9-9-2022

Measurements of $B \to D^{(*)-,0} D^{(*)+,0} K^+ \pi^-$ Branching Fractions in the K^{*0} Mass Window

Harris Bernstein

Abstract

This thesis presents measurements of 11 branching fractions of the form $B \to \overline{D}^{(*)-,0} D^{(*)+,0} K^+ \pi^-$ within the K^{*0} mass window using LHCb data taken in 2016, 2017 and 2018. All 11 branching fractions are measured simultaneously and are reported alongside the covariance and correlation matrices for the final measurement.

Measurements of $B \to \overline{D}^{(*)-,0} D^{(*)+,0} K^+ \pi^-$ Branching Fractions in the K^{*0} Mass Window

by Harris Conan Bernstein

B.S., Pennsylvania State University 2015

Dissertation

Submitted in partial fulfillment of the requirements for a degree of

Doctor of Philosophy in Physics

Syracuse University

December 2022

I would like to thank my advisor Matthew Rudolph, for his support. I am grateful for his willingness to answer questions and patience throughout the last few years

I also appreciate the support of the professors and fellow students at Syracuse University especially Marina Artuso, Steve Blusk, Ray Mountain, Tomasz Skwarnicki as well as Andrew Beiter, Aravindhan Venkateswaran, and Michael Wilkinson, and the late Sheldon Stone.

My deepest thanks to my parents Steven and Toba Bernstein, who supported me throughout my struggles,

Finally a special thanks to Samantha Weigner whose presence and words helped me push through the final months of my graduate school career.

Contents

List of Tables

List of Figures

1 Introduction

The goal of physicists is to describe how the world around us works. Particle physicists attempt to explain the world in terms of indivisible fundamental objects and their interactions. These objects are *particles*, and the way they behave and interact is described by a *model* known as the Standard Model, often times abbreviated as SM. The Standard Model is a robust theory, born from theoretical prediction and experimental validation over the last century. It describes how our electronic devices work (electromagnetism), how certain particles decay (weak nuclear force), and how other composite particles like the nucleus of a atom stay together (strong nuclear force). The SM does not explain several important well known physical phenomena, such as gravity, the asymmetry between matter and antimatter [13], or the accelerating expansion of the universe. These phenomena are explained by different models such as General Relativity or not at all. Therefore, New physics (NP) beyond the SM is required if we are to explain how the world around us works.

To that end, particle physicists perform experiments as a way to test and validate models such as the Standard Model, and discover new phenomena not yet predicted. The Large Hadron[1] Collider (LHC) is the largest of these experiments. It collides particles at conditions that mimic the energies and conditions of the beginning of the universe, so we can better understand the fundamental constituents and interactions of particles. Several experiments measure properties of these collisions at LHC. The LHCb experiment specializes in the measurements of the physics of b-quarks and can measure other phenomena relating to c-quarks and electroweak physics[2]. Recent measurements relating to *Lepton Flavor Universality* are at tension with what the SM predict, and is discussed in section 1.2.1. In order to reduce uncertainties on these measurements and aid in similar searches we can aim to measure certain decays at LHCb. This thesis measures the *branching fraction* of 11 different B Mesons decays (particles containing one b quark and one other quark). In addition to what has

[1] Hadrons are any particle containing a quark, the fundamental building block of protons and neutrons
[2] At high energies, electromagnetism and the weak nuclear force unify

already been mentioned, this measurement is a starting point for any number of amplitude analyses that seek to add to the literature of hadron spectroscopy.

The first part of this dissertation summarizes the theoretical parts of the Standard Model that are relevant to our experimental measurements. The second part is a description of the LHCb detector. The third part is a detailed explanation of our branching fraction measurement, which specifically measures 11 different decays of the form $B \to \overline{D}^{(*)-,0} D^{(*)+,0} K^+ \pi^-$.

1.1 Standard Model

The Standard Model is a Quantum Field Theory, where its particles and interactions are described in terms of fields. We refer to the localized vibrations of fields as particles, and often times use the two words interchangeably. *Spin* is an intrinsic property of particles in the Standard Model. It can be compared to the classical definition of angular momentum, but unlike the wheel of a bicycle, particles do not "spin" around an axis. Instead spin is just one part of the particles *total angular momentum*, the other part being the *orbital angular momentum*. Particles with a spin of half integer values are known as *fermions*. Particles with a spin of integer values are known as *bosons*. The Standard Model is broken up even further by what interactions are allowed between the bosons and fermions. Leptons are fermions that do not interact with the strong nuclear force. The electron is an example of a lepton. Quarks can interact with the strong nuclear force, the electromagnetic force, and the weak nuclear force.

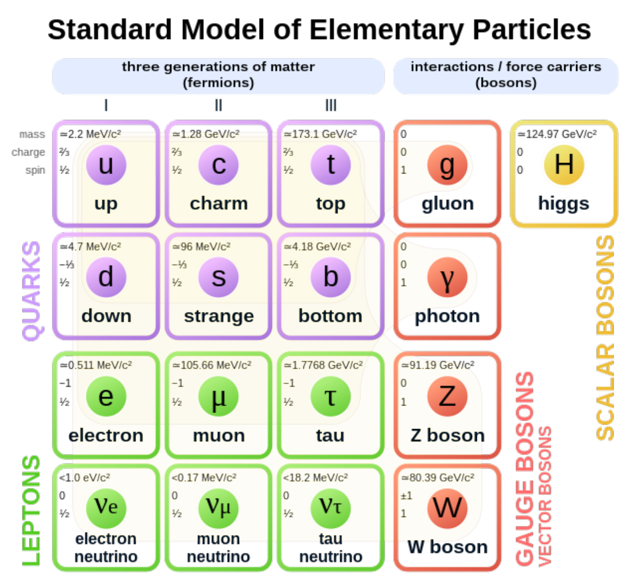


Figure 1: The particle content of the Standard Model is shown. Fermions, spin $\frac{1}{2}$ particles, are broken down into quarks and leptons. Bosons, spin 1 particles, are the mediating particles that describe the interactions of the 3 fundamental forces in the Standard Model. [2]

The fundamental forces the Standard Model describes are mediated by bosons. The strong nuclear force is mediated by the gluon. The gluon, unlike the other bosons, can interact with itself. The strong nuclear force is what keeps quarks confined[3], most commonly in three quark states known as *Baryons*, or two quark states known as *Mesons*. The proton and neutron are examples of baryons. More exotic 4 and 5 quark states have been observed [14]. Mesons and Baryons are collectively called *Hadrons* (any particle with a quark).

Electromagnetism is mediated by the photon. Any particle with electric charge interacts with the electromagnetic force. It acts over long ranges. Alongside gravity, is what we most commonly are aware of on a day to day basis.

The weak nuclear force is mediated by the W boson and the Z boson. All fermions have a *weak* charge, and so interact with the weak force. It is only effective over very short ranges, but is responsible for several phenomena, such as beta decay and electron capture. The Standard Model assumes that, up to differences in the mass, leptons and weak bosons interact equally. This assumption is called *Lepton Flavor Universality* and a potential test of it is the subject of section 1.2.1.

The only stable particles described by the Standard Model are the electron (a lepton) and the proton (a baryon of two up quarks and one down quark) [4] All other particles decay. The probability of a particle decaying in a specific way is known as the *Branching Fraction*. Theoretically a branching fraction can be predicted by summing up the respective amplitudes of each possible decay. A branching fraction can also be measured directly through experimentation. Often times for decays with numerous potential final states, a branching fraction measurement is an important first step in setting constraints on the amplitudes, as well as testing standard model predictions.

Decays of the form $b \to c\bar{c}s$[5] are an extremely common topology for b-hadrons. B meson decays with explicit final states of the form $\overline{D}^{(*)}D^{(*)}K^+$ or $\overline{D}^{(*)}D^{(*)}K^{*0}$ have been

[3] *Confinement* is a phenomenon of Quantum Chromodynamics (QCD) the study of the Strong nuclear force

[4] Some theories beyond the standard model predict proton decay [15] and experiments such as DUNE [16] aim to observe it.

[5] The decay of particle with a b quark to particles with some combination of $c\bar{c}$ s quarks

measured [17] [18] [19]. These three-body decays have relatively large branching fractions from the 10^{-3} level to the percent level. A recent measurement of the $\overline{D}^0 D^0 K^+ \pi^-$ branching fraction has been measured at a value of 3.5×10^{-4} [20]. These decays contain a rich abundance of resonance structure that can be measured in full amplitude analyses, and include the potential for exotic hadron contributions. From an experimental point of view, they are significant backgrounds for analyses involving rare partially-reconstructed decays, such as $B^0 \to K^{*0}\tau\tau$, that are significant for probing beyond the Standard Model. All of this is motivation for the measurement in section 3. section 1.2 explains how the particles mentioned above are defined by various quantum numbers. section 1.2.1 connects the measurement in section 3 to probes for physics beyond the standard model.

1.2 Resonant States and Excited States

The species of a hadron is determined solely by its constituent quarks. However there are several quantum numbers that define the states available to any particle. A particle with quantum numbers corresponding to the lowest energy configuration is said to be in its ground state. *Excited States* have a different configuration of quantum numbers. Excited state particles have more mass then their lower states. These excited states decay either via the electromagnetic force or the strong force to lower energy states. Unstable ground state particles decay via the weak nuclear force. A short lived particle, which is often times simply an excited state, is often called a resonant state. In addition all particles in the standard model have corresponding *anti-particles*. Particles and their anti-particles have the same properties (quantum numbers) with the exception of electric charge which is flipped. As an example the antiparticle of a u quark is written as $\bar{u}$. We provide the quantum numbers[6] that determine a state below, as well as the allowed values for mesons which are the particles studied in section 3. We use [21] as our primary reference:

- *Spin* or $\mathcal{S}$: The individual quarks each add a spin vector of magnitude $\frac{1}{2}$ to the total

[6]Two quantum numbers Hypercharge and G-Parity are not included

spin of a composite particle. For mesons you can have spin 0 where the spin vectors are unaligned, or spin 1 where the spin vectors are aligned.

- Orbital Angular Momentum, $\mathcal{L}$, can take any positive integer value, 0,1,2,3 .... The parity of a state is determined by its orbital angular momentum.

- Parity or $\mathcal{P}$. For mesons $\mathcal{P} = -1^{\mathcal{L}+1}$.

- Charge Conjugation Parity or $\mathcal{C} = -1^{\mathcal{L}+\mathcal{S}}$. $\mathcal{C}$ Parity is a quantum number for mesons whose quark anti-quark pair are anti-particles of each other. We call these unflavored mesons. For flavored particles $\mathcal{C}$ is undefined.

- The total angular momentum $\mathcal{J}$. It takes values between $|\mathcal{L} + \mathcal{S}|$ and $|\mathcal{L} - \mathcal{S}|$ A State with $\mathcal{J} = 0$ is called an S wave. A State with $\mathcal{J} = 1$ is called an P wave. The K^{*0} meson is an excited state that can decay in its S wave or P wave state. Amplitude analyses can measure the relative rates at which these happen

- Isospin, $\mathcal{I}$, is determined by the number of u and d quarks in a composite particle. Depending on the exact q$\bar{q}$ flavors of a meson, they can take an isospin of 0, $\frac{1}{2}$, or 1. For example, heaver mesons such as B's, D's, and Kaons take an isospin of $\frac{1}{2}$. There are additional flavor quantum numbers for Strangeness - S, Charmness - C, Bottomness - B, and Topness - T. By convention the sign of the flavor quantum number is positive for up type quarks (u, c, t) and negative for down type quarks (d, s, b)

- Baryon Number, $\mathcal{B}$, is

$$\mathcal{B} = \frac{1}{3}(N(q) - N(\overline{q}))$$

where N(q) is the number of quarks for a particle and N($\overline{q}$) is the number of antiquarks. For mesons, the Baryon number is 0.

- Electric Charge, $\mathcal{Q}$, is given by

$$\mathcal{Q} = I_z + \frac{\mathcal{B} + S + C + B + T}{2}$$

This also known as the Gell-Mann-Nishijima formula.

The naming and classification of particles is done by examining the possible quantum states, as defined by their quantum numbers. For unflavored mesons, particles are classified in a $\mathcal{J}^{\mathrm{PC}}$ notation; for flavored mesons, $\mathcal{J}^{\mathrm{P}}$ notation is used. The initial and final states of the decays measured in section 3, the B and D mesons and the ground state Kaons and Pions are 0^1. The intermediate resonance, K^{*0} meson is 1^-.

1.2.1 Standard Model Measurements

Recent measurements of various branching fractions ratios have found non-trivial deviations from the standard model predictions that assume lepton flavor universality. As an example, fig. 2 highlights the measurements of $\mathcal{R}(D)$ and $\mathcal{R}(D*)$ defined as

$$R_{D^{(*)}} = \frac{\mathcal{B}(B \to D^{(*)}\tau\nu)}{\mathcal{B}(B \to D^{(*)}l\nu)}, l = e, \mu \tag{1}$$

Searches for explanations for these violations are often times motivated by Effective Field Theories, EFTs. EFTs are model independent; they do not depend on what specific new physics is creating these deviations. EFTs can predict enhancements to measurements of physical decays that the standard model predicts. One such enhancement is predicted to exist in $B^0 \to K^{*0}\tau\tau$ [22]. The predicted magnitude of this enhancement can be seen in fig. 3.

The measurement in section 3 helps to constrain potential backgrounds in any search for $B^0 \to K^{*0}\tau\tau$. For a full discussion of the various experimental measurements that are at odds with the standard model, the EFTs that could potentialy explain these measurments,

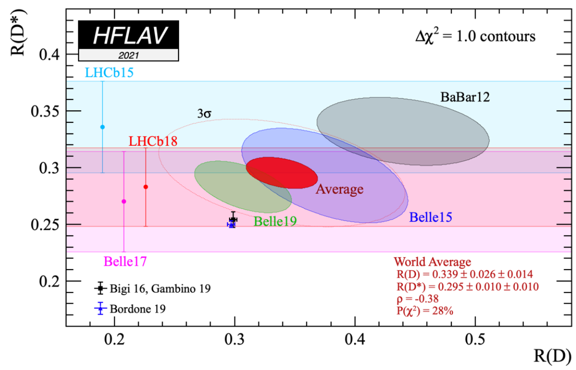


Figure 2: Measurements of R_D and R_{D^*} by BaBar [3], Belle [4] [5] [6], LHCb [7] [8]. The dark red ellipse shows the two dimensional average while the bands and ellipses encompass the various uncertaintys. The SM model predictions are shown as the black and blue points with error bars.

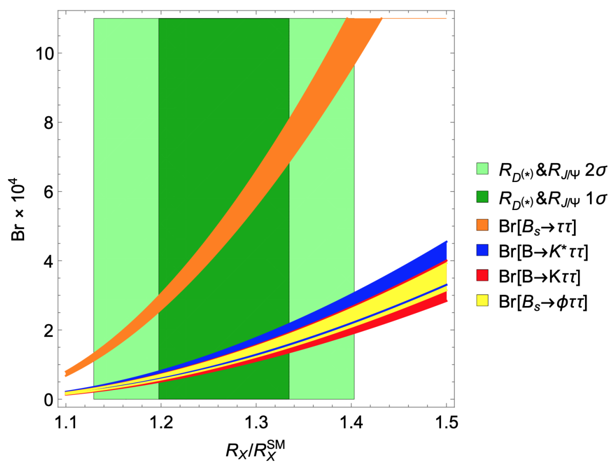


Figure 3: Correlation between different SM processes as a function of potential enhancements to the SM . The 1σ and 2σ intervals are the measured ratios depicted in fig. 2, with X beight either D or D^*.

.

and how they connect to $B^0 \to K^{*0}\tau\tau$, we encourage the reader to examine [23]. [23] used previous values of the branching fractions that arereported in section 3.10.

2 The LHCb Experiment

2.1 The LHC

The Large Hadron Collider[7] is 27 km circumference circular particle collider spanning the borders of Switzerland and France. CERN - the European Organization for Nuclear Research - houses the LHC. Several subaccelerators at CERN feed into the LHC. D A linear accelerator, Linac2(Linac4) accelerates hydrogen anions (H^-) to 50 MeV (160 MeV). They are then accelerated inside the The Proton Synchrotron Booster to 1.4 GeV (2.0 GeV). This process also strips electrons from the hydrogen, leaving us with only protons. These protons are accelerated to 26 GeV and then 450 GeV by the Proton Synchrotron and the Super Proton Synchrotron respectively. At this point the protons are injected into the LHC.

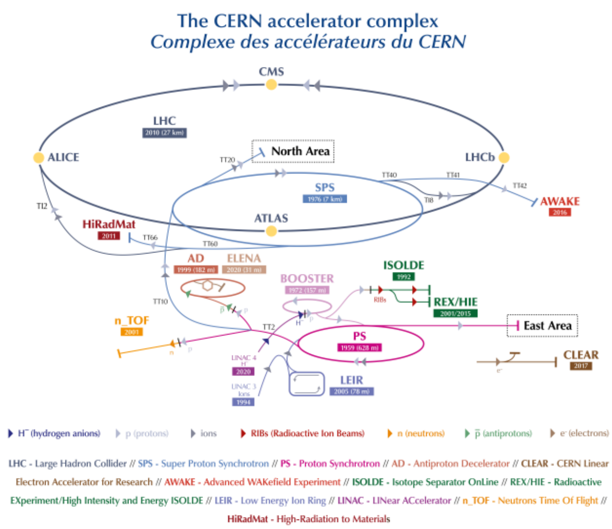


Figure 4: The The CERN accelerator in 2019. [9]

Protons in the LHC are accelerated in two separate beams to energies of 6.5 TeV. These beams go clockwise and counterclockwise, colliding at 4 different points at a center of mass energy of 13 TeV. These points correspond to the ALICE, ATLAS, CMS, and LHCb

[7]During the long shutdown between 2019-2020, many components of the LHC were upgraded. Because this thesis uses data collected during the run2(2016-2018) data taking period, we will quote both the numbers and components used during the run2 period and run3 period

experiment detectors. During the acceleration, the beams are separated further into ”bunches” allowing for more control over collisions. Bunch collisions are referred to as ”events”, and occur at a frequency of nearly 40 MHz.

2.2 The LHCb

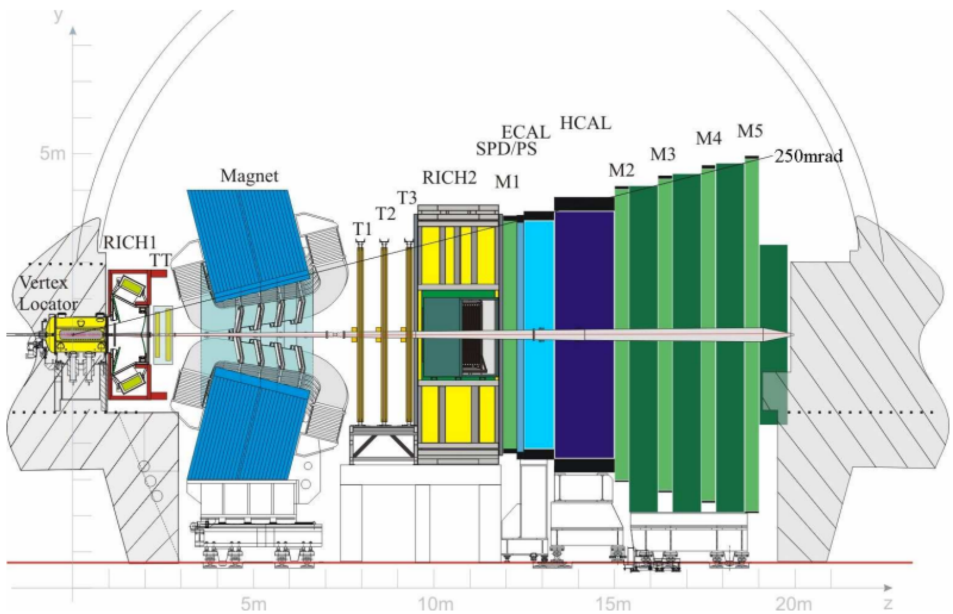
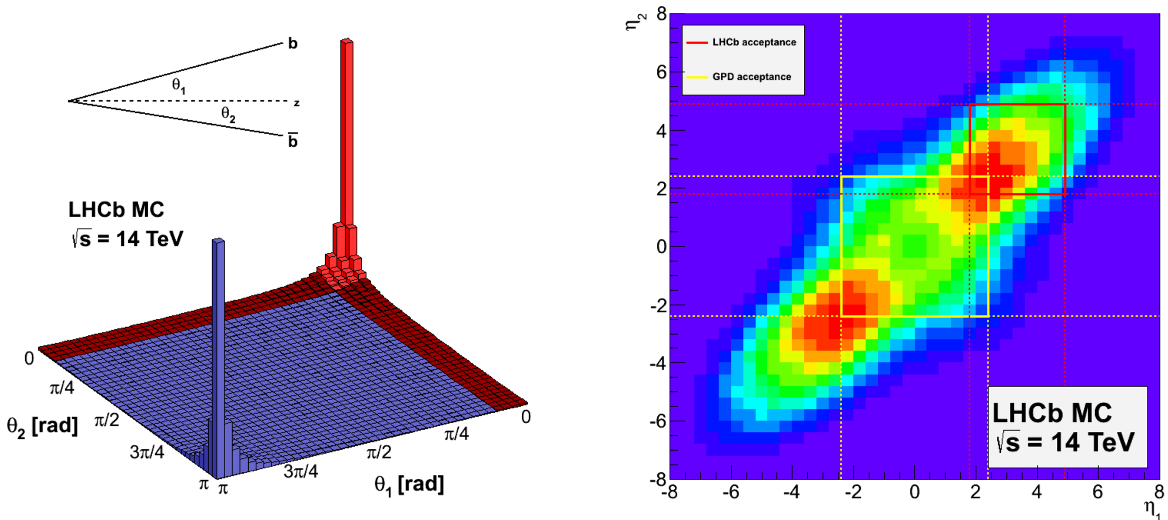


Figure 5: A cross section of the LHCb detector [10]

The LHCb’s primary goal is to measure the physics of particles containing b and c quarks as well as potential charge-parity (CP) violating physics. $b\bar{b}$ quark pair production from pp collisions occurs primarily in tight forward and backward cones, show in simulation fig. 6.



(a) Production angles, *theta*, with respect to the beam line.

(b) LHCb vs GPD η acceptance regions. GPD is a general particle detector like CMS or ATLAS

Figure 6: PYTHIA Simulation of $b\bar{b}$ pair production due to pp collisions. Red areas are the LHCb's acceptance. [11]

3 Branching Fraction Measurement

3.1 Introduction

In this analysis, we perform measurements of 11 of these kinds of branching fractions for decays of the form $B \to \overline{D}^{(*)-,0} D^{(*)+,0} K^+ \pi^-$ in the K^{*0} mass window. We do not separate the resonant K^{*0} contribution from non-resonant $K^+\pi^-$ decays, although we will often refer to the final states as containing a K^{*0}. Excited states with higher mass then then the $D*$ mesons (D^{**} charm states) which decay to $D^{(*)}\pi$, where the pion is used in the $K^+\pi^-$ combination, are also considered signal. Our final states take the form of $\overline{D}^{-,0} D^{+,0} K^+ \pi^-$, without attempting to reconstruct any photons or neutral pions. This takes advantage of the small phase space in D^* meson decays, which allows us to identify the initial decay states. We use the charm meson decays $D^0 \to K^-\pi^+$ and $D^+ \to K^-\pi^+\pi^+$ to reconstruct the the appropriate D species in our signal modes. Throughout this note, we will label the analyzed final states that contain a $K^+\pi^-$ combination, so $\overline{D}$ is used for the anti-charm meson that then decays with a K^+, and D for the charm meson decaying with a K^-. Because many different initial decays can contribute to each $\overline{D}D$ combination, we extract the final branching fractions simultaneously in four primary spectra: D^-D^+, $\overline{D}^0 D^0$, $D^- D^0$, and $\overline{D}^0 D^+$. We also use one additional spectrum with a soft charged pion reconstructed as $\overline{D}^0 D^{*+} K^+ \pi^-$. This spectrum is needed to resolve an ambiguity that will be described in 3.2. We measure these branching fractions using Run 2 data collected in 2016, 2017, and 2018. We measure the branching fractions relative to the previously measured $B^{0,+} \to \overline{D}^{(-,0)} D^0 K^+$ decays with the $D^0 \to K^-\pi^+\pi^+\pi^-$ depending on the final reconstructed track multiplicity.

3.2 Detailed Analysis Strategy

We start by carefully defining for the reader the signal spectra under study, each of which contains several contributing B meson channels. These decays manifest as our signal peaks.

The five possible final state reconstructions in the signal regions of this analysis are:

Z: $D^- D^+ K^{*0}$

ZZ: $\overline{D}^0 D^0 K^{*0}$

P: $\overline{D}^0 D^+ K^{*0}$

M: $D^- D^0 K^{*0}$

ST: $\overline{D}^0 (D^{*+} \rightarrow D^0 \pi^+) K^{*0}$ (We reconstruct the soft pion here)

The peaks in each of these final states fig. 7 can be fed to from up to three different signal B decays and intermediate states. These labeling schemes are used only for the reading of the yield equations presented later on in this section. We use the full description of the appropriate decay in all other places. Each B decay of interest is labeled h a numbered code, with an additional letter to refer to different intermediate states that feed into separate signal spectra. The following scheme drops any soft or neutral particles that are not part of the reconstruction.

1. $B^0 \to D^- D^+ K^{*0}$

2. $B^0 \to D^{*-} D^+ K^{*0}$

    a. $D^- D^+ K^{*0}$

    b. $\overline{D}^0 D^+ K^{*0}$

3. $B^0 \to D^- D^{*+} K^{*0}$

    a. $D^- D^+ K^{*0}$

    b. $D^- D^0 K^{*0}$

4. $B^0 \to D^{*-} D^{*+} K^{*0}$

    a. $D^- D^+ K^{*0}$

    b. $\overline{D}^0 D^+ K^{*0}$

    c. $D^- D^0 K^{*0}$

    d. $\overline{D}^0 D^0 K^{*0}$

    e. $\overline{D}^0 (D^{*+} \to D^0 \pi^+) K^{*0}$

5. $B^+ \to \overline{D}^0 D^+ K^{*0}$

6. $B^+ \to \overline{D}^{*0} D^+ K^{*0}$

7. $B^+ \to \overline{D}^0 D^{*+} K^{*0}$

    a. $\overline{D}^0 D^+ K^{*0}$

    b. $\overline{D}^0 D^0 K^{*0}$

    c. $\overline{D}^0 (D^{*+} \to D^0 \pi^+) K^{*0}$

8. $B^+ \to \overline{D}^{*0} D^{*+} K^{*0}$

    a. $\overline{D}^0 D^+ K^{*0}$

    b. $\overline{D}^0 D^0 K^{*0}$

    c. $\overline{D}^0 (D^{*+} \to D^0 \pi^+) K^{*0}$

9. $B^0 \to \overline{D}^0 D^0 K^{*0}$

10. $B^0 \to \overline{D}^{*0} D^0 K^{*0}$ + $B^0 \to \overline{D}^0 D^{*0} K^{*0}$

11. $B^0 \to \overline{D}^{*0} D^{*0} K^{*0}$

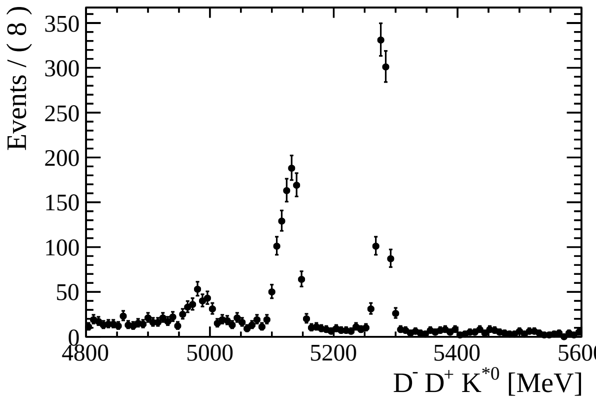
We note that for the numbers, each block of four refer to either a different initial B meson or in the case of the $\overline{D}^0 D^0 K^{*0}$ block the fact that the B^0 decays without a $D^{(*)+}$ meson. In each block of four, increasing number refers to more * mesons. For the letters, the $D^{*+} \to D^+ \pi^0/\gamma$ comes before $D^{*+} \to D^0 \pi^+$.

When selecting channels of the above form using ground state D mesons, from one to three peaks will be visible, where the lower mass peaks correspond to states with one or two D^* mesons in the initial decay. Because the phase space of the D^* is small these peaks are well separated across the invariant mass spectrum of the B candidate. However missing particles from the D^* means we will reconstruct events from multiple decays of interest in individual D^* peaks across multiple spectra. For example, the decay $B^0 \to D^{*-} D^+ K^{*0}$ contributes to $D^- D^+ K^{*0}$ and$\overline{D}^0 D^+ K^{*0}$.

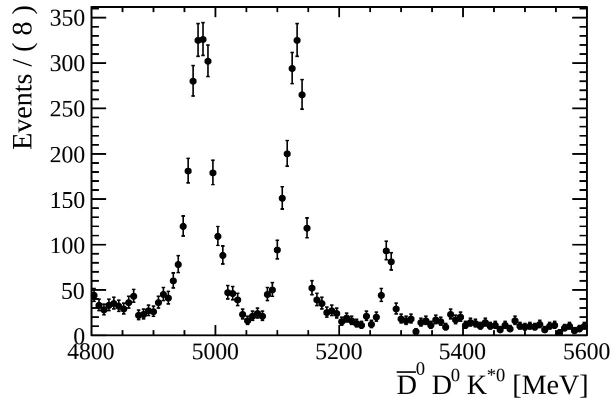
For the $D^- D^0 K^{*0}$ spectrum we do not expect to see the fully reconstructed decay due to the incorrect strange quark (a B^- contains a b, so we would see $D^0 D^- K^- \pi^+$ which is part of the $\overline{D}^0 D^+ K^{*0}$ spectrum). The excited state peaks in $D^- D^0 K^{*0}$ correspond to the reconstruction of $D^{*-}D^{*+}K^{*0}$ and $D^{-}D^{*+}K^{*0}$ decays, when the D^{*+} decays to $D^0\pi^+$. For the $\overline{D}^0(D^{*+}\to D^0\pi^+)K^{*0}$ spectrum the reconstruction of the soft pion means we only miss one pion at most, therefore only two peaks are seen. To distinguish the intermediate states of the B decays, it is necessary to simultaneously examine each of these final states, as the expected shapes of peaks passing through different intermediate channels are expected to be almost identical. All the peaks visible in data can be seen in fig. 7. Each can contain multiple different B decays and intermediate states. Each peak has a measured yield denoted by M and is assigned a label with a superscript of the signal channel label and a subscript of the peak number (0 for no unmeasured * components, 1 for one, 2 for two). In table 1 we label each possible B decay with the scheme so far. For each decay we specify the corresponding simulation sample event ID and the peak id that corresponds to the ones seen in fig. 7. For future reference we also include the normalization modes in table 1.

Table 1: Decay Modes and which peaks they will contribute to in our signal spectrum. Modes with two peak IDs use the same simulation sample twice for efficiency calculations. The exception is for scheme 10 which is used to represent the sum of $B^0\to\overline{D}^{*0}D^0K^{*0}$ and the $B^0\to\overline{D}^0D^{*0}K^{*0}$ decays as this analysis does not separate the two modes. The peak IDs are used to aid in the construction of the yield equations below. The normalization modes are included as well for clarity

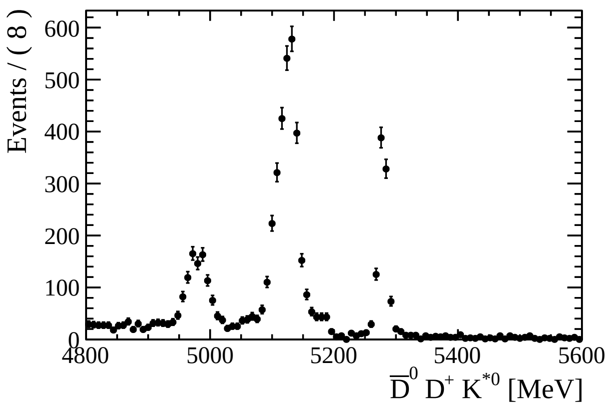
Mode	Event Type	Peak ID
$B^0\to D^-D^+K^{*0}$	11198006	M_0^Z
$B^0\to(D^{*-}\to D^-\pi^0)D^+K^{*0}$	11198400	M_1^Z
$B^0\to(D^{*-}\to\overline{D}^0\pi^-)D^+K^{*0}$	11198005	M_1^P and M_1^M
$B^0\to(D^{*-}\to D^-\pi^0)(D^{*+}\to D^+\pi^0)K^{*0}$	11198401	M_2^Z
$B^0\to(D^{*-}\to\overline{D}^0\pi^-)(D^{*+}\to D^+\pi^0)K^{*0}$	11198410	M_2^P and M_2^M
$B^0\to(D^{*-}\to\overline{D}^0\pi^-)(D^{*+}\to\overline{D}\pi^+)K^{*0}$	11198023	M_2^{ZZ} and M_1^{Pst}
$B^+\to\overline{D}^0D^+K^{*0}$	12197023	M_0^P
$B^+\to\overline{D}^{*0}D^+K^{*0}$	12197410	M_1^P
$B^+\to\overline{D}^0(D^{*+}\to D^+\pi^0)K^{*0}$	12197400	M_1^P
$B^+\to\overline{D}^0(D^{*+}\to\overline{D}\pi^+)K^{*0}$	12197045	M_1^{ZZ} and M_0^{Pst}
$B^+\to\overline{D}^{*0}(D^{*+}\to D^+\pi^0)K^{*0}$	12197401	M_2^P
$B^+\to\overline{D}^{*0}(D^{*+}\to\overline{D}\pi^+)K^{*0}$	12197423	M_2^{ZZ} and M_1^{Pst}
$B^0\to\overline{D}^0D^0K^{*0}$	11196019	M_0^{ZZ}
$B^0\to\overline{D}^{*0}D^0K^{*0}+B^0\to\overline{D}^0D^{*0}K^{*0}$	11196413	M_1^{ZZ}
$B^0\to\overline{D}^{*0}D^{*0}K^{*0}$	11196414	M_2^{ZZ}
$\overline{D}^0(D^0\to K^-\pi^+\pi^+\pi^-)K^+$	12197008	-
$D^-(D^0\to K^-\pi^+\pi^+\pi^-)K^+$	11198007	-



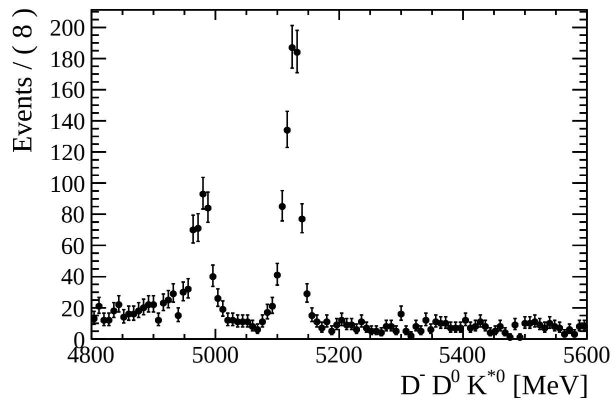
(a) Events reconstructed as $D^{-}D^{+}K^{*0}$. This analysis will distinguish the contributions to the middle peak where we miss a particle from one excited charm



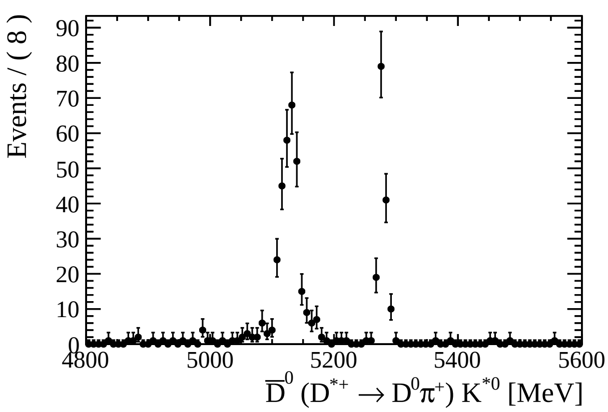
(b) Events reconstructed as $\overline{D}^{0}D^{0}K^{*0}$. We aim to distinguish between six distinct B decays in this spectrum.



(c) Events reconstructed as $\overline{D}^{0}D^{+}K^{*0}$. We aim to distinguish between six distinct B decays in this spectrum.



(d) Events reconstructed as $D^{-}D^{0}K^{*0}$. No peak exists at the B mass as the decay $B^{-}\to D^{-}D^{0}K^{*0}$ has the wrong strange content. The two peaks correspond to missing a charged soft pion.



(e) Events reconstructed as $\overline{D}^{0}(D^{*+}\to D^{0}\pi^{+})K^{*0}$ where we reconstruct the soft pion. At this stage in the event selection no candidates are shared between fig. 7e and fig. 7b.

Figure 7: Signal Distributions of interest for Data collected from 2016 - 2018 after the event selection process. Higher excited D^{**} charm states which decay to $D^{(*)}\pi$, where the pion is used in the $K^{+}\pi^{-}$ combination are still present in the data

The yields of each *decay mode*, denoted with capital letter N, can be computed using the full set of measured *peak yields* M, although we must include corrections for D meson branching fractions and efficiencies. These yields N_i correspond to the numbering of the scheme ID. Each yield term can be defined as:

$$N_i(B \to D\overline{D}K^{*0}) = \mathcal{B}(B \to D\overline{D}K^{*0}) \times \epsilon_{\text{rel}}[\times \prod \mathcal{B}(D)] \times \mathcal{N}(\text{NORM}) \tag{2}$$

Where ϵ_{rel} is the relative efficiency between reconstructing a decay in a specific spectrum and its normalization channel, $\prod \mathcal{B}(D)$ is the product of the relevant D^* and D meson branching fractions, and $\mathcal{N}(\text{NORM})$ is the yield of its normalization channel. Each measurable peak is described as a sum of these different N_i. The need to study the spectra simultaneously can be illustrated by considering how each decay mode contributes to the observed yield in each peak. The final fits for the analysis will also include shape information to improve the separation. For the following equations, we define $f_0 = \mathcal{B}(D^{*+} \to D^0\pi^+)$ and $f_+ = 1 - f_0$.

Each peak yield with no D^* components is then:

$$M_0^Z = N_1$$
$$M_0^{ZZ} = N_5$$
$$M_0^P = N_9$$

The four double starred decay peaks are related to one another:

$$
\begin{aligned}
M_2^Z &= f_+^2 N_4 \\
M_2^{ZZ} &= f_0^2 N_4 + f_0 N_8 + N_{12} \\
M_2^P &= f_0 f_+ N_4 + f_+ N_8 \\
M_2^M &= f_0 f_+ N_4
\end{aligned}
$$

The four yields are determined by only three signal spectra, so this part of the analysis is over-constrained. It is when one analyzes the single starred peaks that we see more information is needed:

$$
\begin{aligned}
M_Z^1 &= f_+ N_2 + f_+ N_3 \\
M_P^1 &= f_0 N_2 + N_6 + f_+ N_7 \\
M_{ZZ}^1 &= f_0 N_7 + N_{10} + N_{11} \\
M_M^1 &= f_0 N_3
\end{aligned}
$$

If we were to analyze only the Z spectrum, we see that we can make a statement about the sum $N_2 + N_3$. Analyzing the M spectrum would give the information needed to distinguish the two components. But to analyze also the B^+ decays to the P spectrum, we add a combination of N_6 and N_7 that cannot be distinguished. Adding ZZ does not help, as it introduces the new sum $(N_{10} + N_{11})$ (this sum will be left as such in the final measurement). To distinguish N_6 and N_7, we introduce the Pst spectrum and the efficiency to reconstruct the soft pion from the D^{*+}, ϵ_π:

$$M^0_{Pst} = \epsilon_\pi f_0 N_7,$$

$$M^1_{Pst} = \epsilon_\pi \big(f_0^2 N_4 + f_0 N_8\big).$$

If we remove all candidates with a reconstructed D^{*+} from the other spectra, then the corresponding terms for N_4, N_7, and N_8 add an additional factor of $(1-\epsilon_\pi)$.

As an example of what a measured yield will look like we fully expand the yields for two of the observed peaks:

$$M^Z_0 = \mathcal{B}\big(D^+\big)^2 \epsilon_1 \frac{\mathcal{B}_1}{\mathcal{B}^0_{\text{norm}}} \frac{M^0_{\text{norm}}}{\epsilon^0_{\text{norm}} \mathcal{B}(D^+) \mathcal{B}(D^0 \to K^-\pi^+\pi^-\pi^+)}. \tag{3}$$

$$M^P_2 = \mathcal{B}\big(D^+\big)\mathcal{B}\big(D^0\big) \Bigg[\epsilon_{4b} f_0 f_+ \frac{\mathcal{B}_4}{\mathcal{B}^0_{\text{norm}}} \frac{M^0_{\text{norm}}}{\epsilon^0_{\text{norm}} \mathcal{B}(D^+) \mathcal{B}(D^0 \to K^-\pi^+\pi^-\pi^+)} \tag{4}$$

$$+ \epsilon_{8a} f_+ \frac{\mathcal{B}_8}{\mathcal{B}^+_{\text{norm}}} \frac{M^+_{\text{norm}}}{\epsilon^+_{\text{norm}} \mathcal{B}(D^0) \mathcal{B}(D^0 \to K^-\pi^+\pi^-\pi^+)}\Bigg]. \tag{5}$$

With the yields of the decays of interest described in terms of the events present in our signal spectra we move on to describing the rest of the analysis. In section 3.3 we describe the variables and mathematical techniques we use in the thesis. In section 3.4 we describe the event selection flow for both Data and Simulation. section 3.5 we describe our choice of normalization channels, the measurement of their yields and the calculation of efficiencies for each simulation sample. In section 3.7 we implement a discrete fit to each MC samples decay tree fitter constrained B mass as well as signal peaks in our data spectra. In section 3.8 we handle the systematic uncertainties present in our analysis. In section 3.10 we describe the process of constructing the full simultaneous fit necessary to measure the 11 different branching fractions across the five distinct B spectra as seen in fig. 7.

3.3 Analysis Techniques and Variables

Here we describe several mathematical techniques and variables of interest that are used throughout the rest of the analysis.

3.3.1 Kinematic Fit

In order to improve the resolution and the best estimate for the track parameters of our final state kaons and pions, a kinematic fit is implemented. We implement this fit via DecayTreeFitter [24]. DecayTreeFitter is a minimization of the χ^2 of a full decay chain, often times called a global least squares fit, where a decay chain can have multiple decay vertices. Our use of the fit produces various track parameters where it constrains the masses of any fully reconstructed intermediate D mesons to their nominal values, while allowing other parameters to vary within their uncertainties. Our analysis contains measurements of decays where we do not reconstruct certain particles, and as such we do not apply any vertex constraints to the B parent for our decays. We often times refer to any variable that employs this fit as a **DTF** variable.

3.3.2 Likelihood Estimation and Probability Density Functions

The fits to the various invariant mass distributions through out the remainder of this analysis are done via a maximum likelihood fit otherwise known as a likelihood estimation. Our likelihood functions are constructed as:

$$\mathcal{L}(x|\theta) = \prod_{i}^{n} f(x_i;\theta) \tag{6}$$

where $f(x_i;\theta)$ is a probability density function (PDF) that described the distribution of data, x, in terms of some function that depends on a set of parameters θ. The full description of the PDFs used in this analysis can be found in appendix B. Our likelihood estimation is done numerically via the RooFit [25] package within ROOT which uses MIGRAD to perform the minimization of the negative log likelihood and uses HESSE to calculate errors and continuity.

3.3.3 sPlot

In order to calculate uncertainties in section 3.8 we use the sPlot technique [26]. This technique unfolds the individual contributions that signal and background process have to a given distribution. In our case we apply it to the invariant masses of of our B mesons in section 3.7.2 using signal and background PDFs where we extract the relative yields of these components via a likelihood fit. The sPlots technique derives from these inputs event weights for our signal components called sWeights. These sWeights are applied later on to event distributions that are independent of the invariant mass of our B mesons. This allows us to compare simulation, which is only signal, to data which represents only the signal contribution.

3.3.4 Boosted Decision Trees

In section 3.8.2 we implement a reweighting scheme in order to estimate systematic uncertainties on our final efficiencies associated with mismodeling in our simulation. This scheme relies on an implementation of Boosted Decision Trees (BDTs) via the hep_ml package [27]. Decision trees are machine learning algorithms that are implemented in classification and regression problems. Decision Trees make a series of binary decisions (if-else) based off a set of input parameters associated with a data set (typically called a training sample), to classify other data (test sample) as either signal or background. However a signal decision tree is a weak learner; it is only slightly better then simply guessing if data is signal or background. It is also sensitive to statistical fluctuations during the training process. To solve these issues, a gradient boosting process is implemented to train an ensemble of trees, one after another, where a reweighted version of one is used to train the next. The hep_ml [27] package trains its trees in the following way:

- A tree is built to maximize the symmetrized binned χ^2 where:

$$\chi^2 = \sum_{bin} \frac{(w_{bin,mc} - w_{bin,data})^2}{(w_{bin,mc} + w_{bin,data})^2} \tag{7}$$

- The tree's final predictions (leaves) are calculated:

$$\text{leaf}_{\text{pred}} = \ln\left(\frac{w_{\text{leaf, data}}}{w_{\text{leaf, mc}}}\right) \tag{8}$$

- The sample is reweighted where:

$$w \longleftarrow \begin{cases} w, & \text{for data distribution} \\ w \times e^{\text{leaf}_{\text{pred}}}, & \text{for mc distribution} \end{cases} \tag{9}$$

This process is implemented a number of times. Each tree also has a depth associated with it that defines the number of nodes or splits the tree makes in its decision process. These two variable our are hyperparmeters and are chosen by optimizing the BDT. Our figure of merit for this optimization is the ROC AUC, or the *Area Under Curve* for the *Receiver Operator Curve*. An example is shown in fig. 8. Since we are attempting to extract simulation weights from a BDT so it is unable to distinguish between data and mc, the auc value we aim for is 0.5.

3.3.5 Variables

Variables commonly used and referenced in this analysis are described in detail in this section.

- ${\chi^2}_{\textbf{track}}$ is the χ^2/ndf of the fit to the track
- $\mathbf{p_t}$ is the transverse momentum of a track. i.e the momentum perpendicular to the beam.
- **p** is the momentum of a track in the direction of the beam.

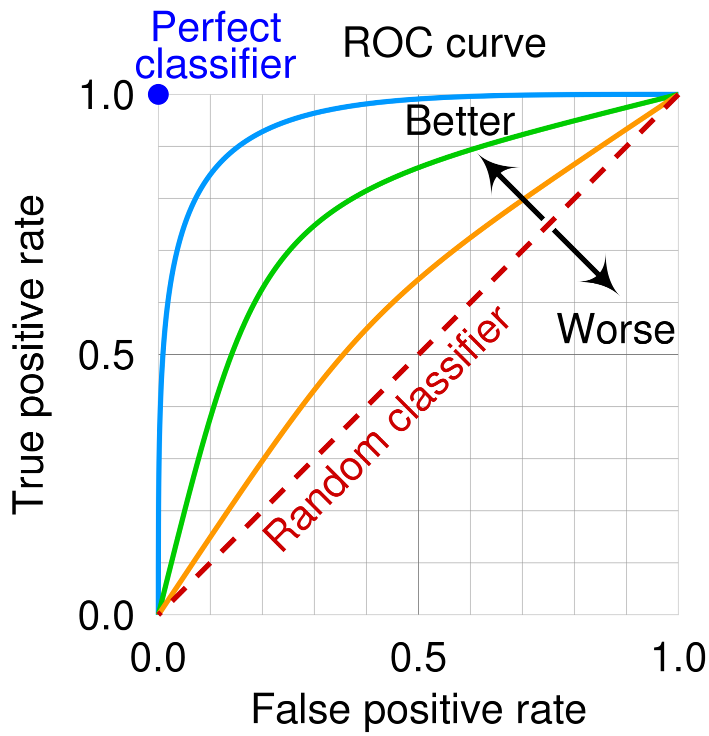


Figure 8: A Receiver Operator Curve [12] where true positive rate is shown vs the false positive rate for potential classifier responses. In the case when a classifer is unable to distinguish between two categories, its AUC will be equal to 0.5

- $\mathbf{Ghost_{track}}$ is the probability of a track being fake. An algorithm [28] calculates the probability of a track being reconstructed with an incorrect combination of hits, which we call a ghost.

- $\mathbf{DLL}_{\mathrm{K}\pi}$ is the difference in log likelihood that a track is a kaon rather than a pion.

- $\mathbf{m_{pdg}}$ is the nominal mass value for a given particle species taken from [29]

- $\mathbf{IP_{PV}}$ is the value of the impact parameter with respect to the primary vertex of a candidate.

- $\mathbf{DOCA}_{\mathrm{k}\pi}$ is the distance of closes approach between two particles

- $\chi^2\mathbf{_{PV}}$ is the significance of the impact parameter with respect to the primary vertex of a candidate. Prompt tracks that originate from the PV will have smaller values.

- $\chi^2_{\mathbf{vtx}}\mathbf{/ndf}$ is the quality of the fit to the decay vertex associated with a track

- $\chi^2_{\mathbf{vtx\ -\ pv}}$ is the significance of the distance from decay vertex and the primary vertex. The larger the value the greater the displacement from the primary vertex.

- $\mathbf{DIRA_{PV}}$ is the cosine of the angle between the vector pointing from the primary vertex to the decay vertex and the momentum vector of the track

- $\tau_{\mathbf{PV}}$ is the lifetime of the of the track with respect to the primary vertex.

- **BBDT** is the output of a bonsai boosted decision tree that helps search for decay topologies compatible with the B meson parent.

3.4 Event Selection

This section describes the order of our event selection process applied to the data and simulation for this analysis. Unless stated otherwise, each section is applied to events that pass the previous section. Unless stated otherwise, each section is applied to both data and simulation in the same way. Due to the large number of simulation (MC) sample present in this analysis we relegate summary tables of the event efficiency for each cut to section 3.5.

3.4.1 Simulation Cuts

We use simulation to model the efficiency of the detector acceptance and the imposed selection requirements, and to model the shape of the invariant mass distributions of our B candidates. The simulated samples used in this analysis are listed in table 1 on page 15. Each simulated sample was produced for each year of data, 2016, 2017, and 2017 using Sim09j and split evenly between the polarity settings of the detector. The three MC samples where we reconstruct the soft pion were produced at a later point in time with Sim09k, but there is no difference between these two simulations settings significant for this analysis. All simulation was generated with LHCb acceptance cuts applied to the final state charged kaons and pions. We also requires that the final state charged kaons and pions, with the exception of eventually reconstructed soft pions, to satisfy the conditions that $10\,\text{mrad} < \theta < 400\,\text{mrad}$ and that the minimum p_{T} is $250\,\text{MeV}$.

3.4.2 Stripping

Data collected by the LHCb detector from 2016 to 2018, comprised of pp collisions at a center-of-mass energy of $13\,\text{TeV}$, corresponding to an integrated luminosity of 5.4 fb-1, is used for this analysis.

For each channel in section 3.2, candidates are selected using individual stripping lines. A summary of which lines are associated with which channel can be seen in table 2.

For each simulation sample in table 1, we require that candidates pass the relevant stripping lines associated with where the samples peak.

Table 2: Spectrum Label and Reconstructed Decay Modes and Stripping Lines

Spectrum	Stripping Line
$B^0 \to D^- D^+ K^{*0}$	B02DDKstBeauty2CharmLine
$B^0 \to \overline{D}^0 D^0 K^{*0}$	B02D0D0KstD02HHD02HHBeauty2CharmLine
$B^+ \to \overline{D}^0 D^+ K^{*0}$	B2DD0KstBeauty2CharmLine
$B^- \to D^- D^0 K^{*0}$	B2DD0KstBeauty2CharmLine
$B^+ \to \overline{D}^0 (D^* + \to D^0 \pi^+) K^{*0}$	B2DstD0KstBeauty2CharmLine
$B^+ \to \overline{D}^0 (D^0 \to K\pi\pi\pi) K^+$	B2D0D0KD02HHD02K3PiBeauty2CharmLine
$B^0 \to D^- (D^0 \to K\pi\pi\pi) K^+$	B02D0DKD02K3PiBeauty2CharmLine

The versions of Stripping used for processing data and simulation for the years 2016, 2017 and 2018 are 28r2, 29r2 and 34 respectively. DaVinci v45r6 is used to process stripped data and MC to make the ROOT tuples that are processed offline. table 3 and table 4 summarize the cuts used in the stripping lines. As a note the stripping also includes selections on the HLT2 Line.

In addition to stripping, we apply a prefilter during the generation of the ntuples for this analysis. The prefilter reduces the size of the ntuples with minimal loss of signal. They are summarized in table 4.

3.4.3 Offline Selection

We apply an offline selection on the Kaon Prob_NN > 0.3 for each kaon track. We apply a second offline section on the B's first Charm Daughters DIRA_ORIVX > 0, effectively requiring the D vertex downstream of the B, to reduce the number of charmless B decays. Finally we apply a cut on a 50 MeV window around the K^{*0} mass to focus on the region dominated by K^{*0}. We do not separate the non-resonant and resonant contributions in this analysis.

Table 3: Stripping selections applied on particles within all decays of the form $B \to \overline{D}^{(*)} D^{(*)+,0} K^{*0}$. Stripping selections here are shared between the spectrum in table 2. Charge conjugation is implied. If a cut is only applied to a specific charm mother it is specified in the subscript

Particle	Cut
π, K	$\chi^2_{\text{track}} < 4.0$
π, K	$p_{\text{T}} > 100\,\text{MeV}/c$
$(\pi, K)_D$	$p > 1000\,\text{MeV}/c$
$(\pi, K)_{K^{*0}}$	$p > 2000\,\text{MeV}/c$
π, K	$\chi^2_{\text{IP}}(\text{Primary}) > 4$
π, K	$Ghost_{\text{track}} < 0.4$
π_D	$\text{DLL}_{K\pi} < 10$
K_D	$\text{DLL}_{K\pi} > -5$
has at least 1 D/K^{*0} daughter	$\chi^2_{\text{track}} < 2.5$
has at least 1 D/K^{*0} daughter	$p_{\text{T}} > 500\,\text{MeV}/c$
has at least 1 D/K^{*0} daughter	$p > 5000\,\text{MeV}/c$
has at least 1 final state track	$p_{\text{T}} > 1.7\,\text{GeV}/c$
has at least 1 final state track	$p > 10\,\text{GeV}/c$
has at least 1 final state track	$IP_{PV} > 0.1mm$
D	$p_{\text{T}} > 1800\,\text{MeV}/c$
K^{*0}	$p_{\text{T}} > 1000\,\text{MeV}/c$
D	$M \in [m_{\text{PDG}}(D) \pm 100\,\text{MeV}/c^2]$
DK^{*0}	$\text{DOCA}_{K\pi} < 0.5\,\text{mm}$
D^0	$\chi^2_{\text{vtx}}/\text{ndf} < 10$
K^{*0}	$\chi^2_{\text{vtx}}/\text{ndf} < 16$
D^0	$\chi^2_{\text{vtx}-\text{PV}} > 36$
K^{*0}	$\chi^2_{\text{vtx}-\text{PV}} > 16$
D^0	$\text{DIRA}_{\text{PV}} > 0$
B	$M \in [4750, 6000]\,\text{MeV}/c^2$
B	$p_{\text{T}} > 5000\,\text{MeV}/c$
B	$\chi^2_{\text{vtx}}/\text{ndf} < 10$
B	$\tau_{\text{PV}} > 0.2\,\text{ps}$
B	$\chi^2_{\text{IP}}(\text{Primary}) < 25$
B	`BBDT` > 0.05
B	`Hlt2Topo or Hlt2IncPhi`
Event	$\#\text{longtracks} < 500$

Table 4: Additional selections applied before the reconstruction of our channels

Particle	Cut
π	$PIDK < 0$
K	$PIDK >= 4$
p	$PIDp >= 0$
D	Vertex position $z_D - z_B > -2millimeter$
K^{*0}	$M \in [750, 1050]$ MeV/c^2
B first daughters	$\chi^2_{\rm vtx}$/ndf $<= 5$

Table 5: A table summarizing how many candidates are removed from the $\overline{D}^0 D^0 K^{*0}$ spectrum in data and the appropriate MC samples. We remove these candidates so we can report and use uncorrelated efficiencies.

Source	Year	Candidates in $\overline{D}^0 D^0 K^{*0}$	Candidates in $\overline{D}^0(D^{*+} \to D^0\pi^+)K^{*0}$	Candidates Removed from $\overline{D}^0 D^0 K^{*0}$
Data	2016	3975	304	255
	2017	3996	362	305
	2018	4793	374	318
MC for $B^0 \to (D^{*-} \to \overline{D}^0\pi^-)(D^{*+} \to \overline{D}\pi^+)K^{*0}$	2016	3010	479	448
	2017	3667	577	537
	2018	3029	508	473
MC for $B^+ \to \overline{D}^0(D^{*+} \to \overline{D}\pi^+)K^{*0}$	2016	3452	633	593
	2017	3672	662	625
	2018	3224	599	574
MC for $B^+ \to \overline{D}^{*0}(D^{*+} \to \overline{D}\pi^+)K^{*0}$	2016	3106	529	497
	2017	3669	679	621
	2018	2916	464	426

3.4.4 Spectra Overlap

During the application of our offline selections we also handle candidates shared between the $\overline{D}^0 D^0 K^{*0}$ spectrum and the $\overline{D}^0(D^{*+} \to D^0\pi^+)K^{*0}$ spectrum, where in the latter case the soft pion is added to the candidate from the former spectrum. In order to properly measure the relative yields of these two modes any candidates shared between the two signal channel are removed from the $\overline{D}^0 D^0 K^{*0}$ spectrum and allowed to remain in the $\overline{D}^0(D^{*+} \to D^0\pi^+)K^{*0}$ spectrum. table 5 summarizes the number of candidates removed from data and the relevant MC samples, which shows that this occurs for almost, but not quite, all of the $\overline{D}^0(D^{*+} \to D^0\pi^+)K^{*0}$ candidates.

3.4.5 Trigger

The trigger requirements for data across all three years are fully summarized in table 6. Each trigger line contains a set of selections that help us eliminate uninteresting events from our samples. L0 lines rely on the hits associated with a track candidate passing a certain threshold of transverse energy within the relevant calorimeters described in **??** or in the case of the L0 Muon, the Muon tracking stations. If it passes this threshold the candidates event passes the trigger. Trigger lines can be broken up into two categories: Trigger on Signal (TOS) and Trigger Independent of Signal (TIS). If the tracks associated with a signal candidate are what pass the trigger, we call that TOS. If the a candidate passes the trigger line even without the tracks associated with our signal candidate it is considered TIS. Sometime a candidate can be classified as both TOS and TIS but this is negligible in our analysis and not considered. At the hardware level we allow candidates that pass either our L0 TOS condition or our L0 TIS condition. How we handle the differences in efficiencies and uncertainties across the data taking years and trigger conditions is described in section 3.5. For the HLT1 trigger, events have to pass either the 1TrackMVA or 2TrackMVA TOS decision lines. The purposed of these lines is to identify tracks with significant displacement from the PV. Both of these lines impose different conditions on the track $\chi^2_{\mathrm{vtx}}/\mathrm{ndf}$, the ghost probability of the track, the χ^2_{IP} and track p_{T}. For the HLT2, 2 and 3 and and 4 body topological TOS triggers are implemented. These lines implement a variation of a Boosted Decision Tree section 3.3.4 that ensures the classifier learns only a general set of traits associated with B-hadrons [30].

Table 6: Trigger Requirements for candidates across all spectra. We note that our stripping requirements contain the HLT2 lines. As such the efficiency reported for this requirement in appendix A is nearly 1

<table>
<tr><th colspan="2">Trigger Line Conditions</th></tr>
<tr><td>L0</td><td>B_L0HadronDecision_TOS || B_L0MuonDecision_TOS ||
B_L0ElectronDecision_TOS || B_L0PhotonDecision_TOS ||
B_L0HadronDecision_TIS || B_L0MuonDecision_TIS ||
B_L0ElectronDecision_TIS || B_L0PhotonDecision_TIS</td></tr>
<tr><td>HLT1</td><td>B_Hlt1TrackMVADecision_TOS || B _Hlt1TwoTrackMVADecision_TOS</td></tr>
<tr><td>HLT2</td><td>B_Hlt2Topo2BodyDecision_TOS || B_Hlt2Topo3BodyDecision_TOS ||
B_Hlt2Topo4BodyDecision_TOS</td></tr>
</table>

Table 7: Summary of our mass windows for the signal and sideband regions extracted from the fits to the D meson masses.The sideband region is used to estimate the yield of remaining candidates in the signal that are charmless background.

D Candidate	Fit Mean [MeV]	Signal Window [MeV, MeV]	Sideband Window [MeV, MeV]
D^-	1869.74	$[1852.24, 1887.24]$	$[1825.99, 1843.49]$ and $[1895.99, 1913.49]$
$\overline{D}^0$	1865.23	$[1846.83, 1883.63]$	$[1819.23, 1837.63]$ and $[1892.83, 1911.23]$
$D^0 \to K\pi\pi\pi$	1865.17	$[1849.17, 1881.17]$	$[1825.17, 1841.17]$ and $[1889.17, 1905.17]$
$D^{*-} - \overline{D}^0$	145.45	$[143.65, 147.25]$	$[140.95, 142.75]$ and $[148.15, 149.95]$

3.4.6 D Mass Window Cuts

We now determine an appropriate cut to apply to each D species reconstructed mass, before any decay tree fitter constraints are looked at in our analysis. For each D species we determine the appropriate mass window cut by preforming an unbinned maximum likelihood fit to the individual D meson masses in data. We distinguish between the D^0 reconstructed as $K^+\pi^-$ and the D^0 reconstructed as $K^+\pi^-\pi^-\pi^+$. The signal shape for the D^- and the $D^0 \to K^+\pi^-\pi^-\pi^+$ is modeled as as a left sided Crystal Ball function. The signal shape for the $D^0 \to K^+\pi^-$ is modeled as a bifurcated Gaussian with an exponential tail. Each fit contains an additional exponential component for the combinatorial background. We then choose our signal window as the region that captures at least 95% of the signal pdf. At this stage we also choose a sideband window separated from the signal window on both sides of the fit which we examine later on in order to estimate certain backgrounds. We also take advantage of the soft pion reconstruction and preform an additional fit with a right sided crystal ball function and cut on the mass of D^* minus the D^0 when we reconstructed the soft pion coming from $D^* decay$. table 7 provides a summary of the relevant parts of the fit and the resulting D window cuts. fig. 9 shows the results of our fit to each D species.

3.4.7 MC Truth Matching

As part of the event selection process, "truth matching" conditions can be applied to our simulation samples. Simulation stores information that matches the reconstructed tracks to

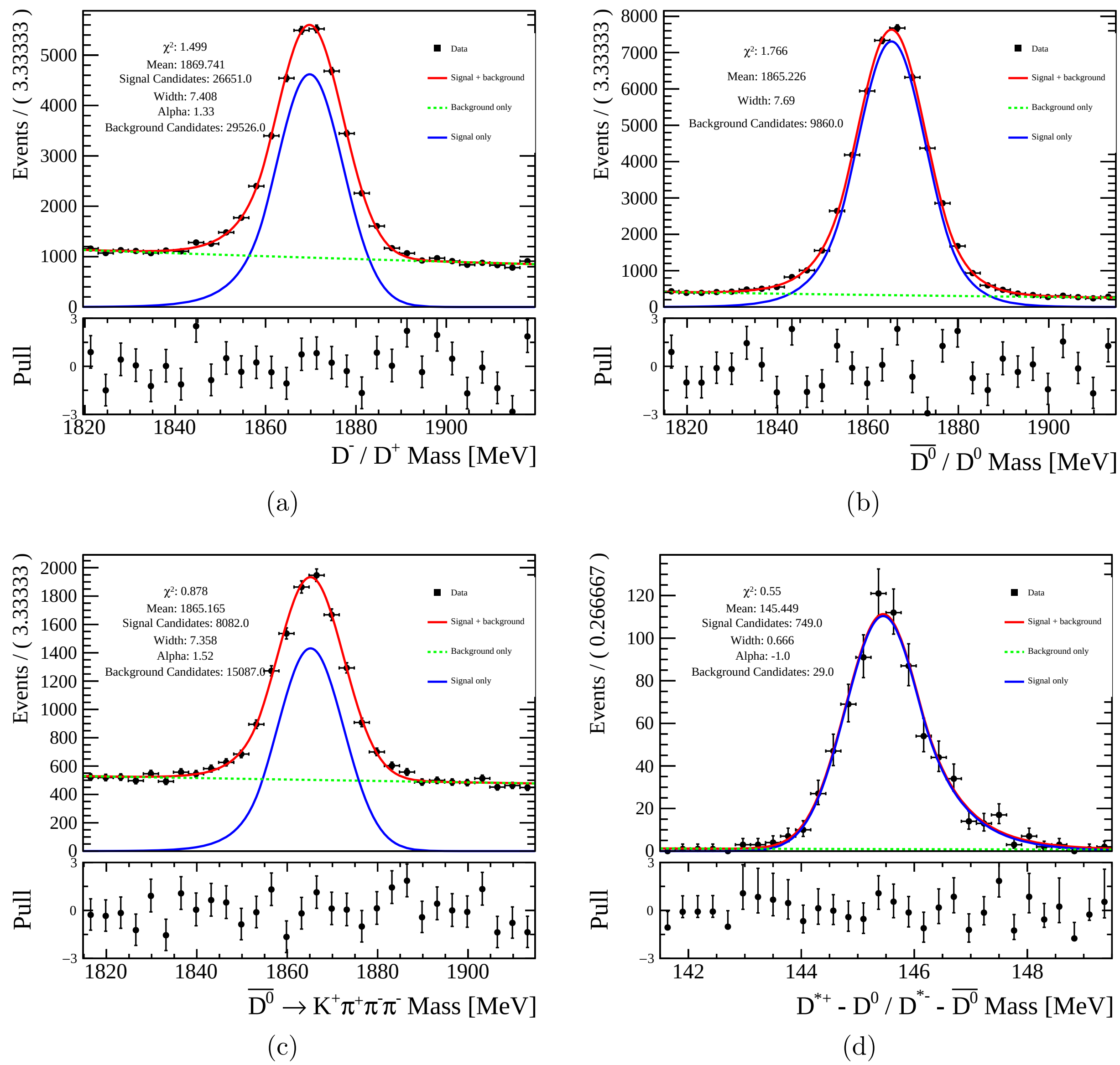


Figure 9: Fits to the reconstructed D meson masses, used to determine the signal and sideband region definitions for signal selection and background estimation that are reported in table 7. The tails help capture the higher rate on the left side due to radiative decays of the D mesons.

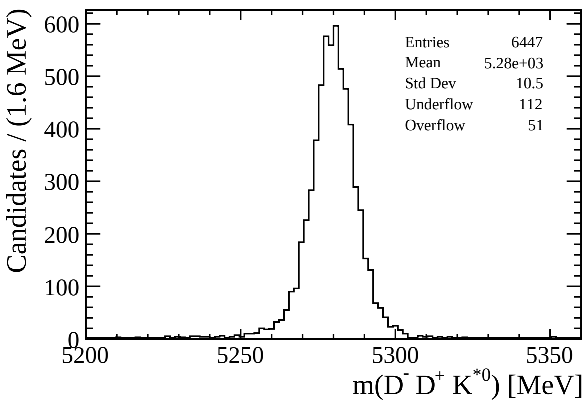
the generated tracks, and a candidates is considered matched when a sufficient number of the detector hits that define a track are shared with the generated track. Typically we want to use this information to eliminate any candidates that are not representative of our signal mode of interest in order to calculate the correct efficiencies. However, we do not apply any truth matching conditions on MC events in the event selection process. At this stage in the event selection nearly all events in simulation that would be considered true background are eliminated. Applying truth matching conditions on the B mothers and reconstructed track ID's removes a significant number of events, ranging from 10% to 15% of a specific simulation sample, that peak at the B mass as illustrated in figure fig. 10. Most of these events are labeled in the BKGCAT variable as ghost candidates because they have failed matching. In addition, applying further truth matching on the D meson ID's and K^{*0} ID serves to only eliminate multiple candidates where pions have been swapped between the D mesons or the Ds and the K^{*0}. This effect is relevant, particularly when estimating some background contributions, and we will deal with the multiple candidates in a consistent fashion between our data and MC samples, as described in section section 3.4.11.

3.4.8 Peaking Background

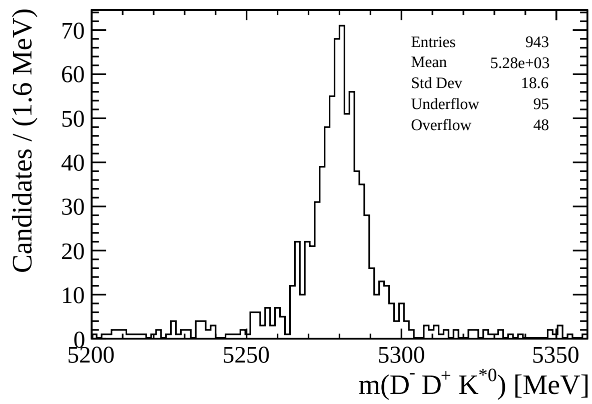
We eliminate several sources of peaking background that remain in our data samples through a series of specific cuts. The most dangerous background sources at this point are decays that peak underneath B peak that we aim to measure, typically from a decay of the form $B \to DDK$.

3.4.9 $B^{0,+} \to (D^{*-} \to \overline{D}^0 \pi^-) D^{0,+} K^+$

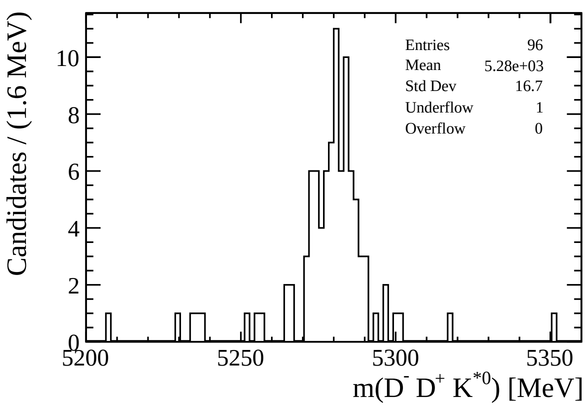
In this mode, we miss the soft pion from the excited D meson, and reconstruct the K*0 with it. This background exists in our $\overline{D}^0 D^0 K^{*0}$ and our $\overline{D}^0 D^+ K^{*0}$ spectrum. We can reconstruct the difference in the invariant masses of the $\overline{D}^0$ and the π^- from the K^{*0} minus the $\overline{D}^0$ in order to take advantage of removing any uncertainty in the reconstruction of the $\overline{D}^0$. In fig. 11 We can clearly see a peak at the $D^{*-} - \overline{D}^0$ mass, and apply a veto of all candidates below 150 MeV. The effect of this cut can be seen in figure fig. 11 where we plot the invariant mass of all tracks minus the pion candidate from the K^{*0}. The efficiency of this cut is $> 99\%$ efficient for all MC samples.



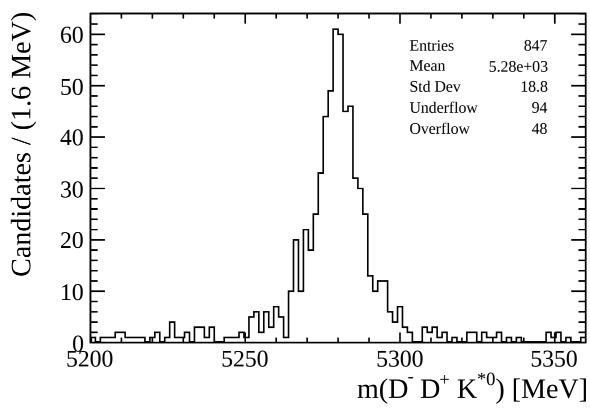
(a) The B mass for the $B^0 \to D^- D^+ K^{*0}$ MC sample at this stage in the event selection process



(b) The B mass for the full $B^0 \to D^- D^+ K^{*0}$ MC sample, with only events that fail at least one of the truth matching conditions on the B ID or one of the track ID's



(c) The B mass for the full $B^0 \to D^- D^+ K^{*0}$ MC sample, with only events that fail at least one of the truth matching conditions on the B ID or one of the track ID's, but pass the truth matching condition on the D's and K^*

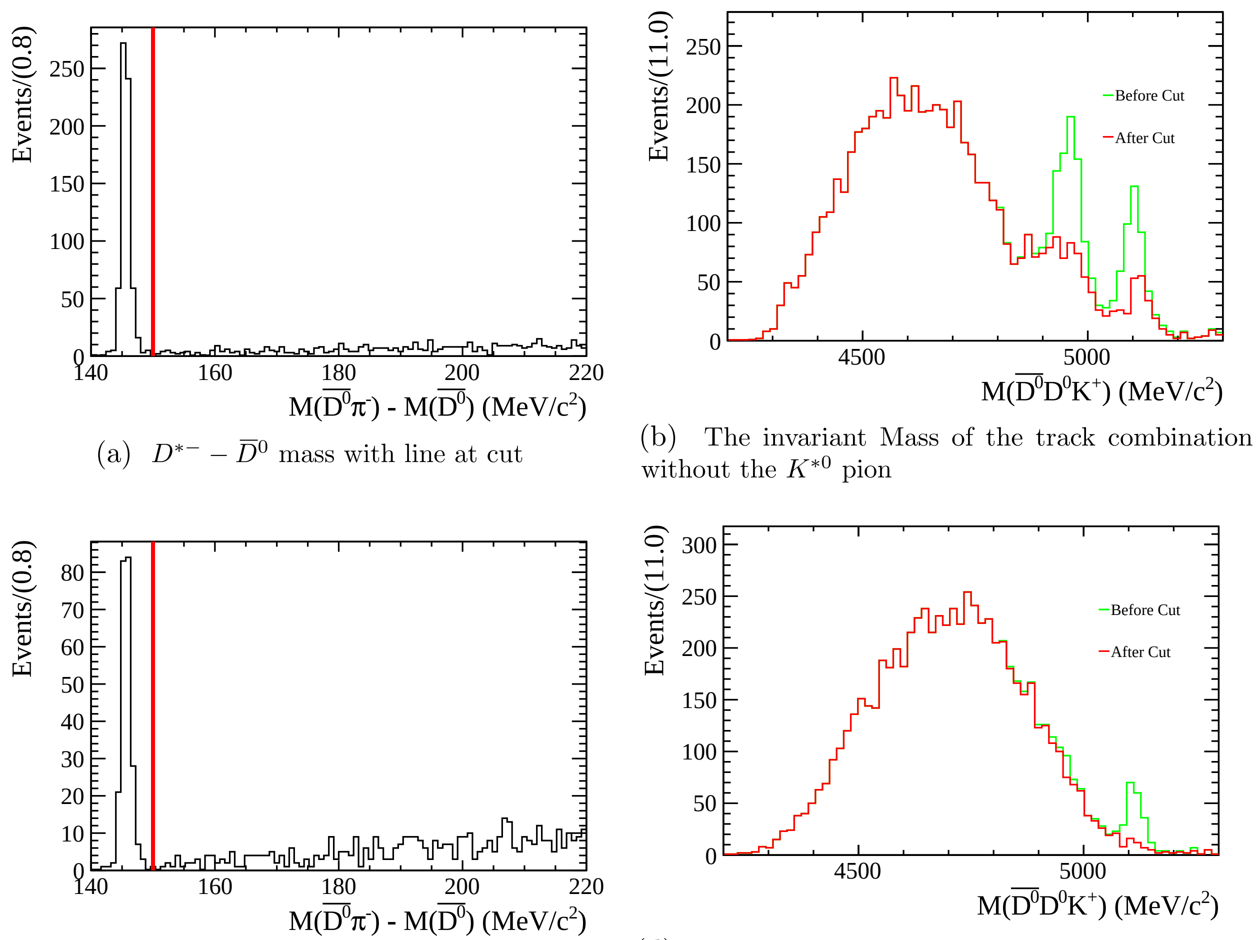


(d) The B mass for the full $B^0 \to D^- D^+ K^{*0}$ MC sample, with only events that fail at least one of the truth matching conditions on the B ID, one of the track IDs, or one of the D IDs or the K^{*0} ID

Figure 10: A breakdown of the effect of applying truth matching conditions to the $B^0 \to D^- D^+ K^{*0}$ MC sample. We see that events that fail truth matching nearly always peak at the B mass, and should be considered signal in our analysis.

3.4.10 Clone Tracks

In this section we remove clone tracks remaining in our Data and MC. Clone tracks occur when the same track is mistakenly used between two or more candidates. We plot the minimum value for the full set of angles between each two track combination in a given candidate in fig. 12. We veto at a value of 0.0005 for theta to eliminate candidates with clone



(a) $D^{*-} - \overline{D}^0$ mass with line at cut

(b) The invariant Mass of the track combination without the K^{*0} pion

(c) $D^{*-} - \overline{D}^0$ mass with line at cut

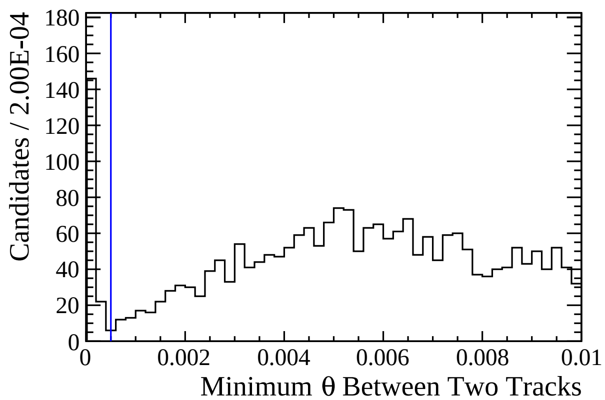
(d) The invariant Mass of the track combination without the K^{*0} pion

Figure 11: The effect of the Dstar veto in (top row) the $\overline{D}^0 D^0 K^{*0}$ and (bottom row) the $\overline{D}^0 D^+ K^{*0}$ spectra. The additional structure in fig. 11b ends up as combinatorial background in our final signal regions

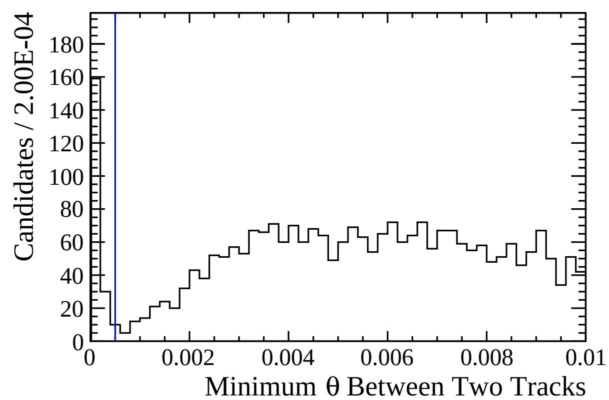
tracks.

3.4.11 Multiple Candidates

The final step in event selection is to handle any remaining multiple candidates in our simulation and data samples. We choose to select our individual candidates for these events randomly. Because we aim to estimate certain backgrounds contributions later on using the D window sideband region, we handle multiple candidates across the D Window signal region and our D window sideband regions simultaneously. This means that during the random



(a) Cut on candidates that contain clone tracks for the Z spectrum



(b) Cut on candidates that contain clone tracks for the N8 spectrum

Figure 12: A non-trivial amount of candidates in our ntuples are clone tracks. A cut on the minimum angles for tracks of candidates in the Z and N8 spectrums with a line at our choice of cut is shown here

selection, if an event exists in both the signal and sideband region, one will be randomly kept and one will be discarded. We summarize the effect of this cut, on our real Data in table 8. With our choice of the sideband window and the difference in efficiency between signal and sideband, less then 1% of the signal remains in the sideband and the majority of multiple candidates exists in the sideband only.

Table 8: Multiple candidate selection on D window signal and sideband regions. The candidate efficiency shows that at most ≈3 percent of events have multiple candidates. We also report the event efficiency of this selection, as removing multiple candidates from our D signal and sideband regions simultaneously means that events in signal can be lost.

Spectrum	Year	ϵ_{can} in Sig	ϵ_{can} in Sb	ϵ_{ev} in Sig	ϵ_{ev} in Sb
$D^-D^+K^{*0}$	2016	0.967 ± 0.006	0.942 ± 0.013	0.9923 ± 0.0029	0.984 ± 0.007
	2017	0.970 ± 0.005	0.926 ± 0.014	0.9906 ± 0.0031	0.971 ± 0.009
	2018	0.966 ± 0.005	0.945 ± 0.011	0.9955 ± 0.0020	0.980 ± 0.007
$\overline{D}^0D^0K^{*0}$	2016	0.9907 ± 0.0022	0.972 ± 0.012	0.9989 ± 0.0008	1.0 ± 0
	2017	0.9930 ± 0.0019	0.989 ± 0.008	0.9995 ± 0.0005	1.0 ± 0
	2018	0.9876 ± 0.0023	0.973 ± 0.011	0.9991 ± 0.0006	0.995 ± 0.005
$\overline{D}^0D^+K^{*0}$	2016	0.9865 ± 0.0026	0.965 ± 0.010	1.0 ± 0	0.9970 ± 0.0030
	2017	0.9830 ± 0.0028	0.973 ± 0.009	0.9981 ± 0.0009	0.991 ± 0.005
	2018	0.9866 ± 0.0023	0.964 ± 0.009	0.9984 ± 0.0008	0.9974 ± 0.0026
$D^-D^0K^{*0}$	2016	0.981 ± 0.005	0.948 ± 0.018	0.9972 ± 0.0020	0.973 ± 0.013
	2017	0.974 ± 0.006	0.950 ± 0.016	0.9956 ± 0.0025	0.983 ± 0.010
	2018	0.966 ± 0.006	0.942 ± 0.015	0.9917 ± 0.0031	0.980 ± 0.009
$\overline{D}^0(D^{*+} \to D^0\pi^+)K^{*0}$	2016	0.989 ± 0.008	1.0 ± 0	1.0 ± 0	1.0 ± 0
	2017	1.0 ± 0	1.0 ± 0	1.0 ± 0	1.0 ± 0
	2018	1.0 ± 0	1.0 ± 0	1.0 ± 0	1.0 ± 0
$\overline{D}^0(D^0 \to K^-\pi^+\pi^+\pi^-)K^+$	2016	0.981 ± 0.004	0.971 ± 0.009	0.9925 ± 0.0028	0.982 ± 0.007
	2017	0.9913 ± 0.0031	0.970 ± 0.009	0.9989 ± 0.0011	0.988 ± 0.006
	2018	0.9905 ± 0.0028	0.973 ± 0.008	0.9948 ± 0.0021	0.984 ± 0.006
$D^-(D^0 \to K^-\pi^+\pi^+\pi^-)K^+$	2016	0.978 ± 0.004	0.946 ± 0.012	0.9973 ± 0.0016	0.970 ± 0.009
	2017	0.982 ± 0.004	0.959 ± 0.010	0.9967 ± 0.0017	0.976 ± 0.008
	2018	0.974 ± 0.004	0.952 ± 0.009	0.9881 ± 0.0030	0.978 ± 0.007

3.5 Monte Carlo Efficiencies

Here we present a breakdown of the Monte Carlo event efficiencies in this analysis. We break down our efficiencies by year and mode in order to account for variations across the data taking periods. efficiencies can be found in appendix A. Systematic uncertainties on these numbers are considered later on.

1. $\epsilon_{generator}$ - This efficiency corresponds to the probability that a full B event described by a given decfile is generated and accepted by the generator level cuts from section section 3.4.1. We obtain this efficiency from MC generator statistics tables produced after the production of the MC samples.

2. $\epsilon_{stripping}$ - This efficiency corresponds to the probability that an event will pass the stripping line requirements mentioned in section 3.4.2. These stripping lines require that a candidate pass any of the Hlt2Topo N Body triggers or any of the Hlt2IncPhi triggers.

3. $\epsilon_{offline}$ - This efficiency corresponds to the probability that an event will pass the offline requirements from table 4 and the additional cuts in section 3.4.3 At this stage we remove the overlapping candidates between ZZ and ST.

4. $\epsilon_{trigger}$ - This efficiency corresponds to the probability of a event passing the combination of the L0, HLT1, and HLT2 trigger selections in section 3.4.5. The efficiency of the individual L0 lines, broken up into disjoint TOS and TIS regions is reported in appendix A.

5. ϵ_{dwin} - This efficiency corresponds to the probability of a event passing the D window cut from section 3.4.6

6. $\epsilon_{peakbkg}$ - This efficiency corresponds to the probability of a event passing the the appropriate bkg vetoes for its decay described in section 3.4.8

7. ϵ_{clone} - This efficiency corresponds to the probability of a event passing the clone tracks cut from section 3.4.10

8. $\epsilon_{multcan}$ - This efficiency corresponds to the probability of event passing the final selection on multiple candidates. Because we apply this selection across the signal region and sideband regions simultaneously, some number of events in the signal region will be cut. Only those present in the signal region can end up in the final signal yield, but it is important to know the rate in the sideband region as well for background estimation.

We report here a small sample of efficiencies in table 9 and table 10.

3.5.1 ReDecay Correction

All MC samples in this analysis were produced using the ReDecay package [31]. While this does speed up simulation time, this does result in an underestimation of the uncertainty in out reconstruction efficiency. In order to compensate for this, we apply a correction in the following way. For a given MC sample, each event has both a ReDecay Event Number and ReDecay Run Number, in such a way that each event can be uniquely identified. For each unique ReDecay iteration (Run Number) the total number of events belonging to that iteration is counted (Event Number) and stored. Each of these sums is assigned a random weight from a Poisson distribution with a mean of one. A final sum of these weighted sums is calculated as the numerator in our reconstruction efficiency calculation, and this process is carried out N = 5000 times. The final result is a reconstruction efficiency with the same nominal value, but a larger uncertainty.

Table 9: Summary of Event Efficiencies for a 6 track, 7 track, and 8 track mode. We point out that the generator and stripping efficiency for 12 are nearly twice as big as 1. The multiplicity of reconstructed tracks (6 vs 8) is mainly responsible for this as our generator cuts and stripping lines on the two samples are nearly identical

Source	Year	Number Accepted	$\epsilon_{Generator}$	$\epsilon_{stripping}$	$\epsilon_{offline}$	$\epsilon_{trigger}$
$B^0 \to D^- D^+ K^{*0}$	2016	655750	5.34 ± 0.09	0.463 ± 0.008	94.1 ± 0.4	94.9 ± 0.4
	2017	603999	5.16 ± 0.09	0.524 ± 0.009	94.5 ± 0.4	94.1 ± 0.4
	2018	689999	5.24 ± 0.09	0.422 ± 0.008	94.3 ± 0.4	94.5 ± 0.5
$B^+ \to \overline{D}^0(D^{*+} \to D^+\pi^0)K^{*0}$	2016	613999	6.82 ± 0.11	0.677 ± 0.010	94.7 ± 0.4	95.73 ± 0.34
	2017	604000	6.79 ± 0.12	0.736 ± 0.011	94.4 ± 0.4	95.53 ± 0.34
	2018	689948	6.99 ± 0.11	0.630 ± 0.010	94.8 ± 0.4	95.46 ± 0.34
$B^0 \to \overline{D}^{*0} D^{*0} K^{*0}$	2016	717998	9.34 ± 0.14	0.956 ± 0.011	94.98 ± 0.28	97.30 ± 0.22
	2017	618000	8.90 ± 0.15	1.073 ± 0.013	94.02 ± 0.32	97.49 ± 0.22
	2018	601848	9.32 ± 0.16	0.945 ± 0.012	95.08 ± 0.31	96.98 ± 0.25

Table 10: Summary of Event Efficiencies for for a 6 track, 7 track, and 8 track mode.

Source	Year	ϵ_D	$\epsilon_{bkgveto}$	ϵ_{clone}	$\epsilon_{MultipleCandidate}$
$B^0 \to D^- D^+ K^{*0}$	2016	82.2 ± 0.8	100.0 ± 0	99.37 ± 0.17	0.9902 ± 0.0022
	2017	83.8 ± 0.7	100.0 ± 0	99.36 ± 0.17	0.9866 ± 0.0025
	2018	80.8 ± 0.8	100.0 ± 0	99.18 ± 0.20	0.9917 ± 0.0021
$B^+ \to \overline{D}^0(D^{*+} \to D^+\pi^0)K^{*0}$	2016	85.4 ± 0.6	99.93 ± 0.05	99.62 ± 0.12	0.9979 ± 0.0009
	2017	85.4 ± 0.6	99.67 ± 0.10	99.34 ± 0.15	0.9993 ± 0.0005
	2018	85.9 ± 0.6	99.90 ± 0.06	99.54 ± 0.12	0.99967 ± 0.00033
$B^0 \to \overline{D}^{*0} D^{*0} K^{*0}$	2016	87.1 ± 0.5	99.979 ± 0.021	99.64 ± 0.09	0.99979 ± 0.00021
	2017	86.8 ± 0.5	100.0 ± 0	99.69 ± 0.08	1.0 ± 0
	2018	85.7 ± 0.5	99.974 ± 0.026	99.68 ± 0.09	0.9995 ± 0.0004

3.6 Normalization Channels

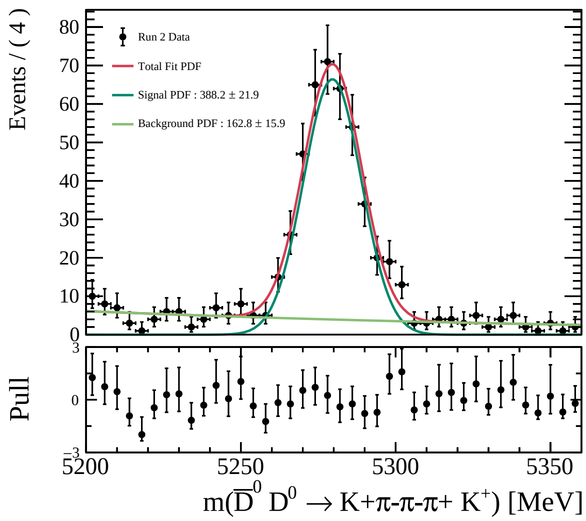
The normalization channels used in this analysis are $B \to D^-(D^0 \to K\pi\pi\pi)K^+$ (N8) which is used for the Z spectrum and the $B^+ \to \overline{D}^0(D^0 \to K\pi\pi\pi)K^+$ (N7) which is used for rest of the signal channels. The event selections applied to each signal channel is applied as well to each normalization channel, with the exception of the window on the K^{*0}. We present a comparison of the known branching fractions to our measured yields for these modes, including the uncertainties present in this analysis, in section 3.9.2

3.6.1 Fits to Normalization Channels

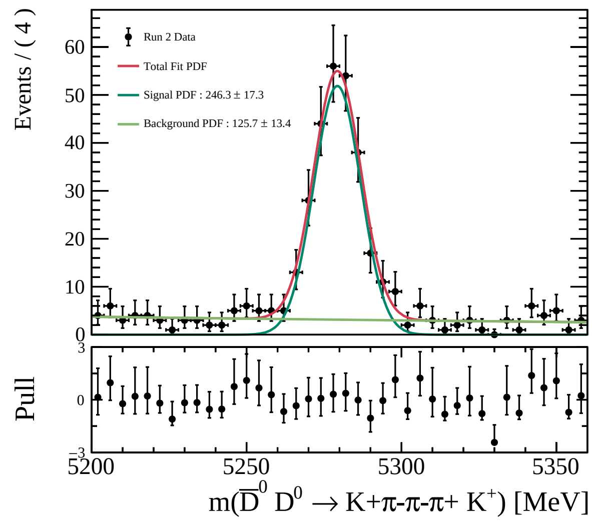
Both normalization modes are broken down by year and L0 trigger condition and an unbinned maximum likelihood fit to the B mass is performed, with DTF constraints applied to the D Mesons masses ensuring that the B meson points at the primary vertex. The signal shape is modeled as a sum of two Gaussian distributions with a shared mean (DG). The combinatorial background is modeled as an exponential function. All parameters in the fit are allowed to float for the fit to Data. The fit range is constrained to (5200, 5360) MeV/c^2.

3.6.2 Normalization Yields and Known Branching Fractions

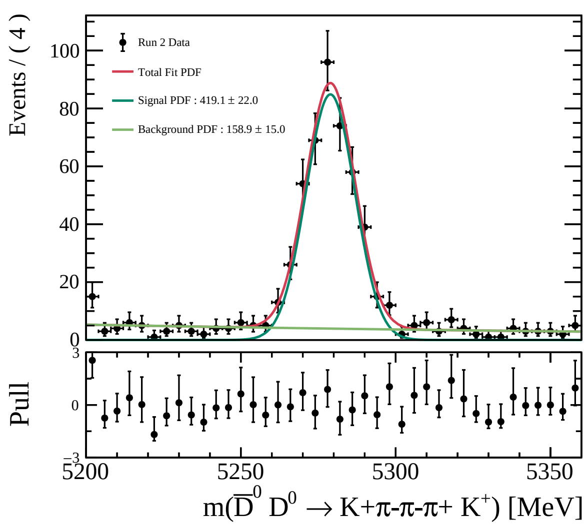
In table 11 we report the final yields of our normalization modes used in this analysis. The TOS and TIS samples are disjoint and are defined as events that pass our L0 TOS line and events that do not pass our L0 TOS line but do pass our L0 TIS line respectively.



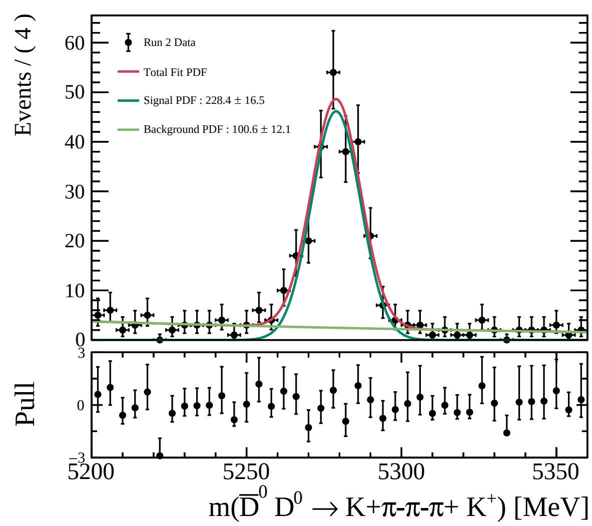
(a) TOS trigger condition in 2016 for our 7 track normalization mode



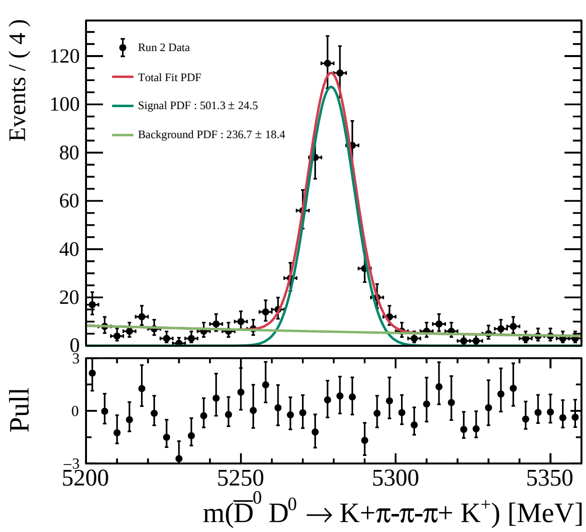
(b) TIS trigger condition in 2016 for our 7 track normalization mode



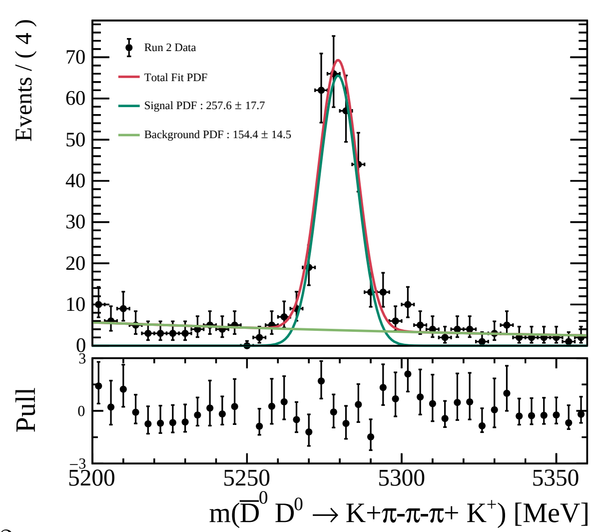
(c) TOS trigger condition in 2017 for our 7 track normalization mode



(d) TIS trigger condition in 2017 for our 7 track normalization mode



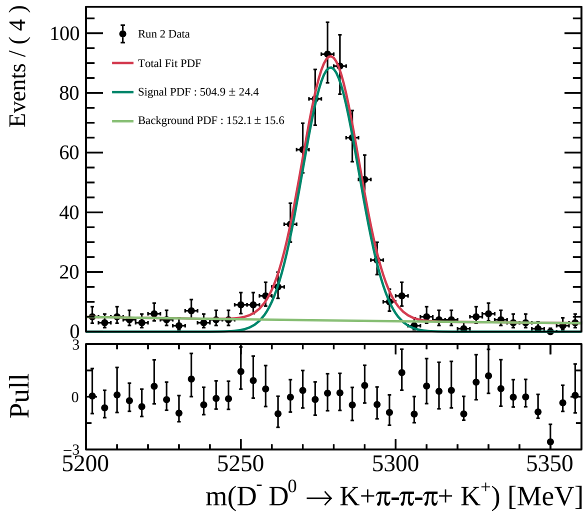
(e) TOS trigger condition in 2018 for our 7 track normalization mode



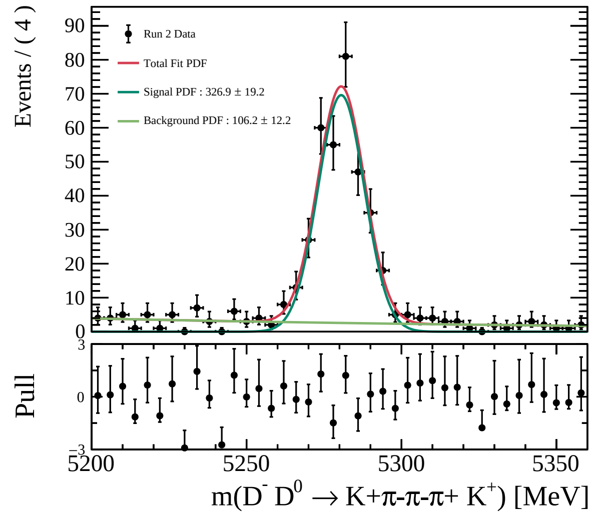
(f) TIS trigger condition in 2018 for our 7 track normalization mode

Table 11: Yields of Normalization Modes broken down by year and L0 trigger condition

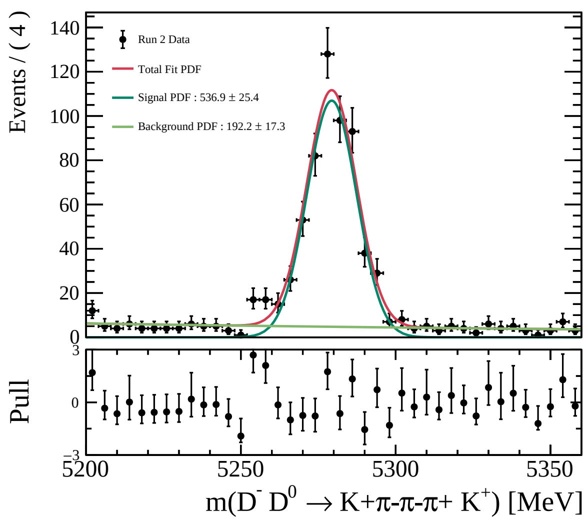
Spectrum	Year	Trigger	Fit Mean [MeV]	Fit Yield
$\overline{D}^0(D^0 \to K^-\pi^+\pi^+\pi^-)K^+$	2016	TOS	5279.6 ± 0.6	388 ± 22
		TIS	5279.3 ± 0.6	246 ± 17
	2017	TOS	5278.9 ± 0.4	419 ± 22
		TIS	5278.9 ± 0.6	228 ± 17
	2018	TOS	5279.2 ± 0.4	501 ± 25
		TIS	5279.4 ± 0.5	258 ± 18
$D^-(D^0 \to K^-\pi^+\pi^+\pi^-)K^+$	2016	TOS	5279.0 ± 0.5	505 ± 24
		TIS	5280.6 ± 0.5	327 ± 19
	2017	TOS	5279.4 ± 0.4	537 ± 25
		TIS	5279.7 ± 0.5	352 ± 20
	2018	TOS	5279.8 ± 0.4	599 ± 26
		TIS	5280.1 ± 0.4	398 ± 21



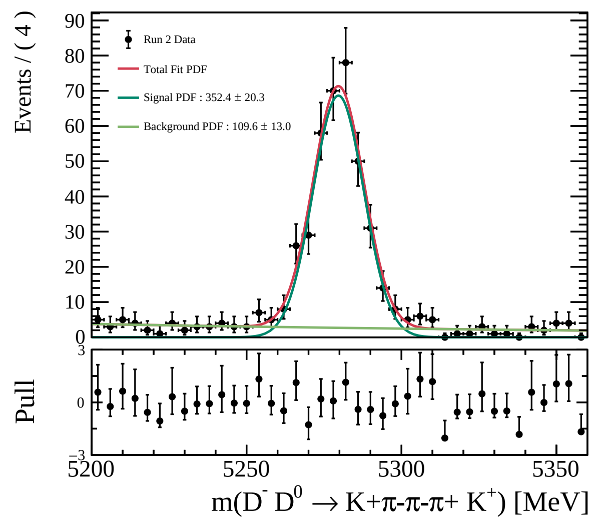
(a) TOS trigger condition in 2016 for our 7 track normalization mode



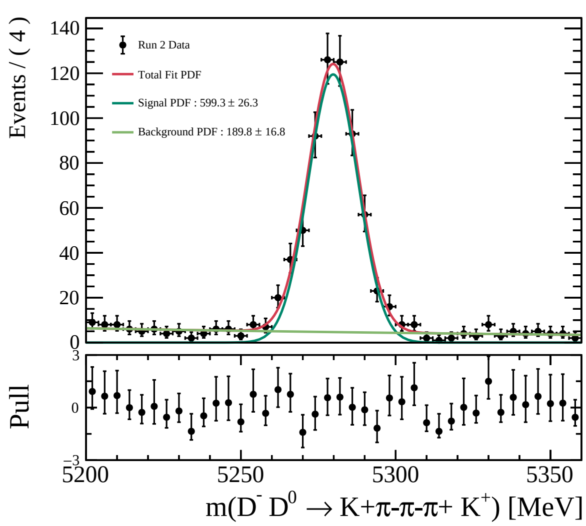
(b) TIS trigger condition in 2016 for our 7 track normalization mode



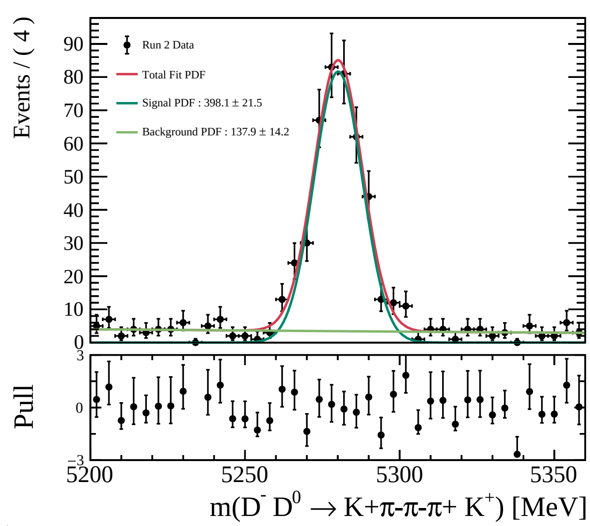
(c) TOS trigger condition in 2017 for our 7 track normalization mode



(d) TIS trigger condition in 2017 for our 7 track normalization mode



(e) TOS trigger condition in 2018 for our 7 track normalization mode



(f) TIS trigger condition in 2018 for our 7 track normalization mode

3.6.3 Normalization Factors

In table 12 we report the known branching fraction values of the D meson decays and known B Meson decays in this analysis. It is important to note that we choose to use the most recent measured value of $(1.31 \pm 0.14) \times 10^{-3}$ for the branching fraction of $B^+ \to \overline{D}^0 D^0 K^+$ [32], as opposed to the PDG average $(1.45 \pm 0.22) \times 10^{-3}$ which includes an uncertainty scale factor of 2.6.

Table 12: Input D meson branching fractions used in this analysis, taken from Ref [1].

Decay	Branching fraction (%)
$B^+ \to D^0(D^0 \to K^-\pi^+\pi^+\pi^-)K^+$	1.31 ± 0.139
$B^0 \to D^-(D^0 \to K^-\pi^+\pi^+\pi^-)K^+$	1.07 ± 0.101
$D^0 \to K^-\pi^+$	3.95 ± 0.031
$D^+ \to K^-\pi^+\pi^+$	9.38 ± 0.16
$D^0 \to K^-\pi^+\pi^+\pi^-$	8.23 ± 0.14

In table table 13 we report the final nominal normalization coefficients, $\mathcal{C}$, that determine the size of our branching fractions of interest. We handle the differences in simulation and data between data taking year and L0 trigger condition by summing the product of the relative efficiencies and normalization yield accross the 6 regions.

$$\mathcal{C}_{XY} = \frac{[\prod \mathcal{B}(D)]}{\mathcal{B}(NORM)} \sum_{year,trigger} \epsilon_{\text{rel(year,trigger)}} \times \mathcal{N}(\text{NORM})_{year,trigger} \tag{10}$$

Where X is one of our data spectra and Y corresponds to to the branching fraction of one our B decays using the labeling scheme from section 3.2. As an example of how $\mathcal{C}$ looks in our final fit, we can rewrite eq. (4) in terms of values of $\mathcal{C}$.

Table 13: Values for our normalization coefficients, corresponding to our simulated decay modes. These values include the correction for our use of ReDecay, but do not include any systematics or our correction for the remaining charmless background. We indicate the normalization coefficients for our reconstruction of the soft pion by bolding the soft pion

Source	Value
$B^0 \to D^- D^+ K^{*0}$	$(2.2 \pm 0.4) \times 10^6$
$B^0 \to (D^{*-} \to D^- \pi^0) D^+ K^{*0}$	$(6.3 \pm 1.0) \times 10^5$
$B^0 \to (D^{*-} \to D^- \pi^0)(D^{*+} \to D^+ \pi^0) K^{*0}$	$(1.6 \pm 0.3) \times 10^5$
$B^0 \to (D^{*-} \to \overline{D}^0 \pi^-)(D^{*+} \to \overline{D} \pi^+) K^{*0}$	$(3.5 \pm 0.6) \times 10^5$
$B^+ \to \overline{D}^0 (D^{*+} \to \overline{D} \pi^+) K^{*0}$	$(5.7 \pm 0.9) \times 10^5$
$B^+ \to \overline{D}^{*0} (D^{*+} \to \overline{D} \pi^+) K^{*0}$	$(5.5 \pm 0.9) \times 10^5$
$B^0 \to \overline{D}^0 D^0 K^{*0}$	$(1.1 \pm 0.2) \times 10^6$
$B^0 \to \overline{D}^{*0} D^0 K^{*0} + B^0 \to \overline{D}^0 D^{*0} K^{*0}$	$(1.1 \pm 0.2) \times 10^6$
$B^0 \to \overline{D}^{*0} D^{*0} K^{*0}$	$(9.7 \pm 1.5) \times 10^5$
$B^0 \to (D^{*-} \to \overline{D}^0 \pi^-) D^+ K^{*0}$	$(7.7 \pm 1.2) \times 10^5$
$B^0 \to (D^{*-} \to \overline{D}^0 \pi^-)(D^{*+} \to D^+ \pi^0) K^{*0}$	$(2.1 \pm 0.3) \times 10^5$
$B^+ \to \overline{D}^0 D^+ K^{*0}$	$(1.4 \pm 0.2) \times 10^6$
$B^+ \to \overline{D}^{*0} D^+ K^{*0}$	$(1.2 \pm 0.2) \times 10^6$
$B^+ \to \overline{D}^0 (D^{*+} \to D^+ \pi^0) K^{*0}$	$(3.9 \pm 0.6) \times 10^5$
$B^+ \to \overline{D}^{*0} (D^{*+} \to D^+ \pi^0) K^{*0}$	$(3.6 \pm 0.6) \times 10^5$
$B^0 \to (D^{*-} \to \overline{D}^0 \pi^-)(D^{*+} \to \overline{D} \pi^+) K^{*0}$	$(8.2 \pm 1.3) \times 10^4$
$B^+ \to \overline{D}^0 (D^{*+} \to \overline{D} \pi^+) K^{*0}$	$(1.6 \pm 0.3) \times 10^5$
$B^+ \to \overline{D}^{*0} (D^{*+} \to \overline{D} \pi^+) K^{*0}$	$(1.3 \pm 0.2) \times 10^5$

$$M_2^P = \mathcal{C}_{P4}\mathcal{B}_4 + \mathcal{C}_{P8}\mathcal{B}_8. \tag{11}$$

The covariance matrix relating these values and there uncertainties is relegated to appendix G. In section 3.8 we handle the systematics present on these values, and in section 3.10 we describe out the correlations between these values are handled in our simultaneous fit. In section 3.9.1 we correct eq. (4) for the reamining charmless background that contributed to the signal peaks.

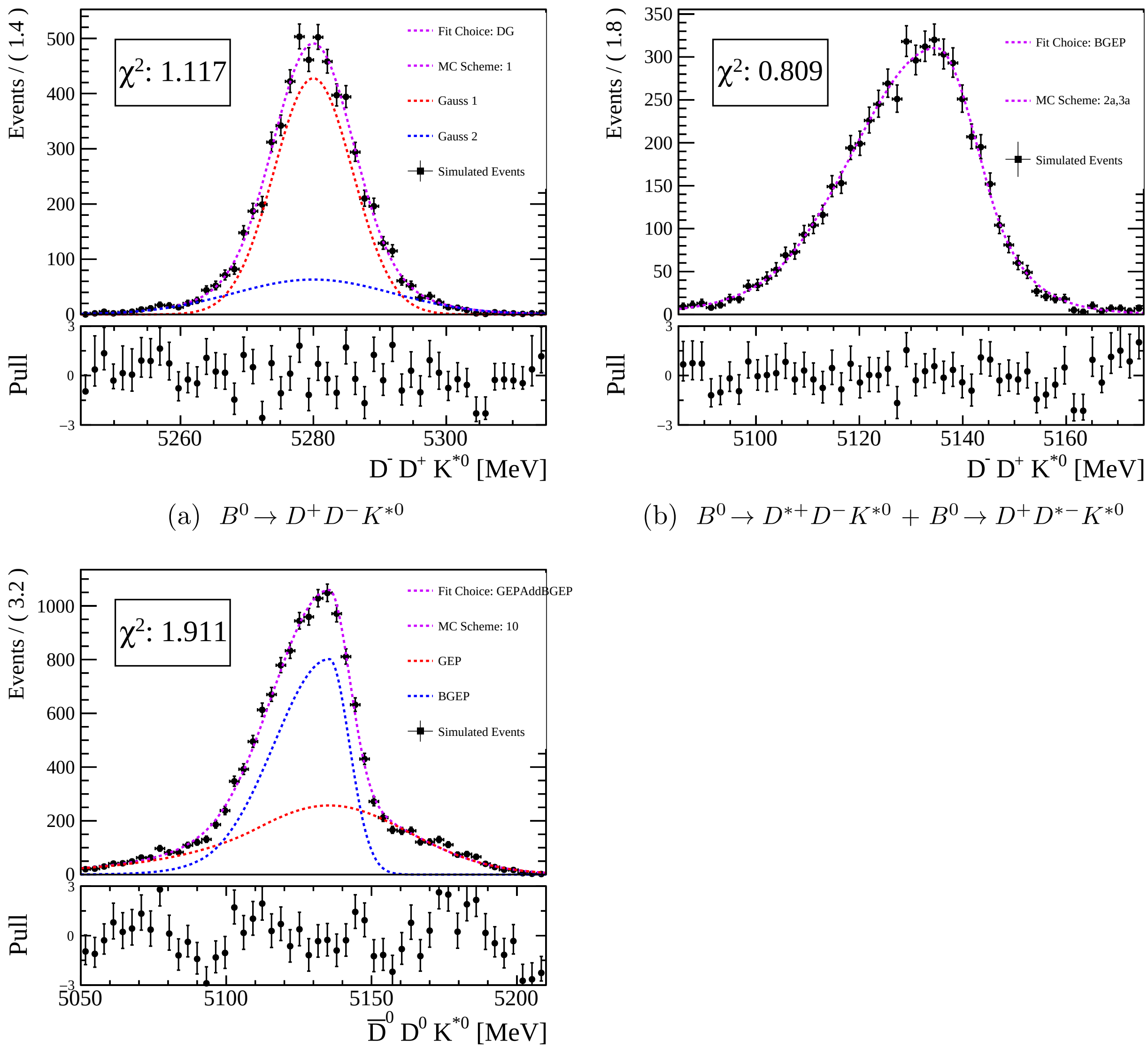
3.7 Fit Studies

Before constructing the full fit to produce the final branching fraction results, we investigate the fit shapes using separate fits to the data and simulated samples for each peak. We note that from this point on we look at the B Mass with the DTF constraints unless otherwise noted. These fits are necessary to study the choice of PDFs that we will use in the final simultaneous fit, as well as certain corrections and systematic uncertainties.

In section section 3.7.1 we examine the invariant B mass spectra of our MC samples and see that, for a given peak in data, there are significant enough differences in the shapes of the contributing decays that we can attempt to describe a given peak in data in terms of a sum of the PDFs extracted from MC. We also choose the PDF shapes for each MC sample that will be used to fit to our signal peaks in data. In section section 3.7.2 we describe the construction of separate fits to each signal peak in our data spectra. These fits are necessary to help us estimate certain corrections and systematic uncertainties which are described in section 3.8. We also use these fits in section section 3.7.3, where we examine the resolution differences in MC and Data, and conclude that best strategy is to constrain ourselves to the parameter values from our MC fits in section 3.7.1 while preforming a convolution of the MC PDFs with a Gaussian smearing function to capture the difference in resolution effects between data and MC.

3.7.1 Fits to MC

We carry out separate fits to each simulated sample listed in table 1. table 14 Summarizes our choice of best PDFs and the relevant results of each fit. A description of each type of PDF utilized is included in appendix . For simulations that include decays with excited D mesons that decay into a D and a soft photon, we describe our full PDFs as a sum of two distinct PDFs with one capturing the effect of the limited phase space of the decay channel to the soft photon. Because of the large number of simulation samples present in this analysis, we relegate the majority of fit plots to appendix C, and instead only show a select sample of the fits in this section.



(a) $B^0 \to D^+ D^- K^{*0}$

(b) $B^0 \to D^{*+} D^- K^{*0} + B^0 \to D^+ D^{*-} K^{*0}$

(c) $B^0 \to \overline{D}^{*0} D^0 K^{*0}$

Figure 15: Fits to select MC samples showing how different decay channels require different choices of our PDFs. Note the additional Gaussian component needed to model the soft particle contribution to fig. 15c

Table 14: Summary of Fit PDFs and parameter values for each fit to our MC samples. Full descriptions of the PDF shapes can be found in appendix B

Source	Shape	Mean	Width 1	Width 2	Alpha 1	Alpha 2
$B^0 \to D^- D^+ K^{*0}$	DG	5279.98 ± 0.09	5.93 ± 0.18	12.0 ± 0.7	–	–
$B^0 \to (D^{*-} \to D^- \pi^0) D^+ K^{*0}$	BGEP	5134.9 ± 0.8	16.0 ± 0.7	7.7 ± 1.0	1.56 ± 0.11	1.02 ± 0.14
$B^0 \to (D^{*-} \to D^- \pi^0)(D^{*+} \to D^+ \pi^0) K^{*0}$	BGEP	4987.5 ± 1.0	15.5 ± 0.9	11.5 ± 1.0	1.42 ± 0.12	1.30 ± 0.14
$B^0 \to (D^{*-} \to \overline{D}^0 \pi^-)(D^{*+} \to D^+ \pi^0) K^{*0}$	BGEP	4983.2 ± 0.7	16.4 ± 0.6	10.8 ± 0.5	1.67 ± 0.11	1.51 ± 0.09
$B^0 \to (D^{*-} \to \overline{D}^0 \pi^-)(D^{*+} \to \overline{D} \pi^+) K^{*0}$	BGEP	4978.5 ± 0.6	16.2 ± 0.5	10.8 ± 0.4	1.90 ± 0.16	2.08 ± 0.18
$B^0 \to (D^{*-} \to \overline{D}^0 \pi^-)(D^{*+} \to \overline{D} \pi^+) K^{*0}$	BGEP	5129.2 ± 0.9	11.8 ± 0.6	9.0 ± 0.7	2.26 ± 0.35	2.1 ± 0.5
$B^+ \to \overline{D}^0 D^+ K^{*0}$	BGEP	5279.43 ± 0.29	6.05 ± 0.29	6.23 ± 0.26	1.29 ± 0.09	1.51 ± 0.11
$B^+ \to \overline{D}^{*0} D^+ K^{*0}$	GEP (45%)	5135.0 ± 0.6	26.8 ± 0.7	–	0.95 ± 0.05	–
	BGEP (55%)	5135.0 ± 0.6	18.5 ± 0.7	7.1 ± 0.5	2.11 ± 0.16	4.0 ± 2.5
$B^+ \to \overline{D}^0 (D^{*+} \to D^+ \pi^0) K^{*0}$	G (10%)	5134.2 ± 0.4	22.2 ± 1.4	–	–	–
	BGEP (90%)	5134.2 ± 0.4	15.84 ± 0.33	7.93 ± 0.27	2.48 ± 0.29	4.00 ± 0.33
$B^+ \to \overline{D}^0 (D^{*+} \to \overline{D} \pi^+) K^{*0}$	BGEP	5131.4 ± 0.5	16.67 ± 0.35	7.0 ± 0.4	2.46 ± 0.35	1.69 ± 0.14
$B^+ \to \overline{D}^0 (D^{*+} \to \overline{D} \pi^+) K^{*0}$	BGEP	5278.7 ± 0.5	5.5 ± 0.5	6.30 ± 0.33	1.31 ± 0.18	2.2 ± 0.4
$B^+ \to \overline{D}^{*0} (D^{*+} \to D^+ \pi^0) K^{*0}$	G (47%)	4984.2 ± 0.6	28.9 ± 0.8	–	–	–
	BGEP (53%)	4984.2 ± 0.6	14.2 ± 0.8	12.7 ± 0.7	3.2 ± 1.5	2.4 ± 0.4
$B^+ \to \overline{D}^{*0} (D^{*+} \to \overline{D} \pi^+) K^{*0}$	GEP (47%)	4982.5 ± 1.2	24.0 ± 1.0	–	1.39 ± 0.26	–
	BGEP (53%)	4982.5 ± 1.2	17.0 ± 2.0	10.1 ± 1.0	1.22 ± 0.21	2.7 ± 0.8
$B^+ \to \overline{D}^{*0} (D^{*+} \to \overline{D} \pi^+) K^{*0}$	BGEP	5135.9 ± 2.5	17.6 ± 2.7	7 ± 5	1.36 ± 0.33	0.6 ± 0.4
$B^0 \to \overline{D}^0 D^0 K^{*0}$	BGEP	5279.49 ± 0.23	5.69 ± 0.24	6.05 ± 0.19	1.20 ± 0.06	1.57 ± 0.08
$B^0 \to \overline{D}^{*0} D^0 K^{*0} + B^0 \to \overline{D}^0 D^{*0} K^{*0}$	GEP (42%)	5135.4 ± 0.5	27.6 ± 0.6	–	0.91 ± 0.05	–
	BGEP (58%)	5135.4 ± 0.5	19.0 ± 0.5	7.0 ± 0.4	2.05 ± 0.21	3.6 ± 1.5
$B^0 \to \overline{D}^{*0} D^{*0} K^{*0}$	G (43%)	4985.8 ± 1.6	29.6 ± 1.1	–	–	–
	BGEP (57%)	4985.8 ± 1.6	19.3 ± 2.2	13.0 ± 2.6	1.03 ± 0.10	0.94 ± 0.19

3.7.2 Discrete Fit PDF

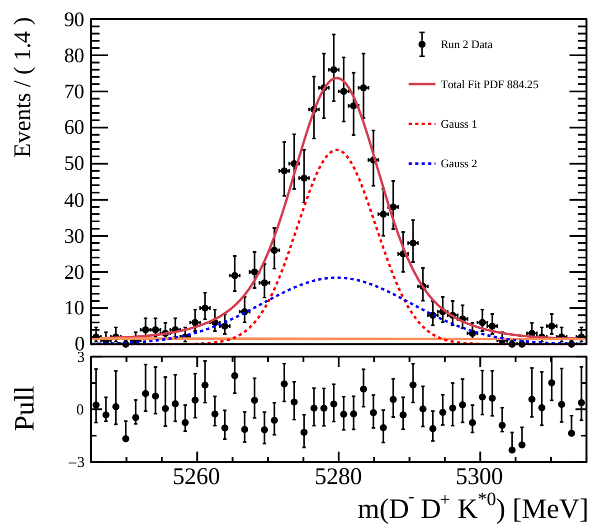
We now construct distinct fits to each of the 11 signal peaks in our 5 spectra separately, ignoring any correlations. For fits to the DTF constrained signal peaks with only one contributing signal decay mode, we constrain ourselves to the PDFs from table table 14. Because we are not preforming a simultaneous fit, we cannot yet separate the different signal contributions to peaks with multiple contributing decay modes. Therefore we fit these peaks with either a Bifurcated Gaussian with an Exponential Tail, or with the sum of a Bifurcated Gaussian with an Exponential Tail and a Gaussian when looking at peaks with a significant photon component. For fits to the mass with no decay tree fitter constraints we fit our signal to a Gaussian. These fits are not used directly in the final computation of the branching fractions.

These fits serve multiple purposes. They are needed to motivate further corrections and systematic uncertainties due to the unmodeled intermediate resonance structure and the contamination of peaking background with charm mesons. The success of these fits also shows that we are not sensitive to the line-shape effects of the individual spin structures and resonances in peaks with missing particles.

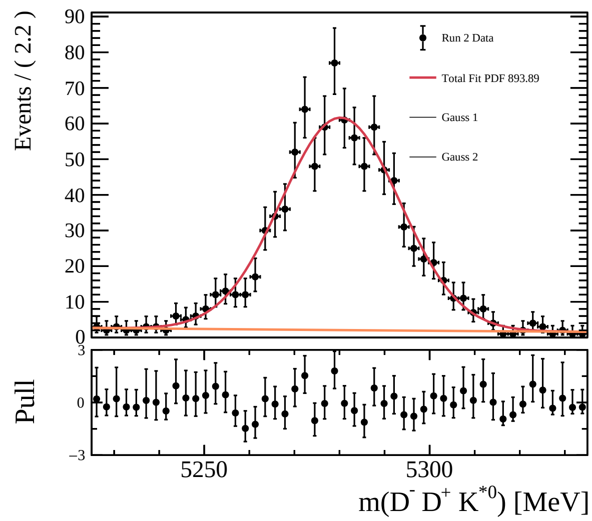
We show several of these fits in fig. 16 and relegate the rest to appendix D.

3.7.3 Resolution Tests

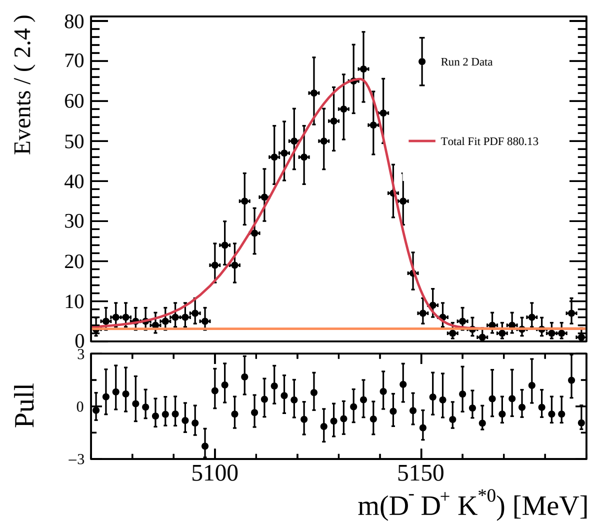
We now more closely examine the resolution differences across simulation and data, and show how we account for them in the final fit. In fig. 17 we examine the resolution difference between simulation and data across certain signal regions and see that if we are to constrain ourselves to the simulation B shapes, we will need to incorporate some additional step to capture the resolution effects in data. Because the final states across are samples are nearly the same, with similar kinematics, the resolution does not depend much on the exact B mass for each peak. A comparison of the simulated resolution, defined as the difference between the reconstructed mass and the true mass without the unreconstructed neutrals, is shown in



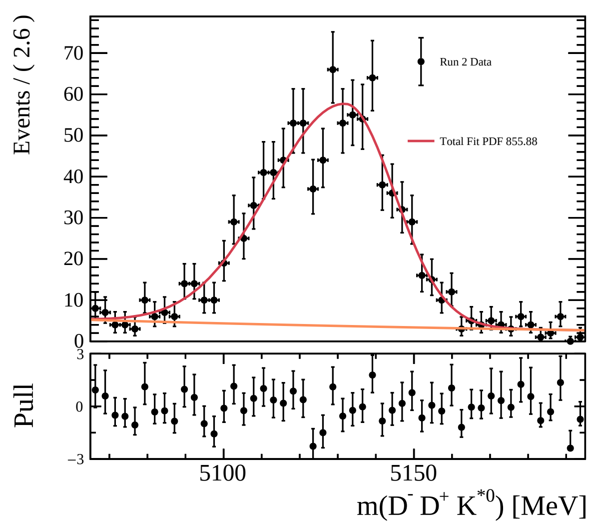
(a) Fit to the decay tree fitter constrained data where the signal is $B^0 \to D^+D^-K^{*0}$. This fit is used in producing sWeights for the reweighting of the MC



(b) Fit to the data without decay tree fitter constraints where the signal is $B^0 \to D^+D^-K^{*0}$. This fit is used in the estimation of our remaining charmless background

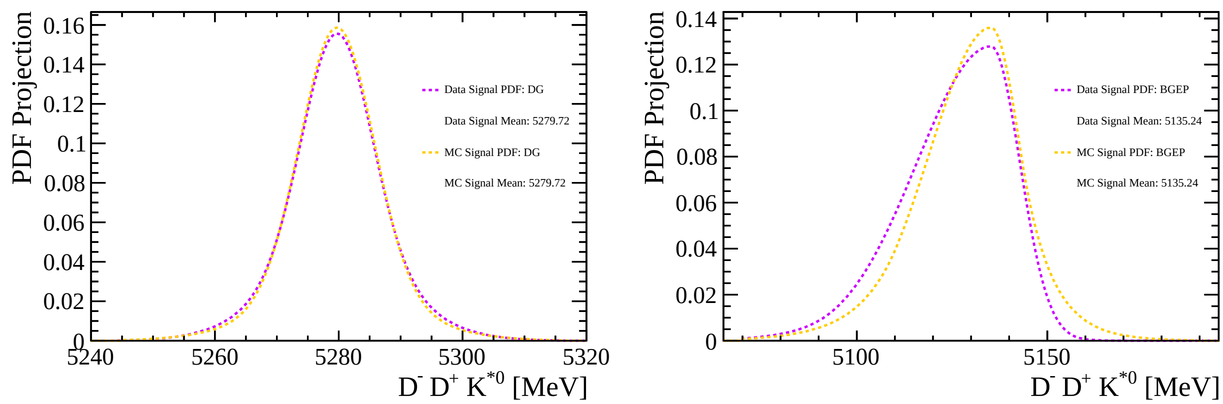


(c) Fit to the decay tree fitter constrained data where the signal is a combination of $B^0 \to D^{*-}D^+K^{*0}$ + $B^0 \to D^-D^{*+}K^{*0}$. This fit is used in producing sWeights for the reweighting of the MC.



(d) Fit to the data without decay tree fitter constraints where the signal is a combination of $B^0 \to D^{*-}D^+K^{*0}$ + $B^0 \to D^-D^{*+}K^{*0}$. This fit is used in the estimation of our remaining charmless background.

Figure 16: Fits to the first two peaks in the $D^-D^+K^{*0}$ spectra both with and without decay tree fitter constraints.

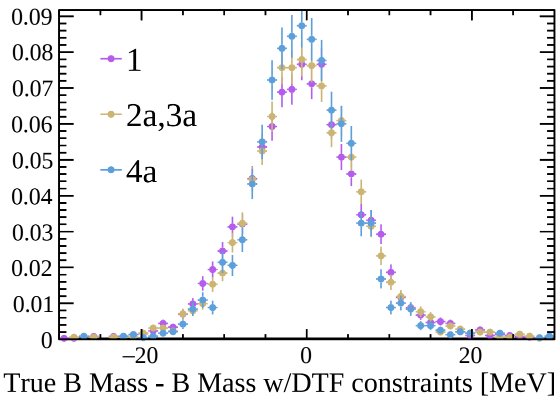


(a) Data Signal pdf from fig. 16a vs MC signal pdf from fig. 15a for peak with no missing particles in $D^-D^+K^{*0}$ spectrum

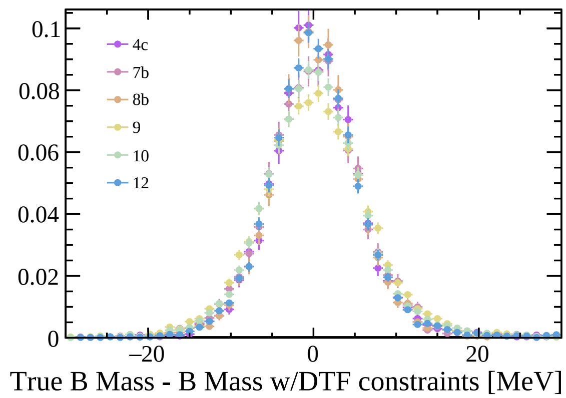
(b) Data Signal pdf from fig. 16c vs MC signal pdf from fig. 15b for peak with 1 missing particles in $D^-D^+K^{*0}$ spectrum

Figure 17: The Signal PDF for both our fits to Data and Simulation for the first two peaks in the $D^-D^+K^{*0}$ spectrum are projected and normalized to unity for comparison. The means are fixed to the same value so we can compare the resolutions. We can see that constraining ourselves to the simulation fits will cause us to miss the true width/resolution in data.

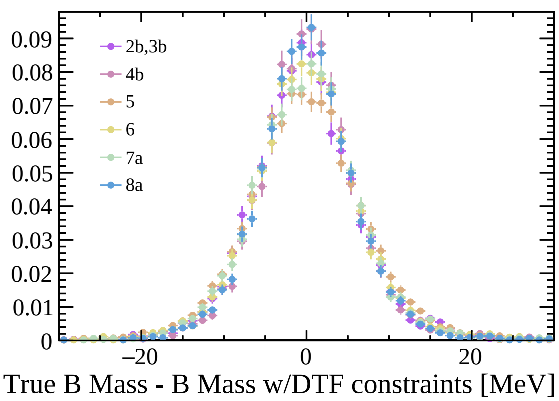
fig. 18. This motivates describing each of the peaks by convoluting the simulated shapes with a single shared smearing function, which we choose as a Gaussian with a mean of 0 and a floating width. We constrain the mean values of the signal shapes to Gaussian's with a mean and width from the MC best fits table 14. We model the background shape as a Bernstein polynomial. The result for an sample implementation of this strategy on the $D^-D^+K^{*0}$ is shown in fig. 19.



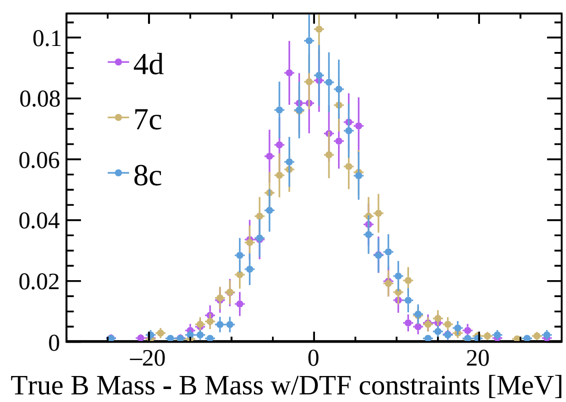
(a) $D^-D^+K^{*0}$ spectrum resolution



(b) $\overline{D}^0D^0K^{*0}$ spectrum resolution



(c) $\overline{D}^0D^+K^{*0}$ spectrum resolution



(d) $\overline{D}^0(D^{*+} \to D^0\pi^+)K^{*0}$ spectrum Spectrum resolution

Figure 18: Resolutions of MC samples shared within data spectra. We omit the $D^-D^0K^{*0}$ spectra for this visual, as the MC there is shared with $\overline{D}^0D^+K^{*0}$. Note that the width across each subplot is similar. Decays with missing particles have resolutions compatible with those that are fully reconstructed.

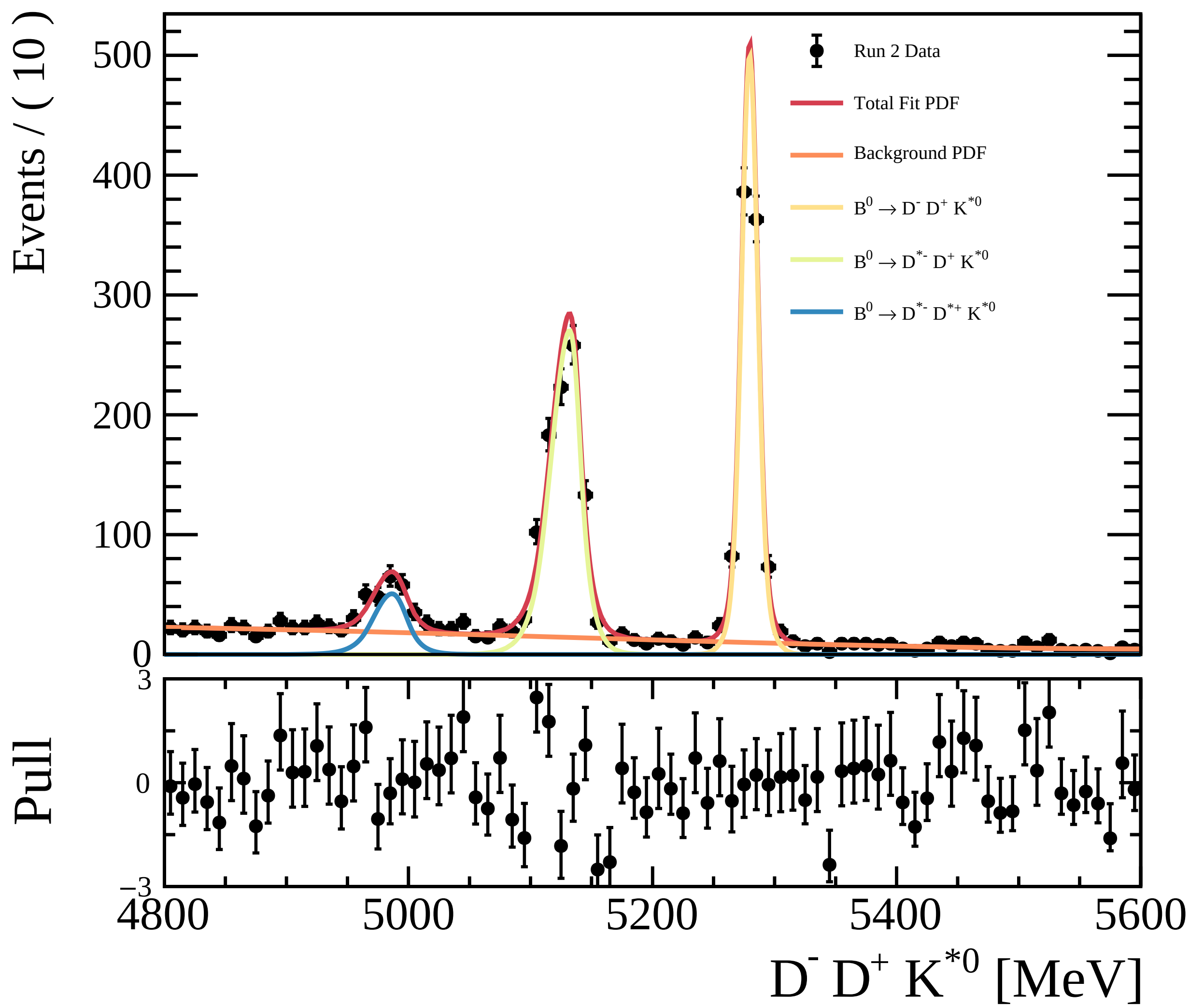


Figure 19: $D^-D^+K^{*0}$ spectrum fit with MC fits convoluted with a Gaussian with a mean of 0 and a floating width. The yields in this fit are allowed to float. The background shape is a Bernstein polynomial and means of the signal shapes are constrained to Gaussian PDFs with means and widths from the nominal MC fit. No separation of the $B^0 \to D^{*-}D^+K^{*0}$ and $B^0 \to D^-D^{*+}K^{*0}$ exists in this fit.

3.8 Systematic Uncertainties

We estimate systematic uncertainties on the final yields and efficiencies coming from the following sources:

- Trigger Efficiency
- Decay mismodeling in Simulation
- Particle Identification Variables (PID)
- Tracking Efficiency

3.8.1 Trigger Efficiency

As a check that simulation of the L0 trigger selections in this analysis are well modeled for this analysis, we implement the data driven TISTOS method in order to compare trigger efficiencies between data and MC for several of the modes present in this analysis. We examine the total efficiencies across our data taking years and L0 trigger condition. The results are summarized in table 15. These efficiencies cannot be compared directly to those quoted in table 27 and table 28, as those MC efficiencies contain trigger selections within the stripping lines.

Based off the differences present between the fully reconstructed simulation and corresponding data regions, we choose to apply a unique and uncorrelated relative systematic to the total efficiency for each value of eq. (10) equal to 10%.

3.8.2 Mis-modeling of Kinematics in Simulation

The B decays in simulation are all produced using a phase space decay model, but we expect that the decays have significant resonant contributions. We estimate a systematic uncertainty on the signal efficiency by weighting the decay kinematics to reproduce what is seen in sWeighted data for each signal peak. For this weight we utilized a $k = 10$ folded gradient

Table 15: Comparison of Trigger Efficiencies calculated using the data driven TISTOS method on peaks in data with no missing particles and the corresponding MC samples. While the absolute TOS efficiencies (e TOS, e TIS) between data and simulation are consistent with each other, some of the relative TIS efficiencies in simulation are larger by nearly 5% in some cases. A closer examination of the relative TOS and TIS efficiencies, shows that the TOS values are consistent within the uncertienties of our Trigger detrmination, while TIS values for Data are nearly always 10% greater.

Source	e TOS	e TIS	Relative e TOS	Relative e TIS
Data for $D^-D^+K^{*0}$	43 ± 5	38.4 ± 2.1	0.95 ± 0.14	0.99 ± 0.06
MC for $B^0 \to D^-D^+K^{*0}$	41.3 ± 0.8	39.5 ± 0.8	0.915 ± 0.022	0.897 ± 0.022
Data for $\overline{D}^0D^0K^{*0}$	47 ± 16	40 ± 5	1.1 ± 0.4	1.12 ± 0.15
MC for $B^0 \to \overline{D}^0D^0K^{*0}$	41.9 ± 0.5	40.2 ± 0.5	0.959 ± 0.017	1.003 ± 0.019
Data for $\overline{D}^0D^+K^{*0}$	38 ± 5	35.8 ± 2.1	0.86 ± 0.15	1.00 ± 0.07
MC for $B^+ \to \overline{D}^0D^+K^{*0}$	42.9 ± 0.6	39.6 ± 0.6	0.980 ± 0.019	0.990 ± 0.020
Data for $\overline{D}^0(D^{*+} \to D^0\pi^+)K^{*0}$	41 ± 12	40 ± 5	0.93 ± 0.28	1.12 ± 0.14
MC for $B^+ \to \overline{D}^0(D^{*+} \to \overline{D}\pi^+)K^{*0}$	49.9 ± 1.7	40.2 ± 1.5	1.14 ± 0.04	1.00 ± 0.04
Data for $\overline{D}^0(D^0 \to K^-\pi^+\pi^+\pi^-)K^+$	44 ± 4	35.7 ± 1.5	-	-
MC for $B^+ \to \overline{D}^0(D^0 \to K^-\pi^+\pi^+\pi^-)K^+$	43.7 ± 0.6	40.0 ± 0.5	-	-
Data for $D^-(D^0 \to K^-\pi^+\pi^+\pi^-)K^+$	45.7 ± 3.4	38.7 ± 1.3	-	-
MC for $B^0 \to D^-(D^0 \to K^-\pi^+\pi^+\pi^-)K^+$	45.1 ± 0.7	44.1 ± 0.7	-	-

boosted decision tree. Its hyper-parameters are optimized to give us a ROCAUC of 0.5 between the reweighted MC and sWeighted Data, meaning that our BDT cannot distinguish between the two samples. The following strategy is implemented:

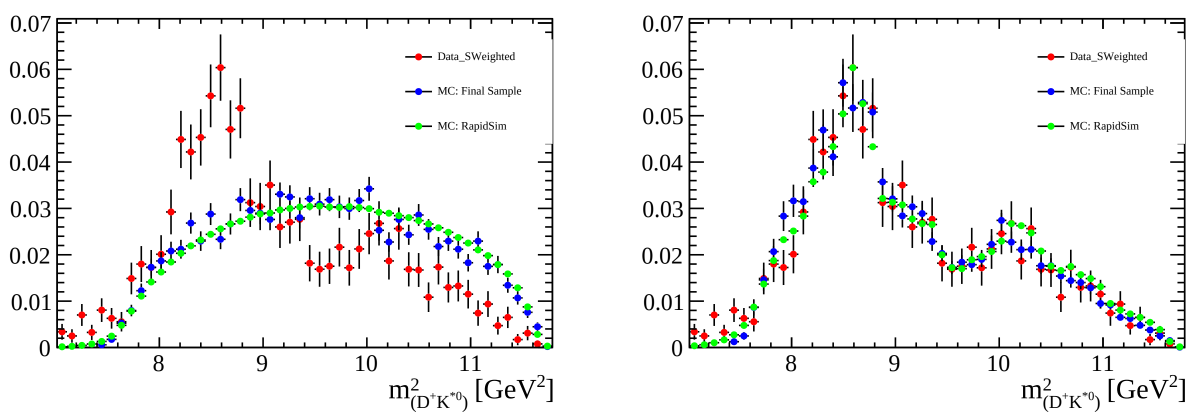
1. Signal weights for each separate signal peak are extracted from the data fits in section 3.7.3 via the SPlots technique.

2. For each peak and the simulation samples that represent the channels contributing to it, a gradient boosted decision tree via the hepml package is trained on the "Dalitz" variables $m_{\overline{D}D}$ and $m_{DK^{*0}}$ of the sWeighted data and the simulation samples. For the normalization modes, the BDT also includes the Dalitz variables representing the four-body D^0 decay which is not modeled correctly in simulation. It is important to note that because of the range of allowed $K\pi$ masses and the missing neutral particles, that these are not true Dalitz variables. However, the effects of the phase space is handled by our use of a machine learning algorithm.

3. We apply this same BDT to a set of generator level MC without applying any cuts, produced via RapidSim, corresponding to the same number of events produced for the relevant MC sample.

4. A weighted efficiency for each MC sample is calculated as a sum over the new weights generated from our BDT for the reconstructed MC divided by the sum over the new weights generated from our BDT for the generator sample.

5. A systematic is chosen based off the relative difference between the reweighted efficiency and the original efficiency.



(a) (b)

Figure 20: The Reweighting of one dalitz variable for the Z spectrum. . Only SWeighted Data is used during the reweighting.

We relegate the plots of this process for the majority of the MC samples to appendix F, and show one example in fig. 20. A summary of the change in efficiencies is in table 16. In general the relative change in our signal mc efficiencies is small when compared to most systematic present in our analysis. We are most sensitive to this effect in our normalization modes, where we not only consider the missed structure in the B Meson decay, but the 4 track D decay. How this systematic is incorporated into the full simultaneous fit can be found in section 3.10. This effect ranges from 1% to 3% across most of our simulation samples.

with the exception of the 8 track normalization mode that requires an additional systematic of 6% on the efficiency.

Table 16: Systematic Uncertainties for MC efficiencies due to Mis-Modeling. We train a Gradient BDT on the sWeighted data and simulation samples for a given peak to capture event by event weights for the simulation. This BDT is also applied to a ReDecay sample that represents a sample with no cuts applied to it. The relative change in the efficiency between unweighted and weighted efficiencies is applied as a systematic.MC samples shared between the extbfP and extbfM spectrum have different systematic uncertainties due to the different sWeighted signal distributions used in the reweighting.

Source	Old Overall Eff	New Overall Eff	Relative Change (Systematic)
$B^0 \to D^- D^+ K^{*0}$	3.11×10^{-3}	3.18×10^{-3}	-1.95×10^{-2}
$B^0 \to (D^{*-} \to D^- \pi^0) D^+ K^{*0}$	2.78×10^{-3}	2.81×10^{-3}	-1.00×10^{-2}
$B^0 \to (D^{*-} \to D^- \pi^0)(D^{*+} \to D^+ \pi^0) K^{*0}$	2.20×10^{-3}	2.19×10^{-3}	3.51×10^{-3}
$B^+ \to \overline{D}^0 D^+ K^{*0}$	6.32×10^{-3}	6.45×10^{-3}	-2.11×10^{-2}
$B^+ \to \overline{D}^{*0} D^+ K^{*0}$	4.93×10^{-3}	4.98×10^{-3}	-9.51×10^{-3}
$B^+ \to \overline{D}^0 (D^0 \to K^- \pi^+ \pi^+ \pi^-) K^+$	6.09×10^{-3}	5.95×10^{-3}	2.32×10^{-2}
$B^0 \to D^- (D^0 \to K^- \pi^+ \pi^+ \pi^-) K^+$	3.81×10^{-3}	3.59×10^{-3}	5.83×10^{-2}
$B^0 \to \overline{D}^0 D^0 K^{*0}$	8.21×10^{-3}	8.15×10^{-3}	8.06×10^{-3}
$B^0 \to \overline{D}^{*0} D^0 K^{*0} + B^0 \to \overline{D}^0 D^{*0} K^{*0}$	8.13×10^{-3}	8.07×10^{-3}	6.42×10^{-3}
$B^0 \to \overline{D}^{*0} D^{*0} K^{*0}$	7.35×10^{-3}	7.26×10^{-3}	1.24×10^{-2}
$B^0 \to (D^{*-} \to \overline{D}^0 \pi^-)(D^{*+} \to \overline{D} \pi^+) K^{*0}$	2.76×10^{-3}	2.73×10^{-3}	1.19×10^{-2}
$B^+ \to \overline{D}^0 (D^{*+} \to \overline{D} \pi^+) K^{*0}$	2.92×10^{-3}	2.88×10^{-3}	1.54×10^{-2}
$B^+ \to \overline{D}^{*0} (D^{*+} \to \overline{D} \pi^+) K^{*0}$	2.88×10^{-3}	2.86×10^{-3}	5.96×10^{-3}
$B^+ \to \overline{D}^0 (D^{*+} \to D^+ \pi^0) K^{*0}$	5.01×10^{-3}	5.08×10^{-3}	-1.30×10^{-2}
$B^+ \to \overline{D}^{*0} (D^{*+} \to D^+ \pi^0) K^{*0}$	5.01×10^{-3}	5.01×10^{-3}	5.54×10^{-4}
$B^0 \to (D^{*-} \to \overline{D}^0 \pi^-)(D^{*+} \to D^+ \pi^0) K^{*0}$	3.99×10^{-3}	3.99×10^{-3}	-1.10×10^{-3}
$B^0 \to (D^{*-} \to \overline{D}^0 \pi^-) D^+ K^{*0}$ for $D^- D^0 K^{*0}$	4.11×10^{-3}	4.13×10^{-3}	-5.69×10^{-3}
$B^0 \to (D^{*-} \to \overline{D}^0 \pi^-)(D^{*+} \to D^+ \pi^0) K^{*0}$ for $D^- D^0 K^{*0}$	4.11×10^{-3}	4.15×10^{-3}	-9.73×10^{-3}
$B^0 \to (D^{*-} \to \overline{D}^0 \pi^-)(D^{*+} \to D^+ \pi^0) K^{*0}$	3.99×10^{-3}	4.01×10^{-3}	-5.04×10^{-3}
$B^0 \to (D^{*-} \to \overline{D}^0 \pi^-)(D^{*+} \to \overline{D} \pi^+) K^{*0}$	6.56×10^{-4}	6.64×10^{-4}	-1.33×10^{-2}
$B^+ \to \overline{D}^0 (D^{*+} \to \overline{D} \pi^+) K^{*0}$	8.15×10^{-4}	8.05×10^{-4}	1.12×10^{-2}
$B^+ \to \overline{D}^{*0} (D^{*+} \to \overline{D} \pi^+) K^{*0}$	7.13×10^{-4}	7.05×10^{-4}	1.13×10^{-2}

3.8.3 PID Variables

It is well-known that the simulation does not perfectly model the response of PID variables. Because our signal and normalization modes use the same particle species with similar kinematics, these effects will mostly cancel in the efficiency ratios. To check our sensitivity to the efficiency difference, the PID distributions for the $B^0 \to D^- D^+ K^{*0}$ MC sample and the $B^0 \to D^-(D^0 \to K^-\pi^+\pi^+\pi^-)K^+$ are transformed using the PIDCorr package [33] to account for discrepancies with data. We summarize the ratio of the final efficiencies before and after PIDCorr is applied, broken down by year and trigger condition both in table 17. At the present level of total systematic uncertainty, we find that the we are not sensitive to the PID correction and no additional systematic is applied.

Table 17: A breakdown of the ratio of Efficiencies for $B^0 \to D^- D^+ K^{*0}$ with and without PIDCorr applied to the appropriate Kaon PID variables

	Year	Ratio TOS	Ratio TIS
Eff Ratio No PIDCorr / With PIDCorr	2016	1.00 ± 0.09	1.00 ± 0.10
	2017	1.00 ± 0.09	1.00 ± 0.11
	2018	1.00 ± 0.10	1.00 ± 0.12

3.8.4 Tracking Efficiency

For our measurement of the branching fractions for, $B^0 \to \overline{D}^0 D^0 K^{*0}$, $B^0 \to \overline{D}^{*0} D^0 K^{*0} + B^0 \to \overline{D}^0 D^{*0} K^{*0}$, and $B^0 \to \overline{D}^{*0} D^{*0} K^{*0}$ we reconstruct only 6 tracks but normalize to $B^+ \to \overline{D}^0(D^0 \to K^-\pi^+\pi^+\pi^-)K^+$, where we reconstruct 7 tracks. As such we choose to apply a relative 2% systematic on the associated normalization coefficient to account for differences in the tracking efficiency. You can see the effect of this on the $B^0 \to \overline{D}^{*0} D^{*0} K^{*0}$ branching fraction in table 18.

Table 18: Sources of uncertainty expressed as a percentage of the normalization factors C for $B^0 \to D^+D^-K^{*0}$ and $B^0 \to \overline{D}^{*0}D^{*0}K^{*0}$ with respect to their normalization modes. We include both systematic and statistical uncertainties from simulation as well as uncertainties from the normalization yields and known branching fractions for comparison.

Sources of Uncertainty (%)	$B^0 \to D^+D^-K^{*0}$	$B^0 \to \overline{D}^{*0}D^{*0}K^{*0}$
D Branching Fractions	2.41	1.87
Normalization B Branching Fraction	10.28	10.61
Normalization Yield TOS	1.80	2.17
Normalization Yield TIS	1.49	1.88
Relative Efficiency TOS (Statistical)	2.34	2.10
Relative Efficiency TIS (Statistical)	1.79	1.54
Trigger Systematic	10.00	10.00
Charmless Background in Normalization	1.29	1.69
Kinematic Mis-Modeling Systematic	6.15	2.63
Tracking Systematic	0.00	2.00
Total Uncertainty	16.28	15.65

3.8.5 Uncertainty Summary

Here we present a compassion of the systematic uncertainties from above with the uncertainties that come our mc statistics, efficiency, and normalization yields. Our largest uncertainty comes from the known B meson branching fractions.

3.9 Corrections and Consistency Checks

3.9.1 Charmless Background Estimation

Decays of the form $B \to DK\pi(\pi)K^{*0}$ - "single charmless" - and $B \to K\pi(\pi)K\pi(\pi)K^{*0}$ - "double charmless" - may remain in our spectra and contribute some nontrivial amount of background to our signal regions, for both the signal and normalization spectra.

We work under the assumption that the charmless events, single and double, peak at the same reconstructed B mass values, without DTF constraints, in both the signal region and sideband regions of our D window cuts, summarized in table section 3.4.6. Because it is the limited phase space of our excited D mesons that produces narrow peaks, we do not expect any charmless contribution to the peak where we are missing two particles. fig. 21 shows the region where only double charmless background would exsists both in the DD plane and the B mass. In addition, we do not expect any charmless contribution in the $\overline{D}^0(D^{*+} \to D^0\pi^+)K^{*0}$ spectrum, due to the soft pion requirement. As such, we cut at 5000 MeV on the B Mass for this section. The 2D sideband regions we examine are visualized in fig. 22.

We estimate the charmless contribution to our signal regions by fitting to the peaks in the B mass region of our D sidebands. However, this fit is complicated by the fact that some fraction of signal B decays will contaminate the sideband, from radiative decays or the possible misconstruction of the intermediate D mesons where tracks are swapped between them. When this happens, a distorted peak shape appears in the reconstructed B mass in MC. These events are labeled as "shifted signal (SS)" and are present in both data and simulation. We estimate the amount of SS by extracting the ratio of sideband events to signal events in simulation. This ratio, times the yield of the signal events in the discrete fits in section 3.7.3, constrains the amount of "SS" in our sideband. We use a non-parametric kernel estimation pdf taken from the sideband region in MC to represent the SS shape in data; plots of the simulation with these shapes can be found in appendix E.

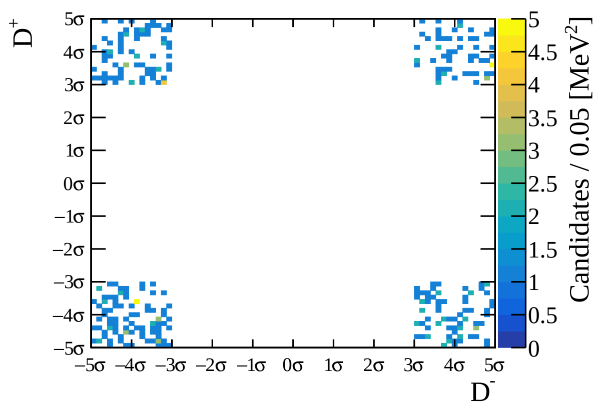
The PDF for the charmless background is pulled from our discrete fits to the B meson

Table 19: Summary of Estimation for remaining charmless background in signal spectra. % for Peak 0, 1 is the percentage of the sWeighted Data that the estimated charmless yield is.

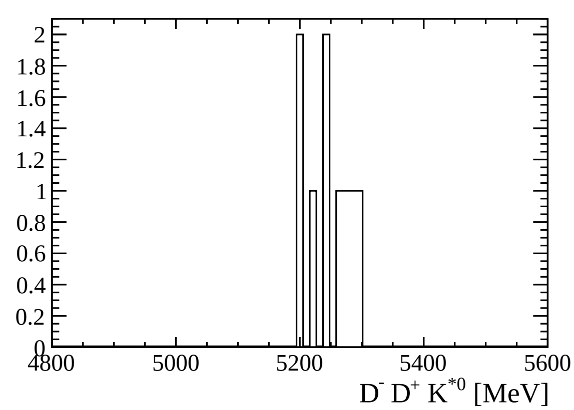
Spectrum	Charmless Yield Peak 0	% for Peak 0	Charmless Yield Peak 1	% for Peak 1
$D^- D^+ K^{*0}$	55 ± 12	6.1 ± 1.4	28 ± 15	3.3 ± 1.8
$\overline{D}^0 D^0 K^{*0}$	19 ± 8	10 ± 4	13 ± 11	0.8 ± 0.7
$\overline{D}^0 D^+ K^{*0}$	51 ± 11	5.0 ± 1.1	83 ± 18	2.9 ± 0.6
$D^- D^0 K^{*0}$	–	–	10 ± 10	1.3 ± 1.3
$\overline{D}^0 (D^0 \to K^- \pi^+ \pi^+ \pi^-) K^+$	309 ± 29	14.5 ± 1.4	–	–
$D^- (D^0 \to K^- \pi^+ \pi^+ \pi^-) K^+$	306 ± 31	10.7 ± 1.1	–	–

masses with no DTF constraints, which are shown in appendix D. We allow the parameters of these PDFs to float within their uncertainties when fitting the sidebands. Our final fit to the sideband is thus a sum of the two PDFs that represent our charmless background, with yields allowed to freely float and the SS PDFs extracted from the fits to MC whose yields are constrained to the ratios extracted earlier and a final polynomial PDF to handle combinatorial background.

The charmless yields in the sidebands are extrapolated, based on the area in the D mass plane, to estimate their contribution in the signal region for each peak. The results of this estimation are given in table 19. In most cases the background contamination is a few percent of the total peak yield in the signal region. This estimation is applied as a correction to the signal yield in the final fit, with a corresponding systematic uncertainty on its size.

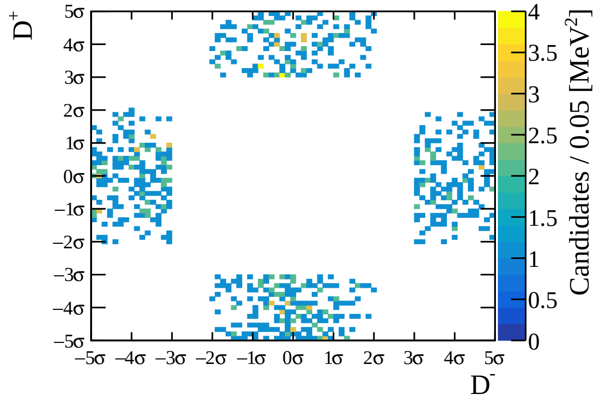


(a) Double Charmless Sideband Region for $D^-D^+K^{*0}$

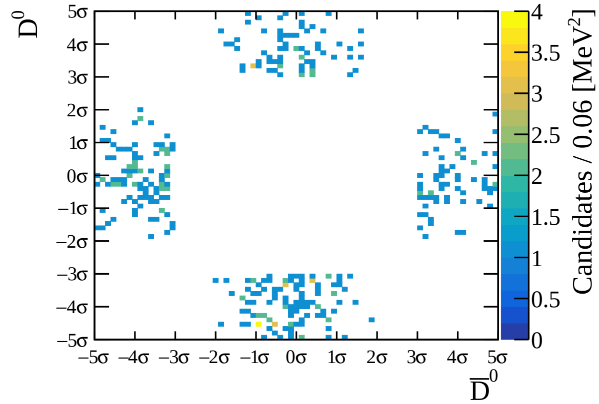


(b) B invariant Mass with only candidates that exist in both D sidebands.

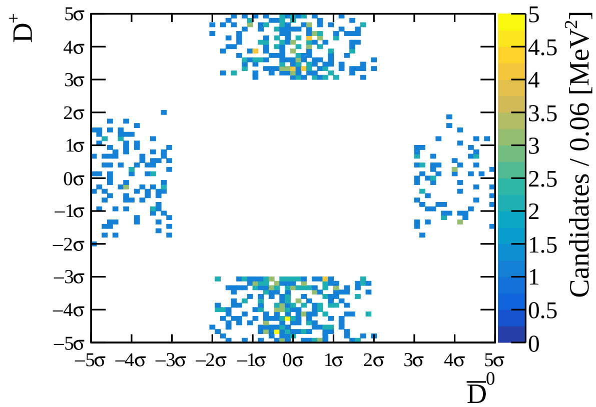
Figure 21: The double charmless sideband region in the D Mass windows for the $D^-D^+K^{*0}$ spectrum alongside the corresponding B mass. Candidates in this region are negligible in our charmless background estimation. As such we do not consider charmless background contribution to our signal peaks with two missing particles.



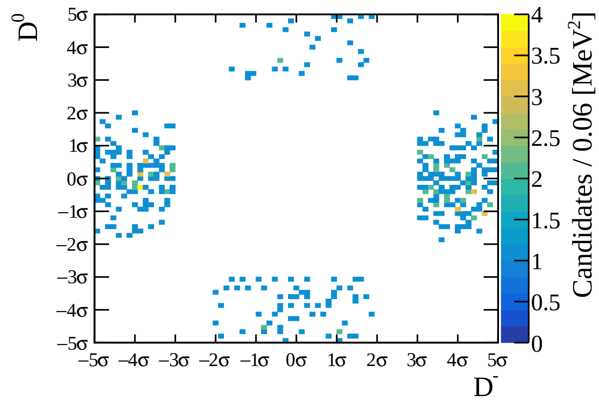
(a) Sideband Region for $D^-D^+K^{*0}$



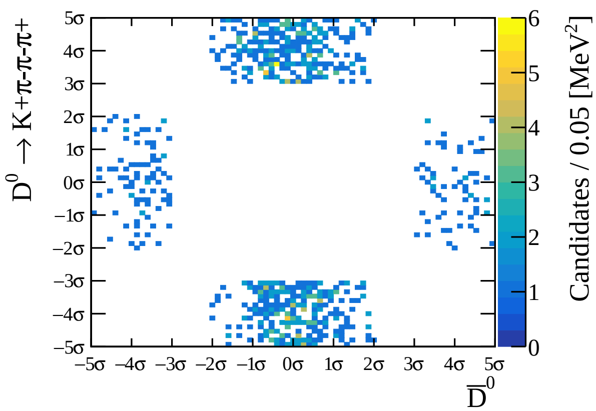
(b) Sideband Region for $\overline{D}^0D^0K^{*0}$



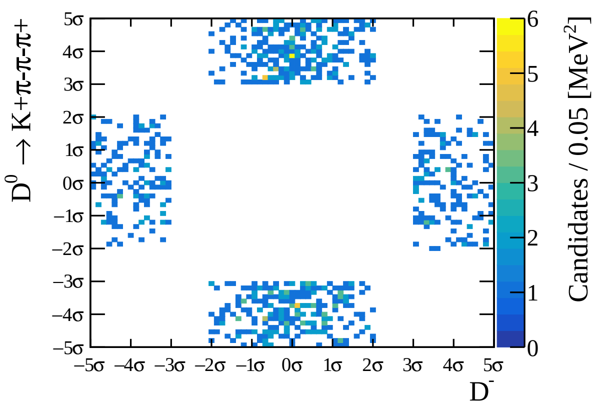
(c) Sideband Region for $\overline{D}^0D^+K^{*0}$



(d) Sideband Region for $D^-D^0K^{*0}$

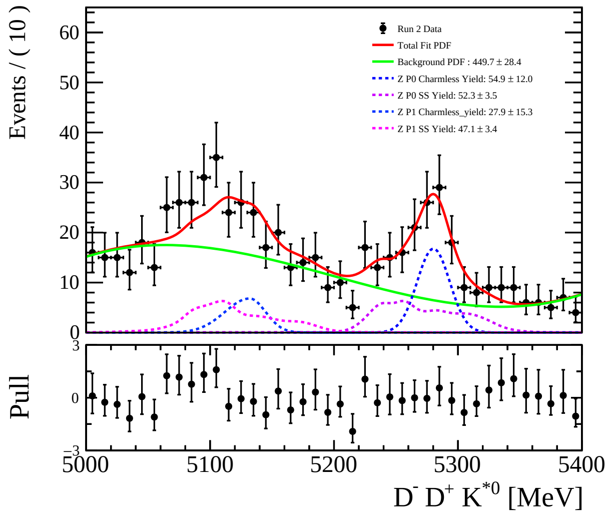


(e) Sideband Region for $\overline{D}^0(D^0 \to K^-\pi^+\pi^+\pi^-)K^+$

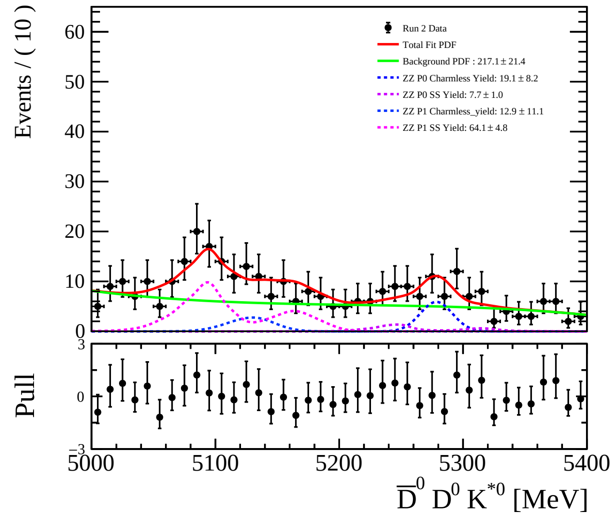


(f) Sideband Region for $D^-(D^0 \to K^-\pi^+\pi^+\pi^-)K^+$

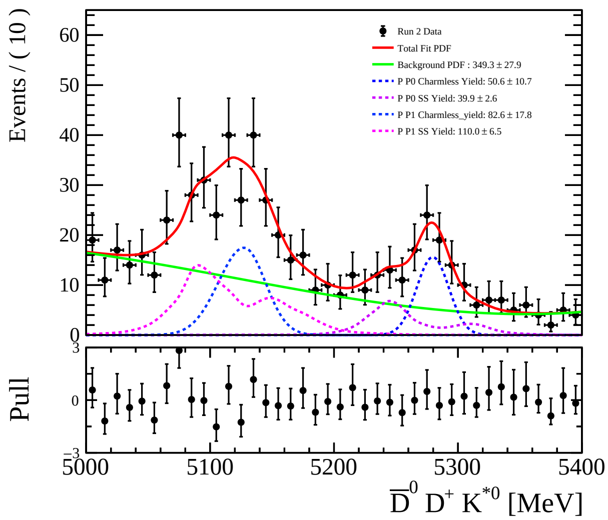
Figure 22: The sideband region in the D Mass windows for each signal and normalization spectra. We separate sideband from signal by 1 σ in order to minimize signal contamination in the sideband during the estimation of charmless background.



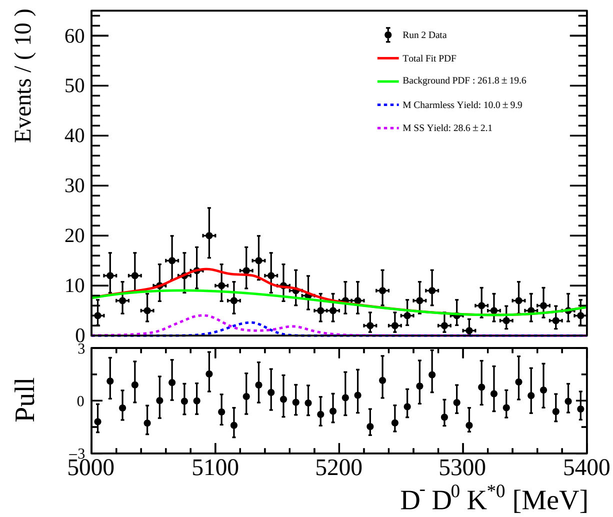
(a) Sideband Region for $D^-D^+K^{*0}$



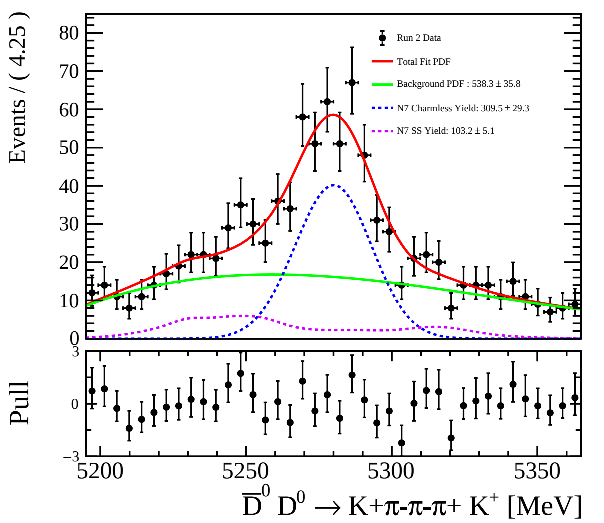
(b) Sideband Region for $\overline{D}^0D^0K^{*0}$



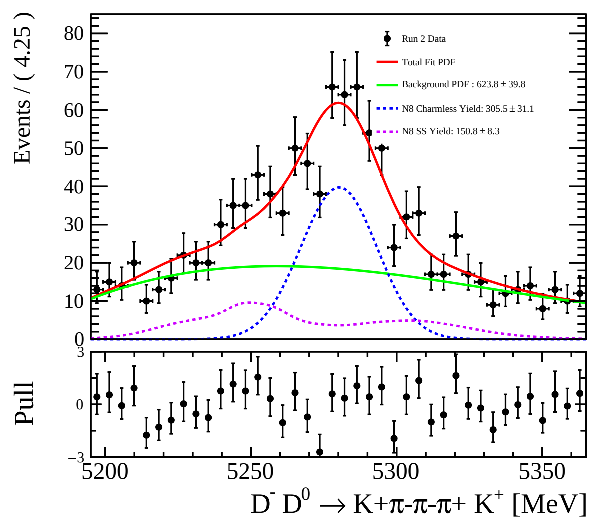
(c) Sideband Region for $\overline{D}^0D^+K^{*0}$



(d) Sideband Region for $D^-D^0K^{*0}$



(e) Sideband Region for $\overline{D}^0(D^0 \to K^-\pi^+\pi^+\pi^-)K^+$



(f) Sideband Region for $D^-(D^0 \to K^-\pi^+\pi^+\pi^-)K^+$

Figure 23: Fits to our D Sideband Regions capturing the expected amount of charmless background that remains in our peaking signal regions for the fully reconstructed modes and modes missing one particles

3.9.2 Validation of Normalization Measurements

As a check, we compare the nominal value of the ratio of the normalization mode branching fractions, to our efficiency corrected yields across data taking years and L0 trigger conditions.

Our nominal values for a given year and trigger condition are calculated as

$$\mathcal{R}_{year,trigger} = \frac{\epsilon(\overline{D}^0(D^0 \to K^-\pi^+\pi^+\pi^-)K^+) \times \mathcal{N}(D^-(D^0 \to K^-\pi^+\pi^+\pi^-)K^+) \times B(D^0 \to K^+\pi^-)}{\epsilon(D^-(D^0 \to K^-\pi^+\pi^+\pi^-)K^+) \times \mathcal{N}(\overline{D}^0(D^0 \to K^-\pi^+\pi^+\pi^-)K^+) \times B(D^+ \to K^-\pi^+\pi^+)} \tag{12}$$

where the yields come from table 11 and are corrected for the charmless background contamination from table 19, the efficiencies come from table 28, and the D branching fractions from table 12. We preform this calculation across the data taking years and L0 trigger conditions, as well as across the full range in table 20.

Table 20: Ratio of the Branching Fractions $\frac{B(B \to \overline{D}^0 D^0 K^+)}{B(B \to D^- D^0 K^+)}$

Source	Values
2016 TOS	1.26 ± 0.13
2016 TIS	1.29 ± 0.15
2017 TOS	1.16 ± 0.12
2017 TIS	1.26 ± 0.15
2018 TOS	1.19 ± 0.12
2018 TIS	1.47 ± 0.18
Total	1.27 ± 0.07
PDG	1.22 ± 0.18

3.10 The Simultaneous Fit and Results

We now construct the single likelihood function that will be used to fit all the data spectra simultaneously. For each data spectrum we constrain ourselves to the relevant signal PDF's from table 14, preforming a convolution with a Gaussian smearing term:

$$f_{SIG,i}(m) = f_{MC,i}(m) \circledast \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{m}{\sigma}\right)^2\right) \tag{13}$$

where σ is allowed to float, and is shared between all $f_i(x)$ across each data spectra, with the exception of the ST spectra, which is expected to have a different resolution term given the choice to reconstruct the soft pion.

We model the combinatorial backgrounds $\overline{D}^0(D^{*+} \to D^0\pi^+)K^{*0}$ spectra as an exponential PDF in the nominal fit:

$$f_{BKG}(m) \propto e^{-\lambda m} \tag{14}$$

We model the combinatorial backgrounds for the remaining spectra as 3rd order Bernstein Polynomial PDFs in the nominal fit:

$$f_{BKG}(m) \propto \mathcal{N} \cdot \sum_{i=0}^{3} c_i \cdot B_{i,3}(m). \tag{15}$$

where

$$B_{i,3}(m) = \left(3i\right) x^i \cdot (1-m)^{3-i} \tag{16}$$

Our final PDF for each data spectrum is then:

$$f(m) = \sum_i [N_{SIG,i} f_{SIG,i}(m)] + N_B f_{BKG}(m), \tag{17}$$

where the sum goes over all signal contributions to that spectrum. Note that there may be multiple signal PDFs contributing to a single visible peak with different shapes; for instance in a case where a D^0 could come from either D^{*+} or D^{*0}. The signal yields $N_{SIG,i}$ are determined from the normalization constants C_i derived from the efficiencies, branching fractions, and the normalization yields and corrected for the estimated charmless background contribution d_i:

$$N_{SIG,i} = C_i \mathcal{B}_i - d_i. \tag{18}$$

Because the charmless corrections are derived on a peak-by-peak basis, we need to make an ansatz about how to "distribute" the correction across overlapping signals. We choose to assign a fraction of the charmless correction to each signal based on its fraction of the total yield in that peak.

The normalization constants among the modes are correlated through their systematic uncertainties, which are included in the fit as floating nuisance parameters. We diagonalize the full 18×18 covariance matrix for the C_i to determine the eigenvalues σ_j^2 and eigenvectors $\vec{V}_j$. The charmless corrections are also included in this manner, with nuisance parameters θ_i. The full yields, as a function of the nuisance parameters ν_j are thus given by

$$N_{SIG,i} = \left[C_{i,0} + \sum_j \sigma_j \nu_j V_{j,i} \right] \mathcal{B}_i - d_i + \sigma_{d,i} \theta_i. \tag{19}$$

The nominal expected yield for a given branching fraction occurs when each nuisance parameter's value is zero. In the fit, an additional Gaussian constraint PDF with a mean of zero and width of one is multiplied on the full PDF for each parameter ν_j, θ_i.

We report the eigenvalues of our diagonalized covariance matrix in fig. 24. Due to the differences between D^0 and D^+ branching fractions as well as the effect of the D^* branching fractions, our eigenvalues exhibit a hierarchy. We also report a 11x11 submatrix of the final covariance and correlation matrices, only examining the final branching fractions that we are interested in.

$$\begin{bmatrix} 1.35439 \times 10^{11} \\ 1.25205 \times 10^{11} \\ 2.01383 \times 10^{10} \\ 1.58867 \times 10^{10} \\ 1.39725 \times 10^{10} \\ 1.14482 \times 10^{10} \\ 8.08353 \times 10^{09} \\ 7.44445 \times 10^{09} \\ 4.51848 \times 10^{09} \\ 3.61854 \times 10^{09} \\ 2.02445 \times 10^{09} \\ 1.55037 \times 10^{09} \\ 1.45619 \times 10^{09} \\ 5.49669 \times 10^{08} \\ 4.69989 \times 10^{08} \\ 3.29222 \times 10^{08} \\ 2.30169 \times 10^{08} \\ 8.96968 \times 10^{07} \end{bmatrix}$$

Figure 24: Eigenvalues extracted from covariance matrix for the full simultaneous fit. Each term can be thought of as one part of the magnitude by which each yield depends on the nuisance parameters

The results of the simultaneous fit are shown in fig. 25 to fig. 29. The branching fractions for each signal mode extracted from the fit are given in table 21. A submatrix of the full covariance matrix only relating the branching fraction parameters, and a visualization of the corresponding correlation among the modes are given in section 3.10.1 and section 3.10.2 respectively. Because the systematic uncertainties are included in the fit already, the quoted uncertainties give the total statistical plus systematic uncertainty. To estimate their separation, we repeat the fit with all nuisance parameters fixed to zero; the result is given in appendix H.

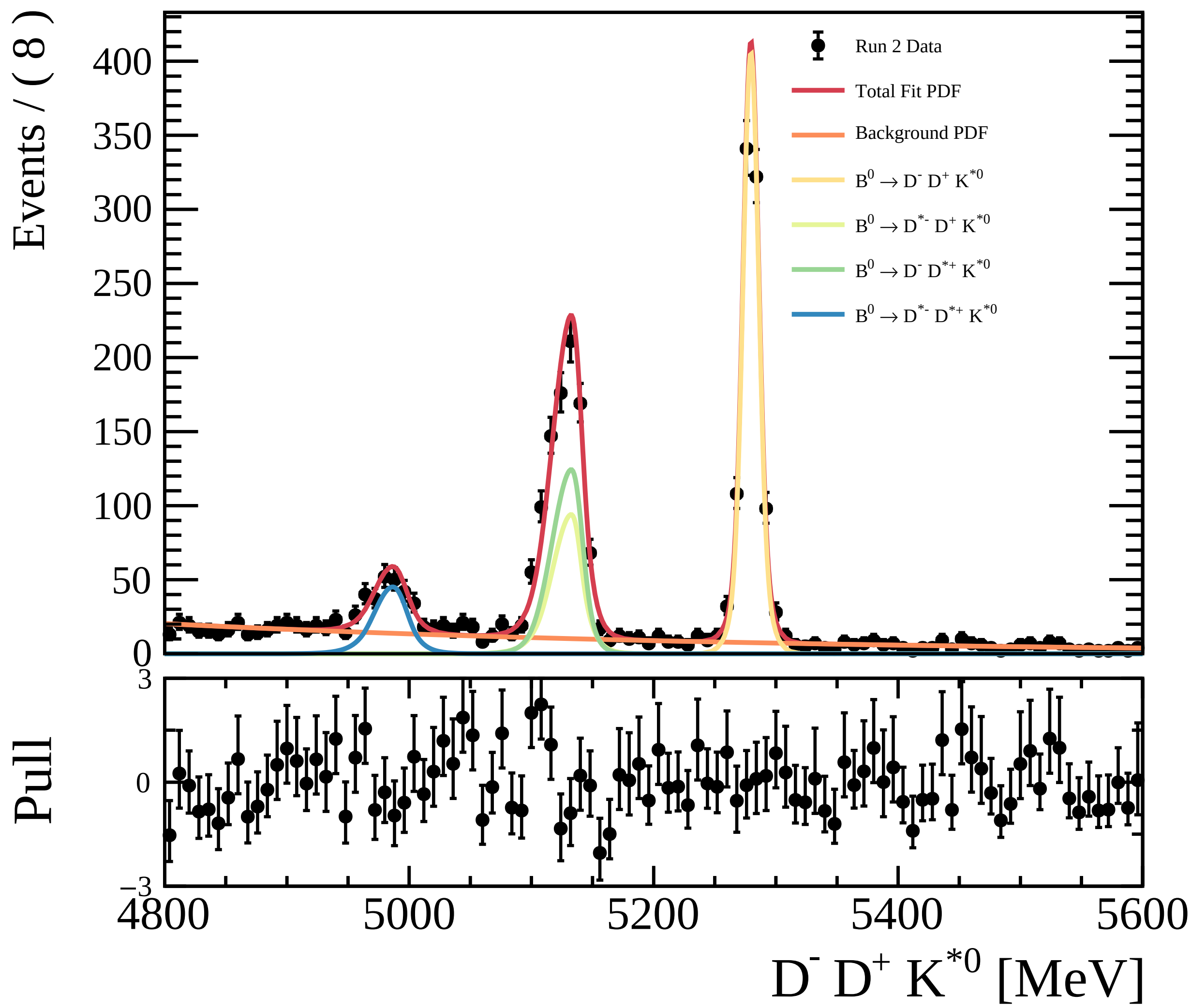


Figure 25: The final simultaneous fit to the $D^{-}D^{+}K^{*0}$ spectrum, incorporating all known uncertainties and correlations

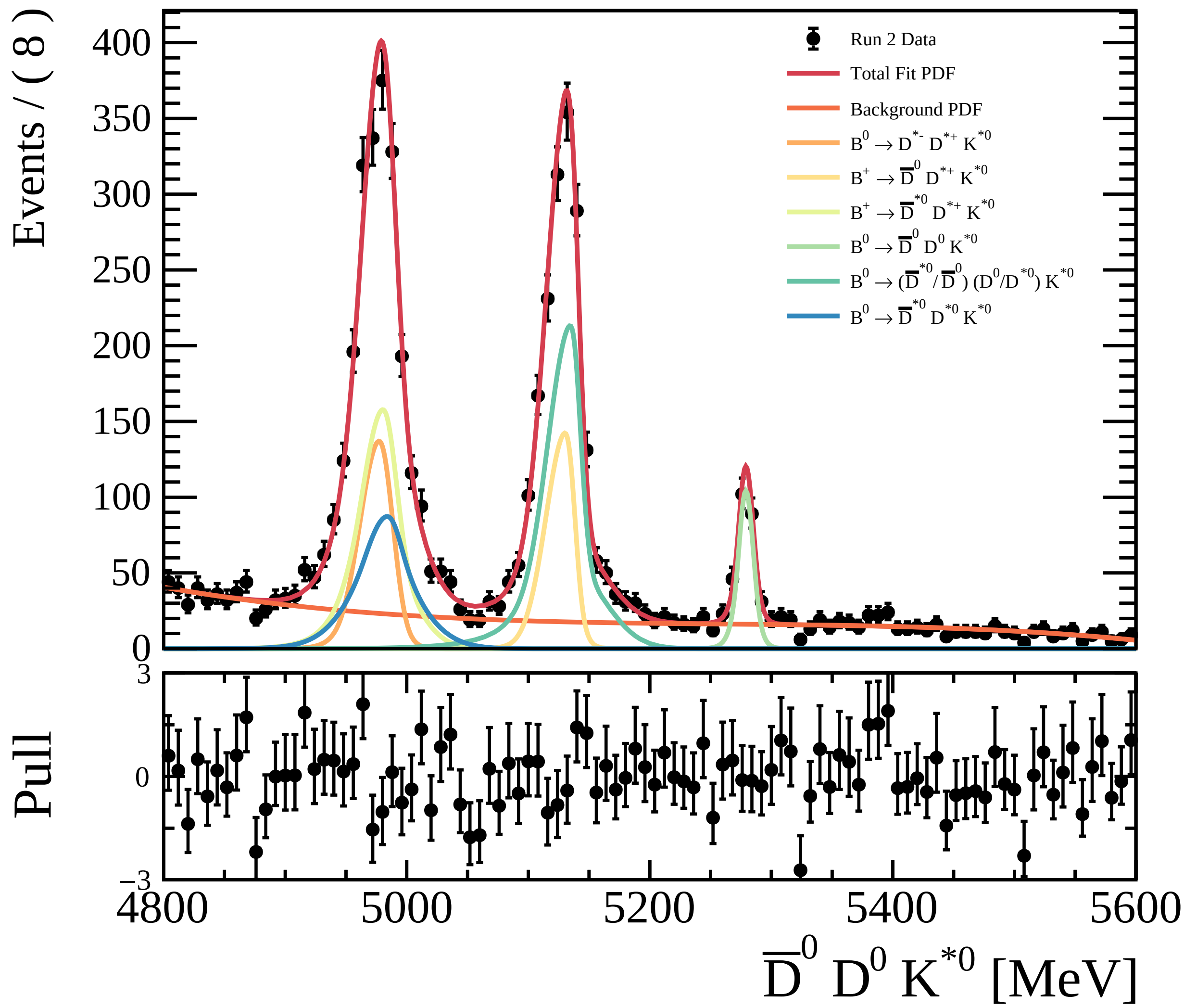


Figure 26: The final simultaneous fit to the $\overline{D}^0 D^0 K^{*0}$ spectrum, incorporating all known uncertainties and correlations.

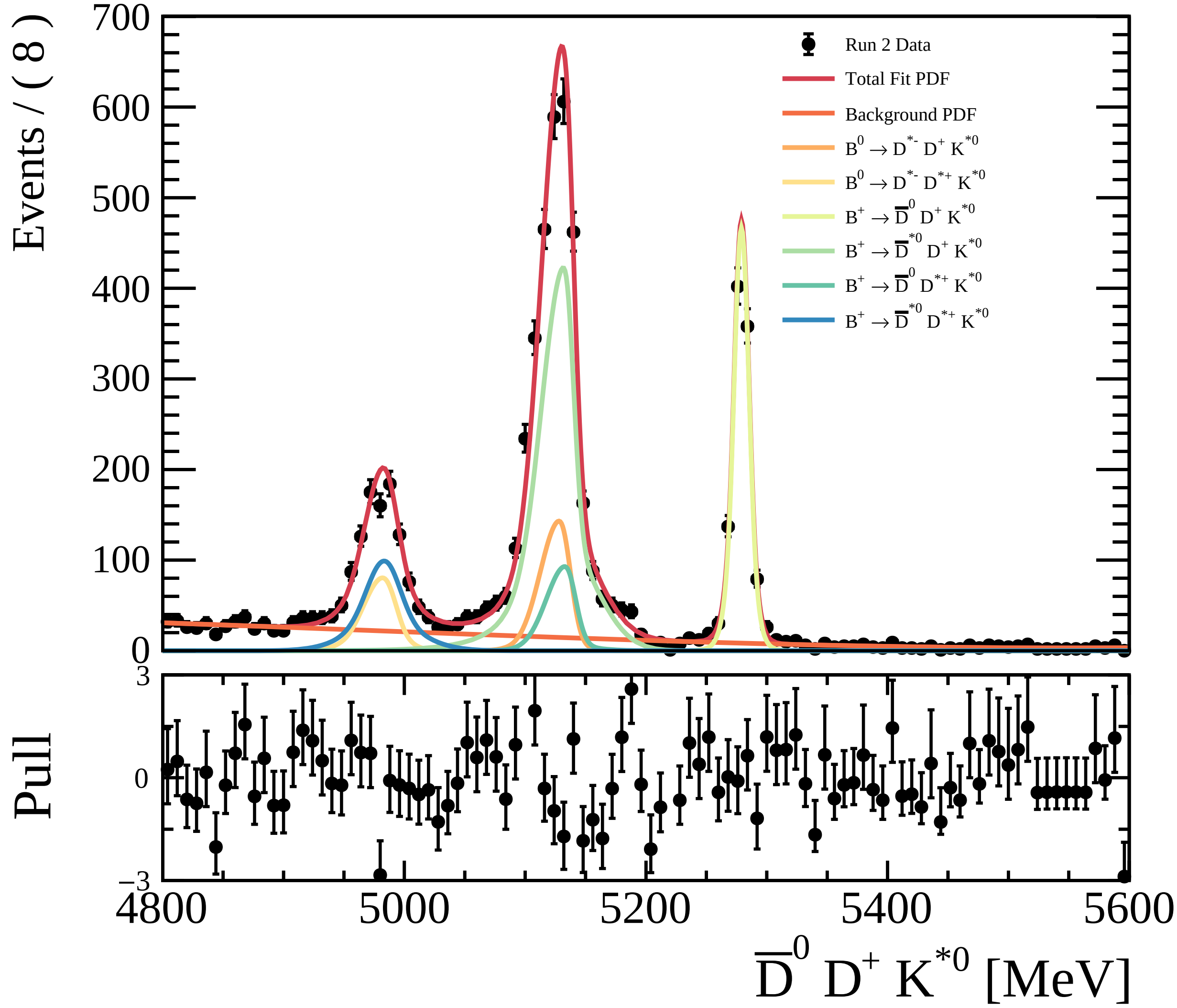


Figure 27: The final simultaneous fit to the $\overline{D}^0 D^+ K^{*0}$ spectrum, incorporating all known uncertainties and correlations.

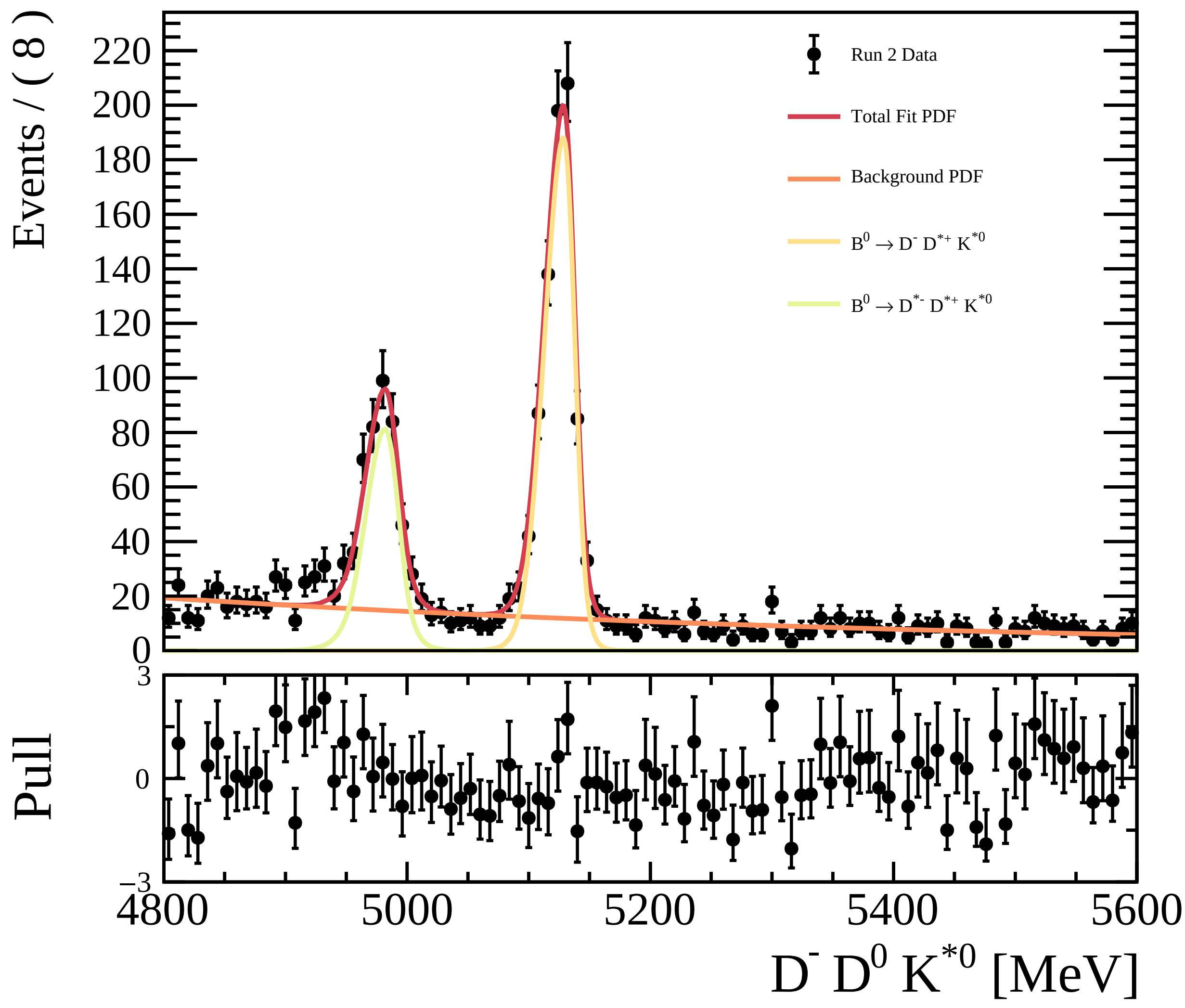


Figure 28: The final simultaneous fit to the $D^-D^0K^{*0}$ spectrum, incorporating all known uncertainties and correlations.

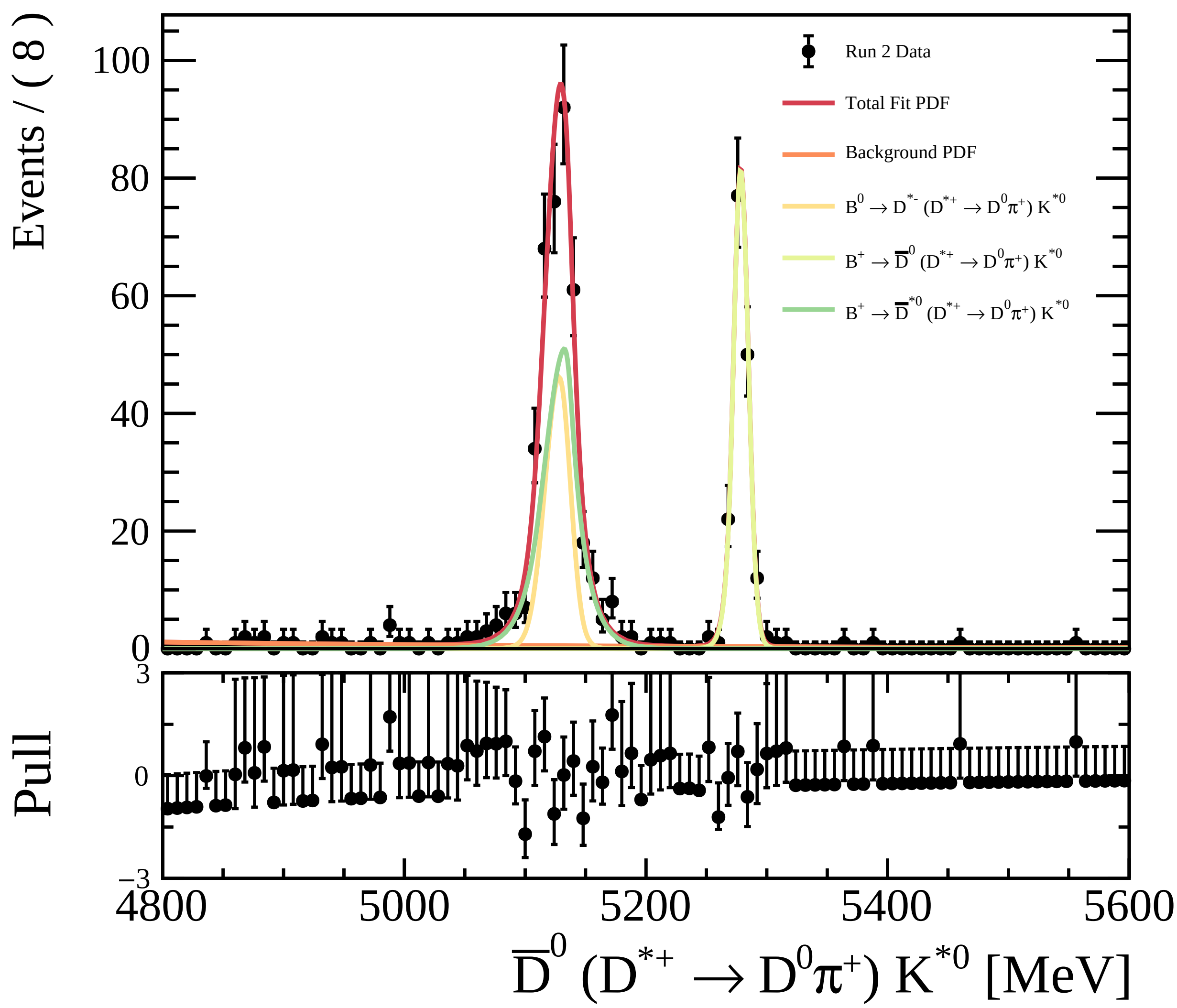


Figure 29: The final simultaneous fit to the $\overline{D}^0(D^{*+} \to D^0\pi^+)K^{*0}$ spectrum, incorporating all known uncertainties and correlations.

Table 21: Final Branching Fractions

Label	Decay	Branching Fraction (Statistical + Systematic)	Branching Fraction (Statistical)
1	$B^0 \to D^- D^+ K^{*0}$	$(4.26 \pm 0.56) \times 10^{-4}$	$(3.75 \pm 0.14) \times 10^{-4}$
2	$B^0 \to D^{*-} D^+ K^{*0}$	$(6.74 \pm 1.52) \times 10^{-4}$	$(5.23 \pm 0.60) \times 10^{-4}$
3	$B^0 \to D^- D^{*+} K^{*0}$	$(8.88 \pm 0.93) \times 10^{-4}$	$(9.06 \pm 0.35) \times 10^{-4}$
4	$B^0 \to D^{*-} D^{*+} K^{*0}$	$(1.59 \pm 0.17) \times 10^{-3}$	$(1.58 \pm 0.08) \times 10^{-3}$
5	$B^+ \to \overline{D}^0 D^+ K^{*0}$	$(6.35 \pm 0.68) \times 10^{-4}$	$(6.75 \pm 0.23) \times 10^{-4}$
6	$B^+ \to \overline{D}^{*0} D^+ K^{*0}$	$(1.75 \pm 0.23) \times 10^{-3}$	$(1.97 \pm 0.05) \times 10^{-3}$
7	$B^+ \to \overline{D}^0 D^{*+} K^{*0}$	$(8.84 \pm 1.07) \times 10^{-4}$	$(9.46 \pm 0.67) \times 10^{-4}$
8	$B^+ \to \overline{D}^{*0} D^{*+} K^{*0}$	$(1.52 \pm 0.18) \times 10^{-3}$	$(1.72 \pm 0.10) \times 10^{-3}$
9	$B^0 \to \overline{D}^0 D^0 K^{*0}$	$(1.58 \pm 0.22) \times 10^{-4}$	$(1.69 \pm 0.15) \times 10^{-4}$
10	$B^0 \to \overline{D}^{*0} D^0 K^{*0} B^0 \to \overline{D}^0 D^{*0} K^{*0}$	$(9.90 \pm 1.14) \times 10^{-4}$	$(1.06 \pm 0.06) \times 10^{-3}$
11	$B^0 \to \overline{D}^{*0} \overline{D}^* K^{*0}$	$(5.95 \pm 1.03) \times 10^{-4}$	$(6.07 \pm 0.74) \times 10^{-4}$

3.10.1 Covariance Matrix

Table 22: Covariance Matrix of the 11 Branching Fractions present in our simultaenous fit as floating parameters

	BF 1	BF 2	BF 3	BF 4	BF 5	BF 6	BF 7	BF 8	BF 9	BF 10	BF 11
BF 1	0.32	0.38	0.22	0.45	0.15	0.26	0.18	0.28	0.03	0.22	0.13
BF 2	0.38	2.31	0.14	0.74	0.16	-0.94	0.11	0.15	0.04	0.30	0.16
BF 3	0.22	0.14	0.86	1.13	0.47	1.58	0.64	1.16	0.11	0.71	0.40
BF 4	0.45	0.74	1.13	2.95	0.81	2.33	1.10	1.45	0.20	1.23	0.67
BF 5	0.15	0.16	0.47	0.81	0.46	1.11	0.47	0.87	0.08	0.53	0.30
BF 6	0.26	-0.94	1.58	2.33	1.11	5.10	1.38	2.91	0.26	1.70	0.90
BF 7	0.18	0.11	0.64	1.10	0.47	1.38	1.14	1.18	0.11	0.46	0.40
BF 8	0.28	0.15	1.16	1.45	0.87	2.91	1.18	3.32	0.21	1.31	0.31
BF 9	0.03	0.04	0.11	0.20	0.08	0.26	0.11	0.21	0.05	0.14	0.08
BF 10	0.22	0.30	0.71	1.23	0.53	1.70	0.46	1.31	0.14	1.29	0.56
BF 11	0.13	0.16	0.40	0.67	0.30	0.90	0.40	0.31	0.08	0.56	1.06

3.10.2 Correlation Matrix

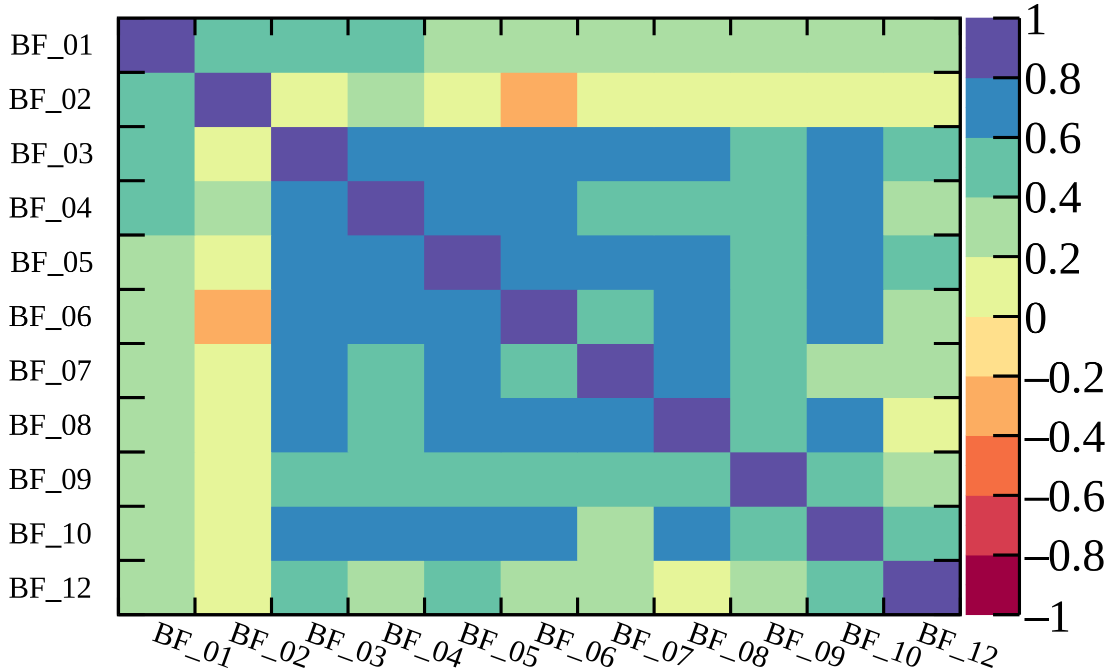
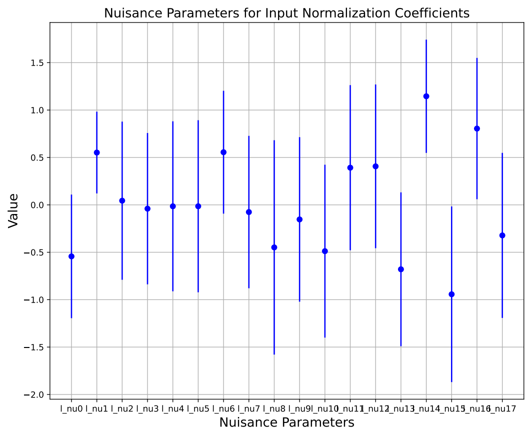


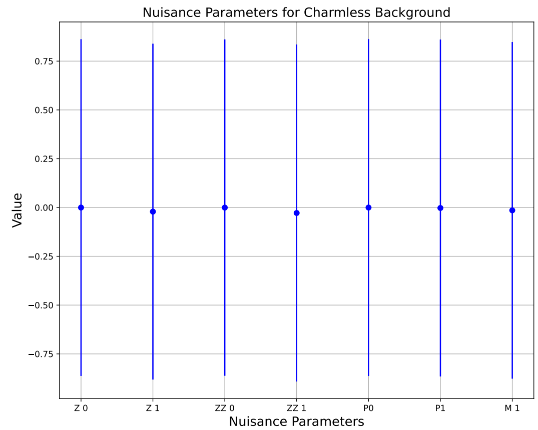
Figure 30: The correlation matrix of our 11 branching fraction parameters. The largest correlations, other then the diagonal, come from the sharing of the normalization mode

3.10.3 Nuisance Parameters



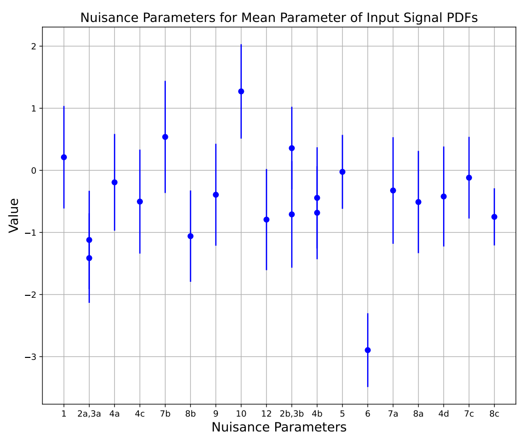
(a) The nuisance parameters of the 18 normalization coefficients. The labeling follows the order given in table 13

.



(b) The nuisance parameters of the charmless backgrounds in the relevant peaks

.



(c) The nuisance parameters of the means from the input MC signal PDFs. The labeling follows the order given in table 13

.

3.10.4 Comparison with Previous Measurement

As another check, we scale the previously measured branching fraction of $B^0 \to \overline{D}^0 D^0 K^{*0}$, $(3.5 \pm 0.5) \times 10^{-4}$ [20], by the integral of their background subtracted mass of their $K^+\pi^-$ invariant mass within our K^{*0} window by the total integral of their their $K^+\pi^-$ region. The scale factor is 0.32 ± 0.05, and the scaled value is $(1.12 \pm 0.24) \times 10^{-4}$. Our value of $(1.58 \pm 0.22) \times 10^{-4}$ is greater then one σ of the scaled value. We are currently investigating whether the difference in normalization modes between the two analysis or other possible systematic are responsible for the large difference between values.

4 Conclusion

The need for precision measurements of b-physics to study extensions to the standard model requires initial analysis of decays of the form $B \to \overline{D}^{(*)-,0} D^{(*)+,0} K^+ \pi^-$. Using the LHCb run 2 data set (2016-2018) we measure 11 different branching fractions of this form. These are the first measurements of 10 of these branching fractions, with the exception of $B^0 \to \overline{D}^0 D^0 K^+ \pi^-$, with a comparison of the two measurements in section 3.10.4. These branching fractions will help set benchmarks for future measurements within and outside the K^{*0} window, and aid in amplitude studies of the various resonances that contribute to these final states.

Appendices

A All MC Efficiencies

The breakdown of the event efficiencies for each selection on signal and normalization mode MC is reported in table 23 through table 26. The quoted uncertainties are statistical. These tables do not include the correction to the uncertainty for our use of ReDecay.The final efficiencies broken up by both year and disjoint L0 trigger conditions are reported in table 27 and table 28. These tables do include the correction to the uncertainty for our use of ReDecay.

Table 23: Summary of Event Efficiencies

Source	Year	Number Accepted	$\epsilon_{Generator}$	$\epsilon_{stripping}$	$\epsilon_{offline}$	$\epsilon_{trigger}$
$B^0 \to D^- D^+ K^{*0}$	2016	655750	5.34 ± 0.09	0.463 ± 0.008	94.1 ± 0.4	94.9 ± 0.4
	2017	603999	5.16 ± 0.09	0.524 ± 0.009	94.5 ± 0.4	94.1 ± 0.4
	2018	689999	5.24 ± 0.09	0.422 ± 0.008	94.3 ± 0.4	94.5 ± 0.5
$B^0 \to (D^{*-} \to D^- \pi^0) D^+ K^{*0}$	2016	643848	5.12 ± 0.09	0.439 ± 0.008	94.0 ± 0.5	92.4 ± 0.5
	2017	715999	5.11 ± 0.09	0.439 ± 0.008	93.2 ± 0.5	92.0 ± 0.5
	2018	698000	5.13 ± 0.08	0.393 ± 0.007	93.8 ± 0.5	91.1 ± 0.6
$B^0 \to (D^{*-} \to D^- \pi^0)(D^{*+} \to D^+ \pi^0) K^{*0}$	2016	631999	4.93 ± 0.09	0.354 ± 0.007	93.8 ± 0.5	89.1 ± 0.7
	2017	609799	5.08 ± 0.09	0.354 ± 0.008	93.1 ± 0.6	89.8 ± 0.7
	2018	619599	5.09 ± 0.09	0.341 ± 0.007	94.1 ± 0.5	90.2 ± 0.7
$B^0 \to (D^{*-} \to \overline{D}^0 \pi^-)(D^{*+} \to D^+ \pi^0) K^{*0}$	2016	615799	6.72 ± 0.12	0.588 ± 0.010	94.1 ± 0.4	95.4 ± 0.4
	2017	607999	6.75 ± 0.12	0.620 ± 0.010	93.8 ± 0.4	95.1 ± 0.4
	2018	608000	6.86 ± 0.12	0.544 ± 0.009	93.6 ± 0.5	94.0 ± 0.5
$B^0 \to (D^{*-} \to D^- \pi^0)(D^{*+} \to \overline{D} \pi^+) K^{*0}$	2016	308518	9.30 ± 0.25	0.778 ± 0.016	92.9 ± 0.6	96.8 ± 0.4
	2017	313901	9.29 ± 0.22	0.938 ± 0.017	94.0 ± 0.5	96.1 ± 0.4
	2018	317305	9.35 ± 0.22	0.754 ± 0.015	92.9 ± 0.6	96.9 ± 0.4
$B^0 \to (D^{*-} \to \overline{D}^0 \pi^-)(D^{*+} \to \overline{D} \pi^+) K^{*0}$	2016	308518	9.30 ± 0.25	0.182 ± 0.008	94.9 ± 1.0	99.1 ± 0.4
	2017	313901	9.29 ± 0.22	0.225 ± 0.008	93.5 ± 1.0	97.6 ± 0.7
	2018	317305	9.35 ± 0.22	0.187 ± 0.008	94.7 ± 1.0	98.6 ± 0.5
$B^+ \to \overline{D}^0 D^+ K^{*0}$	2016	619320	6.83 ± 0.12	0.749 ± 0.011	95.14 ± 0.33	96.59 ± 0.29
	2017	767934	6.80 ± 0.11	0.838 ± 0.010	94.71 ± 0.29	96.97 ± 0.23
	2018	689998	7.07 ± 0.11	0.723 ± 0.010	95.04 ± 0.33	96.69 ± 0.28
$B^+ \to \overline{D}^{*0} D^+ K^{*0}$	2016	632950	6.81 ± 0.11	0.657 ± 0.010	93.7 ± 0.4	96.60 ± 0.31
	2017	611949	6.83 ± 0.12	0.758 ± 0.011	94.3 ± 0.4	96.27 ± 0.30
	2018	610000	6.92 ± 0.12	0.634 ± 0.010	93.0 ± 0.4	95.6 ± 0.4

Table 24: Summary of Event Efficiencies

Source	Year	Number Accepted	$\epsilon_{Generator}$	$\epsilon_{stripping}$	$\epsilon_{offline}$	$\epsilon_{trigger}$
$B^0 \to D^- D^+ K^{*0}$	2016	655750	5.34 ± 0.09	0.463 ± 0.008	94.1 ± 0.4	94.9 ± 0.4
	2017	603999	5.16 ± 0.09	0.524 ± 0.009	94.5 ± 0.4	94.1 ± 0.4
	2018	689999	5.24 ± 0.09	0.422 ± 0.008	94.3 ± 0.4	94.5 ± 0.5
$B^0 \to (D^{*-} \to D^- \pi^0) D^+ K^{*0}$	2016	643848	5.12 ± 0.09	0.439 ± 0.008	94.0 ± 0.5	92.4 ± 0.5
	2017	715999	5.11 ± 0.09	0.439 ± 0.008	93.2 ± 0.5	92.0 ± 0.5
	2018	698000	5.13 ± 0.08	0.393 ± 0.007	93.8 ± 0.5	91.1 ± 0.6
$B^0 \to (D^{*-} \to D^- \pi^0)(D^{*+} \to D^+ \pi^0) K^{*0}$	2016	631999	4.93 ± 0.09	0.354 ± 0.007	93.8 ± 0.5	89.1 ± 0.7
	2017	609799	5.08 ± 0.09	0.354 ± 0.008	93.1 ± 0.6	89.8 ± 0.7
	2018	619599	5.09 ± 0.09	0.341 ± 0.007	94.1 ± 0.5	90.2 ± 0.7
$B^0 \to (D^{*-} \to \overline{D}^0 \pi^-)(D^{*+} \to D^+ \pi^0) K^{*0}$	2016	615799	6.72 ± 0.12	0.588 ± 0.010	94.1 ± 0.4	95.4 ± 0.4
	2017	607999	6.75 ± 0.12	0.620 ± 0.010	93.8 ± 0.4	95.1 ± 0.4
	2018	608000	6.86 ± 0.12	0.544 ± 0.009	93.6 ± 0.5	94.0 ± 0.5
$B^0 \to (D^{*-} \to D^- \pi^0)(D^{*+} \to \overline{D} \pi^+) K^{*0}$	2016	308518	9.30 ± 0.25	0.778 ± 0.016	92.9 ± 0.6	96.8 ± 0.4
	2017	313901	9.29 ± 0.22	0.938 ± 0.017	94.0 ± 0.5	96.1 ± 0.4
	2018	317305	9.35 ± 0.22	0.754 ± 0.015	92.9 ± 0.6	96.9 ± 0.4
$B^0 \to (D^{*-} \to \overline{D}^0 \pi^-)(D^{*+} \to \overline{D} \pi^+) K^{*0}$	2016	308518	9.30 ± 0.25	0.182 ± 0.008	94.9 ± 1.0	99.1 ± 0.4
	2017	313901	9.29 ± 0.22	0.225 ± 0.008	93.5 ± 1.0	97.6 ± 0.7
	2018	317305	9.35 ± 0.22	0.187 ± 0.008	94.7 ± 1.0	98.6 ± 0.5
$B^+ \to \overline{D}^0 D^+ K^{*0}$	2016	619320	6.83 ± 0.12	0.749 ± 0.011	95.14 ± 0.33	96.59 ± 0.29
	2017	767934	6.80 ± 0.11	0.838 ± 0.010	94.71 ± 0.29	96.97 ± 0.23
	2018	689998	7.07 ± 0.11	0.723 ± 0.010	95.04 ± 0.33	96.69 ± 0.28
$B^+ \to \overline{D}^{*0} D^+ K^{*0}$	2016	632950	6.81 ± 0.11	0.657 ± 0.010	93.7 ± 0.4	96.60 ± 0.31
	2017	611949	6.83 ± 0.12	0.758 ± 0.011	94.3 ± 0.4	96.27 ± 0.30
	2018	610000	6.92 ± 0.12	0.634 ± 0.010	93.0 ± 0.4	95.6 ± 0.4

Table 25: Summary of Event Efficiencies

Source	Year	ϵ_D	$\epsilon_{bkgveto}$	ϵ_{clone}	$\epsilon_{MultipleCandidate}$
$B^0 \to D^- D^+ K^{*0}$	2016	82.2 ± 0.8	100.0 ± 0	99.37 ± 0.17	0.9902 ± 0.0022
	2017	83.8 ± 0.7	100.0 ± 0	99.36 ± 0.17	0.9866 ± 0.0025
	2018	80.8 ± 0.8	100.0 ± 0	99.18 ± 0.20	0.9917 ± 0.0021
$B^0 \to (D^{*-} \to D^- \pi^0) D^+ K^{*0}$	2016	83.8 ± 0.8	100.0 ± 0	99.27 ± 0.20	0.9884 ± 0.0025
	2017	84.7 ± 0.7	100.0 ± 0	99.25 ± 0.19	0.9901 ± 0.0022
	2018	82.5 ± 0.8	100.0 ± 0	99.21 ± 0.21	0.9909 ± 0.0023
$B^0 \to (D^{*-} \to D^- \pi^0)(D^{*+} \to D^+ \pi^0) K^{*0}$	2016	81.3 ± 0.9	100.0 ± 0	99.01 ± 0.26	0.9857 ± 0.0032
	2017	83.0 ± 0.9	100.0 ± 0	98.75 ± 0.30	0.9941 ± 0.0021
	2018	82.4 ± 0.9	100.0 ± 0	98.95 ± 0.28	0.9894 ± 0.0028
$B^0 \to (D^{*-} \to \overline{D}^0 \pi^-)(D^{*+} \to D^+ \pi^0) K^{*0}$	2016	85.4 ± 0.7	98.98 ± 0.20	99.14 ± 0.19	0.9983 ± 0.0008
	2017	84.7 ± 0.7	98.76 ± 0.22	99.56 ± 0.13	0.9996 ± 0.0004
	2018	85.1 ± 0.7	98.40 ± 0.27	99.30 ± 0.18	0.9991 ± 0.0007
$B^0 \to (D^{*-} \to D^- \pi^0)(D^{*+} \to \overline{D} \pi^+) K^{*0}$	2016	87.3 ± 0.8	98.84 ± 0.27	99.87 ± 0.09	1.0 ± 0
	2017	86.3 ± 0.7	98.24 ± 0.30	99.62 ± 0.14	0.9995 ± 0.0005
	2018	85.6 ± 0.8	98.68 ± 0.29	99.20 ± 0.23	1.0 ± 0
$B^0 \to (D^{*-} \to \overline{D}^0 \pi^-)(D^{*+} \to \overline{D} \pi^+) K^{*0}$	2016	76.7 ± 2.0	100.0 ± 0	98.5 ± 0.6	1.0 ± 0
	2017	81.9 ± 1.7	100.0 ± 0	98.9 ± 0.5	1.0 ± 0
	2018	80.3 ± 1.8	100.0 ± 0	99.2 ± 0.5	1.0 ± 0
$B^+ \to \overline{D}^0 D^+ K^{*0}$	2016	86.0 ± 0.6	99.79 ± 0.08	99.57 ± 0.11	0.9978 ± 0.0008
	2017	86.1 ± 0.5	99.80 ± 0.07	99.49 ± 0.11	0.9987 ± 0.0005
	2018	86.1 ± 0.5	99.80 ± 0.07	99.52 ± 0.12	0.9986 ± 0.0006
$B^+ \to \overline{D}^{*0} D^+ K^{*0}$	2016	85.8 ± 0.6	99.965 ± 0.035	99.31 ± 0.15	0.9986 ± 0.0007
	2017	86.3 ± 0.6	99.969 ± 0.031	99.35 ± 0.14	0.9994 ± 0.0004
	2018	87.4 ± 0.6	100.0 ± 0	99.51 ± 0.13	0.9989 ± 0.0006

Table 26: Summary of Event Efficiencies

Source	Year	ϵ_D	$\epsilon_{bkgveto}$	ϵ_{clone}	$\epsilon_{MultipleCandidate}$
$B^0 \to D^- D^+ K^{*0}$	2016	82.2 ± 0.8	100.0 ± 0	99.37 ± 0.17	0.9902 ± 0.0022
	2017	83.8 ± 0.7	100.0 ± 0	99.36 ± 0.17	0.9866 ± 0.0025
	2018	80.8 ± 0.8	100.0 ± 0	99.18 ± 0.20	0.9917 ± 0.0021
$B^0 \to (D^{*-} \to D^- \pi^0) D^+ K^{*0}$	2016	83.8 ± 0.8	100.0 ± 0	99.27 ± 0.20	0.9884 ± 0.0025
	2017	84.7 ± 0.7	100.0 ± 0	99.25 ± 0.19	0.9901 ± 0.0022
	2018	82.5 ± 0.8	100.0 ± 0	99.21 ± 0.21	0.9909 ± 0.0023
$B^0 \to (D^{*-} \to D^- \pi^0)(D^{*+} \to D^+ \pi^0) K^{*0}$	2016	81.3 ± 0.9	100.0 ± 0	99.01 ± 0.26	0.9857 ± 0.0032
	2017	83.0 ± 0.9	100.0 ± 0	98.75 ± 0.30	0.9941 ± 0.0021
	2018	82.4 ± 0.9	100.0 ± 0	98.95 ± 0.28	0.9894 ± 0.0028
$B^0 \to (D^{*-} \to \overline{D}^0 \pi^-)(D^{*+} \to D^+ \pi^0) K^{*0}$	2016	85.4 ± 0.7	98.98 ± 0.20	99.14 ± 0.19	0.9983 ± 0.0008
	2017	84.7 ± 0.7	98.76 ± 0.22	99.56 ± 0.13	0.9996 ± 0.0004
	2018	85.1 ± 0.7	98.40 ± 0.27	99.30 ± 0.18	0.9991 ± 0.0007
$B^0 \to (D^{*-} \to D^- \pi^0)(D^{*+} \to \overline{D} \pi^+) K^{*0}$	2016	87.3 ± 0.8	98.84 ± 0.27	99.87 ± 0.09	1.0 ± 0
	2017	86.3 ± 0.7	98.24 ± 0.30	99.62 ± 0.14	0.9995 ± 0.0005
	2018	85.6 ± 0.8	98.68 ± 0.29	99.20 ± 0.23	1.0 ± 0
$B^0 \to (D^{*-} \to \overline{D}^0 \pi^-)(D^{*+} \to \overline{D} \pi^+) K^{*0}$	2016	76.7 ± 2.0	100.0 ± 0	98.5 ± 0.6	1.0 ± 0
	2017	81.9 ± 1.7	100.0 ± 0	98.9 ± 0.5	1.0 ± 0
	2018	80.3 ± 1.8	100.0 ± 0	99.2 ± 0.5	1.0 ± 0
$B^+ \to \overline{D}^0 D^+ K^{*0}$	2016	86.0 ± 0.6	99.79 ± 0.08	99.57 ± 0.11	0.9978 ± 0.0008
	2017	86.1 ± 0.5	99.80 ± 0.07	99.49 ± 0.11	0.9987 ± 0.0005
	2018	86.1 ± 0.5	99.80 ± 0.07	99.52 ± 0.12	0.9986 ± 0.0006
$B^+ \to \overline{D}^{*0} D^+ K^{*0}$	2016	85.8 ± 0.6	99.965 ± 0.035	99.31 ± 0.15	0.9986 ± 0.0007
	2017	86.3 ± 0.6	99.969 ± 0.031	99.35 ± 0.14	0.9994 ± 0.0004
	2018	87.4 ± 0.6	100.0 ± 0	99.51 ± 0.13	0.9989 ± 0.0006

Table 27: Summary of final efficiencies for our simulaton with corrections for ReDecay, but before any systematics are applied.

Source	Year	Final TOS Eff	Final TIS Eff
$B^0 \to D^- D^+ K^{*0}$	2016	$(1.030 \pm 0.050) \times 10^{-4}$	$(6.181 \pm 0.339) \times 10^{-5}$
	2017	$(1.187 \pm 0.058) \times 10^{-4}$	$(6.389 \pm 0.350) \times 10^{-5}$
	2018	$(9.462 \pm 0.456) \times 10^{-5}$	$(5.087 \pm 0.278) \times 10^{-5}$
$B^0 \to (D^{*-} \to D^- \pi^0) D^+ K^{*0}$	2016	$(9.013 \pm 0.460) \times 10^{-5}$	$(5.854 \pm 0.320) \times 10^{-5}$
	2017	$(9.698 \pm 0.454) \times 10^{-5}$	$(5.208 \pm 0.282) \times 10^{-5}$
	2018	$(7.654 \pm 0.385) \times 10^{-5}$	$(5.110 \pm 0.296) \times 10^{-5}$
$B^0 \to (D^{*-} \to D^- \pi^0)(D^{*+} \to D^+ \pi^0) K^{*0}$	2016	$(6.537 \pm 0.370) \times 10^{-5}$	$(4.248 \pm 0.263) \times 10^{-5}$
	2017	$(6.977 \pm 0.382) \times 10^{-5}$	$(4.189 \pm 0.259) \times 10^{-5}$
	2018	$(6.230 \pm 0.352) \times 10^{-5}$	$(4.559 \pm 0.296) \times 10^{-5}$
$B^0 \to (D^{*-} \to \overline{D}^0 \pi^-)(D^{*+} \to D^+ \pi^0) K^{*0}$	2016	$(1.566 \pm 0.075) \times 10^{-4}$	$(1.062 \pm 0.057) \times 10^{-4}$
	2017	$(1.686 \pm 0.077) \times 10^{-4}$	$(1.052 \pm 0.057) \times 10^{-4}$
	2018	$(1.469 \pm 0.073) \times 10^{-4}$	$(9.363 \pm 0.532) \times 10^{-5}$
$B^0 \to (D^{*-} \to \overline{D}^0 \pi^-)(D^{*+} \to \overline{D} \pi^+) K^{*0}$	2016	$(2.802 \pm 0.173) \times 10^{-4}$	$(1.809 \pm 0.126) \times 10^{-4}$
	2017	$(3.310 \pm 0.191) \times 10^{-4}$	$(2.124 \pm 0.139) \times 10^{-4}$
	2018	$(2.541 \pm 0.152) \times 10^{-4}$	$(1.841 \pm 0.123) \times 10^{-4}$
$B^0 \to (D^{*-} \to \overline{D}^0 \pi^-)(D^{*+} \to \overline{D} \pi^+) K^{*0}$	2016	$(7.498 \pm 0.703) \times 10^{-5}$	$(2.686 \pm 0.341) \times 10^{-5}$
	2017	$(9.042 \pm 0.782) \times 10^{-5}$	$(3.766 \pm 0.416) \times 10^{-5}$
	2018	$(7.732 \pm 0.666) \times 10^{-5}$	$(3.473 \pm 0.392) \times 10^{-5}$
$B^+ \to \overline{D}^0 D^+ K^{*0}$	2016	$(2.296 \pm 0.101) \times 10^{-4}$	$(1.279 \pm 0.064) \times 10^{-4}$
	2017	$(2.728 \pm 0.104) \times 10^{-4}$	$(1.261 \pm 0.057) \times 10^{-4}$
	2018	$(2.263 \pm 0.097) \times 10^{-4}$	$(1.335 \pm 0.065) \times 10^{-4}$
$B^+ \to \overline{D}^{*0} D^+ K^{*0}$	2016	$(1.966 \pm 0.091) \times 10^{-4}$	$(1.106 \pm 0.058) \times 10^{-4}$
	2017	$(2.341 \pm 0.100) \times 10^{-4}$	$(1.267 \pm 0.065) \times 10^{-4}$
	2018	$(1.880 \pm 0.087) \times 10^{-4}$	$(1.137 \pm 0.060) \times 10^{-4}$

Table 28: Summary of final efficiencies for our simulaton with corrections for ReDecay, but before any systematics are applied.

Source	Year	Final TOS Eff	Final TIS Eff
$B^+ \to \overline{D}^0(D^{*+} \to D^+\pi^0)K^{*0}$	2016	$(1.971 \pm 0.090) \times 10^{-4}$	$(1.194 \pm 0.061) \times 10^{-4}$
	2017	$(2.152 \pm 0.095) \times 10^{-4}$	$(1.240 \pm 0.063) \times 10^{-4}$
	2018	$(1.850 \pm 0.082) \times 10^{-4}$	$(1.203 \pm 0.061) \times 10^{-4}$
$B^+ \to \overline{D}^0(D^{*+} \to \overline{D}\pi^+)K^{*0}$	2016	$(3.187 \pm 0.178) \times 10^{-4}$	$(2.097 \pm 0.137) \times 10^{-4}$
	2017	$(3.643 \pm 0.204) \times 10^{-4}$	$(2.221 \pm 0.140) \times 10^{-4}$
	2018	$(2.930 \pm 0.177) \times 10^{-4}$	$(1.961 \pm 0.135) \times 10^{-4}$
$B^+ \to \overline{D}^0(D^{*+} \to \overline{D}\pi^+)K^{*0}$	2016	$(1.057 \pm 0.088) \times 10^{-4}$	$(4.288 \pm 0.492) \times 10^{-5}$
	2017	$(1.194 \pm 0.095) \times 10^{-4}$	$(3.805 \pm 0.414) \times 10^{-5}$
	2018	$(9.428 \pm 0.819) \times 10^{-5}$	$(4.651 \pm 0.494) \times 10^{-5}$
$B^+ \to \overline{D}^{*0}(D^{*+} \to D^+\pi^0)K^{*0}$	2016	$(1.812 \pm 0.082) \times 10^{-4}$	$(1.084 \pm 0.057) \times 10^{-4}$
	2017	$(2.035 \pm 0.083) \times 10^{-4}$	$(1.204 \pm 0.057) \times 10^{-4}$
	2018	$(1.694 \pm 0.080) \times 10^{-4}$	$(1.066 \pm 0.059) \times 10^{-4}$
$B^+ \to \overline{D}^{*0}(D^{*+} \to \overline{D}\pi^+)K^{*0}$	2016	$(2.668 \pm 0.166) \times 10^{-4}$	$(2.123 \pm 0.139) \times 10^{-4}$
	2017	$(3.479 \pm 0.197) \times 10^{-4}$	$(2.373 \pm 0.151) \times 10^{-4}$
	2018	$(2.620 \pm 0.166) \times 10^{-4}$	$(2.017 \pm 0.136) \times 10^{-4}$
$B^+ \to \overline{D}^{*0}(D^{*+} \to \overline{D}\pi^+)K^{*0}$	2016	$(8.528 \pm 0.785) \times 10^{-5}$	$(3.323 \pm 0.375) \times 10^{-5}$
	2017	$(1.123 \pm 0.092) \times 10^{-4}$	$(4.371 \pm 0.479) \times 10^{-5}$
	2018	$(6.849 \pm 0.664) \times 10^{-5}$	$(3.601 \pm 0.411) \times 10^{-5}$
$B^0 \to \overline{D}^0 D^0 K^{*0}$	2016	$(4.492 \pm 0.186) \times 10^{-4}$	$(2.635 \pm 0.123) \times 10^{-4}$
	2017	$(5.089 \pm 0.193) \times 10^{-4}$	$(2.607 \pm 0.110) \times 10^{-4}$
	2018	$(4.439 \pm 0.189) \times 10^{-4}$	$(2.498 \pm 0.120) \times 10^{-4}$
$B^0 \to \overline{D}^{*0} D^0 K^{*0} + B^0 \to \overline{D}^0 D^{*0} K^{*0}$	2016	$(4.150 \pm 0.172) \times 10^{-4}$	$(2.516 \pm 0.119) \times 10^{-4}$
	2017	$(4.668 \pm 0.178) \times 10^{-4}$	$(2.517 \pm 0.109) \times 10^{-4}$
	2018	$(4.255 \pm 0.169) \times 10^{-4}$	$(2.464 \pm 0.114) \times 10^{-4}$
$B^0 \to \overline{D}^{*0} D^{*0} K^{*0}$	2016	$(3.703 \pm 0.149) \times 10^{-4}$	$(2.403 \pm 0.107) \times 10^{-4}$
	2017	$(4.056 \pm 0.173) \times 10^{-4}$	$(2.333 \pm 0.109) \times 10^{-4}$
	2018	$(3.645 \pm 0.161) \times 10^{-4}$	$(2.206 \pm 0.109) \times 10^{-4}$
$B^+ \to \overline{D}^0(D^0 \to K^-\pi^+\pi^+\pi^-)K^+$	2016	$(2.597 \pm 0.110) \times 10^{-4}$	$(1.419 \pm 0.067) \times 10^{-4}$
	2017	$(2.851 \pm 0.121) \times 10^{-4}$	$(1.363 \pm 0.066) \times 10^{-4}$
	2018	$(2.517 \pm 0.104) \times 10^{-4}$	$(1.315 \pm 0.063) \times 10^{-4}$
$B^0 \to D^-(D^0 \to K^-\pi^+\pi^+\pi^-)K^+$	2016	$(1.180 \pm 0.056) \times 10^{-4}$	$(6.440 \pm 0.334) \times 10^{-5}$
	2017	$(1.387 \pm 0.065) \times 10^{-4}$	$(7.358 \pm 0.384) \times 10^{-5}$
	2018	$(1.116 \pm 0.054) \times 10^{-4}$	$(6.104 \pm 0.349) \times 10^{-5}$

B PDF Descriptions

Here we describe the PDF's used in this analysis. PDFs are used for the unbinned maximum likelihood fits to our B masses as well as several other distributions. When necessary we construct PDFs as a sum of two or more PDFs where PDFs are added as:

$$M(x) = \sum_{i=1}^{N-1} f_1 F_1(x) + (1 - \sum_{i=1}^{N-1} f_i) F_N(x) \tag{20}$$

within a RooFit Workspace. Normalization coefficients are dropped from the construction of our PDFs as normalization over our fit ranges and parameters is handled numerically by RooFit.

B.1 Gaussian (G)

The general form of the Gaussian PDF is:

$$F(x; \mu, \sigma) = e^{\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)} \tag{21}$$

where σ and μ are the width and mean parameters of the distribution.

B.2 Double Gaussian (DG)

A Double Gaussian is the sum of two Gaussian PDFs eq. (21), using eq. (20). They share a mean parameter but the two widths allowed to be different.

B.3 Gaussian with an Exponential Tail (GEP)

A Gaussian with an exponential tail [34] is an alternative to the Crystal Ball Function [35].

The general form is:

$$F(x;\alpha,\eta,\mu,\sigma) = \begin{cases} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma_1}\right)^2}, & \text{for } \frac{x-\mu}{\sigma} > -\alpha \\ (\frac{\eta}{|\alpha|})e^{-\frac{|\alpha|^2}{2}}\left(\frac{\eta}{|\alpha|}-|\alpha|-\frac{x-\mu}{\sigma}\right)^{-\eta}, & \text{for } \frac{x-\mu}{\sigma} \leq -\alpha \end{cases} \tag{22}$$

B.4 Bifurcated Gaussian (BG)

The general form of the Bifurcated Gaussian PDF is:

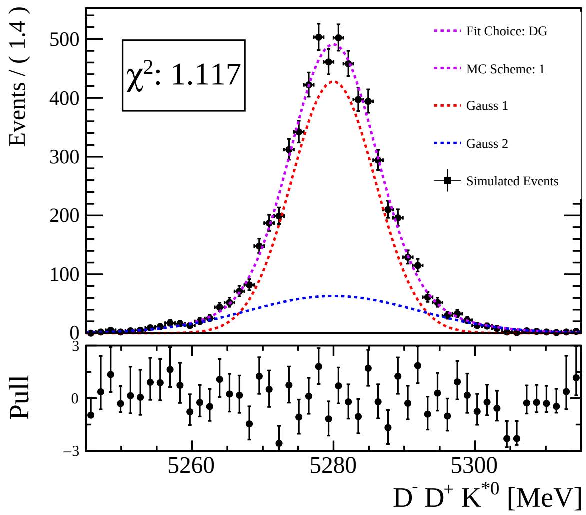
$$F(x;\mu,\sigma_1,\sigma_2) = \begin{cases} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma_1}\right)^2}, & \text{for } x < \mu \\ e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma_2}\right)^2}, & \text{for } x \geq \mu \end{cases} \tag{23}$$

B.5 Bifurcated Gaussian with an Exponential Tail (BGEP)

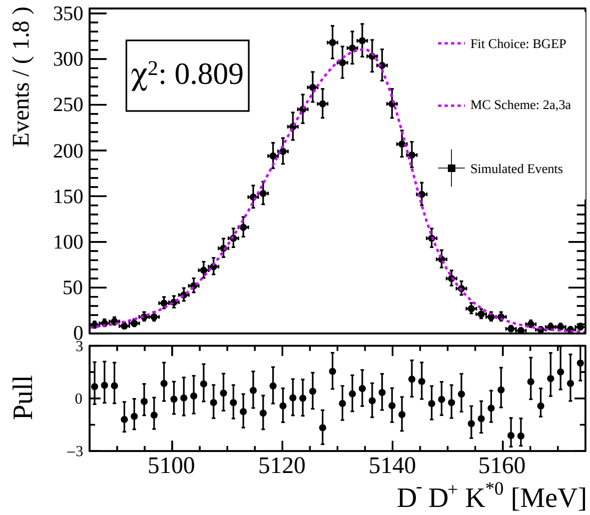
A Bifurcated Gaussian with an Exponential tail is an extension of eq. (22) to the case where two width parameters exists on either side of the mean, with an additional alpha parameter.

C Fits to MC

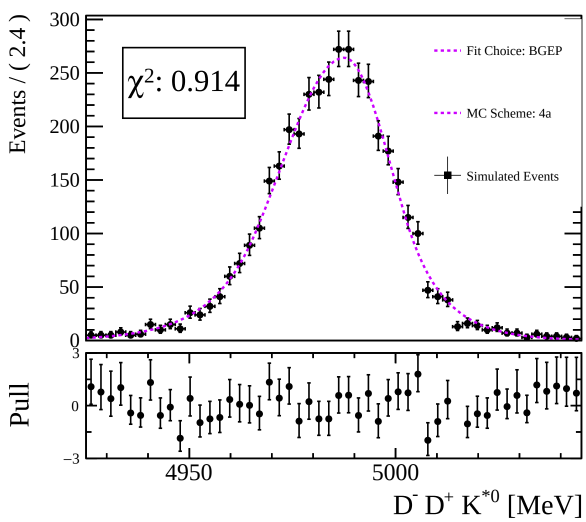
In this section we provide the plots of our best fits to all the MC samples present in this analysis. These choices are summarized in table 14



(a) $B^0 \to D^+ D^- K^{*0}$



(b) $B^0 \to D^{*+} D^- K^{*0} + B^0 \to D^+ D^{*-} K^{*0}$



(c) $B^0 \to D^{*+} D^{*-} K^{*0}$

Figure 32: MC Samples Reconstructed as $D^+ D^- K^{*0}$

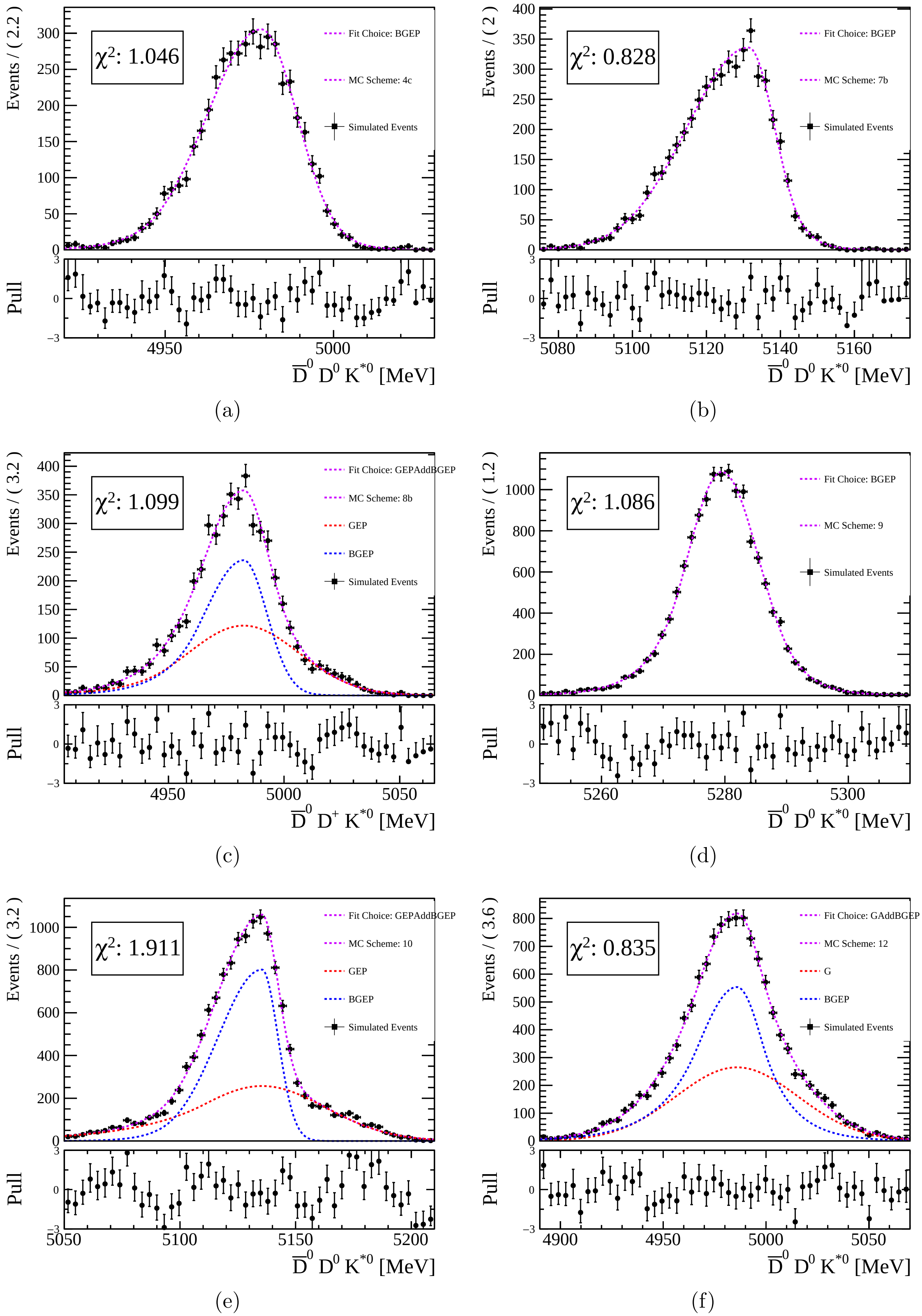
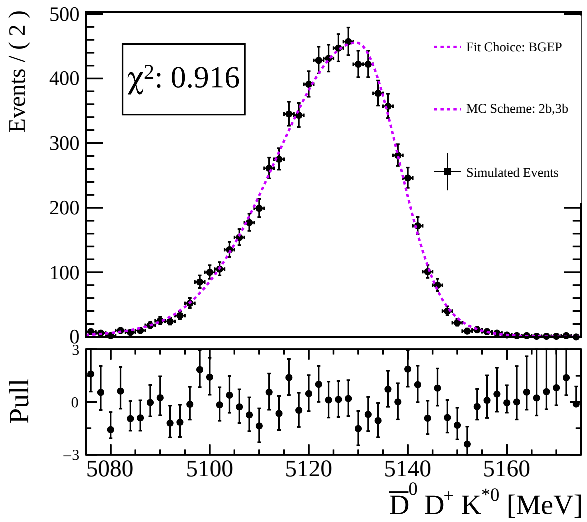
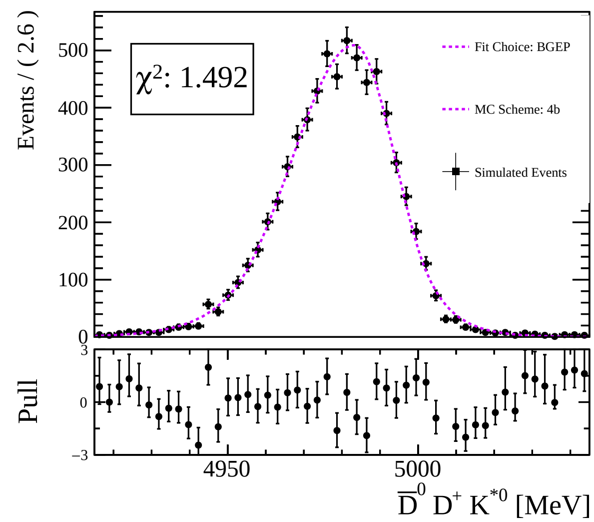


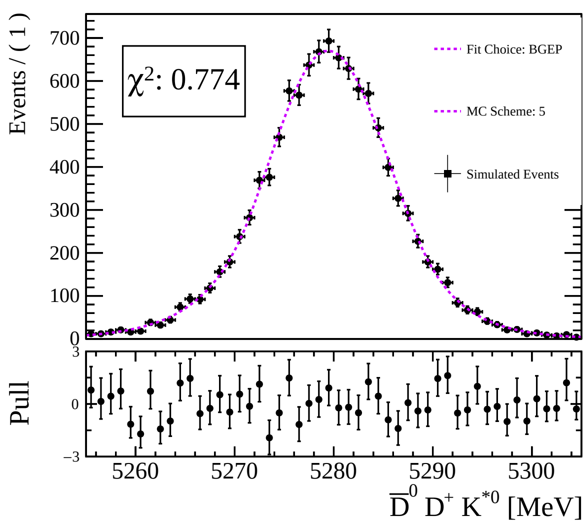
Figure 33: MC Samples Reconstructed as $\overline{D}^0 D^0 K^{*0}$



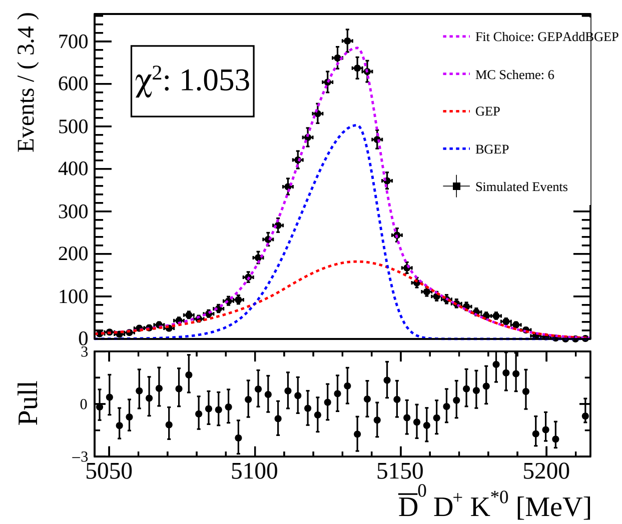
(a) $B^0 \rightarrow D^{*+}D^-K^{*0}$



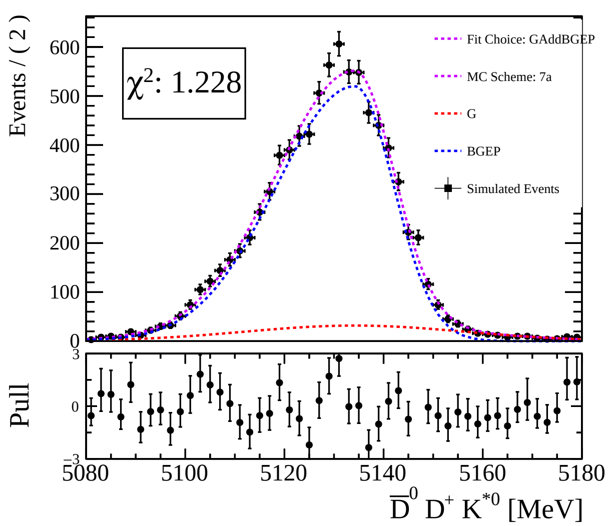
(b)



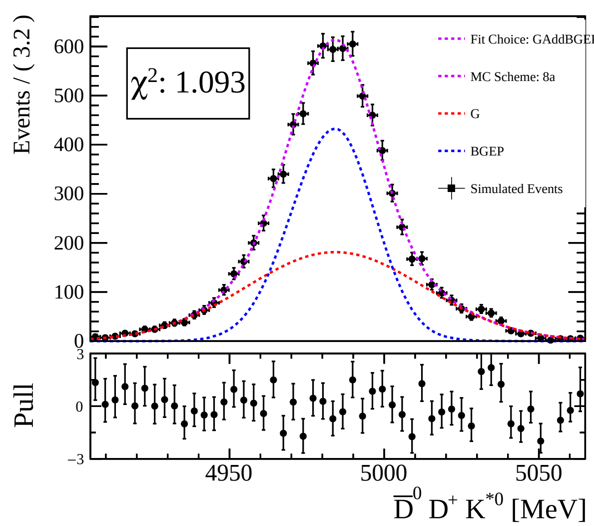
(c)



(d)



(e)



(f)

Figure 34: MC Samples Reconstructed as $\overline{D}^0 D^+ K^{*0}$

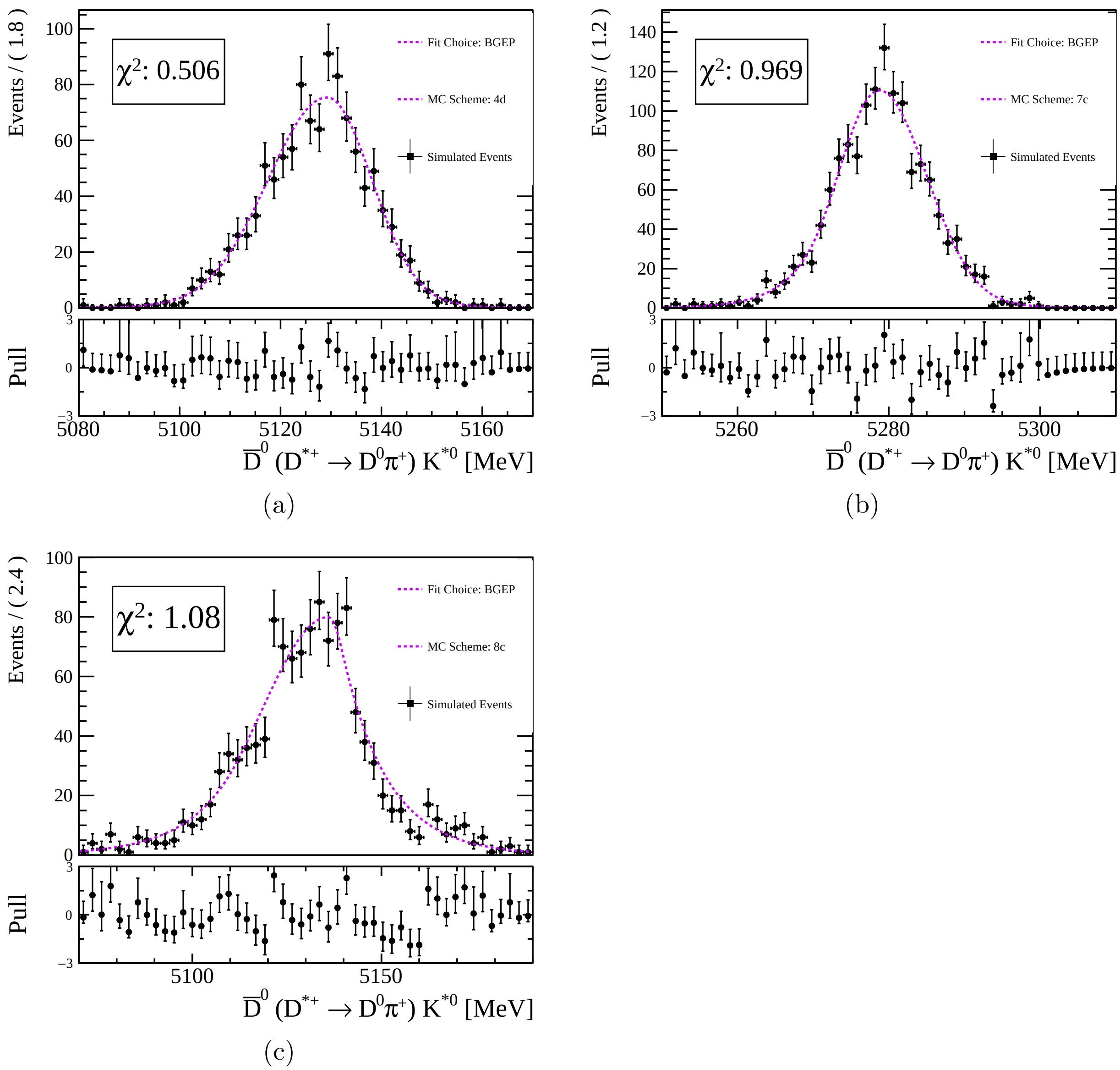
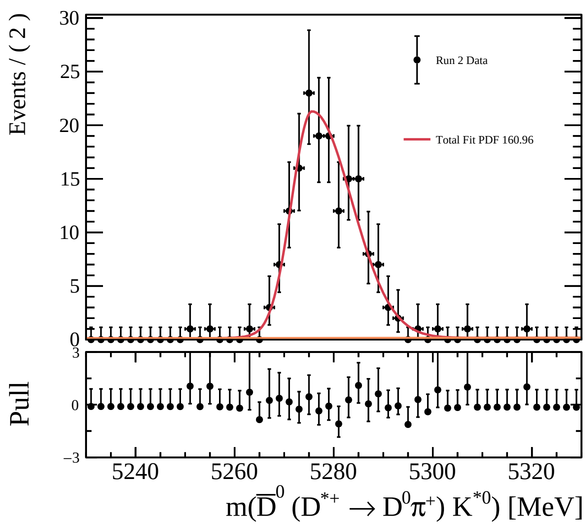


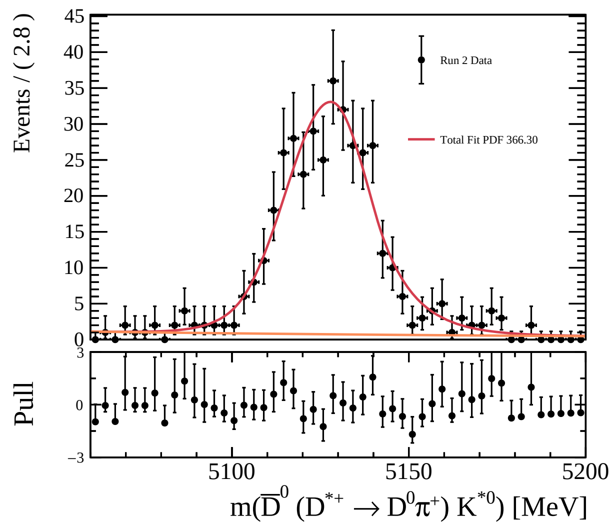
Figure 35: MC Samples Reconstructed as $\overline{D}^0$ $D^{*+} \rightarrow D^0\pi^+K^{*0}$

D Discrete Fits

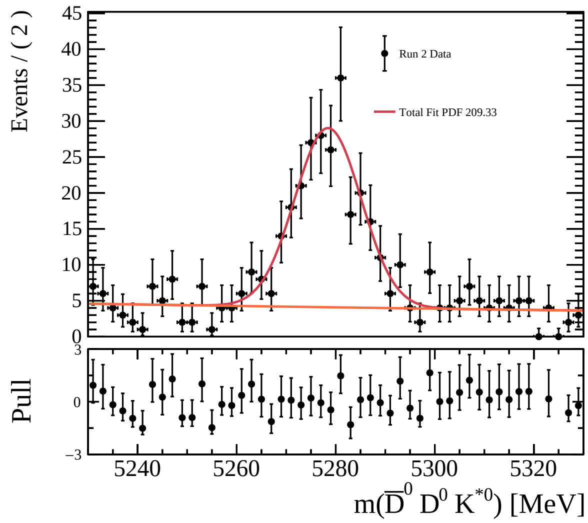
Fits to each B Mass for each data spectra. These fits do not incorporate any correlations or attempt to break up the individual contributions to each fit, and are only used to estimate certain systematic uncertainties.



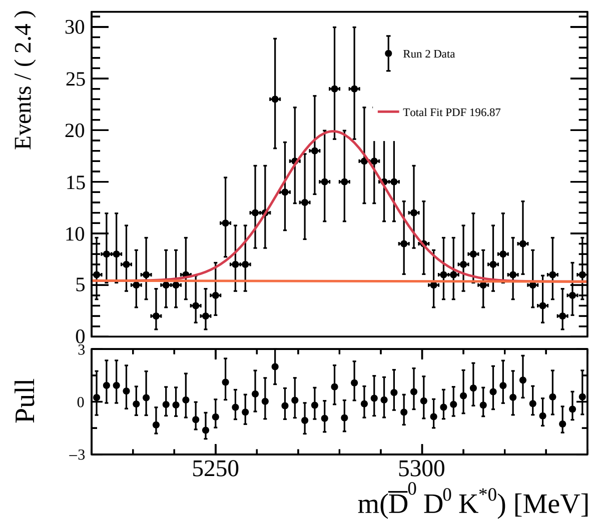
(a) Fit to the decay tree fitter constrained data where we miss zero particles in $\overline{D}^0(D^{*+} \to D^0\pi^+)K^{*0}$.



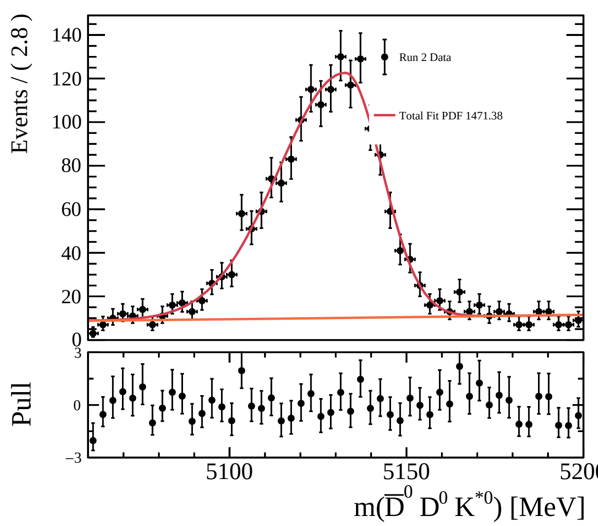
(b) Fit to the decay tree fitter constrained data where we miss one particle in $\overline{D}^0(D^{*+} \to D^0\pi^+)K^{*0}$.



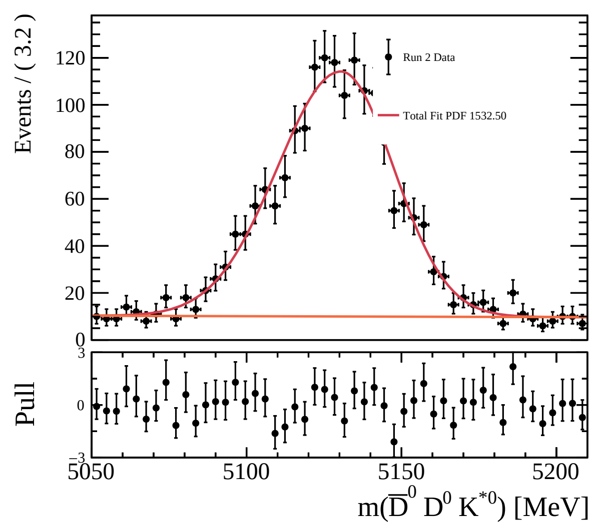
(a) Fit to the decay tree fitter constrained data where we miss zero particles in $\bar{D}^0 D^0 K^{*0}$.



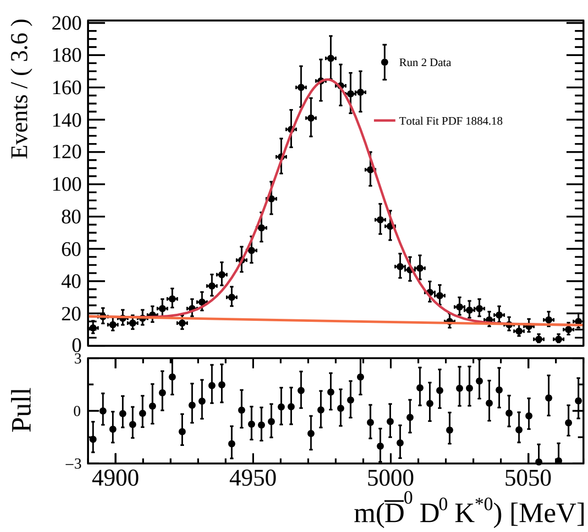
(b) Fit to the data with no decay tree fitter constraints where we miss zero particles in $\bar{D}^0 D^0 K^{*0}$.



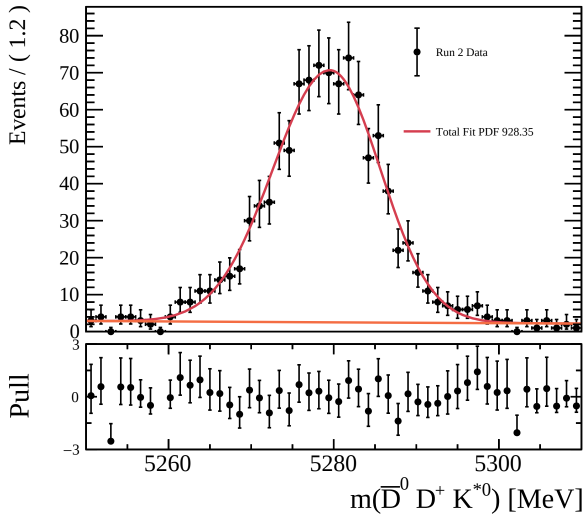
(c) Fit to the decay tree fitter constrained data where we miss one particle in $\bar{D}^0 D^0 K^{*0}$.



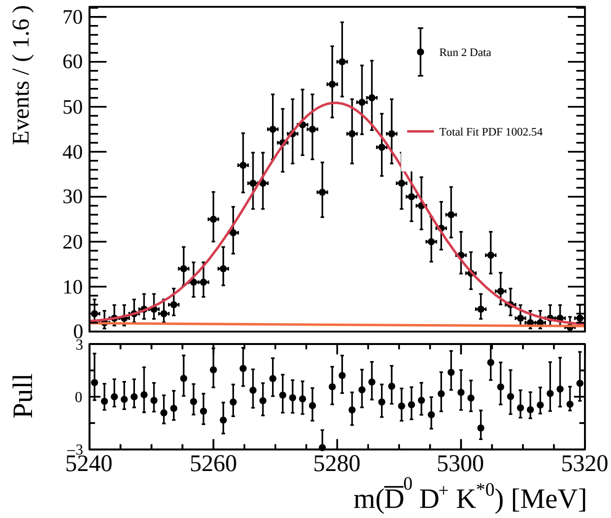
(d) Fit to the data with no decay tree fitter constraints where we miss one particle in $\bar{D}^0 D^0 K^{*0}$.



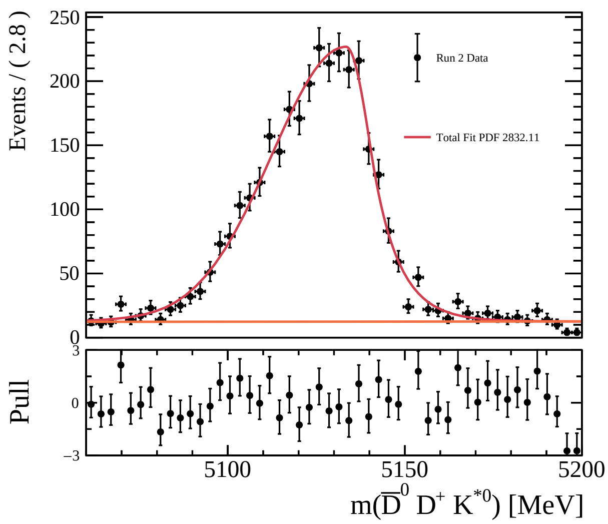
(e) Fit to the decay tree fitter constrained data where we miss two particles in $\bar{D}^0 D^0 K^{*0}$.



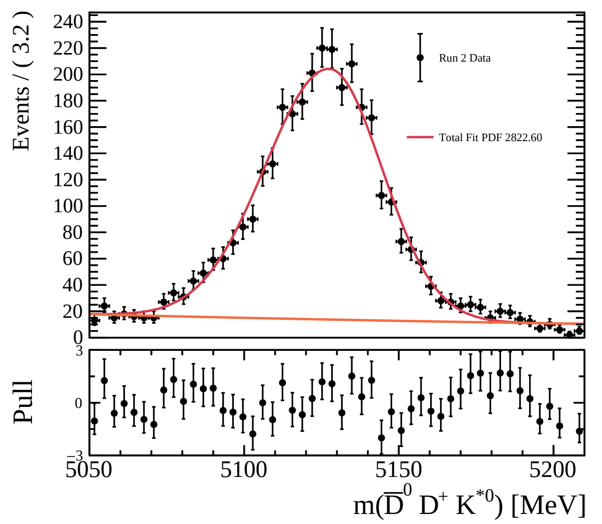
(a) Fit to the decay tree fitter constrained data where we miss zero particles in $\overline{D}^0 D^+ K^{*0}$.



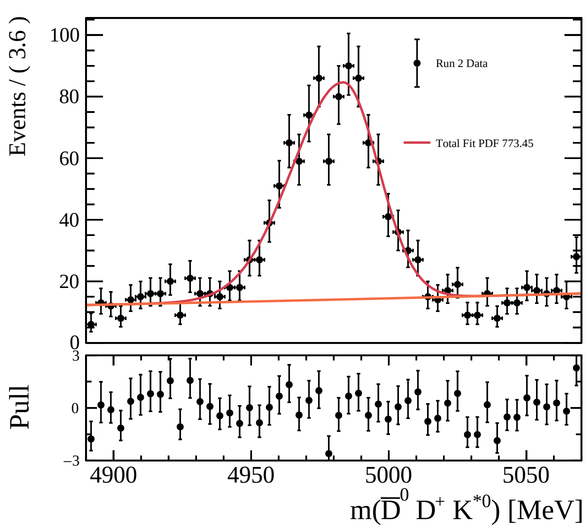
(b) Fit to the data with no decay tree fitter constraints where we miss zero particles in $\overline{D}^0 D^+ K^{*0}$.



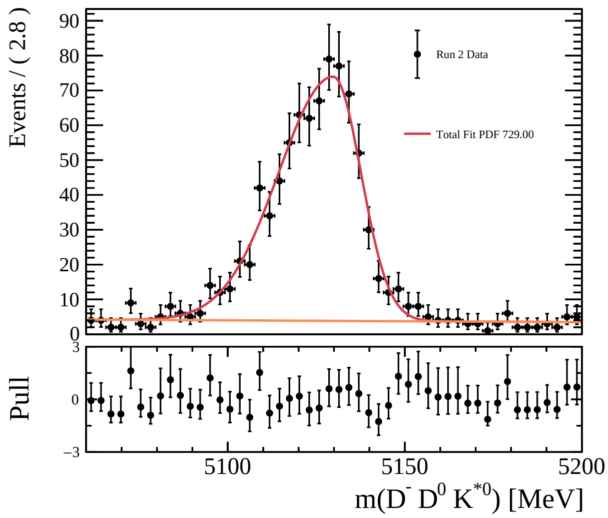
(c) Fit to the decay tree fitter constrained data where we miss one particle in $\overline{D}^0 D^+ K^{*0}$.



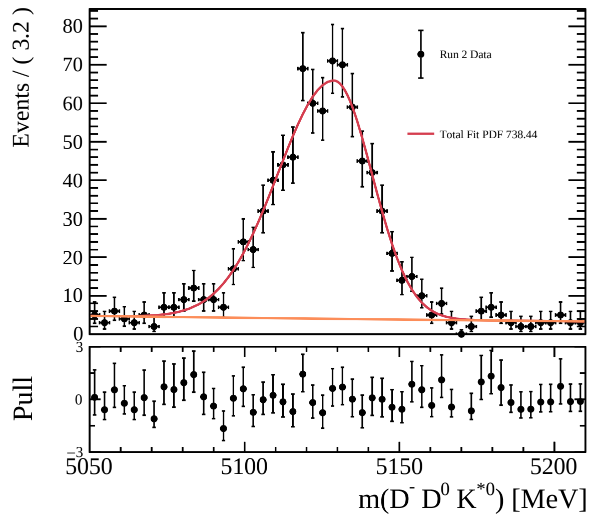
(d) Fit to the data with no decay tree fitter constraints where we miss one particle in $\overline{D}^0 D^+ K^{*0}$.



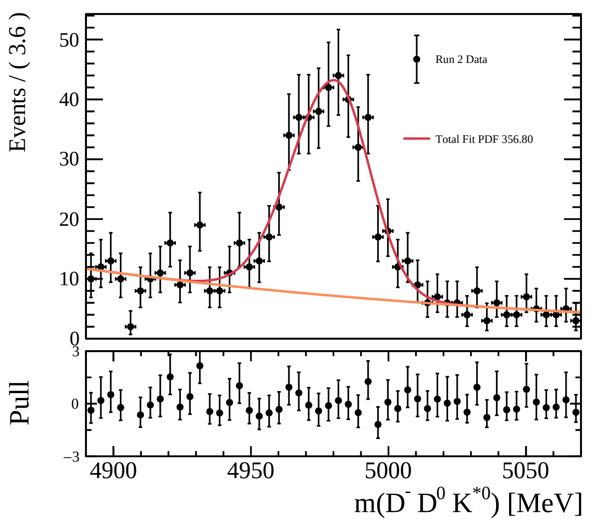
(e) Fit to the decay tree fitter constrained data where we miss two particles in $\overline{D}^0 D^+ K^{*0}$.



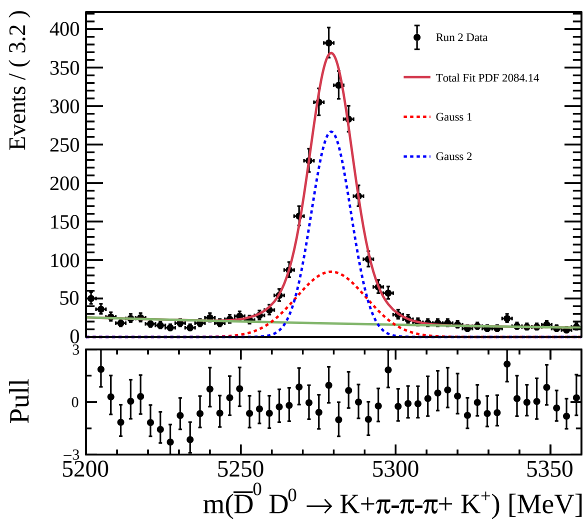
(a) Fit to the decay tree fitter constrained data where we miss one particle in $D^- D^0 K^{*0}$.



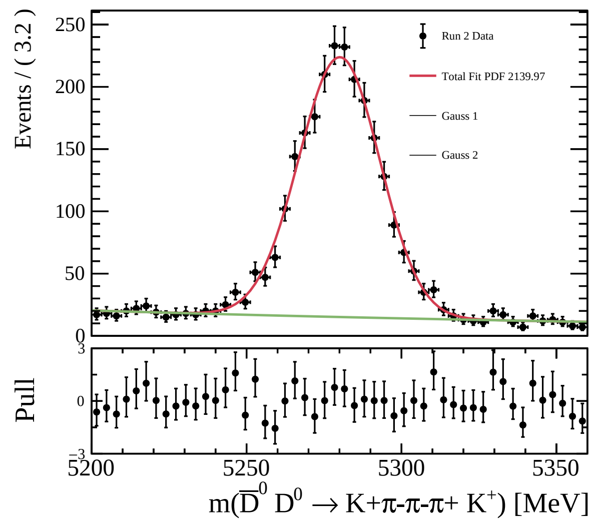
(b) Fit to the data with no decay tree fitter constraints where we miss one particle in $D^- D^0 K^{*0}$.



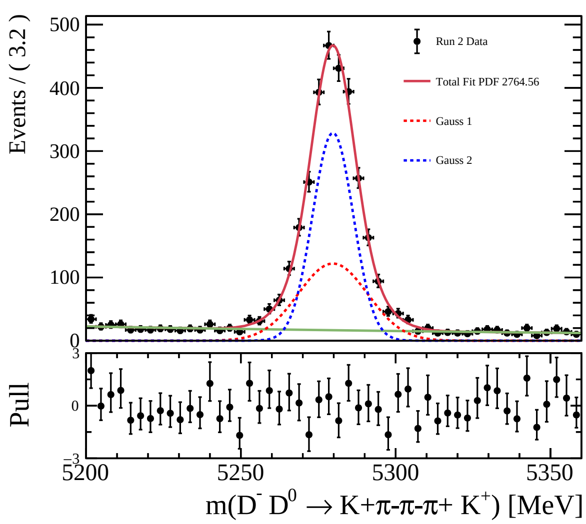
(c) Fit to the decay tree fitter constrained data where we miss two particles in $D^- D^0 K^{*0}$.



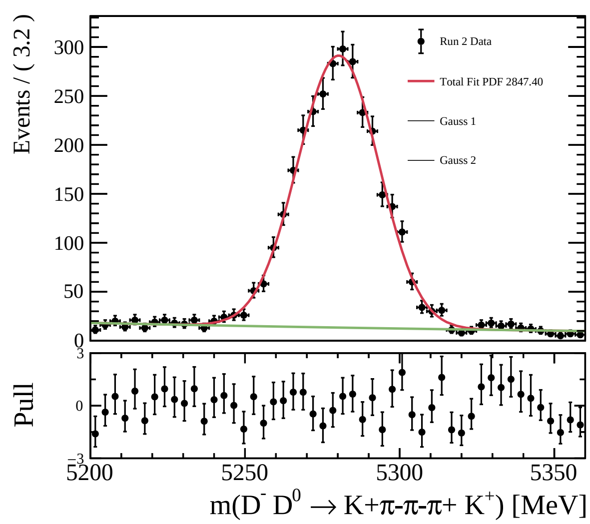
(a) Fit to the decay tree fitter constrained data where we miss zero particles in $\bar{D}^0(D^0 \to K^-\pi^+\pi^+\pi^-)K^+$.



(b) Fit to the data with no decay tree fitter constraints where we miss zero particles in $\bar{D}^0(D^0 \to K^-\pi^+\pi^+\pi^-)K^+$.



(c) Fit to the decay tree fitter constrained data where we miss zero particles in $D^-(D^0 \to K^-\pi^+\pi^+\pi^-)K^+$.



(d) Fit to the data with no decay tree fitter constraints where we miss zero particles in $D^-(D^0 \to K^-\pi^+\pi^+\pi^-)K^+$.

E Non-Parametric Fits to Partially truth matched MC

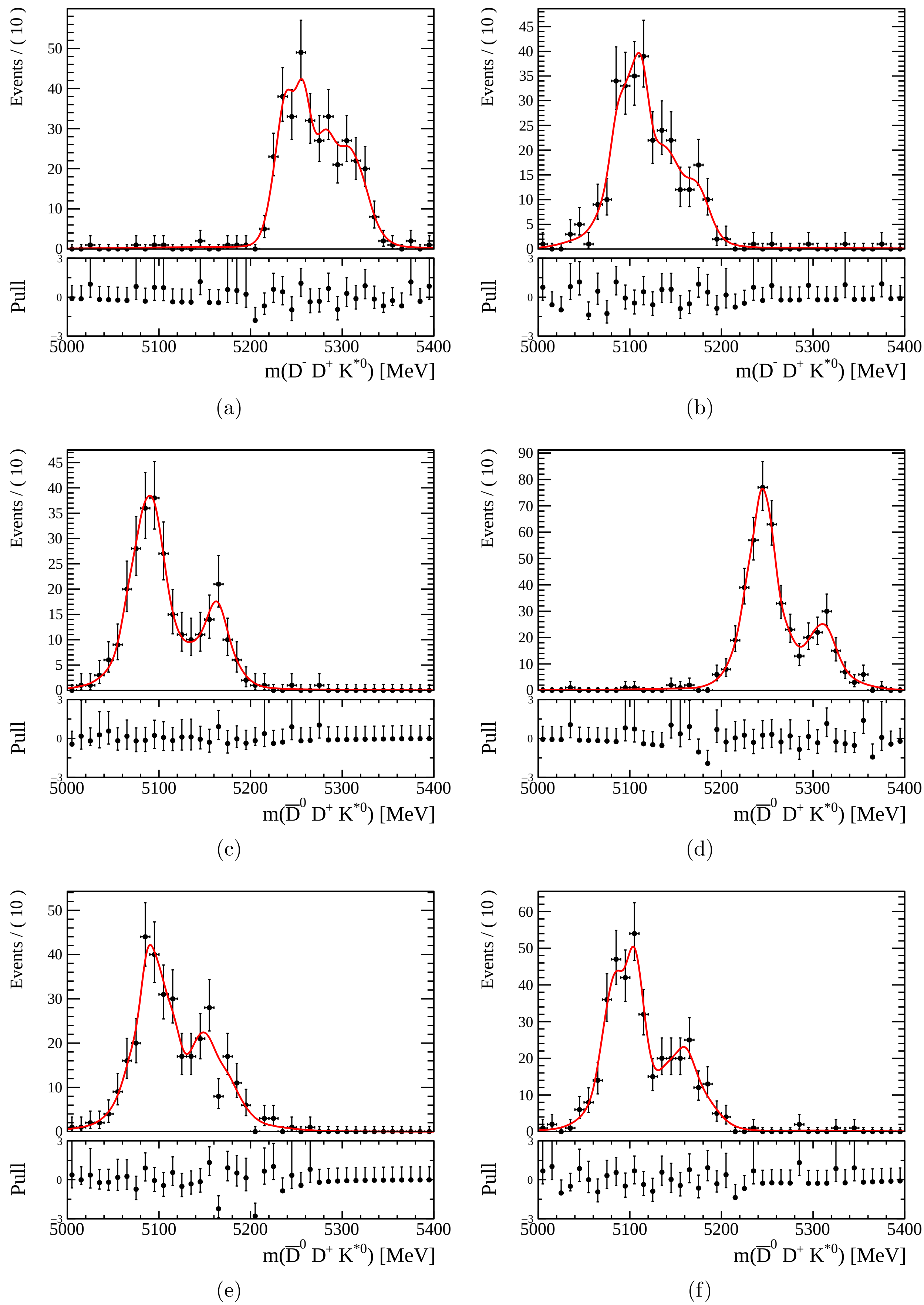


Figure 41: Non-Parametric Keys PDF on MC used in $D^-D^+K^{*0}$ and $\overline{D}^0D^+K^{*0}$ for charmless estimation in the D sideband

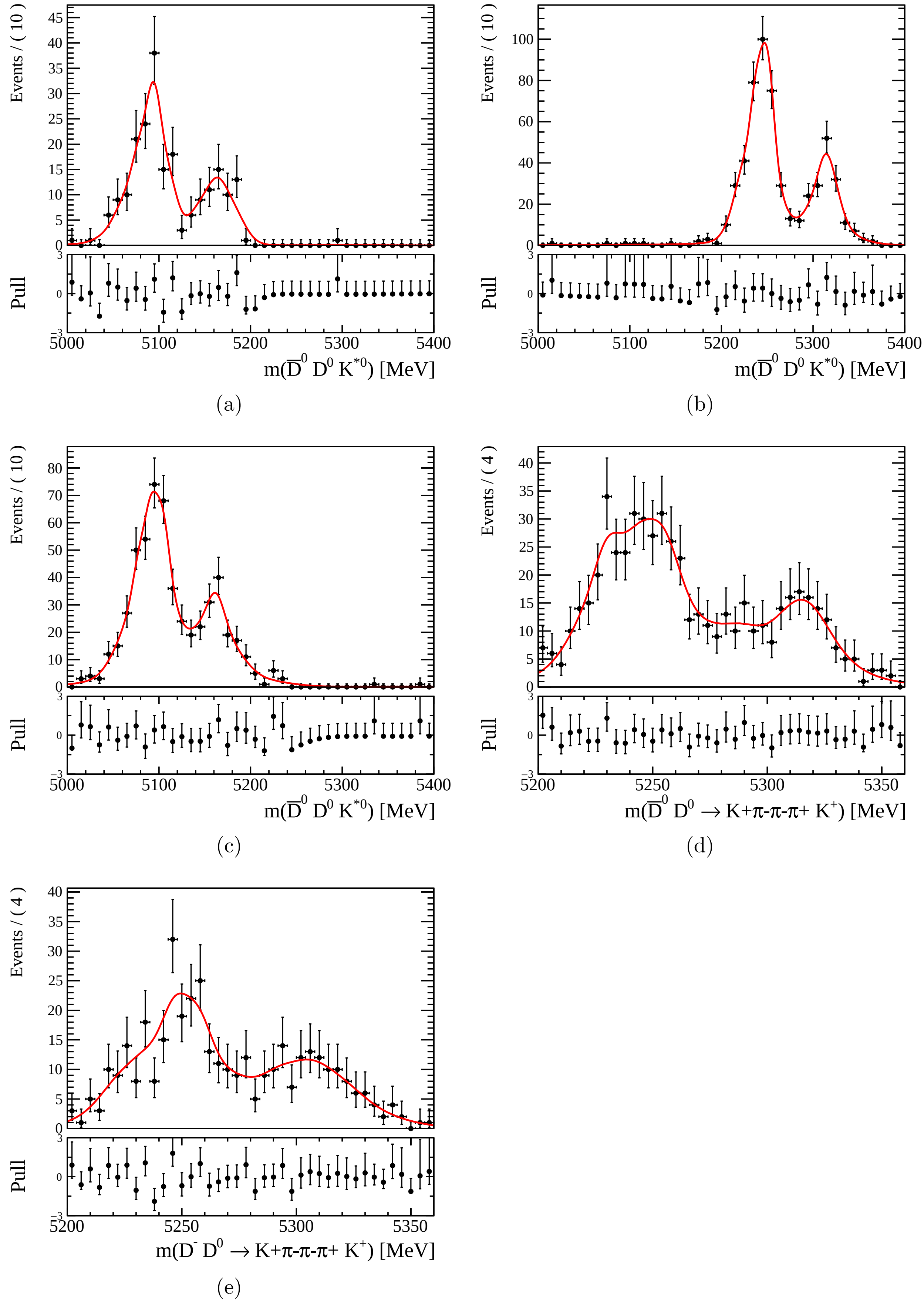


Figure 42: Non-Parametric Keys PDF on MC used in $\overline{D}^0 D^0 K^{*0}$ and $\overline{D}^0(D^0 \to K^-\pi^+\pi^+\pi^-)K^+$ and $D^-(D^0 \to K^-\pi^+\pi^+\pi^-)K^+$ for charmless estimation in in the D sideband

F Reweighting of MC Dalitz Variables

In this section we provide plots of how the reweighting scheme to the MC samples manifests in relevant variables. Plots on the left side are before the reweighting is applied. Plots on the right show the reweighted simulation both for our signal samples and ReDecay samples. The systematic uncertainty on the MC efficiencies for the mismodeling of these variables is extracted from the reweighted mc. These reweighted samples are not used in the calculation of our shapes for the simulation samples.

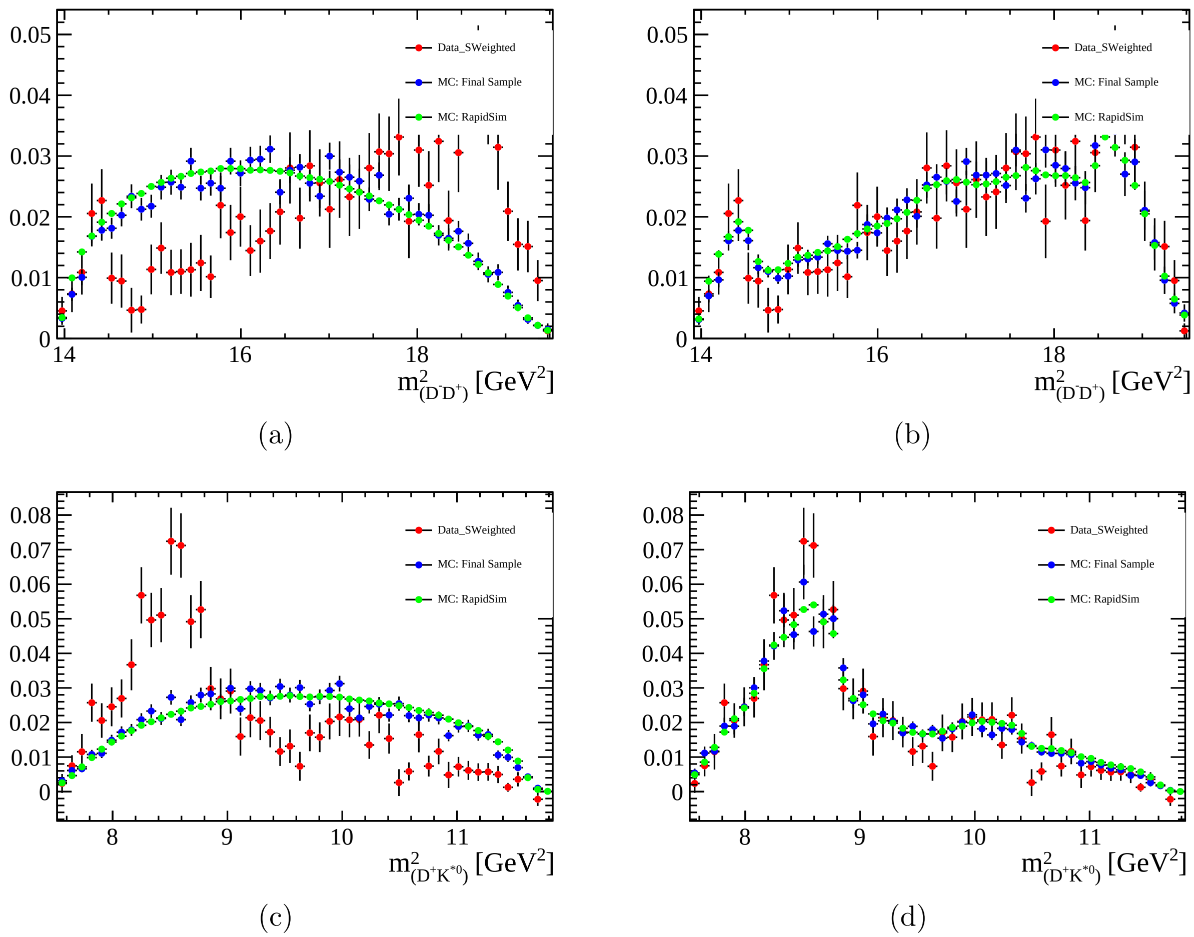


Figure 43: MC Reweighting of Simulation for 01 - $D^-D^+K^{*0}$

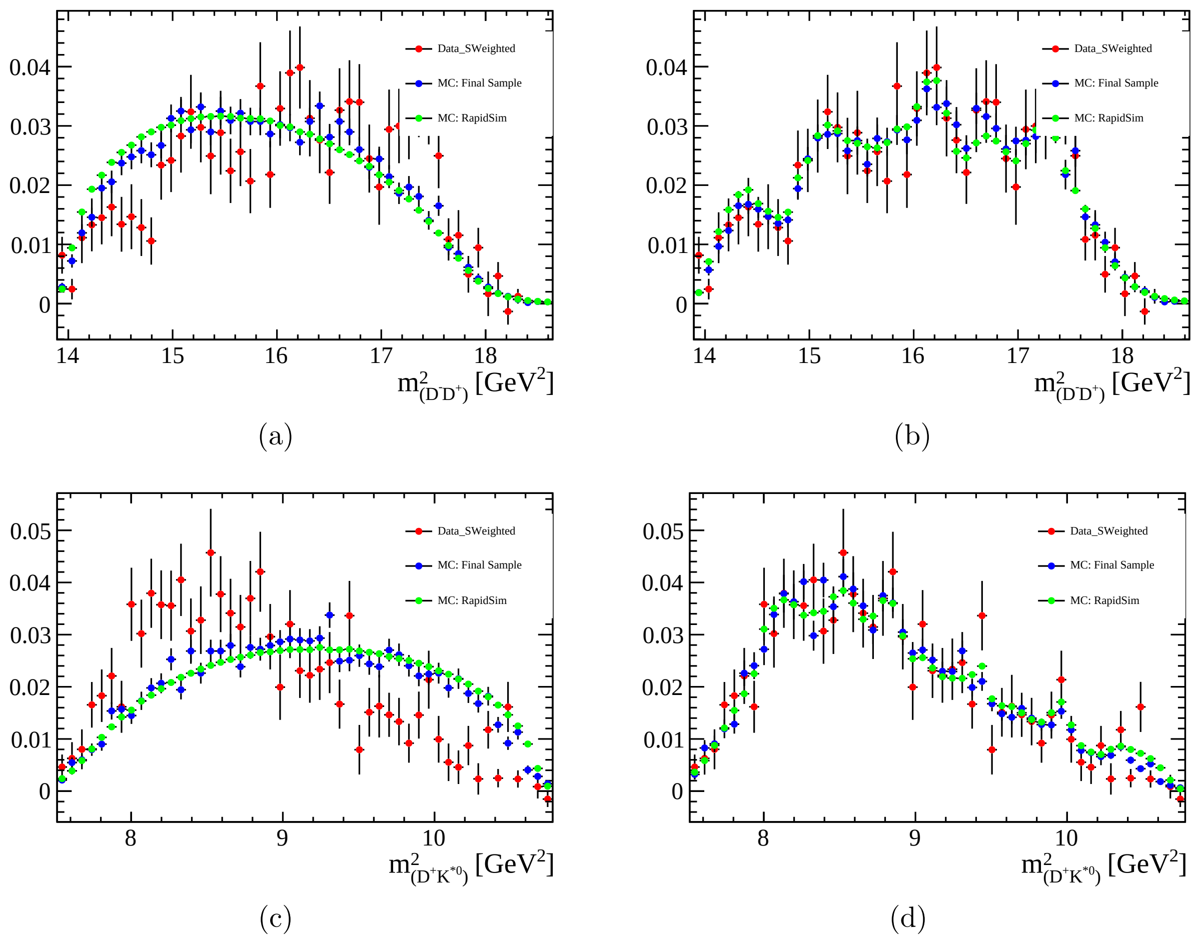


Figure 44: MC Reweighting of Simulation for 02 - $D^{*-}D^{+}K^{*0}$

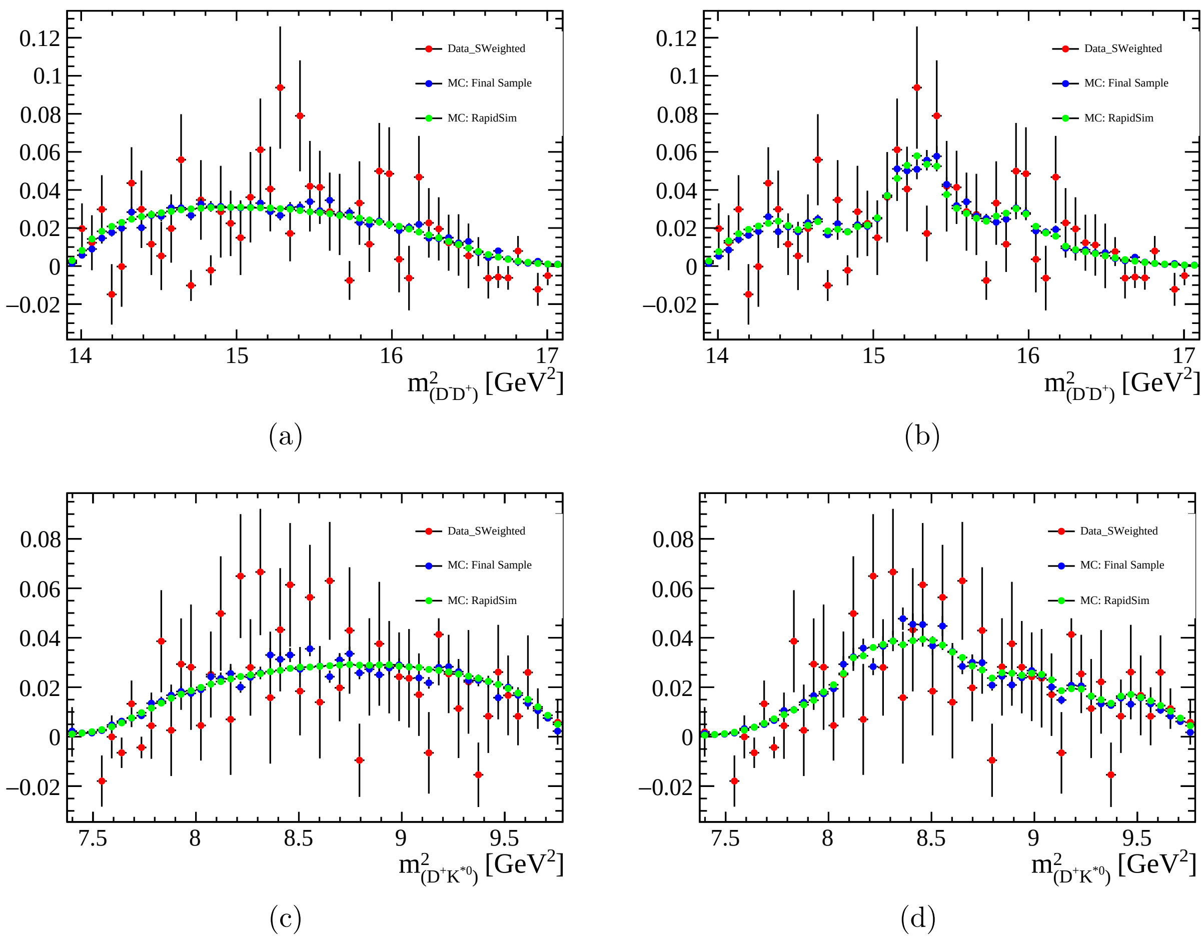
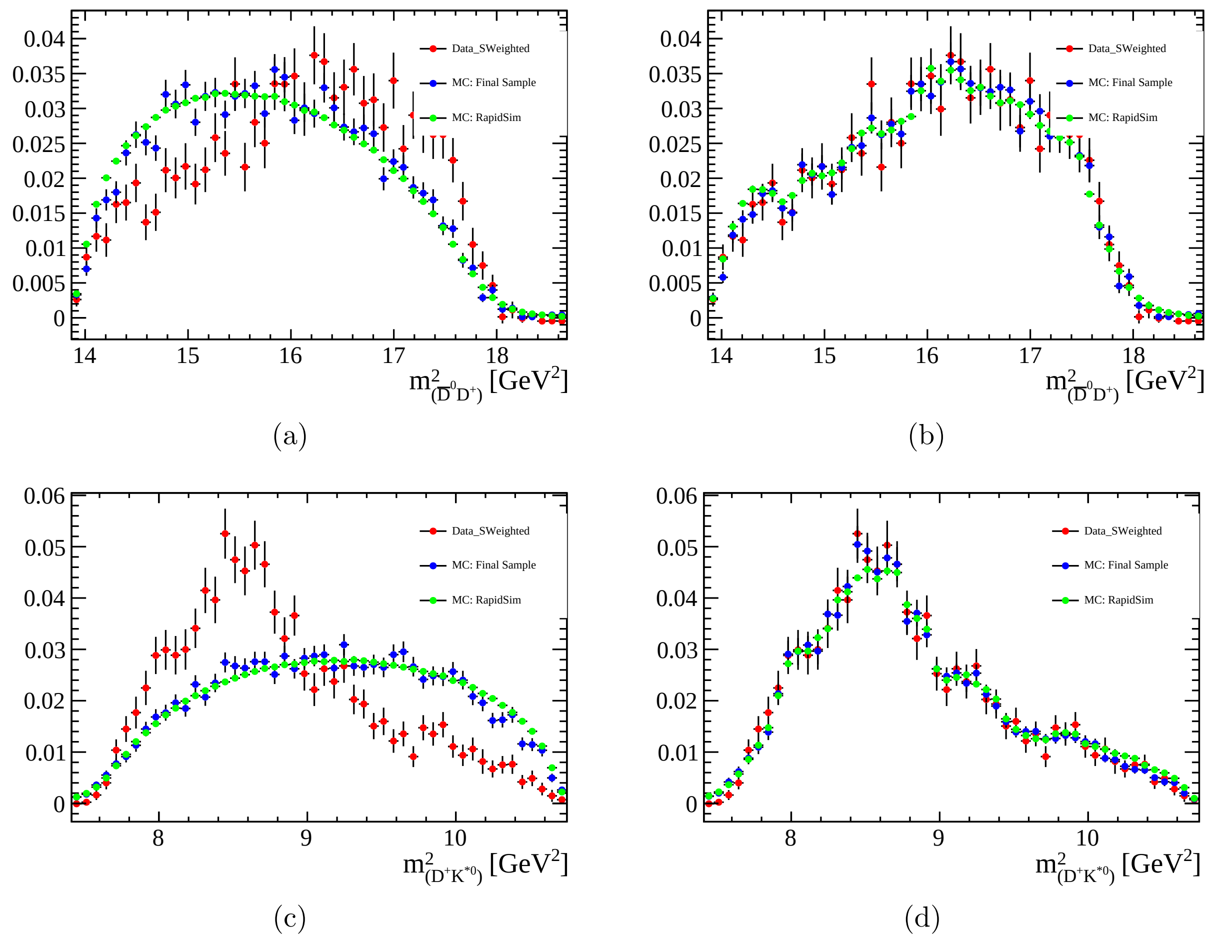


Figure 45: MC Reweighting of Simulation for 04 - $D^{*-}D^{*+}K^{*0}$



(a) (b) (c) (d)

Figure 46: MC Reweighting of Simulation for 02 - $\overline{D}^0 D^+ K^{*0}$

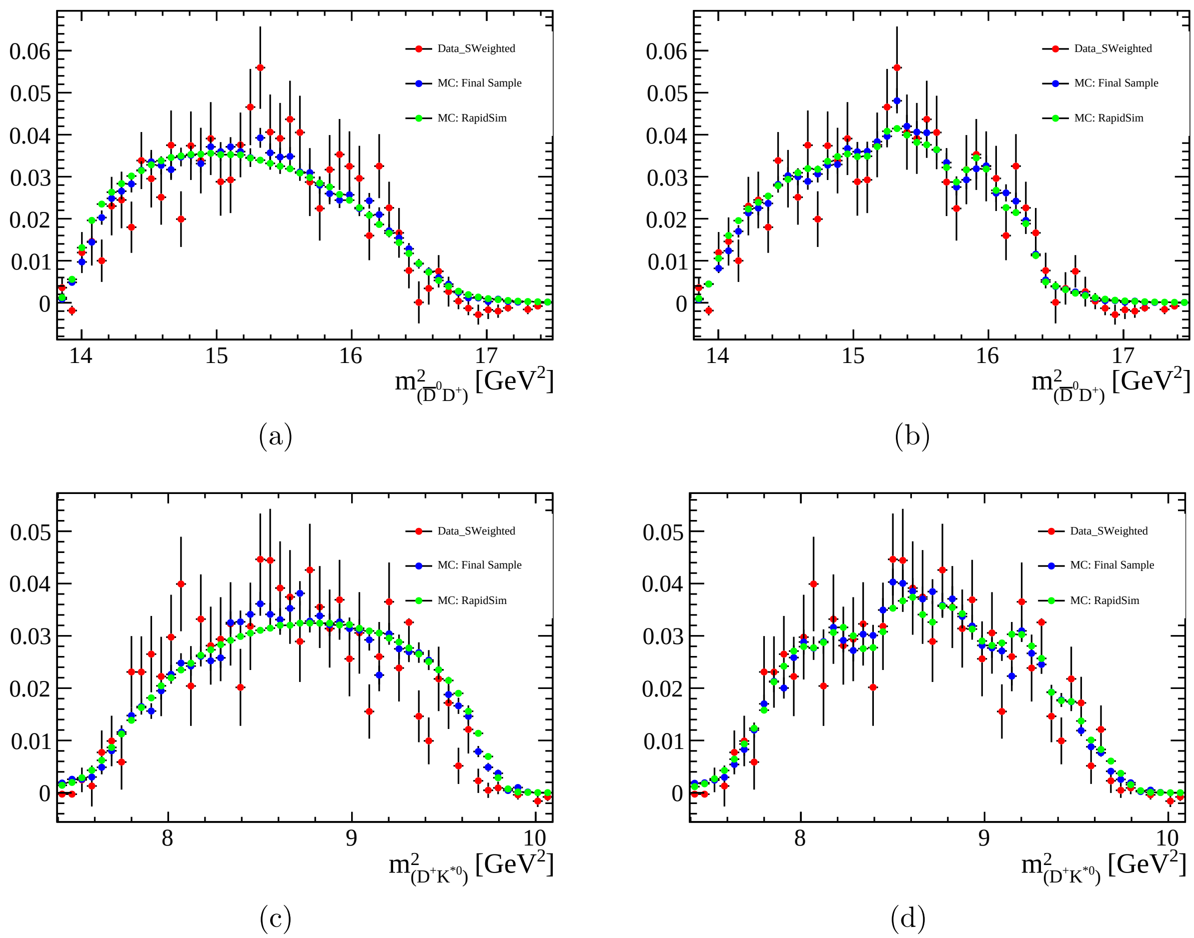


Figure 47: MC Reweighting of Simulation for 04 - $D^{*-}D^{*+}K^{*0}$

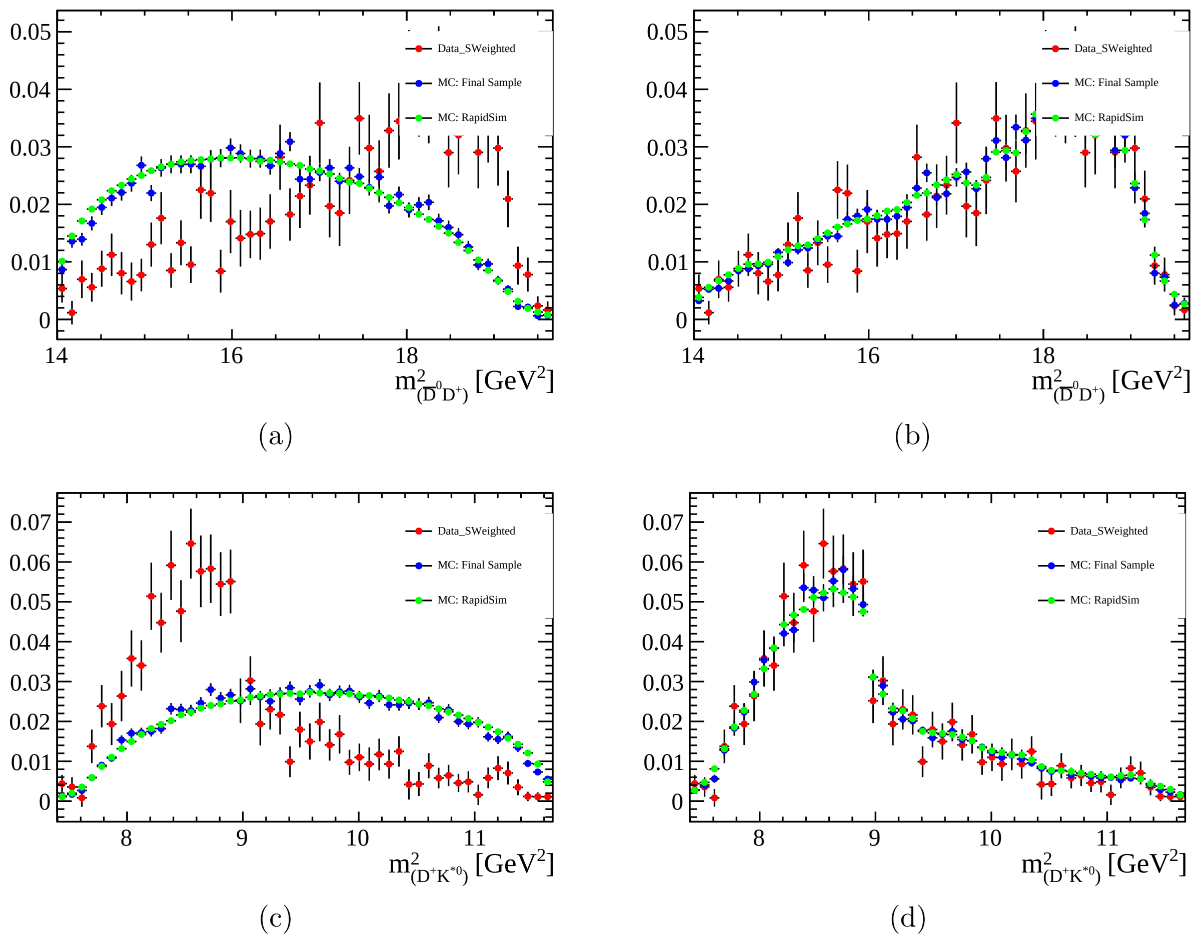


Figure 48: MC Reweighting of Simulation for 05 - $\overline{D}^0 D^+ K^{*0}$

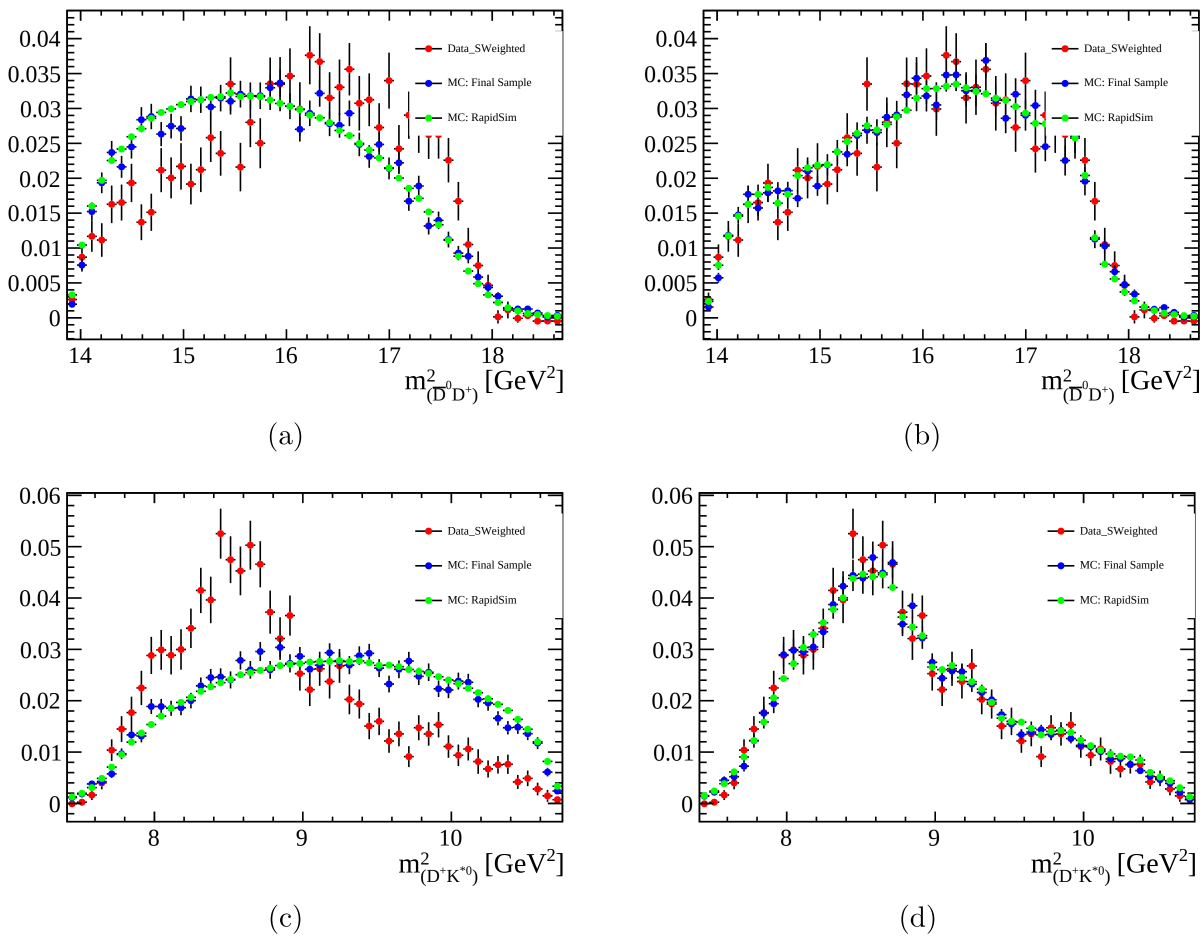


Figure 49: MC Reweighting of Simulation for 06 - $\overline{D}^0 D^+ K^{*0}$

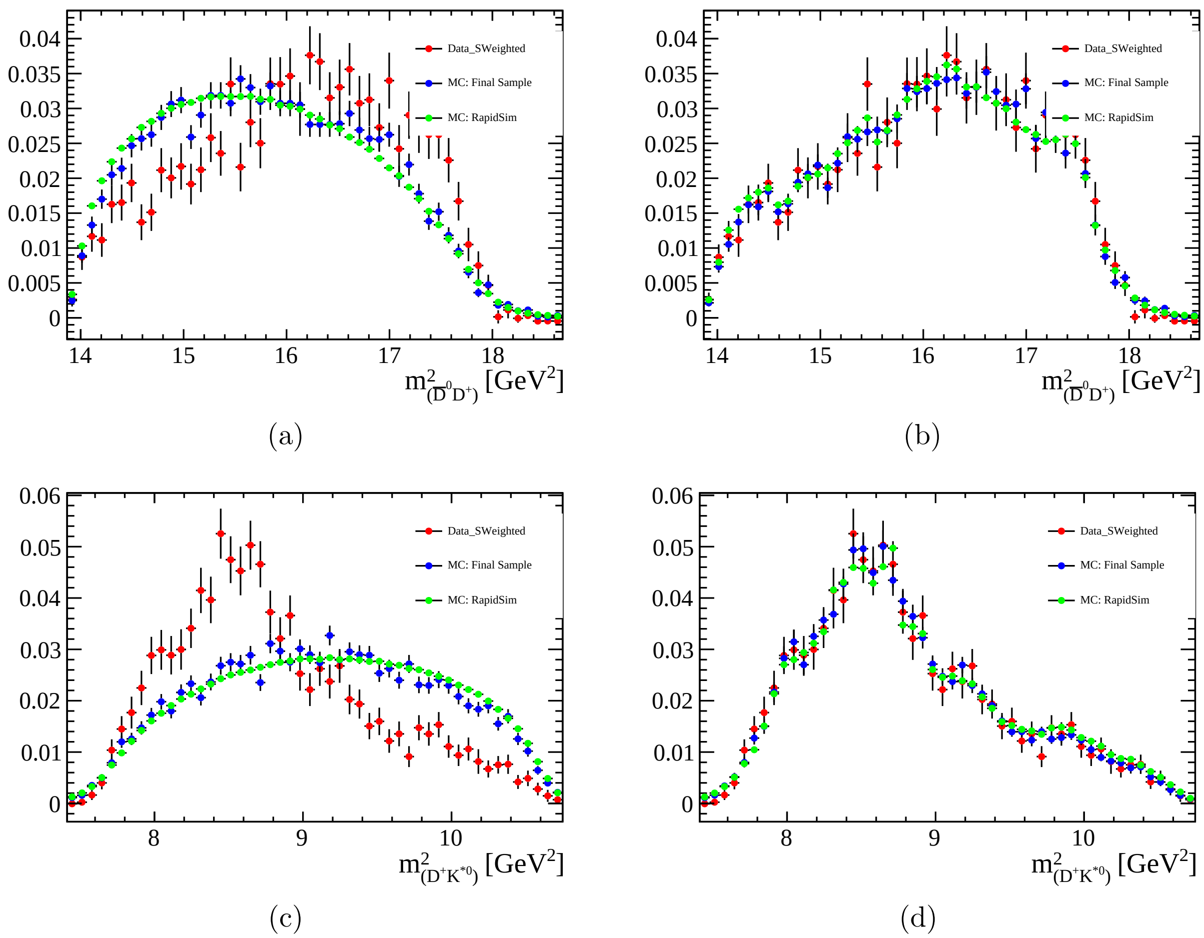


Figure 50: MC Reweighting of Simulation for 07 - $\overline{D}^0 D^+ K^{*0}$

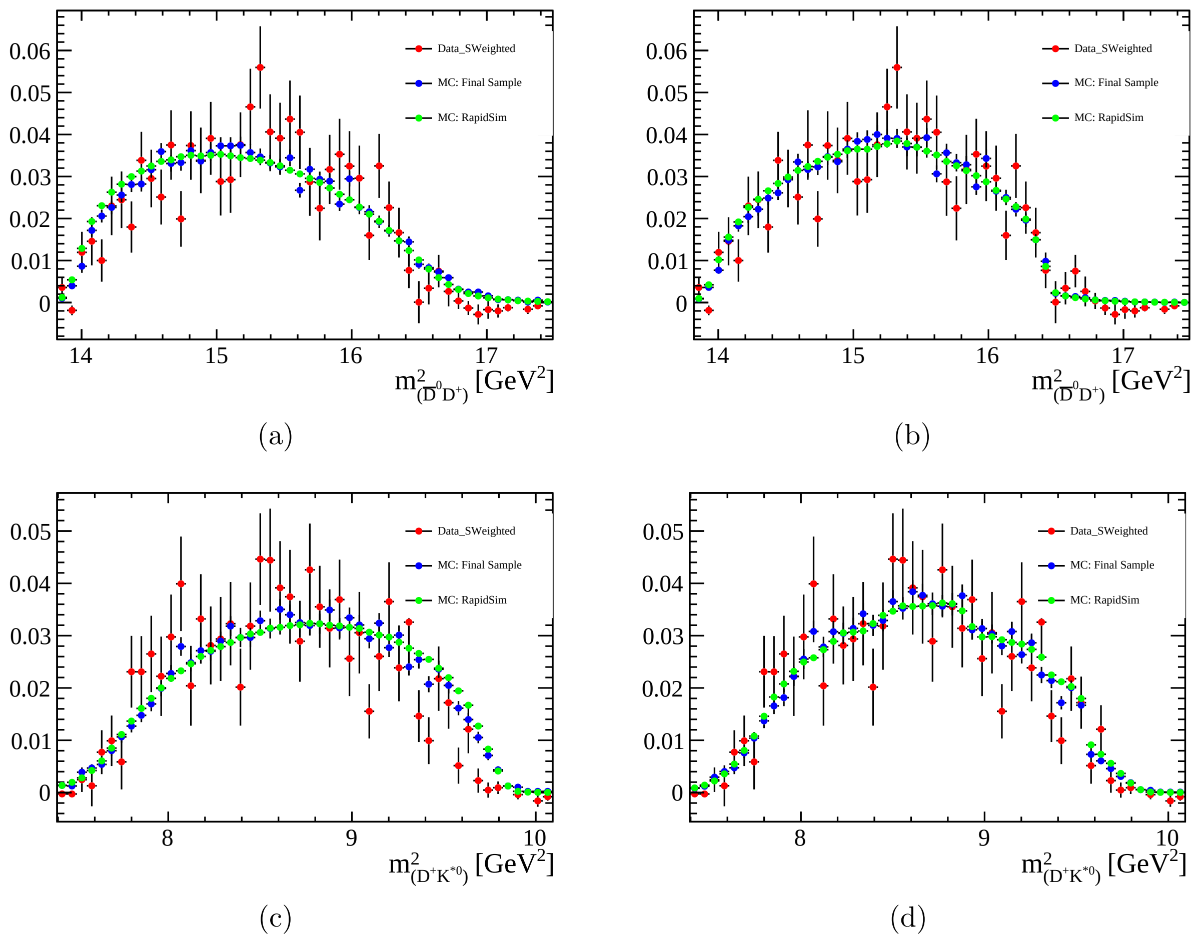


Figure 51: MC Reweighting of Simulation for 08 - $\overline{D}^0 D^+ K^{*0}$

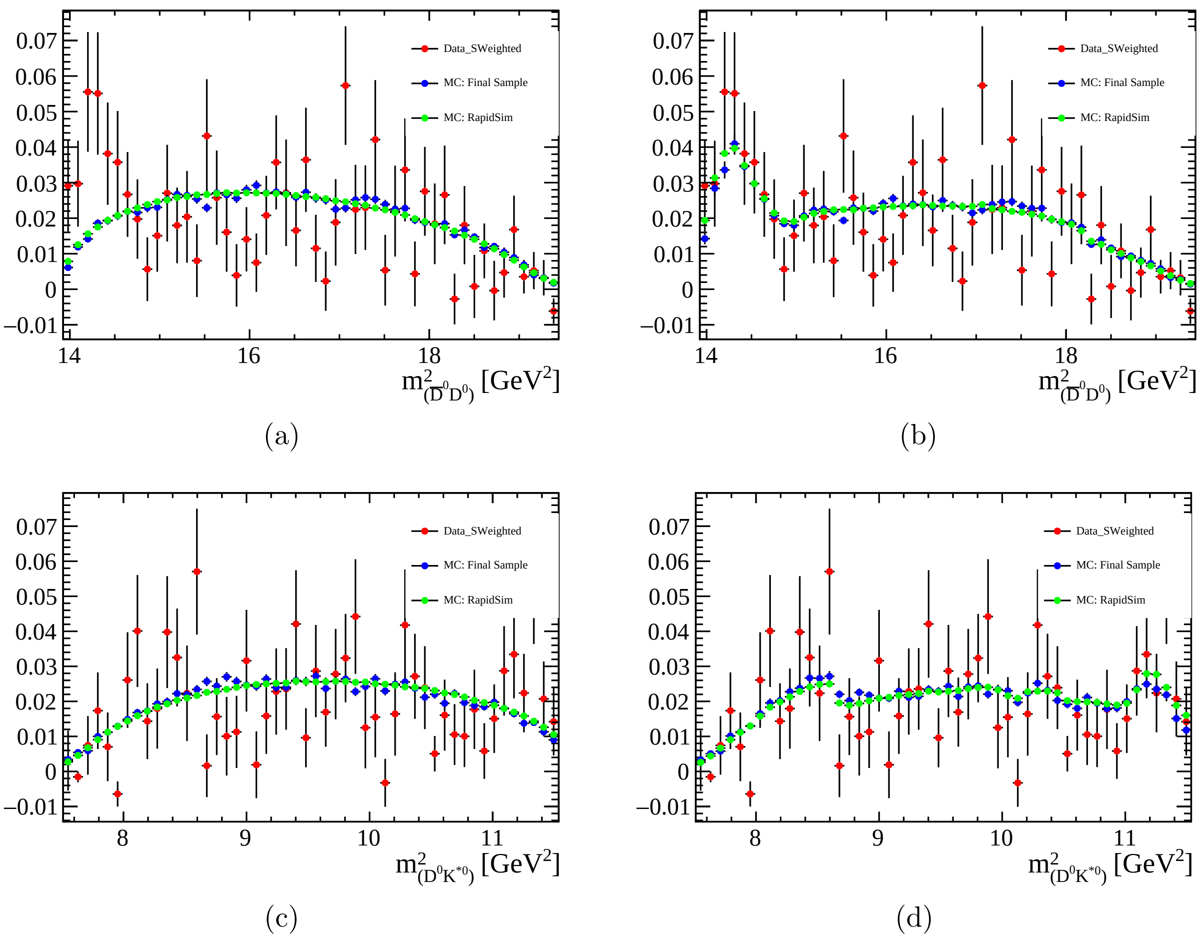


Figure 52: MC Reweighting of Simulation for 09 - $\overline{D}^0 D^0 K^{*0}$

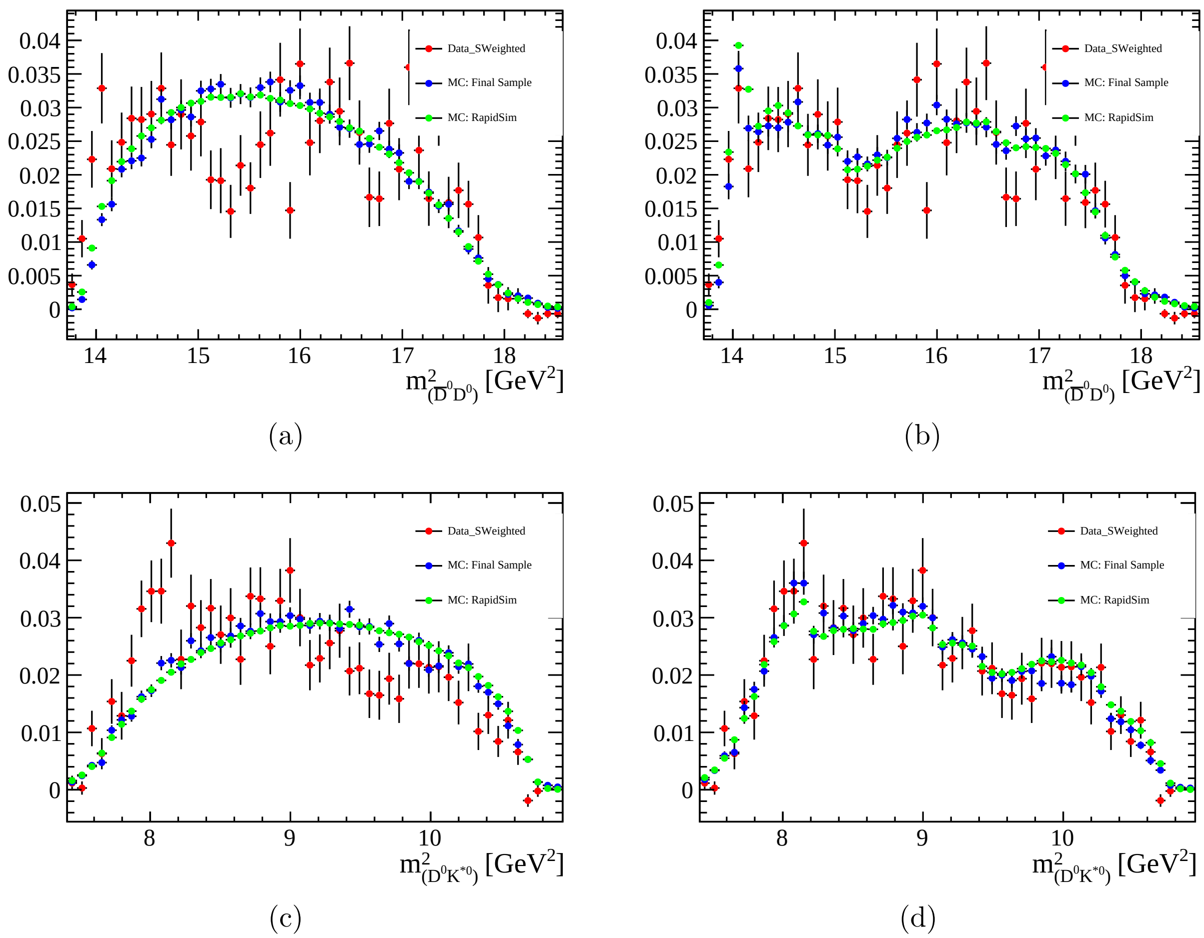


Figure 53: MC Reweighting of Simulation for 10 - $\overline{D}^{*0} D^0 K^{*0}$

G Input covariance matrix for simulaneous fit

In this section we provide covariance matrix for the values listed in table 13.

Table 29: Covariance Matrix of the 18 normalization coefficients present in our simultaneous fit

	$B^0 \to D^- D^+ K^{*0}$	$B^0 \to (D^{*-} \to D^- \pi^0) D^+ K^{*0}$	$B^0 \to (D^{*-} \to D^- \pi^0)(D^{*+} \to D^+ \pi^0) K^{*0}$
$B^0 \to D^- D^+ K^{*0}$	1.32×10^{11}	1.72×10^{1}	4.32×10^{9}
$B^0 \to (D^{*-} \to D^- \pi^0) D^+ K^{*0}$	1.72×10^{1}	1.04×10^{1}	1.26×10^{9}
$B^0 \to (D^{*-} \to D^- \pi^0)(D^{*+} \to D^+ \pi^0) K^{*0}$	4.32×10^{9}	1.26×10^{9}	6.78×10^{8}
$B^0 \to (D^{*-} \to \overline{D}^0 \pi^-)(D^{*+} \to \overline{D} \pi^+) K^{*0}$	2.25×10^{8}	6.31×10^{7}	1.59×10^{7}
$B^+ \to \overline{D}^0 (D^{*+} \to \overline{D} \pi^+) K^{*0}$	3.69×10^{8}	1.04×10^{8}	2.61×10^{7}
$B^+ \to \overline{D}^{*0} (D^{*+} \to \overline{D} \pi^+) K^{*0}$	3.52×10^{8}	9.89×10^{7}	2.49×10^{7}
$B^0 \to \overline{D}^0 D^0 K^{*0}$	7.40×10^{8}	2.08×10^{8}	5.23×10^{7}
$B^0 \to \overline{D}^{*0} D^0 K^{*0} + B^0 \to \overline{D}^0 D^{*0} K^{*0}$	7.01×10^{8}	1.97×10^{8}	4.95×10^{7}
$B^0 \to \overline{D}^{*0} D^{*0} K^{*0}$	6.25×10^{8}	1.76×10^{8}	4.42×10^{7}
$B^0 \to (D^{*-} \to \overline{D}^0 \pi^-) D^+ K^{*0}$	9.99×10^{8}	2.81×10^{8}	7.06×10^{7}
$B^0 \to (D^{*-} \to \overline{D}^0 \pi^-)(D^{*+} \to D^+ \pi^0) K^{*0}$	2.76×10^{8}	1.09×10^{8}	3.55×10^{7}
$B^+ \to \overline{D}^0 D^+ K^{*0}$	1.81×10^{9}	5.07×10^{8}	1.28×10^{8}
$B^+ \to \overline{D}^{*0} D^+ K^{*0}$	1.57×10^{9}	4.40×10^{8}	1.11×10^{8}
$B^+ \to \overline{D}^0 (D^{*+} \to D^+ \pi^0) K^{*0}$	5.03×10^{8}	2.00×10^{8}	6.49×10^{7}
$B^+ \to \overline{D}^{*0} (D^{*+} \to D^+ \pi^0) K^{*0}$	4.65×10^{8}	1.84×10^{8}	6.00×10^{7}
$B^0 \to (D^{*-} \to \overline{D}^0 \pi^-)(D^{*+} \to \overline{D} \pi^+) K^{*0}$	5.30×10^{7}	1.49×10^{7}	3.75×10^{6}
$B^+ \to \overline{D}^0 (D^{*+} \to \overline{D} \pi^+) K^{*0}$	1.02×10^{8}	2.86×10^{7}	7.20×10^{6}
$B^+ \to \overline{D}^{*0} (D^{*+} \to \overline{D} \pi^+) K^{*0}$	8.56×10^{7}	2.41×10^{7}	6.05×10^{6}

Table 30: Covariance Matrix of the 18 normalization coefficients present in our simultaneous fit

	$B^0 \to (D^{*-} \to \overline{D}^0\pi^-)(D^{*+} \to \overline{D}\pi^+)K^{*0}$	$B^+ \to \overline{D}^0(D^{*+} \to \overline{D}\pi^+)K^{*0}$	$B^+ \to \overline{D}^{*0}(D^{*+} \to \overline{D}\pi^+)K^{*0}$
$B^0 \to D^-D^+K^{*0}$	2.25×10^8	3.69×10^8	3.52×10^8
$B^0 \to (D^{*-} \to D^-\pi^0)D^+K^{*0}$	6.31×10^7	1.04×10^8	9.89×10^7
$B^0 \to (D^{*-} \to D^-\pi^0)(D^{*+} \to D^+\pi^0)K^{*0}$	1.59×10^7	2.61×10^7	2.49×10^7
$B^0 \to (D^{*-} \to \overline{D}^0\pi^-)(D^{*+} \to \overline{D}\pi^+)K^{*0}$	3.03×10^9	2.62×10^9	2.50×10^9
$B^+ \to \overline{D}^0(D^{*+} \to \overline{D}\pi^+)K^{*0}$	2.62×10^9	8.14×10^9	4.09×10^9
$B^+ \to \overline{D}^{*0}(D^{*+} \to \overline{D}\pi^+)K^{*0}$	2.50×10^9	4.09×10^9	7.37×10^9
$B^0 \to \overline{D}^0D^0K^{*0}$	5.20×10^9	8.55×10^9	8.16×10^9
$B^0 \to \overline{D}^{*0}D^0K^{*0} + B^0 \to \overline{D}^0D^{*0}K^{*0}$	4.93×10^9	8.10×10^9	7.74×10^9
$B^0 \to \overline{D}^{*0}D^{*0}K^{*0}$	4.40×10^9	7.22×10^9	6.90×10^9
$B^0 \to (D^{*-} \to \overline{D}^0\pi^-)D^+K^{*0}$	3.52×10^9	5.75×10^9	5.50×10^9
$B^0 \to (D^{*-} \to \overline{D}^0\pi^-)(D^{*+} \to D^+\pi^0)K^{*0}$	9.70×10^8	1.59×10^9	1.52×10^9
$B^+ \to \overline{D}^0D^+K^{*0}$	6.30×10^9	1.04×10^1	9.88×10^9
$B^+ \to \overline{D}^{*0}D^+K^{*0}$	5.47×10^9	8.98×10^9	8.58×10^9
$B^+ \to \overline{D}^0(D^{*+} \to D^+\pi^0)K^{*0}$	1.76×10^9	2.89×10^9	2.76×10^9
$B^+ \to \overline{D}^{*0}(D^{*+} \to D^+\pi^0)K^{*0}$	1.62×10^9	2.67×10^9	2.55×10^9
$B^0 \to (D^{*-} \to \overline{D}^0\pi^-)(D^{*+} \to \overline{D}\pi^+)K^{*0}$	3.79×10^8	6.18×10^8	5.89×10^8
$B^+ \to \overline{D}^0(D^{*+} \to \overline{D}\pi^+)K^{*0}$	7.22×10^8	1.18×10^9	1.13×10^9
$B^+ \to \overline{D}^{*0}(D^{*+} \to \overline{D}\pi^+)K^{*0}$	6.07×10^8	9.93×10^8	9.48×10^8

Table 31: Covariance Matrix of the 18 normalization coefficients present in our simultaneous fit

	$B^0 \to \overline{D}^0 D^0 K^{*0}$	$B^0 \to \overline{D}^{*0} D^0 K^{*0} + B^0 \to \overline{D}^0 D^{*0} K^{*0}$	$B^0 \to \overline{D}^{*0} D^{*0} K^{*0}$
$B^0 \to D^- D^+ K^{*0}$	7.40×10^8	7.01×10^8	6.25×10^8
$B^0 \to (D^{*-} \to D^- \pi^0) D^+ K^{*0}$	2.08×10^8	1.97×10^8	1.76×10^8
$B^0 \to (D^{*-} \to D^- \pi^0)(D^{*+} \to D^+ \pi^0) K^{*0}$	5.23×10^7	4.95×10^7	4.42×10^7
$B^0 \to (D^{*-} \to \overline{D}^0 \pi^-)(D^{*+} \to \overline{D} \pi^+) K^{*0}$	5.20×10^9	4.93×10^9	4.40×10^9
$B^+ \to \overline{D}^0 (D^{*+} \to \overline{D} \pi^+) K^{*0}$	8.55×10^9	8.10×10^9	7.22×10^9
$B^+ \to \overline{D}^{*0} (D^{*+} \to \overline{D} \pi^+) K^{*0}$	8.16×10^9	7.74×10^9	6.90×10^9
$B^0 \to \overline{D}^0 D^0 K^{*0}$	3.20×10^1	1.63×10^1	1.45×10^1
$B^0 \to \overline{D}^{*0} D^0 K^{*0} + B^0 \to \overline{D}^0 D^{*0} K^{*0}$	1.63×10^1	2.87×10^1	1.37×10^1
$B^0 \to \overline{D}^{*0} D^{*0} K^{*0}$	1.45×10^1	1.37×10^1	2.30×10^1
$B^0 \to (D^{*-} \to \overline{D}^0 \pi^-) D^+ K^{*0}$	1.15×10^1	1.09×10^1	9.71×10^9
$B^0 \to (D^{*-} \to \overline{D}^0 \pi^-)(D^{*+} \to D^+ \pi^0) K^{*0}$	3.17×10^9	3.00×10^9	2.68×10^9
$B^+ \to \overline{D}^0 D^+ K^{*0}$	2.08×10^1	1.97×10^1	1.75×10^1
$B^+ \to \overline{D}^{*0} D^+ K^{*0}$	1.80×10^1	1.71×10^1	1.52×10^1
$B^+ \to \overline{D}^0 (D^{*+} \to D^+ \pi^0) K^{*0}$	5.79×10^9	5.48×10^9	4.89×10^9
$B^+ \to \overline{D}^{*0} (D^{*+} \to D^+ \pi^0) K^{*0}$	5.35×10^9	5.07×10^9	4.52×10^9
$B^0 \to (D^{*-} \to \overline{D}^0 \pi^-)(D^{*+} \to \overline{D} \pi^+) K^{*0}$	1.23×10^9	1.17×10^9	1.04×10^9
$B^+ \to \overline{D}^0 (D^{*+} \to \overline{D} \pi^+) K^{*0}$	2.36×10^9	2.24×10^9	2.00×10^9
$B^+ \to \overline{D}^{*0} (D^{*+} \to \overline{D} \pi^+) K^{*0}$	1.98×10^9	1.88×10^9	1.68×10^9

Table 32: Covariance Matrix of the 18 normalization coefficients present in our simultaneous fit

	$B^0 \to (D^{*-} \to \overline{D}^0\pi^-)D^+K^{*0}$	$B^0 \to (D^{*-} \to \overline{D}^0\pi^-)(D^{*+} \to D^+\pi^0)K^{*0}$	$B^+ \to \overline{D}^0D^+K^{*0}$
$B^0 \to D^-D^+K^{*0}$	9.99×10^8	2.76×10^8	1.81×10^9
$B^0 \to (D^{*-} \to D^-\pi^0)D^+K^{*0}$	2.81×10^8	1.09×10^8	5.07×10^8
$B^0 \to (D^{*-} \to D^-\pi^0)(D^{*+} \to D^+\pi^0)K^{*0}$	7.06×10^7	3.55×10^7	1.28×10^8
$B^0 \to (D^{*-} \to \overline{D}^0\pi^-)(D^{*+} \to \overline{D}\pi^+)K^{*0}$	3.52×10^9	9.70×10^8	6.30×10^9
$B^+ \to \overline{D}^0(D^{*+} \to \overline{D}\pi^+)K^{*0}$	5.75×10^9	1.59×10^9	1.04×10^1
$B^+ \to \overline{D}^{*0}(D^{*+} \to \overline{D}\pi^+)K^{*0}$	5.50×10^9	1.52×10^9	9.88×10^9
$B^0 \to \overline{D}^0D^0K^{*0}$	1.15×10^1	3.17×10^9	2.08×10^1
$B^0 \to \overline{D}^{*0}D^0K^{*0} + B^0 \to \overline{D}^0D^{*0}K^{*0}$	1.09×10^1	3.00×10^9	1.97×10^1
$B^0 \to \overline{D}^{*0}D^{*0}K^{*0}$	9.71×10^9	2.68×10^9	1.75×10^1
$B^0 \to (D^{*-} \to \overline{D}^0\pi^-)D^+K^{*0}$	1.45×10^1	2.19×10^9	1.43×10^1
$B^0 \to (D^{*-} \to \overline{D}^0\pi^-)(D^{*+} \to D^+\pi^0)K^{*0}$	2.19×10^9	1.11×10^9	3.94×10^9
$B^+ \to \overline{D}^0D^+K^{*0}$	1.43×10^1	3.94×10^9	4.79×10^1
$B^+ \to \overline{D}^{*0}D^+K^{*0}$	1.24×10^1	3.42×10^9	2.24×10^1
$B^+ \to \overline{D}^0(D^{*+} \to D^+\pi^0)K^{*0}$	3.98×10^9	1.12×10^9	7.20×10^9
$B^+ \to \overline{D}^{*0}(D^{*+} \to D^+\pi^0)K^{*0}$	3.68×10^9	1.03×10^9	6.65×10^9
$B^0 \to (D^{*-} \to \overline{D}^0\pi^-)(D^{*+} \to \overline{D}\pi^+)K^{*0}$	8.30×10^8	2.29×10^8	1.49×10^9
$B^+ \to \overline{D}^0(D^{*+} \to \overline{D}\pi^+)K^{*0}$	1.59×10^9	4.38×10^8	2.86×10^9
$B^+ \to \overline{D}^{*0}(D^{*+} \to \overline{D}\pi^+)K^{*0}$	1.33×10^9	3.68×10^8	2.40×10^9

Table 33: Covariance Matrix of the 18 normalization coefficients present in our simultaneous fit

	$B^+ \to \overline{D}^{*0} D^+ K^{*0}$	$B^+ \to \overline{D}^0 (D^{*+} \to D^+ \pi^0) K^{*0}$	$B^+ \to \overline{D}^{*0} (D^{*+} \to D^+ \pi^0) K^{*0}$
$B^0 \to D^- D^+ K^{*0}$	1.57×10^9	5.03×10^8	4.65×10^8
$B^0 \to (D^{*-} \to D^- \pi^0) D^+ K^{*0}$	4.40×10^8	2.00×10^8	1.84×10^8
$B^0 \to (D^{*-} \to D^- \pi^0)(D^{*+} \to D^+ \pi^0) K^{*0}$	1.11×10^8	6.49×10^7	6.00×10^7
$B^0 \to (D^{*-} \to \overline{D}^0 \pi^-)(D^{*+} \to \overline{D} \pi^+) K^{*0}$	5.47×10^9	1.76×10^9	1.62×10^9
$B^+ \to \overline{D}^0 (D^{*+} \to \overline{D} \pi^+) K^{*0}$	8.98×10^9	2.89×10^9	2.67×10^9
$B^+ \to \overline{D}^{*0} (D^{*+} \to \overline{D} \pi^+) K^{*0}$	8.58×10^9	2.76×10^9	2.55×10^9
$B^0 \to \overline{D}^0 D^0 K^{*0}$	1.80×10^1	5.79×10^9	5.35×10^9
$B^0 \to \overline{D}^{*0} D^0 K^{*0} + B^0 \to \overline{D}^0 D^{*0} K^{*0}$	1.71×10^1	5.48×10^9	5.07×10^9
$B^0 \to \overline{D}^{*0} D^{*0} K^{*0}$	1.52×10^1	4.89×10^9	4.52×10^9
$B^0 \to (D^{*-} \to \overline{D}^0 \pi^-) D^+ K^{*0}$	1.24×10^1	3.98×10^9	3.68×10^9
$B^0 \to (D^{*-} \to \overline{D}^0 \pi^-)(D^{*+} \to D^+ \pi^0) K^{*0}$	3.42×10^9	1.12×10^9	1.03×10^9
$B^+ \to \overline{D}^0 D^+ K^{*0}$	2.24×10^1	7.20×10^9	6.65×10^9
$B^+ \to \overline{D}^{*0} D^+ K^{*0}$	3.56×10^1	6.25×10^9	5.77×10^9
$B^+ \to \overline{D}^0 (D^{*+} \to D^+ \pi^0) K^{*0}$	6.25×10^9	3.72×10^9	1.89×10^9
$B^+ \to \overline{D}^{*0} (D^{*+} \to D^+ \pi^0) K^{*0}$	5.77×10^9	1.89×10^9	3.15×10^9
$B^0 \to (D^{*-} \to \overline{D}^0 \pi^-)(D^{*+} \to \overline{D} \pi^+) K^{*0}$	1.29×10^9	4.15×10^8	3.83×10^8
$B^+ \to \overline{D}^0 (D^{*+} \to \overline{D} \pi^+) K^{*0}$	2.48×10^9	7.97×10^8	7.36×10^8
$B^+ \to \overline{D}^{*0} (D^{*+} \to \overline{D} \pi^+) K^{*0}$	2.09×10^9	6.69×10^8	6.19×10^8

Table 34: Covariance Matrix of the 18 normalization coefficients present in our simultaneous fit

	$B^0 \to (D^{*-} \to \overline{D}^0\pi^-)(D^{*+} \to \overline{D}\pi^+)K^{*0}$	$B^+ \to \overline{D}^0(D^{*+} \to \overline{D}\pi^+)K^{*0}$	$B^+ \to \overline{D}^{*0}(D^{*+} \to \overline{D}\pi^+)K^{*0}$
$B^0 \to D^-D^+K^{*0}$	5.30×10^7	1.02×10^8	8.56×10^7
$B^0 \to (D^{*-} \to D^-\pi^0)D^+K^{*0}$	1.49×10^7	2.86×10^7	2.41×10^7
$B^0 \to (D^{*-} \to D^-\pi^0)(D^{*+} \to D^+\pi^0)K^{*0}$	3.75×10^6	7.20×10^6	6.05×10^6
$B^0 \to (D^{*-} \to \overline{D}^0\pi^-)(D^{*+} \to \overline{D}\pi^+)K^{*0}$	3.79×10^8	7.22×10^8	6.07×10^8
$B^+ \to \overline{D}^0(D^{*+} \to \overline{D}\pi^+)K^{*0}$	6.18×10^8	1.18×10^9	9.93×10^8
$B^+ \to \overline{D}^{*0}(D^{*+} \to \overline{D}\pi^+)K^{*0}$	5.89×10^8	1.13×10^9	9.48×10^8
$B^0 \to \overline{D}^0D^0K^{*0}$	1.23×10^9	2.36×10^9	1.98×10^9
$B^0 \to \overline{D}^{*0}D^0K^{*0} + B^0 \to \overline{D}^0D^{*0}K^{*0}$	1.17×10^9	2.24×10^9	1.88×10^9
$B^0 \to \overline{D}^{*0}D^{*0}K^{*0}$	1.04×10^9	2.00×10^9	1.68×10^9
$B^0 \to (D^{*-} \to \overline{D}^0\pi^-)D^+K^{*0}$	8.30×10^8	1.59×10^9	1.33×10^9
$B^0 \to (D^{*-} \to \overline{D}^0\pi^-)(D^{*+} \to D^+\pi^0)K^{*0}$	2.29×10^8	4.38×10^8	3.68×10^8
$B^+ \to \overline{D}^0D^+K^{*0}$	1.49×10^9	2.86×10^9	2.40×10^9
$B^+ \to \overline{D}^{*0}D^+K^{*0}$	1.29×10^9	2.48×10^9	2.09×10^9
$B^+ \to \overline{D}^0(D^{*+} \to D^+\pi^0)K^{*0}$	4.15×10^8	7.97×10^8	6.69×10^8
$B^+ \to \overline{D}^{*0}(D^{*+} \to D^+\pi^0)K^{*0}$	3.83×10^8	7.36×10^8	6.19×10^8
$B^0 \to (D^{*-} \to \overline{D}^0\pi^-)(D^{*+} \to \overline{D}\pi^+)K^{*0}$	1.73×10^8	1.71×10^8	1.44×10^8
$B^+ \to \overline{D}^0(D^{*+} \to \overline{D}\pi^+)K^{*0}$	1.71×10^8	6.31×10^8	2.75×10^8
$B^+ \to \overline{D}^{*0}(D^{*+} \to \overline{D}\pi^+)K^{*0}$	1.44×10^8	2.75×10^8	4.48×10^8

H Simultaneous Fit: No Systematics

Here we present plots of the simultaneous fit where we constrain ourselves to the nominal values of our uncertainties, and do not incorporate any correlations in the form of our nuisance parameter vectors.

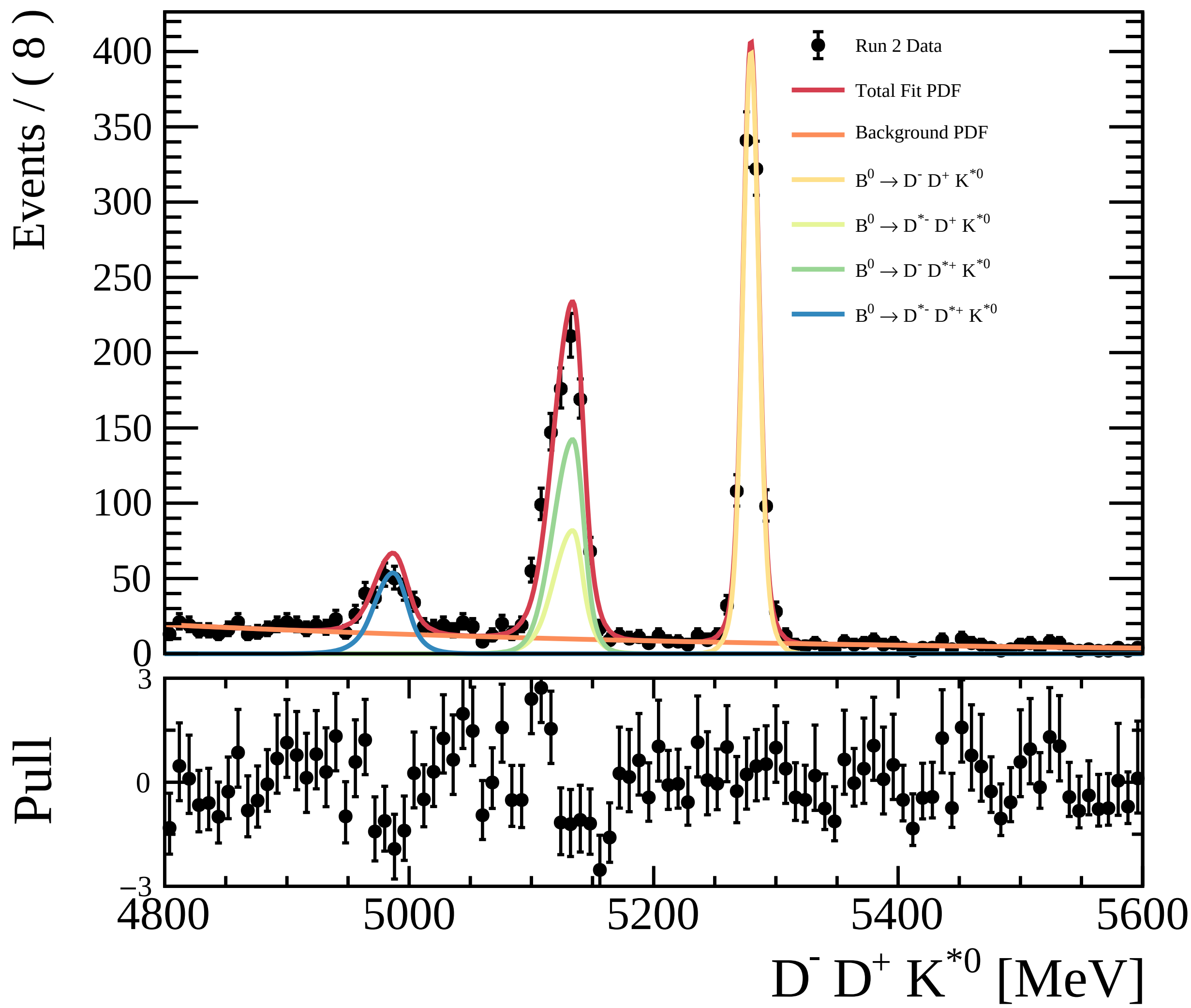


Figure 54: The final simultaneous fit to the $D^-D^+K^{*0}$ spectrum with no uncertainties or correlations Incorporated into the fit.

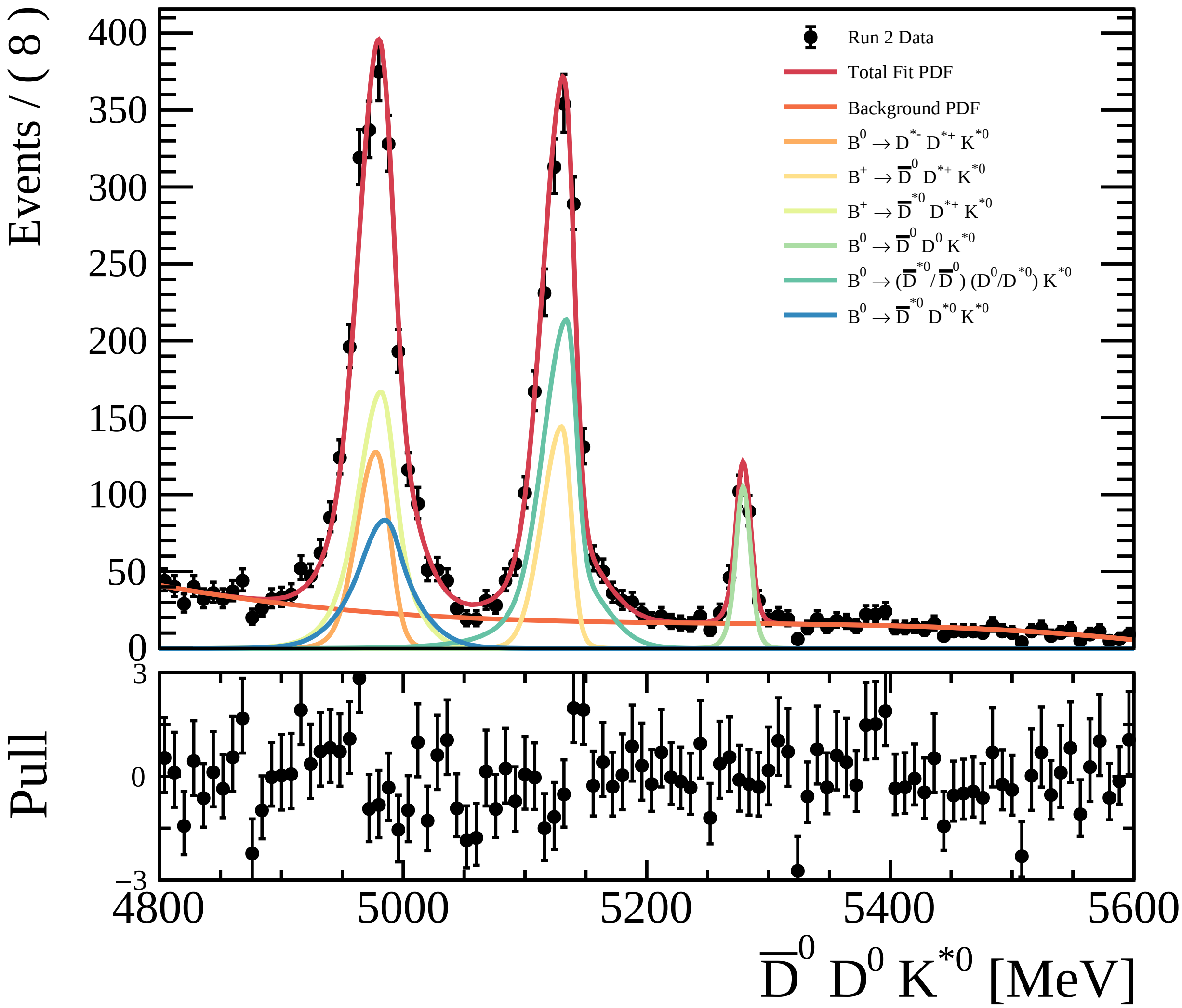


Figure 55: The final simultaneous fit to the $\overline{D}^0 D^0 K^{*0}$ spectrum with no uncertainties or correlations Incorporated into the fit.

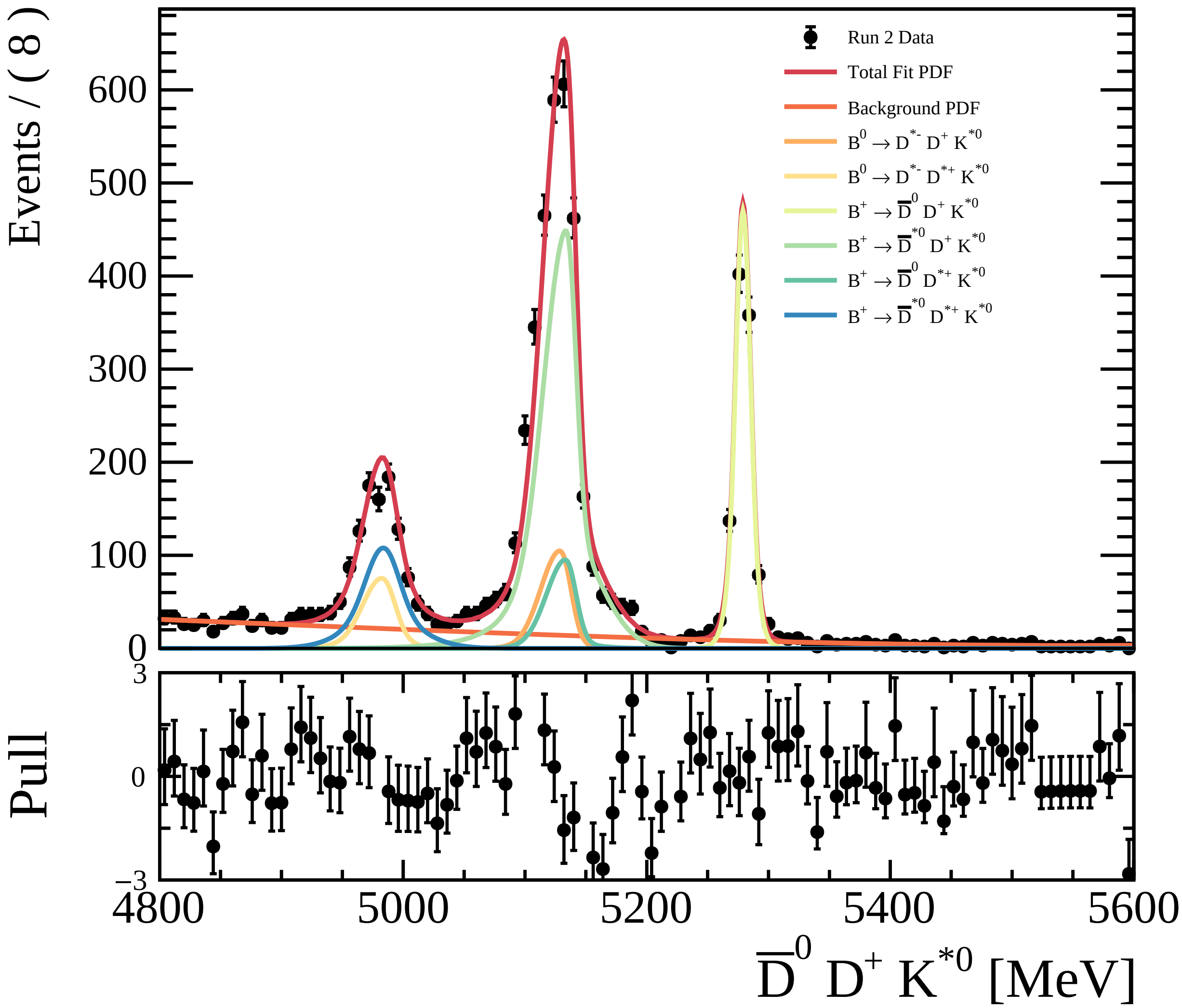


Figure 56: The final simultaneous fit to the $\overline{D}^0 D^+ K^{*0}$ spectrum with no uncertainties or correlations Incorporated into the fit.

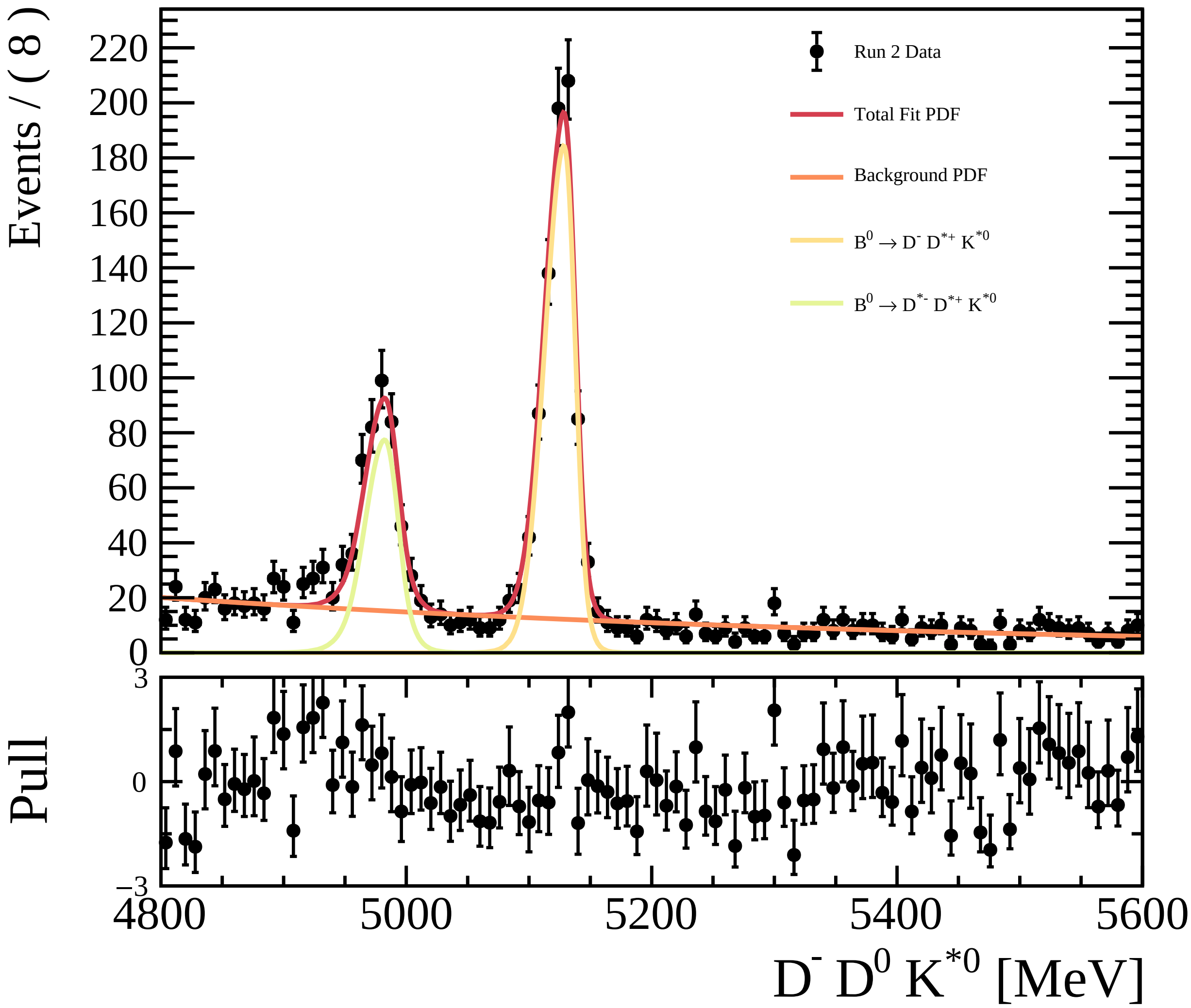


Figure 57: The final simultaneous fit to the $D^-D^0K^{*0}$ spectrum with no uncertainties or correlations Incorporated into the fit.

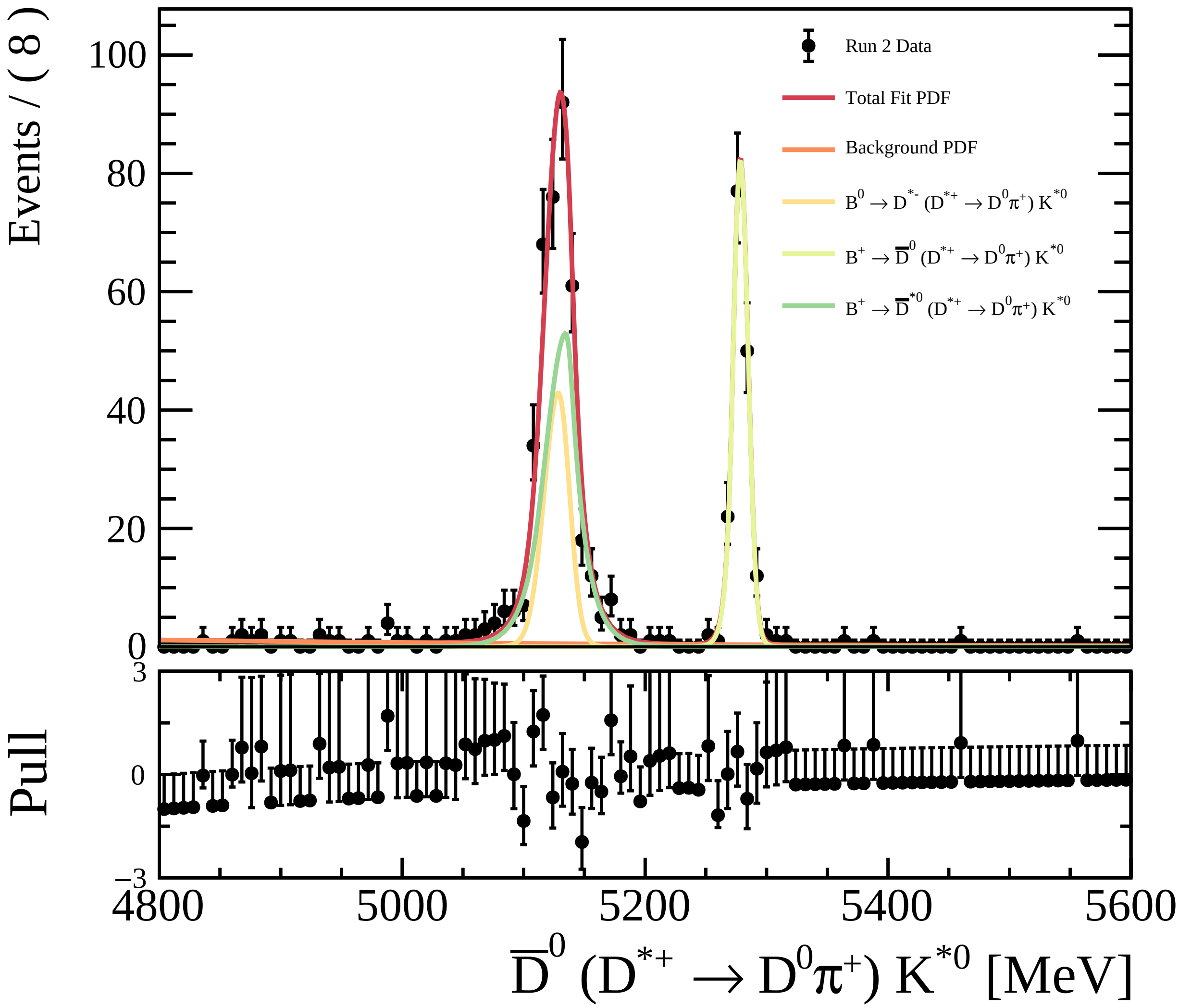


Figure 58: The final simultaneous fit to the $\overline{D}^0(D^{*+} \to D^0\pi^+)K^{*0}$ spectrum with no uncertainties or correlations incorporated into the fit.

References

[1] Particle Data Group, M. Tanabashi *et al.*, *Review of particle physics*, Phys. Rev. **D98** (2018) 030001.

[2] *Standard model of elementary particles.*, `https://en.wikipedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg`. Accessed: 2022-07-12.

[3] BaBar collaboration, J. P. Lees *et al.*, *Measurement of an Excess of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ Decays and Implications for Charged Higgs Bosons*, Phys. Rev. D **88** (2013) 072012, `arXiv:1303.0571`.

[4] Belle collaboration, M. Huschle *et al.*, *Measurement of the branching ratio of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ relative to $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ decays with hadronic tagging at Belle*, Phys. Rev. D **92** (2015) 072014, `arXiv:1507.03233`.

[5] Belle collaboration, S. Hirose *et al.*, *Measurement of the τ lepton polarization and $R(D^*)$ in the decay $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ with one-prong hadronic τ decays at Belle*, Phys. Rev. D **97** (2018) 012004, `arXiv:1709.00129`.

[6] Belle collaboration, G. Caria *et al.*, *Measurement of $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ with a semileptonic tagging method*, Phys. Rev. Lett. **124** (2020) 161803, `arXiv:1910.05864`.

[7] LHCb collaboration, R. Aaij *et al.*, *Measurement of the ratio of branching fractions $\mathcal{B}(\bar{B}^0 \to D^{*+}\tau^-\bar{\nu}_\tau)/\mathcal{B}(\bar{B}^0 \to D^{*+}\mu^-\bar{\nu}_\mu)$*, Phys. Rev. Lett. **115** (2015) 111803, `arXiv:1506.08614`, [Erratum: Phys.Rev.Lett. 115, 159901 (2015)].

[8] LHCb collaboration, R. Aaij *et al.*, *Test of Lepton Flavor Universality by the measurement of the $B^0 \to D^{*-}\tau^+\nu_\tau$ branching fraction using three-prong τ decays*, Phys. Rev. D **97** (2018) 072013, `arXiv:1711.02505`.

[9] E. Mobs, *The CERN accelerator complex in 2019. Complexe des accélérateurs du CERN en 2019*, , General Photo.

[10] LHCb collaboration, A. A. Alves Jr. *et al.*, *The LHCb detector at the LHC*, JINST **3** (2008) S08005.

[11] C. Elsasser, *Bb production angle plots*, .

[12] M. cmglee, *File:roc-draft-xkcd-style.svg*, .

[13] M. B. Gavela, P. Hernandez, J. Orloff, and O. Pène, *Standard model CP violation and baryon asymmetry*, Mod. Phys. Lett. **A9** (1994) 795, `arXiv:hep-ph/9312215`.

[14] LHCb collaboration, R. Aaij *et al.*, *Observation of $J/\psi p$ resonances consistent with pentaquark states in $\Lambda_b^0 \to J/\psi K^- p$ decays*, Phys. Rev. Lett. **115** (2015) 072001, `arXiv:1507.03414`.

[15] P. Nath and P. F. Pérez, *Proton stability in grand unified theories, in strings and in branes*, Physics Reports **441** (2007) 191.

[16] D. C. B. Abi, R. Acciarri *et al.*, *Prospects for beyond the standard model physics searches at the deep underground neutrino experiment*, The European Physical Journal. C, Particles and Fields **81** (2021).

[17] BaBar collaboration, P. del Amo Sanchez *et al.*, *Measurement of the*, Phys. Rev. **D83** (2011) 032004, `arXiv:1011.3929`.

[18] Belle collaboration, J. Brodzicka *et al.*, *Observation of a new D(sJ) meson in B+ —¿ anti-D0 D0 K+ decays*, Phys. Rev. Lett. **100** (2008) 092001, `arXiv:0707.3491`.

[19] BaBar collaboration, B. Aubert *et al.*, *Measurement of the branching fractions for the exclusive decays of B^0 and B^+ to $\bar{D}^{(*)}D^{(*)}K$*, Phys. Rev. D **68** (2003) 092001, `arXiv:hep-ex/0305003`.

[20] LHCb collaboration, R. Aaij *et al.*, *First observation of the decay $B^0 \to D^0\overline{D}^0K^+\pi^-$*, Phys. Rev. D **102** (2020) 051102, arXiv:2007.04280.

[21] C. Amsler, *15. quark model - particle data group*, 2021.

[22] R. Alonso, B. Grinstein, and J. M. Camalich, *Lepton universality violation and lepton flavor conservation in b-meson decays*, 2015. doi: 10.48550/ARXIV.1505.05164.

[23] A. Venkateswaran, *Probing Lepton Flavour Universality*, PhD thesis, 2022, Copyright - Database copyright ProQuest LLC; ProQuest does not claim copyright in the individual underlying works; Last updated - 2022-07-18.

[24] W. D. Hulsbergen, *Decay chain fitting with a kalman filter*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **552** (2005) 566.

[25] W. Verkerke and D. P. Kirkby, *The RooFit toolkit for data modeling*, eConf **C0303241** (2003) MOLT007, arXiv:physics/0306116, [,186(2003)].

[26] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, arXiv:physics/0402083.

[27] A. Rogozhnikov, *Reweighting with Boosted Decision Trees*, J. Phys. Conf. Ser. **762** (2016) , arXiv:1608.05806, https://github.com/arogozhnikov/hep_ml.

[28] M. Needham, *Identification of Ghost Tracks using a Likelihood Method* , CERN, Geneva, 2008.

[29] Particle Data Group, P. A. Zyla *et al.*, *Review of particle physics*, Prog. Theor. Exp. Phys. **2020** (2020) 083C01.

[30] V. V. Gligorov and M. Williams, *Efficient, reliable and fast high-level triggering using a bonsai boosted decision tree*, JINST **8** (2013) P02013, arXiv:1210.6861.

[31] D. Müller, M. Clemencic, G. Corti, and M. Gersabeck, *ReDecay: a novel approach to speed up the simulation at LHCb*, The European Physical Journal C **78** (2018) .

[32] BaBar collaboration, P. del Amo Sanchez *et al.*, *Measurement of the B —> D-bar(*)D(*)K branching fractions*, Phys. Rev. D **83** (2011) 032004, `arXiv:1011.3929`.

[33] A. Poluektov, *Kernel density estimation of a multidimensional efficiency profile*, Journal of Instrumentation **10** (2015) P02011.

[34] S. Das, *A simple alternative to the Crystal Ball function*, `arXiv:1603.08591`.

[35] T. Skwarnicki, *A study of the radiative cascade transitions between the Upsilon-prime and Upsilon resonances*, PhD thesis, Institute of Nuclear Physics, Krakow, 1986, DESY-F31-86-02.

Vita

NAME OF AUTHOR: Harris Bernstein

PLACE OF BIRTH: Redwood CA, USA

DATE OF BIRTH: August 25, 1993

DEGREES AWARDED: Bachelors of Science in Physics. Pennsylvania State University, USA