Measurement of the $P_1$ and $P'_5$ angular parameters of the decay $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ in proton-proton collisions at $\sqrt{s} = 8$ TeV

The CMS Collaboration

Abstract

The angular distributions of the decay $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ are studied using data corresponding to an integrated luminosity of 20.5 fb$^{-1}$ collected with the CMS detector at the LHC in proton-proton collisions at $\sqrt{s} = 8$ TeV. An angular analysis gives access to many physical observables, including $P'_5$, which is of particular interest due to measurements from the LHCb Collaboration, that exhibit potential discrepancy with the standard model. Based on a sample of 1397 signal events, the $P_1$ and $P'_5$ angular parameters are determined as a function of the dimuon invariant mass squared. The measurement are in agreement with standard model predictions.
1 Introduction

Phenomena beyond the standard model (SM) of particle physics can become manifest directly, via the production of new particles, or indirectly, by affecting the production and decay of SM particles. Analyses of flavor-changing neutral current (FCNC) decays are particularly sensitive to the effects of new physics, since such decays are highly suppressed in the SM. The FCNC decay $B^0 \rightarrow K^{*0} \mu^+ \mu^-$, where $K^{*0}$ indicates the $K^+(892)^0$ meson and charge-conjugate states are implied for all particles unless stated, provides many opportunities to search for new phenomena. An angular analysis as a function of the dimuon invariant mass squared ($q^2$) allows the decay to be thoroughly investigated. New physics may modify any of the angular variables [1–17] relative to their SM values [1, 18–24]. While previous measurements of some of these quantities by the BaBar, Belle, CDF, LHCb, and CMS experiments are consistent with the SM [25–29], the LHCb and Belle Collaborations recently reported a discrepancy larger than 3 standard deviations ($\sigma$) with respect to the SM for the so-called $P'_5$ variable [30–32].

This Note presents a precise measurement of the $P'_5$ variable, together with the $P_1$ variable, using a sample of $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ events collected in proton-proton (pp) collisions at center-of-mass energy $\sqrt{s} = 8 \text{ TeV}$ with the CMS detector at the CERN LHC. The data correspond to an integrated luminosity of $20.5 \pm 0.5 \text{ fb}^{-1}$ [33]. The $K^{*0}$ is reconstructed through its decay to $K^+ \pi^-$, and the $B^0$ is reconstructed by fitting to a common vertex the tracks from two identified muon candidates and the tracks from the two hadrons in the $K^{*0}$ decay. The values of $P_1$ and $P'_5$ are measured by fitting the distribution of events as a function of the three angular variables: the angle between the positively charged muon and the $B^0$ in the dimuon rest frame, the angle between the $K^+$ and the $B^0$ in the $K^{*0}$ rest frame, and the angle between the dimuon and the $K\pi$ decay planes in the $B^0$ rest frame. All measurements are performed in $q^2$ bins from 1 to 19 GeV$^2$. The $q^2$ bins $8.68 < q^2 < 10.09 \text{ GeV}^2$ and $12.90 < q^2 < 14.18 \text{ GeV}^2$, corresponding to $B^0 \rightarrow J/\psi K^{*0}$ and $B^0 \rightarrow \psi' K^{*0}$ decays, respectively, are used to validate the analysis, where $\psi'$ denotes the $\psi(2S)$ meson.

2 CMS detector

A detailed description of the CMS detector, together with a definition of the coordinate system used and the standard kinematic variables, can be found in Ref. [34]. The main detector components used in this analysis are the silicon tracker and the muon detection systems. The silicon tracker, located in the 3.8 T field of a superconducting solenoid, consists of three pixel layers and ten strip layers (four of which have a stereo view) in the barrel region accompanied by similar endcap pixel and strip detectors on each side that extend coverage out to a pseudorapidity range of $|\eta| < 2.5$. For tracks with transverse momenta $1 < p_T < 10 \text{ GeV}$ and $|\eta| < 1.4$, the resolutions are typically 1.5% in $p_T$ and 25−90 (45−150) $\mu m$ in the transverse (longitudinal) impact parameter [35]. Muons are measured in the range $|\eta| < 2.4$ with detection planes made using three technologies: drift tubes, cathode strip chambers, and resistive plate chambers [36]. In addition to the tracker and muon detectors, CMS is equipped with electromagnetic and hadronic calorimeters that cover $|\eta| < 5$.

Events are selected using a two-level trigger system. The first level has specialized hardware processors that use information from the calorimeters and muon systems that select events of interest. A high-level trigger processor farm further decreases the event rate from around 90 kHz to around 400 Hz, before data storage.
3 Reconstruction, event selection, and efficiency

The criteria used to select the candidate events during data taking (trigger) and after full event reconstruction (offline) take advantage of the fact that $B^0$ mesons have relatively long lifetimes and therefore decay on average about 1 mm from their production point. The trigger only uses muons to select events, while the offline selection includes the full reconstruction of all decay products.

All events used in this analysis were recorded with the same trigger, requiring two identified muons of opposite charge to form a vertex that is displaced from the pp collision region (beamspot). The beamspot position (most probable collision point) and size (the extent of the luminous region covering 68% of the collisions in each dimension) were continuously measured through Gaussian fits to reconstructed vertices as part of the online data quality monitoring. The trigger required each muon to have $p_T > 3.5$ GeV, $|\eta| < 2.2$, and to pass within 2 cm of the beam axis. The dimuon system was required to have $p_T > 6.9$ GeV, a vertex fit $\chi^2$ probability larger than 10%, and a separation of the vertex relative to the beamspot in the transverse plane of at least 3$\sigma$, where $\sigma$ includes the calculated uncertainty in the vertex position and the measured size of the beamspot. In addition, the cosine of the angle, in the transverse plane, between the dimuon momentum vector and the vector from the beamspot to the dimuon vertex was required to be greater than 0.9.

The offline reconstruction requires two muons of opposite charge and two oppositely charged hadrons. The muons are required to match those that triggered the event readout, and also to pass general muon identification requirements. These include a track matched to at least one muon segment (collection of hits in a muon chamber consistent with the passage of a charged particle), a track fit $\chi^2$ per degree of freedom less than 1.8, hits in at least six tracker layers with at least two from the pixel detector, and a transverse (longitudinal) impact parameter with respect to the beamspot less than 3 (30) cm. The reconstructed dimuon system must also satisfy the same requirements that were applied in the trigger.

The hadron tracks are required to fail the muon identification criteria, have $p_T > 0.8$ GeV, and have an extrapolated distance of closest approach to the beamspot in the transverse plane greater than twice the sum in quadrature of the distance uncertainty and the beamspot transverse size. The two hadrons must have an invariant mass within 90 MeV of the nominal $K^{*0}$ mass [37] for either the $K^+\pi^-$ or $K^-\pi^+$ hypothesis. To remove contamination from $\phi(1020) \rightarrow K^+K^-$ decays, the invariant mass of the hadron pair must be greater than 1.035 GeV when the charged kaon mass is assigned to both hadrons. The $B^0$ candidates are obtained by fitting the four charged tracks to a common vertex, and applying a vertex constraint to improve the resolution of the track parameters. The $B^0$ candidates must have $p_T > 8$ GeV, $|\eta| < 2.2$, vertex fit $\chi^2$ probability larger than 10%, vertex transverse separation from the beamspot greater than 12 times the sum in quadrature of the separation uncertainty and the beamspot transverse size, and $\cos \alpha_{xy} > 0.9994$, where $\alpha_{xy}$ 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. The invariant mass $m$ of the $B^0$ candidate must also be within 280 MeV of the nominal $B^0$ mass $m_{B^0}$ [37] for either the $K^-\pi^+\mu^+\mu^-$ or $K^+\pi^-\mu^+\mu^-$ hypothesis. The selection criteria are optimized using signal event samples from simulation and background event samples from sideband data in the $B^0$ mass. After applying the selection criteria, events in which at least one candidate is found contain on average 1.05 candidates. A single candidate is chosen from each event based on the best $B^0$ vertex fit $\chi^2$.

From the selected data, the dimuon invariant mass $q$ and its calculated uncertainty $\sigma_q$ are used to distinguish signal events from events in the $B^0 \rightarrow J/\psi K^{*0}$ and $B^0 \rightarrow \psi' K^{*0}$ control samples,
where these latter samples are defined by $|q - m_{\psi K^{*0}}| < 3\sigma_q$ and $|q - m_{\psi'}| < 3\sigma_q$, respectively, where $m_{\psi K^{*0}}$ and $m_{\psi'}$ are the nominal masses [37] of the indicated meson. The average value of $\sigma_q$ is about 26 MeV. The signal sample is composed of the events that are not assigned to the $J/\psi$ and $\psi'$ control samples.

The signal sample at this point still contains contributions from $B^0 \to J/\psi K^{*0}$ and $B^0 \to \psi' K^{*0}$ events, mainly due to unreconstructed soft photons in the charmonium decay. These events have a low $q$ value and fall outside the selection described above. These events also have a low reconstructed $B^0$ mass value ($m$) and therefore they can be selectively removed using a combined selection on $q$ and $m$. For $q < m_{\psi K^{*0}}$ ($q > m_{\psi K^{*0}}$), we require $|(m - m_{B^0}) - (q - m_{\psi K^{*0}})| > 160 (60)$ MeV. For $q < m_{\psi'}$ ($q > m_{\psi'}$), we require $|(m - m_{B^0}) - (q - m_{\psi'})| > 60 (30)$ MeV. The requirements are such that less than 10% of the background events originate from the control channels. In total, 1397 signal event candidates are selected.

The four-track vertex candidate is identified as a $B^0$ or $\bar{B}^0$ depending on whether the $K^+\pi^-$ or $K^-\pi^+$ invariant mass is closest to the nominal $K^{*0}$ mass. The fraction of candidates assigned to the incorrect state is estimated from simulation to be 12–14%, depending on $q^2$.

The global efficiency, $\epsilon$, is the product of the acceptance and the combined trigger, reconstruction, and selection efficiency, both of which are obtained from Monte Carlo (MC) simulations. The pp collisions are simulated using PYTHIA [38] version 6.424, the unstable particles are decayed by EVTGEN [39] version 9.1 (using the default matrix element for the signal), and the particles are propagated through a detailed model of the detector with GEANT4 [40]. The reconstruction and selection of the generated events proceed as for data. Three simulated samples were created in which the $B^0$ was forced to decay to $K^{*0}(K^+\pi^-)\mu^+\mu^-$, $J/\psi(\mu^+\mu^-)K^{*0}(K^+\pi^-)$, or $\psi'(\mu^+\mu^-)K^{*0}(K^+\pi^-)$. The samples were constructed to ensure that the number and spatial distribution of pp collision vertices in each event match the distributions found in data. The acceptance is obtained from generated events, before the particle propagation with GEANT4, and is calculated as the fraction of events passing the single-muon requirement of $p_T(\mu) > 3.3$ GeV and $|\eta(\mu)| < 2.3$ relative to all events with $p_T(B^0) > 8$ GeV and $|\eta(B^0)| < 2.2$. As the acceptance requirements are placed on the generated quantities, they are less restrictive than the final selection requirements, which are based on the reconstructed quantities, to allow for the effect of finite resolution. Only events passing the acceptance criteria are processed through the GEANT simulation, the trigger simulation, and the reconstruction software. The combined trigger, reconstruction, and selection efficiency is the ratio of the number of events that pass the trigger and selection requirements and have a reconstructed $B^0$ compatible with the generated $B^0$ in the event, relative to the number of events that pass the acceptance criteria. The compatibility of generated and reconstructed particles is enforced by requiring the reconstructed $K^+$ to appear within a distance $\Delta R = \sqrt{\Delta \eta^2 + \Delta \phi^2}$ of the generated $K^+$ with $\phi$ the azimuthal angle, and analogously for the $\pi^-$, $\mu^+$, and $\mu^-$, where $\Delta R = 0.3$ for the hadrons and 0.004 for the muons. Requiring all four particles in the $B^0$ decay to be so-matched results in an efficiency of 99.6% (0.4% of the events have a correctly reconstructed $B^0$ that is not matched to a generated $B^0$) and a purity of 99.5% (0.5% of the matched candidates are not a correctly reconstructed $B^0$). Efficiencies are determined for both correctly tagged (the $K$ and $\pi$ have the correct charge) and mistagged (the $K$ and $\pi$ charges are reversed) candidates.

Possible peaking backgrounds are studied with simulation. The event selection is applied to inclusive samples of $B^0 \to \psi(\mu^+\mu^-)X$, $B_s \to \psi(\mu^+\mu^-)X$, $B^+ \to \psi(\mu^+\mu^-)X$, and $\Lambda_b \to \psi(\mu^+\mu^-)X$, where “$X$” denotes all allowed SM final states. No evidence for a peaking structure around the $B^0$ mass is found. The distributions of the few events that pass the selection mimic the combinatorial background shape. Studies of simulated $B_s \to K^{*0}(K^+\pi^-)\mu^+\mu^-$ events, gen-
erated with the same branching fraction as $B^0 \to K^{*0}(K^+\pi^-)\mu^+\mu^-$ events, reveal that around 70 such events are expected to peak near the $B_s$ mass, integrated over the entire $q^2$ signal region. We consider this background to be negligible in comparison to the 1397 signal events.

4 Analysis method

This analysis measures the $P_1$ and $P'_5$ variables of the decay $B^0 \to K^{*0}\mu^+\mu^-$ as a function of $q^2$. Figure 1 shows the angular variables needed to define the decay: $\theta_1$ is the angle between the positive (negative) muon momentum and the direction opposite to the $B^0$ ($\bar{B}^0$) in the dimuon rest frame, $\theta_K$ is the angle between the kaon momentum and the direction opposite to the $B^0$ ($\bar{B}^0$) in the $K^{*0}$ ($\bar{K}^{*0}$) rest frame, and $\phi$ is the angle between the plane containing the two muons and the plane containing the kaon and pion in the $B^0$ rest frame. Although the $K^+\pi^-$ invariant mass must be consistent with that of a $K^{*0}$, there can be a contribution from spinless (S-wave) $K^+\pi^-$ combinations [24, 41–43]. This is parametrized with three terms: $F_S$, which is related to the S-wave fraction, and $A_S$ and $A^2_P$, which are the interference amplitudes between the S-wave and P-wave decays. Including these components, the angular distribution of $B^0 \to K^{*0}\mu^+\mu^-$ can be written as [24]:

$$\frac{1}{\text{d}^4\Gamma} \frac{\text{d}^4\Gamma}{\text{d}q^2\text{d}\cos\theta_1\text{d}\cos\theta_K\text{d}\phi} = \frac{9}{8\pi} \left\{ \frac{2}{3} \left[ (F_S + A_S \cos\theta_K) (1 - \cos^2\theta_1) + A^2_P \sqrt{1 - \cos^2\theta_K} \right. \right.$$

$$\left. \sqrt{1 - \cos^2\theta_1 \cos\phi} + (1 - F_S) \left[ 2F_L \cos^2\theta_K (1 - \cos^2\theta_1) \right. \right.$$

$$\left. + \frac{1}{2} (1 - F_L) (1 - \cos^2\theta_K) (1 + \cos^2\theta_1) + \frac{1}{2} P_1 (1 - F_L) \right. \right.$$

$$\left. (1 - \cos^2\theta_K)(1 - \cos^2\theta_1) \cos 2\phi + 2P'_5 \cos\theta_K \sqrt{F_L (1 - F_L)} \right. \right.$$

$$\left. \sqrt{1 - \cos^2\theta_K} \sqrt{1 - \cos^2\theta_1 \cos\phi} \right\}. \tag{1}$$

The expression is an exact simplification of the full angular distribution, obtained by folding the $\phi$ and $\theta_1$ angles around zero and $\pi/2$, respectively. Specifically, if $\phi < 0$, then $\phi \to -\phi$, and the new $\phi$ domain is $[0, \pi]$. If $\theta_1 > \pi/2$, then $\theta_1 \to \pi - \theta_1$, and the new $\theta_1$ domain is $[0, \pi/2]$. Fitting the data with the full angular distribution would cause fit convergence problems due to the limited number of signal candidate events, which is why we adopt the folding procedure. It exploits the odd symmetry of the angular variables with respect to $\phi = 0$ and $\theta_1 = \pi/2$ in such a manner that the cancellation about these angular values is exact.

For each $q^2$ bin, the observables of interest are extracted from an unbinned extended maximum-likelihood fit to four variables: the $K^+\pi^-\mu^+\mu^-$ invariant mass $m$ and the three angular variables $\theta_1, \theta_K,$ and $\phi$. For each $q^2$ bin, the unnormalized probability density function (pdf) has the following expression:

$$\text{pdf}(m, \theta_K, \theta_1, \phi) = Y^S_F \left[ S^C(m) S^a(\theta_K, \theta_1, \phi) \epsilon^C(\theta_K, \theta_1, \phi) \right.$$  

$$+ \frac{f^M}{1 - f^M} S^M(m) S^a(-\theta_K, -\theta_1, \phi) \epsilon^M(\theta_K, \theta_1, \phi) \right]$$

$$+ Y_B B^m(m) B^p(\theta_K) B^\phi(\theta_1) B^\phi(\phi), \tag{2}$$

where the contributions correspond to correctly tagged signal events, mistagged signal events, and background events. The parameters $Y^S_F$ and $Y_B$ are the yields of correctly tagged signal
events and background events, respectively, and are free parameters in the fit. The parameter $f^M$ is the fraction of signal events that are mistagged and is determined from MC simulation. The signal mass probability functions $S^C(m)$ and $S^M(m)$ are each the sum of two Gaussian functions sharing the same mean, and describe the mass distribution for correctly tagged and mistagged signal events, respectively. In the fit, the mean, the four Gaussian $\sigma$ parameters, and two fractions relating the contribution of each Gaussian, are determined from simulation, which has been found to accurately reproduce the data. The function $S^O(\theta_K, \theta_l, \phi)$ describes the signal in the three-dimensional (3D) space of the angular variables and corresponds to Eq. (1). The combination $B^m(m) B^K_K(\theta_K) B^K_l(\theta_l) B^K^O(\phi)$ is obtained from $B^0$ sideband data and describes the background in the space of $(m, \theta_K, \theta_l, \phi)$, where the mass distribution is an exponential function and the angular distributions are polynomials ranging from second to fourth degree, for both $B^K_K(\theta_K)$ and $B^K_l(\theta_l)$, depending on the $q^2$ bin, while the term $B^K^O(\phi)$ is of first degree for all $q^2$ bins. The functions $e^C(\theta_K, \theta_l, \phi)$ and $e^M(\theta_K, \theta_l, \phi)$ are the efficiencies in the 3D space of $-1 \leq \cos \theta_K \leq 1, 0 \leq \cos \theta_l \leq 1$, and $0 \leq \phi \leq \pi$ for correctly tagged and mistagged signal events, respectively. The numerator and denominator of the efficiency are separately described with a nonparametric technique, which is implemented with a kernel density estimator [44, 45]. The final efficiency distributions used in the fit are obtained from the ratio of 3D histograms derived from the sampling of the kernel density estimators. The histograms have 40 bins in each dimension. A consistency check of the procedure used to determine the efficiency is performed by dividing the simulated data sample into two independent subsets, and extracting the angular variables from the first subset using the efficiency computed from the second subset.

The efficiencies for both correctly tagged and mistagged events peak at $\cos \theta_l$ near 0 for $q^2 < 10 \text{GeV}^2$, becoming flat for larger values of $q^2$, and are relatively flat in $\phi$. The efficiency for correctly tagged events tends to decrease with increasing $\cos \theta_K$, and for $q^2 > 14 \text{GeV}^2$ a small decrease is seen for $\cos \theta_K$ near 0. The efficiency for mistagged events is maximal near $\cos \theta_K = 0$ for $q^2 < 9 \text{GeV}^2$, while for $q^2 > 14 \text{GeV}^2$ a mild maximum appear near $\cos \theta_K = 1$.

The fit is performed in two steps. The initial fit uses the data from the sidebands of the $B^0$ mass to obtain the $B^m(m)$, $B^K_K(\theta_K)$, $B^K_l(\theta_l)$, and $B^K^O(\phi)$ distributions (the signal component is absent from this fit). The sideband regions are $3\sigma_m < |m - m_{\text{true}}| < 5.5\sigma_m$, where $\sigma_m$ is the average mass resolution ($\approx 45 \text{MeV}$), obtained from fitting the MC simulation signal to a sum of two Gaussians with a common mean. The distributions obtained in this step are then fixed for the second step, which is a fit to the data over the full mass range. The free parameters in this fit are the angular parameters $P_l$, $P_\phi$, and $A^K_\phi$, and the yields $Y^K_{\text{S}}$ and $Y^K_{\text{B}}$. In order to avoid fit convergence problems due to the limited number of signal candidate events, the angular...
parameters $F_L$, $F_S$, and $A_S$ are fixed to previous CMS measurements performed on the same data set with the same event selection criteria [29].

The expression describing the angular distribution of $B^0 \to K^{*0}\mu^+\mu^-$, Eq. (1) and also its more general form in Ref. [24], can become negative for certain values of the angular parameters. In particular the pdf in Eq. (2) is only guaranteed to be nonnegative for a particular subset of the parameter space $P_1$, $P'_5$, and $A^S_5$, whose mathematical expression is nontrivial. The presence of such a physical region greatly complicates the numerical maximisation process of the likelihood by MINUIT [46] and especially the error determination by MINOS [46], in particular next to the boundary between physical and unphysical regions. Therefore the second fit step is performed by discretizing the bidimensional space $P_1$–$P'_5$, and by maximizing the likelihood as a function of the nuisance parameters $Y^S_5$, $Y^B_5$, and $A^S_5$ at fixed values of $P_1$ and $P'_5$. Finally the distribution of the likelihood values is fit with a bivariate Normal distribution whose position of the maximum inside the physical region corresponds to the best estimate of the angular parameters $P_1$ and $P'_5$.

The interference terms $A_S$ and $A^S_5$ must vanish if either of the two interfering components vanish. From Ref. [24], these constraints are implemented as $|A_S| < \sqrt{12F_S(1-F_S)F_L} R$ and as $|A^S_5| < \sqrt{3F_S(1-F_S)(1-F_L)(1+P_1)} R$, where $R$ is a ratio related to the S-wave and P-wave line shapes, estimated to be 0.89 near the $K^{*0}$ mass. The constraint on $A_S$ is naturally satisfied since the measurement of the parameters $F_S$, $F_L$, and $A_S$ is inherited from the previous CMS analysis [29].

To ensure correct coverage for the uncertainties of the angular parameters, the Feldman-Cousins (FC) method [47] is used with nuisance parameters. Two main sets of pseudo-experimental samples are generated to compute the coverage for the two angular observables $P_1$ and $P'_5$, respectively. The first (second) set, used to compute the coverage for $P_1$ ($P'_5$), is generated by assigning values to the other parameters as obtained by profiling the bivariate Normal distribution description of the likelihood on data at fixed $P_1$ ($P'_5$) values. When fitting the pseudo-experimental samples, the same fit procedure as applied to data is used.

The fit formalism and results are validated through fits to pseudo-experimental samples, MC simulation samples, and control channels. Additional details, including the sizes of the systematic uncertainties assigned from these fits, are described in Section 5.

5 Systematic uncertainties

Since the efficiency is computed with simulated events, it is essential that the MC simulation program correctly reproduce the data, and extensive checks have been performed to verify its accuracy. The systematic uncertainties associated with the efficiencies and other effects are described below and summarized in Table 1.

The correctness of the fit function, of the folding procedure, and of the procedure for measuring the variables of interest are verified in three ways. Firstly, a large, statistically precise MC signal sample, with approximately 400 times the number of events as the data, is used to verify that the fitting procedure produces results consistent with the input values to the simulation. The discrepancy between the input and output values in this check is assigned as a simulation mismodeling systematic uncertainty. It is also verified that fitting a sample with only either correctly tagged or mistagged events gives the correct results. Secondly, the large MC signal sample is divided into 200 subsamples and combined with background events obtained from the pdf in Eq. (2) to mimic 200 independent data sets of similar size to the data. These are used
to estimate a fit bias, by comparing the average value of the results obtained by fitting the 200 samples to the result obtained using the full MC signal sample. Much of the observed bias is a consequence of the fitted parameters lying close to the boundaries of the physical region. Third, 200 pseudo-experiments, each with the same number of events as the data sample, are generated in each $q^2$ bin using the pdf in Eq. (2), with parameters obtained from the fit to the data. Fits to these 200 samples do not reveal any additional systematic uncertainty.

Because the efficiency functions are estimated from a finite number of simulated events, there is a corresponding statistical uncertainty in the efficiency. The efficiency functions are obtained from fits to simulated data. Alternatives to the default efficiency function are obtained by generating 100 new distributions of both for the numerator and the denominator of the efficiency ratio using the default kernel density estimators as pdfs. The effect of these different efficiency functions on the final result is used to estimate the systematic uncertainty.

The principal check of the efficiency is obtained by comparing the efficiency-corrected results for the control channels with the corresponding world-average values. Since the $B^0 \rightarrow J/\psi K^{*0}$ control channel has reduced uncertainties with respect to $B^0 \rightarrow \psi' K^{*0}$, the efficiency as a function of the angular variables is checked by comparing the $F_L$ measurements from the former sample, composed of 165 000 signal events. The value of $F_L$ obtained in this analysis is $0.537 \pm 0.002$ (stat), compared with the world-average value of $0.571 \pm 0.007$ (stat+syst) [37], indicating a discrepancy of 0.034, which is then propagated to $P_1$ and $P'_5$ by taking the RMS of their distributions resulting from refitting the data 200 times varying $F_L$ within the discrepancy. For completeness, the $F_L$ variable is also measured with the $B^0 \rightarrow \psi' K^{*0}$ sample, resulting in the value $0.538 \pm 0.008$ (stat) to be compared with the world-average value of $0.463^{+0.028}_{-0.040}$ (stat+syst) [37].

The pdf used in the analysis accommodates cases in which the kaon and pion charges are correctly or incorrectly assigned. Both of these contributions are treated as signal. The mistag fraction is fixed to the value obtained from MC simulation. In the statistically precise $B^0 \rightarrow J/\psi K^{*0}$ control channel, the mistag fraction is allowed to vary in the fit and a value of $f_M = (14.5 \pm 0.005)$% is found, to be compared to the simulated value of $(13.7 \pm 0.1)$%. The 0.8% discrepancy is then propagated to $P_1$ and $P'_5$ by taking the RMS of their distribution resulting from refitting the data 10 times varying $f_M$ within the discrepancy.

The systematic uncertainty associated with the functions used to model the angular distribution of the background is obtained from the statistical uncertainty in the background shape, as these shapes are fixed in the final fit. This uncertainty is obtained by fitting the data 200 times after varying the background parameters within their Gaussian uncertainties and taking the RMS of the angular variable values as systematic uncertainty. The factorization of the background pdf in Eq. (2), $B^0 \theta_K (\theta_K) B^0 \theta_l (\theta_l) B^0 (\phi)$, is verified by dividing the range of an angular variable into two, using the centerpoint of the range, and comparing the distributions of events from the two halves in the other angular variables.

The mass distributions for the correctly tagged and mistagged events are each described by the sum of two Gaussian functions, with a common mean for all four Gaussian functions. All parameters (the mean, four $\sigma$, and two ratios) are obtained from fits to MC-simulated events. For the statistically precise control channels, it is possible to fit the data while allowing some of the parameters to vary. The maximum change in the measured values in the two $q^2$ bins corresponding to the two control channels when the parameters are allowed to vary is taken as the systematic uncertainty for all $q^2$ bins.

The $q^2$ bins just below and above the $J/\psi (\psi')$ region may be contaminated with $B^0 \rightarrow J/\psi K^{*0}$.
(B⁰ → ψ′K*⁰) “feed-through” events that are not removed by the selection criteria. A special fit in these two bins is made, in which an additional background term is added to the pdf. This background distribution is obtained from the MC simulation of B⁰ → J/ψK*⁰ and B⁰ → ψ′K*⁰ decays, and the background yield is a free parameter. The resulting changes in the fit parameters are used as estimates of the systematic uncertainty associated with this contribution.

In the final fit, the Fₗ, Fₛ, and Aₛ parameters are fixed to values found in previous CMS measurements [29]. To propagate their uncertainty, one pseudo-experiment per q² bin is generated using the pdf parameters determined from the fit to data. The number of events in these pseudo-experiments is 100 times that of the data. The pseudo-experiments are then fitted twice, once with the same procedure as in data and once with all angular parameters free to vary. The ratio ρ of the statistical uncertainties in P₁ and P₅ determined from the two fits is used to compute the systematic uncertainty, which is proportional to the confidence interval determined with the FC method through the coefficient √ρ² − 1. The stability of ρ as a function of the number of events is also verified.

The effects of angular resolution on the reconstructed values of θK and θl are estimated by performing two fits on the same MC-simulated events. One fit uses the true values of the angular variables and the other fit their reconstructed values. The discrepancy in the fitted parameters between the two fits is taken as an estimate of the systematic uncertainty.

The systematic uncertainties are measured and applied in each q² bin, with the total systematic uncertainty obtained by adding the individual contributions in quadrature.

As a note for future possible global fits of our P₁ and P₅ data, the systematic uncertainties associated with the efficiency, Kπ mistagging, mass distribution, and angular resolution can be assumed to be fully correlated bin-by-bin, while the remaining uncertainties can be assumed to be uncorrelated.

Table 1: Systematic uncertainty contributions for the measurements of P₁ and P₅. The total uncertainty in each q² bin is obtained by adding each contribution in quadrature. For each item, the range indicates the variation of the uncertainty in the q² bins.

<table>
<thead>
<tr>
<th>Systematic uncertainty</th>
<th>P₁ (10⁻³)</th>
<th>P₅ (10⁻³)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simulation mismodeling</td>
<td>1–33</td>
<td>10–23</td>
</tr>
<tr>
<td>Fit bias</td>
<td>5–78</td>
<td>10–119</td>
</tr>
<tr>
<td>MC statistical uncertainty</td>
<td>29–73</td>
<td>31–112</td>
</tr>
<tr>
<td>Efficiency</td>
<td>17–100</td>
<td>5–65</td>
</tr>
<tr>
<td>Kπ mistagging</td>
<td>8–110</td>
<td>6–66</td>
</tr>
<tr>
<td>Background distribution</td>
<td>12–70</td>
<td>10–51</td>
</tr>
<tr>
<td>Mass distribution</td>
<td>12</td>
<td>19</td>
</tr>
<tr>
<td>Feed-through background</td>
<td>4–12</td>
<td>3–24</td>
</tr>
<tr>
<td>Fₗ, Fₛ, Aₛ uncertainty propagation</td>
<td>0–126</td>
<td>0–200</td>
</tr>
<tr>
<td>Angular resolution</td>
<td>2–68</td>
<td>0.1–12</td>
</tr>
<tr>
<td>Total systematic uncertainty</td>
<td>60–220</td>
<td>70–230</td>
</tr>
</tbody>
</table>

6 Results

The signal data, corresponding to 1397 events, are fit in seven q² bins from 1 to 19 GeV². As an example, the individual mass and angular distributions for the second and third q² bins of the K⁺π⁻μ⁺μ⁻ channel, along with the fit projections, are shown in Fig. 2. The fitted values of the
signal yields, $P_1$, and $P_2'$, along with their associated uncertainties, are given for each of the $q^2$ regions in Table 2. These results are also shown in Fig. 3, along with the SM predictions. The fitted values for $A_3^2$ vary from $-0.052$ to $+0.057$.

Figure 2: $K^+\pi^-\mu^+\mu^-$ invariant mass and angular distributions for the second and third $q^2$ bin (top four plots) $2.00 < q^2 < 4.30$ GeV$^2$, and (bottom four plots) $4.30 < q^2 < 6.00$ GeV$^2$. Overlaid on each plot is the projection of the results for the total fit, as well as for the three components: correctly tagged signal, mistagged signal, and background. The vertical bars indicate the statistical uncertainties.

Two SM predictions, SM-DHNV and SM-HEPfit, are available for comparison with the measured angular parameters. The SM-DHNV result, derived from Refs. [17, 24], updates the calculations from Ref. [50] to account for the known correlation between the different form factors [51]. Light-cone sum rule predictions, which are valid in the low-$q^2$ region, are also combined with lattice determinations at high $q^2$ [52] to yield more precise determinations of the form factors over the full $q^2$ range. The hadronic charm-loop contribution is derived from Ref. [53]. The SM-HEPfit result, derived from the calculation reported in Refs. [48, 49], uses full QCD form factors [51] and derives the hadronic contribution from LHCb data [30]. Reliable
Table 2: The measured signal yields, which include both correctly tagged and mistagged events, and the $P_1$ and $P'_s$ values, in bins of $q^2$, for the decay $B^0 \rightarrow K^{*0}\mu^+\mu^-$. The first uncertainty is statistical and the second is systematic. The bin ranges are selected to allow comparisons to previous measurements.

<table>
<thead>
<tr>
<th>$q^2$ (GeV$^2$)</th>
<th>Signal yield</th>
<th>$P_1$</th>
<th>$P'_s$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00–2.00</td>
<td>80 ± 12</td>
<td>+0.12+0.46</td>
<td>−0.47</td>
</tr>
<tr>
<td>2.00–4.30</td>
<td>145 ± 16</td>
<td>−0.69+0.58</td>
<td>−0.27</td>
</tr>
<tr>
<td>4.30–6.00</td>
<td>119 ± 14</td>
<td>+0.53+0.24</td>
<td>−0.33</td>
</tr>
<tr>
<td>6.00–8.68</td>
<td>247 ± 21</td>
<td>−0.47+0.27</td>
<td>−0.23</td>
</tr>
<tr>
<td>10.09–12.86</td>
<td>354 ± 23</td>
<td>−0.53+0.20</td>
<td>−0.14</td>
</tr>
<tr>
<td>14.18–16.00</td>
<td>213 ± 17</td>
<td>−0.33+0.24</td>
<td>−0.23</td>
</tr>
<tr>
<td>16.00–19.00</td>
<td>239 ± 19</td>
<td>−0.53+0.19</td>
<td>−0.19</td>
</tr>
</tbody>
</table>

Figure 3: Measured values of $P_1$ and $P'_s$ versus $q^2$ for $B^0 \rightarrow K^{*0}\mu^+\mu^-$ from CMS, compared with LHCb [30] and Belle [32] results. The statistical uncertainty is shown by the inner vertical bars, while the outer vertical bars give the total uncertainty. The horizontal bars show the bin widths. The vertical shaded regions correspond to the $J/\psi$ and $\psi'$ resonances. The red and blue hatched regions show two SM predictions averaging over each $q^2$ bin to provide a direct comparison to the data. The SM-DHMV result is derived from Refs. [17, 24], while SM-HEPfit result from Refs. [48, 49]. Reliable theoretical predictions are not available near the $J/\psi$ and $\psi'$ resonances.

Theoretical predictions are not available near the $J/\psi$ and $\psi'$ resonances. The two SM predictions are shown in comparison to the data in Fig. 3. Both are seen to be in agreement with the CMS results, although the agreement with SM-DHMV is somewhat better. Thus we do not obtain evidence for physics beyond the SM. Qualitatively, the LHCb data appear to be in better agreement with the SM-HEPfit prediction than with SM-DHMV result, but the uncertainties are too large to allow a definite conclusion.

7 Summary

Using pp collision data recorded at $\sqrt{s} = 8$ TeV with the CMS detector at the LHC, corresponding to an integrated luminosity of 20.5 fb$^{-1}$, an angular analysis has been performed for the decay $B^0 \rightarrow K^{*0}\mu^+\mu^-$. In total, 1397 signal events are obtained. For each bin of the dimuon invariant mass squared ($q^2$), unbinned maximum-likelihood fits were performed to the distri-
butions of the $K^+\pi^-\mu^+\mu^-$ invariant mass and the three decay angles, to obtain values of the $P_1$ and $P_5'$ parameters. The results are among the most precise to date and are consistent with standard model predictions and previous measurements.

**References**


