Measurement of the Production Asymmetry between D^0 and $\bar{D}^0$ mesons with proton-proton collisions at the LHCb experiment

Laxman Seelan

Student ID: 10717429

Department of Physics and Astronomy

The University of Manchester

Performed in collaboration with Samuel Taylor

Supervised by Dr Evelina Gersabeck and Aodhan Burke

(Dated: January 2024)

Abstract: This report presents a measurement of the pseudo-bin-integrated production asymmetry between D^0 and $\bar{D}^0$ in the Cabbibo-favoured decay mode $D^0 \rightarrow K^- \pi^+$. This analysis uses data collected from the Large Hadron Collider beauty experiment in the years 2016, 2017 and 2018 to find the production asymmetry for each year and magnet polarity. These results are averaged to find the overall production asymmetry, which is calculated to be

$$A_{prod}(D^0) = -0.546 \pm 0.122 \pm 0.018 \text{ \%},$$

where the first uncertainty is statistical, and the second is the calculated systematic uncertainty. These results agree with the production asymmetry calculated by a previous analysis, which was found to be $A_{prod}(D^0) = -0.62 \pm 0.11 \pm 0.11$ %. The measurement was repeated in bins of the D^0 p_T η phase space in the range 0 GeVc^{-1} < p_T < 10 GeVc^{-1} and 0 < η < 6. . Fits for the kinematic dependence on production asymmetry were explored but with inconclusive evidence. An analysis of a larger dataset may reveal a dependence.

Contents

1 Introduction

1.1 Background and motivation

After J.J. Thompson's electron discovery in 1897, many more particles were discovered, leading to the Standard Model (SM) formulation [1]. The Large Hadron Collider beauty (LHCb) experiment uses a single-arm forward spectrometer [2]; it probes the laws of physics at the TeV energy scale and searches for physics beyond the SM [3]. Furthermore, the detector investigates the slight differences between matter and antimatter by studying the decays of the bottom (b) and charm (c) quarks [4]. The high integrated luminosity, large production cross-section for heavy hadrons, and a series of detectors make an ideal system for production measurements [2]. Unlike other general-purpose detectors that surround the collision point with an enclosed detector, the LHCb experiment focuses on forward particles that have a high pseudorapidity, η [4]. The pseudorapidity, η, is a coordinate describing the angle the particle is emitted relative to the beam axis. It is observed that the produced B and D mesons show a strong boost in the direction of the beam, corresponding to a high pseudorapidity [5]. Hence, the forward detector is practical for isolating events of interest with greater precision and reducing the impact of background noise. Furthermore, by isolating events containing b or c quarks, researchers can improve their knowledge of rare decays and flavour physics by exploring phenomena such as Charge Parity (CP) violation. Rare b and c hadron decays play a vital role in comprehending the underlying patterns of SM physics and give new directions in model building for unique physics contributions [5].

The experiment aims to calculate the differences in the production rates (production asymmetry) between the D^0 and $\bar{D}^0$ mesons in the proton-proton (pp) collisions at the LHC by using the decay mode of the neutral D meson decaying to a kaon and a pion. Additionally, a kinematic dependence on production asymmetry is explored. Determining production asymmetries (A_{prod}) is paramount since, in CP Violation measurements, one needs to remove the physical asymmetry from other effects that could resemble the effect of CP violation [6]. One of the main goals of this analysis is to significantly increase the amount of data used.

2 Theory

2.1 The Standard Model

The SM of particle physics provides a good explanation for the vast amount of data gathered by various high-energy physics over the 20th century [3], including the very precise tests of the unified electroweak model performed at the Large Electron Positron collider (LEP). The SM uses at least 19 free parameters to characterise particles and their interactions, such as the mass of the Higgs boson [3]. In Quantum Field Theory, particles are described by the excitation of quantum fields, ϕ, which satisfy the corresponding quantum mechanical field equations. These excitations are assigned to different types of particles (for example, electrons, quarks, photons) and are defined throughout spacetime, with coordinates indexed by μ. The interactions and masses of the particles are obtained by requiring the Lagrangian density ($\mathcal{L}$) to be invariant under a local phase transformation (ϕ to ϕ'); this idea is the principle of gauge invariance. The Lagrangian density is the difference between the kinetic and potential energies and its derivatives. An example of a Lagrangian density for a scalar field is shown in Equation 1,

$$\mathcal{L} = \frac{1}{2}\partial_\mu \phi \partial^\mu \phi - \frac{1}{2}m^2\phi^2, \tag{1}$$

where (∂_μ) is the partial derivative with respect to spacetime coordinates, ϕ is the scalar field and m is the scalar particle's mass.
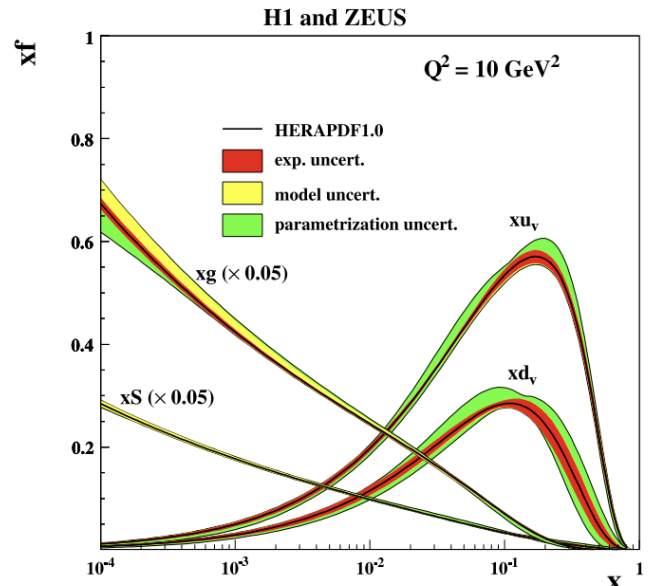
The present-day SM of particle physics was developed from Quantum Electrodynamics (QED), Quantum Chromodynamics (QCD), electroweak theory and other discoveries. The SM contains three families of quarks and leptons, which are spin-1/2 fermions, as well as spin-1 (vector) bosons ($W^{\pm}$, Z, gluon and photons), which are force carriers that mediate the weak, strong and electromagnetism interactions [3] [7]. Quark families consist of two categories: up-type and down-type quarks. The up-type quarks (u, c, t) carry an electric charge of Q = $+2/3e$, where e represents the magnitude of the charge on an electron; the quarks also carry a colour charge. Meanwhile, the down-type quarks (d, s, b) have Q = $-1/3e$. Similarly, the lepton family consists of charged particles (e^{-}, μ^{-}, τ^{-}) and neutrinos (ν_e, ν_μ, ν_τ). Corresponding antiparticles exist for each quark and lepton, characterised by the same mass but opposite quantum numbers. Mesons contain a quark-antiquark bound state, while baryons are a bound state with three quarks [8]. QCD also reveals gluon self-interaction. Consequently, coloured objects such as free quarks and gluons cannot be observed - only colourless bound states can be observed [3] [9]. The latest fundamental particle confirmed by experiments is the spin-0 (scalar) boson, the Higgs. The particle was theoretically predicted 50 years before its discovery by the SM to accommodate a spontaneous symmetry breaking in gauge theories that explains the mass of the weak force carriers [10].

The discovery of the Higgs boson marked a significant validation of the SM framework. However, it is recognised that the SM remains incomplete. The SM successfully explains 4% of the universe's energy content (baryonic matter) but does not explain dark matter or energy. Furthermore, the fourth fundamental force, gravity, is not described well by QFT, and the SM has a hierarchy problem where the quantum corrections to the Higgs boson are too large. These limitations collectively suggest that the Standard Model may not represent the final theory. Exploring a grand unified theory incorporating gravity may offer a more comprehensive explanation for these outstanding issues, including the observed matter-antimatter asymmetry in the universe.

2.2 Parton Distribution Function
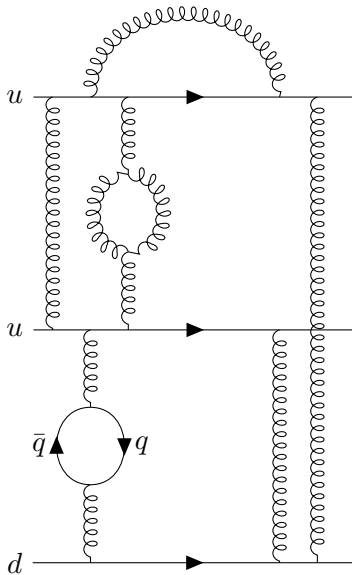
(a) The quark-parton model of the proton.   **(b)** The PDF for a proton.

Fig. 1. a) A Feynman diagram illustrating the three valence quarks of the proton interacting via the strong force, resulting in the 'sea quarks' [3]. b) A graph of the proton's PDF encoding the structure of strongly interacting particles, probed at 10GeV by ZEUS. The gluon (xG) and sea distribution (xS) are scaled down by a factor of 20 [11]. The probability of observing a type of particle is depicted to vary with the fractional momentum of the proton that the particle carries.

The interaction is elastic at low energy collision between protons; meanwhile, at high energies, they are inelastic, as explained by the quark-parton model. The quark-parton model theorises that the proton consists

of three valence quarks, uud, and a sea of quarks and gluons mediating the strong interaction. This process is shown in Figure 1a for the proton. The quark loop must have identical flavour quarks as the strong force conserves flavour in interactions. The valence quarks determine the quantum numbers cf the proton, while the other interactions within contribute to the proton's mass. For high-energy interactions, the partons are considered as being free, and the momentum of the parton (p_i) is expressed as a fraction (x) of the total momentum of the proton (p_{proton}) [12]:

$$p_i = x p_{proton}.$$ (2)

The parton distribution function (PDF) provides information on the probability of extracting a specific parton from the colliding hadrons [13]. Figure 1b shows the PDF for a proton fitted from the data obtained at the ZEUS particle accelerator. If the parton has a low x, the probability that it is a gluon or a sea quark is high. On the other hand, the probability of extracting a valence quark is high if the parton has a high x. An accurate calculation of the PDF is essential for perturbative cross-section calculations; during the pp collision, only a selected quark within the proton will collide. This results in an unknown initial momentum of the parton and has to be predicted by the PDF. The remaining quarks in the protons are spectator quarks that are not involved in the collision.
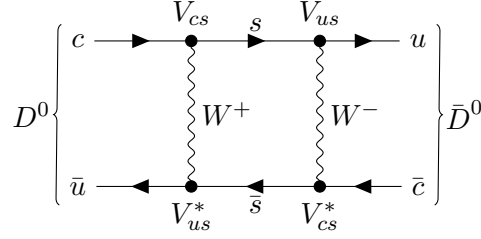
2.3 CP Violation

Among the discrete symmetries in particle physics, the combined symmetry under charge conjugation (C) and parity (P) is particularly significant [12]. C is the transformation associated with exchanging particles and antiparticles [14] by conjugating all internal quantum numbers, such as the electric charge [12]. Similarly, P transformation is defined as the inversion of the spatial coordinates [15]. The Time Reversal (T) process changes a physical system with a given sense of time evolution into another with the opposite sense [16]. Violation of CP symmetry (CPV) is one of Sakharov's conditions for the evolution of a matter-dominated universe [12]. The Big Bang theorises that an equal amount of matter and antimatter were created; the expansion and subsequent cooling of the universe allowed the production of photons via matter-antimatter annihilation to fall out of the thermal bath. If the interactions were symmetric concerning matter and antimatter, and if the baryon number was conserved, all particles would convert to photons [17].

CPV can be partially explained by the interactions propagated by the $W^{\pm}$ bosons - the only tree-level interactions in the SM that change flavour [17]. In the electroweak sector, the quarks can mix via the Cabibbo-Kobayashi-Maskawa (CKM) mixing matrix [3], which quantifies the matrix element (V_{ij}) of the transition. Mixing is also possible in the QCD sector, but the parameters are nearly zero [3]. The CKM matrix is obtained by extending the Cabbibo mixing angle concept to include three pairs and a single CP-odd phase [12] [18]. By convention, the CKM matrix is given by a 3 x 3 unitary matrix [19],

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \cdot \begin{pmatrix} d \\ s \\ b \end{pmatrix},$$ (3)

where the weak eigenstates (d', s', b') are a linear rotation of the mass eigenstates (d, s, b) governed by the CKM matrix. The transition probability between quark i and j is proportional to the $|V_{ij}|^2$; the diagonal terms in the CKM matrix have the highest probability. The complex phase in the CKM matrix explains direct CPV. After performing the $\hat{C}$ on the coupling for a process, the matrix element is altered into V_{ij}^*. If $V_{ij}^* \neq V_{ij}$, then the process violates CP symmetry, as the partial decay widths of the particle and its antiparticle will be different. However, in a local field theory, all processes are invariant under the combined operation CPT [17]; this implies that the full decay width will be the same [20].

In 1983, the Wolfenstein characterisation of the CKM was developed [18], where the complex (phase) elements can only appear in V_{td} and V_{ub}, as seen in Ref [20]. However, a detailed expression up to the order of λ^6 will show phase elements in other transitions but not in the V_{cs} matrix element [18]. This reveals that a quark flavour change from charm quark to the strange quark will not contribute to the CP Violation as $V_{cs} = V_{cs}^*$.

Fig. 2. A Feynman diagram illustrating the D^0 oscillation to $\bar{D}^0$ [21]. Indirect CP violation is observed if the process is not time-symmetric.

Quarks can only be observed bound into hadrons due to the quantum number colour. This allows a neutral meson to oscillate into its antiparticle. A Feynman diagram of this process is shown in Figure 2, illustrating the process proceeding via the $W^\pm$ bosons. Mixing occurs as the physical neutral meson states are a superposition of its matter and antimatter eigenstates,

$$D^0_H = \left(pD^0 + q\overline{D^0} \right), \tag{4a}$$

$$D^0_L = \left(pD^0 - q\overline{D^0} \right), \tag{4b}$$

where the physical states are labelled D^0_H and D^0_L to distinguish the heavier (H) from the lighter (L) [12] [21]. Moreover, in the equation, p and q are complex coefficients that satisfy $|p|^2 + |q|^2 = 1$, so that the total probability equals one. CP Violation in mixing is observed if $|q/p| \neq 1$, implying the oscillation has a preferred final state [12].

(a) $c\bar{u} \to s\bar{u} + u\bar{d}$

(b) $\bar{c}u \to \bar{s}u + \bar{u}d$

Fig. 3. Feynman diagrams showing the leading order decays of the neutral D meson decaying to a kaon and pion.

The leading order (LO) decay channels $D^0 \to K^-\pi^+$ and $\bar{D}^0 \to K^+\pi^-$ are the focus of this analysis and are shown in Figure 3. The relatively large cross-section of the decay, due to being a Cabbibo-favoured decay as the flavour change only occurs between the same generation of quarks, makes the decay mode an excellent probe to measure A_{prod}. Moreover, the D^0 oscillation parameter $|q/p|$ is measured to be 0.995 $\pm$0.016; therefore, indirect CPV can be assumed to be negligible in this decay mode [22]. As shown by the Wolfenstein characterisation of the CKM, we expect no CPV for the leading order decay mode [12]; however, a small amount of CPV could be found in the background peaks mistaken as signals [23]. These background events are Cabbibo-suppressed decays of the D^0 meson, such as the $D^0 \to \pi^+\pi^-$ decay mode, as these proceed through an off-diagonal CKM matrix element [22].

2.4 Production Asymmetry

Production asymmetry arises from the different hadronisation probabilities between D^0 and $\bar{D}^0$ mesons in pp collisions: $\sigma(D^0) \neq \sigma(\bar{D}^0)$ [24]. A_{prod} for the neutral D meson is defined as,

$$A_{prod}(D^0) = \frac{\sigma(D^0) - \sigma(\bar{D}^0)}{\sigma(D^0) + \sigma(\bar{D}^0)}, \tag{5}$$

where $\sigma(D^0)$ and $\sigma(\bar{D}^0)$ are the production cross sections of the meson. However, QCD calculations of the cross sections of decays are complex due to the difficulty in modelling jets and non-perturbative QCD effects. Therefore, asymmetry in the signal yield, A_{raw}, is measured by performing an extended maximum-likelihood fit to the invariant mass of the neutral D mesons,

$$A_{raw}(D^0) = \frac{N(D^0) - N(\bar{D}^0)}{N(D^0) + N(\bar{D}^0)}, \tag{6}$$

where $N(D^0)$ and $N(\bar{D}^0)$ are the signal yield of the respective meson. A_{prod} can also be written as a sum of its contributions [6],

$$A_{prod}(D^0) = A_{raw}(D^0) - A_{Det}(D^0) - A_{CP}(D^0), \tag{7}$$

where $A_{Det}(D^0)$ is the asymmetry due to the detector and $A_{CP}(D^0)$ is the asymmetry related to CPV. However, as previously explained, $A_{CP}(D^0)$ is negligible as the decay mode is Cabbibo favoured. The error on $A_{prod}(D^0)$ is calculated by propagating the errors through quadrature. To reduce residual detection asymmetries, the final results are determined from the arithmetic mean of the results obtained for each bending magnet polarity, denoted MagUp and MagDown [25].

2.5 Leading Particle Effect

(a) Gluons combine to give two charm quarks.   (b) Quarks combine to give two charm quarks.

Fig. 4. Leading order diagrams for $c\bar{c}$ production with (a) the gg contribution and (b) the $q\bar{q}$ annihilation. Two more LO diagrams have gg initial states, which also show that the c and $\bar{c}$ are produced symmetrically [26].

Fixed-target experiments have observed considerable asymmetries in producing charmed mesons and baryons [27]. The leading particle effect explains the observed A_{prod}. The D^0 meson is composed of a $\bar{u}$ and c quark, while the $\bar{D}^0$ consists of a $\bar{c}$ and u quark. If the charmed hadron has a valence parton in common with the proton beam, the hadron is produced in more significant numbers [27]. The antiparticle cross section for the $\bar{D}^0$ will be greater as the u quark contained matches the valence quarks in the proton, unlike the $\bar{u}$ found in the D^0 meson.

There are different models for charm quark hadronisation, one of which is fragmentation [9]. In the fragmentation process, a $q\bar{q}$ pair is generated for each initial c quark from the parton shower [9] [3]; the charm quarks hadronises with q to form the D^0 meson, which results in the $\bar{q}$ being a free quark. The

fragmentation process is repeated for every free quark/antiquark until the available momentum is used up [3].

According to LO perturbative QCD, there is no predicted production asymmetry between the neutral D mesons. This is due to LO diagrams of charm production in pp collision, shown in Figure 4, producing c and $\bar{c}$ symmetrically [27]. In Figure 5, the hadronisation of c and $\bar{c}$ are depicted to be independent of each other. In addition, charge conjugation, C, is a good symmetry in QCD, and no A_{prod} is expected [28]. Production asymmetries are generated at the next-to-leading-order (NLO) process but are too small to explain the experimental results [27]. To explain the experimental results, different models for charm quark hadronisation were theorised, such as the beam drag effect, meson clouds and heavy quark recombination [28].

(a) $\bar{c} \to \bar{D}^0 \bar{u}$ **(b)** $c \to D^0 u$

Fig. 5. The Feynman diagrams show the charm and anticharm quarks' hadronisation via fragmentation [29].

2.6 Heavy Quark Recombination

The Heavy quark recombination (HQR) mechanism can be used to explain the leading particle effect in the forward regime [27]. The mechanism, shown in Figure 6, has explained the asymmetries observed using Fermilab's data in charm quarks using one fit parameter [27]. The HQR mechanism proposes that a heavy quark or gluon, originating from the hard scattering process between protons, combines with a spectator parton from the proton to form the final meson [28]. HQR provides a natural explanation for the leading particle effect because in the forward region, $\sigma(D^0 \to K^- \pi^+) < \sigma(\bar{D}^0 \to K^- \pi^+)$ since the PDF of the u quark will be much higher than of the sea quark ($\bar{u}$) [27], as shown in Figure 1b. Moreover, the variation of production asymmetry with respect to p_T can be explained using the HQR mechanism for the $D^\pm$ mesons as outlined in the references [27] [28], suggesting a similar relationship may be found for a neutral D meson.
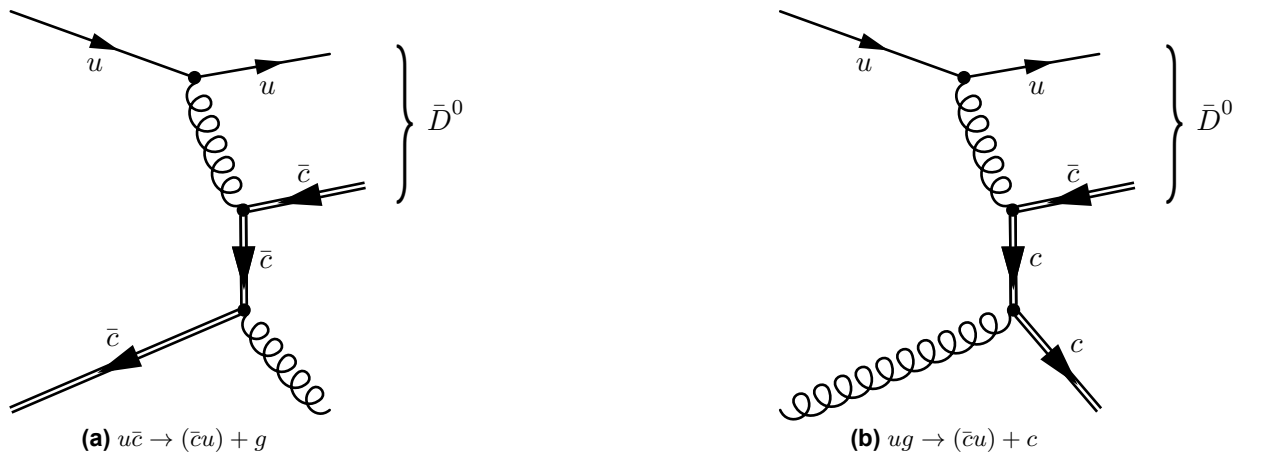

(a) $u\bar{c} \to (\bar{c}u) + g$ **(b)** $ug \to (\bar{c}u) + c$

Fig. 6. The figure shows some Feynman diagrams for the HQR for a $\bar{D}^0$. A spectator quark from the proton combines with (a) a charm quark and (b) gluon produced in pp collision to form the D^0 and an extra parton. Each process has five diagrams [28]. Similar Feynman diagrams exist for the D^0.

3 The LHCb Experiment

Identifying the decay products is the only way for the LHCb experiment to detect D mesons, given the neutral D meson's lifetime of $(410.3 \pm 1.0) \times 10^{-15}$s [22]. LHCb uses a combination of silicon detectors and calorimeters to reconstruct the decay products precisely; this process requires excellent vertex, momentum
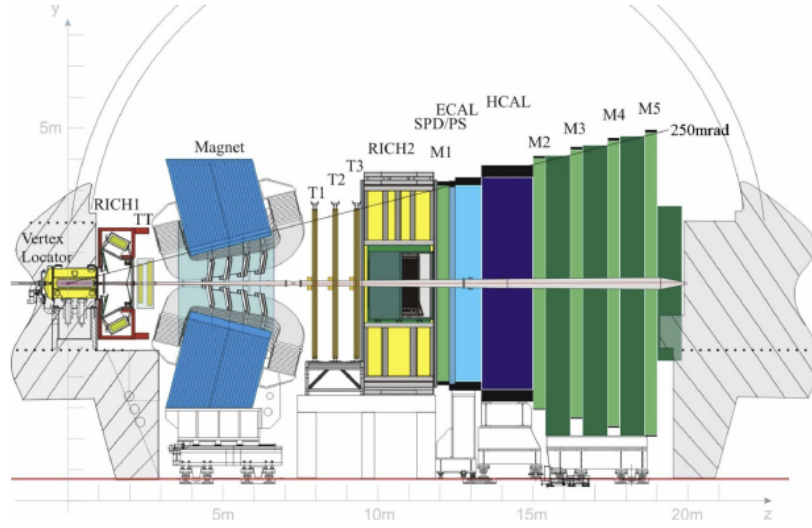
Fig. 7. The experimental layout of the LHCb spectrometer. The figure is obtained from LHC Colabration's report, 'The LHCb detector at the LHC' [30]. The tracking and PID systems are shown and explained in the text.

and invariant mass resolutions [30]. The spectrometer is divided into the tracking and particle identification (PID) systems. The trigger is required to save processing power and storage by quickly selecting the events of interest [8]. The layout of the LHCb spectrometer is shown in Figure 7. The particle reconstruction process is slightly different for neutral particles [30], as these do not interact with the tracking system.

3.1 Tracking System

The tracking system consists of the Vertex Locator (VELO), magnet and silicon trackers to measure the momentum and charge of the charged particle.

Heavy mesons, like the neutral D meson, can have a range of about a few mm; measuring the separation from the primary vertex is critical for selecting the heavy meson decay for further analysis [5]. The VELO, situated in a vacuum tank, surrounds the interaction point with a series of silicon modules [31]; the silicon modules precisely measure the r and ϕ coordinates of a charged particle [30]. The measurements of the track coordinates are used to measure the pp primary vertex and possible secondary vertex with a resolution of 10 μm [5] [8].

The charged particles then pass through dipole magnets to measure their momentum from the particle's trajectory [30] [8]. Three magnets are used to compensate for the magnets' effect on the pp LHC beam [31]. The magnet's polarity is regularly reversed to estimate and control the effect of an asymmetric detector [5].

The detector contains five tracking stations: Trigger Turicensis (TT) tracker contains two tracking stations separated by 27cm before the dipole magnet and three trackers referred to as T trackers. To achieve excellent accuracy, the T trackers are made of two different detector types: Inner and Outer Trackers (IT and OT) [8]. The TT and IT are silicon micro-strip detectors. As the charged particle passes through the silicon detector, an electron-hole pair is created, drifting to a nearby electrode; this signal is used to track the particle's trajectory. Meanwhile, the OT use gas-filled straw tubes to detect the charged particles. The TT tracker is used in the trigger software, while the T trackers are predominantly used to measure a particle's momentum [5]. The resolution for the tracking system is measured to be 0.6% [5].

3.2 Particle Identification System

The PID uses a combination of Ring-imaging-Cherenkov (RICH) detectors, a calorimeter system and muon chambers to determine the composition of the particle.

The Cherenkov effect is the emission of a photon after the particle transitions to a different medium while travelling at a phase velocity, v_p, that is greater than the speed of light in the medium (v_m); the emission of the photon is caused by the charged particles polarising the nearby atoms [1]. Cherenkov radiation is emitted in a cone with a polar angle (θ_c) and can be related to the velocity and mass of the particle m,

$$\cos(\theta_c) = \frac{c}{v_p n} = \frac{1}{n}\sqrt{1 + \left(\frac{mc}{p}\right)^2} \tag{8}$$

where n is the material's refractive index, and p is the measured momentum. The RICH detector uses spherical and flat mirrors to focus the Cherenkov light on the photon detector within [30].

The Calorimeter system provides data regarding the energy and spatial locations of hadrons, electrons, and muons [8], which allows LHCb to distinguish between these particles for the trigger decisions. The Scintillating Pad Detector (SPD) allows the spectrometer to distinguish between electrons and photons, as only the electrons slightly interact with the detector [8]; Utilising the information gathered from the SPD, background processes such as the $\pi^0 \rightarrow \gamma\gamma$ can be rejected. The pre-shower (PS) detectors and Electromagnetic Calorimeter (ECAL) measure the energy of photons and electrons, while the Hadron Calorimeter (HCAL) measures the energy of hadrons. A combination of the calorimeters mentioned is used to differentiate between the hadrons and photons or electrons [5].

Muons minimally interact with the mentioned systems, and the muon system (M2-M5) ensures that only the muons are detected after the calorimeter system by introducing iron absorbers between them. The first muon system, M1, is placed before the calorimeter system for trigger purposes and to improve the momentum resolution [5] [8].

3.3 Trigger

The crossing frequency with interactions visible to the spectrometer is 10 MHz; the trigger has to reduce this to 2 KHz so that the system can save the events for future analysis [30]. The trigger is separated into three systems: a hardware trigger, L0, and two software triggers, HLT1 and HLT2. The L0 trigger reduces the information to 1 MHz by using information from the calorimeter and muon systems at a fixed latency of 4 μs [8]. The software trigger further reduces the frequency by calculating combined quantities like the invariant mass on a processing farm. The HLT1 system adds information from the tracking system to create track candidates, while the HLT2 trigger creates a near-complete decay reconstruction using all the event information.

4 Previous Analysis

The prior analysis measured the production asymmetry between the neutral D mesons by using the decay channels of the D^0 decaying to a kaon and pion. The previous analyses are found in Ref [32] and Ref [33], and a summary of the methodologies and critical findings employed in the analysis are shown below.

4.1 Selection Criteria

The Monte Carlo (MC) framework offers simulated data samples that are accurately matched to both signal and background. This means the events related to correctly reconstructed particles and backgrounds are identifiable. In the preceding analysis, two MC data samples were created: one for the decays originating from the pp collision and another for the secondary decays that occur via an intermediate state.

During the HLT2 process, initial PID selection criteria are implemented for kaons and pions. These criteria are designed to effectively filter out events featuring a kaon candidate with a low likelihood of being a kaon

and a pion candidate that is likely to be misidentified as a kaon. To achieve this, a variable called delta log-likelihoods (DLL) is employed to differentiate between particles based on their distinct likelihoods and to filter out the background. The equation for a DLL between a kaon and pions is,

$$DLL_{K-\pi} = \log(\mathcal{L}_K) - \log(\mathcal{L}_\pi), \tag{9}$$

where $\mathcal{L}_K$ and $\mathcal{L}_\pi$ are the likelihoods that the particle is a kaon and pion. The HLT2 imposes specific constraints, requiring a $DLL_{K-\pi}$ value > 5 for kaons and < 5 for pions. The previous analysis tightens the selection applied by HLT2 to remove more background events. The truth-matched simulation data provided by a MC simulation is used to optimise the selection criteria by maximising the significance, S, while varying $DLL_{K-\pi}$ selection criteria for the pion and kaon. The significance is defined as,

$$S = \frac{N(D^0)}{\sqrt{N(D^0) + N_B(D^0)}}, \tag{10}$$

where $N_B(D^0)$ is the background yield calculated from a fit of the invariant mass distribution of the D meson. Subsequently, the previous analysis found the $DLL_{K-\pi} < 0$ for the pion to result in the least background and no further restrictions were placed on the kaon.

For the neutral D meson decays used in this analysis, an event may have multiple candidates for the reconstructed particle. Techniques, such as "random killing", can be applied to remove multiple candidate events, leading to a more sensitive measurement. The 'random killing' technique does not discriminate between the corresponding events to the signal and combinatorial background and is applied in this analysis.
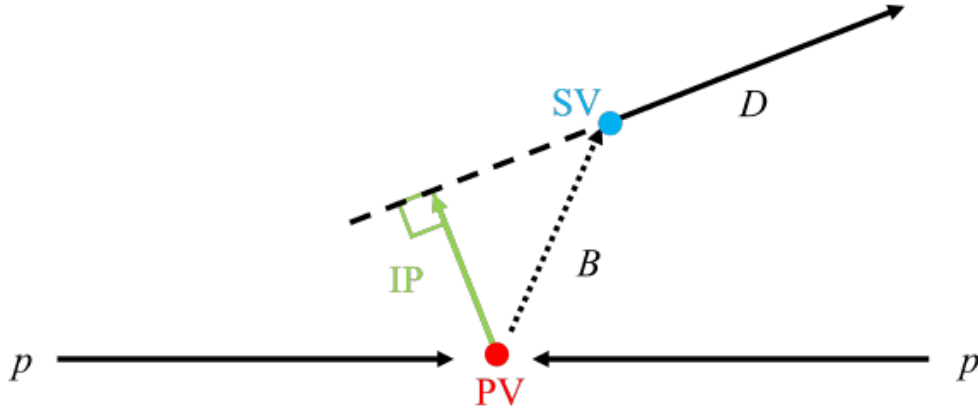
Fig. 8. A diagram illustrating the impact parameter between the primary vertex and the secondary vertex.

Excluding events that have decayed to the neutral D mesons through an intermediary state is vital since the aim is to quantify the production asymmetry of D^0 and $\bar{D}^0$ mesons in pp collisions. The secondary vertex is the point where the intermediary state decays. The impact parameter (IP) is the shortest distance between the primary vertex (PV) and the particle track. The truth-matched MC data is used to find the optimal $\ln(\chi^2_{IP})$ by observing the number of secondary vertex events removed. χ^2_{IP} is defined as the "chi-squared significance of the fit of the impact parameter of the PV and D^0 meson tracks, and the proper lifetime of the particle", where a more significant χ^2_{IP} implies a higher probability of the tracks originating from a secondary vertex.

4.2 Detector Asymmetry

The LHCb has several sources of inefficiency, resulting in A_{Det}. One source is the asymmetrical nature of the detector, which is reduced to first order by the periodic reversal of the dipole magnet. The previous analysis computes an additional form of detection asymmetry arising from the distinct interactions of the decay products of neutral D mesons with the detector material. The detector has an approximately equal distribution of two

types of baryonic matter: protons (p) with valence quarks of uud and neutrons (n) with valence quarks of udd. The HCAL interact with detected particles' quarks via processes like annihilation. Annihilation of a quark and antiquark is possible through the weak, EM, and strong forces, but the energy scale of LHCb implies that the strong force dominates the interaction. The decay product, $\pi^{\pm}$, interacts symmetrically via annihilation with the HCAL as the composition of both contains an antiparticle with a valence quark that matches the valence quark of the nucleons. Meanwhile, the $K^{\pm}$ is detected asymmetrically; the K^+ and K^- mesons have a quark content of $(u\bar{s})$ and $(\bar{u}s)$, where the K^- is more likely to interact with the spectrometer. This is the result of the valence antiquark of K^-, $\bar{u}$ being able to annihilate with a u quark in the detector; meanwhile, this type of annihilation process for K^+ is not permitted via the strong and EM force due to flavour not being conserved in the process.

In experimental observations, obtaining the combined asymmetry of $K^-\pi^+$ pairs was more feasible than measuring the detection asymmetry of $K^{\pm}$ [34]. This combined asymmetry is the difference between the detection efficiencies between $K^-\pi^+$ and $K^+\pi^-$.

$$A_{Det}(K^-\pi^+) = \frac{\varepsilon_{Det}(K^-\pi^+) - \varepsilon_{Det}(K^+\pi^-)}{\varepsilon_{Det}(K^-\pi^+) + \varepsilon_{Det}(K^+\pi^-)}, \tag{11}$$

where the ε_{Det} is the efficiency of the detector measuring the track. Therefore, the detector asymmetry is predicted to be negative as K^+ particles will interact less with the material and will allow the particle to be tracked by the tracking system in the LHCb. The previous analysis used decay modes with known detector asymmetries to calculate $A_{Det}(K^-\pi^+)$. By considering the raw asymmetries of $D^+ \to K^-\pi^+\pi^+$ and $D^+ \to K_s^0\pi^+$ the $A_{Det}(K^-\pi^+)$ can be calculated to be,

$$A_{Det}(K^-\pi^+) = A_{raw}(D^+ \to K^-\pi^+\pi^+) - A_{raw}(D^+ \to K_s^0\pi^+) + A_{Det}(K_s^0), \tag{12}$$

where $A_{Det}(K_s^0)$ is a small asymmetry known to be 0.054 ± 0.014 % [35]. The decay channel of $D^+ \to K_s^0\pi^+$ is utilised to eliminate unnecessary terms, such as the detection asymmetry of π^+. The phase-space-dependent detection and production asymmetries affect the measured asymmetry in D^+ decays [34]. Therefore, to calculate the raw asymmetries in Equation 12, the kinematic variables of the two control modes must be reweighted using the p_T and the pseudorapidity of $K^-\pi^+$ pair using a public tool [34]. A comprehensive derivation of Equation 12 and an in-depth explanation of the reweighting procedure can be found in the earlier analysis documented in Ref [33], as well as the creators of the tool [34].

Year	Polarity	$A_{Det}(K^-\pi^+)[\%]$
2016	MagUp	-0.82 ± 0.26
	MagDown	-0.10 ± 0.25
2017	MagUp	-0.76 ± 0.27
	MagDown	-0.99 ± 0.27
2018	MagUp	-0.25 ± 0.27
	MagDown	-0.97 ± 0.28

Table 1. A table showing the results of the detector asymmetry in the decay mode of the neutral D meson to a kaon and pion. The results were obtained by using multiple decay modes by the prior analysis.

The outcomes of the calculation of Equation 11, as conducted by the preceding research, containing all years and polarities, are displayed in Table 1. This analysis will use these values to calculate the production asymmetry using Equation 7.

4.3 Blinding

A blind analysis is performed to reduce the bias from the experimenter [36]. During the blind analysis, the result is hidden using an offset,

$$A_{raw}(blind) = \begin{bmatrix} 1 \\ -1 \end{bmatrix} A_{raw} + C, \tag{13}$$

where C is a constant [36] and a matrix element that randomly changes the sign of the asymmetry. A Gaussian distribution was then generated with the mean equal to zero and the standard deviation equal to the uncertainty on the raw asymmetry. C was chosen randomly from this distribution within three standard deviations of the mean. The previous analysis utilised a blind analysis approach until completion, and this same method will also be used for this analysis.

5 Modelling the Invariant Mass Distribution

In the previous analysis, the model employed for the extended likelihood fit of the invariant mass distribution was a combination of three functions. The signal was modelled as a Gaussian and Crystal Ball function, while the background was modelled as a Chebychev function (Model 1). However, it was discovered that the background function required a more accurate fitting to the distribution after the selection criteria were modified. This report outlines the procedure to find the best model.

The likelihood function, $\mathcal{L}(\vec{x}, \vec{\lambda})$, is used to describe the observed distribution,

$$\mathcal{L}(\vec{x}, \vec{\lambda}) = \prod_{i=1}^{N} P(\vec{x}_i | \vec{\lambda}_i), \tag{14}$$

where the distribution of measured observables is $\vec{x}_i = x_1, x_2, ...,$ and $\vec{\lambda}_i = \lambda_1, \lambda_2, ...$ are the free parameters of the probability distribution function. The probability distribution, $P(\vec{x}_i | \vec{\lambda}_i)$, function gives the probability for the measured observables for a given set of parameters [5]. However, an issue with using $\mathcal{L}(\vec{x}, \vec{\lambda})$ is that the observed number of events of one kind is not fixed. It is an unknown parameter subject to Poisson fluctuations; the expected value of the number of events depends on the parameters to be estimated [37]. $\mathcal{L}(\vec{x}, \vec{\lambda})$ is extended,

$$\mathcal{L}(\vec{x}, \vec{\lambda})_{extended} = e^{\int P(\vec{x}|\vec{\lambda})\,d\vec{x}} \prod_{i=1}^{N} P(\vec{x}_i | \vec{\lambda}), \tag{15}$$

to include a Poisson fluctuation term, $e^{\int P(\vec{x}|\vec{\lambda})\,d\vec{x}}$, to correct for this issue. $\mathcal{L}(\vec{x}, \vec{\lambda})_{extended}$ is maximised to find the optimal parameters, $\vec{\lambda}_i$, for the given model and is generated using `RooFit`, a package for statistical modelling distributed with Root [38].

To estimate the quality of the model, a pull for each data, δ_i, can be calculated as,

$$\delta_i = \frac{y_i^{data}(x_i) - y_i^{expected}(x_i)}{\sigma_i^{expected}}, \tag{16}$$

with $y_i^{expected}$ and $\sigma_i^{expected}$ being the value and uncertainty of the probability distribution function at the given x_i, and the $y_i^{data}(x_i)$ is the value of the ith bin. A pull distribution is plotted with values obtained for each y_i, and a Gaussian fit is plotted on the pull distribution using the `curve_fit` function from SciPy [39]. The best model is determined by the pull distribution resembling closest to a normal distribution. This is quantified by the mean

(μ) and standard deviation (σ) of the Gaussian, as a better model would result in a normal distribution with μ equal to zero and σ equal to one.

5.1 Unbinned Model
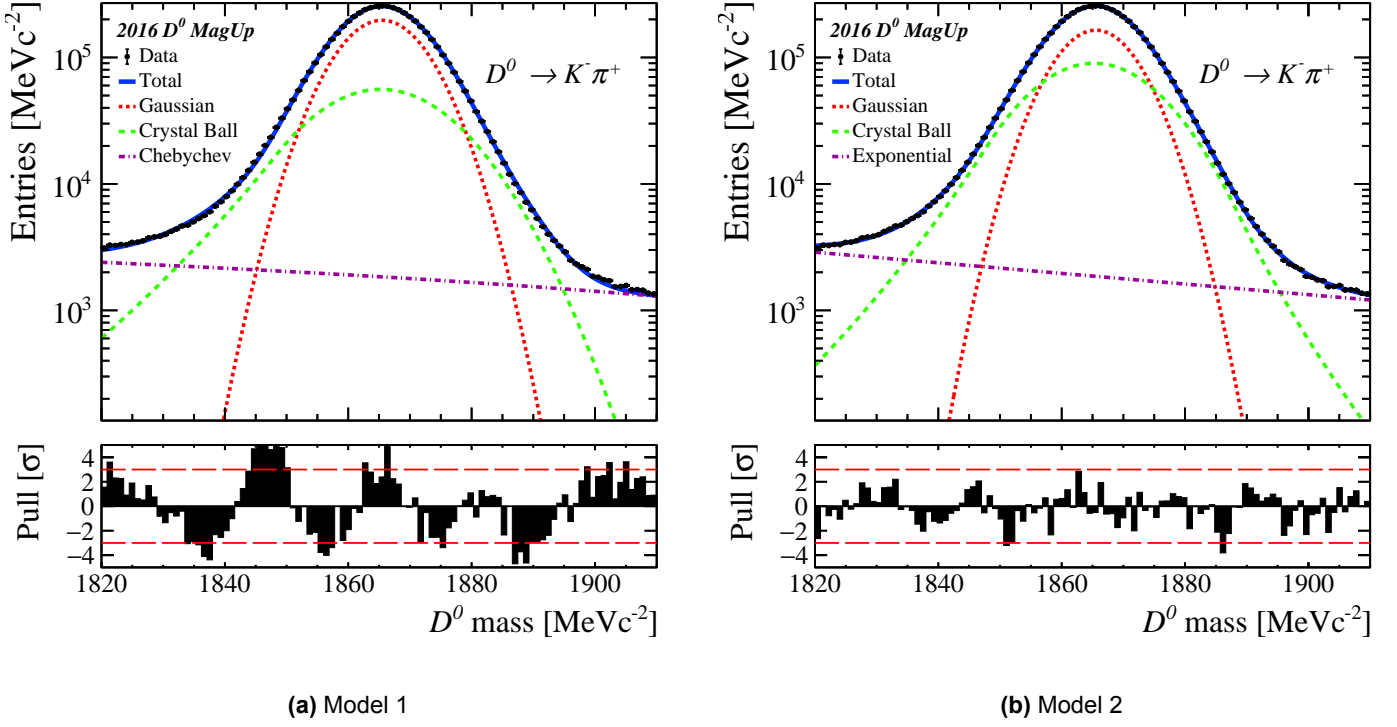
(a) Model 1

(b) Model 2

Fig. 9. Invariant mass distributions of D^0 meson, using the 2016 MagUp data and fitted using a) Model 1 and b) Model 2. The pull distribution shows an initial observation that Model 2 is better fitted to the data. Statistical error bars are present for all points but are too small to be seen.
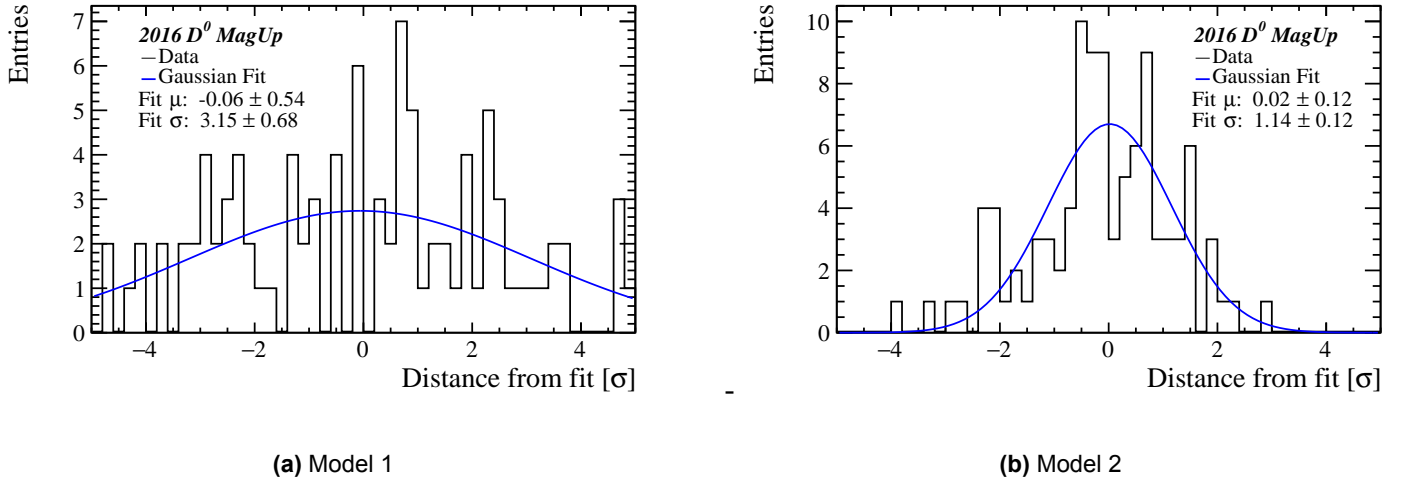

(a) Model 1

(b) Model 2

Fig. 10. Pull distributions of D^0 meson, using the 2016 MagUp data and modelled using Model 1 (a) and Model 2 (b) with an unbinned extended maximum likelihood fit. These were used to assess the model, graphically and analytically, using the μ and σ of the Gaussian fit.

An unbinned extended likelihood fit performed initial modelling of the invariant mass distribution (IMD). A second model (Model 2) was developed using the procedure explained previously, which modelled the combinatorial background using an exponential function and the signal as a Gaussian and Crystal Ball function. Models 1 and 2's resulting fits are shown in Figure 9 and 10. Note that all data distributions shown have statistical error bars unless otherwise specified. The mean and sigma calculated for all years and polarities for Models 1 and 2 are shown in Appendix A.1 and Appendix A.2. For Model 1, the average μ and σ were propagated to be -0.478 ± 2.75 and 2.91 ± 3.71, respectively. Similarly, for Model 2, the average μ and σ were found to be -0.161 ± 0.442 and 1.27 ± 0.412. These results suggest that the pull distribution of Model 2

resembles more of a normal distribution when compared with the pull distribution of Model 1; this implies that the combinatorial background is modelled better with an exponential function.

5.2 Binned Model

To perform a binned fit on the IMD, the fit observable - the mass of D^0 - must be histogrammed. The binned fit is the preferred method for the rest of the analysis due to its computational advantages over the unbinned fit. While the unbinned fit compares each event in the data set with the probability distribution function, the binned fit compares the bin value with the probability distribution function, resulting in a faster fit. The computational advantages facilitate the analysis of more data, resulting in a higher accuracy of the result. However, it is necessary to note that due to Poisson statistics, the binned fit may be less accurate than the unbinned fit as it will perform the fit on fewer data points.
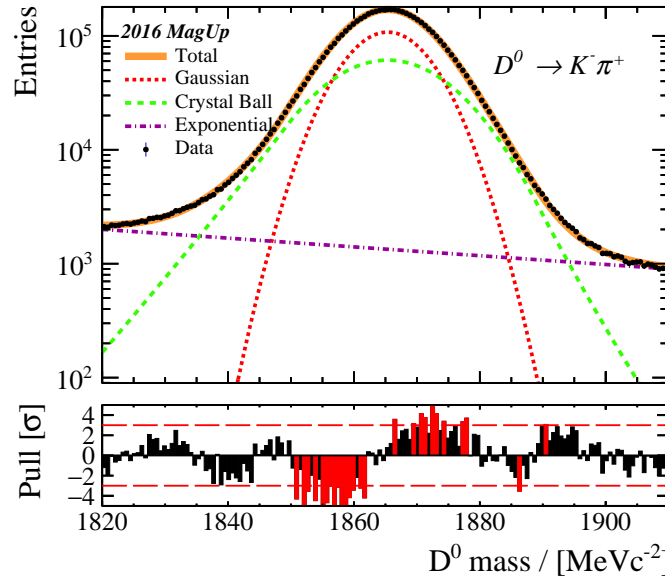
Fig. 11. Invariant mass distributions of D^0 meson, using the 2016 MagUp data with a binned fit, using Model 2, fitted to the data. The red pulls represent pulls that exceed 3σ.

To obtain the production asymmetry, the data was split into data sets for each combination of polarity, meson, and year. The parameters were obtained through a simultaneous binned fit for the four datasets. As illustrated in Figure 11, the pulls around the Gaussian and Crystal Ball peaks are large. Attempts were made to modify the initial guesses for the parameters to improve the fit. However, it was determined that the signal model no longer reflected the data. Nevertheless, the exponential still represents a good fit for the combinatorial background.

	Average μ	Average σ
Model 3	-0.12 ± 0.24	1.23 ± 0.24
Model 4	-0.08 ± 0.26	1.27 ± 0.26
Model 3 with 200 Bins	-0.09 ± 0.08	1.25 ± 0.08
Model 3 with 240 Bins	-0.11 ± 0.05	1.23 ± 0.10

Table 2. Results of the μ and σ of the pull distribution for the different procedures used to fit the IMD.

Numerous models underwent testing to identify the optimal. Several models, such as the three Gaussian models, were disregarded, constantly producing subpar results with the pulls frequently exceeding three sigma. This report will focus on two models: one with the signal modelled as a Johnson and two Bifurcated Gaussian functions (Model 3), and another with the signal modelled as a Gaussian, Johnson and a

Bifurcated Gaussian function (Model 4) as shown in Figure 12. Appendix A.3 and Appendix A.4 display the calculated mean and sigma for all years and polarities for Models 3 and 4. The averages of the μ and σ were then determined for each model. These results, shown in Table 2, suggest that both models' pulls are similar, but the means of the three peaks representing the signal in Model 4 in Figure 12 do not align. As a result, it was decided to adopt Model 3 as the optimal model for the subsequent analysis of production asymmetry and to use Model 4 to calculate the systematic uncertainty in choosing the model.

(a) Model 3   **(b)** Model 4
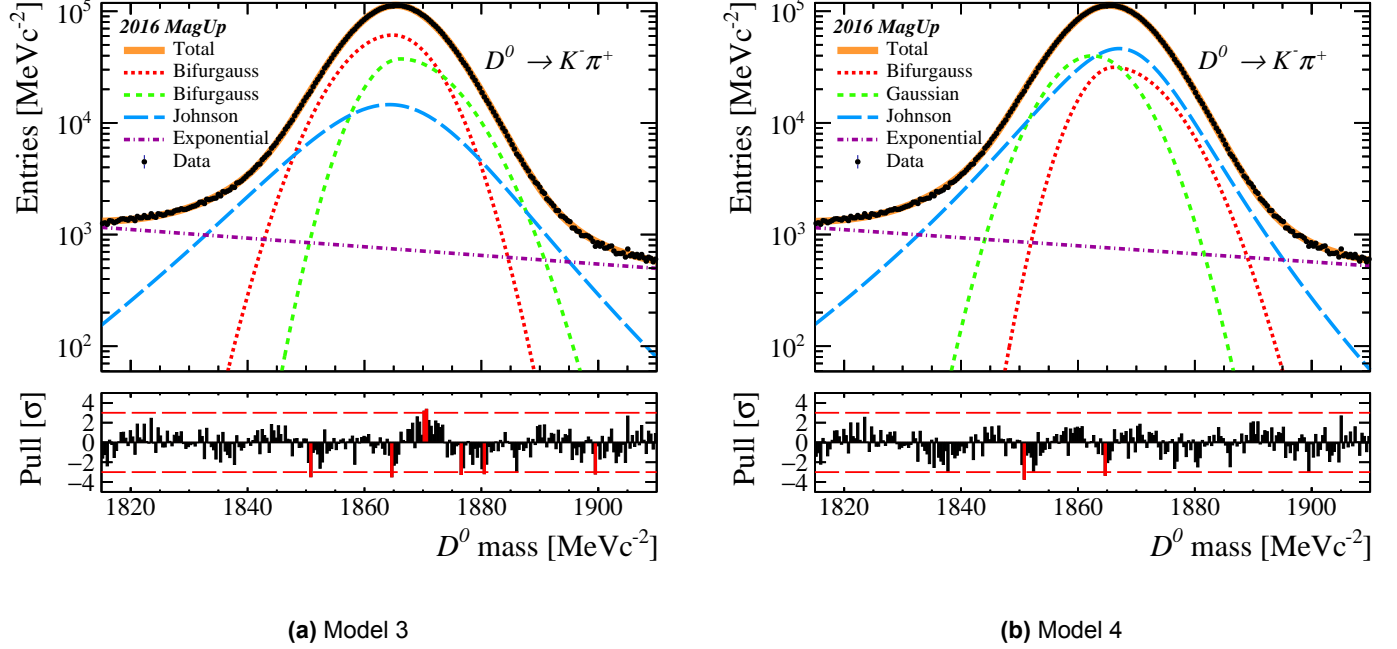
Fig. 12. Invariant mass distributions of D^0 meson, using the 2016 MagUp data and fitted using Model 3 (a) and Model 4 (b). These were used to extract the number of events under the signal, N, to calculate A_{raw} and A_{prod}.

Experimenting with bin numbers between 130 and 260 revealed that most bin numbers led to suboptimal results. However, 200 and 240 bins resulted in good fits and were subjected to further analysis. Comparing the μ and σ values of the pull distribution between these two bin numbers revealed similar results, as displayed in Table 2. The fit is chosen to be performed with 240 bins for the rest of the analysis, and a systematic uncertainity was calculated for this choice.

6 Production Asymmetry

The production asymmetry is calculated for every combination of year and polarity using the procedure discussed in Subsection 2.4. A binning scheme must be applied to the data to explore any kinematic relationships with the production asymmetry in the range of 0 GeV < p_T < 10 GeV and 0 < η < 6. Additionally, the uncertainty calculated through quadrature does not include systematic uncertainties. This section outlines the procedure to create the binning scheme and find the effect of the systematic uncertainties.

6.1 Binning Scheme

This analysis uses three different binning schemes, which contain a comparable amount of data in each bin. The first binning scheme creates 100 bins in the p_T and η phase space, as seen in Figure 13a. The figure illustrates some bins with a higher density of data; to ensure the number of events in each bin is equal, the area of the bins is varied. The bin boundaries are calculated using data from both mesons and both polarities. Following the implementation of the binning method, an extended maximum-likelihood fit is performed for every bin for each meson and polarity combination, as portrayed in Figure 13b. As a result, the signal yield can be obtained to find the production asymmetry in each bin, which is used to explore a potential dependence on the kinematic variables. The procedure for implementing the second and third binning schemes is similar

to the first. The second binning scheme creates ten bins in the phase space of p_T, while the third creates ten bins in the phase space of η.

(a) 100 Bins in p_T and η

(b) A fit for the IMD using Model 3 for D^0 MagUp Bin 25 data.
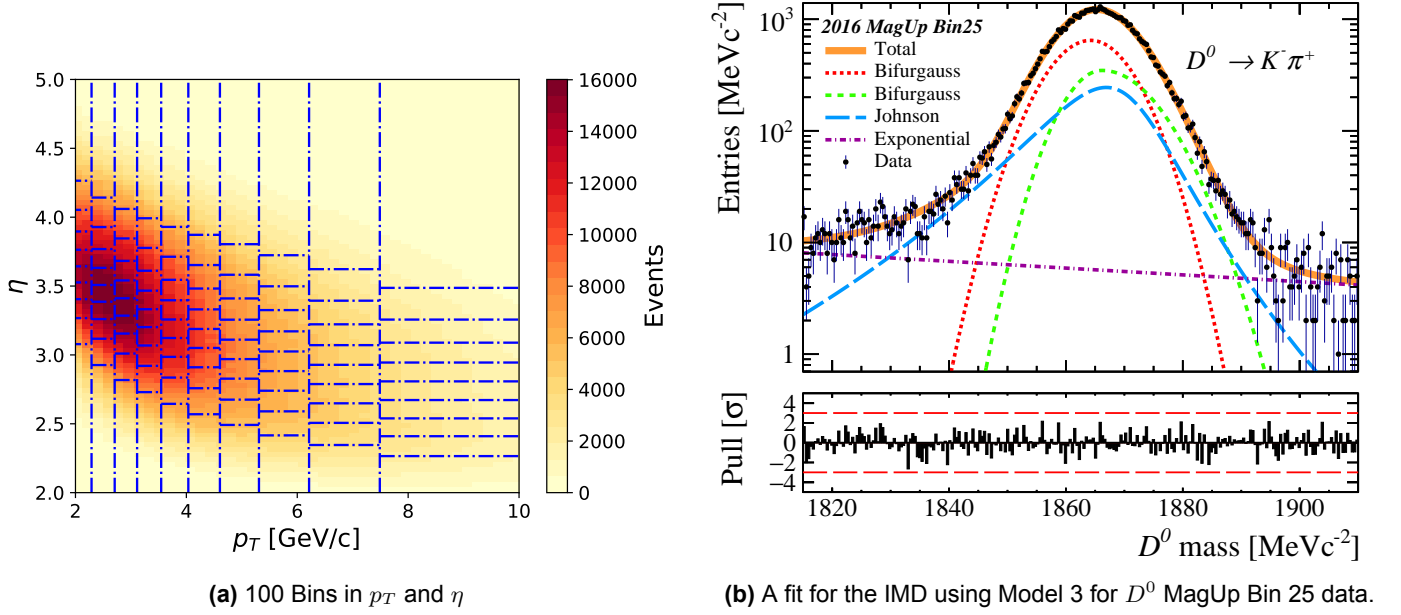
Fig. 13. (a) The binning scheme for pT and eta displays the bin boundaries over the data of D^0 MagUp. (b) An extended maximum-likelihood fit on the IMD is portrayed for the data in Bin 25 created by the binning scheme.

However, due to the smaller data set, the Poisson errors of the background are relatively large. As a result, a bad fit of the IMD's background can still lead to a good pull distribution. Furthermore, the model was fitted as being discontinuous at the peaks in a few bins, as seen in Appendix C. These issues can be addressed if the initial data set, before applying the binning scheme, had more events.

6.2 Systematic Uncertainties

This analysis explores four different systematic uncertainties: σ_{Model}, $\sigma_{Residual}$, $\sigma_{Bin-280}$, and $\sigma_{Bin-200}$. The systematic uncertainty is acquired by finding the absolute difference between the production asymmetry obtained from a set of assumptions and its A'_{prod} calculated using a different set of assumptions:

$$\sigma_{systematic} = \left| A_{prod} - A'_{prod} \right|. \tag{17}$$

The sensitivity of the results to different aspects of the fitting process can be assessed using this approach. The choice of model introduces a systematic uncertainty, σ_{Model}, and is probed by using Model 3 and Model 4. Additionally, a systematic uncertainty for the number of bins can be calculated by comparing the use of 240, 280, and 200 bins for the binned fit. The phase-space bin-integrated asymmetry is found by obtaining the signal yield for D^0 and $\bar{D}^0$ for the total data in a given dataset. Alternatively, the production asymmetry can be computed by averaging the asymmetry values for each phase-space bin after applying a binning scheme. This analysis employs the bin-integrated approach to measure the main result, while the second method is used to estimate a systematic uncertainty, $\sigma_{Residual}$.

The production asymmetry and the systematic uncertainties are calculated for each year and polarity, as shown in Appendix D. Therefore, an arithmetic average was performed to obtain a pseudo-bin-integrated for all three years combined and for each year separately. The uncertainty for each year's arithmetic averaged A_{prod} was calculated using quadrature by taking the uncertainty for the integrated A_{prod} for each polarity. The uncertainty for the total pseudo-bin-integrated A_{prod} is calculated by using the uncertainty of the A_{prod} for each year.

Year	Psuedo-Bin-Integrated A_{prod} [%]	σ_{Stat} [%]	Systematic Uncertainties [%]			
			σ_{Model}	$\sigma_{Residual}$	$\sigma_{Bin-200}$	$\sigma_{Bin-280}$
2016	-0.866	$\pm$ 0.204	$\pm$ 0.011	$\pm$ 0.020	$\pm$ 0.001	$\pm$ 0.004
2017	-0.376	$\pm$ 0.213	$\pm$ 0.008	$\pm$ 0.018	$\pm$ 0.001	$\pm$ 0.006
2018	-0.397	$\pm$ 0.219	$\pm$ 0.003	$\pm$ 0.018	$\pm$ 0.004	$\pm$ 0.009
Total	-0.546	$\pm$ 0.122	$\pm$ 0.005	$\pm$ 0.011	$\pm$ 0.001	$\pm$ 0.004

Table 3. A table displaying the results from the calculation of A_{prod} for each year were obtained by modelling the invariant mass distribution using Model 3.

7 Results and Analysis

The results computed for the production asymmetry and the systematic and statistical uncertainties are compared with those obtained by the previous analysis in Table 4. The values agree, but the statistical variation has slightly increased. The change in the fit methodology explains the increase. Specifically, the previous analysis used an unbinned fit, while a binned fit was chosen for the current analysis. The results were used to explore a dependence on kinematic variables.

Year	Pseudo-Bin-Intergrated A_{prod} [%]	Pseudo-Bin-Intergrated A_{prod}' [%]
2016	-0.866 $\pm$ 0.204 $\pm$ 0.013	-0.84 $\pm$ 0.18 $\pm$ 0.01
2017	-0.376 $\pm$ 0.213 $\pm$ 0.020	-0.39 $\pm$ 0.19 $\pm$ 0.01
2018	-0.397 $\pm$ 0.219 $\pm$ 0.013	-0.64 $\pm$ 0.20 $\pm$ 0.01
Total	-0.546 $\pm$ 0.122 $\pm$ 0.018	-0.62 $\pm$ 0.11 $\pm$ 0.01

Table 4. A table showing results from calculating A_{prod} for each year that was obtained by modelling the invariant mass distribution using Model 3 and Model 1. The results for Model 1 on the right are acquired from the previous analysis. An arithmetic average was performed on the results for each year to find the total asymmetry for both models.

7.1 Kinematic Relationship
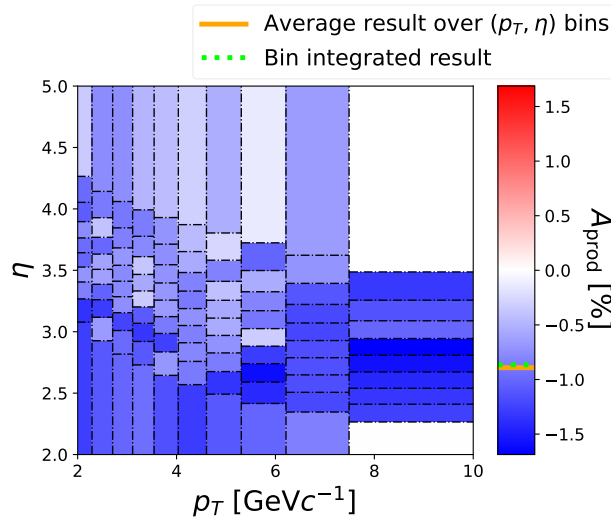
Fig. 14. A 2D histogram showing the pseudo production asymmetry using a colour scale in the 100 bins in the phase space of p_T and η for the 2016 data. The indicators on the colour bar show the result obtained for the bin-integrated result and the average result over the bins.

A 2D histogram was plotted for the production asymmetry values for the binning scheme with 100 bins in the phase space of p_T and η, as shown in Figure 14. The graphs illustrate that the production asymmetry is primarily negative and has a continuous look from white to blue. This correlates to values less than zero for the production asymmetry, as expected from the leading-particle effect.
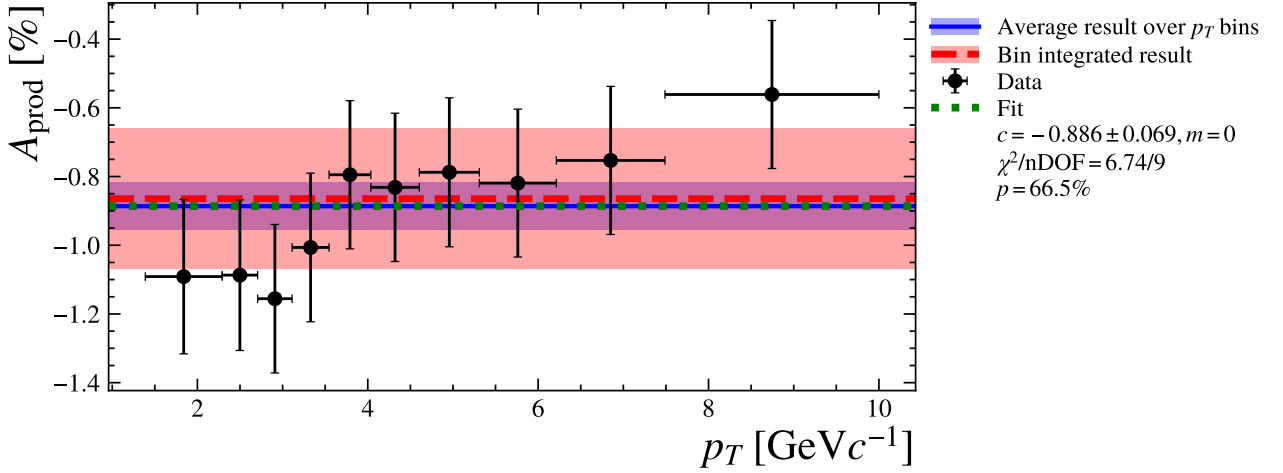
Fig. 15. A graph portraying the pseudo-bin-integrated A_{prod} for 2016 calculated in each bin after applying the binning scheme to create 10 p_T bins. The value for p_T is found by calculating the bin centre, and the error is half the bin width.

To investigate the relationship between the production asymmetry and p_T, a binning scheme for ten p_T bins was applied. The production asymmetry was calculated for each bin, and a constant and a linear function were explored to fit the data. The chi-squared, χ^2, is minimised to find the optimal parameters for the given model and is generated using `minuit`, a package for statistical modelling [40]. A hypothesis test was conducted to find which function fitted the data the best, with the null hypothesis, H_0, being that the data agrees with the fit. Using chi-squared statistics, a probability (p-value) is found for the likelihood of the observed data if the H_0 is valid. If the p-value is less than 0.05, we can reject the null hypothesis and conclude that the function does not represent the data.

Figure 15 illustrates the pseudo-bin-integrated production asymmetry calculated for each bin, and a constant function was fitted to the data. This procedure was repeated for both fit functions and all three years. A similar process was conducted to find a relationship in the phase space of η, as seen in Figure 16.
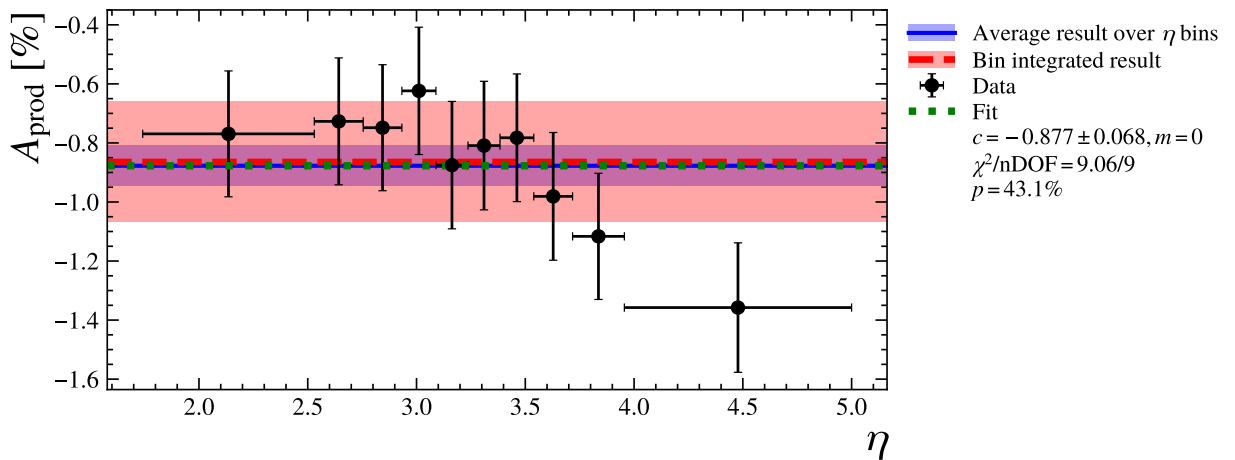
Fig. 16. A graph portraying the pseudo-bin-integrated A_{prod} for 2016 calculated in each bin after applying the binning scheme to create 10 η bins. The value for η is found by calculating the bin centre, and the error is half the bin width.

The result for the p-value (p) and reduced chi-squared χ_R^2 for each binning scheme using both fit functions are shown in Table 5. The table shows that the null hypothesis for all the tested relationships can not be rejected because the p-value is much greater than the significance level of 0.05. The p-value is

consistently higher for a linear relationship, but the much lower χ^2_R generated by the significant uncertainty on the data points suggests the linear model overfits the data. As a result, there may be a dependence, and an investigation conducted with more data could yield a valid fit for the kinematic dependence.

Year	Parameter	Constant Fit Function		Linear Fit Function	
		p	χ^2_R	p	χ^2_R
2016	p_T	66.5%	0.75	99.6%	0.15
2017	p_T	88.4%	0.49	100%	0.07
2018	p_T	93.5%	0.40	99.9%	0.10
2016	η	43.1%	1.01	94.3%	0.36
2017	η	90.6%	0.45	98.5%	0.23
2018	η	79.2%	0.61	93.3%	0.38

Table 5. A table displaying the χ^2_R and p-values for the tested relationships.

8 Conclusions and Future Work

This report presents the findings of the total pseudo-bin-integrated production asymmetry from the data collected from 2016, 2017 and 2018 by the LHCb detector. The key objective of computing the production asymmetry is accomplished by calculating the raw and detector asymmetry. The raw asymmetry was estimated by performing a maximum extended likelihood fit to find the signal yield. The fit is performed by representing the signal as a Johnson and two Bifurcated Gaussian functions and the combinatorial background as a decaying exponential. Meanwhile, the detector asymmetry was obtained by using the prior analysis result. The bin-integrated production asymmetry was computed for both polarities in each year. An arithmetic average was performed to calculate the pseudo-bin-integrated result for each year and for the total data analysed:

$$A_{\text{prod}}(D^0) = (-0.866 \pm 0.204 \pm 0.012)\% \quad (2016),$$
$$A_{\text{prod}}(D^0) = (-0.376 \pm 0.213 \pm 0.010)\% \quad (2017),$$
$$A_{\text{prod}}(D^0) = (-0.397 \pm 0.219 \pm 0.012)\% \quad (2018),$$
$$A_{\text{prod}}(D^0) = (-0.546 \pm 0.122 \pm 0.010)\% \quad (\text{Total}),$$

where the first uncertainty is statistical and the second is systematic uncertainty. The next step is to perform a bin-integrated analysis for each year and total rather than relying on the averaging of results obtained from the MagUp and MagDown datasets. The systematic uncertainty is generated by the model, the number of bins and the methodology used to calculate production asymmetry. These results indicate that more $\bar{D}^0$ mesons were produced relative to the D^0 meson, as expected from the leading particle effect.

A search for the dependence on production asymmetry on the kinematic variables using a linear fit did not yield an answer, as the reduced chi-squared was generally too low due to the large errors. However, reducing the uncertainty of detector asymmetry and using more data decreases the uncertainty of production asymmetry. This process enables us to explore a linear fit for the relationship. Furthermore, detector asymmetry substantially varies with phase space and could be used to vary detector asymmetry per bin for a more accurate result.

Moreover, the signal may consist of peaking Cabibo-suppressed decay modes, which could be fitted to the background to improve the quality of the measurements. Furthermore, implementing a binned fit technique permits future analysis to include substantially more events so that the statistical uncertainty is reduced and the sensitivity of the result is increased.

References

[1] N. A. Lowrey, "Analysis of the neutral D-meson decay to a neutral kaon and two neutral pions," Ph.D. dissertation, University of Illinois at Urbana-Champaign, 2010.

[2] L. Xu, "Heavy flavour production in pp collisions at LHCb," *Nuclear and Particle Physics Proceedings*, vol. 318-323, pp. 7–11, 2022.

[3] L. M. Kashif, "Measurement of the Z boson cross-section in the dimuon channel in pp collisions at $\sqrt{s}$ = 7 TeV," Ph.D. dissertation, Harvard U., 2010.

[4] A. Trišović, "Measuring the D^0 lifetime at the LHCb Masterclass," *Nuclear and Particle Physics Proceedings*, vol. 273-275, pp. 1215–1220, 2016.

[5] S. Stemmle, "Measurement of the time-integrated CP asymmetry in prompt D0 → K−K+ decays with the LHCb experiment ," University of Heidelberg, 2015.

[6] F. Ferrari, "Production asymmetries of b and c hadrons at LHCb," *Journal of Physics: Conference Series*, vol. 770, no. 1, Nov. 2016.

[7] G. Altarelli, "The Standard Model of Particle Physics," in *Encyclopedia of Mathematical Physics*, 2005.

[8] F. D. Lorenzi, "Parton Distribution Function Studies and a Measurement of Drell-Yan Produced Muon Pairs at LHCb ," Ph.D. dissertation, University College Dublin, 2011.

[9] C. Avila, *Charm particle production in hadronic collisions*, 2006.

[10] S. Chatrchyan, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC," *Physics Letters B*, vol. 716, no. 1, pp. 30–61, 2012.

[11] A. Petrukhin, "Measurements of the proton structure at HERA and their impact for LHC," *Nuclear Physics B - Proceedings Supplements*, vol. 207–208, pp. 45–48, Oct. 2010.

[12] T. Gershon, "CP violation in the Bsystem," *Reports on Progress in Physics*, vol. 80, no. 4, 2017.

[13] S. Forte and S. Carrazza, *Parton distribution functions*, 2020.

[14] M. S. Sozzi, "Charge Conjugation," in *Discrete Symmetries and CP Violation: From Experiment to Theory*, 2007.

[15] Sozzi, M. S., "Parity," in *Discrete Symmetries and CP Violation: From Experiment to Theory*, 2007.

[16] J. Bernabéu and F. Martínez-Vidal, "Time-Reversal Violation," *Annual Review of Nuclear and Particle Science*, vol. 65, no. 1, pp. 403–427, 2015.

[17] K. Kleinknecht, "Symmetry Violations and Quark Flavour Physics," pp. 55–61, 519–523, 2020.

[18] Y. Ahn, H.-Y. Cheng, and S. Oh, "Wolfenstein parametrization at higher order: Seeming discrepancies and their resolution," *Physics Letters B*, vol. 703, no. 5, pp. 571–575, 2011.

[19] K. Kleinknecht, "Weak Quark Mixing and the CKM Matrix," pp. 123–138, 2003.

[20] J. Bernabeu, "Discrete Symmetries CP, T, CPT," *Journal of Physics: Conference Series*, vol. 631, no. 1, p. 012 015, Jul. 2015.

[21] T. D. K. Jr., "Analysis of Neutral D Meson Two-Body Decays to a Neutral Kaon and a Neutral Pion," Ph.D. dissertation, Virginia Polytechnic Institute and State University, 2021.

[22] R. L. Workman and Burkert, "Review of Particle Physics," *Progress of Theoretical and Experimental Physics*, vol. 2022, no. 8, p. 083C01, Aug. 2022.

[23] R. Aaij, "Observation of CP Violation in Charm Decays," *Phys. Rev. Lett.*, vol. 122, 21 2019.

[24] A. R., "Measurement of the Time-Integrated CP Asymmetry in $D^0 \to K^- K^+$ Decays," *Phys. Rev. Lett.*, vol. 131, p. 091 802, 9 Aug. 2023.

[25] R. Aaij, "Measurement of CP asymmetry in D 0 → K − K + and D 0 → π − π + decays," *Journal of High Energy Physics*, vol. 2014, no. 7, Jul. 2014.

[26] R. Vogt, "Heavy flavor azimuthal correlations in cold nuclear matter," *Physical Review C*, vol. 98, 2018.

[27] E. Braaten, Y. Jia, and T. Mehen, "Leading-Particle Effect from Heavy-Quark Recombination," *Phys. Rev. Lett.*, vol. 89, p. 122 002, 12 Aug. 2002.

[28] W. Lai, "D± Production Asymmetry at the LHC from Heavy-Quark Recombination," *Physical Review D*, vol. 90, no. 5, Sep. 2014.

[29] C. Peterson, D. Schlatter, I. Schmitt, and P. M. Zerwas, "Scaling Violations in Inclusive e+ e- Annihilation Spectra," *Phys. Rev. D*, vol. 27, p. 105, 1983.

[30] The LHCb Collaboration, "The LHCb Detector at the LHC," *Journal of Instrumentation*, vol. 3, 2008.

[31] L. Collaboration, "LHCb detector performance," *International Journal of Modern Physics A*, vol. 30, no. 07, 2015.

[32] C. Jarvis-Stiggants, "Measurement of the $\bar{D}^0 - D^0$ production asymmetry in pp collisions at $\sqrt{s} = 13\,\mathrm{TeV}$," University of Manchester, 2023.

[33] M. England, "Measurement of the production asymmetry of D^0 mesons in pp collisions at the lhcb experiment," University of Manchester, 2023.

[34] A. Davis, "Measurement of the instrumental asymmetry for $K^- \pi^+$-pairs at LHCb in Run 2," 2018.

[35] S. Stahl, "Search for CP violation in $D^0 \to K^- K^+, \pi^- \pi^+$ using semileptonic B decays on 3 fb$^-1$," 2014.

[36] A. Roodman, *Blind Analysis in Particle Physics*, 2003.

[37] R. Barlow, "Extended maximum likelihood," *Nuclear Instruments and Methods in Physics Research*, vol. 297, no. 3, pp. 496–506, 1990.

[38] W. Verkerke and D. P. Kirkby, "The RooFit toolkit for data modeling," 2003.

[39] P. Virtanen, "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python," *Nature Methods*, vol. 17, pp. 261–272, 2020.

[40] F. James, "MINUIT: Function Minimization and Error Analysis Reference Manual," 1998.

Appendices

A Results of the Models of the Invariant Mass Distribution.

A.1 Model 1

Year	Meson	Polarity	μ	σ
2016	D0	MagUp	-0.06 ± 0.54	3.15 ± 0.68
		MagDown	-0.33 ± 0.13	1.30 ± 0.13
	D0bar	MagUp	-0.91 ± 0.97	3.28 ± 1.55
		MagDown	-0.14 ± 0.16	1.51 ± 0.16
2017	D0	MagUp	-0.09 ± 0.19	1.72 ± 0.20
		MagDown	-0.52 ± 0.53	3.21 ± 0.68
	D0bar	MagUp	-0.78 ± 0.69	3.75 ± 1.89
		MagDown	-0.89 ± 0.97	4.15 ± 1.47
2018	D0	MagUp	-0.53 ± 1.90	4.07 ± 1.72
		MagDown	-0.28 ± 0.21	1.92 ± 0.21
	D0bar	MagUp	-0.57 ± 0.66	3.66 ± 0.92
		MagDown	-0.62 ± 0.69	3.24 ± 0.87
Average			-0.478 ± 2.75	2.91 ± 3.71

Table 6. Table showing the μ and σ of the fit for each year, and polarity obtained by modelling the invariant mass distribution using an unbinned fit of Model 1

A.2 Model 2

Year	Meson	Polarity	μ	σ
2016	D0	MagUp	0.02 ± 0.12	1.14 ± 0.12
		MagDown	-0.21 ± 0.12	1.36 ± 0.12
	D0bar	MagUp	-0.22 ± 0.15	1.51 ± 0.15
		MagDown	- 0.21 ± 0.13	1.23 ± 0.13
2017	D0	MagUp	-0.12 ± 0.11	1.28 ± 0.11
		MagDown	-0.17 ± 0.09	1.28 ± 0.09
	D0bar	MagUp	-0.17 ± 0.09	1.30 ± 0.09
		MagDown	-0.39 ± 0.13	1.23 ± 0.13
2018	D0	MagUp	-0.04 ± 0.09	1.01 ± 0.09
		MagDown	-0.22 ± 0.13	1.41 ± 0.13
	D0bar	MagUp	-0.15 ± 0.13	1.25 ± 0.13
		MagDown	-0.05 ± 0.12	1.34 ± 0.12
Average			-0.161 ± 0.442	1.27 ± 0.412

Table 7. Table showing the μ and σ of the fit for each year, and polarity obtained by modelling the invariant mass distribution using an unbinned fit of Model 2.

A.3 Model 3

Year	Meson	Polarity	μ	σ
2016	D0	MagUp	-0.06 ± 0.07	1.19 ± 0.07
		MagDown	-0.20 ± 0.09	1.40 ± 0.09
	D0bar	MagUp	-0.16 ± 0.07	1.32 ± 0.07
		MagDown	-0.14 ± 0.05	1.18 ± 0.05
2017	D0	MagUp	-0.11 ± 0.07	1.34 ± 0.07
		MagDown	-0.06 ± 0.06	1.30 ± 0.06
	D0bar	MagUp	-0.07 ± 0.07	1.23 ± 0.07
		MagDown	-0.08 ± 0.06	1.17 ± 0.06
2018	D0	MagUp	-0.13 ± 0.09	1.28 ± 0.09
		MagDown	-0.19 ± 0.06	1.17 ± 0.06
	D0bar	MagUp	0.02 ± 0.06	1.13 ± 0.06
		MagDown	-0.17 ± 0.06	1.03 ± 0.06
Average			-0.12 ± 0.24	1.23 ± 0.24

Table 8. Table showing the μ and σ of the fit for each year, and polarity obtained by modelling the invariant mass distribution using a binned fit of Model 3. The fit was also run over a data set containing all the years.

A.4 Model 4

Year	Meson	Polarity	μ	σ
2016	D0	MagUp	0.05 ± 0.08	1.19 ± 0.08
		MagDown	-0.12 ± 0.08	1.31 ± 0.08
	D0bar	MagUp	-0.05 ± 0.08	1.37 ± 0.08
		MagDown	-0.18 ± 0.07	1.28 ± 0.07
2017	D0	MagUp	0.00 ± 0.06	1.32 ± 0.06
		MagDown	0.03 ± 0.08	1.22 ± 0.08
	D0bar	MagUp	-0.05 ± 0.08	1.20 ± 0.08
		MagDown	-0.10 ± 0.06	1.23 ± 0.06
2018	D0	MagUp	-0.10 ± 0.08	1.49 ± 0.08
		MagDown	-0.19 ± 0.06	1.16 ± 0.06
	D0bar	MagUp	-0.02 ± 0.09	1.32 ± 0.09
		MagDown	-0.28 ± 0.08	1.14 ± 0.08
Average			-0.08 ± 0.26	1.27 ± 0.26

Table 9. Table showing the μ and σ of the fit for each year, and polarity obtained by modelling the invariant mass distribution using a binned fit of Model 4.

B Graphs of the Invariant Mass Distribution
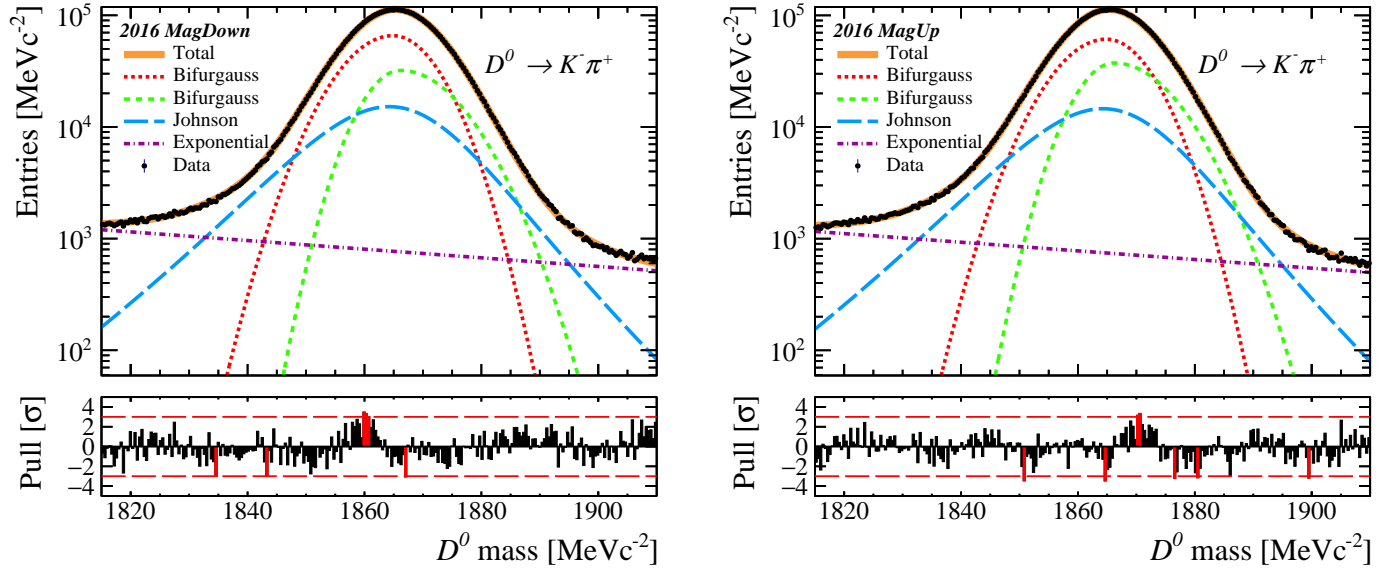
Fig. 17. Invariant mass distributions of D^0 meson, using the 2016 data and fitted using Model 3. These were used to extract the number of events under the signal, N, to calculate A_{raw} and A_{prod}.
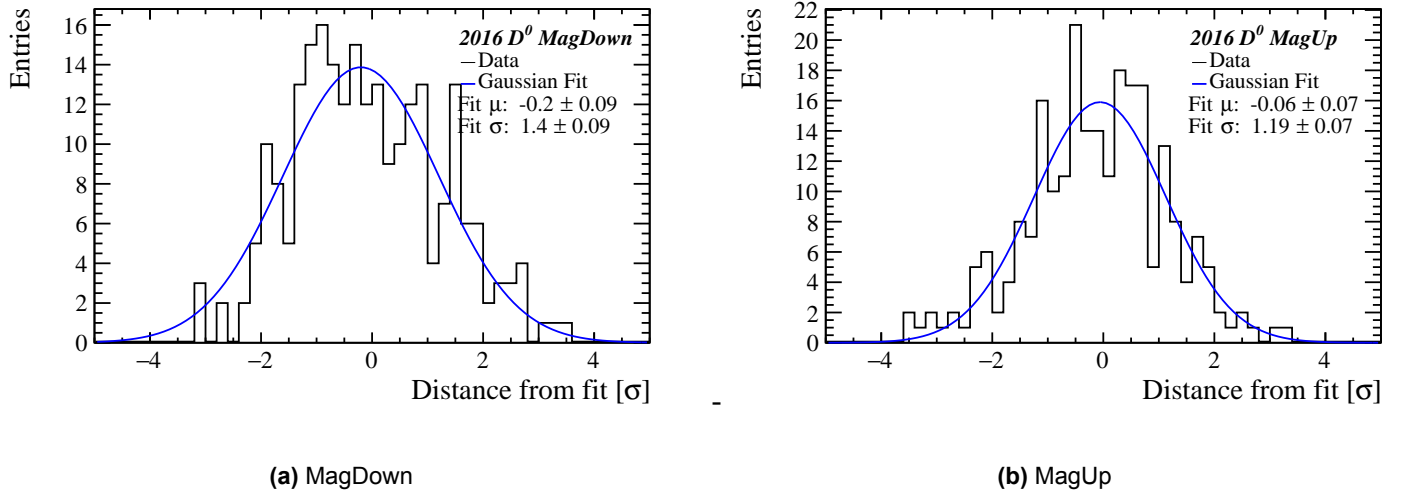
Fig. 18. Pull distributions of D^0 meson, using the 2016 data and modelled using Model 3. These were used to asses the model, graphically and analytically using the Gaussian fit μ and σ.
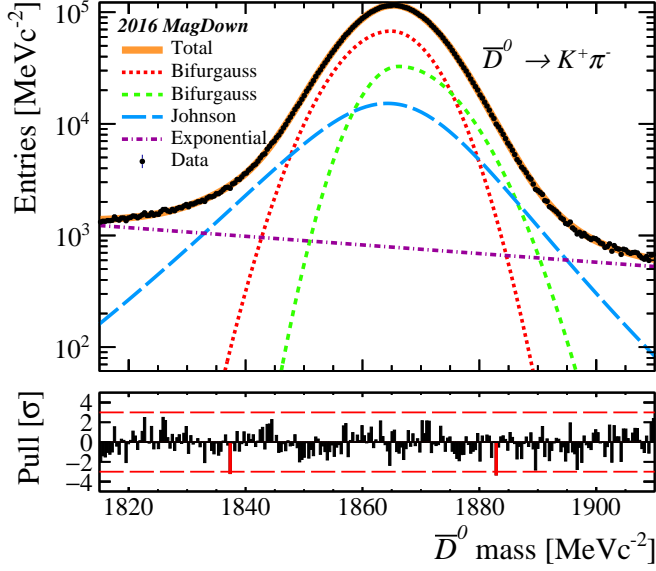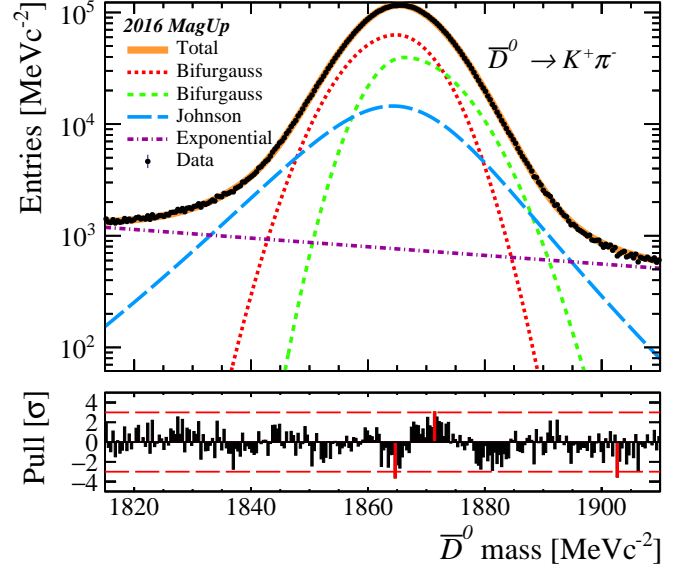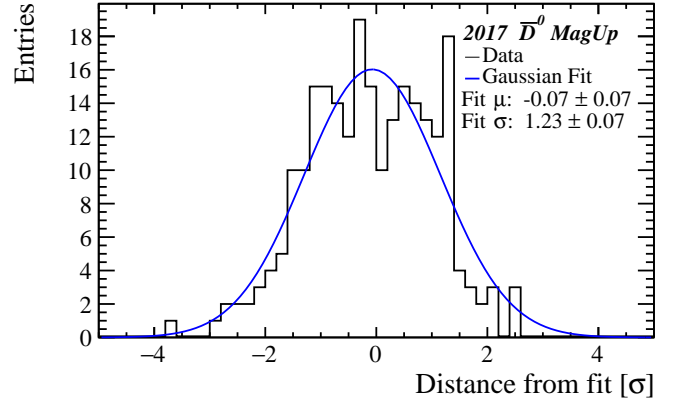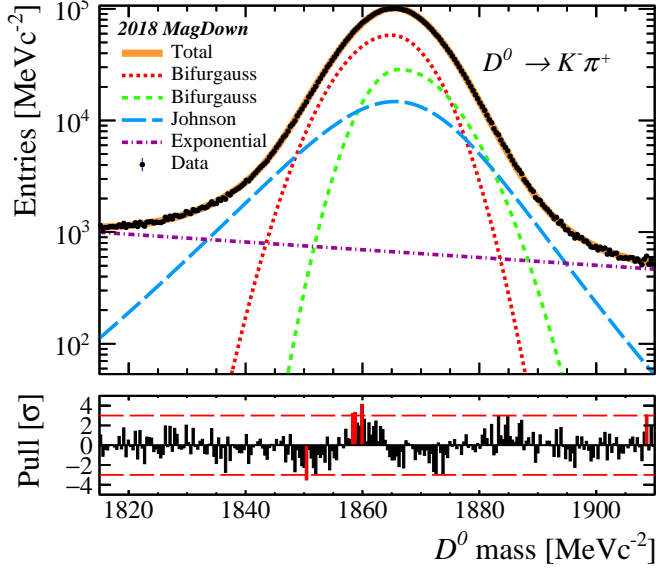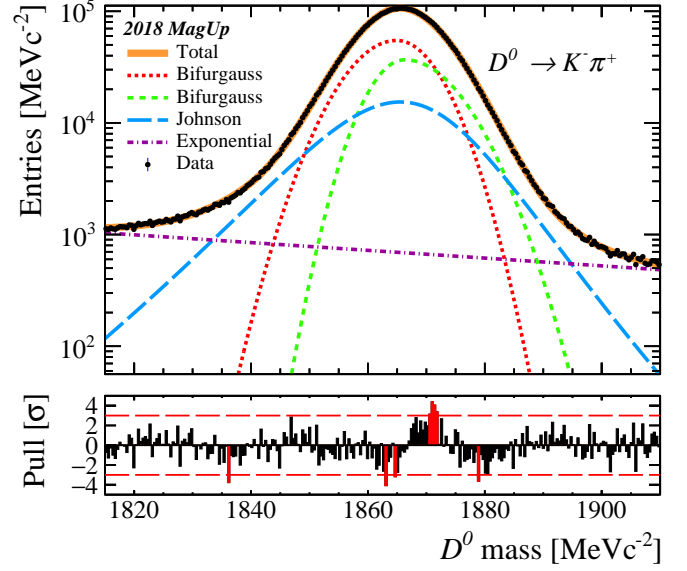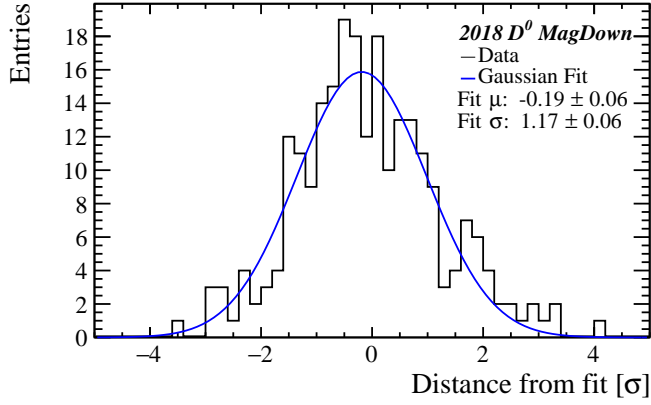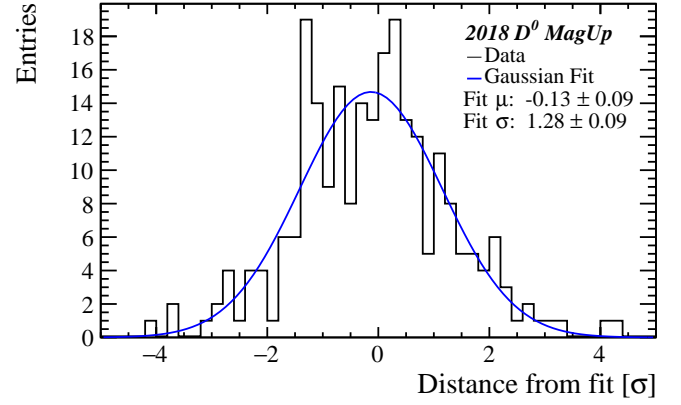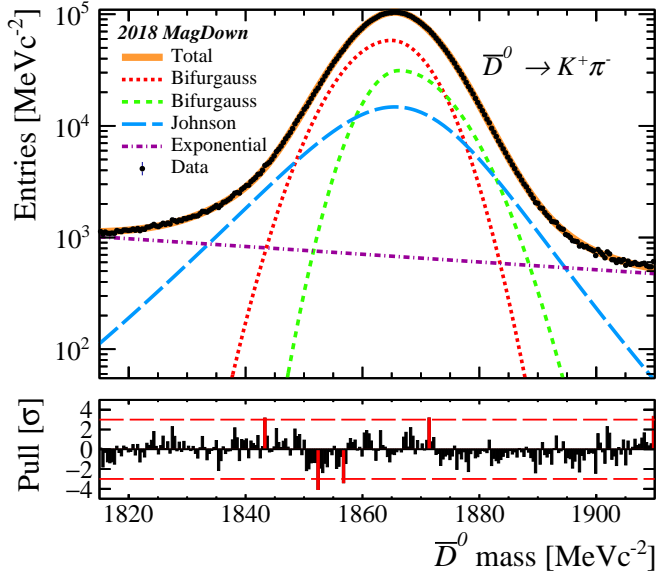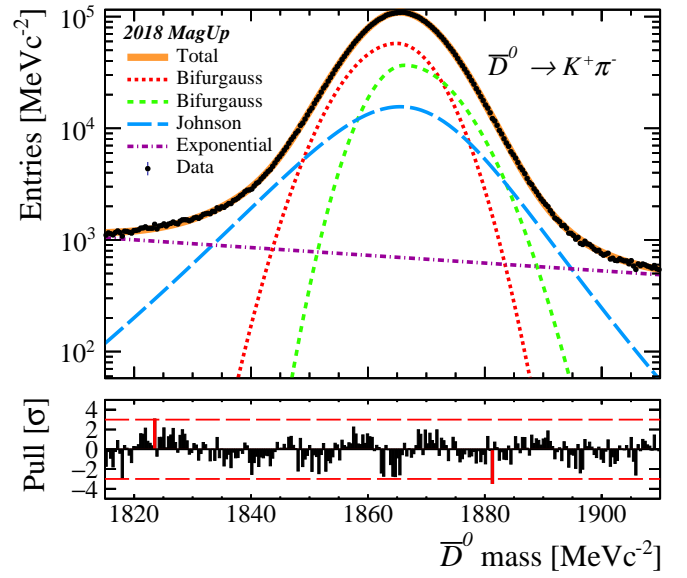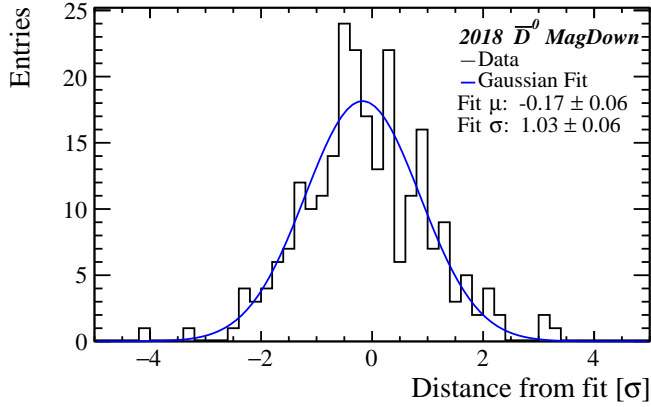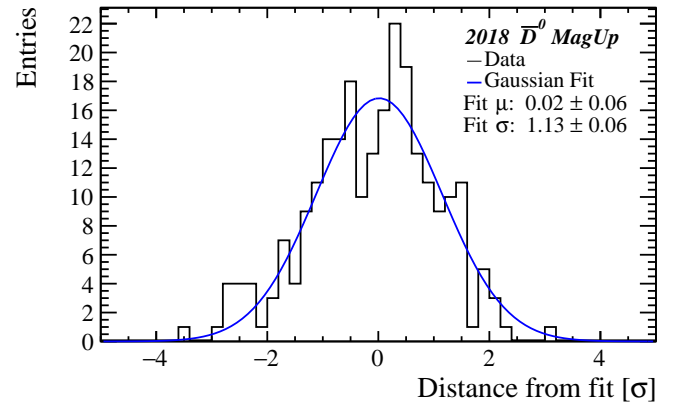
(a) MagDown

(b) MagUp

Fig. 19. Invariant mass distributions of $\bar{D}^0$ meson, using the 2016 data and fitted using Model 3. These were used to extract the number of events under the signal, $\bar{N}$, to calculate A_{raw} and A_{prod}.
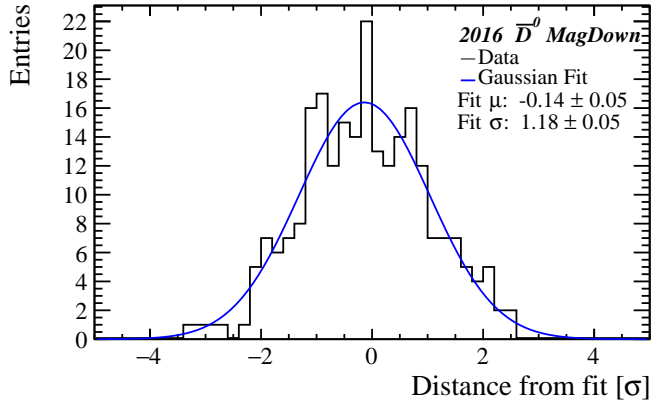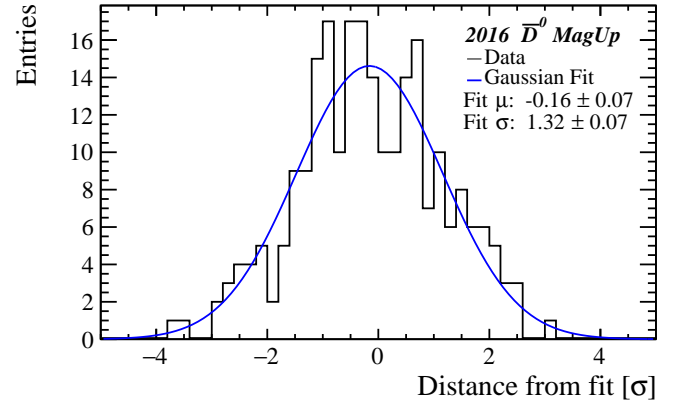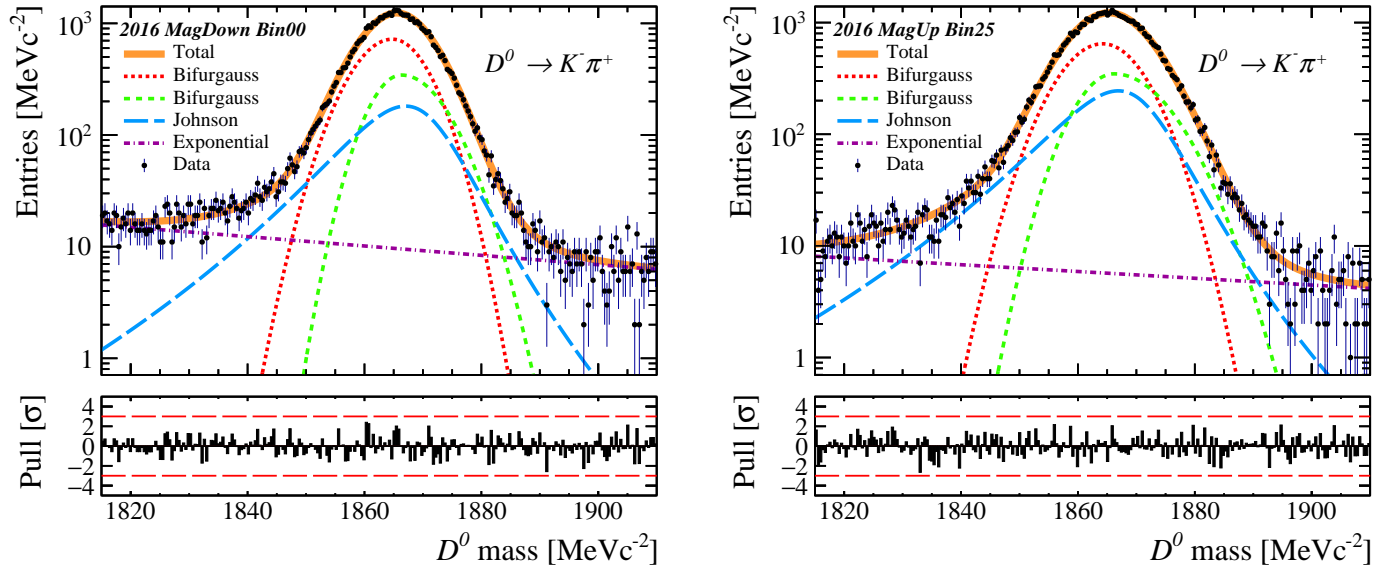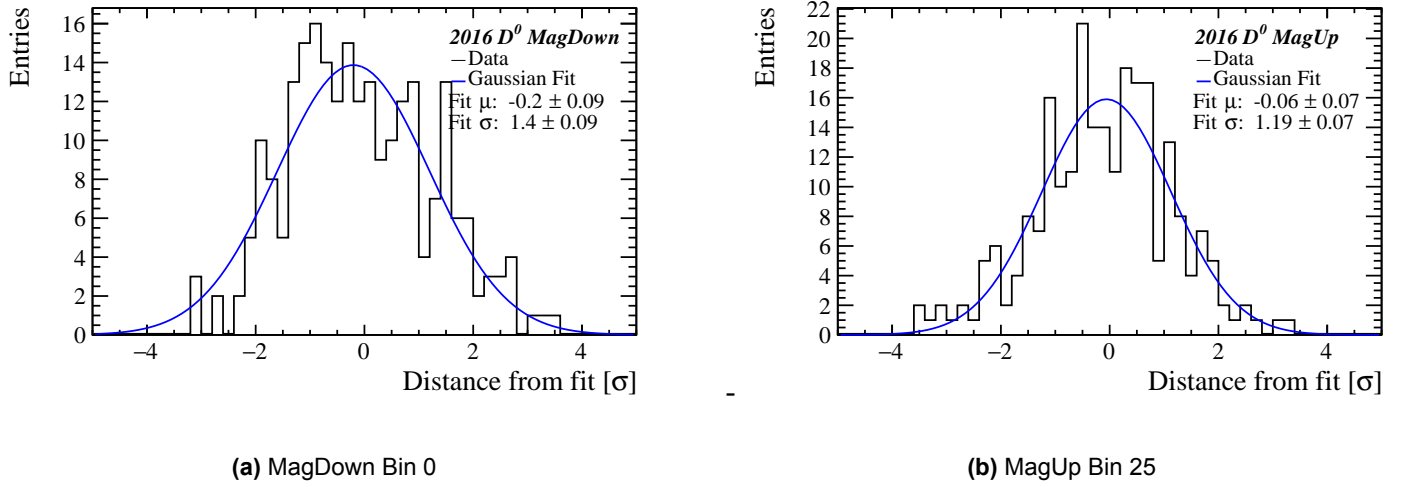

(a) MagDown

(b) MagUp

Fig. 20. Pull distributions of $\bar{D}^0$ meson, using the 2016 data and modelled using Model 3. These were used to asses the model, graphically and analytically using the Gaussian fit μ and σ.
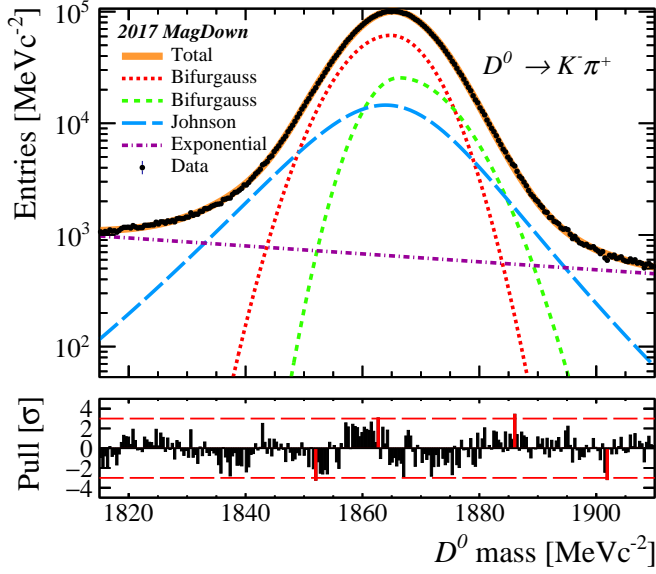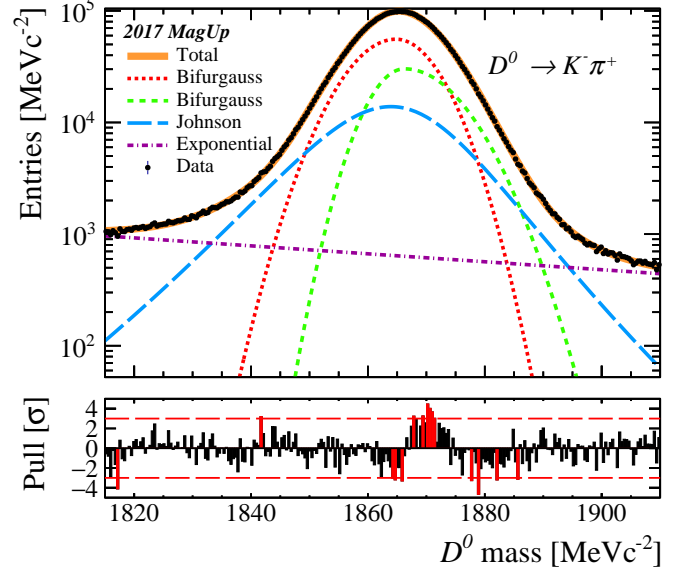
(a) MagDown

(b) MagUp

Fig. 21. Invariant mass distributions of D^0 meson, using the 2017 data and fitted using Model 3. These were used to extract the number of events under the signal, N, to calculate A_{raw} and A_{prod}.
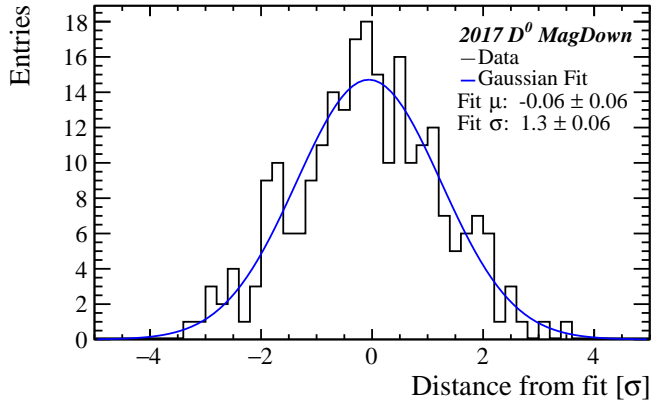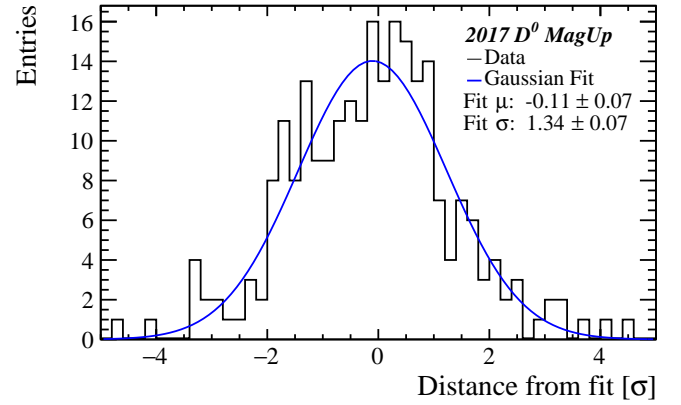
(a) MagDown

(b) MagUp

Fig. 22. Pull distributions of D^0 meson, using the 2017 data and modelled using Model 3. These were used to asses the model, graphically and analytically using the Gaussian fit μ and σ.
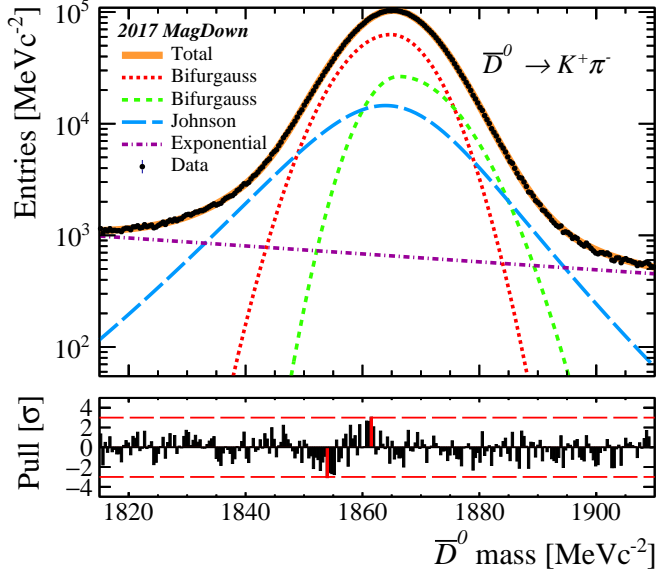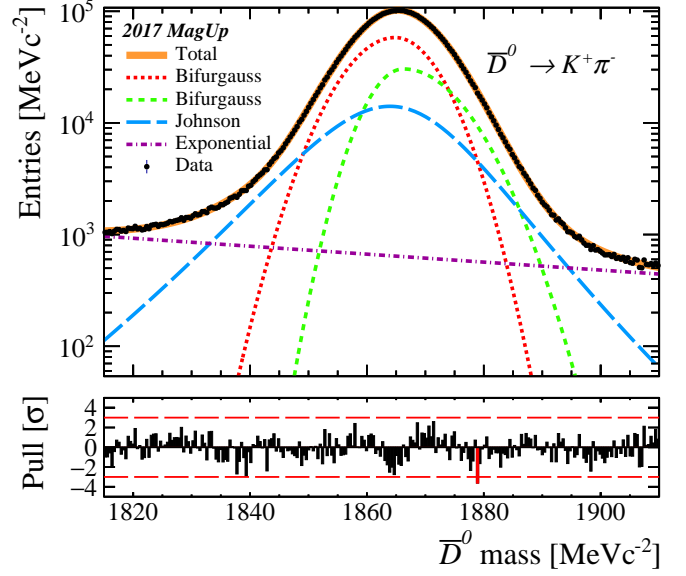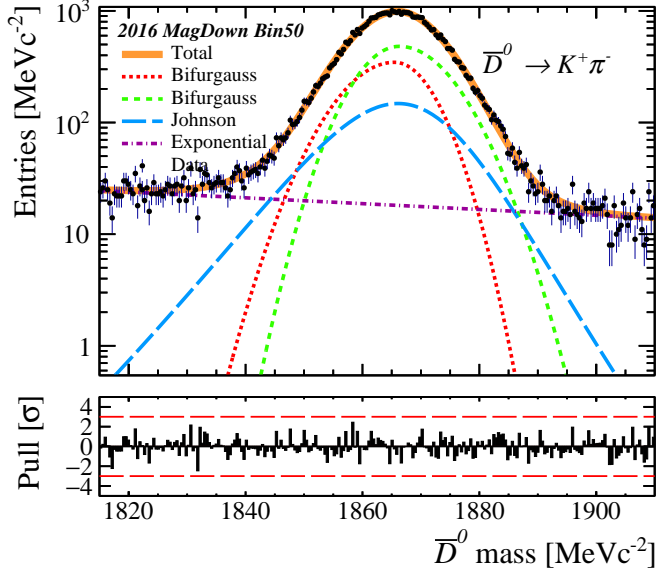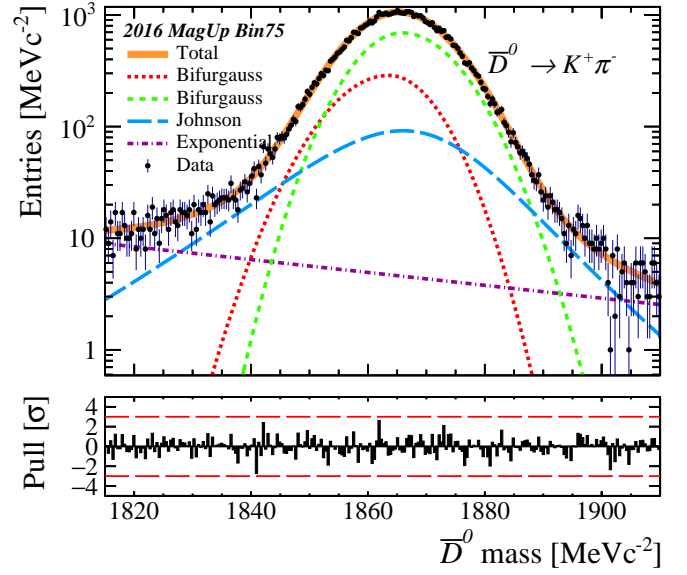
(a) MagDown

(b) MagUp

Fig. 23. Invariant mass distributions of $\bar{D}^0$ meson, using the 2017 data and fitted using Model 3. These were used to extract the number of events under the signal, $\bar{N}$, to calculate A_{raw} and A_{prod}.
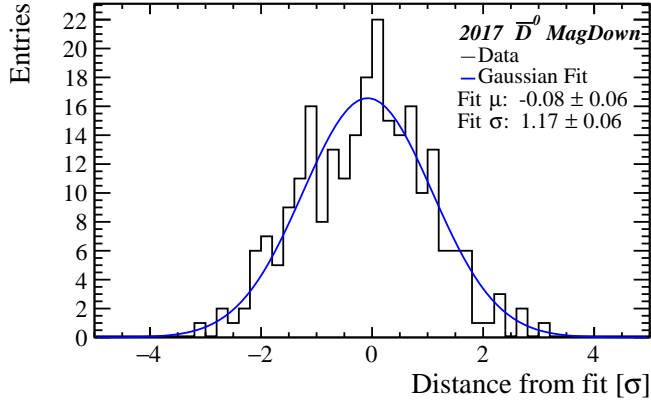
(a) MagDown

(b) MagUp

Fig. 24. Pull distributions of $\bar{D}^0$ meson, using the 2017 data and modelled using Model 3. These were used to asses the model, graphically and analytically using the Gaussian fit μ and σ.
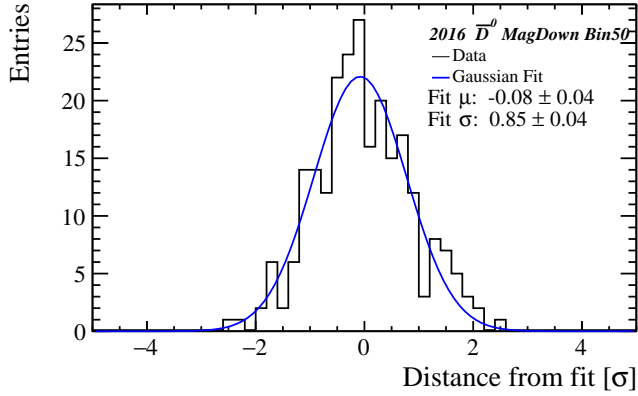
(a) MagDown

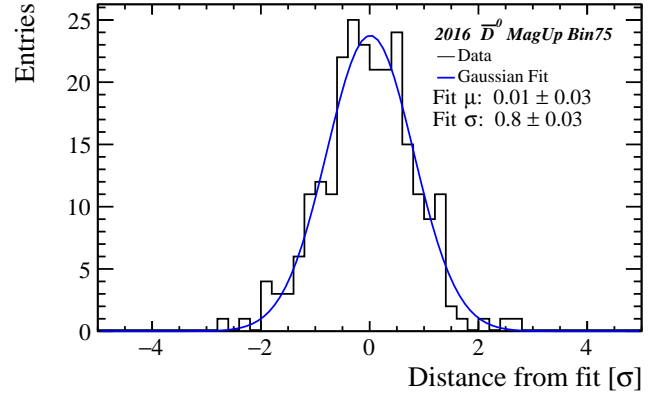**(b)** MagUp

Fig. 25. Invariant mass distributions of D^0 meson, using the 2018 data and fitted using Model 3. These were used to extract the number of events under the signal, N, to calculate A_{raw} and A_{prod}.

(a) MagDown

(b) MagUp

Fig. 26. Pull distributions of D^0 meson, using the 2018 data and modelled using Model 3. These were used to asses the model, graphically and analytically using the Gaussian fit μ and σ.

(a) MagDown

(b) MagUp

Fig. 27. Invariant mass distributions of $\bar{D}^0$ meson, using the 2018 data and fitted using Model 3. These were used to extract the number of events under the signal, $\bar{N}$, to calculate A_{raw} and A_{prod}.
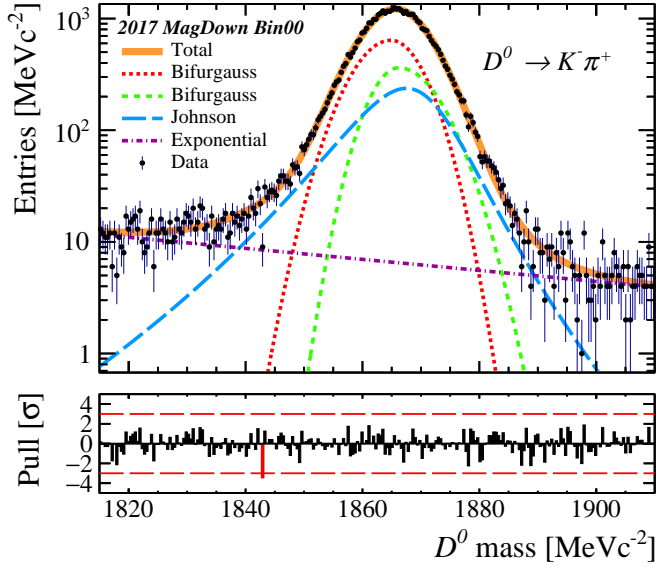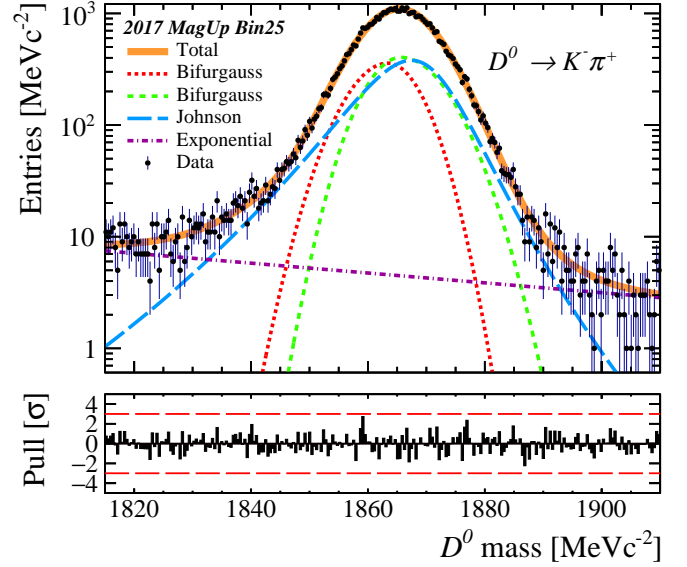

(a) MagDown

(b) MagUp

Fig. 28. Pull distributions of $\bar{D}^0$ meson, using the 2018 data and modelled using Model 3. These were used to asses the model, graphically and analytically using the Gaussian fit μ and σ.
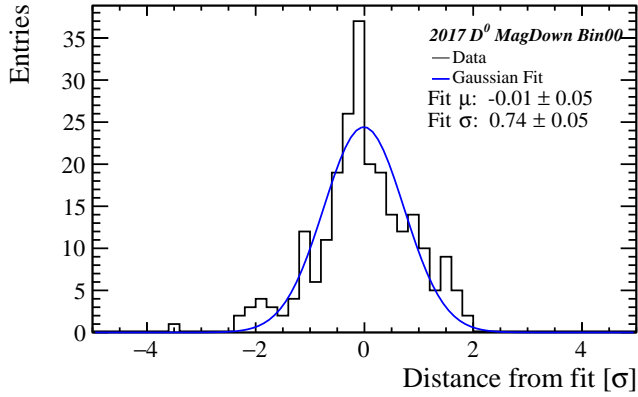
C Models of the Invariant Mass Distribution for Bins
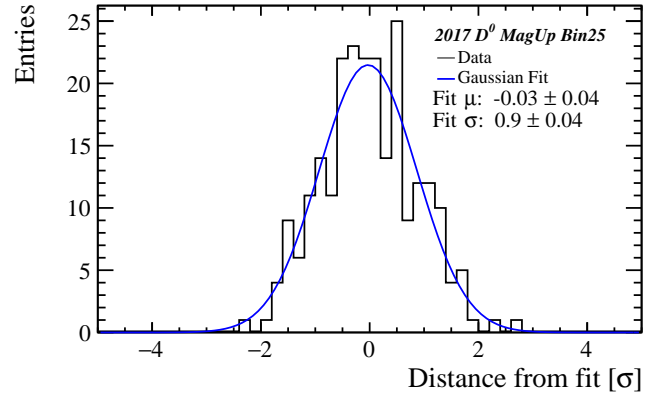

(a) MagDown Bin 0

(b) MagUp Bin 25

Fig. 29. Invariant mass distributions of D^0 meson, using the 2016 data and fitted using Model 1. These were used to extract the number of events under the signal, N, to calculate A_{raw} and A_{prod} in Bin 0 and Bin 25.
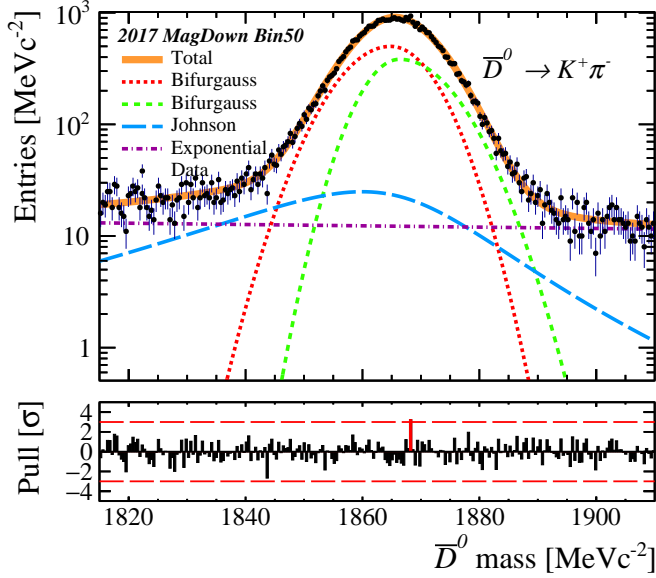

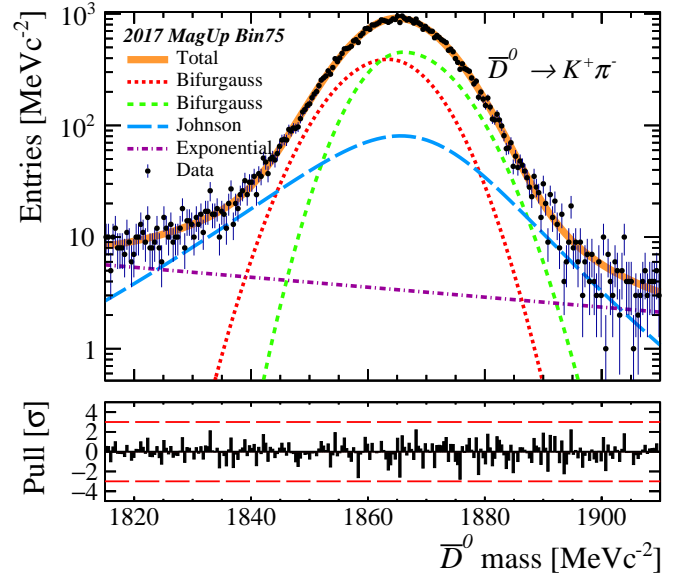**(a)** MagDown Bin 0

(b) MagUp Bin 25

Fig. 30. Pull distributions of D^0 meson, using the 2016 data and modelled using Model 1. These were used to assess the model, graphically and analytically, using the Gaussian fit μ and σ.

(a) MagDown Bin 50

(b) MagUp Bin 75

Fig. 31. Invariant mass distributions of $\bar{D}^0$ meson, using the 2016 data and fitted using Model 1. These were used to extract the number of events under the signal, $\bar{N}$, to calculate A_{raw} and A_{prod}.

(a) MagDown Bin 50

(b) MagUp Bin 75

Fig. 32. Pull distributions of $\bar{D}^0$ meson, using the 2016 data and modelled using Model 1. These were used to asses the model, graphically and analytically using the Gaussian fit μ and σ.
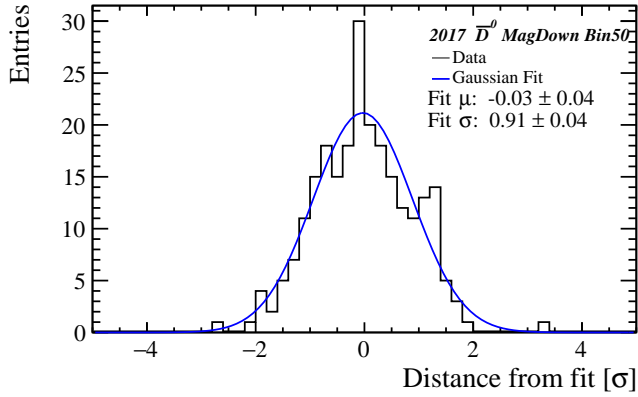
(a) MagDown Bin 0

(b) MagUp Bin 25

Fig. 33. Invariant mass distributions of D^0 meson, using the 2017 data and fitted using Model 1. These were used to extract the number of events under the signal, N, to calculate A_{raw} and A_{prod} in Bin 0 and Bin 25.

(a) MagDown Bin 0

(b) MagUp Bin 25

Fig. 34. Pull distributions of D^0 meson, using the 2017 data and modelled using Model 1. These were used to asses the model, graphically and analytically using the Gaussian fit μ and σ.
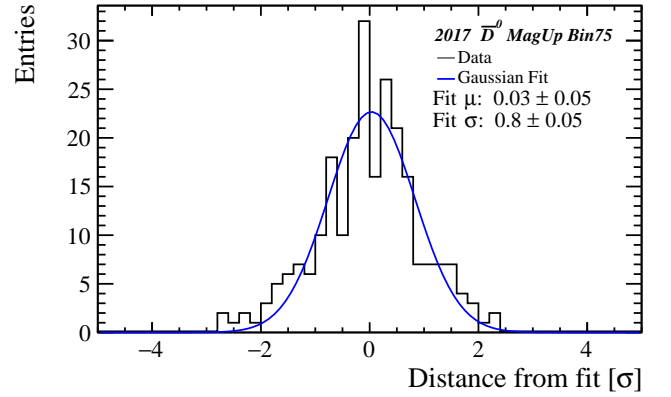
(a) MagDown Bin 50

(b) MagUp Bin 75

Fig. 35. Invariant mass distributions of $\bar{D}^0$ meson, using the 2017 data and fitted using Model 1. These were used to extract the number of events under the signal, $\bar{N}$, to calculate A_{raw} and A_{prod}.
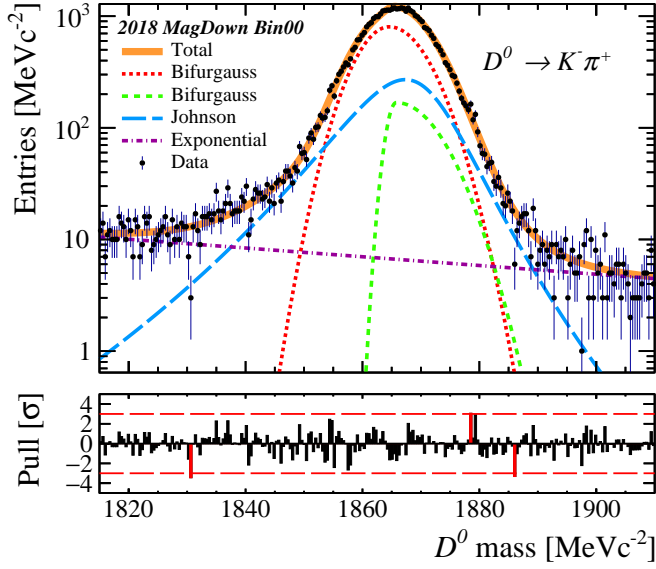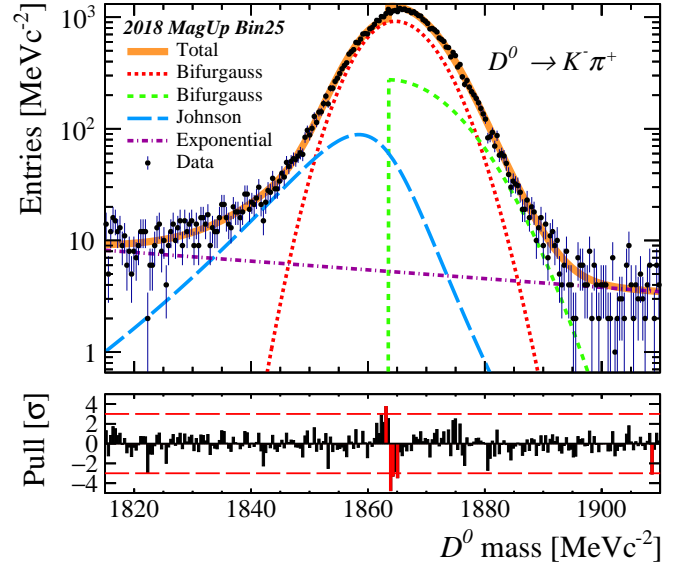

(a) MagDown Bin 50

(b) MagUp Bin 75

Fig. 36. Pull distributions of $\bar{D}^0$ meson, using the 2017 data and modelled using Model 1. These were used to asses the model, graphically and analytically using the Gaussian fit μ and σ.
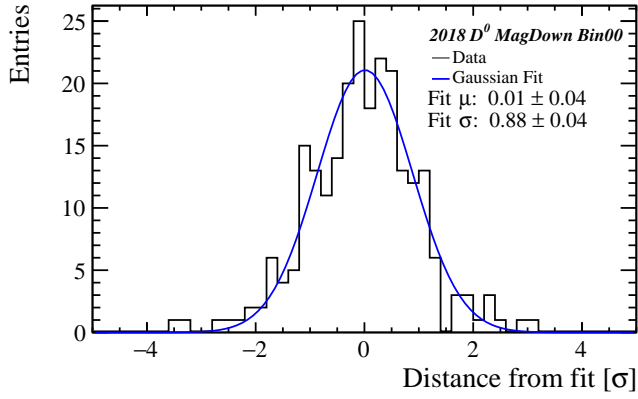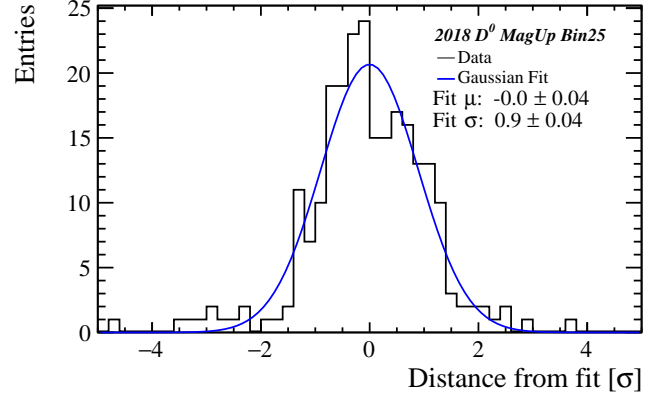
(a) MagDown Bin 0

(b) MagUp Bin 25

Fig. 37. Invariant mass distributions of D^0 meson, using the 2018 data and fitted using Model 1. These were used to extract the number of events under the signal, N, to calculate A_{raw} and A_{prod} in Bin 0 and Bin 25.
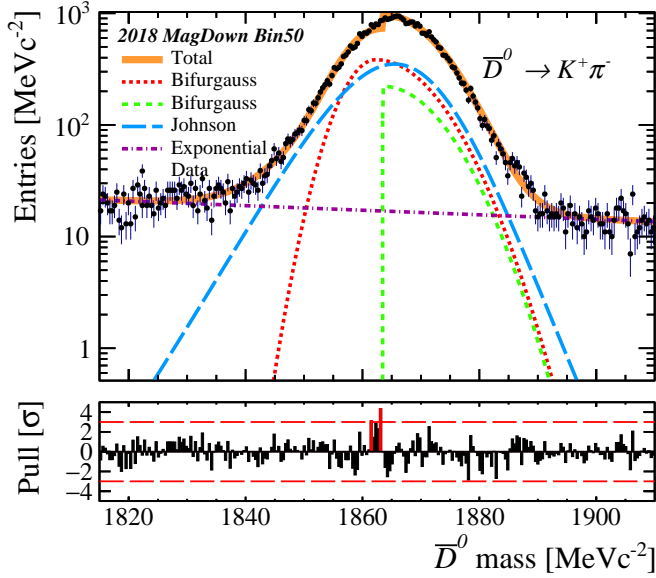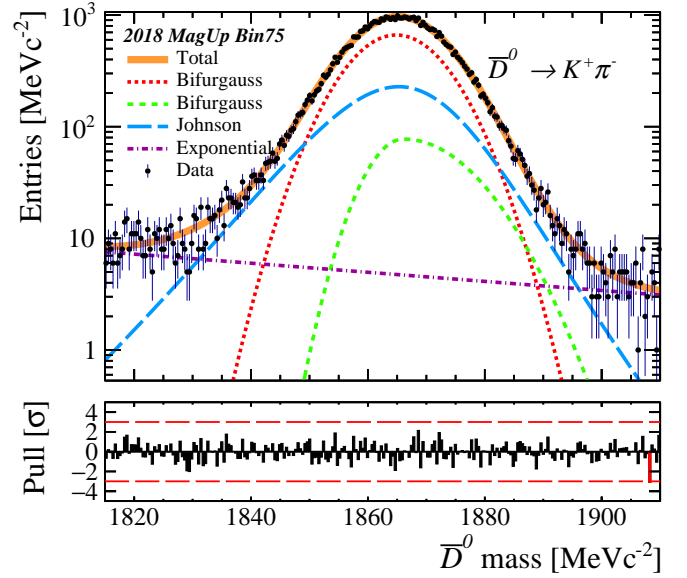

(a) MagDown Bin 0

(b) MagUp Bin 25

Fig. 38. Pull distributions of D^0 meson, using the 2018 data and modelled using Model 1. These were used to asses the model, graphically and analytically using the Gaussian fit μ and σ.

(a) MagDown Bin 50

(b) MagUp Bin 75

Fig. 39. Invariant mass distributions of $\bar{D}^0$ meson, using the 2018 data and fitted using Model 1. These were used to extract the number of events under the signal, $\bar{N}$, to calculate A_{raw} and A_{prod}.

(a) MagDown Bin 50

(b) MagUp Bin 75

Fig. 40. Pull distributions of $\bar{D}^0$ meson, using the 2018 data and modelled using Model 1. These were used to asses the model, graphically and analytically using the Gaussian fit μ and σ.
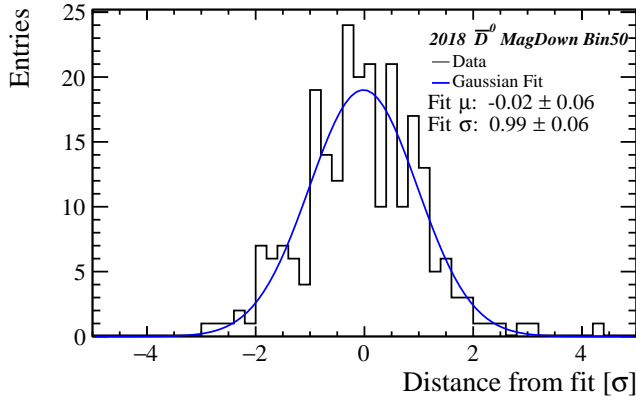
D Production Asymmetry
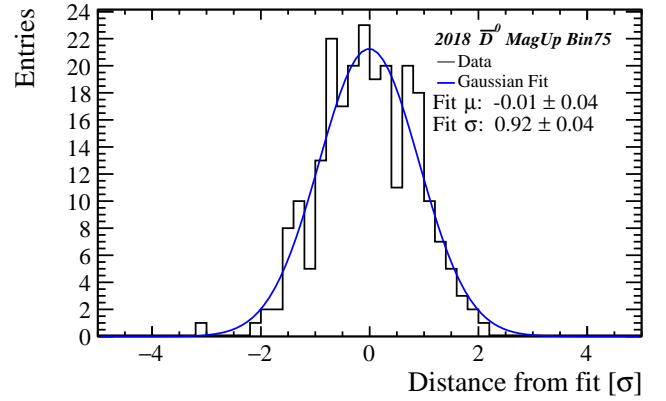
Year	Polarity	Intergrated A_{prod} [%]	Intergrated A_{prod}' [%]	$\sigma_{Residual}$
2016	MagUp	-0.983 ± 0.298	-1.020 ± 0.043	0.037
	MagDown	-0.746 ± 0.279	-0.763 ± 0.042	0.017
2017	MagUp	-0.595 ± 0.302	-0.592 ± 0.045	0.003
	MagDown	-0.157 ± 0.301	-0.193 ± 0.044	0.036
2018	MagUp	-0.652 ± 0.303	-0.661 ± 0.044	0.009
	MagDown	-0.156 ± 0.318	-0.190 ± 0.046	0.034

Table 10. Results from calculating A_{prod} for each year and polarity, obtained by modelling the invariant mass distribution using Model 4. The integrated A_{prod} and Average A_{prod} are used to calculate a systematic error.

Year	Polarity	240 Bins A_{prod} [%]	200 Bins A_{prod} [%]	$\sigma_{Bins-200}$
2016	MagUp	-0.983 ± 0.298	-0.984 ± 0.298	0.001
	MagDown	-0.746 ± 0.279	-0.747 ± 0.279	0.001
2017	MagUp	-0.595 ± 0.302	-0.596 ± 0.302	0.001
	MagDown	-0.157 ± 0.301	-0.158 ± 0.300	0.001
2018	MagUp	-0.652 ± 0.303	-0.645 ± 0.303	0.007
	MagDown	-0.156 ± 0.318	-0.160 ± 0.318	0.004

Table 11. Results from calculating A_{prod} for each year and polarity obtained by modelling the invariant mass distribution using Model 4 with 240 and 200 bins. The integrated A_{prod} calculated using Model 4 with 240 and 200 bins are used to calculate a systematic error.

Year	Polarity	240 Bins A_{prod} [%]	280 Bins A_{prod} [%]	$\sigma_{Bins-280}$
2016	MagUp	-0.983 ± 0.298	-0.975 ± 0.298	0.008
	MagDown	-0.746 ± 0.279	-0.747 ± 0.279	0.001
2017	MagUp	-0.595 ± 0.302	-0.583 ± 0.302	0.012
	MagDown	-0.157 ± 0.301	-0.153 ± 0.301	0.004
2018	MagUp	-0.652 ± 0.303	-0.635 ± 0.303	0.017
	MagDown	-0.156 ± 0.318	-0.150 ± 0.318	0.006

Table 12. Results from calculating A_{prod} for each year and polarity obtained by modelling the invariant mass distribution using Model 4 with 240 and 280 bins. The integrated A_{prod} calculated using Model 4 with 240 and 280 bins are used to calculate a systematic error.

Year	Polarity	Intergrated A_{prod} [%]	Intergrated A_{prod}' [%]	σ_{Model}
2016	MagUp	-0.983 ± 0.298	-0.988 ± 0.298	0.005
	MagDown	-0.746 ± 0.279	-0.767 ± 0.279	0.021
2017	MagUp	-0.595 ± 0.302	-0.600 ± 0.302	0.005
	MagDown	-0.157 ± 0.301	-0.172 ± 0.300	0.015
2018	MagUp	-0.652 ± 0.303	-0.648 ± 0.303	0.004
	MagDown	-0.156 ± 0.318	-0.162 ± 0.318	0.006

Table 13. Results from calculating A_{prod} for each year and polarity, obtained by modelling the invariant mass distribution using Models 3 and 4. A_{prod} is the result from using Model 3 and A_{prod}' is the result from using Model 4. The integrated A_{prod} and integrated A_{prod}' are used to calculate a systematic error.