New Physics and for precision measurements of SM processes. CMS has been designed for a robust detection of muons over the entire kinematic range of the LHC and in a condition of very high background. The muon system allows an efficient and pure identification

of muons, while the inner tracker provides a very precise measurement of their properties. An excellent muon momentum resolution is made possible by the high-field solenoidal magnet. The steel flux return yoke provides additional bending power in the spectrometer, and serves as hadron absorber to facilitate the muon identification. Several muon reconstruction strategies are available in CMS, in order to fulfil the specific needs of different analyses. The muon reconstruction consists of three main stages:

1. local reconstruction: in each muon chamber, the raw data from the detector read-out are reconstructed as individual points in space; in CSC and DT chambers, such points are then fitted to track segments;

2. stand-alone reconstruction: points and segments in the muon spectrometer are collected and fitted to tracks, referred to as "stand-alone muon tracks";

3. global reconstruction: stand-alone tracks are matched to compatible tracks in the inner tracker and a global fit is performed using the whole set of available measurements: the resulting tracks are called "global muon tracks".

Muon identification represents a complementary approach with respect to global reconstruction: it starts from the inner tracker tracks and flags them as muons by searching for matching segments in the muon spectrometer. The muon candidates produced with this strategy are referred to as "tracker muons". After the completion of both algorithms, the reconstructed stand-alone, global and tracker muons are merged into a single software object, with the addition of further information, like the energy collected in the matching calorimeter towers. This information can be used for further identification, in order to achieve a balance between efficiency and purity of the muon sample.

2.2.4 Calorimetry

Identification of electrons, photons, and hadrons relies on accurate calorimetry, which is a destructive measurement of the energy of a particle. As in most of the particle physics experiments, a distinction is made between electromagnetic calorimetry and hadron calorimetry. Electromagnetic calorimetry is based on the production of EM showers inside a high-Z absorber, while hadron calorimetry measures the effects of hadron inelastic scattering with heavy nuclei, including production of photons from neutral pions and muons, and neutrinos from weak decays. Calorimetry must be precise and hermetic also to measure any imbalance of momenta in the transverse plane which can signal the presence of undetected particles such as high-p_T neutrinos.

The electromagnetic calorimeter of CMS, ECAL, is a homogeneous calorimeter, where the absorber material is the same as the sensitive one [28]. ECAL

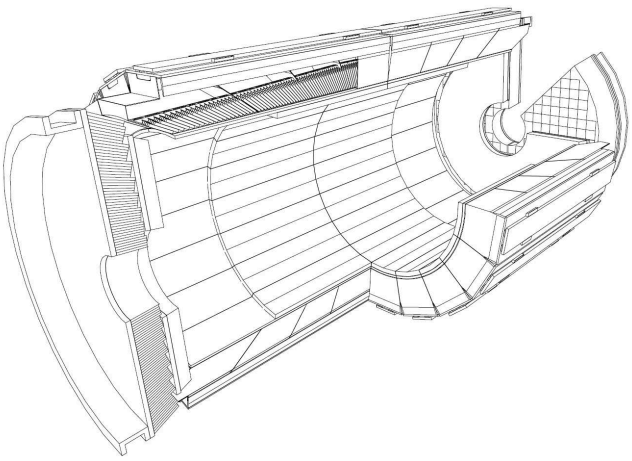


Figure 2.12: Cut-away view of the CMS ECAL showing the hierarchical structure of crystals arranged in supercystals and modules and the orientation of crystals whose major axis is always directed to the origin of the reference frame.

is composed of 61200 lead tungstate ($PbWO_4$) crystals in the barrel region and 7324 crystals in the end-caps, as shown in figure 2.12. The crystal cross-section is $22 \times 22\,\mathrm{mm}^2$ at the front face, while the length is 230 mm. End-caps are equipped with a preshower detector. Lead tungstate was chosen because of its high density, 8.28 g cm^{-3}, short radiation length, 0.89 cm, and small Molire radius, 2.2 cm. This way, the calorimeter can be kept compact with fine granularity, while scintillation and optical properties of $PbWO_4$ make it fast and radiation tolerant. Signal transmission exploits total internal reflection. Scintillation light detection relies on two different technologies. Avalanche photodiodes (APD) are used in the barrel region, mounted in pairs on each crystals, while vacuum phototriodes (VPT) are used in the end-caps. The preshower detector is a sampling calorimeter composed of lead radiators and silicon strips detectors, and it is used to identify neutral pions in the forward region. The nominal energy resolution, measured with electron beams having momenta between 20 and 250 GeV, is

$$\left(\frac{\sigma_E}{E}\right)^2 = \left(\frac{2.8\%}{\sqrt{E}}\right)^2 + \left(\frac{0.12}{E}\right)^2 + (0.30\%)^2 \quad ,$$

where all the energies are in GeV and the different contributions are respectively: the stochastic one (S), due to fluctuations in the lateral shower containment and in the energy released in the preshower, that due to electronics (N), digitization and pile-up, and the constant term (C), due to intercalibration errors, energy leakage from the back of the crystal and non-uniformity in light collection.

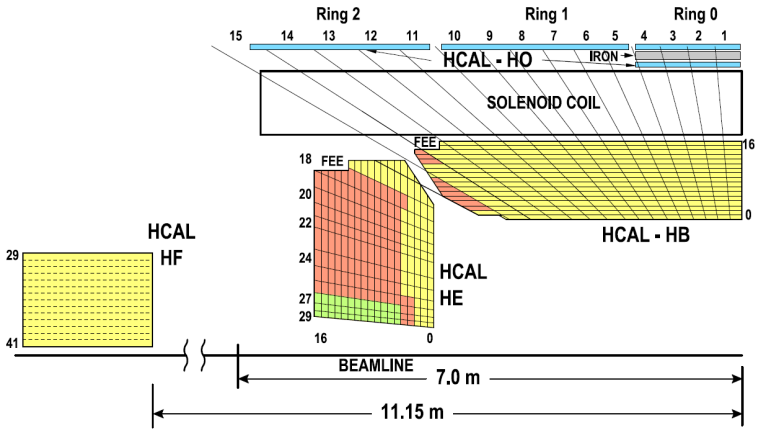


Figure 2.13: Cross section of the CMS HCAL showing the tower segmentation. Figure from [29].

The hadron calorimeter of CMS, HCAL, is a sampling calorimeter employed for the measurement of hadron jets and neutrinos or exotic particles resulting in apparent missing transverse energy [29]. A longitudinal view of HCAL is shown in figure 2.13. The hadron calorimeter size is constrained in the barrel region, $|\eta| < 1.3$, by the maximum radius of ECAL and the inner radius of the solenoid coil. Because of this, the total amount of the absorber material is limited and an outer calorimeter layer is located outside of the solenoid to collect the tail of the showers. The pseudorapidity coverage is extended in the $3 < |\eta| < 5.2$ by forward Cherenkov-based calorimeters. The barrel part, HB, consists of 36 wedges, segmented into 4 azimuthal sectors each, and made out of flat brass absorber layers, enclosed between two steel plates and bolted together without any dead material on the full radial extent. There are 17 active plastic scintillator tiles interspersed between the stainless steel and brass absorber plates, segmented in pseudorapidity to provides an overall granularity of $\Delta\phi \times \Delta\eta = 0.087 \times 0.087$. The same segmentation is maintained in end-cap calorimeters, HE, up to $|\eta| < 1.6$, while it becomes two times larger in the complementary region. The maximum material amount in both HB and HE corresponds to approximately 10 interaction lengths λ_I. The energy resolution on single electron and hadron jets is shown in figure 2.14.

2.2.5 Trigger and Data Acquisition

High bunch crossing rates and design luminosity at LHC correspond to approximately 20–25 superimposed events every 25 ns, for a total of 10^9 events per second. The large amount of data associated to them is impossible to

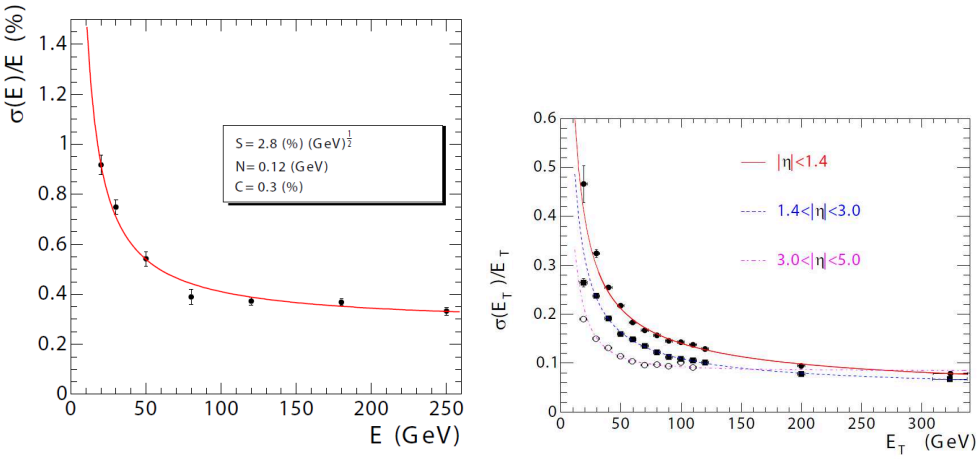


Figure 2.14: Left: ECAL energy resolution as a function of the electron energy as measured from a beam test. The energy was measured in a 3×3 crystals array with the electron impacting the central one. The stochastic, noise and constant terms are given. Right: the jet transverse energy resolution as a function of the transverse energy for barrel jets, end-cap jets and very forward jets reconstructed with an iterative cone algorithm with cone radius $R = 0.5$. Figures from [23].

store and process, therefore a dramatic rate reduction has to be achieved. This is obtained with two steps: the Level 1 Trigger [17] and the High Level Trigger, HLT [26].

The Level 1 Trigger is based on custom and programmable electronics, while HLT is a software system implemented on a $\sim$ 1000 commercial processors farm. The maximum allowed output rate for Level 1 Trigger is 100 kHz, which should be even kept lower, about 30 kHz, for safe operation. Level 1 Trigger uses rough information from coarse segmentation of calorimeters and Muon Detectors and holds the high-resolution data in a pipeline until acceptance/rejection decision is made. HLT exploits the full amount of collected data for each bunch crossing accepted by Level 1 Trigger and is capable of complex calculations such as the off-line ones. HLT algorithms are those expected to undergo major changes in time, particularly with increasing luminosity. Configuration and operation of the trigger components are handled by a software system called Trigger Supervisor.

The Level 1 Trigger relies on local, regional and global components. The Global Calorimeter and Global Muon Triggers determine the highest-rank calorimeter and muon objects across the entire experiment and transfer them to the Global Trigger, the top entity of the Level 1 hierarchy. The latter takes the decision to reject an event or to accept it for further evaluation by the HLT. The total allowed latency time for the Level 1 Trigger is 3.2 μs.

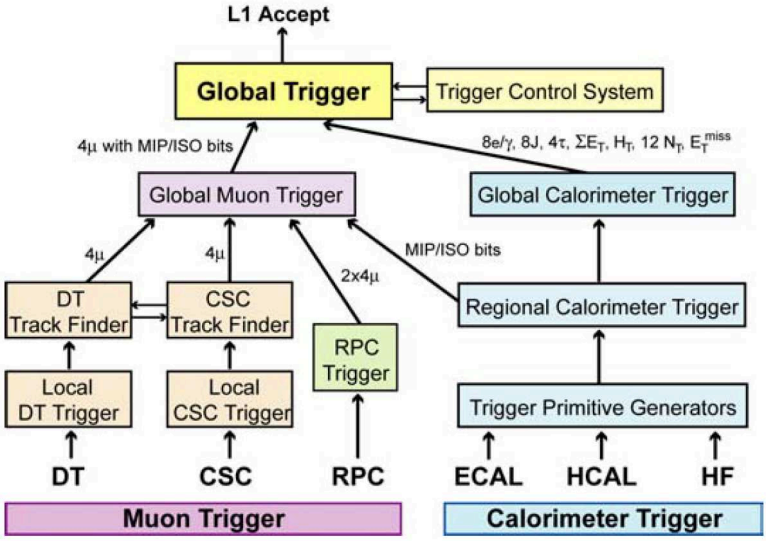


Figure 2.15: Schematic representation of the Level 1 Trigger data flow.

A schematic representation of the Level 1 Trigger data flow is presented in figure 2.15.

Muon Trigger

All Muon Detectors – DT, CSC and RPC – contribute to the Trigger. Barrel DTs provide Local Trigger in the form of track segments in ϕ and hit patterns in η. End-cap CSCs provide 3-dimensional track segments. Both CSCs and DTs provide also timing information to identify the bunch crossing corresponding to candidate muons. The Local DT Trigger is implemented in custom electronics. BTIs, Bunch and Track Identifiers, search for coincidences of aligned hits in the four equidistant planes of staggered drift tubes in each chamber superlayer. From the associated hits, track segments defined by position and angular direction are determined. TRACOs, Track Correlators, attempt to correlate track segments measured in the two ϕ superlayers of each DT chamber, enhancing the angular resolution and producing a quality hierarchy.

The requirement of robustness implies redundancy, which introduces, however, a certain amount of noise or duplicate tracks giving rise to false Triggers. Therefore the BTIs, the TRACOs and the different parts of the Local Trigger contain complex noise and ghost reduction mechanisms. The position, transverse momentum and quality of tracks are coded and transmitted to the DT regional Trigger, called the Drift Tube Track Finder (DTTF), through high-speed optical links.

The Global Muon Trigger (GMT) combines the information from DTs,

CSCs and RPCs, achieving an improved momentum resolution and efficiency compared to the stand-alone systems. It also reduces the Trigger rate and suppresses backgrounds by making use of the complementarity and redundancy of the three Muon Systems. The Global Muon Trigger also exploits MIP/ISO bits[4] from the Regional Calorimeter Trigger. A muon is considered isolated if its energy deposit in the calorimeter region from which it emerged is below a defined threshold. DT and CSC candidates are first matched with barrel and forward RPC candidates based on their spatial coordinates. If a match is possible, the kinematic parameters are merged. Several merging options are possible and can be selected individually for all track parameters, taking into account the strengths of the individual Muon Systems. Muons are back-extrapolated through the calorimeter regions to the vertex, in order to retrieve the corresponding MIP and ISO bits, which are then added to the GMT output and can be taken into account by the Global Trigger (GT). Finally, the muons are sorted by transverse momentum and quality to deliver four final candidates to the GT. The Muon Trigger is designed to cover up to $|\eta| < 2.4$.

Global Trigger

The Global Trigger takes the decision to accept or reject an event at Level 1, based on candidate e/γ, muons, jets, as well as global quantities such as the sums of transverse energies (defined as $E_T = E \sin\theta$), the missing transverse energy and its direction, the scalar transverse energy sum of all jets above a chosen threshold (usually identified by the symbol H_T), and several threshold-dependent jet multiplicities. Objects representing particles and jets are ranked and sorted. Up to four objects are available and characterized by their p_T or E_T, direction and quality. Charge, MIP and ISO bits are also available for muons. The Global Trigger has five basic stages implemented in Field-Programmable Gate-Arrays (FPGAs): input, logic, decision, distribution and read-out. If the Level 1 Accept decision is positive, the event is sent to the Data Acquisition stage.

High Level Trigger and Data Acquisition

The CMS Trigger and DAQ system is designed to collect and analyse the detector information at the LHC bunch crossing frequency of 40 MHz. The DAQ system must sustain a maximum input rate of 100 kHz, and must provide enough computing power for a software filter system, the High Level Trigger (HLT), to reduce the rate of stored events by a factor of 1000. In CMS all events that pass the Level 1 Trigger are sent to a computer farm

[4]The MIP bit is set if the calorimeter energy is consistent with the passage og a minimum ionizing particle, the isolation bit is set if a certain energy threshold in the trigger towers surrounding the muon is not exceeded.

(Event Filter) that performs physics selections, using faster versions of the offline reconstruction software, to filter events and achieve the required output rate. The various subdetector front-end systems store data continuously in 40 MHz pipelined buffers. Upon arrival of a synchronous Level 1 Trigger Accept via the Timing, Trigger and Control System (TTCS) the corresponding data are extracted from the front-end buffers and pushed into the DAQ system by the Front-End Drivers (FEDs). The event builder assembles the event fragments belonging to the same Level 1 Trigger from all FEDs into a complete event, and transmits it to one Filter Unit (FU) in the Event Filter for further processing. The DAQ system includes back-pressure from the filter farm through the event builder to the FEDs. During operation, Trigger thresholds and pre-scales will be optimized in order to fully utilize the available DAQ and HLT throughput capacity.

2.3 Monte Carlo Event Generator

Monte Carlo (MC) event generators provide an event-by-event prediction of complete hadronic final states based on QCD calculation. They allow to study the topology of events generated in hadronic interactions and are used as input for detector simulation programs to investigate detector effects. The event simulation is divided into different stages as illustrated in figure 2.16. First, the partonic cross section is evaluated by calculating the matrix element in fixed order pQCD. The event generators presently available for the simulation of proton-proton collisions provide perturbative calculations for beauty production up to NLO. Higher order corrections due to initial and final state radiation are approximated by running a parton shower algorithm. The parton shower generates a set of secondary partons originating from subsequent gluon emission of the initial partons. It is followed by the hadronization algorithm which clusters the individual partons into colour-singlet hadrons. In a final step, the short lived hadrons are decayed. In the framework of the analysis presented here, the MC event generator PYTHIA 6.4 [43] is used to compute efficiencies, kinematic distributions, and for comparisons with the experimental results. This programs were run with its default parameter settings, except when mentioned otherwise.

PYTHIA

In the PYTHIA program, the matrix elements are calculated in LO pQCD and convoluted with the proton PDF, chosen herein to be CTEQ6L1. The mass of the b-quark is set to $m_b = 4.8$ GeV. The underlying event is simulated with the D6T tune. Pile-up events were not included in the simulation. The parton shower algorithm is based on a leading-logarithmic approximation for QCD radiation and a string fragmentation model (implemented in JETSET) is applied. The longitudinal fragmentation is described by the

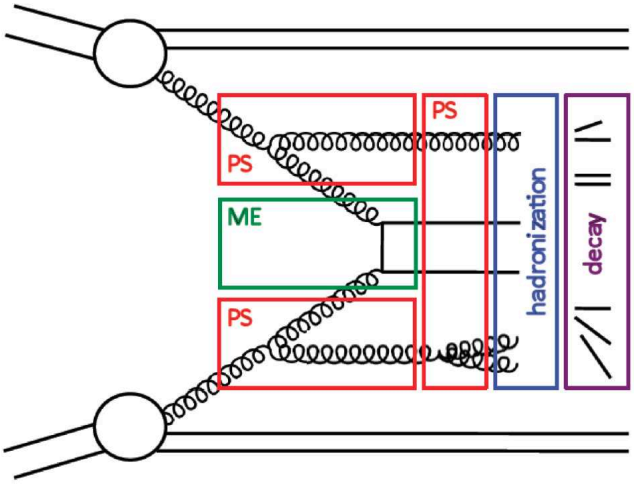


Figure 2.16: Schematic view of the subsequent steps of a MC event generator: matrix element (ME), parton shower (PS), hadronization and decay.

Lund symmetric fragmentation function for light quarks and by the Peterson fragmentation function for charm and beauty quarks, that is

$$f(z) \propto \frac{1}{z\left[1 - \frac{1}{z} - \frac{\varepsilon_Q}{(1-z)}\right]^2} \quad ,$$

where z is defined as

$$z = \frac{(E + p_{\parallel})_{\text{hadron}}}{(E + p)_{\text{quark}}} \quad ,$$

$(E + p_{\parallel})_{\text{hadron}}$ is the sum of the energy and momentum component parallel to the fragmentation direction carried by the primary hadron, $(E + p)_{\text{quark}}$ is the energy-momentum of the quark after accounting for initial state radiation, gluon bremsstrahlung and photon radiation in the final state. The parameters of the Peterson fragmentation function are set to $\epsilon_c = 0.05$ and $\epsilon_b = 0.005$. In order to estimate the systematic uncertainty introduced by the choice of the fragmentation function, samples generated with different values of ϵ_b are studied. The hadronic decay chain used in PYTHIA is also implemented by the JETSET program. For comparison, additional event samples are generated where the EvtGen program is used to decay the b-hadrons. EvtGen is an event generator designed for the simulation of the physics of b-hadron decays, and in particular provides a framework to handle complex sequential decays and CP violating decays.

Chapter 3

Data collection and event selection

The data used for this analysis have been collected by CMS detector during 2012 pp run, at a centre-of-mass energy $\sqrt{s} = 8\,\text{TeV}$. The integrated luminosity collected and certified is $20.5\,\text{fb}^{-1}$.

The events have been selected by two-level criteria: firstly an online trigger selection is used during data taking, and then an offline selection and candidate identification is performed after the events are fully reconstructed.

3.1 Online event selection

All the events used in this analysis, both for signal regions and control regions, are selected by a single trigger. This requires the presence of at least one pair of reconstructed muons in the event with opposite charge, each of them with a transverse momentum greater than 3.5 GeV, a pseudo-rapidity smaller, in module, than 2.2, and a distance of closest approach with respect to the beam axis smaller than 2 cm. The dimuon system is required to have a transverse momentum greater than 6.9 GeV, an invariant mass between 1 GeV and 4.8 GeV, and a distance of closest approach (DCA) between the muons smaller than 0.5 cm. In addition, the two muons are required to form a common vertex, with fit χ^2 probability greater than 10%, a flight distance significance with respect to the beamspot, measured in the plane transverse to the beam axis, greater than 3, and $\cos\alpha > 0.9$, where α is the angle in the transverse plane between the dimuon momentum vector and the vector from the beamspot to the dimuon vertex.

No requirements on the hadronic particles in the final state are present at trigger level.

3.2 Offline candidate identification

In the offline selection the full final state, composed by two muons and two hadrons, is reconstructed. A set of four reconstructed objects compatible with those four particles is considered as a candidate. The offline cuts are applied independently to each candidate in an event.

Candidate pre-selection cuts

The two reconstructed muons are required to have opposite charge and match those that triggered the event. This is done by requiring $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.1$, where $\Delta\eta$ and $\Delta\phi$ are the pseudorapidity and azimuthal angle differences between the directions of the muons reconstructed at trigger level and in the offline analysis. In addition, they have to satisfy general muon identification requirements: the muon track candidate from the silicon tracker must match a track segment from the muon detector, the χ^2 per degree of freedom in a global fit to the silicon tracker and muon detector hits must be less than 1.9, there must be at least 6 silicon tracker hits, including at least 2 from the pixel detector, and the transverse and longitudinal impact parameters with respect to the beamspot must be less than 3 cm and 30 cm, respectively. The same requirements applied at trigger level and described in Section 3.1 are also applied to the offline reconstructed dimuon system.

The two charged hadron candidates are required to have opposite charge and each of them must fail the muon identification criteria.

The B^0 candidates are obtained by fitting a second time the four charged tracks after applying a common vertex constraint. This operation is used to improve the resolution of the track parameters. The B^0 candidates is required to have a transverse momentum greater than 8 GeV, and a pseudorapidity smaller, in module, than 2.2.

Since the CMS detector does not have particle identification capability, in each candidate is still present an ambiguity: the mass of the kaon can be assigned to the positive charged hadron track and the mass of the pion to the negative one to reconstruct a K^{*0} candidate, or viceversa to reconstruct a $\overline{K}^{*0}$ candidate. The invariant mass m of the B^0 candidate is required to be within 280 MeV of the nominal B^0 mass (m_{B^0}) [39], for at least one of the mass assignment hypothesis, $K^-\pi^+\mu^+\mu^-$ or $K^+\pi^-\mu^+\mu^-$.

The mass sideband is defined as the set of B^0 candidates with $3\sigma_m < |m - m_{B^0}| < 280$ MeV, where σ_m is the average mass resolution ($\approx$45 MeV) as obtained from fitting the m distribution of simulated signal events with a sum of two Gaussian functions with a common mean.

Candidate optimised selection cuts

The two charged hadrons of the candidate are required to have a transverse momentum greater than 0.8 GeV, and a significance of the extrapolated distance d of closest approach to the beamspot in the transverse plane greater 2. The uncertainty associated to d is defined as the sum in quadrature of the uncertainty of the track position and the beamspot transverse size.

For at least one of these two identity assignment hypotheses, the hadron pair invariant mass is requested to be within 90 MeV of the nominal K^{*0} mass [39].

The B^0 vertex fit χ^2 probability must be larger than 10%, while the distance from the beamspot in the transverse plane, L, must be greater than 12 times the sum in quadrature of the uncertainty on L and the beamspot transverse size. Then, $\cos\alpha_{xy}$, where α_{xy} is the angle in the transverse plane between the B^0 momentum vector and the line-of-flight between the beamspot and the B^0 vertex, is required to be larger than 0.9994.

The values of the cuts on the hadronic track transverse momentum and d significance, the K^{*0} mass window, and the B^0 candidate vertex fit probability, displacement significance and pointing angle are optimised by maximising the signal significance in the region $|m - m_{\mathrm{B}^0}| < 2.5\sigma_m$, using signal event samples from simulation and background event samples from sideband data in m.

After applying these selection criteria, about 5% of the events have more than one candidate. For these events, a single candidate is chosen based on the best B^0 vertex χ^2 probability.

Additional selection cuts

For each candidate, the dimuon invariant mass q and its uncertainty σ_q are calculated. Two control samples, corresponding to the $\mathrm{B}^0 \to \mathrm{J}/\psi\mathrm{K}^{*0}$ and $\mathrm{B}^0 \to \psi'\mathrm{K}^{*0}$ decay channels, are defined by the requirements $|q - m_{\mathrm{J}/\psi}| < 3\sigma_q$ and $|q - m_{\psi'}| < 3\sigma_q$, respectively, where $m_{\mathrm{J}/\psi}$ and $m_{\psi'}$ are the nominal masses [39] of the indicated meson. On average, the value of σ_q is about 26 MeV.

The distribution of the B^0 invariant mass for the events in signal q^2 regions is shown in Figure 3.1, while the same distributions for events in the control regions are shown in Figure 3.2. The peaks corresponding to the B^0 decays are clearly visible in all three distributions, and can be distinguished from the exponential background shape.

A strong contamination from $\mathrm{B}^0 \to \mathrm{J}/\psi\mathrm{K}^{*0}$ and $\mathrm{B}^0 \to \psi'\mathrm{K}^{*0}$ decays is still present in the sample of events passing the selections, mainly because of unreconstructed soft photons in the charmonium decay, i.e., J/ψ or $\psi' \to \mu^+\mu^-\gamma$. These events have $q < m_{\mathrm{J}/\psi}$ and $q < m_{\psi'}$, respectively, and are not included in the control sample described above. They also have m

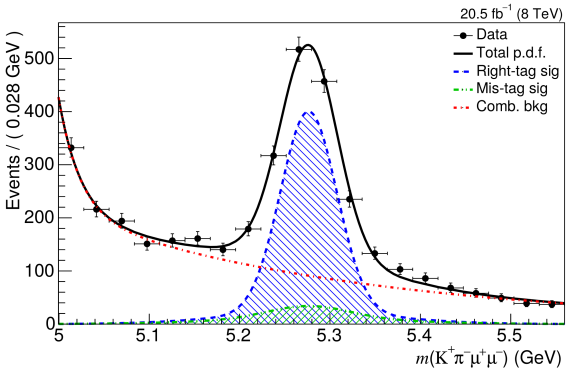


Figure 3.1: B^0 invariant mass from data, computed from the whole q^2 spectrum excluding the J/ψ and ψ' ranges as described in the text. Just to guide the eye the plot is fitted with a double Gaussian function with unique mean to measure the signal yield (1232 ± 44 events) and with two Gaussian functions and a double exponential to distinguish the background.

value lower than m_{B^0}, and they can be efficiently removed by a combined requirement on q and m. For $q < m_{J/\psi}$ ($q > m_{J/\psi}$), it is required that $|(m - m_{B^0}) - (q - m_{J/\psi})| > 160\,(60)$ MeV. For $q < m_{\psi'}$ ($q > m_{\psi'}$), it is required that $|(m - m_{B^0}) - (q - m_{\psi'})| > 60\,(30)$ MeV. These cuts are tuned using MC simulations, in such a way that less than 10% of the background events with q^2 values close to the control regions originate from the control channels.

To avoid the contamination from $\phi \to K^+K^-$ decays, we additionally require that the invariant mass of the hadron pair, in the hypothesis that both tracks have the kaon mass, $m(K^+K^-)$, is larger than 1.035 GeV. This cut has been tuned using the data/MC comparison of the $m(K^+K^-)$ distribution in the $B^0 \to J/\psi(\mu^+\mu^-)K^{*0}$ control channel, as shown in Figure 3.3.

After applying the full set of requirements described here 3191 events remain in the data sample, including the sideband region.

CP-state assignment

The selected four-track candidate is identified as a B^0 or $\overline{B}^0$, and the corresponding masses are assigned to the hadronic tracks, depending on whether the $K^+\pi^-$ or $K^-\pi^+$ invariant mass hypothesis is closest to the nominal K^{*0} mass. The candidates assigned to the correct state will be called right-tagged, while we will refer to the candidates assigned to the incorrect state as mis-tagged. The fraction of mis-tagged events is estimated from simulation to be in the range 12–14%, depending on q^2.

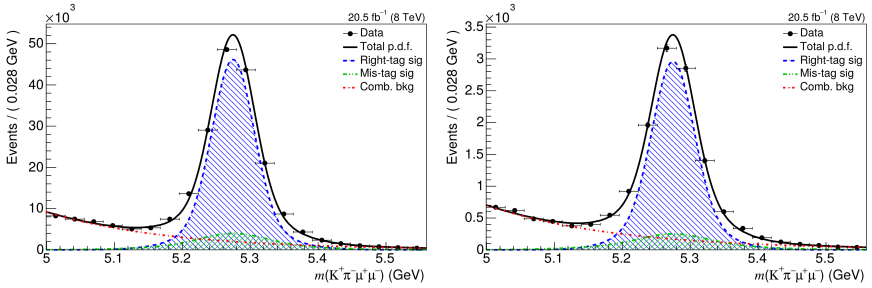


Figure 3.2: B^0 invariant mass for both control channels, $B^0 \to J/\psi(\mu^+\mu^-)K^{*0}$ (left) and $B^0 \to \psi'(\mu^+\mu^-)K^{*0}$ (right), from data. Just to guide the eye the plot is fitted with a double Gaussian function with unique mean and an exponential to describe the signal and the background respectively.

3.3 Dimuon mass square binning

The q^2 range used in this analysis extends from 1 GeV ^{2} to 19 GeV ^{2} and it is divided in nine bins, defined in Table 3.1. The bin definition is the same used in the previous CMS angular analysis on the same dataset and it is the result of a compromise between being coherent with the definition used in the previous measurements and having an expected signal yield homogeneously distributed over the q^2 bins.

Table 3.1: Range definition of the dimuon invariant mass bins.

Bin index	q^2 range (GeV^2/c^4)
0	1-2
1	2-4.3
2	4.3-6
3	6-8.68
4	8.68-10.09 (J/Ψ region)
5	10.09-12.86
6	12.86-14.18 (Ψ' region)
7	14.18-16
8	16-19

The selection criteria and the analysis techniques are identical for any q^2 bin and in each of them the analysis is performed independently. The two q^2 bins containing the control channel regions are not used to fit the signal events.

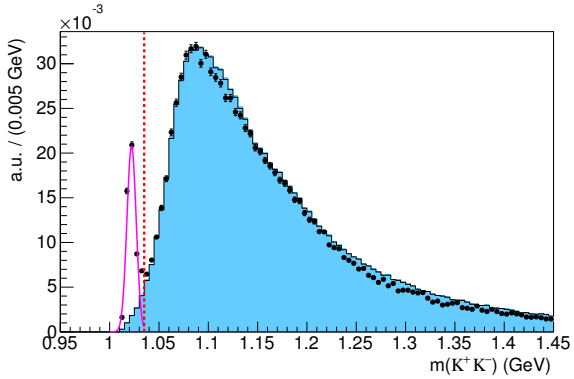


Figure 3.3: Invariant mass of the two hadron tracks when the kaon mass is assigned to both hadrons. The plot is obtained after applying all selections but the one on the invariant mass of the two hadron tracks with kaon mass assigned. The two superimposed plots are obtained from simulation and data, the former with the control channel $B^0 \to J/\psi(\mu^+\mu^-)K^{*0}$, the latter is background subtracted (the plot of the sidebands is subtracted from the plot of the signal region), and no J/ψ nor ψ' rejections are applied, therefore the spectrum is dominated by events from the decay $B^0 \to J/\psi(\mu^+\mu^-)K^{*0}$. The first peak on the left corresponds precisely to the ϕ particle ($m(\phi) = 1020\,\text{MeV}$). Superimposed to the ϕ peak there is shown the Gaussian fit ($\sigma = 4.7 \pm 0.1\,\text{MeV}$). The vertical dashed line corresponds to the 1.035 GeV selection cut.

Chapter 4

Analysis strategy

The angular parameters are extracted through an unbinned fit to four variables, the $K^+\pi^-\mu^+\mu^-$ invariant mass and the three angular variables, using an extended maximum likelihood estimator.

A *blinding* procedure has been used for this analysis, to avoid any bias due to human decisions during the set up of the analysis strategy. During the steps of event selection, fit strategy definition and validation, and efficiency measurement, only the MC samples and data from the mass sidebands have been used, while the distributions of the events in the signal region were not considered. Most of the systematic uncertainty studies have been performed with *blinded* datasets. Some of these studies have been firstly defined and tested on MC samples and data sidebands, then propagated on data signal after the *unblinding* step. Some other systematic uncertainties, when specified in the description, have been studied directly on data, after the *unblinding* step. After the analysis strategy and all the steps described here were defined and approved by a committee of reviewer inside the Collaboration, the data in the signal region were *unblinded* and the fit procedure run on them. The estimation of the statistical uncertainty has been performed entirely after the data *unblinding*.

In the following sections I will give a detailed description of the probability density function used in the fit, and of the methods used to estimate its parameters.

4.1 Probability density function

The probability density function (`pdf`) used in the fit has the following expression:

$$\begin{aligned}\mathrm{pdf}(m,\theta_K,\theta_l,\phi) = Y_S^C \Big[& S^C(m)\, S^a(\cos\theta_K,\cos\theta_l,\phi)\, \epsilon^C(\theta_K,\theta_l,\phi) \\ & + \frac{f^M}{1-f^M}\, S^M(m)\, S^a(-\cos\theta_K,-\cos\theta_l,\phi)\, \epsilon^M(\theta_K,\theta_l,\phi) \Big] \\ & + Y_B\, B^m(m)\, B^{\theta_K}(\cos\theta_K)\, B^{\theta_l}(\cos\theta_l)\, B^{\phi}(\phi),\end{aligned} \tag{4.1}$$

where its three terms correspond to the `pdf`s for right-tagged signal, mis-tagged signal, and background events, respectively.

The pure-physics information about the angular distribution, derived from the theoretical description of the decay presented in Section 1.2, are contained in $S^a(\cos\theta_K,\cos\theta_l,\phi)$. This term is defined as the differential decay distribution, which has been formulated, after the angular folding application, in Equation 1.25.

In order to describe the data collected in the collisions, we should adapt the theoretical differential distribution to the experimental conditions. They include the finite resolution of the measured kinematical variables and the distortion of the angular distributions introduced by the geometric acceptance of the detector and the selection criteria. All these effect are applied by multiplying the decay distribution by the efficiency function $\epsilon^C(\theta_K,\theta_l,\phi)$.

As explained in Section 3.2, there is a non-negligible probability that the wrong CP-state is assigned to the decay and the wrong mass hypothesis is assigned to the hadrons. However, these mis-tagged events still contain information about the decay distribution and their contribution can be taken into account in the `pdf`. The choice of the wrong K and π mass hypothesis and CP-state causes some differences in the angle definition; using the Equations 1.10, 1.11, and 1.12, one can obtain that the correct term to use in the `pdf`, as a function of the reconstructed variables, is $S^a(-\cos\theta_K,-\cos\theta_l,\phi)$. The efficiency function for the mis-tagged events has been computed separately: $\epsilon^M(\theta_K,\theta_l,\phi)$. Note that ϵ^M is defined as a function of the reconstructed variables, so there is no need to correct the signs of the angles. The term describing the mis-tagged event contribution is multiplied by $f^M/1-f^M$ to scale it correctly with respect to the right-tagged event contribution.

As a last point, the `pdf` must be able to describe the background contamination and distinguish it from the signal events. The set of functions $B^{\theta_K}(\cos\theta_K)\, B^{\theta_l}(\cos\theta_l)\, B^{\phi}(\phi)$ describe the angular distribution of the background events. However, to distinguish signal and background events the three angular variables are not useful, and a new variable is needed; the

best kinematic variable for this use is the mass of the B^0 candidate. Each of the three terms in the pdf is then multiplied by a function describing the mass distribution: $B^m(m)$ for the background events, $S^C(m)$ for the right-tagged signal events, and $S^M(m)$ for the mis-tagged signal events. Even if mis-tagged events are events produced by a B^0 decay, like the right-tagged events, they need a different mass-shape function because the wrong mass assignment for the hadronic tracks produce a different four-body invariant mass value.

A factor Y_B is used to multiply the background term, while a factor Y_S^C is used with the signal terms, to correctly scale the two contributions. To give a physical meaning of number of events to these two parameters, the background term multiplying Y_B and the right-tagged signal term multiplying Y_S^C are normalised to 1 over the full phase-space. Note that, with this definition, Y_S^C stands for the number of right-tagged events, not for the total number of signal events.

4.2 Parameterisation of the pdf terms

Each of the terms $S^C(m)$ and $S^M(m)$ is composed as the sum of two Gaussian functions with different normalisation factors. For each q^2 bin, all these four Gaussian functions are constrained to have the same mean value.

The determination of the mistag fraction f^M will be presented in Section 4.2.1.

The construction of the three-dimensional functions $\epsilon^C(\theta_K, \theta_l, \phi)$ and $\epsilon^M(\theta_K, \theta_l, \phi)$ will be described in Section 5.

The background mass shape $B^m(m)$ is parameterised by an exponential function, while the angular shapes $B^{\theta_K}(\cos\theta_K)$ and $B^{\theta_l}(\cos\theta_l)$ are polynomials functions, ranging from second to fourth degree depending on the q^2 bin, and $B^{\phi}(\phi)$ is a first-order polynomial function. Some more details on the background description are given in Section 4.2.2. The factorisation assumption of the background term in Equation (4.1) is discussed in Section 4.2.2.

4.2.1 Fraction of mis-tagged signal events

The algorithm used to tag the CP-state of the B^0 candidate, described in Section 3, has an intrinsic percentage of failure which is referred to as mistag fraction, f^M, defined as the ratio of mis-tagged signal events divided by the total number of signal events. The mistag fraction is determined from simulated events, comparing the results of the tag method to the MC truth, and by counting the number of correctly and wrongly tagged events. The resulting mistag fractions, for each q^2 bin, are shown in Table 4.1.

The mistag fraction parameter is fixed in the fit to the values evaluated on MC simulation. Any differences between simulated and real events could

Table 4.1: Mistag fraction as determined from simulated MC samples, for each q^2 bin. The values in q^2 bins 4 and 6 are evaluated on $B^0 \to J/\psi(\mu^+\mu^-)K^{*0}$ and $B^0 \to \psi'(\mu^+\mu^-)K^{*0}$ MC samples, respectively. For the other q^2 bins the values are evaluated on signal MC sample.

q^2 bin index	Mistag fraction f^M	Statistical uncertainty
0	0.124	0.002
1	0.129	0.001
2	0.134	0.001
3	0.132	0.001
4	0.1373	0.0005
5	0.132	0.001
6	0.140	0.002
7	0.132	0.001
8	0.137	0.001

lead to a bias in the mistag fraction values used in the `pdf`, that could propagate to a bias in the analysis results. The contribution of this effect to the systematic uncertainty is discussed in Section 7.5.

4.2.2 Background parameterisation

Several kind of background events can contaminate the dataset used for the fit. In this section I will present the studies performed to describe and evaluate the many sources of this contamination.

The main contribution derives from the combinatorial background, i.e. events in which the four particles of the final state do not come from the same decay vertex. Since the four-body invariant mass distribution of these events does not show any structure, their contribution can be evaluated from the mass sidebands, and extrapolated to the full mass range.

In addition, no correlation is expected between the four-body mass distribution and the angular variable distributions. This allows to estimate the shapes of the angular distributions on the sidebands, and assume them valid also to describe the combinatorial background contamination in the B^0 mass region.

The `pdf` used to model the angular shape of this background is the product of three uncorrelated polynomial functions, $B^{\theta_K}(\cos\theta_K)\, B^{\theta_l}(\cos\theta_l)\, B^{\phi}(\phi)$. The orders of the polynomial functions are chosen individually for each q^2 bin, in order to have them successfully describe any shape feature, while keeping the degrees of freedom as low as possible, to have a converging fit even with the low sideband statistics. The polynomial degrees that we use are listed in the Table 4.2.

Table 4.2: Degree of the polynomial functions used to described the angular shape of the combinatorial background distributions, for each q^2 bin.

q^2 bin index	B^{θ_K} degree	B^{θ_l} degree	B^{ϕ} degree
0	3	2	1
1	4	2	1
2	4	3	1
3	2	4	1
5	4	2	1
7	2	3	1
8	2	2	1

An example of the sideband event distributions, for q^2 bin 0, are plotted in Figure 4.1-4.3, together with the projections of the combinatorial background `pdf` on the angular variables, after the fit to the sideband distributions described in Section 4.3.

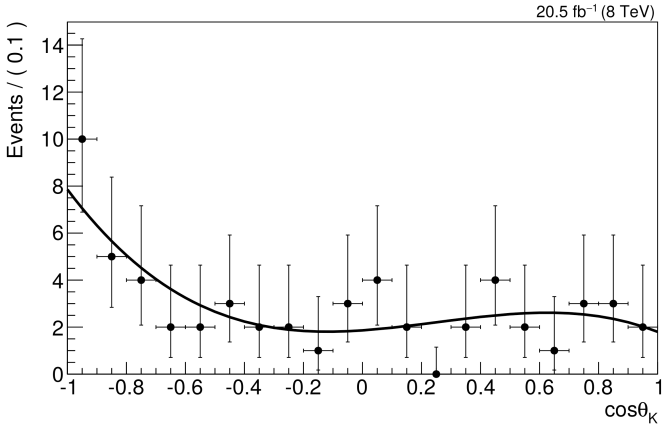


Figure 4.1: The sideband event distribution and the projection of the background `pdf` as a function of $\cos\theta_K$, for q^2 bin 0.

Other background sources has been investigated. None of them has been found relevant enough to be require a dedicated component in the fit `pdf`. In the following paragraphs I will present the studies performed on these possible sources of contamination and the estimates on their impact on the analysis results.

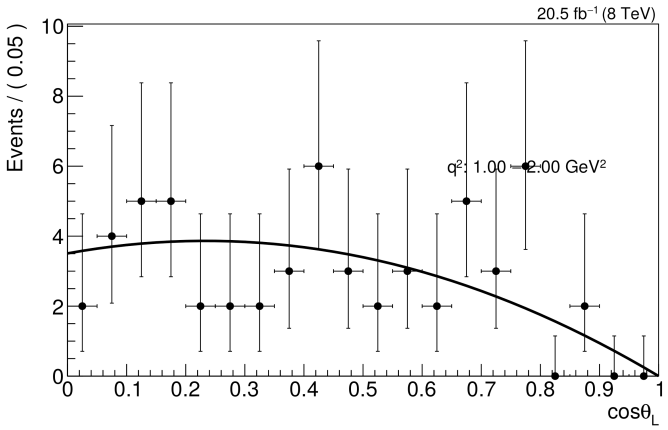


Figure 4.2: The sideband event distribution and the projection of the background `pdf` as a function of $\cos\theta_l$, for q^2 bin 0.

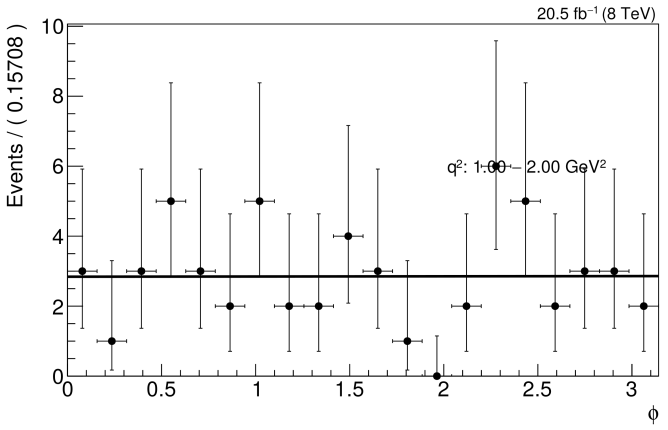


Figure 4.3: The sideband event distribution and the projection of the background `pdf` as a function of ϕ, for q^2 bin 0.

Feed-through background from control channels

The second background contamination for relevance is due to the "feed-through" of events from the resonant control channels, $B^0 \to J/\psi K^{*0}$ and $B^0 \to \psi' K^{*0}$, that are not rejected by the diagonal $q - m$ cuts and populate the q^2 bins adjacent to the resonance regions, i.e. bin 3, 5, and 7.

The `pdf` component needed to describe the distributions of these events is not included in the fit, because of its complex form and the fit general instability. A systematic uncertainty has been set up to describe the bias introduced by this omission on the fit results. It is described in Section 7.10.

Peaking background

Using a set of simulated MC samples, possible background sources that peak in the B^0 mass region are investigated.

A set of inclusive MC samples of B^0, B_s^0, B^+, and Λ_b hadrons decaying to $J/\psi(\mu^+\mu^-)X$ and $\psi'(\mu^+\mu^-)X$, have been probed, after applying the full set of selection criteria, but the control region removal and the diagonal cuts. No hints for peaking structures in the region of the B^0 mass are found. The mass distribution of these events is similar to the shape of the combinatorial background.

The same checks are performed with a MC sample of $B_s^0 \to K^{*0}(K^+\pi^-)\mu^+\mu^-$ decays. Normalising the area of the cluster near the B_s^0 mass to the correct luminosity and assuming the same branching fraction as for $B^0 \to K^{*0}(K^+\pi^-)\ \mu^+\mu^-$, about 70 such events, integrated over q^2, are present. Since the branching fraction of the $B_s^0 \to K^{*0}\mu^+\mu^-$ decay has never been measure, we can assume that the same ratio of branching fractions measured for the resonant channels, $\mathcal{B}(B_s^0 \to J/\psi K^{*0})/\mathcal{B}(B^0 \to J/\psi K^{*0}) \approx 10^{-2}$, is valid for the non-resonant channels as well. In this case, this background can be considered negligible since we expect less than one event in the whole q^2 range.

In the same way, the MC samples containing $B^+ \to K^+\mu^+\mu^-$ and $\Lambda_b \to pK\mu^+\mu^-$ events are tested. In this case the mass values assigned to the hadrons in the final state are changed with respect to the original ones, to one kaon mass and one pion mass, assigned according to the tagging criterion described in Section 3.2. Both these potential sources of background are found to be negligible.

Background from muon misidentification

A possible background contamination can come from events with hadrons misidentified as muons, or vice-versa. To peak under in the signal mass region, these events can be either a four-hadron final states from B meson decay, like the D meson mediated decays $B^0 \to DX$, with a pair of opposite-sign hadrons misidentified as muons, or a control-channel decay $B^0 \to J/\psi(\mu^+\mu^-)K^{*0}(K^+\pi^-)$ in which both one of the two muons has been misidentified as hadron and one of the two hadrons has been misidentified as muon. These contributions are calculated to be negligible because of the good muon identification capabilities of the CMS detector [24].

Test of factorisable background hypothesis

The result of a test on the correlation of the variable distributions for the combinatorial background events is shown in this section.

For each angular variable, the mass sideband sample has been divided in two sub-samples, cutting in the middle of the variable range. Then, the two

sub-sample distributions as functions of the other variables are compared. Furthermore, the angular variable distributions of the lower and higher sidebands are compared.

These comparisons are performed for each signal bin, for each control sample, and for two "special bins" that merge events in bin ranges [0-3] and [8-9], respectively. As example, the distributions for the special q^2 bin [0-3] are showed in Figure 4.4, where four plots are shown, one for each variable, and in each plot six distributions are compared, corresponding to the three pair of sub-samples obtained cutting on the not-plotted variables.

For some bins the low statistics don't allow to compare the distributions, but in the other bins and in the merged ones the good compatibility is probed.

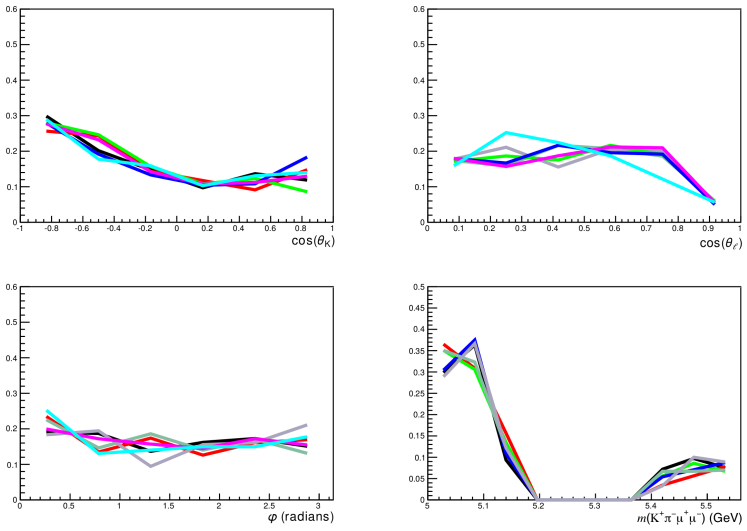


Figure 4.4: Distributions of the sideband events as a function of $\cos\theta_K$ (top left), $\cos\theta_L$ (top right), ϕ (bottom left), $m_{K\pi\mu\mu}$ (bottom right), in the q^2 bin range [0,3]. As described in Section 4.2.2, each distribution corresponds to a different sub-sample of the sideband events: $\cos\theta_K$ lower half (dim green) and upper half (grey), $\cos\theta_L$ lower half (black) and upper half (red), ϕ lower half (bright green) and upper half (blue), and low and high mass sidebands (magenta and light blue, respectively).

4.3 The fitting sequence, components and strategy

Before applying the fit procedure on data, some parameters of the `pdf` in Equation 4.1 are estimated, for each q^2 bin, on the simulated signal MC

sample and kept fixed for the full fitting process. These parameters are the mistag fraction f^M, as described in Section 4.2.1, and the seven parameters of the signal mass component of the `pdf`: the two widths and the relative abundance for the double Gaussian describing the right-tagged events, the same for the double Gaussian describing the mis-tagged events, and the common mean. To take into account the effect of any difference between simulated and real events, a specific systematic uncertainty have been computed and will be described in Section 7.

The angular component of the signal `pdf`, as described by Equation (1.25), depends on six parameters, F_{L}, F_{S}, A_{S}, P_1, P_5', and A_{S}^5. In order to facilitate the convergence of the fit process, and avoid problems related to the limited number of events and the presence of the physical boundary in the parameter phase space, the angular parameters that have already been measured by the previous CMS analysis on the same dataset, F_{L}, F_{S}, and A_{S}, have been fixed to the results of that measurement. To take into account the effect of fixing these parameters, a systematic uncertainty has been computed and will be described in Section 7.

The fit is performed in two steps. In the first one, the sidebands events are fitted, using only the background component of the `pdf`, to obtain the parameters of the $B^m(m)$, $B^{\theta_K}(\theta_K)$, $B^{\theta_l}(\theta_l)$, and $B^{\phi}(\phi)$ distributions. These parameters are then kept fixed in the second step of the fit. To correctly propagate the uncertainties on the background parameters to the analysis results, a specific systematic uncertainty has been computed and will be described in Section 7.

In the second step, the full set of events is fitted using the whole `pdf`. The free parameters in this fit are the angular parameters P_1, P_5', and A_{S}^5, and the yields Y_S^C and Y_B.

The separation of the fit in these two steps is needed to allow a stable fit to the signal region. As described in Section 4.2.2 the angular shape of the background events contains a high number of parameters and evaluating them together with the signal components would lead to a very unstable fit. Furthermore, using this two-step fit allows to keep the background parameters fixed for any fit procedure applied on data; this has the benefit of saving a huge amount of computing time, especially when large series of fits are run for systematic uncertainty estimations or for finding the best-fit values of the parameters, as described in the following paragraph.

The second step of the fit cannot be run at once, since the presence of a physical boundary for the validity of the fitted parameters complicates the process of numerical maximisation of the likelihood performed by MINUIT [33]. Especially in the q^2 bins in which the likelihood maximum is close to this boundary, the maximisation results tend to be unstable and strongly dependent on the values of the parameters at the begin of the fit. A strategy was then developed to avoid the effect of the physical boundary: the bi-dimensional space P_1 – P_5' is discretised by building a 90×90 rect-

angular grid, and for each of its points the values of P_1, P_5' are fixed in the fit, and the likelihood is maximised as a function of the nuisance parameters Y_S^C, Y_B, and A_S^5. Once the likelihood has been minimised for each point of the grid, it is fit with a bi-variate Normal distribution. The position of the maximum of this function, limited to the physical region, corresponds to the best estimate of the angular parameters P_1, P_5'. To avoid that any eventual non-Gaussian behaviour of the likelihood distribution in regions far from the maximum could introduce a bias in the results, the fit with the bi-variate Normal distribution is limited to the grid points (P_1^i, $P_5'^i$) for which is valid the following request:

$$\log \mathcal{L}(P_1^i, P_5'^i) > \log \mathcal{L}(P_1^{\max}, P_5'^{\max}) - 0.5$$

where ($P_1^{\max}$, $P_5'^{\max}$) is the grid point for which the likelihood is maximum. In this way the grid points fitted are limited to a region around the maximum position. The dependence of the fit result as a function of the region width, has been tested and found to be negligible, as will be discussed in Section 7.11.

Chapter 5

Efficiency

The efficiency for signal and for control channels is defined as the ratio of number of events passing the selection and whose selected candidate is matched with the generated final state, over the total number of events generated. It includes the effects of detector geometric acceptance, the trigger selection efficiency and the offline selection efficiency, and is entirely computed from MC simulation. The efficiency is built as a three-dimensional function of the angular observables θ_l, θ_K, and ϕ, and computed independently for each bin of q^2, both for signal and for the control regions. The use of such a function allows to account for any possible correlation among the variables introduced by the selection cuts.

In the CMS official MC samples, not all the generated events are reconstructed, to save computing resources. Before the simulation of the detector response, some basic cuts on the generated kinematic variables, p_T and η, of the signal final state muon pair are applied to remove the majority of events for which the final state is not completely in the geometric detector acceptance. In this thesis I will refer to this cuts as GEN-filter.

Only the events passing the GEN-filter are then reconstructed and compose the MC samples. To correctly take this into account, the efficiency is split into two different terms:

$$\epsilon^{R/M}(q^2, \theta_L, \theta_K, \phi) = \mathcal{A}(q^2, \theta_L, \theta_K, \phi) \times \epsilon_{reco}(q^2, \theta_L, \theta_K, \phi) = \frac{N_{gen}}{D_{gen}} \times \frac{N_{reco}}{D_{reco}} \tag{5.1}$$

where $\mathcal{A}$ is called acceptance and is the fraction of generated events that pass the GEN-filter, and ϵ_{reco} is the selection efficiency, namely the fraction of events which pass the selection with respect to those which have been reconstructed.

More precisely, the acceptance terms are defined as:

D_{gen} is the number of generated events which pass a selection $p_T(\mathrm{B}^0) > 8$ GeV and $|\eta(\mathrm{B}^0)| < 2.2$, in a given bin of q^2;

N_{gen} is the number of generated events with pass the above selection and, in addition, requires both muons to have $p_T(\mu_{GEN}) > 3.3\,\text{GeV}$ and $|\eta(\mu_{GEN})| < 2.3$;

This quantity is computed using only generator level quantities, as a function of generator-level angular observables.

The selection efficiency instead, is defined as:

$$\epsilon_{reco}(q^2, \theta_L, \theta_K, \phi) = \frac{N_{reco}}{D_{reco}} \tag{5.2}$$

where

D_{reco} is the number of events in the MC samples that pass the GEN-filter and on which the reconstruction is run;

N_{reco} is the number of reconstructed events passing all selection cuts defined in Sec. 3;

The D_{reco} is computed as a function of the generator level observables, while N_{reco} is computed as a function of the reconstructed quantities. The rationale for this choice is that the effect of the detector resolution, passing from the generated observables to the reconstructed one, is taken into account directly into the efficiency, without the need to add an additional term in the fitting function.

Since we cannot assume that the efficiency function is the same for events where the correct flavour has been identified and events where it is not, it is computed separately for candidates with correct tag (ϵ^R) and wrong tag (ϵ^M). The classification of an event as wrong or correct tagged is defined by comparing the MC truth with the result of the tagging algorithm described in Sec. 3.

Since the θ_K and θ_l variables are defined in the range $[0, \pi]$, there is no loss in information in building the efficiency as a function of θ_K, θ_l and ϕ, or as a function of $\cos\theta_K$, $\cos\theta_l$ and ϕ. Since this second choice leads to slightly more smooth functions, it has been chosen for the construction of the efficiency.

5.1 Parameterisation

In the first angular analysis based on this dataset, the efficiency was built as a two-dimensional function against the angular variables θ_K and θ_l. The method used to parameterise it was composed by two steps. Firstly, a two-dimensional function is constructed by performing the bin-by-bin ratios of the binned distributions of numerators and denominators. Then, this binned efficiency is fitted with a polynomial function, which is used as parameterisation of the efficiency. Some complex procedures were needed in order to

have this fit converging and to grant that the final efficiency functions were positive in the whole range of definition.

Most of the problems in this technique were due to both the low MC event statistics in some q^2 bins, especially for the mis-tagged event efficiency, and the large number of free parameters in the two-dimensional polynomial function. Extending this procedure to a three dimensional efficiency would imply the usage of three-dimensional binned distributions, with a global decrease of the bin statistics, and the usage of a three-dimensional polynomial function, with a larger number of free parameters to fit. For this reason, the usage of this method was not considered for this analysis, but new parameterisation techniques were tested.

Two independent approaches have been tested: a two dimensional reduction of the binned method and a parameterisation based on Kernel Density Estimator (KDE) distributions. In this section both of them are described and some example functions are showed. In the next section a closure test developed to determine their accuracy is described and the better performances of the KDE-based method will be highlighted.

5.1.1 Two-dimensional binned method

The first method is based on binned efficiency functions. Instead of building directly a three-dimensional binned efficiency, facing the low per-bin statistics, we use two-dimensional functions to compose the final efficiency. To model the efficiency correlation between each couple of parameters, three functions are created by integrating out, recursively, one angular variable and building the efficiency as a function of the remaining two. The final three dimensional efficiency is then defined as the product of the three two dimensional ones. In formula:

$$\epsilon(\cos\theta_l, \cos\theta_K, \phi) = F \times \epsilon(\cos\theta_l, \cos\theta_K) \times \epsilon(\cos\theta_l, \phi) \times \epsilon(\cos\theta_K, \phi) \quad (5.3)$$

Since the product of the three functions results not correctly normalised, it is rescaled by a factor F, in such a way that the global efficiency, averaged over all the angular variables, matches the simple ratio of the number of events in the numerator and denominator distributions.

This approach has the advantage to reduce the three-dimensional problem to three two-dimensional ones, easier to deal with and less prone to lack of statistics. The correlation among the three angular variables are taken into account by the final product. A disadvantage is that the efficiency is computed from binned distributions, where the bin width is determined by the statistics of events available in the MC samples used. Any structure or behaviour inside a given bin is not resolved, and some kind of smoothing is needed.

The solution used to mitigate this binning problem is to use an interpolation between the bins of each two-dimensional efficiencies of Equation 5.3.

Both a linear and parabolic interpolations have been tested.

An example of the two dimension binned efficiencies after a linear interpolation is applied, for the q^2 bin 1, are shown in Fig. 5.1.

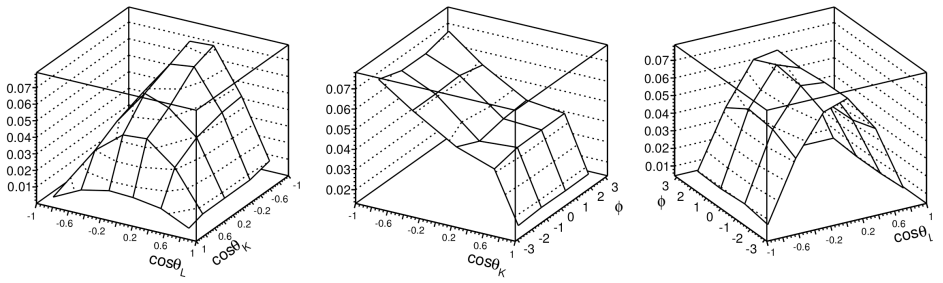


Figure 5.1: Distribution of the two-dimensional binned efficiency of equation 5.3 after the application of a linear interpolation between the bins, for q^2 bin 1: $\epsilon(\cos\theta_l, \cos\theta_K)$ (left), $\epsilon(\cos\theta_l, \phi)$ (centre), and $\epsilon(\cos\theta_K, \phi)$ (right).

5.1.2 Kernel Density Estimator method

The second approach uses a non-parametric description of the efficiency based on a Kernel Density Estimator (KDE) [32, 41].

The general idea behind this method, in its simplest form for a unidimensional problem, is to start from an unbinned distribution of events as a function of a given variable x, and describe its true distribution $\texttt{pdf}_{\rm TRUE}$ by building a kernel, namely a function $K(x)$ with unitary integral over the range of definition of x, on top of every event, and use the sum of all kernels, with proper normalisation, as a non-parametric description of $\texttt{pdf}_{\rm TRUE}$:

$$\texttt{pdf}_{\rm KDE}(x) = \frac{1}{N}\sum_{i=1}^{N} K(x - x_i) \tag{5.4}$$

where i is an index running over the set of events and N is the total number of events. The kernel $K(x)$ can be any non-negative function with unitary integral, and can be of different type; some common examples are uniform, triangular, or Gaussian kernels. A remarkable property of the KDE technique is that the quality of the PDF description depends rather weakly on the kernel used. The resulting $\texttt{pdf}_{\rm KDE}(x)$ in the limit $N \to \infty$ is a convolution of $\texttt{pdf}_{\rm TRUE}$ with the kernel $K(x)$. In this analysis, we decided to use a Gaussian kernel, because of its ability to produce a smooth `pdf` even in regions where a low event statistics is available.

The KDE method can be extended to multi-dimensional problems, building a `pdf` as a function of a set of n variables $x_1, \ldots, x_n$). In this case, the kernel used should be n-dimensional as well.

An implementation of KDE method using a Gaussian kernel is available, for both the uni-dimensional and the multi-dimensional application, within the RooFit package [45].

The advantage of the KDE over a parametric description is that no prior assumption on the actual $\texttt{pdf}_{\text{TRUE}}$ is needed, and the structure of the original `pdf` is described as accurately as the statistics allow. The last point is particularly relevant in the case under study, since in spite of the large statistics of events in the MC samples used to compute the efficiency, the phase space is very large, given the number of variables considered for the efficiency description. A disadvantage is that an unbinned distribution of events is needed, so it is not possible to use the KDE method directly on the efficiency, which is binned by construction, but it needs to be used on the numerator and denominator distributions, and then the efficiency will be defined through the ratios of these `pdf`s.

The efficiency at the observable boundaries is not expected to be null, especially for ϕ, given its periodic nature. In such cases, the kernel estimator would introduce a significant decreasing of the `pdf`, since it would behave, close to the border, like a convolution between a step function and a Gaussian function. In order to improve the behaviour of the modelling close the observable boundaries, data are mirrored across the boundaries. It means that, for any Gaussian function added for a data point next to a boundary, the tail of this function that exceeds it is reflected inside the boundary.

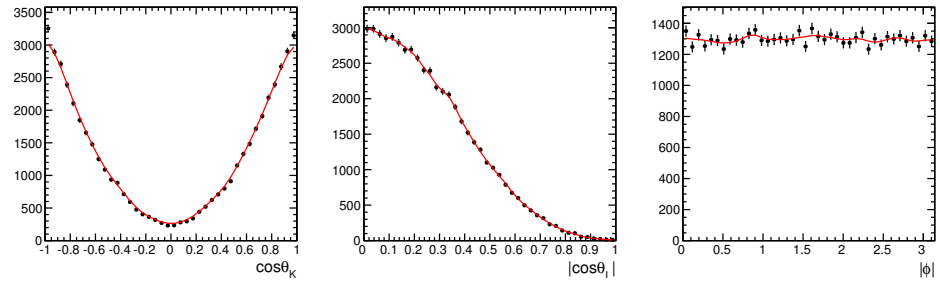
Even if the KDE with Gaussian kernel is a non-parametric method, the values of the widths and correlation terms of the multi-dimensional kernel need to be determined. For simplicity, the correlation terms are kept equal to zero; the usage of these terms could be useful when the distribution to model shows a strong correlation between two or more of its variables, but it is not the case for these numerator and denominator distributions. The RooFit class used to build the $\texttt{pdf}_{\text{KDE}}$ uses a set of standard values for the width parameters, determined as a function of the variables range of definition. An allowed tuning is the definition of an global scale factor, applied to the whole set of width parameters. On one hand, a wide kernel is less sensitive to the limited statistic of the data sample and allows to produce smoother `pdf`s. On the other hand, a narrow kernel reproduces in a more accurate way the fine structures of the original distribution, especially when this distribution is steep. Several width scale factors (1.0, 0.5, 0.3) have been tested and the choice of the value used for the efficiency is based on the goodness of the results of the closure test described in Section 5.2.

The procedure used to obtain the efficiency function is to use the KDE method on each of the four distributions $(N_{gen}, D_{gen}, N_{reco}, D_{reco})$ used in Equation 5.1, normalise each of the resulting `pdf`s to the number of events in the original distributions, and then combine the four functions into final efficiency. For technical reason, it is not possible, in ROOT, to save the output of a multidimensional KDE `pdf` into an output file and the time

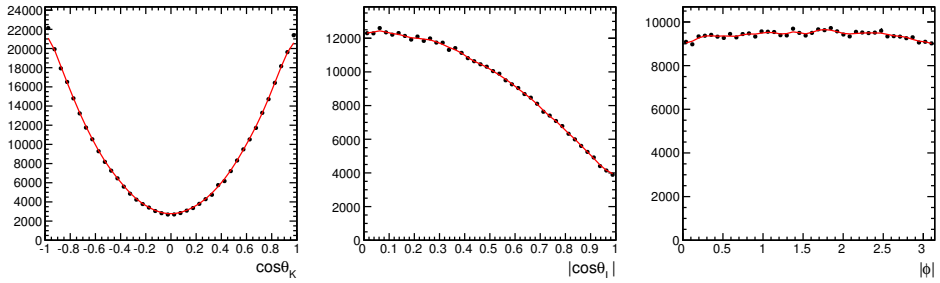
consumption to create the efficiency function, starting from the MC event distributions, any time the fit procedure is run is too large. For this reason, it was decided to save a binned distribution obtained from the sampling of the KDE `pdfs`: numerator and denominator functions, obtained via the KDE algorithm, have been sampled into a three dimensional histogram, with $40 \times 40 \times 40$ bins, then the four histograms were combined with a bin-per-bin application of Equation 5.1 in an efficiency function, which has been finally saved in the output file. The granularity of the bins have been chosen as a compromise between the fine structure of the efficiency distribution and computing time needed to perform the `pdf` sampling.

An example of the event distribution of the four terms of Equation 5.1, D_{gen}, N_{gen}, D_{reco}, and N_{reco}, for right-tagged events in q^2 bin 1, as functions of each of the angular variables, together with the projections of the correctly-normalised `pdf`s, as obtained by the KDE method with width scale factor equal to 0.5, are shown in Fig. 5.2.

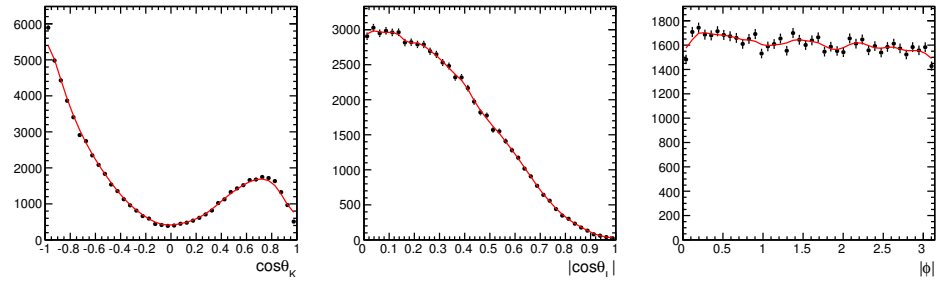
An example of the projections of the efficiency, as obtained by the KDE method with width scale factor equal to 0.5, on each single and on each couple of angular variables, for right-tagged events in q^2 bin 1, are shown in Fig. 5.3.



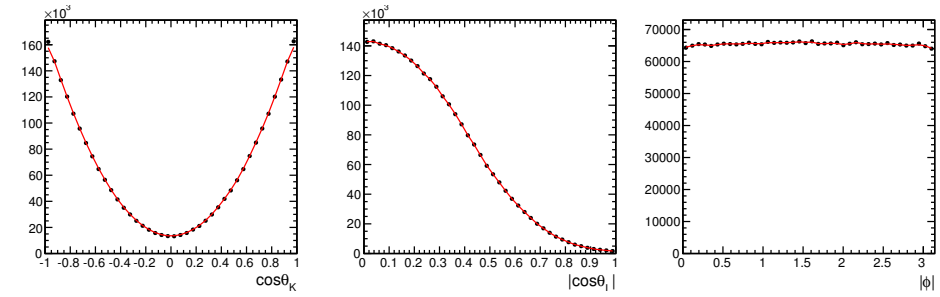
(a) N_{gen}



(b) D_{gen}



(c) N_{reco}



(d) D_{reco}

Figure 5.2: Event distributions of the four efficiency terms of Equation 5.1, D_{gen}, N_{gen}, D_{reco}, and N_{reco} (black dots), for right-tagged events in q^2 bin 1, together with the projections of the correctly-normalised `pdf`s (red line), as obtained by the KDE method with width scale factor equal to 0.5.

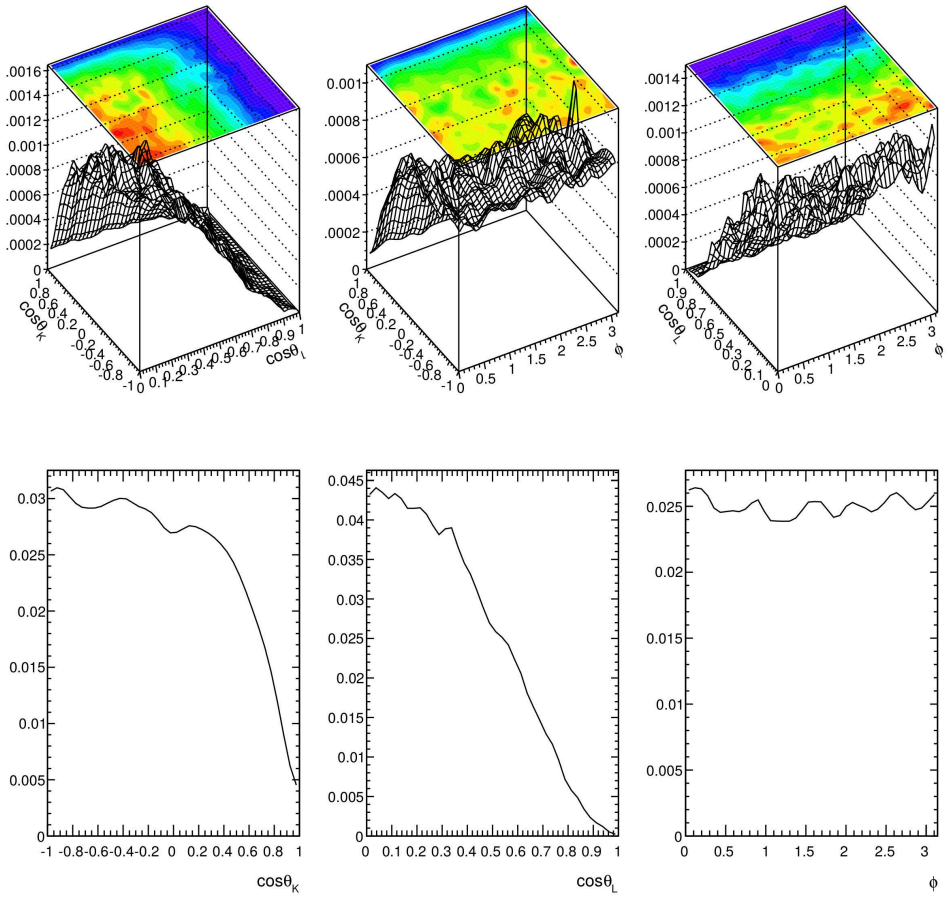


Figure 5.3: Projections of the efficiency, as obtained by the KDE method with width scale factor equal to 0.5, on two (upper row) and one angular variable (lower row), for right-tagged events in q^2 bin 1

5.2 Closure test

A validation of the efficiency parameterisation is performed via a closure test based on MC samples. We compare the event distribution of the three angular observables, $\cos\theta_K$, $\cos\theta_l$, and ϕ, as reconstructed after the full simulation and after the application of the whole set of selections described in Section 3, with the full set of generated event distribution, before the application of the GEN-filter, weighted, event-by-event, with the tested efficiency function. In order to remove any statistical correlation in the closure test, half of each sample have been used to estimated the efficiency, and the closure has been performed on the other half.

The equality to test is the following:

$$\texttt{pdf}(q^2, \cos\theta_K, \cos\theta_l, \phi)^{R/M}_{reco} = \texttt{pdf}(q^2, \cos\theta_K, \cos\theta_l, \phi)_{gen} \otimes \epsilon^{R/M}(q^2, \cos\theta_K, \cos\theta_l, \phi) \tag{5.5}$$

The closure test has been performed for any efficiency function produced: for each q^2 bin, separately for right (R) and mis-tag (M), and for signal and control samples.

An example of the closure test results of the first method efficiencies, as described in Section 5.1.1, is shown, for right-tagged events in q^2 bin 1, for the linear interpolation in Fig. 5.4, and for the parabolic interpolation in Fig. 5.5. The result quality for the linearly-interpolated efficiency does not follow very well the original distribution where it is very steep and it introduces a ϕ dependence where none is present. The quality improvement due to the second-order interpolation is not enough to have an accurate description of the shape in the $\cos\theta_K$ projection and the modulation against ϕ is still present. An improvement to this technique could come from an increase in the number of bins used to build the two-dimensional components of the efficiency, but the available statistics of the MC samples, especially for the mis-tagged events, prevent to go further with this.

An example of the closure test results for the KDE method is shown in Fig. 5.6 for right-tagged events in q^2 bin 0, and in Fig. 5.7 for right-tagged events in q^2 bin 7, comparing three different width scale factors: 1.0, 0.5, and 0.3. It is possible to notice that using smaller widths improves the result quality where some steep features are present, in particular near the boundary of $\cos\theta_K$ for q^2 bin 0, and of ϕ for q^2 bin 7. On the other hand, a small width tends to create fine structures due to statistic fluctuations, in particular near the boundary of ϕ for q^2 bin 0.

From the overall results of the closure test, the optimal parameterisation for this analysis in the KDE method. The width scale factor is chosen to be 0.5 for the efficiency for right-tagged events, while the scale factor of 1.0 was found to be optimal for the efficiency of mis-tagged events, for which less statistic is available in the MC samples. An example of the closure test results for mis-tagged events is shown in Fig. 5.8, for q^2 bin 0 and q^2 bin 7.

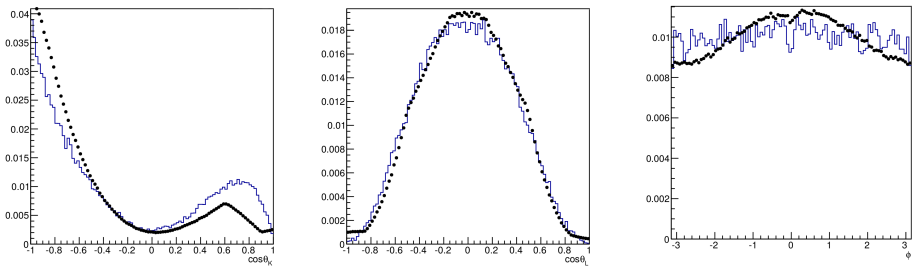


Figure 5.4: Comparison of the reconstructed and selected MC event distributions (histogram) with generated event distributions (dots) weighted with the efficiency, parameterised as in Equation 5.3 with linear interpolation of the two-dimension efficiency functions. Both the distributions and the efficiency refer to right-tagged events in q^2 bin 1.

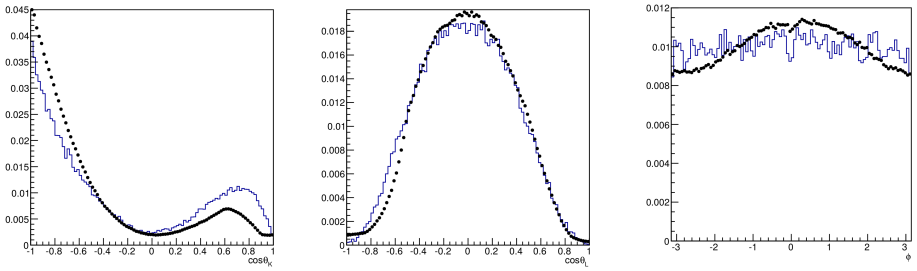


Figure 5.5: Comparison of the reconstructed and selected MC event distributions (histogram) with generated event distributions (dots) weighted with the efficiency, parameterised as in Equation 5.3 with parabolic interpolation of the two-dimension efficiency functions. Both the distributions and the efficiency refer to right-tagged events in q^2 bin 1.

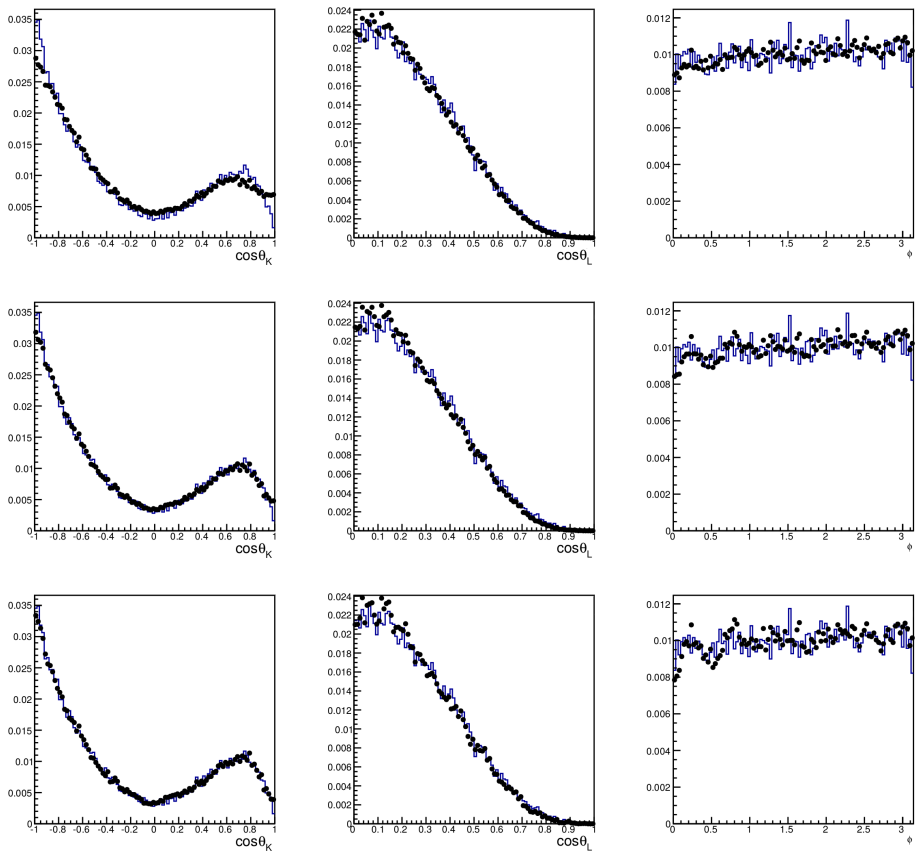


Figure 5.6: Comparison of the reconstructed and selected MC event distributions (histogram) with generated event distributions (dots) weighted with the efficiency, built through the KDE method, with width scale factor set to 1.0 (top), 0.5 (middle), and 0.3 (bottom). Both the distributions and the efficiency refer to right-tagged events in q^2 bin 0.

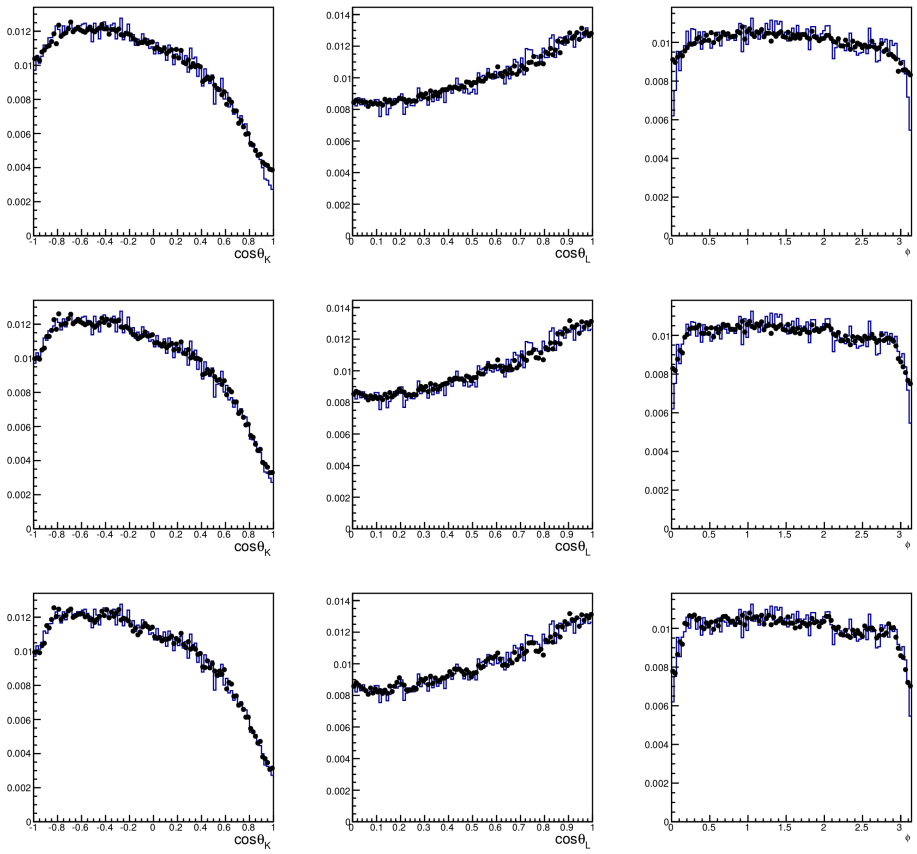


Figure 5.7: Comparison of the reconstructed and selected MC event distributions (histogram) with generated event distributions (dots) weighted with the efficiency, built through the KDE method, with width scale factor set to 1.0 (top), 0.5 (middle), and 0.3 (bottom). Both the distributions and the efficiency refer to right-tagged events in q^2 bin 7.

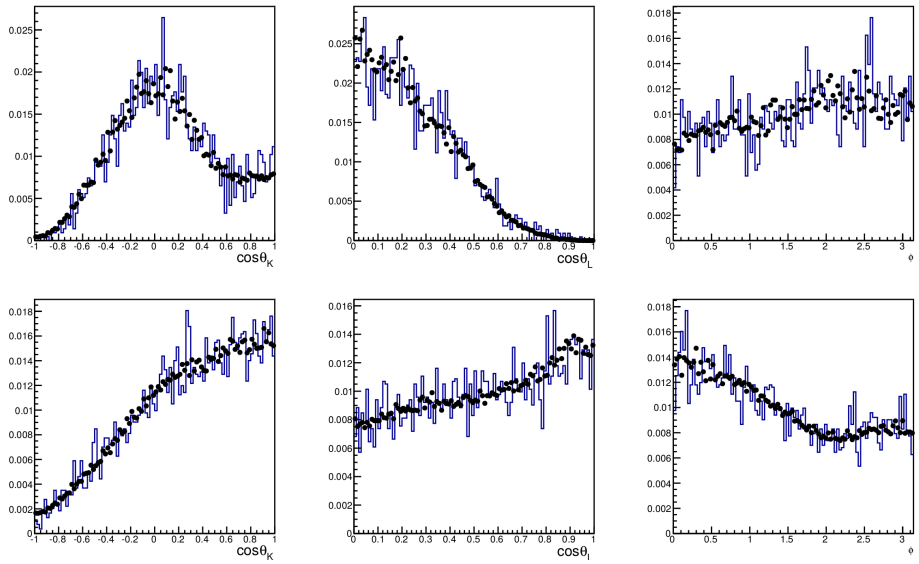


Figure 5.8: Comparison of the reconstructed and selected MC event distributions (histogram) with generated event distributions (dots) weighted with the efficiency, built through the KDE method. Both the distributions and the efficiency refer to mis-tagged events in q^2 bin 0 (top) and in q^2 bin 0 (bottom).

Chapter 6

Validation of the fit algorithm

Since the whole fit procedure is very complex, many validation checks are performed to verify the robustness of the final result.

In this section I will present the studies performed on the fit procedure, both on simulated MC events and on data control channels. Since all the signal MC samples contain only events with resonant P-wave Kπ state, it will be implicit in this section that the angular decay rate used when fitting MC events does not have the S-wave and interference terms, while the full decay rate is used when fitting on data control channels.

6.1 Generator-level fit to simulated events

The first validation step, presenting the minimum level of complexity, is performed by fitting the generator-level angular distributions of the signal MC sample, using the pure decay rate as `pdf`: $S^a(\theta_K, \theta_l, \phi)$.

The good status of the fit is considered a successful probe of the the minimisation process and of the correct description of the generated angular distribution through the decay rate. The angular folding operations are applied to distributions and `pdf`s, but this has no effect on the result, since the terms of the `pdf` are either odd or even with respect to them.

Since the "true" values of the angular parameters used to simulate the events of the MC samples are not defined, the results of this fit will be used as a reference value for any comparison of the following results. This fit is clean from any effect due to finite experimental resolution, so I expect that the fit results do not have any bias.

An example of the generator-level angular distributions and the projections of the fitted decay rate, for q^2 bin 0, are shown in Fig. 6.1.

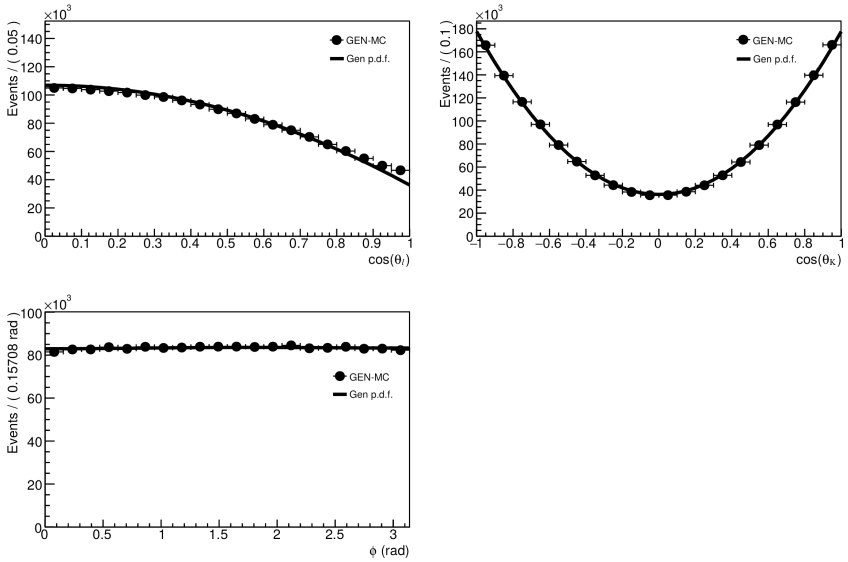


Figure 6.1: Generator-level variable distributions of the MC sample before the GEN-filter, together with the projections of the fitted `pdf`, as described in Section 6.1, for q^2 bin 0.

6.2 Reconstruction-level fit to simulated events

The second validation step is performed by fitting the reconstruction-level angular distributions of the signal MC sample, after applying the criteria of candidate selection and CP-state tag described in Section 3. This fit is performed using only the angular terms of the signal components in the `pdf` described in Equation 4.1.

The simulated sample contains both right-tagged and mis-tagged events, which can be distinguished by using the MC truth information. As a first step, these two categories are fitted individually, using for each of them the corresponding angular component of the `pdf`, which is the decay rate times the efficiency function. The results of the fit show that the inclusion of the efficiency function in the fit procedure is correctly implemented.

An example of the reconstruction-level angular distributions of the signal MC sample and the projections of the angular component of the `pdf`, for q^2 bin 0, is shown in Figure 6.2 for right-tagged events and in Figure 6.3 for mis-tagged events.

As a second step, the full MC sample, containing both right-tagged and mis-tagged events, is used in the fit, and both the signal terms of the `pdf` are used. The mistag fraction parameter in the `pdf` is fixed to the values reported in Section 4.2.1. Since the mistag fraction values are computed on the same MC sample used for this fit, there is a statistical correlation

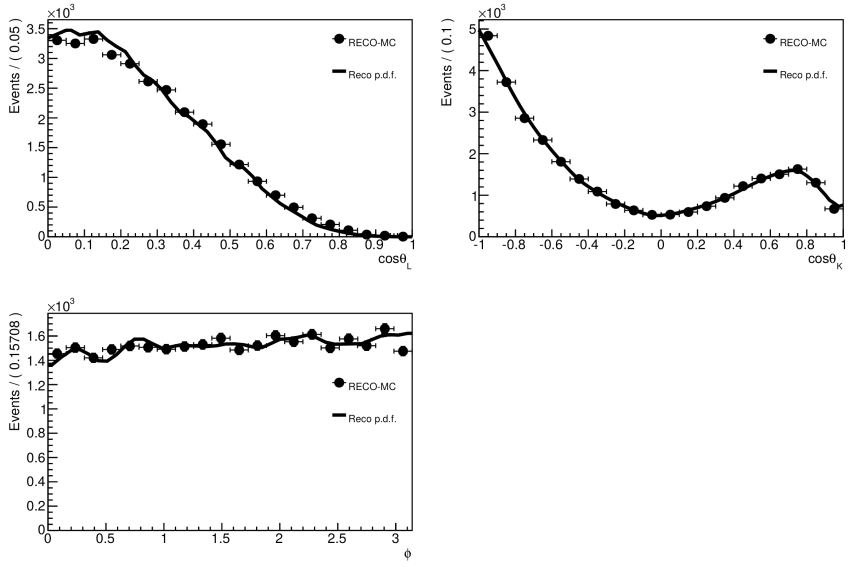


Figure 6.2: Angular variable distributions of the right-tagged MC events after the selection criteria application, together with the projections of the fitted `pdf`, as described in Section 6.2, for q^2 bin 0.

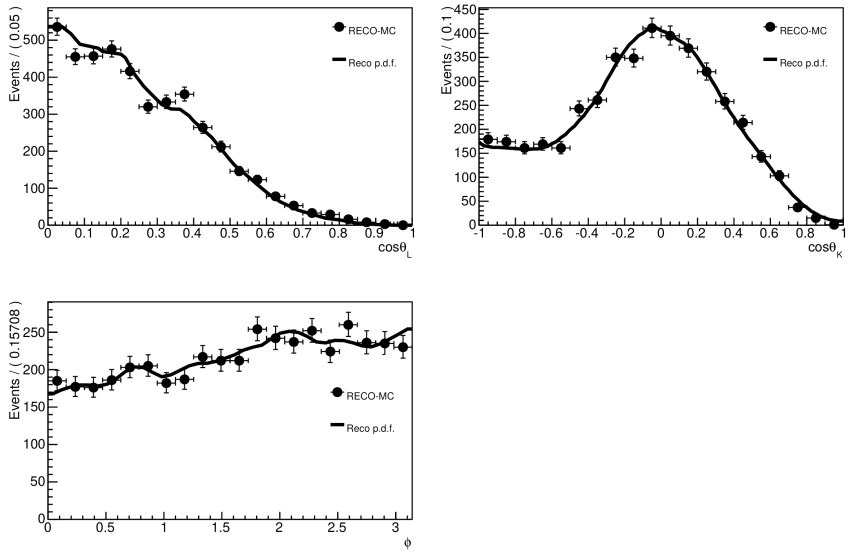


Figure 6.3: Angular variable distributions of the mis-tagged MC events after the selection criteria application, together with the projections of the fitted `pdf`, as described in Section 6.2, for q^2 bin 0.

between it and the fitted distributions; anyway its effects on the fit results are negligible because of the extremely small statistical uncertainty on this parameter.

The fit is still limited to the three angular variables, because all the parameters of the mass shapes and the relative abundances are fixed in the fit and there is no information to extract from the mass distributions.

An example of the reconstruction-level angular distributions of the signal MC sample and the projections of the angular component of the `pdf`, for both right-tagged and mis-tagged events in q^2 bin 0, is shown in Figure 6.4.

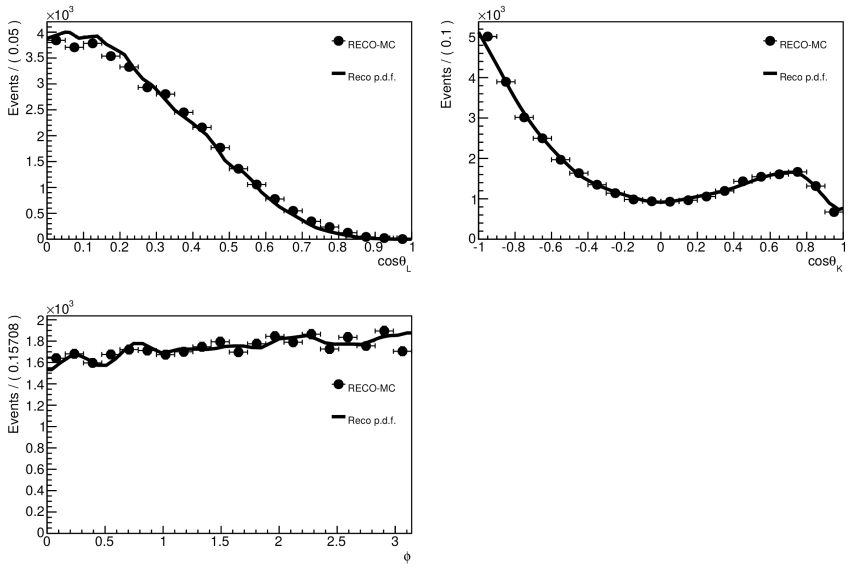


Figure 6.4: Angular variable distributions of all the MC events after the selection criteria application, together with the projections of the fitted `pdf`, as described in Section 6.2, for q^2 bin 0.

The results of the fits performed in this section are compared with the results of the generator-level fits.

The results of right-tagged event fit and of the generator-level fit, are shown in Figure 6.5 for the F_L parameter, in Figure 6.6 for the P_5' parameter, and in Figure 6.7 for the P_1 parameter. The results of mis-tagged event fit and of the generator-level fit, are shown in Figure 6.8 for the F_L parameter, in Figure 6.9 for the P_5' parameter, and in Figure 6.10 for the P_1 parameter. The results of right-tagged event fit and of the generator-level fit, are shown in Figure 6.11 for the F_L parameter, in Figure 6.12 for the P_5' parameter, and in Figure 6.13 for the P_1 parameter.

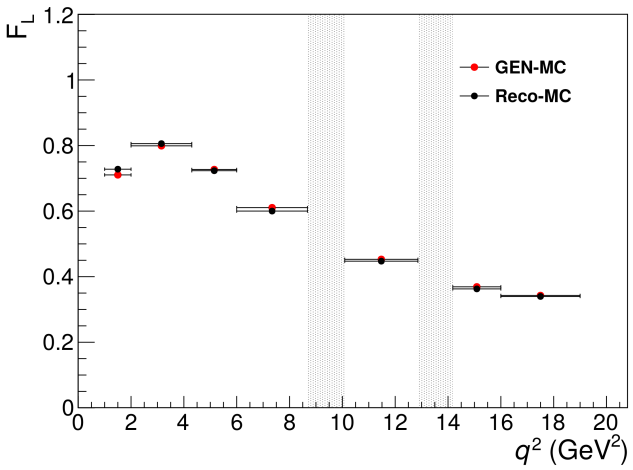


Figure 6.5: Results for the F_L parameter from the reconstruction-level fit to right-tagged event distributions (black) and from the generator-level fit (red), for each q^2 bin. The vertical shaded regions correspond to the q^2 bins dedicated to the J/ψ and ψ' control channels.

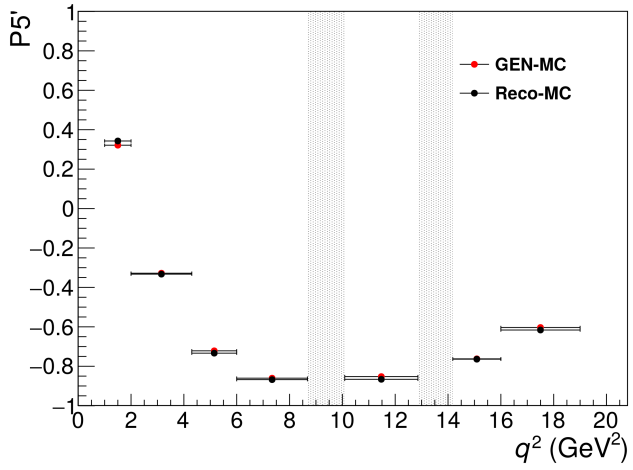


Figure 6.6: Results for the P_5' parameter from the reconstruction-level fit to right-tagged event distributions (black) and from the generator-level fit (red), for each q^2 bin. The vertical shaded regions correspond to the q^2 bins dedicated to the J/ψ and ψ' control channels.

6.3 Reconstruction-level fit to low statistics simulated samples

The fit algorithm is also validated with simulated samples having the same statistics of the real data sample, and containing both the signal and the

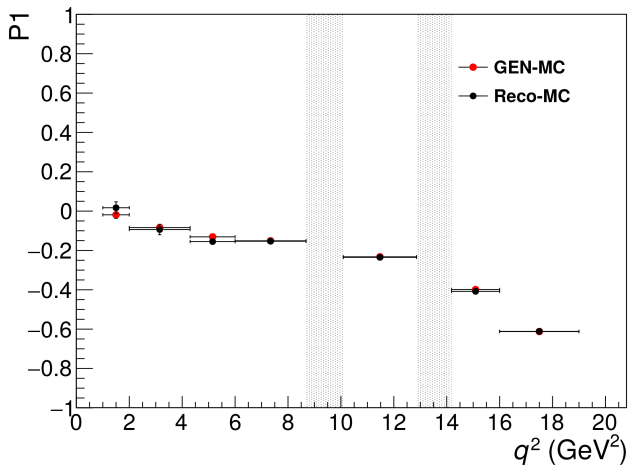


Figure 6.7: Results for the P_1 parameter from the reconstruction-level fit to right-tagged event distributions (black) and from the generator-level fit (red), for each q^2 bin. The vertical shaded regions correspond to the q^2 bins dedicated to the J/ψ and ψ' control channels.

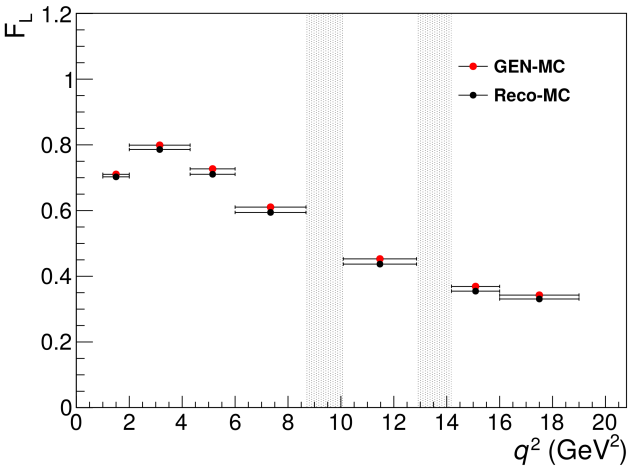


Figure 6.8: Results for the F_L parameter from the reconstruction-level fit to mis-tagged event distributions (black) and from the generator-level fit (red), for each q^2 bin. The vertical shaded regions correspond to the q^2 bins dedicated to the J/ψ and ψ' control channels.

background components. The goal is to verify whether the analysis is able to measure the interesting observables, in conditions as close as possible to the real data sample.

The data events are obtained by dividing the MC samples in sub-samples with a number of events exactly equal to the signal yield, as obtained from

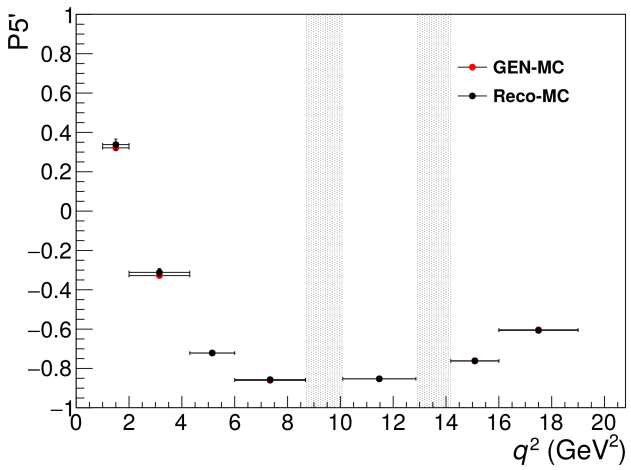


Figure 6.9: Results for the P_5' parameter from the reconstruction-level fit to mis-tagged event distributions (black) and from the generator-level fit (red), for each q^2 bin. The vertical shaded regions correspond to the q^2 bins dedicated to the J/ψ and ψ' control channels.

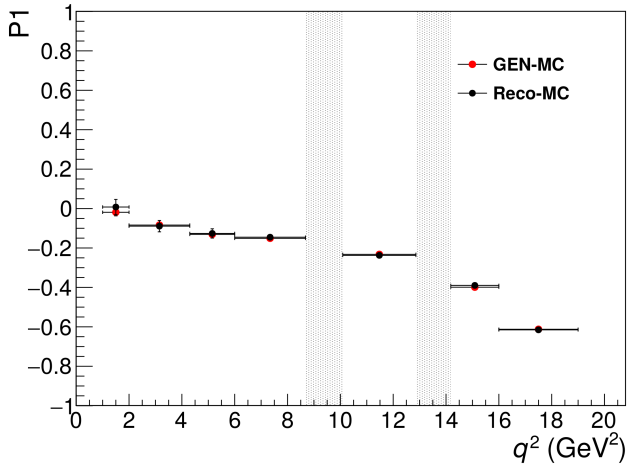


Figure 6.10: Results for the P_1 parameter from the reconstruction-level fit to mis-tagged event distributions (black) and from the generator-level fit (red), for each q^2 bin. The vertical shaded regions correspond to the q^2 bins dedicated to the J/ψ and ψ' control channels.

fitting on data the B^0 mass distribution. The number of sub-samples that we can produced is limited by the statistics available in the MC sample, and for simplicity has been rounded down to 200 sub-samples, for each q^2 bin.

The background distributions are generated with pseudo-experiments, using the `pdf` described in Section 4.2.2, and parameter values measured

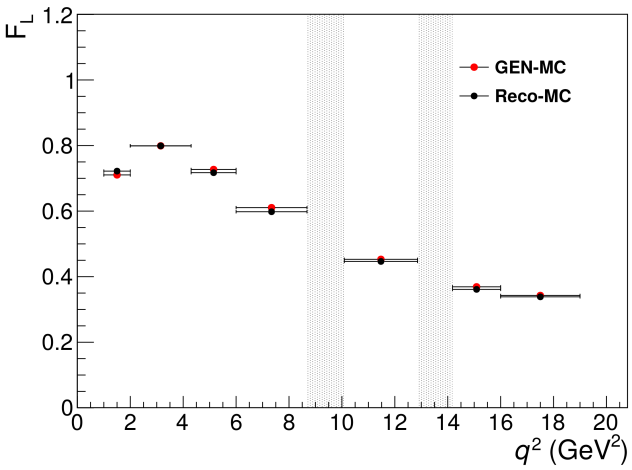


Figure 6.11: Results for the F_L parameter from the reconstruction-level fit (black) and from the generator-level fit (red), for each q^2 bin. The vertical shaded regions correspond to the q^2 bins dedicated to the J/ψ and ψ' control channels.

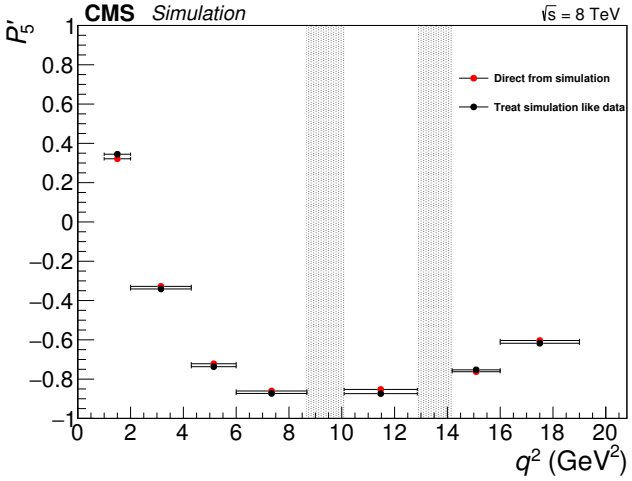


Figure 6.12: Results for the P_5' parameter from the reconstruction-level fit (black) and from the generator-level fit (red), for each q^2 bin. The vertical shaded regions correspond to the q^2 bins dedicated to the J/ψ and ψ' control channels.

with data sidebands. The number of events generated for each sub-sample are equal to the background yield, as obtained from fitting on data the B^0 mass distribution. To match the signal MC sub-samples, also for background a total of 200 sets of events have been generated, for each q^2 bin.

These samples have been used to validate the fitting procedure, first us-

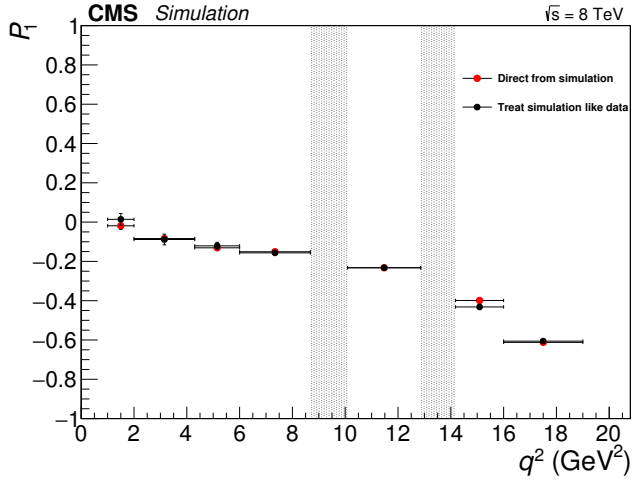


Figure 6.13: Results for the P_1 parameter from the reconstruction-level fit (black) and from the generator-level fit (red), for each q^2 bin. The vertical shaded regions correspond to the q^2 bins dedicated to the J/ψ and ψ' control channels.

ing only the signal, then merging each signal sub-sample with a background one. I will refer to a merged sub-sample as "cocktail" MC sample.

6.3.1 Data-like samples of signal MC events

As a first step, the angular distribution of the signal sub-samples is fitted, using only the signal components of the `pdf`. As for the fit to the full MC sample, including the mass in the fitted distributions would not change the result.

Each of the 200 MC sub-samples are fitted. An example of the distributions of the resulting parameters, for q^2 bin 3, is shown in Figure 6.14.

The distributions of these results are compared with the results of the fit to the full MC sample. Any significant difference can be an hint of biases on the results introduced by the fit procedure.

The mean value of the result distributions are shown in Figure 6.15, in Figure 6.16, and in Figure 6.17, for the F_L, P_1, and P_5' parameters, respectively. Only the results from converging fits are included in this distributions, so in general their number is lower than 200. The error bars assigned to the mean values are the standard deviations of the result distributions, divided by the square root of the number of results in them.

6.3.2 Data-like "cocktail" MC samples

As a second step, the mass and angular distributions of the 200 "cocktail" MC samples are fitted using the full `pdf` function. This fit is very similar

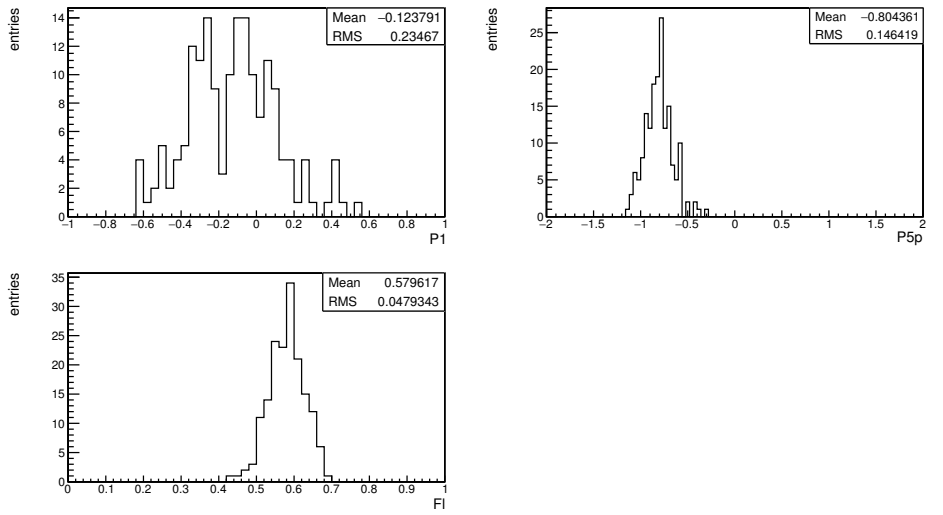


Figure 6.14: Distributions of the results of the fits to the 200 signal MC sub-samples, for q^2 bin 3.

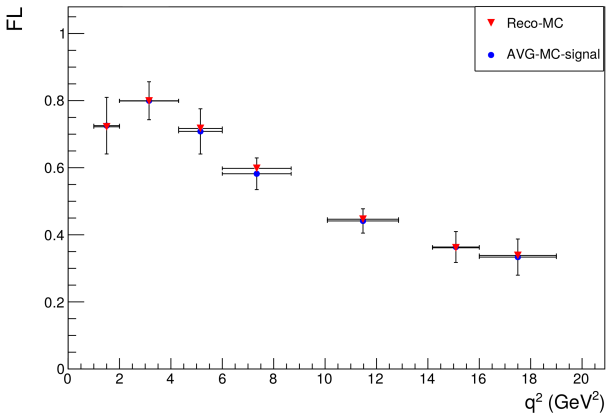


Figure 6.15: Average values of the F_L result distribution from the fit to 200 signal MC sub-samples (blue), together with the F_L results of the fit to the full MC sample (red). The error bars associated to the sub-sample fit results represent the statistical uncertainty associated to the arithmetic average of the results, as descibed in Section 6.3.1.

to the one performed on data, even if there is no S-wave component in the signal events. For this reason, this validation check is one of the most important steps of the procedure. As for the signal MC sub-samples, also here the mean and the standard deviation, divided by the square root of fits, are used to represent the "cocktail" fit results.

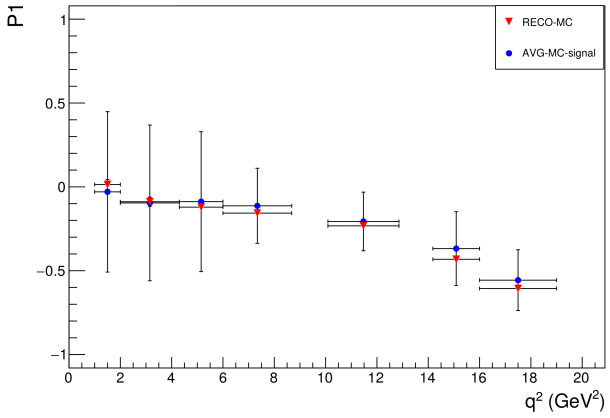


Figure 6.16: Average values of the P_1 result distribution from the fit to 200 signal MC sub-samples (blue), together with the P_1 results of the fit to the full MC sample (red). The error bars associated to the sub-sample fit results represent the statistical uncertainty associated to the arithmetic average of the results, as descibed in Section 6.3.1.

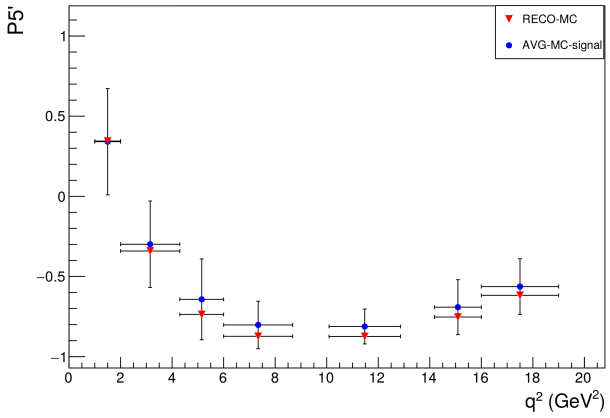


Figure 6.17: Average values of the P_5' result distribution from the fit to 200 signal MC sub-samples (blue), together with the P_5' results of the fit to the full MC sample (red). The error bars associated to the sub-sample fit results represent the statistical uncertainty associated to the arithmetic average of the results, as descibed in Section 6.3.1.

The comparison between the average results of the 200 "cocktails" fit and the fit results of Reco MC, is shown in Figure 6.18 and Figure 6.19, for the P_1 and P_5' parameters, respectively.

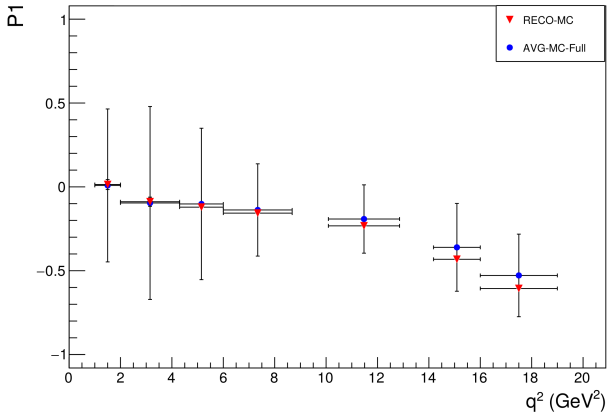


Figure 6.18: Average values of the P_1 result distribution from the fit to 200 "cocktail" samples (blue), together with the P_1 results of the fit to the full MC sample (red). The error bars associated to the cocktail-MC fit results represent the statistical uncertainty associated to the arithmetic average of the results, as descibed in Section 6.3.2.

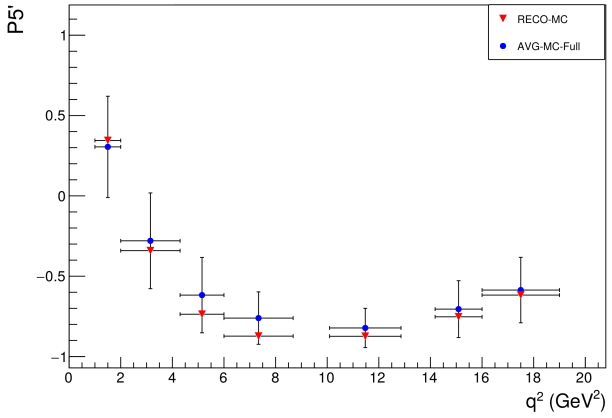


Figure 6.19: Average values of the P_5' result distribution from the fit to 200 "cocktail" samples (blue), together with the P_5' results of the fit to the full MC sample (red).The error bars associated to the cocktail-MC fit results represent the statistical uncertainty associated to the arithmetic average of the results, as descibed in Section 6.3.2.

6.4 Validation with data control channels

The analysis technique is validated with the data by means of the control channels. In this way, the S-wave component of the PDF is tested, and we have a check of the efficiency behaviour on real data.

6.4.1 Sideband fit

The background shape is determined as for the other bins, by using the data sidebands as a function of the angular observables. Each of the three angular observables is fit with a polynomial with different degrees, specified in Table 6.1. The background distributions and `pdf` projections are plotted in Figure 6.20 to Figure 6.22, for the $B^0 \to J/\psi K^{*0}$ control channel, and in Figure 6.23 to Figure 6.25, for the $B^0 \to \psi' K^{*0}$ control channel.

Table 6.1: Degrees of the polynomial functions used for control channel `pdf`s.

q^2 bin index	$B^{\cos\theta_K}$ degree	$B^{\cos\theta_l}$ degree	B^{ϕ} degree
$B^0 \to J/\psi K^{*0}$	4	4	5
$B^0 \to \psi' K^{*0}$	4	3	5

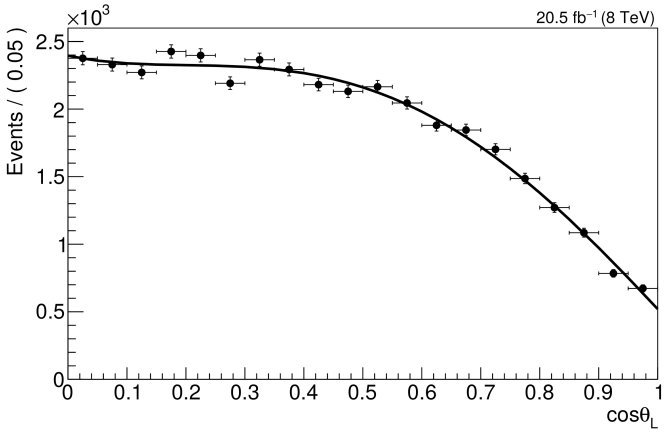


Figure 6.20: Distribution of the $\cos\theta_l$ variable in mass sidebands of the $B^0 \to J/\psi K^{*0}$ control channel and the background `pdf` projection.

In the sideband $\cos\theta_K$ distribution of the $B^0 \to J/\psi K^{*0}$ channel, it is possible to note a feature around the value of $\cos\theta_K = 0.7$ that cannot be described by the background `pdf` used. The origin of this shape has been attributed to a contamination from a source of non-combinatorial background in the sidebands, like partially reconstructed background events. Since in the sideband distributions of the signal q^2 bins the statistics is not sufficient to appreciate whether this feature is present or not, as can been noticed in the fits shown in Section /refsec:backg, a dedicated component in the `pdf` can not be used. Adding such a component only in the fit to the control regions would spoil the assumption that the control regions are treated in

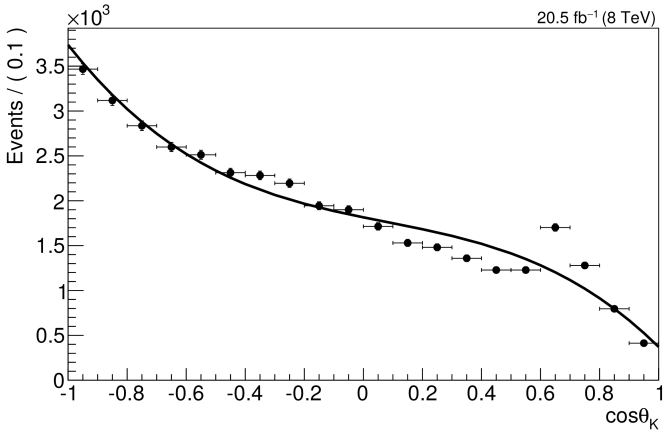


Figure 6.21: Distribution of the $\cos\theta_K$ variable in mass sidebands of the $B^0 \to J/\psi K^{*0}$ control channel and the background `pdf` projection.

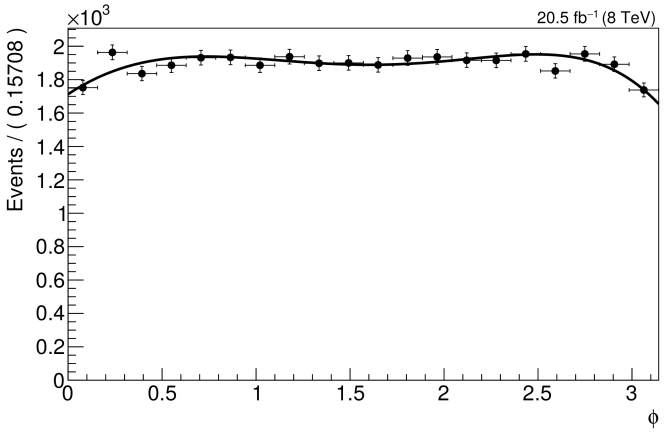


Figure 6.22: Distribution of the ϕ variable in mass sidebands of the $B^0 \to J/\psi K^{*0}$ control channel and the background `pdf` projection.

the same way as the signal regions and would make unreliable any study on sources of systematic uncertainty performed on them.

6.4.2 Fit to data control channel

The two control channels are fitted, with F_s and A_s fixed, as described in Section 4.3, while F_L is kept floating. The projection plots of the fit results are shown Figure 6.26 and Figure 6.27, for $B^0 \to J/\psi K^{*0}$ and $B^0 \to \psi' K^{*0}$ channels respectively.

The results of the measurements for the control channels $B^0 \to J/\psi K^{*0}$

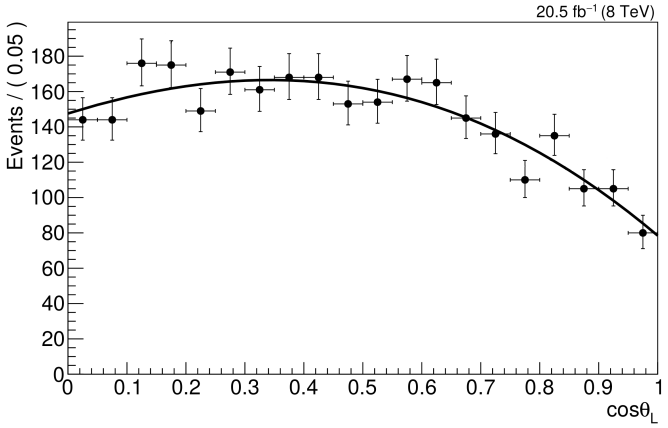


Figure 6.23: Distribution of the $\cos\theta_l$ variable in mass sidebands of the $B^0 \to \psi' K^{*0}$ control channel and the background `pdf` projection.

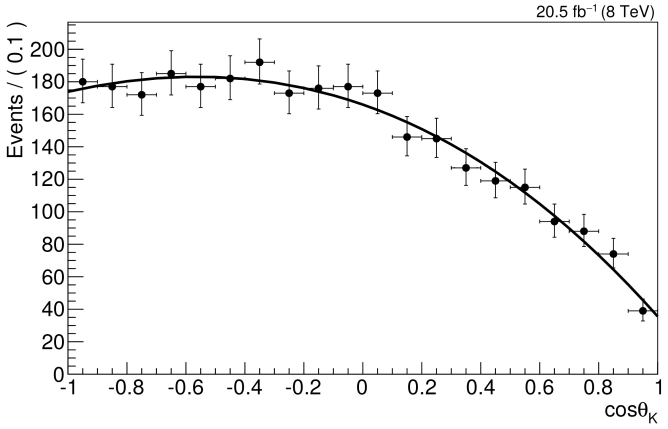


Figure 6.24: Distribution of the $\cos\theta_K$ variable in mass sidebands of the $B^0 \to \psi' K^{*0}$ control channel and the background `pdf` projection.

and $B^0 \to \psi' K^{*0}$ are summarised in Table 6.2.

Table 6.2: Results of the fit to the data control channels, as described in Section 6.4.2. The reported uncertainty is fully statistical.

control channel	F_L	P_1	P_5'	A_s^5
$B^0 \to J/\psi K^{*0}$	0.537 ± 0.002	-0.081 ± 0.011	-0.024 ± 0.007	-0.002 ± 0.002
$B^0 \to \psi' K^{*0}$	0.538 ± 0.008	-0.031 ± 0.001	-0.039 ± 0.001	0.005 ± 0.001

For both the control channels, the F_L parameter has been measured in

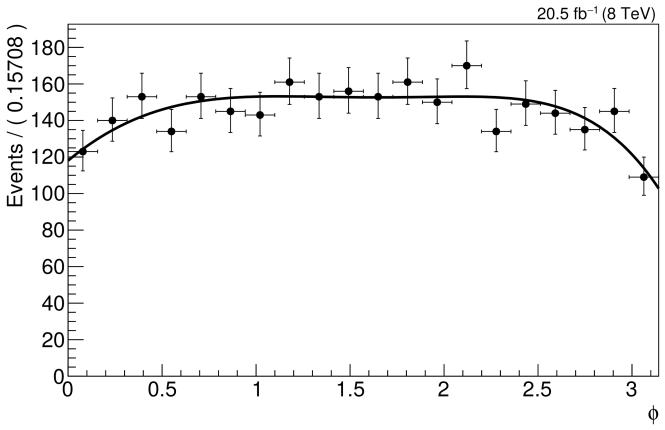


Figure 6.25: Distribution of the ϕ variable in mass sidebands of the $B^0 \to \psi' K^{*0}$ control channel and the background `pdf` projection.

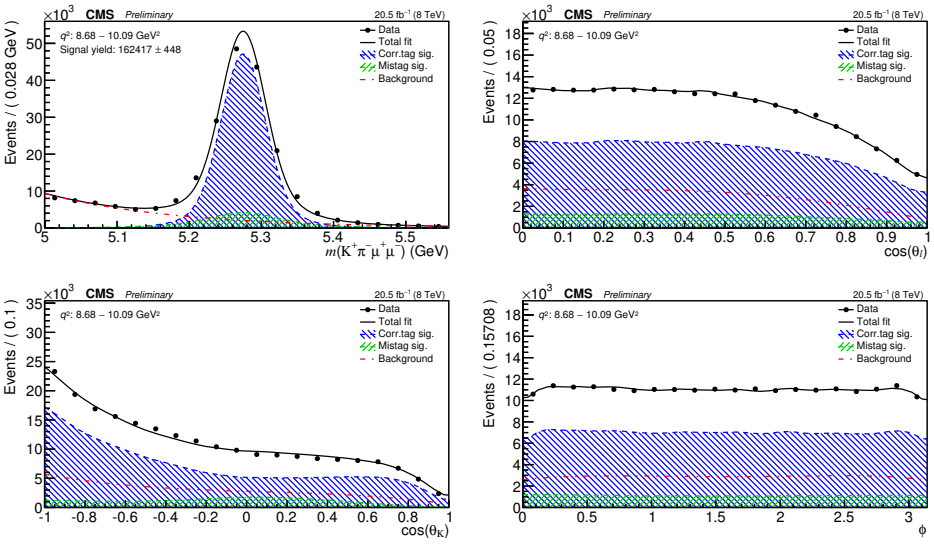


Figure 6.26: The fitting results of the control channel J/ψ on data. The plots show the projections of the fitting results on three different angular variables: B^0 mass, $\cos\theta_l$, $\cos\theta_K$ and ϕ.

the previous CMS analysis and also by other experiments. The results from this work, from the previous CMS analysis and from other experiments are in good agreement, as shown in Table 6.3.

As further test, the fit on the two control channels has been repeated with the F_L parameter fixed, coherently with the procedure used on data and described in Section 4.3. The results of the two kind of fits, with F_L

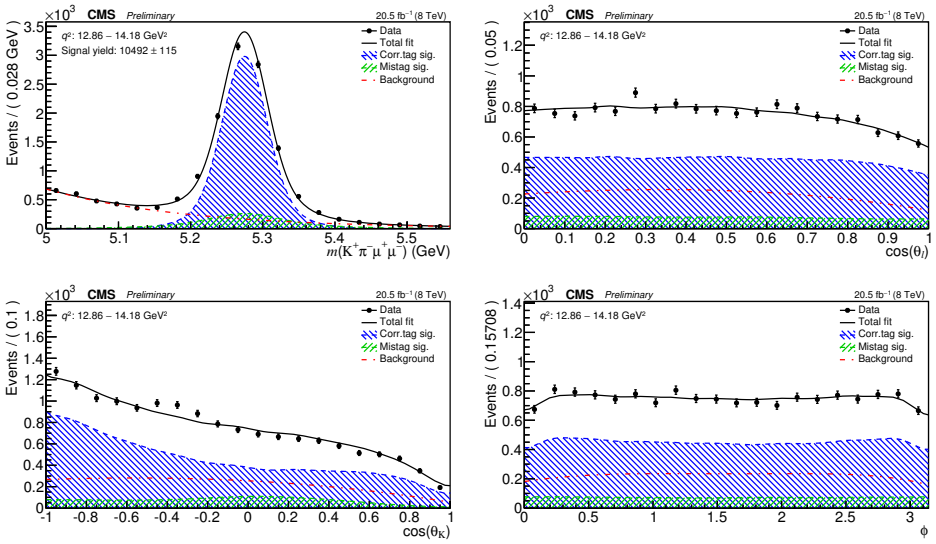


Figure 6.27: The fitting results of the control channel Ψ' on data. The plots show the projections of the fitting results on three different angular variables: B^0 mass, $\cos\theta_l$, $\cos\theta_K$ and ϕ.

Table 6.3: Measurements from CMS (both in this and in the previous analysis), PDG, and BaBar [15] of F_L in the control channels. The first uncertainty is statistical and the second is systematic.

control channel	$B^0 \to K^{*0}(K^+\pi^-)J/\psi(\mu^+\mu^-)$			$B^0 \to K^{*0}(K^+\pi^-)\psi'(\mu^+\mu^-)$		
Experiment	F_L	Err(stat)	Err(syst)	F_L	Err(stat)	Err(syst)
This work	0.537	±0.002	–	0.538	±0.008	–
CMS	0.537	±0.002	–	0.538	±0.008	–
PDG	0.571	±0.007	–	0.463	$^{+0.028}_{-0.040}$	–
BaBar	0.556	±0.009	±0.010	0.48	±0.005	±0.002

either free to float or fixed, are compared in Table 6.4.

Table 6.4: Results of the fit to the data control channels, as described in Section 6.4.2. The reported uncertainty is fully statistical.

control channel	F_L state	P_1	P_5'	A_s^5
$\mathrm{B}^0 \to \mathrm{J}/\psi\mathrm{K}^{*0}$	floating	-0.081 ± 0.011	-0.024 ± 0.007	-0.002 ± 0.002
	fixed	-0.082 ± 0.003	-0.024 ± 0.002	-0.001 ± 0.002
$\mathrm{B}^0 \to \psi'\mathrm{K}^{*0}$	floating	-0.031 ± 0.001	-0.039 ± 0.001	0.005 ± 0.001
	fixed	-0.033 ± 0.025	-0.040 ± 0.031	0.005 ± 0.011

The difference between the fit results of the P_1 and P_5' parameters are reported in the Table 6.5. Since these differences are very small compared to the statistical errors of the results, this test shows that the bias in the results introduced by the choice of fixing some parameters in the `pdf` is negligible in the final result.

Table 6.5: Difference between the P_1 and P_5' results obtained from a fit sequence with the F_L parameter fixed or free to float.

control channel	$P1$	P_5'
$B^0 \to J/\psi K^{*0}$	0.001	< 0.001
$B^0 \to \psi' K^{*0}$	0.002	0.001

Chapter 7

Systematic uncertainties

In this section the systematic uncertainties considered for this analysis are discussed. Some of the systematics are handled in similar methods as those discussed in the previous CMS analysis [36].

The sources of systematic uncertainties investigated are:

limited amount of simulated events: the propagation of the statistical uncertainty of the MC sample used to compute the efficiency;

simulation mismodelling: the effect of eventual mis-modelling in the simulated angular shape;

efficiency shape: the effect of mis-modelling in the efficiency functions, computed using the control channels;

fitting bias: the possible biases from the fitting procedures, evaluated on data-like "cocktail" MC samples;

wrong CP assignment: the effect of wrong CP assignment on fit results;

background distributions: the effects of the background `pdf` uncertainties, due to finite sideband statistics, on the fit results;

uncertainty from fixed `pdf` parameters: the propagation of the uncertainties on the angular parameters F_L, A_S, F_S;

angular resolution: the effect of the finite reconstruction resolution on the fit results;

feed-through background: the effect of the contamination with $B^0 \to \mathrm{J}\psi K^{*0}$ and $B^0 \to \psi' K^{*0}$ feed-through events, in q^2 bins just below and above the resonance regions;

bivariate Gauss fit range: the dependency of the results on the range of the bivariate fit to the likelihood distribution in the P_1, P_5' plane, when estimating the best-fit value.

In the following sections, the systematic uncertainties sources are discussed and estimated.

7.1 Limited amount of simulated events

The use of kernel density estimator to determine the numerator and denominator of the efficiency is based on a sample of simulated events, and a systematic uncertainty is expected from the limited size of the sample used. The unbinned approach prevents the use of a simple binomial error estimation to be propagated to the parameterisation itself.

The method used to evaluate this systematic contribution makes use of a set of 101 efficiency function based on pseudo-experiments. For each efficiency, we generate pseudo-experiment datasets for numerator and denominator terms of the efficiency, with the same number of events as the original samples, based on the `pdf` returned by the KDE description of numerator and denominator, respectively. These additional datasets are then used to compute efficiency using the same KDE approach used in the original sample. Finally, 101 fits are performed using each of the new efficiency functions. The spread of the fit result distributions obtained by these fits is used as systematic uncertainty.

This method is first tested on a cocktail-MC sample, as the ones used for the validation of the fit described in Section 6.3.1. The results of the fits with the toy-efficiencies are compared to the result of the fit with the original efficiency, computed both on the same cocktail-MC sample and on the full MC sample. This comparison validates the procedure of generating toy-efficiencies, showing that the results obtained with them are compatible with the result obtained with the original efficiency. Then, the method is applied to the data, after the unbinding of the signal region .

An example of the fit result distributions, for q^2 bin 3, is shown in Figure 7.1, for the fits on a cocktail-MC sample, and in Figure 7.2, for the fits on the data sample.

The spread of the fit results are summarised in Table 7.1, for the fit on the cocktail-MC sample, and in Table 7.2 for the fit to the data sample. The latter are used as systematic uncertainty.

7.2 Simulation mismodelling

The effects of the simulation mismodelling is measured through the capability of the analysis to retrieve the interesting observables in extremely favourable conditions, using pure-signal simulation with high statistics.

The fit results on generator-level MC sample are compared with the one on reconstruction-level MC sample, as described in Section 6.2. The discrepancies between them are considered as symmetric systematic uncertainties.

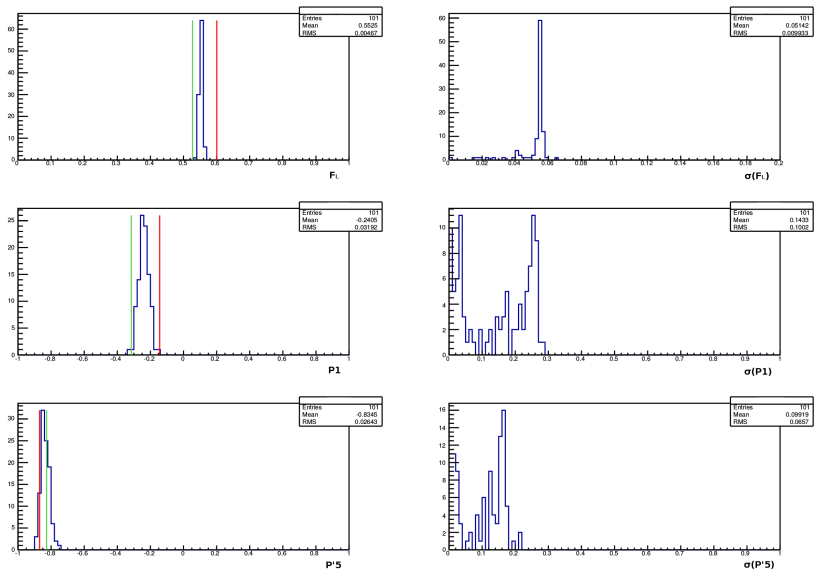


Figure 7.1: Results of the fits on a cocktail-MC sample with toy-efficiency functions, for q^2 bin 3. On the left, the distributions of the F_L (top), P_1 (middle), and P_5' (bottom) parameter results are compared with the results from the fits with the original efficiency: on the same cocktail-MC sample (green line) and on the full MC sample (red line). On the right: the distributions of the `HESSE` errors on the F_L(top), P_1 (middle), and P_5' (bottom) parameters are shown.

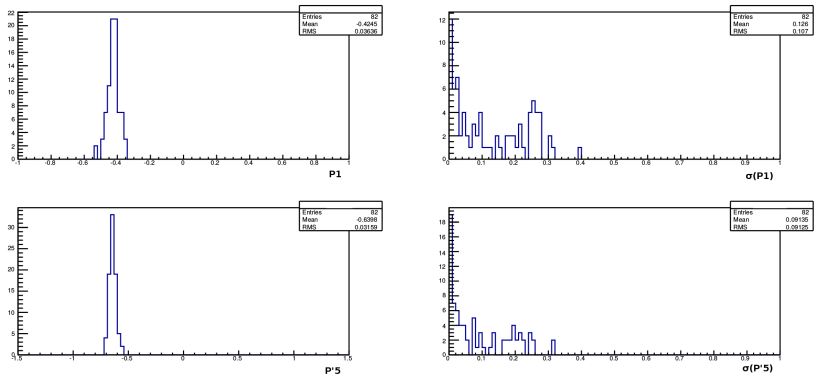


Figure 7.2: Results of the fits on the data sample with toy-efficiency functions, for q^2 bin 3. On the left, the distributions of the P_1 (top) and P_5' (bottom) parameter results are shown. On the right: the distributions of the `HESSE` errors on the P_1 (top) and P_5' (bottom) parameters are shown.

This systematic uncertainty also evaluates the impact of the non perfect

Table 7.1: Spread values of the toy-efficiency result distributions, computed fitting a data-like statistics cocktail-MC sample.

q^2 bin index	F_L	P_1	P_5'
0	±0.0157	±0.0720	±0.0859
1	±0.0110	±0.1208	±0.1070
2	±0.0140	±0.0726	±0.0627
3	±0.0047	±0.0319	±0.0264
5	±0.0051	±0.0170	±0.0099
7	±0.0043	±0.0440	±0.0384
8	±0.0070	±0.0790	±0.0622

Table 7.2: Spread values of the toy-efficiency result distributions, computed fitting real data.

q^2 bin index	P_1	P_5'
0	±0.050	±0.046
1	±0.062	±0.066
2	±0.057	±0.031
3	±0.036	±0.032
5	±0.068	±0.049
7	±0.073	±0.112
8	±0.029	±0.036

symmetry of the efficiency with respect to the angular folding applied. The folding procedure cancels some of the angular parameters only if applied at `pdf` level. The actual fit is performed at reconstruction level, taking into account the efficiency as a function of the angular variables. The efficiency does not have the same exact symmetries as the `pdf`, and this might cause the cancellation to be incomplete. By comparing the results of the fit at generator level, when the cancellation is exact, with that at reconstruction level, when it is not, we evaluate the uncertainty related to the non-perfect cancellation.

The results are summarised for each q^2 bin in Table 7.3.

Table 7.3: Systematic uncertainties: simulation mismodelling.

q^2 bin index	F_L	P_1	P_5'
0	±0.011	±0.005	±0.023
1	±0.001	±0.005	±0.013
2	±0.009	±0.001	±0.015
3	±0.012	±0.006	±0.012
5	±0.006	±0.001	±0.021
7	±0.008	±0.033	±0.010
8	±0.004	±0.006	±0.014

7.3 Efficiency shape

The main validation of the correctness of the efficiency is performed by comparing the efficiency-corrected results for the control channels with the corresponding world-average values. The efficiency as a function of the angular variables is checked by comparing the $F_{\rm L}$ measurements from the $\rm B^0 \to J/\psi K^{*0}$ sample, composed of 165 000 signal events. The value of $F_{\rm L}$ obtained in this analysis is $0.537 \pm 0.002\,(\text{stat})$, compared with the world-average value of $0.571 \pm 0.007\,(\text{stat+syst})$, indicating a discrepancy of 0.034, which is used in the other q^2 bins and propagated to the P_1 and P_5' parameters. In each q^2 bin, a total of 200 values for the F_L parameter is randomly generated from a Gaussian distribution, with mean the value of F_L used in the fit to the data and with width the discrepancy to propagate, 0.034. Finally, the data sample is fitted fixing the F_L parameter to each of these 200 values and the RMS of the fit results is taken as systematic uncertainty. The results are summarised in Table 7.4.

For completeness the $F_{\rm L}$ variable is also measured with the $\rm B^0 \to \psi' K^{*0}$ sample, obtaining a value $0.538 \pm 0.008\,(\text{stat})$ to be compared with the world-average value of $0.463^{+0.028}_{-0.040}\,(\text{stat+syst})$.

Table 7.4: Systematic uncertainties: efficiency shape.

q^2 bin index	P_1	P_5'
0	± 0.017	± 0.005
1	± 0.048	± 0.060
2	± 0.093	± 0.065
3	± 0.094	± 0.045
5	± 0.083	± 0.059
7	± 0.100	± 0.060
8	± 0.068	± 0.041

Cross check: branching fraction of $\mathbf{B^0 \to K^* \psi'}$ and $\mathbf{B^0 \to K^* J/\psi}$

A further test to validate the efficiency shape obtained from MC is to compare the branching fraction of the two control samples: $\rm B^0 \to K^* \psi'$ and $\rm B^0 \to K^* J/\psi$.

The ratio of the two BR can be computed as follow:

$$\frac{\mathcal{B}(\mathrm{B^0} \to \mathrm{K^*}\psi')}{\mathcal{B}(\mathrm{B^0} \to \mathrm{K^*}J/\psi)} = \frac{Y_{\psi'}}{\epsilon_{\psi'}} \frac{\epsilon_{\mathrm{J}/\psi}}{Y_{\mathrm{J}/\psi}} \frac{\mathcal{B}(J/\psi \to \mu^+\mu^-)}{\mathcal{B}(\psi' \to \mu^+\mu^-)} \tag{7.1}$$

where $Y_{\psi'}$ and $\epsilon_{\psi'}$ are the yield and the efficiency for the ψ' channel, and likewise for $Y_{\mathrm{J}/\psi}$ and $\epsilon_{J/\psi}$ for J/ψ one.

The ratio of $\mathcal{B}$ for $J/\psi(\psi') \to \mu^+\mu^-$ is 7.54 ± 0.86 (PDG). The value of the ratio is (PDG): $\frac{\mathcal{B}(\mathrm{B^0} \to \mathrm{K^*}\psi')}{\mathcal{B}(\mathrm{B^0} \to \mathrm{K^*}J/\psi)} = 0.484 \pm 0.018_{(\text{stat})} \pm 0.011_{(\text{syst})} \pm 0.012_{(R^{ee}_{\psi})}$.

A first way to compute the ratio is to use directly the absolute efficiency for the two channels as obtained from MC. The ratio we obtain with this method is $0.476 \pm 0.008(\text{stat}) \pm 0.055(\text{R}_{\psi}^{\mu\mu})$, in very good agreement with PDG.

However, the computation above assumes that the angular shape of the control channel is correctly simulated in MC. To take properly into account the real signal and efficiency shape, we repeated the same computation using as efficiency $\varepsilon_{J/\psi/\psi'} = \int_{phase\ space} S(\vec{x};\vec{p}) \times \varepsilon(\vec{x}) d\vec{x}$, where $S(\vec{x};\vec{p})$ is the signal `pdf`, $\vec{p}$ is the set of angular parameter we got from the fit on the data on each control region, and $\varepsilon(\vec{x})$ is the efficiency for each region as a function of angular variable. The result of this more detailed computation is: $0.480 \pm 0.008(\text{stat}) \pm 0.055(\text{R}_{\psi}^{\mu\mu})$, again in very good agreement with PDG.

The same ratio can be computed using only right tag events, only wrong tag events, or both: in all the cases the agreement is very good.

7.4 Fitting Bias

The fitting procedure itself could generate biases in the results, in addition to the uncertainties from the fitting components (efficiency, background shape, `pdf`, etc) described above. The fitting procedure uncertainties thus arise from the possible biases from the fitting methods and procedures.

We estimate this contribution from the fit results of the data-like simulated samples, which are described in Section 6.3.2.

To evaluate the fitting bias, we compare the average result of the fits to cocktail MC samples and the fitting result of full statistics MC sample. These differences are used as an estimation of this systematic uncertainty. The results are summarised in Table 7.5.

Table 7.5: Systematic uncertainties: fitting bias.

q^2 bin index	P_1	P_5'
0	± 0.005	± 0.040
1	± 0.007	± 0.010
2	± 0.019	± 0.119
3	± 0.019	± 0.106
5	± 0.041	± 0.052
7	± 0.071	± 0.048
8	± 0.078	± 0.031

7.5 Wrong CP assignment

The error on mistag fraction has been estimated using the control channel $B^0 \to J/\psi K^{*0}$, in q^2 bin 4. A full fit on this channel has been performed, leaving also the mistag fraction free to float. This fit is possible on the control channel thanks to its large statistics, and the result has been compared with the mistag fraction estimated from MC sample. The difference between these values is 0.008 and, compared with the statistical uncertainty of the MC estimate in this bin, $\sigma \lesssim 0.002$, it results to be the dominant uncertainty related to the mistag fraction.

To assess the propagation of this systematic effect to the physical parameter of the final fit on data, a set of fits has been performed, using as free parameters A_s^5, P_1, P_5', the signal and background yields, and using a fixed mistag fraction. A set of 10 different values of mistag fractions has been used, randomly generated with Gaussian distribution around the mistag fraction estimated from MC for each q^2 bin, and with σ equal to the difference of 0.008 computed on the control channel. For each q^2 bin, the RMS of the results of the 10 fits is used as systematic uncertainty for the fitted angular parameters, as summarised in Table 7.6.

Table 7.6: Systematic uncertainties: mistag fraction.

q^2 bin index	A_s^5	P_1	P_5'
0	0.000035	0.014	0.013
1	0.000014	0.022	0.015
2	0.0050	0.016	0.014
3	0.0025	0.0084	0.0058
5	0.0037	0.043	0.032
7	0.015	0.11	0.066
8	0.0050	0.025	0.016

7.6 Background distributions

The parameters of the background component in the `pdf` are estimated on the mass sidebands and kept fixed in the fit to the full mass range, as described in Section 4.2.2. Due to the limited statistics of the mass sideband samples, the background parameters have a not-negligible uncertainty, that should be propagated to the analysis results as systematic uncertainty.

The errors of the background `pdf` parameters are evaluated by the `HESSE` algorithm running during the fit to the mass sidebands. A set of 200 background toy-functions are then generated by using the same polynomial expression described in Section 4.2.2, and randomly generating the value of

each parameter from a Gaussian distribution, with mean and width equal to the HESSE results of the sideband fit.

To propagate this uncertainty to the signal parameters, the fit to the data is run 200 times, by using the new background toy-functions. The widths of the result distributions are used as systematic uncertainties.

An example of the distributions of the toy-background fit results, for q^2 bin 2, is shown in Figure 7.3.

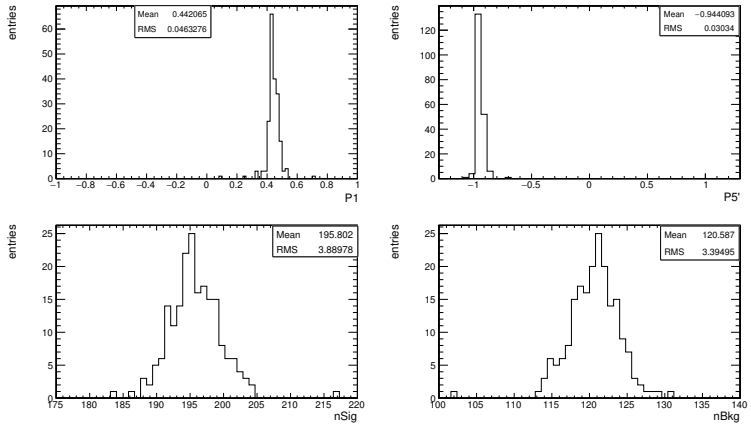


Figure 7.3: The results of the fits on data performed with background toy-functions, for the q^2 bin 2. The distributions of P_1 (top left), P_5' (top right), signal yield (bottom left), and background yield (bottom right) are shown.

The resulting uncertainties are summarised in the Table 7.7.

Table 7.7: Systematic uncertainties: background shape.

q^2 bin index	P_1	P_5'
0	±0.013	±0.010
1	±0.031	±0.024
2	±0.046	±0.030
3	±0.022	±0.013
5	±0.070	±0.049
7	±0.069	±0.051
8	±0.012	±0.015

Partially reconstructed background Since the lower sideband in the candidate B^0 mass has a value range starting at about 5 GeV, it can contain a contamination from partially reconstructed events, i.e. events in which a B^0 meson decays in a final state with more than four detectable particles and the candidate B^0 is built using only four of them. The kinematic end-point for this kind of events is between 5.1 GeV and 5.15 GeV. Since

the angular distribution of these background events is expected to be different with respect to the distribution of the combinatorial background, a contamination of the lower sideband would bias the fit of the background parameters of the `pdf`, estimated in the first step of the fit, according to the fit sequence described in Section 4.3. Then, the biased background `pdf` is used in the second step of the fit, even though under the resonance there is no contamination, and the fitted values of P_1 and P_5' can be affected.

In order to check whether an eventual contamination has introduced a significant bias in the fitted values of P_1 and P_5', the two steps of the fit are repeated, but with the range of the low-mass sideband starting at about 5.1 GeV (exactly 1 GeV more narrow than the original sideband), and the results of the P_1 and P_5' parameters are compared with the original fit. A natural difference in these results can derive from the statistical fluctuation due to the use of a sub-sample of the original dataset. For this reason, the differences observed in this test are compared with the values of the *Background Distribution* systematic uncertainty.

The results of the fits with the reduced sideband are reported in Table 7.8, together with their absolute difference with respect to the original results. The values of the *Background Distribution* systematic uncertainty are reported in this table as well.

In the table are highlighted the differences that exceed twice the value of the systematic uncertainty. Globally, the differences are considered compatible with the systematic uncertainty. In the few bins with a larger difference, the results of the fit with reduced sidebands are affected by the very low statistics left in the sideband region. The worst case is bin 7, where the sideband reduction remove about 75% of the background events.

Table 7.8: Results of the fits with the reduced sideband and their absolute difference with respect to the results with the full sideband, as described in Section 7.6. The values of the *Background Distribution* systematic uncertainty, as described in Section 7.6, are reported here for comparison sake.

q^2 bin	P_1			P_5'		
index	result	difference	syst	result	difference	syst
0	0.124	0.005	0.013	0.077	*0.024*	0.010
1	-0.675	0.010	0.031	-0.541	0.026	0.024
2	0.505	0.028	0.046	-0.888	*0.069*	0.030
3	-0.457	0.013	0.022	-0.689	*0.046*	0.013
5	-0.490	0.041	0.070	-0.714	0.024	0.049
7	-0.502	*0.173*	0.069	-0.705	0.069	0.051
8	-0.520	0.013	0.012	-0.565	0.006	0.015

7.7 Mass Distribution

The parameters of the mass shape in the signal component of the `pdf` are estimated through a fit to the MC mass distribution and kept fixed in the data fit, as described in Section 4.3. To evaluate the bias due to mis-modelling of the MC description of the mass shapes, and propagate it to the analysis results as a systematic uncertainty, the control channels are used. Thanks to the high-statistics of these samples, it is possible to perform a fit with some mass parameters free to float. The angular parameters are measured on both control channels, at first with signal mass shape for right-tagged events free to float, and then with signal mass shape for mis-tagged events free to float. The difference of the angular parameter values obtained from these fits, with respect to the results with the standard fit sequence, are calculated. For each parameter, the largest of these differences is used as systematic uncertainty for all the q^2 bins.

The maximum difference of $P1$ is 0.012, of $P5'$ is 0.019.

7.8 Uncertainty from fixed `pdf` parameters

Fixing three of the six angular parameters in the fit could have an effect of modifying the statistical uncertainties for the fitted parameters. This is true, in general, whenever there is a correlation between any fixed parameter and a free one. Since this effect could lead to an underestimation of the statistical error, a specific systematic uncertainty has been created to compensate this effect. For each measured parameter we define a *scale factor* as the correction factor, greater than or equal to one, that should be applied to the underestimated statistical error to compensate this effect.

Since it is not possible to fit the data with the F_L, F_s, A_s parameters free to float, this scale factor is computed on pseudo-experiments with one hundred times the statistics of the data sample. Ten pseudo-experiments are used, for each q^2 bin, and the average of the ten resulting scale factors is used to compute the uncertainty. The same procedure described in Sec. 8.1 to fit a pseudo-experiment for the Feldman-Cousins (FC) method is used here to fit each high-statistics pseudo-experiment. The only difference is that the number of points where the 2D-likelihood is probed here is around one hundred, in order to guarantee a more robust result. This fitting procedure is applied twice, both fixing the F_L, F_s, A_s parameters and leaving them free to float. Then, two scale factors are defined, one for P_1 and one for P_5', as the ratio between the confidence intervals obtained with the two fits (fit with floating parameters over fit with fixed parameters). Note that here the `custom MINOS` method, defined in Sec. 8.1, is used to define the parameters' confidence interval.

Finally, each scale-factor average was used to compute the systematic

uncertainty. If it is smaller than one, it means that the estimated effect of freeing the parameters is to reduce the statistical error. In this case, no correction is applied. Otherwise, if it is larger than one, we define the uncertainty value as the one that, if added in quadrature with the statistical error, would increase it by this estimated factor. Thus, it is

$$S = \sigma\sqrt{f^2 - 1} \tag{7.2}$$

where σ is the statistical error and f is the scale-factor average.

The scale-factor averages and the systematic uncertainty values are reported in table 7.9.

Table 7.9: Scale-factor average values and systematic uncertainties computed to compensate the statistical error reduction introduced fixing some PDF parameters

q^2 bin index	SF(P_1)	P_1	SF(P'_5)	P'_5
0	1.014	0.077	1.003	0.025
1	1.116	0.211	1.099	0.148
2	1.113	0.139	1.385	0.206
3	1.082	0.103	1.028	0.041
5	1.048	0.053	1.143	0.069
7	0.982	0.000	1.090	0.072
8	1.091	0.083	0.989	0.000

To verify that the procedure of fixing those parameters in the fit is only affecting the statistical uncertainty, but it is not introducing any significant bias in the best-fit values, we compared the results of the fits to the sets of pseudo-experiments, both with floating and fixed parameters, with the input values used to generate them. An example of the plots used for this comparison, for q^2 bin 2, is shown in Figure 7.4. No significant bias are present.

7.8.1 Test of scale-factor dependency on toy statistics

The procedure used to compute the systematic uncertainty due to the fixed parameters, as described above, relies on the assumption that the scale factors do not depend on the statistics of the pseudo-experiment used. This assumption has been tested for two q^2 bins, for two different q^2 regimes, bin 0 and bin 8.

The test consists in repeating the procedure to extract the scale factor on a set of pseudo-experiments with different statistics. Defining R as the ratio between the statistics of a pseudo-experiment and the data sample, the procedure used for the estimation of the systematic uncertainty used

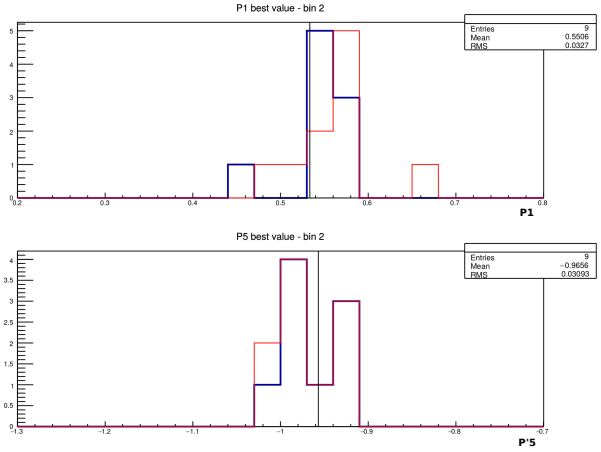


Figure 7.4: Distributions of the best-fit values of the fits performed, in q^2 bin 2, on the high-statistics pseudo-experiments used to compute the systematic uncertainty for the fixed parameters, as described in Section 7.8. The blue (red) histogram is the distribution of the results of the fit with fixed (floating) F_L, F_s, A_s parameters. The black line marks the input value used to generate the pseudo-experiments. Top: results for P_1. Bottom: results for P_5'.

a sample with $R = 100$. The test is performed with a set of 9 samples, with $R = \{10, 20, \ldots, 90\}$. The same procedure used for the systematic uncertainty is applied on these pseudo-experiments, to extract the scale factor for both P_1 and P_5'.

The set of scale factors obtained is shown in Figure 7.5 and Figure 7.6, for bin 0 and bin 8 respectively. From these results, one can notice that, besides few isolated points, all the scale factors are in agreement and there are no evident trends as a function of the statistics.

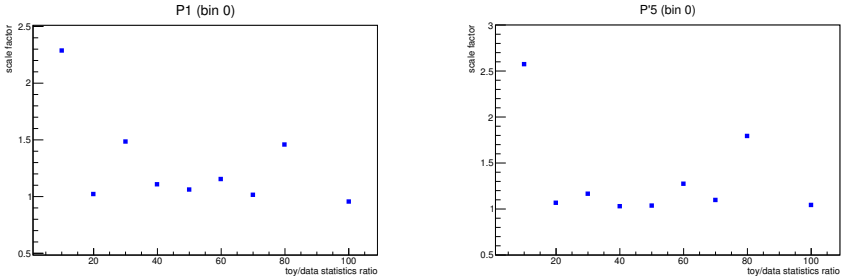


Figure 7.5: The error scale factors, obtained with the procedure described in Section 7.8, are shown for P_1 (left) and P_5' (right), in q^2 bin 0, as a function of the statistics of the pseudo-experiment used.

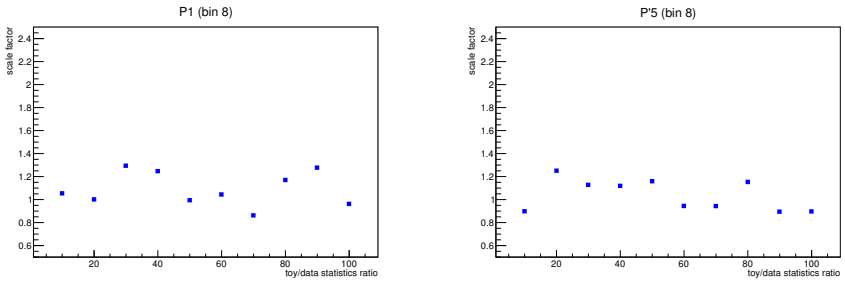


Figure 7.6: The error scale factors, obtained with the procedure described in Section 7.8, are shown for P_1 (left) and P_5' (right), in q^2 bin 8, as a function of the statistics of the pseudo-experiment used.

7.9 Angular Resolution

This systematic uncertainty is due to the limited detector resolution on angular distributions. To evaluate its impact, the likelihood fit has been performed on the full simulated dataset on reconstructed quantities after all selections, considering only the right-tagged candidates. The same fit has been performed on the same sample, but using the generated angular quantities in place of the reconstructed ones. The difference between the two results, shown in table 7.10, is used as a systematic uncertainty for each q^2 bin.

Table 7.10: Difference on the target physics observables (F_L, P_1, and P_5') when obtained via a fit on reconstructed and generated angular distributions after all selection, performed on full MC sample. $\Delta x = |x_{RECO} - x_{GEN}|$

q^2 bin index	ΔF_L	ΔP_1	$\Delta P_5'$
0	$-3.98\cdot10^{-4}$	$-1.50\cdot10^{-3}$	$1.05\cdot10^{-4}$
1	$-7.83\cdot10^{-4}$	$-3.30\cdot10^{-3}$	$1.03\cdot10^{-3}$
2	$-8.83\cdot10^{-3}$	$-6.88\cdot10^{-3}$	$1.33\cdot10^{-3}$
3	$-4.94\cdot10^{-4}$	$-7.38\cdot10^{-3}$	$8.10\cdot10^{-4}$
5	$1.16\cdot10^{-3}$	$-1.51\cdot10^{-2}$	$2.36\cdot10^{-3}$
7	$2.90\cdot10^{-4}$	$-7.26\cdot10^{-3}$	$-8.28\cdot10^{-3}$
8	$-1.93\cdot10^{-3}$	$-6.81\cdot10^{-2}$	$1.23\cdot10^{-2}$

7.10 Feed-through background

The q^2 bins just below and above the resonance regions may be contaminated by $B^0 \to \mathrm{J}/\psi K^{*0}$ and $B^0 \to \psi' K^{*0}$ feed-through events that are not removed by the selection criteria. The potential uncertainties due to this contamination have been evaluated.

The distribution of residual feed-through events are described by a special `pdf`. The distributions in q^2 bins 3 and 5 are evaluated from $B^0 \to \mathrm{J}/\psi K^{*0}$ simulation, as shown in Figure 7.7 and Figure 7.8. Similarly, the distributions in q^2 bin 5 are evaluated from $B^0 \to \psi' K^{*0}$ simulated sample

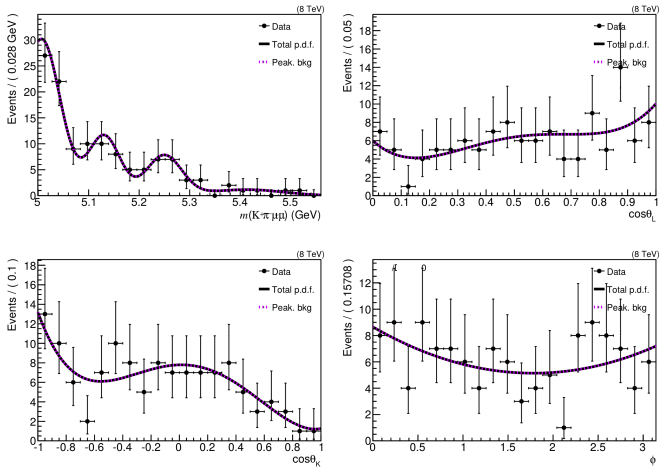


Figure 7.7: Distributions of the feed-through events of the $\mathrm{B}^0 \to \mathrm{J}/\psi\mathrm{K}^{*0}$ MC sample and the projections of the `pdf` used to describe it, for q^2 bin 3.

Data are then fitted with this additional component of feed through backgrounds, as shown in Fig 7.9 and Fig 7.10.

Discrepancies on the measured observables are then conservatively considered as symmetric systematic uncertainties. There are two results for q^2 bin 5, one from $B^0 \to \mathrm{J}/\psi K^{*0}$ and the other from $B^0 \to \psi' K^{*0}$. Then we choose the bigger one as the systematic uncertainty. The result are in the Table 7.11.

Table 7.11: Systematic uncertainties from the feed through backgrounds.

q^2 bin index	P_1	P_5'
3	± 0.004	± 0.012
5	± 0.012	± 0.020
7	± 0.011	± 0.024

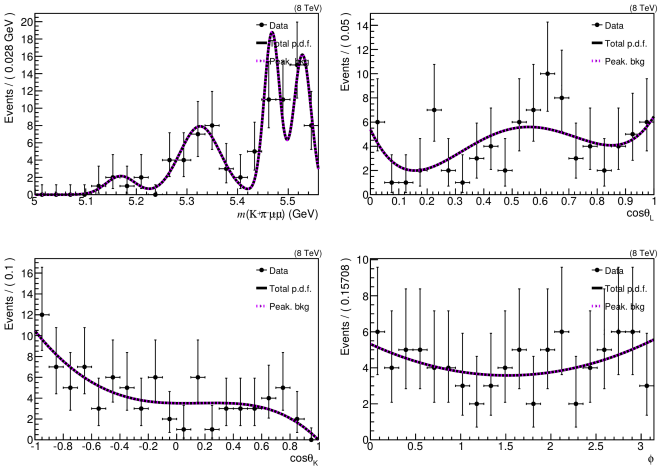


Figure 7.8: Distributions of the feed-through events of the $B^0 \to J/\psi K^{*0}$ MC sample and the projections of the `pdf` used to describe it, for q^2 bin 5.

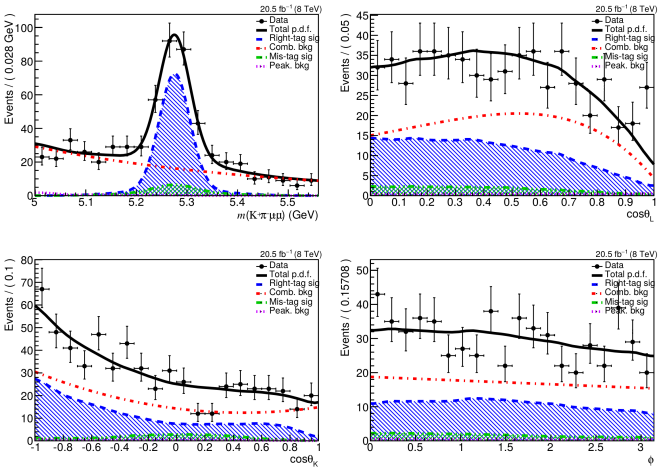


Figure 7.9: Distributions of the data events, and the projections of the `pdf` with the additional component for feed-through background, for q^2 bin 3.

7.11 Bivariate Gaussian fit range

As described in Section 4.3, the procedure to get the best fit from the data is to discretise the P_1, P_5' space, and maximise the likelihood as a function of nuisance parameters at fixed values of P_1, P_5'. So we have a scan of the likelihood in the P_1, P_5' plane. Finally, the likelihood distribution is fitted with a bi-variate normal distribution.

In order to check the stability of the fit, as well as any possible source

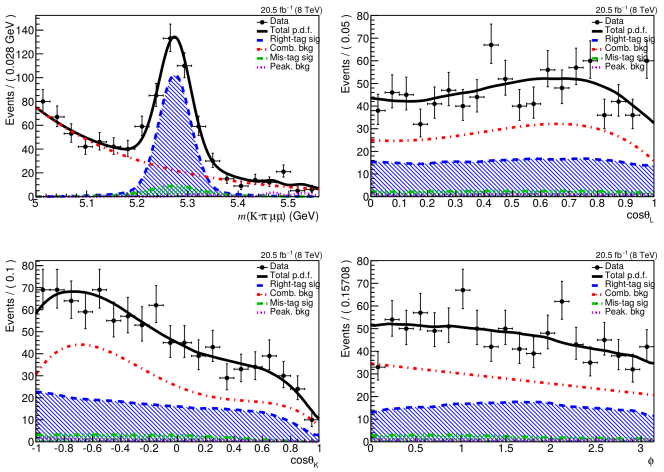


Figure 7.10: Distributions of the data events, and the projections of the `pdf` with the additional component for feed-through background, for q^2 bin 5.

of bias due to the procedure, the angular parameters are evaluated varying the fit range. The range of the fit is defined as $\pm 1\sigma$ (as computed by the FC procedure described in Section 8.1), multiplied by a scale factor (in the range $[0.1, 2]$), around the centre defined as the value of P_1 or P_5', among the points of the grid defined in Section 4.3, where the likelihood has an absolute maximum.

An example of the results of these scans is shown in Figure 7.11, for q^2 bin 3. To give an element of comparison, the y-axis range in the plots of the results is set equal to the confidence interval, as computed by the FC procedure described in Section 8.1. The difference between the value of P_1 and P_5' from the absolute maximum and the normal fit is small, when compared with statistical uncertainties and to the *fit bias* uncertainty. Also the trend with respect to the fit range is small.

A quantitative difference has been extracted by comparing the value of the position of the absolute maximum with that of a bi-variate Gaussian fit, via linear fit in the scale factor range $[0.4, 1.6]$, evaluated for scale factor equal to 1. The differences are reported in table 7.12.

Since these differences have the same nature of the *fit bias* uncertainty, but are smaller than it, no specific systematic uncertainty has been introduced for them.

7.12 Total Systematic Uncertainties

The summary of all systematic uncertainties are shown in Table 7.13.

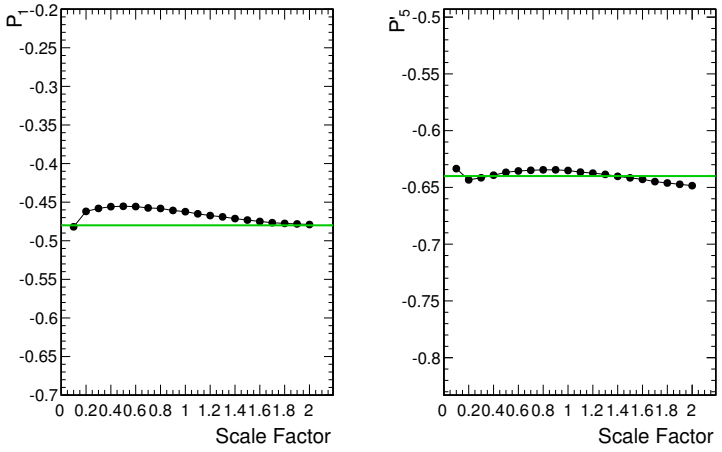


Figure 7.11: Comparison between the position of the absolute maximum of the likelihood (green line) and the results of the bi-variate Gaussian fit, as a function of the fit range (black dots), for the P_1 (left) and P_5' (right) parameters, in q^2 bin 3.

Table 7.12: Systematic uncertainties: bias from bi-variate Gaussian fit to the likelihood.

q^2 bin index	P_1	P_5'
0	± 0.012	± 0.009
1	± 0.005	± 0.005
2	± 0.002	± 0.007
3	± 0.016	± 0.0002
5	± 0.009	± 0.020
7	± 0.010	± 0.012
8	± 0.004	± 0.001

Table 7.13: Systematic uncertainty contributions for the measurements of $P1$ and $P5'$. The total uncertainty in each q^2 bin is obtained by adding each contribution in quadrature. For each item, the range indicates the variation of the uncertainty in the signal q^2 bins.

Source	$P_1(\times 10^{-3})$	$P_5'(\times 10^{-3})$
Simulation mismodeling	1–33	10–23
Fit bias	5–78	10–120
Finite size of simulated samples	29–73	31–110
Efficiency	17–100	5–65
Kπ mistagging	8–110	6–66
Background distribution	12–70	10–51
Mass distribution	12	19
Feed-through background	4–12	3–24
$F_{\rm L}$, $F_{\rm S}$, $A_{\rm S}$ uncertainty propagation	0–210	0–210
Angular resolution	2–68	0.1–12
Total	100–230	70–250

Chapter 8

Fit results

In this section I will present the results of the fit to the data sample, after removing the *blind* status to the analysis. Firstly, the methodology used to extract the statistical uncertainty is presented. Then the best-fit values of the parameters are reported, together with the total statistical and systematical uncertainties. Finally the distributions and the `pdf` profiles for each q^2 bin are shown.

If not otherwise specified, we are using the fitting procedures described in Sect.4.3 and other relevant sections.

8.1 Statistical uncertainties determination

The determination of the statistical uncertainties for the measured parameters cannot just be delegated to the fit software, and in particular to `MINOS`, due to the presence of physical boundaries on the parameters. These boundaries are discussed in Section 1.2.3, and comes from the requirement that the `pdf` is positive defined everywhere.

Different approaches to the problem have been explored: in the following we describe three of them: `custom MINOS`, `hybrid frequentist-Bayesian`, and `Feldman-Cousins`. For the final determination of the statistical uncertainties we have used the latter, despite of its complexity and the huge CPU time required.

In this thesis, only the `Feldman-Cousins` approach is described.

8.1.1 Feldman-Cousins method

The approach used for the computation of statistical uncertainties, strongly suggested by the CMS Statistical Committee, is to apply the Feldman-Cousins method [34] with nuisance parameters. Given the enormous time that would be needed to build a full bi-dimensional confidence interval in the P_1-P_5' parameter space, we decided to perform the F-C procedure only

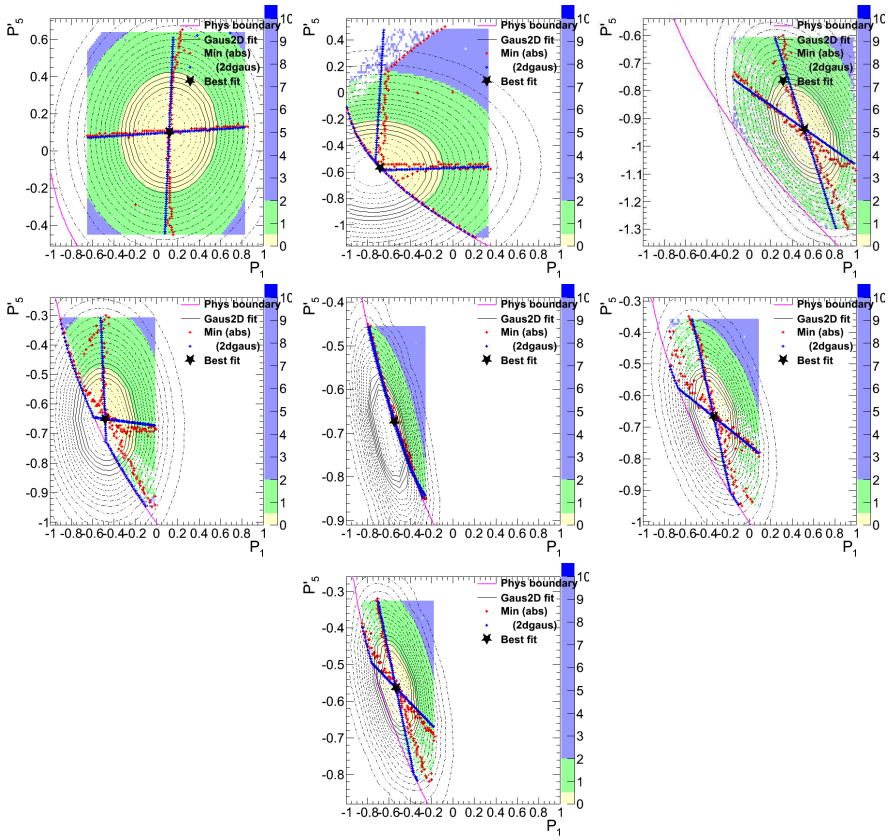


Figure 8.1: Distribution of $\mathcal{L}$ in the P_1, P_5', with contour curves at $\Delta \log \mathcal{L} = 0.5$ and 2.0, with the indication of the physical region. Superimposed is a bivariate Gaussian fit, and the position of the maximums of the profiled $\mathcal{L}$ (red) and that of the Gaussian fit (blue).

along the maximum of the profile along P_1 and P_5', respectively, of the $\mathcal{L}$ distribution, and extract two mono-dimensional confidence intervals.

We consider the $\mathcal{L}$ distribution built fitting the data, as described in Section 4.3, and look for the maximum of that distribution along one variable, while the other one is fixed. In general, an evaluation of the likelihood is not available for all the points in the scan of the P_1, P_5' plane, since, for some P_1, P_5' points, the fit is not converging. Thus, building the profiles of the likelihood just looking for the maximum among the valid points leads often to unstable results. So, we perform a bivariate Gaussian fit to the $\mathcal{L}$ points available, similar to the one used to estimate the best-fit values of the parameters, and take the two profiles of the fitted function, instead. The profiles of the function are always restricted inside the physical region: the results are shown in fig 8.1.

This procedure is more robust than just looking for the maximum of the

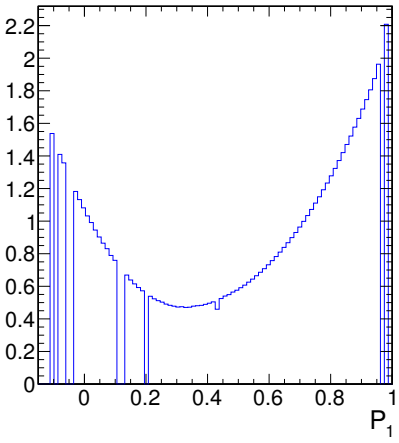


Figure 8.2: An example for profile of $\log \mathcal{L}$, where it is possible to see the two possible issues to find the minimum just scanning the results, instead of performing a fit. Some bins are empty since the fit for those points fails, and there is a downward fluctuation of $\log \mathcal{L}$ which creates a fake minimum.

$\mathcal{L}$ on a profile, since we do not have a determination of $\mathcal{L}$ for every bin. This is typically due to the fact that the physical boundary in the $P_1,\ P_5'$ region depends on A_s^5, so the fit of a particular point might converge to a set of values outside the physical region, and so it fails. Moreover, there is the chance that a determination of the $\log \mathcal{L}$ has a downward fluctuation which can be the minimum even though the shape of the $\log \mathcal{L}$ shows a parabolic behaviour, with a different vertex. This kind of behaviour can be seen in Figure 8.2.

For each of the point in the scan of the $P_1,\ P_5'$ plane defined above, thereafter referred to as GEN-points, we generate a set of 100 toy MC with data-like statistics, using the full `pdf`, with P_1 and P_5' parameters defined by the coordinates of the GEN-point, and with nuisance parameter A_s^5 which guarantees the wider physical region for $P_1,\ P_5'$, namely 0.99 or -0.99, depending on the bin.

Each toy is fitted with a procedure identical to that used for the data, the only difference is that instead of evaluating the likelihood on a 90×90 grid in the $P_1,\ P_5'$ plane, only 20 points are scanned, in order to reduce the time needed to perform the whole procedure. These 20 points are chosen randomly, according to a bivariate Gaussian, around a central value. The width of the Gaussian is the one returned by the fit on the data on the bin under consideration, while correlation is not considered. The central value is the result of a fit on the toy with $P_1,\ P_5'$ free to float, or, if the fit does not converge, the GEN-point itself. If one of these points falls outside the physical boundary, it is discarded and an additional point is generated to

replace it. The $\mathcal{L}$ is evaluated in these 20 points and it is finally fitted with a bivariate Gaussian. An example of such a fit is shown in Figure 8.3.

Eventually, for each toy of any GEN-point, a determination of the $\mathcal{L}$ function is available. In order to check if the GEN-point is inside or outside the 68% CL region, we compare the $\Delta \log \mathcal{L}$ of the toys with the one of the data. In particular, when computing the boundary of the parameter P_1 (P_5'), $\Delta \log \mathcal{L}_{toy_i}$ is defined, on the projection of $\log \mathcal{L}$ of the toy_i on the P_1 (P_5') axis, as the difference between the maximum of this projection and its value at $P_1 = P_{1\ \mathrm{GEN}}$ ($P_5' = P_{5\ \mathrm{GEN}}'$), where ($P_{1\ \mathrm{GEN}}$, $P_{5\ \mathrm{GEN}}'$) are the coordinates of the GEN-point. $\Delta \log \mathcal{L}_{Data}$ is defined as the difference between $\log \mathcal{L}$ of the best fit on data and that of the GEN-point.

The *ratio* value is computed as the fraction of toys with $\Delta \log \mathcal{L}$ lesser than that of the data. A GEN point is considered inside the 1σ region, namely the 68.27% CL region, if the ratio is lower than 68.27%. An example of this comparison is shown in fig. 8.3.

The final 68% confidence interval for P_1 and P_5' is found by looking at the distribution of the ratio as a function of P_1 and P_5', respectively, and fitting this distribution with a linear function. The crossing of the linear function with the 68.27% line defines the $\pm 1\sigma$ value. In total four directions were scanned, corresponding to lower and upper bound for P_1 and P_5'.

In order to reduce the time consumption of this procedure, we started scanning GEN-points around the value of $\Delta \log \mathcal{L} = 0.5$, and then extend the scanned region inside or outside depending on the results. In case the slope of the linear fit was not very steep, more GEN-points were scanned to make the result more robust.

The results of the FC procedure described above are shown in fig 8.4 to 8.10 for all the seven q^2 bins considered. The intervals, as well as the central values, are summarised in table 8.1.

Table 8.1: Summary of results of P_1 and P_5' in different q^2 bins with the $\pm 1\sigma$ statistical uncertainties as computed with the FC procedure.

	P_1		P_5'	
Bin	Fit		Fit	
0	0.119	$^{+0.46}_{-0.47}$	0.101	$^{+0.32}_{-0.31}$
1	−0.685	$^{+0.58}_{-0.27}$	−0.567	$^{+0.34}_{-0.31}$
2	0.533	$^{+0.24}_{-0.33}$	−0.957	$^{+0.22}_{-0.21}$
3	−0.470	$^{+0.27}_{-0.23}$	−0.643	$^{+0.15}_{-0.19}$
5	−0.531	$^{+0.2}_{-0.14}$	−0.690	$^{+0.11}_{-0.14}$
7	−0.329	$^{+0.24}_{-0.23}$	−0.664	$^{+0.13}_{-0.2}$
8	−0.533	$^{+0.19}_{-0.19}$	−0.559	$^{+0.12}_{-0.12}$

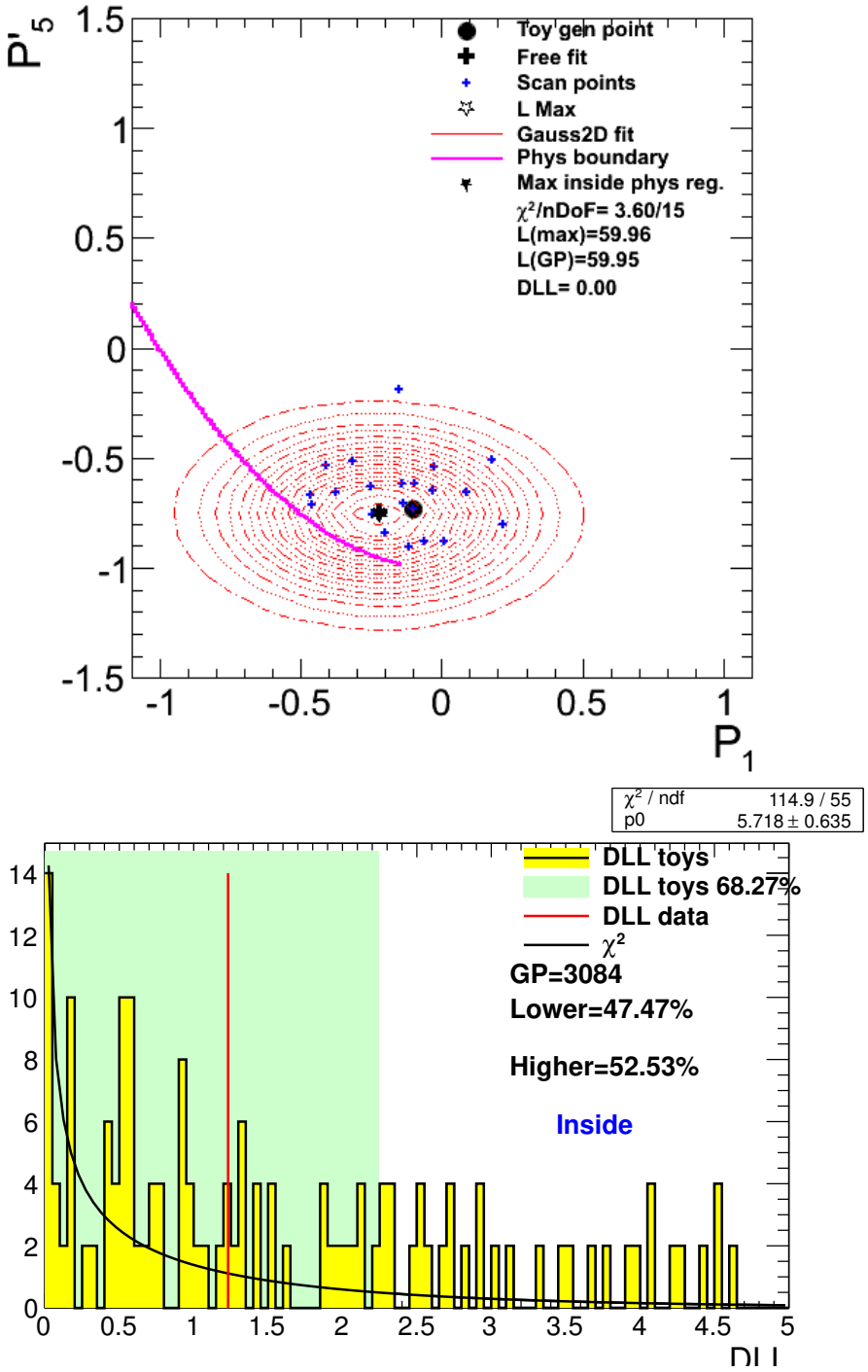


Figure 8.3: An example of a bivariate fit on the 20 scan points for one GEN-point (top). Distribution of the $\Delta \log \mathcal{L}$ values for the 100 toys for one GEN-point and comparison with $\Delta \log \mathcal{L}$ for data, computed for that GEN-point (bottom).

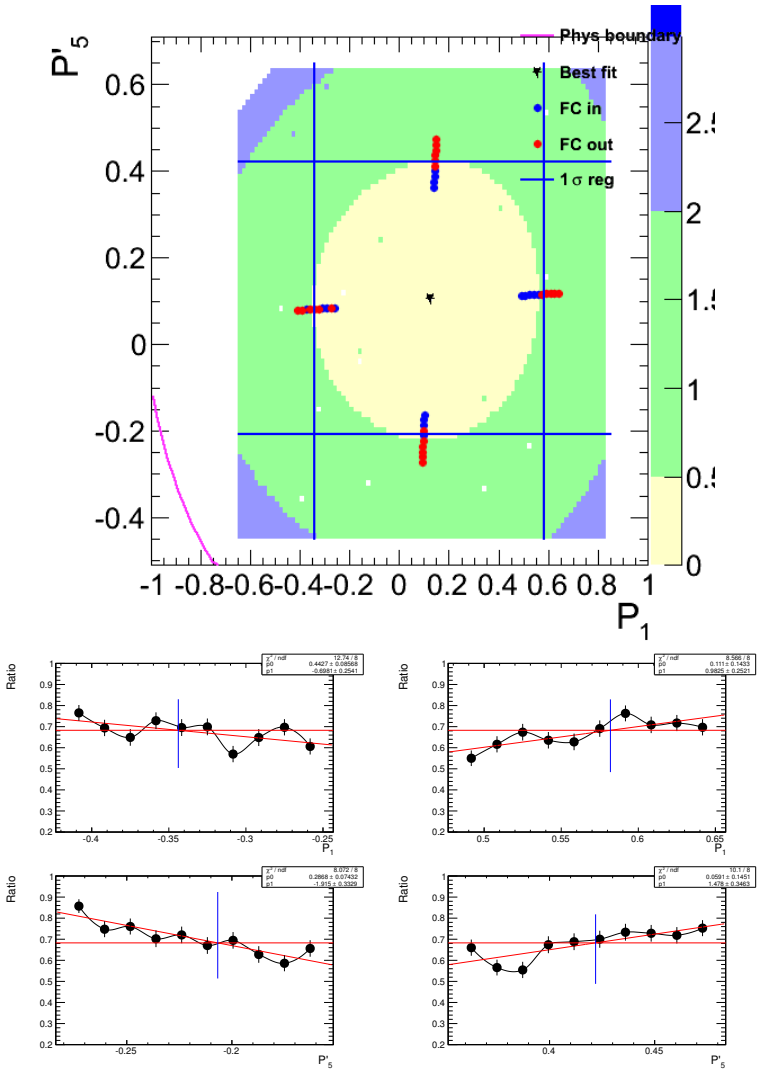


Figure 8.4: F-C results for bin 0. Scan of the GEN-points, superimposed to data $\mathcal{L}$ distribution: red points are outside the 68% CL region, the blue ones are inside. The blue lines defines the $\pm 1\sigma$ region (top). Ratio distribution as a function of P_1 for lower and upper bounds, with linear fit and 68% horizontal line (middle left, right). Likewise for P_5' (bottom).

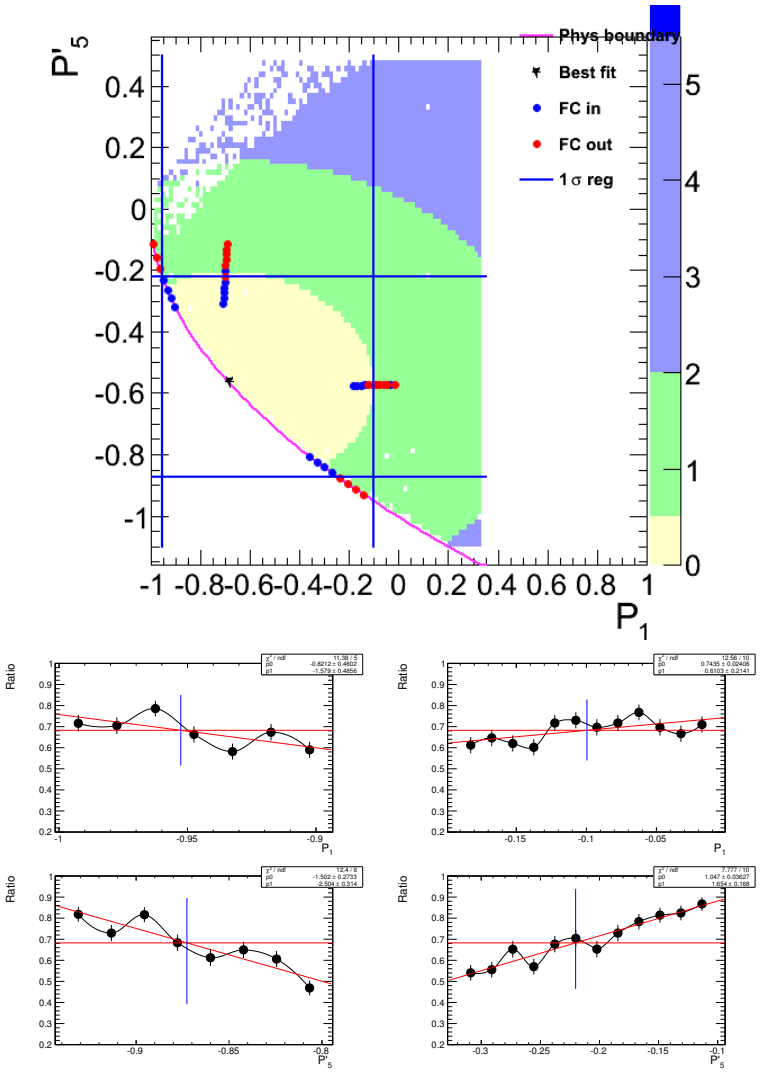


Figure 8.5: F-C results for bin 1. Scan of the GEN-points, superimposed to data $\mathcal{L}$ distribution: red points are outside the 68% CL region, the blue ones are inside. The blue lines defines the $\pm 1\sigma$ region (top). Ratio distribution as a function of P_1 for lower and upper bounds, with linear fit and 68% horizontal line (middle left, right). Likewise for P_5' (bottom).

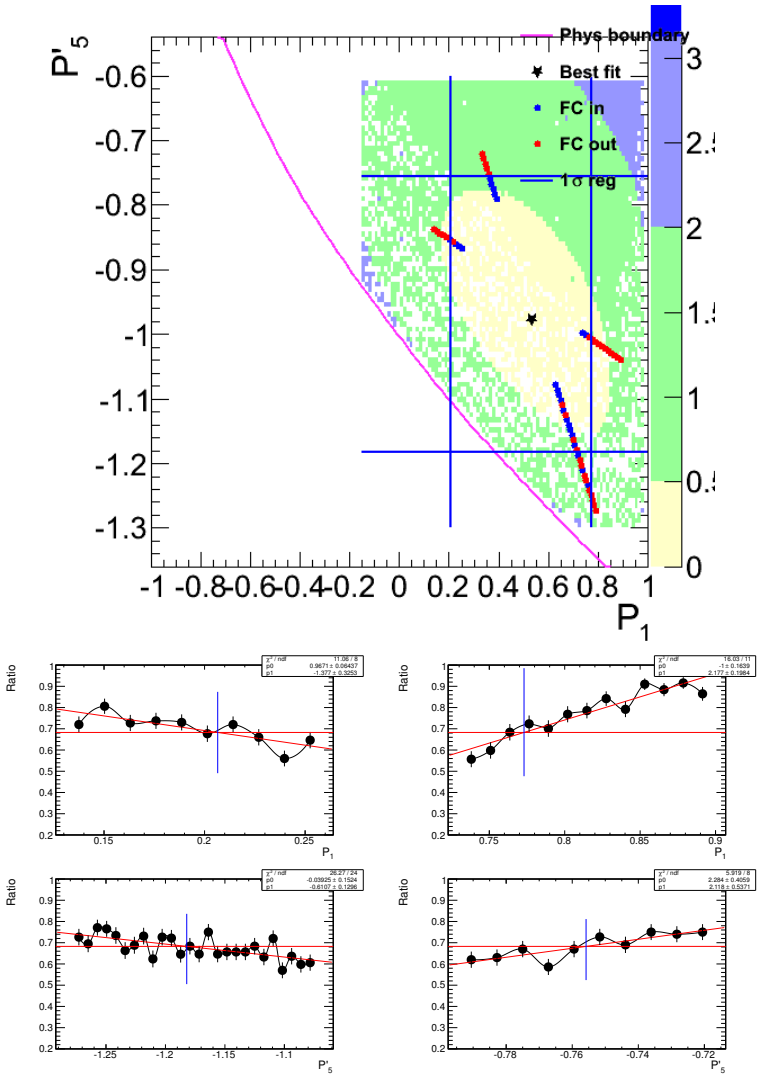


Figure 8.6: F-C results for bin 2. Scan of the GEN-points, superimposed to data $\mathcal{L}$ distribution: red points are outside the 68% CL region, the blue ones are inside. The blue lines defines the $\pm 1\sigma$ region (top). Ratio distribution as a function of P_1 for lower and upper bounds, with linear fit and 68% horizontal line (middle left, right). Likewise for P_5' (bottom).

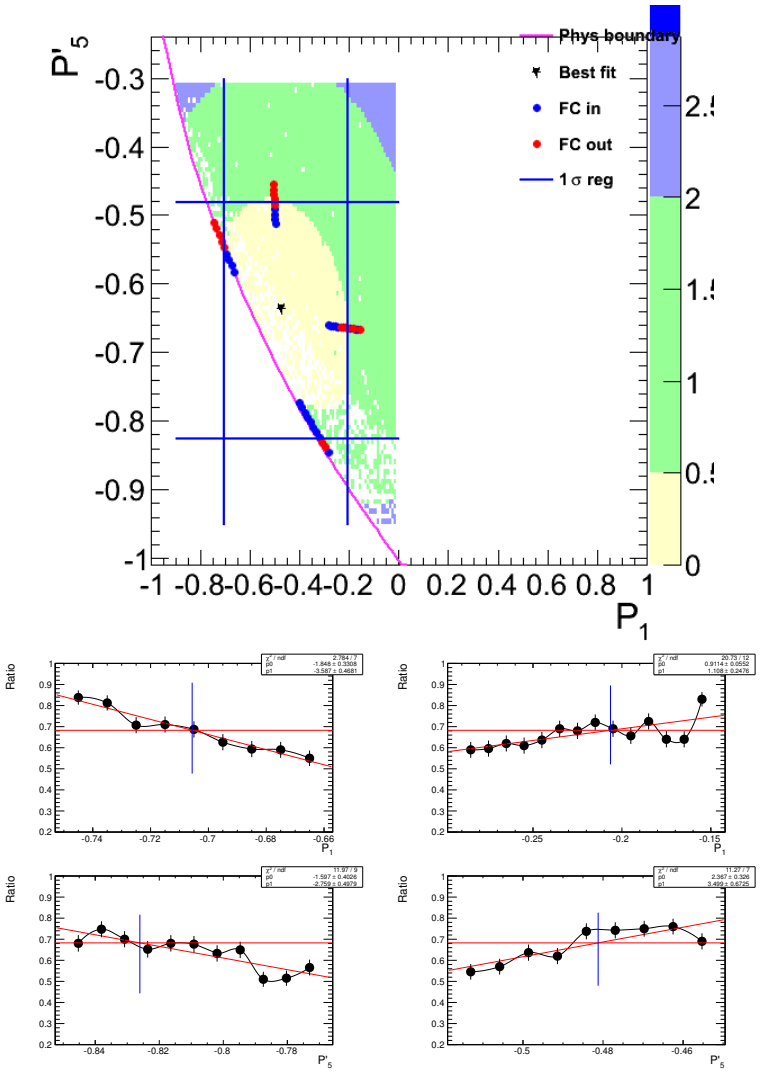


Figure 8.7: F-C results for bin 3. Scan of the GEN-points, superimposed to data $\mathcal{L}$ distribution: red points are outside the 68% CL region, the blue ones are inside. The blue lines defines the $\pm 1\sigma$ region (top). Ratio distribution as a function of P_1 for lower and upper bounds, with linear fit and 68% horizontal line (middle left, right). Likewise for P_5' (bottom).

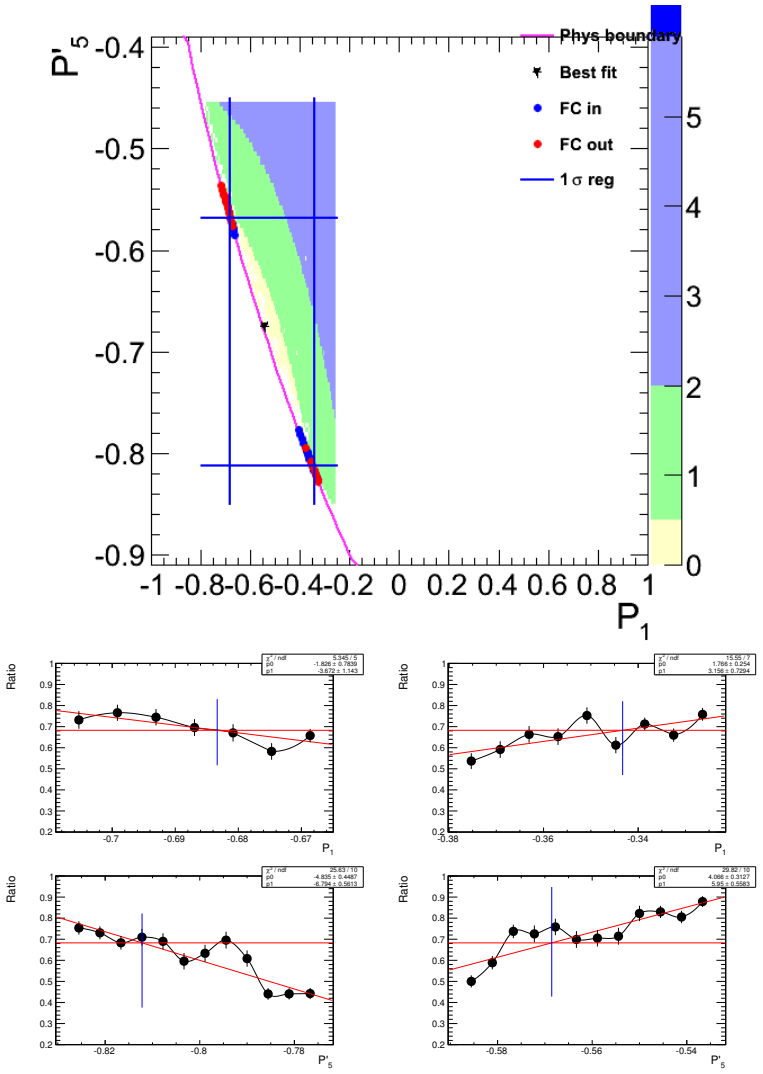


Figure 8.8: F-C results for bin 5. Scan of the GEN-points, superimposed to data $\mathcal{L}$ distribution: red points are outside the 68% CL region, the blue ones are inside. The blue lines defines the $\pm 1\sigma$ region (top). Ratio distribution as a function of P_1 for lower and upper bounds, with linear fit and 68% horizontal line (middle left, right). Likewise for P_5' (bottom).

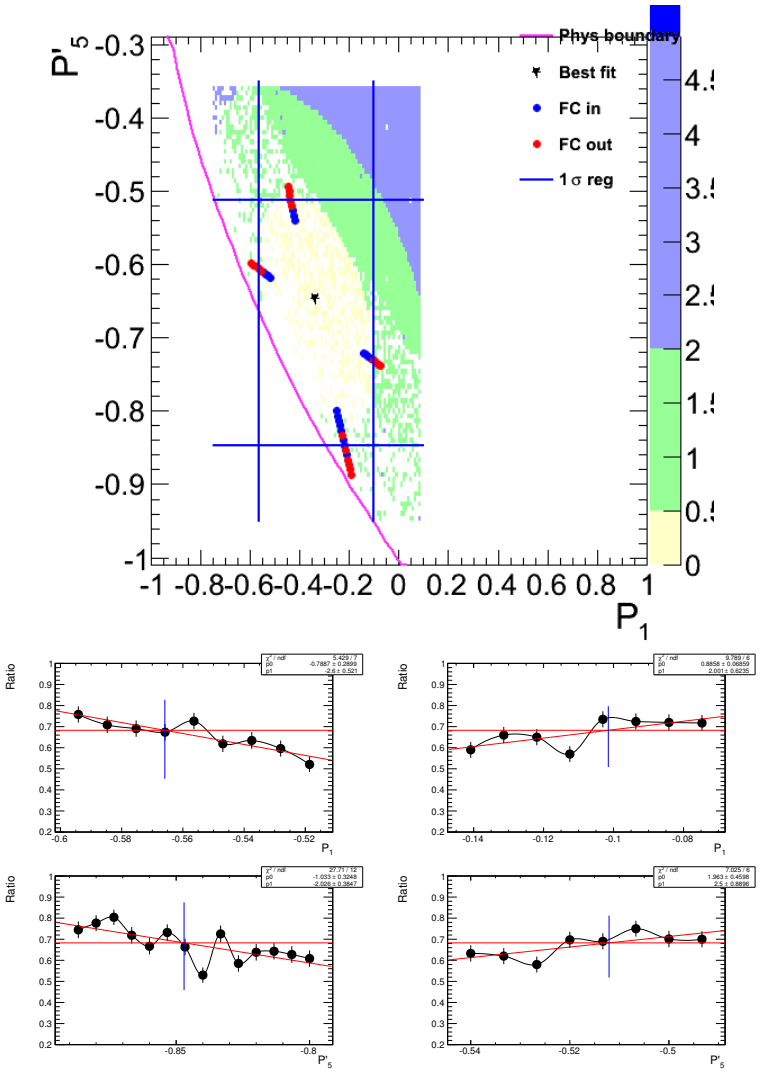


Figure 8.9: F-C results for bin 7. Scan of the GEN-points, superimposed to data $\mathcal{L}$ distribution: red points are outside the 68% CL region, the blue ones are inside. The blue lines defines the $\pm 1\sigma$ region (top). Ratio distribution as a function of P_1 for lower and upper bounds, with linear fit and 68% horizontal line (middle left, right). Likewise for P_5' (bottom).

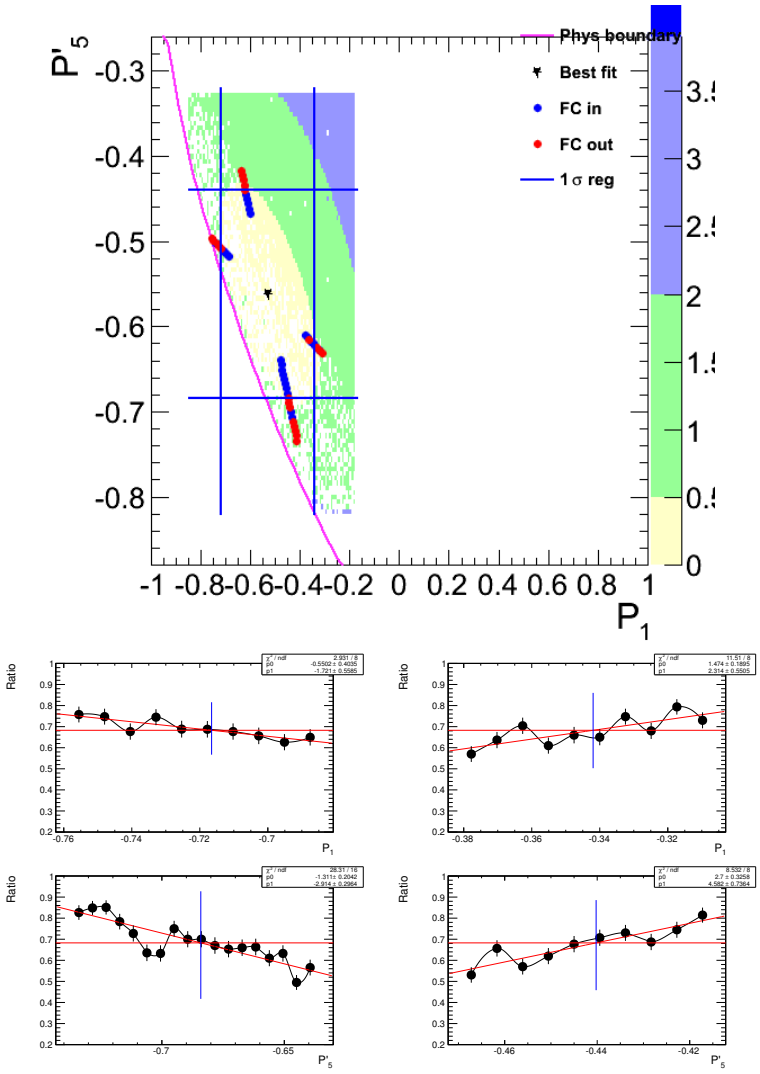


Figure 8.10: F-C results for bin 8. Scan of the GEN-points, superimposed to data $\mathcal{L}$ distribution: red points are outside the 68% CL region, the blue ones are inside. The blue lines defines the $\pm 1\sigma$ region (top). Ratio distribution as a function of P_1 for lower and upper bounds, with linear fit and 68% horizontal line (middle left, right). Likewise for P_5' (bottom).

8.1.2 Correlation coefficient

From the distribution of the likelihood, in Figure 8.1, it is also possible to get the correlation coefficient between the two measured parameters, P_1, P_5'. The $\mathcal{L}$ in the P_1, P_5' plane is fitted with a bivariate Gaussian, with the following expression:

$$f(x,y) = \frac{\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2 - 2\rho\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right) + \left(\frac{y-\mu_y}{\sigma_y}\right)^2\right]\right\}}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \tag{8.1}$$

where $x = P_1$, $y = P_5'$ etc. The ρ parameter in Equation 8.1 is the correlation coefficient among the two parameters and it is reported in Table 8.2 for all the bins.

Table 8.2: Correlation coefficient between P_1 and P_5' in different q^2 bins.

Bin	ρ
0	-0.0526
1	-0.0452
2	$+0.4715$
3	$+0.0761$
5	$+0.6077$
7	$+0.4188$
8	$+0.4621$

8.2 Results of central values

The results of the fit on data are summarised in Table 8.3. The table presents results for all five parameters that were floating in the fit, both the analysis targets, $P1$ and P_5', and the nuisance, Y_S^C, Y_B, and $A_{\rm S}^5$. The errors reported for $P1$ and P_5' are obtained from the FC procedure, while the errors returned by MINOS are quoted for the other three parameters.

The results for $P1$, P_5' and $As5$ are shown in Fig 8.11, Fig 8.12, Fig 8.13. On the plots, the predictions from two theoretical groups are also shown. The orange band shows the predictions from Matias et al [40]. and the pink band shows the predictions from Paul et al [27]. Both predictions are adapted to our q^2 binning scheme.

We also put the latest LHCb results [4] on the plots for comparison, also with only statistical errors.

The results of detailed distributions in each q^2 bin, together with the projections of the `pdf`, are shown in the following figures from Fig. 8.14 to Fig. 8.20.

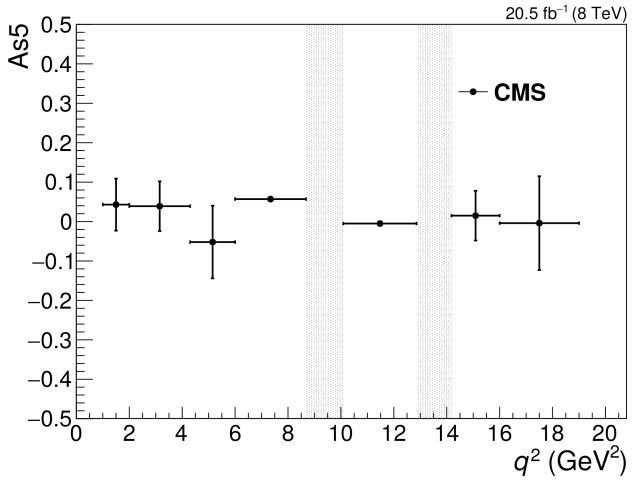


Figure 8.11: Fit results of the A_S^5 angular parameter versus q^2. Only the statistical uncertainties are shown. The vertical shaded regions correspond to the J/ψ and ψ' resonances.

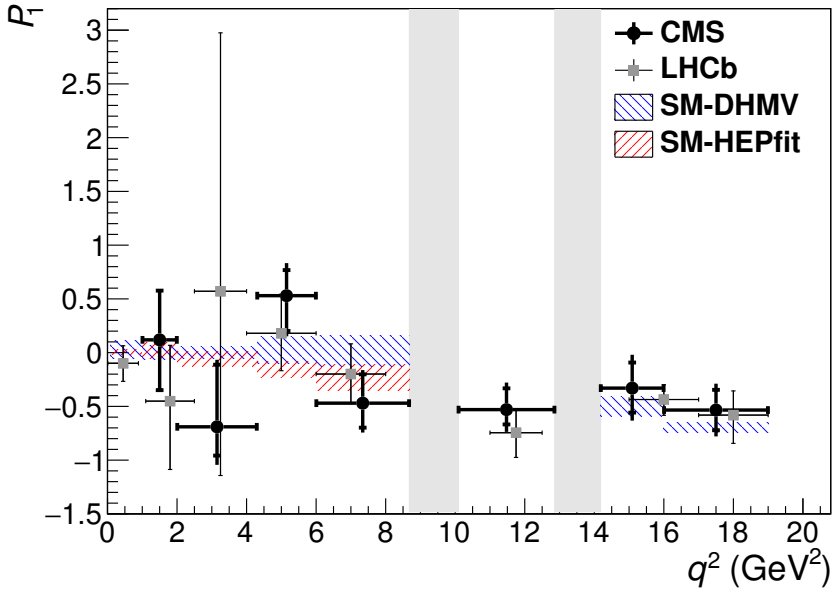


Figure 8.12: CMS measurements of the P_1 angular parameter versus q^2 for $B^0 \to K^{*0}\mu^+\mu^-$ decays, in comparison to results from the LHCb [5] Collaboration. The statistical uncertainties are shown by the inner vertical bars, while the outer vertical bars give the total uncertainties. The horizontal bars show the bin widths. The vertical shaded regions correspond to the J/ψ and ψ' resonances. The hatched regions show the predictions from two SM calculations described in the text, averaged over each q^2 bin.

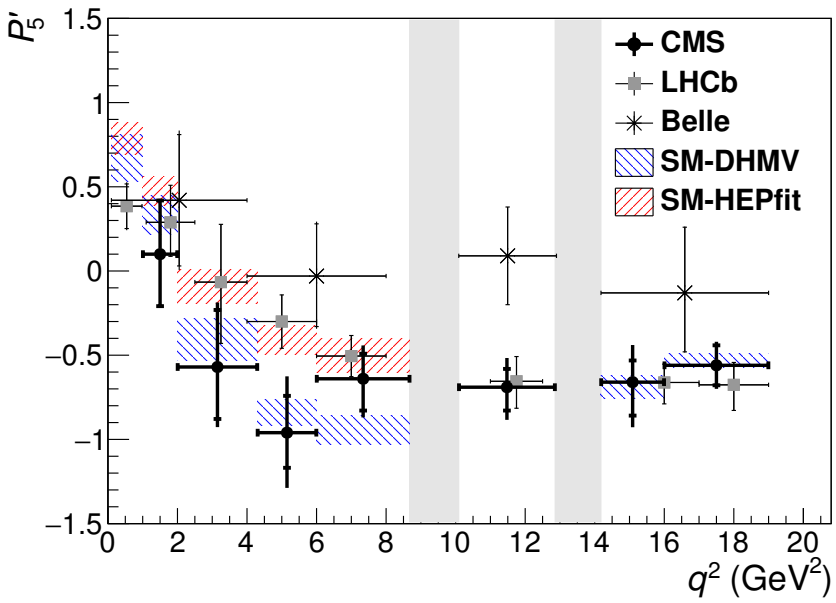


Figure 8.13: CMS measurements of the P_5' angular parameter versus q^2 for $B^0 \to K^{*0}\mu^+\mu^-$ decays, in comparison to results from the LHCb [5] and Belle [10] Collaborations. The statistical uncertainties are shown by the inner vertical bars, while the outer vertical bars give the total uncertainties. The horizontal bars show the bin widths. The vertical shaded regions correspond to the J/ψ and ψ' resonances. The hatched regions show the predictions from two SM calculations described in the text, averaged over each q^2 bin.

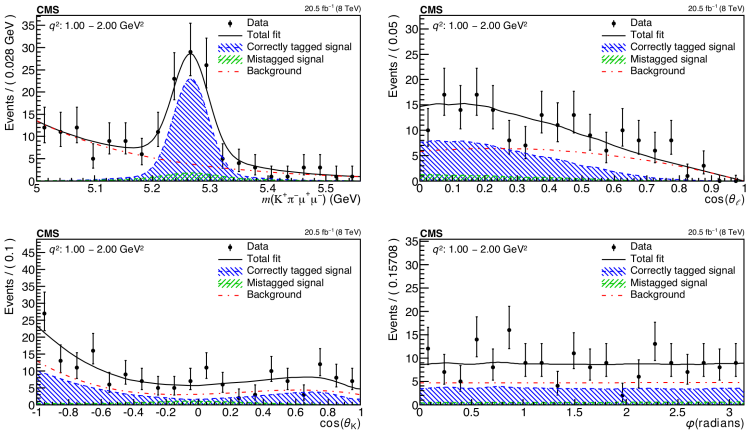


Figure 8.14: Data distributions (black points) and the projections of the fitted `pdf` (black curves), of its signal right-tagged component (blue curves), of its signal mis-tagged component (green curves), and of its background component (red curves), for q^2 bin 0. The data distribution and `pdf` projections are shown as functions of B^0 invariant mass, $\cos\theta_l$, $\cos\theta_K$, and ϕ.

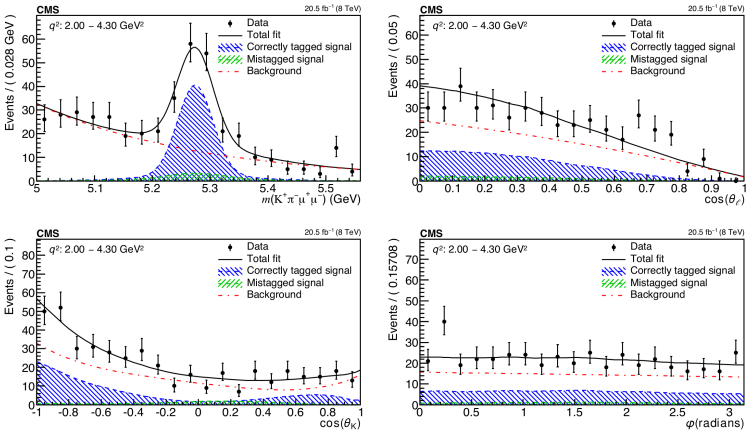


Figure 8.15: Data distributions (black points) and the projections of the fitted `pdf` (black curves), of its signal right-tagged component (blue curves), of its signal mis-tagged component (green curves), and of its background component (red curves), for q^2 bin 1. The data distribution and `pdf` projections are shown as functions of B^0 invariant mass, $\cos\theta_l$, $\cos\theta_K$, and ϕ.

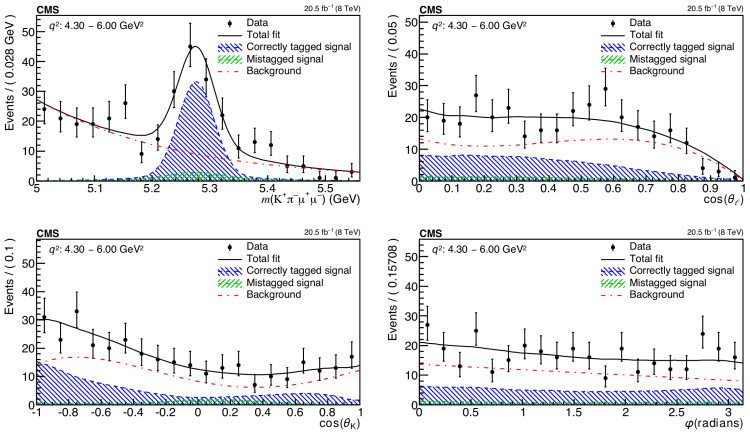


Figure 8.16: Data distributions (black points) and the projections of the fitted `pdf` (black curves), of its signal right-tagged component (blue curves), of its signal mis-tagged component (green curves), and of its background component (red curves), for q^2 bin 2. The data distribution and `pdf` projections are shown as functions of B^0 invariant mass, $\cos\theta_l$, $\cos\theta_K$, and ϕ.

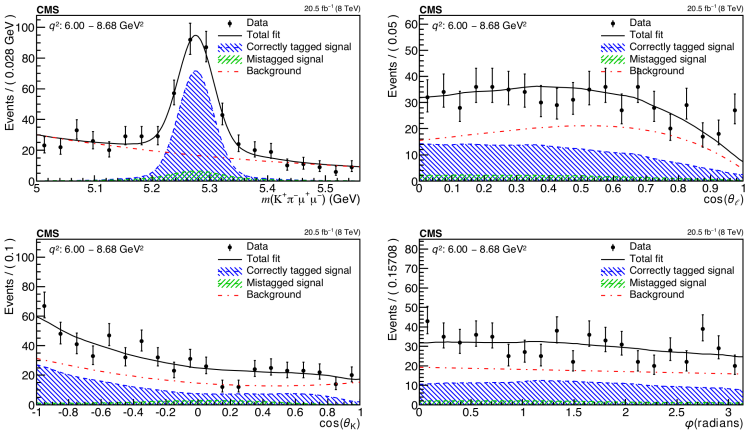


Figure 8.17: Data distributions (black points) and the projections of the fitted `pdf` (black curves), of its signal right-tagged component (blue curves), of its signal mis-tagged component (green curves), and of its background component (red curves), for q^2 bin 3. The data distribution and `pdf` projections are shown as functions of B^0 invariant mass, $\cos\theta_l$, $\cos\theta_K$, and ϕ.

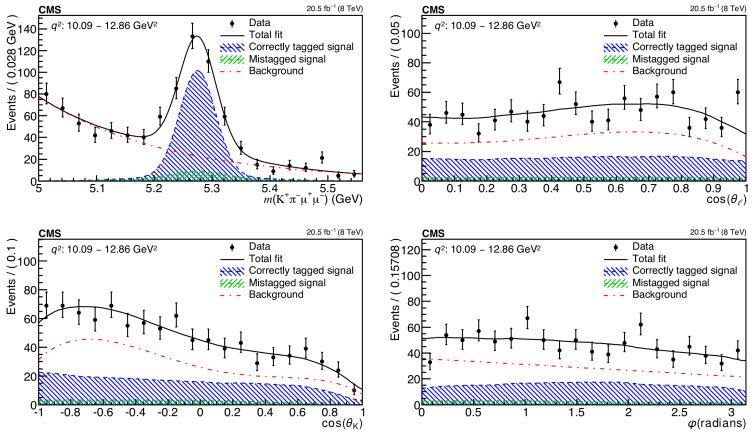


Figure 8.18: Data distributions (black points) and the projections of the fitted `pdf` (black curves), of its signal right-tagged component (blue curves), of its signal mis-tagged component (green curves), and of its background component (red curves), for q^2 bin 5. The data distribution and `pdf` projections are shown as functions of B^0 invariant mass, $\cos\theta_l$, $\cos\theta_K$, and ϕ.

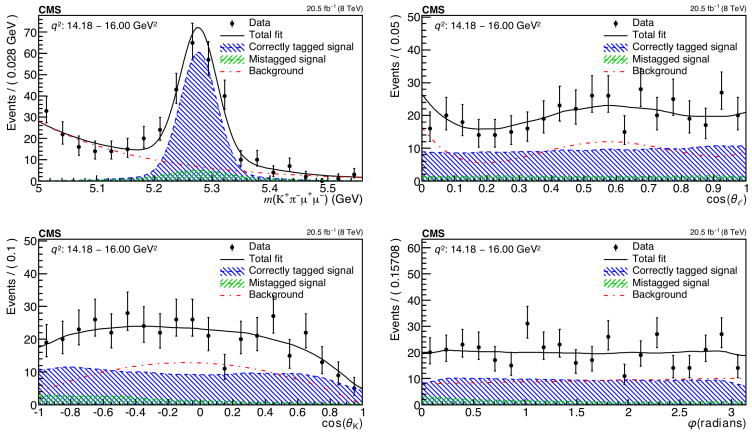


Figure 8.19: Data distributions (black points) and the projections of the fitted `pdf` (black curves), of its signal right-tagged component (blue curves), of its signal mis-tagged component (green curves), and of its background component (red curves), for q^2 bin 7. The data distribution and `pdf` projections are shown as functions of B^0 invariant mass, $\cos\theta_l$, $\cos\theta_K$, and ϕ.

Table 8.3: The measured values of signal yield Y_S^C, background yield Y_B, A_S^5, $P1$, and $P5'$ for the decay $B^0 \to K^{*0}(\to K^+\pi^-)\mu\mu$ in bins of q^2. The first uncertainty is statistical and the second (when present) is systematic.

q^2 (GeV^2)	Y_S^C	Y_B	A_s^5	P_1	P_5'
1.00–2.00	80 ± 12	95 ± 11	0.043 ± 0.066	$0.119^{+0.46}_{-0.47} \pm 0.058$	$0.101^{+0.32}_{-0.31} \pm 0.116$
2.00–4.30	145 ± 16	290 ± 20	0.039 ± 0.063	$-0.685^{+0.58}_{-0.27} \pm 0.088$	$-0.567^{+0.34}_{-0.31} \pm 0.153$
4.30–6.00	119 ± 14	216 ± 17	-0.052 ± 0.092	$0.533^{+0.24}_{-0.33} \pm 0.175$	$-0.957^{+0.22}_{-0.21} \pm 0.161$
6.00–8.68	247 ± 21	351 ± 23	0.057 ± 0.005	$-0.470^{+0.27}_{-0.23} \pm 0.131$	$-0.643^{+0.15}_{-0.19} \pm 0.138$
10.09–12.86	354 ± 23	575 ± 1	-0.005 ± 0.008	$-0.531^{+0.2}_{-0.14} \pm 0.215$	$-0.690^{+0.11}_{-0.14} \pm 0.246$
14.18–16.00	213 ± 17	185 ± 16	0.015 ± 0.063	$-0.329^{+0.24}_{-0.23} \pm 0.245$	$-0.664^{+0.13}_{-0.2} \pm 0.188$
16.00–19.00	239 ± 19	82 ± 0	-0.004 ± 0.119	$-0.533^{+0.19}_{-0.19} \pm 0.131$	$-0.559^{+0.12}_{-0.12} \pm 0.072$

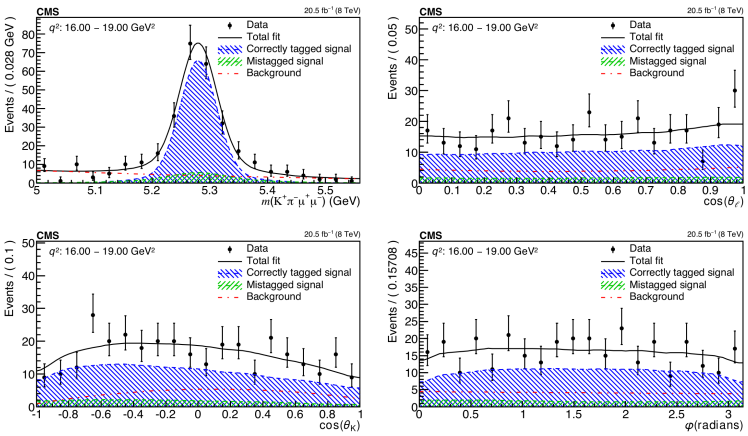


Figure 8.20: Data distributions (black points) and the projections of the fitted `pdf` (black curves), of its signal right-tagged component (blue curves), of its signal mis-tagged component (green curves), and of its background component (red curves), for q^2 bin 8. The data distribution and `pdf` projections are shown as functions of B^0 invariant mass, $\cos\theta_l$, $\cos\theta_K$, and ϕ.

8.2.1 Validation of the yield results

The yield of signal and background events in each q^2 bin is compared with the previous CMS analysis, which used the same dataset, on Table 8.4.

Table 8.4: Comparison of values of signal yield Y_S^C and background yield Y_B with the same values found in previous CMS analysis.

	Y_S^C		Δ	Y_B		Δ
q^2 (GeV^2)	this analysis	previous CMS		this analysis	previous CMS	
1.00–2.00	80 ± 12	84 ± 11	-4 ± 16	95 ± 11	91 ± 12	4 ± 16
2.00–4.30	145 ± 16	145 ± 16	0 ± 23	290 ± 20	289 ± 20	1 ± 28
4.30–6.00	119 ± 14	117 ± 15	2 ± 21	216 ± 17	218 ± 18	-2 ± 25
6.00–8.68	247 ± 21	254 ± 21	-7 ± 30	351 ± 23	344 ± 23	7 ± 33
10.09–12.86	354 ± 23	362 ± 25	-8 ± 34	575 ± 28	567 ± 29	8 ± 40
14.18–16.00	213 ± 17	225 ± 18	-12 ± 25	185 ± 16	175 ± 17	10 ± 23
16.00–19.00	239 ± 19	239 ± 18	0 ± 26	82 ± 12	82 ± 12	0 ± 17

Chapter 9

Future perspective and conclusions

9.1 Perspective at LHC Run II

The puzzle originated by the discrepancy observed by the LHCb collaboration in the P_5' angular parameter is far to be resolved, both from the theoretical side, where a better understanding of the effects of the hadronic uncertainties is essential to have a reliable Standard Model prediction, and from the experimental side.

In particular regarding this latter aspect, in the last few years the analyses from many experiments added their contribution to the LHCb result, to validate it and, by combining the results, improve the precision of the experimental measurement.

Despite the large statistics collected during LHC Run I by CMS, LHCb and ATLAS experiments, the uncertainty of the results of these analyses is still statistically dominated. This makes this angular analysis an hot topic for the near future.

Since 2015 LHC Run II is ongoing, and the machine is showing very good performance, delivering an instantaneous luminosity that is increasing year-by-year. In particular for CMS, the integrated luminosity expected to be collected in the full Run II period, ranging from 2015 to 2018, is $150\,\text{fb}^{-1}$, which is a huge value if compared to the statistics collected in Run I, considering also that the B^0 cross section increases by a factor of about two, thanks to the higher centre-of-mass energy of the collisions: $\sqrt{s} = 13\,\text{TeV}$.

Performing the analysis on Run II data can give a large improvement to the experimental precision reached in the measurement of the angular parameters. In this section I will describe some new issues that an analysis on CMS Run-II data should face, and some upgrades that could improve the results and make them more robust.

9.1.1 Trigger developments

The main negative aspect, at analysis level, of the increased LHC luminosity is the trigger selections. Due to the limited availability of computing resources, the event reconstruction can be run only on a certain amount of data. Since it is pointless to collect data if they can never be reconstructed, the limited number of reconstructed events can be translated in an upper limit of the rate of events that pass the HLT system. This limit, in the first three years of Run II, is about 1000 Hz, for the data stream dedicated to physics, and it is calculated taking into account the LHC down-times, when the detector is not collecting data but the reconstruction process continues.

In addition, the maximum frequency at which the silicon detectors can be read sets a limit in the L1T output rate to 100 kHz.

When LHC increases the instantaneous luminosity of the collisions, the rate would increase as well, because the probability of having a proton-proton interaction with final state that fires the trigger is proportional to the luminosity. The situation is even worse when the average pileup is increased, because some HLT algorithms, and most of the L1T ones, can sum up the contribution of final states originated in different proton-proton interactions. The firing probability in this case increases more than linearly, as a function of the luminosity. Due to the rate upper limit, when the delivered luminosity increases, we need to set tighter requirements in the trigger algorithms.

Focusing now on the trigger selection used by the $B^0 \to K^{*0}\mu^+\mu^-$ angular analysis, in the 2012 run there was a simple requirement on the presence of two muons with $p_T > 3.5\,\text{GeV}$ and $|\eta| < 2.2$, and forming a common displaced vertex with some quality requirements, as described in details in Section 3.1. Since a simple increase of the p_T and η threshold, to keep the rate stable, would have led to a large drop in the signal efficiency, many studies have been performed in the selection optimisation, to achieve a sufficient rate reduction without affecting largely the signal efficiency.

2015 and 2016 trigger

The L1T seed used during 2015 and 2016 runs is not containing any specific improvement, but only tighter thresholds are applied on the two muons. Two algorithms are contributing to the trigger:

- a seed with no cuts on the muon transverse momenta, and a cut on their pseudorapidity $|\eta_{L1\mu}| < 1.4\,(1.6)$. The two thresholds were used for collisions with higher (1.4) and lower (1.6) values of instantaneous luminosity;

- a seed with no cuts on the pseudorapidity, and a cut on muon transverse momenta. During collisions at higher (lower) instantaneous lu-

minosity, the cut on the leading muon is $p_{T\,L1\mu} > 12\,(11)\,\text{GeV}$, and the cut on the second muon is $p_{T\,L1\mu} > 5\,(4)\,\text{GeV}$.

The HLT path used has a raised requirement on the transverse momenta of the two muons, with respect to the 2012 version, $p_{T\,L3\mu} > 4\,\text{GeV}$. There are not cuts on the muon pseudorapidity, while the vertexing requirements are still present. In addition, it constraints the requirement of an additional hadronic track, with $p_{T\,h} > 0.8\,\text{GeV}$ and forming a common vertex with the muon pair. Thanks to the latter requirement, the trigger manages to deal with the higher luminosity values reached without needing to increase dramatically the cuts on p_T.

A preliminary and underestimated prevision on the number of signal events expected in 2016 dataset can be obtained by applying the same offline selection criteria used in Run 1 analysis to the new data sample, and adding the matching requirement between the hadronic track firing the trigger and one of the two tracks used to build the candidate at reconstructed-level. The candidate-B^0 invariant mass distribution of the selected sample, obtained from the full 2016 dataset, is shown in Figure 9.1, for the signal and J/ψ q^2 regions. Note that the selection cuts used were optimised on the 2012 data and re-optimising them on the Run 2 data would improve the signal over background ratio. In addition, the analysis upgrades described in Section 9.1.2 are expected to further improve it.

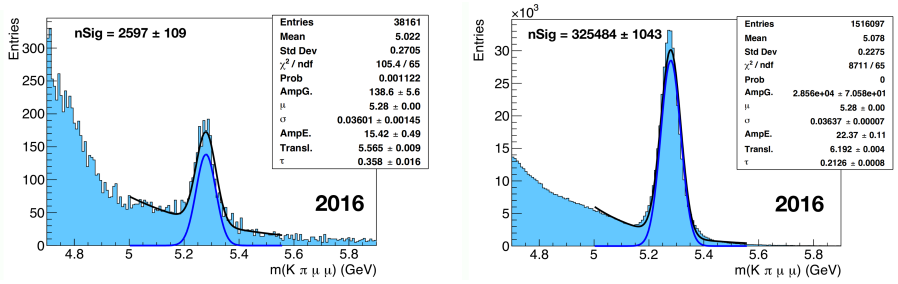


Figure 9.1: Distribution of the candidate-B^0 invariant mass of the 2016 dataset, for the signal (left) and J/ψ (right) q^2 regions. The selection criteria applied to the dataset are described in Section 9.1.1. The rough fits are performed with a `pdf` obtained from the sum of a Gaussian and an exponential function.

2017 trigger

The L1T firmware used during 2017 run allowed to exploit the correlations between pairs of objects, giving the possibility to cut on di-object quantities, like invariant mass and $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$. The two L1T seeds used for the $B^0 \to K^{*0}\mu^+\mu^-$ channel has been developed to contain cuts on the distance between the two muons:

- a seed with no cut on the muon transverse momenta, used a stable cut on their pseudorapidity $|\eta_{L1\mu}| < 1.5$ and a cut $\Delta R_{\mu\mu} < 1.4$;
- a seed with no cut on the pseudorapidity, and a cut on muon transverse momenta, $p_{T\,L1\mu} > 4\,\text{GeV}$ for both the leading and second muons, and a cut on the dimuon distance, $\Delta R_{\mu\mu} < 1.2$.

In addition the quality requirement on the two muons has been increased, with respect to 2016 version, from an intermediate quality to the highest quality possible. These new cuts on the dimuon distance and quality allows to keep looser thresholds on muon transverse momenta and pseudorapidity, even if the instantaneous luminosity has been increased with respect to 2016 run.

The HLT path used in 2017 is very similar to the 2016 version. The only difference, introduced to keep the rate at affordable values, is that the cut on hadronic track transverse momentum is increased, $p_{T\,h} > 1.2\,\text{GeV}$, and a new requirement has been introduced on the significance of the hadronic track impact parameter on the transverse plane with respect to the beamspot, $d_{xy}/\sigma(d_{xy}) > 2$. The latter cut is identical to the one applied at reconstructed level in the Run 1 analysis, as described in Section 3.2.

The same selections described in the previous section and applied to the 2016 dataset, are now applied to a partial 2017 dataset, containing the events corresponding to an integrated luminosity of $28\,\text{fb}^{-1}$. The candidate-B^0 invariant mass distribution of the selected sample is shown in Figure 9.2, for the signal and J/ψ q^2 regions.

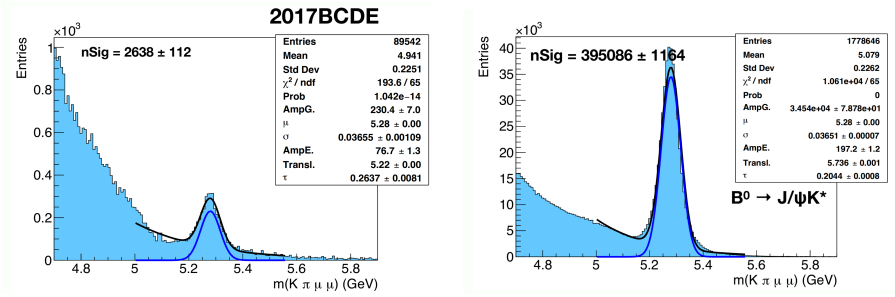


Figure 9.2: Distribution of the candidate-B^0 invariant mass of a partial 2017 dataset, corresponding to $28\,\text{fb}^{-1}$, for the signal (left) and J/ψ (right) q^2 regions. The selection criteria applied to the dataset are described in Section 9.1.1. The rough fits are performed with a `pdf` obtained from the sum of a Gaussian and an exponential function.

9.1.2 Analysis upgrades

On the analysis methods, there are many aspects that can be improved with respect to the version presented in this thesis.

The first upgrade will aim to extend the analysis to measure the full set of angular parameters present in the decay rate. An increased number of events in the dataset and a better or equal signal-to-noise fraction will be crucial for this purpose. Also a better handling of the physical boundary can help with improving the fit stability, needed to extend the measurement.

A second goal is to improve the fit performances, both in terms of time consumption, for which parallel GPU-based computing can be used, and in term of stability, for which a simpler efficiency parameterisation can be tested.

Furthermore, a great improvement to the signal-to-noise ratio can be achieved by using a selection based on multi-variate analysis, like a Neural Network. This techniques would allow to maximise the background rejection by a full exploitation of the information contained in the variables.

9.2 Summary

In this thesis, I have presented an important result of the CMS Collaboration in the Flavour Physics sector. The angular analysis of the $B^0 \to K^{*0}\mu^+\mu^-$ decay has been performed with the data collected by the CMS Experiment in the 2012 run of pp collisions at $\sqrt{s} = 8\,\text{TeV}$, corresponding to an integrated luminosity of $20.5\,\text{fb}^{-1}$.

After presenting a general status of the theoretical description of this analysis and describing the LHC machine and the CMS detector, the details of the analysis are reported.

Firstly, the selection criteria applied to the collected data, and the parameterisation of their efficiency and of the detector acceptance, evaluated on signal simulated samples, have been described.

A complex fitting algorithm has been set up, to extract in a stable and reliable way the P_1 and P_5' parameters from the distributions of the K^+ π^- μ^+ μ^- invariant mass and of the three angular variables. This fitting algorithm has been validated in many ways, by testing it on MC samples and on data control channels, $B^0 \to J/\psi(\mu^+\mu^-)K^{*0}$ and $B^0 \to \psi'(\mu^+\mu^-)K^{*0}$.

In order to make the results as robust as possible, many sources of possible systematic uncertainty has been studied. The statistical uncertainties have been evaluated using a simplified form of a bi-dimensional Feldman-Cousins approach, to guarantee the correct coverage even when the result is close to an nonphysical region in the parameter phase-space.

Finally, the fit procedure has been applied on data, and the results extracted. Currently, they are among the most precise measurements of these parameters, they are compatible with the results from the other experiments and they show no discrepancies to the Standard Model predictions.

Bibliography

[1] G. *et al.* (ATLAS Collab.) Aad. *Journal of Instrumentation*, 3:8003, August 2008.

[2] R. Aaij et al. Differential branching fraction and angular analysis of the decay $B^0 \to K^{*0}\mu^+\mu^-$. *JHEP*, 08:131, 2013.

[3] R Aaij et al. Measurement of Form-Factor-Independent Observables in the Decay $B^0 \to K^{*0}\mu^+\mu^-$. *Phys. Rev. Lett.*, 111:191801, 2013.

[4] R. Aaij et al. Angular analysis of the $B^0 \to K^*\mu^+\mu^-$ decay using 3 fb^{-1} of integrated luminosity. *JHEP*, 02:104, 2016.

[5] R. Aaij et al. Angular analysis of the $B^0 \to K^*\mu^+\mu^-$ decay using 3 fb^{-1} of integrated luminosity. *JHEP*, 02:104, 2016.

[6] R. Aaij et al. Test of lepton universality with $B^0 \to K^{*0}\ell^+\ell^-$ decays. *JHEP*, 08:055, 2017.

[7] Roel Aaij et al. Angular analysis of the $B^0 \to K^{*0}\mu^+\mu^-$ decay using 3 fb^{-1} of integrated luminosity. *JHEP*, 02:104, 2016.

[8] T. Aaltonen et al. Measurements of the Angular Distributions in the Decays $B \to K^{(*)}\mu^+\mu^-$ at CDF. *Phys. Rev. Lett.*, 108:081807, 2012.

[9] K. *et al.* (ALICE Collab.) Aamodt. *Journal of Instrumentation*, 3:8002, August 2008.

[10] A. Abdesselam et al. Lepton-flavor-dependent angular analysis of $B \to K^*\ell^+\ell^-$. *Phys. Rev. Lett.*, 118:111801, 2017.

[11] W. *et al.* (CMS Collab.) Adam. Track Reconstruction in the CMS tracker. Technical Report CMS-NOTE-2006-041, CERN, Geneva, Dec 2006.

[12] Wolfgang Altmannshofer, Patricia Ball, Aoife Bharucha, Andrzej J. Buras, David M. Straub, and Michael Wick. Symmetries and Asymmetries of $B \to K^*\mu^+\mu^-$ Decays in the Standard Model and Beyond. *JHEP*, 01:019, 2009.

[13] A. A. *et al.* (LHCb Collab.) Alves, Jr. *Journal of Instrumentation*, 3:8005, August 2008.

[14] Bernard Aubert et al. Angular Distributions in the Decays B —¿ K* l+ l-. *Phys. Rev.*, D79:031102, 2009.

[15] B. Aubert et al. Time-integrated and time-dependent angular analyses of $B \to J/\psi K\pi$: A measurement of $cos2\beta$ with no sign ambiguity from strong phases. *Phys. Rev. D*, 71:032005, 2005.

[16] G. L. *et al.* (CMS Collab.) Bayatian. *CMS Physics: Technical Design Report Volume 1: Detector Performance and Software*. Technical Design Report CMS. CERN, Geneva, 2006.

[17] G. L. *et al.* (CMS Collab.) Bayatyan. *CMS TriDAS project: Technical Design Report, Volume 1: The Trigger Systems*. Technical Design Report CMS.

[18] M. Beneke and T. Feldmann. Symmetry breaking corrections to heavy to light B meson form-factors at large recoil. *Nucl. Phys.*, B592:3–34, 2001.

[19] M. Beneke, T. Feldmann, and D. Seidel. Systematic approach to exclusive $B \to V l^+ l^-$, $V\gamma$ decays. *Nucl. Phys.*, B612:25–58, 2001.

[20] M. Beneke, Th. Feldmann, and D. Seidel. Exclusive radiative and electroweak $b \to d$ and $b \to s$ penguin decays at NLO. *Eur. Phys. J.*, C41:173–188, 2005.

[21] Christoph Bobeth, Mikolaj Misiak, and Jorg Urban. Photonic penguins at two loops and m_t dependence of $BR[B \to X_s l^+ l^-]$. *Nucl. Phys.*, B574:291–330, 2000.

[22] O. S. *et al.* Bruning. 2004.

[23] S. *et al.* (CMS Collab.) Chatrchyan. *Journal of Instrumentation*, 3:8004, August 2008.

[24] Serguei Chatrchyan et al. Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV. *JINST*, 7:P10002, 2012.

[25] Serguei Chatrchyan et al. Angular analysis of the decay $B^0 \to K^{*0}\mu^+\mu^-$ from pp collisions at $\sqrt{s} = 8$ TeV. *Phys. Lett. B*, 753:424, 2016.

[26] S. *et al.* Cittolin. *CMS The TriDAS Project: Technical Design Report, Volume 2: Data Acquisition and High-Level Trigger. CMS trigger and data-acquisition project*. Technical Design Report CMS. CERN, Geneva, 2002.

[27] M. Ciuchini, M. Fedele, E. Franco, S. Mishimad, A. Paul, L. Silvestrini, and M. Valli. $B \to K^* \ell^+ \ell^-$ decays at large recoil in the Standard Model: a theoretical reappraisal. *Submitted to JHEP*.

[28] CMS Collab. *The CMS electromagnetic calorimeter project: Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997.

[29] CMS Collab. *The CMS hadron calorimeter project: Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997.

[30] CMS Collab. *The CMS magnet project: Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997.

[31] CMS Collab. *The CMS muon project: Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997.

[32] Kyle S. Cranmer. Kernel estimation in high-energy physics. *Comput. Phys. Commun.*, 136:198–207, 2001.

[33] F. James, and M. Winklerc. Minuit user's guide. http://seal.web.cern.ch/seal/work-packages/mathlibs/minuit/doc/doc.html, 2004.

[34] G. J. Feldman and R. D. Cousins. Unified approach to the classical statistical analysis of small signals. *Phys. Rev. D*, 57:3873, 1998.

[35] V. *et al.* (CMS Collab.) Karimki. *The CMS tracker system project: Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997.

[36] Vardan Khachatryan et al. Angular analysis of the decay $\mathrm{B}^0 \to \mathrm{K}^{*0} \mu^+ \mu^-$ from pp collisions at $\sqrt{s} = 8$ TeV. 2015. Submitted to to Phys. Lett. B.

[37] Alexander Khodjamirian. QCD sum rules for heavy flavor physics. *AIP Conf. Proc.*, 602:194–205, 2001. [,194(2001)].

[38] Joaquim Matias and Nicola Serra. Symmetry relations between angular observables in $B^0 \to K^* \mu^+ \mu^-$ and the LHCb P_5' anomaly. *Phys. Rev.*, D90(3):034002, 2014.

[39] K.A. Olive. Review of particle physics, 2014-2015. *Chin. Phys. C*, 38:090001, 2014.

[40] S. Descotes-Genon, T. Hurth, J. Matias, and J. Virto. Optimizing the basis of $\mathrm{b} \to \mathrm{k}^* \ell^+ \ell^-$ observables in the full kinematic range. *JHEP*, 1305:137, 2013.

[41] David W. Scott. *Multivariate density estimation : theory, practice, and visualization.* Wiley series in probability and mathematical statistics : Applied probability and statistics section. Wiley-Interscience, New York, Chichester, Brisbane, 1992.

[42] Albert M Sirunyan et al. Measurement of angular parameters from the decay $B^0 \to K^{*0}\mu^+\mu^-$ in proton-proton collisions at $\sqrt{s} = 8$ TeV. 2017.

[43] T. Sjöstrand, S. Mrenna, and P. Skands. PYTHIA 6.4 physics and manual. 5:26, May 2006.

[44] T. *et al.* (CMS Collab.) Speer. Vertex Fitting in the CMS Tracker. Technical Report CMS-NOTE-2006-032, CERN, Geneva, Feb 2006.

[45] W. Verkerke and D. Kirkby. The RooFit toolkit for data modeling. *ArXiv Physics e-prints*, June 2003.

[46] J. T. Wei et al. Measurement of the Differential Branching Fraction and Forward-Backword Asymmetry for $B \to K^{(*)}\ell^+\ell^-$. *Phys. Rev. Lett.*, 103:171801, 2009.